

Peptoid Residues Make Diverse, Hyperstable Collagen Triple Helices

Julian Kessler, Grace Kang, Zhao Qin, Helen Kang, Frank Whitby, Thomas Cheatham, Christopher Hill, Yang Li, S. Michael Yu

Submitted date: 05/09/2020 • Posted date: 08/09/2020

Licence: CC BY-NC-ND 4.0

Citation information: Kessler, Julian; Kang, Grace; Qin, Zhao; Kang, Helen; Whitby, Frank; Cheatham, Thomas; et al. (2020): Peptoid Residues Make Diverse, Hyperstable Collagen Triple Helices. ChemRxiv. Preprint. <https://doi.org/10.26434/chemrxiv.12921995.v1>

The triple-helical structure of collagen, responsible for collagen's remarkable biological and mechanical properties, has inspired both basic and applied research in synthetic peptide mimetics for decades. Since non-proline amino acids weaken the triple helix, the cyclic structure of proline has been considered necessary, and functional collagen mimetic peptides (CMPs) with diverse sidechains have been difficult to produce. Here we show that N-substituted glycines (N-glys), also known as peptoid residues, exhibit a general triple-helical propensity similar to or greater than proline, allowing synthesis of thermally stable triple-helical CMPs with unprecedented sidechain diversity. We found that the N-glys stabilize the triple helix by sterically promoting the preorganization of individual CMP chains into the polyproline-II helix conformation. Our findings were supported by the crystal structures of two atomic-resolution N-gly-containing CMPs, as well as experimental and computational studies spanning more than 30 N-gly-containing peptides. We demonstrated that N-gly sidechains with diverse exotic moieties including a 'click'-able alkyne and a photo-sensitive sidechain can be incorporated into stable triple helices, enabling functional applications such as spatio-temporal control of cell adhesion and migration on a gelatin matrix. The folding principles discovered in this study open up opportunities for a new generation of collagen mimetic therapeutics and materials with extraordinary properties.

File list (2)

ChemRxiv PEPTOID-CMP_manuscript 20200905_submitt... (1.52 MiB)	view on ChemRxiv • download file
--	--

ChemRxiv-SI-Peptoid-CMP_submitted (20200905).pdf (4.82 MiB)	view on ChemRxiv • download file
---	--

Peptoid Residues Make Diverse, Hyperstable Collagen Triple Helices

Julian L. Kessler¹, Grace Kang¹, Zhao Qin², Helen Kang¹, Frank G. Whitby³, Thomas E.

Cheatham III⁴, Christopher P. Hill³, Yang Li^{1,*}, and S. Michael Yu^{1,5}

¹Department of Biomedical Engineering, University of Utah, Salt Lake City, Utah 84112, USA

²Department of Civil & Environmental Engineering, College of Engineering & Computer Science, Syracuse University, Syracuse, New York 13244, USA

³Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

⁴Department of Medicinal Chemistry, College of Pharmacy, L. S. Skaggs Pharmacy Research Institute, University of Utah, Salt Lake City, Utah 84112, USA

⁵Department of Pharmaceutics and Pharmaceutical Chemistry, University of Utah, Salt Lake City, Utah 84112, USA

***Corresponding Author:** Yang Li (yang.d.li@utah.edu)

Abstract

The triple-helical structure of collagen, responsible for collagen's remarkable biological and mechanical properties, has inspired both basic and applied research in synthetic peptide mimetics for decades. Since non-proline amino acids weaken the triple helix, the cyclic structure of proline has been considered necessary, and functional collagen mimetic peptides (CMPs) with diverse sidechains have been difficult to produce. Here we show that *N*-substituted glycines (N-glys), also known as peptoid residues, exhibit a general triple-helical propensity similar to or greater than proline, allowing synthesis of thermally stable triple-helical CMPs with unprecedented sidechain diversity. We found that the N-glys stabilize the triple helix by sterically promoting the preorganization of individual CMP chains into the polyproline-II helix conformation. Our findings were supported by the crystal structures of two atomic-resolution N-gly-containing CMPs, as well as experimental and computational studies spanning more than 30 N-gly-containing peptides. We demonstrated that N-gly sidechains with diverse exotic moieties including a 'click'-able alkyne and a photo-sensitive sidechain can be incorporated into stable triple helices, enabling functional applications such spatio-temporal control of cell adhesion and migration on a gelatin matrix. The folding principles discovered in this study open up opportunities for a new generation of collagen mimetic therapeutics and materials with extraordinary properties.

Introduction

Among the twenty canonical amino acids, proline (Pro), formally an imino acid, stands out as the single residue that features an *N*-substitution. Pro's unique cyclic sidechain mandates restricted backbone dihedral angles and, when incorporated into a protein chain, its tertiary amide group lacks the *N*-hydrogen atom to donate a hydrogen bond. As a result, Pro disrupts the regular secondary conformations favored by most amino acids, such as α -helices and β -sheets, and promotes distinct folding patterns, such as β -turns and polyproline helices. From a chemical perspective, there are numerous *N*-substituted α -amino acids as possible protein building blocks¹, of which evolution has sampled only *one* candidate for ribosomal protein expression: Pro. The rules governing the folding of *N*-substituted amino acids and their relationship to Pro's unique conformation are intriguing topics for both fundamental science and practical applications of *de novo* protein designs².

Collagen, the most abundant protein in vertebrates, best exemplifies a natural protein structure dictated by Pro's folding. The collagen chain comprises a repetitive sequence of Gly-X-Y triplets, where the X and Y positions are often occupied by Pro and its post-translational hydroxylation product, 4(*R*)-hydroxyproline (Hyp)³. As many as 22% of all residues in human collagen are either Pro or Hyp⁴ and, as a result, the defining structural motif of collagen is the intertwining of three polyproline II-type helices^{5,6}. This unique *triple helix* motif is responsible for collagen's remarkable properties, including higher-order assembly⁷, mechanical strength⁸, resistance to proteases, and binding with numerous cell receptors⁹. The triple helix has also inspired many synthetic designs aimed at recapitulating nature's supramolecular chemistry^{10,11}.

Because collagens are large, insoluble proteins that are difficult to study holistically, many research groups have turned to synthetic collagen mimetic peptides (CMPs)^{3,6,12}. CMPs featuring GlyProPro or GlyProHyp triplets have been synthesized as models of the collagen triple helix since the late 1960s¹³. Studies of CMPs revealed that all natural amino acids destabilize the triple helix when replacing Pro in the GlyProHyp triplet¹⁴, while hydroxylation of Pro at the Y position stabilizes the structure³. Throughout the 1990s¹² and more recently^{15,16}, a substantial portion of CMP studies have focused on uncovering and exploiting the stabilizing effects of the

post-translational modification of Pro on the triple helix by leveraging synthetic Pro derivatives^{3,15-19}.

Despite more than five decades of vigorous research, there have been few attempts to produce a stable collagen triple helix without Pro, and the structural forces imposed on the triple-helical conformation by unnatural *N*-substituted amino acids remain almost entirely unexplored. As probably the only precedent, Goodman and coworkers showed that replacing the central Pro with *N*-isobutylglycine (Nleu) within the GlyProPro triplet of a CMP results in a more stable triple helix²⁰. They suggested that the stabilization results from interchain interactions between the hydrophobic sidechains of Nleu and the adjacent Pro^{21,22}; however this hypothesis was not tested further. Meanwhile, peptoids, which are synthetic oligomers of *N*-substituted glycines (the simplest subset of *N*-substituted α -amino acids, e.g., Nleu), have been studied for decades as a promising class of synthetic peptidomimetics in biomedicine and material science^{23,24}. Repetitive peptoid sequences with various sidechains have been shown to form robust conformations resembling polyproline helices²⁵⁻²⁹, as well as a multitude of secondary and higher order structures including helical bundles³⁰, superhelices³¹, and two-dimensional sheets^{32,33}. Yet, so far, the structural influences of peptoid residues upon collagen triple helicity remain unclear.

Here we present a systematic study on the triple-helical folding propensity of *N*-substituted glycines (N-glys, **Fig. 1a**), and demonstrate their potential for making stable collagen triple helices with extraordinarily diverse sidechain structures. Using host-guest peptide systems for both collagen-mimetic and polyproline peptides, we examined more than 30 different peptides featuring N-glys with a spectrum of sidechain chemistries, for their triple-helical and polyproline-II-helix (PPII) propensities. We also acquired atomic-resolution crystal structures of two N-gly containing CMPs and performed molecular dynamics (MD) simulations, which together provide insights into the molecular structure of peptoid residues within collagen triple helices and their folding mechanism. Furthermore, to demonstrate utility, we present stable triple helices featuring exotic N-gly sidechains including light sensitive ones, which enabled spatiotemporal control of cell adhesion and migration on a collagen substrate. So far, construction of synthetic collagen mimetics has critically relied on Pro and its derivatives; the

design principles uncovered in this study drastically expand the library of residues with high triple-helical propensity, which has immense implications for a new generation of collagen-mimetic therapeutics and materials.

Results and Discussion

Triple-helical stability of CMPs with N-glys. To investigate the triple-helical folding propensity of peptoids residues, we inserted a series of N-gly guests into the central X position of a conventional CMP host peptide with the sequence: Ac-(GlyProHyp)₃-Gly-X-Hyp-(GlyProHyp)₃-NH₂ (designated as X-CMP, **Fig. 1b**, and supplementary Materials and Methods)³⁴. We measured the X-CMPs' triple-helical stability *via* thermal unfolding experiments monitored under circular dichroism (CD, **Fig. 1c**, Supplementary Section 3 and **Table S1**), first in a group of N-glys with sidechains selected from the canonical amino acids (**Fig. 1d**). It is known that Pro is the most stabilizing amino acid at position X, and substitution from Pro to another canonical amino acid clearly reduces T_m by 4-17 °C (**Fig. 1d**, column AA)¹⁴. Surprisingly, we found that many of the N-glys with canonical sidechains were at least as stable as Pro, and almost *all* N-gly residues were more stable than their amino acid counterparts (**Fig. 1d**, exception: Nval), with the biggest T_m difference seen between Nphe- and Phe-CMP (**Fig. 1c**). These results demonstrate that, in the X position of the CMP, shifting the sidechain from the C α carbon to the nitrogen (i.e., transforming a canonical amino acid to its peptoid analogue) may improve triple-helical stability.

Crystal structures of N-gly-CMPs. To confirm the triple-helical folding of N-gly-CMPs and to decipher the precise molecular structure of the N-glys within the peptide conformation, we determined the X-ray crystal structures of two compounds. We chose Nlys-CMP and Nphe-CMP for crystallographic analysis because of their high stabilities and differing N-gly sidechain moieties, and determined their structures at resolutions of 0.95 Å and 1.10 Å, respectively (**Fig. 2a, 2b**; **Supplementary Table S2, Fig. S1, S2**). The overall conformation in both crystal structures was consistent with those reported for conventional CMPs found in the Protein Data Bank (PDB, **SI Table S3**), with the average (φ , ψ) angles and standard deviations of all non-

terminal Pro, Hyp, and Gly residues being ($-70 \pm 6^\circ$, $162 \pm 6^\circ$), ($-58 \pm 5^\circ$, $147 \pm 7^\circ$), and ($-69 \pm 3^\circ$, $174 \pm 5^\circ$), respectively (**Fig. 2c**, **SI Fig. S3**). The dihedral angles of the two N-glys [Nlys: ($-86 \pm 3^\circ$, $176 \pm 3^\circ$), Nphe: ($-82 \pm 3^\circ$, $177 \pm 3^\circ$)] were also within the range populated by the X-position residues of collagen peptides in the PDB (**Fig. 2d**), though slightly shifted from the PDB's average angles ($\Delta\phi = -13^\circ$, $\Delta\psi = +21^\circ$). This deviation did not significantly alter the backbone structure and there was neither kinking nor bending of the triple helix near the N-gly in either crystal structure. Furthermore, the inter-strand hydrogen bonding pattern, formed along the length of the triple helix, agreed well with that seen in other collagen peptides in the PDB (**Fig. 2e**), although the hydrogen bonds formed with N-glys appeared to adopt a slightly less ideal geometry (**Fig. 2f**). This suggests that the triple-helical stability of N-gly-CMPs does not originate from improved hydrogen bonding³⁵ of N-glys.

Probing the structural factors for N-glys' triple-helical stability. Next, we introduced a series of guest N-glys into the X-CMPs to understand how sidechain structures affect triple-helical stability (**Fig. 3**). We first focused on hydrophobic sidechains, which previous reports suggested to be the main stabilizing factor in N-gly-CMPs^{21,22}. We found that the aliphatic Nchx as well as the aromatic Nphe produced triple helices with stability notably higher than Pro (**Fig. 3a**). In addition, Nphe derivatives having an electron-rich (Ntyr) or poor (Nnbz) aromatic ring generated triple helices with similar stability, suggesting that neither the aromaticity nor its electron density is crucial to triple helix stabilization. Meanwhile, although structurally similar, the hydrophilic Ndxn and the deprotonated Ntyr (pH 12) were both more stabilizing than Pro and only slightly less stabilizing than their hydrophobic analogs in the group (**Fig. 3a**). These results suggest that the hydrophobicity of the peptoid residue might not be the main driving force for the triple helix stabilization. Furthermore, we synthesized a pair of CMPs featuring central guest triplets of GlyNpheHyp and GlyNphelle to examine the hydrophobic interactions between adjacent residues (SI Table S4). Introduction of Ile did not provide any additional stabilizing effect, further de-emphasizing hydrophobicity in triple helix stabilization.

Next, we investigated the charged and hydrophilic peptoid sidechains (**Fig. 3b**). Although charged Nasp and Nlys destabilized the triple helices modestly through electrostatic repulsions,

when we altered the pH so their sidechains were uncharged, both X-CMPs were as stable as Pro-CMP. Also, we noted that peptoid residues with similar sidechain configurations conferred highly similar T_m values (**Fig. 3b**: Nasp, Nasn, and Nleu; **Fig. 3a**: Nchx, Nphe, Nnbz, Ntyr, Ndxn), while within each group, T_m values of the hydrophobic residues were only higher by 2-4 °C.

Lastly, we found that progressively increasing the size of the aliphatic side chains produced a concomitant increase in T_m of the corresponding X-CMP (**Fig. 3c**). Knowing that the hydrophobicity has modest stabilizing effects for triple helices, we reasoned that the large increase in stability from Gly-CMP (T_m , 37 °C) to Nchx-CMP (T_m , 63 °C) in this series cannot be explained by the hydrophobicity alone, and were prompted to look into other mechanisms by which peptoid residues stabilize the collagen triple helix.

Polyproline-II-helix (PPII) propensity of N-glys. It is generally accepted that Pro and Hyp residues stabilize the collagen triple helix by pre-organizing individual collagen strands into the PPII conformation³⁶. Therefore, we hypothesized that N-glys, with their Pro-like tertiary amide structure, could induce similar pre-organizational effects in N-gly-CMPs. To test this hypothesis, we first surveyed a large set of small molecule crystal structures containing N-glys from the Cambridge Structural Database (SI **Table S5, S6**). We found that the dihedral angles of these peptoid residues largely fall within the most probable range for Pro, with the highest density present in the area corresponding to a PPII helix (**Fig. 4a**). This suggests that similar to Pro, N-glys have a high natural propensity for the PPII conformation.

To systematically study the N-glys' propensity to adopt the PPII conformation, we employed another host-guest peptide sequence: Ac-GlyProPro-X-ProProGlyTyr-NH₂, designated as X-PP5, where X represents the N-gly or amino acid residue replacing the central Pro within the (Pro)₅ sequence. This short, proline-rich host peptide was previously utilized to measure the PPII propensity of the twenty canonical amino acids based on the CD intensity near 228 nm^{37,38}. Except for Sar-PP5, all N-gly-CMPs showed PPII CD intensity higher than Pro-PP5 (**Fig. 4b**) and their amino acid counterparts³⁷. Notably, the PPII propensities of Nphe, Nlys, and Nleu were higher than that of Hyp, which is known for having the highest PPII propensity among all

natural amino acids³⁹. This systematic investigation supports that N-glys promote the intrinsic PPII folding of individual peptide strands, which can contribute to triple-helical stabilization. Energy-wise, the free energy bonus arising from the PPII preorganization by just a few bulky N-glys in a triple helix (**Fig. 4b**, $\Delta \Delta G$) can add up to be roughly an extra amide-amide hydrogen bond (-2.0 kcal/mol)⁴⁰.

Previous studies have shown that peptoids with specific *N*-substitutions can form stable polyproline helices^{25-29,41}. So far, the frontier of peptoid research has largely focused on control over the *cis-trans* isomerization of peptoid amide bonds (i.e., adoptions of PPI or PPII conformation) through specific steric or stereoelectronic effects induced by *N*-aryl or *N*-C α -chiral sidechains^{25-29,42}. In this work, we systematically quantitated the general PPII propensity of peptoid residues. We learned that although Pro has the highest PPII propensity among all canonical amino acids, it is on the low side among the peptoid residues (**Fig. 4b**). The strong PPII propensity of peptoid residues as compared to the amino acids implies that Pro may be the an imperfect, biosynthetic surrogate for N-glys⁴³ where PPII folding is needed, but its moderate folding propensity can be improved *via* post-translational hydroxylation (e.g., in collagen).

MD simulation and steric effects of the *N*-substitution. We conducted MD simulations for triple-helical X-CMPs incorporating guest residues Gly, Sar, Nleu, Nchx, or Pro (**Fig. 4c**). We based our simulations on the crystal structures of Nphe- and Nlys-CMPs (**Fig. 2**) and chose N-glys with simple aliphatic side chains that lack potential for electrostatic and hydrogen bonding interactions. The simulation results demonstrated that the Nleu and Nchx residues displayed energy landscapes very similar to Pro, with a low-energy valley confined to the regions (φ , ψ angles) for PPII and the triple helix. In comparison, Sar which has a smaller sidechain, and to a greater extent Gly, showed more conformational flexibility (**Fig. 4c**). These simulation results are in agreement with our CD experiments, which showed that an N-gly's PPII propensity (i.e., Gly < Sar < Nleu) and the triple-helical stability of its corresponding X-CMP correlate with the size of its *N*-substitution (**Fig. 3c & 4c**). As a plausible explanation for this steric effect, a bulky sidechain can impose a steric restraint to the N-gly's backbone carbonyl, thereby hindering the rotation around its nitrogen-carbon bond and constraining the φ angle to values appropriate for

triple-helical folding (**Fig. 4d**). This is evident in the most flexible guest residue Gly, where addition of a single *N*-methyl group from Sar increases the T_m by 13 °C (**Fig. 3c**). Additionally, we believe that an N-gly's ψ angle is less sterically influenced by the *N*-substitution because it is located further from the sidechain, and that it is more likely to assume a value close to $\pm 180^\circ$ because the α -carbon has no sidechain.

The pyrrolidine ring of Pro has been assumed to be a structural requisite for collagen mimetics for decades^{3,44}. In an attempt to replace Pro in CMPs, Raines and coworkers found that converting Pro to *N*-methyl-L-alanine (NMe-Ala), which removes only the γ carbon and eliminates the ring structure, substantially destabilizes the triple helix⁴⁵ (**SI Fig. S4**). From this observation, they concluded that the conformational restrictions imposed by the Pro ring are more important for triple-helical stability than the presence of an *N*-substitution⁴⁵. In contrast, we discovered that N-glys with bulky sidechains can form well-folded collagen triple helices without Pro's cyclic sidechain. We believe that the low triple helicity of NMe-Ala⁴⁵ is due to its poor PPII propensity (**Fig. 4b**). Without the ring, steric repulsion between the two adjacent methyl groups of NMe-Ala precludes formation of the dihedral angles necessary to form the polyproline-II helix, thereby nullifying any stabilizing effect from the *N*-methyl group (**SI Fig. S4**).

***N*-C α chiral branching.** To understand why Nval, the only peptoid residue with an *N*-C α -branched sidechain in our library, has unexpected, low triple-helical stability (T_m : 36 °C, **Fig. 1d**), we created a group of X-CMPs featuring guest residues NEt, Nval, Nphe, Nspe and Nrpe (**Fig. 5**). The T_m differences (**Fig. 5a**) indicated that adding an *N*-C α methyl group to an N-gly destabilizes the triple helix, and that the triple-helical folding was greatly affected by the chirality of the *N*-C α -branching (**Fig. 5a**, Nspe versus Nrpe). For each chirality, 3D modeling based on the Nphe-CMP crystal structure revealed a specific location of possible steric clash between the branching methyl group and an adjacent peptide strand (**Fig. 5b,c**), which provides an explanation for Nval's low triple-helicity despite its higher-than-Pro PPII propensity (**Fig. 4b**). Although additional confirmatory studies are needed, this is the first time that a

stereoselective steric interaction at the *N*- α -carbon is presented in the context of collagen triple helix.

Sidechain functionalization. After a series of mechanistic investigations into triple-helical folding, we turned to leveraging the diverse pool of peptoid residues to produce functionalized collagen mimetics for potential applications. As one example, we demonstrated facile sidechain functionalization through ‘click’ chemistry, utilizing an alkyne-bearing residue (Nakn, **Fig. 6a**, **SI Methods** and **SI Discussions**). In another example, to demonstrate precise control over the structure and function of CMP-based molecules and materials, we synthesized Nnbz-CMP, an X-CMP featuring a UV-cleavable *N*-*o*-nitrobenzyl (nbz) sidechain⁴⁶. We hypothesized that Nnbz-CMP would form a robust triple helix, but UV irradiation would convert the stabilizing Nnbz residue into Gly, thereby dramatically weakening the triple helix (**Fig. 6b**). CD measurements confirmed this hypothesis (**Fig. 6c**). Moreover, characterizations of CMPs incorporating more than one Nnbz residues (e.g., Nnbz2,3, SI Table S7) indicated that (i) the stabilizing effect of N-glys may be additive, and that (ii) the increased number of Pro \rightarrow Nnbz substitutions resulted in even wider T_m differences before and after UV irradiation (**Fig. 6c**).

Next, we examined the Nnbz-substituted CMP’s capacity to hybridize to denatured collagen and tested their release afterwards by UV irradiation. Previously, we reported that single-strand CMPs [e.g., (GlyProHyp)₉] can form hybrid triple helices with collagens that are denatured by heat, proteases, or mechanical damage^{9,69}. Using Nnbz2-CMP labeled with fluorescein (F-Nnbz2-CMP), We found that Nnbz2-CMP displayed a drastically higher level of binding to gelatin (heat-denatured collagen) than the host peptide Pro-CMP, consistent with its stronger triple-helical propensity (**Fig. 6d**). Furthermore, upon brief UV irradiation, over 80% of the bound F-Nnbz2-CMP was released from the gelatin matrix (**Fig. 6d**).

Spatio-temporal modification of cell culture substrates. To further demonstrate X-CMP’s applications in biomaterials, we conjugated Nnbz2-CMP to a multi-arm poly(ethylene glycol) (MA-PEG) polymer (**SI Fig. S5a,b**). The resulting conjugate (MA-PEG-CMP) readily formed a hydrogel that was physically crosslinked *via* its Nnbz2-CMP triple helices; this stable hydrogel

showed the ability to dissolve and release its bound contents on demand as the CMP crosslinks unfold upon UV irradiation (**SI Fig. S5**). Next, taking advantage of PEG's ability to repel cell adhesion and Nnbz2-CMP's UV-triggered unbinding from gelatin (**Fig. 6d**), we demonstrated that a crosslinked gelatin film can be coated with MA-PEG-CMP to accurately photo-pattern cell attachment in selected areas (**SI Fig. S6**), or to trigger cell migration into defined regions at a designated time (**Fig. 6e,f**). These proof-of-concept experiments showcase the practical use of X-CMPs for creating complex cell culture substrates with spatio-temporal controls.

Billions of stable, Pro-free triple helices. Lastly, we synthesized **X7-CMP**, a (Gly-X-Hyp)₇ sequence where each X position was populated with a different peptoid residue (**Fig. 6g**). With a T_m 1 °C higher than (GlyProHyp)₇, **X7-CMP** represents a hyperstable, synthetic collagen triple-helix with the greatest sidechain-diversity to date. With hundreds of amines commercially available for efficient solid-phase peptoid synthesis, this work expands the amino-acid library for stable CMPs by more than an order of magnitude. Even with a modest library of 20 N-glycyls for X-substitution, for a short sequence of (Gly-X-Hyp)₇, over 1.2 billion (20^7) triple-helical peptides can be generated. The design principles discovered in this study open the door for not only deciphering new fundamental molecular interactions (e.g., $n \rightarrow \pi^*$)⁴⁷ in polyproline and triple-helical folding, but also developing a new class of collagen-mimetic therapeutics and biomaterials with remarkable functionalities. With the unprecedented structural diversity offered by X-CMPs, we anticipate the long-awaited new era of *de novo* design of functional collagen peptidomimetics.

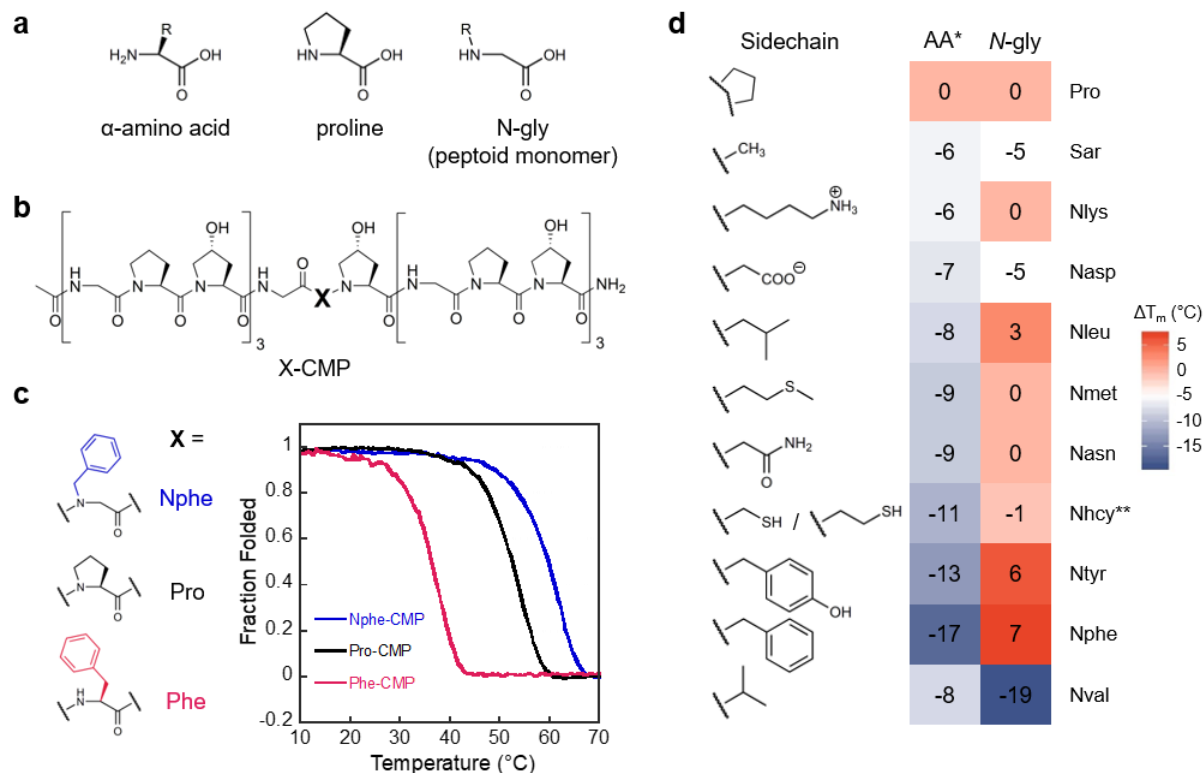


Fig. 1 | A single N-substituted glycine (N-gly), also known as a peptoid residue, in the central Pro position within (GlyProHyp)₇ shows high triple-helical propensity. **a**, General structures of a canonical amino acid, Pro, and an N-gly. **b**, Chemical structure of the X-CMP, a host-guest collagen mimetic peptide with the sequence of a (GlyProHyp)₇, where the Pro residue in the central X position is substituted, for example, with an N-gly residue. **c**, An X-CMP triple helix dissociates into single strands under gradual heating; monitored under CD, the middle point of this two-state transition is defined as the X-CMP's melting temperature (T_m). The Nphe sidechain at the X position resulted in a CMP triple helix considerably more stable than Pro (ΔT_m : +7 $^{\circ}\text{C}$). In comparison, the amino acid with the same sidechain was highly destabilizing (ΔT_m : -17 $^{\circ}\text{C}$). **d**, The triple helical stabilities of X-CMPs featuring N-glys with canonical sidechains, quantitated in ΔT_m (in comparison to the host peptide: Pro-CMP). Almost universally, peptoid residues (right column) have higher stability than their amino acid counterparts (left column) except for Nval, and many have stabilities similar to or better than Pro. * ΔT_m for amino acid residues taken from reference²¹. **Nhcy (homocysteine) is a similar but not exact equivalent to cysteine.

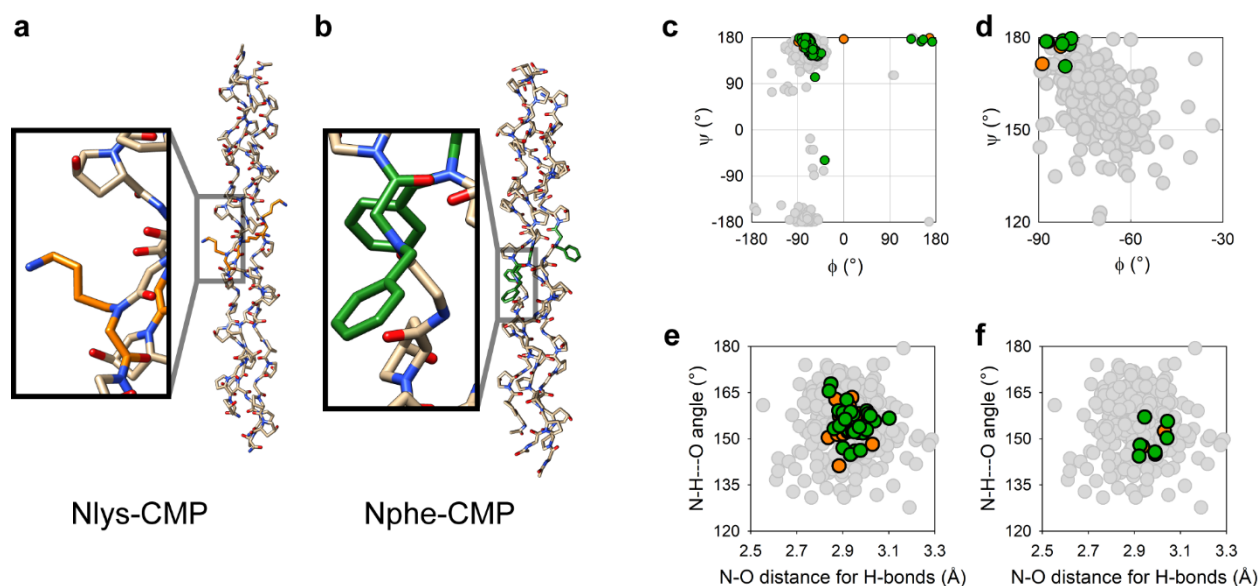


Fig. 2 | CMPs containing peptoid residues adopt a triple-helical conformation nearly identical to the native collagen. a,b Crystal structures of Nlys-CMP (a) and Nphe-CMP (b); each shows a triple helix with the N-gly residue uniquely colored (Nlys in orange and Nphe in green). **c**, A Ramachandran plot showing the ϕ and ψ angles of all amino acid residues in the crystal structure of Nlys-CMP (orange) and Nphe-CMP (green), overlaid with the ϕ and ψ angles from the reported CMP crystal structure in the PDB (grey, see SI for structure IDs used for comparison). **d**, A Ramachandran plot showing the ϕ and ψ angles for only the N-gly residues (Nlys in orange, Nphe in green), overlaid with the ϕ and ψ angles (in grey) from the X position of Gly-X-Y repeats of reported CMP crystal structures. The peptoid residues adopt a conformation common to the X position in previously reported Gly-X-Y collagen sequences, yet with slightly shifted average angles ($\Delta\phi$: -13° , $\Delta\psi$: $+21^\circ$). **e**, A plot of the N-H...O angle versus the N-O distance of all the interstrand amide-amide H-bonds in the crystal structures of Nlys-CMP (orange) and Nphe-CMP (green) overlaid with that (in grey) of all the H-bonds in the reported crystal structure of CMPs. **f**, A plot, same as **e** but showing H-bonds associated only with N-gly residues in the crystal structures (Nlys in orange, and Nphe in green). The H-bonds parameters for both residues are consistent with previously reported crystal structures of collagen peptides.

