

Supporting Information

PEGylation near a patch of non-polar surface residues increases the conformational stability of the WW domain

Steven R. E. Draper, Dallin S. Ashton, Benjamin M. Conover, Anthony J. Carter, Kimberlee L. Stern, Qiang Xiao and Joshua L. Price*

Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, United States

jlprice@chem.byu.edu

1. ESI-TOF MS Spectra.....	S2
2. Analytical HPLC Chromatograms.....	S10
3. Variable Temperature CD Experiments	S34

1. ESI-TOF MS spectra

Sequences, molecular formulas, and expected and observed m/z ratios for each WW variant are given in Table S1.

ESI-TOF MS spectra for each variant appear in Figures S1–S24.

Table S1. Sequences, molecular formulas, expected and observed m/z ratios for WW variants.

Protein	Sequence	Molecular Formula	z	Expected [M+z·H]/z	Observed [M+z·H]/z
LL	H ₂ N-KLPPGWL LK LMSRSSGRV Z YFNHITNASQFERPSG-OH	C ₁₈₁ H ₂₇₃ N ₅₁ O ₄₈ S	4	991.2633	991.2594
LA	H ₂ N-..... LK A M SRSSGRV Z-OH	C ₁₇₈ H ₂₆₇ N ₅₁ O ₄₈ S	4	980.7516	980.7500
AL	H ₂ N-..... AK L M SRSSGRV Z-OH	C ₁₇₈ H ₂₆₇ N ₅₁ O ₄₈ S	4	980.7516	980.7511
AA	H ₂ N-..... AK A M SRSSGRV Z-OH	---	---	---	---
pLL	H ₂ N-..... LK L M SRSSGRV Z-OH	C ₁₉₀ H ₂₉₂ N ₅₄ O ₅₂ S	4	1049.5477	1049.5450
pLA	H ₂ N-..... LK A M SRSSGRV Z-OH	C ₁₈₇ H ₂₈₆ N ₅₄ O ₅₂ S	4	1039.0360	1039.0326
pAL	H ₂ N-..... AK L M SRSSGRV Z-OH	C ₁₈₇ H ₂₈₆ N ₅₄ O ₅₂ S	4	1039.0360	1039.0357
pAA	H ₂ N-..... AK A M SRSSGRV Z-OH	---	---	---	---
FF	H ₂ N-..... FK F M SRSSGRV Z-OH	C ₁₈₇ H ₂₆₉ N ₅₁ O ₄₈ S	4	1008.2555	1008.2572
FA	H ₂ N-..... FK A M SRSSGRV Z-OH	C ₁₈₁ H ₂₆₅ N ₅₁ O ₄₈ S	4	989.2476	989.2472
AF	H ₂ N-..... AK F M SRSSGRV Z-OH	C ₁₈₁ H ₂₆₅ N ₅₁ O ₄₈ S	4	989.2476	989.2480
pFF	H ₂ N-..... FK F M SRSSGRV Z-OH	C ₁₉₆ H ₂₈₈ N ₅₄ O ₅₂ S	4	1066.5399	1066.5338
pFA	H ₂ N-..... FK A M SRSSGRV Z-OH	C ₁₉₀ H ₂₈₄ N ₅₄ O ₅₂ S	4	1047.5321	1047.5375
pAF	H ₂ N-..... AK F M SRSSGRV Z-OH	C ₁₉₀ H ₂₈₄ N ₅₄ O ₅₂ S	4	1047.5321	1047.5337
XX	H ₂ N-..... XK X M SRSSGRV Z-OH	C ₁₈₇ H ₂₈₁ N ₅₁ O ₄₈ S	4	1011.2790	1011.2825
XA	H ₂ N-..... XK A M SRSSGRV Z-OH	C ₁₈₁ H ₂₇₁ N ₅₁ O ₄₈ S	4	990.7594	990.7580
AX	H ₂ N-..... AK X M SRSSGRV Z-OH	C ₁₈₁ H ₂₇₁ N ₅₁ O ₄₈ S	4	990.7594	990.7629
pXX	H ₂ N-..... XK X M SRSSGRV Z-OH	C ₁₉₆ H ₃₀₀ N ₅₄ O ₅₂ S	4	1069.5634	1069.5557
pXA	H ₂ N-..... XK A M SRSSGRV Z-OH	C ₁₉₀ H ₂₉₀ N ₅₄ O ₅₂ S	4	1049.0438	1049.0425
pAX	H ₂ N-..... AK X M SRSSGRV Z-OH	C ₁₉₀ H ₂₉₀ N ₅₄ O ₅₂ S	4	1049.0438	1049.0381
XL	H ₂ N-..... XK L M SRSSGRV Z-OH	C ₁₈₄ H ₂₇₇ N ₅₁ O ₄₈ S	4	1001.2711	1001.2705
FL	H ₂ N-..... FK L M SRSSGRV Z-OH	C ₁₈₄ H ₂₇₁ N ₅₁ O ₄₈ S	4	999.7594	999.7586
LF	H ₂ N-..... LK F M SRSSGRV Z-OH	C ₁₈₄ H ₂₇₁ N ₅₁ O ₄₈ S	4	999.7594	999.7577
pXL	H ₂ N-..... XK L M SRSSGRV Z-OH	C ₁₉₃ H ₂₉₆ N ₅₄ O ₅₂ S	4	1059.5555	1059.5534
pFL	H ₂ N-..... FK L M SRSSGRV Z-OH	C ₁₉₃ H ₂₉₀ N ₅₄ O ₅₂ S	4	1058.0438	1058.0419
pLF	H ₂ N-..... LK F M SRSSGRV Z-OH	C ₁₉₃ H ₂₉₀ N ₅₄ O ₅₂ S	4	1058.0438	1058.0412

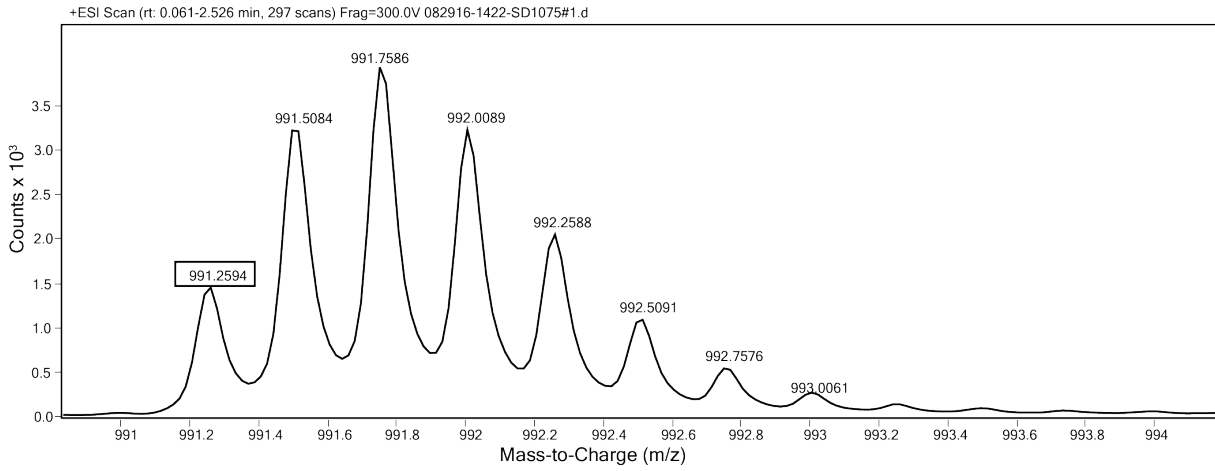


Figure S1. ESI-TOF MS data for LL (SD1075#1). Expected $[M+4H^+]/4 = 991.2633$.

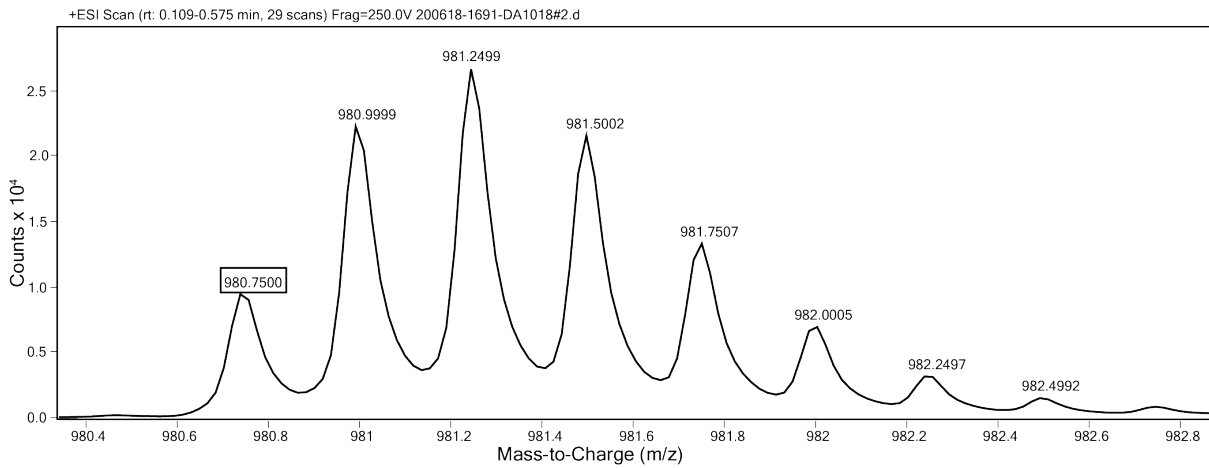


Figure S2. ESI-TOF MS data for LA (DA1018#2). Expected $[M+4H^+]/4 = 980.7516$.

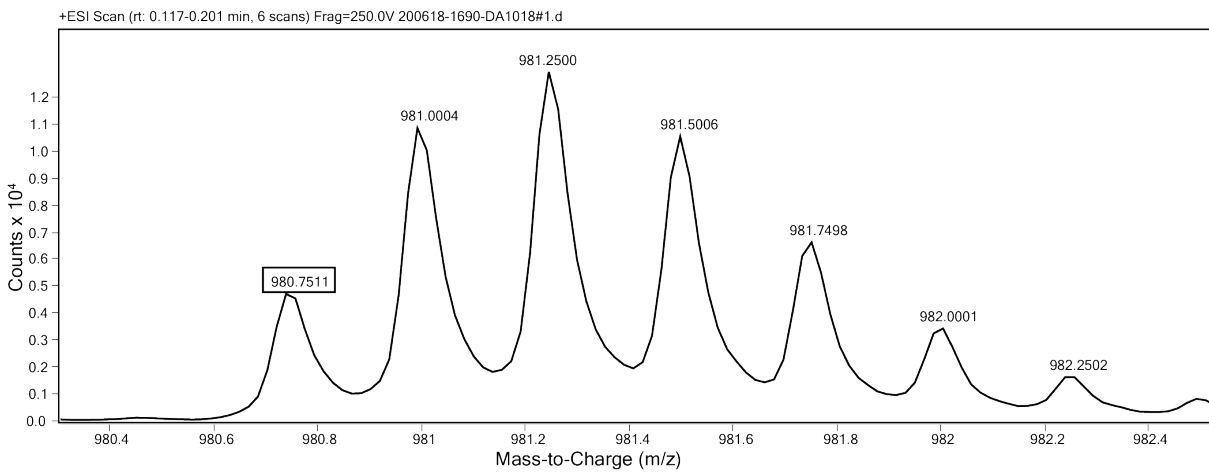


Figure S3. ESI-TOF MS data for AL (DA1018#1). Expected $[M+4H^+]/4 = 980.7516$

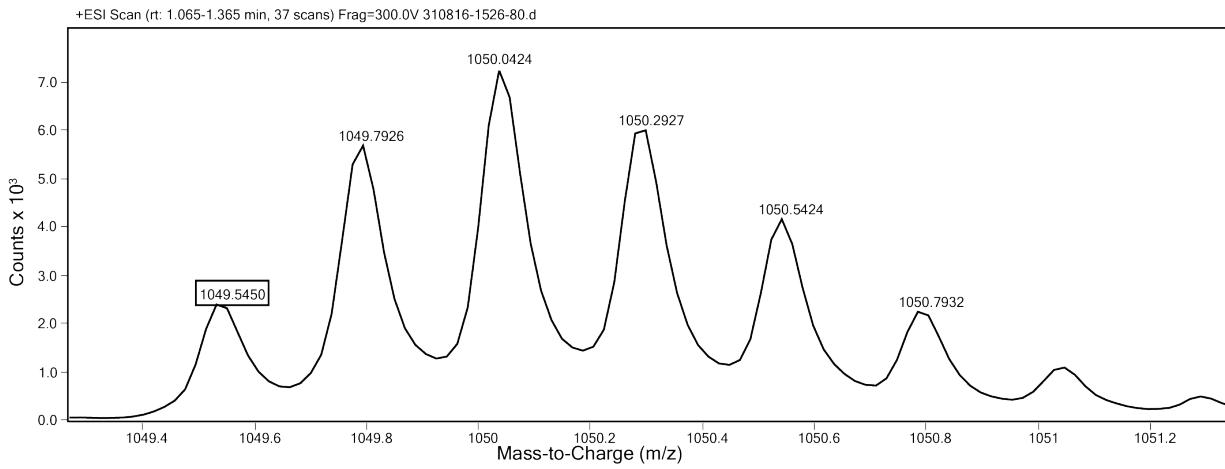


Figure S4. ESI-TOF MS data for **pLL** (SD1075#1C). Expected $[M+4H^+]/4 = 1049.5477$.

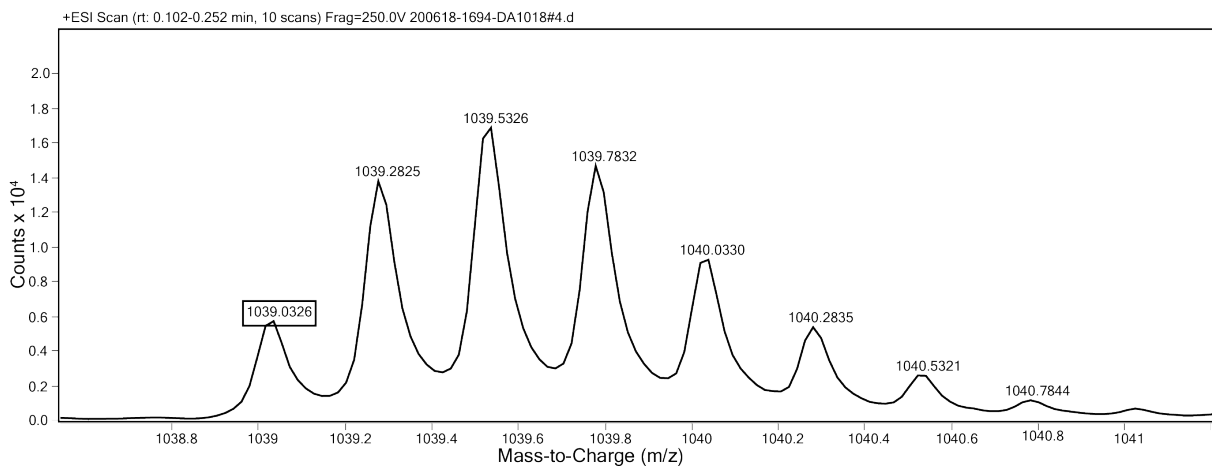


Figure S5. ESI-TOF MS data for **pLA** (DA1018#4). Expected $[M+4H^+]/4 = 1039.0360$.

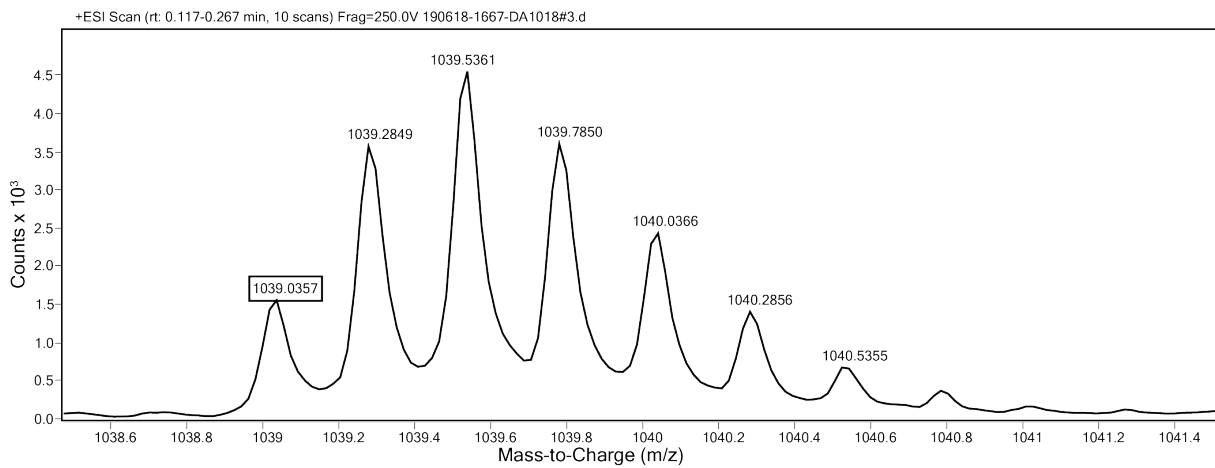


Figure S6. ESI-TOF MS data for **pAL** (DA1018#3). Expected $[M+4H^+]/4 = 1039.0360$.

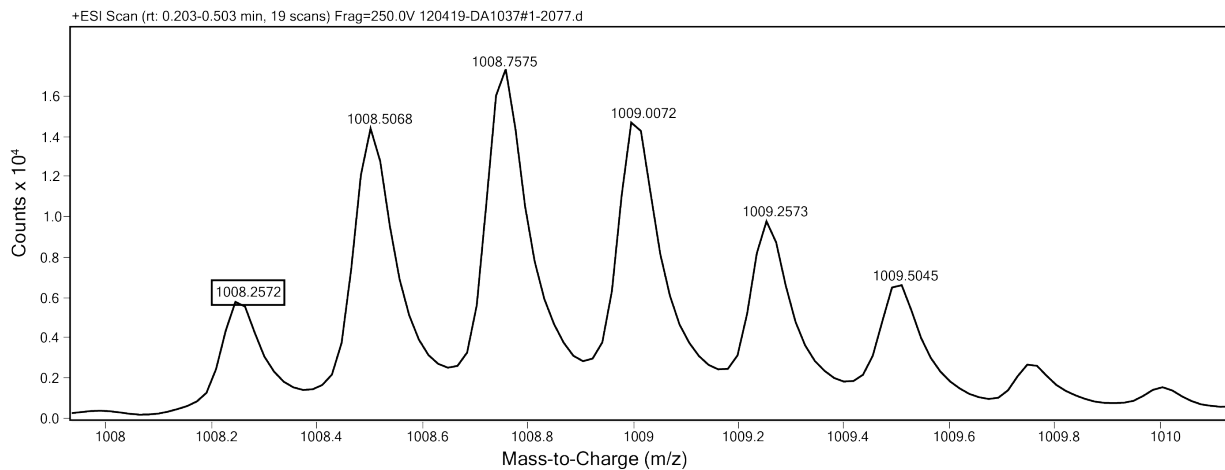


Figure S7. ESI-TOF MS data for **FF** (DA1037#1). Expected $[M+4H^+]/4 = 1008.2555$.

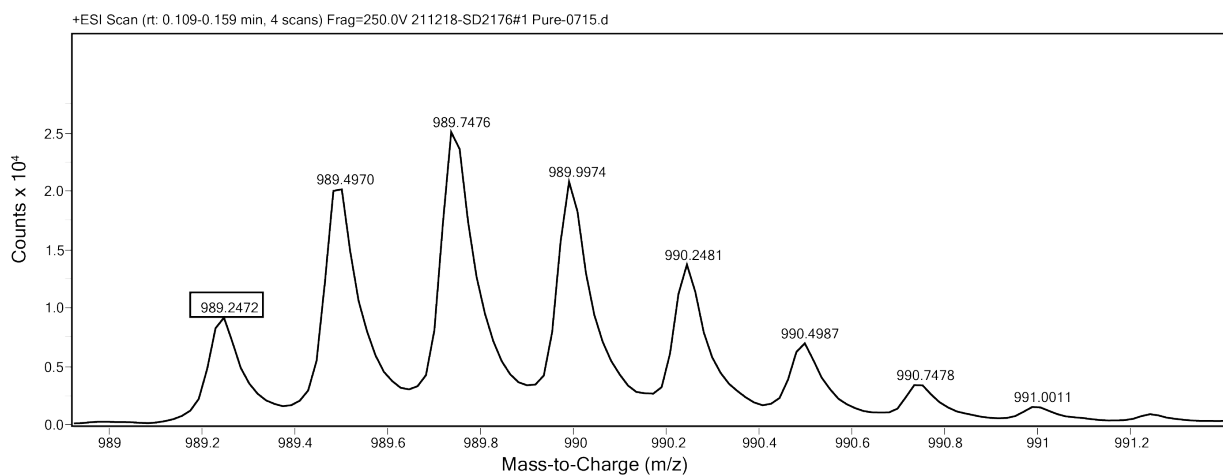


Figure S8. ESI-TOF MS data for **FA** (SD2176#1). Expected $[M+4H^+]/4 = 989.2476$.

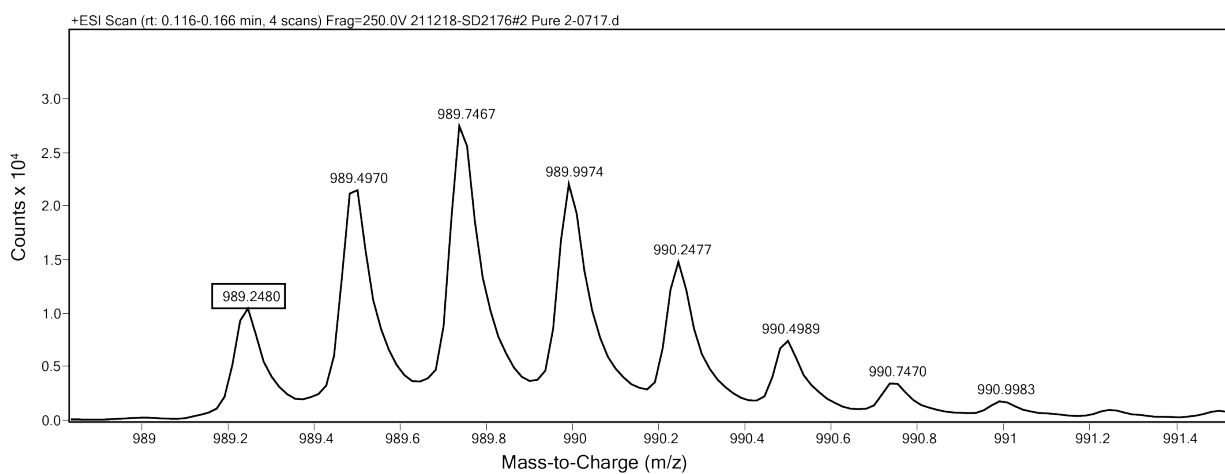


Figure S9. ESI-TOF MS data for **AF** (DS2176#2). Expected $[M+4H^+]/4 = 989.2476$.

