Electronic Supplementary Material (ESI) for Soft Matter.

Electronic Supporting Information

Combining simulations and experiments for the molecular engineering of multifunctional collagen mimetic peptide-based materials

Amber M. Hilderbrand[1], Phillip A. Taylor[1], Francesca Stanzione[1], Mark LaRue[3], Chen Guo[1], Arthi Jayaraman[1,2]*, April M. Kloxin[1,2]*

Affiliations:
[1]Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716 USA
[2]Department of Materials Science and Engineering, University of Delaware, Newark, DE 19716 USA
[3]Department of Biomedical Engineering, University of Delaware, Newark, DE 19716 USA

*Corresponding Authors
Email: arthij@udel.edu
Email: akloxin@udel.edu

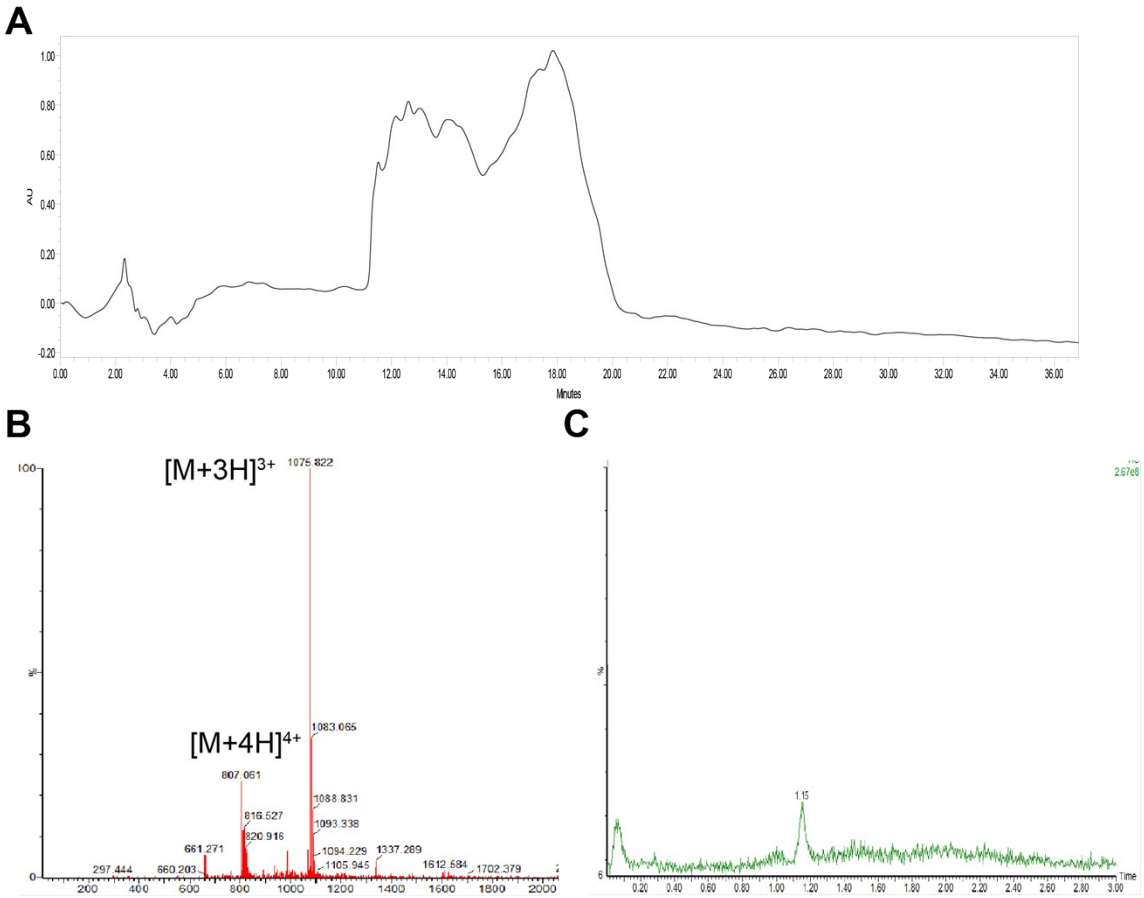


Figure S1: Purification and characterization of $(POG)_{12}$. (A) Preparative high performance liquid chromatography (HPLC) for purification of crude $(POG)_{12}$, where samples were collected from approximately 11 to 20 minutes and analyzed for the desired product. Purified $(POG)_{12}$ product (B) electrospray ionization (ESI+) mass spectrometry with expected molecular weight = 3224 g/mol ($[M+3H]^{3+}$ = 1075 g/mol, $[M+4H]^{4+}$ = 807 g/mol) and (C) ultra performance liquid chromatography (UPLC) trace showing singular product peak.