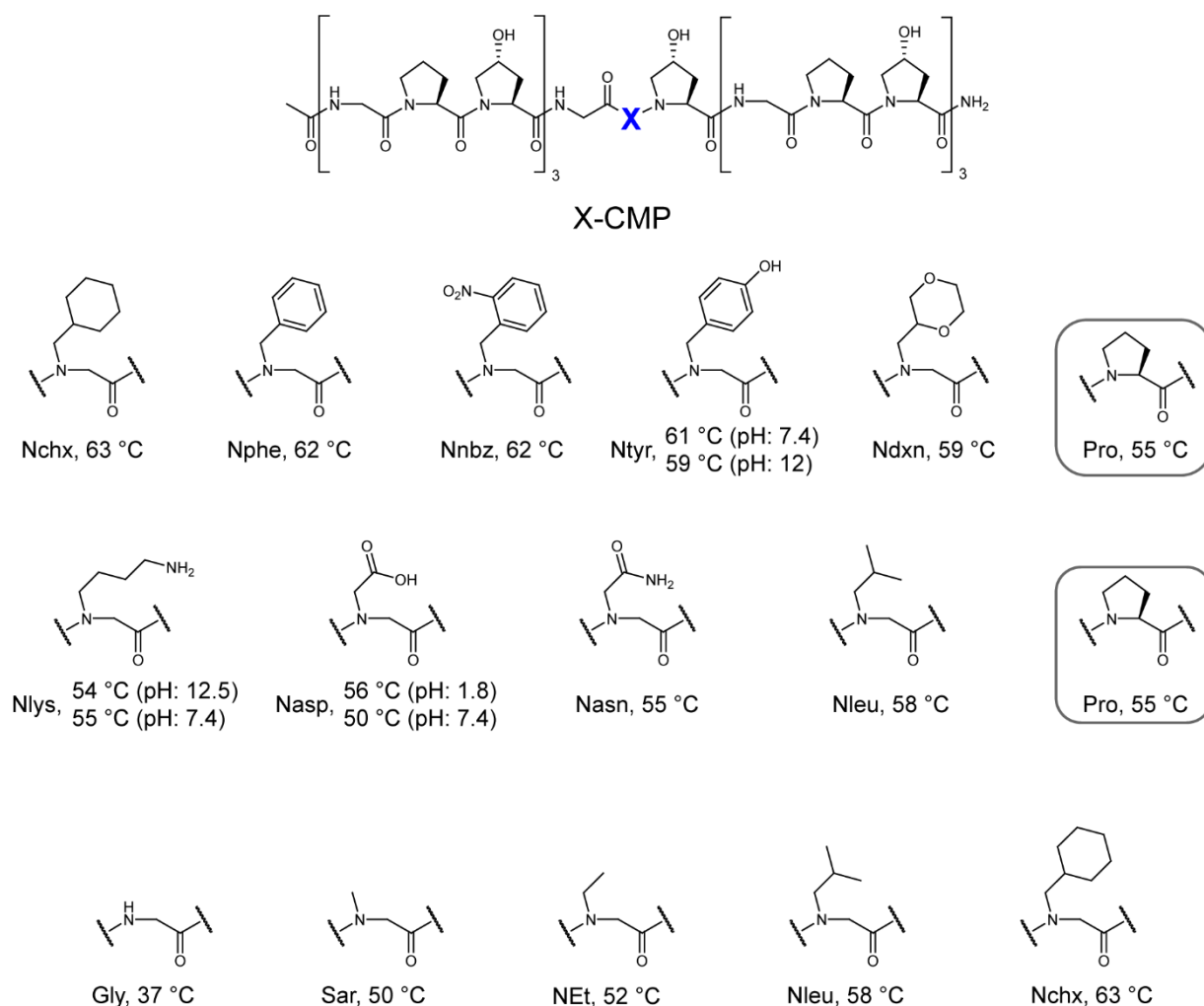


Fig. 3 | The structure of the peptoid sidechain affects triple-helical propensity. **a**, T_m values for a series of X-CMPs with N- α -substituted 6-membered rings in the sidechains. There is little difference in triple-helical stability among the aliphatic, aromatic, and hydrophilic rings, and every peptide had a T_m higher than the host: Pro-CMP. **b**, Although charged groups destabilized the triple helix at physiological pH, when the sidechains were uncharged, both Nlys and Nasp, as well as the hydrophilic Nasn, were as stabilizing as Pro. **c**, A series of X-CMPs with aliphatic sidechains of increasing size showing concomitantly increasing T_m values, implying an robust steric effect. Although the hydrophobicity in this series increases as the aliphatic sidechains increase in size, according to series **a** and **b**, hydrophobicity has only a minor effect on stability. The results suggest that the tertiary amide structure N-gly (which hinders the α helix and β sheet conformations) and the size of the N-substituted sidechain play the most critical roles in stabilization of N-gly-CMP triple helices.

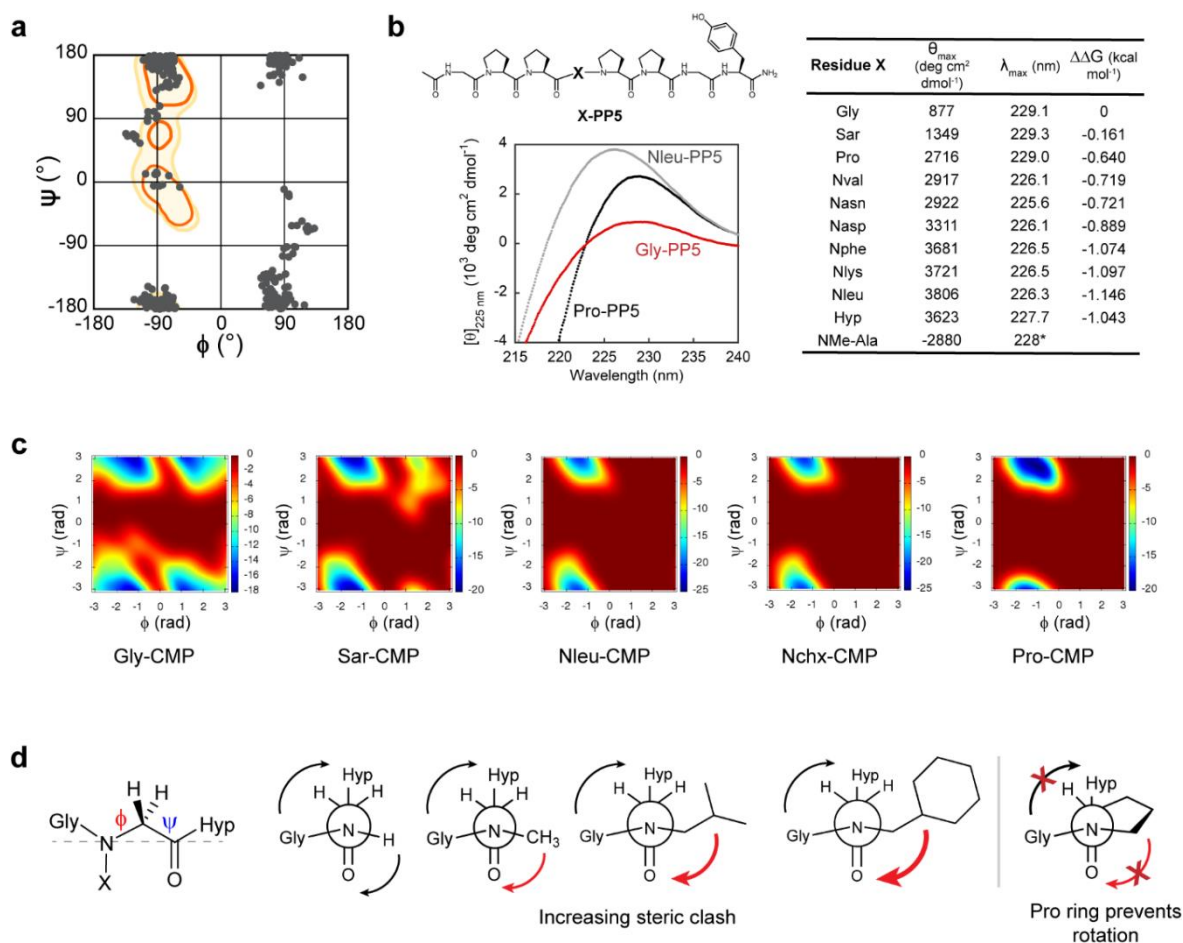


Fig. 4 | N-gly residues have a strong polyproline-II helix (PPII) propensity which improve the triple-helical stability of X-CMPs. **a**, A survey of ϕ and ψ angles of N-glys from high resolution crystal structures of small molecules (from Cambridge Structure Database). N-glys have a high natural propensity to fold into either the left- or right-handed polyproline conformation ($\phi \pm 90^\circ$, $\psi 180^\circ$). Watermarks show the regions highly populated by Pro residues. **b**, Structure of the PPII host-guest peptide **X-PP5** and the comparative CD spectra of Gly-PP5, Pro-PP5, and Nleu-PP5. The PPII propensity of each X residue within the X-PP5 model was assessed by the ellipticity value of the characteristic CD signal near 228 nm. All N-glys, with the exception of Sar, had a θ_{\max} higher than Pro indicating peptoids' high intrinsic PPII propensity. *No peak was observed near 228 nm for NMe-Ala (Supplementary data). **c**, Metadynamics calculations of the free energy landscape of the guest residues in Gly-, Sar-, Nleu-, Nchx-, and Pro-CMP. Each model yielded a native state consistent with the expected triple helix and PPII conformations (-70° , 170°); however, peptides with small guest residues (Gly and Sar) showed a secondary non-PPII energy well. **d**, Newman-like projections of N-glys with the grey dashed line indicating the line of sight for the projection. N-glys with large side chains can experience major steric clashes with its carbonyl group. Such reduced conformational flexibility of the N-gly may help preorganize the individual chains into a PPII conformation, thereby stabilizing the triple-helical assembly.

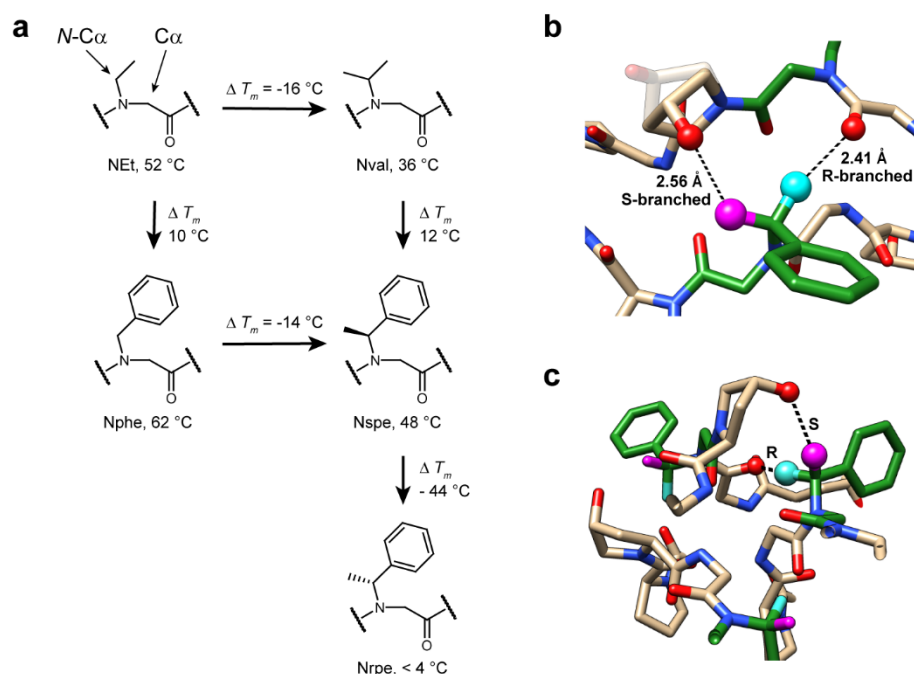


Fig. 5 | *N*-C α branching of peptoid residues affects triple-helical folding. **a**, With a *N*-C α branch, Nval- and Nspe-CMPs, had T_m values 14–16 °C lower than their unbranched analogs, NEt and Nphe (ΔT_m : -14 ~ -16 °C). Moreover, while residue Nspe conferred a triple helix with a T_m of 48 °C, its enantiomer, Nrpe completely abolished the triple-helical folding. **b,c**, Molecular modeling based on solved crystal structure of Nphe-CMP showing potential steric clashes within the triple helix introduced by the *N*-C α methyl branch of Nspe and Nrpe. The *R*-branch of Nrpe (cyan), pointing directly toward the inner core of the triple helix, may clash with the backbone carbonyl of the cross-chain Gly and interfere with proper backbone assembling; in contrast, the *S*-branch (magenta) in Nspe may clash with the hydroxyl group of a cross-chain Hyp, which may be less destabilizing since the steric repulsion may be alleviated by changing the Hyp's ring pucker (shown in half-transparency in **b**).

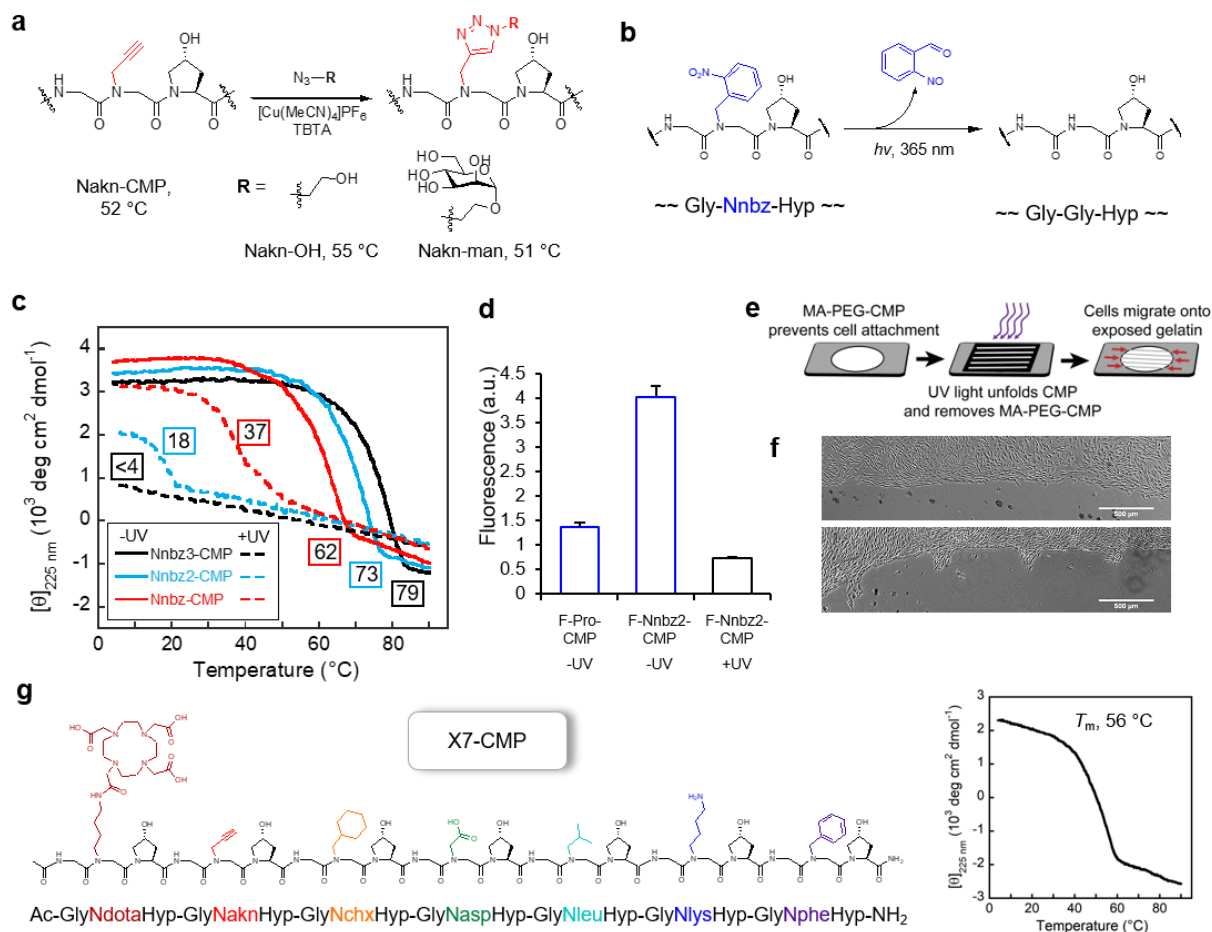


Fig. 6 | X-CMPs enable new functional design strategies. **a**, Facile modification of the alkyne-sidechain through alkyne-azide ‘click’ chemistry. Nakn-CMP retained strong triple-helical propensity after the conjugation, indicating that the N-substitution can well tolerate the spatially demanding triazole units within the triple helix (Supplementary Discussions). **b**, Reaction scheme of photo-conversion of the Nnbz residue to Gly by UV light. **c**, CD melting curves showing UV-induced triple helix destabilization of Nnbz-CMPs. A single Pro \rightarrow Nnbz substitution (Nnbz-CMP) in the central triad stabilized the triple helix (T_m : 55 $^{\circ}\text{C} \rightarrow$ 62 $^{\circ}\text{C}$), while photo-cleavage of the N-*o*-nitrobenzyl group destabilized the triple helix (T_m : 62 $^{\circ}\text{C} \rightarrow$ 37 $^{\circ}\text{C}$). Additional substitutions of Pro with Nnbz within the CMP (Nnbz2- and Nnbz3-CMPs, SI Table S7) further increased the ΔT_m of the peptide before and after UV irradiation. **d**, Comparative fluorescence intensity of gelatin films treated with fluorescently labeled X-CMPs. F-Nnbz2-CMP demonstrated higher affinity to gelatin than F-Pro-CMP, as well as triple-helical unfolding and releasing after UV irradiation. **e**, Schematics of UV-patterning on a gelatin substrate mediated by Nnbz2-CMP: MA-PEG-CMPs bound onto the crosslinked gelatin film prevented cell attachment; however, UV irradiation released the PEG-CMP conjugate, exposing gelatin that supports cell adhesion. **f**, Light micrographs of MDA-MB-231 cells grown to a confluency only in regions devoid of MA-PEG-CMP on a gelatin film (top). After exposing this

cell culture to UV light through a photomask (**e**), which unbound the MA-PEG-CMPs in selected areas, the cells began to infiltrate the patterned areas with exposed gelatin (bottom, taken 1 d after UV treatment). **g**, Pro-free X7-CMP hosting seven peptoid residues with diverse sidechain structures. Its high T_m of 56 °C demonstrates a hyper-stable triple-helical collagen peptide with the greatest sidechain diversity reported to date.

References

1. Horne, W.S. & Grossmann, T.N. Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns. *Nat. Chem.* **12**, 331-337 (2020).
2. Huang, P., Boyken, S.E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320-327 (2016).
3. Shoulders, M.D. & Raines, R.T. Collagen structure and stability. *Annu. Rev. Biochem.* **78**, 929-958 (2009).
4. Ramshaw, J.A.M., Shah, N.K. & Brodsky, B. Gly-X-Y tripeptide frequencies in collagen: A context for host-guest triple-helical peptides. *J. Struct. Biol.* **122**, 86-91 (1998).
5. Ramachandran, G.N. & Kartha, G. Structure of collagen. *Nature* **176**, 593-595 (1955).
6. 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 (1994).
7. Orgel, J.P.R.O., Irving, T.C., Miller, A. & Wess, T.J. Microfibrillar structure of type I collagen *in situ*. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9001 (2006).
8. Zitnay, J.L. et al. Molecular level detection and localization of mechanical damage in collagen enabled by collagen hybridizing peptides. *Nature Communications* **8**, 14913 (2017).
9. Leitinger, B. Transmembrane collagen receptors. *Annu. Rev. Cell. Dev. Biol.* **27**, 265-290 (2011).
10. O'Leary, L.E.R., Fallas, J.A., Bakota, E.L., Kang, M.K. & Hartgerink, J.D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nat. Chem.* **3**, 821 (2011).
11. Tanrikulu, I.C., Forticaux, A., Jin, S. & Raines, R.T. Peptide tessellation yields micrometre-scale collagen triple helices. *Nat. Chem.* **8**, 1008 (2016).
12. Holmgren, S.K., Taylor, K.M., Bretscher, L.E. & Raines, R.T. Code for collagen's stability deciphered. *Nature* **392**, 666-667 (1998).
13. Sakakibara, S., Kishida, Y., Kikuchi, Y., Sakai, R. & Kakiuchi, K. Synthesis of poly-(L-prolyl-L-prolylglycyl) of defined molecular weights. *Bull. Chem. Soc. Jpn.* **41**, 1273-1273 (1968).
14. Persikov, A.V., Ramshaw, J.A.M., Kirkpatrick, A. & Brodsky, B. Amino acid propensities for the collagen triple-helix. *Biochemistry* **39**, 14960-14967 (2000).
15. Maaßen, A. et al. Triple-helix-stabilizing effects in collagen model peptides containing PPII-helix-preorganized diproline modules. *Angew. Chem. Int. Ed.* **59**, 5747-5755 (2020).
16. Aronoff, M.R., Egli, J., Schmitt, A. & Wennemers, H. Alkylation of γ -Azaproline Creates Conformationally Adaptable Proline Derivatives for pH-Responsive Collagen Triple Helices. *Chemistry – A European Journal* **26**, 5070-5074 (2020).
17. Erdmann, R.S. & Wennemers, H. Functionalizable collagen model peptides. *J. Am. Chem. Soc.* **132**, 13957-13959 (2010).

18. Siebler, C., Erdmann, R.S. & Wennemers, H. Switchable proline derivatives: Tuning the conformational stability of the collagen triple helix by pH changes. *Angew. Chem. Int. Ed.* **53**, 10340-10344 (2014).
19. Hentzen, N.B., Smeenk, L.E.J., Witek, J., Riniker, S. & Wennemers, H. Cross-linked collagen triple helices by oxime ligation. *J. Am. Chem. Soc.* **139**, 12815-12820 (2017).
20. Goodman, M., Melacini, G. & Feng, Y. Collagen-like triple helices incorporating peptoid residues. *J. Am. Chem. Soc.* **118**, 10928-10929 (1996).
21. Feng, Y., Melacini, G. & Goodman, M. Collagen-based structures containing the peptoid residue N-Isobutylglycine (Nleu): synthesis and biophysical studies of Gly-Nleu-Pro sequences by circular dichroism and optical rotation. *Biochemistry* **36**, 8716-8724 (1997).
22. Melacini, G., Feng, Y. & Goodman, M. Collagen-based structures containing the peptoid residue N-Isobutylglycine (Nleu): conformational analysis of gly-nleu-pro sequences by ¹H-NMR and molecular modeling. *Biochemistry* **36**, 8725-8732 (1997).
23. Simon, R.J. et al. Peptoids: a modular approach to drug discovery. *Proceedings of the National Academy of Sciences* **89**, 9367 (1992).
24. Robertson, E.J. et al. Design, synthesis, assembly, and engineering of peptoid nanosheets. *Acc. Chem. Res.* **49**, 379-389 (2016).
25. Wu, C.W., Sanborn, T.J., Huang, K., Zuckermann, R.N. & Barron, A.E. Peptoid oligomers with α -chiral, aromatic side chains: sequence requirements for the formation of stable peptoid helices. *J. Am. Chem. Soc.* **123**, 6778-6784 (2001).
26. Wu, C.W. et al. Structural and spectroscopic studies of peptoid oligomers with α -chiral aliphatic side chains. *J. Am. Chem. Soc.* **125**, 13525-13530 (2003).
27. Stringer, J.R., Crapster, J.A., Guzei, I.A. & Blackwell, H.E. Extraordinarily robust polyproline type I peptoid helices generated via the incorporation of α -chiral aromatic N-1-Naphthylethyl side chains. *J. Am. Chem. Soc.* **133**, 15559-15567 (2011).
28. Roy, O. et al. Homogeneous and robust polyproline type I helices from peptoids with nonaromatic α -chiral side chains. *J. Am. Chem. Soc.* **139**, 13533-13540 (2017).
29. Shah, N.H. et al. Oligo(N-aryl glycines): a new twist on structured peptoids. *J. Am. Chem. Soc.* **130**, 16622-16632 (2008).
30. Lee, B.-C., Zuckermann, R.N. & Dill, K.A. Folding a nonbiological polymer into a compact multihelical structure. *J. Am. Chem. Soc.* **127**, 10999-11009 (2005).
31. Murnen, H.K., Rosales, A.M., Jaworski, J.N., Segalman, R.A. & Zuckermann, R.N. Hierarchical self-assembly of a biomimetic diblock copolypeptoid into homochiral superhelices. *J. Am. Chem. Soc.* **132**, 16112-16119 (2010).
32. Mannige, R.V. et al. Peptoid nanosheets exhibit a new secondary-structure motif. *Nature* **526**, 415-420 (2015).
33. Nam, K.T. et al. Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nature Materials* **9**, 454-460 (2010).
34. Zuckermann, R.N., Kerr, J.M., Kent, S.B.H. & Moos, W.H. Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* **114**, 10646-10647 (1992).
35. Zhang, Y., Malamakal, R.M. & Chenoweth, D.M. Aza-glycine induces collagen hyperstability. *J. Am. Chem. Soc.* **137**, 12422-12425 (2015).
36. Adzhubei, A.A., Sternberg, M.J.E. & Makarov, A.A. Polyproline-II helix in proteins: Structure and function. *J. Mol. Biol.* **425**, 2100-2132 (2013).
37. Brown, A.M. & Zondlo, N.J. A propensity scale for type II polyproline helices (PPII): aromatic amino acids in proline-rich sequences strongly disfavor PPII due to proline–aromatic interactions. *Biochemistry* **51**, 5041-5051 (2012).

38. Pandey, A.K., Thomas, K.M., Forbes, C.R. & Zondlo, N.J. Tunable control of polyproline helix (PPII) structure via aromatic electronic effects: an electronic switch of polyproline helix. *Biochemistry* **53**, 5307-5314 (2014).
39. Horng, J.-C. & Raines, R.T. Stereoelectronic effects on polyproline conformation. *Protein Sci.* **15**, 74-83 (2006).
40. Jenkins, C.L., Vasbinder, M.M., Miller, S.J. & Raines, R.T. Peptide bond isosteres: Ester or (E)-alkene in the backbone of the collagen triple helix. *Org. Lett.* **7**, 2619-2622 (2005).
41. Butterfoss, G.L., Renfrew, P.D., Kuhlman, B., Kirshenbaum, K. & Bonneau, R. A preliminary survey of the peptoid folding landscape. *J. Am. Chem. Soc.* **131**, 16798-16807 (2009).
42. Crapster, J.A., Guzei, I.A. & Blackwell, H.E. A peptoid ribbon secondary structure. *Angewandte Chemie (International ed. in English)* **52**, 5079-5084 (2013).
43. Delauney, A.J. & Verma, D.P.S. Proline biosynthesis and osmoregulation in plants. *The Plant Journal* **4**, 215-223 (1993).
44. Egli, J., Schnitzer, T., Dietschreit, J.C.B., Ochsenfeld, C. & Wennemers, H. Why proline? Influence of ring-size on the collagen triple helix. *Org. Lett.* **22**, 348-351 (2020).
45. Cram, D.J. The design of molecular hosts, guests, and their complexes. *Science* **240**, 760 (1988).
46. Li, Y. et al. Targeting collagen strands by photo-triggered triple-helix hybridization. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14767-14772 (2012).
47. Gorske, B.C., Stringer, J.R., Bastian, B.L., Fowler, S.A. & Blackwell, H.E. New strategies for the design of folded peptoids revealed by a survey of noncovalent interactions in model systems. *J. Am. Chem. Soc.* **131**, 16555-16567 (2009).

Acknowledgements

The authors thank Rodrigo Galindo-Murillo and Xiaolei Zhu for consultation on the simulations of X-CMP structures and Hendra Wahyudi for assistance in the synthesis of Fmoc-Nnbz-OH. This research was funded by grants from the National Institutes of Health (R01AR071358, R21EY029430, and R21OD026618) awarded to SMY.

Author contributions

Y.L. and J.L.K. designed the studies. J.L.K., Y.L., G.K., and H.K. performed the syntheses. J.L.K and Y.L. performed the experiments. J.L.K. and F.G.W. solved the crystal structures. C.P.H. oversaw the crystallography. Z.Q. and T.E.C. performed the simulations. J.L.K., S.M.Y., and Y.L. wrote the paper. All authors were involved in the completion of this work.

Competing interests

The authors declare no competing interests.

ChemRxiv PEPTOID-CMP_manuscript 20200905_submitt... (1.52 MiB) [view on ChemRxiv](#) • [download file](#)

Peptoid Residues Make Diverse, Hyperstable Collagen Triple Helices

Julian L. Kessler¹, Grace Kang¹, Zhao Qin², Helen Kang¹, Frank G. Whitby³, Thomas E. Cheatham III⁴, Christopher P. Hill³, Yang Li^{1,*} and S. Michael Yu^{1,5}

¹Department of Biomedical Engineering, University of Utah, Salt Lake City, Utah 84112, USA

²Department of Civil & Environmental Engineering, College of Engineering & Computer Science, Syracuse University, Syracuse, New York 13244, USA

³Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

⁴Department of Medicinal Chemistry, College of Pharmacy; University of Utah, Salt Lake City, 84112, USA

⁵Department of Pharmaceutics and Pharmaceutical Chemistry, University of Utah, Salt Lake City, Utah 84112, USA

*Correspondence to: Yang Li (yang.d.li@utah.edu)

Table of Contents

Supplementary Section 1: Materials and Methods	S2
Materials	S2
Solid phase synthesis	S3
On-resin functionalization <i>via</i> “click” chemistry	S5
Cleavage protocol	S5
Purification and mass spectrometry	S5
Circular dichroism spectroscopy	S6
X-ray Crystallography	S6
MD Simulations	S7
Gelatin binding assays	S7
UV triggered hydrogel dissolution and spatial control of cell attachment	S8
Supplementary Section 2: Figures and Tables	S10
Supplementary Table S1. Chemical structures of all guest residues and their T_m .	S10
Supplementary Table S2. X-ray crystallography parameters.	S11
Supplementary Table S3. Existing PDB entries of collagen model peptides.	S12
Supplementary Table S4. Primary structures and T_m of X-CMP with double guests.	S13
Supplementary Table S5. Existing CSD entries of peptoid-containing small molecules	S14
Supplementary Table S6. Examples of crystal structures of peptoids in CSD	S15
Supplementary Table S7. Primary structures and T_m of Nnbz-, Nnbz2-, and Nnbz3-CMP.	S16
Supplementary Fig. S1. Crystal structure of Nlys-CMP.	S17
Supplementary Fig. S2. Crystal structure of Nphe-CMP.	S18
Supplementary Fig. S3. Ramachandran plots for Nlys- and Nphe-CMP.	S19
Supplementary Fig. S4. Newman-like projection of NMe-Ala.	S20
Supplementary Fig. S5. MA-PEG-CMP and UV-induced unfolding.	S21
Supplementary Fig. S6. Spatial control of cell adhesion.	S22
Supplementary Section 3: HPLC, MALDI and CD of all peptides	S23
Analyses of X-CMP peptides (in alphabetical order)	S23
Analyses of X-PP5 peptides (in alphabetical order)	S52
Supplementary Section 4: Supplementary Discussion	S58
References	S59

Supplementary Section 1: Materials and methods

Materials

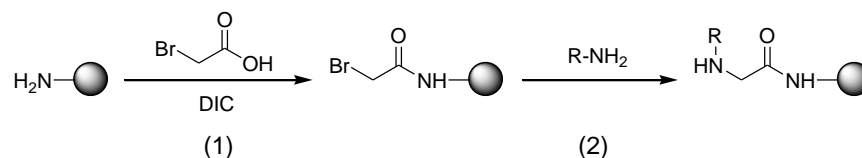
All commercial chemicals and solvents were used as received. The solvents, resin, Fmoc amino acids, and reagents used in the synthesis and purification of the peptides and peptoids were purchased from the suppliers listed in the table below.

General reagents	Supplier	Catalog number
<i>N</i> -Methylpyrrolidinone (NMP)	Fisher Scientific	BP1172
<i>N,N</i> -Dimethylformamide (DMF)	Fisher Scientific	BP1160
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	276855
Methylene Chloride (DCM)	Fisher Scientific	D37
Acetonitrile	Fisher Scientific	A998
TentaGel R RAM resin	Peptides International	RTS-9995-PI
Fmoc-Gly-OH	EMD Millipore	852001
Fmoc-Pro-OH	EMD Millipore	852017
Fmoc-Hyp(tBu)-OH	EMD Millipore	852036
Fmoc-Lys(Boc)-OH	AAPPTec	AFK105
Fmoc-Phe-OH	AAPPTec	AFF101
Fmoc-Tyr(tBu)-OH	EMD Millipore	852020
Fmoc-Ile-OH	Advanced ChemTech	FI2326
Piperidine	Sigma-Aldrich	104094
Bromoacetic acid	Sigma-Aldrich	17000
<i>N,N'</i> -Diisopropylcarbodiimide (DIC)	Sigma-Aldrich	D125407
O-(7-Azabenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate (HATU)	Chem-Impex	12881
1-Hydroxy-7-azabenzotriazole (HOAT)	AAPPTec	CXZ012
(7-Azabenzotriazol-1-yl)oxy)trispyrrolidinophosphonium hexafluorophosphate (PyAOP)	AAPPTec	CXZ070
<i>N</i> -Ethyl-diisopropylamine (DIEA)	EMD Millipore	845017
Acetic anhydride	Alfa Aesar	L04295
Trifluoroacetic acid (TFA)	Fisher Scientific	BP618
Triisopropylsilane (TIS)	Sigma-Aldrich	233781
1,2-Ethanedithiol (EDT)	Sigma-Aldrich	02390
DL-Dithiothreitol	Sigma-Aldrich	D0632
Anisole	Sigma-Aldrich	123226
Ethyl ether	Fisher Scientific	E138

Solid phase synthesis

All sequences were prepared on TentaGel R RAM resin (substitution level: 0.2 mmol/g) *via* solid phase synthesis. Typically, 50 mg of resin (containing 10 μ mol of reaction sites) was used to prepare one host-guest peptide/peptoid sequence. Prior to the first coupling, resin was swelled in 0.4 mL of NMP for 30 min. Each amino acid residue was coupled by agitating the resin with a solution of Fmoc protected residue (50 μ mol), HATU (50 μ mol, 19 mg), HOAT (50 μ mol, 6.8 mg) and DIEA (75 μ mol, 13 μ L) in 0.4 mL of NMP for over 3 hr. The Fmoc protective group was removed by treating the resin with piperidine (20% by volume in NMP) for 20 min. Following each reaction, the resin was drained and washed with NMP (3 \times 10 mL). All coupling and deprotection reactions were monitored by the standard Kaiser (for primary amines) and chloranil (for secondary amines) tests.

