The Kinetics, Thermodynamics and Mechanisms of Short Aromatic Peptide Self-Assembly

Mason, Thomas O.; Buell, Alexander K.

Published in:
Biological and Bio-inspired Nanomaterials

Link to article, DOI:
10.1007/978-981-13-9791-2_3

Publication date:
2019

Document Version
Peer reviewed version

The kinetics, thermodynamics and mechanisms of short aromatic peptide self-assembly

Thomas O. Mason[†] and Alexander K. Buell[*,‡]

[†]*Department of Materials and Interfaces, Weizmann Institute of Science, Rehovot, 76100, Israel*

[‡]*Institute of Physical Biology, Heinrich-Heine-University, 40225 Düsseldorf, Germany*

E-mail: alexander.buell@uni-duesseldorf.de

Abstract

The self-assembly of short aromatic peptides and peptide derivatives into a variety of different nano- and microstructures (fibrillar gels, crystals, spheres, plates) is a promising route toward the creation of bio-compatible materials with often unexpected and useful properties. Furthermore, such simple self-assembling systems have been proposed as model systems for the self-assembly of longer peptides, a process that can be linked to biological function and malfunction. Much effort has been made in the last 15 years to explore the space of peptide sequences, chemical modifications and solvent conditions in order to maximise the diversity of assembly morphologies and properties. However, quantitative studies of the corresponding mechanisms of, and driving forces for, peptide self-assembly have remained relatively scarce until recently. In this chapter we review the current state of understanding of the thermodynamic driving forces and self-assembly mechanisms of short aromatic peptides into supramolecular structures. We will focus on experimental studies of the assembly process and our perspective will be centered around diphenylalanine (FF), a key motif of the amyloid β sequence and a paradigmatic self-assembly building block. Our main focus is the basic physical

chemistry and key structural aspects of such systems, and we will also compare the mechanism of dipeptide aggregation with that of longer peptide sequences into amyloid fibrils, with discussion on how these mechanisms may be revealed through detailed analysis of growth kinetics, thermodynamics and other fundamental properties of the aggregation process.

Introduction

It has been known for decades that (poly)peptides can assemble into ordered supramolecular structures, such as crystals and fibres.[1] Such assemblies were mostly studied in the context of structural biology (e.g. protein crystals[2]) or cellular biology (e.g. filaments of the cytoskeleton). Only in the second half of the 20th century it became clear that filamentous protein assemblies can also be linked to disease processes.[3] A particular type of filamentous assembly, so-called amyloid fibrils, were found in a range of different disorders, ranging from Alzheimer's disease to type 2 diabetes.[4] Amyloid fibrils can be formed by a large variety of unrelated polypeptides and share a common morphology as linear, mostly unbranched structures of several nanometers in diameter and several micrometers in length. These assemblies are characterised by a high β-sheet content and characteristic tinctorial properties and X-ray diffraction pattern.[5] Until recently, mostly low resolution structural information was available on assemblies from sequences longer than about 10 residues, e.g. from limited proteolysis experiments[6] whereas the possibility to crystallise short (<10 aa) peptides allowed access to atomically resolved structures.[7] The recent progress in the production of homogeneous samples of amyloid fibrils (*in vitro*[8,9] and *ex vivo*[10]) as well as technical progress in solid state NMR spectroscopy and cryo-electron microscopy now allows to define the structures of amyloid fibrils from longer peptide sequences at the individual amino acid residue level of detail.

The finding that such supramolecular protein structures play a role in diseases triggered a large scale effort to elucidate the mechanisms and driving forces that lead to the formation

of amyloid fibrils, as well as a search for potential inhibitors of their formation. In the course of such studies, Gazit and co-workers aimed at identifying minimal amino acid sequence motifs that would be responsible for the self-assembly of the proteins implicated in disease.[11] These investigations ultimately led to the isolated study of the diphenylalanine (FF) peptide,[12] a central motif (position 19 and 20) of the sequence of the amyloid β peptide, the aggregation of which is a hallmark of Alzheimer's disease.[13] It was found that FF, rather than assemble into amyloid fibrils, formed crystalline hollow tubes,[12] which were later shown to have the same molecular arrangement as the crystals of the FF peptide the structure of which had been solved a few years earlier.[14] The study of these structures really gained momentum when it was shown that the resulting nano- to microscale structures could be functionalised in a variety of ways and displayed interesting mechanical, optical and electrical properties.[15–18] Ultimately, this line of research converged with the field of short peptide gelators. It had been known for some time that, in addition to many other small molecules, short peptides or even amino acids (aromatic or aliphatic), often terminally capped or functionalised with aromatic protection groups, were capable of self-assembly and gelation.[19–21] To-date, hundreds of studies have been published that report on the assembly and structure formation of short (aromatic) peptides in order to create assemblies with potentially useful properties. Most of the experimental work in this area in the last 15 years has focused on the creation of new types of materials, whereas fundamental mechanistic studies have received less attention. The lack of basic physico-chemical characterisation has in some cases led to misinterpretation of observed phenomena and, more generally, inhibited quantitative studies of the dipeptide systems. This situation has changed in the last five years or so, where significant efforts and progress have been made in the elucidation of the fundamental driving forces responsible for the assembly of aromatic dipeptides and their mechanisms of assembly. We attempt to give a summary of these studies and the advances in the understanding of this class of assembly processes. In addition, we will address the question as to what we can learn from the self-assembly of short aromatic peptides about the assembly characteristics

of longer peptide sequences into (disease-related or functional) amyloid fibrils. We will start the chapter by reviewing the most important types of intermolecular interactions responsible for peptide self-assembly. The reader familiar with these basic physico-chemical concepts is invited to skip the first section.

The nature of the interactions responsible for peptide assembly

A key concept in biologically-relevant self-assembly processes is the reversibility of the interactions[22,23]- in contrast to reactions where bonds are formed or broken by external stimulus, self-assembly relies on the sum total of a number of lower energy interactions being favourable in order to achieve a stable form. The interactions in a self-assembly process depend on multipolar interactions, hydrogen bonding, steric effects, dispersion forces, in addition to 'forces' arising from the environment in which assembly takes place, most notably for the purposes of this chapter, the hydrophobic effect.

Hydrogen bonding

Energetically, the strongest interactions beyond the covalent bonds in biological systems are electrostatic in nature - variations in electron density giving rise to a force between two building blocks that may be attractive or repulsive, in contrast to the always-attractive dispersion forces. In polypetides, major structural elements are defined by the pattern of hydrogen bonding between amide linkages they display.[24,25] The hydrogen bond is a peculiar case of an electrostatic attraction for both its ubiquity and its strength relative to other intermolecular forces. It occurs between a dipole of the form $^{\delta^-}A - H^{\delta^+}$ and an acceptor B, a region of high electron density typically (but not always) associated with an electronegative atom in a dipole or an anion.[26] B donates electron density to the partially positively charged hydrogen atom, resulting in an attractive potential. Steiner in a 2002 review highlighted the

concept of a hydrogen bond as being describable as a "frozen" proton transfer.[27] The wide range of possible species A and B, together with geometric restraints imposed by the local covalent bonds and resultant steric effects leads to a range of energies for the hydrogen bond, but in biomolecules these rarely exceed 10 $k_B T$[28] as isolated interactions. Importantly, in aqueous solution, the competition from hydrogen-bonding solvent water lowers the overall free energy of formation of intramolecular hydrogen bonds[29] to a range in which they may be reversible *in vivo*, as expected and required in self-assembly processes.[23,30]

An important implication of hydrogen bonding systems in structure formation is geometry.[24,31] The hydrogen bond, as an electrostatic multipole, has an orientation dependent energy, with a minimum for the three nuclei involved being collinear. In systems where external constraints on the geometry of the hydrogen bond exist (as is the case in biomacromolecules), the hydrogen bond can be observed at less-favourable angles. The second geometric implication of the hydrogen bond is its range - the interaction energy between the dipole (A-H) and monopolar acceptors scales as r^{-2}, and the dipole-dipole contributions scale as r^{-3}. Monopole-monopole interactions are most important in the case of interactions between formal charges, such as in salt bridges and these scale as r^{-1}.[32] In short peptides terminated by the free amine and carboxylic acid, these interactions commonly occur at the termini, and also occur between charged side chains, the high energy of the attraction allowing self-assembling peptides to be designed based simply on the predictability of the proton transfer occurring.[33–35] Indeed, in native proteins, in the vast majority of cases, species capable of hydrogen bonding are observed to be involved in a hydrogen bond,[36] and these bonding interactions have favoured bond angles and bond lengths. The hydrogen bond tends towards a co-linear geometry for the three atoms - the $A-H\cdots B$ angle tending to 180° with increasing bond strength,[37] and the angle $H\cdots CO$ at carbonyl acceptors has an energy minimum at 120°, around which there is clustering in the distribution of observed angles in the solid state.[38] Alkyl protons, both simple and activated, can also play a role as secondary hydrogen bond donors in protein self-assembly and molecular recognition. These

weak interactions are explained relatively conventionally for what has been said thus far - the feebly polarised C-H bonds can, in the presence of a relatively strong acceptor, form a hydrogen bond. No capacity for acceptance of hydrogen bonds is observed for alkanes, and hence the side chains of the amino acids alanine, valine, leucine, isoleucine and proline do not operate as hydrogen bond acceptors. The hierarchy of hydrogen bond strength has greater structural implication in short peptides, with each interaction contributing a proportionally greater of the total than in large biomolecules, where the weak bonds tend to have a greater significance in structure and recognition.[39] Of particular present interest is the capacity of the aromatic phenylalanine and tyrosine to act as both aromatic C-H donors and as acceptors of hydrogen bonds perpendicular to the ring centre.[39]

Backbone-backbone interactions represent a significant proportion of hydrogen bonding interactions in natively-folded proteins,[36] with other interactions of the conventional form such as (N, NH, NH_2^+, O)H$\cdots$(OH, O=, N). However, hydrogen bonding is not limited to hydrogen covalently bonded to one of two closely-approaching nitrogen or oxygen atoms.[40] The interaction is fairly general provided there exists an A-H dipole with a partial positive charge on the hydrogen, and a lone pair on B.[41,42] Phenylalanine for example can, and frequently does act as a hydrogen bond acceptor.[43] The reason for the observed dual functionality is the aromaticity phenomenon (see below), which leads to non-isotropic electron density in much the same manner as more familiar dipolar species, but with different symmetry and hence interaction geometry.

Hydrophobicity

The previously described effects are inherent to the molecules themselves - hydrogen bonding and electrostatic interactions will occur between suitable species in the gas phase, and similarly steric effects limit the allowable conformations of isolated molecules and those contacting others. In dealing with biologically-relevant systems, however, consideration of the solvent environment is essential to any analysis of the mechanism and thermodynamics of

the assembly process. The hydrophobic effect is the name given to the familiar phenomenon whereby non-polar species display far lower solubility in water than they do in organic solvents, tending to phase separate despite only feeble attractive forces between the non-polar species that form the new phase, and operates on both apolar and amphiphilic solutes, the latter giving rise to micellar and bilayer self-assembly.[44,45] Water structure, characterised by extensive and highly fluxional chains of (directional) hydrogen bonds, is frustrated by the presence of the solute, which acts as a 'cavity' to which no hydrogen bonds may be made.[46,47] Small solutes, those with radii of curvature on the molecular scale such as linear hydrocarbons, are less disruptive to water structure than are flatter/larger species which do not afford the water molecules the ability to form hydrogen bonding chains around the solute, instead forming an interface.[48–50] The contribution of this process to the free energy of solvation naturally has a linear dependence on surface area, and for solute radii larger than a few nanometres, the surface energy is the dominant term and tends towards the value observed in bulk.[51,52] The small solutes, through their effects as cavities in the hydrogen bond network, affect the solvent in the hydration layer most strongly, the number of water molecules in this layer scaling with solute volume.[52]

The entropic and enthalpic signature of small-molecule hydrophobic solvation at low temperature is a free energy composed of a favourable enthalpic (from intermolecular interactions) component and an unfavourable entropic (reduction in the number of energetically favourable conformations available to the water molecules in the presence of the hydrophobe) component. The ordering of the water around the solute is a compensated process. The restrictions on the conformations available to the water molecules results in there being a significant change in the heat capacity of the solvation, where the effects due to this localised ordering diminish rapidly with temperature.[53] The compensated ordering process, yielding a small negative free energy, can be seen as a mechanism that enhances the solubility of hydrophobic species at low temperatures.[54] The magnitude of fluctuations in water density, a relevant concern for the creation of a cavity, at the microscale are fairly indepen-

dent of temperature,[55] certainly when compared to less 'structured' organic media, and a cavity-insertion study demonstrates that this invariance in compressibility of solvent determines both the dominance of excluded-volume effects in the entropy of solvation for small species, and the observed high partial molar heat capacity associated with hydrophobic solvation.[52,56] Desolvation and the formation of hydrophobic aggregates in a new phase is driven by the differential scaling of the entropic excluded-volume type contributions and the enthalpic interface contributions. Above a critical radius, which for short hydrocarbons in water is around 1 nm, clusters display surface-area energy scaling (sublinear in aggregation number), lower in energy than the separate excluded-volume cavities,[51,57] and so the new phase tends to grow. The early emergence of rudimentary structure in protein folding has been ascribed to the desolvation and interior segregation of hydrophobic residues in the "hydrophobic collapse" model of protein folding, in which desolvation, with or without secondary structure formation is on the main folding pathway.[58,59] Buried residues typically pack extremely closely in their native state- the packing efficiency in some cases being found to be as high as in organic crystalline solids.[60]

As a structural determinant in short-peptide self-assembled aggregates, hydrophobicity is similarly a phenomenon associated with the absence of strong hydrogen bonding or other polar interaction. Segregation of the polar and apolar components on adoption of an ordered structure is usually achieved in the solid state, and those dipeptides which reliably crystallise have remarkable commonalities in the arrangement of their hydrogen bonded and charged moieties, permitting groupings into a comparatively small number of structural classes.[61] The higher-energy electrostatic intermolecular interactions are present in all the surveyed crystalline structures, but their connectivity and spatial arrangement are moderated by the packing requirements of the side chains.

Aromaticity in proteins and short peptides

Aromaticity is a property of cyclic species with delocalised electron density above and below the plane of the ring. All bonding molecular orbitals of the ring system are filled with electron pairs, while nonbonding and antibonding orbitals are unfilled. It is a stable phenomenon, and reactions commonly favour its retention rather more than comparable systems (for example, linear conjugated bonds) without aromatic character. The electron density distribution has a significant quadrupolar component. The symmetry is that of a spherical harmonic of degree 2 (vs. 1 for the dipole, 0 for the isolated charge) and so, as in the case of the dipole, the quadrupole interaction is directional, its sign depending on the approaching species and the orientation of the quadrupole. In proteins, of course, even phenylalanine will display lower order multipoles due to the asymmetry induced by the bonds to the backbone, this being obviously true for the substituted tyrosine and the heterocyclic tryptophan side chains. The interactions of the quadrupole have a steeper fall-off with distance than comparable dipolar interactions, the dipole-quadrupole having r^{-4} dependence and inter-quadrupole interactions varying as r^{-5}. Despite the tendency of quadrupolar interactions to be significant only at short ranges, the ubiquity of inter-aromatic interactions between the side chains of the amino acids phenylalanine, tyrosine and tryptophan suggests that the peculiar directionality[62] of the quadrupolar interactions may play an important role in the generation or stability of protein secondary and tertiary structure. The aromatic residues have a significant tendency to co-locate in folded sequences, comparable to that of oppositely-charged side chains,[63] and aromatic pairing was found in 89% of a representative sample of proteins (including the aromatic neutral histidine side chain) for which high-resolution data was available in the Protein Data Bank.[64] Aromatic-aromatic pairs adopted both stacked-offset or 'herringbone' geometry and T-shaped C-H$\cdots\pi$ 'hydrogen bond' pairs.[64] A distinction from general hydrophobic interactions in the folding process is suggested by analysis of the sequence separation of the interacting pairs. 74% of aromatic-aromatic interactions in a statistical sample were found to occur between distant pairs, suggesting that formation of

the π interactions follows folding of secondary structure.[65] Both of the favoured geometries involve favourable quadrupole-quadrupole interactions, while the T-shaped arrangement has a significant dipole-quadrupole character.[66,67] The effect of the quadrupole is to slightly favour the geometry where the C-H bond vector points to the acceptor ring centre, whereas non-aromatic donors are found to 'aim' at centres and edges with little energy difference.[27] Both appear in native protein structures, with stacked-offset arrangements more common in isolated pairs not part of a network.[68] A tendency has been observed for aromatics and certain other hydrophobic amino acids to be encoded in β-sheet forming regions of native structure, a phenomenon that has been linked to the steric effect of branching on carbon 3 of the side chain[69] or from the presence of the bulky (and conformationally rigid) aromatic side chains of Trp, Phe and Tyr.[70] The large accessible surface area of the side chains is also significant in their pairing. An exponential relationship was found between side chain solvent accessible surface area and the average number of atoms within a contact radius of the atoms of the side chain. The aromatic species, with a large and hydrophobic side chain surface area, were found to have significantly more neighbouring atoms in folded structures than any other amino acid.[71]

The role of phenylalanine residues in peptide self-assembly into amyloid fibrils

The hexapeptide human islet amyloid polypeptide fragment hIAPP (22-27), sequence NFGAIL, was found to assemble *in vitro* into amyloid fibrils in a similar fashion to its parent polypeptide.[73] An alanine-scanning study showed the importance of the phenylalanine residue to the formation of fibrils by the fragment and it was found that only the mutation F2A completely abolished the capacity of the species to form fibrils.[11] Other studies on fragments of the Aβ peptide showed that KLVFF (Aβ (16-20)) was capable of inhibiting Aβ fibril growth *in vitro*,[74] as were cholylated sequences cho-QKLVFF, cho-KLVFF and

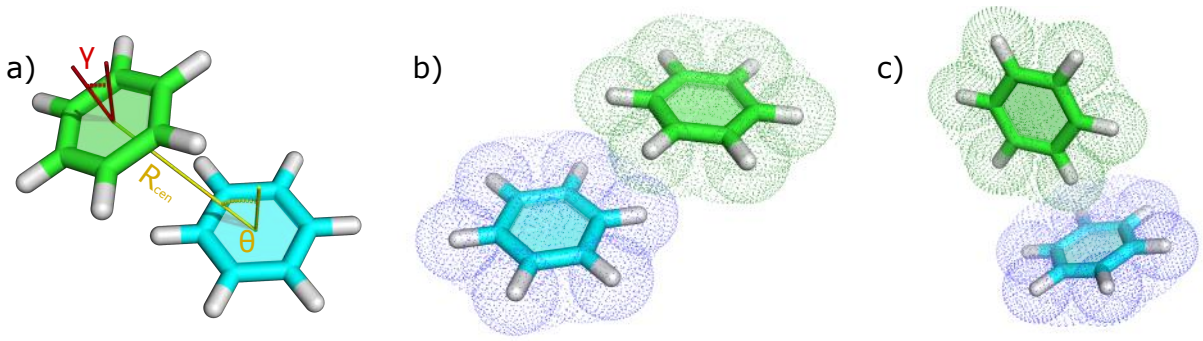

Figure 1: The geometry of the π-π interactions can be described by a variety of parameters, in a) two angles and a distance are defined, the angle θ being from the normal of one ring to the inter-centroid vector R_{cen}, and γ is the angle between the plane normals. b) A representative favourable[72] stacked-offset interaction is shown - $20° < \theta < 50°$, $\gamma < 30°$.[62] In c) the T-shaped edge-to-face geometry is shown, these being classified by $60° < \theta$, $\gamma < 90°$. The intercentroid distance is below 6.5 Å, peaking around 5.3 Å for F-F pairs,[68] naturally with stacked offset dominating at lower R_{cen} and T-shaped at greater range.[64] Detection of interaction by finding the closest intermolecular carbon-carbon contact (R_{CC}) gives a sharper range distribution, peaking around R_{CC}=3.8 Å.

cho-LVFFA.[75] This prompted further research into the possible role of phenylalanine specifically[76] in the processes of amyloid formation and deposition, with a focus on π-stacking interactions as a driving force for aggregation. Still shorter polypeptides, the tetrapeptides KFFE and KVVE, were also shown to be capable of self-organisation into amyloid-like β-sheet structures.[77] This remarkable aggregation behaviour, ascribed to sequence-dependent β-sheet propensity and terminal charge-charge interactions suggested that the central dipeptide FF, Aβ (19-20), could display aggregation properties and processes relevant to those of longer chains. The impact of two consecutive phenylalanine residues was later highlighted in a study that compared the assembly behaviour of extended versions of KLVFF/KLVF (capped and uncapped analogs).[78] However, the special role of aromatic residues in peptide self-assembly has also been challenged in a study that demonstrated similar self-assembly of aliphatic and disease-amyloid derived short peptides.[79] The debate as to the role of aromatic residues in the formation of amyloid fibrils is still ongoing, with the most recent support of an important role of phenylalanine provided by a study that substituted the phenylalanines of the amyloid-β peptide by cyclohexane and showed that substituton of any of the three F

residues abolished amyloid fibril formation.[80] However, in this study, recombinant wild type peptide was compared with chemically synthesized substitution variants, rendering direct comparison of self-assembly behaviour not straightforward.

