De Novo Designed Peptide and Protein Hairpins Self-Assemble into Sheets and Nanoparticles

Johanna M. Galloway,* Harriet E. V. Bray, Deborah K. Shoemark, Lorna R. Hodgson, Jennifer Coombs, Judith M. Mantell, Ruth S. Rose, James F. Ross, Caroline Morris, Robert L. Harniman, Christopher W. Wood, Christopher Arthur, Paul Verkade, and Derek N. Woolfson

1. Introduction

Self-assembling proteins form the building blocks of life to control many, if not all, cellular processes. Natural self-assembling proteins include: 1D actin filaments for cell structure and motility,[1] 2D S-layers that form protective outer barriers in some bacteria,[2] 3D capsids that protect viral DNA/RNA,[3] and the shells of bacterial microcompartments, which are natural nanoreactors.[4] These natural scaffolds can be engineered to display other proteins to create assemblies for biotechnology.[5] For example, enzyme pathways fused to viral capsid proteins self-assemble to create nanoreactors.[6] Virus-like particles have also been engineered to display antigenic epitopes for use in vaccination,[7] or adorned with targeting motifs to direct them to diseased cells.[8] The toolbox of...

DOI: 10.1002/smll.202100472
useful self-assembling protein structures has been expanded by mutating natural protein interfaces to induce controlled self-assembly and through de novo design.\[9\] Such structures are usually computationally designed to form closely packed 2D arrays, tubes,\[10\] or 3D icosahedral particles.\[12\] However, these beautifully ordered and near crystalline assemblies may not be amenable to decoration with large cargos, or be permeable to small molecules due to their close packed nature. Also, these engineered or designed arrays can require thermal annealing to assemble,\[10d,e\] which may reduce or even destroy the activity of any appended proteins. Alternatively, 1D fiber assemblies can form extensive 3D gels,\[13\] but controlling the localization of appended motifs and/or functions is limited because network formation is a stochastic process. Between these two extremes of close-packed order and 3D entangled networks, there is space for the development of room-temperature self-assembling biomolecular systems that can display functional cargos and be permeable to small molecules.

Previously, we have used de novo \(\alpha\)-helical coiled-coil peptides to make self-assembled peptide cages (SAGEs).\[14\] In SAGEs a homotrimeric coiled-coil peptide (CC-Tri3)\[15\] is joined back-to-back with a disulfide bond to one of two halves of a heterodimeric coiled-coil pair (CC-DiA and CC-DiB).\[15\] This generates two complementary building-blocks or hubs, A and B.\[14\] When mixed, these self-assemble into a hexagonal lattice that curves,\[14\] and forms closed structures with the aid of defects.\[16\] SAGEs have been adapted for uptake by mammalian cells as potential drug delivery vehicles,\[17\] and decorated with immunogenic peptide sequences to make a modular vaccine platform.\[18\] The peptide cages are permeable to small molecules, so are ideal scaffolds for nanoreactors.\[19\] The above examples all employ synthetically derived peptides and hubs. Natural proteins, such as enzymes and whole protein antigens, are too large to incorporate into SAGEs in this way. However, the system can be adapted to make peptide-protein fusions for recombinant production, leading to functionalized pSAGEs.\[20\]

Others have pioneered approaches to construct large coiled-coil based nanoparticles. For example, Burkhard and co-workers have linked a de novo homotrimeric and modified natural homopentameric coiled-coil sequences with a short GG linker and an inter-helix disulfide bridge proximal to the loop which self-assemble into polyhedral nanoparticles.\[12d-e\] The group has been particularly successful at developing these as vaccine platforms for the presentation of antigenic peptides.\[12d-e\]

In a different concept, Marsh and co-workers combine de novo coiled-coil units and natural oligomeric proteins to render defined protein nanocages.\[21\]

Here we present self-assembling hairpin designs that form two distinct supramolecular assemblies: namely, closed nanoparticles and extended sheets. These designs combine ideas from the SAGEs and Burkhard’s nanoparticles. Specifically, homodimeric (CC-Di) and homotrimeric (CC-Tri3) blocks\[15\] are joined by a flexible linker, and pinned back-to-back with a disulfide bond (Figure 1). Computational modeling indicates that these should fold into stable hairpins that can be arrayed hexagonally. Moreover, and distinctively, the models suggest that the position of the disulfide may influence the supramolecular assembly profoundly: a pin proximal to the loop leads to curved arrays, which could close to form particles, whereas a distal pin restricts curvature, potentially leading to extended sheets. These extremes are confirmed experimentally using a variety of advanced microscopy methods for hairpins made by peptide synthesis and when produced recombinantly. The validated designs offer de novo scaffolds with potential to display a range of functionalities for application in imaging, cell targeting, nanoreactors, drug delivery, and modular vaccines.

2. Results and Discussion

2.1. Design of \(\alpha\)-Helical Hairpin Peptides

Our design strategy used a GSGSG sequence to link two self-assembling coiled-coil 3 heptad building blocks: homodimeric (CC-Di) and homotrimeric (CC-Tri3).\[15\] This gives two possible linear peptide sequences: CC-Di–link–CC-Tri3 and CC-Tri3– link–CC-Di, abbreviated to DT and TD respectively. These were pinned back-to-back with a disulfide bond between cysteines placed at one of two complementary positions either proximal (Cys15 & Cys35) or distal (Cys8 & Cys41) to the linking loop. In combination, this gives four possible constructs illustrated in Figure 1 and their sequences listed in Table 1. To assess any differences in the placement of the disulfide pins, we constructed all-atom models for 19-hexagon patches of assembled arrays for the two different pin positions and subjected these to molecular dynamics (MD) simulations in water (Figure 1c–f, Movies S1 and S2, Supporting Information). Intriguingly, the patch of the proximally pinned hairpins (pepHP–DTprox) curved in the 100 ns simulation. This resulted in the N and C termini of the peptides being presented on the convex face and the loops on the concave side. Projection of this curvature suggested that the arrays could close to form a nanoparticle 71 ± 11 nm in diameter, i.e., like the SAGEs (Figure S1a, Supporting Information). In contrast, simulations of arrays of distally pinned hairpins (pepHP–DTdist) the 19-hexagon patches remained relatively flat throughout the trajectories, with no preferred direction or magnitude of curvature (Figure S1b, Supporting Information). This suggested that these peptides might form flat extended sheets. The cysteines in the model assembly were replaced with glutamine, so that the hairpin building blocks could no longer form the disulfide pin. These modeled “C to Q” hairpin constructs did not remain flat, but instead flexed out of plane with no preferred direction of curvature, indicating that the assembly was unstable in the absence of the disulfide pins (Movie S3 and Figure S1c,d, Supporting Information).

