

Conformation Dependence of Diphenylalanine Self-Assembly Structures and Dynamics#Insights from Hybrid-Resolution Simulations

Qinsi Xiong, Yixiang Jiang, Xiang Cai, Fadeng Yang, Zigang Li, and Wei Han

ACS Nano, Just Accepted Manuscript • DOI: 10.1021/acs.nano.8b09741 • Publication Date (Web): 14 Mar 2019

Downloaded from <http://pubs.acs.org> on March 19, 2019

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

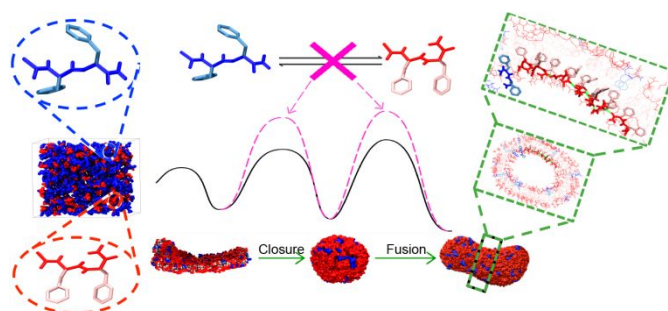
Conformation Dependence of Diphenylalanine Self-Assembly Structures and Dynamics : Insights from Hybrid-Resolution Simulations

Qinsi Xiong,^{‡,†} Yixiang Jiang,^{‡,†} Xiang Cai,[†] Fadeng Yang,[†] Zigang Li,^{*,†} Wei Han^{*,†}

[†]State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen, 518055, China

*Corresponding author: email: hanw@pkusz.edu.cn; lizg@pkusz.edu.cn

TOC Graphics



ABSTRACT: The molecular design of peptide-assembled nanostructures relies on extensive knowledge pertaining to the relationship between conformational features of peptide constituents and their behavior regarding self-assembly, and characterizing the conformational details of peptides during their self-assembly is experimentally challenging. Here we demonstrate that a hybrid-resolution modelling method can be employed to investigate the role that conformation plays during the assembly of terminally-capped diphenylalanines (FF) through microsecond simulations of hundreds or thousands of peptides. Our simulations discovered tubular or vesicular nanostructures that were consistent with experimental observation while reproducing critical self-assembly concentration and secondary structure contents in the assemblies that were measured in our experiments. The atomic details provided by our method allowed us to uncover diverse FF conformations and conformation dependence of assembled nanostructures. We found that the assembled morphologies and the molecular packing of FFs in the observed assemblies are linked closely with side-chain angle and peptide bond orientation, respectively. Of various conformations accessible to soluble FFs only a select few are compatible with the assembled morphologies in water. A conformation resembling a FF crystal, in particular, became predominant due to its ability to permit highly ordered and energetically favorable FF packing in aqueous assemblies. Strikingly, several conformations incompatible with the assemblies arose transiently as intermediates, facilitating key steps of the assembly process. The molecular rationale behind the role of these intermediate conformations were further explained. Collectively, the structural details reported here advance the understanding of the FF self-assembly mechanism, and our method shows promise for studying peptide-assembled nanostructures and their rational design.

KEYWORDS: *diphenylalanine self-assembly, hybrid-resolution force field, molecular dynamic simulation, peptide nanotube, peptide secondary structure, conformational dynamics, flip-flop motion*

1
2
3 Peptide-assembled materials have garnered great attention in recent years due to their simplicity,
4 availability, excellent biocompatibility and biodegradability, and have gradually aided in great
5 achievements in material science.¹⁻⁵ Diphenylalanine (FF), derived originally from the core motif of
6 disease-related amyloid- β peptides, is one of the simplest molecules that can form these types of
7 materials.⁶ In water, FFs self-assemble into nanotubes, but other assembled structures such as
8 nanospheres, nanofibers, nanowires, nanosheet, and *etc.* are also attainable depending on assembly
9 conditions and chemical modifications introduced.⁶⁻¹⁴ Due to their excellent optoelectrical and
10 mechanical properties, these FF-based materials have various applications ranging from batteries to
11 tissue repair.^{3,4,15,16}

12
13 A central task of the molecular design for FF-based materials is establishing the relationships
14 between structures of constituent peptides and assembled supramolecular structures with the aim of
15 applying these relationships to design materials with targeted assembled structures and properties.^{17,18}
16 When achieving this goal, it is necessary to elucidate molecular factors underlying the self-assembly
17 process. The importance of non-covalent forces between peptide molecules is evident. The knowledge
18 regarding the roles of these forces in the self-assembly of FF-based molecules have been greatly
19 advanced thanks to numerous insightful experimental studies and simulations.¹⁹⁻²⁹ However, one of
20 the main hurdles that prevent further understanding of the self-assembly process lies in the
21 conformational flexibility inherent to FF peptides.^{30,31}

22
23 There have been many reported examples of various secondary structures in FF-based assemblies,
24 including β -sheets, α -helices, and turns when exposed to distinct external conditions in pH,
25 temperature, and solvent environment.^{7,13,32,33} Although the exact nature of the corresponding
26 conformations of FF building blocks remain unknown, it is conceivable that there could be many
27 diverse peptide backbone conformations. Different conformations may exhibit distinct side-chain
28 arrangements and hydrogen bond donor and acceptor orientations for backbone. These local
29 geometric features dictate how the building blocks may interact with their neighbors, which can
30 sometimes result in a long range structural order and allow for particular supramolecular structure
31 formation.^{31,34-38} It is important, therefore, to explore internal conformations accessible to FF-based
32 building blocks and investigate the conformational-dependence of molecular packing and
33 supramolecular structures.

34
35 The diverse conformations present in FF-based molecules further implies that these molecules
36 would undergo conformational changes when exposed to thermal forces, thereby altering the way they
37 would interact with their surroundings. Presented in a recent kinetic study using Raman spectroscopy,
38 heterogeneous conformational transitions of FFs were indeed observed during their self-assembly into
39 nanotubes.³⁹ In general, conformational fluctuation of protein-like molecules can be crucial for self-
40 assembly processes.⁴⁰ It has been suggested that a reconfiguration of protein-like building blocks may
41 assist systems in escaping trapped states, thereby facilitating either the formation of correctly
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 assembled structures or the transformation of assemblies between different structures.^{41,42} The role of
4 conformational fluctuation in the assembly process of FF-based systems, however, still awaits
5 elucidation.
6
7

8 Characterizing peptide conformations in assemblies remains challenging. Methods that rely on
9 ensemble average such as infrared absorption, circular dichroism and X-ray diffraction oftentimes
10 cannot resolve conformations of individual amino acids.¹⁷ X-ray crystallography has been applied in
11 resolving atomic structures of FF-based molecules,^{43,44} but it is still unclear if crystalized molecules
12 assume the exact same conformation as they do in the assembly.^{45,46}
13
14
15

16 Molecular dynamics simulations have provided another means of investigating molecular details of
17 FF self-assembly.^{22,24,25,28,29,47-50} All-atom simulations, although arguably accurate, usually can
18 examine the self-assembly of only a small number of peptide molecules in relatively short timescales.
19 FF-based molecular conformations were typically examined at a monomer level⁴⁷ or in small self-
20 assembled structures that were composed of dozens of peptide molecules.^{24,25,48} Alternatively, coarse-
21 grained (CG) models can simplify spatial representations of systems and reduce the computational
22 cost of simulating the self-assembly of hundreds or even thousands of FF-based molecules.^{22,26,28,29,49}
23 However, most of these models neglect atomic details of peptide molecules and, oftentimes cannot be
24 used to study the conformations of peptide building blocks.
25
26
27
28
29
30

31 In this study, a hybrid-resolution model (termed PACE) was employed to simulate peptide self-
32 assembly. In PACE, solvent representation is simplified using the MARTINI CG model⁵¹ while
33 retaining most of atomic details of proteins.⁵² The PACE model has been successfully applied to
34 obtain mechanistic insight into protein folding,⁵³ protein aggregation,⁵⁴ and protein-peptide
35 interaction.⁵⁵ Here, the model has been improved for simulating the assembly of FF-based systems.
36 The improved model allows for long-time simulation of hundreds or even thousands of FF molecules
37 while at the same time capturing important details such as hydrogen bonds (HB) and peptide
38 conformations.
39
40
41
42
43

44 With this approach, we investigate the self-assembly mechanism of charged-termini-capped FFs.
45 The N-terminally acetylated and C-terminally amidated FFs comprise two repeating peptide units and
46 structurally resemble typical polypeptides more than the uncapped FFs. Since the capped FFs can also
47 self-assemble into nanotubes,^{7,56} this FF derivative (hereinafter referred to as FF instead) was used
48 here as a model system for conformational study of peptide self-assembly.
49
50
51

52 Extensive simulations enabled by PACE allowed us to observe the spontaneous assembly of FFs
53 into doubled layered patch structures that would further curve up into hollow vesicles and tubes,
54 corroborating what was proposed previously based on CG simulations.^{14,26,28} However, our
55 simulations revealed extensive details of FF conformations that have not been observed thus far.
56 While there are numerous conformations accessible to FFs in water, only a select few are
57
58
59
60

1
2
3 geometrically compatible with the assembled structures. In particular, the crystal-like conformation,
4 which is rarely observed in water, became the dominant form in the assemblies, permitting a highly
5 ordered packing of FFs that is energetically favorable. We found that the angle between the two
6 aromatic side chains of FFs is a key geometric parameter associated with assembled morphologies
7 whereas the orientation of the backbone HB donors and receptors instead correlated with the local
8 packing of FFs. While the conformations with a larger side-chain angle are not preferred in hollow
9 vesicles or tubes, they still need to arise transiently in order to facilitate the formation of vesicles and
10 tubes during the assembly process. These findings provide additional understanding of the
11 relationship between conformational complexity of FFs and their self-assembly behavior, aiding to
12 furnish guidance for the design of FF-based materials.
13
14
15
16
17
18
19
20

21 **RESULTS AND DISCUSSION**

22 **Parameterization of PACE for diphenylalanine self-assembly**

23
24
25
26 Previous representation of FFs with PACE captures important structural features of peptide
27 molecules such as the packing of side chains, directionality of HBs and the secondary structures of the
28 backbone (**Figures 1a and S1 in the supplementary information, SI**). While this representation has
29 permitted the application of PACE in systems like protein folding and protein-protein interactions,
30 here we fine-tuned the parameters of PACE to further increase its accuracy in modeling the FF self-
31 assembly. More specifically, the parameters associated with several types of non-covalent interactions
32 were adjusted to fit the all-atom simulation results of the FF assembly system (see Method and Table
33 **S1**). Since this system contains eight FFs, it is small enough for equilibrium sampling through replica
34 exchange molecular dynamics (REMD) simulations.⁵⁷
35
36
37
38
39

40
41 In parameter fitting, focus was made on a few structural properties. The first such property was the
42 intermolecular HB contacts between FFs, with the HB interactions involved with individual peptide
43 bonds considered separately for the fitting. A second property was the aromatic-aromatic contacts
44 involved with each aromatic side chain, and the third property was the spatial distributions of different
45 parts of FFs with respect to the centers of the assembled structures. We found that where a particular
46 part of a FF would stay in an assembly is sensitive to the balance between its interactions with other
47 solutes and with solvent. **Figures 1b-d** shows these properties calculated with PACE before and after
48 the optimization, clearly indicating that the optimized PACE had been improved.
49
50
51
52
53

54 Since it is important to test if the parameters optimized for a small FF system can be transferred to
55 large systems, we performed simulations for two large systems (one with 810 FFs and the other with
56 2000 FFs). The detailed results of these simulations will be discussed later. The ability of the
57 improved PACE to reproduce the critical concentration (C_s) of self-assembly of FFs was of greater
58 interest here, as this quantity denotes the concentration at which soluble FFs reach an equilibrium
59
60

with FFs in large aggregates and represents one of the most important thermodynamic parameters of aggregation.