Except residues Sar, NMe-Ala, and Nnbz, which were coupled by using the respective Fmoc-protected compounds and the above-mentioned HATU chemistry, all other peptoid residues were incorporated on-resin using the two-step sub-monomer method reported by Zuckermann and others¹.



(1) Acylation reactions were performed by addition of a solution of bromoacetic acid (100 μmol , 13.9 mg) and DIC (98 μmol , 15.2 μL) in DMF (0.5 mL) to 10 μmol of resin-bound amine. Reaction mixtures were agitated at room temperature for 30 min. Each acylation was repeated once before the resin was drained and washed with NMP (3 \times 10 mL). (2) Each displacement reaction was performed by addition of a primary amine (200 μmol) in 0.5 mL of NMP, followed by agitation overnight at room temperature. Following the displacement, the resin was drained and washed with DMF (3 \times 10 mL). The reaction was monitored by the chloranil test. Typically, resin beads turned strongly green in a chloranil test following a displacement reaction of 2 hr, suggesting successful coupling of a peptoid; we chose to run the reactions overnight to ensure completion. In the table below, the primary amine used to create each peptoid residue is listed, and special reaction conditions and observations are noted. In general, the next amino acid following a peptoid residue (e.g., Fmoc-Gly-OH or Fmoc-Pro-OH) was coupled with full completion using the above-mentioned 5 eq HATU protocol.

Residue	R-NH ₂ Reagent	Supplier	Catalog number	Note
Sar	Fmoc-Sar-OH	EMD Millipore	852055	Coupled by Fmoc & HATU chemistry
NMe-Ala	Fmoc-N-methyl-L-alanine	Chem-Impex	2650	Coupled by Fmoc & HATU chemistry
Nnbz	Fmoc(<i>N</i> -o-nitrobenzyl)Gly-OH	Synthesized in-house ^{2,3}		Coupled by Fmoc & HATU chemistry
Nakn	Propargylamine	Sigma-Aldrich	P50900	H ₂ N-Gly-Nakn-peptide: in Kaiser test, resin beads and the solution do not turn dark blue despite the presence of primary amine, perhaps because the Nakn side chain affects the ninhydrin reactions.
Nasn	Glycinamide	Combi-Blocks	QA-8748	Glycinamide is not soluble in NMP. The displacement reaction was carried out in 0.5 mL DMSO.
Nasp	Glycine tert-butyl ester	Alfa Aesar	L16258	
Nchx	Cyclohexanemethylamine	Sigma-Aldrich	101842	
Ncys	2-(tert-butylsulfanyl)ethan-1-amine	Enamine	EN300-110537	
Ndxn	C-[1,4]Dioxan-2-yl-methylamine	Matrix Scientific	008403	
NEt	Ethylamine solution (2.0 M in THF)	Sigma-Aldrich	395072	
Nleu	Isobutylamine	Sigma-Aldrich	I14150	
Nlys	<i>N</i> -Boc-1,4-butanediamine	Chem-Impex	31317	
Nmet	2-(Methylthio)ethylamine	Sigma-Aldrich	632929	
Nphe	Benzylamine	Sigma-Aldrich	185701	
Nrpe	(R)-(+)- α -Methylbenzylamine	Sigma-Aldrich	115541	The Gly residue following Nrpe/Nspe was coupled by agitating the resin (10 μ mol) in a solution of Fmoc-Gly-OH (100 μ mol) and DIC (100 μ mol) in 0.5 mL of DMF.
Nspe	(S)-(-)- α -Methylbenzylamine	Sigma-Aldrich	115568	
Ntyr	(4-tert-butoxyphenyl)methanamine	Enamine	EN300-55516	
Nval	Isopropylamine	Sigma-Aldrich	471291	The Gly residue following Nval was coupled by agitating the resin (10 μ mol) in a solution of Fmoc-Gly-OH (100 μ mol) and DIC (100 μ mol) in 0.5 mL of DMF.
Ndota	1,4,7,10-Tetraazacyclododecane-1,4,7-tris(<i>t</i> -butyl acetate)-10-(4-aminobutyl)acetamide	Macrocyclics	B-279	The displacement reaction was completed by agitating the resin (10 μ mol) in 0.5 mL of NMP containing the primary amine (63 μ mol) and DIEA (115 μ mol) overnight.
Nman	2-Aminoethyl 2,3,4,6-tetra-O-acetyl- α -D-mannopyranoside hydrochloride	Synthese	AM876	DIEA (250 μ mol, 43.5 μ L) was added to the reaction mixture in the displacement reaction.

On-resin functionalization via “click” chemistry

The azido compound (60 μ mol), such as 2-Azidoethanol (Carbosynth, FA07084) or 2-Azidoethyl α -D-mannopyranoside (Synthos, AM482), was dissolved in 0.5 mL of DMF along with [Cu(MeCN)₄]PF₆ (10 μ mol, 3.7 mg, Sigma-Aldrich, 346276) and TBTA (20 μ mol, 10.6 mg, Sigma-Aldrich, 678937). This solution was added to the resin-bound alkyne-containing peptide Nakn-CMP (5 μ mol), and the mixture was agitated at room temperature overnight. The resin was drained and washed with DMF (5 \times 10 mL) before the “click”-functionalized product was cleaved.

Cleavage protocol

Following the coupling and Fmoc-deprotection of the last Gly residue, the *N*-terminal amine of each sequence was acetylated by mixing the resin (5 μ mol) with a solution of acetic anhydride (500 mM), HOAT (14.7 mM), and DIEA (129 mM) in 0.5 mL of NMP at room temperature for 30 min. The resin was drained, washed with NMP (3 \times 10 mL) and DCM (5 \times 10 mL). Unless otherwise mentioned below, the resin (5 μ mol) was treated with 1 mL mixture of TFA / TIS / water (95:2.5:2.5) for over 3 hr with stirring. Subsequently, the resin was washed with 0.5 mL of TFA twice. After the resin was filtered, the TFA cleavage solution was collected and evaporated under a stream of nitrogen down to approximately 0.5 mL. The crude peptide/peptoid products were precipitated by adding 5 mL of cold ethyl ether to the TFA solution on ice.

Sequence	Special cleavage protocol
Nmet	Treat the resin with a mixture of TFA/EDT/water/TIS (94:2.5:2.5:1) for over 3 h.
Nhcy	<i>Step 1:</i> for 5 μ mol of crude product, treat the resin with a mixture (1 mL) of TFA/TIS/water (95:2.5:2.5) for over 2 hr, and collect the crude product.
	<i>Step 2:</i> Remove the tBu protective group from the <i>N</i> cys residue by stirring the crude product in a mixture (3 mL) of TFA/DMSO/anisole (97.9:2:0.1) at room temperature for 1 hr, precipitate and collect the product again.
	<i>Step 3:</i> dry the crude product and dissolve it in 2 mL of water containing 15 mg of dithiothreitol to reduce any oxidized thiol groups before purification.

Purification and mass spectrometry

All peptides were purified by reverse-phase high performance liquid chromatography (HPLC) on an Agilent Zorbax SB C-18 column, using a mixture of water (A) and acetonitrile (B) (linear gradient: 5% to 35% acetonitrile over 30 min). Both eluents A and B contain 0.1% TFA. The column oven was heated to 70 °C to prevent triple helix formation. After semi-preparative purifications, the collected fractions containing the target peptides were analyzed by analytical HPLC again using the following gradient parameters.

Time (min)	0	2	4	25	25.5	30
B%	5%	5%	10%	40%	5%	5%

The flow rates used for semi-preparative and analytical HPLC were 4 and 1.5 mL/min, respectively. The pure fractions were lyophilized. The lyophilized products were reconstituted in pure water as stock solutions to be used for structural characterizations. Peptide concentrations in the stock solutions were determined by measuring the ultraviolet (UV) absorbance of the solutions at 214 nm (for the collagen mimetic host-guest sequences, extinction coefficient: 2200 M⁻¹cm⁻¹ per peptide bond) or 280 nm (for the Tyr-containing polyproline host-guest sequences, extinction coefficient: 1280 M⁻¹cm⁻¹ per tyrosine residue) on a SpectraMax M2e microplate reader (Molecular Devices) using a quartz cell with 1 cm cell path length.

All purified peptides/peptoids were verified by mass spectrometry. Mass spectra were obtained at the University of Utah Mass Spectrometry and Proteomics core facility on a Bruker UltrafleXtreme matrix-assisted laser desorption/ionization time-of-flight (MALDI/TOF) mass spectrometer.

Circular dichroism (CD) spectroscopy

CD measurements of the peptide solutions were recorded in quartz cells with a path length of 0.1 cm, on a JASCO J-1500 CD spectrophotometer. Prior to CD spectra or melting experiments, peptide stock solutions were diluted to 150 μM in 1×PBS (pH: 7.4 for triple helical peptides) or in 5 mM phosphate buffer (pH: 7.0 for polyproline host-guest peptides) and heated at 80 °C for 5 min followed by incubation at 4 °C for at least 48 hr. Sequences containing guest residues with acidic or basic side chains were also diluted to 150 μM in NaOH (10 or 35 mM) or HCl (15 mM) solutions.

CD spectra were scanned at 4 °C (for triple helical peptides) or 25 °C (for polyproline host-guest peptides) using the following parameters: bandwidth, 5 nm; digital integration time, 16 s; scanning speed, 20 nm/min; data pitch, 0.1 nm. For the polyproline host-guest series, all reported spectra were the average of three independent scans, and were corrected from the blank buffer background.

Thermal melting curves were obtained by monitoring the ellipticity at 225 nm from 4 °C to 90 °C at a heating rate of 0.5 °C/min. The mean residue ellipticity (MRE, $[\theta]$) was calculated using the equation $[\theta] = (\theta \times m) / (c \times l \times n)$, where θ is measured ellipticity (mdeg), m is molecular weight (g/mol), c is concentration (mg/mL), l is path length of the cuvette (mm), and n is the number of amino acid residues in the peptide. The derivative of a melting curve was generated using the JASCO Spectra Manager software (Version 2.10.05), and the temperature at the minimum of the derivative curve was defined as the melting temperature (T_m). Each T_m value reported in this study was averaged from two CD thermal unfolding experiments, in which the difference between the two measured T_m values was less than 1 °C for all triple helical host-guest peptides.

X-ray crystallography

Purified, lyophilized Nlys-CMP [Ac-(GlyProHyp)₃-GlyNlysHyp-(GlyProHyp)₃] was dissolved in water to a final concentration of 5.0 mg/mL. Crystals were grown by vapor diffusion in sitting drops of 0.3 μL of the peptide

solution and 0.3 μ L of 2.1 M DL-malic acid (pH 7.0) corresponding to well F8 in the commercially-available crystal screen JCSG-*plus* (Molecular Dimensions) at 21 °C. Crystals appeared within 10 days. Purified, lyophilized Nphe-CMP [Ac-(GlyProHyp)₃-GlyNpheHyp-(GlyProHyp)₃] was dissolved in water to a final concentration of 10.0 mg/mL. Crystals were grown by vapor diffusion in sitting drops of 0.3 μ L of the peptide solution and 0.3 μ L of 2.4 M sodium malonate dibasic monohydrate (pH 7.0) corresponding to well F9 in the commercially available crystal screen JCSG-*plus* (Molecular Dimensions) at 4 °C. Crystals appeared within 1 week.

In preparation for data collection, crystals were suspended in a small rayon loop attached to a mounting pin, immersed in 20 μ L crystallization buffer with 25% added glycerol, then cryocooled by plunging into liquid nitrogen. Data were collected from crystals maintained at 100 K. Diffraction data were collected on beam lines 12-2 (Nlys-CMP) and 9-2 (Nphe-CMP) at the Stanford Synchrotron Radiation Lightsource (SSRL). The resulting data were integrated and scaled using HKL2000⁴. Phases were determined by molecular replacement with Phaser-MR⁵ using 1G9W as a search model. Models were built with COOT⁶ and refined with Refmac5⁷. UCSF Chimera version 1.14 was used to render molecular structure figures. Data and refinement statistics are indicated in Supplementary Table S3.

Molecular Dynamics (MD) Simulations

We built the molecular structure and atomic interaction of different peptoids by combining DFT calculation with CHARMM General Force field (CHARMM GenFF)⁸. The standard CHARMM27 force field⁹ was updated to enable the modeling and simulation of peptide-peptoid hybrid system. To be more specific, the DFT calculation for each peptoid was performed by using HyperChem (Hypercube, Inc. Gainesville, FL, USA), which uses the 3-21G basis set¹⁰ to represent the electronic wave function. Geometric optimization with the conjugate gradient method was used to find the minimized coordinates of all the atoms. These optimized structures were used to make automatic analogy with the CHARMM GenFF to define the charge distribution within the molecule, and parameters of bond, angle and dihedral angle and vdw parameters. All the parameters had penalty score generally lower than 10, indicating the analogy was fair⁸. The intrinsic coordinate of each peptoid molecule was also defined according to the optimized structure, ensuring the initial geometry had optimum energy.

Gelatin binding assays

Fluorescently labeled peptides F-Pro-CMP and F-Nnbz2-CMP [sequences: F-GlyGlyGly-(GlyProHyp)₇, F-Ahx-(GlyProHyp)₂(GlyNnbzHyp)₂(GlyProHyp)₃; GlyGlyGly: spacer, Ahx: aminohexanoic acid, spacer] were prepared by reacting the N-terminal amines of the sequences on-resin with 6 molar equivalents of 5(6)-carboxyfluorescein (designated as F-, Sigma-Aldrich, 21877) activated by 6 molar equivalents of PyAOP with 12 molar equivalents of DIEA in NMP for over 24 hr. The labeled peptides were cleaved from the resin, purified by HPLC with protection

from ambient light, and analyzed by MALDI [F-Pro-CMP MALDI predicted mass: 2417, observed mass: 2440 (+Na); F-Nnbz2-CMP predicted mass: 2549, observed mass: 2572 (+Na)]. The pure fluorescent peptides were lyophilized and dissolved in 1×PBS.

Wells of a 96-well plate were coated with approximately 6 μ L of 85 °C gelatin (Sigma-Aldrich, G2500) solution in 1×PBS (10% w/v) and incubated at 4 °C for 15 min to allow gelation. The thin gelatin hydrogel films were crosslinked with a MES buffered solution (pH 4.7, 100 μ L/well) containing 2 mM of NHS and 10 mM EDC with shaking overnight at room temperature. The crosslinked gelatin films were washed with 200 μ L of 25 °C 1×PBS 10 times. PBS solution (50 μ L) containing 10 μ M of F-Pro- or F-Nnbz2-CMP was heated at 85 °C for 15 min to dissociate the triple helices before being immediately added to each well and allowed to bind for 1 hr at room temperature. Following binding, the plate was washed three times with 1×PBS at room temperature before the fluorescence (ex: 489 nm, em: 533 nm) was measured on a SpectraMax M2e microplate reader. Subsequently, half the wells treated with F-Nnbz2-CMP were exposed to UV light for 15 min (mercury arc lamp, 365 nm, 15.5 mW/cm²). Afterwards, the gelatin hydrogels were washed with 1×PBS at room temperature (3×30 sec, 4×15 min) before their fluorescence was measured again. The binding experiment for each peptide was run three times with protection from ambient light.

UV-triggered hydrogel dissolution and spatial control of cell attachment

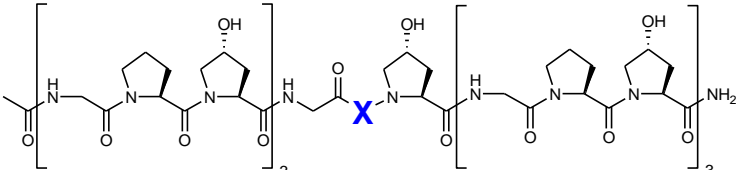
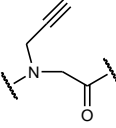
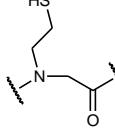
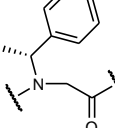
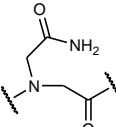
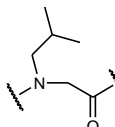
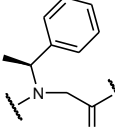
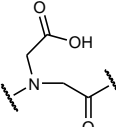
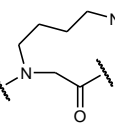
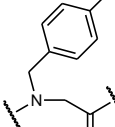
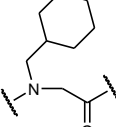
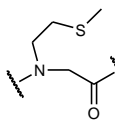
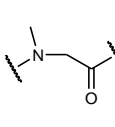
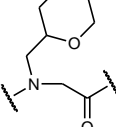
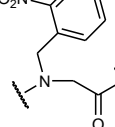
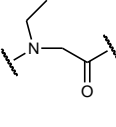
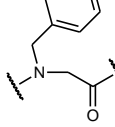
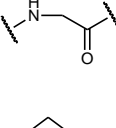
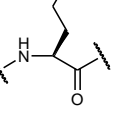
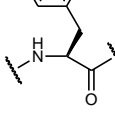
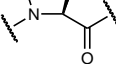
To prepare multi-arm (MA) PEG-CMPs, 8-arm PEG-Mal-40000 (Jenkem) was dissolved in 1×PBS to produce an 8% solution and added directly to CysGlyGlyGly-(GlyProHyp)₂-(GlyNnbzHyp)₂-(GlyProHyp)₃ [MALDI predicted mass: 2352.3 Da, observed mass: 2353.1 Da] (8 eq, dry powder). Gelation was observed almost immediately. Conjugation was confirmed on HPLC (70% X-CMP conjugated to polymer, Supplementary Fig. S5). Fluorescently labeled gels were prepared by dissolving the dry MA-PEG-CMP in a solution containing 650 μ M F-Nnbz2-CMP and 1×PBS. To observe UV-induced unfolding and dissolution of the hydrogel, the Nnbz2^X-CMP containing MA-PEG-CMP hydrogel was melted by heating to 80 °C and 50 μ L of the melted gel was added to the bottom of a 1.7 mL conical tube and allowed to gel at 4 °C for 15 min. After gelation, 250 μ L of 1×PBS (25 °C) was added atop the green colored gel. No diffusion of fluorescence into the PBS was observed. The gel was then exposed to UV light (Mercury arc lamp, 365 nm, 33 mW/cm²) for 30 min until the gel was no longer visible at the bottom of the tube. After UV irradiation, the green color fully permeated the solution indicating release of the Nnbz2^X-CMP and dissolution of the gel (Supplementary Fig. S5).

To utilize the MA-PEG-CMP to spatially control cell attachment and migration, we prepared a thin gelatin film on the bottom of a 6-well plate (size: ~1 cm², thickness: ~50 μ m) and crosslinked it with EDC/NHS. MA-PEG-CMP was melted by heating to 80 °C and diluted in hot 1×PBS to a final concentration of 31.25 μ M. The MA-PEG-CMP solution (85 μ L, 80 °C) was applied to the film surface and allowed to dry completely. Wells patterned before cell addition (Supplementary Fig. S6) were exposed to UV light through a transparency mask placed directly onto the

film surface for 10 min (mercury arc lamp, 365 nm, 15.5 mW/cm²). Patterns were produced using CorelDraw and printed onto transparencies by CAD/Art Services (Brandon, OR). Films were then washed with 1×PBS and sterilized with 70% ethanol by immersion in water for 1 hr and washed with sterile 1×PBS. MDA-MB-231 cells (1.4×10^4 cells/mL) in 3 mL of Dulbecco's modified eagle medium (supplemented with 10% fetal bovine serum and 1% mix of penicillin-streptomycin) were added to each well and were incubated at 37 °C in a 5% CO₂ atmosphere. Growth media was changed every 2 days. Cell attachment was monitored with an EVOS light microscope. Once cells grew to 80% confluency (2 days), unpatterned wells were drained of growth media and exposed to UV light (Mercury arc lamp, 365 nm, 33 mW/cm²) through a photomask elevated by thin spacers (0.9 mm, to prevent contact with the film surface) for 8 min. After patterning, wells were briefly washed with 1×PBS, growth media was replaced, and cells were incubated for an additional 24 hr.

Supplementary Section 2: Supplementary Tables and Figures

Supplementary Table S1. Name, structure, and T_m of the host-guest CMPs used in this study.

								
Name		Structure	T_m (°C)					
N-gly residues								
Nakn		52	Nhcy		54	Nrpe		< 4
Nasn		55	Nleu		58	Nspe		48
Nasp		50 (PBS) 56 (HCl)	Nlys		55 (PBS) 54 (NaOH)	Ntyr		61 (PBS) 59 (NaOH)
Nchx		63	Nmet		55	Sar		50
Ndxn		59	Nnbz		62			
NEt		52	Nphe		62			
Amino Acids								
Gly		37	Lys		47 (PBS) 46 (NaOH)	Phe		38
Pro		55						

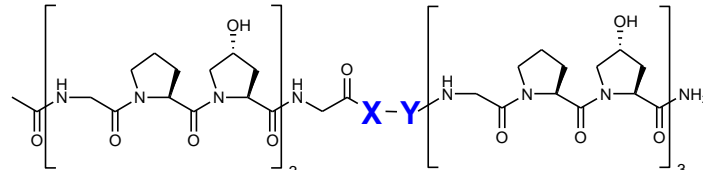
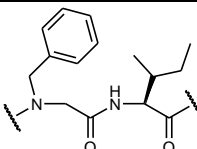
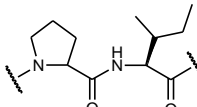
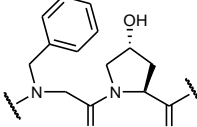
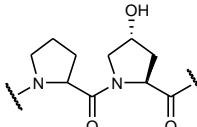
Supplementary Table S2. X-CMP crystallographic data and refinement statistics

Data		
Peptoid guest residue	N-Lysine	N-Phenylalanine
X-ray Source/Wavelength	SSRL 12-2 / 0.7500	SSRL 9-2 / 0.9795
Data processing software	HKL2000	XDS
Space Group	C2	P1
Unit cell dimensions		
a, b, c (Å)	72.34, 24.76, 25.36	20.44, 31.33, 35.41
α , β , γ (°)	90.00, 98.72, 90.00	88.72, 74.96, 89.79
Resolution (Å)	40.0 – 0.95	34.19 – 1.10
Resolution (Å) - (high-resolution shell)	(0.98 – 0.95)	(1.12 – 1.10)
# Reflections measured	579681	446025
# Unique reflections	24848	31530
Redundancy	23	14
Completeness (%)	88.9 (72.3)	90.6 (66.3)
$\langle I/\sigma I \rangle$	8.0 (0.9)	17.8 (1.4)
CC(1/2)	0.999 (0.734)	0.999 (0.459)
Mosaicity (°)	0.44	0.30
R _p im	0.019 (0.344)	0.029 (1.237)
Refinement		
Refinement software program	Refmac5	Refmac5
Resolution (Å)	13.62 – 0.95	34.2 – 1.10
Resolution (Å) - (high-resolution shell)	(0.977 – 0.952)	(1.125 – 1.096)
# Reflections used for refinement	23613	29490
# Reflections in R _{free} set	1221	2021
R _{cryst}	0.139 (0.319)	0.154 (0.289)
R _{free}	0.158 (0.329)	0.198 (0.371)
RMSD: bonds (Å) / angles (°)	0.013 / 1.848	0.014 / 2.010
$\langle B \rangle$ (Å ²): all atoms / # atoms	10.7 / 532	15.4 / 1129
$\langle B \rangle$ (Å ²): water molecules / #water	16.5 / 115	23.6 / 254

Supplementary Table S3. Protein Data Bank (PDB) IDs for collagen crystal structures used in backbone analysis.

1BKV	1V6Q	3WN8
1CAG	2D3H	4AXY
1CGD	2V53	4DMT
1EI8	2V53	4LOR
1G9W	3A1H	4Z1R
1V4F	3P46	

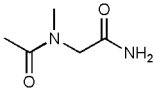
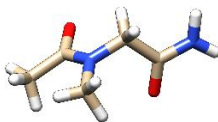
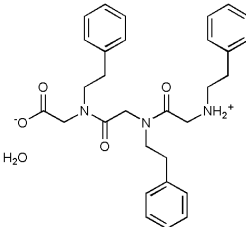
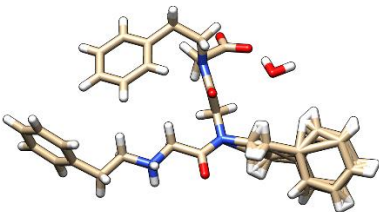
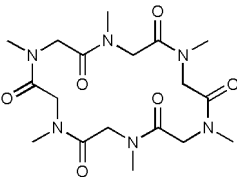
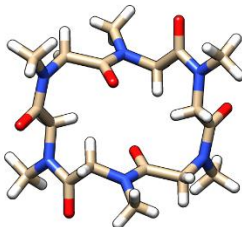
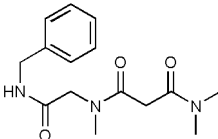
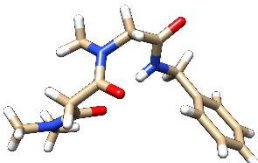
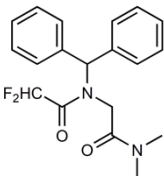
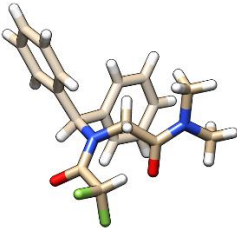
Supplementary Table S4. X-CMP sequences with X and Y substitutions at the central repeat designed to investigate the effect of hydrophobic neighbors on triple helical stability. For these constructs, we selected Nphe and Ile to replace Pro and Hyp. Following Goodman's hypothesis that hydrophobic effects are the critical factor for inducing triple helix formation, we expected that the destabilization resulting from the Y position Hyp \rightarrow Ile substitution would be attenuated when X = Nphe, since hydrophobic interactions between the Nphe-Ile pair are stronger than Pro-Ile. Ile \rightarrow Hyp increased T_m by 10 or 11 °C and Pro \rightarrow Nphe increased T_m by 6 or 7 °C. Since individual substitutions resulted in changes that are nearly identical regardless of neighboring residues, we believe that adjacent hydrophobic interactions do not contribute significantly to the triple helix stability. These results further support the notion that hydrophobicity of peptoid is not the most important factor in stabilizing the collagen triple helix.

					
Name	Structure	T_m (°C)	Name	Structure	T_m (°C)
Nphe-Ile		51	Pro-Ile		45
Nphe-Hyp		62	Pro-Hyp		55

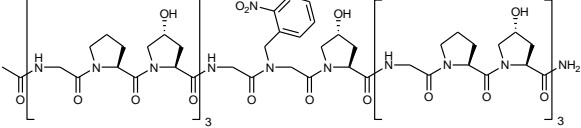
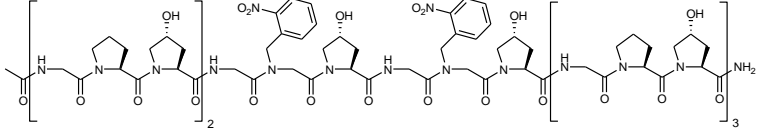
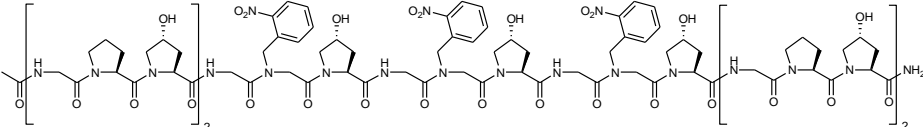
Supplementary Table S5. Reference codes for peptoid structures analyzed from the Cambridge Molecular Database (CSD).

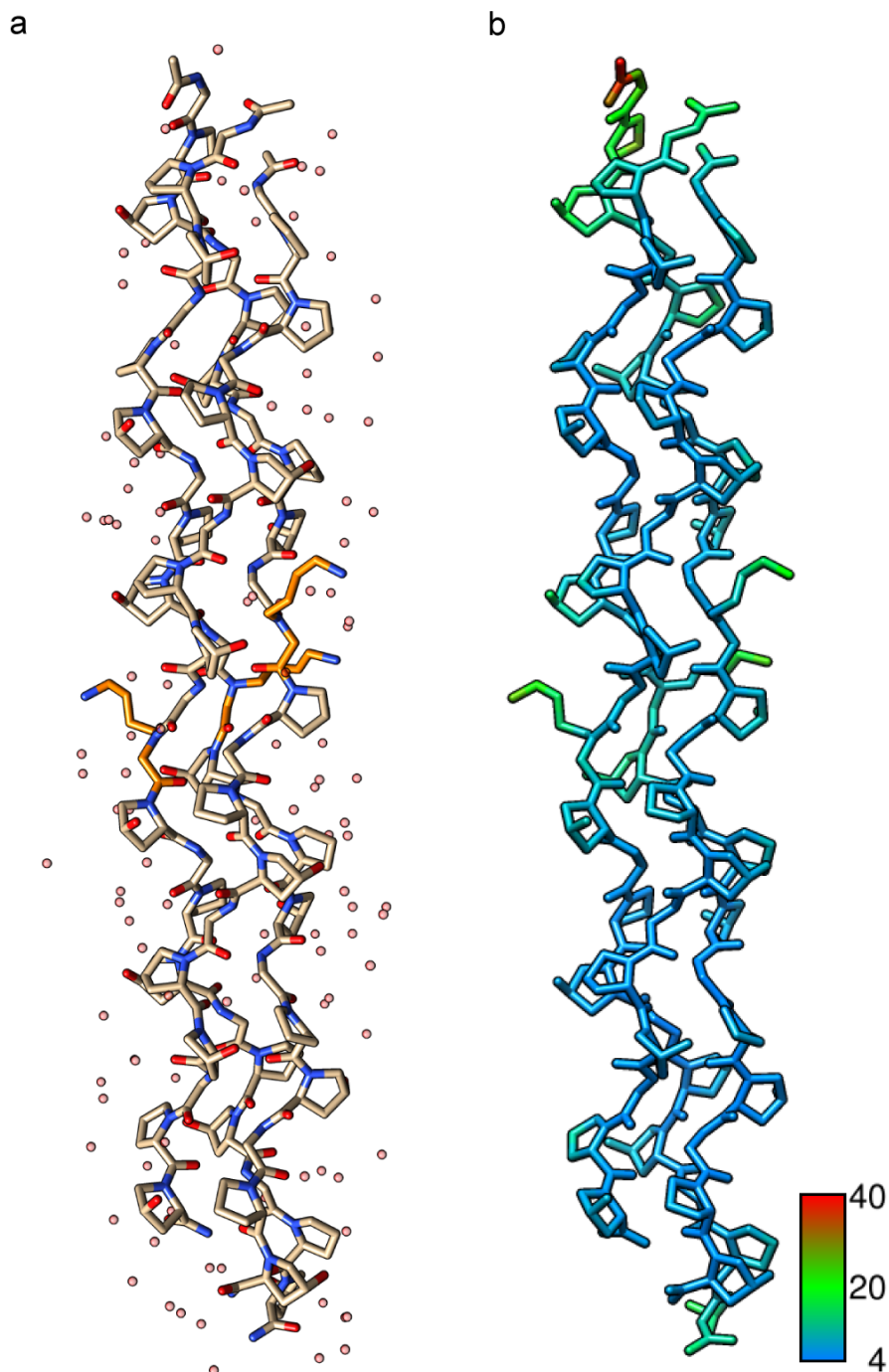
ACEVEM	CUCPAR	GIKBEJ	KALWAX	QECXEE	XEDKOH
ACEVIQ	CUCPEV	GLSARM	KEPNAU	QETQEO	YIYJUN
ACTDGU10	CUKGIA	HAXMOI	KEXZAQ	QETQIS	YOCWUK
ACUMOC	CUQRUB	HAXMOI10	LEBDEB	QICFEP	YOJNUH
ACUMUI	CUQSAI	HEYZAN	LIPDOD	QODTIO	YUHDIQ
ALASAR	CYDSAR	HEYZER	LOSLEL	ROLDUR	YUHDOW
ASUXOB	CYGSXS	HUCVIK	LOWGOT	TBXSAG	YUHDOC
BEHTEN	CYTSAK	HUCVOQ	MIJVEH	UKUTUP	YUHFAK
BIPWAY	DEDQAU	HXSARM	NAFJEJ	UMAYEM	YUHGIT
BIPWAY10	DEKSAN	ICYSPA	NAMPIA	UWUFUO	YUHGOZ
BIPWEC10	DETWUU	IGEJUB	NIZHIO	UXOYEN	ZADREB
BRAXGU	FIFSOE	IJAQIV	NUWNEY	UXOYOX	ZADRIF
CALSAR	FOLMUQ	IJAQIV10	OGAZOM	VACLOD	ZAIDUJ
CAZSEB	GABHEX	IPUQAN	OGAZUS	VACLUJ	ZAIZAO
CBBLPB10	GABXEN	IPUQER	PAGTIA	VEJGEY	ZATJEL
CHPSAR	GEDYOG	JOKMAX	PAGTOG	VEJGIC	ZEWREA
COSARC10	GEDYUM	KADFOL	PAGTUM	VEJGUO	ZEWRIE
CPSAYL10	GIDNUC	KALVIE	PIJPAZ	WEBZEI	
CTSARC	GIJZUW	KALVOK	POHUUU	WEXPOE	
CUCNUJ	GIKBAF	KALVUQ	POXPAT	WEXPUK	

Supplementary Table S6. Examples of peptoid-containing small molecule structures from CSD.