These simulations gave the first indication of differences between the proximally and distally pinned systems. We rationalize this in terms of different flexibilities in the constructs. For the proximal pin, the helical N and C termini have more freedom to explore space and move apart, resulting in wedge-shaped building-blocks, narrow at the loop and wider at the termini. A distal pin reduces this freedom and constrains the assembly to a flatter topology.

2.2. Peptide Hairpins Assemble into Particles and Sheets as Predicted

All four peptide hairpin designs (Figure 1a and Table 1) were synthesized and purified by HPLC. The constructs were oxidized with iodine to form the disulfide pin, and subject to Ellman’s test\[22\] to confirm disulfide bond formation. The peptides were confirmed as monomers by nanospray ionization mass spectrometry
As judged by a change from clear to cloudy, all four peptides began to assemble as soon as they were hydrated in aqueous buffer. As a result, the samples scattered light and it was not possible to record reliable circular dichroism spectra to confirm α-helical assemblies. Instead, we visualized the assemblies by negative-stain transmission electron microscopy (TEM) and cryogenic transmission electron microscopy (cryo-TEM) (Figure 2). With either CC-Di or CC-Tri3 as the N-terminal block, the proximally pinned peptides pepHP–DTprox and pepHP–TDprox respectively, formed closed nanoparticle structures (Figure 2a–h). In TEM, the particles seemed to have a nonuniform size distribution. As the particle edges were difficult to discern in both the negatively stained and cryo-TEM images, we assessed particle size distribution by atomic force microscopy (AFM) (see below). All of the analogous peptides with the distal pin, pepHP–DTdist and pepHP–TDdist, formed sheet-like structures, extending for >100 nm in the xy dimension (Figure 2i–p). Thus, these initial experimental data support the MD simulations. Further, as there was no discernible difference in assemblies produced by the constructs with CC-Di or CC-Tri3 units placed first in the sequence, all further data presented had the CC-Di sequence at the N terminus, i.e., HP-DT constructs. (Figures S2 and S3, Supporting Information, sequences and mass data in Table S1, Supporting Information). As judged by a change from clear to cloudy, all four peptides began to assemble as soon as they were hydrated in aqueous buffer. As a result, the samples scattered light and it was not possible to record reliable circular dichroism spectra to confirm α-helical assemblies.

Instead, we visualized the assemblies by negative-stain transmission electron microscopy (TEM) and cryogenic transmission electron microscopy (cryo-TEM) (Figure 2). With either CC-Di or CC-Tri3 as the N-terminal block, the proximally pinned peptides pepHP–DTprox and pepHP–TDprox respectively, formed closed nanoparticle structures (Figure 2a–h). In TEM, the particles seemed to have a nonuniform size distribution. As the particle edges were difficult to discern in both the negatively stained and cryo-TEM images, we assessed particle size distribution by atomic force microscopy (AFM) (see below). All of the analogous peptides with the distal pin, pepHP–DTdist and pepHP–TDdist, formed sheet-like structures, extending for >100 nm in the xy dimension (Figure 2i–p). Thus, these initial experimental data support the MD simulations. Further, as there was no discernible difference in assemblies produced by the constructs with CC-Di or CC-Tri3 units placed first in the sequence, all further data presented had the CC-Di sequence at the N terminus, i.e., HP-DT constructs.

Table 1. Peptide sequences synthesized using solid-phase peptide synthesis. Peptide name and corresponding amino acid sequence aligned with numbering of amino acid position (1 to 51) and coiled-coil register (a to g). Sequences based on a basis set of coiled-coils,\textsuperscript{[15]} with the sequence based on the homodimer (CC-Di) shown in grey, and the sequence based on the homotrimer (CC-Tri3) shown in green. Cysteine mutations introduced at f positions in order to form a disulfide bonded hairpin are underlined. Linkers (C, GSGSG) and chromophore tags (GY) are shown in black.

| Name         | Sequence                                                                 |
|--------------|---------------------------------------------------------------------------|
| pepHP–DTprox | Ac–GEIAALK\textsubscript{C}K\textsubscript{N}AALE\textsubscript{E}IAALK\textsubscript{Q}GS\textsubscript{G}\textsubscript{E}IAAKKEI\textsubscript{A}A\textsubscript{K}CEIA\textsubscript{A}IKQGY–Am |
| pepHP–TDprox | Ac–GEIAIK\textsubscript{E}IAAIKKEIA\textsubscript{A}IALK\textsubscript{Q}GS\textsubscript{G}\textsubscript{E}IAALK\textsubscript{K}N\textsubscript{A}AE\textsubscript{C}IAALKQGY–Am |
| pepHP–DTdist | Ac–GEIAALK\textsubscript{C}K\textsubscript{N}AALE\textsubscript{E}IAALK\textsubscript{Q}GS\textsubscript{G}\textsubscript{E}IAAKKEI\textsubscript{A}A\textsubscript{K}CEIA\textsubscript{A}IKQGY–Am |
| pepHP–TDdist | Ac–GEIAIK\textsubscript{E}IAAIKKEIA\textsubscript{A}IALK\textsubscript{Q}GS\textsubscript{G}\textsubscript{E}IAALK\textsubscript{K}N\textsubscript{A}AE\textsubscript{C}IAALKQGY–Am |
Next, we used AFM to probe the nature and dimensions of the assemblies in more detail. Figure 3. First, at 1 h postassembly, the proximally pinned peptide pepHP–DTprox formed nanoparticles with an average maximum height of 27 ± 22 nm and average diameter 101 ± 58 nm (n = 5400) Figure 3a–c. The large ranges reflected a bimodal distribution of the aspect ratios, centered on ≈0.1 and ≈0.4. After 24 h, the particles had approximately the same diameter (104 ± 79 nm), but only the thicker particles were apparent, with a height average of 39 ± 38 nm and aspect ratio 0.37, Figure 3d–f. Though the distributions were broad, these data were consistent with the particle sizes seen in TEM and cryo-TEM, and estimated from the modeling (71 ± 11 nm), and showed that the particle diameters were stable over time. In addition, the experimental data indicated maturation of the particles between 1 and 24 h. We cannot be sure what this is due to, but posit that recruitment of peptides to the structures over time may result in organization into a multilamellar, stiffened particles, with the same diameter as seen in the earlier assembly.

