Stereoelectronic and steric effects in side chains preorganize a protein main chain

Matthew D. Shoulders, Kenneth A. Satyshur, Katrina T. Forest, and Ronald T. Raines

Page	Contents

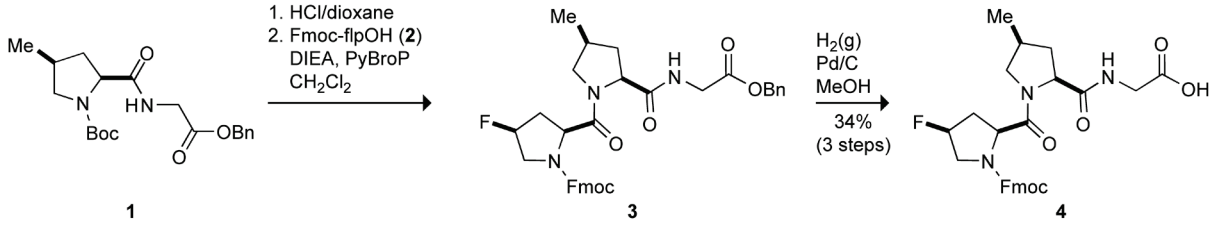
General Experimental Procedures. Commercial chemicals were of reagent grade or better, and were used without further purification. Anhydrous CH_2Cl_2 was obtained from a CYCLE-TAINER® solvent delivery system (J. T. Baker, Phillipsburg, NJ). In all reactions involving anhydrous solvents, glassware was either oven- or flame-dried. $NaHCO_3$(aq) and brine (NaCl) refer to saturated aqueous solutions. Flash chromatography was performed with columns of silica gel 60, 230–400 mesh (Silicycle, Québec City, Canada). Semi-preparative HPLC was performed with a Zorbax C-8 reversed-phase column. Analytical HPLC was performed with a Varian C-18 reversed-phase column. HPLC purifications and analyses employed linear gradients of solvent A (H_2O with 0.1% v/v TFA) and solvent B (CH_3CN with 0.1% v/v TFA).

The term "concentrated under reduced pressure" refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 torr) while maintaining the water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 torr). The term "high vacuum" refers to vacuum achieved by a mechanical belt-drive oil pump.

NMR spectra were acquired with a Bruker DMX-400 Avance spectrometer (^{1}H, 400 MHz; ^{13}C, 100.6 MHz) at the National Magnetic Resonance Facility at Madison (NMRFAM). NMR spectra were obtained at ambient temperatures on samples dissolved in MeOH-d_4. Coupling constants J are provided in Hertz. Compounds with a Fmoc-protecting group exist as mixtures of Z and E isomers that do not interconvert on the NMR time scale at ambient temperatures. Accordingly, these compounds exhibit two sets of NMR signals.

Mass spectrometry was performed with either a Micromass LCT (electrospray ionization, ESI) in the Mass Spectrometry Facility in the Department of Chemistry or an Applied Biosystems Voyager DE-Pro (matrix-assisted laser desorption/ionization, MALDI) mass spectrometer in the University of Wisconsin Biophysics Instrumentation Facility.

Scheme 1. Synthesis of Fmoc-flpMepGlyOH (4)

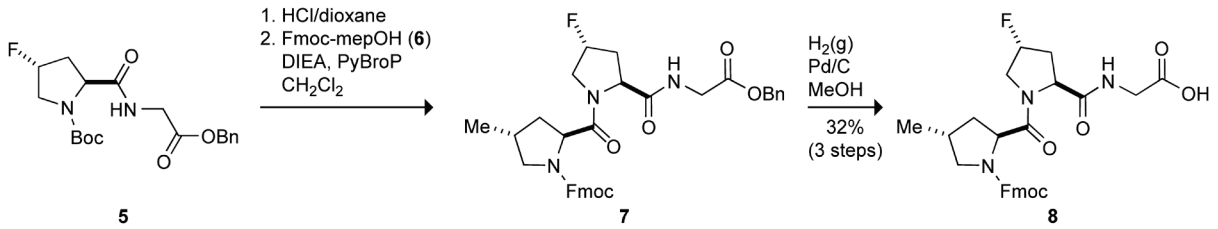


***N*-9-Fluorenylmethoxycarbonyl–(2*S*,4*S*)-4-fluoroprolyl–(2*S*,4*S*)-4-methylprolyl–glycine Benzyl Ester (3).** *N*-*tert*-Butyloxycarbonyl–(2*S*,4*S*)-4-methylprolyl–glycine benzyl ester **1** (0.80 g, 2.2 mmol), synthesized as described previously (1), was dissolved in 4 N HCl in dioxane (15 mL) under Ar(g). The resulting solution was stirred for 1.5 h. The solution was then concentrated under reduced pressure and dried under high vacuum, and the residue was dissolved in anhydrous CH_2Cl_2 (40 mL). *N*-9-Fluorenylmethoxycarbonyl–(2*S*,4*S*)-4-fluoroproline (**2**) (0.75 g, 2.2 mmol), synthesized by Dr. F. W. Kotch (University of Wisconsin–Madison) using established synthetic methodology (2), was added, and the resulting solution was cooled to 0 °C. PyBroP (1.03 g, 2.2 mmol) and DIEA (0.91 g, 7.0 mmol) were added. The resulting solution was allowed to warm slowly to room temperature, and then stirred for 4 h. The reaction mixture was diluted with CH_2Cl_2 (50 mL), washed with 10% w/v aqueous citric acid (100 mL), $NaHCO_3$(aq) (100 mL), and brine (100 mL), dried over anhydrous $MgSO_4$(s), and concentrated under reduced pressure. Flash chromatography (70% v/v EtOAc in hexane) afforded **3** (990 mg, 1.6 mmol)

containing a slight impurity that was removed in the subsequent step. HRMS–ESI (*m*/*z*): [M + Na]$^+$ calcd for $C_{35}H_{36}FN_3O_6Na$, 636.2486; found, 636.2473.

***N*-9-Fluorenylmethoxycarbonyl–(2*S*,4*S*)-4-fluoroprolyl–(2*S*,4*S*)-4-methylprolyl–glycine (4).** MeOH (75 mL) was added carefully to a mixture of compound **3** (990 mg, 1.6 mmol) and Pd/C (10% w/w, 160 mg, 0.3 mmol) under Ar(g), and the resulting black suspension was stirred under H_2(g) for 3 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The suspension was filtered through a Celite pad and concentrated under reduced pressure. The crude product was purified by flash chromatography (1.5% v/v MeOH in CH_2Cl_2 containing 0.1% v/v formic acid). The fractions containing **4** were concentrated under reduced pressure. The formic acid was removed by dissolving the residue in 10% v/v MeOH in toluene and concentrating under reduced pressure to afford **4** (390 mg, 0.7 mmol, 34%, 3 steps) as a white solid. ^{1}H NMR δ: 1.10 and 1.15 (d, J = 6.5, 3H), 1.49–1.64 (m, 1H), 2.20–2.62 (m, 4H), 3.00–3.15 (m, 1H), 3.32–4.06 (m, 5.6H), 4.07–4.71 (m, 3.4H), 5.04–5.38 (m, 1H), 7.25–7.46 (m, 4H), 7.47–7.66 (m, 2H), 7.67–7.77 (m, 2H); ^{13}C NMR δ: 17.0, 17.0, 35.4, 35.5, 36.4, 36.6, 37.0, 37.2, 38.5, 38.7, 41.9, 50.1, 54.1, 54.3, 54.6, 54.8, 55.0, 55.4, 55.7, 58.9, 59.2, 62.7, 62.7, 71.0, 71.8, 81.3, 81.5, 92.4 and 93.2 (d, $J_{C–F}$ = 178.4), 121.1, 121.2, 125.4, 125.7, 125.9, 125.9, 129.1, 129.2, 130.6, 130.7, 141.4, 141.5, 147.7, 147.8, 148.0, 148.2, 156.1, 171.8, 172.9, 174.8, 174.9; HRMS–ESI (*m*/*z*): [M + Na]$^+$ calcd for $C_{28}H_{30}FN_3O_6Na$, 546.2016; found, 546.2000.

