

Supporting Information

Influence of PEGylation on the strength of nearby salt bridges

Qiang Xiao,^a Steven R. E. Draper,^a Mason S. Smith, Nathan P. Brown, Natalie A. B. Pugmire, Dallin S. Ashton, Anthony J. Carter, Eliza E. K. Lawrence, and Joshua L. Price*

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

^aThese authors contributed equally.

1. Protein synthesis, purification and characterization.....	3
2. Biophysical characterization of peptide variants.....	55
3. Crystallographic characterization of $\alpha 18$ and $p\alpha 18$	89
4. References.....	90

1. Protein Synthesis, Purification and Characterization

All peptide variants were synthesized as C-terminal acids by Fmoc-based solid-phase peptide synthesis as described previously.¹ $\beta 18$, $p\beta 18$, $\beta 18$ -SR, $p\beta 18$ -SR and $\beta 23$ and $p\beta 23$ were synthesized previously.^{1,2} Peptides with carboxyl termini were synthesized on Fmoc-Gly-Wang resin (EMD Biosciences) and those with amidated termini were synthesized on Rink amide resin (Advanced Chem Tech). We used standard Fmoc-protected amino acids with acid-labile sidechain protecting groups. We used previously synthesized Fmoc-L-GlnPEG4-OH [18-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-15-oxo-2,5,8,11-tetraoxa-14-azanonadecan-19-oic acid],¹ previously synthesized Fmoc-L-PrF-OH N-[(9H-Fluoren-9-ylmethoxy)-O-2-propyn-1-yl-L-tyrosine,³ and previously synthesized PEG-azide 13-azido-2,5,8,11-tetraoxatridecane⁴ for PEGylating the PrF peptides via the copper (I) catalyzed azide-alkyne cycloaddition.¹ For the asparagine PEG peptides we used previously synthesized Fmoc-L-(AsnPEG4)2-OH [(S)-17-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-15-oxo-14-(2,5,8,11-tetraoxatridecan-13-yl)-2,5,8,11-tetraoxa-14-azaoctadecan-18-oic acid].¹ Other reagents we used were: 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and N-hydroxybenzotriazole hydrate (HOBT) from Advanced ChemTech for amino acid activation; 20% piperidine in N,N-dimethylformamide for removal of the Fmoc protecting group from the N-terminal α - amine; a solution of a solution of phenol (0.0625 g), water (62.5 μ L), thioanisole (62.5 μ L), ethanedithiol (31 μ L) and triisopropylsilane (12.5 μ L) in trifluoroacetic acid (TFA, 1 mL)

for cleaving the protein from resin and globally removing acid-labile side-chain protecting groups. Proteins were precipitated from the TFA solution by addition of diethyl ether (~40 mL). Following centrifugation, the ether was decanted, and the pellet was dissolved in ~40mL 1:1 H₂O/MeCN, then flash frozen over dry ice in acetone and lyophilized to remove volatile impurities. The resulting powder was stored at -20°C until purification.

Proteins were purified by preparative reverse-phase high performance liquid chromatography (HPLC) on a C18 column using a linear gradient of water in acetonitrile with 0.1% v/v TFA. Fractions containing the desired protein product were pooled, frozen, and lyophilized. Proteins were identified by electrospray ionization time of flight mass spectrometry (ESI-TOF); expected and observed exact masses mass spectra appear in Table S1, S2 and S3 and spectra appear in Figures S1-S63. Protein purity was assessed by Analytical HPLC (Figures S64-S127).

Table S2. Sequences, molecular formulas, expected and observed m/z ratios for 1CW variants.

Name	Sequence	Molecular Formula	z	Expected [M+z·H]/z	Observed [M+z·H]/z
3a1	Ac-QV EALEK KVAAL ES KVQALEKKVEALEY-NH ₂	C ₁₄₃ H ₂₄₂ N ₃₆ O ₄₅	3	1062.267	1062.264
3a1-ER	Ac-QV EALER-NH ₂	C ₁₄₃ H ₂₄₂ N ₃₈ O ₄₅	3	1071.602	1071.615
3a1-EA	Ac-QV EALEA-NH ₂	C ₁₄₀ H ₂₃₅ N ₃₅ O ₄₅	3	1043.247	1043.245
3a1-AK	Ac-QV AALEK-NH ₂	C ₁₄₁ H ₂₄₀ N ₃₆ O ₄₃	3	1042.931	1042.929
3a1-AR	Ac-QV AALER-NH ₂	C ₁₄₁ H ₂₄₀ N ₃₈ O ₄₃	3	1052.267	1052.280
3a1-AA	Ac-QV AALEA-NH ₂	C ₁₃₈ H ₂₃₃ N ₃₅ O ₄₃	3	1023.912	1023.908
p3a1	Ac-QV EALEK-NH ₂	C ₁₅₂ H ₂₆₀ N ₃₆ O ₄₉	3	1125.640	1125.636
p3a1-ER	Ac-QV EALER-NH ₂	C ₁₅₂ H ₂₆₀ N ₃₈ O ₄₉	3	1134.976	1134.986
p3a1-EA	Ac-QV EALEA-NH ₂	C ₁₄₉ H ₂₅₃ N ₃₅ O ₄₉	3	1106.621	1106.616
p3a1-AK	Ac-QV AALEK-NH ₂	C ₁₅₀ H ₂₅₈ N ₃₆ O ₄₇	3	1106.305	1106.301
p3a1-AR	Ac-QV AALER-NH ₂	C ₁₅₀ H ₂₅₈ N ₃₈ O ₄₇	3	1115.640	1115.653
p3a1-AA	Ac-QV AALEA-NH ₂	C ₁₄₇ H ₂₅₁ N ₃₅ O ₄₇	3	1087.286	1087.284
3a6	Ac-EVEAL Q KVAAL ES KVQALEKKVEALEY-NH ₂	C ₁₄₃ H ₂₄₂ N ₃₆ O ₄₅	3	1062.267	1062.260
3a6-KA	Ac-..... Q KVAAL A-NH ₂	C ₁₄₁ H ₂₄₀ N ₃₆ O ₄₃	3	1042.931	1042.933
3a6-AE	Ac-..... Q KVAAL E-NH ₂	C ₁₄₀ H ₂₃₅ N ₃₅ O ₄₅	3	1043.247	1043.241
3a6-AA	Ac-..... Q KVAAL A-NH ₂	C ₁₃₈ H ₂₃₃ N ₃₅ O ₄₃	3	1023.912	1023.914
p3a6	Ac-..... Q KVAAL A-NH ₂	C ₁₅₂ H ₂₆₀ N ₃₆ O ₄₉	3	1125.640	1125.633
p3a6-KA	Ac-..... Q KVAAL E-NH ₂	C ₁₅₀ H ₂₅₈ N ₃₆ O ₄₇	3	1106.305	1106.290
p3a6-AE	Ac-..... Q KVAAL A-NH ₂	C ₁₄₉ H ₂₅₃ N ₃₅ O ₄₉	3	1106.621	1106.620
p3a6-AA	Ac-..... Q KVAAL A-NH ₂	C ₁₄₇ H ₂₅₁ N ₃₅ O ₄₇	3	1087.286	1087.287

Table S3. Sequences, molecular formulas, expected and observed m/z ratios for WW variants.

Name	Sequence	Molecular Formula	z	Expected [M+z·H]/z	Observed [M+z·H]/z
β18	H ₂ N-KLPPGW E KRMSR S SGRV X YFNHITNASQFERPSG-OH	C ₁₇₃ H ₂₅₇ N ₅₁ O ₄₉ S	4	967.231	967.226
β18-DA	H ₂ N-..... D ANG A-OH	C ₁₇₀ H ₂₅₀ N ₄₈ O ₄₉ S	4	945.965	945.948
β18-SR	H ₂ N-..... S ANG R-OH	C ₁₇₂ H ₂₅₇ N ₅₁ O ₄₈ S	3	1279.973	1279.973
β18-SA	H ₂ N-..... S ANG A-OH	C ₁₆₉ H ₂₅₀ N ₄₈ O ₄₈ S	3	1251.619	1251.607
pβ18	H ₂ N-..... D ANG R-OH	C ₁₈₂ H ₂₇₅ N ₅₁ O ₅₃ S	4	1014.761	1014.747
pβ18-DA	H ₂ N-..... D ANG A-OH	C ₁₇₉ H ₂₆₈ N ₄₈ O ₅₃ S	4	993.495	993.483
pβ18-SR	H ₂ N-..... S ANG R-OH	C ₁₈₁ H ₂₇₅ N ₅₁ O ₅₂ S	4	1007.762	1007.751
pβ18-SA	H ₂ N-..... S ANG A-OH	C ₁₇₈ H ₂₆₈ N ₄₈ O ₅₂ S	4	986.496	986.482
β23	H ₂ N-KLPPGW E KRMSR S SGRV X YFNHITNASQFERPSG-OH	---	---	---	---
β23-EA	H ₂ N-..... E KAMSRSSGRV X-OH	C ₁₇₇ H ₂₆₃ N ₅₁ O ₅₀ S	4	984.741	984.735
β23-AR	H ₂ N-..... A KRMSRSSGRV X-OH	C ₁₇₈ H ₂₆₈ N ₅₄ O ₄₈ S	4	991.506	991.503
β23-AA	H ₂ N-..... A KAMSRSSGRV X-OH	C ₁₇₅ H ₂₆₁ N ₅₁ O ₄₈ S	4	970.240	970.232
pβ23	H ₂ N-..... E KRMSRSSGRV X-OH	---	---	---	---
pβ23-EA	H ₂ N-..... E KAMSRSSGRV X-OH	C ₁₈₆ H ₂₈₂ N ₅₄ O ₅₄ S	4	1043.026	1043.020
pβ23-AR	H ₂ N-..... A KRMSRSSGRV X-OH	C ₁₈₇ H ₂₈₇ N ₅₇ O ₅₂ S	4	1049.790	1049.787
pβ23-AA	H ₂ N-..... A KAMSRSSGRV X-OH	C ₁₈₄ H ₂₈₀ N ₅₄ O ₅₂ S	4	1028.524	1028.517