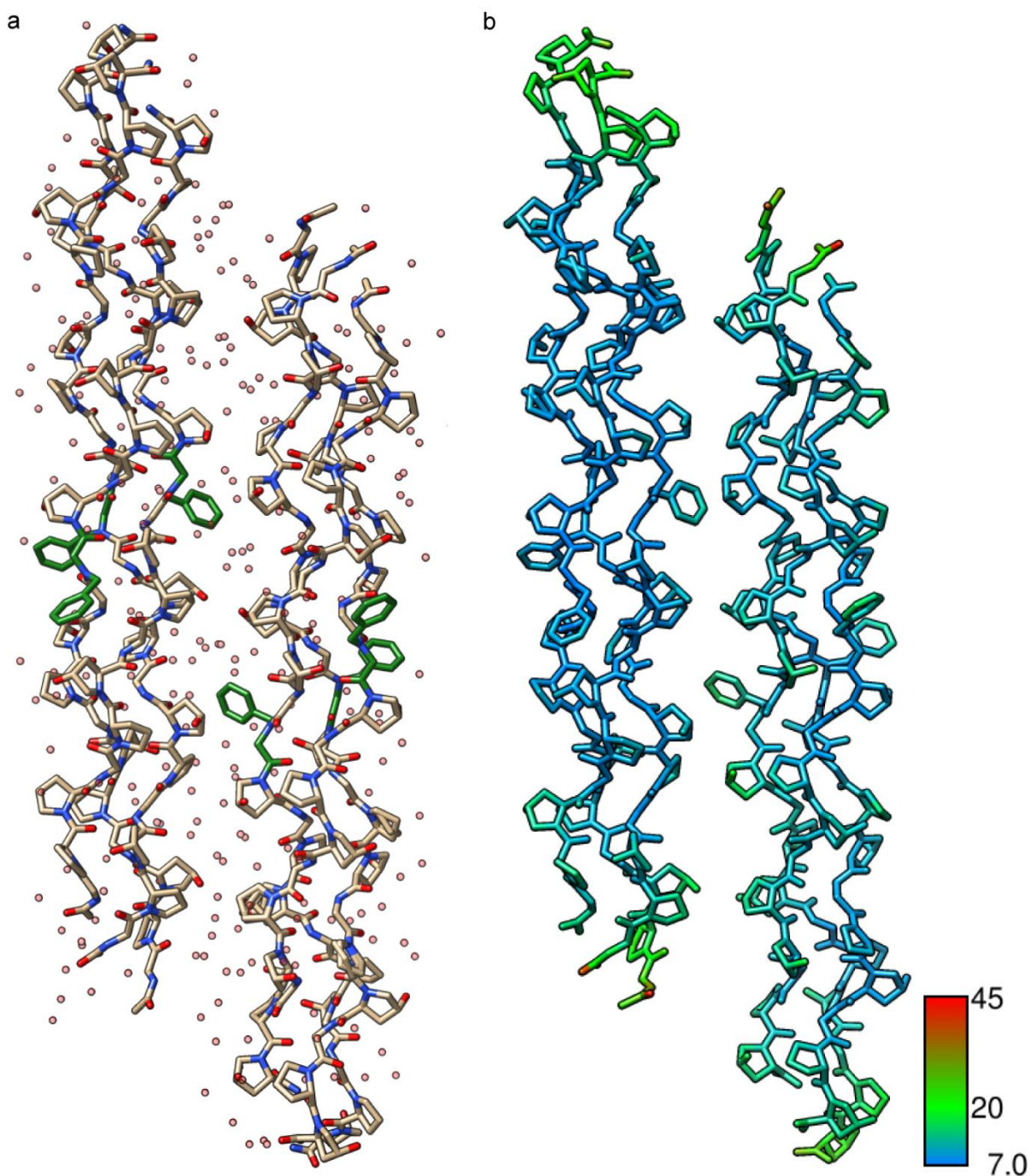
CSD- REFCODE	Chemical structure	Crystal structure
POXPAT		
ACUMUI		
HXSARM		
PAGTIA		
YUHGIT		

Supplementary Table S7. UV-responsive X-CMPs. Additional Nnbz groups increase T_m which indicates that peptoid stabilization by N-gly is additive. When nitrobenzyl side chain is removed by exposure to UV light, the triple helix becomes less stable as the Nnbz residue is converted to glycine.

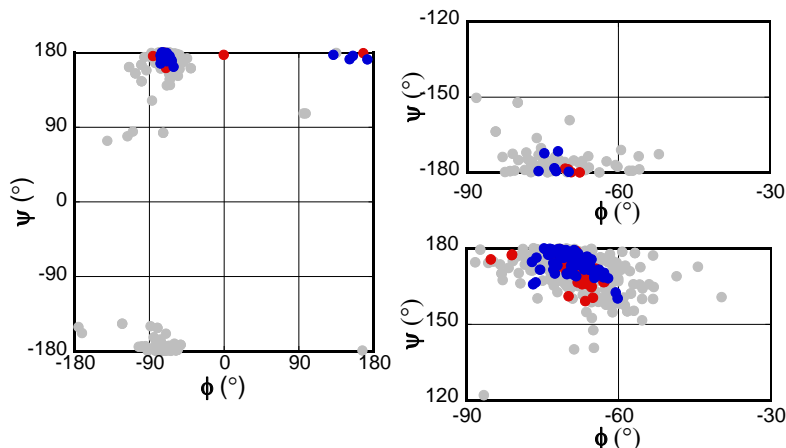
Name	Primary structure	T_m (°C)	T_m (°C) post UV
Nnbz-CMP		62	37
Nnbz2-CMP		73	18
Nnbz3-CMP		79	< 4



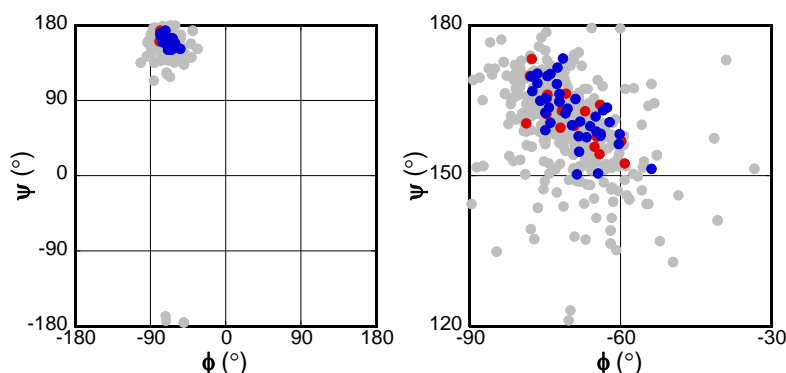
Supplementary Fig. S1 | Crystal structure of Nlys-CMP. **a**, Overall triple-helical structure, showing the one trimer in the asymmetric unit, including complete hydration shell, determined at 0.95 Å resolution. Central Nlys residues are colored in orange. **b**, Illustration of B-factors in the crystal structure: Overall B-factors are low through most of the triple-helical subunits with higher B-factors only appearing near the ends, indicating that the structure is less ordered near the ends of the helix.



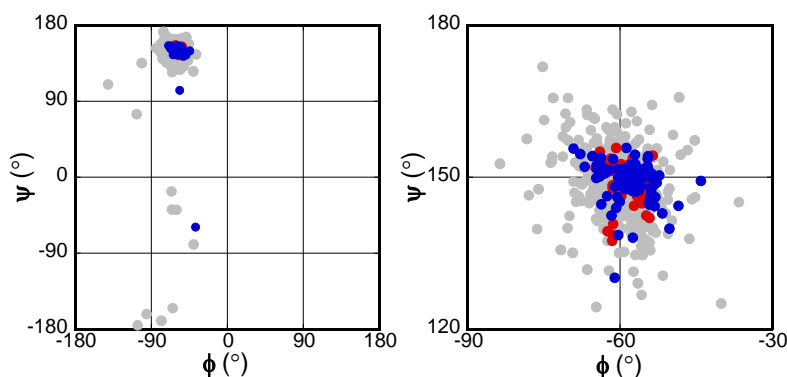
Supplementary Fig. S2 | Crystal structure of Nphe-CMP. **a**, Overall triple-helical structure, showing the two trimers in the asymmetric unit, including the complete hydration shell, determined at 1.10 Å resolution. Central Nphe residues are colored in green. **b**, Illustration of B-factors in the crystal structure: Overall B-factors are low through most of the triple helical subunits with higher B-factors only appearing near the ends, indicating that the structure is less ordered near the ends of the helix.



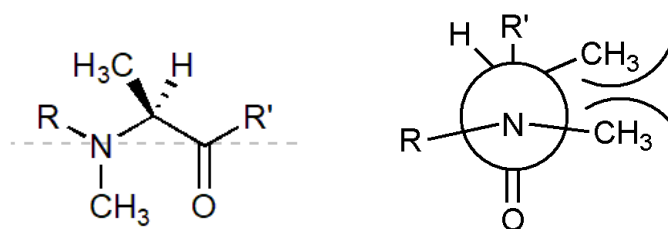
Supplementary Fig. S3a | Distribution of glycine ϕ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs glycine residues in collagen structures from literature (grey). Average dihedral angles (ϕ , ψ) for all non-terminal glycine residues are almost identical to previously reported structures: $(-69 \pm 3^\circ, 174 \pm 5^\circ)$. We removed the terminal glycines because they have very high B-factors. The collagen structures used for comparison are listed in Supplementary Table S4.



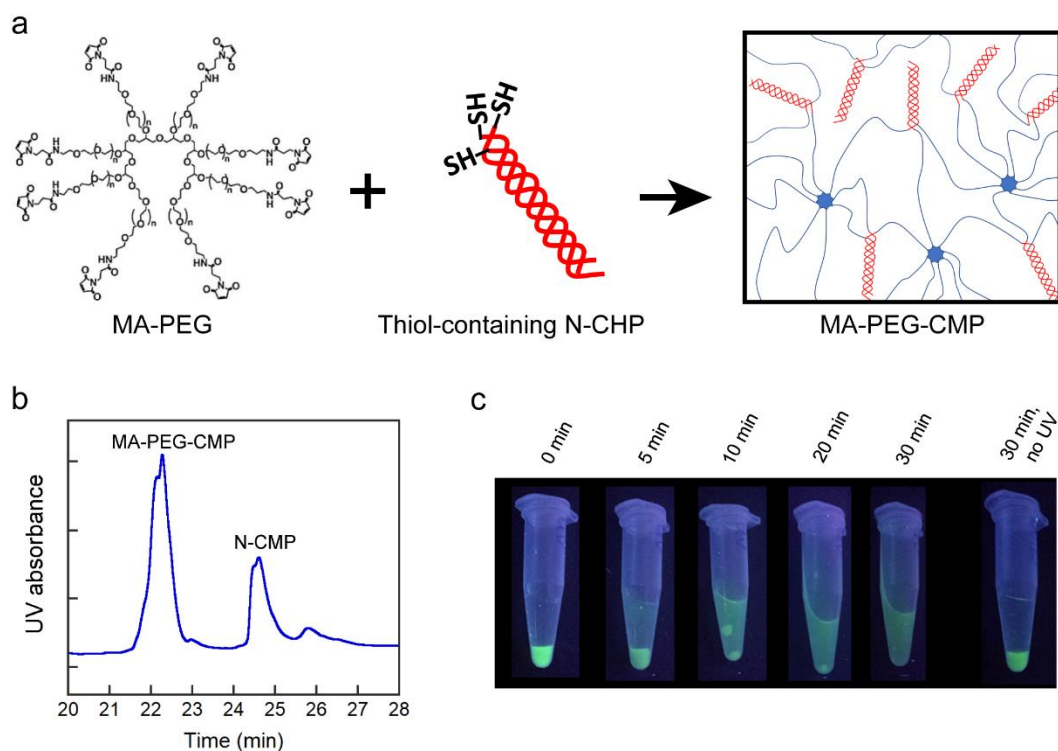
Supplementary Fig. S3b | Distribution of proline ϕ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs X position amino acids in collagen structures from literature (grey). Average dihedral angles for all proline residues are $(-70 \pm 6^\circ, 162 \pm 6^\circ)$. The collagen structures used for comparison are listed in Supplementary Table S4.



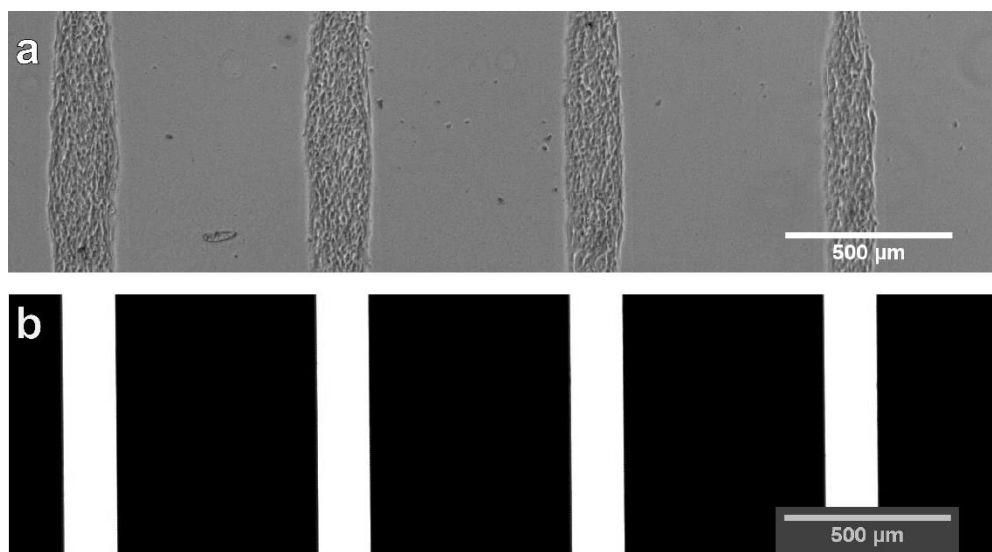
Supplementary Fig. S3c | Distribution of hydroxyproline ϕ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs Y position amino acids in collagen structures from literature (grey). Average dihedral angles for all proline residues are $(-58 \pm 5^\circ, 147 \pm 7^\circ)$. The collagen structures used for comparison are listed in Supplementary Table S4.



Supplementary Fig. S4 | Newman-like Projection of NMe-Ala showing the steric clash of adjacent methyl groups in the polyproline backbone conformation.



Supplementary Fig. S5 | MA-PEG-CMPs produce a stable hydrogel capable of UV-induced dissolution. a, Schematic of hydrogel formation by reacting an 8-arm PEG hydrogel with X-CMPs. Maleimide groups at the termini of the 8-arm PEG react with the thiol-containing X-CMPs to form a triple helix crosslinked hydrogel. **b,** HPLC of MA-PEG-CMP showing 70% of the X-CMP conjugated to the 8-arm PEG. **c,** MA-PEG-CMP mixed with F-Nnbz2-CMP forms a green hydrogel that traps the green colored F-Nnbz2-CMP via triple helical crosslinks. Upon UV exposure, the triple helical crosslinks unfold, releasing the bound F-Nnbz2^X-CMP into solution.

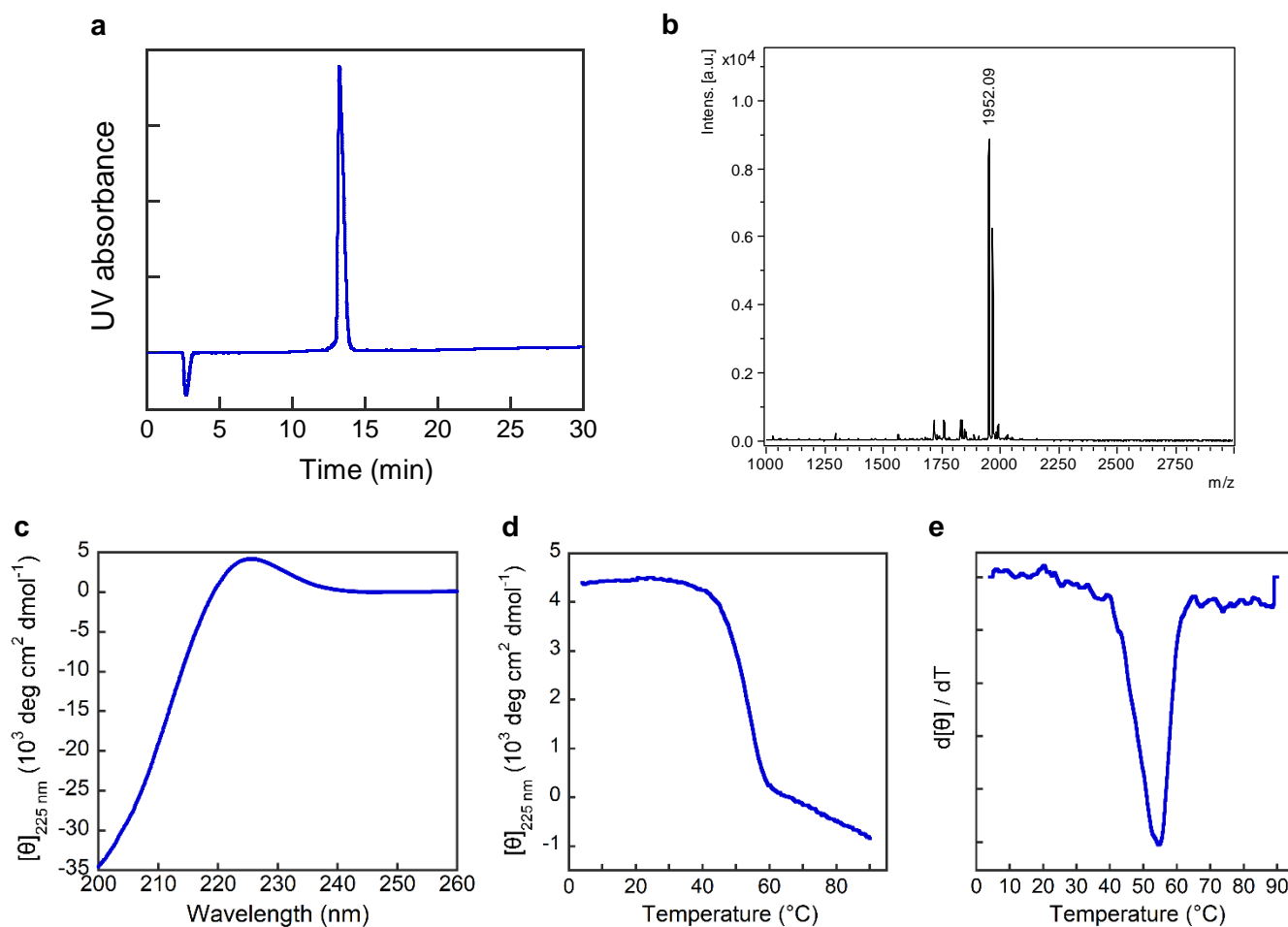
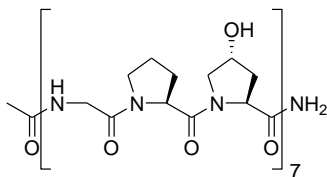


Supplementary Fig. S6 | Light micrographs showing spatial control of cell attachment by MA-PEG-CMP. MDA-MB-231 cells seeded onto a gelatin substrate coated with the MA-PEG-CMP hydrogel (**a**). The coated hydrogel had previously been exposed to UV light through a photomask (**b**) and cells bound only to areas which had been exposed. The results indicate that the MA-PEG-CMP binds to gelatin preventing cell attachment to the gelatin substrate and can be photo-removed to recover cell adherence.

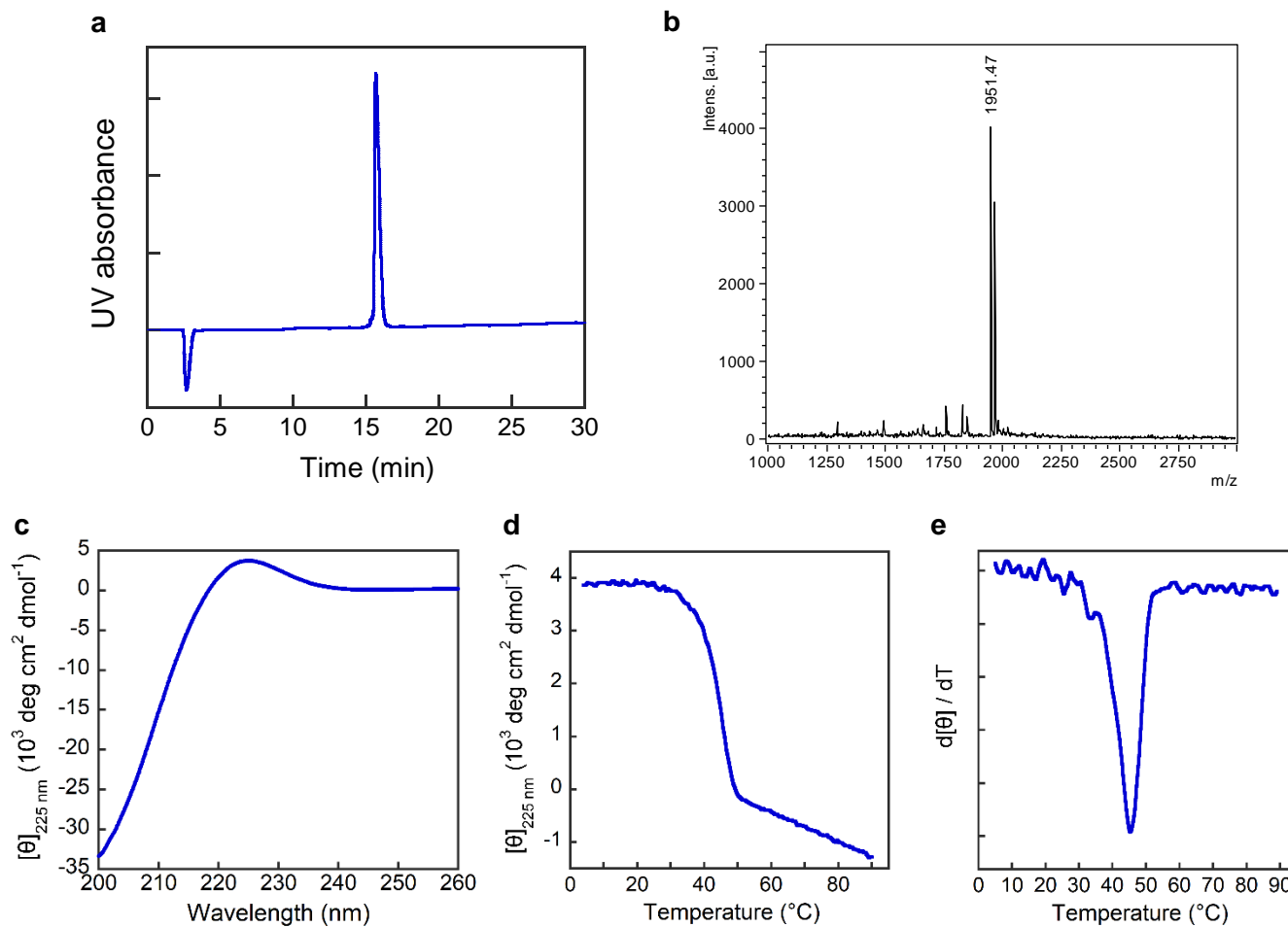
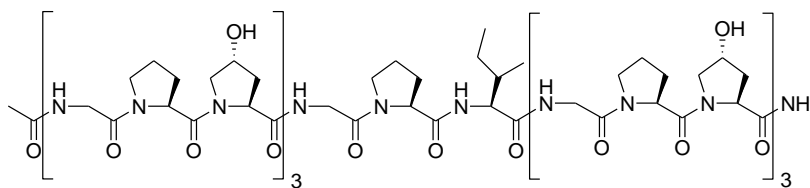
Supplementary Section 3: HPLC, MALDI and CD of all peptides

X-CMP peptides

Pro-CMP



Pro-Ile-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 15.7$ min.

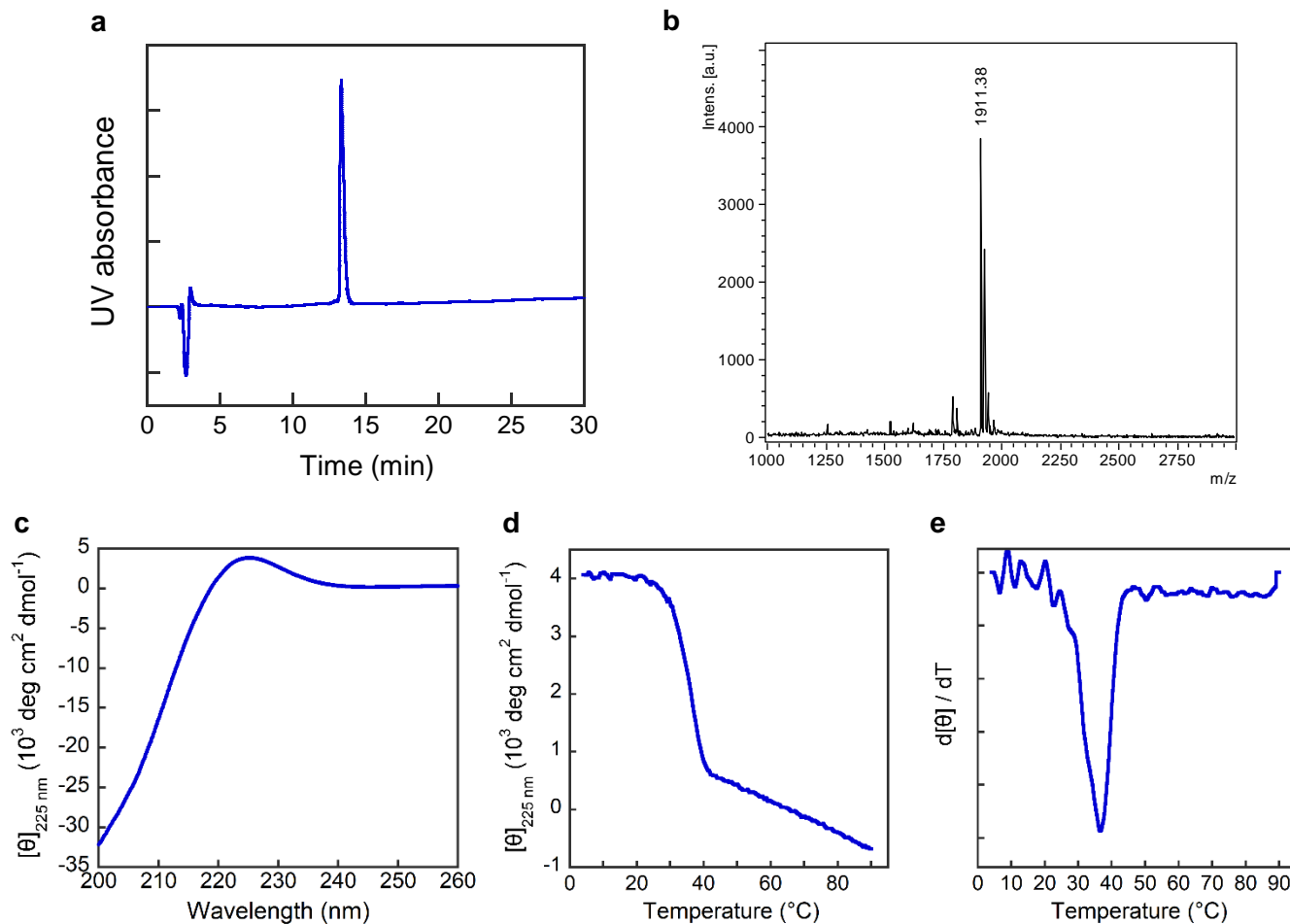
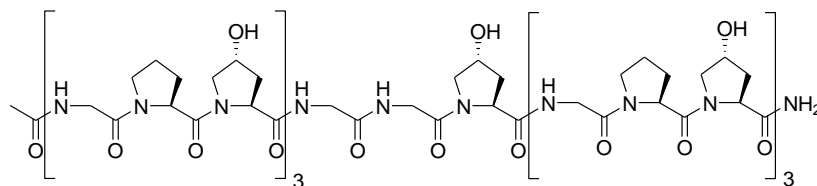
b, MALDI-MS, calculated: 1951.9 $[M+\text{Na}]^+$, observed: 1951.5 $[M+\text{Na}]^+$, 1967.5 $[M+\text{K}]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 45.5$ $^{\circ}\text{C}$.

Gly-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 13.3$ min.

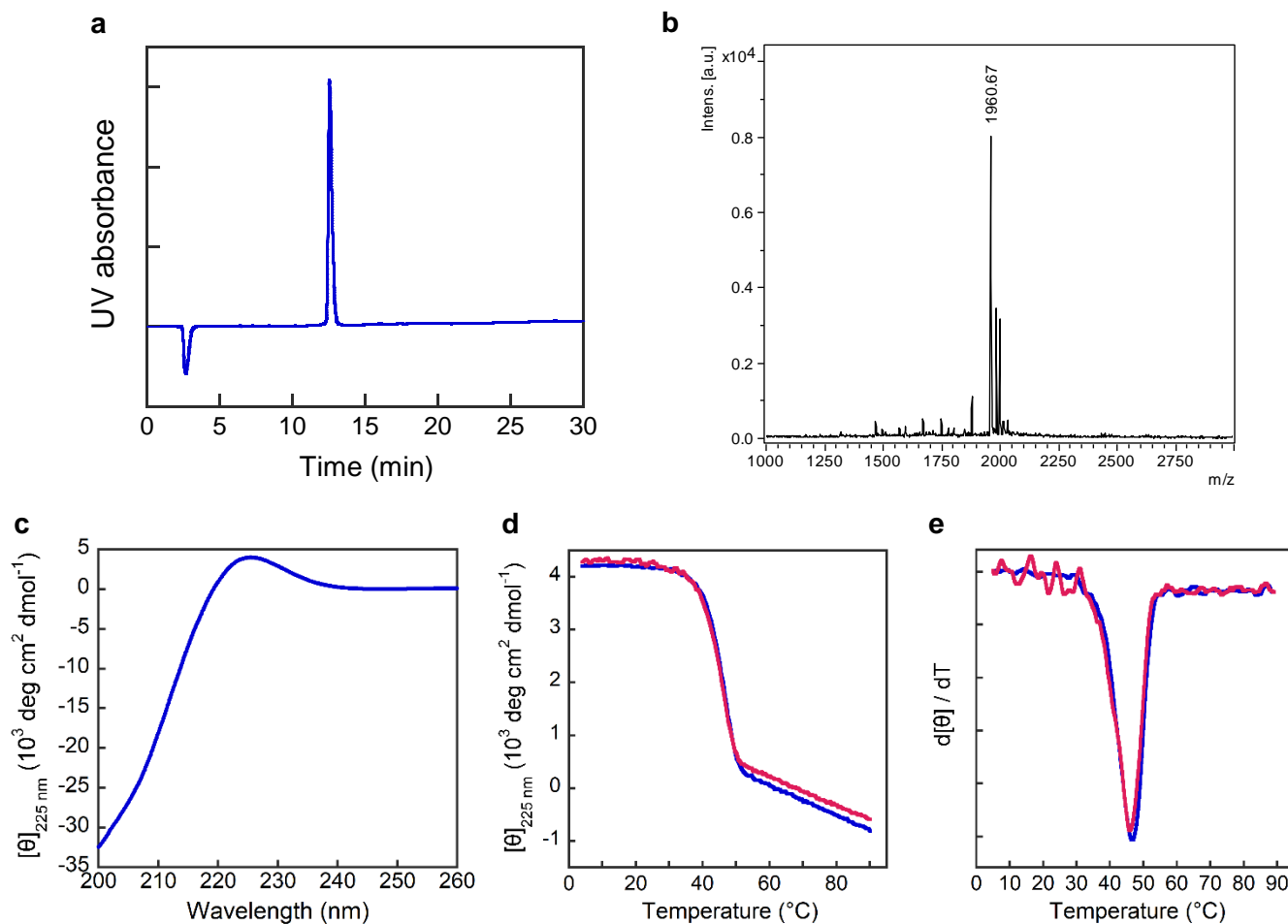
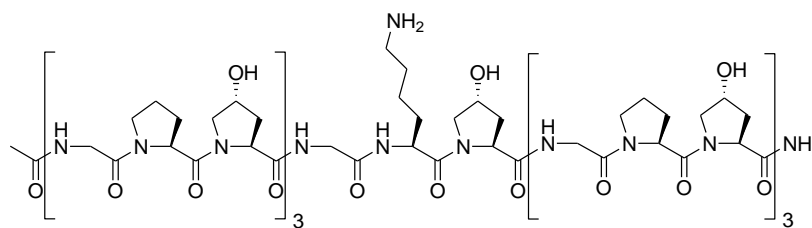
b, MALDI-MS, calculated: 1911.9 $[M+Na]^+$, observed: 1911.4 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 37$ °C.

Lys-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 12.5$ min.

