Published in final edited form as:
*Biomacromolecules*. 2009 September 14; 10(9): 2626–2631. doi:10.1021/bm900551c.

# Supramolecular assembly of electrostatically stabilized, hydroxyproline-lacking collagen-mimetic peptides

**Ohm D. Krishna** and **Kristi L. Kiick***
Department of Materials Science and Engineering, University of Delaware, Newark, DE 19716

Delaware Biotechnology Institute, 15 Innovation Way, Newark, DE 19711

## Abstract

The mechanical and biological functions of the native collagens remain an inspiration in materials design, but widespread application of de novo collagens has been limited in part by the need for hydroxylated proline in the formation of stable triple helical structures. In order to address this continued need and to expand the potential for recombinant expression of functional, hydroxyproline-lacking collagen-mimetic peptides, we have designed a hydrophilic, non-repetitive, and thermally stable collagen-mimetic peptide via the incorporation of triple-helix-stabilizing charged triplets. The peptide sequence is also equipped with a type III-collagen-mimetic cystine knot at the C-terminus to facilitate covalent crosslinking of the triple helix via simple air oxidation. Circular dichroic (CD) studies of this collagen-mimetic peptide revealed a typical, thermally stable, collagen triple helix signature, with a weak positive maximum at 225 nm, and a triple helix melting temperature ($T_m$) of 35 °C and 43 °C for the reduced and oxidized forms respectively. The thermal behavior was confirmed via analysis by differential scanning calorimetry. Interestingly, this hydroxyproline-lacking, collagen-mimetic peptide also assembles into nanorods and microfibrillar structures as observed via transmission electron microscopy. The identification and demonstrated useful collagen-mimetic properties of this peptide suggests important opportunities in the recombinant design of new collagen-based biomaterials.

### Keywords

Collagen peptide; collagen triple helix; self-assembly; nanorods; microfibrils

## Introduction

Collagens are one of the most abundant fibrous proteins found in body tissues and organs, endowing structural integrity, mechanical strength and multiple biological functions. This has continued to motivate the design of peptides and recombinant polypeptides to closely mimic native collagens. The collagens consist of regularly repeated G-X-Y triplet amino acid sequences, with every third residue comprising glycine; X and Y are most often the imino acids proline (P) and hydroxyproline (O). This primary structure allows three such polyproline-II-type helical polypeptide chains to form a distinctive triple helical conformation.[1] The thermal stability of the collagen triple helices is greatly enhanced by the presence of the $(GPO)_n$ sequence, via origins ascribed to both stereoelectronic effects and potential hydrogen bonding

*To whom correspondence should be addressed. kiick@udel.edu, Tel.: +302-831-0201. Fax: +302-831-4545.

Supporting Information Available.
Data for peptide purification via HPLC, Molecular weight characterization via ESI-MS, SDS-PAGE methods and characterization and additional data for CD, DLS, and TEM are available free of charge at http://pubs.acs.org.

of the hydroxyproline residue.[1] For example, simple $(GPP)_{10}$ collagen-like peptides exhibit melting temperature ($T_m$) of ~28 °C, while $(GPO)_{10}$ exhibits a $T_m$ of ~60 °C.[1]

Of the over 25 human collagens, several types show additional lateral and end-to-end assembly, forming fibrillar rods via secondary forces determined by the sequence of polar, charged and hydrophobic amino acids. These fibrils are further crosslinked to achieve desired mechanical properties. Since tissue-extracted cross-linked collagen is always altered, precise biochemical studies and safe biomaterial application has been difficult. Thus the production of *de novo* collagen mimetic sequences (either chemically or biosynthetically) that can mimic the properties of native collagens will aid in our understanding of their higher order assembly behavior and expand the options for the design of collagen-mimetic materials. Essentially all reported fibril-forming collagen-mimetic peptides to date have exploited hydroxyproline-based, thermally stabilized sequences,[2–7] while $(GPP)_{10}$ peptides have been reported to *not* form such higher order assemblies.[8] In addition, the necessity of hydroxyproline for imparting thermal stability has made the synthesis of stable collagens in *E. coli* difficult, as the bacterial expression host lacks the enzyme responsible for the post-translational modification of proline to hydroxyproline. The development of simple bacterial expression protocols would broaden the utility and application of recombinant collagens, therefore there have been several efforts to stabilize collagen sequences lacking hydroxyproline.[9–14] Addressing these challenges, we report the design of a hydrophilic and non-repetitive collagen-mimetic peptide which is thermally stable under physiological conditions even in the absence of hydroxyproline, shows higher-order assembly behaviors to form fibrils, and contains a biologically relevant peptide motif. The three collagen-based chains of the designed peptide (CLP-Cys; Figure 1) are staggered with respect to the adjacent chain by one amino acid, and contain a C-terminal cystine-knot for oxidation via simple air oxidation to form covalently cross-linked homotrimers. The formation and stability of the triple helical structures adopted by the designed peptide in both oxidized and reduced form were analyzed via circular dichroic spectroscopy (CD) experiments and confirmed via differential scanning calorimetry (DSC). These results suggest the formation of a stable triple helix. Interestingly, this hydroxyproline-free, collagen-mimetic peptide also assembles into nanorods and microfibrillar structures as observed via a combination of dynamic light scattering (DLS) and transmission electron microscopy (TEM), despite the absence of hydroxyproline or other specific assembly-directing motifs in its design.

## Experimental Section

### Materials

Fmoc-protected amino acids, HBTU (O-Benzotriazole-N, N, N′, N′-tetramethyl-uronium-hexafluoro-phosphate) and the rink amide MBHA (4-methylbenzhydrylamine) resin for solid phase peptide synthesis were purchased from Novabiochem (San Diego, CA). HPLC grade DMF, acetonitrile, and TFA (trifluoroacetic acid) were purchased from Fisher Scientific (Fairlawn, NJ). Piperidine, 4-methylmorpholine and DTT (dithiothreitol) were purchased from Sigma-Aldrich.

### Peptide synthesis

The peptide was synthesized via automated solid-phase peptide synthesis (SPPS) procedures, using a PS3 peptide synthesizer (Protein Technologies Inc., Tuscon, AZ) and employing Fmoc-protected amino acids on a rink amide MBHA resin. The amino acid residues were activated for coupling with HBTU in the presence of 0.4 M methylmorpholine in DMF (N, N-dimethyl formamide). Fmoc was deprotected using 20% piperidine in DMF for 10 min. 1 hour coupling cycles were used for the first 14 residues. The remaining residues were attached in double coupling cycles. Cleavage of the peptide from the resin was performed in 94:1:2.5:2.5::trifluoroacetic acid (TFA):triisopropylsilane (TIS):water:1,2-ethanedithiol

(EDT) for 6 h. TFA was evaporated and the cleaved peptide was precipitated in ether. The peptide was extracted with water and lyophilized.