Experimental methods to study short peptide assembly

The self-assembly of short aromatic peptides can be induced in a variety of ways and the self-assembly process monitored *in situ* or *ex situ* by a large array of experimental methods, that we summarise in this section, with a particular focus on studies of diphenylalanine (FF).

The choice of the assembly conditions for self-assembly

The self-assembly of FF and its derivatives into ordered structures is most commonly, but not exclusively, induced and investigated in solution. In many studies, FF was dissolved at high concentration in an organic solvent (mostly Hexafluoroisopropanol, HFIP) and the self-assembly was then induced by injecting the concentrated stock solution into a low solubility solvent (mostly water).[12] A large number of different combinations of high solubility and low solubility solvents have been investigated[81–87] and it has been shown that very different morphologies can be obtained, depending on the choice of solvents. Indeed, it was shown that even small impurities of solvents can have a significant effect on the self-assembly process.[28] Mason et al. have determined the solubilities of FF in a variety of solvents at room temperature,[87] allowing a rational choice of solvent combination based on their relative FF-solubilising abilities. In the context of the creation and characterisation of novel materials, the choice of appropriate solvent for FF assembly is mostly determined by the desired type of structure. However, assembly in water has a privileged role, due to the relevance of self-assembly studies in water for biology and also biocompatibility of the so-created structures. Indeed, in some of the earliest studies of FF, the peptide was dissolved at elevated temperatures in pure water, followed by cooling down of the solution, which led to

self-assembly.[14,88] It has also been demonstrated that a change in pH is a suitable method to induce FF assembly, based on the strong pH-dependence of FF solubility.[89] A concentrated solution of FF in trifluoroacetic acid (TFA) can be titrated with ammonia solution, which leads to the formation of FF nanowires.[90] A particularly elegant method of pH-change-induced self-assembly has been developed in the context of short peptide hydrogel formation, whereby the slow hydrolysis of an acid anhydride (most notably glucono-δ-lactone, GdL) to the corresponding acid (gluconic acid) slowly and homogeneously acidifies the peptide solution, leading to significantly more homogeneous gel appearance compared to titration with a strong acid.[91] An alternative method for homogeneous induction of self-assembly is enzymatic activation that has been presented for FF derivatives,[92,93] whereby the precursor molecule is designed to be the substrate of an enzyme, such as a kinase. Self-assembly can also be induced through enzymatic action if the dipeptide is created *in situ* from an activated amino acid derivative (methanoate ester) and an amino acid amide, whereby self-assembly of the formed dipeptide immediately follows.[94]

Assembly of FF from the gas phase has also been investigated, using physical vapour deposition (PVD)[96] and it has been demonstrated that at sublimation temperatures above 190°C, the FF peptide cyclizes.[97] The use of plasma enhanced CVD at 200°C was also reported,[98] and most likely also yields assemblies of chemically modified species rather than intact FF. Vapour deposition at moderate temperatures initially only yields amorphous films. Hollow tubes of intact FF can be obtained by hydrating the amorphous layer or by depositing the FF at lower grade vacuum, illustrating the importance of water in the formation of the crystalline tubes.[97]

Microscopic methods

The observation of the peptide assemblies with different forms of microscopy is the most direct method of detection and is often used to define and distinguish different morphologies.[87,99] In many cases, in particular in the case of gel formation, the self-assembled struc-

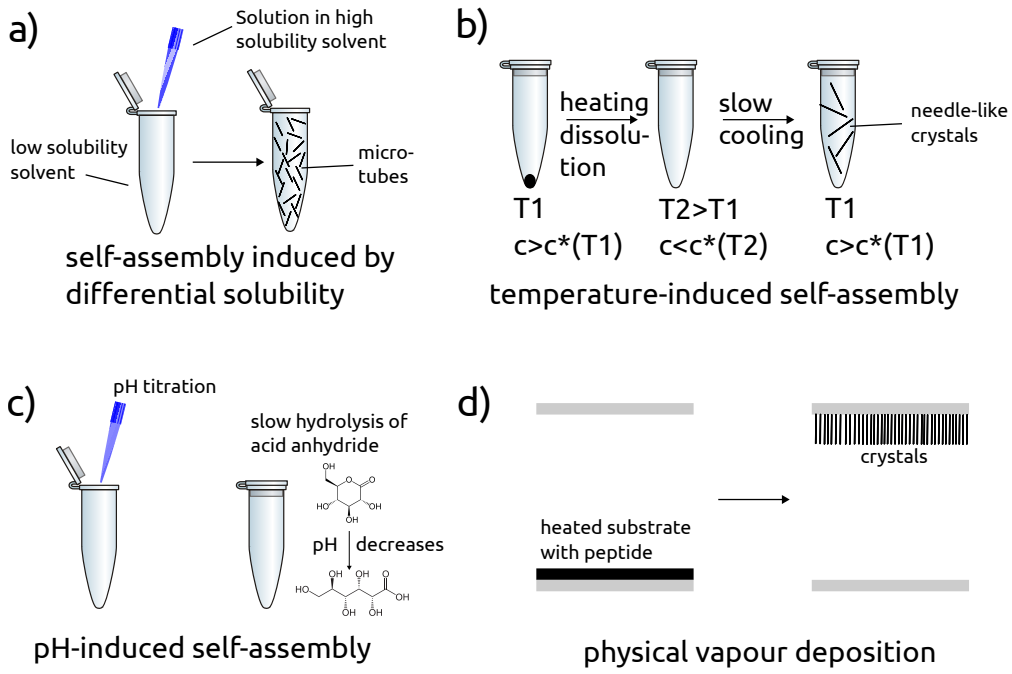

Figure 2: **Different methods to induce small aromatic peptide assembly a)** Introduction of a concentrated solution of dipeptide in a high solubility solvent, such as acetic acid or HFIP into a low solubility solvent, such as water or methanol.[12,87] **b)** Heating of a dipeptide solution at a concentration c to a temperature T2 where the solubility of the peptide, c*, is higher than the peptide concentration, followed by slow cooling.[88,95] **c)** pH titration[90] or *in situ* pH change through hydrolysis of an acid anhydride (here illustrated with gluconо-δ-lactone,[91]) exploit the pH-dependence of peptide solubility[89] to induce self-assembly. **d)** Self-assembly can also be induced by physical vapour deposition in vacuum,[96–98] leading to the formation of arrays of crystals oriented perpendicularly to the surface.

tures are too small to be observed by optical brightfield microscopy, and therefore atomic force microscopy (AFM),[16,100,101] scanning electron microscopy (SEM)[100,102,103] or transmission electron microscopy (TEM), either in the form of negative staining TEM[102,104] or in the form of cryo-TEM[105] can be used. The use of atomic force microscopy has the additional benefit of allowing to probe the mechanical properties of the assemblies.[16,103] For fibrillar structures of diameters in the nm range and lengths in the μm range, various fluorescence microscopy techniques can also be employed, such as confocal laser scanning[106–108] or super-resolution microscopy.[107] In particular, these direct nanoscopic imaging techniques have enabled highly detailed mechanistic studies of the assembly mechanisms[107] (see below).

Spectroscopic methods

A wide range of different spectroscopic methods have been used to detect and follow the self-assembly process. In particular spectroscopic tools have been employed in order to monitor the change in peptide conformation between the soluble and assembled states. Fluorescence spectroscopy can be used, both in the form of measurements of the intrinsic fluorescence of the aromatic residues,[109] or through the binding of a fluorescent molecule, such as the amyloid fibril-binding dye Thioflavin-T[102,104] or fluorescently labeled DNA.[110] Other types of spectroscopy that have been shown to be useful for structural characterisations are circular dichroism (CD)[100,104,105] and Fourier transform infrared (FT-IR)[102,104,106] spectroscopy, that are able to inform about secondary structure motifs formed by the peptides. Furthermore, NMR spectroscopy has also been used, either to follow the decrease in soluble peptide,[99,111] to monitor structural changes of assemblies,[105] to characterize the charge state of the peptide, e.g. as a function of concentration,[78] or to study the interactions between assemblies and other species, such as metal ions.[112]

Despite not being a spectroscopic method, we include here also mass spectrometry, in the form of ion mobility mass spectrometry, as it has been used in order to detect the presence of clusters of aromatic dipeptides in solution.[113]

Scattering, rheological, calorimetric and conductivity-based methods

The often dramatic change in size upon self-assembly makes it possible to follow the self-assembly process by a variety of scattering methods. Dynamic light scattering (DLS),[89,99] small angle X-ray scattering (SAXS),[114] as well as small angle neutron scattering (SANS)[106,115,116] have been employed to follow the assembly process. In the case of small angle scattering (SAXS and SANS), the data can sometimes be fitted to a model corresponding to the emergence of rod-like micelles/fibrils upon the induction of self-assembly.[106]

The self-assembly is often associated with a considerable change in mechanical properties, in particular in the case of gel-formation. This can be followed by the measurement of the rheological properties, in particular the ratio between elastic and dissipative behaviour.[93,104,106,116]

In the case of orthogonally assembling, self-sorting multi-component gels, the rheological properties provide a measure of the assembly state of the individual components and can be selectively tuned through external stimuli.[117]

In some cases, the self-assembly reaction can also be directly carried out inside a calorimeter, with the aim to study the thermodynamic signatures of the assembly reaction. In the case of Boc-FF, the peptide initially forms amorphous spherical structures when a concentrated solution in ethanol is diluted into water (see below). These structures can be metastable against transformation into crystals for sufficient amounts of time such that the heat exchange linked to this conversion can be measured with a differential scanning calorimeter.[99]

Inspired by classical methods for the determination of the critical micelle concentrations of surfactants, it is also possible to follow different levels of peptide self-assembly by measuring the concentration-dependence of the conductivity of the peptide solution.[106] This type of method is based on the fact that the electrophoretic mobility of assemblies can differ significantly from those of their building blocks.

Microfluidics

A relatively recent addition to the toolbox of methods for the study of short aromatic peptide assembly is microfluidics, which is particularly well-suited for self-assembled structures that can be directly observed by optical microscopy, see Figure 3. Microfluidics allows the study of peptide assembly in real time by optical time lapse microscopy (Figure 3 c and e) while maintaining well-defined solution conditions, by constant flow of a solution with defined concentration over the growing assemblies, eliminating depletion effects.[95] The concentration can be adjusted by changing the degree of supersaturation of the stock solution that is drawn into the device (Figure 3 b and reference 95) or else by changing the relative proportions of supersaturated peptide solution and water/buffer in an on-chip mixing device (Figure 3 d and reference 118). This methodology has recently been applied in a series of studies that investigated the growth rate of FF microcrystals as a function of concentration[95,118] and temperature.[89] This approach allows extensive and accurate measurements to be made and the resulting data has contributed significantly to a better understanding of assembly mechanism of FF (see below).

Thermodynamic stability of peptide assemblies

The thermodynamics of short aromatic peptide assembly depends crucially on the exact sequence of the peptide, including capping and protection groups, as well as the solution conditions, i.e. temperature, pH and salt concentration. In the following, we will discuss the thermodynamic stability of crystalline and fibrillar structures and how they are influenced by various factors.

Thermal stability of FF crystals

It has been reported that FF microcrystals are very stable against heating in the dry state.[100] In the dry state, these structures can withstand temperatures above 100°C and thermogravi-

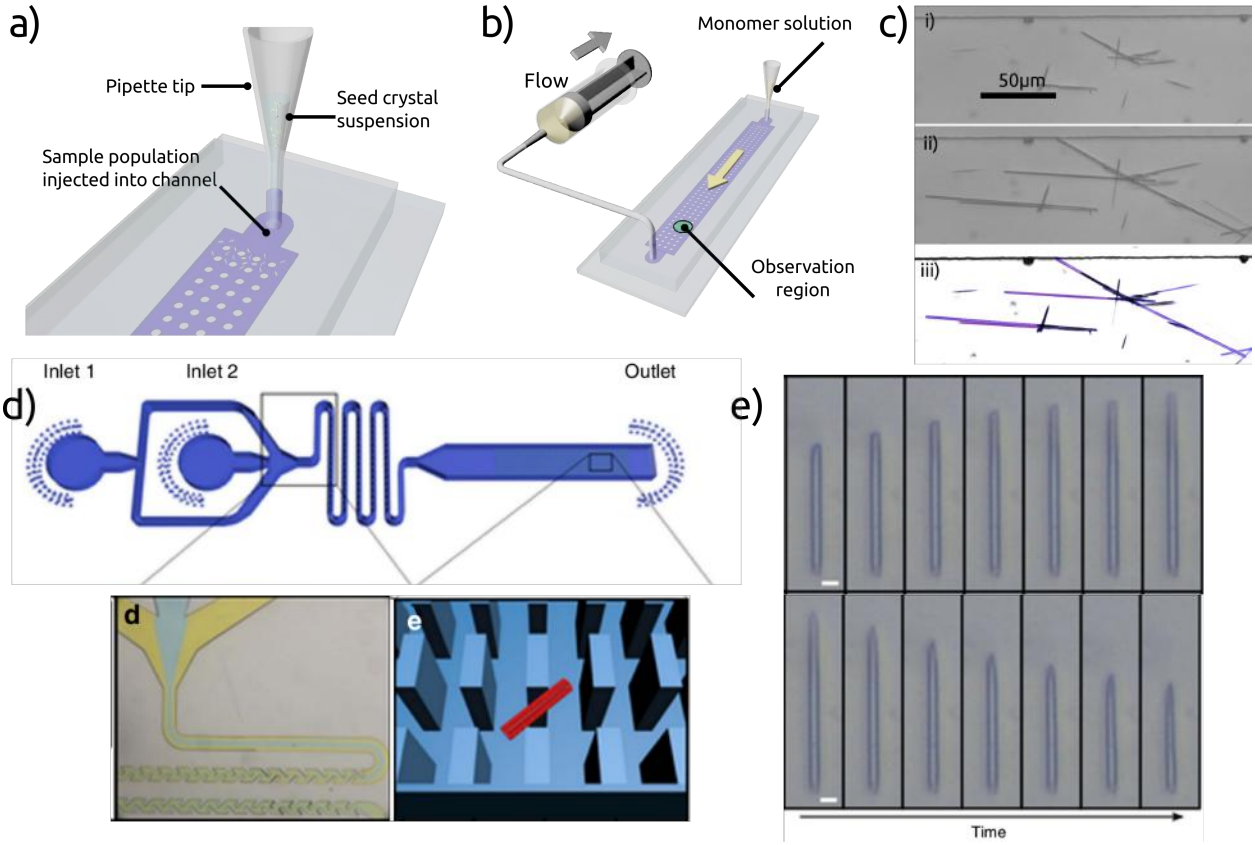

Figure 3: **Microfluidic methods to study diphenylalanine self-assembly a)** Preformed FF crystals can be injected into a wide microfluidic channel and trapped by support columns that avoid collapse of the channel. **b)** The trapped seed crystals can be flushed with a constant stream of a peptide solution at well-defined concentration. **c)** Optical time lapse microscopy can be used to follow the growth of the crystals. **d)** A mixing device can also be used, whereby the relative proportions of supersaturated peptide solution and water/buffer can be varied, leading to an adjustable degree of saturation of the peptide solution in contact with the seed crystals. **e)** If the seed crystals are flushed with a supersaturated solution, they are observed to grow (top), whereas they shrink when incubated with a sub-saturated solution (bottom). a)-c) reproduced with permission from reference 95 ©2016 American Chemical Society, and d)-e) reproduced from reference 118 under the Creative Commons Licence (CC-BY-4.0)

metric analysis reveals the loss of water from the crystal structure.[100] However, at temperatures above 150°C, it was shown that chemical decomposition sets in that can lead to the loss of phenylalanine fragments.[101] Furthermore, the FF peptide can cyclize under loss of water at such high temperatures, which leads to an irreversible collapse of the tubular crystals.[119,120] Full thermal decomposition of the peptide film finally occurs at temperatures above 300°C.[100,121] On the other hand, amorphous films of FF can be converted into fibrillar morphology through dry heating at 150°C,[121] possibly through rearrangement or, more likely, sublimation of chemically degraded/cyclized FF, similar to the observed formation of deposits on AFM cantilever observed in an earlier study.[101] This is corroborated by the finding that pre-cyclized, phenylalanine-containing dipeptides can form fibrils and gels.[122] High thermal stability of FF assemblies above 100°C (in an autoclave) has also been reported in aqueous solution.[100] Only recently, the solubility of FF in water has been systematically studied as a function of temperature, by measuring the FF concentration in the supernatant in equilibrium with FF crystals.[89] It was found that the solubility of FF dramatically increases with temperature, from about 0.5 g/L (1.6 mM) at 10°C to about 1.9 g/L (6 mM) at 70°C. Interestingly, phenylalanine shows a similar temperature dependence of its solubility, the absolute solubility being much higher,[123] whereas benzene shows a very weak temperature dependence of its solubility in water.[124] It is interesting to note that benzene, being a prototypical hydrophobic molecule, has a higher solubility in water than the diphenylalanine molecule. The temperature-dependent solubilities of F, FF and benzene are compared in Figure 4. In the light of the reported temperature dependence of FF solubility, it is likely that FF crystals at 2 g/L dissolve completely upon autoclaving and reform upon cooling of the sample.[100] At the same time it is plausible that the individual FF molecules will be more stable against chemical degradation in aqueous solution compared to the dry state, given that the major degradation process is the cyclization reaction which, being associated with the loss of water, is unfavourable in aqueous solution.