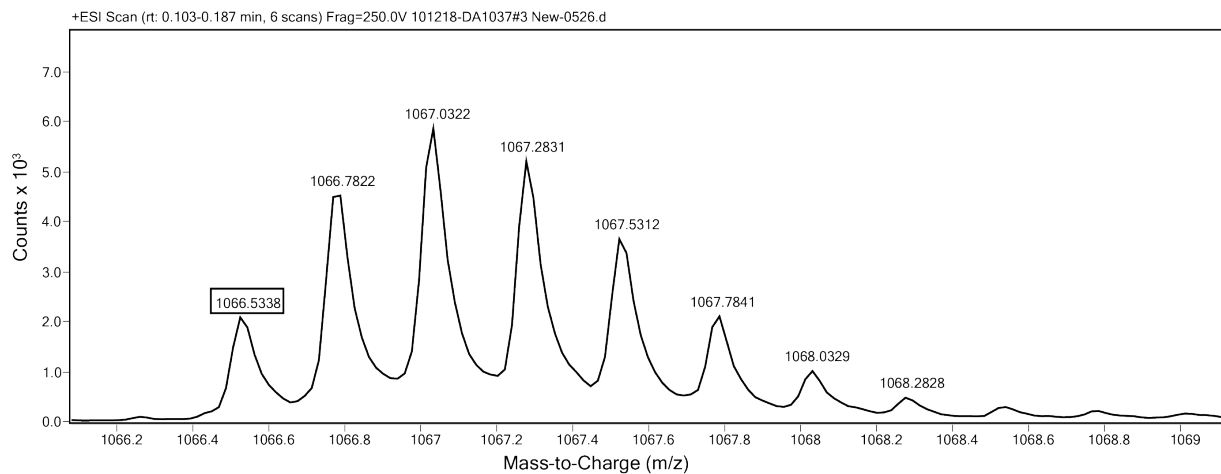


Figure S10. ESI-TOF MS data for **pFF** (DA1037#3). Expected $[M+4H^+]/4 = 1066.5399$.

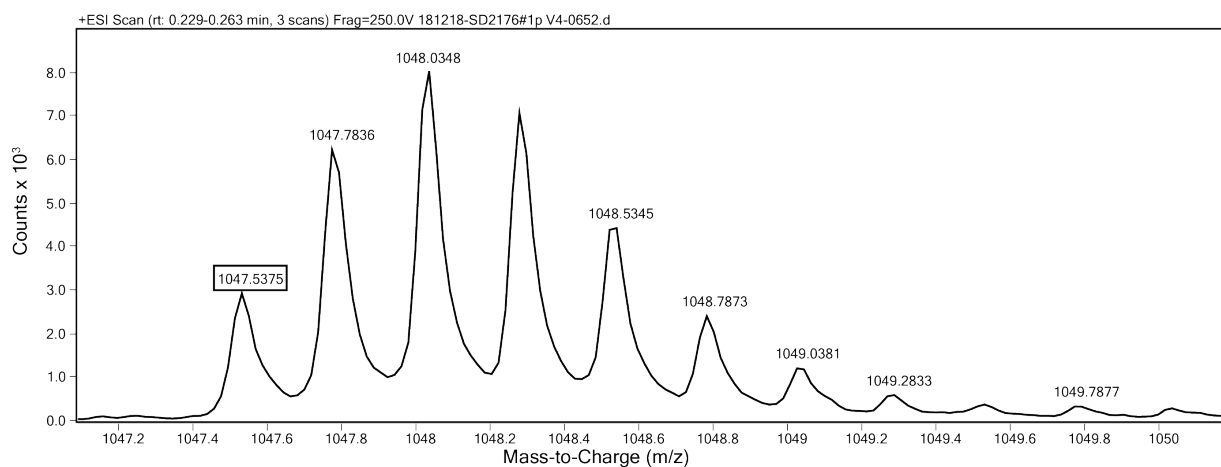


Figure S11. ESI-TOF MS data for **pFA** (SD2176#1p). Expected $[M+4H^+]/4 = 1047.5321$.

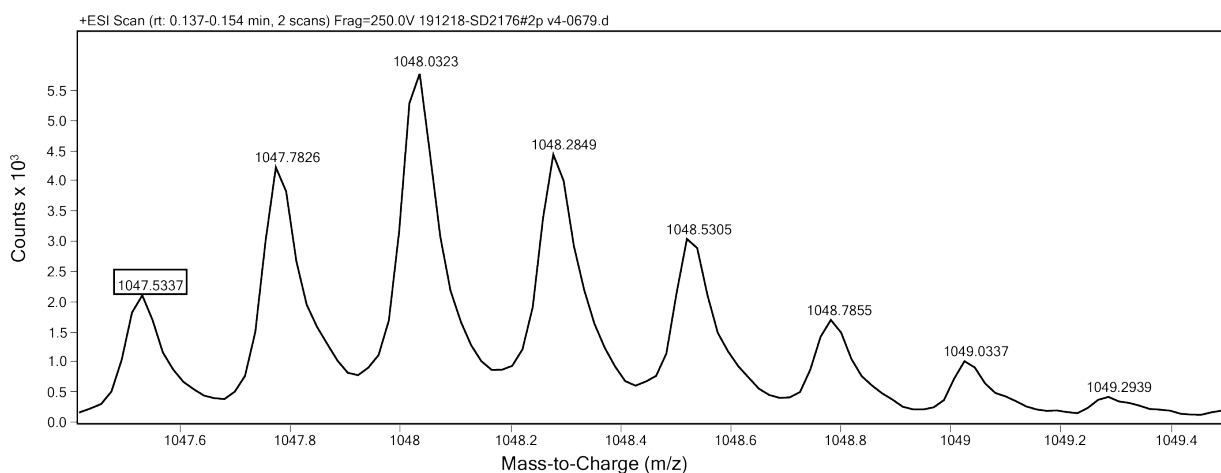


Figure S12. ESI-TOF MS data for **pAF** (SD2176#2p). Expected $[M+4H^+]/4 = 1047.5321$.

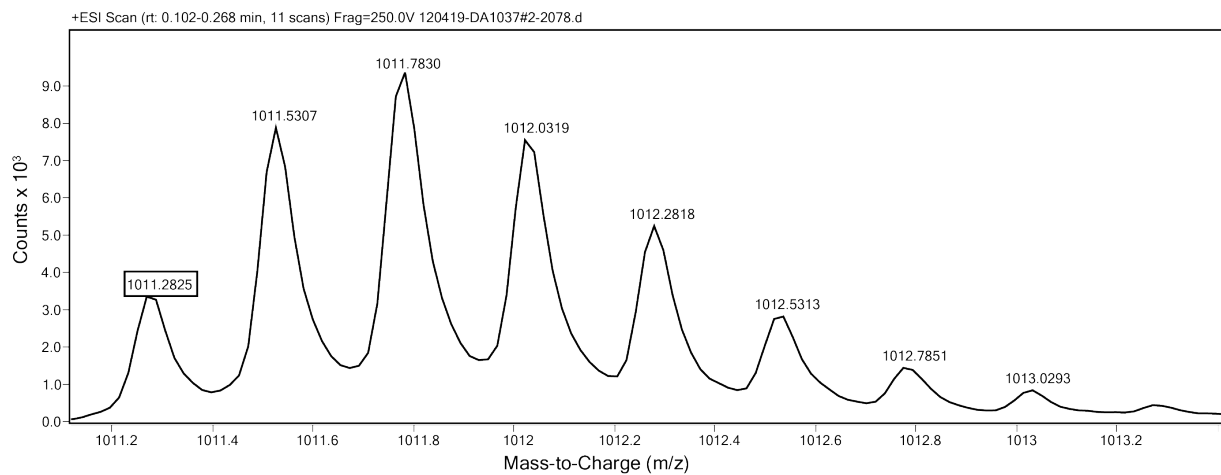


Figure S13. ESI-TOF MS data for **XX** (DA1037#2). Expected $[M+4H^+]/4 = 1011.2790$.

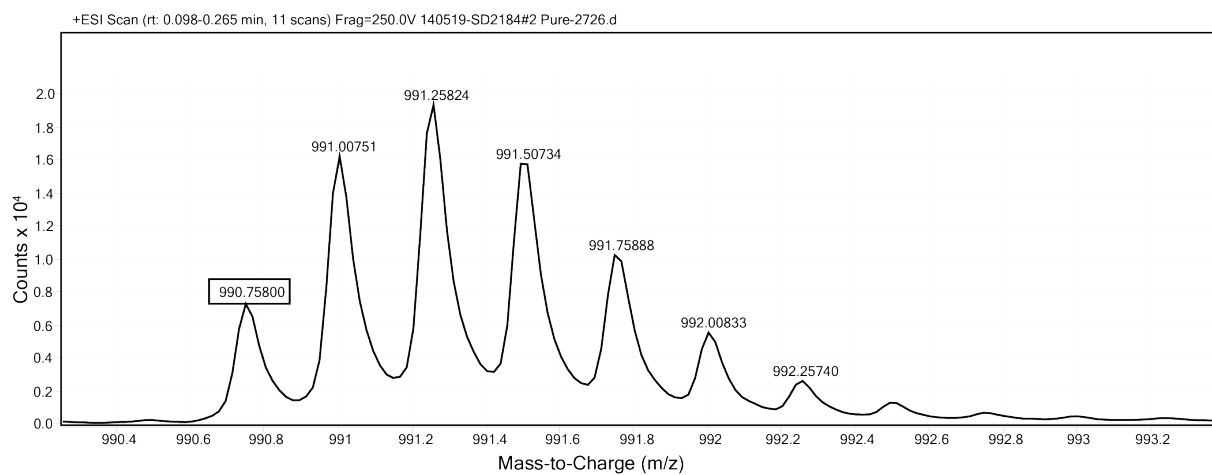


Figure S14. ESI-TOF MS data for **XA** (SD2184#2). Expected $[M+4H^+]/4 = 990.7594$.

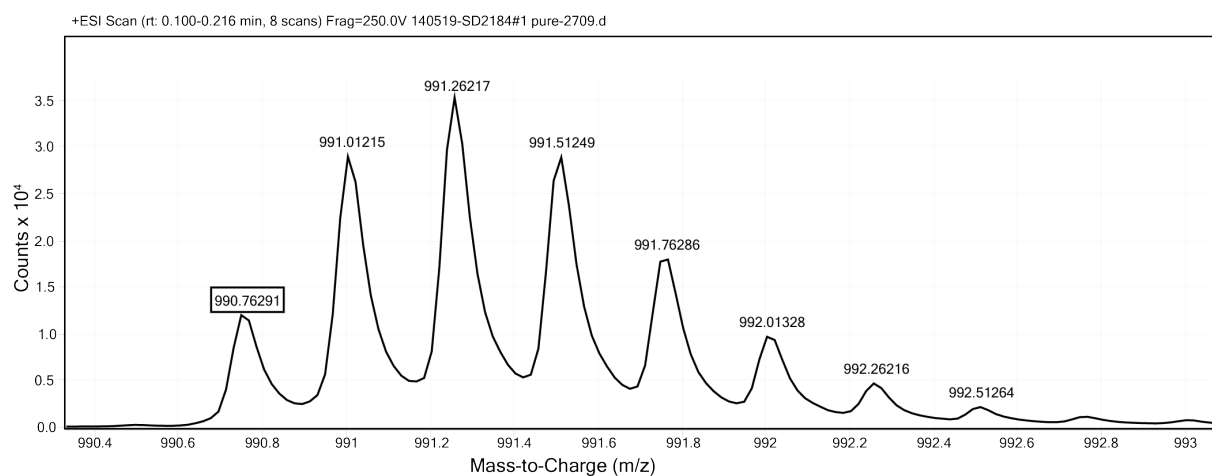


Figure S15. ESI-TOF MS data for **AX** (SD2184#1). Expected $[M+4H^+]/4 = 990.7594$.

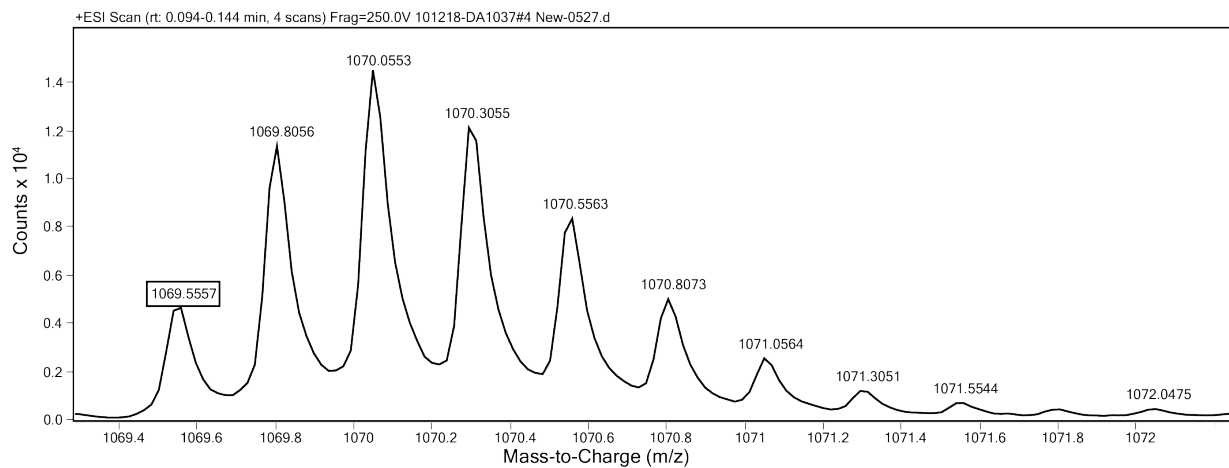


Figure S16. ESI-TOF MS data for pXX (DA1037#4). Expected $[M+4H^+]/4 = 1069.5634$.

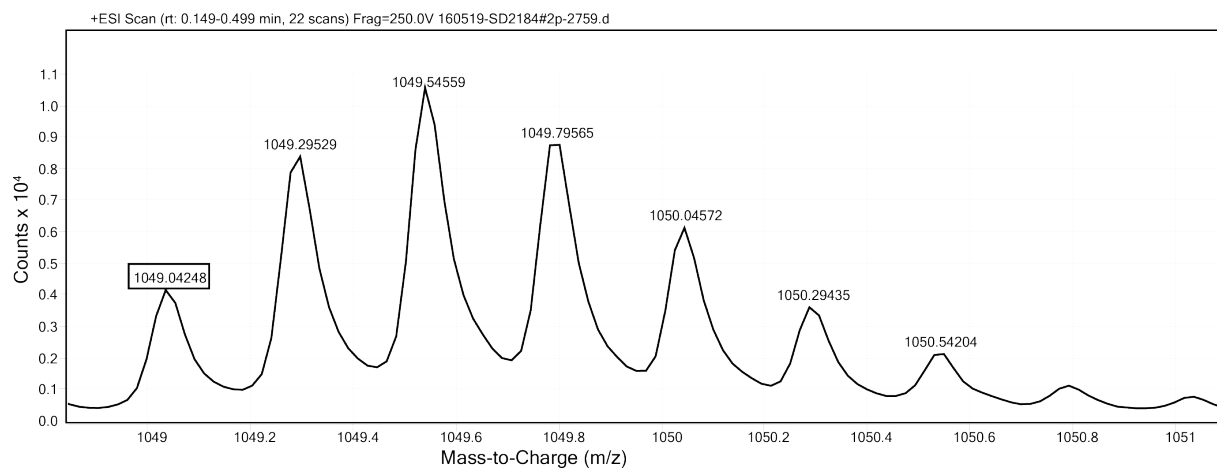


Figure S17. ESI-TOF MS data for pXA (SD2184#2p). Expected $[M+4H^+]/4 = 1049.0438$.

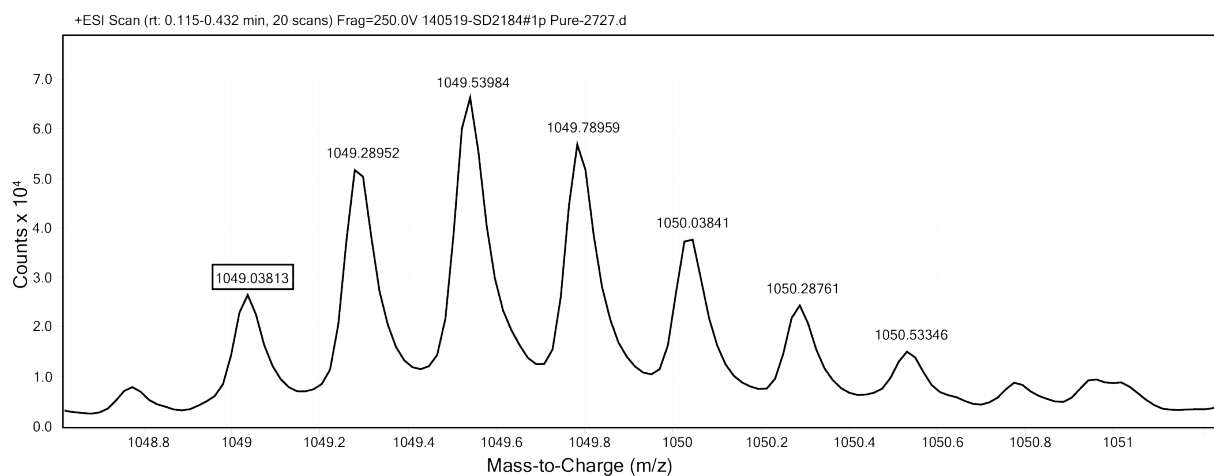


Figure S18. ESI-TOF MS data for pAX (SD2184#1p). Expected $[M+4H^+]/4 = 1049.0438$.

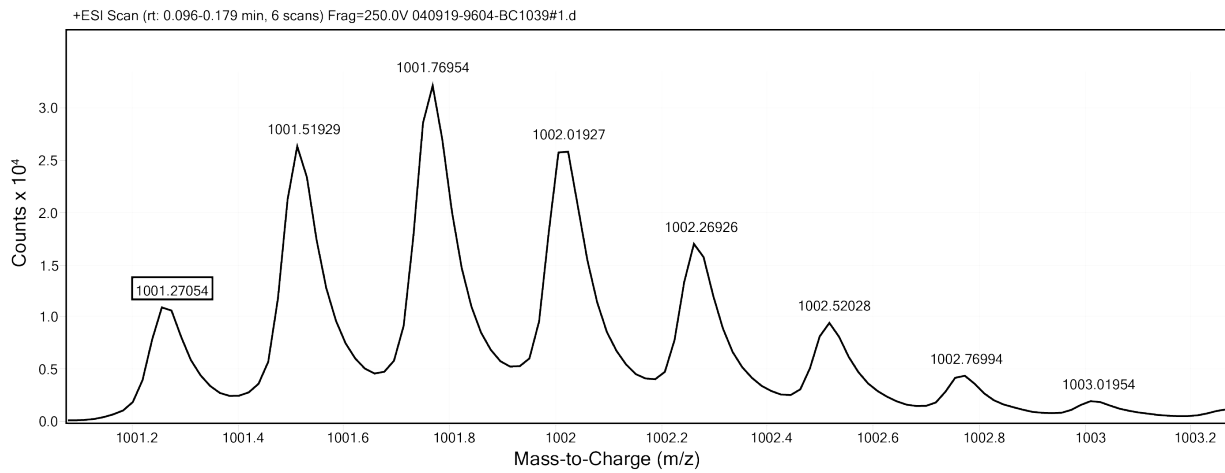


Figure S19. ESI-TOF MS data for **XL** (BC1039#1). Expected $[M+4H^+]/4 = 1001.2711$.

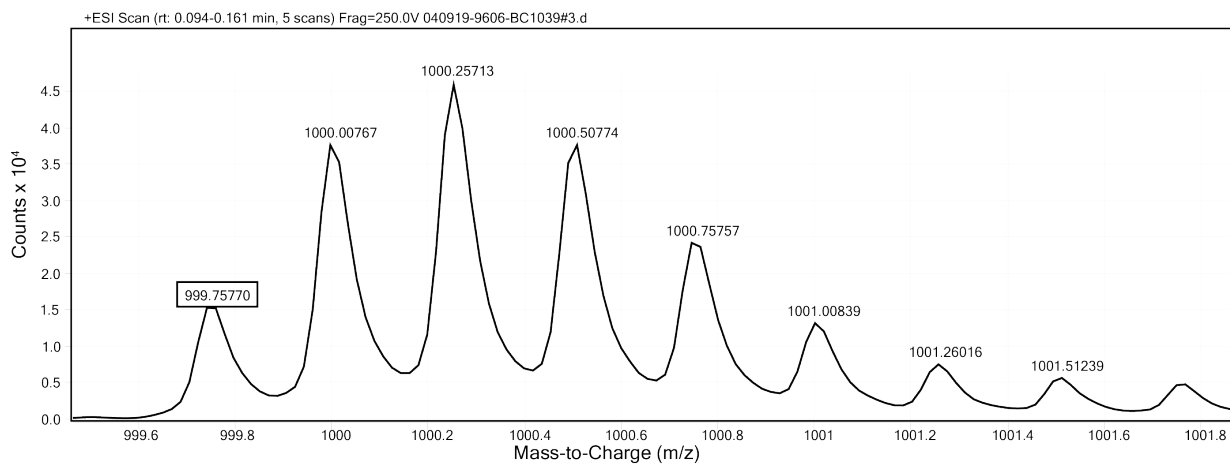


Figure S20. ESI-TOF MS data for **LF** (BC1039#3). Expected $[M+4H^+]/4 = 999.7594$.

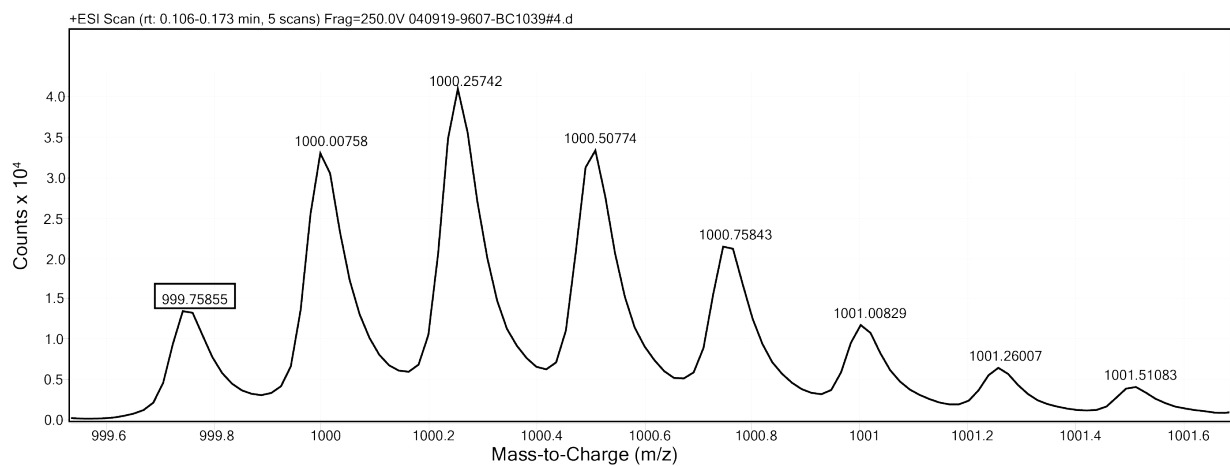


Figure S21. ESI-TOF MS data for **FL** (BC1039#4). Expected $[M+4H^+]/4 = 999.7594$.

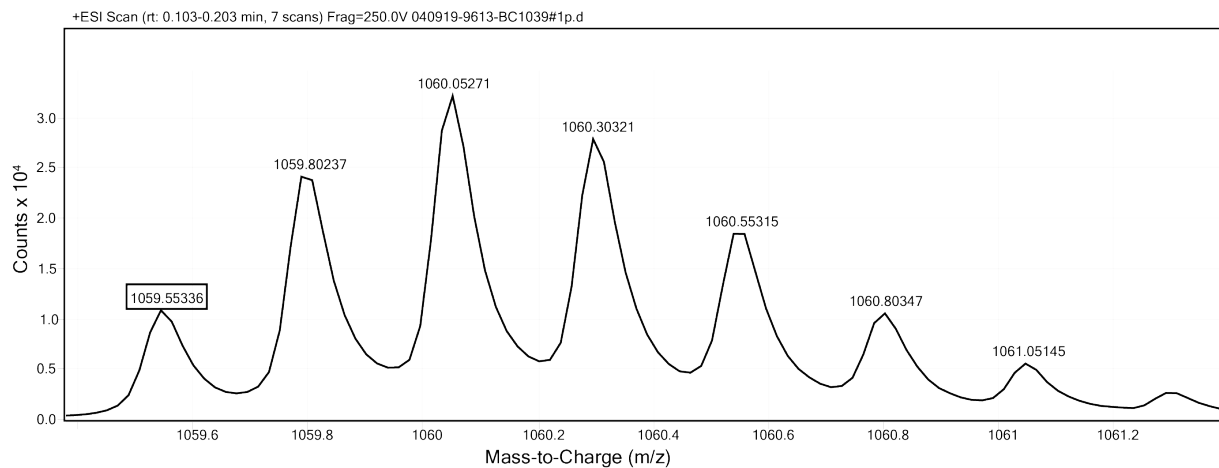


Figure S22. ESI-TOF MS data for **pXL** (BC1039#1p). Expected $[M+4H^+]/4 = 1059.5555$.