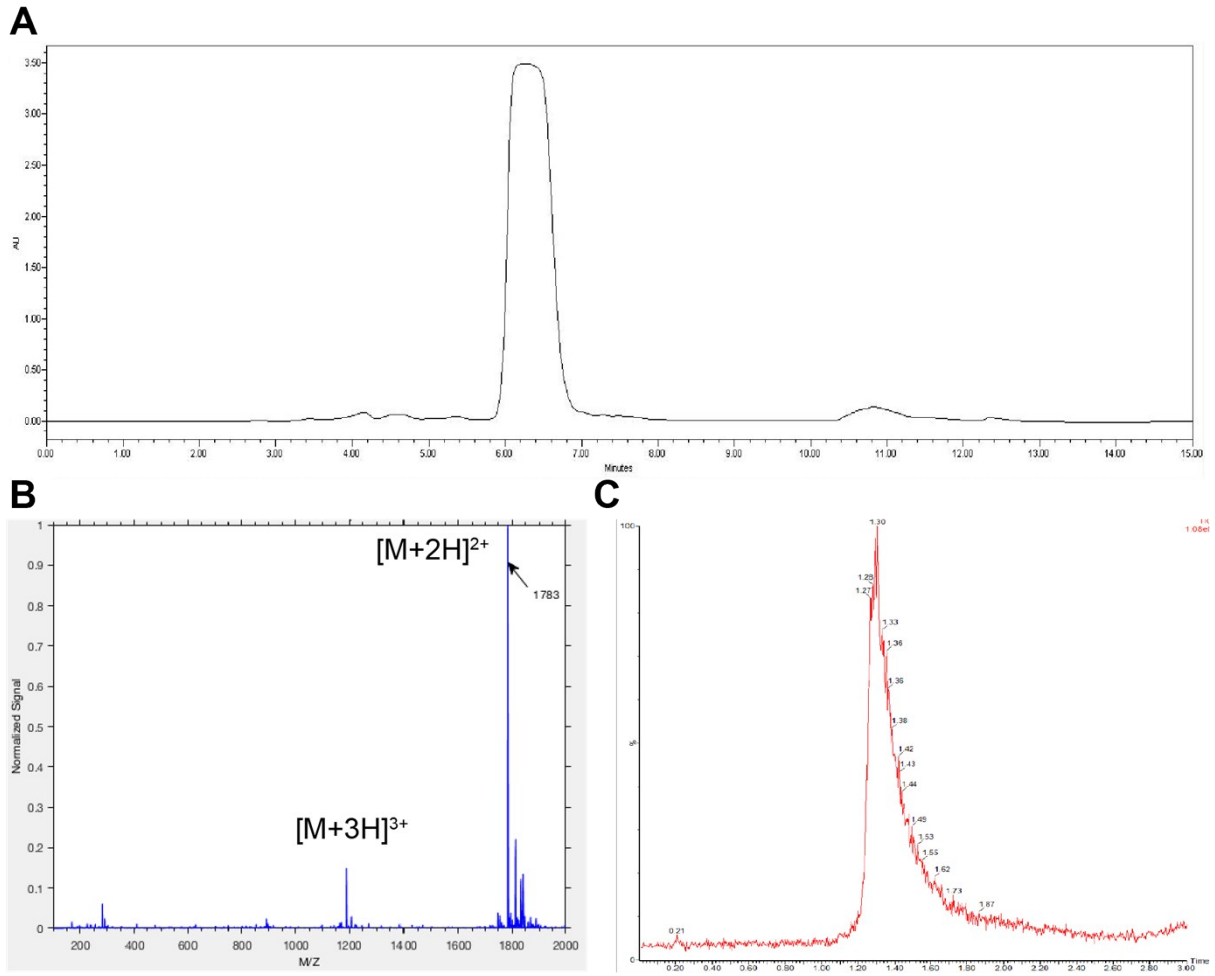


Figure S2: Purification and characterization of $Ka(PKG)_4(POG)_4(DOG)_4$. (A) Preparative HPLC of $Ka(PKG)_4(POG)_4(DOG)_4$ showed 1 main peak, which was collected and analyzed using mass spectrometry. Purified $Ka(PKG)_4(POG)_4(DOG)_4$ product (B) ESI+ mass spectrometry with expected molecular weight 3568 g/mol ($[M+2H]^{2+}$ = 1783 g/mol, $[M+3H]^{3+}$ = 1190 g/mol) and (C) UPLC trace, which also showed singular peak.

