Peptoid Residues Make Diverse, Hyperstable Collagen Triple-Helices

Julian L. Kessler[†], Grace Kang[†], Zhao Qin[‡], Helen Kang[†], Frank G. Whitby[§], Thomas E. Cheatham III[^], Christopher P. Hill[§], Yang Li[†,*], and S. Michael Yu[†,‖]

[†]Department of Biomedical Engineering, University of Utah, Salt Lake City, Utah 84112, USA
[‡]Department of Civil & Environmental Engineering, Collagen of Engineering & Computer Science, Syracuse University, Syracuse, New York 13244, USA
[§]Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[^]Department of Medicinal Chemistry, College of Pharmacy, L. S. Skaggs Pharmacy Research Institute, University of Utah, Salt Lake City, Utah 84112, USA
[‖]Department 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

Section 1: 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 (Figure 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 (Figure 6a, see Materials and Methods for synthesis protocols). 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-Cα-triazole ring (Figure 6a). Conjugation with 2-azidoethylmannose had almost no effect on T_m (Figure 6a). Previously, click reactions were used to functionalize CMP *via* (*4R*)azidoproline; however the formation of the triazole-proline units resulted in T_m drop of at least 7 °C[1]. Our results demonstrate that the N-appended sidechains at the X position can better tolerate spatially demanding triazole units within the triple-helix than Cγ-substituted proline derivatives[1], and it showcase a perhaps more stable approach for click-functionalization of triple-helical CMPs.

Section 2: 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
N-Methylpyrrolidinone (NMP)	Fisher Scientific	BP1172
N,N-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
N,N'-Diisopropylcarbodiimide (DIC)	Sigma-Aldrich	D125407
O-(7-Azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU)	Chem-Impex	12881
1-Hydroxy-7-azabenzotriazole (HOAT)	AAPPTec	CXZ012
(7-Azabenzotriazol-1-yloxy)trispyrrolidinophosphonium hexafluorophosphate (PyAOP)	AAPPTec	CXZ070
N-Ethyldiisopropylamine (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 µmol 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 µmol), HATU (50 µmol, 19 mg), HOAT (50 µmol, 6.8 mg) and DIEA (75 µmol, 13 µL) 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 (3×10 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 others[2].

(1)   (2)
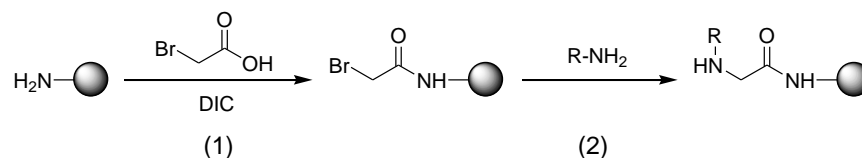
(1) Acylation reactions were performed by addition of a solution of bromoacetic acid (100 µmol, 13.9 mg) and DIC (98 µmol, 15.2 µL) in DMF (0.5 mL) to 10 µmol of resin-bound amine. Reaction mixtures were agitated at room temperature for 30 min. Each acylation was repeated once before the resin was drained and washed with NMP (3×10 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 (3×10 mL). The reaction was monitored by the chloranil test. Typically, resin beads turned dark green in a chloranil test following a displacement reaction of 2 hr, suggesting successful coupling of a peptoid residue; 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(*N*-o-nitrobenzyl)Gly-OH	Synthesized in-house[3,4]		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 sidechain 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	
Nhcy	2-(tert-butylsulfanyl)ethan-1-amine	Enamine	EN300-110537	
Ndxn	C-[1,4]Dioxan-2-yl-methylamine	Matrix Scientific	008403	a racemic mixture of *R* and *S* (1,4)dioxan-2-yl-methylamine
NEt	Ethylamine solution (2.0 M in THF)	Sigma-Aldrich	395072	
Nleu	Isobutylamine	Sigma-Aldrich	I14150	
Nlys	*N*-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 µmol) in a solution of Fmoc-Gly-OH (100 µmol) and DIC (100 µmol) 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 µmol) in a solution of Fmoc-Gly-OH (100 µmol) and DIC (100 µmol) in 0.5 mL of DMF.
Ndota	1,4,7,10-Tetraazacyclododecane-1,4,7-tris(t-butyl acetate)-10-(4-aminobutyl)acetamide	Macrocyclics	B-279	The displacement reaction was completed by agitating the resin (10 µmol) in 0.5 mL of NMP containing the primary amine (63 µmol) and DIEA (115 µmol) overnight.

On-resin functionalization *via* "click" chemistry

The azido compound (60 µmol), such as 2-Azidoethanol (Carbosynth, FA07084) or 2-Azidoethyl α-D-mannopyranoside (Synthose, AM482), was dissolved in 0.5 mL of DMF along with [Cu(MeCN)$_4$]PF$_6$ (10 µmol, 3.7 mg, Sigma-Aldrich, 346276) and TBTA (20 µmol, 10.6 mg, Sigma-Aldrich, 678937). This solution was added to the resin-bound alkyne-containing peptide Nakn-CMP (5 µmol), and the mixture was agitated at room temperature overnight. The resin was drained and washed with DMF (5×10 mL) before the "click"-functionalized product was cleaved.

Cleavage protocols

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 µmol) 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 (3×10 mL) and DCM (5×10 mL). Unless otherwise mentioned below, the resin (5 µmol) 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	The resin was treated with a mixture of TFA/EDT/water/TIS (94:2.5:2.5:1) for over 3 h.
Nhcy	*Step 1:* For 5 µmol of crude product, the resin was treated with a mixture (1 mL) of TFA/TIS/water (95:2.5:2.5) for over 2 hr, then the crude product was collected.
	Step 2: The tBu protective group was removed from the Nhcy 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, then the product was precipitated and collected again.
	Step 3: The crude product was dried then dissolved 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 contained 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 $M^{-1}cm^{-1}$ per peptide bond) or 280 nm for the Tyr-containing polyproline host-guest sequences (extinction coefficient: 1280 $M^{-1}cm^{-1}$ 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 sidechains 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 μL 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 μL of the peptide solution and 0.3 μL 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 μL 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 HKL2000[5]. Phases were determined by molecular replacement with Phaser-MR[6] using PDB structure 1G9W as a search model. Models were built with COOT[7] and refined with Refmac5[8]. UCSF Chimera version 1.14 was used to render molecular structure figures. Data and refinement statistics are indicated in Table S3.

Metadynamics calculation of energy landscape

We applied well-tempered metadynamics to study the thermal stability of the peptoid conformation within the X-CMP system. This technique enhances the sampling in molecular simulations and enabled us to measure the free energy landscape as a function of the two backbone dihedral angles φ (defines how the peptoid unit folds against the previous amino acid) and ψ (defines how the peptoid unit folds against the following amino acid). The metadynamics simulations are performed by NAMD and the PLUMED package[9–11], with molecular modeling details given in the method below. This method has been applied before to study the energy landscape for adhesion[12] and the thermal stability of protein[13]. The energy deposition event occured every 200 fs in the form of a Gaussian with height 0.01 kcal mol^{-1} and width of 0.35 rad for each dihedral angle. The resulting free energy landscape as a function of time is obtained by summing all Gaussians deposited up to a specified time. We summarized the calculation results in Figure 4c, which include the results after depositing and summing 200,000 Gaussians for different molecular models. This number of Gaussians was used because it yields the convergence of the free energy landscape. It is shown in Figure 4c that different center acid yields a different energy landscape. The native state is highly similar to the expected polyproline conformation and the (φ, ψ) values corresponding to the lowest free energy in each X-CMP are (-63.1°, 180°) for Gly, (-86.0°, -166.2°) for Sar, (-74.5°, 180°) for Nleu, and (-45.9°, 143.3°) for Pro. Besides the native state, Gly and Sar have a second energy well with slightly higher energy level and an energy barrier (-16.9 kcal/mol for Gly and -19.4 kcal/mol for Sar) to cross. However, the energy landscapes of Nleu (well depth of -21.7 kcal/mol) and Pro (well depth of -19.9 kcal/mol) do not give a clear second energy

well. The energy minimum of Nleu is apparently more localized than Pro, indicating that an X-CMP containing Nleu may be more thermally stable than Pro.

Molecular dynamics simulations of X-CMPs

We built the molecular structure and atomic interaction of X-CMPs by combining DFT calculation with CHARMM General Force field (CHARMM GenFF)[14]. The standard CHARMM27 force field[15] 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 set[16] 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 fair[14]. The intrinsic coordinate of each peptoid molecule was also defined according to the optimized structure, ensuring the initial geometry had optimum energy.

Classical full atomistic molecular dynamics (MD) simulations were used to fully equilibrate the X-CMP system in the explicit solvent composed of 100 mM NaCl and TIP3P water prior to metadynamics using the updated CHARMM27 all atom energy force field[15]. The simulation system was $11 \times 5 \times 5$ nm^3 in size. The net charge of the system was set to zero by adjusting the ratio between cation and anion; each ion was initially distributed randomly in water and at least 5 Å from the peptide-peptoid structure. For the initial run, each of the backbone atoms were constrained in space by elastic springs with stiffness of 5 kcal/mol/Å^2 in all 3 directions, while the sidechain atoms and solvent molecules were free to move. The simulation time step was 2 fs with rigid bonds modeled for all the covalent bonds between hydrogen atoms and other heavy atoms. We used particle mesh Ewald (PME) function with a grid width <1 Å to calculate the electronic interaction as an efficient method to accurately include all the long-distance electrostatics interactions. We ran 500,000 steps of energy minimization by using a conjugate gradient and line search algorithm followed by a 1 ns dynamics run for initial equilibrium. The dynamics simulation was performed in the NPT (constant pressure of 1 atm and constant temperature of 310 K) ensemble controlled by the Langevin dynamics to reach a constant pressure and temperature (with 5 ps^{-1} damping coefficient for temperature, 100 fs oscillation period and 50 fs damping time scale for pressure control) with the shrinkable volume during the run. After the initial run, we restarted the simulation, released the constraints on the backbone atoms, and continued using NPT ensemble for 20 ns for the purpose of getting a fully equilibrated X-CMP structure. The simulations were performed by using a NAMD package[2] (v2.12) in the local Linux workstation with multiple graphics processing units (GeForce GTX 1080) support through Compute Unified Device Architecture (CUDA v8.0).

Gelatin binding assays

Fluorescein-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-MS, calculated and observed: 2440 [M+Na]$^+$; F-Nnbz2-CMP, MALDI-MS, calculated and observed: 2572 [M+Na]$^+$). The pure fluorescent peptides were lyophilized and dissolved in 1×PBS.

Wells of a 96-well plate were coated with approximately 6 µL 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 µL/well) containing 2 mM of NHS (N-hydroxysuccinimide) and 10 mM EDC [1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide] with shaking overnight at room temperature. The crosslinked gelatin films were washed with 200 µL of 25 °C 1×PBS 10 times. PBS solution (50 µL) containing 10 µM 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/cm^2). 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 a powder of Cys-GlyGlyGly-(GlyProHyp)$_2$-(GlyNnbzHyp)$_2$-(GlyProHyp)$_3$ (MALDI-MS, calculated 2352.3 [M+H]$^+$, observed: 2353.1 [M+H]$^+$, 8 eq). Gelation was observed almost immediately. Conjugation efficiency was measured by HPLC, which showed that 70% of the Cys-Nnbz2-CMP was conjugated to the multi-arm polymer (Figure S5). Fluorescently labeled gels were prepared by dissolving dry MA-PEG-CMP in a PBS (1×) solution containing 650 µM of F-Nnbz2-CMP. The Nnbz2-CMP conjugated MA-PEG-CMP hydrogel (50 µL) was melted by heating to 80 °C, added to the bottom of a microtube, and allowed to gel at 4 °C for 15 min. After gelation, 250 µL 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/cm^2) 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 F-Nnbz2-CMP and dissolution of the gel (Figure S5).

To utilize the MA-PEG-CMP (Nnbz2-CMP) to spatially control cell attachment (Figure S6) and migration (Figure 6f), we prepared a thin gelatin film on the bottom of a 6-well plate (size: ~1 cm^2, thickness: ~50 μm) and crosslinked it with EDC/NHS as described above. MA-PEG-CMP was melted by heating to 80 °C and diluted in hot 1×PBS to a final concentration of 31.25 μM. The MA-PEG-CMP solution (85 μL, 80 °C) was applied to the film surface and allowed to dry completely. Wells patterned before cell addition (Figure S6) were exposed to UV light (mercury arc lamp, 365 nm, 15.5 mW/cm^2) through a transparency mask placed directly onto the film surface for 10 min. Patterns were produced using CorelDraw and printed onto transparences 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% CO$_2$ atmosphere. Growth media was changed every 2 days. Cell attachment was monitored with an EVOS light microscope (Thermofisher Scientific, Figure S6). In another experiment, after cells grew to about 80% confluency on the gelatin films modified with MA-PEG-CMP, wells were drained of growth media, and areas at the boundary between gelatin and MA-PEG-CMP were exposed to UV light (Mercury arc lamp, 365 nm, 33 mW/cm^2) for 8 min through a photomask elevated by thin spacers (0.9 mm in height, which prevented cells from contacting the film surface). After UV patterning, wells were briefly washed with 1×PBS, and growth media was replaced. The cells were incubated for an additional 24 hr before their migration were observed with the EVOS light microscope (Figure 6f).

Section 3: Supplementary Tables and Figures

Table S1. Residues, structures, and T_m values of the host-guest X-CMPs used in this study.
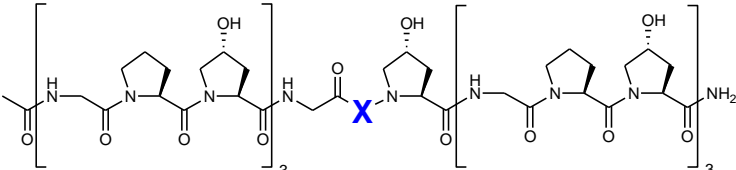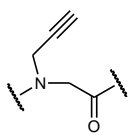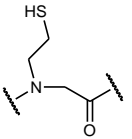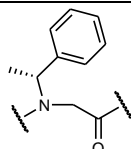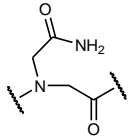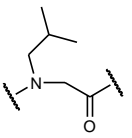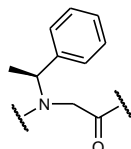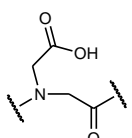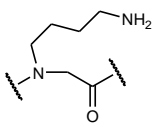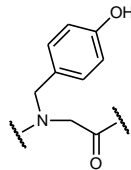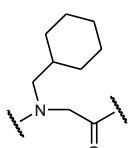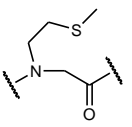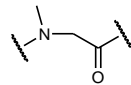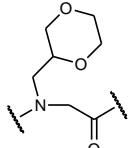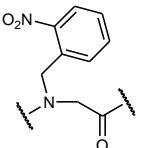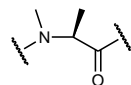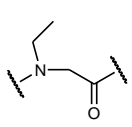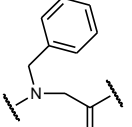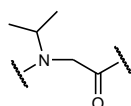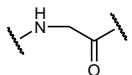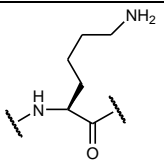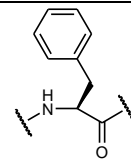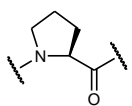
Residue	Structure	T_m (°C)
N-gly residues		

Residue	T_m (°C)	Residue	T_m (°C)	Residue	T_m (°C)
Nakn	52	Nhcy	54	Nrpe	< 4
Nasn	55	Nleu	58	Nspe	48
Nasp	50 (PBS) 56 (HCl)	Nlys	55 (PBS) 56 (NaOH)	Ntyr	61 (PBS) 59 (NaOH)
Nchx	63	Nmet	55	Sar	50
Ndxn	59	Nnbz	62	NMe-Ala	40
NEt	52	Nphe	62	Nval	36
Amino Acids					
Gly	37	Lys	47 (PBS) 46 (NaOH)	Phe	38
Pro	55				

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
Rpim	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 Rfree set	1221	2021
Rcryst	0.139 (0.319)	0.154 (0.289)
Rfree	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

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	

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[17,18], we expected that the destabilization resulting from the Y position Hyp → Ile substitution would be attenuated when X = Nphe, since hydrophobic interactions between the Nphe-Ile pair are stronger than Pro-Ile. Our results showed that Hyp → Ile decreased T_m by 10 or 11 °C regardless of the X residue, suggesting that adjacent hydrophobic interactions do not contribute significantly to the triple-helix stability. These results further support the notion that hydrophobicity of peptoid residues is not the most important factor in stabilization the collagen triple-helix.
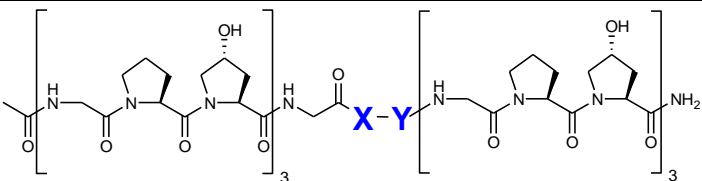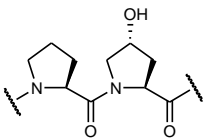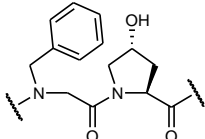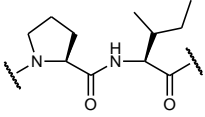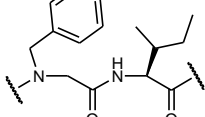
X-Y	Structure	T_m (°C)	X-Y	Structure	T_m (°C)
Pro-Hyp		55	Nphe-Hyp		62
Pro-Ile		45	Nphe-Ile		51

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	CYGSGS	HUCVIK	LOWGOT	TBXSAG	YUHDUC
BEHTEN	CYTSAR	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	ZAJDUJ
CAZSEB	GABHEX	IPUQAN	OGAZUS	VACLUJ	ZAJZAO
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	POHMUU	WEXPOE	
CUCNUJ	GIKBAF	KALVUQ	POXPAT	WEXPUK	

Table S6. Examples of N-gly-containing small molecule structures from CSD (Figure 4a).
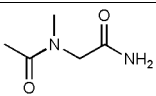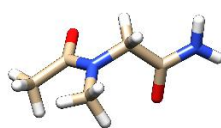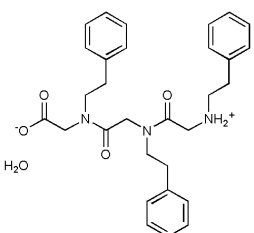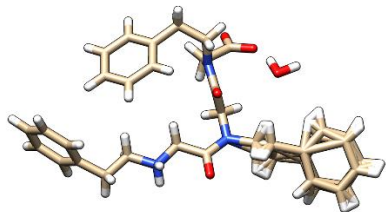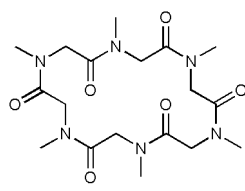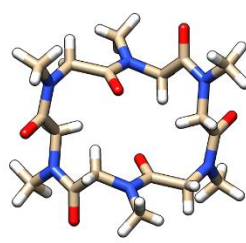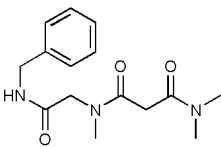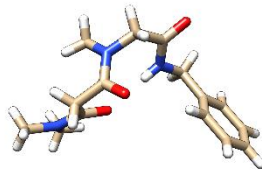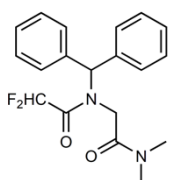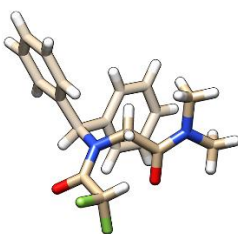
CSD-REFCODE	Chemical structure	Crystal structure
POXPAT		
ACUMUI		
HXSARM		
PAGTIA		
YUHGIT		

Table S7. UV-responsive X-CMPs. Additional Nnbz groups increase T_m indicating that peptoid stabilization by N-gly is additive. When the nitrobenzyl sidechain is removed by exposure to UV light, the triple-helix becomes dramatically less stable as the Nnbz residue is converted to Gly (Figure 6b).
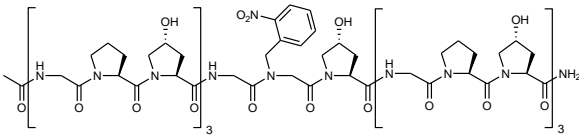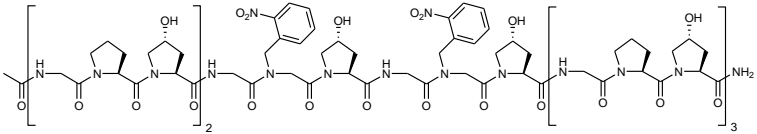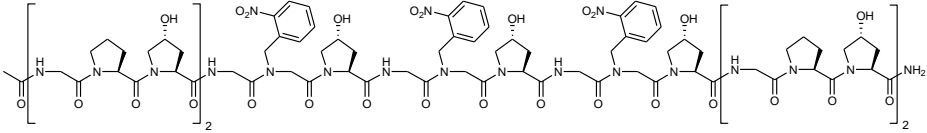
Name	Primary structure	T_m (°C)	T_m (°C) post UV
Nnbz-CMP		62	37
Nnbz2-CMP		73	18
Nnbz3-CMP		79	< 4

Table S8. The *cis-trans* ratios of the amide bond of selected peptoid residues in this study.
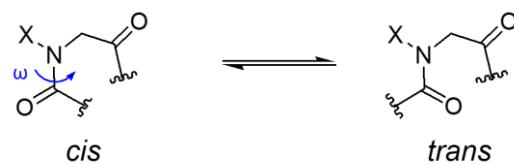
cis   trans