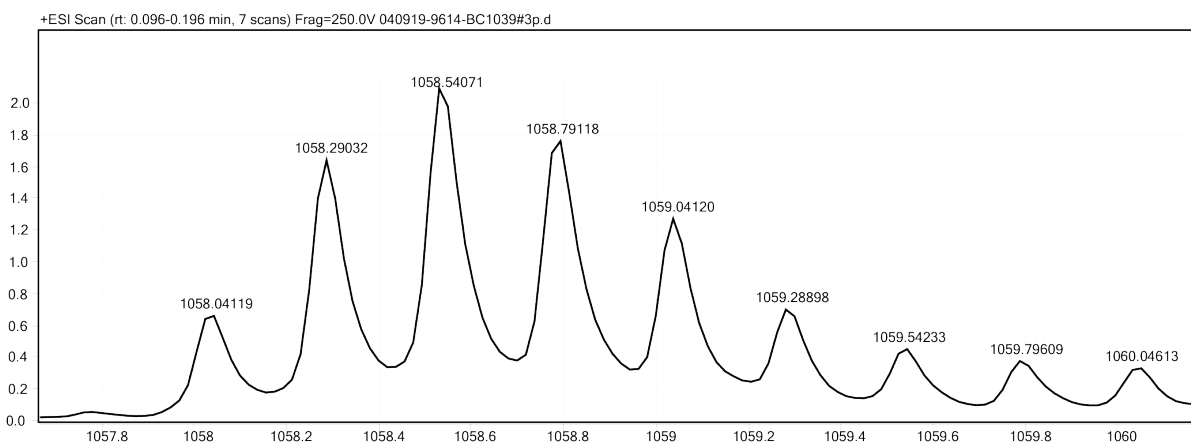


Figure S23. ESI-TOF MS data for **pLF** (BC1039#3p). Expected $[M+4H^+]/4 = 1058.0438$.

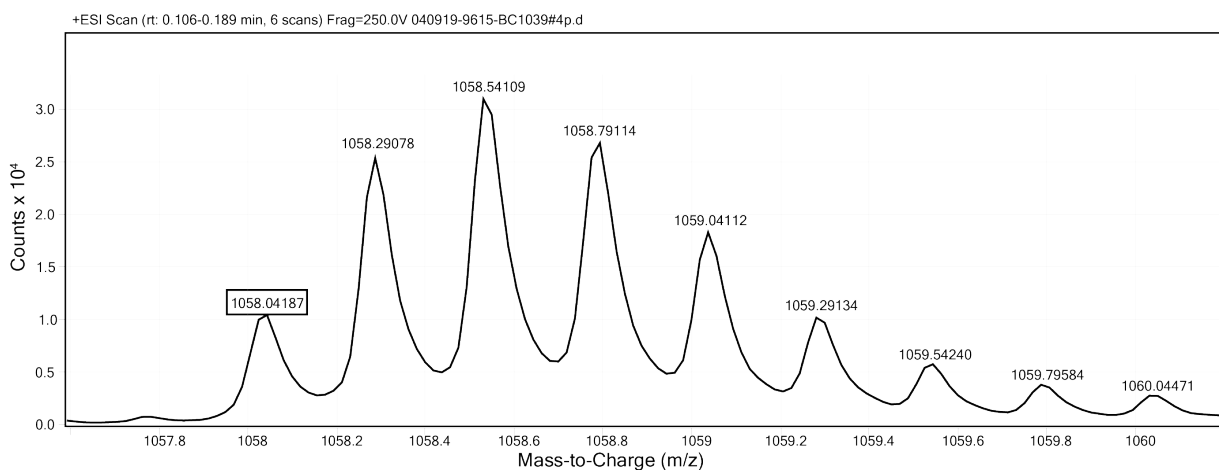


Figure S24. ESI-TOF MS data for **pFL** (BC1039#4p). Expected $[M+4H^+]/4 = 1058.0438$.

2. Analytical HPLC Chromatograms

Peptide solution was injected onto a C18 analytical column and eluted with a linear gradient of 10–60% B (A = H₂O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 min.; 10-min. rinse (95% B); and 10-min. column re-equilibration. Chromatograms appear in Figures S25–S48.

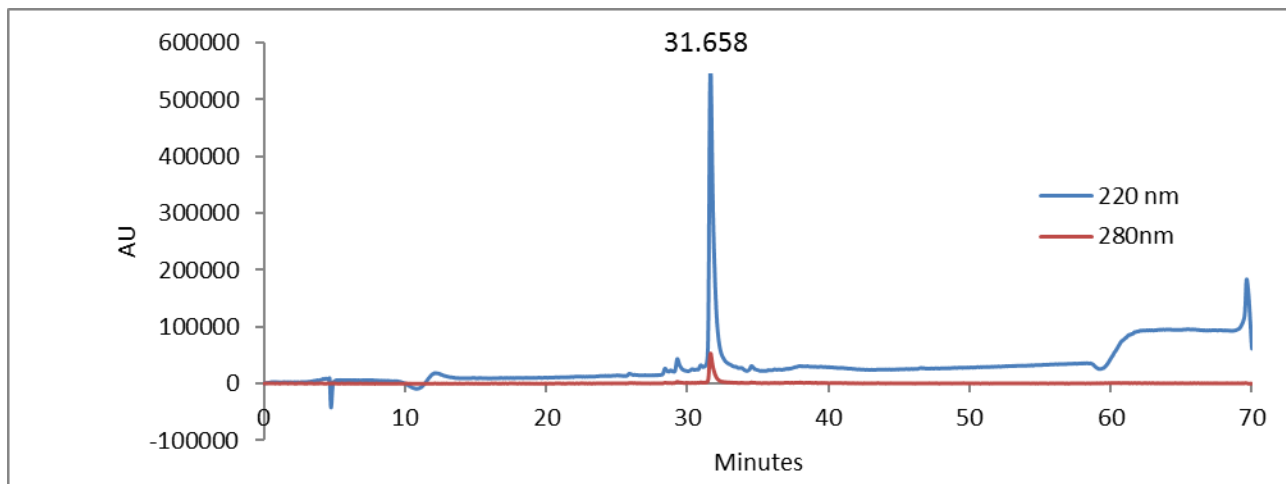


Figure S25. Analytical HPLC data for **LL** (SD1075#1). Retention time = 31.658 minutes.

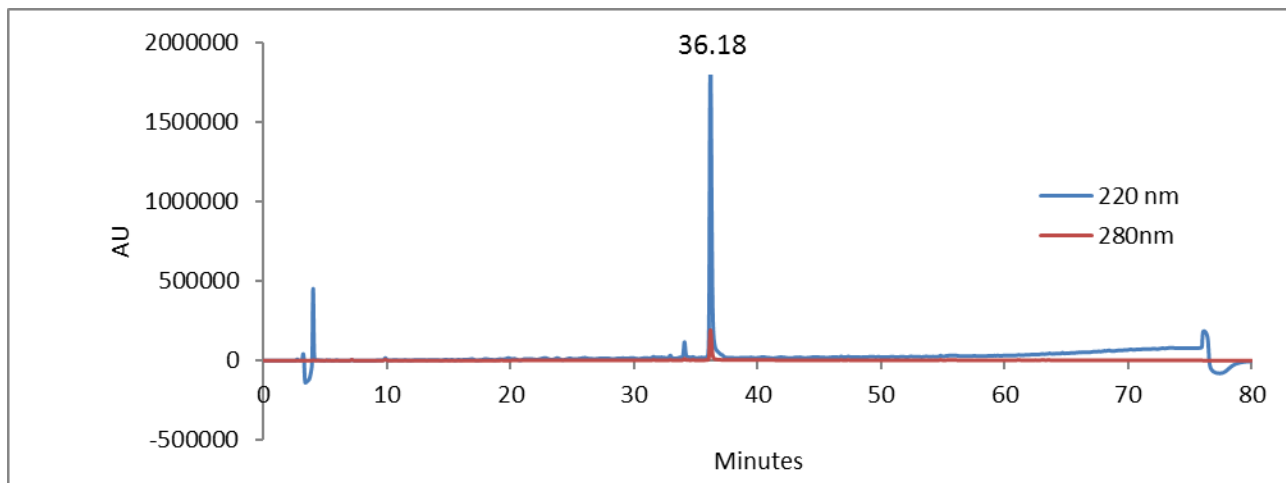


Figure S26. Analytical HPLC data for **LA** (DA1018#2). Retention time = 36.18 minutes.

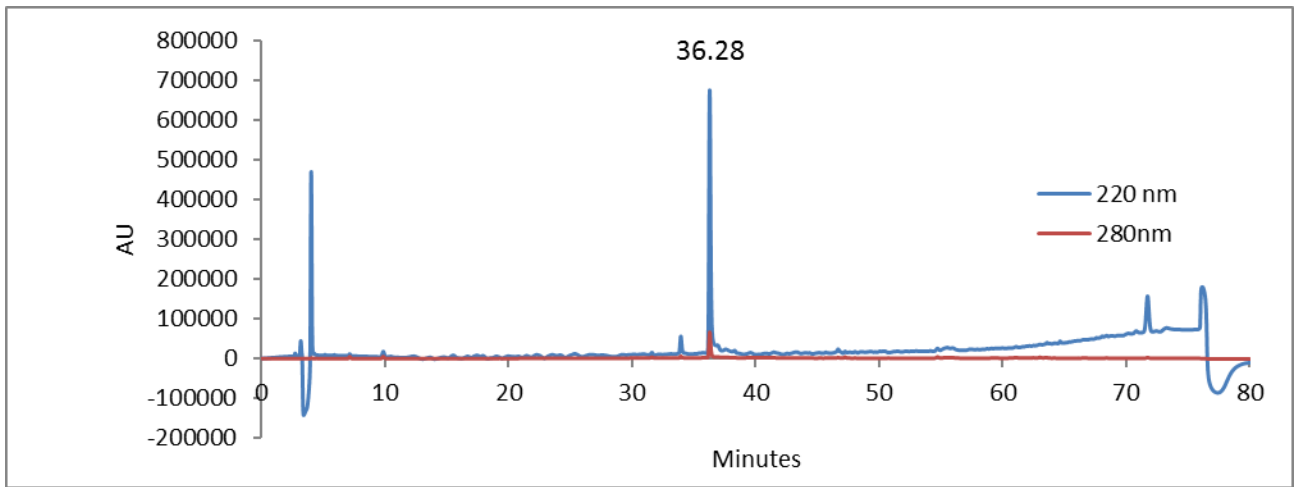


Figure S27. Analytical HPLC data for **AL** (DA1018#1). Retention time = 36.28 minutes.

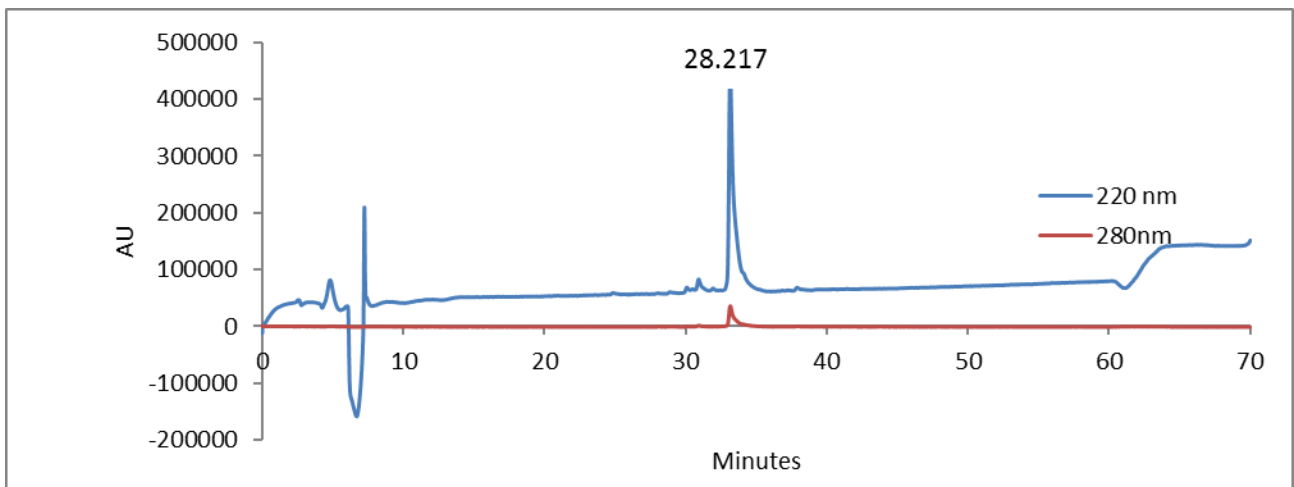


Figure S28. Analytical HPLC data for **pLL** (SD1075#1C). Retention time = 28.217 minutes.

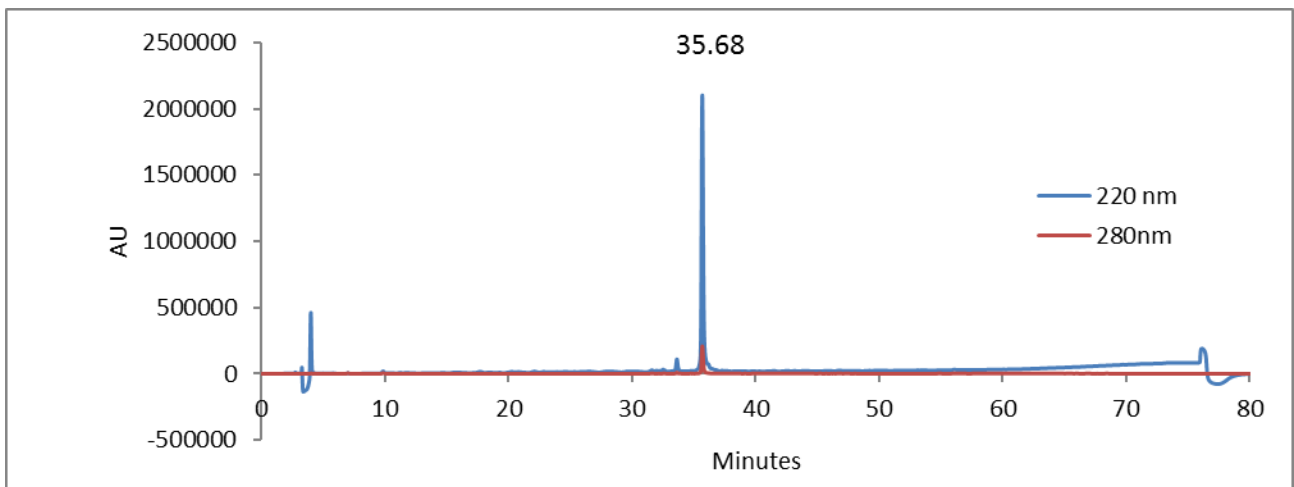


Figure S29. Analytical HPLC data for **pLA** (DA1018#4). Retention time = 35.68 minutes.

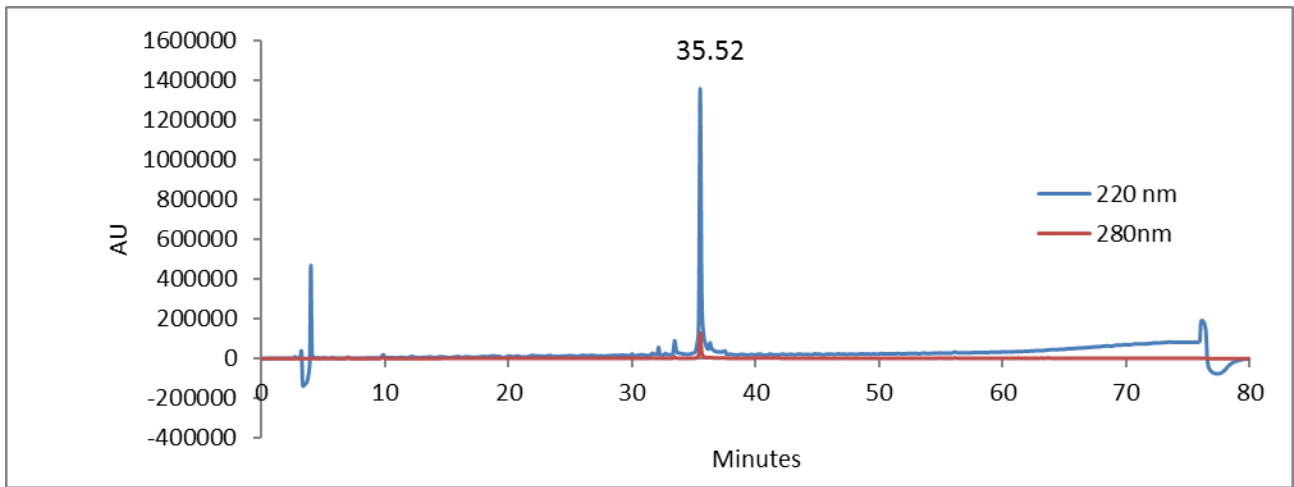


Figure S30. Analytical HPLC data for **pAL** (DA1018#3). Retention time = 35.52 minutes.

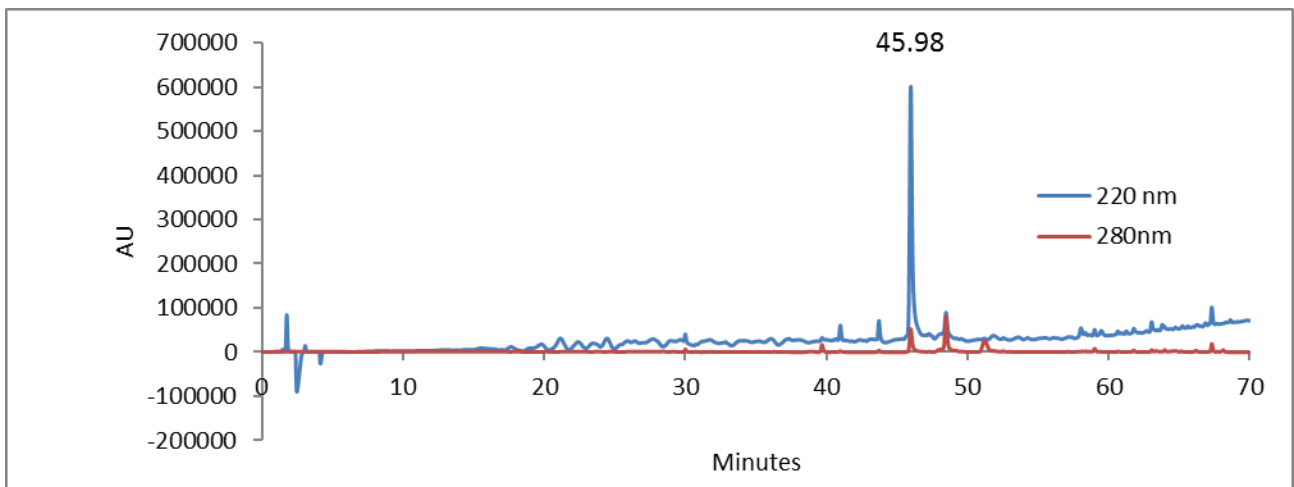


Figure S31. Analytical HPLC data for **FF** (DA1037#1). Retention time = 45.98 minutes.

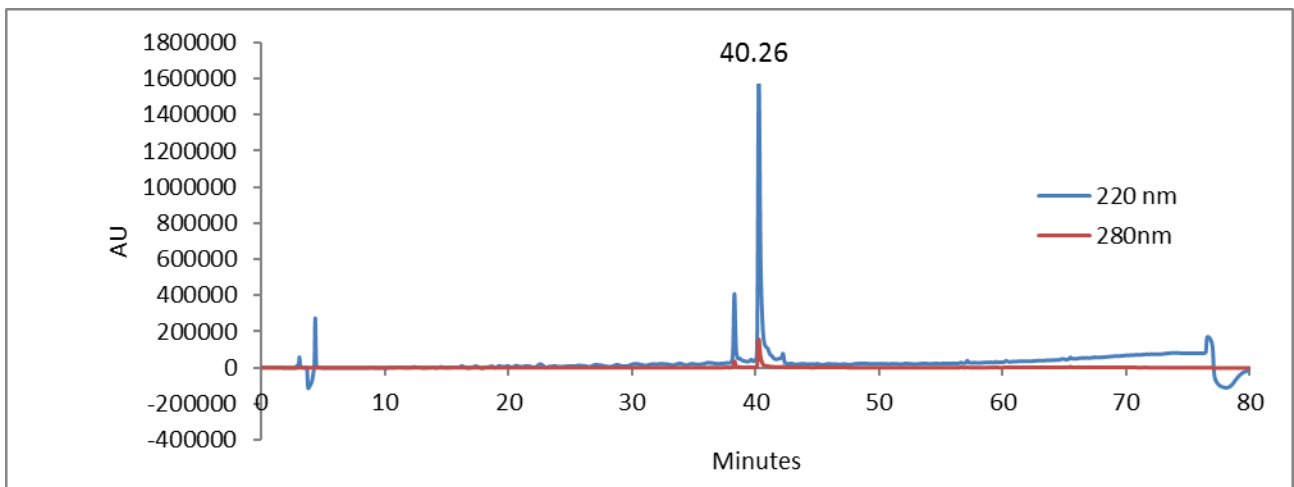


Figure S32. Analytical HPLC data for **FA** (SD2176#1). Retention time = 40.26 minutes.

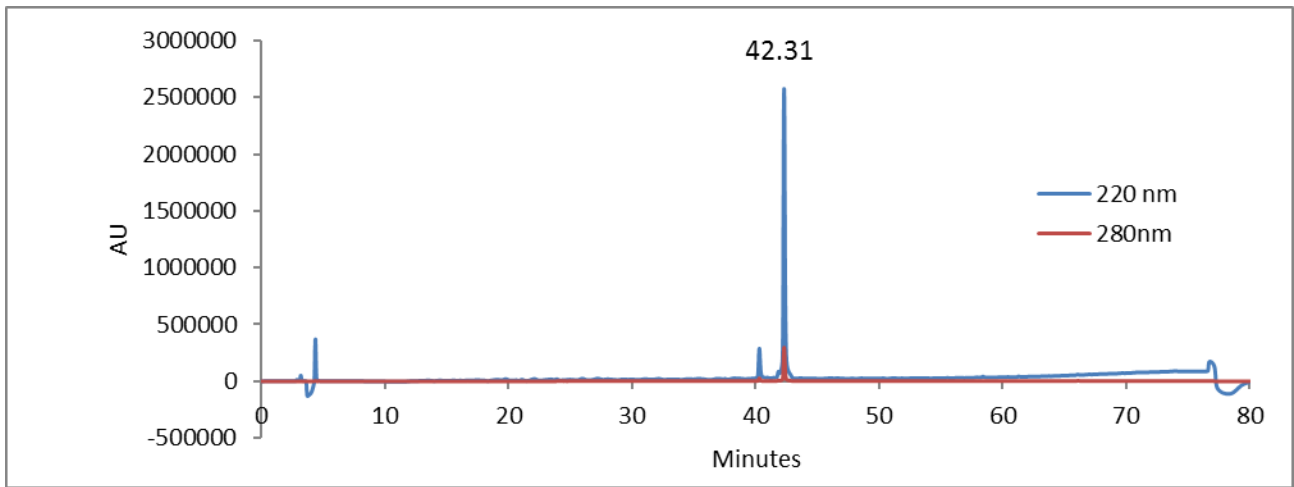


Figure S33. Analytical HPLC data for **FA** (SD2176#2). Retention time = 42.31 minutes.