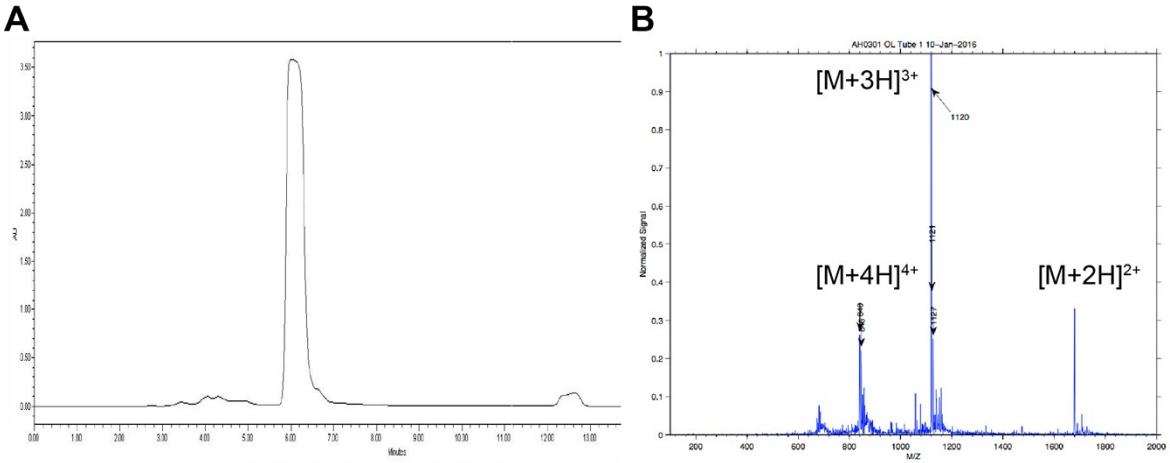


Figure S3: Purification and characterization of $(PKG)_4(POG)_4(DOG)_4$. (A) Preparative HPLC of $(PKG)_4(POG)_4(DOG)_4$ showed singular main peak that was collected and analyzed using mass spectrometry. Purified $(PKG)_4(POG)_4(DOG)_4$ product (B) ESI+ mass spectrometry with expected molecular weight 3356 g/mol ($[M+2H]^{2+}$ = 1677g/mol, $[M+3H]^{3+}$ = 1120 g/mol, $[M+4H]^{4+}$ = 840 g/mol).