Scheme 2. Synthesis of Fmoc-flpMepGlyOH (8)



N-9-Fluorenylmethoxycarbonyl–(2*S*,4*R*)-4-methylprolyl–(2*S*,4*R*)-4-fluoroprolyl–glycine Benzyl Ester (7). *N*-*tert*-Butyloxycarbonyl–(2*S*,4*R*)-4-fluoroprolyl–glycine benzyl ester **5** (0.71 g, 1.9 mmol), synthesized by Dr. Frank W. Kotch using established synthetic methodology (2), was dissolved in 4 N HCl in dioxane (25 mL) under Ar(g) and stirred for 75 min. The resulting solution was concentrated under reduced pressure and the residue dissolved in anhydrous CH_2Cl_2 (35 mL). *N*-9-Fluorenylmethoxycarbonyl–(2*S*,4*R*)-4-methylproline (**6**; 0.69 g, 1.9 mmol), synthesized as described previously (1), PyBop (0.99 g, 1.9 mmol) and DIEA (0.87 g, 6.7 mmol) were added. The resulting solution was stirred for 20 h. The reaction mixture was diluted with CH_2Cl_2 (50 mL), washed with 10% w/v aqueous citric acid (100 mL), $NaHCO_3$(aq) (100 mL), and brine (100 mL), dried over anhydrous $MgSO_4$(s), and concentrated under reduced pressure. Flash chromatography (70% v/v EtOAc in hexane) afforded slightly impure **7** (990 mg, 1.6 mmol) as a white solid. HRMS–ESI (*m*/*z*): [M + Na]$^+$ calcd for $C_{35}H_{36}FN_3O_6Na$, 636.2486; found, 636.2463.

***N*-9-Fluorenylmethoxycarbonyl–(2*S*,4*R*)-4-methylprolyl–(2*S*,4*SR*)-4-fluoroprolyl–glycine (8).** MeOH (75 mL) was added carefully to a mixture of compound **7** (560 mg, 0.9 mmol) and Pd/C (10% w/w, 130 mg, 0.2 mmol) under Ar(g), and the resulting black suspension was stirred under H_2(g) for 3 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc

group. The suspension was filtered through a pad of Celite, and concentrated under reduced pressure. The crude product was purified by flash chromatography (EtOAc to elute byproducts, then 10% v/v MeOH in CH_2Cl_2 containing 0.1% v/v formic acid). The fractions containing **8** were concentrated under reduced pressure, and the formic acid was removed by dissolving the residue in 10% v/v MeOH in toluene and concentrating under reduced pressure to afford **8** (327 mg, 0.6 mmol, 32% two-step yield) as a white solid. ^{1}H NMR δ: 1.03 and 1.04 (d, J = 6.6, 3H), 1.78–1.91 (m, 1H), 2.07–2.28 (m, 2H), 2.32–2.64 (m, 2H), 2.90–2.98 (m, 1H), 3.28–3.62 (m, 6H), 4.32–4.47 (m, 2H), 4.52–4.68 (m, 2H), 5.23–5.44 (m, 1H), 7.27–7.43 (m, 4H), 7.53–7.68 (m, 2H), 7.76–7.83 (m, 2H); ^{13}C NMR δ: 17.6, 17.7, 32.0, 33.1, 36.8, 36.9, 37.0, 37.1, 37.6, 38.6, 41.8, 49.9, 54.7, 54.8, 55.1, 59.8, 60.2, 68.6, 68.7, 93.5 (d, $J_{C–F}$ = 178.5), 120.9, 126.1, 126.1, 126.3, 128.2, 128.2, 128.4, 128.8, 142.5, 142.6, 145.0, 145.1, 145.5, 145.6, 156.2, 156.7, 172.5, 173.3, 174.0, 174.1; HRMS–ESI (m/z): $[M – H]^-$ calcd for $C_{28}H_{29}FN_3O_6$, 523.2119; found, 522.2015.

Solid-Phase Peptide Synthesis. $(ProProGly)_7$ and $(ProHypGly)_7$ were prepared via coupling of the appropriate Fmoc-protected amino acids and H-Gly-2-chlorotrityl resin using an Applied Biosystems Synergy 432A Peptide Synthesizer at the University of Wisconsin Biotechnology Center. Fmoc-deprotection was achieved by treatment with 20% (v/v) piperidine in DMF. The trimers (3 equivalents) were converted to active esters by treatment with HBTU, DIEA, and HOBt. Extended couplings (120 min) were employed at room temperature.

$(flpMepGly)_7$ and $(mepFlpGly)_7$ were prepared as follows. Fmoc-tripeptides **4** and **8** were loaded onto 2-chlorotrityl resin as described previously for similar amino acid trimers (1). Loadings were measured by ultraviolet spectroscopy (3) to be 0.52 mmol/g for **4** and 0.61 mmol/g for **8**. The two 21-mer peptides were synthesized on a 25 μmol-scale by segment condensation of their corresponding Fmoc-tripeptides (**4** and **8**) on solid phase using an Applied Biosystems Synergy 432A Peptide Synthesizer at the University of Wisconsin Biotechnology Center. Fmoc-deprotection was achieved by treatment with 20% (v/v) piperidine in DMF. The trimers (3 equiv) were converted to active esters by treatment with HBTU, DIEA, and HOBt. Extended couplings (60–90 min) were employed at room temperature.

Peptides were cleaved from the resin in 95:3:2 TFA:triisopropylsilane:H_2O (2 mL), precipitated from *tert*-butylmethylether at 0 °C, and isolated by centrifugation. Semi-preparative HPLC was used to purify the peptides $(ProProGly)_7$ (gradient: 10% B to 80% B over 50 min), $(ProHypGly)_7$ (gradient: 10% B to 60% B over 50 min), $(flpMepGly)_7$ (gradient: 3% B to 50% B over 50 min), and $(mepFlpGly)_7$ (gradient: 5% B to 60% B over 60 min). All four peptides were >90% pure by analytical HPLC and MALDI–TOF mass spectrometry (m/z) $[M + H]^+$ calcd for $C_{84}H_{122}N_{21}O_{22}$ 1776.9; found 1776.5 for $(ProProGly)_7$, calcd for $C_{84}H_{122}N_{21}O_{29}$ 1888.9; found 1888.1 for $(ProHypGly)_7$, and calcd for $C_{91}H_{129}F_7N_{21}O_{22}$ 2000.9; found 2001.1 for $(flpMepGly)_7$ and 2001.4 for $(mepFlpGly)_7$.