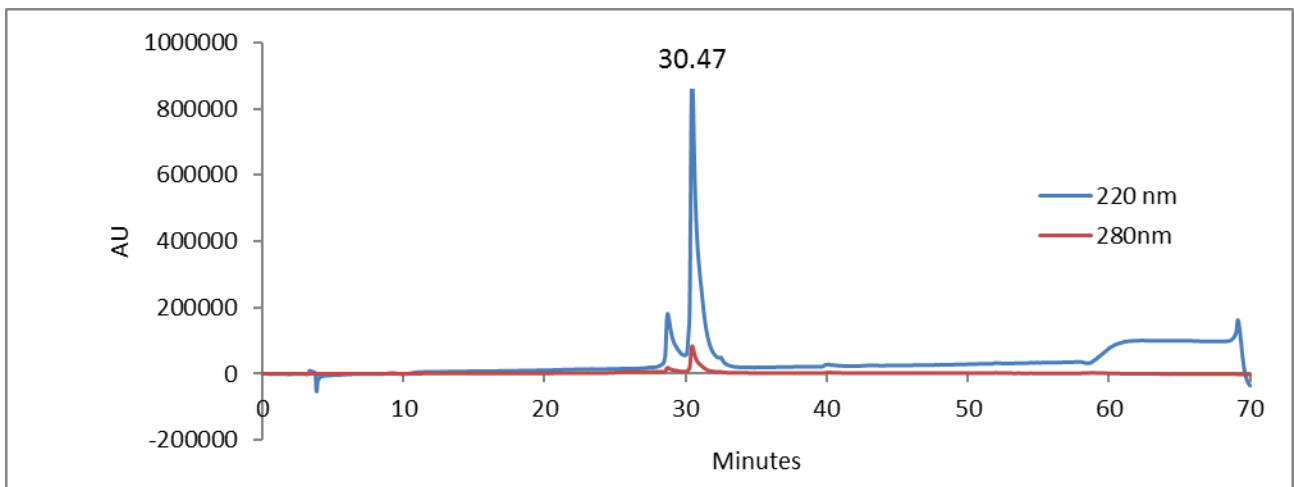


Figure S34. Analytical HPLC data for **pFF** (DA1037#3). Retention time = 30.47 minutes.

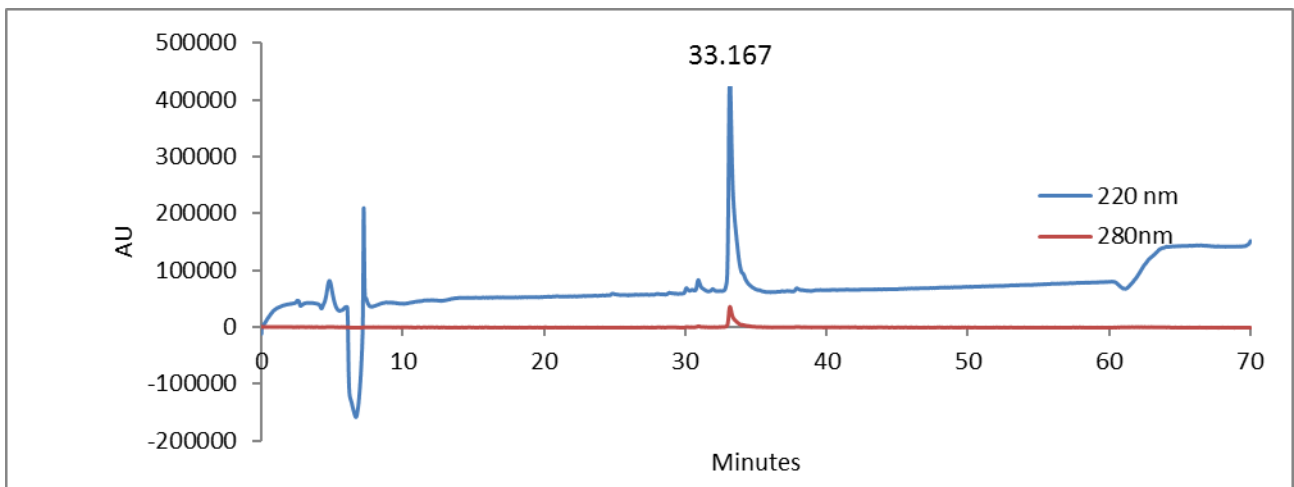


Figure S35. Analytical HPLC data for **pFA** (SD2176#1p). Retention time = 33.167 minutes.

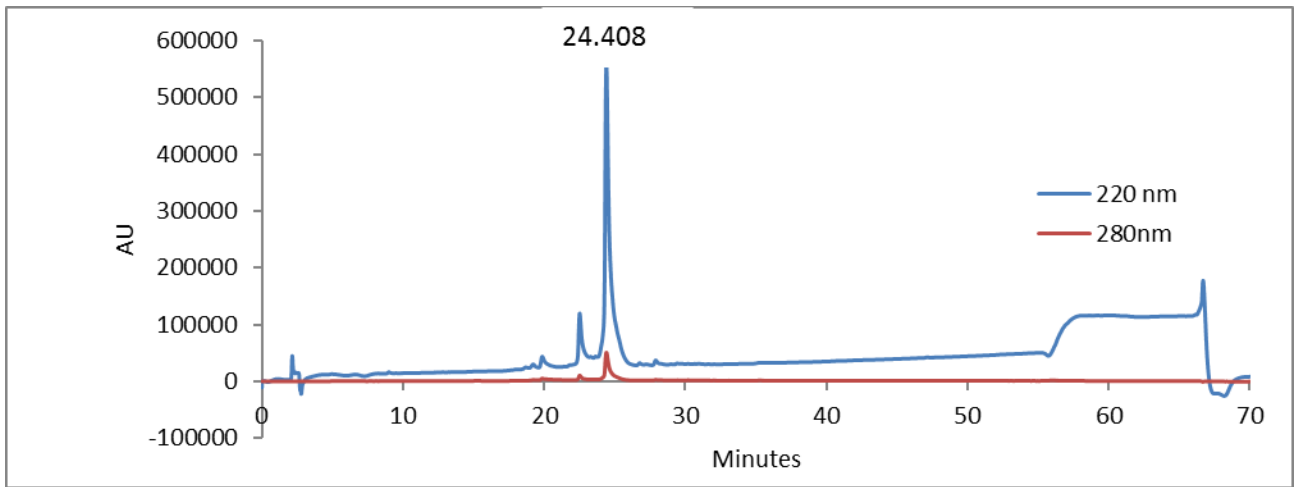


Figure S36. Analytical HPLC data for **pAF** (SD2176#2p). Retention time = 24.408 minutes.

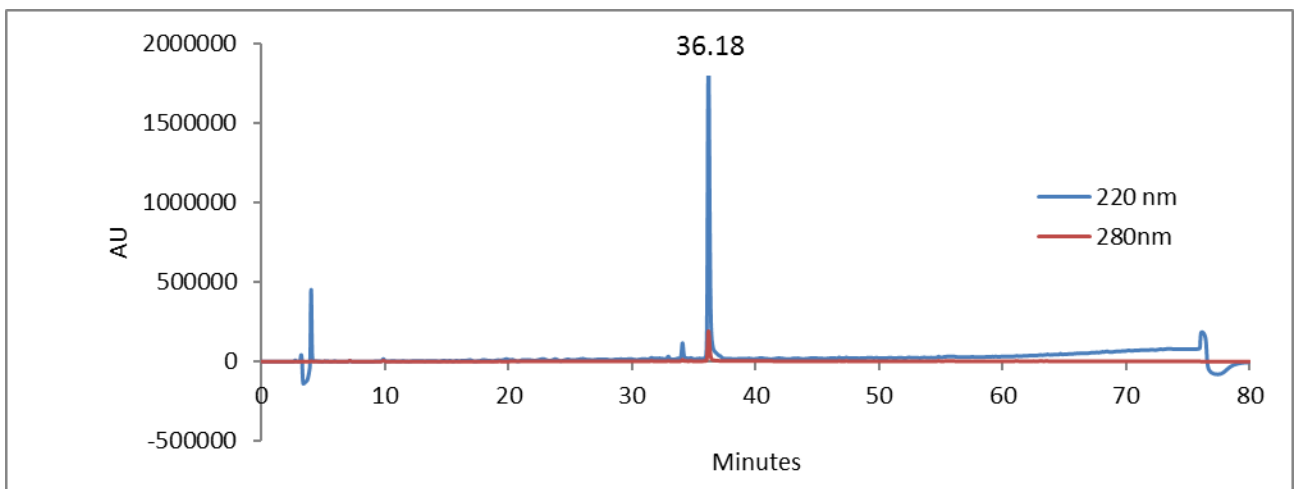


Figure S37. Analytical HPLC data for **XX** (DA1037#2). Retention time = 36.18 minutes.

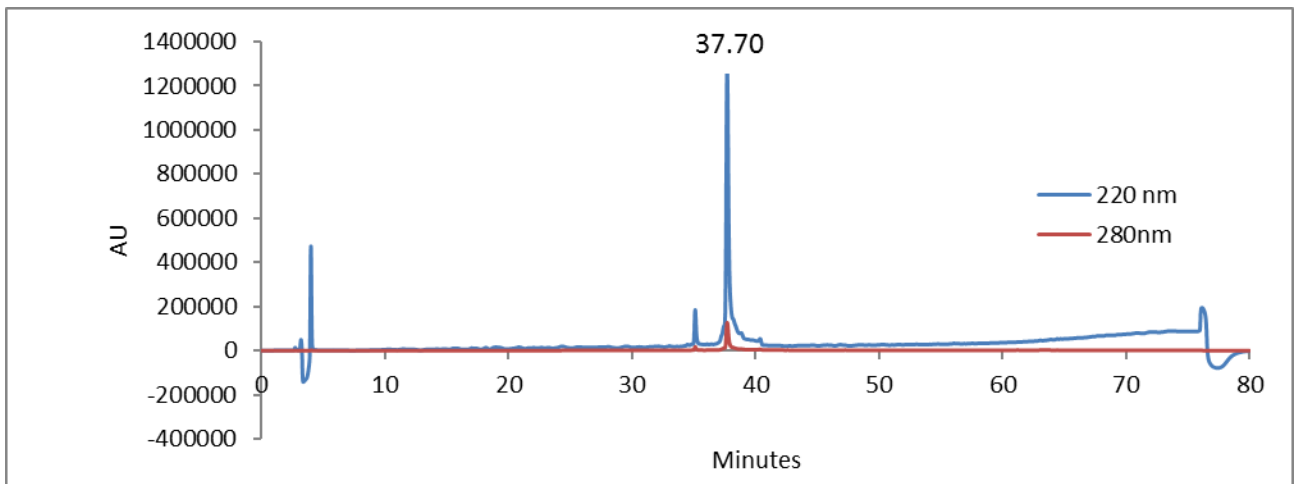


Figure S38. Analytical HPLC data for **XA** (SD2184#2). Retention time = 37.70 minutes.

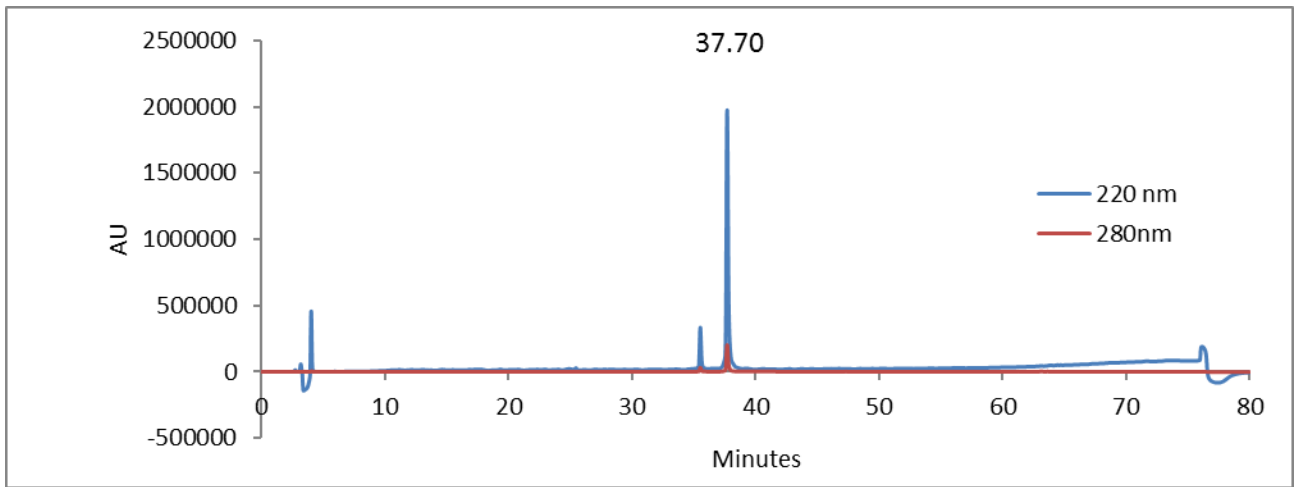


Figure S39. Analytical HPLC data for **AX** (SD2184#1). Retention time = 37.70 minutes.

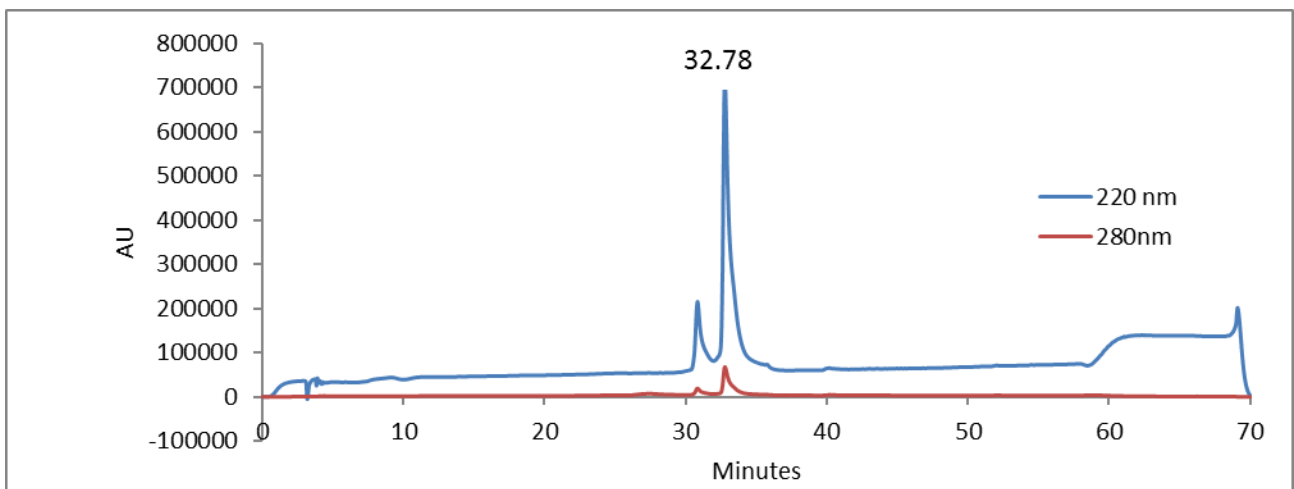


Figure S40. Analytical HPLC data for **pXX** (DA1037#4). Retention time = 32.78 minutes.

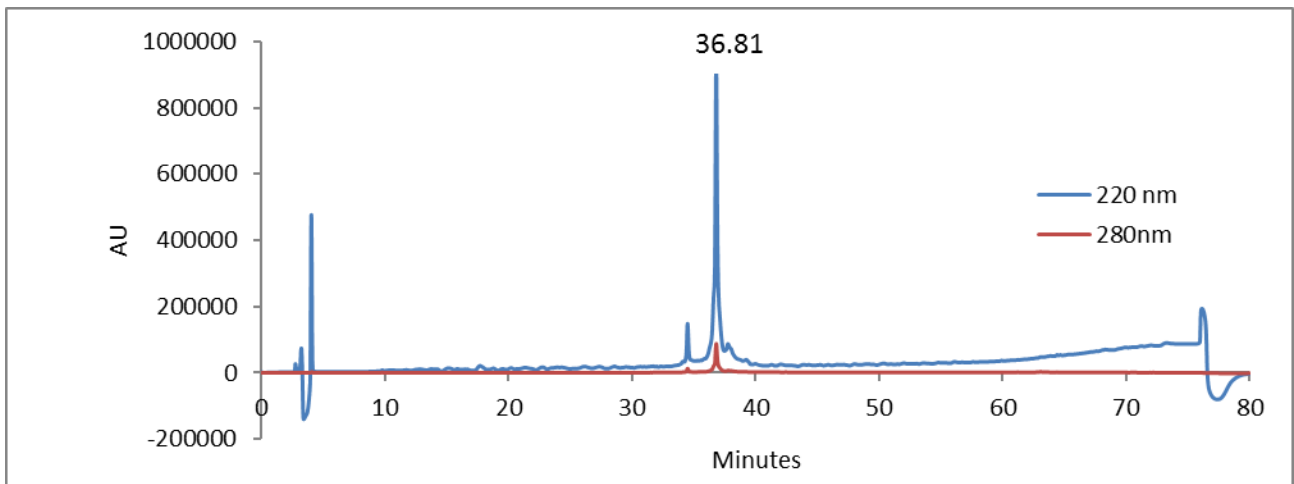


Figure S41. Analytical HPLC data for **pXA** (SD2184#2p). Retention time = 36.81 minutes.

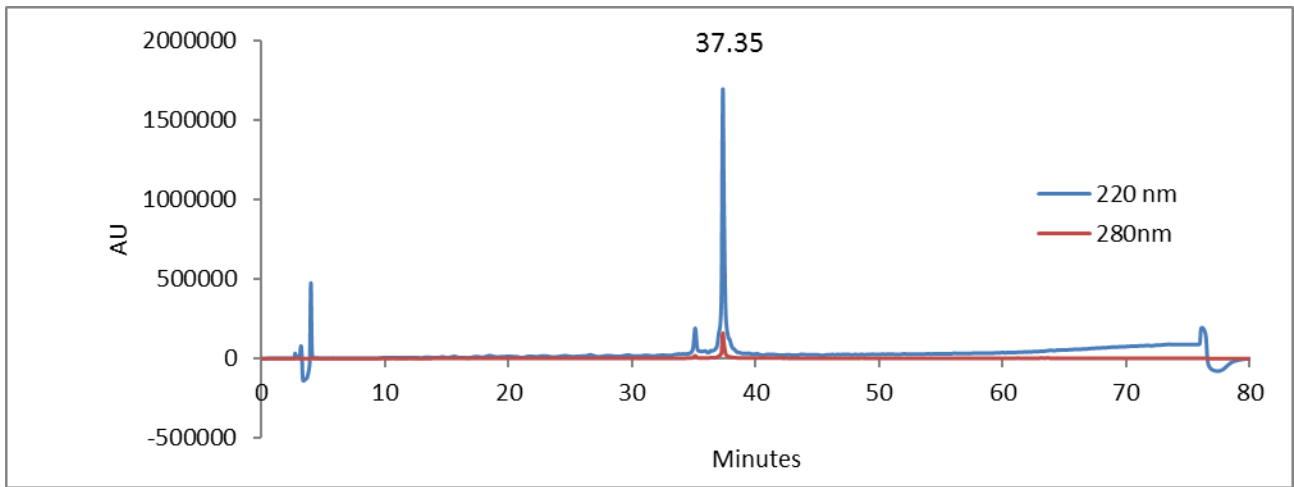


Figure S42. Analytical HPLC data for **pAX** (SD2184#1p). Retention time = 37.35 minutes.

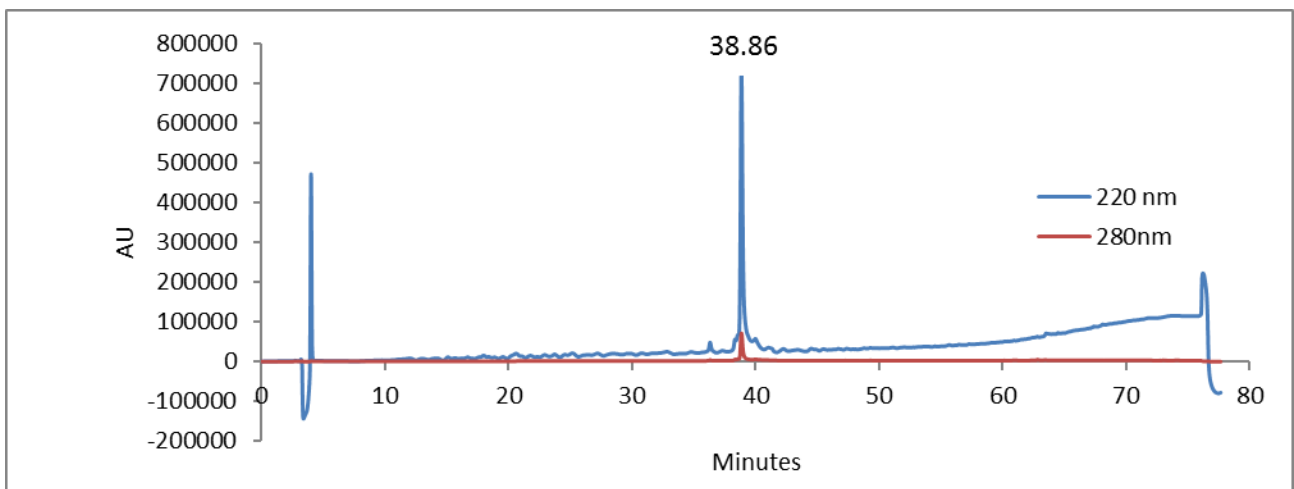


Figure S43. Analytical HPLC data for **XL** (BC1039#1). Retention time = 38.86 minutes.

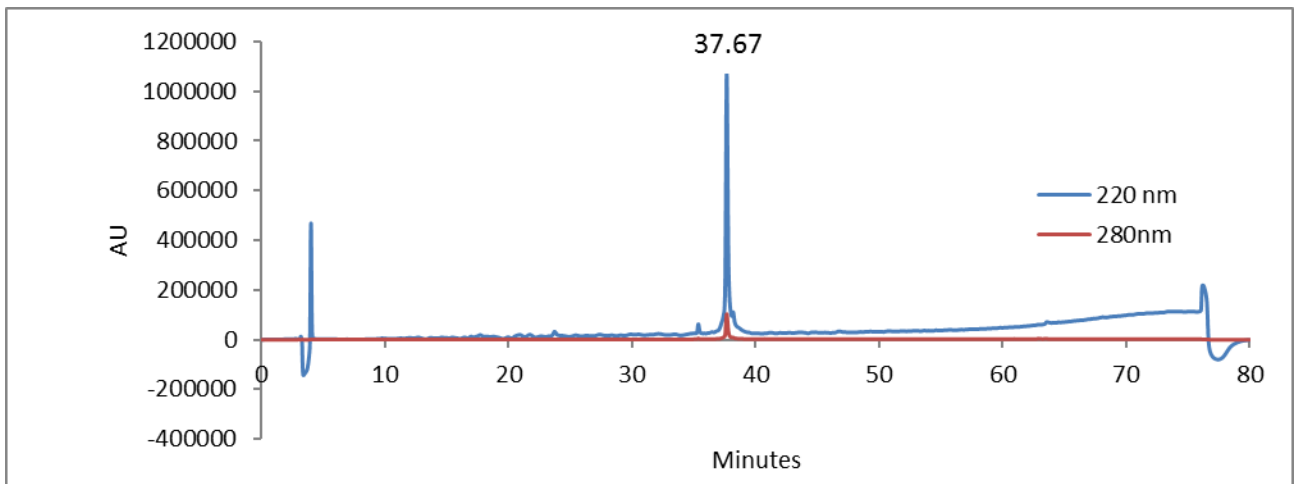


Figure S44. Analytical HPLC data for **FL** (BC1039#4). Retention time = 37.67 minutes.