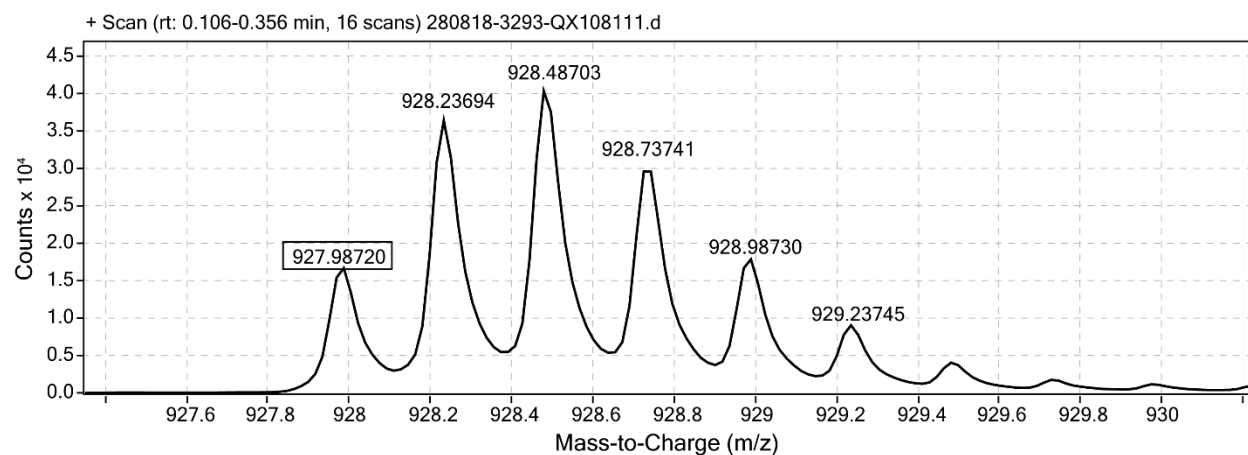


Figure S1. ESI-TOF MS data for **2a1** (QX108111). Expected $[M+4H^+]/4 = 928.009$.

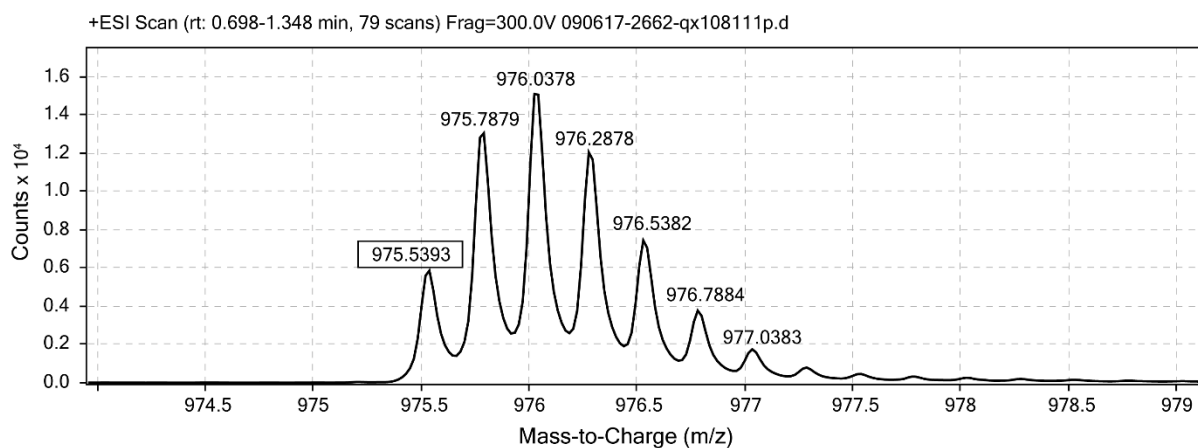


Figure S2. ESI-TOF MS data for **p2a1** (QX108111p). Expected $[M+4H^+]/4 = 975.539$.

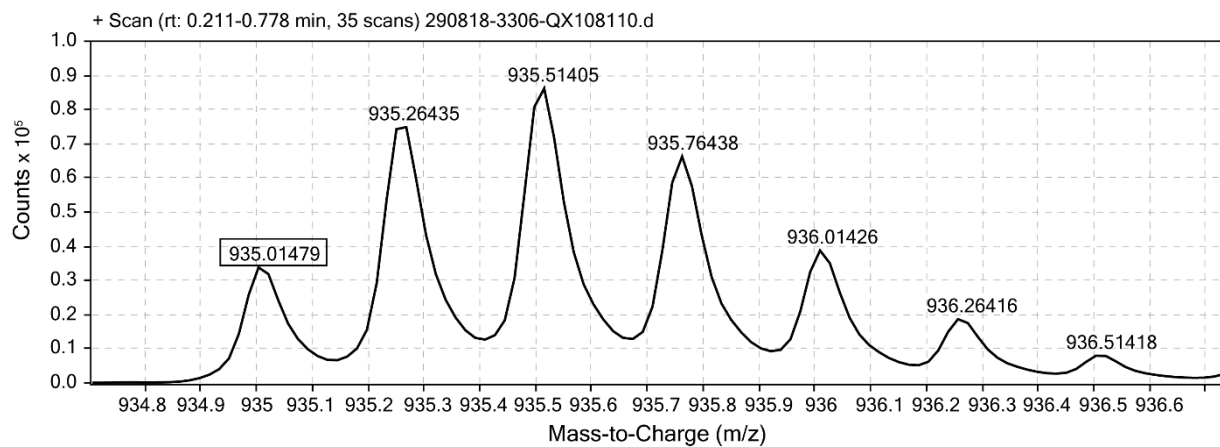


Figure S3. ESI-TOF MS data for **2a3** (QX108110). Expected $[M+4H^+]/4 = 935.010$.

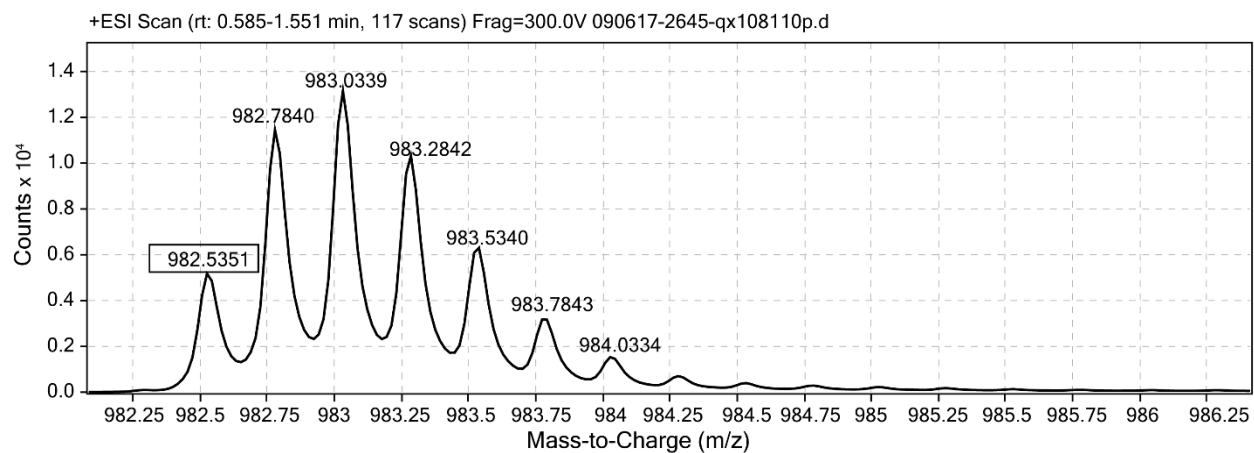


Figure S4. ESI-TOF MS data for **p2a3** (QX108110p). Expected $[M+4H^+]/4 = 982.541$.

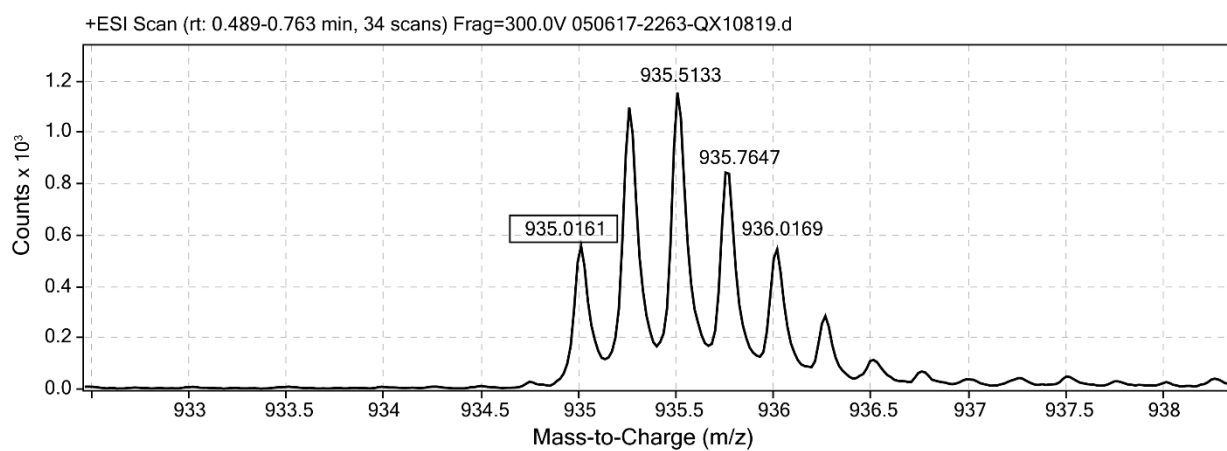


Figure S5. ESI-TOF MS data for **2a4** (QX10819). Expected $[M+4H^+]/4 = 935.020$.