### Peptide purification via RP-HPLC and mass spectrometric analysis via ESI-MS

The crude peptide was purified via reverse-phase HPLC (Waters 2996) on a Waters Symmetry 300™, C-18 column (19×150 mM, pore size 5μm) at room temperature. Solvent A (water/0.1% TFA) and solvent B (acetonitrile/0.1% TFA) were employed as HPLC solvents in a programmed gradient of solvent B (acetonitrile/0.1% TFA) from 25–35% over 30 min; and the eluent was monitored at 214 nm. The purity of peptide obtained via RP-HPLC was confirmed via electrospray ionization mass spectrometry (ESI-MS) (AutospecQ, VG Analytical, Manchester, UK), m/z = 1610.6[$(M+2H)^{2+}$, calcd 1610.85], m/z = 1073.8[$(M+3H)^{3+}$, calcd 1074.23], m/z = 805.8[$(M+4H)^{4+}$, calcd 805.92], m/z = 644.8[$(M+5H)^{5+}$, calcd 644.94].

### Triple helix formation and characterization

Stock solutions of the RP-HPLC purified peptide (1 mM, gravimetric) were prepared by dissolving the peptide in 10 mM phosphate buffered saline (150 mM NaCl; pH 7.4). The stock solution was pre-equilibrated at 4 °C under inert atmosphere for at least 15h to allow triple helix formation. Then, the cysteine residues of the triple helices were either air-oxidized (by exposing to atmospheric $O_2$ at 4 °C for 15h) or reduced (by adding 10 mM of dithiothreitol (DTT) and then incubating overnight at 4 °C for 15h) prior to the collection of CD spectra. This procedure for oxidation and reduction of the cysteine residues on collagen peptides was based on similar and previously described protocols.[11,16,17]

To verify the formation and purity of the air-oxidized trimeric (covalently cross-linked) form of the CLP-Cys peptide, RP-HPLC, ESI-Q-TOF-MS analysis (Q-TOF Micro, Waters Associates) and SDS-PAGE (Figure S1 and S2) were performed on the air-oxidized trimer. RP-HPLC performed on the air-oxidized trimer showed a single peak. The trimeric form of the CLP-Cys upon air oxidation was clearly identified via ESI-Q-TOF-MS analysis of the multiple charged states of the trimeric species with m/z = 1610.28 ($[M+6H]^{6+}$, calcd 1609.85), m/z = 1380.1 ($[M+7H]^{7+}$, calcd 1379.87), m/z = 1207.73 ($[M+8H]^{8+}$, calcd 1207.3875). The mass separation of the isotope pattern of these peaks unambiguously confirmed the charged state of the trimeric species. However, the presence of low intensity peaks at m/z = 1288.25 ($[M+5H]^{5+}$) and m/z = 1073.89 ($[M+3H]^{3+}$) corresponded respectively to the fifth charged state of the dimeric species and the triply charged state of the self-oxidized monomer as confirmed via the mass separation of the isotope pattern of these peaks. Densitometry analysis of the band intensities (TotalLab TL100 software, Nonlinear USA Inc.) in the SDS-PAGE of the CLP-Cys trimer indicated approximately 90% trimeric species. These observations are consistent with previously reported studies, in which it has been difficult to obtain 100% yields of correctly oxidized trimeric peptide with three interchain disulfide bonds, either via GSH/GSSG catalysis in the presence of glutathione or via simple air oxidation.[11,16–19]

### Circular dichroic spectroscopy

Characterization of the secondary structure of the peptide was conducted via circular dichroic spectroscopy (Jasco 810 circular dichroism spectropolarimeter, Jasco, Inc., Easton, MD, USA). CD spectra were recorded using quartz cells with a 0.1 cm optical path length. The spectra reported are averages of three scans. The scanning rate was 50 nm/min, with a response time of 4 seconds. The wavelength scans were obtained from 197 to 260 nm and were recorded every 1 nm with a 1 nm bandwidth. Data points for the thermal unfolding experiments were recorded at 225 nm, at different scanning rates (from 1 °C/min – 10 °C/hr) with an equilibration time of 3 min at each recorded temperature. The van't Hoff enthalpy ($\Delta H_{vH}$) was estimated by

fitting the entire transition curve obtained from the CD measurements with a two-state first order model as described previously[20] and as detailed in the supporting information.

### Differential scanning calorimetry

The temperature dependence of partial heat capacity was analyzed via the use of a VP-DSC microcalorimeter (Microcal, Northampton, USA). The instrument is equipped with reference and sample cells of approximate volume of 500 μl. First, both the sample and reference cells were filled with properly degassed 10 mM PBS (150 mM NaCl) and scans were conducted from 5 – 80 °C at 40 °C/h until at least three consecutive buffer-buffer reference scans overlapped. Then the PBS buffer was carefully removed from the sample cell and oxidized peptide sample (500 μM, gravimetric) was loaded, and the scan was continued. The buffer-buffer reference scan was subtracted from the DSC scan with the sample and normalized with respect to the concentration to generate the data for excess heat capacity ($C_p$) as a function of temperature using the Origin software package provided with the VP-DSC microcalorimeter instrument. The calorimetric enthalpy $\Delta H_{cal}$ was calculated from the area under the curve of the concentration-normalized and baseline-subtracted DSC endotherm.

### Particle size analysis via DLS

The average particle size in solutions of the collagen-mimetic peptide was estimated using dynamic light scattering on a ZetaSizer Nano Series (Nano ZS, Malvern Instruments, UK) at a scattering angle of 90° and laser wavelength of 532 nm and data fitting via the cumulant method. Plots of the change in autocorrelation function as a function of delay time represented by experimental data points and the cumulant fit are provided in Figure S7. The peptide samples (1.6 wt %) were prepared in 10 mM phosphate buffered saline (150 mM NaCl; pH 7.4) and incubated at 4 °C to allow formation of the collagen triple helix and were then air oxidized and incubated for 1 week at 4 °C. Particle size measurements were conducted on both unfiltered and filtered (filtered with 0.45 μm filter prior to particle size analysis) samples of the oxidized peptide. 500 μl of the samples were added to prescribed low volume disposable polystyrene cuvettes. The measurements were taken at a pre-equilibrated temperature of 25 °C. The reported data represent an average of 3 scans.

### Transmission electron microscopy

TEM images of peptide samples were obtained on a Tecnai G2 12 TEM (FEI Company, Hillsboro, OR) at an acceleration voltage of 120 keV. For each sample, 3 μL of 1.6 wt% oxidized peptide solution (the same as that used for particle size analysis) was placed on a carbon-coated copper grid and blotted with a filter paper after approximately 2 min at room temperature. The samples were stained with 2 μL of 1 w/v % uranyl acetate staining solution for 30 s before blotting with a filter paper. After blotting the excess staining solution, the TEM grids were dried at room temperature overnight prior to imaging. Images were obtained from TEM at room temperature.