AFM of the distally pinned peptide pepHP–DTdist revealed thin (5.6 ± 0.2 nm, Table S2, Supporting Information n = 5 ± standard deviation) sheets (Figure S2). This is similar to the expected height of one hairpin from termini to loop, which is ≈5 nm. This value, and the tight distribution of the experimental data support the hypothesis that these peptides form unilamellar sheets. It was not possible to discern the lattice clearly in AFM (Figure S5, Supporting Information), which could be due to tip resolution, or flexibility and thermal motions in the assembly. However, fast Fourier transform (FFT) analysis of further TEM and cryo-TEM images (Figure 4b,d) revealed regular structures with spacings of 4.8 nm. This corresponds closely to the expected inter-vertex (inter-trimer) distances of the hexagonal lattice, which span four helices each just over ≈1 nm in diameter (Figure 1b).

Figure 2. Negative stained TEM, FFTs, and cryo-TEM of peptide hairpin assemblies. Pin proximal to the loop: a–d) pepHP–TDprox and e–h) pepHP–DTprox formed closed structures. Pin distal to the loop: i–l) pepHP–TDdist and m–p) pepHP–DTdist formed flat structures. Negative stained samples at a,e) low, b,f,i,m) intermediate, and c,g,j,n) high magnification. Fourier transforms (FFT) from distally pinned hairpins, area highlighted in j) FFT shown k) from pepHP–DTdist and and area highlighted in n) FFT shown in o) for pepHP–TDdist. d,h,l, p) Cryo-TEM images (unannotated on the left, peptide structure edges highlighted in yellow dotted lines on the right to distinguish them from the lacey carbon support delineated in pink). Hairpin peptide aliquots (100 × 10–6 m) were assembled for 1 h in HBS (25 × 10–3 m HEPES, 25 × 10–3 m NaCl, pH 7.2).
2.3. Mixed Assemblies Can Be Made, but Preassembled Structures Do Not Exchange Peptide Components

To test for peptide mixing during assembly and for exchange postassembly, we made two fluorescently labeled variants of pepHP–DTdist and performed correlative light and electron microscopy (CLEM). We focused on this design as the sheets were much larger than the nanoparticles formed by the proximally pinned constructs, which made correlating the light and electron microscopy for individual sheets clearer than for the particles. The peptides had N-terminal 5(6)-carboxyfluorescein (green) and 5(6)-carboxytetramethylrhodamine (red) labels, Table S2 (Supporting Information). First, the labeled peptides were mixed when unfolded in 50% aqueous acetonitrile and

Figure 3. AFM grainsizing and height measurements of hairpin peptides. Proximally pinned pepHP–DTprox ($100 \times 10^{-6}$ M) were assembled in HBS (25 $\times$ $10^{-3}$ M HEPES, 25 $\times$ $10^{-3}$ M NaCl, pH 7.2) for a–c) 1 h and d–f) 24 h, then deposited onto mica. Images were recorded using peak force atomic force microscopy (PF-AFM, a,d). Particle height and diameter data were extracted using a particle analysis script (link available in Supporting Information) and plotted to show height versus diameter b,e) and diameter versus aspect ratio c,f) (areas colored in red were the highest populated, those in blue the lowest populated). In b,c) $n = 5400$ particles, in e,f) $n = 613$ particles. Distally pinned pepHP–DTdist was assembled as above for 1 h, deposited on mica and imaged using tapping mode AFM (TM-AFM, g). Height profiles were measured across a tear using Nanoscope analysis v1.5. (positions annotated on (h) and shown in (i)) where dotted lines indicate the area averaged to get the height difference between the mica substrate and the assembled peptide sheet, shown in Table S2 (Supporting Information), $n = 5$, ± standard deviation.
freeze dried. When hydrated, these samples formed mixed assemblies as judged by the coincidence of the fluorescence from the two labels and electron density in CLEM, Figure S6a–f (Supporting Information). This demonstrates that differently decorated hairpin peptides can be combined into the same assembly. Second, green- and red-labeled hairpins were hydrated for 1 h separately before mixing, incubated for 1 h further, and then prepared for microscopy. In this case, the CLEM revealed distinct regions of green and red fluorescence, indicating that once assembled, the structures were stable and did not exchange peptide modules. Thus, despite their flexible construction, the hairpin peptides form robust and stable sheet assemblies from solution.

2.4. Hairpin Redesign for Protein Expression

Next, we sought to add functional proteins to the hairpin constructs. For this, we turned to the expression of synthetic genes in *Escherichia coli*. We designed genes for two constructs, pepHP–DTdist and proHP–DTdist, that harbored a 5’ multiple cloning site and a His-tag at the 3’ end (Figure S7, Tables S1 and S5, Supporting Information); the prescript “pro” refers to the recombinantly expressed protein constructs. These were overexpressed in SHuffle T7 cells, which are engineered to allow disulfide bond formation in their cytoplasm, then purified and characterized by SDS-PAGE and confirmed as monomers by nanospray ionization mass spectrometry (Figure S8a–f, Supporting Information). Ellman’s test confirmed that the molecules contained disulfide bonds. CLEM imaging of cell sections immunolabelled for the His-tag revealed protein dispersed within the cells rather than forming inclusion bodies (Figure S9, Supporting Information).