Systematic analysis of the temperature-dependent solubility of FF is not only useful for

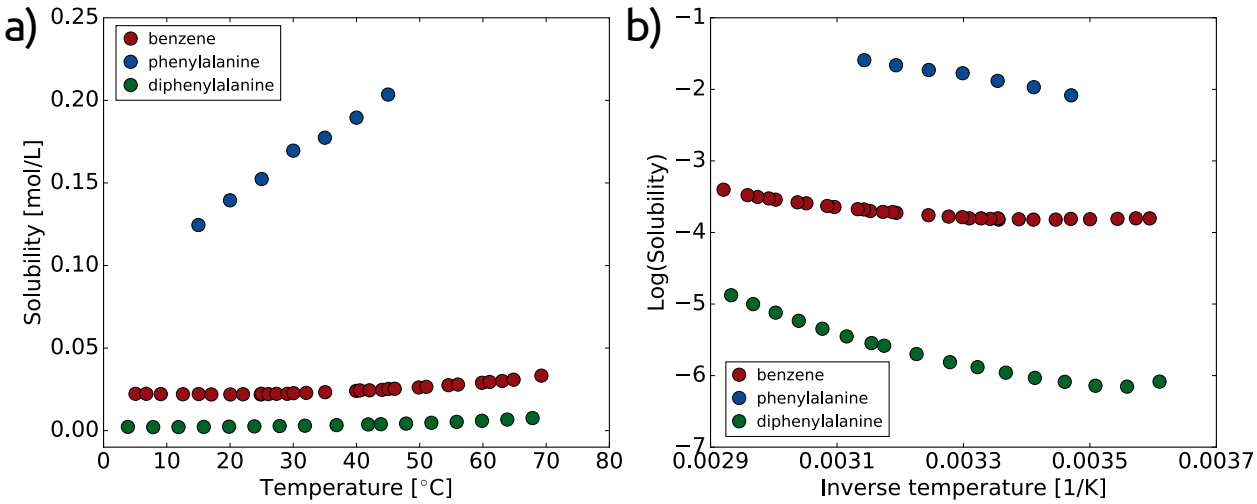

Figure 4: **Temperature-dependent solubilities of benzene, phenylalanine and diphenylalanine.** **a)** The equilibrium solubility of benzene,[124] phenylalanine[123] and diphenylalanine[89] are plotted against temperature. **b)** The logarithms of the solubility data are plotted against the inverse absolute temperature in a van't Hoff-like plot. The sign of the slope at any one point corresponds to the sign of the enthalpy of self-assembly (F, FF) or phase separation (benzene).

rationalising experiments that probe the thermal stability of its assemblies, but also allows to decompose the thermodynamic driving force of the assembly into enthalpic and entropic contributions, through a van't Hoff analysis (see Figure 4 for van't Hoff plots of F, FF and benzene). This type of analysis has been carried out for FF and it was found that while the free energy of assembly is virtually invariant throughout the investigated temperature range (4-68°C), the balance of enthalpy and entropy changes significantly.[89] While the assembly of FF into crystalline structures is purely entropy-driven below 10°C, it becomes enthalpically favourable above 10°C and entropically unfavourable above 30°C, a manifestation of entropy-enthalpy compensation. Overall, this temperature dependent solubility follows the behaviour expected for a peptide system dominated by the hydrophobic effect.[125]

Chemical stability of FF crystals

The stability of FF crystals against non-polar solvents has also been highlighted.[100] Systematic analysis of the solubilities of FF in a wide variety of solvents reveals that the solubility

varies by a factor of more than 200 000 between the lowest solubility solvent (hexane, 0.002 g/L) and the highest solubility solvent (acetic acid, 430 g/L) investigated.[87] It was also shown that the thermodynamic stability of FF crystals is strongly pH-dependent, with the solubility minimum coinciding with the pH range where FF occurs as a globally neutral zwitterion,[89] and the solubility at pH 1 being approximately ten times higher than at the solubility minimum.[126] Interestingly, addition of NaCl only marginally changed the solubility of FF in pure water; only at NaCl concentrations of 1 M and above, the solubility decreases appreciably,[89] suggesting that electrostatic screening effects do not play important roles in the assembly thermodynamics. The effect of high NaCl concentrations is likely related to the well-known phenomenon of salting out, that can for example be observed for aqueous solutions of benzene.[127]

Non-crystalline short aromatic peptide assemblies

Fibrils and gels

Structurally, one of the most interesting features of short aromatic peptides is their ability to form fibrils and gels in addition to crystalline structures. In this context, fibrils are defined as highly anisotropic structures with two dimensions of the order of (a few times) the molecular size and one dimension that can be of the order of μm, and which do not show long range crystalline order. Fibril/crystal dimorphism is of course not restricted to (aromatic) dipeptides, but is also observed for longer sequences, in particular those derived from amyloid fibrils.[128,129] While uncapped zwitterionic dipeptides easily assemble into crystalline structures, capped versions display a clear preference for assembly into fibrils, even though they can be crystallised under some conditions.[92,130] In particular large, bulky aromatic capping groups, such as fluorenylmethyloxycarbonyl (Fmoc),[131] carboxybenzyl (Cbz) and naphtalene, direct the FF peptide into fibrillar structures,[132–134] whereas FF with smaller capping groups (Boc, acetate, amide) is still able to form crystals.[99,132] The fibril formation and subsequent gelation of N-terminally capped aromatic dipeptides has been studied particu-

larly extensively.[20,115,133–135] However, whether a gel or a crystal is formed is not exclusively determined by chemical modifications of the FF peptide, but also strongly depends on the solvent conditions. In the case of uncapped FF, polar, hydrogen-bonding solvents are found to favour the formation of crystals[14,87] and non-polar solvents favouring the formation of gels.[81,84,85] A theoretical thermodynamic framework to decide whether the gel or the crystal is the equilibrium structure has been developed and is based on the geometry of the self-assembling molecule and its solvophilic or solvophobic properties.[136] It is predicted by this model that the gel, usually considered a metastable phase compared to the crystal, can in fact correspond to the thermodynamic minimum, depending on the solution conditions. In cases, however, where a system under the same conditions can form both a gel and a crystal, it is usually found that the crystalline state appears after the gel state and corresponds to the final thermodynamic minimum,[99,105,134] in agreement with Ostwald's rule of stages.[99] Structurally, gel and crystal are often distinct,[99,134] suggesting that they do not interconvert directly, but that the thermodynamically more stable crystal phase nucleates either independently or catalysed by the gel (heterogeneous nucleation).

The most prominent example of an N-terminally capped FF peptide that forms fibrillar gels in aqueous solution is the Fmoc-FF system, which has been characterised in detail.[20] It was proposed that the fibrils are stabilised through aromatic stacking interactions, a design principle that has been successfully mimicked with molecules containing both a phenyl and fluorophenyl group.[137] Notwithstanding the fact that Fmoc-FF displays very robust fibril formation, it was shown that peptoid substitutions (displacement of the phenyl ring from the side chain to the amide nitrogen) of one or both of the phenylalanine residues impede (if one F is changed) or completely abolish (if both F are changed) gel formation.[138]

Gel formation in N-terminally capped dipeptides is usally induced by dissolving the peptide at elevated pH and then lowering the pH, either through titration with acid or *in situ* through the hydrolysis of an anhydride (Figure 2 and references 91 and 130). Gel formation occurs as soon as the pK_a value of the C-terminal carboxy-group is reached, suggesting

that it is mainly the globally neutral species that is incorporated into the fibrils. However, through pioneering and detailed pH-titration studies, it could be shown that the initial fibrillar assemblies are not globally neutral, but consist of molecules with an average degree of ionisation of 0.66.[135] Upon further acidification, subsequent transitions into higher order structures occur that accompany increasing neutralisation of the peptides. Through such titration studies,[130,135] as well as solution state NMR experiments[78] it became clear that the process of self-assembly actually modulates the pK_a of the peptide, with the pK_a of the assembled structures being significantly higher than that of the isolated peptide, whereby the difference can correspond to up to 5 pH units.[135] The protonation state of the peptides is also strongly dependent on the total peptide concentration.[78] If a pH titration is performed on such a peptide system, the self-assembly sets in as soon as appropriate pH values are reached, but the process of self-assembly is slow on the time scale of the titration experiments and therefore the system will not be at thermodynamic equilibrium during the titration. Indeed, the titration of self-assembling peptide systems is an ingenious way to access the thermodynamics of gel-forming systems that are otherwise difficult to study. This method was applied to characterize the thermodynamics of fibril and gel formation of the FFDD tetrapeptide and its sequence permutations.[109] Free energies of self-assembly were determined from titration curves under the assumption that assemblies only form from neutral species, an assumption, however, which might not be fully valid.[112,114,135] Nevertheless, from a comparison of the titration behaviour of a series of closely related peptides, a relative ordering of the driving forces for assembly can be deduced.[109] This type of data would otherwise be difficult to obtain, given the high stability of these fibrils and the resulting low equilibrium concentration of soluble peptide that are challenging to measure directly. In the same study,[109] one of the very few quantitative physico-chemical investigations of aromatic peptide fibril formation published to-date, the kinetics of assembly was followed by NMR and fluorescence spectroscopy, as well as small angle neutron scattering and the important conclusion was drawn that the kinetics and the thermodynamics of the assembly of the six

homologous peptides are not correlated.[109]

The thermodynamic stability of different gel-like N-terminally naphthoxyaetyl (NAP)-protected dipeptides was also evaluated qualitatively in the context of a study that aimed at the creation of a self-assembling system that displays dynamic instability.[94] The motivation behind this fascinating type of study is to artificially create a system that displays properties similar to biological polymers, in particular tubulin.[139] The competition between enzymatic dipeptide formation, peptide self-assembly and enzymatic peptide hydrolysis was investigated and it was found that only in the case of the YF peptide, a gel persisted at equilibrium, due to the high thermodynamic stability of the assemblies of this peptide, whereas YY and YL dissolved completely.[94]

In addition to pH-drop induced fibril- and gel-formation, it was also shown that N-terminally naphtalene (as well as Fmoc)-capped peptides can be gelled at high pH by adding salt ions that cross-link the rod-like structures, present at elevated pH (>10).[114] It was proposed that at high pH, it is not fibrils that form, but rather worm-like micelles with the deprotonated carboxylic acid groups pointing towards the water. These micelles can then be cross-linked by the addition of metal ions (such as Ca^{2+}). The direct interactions of calcium ions with such peptide micelles is confirmed with solution state NMR spectroscopy.[112] The concentration-dependent formation of such micelles at pH 10.5 and different peptide concentrations has been characterized in detail, using a variety of experimental techniques.[106] In this context it should be noted that the distinction between fibrils and elongated micelles is not simply a matter of semantics; indeed it has recently been shown that while peptide fibrils display a formation mechanism characterised by nucleated growth, micelles (of a different self-assembling system based on lipid-like molecules) form without an appreciable nucleation step and do also not seed further growth.[107] In has been demonstrated that it is possible to trigger the conversion of a micellar aromatic peptide system into a fibrillar system through an external stimulus, such as the dephosphorylation of a tyrosine residue in the peptide.[140] The removal of the phosphate group renders the peptide derivative less amphiphilic and hence the micelles

less stable. Detailed spectroscopic characterisation of this system suggests that the micelles are not directly converted into fibrils, but that they rather dissolve first, followed by a rapid *de novo* nucleation of fibrils.[140]

While most of the hydrogel-forming peptides and peptide derivatives so far reported have been found serendipitously, it has also been shown that coarse grained simulations allow to explore the sequence space, at least of tripeptides.[141] In this way, several previously unknown gel-forming (in the absence of organic solvent) tripeptides were predicted and experimentally verified (KYF, KYY, KFF and KYW). The important role of aromatic residues is obvious from this list.

Amorphous materials

In addition to the formation of ordered crystalline and fibrillar structures, short aromatic peptides and peptide derivatives can also form disordered structures. When screening a range of aromatic dipeptides, such as FW, WY, WF and WW, Reches and Gazit noticed that only FW was capable of ordered crystalline assembly, in agreement with a later large scale survey of published crystal structures,[61] even though the crystals were mixed with amorphous assemblies.[12] Interestingly, it has been found that FF itself at highly acidic pH values, where its solubility is very high (see above), also does not form crystalline, but rather dendritic assemblies that are likely to be disordered at the molecular level.[126] Both the high solubility and the inability to form crystalline assemblies is likely to be linked to the net charge carried by the peptide at this pH, which would need to be neutralised by the incorporation of a counter ion into the crystal. Intriguingly, the individual phenylalanine molecule can be crystallised from acidic solution (dilute formic acid) by incorporating two chemically distinct F molecules into the asymmetric unit cell: a neutral F zwitterion and a positively charged F ion associated with a formate molecule.[142]

The formation of other types of regularly shaped assemblies that are molecularly disordered can be observed for a range of N-terminally capped FF derivatives (Boc-FF and Azobenzene-

FF (PPA)). The introduction of an ethanol solution of Boc-FF into water leads to the immediate formation of rather uniform spherical aggregates (Figure 5 and references 103 and 99), probably as a result of a de-mixing process. The dependence of the formation of these spherical aggregates on the ethanol and peptide content of the solution was studied systematically (Figure 5 c). It was shown that the spheres are amorphous in structure (Figure 5 d) and represent metastable assemblies that evolve further into fibrils and finally into crystals (Figure 5 e, f and reference 99). The transformation of the spheres can be monitored directly through optical microscopy, dynamic light scattering, NMR spectroscopy, and differential scanning calorimetry (DSC)[99] and the ensemble of observations suggests that the spheres are dynamic structures in rapid exchange with monomeric peptide that dissolve once the thermodynamically more stable fibrillar phase has formed, which then grows at the expense of the spheres (Figure 5 e). Thermodynamic analysis shows that the system evolves down a gradient in free energy (Figure 5 f), as expected for a spontaneous process and that the enthalpy change associated with the dissolution of spheres and simultaneous formation of the gel or crystals shows little temperature dependence.[99] How the net negative charge of the Boc-FF is accommodated in the final crystal structure is at present unknown. A related system is the N-terminally-azobenzene-capped FF (PPA), which forms amorphous spherical structures upon evaporation from an HFIP solution.[143] Interestingly, the layer of spherical structures converts into fibrils upon incubation with water, as was studied by AFM.[143] It is unclear whether the spheres assemble directly into fibrils or whether the latter nucleate *de novo* and grow at the expense of the dynamically dissolving spheres, as in the case of Boc-FF.[99] The observation that upon contact with water the spheres become significantly smaller[143] suggests the latter possibility as the likely mechanism. Related observations have been made for layers of FF that are spin-coated onto various substrates directly from HFIP solution, leading to the formation of dendritic fibrillar structures, which upon prolonged incubation at 100% humidity convert into more ordered, crystalline-appearing structures.[144] Recently, it was shown that also the Fmoc-DOPA molecule undergoes a series of transitions

from spherical amorphous structures to a fibrillar gel and finally to a crystal upon rapid dilution of an ethanol stock solution into water.[105] In this study, the various phases of the system were structurally characterised in detail through a range of spectroscopic (CD, NMR, IR) and microscopic (cryo-TEM) methods, but no thermodynamic characterisation of the relative stability of the different phases was given.

A very interesting phenomenon in the framework of the thermodynamic stability of FF-derived assemblies is the observation that fibrils formed from C-terminally capped, and therefore cationic FF peptide transform into spherical structures upon dilution of the solution below 7 g/L.[110,145] This structural transition is partly reversible upon re-concentration of the sample.[145] A theoretical description has been presented that attempts to explain this behaviour as a result of a competition between unfavourable interfacial and favourable volume terms that are concentration dependent.[145] It is tempting to contrast this fluid-like model of C-terminally capped FF spherical structures (in solution) with the observation that (dried) N-terminally capped FF displays metal-like stiffness.[103] In our view, these observations of seemingly contradictory mechanical properties of very similar types of structures illustrate the necessity for direct measurements of the mechanical properties in solution.

Remarkably, it has also been reported that uncapped FF, albeit the D-enantiomeric variant, transforms into spherical or vesicular structures upon dilution.[88] This behaviour is very different from what is observed for the L-enantiomeric form of FF, tubes of which are simply found to dissolve upon dilution below the critical concentration.[89,118] This difference in observed behaviour could be due to the difference in chirality, but that seems rather unlikely. Therefore, this conflicting behaviour remains at present unresolved, but future detailed microfluidic studies[95] have the potential to resolve this question.

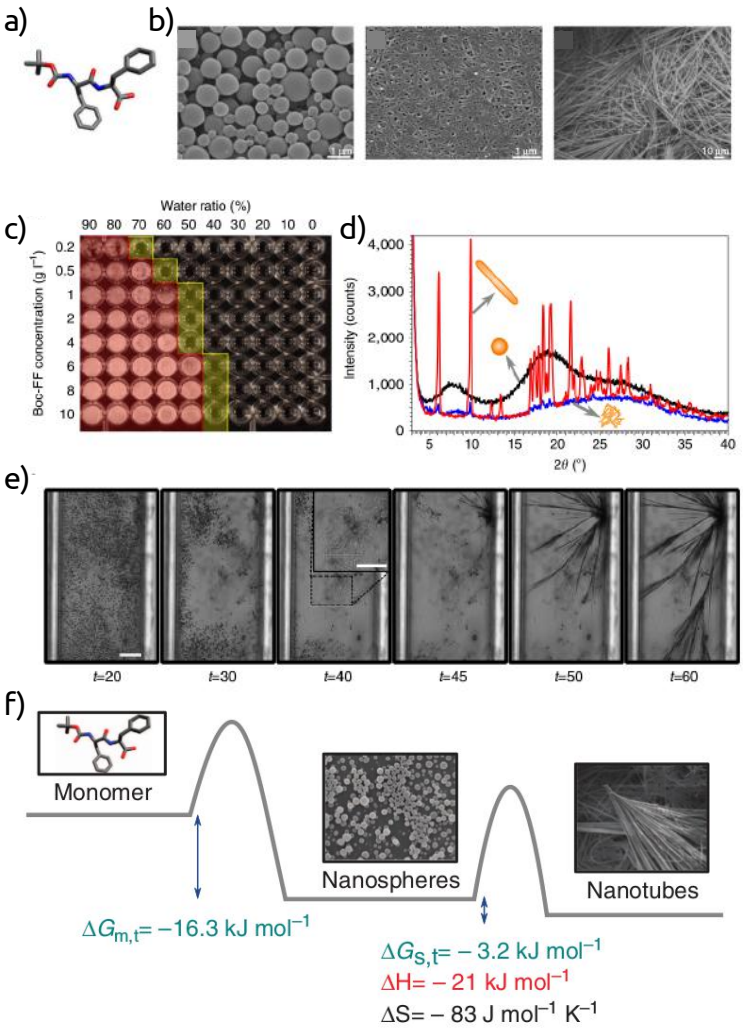

Figure 5: **The thermodynamics of self-assembly of Boc-FF a)** The structure of the Boc-FF molecule. **b)** SEM images of the different self-assembled phases that Boc-FF can form. Left: spheres, centre: gel, right: crystals **c)** Illustration of the solubility of Boc-FF as a function of peptide concentration and ethanol content. The red zone corresponds to the crystalline phase, the yellow zone to spherical structures and the black zone to complete dissolution of the Boc-FF. **d)** Powder diffraction spectra demonstrate that the spheres are amorphous and the gel is mostly amorphous as well. **e)** Self-assembly experiments inside glass microcapillaries illustrate the step-wise assembly mechanism from spheres over gel to crystals and suggest that the formation of a new phase leads to the dissolution of the less stable previous phase. **f)** Proposed energy landscape of the Boc-FF peptide, with determined free energy and enthalpy differences indicated. The system follows a down-hill free energy trajectory, as described by Ostwald's rule of stages. Adapted with permission from reference 99, ©2014 Springer Nature.

Mechanistic and kinetic description of aromatic peptide assembly

Ordered self-assembly of aromatic dipeptides leads to the formation of elongated structures, either fibrils or needle-like crystals, the latter with diameters in the range of hundreds of nanometer to several μm and lengths that can attain many μm. In the following, we will discuss the current state of knowledge of how such structures form, with, as in the remainder of this chapter, a particular focus on crystalline assemblies of FF.

Growth processes

An early model of how microcrystalline FF assemblies might form is based on the idea that sheet structures are first formed which then either roll up along one dimension to form tubes, or along two dimensions to form spherical structures.[146] However, recent experimental results from the direct, real-time study of FF microcrystals in contact with monomer solutions of varying concentrations inside microfluidic channels, using time lapse optical microscopy, clearly demonstrate the growth of the crystals through the addition of soluble peptide (Figure 5 and references 95, 118 and 89). These observations render the formation, at least of fully formed microcrystals, through the rolling up of sheets extremely unlikely. The dependence of the axial growth rate of the assemblies on the concentration of soluble FF was measured and it was found to be approximately linear at low degrees of supersaturation (see footnote [1]), while the assemblies were observed to shrink when washed with a sub-saturated solution.[118] A study which extended into a regime of higher supersaturation found that the supersaturation dependence of both the axial and radial growth rates, in a regime limited by surface processes, is exponential (Figure 6 a and reference 89). The fact that the concentration dependence of the growth rate in both dimensions was precisely known, combined with

[1]The dimensionless supersaturation σ of the peptide solution is defined as $\sigma = \frac{c-c^*}{c^*}$, where c^* is the critical concentration, see Figure 11 a for an illustration of how the critical concentration can be determined.

the ability to maintain the soluble FF concentration constant at a specifically adjusted value inside the microfluidic flow reactor, allowed tuning of the aspect ratio of the FF structures.[95] This was achieved by allowing the structures to evolve at a concentration where they grow almost exclusively longer and not wider (compare absolute rates of axial and radial growth in Figures 6 a and b), and represents one of the first examples where detailed knowledge of the assembly kinetics has been used for morphological control of assembled peptide structures. The exponential dependence of the axial growth rate of the crystals on the concentration of soluble peptide reflects the fact that the growth reaction is a nucleated process (Figure 6 c), whereby new crystal layers nucleate on the advancing phase (2D-nucleation) and spread across the face (Figure 6 d and reference 95). The absolute growth rates reported in these two microfluidic studies differ significantly; at an equivalent degree of supersaturation of ca. 0.7, the absolute growth rates are reported to be 5 μm/min[118] or 24 μm/min.[95] The reason for the discrepancy is likely to lie in the differing limiting regimes that operated in each study. Arnon *et. al.* operated in a flow regime where arrival at the growth face was transport limited. Both studies used flow reactor chambers of comparable cross section, but the flow rates varied by a factor of 200. At low flow rates (below 100 μl/h), the rate of crystal face growth was observed to be approximately linearly dependent on flow rate,[95] in accordance with diffusion-limited growth.[147]

In addition to the growth of assemblies formed from unmodified FF, the growth of cyclized FF from DMSO is also reported.[118] Interestingly, it was noted that while cyclo-FF assemblies grow bidirectionally, crystalline assemblies formed from unmodified FF display only unidirectional growth, a difference that is explained through the different crystal symmetry.