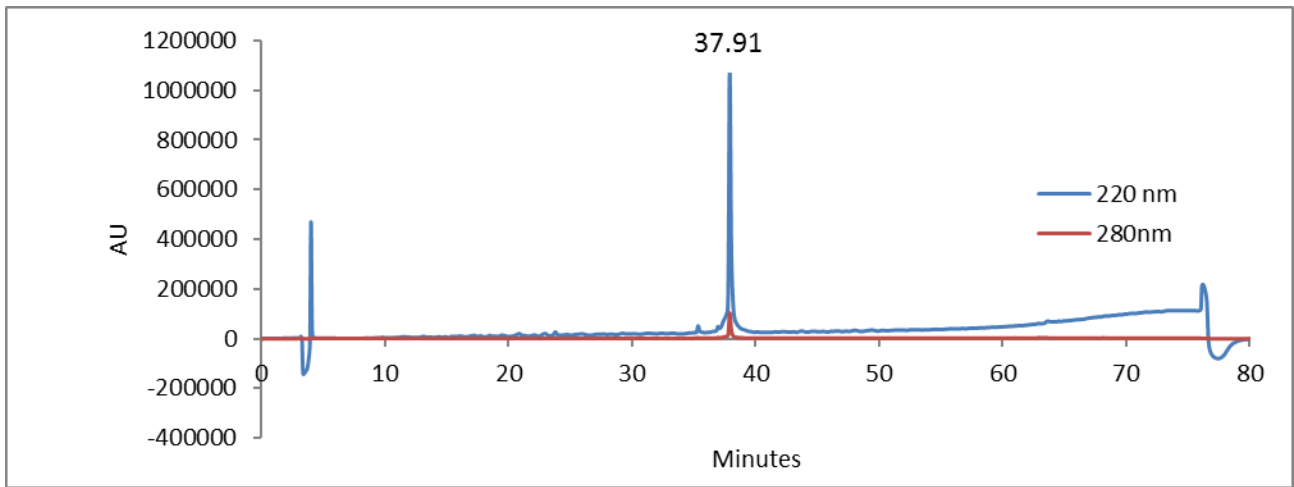


Figure S45. Analytical HPLC data for **LF** (BC1039#3). Retention time = 37.91 minutes.

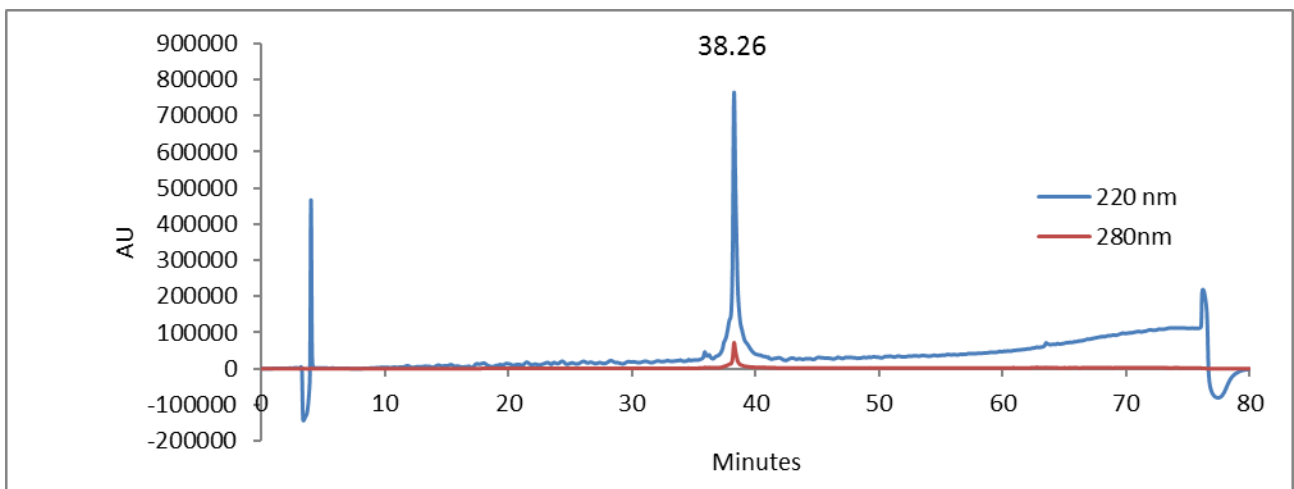


Figure S46. Analytical HPLC data for **pXL** (BC1039#1p). Retention time = 38.26 minutes.

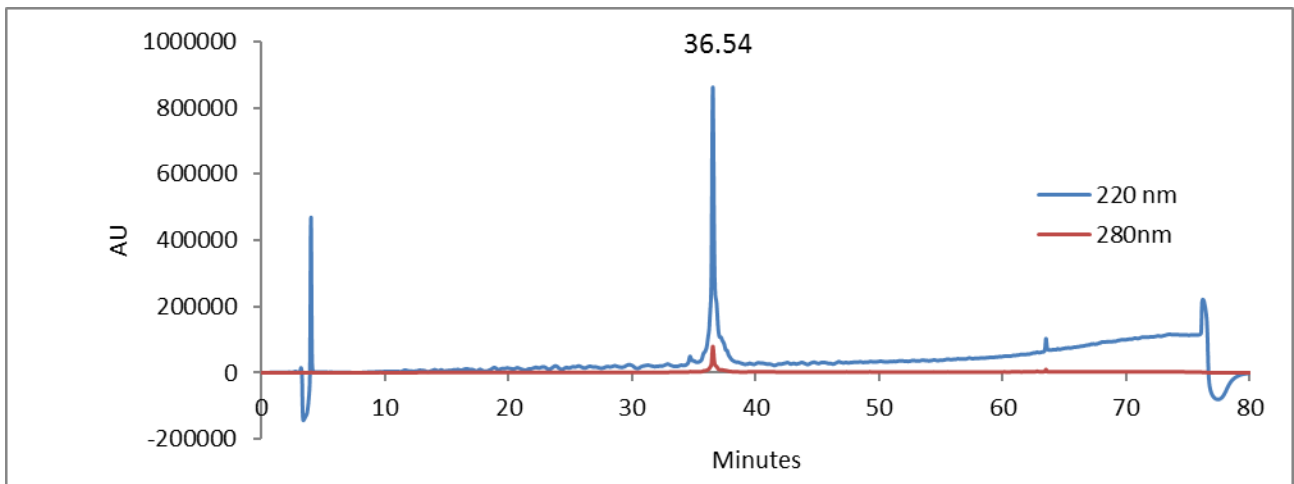


Figure S47. Analytical HPLC data for **pFL** (BC1039#4p). Retention time = 36.54 minutes.

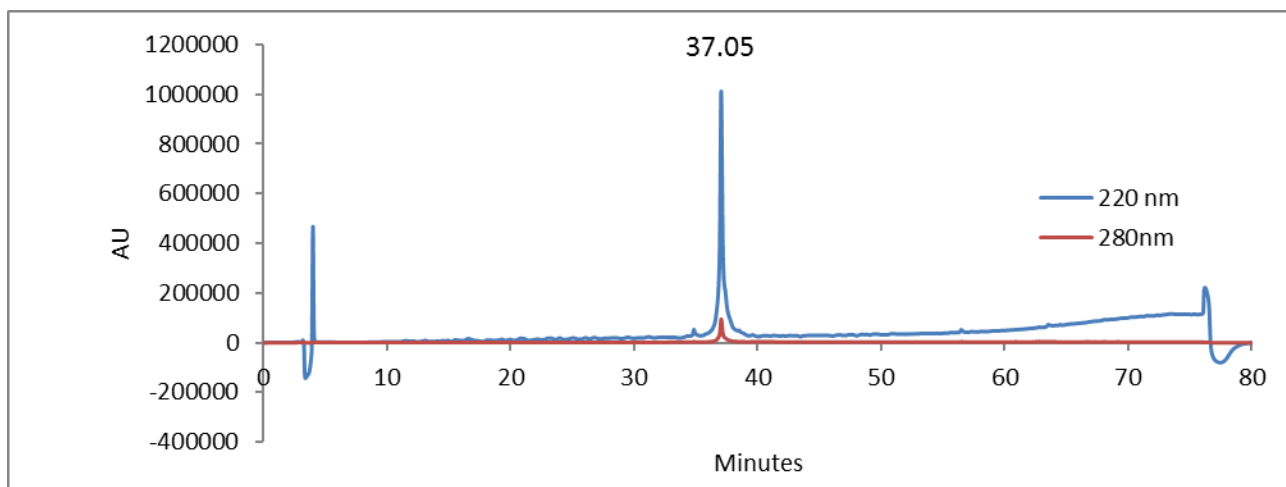


Figure S48. Analytical HPLC data for **pLF** (BC1039#3p). Retention time = 37.05 minutes.

3. Fitting of Variable Temperature CD data

Variable temperature CD data were obtained at least in triplicate (one sample was made and then aliquoted into three different cuvettes) by monitoring the molar ellipticity $[\theta]$ at 227 nm of 50 μM solutions of each WW variant in 20 mM sodium phosphate (pH 7) from 1 to 95 $^{\circ}\text{C}$ at 2 $^{\circ}\text{C}$ intervals, with 120 s equilibration time between data points and 30 s averaging time. Triplicate variable temperature CD data for each peptide were fit globally to a two-state model for thermally induced unfolding. This approach treats the observed $[\theta]$ of a peptide solution at a given temperature as the average of the $[\theta]$ values for the folded state and the unfolded ensemble, weighted according to their relative concentrations at that temperature, as shown in the following equation:

$$[\theta] = (a + bT)(F_{\text{folded}}) + (c + dT)(1 - F_{\text{folded}}) \quad (\text{S1})$$

In equation S1, T is the temperature in Kelvin; a and b are the intercept and slope of the pre-transition baseline (which represents the linear dependence of the folded ensemble CD signal $[\theta]$ on temperature); c and d are the intercept and slope of the post-transition baseline (which represents the linear dependence of the unfolded ensemble CD signal $[\theta]$ on temperature); and F_{folded} is the fraction of the total protein concentration that is folded as at temperature T .

F_{fit} is a function of the folding equilibrium constant; folding of WW involves an equilibrium between an unfolded monomer (**U**) and a folded monomer (**F**) with the temperature-dependent equilibrium constant **K** as defined below:



The constant total concentration of peptide in solution **P** is defined by equation S3:

$$P = [\mathbf{U}] + [\mathbf{F}] = [\mathbf{U}] + \mathbf{K}[\mathbf{U}] \quad (\text{S3})$$

F_{folded} of the monomer folding equilibrium is defined as follows:

$$F_{\text{folded}} = \frac{\mathbf{K}}{1+\mathbf{K}} \quad (\text{S4})$$

In each of these cases, \mathbf{K} is related to the change in free energy upon folding (ΔG_f):

$$\mathbf{K} = e^{-\frac{\Delta G_f}{RT}} \quad (\text{S5})$$

In turn, the temperature-dependence of ΔG_f for can be expressed as a polynomial expansion about the melting temperature T_m (i.e., the temperature at which $F_{\text{folded}} = 0.5$, $\mathbf{K} = 1$, and $\Delta G_f = 0$):

$$\Delta G_f = \Delta G_1(T - T_m) + \Delta G_2(T - T_m)^2 \quad (\text{S6})$$

where ΔG_1 , ΔG_2 , and T_m are parameters to be determined via least-squares regression (though ΔG_2 was excluded from the fits of all the WW variants described here, because attempts to use it resulted in corresponding parameter standard error values that were too high). We used least-squares regression to fit the variable temperature CD for each variant to these equations. Far-UV CD spectra and variable temperature CD data for these compounds are shown below in Figures S49-S74, along with the parameters of the fits (and their standard errors and p-values) and fit statistics (including R^2 and sum of the squared residuals).

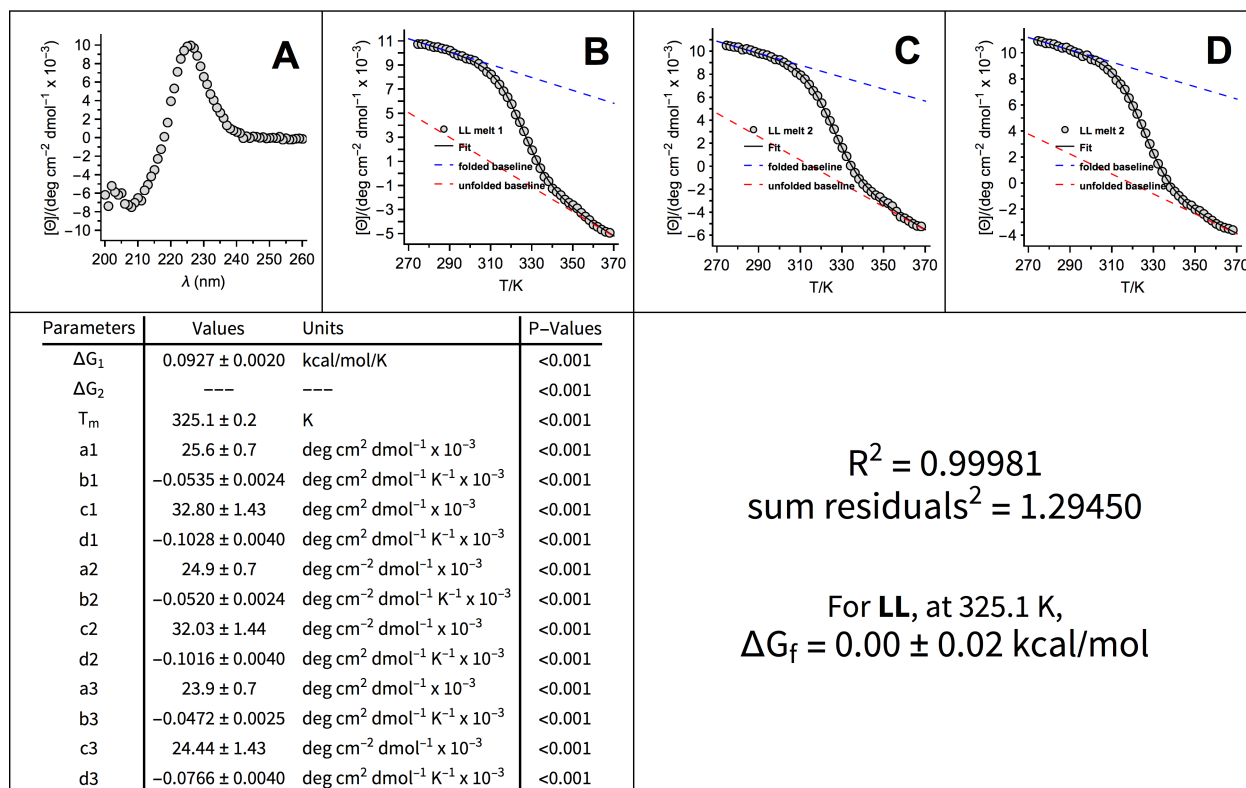


Figure S49. CD data for LL (SD1075#1).

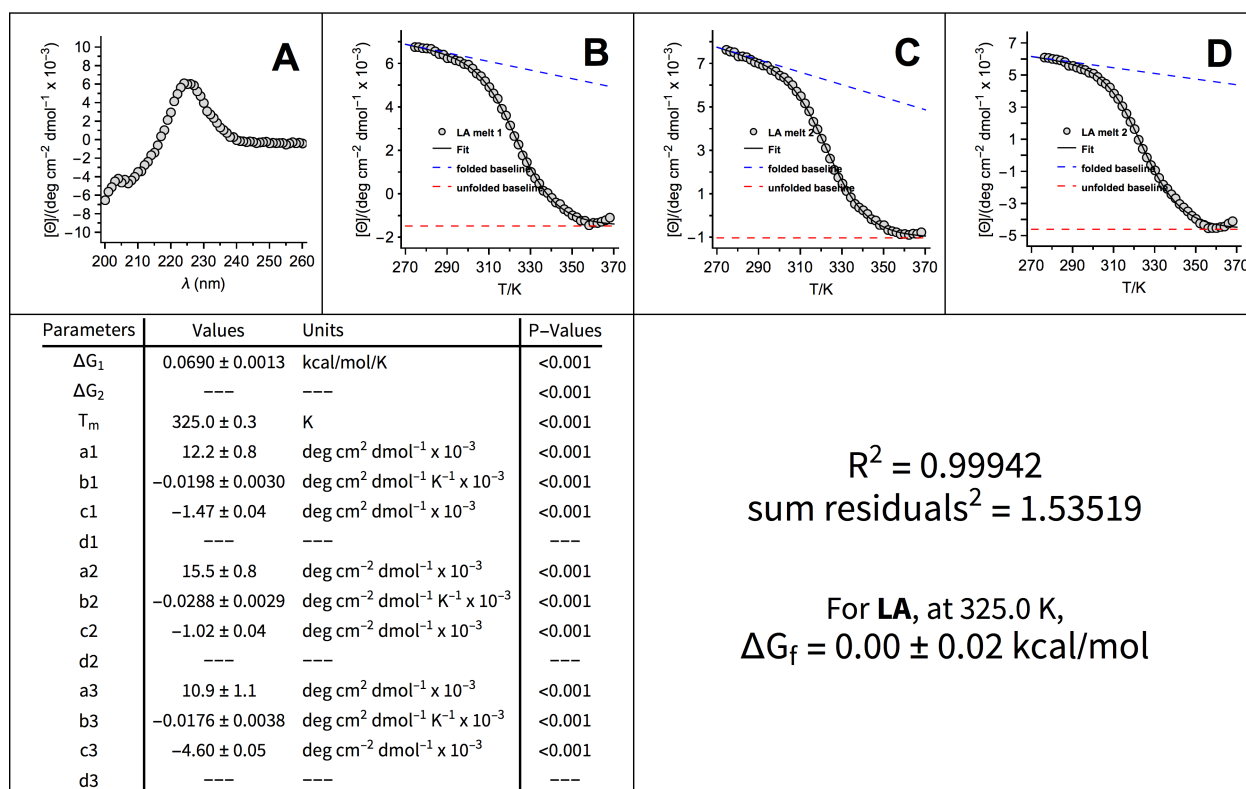


Figure S50. CD data for LA (DA1018#2).

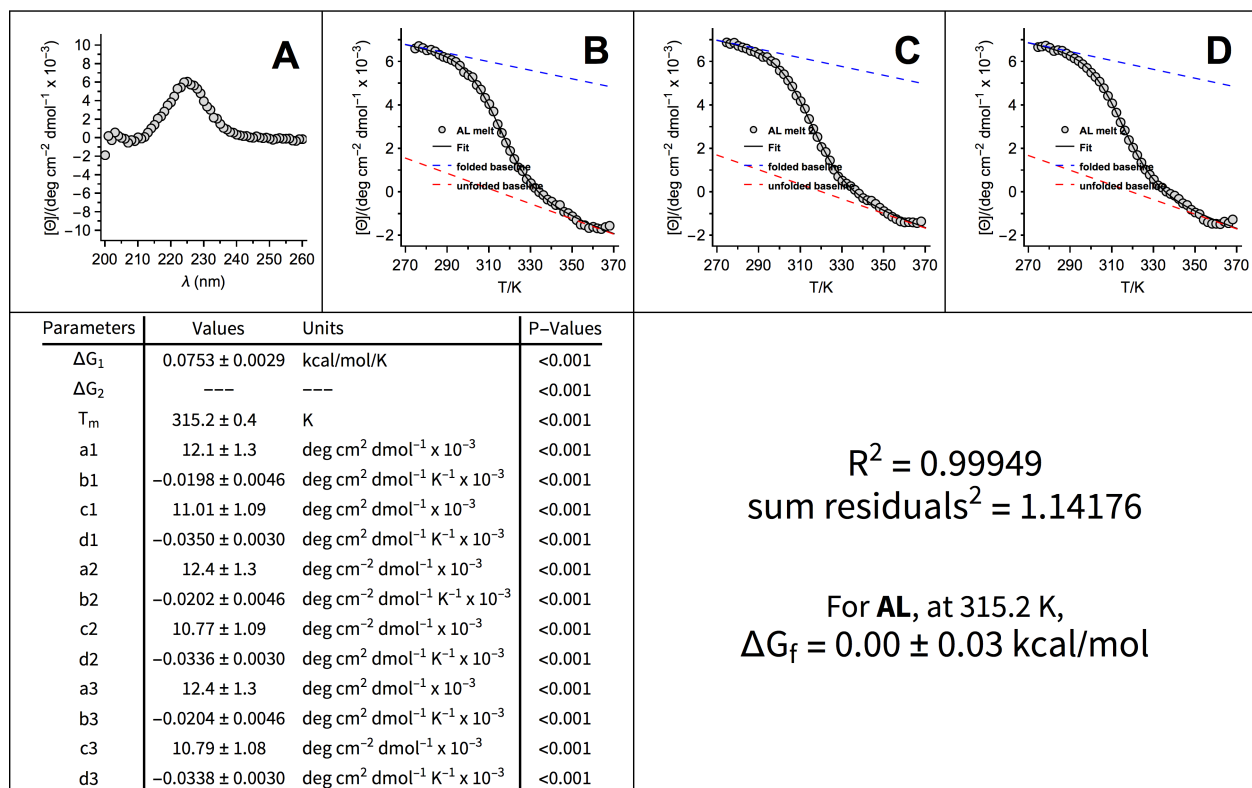


Figure S51. CD data for AL (DA1018#1).

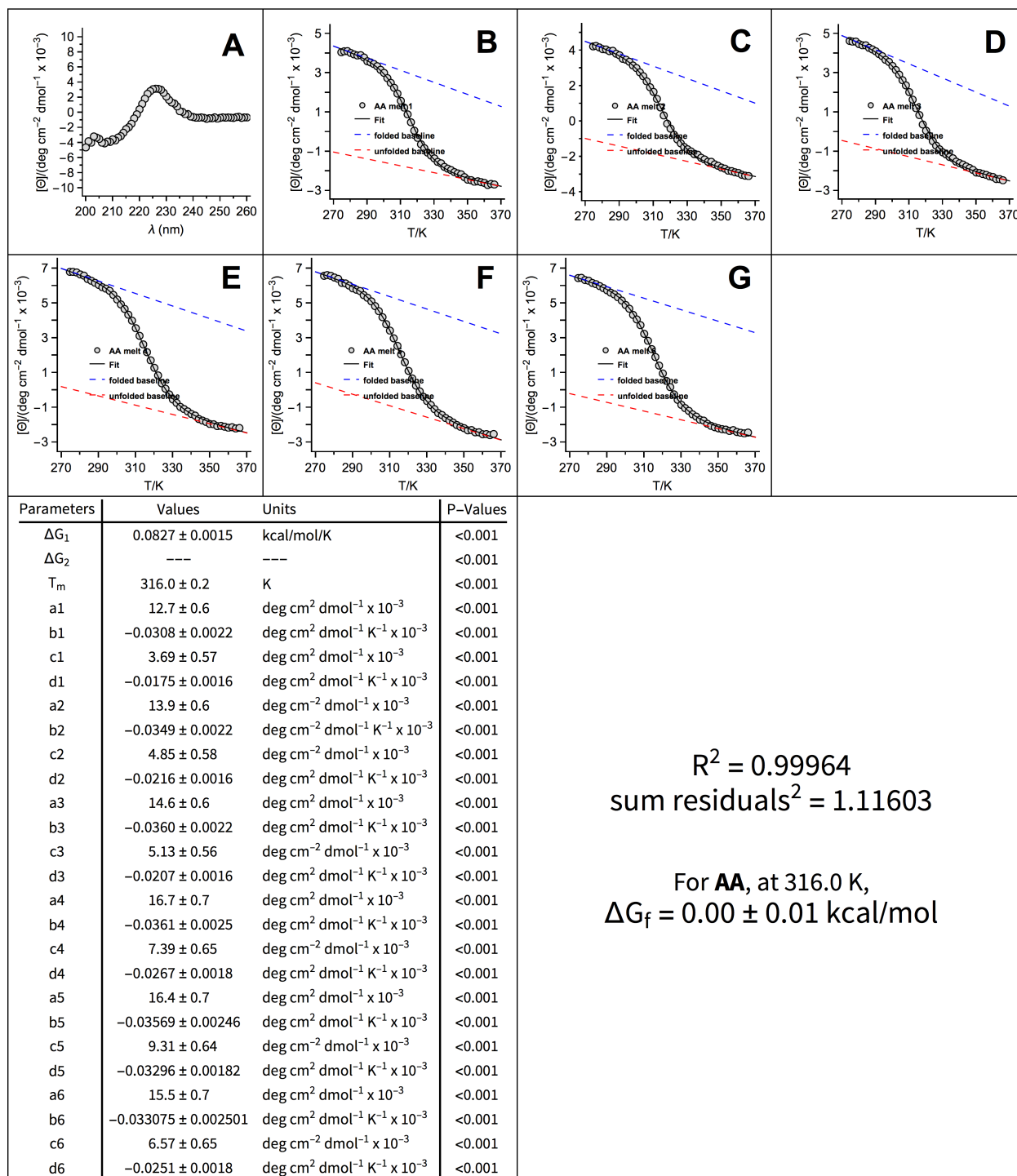


Figure S52. CD data for AA.