Whereas the proximally pinned synthetic peptide hairpins formed spherical structures (Figure 2a–h), the recombinant variant, proHP–DTdist formed aggregates (Figure S10a, Supporting Information pH 7.2). However, like the peptide variant (Figure 4a–f), the distally pinned protein hairpin, proHP–DTdist formed sheets (Figure 4g–l). Interestingly, these were smaller and had a clear ultrastructure, in the form of stripes. These stripes were apparent in both negative stain TEM (Figure 4e,k) and cryo-TEM (Figure 4f,l), and thus they are not a drying...
The loop on the convex side in unprotonated proHP–DTdist rather opposite to that seen for the proximally pinned peptide, with curved after 100 ns simulation (Figure 5b). The curvature was S5, Supporting Information). When unprotonated, the patch protonated His-tags as described above (Figure 5 and Movies S4 and S5, Supporting Information). When unprotonated, the patch curved after 100 ns simulation (Figure 5b). The curvature was opposite to that seen for the proximally pinned peptide, with the loop on the convex side in unprotonated proHP–DTdist rather than on the concave side as was seen for pepHP–DTprox. The protonated constructs maintained a flat trajectory throughout the 100 ns simulation (Figure 5d), as was seen for the distally pinned peptide pepHP–DTdist.

The assembly behavior of the protein hairpins was tested by hydration in HBS at pH values between 6.0 and 8.0 (Figure S10, Supporting Information). The distally pinned protein formed extensive sheet like structures at low pH (pH 6.5), and 3D aggregates at higher pH (pH 8.0). Stripes appeared in the images of the proHP–DTprox structures, but at slightly lower pH (6.5) when compared to the proHP–DTdist (7.2). Aggregates of proHP–DTprox were formed at low and high pH, indicating that the proximally pinned protein was not able to form stable structures. Modeling indicated that deprotonation of the His-tag induced the opposite sense of curvature to that caused by the proximal pin position. Thus proHP–DTdist may be too flexible, and the opposite sense of curvature induced by protonation versus pin position destabilized the assembly leading to an aggregation-based failure mechanism. The striped assemblies may represent a transition between the low pH “peptide-like” structure and the high pH “unprotonated curved” structures, probably induced by partial protonation of the His-tag. As its assembly was more stable, only proHP–DTprox was investigated further.

2.5. pH Alters Protein Hairpin Assembly Structure

We constructed all-atom models for 19 hexagon patches of assembled arrays for the distally pinned and His-tagged protein pepHP–DTdist. The pKa of the histidine side chain is near physiological pH. MD simulations were run for the distally pinned protein hairpin proHP–DTdist with unprotonated and protonated His-tags as described above (Figure 5 and Movies S4 and S5, Supporting Information). When unprotonated, the patch curved after 100 ns simulation (Figure 5b). The curvature was opposite to that seen for the proximally pinned peptide, with the loop on the convex side in unprotonated proHP–DTdist rather than on the concave side as was seen for pepHP–DTprox. The protonated constructs maintained a flat trajectory throughout the 100 ns simulation (Figure 5d), as was seen for the distally pinned peptide pepHP–DTdist.

The assembly behavior of the protein hairpins was tested by hydration in HBS at pH values between 6.0 and 8.0 (Figure S10, Supporting Information). The distally pinned protein formed extensive sheet like structures at low pH (<pH 6.5), and 3D aggregates at higher pH (pH 8.0). Stripes appeared in the images of the proHP–DTprox structures, but at slightly lower pH (6.5) when compared to the proHP–DTdist (7.2). Aggregates of proHP–DTprox were formed at low and high pH, indicating that the proximally pinned protein was not able to form stable structures. Modeling indicated that deprotonation of the His-tag induced the opposite sense of curvature to that caused by the proximal pin position. Thus proHP–DTdist may be too flexible, and the opposite sense of curvature induced by protonation versus pin position destabilized the assembly leading to an aggregation-based failure mechanism. The striped assemblies may represent a transition between the low pH “peptide-like” structure and the high pH “unprotonated curved” structures, probably induced by partial protonation of the His-tag. As its assembly was more stable, only proHP–DTprox was investigated further.

2.6. Successful Self-Assembly of Hairpins When Fused to Large Cargos

Fusions of the hairpin to the fluorescent protein sfGFP (~28 kDa) (Figure S7 and Tables S1 and S5, Supporting Information) were expressed in SHuffle T7 cells and purified as described above (Figure S8, Supporting Information). The fluorescence in CLEM images of cell sections indicates that the sfGFP cargo folds in the cytoplasm (Figure S11, Supporting Information). Again, we were not able to see any assembled structures within the cytoplasm. If the hairpins were able to self-assemble in cells, it may be that their open network structure is not electron dense enough to resolve in thin-section in TEM, so these constructs were also characterized in vitro. Ellman’s test confirmed that the molecules contained disulfide bonds after purification. Nanospray mass spectrometry and SDS-PAGE indicated that proGFP–HP–DTdist was monomeric (Figure S8g–i, Supporting Information). When hydrated in HBS, we observed flat sheets formed by the cargo laden hairpins (Figure S12, Supporting Information). This demonstrates that distally pinned hairpins are amenable to being decorated with large proteins.