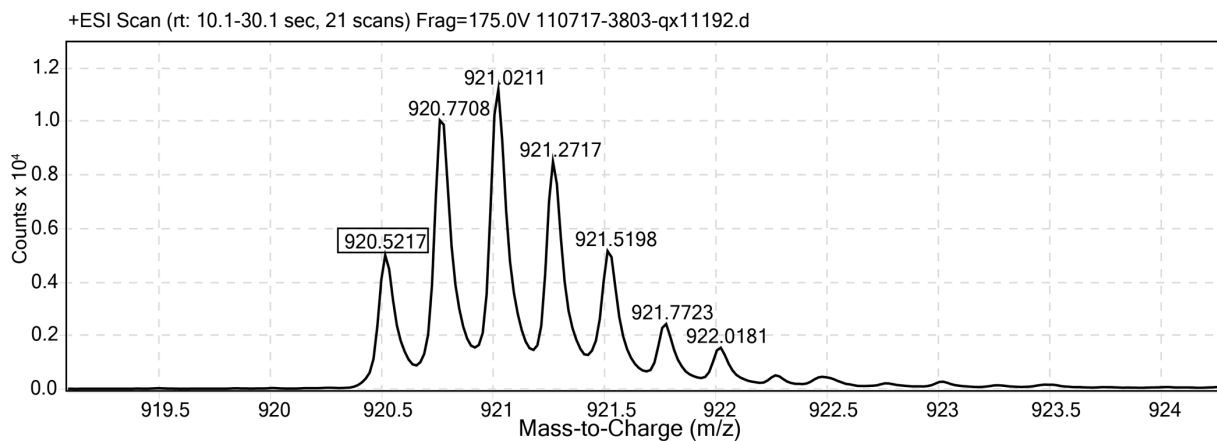


Figure S6. ESI-TOF MS data for **2a4-KA** (QX11192). Expected $[M+4H^+]/4 = 920.518$.

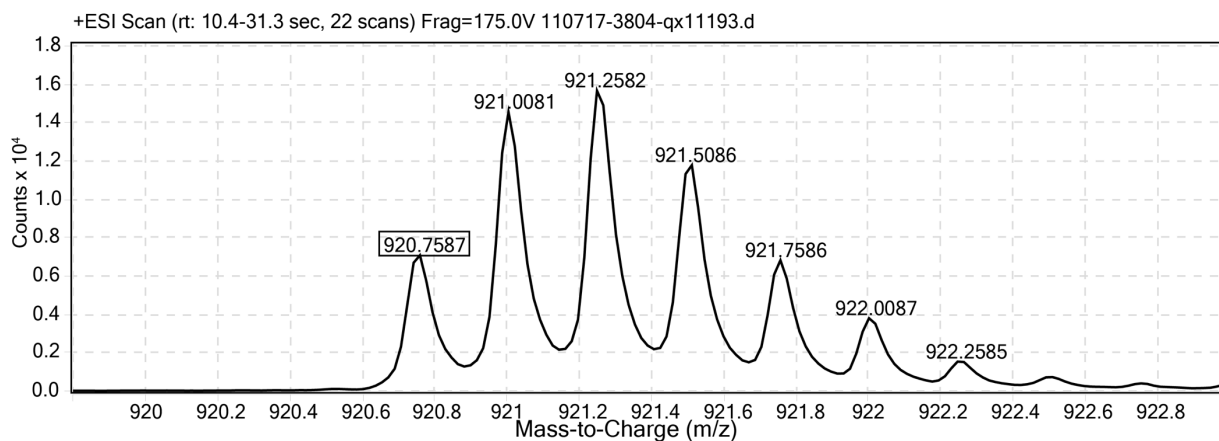


Figure S7. ESI-TOF MS data for **2a4-AE** (QX11193). Expected $[M+4H^+]/4 = 920.755$.

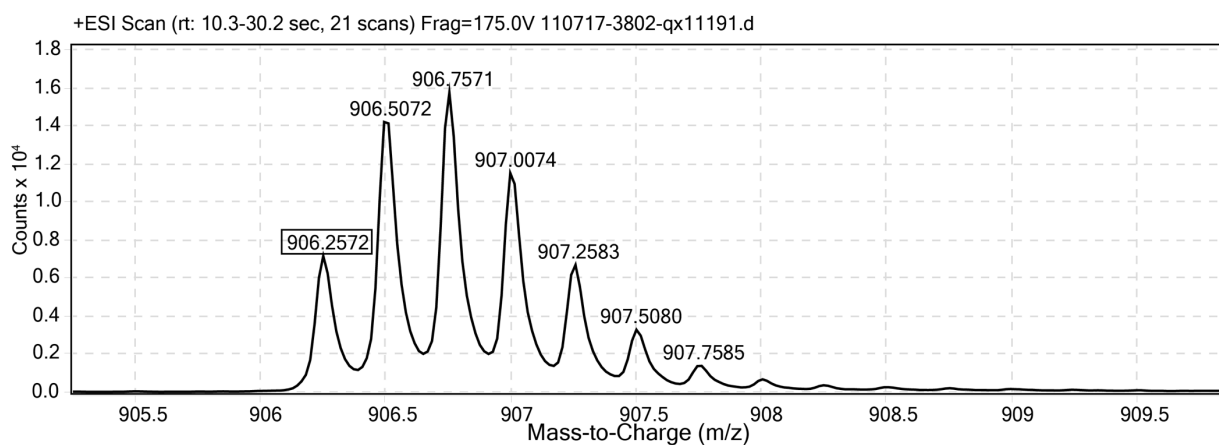


Figure S8. ESI-TOF MS data for **2a4-AA** (QX11191). Expected $[M+4H^+]/4 = 906.254$.

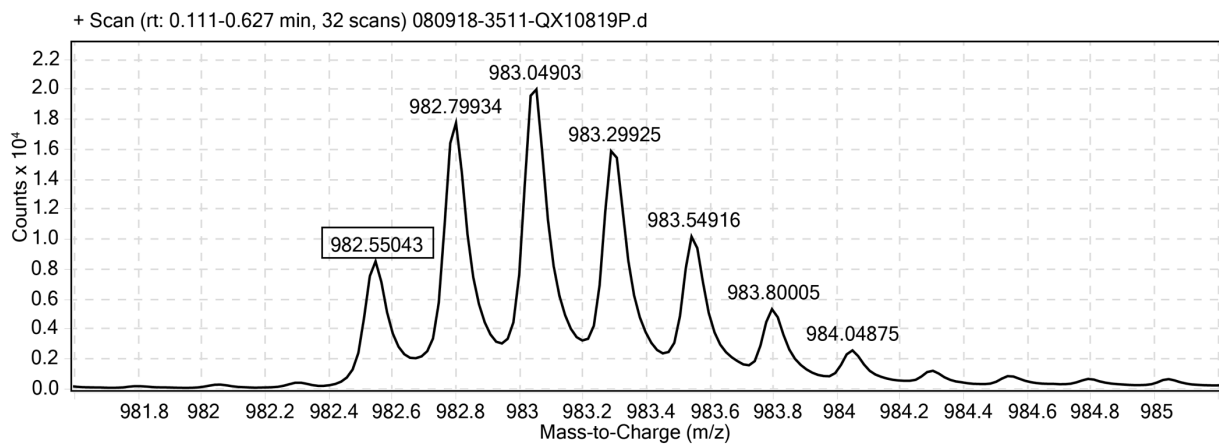


Figure S9. ESI-TOF MS data for **p2a4** (QX10819p). Expected $[M+4H^+]/4 = 982.550$.

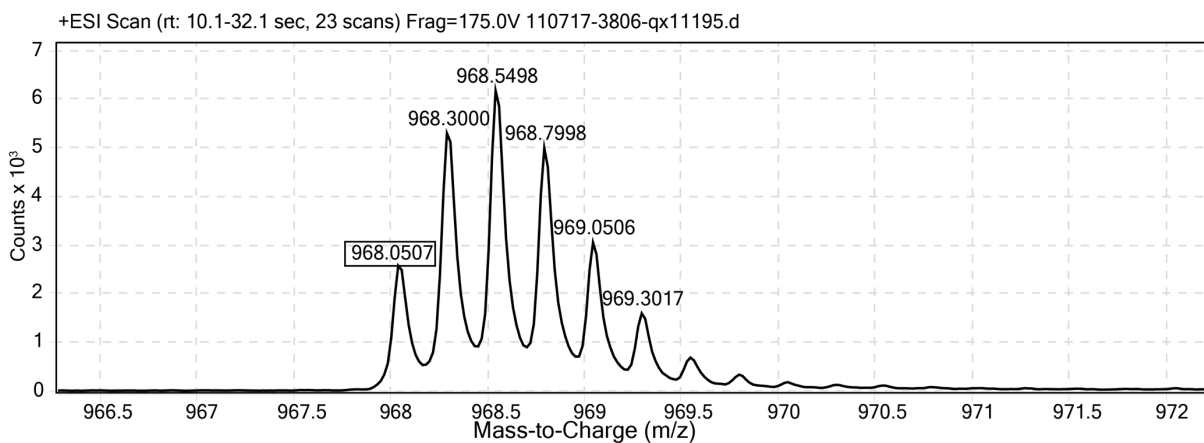


Figure S10. ESI-TOF MS data for **p2a4-KA** (QX11195). Expected $[M+4H^+]/4 = 968.051$.