Residue	Sidechain X	$K_{cis/trans}$ (solvent)	Reference	T_m of X-CMP (°C)
Pro		0.22 (D$_2$O) 0.26 (D$_2$O) 0.33 (dioxane)	*J. Am. Chem. Soc.* 2001, 123, 4, 777–778 *Org. Lett.* 2020, 22, 2, 348–351 *J. Am. Chem. Soc.* 2002, 124, 11, 2497–2505	55
Sar	CH$_3$	0.42 (CD$_3$OD) 0.60 (CD$_3$CN)	*J. Am. Chem. Soc.* 2009, 131, 45, 16555–16567	50
NEt		0.47 (CD$_3$OD) 0.66 (CD$_3$CN)		52
Nphe		0.62 (CD$_3$OD) 1.15 (CD$_3$CN)		62
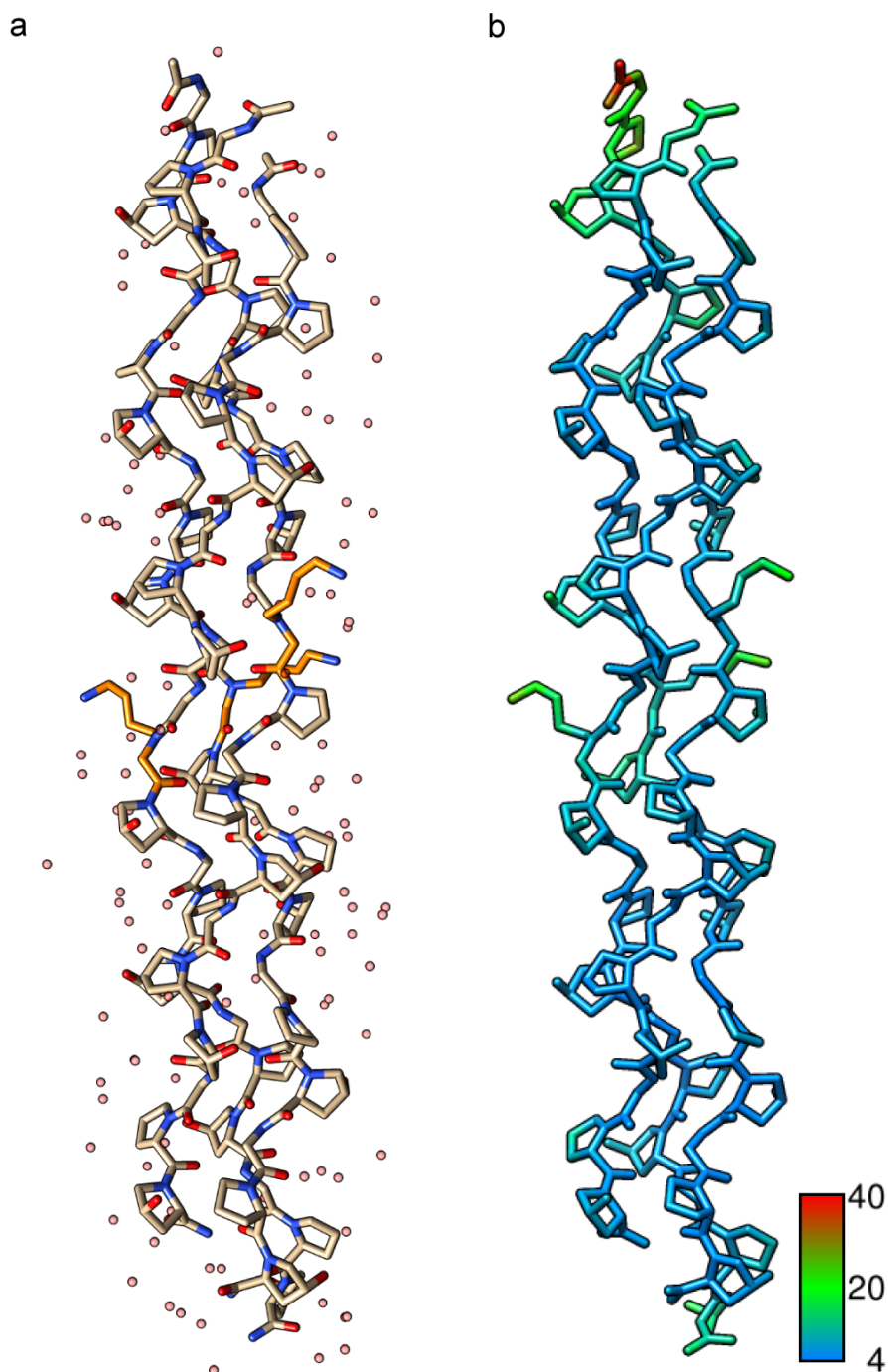
Figure S1. Crystal structure of Nlys-CMP. **a**, Overall triple-helical structure, showing one trimer in the asymmetric unit, including complete hydration shell, determined at 0.95 Å resolution. Central Nlys residues are colored in orange. **b**, Illustration of B-factors in the crystal structure: Overall 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.
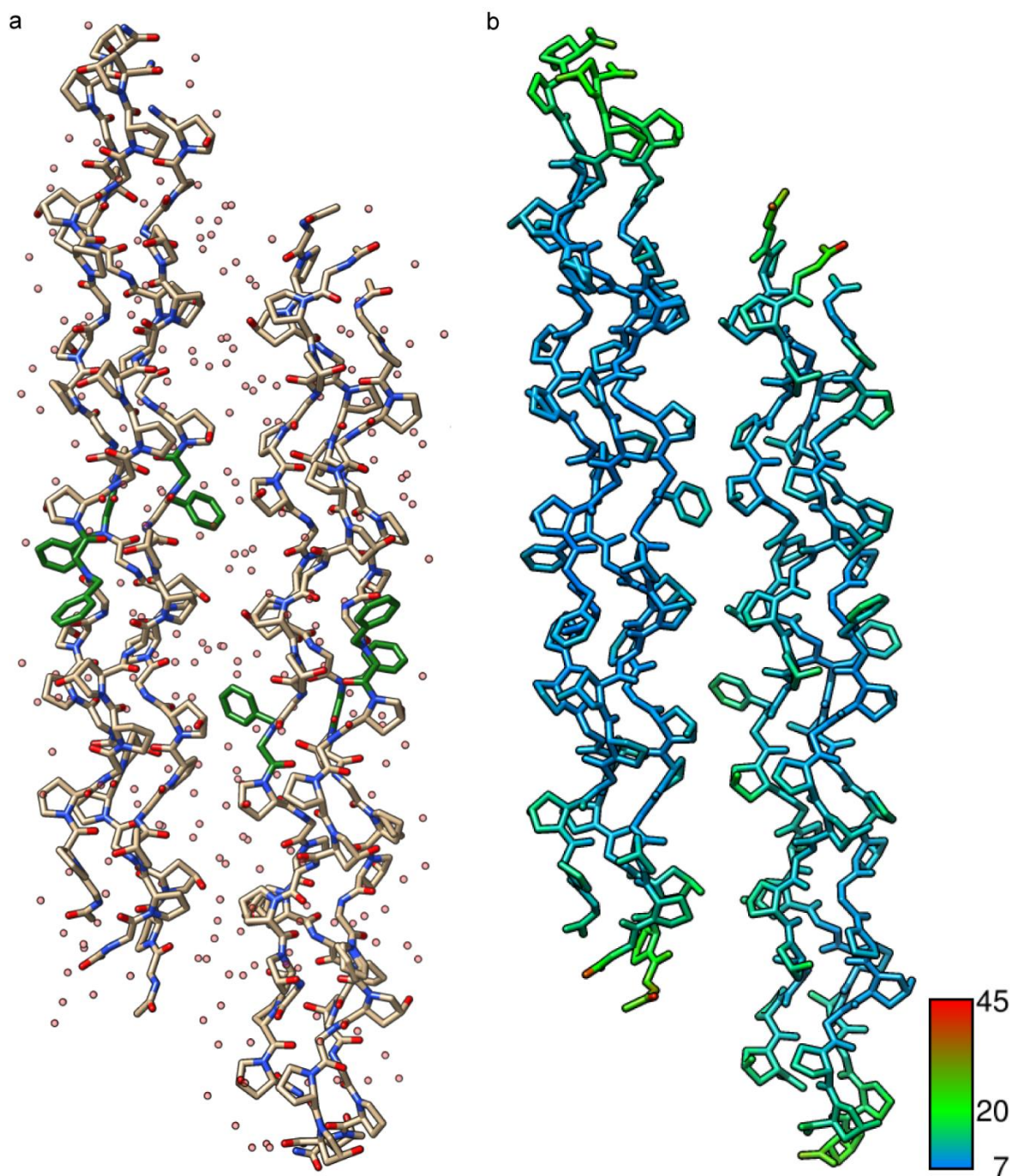
Figure S2. Crystal structure of Nphe-CMP. **a**, Overall triple-helical structure, showing two trimers in the asymmetric unit, including the complete hydration shell, determined at 1.10 Å resolution. Central Nphe residues are colored in green. **b**, Illustration of B-factors in the crystal structure: 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.
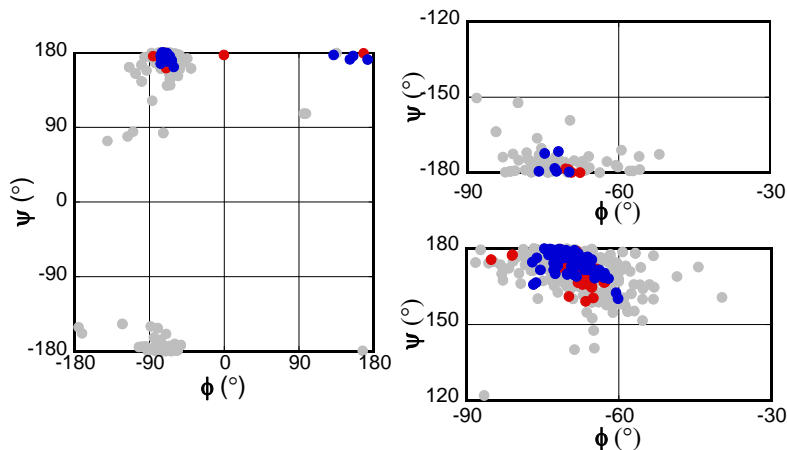
Figure S3a. Distribution of glycine φ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs glycine residues in collagen structures from literature (grey). Average dihedral angles (φ, ψ) for all non-terminal glycine residues are almost identical to previously reported structures: (-69 ± 3°, 174 ± 5°). Terminal glycines were removed because they have very high B-factors. The collagen structures used for comparison are listed in Table S4.
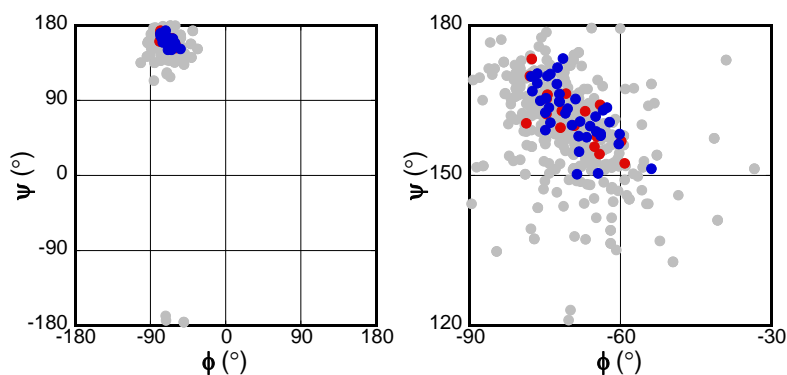
Figure S3b. Distribution of proline φ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs X position amino acids in collagen structures from literature (grey). Average dihedral angles for all proline residues are (-70 ± 6°, 162 ± 6°). The collagen structures used for comparison are listed in Table S4.
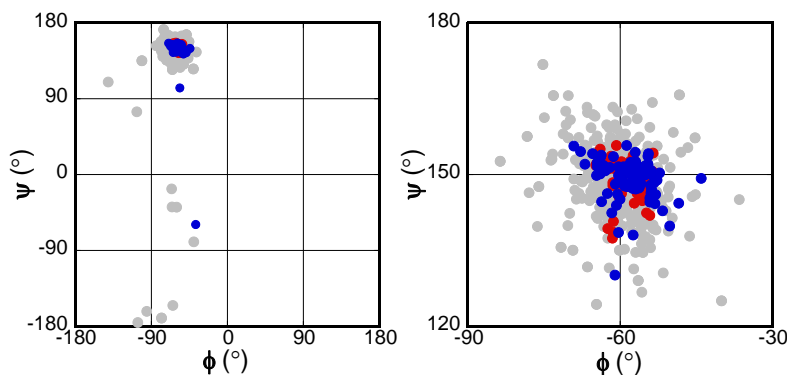
Figure S3c. Distribution of hydroxyproline φ and ψ angles in the solved crystal structures for Nlys- (red) and Nphe-CMPs (blue) vs Y position amino acids in collagen structures from literature (grey). Average dihedral angles for all hydroxyproline residues are (-58 ± 5°, 147 ± 7°). The collagen structures used for comparison are listed in Table S4.

a

Pro, 55 °C   NMe-Ala, 40 °C   Sar, 50 °C

b
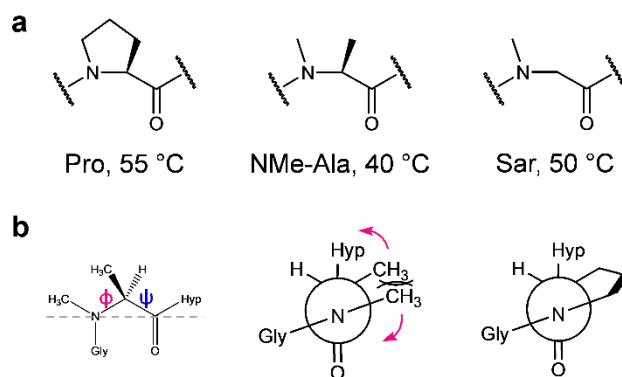
Figure S4. Guest residue *N*-methyl-L-alanine (NMe-Ala). **a**, Structures of Pro, NMe-Ala, and Sar, as well as the T_m values of their corresponding X-CMP triple-helices. **b**, Newman-like Projection of NMe-Ala showing the plausible steric clash of adjacent methyl groups in the polyproline backbone conformation.
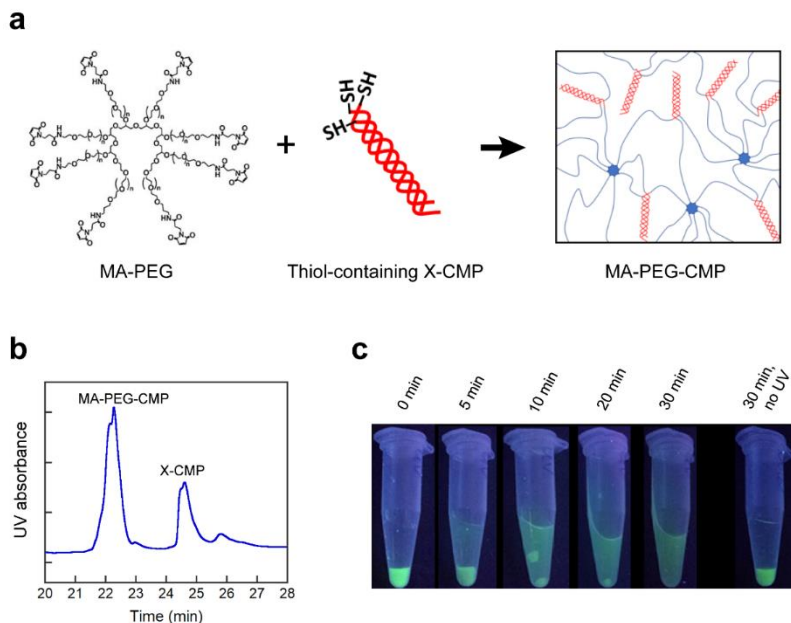
Figure S5. MA-PEG-CMPs produce a stable hydrogel capable of UV-induced dissolution. **a**, Schematic of maleimide groups at the termini of the 8-arm PEG reacting with the thiol-containing X-CMPs (Cys-Nnbz2-CMP) to form a triple-helix crosslinked hydrogel. **b**, HPLC of MA-PEG-CMP showing 70% of Cys-Nnbz2-CMP conjugated to the 8-arm PEG. **c**, MA-PEG-CMP mixed with fluorescein labeled F-Nnbz2-CMP formed 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 fluorescent peptide into solution.
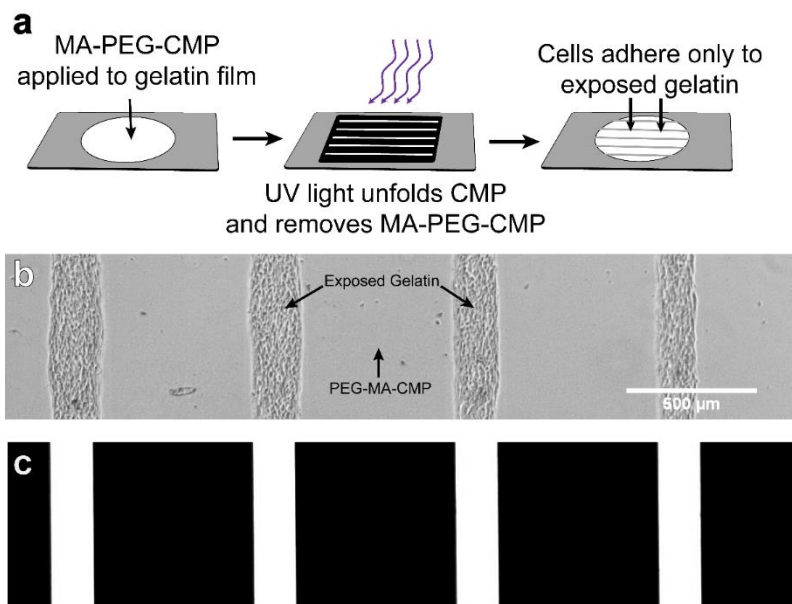
Figure S6. Spatial control of cell attachment using MA-PEG-CMP. **a**, Schematic of a gelatin substrate spatially modified with MA-PEG-CMP hydrogels through photo-patterning. Gelatin and collagen are excellent natural matrices for cell adhesion and migration; surface modification with a bulky, hydrophilic polymer can inhibit cell attachment. Nnbz2-CMP-conjugated MA-PEG-CMP hydrogel was added to the center of a gelatin film and allowed to bind. Next, we exposed this film to UV light through a photomask printed with a line pattern to remove the MA-PEG-CMP conjugate and expose the gelatin substrate in selected areas. Finally, MDA-MB-231 cells were seeded atop the MA-PEG-CMP modified gelatin and were found to adhere only to the exposed gelatin regions. **b**, Light micrograph showing cell attachment in a spatial pattern similar to the photomask used in this experiment (**c**).

Supplementary Section 4: HPLC, MALDI and CD of all peptides

X-CMP peptides

Pro-CMP
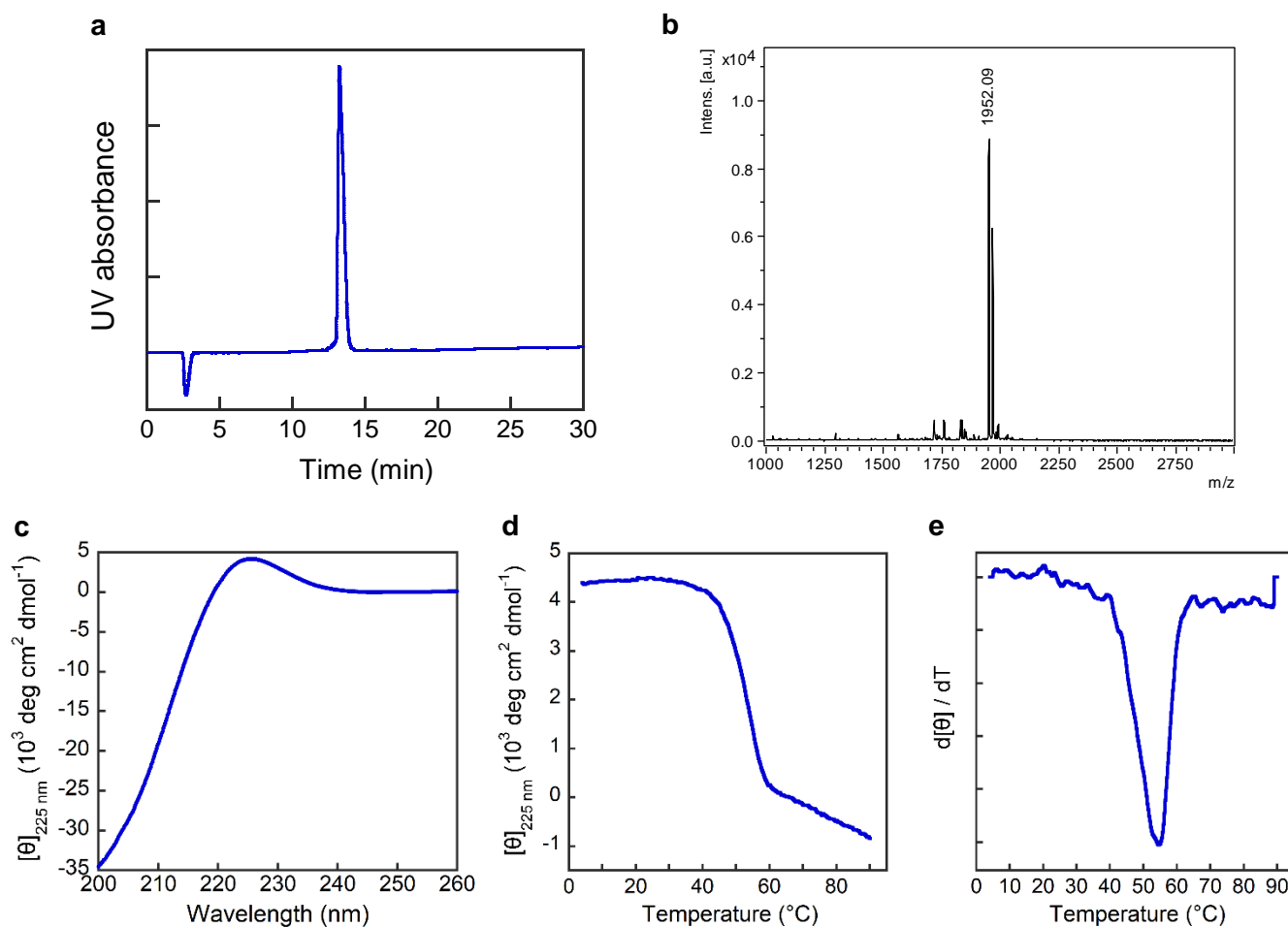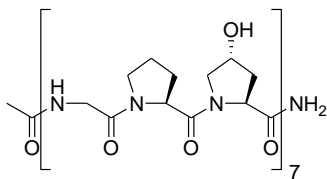
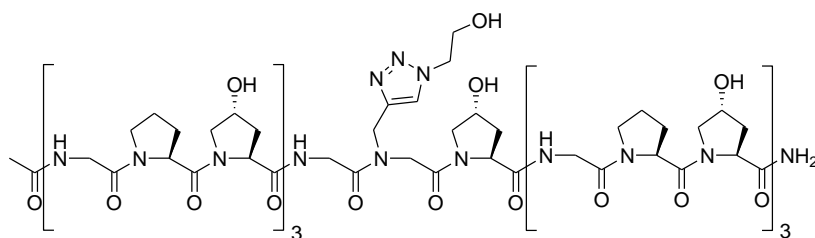**a**, The HPLC chromatogram of purified peptide, t_R = 13.2 min.

b, MALDI-MS, calculated: 1951.9 [M+Na]$^+$, observed: 1952.1 [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.