3. Conclusion and Future Outlook

In summary, we have described the design, modeling, synthesis, assembly and characterization of two types of de novo designed peptide hairpins. In these, sequences for coiled-coil dimers and trimers are joined through a flexible linker. The hairpins are then pinned, using disulfide bridges either proximal to or distal from the linker. The aim being to expose the hydrophobic seams of the coiled-coil segments and promote assembly of the hairpins into hexagonal arrays. The two pin positions have profoundly different effects on the topology of the supramolecular assemblies formed. The proximal pins lead to closed, spherical objects on the order of ~100 nm in diameter, whereas the distal pin results in sheet-like assemblies consistent with a monolayer of folded, self-assembled hairpins. These structures are observed in a range of advanced microscopies, and the proposed mechanism of formation are supported by extended molecular dynamics simulations. Specifically, the proximal pin allows spaying of the helical termini, which in turn leads to curved arrays that can close, whereas the distal pins give a more tightly constrained hairpin structure, consistent with building blocks required to make flat sheets.

As noted above, others have developed similar self-assembling peptide- and protein-based nanoparticles or sheets. So what are the differences and advantages to our hairpin system? First, by including two points (loop and pin) that define the relative positions and rotational freedom of the two coiled-coil components, we are able to control the topology of the self-assembled structures by design in a single system to render closed nanoparticles or extended sheets. This level of dual control is encouraging for future biomaterials design using relatively short linear peptides. Second, although this is true for some other systems, the relatively short lengths of our hairpin designs and the minimal postsynthesis processing allows them to be made both synthetically and recombinantly. This has allowed us to decorate assemblies chemically, and through fusions to functional natural proteins; and to explore additional properties of the system. For example, experiments with fluorophore-labeled synthetic hairpin peptides showed that coassemblies can be made when starting with mixtures
of unfolded peptide variants. However, once formed, there is no interchange between assembled structures. This approach could be used to add other functional moieties, e.g., for drug delivery and targeting diseased cells. Extending this to recombinantly produced hairpin–protein fusions, which express and purify well, could allow the production of nanoparticles or sheets with combinations of small molecule and protein cargoes. Third, the fabric of the system is potentially modular, and we envisage that other de novo coiled-coil units could be swapped in to change the homomeric components used here.\(^{[27]}\) For instance, as we have demonstrated for the foregoing SAGE system,\(^{[14]}\) the homodimer sequence on the hairpins could be changed to a heterodimer pair.\(^{[27,28]}\) This would afford an additional level of control on the assembly.

In all of these respects, the linear peptide hairpins that we describe provide de novo modules to add to the growing toolkit of components for the rational construction of biomaterials.\(^{[9,12c,29]}\) That said, how subtle differences in peptide design translate to relatively small changes in module structure and then to completely different assembly topologies are both alarming and encouraging. It is of concern because it highlights how careful the design process must be in order to arrive at a targeted

---

**Figure 5.** Snapshots from all atom MD simulations of 19 hexagon protein hairpin patches. Distally pinned protein hairpins (\(\mu_{\text{HP–DT}_{\text{dist}}}\)) with a,b) unprotonated (Movie S4, Supporting Information) and c,d) protonated (Movie S5, Supporting Information) His-tags. Snapshots taken (a,c) before (blue, Movie S4, Supporting Information) and b,d) after (magenta, Movie S5, Supporting Information) 100 nm of MD simulation under periodic boundary conditions using explicit TIP3P water, pH 7.0, 298 K, 0.15 m NaCl.
design, and that designs must be tested when appending functionality to any self-assembling building-block. However, it is also exciting, as it opens considerable possibilities for designing a wider range of biomaterial structures and functions. Whatever your stance, the rational and computational de novo design of biomaterials remains both challenging and full of potential.

4. Experimental Section

A detailed description of the materials and methods is included in the Supporting Information.

MD Simulation: The coiled-coil helices were arranged in a 19 hexagon patch. [24] The disulfide pins were constructed by hand and aligned, and a His-tag added if necessary, in InsightII (Accelrys), and minor discrepancies in disulfide bond lengths fixed to 2.0 Å. This was parameterized with the Amber-99SBildn force field, before undergoing 5000 steps of energy minimization. The GROMACS-4.6.7 suite was used to perform simulations as NPT ensembles at 298 K using periodic boundary conditions. The simulations were integrated with a leap-frog algorithm over a 2 fs time-step, constraining bond vibrations with the P-LINCs method. Structures were saved every 0.1 ns for analysis and run over 100 ns. Molecular graphics manipulations and visualizations were performed using InsightII, VMD-1.9.1 and Chimera-1.10.2 and movies made using VMD.

Peptide Synthesis and Purification: A low-loading rink amide resin support was used in a Liberty microwave peptide synthesizer for solid phase peptide synthesis. [25] Fmoc-protected amino acids (0.1 × 10−3 m) were coupled (4.5 eq chloro-1-hydroxybenzotriazole (Cl-HOBt), 5 eq. N,N′-disopropylcarbodiimide (DIC) in 7 mL dimethylformamide (DMF) at 25 °C for 2 min, then at 25 W microwave irradiation at 50 °C for 5 min), washed (3 × 7 mL DMF), deprotected (7 mL 20% (v/v) morpholine in DMF at 20 W and 75 °C for 5 min), and washed (3 × 7 mL DMF) before adding the next Fmoc protected amino acid.

Peptides were capped using 3 eq. acetic anhydride and 4.5 eq. N,N-diisopropylethylamine (DIPEA) in 10 mL DMF for 30 min at room temperature (see Supporting Information for details on fluorescent labeling). The peptides were washed (3 × 5 mL DMF, 3 × 5 mL dichloromethane (DCM)), dried, and cleaved (9.5 mL trifluoroacetic acid (TFA), 250 µL water, 250 µL triisopropylsilane (TIPS), 3 h). The resin was removed by filtration and excess TFA evaporated. The crude peptide was precipitated (30 mL diethyl ether (Et2O), 4 × 5 mL water) before TEM. After centrifugation (10 000 × C, 10 min) the supernatant was applied to an immobilized phospho-anti-phospho-antibody (BAS) in phosphate buffered saline (PBS, 140 × 10−3 m NaCl, 2.7 × 10−3 m KCl, 10.1 × 10−3 m Na2HPO4, 1.8 × 10−3 m KH2PO4, pH 7.4, 2 × 5 min). Blocked sections were labeled with a primary antibody (1:500 6×His-tag monoclonal mouse IgG2b antibody in 1% BSA in PBS, 1 h), washed (3 × 5 min in 1% BSA in PBS), incubated with a fluorescent secondary antibody (DyLight 488 labelled secondary polyclonal goat anti mouse IgG antibody, 1 µg mL−1 in 1% BSA in PBS, 1 h), then washed in 1% BSA in PBS followed by water before drying. Cryo-TEM imaging (described above) was done before TEM.