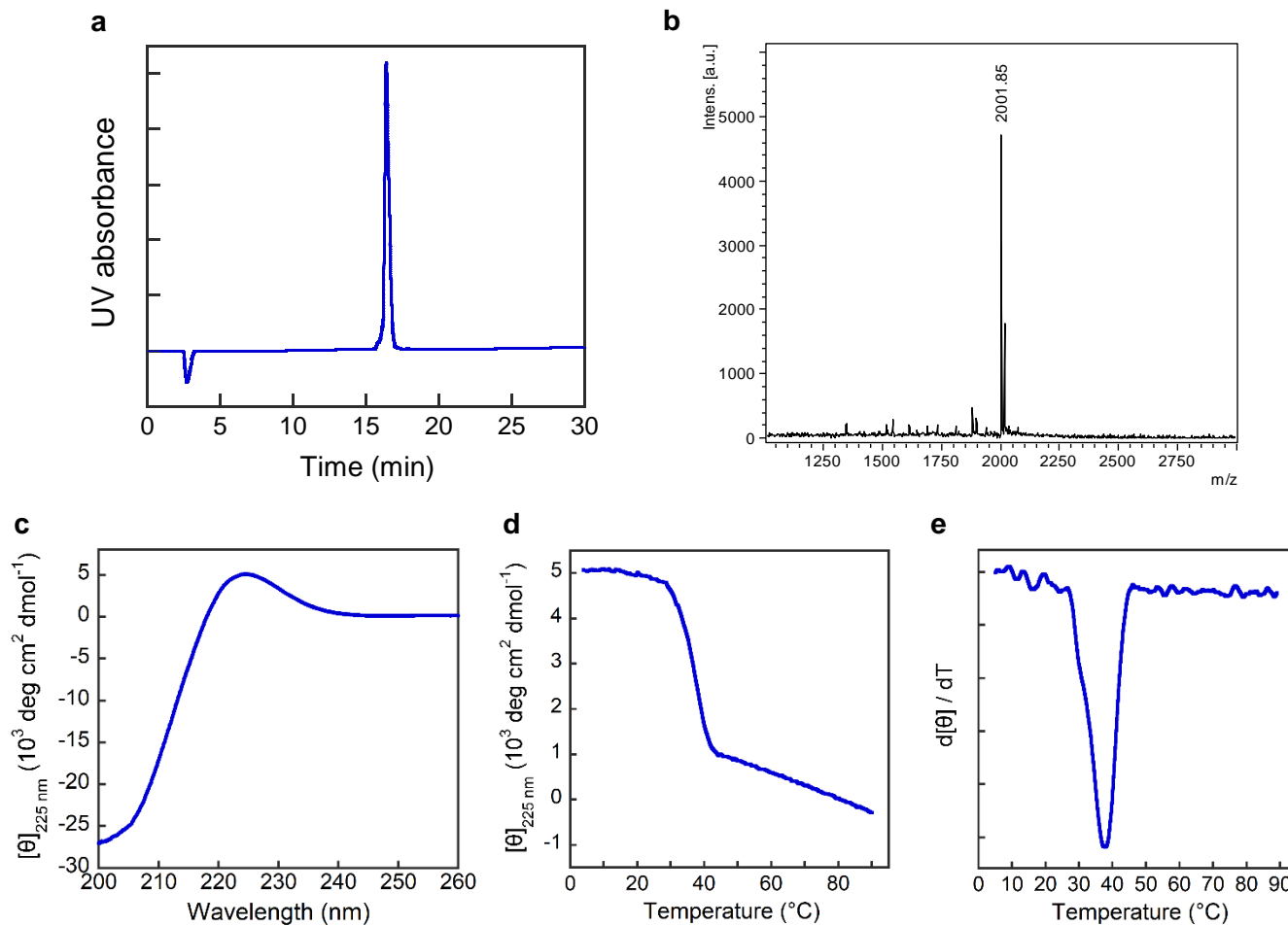
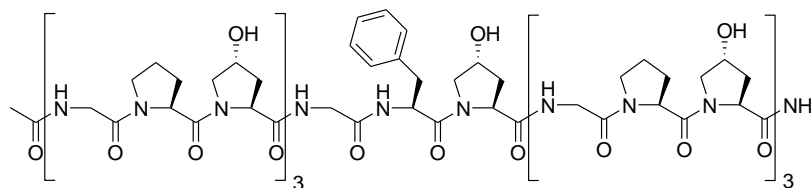
b, MALDI-MS, calculated: 1960.9 $[M+H]^+$, observed: 1960.7 $[M+H]^+$, 1982.7 $[M+Na]^+$, 1998.7 $[M+K]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curves in PBS buffer (blue) and 32 mM NaOH solution (pH 12.51, red).

e, The first derivatives of the melting curves, $T_m = 47$ °C (PBS, blue) or 46 °C (NaOH, red).

Phe-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 16.4$ min.

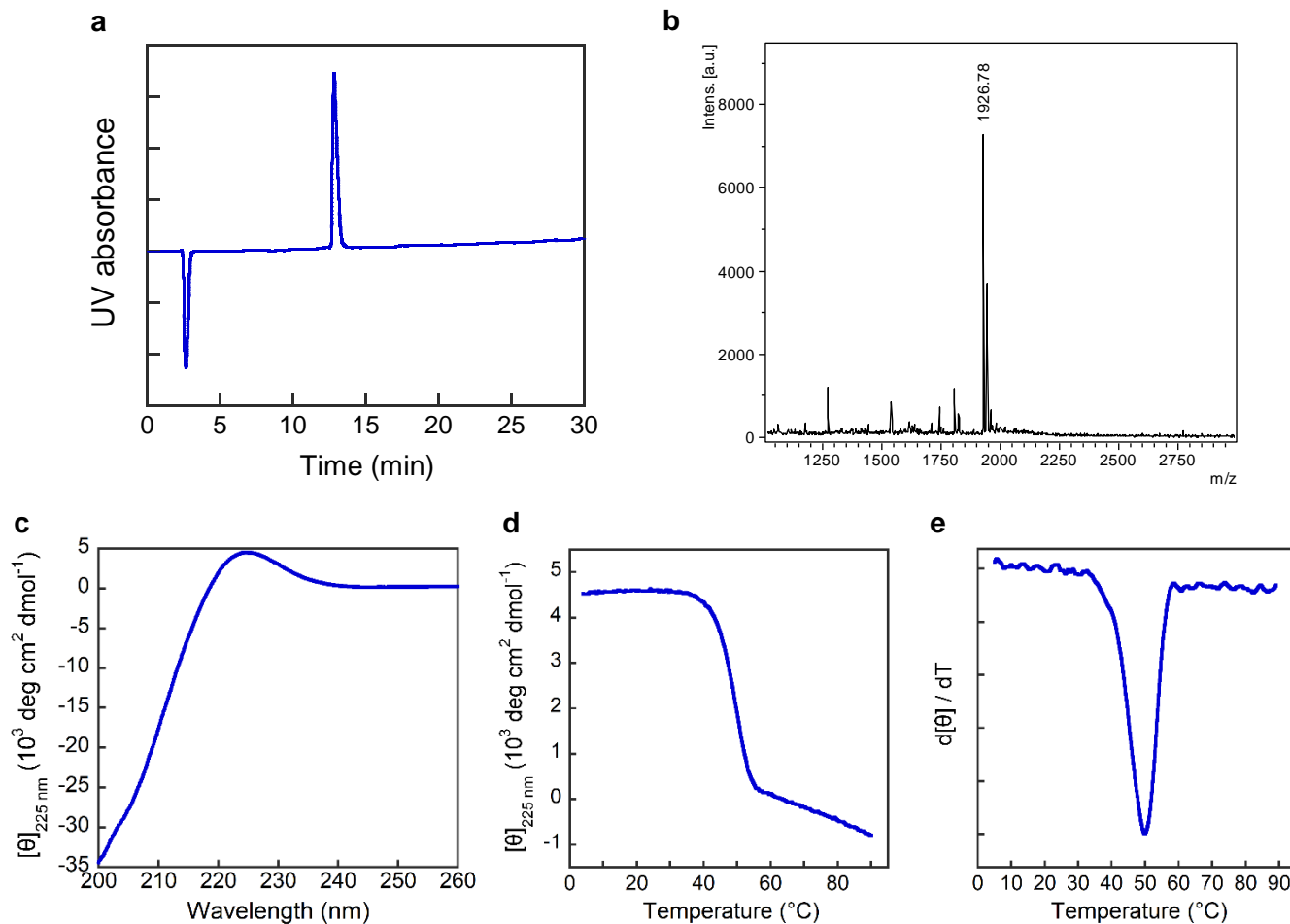
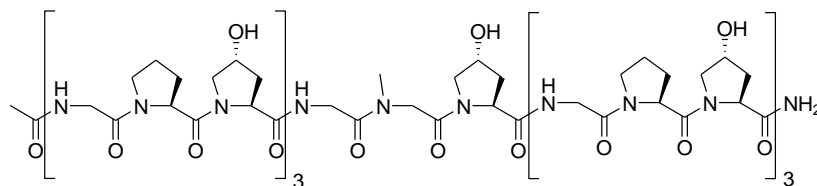
b, MALDI-MS, calculated: 2001.9 $[M+Na]^+$, observed: 2001.9 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 38$ $^{\circ}\text{C}$.

Sar-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 12.8$ min.

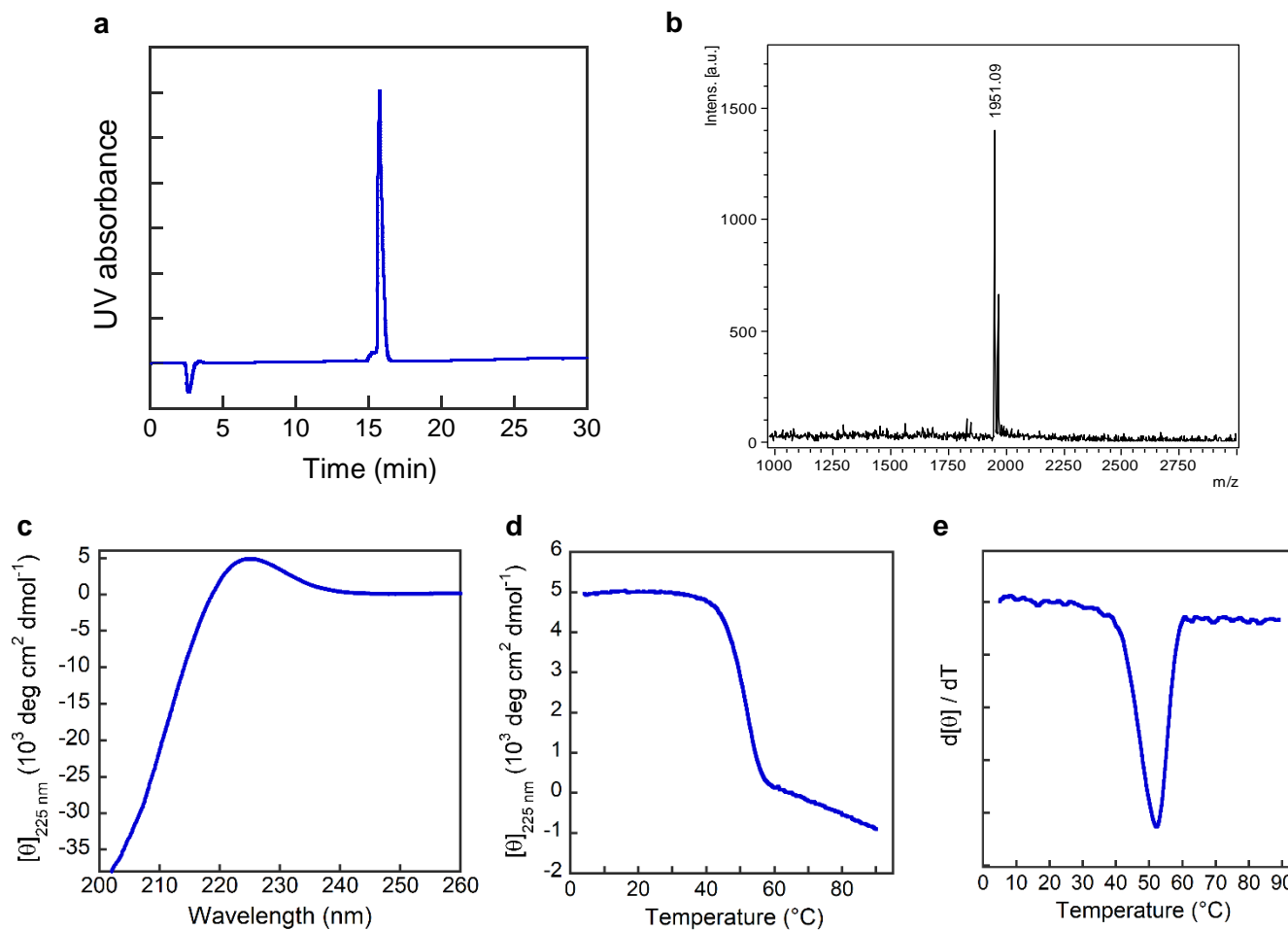
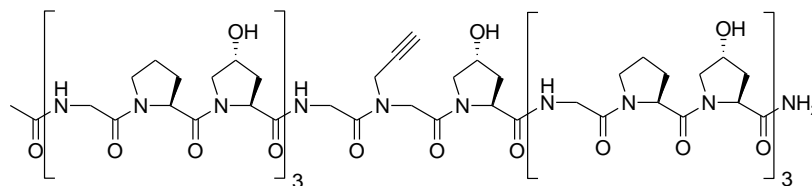
b, MALDI-MS, calculated: 1925.9 $[\text{M}+\text{Na}]^+$, observed: 1926.8 $[\text{M}+\text{Na}]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 50$ $^{\circ}\text{C}$.

Nakn-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 15.7$ min.

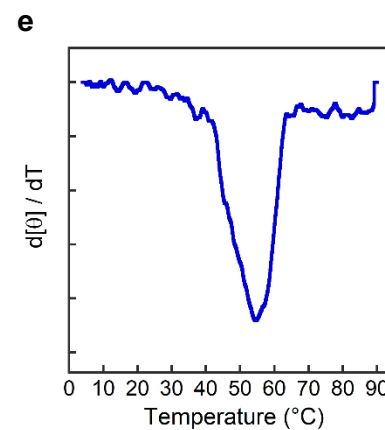
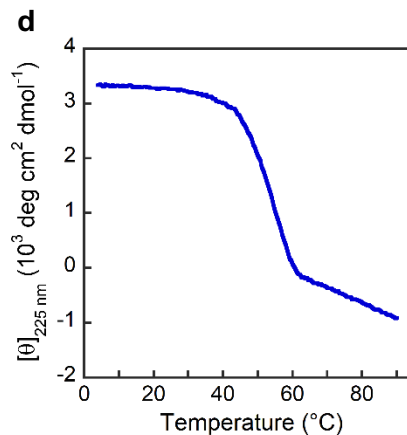
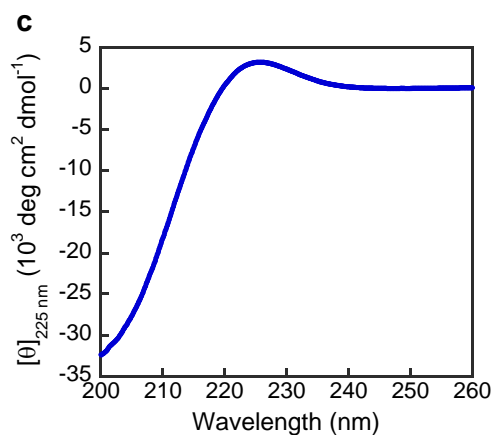
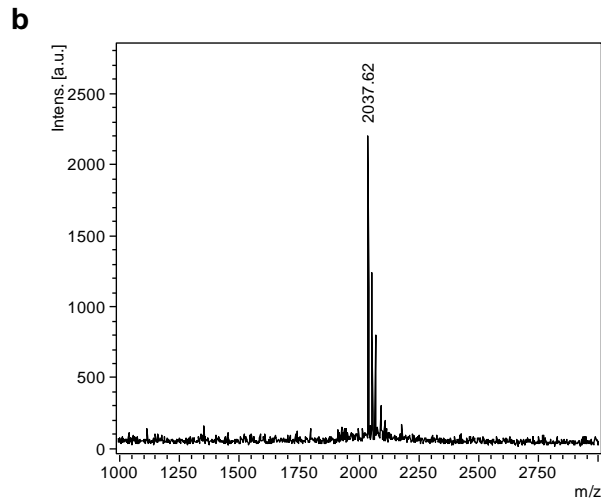
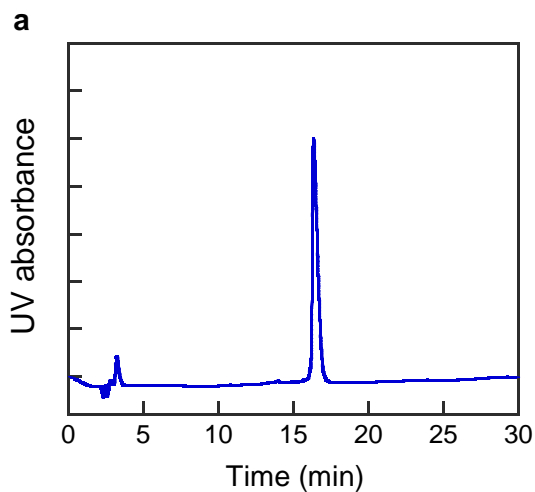
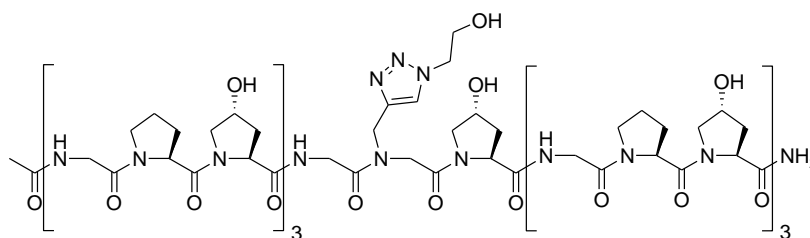
b, MALDI-MS, calculated: 1949.9 $[M+Na]^+$, observed: 1951.1 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 52$ °C.

Nakn-OH-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 16.3$ min.

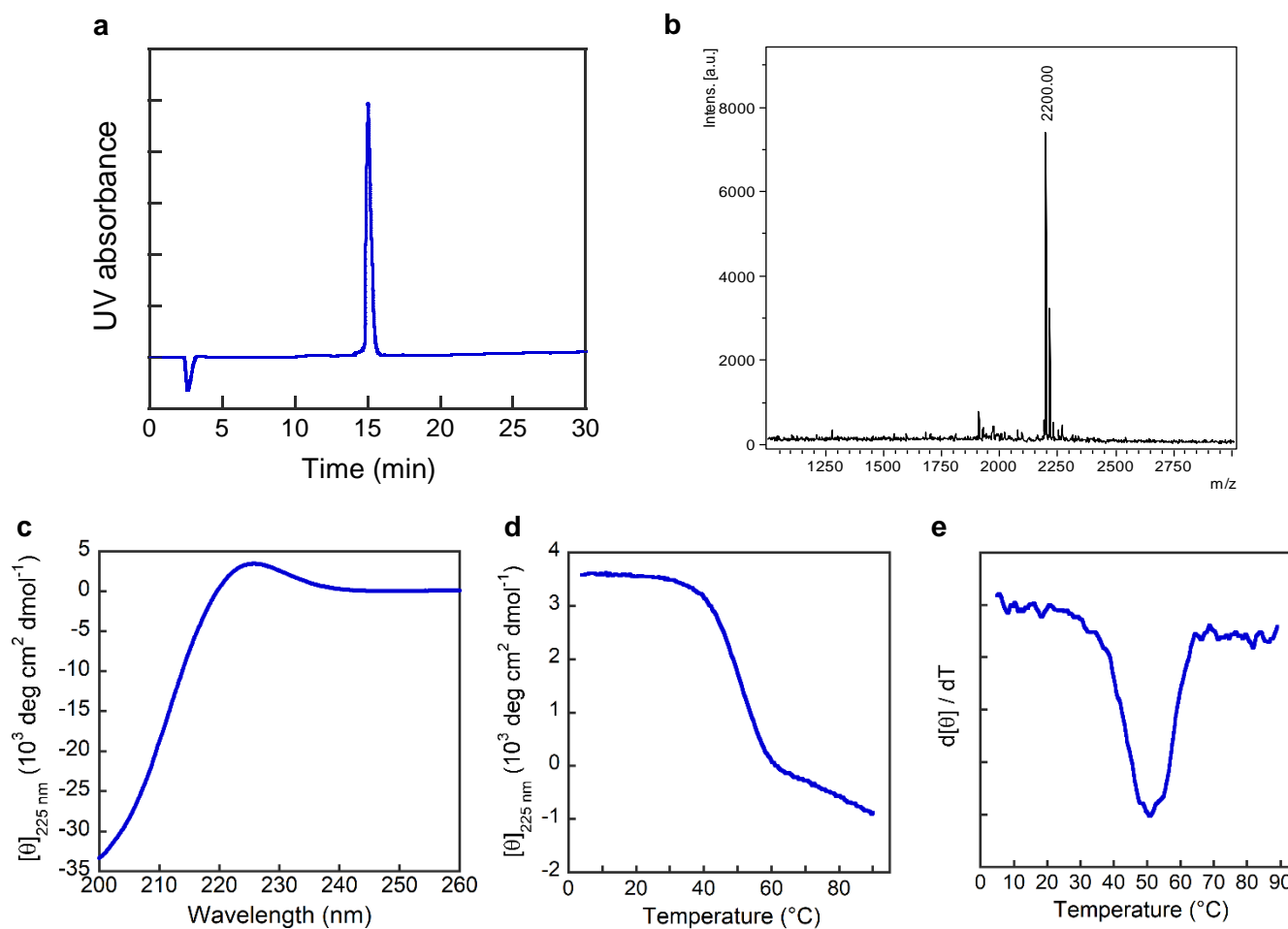
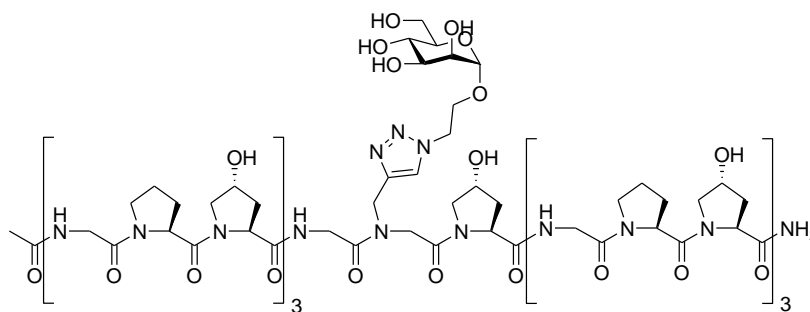
b, MALDI-MS, calculated: 2036.9 $[M+Na]^+$, observed: 2037.6 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

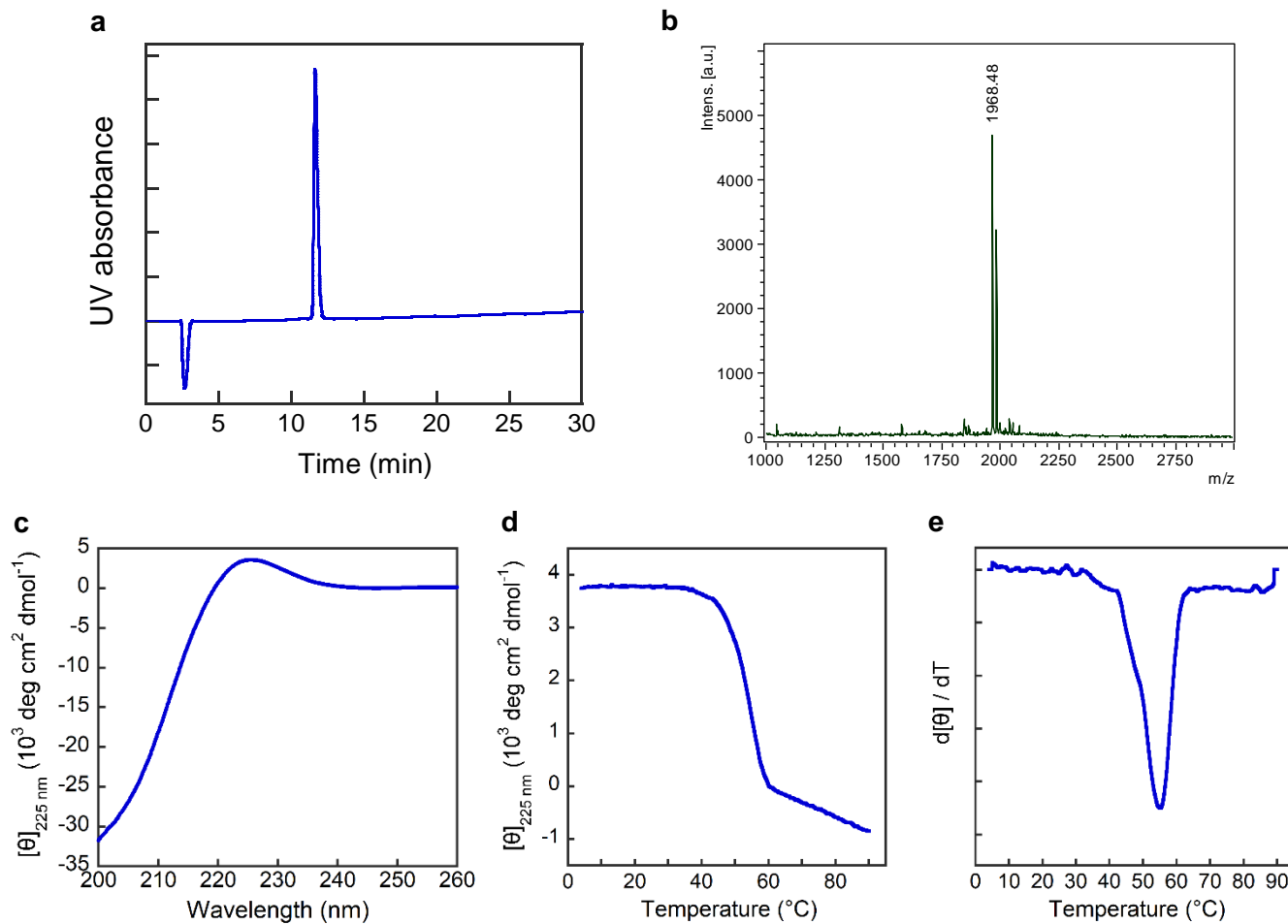
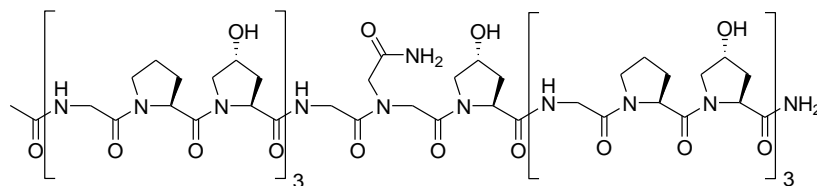
d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 55$ °C.

Nakn-maX-CMP



Nasn-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 11.6$ min.

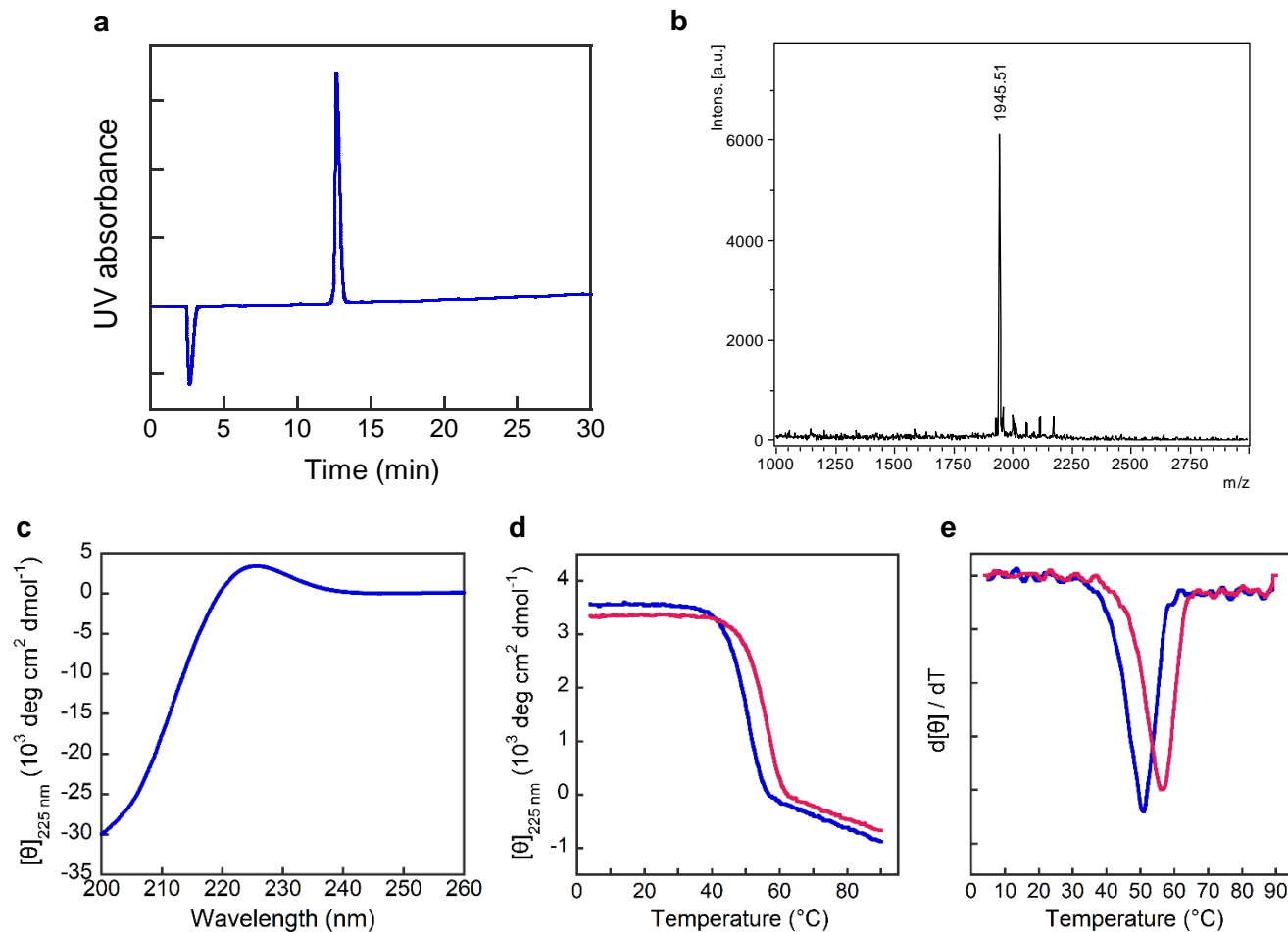
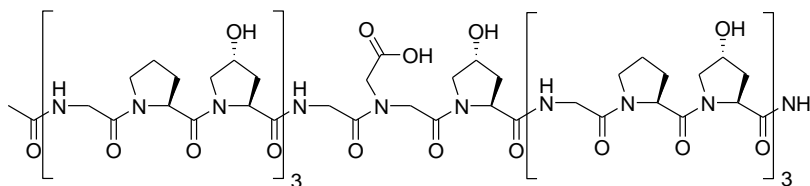
b, MALDI-MS, calculated: 1968.9 $[M+Na]^+$, observed: 1968.5 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 55$ °C.

Nasp-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 12.7$ min.

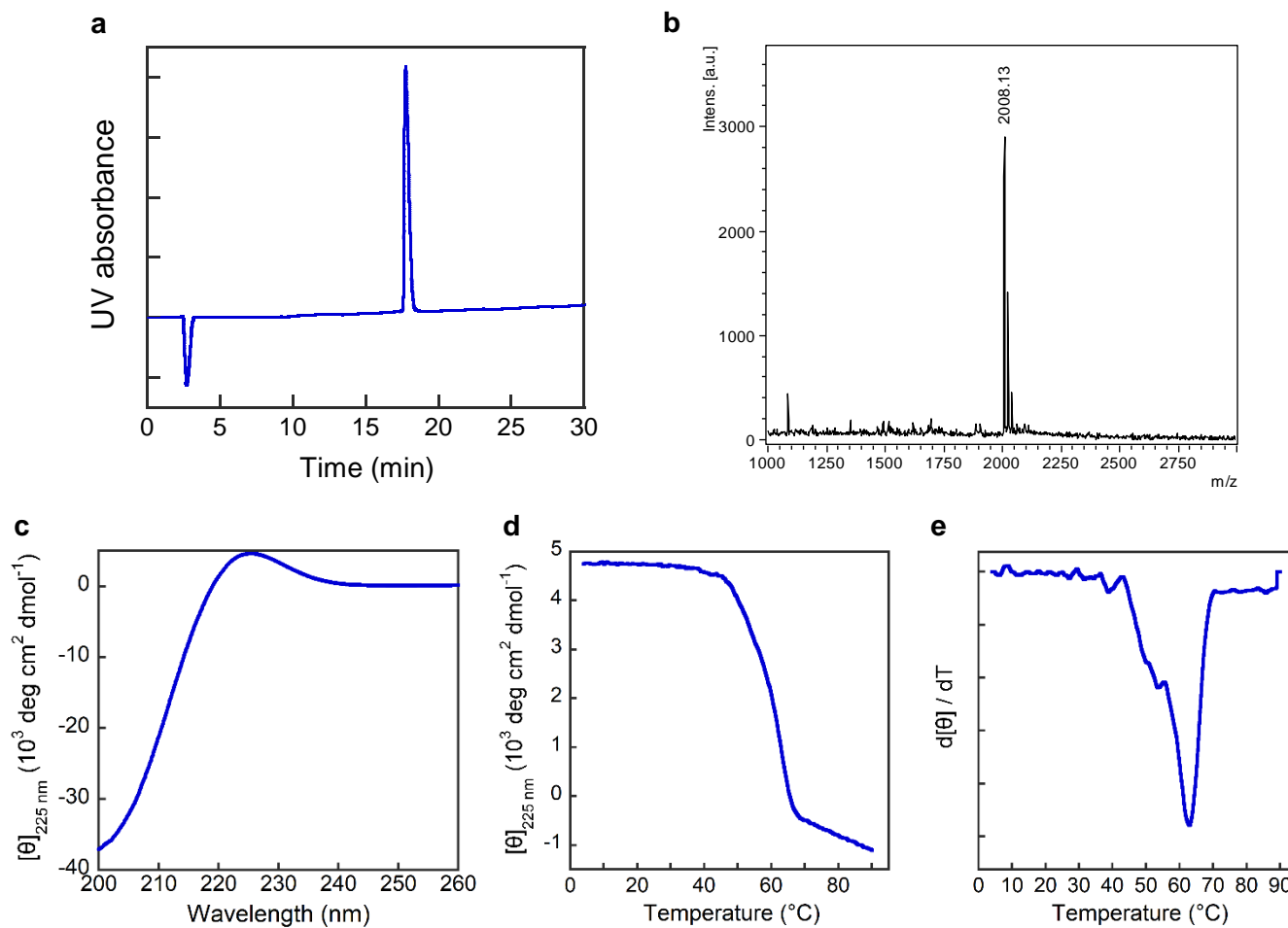
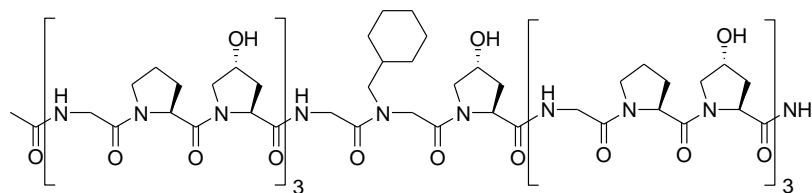
b, MALDI-MS, calculated: 1945.9 [M-H]⁺, observed: 1945.5 [M-H]⁺.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curves in PBS buffer (blue) and 15 mM HCl solution (pH 1.82, red).

e, The first derivatives of the melting curves, $T_m = 50$ °C (PBS, blue) or 56 °C (HCl, red).