## Results and Discussion

### Design and characterization of triple-helix forming peptide

We designed the sequence $(GPP)_3$ GPR GEK GER GPR $(GPP)_3$ GPCCG (hereafter designated as CLP-Cys) based on the following principles. The inclusion of triple helix-promoting $(GPP)_3$ triplets can act as efficient folding nuclei for initiating triple helix formation.[21] In addition, the GPR triplet has been shown to confer stability similar to GPO in collagen triple helix host-guest peptides.[22] Furthermore, sequence motifs such as G_KGE_ have been shown to impart stabilizing electrostatic interactions between residues in adjacent tripeptides.[23] Such charged motifs are also found in certain bacterial collagens, which are thermally stable even

in the absence of hydroxyproline.[13] A GER triplet was therefore included in the sequence because it occurs abundantly in collagens I, II, III, V and XI and is an essential triplet for binding to various integrins expressed on cell surfaces and blood platelets.[16,24] Finally, the reported peptide sequence also utilizes a type III collagen mimetic cystine knot at the C terminus, known to act as nucleating and stabilizing motif by covalently crosslinking (via oxidation) the three peptide chains.[9,16,17,25]

The peptide was synthesized via automated solid-phase peptide synthesis procedures employing standard Fmoc chemistry, purified via RP-HPLC, and characterized via ESI-MS and SDS-PAGE (Figure S1, S2). The HPLC- purified peptides were of >95% purity. The formation and purity of the air-oxidized trimeric (covalently cross-linked) form of the CLP-Cys peptide was also verified via RP-HPLC, ESI-MS and SDS-PAGE (see experimental section). The single peak for the trimer in the RP-HPLC (Figure S1); the mass separation of the isotope pattern for the multiple charged states of the trimer observed via ESI-Q-TOF-MS (Figure S1); and the presence of a single band for the oxidized peptide with approximately thrice the apparent molecular weight of the reduced form in SDS-PAGE (Figure S2) confirmed the presence of the trimer. Densitometry analysis of the band intensities in the SDS-PAGE of the trimer indicated a good yield of approximately 90% of the trimeric species, which is consistent with the low intensity of non-trimeric species observed in the ESI-Q-TOF-MS (Figure S1). As mentioned above, these results are also consistent with previously reported studies, in which it has been difficult to obtain 100% yields of correctly oxidized trimeric peptide with three interchain disulfide bonds, either via GSH/GSSG catalysis in the presence of glutathione or via simple air oxidation.[11,16–19]

## Conformational behavior and thermodynamic analysis via CD and DSC

The secondary structure of the CLP-Cys peptide was investigated via far-ultraviolet CD spectroscopy in 10 mM phosphate buffered saline (150 mM NaCl; pH 7.4). CD spectra for the peptide in both oxidized (Figure 1a) and reduced forms (essentially identical to the spectra of the oxidized peptide, Figure S3), showed a positive peak at 225 nm and a minimum just below 200 nm, indicative of the collagen triple helical structure.[1] In addition, the isodichroic point near 213 nm suggests a two-state equilibrium between the triple helical and random-coil state.[26]

The thermal stability of the triple helical conformation was assessed by monitoring the decrease in mean residue ellipticity at 225 nm ($[\Theta]_{225}$) as a function of increasing temperature (Figure 2b). This yields a typical sigmoidal curve indicating cooperativity in unfolding of the triple helix to random coil, and a melting temperature ($T_m$) of approximately 35 °C and 43 °C for the reduced and oxidized forms of the peptide, respectively. The increased stability of these sequences compared to simple $(GPP)_{10}$ peptides[1] substantiates the expected stabilizing effect of the charged triplets. These results also indicate that the covalent link between the three peptide chains leads to significant increase in the thermal stability which is consistent with previous reports.[11] It is also worthwhile to note that although the CD spectral features for both the oxidized and reduced forms in full wavelength scans are similar (Figures 2a and S3), the positive peak at 225 nm exhibited a higher mean residue ellipticity at all temperatures for the oxidized form as compared to the reduced form (Figure 2b), which is also consistent with previous reports.[11]

CD studies were conducted to determine the effect of concentration (50, 100, 200, and 500 μM) on the $T_m$ of only the oxidized form of CLP-Cys given its increased thermal stability, and results are given as normalized mean residue ellipticity at 225 nm ($[\Theta]_{225}$) as a function of temperature in Figure 3. These results yield nearly identical $T_m$ values (~ 43 °C) that were independent of the concentration. Hence the triple helix melting can be described as a pseudomonomolecular or first-order unfolding process[20] in which the covalently linked trimer

remains in equilibrium between its unfolded and folded state, and in which the presence of aggregates of multiple sizes is not suggested under these conditions. These results are consistent with the data from SDS-PAGE, RP-HPLC and ESI-Q-TOF-MS on the trimers as described above, which suggested the presence (>90%) of primarily the trimeric form of the peptide. The more commonly studied un-crosslinked collagen model peptides such as $(GPP)_{10}$ or $(GPO)_{10}$ exhibit trimolecular melting transitions, such that the three unfolded peptide chains remain in equilibrium with its associated trimer. Previous work by Henkel et al.,[20] based on a collagen III model peptide with a C-terminal crosslinking trimeric branch and an N-terminus equipped with 7 repeats of GPO triplets, exhibited a concentration independent $T_m$ near 44°C. The fact that the hydroxyproline-lacking crosslinked CLP-Cys peptide sequence exhibited a similar melting temperature substantiates the role of favorable electrostatic interactions in stabilizing the CLP-Cys triple helices. The importance of the triple-helix stabilizing charged triplets in stabilizing the triple helical structures of these peptides, over that provided by the disulfide crosslinking, is suggested by the similarity of the observed $T_m$ values for the reduced and oxidized forms of the CLP-Cys. In addition, in our studies with the electrostatically stabilized peptide sequence $H_3G_3(GPP)_3GPRGEKGERGPR(GPP)_3)G_3H_3$ lacking a cystine knot, a $T_m$ value of 36 °C was observed in PBS,[27] further substantiating the role of the electrostatic stabilization over that provided by the cystine knot. The utility of electrostatic stabilization is also suggested by considering the results of Boudko et al.,[11] who reported a concentration-independent $T_m$ of only 14 °C for a recombinant, C-terminal disulfide cross-linked collagen III fragment lacking any hydroxyproline residues, as well as by the low $T_m$ values observed for other non-repetitive collagen-mimetic peptide designs containing cystine knots but lacking electrostatic stabilization.[28] Electrostatically stabilized formation of hydroxyproline-based heterotrimeric collagen triple helices has also been recently demonstrated in the work of Gauba and Hartgerink.[29–32] It is also apparent in Figure 3 that significant triple-helical character (~75%) for the oxidized peptide is retained at 37 °C, while the reduced form of the peptide is significantly less stable under these conditions (Figures 2b and S4). The difference in stability of the two forms of the peptide from above (for the oxidized form) to below (for the reduced form) physiological temperature could be utilized as an assembly trigger in vivo where reducing conditions exist. We also investigated the effect of different scanning rate (1 °C/min – 10 °C/hr, Figure S5) which resulted in similar value of $T_m$ for the oxidized peptide at all scanning rates and provided a reliable basis for employing a heating rate of 40 °C/hr in DSC (see below).