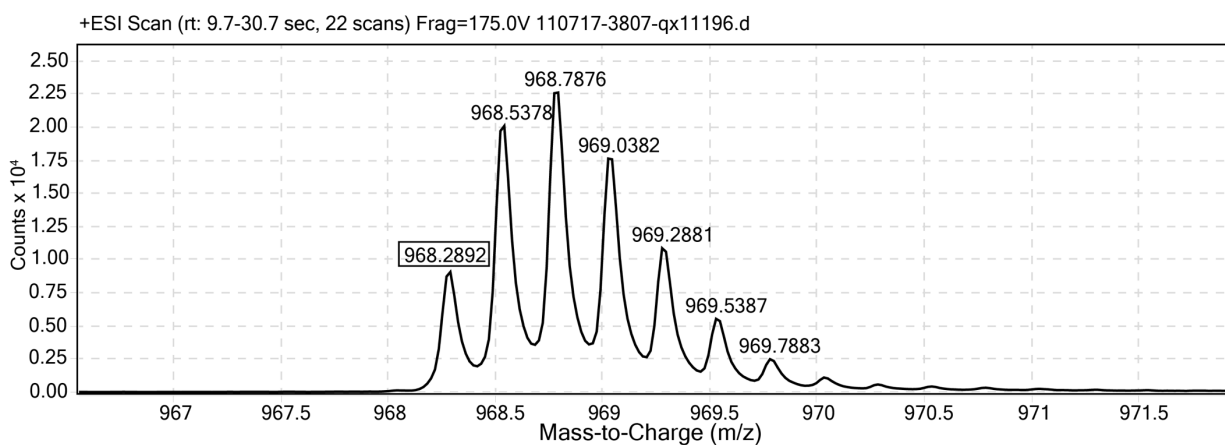


Figure S11. ESI-TOF MS data for **p2a4-AE** (QX11196). Expected $[M+4H^+]/4 = 968.285$.

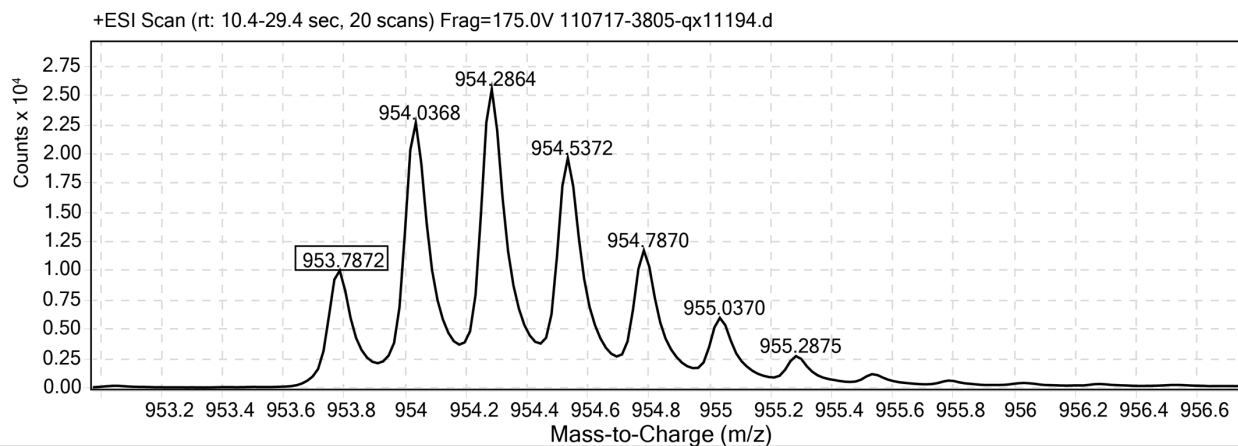


Figure S12. ESI-TOF MS data for **p2a4-AA** (QX11194). Expected $[M+4H^+]/4 = 953.787$.

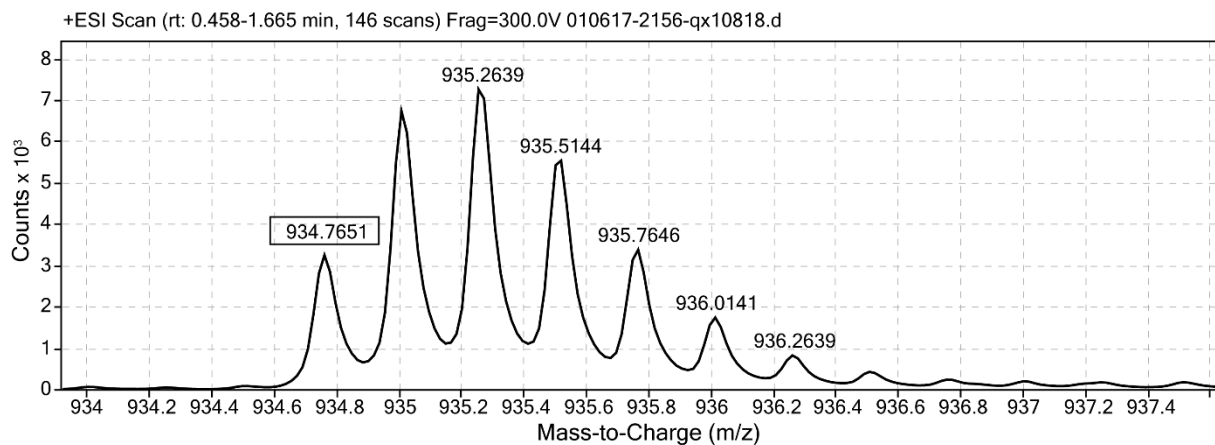


Figure S13. ESI-TOF MS data for **2a6** (QX10818). Expected $[M+4H^+]/4 = 934.774$.

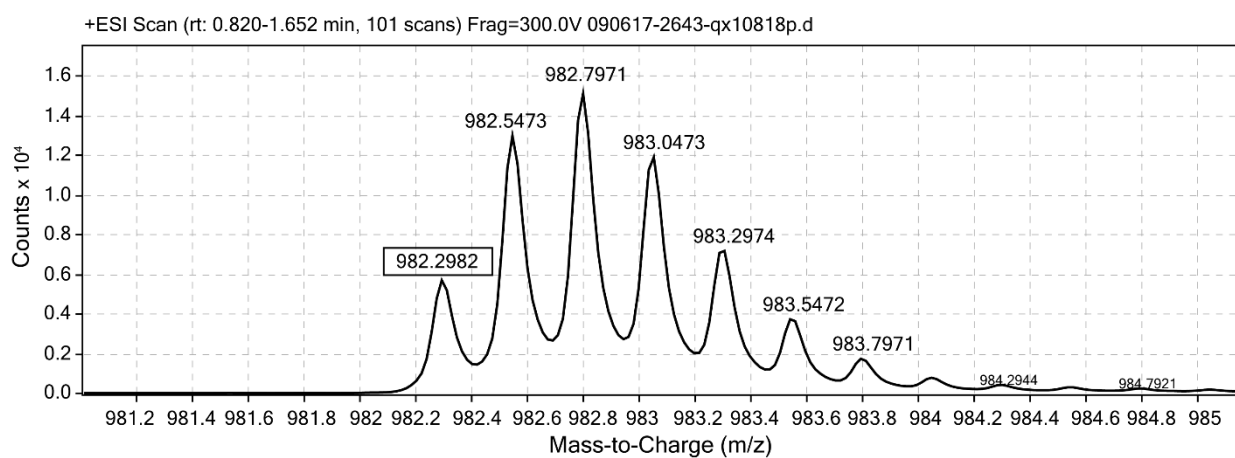


Figure S14. ESI-TOF MS data for **p2a6** (QX10818p). Expected $[M+4H^+]/4 = 982.304$.

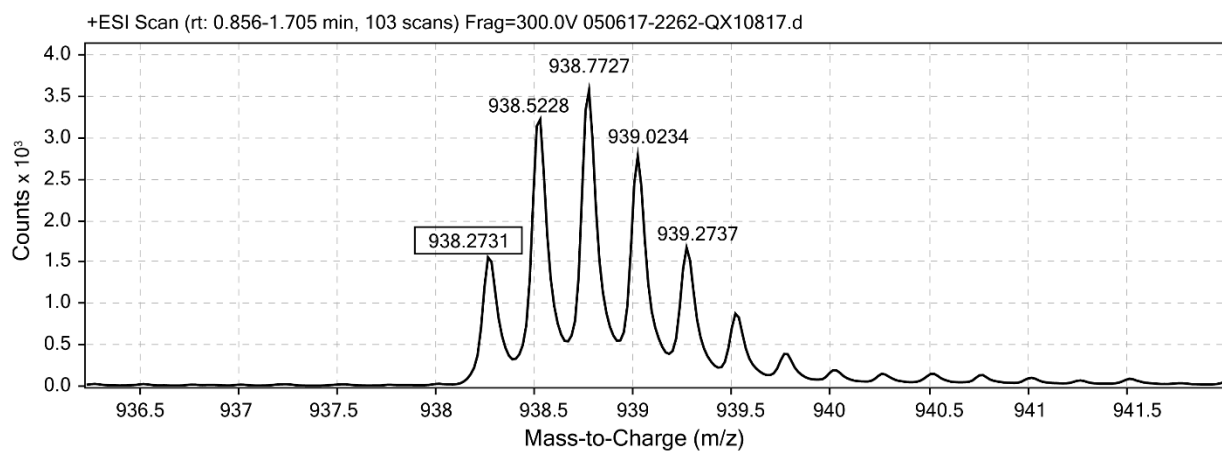


Figure S15. ESI-TOF MS data for **2a7** (QX10817). Expected $[M+4H^+]/4 = 938.277$.