Pro-Ile-CMP
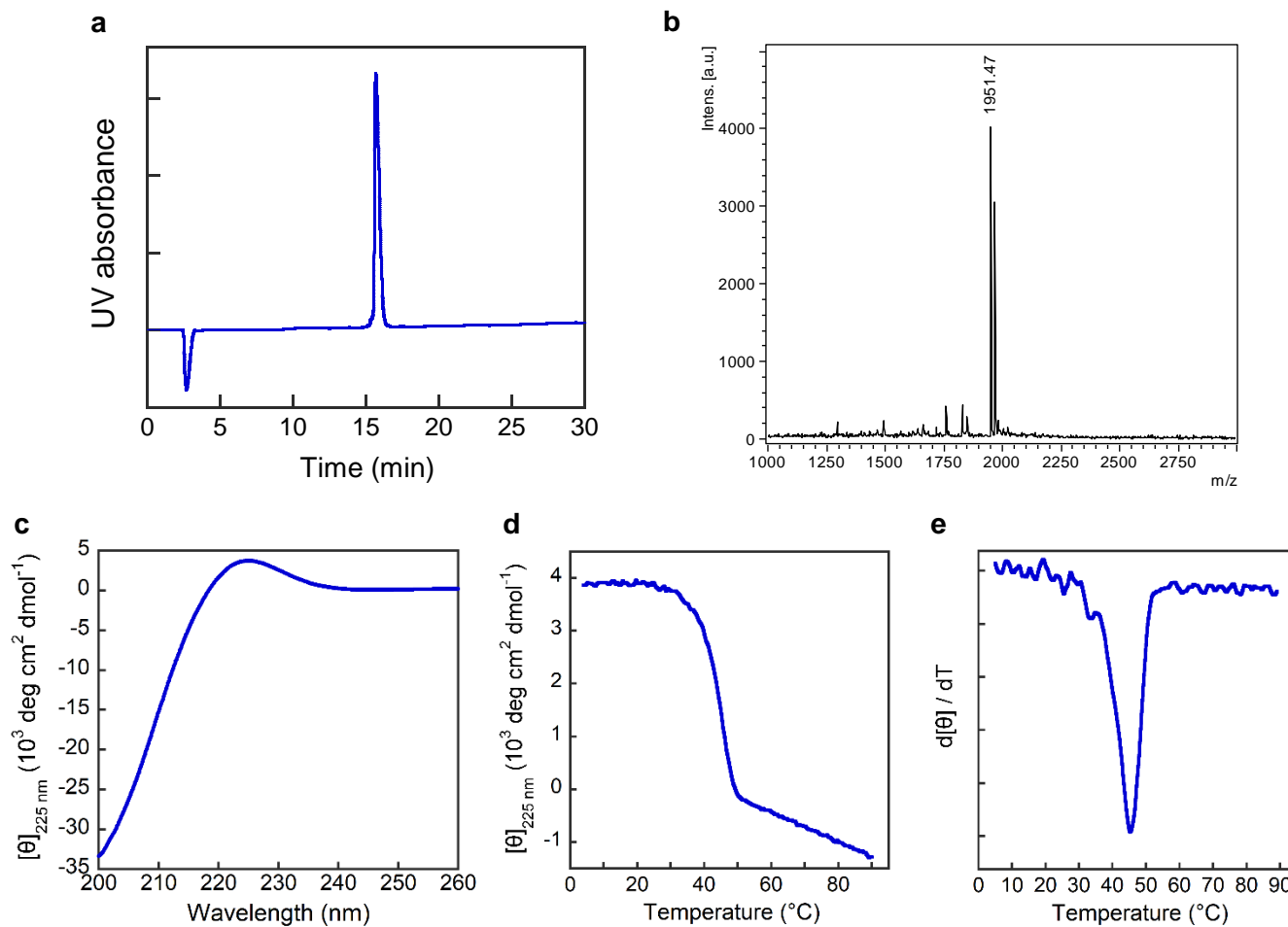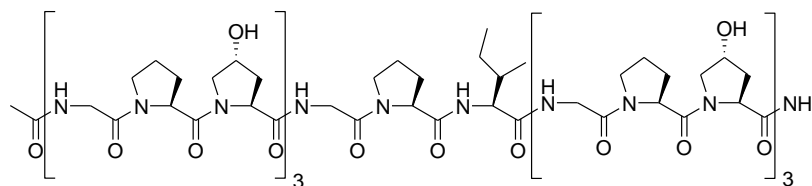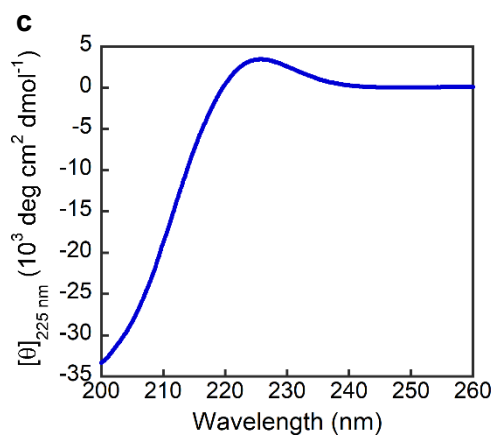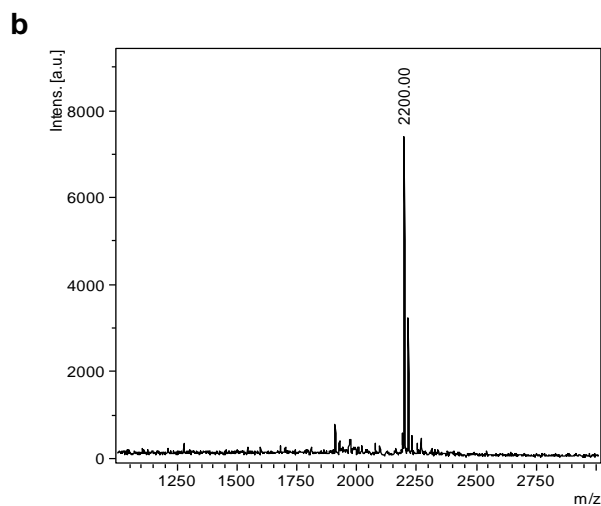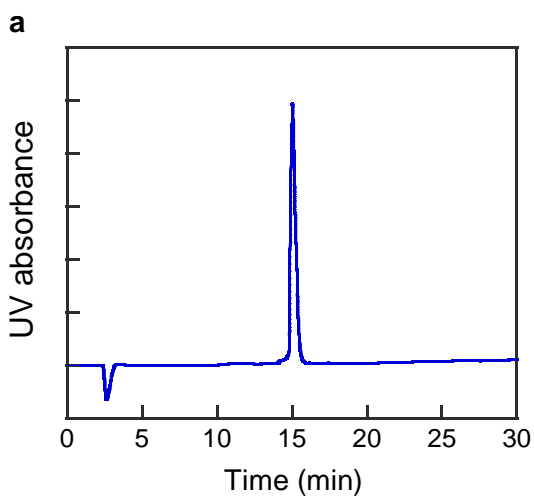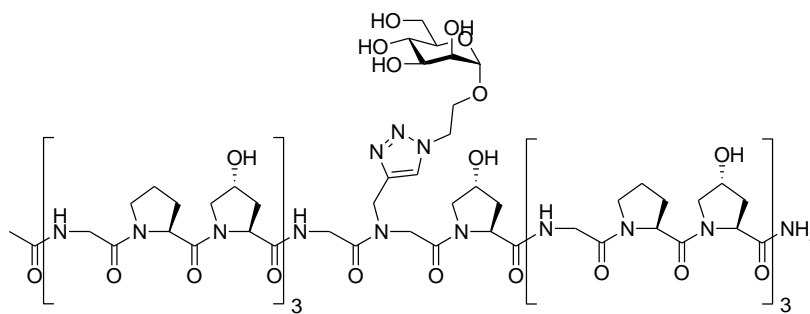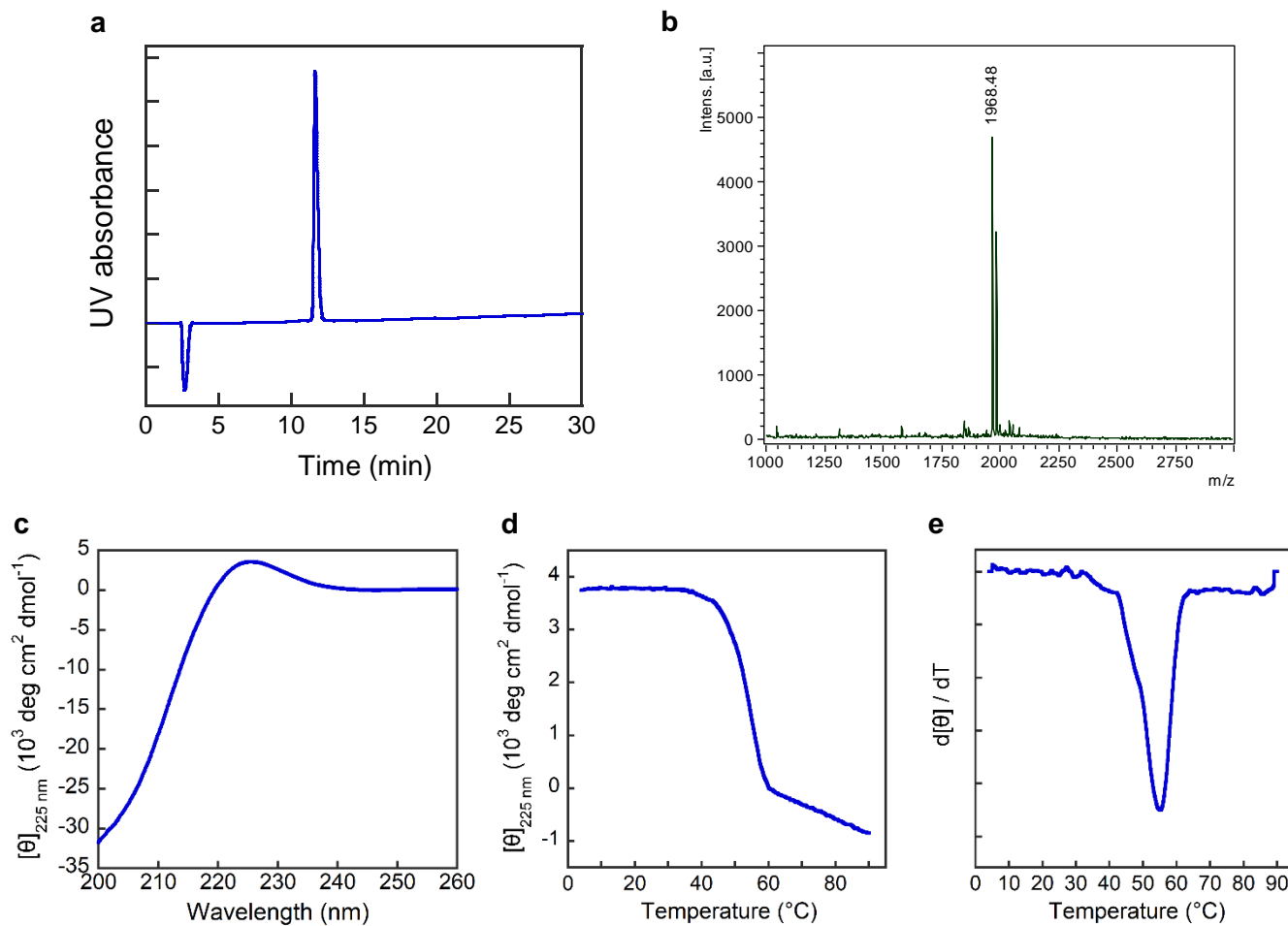

a, The HPLC chromatogram of purified peptide, t_R = 15.7 min.

b, MALDI-MS, calculated: 1951.9 [M+Na]$^+$, observed: 1951.5 [M+Na]$^+$, 1967.5 [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 = 45.5 °C.

Gly-CMP
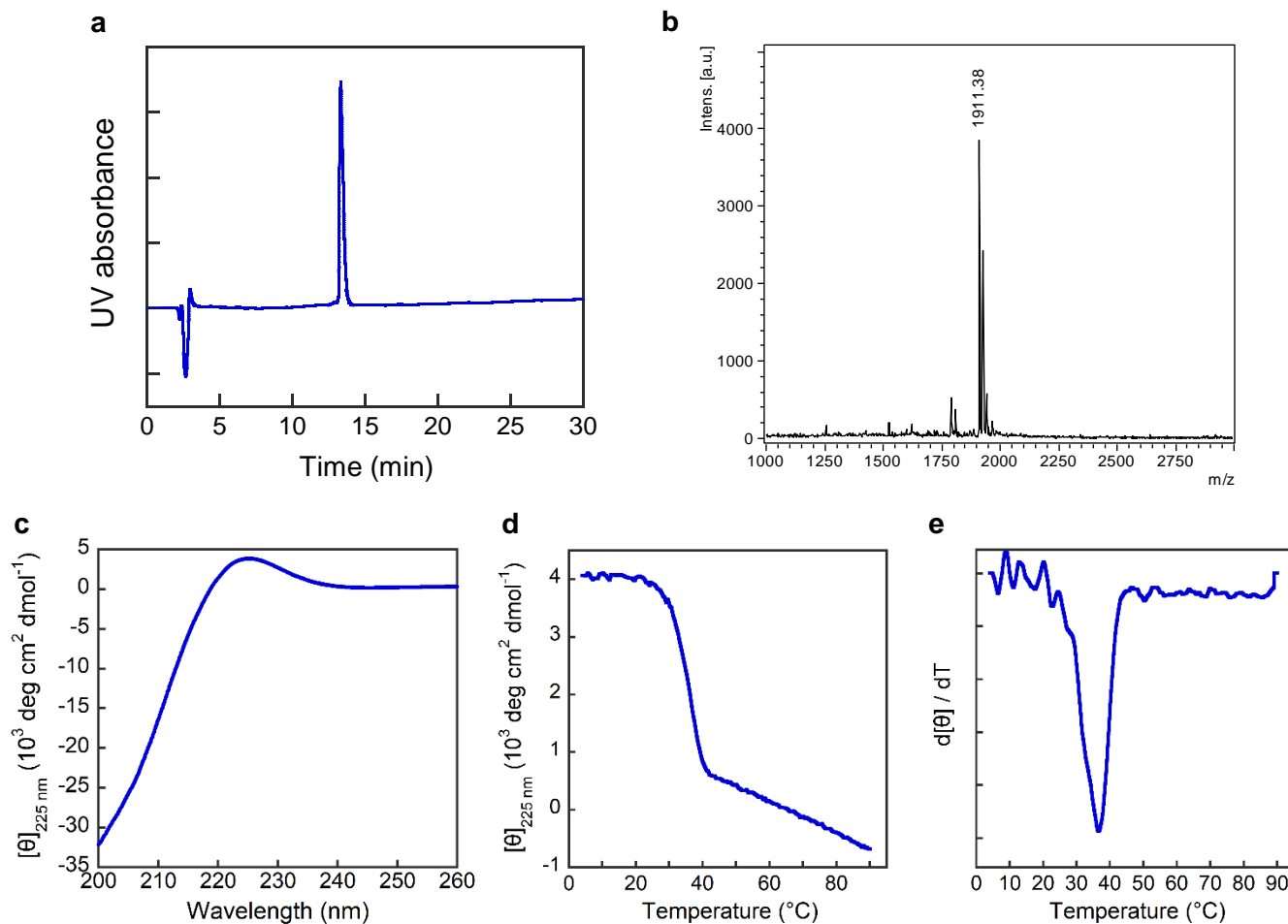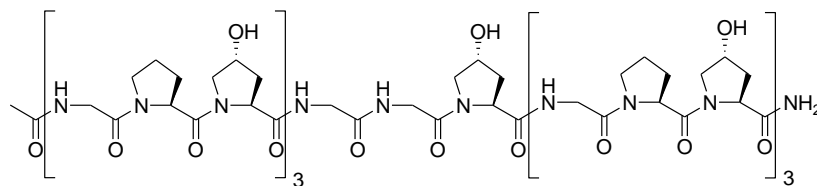
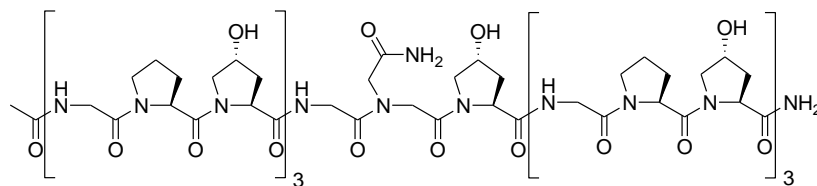**a**, The HPLC chromatogram of purified peptide, t_R = 13.3 min.

b, MALDI-MS, calculated: 1911.9 [M+Na]⁺, observed: 1911.4 [M+Na]⁺.
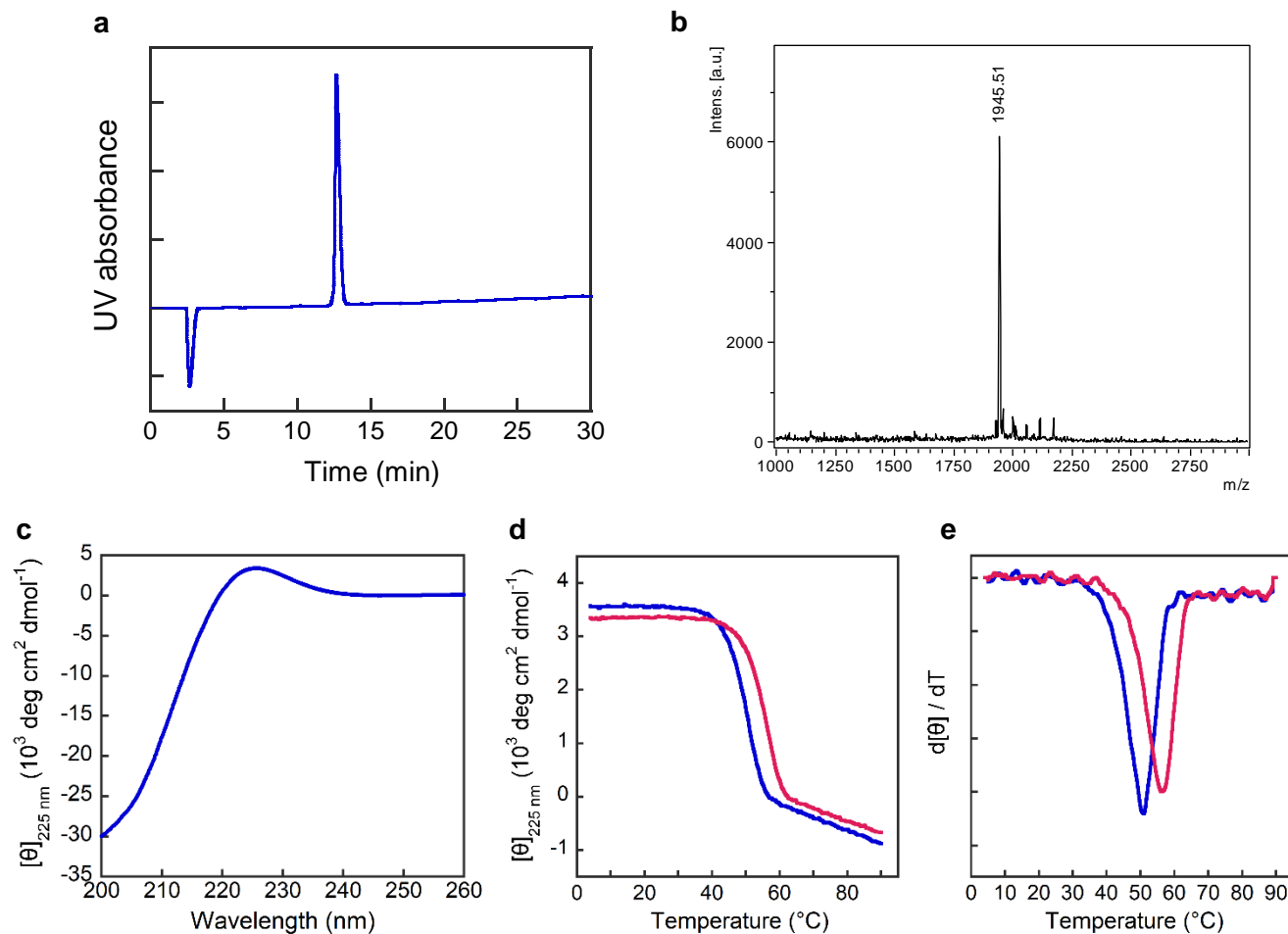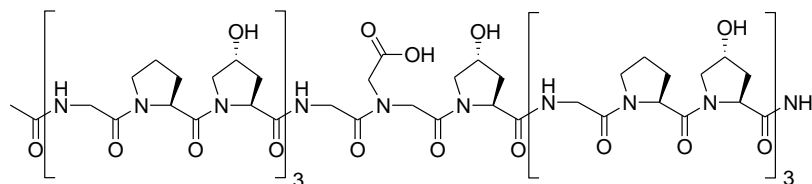
c, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

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

Lys-CMP

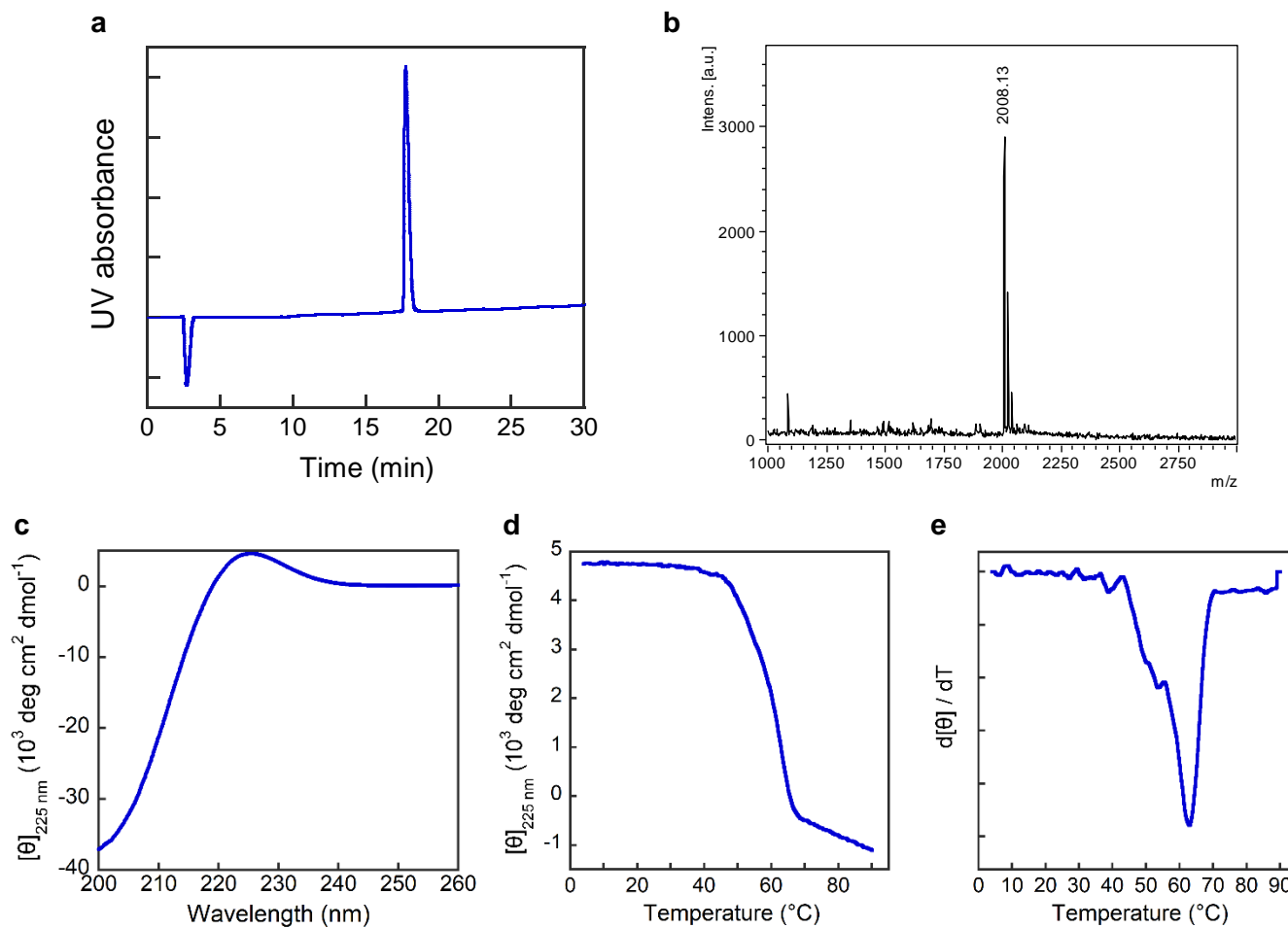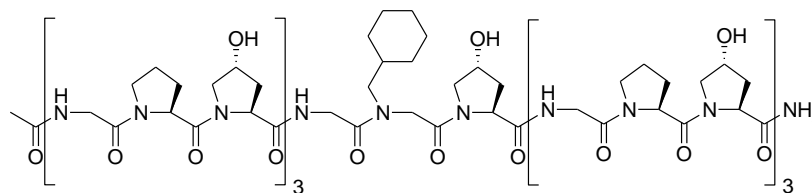**a**, The HPLC chromatogram of purified peptide, t_R = 12.5 min.