Owing to the stability of the oxidized form of the CLP-Cys at physiological temperature and its intended use in application, we characterized the thermodynamic stability of the oxidized form of the CLP-Cys via both CD and DSC. Thermodynamic parameters were extracted from the CD melting curves ($T_m$ and the van't Hoff enthalpy $\Delta H_{vH}$), by fitting the transition curve with a two-state first-order model. The model provides excellent fits ($r^2 > 0.999$) to the data for the transition curves at all concentrations and yields values for $T_m$ of ~ 44 °C and for $\Delta H_{vH}$ of −48.8 ± 4.8 kcal/mole (where mole refers to mole of triple helix) for the oxidized form of the peptide. These CD-extracted thermodynamic parameters are listed in Table 1.

The thermodynamic results obtained from CD were further confirmed via differential scanning calorimetry (DSC). Figure 4 shows the concentration-normalized and baseline-subtracted DSC scan of the oxidized peptide at a concentration of 500 μM in PBS buffer at a heating rate of 40 °C/hr. The DSC trace exhibits both a $T_m$ (45.2 °C±0.3) and a transition breadth that match the CD transition curves. The calorimetric enthalpy $\Delta H_{cal}$ calculated from the area under the curve of the DSC trace is −43.68 ± 1.6 kcal/mole of oxidized triple helix, which is comparable to previous reports of the thermal stabilities of collagen types I, III and charged bacterial collagens.[1,33,34] The similar values of van't Hoff enthalpy $\Delta H_{vH}$ calculated from the CD melting transition curves and the calorimetric enthalpy $\Delta H_{cal}$ obtained from the heat capacity curves of DSC, yields a cooperative ratio of $\Delta H_{vH}/\Delta H_{cal} = 1.1$ and suggests an approximate

two-state transition mechanism for the unfolding of the CLP-Cys triple helices.[33,35] The observed 10% error associated with $\Delta H_{vH}$ may explain the slight difference between the $\Delta H_{vH}$ and the $\Delta H_{cal}$ values. Nevertheless, these results taken together – the concentration independence of $T_m$ determined via CD and the similar values of $\Delta H_{vH}$ and the $\Delta H_{cal}$ determined via CD and DSC – validates a first-order two-state transition mechanism for unfolding of the oxidized CLP-Cys. The similarities of these results with previous studies of collagen-mimetic peptides suggest their biomimetic folding and stability as well as their potential for use in materials assembly.

## Investigation of supramolecular assembly via DLS and TEM

Collagen in vivo is organized supramolecularly at multiple length scales in highly concentrated solution, providing the basis for its diverse mechanical and functional roles in different tissues.[2] To mimic this assembly process, most previous reports on the design of fibril-forming collagen-mimetic peptides have been based on the hydroxyproline (O)-based thermally stabilized repeating GPO triplets.[2–7] In addition, many of these studies employ multistep chemical transformations for obtaining pre-staggered chains linked via cystines,[3,4] non-natural moieties,[5] and metal chelation,[7] thus deviating from the natural collagen sequence. Unfortunately, $(GPP)_{10}$ peptides have failed to exhibit fibril formation.[8] Based on these challenges, coupled with the observations of collagen-like triple helix formation and stability in the CLP-Cys, we investigated via DLS and TEM the competence of our hydroxyproline-free, collagen-mimetic peptide sequence (air-oxidized CLP-Cys trimer) for higher-order assembly.

Dynamic light scattering studies were conducted to analyze the sizes and homogeneity of any supramolecular assemblies formed by the air-oxidized CLP-Cys trimer. An unfiltered sample (1.6 wt % samples were prepared in 10 mM phosphate buffered saline as described in the experimental section), analyzed in order to observe all species present at room temperature, showed three species with diameters of 5.2 nm, 226 nm, and 5.6 μm as shown in Figure 5. A filtered sample (prepared by filtration through a 0.45 μm pore-size filter) was also analyzed and provided similar results with the exception of the removal of the micron-size species (Figure S6). The presence of particles with sizes greater than 10 nm indicates the formation of larger-order assemblies, while the small species is attributed to the soluble trimer.[3,5,7]

To investigate the morphology of the assembled structures, these samples were further characterized via TEM. TEM samples were prepared on a carbon-coated copper grid and stained via uranyl acetate. Figure 6 presents TEM images for these samples, which revealed the presence of supramolecular assemblies existing as nanorods (Figure 6a, b, Figure S8a–d) throughout the sample and micron-scale fibrils (Figure 5c, d, Figure S8e–f) at lower abundance. Control samples prepared on the carbon-coated copper grid with just the buffer and uranyl acetate stain showed no regular structures in contrast to those observed for the samples containing the peptide. It is noteworthy that the TEM images appear to possess a large positive stain (appearing dark), likely due to the presence of abundant charged residues in the sequence. This observation is consistent with previous observations that uranyl ions bind to both positively and negatively charged amino acids in collagen type I, generating a positive stain.[36] Similarly, previous studies on type II and type III collagens have indicated that negatively stained areas were located at the clusters of hydroxyproline and large hydrophobic amino acid residues while positively stained regions correlated well with the distribution of charged polar amino acids along the collagen molecule.[37]

The observation of assembled structures of these morphologies and lengthscales is consistent with previous studies of fibril-forming collagen-mimetic peptides, which also report assembly from the nano- to the micro-scale. Cejas et al.[5] reported micron-long fibrils with average diameters of 0.26 μm, while Rele et al.[6] reported 3–4 μm long fibrils with diameters of 12–15

nm; after annealing for several hours, fibrils of hundreds of nanometers in length and tens of nanometers in diameter were observed, while longer annealing periods resulted in micron-length fibrils with ~70 nm diameter. Kotch et al.[4] reported one-dimensional fibrillar structures of length varying from 30 nm to >400 nm with diameters of 0.5–1.0 nm, and Przybyla and Chmielewski reported fibrils with lengths on the order of 3–5 μm with extensive branching. In the present case of the assembly of CLP-Cys, analysis of the DLS data and the observed dimensions of the nanorods and microfibrils suggest that the assembled structures evolve from those with dimensions smaller than 60 nm in length and 5.5 nm in width to those greater than 6 μm in length and 130 nm in width. Considering the length (*ca*. 10 nm) and width (*ca*. 1.5 nm) of the isolated collagen triple helical peptide,[1] both linear and lateral assembly of these peptides is indicated.