In order to gain insight into the energy barriers of the assembly process, the growth of FF assemblies was measured at different temperatures.[89] When pre-formed FF assemblies are incubated with a constant soluble peptide concentration at different temperatures, a decreasing rate of growth with increasing temperature is observed (Figure 6 e). This unexpected

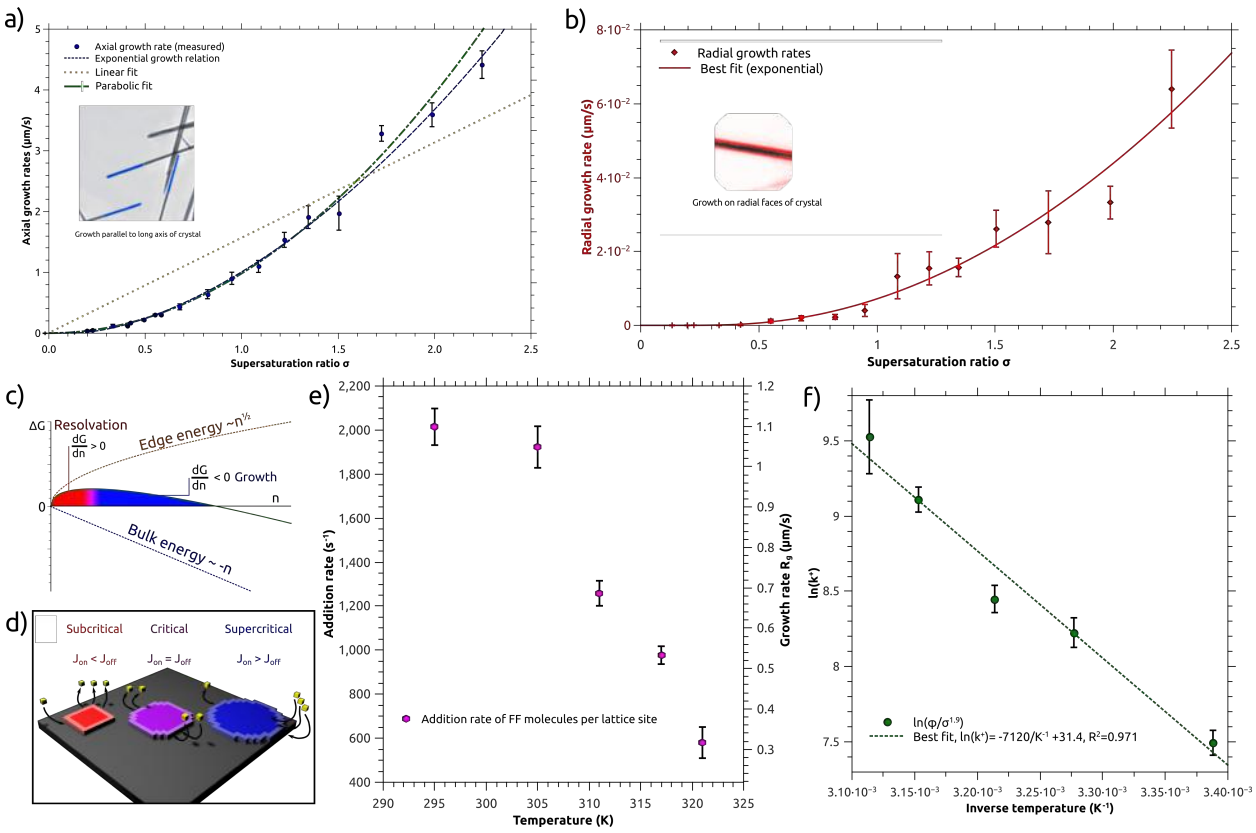

Figure 6: **The growth kinetics of FF crystals from microfluidics experiments a)** The axial growth kinetics as a function of solution supersaturation. Fits to different models (exponential, polynomic, linear) are shown. **b)** The radial growth kinetics as a function of solution supersaturation. An exponential fit is shown. **c)** Illustration of the origin of the energy barrier for two-dimensional nucleation, i.e. the competition between favourable bulk energy and the unfavourable edge energy of the two-dimensional nucleus. **d)** Illustration of sub-critical, critical and supercritical sizes of nuclei. **e)** The axial growth rates of FF crystals measured at different temperatures at a constant concentration of soluble FF (1.2 g/L). **f)** The growth rate is corrected for the different degrees of supersaturation and the growth rate constant has been determined and plotted in the form of an Arrhenius plot, which illustrates that the crystal growth is an activated process. a)-d) adapted with permission from reference 95 ©2016 American Chemical Society, and e)-f) adapted with permission from reference 89 ©2017 American Chemical Society.

result can be understood when the strong increase in solubility of FF with increasing temperature is considered (Figure 4 and reference 89). The driving force for net growth of the assemblies is the supersaturation σ, i.e. the difference between actual concentration of the soluble peptide and the critical concentration.[89,118] Therefore, at constant total concentration, the net driving force for growth will strongly decrease with increasing temperature. Given the exponential dependence of the growth rate on the supersaturation,[95] it can be understood why the observed growth rate decreases. However, if this effect is taken into account, it is found that the crystal growth process is actually a thermally activated process (Figure 4 f), as has been found for the growth of amyloid fibrils.[148]

It was recently shown that FF and Boc-FF can co-assemble and that the presence of Boc-FF slows down the assembly compared to pure FF.[149] Furthermore, the structures formed in the presence of Boc-FF are shorter, which was explained through their higher frangibility under the assembly conditions of vigorous stirring. Co-assembly, rather than separate assembly was indicated by spatially resolved time-of-flight secondary ion mass spectrometry (TOF-SIMS) measurements. Furthermore, in another study by the same authors, the mechanical properties (stiffness and Young's modulus) of FF-Boc-FF co-assembled structures were determined by AFM and it was found that the co-assemblies display a significantly lower Young's modulus than the pure FF assemblies.[150] In this study, the assemblies from different ratios of the two peptide variants were also spectroscopically characterised and a change in π-stacking and hydrogen bonding was proposed. These results show that despite the crystalline nature of the assemblies, the incorporation of foreign molecules is possible and likely leads to the formation of defects in the crystal structure.

The growth kinetics of fibrils formed by the six sequence permutations of the DDFF peptide have recently been investigated by fluorescence spectroscopy and SANS[109] and the second order growth rate constants have been determined. It was shown that depending on the sequence, the fibril growth rates differed by up to a factor of 10, assuming the same number of growing fibrils for all sequences. Given the uncertainty on the absolute nucleation rate

(see below), the latter assumption could be the source of some error. However, this study[109] represents one of the first attempts to quantitatively describe the growth rates of short aromatic peptide assemblies in a size range where bright field time lapse microscopy[89,95,118] is not possible, and where therefore ensemble techniques are more easily employed. It has, however, recently been shown that even the growth of thin aromatic peptide fibrils can be studied by confocal laser scanning microscopy (CLSM) when the fibrils can be appropriately fluorescently stained.[107] In this study it was demonstrated that an N-terminally protected triphenylalanine (FFF) peptide displayed seeded fibril growth, whereas no seeding was observed for fibrils made from a lipid-like micelle-forming hydrogelator.[107]

Nucleation processes

Nucleation is in general more difficult to study than growth, because nucleation is a rare process and because the nucleated structures rapidly evolve through growth,[151] leading in the case of short aromatic dipeptides typically to high aspect ratio structures, such as fibrils and needle-like crystals. Nucleation can not easily be directly observed, due to the often nanoscopic nature of the nuclei and the stochastic occurrence of nucleation. A phenomenological approach to describing the relative rates of nucleation and growth of crystalline structures could consider the distributions of absolute lengths, diameters and wall thicknesses (in the case of tubular crystals) of the final structures. It has been shown that these properties can be tuned in the case of FF tubes by adjusting the relative proportions of different solvents,[152] probably reflecting different ratios of nucleation and growth rates (kinetic factors), but probably also varying free energies of the different crystal-solution interfaces (thermodynamic factors). Along similar lines, the introduction of concentrated stock solutions of FF in high solubility solvents, such as HFIP[12] or acetic acid,[87] into water leads to many small crystals, whereas the slow cooling to room temperature of an aqueous solution of FF that was saturated at 80°C leads to the formation of much fewer and larger crystals.[89] The nucleation rate is therefore, as expected, highly dependent on the (local) degree of supersaturation.

This could explain why at high salt concentration, where the solubility of FF is reduced (see above and reference 89), smaller crystals are formed, but this could also be partly due to an increased nucleation rate constant in the presence of salt or indeed inhibition of growth by association of counterions at growth faces. While, therefore, general and qualitative conclusions about nucleation processes can be drawn from such observations, nucleation itself remains difficult to capture. The elusive nature of the nucleation process might explain why a range of competing models have been proposed for how aromatic dipeptides and related molecules nucleate their self-assembly. As mentioned above, it has been proposed that FF tube-like crystals and spherical structures might form through the rolling up of sheets.[146] While this is unlikely to explain the formation of fully growth structures, it cannot be excluded that the formation of the initial nanoscopic nucleus resembles such a process.

Among the different types of nucleation scenarios, the formation of dimers has been proposed on various occasions to play a prominent role.[109,113,153] Especially mass spectrometry, in the form of time-of-flight or ion mobility mass spectrometry, has been employed to demonstrate the presence of dimeric structures in solution.[113,153] Ion mobility mass spectrometry is probably the highest resolution molecular size measurement technique (short of directly determining a crystal-, NMR- or cryo-TEM-structure), and is able to distinguish between different conformations of small molecules, such as dipeptides, as well as between different degrees of water clustering around such molecules, as shown in the case of end-capped FF.[113] On the other hand, this method has the caveat of requiring transfer of the molecule or complex of interest into the gas phase. Depending on the dominant type of interaction in the complex, transfer to the gas phase can change the relative interaction strength compared to measurement in the condensed phase, which can lead to a biased distribution of species. Indeed, the question as to whether dipeptides or even single amino acids occur in solution in the form of dimers or larger clusters has been a long standing one and a summary and discussion of the debate surrounding this question can be found in.[154] In a recent study employing dynamic light scattering (DLS), it has been shown that at least in the case of FF,

no larger pre-nucleation clusters exist, as the scattering intensity-weighted size distribution was found to be dominated by a species of approximately 1 nm hydrodynamic diameter.[89] The resolution of these measurements is not high enough to decide whether this dominant population corresponds to monomers, dimers or even slightly larger species. However, it is interesting to note that the size distributions from DLS are essentially identical for strongly sub-saturated (in methanol) and strongly super-saturated solutions (in water).[89] If an equilibrium population of multimers with a finite binding affinity existed, it would be expected that its population depends strongly on the total peptide concentration. The strongly sub-saturated conditions (2 g/L in methanol) correspond to conditions where it has previously been proposed that FF occurs in the form of dimers,[153] which is compatible with the size distribution measured in reference 89.

On a more conceptual level, studies that show a significant fraction of the peptide to occur in the form of dimers or multimers[113,153] are unlikely to describe nuclei, given that nuclei, by their very definition, are transient, rare species because they correspond to the species with highest free energy along the reaction coordinate. In a study that investigated the kinetics and thermodynamics of six tetrapeptides (all permutations of the FFDD peptide), it was proposed that dimer formation is slow and therefore corresponds to the rate-determining step.[109] This conclusion is based on the observation that the fluorescence intensity of the phenylalanine residues initially increases steeply after a lag time and then slowly decreases, whereby the decrease corresponds to the formation of fibrils. However, the fluorescence signal in this case cannot unambiguously be ascribed to dimeric species and it is difficult to estimate how large the population of dimers/oligomers is, given that their fluorescence properties cannot be studied in isolation. Nevertheless, the data do strongly suggest that this system is initially dominated by the formation of unstable/metastable smaller species, followed by the further evolution of these species through growth, which corresponds to an energetic down-hill process.

Information about the thermodynamic stability of small species can also be obtained by

analysing the dependence of soluble peptide on the total peptide concentration.[89] It is found that in the case of FF, the soluble concentration corresponds to the total concentration up to the critical concentration and remains constant at higher total concentrations. Such a behaviour, analysed within the framework of linear polymerisation, suggests a large critical nucleus size and hence a highly cooperative nucleation process.[89]

A very different approach to studying the nucleation of FF microtubes has recently been taken, whereby amorphous films of FF, formed from HFIP, have been hydrated and studied by highly time-resolved Raman spectroscopy.[155] It was thereby found that the transition from amorphous to fully crystalline material involves the formation of an intermediate species. It is proposed that the entire peptide population is converted into this intermediate initially, followed by full conversion into the final state. This model is compatible with the observed growth of FF microtubes by addition of soluble peptide,[89,95,118] if one assumes a dissolution of the intermediate, followed by growth of the final structures, rather than a direct interconversion.

Overall, the currently available data suggests that gel-forming systems might display rather small nucleus sizes of the order of a dimer,[109] leading to rapid nucleation (time scales of seconds to minutes). Another aromatic peptide system that supports this conclusion is given by the N-terminally protected triphenylalanine peptides developed by Hamachi.[107] Here, a solution that is left to cool to room temperature displays rapid *de novo* nucleation of fibrils within minutes.[107] On the other hand, crystallizing systems are more likely to feature significantly larger nuclei, as indicated by the ability to create supersaturated solutions of FF that can be stable for hours.[95] Such systems are more likely to be appropriately described within the framework of classical nucleation theory.[151] A note of caution is appropriate here, however, as these statements are likely to only hold for very short sequences. For significantly longer sequences, such as the amyloid fibril-forming intrinsically disordered protein α-synclein, monomer solutions can be kinetically stable for days.[156] While this behaviour could also be indicative of a large nucleus size, it is more likely that it stems from a high

energetic/entropic cost of forming even a small nucleus.

Structure of dipeptide crystals with particular emphasis on FF

The structural analysis of assemblies of dipeptides, in particular also aromatic ones, was significantly advanced by Carl-Henrik Gørbitz.[14,61,157] In particular, he solved the crystal structure of FF crystallized from water and highlighted the importance of water inside nanochannels for the structural integrity of the crystal. In the case of FF, the occupancy of the water sites is found to be between 0.1-0.5, whereas in other dipeptide crystals, water molecules can play an even more well-defined structural role.[14] The importance of water in such crystal structures is rather general, but the water is not always found inside channels - it can also play a bridging role, such as in YW crystals.[157] This layer structure is also interesting in that it displays hydrogen bonding (albeit weak) between the ammonium terminus and the aromatic tyrosine ring. The establishment of continuous hydrogen-bonded networks in short peptide structures is a hallmark of their self-assembly, and notably all backbone hydrogen bonding interactions of uncharged dipeptides can easily adopt planar conformation, leading to layered structures and the potential for polar/nonpolar segregation.[158,159] In cases where the drive to segregate nonpolar species is strong, and, not unconnected, where the species are notably bulky, the hydrogen-bond 'planes' may adopt inward curvature, maintaining the critical connectivity while allowing the large R-groups to segregate and pack efficiently.[14,160] This can lead to the formation of nanochannels, unit-cell-level voids occupied by both solvation species and indeed non-associated solvent or other species. The energetic cost of porous structures due to loss of van der Waals interactions can be steep, so it is perhaps notable that the hydrophobic dipeptide family of molecules display such arrangements with a degree of frequency.

Hydrophobic structures in aromatic dipeptides

Two major structural classes of porous hydrophobic dipeptide crystals are known, each named for the dipeptide forming the archetypical structure. The Val-Ala class[159,161] displays hexagonal hydrophobic pores ringed by six left-handed dipeptide double helices, interacting with each other by an unusual bridging interaction involving hydrogen bond donation by the amide and C_α to a carboxylate group pointed end-on at the two donors. The structure is known for seven of the nine dipeptides containing Val, Ala and Ile, with di-alanine and di-isoleucine being the exceptions. The second major structural family is the Phe-Phe family, and it is naturally this family on which this section focuses. The key property of this family is the total segregation of the hydrogen bonding chains into mutually non-interacting tubular regions within the structure, the tubular regions being described by helices of hydrogen bonded, hydrophilic dipeptide backbones surrounding a solvated void. The tubes are completely surrounded by interacting 'sleeves' formed by the bulky side chains, and the terminal ammonium projects a proton into the channel to a solvent hydrogen bond acceptor. It is known for the dipeptides IL, FL, LL, LF, FW, WG and of course FF, the most dramatic demonstration.[14,159]

Hydrogen bond connectivity in aromatic dipeptides

A statistical study[61] of hydrogen bond connectivity in dipeptide structures showed that out of the 160 investigated crystal structures of dipeptides, 152 are uncharged zwitterions, and only 8 are charged and crystallize with associated counter ions. This finding shows that the presence of charges on zwitterionic peptides can be well-accommodated in crystals in a head-to-tail arrangement. In the survey, the hydrogen bonding is classed in terms of the unbounded chains formed by hydrogen bonding in the crystal - these chains consist of a repeating unit that can be mapped on to itself by a simple translation along unit cell vectors,[165,166] and are expressed in terms of graph set terminology,[167] in this case specifically the chain patterns denoted as $C(n)$, with n the number of atoms in the repeating