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).
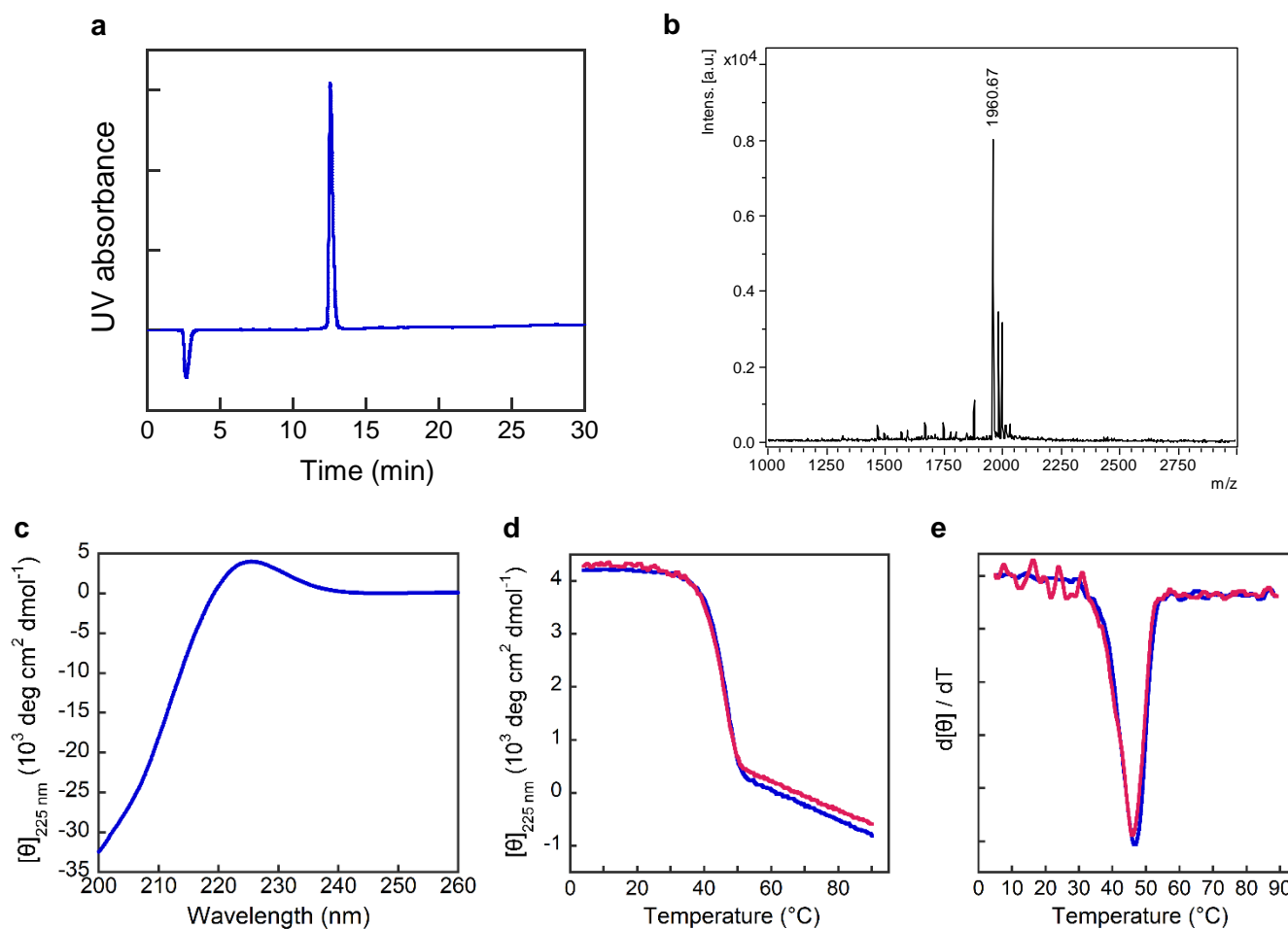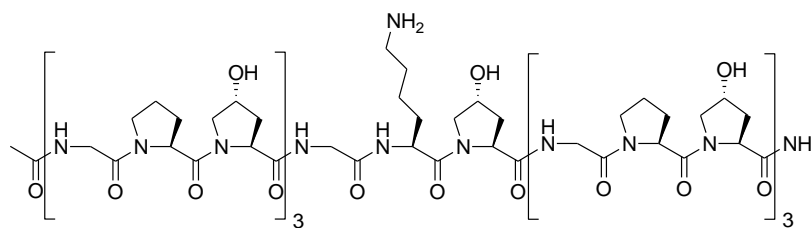
Phe-CMP
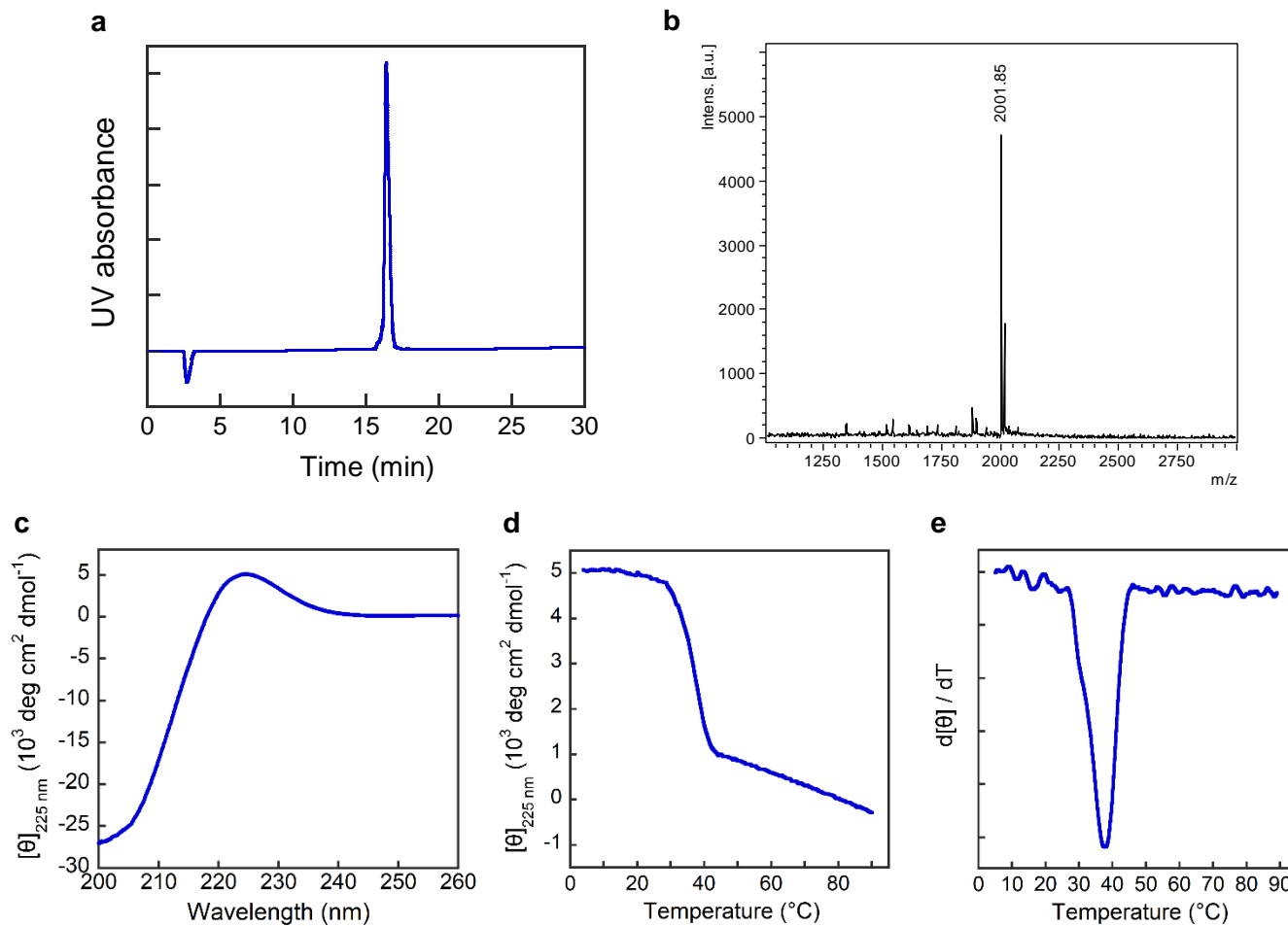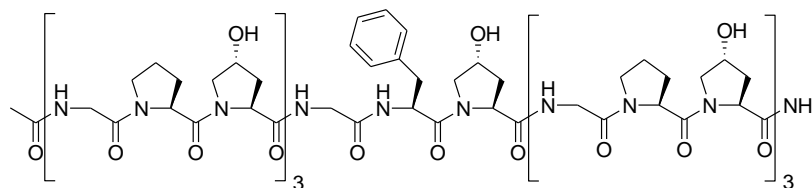
a, The HPLC chromatogram of purified peptide, t_R = 16.4 min.

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

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

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, T_m = 38 °C.

Sar-CMP

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

b, MALDI-MS, calculated: 1925.9 [M+Na]$^+$, observed: 1926.8 [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 = 50 °C.
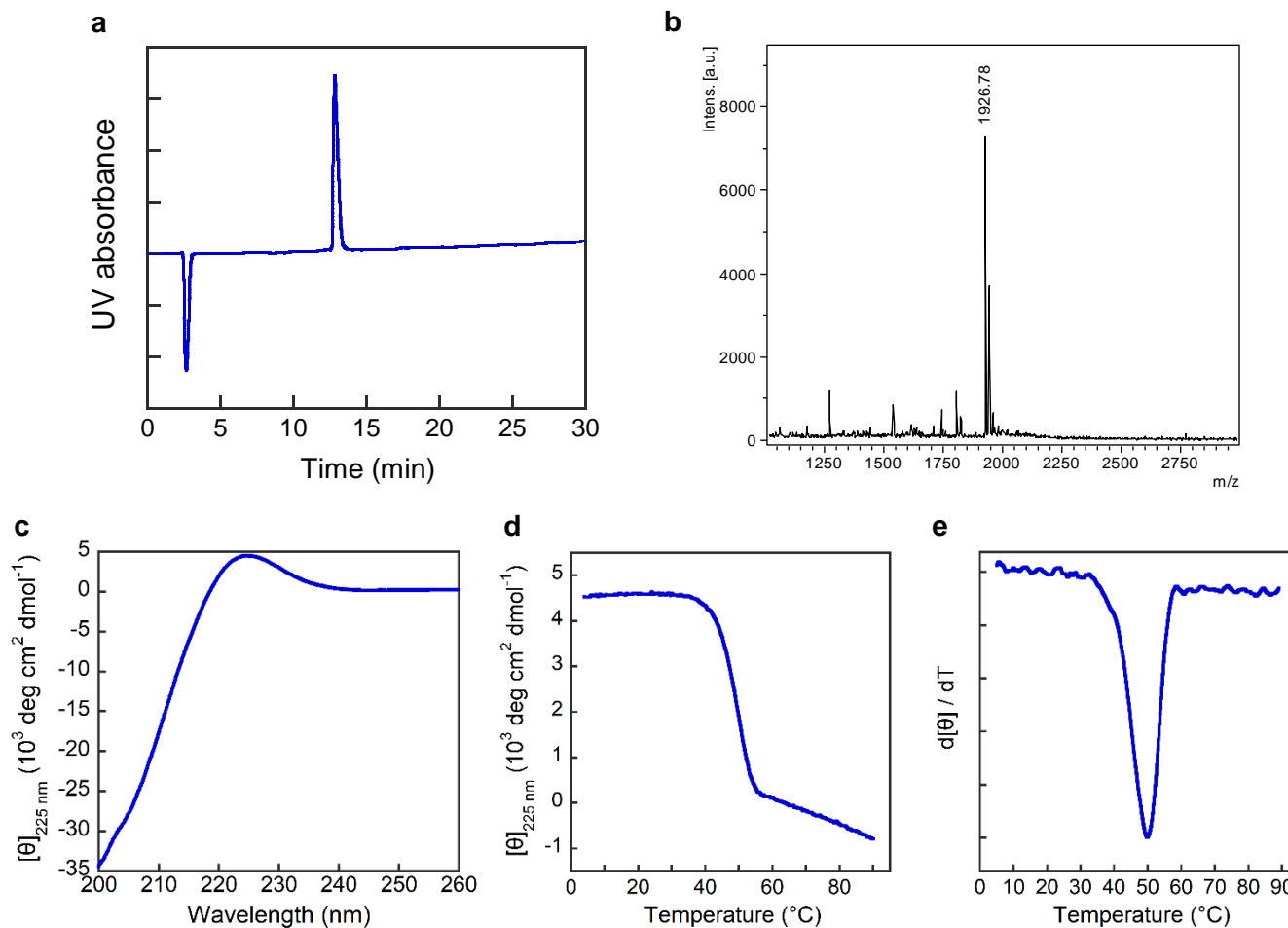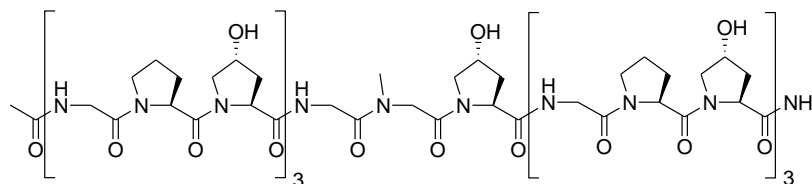
Nakn-CMP
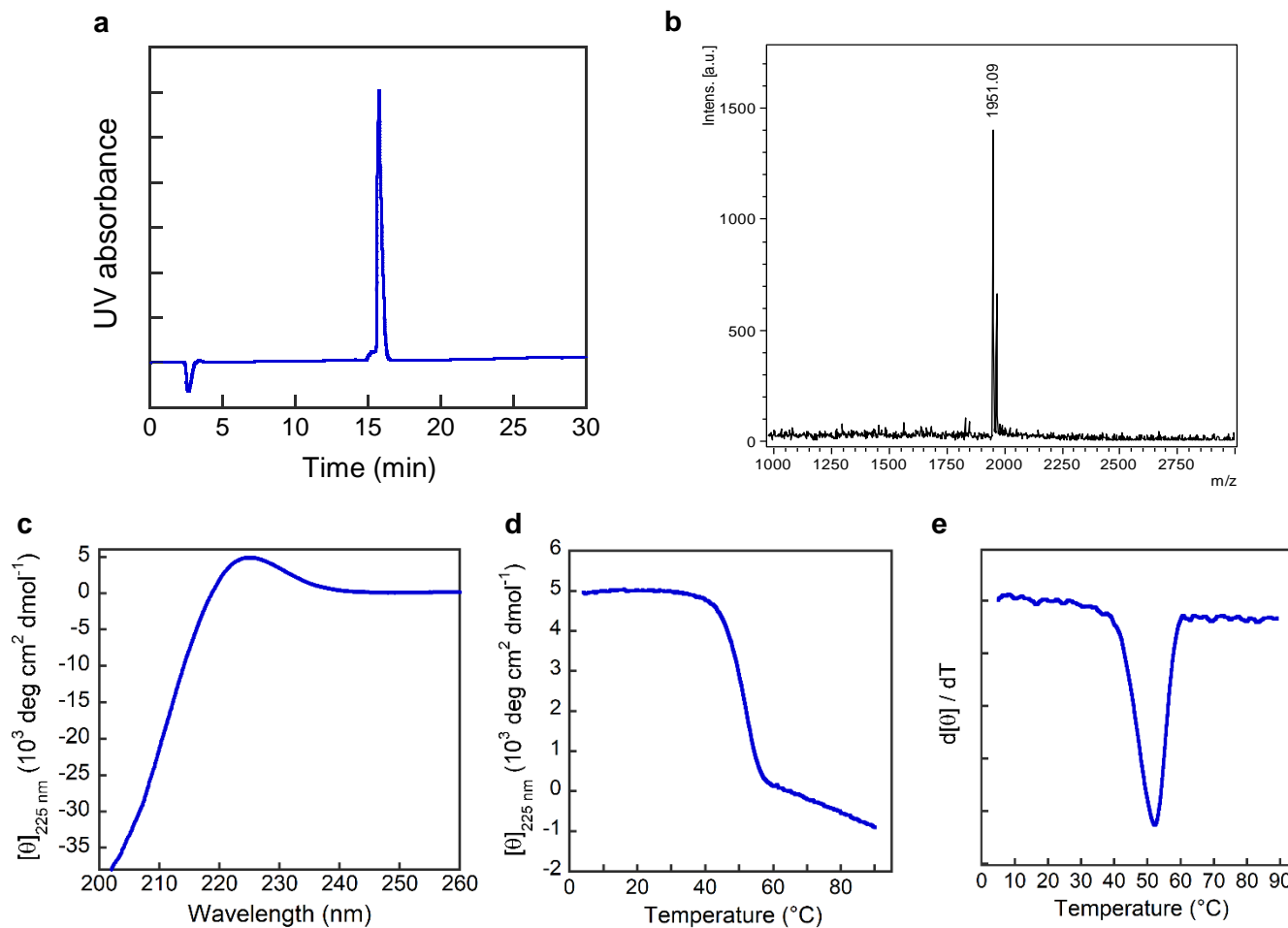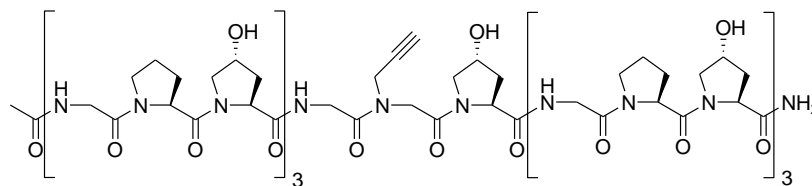
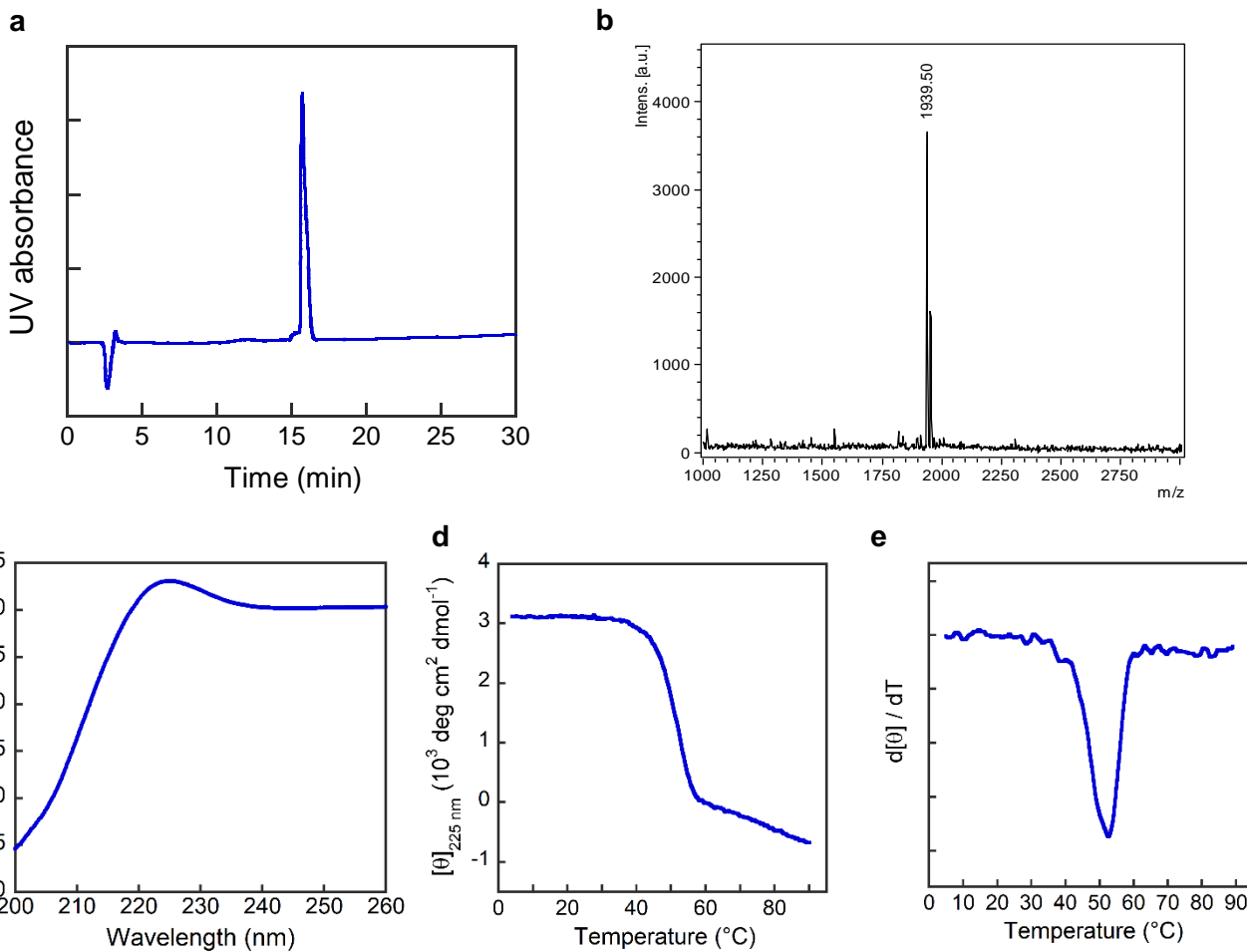**a**, The HPLC chromatogram of purified peptide, t_R = 15.7 min.

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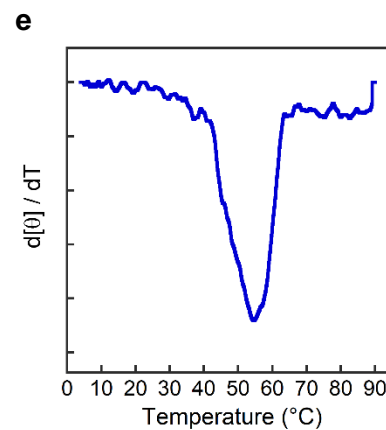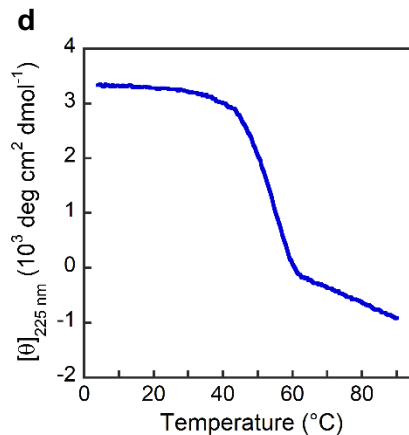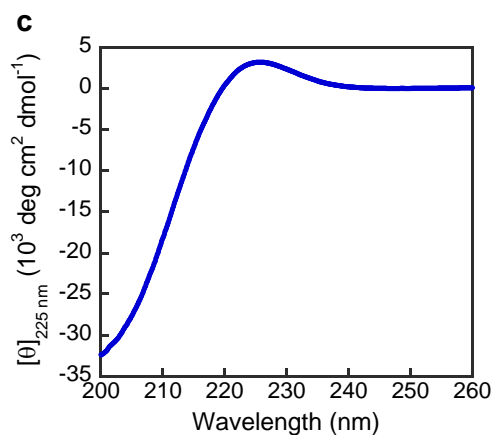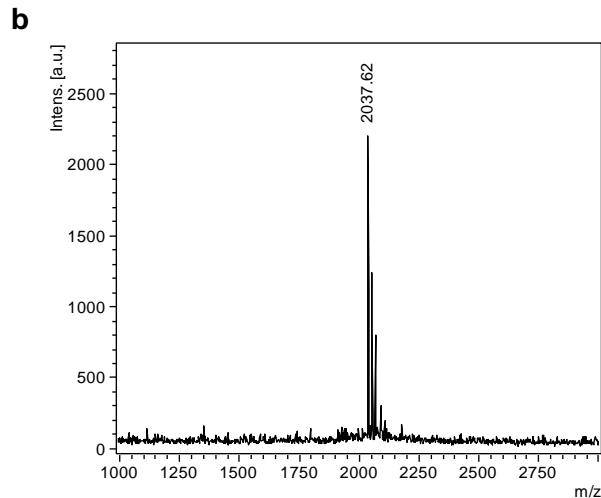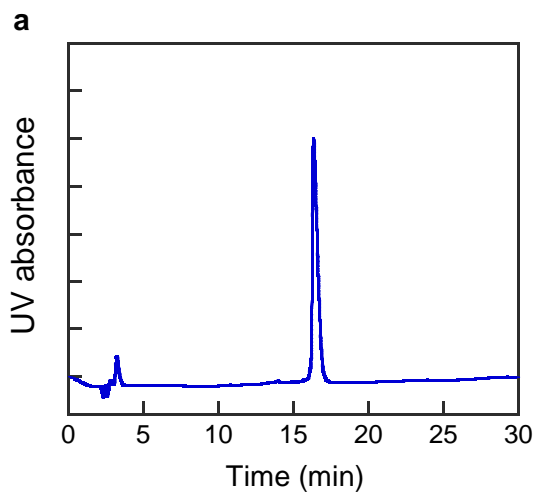
a, The HPLC chromatogram of purified peptide, t_R = 16.3 min.

