nature chemistry

ARTICLES

PUBLISHED ONLINE: 28 AUGUST 2011 | DOI: 10.1038/NCHEM.1123

Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel

Lesley E. R. O'Leary, Jorge A. Fallas, Erica L. Bakota, Marci K. Kang and Jeffrey D. Hartgerink*

Replicating the multi-hierarchical self-assembly of collagen has long-attracted scientists, from both the perspective of the fundamental science of supramolecular chemistry and that of potential biomedical applications in tissue engineering. Many approaches to drive the self-assembly of synthetic systems through the same steps as those of natural collagen (peptide chain to triple helix to nanofibres and, finally, to a hydrogel) are partially successful, but none simultaneously demonstrate all the levels of structural assembly. Here we describe a peptide that replicates the self-assembly of collagen through each of these steps. The peptide features collagen's characteristic proline–hydroxyproline–glycine repeating unit, complemented by designed salt-bridged hydrogen bonds between lysine and aspartate to stabilize the triple helix in a sticky-ended assembly. This assembly is propagated into nanofibres with characteristic triple helical packing and lengths with a lower bound of several hundred nanometres. These nanofibres form a hydrogel that is degraded by collagenase at a similar rate to that of natural collagen.

Collagen, the most abundant protein in the human body, exemplifies multi-hierarchical self-assembly. In the case of type I collagen, self-assembly begins with three 1,000 amino acid peptide strands that adopt a polyproline type II helical structure and wind around one another to form a superhelical trimer that gives the well-known collagen triple helix. These triple helices then pack against one another in a quasihexagonal and staggered fashion to form nanofibrous structures known as collagen fibrils[1,2]. Collagen fibrils continue to self-assemble both linearly and laterally to give collagen fibres and a hydrogel network (Fig. 1). Together, the multiple levels of collagen's structural hierarchy play a major role in the structural integrity of the extracellular matrix and provide binding sites for other proteins and cells.

Collagen has been the target of biomimetic design for decades because of the many difficulties associated with the use and characterization of collagen from natural sources and by expression. The use of recombinant systems requires either genetic modifications or a novel biosynthetic pathway in *Escherichia coli* to express hydroxyproline-containing collagens[3–8]. There are many successes in the recapitulation of the collagen triple helix in short peptides, both as a homotrimer[9–17] and, more recently, as a heterotrimer[18–24]. However, examples that take these collagen-like peptides and use them to mimic the higher order assembly of collagen have faced a great deal of difficulty. In all previously reported systems, none has demonstrated discretely each level of collagen self-assembly (triple helix, nanofibre and hydrogel) within the same system. There are many examples of peptides that form organized nanostructures without gelation[10,12,25–29], and a few that show gelation without proof of the presence of triple helices or nanofibres[30,31]; however, no system with triple-helix formation, nanofibre formation and gelation is reported.

Perhaps the best example of a fibre-forming collagen-like peptide was demonstrated by Chaikof, Conticello and co-workers, who prepared a 36 amino acid peptide with the sequence $(Pro\text{-}Arg\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Glu\text{-}Hyp\text{-}Gly)_4$ (ref. 32). This zwitterionic peptide assembled into large organized fibres. However, even these collagen mimetic fibres have some drawbacks, which include (i) a mixed composition of fibres associated with a significant quantity of other amorphous material, (ii) the requirement for specific concentration and buffer composition outside of which the quality of assembly degrades or fails completely and (iii) phase separation and precipitation of the formed fibres as opposed to the formation of a hydrogel[32].

Recently, in our laboratory we investigated the structure of several heterotrimeric collagen helices using NMR spectroscopy[18,24],

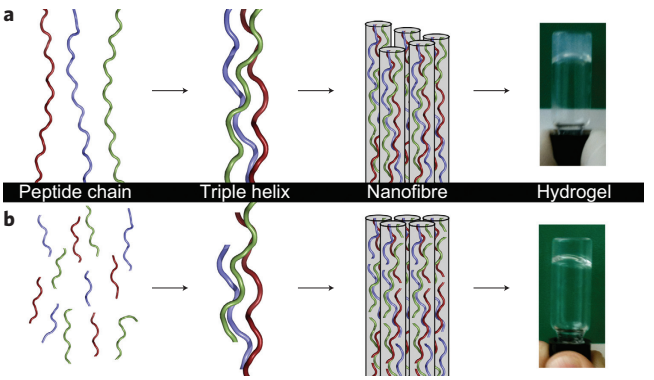

Figure 1 | Self-assembly of collagen type I compared to that of collagen mimetic peptides. a, Scheme of type I collagen assembly in which the peptide chains (shown in red, blue and green) consist of 1,000 amino acids, the triple helices are 100 nm in length and the blunt-ended nanofibres (shown in grey) assemble via the staggered lateral packing of the triple helices. The hydrogel pictured is from a rat-tail collagen sample. **b**, Scheme for the self-assembly of collagen mimetic peptides in which the peptides consist of 36 amino acids (shown in red, blue and green), the triple helix is staggered with a length of 10 nm and the nanofibres (shown in grey) result from triple helical elongation as well as from lateral packing. The hydrogel pictured is the designed peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$.

Departments of Chemistry and Bioengineering Rice University, Houston, Texas 77005, USA. *e-mail: jdh@rice.edu