As these sequences do not carry staggered ends,[4] electrostatic directionality,[6] or chelation sites[7] to promote fibril formation as other reported fibril-forming collagen-like peptides, their assembly is likely mediated by liberation of bound water as postulated for naturally-based collagen peptide sequences.[8] We hypothesize that a synergistic effect of the hydrophobic GPP clusters at both the ends, coupled with the electrostatic interactions mediated via the charged residues, guides the linear and lateral assembly; the presence of the cystine knot is not suggested to be critical to assembly, as the sequence $H_3G_3(GPP)_3GPRGEKGERGPR(GPP)_3)G_3H_3$ mentioned above also demonstrates higher-order assembled structures.[27] The proposed hypothesis is reinforced by a recent study on crystal structure of a human type III collagen peptide fragment which suggested the role of numerous direct or water-mediated ionic and polar contacts between triple helices for staggered helical packing and fibril formation.[18]

In the present study we investigated the non-repetitive CLP-Cys owing to a similar lack of repetition found in bacterial collagens that is expected to facilitate expression of such sequences in *E. coli*; our results indicate the potential for these sequences in their intended design. Comparative studies of various repetitive peptide sequences comprising individual triplets and combinations of triplets such as GPR, GEK, GER and GEKGER, are planned and will yield a better understanding of the contribution of these individual triplets toward their thermal stability and assembly behavior; as well as their role in the biological properties of the assembled materials. The observed stability and assembly behavior of these simple, hydroxylproline-free sequences suggests the potential for their recombinant production in bacterial hosts, yielding collagen-like polypeptides with improved thermal stability and biological activity over those currently available.

## Conclusions

We present a new collagen-mimetic peptide that exploits a set of triple helix-stabilizing triplets and minimizes the repetitiveness of the sequence, and have demonstrated that this sequence forms stable triple helices as determined via CD and DSC studies. The peptide was also shown to form higher order assemblies as determined via DLS and TEM. The collagen mimetic sequence investigated in the present study highlights the role of charged residues for thermal stability; evaluation of the assembly of these sequences also suggests that rationally designed hydroxyproline-free collagen mimetic sequences can elicit higher order assembly behavior. We speculate that these sequences may be useful for the recombinant expression in *E. coli* of functional collagens, which would considerably expand the use of such sequences in biomaterials design.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported in part by the National Science Foundation (DMR0239744) and the National Center for Research Resources, a component of the National Institutes of Health (1-P20-RR017716 and 5-P20-RR015588 for instrument facilities). Its contents are solely the responsibility of the authors. We thank Dr. Chaoying Ni & Frank Kriss (W.M. Keck Electron Microscopy Facility, UD) for assistance with TEM analysis. We thank Daphne A. Salick (Department of Chemistry and Biochemistry, UD) for assistance with DSC. Amit K. Jha and Hao Zhu (Department of Materials Science & Engineering, and Physics and Astronomy, UD) are thanked for assistance with particle size analysis. We thank John Dykins (Mass Spectrometry Facility, UD) for assistance with the ESI-Q-TOF-MS analysis.

## References

1. Engel J, Bächinger HP. Top Curr Chem 2005;247:7–33.
2. Martin R, Waldmann L, Kaplan DL. Biopolymers 2003;70:435–444. [PubMed: 14648755]
3. Koide T, Homma DL, Asada S, Kitagawa K. Bioorg Med Chem Lett 2005;15:5230–5233. [PubMed: 16185864]
4. Kotch FW, Raines RT. Proc Natl Acad Sci USA 2006;103:3028–3033. [PubMed: 16488977]
5. Cejas MA, Kinney WA, Chen C, Leo GC, Tounge BA, Vinter JG, Joshi PP, Maryanoff BE. J Am Chem Soc 2007;129:2202–2203. [PubMed: 17269769]
6. Rele S, Song YH, Apkarian RP, Qu Z, Conticello VP, Chaikof EL. J Am Chem Soc 2007;129:14780–14787. [PubMed: 17985903]
7. Przybyla DE, Chmielewski J. J Am Chem Soc 2008;130:12610–12611. [PubMed: 18763780]
8. Kar K, Amin P, Bryan MA, Persikov AV, Mohs A, Wang YH, Brodsky B. J Biol Chem 2006;281:33283–33290. [PubMed: 16963782]
9. Frank S, Boudko S, Mizuno K, Schulthess T, Engel J, Bächinger HP. J Biol Chem 2003;278:7747–7750. [PubMed: 12540847]
10. Frank S, Kammerer RA, Mechling D, Schulthess T, Landwehr R, Bann J, Guo Y, Lustig A, Bächinger HP, Engel J. J Mol Biol 2001;308:1081–1089. [PubMed: 11352592]
11. Boudko SP, Engel J. J Mol Biol 2004;335:1289–1297. [PubMed: 14729344]
12. Bai HY, Xu K, Xu YJ, Matsui H. Angew Chem Int Ed Engl 2007;46:3319–3322. [PubMed: 17352428]
13. Xu Y, Keene DR, Bujnicki JM, Hook M, Lukomski S. J Biol Chem 2002;277:27312–27318. [PubMed: 11976327]
14. Yao JM, Yanagisawa S, Asakura T. J Biochem (Tokyo) 2004;136:643–649. [PubMed: 15632304]
15. Bruckner P, Bächinger HP, Timpl R, Engel J. Eur J Biochem 1978;90:595–603. [PubMed: 710449]
16. Mechling DE, Bächinger HP. J Biol Chem 2000;275:14532–14536. [PubMed: 10799537]
17. Barth D, Kyrieleis O, Frank S, Renner C, Moroder L. Chem Eur J 2003;9:3703–3714.
18. Boudko SP, Engel J, Okuyama K, Mizuno K, Bächinger HP, Schumacher MA. J Biol Chem 2008;283:32580–32589. [PubMed: 18805790]
19. Boudko SP, Engel J, Bächinger HP. J Biol Chem 2008;283:34345–34351. [PubMed: 18845531]
20. Henkel W, Vogl T, Echner H, Voelter W, Urbanke C, Schleuder D, Rauterberg J. Biochemistry 1999;38:13610–13622. [PubMed: 10521268]
21. Germann HP, Heidemann E. Biopolymers 1988;27:157–163. [PubMed: 3342275]
22. Yang W, Chan VC, Kirkpatrick A, Ramshaw JAM, Brodsky B. J Biol Chem 1997;272:28837–28840. [PubMed: 9360948]
23. Persikov AV, Ramshaw JAM, Kirkpatrick A, Brodsky B. Biochemistry 2005;44:1414–1422. [PubMed: 15683226]
24. Reyes CD, Garcia AJ. J Biomed Mater Res A 2003;65A:511–523. [PubMed: 12761842]
25. Barth D, Musiol H, Schutt M, Fiori S, Milbradt AG, Renner C, Moroder L. Chem Eur J 2003;9:3692–3702.
26. Babu IR, Ganesh KN. J Am Chem Soc 2001;123:2079–2080. [PubMed: 11456840]
27. Krishna OD, Kiick KL. PMSE Prepr 2008;99:475–476.
28. Bachmann A, Kiefhaber T, Boudko S, Engel J, Bächinger HP. Proc Natl Acad Sci USA 2005;102:13897–13902. [PubMed: 16172389]