Nchx-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 17.7$ min.

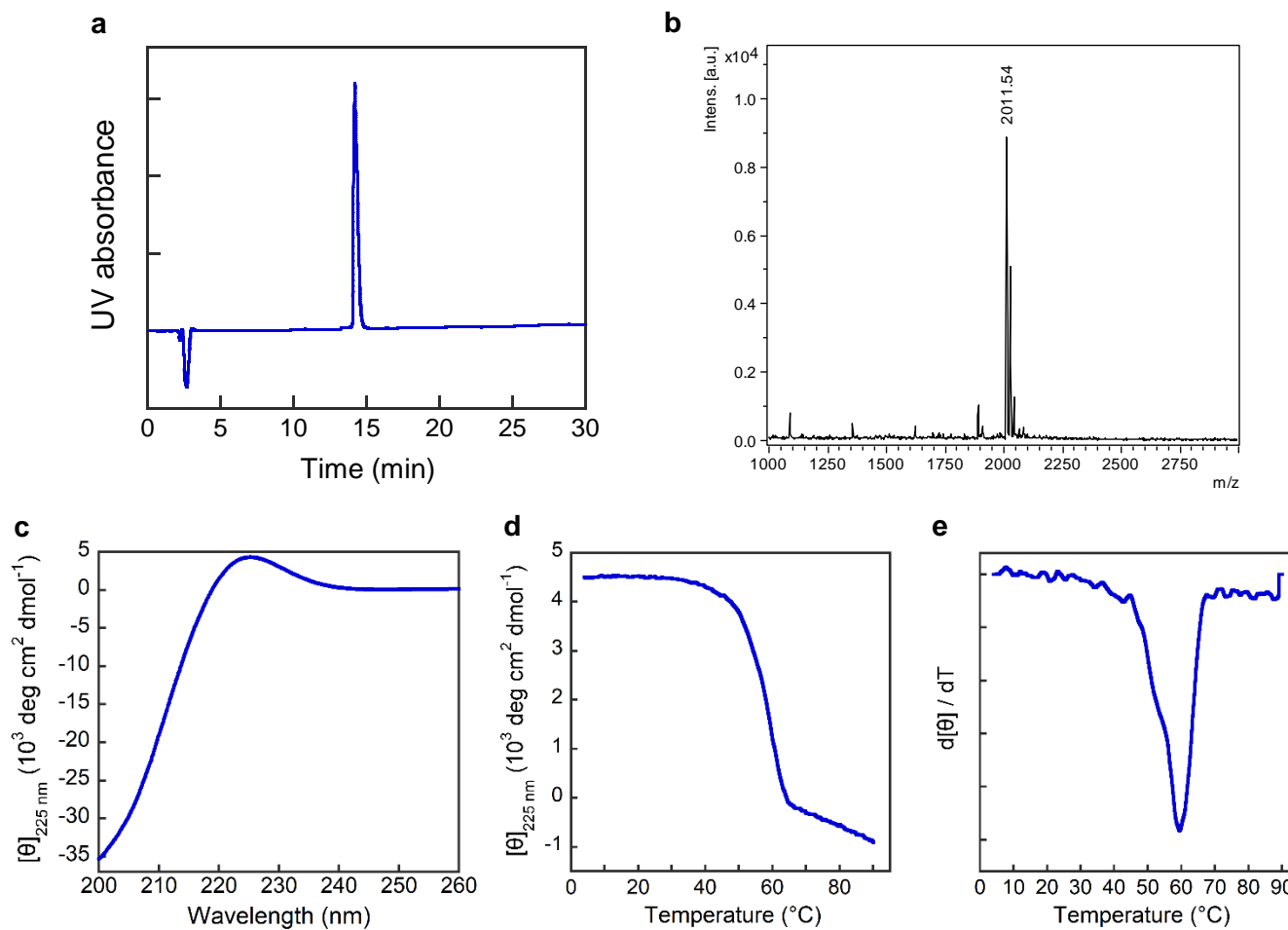
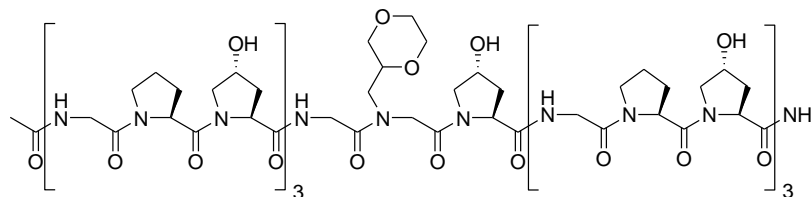
b, MALDI-MS, calculated: 2008.0 $[\text{M}+\text{Na}]^+$, observed: 2008.1 $[\text{M}+\text{Na}]^+$, 2024.1 $[\text{M}+\text{K}]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

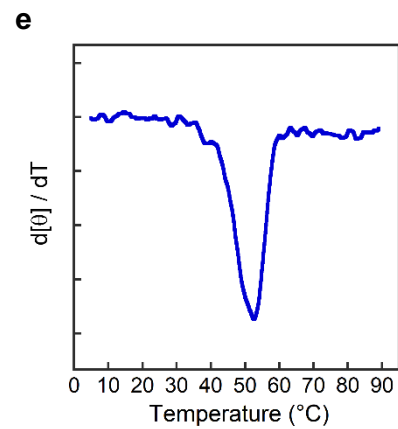
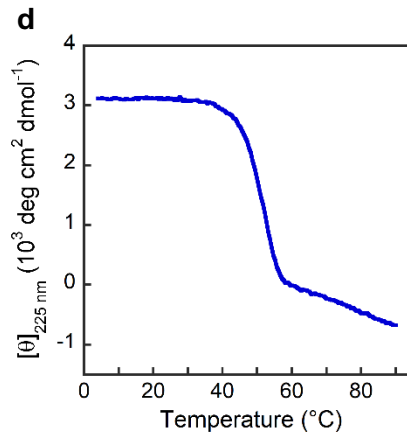
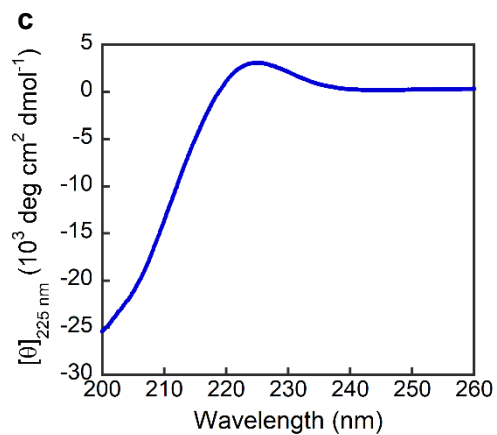
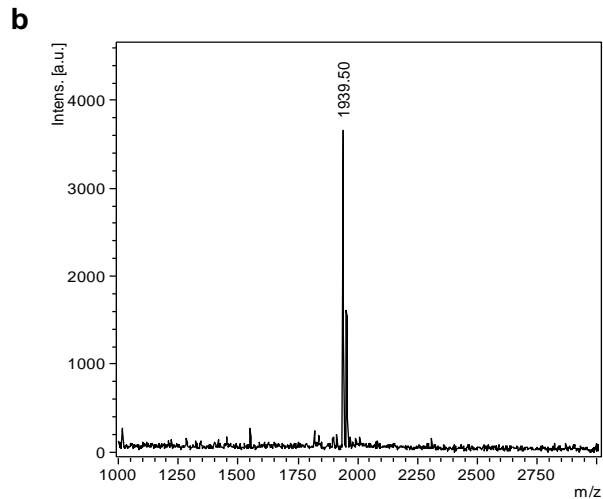
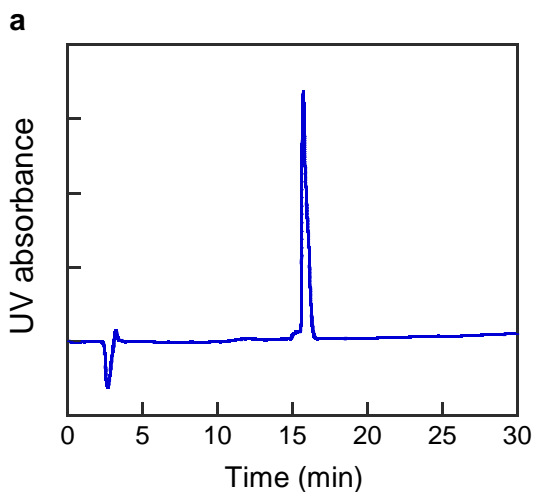
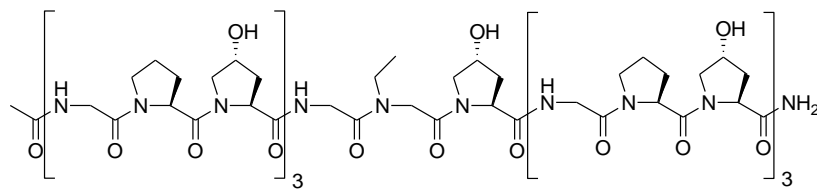
d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 63$ $^{\circ}\text{C}$.

Ndxn-CMP



NEt-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 15.7$ min.

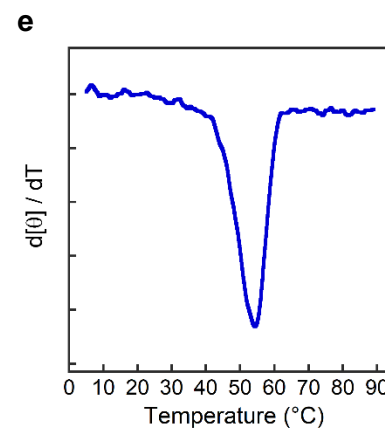
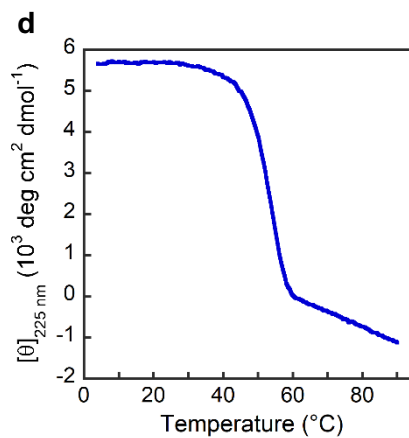
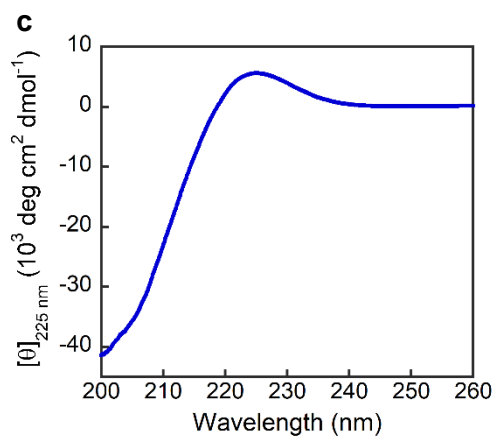
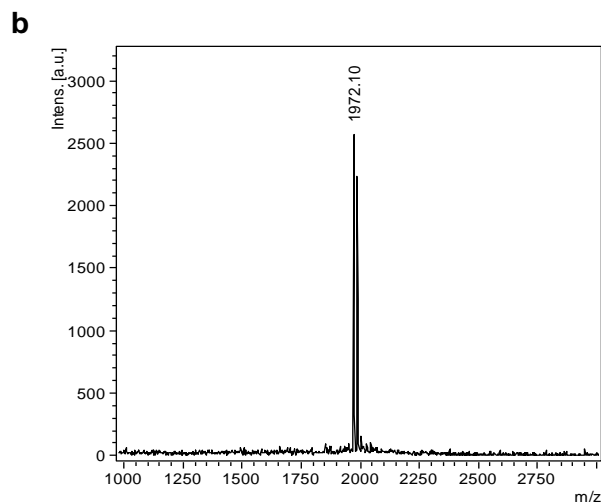
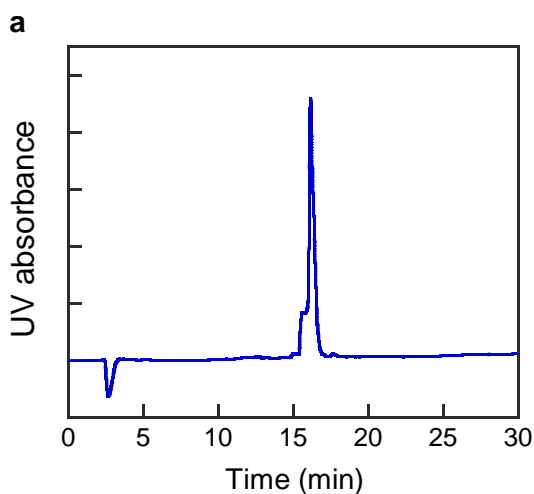
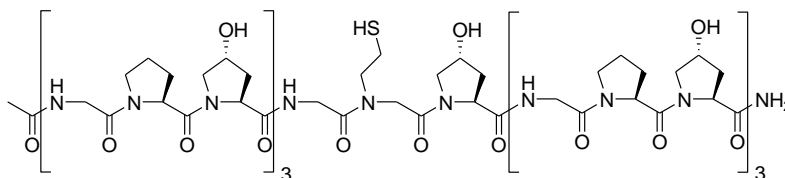
b, MALDI-MS, calculated: 1939.9 $[M+Na]^+$, observed: 1939.5 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 52$ °C.

Nhcy-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 16.2$ min.

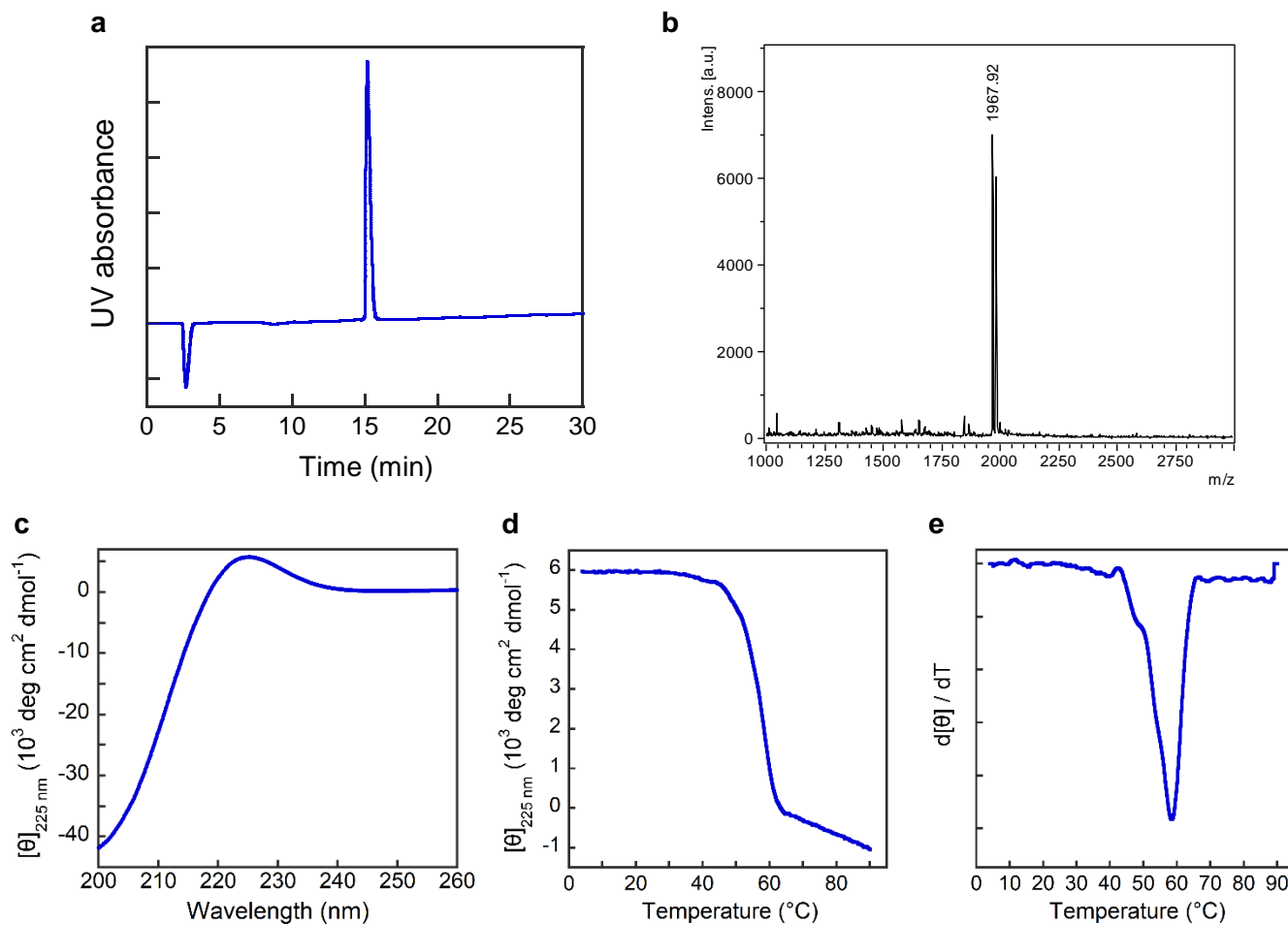
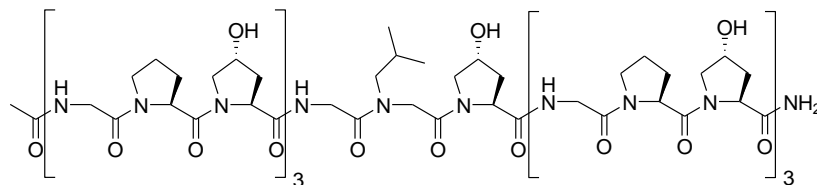
b, MALDI-MS, calculated: 1971.9 $[M+Na]^+$, observed: 1972.1 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

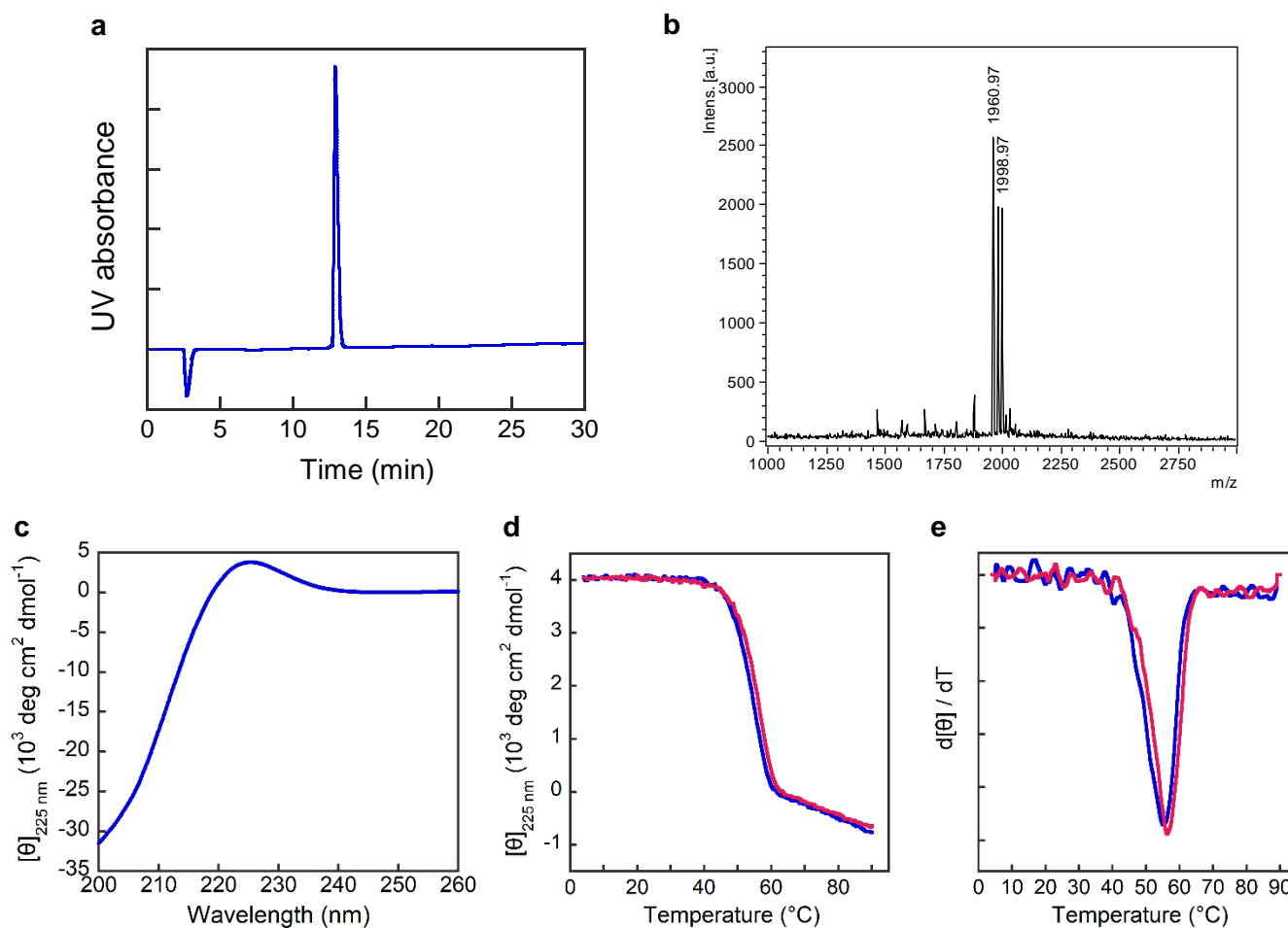
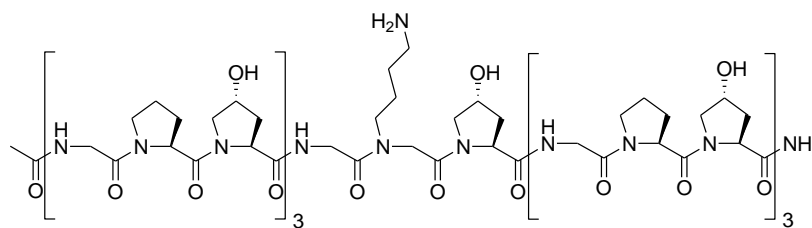
d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 54$ °C.

Nleu-CMP



Nlys-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 12.9$ min.

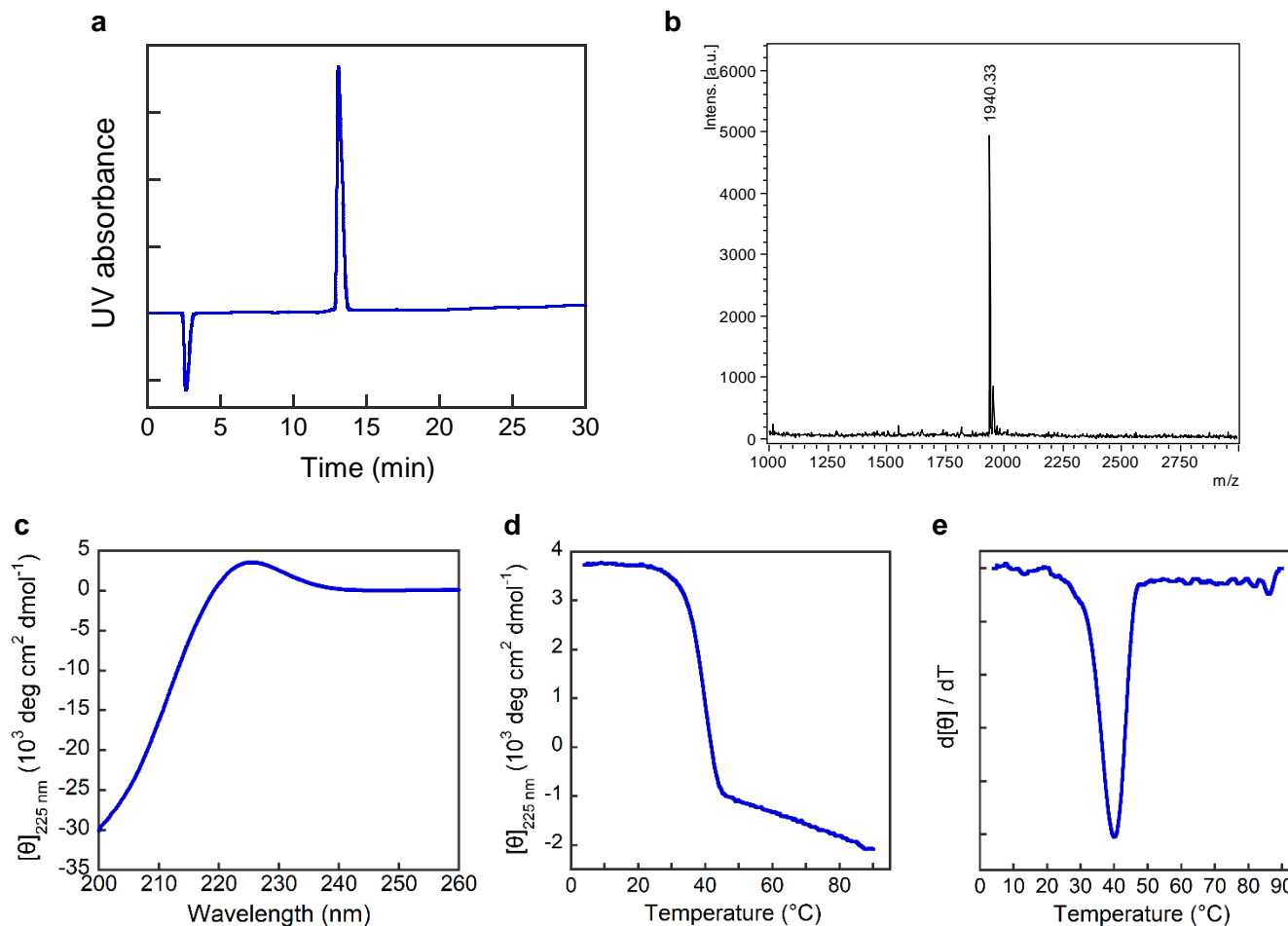
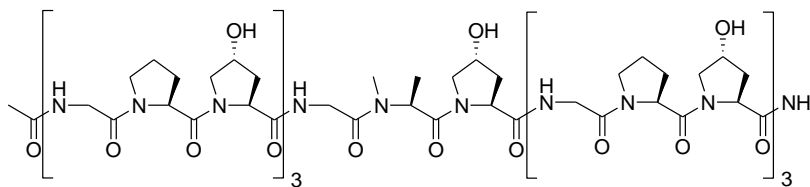
b, MALDI-MS, calculated: 1960.9 $[M+H]^+$, observed: 1961.0 $[M+H]^+$, 1983.0 $[M+Na]^+$, 1999.0 $[M+K]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curves in PBS buffer (blue) and 35 mM NaOH solution (pH 12.54, red).

e, The first derivatives of the melting curves, $T_m = 55$ °C (PBS, blue) or 56 °C (NaOH, red).

NMe-Ala-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 13.1$ min.

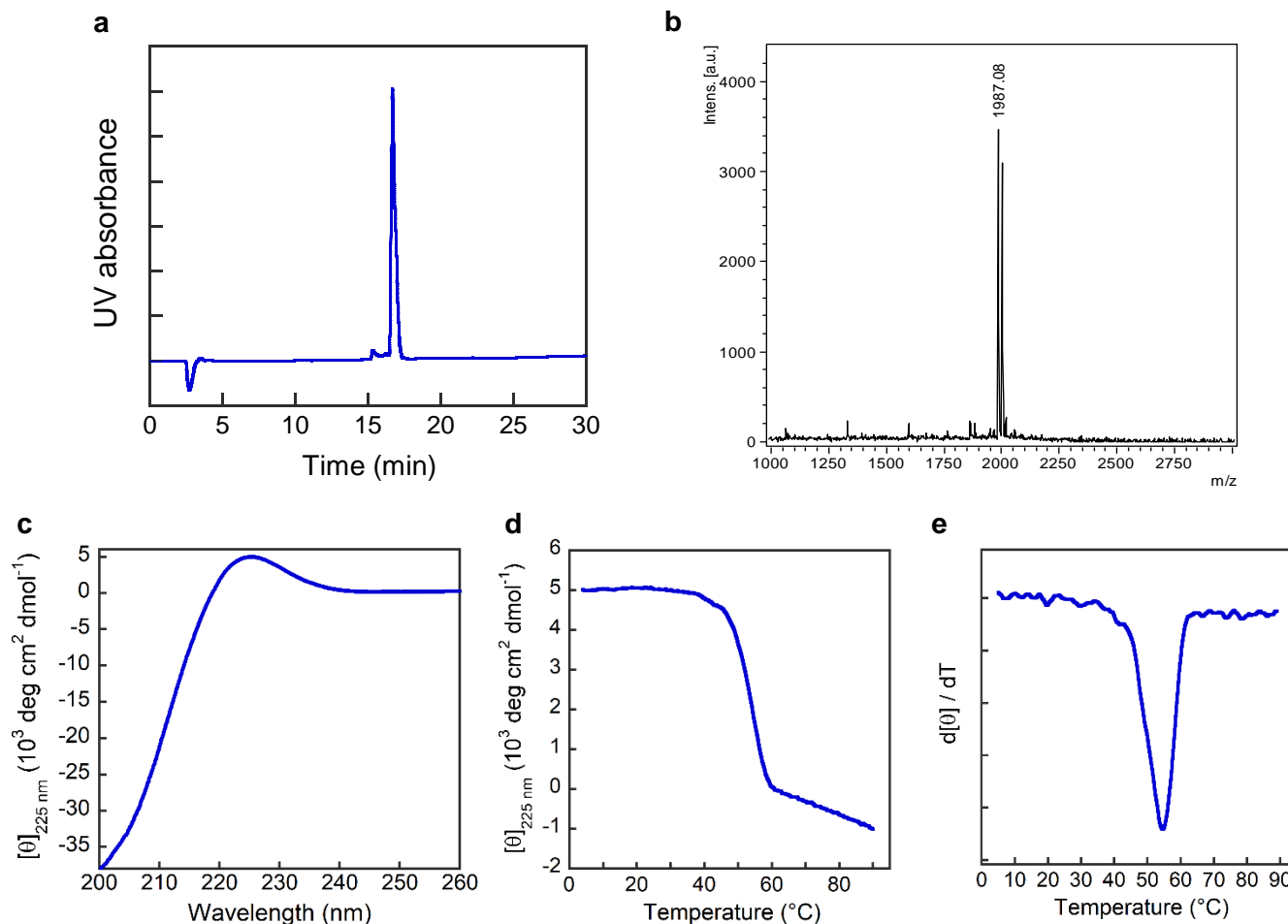
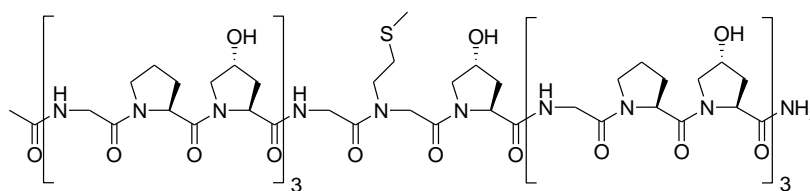
b, MALDI-MS, calculated: 1939.9 $[M+Na]^+$, observed: 1940.3 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve of in PBS buffer.

e, The first derivative of the melting curve, $T_m = 40$ $^{\circ}\text{C}$.

Nmet-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 16.7$ min.