CD Spectroscopy of $(ProProGly)_7$, $[(ProHypGly)_7]_3$, $[(flpMepGly)_7]_3$, and $[(mepFlpGly)_7]_3$. Peptides were dried under vacuum for at least 24 h before being weighed and dissolved to 0.2 mM in 50 mM HOAc(aq) (pH 2.9). The solutions were incubated at ≤4 °C for ≥24 h before CD spectra were acquired with an Aviv Associates (Lakewood, NJ) 202SF CD spectrometer. Spectra were measured with a 1-nm band-pass in cuvettes with a 0.1-cm pathlength. The signal was averaged for 3 s during wavelength scans, and either 5 or 15 s during denaturation experiments using a 0.6-°C temperature dead band. (Changing the averaging time did not alter the T_m value.) During

denaturation experiments, CD spectra were acquired at intervals of 3 °C. At each temperature, solutions were equilibrated for 5 min before data acquisition. Values of T_m were determined in triplicate by fitting the molar ellipticity at 225 nm to a two-state model (4).

Differential Scanning Calorimetry of (ProProGly)$_7$, [(ProHypGly)$_7$]$_3$, [(flpMepGly)$_7$]$_3$, and [(mepFlpGly)$_7$]$_3$. DSC measurements were conducted on a VP-DSC instrument (MicroCal, Northampton, MA). Instrument baselines were established by filling both the sample and the reference cells with degassed 50 mM HOAc(aq) and scanning from 5–98 °C at a scan rate of 6 °C/h until at least two consecutive overlaying buffer-buffer scans were observed. The final buffer versus buffer scan was used as the baseline for subsequent scans.

Peptide solutions (~50 μM in 50 mM HOAc(aq)) were incubated at ≤4 °C for ≥24 h prior to degassing and loading in the sample cell (the reference cell solution was not replaced) during cooling from the previous buffer versus buffer scan (at ~10 °C). Samples were scanned from 5–98 °C at a scan rate of 6 °C/h; the first scan of each sample was used in the analysis. Data could be reproduced reliably by re-incubating the same samples at ≤4 °C for ≥24 h, and then rescanning.

After DSC measurements, the concentration of each peptide that formed a stable triple helix was determined by quantitative amino-acid analysis at the Protein Chemistry Laboratory, University of Texas Medical Branch (Galveston, TX); the concentration of (ProProGly)$_7$ was estimated by high-performance liquid chromatography analysis. Peptide concentrations of 100, 22, 51, and 83 μM were obtained for (ProProGly)$_7$, (ProHypGly)$_7$, (flpMepGly)$_7$, and (mepFlpGly)$_7$, respectively.

For each sample, the appropriate reference scan was subtracted from the sample scan. The data were normalized to monomer quantity (P_{Tot}) and then treated using the simplest, physically valid model as a two-state, association–dissociation equilibrium with ΔC_p dependent on temperature (5). Relevant equations and the methodology for data-fitting are described below.

The equilibrium between a triple helix and single coil for a single, two-state system is represented as

$$\frac{1}{3}t \underset{\leftarrow}{\overset{\rightarrow}{}} s \qquad \textbf{(S.1)}$$

For the equilibrium in equation **S.1**,

$$\Delta G(T) = \Delta H(T) - T\Delta S(T) \qquad \textbf{(S.2)}$$

$\Delta H(T)$ is defined as

$$\Delta H(T) = \Delta H(T°) + \int_{T°}^{T} \Delta C_P(T')dT' \qquad \textbf{(S.3)}$$

and $\Delta S(T)$ is defined as

$$\Delta S(T) = \Delta S(T°) + \int_{T°}^{T} \frac{\Delta C_P(T')}{T'} dT' \qquad \textbf{(S.4)}$$

In our model, $T°$ is the reference temperature and is defined as the temperature where $\Delta G = 0$, in contrast to the T_m which is defined as the temperature where the transition is half-completed. Therefore,

$$\Delta G(T°) = \Delta H(T°) - T°\Delta S(T°) = 0 \Rightarrow \Delta S(T°) = \frac{\Delta H(T°)}{T°} \quad \textbf{(S.5)}$$

$\Delta C_P(T)$ is the difference in molar heat capacities of peptides in the single-coil form and in the triple-helix form. Thus,

$$\Delta C_P(T) = C_{P,s}(T) - C_{P,t}(T) \quad \textbf{(S.6)}$$

where $C_{P,s}(T)$ is the portion of $C_{P,base}(T)$ (defined in equation **S.14**) attributable to the single-coil state and $C_{P,t}(T)$ is the portion of $C_{P,base}(T)$ attributable to the triple-helix state. If we approximate $C_{P,s}(T)$ and $C_{P,t}(T)$ as linear functions of temperature, then

$$C_{P,s}(T) = B_s + D_s T \quad \textbf{(S.7)}$$

where B_s is the intercept and D_s is the slope of the line $C_{P,s}(T)$ and

$$C_{P,t}(T) = B_t + D_t T \quad \textbf{(S.8)}$$

where B_t is the intercept and D_t is the slope of the line $C_{P,t}(T)$. B_s, D_s, B_t, and D_t are, therefore, independent of temperature. Equation **S.6** then becomes

$$\Delta C_P(T) = B_s - B_t + (D_s - D_t)T \quad \textbf{(S.9)}$$

and $\Delta H(T)$ and $\Delta S(T)$ can be expressed as

(S.10)

and

(S.11)

Based on equation **S.2**, integrating and solving for $\Delta G(T)$ yields

$$\Delta G(T) = \Delta H(T°) + (B_s - B_t)(T - T°) + 0.5(D_s - D_t)(T^2 - T°^2) - T\left[\frac{\Delta H(T°)}{T°} + (B_s - B_t)\ln\left(\frac{T}{T°}\right) + (D_s - D_t)(T - T°)\right] \quad \textbf{(S.12)}$$

Hence, with known values of $T°$, $\Delta H(T°)$, B_s, D_s, B_t, and D_t, we can calculate ΔC_P, ΔH, ΔS, and ΔG at any temperature T. To retrieve $T°$, $\Delta H(T°)$, B_s, D_s, B_t, and D_t, we first derive the expression of the observed excess heat capacity function with these parameters. The observed excess heat capacity function, $C_{P,obs}(T)$, can be deconvoluted as

$$C_{P,obs}(T) = C_{P,trans}(T) + C_{P,base}(T) \quad \textbf{(S.13)}$$

where $C_{P,trans}(T)$ is the portion of $C_{P,obs}(T)$ attributable to the transition and $C_{P,base}(T)$ is the portion of $C_{P,obs}(T)$ attributable to the triple-helix and single-coil states (baseline function). $C_{P,base}(T)$ can be written as

$$C_{P,base}(T) = C_{P,t}(T)F_t(T) + C_{P,s}(T)F_s(T) \quad \textbf{(S.14)}$$

$C_{P,trans}(T)$ can be written as

$$C_{P,trans}(T) = \Delta H(T)\frac{dF_s(T)}{dT} = \left[\Delta H(T°) + (B_s - B_t)(T - T°) + 0.5(D_s - D_t)(T^2 - T°^2)\right]\frac{dF_s(T)}{dT} \quad \textbf{(S.15)}$$

In these equations, $F_s(T)$ is the fraction of peptides in the single-coil form

$$F_s(T) = \frac{s(T)}{P_{Tot}} \quad \textbf{(S.16)}$$

and $F_t(T)$ is the fraction of peptides in the triple-helix form

$$F_t(T) = \frac{3t(T)}{P_{Tot}} \quad \textbf{(S.17)}$$

with P_{Tot} (the overall peptide concentration determined independently by amino acid analysis) given by

$$P_{Tot} = s(T) + 3t(T) \quad \textbf{(S.18)}$$

$F_t(T)$ can then be expressed in terms of $F_s(T)$ as

$$F_t(T) = 1 - F_s(T) \quad \textbf{(S.19)}$$

The function $F_s(T)$ is obtained as follows: The equilibrium constant $K(T)$ for equation **S.1** is defined as

$$K(T) = \frac{s(T)}{t(T)^{1/3}} \quad \textbf{(S.20)}$$

$K(T)$ can be expressed in terms of $F_s(T)$ and P_{Tot} as

$$K(T) = \frac{\left(F_s(T)P_{Tot}\right)}{\left[\frac{\left(1-F_s(T)\right)P_{Tot}}{3}\right]^{1/3}} \quad \textbf{(S.21)}$$