29. Gauba V, Hartgerink JD. J Am Chem Soc 2007;129:15034–15041. [PubMed: 17988128]
30. Gauba V, Hartgerink JD. J Am Chem Soc 2007;129:2683–2690. [PubMed: 17295489]
31. Gauba V, Hartgerink JD. J Am Chem Soc 2008;130:7509–7515. [PubMed: 18481852]
32. Brodsky B, Baum J. Nature 2008;453:998–999. [PubMed: 18563144]
33. Engel J, Bächinger HP. Matrix Biol 2000;19:235–244. [PubMed: 10936448]
34. Mohs A, Silva T, Yoshida T, Amin R, Lukomski S, Inouye M, Brodsky B. J Biol Chem 2007;282:29757–29765. [PubMed: 17693404]
35. Persikov AV, Xu YJ, Brodsky B. Protein Sci 2004;13:893–902. [PubMed: 15010541]
36. Bender E, Silver FH, Hayashi K, Trelstad RL. J Biol Chem 1982;257:9653–9657. [PubMed: 6179935]
37. Kobayashi K, Hashimoto Y, Hayakawa T, Hoshino T. J Ultrastruct Mol Struct Res 1988;100:255–262. [PubMed: 2468721]

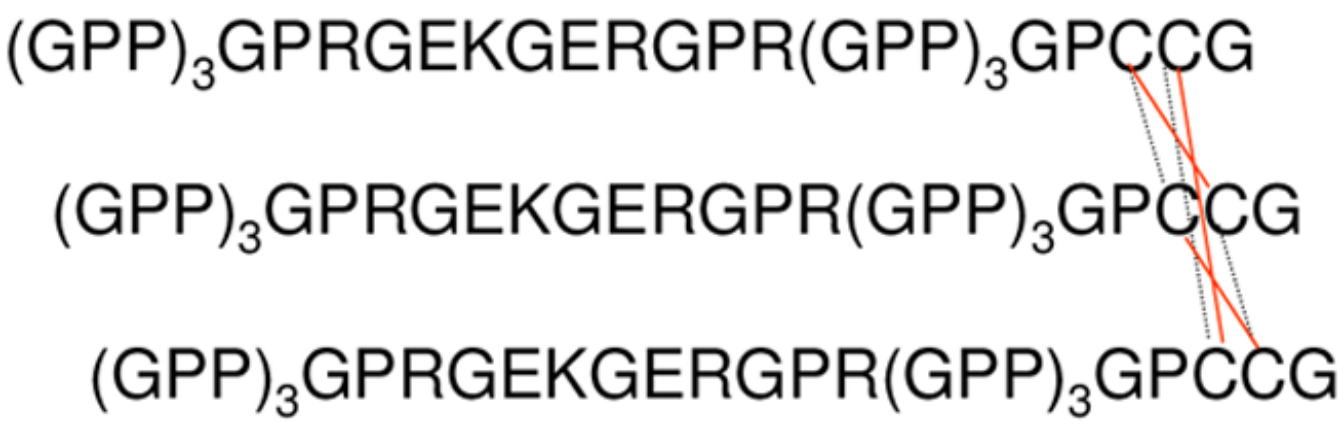
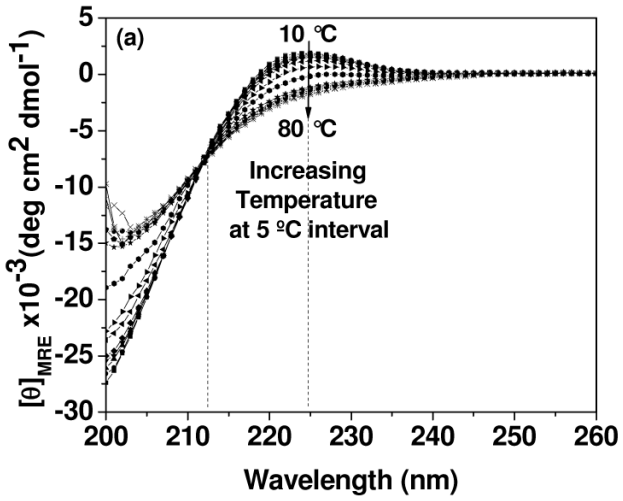



**Figure 1.**
Peptide sequence of the homotrimer (CLP-Cys) showing the two most plausible cystine knot connectivities (dotted black and solid red lines) formed via disulfide linkages and based on the model proposed by Bruckner et al.[15] and Barth et al..[17]



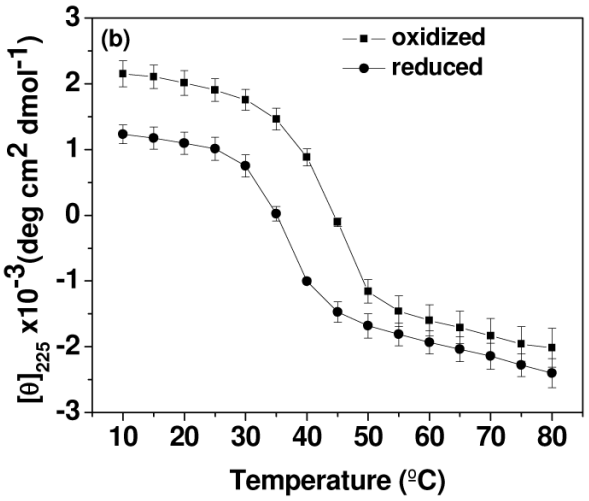




**Figure 2.**

(a) CD spectra showing the wavelength scans for the peptide (100 μM, oxidized) in 10 mM PBS (150 mM NaCl, pH 7.4). (b) Thermal transition curves, plotted as ($[\Theta]_{225}$) versus temperature for the oxidized and reduced peptide (100 μM) reported as mean±standard error for three separate experiments.