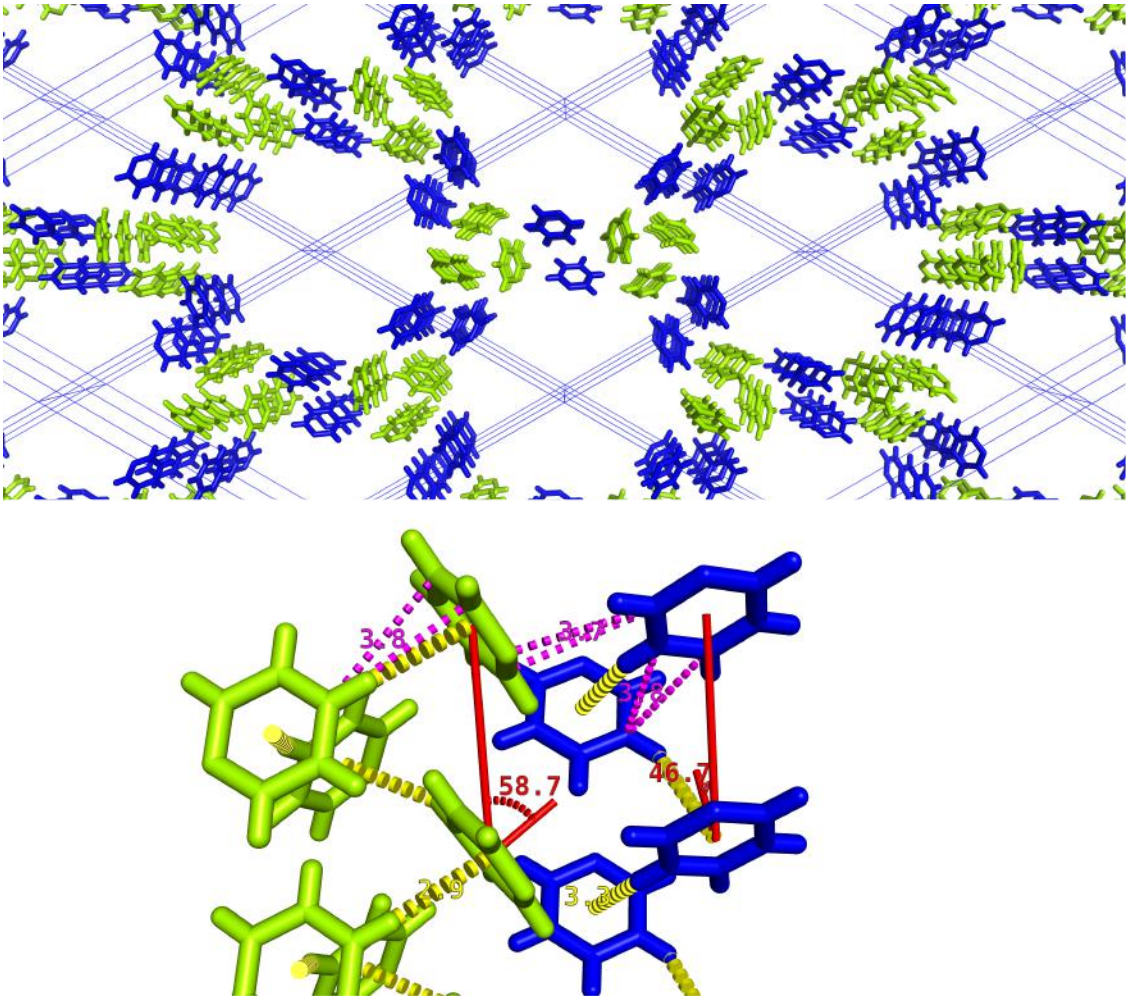

Figure 7: Top: The continuous hydrophobic domain within FF arises partly from the bulk of the side chains, shown here within the hexagonal unit cell. The polar backbones line the channel, and form essentially one-dimensional structures. F1, the N-terminal amino acid, is in blue, and F2 is in green. The asymmetric unit cell contains only one FF, and so there is one F1 environment and one F2 environment, each F1 or F2 site being interchangeable with other F1 or F2 sites by symmetry operations of the $P6_1$ space group. Bottom: Geometry of interactions within the hydrophobic domain shows a remarkable helix structure for F2 sites akin to McGaughey's energy-minimised pinwheel for the benzene trimer,[68] linked by T-shape type aromatic-aromatic interactions (R_{cen} = 5.1 Å, close to the empirical Phe-Phe pair minimum[68] and the calculated minimum[66]), with closest approach through the carbon 4 C-H$\cdots$ π bond (yellow) at a distance from H to the ring centroid of 2.9Å. The comparable interaction for F1 does not have the proton pointed directly at the ring, but retains an offset T-shape geometry. The measurement for closest carbon-carbon approach (magenta), often used in automated systems for the discovery of interaromatic contact, is 3.8Åon both the F1 and F2 sites. The F1-F2 interactions are shown at 3.7Åbetween C2 of F1 and C1 and C2 of F2. The vertical spacing of each site is naturally the lattice parameter c, 5.46Å, at angle θ (in red) 58.7° on F2 and 46.7° on F1. This geometry suggests a lack of $\pi - \pi$ interaction along the long axis of the crystal.[162–164]

unit. The eight-membered repeat unit is by a long way the most common motif, and in this case the eight atoms involved are those along the dipeptide backbone, corresponding to the "head-to-tail" charge-charge interaction pattern, with surveyed dipeptides showing up to three independent eight-membered chains. Of the 160 structures surveyed, only 15 did not display a C(8) chain. FF displays two, remarkable for their helical shape - one helix is right handed, and one unit cell in length, the other left handed and five unit cells in length. A critical characteristic of FF in terms of its relevance to amyloid fibrils is that its crystal grown in aqueous solution *does not show* a C(4) chain corresponding to inter-amide hydrogen bonding. The central amide is instead involved in the more common C(5) pattern, with the amide proton closely approaching the carboxylate terminus rather than a neighbouring amide. A more conventional 'amyloid-like' hydrogen bond structure is achieved by the FF-methanol solvate,[87] which displays C(4) inter-amide and C(8) head-to-tail patterns (the C(8) displaying a twofold rotational axis normal to the backbone plane). Notably, the short C$\alpha\cdots$O=C distance, demonstrating a hydrogen bond,[168,169] is a familiar feature in parallel β-sheet structures. This connectivity is, interestingly, exactly that observed for the dipeptide di(phenylglycine) in its inclusion compounds with certain chiral sulphoxides.[170] Its aqueous self-assembly process yields a spherical aggregate, suggesting a lack of long-range crystalline structure.[146] Conversely, successive prolonging of the aliphatic linker between the phenyl ring and the peptide backbone yields non-hollow elongated structures, as well as plate-like structures, which reflects presumably the different crystal packing, as well as a different balance of crystal nucleation and growth rates.[171] The chaotic response of dipeptide structures to even small changes in molecular identity has been observed frequently,[159,172] and a number of ring-substituted FF analogues are known: substitution at the 4 positions of the ring with fluoro, iodo and nitro groups yield extremely narrow tubular structures, similar in appearance but remarkably different in their structure. FTIR indicates antiparallel β-sheet structure for the halo-substituted dipeptides, while di-(4-nitrophenyl)alanine has a distinct signal consistent with random coil connectivity, possibly suggesting involve-

ment of the nitro group (a strong hydrogen bond acceptor) in the polar/hydrogen bonded network of the aggregate. Di-(1-naphthyl)alanine and di-(2-naphthyl)alanine yield flexible tubular structures showing β-sheet FTIR signals, and di-(1-biphenyl)alanine crystallises as square plates.[173] The formation of plate-like structures has also been observed for the FFF tripeptide.[102] Based on molecular simulations, those structures are reported to be thermodynamically more stable than FF tubes, but this result is not experimentally verified. It is, however, plausible, given the trend in solubility observed between F and FF (Figure 4).

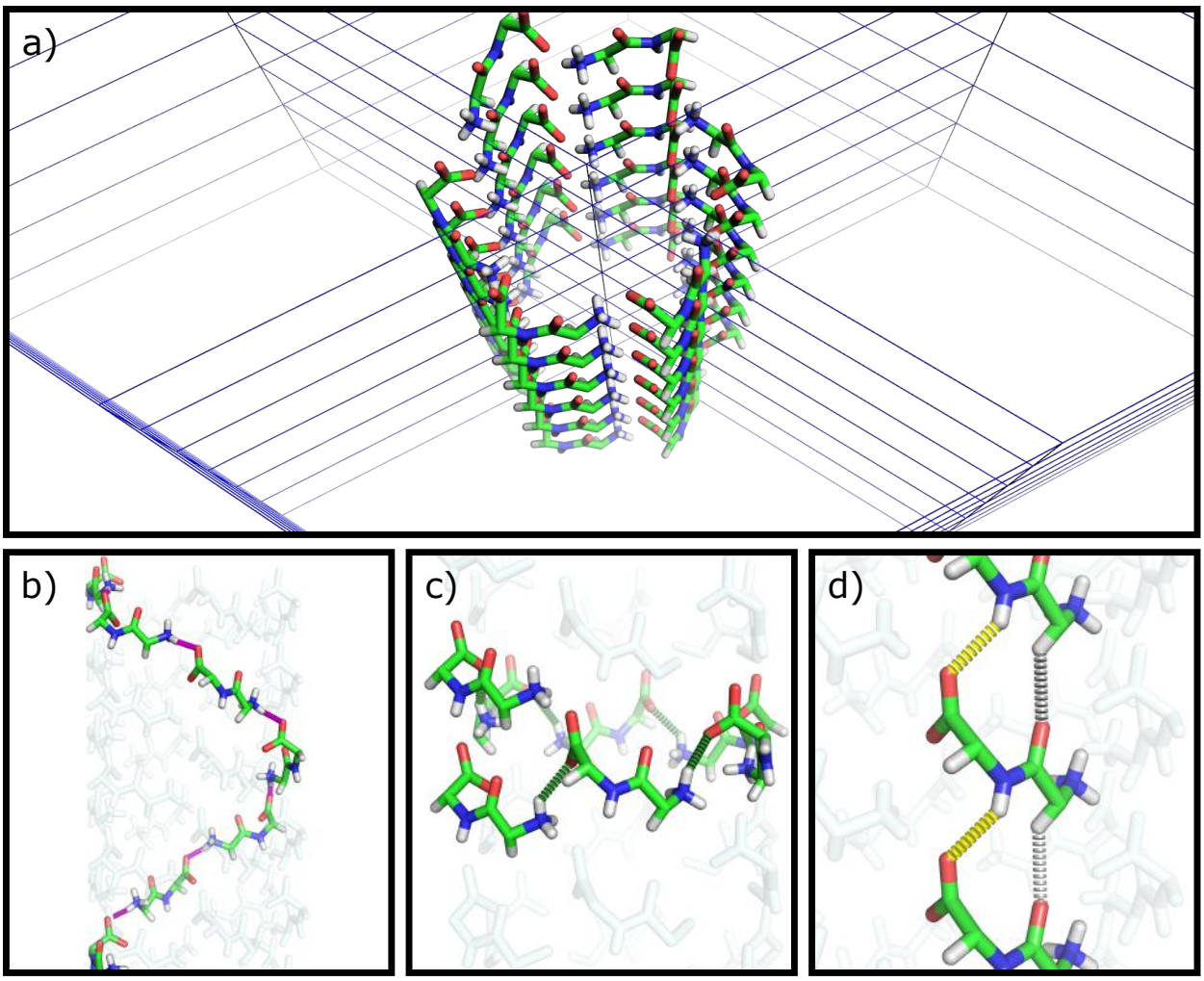

Figure 8: a) Hydrogen bonded network in FF, superimposed on the unit cell. The hydrogen bonds form closed structures due to the curvature imposed by the bulky side chains. b) Six-membered left-handed helix, C(8). c) Six-membered right-handed helix, C(8). d) Axial chains, the strongly-bonded C(5) chain in yellow and the weaker Cα-carbonyl, a C(4).

Importantly, by comparing experimental and simulated powder diffraction data, it was

also shown by Gørbitz that the molecular arrangement of large FF crystals are identical to the much smaller ones obtained by introducing concentrated solutions in organic solvents, such as HFIP,[12] into water.[174] This conclusion gained further support when FF assemblies of different sizes (sub μm thickness and wider than μm) were structurally analysed by AFM and polarized Raman microspectroscopy and no differences were found.[175]

Macroscale aggregate structure

The overall tubular appearance of FF crystals grown from water is one of the incredible traits of the diphenylalanine aggregate, almost seeming as though its habit recapitulates its unit cell down to the incorporation of a solvated void! Naturally, this phenomenon was the cause of much early discussion[12,14] and was exploited by, for example, filling the core with elemental silver.[12] It is currently still unresolved why FF forms hollow crystals, but a potential explanation can be found in reference 176, where it was proposed that hollow crystals from small molecules can be formed when the growth rate of the crystal is faster than the arrival of peptide building block through diffusion. In this case the hollow core simply reflects the insufficient delivery of peptide building blocks in the centre of the growing crystal. In the light of this hypothesis, it can also be understood that under certain solution conditions, solid rods rather than hollow tubes can be formed from FF for example by titrating a solution of FF in TFA with ammonia solution,[90,177] or by sonication.[177,178] The sonication process is likely to yield a more homogeneous peptide solution and prevents the built-up of concentration gradients that can lead to the formation of hollow tubes.[176] Interestingly, it was reported that the molecular arrangement of tubes and rods is not identical.[177] It is suggested that the different types of structures can be inter-converted,[177] but a more detailed look at the protocols involved ("annealing" at 80°C, 2 g/L) suggests, based on the solubility of FF (see above) that the rods are dissolving and tubes are subsequently nucleating *de novo*. That differences in solution conditions should lead to the nucleation of different structures and morphologies (high peptide concentration and high ionic strength in

the case of the rods and lower concentration and slow increase in supersaturation through cooling in the case of the tubes) seems overall more likely than a large scale, concerted rearrangement of an entire crystal. That said, it has been proposed elsewhere that transient structural changes can be induced by ultrasonication,[179,180] and these changes are not attributed to heating of the sample. It has to be noted, however, that sonication can lead to significant local heating and that the overall temperature increase of the sample is not necessarily representative of the maximum temperatures locally reached inside the sample.[181]

Comparison with the assembly of longer sequences into amyloid fibrils

Structural comparison

In light of the remarkable ability of the tetrapeptides KFFE and KVVE to form unambiguous amyloid fibrils,[77] a property shared with a number of short, hydrophobic peptides containing phenylalanine,[182–184] and the presence of neighbouring phenylalanine residues in an aggregation-critical sequence of the amyloidogenic Aβ peptides, the possibility that FF represents an instructive model system in which interactions key to the early stages of amyloid aggregation could be simulated is a tantalising one, and indeed was a major driving force in much of the research into FF and similar short aromatic peptides.

There are certain superficial structural similarities between the water-grown crystal of FF and particular amyloid systems, outside of their compositional similarities - side chain interaction and close packing/interdigitation is a common property of cross-β amyloid structures,[185] and is clearly a strong influence in FF self-assembly. We have already discussed the tendency of phenylalanine to occur in β-sheet-forming regions of native chains and the β-sheet is, of course, the fundamental structural element of the amyloid fibril, and any tendency to pro-

mote β-sheet formation can, with a few caveats, be seen as a tendency to stabilise amyloid structures. Inspection of the Ramachandran plots for the aromatic amino acids shows the extent of the steric restrictions imposed by the involvement of the γ-carbons in a rigid ring system. Application of the Ramachandran plot to dipeptide conformation is problematic, however - the angles are dihedral angles between amide CO-NH bonds and the two backbone bonds of the C_α; ψ is defined as the C_o^{i-1}-N-C_α-C_o dihedral angle, while ϕ is the N-C_α-C_o-N^{i+1} dihedral angle (Figure 9, a, b). Both are measured clockwise around the central bond in the N-terminal to C-terminal direction. Obviously, FF only has one amide bond, and the C-terminal is a carboxylate and the N-terminal an ammonium, strictly meaning that the N-terminal Phe only has a defined ψ (156.9°), and the C-terminal only a defined ϕ (54.7°). At the cost of some assumptions, replacing a carboxylate oxygen with an imaginary amide yields two potential ψ values for F2, at 44.5° and -138.8°. A ϕ angle of 156.9° (F1) is associated nearly exclusively with β-sheet secondary structure, while the situation for F2 is more unusual - the positive ϕ angle, a consequence of the cis arrangement of side chains, is more rarely seen in native protein structures and is typically associated with a left-handed α-helix[186] (for ψ=44.5°) or an unclassified structure[187] for ψ= -138.8°. Inspection of Figure 8 shows that the conformation where the first F is in a sheet-like conformation and the second in a left-hand helix or the unclassified conformation is relatively rare - of the more than 16000 phenylalanine residues in FF or FFF sequences sampled (see footnote[2]), only 40 show geometries approaching the ones observed in diphenylalanine crystals (Figure 9 c). A cyclic peptide RACAFFC containing the FF motif was the subject of a study into its activity as an antagonist at a chemokine receptor,[190] and measurement of the ϕ, ψ angles of the two phenylalanines, constrained as part of a five-membered ring closed by a disulphide bond shows they are similarly not in a conformation associated with secondary structure (F6 does show angles typical of a β-sheet, but F5 is well outside the conventional limits). The

[2]A search was performed on the PDB[188] for the "FF" sequence motif, where it occurs as part of an unmodified sequence, at resolution $< 2.5\mathring{A}$ and without ligands. The sample was analysed using Pymol 2.1.0[189] via the get_phipsi command and the output plotted using Gnuplot.

amyloids formed by KLVFFA, however, show that both phenylalanines are part of a continuous β-sheet structure in all polymorphs and have dihedral angles well within the expected bounds,[191] and the same applies for the oligomerising β-hairpin structures formed by longer core fragments of Aβ.[192,193]

The helical connectivity of the hydrogen bonding chains has some similarities to a proposed model of polyglutamine aggregation advanced by Perutz et. al.[195] contemporaneously with early, but highly influential, reports on the properties of FF,[12] though the β-helix model for polyglutamine species has not stood the test of time.[196,197] The parallel arrangement of the two-membered chains might be seen as something of a departure from expected behaviour for *short* polypeptides forming amyloid fibrils, due to the increased importance of terminal charge complementarity, which forces antiparallel organisation in KFFE and other asymmetric short peptides with charged residues.[77,198] A critical factor in the topology of the aggregate structure is the twist observed for amyloid fibrils - this phenomenon is due to the steric effects of the chiral backbone.[199] Only if $\phi + \psi = 0$ on average along the extended chain in the sheet are the amide groups parallel. In general, the average has a small finite value, resulting in a tendency towards a progressive deflection of the polar direction of the amide groups along the chain, and hence to a helical twist around the axial direction of the polymeric fibril or β-sheet.[200] This helical twist effectively limits the radial dimension of the fibril, as fibril addition sites further from the central position experience progressively greater elastic strain. The equilibrium radius of the fibrils arises from the balance between this elastic deformation cost and the free energy of radial addition to the fibril.[201] Polymorphism in radial dimension, helical twist pitch, handedness and indeed in coiling of the fibril into nanotube-like ribbons have been observed.[202,203] The limited width, due to mechanical effects, is a key property of amyloid fibrils; indeed, it is a major reason why they have a fibrillar morphology and critically it distinguishes the fibril from the crystal. This was shown in a theoretical study, where the free energy as a function of the twist angle was treated

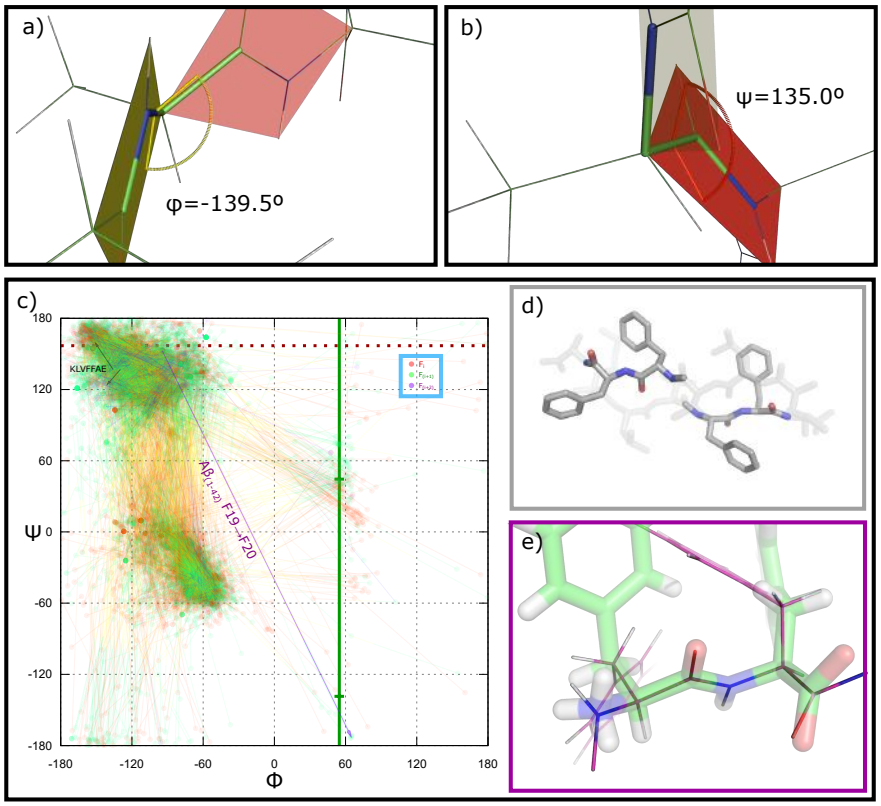

Figure 9: **The Ramachandran angles of the FF sequence in a statistical sample a)** Ramachandran angle ϕ, measured in a right-handed fashion in the N-to-C terminal direction. The angle shown is typical of an antiparallel β-sheet. **b)** The angle ψ, again with a value typically seen in the antiparallel β-sheet. ϕ and ψ values are restricted by steric effects to ranges of 'allowed' values. **c)** A statistical sample from the Protein Data Bank of 8044 occurrences of the unmodified FF sequence in 3087 PDB entries, collected by the authors for this work. Direction of the chain is shown by the colour gradation. In nature, the sequence has a relatively typical Ramachandran distribution, with a tendency for a first residue in a helical-type structure to result in a second residue with sheet-like conformation. Two black arrows represent the pair of FF sequences in PDB 3ow9,[191] a polymorph of KLVFFA with typical anti-parallel β-sheet conformation (Panel **d**) The dipeptide crystal has ϕ_1=156.9°, with ψ_1 undefined. The dotted red line shows the ϕ angle on F1. The carboxylate terminus of F2 allows the guessing of two possible "ψ angles", for ϕ_2=64.7°. Recognition of a similar structure with a sheet-like first residue, and a second in a left-handed conformation, reveals the curious fact that FF (Aβ_{19-20}) adopts a similar backbone structure to PDB entry 5NAO, Aβ_{1-42}, as solved by Wälti.[194] **e)** Superposition of the FF crystalline conformation (green) over the conformation of the sequence in Aβ_{1-42}.

as a perturbation in accordance with the Landau theory of phase transitions, a mean-field theory whereby the perturbation can be expanded as a Taylor series.[204] In the case of the twisting of amyloid fibrils, the chirality of the individual chains requires the retention of the odd-powered terms in the expansion which would otherwise cause asymmetry about the transition point, though the linear term is shown to have a zero coefficient. The analysis in the paper proceeds to identify the stationary points of the series expansion, with a minimum at zero twist (a crystalline state), and a minimum at a specific twist angle, this twist angle being dependent on the equilibrium twist angle (the twist angle at zero fibril thickness), and the ratio of the torsional spring constant of the fibril (a function of the square of fibril thickness) to the spring constant for rotations between individual chains in the fibril (going as the cube of the chain length for long chains). A critical fibril thickness was established, above which the crystalline form would be stable.[204] This critical width scales linearly with chain length, and is of course also dependent on other material parameters, many of which are common to all amyloid structures. It is to be noted that the origin of the asymmetry, the equilibrium chiral twist of the amide backbone, cannot exert influence in the dipeptides - only one amide bond exists, and so it is impossible for there to be an equilibrium series of amide rotations about the backbone. The ammonium and carboxylate termini are freely rotated, and interact by a monopole-monopole force that, inherently, only has a distance and no directional dependence.