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-man-CMP
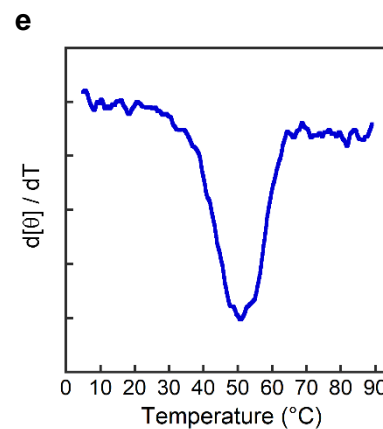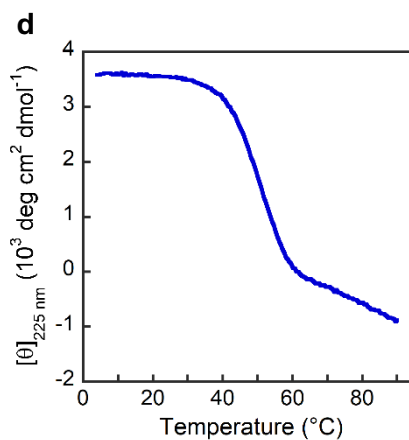
a, The HPLC chromatogram of purified peptide, $t_R = 15.0$ min.
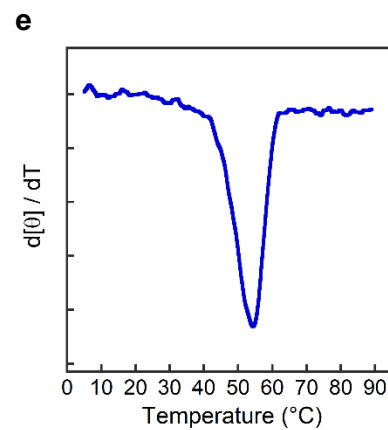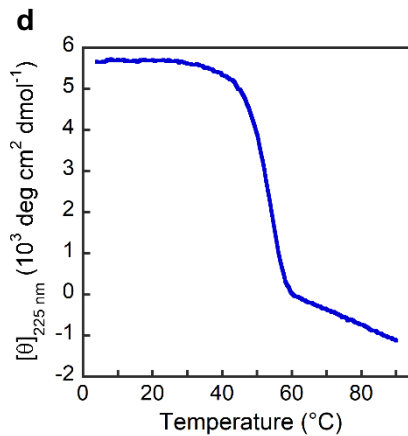
b, MALDI-MS, calculated: 2199.0 $[M+Na]^+$, observed: 2200.0 $[M+Na]^+$.

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, $T_m = 51$ °C.

Nasn-CMP

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

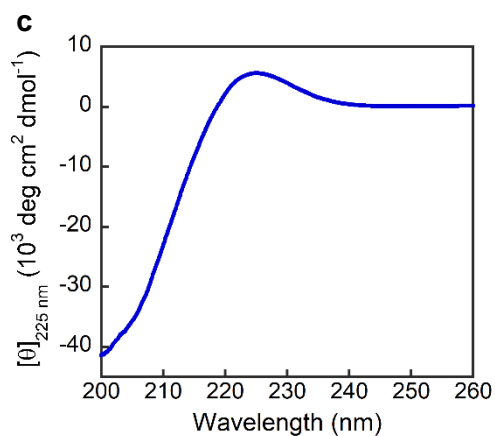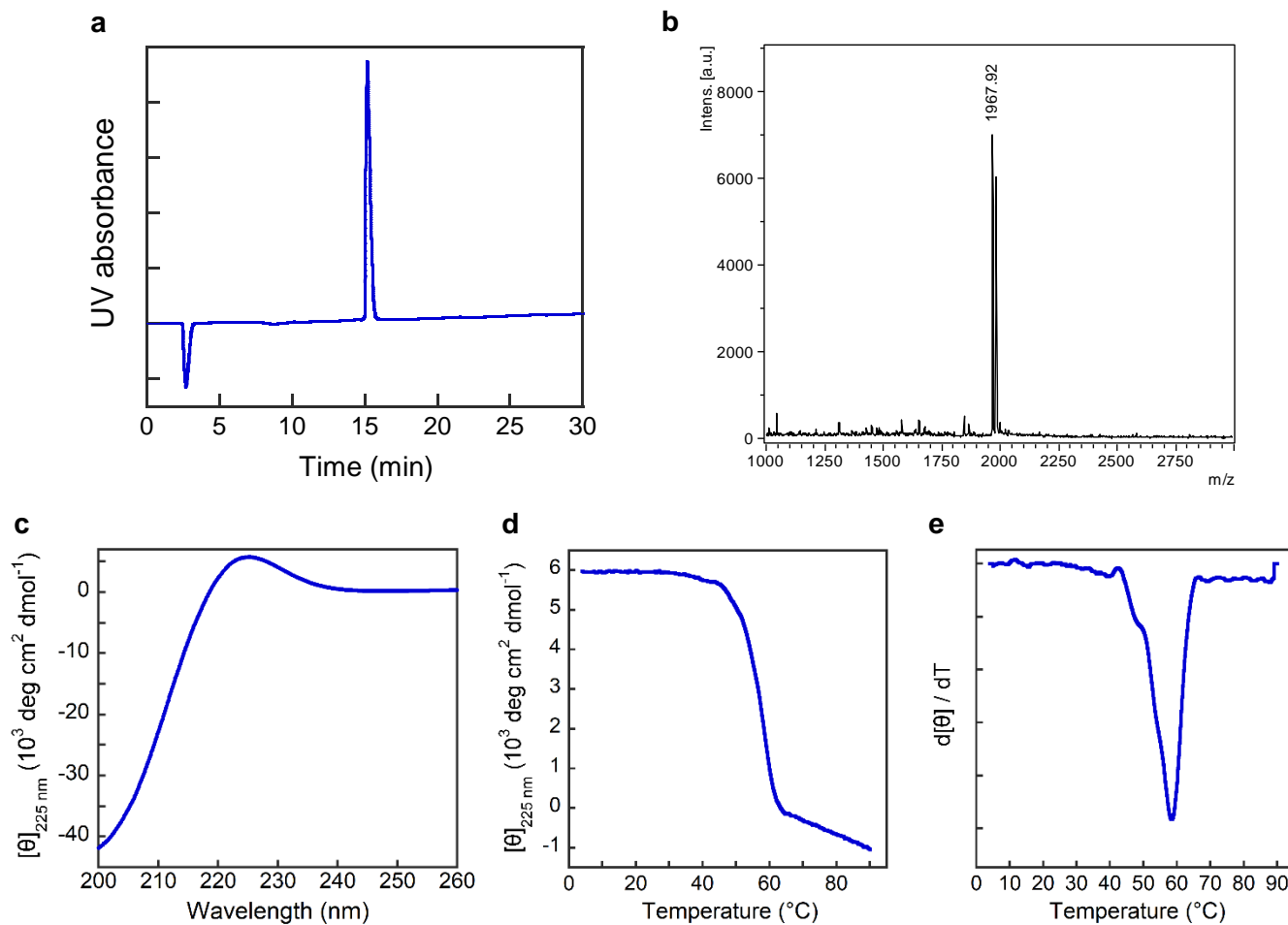**b**, MALDI-MS, calculated: 1968.9 [M+Na]$^+$, observed: 1968.5 [M+Na]$^+$.

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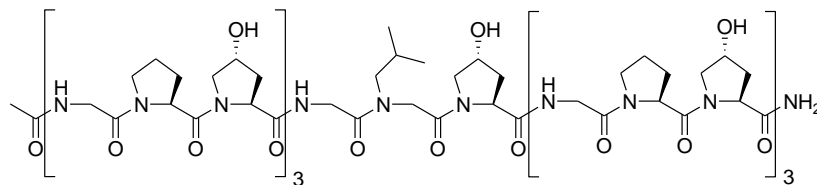
b, MALDI-MS, calculated: 1945.9 [M-H]⁻, observed: 1945.5 [M-H]⁻.

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 [M+Na]+, observed: 2008.1 [M+Na]+, 2024.1 [M+K]+.

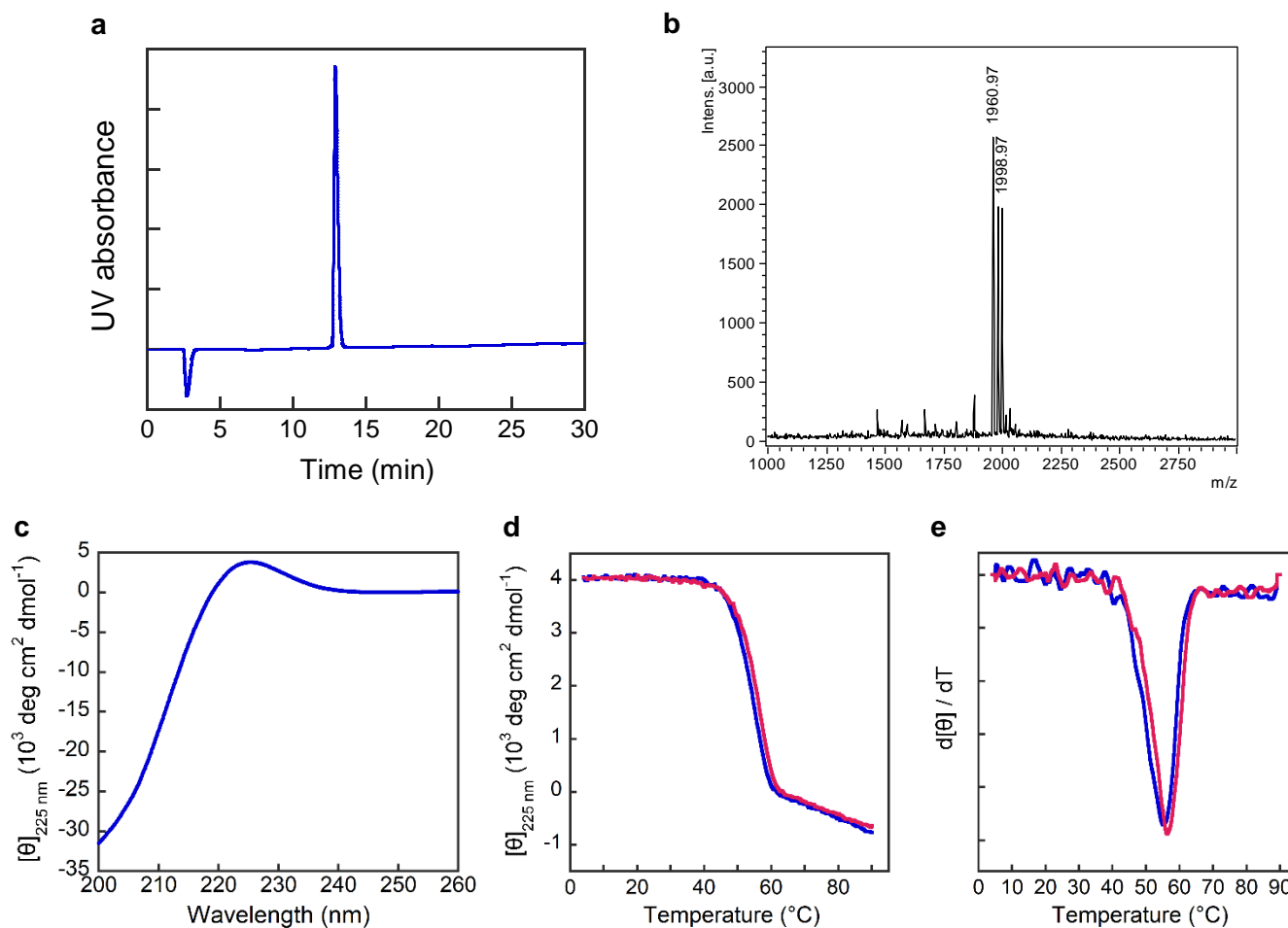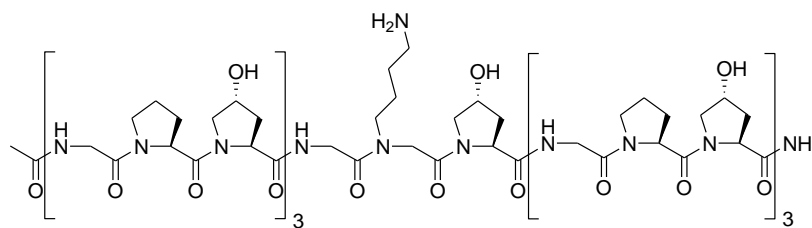**c**, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, T_m = 63 °C.

Ndxn-CMP
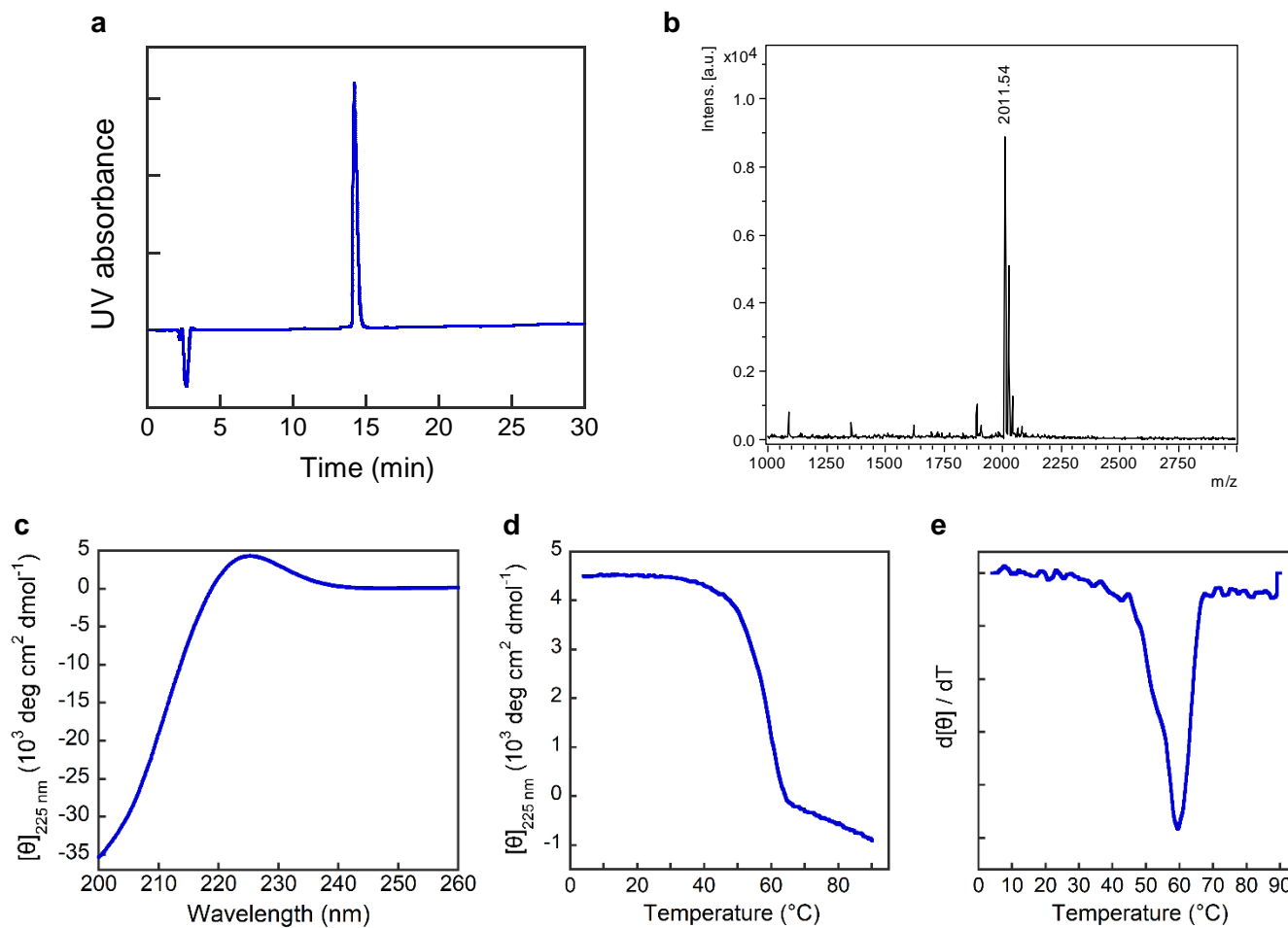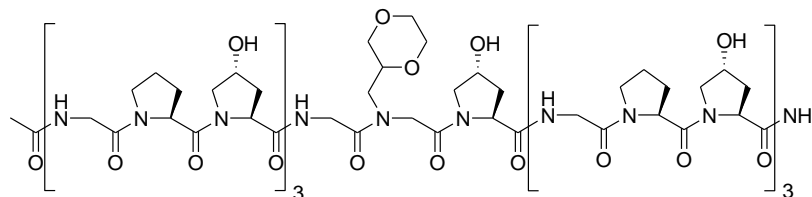
a, The HPLC chromatogram of purified peptide, t_R = 14.2 min.

b, MALDI-MS, calculated: 2011.9 [M+Na]$^+$, observed: 2011.5 [M+Na]$^+$, 2027.5 [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 = 59 °C.

NEt-CMP
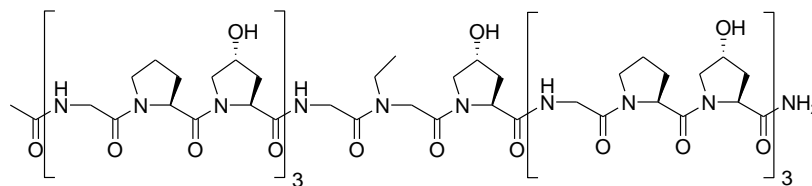
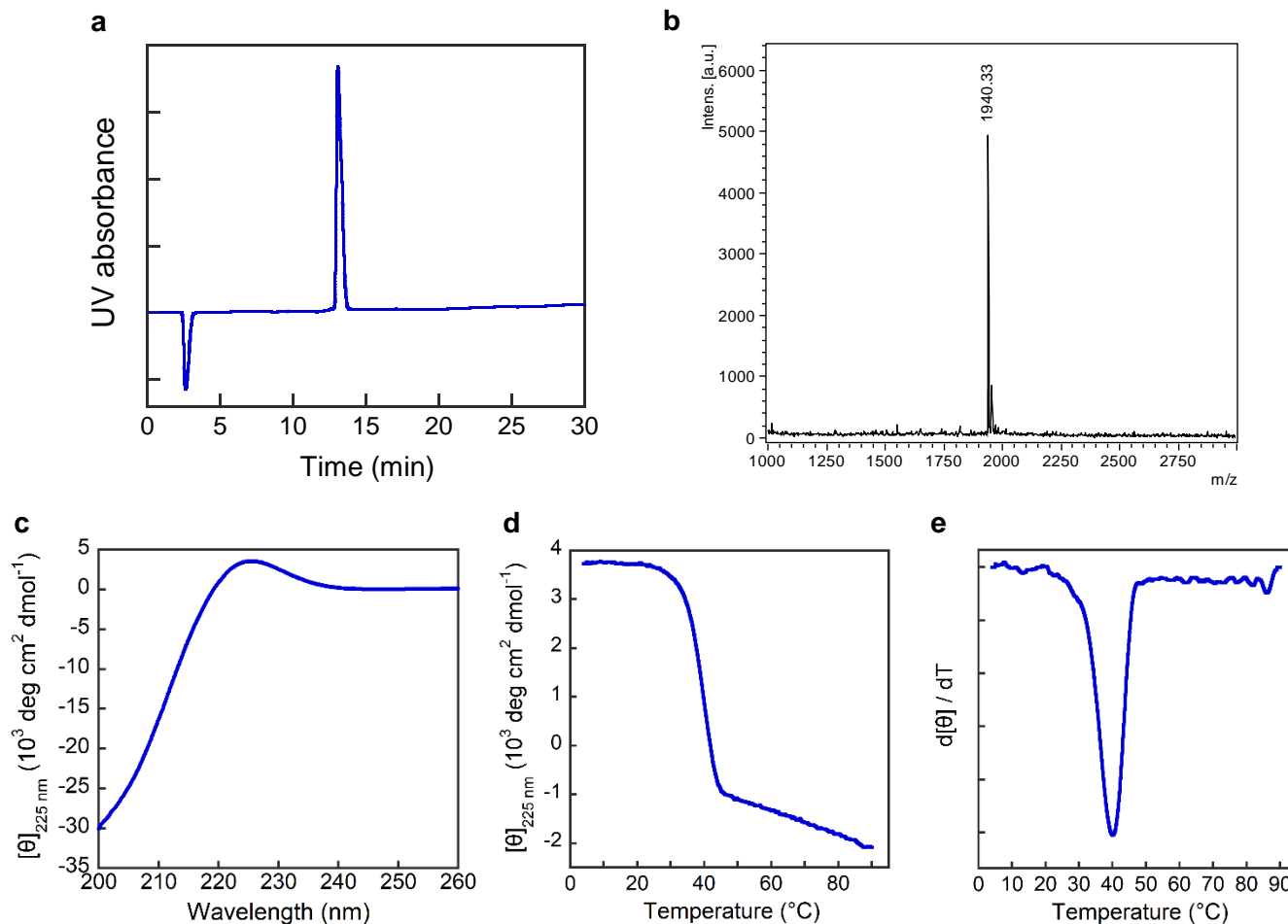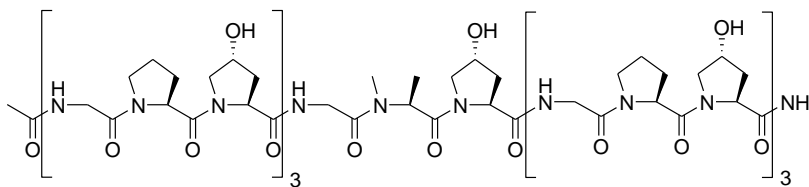**a**, The HPLC chromatogram of purified peptide, t_R = 15.7 min.