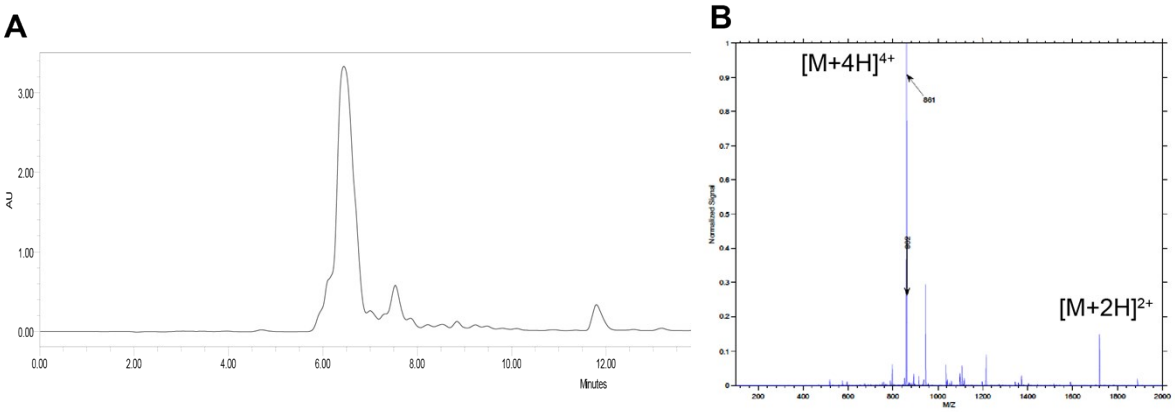


Figure S4: Purification and characterization of $(PKG)_3(PKaG)(POG)_4(DOG)_4$. (A) Preparative HPLC of $(PKG)_3(PKaG)(POG)_4(DOG)_4$ showed a main peak (from approximately 6 minutes to 7 minutes) along with 2 minor peaks that were found not to contain the desired product using mass spectrometry; the purified product isolated from the main peak was used. Purified $(PKG)_3(PKaG)(POG)_4(DOG)_4$ product (B) ESI+ mass spectrometry with expected molecular weight 3441 g/mol ($[M+2H]^{2+}$ = 1720g/mol, $[M+4H]^{4+}$ = 861 g/mol).

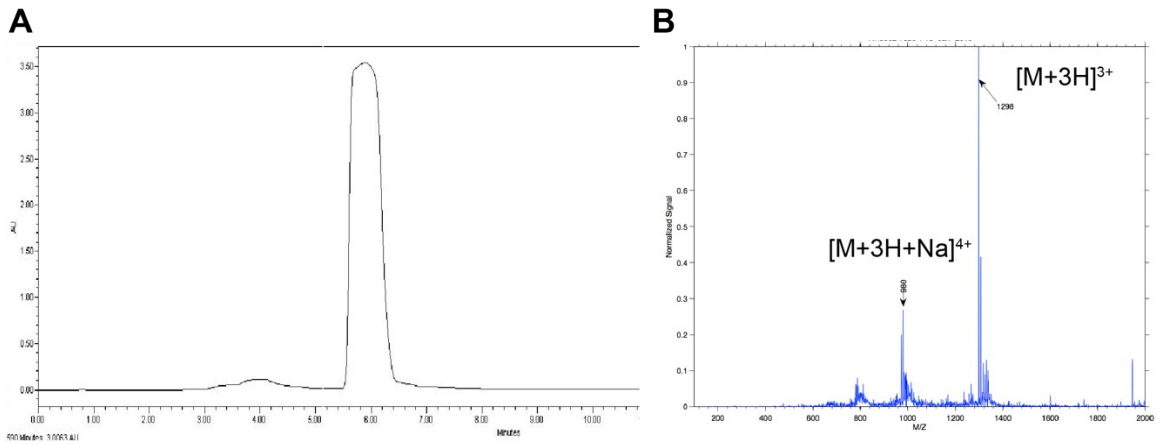


Figure S5: Purification and characterization of $(PKG)_4(POG)_6(DOG)_4$. (A) Preparative HPLC of $(PKG)_4(POG)_6(DOG)_4$ yielded 1 main peak that was collected and analyzed using mass spectrometry. Purified $(PKG)_4(POG)_6(DOG)_4$ product (B) ESI+ mass spectrometry with expected molecular weight 3891 g/mol ($[M+3H]^{3+}$ = 1298 g/mol, $[M+3H+Na]^{4+}$ = 980 g/mol).

