Supramolecular Self-assembly of Engineered Polyproline Helices

Dominic F Brightwell, Giada Truccolo, Kushal Samanta, Helena J Shepherd and Aniello Palma*

a) School of Chemistry and Forensic Science, Supramolecular Interfacial and Synthetic Chemistry, Ingram building, The University of Kent, Canterbury CT2 7NZ, Kent, United Kingdom

b) School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland

Emails: a.palma@kent.ac.uk; aniello.palma@ucd.ie

Homepage: https://www.palmaresearchgroup.com/

Abstract

The ability to rationally design biomaterials to form desired supramolecular constructs presents an ever-growing research field, with many burgeoning works within recent years providing exciting results, however, there exists a broad expanse of promising avenues of research yet to be investigated. As such we have set out to make use of the polyproline helix as a rigid, tuneable, and chiral ligand for the design and synthesis of supramolecular constructs. In this investigation we show how an oligoproline tetramer can be specifically designed and functionalised, allowing predictable tuning of supramolecular interactions to engineer the formation of supramolecular peptide frameworks with varying properties. Consequently, laying the groundwork for further studies utilising the polyproline helix, with the ability to design desired supramolecular structures, having tuneable structural features and functionalities.

Introduction

The synthesis of hierarchical supramolecular functional materials is an exciting field of research with applications in biomedicine,1,2 separation and catalysis,3–5 and sensing.6 Pivotal to the successful design of these supramolecular constructs is the ability to synthesise building blocks with specific topologies. Pioneering work by Raymond and Stang in the field of extended and discrete metal organic frameworks were instrumental in demonstrating the importance not only of the nature of the chemical handles but also of their relative position in space.7,8 While a high level of positional control
of these handles can be achieved with relative ease on classical (poly)aromatic building blocks, the same cannot be stated if one intends to use structured peptides as supramolecular building blocks. Recently, biomolecules such as peptides, lipids and DNA/RNA, have appeared in a number of reports as interesting building blocks in the synthesis of novel 2D and 3D biomaterials which assemble using supramolecular interactions.\textsuperscript{6,9–13} We are particularly interested in the use of structured peptides as supramolecular building blocks. Peptides can be prepared at scale with high purity, have canonical and non-canonical amino acids incorporated into their primary structure with high accuracy and are biocompatible.\textsuperscript{14} This creates a vast array of accessible structures that are easily tuneable, thanks to the stepwise synthesis of peptides on a solid support, creating an expanse of chemical space yet to be explored, with a broad chemical diversity of potential building-blocks and secondary structures. As such, the efforts to investigate this class of compounds as chiral, tuneable ligands has seen a surge in recent years.\textsuperscript{4,5,15–19}

Peptides are typically flexible, chiral, and present a multitude of chemical side chains, which results in complex intermolecular interactions and packing within the solid state. This often means the accessibility of good quality single-crystals for solid state analysis is difficult.\textsuperscript{6} Moreover, when structured peptides such as $\alpha$-helices and $\beta$-sheets are used as supramolecular building blocks, it is known that they can suffer perturbation of their periodicity upon functionalisation (\textit{i.e.}, deterioration of secondary structures) which leads to the inhibition of predictable self-assembly. The nature of the amino acids, their position within the sequence, and side-chain-to-side-chain interactions are all aspects that need to be carefully considered in order to minimise the risk of perturbation of the secondary structures.\textsuperscript{20} With these challenges in hand, it is essential to have a thorough understanding of the peptide secondary structure, and the resulting interacting moieties, to predict the assembly of the peptide ligands within supramolecular constructs. It is within this context that we propose the use of polyproline helices as supramolecular building blocks. The polyproline II helix is both rigid and stable in short sequences, and has three repeating helical faces, creating predictable and accessible handles
for functionalisation and supramolecular assembly.\textsuperscript{21–23} The lack of internal hydrogen bonding further simplifies the potential assembly principles compared to other helices.

We have recently demonstrated that super short polyproline helices (tetrameric peptides) can assemble into a reversibly porous supramolecular peptide framework (SPF) capable of engaging in stereoselective host encapsulation.\textsuperscript{24} Herein, we demonstrate the ability to utilise functionalised short polyproline helices as predictable ligands with an exceptional level of control, for the rational design of series of H-bonding driven supramolecular peptide frameworks. The design principles successfully applied to these peptides, can be used to drive the design of more complex peptide-based materials.