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
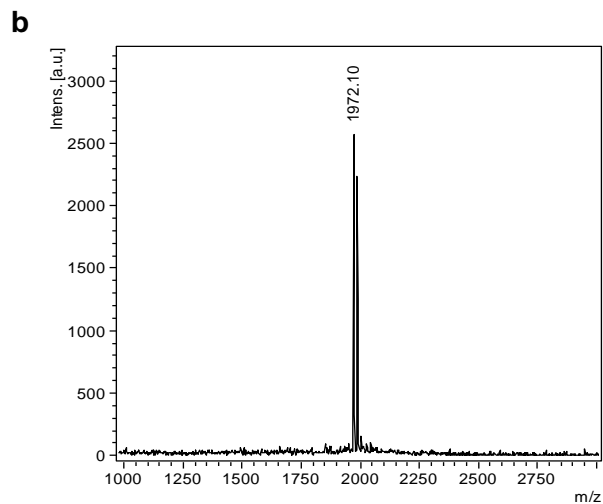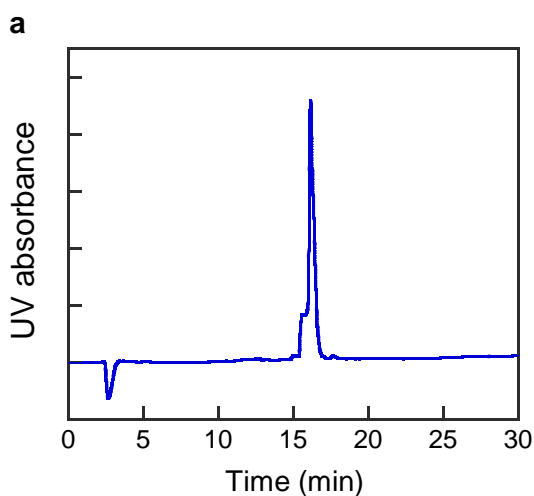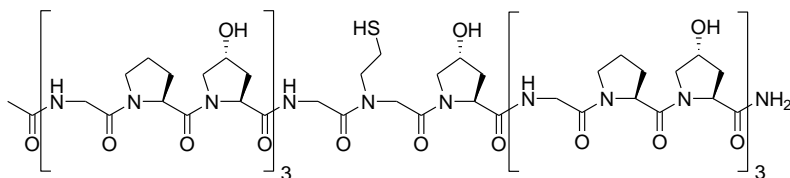
a, The HPLC chromatogram of purified peptide, t_R = 16.2 min.
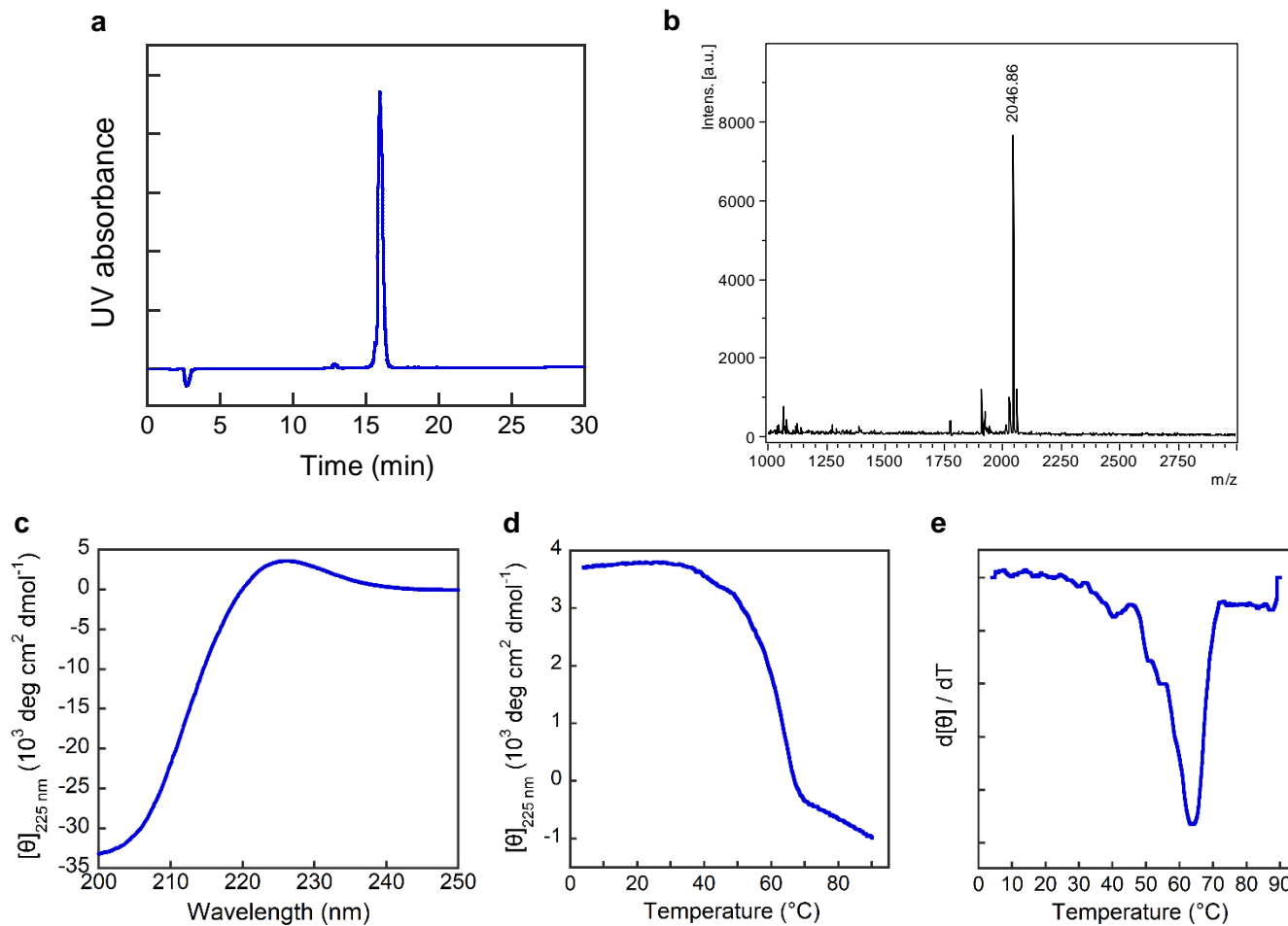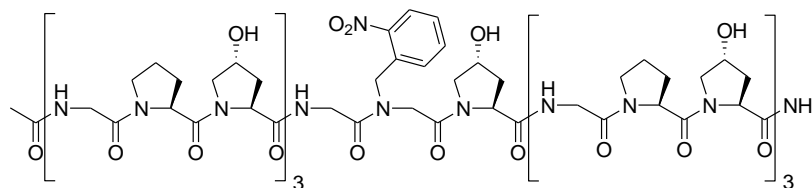
b, MALDI-MS, calculated: 1971.9 [M+Na]+, observed: 1972.1 [M+Na]+.

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

a, The HPLC chromatogram of purified peptide, t_R = 15.2 min.

b, MALDI-MS, calculated: 1967.9 [M+Na]$^+$, observed: 1967.9 [M+Na]$^+$, 1983.9 [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 = 58 °C.

Nlys-CMP

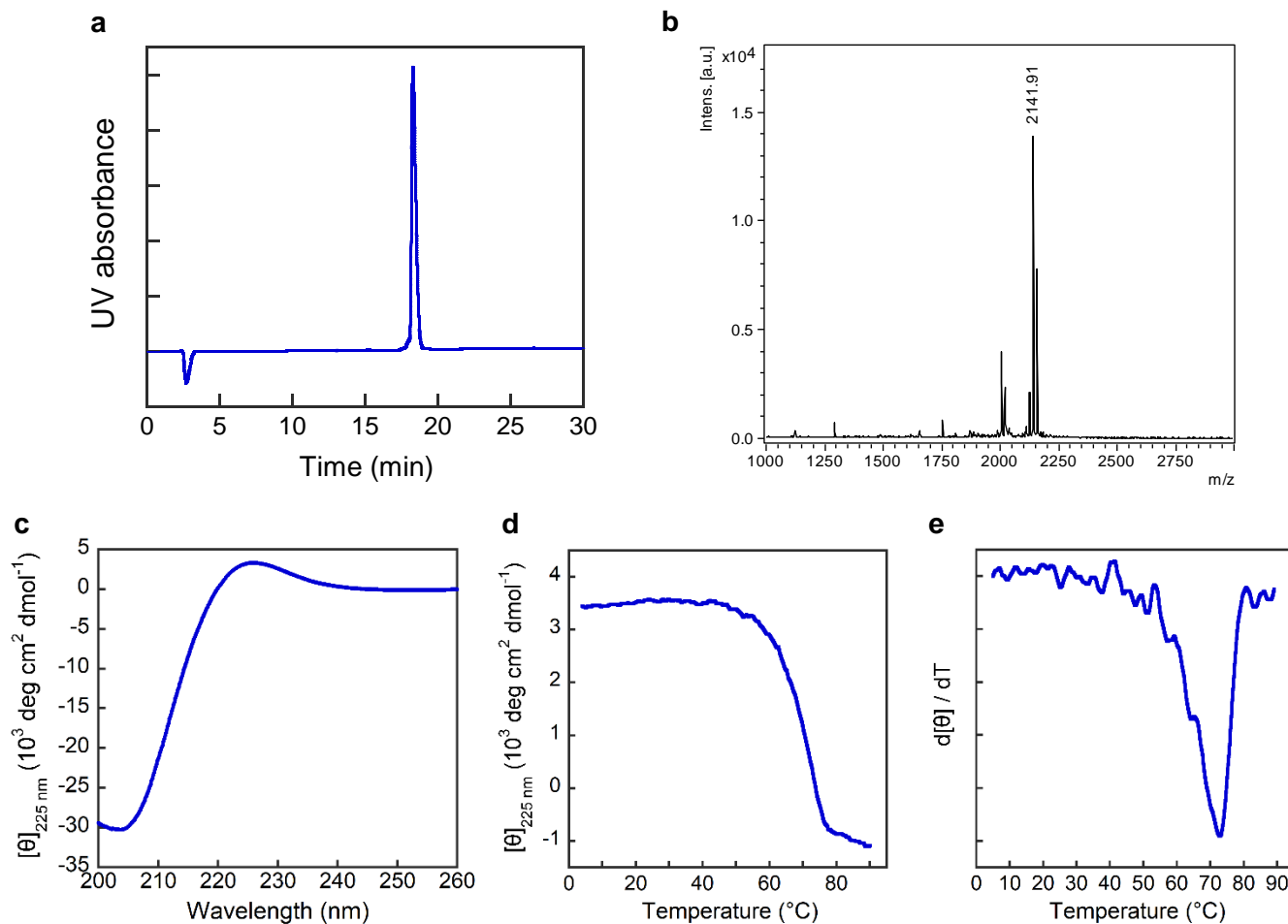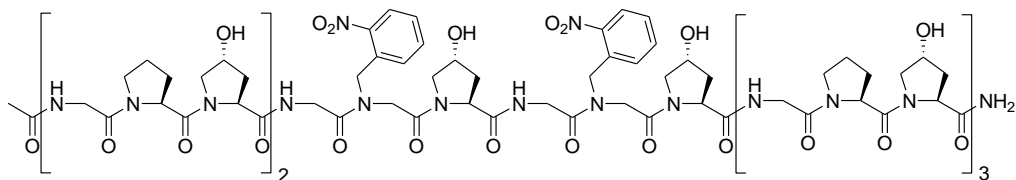**a**, The HPLC chromatogram of purified peptide, t_R = 12.9 min.

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 °C.

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

e, The first derivative of the melting curve, T_m = 40 °C.

Nmet-CMP
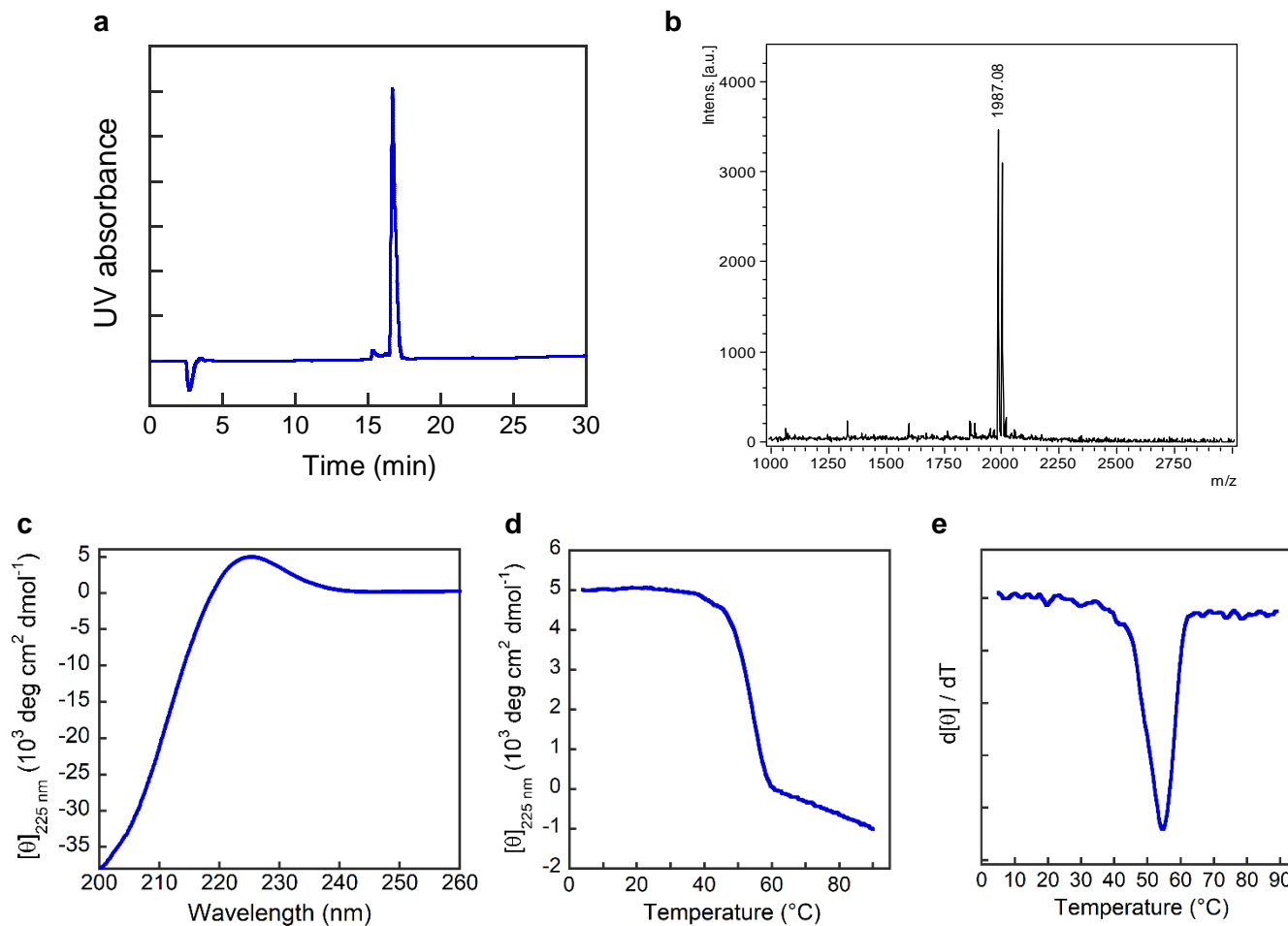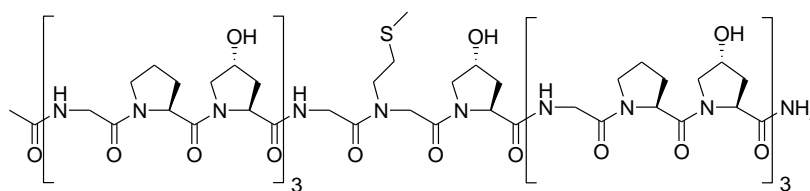
a, The HPLC chromatogram of purified peptide, t_R = 16.7 min.

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 °C.

d, The CD thermal melting curve in PBS buffer.

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

Nnbz2-CMP

a, The HPLC chromatogram of purified peptide, t_R = 18.3 min.

b, MALDI-MS, calculated: 2141.9 [M+Na]⁺, observed: 2141.9 [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 = 73 °C.

Nnbz3-CMP
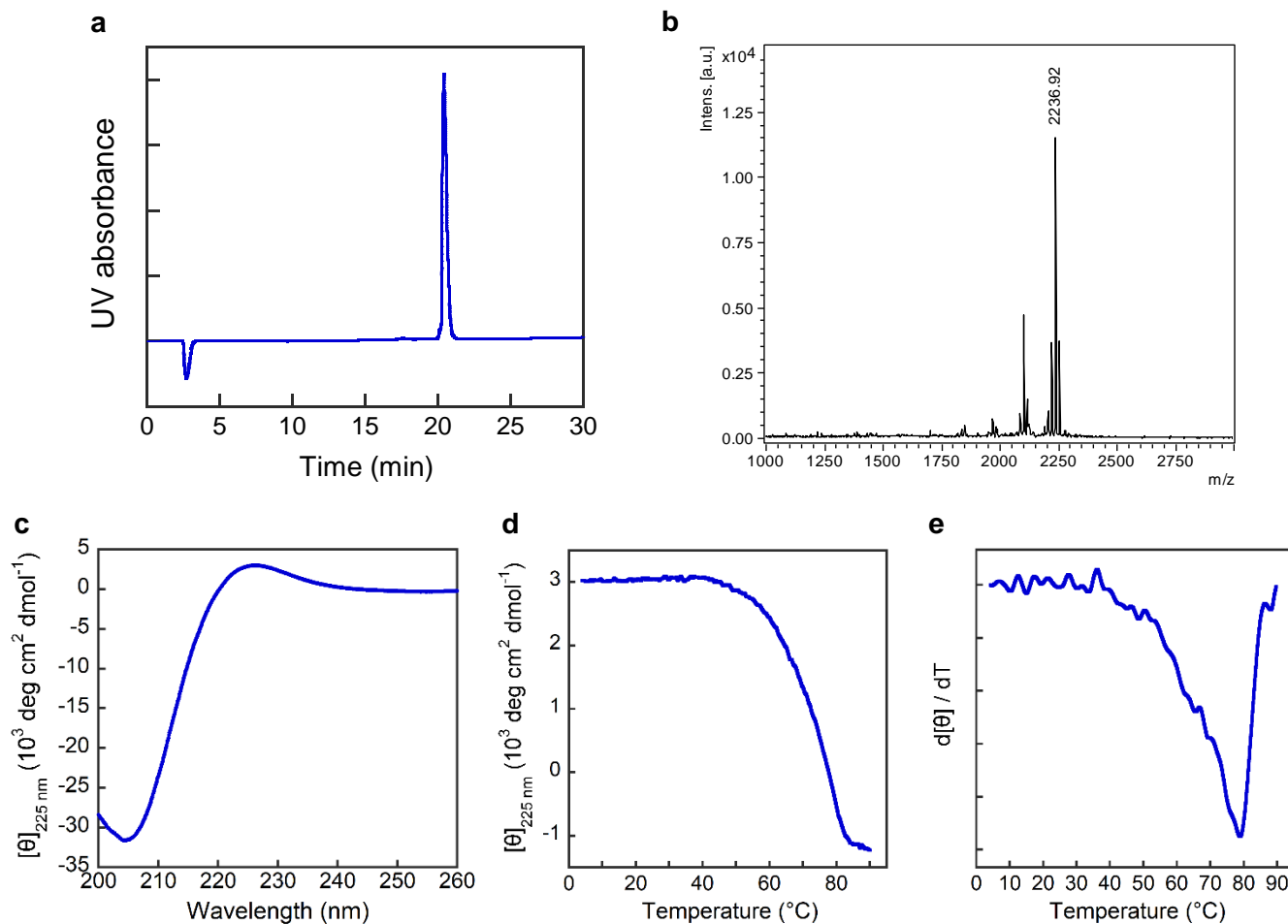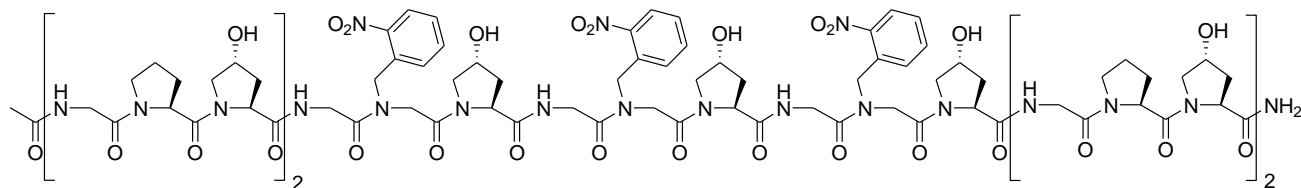
a, The HPLC chromatogram of purified peptide, t_R = 20.4 min.

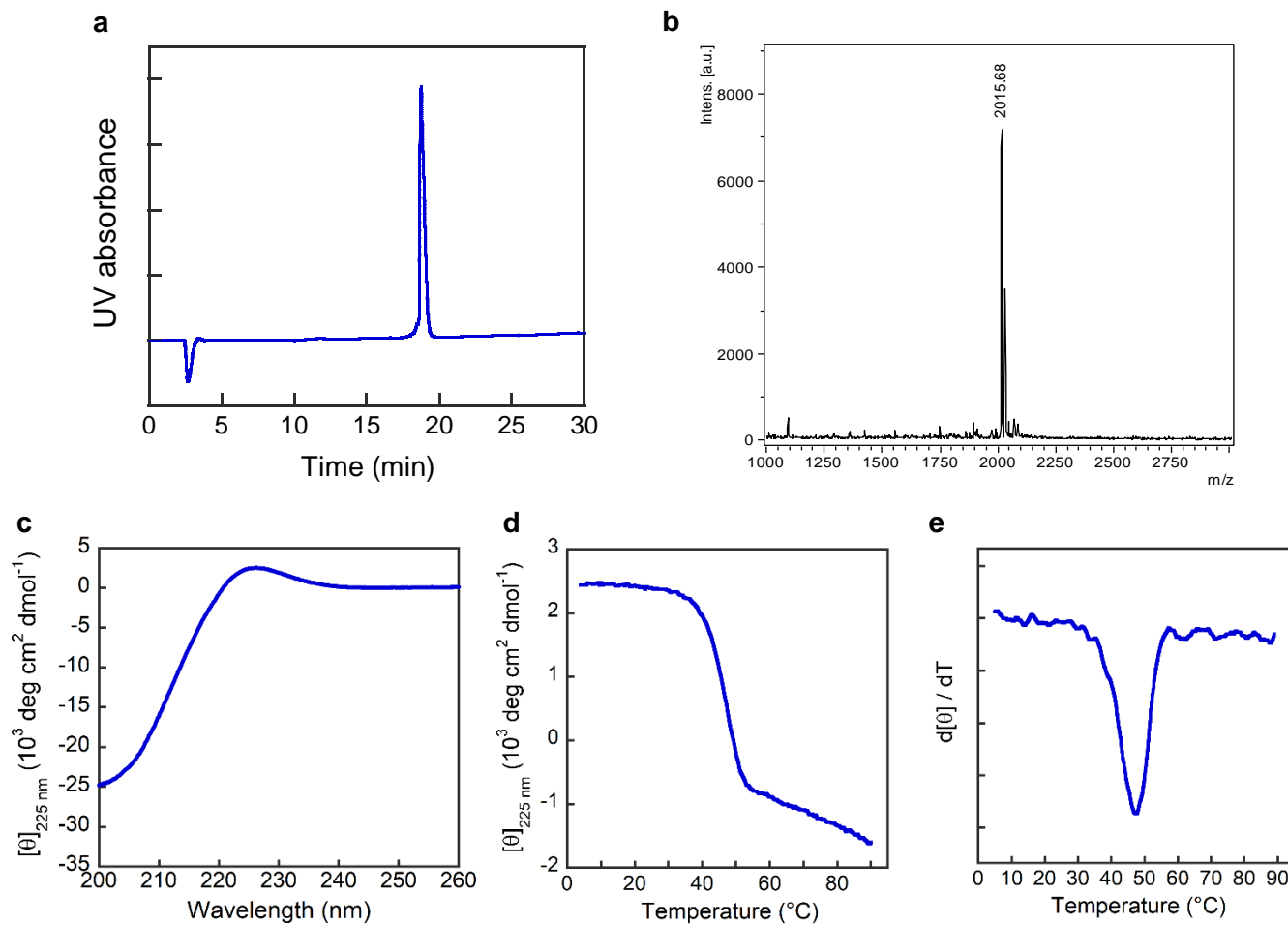**b**, MALDI-MS, calculated: 2236.9 [M+Na]$^+$, observed: 2236.9 [M+Na]$^+$.

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 = 79 °C.