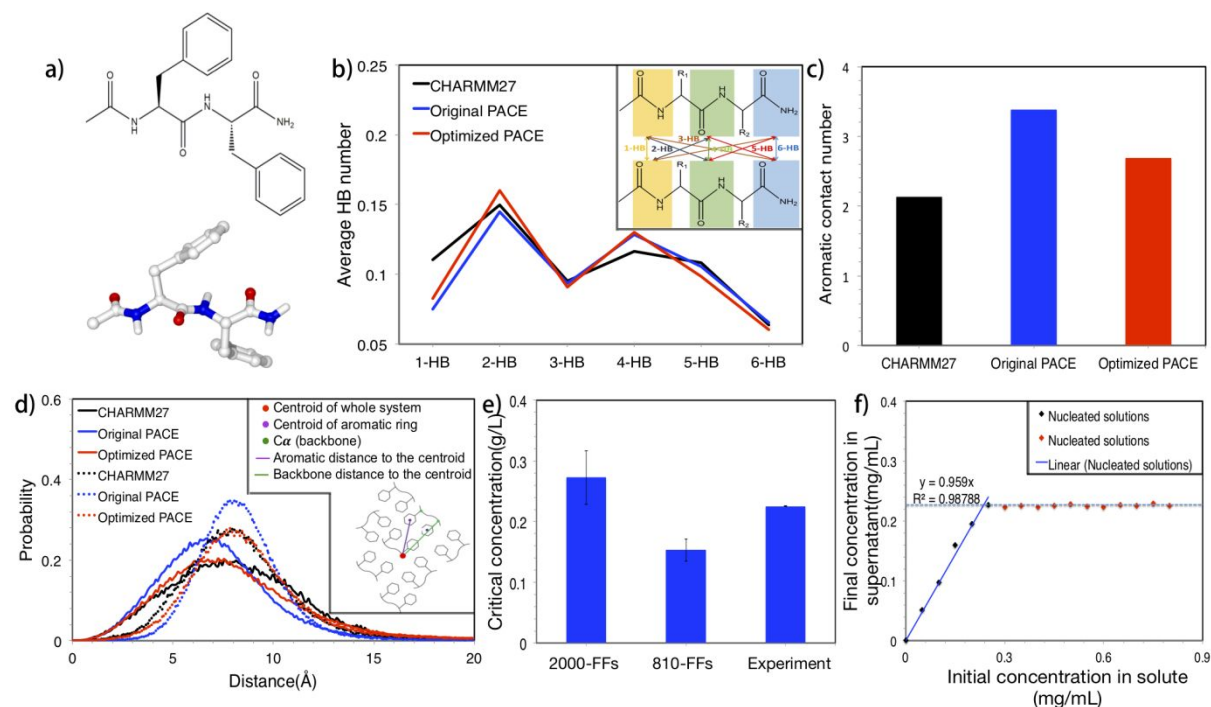


Figure 1. a) Chemical structure of FF and its representation using PACE. Color code: red for oxygen; gray for carbon; blue for nitrogen; white for hydrogen. b) Average count of HBs formed between any two FFs calculated according to simulations of a 8-FF system using all-atom models (black) and original (blue) and optimized PACE (red). HB contacts were analyzed for each possible pair of peptide bonds from any two FFs (see the inset). A HB between two peptide bonds was thought to arise if the donor-acceptor distance is shorter than 0.35 nm and the donor-hydrogen-acceptor angle is greater than 120°. c) Average aromatic contacts involved with each aromatic side chain. Two aromatic rings form a contact if their centroid distance is shorter than 0.65 nm. d) Distributions of distance of aromatic side chains (solid curves) or C α atoms (dotted curves) to the centroid of the 8-FF system. The inset illustrates the definitions of these distances. e) Critical concentration of FF self-assembly obtained from the simulations and the absorbance experiment. Error bars of the simulated results were obtained as block average by dividing the last one third of the simulations. f) Supernatant concentration of FFs measured at 330 K as a function of initial concentrations of FFs added.

We compared the critical concentration of FFs derived from our simulations and experiments. Computationally, the critical concentration of FFs was determined by counting the average number of dissociated FFs observed in microsecond-long simulations (**Figure 1e**). The calculated critical concentrations of the two FF systems tested were similar, both within a range of 0.15-0.3 mg/mL. In addition, following previous studies,¹⁹ we also experimentally established the critical concentration of FFs by monitoring the dependence of concentration of soluble FFs as a function of total FF concentration (See Method). As shown in **Figure 1f**, the peak absorption originating from the free FFs leveled off after a steady rise, indicating a C_s of 0.23 mg/mL for the FFs. The good agreement of C_s between both the experiments and our simulations of the two systems with different sizes (810 FFs and 2000 FFs) suggests that the improved PACE may be applicable for larger FF systems.

Diphenylalanine self-assembles into vesicles and tubes with double-layered structures

Multiple self-assembly simulations were conducted with the improved PACE, starting with 810 randomly dispersed FFs (**Tables 1 and S2**). A repeating simulated annealing procedure was employed to ensure that thermodynamically stable structures were sampled as well (see Method).

Table 1. The observed assembled morphologies in simulated annealing simulations of 810 FFs.

concentration (mg/ml)	total no. of MD runs	vesicle-like	tubular	sheet
107	5	3	1	1
173	5	0	4	1

In all the simulations, we observed the assembly of FFs into double-layered structures with a thickness of 1.4 nm with the backbones of FFs constituting the surfaces of the double-layered structures and the side chains buried inside. This supramolecular structure was reminiscent of a lipid bilayer. Similar assembled structures have also been reported previously in several CG simulations for uncapped FFs.^{14,26,28}

At a FF concentration of 107mg/mL, the double-layered structures primarily rolled up into hollow vesicles and at a higher concentration (173mg/mL), hollow tubes became dominant, extending infinitely in one dimension due to the periodic boundary conditions. At both concentrations, extended 2D sheets were occasionally observed (**Figure 2a**). This phenomenon, where morphology is concentration-dependent, is largely in line with previous CG simulations and experimentations.^{14,26,58}

Previous scanning electron and transmission electron microscopy (SEM/TEM) studies indicate that capped FFs in water mainly assemble into long hollow nanotubes.^{7,56} Our SEM measurement showed very similar assembled structures to the previous experimental observations (**Figure 2b**), and thus the results from the assembly simulations of 810 FFs agree qualitatively with the experimental studies. Of note, the tubes found in both previous experiments and ours were ~100 nm up to several microns in diameter and were much wider than observed (~6-7 nm) in the simulations (**Figure S2**). In addition, the tubes and vesicles observed experimentally usually have thick walls that likely involves multi-laminar structures.^{14,56} The smaller size of the tubes observed in the simulations could be attributed to far fewer FFs in simulations than needed for the formation of tubes with sizes comparable to what were observed in experiments. Therefore, more FFs may need to be included in simulations in order to capture large tubular structures. Indeed, as will be shown later, a microsecond long simulation of FF self-assembly for a bigger system that contains 2000 FFs allowed us to observe a tube-like structure with a larger diameter (~10 nm, **Figure S2**). However, the number of FFs needed for simulations of realistic tube structures observed experimentally is currently prohibitive. This challenge is also common for many other simulation studies.^{23,28,50}

We used Fourier transform infrared spectroscopy to analyze secondary structures of the FF assemblies to further validate the simulation results. The measured spectra exhibited strong absorption at 1500-1580 cm^{-1} and 1600-1700 cm^{-1} , corresponding to amide II and I regions associated with C=O vibrations, respectively. The absorption band in the amide I region mainly was centered at $\sim 1641 \text{ cm}^{-1}$ while also showing considerable absorption at 1620-1640 cm^{-1} , 1660-1685 cm^{-1} and $\sim 1690 \text{ cm}^{-1}$, indicating a coexistence of several types of secondary structures including coils, β -sheet and β -turns.^{59,60} Following a recent study of secondary structure characterization of short peptide assemblies,⁶¹ we assessed fractions of different types of secondary structures by fitting the amide I band to multiple Gaussians distributions (see **Figure 2c** and **SI**). The fitting indicated that the coils are the dominant structural type, but there are still $\sim 34\%$ and $\sim 20\%$ chances to observe β -sheet and β -turns, respectively. We then calculated secondary structural contents from our simulations of the assembled structures formed by 810 FFs, which agreed well with those from the FTIR measurements (**Figure 2d**). This suggests that, despite the limitations mentioned above, the self-assembled structures in our simulations may still provide a reasonable model for identifying conformational details of FF assemblies.

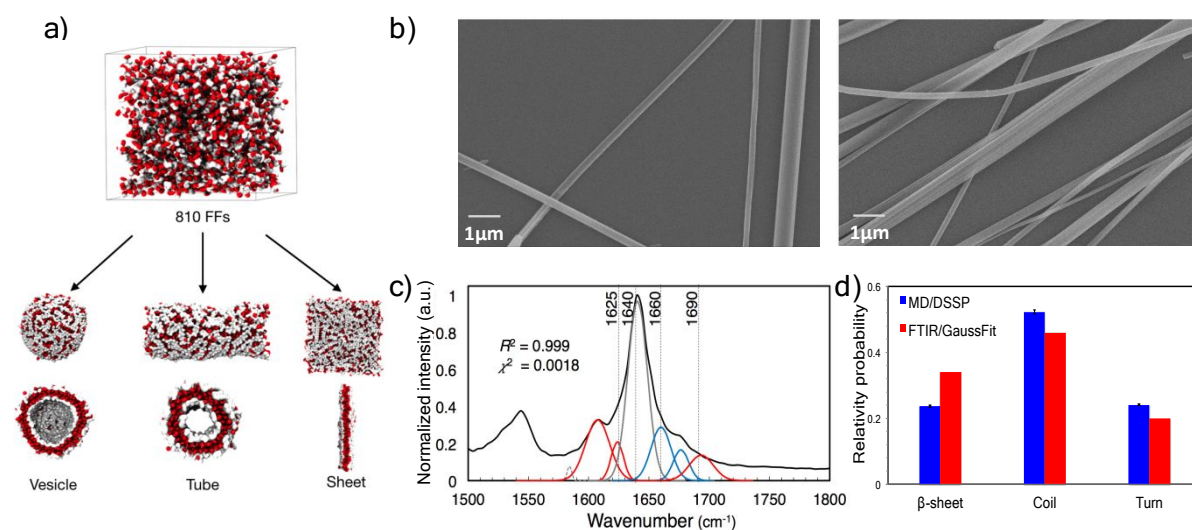


Figure 2. a) Representative assembly structures observed from annealing simulations. The backbones and side chains of FFs are shown as white and red ellipsoids, respectively. b) SEM images of FF assemblies. c) Normalized FTIR spectrum of FF assemblies at 1500-1800 cm^{-1} (black curve). The results of the Gaussian fitting are shown as red (β -sheet), blue (turn) and gray peaks (coil). The coefficients (χ^2 and R^2) of the fitting are also shown. d) Secondary structural contents of FF assemblies calculated from simulations (blue) and experiment (red).