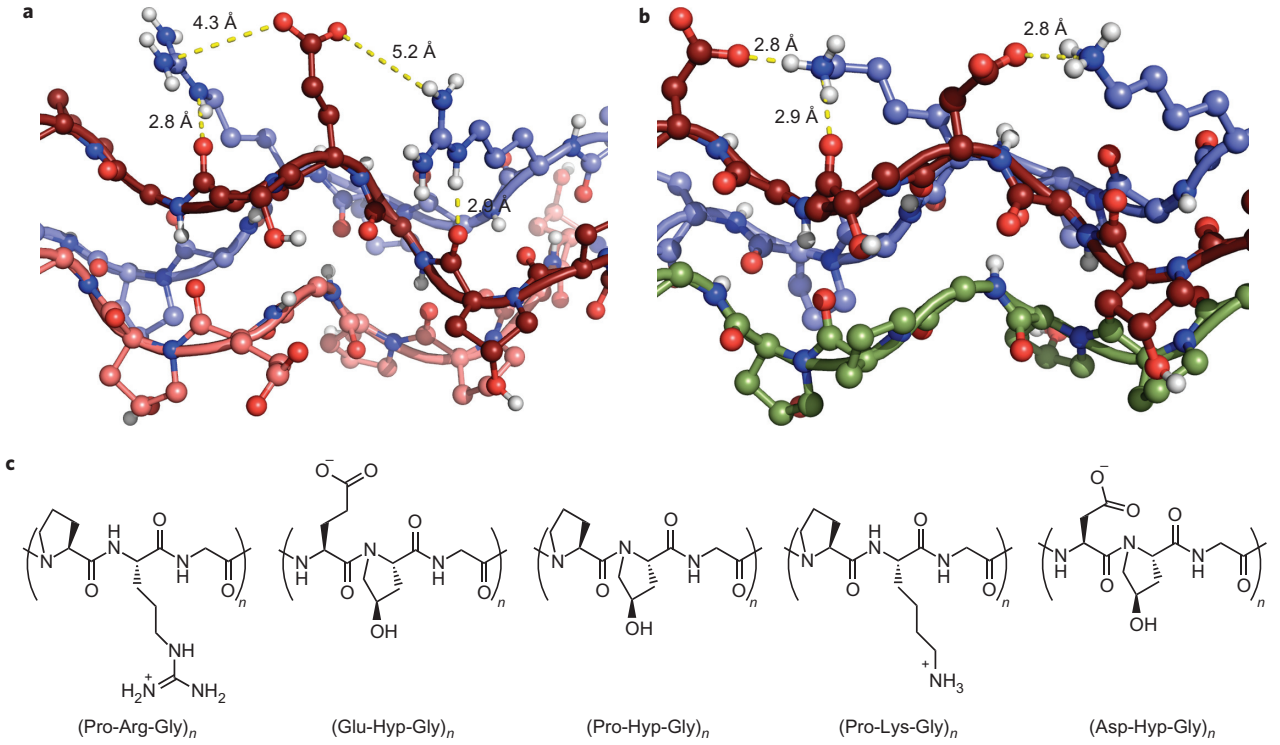

Figure 2 | Models of electrostatic interactions between charged amino acids in collagen mimetic peptides. a,**b**, Models of Arg–Glu (**a**) and Lys–Asp (**b**) charged pairs in collagen triple helices[18,24]. The peptide chains are shown in red, blue and pink for (**a**) and red, blue and green for (**b**), with the hydrogen atoms highlighted in white, oxygen in pink and nitrogen in blue. The hydrogen-bond lengths shown are measured from N to O. Arg–Glu pairs do not appear to form high-quality interactions because of the strong hydrogen bonding between Arg and a cross-strand carbonyl oxygen, which locks the side chain into place. In contrast, two conformers of lysine are present and both allow excellent hydrogen bonding to aspartic acid despite one of them displaying a similar hydrogen bond to a cross-strand carbonyl. **c**, Chemical structures of the common amino-acid triplets $(Pro\text{-}Arg\text{-}Gly)_n$, $(Glu\text{-}Hyp\text{-}Gly)_n$, $(Pro\text{-}Hyp\text{-}Gly)_n$, $(Pro\text{-}Lys\text{-}Gly)_n$ and $(Asp\text{-}Hyp\text{-}Gly)_n$.

in these systems, the arginine–glutamate interactions on which they rely are distant from one another and interact primarily by charge screening rather than by a specific salt-bridged hydrogen bond (Fig. 2a)[24]. One of the reasons for this is that the arginine side chain forms a tight hydrogen bond with the backbone carbonyl of an adjacent peptide chain, which restrains it from making a more intimate contact with glutamate. In contrast, we observed very high-quality formation of lysine–aspartate salt-bridged hydrogen bonds (Fig. 2b)[18]. Based on these charge-pair observations and the work of Chaikof and Conticello, we prepared a new peptide in which the arginine residues are replaced with lysine and the glutamate residues with aspartate, to give the sequence $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$. Our hypothesis was that the more effective interactions between lysine and aspartate, previously observed, would result in superior fibre- and hydrogel-forming characteristics. Indeed, this is what we observed.

Here we report the synthesis and multi-hierarchical assembly of the collagen mimetic peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ through each level of assembly, as depicted in Fig. 1b. This peptide demonstrated successfully the formation of a stable triple helix with a melting temperature of 40–41 °C. It exhibited nanofibre morphologies, as observed in atomic force microscopy (AFM), scanning electron microscopy (SEM) and transmission electron microscopy (TEM), including both dry and hydrated techniques, and the nanofibres formed were quite uniform, with virtually no other aggregations or morphologies observed. Furthermore, nanofibrous self-assembly was observed easily under a wide range of buffers and ionic strengths, which indicates the robust nature of the self-assembly process. The nanofibres displayed characteristic triple helical packing, as confirmed by fibre diffraction, and self-assembled into hydrogels with good viscoelastic properties, as measured by oscillatory rheology and through comparisons to both natural and synthetic hydrogels. Finally, the prepared hydrogels were broken down by collagenase type IV at a similar rate to rat-tail collagen in a simple functionality test[33]. As the system demonstrated control at each of level of collagen assembly (triple helicity, fibre formation and hydrogel formation), we believe this peptide, as well as future systems based on it, have a large potential for use as tissue-engineering scaffolds.

Results and discussion

Once the peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ was synthesized and purified successfully (complete details of the procedures are given in Methods and in the Supplementary Information), samples were made at specified concentrations between 0.2% (0.6 mM) and 1.0% (3 mM) by weight. Although many buffer systems with varying ionic strengths were explored (see Supplementary Information for further details), here we discuss primarily the results using 10 mM sodium phosphate buffer at pH 7 (referred to as phosphate). In this buffer system, all the samples made at concentrations of 0.5% (1.5 mM) by weight or more formed hydrogels within a few hours. Once hydrogel formation was observed, we began systematically to analyse the peptide at each level of self-assembly: triple helix, nanofibre and hydrogel.

Triple helix. To determine whether a collagen mimetic peptide forms a triple helix, two circular dichroism (CD) experiments must be performed: a wavelength spectrum and a thermal unfolding curve. Collagen triple helices have a signature CD

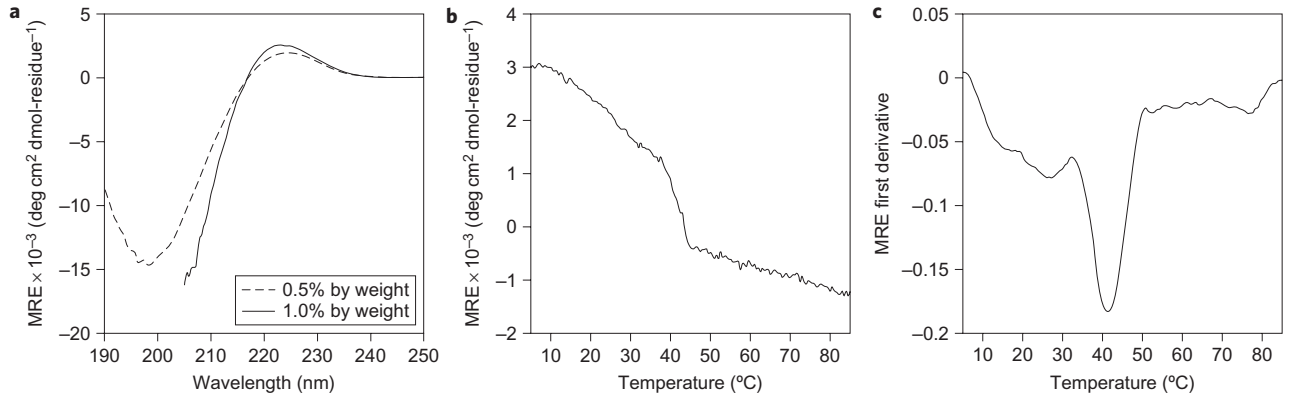

Figure 3 | Spectroscopy graphs illustrating the triple helical nature of the designed collagen mimetic peptide. a, CD spectrum of the fibre-forming collagen-like peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ at 0.5% and 1.0% by weight concentrations in phosphate at a temperature of 5 °C. The maximum at 225 nm and minimum near 200 nm are characteristic of polyproline type II helices. **b**, Thermal unfolding analysis curve for the peptide at 1.0% by weight concentration in phosphate. **c**, First derivative of the mean residue ellipticity (MRE) versus temperature curve shown in (**b**). The cooperative transition at 41 °C demonstrates that the designed peptide forms a triple helix.