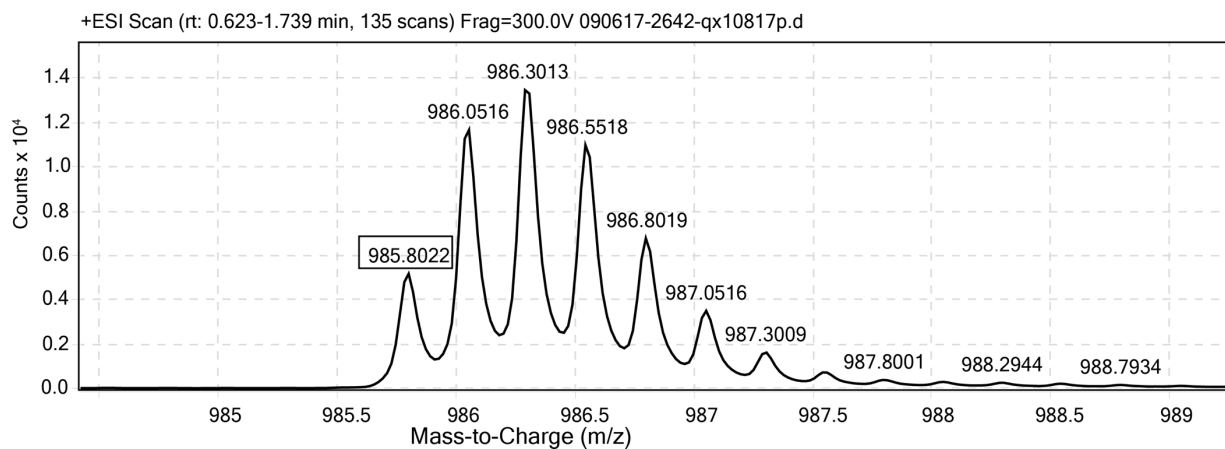


Figure S16. ESI-TOF MS data for **p2a7** (QX10817p). Expected $[M+4H^+]/4 = 985.808$.

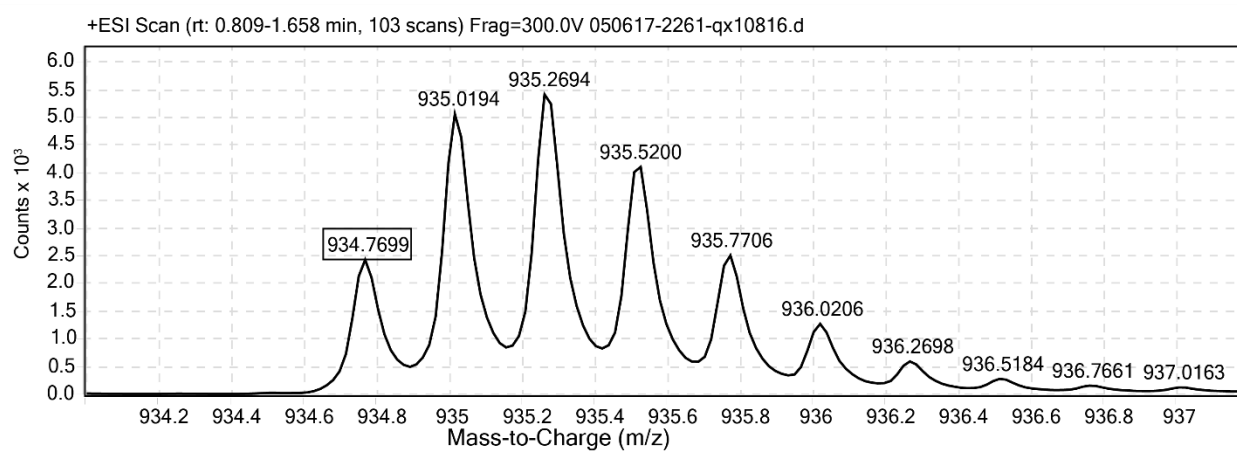


Figure S17. ESI-TOF MS data for **2a10** (QX10816). Expected $[M+4H^+]/4 = 934.774$.

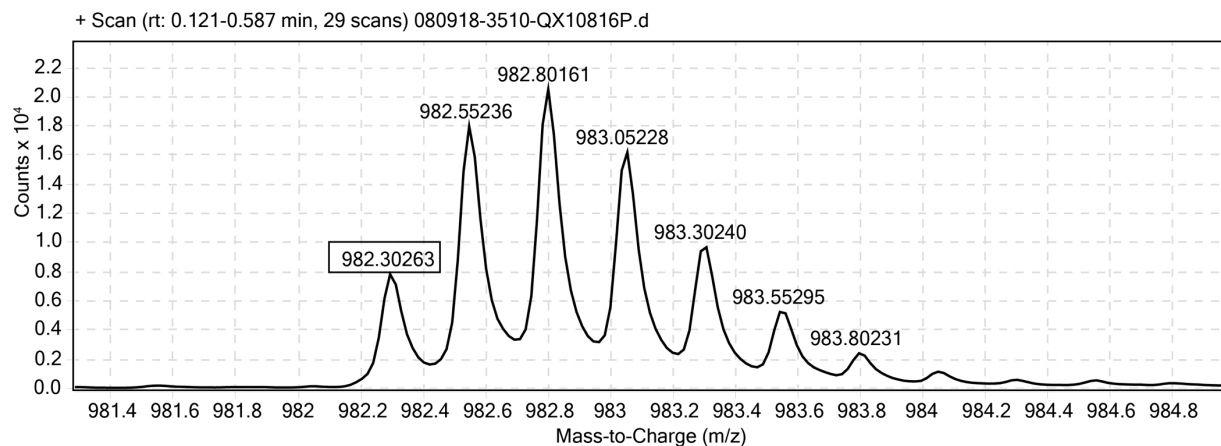


Figure S18. ESI-TOF MS data for **p2a10** (QX10816p). Expected $[M+4H^+]/4 = 982.304$.

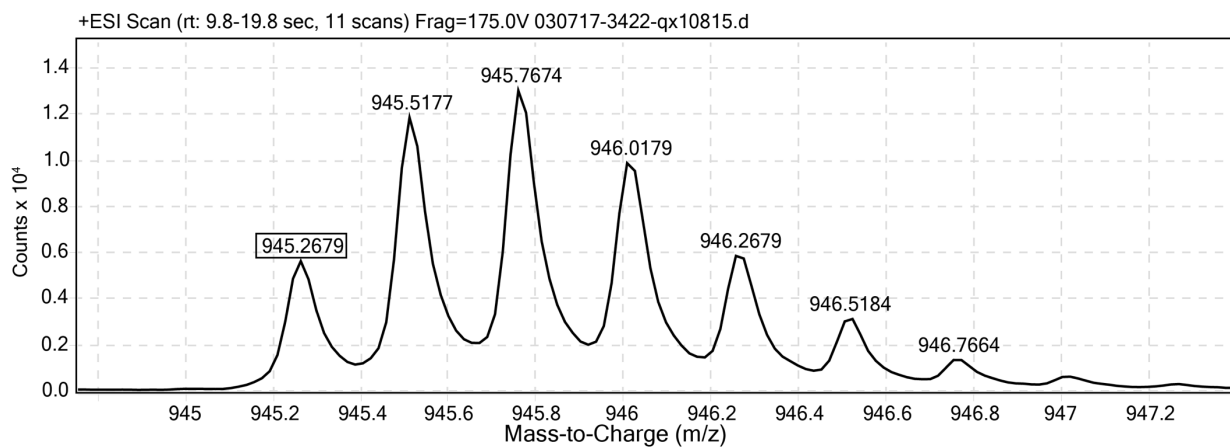


Figure S19. ESI-TOF MS data for **2a14** (QX10815). Expected $[M+4H^+]/4 = 945.276$.

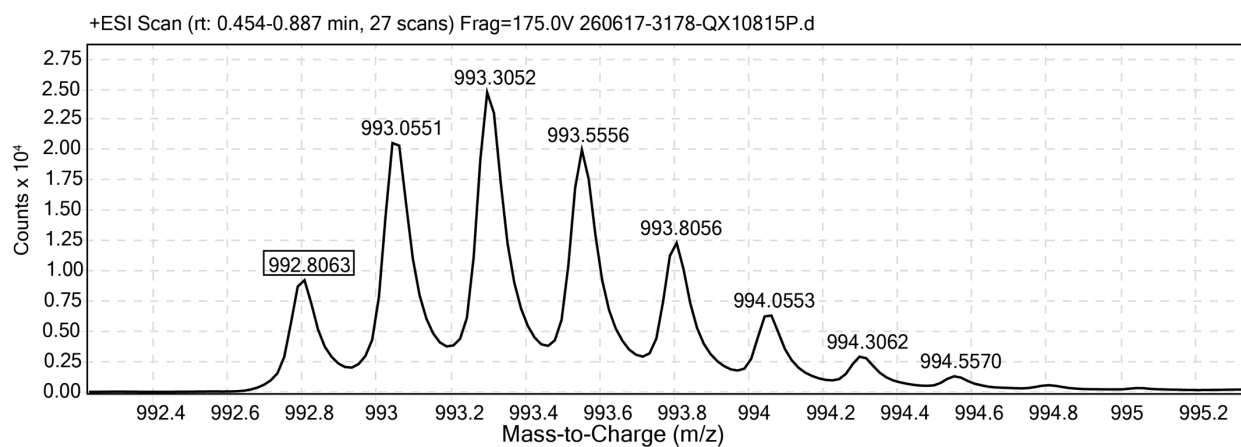


Figure S20. ESI-TOF MS data for **p2a14** (QX10815p). Expected $[M+4H^+]/4 = 992.806$.