Equation **S.21** can be rearranged to afford the cubic equation

$$F_s^3(T) + \frac{F_s(T)K^3(T)}{3P_{Tot}^2} - \frac{K^3(T)}{3P_{Tot}^2} = 0 \quad \textbf{(S.22)}$$

Equation **S.22** can be solved numerically if $K(T)$ is known. $K(T)$ can also be expressed as

$$K(T) = e^{-\left[\Delta G(T)/RT\right]} \quad \textbf{(S.23)}$$

$\Delta G(T)$ can be calculated by using equation **S.12**. In our fitting procedure (using code written by Y.-S. Lin), $\Delta G(T)$ and $K(T)$ are calculated with a set of $T°$, $\Delta H(T°)$, B_s, D_s, B_t, and D_t. Then, equation **S.22** is numerically solved to obtain $F_s(T)$ and numerically differentiated to obtain $dF_s(T)/dT$. Equations **S.13**, **S.14**, and **S.15** are then utilized to generate a C_P fitting curve. The six parameters $T°$, $\Delta H(T°)$, B_s, D_s, B_t, and D_t, are optimized to generate the best fit to $C_{P,obs}(T)$ by implementing the Fletcher-Reeves-Polak-Ribiere minimization of the square deviation between the calculated and empirical curves. P_{Tot} and optimized values for the six fitting parameters for each peptide are displayed in Table S1.

Crystallization of Ac-Mep-OMe. [$^{13}CH_3$]Ac-Mep-OMe (15 mg), prepared as described previously (1), was dissolved in EtOAc (1.5 mL) and the solution was allowed to stand at –20 °C in a loosely capped vial. Slow evaporation afforded crystals suitable for X-ray analysis after ~1 month (by then, all solvent had evaporated).

X-Ray Crystallography Data Collection for Ac-Mep-OMe. A colorless crystal with approximate dimensions 0.41 × 0.36 × 0.29 mm^3 was selected under oil under ambient conditions and attached to the tip of a nylon loop. The crystal was mounted in a stream of cold N_2(g) at 100(2) K and centered in the X-ray beam by using a video camera. The crystal evaluation and data collection were performed on a Bruker CCD-1000 diffractometer with Mo K_α (λ = 0.71073 Å) radiation and the diffractometer to crystal distance of 4.7 cm.

The initial cell constants were obtained from three series of ω scans at different starting angles. Each series consisted of 20 frames collected at intervals of 0.3° in a 6° range about ω with the exposure time of 10 s per frame. A total of 52 reflections were obtained. The reflections were successfully indexed by an automated indexing routine built in the SMART program. The final cell constants were calculated from a set of 5781 strong reflections from the actual data collection.

The data were collected by using the full sphere data collection routine to survey the reciprocal space to the extent of a full sphere to a resolution of 0.80 Å. A total of 6859 data were harvested by collecting four sets of frames with 0.36° scans in ω and one set with 0.45° scans in

φ with an exposure time of 20 s per frame. These highly redundant datasets were corrected for Lorentz and polarization effects. The absorption correction was based on fitting a function to the empirical transmission surface as sampled by multiple equivalent measurements (6).

Structure Solution and Refinement for Ac-Mep-OMe. The systematic absences in the diffraction data were uniquely consistent for the space group $P2_12_12_1$ that yielded chemically reasonable and computationally stable results of refinement (6). A successful solution by the direct methods provided most non-hydrogen atoms from the *E*-map. The remaining non-hydrogen atoms were located in an alternating series of least-squares cycles and difference Fourier maps. All non-hydrogen atoms were refined with anisotropic displacement coefficients. All hydrogen atoms were included in the structure factor calculation at idealized positions and were allowed to ride on the neighboring atoms with relative isotropic displacement coefficients.

The final least-squares refinement of 122 parameters against 1750 data resulted in residuals R (based on F^2 for $I \geq 2\sigma$) and wR (based on F^2 for all data) of 0.0779 and 0.2120, respectively. The final difference Fourier map was featureless.

References

1. Shoulders MD, Hodges JA, Raines RT (2006) Reciprocity of steric and stereoelectronic effects in the collagen triple helix. *J Am Chem Soc* 128:8112–8113.
2. Hodges JA, Raines RT (2005) Stereoelectronic and steric effects in the collagen triple helix: Toward a code for strand association. *J Am Chem Soc* 127:15923–15932.
3. AppliedBiosystems. Determination of the amino acid substitution level via an Fmoc assay. http://docs.appliedbiosystems.com/search/taf.
4. Becktel WJ, Schellman JA (1987) Protein stability curves. *Biopolymers* 26:1859–1877.
5. Nishi Y, Uchiyama S, Doi M, Nishiuchi Y, Nakazawa T, Ohkubo T, Kobayashi Y (2005) Different effects of 4-hydroxyproline and 4-fluoroproline on the stability of the collagen triple helix. *Biochemistry* 44:6034–6042.
6. Bruker-AXS (2000–2003) *SADABS v.2.05, SAINT v.6.22, SHELXTL v.6.10 & SMART 5.622 Software Reference Manuals* (Bruker–AXS, Madison, WI, USA).
7. Panasik N, Jr., Eberhardt ES, Edison AS, Powell DR, Raines RT (1994) Inductive effects on the structure of proline residues. *Int J Pept Protein Res* 44:262–269.
8. DeRider ML, Wilkens SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT, Markley JL (2002) Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. *J Am Chem Soc* 124:2497–2505.
9. Improta R, Benzi C, Barone V (2001) Understanding the role of stereoelectronic effects in determining collagen stability. 1. A quantum mechanical study of proline, hydroxyproline, and fluoroproline dipeptide analogues in aqueous solution. *J Am Chem Soc* 123:12568–12577.

Table S1. Summary of structural parameters for Pro, Hyp, flp, Flp, mep, and Mep

Residue (reference)	Crystal structure	Ring pucker	$(E_{endo} - E_{exo})$[a] (kcal/mol)	ϕ[b] (deg)	ψ[b] (deg)
Mep (1)		C^{γ}-exo	1.7	–62	153
Flp (7, 8)		C^{γ}-exo	0.85	–55	141
Hyp (7, 9)		C^{γ}-exo	—	–57	151
Pro (7, 8)		C^{γ}-endo	–0.41	–79	177
mep (1)	—	C^{γ}-endo	–1.4	—	—
flp (8)	—	C^{γ}-endo	–0.61	–76[a]	172[a]

[a]From density functional theory calculations. [b]Values of ϕ and ψ are from crystal structures of Ac-AA-OMe, unless indicated otherwise.