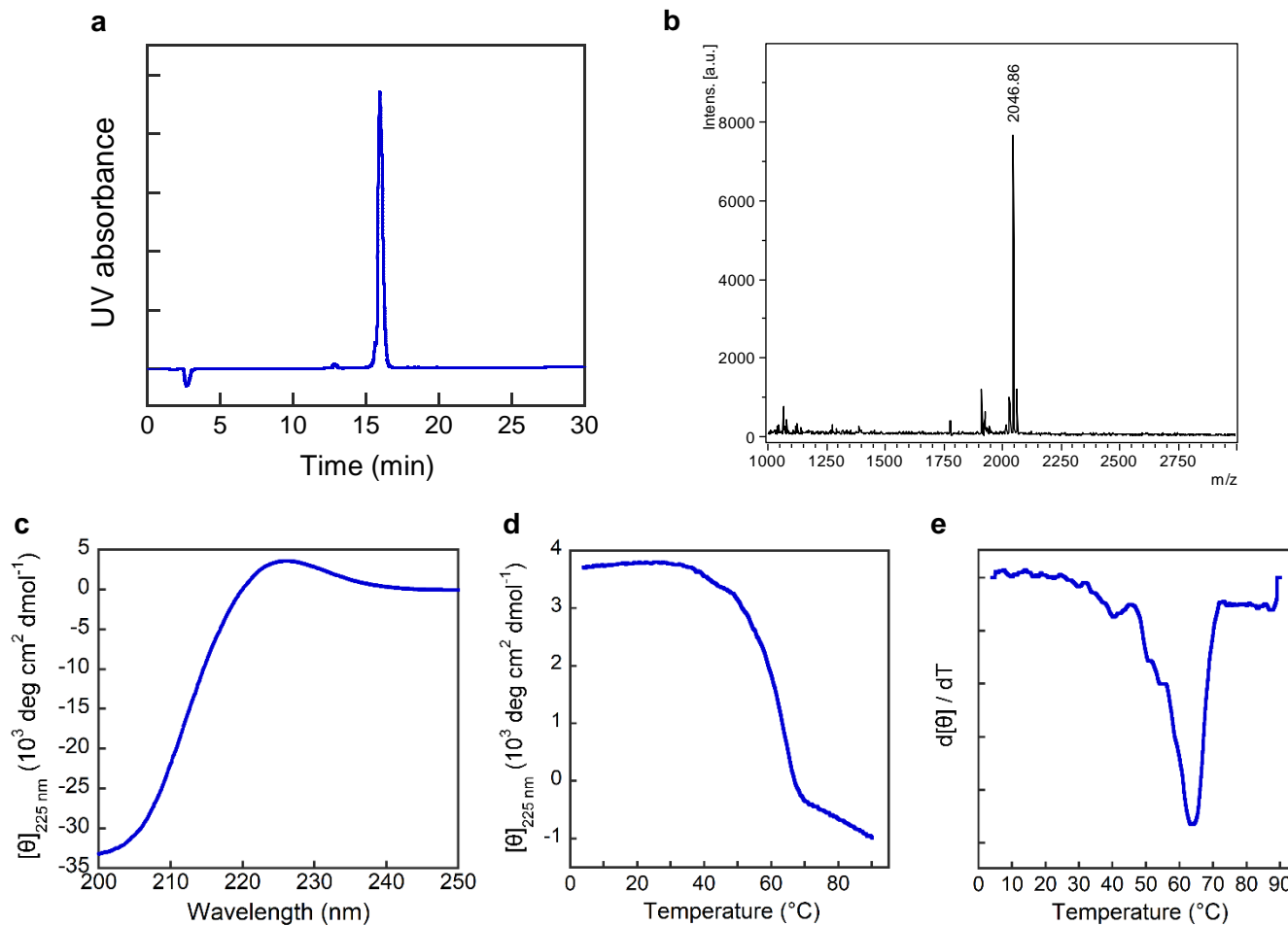
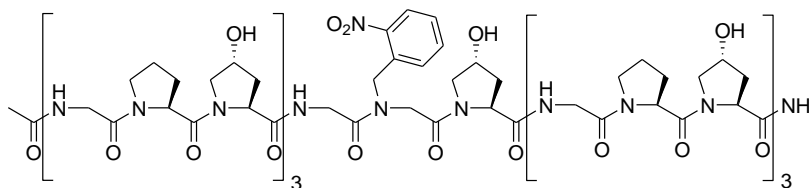
b, MALDI-MS, calculated: 1986.9 $[M+Na]^+$, observed: 1987.1 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 55$ °C.

Nnbz-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 15.9$ min.

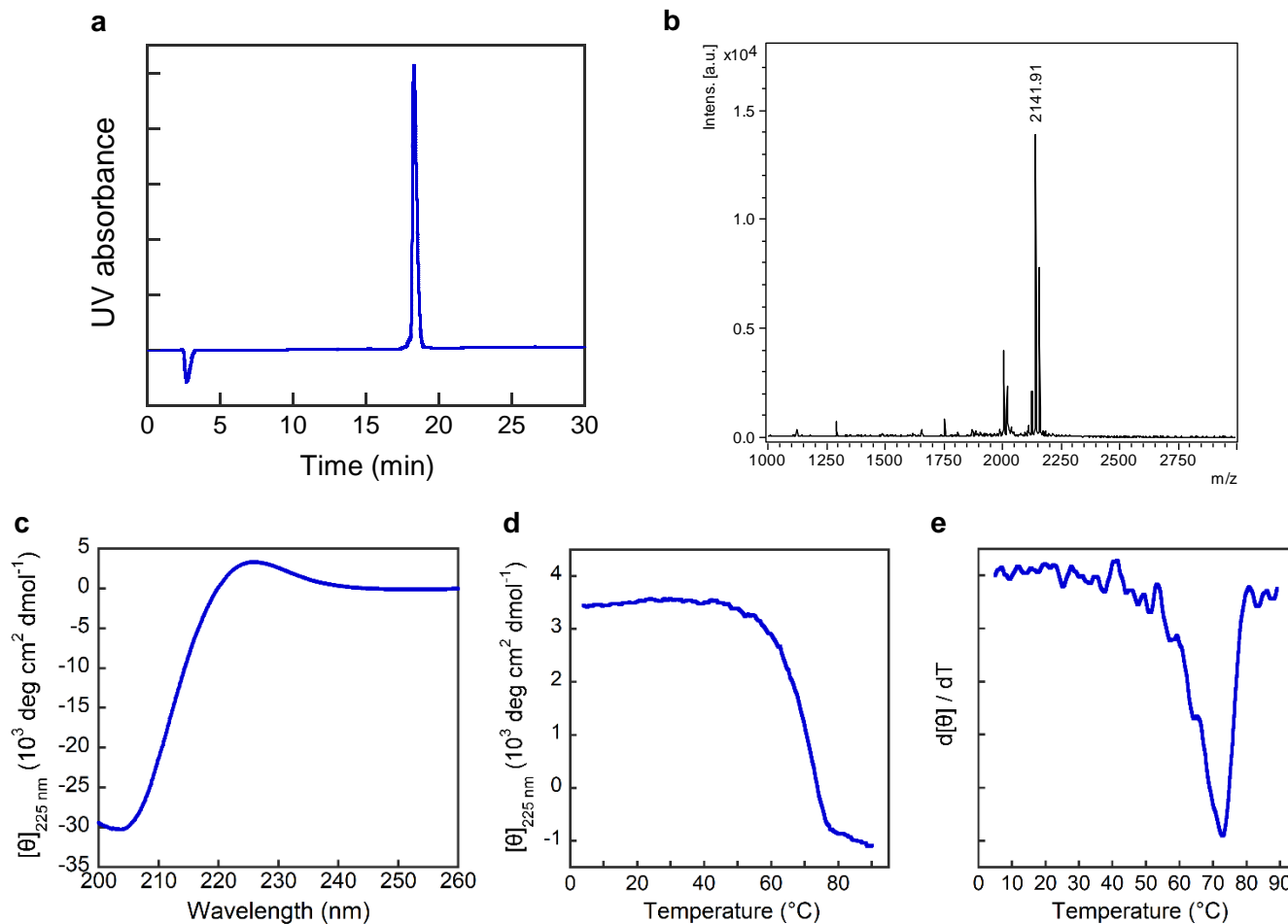
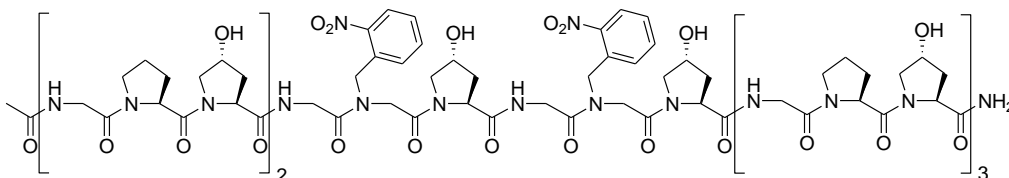
b, MALDI-MS, calculated: 2046.9 $[M+Na]^+$, observed: 2046.9 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

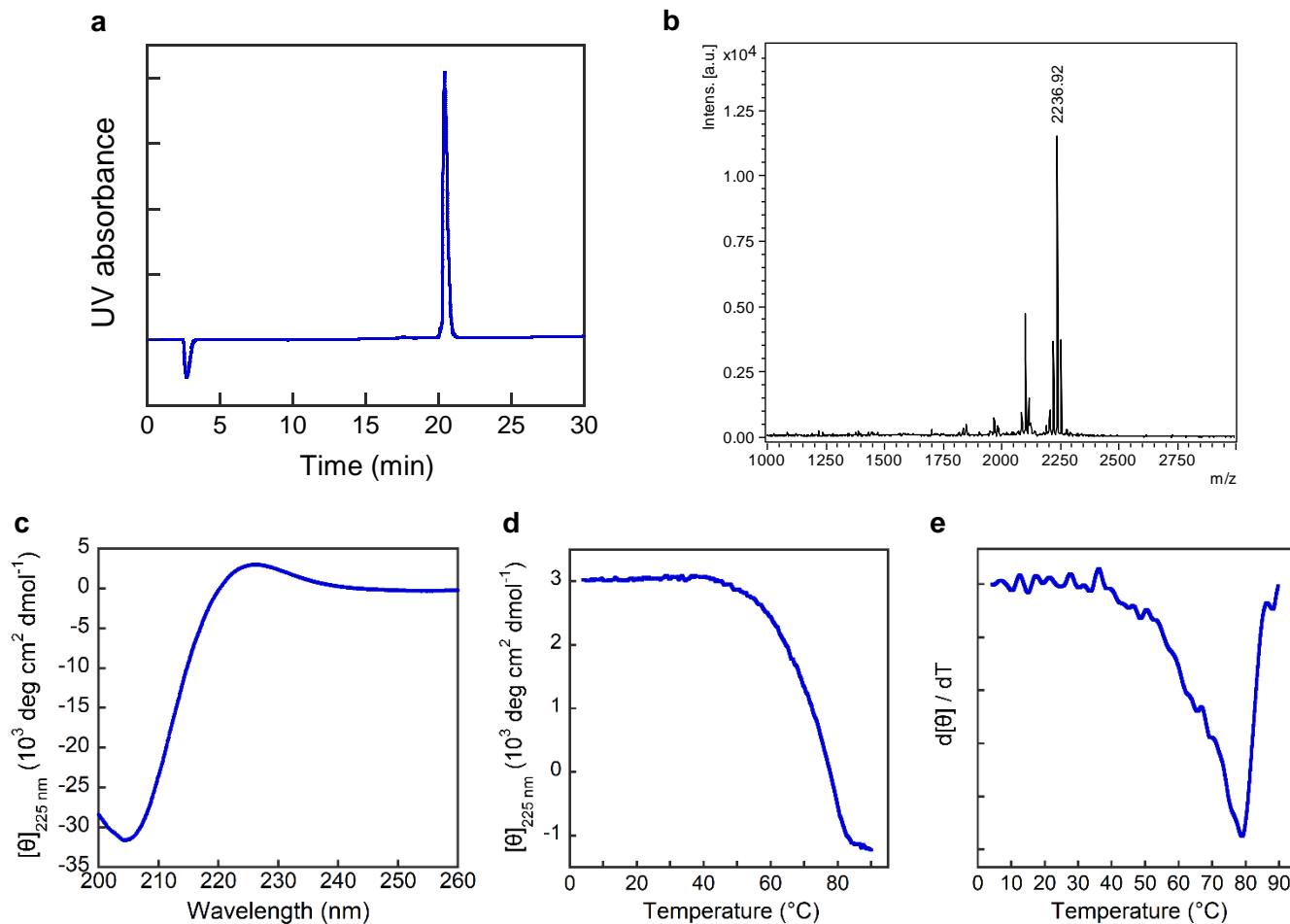
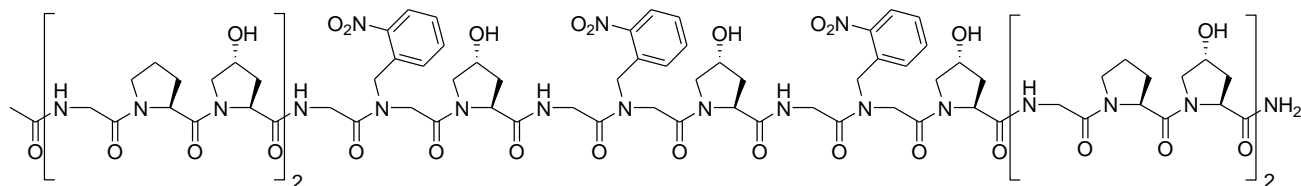
d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 62$ $^{\circ}\text{C}$.

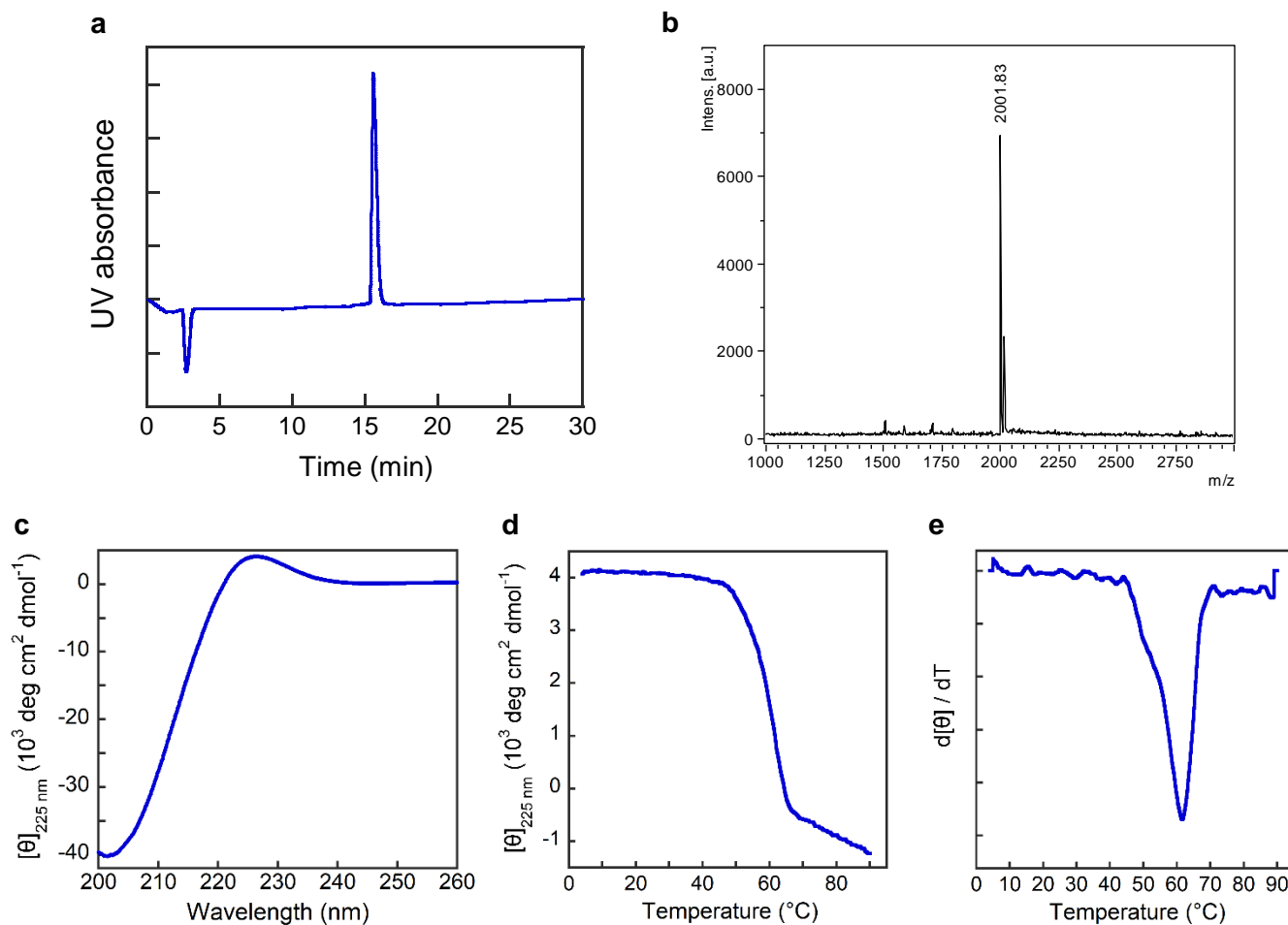
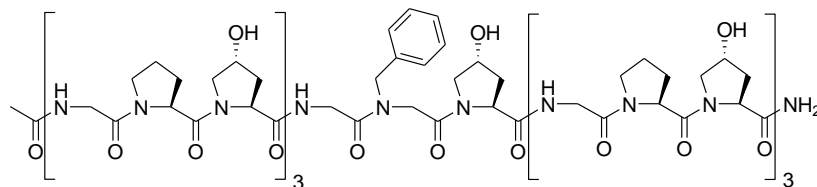
Nnbz2-CMP



Nnbz3-CMP



Nphe-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 15.6$ min.

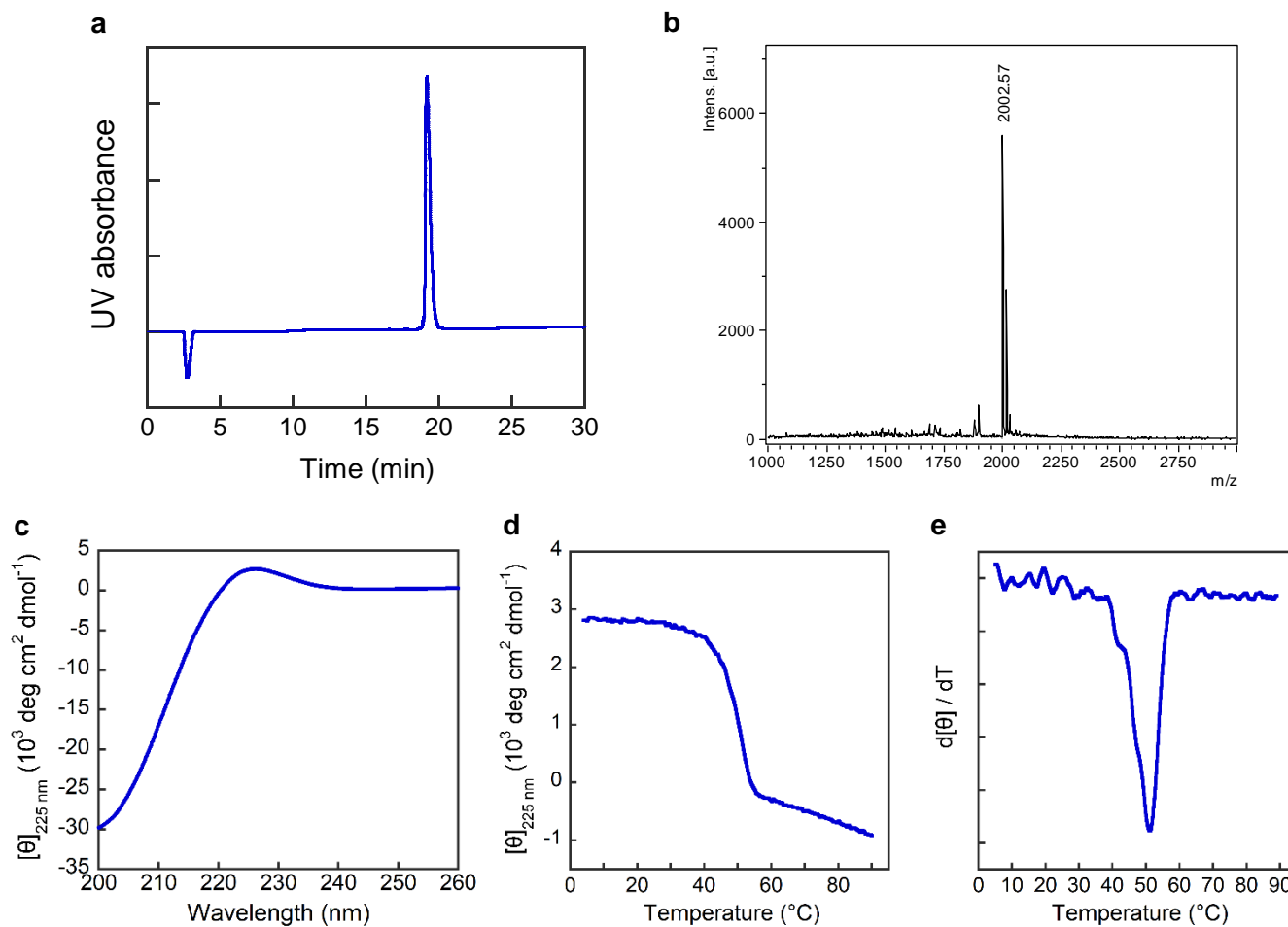
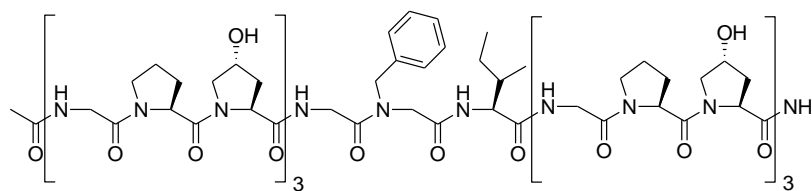
b, MALDI-MS, calculated: 2001.9 $[M+Na]^+$, observed: 2001.8 $[M+Na]^+$, 2017.8 $[M+K]^+$.

c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 62$ °C.

Nphe-Ile-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 19.2$ min.

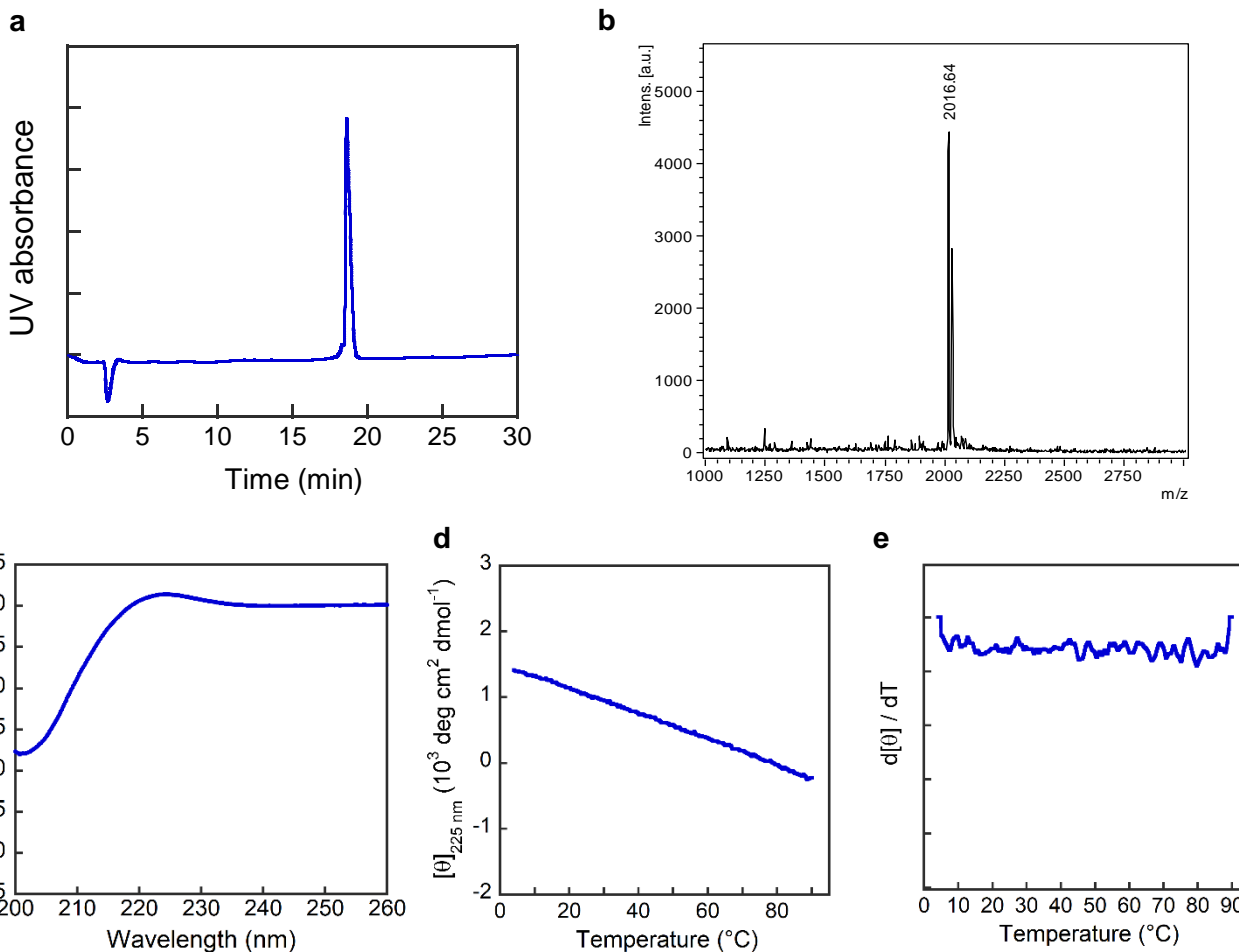
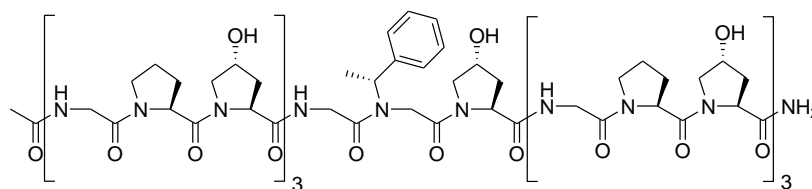
b, MALDI-MS, calculated: 2001.9 $[M+Na]^+$, observed: 2002.6 $[M+Na]^+$.

c, The CD spectrum -Ile in PBS buffer at 4°C .

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 51^{\circ}\text{C}$.

Nrpe-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 18.6$ min.

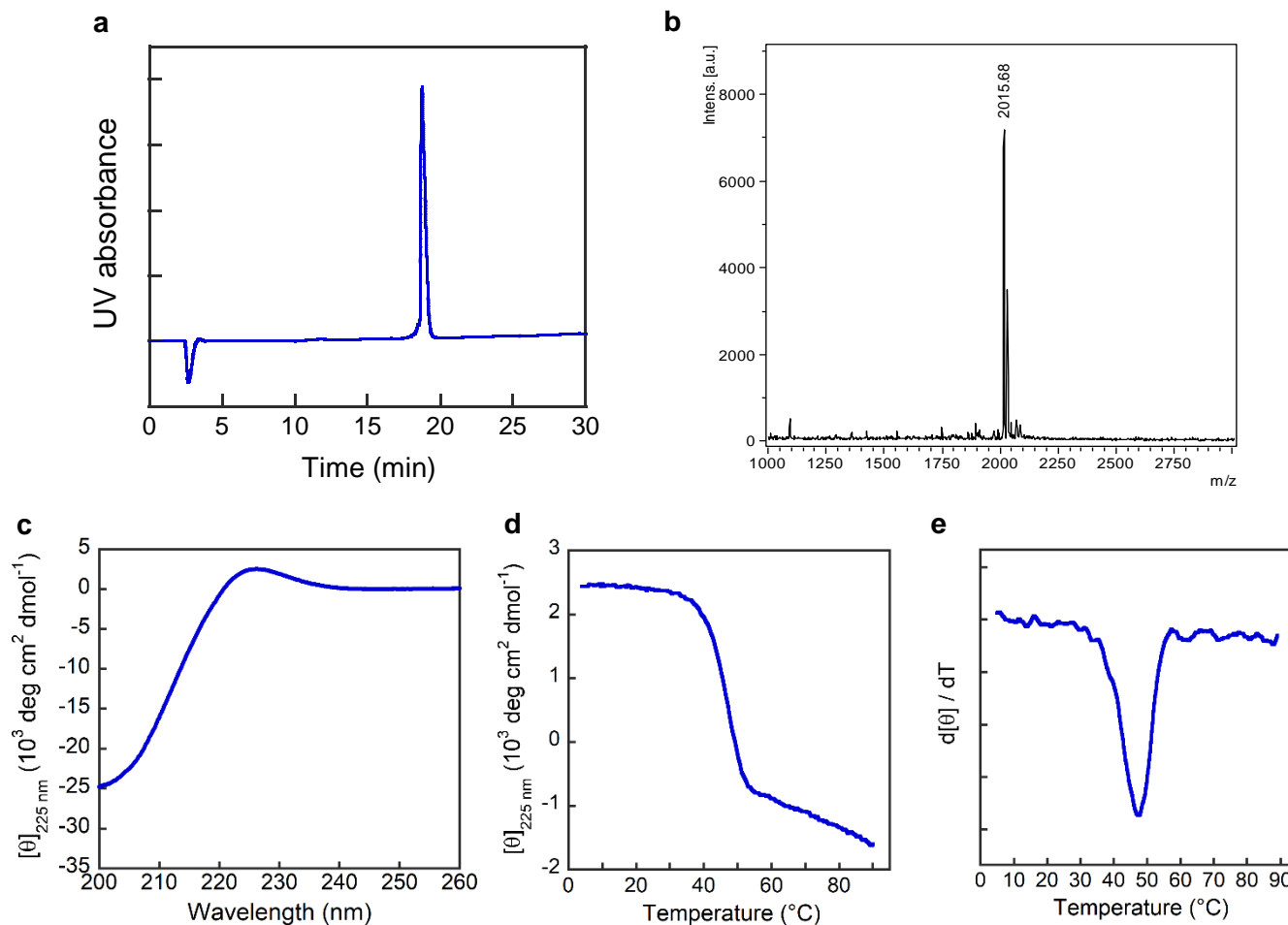
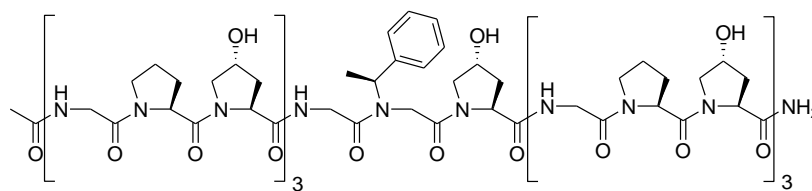
b, MALDI-MS, calculated: 2015.9 $[M+Na]^+$, observed: 2016.6 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve.

Nspe-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 18.7$ min.

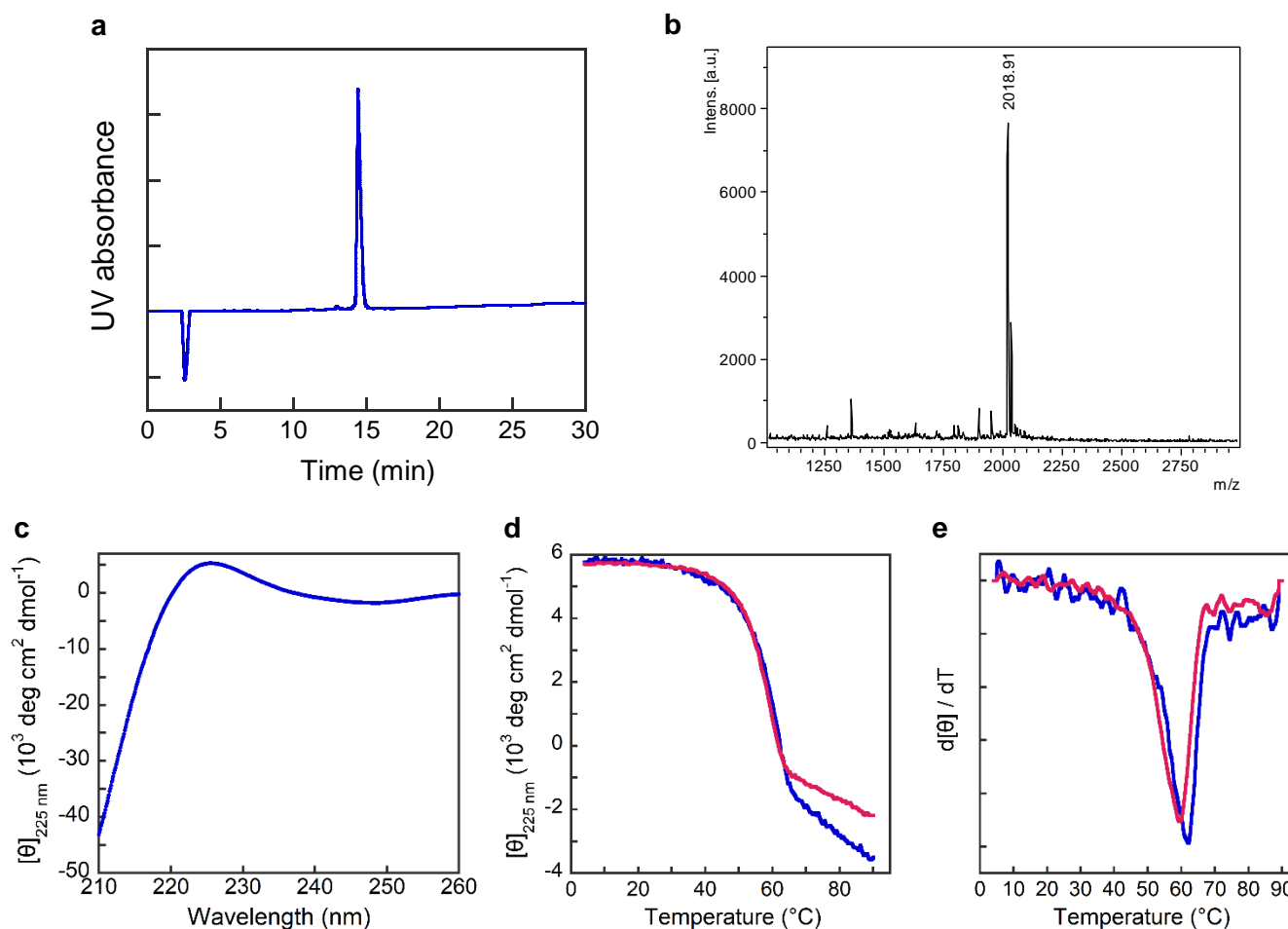
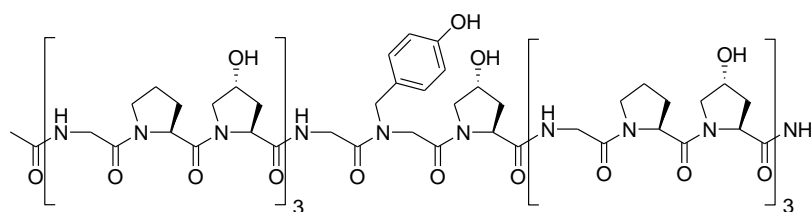
b, MALDI-MS, calculated: 2015.9 $[M+Na]^+$, observed: 2015.7 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 48$ $^{\circ}\text{C}$.