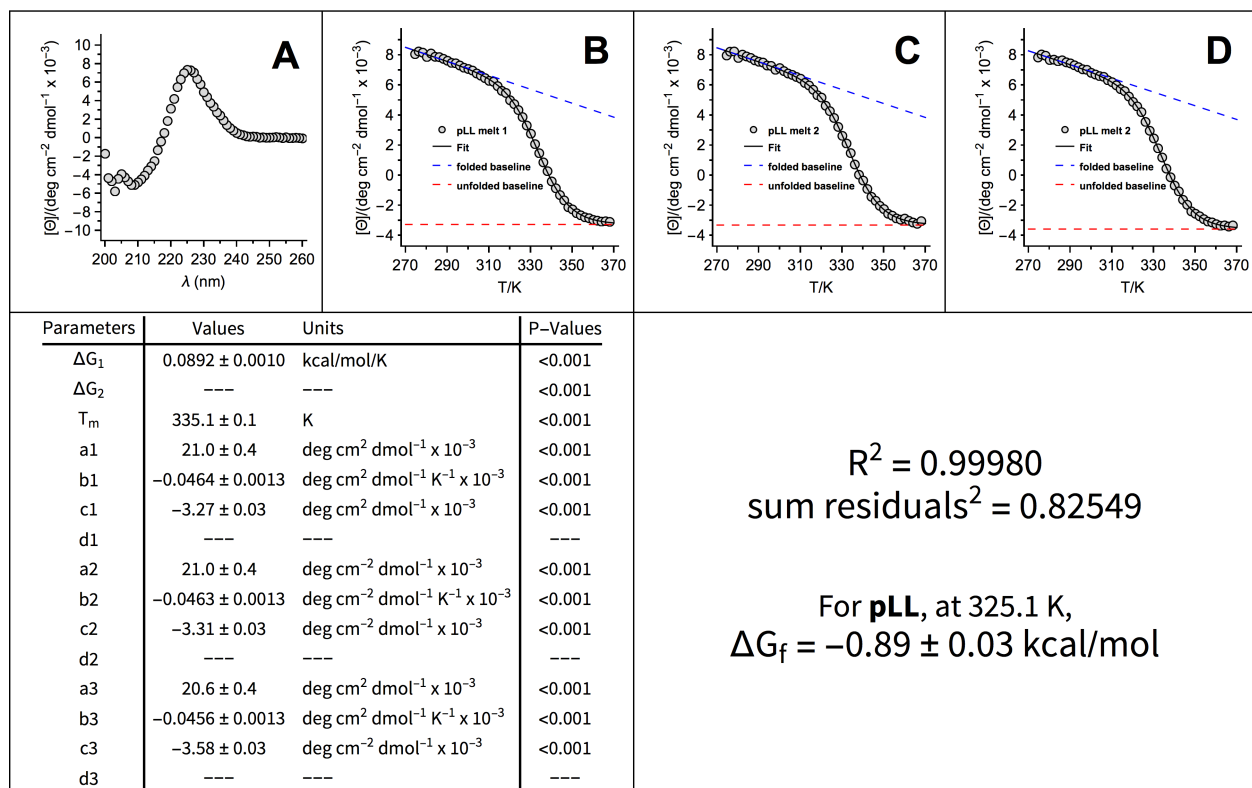


Figure S53. CD data for pLL (SD1075#1C).

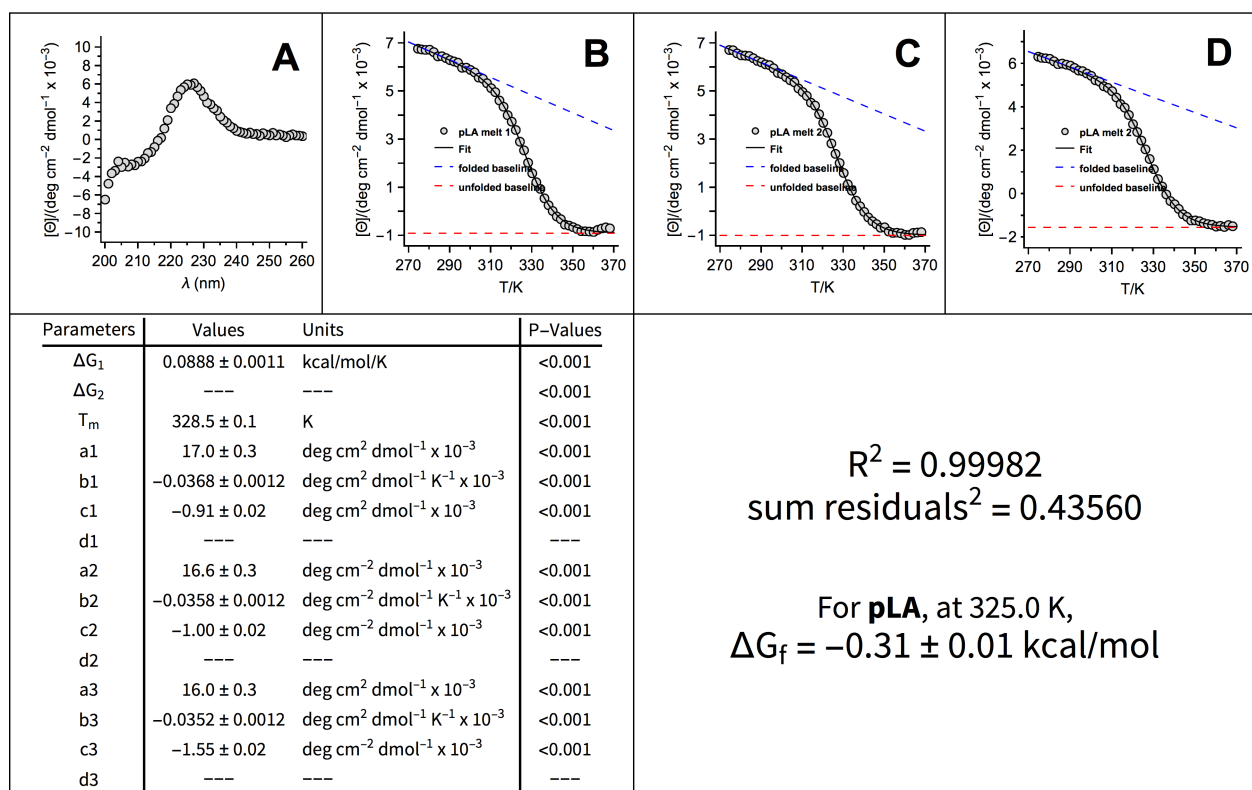


Figure S54. CD data for pLA (DA1018#4).

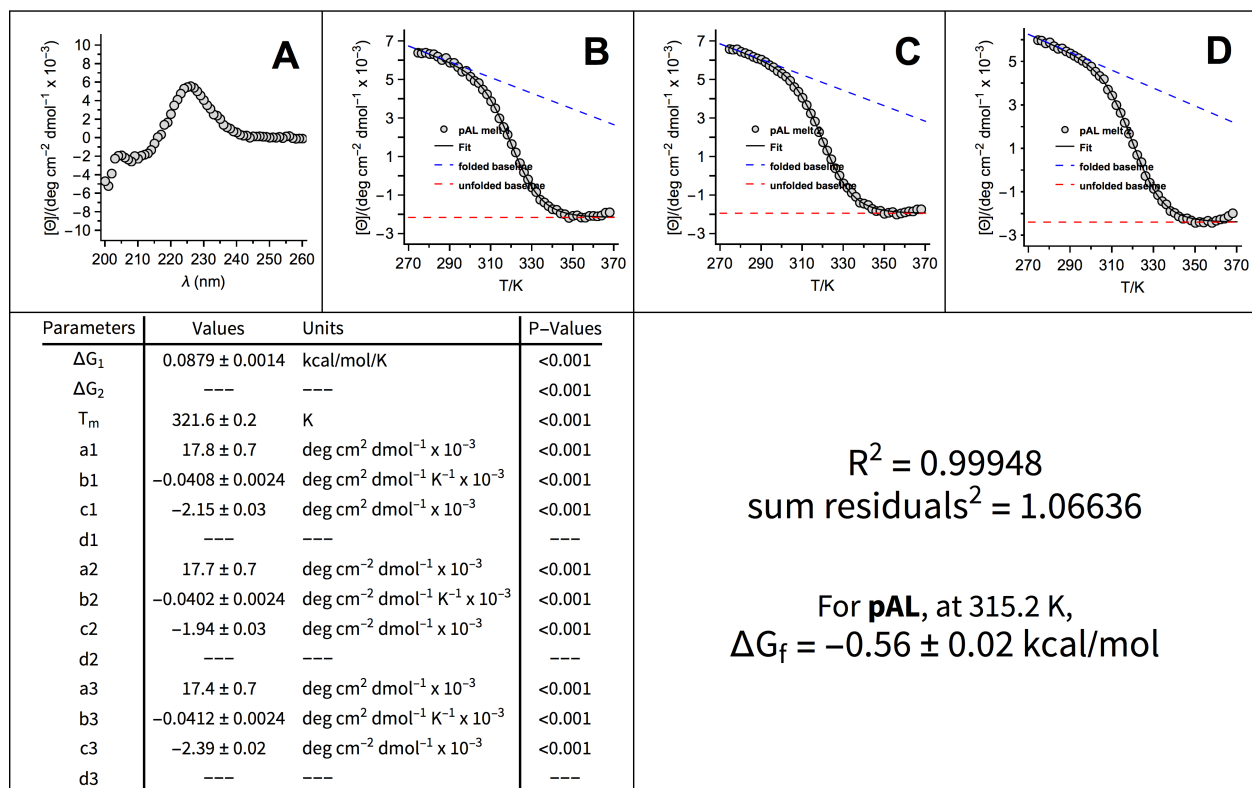


Figure S55. CD data for pAL (DA1018#3).

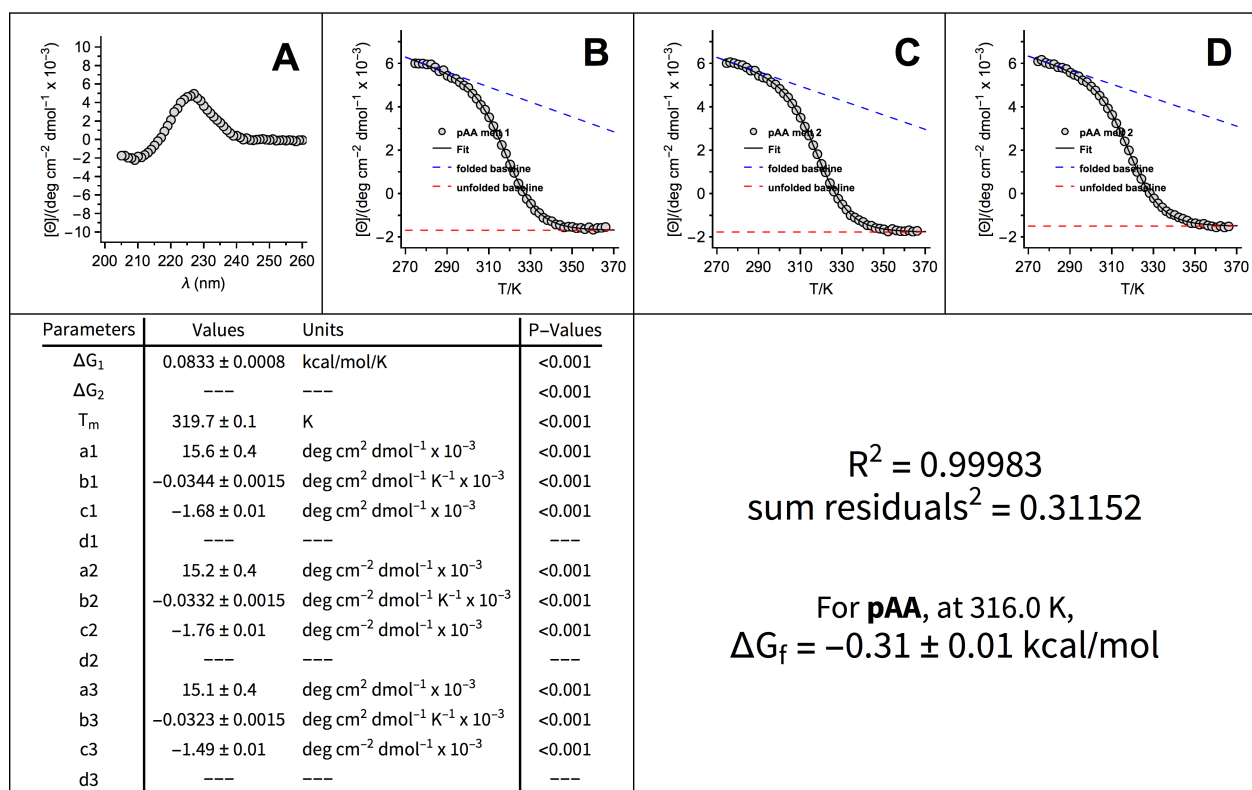


Figure S56. CD data for pAA.

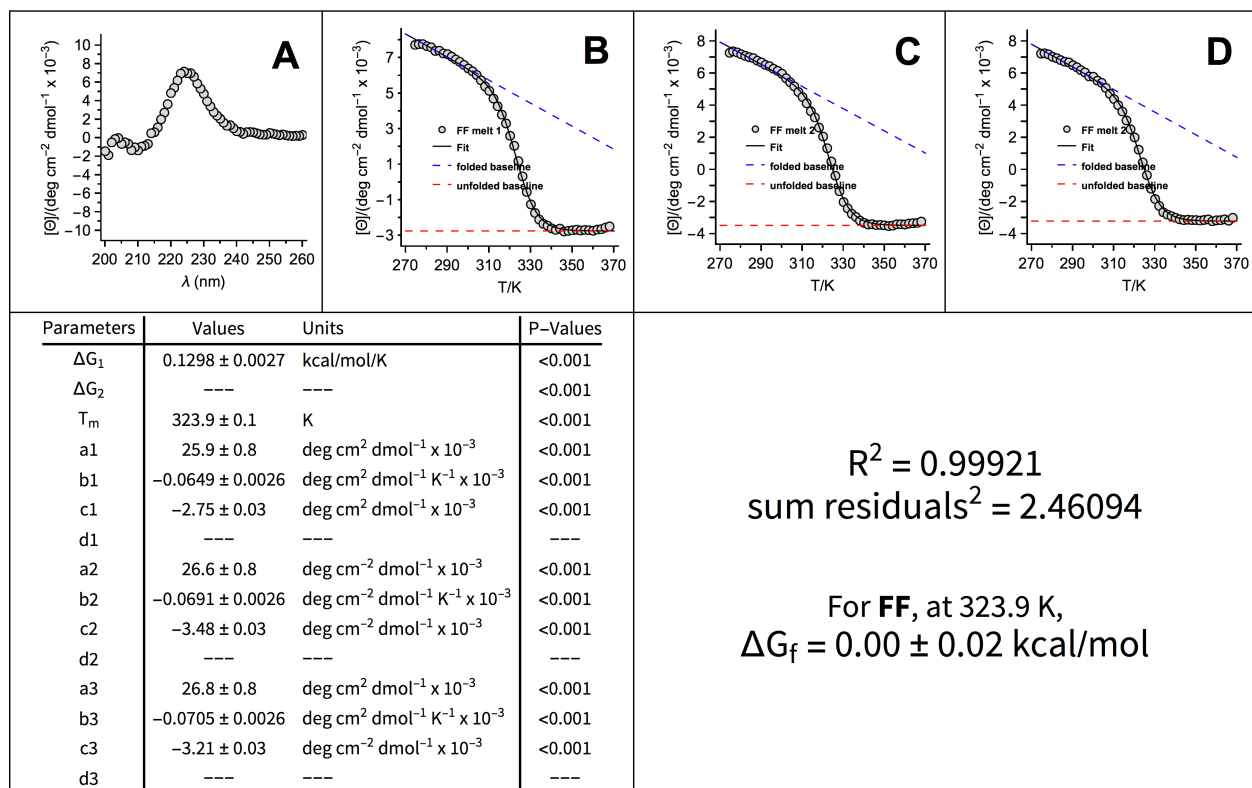


Figure S57. CD Data for FF (DA1037#1).

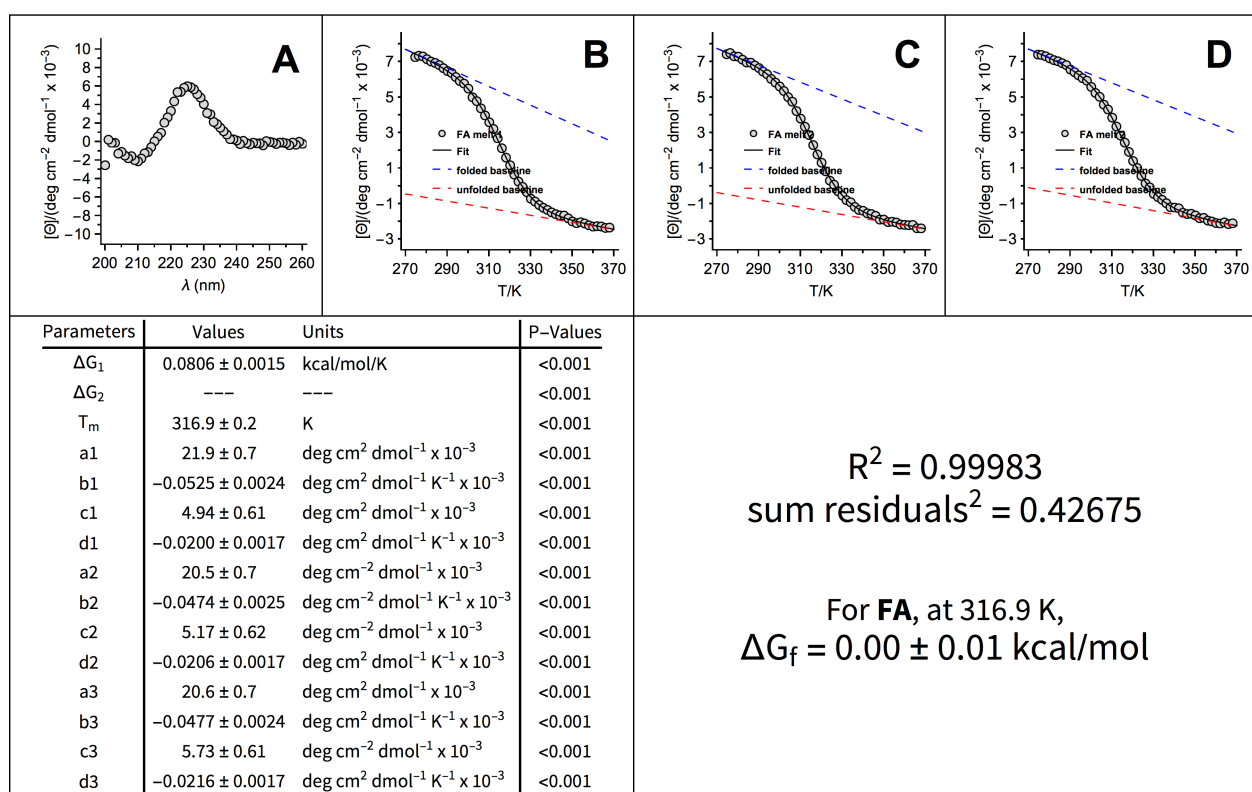


Figure S58. CD Data for FA (SD2176#1).

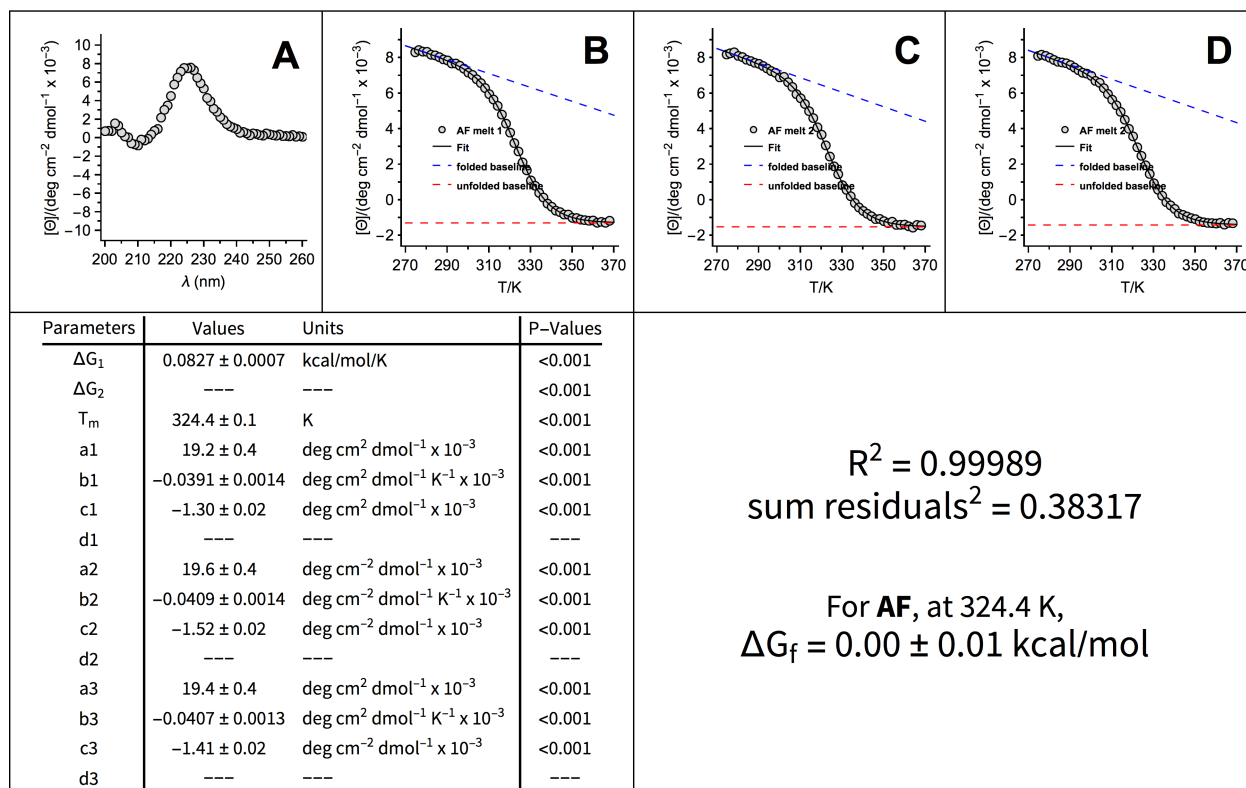


Figure S59. CD data for AF (DS2176#2).