spectral profile that consists of a maximum at 225 nm and a minimum near 200 nm, indicative of a polyproline type II helix. The thermal unfolding experiment monitors the spectral maximum as temperature is increased which, when a triple helix is present, shows a cooperative transition. For the peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$, CD spectra taken at all concentrations showed a strong maximum at 225 nm. The spectra for 0.5% and 1.0% by weight in phosphate are shown in Fig. 3a. Note the size difference between the maximum in the spectrum for 1.0% by weight compared to that for 0.5% by weight, even though the data is normalized for concentration. This indicates an increased percentage of the peptide is folded at the higher concentration. When melting experiments were performed from 5 to 85 °C on samples at concentrations of 0.2%, 0.5% and 1.0% by weight, all the samples exhibited a cooperative transition in the melting profile. Additionally, transitions for samples at higher peptide concentrations were stronger and more obvious than those for lower concentrations, which indicates that the higher concentration of peptide helps to drive triple-helix formation. The thermal unfolding curve and the first derivative of the curve for a 1.0% by weight sample in phosphate are shown in Fig. 3b,c. (Melting studies for 0.2% and 0.5% by weight are given in the Supplementary Information.) A major transition occurs in the first derivative curve at 40–41 °C, which corresponds to the melting temperature of the peptide. However, a broad, minor transition is also visible between 10 and 30 °C. The minor transition may result from increased helicity on fibre elongation and lateral packing. A more detailed explanation for this is given below.

Nanofibre. After confirmation of the triple helical nature of the peptide, the next step was to understand the nanostructure of the self-assembled peptide. The microscopy techniques used included TEM, AFM and SEM. TEM is an integral technique used to view the morphology and measure the length and width of structures on the nanoscale. It is most commonly a dry technique that, to view carbon-based materials, requires the sample to be stained with a heavy metal, such as phosphotungstic acid (PTA). For this peptide, a 1.0% by weight concentration sample in phosphate was prepared, using the previously described procedure, and negatively stained with PTA. Images of these stained samples (Fig. 4a,b and Supplementary Information) reveal long nanofibres, present both as single fibres and as fibre bundles. Figure 4a shows the variety of fibre widths present within this system when dried and stained. These fibres are the major species within the TEM sample, in contrast to previous reports of collagen mimetic nanofibres that showed large aggregates and a variety of other non-fibrous structures in the TEM images. Figure 4b reveals the twisting nature of some of the nanofibres in contrast to fibres with a smoother morphology. Although the negatively stained TEM images show the presence of nanofibres for this peptide system, drying artefacts can cause samples to appear more densely packed or with a completely different structure than that present in the hydrated state. In addition, the use of a heavy-metal stain adds an additional level of uncertainty in assessing fibre size and morphology. For these reasons, we believe the presence of nanofibres in a solution state can only be proved by imaging the system in a hydrated environment, specifically using vitreous ice cryo-TEM.

The sample preparation for cryo-TEM differs greatly from that for dry TEM because cryo-TEM requires a thin aqueous film of sample on the TEM grid before it is flash frozen in ethane slush. A representative TEM image from this preparation is given in Fig. 4c (for more images see the Supplementary Information). In contrast to the dry TEM images, especially Fig. 4a, the fibres seen in the vitreous ice cryo-TEM sample have uniform widths from 4 to 5 nm and fibre lengths from several hundred nanometres to many microns. However, similar to the dry cryo-TEM images, the fibres observed in cryo-TEM form the majority of the peptide population in the sample. (The spherical species seen in the cryo-TEM image are ethane artefacts that result from sample preparation, not peptide aggregates.) Therefore, both dry and cryo-TEM confirmed the presence of nanofibres and that they were the major species within the system. Once the length and width of the nanofibres formed from the peptide was determined by cryo-TEM, the height of the fibres was needed to understand the mechanism of fibre formation. Tapping-mode AFM is the most efficient method for acquiring this data. Figures 4d,e are AFM images taken of 1.0% and 0.5% by weight samples in phosphate buffer. Nanofibres are seen in both images; the sample of higher concentration exhibited a denser network of nanofibres. Similar fibres were seen in all the buffers examined, including higher ionic strength buffers such as phosphate-buffered saline. The measured height profile in phosphate buffer from the AFM images was 1.2 ± 0.3 nm. This value is much lower than the fibre width of 4–5 nm measured by cryo-TEM, and the fibre lengths observed by AFM also appear smaller than those observed by TEM. A hypothesis for this difference is discussed below. One advantage of these images is that, as a result of their lower magnification, a larger area is visible and the uniformity of the population of self-assembled nanofibres is more apparent.