FFs display different conformational preference in water and in assemblies

Next, we analyzed the conformational features of FFs in the self-assembled structures and compared these features with those found in soluble states. In the conformational analysis, all observed FF conformations were categorized into 12 different conformational states (see **Table S3** and **SI**). Each conformational state of a FF was defined by the backbone conformational states of the

two amino acid residues of the FF. The backbone conformational states of an amino acid residue represented free energy minima in the Ramachandran map of backbone dihedrals (ϕ , ψ), typically including a right-handed (α_R) and a left-handed α helical minima (α_L), along with a shallow minimum (B) comprising a β strand substate and a polyproline II substate (PPII) separated by a small free energy barrier (**Figure 3a**).⁶² We assigned a different name to each of the conformational states of the FF. For example, when the first and the second residues of a FF are in the α_R and β minima, respectively, the conformation of the FF was named $\alpha_R\beta$.

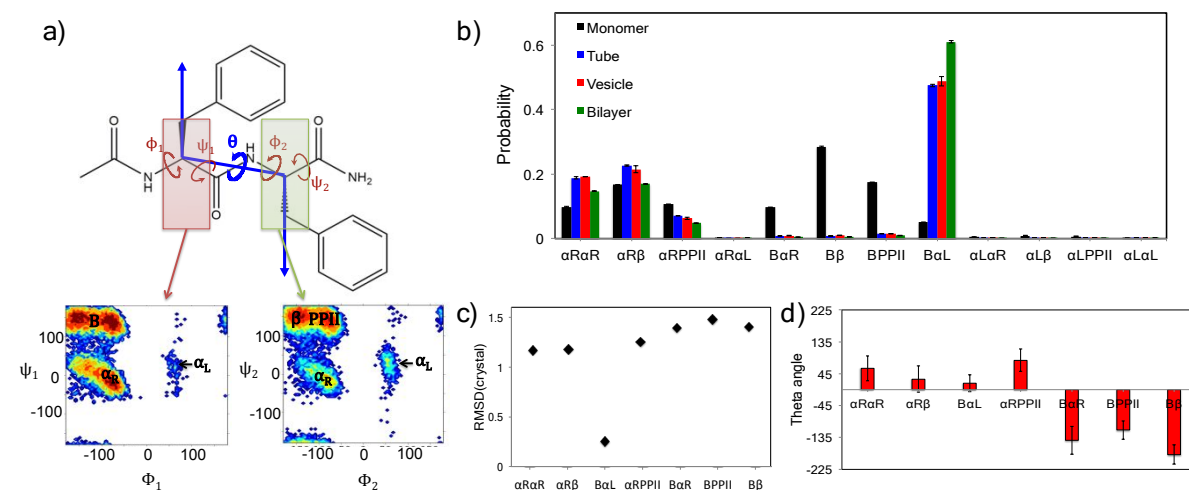


Figure 3. a) Illustration of conformations of each amino acid of a FF. The Ramachandran plots were obtained from a simulation of a FF in water using PACE. b) Conformational distributions of the FF in water (black) and in double-layered assemblies with tubular (blue), vesicular (red) and sheet-like (green) morphologies. c) RMSD of representative capped FF structures to that of crystalized uncapped FFs. The RMSDs were calculated based on the overlapping backbone atoms of the capped and uncapped FFs and their C_β atoms. d) θ angle of the major FF conformations. All error bars were obtained using three independent simulations.

We first examined conformational features of soluble states of FFs by performing a 2- μ s simulation of a single FF in water. Our simulation indicated that the conformation of a FF in water is rather diverse, with seven of the 12 conformational states having a non-negligible probability ($\sim 5\%$ - 30%). The most favorable conformation was an extended β -strand (B β) conformation, which is consistent with the high tendency of phenylalanine to adopt a β conformation.⁶³ To confirm this finding, we also conducted conformational analyses on structures of FF segments taken from a coil library in the protein data bank (PDB) (see **SI**). PDB coil libraries have been proven to be accurate in characterizing intrinsic conformational propensities of amino acids,⁶⁴⁻⁶⁶ and as shown in **Figure S3**, there is good agreement between the conformational distribution of FFs obtained from our soluble-state simulation and those from the coil library. This finding also demonstrated the ability of PACE to describe conformations of soluble FFs.

1
2
3 **Figure 3b** shows the conformational distributions of FFs in various double-layered assemblies,
4 which include vesicles, tubes and sheets. These assembly structures were obtained from the self-
5 assembly simulations of 810 FFs mentioned earlier. The conformational preferences of FFs in these
6 assemblies are similar regardless of their morphologies, and the comparison with the data from the
7 simulation of a free FF demonstrated a marked difference between the conformational distributions of
8 FFs in water and in the assemblies. In particular, three major conformational states ($B\alpha_R$, $B\beta$ and
9 $BPPII$) accessible to soluble FFs rarely formed in the assembled structures while the $B\alpha_L$ conformer
10 that was scarcely populated ($\sim 5\%$) in water became the predominant conformation ($>50\text{-}60\%$) in the
11 assembled structures. Interestingly, this conformation was very similar to the crystal structure of
12 uncapped FFs. The root mean square distance (RMSD) between the representative structure of $B\alpha_L$
13 and the crystal structure was $\sim 0.2 \text{ \AA}$ and was much lower than the RMSD values ($>1.1 \text{ \AA}$) obtained for
14 the other conformations. This result supports the previous assumption that FFs adopt the crystal-like
15 conformation in aqueous assemblies (**Figure 3c**).^{8,25,43,44,67}

24 **Molecular basis of conformational preference of FFs in self-assembled structures**

26 The results above suggest that assembled structures of FFs are conformationally specific, and to
27 understand the molecular basis of this conformational specificity better, we performed a series of
28 conformationally constrained simulations of 810 FFs in which constraints were applied to maintain all
29 the FFs in the system in single conformational states (see Method). This approach provides a useful
30 means by which to investigate the correlation between a given FF conformation and assembled
31 structures under idealized circumstances. The conformational states examined were those with an
32 intrinsic probability greater than 5% (**Figure 3b**).

34 The constrained simulations clearly indicated that the FFs constrained to different conformations
35 could self-assemble into nanostructures with distinct morphologies (**Figure 4a**). We analyzed the
36 geometric features of FF conformations to provide insight into the relationship between each FF
37 conformation and its corresponding assembled morphology. We first examined the angle θ between
38 the two side chains of a FF, a key geometric parameter of FFs that have been previously
39 demonstrated.^{23,28,43} When θ is small (closer to 0°), the two side chains are positioned on the same side
40 of peptide backbone (called a *syn* arrangement); when θ is large (closer to 180°), the two side chains
41 are positioned on the opposite sides of the backbone (called an *anti* arrangement).

43 Our analysis revealed that when FFs were constrained to the conformations $B\alpha_L$, $\alpha_R\alpha_R$, or $\alpha_R\beta$,
44 double-layered structures such as hollow tubes were formed (**Figures 4a** and **S4**). All three of these
45 conformations exhibited small θ ($|\theta| < 60^\circ$, **Figure 3d**), and when the FFs were fixed to $\alpha_R PPII$ whose θ
46 was a right angle, we observed the formation of irregular porous assemblies containing ring-shape
47 substructures. For $B\alpha_L$, $B\beta$ and $BPPII$ that had large θ angles ($|\theta| > 115^\circ$), the constrained simulations
48 always gave rise to aggregates in spherical or rod-like shapes. These aggregates lacked any large
49
50
51
52
53
54
55
56
57
58
59
60

internal hollow chambers as were also seen in the double-layered assemblies (**Figure 2a**). Of note, a very thin tunnel with a diameter of 1-1.5 nm was observed in the elongated aggregates formed by FFs when fixed to a B β conformation (**Figure S5**). Collectively, these findings suggest that a correlation exists between the side-chain arrangement of FFs and assembled morphologies.

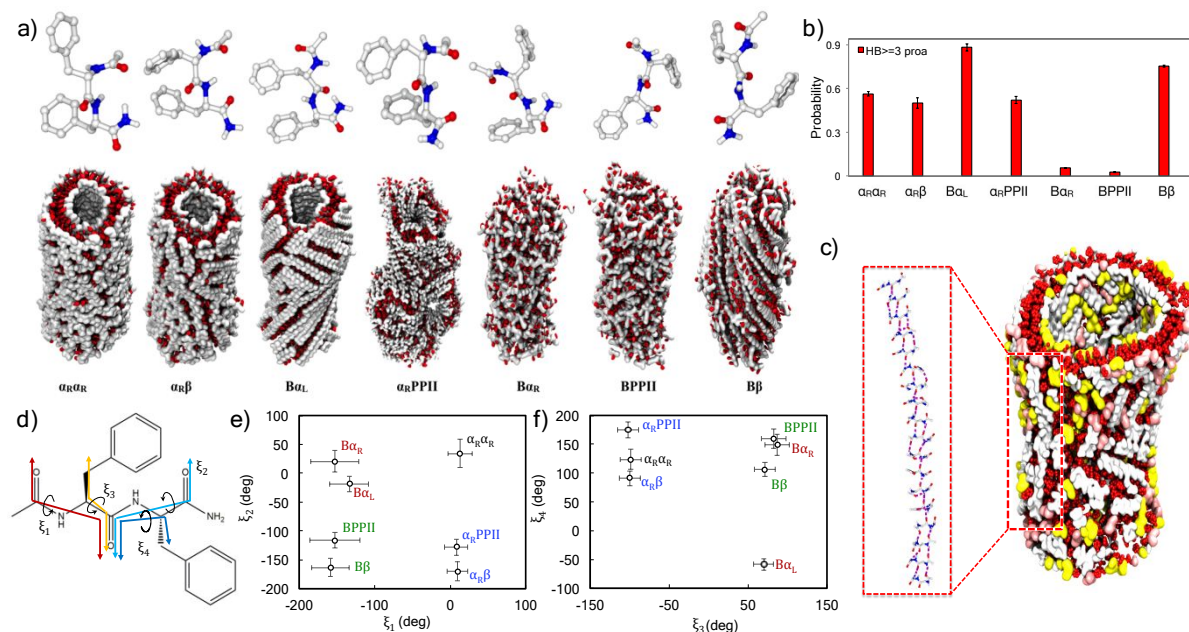


Figure 4. a) Representative structures of the major FF conformations (top) and corresponding assembled supramolecular structures (bottom) when FFs were constrained to each of these conformations during self-assembly simulations. The backbones and side chains of FFs are shown as white and red ellipsoids, respectively. b) Probability of finding FFs involved in continuous HB chains when FFs were fixed to different conformations. A chain is formed through repeated parallel stacking of at least three FFs (see SI). c) Representative tubular structure from an unconstrained simulation. FFs in $\alpha_R\alpha_R$, $\alpha_R\beta$, or B α_L are colored yellow, pink, or white, respectively. Inset is a close-up view of one of the HB chains. d) Definition of ξ_1 , ξ_2 , ξ_3 and ξ_4 . Shown in panels e) and f) are the (ξ_1, ξ_2) and (ξ_3, ξ_4) plots of the major FF conformational states. The circles denote the representative structures of the conformational states. Error bars show the standard deviation of structures that belong to each conformational state. The conformational states that are close to each other in the (ξ_1, ξ_2) map are marked in the same color. No two conformational states are close to each other in both plots except for B β and BPPII.

Aside from the overall morphologies, the packing of FFs in the assemblies observed in the constrained simulations was also investigated. Different patterns of molecular packing were observed in the same types of assembled morphologies (either double-layered assemblies or aggregates) depending on the conformational states of FF building blocks (**Figures 4a** and **S4**). Highly ordered packing of FFs in the double-layered assemblies and the aggregates was observed only when FFs were constrained to B α_L or B β , respectively, and involved repeated parallel stacking of FFs that allowed continuous HB chain structures to form. There is a considerably lower probability of the involvement of the other conformational states in the HB chain structures (**Figure 4b**).