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.
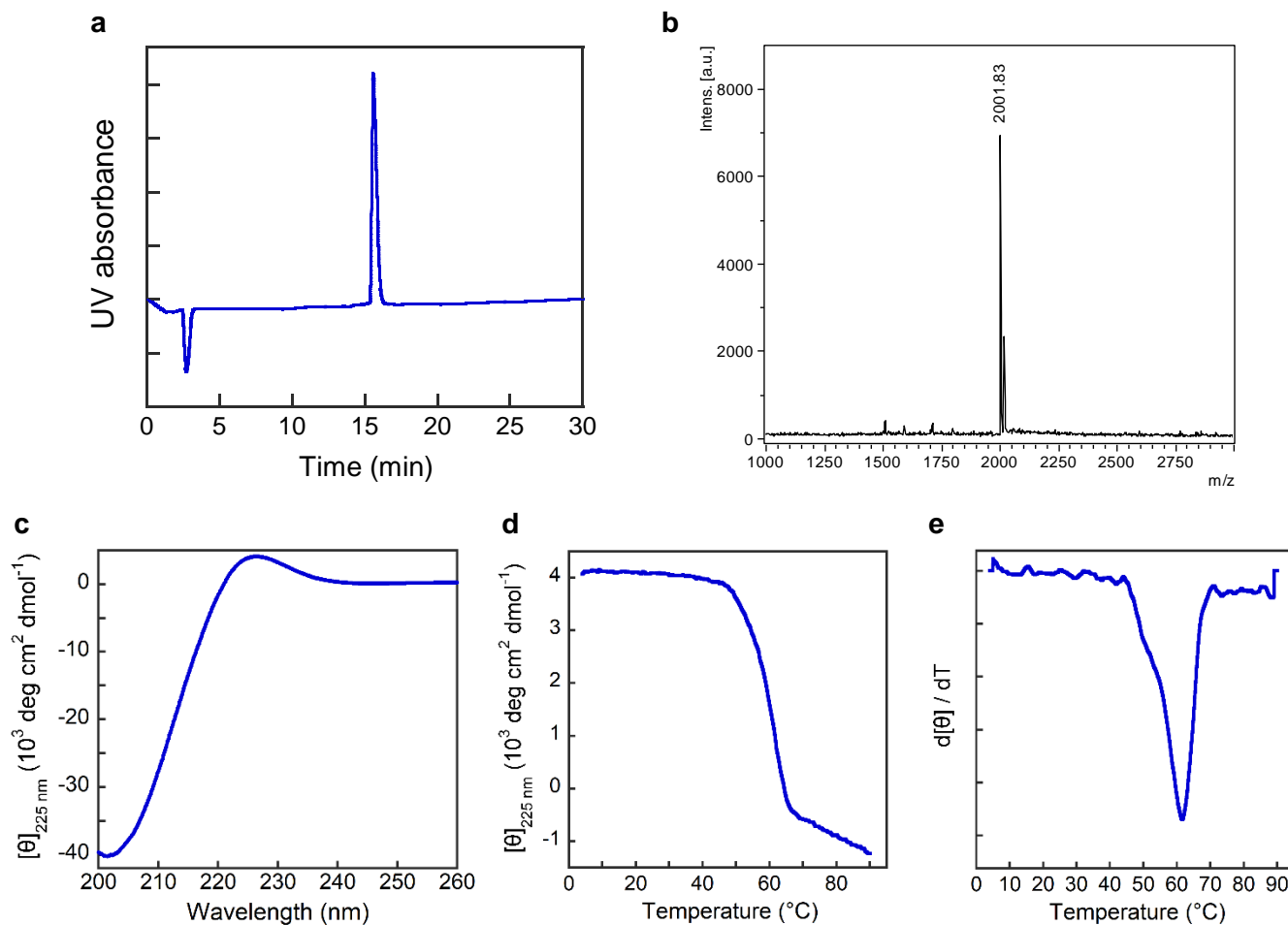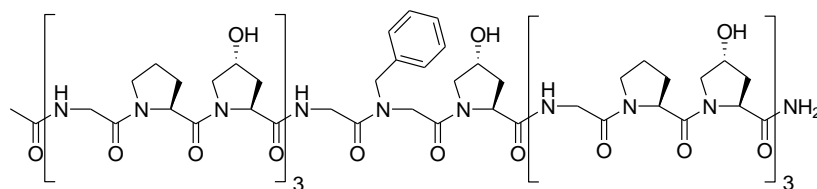
Nphe-Ile-CMP
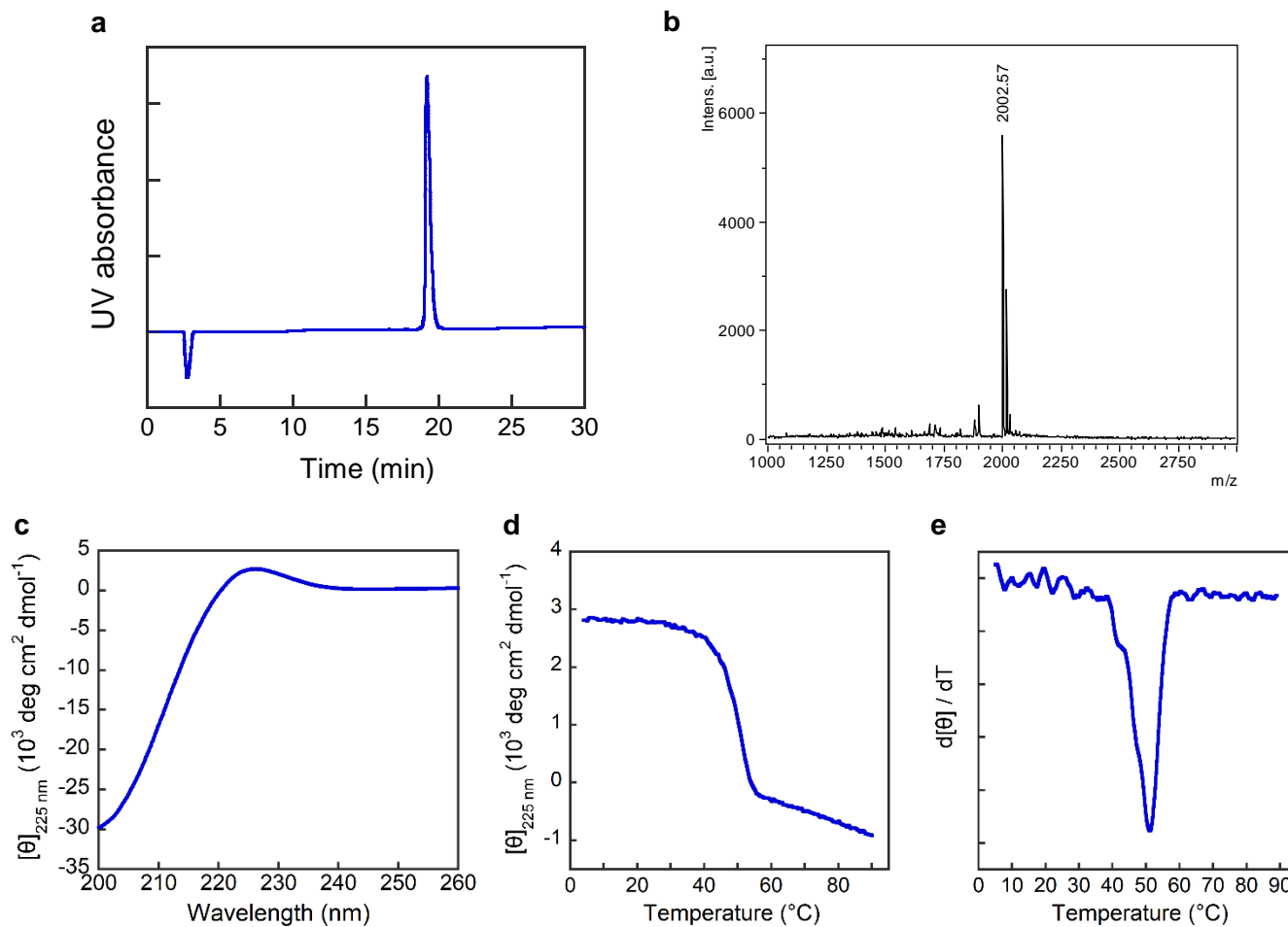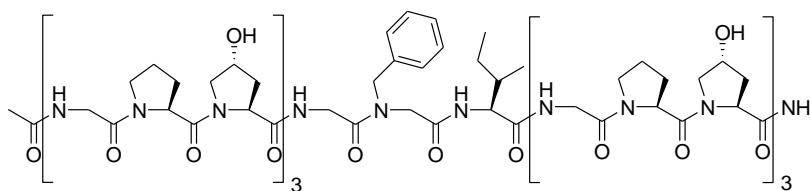
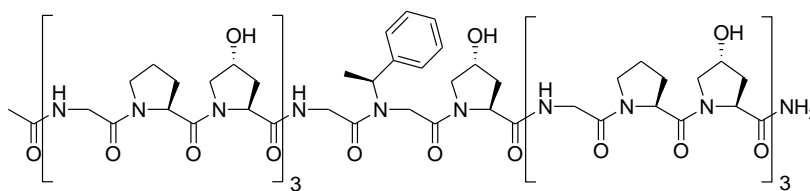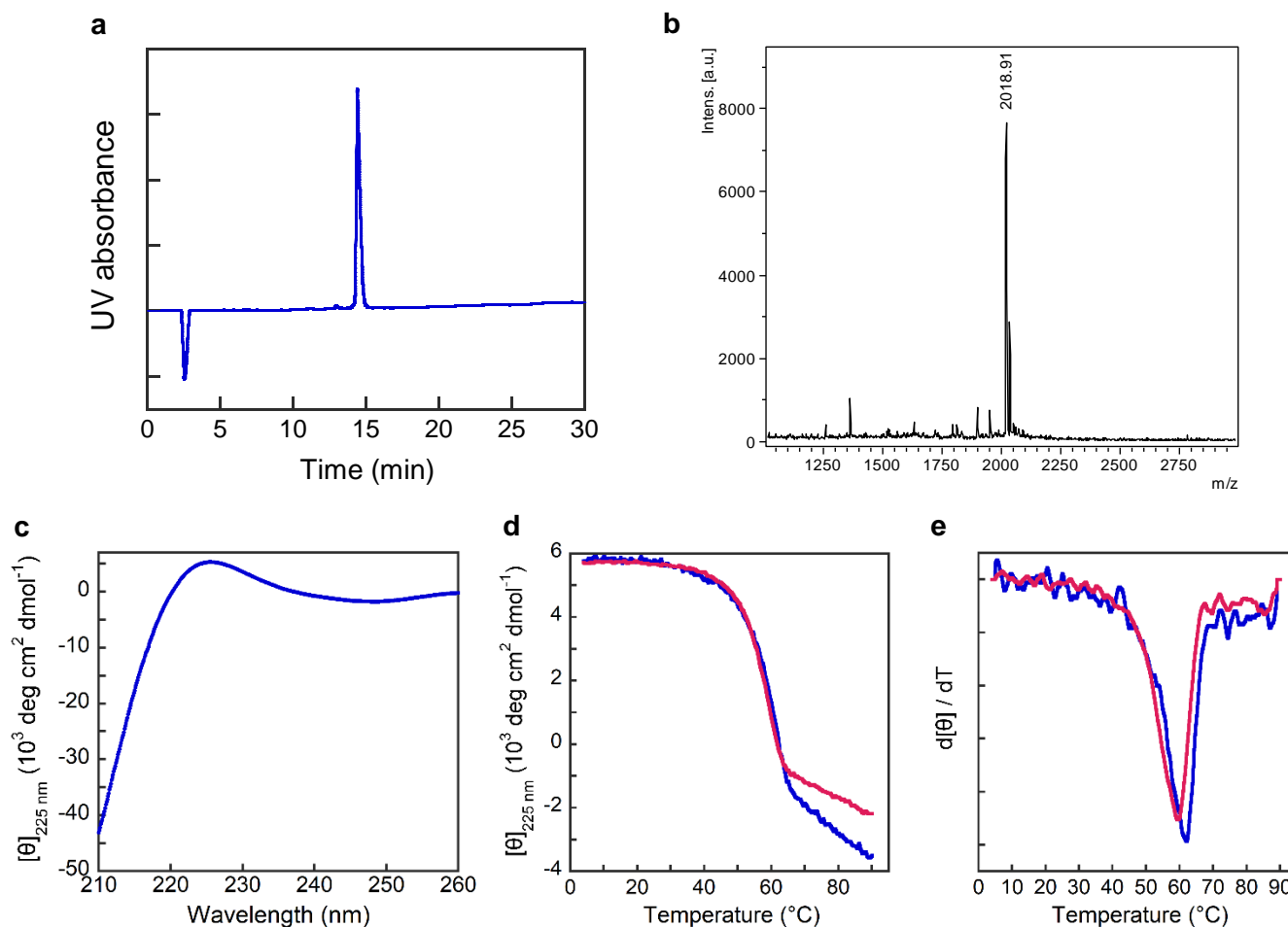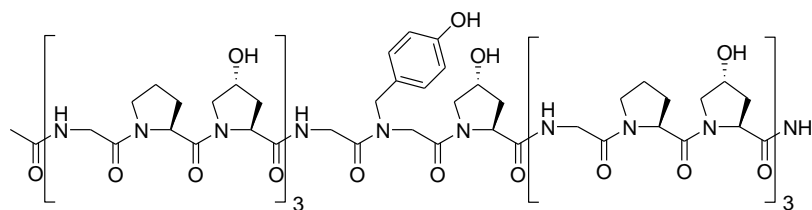**a**, The HPLC chromatogram of purified peptide, $t_R = 19.2$ min.

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, $T_m = 51$ °C.

Nrpe-CMP
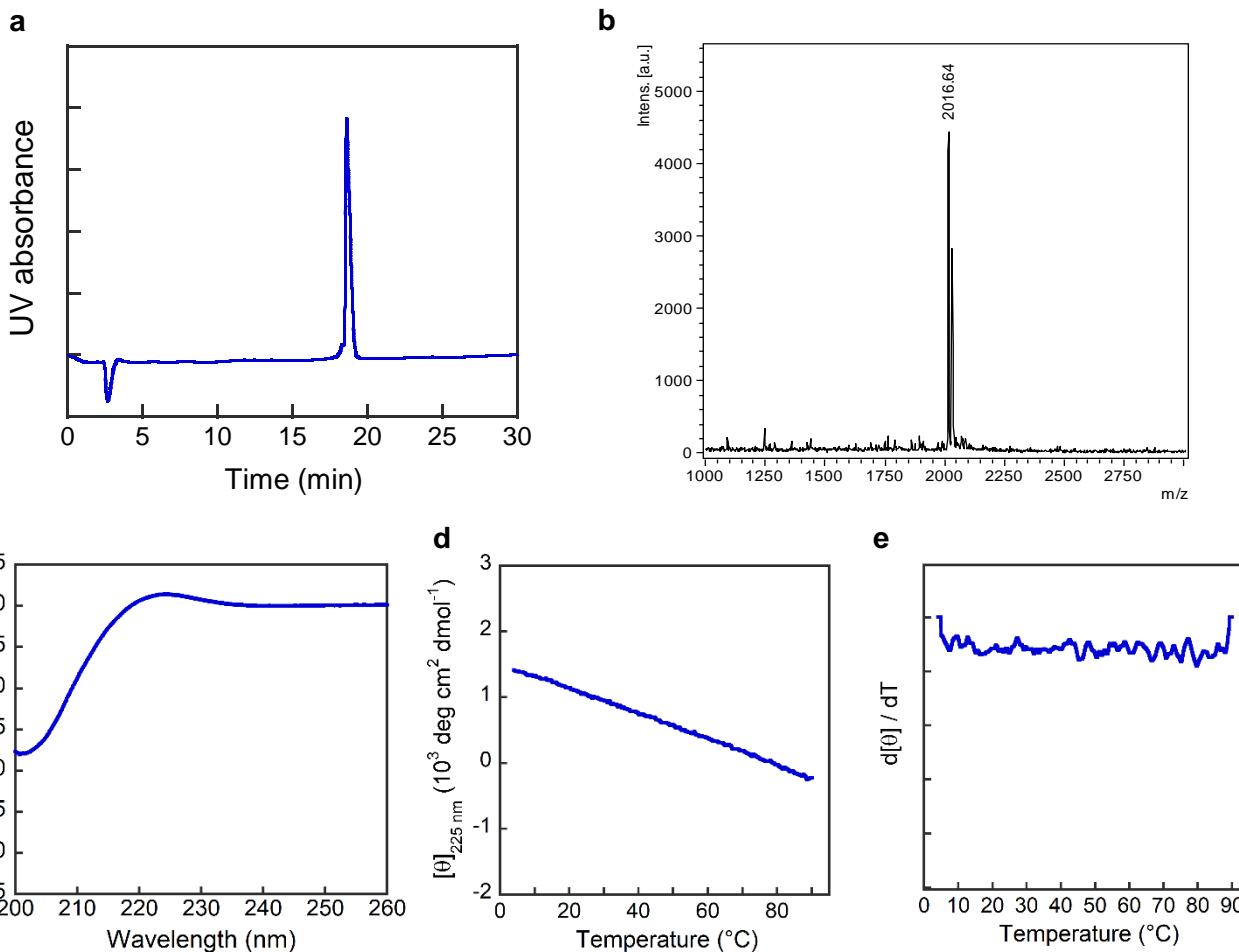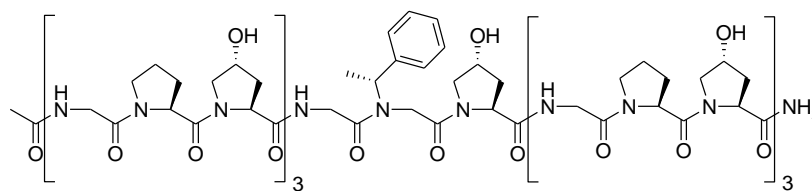
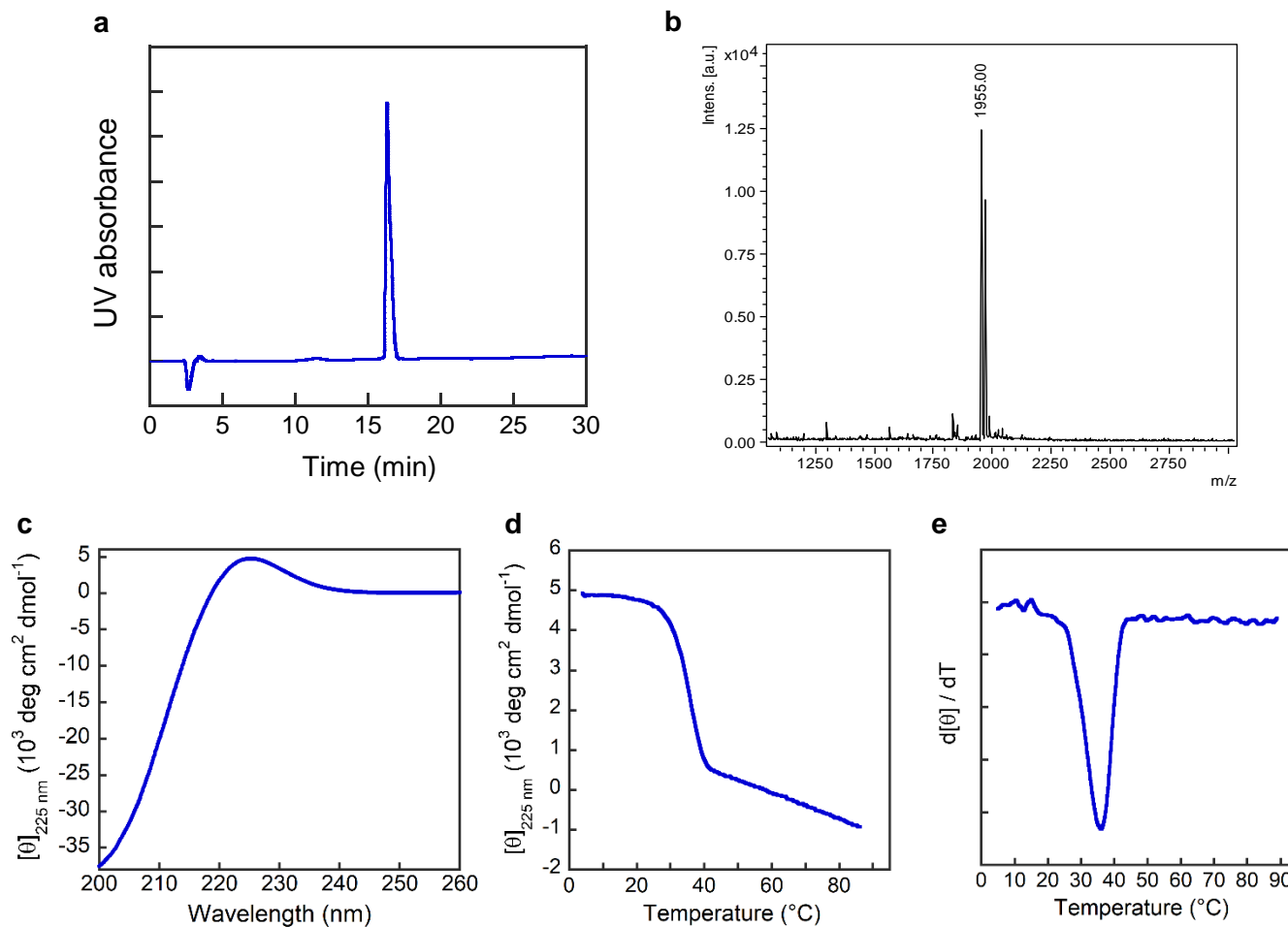**a**, The HPLC chromatogram of purified peptide, t_R = 18.6 min.

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

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

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve.

Nspe-CMP

a

b

c

d

e

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 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, T_m = 48 °C.

Ntyr-CMP

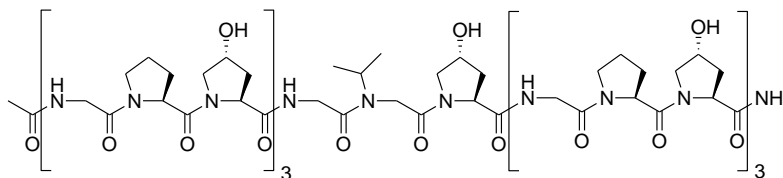**a**, The HPLC chromatogram of purified peptide, $t_R = 14.4$ min.

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]$^+$.

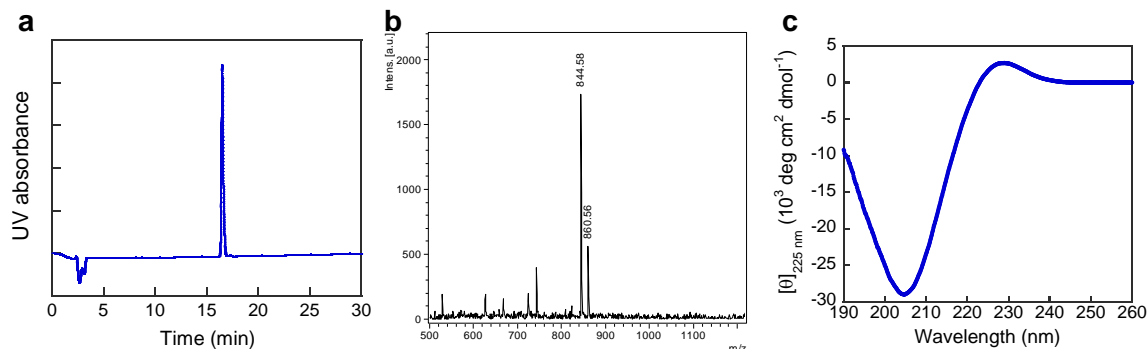**c**, The CD spectrum in PBS buffer at 4 °C.

d, The CD thermal melting curve in PBS buffer.

e, The first derivative of the melting curve, T_m = 36 °C.