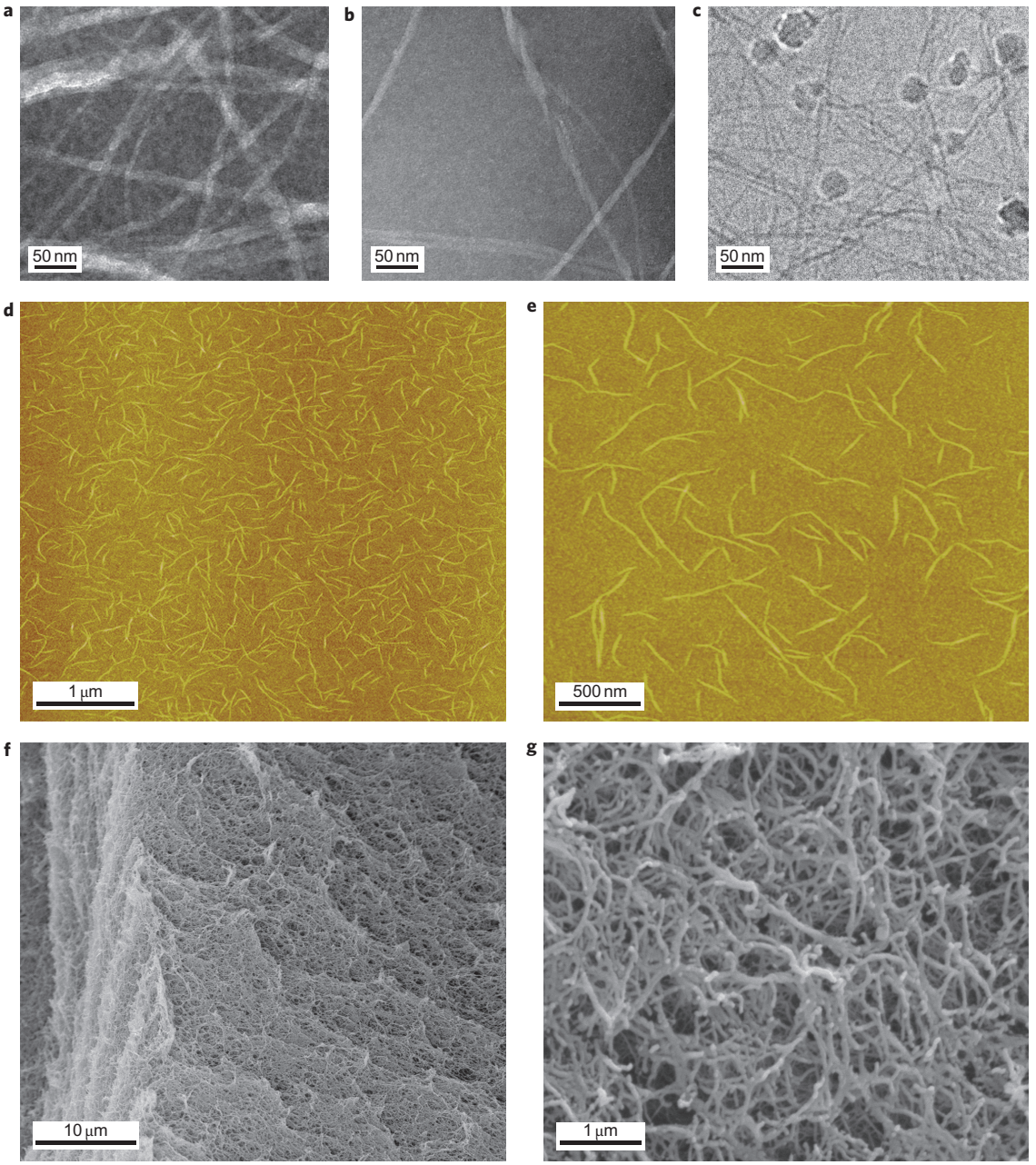

Figure 4 | Microscopy images of (Pro-Lys-Gly)$_4$(Pro-Hyp-Gly)$_4$(Asp-Hyp-Gly)$_4$ that show the fibrillar assembly of the system in phosphate buffer. a–c, TEM images of collagen-like nanofibres in phosphate taken at ×40,000. **a**,**b**, Negatively stained images of the peptide in phosphate at a concentration of 1.0% by weight and stained with PTA, pH 6. **c**, Vitreous ice cryo-TEM image of collagen-like nanofibres taken in phosphate at a concentration of 0.25%, which was diluted from a 0.5% by weight sample. **d**,**e**, AFM of collagen-like nanofibres in phosphate as observed after spin coating onto freshly cleaved mica from solutions of peptide at concentrations of 1.0% (**d**) and 0.5% (**e**) by weight. **f**,**g**, SEM images of critical point dried hydrogel with a peptide concentration of 1.0% by weight that show the interconnected fibrous structure responsible for the gel forming properties at ×3,100 (**f**) and ×30,000 (**g**).

One final microscopy method, SEM, is important for understanding the qualitative long-range nanoscale behaviour of the system. Samples imaged by SEM were 1.0% by weight in phosphate buffer. In Fig. 4f, a dense fibre network that is homogeneous and extends across tens of microns is apparent. When the magnification is increased (Fig. 4g), the uniform nature of the nanofibres within the network is more obvious. These results directly complement the fibre morphologies observed by TEM and AFM and also indicate the three-dimensional structure of the hydrogel.

With the use of multiple microscopy techniques, the nanomorphology of (Pro-Lys-Gly)$_4$(Pro-Hyp-Gly)$_4$(Asp-Hyp-Gly)$_4$ was found to be nanofibres of relatively uniform dimensions with observed lengths of at least several hundred nanometres, widths of 4–5 nm, measured heights of 1.2 ± 0.3 nm and a uniform long-range behaviour visible in the hydrated state.

Hydrogel. With the first two levels of self-assembly confirmed, the final layer of analysis needed to describe the multi-hierarchical assembly of (Pro-Lys-Gly)$_4$(Pro-Hyp-Gly)$_4$(Asp-Hyp-Gly)$_4$ was the assessment of the viscoelastic properties of the formed hydrogel (Fig. 5). Visually, the gels maintained their shape when they were removed from their containers, including the visible sustainability of the gel's sharp edges. The image in Fig. 5d depicts the visual properties of the hydrogel. To analyse

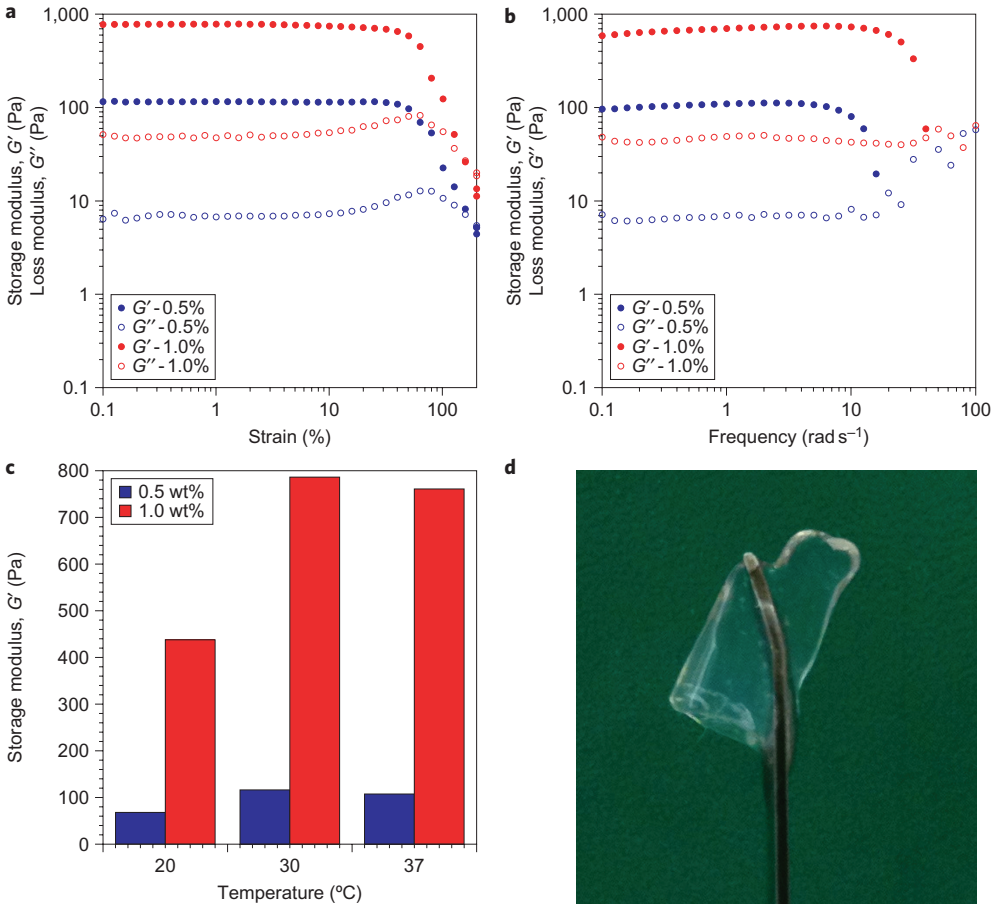

Figure 5 | Rheology of the collagen-like peptide that demonstrates the temperature-dependent strength of the hydrogel. a, Strain sweep at 0.5% and 1.0% by weight peptide concentration in phosphate buffer at a temperature of 30 °C and a frequency of 1 rad s^{-1} shown as G' and G''. **b**, Frequency sweep at 0.5% and 1.0% by weight peptide concentration in phosphate at a temperature of 30 °C and 1% strain shown as G' and G''. **c**, Temperature dependence of rheological properties at 0.5 and 1.0% by weight peptide concentrations shown as G'. Data points were acquired at 1 rad s^{-1} and 1% strain. **d**, Photo of the shape-persistent nature of the gel with a concentration of 1.0% by weight in phosphate. The gel was prepared at a volume of 0.5 ml. Note the sustainability of the sharp gel edges.