Cryo-TEM: 5 µL in vitro assembled sample was blotted and plunged frozen onto Lacey carbon grids using a Leica EM CP GA automatic plunge freezer and stored in liquid N2. Frozen samples were loaded into a Gatan CryoTransfer specimen holder and imaged using a Tecnai 20 FEG 200 kV Twin Lens scanning transmission electron microscope (LaB6 filament, 200 keV) on a FEI Eagle 4k 4k CCD camera. Higher resolution images were recorded on an FEI Talos Arctica equipped with a 200 kV X-FEG and a Gatan CFI Quantum LS energy filter on a Ceta 16 M CCD detector or Gatan K2 DED. In addition to ImageJ, cryo-TEM images were processed using the Gatan Microscopy Suite (GMS) 3 software.

Protein Expression and Purification: Sequences were optimized for production in E. coli (Table S5, Supporting Information), ordered from GeneART, and transferred to pET3a (lactose control) and TBAD (arbinosine control) expression vectors (Figure S7, Supporting Information). These were transformed into SHuffle T7 Express competent cells and expressed by autoinduction [26] in Luria-Bertani (LB) medium (10 g L−1 NaCl, 10 g L−1 tryptone, 5 g L−1 yeast extract) 400 mL plus 8 mL 50X auto induction solution (for pET3a 100 mL contains 25 mL glycerol, 2.5 g D-glucose and 10 g galactose; for TBAD 100 mL contains 25 mL glycerol, 5 g D-glucose plus 5.0 g L-arabinose) at 30 °C and 200 rpm for 24 h. Cells were pelleted (10 000 × g, 4 °C, 10 min) and frozen at -20 °C.

Purified samples were analyzed by SDS-PAGE before being analyzed by mass spectrometry. [27]
and freeze dried. Smaller proteins (proHP–DTxxx) could also be purified by extracted the maximum height (above a 5 nm cutoff) and diameter (above wells-wood-research/galloway-jg-hairpin-self-assembly-2020. This script and diameters extracted using a script available here https://github.com/MaxEnt1 to 1 Da accuracy. AFM height data were flattened using the first computation modeled patches by extrapolating the curvature of the patch and assuming this would form a closed sphere. An average size was calculated from the recorded nanospray spectra by fitting protein was calculated for protein concentration, and eluted proteins were collected in 2 mL fractions. Fractions were analyzed for purity by sodium dodecyl sulfide polyacrylamide gel electrophoresis (SDS-PAGE, Figure 5B, Supporting Information), buffer exchanged (50 × 10−3 m ammonium bicarbonate) and freeze dried. Smaller proteins (proHP–DTxxx) could also be purified by RP-HPLC; and protein purity, mass and oligomeric state was confirmed by nanospray mass spectrometry as described for peptides above.

Statistical Analysis: Particle sizes were determined from the computationally modeled patches by extrapolating the curvature of the patch and assuming this would form a closed sphere. An average size was calculated by measuring the extrapolated diameter at 19 points equally spaced in time between 5 and 50 nm, and ± one standard deviation of these diameters is the quoted error. The mass of the purified peptide or protein was calculated from the recorded nanospray spectra by fitting using the Micromass Masslynx software maximum entropy analysis MaxEnt1 to 1 Da accuracy. AFM height data were flattened using the first order flattening tool in the Nanoscope software (v1.5), and particle height and diameters extracted using a script available here https://github.com/wells-wood-research/galloway-jg-hairpin-self-assembly-2020. This script extracted the maximum height (above a 5 nm cutoff) and diameter (above a 50 nm cutoff) of each particle in an AFM topography plot.

Supporting Information
Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements
This research was supported by funding from a LoLa (Strategic Longer and Larger) grant from the BBSRC “Development of supramolecular assemblies for enhancing cellular productivity and the synthesis of chemicals and biotherapeutics,” (BB/M002969/1). Dan Mulvihill and Martin Warren, who are based at the University of Kent, and Richard Pickersgill based at Queen Mary’s University of London, contributed to discussion and development of this work as part of this consortium. The authors also thank Evelyn Deery and Maria Stanley for supplying the pET3a and TBAD plasmids, (University of Kent) and Richard Sessions for support with the computational modeling (University of Bristol). The authors thank the EPSRC for awarding HECBiosym and an Archer Leadership Award for providing computer time on the U.K. Supercomputer Archer. The authors thank the EPSRC/EPSRC funded Synthetic Biology Research Centre (BrisSynBio, BB/L01386X/1) for providing funding for researchers and Bluegem. The authors thank the EPSRC for funding the Chemical Imaging Facility (PF-AFM, EP/ K035746/1), the mass spectrometry equipment (EP/K03927X/1), and the Wolfson Bioimaging Facility (fluorescence and electron microscopy, BB/ L014181/1) at the University of Bristol.

Conflict of Interest
The authors declare no conflict of interest.

Data Availability Statement
The data that support the findings of this study are openly available from the Leeds University Data Repository at https://doi.org/10.5518/955, reference [35].