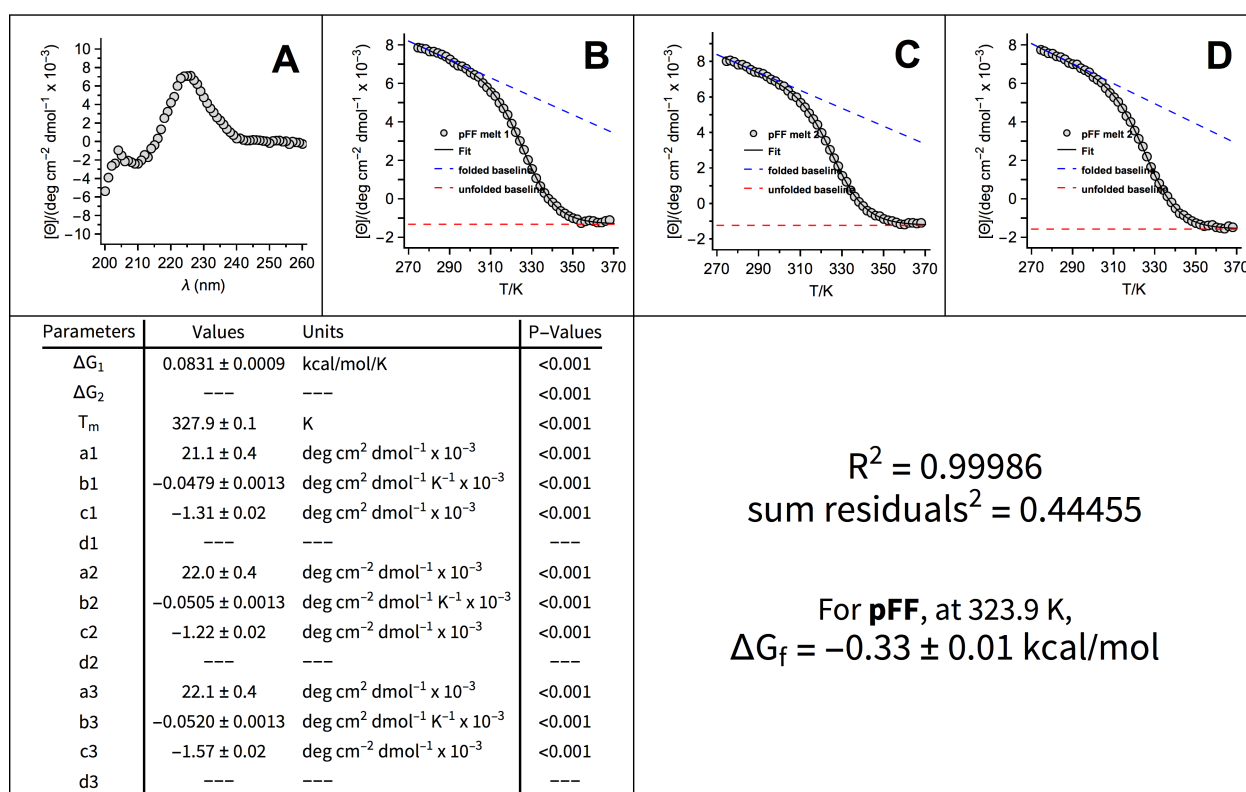


Figure S60. CD data for pFF (DA1037#3).

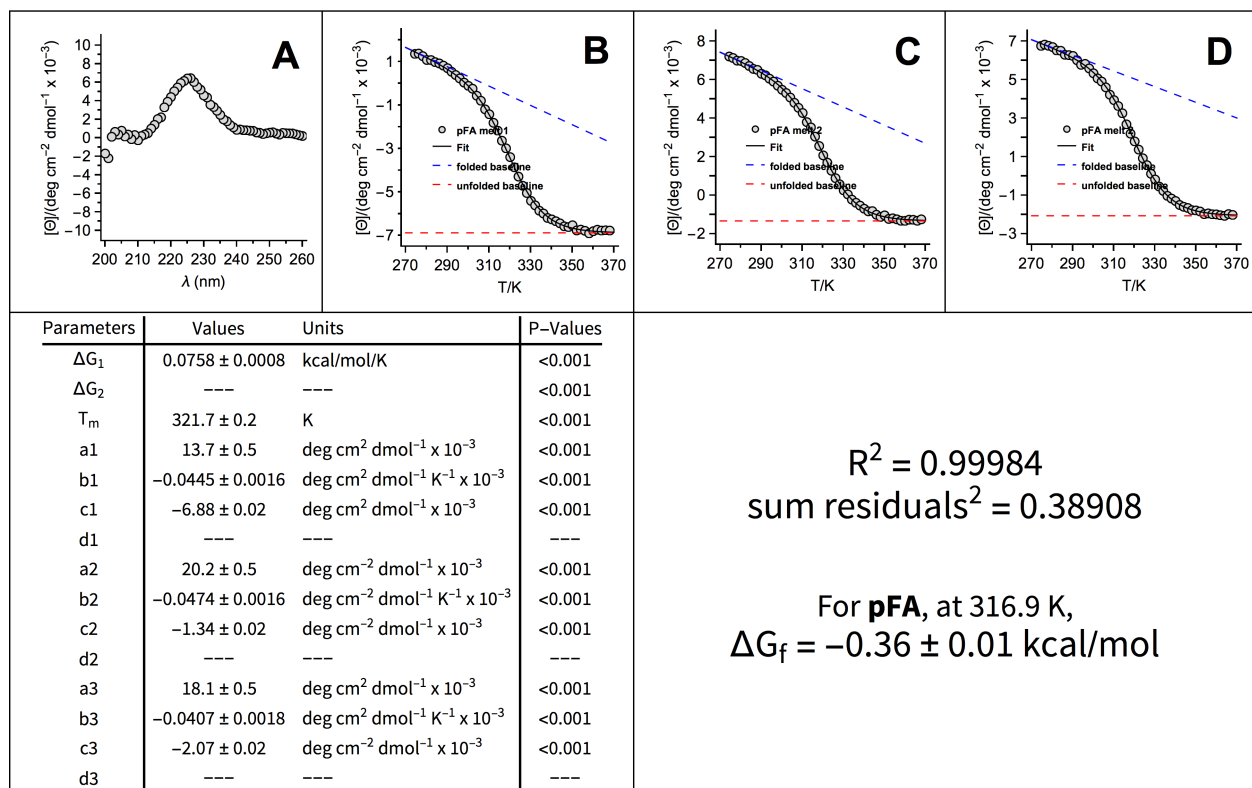


Figure S61. CD data for pFA (SD2176#1p).

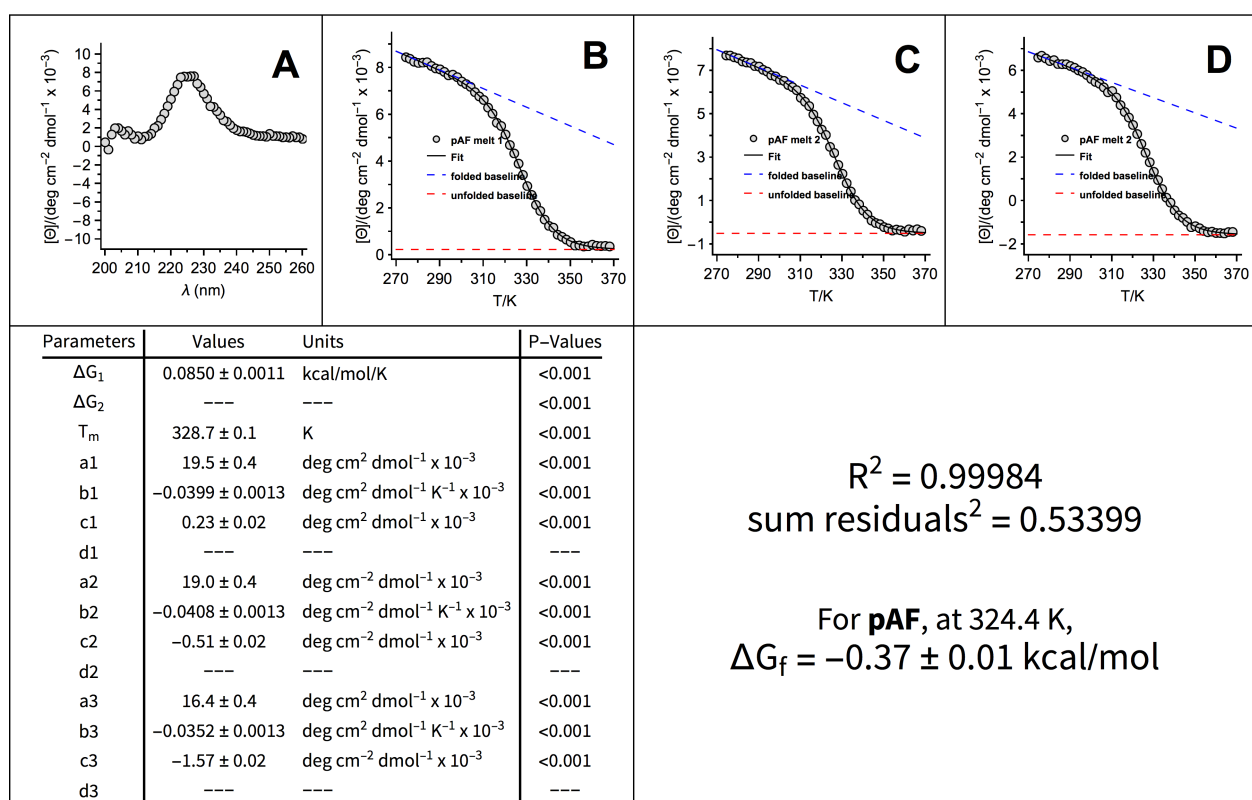


Figure S62. CD data for pAF (SD2176#2p).

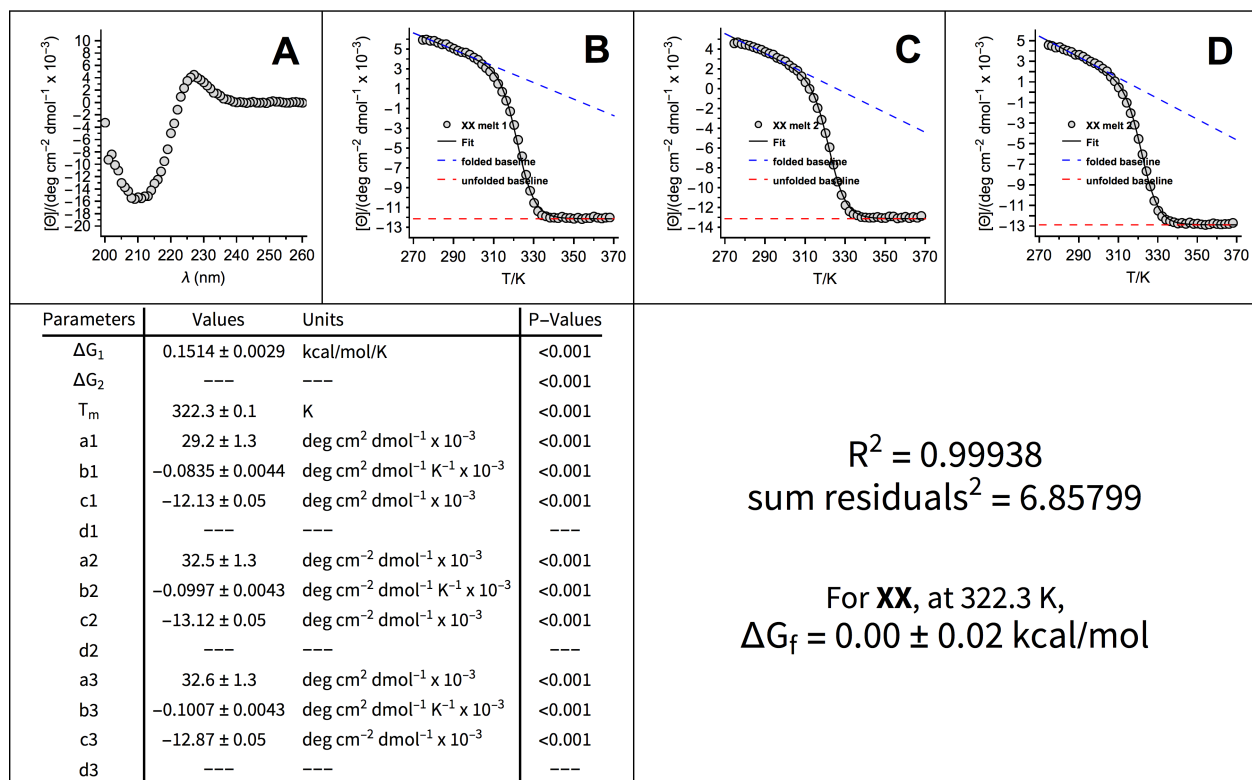


Figure S63. CD data for **XX** (DA1037#4).

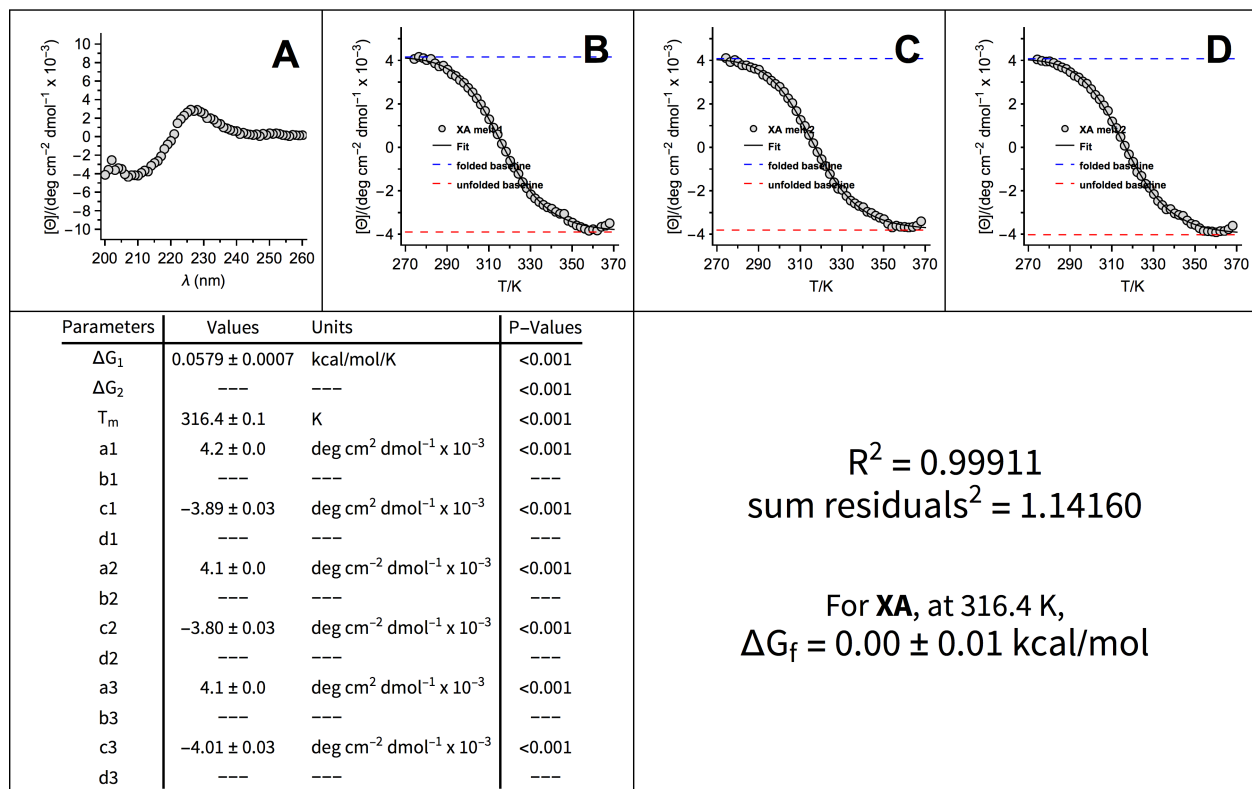


Figure S64. CD data for **XA** (SD2184#2).

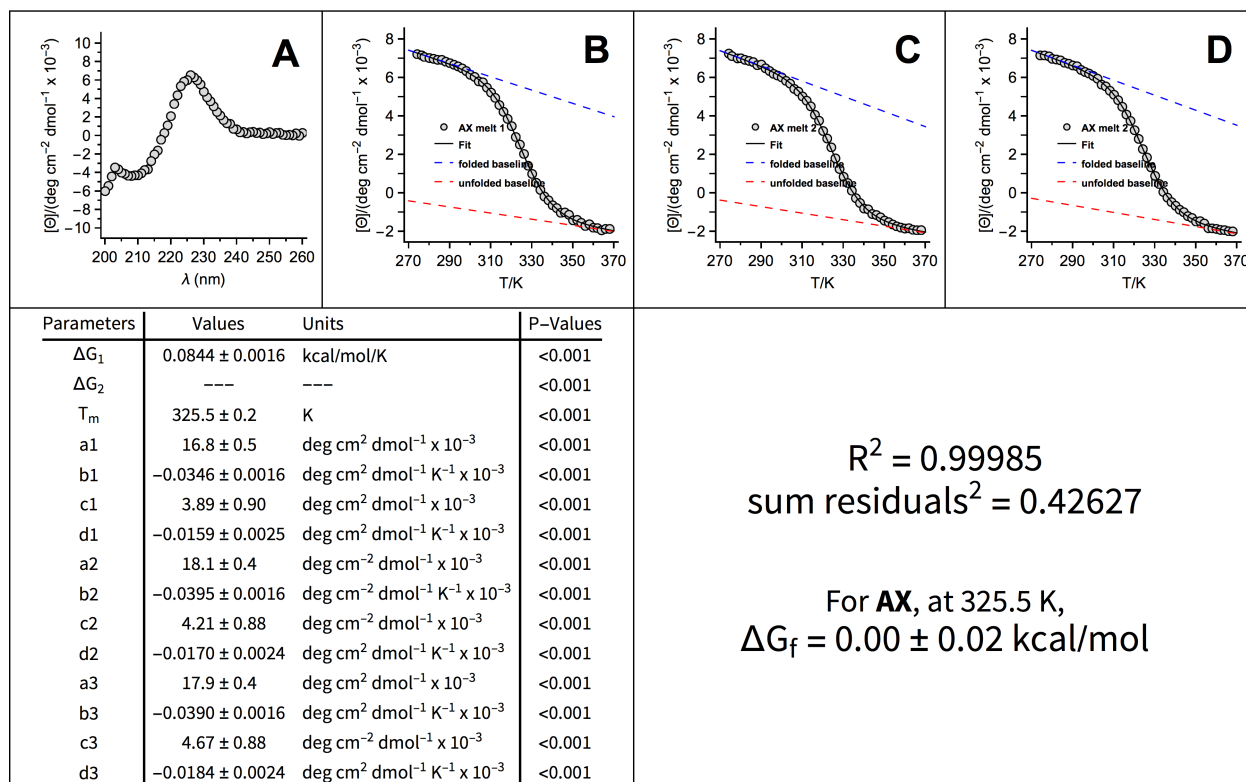


Figure S65. CD data for AX (SD2184#1).

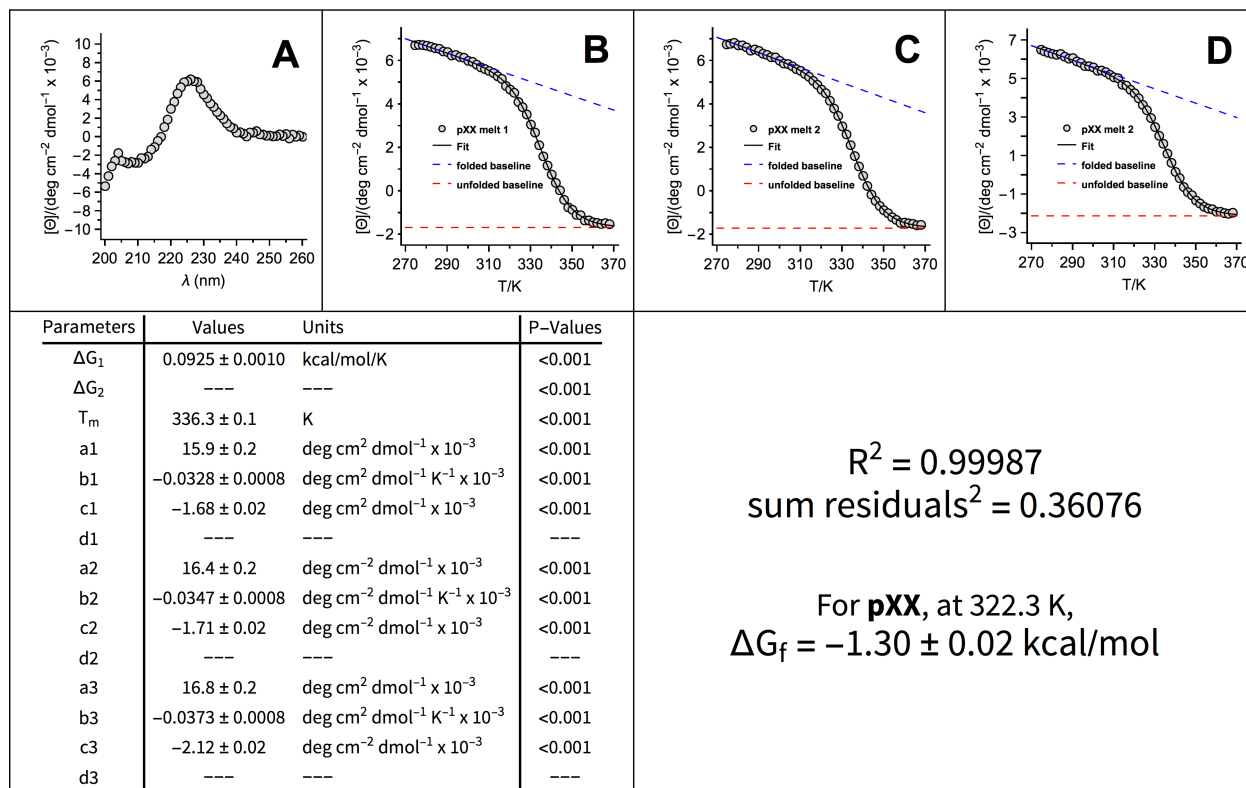


Figure S66. CD data for pXX (DA1037#4).

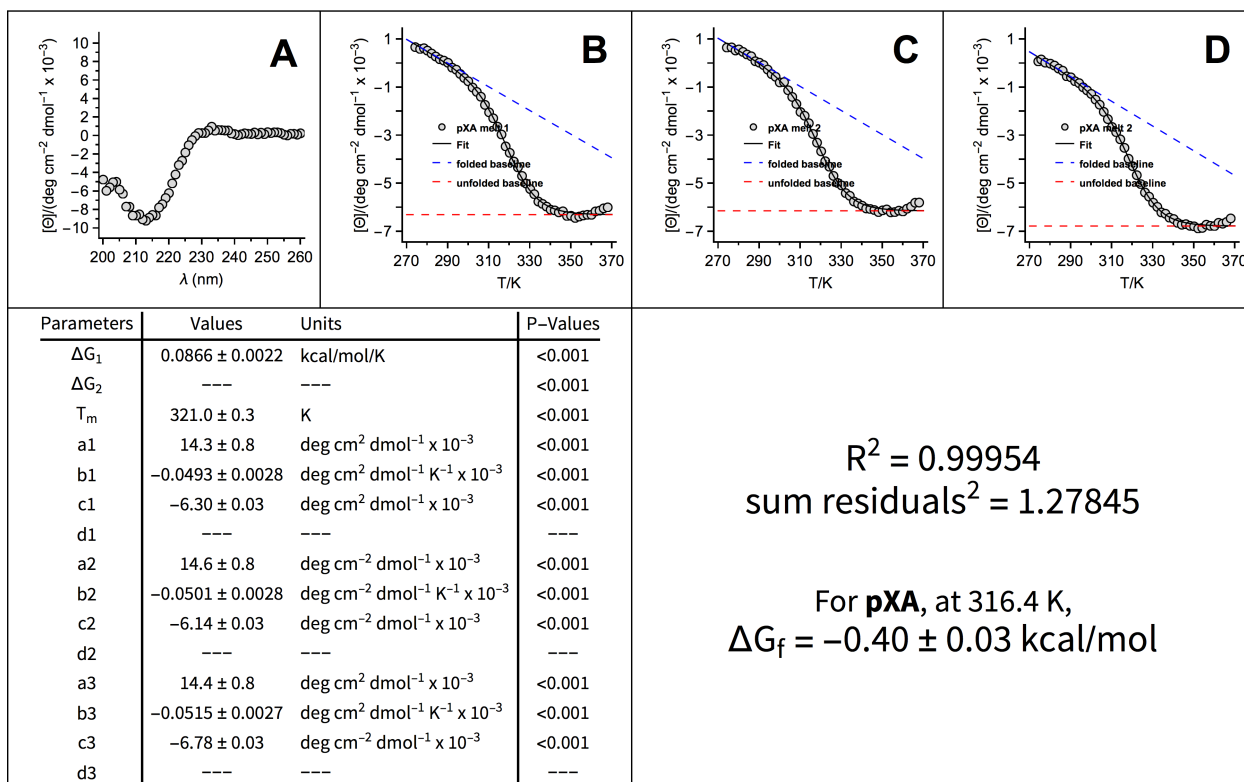


Figure S67. CD data for pXA (SD2184#2p).