quantitatively the peptide hydrogels, rheological studies were performed. Strain and frequency-sweep experiments were performed to assess the gel properties and, specifically, the storage modulus (G') and loss modulus (G''), which measure, respectively, the elastically stored energy and the energy lost as heat within the hydrogel on application of a shearing force. Representative graphs of each type of experiment are shown in Fig. 5a,b and the first observation is that the G' is substantially larger than G'' for both 0.5% and 1.0% by weight concentrations of the peptide in phosphate buffer. Therefore, $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ forms a hydrogel in phosphate buffer at 0.5% by weight concentration and higher. The observed G' of this collagen mimetic system is similar to that typically observed for a collagen hydrogel formed from natural sources, such as rat-tail collagen, even though our peptide is approximately 30 times shorter (36 amino acids as compared to 1,000) (ref. 34). It is also higher than that of Matrigel[35] and comparable to those of the popular β-sheet hydrogels described in the literature[36–41].

The collagen mimetic hydrogel was temperature sensitive. From the CD melting studies, we know that the triple helix unfolds at 40–41 °C, and therefore a temperature-ramp rheological experiment from 20 to 60 °C was used to demonstrate the melting of the hydrogel. Indeed, the G' values decreased above 40 °C, and by 50 °C the G'' values exceeded the G' values, which indicates that the gel disassembled (see Supplementary Information). Figure 5c is a bar graph of the G' values for 0.5% and 1.0% by weight gels in phosphate at 20, 30 and 37 °C. The system was examined at different temperatures to gain insight into its behaviour before the gel melted. As shown in Fig. 5c, the gels have their highest G' at 30 °C and 37 °C, and a substantially lower observed storage modulus at 20 °C. The CD melting profile shows a minor transition of the peptide between 10 and 30 °C prior to the actual triple-helix unfolding of the system. The temperature-dependent rheological results combined with the CD data suggest that, as the peptide slightly unfolds between 10 and 30 °C, the unfolded regions of fibre may interdigitate with other nanofibres and result in a strengthened hydrogel.

As a simple functional test of the collagen hydrogel mimetic, we compared its ability to be broken down by collagenase (type IV, Invitrogen), the primary component of which is MMP2, a protease known to cleave specifically between the X and Gly residues of an X-Y-Gly repeat found in a triple helix[42]. $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ hydrogels were prepared at a concentration of 2.0% by weight in phosphate buffer and treated with either collagenase in Hank's Balanced Salt Solution (HBSS) or with HBSS alone. The samples were allowed to incubate at room temperature (approximately 20 °C), 30 °C and 37 °C. Hydrogels of rat-tail collagen were prepared in the same fashion, with and without collagenase. As shown in Supplementary Table 1, hydrogels prepared from our self-assembling peptide and rat-tail collagen degraded at similar

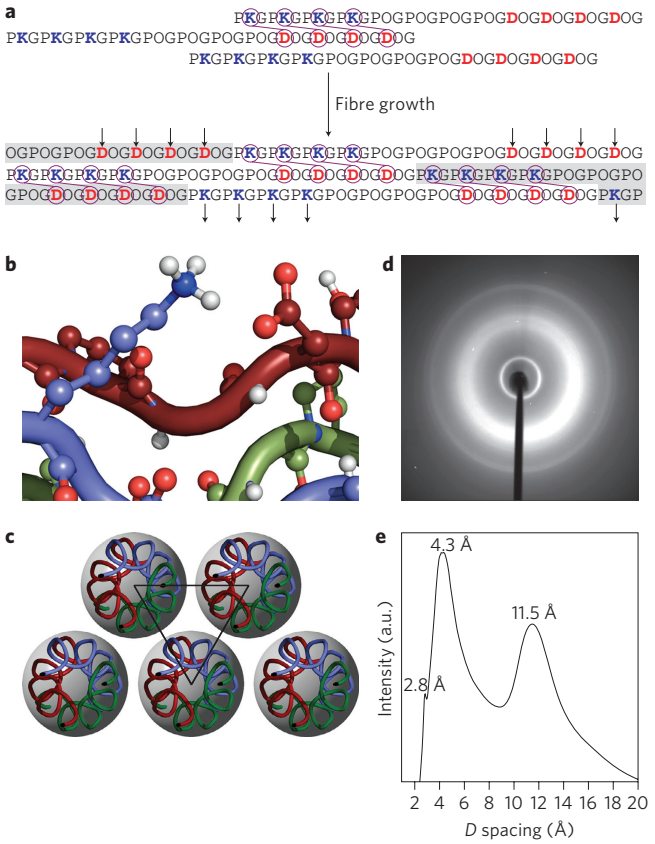

Figure 6 | Proposed mechanism of fibre self-assembly. a, Peptide sequence shown as single-letter amino acid code with P for proline, K for lysine, G for glycine, O for hydroxyproline and D for aspartate. The minimum repeating unit of the triple helical fibre has extensive 'sticky' ends. As additional peptides (shaded grey) add to the minimum repeating unit, the percentage of amino acids that forms a high-quality triple helical structure increases rapidly. Positively charged lysine residues are in blue, negatively charged aspartates are in red and satisfied intrahelical electrostatic interactions are indicated by purple lassoes. Available interhelical charged-pair hydrogen bonds are shown by small arrows. **b**, Lysine–aspartate interaction between the *i* and *i* + 3 amino acids of adjacent peptide strands. **c**, Quasihexagonal packing of growing fibres results in a bundle approximately 2 by 4 nm based on a triple helical cross-section of 1.2 nm. **d**,**e**, Fibre diffraction pattern (**d**) and its radially averaged intensity (**e**). Characteristic bands at 2.8 Å, 4.3 Å and 11.5 Å match well with previously reported fibre diffraction from natural collagen. a.u. = arbitrary units.

rates: samples of both types of hydrogels treated with collagenase dissolved fully after one hour (37 °C) or four hours (30 °C and room temperature), whereas untreated controls did not.

Fibre diffraction. To learn more about the packing morphology of the self-assembled nanofibres and, in turn, the fibre self-assembly process (Fig. 6), X-ray fibre diffraction studies were carried out on a dried peptide sample (see Methods). As is apparent from the microscopy images (Fig. 4), neighbouring fibres lack a common orientation axis. To align the fibres partially, the drying peptide solution was placed in a strong magnetic field to promote alignment during the drying process. This methodology has been shown to produce highly aligned protein fibres[43], but had only limited success in our system. Figure 6d shows the recorded diffraction pattern. The dried pellet exhibits some alignment, as evidenced by the pseudo two-fold symmetry observed in the intensity versus azimuthal angle scan of the diffraction pattern (see Supplementary Information). However, no clear equatorial or meridional axis could be determined and thus the data were analysed by performing a radial integration of the diffraction pattern to yield a plot of the observed intensities as a function of *D* spacing (Fig. 6e). The plot shows three distinct features: a weaker, sharp line near 2.8 Å, a diffuse intense reflection near 4.3 Å and a strong well-defined band near 11.5 Å. The spacing of the observed lines agrees well with that observed for collagen from stretched kangaroo-tail tendons[44]. Based on this, we assigned the 11.5 Å band to the distance between two triple helices within the nanofibres, the diffuse reflection at 4.3 Å to the distance between peptide chains inside a triple helix and the reflection at 2.8 Å to the translation per triple helical triplet. This suggests that our collagen-like peptide fibres pack in a fashion similar to that of natural collagen.