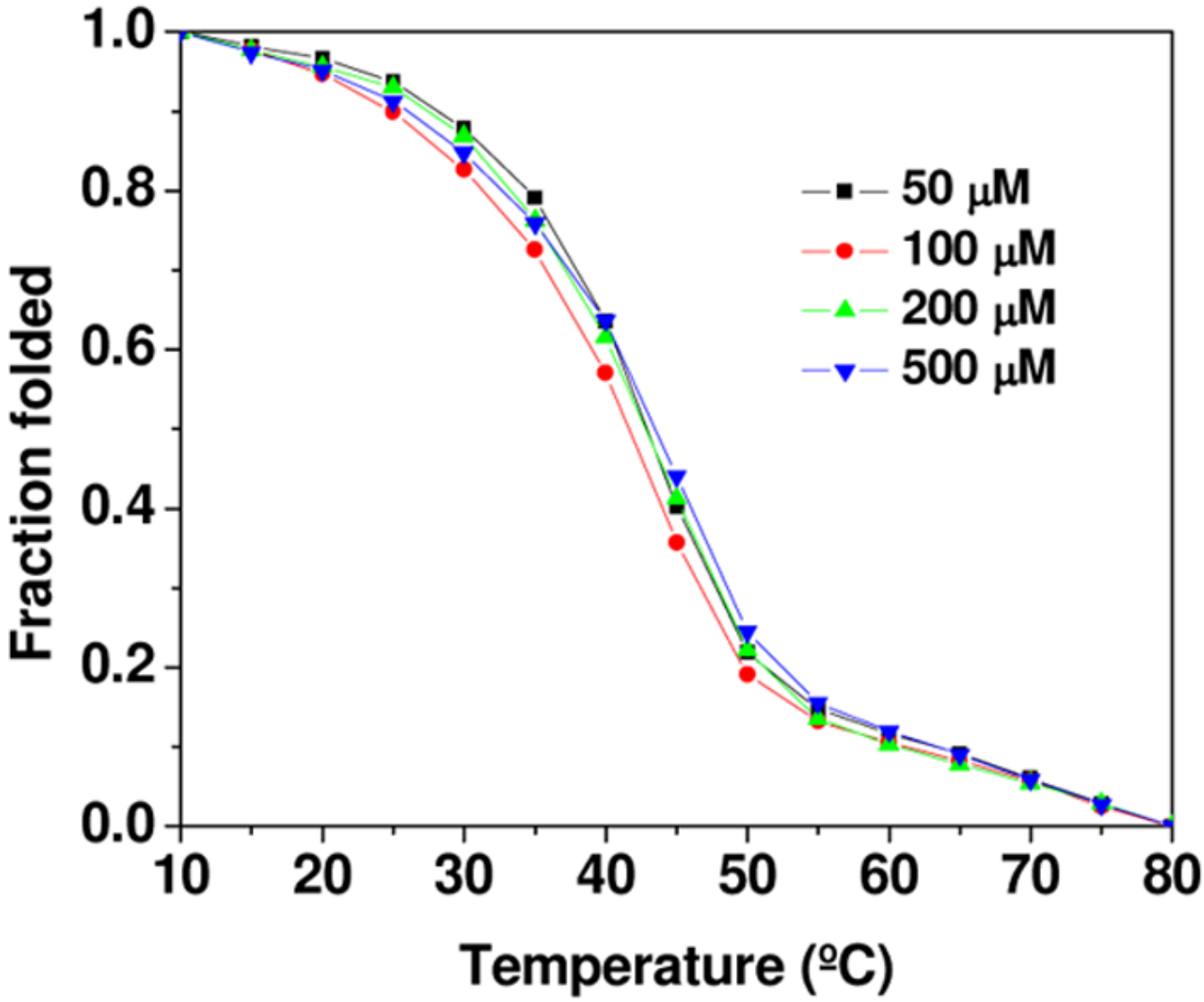



**Figure 3.**
The thermal transition curves for the oxidized peptide at different concentrations (50 – 500 μM) in 10 mM PBS (150 mM NaCl, pH 7.4).