**Results and Discussion**

*Figure 1* – Chemical structures of synthesised oligoproline peptides, helical faces are highlighted in colour, H indicates the position of the hydroxyproline, starting from the N-terminus, in the sequence.
Minimalistic peptides have the potential to play a key role in the emerging field of bionanomaterials.\textsuperscript{25–27} We recently reported the first SPF formed by the self-assembly of a polyproline helical tetramer, Fmoc-(Pro)\textsubscript{4}-NH\textsubscript{2}, \textit{P\textsubscript{4}}. Despite the short length of this tetrapeptide, \textit{P\textsubscript{4}} crystallised in the polyproline II helical form, a common secondary structure found in nature.\textsuperscript{28} This SPF showed remarkable reversible porosity, the ability to reversibly host chemical guests in its channels and exhibited enantioselectivity for this adsorption process.\textsuperscript{24} The formation of this porous framework was driven by hydrogen-bond donor and acceptor (H\textsubscript{D}-H\textsubscript{A}) interactions between the hydrogens of the primary amide at the C-terminus and the carbonyl groups on the peptide backbone. These results gave us an insight into the potential of short polyproline helices as supramolecular building blocks.

We aim to exploit the full potential of polyprolines as minimalistic peptides in the construction of emerging bionanomaterials. We are particularly interested in the functionalisation of position 4 (or \(\gamma\)-carbon) of the proline amino acid as this position is exposed on the exterior of the helix\textsuperscript{22} and is capable, upon functionalisation,\textsuperscript{29} of engaging in the formation of supramolecular interactions. Remarkably, analysing the \textit{P\textsubscript{4}} supramolecular framework, we were able to successfully predict the effect of H\textsubscript{D} (i.e. -OH) interactions on the supramolecular assembly for a series of polyproline peptides. Led by these design principles, a series of seven hydroxy-functionalised derivatives were synthesised using Fmoc-based solid-phase peptide (SPPS) techniques (\textit{Figure 1, SI 1}). We anticipated that if the polyproline II conformation was retained for these peptides, the spatial orientation of the hydroxyl group would be highly predictable, thereby enabling future endeavours in the rational design of polyproline based ligands to assemble into supramolecular bio-constructs. If successful, we would also demonstrate the resilience and high level of positional control achievable using proline based minimalistic peptides.

Analysing the crystal structure of \textit{P\textsubscript{4}} we were able to observe that the terminal prolines, Pro1 (N-terminus) and Pro4 (C-terminus), clearly have close contacts with the neighbouring peptide along the \(b\)-axis (\textit{Figure 2}),\textsuperscript{24} suggesting hydroxyl groups at these positions are likely to extend new hydrogen-
Figure 2 – Crystal structures of peptides P₄ (a), highlighting close contacts between Cγ (Pro1 and Pro4) and the adjacent peptide’s closest carbonyl groups, HP₃ (b), P₃H (c), HP₂H (d), Cis-HP₂H (e), and AcHP₂H (f), showing the new hydrogen bond interactions formed by the additional hydroxyl moieties. Cis-HP₂H C-terminus hydroxyl has an undefined hydrogen bond to disordered solvent within the channels of the framework.

bonding interactions along the axis without significantly impacting the packing topology. As such a series of peptides were synthesised by replacing the terminal prolines with hydroxyprolines, these were peptides; HP₃, P₃H, HP₂H, Cis-HP₂H, and AcHP₂H (Figure 1). Due to the increased solubility of the functionalised peptides in polar solvents compared to peptide P₄, they were largely crystallised from a mixture of ethyl acetate or acetonitrile and ethanol, while peptide AcHP₂H, with the Fmoc group replaced with an acetyl group, crystallised from an acetonitrile solution (SI 3).

The first peptide tested, (Fmoc-Hyp-Pro3-NH₂) HP₃, crystallised to form a non-porous structure, driven by the addition of a hydrogen-bond between the Pro1 hydroxyl to the neighbouring peptide’s Pro1 carbonyl, forming 1D hydrogen bonded tapes of the peptide along the b-axis (Figure 2b), similar to the tapes of peptide along the b-axis already present in P₄ (Figure 2a). The amidated C-terminal formed the same hydrogen-bond interactions as found for the P₄ framework, with interactions with the Pro2 and Pro3 carbonyls of adjacent antiparallel peptides. The assembly of the peptides in these hydrogen-bonded layers into 2D layers via Fmoc-Fmoc interactions is very similar to the crystal structure of P₄, however, these 2D sheets then stack with a slight displacement to form the 3D framework,
significantly changing the unit cell and reducing the void volume, resulting in no solvent-accessible
channels present within the framework.