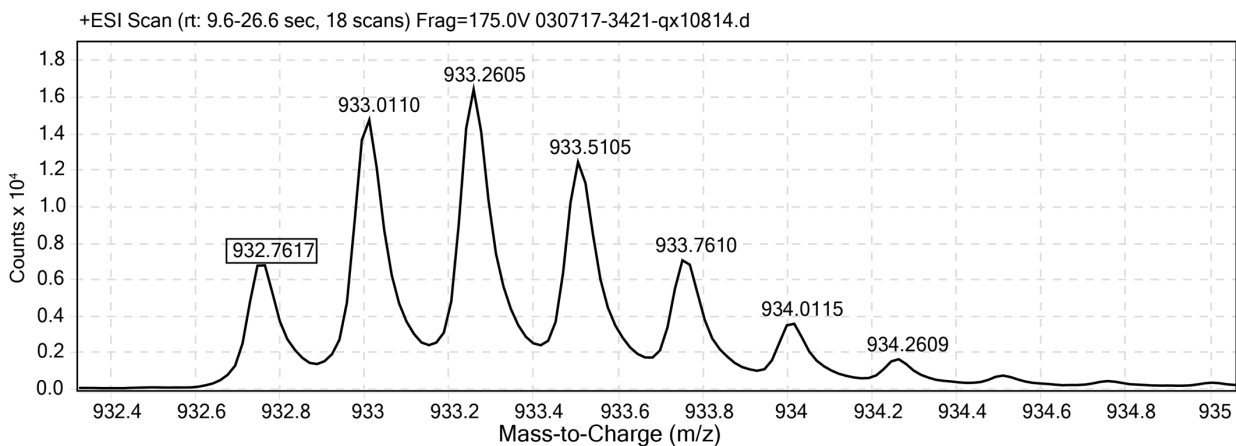


Figure S21. ESI-TOF MS data for **2a18** (QX10814). Expected $[M+4H^+]/4 = 932.769$.

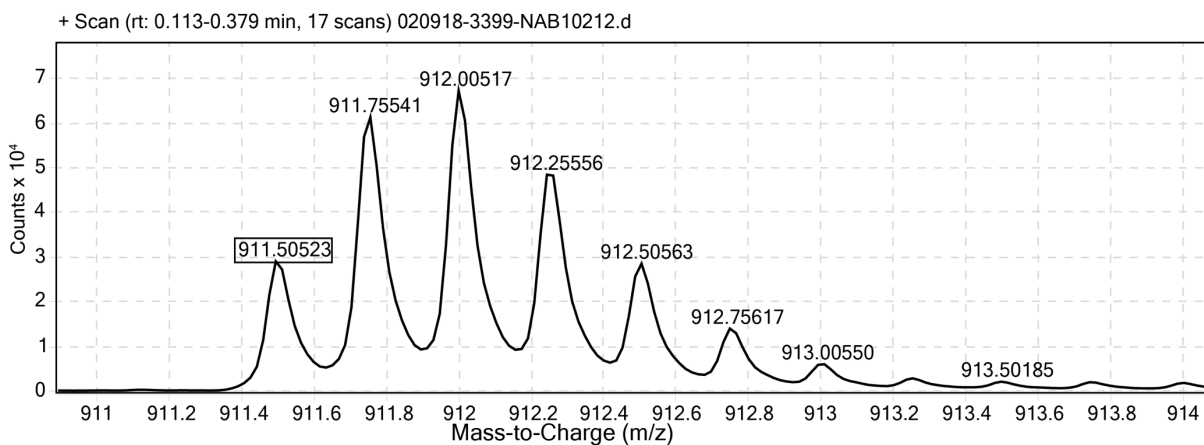


Figure S22. ESI-TOF MS data for **2a18-EA** (NAB10212). Expected $[M+4H^+]/4 = 911.503$.

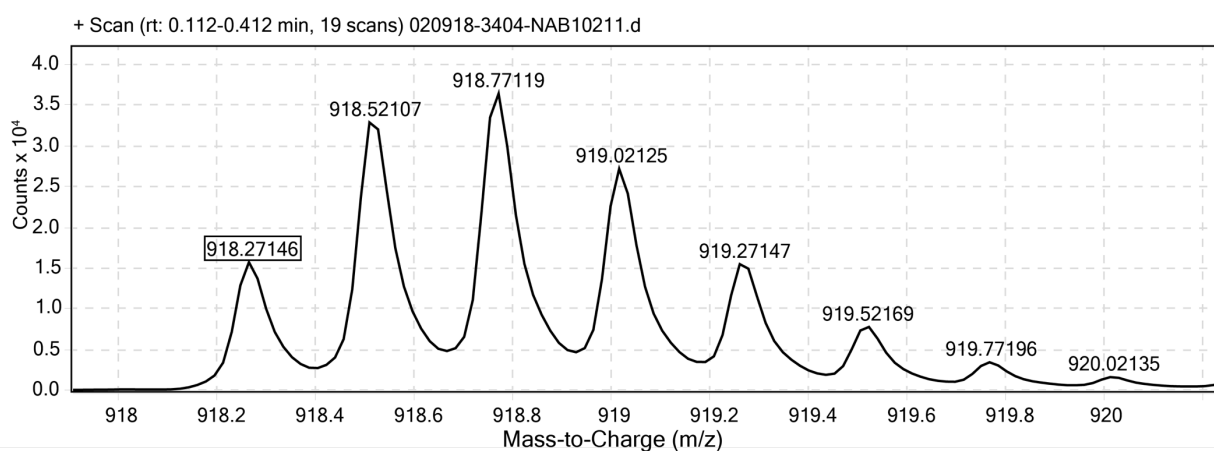


Figure S23. ESI-TOF MS data for **2a18-AR** (NAB10211). Expected $[M+4H^+]/4 = 918.268$.

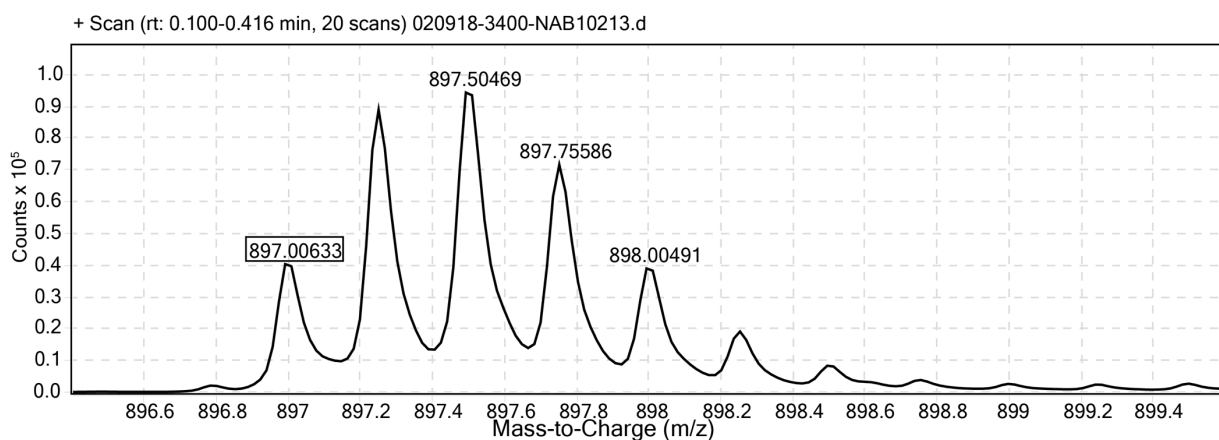


Figure S24. ESI-TOF MS data for **2a18-AA** (NAB10213). Expected $[M+4H^+]/4 = 897.002$.

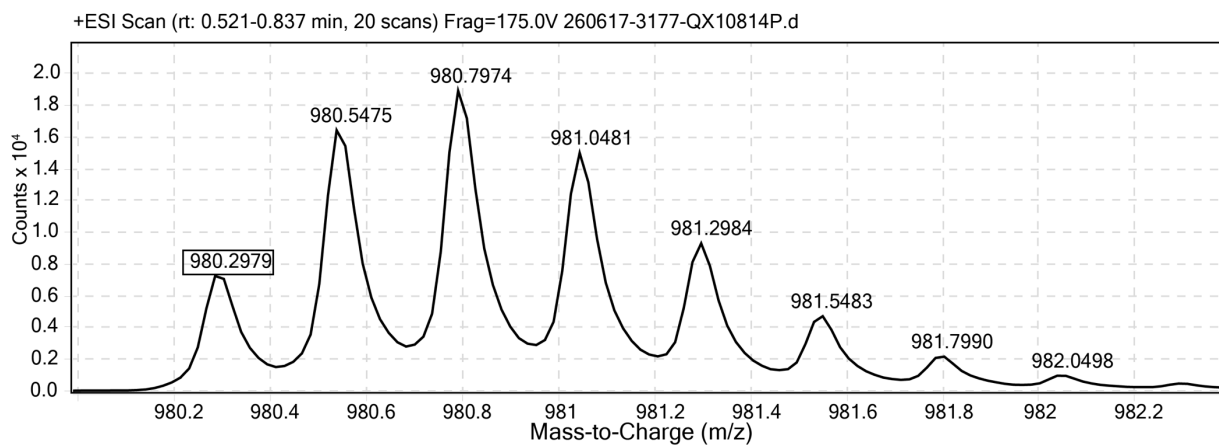


Figure S25. ESI-TOF MS data for **p2a18** (QX10814p). Expected $[M+4H^+]/4 = 980.300$.