Proposed mechanism of assembly. As mentioned above, the charge pairing of lysine and aspartate was shown previously to form direct electrostatic interactions in collagen mimetic peptides[18]. Specifically, lysine's side chain reaches in a C-terminal direction to make an intimate salt-bridge hydrogen bond with an aspartate on an adjacent lagging peptide offset by three amino acids (Fig. 6b). Our peptide forms a homotrimer, so there is a potential for these charged amino-acid salt bridges to form between peptide strands and create an offset, sticky-ended triple helix. Similar sticky-ended assemblies have been designed and reported, particularly for alpha-helical coiled coils[45,46]. Figure 6a shows the proposed repeating unit of peptide self-assembly. Lysine–aspartate interactions are highlighted with purple lassoes. This favourable interaction forces a dramatic sticky-ended triple helix in which only one-third of the possible lysine–aspartate pairs are satisfied. However, as additional peptides are added to extend the triple helical system, the fraction of satisfied charge pairs increases. For example, adding just one more peptide increases the fraction of satisfied charge pairs to one-half and an infinite-length triple helical fibre will have two-thirds of the salt-bridges satisfied through intrahelical interactions. In addition, for our collagen mimetic system, fibre elongation satisfies a larger percentage of interpeptide backbone hydrogen bonds donated from glycine, which are known to stabilize collagen triple helices[47–52]. In the three-peptide nucleation centre, only 50% of the glycine residues are capable of forming these interpeptide interactions; however, as the fibre grows, the percentage of glycines that participate in hydrogen bonds approaches 100%.

As observed by TEM, SEM and AFM, the nanofibres formed have dimensions greater than those of a single collagen triple helix. Therefore, several triple helices must bundle together to form the observed nanofibres. This is supported by the fibre diffraction data, which clearly display the characteristic triple-helix packing band at 11.5 Å (Fig. 6d,e). The lysine and aspartate side chains not participating in intrahelix salt bridges (indicated by small arrows in Fig. 6a) are available for interhelix interactions, which promote helix bundling. In natural collagen, five helices are believed to pack in a quasihexagonal fashion to form fibrils that continue to assemble into mature fibres[1,2]. Based on the height and width for the $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ nanofibres measured by AFM and cryo-TEM, respectively, and a helix-packing distance from fibre diffraction, we hypothesize that our peptide system assembles in a similar fashion. Figure 6c illustrates this packing. The AFM-calculated nanofibre height was 1.2 ± 0.3 nm and the nanofibre width observed by cryo-TEM was 4–5 nm. Both of these measured values are within reason for our proposed quasihexagonal packing. However, for the fibre height the value measured by AFM appears to be significantly less than expected. There are several possible explanations for this. First, it is known that in AFM the measured heights of soft organic materials are often less than expected because of flattening from

surface forces or from the AFM tip itself[53]. Another possible explanation is that the triple helices not in direct contact with the mica surface are removed during the washing step and leave behind collagen ribbons only one triple helix high and shorter in length. In fact, our AFM-measured height is very nearly exactly that expected from a single triple helix. Nevertheless, the bundled fibrous structure is well supported by our X-ray diffraction data and the variances between cryo-TEM, stained TEM, AFM, SEM and X-ray diffraction can be attributed to necessary differences in sample preparation.

Conclusion

In this report, we describe the design and synthesis of a self-assembling peptide that forms a sticky-ended collagen-like triple helix. At sufficient concentrations, these triple helices elongate and bundle into a homogeneous population of nanofibres with triple helical packing similar to that of natural collagen, and these nanofibres interact to form high-quality hydrogels that are degraded at a similar rate to that of rat-tail collagen. This collagen-based system simultaneously demonstrates triple-helix, nanofibre and hydrogel formation and, as such, substantially recapitulates the multi-hierarchical self-assembly of natural collagen. As a result of collagen's major role in critical functions, such as tissue structure, repair and regeneration, we expect this peptide, and those based on its design, will play an important role in regenerative medicine and drug delivery.

Methods

Peptide synthesis. $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ was synthesized using standard 9-fluorenylmethoxycarbonyl chemistry for solid-phase peptide synthesis on an Advanced Chemtech Apex 396 multipeptide automated synthesizer at a scale of 0.15 mM on a glycine-preloaded Wang resin. Once synthesized, the peptide was purified on a Varian PrepStar220 high-performance liquid chromatograph (HPLC) using a preparative reverse-phase C-18 column and then dialysed against deionized water to remove salts. Once dialysed, the peptide was analysed by electron-spray ionization time of flight (TOF) mass spectrometry on a Bruker microTOF. The HPLC and mass spectrum are given in the Supplementary Information.

Sample preparation. All peptide concentrations were measured by weight. All samples were adjusted to pH 7 with sodium hydroxide prior to the addition of buffer and then annealed for 15 minutes at 85 °C. Finally, the samples were incubated at room temperature for at least 12 hours prior to characterization to ensure complete assembly. Time-course rheological studies are given in the Supplementary Information to support this time scale.

Circular dichroism. All spectra and thermal unfolding studies were performed on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature-control system. Quartz cells were used with path lengths of 0.01 cm and 0.1 cm depending on the peptide concentration and buffer. Spectra were collected from 190 to 250 nm. Melting experiments were performed from 5 to 85 °C, monitoring at 225 nm, and the first derivative of the thermal unfolding curve was taken to determine the melting temperature of the sample. The molar residual ellipticity (MRE) is calculated from the measured ellipticity using the equation:

$$[\theta] = \frac{\theta \times m}{C \times l \times n_r} \quad (1)$$

where θ is the ellipticity in millidegrees, m is the molecular weight in g mol^{-1}, c is the concentration in mg ml^{-1}, l is the path length of the cuvette in cm and n_r is the number of amino acids in the peptide. The spectrum for 1.0% by weight (Fig. 3a) is only shown from 250 to 205 nm because of the increase in background noise for samples of higher concentration at lower wavelengths.

Atomic force microscopy. Samples were prepared and dropped onto freshly cleaved mica spinning on a Headway Research photoresist spinner. The sample was quickly rinsed with deionized water for 4–5 seconds and then spun for an additional ten minutes. AFM images were collected on a Digital Instruments Nanoscope IIIa AFM in tapping mode under ambient conditions. Height profiles were obtained using Nanoscope software (20 measurements were taken per peptide concentration and buffer, then averaged and the standard deviation calculated).