\( \text{P}_3\text{H} \) crystallised as a porous structure with channels (Volume 342.6 Å\(^3\), 10 % / unit cell, Probe \( r = 1.2 \) Å, Grid spacing 0.4 Å; SI 3.1.5), isostructural to the \( \text{P}_4 \) framework except for a doubling of the \( c \)-axis, to be expected from the doubling of \( Z' \). These pores are a slightly larger size than those found in the original framework (Volume 245.09 Å\(^3\), 13.8 % / unit cell, Probe \( r = 1.2 \) Å, Grid spacing 0.4 Å; Figure 3d), which presents a good example of the accessibility of alternative structures via editing of the peptide monomers, allowing tuning of the pore environment for various functionalities. In this case the asymmetric unit was comprised of two N-terminally hydrogen-bonded peptides. The new hydroxyl group has a hydrogen bond interaction with the Pro4 carbonyl of the adjacent parallel peptide, occurring similarly for both peptides within the asymmetric unit (\( \text{OH}---\text{O} = \text{C}, 2.693(3) \) and 2.762(5) Å, Figure 2c). The precise placement of the -OH groups along the same helical face, i.e., Pro1 and Pro4, produces H-bonding interactions towards adjacent carbonyls of Pro1 and Pro4 units respectively, thus highlighting the exceptional control possible, producing interactions with specific predictable geometries.

As both \( \text{HP}_2 \) and \( \text{P}_3\text{H} \) peptides form similar hydrogen bonded layers within their frameworks (Figure 2b-c), we predicted functionalisation of both positions would allow for both hydrogen bonds to be present simultaneously within the framework without significant disruption of the packing. Single crystals of \( \text{HP}_2\text{H} \) were successfully obtained in the same manner as \( \text{HP}_3/\text{P}_3\text{H} \), and subsequently used for single crystal analysis. The crystal structure contained both hydrogen bonds present in the previous structures as predicted, with similar hydrogen bond distances (Pro1 \( \text{OH}---\text{O} = \text{C}, 2.98(2) \) Å vs \( \text{HP}_3; 2.92(1) \) Å. Pro4; 2.71(1) Å vs \( \text{P}_3\text{H}; 2.693(3) \) Å, Figure 2d). However, the extended structure matched very closely with peptide \( \text{HP}_3 \), with only small differences in unit cell parameters (SI 3.1.7). The successful formation of this framework and retention of the polyproline II helix, despite 50 % functionalisation of the peptide clearly highlights the resilience of the polyproline helix and how the
Figure 3 – Crystal structures of peptides $\text{HP}_3$ (a), $\text{P}_3\text{H}$ (b), $\text{HP}_2\text{H}$ (c), $\text{Cis-HP}_2\text{H}$ (d), and $\text{AcHP}_2\text{H}$ (e), showing their extended structures viewed along the rows of peptides, atomic displacement parameters are shown at 50 % probability.

ability to predict the geometry of new functionalities can be utilised to rationally design supramolecular constructs.

To observe the impact of cis-hydroxy, versus the trans-hydroxy previously used, on the packing topology, the peptide $\text{Cis-HP}_2\text{H}$ was synthesised, with both hydroxyprolines cis rather than trans. In this case we expected the 4S-hydroxyproline to prefer the endo conformation and internally hydrogen bond to the hydroxyproline’s amide carbonyl, thus restricting the formation of intermolecular interactions. This peptide was successfully crystallised from a mixture of acetonitrile and ethanol, forming colourless plank crystals ($\text{SI 3.1.8}$). The crystal structure obtained contains channels filled with disordered solvent (volume 628.9 Å$^3$, 17.1 % / unit cell, Probe $r = 1.2$ Å, Grid spacing 0.4 Å, Figure 3e, $\text{SI 3.1.8}$), due to the presence of both acetonitrile and ethanol solvent molecules, and significant disorder, the solvent could not be accurately modelled as such solvent masking was used. However, the last hydroxyproline’s (C-terminus) hydroxyl group is aligned into the pore space, clearly hydrogen
bonding to a solvent molecule, with significant electron density adjacent to this group. While this hydroxyl was in the typical exo ring puckering, thus aligned mostly perpendicular to the helix, the N-terminal hydroxyl adopted endo ring puckering. The endo hydroxyl, now facing into the helix, was then able to satisfy the H₆ via intramolecular hydrogen bonding (OH---O=C, 2.793(4) Å, Figure 2e) towards the same hydroxyproline’s amide carbonyl (Figure 2e). Therefore, the only intermolecular hydrogen-bonding between peptides is the typical C-terminal NH₂ amide bonding present in all the structures seen previously. This highlights how small changes in the placement of functional groups can be used to affect the assembly process, with control over even the flexible pyrrolidine ring endo/exo conformations possible by use of 4S versus 4R functional groups, while the polyproline II helix remains as a rigid ligand for placement of these functional groups.