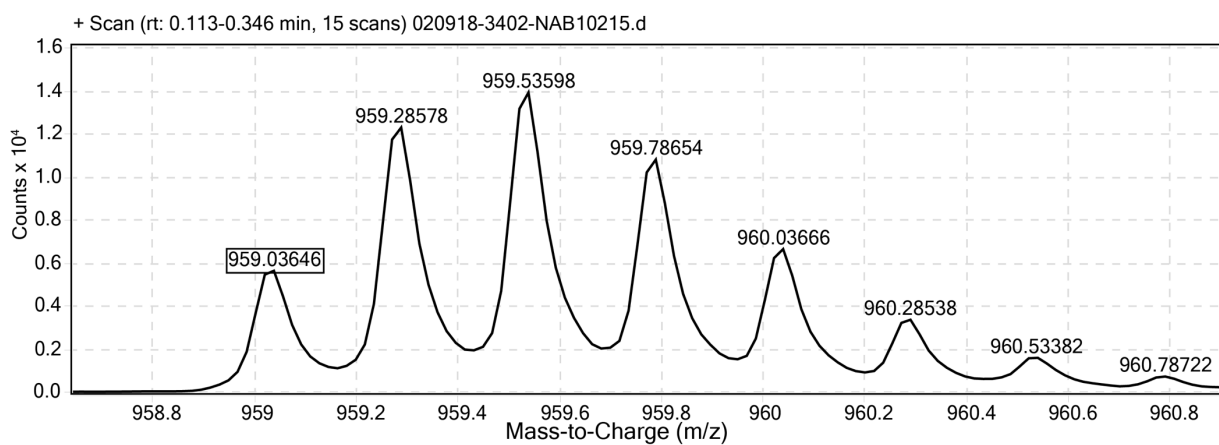


Figure S26. ESI-TOF MS data for **p2a18-EA** (NAB10215). Expected $[M+4H^+]/4 = 959.034$.

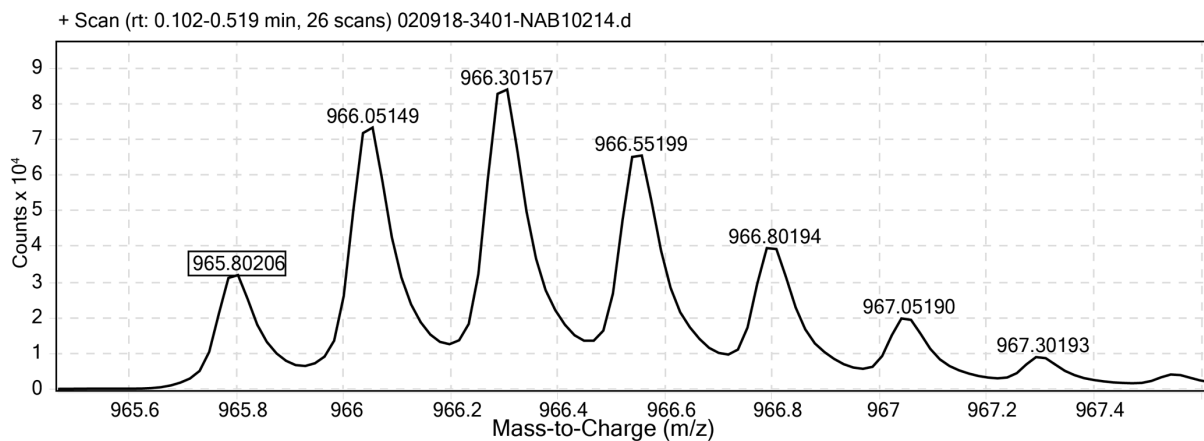


Figure S27. ESI-TOF MS data for **p2a18-AR** (NAB10214). Expected $[M+4H^+]/4 = 965.798$.

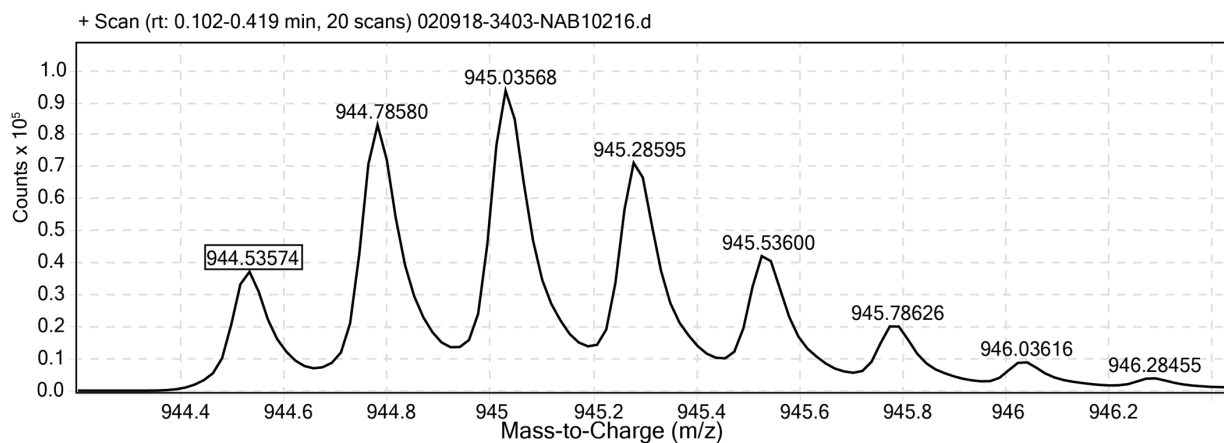


Figure S28. ESI-TOF MS data for **p2a18-AA** (NAB10216). Expected $[M+4H^+]/4 = 944.532$.

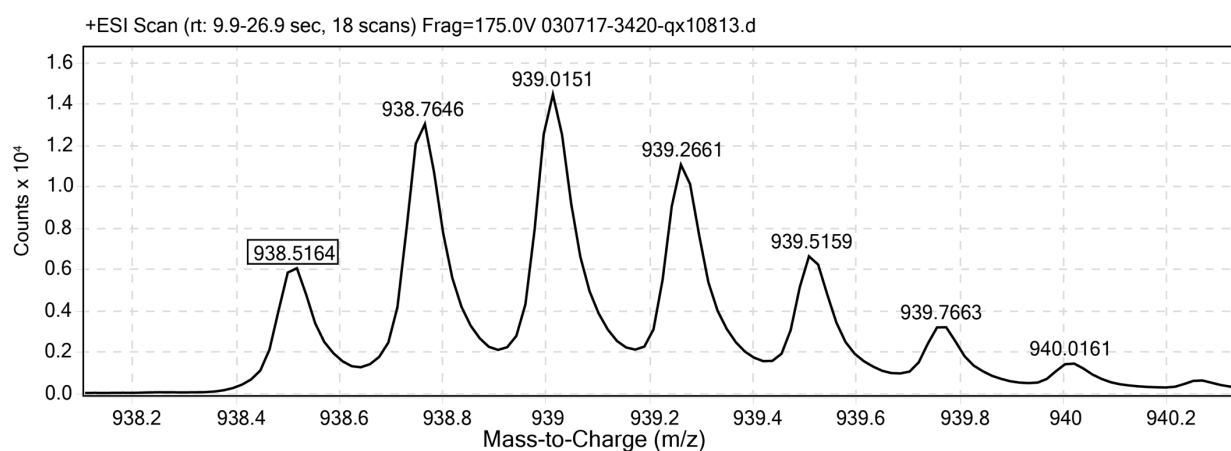


Figure S29. ESI-TOF MS data for **2a21** (QX10813). Expected $[M+4H^+]/4 = 938.523$.

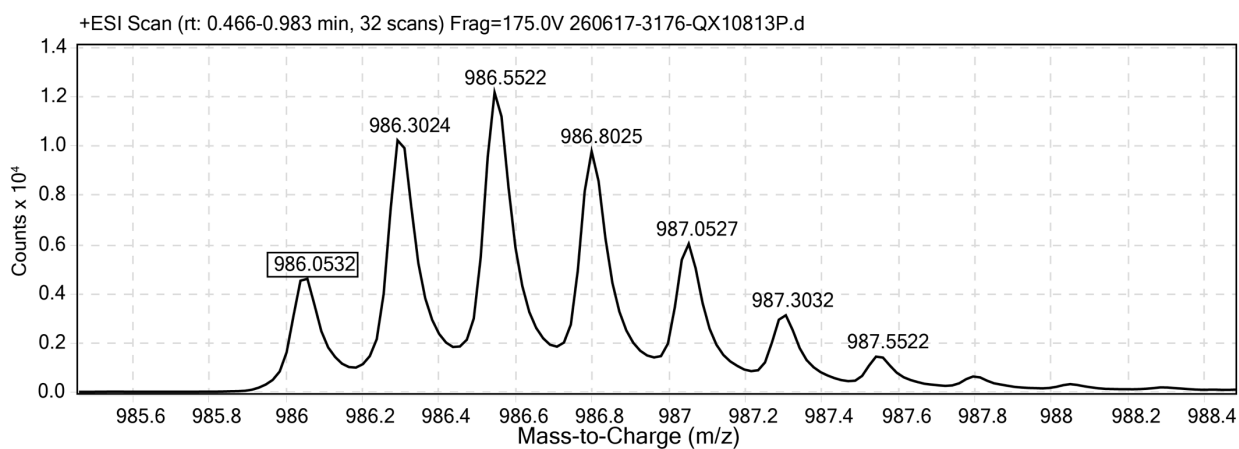


Figure S30. ESI-TOF MS data for **p2a21** (QX10813p). Expected $[M+4H^+]/4 = 986.054$.