Table S2. Optimized fitting parameters for DSC analyses

Parameter	$[(ProHypGly)_7]_3$	$[(flpMepGly)_7]_3$	$[(mepFlpGly)_7]_3$
P_{Tot} (μM)	22	51	83
$T°$ (°C)	331.839201	368.955533	393.498713
$\Delta H(T°)$ (kcal/mol)	56.469197	19.6002838	10.9979831
B_s (cal K^{-1} mol^{-1})	–6569.4908	–10896.0047	–4400.00008
D_s (cal mol^{-1})	16.4797743	28.54226	10.0509202
B_t (cal K^{-1} mol^{-1})	–10350.9977	–2420.49211	–999.999932
D_t (cal mol^{-1})	29.2023114	2.8439789	–0.0301716578

Table S3. Crystal Data and Structure Refinement for Crystalline Ac-Mep-OMe

Empirical formula	$C_9H_{15}NO_3$	
Formula weight	185.22	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	$P2_12_12_1$	
Unit cell dimensions	a = 5.990(2) Å	α = 90°
	b = 6.457(3) Å	β = 90°
	c = 25.735(10) Å	γ = 90°
Volume	995.3(7) Å^3	
Z	4	
Density (calculated)	1.236 Mg/m^3	
Absorption coefficient	0.092 mm^{-1}	
F(000)	400	
Crystal size	0.41 × 0.36 × 0.29 mm^3	
Theta range for data collection	1.58 to 24.99°	
Index ranges	$-7 \le h \le 7$, $-7 \le k \le 7$, $-30 \le l \le 30$	
Reflections collected	6859	
Independent reflections	1750 [R(int) = 0.0431]	
Completeness to θ = 24.99°	99.3%	
Absorption correction	Multi-scan with SADABS	
Max. and min. transmission	0.9737 and 0.9631	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	1750 / 0 / 122	
Goodness-of-fit on F^2	1.278	
Final R indices [$I > 2\sigma(I)$]	$R1$ = 0.0779, $wR2$ = 0.2025	
R indices (all data)	$R1$ = 0.0806, $wR2$ = 0.2120	
Absolute structure parameter	N/A	
Largest diff. peak and hole	0.590 and –0.393 e.Å^{-3}	

Table S4. Atomic Coordinates (×10^4) and Equivalent Isotropic Displacement Parameters (Å^2 × 10^3) for Crystalline Ac-Mep-OMe

	x	*y*	*z*	*U*(eq)
O(1)	-1297(6)	3192(6)	1950(2)	22(1)
O(2)	42(6)	3529(6)	718(2)	25(1)
O(3)	-3225(6)	5267(6)	707(2)	22(1)
N(1)	1062(7)	5651(6)	1673(2)	16(1)
C(1)	3050(9)	6977(8)	1717(2)	18(1)
C(2)	2334(9)	8988(8)	1469(2)	17(1)
C(3)	735(9)	8287(9)	1043(2)	21(1)
C(4)	-537(8)	6459(7)	1292(2)	14(1)
C(5)	515(9)	4058(8)	1987(2)	17(1)
C(6)	2297(9)	3423(8)	2376(2)	19(1)
C(7)	4252(9)	10340(8)	1270(2)	22(1)
C(8)	-1180(8)	4853(8)	888(2)	15(1)
C(9)	-3915(9)	4016(9)	269(2)	24(1)

[a] *U*(eq) is defined as one-third of the trace of the orthogonalized U^{ij} tensor.

Table S5. Bond Lengths [Å] and Angles [°] for Crystalline Ac-Mep-OMe

O(1)-C(5)	1.225(6)
O(2)-C(8)	1.208(6)
O(3)-C(8)	1.338(6)
O(3)-C(9)	1.447(6)
N(1)-C(5)	1.347(7)
N(1)-C(4)	1.466(6)
N(1)-C(1)	1.471(6)
C(1)-C(2)	1.509(7)
C(1)-H(1A)	0.9900
C(1)-H(1B)	0.9900
C(2)-C(3)	1.523(7)
C(2)-C(7)	1.531(7)
C(2)-H(2)	1.0000
C(3)-C(4)	1.545(7)
C(3)-H(3A)	0.9900
C(3)-H(3B)	0.9900
C(4)-C(8)	1.518(7)
C(4)-H(4)	1.0000
C(5)-C(6)	1.521(7)
C(6)-H(6A)	0.9800
C(6)-H(6B)	0.9800
C(6)-H(6C)	0.9800
C(7)-H(7A)	0.9800
C(7)-H(7B)	0.9800
C(7)-H(7C)	0.9800
C(9)-H(9A)	0.9800
C(9)-H(9B)	0.9800
C(9)-H(9C)	0.9800
C(8)-O(3)-C(9)	114.9(4)
C(5)-N(1)-C(4)	120.9(4)
C(5)-N(1)-C(1)	126.5(4)
C(4)-N(1)-C(1)	111.9(4)
N(1)-C(1)-C(2)	103.8(4)
N(1)-C(1)-H(1A)	111.0
C(2)-C(1)-H(1A)	111.0
N(1)-C(1)-H(1B)	111.0
C(2)-C(1)-H(1B)	111.0
H(1A)-C(1)-H(1B)	109.0
C(1)-C(2)-C(3)	103.1(4)
C(1)-C(2)-C(7)	114.7(4)
C(3)-C(2)-C(7)	113.6(4)
C(1)-C(2)-H(2)	108.4
C(3)-C(2)-H(2)	108.4
C(7)-C(2)-H(2)	108.4
C(2)-C(3)-C(4)	103.8(4)
C(2)-C(3)-H(3A)	111.0
C(4)-C(3)-H(3A)	111.0
C(2)-C(3)-H(3B)	111.0
C(4)-C(3)-H(3B)	111.0
H(3A)-C(3)-H(3B)	109.0
N(1)-C(4)-C(8)	112.4(4)
N(1)-C(4)-C(3)	103.1(4)

C(8)-C(4)-C(3)	111.2(4)
N(1)-C(4)-H(4)	110.0
C(8)-C(4)-H(4)	110.0
C(3)-C(4)-H(4)	110.0
O(1)-C(5)-N(1)	121.2(5)
O(1)-C(5)-C(6)	123.3(5)
N(1)-C(5)-C(6)	115.5(4)
C(5)-C(6)-H(6A)	109.5
C(5)-C(6)-H(6B)	109.5
H(6A)-C(6)-H(6B)	109.5
C(5)-C(6)-H(6C)	109.5
H(6A)-C(6)-H(6C)	109.5
H(6B)-C(6)-H(6C)	109.5
C(2)-C(7)-H(7A)	109.5
C(2)-C(7)-H(7B)	109.5
H(7A)-C(7)-H(7B)	109.5
C(2)-C(7)-H(7C)	109.5
H(7A)-C(7)-H(7C)	109.5
H(7B)-C(7)-H(7C)	109.5
O(2)-C(8)-O(3)	124.7(5)
O(2)-C(8)-C(4)	125.3(4)
O(3)-C(8)-C(4)	109.5(4)
O(3)-C(9)-H(9A)	109.5
O(3)-C(9)-H(9B)	109.5
H(9A)-C(9)-H(9B)	109.5
O(3)-C(9)-H(9C)	109.5
H(9A)-C(9)-H(9C)	109.5
H(9B)-C(9)-H(9C)	109.5

Table S6. Anisotropic Displacement Parameters (Å^2 × 10^3) for Crystalline Ac-Mep-OMe[a]

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
O(1)	11(2)	16(2)	38(2)	-1(2)	-1(2)	-2(2)
O(2)	18(2)	21(2)	35(2)	-10(2)	-4(2)	9(2)
O(3)	13(2)	23(2)	29(2)	-5(2)	-5(2)	6(2)
N(1)	11(2)	11(2)	24(2)	-2(2)	-3(2)	-1(2)
C(1)	11(3)	15(3)	29(3)	5(2)	-2(2)	-2(2)
C(2)	15(3)	9(2)	27(3)	-1(2)	1(2)	0(2)
C(3)	15(3)	17(3)	30(3)	2(2)	-1(2)	-1(2)
C(4)	12(2)	6(2)	25(2)	1(2)	-1(2)	0(2)
C(5)	18(3)	9(2)	23(3)	-2(2)	2(2)	2(2)
C(6)	20(3)	12(3)	26(3)	4(2)	-1(2)	0(2)
C(7)	16(3)	10(2)	39(3)	2(2)	3(2)	-5(2)
C(8)	9(2)	12(2)	24(3)	4(2)	-1(2)	-2(2)
C(9)	13(3)	29(3)	30(3)	-2(2)	-6(2)	0(3)

[a]The anisotropic displacement factor exponent takes the form: $-2p^2[h^2\, a^{*2} U^{11} + ... + 2\, h\, k\, a^*\, b^*\, U^{12}]$.