It would be helpful to understand what geometric features of FFs could be responsible for such diversity of molecular packing in the assembled structures. As the major forces driving molecular

packing of FFs include aromatic attraction between side chains and HB interactions between backbone, the correlation between the observed packing patterns and the geometric features pertaining to these interaction types was investigated. We first examined the relative arrangement of the two side chains of FFs as it could affect how the FFs form aromatic contacts with their nearest neighbors. However, since the conformations compatible with the same type of assembled morphology exhibit a similar side-chain arrangement, this geometric factor alone cannot explain the observed difference in patterns of FF packing. Therefore, we focused instead on the geometric features pertaining to intermolecular HB interactions. Specifically, we analyzed the relative orientation of backbone HB donors and acceptors that were described by the angles (ξ_1 and ξ_2 , shown in **Figure 4d**) between the three peptide bonds of a FF and those (ξ_3 and ξ_4) between the side chains and the central peptide bond. The internal orientation distribution of HB donors and acceptors of FFs can affect their coordination propensities, and the analysis revealed a large variation in ξ s among the different conformations. No two conformations exhibited similar values ($|\Delta\xi| < 90^\circ$) in all ξ angles except for B β and BPPII (**Figures 4e** and **4f**). This large variation in the orientation of HB donors and acceptors may have accounted for the diversity observed for the FF packing in the assemblies.

Table 2. Non-covalent contacts formed in the assembled structures observed in conformationally constrained simulations.^a

	Backbone/water	Side chain/side chain	HB
$\alpha_R\alpha_R$	5.5 \pm 0.1	160 \pm 2	1.52 \pm 0.01
$\alpha_R\beta$	6.8 \pm 0.1	158.2 \pm 0.2	1.47 \pm 0.01
B α_L	7.4 \pm 0.2	168.4 \pm 0.2	1.69 \pm 0.03

a) All the contacts were averaged over all FFs. Two particles were thought to form a contact if they are closer than 0.4 nm.

The insights derived from the ideal constrained simulations helped us in understanding the results of the more realistic unconstrained simulations for the 810-FF systems. The three conformational states, including B α_L , $\alpha_R\alpha_R$ and $\alpha_R\beta$, that were predicted by the constrained simulations compatible with the double-layered structures were also conformations that were frequently observed (with a chance >15%) in the unconstrained simulations (**Figure 3b**). The other conformations that were predicted to be favored in aggregates, such as B β , B α_R and BPPII, had a negligible chance (<1%) to occur in the unconstrained simulations.

In addition, the results from the constrained simulations also aided in explaining why the crystal-like conformation (particularly, B α_L) was more favorable than $\alpha_R\alpha_R$ and $\alpha_R\beta$ for the double-layered assemblies (**Figure 3b**). The constrained simulations allowed us to analyze the degree of stabilization the assemblies would gain if FFs predominantly adopted a particular conformation state. Our analyses covered several types of stabilizing forces that were regarded previously as crucial for peptide assembly, including inter-peptide HB, aromatic interactions and interactions between solvent and the FF backbone.^{25,48,49} FFs adopting B α_L consistently provided 5-10% more stabilization to the double-

1
2
3 layered assemblies than they did when adopting $\alpha_R\alpha_R$ or $\alpha_R\beta$ for all interaction types examined (**Table**
4 **2**). This additional stability renders $B\alpha_L$ more favorable than $\alpha_R\alpha_R$ and $\alpha_R\beta$ in the double-layered
5 assemblies, even though $B\alpha_L$ is much less populated in water.
6
7

8 According to the constrained simulations, when FFs adopts $B\alpha_L$, the doubled-layered structures
9 with continuous HB chain structures would arise and, since $B\alpha_L$ was determined to be the most
10 abundant conformation in the unconstrained simulations, one would expect that the HB chains would
11 be prevalent structural features of the doubled-layered assemblies. The representative structure of the
12 double layered assemblies taken from the unconstrained simulations indeed showed that there are
13 many HB chain structures mainly formed by the FFs in the $B\alpha_L$ conformation (**Figure 4c**).
14
15
16
17
18

19 **Conformational transitions during the self-assembly of FFs**

20

21 The conformational difference between soluble and assembled FFs observed in the present study
22 corroborates the suggestion that there is a large conformational transition of FFs during the self-
23 assembly process.³⁹ In order to gain insight into the nature by which this transition takes place, we
24 conducted an unbiased simulation of FF self-assembly. As suggested by a previous simulation study,²⁸
25 although simulations of hundreds of FFs are efficient in sampling basic nanostructures assembled by
26 FFs, large simulation systems containing several thousand FFs may often be needed for more realistic
27 modeling of complex processes of FF self-assembly. As such, our unbiased simulation included 2000
28 FFs in a box with a size of $30\times 30\times 30$ nm³ and the length of the simulation reached 2.2 μ s.
29
30
31
32
33

34 Shortly after the simulation started ($t\sim 10$ ns), the FFs that were initially dispersed had quickly
35 assembled into small clusters containing 20-30 monomers (**Figures 5a** and **b**). During the next 150 ns
36 of simulation, these small clusters associated into small double-layered patches (at $t\sim 50$ -100 ns) that
37 continued to merge into several large patches. The two largest patches then curved up into vesicles,
38 and the two resulting vesicles collided at $t\sim 500$ ns, eventually fusing into a capsule. Although a
39 tubular structure was not observed, the cylindrical part of the capsule displayed a morphology very
40 similar to the tubular assemblies that had been observed in our annealing simulations (**Figure 2a**). As
41 this capsule would have continued to grow into a tube if it absorbed more vesicles, our results agree
42 with the previous suggestion that the FF tubes may result from the fusion of vesicles.^{11,14,58}
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

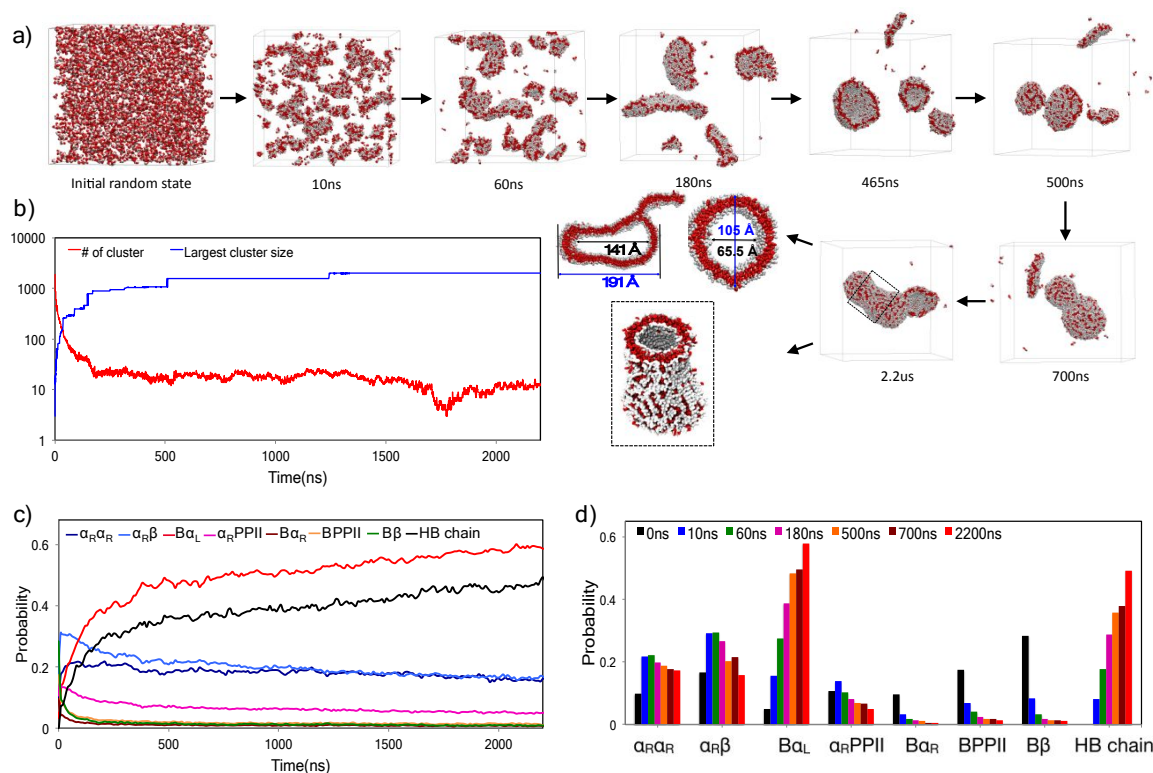


Figure 5. a) Snapshots taken from self-assembly simulation of 2000 FFs. b) Time evolution of the number of clusters (red curves) and the number of FFs contained in the largest cluster (blue lines) during the simulation. A group of FFs was thought to form a cluster if any of these FFs has at least one atomic contact with the others. Two atoms were thought to contact each other if their distance is shorter than 0.45 nm. c) Time evolution of conformational distribution of FFs and probability of finding a FF in HB chains during the simulation. d) Conformational distribution of FFs and probability of finding HB chains at select time points of the assembly simulation.

We monitored the change in probability of the major conformational states of FFs during the assembly process of 2000 FFs. The probabilities of detecting $B\beta$, $B\alpha_R$ and $BPPII$ (as shown in **Figures 5c** and **d**) diminished rapidly within the first 50-100 ns of the simulation and remained negligible beyond this time period. Conversely, the probabilities of $\alpha_R\alpha_R$, $\alpha_R\beta$ and $\alpha_R PPII$ rapidly increased at the beginning of the simulation and then peaked early ($t \sim 50-100$ ns), gradually decreasing until the end of the simulation. As for $B\alpha_L$, its probability quickly increased initially and then continued to increase more slowly throughout the remaining simulation. At the early stage of the assembly, it seems that FFs rapidly redistributed their conformational states to eliminate the conformations with large θ and instead populate those with small θ . This conformational transition appears necessary for initial formation of small clusters and double-layered patches. At the later stages of the assembly, a slow conformational convention took place, shifting the population towards the $B\alpha_L$ conformation. Since the majority of free FFs had already been absorbed after the initial stages of the assembly, the conformational changes afterwards occurred within the aggregates. This means that the aggregated environment did not suppress the conformational transitions of FFs, and our analysis of state-to-state transition over the second half of the simulation of 2000 FFs indeed revealed

1
2
3 that the transition times between different FF conformations in the assemblies was at most ~ 10 ns
4 (**Figure S6**). These transitions were much faster compared to the assembly process.