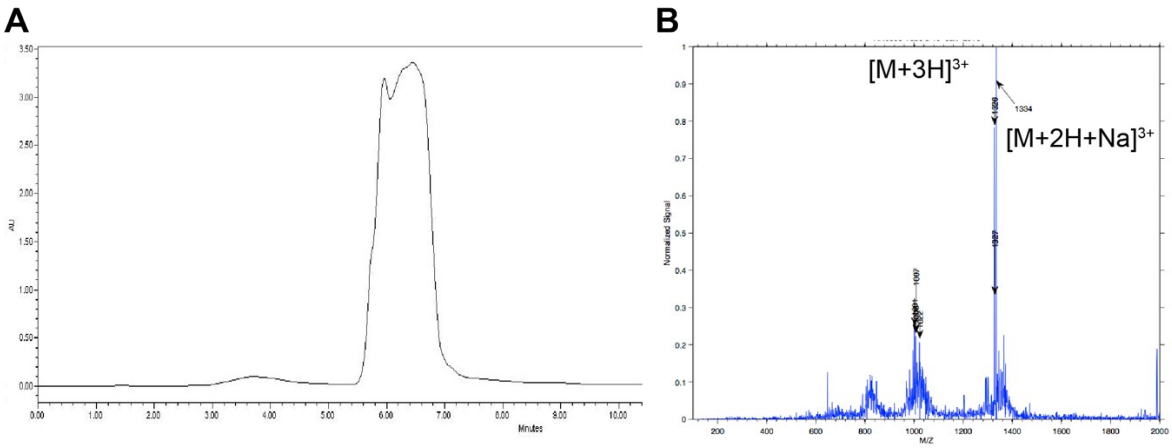


Figure S6: Purification and characterization of $(PKG)_3(PKaG)(POG)_6(DOG)_4$. (A) Preparative HPLC of $(PKG)_3(PKaG)(POG)_6(DOG)_4$ showed 1 main peak with a small shoulder, which were collected and analyzed using mass spectrometry; while both the main peak and the shoulder contained the desired product, the purified product isolated from the main peak was used. Purified $(PKG)_3(PKaG)(POG)_6(DOG)_4$ product (B) ESI+ mass spectrometry with expected molecular weight 3975 g/mol ($[M+3H]^{3+}$ = 1326 g/mol, $[M+2H+Na]^{3+}$ = 1334 g/mol).