Table S7. Hydrogen Coordinates (×10^4) and Isotropic Displacement Parameters (Å^2 × 10^3) for Crystalline Ac-Mep-OMe

	x	*y*	*z*	*U*(eq)
H(1A)	4332	6363	1529	22
H(1B)	3469	7188	2085	22
H(2)	1468	9803	1730	20
H(3A)	1562	7833	730	25
H(3B)	-302	9416	945	25
H(4)	-1903	6978	1474	17
H(6A)	2259	4365	2675	29
H(6B)	3769	3491	2211	29
H(6C)	2011	2004	2494	29
H(7A)	5258	10672	1558	32
H(7B)	3648	11623	1123	32
H(7C)	5078	9588	1000	32
H(9A)	-3008	4371	-35	36
H(9B)	-5492	4283	193	36
H(9C)	-3712	2548	353	36
H(16C)	4278	12159	9339	43

Table S8. Torsion angles [°] for crystalline Ac-Mep-OMe

C(5)-N(1)-C(1)-C(2)	-153.6(5)
C(4)-N(1)-C(1)-C(2)	16.7(6)
N(1)-C(1)-C(2)-C(3)	-33.3(5)
N(1)-C(1)-C(2)-C(7)	-157.4(4)
C(1)-C(2)-C(3)-C(4)	37.8(5)
C(7)-C(2)-C(3)-C(4)	162.6(4)
C(5)-N(1)-C(4)-C(8)	-62.4(6)
C(1)-N(1)-C(4)-C(8)	126.7(4)
C(5)-N(1)-C(4)-C(3)	177.7(4)
C(1)-N(1)-C(4)-C(3)	6.8(5)
C(2)-C(3)-C(4)-N(1)	-27.4(5)
C(2)-C(3)-C(4)-C(8)	-148.1(4)
C(4)-N(1)-C(5)-O(1)	2.7(7)
C(1)-N(1)-C(5)-O(1)	172.1(5)
C(4)-N(1)-C(5)-C(6)	-177.9(4)
C(1)-N(1)-C(5)-C(6)	-8.5(7)
C(9)-O(3)-C(8)-O(2)	-1.2(7)
C(9)-O(3)-C(8)-C(4)	172.1(4)
N(1)-C(4)-C(8)-O(2)	-34.0(7)
C(3)-C(4)-C(8)-O(2)	81.0(6)
N(1)-C(4)-C(8)-O(3)	152.8(4)
C(3)-C(4)-C(8)-O(3)	-92.2(5)

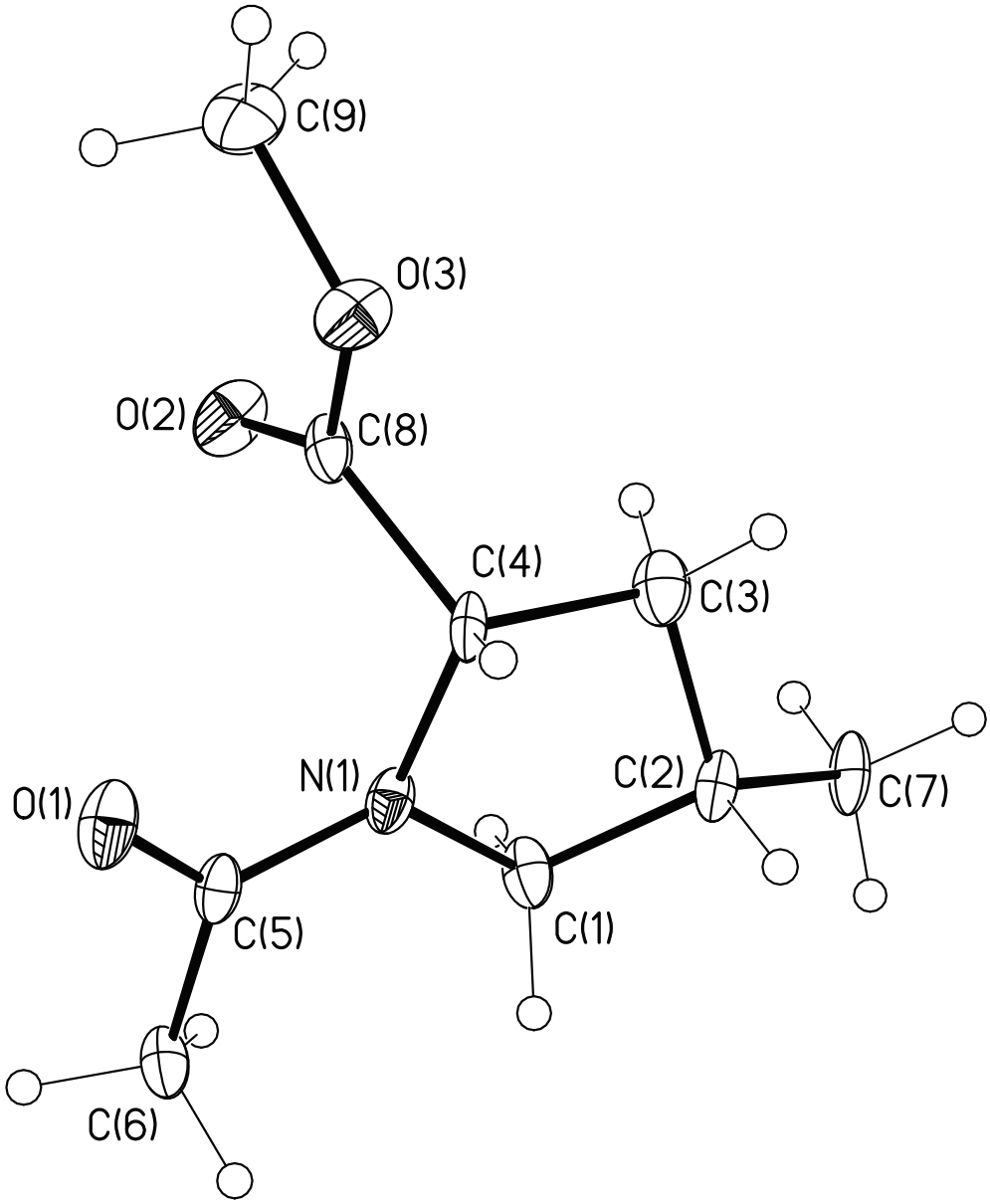


Fig. S1. Molecular drawing of the asymmetric unit of crystalline Ac-Mep-OMe depicted with 50% probability ellipsoids.