5
6
7 The probability of observing FFs in $B\alpha_L$ throughout the simulation was closely correlated with the
8 probability of observing FFs involved in the continuous HB chain structures (**Figures 5c and d**). This
9 result corroborates our hypothesis that $B\alpha_L$ is needed for FFs to provide ordered intermolecular
10 packing in the double-layered assemblies, therefore indicating that the conformational conversion of
11 FFs into $B\alpha_L$ could be an important type of conformational transition that leads to the ordered packing
12 of FFs.
13
14
15

16 **Importance of transient conformational change for self-assembly of FFs**

17
18
19 Our simulation of self-assembly of 2000 FFs revealed that even in aggregates, there were rapid
20 conformational fluctuations of FFs that were responsible for the conversion of FFs from soluble to
21 assembled states. We performed a series of conformationally constrained simulations to determine
22 how the dynamics of several key steps of the self-assembly process might be affected by the
23 suppression of conformational fluctuations of FFs. The steps examined included the initial formation
24 of double layered patches, the closure of the patches into vesicles and the fusion of the vesicles into
25 capsules.
26
27
28
29

30
31 Specifically, we performed multiple test simulations containing randomly dispersed FFs to examine
32 the assembly dynamics of flat patch structures formed at the initial stage of the assembly process.
33 According to the simulation of 2000 FF, this type of patch structures normally contained about 100
34 FFs at least. Hence, 100 FFs were included in these test simulations. In half of the test simulations, the
35 conformations of FFs were constrained in such a way that the resulting conformational distribution
36 was consistent with that obtained at the time point ($t=200$ ns) of the simulation of 2000 FFs when the
37 patches had already occurred (see Method). In the other half of the test simulations, no such
38 constraints were applied. In all of the simulations, we observed the formation of a small double-
39 layered patch at $t\sim 100$ ns (**Figure S7**), indicating that the conformation fluctuation of FFs is not
40 essential for the formation of these patch structures.
41
42
43
44
45

46
47 After the assembly dynamics of small patches, we then examined the closure of the double-layered
48 patches. The test simulations began with a flat patch containing ~ 500 FFs taken from a frame at $t=$
49 180 ns of the simulation of 2000 FFs. As done previously, these simulations were performed in the
50 absence or in the presence of conformational constraints, and in the constrained test simulations, each
51 FF was constrained to its starting conformation. In the constraint-free test simulations, vesicles
52 formed within 350 ns, as indicated by a large reduction in the solvent accessible surface area of
53 aromatic side chains that were exposed at the patch edges. In contrast, the patches in the 1-us
54 constrained simulations were far from being closed, leaving a large number of aromatic side chains
55
56
57
58
59
60

exposed at the patch edges (**Figures 6a** and **S8**), indicating that the suppression of the conformational fluctuation of FFs impeded vesicle closure.

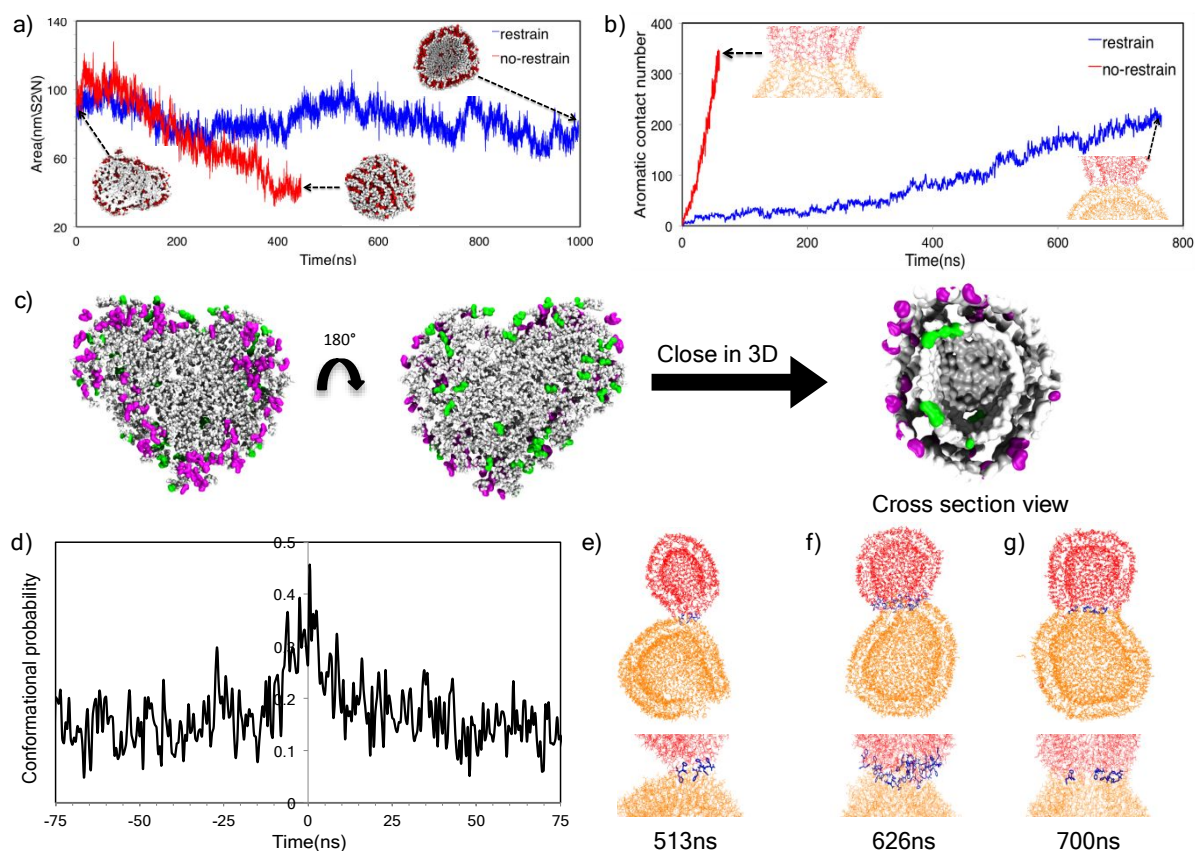


Figure 6. Simulation tests of impact of conformational fluctuation on closure of the double-layered patches (a) and fusion of vesicles (b). Blue and red curves denote the results from the simulations with and without conformational constraints, respectively. The simulations of patch closure and vesicle fusions began with a patch containing ~ 500 FFs and two vesicles containing in total ~ 1500 FFs, respectively. Shown in (a) is the time evolution of total accessible surface areas of side chains which mainly arose from the exposed edges of the patch. Shown in (b) is the time evolution of aromatic contacts formed between aromatic rings from different vesicles (colored red and orange, respectively). c) Repartitioning of FFs between the two leaflets of a double-layered patch during the closure of the patch. The patch and the vesicle evolved from the patch were taken from the simulation of 2000 FFs at $t=180$ ns and $t=510$ ns, respectively. The FFs in each leaflet that traversed the double-layered structure are marked purple and green, respectively. d) Probability of finding FFs in $B\beta$, $B\alpha_R$, $BPPII$ and $\alpha_R PPII$ during their traverse movement across the double-layered structure. The results were averaged over the trajectories of all the FFs that had traversed the double-layered structures. The time zero corresponds to the moment that a FF was positioned between the two leaflets during its traverse movement. e-g) Snapshots during fusion of two vesicles taken from the 2000-FF system. The two vesicles are colored red and orange, respectively. Marked blue are the FFs that form aromatic contacts with both vesicles while adopting one of $B\beta$, $B\alpha_R$, $BPPII$ and $\alpha_R PPII$ conformations.

Similarly, our test simulations of vesicle fusion began with a system containing two vesicles that were about to form initial contact. The coordinates of the system (containing ~ 1500 FFs) were taken from a frame of $t=502$ ns of the simulation of 2000 FFs, and the fusion event usually occurred within the first 200 ns of the unconstrained test simulations, but never took place during the 600-800 ns constrained simulations (**Figures 6b** and **S9**). At the end of the constrained test simulations, the two

1
2
3 vesicles partly shared their walls at best, but by no means joined together to form a connected internal
4 chamber and thus indicated that the conformational fluctuation of FFs is also important for the
5 process of vesicle fusion.
6
7

8 We sought to understand further the importance of the conformational fluctuation of FFs to the
9 formation and fusion of vesicles. Since the double-layered patches had very similar conformational
10 distributions of FFs to those of the vesicles (**Figure 3b**), it was supposed that there would be little
11 need for FFs to change their conformational distribution when the patches curved into vesicles.
12 Despite this, we noticed that the partition of FFs between the two leaflets of the double-layered
13 structures was different for the patches and vesicles. For instance, the large flat patches observed in
14 our 2000-FF simulation contained almost equal numbers of FFs in each of their two leaflets (**Table S4**)
15 and once they turned into vesicles, the outer leaflets of the vesicles contained 100-150 more FFs than
16 their inner leaflets (probably due to the smaller surface area of the inner leaflet of a vesicle being able
17 to accommodate fewer FFs). Based on this information, the transformation of the patches to the
18 vesicles must involve a redistribution of FFs for both leaflets.
19
20
21
22
23
24
25

26 The simulation of 2000 FFs also allowed us to trace the movement of every single FF during the
27 formation of vesicles, and we found that tens of FFs in the inner leaflet of the vesicle actually came
28 from the edge part of the patch leaflet that had eventually evolved into the outer leaflet of the vesicle
29 (**Table S4 and Figure 6c**). Conversely, twice as many FFs in the outer leaflet of the vesicles came
30 from the edge part of the other leaflet from the flat patches. This finding indicates that the
31 repartitioning of FFs between the two leaflets of the patches was mainly achieved by the flip-flop of
32 FFs in the patches.
33
34
35
36
37

38 We then examined why the conformational fluctuation of FFs was important for the flip-flop of
39 these molecules. We calculated a time series of the conformational states during the vesicle formation
40 for each FF involved in flip-flop and all the resulting time series were aligned such that their time
41 origins coincide to when the FFs happened to be located between the two leaflets. These aligned time
42 series provided a statistical description of conformational change associated with the flip-flop of FFs.
43
44
45

46 As shown in **Figure 6d**, the population of the conformations with large θ increased transiently
47 (from 10-15% to 45-50%) and peaked half way during the flip-flop of the FFs even though the
48 conformations with small θ were dominant conformations of FFs in the double-layered structures.
49 This finding suggests that a transient conformational transition of FFs to those with large θ seemed
50 crucial for the flip-flop of FFs and subsequently, for the transformation of the patches into the vesicles.
51
52
53

54 Additionally, we identified all the FFs adopting this type of conformation during the fusion process
55 in the interfacial regions of the two vesicles (**Figures 6e-g**) to test if a similar transient conversion of
56 FFs may also be needed for the fusion process. Again, the results from the simulation of 2000 FFs
57 were used for the analysis. Upon the initial contacts of the two vesicles, most of FFs at the interface
58
59
60

1
2
3 needed to assume the conformations with a larger θ (**Figure 6e**), since a FF in these conformations
4 can keep one of its side chains in the vesicle where the FF originally stayed while inserting its other
5 side chain into a second vesicle. At the middle and late stages of the fusion process, there was also a
6 largely increased chance to observe the FFs assuming the large- θ conformations in the junction region
7 of the two vesicles (**Figure 6f**), but once the fusion process ended, the chance to observe the large θ
8 conformations in this region diminished (**Figure 6g**). Based on these findings, a transient conversion
9 of FFs into the conformations with large θ is most likely essential at various stages of the vesicle
10 fusion process.
11
12
13
14
15

16 CONCLUSIONS

17
18 The rational design of self-assembled materials relies on a deep understanding of the relationship
19 between the building block structures and resulting assembled structures. In this study, we explored
20 through extensive simulations the conformational complexity of diphenylalanine (FF) molecules and
21 elucidated upon the influences of the conformational complexity regarding both morphologies and
22 molecular packing associated with assembled structures as well as assembly dynamics.
23
24
25