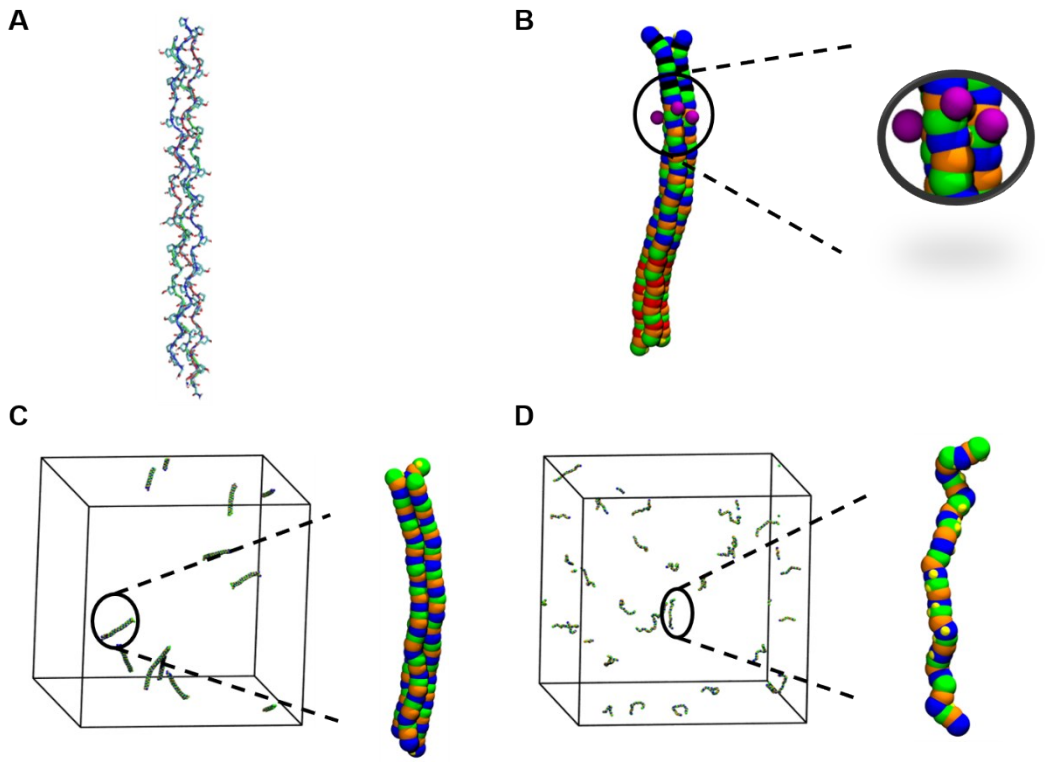


Figure S7: Simulation snapshots of atomistic and coarse-grained CMPs. (A) An atomistic representation of the $(POG)_{12}$ triple helix. (B) A snapshot of $(PKG)_3(PKaG)(POG)_6(DOG)_4$ including a zoomed-in perspective of three alloc beads. (C) Representative image of a simulation box of coarse-grained $(POG)_{12}$ triple helices below melting temperature; also included is a zoomed-in image of a single $(POG)_{12}$ triple helix. (D) Representative image of a simulation box containing coarse-grained melted $(POG)_{12}$ strands above melting temperature; also included is a zoomed-in perspective of a single $(POG)_{12}$ strand.

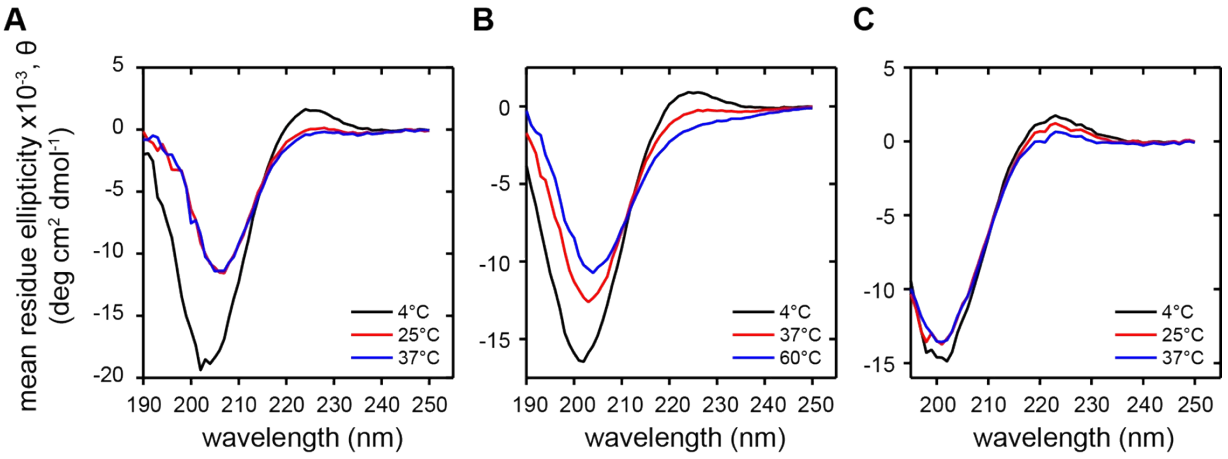
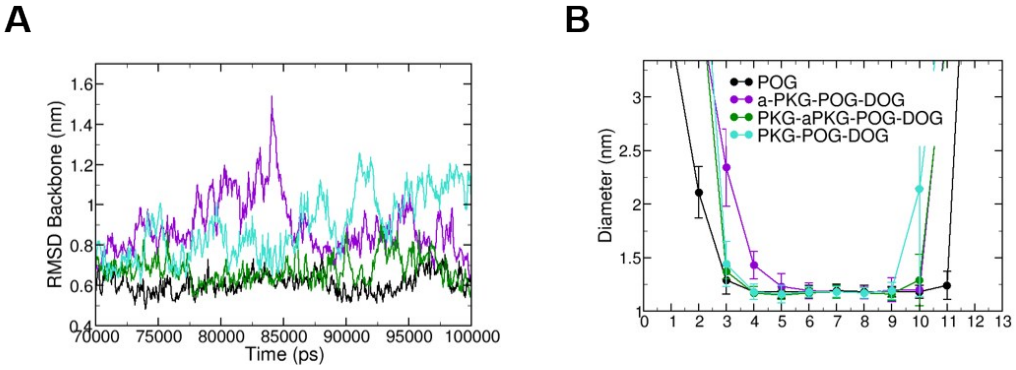