Keywords
coil, coiled coil, computational modeling, peptide design, protein design, self-assembly

Received: January 25, 2021
Published online: February 15, 2021

[1] T. D. Pollard, J. A. Cooper, Science 2009, 326, 1208.
[2] R. P. Fagan, N. F. Fairweather, Nat. Rev. Microbiol. 2014, 12, 211.
[3] J. E. Johnson, W. Chiu, Curr. Opin. Struct. Biol. 2000, 10, 229.
[4] C. A. Kerfeld, C. Aussignargues, J. Zarzyski, F. Cai, M. Sutter, Nat. Rev. Microbiol. 2018, 16, 277.
[5] a) W. M. Aumiller, M. Uchida, T. Douglas, Chem. Soc. Rev. 2018, 47, 3433; b) A. M. Wen, N. F. Steinmetz, Chem. Soc. Rev. 2016, 45, 4074.
[6] D. P. Patterson, B. Schwarz, R. S. Waters, T. Gedeon, T. Douglas, ACS Chem. Biol. 2014, 9, 359.
[7] B. Schwarz, K. M. Morabito, T. J. Ruckwardt, D. P. Patterson, J. Avera, H. M. Miettinen, B. S. Graham, T. Douglas, ACS Biomater. Sci. Eng. 2016, 2, 2324.
[8] M. J. Rohovoi, M. Nagasawa, J. R. Swartz, Biogeoch. Transl. Med. 2017, 2, 43.
[9] J. L. Beesley, D. N. Woolfson, Curr. Opin. Biotechnol. 2019, 58, 175.
[10] a) H. Pyles, S. Zhang, J. J. D. Yoreo, D. Baker, Nature 2018, 571, 251; b) A. Thomas, J. F. Matteahe, F. Baneyx, Biotechnol. J. 2018, 13, ei800141; c) S. Gonen, F. DiMaio, T. Gonen, D. Baker, Science 2015, 348, 1365; d) H. V. Zhang, F. Polzer, M. J. Haider, Y. Tian, J. A. Villegas, K. L. Kiick, D. J. Pochan, J. G. Saven, Sci. Adv. 2016, 2, e1600307; e) A. D. Merg, G. Toupone, E. V. Genderen, X. Zuo, A. Bazrafshan, T. Blum, S. Hughes, K. Salata, J. P. Abravamas, V. P. Conticello, Angew. Chem., Int. Ed. 2019, 58, 13507; f) J. F. Matteahe, F. DiMaio, J. J. Richards, L. D. Pozzo, D. Baker, F. Baneyx, Nano Lett. 2015, 15, 5235; g) Z. Chen, C. M. Johnson, J. Chen, M. J. Bick, S. E. Boyken, B. Lin, J. J. D. Yoreo, J. M. Kollman, D. Baker, F. DiMaio, J. Am. Chem. Soc. 2019, 141, 8891; h) A. J. Ben-Sasson, J. L. Watson, W. Sheffler, M. C. Johnson, A. Bittleston, L. Somasundaram, J. Decarreau, F. Jiao, J. Chen, I. Mela, A. A. Drabek, S. M. Jarrett, S. C. Blacklow, C. F. Kamiński, G. L. Hura, J. J. De Yoreo, J. M. Kollman, H. Ruohola-Baker, E. Derivery, D. Baker, Nature 2021, 589, 648.
[11] a) J. F. Ross, G. C. Wildsmith, M. Johnson, D. L. Hurdiss, K. Hollingsworth, R. F. Thompson, M. Mosayebi, C. H. Trinh, E. Paci, A. R. Pearson, M. E. Webb, W. B. Turnbull, J. Am. Chem. Soc. 2019, 141, 5211; b) I. Uddin, S. Frank, M. J. Warren, R. W. Pickersgill, Small 2018, 14, 1704020; c) A. Pang, S. Frank, I. Brown, M. J. Warren, R. W. Pickersgill, J. Biol. Chem. 2014, 289, 22377.
[12] a) J. J. B. Bala, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, D. Baker, Science 2016, 353, 389; b) N. P. King, J. J. Bala, W. Sheffler, D. E. McNamara, S. Gonen, T. Gonen, T. O. Yeates, D. Baker, Nature 2014, 510, 103; c) P.-S. Huang, S. E. Boyken, D. Baker, Nature 2014, 537, 320; d) S. Raman, G. Machaidze, A. Lustig, U. Aebi, P. Burkhard, Nanomed.: Nanotechnol. Biol. Med. 2006, 2, 95; e) R. Dey, Y. Xia, M.-P. Nieh, P. Burkhard, ACS Biomater. Sci. Eng. 2017, 3, 724; f) C. P. Karch, P. Burkhard, Biochem. Pharmacol. 2016, 120, 1; g) T. A. P. F. Pimentel, Z. Yan, S. A. Jeffers, K. V. Holmes, R. S. Hodges, P. Burkhard, Chem. Biol. Drug Des. 2009, 73, 53; h) C. P. Karch, P. Burkhard, G. R. Matyas, Z. Beck, J. Vis. 2019, e60103; i) Q. Guo, D. Dasgupta, T. A. P. F. Doll, P. Burkhard, D. E. Laran, Methods 2013, 60, 242.
[13] a) N. Mehrban, B. Zhu, F. Tamagnini, F. I. Young, A. Wasmuth, K. L. Hudson, A. R. Thomson, M. A. Birchall, A. D. Randall,
B. Song, D. N. Woolfson, ACS Biomater. Sci. Eng. 2015, 1, 431;
b) E. F. Banwell, E. S. Abelardo, D. J. Adams, M. A. Birchall, A. Corrigan, A. M. Donald, M. Kirkland, L. C. Serpell, M. F. Butler, D. N. Woolfson, Nat. Mater. 2009, 8, 596; c) P. Worthington, S. Langhans, D. Pochan, Adv. Drug Delivery Rev. 2017, 110-111, 127;
d) K. Nagy-Smith, Y. Yamada, J. P. Schneider, J. Mater. Chem. B 2016, 4, 1999; e) R. V. Rughani, J. P. Schneider, MRS Bull. 2008, 33, 530.