26 Our simulations demonstrated that there are numerous conformations accessible to FFs that exhibit
27 diverse molecular geometries. The conformations involving a *syn* arrangement of the two aromatic
28 side chains best fit in double-layered structures that can further roll up into hollow vesicles and tubes
29 similar to those observed experimentally. While a similar role of this type of side chain arrangement
30 has also been reported by others,^{23,28,43} we showed that the FFs with a large angle between the side
31 chains tended to assemble into ordered or partially ordered aggregates lacking in any of the large
32 internal cavities that was normally observed in the double-layered assemblies. Although these types of
33 aggregates have not been reported for FF assemblies in water, they are attainable in other solvent
34 environment.^{33,68} Additionally, a previous SEM/TEM study showed that a FF derivative
35 conformationally locked by chemical modification into an *anti* arrangement of side chains could
36 assemble into thin linear assemblies in water. Unlike those hollow assemblies formed by unmodified
37 FFs, these assembled structures exhibited very thin tunnels (with diameters < 2 nm) at best.³⁵ A similar
38 morphology was observed in our assembly simulations when FFs were constrained to one of the
39 conformations with large side-chain angles (**Figures S5**), which may represent an atomic model for
40 the assemblies that were observed experimentally. Collectively, our study reveals the critical role of
41 side-chain arrangement in determining assembled morphologies.
42
43
44
45
46
47
48
49
50
51

52 Furthermore, we have identified other geometric factors aside from the side chain arrangement that
53 also can affect molecular architectures of the assemblies. While multiple conformations that share the
54 similar side chain arrangement are compatible with assemblies with similar morphology (either
55 doubled-layered assemblies or aggregates), they usually differ from one other in terms of geometric
56 details like the relative orientation of backbone HB donors and acceptors. As a result, the packing of
57 FFs in the assemblies can be conformationally specific and only a select few conformations permit a
58
59
60

1
2
3 highly ordered packing of FFs in the assemblies. These results shed light on the intricate dependencies
4 of molecular architecture of assemblies on multiple FF geometric factors and highlight the need in
5 considering the full conformational details of peptide building blocks in their molecular design.
6
7

8 Although the FF conformation obtained from crystal has been widely accepted as a model for FF
9 conformation in assemblies,^{25,43,44,67} its relevance to these assemblies (which are formed in a fully
10 hydrated environment) remains ambiguous.^{17,45,46} To date, all-atom simulation studies have failed to
11 observe the conformation during the self-assembly of small systems containing tens of FFs in
12 water.^{25,48,49} Our simulations demonstrate that this conformation, although unstable in isolation or in
13 small clusters of FFs, can be gradually enriched and eventually become dominant during the
14 spontaneous assembly of many FFs into double-layered assemblies like hollow vesicles and tubes.
15 When compared to the other conformations, the special geometry of this conformation allows FFs to
16 pack into assemblies in an energetically favorable, highly ordered manner. Therefore, our results
17 provide compelling evidence that supports the relevance of the crystal conformation to the assembled
18 hollow tubes and vesicles.
19
20
21
22
23
24
25

26 It has been supposed that the dynamics of internal conformations of protein-like building blocks
27 play crucial roles in the assembly mechanism,^{40,69} with possible roles of the conformational
28 fluctuations having been probed computationally at CG level for several other assembly systems.^{41,42}
29 In the study presented here, our computational investigation of atomic details of FF assembly revealed
30 at least two possible roles of conformational fluctuations in the self-assembly of FFs: 1) the
31 fulfillment of the conformational transformation of FFs from disordered monomer states into ordered
32 aggregation states, which basically reiterate what was concluded in those CG studies,^{41,42} and 2) the
33 transient conversion of FFs into particular intermediate conformations that are required for the closure
34 of double layered patches into vesicles and fusion of vesicles into tubes. The observed intermediate
35 conformations usually displayed large side-chain angles, which does not match the structures that
36 were assembled eventually. Although the significance of this transient conformational conversion
37 needs to be confirmed by further experimentation, a similar phenomenon has already been reported in
38 several cases of peptide aggregation during which the transient formation of helical structures was
39 discovered to expedite the formation of amyloid fibrils rich in β sheet.^{70,71}
40
41
42
43
44
45
46
47
48

49 Our simulations explain the molecular rationale behind the role played by these intermediate
50 conformations further. During the transformation from a flat patch to a vesicle, a double-layered
51 structure must create an imbalance of FFs between its two leaflets achieved mainly through the flip-
52 flop of FFs in the patch with large side-chain angle conformations. Intriguingly, it was found in
53 previous simulation work that lipid molecules needed to open up their two tails when traversing
54 bilayer membrane.⁷² During the fusion of two vesicles, the FFs in the interfacial region also tended to
55 adopt similar conformations in order to facilitate the initial attachment of the vesicles along with the
56 subsequent opening of vesicle walls. This molecular explanation adds a crucial layer of mechanistic
57
58
59
60

1
2
3 details of the self-assembly of FFs and would be helpful for understanding self-assembly of protein-
4 like molecules in general.
5

6
7 We were able to observe the atomic details of the self-assembly process of hundreds and even
8 thousands of peptide molecules into experimentally observed nanostructures using the hybrid-
9 resolution model presented in this study. When combined with conformationally constrained
10 simulations, the model can further be utilized into assessing the conformational impact on the
11 assembly process and can predict possible assembled structures for conformationally restrained
12 peptides. In conclusion, this hybrid-resolution model shows potential as a valuable computational tool
13 for studying the self-assembly of peptide-based molecules like FFs and providing aid in the design of
14 related materials.
15
16
17
18
19

20 METHODS

21 **Peptide Synthesis.** Peptide synthesis was performed manually on Rink Amide MBHA resin (loading
22 capacity: 0.54 mmol/g) (GL Biochem Co. Ltd) by standard Fmoc-based solid-phase peptide synthesis.
23 Generally, Rink amide AM resin was pre-swelled with DCM/NMP (1/1) for 30 min. Fmoc
24 deprotection was performed with morpholine (50% in NMP) for 30min × 2. Then the resin was
25 washed with NMP (5 times), DCM (5 times), and NMP (5 times). Fmoc-protected amino acids (6.0
26 equiv according to initial loading of the resin) and HCTU (5.9 equiv) were dissolved in NMP,
27 followed by DIPEA (12.0 equiv). The mixture was pre-activated for 1 min and added to the resin for
28 1-2 h, then the resin was washed with NMP (5 times), DCM (5 times), and NMP (5 times). Upon
29 completion of peptide synthesis, peptides were N-terminally acetylated with a solution of acetic
30 anhydride and DIPEA in NMP (1:1:8 in volume) for 1 h. Peptides were cleaved from the resin with a
31 mixture of TFA/H₂O/EDT/TIS (94:2.5:2.5:1) for 2 h and concentrated under a stream of nitrogen.
32 The crude peptides were then precipitated with Hexane/Et₂O (1:1 in volume) at -20°C, isolated by
33 centrifugation then dissolved in water/acetonitrile, purified by semi-preparative HPLC and analyzed
34 by LC-MS.
35

36 **Preparation.** Fresh peptide stock solution was prepared by dissolving lyophilized dipeptide powder
37 in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP, Sigma) at a concentration of 100 mg/mL. The stock
38 solutions were then diluted to a desired concentration in deionized water.
39

40 **Scanning Electron Microscopy (SEM).** Immediately after dilution in deionized water (Final
41 concentration of 0.5mg/ml), peptide solution was placed on silicon slides and were left to dry at room
42 temperature, and coated with gold. Samples were then viewed using a scanning electron microscope
43 (ZEISS Supra 55, Oxford X-Max 20, 20 kV).
44

45 **Fourier-Transform Infrared Spectroscopy (FTIR).** Peptide solution taken from SEM assay and
46 dried by vacuum. KBr pellet was prepared prior to measurement. Qualitative analysis of the samples
47 was performed. Bruker Vertex 70 FT-IR spectrometer was used for FT-IR analysis with wavenumber
48 range from 4000 to 400 cm⁻¹.
49
50
51
52
53
54
55
56
57
58
59
60

Critical concentration. Critical concentration was measured at 258 nm at 330K, using a molar absorption coefficient of $390 \text{ M}^{-1} \text{ cm}^{-1}$, calculated from the molar absorption coefficient of phenylalanine. To explore the potential subcritical association in deionized water, solutions at concentrations 0.05–0.80 mg/mL were prepared. These were incubated for 1 h, and then ultracentrifuged for 1 h at 62000g. The supernatant was measured, and the absorption at 258 nm was plotted against concentration.

Parameterization of PACE model. The potential energy function of PACE is expressed as:

$$E = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}} + E_{\text{improper}} + E_{\phi,\psi,\chi1} \\ + E_{\text{CGW-CGW}} + E_{\text{CGW-UA}} + E_{\text{nonpolar}} + E_{\text{polar}} \quad [1]$$

where the first five terms describe the bonded energy and geometry and rotameric preference in proteins and the last four terms describe non-covalent interactions between protein sites, between CG water sites and between the two types of sites at different resolutions. These energy terms and their parameterization have been thoroughly explained in our previous studies and the SI.^{53,73}

In this study, the non-covalent parameters of PACE (version 1.3) were further fine-tuned to fit the results from simulations of eight FFs using the CHARMM27 force field⁷⁴ and the TIP3P solvent model⁷⁵ that has been validated for studying proteins and peptides self-assembly systems.^{48,76} Specifically, we adjusted the parameters for HB interaction strength, described as the well depth of Lennard-Jones potential between HB donor and acceptor atoms, to reproduce HB patterns formed between FFs (**Figure S1** and **Table S1**). In addition, we also adjusted the parameters associated with the interaction between peptide sites and CG water to correctly model aromatic-aromatic contacts and the distance-to-center distributions. We conducted an iterative trial-and-error procedure to optimize these parameters against the all-atom results while attempting to minimize change made to PACE. The optimized parameters are listed in **Table 3**. Of note, despite these changes, PACE can still be used to simulate *ab initio* folding of the mini-protein systems (**Figure S10**) that were employed previously to validate PACE.⁵³ This suggests that the changes made to the force field, though important for correct modeling of self-assembly, are fairly small and do not affect simulation results at a single molecular level.

Table 3. The parameters adjusted for self-assembly simulations in this study.^a

Interaction type	N-O	HA-O	CZ-W	CA-W	C-W
Original PACE	17.7	20	0.81	1.15	0.8
Optimized PACE	17.2	18.82	0.926	0.805	0.56

a) All the parameters adjusted are those associated with well depth (ϵ , in unit of kJ mol^{-1}) of LJ potentials describing nonbonded interactions between several pairs of atom types: namely N, O, HA, and C for amide nitrogen, oxygen, hydrogen and carbon, respectively, CA for C_{α} of backbone, CZ for carbon atoms in aromatic rings and W for coarse-grained water particles.

Models and Simulation Setup. All simulations in this study were conducted using the GROMACS 4.5.4 package.⁷⁷ In all atom systems, either a single FF or eight FFs were placed in a cubic box

1
2
3 containing 1181 and 1026 TIP3P water molecules, respectively. The particle mesh Ewald⁷⁸ (PME)
4 method was employed to calculate electrostatic interactions with a short range cutoff of 1.2 nm. The
5 REMD simulations were performed with 16 replicas at temperatures of 289-405 K. Temperature in
6 simulations was kept using the Nose-Hoover method. Simulation timestep was 2 fs with the LINCS
7 algorithm. Exchange between replicas was attempted every 10 ps with average acceptance ratio
8 of >20 %. The total simulation length was 750 ns for the monomer system and 3.2 μ s for the octamer
9 system. The data collected at 310 K were used for the fitting.

14 The simulation setup for each system modeled with PACE is summarized in **Table S2**. In general, a
15 number of FFs were initially randomly dispersed in a simulation box with a clearance of 0.5 nm. The
16 MARTINI CG water was used to solvate the systems. All the systems were subject to a 5000-step
17 energy minimization followed by a 5-ns pre-equilibration at 310K with system volume being fixed.
18 Then production runs were conducted at 1 atm. The Nose-Hoover method and the Parrinello-Rahman
19 method were used to maintain simulation temperature and pressure, respectively.