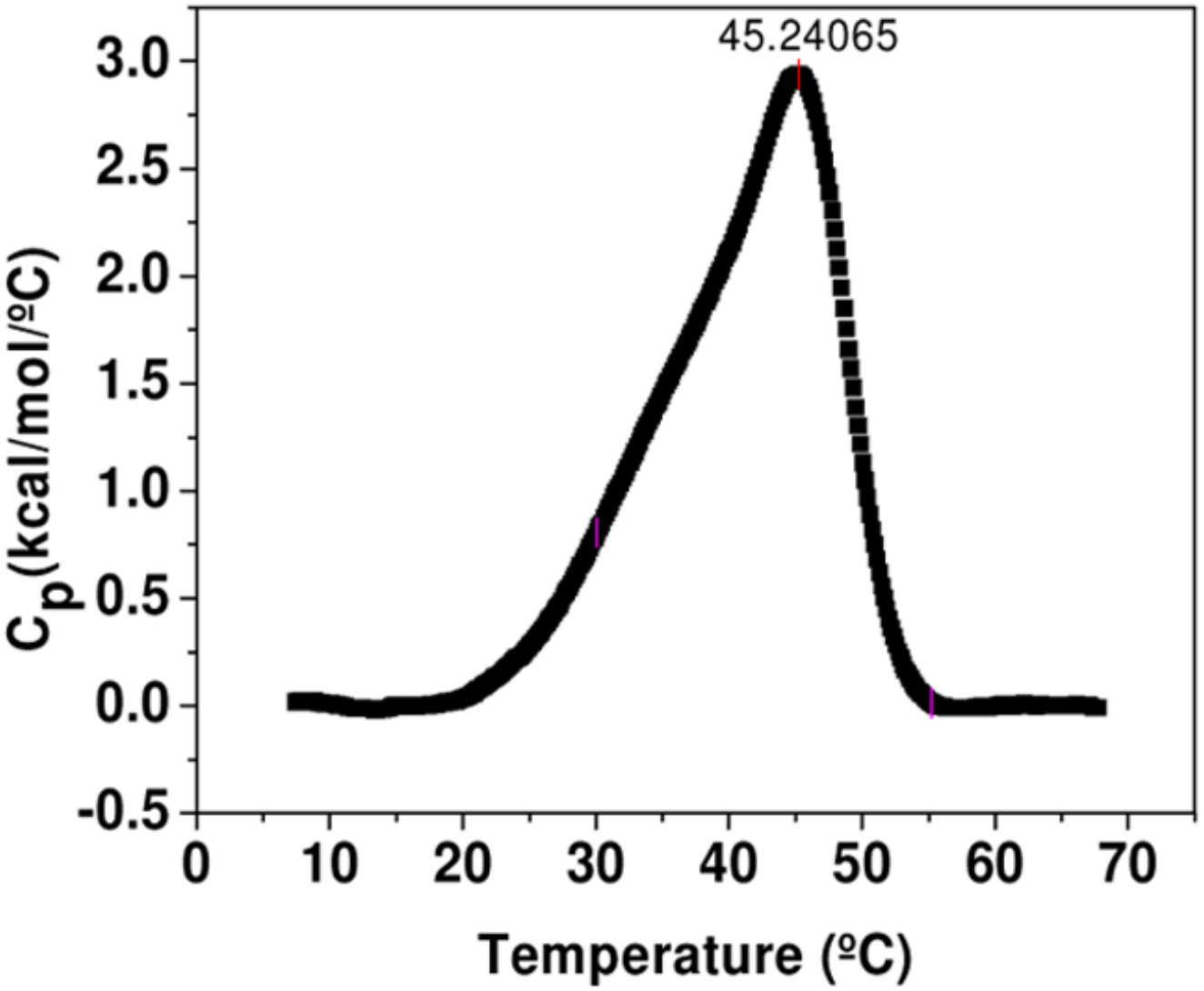



**Figure 4.**
Concentration-normalized and baseline-subtracted DSC scan of the oxidized peptide at a concentration of 500 μM in 10 mM PBS (150 mM NaCl, pH 7.4) at a heating rate of 40 °C/hr.

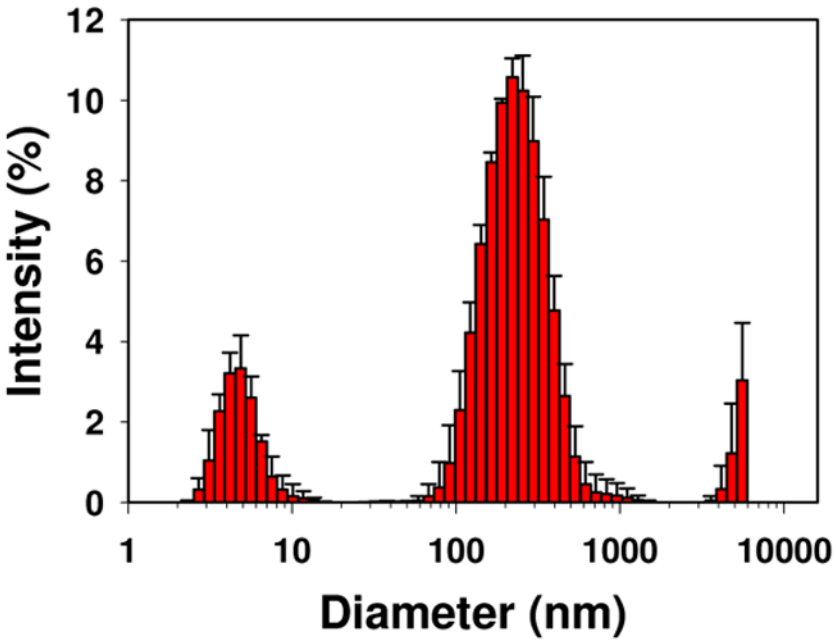



**Figure 5.**
Particle size analysis of the oxidized peptides (1.6 wt %) in 10 mM PBS (150 mM NaCl, pH 7.4) via DLS. Results presented are average of three experiments with standard deviation.

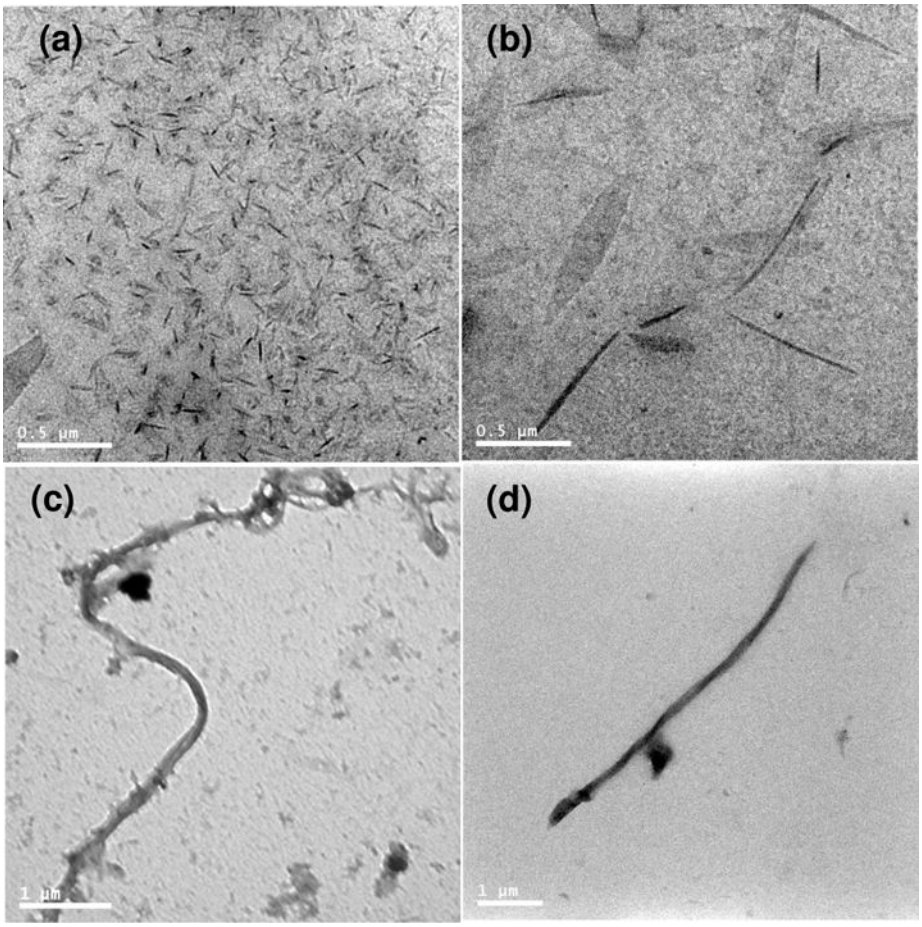



**Figure 6.**
TEM images of peptide assemblies stained with uranyl acetate with (a, b) 0.5 μm and (c, d) 1 μm scale bars.

**Table 1**

Comparison of the thermodynamic data obtained from CD and DSC for the oxidized CLP-Cys peptide at 500 μM concentration in 10 mM PBS (150 mM NaCl, pH 7.4).

| Sample | $T_m$ (°C) by CD | $T_m$ (°C) by DSC | $\Delta H_{vH}$ by CD (kcal/mole) | $\Delta H_{cal}$ by DSC (kcal/mole) |
|---|---|---|---|---|
| CLP-Cys | 43.9 ± 0.7 | 45.2 ± 0.3 | −48.8 ± 4.8 | −43.68 ± 1.6 |