With the addition of two hydrogen bonding interactions in the HP₂H structure it seemed apparent that these hydrogen bonding interactions should still produce an extended 2D structure without the Fmoc-Fmoc interactions. As such the peptide AcHP₂H was synthesised, whereby the Fmoc protecting group on the N-terminus was replaced with an acetyl group (Figure 1). Initial attempts to crystallise the peptide indicated the clear formation of an extended network as the peptide formed an organogel upon sonication in a saturated solution (SI 3). Heating and slowly cooling a supersaturated solution of the peptide in acetonitrile, or slow evaporation of an acetonitrile solution, produced clusters of long fibre/needle-like crystals of the peptide, unlike those produced in any of the Fmoc containing structures. These largely 1-dimensional crystals were expectedly poorly diffracting; however, a crystal structure was obtained from a suitable crystal (SI 3.1.9), with the data giving a predicted powder pattern that matched well with PDXRD analysis of the bulk material (SI 3.2.9). As anticipated, analysis of the single crystal data showed the retention of the same hydrogen bonding interactions present in the HP₂H structure (Figure 2d+f) but, with the loss of the Fmoc interactions, there are no other significant interactions extending along two of the axes (Figure 3c). Channels within the framework contain well-ordered acetonitrile molecules with no apparent strong interactions between the solvent
Figure 4 – Crystal structure of P₄, highlighting the C4 positions on the Pro2 residue (green) and the lack of close contact hydrogen bond acceptors.

and the peptide within the structure. Remarkably, comparing the structures of AcHP₂H and HP₂H we can clearly conclude that, while not necessary to the formation the peptide framework, the terminal groups can also be functionalised with chemical handles (e.g., Fmoc/Acetyl) to increase the level of control of the supramolecular assembly process for these short peptides.

We then analysed the other positions on the P₄ backbone. The Pro2 residue on the P₄ backbone is aligned with adjacent peptides’ aromatic groups and the closest hydrogen bond acceptor (Hₐ) is 4.4 Å from the C4 position, which suggests the current crystal structure is unlikely to satisfy the hydrogen bond donation of a new hydroxyl at this position (Figure 4). Unfortunately, the PHP₂ peptide did not crystallise effectively. We hypothesise that the absence of favourably positioned Hₐ groups close to the Pro2 C4 position (closest carbonyl Hₐ, C–O 4.442 Å) in the P₄ framework is impacting the assembly process. However, PDXRD of the precipitate, from the attempted crystallisation of PHP₂ in EtOAc/EtOH, showed partial crystallinity (SI 3.2.3).

The Pro3 residues of the P₄ peptide, face into the channels of the framework and as such are prime candidates for functionalisation of the channels to affect host-guest interactions (Figure 5a). As such peptide P₃HP was synthesised and successfully crystallised with the same packing as the P₄ framework, retaining polyproline II helicity. The hydroxyl groups line the previous channels of the framework, with the main structural difference being the hydroxyl forcing exo puckering of the attached pyrrolidine ring, which is typical for this functionality,³¹ and the subsequent reduction in pore volume (Volume
Figure 5 – Crystal structures of \( P_4 \) (a), \( P_2 \text{HP} \) (b) and \( P_4+P_2 \text{HP} \) (c) mixed crystal, depicting two peptide units adjacent to the framework channels, view along the \( b \)-axis, Mercury. Hydrogen bond interactions are shown, (c) is slightly off the \( b \)-axis to show both hydrogen bonds to different ethanol molecules.

161.8 \( \text{Å}^3 \), 9.2 \% / unit cell, Probe \( r = 1.2 \text{ Å} \), Grid spacing 0.4 \( \text{Å} \), due to the steric bulk from the additional functional group (Figure 5b, SI 3.1.4). This reduced pore size is too small to accommodate EtOH molecules, as such the framework is selective towards \( \text{H}_2\text{O} \) molecules in the wet solvent during crystallisation, which can be clearly modelled in the crystal structure (Figure 5b). This change in properties of the framework was exemplified through the repetition of thermal activation studies carried out on the \( P_4 \) framework previously, whereby encapsulated solvent could be removed resulting in the collapse of the channels forming a non-porous structure.\(^{24} \) In this case heating the \( P_2 \text{HP} \) framework under reduced pressure at 45 °C did not show the formation of a second crystalline phase (Figure S42, SI 3.2.4), showing how the solvent filled voids are now trapped in place by the additional hydroxyl moieties. This ability to functionalise the pores of the framework with no disruption of the peptide helix, thereby allowing tuning of the selectivity of the pores has clear applications towards specific host-guest interactions.