24 For simulated annealing simulations, the following temperature scheme was adopted. The systems
25 were first maintained at the highest temperature for 50-200 ns and then cooled down at a speed of 0.5
26 K per ns. The systems stayed at the lowest temperature for 20 ns before another annealing cycle starts
27 over. Typically, two to five annealing cycles were conducted for each simulation, followed by a long
28 time equilibrium run at 330K that were employed for analysis.

31 In conformationally constrained simulations, each FF was constrained to one of the 12 possible
32 conformational states *via* harmonic forces applied to ϕ_1 , ψ_1 , ϕ_2 and ψ_2 of the backbone of the FF. The
33 force constants and the equilibrium values of the harmonic forces were optimized against the results
34 of the unconstrained self-assembly simulations (**Figure S11**). These parameters are listed in **Table S5**.

38 ASSOCIATED CONTENT

39 Supporting Information

41 Tables S1-5, Figures S1-S11 and detailed descriptions of PACE force field and structural analysis
42 methods. This material is available free of charge *via* the Internet at <http://pubs.acs.org>.

46 The authors declare no competing financial interest.

49 AUTHOR INFORMATION

51 Corresponding Author

52 hanw@pkusz.edu.cn

53 lizg@pkusz.edu.cn

55 ORCID

56 Wei Han: 0000-0003-0759-1766

57 Zigang Li: 0000-0002-3630-8520

1
2
3 **Author Contributions**
4

5 Q.X. and Y.J. contributed equally to this work.
6
7

8 **ACKNOWLEDGMENTS**
9

10 We thank financial supports from the National Science Foundation of China (21673013 to W.H.,
11 21778009 to Z.G.L.) and the Shenzhen STIC (KQTD2015032709315529 to W.H.,
12 JCYJ20160330095839867 to W.H., JCYJ20170412150719814 to Z.G.L.). The computational work is
13 supported by High-Performance Computing Platform of Peking University.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

REFERENCES

1. Kim, S.; Kim, J. H.; Lee, J. S.; Park, C. B. Beta-Sheet-Forming, Self-Assembled Peptide Nanomaterials towards Optical, Energy, and Healthcare Applications. *Small* **2015**, *11*, 3623-3640.
2. Adler-Abramovich, L.; Gazit, E. The Physical Properties of Supramolecular Peptide Assemblies: from Building Block Association to Technological Applications. *Chem. Soc. Rev.* **2014**, *43*, 6881-6893.
3. Tao, K.; Makam, P.; Aizen, R.; Gazit, E. Self-Assembling Peptide Semiconductors. *Science* **2017**, *358*, No. eaam9756.
4. Tao, K.; Levin, A.; Adler-Abramovich, L.; Gazit, E. Fmoc-Modified Amino Acids and Short Peptides: Simple Bio-Inspired Building Blocks for The Fabrication of Functional Materials. *Chem. Soc. Rev.* **2016**, *45*, 3935-3953.
5. Lee, O. S.; Stupp, S. I.; Schatz, G. C. Atomistic Molecular Dynamics Simulations of Peptide Amphiphile Self-Assembly into Cylindrical Nanofibers. *J. Am. Chem. Soc.* **2011**, *133*, 3677-3683.
6. Reches, M.; Gazit, E. Casting Metal Nanowires within Discrete Self-Assembled Peptide Nanotubes. *Science* **2003**, *300*, 625-627.
7. Reches, M.; Gazit, E. Self-Assembly of Peptide Nanotubes and Amyloid-Like Structures by Charged-Termini-Capped Diphenylalanine Peptide Analogues. *Isr. J. Chem.* **2005**, *45*, 363-371.
8. Li Q.; Jia Y.; Dai L.; Yang Y.; Li J. Controlled Rod Nanostructured Assembly of Diphenylalanine and Their Optical Waveguide Properties. *ACS Nano* **2015**, *9*, 2689-2695.
9. 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.; Gazit, E. Ostwald's Rule of Stages Governs Structural Transitions and Morphology of Dipeptide Supramolecular Polymers. *Nat. Commun.* **2014**, *5*, 5219.
10. Smith, A. M.; Williams, R. J.; Tang, C.; Coppo, P.; Collins, R. F.; Turner, M. L.; Saiani, A.; Ulijn, R. V. Fmoc-Diphenylalanine Self Assembles to a Hydrogel via a Novel Architecture Based on π - π Interlocked β -Sheets. *Adv. Mater.* **2008**, *20*, 37-41.
11. Yan, X.; He, Q.; Wang, K.; Duan, L.; Cui, Y.; Li, J. Transition of Cationic Dipeptide Nanotubes into Vesicles and Oligonucleotide Delivery. *Angew. Chem.* **2007**, *119*, 2483-2486.
12. Reches, M.; Gazit, E. Formation of Closed-Cage Nanostructures by Self-Assembly of Aromatic Dipeptides. *ACS Nano* **2004**, *4*, 581-585.
13. Ryu, J.; Park, C. B. High Stability of Self-Assembled Peptide Nanowires Against Thermal, Chemical, and Proteolytic Attacks. *Biotechnol. Bioeng.* **2009**, *105*, 221-230.
14. Song, Y.; Challa, S. R.; Medforth, C. J.; Qiu, Y.; Watt, R. K.; Pena, D.; Miller, J. E.; van Swol, F.; Shelnutt, J. A. Synthesis of Peptide- Nanotube Platinum-Nanoparticle Composites. *Chem. Commun.* **2004**, *9*, 1044-1045.
15. Kol, N.; Adler-Abramovich, L.; Barlam, D.; Shneck, R. Z.; Gazit, E.; Rousso, I. Self-Assembled Peptide Nanotubes are Uniquely Rigid Bioinspired Supramolecular Structures. *Nano Lett.* **2005**, *5*, 1343-1346.
16. Yan, X.; Zhu, P.; Li, J. Self-Assembly and Application of Diphenylalanine-Based Nanostructures. *Chem. Soc. Rev.* **2010**, *39*, 1877-1890.
17. Frederix, P. W. J. M.; Patmanidis, I.; Marrink, S. J. Molecular Simulations of Self-Assembling Bio-Inspired Supramolecular Systems and Their Connection to Experiments. *Chem. Soc. Rev.* **2018**, *47*, 3470-3489.
18. Lampel, A.; Ulijn, R. V.; Tuttle, T. Guiding Principles for Peptide Nanotechnology through Directed Discovery. *Chem. Soc. Rev.* **2018**, *47*, 3737-3758.
19. Mason, T. O.; Michaels, T. C. T.; Levin, A.; Dobson, C. M.; Gazit, E.; Knowles, T. P. J.; Buell, A. K. Thermodynamics of Polypeptide Supramolecular Assembly in the Short-Chain Limit. *J. Am. Chem. Soc.* **2017**, *139*, 16134-16142.
20. Reddy, S. M. M.; Shanmugam, G. Role of Intramolecular Aromatic p-p Interactions in the Self-Assembly of Di-l-Phenylalanine Dipeptide Driven by Intermolecular Interactions: Effect of Alanine Substitution. *ChemPhysChem* **2016**, *17*, 2897 - 2907.
21. Mishra, A.; Chauhan, V. S. Probing the Role of Aromaticity in the Design of Dipeptide Based Nanostructures. *Nanoscale* **2011**, *3*, 945-949.

- 1
2
3 22. Brown, N.; Lei, J.; Zhan, C.; Shimon, L. J. W.; Adler-Abramovich, L.; Wei, G.; Gazit, E. Structural Polymorphism in a Self-Assembled Tri-Aromatic Peptide System. *ACS Nano* **2018**, *12*, 3253-3262.
- 4
5
6 23. Guo, C.; Luo, Y.; Zhou, R.; Wei, G. Triphenylalanine Peptides Self-Assemble into Nanospheres and Nanorods That Are Different from the Nanovesicles and Nanotubes Formed by Diphenylalanine Peptides. *Nanoscale* **2014**, *6*, 2800-2811.
- 7
8
9 24. Rissanou, A. N.; Georgilis, E.; Kasotakis, E.; Mitraki, A.; Harmandaris, V. Effect of Solvent on the Self-Assembly of Dialanine and Diphenylalanine Peptides. *J. Phys. Chem. B* **2013**, *117*, 3962-3975.
- 10
11
12 25. Jeon, J.; Mills, C. E.; Shell, M. S. Molecular Insights into Diphenylalanine Nanotube Assembly: All-Atom Simulations of Oligomerization. *J. Phys. Chem. B* **2013**, *117*, 3935-3943.
- 13
14
15 26. Guo, C.; Luo, Y.; Zhou, R.; Wei, G. Probing the Self-Assembly Mechanism of Diphenylalanine-Based Peptide Nanovesicles and Nanotubes. *ACS Nano* **2012**, *6*, 3907-3918.
- 16
17
18 27. Wang, J.; Liu, K.; Xing, R.; Yan, X. Peptide Self-Assembly: Thermodynamics and Kinetics. *Chem. Soc. Rev.* **2016**, *45*, 5589-5604.
- 19
20
21 28. Frederix, P. W. J. M.; Ulijn, R. V.; Hunt, N. T.; Tuttle, T. Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J. Phys. Chem. Lett.* **2011**, *2*, 2380-2384.
- 22
23
24 29. Frederix, P. W. J. M.; Scott, G. G.; Abul-Haija, Y. M.; Kalafatovic, D.; Pappas, C. G.; Javid, N.; Hunt, N. T.; Ulijn, R. V.; Tuttle, T. Exploring the Sequence Space for (Tri-)Peptide Self-Assembly to Design and Discover New Hydrogels. *Nat. Chem.* **2015**, *7*, 30-37.
- 25
26
27 30. Hamley, I. W. Peptide Nanotubes. *Angew. Chem., Int. Ed.* **2014**, *53*, 6866-6881.
- 28
29
30 31. Dinesh, B.; Squillaci, M. A.; Menard-Moyon, C.; Samori, P.; Bianco, A. Self-Assembly of Diphenylalanine Backbone Homologues and Their Combination with Functionalized Carbon Nanotubes. *Nanoscale* **2015**, *7*, 15873-15879.
- 31
32
33 32. Xing, R.; Yuan, C.; Li, S.; Song, J.; Li, J.; Yan, X. Charge-Induced Secondary Structure Transformation of Amyloid-Derived Dipeptide Assemblies from Beta-Sheet to Alpha-Helix. *Angew. Chem., Int. Ed.* **2018**, *57*, 1537-1542.
- 34
35
36 33. Zhu, P.; Yan, X.; Su, Y.; Yang, Y.; Li, J. Solvent-Induced Structural Transition of Self-Assembled Dipeptide: From Organogels to Microcrystals. *Chem. —Eur. J.* **2010**, *16*, 3176-3183.
- 37
38
39 34. Erdogan, H.; Babur, E.; Yilmaz, M.; Candas, E.; Gordesel, M.; Dede, Y.; Oren, E. E.; Demirel, G. B.; Ozturk, M. K.; Yavuz, M. S.; Demirel, G. Morphological Versatility in the Self-Assembly of Val-Ala and Ala-Val Dipeptides. *Langmuir* **2015**, *31*, 7337-7345.
- 40
41
42 35. Gupta, M.; Bagaria, A.; Mishra, A.; Mathur, P.; Basu, A.; Ramakumar, S.; Chauhan, V. S. Self-Assembly of a Dipeptide-Containing Conformationally Restricted Dehydrophenylalanine Residue to Form Ordered Nanotubes. *Adv. Mater.* **2007**, *19*, 858-861.
- 43
44
45 36. Amdursky, N.; Molotskii, M.; Gazit, E.; Rosenman, G. Elementary Building Blocks of Self-Assembled Peptide Nanotubes. *J. Am. Chem. Soc.* **2010**, *132*, 15632-15636.
- 46
47
48 37. Edison, J. R.; Spencer, R. K.; Butterfoss, G. L.; Hudson, B. C.; Hochbaum, A. I.; Paravastu, A. K.; Zuckermann, R. N.; Whitelam, S. Conformations of Peptoids in Nanosheets Result from the Interplay of Backbone Energetics and Intermolecular Interactions. *Proc. Nat. Acad. Sci. U.S.A.* **2018**, *115*, 5647-5651.
- 49
50
51 38. Johny, M.; Vijayalakshmi, K.; Das, A.; Roy, P.; Mishra, A.; Dasgupta, J. Modulating the Phe-Phe Dipeptide Aggregation Landscape via Covalent Attachment of an Azobenzene Photoswitch. *Chem. Commun.* **2017**, *53*, 9348-9351.
- 52
53
54 39. Ishikawa, M. S.; Busch, C.; Motzkus, M.; Martinho, H.; Backup, T. Two-Step Kinetic Model of the Self-Assembly Mechanism for Diphenylalanine Micro/Nanotube Formation. *Phys. Chem. Chem. Phys.* **2017**, *19*, 31647-31654.
- 55
56
57 40. Whitelam, S.; Jack, R. L. The Statistical Mechanics of Dynamic Pathways to Self-Assembly. *Annu. Rev. Phys. Chem.* **2015**, *66*, 143-163.
- 58
59
60 41. Nguyen, T. D.; Glotzer, S. C. Reconfigurable Assemblies of Shape-Changing Nanorods. *ACS Nano* **2010**, *4*, 2585-2594.
42. Whitelam, S.; Rogers, C.; Pasqua, A.; Paavola, C.; Trent, J.; Geissler, P. L. The Impact of Conformational Fluctuations on Self-Assembly: Cooperative Aggregation of Archaeal Chaperonin Proteins. *Nano Lett.* **2009**, *9*, 292-297.