Figure S8: Wavelength scans of charged and charged and reactive sequences measured by circular dichroism. (A) $(PKG)_4(POG)_4(DOG)_4$, (B) $Ka(PKG)_4(POG)_4(DOG)_4$, (C) $(PKG)_3(PKaG)(POG)_4(DOG)_4$.



C

Sequence	Average end-to-end distance of backbone, $\langle R_{ee} \rangle$ (nm)
$(POG)_{12}$	8.05 ± 0.29
$(PKG)_4(POG)_4(DOG)_4$	7.9 ± 0.56
$Ka(PKG)_4(POG)_4(DOG)_4$	7.56 ± 0.36
$(PKG)_3(PKaG)(POG)_4(DOG)_4$	8.57 ± 0.48

Figure S9: Root mean square deviation (RMSD) of peptide backbone, diameter, and end-to-end distance analyses for atomistic CMPs. (A) Larger RMSDs were observed for $Ka(PKG)_4(POG)_4(DOG)_4$ relative to $(PKG)_3(PKaG)(POG)_4(DOG)_4$ indicating that the placement of the Kalloc residue in at the ends of the sequence destabilized the triple helix. (B**)** Greater splay was observed at the ends of the triple helix for Kalloc-containing CMP sequences such as $Ka(PKG)_4(POG)_4(DOG)_4$ and $(PKG)_3(PKaG)(POG)_4(DOG)_4$ as compared to $(POG)_{12}$. (C) Shorter average end-to-end distances of the peptide backbone ($\langle R_{ee} \rangle$) were observed for $Ka(PKG)_4(POG)_4(DOG)_4$ vs. $(PKG)_3(PKaG)(POG)_4(DOG)_4$ with a higher R_{ee} indicating a more stable triple helix.