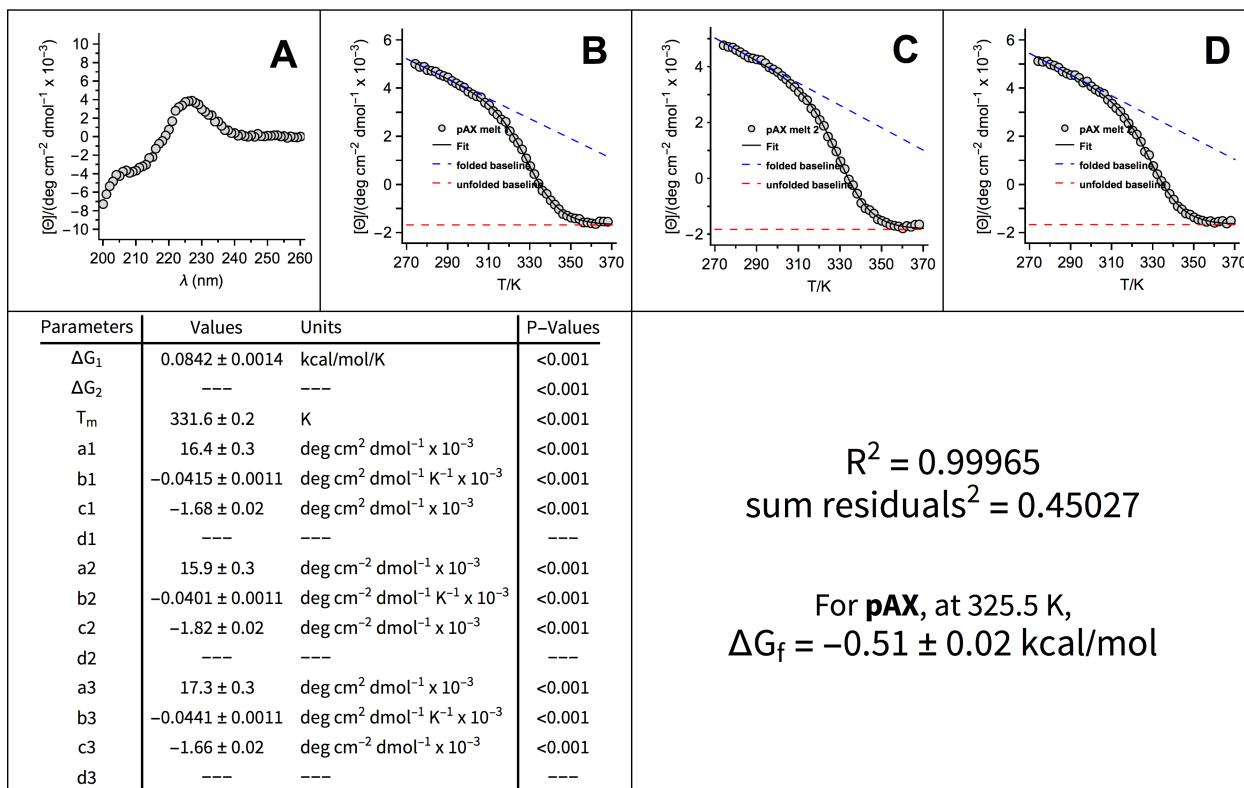


Figure S68. CD data for pAX (SD2184#1p).

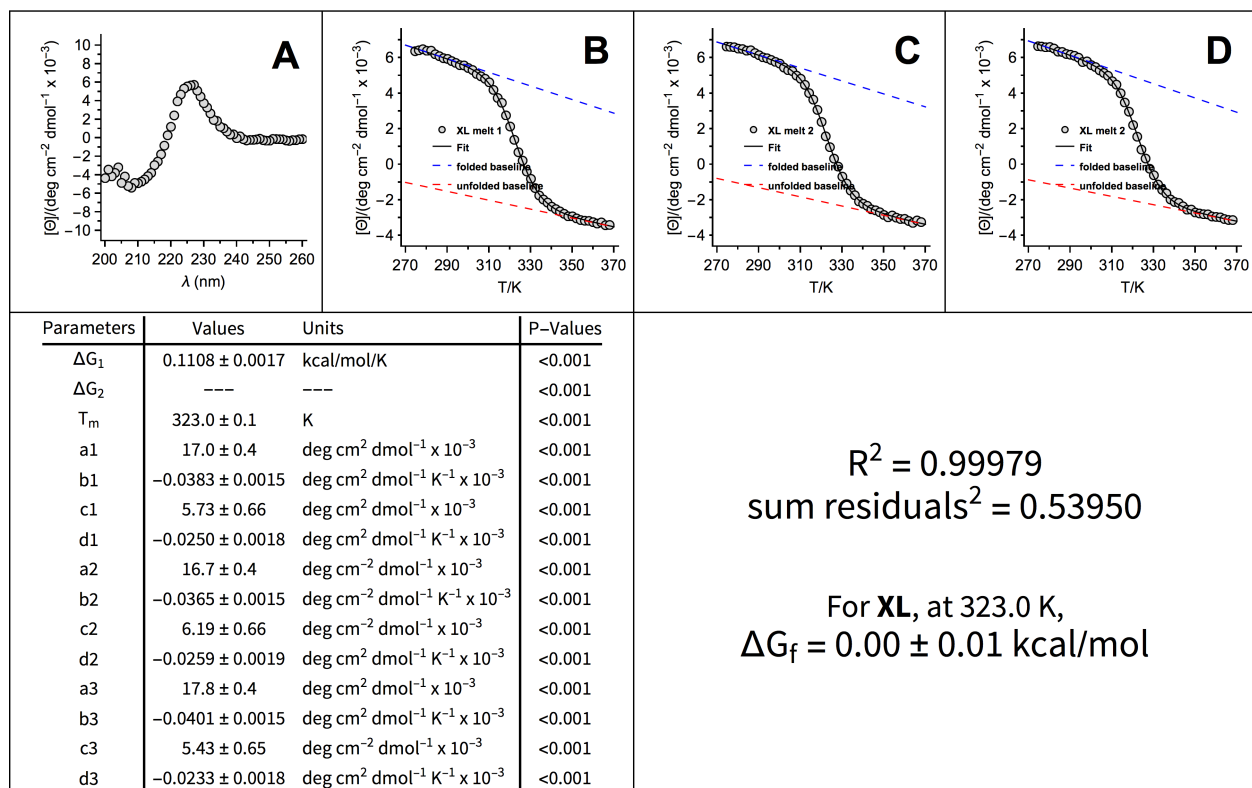


Figure S69. CD data for XL (BC1039#1).

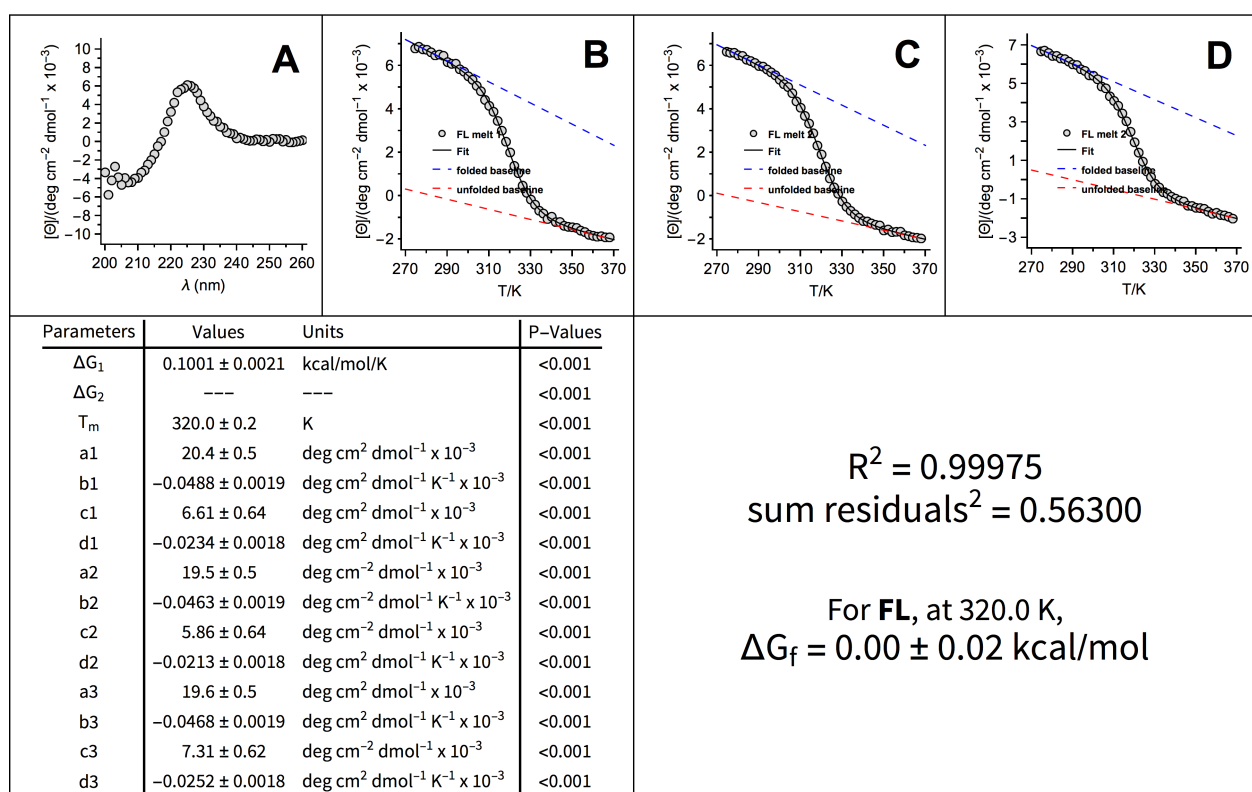


Figure S70. CD data for FL (BC1039#4).

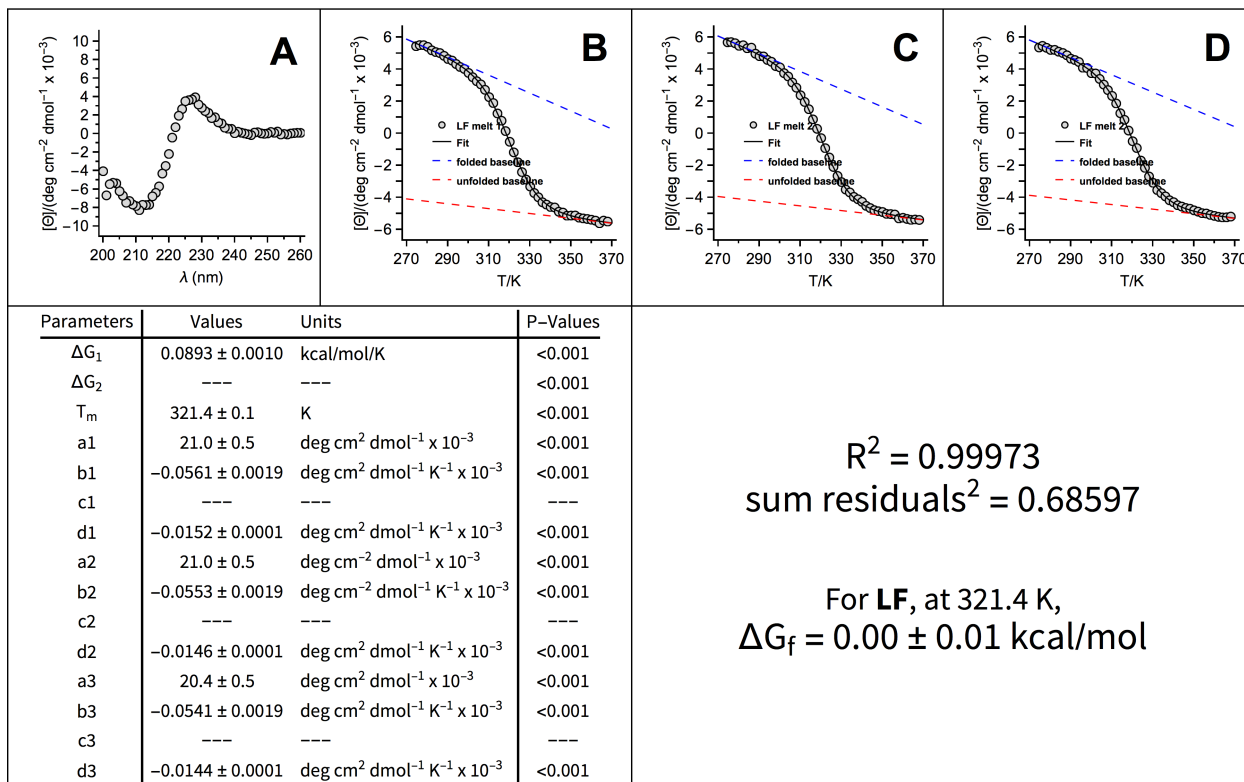


Figure S71. CD data for LF (BC1039#3).

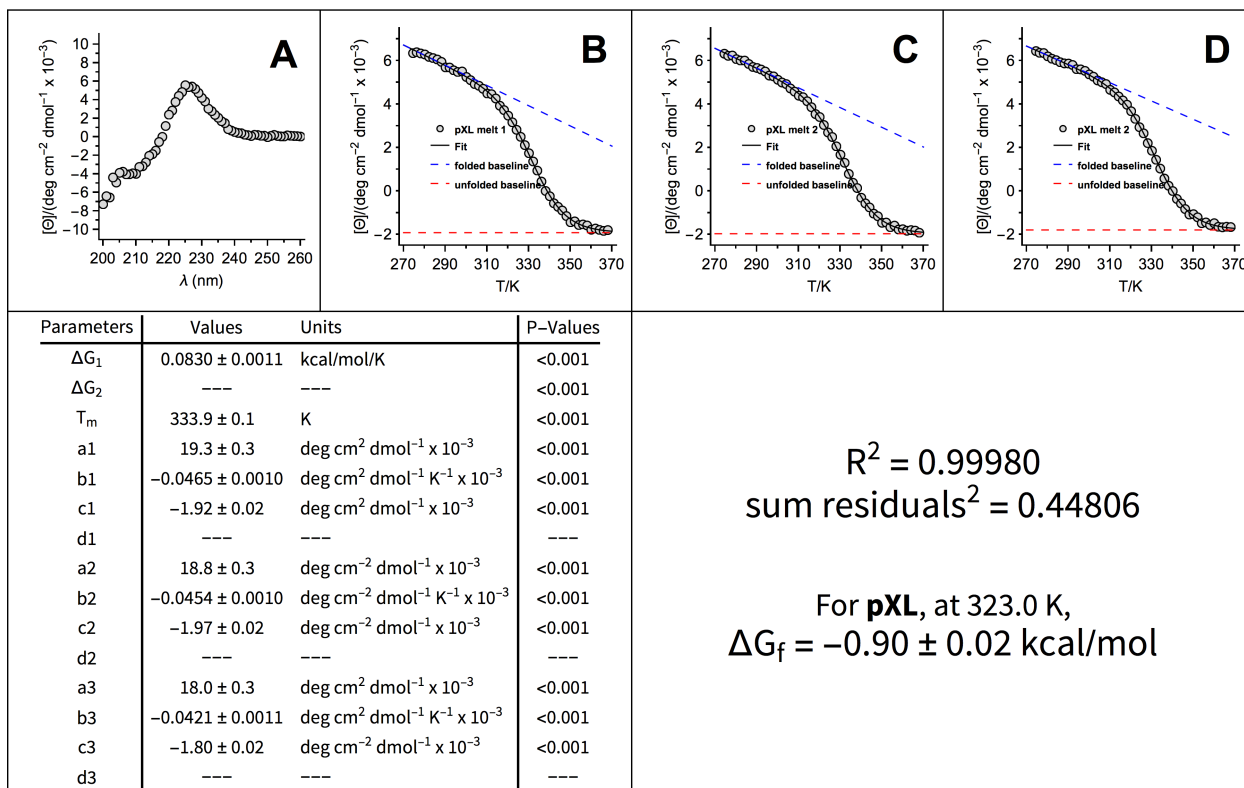


Figure S72. CD data for pXL (BC1039#1p).

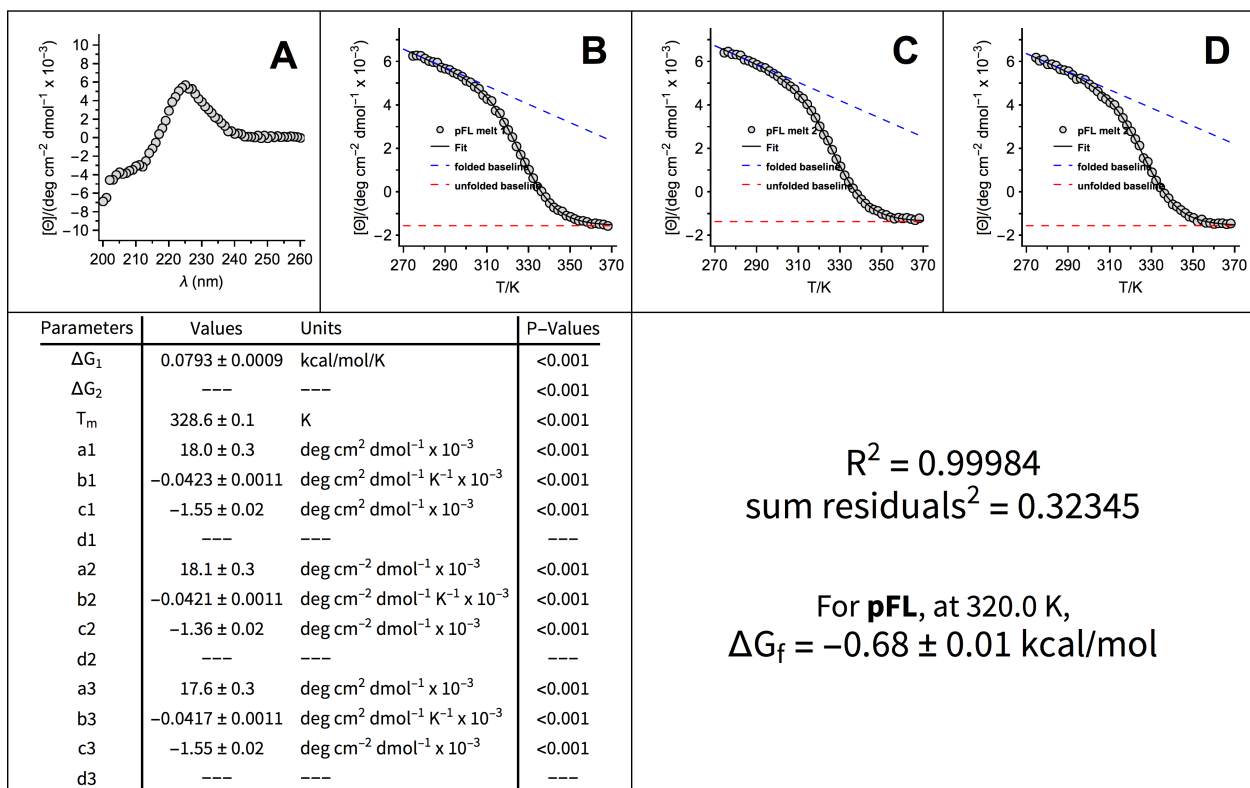


Figure S73. CD data for pFL (BC1039#4p).

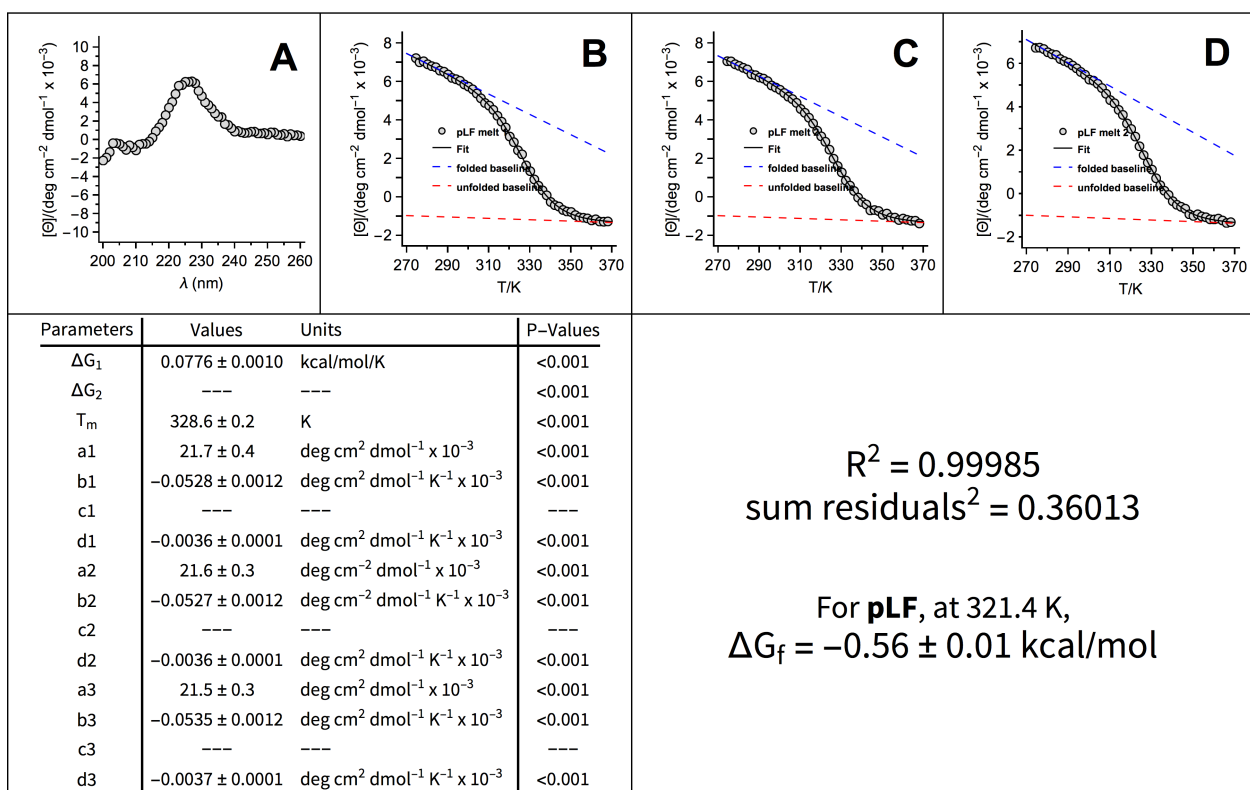


Figure S74. CD data for pLF (BC1039#3p).