A crystal is defined by its repeating, three-dimensional motif and in principle, any translation along an integer number of unit cell vectors will lead to an identical environment. This property has been confirmed for two distinct FF solvomorphs[87,174] among many other short peptides[61] by virtue of the analysis of X-ray diffractograms. Diphenylalanine crystallises from water into the space group $P6_1$, which on the face of it has a helical appearance, borne out by the angled C(8) chains which spiral the inside of the nanochannels. What is not helical, though, is the vertical (axial) connectivity, the C(5) chain- each step along a

multiple of the short axial unit cell vector brings you to an identical diphenylalanine. This argument cannot be made for the amyloid axial translation unless the pitch of the helix was a precise multiple of the hydrogen bond separation. Even though the repeat distance is highly consistent along a fibril,[205] it cannot be said to represent a crystalline repeat due to the far lower energy of minute fluctuations over the repeat distance. Still harder to equate to the crystalline regime are the radial 'vectors' in amyloid fibrils. For these sites, as has been discussed, energy and concomitantly equilibrium shape change with distance from the fibril centre, and there exist defined upper limits for the radial dimension. Nevertheless, the attachment of monomeric peptide onto the surfaces of amyloid fibrils of the Aβ peptide has been demonstrated to be the first step of the autocatalytic secondary nucleation process and the affinity has been determined from quantitative surface plasmon resonance (SPR) biosensing studies to be up to two orders of magnitude weaker compared to the affinity for the fibril end.[206,207]

Comparison of assembly kinetics and thermodynamics of short aromatic peptides and longer amyloid forming sequences

It has been shown that in diphenylalanine, the radial growth rate is a continuous, monotonic function of excess concentration of the dipeptide in water (Figure 6 b and reference 95, while in the amyloid fibril continuous radial growth is not normally observed.[208,209] The aggregation process of FF, in common with other crystalline species, is highly cooperative. Concerns of edge and surface energies determine the stability of 'islands' of a new layer at a surface (Figure 6 d) and give rise to complex, higher-order dependence of the incorporation rate on the excess arrivals of monomer at the surface over the off-rate of resolvating species (Figures 6 a and 10 a). In contrast, the kinetics of elongation of amyloid fibrils have been shown in a range of studies to depend linearly on the concentration of soluble peptide at low concentrations and saturate at higher concentrations (Figure 10 b and references 210, 211), and this distinct concentration dependence has been interpreted as strong evidence for a mechanism

whereby the fibrils elongate through the addition of monomeric, rather than oligomeric building blocks.[210] The effect of monomer concentration on axial growth rate is a readily-identified difference between crystalline and fibrillar systems- crystalline systems display cooperativity at moderate supersaturation, while fibrillar systems transition from non-cooperative (linear in supersaturation) to anti-cooperative (sublinear) regimes at a supersaturation determined by internal dynamics of the attached monomer during incorporation.[211]

The difference in growth mechanisms between amyloid fibrils and crystals renders a direct comparison of the absolute growth rates and in particular of the underlying free energy barriers not straightforward. Nevertheless, a first idea can be gained by comparing the absolute rates of growth of these structures with the theoretical maximum of the assembly rate, the diffusional arrival of the peptide into a reaction volume of size comparable to the soluble building block. In Figure 10 c such a comparison is shown for the FF peptide and the Aβ(1-40) peptide. The difference between the actual rate and the theoretical maximum is approximately a factor of 10^3 in the case of Aβ(1-40) across a range of concentrations, and due to the saturation of fibril elongation this factor increases at high concentrations. On the other hand, the exponential dependence of the axial growth rate of FF leads to a strongly decreasing difference between the rates of diffusional arrival and growth. If the difference between maximal and actual rates is interpreted as stemming from a free energy barrier, the latter is therefore concentration dependent.[89]

All faces of the crystal have been shown to grow continuously at finite rates in suitable conditions, but another comparison with the amyloid system may be made in that the crystals of FF have extreme aspect ratios, and through careful control of solution conditions, can be made to grow axially with extremely slow radial growth,[95] a phenomenon arising from the exponential dependence of growth rate on supersaturation on each face. The observation is explained in terms of the intermolecular interaction chains coplanar with the initial deposition of a new face. In FF, the radial faces may either intersect the helical hydrogen bonded channels, which will readily stabilise incoming FF, or they may present a simple hydrophobic

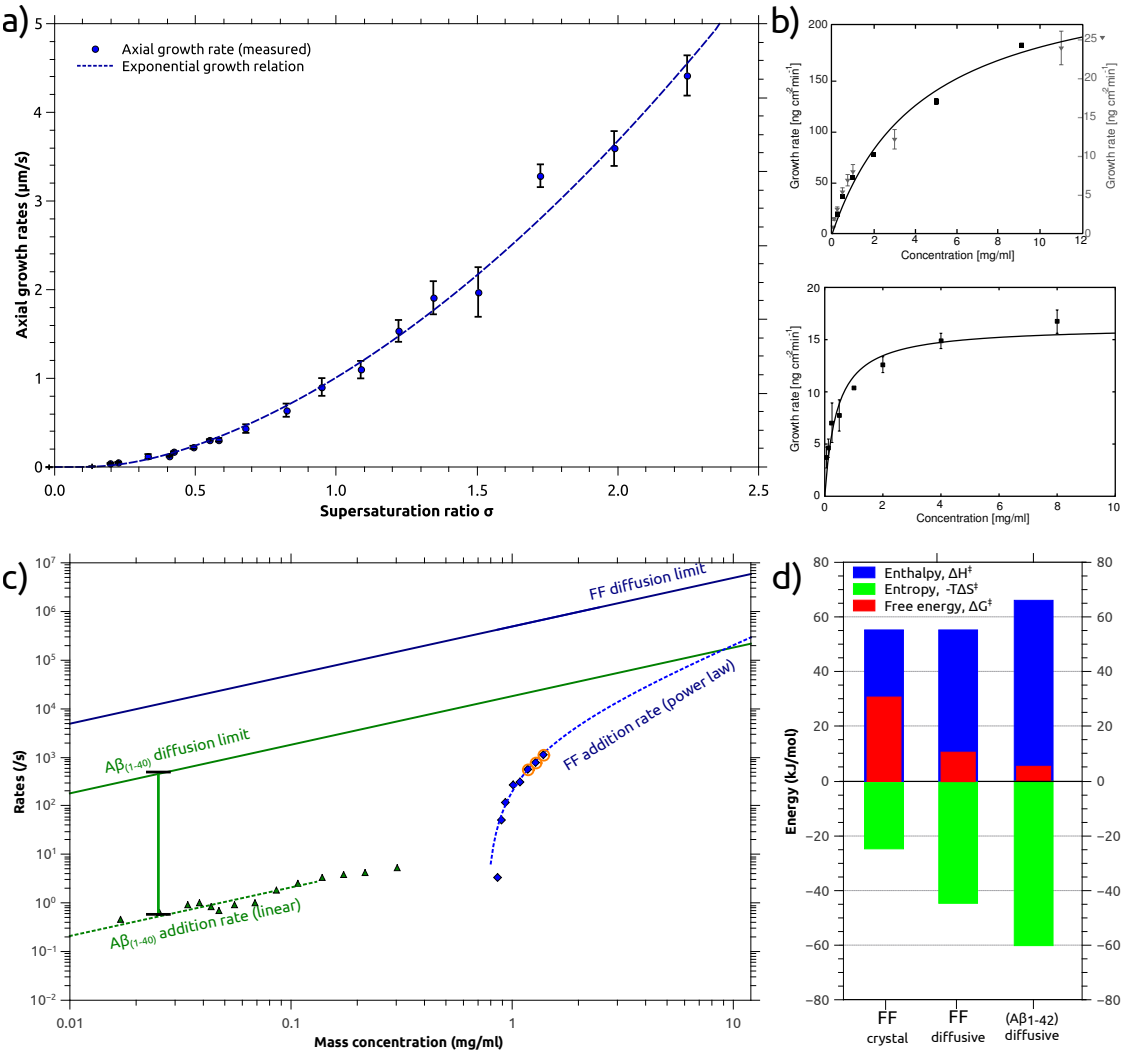

Figure 10: **Comparison of the concentration-dependence and free energy barriers of peptide self-assembly into amyloid fibrils and crystals a)** The axial growth rates of FF crystals as a function of solution supersaturation σ displays an exponential dependence. **b)** The dependence of amyloid fibril elongation rate on the solution concentration of insulin under two different solution conditions (top)- the growth rates in 100 mM NaCl / 20% acetic acid (squares) and in 10 mM HCl with no added salt[211] can be scaled to the same master curve, despite an order of magnitude difference in absolute rate and the concentration dependence of fibril elongation rate in the longer chain α-lactalbumin amyloid system (bottom), showing onset of saturation at lower concentration.[211] **c)** Loglog graph comparing the aggregate growth rates of amyloid fibrils of the Aβ(1-40) peptide with those of FF crystals. The solid lines represent the diffusional fluxes of monomers into a reaction volume on the end of the aggregate and correspond to the theoretical maximum rates. **d)** Comparison of the free energies, entropies and enthalpies of activation of FF crystal growth with those of amyloid fibril growth of the Aβ(1-42) peptide. The kinetic data of FF has been analysed in two different ways, using a diffusive model for direct comparability with amyloid fibril growth and also by taking the crystalline nature of the assemblies explicitly into account. Details of this analysis can be found in reference 89. a), c) and d) adapted with permission from reference 89 ©2017 American Chemical Society and b) adapted with permission from reference 211 ©2010 American Physical Society.

surface, compare Figures 7 and 8. The close-packed hydrophobic surface is a relatively low energy face, and so the borders of new layers on the axial face come at a rather lower cost than their counterparts on the radial face, which must always expose unsatisfied pendant hydrogen bonds. As a result of the greater nucleation rate of new layers, the axial face grows faster. Certain theoretical models have seen success in explaining amyloid nucleation and growth rates in terms of pseudo-one-dimensional crystallisation[212,213]- crystals where the energy difference between the faces is such to render radial growth impossible for wide ranges of conditions, though it should be mentioned that the Gibbs-Wulff theorem, to which analogy is made for the equilibrium aggregate shape, is only demonstrable in aggregates capable of rapid equilibration, i.e. very small ones. Kinetic processes are dominant determinants of crystal shape at the macroscale, and the early stages of polypeptide aggregation frequently involve non-fibrillar oligomeric species, for which the crystal theories of amyloid formation do not account.

For shorter species, the line between a crystal and an amyloid fibril can be a narrow one, and the classic examples are the Sup35 fragments GNNQQNY and NNQQNY, a heptapeptide and hexapeptide which, like their parent chain, readily form amyloid fibrils.[128,214] Unusually, though, a crystalline polymorph is also observed in each case at slightly higher monomer concentration, this crystal having very similar structure to the fibril as determined by X-ray diffractometry (the 4.7 Å meridional reflection being particularly notable), and an interesting point is made that the crystals were also highly acicular and despite the efforts of the researchers, crystals could not be grown continuously in at least the radial dimension. A microcrystal of NNQQNY is shown, with a width slightly over 1 μm, corresponding to approximately 500 unit cells, and appears to have a somewhat irregular surface, suggesting a high density of grain boundaries parallel to the long axis.[7] Other short peptide polymorphic systems have been prepared, and they have a notably wide range of side chains[215]- NNQQ, a tetrapeptide with purely polar residues, fibrillises, as does penta-glutamine, results in keeping with observations of the pathology of Huntington's chorea, a neuropathological

amyloidosis triggered by extended polyglutamine repeats.[216] The same work established a fibrillar state for the hexapeptide SSTSAA, a fairly hydrophilic fragment with short side chains, as well as larger, more hydrophobic species. A key finding of this extensive survey into amyloidogenic fragments is the concept of "steric zippers". These fragments of longer chains are computationally predicted and experimentally confirmed to be amyloidogenic in their own right, and as such are considered to be potential candidates for "trigger" regions of the parent chains, whose intermolecular interaction can lead to the nucleation of the fibrils.[217,218] The fragments proved capable in many cases of enhancing fibrillisation of the parent polypeptide, and vice versa. In the diphenylalanine case, the identity of the dipeptide as Aβ_{19-20} is frequently remarked upon in studies of the species, and recent NMR-derived structures of Aβ_{1-42} show the motif buried in a hydrophobic "pocket", interestingly in a similar 'cis' configuration of the side chains.[194,219] However, a recent cryo-EM study showed F20 oriented out of the hydrophobic 'core' containing F19.[9] In this study, the fibrils were formed in 30% acetonitrile solution. This alteration of the solvent environment may render the projection of F20 at the surface less unfavourable than in the aqueous/salt systems of references 194 and 219. The Ramachandran angles are ϕ_{F1} = -94.7°, ψ_{F1} = 152.7°, ϕ_{F2} = 64.7°, ψ_{F2} = -172.5° in the structure as given by Waelti *et. al.*. Interestingly, these correspond very closely to an equivalent FF conformation in the crystal, and a glance at the structural solution shows the effect of the same forces. The hydrophobic residues are buried, and indeed interact in a clear herringbone-type chain typical of the π-π interaction, even at significant cost to the continuity of the hydrogen bonding between successive chains.[194] The effects of hydrophobicity are readily discerned from the convoluted structure of the chain in the fibril. Two separate hydrophobic pockets are formed by each of the chains forming the dimeric 'unit cell' in the axial direction, one involving the residues 16-22 (KLVFFA), and the second involving interdigitating aliphatic side chains from residue 31 onwards.

Comparison of the thermodynamics of dipeptide and polypeptide aggregation is an important step towards rationalising or discounting the employment of short peptides as model

systems for amyloid propagation. In particular, the desolvation of the aromatic benzyl groups is predicted to contribute a significant negative free energy to the process of amyloid fibril formation. In polypeptide systems with hydrophobic residues, the enthalpic component of the free energy of desolvation (upon folding and burial of the hydrophobic residues) is positive at low temperature, transitioning through zero to give a negative ΔH at high temperature.[220] The free energy of solvation remains relatively unchanged, however, indicating an entropic contribution to the free energy that is nearly equal in magnitude and opposite in sign. This compensated relationship is diagnostic of the influence of the hydrophobic effect.[221] In the case of FF, the signature of the free energy, as well as its compensated nature correspond very well to that of a system dominated by the hydrophobic effect (Figure 11 b).

The dependence of the free energy of fibril elongation, per peptide unit, on the length of the polypeptide has been investigated for a range of amyloid-forming systems, and an empirical power-law relationship governed by contact area between chains was proposed to describe the scaling of this property (Figure 11 c and reference 222). The free energy was established by depolymerisation studies with a strong denaturant and analysis of the data with the linear polymerisation model,[223] where the equilibrium constant for the addition to fibrils of all lengths reduces to that between fibril ends and monomers. Diphenylalanine displays markedly greater non-polar side chain surface area per peptide than typical biologically relevant polypeptides, and the π-stacking interactions will represent a significant portion of the side chain interaction energy, explaining the comparably higher stability of FF assemblies shown in Figure11 c.

Finally, the temperature dependence of the axial growth rate of crystals formed by the FF peptide (Figure 6 f), combined with theoretical analysis in the framework of both diffusive and surface nucleation models has also been used in order to define the enthalpic and entropic character of the free energy barriers of this process.[89] As mentioned above, within the framework of crystal growth theory, the free energy barrier is concentration dependent. The free energy barrier shown in Figure 10 d has been calculated from the measured absolute

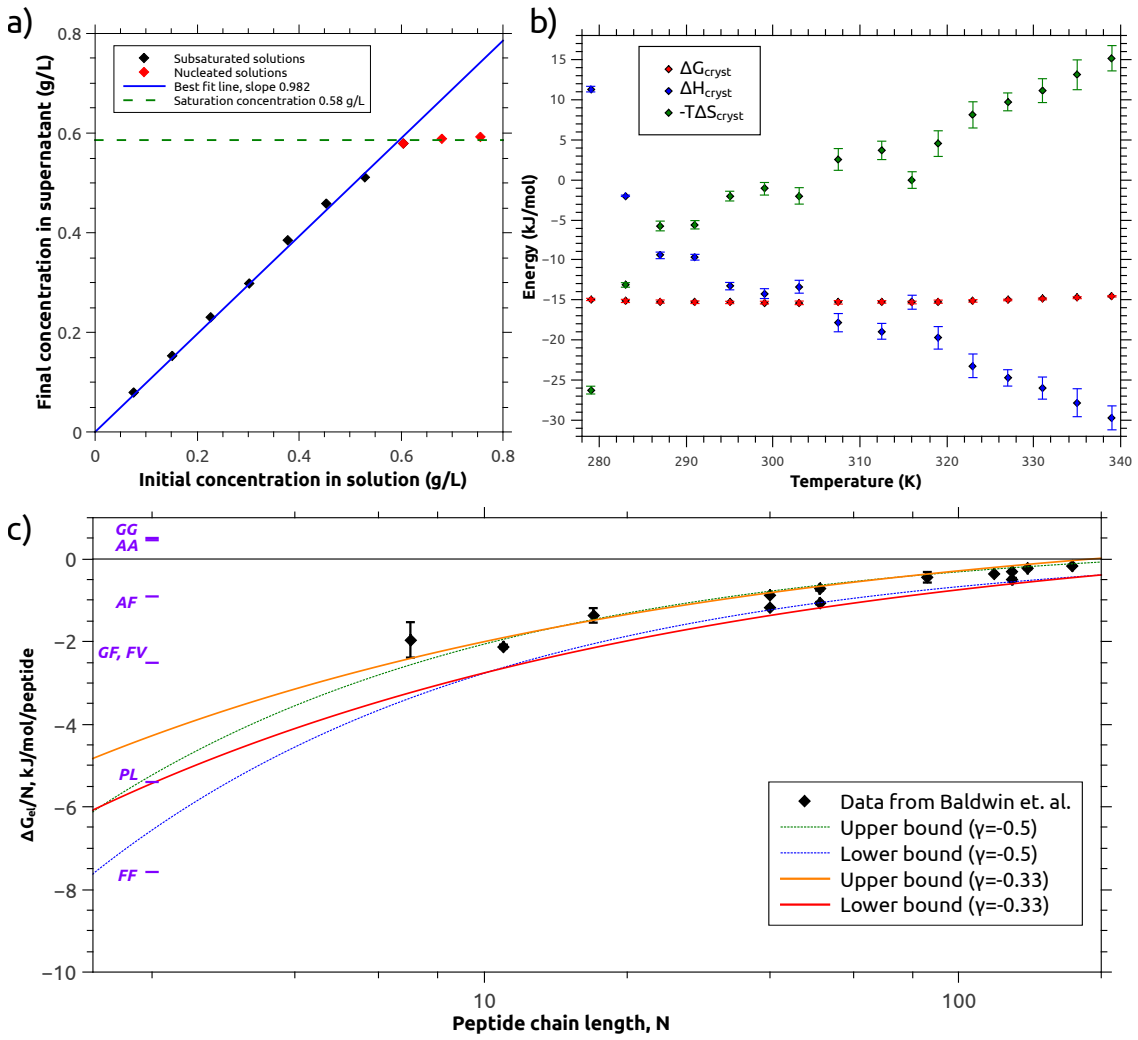

Figure 11: **Comparison of the thermodynamics of crystallisation with that of amyloid fibril formation a)** Supernatant concentrations, following centrifugation of samples equilibrated at 293 K, are plotted as a function of varying initial peptide concentration. The abrupt change in behavior, from linear with slope 1 to independent of initial concentration, at a value of 0.58 g/L demonstrates the absence of association below the critical concentration. **b)** Overall thermodynamics of FF crystallization as a function of temperature, derived from the solubility data in Figure 4. The entropy-enthalpy compensation of the crystallisation process is clearly apparent. **c)** Peptide self-assembly thermodynamics and the limit of short sequences. Data from reference 222 with addition of the dipeptides FF, AF, FV, AA, GF, GG and PL. Figure adapted with permission from reference 89 ©2017 American Chemical Society.

growth rate at 295 K shown in Figure 6 e (1.2 g/L soluble FF, corresponding to $\sigma \sim 1$ at this temperature). Interestingly, despite the fact that the overall mechanism of assembly is clearly distinct between amyloid fibrils and crystals (see above), the magnitude and energetic signature of the free energy barriers are remarkably similar, with both types of barriers displaying a significant unfavourable enthalpic contribution, which is largely compensated by a favourable entropy of activation (Figure 10 d). For amyloid fibrils this structure of the free energy barriers has been explained through the interactions that need to be broken in order to reach the transition state (intramolecular secondary structure elements, as well as hydrogen bonds with the solvent).[148] On the other hand, the desolvation of hydrophobic regions of the peptide upon the first contacts with the fibril end is associated with a favourable entropy of activation.[148]

The comparison between amyloid fibrils and crystalline peptide systems therefore shows that while the assembly mechanisms can be quite distinct, the fundamental forces responsible for the assembly, and which determine both its thermodynamics and kinetics, are highly similar.