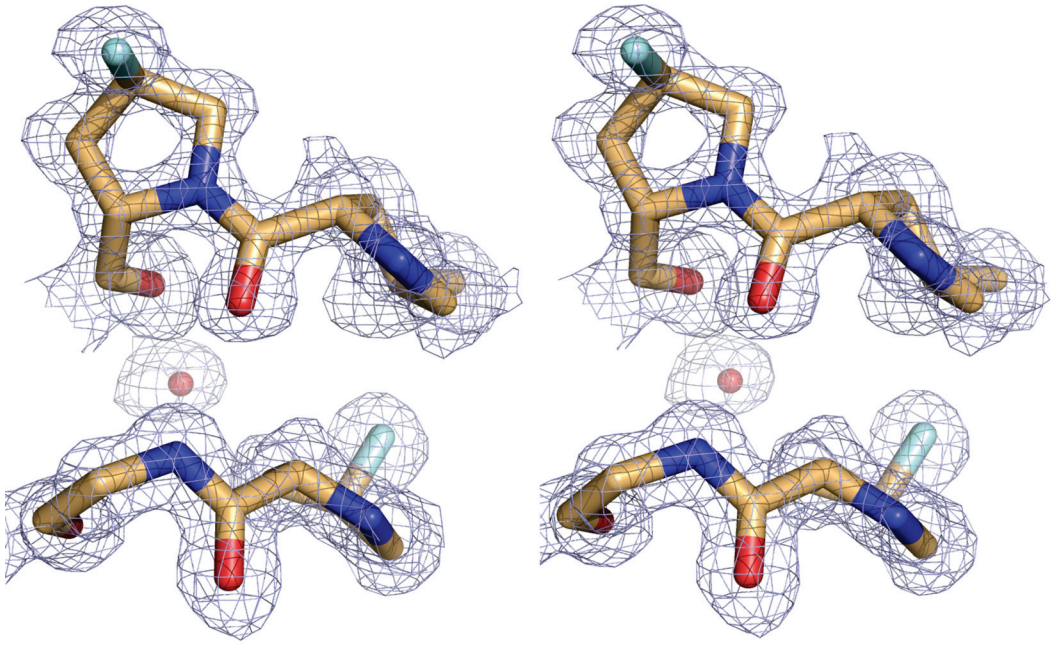
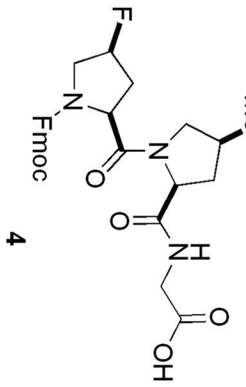
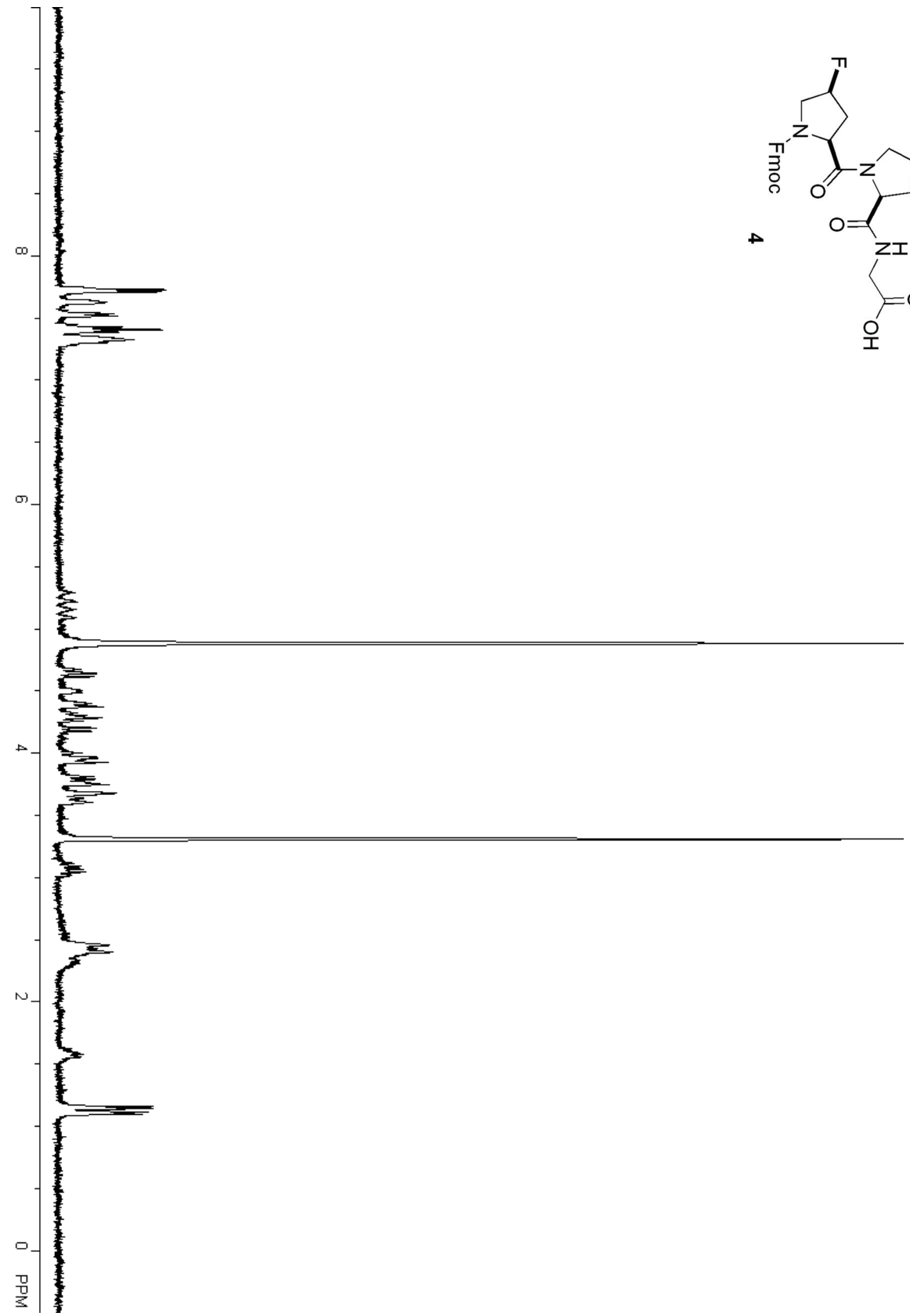


Fig. S2. Central portion of final refined $[(mepFlpGly)_7]_3$ model (wheat C, blue N, red O, pale cyan F) shown in stereo with $2F_o–F_c$ electron density map contoured at 1.0 standard deviation above the mean. Two adjacent residues from each of two neighboring strands (E10, E11, F11, and F12) and a nearby water molecule are depicted.

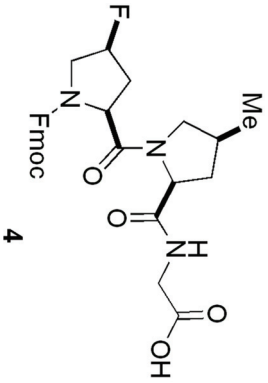
400 MHz ^{1}H NMR Spectrum of Compound **4**