[14] J. M. Fletcher, R. L. Harniman, F. R. H. Barnes, A. L. Boyle, A. Collins, J. Mantell, T. H. Sharp, M. Antognozzi, P. J. Booth, N. Linden, M. J. Miles, R. B. Sessions, P. Verkade, D. N. Woolfson, Science 2013, 340, 595.

[15] J. M. Fletcher, A. L. Boyle, M. Bruning, G. J. Bartlett, T. L. Vincent, N. R. Zaccai, C. T. Armstrong, E. H. C. Bromley, P. J. Booth, R. L. Brady, A. R. Thomson, D. N. Woolfson, ACS Synth. Biol. 2012, 1, 240.

[16] a) J. M. Galloway, L. Senior, J. M. Fletcher, J. L. Beesley, L. R. Hodgson, R. L. Harniman, J. M. Mantell, J. Coombs, G. G. Rhys, W. F. Xue, M. Mosayebi, N. Linden, T. B. Liverpool, P. Curnow, P. Verkade, D. N. Woolfson, ACS Nano 2018, 12, 1420; b) M. Mosayebi, D. K. Shoemark, J. M. Fletcher, R. B. Sessions, N. Linden, D. N. Woolfson, T. B. Liverpool, Proc. Natl. Acad. Sci. USA 2017, 114, 9014.

[17] J. L. Beesley, H. E. Baum, L. R. Hodgson, P. Verkade, G. S. Banting, D. N. Woolfson, Nano Lett. 2018, 18, 5933.

[18] C. Morris, S. J. Glennie, H. S. Lam, H. E. Baum, D. Kandage, N. A. Williams, D. J. Morgan, D. N. Woolfson, A. D. Davidson, Adv. Funct. Mater. 2019, 29, 1807357.

[19] D. K. Shoemark, A. Avila Ibarra, J. F. Ross, J. L. Beesley, H. E. V. Bray, M. Mosayebi, N. Linden, T. B. Liverpool, S. N. McIntosh-Smith, D. N. Woolfson, R. B. Sessions, Phys. Chem. Chem. Phys. 2019, 21, 137.

[20] J. F. Ross, A. Bridges, J. M. Fletcher, D. Shoemark, D. Alibhai, H. E. V. Bray, J. L. Beesley, W. M. Dawson, L. R. Hodgson, J. Mantell, P. Verkade, C. M. Edge, R. B. Sessions, D. Tew, D. N. Woolfson, ACS Nano 2017, 11, 7901.

[21] a) A. Sciore, M. Su, P. Koldewey, J. D. Eschweiler, K. A. Difflay, B. M. Linhares, B. T. Ruotolo, J. C. A. Bardwell, G. Skiniotis, E. N. G. Marsh, Proc. Natl. Acad. Sci. USA 2016, 113, 8681; b) D. P. Patterson, A. M. Desai, M. M. B. Holl, E. N. G. Marsh, RSC Adv. 2011, 1, 1004; c) D. P. Patterson, M. Su, T. M. Franzmann, A. Sciore, G. Skiniotis, E. N. G. Marsh, Protein Sci. 2014, 23, 190; d) A. S. Cristie-David, J. Chen, D. B. Nowak, A. L. Bondy, K. Sun, S. I. Park, M. M. Banaszak Holl, M. Su, E. N. G. Marsh, J. Am. Chem. Soc. 2019, 141, 9207.

[22] G. L. Ellman, Arch. Biochem. Biophys. 1959, 82, 70.

[23] J. Porath, Protein Expression Purif. 1992, 3, 263.

[24] J. Wu, M. Filutowicz, Acta Biochim. Pol. 1999, 46, 591.

[25] D. B. Williams, C. B. Carter, in Transmission Electron Microscopy: A Textbook for Materials Science, Springer, New York 2009, p. 389.

[26] A. L. Hansen, L. E. Kay, Proc. Natl. Acad. Sci. USA 2014, 111, E1705.

[27] F. Thomas, A. L. Boyle, A. J. Burton, D. N. Woolfson, J. Am. Chem. Soc. 2013, 135, 5161.

[28] J. Jin, E. G. Baker, C. W. Wood, J. Bath, D. N. Woolfson, A. J. Turberfield, ACS Nano 2019, 13, 9927.

[29] a) D. N. Woolfson, Z. N. Mahmoud, Chem. Soc. Rev. 2010, 39, 3464; b) Q. Luo, C. Hou, Y. Bai, R. Wang, J. Liu, Chem. Rev. 2016, 116, 13571; c) W. Zhang, S. Mo, M. Liu, L. Liu, L. Yu, C. Wang, Front. Chem. 2020, 8, 587975.

[30] G. B. Fields, R. L. Noble, Int. J. Pept. Protein Res. 1990, 35, 161.

[31] D. Andreu, F. Albericio, N. A. Solé, M. C. Munson, M. Ferrer, G. Barany, in Methods in Molecular Biology: Peptide Synthesis Protocols, Vol. 35 (Eds: M. W. Pennington, B. M. Dunn), Humana Press Inc., Totowa, NJ 1994, p. 91.

[32] a) M. D. Abramoff, P. J. Magalhaes, S. J. Ram, Biophotonics Int. 2004, 11, 36; b) J. Schindelin, C. T. Rueden, M. C. Hiner, K. W. Eliceiri, Mol. Reprod. Dev. 2015, 82, 518.

[33] M. J. Lee, J. Mantell, L. Hodgson, D. Alibhai, J. M. Fletcher, I. R. Brown, S. Frank, W.-F. Xue, P. Verkade, D. N. Woolfson, M. J. Warren, Nat. Chem. Biol. 2018, 14, 142.

[34] F. W. Sturdivant, Protein Expression Purif. 2005, 41, 207.

[35] J. M. Galloway, H. E. V. Bray, D. K. Shoemark, L. R. Hodgson, J. Coombs, J. M. Mantell, R. S. Rose, J. F. Ross, C. Morris, R. L. Harniman, C. W. Wood, C. Arthur, P. Verkade, D. N. Woolfson, Dataset 2021, https://doi.org/10.5518/955.