Conclusions and future perspectives

In this chapter, we have given a comprehensive overview over the state of knowledge of the physico-chemical and mechanistic aspects of short aromatic peptide assembly into fibrils, crystals and various other morphologies. It is rewarding to see that quantitative mechanistic studies have seen a steep rise in the last 5 years or so, and that our understanding of the driving forces responsible for these fascinating self-assembly processes is slowly, but surely, catching up with the enormous body of work that has so far mostly been based on empirical findings. A crucial remaining question is if, and how, the newly gained mechanistic understanding can be translated into rational control of the morphology and improvement of the material properties of (aromatic) peptide assemblies. The motivation to achieve this aim has the potential to be the major driving force for further fundamental studies in the coming

years. While we think that the fully *de novo* design of assembly mechanisms and properties is not yet routinely possible, the experimental and theoretical toolbox now at our disposal allows a highly detailed characterization of any given self-assembling peptide system, which in turn enables systematic control of its assembly process. The chaotic nature of the relationship between molecular structure of the short peptide monomer and the properties of its aggregate as yet precludes direct prediction of the latter from the former, although advances in areas outside traditional biophysics, for example crystal structure prediction, continue to expand the range of methods by which short peptide structures can be rationally designed. The amyloid cross-β structure, where it is experimentally found to occur in the short peptides, provides a biomedically-relevant model system amenable to detailed computation.

The finding that such simple peptide systems can be shown to adopt such a wide range of structures and properties is both a driving force for new research and a long-standing challenge to researchers seeking to rationalise their findings and base predictions on them. Recent quantitative and theoretical studies, seeking to establish the basic parameters of the assembly processes, should provide a base for such predictions, and for the intelligent design of new self-assembling systems based on these monomers. The field, more than twenty years on from the first organised studies into aromatic dipeptides, continues to provide new and often surprising results. We have also aimed to provide an overview of the relationship between these fascinating systems and the self-assembly of full-length polypeptides, and the extent to which understanding of one system - the forces driving and the mechanisms describing its aggregation, the properties and applications of the self-assembled structure, and the experimental techniques for its study - can be used to understand the other.

Acknowledgement

TOM thanks the Newman Foundation and the Weizmann Institute for funding. AKB thanks the Turnberg Foundation for a travel grant to Tel Aviv (2011), that enabled to start the

mechanistic studies of short aromatic peptide self-assembly.

References

(1) Waugh, D. F. *J. Am. Chem. Soc.* **1944**, *66*, 663–663.

(2) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662–666.

(3) Kidd, M. *Nature* **1963**, *197*, 192–193.

(4) Dobson, C. M. *Nature* **2003**, *426*, 884–890.

(5) Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. *J. Mol. Biol.* **1997**, *273*, 729–739.

(6) Polverino de Laureto, P.; Taddei, N.; Frare, E.; Capanni, C.; Costantini, S.; Zurdo, J.; Chiti, F.; Dobson, C. M.; Fontana, A. *J. Mol. Biol.* **2003**, *334*, 129–141.

(7) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. .; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773–778.

(8) Colvin, M. T.; Silvers, R.; Frohm, B.; Su, Y.; Linse, S.; Griffin, R. G. *J. Am. Chem. Soc.* **2015**, *137*, 7509–7518.

(9) Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; Willbold, D.; Schröder, G. F. *Science* **2017**, *358*, 116–119.

(10) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. W. *Nature* **2017**, *547*, 185–190.

(11) Azriel, R.; Gazit, E. *J. Biol. Chem.* **2001**, *276*, 34156–34161.

(12) Reches, M.; Gazit, E. *Science* **2003**, *300*, 625–627.

(13) Masters, C. L.; Simms, G.; Weinman, N. A.; Multhaup, G.; McDonald, B. L.; Beyreuther, K. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4245–4249.

(14) Görbitz, C. H. *Chemistry* **2001**, *7*, 5153–5159.

(15) Yemini, M.; Reches, M.; Gazit, E.; Rishpon, J. *Anal. Chem.* **2005**, *77*, 5155–5159.

(16) Kol, N.; Adler-Abramovich, L.; Barlam, D.; Shneck, R. Z.; Gazit, E.; Rousso, I. *Nano Lett.* **2005**, *5*, 1343–1346.

(17) Kholkin, A.; Amdursky, N.; Bdikin, I.; Gazit, E.; Rosenman, G. *ACS Nano* **2010**, *4*, 610–614.

(18) Wang, M.; Xiong, S.; Wu, X.; Chu, P. K. *Small* **2011**, *7*, 2801–2807.

(19) Mitra, R. N.; Das, D.; Roy, S.; Das, P. K. *J. Phys. Chem. B* **2007**, *111*, 14107–14113.

(20) Smith, A. M.; Williams, R. J.; Tang, C.; Coppo, P.; Collins, R. F.; Turner, M. L.; Saiani, A.; Ulijn, R. V. *Adv. Mater.* **2008**, *20*, 37–41.

(21) Johnson, E. K.; Adams, D. J.; Cameron, P. J. *J. Mater. Chem.* **2011**, *21*, 2024–2027.

(22) Whitesides, G.; Mathias, J.; Seto, C. *Science* **1991**, *254*, 1312–1319.

(23) Whitesides, G. M.; Grzybowski, B. *Science* **2002**, *295*, 2418–2421.

(24) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.

(25) Frishman, D.; Argos, P. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 566–579.

(26) Perrin, C. L.; ; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511–544, PMID: 9348662.

(27) Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.

(28) Wang, J.; Liu, K.; Xing, R.; Yan, X. *Chem. Soc. Rev.* **2016**, *45*, 5589–5604.

(29) Bowie, J. U. *Nature* **2005**, *438*, 581.

(30) Sheu, S.-Y.; Yang, D.-Y.; Selzle, H.; Schlag, E. *Proceedings of the National Academy of Sciences* **2003**, *100*, 12683–12687.

(31) Grzybowski, B. A.; Ishchenko, A. V.; DeWitte, R. S.; Whitesides, G. M.; Shakhnovich, E. I. *J. Phys. Chem. B.* **2000**, *104*, 7293–7298.

(32) Stone, A. *The Theory of Intermolecular Forces*; OUP Oxford, 2013.

(33) Gazit, E. *Chem. Soc. Rev.* **2007**, *36*, 1263–1269.

(34) Holmes, T. C.; de Lacalle, S.; Su, X.; Liu, G.; Rich, A.; Zhang, S. *Proc. Natl. Acad. Sci. (U.S.A)* **2000**, *97*, 6728–6733.

(35) Caplan, M. R.; Schwartzfarb, E. M.; Zhang, S.; Kamm, R. D.; Lauffenburger, D. A. *Biomaterials* **2002**, *23*, 219 – 227.

(36) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777 – 793.

(37) Kroon, J.; Kanters, J. *Nature* **1974**, *248*, 667.

(38) Allen, F. H.; Bird, C. M.; Rowland, R. S.; Raithby, P. R. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **1997**, *53*, 696–701.

(39) Sarkhel, S.; Desiraju, G. R. *Proteins: Structure, Function, and Bioinformatics* **2004**, *54*, 247–259.

(40) Pogorelyi, V. K. *Russ. Chem. Rev.* **1977**, *46*, 316–336.

(41) Desiraju, G. R. *J. Chem. Soc., Chem. Commun.* **1991**, 426–428.

(42) Derewenda, Z. S.; Lee, L.; Derewenda, U. *J. Mol. Biol.* **1995**, *252*, 248–262.

(43) Perutz, M. *Phil. Trans. R. Soc. A* **1993**, *345*, 105–112.

(44) Kauzmann, W. In *Some Factors in the Interpretation of Protein Denaturation*; Anfinsen, C., Anson, M., Bailey, K., Edsall, J. T., Eds.; Advances in Protein Chemistry; Academic Press, 1959; Vol. 14; pp 1 – 63.

(45) Tanford, C. *Science* **1978**, *200*, 1012–1018.

(46) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683–3704.

(47) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1980**, *73*, 3430–3433.

(48) Stillinger, F. H. In *Structure in Aqueous Solutions of Nonpolar Solutes from the Standpoint of Scaled-Particle Theory", bookTitle="The Physical Chemistry of Aqueous System: A Symposium in Honor of Henry S. Frank on His Seventieth Birthday"*; Kay, R. L., Ed.; Springer US: Boston, MA, 1973; pp 43–60.

(49) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B.* **1999**, *103*, 4570–4577.

(50) Huang, D. M.; Chandler, D. *J. Phys. Chem. B.* **2002**, *106*, 2047–2053.

(51) Huang, D. M.; Geissler, P. L.; Chandler, D. *J. Phys. Chem. B.* **2001**, *105*, 6704–6709.

(52) Chandler, D. *Nature* **2005**, *437*, 640–647.

(53) Patterson, D.; Barbe, M. *The Journal of Physical Chemistry* **1976**, *80*, 2435–2436.

(54) Shinoda, K.; Fujihira, M. *Bull. Chem. Soc. Jpn.* **1968**, *41*, 2612–2615.

(55) Fine, R. A.; Millero, F. J. *J. Chem. Phys.* **1973**, *59*, 5529–5536.

(56) Garde, S.; Hummer, G.; García, A. E.; Paulaitis, M. E.; Pratt, L. R. *Physical review letters* **1996**, *77*, 4966.

(57) Raschke, T. M.; Tsai, J.; Levitt, M. *Proc. Natl. Acad. Sci. (U.S.A)* **2001**, *98*, 5965–5969.

(58) Dill, K. A. *Biochemistry* **1985**, *24*, 1501–1509.

(59) Agashe, V. R.; Shastry, M.; Udgaonkar, J. B. *Nature* **1995**, *377*, 754.

(60) Richards, F. M. *Annual review of biophysics and bioengineering* **1977**, *6*, 151–176.

(61) Görbitz, C. H. *Acta Crystallogr. B* **2010**, *66*, 84–93.

(62) Samanta, U.; Pal, D.; Chakrabarti, P. *Acta crystallographica. Section D, Biological crystallography* **1999**, *55*, 1421–1427.

(63) Chakrabarti, P.; Bhattacharyya, R. *Progress in biophysics and molecular biology* **2007**, *95*, 83–137.

(64) Chourasia, M.; Sastry, G. M.; Sastry, G. N. *Int. J. Biol. Macromol.* **2011**, *48*, 540 – 552.

(65) Thomas, A.; Meurisse, R.; Brasseur, R. *Proteins* **2002**, *48*, 635–644.

(66) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790–18794.

(67) Ninkovic, D. B.; Andric, J. M.; Malkov, S. N.; Zaric, S. D. *Phys. Chem. Chem. Phys.* **2014**, *16*, 11173–11177.

(68) McGaughey, G. B.; Gagn?, M.; Rapp?, A. K. *The Journal of biological chemistry* **1998**, *273*, 15458–15463.

(69) Street, A. G.; Mayo, S. L. *Proc. Natl. Acad. Sci. (U.S.A)* **1999**, *96*, 9074–9076.

(70) Malkov, S. N.; Živković, M. V.; Beljanski, M. V.; Hall, M. B.; Zarić, S. D. *J. Mol. Model.* **2008**, *14*, 769–775.

(71) Samanta, U.; Bahadur, R. P.; Chakrabarti, P. *Protein Engineering, Design and Selection* **2002**, *15*, 659–667.

(72) Hunter, C. A.; Singh, J.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 837–846.

(73) Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M.-L.; Voelter, W.; Brunner, H.; Kapurniotu, A. *J. Mol. Biol.* **2000**, *295*, 1055 – 1071.

(74) Tjernberg, L. O.; N?slund, J.; Lindqvist, F.; Johansson, J.; Karlstr?m, A. R.; Thyberg, J.; Terenius, L.; Nordstedt, C. *The Journal of biological chemistry* **1996**, *271*, 8545–8548.

(75) Findeis, M. A.; Musso, G. M.; Arico-Muendel, C. C.; Benjamin, H. W.; Hundal, A. M.; Lee, J.-J.; Chin, J.; Kelley, M.; Wakefield, J.; Hayward, N. J.; Molineaux, S. M. *Biochemistry (Mosc.)* **1999**, *38*, 6791–6800, PMID: 10346900.

(76) Gazit, E. *FASEB J.* **2002**, *16*, 77–83.

(77) Tjernberg, L.; Hosia, W.; Bark, N.; Thyberg, J.; Johansson, J. *J. Biol. Chem.* **2002**, *277*, 43243–43246.

(78) Castelletto, V.; Hamley, I. W.; Cenker, C.; Olsson, U.; Adamcik, J.; Mezzenga, R.; Miravet, J. F.; Escuder, B.; Rodriguez-Llansola, F. *J. Phys. Chem. B* **2011**, *115*, 2107–2116.

(79) Lakshmanan, A.; Cheong, D. W.; Accardo, A.; Di Fabrizio, E.; Riekel, C.; Hauser, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 519–524.

(80) Genji, M.; Yano, Y.; Hoshino, M.; Matsuzaki, K. *Chem. Pharm. Bull.* **2017**, *65*, 668–673.

(81) Yan, X.; Cui, Y.; He, Q.; Wang, K.; Li, J. *Chem. Mater.* **2008**, *20*, 1522–1526.

(82) Demirel, G.; Malvadkar, N.; Demirel, M. C. *Langmuir* **2009**, *26*, 1460–1463.

(83) Su, Y.; Yan, X.; Wang, A.; Fei, J.; Cui, Y.; He, Q.; Li, J. *J. Mater. Chem.* **2010**, *20*, 6734–6740.

(84) Zhu, P.; Yan, X.; Su, Y.; Yang, Y.; Li, J. *Chem. Eur. J.* **2010**, *16*, 3176–3183.

(85) Huang, R.; Qi, W.; Su, R.; Zhao, J.; He, Z. *Soft Matter* **2011**, *7*, 6418–6421.

(86) Huang, R.; Wang, Y.; Qi, W.; Su, R.; He, Z. *Nanoscale Res. Lett.* **2014**, *9*, 653.

(87) Mason, T. O.; Chirgadze, D. Y.; Levin, A.; Adler-Abramovich, L.; Gazit, E.; Knowles, T. P. J.; Buell, A. K. *ACS Nano* **2014**, *8*, 1243–53.

(88) Song, Y.; Challa, S. R.; Medforth, C. J.; Qiu, Y.; Watt, R. K.; Pena, D.; Miller, J. E.; van Swol, F.; Shelnutt, J. A. *Chem. Commun. (Cambridge)* **2004**, 1044–1045.

(89) Mason, T. O.; Michaels, T. C. T.; Levin, A.; Dobson, C. M.; Gazit, E.; Knowles, T. P. J.; Buell, A. K. *J. Am. Chem. Soc.* **2017**, *139*, 16134–16142.

(90) Park, J. S.; Han, T. H.; Oh, J. K.; Kim, S. O. *J. Nanosci. Nanotech.* **2010**, *10*, 6954–6957.

(91) Adams, D. J.; Butler, M. F.; Frith, W. J.; Kirkland, M.; Mullen, L.; Sanderson, P. *Soft Matter* **2009**, *5*, 1856–1862.

(92) Zhang, Y.; Kuang, Y.; Gao, Y.; Xu, B. *Langmuir* **2010**, *27*, 529–537.

(93) Zhou, J.; Du, X.; Gao, Y.; Shi, J.; Xu, B. *J. Am. Chem. Soc.* **2014**, *136*, 2970–2973.

(94) Debnath, S.; Roy, S.; Ulijn, R. V. *J. Am. Chem. Soc.* **2013**, *135*, 16789–16792.

(95) Mason, T. O.; Michaels, T. C. T.; Levin, A.; Gazit, E.; Dobson, C. M.; Buell, A. K.; Knowles, T. P. J. *J. Am. Chem. Soc.* **2016**, *138*, 9589–9596.

(96) Adler-Abramovich, L.; Aronov, D.; Beker, P.; Yevnin, M.; Stempler, S.; Buzhansky, L.; Rosenman, G.; Gazit, E. *Nature Nanotechnology* **2009**, *4*, 849.

(97) Bank-Srour, B.; Becker, P.; Krasovitsky, L.; Gladkikh, A.; Rosenberg, Y.; Barkay, Z.; Rosenman, G. *Polymer J.* **2013**, *45*, 494.

(98) Vasudev, M. C.; Koerner, H.; Singh, K. M.; Partlow, B. P.; Kaplan, D. L.; Gazit, E.; Bunning, T. J.; Naik, R. R. *Biomacromolecules* **2014**, *15*, 533–540.

(99) Levin, A.; Mason, T. O.; Adler-Abramovich, L.; Buell, A. K.; Meisl, G.; Galvagnion, C.; Bram, Y.; Stratford, S. A.; Dobson, C. M.; Knowles, T. P. J.; Gazit, E. *Nat. Commun.* **2014**, *5*, 5219.

(100) Adler-Abramovich, L.; Reches, M.; Sedman, V. L.; Allen, S.; Tendler, S. J. B.; Gazit, E. *Langmuir* **2006**, *22*, 1313–1320.

(101) Sedman, V. L.; Adler-Abramovich, L.; Allen, S.; Gazit, E.; Tendler, S. J. B. *J. Am. Chem. Soc.* **2006**, *128*, 6903–6908.

(102) Tamamis, P.; Adler-Abramovich, L.; Reches, M.; Marshall, K.; Sikorski, P.; Serpell, L.; Gazit, E.; Archontis, G. *Biophys. J.* **2009**, *96*, 5020–5029.

(103) Adler-Abramovich, L.; Kol, N.; Yanai, I.; Barlam, D.; Shneck, R. Z.; Gazit, E.; Rousso, I. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 9939–9942.

(104) Chen, L.; Morris, K.; Laybourn, A.; Elias, D.; Hicks, M. R.; Rodger, A.; Serpell, L.; Adams, D. J. *Langmuir* **2010**, *26*, 5232–5242.