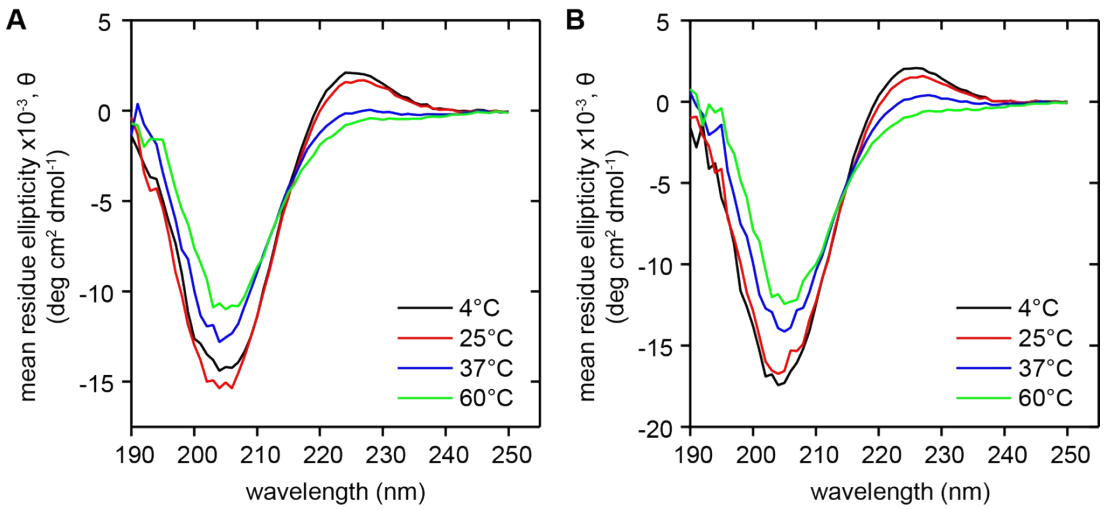


Figure S10: Wavelength scans of sequences with lengthened (POG) block measured by circular dichroism. (A) $(PKG)_4(POG)_6(DOG)_4$ and (B) $(PKG)_3(PKaG)(POG)_6(DOG)_4$.

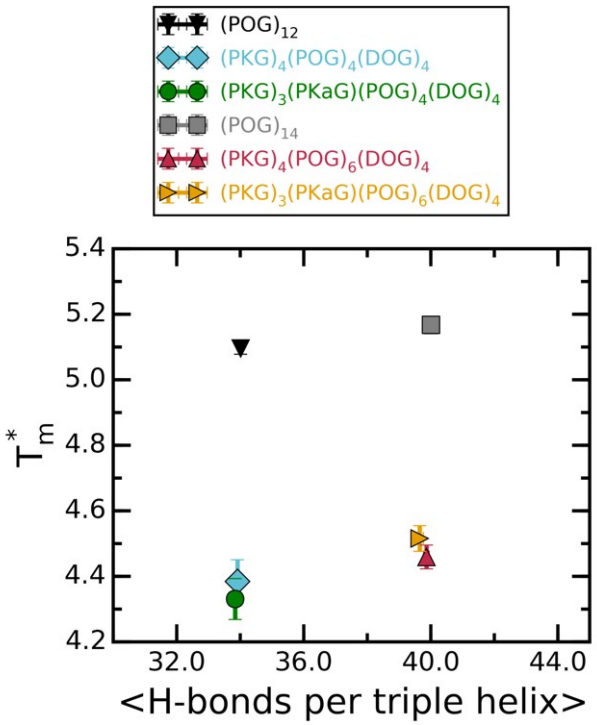


Figure S11: Calculated melting temperature, T^*_m, (in reduced units) versus average number of H-bonds in each triple helix, <H-bonds per triple helix>, obtained at T* = 3.0 (below the melting temperature for all sequences) from CG MD simulations. The error bars are the standard deviations calculated for the three mean values from three independent trials.

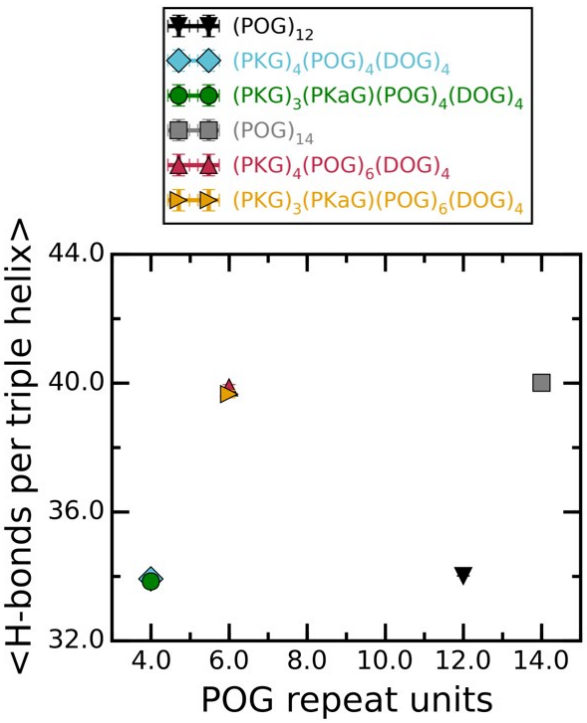


Figure S12: Average number of H-bonds per triple helix <H-bonds per triple helix> versus CMP sequence design, obtained at T* = 3.0 (below the melting temperature for all sequences). from CG MD simulations. The error bars are smaller than the size of the symbol and are the standard deviations calculated for the three mean values from three independent trials.