Transmission electron microscopy. Samples for TEM were prepared on Quantifoil R1.2/1.3 holey carbon mesh on copper grids. For dry TEM, PTA was used to stain the TEM grids using negative-staining techniques. A 2.0% by weight solution of PTA was prepared and adjusted to pH 6 with sodium hydroxide. All stains were made fortnightly and syringe filtered prior to use. For dry TEM sample preparation, the peptide solution was added to the carbon side of a TEM grid, allowed to dry for one minute and then indirectly blotted with filter paper to remove excess solution. The grid was allowed to dry for five minutes before it was inverted onto an aliquot of PTA solution, where it remained for ten minutes. The grid was then placed on filter paper to dry overnight.

Vitreous ice TEM samples were prepared as follows. First, the TEM grids were glow discharged for one minute with a 5 mA discharge on a EMS 100 Glow Discharge Unit. The next stages of sample preparation were performed using a FEI Vitrobot type FP5350/60. The peptide solution (a diluted sample with a concentration of 0.25% by weight made from a 0.5% by weight sample) was added to the grid and immediately blotted for two seconds before being immersed in liquid ethane. Then the grid was transferred manually from liquid ethane to liquid nitrogen, in which it was stored until imaging. All TEM imaging was performed on a JEOL 2010 microscope (200 kV) and cryoimaging was taken at a temperature of −176 °C using low-dose conditions.

Scanning electron microscopy. 100 μl aliquots of each gel were placed in a 24-well plate. Gels were dehydrated in a series of ethanol–water solutions, progressing from 30% to 100% ethanol over the course of 24 hours. The dehydrated gels were critical point dried using an Electron Microscopy Sciences 850 critical point drier. They were then affixed to SEM pucks using conductive carbon tape. The pucks were sputter coated with 10 nm gold, rotated and then sputter coated with an additional 5 nm gold using a CRC-150 sputter coater. Samples were imaged using a FEI Quanta 400 ESEM at 20.00 kV.

Rheology. All rheological studies were performed on a TA AR-G2 rheometer. Strain and frequency experiments were carried out using 12 mm stainless-steel parallel-plate geometry with a gap size of 500 mm. Strain sweeps maintained a fixed frequency (1 rad s^{-1}) and a variable strain (0.01–200%). Frequency sweeps utilized a fixed strain (1%) and varying frequencies (0.1–200 rad s^{-1}).

X-ray fibre diffraction. A freshly annealed 1.0% by weight sample was dried by placing 10 μl droplets between two capillaries held in the centre of a custom magnet assembly, as described by Sunde and co-workers, over a period of several days[43]. A dried peptide pellet attached to the end of the capillary was used for data collection. Data were collected at 1.54 Å using a Rigaku RUH3R rotating anode X-ray generator with a Rigaku R-axis IV ++ detector. The detector was placed at a distance of 180.0 mm from the sample, which was cooled using a N_2 stream to 100 K. Diffraction patterns were acquired with exposure times that ranged from one to 40 minutes; the highest exposure time yielded the best pattern. The data were analysed using the Fit2D software package[54]. The position of the beam stop was calculated using the ring 11.5 Å and a median filter was applied to the data. Radial integration was carried out to produce a one-dimensional profile of the observed intensities as a function of D spacing (Å) and angular integration to generate a plot of the observed intensities as a function of D spacing (Å) and azimuthal angle.

Received 11 February 2011; accepted 25 July 2011; published online 28 August 2011

References

1. Ottani, V., Martini, D., Franchi, M., Ruggeri, A. & Raspanti, M. Hierarchical structures in fibrillar collagens. *Micron* **33,** 587–596 (2002).
2. Ottani, V., Raspanti, M. & Ruggeri, A. Collagen structure and functional implications. *Micron* **32,** 251–260 (2001).
3. Pinkas, D. M., Ding, S., Raines, R. T. & Barron, A. E. Tunable, post-translational hydroxylation of collagen domains in *Escherichia coli*. *ACS Chem. Biol.* **6,** 320–324 (2011).
4. Buechter, D. D. *et al.* Co-translational incorporation of *trans*-4-hydroxyproline into recombinant proteins in bacteria. *J. Biol. Chem.* **278,** 645–650 (2003).
5. Kohrer, C., Xie, L., Kellerer, S., Varshney, U. & Rajbhandary, U. L. Import of amber and ochre suppressor tRNAs into mammalian cells: a general approach to site-specific insertion of amino acid analogues into proteins. *Proc. Natl Acad. Sci. USA* **98,** 14310–14315 (2001).
6. Liu, D. R., Magliery, T. J., Pasternak, M. & Schultz, P. G. Engineering a tRNA and aminoacyl–tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo*. *Proc. Natl Acad. Sci. USA* **94,** 10092–10097 (1997).
7. Liu, D. R. & Schultz, P. G. Progress toward the evolution of an organism with an expanded genetic code. *Proc. Natl Acad. Sci. USA* **96,** 4780–4785 (1999).
8. Mendel, D., Cornish, V. W. & Schultz, P. G. Site-directed mutagenesis with an expanded genetic-code. *Annu. Rev. Biophys. Biomol. Struct.* **24,** 435–462 (1995).
9. Boudko, S. P. *et al.* Crystal structure of human type III collagen Gly991-Gly1032 cystine knot-containing peptide shows both 7/2 and 10/3 triple helical symmetries. *J. Biol. Chem.* **283,** 32580–32589 (2008).
10. Kar, K. *et al.* Aromatic interactions promote self-association of collagen triple-helical peptides to higher-order structures. *Biochemistry* **48,** 7959–7968 (2009).
11. Kramer, R. Z., Bella, J., Brodsky, B. & Berman, H. M. The crystal and molecular structure of a collagen-like peptide with a biologically relevant sequence. *J. Mol. Biol.* **311,** 131–147 (2001).