S53

X7-CMP
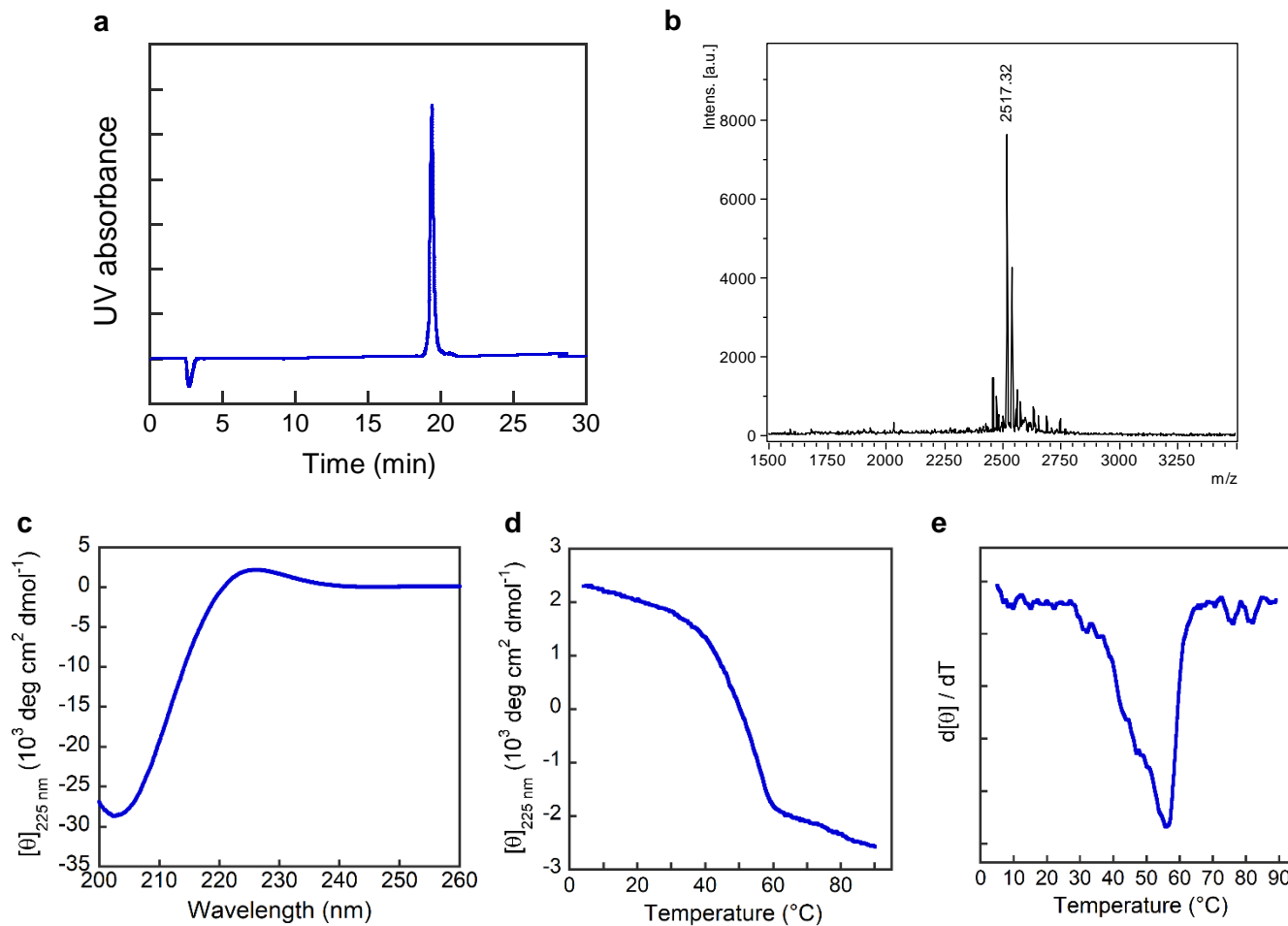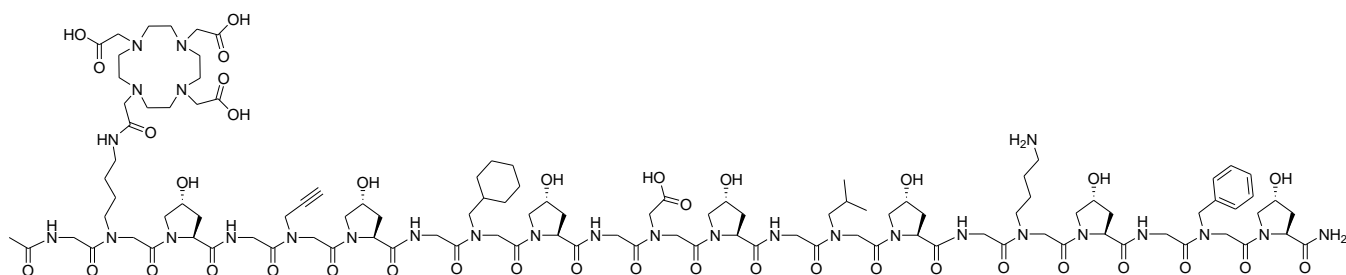
a, The HPLC chromatogram of purified peptide, t_R = 19.4 min.

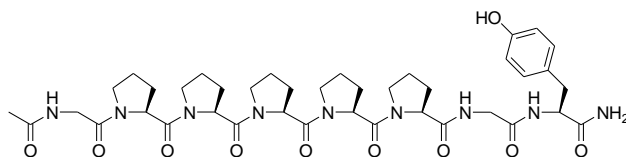**b**, MALDI-MS, calculated: 2516.2 $[M+H]^+$, observed: 2517.3 $[M+H]^+$.

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 = 56 °C.
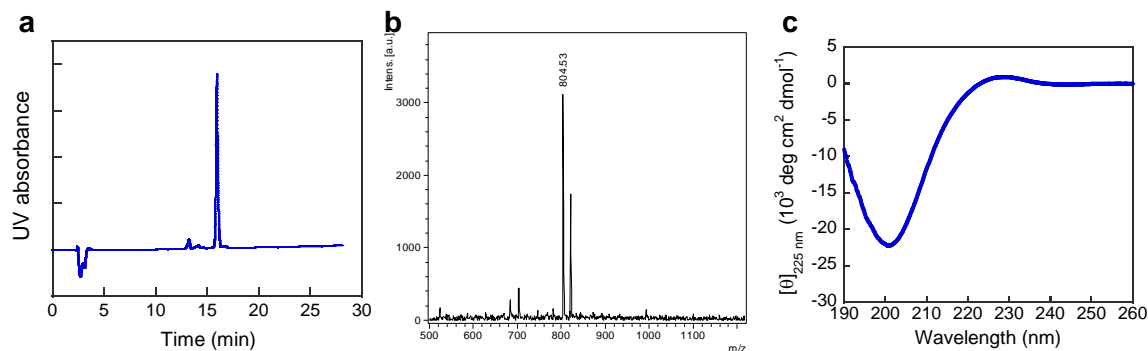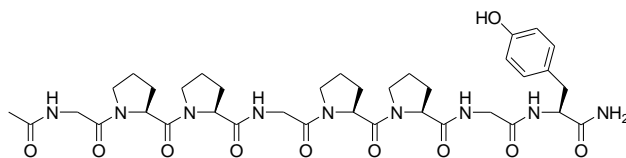
X-PP5 peptides

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.
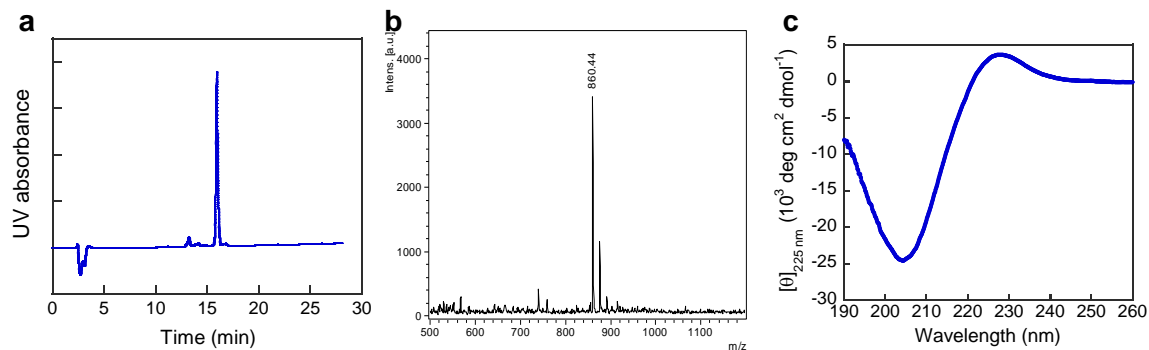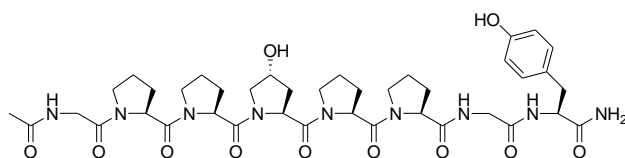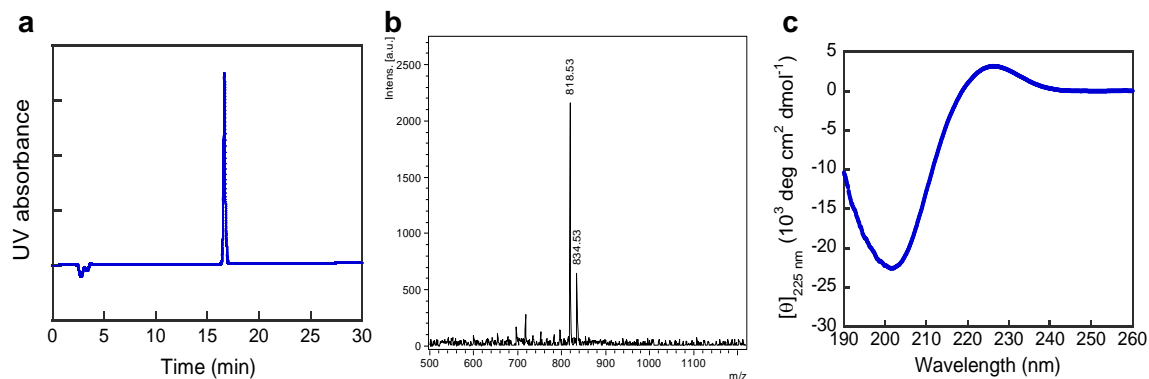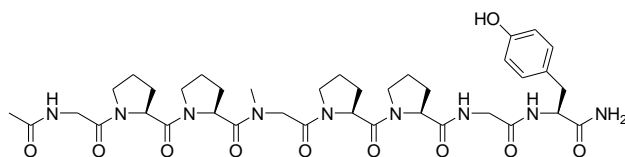
Hyp-PP5

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

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

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

Sar-PP5

a, The HPLC chromatogram of purified peptide, t_R = 16.6 min.

b, MALDI-MS, calculated: 818.4 [M+Na]⁺, observed: 818.5 [M+Na]⁺, 834.5 [M+K]⁺.

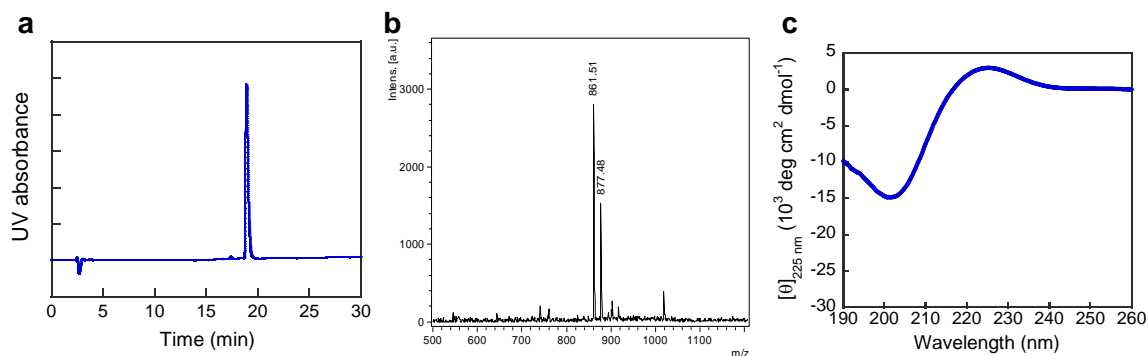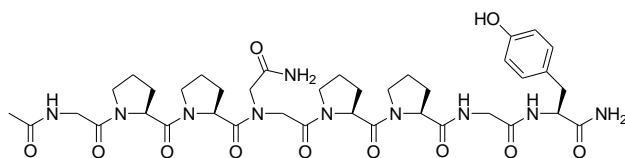**c**, The CD spectrum in 5 mM phosphate buffer at 25 °C.

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]$^+$.

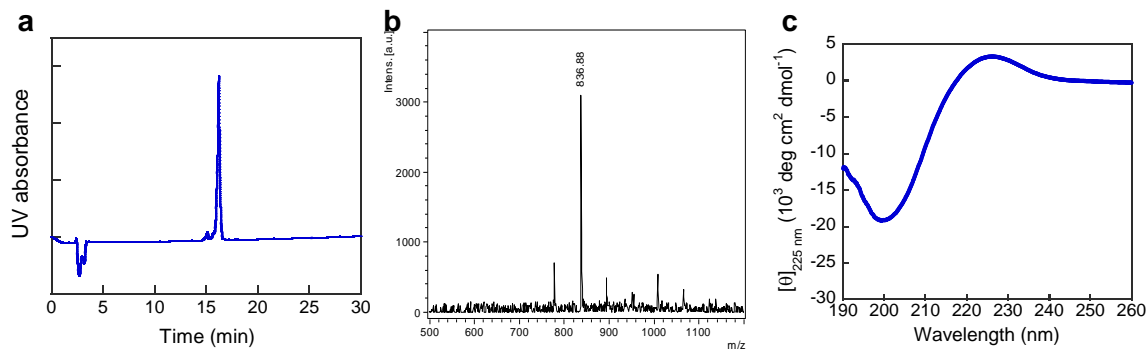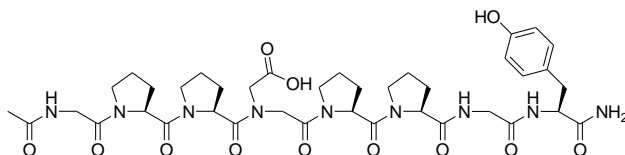**c**, The CD spectrum in 5 mM phosphate buffer at 25 °C.

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]⁺.

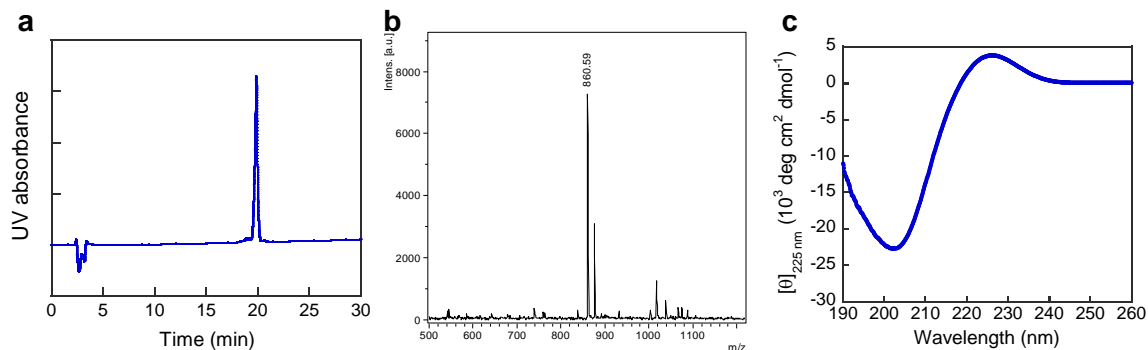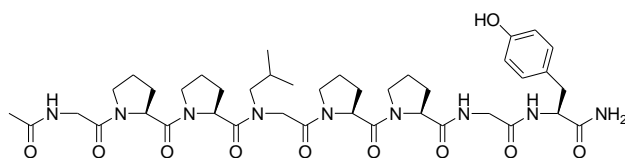**c**, The CD spectrum in 5 mM phosphate buffer at 25 °C.

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.
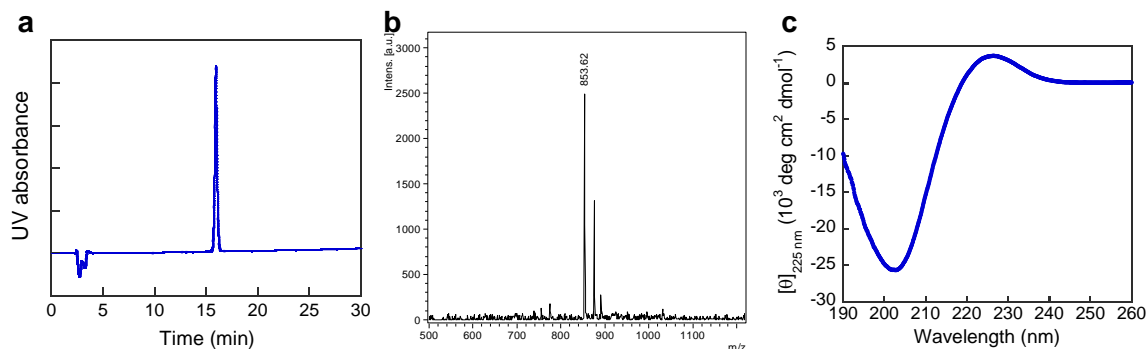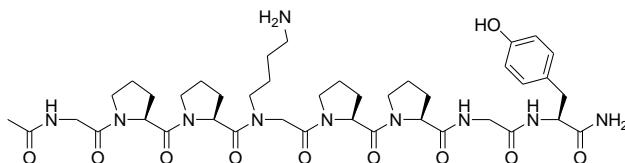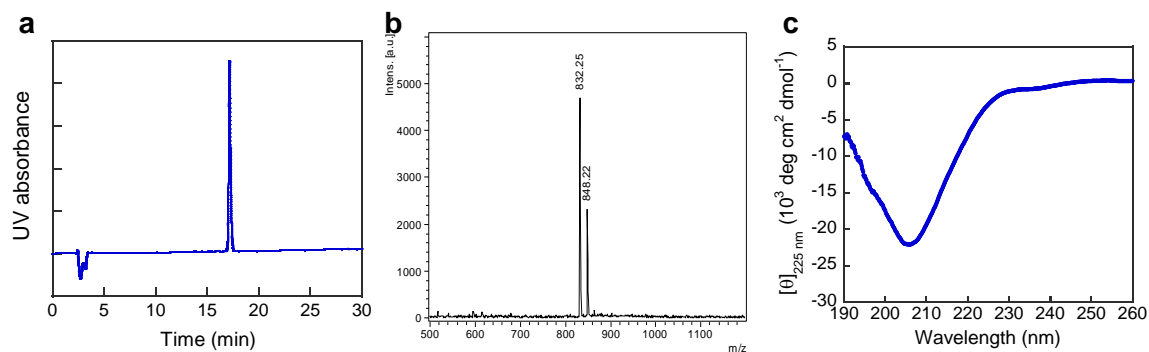
Nval-PP5

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]$^+$.

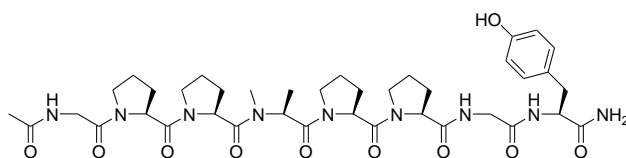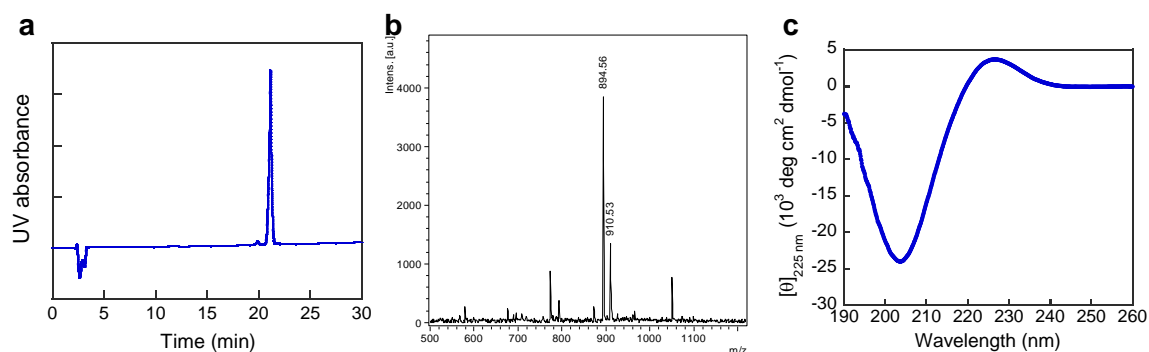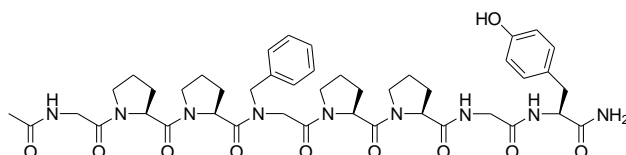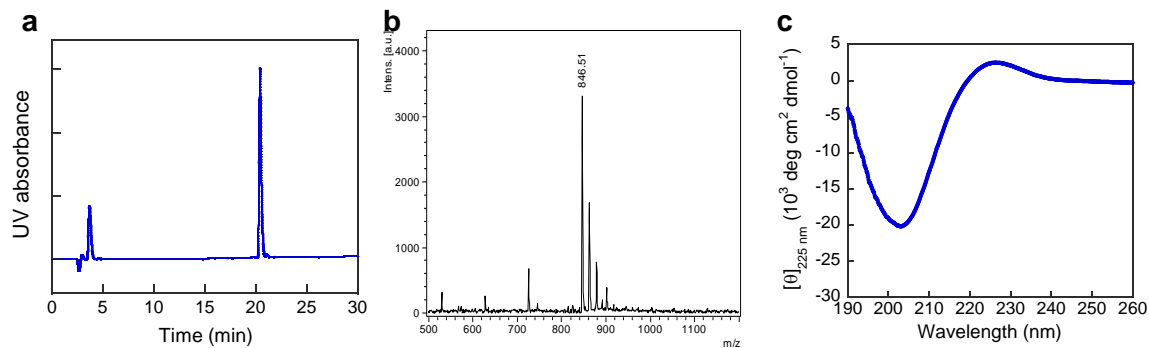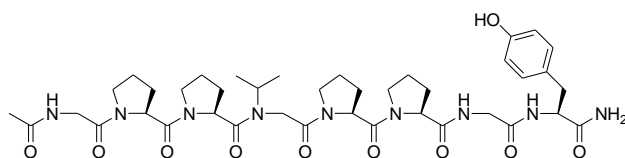**c**, The CD spectrum in 5 mM phosphate buffer at 25 °C.

References

(1)   Erdmann, R. S.; Wennemers, H. Functionalizable Collagen Model Peptides. *J. Am. Chem. Soc.* **2010**, *132* (40), 13957–13959.

(2)   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.* **1992**, *114* (26), 10646–10647.

(3)   Li, Y.; Foss, C. a.; Summerfield, D. D.; Doyle, J. J.; Torok, C. M.; Dietz, H. C.; Pomper, M. G.; Yu, S. M. Targeting Collagen Strands by Photo-Triggered Triple-Helix Hybridization. *Proc. Natl. Acad. Sci.* **2012**, *109* (37), 14767–14772.

(4)   Li, Y.; San, B. H.; Kessler, J. L.; Kim, J. H.; Xu, Q.; Hanes, J.; Yu, S. M. Non-Covalent Photo-Patterning of Gelatin Matrices Using Caged Collagen Mimetic Peptides. *Macromol. Biosci.* **2015**, *15*, 52–62.

(5)   Otwinowski, Z.; Minor, W. Processing of X-Ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*; Carter Jr., C. W., Sweet, R. M., Eds.; Academic Press: New York, 1997; Vol. 276, pp 307–326.

(6)   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(7)   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501.

(8)   Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC 5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *D67*, 355–367.

(9)   Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100* (2), 1–4.

(10)   Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(11)   Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180* (10), 1961–1972.

(12)   Qin, Z.; Buehler, M. J. Molecular Mechanics of Mussel Adhesion Proteins. *J. Mech. Phys. Solids* **2014**, *62*, 19–30.

(13)   Qin, Z.; Fabre, A.; Buehler, M. J. Structure and Mechanism of Maximum Stability of Isolated Alpha-Helical Protein Domains at a Critical Length Scale. *Eur. Phys. J. E* **2013**, *36* (5).

(14)   Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31* (4), 671–690.

(15)   MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.

(16)   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.* **1980**, *102* (3), 939–947.

(17)   Goodman, M.; Melacini, G.; Feng, Y. Collagen-like Triple Helices Incorporating Peptoid Residues. *J. Am. Chem. Soc.* **1996**, *118* (44), 10928–10929.

(18)   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* **1997**, *36* (29), 8725–8732.