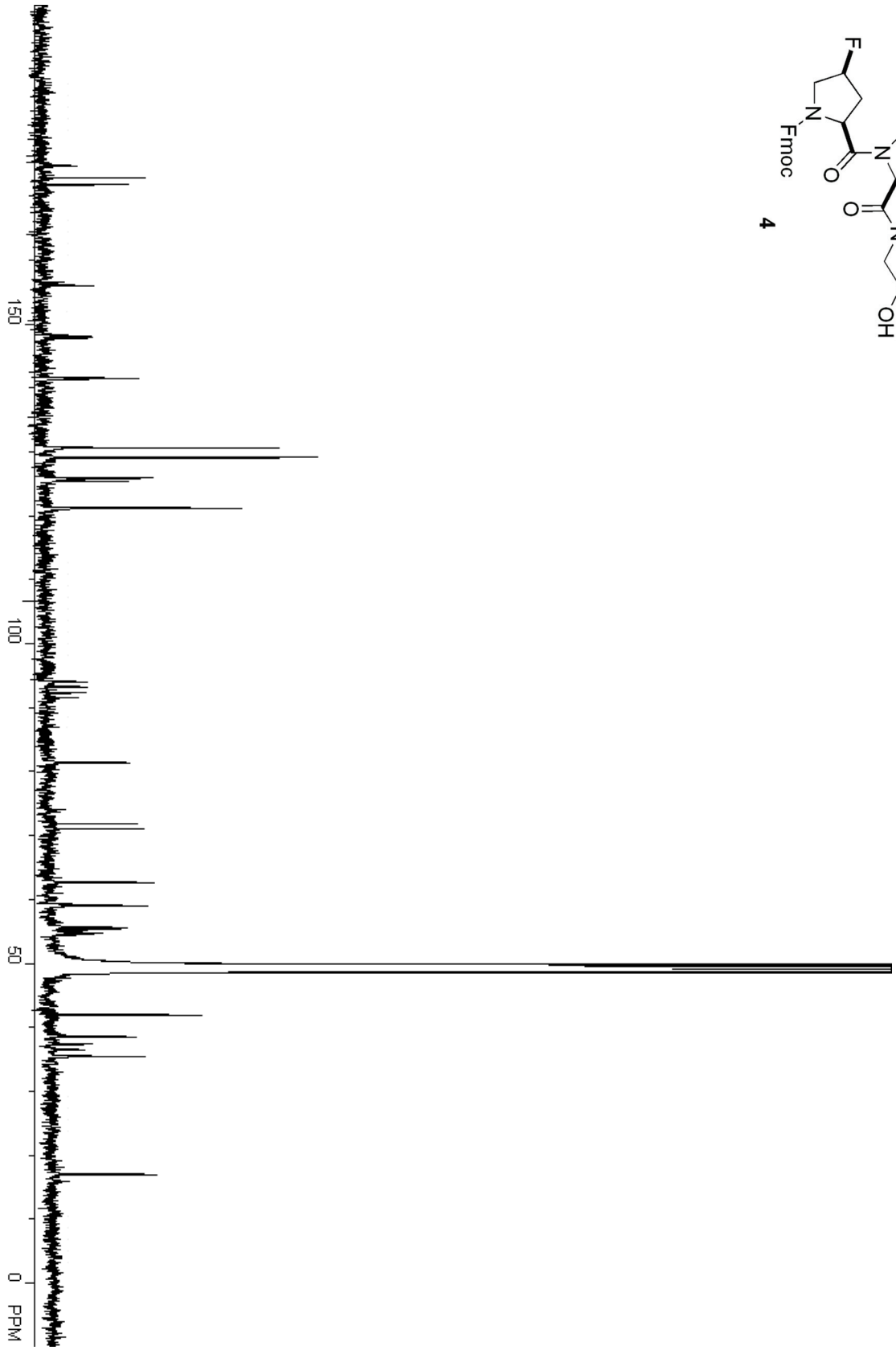
4



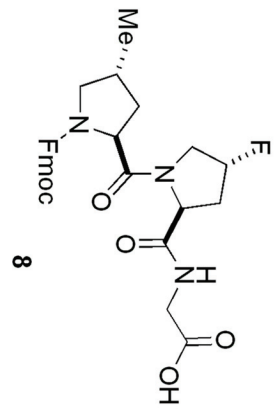
125 MHz ^{13}C NMR Spectrum of Compound **4**



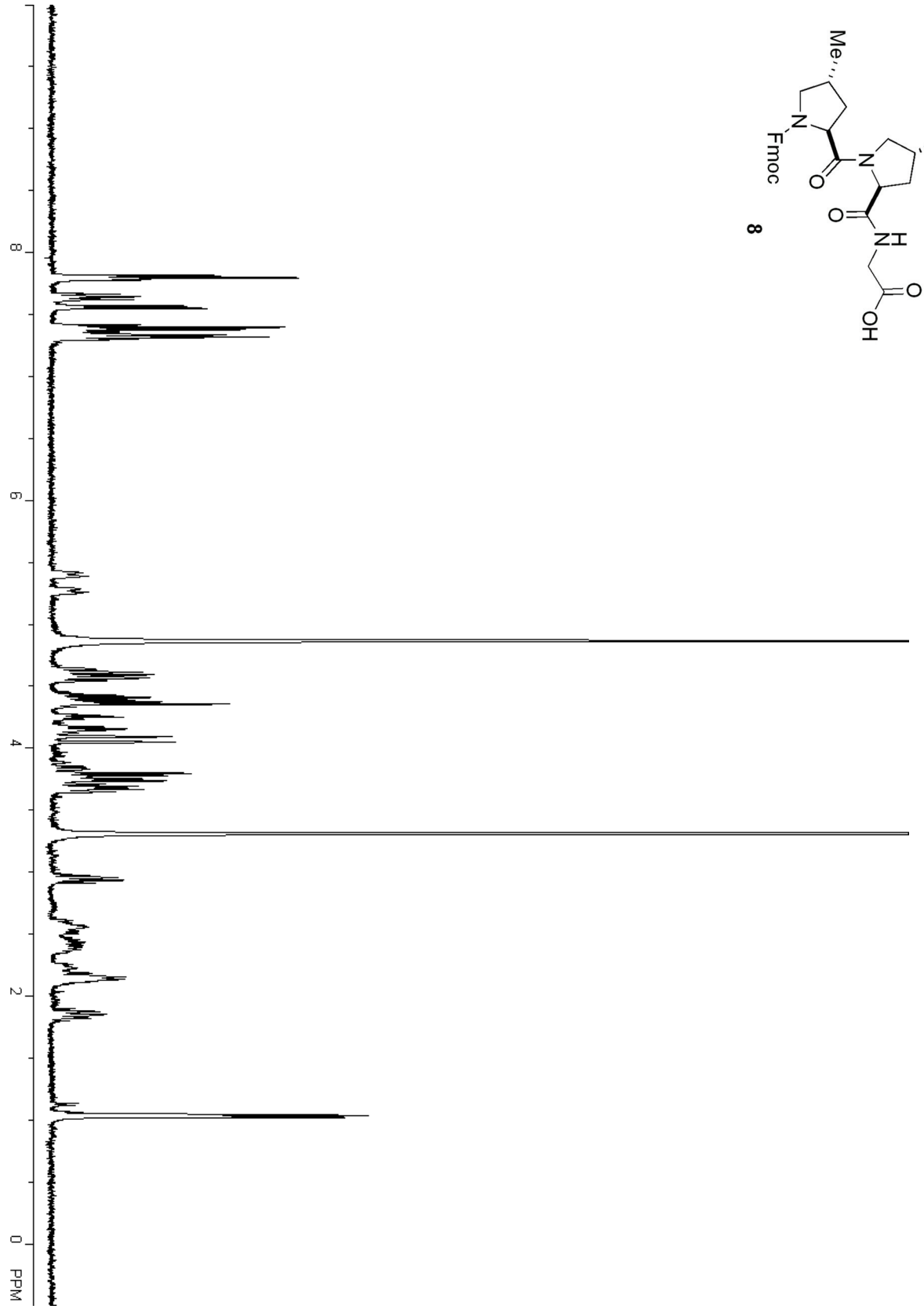
4



400 MHz ^{1}H NMR Spectrum of Compound **8**



8



125 MHz ^{13}C NMR Spectrum of Compound **8**