12. Krishna, O. D. & Kiick, K. L. Supramolecular assembly of electrostatically stabilized, hydroxyproline-lacking collagen-mimetic peptides. *Biomacromolecules* **10,** 2626–2631 (2009).
13. Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A. & Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **39,** 14960–14967 (2000).
14. Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability. *Biochemistry* **44,** 1414–1422 (2005).
15. Sakakibara, S. *et al.* Synthesis of $(Pro\text{-}Hyp\text{-}Gly)_n$ of defined molecular-weights – evidence for stabilization of collagen triple helix by hydroxyproline. *Biochim. Biophys. Acta* **303,** 198–202 (1973).
16. Shah, N. K., Ramshaw, J. A., Kirkpatrick, A., Shah, C. & Brodsky, B. A host–guest set of triple-helical peptides: stability of Gly-X-Y triplets containing common nonpolar residues. *Biochemistry* **35,** 10262–10268 (1996).
17. Venugopal, M. G., Ramshaw, J. A., Braswell, E., Zhu, D. & Brodsky, B. Electrostatic interactions in collagen-like triple-helical peptides. *Biochemistry* **33,** 7948–7956 (1994).
18. Fallas, J. A., Gauba, V. & Hartgerink, J. D. Solution structure of an ABC collagen heterotrimer reveals a single-register helix stabilized by electrostatic interactions. *J. Biol. Chem.* **284,** 26851–26859 (2009).
19. Gauba, V. & Hartgerink, J. D. Self-assembled heterotrimeric collagen triple helices directed through electrostatic interactions. *J. Am. Chem. Soc.* **129,** 2683–2690 (2007).
20. Gauba, V. & Hartgerink, J. D. Surprisingly high stability of collagen ABC heterotrimer: evaluation of side chain charge pairs. *J. Am. Chem. Soc.* **129,** 15034–15041 (2007).
21. Gauba, V. & Hartgerink, J. D. Synthetic collagen heterotrimers: structural mimics of wild type and mutant collagen type I. *J. Am. Chem. Soc.* **130,** 7509–7515 (2008).
22. Madhan, B., Xiao, J. X., Thiagarajan, G., Baum, J. & Brodsky, B. NMR monitoring of chain-specific stability in heterotrimeric collagen peptides. *J. Am. Chem. Soc.* **130,** 13520–13521 (2008).
23. Ottl, J. *et al.* Design and synthesis of heterotrimeric collagen peptides with a built-in cystine-knot. Models for collagen catabolism by matrix-metalloproteases. *FEBS Lett.* **398,** 31–36 (1996).
24. Russell, L. E., Fallas, J. A. & Hartgerink, J. D. Selective assembly of a high stability AAB collagen heterotrimer. *J. Am. Chem. Soc.* **132,** 3242–3243 (2010).
25. Cejas, M. A. *et al.* Thrombogenic collagen-mimetic peptides: self-assembly of triple helix-based fibrils driven by hydrophobic interactions. *Proc. Natl Acad. Sci. USA* **105,** 8513–8518 (2008).
26. Kar, K. *et al.* Self-association of collagen triple helix peptides into higher order structures. *J. Biol. Chem.* **281,** 33283–33290 (2006).
27. Kar, K., Wang, Y. H. & Brodsky, B. Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures. *Protein Sci.* **17,** 1086–1095 (2008).
28. Kotch, F. W. & Raines, R. T. Self-assembly of synthetic collagen triple helices. *Proc. Natl Acad. Sci. USA* **103,** 3028–3033 (2006).
29. Paramonov, S. E., Gauba, V. & Hartgerink, J. D. Synthesis of collagen-like peptide polymers by native chemical ligation. *Macromolecules* **38,** 7555–7561 (2005).
30. Yamazaki, C. M., Asada, S., Kitagawa, K. & Koide, T. Artificial collagen gels via self-assembly of *de novo* designed peptides. *Biopolymers* **90,** 816–823 (2008).
31. Skrzeszewska, P. J. *et al.* Physical gels of telechelic triblock copolymers with precisely defined junction multiplicity. *Soft Matter* **5,** 2057–2062 (2009).
32. Rele, S. *et al.* D-periodic collagen-mimetic microfibers. *J. Am. Chem. Soc.* **129,** 14780–14787 (2007).
33. Banwell, E. F. *et al.* Rational design and application of responsive alpha-helical peptide hydrogels. *Nature Mater.* **8,** 596–600 (2009).
34. Yang, Y. L., Leone, L. M. & Kaufman, L. J. Elastic moduli of collagen gels can be predicted from two-dimensional confocal microscopy. *Biophys. J.* **97,** 2051–2060 (2009).
35. Mi, K. *et al.* Influence of a self-assembling peptide, RADA16, compared with collagen I and Matrigel on the malignant phenotype of human breast cancer cells in 3D cultures and *in vivo*. *Macromol. Biosci.* **9,** 437–443 (2009).
36. Greenfield, M. A., Hoffman, J. R., de la Cruz, M. O. & Stupp, S. I. Tunable mechanics of peptide nanofiber gels. *Langmuir* **26,** 3641–3647 (2010).
37. Yokoi, H., Kinoshita, T. & Zhang, S. Dynamic reassembly of peptide RADA16 nanofiber scaffold. *Proc. Natl Acad. Sci. USA* **102,** 8414–8419 (2005).
38. Zhang, S. G. *et al.* Self-complementary oligopeptide matrices support mammalian-cell attachment. *Biomaterials* **16,** 1385–1393 (1995).
39. Lamm, M. S., Rajagopal, K., Schneider, J. P. & Pochan, D. J. Laminated morphology of nontwisting beta-sheet fibrils constructed via peptide self-assembly. *J. Am. Chem. Soc.* **127,** 16692–16700 (2005).
40. Ozbas, B., Kretsinger, J., Rajagopal, K., Schneider, J. P. & Pochan, D. J. Salt-triggered peptide folding and consequent self-assembly into hydrogels with tunable modulus. *Macromolecules* **37,** 7331–7337 (2004).
41. Aulisa, L., Dong, H. & Hartgerink, J. D. Self-assembly of multidomain peptides: sequence variation allows control over cross-linking and viscoelasticity. *Biomacromolecules* **10,** 2694–2698 (2009).
42. Ottl, J. *et al.* Recognition and catabolism of synthetic heterotrimeric collagen peptides by matrix metalloproteinases. *Chem. Biol.* **7,** 119–132 (2000).
43. Serpell, L. C., Fraser, P. E. & Sunde, M. X-ray fiber diffraction of amyloid fibrils. *Methods Enzymol.* **309,** 526–536 (1999).
44. Okuyama, K. Revisiting the molecular structure of collagen. *Connect. Tissue Res.* **49,** 299–310 (2008).
45. Pandya, M. J. *et al.* Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry* **39,** 8728–8734 (2000).
46. Woolfson, D. N. Building fibrous biomaterials from alpha-helical and collagen-like coiled-coil peptides. *Biopolymers* **94,** 118–127 (2010).
47. Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* **266,** 75–81 (1994).
48. Okuyama, K., Okuyama, K., Arnott, S., Takayanagi, M. & Kakudo, M. Crystal and molecular structure of a collagen-like polypeptide $(Pro\text{-}Pro\text{-}Gly)_{10}$. *J. Mol. Biol.* **152,** 427–443 (1981).
49. Pauling, L. & Corey, R. B. The structure of fibrous protein of the collagen-gelatin group. *Proc. Natl Acad. Sci. USA* **37,** 272–281 (1951).
50. Ramachandra, G. N. & Kartha, G. Structure of collagen. *Nature* **176,** 593–595 (1955).
51. Ramachandran, G. N. & Kartha, G. Structure of collagen. *Nature* **174,** 269–270 (1954).
52. Rich, A. & Crick, F. H. C. Molecular structure of collagen. *J. Mol. Biol.* **3,** 483–506 (1961).
53. Ruozi, B., Tosi, G., Leo, E. & Vandelli, M. A. Application of atomic force microscopy to characterize liposomes as drug and gene carriers. *Talanta* **73,** 12–22 (2007).
54. Wess, T. J., Hammersley, A., Wess, L. & Miller, A. Type-I collagen packing, conformation of the triclinic unit-cell. *J. Mol. Biol.* **248,** 487–493 (1995).

Acknowledgements

This work was funded in part by National Science Foundation CAREER Award (DMR-0645474), the Robert A. Welch Foundation (Grant No. C1557) and the Norman Hackerman Advanced Research Program of Texas.

Author contributions

L.E.R.O. designed and performed the experiments (except SEM and fibre diffraction) and co-wrote the manuscript. J.A.F. performed and analysed fibre-diffraction experiments. E.L.B. performed the SEM experiments. M.K.K. performed the collagenase experiments. J.D.H. supervised the research, evaluated all the data and co-wrote the manuscript.

Additional information

The authors declare no competing financial interests. Supplementary information accompanies this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.D.H.