Ntyr-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 14.4$ min.

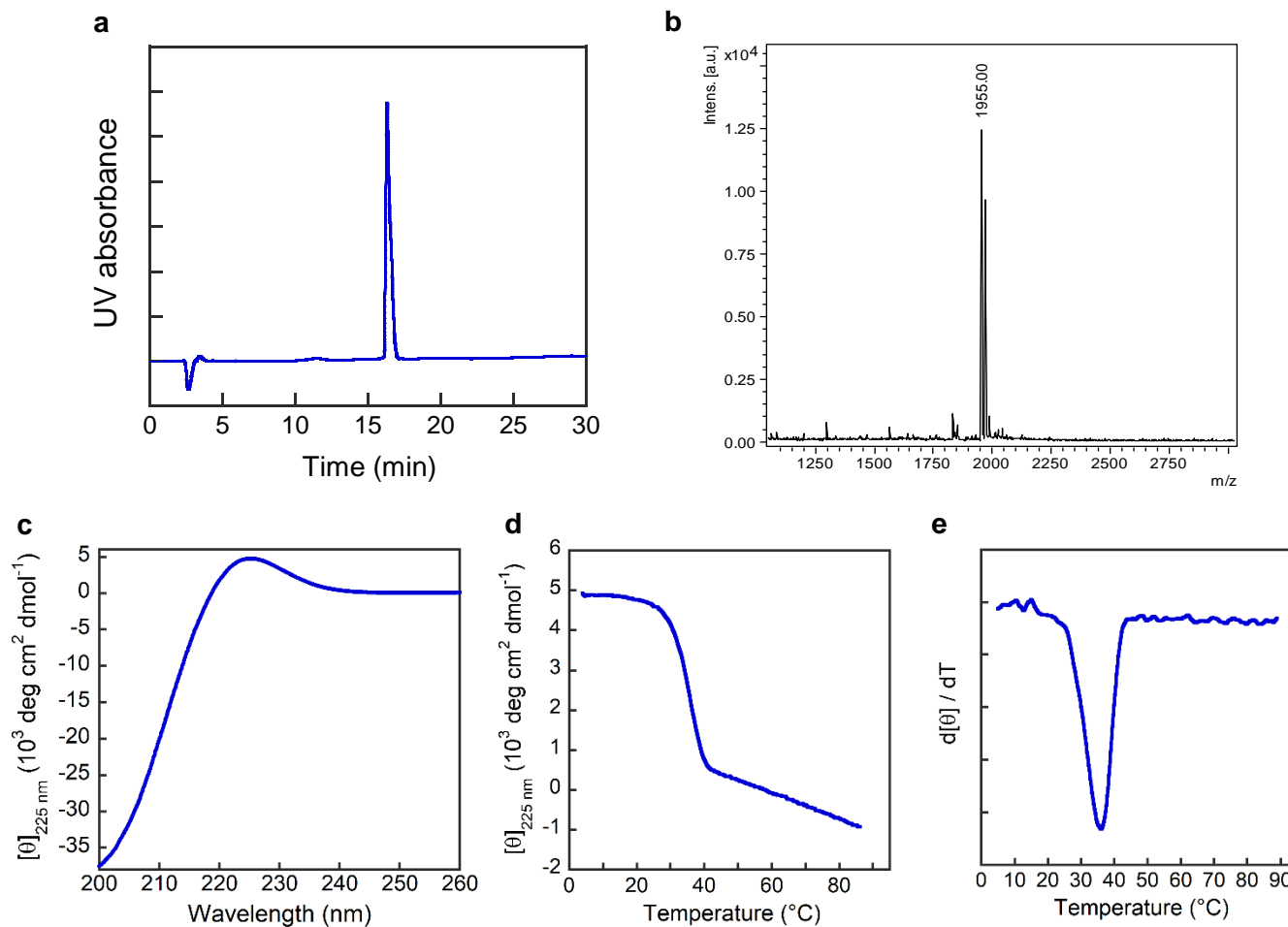
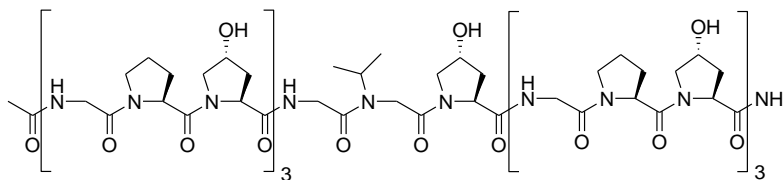
b, MALDI-MS, calculated: 2017.9 $[M+Na]^+$, observed: 2018.9 $[M+Na]^+$.

c, The CD spectrum in 10 mM NaOH solution (pH 12) at 4 °C.

d, The CD thermal melting curves in PBS buffer (blue) and 10 mM NaOH solution (pH 12, red).

e, The first derivatives of the melting curves, $T_m = 61$ °C (PBS, blue) or 59 °C (NaOH, red).

Nval-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 16.3$ min.

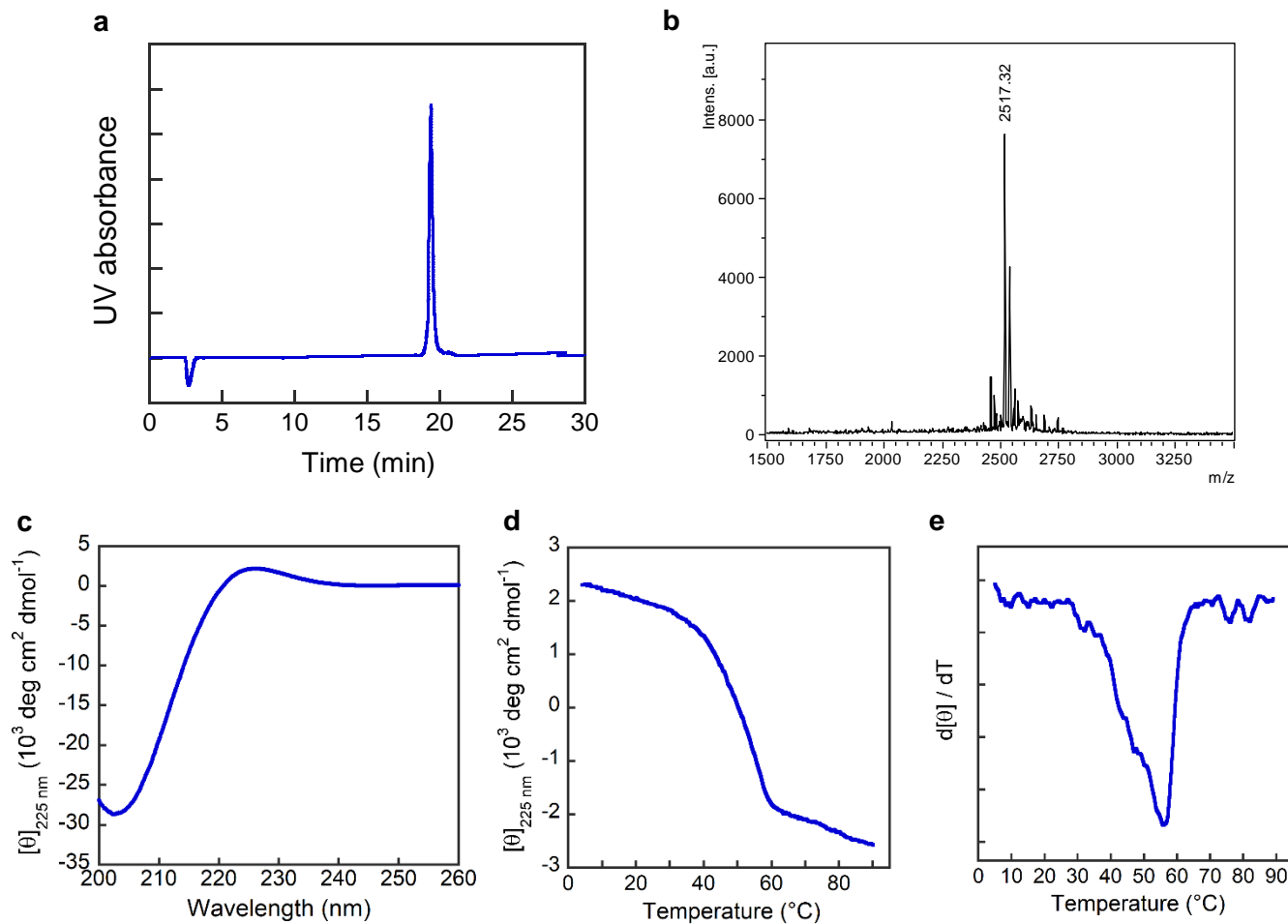
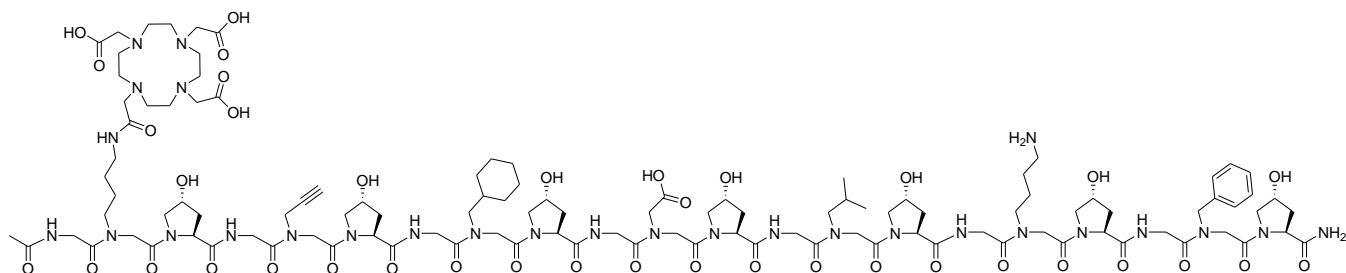
b, MALDI-MS, calculated: 1953.9 $[M+Na]^+$, observed: 1955.0 $[M+Na]^+$.

c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 36$ $^{\circ}\text{C}$.

X7-CMP



a, The HPLC chromatogram of purified peptide, $t_R = 19.4$ min.

b, MALDI-MS, calculated: 2516.2 $[\text{M}+\text{H}]^+$, observed: 2517.3 $[\text{M}+\text{H}]^+$.

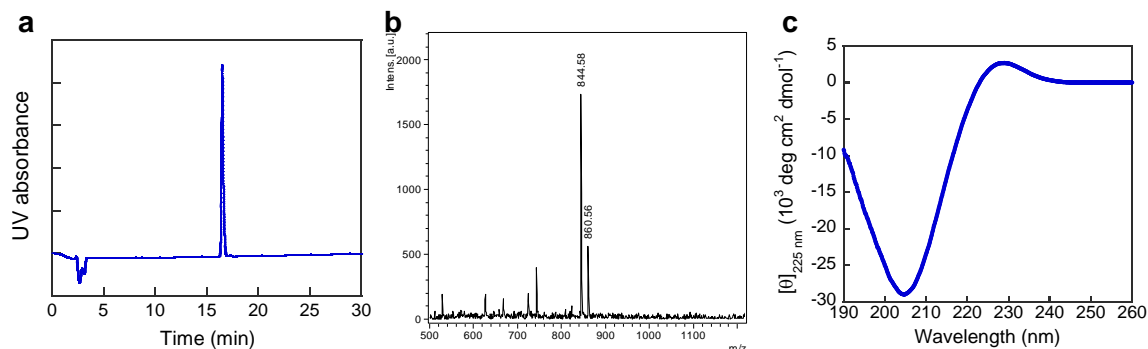
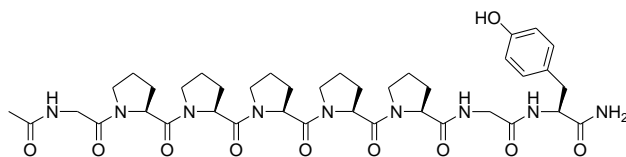
c, The CD spectrum in PBS buffer at 4 $^{\circ}\text{C}$.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, $T_m = 56$ $^{\circ}\text{C}$.

X-PP5 peptides

Pro-PP5

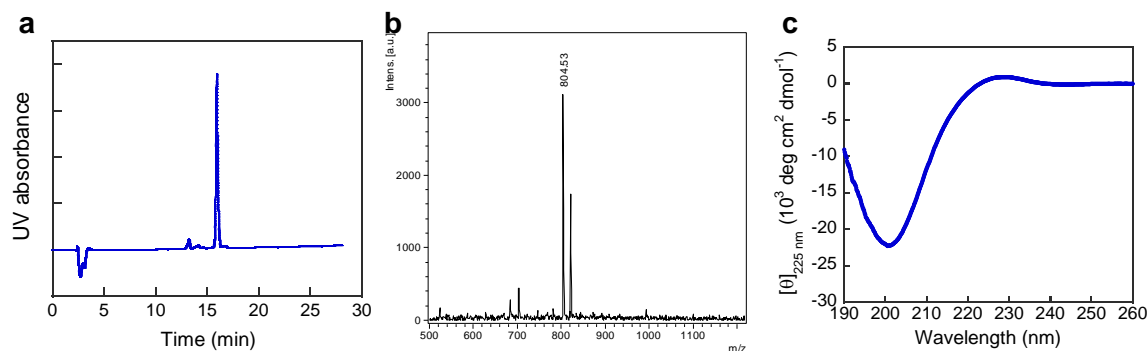
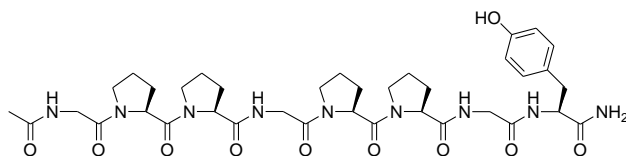


a, The HPLC chromatogram of purified peptide, $t_R = 16.5$ min.

b, MALDI-MS, calculated: 844.4 $[M+Na]^+$, observed: 844.6 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Gly-PP5

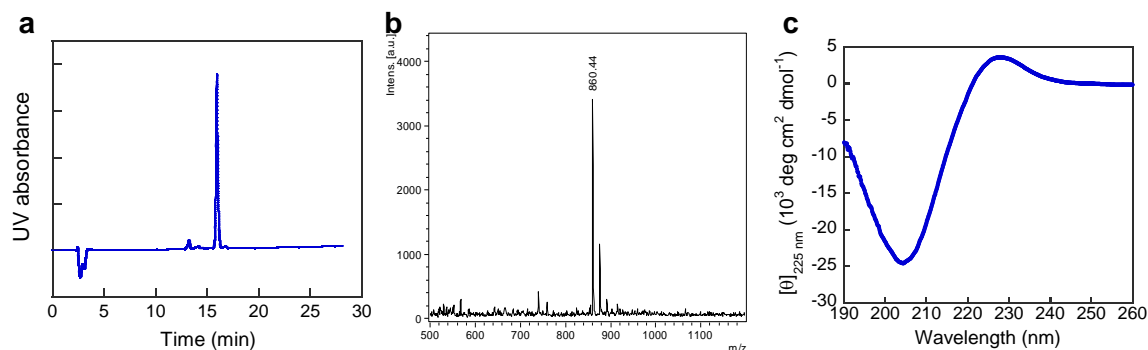
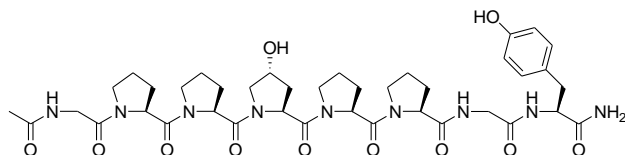


a, The HPLC chromatogram of purified peptide, $t_R = 15.9$ min.

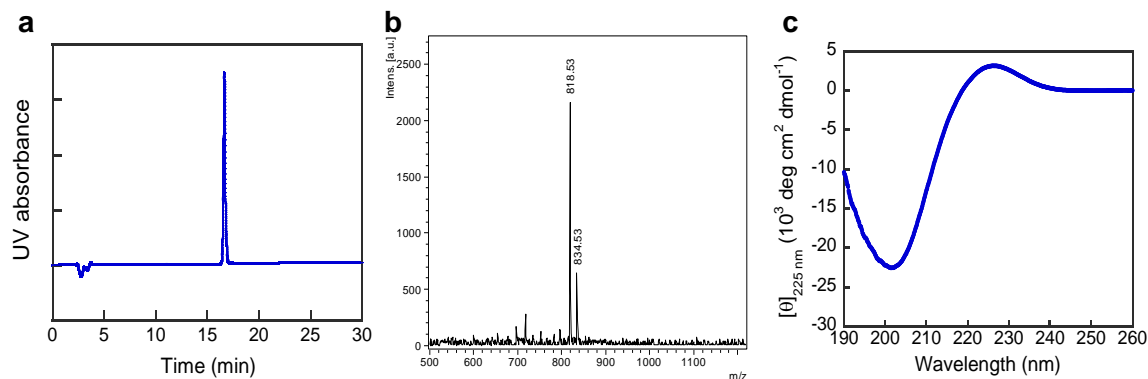
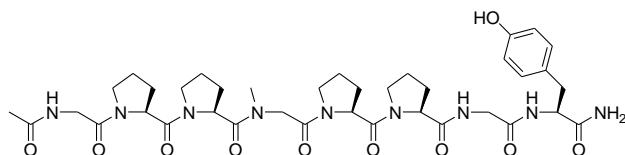
b, MALDI-MS, calculated: 804.4 $[M+Na]^+$, observed: 804.5 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

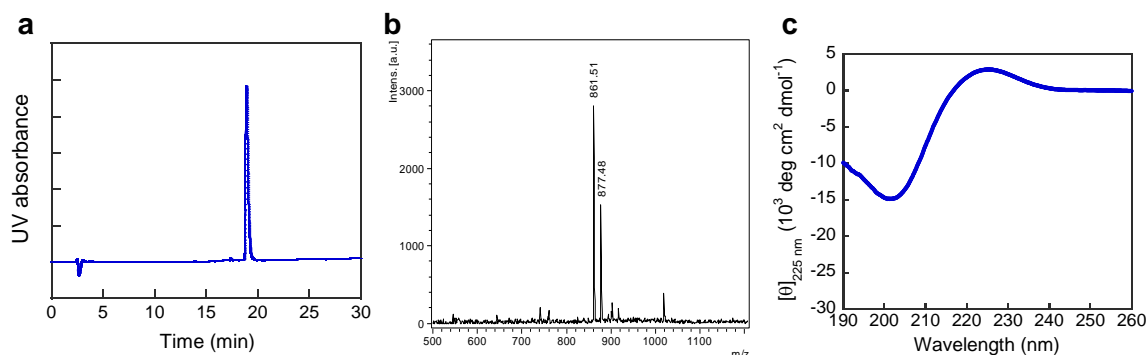
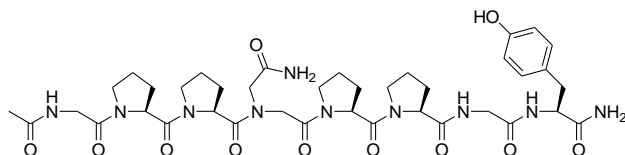
Hyp-PP5



Sar-PP5



Nasn-PP5

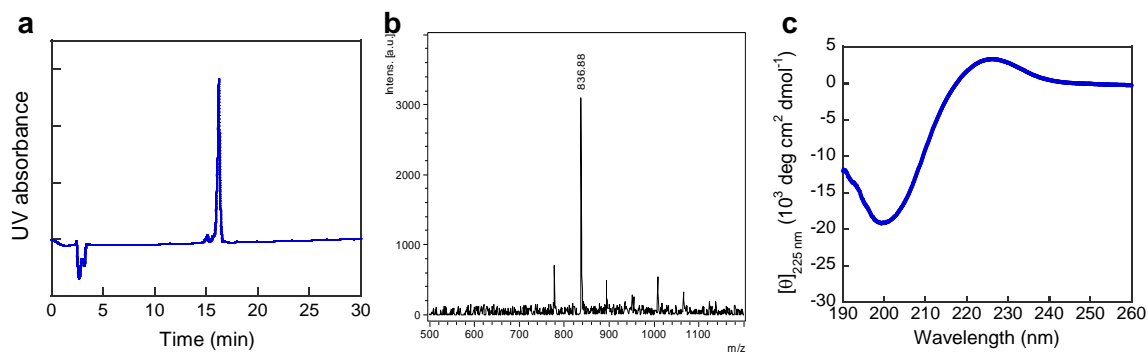
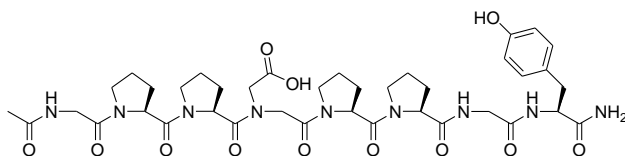


a, The HPLC chromatogram of purified peptide, $t_R = 18.9$ min.

b, MALDI-MS, calculated: 861.4 $[M+Na]^+$, observed: 861.5 $[M+Na]^+$, 877.5 $[M+K]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Nasp-PP5

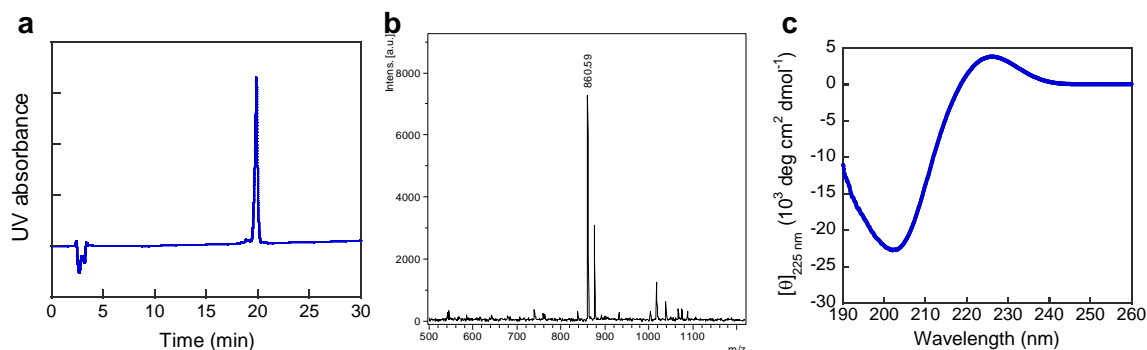
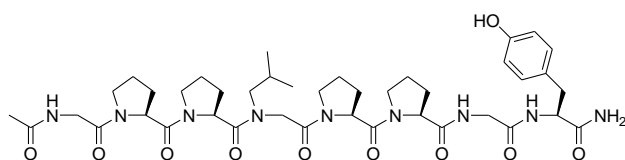


a, The HPLC chromatogram of purified peptide, $t_R = 16.2$ min.

b, MALDI-MS, calculated: 838.4 $[M-H]^-$, observed: 836.9 $[M-H]^-$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Nleu-PP5

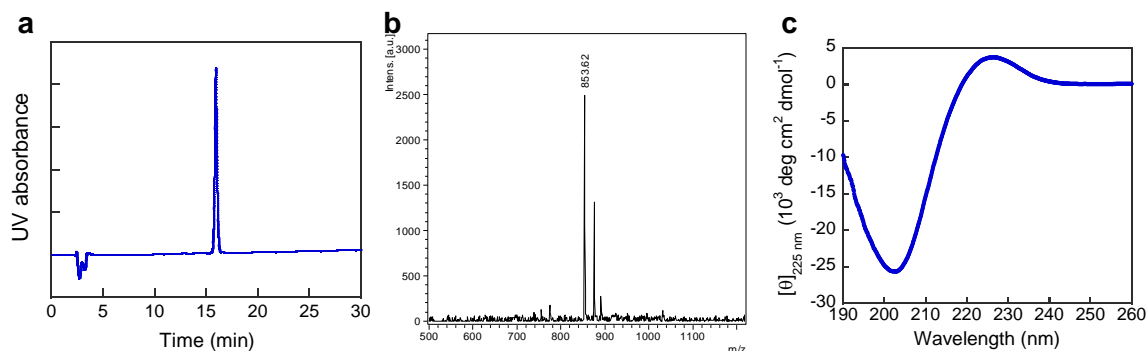
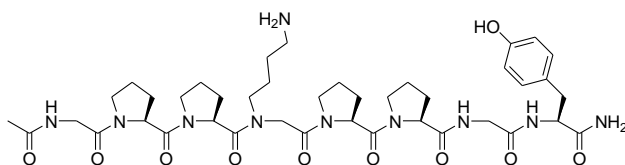


a, The HPLC chromatogram of purified peptide, $t_R = 19.9$ min.

b, MALDI-MS, calculated: 860.4 $[M+Na]^+$, observed: 860.6 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Nlys-PP5

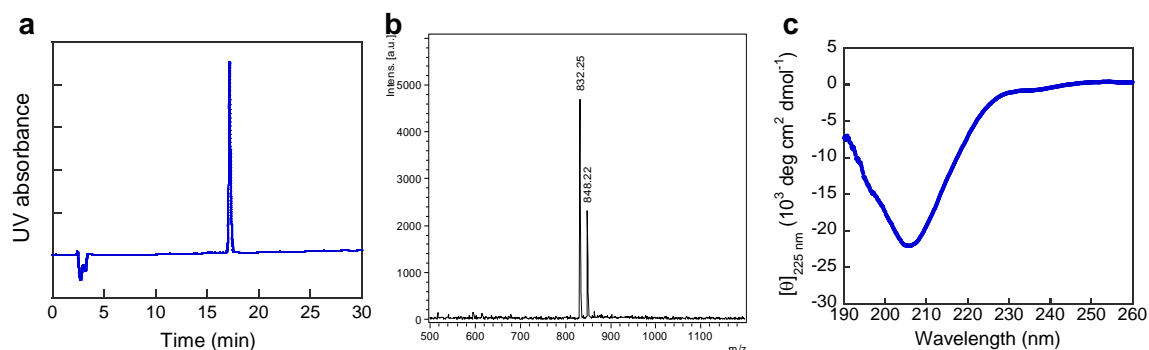
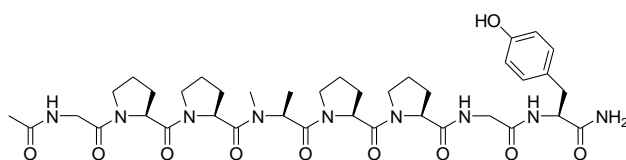


a, The HPLC chromatogram of purified peptide, $t_R = 15.9$ min.

b, MALDI-MS, calculated: 853.4 $[M+H]^+$, observed: 853.6 $[M+H]^+$, 875.4 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

NMe-Ala-PP5

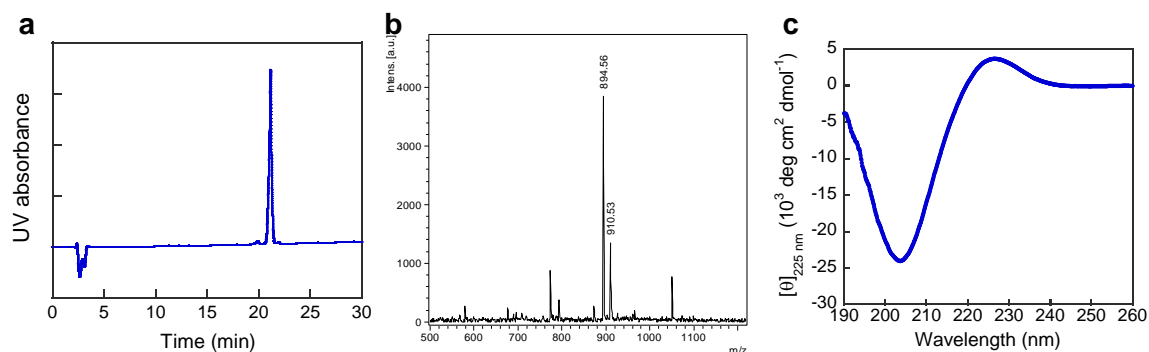
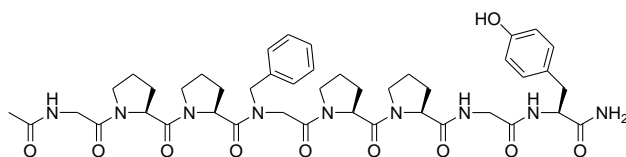


a, The HPLC chromatogram of purified peptide, $t_R = 17.1$ min.

b, MALDI-MS, calculated: 832.4 $[M+Na]^+$, observed: 832.3 $[M+Na]^+$, 848.2 $[M+K]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Nphe-PP5

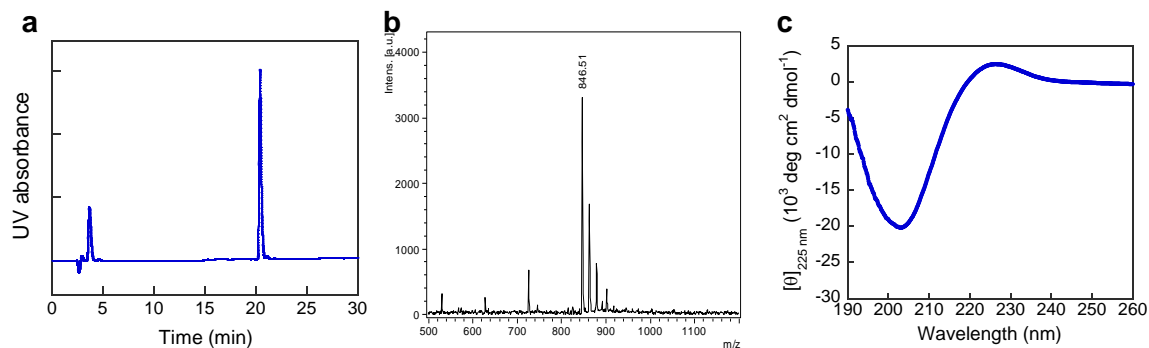
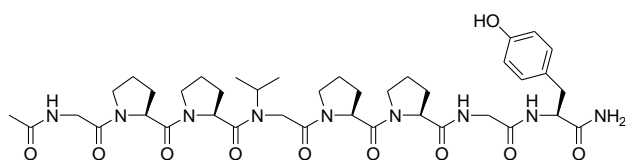


a, The HPLC chromatogram of purified peptide, $t_R = 21.1$ min.

b, MALDI-MS, calculated: 894.4 $[M+Na]^+$, observed: 894.6 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Nval-PP5



a, The HPLC chromatogram of purified peptide, $t_R = 21.0$ min.

b, MALDI-MS, calculated: 846.4 $[M+Na]^+$, observed: 846.5 $[M+Na]^+$.

c, The CD spectrum in 5 mM phosphate buffer at 25 °C.

Supplementary Section 4: Supplementary Discussion

X-CMP functionalization through click chemistry. One simple way to introduce functionality into a CMP is to install an alkyne or azido group, which allows facile derivatization through ‘click’ chemistry. To prepare an X-CMP capable of undergoing such reactions, we prepared Nakn-CMP which features a central N-gly residue bearing an alkyne moiety. Nakn’s unbranched aliphatic sidechain is similar to that of NEt (**Fig. 6a**) and the two residues produced almost identical CD melting temperatures. Click reactions between the sidechain alkyne and azides bearing unprotected hydroxyl or mannose moieties proceeded efficiently on resin in the presence of Cu(I)/TBTA (**Fig. 6a and SI Methods**). Following click conjugation with 2-azidoethanol, the T_m of the Nakn triple helix increased by 3 °C, presumably due to the increased size of the *N*-C α -triazole ring (**Fig. 6a**). Conjugation with 2-azidoethylmannose had almost no effect on T_m (**Fig. 6a**). Previously, click reactions were used to functionalize CMP via (4*R*)azidoproline; however the formation of the triazole-proline units resulted in T_m drop of at least 7 °C¹¹. Our results demonstrate that the *N*-appended sidechain can better tolerate spatially demanding triazole units within the triple helix than C γ -substituted proline derivatives¹¹ and showcase a superior approach for click-functionalization of CMPs.

References

1. Zuckermann, R. N., Kerr, J. M., Moosf, W. H. & Kent, S. B. H. Efficient Method for the Preparation of Peptoids [Oligo(N-substituted glycines)] by Submonomer Solid-Phase Synthesis. *J. Am. Chem. Soc.* **114**, 10646–10647 (1992).
2. Li, Y. *et al.* Targeting collagen strands by photo-triggered triple-helix hybridization. *Proc. Natl. Acad. Sci.* **109**, 14767–14772 (2012).
3. Li, Y. *et al.* Non-Covalent Photo-Patterning of Gelatin Matrices Using Caged Collagen Mimetic Peptides. *Macromol. Biosci.* **15**, 52–62 (2015).
4. Otwinowski, Z. & Minor, W. Processing of X-Ray Diffraction Data Collected in Oscillation Mode. in *Methods in Enzymology* (eds. Carter Jr., C. W. & Sweet, R. M.) vol. 276 307–326 (Academic Press, 1997).
5. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
6. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
7. Murshudov, G. N. *et al.* REFMAC 5 for the refinement of macromolecular crystal structures research papers. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D67**, 355–367 (2011).
8. Vanommeslaeghe, K. *et al.* CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **31**, 671–690 (2010).
9. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
10. Binkley, J. S., Pople, J. A. & Hehre, W. J. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.* **102**, 939–947 (1980).
11. Erdmann, R. S. & Wennemers, H. Functionalizable collagen model peptides. *J. Am. Chem. Soc.* **132**, 13957–13959 (2010).

ChemRxiv-SI-Peptoid-CMP_submitted (20200905).pdf (4.82 MiB)

[view on ChemRxiv](#) • [download file](#)