- 1
2
3 43. Görbitz, C. H. Nanotube Formation by Hydrophobic Dipeptides. *Chem.—Eur. J.* **2001**, *7*,
4 5153-5159.
- 5 44. Görbitz, C. H. The Structure of Nanotubes Formed by Diphenylalanine, the Core Recognition
6 Motif of Alzheimer's β -amyloid Polypeptide. *Chem. Commun.* **2006**, *22*, 2332-2334.
- 7 45. Houton, K. A.; Morris, K. L.; Chen, L.; Schmidtman, M.; Jones, J. T. A.; Serpell, L. C.;
8 Lloyd, G. O.; Adams, D. J. On Crystal *versus* Fiber Formation in Dipeptide Hydrogelator Systems.
9 *Langmuir* **2012**, *28*, 9797-9806.
- 10 46. Adams, D. J.; Morris, K.; Chen, L.; Serpell, L. C.; Bacsá, J.; Day, G. M. The Delicate
11 Balance between Gelation and Crystallisation: Structural and Computational Investigations. *Soft*
12 *Matter* **2010**, *6*, 4144-4156.
- 13 47. Lampel, A.; McPhee, S. A.; Park, H.-A.; Scott, G. G.; Humagain, S.; Hekstra, D. R.; Yoo, B.;
14 Frederix, P. W. J. M.; Li, T.-D.; Abzalimov, R. R.; Greenbaum, S. G.; Tuttle, T.; Hu, C.; Bettinger, C.
15 J.; Ulijn, R. V. Polymeric Peptide Pigments with Sequence-Encoded Properties. *Science* **2017**, *356*,
16 1064-1068.
- 17 48. Tamamis, P.; Adler-Abramovich, L.; Reches, M.; Marshall, K.; Sikorski, P.; Serpell, L.; Gazit,
18 E.; Archontis, G. Self-Assembly of Phenylalanine Oligopeptides: Insights from Experiments and
19 Simulations. *Biophys J.* **2009**, *96*, 5020-5029.
- 20 49. Guo, C.; Arnon, Z. A.; Qi, R.; Zhang, Q.; Adler-Abramovich, L.; Gazit, E.; Wei, G.
21 Expanding the Nanoarchitectural Diversity Through Aromatic Di- and Tri-Peptide Coassembly:
22 Nanostructures and Molecular Mechanisms. *ACS Nano* **2016**, *10*, 8316-8324.
- 23 50. Mishra, A.; Chauhan, V. S. Probing the Role of Aromaticity in the Design of Dipeptide Based
24 Nanostructures. *Nanoscale* **2011**, *3*, 945-949.
- 25 51. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI
26 Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812-
27 7824.
- 28 52. Han W.; Wan C.; Jiang F.; Wu Y. PACE Force Field for Protein Simulations. 1. Full
29 Parameterization of Version 1 and Verification. *J. Chem. Theory Comput.* **2010**, *6*, 3373-3389.
- 30 53. Han, W.; Schulten, K. Further Optimization of a Hybrid United-Atom and Coarse-Grained
31 Force Field for Folding Simulations: Improved Backbone Hydration and Interactions between
32 Charged Side Chains. *J. Chem. Theory Comput.* **2012**, *8*, 4413-4424.
- 33 54. Cao, Y.; Jiang, X.; Han, W. Self-Assembly Pathways of Beta-Sheet-Rich Amyloid-Beta(1-40)
34 Dimers: Markov State Model Analysis on Millisecond Hybrid-Resolution Simulations. *J. Chem.*
35 *Theory Comput.* **2017**, *13*, 5731-5744.
- 36 55. Jiang, X.; Cao, Y.; Han, W. In Silico Study of Recognition between A β 40 and Abeta40 Fibril
37 Surfaces: An N-Terminal Helical Recognition Motif and Its Implications for Inhibitor Design. *ACS*
38 *Chem. Neurosci.* **2018**, *9*, 935-944.
- 39 56. Reches, M.; Gazit, E. Controlled Patterning of Aligned Self-Assembled Peptide Nanotubes.
40 *Nat. Nanotechnol.* **2006**, *1*, 195-200.
- 41 57. Hukushima, K.; Nemoto, K. Exchange Monte Carlo Method and Application to Spin Glass
42 Simulations. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604-1608.
- 43 58. Yan, X. H.; Cui, Y.; He, Q.; Wang, K. W.; Li, J. B.; Mu, W. H.; Wang, B.; Ou-yang, Z. C.
44 Reversible Transitions between Peptide Nanotubes and Vesicle-Like Structures including Theoretical
45 Modeling Studies. *Chem. Eur. J.* **2008**, *14*, 5974-5980.
- 46 59. Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2007**, *1767*,
47 1073-1101.
- 48 60. Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination
49 of Protein Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95-120.
- 50 61. Seo, J.; Hoffmann, W.; Warnke, S.; Huang, X.; Gewinner, S.; Schollkopf, W.; Bowers, M. T.;
51 von Helden, G.; Pagel, K. An Infrared Spectroscopy Approach to Follow Beta-Sheet Formation in
52 Peptide Amyloid Assemblies. *Nat. Chem.* **2017**, *9*, 39-44.
- 53 62. Pizzanelli, S.; Forte, C.; Monti, S.; Zandomenighi, G.; Hagarman, A.; Measey, T. J.;
54 Schweitzer-Stenner, R. Conformations of Phenylalanine in the Tripeptides AFA and GFG Probed by
55 Combining MD Simulations with NMR, FTIR, Polarized Raman, and VCD Spectroscopy. *J. Phys.*
56 *Chem. B* **2010**, *114*, 3965-3978.
- 57
58
59
60

- 1
2
3 63. Grdadolnik, J.; Mohacek-Grosev, V.; Baldwin, R. L.; Avbelj, F. Populations of the Three
4 Major Backbone Conformations in 19 Amino Acid Dipeptides. *Proc. Natl. Acad. Sci. U.S.A.* **2010**,
5 108, 1794–1798.
- 6 64. Jiang, F.; Han, W.; Wu, Y. D. The Intrinsic Conformational Features of Amino Acids from a
7 Protein Coil Library and Their Applications in Force Field Development. *Phys. Chem. Chem. Phys.*
8 **2013**, *15*, 3413-3428.
- 9 65. Fitzkee, N. C.; Fleming, P. J.; Rose, G. D. The Protein Coil Library: A Structural Database of
10 Nonhelix, Nonstrand Fragments Derived from the PDB. *Proteins: Struct., Funct., Genet.* **2005**, *58*,
11 852-854.
- 12 66. Perskie, L. L.; Street, T. O.; Rose, G. D. Structures, Basins, and Energies: A Deconstruction
13 of the Protein Coil Library. *Protein Sci.* **2008**, *17*, 1151-1161.
- 14 67. Yan, X.; Li, J.; Mohwald, H. Self-Assembly of Hexagonal Peptide Microtubes and Their
15 Optical Waveguiding. *Adv. Mater.* **2011**, *23*, 2796-2801.
- 16 68. Wang, J.; Liu, K.; Yan, L.; Wang, A.; Bai, S.; Yan, X. Trace Solvent As a Predominant
17 Factor To Tune Dipeptide Self-Assembly. *ACS Nano* **2016**, *10*, 2138-2143.
- 18 69. Lee E.; Kim J.; Lee M. Reversible Scrolling of Two-Dimensional Sheets from the
19 Self-Assembly of Laterally Grafted Amphiphilic Rods. *Angew. Chem., Int. Ed.* **2009**, *48*, 3657-3660.
- 20 70. Misra, P.; Kodali, R.; Chemuru, S.; Kar, K.; Wetzal, R. Rapid α -Oligomer Formation
21 Mediated by the Abeta C Terminus Initiates an Amyloid Assembly Pathway. *Nat. Commun.* **2016**, *7*,
22 12419.
- 23 71. Hauser, C. A.; Deng, R.; Mishra, A.; Loo, Y.; Khoe, U.; Zhuang, F.; Cheong, D. W.; Accardo,
24 A.; Sullivan, M. B.; Riek, C.; Ying, J. Y.; Hauser, U. A. Natural Tri- to Hexapeptides Self-
25 Assemble in Water to Amyloid Beta-Type Fiber Aggregates by Unexpected Alpha-Helical
26 Intermediate Structures. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1361-1366.
- 27 72. Song, B.; Yuan, H.; Pham, S. V.; Jameson, C. J.; Murad, S. Nanoparticle Permeation Induces
28 Water Penetration, Ion Transport, and Lipid Flip-Flop. *Langmuir* **2012**, *28*, 16989-17000.
- 29 73. Han W.; Wan C.; Wu Y. PACE Force Field for Protein Simulations. 2. Folding Simulations
30 of Peptides. *J. Chem. Theory Comput.* **2010**, *6*, 3390–3402.
- 31 74. MacKerell, A. D., Jr; Feig, M.; Brooks, C. L. 3rd. Extending the Treatment of Backbone
32 Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing
33 Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**,
34 *25*, 1400-1415.
- 35 75. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison
36 of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.
- 37 76. Ramos Sasselli, I.; Ulijn, R. V.; Tuttle, T. CHARMM Force Field Parameterization Protocol
38 for Self-Assembling Peptide Amphiphiles: The Fmoc Moiety. *Phys. Chem. Chem. Phys.* **2016**, *18*,
39 4659-4667.
- 40 77. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly
41 Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*,
42 435-447.
- 43 78. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth
44 Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- 45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60