(105) Fichman, G.; Guterman, T.; Damron, J.; Adler-Abramovich, L.; Schmidt, J.; Kesselman, E.; Shimon, L. J.; Ramamoorthy, A.; Talmon, Y.; Gazit, E. *Science advances* **2016**, *2*, e1500827.

(106) Cardoso, A. Z.; Mears, L. L. E.; Cattoz, B. N.; Griffiths, P. C.; Schweins, R.; Adams, D. J. *Soft Matter* **2016**, *12*, 3612–3621.

(107) Onogi, S.; Shigemitsu, H.; Yoshii, T.; Tanida, T.; Ikeda, M.; Kubota, R.; Hamachi, I. *Nature chemistry* **2016**, *8*, nchem–2526.

(108) Kubota, R.; Liu, S.; Shigemitsu, H.; Nakamura, K.; Tanaka, W.; Ikeda, M.; Hamachi, I. *Bioconjugate chemistry* **2018**,

(109) Tena-Solsona, M.; Escuder, B.; Miravet, J. F.; Casttelleto, V.; Hamley, I. W.; Dehsorkhi, A. *Chem. Mater.* **2015**, *27*, 3358–3365.

(110) Yan, X.; He, Q.; Wang, K.; Duan, L.; Cui, Y.; Li, J. *Angew. Chem. Int. Ed. Engl.* **2007**, *119*, 2483–2486.

(111) Wallace, M.; Iggo, J. A.; Adams, D. J. *Soft Matter* **2015**, *11*, 7739–7747.

(112) Wallace, M.; Iggo, J. A.; Adams, D. J. *Soft Matter* **2017**, *13*, 1716–1727.

(113) Do, T. D.; Bowers, M. T. *Anal. Chem.* **2015**, *87*, 4245–4252.

(114) Chen, L.; Pont, G.; Morris, K.; Lotze, G.; Squires, A.; Serpell, L. C.; Adams, D. J. *Chem. Commun.* **2011**, *47*, 12071–12073.

(115) Martin, A. D.; Wojciechowski, J. P.; Robinson, A. B.; Heu, C.; Garvey, C. J.; Ratcliffe, J.; Waddington, L. J.; Gardiner, J.; Thordarson, P. *Scientific Reports* **2017**, *7*.

(116) Colquhoun, C.; Draper, E. R.; Schweins, R.; Marcello, M.; Vadukul, D.; Serpell, L. C.; Adams, D. J. *Soft Matter* **2017**, *13*, 1914–1919.

(117) Shigemitsu, H.; Fujisaku, T.; Tanaka, W.; Kubota, R.; Minami, S.; Urayama, K.; Hamachi, I. *Nature nanotechnology* **2018**, 1.

(118) Arnon, Z. A.; Vitalis, A.; Levin, A.; Michaels, T. C.; Caflisch, A.; Knowles, T. P.; Adler-Abramovich, L.; Gazit, E. *Nat. Commun.* **2016**, *7*, 13190.

(119) Amdursky, N.; Beker, P.; Koren, I.; Bank-Srour, B.; Mishina, E.; Semin, S.; Rasing, T.; Rosenberg, Y.; Barkay, Z.; Gazit, E.; Rosenman, G. *Biomacromolecules* **2011**, *12*, 1349–1354.

(120) Handelman, A.; Natan, A.; Rosenman, G. *J. Pept. Sci.* **2014**, *20*, 487–493.

(121) Ryu, J.; Park, C. B. *Chem. Mater.* **2008**, *20*, 4284–4290.

(122) Kleinsmann, A. J.; Nachtsheim, B. J. *Chem. Commun.* **2013**, *49*, 7818–7820.

(123) Zhou, X.; Fan, J.; Li, N.; Du, Z.; Ying, H.; Wu, J.; Xiong, J.; Bai, J. *Fluid Phase Equilibria* **2012**, *316*, 26–33.

(124) Franks, F.; Gent, M.; Johnson, H. *J. Chem. Soc. (Resumed)* **1963**, 2716–2723.

(125) Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8069–8072.

(126) Han, T. H.; Oh, J. K.; Lee, G.-J.; Pyun, S.-I.; Kim, S. O. *Colloids and Surfaces B: Biointerfaces* **2010**, *79*, 440–445.

(127) McDevit, W.; Long, F. *J. Am. Chem. Soc.* **1952**, *74*, 1773–1777.

(128) Marshall, K. E.; Hicks, M. R.; Williams, T. L.; Hoffmann, S. r. V. n.; Rodger, A.; Dafforn, T. R.; Serpell, L. C. *Biophys. J.* **2010**, *98*, 330–338.

(129) Reynolds, N. P.; Adamcik, J.; Berryman, J. T.; Handschin, S.; Zanjani, A. A. H.; Li, W.; Liu, K.; Zhang, A.; Mezzenga, R. *Nat. Commun.* **2017**, *8*, 1338.

(130) Chen, L.; Revel, S.; Morris, K.; C. Serpell, L.; Adams, D. J. *Langmuir* **2010**, *26*, 13466–13471.

(131) Fleming, S.; Debnath, S.; Frederix, P. W.; Tuttle, T.; Ulijn, R. V. *Chemical Communications* **2013**, *49*, 10587–10589.

(132) Reches, M.; Gazit, E. *Isr. J. Chem.* **2005**, *45*, 363–371.

(133) Tang, C.; Ulijn, R. V.; Saiani, A. *Langmuir* **2011**, *27*, 14438–14449.

(134) Houton, K. A.; Morris, K. L.; Chen, L.; Schmidtmann, M.; Jones, J. T. A.; Serpell, L. C.; Lloyd, G. O.; Adams, D. J. *Langmuir : the ACS journal of surfaces and colloids* **2012**, *28*, 9797–9806.

(135) Tang, C.; Smith, A. M.; Collins, R. F.; Ulijn, R. V.; Saiani, A. *Langmuir* **2009**, *25*, 9447–9453.

(136) Ramos Sasselli, I.; Halling, P. J.; Ulijn, R. V.; Tuttle, T. *ACS Nano* **2016**, *10*, 2661–2668.

(137) Hsu, S.-M.; Lin, Y.-C.; Chang, J.-W.; Liu, Y.-H.; Lin, H.-C. *Angew. Chem. Intl. Ed. Engl.* **2014**, *53*, 1921–1927.

(138) Rajbhandary, A.; Nilsson, B. L. *Pept. Sci.* **2017**, *108*.

(139) Brouhard, G. J. *Molecular biology of the cell* **2015**, *26*, 1207–1210.

(140) Sadownik, J. W.; Leckie, J.; Ulijn, R. V. *Chemical communications (Cambridge, England)* **2011**, *47*, 728–730.

(141) Frederix, P. W.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. *Nature Chemistry* **2015**, *7*, 30–37.

(142) Görbitz, C.; Etter, M. *Acta Crystallographica Section C* **1992**, *48*, 1317–1320.

(143) Zeng, G.; Liu, L.; Xia, D.; Li, Q.; Xin, Z.; Wang, J.; Besenbacher, F.; Skrydstrup, T.; Dong, M. *RSC Advances* **2014**, *4*, 7516–7520.

(144) Korolkov, V. V.; Allen, S.; Roberts, C. J.; Tendler, S. J. *Faraday discussions* **2013**, *166*, 257–267.

(145) Yan, X.; Cui, Y.; He, Q.; Wang, K.; Li, J.; Mu, W.; Wang, B.; Ou-yang, Z.-c. *Chem.–Eur. J.* **2008**, *14*, 5974–5980.

(146) Reches, M.; Gazit, E. *Nano Lett.* **2004**, *4*, 581–585.

(147) Nielsen, A. E. *J. Cryst. Growth* **1984**, *67*, 289–310.

(148) Buell, A. K.; Dhulesia, A.; White, D. A.; Knowles, T. P.; Dobson, C. M.; Welland, M. E. *Angew. Chem. Intl. Ed. Engl.* **2012**, *51*, 5247–5251.

(149) Adler-Abramovich, L.; Marco, P.; Arnon, Z. A.; Creasey, R. C. G.; Michaels, T. C. T.; Levin, A.; Scurr, D. J.; Roberts, C. J.; Knowles, T. P. J.; Tendler, S. J. B.; Gazit, E. *ACS Nano* **2016**, *10*, 7436–7442.

(150) Creasey, R. C. G.; Louzao, I.; Arnon, Z. A.; Marco, P.; Adler-Abramovich, L.; Roberts, C. J.; Gazit, E.; Tendler, S. J. B. *Soft Matter* **2016**, *12*, 9451–9457.

(151) Buell, A. K. *Int. Rev. Cell Mol. Biol.* **2017**, *329*, 187–226.

(152) Wang, Y.; Huang, R.; Qi, W.; Xie, Y.; Wang, M.; Su, R.; He, Z. *Small* **2015**, *11*, 2893–2902.

(153) Amdursky, N.; Molotskii, M.; Gazit, E.; Rosenman, G. *J. Am. Chem. Soc.* **2010**, *132*, 15632–15636.

(154) Gebauer, D.; Kellermeier, M.; Gale, J. D.; Bergström, L.; Cölfen, H. *Chem. Soc. Rev.* **2014**, *43*, 2348–2371.

(155) Ishikawa, M.; Busch, C.; Motzkus, M.; Martinho, H.; Buckup, T. *Phys. Chem. Chem. Phys.* **2017**,

(156) Buell, A. K.; Galvagnion, C.; Gaspar, R.; Sparr, E.; Vendruscolo, M.; Knowles, T. P. J.; Linse, S.; Dobson, C. M. *Proc Natl Acad Sci U S A* **2014**, *111(21)*, 7671–7676.

(157) Görbitz, C. H.; Hartviksen, L. M. *Acta Crystallogr. C* **2008**, *64*, 171–6.

(158) Görbitz, C. H.; Etter, M. C. *Int. J. Pept. Protein Res.* **1992**, *39*, 93–110.

(159) Görbitz, C. H. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2002**, *58*, 512–518.

(160) Emge, T. J.; Agrawal, A.; Dalessio, J. P.; Dukovic, G.; Inghrim, J. A.; Janjua, K.; Macaluso, M.; Robertson, L. L.; Stiglic, T. J.; Volovik, Y.; Georgiadis, M. M. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **2000**, *56*, e469–e471.

(161) Görbitz, C. H.; Gundersen, E. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1996**, *52*, 1764–1767.

(162) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893.

(163) Lee, E. C.; Kim, D.; Jurecka, P.; Tarakeshwar, P.; Hobza, P.; Kim, K. S. *The journal of physical chemistry. A* **2007**, *111*, 3446–3457.

(164) Prampolini, G.; Livotto, P. R.; Cacelli, I. *Journal of chemical theory and computation* **2015**, *11*, 5182–5196.

(165) Suresh, C.; Vijayan, M. *Journal of Biosciences* **1987**, *12*, 13–21.

(166) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N.-L. *Angewandte Chemie International Edition in English* **1995**, *34*, 1555–1573.

(167) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **1990**, *46*, 256–262.

(168) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750–4755.

(169) Fabiola, G. F.; Krishnaswamy, S.; Nagarajan, V.; Pattabhi, V. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 316–320.

(170) Akazome, M.; Ueno, Y.; Ooiso, H.; Ogura, K. *J. Org. Chem.* **2000**, *65*, 68–76.

(171) Pellach, M.; Mondal, S.; Shimon, L. J.; Adler-Abramovich, L.; Buzhansky, L.; Gazit, E. *Chem. Mater.* **2016**, *28*, 4341–4348.

(172) Eggleston, D. S.; Hodgson, D. J. *Chemical Biology & Drug Design* **1985**, *26*, 509–517.

(173) Reches, M.; Gazit, E. *Nature Nanotechnology* **2006**, *1*, 195–200.

(174) Görbitz, C. H. *Chem. Commun. (Camb.)* **2006**, 2332–2334.

(175) Lekprasert, B.; Korolkov, V.; Falamas, A.; Chis, V.; Roberts, C. J.; Tendler, S. J. B.; Notingher, I. *Biomacromolecules* **2012**, *13*, 2181–2187.

(176) Eddleston, M. D.; Jones, W. *Cryst. Growth & Design* **2009**, *10*, 365–370.

(177) Kim, J.; Han, T. H.; Kim, Y.-I.; Park, J. S.; Choi, J.; Churchill, D. G.; Kim, S. O.; Ihee, H. *Adv. Mater.* **2010**, *22*, 583–587.

(178) Li, Q.; Jia, Y.; Dai, L.; Yang, Y.; Li, J. *ACS Nano* **2015**, *9*, 2689–2695.

(179) Pappas, C. G.; Frederix, P. W.; Mutasa, T.; Fleming, S.; Abul-Haija, Y. M.; Kelly, S. M.; Gachagan, A.; Kalafatovic, D.; Trevino, J.; Ulijn, R.; Bai, S. *Chem. Commun.* **2015**, *51*, 8465–8468.

(180) Pappas, C. G.; Mutasa, T.; Frederix, P. W.; Fleming, S.; Bia, S.; Debnath, S.; Kelly, S. M.; Gachagan, A.; Ulijn, R. V. *Mater. Horiz.* **2015**, *2*, 198–202.

(181) Moholkar, V. S.; Sable, S. P.; Pandit, A. B. *AIChE J* **2000**, *46*, 684–694.

(182) Tjernberg, L. O.; Callaway, D. J.; Tjernberg, A.; Hahne, S.; Lilliehöök, C.; Terenius, L.; Thyberg, J.; Nordstedt, C. *J. Biol. Chem.* **1999**, *274*, 12619–12625.

(183) Mazor, Y.; Gilead, S.; Benhar, I.; Gazit, E. *J. Mol. Biol.* **2002**, *322*, 1013–1024.

(184) Gazit, E. *FEBS J.* **2005**, *272*, 5971–5978.

(185) Jahn, T. R.; Makin, O. S.; Morris, K. L.; Marshall, K. E.; Tian, P.; Sikorski, P.; Serpell, L. C. *J. Mol. Biol.* **2010**, *395*, 717–727.

(186) Carugo, O.; Djinović-Carugo, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1333–1341.

(187) Hollingsworth, S. A.; Karplus, P. A. *Biomolecular Concepts* **2010**, *1*, 271–283.

(188) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Research* **2000**, *28*, 235–242.

(189) Schrödinger, LLC,

(190) Di Maro, S. et al. *J. Med. Chem.* **2016**, *59*, 8369–8380, PMID: 27571038.

(191) Colletier, J.-P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D. *Proc. Natl. Acad. Sci. (U.S.A)* **2011**, *108*, 16938–16943.

(192) Spencer, R. K.; Li, H.; Nowick, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 5595–5598.

(193) Kreutzer, A. G.; Hamza, I. L.; Spencer, R. K.; Nowick, J. S. *J. Am. Chem. Soc.* **2016**, *138*, 4634–4642, PMID: 26967810.

(194) Wälti, M. A.; Ravotti, F.; Arai, H.; Glabe, C. G.; Wall, J. S.; Böckmann, A.; Güntert, P.; Meier, B. H.; Riek, R. *Proc. Natl. Acad. Sci. (U.S.A)* **2016**, *113*, E4976–E4984.

(195) Perutz, M. F.; Finch, J. T.; Berriman, J.; Lesk, A. *Proc Natl Acad Sci U S A* **2002**, *99*, 5591–5595.

(196) Sikorski, P.; Atkins, E. *Biomacromolecules* **2005**, *6*, 425–432.

(197) Buchanan, L. E.; Carr, J. K.; Fluitt, A. M.; Hoganson, A. J.; Moran, S. D.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. *Proceedings of the National Academy of Sciences* **2014**, *111*, 5796–5801.

(198) Makin, O. S.; Serpell, L. C. *The FEBS journal* **2005**, *272*, 5950–5961.

(199) Harper, J. D.; Lieber, C. M.; Lansbury, P. T. *Chemistry & Biology* **1997**, *4*, 951 – 959.

(200) Chothia, C. *J. Mol. Biol.* **1973**, *75*, 295–302.

(201) Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C.; Semenov, A. N.; Boden, N. *Proc. Natl. Acad. Sci. (U.S.A)* **2001**, *98*, 11857–11862.

(202) Usov, I.; Adamcik, J.; Mezzenga, R. *ACS Nano* **2013**, *7*, 10465–10474.

(203) Wang, S.-T.; Lin, Y.; Spencer, R. K.; Thomas, M. R.; Nguyen, A. I.; Amdursky, N.; Pashuck, E. T.; Skaalure, S. C.; Song, C. Y.; Parmar, P. A.; Morgan, R. M.; Ercius, P.; Aloni, S.; Zuckermann, R. N.; Stevens, M. M. *ACS Nano* **2017**, *11*, 8579–8589, PMID: 28771324.

(204) Knowles, T. P. J.; Simone, A. D.; Fitzpatrick, A. W.; Baldwin, A.; Meehan, S.; Rajah, L.; Vendruscolo, M.; Welland, M. E.; Dobson, C. M.; Terentjev, E. M. *Phys. Rev. Lett.* **2012**, *109*, 158101.

(205) Knowles, T. P.; Smith, J. F.; Craig, A.; Dobson, C. M.; Welland, M. E. *Phys. Rev. Lett.* **2006**, *96*, 238301.

(206) Šarić, A.; Buell, A. K.; Meisl, G.; Michaels, T. C.; Dobson, C. M.; Linse, S.; Knowles, T. P.; Frenkel, D. *Nature Physics* **2016**, *12*, 874–880.

(207) Cohen, S. I. A.; Cukalevski, R.; Michaels, T. C. T.; Saric, A.; Törnquist, M.; Vendruscolo, M.; Dobson, C. M.; Buell, A. K.; Knowles, T. P. J.; Linse, S. *Nature Chemistry* **2018**, *10*, pages523?531.

(208) Cabriolu, R.; Auer, S. *J. Mol. Biol.* **2011**, *411*, 275–285.

(209) Fitzpatrick, A. W. P. et al. *Proc. Natl. Acad. Sci. (U.S.A)* **2013**, *110*, 5468–5473.

(210) Collins, S. R.; Douglass, A.; Vale, R. D.; Weissman, J. S. *PLoS Biol.* **2004**, *2*, e321.

(211) Buell, A. K.; Blundell, J. R.; Dobson, C. M.; Welland, M. E.; Terentjev, E. M.; Knowles, T. P. J. *Phys. Rev. Lett.* **2010**, *104*, 228101.

(212) Kashchiev, D.; Auer, S. *J Chem Phys* **2010**, *132*, 215101.

(213) Cabriolu, R.; Kashchiev, D.; Auer, S. *J. Chem. Phys.* **2010**, *133*, 12B602.

(214) Balbirnie, M.; Grothe, R.; Eisenberg, D. S. *Proc. Natl. Acad. Sci. (U.S.A)* **2001**, *98*, 2375–2380.

(215) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453 EP.

(216) Scherzinger, E.; Lurz, R.; Turmaine, M.; Mangiarini, L.; Hollenbach, B.; Hasenbank, R.; Bates, G. P.; Davies, S. W.; Lehrach, H.; Wanker, E. E. *Cell* **1997**, *90*, 549–558.

(217) Han, H.; Weinreb, P. H.; Lansbury Jr, P. T. *Chemistry & Biology* **1995**, *2*, 163–169.

(218) Rodriguez, J. A. et al. *Nature* **2015**, *525*, 486 EP.

(219) Colvin, M. T.; Silvers, R.; Ni, Q. Z.; Can, T. V.; Sergeyev, I.; Rosay, M.; Donovan, K. J.; Michael, B.; Wall, J.; Linse, S.; Griffin, R. G. *J. Am. Chem. Soc.* **2016**, *138*, 9663–9674, PMID: 27355699.

(220) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639–659.

(221) Muller, N. *J. Solution Chem.* **1988**, *17*, 661–672.

(222) Baldwin, A. J.; Knowles, T. P. J.; Tartaglia, G. G.; Fitzpatrick, A. W.; Devlin, G. L.; Shammas, S. L.; Waudby, C. A.; Mossuto, M. F.; Meehan, S.; Gras, S. L.; Christodoulou, J.; Anthony-Cahill, S. J.; Barker, P. D.; Vendruscolo, M.; Dobson, C. M. *J. Am. Chem. Soc.* **2011**, *133*, 14160–14163, PMID: 21650202.

(223) Oosawa, F.; Kasai, M. *J. Mol. Biol.* **1962**, *4*, 10–21.