Due to the similarity between the two crystal structures of both \( P_4 \) and \( P_2 \text{HP} \) we theorised that both peptides would crystallise together with little differentiation between the peptides in the self-assembly process. Therefore, the two could be combined to create a porous framework spiked with the hydroxyproline moiety, reducing the effect of the limiting reduction of the channel diameter found in the \( P_2 \text{HP} \) structure. When combined the two peptides crystallised by cooling a supersaturated
Figure 6 – Crystal structures of P₄ (a), P₂HP (b) and P₄+P₂HP (c) mixed crystal, showing their closely matched extended structures viewed along the b-axis, atomic displacement parameters are shown at 50 % probability.

solution of the peptides (1:1 molar ratio) in ethanol. We were pleased to see that SC-XRD analysis of the crystals formed clearly showed the presence of both peptides (Figure 6c, SI 3.1.6). Within the pores of this SPF, the hydroxyl moiety had a reduced chemical occupancy of the oxygen atom (0.375), indicating the lack of differentiation between the two peptides during the self-assembly process resulting in the formation of a “solid solution”, with P₂HP randomly dispersed throughout the extended P₄ framework. Interestingly this framework adopted the endo conformation similarly to P₄ alone, having significantly less impact on the channel volume, and thus contained EtOH within the pores rather than H₂O (Figure 5-6c). As mentioned previously the P₄ peptide framework could be activated by heating (45 °C) under reduced pressure, with characterisation through both PDXRD and SCXRD showing encapsulated EtOH had been removed, forming a collapsed non-porous structure. As the mixed framework exhibited a very similar structure to the P₄ framework, with EtOH encapsulated in accessible channels(Figure 5-6), while the P₂HP peptide framework contained water trapped in isolated pores (Figure 5b-6b), and significantly showed no change after attempted activation (SI 3.2.4, Figure S42), we anticipated the mixed framework would exhibit different behaviour to the peptides alone. Interestingly, the mixed framework P₄+P₂HP showed hybrid behaviour compared to the two single peptide frameworks, with the powder pattern upon activation for P₄+P₂HP bearing similarities to both activated single peptide frameworks (i.e. P₄ and P₂HP) with some formation of the new characteristic peaks seen in the P₄ “activated” diffractogram (Figure S47, SI 3.2.6) while a portion of
the original peaks remained. This suggested only partial “activation” and collapse of the pores (SI 3.2.6, Figure S47).

Conclusion

In conclusion, we have synthesised a series of hydroxyproline-based derivatives of an oligoproline tetramer, successfully forming novel H-bonding driven supramolecular peptide frameworks. Using rational design, based on the original supramolecular framework, for the placement of additional functional groups we were able to consistently predict the position of the hydroxyl groups in relation to the polyproline helix and therefore anticipate their effect on the topology of the H-bonding supramolecular network formed. We clearly demonstrate that we can modulate the nature of the frameworks (i.e., porous or non-porous) as well as the size and the properties of the pores (e.g., solid solution framework), while functionalising up to 50% of the peptide’s backbone. The resilience of the polyproline II helix, shows how their predictable geometries can be utilised for the rational design of discrete and extended supramolecular three-dimensional structures. This work lays the groundwork for further studies focusing on the polyproline helix, to rationally design structural units capable of forming desired supramolecular structures with tuneable structural features and functionalities.

Associated Content

Supporting information is available free of charge via the Internet. CCDC-2127751,24 2234312, 2238152, 2238155, 2238160-1, 2238180 and 2238252 contain the supplementary crystallographic data for this paper, including structure factors and refinement instructions, and can be obtained free of charge from The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (e-mail: deposit@ccdc.cam.ac.uk), or via https://www.ccdc.cam.ac.uk/getstructures.

Acknowledgments

We thank the EPSRC (grant no. EP/T016140/1) and the Royal Society of Chemistry (grant E21-9299054940) for their generous financial support. Also, thanks to the University of Kent (Vice Chancellor’s fellowship).

Conflict of Interest

There are no conflicts of interest to declare.

Abbreviations

SCXRD, single-crystal X-ray diffraction; PDXRD, powder x-ray diffraction; SPPS, solid-phase peptide synthesis; SPF, supramolecular peptide framework
References