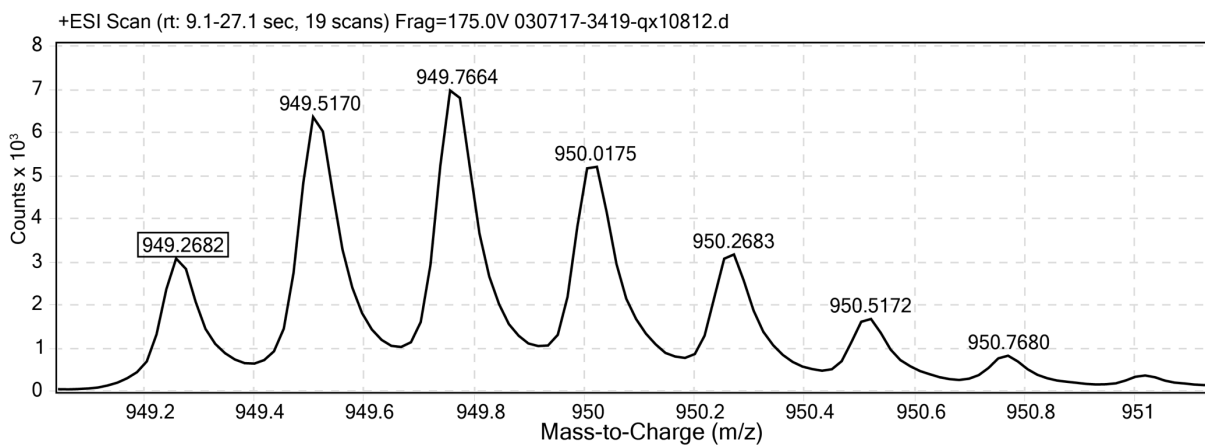


Figure S31. ESI-TOF MS data for **2a25** (QX10812). Expected $[M+4H^+]/4 = 949.275$.

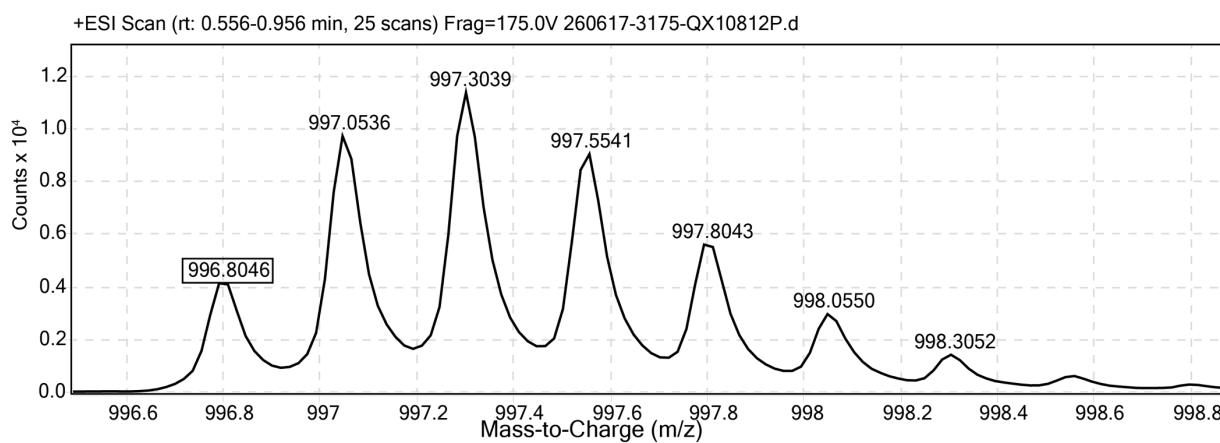


Figure S32. ESI-TOF MS data for **p2a25** (QX10812p). Expected $[M+4H^+]/4 = 996.805$.

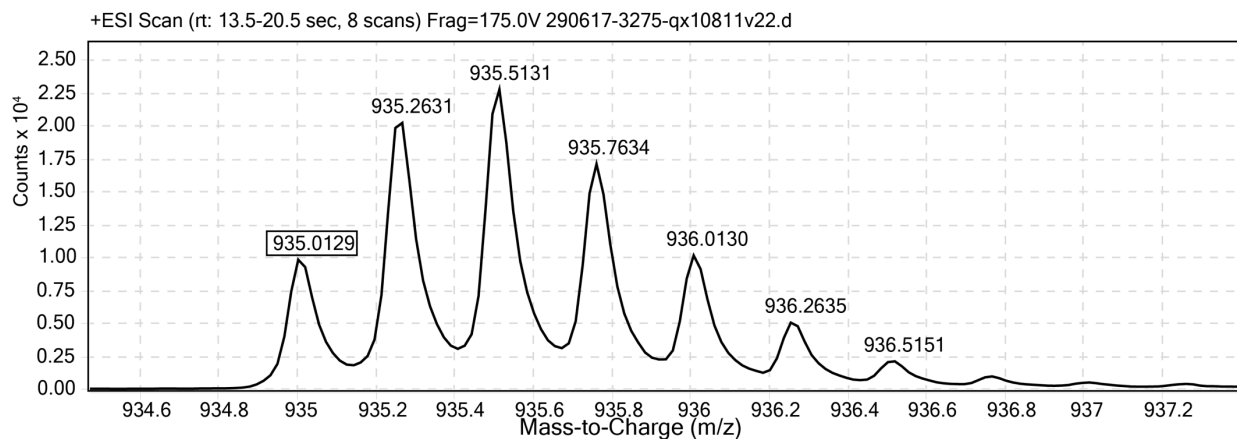


Figure S33. ESI-TOF MS data for **2a28** (QX10811). Expected $[M+4H^+]/4 = 935.010$.

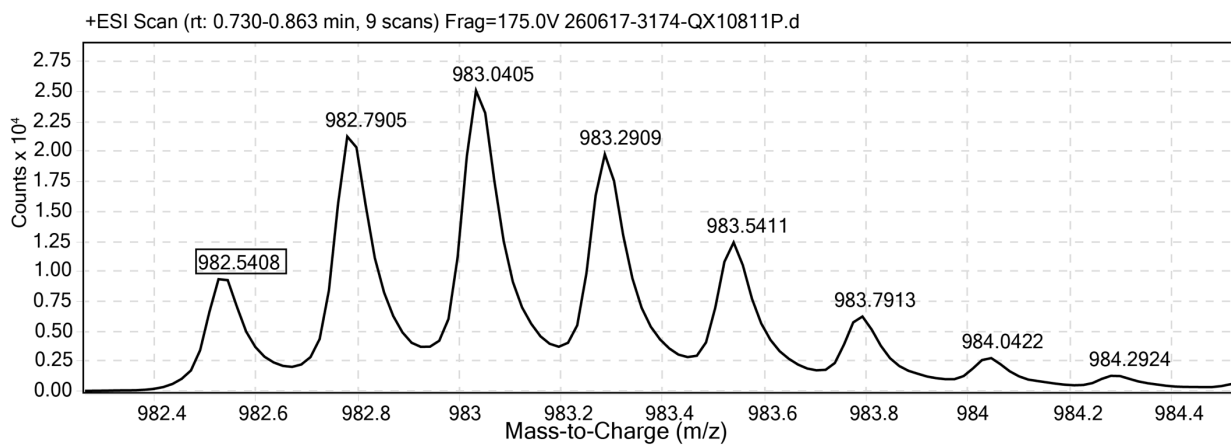


Figure S34. ESI-TOF MS data for **p2a28** (QX10811p). Expected $[M+4H^+]/4 = 982.541$.

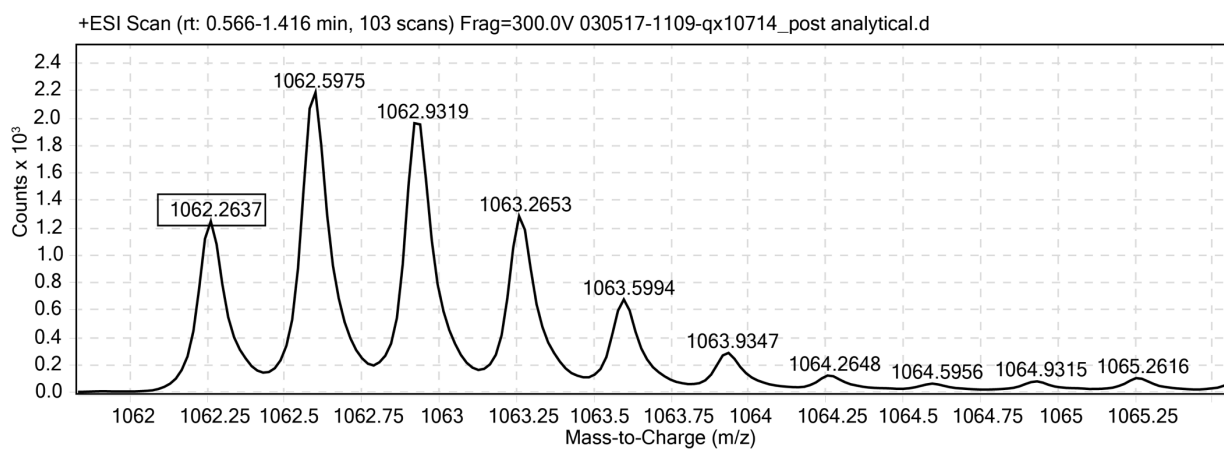


Figure S35. ESI-TOF MS data for **3a1** (QX10714). Expected $[M+3H^+]/3 = 1062.267$.

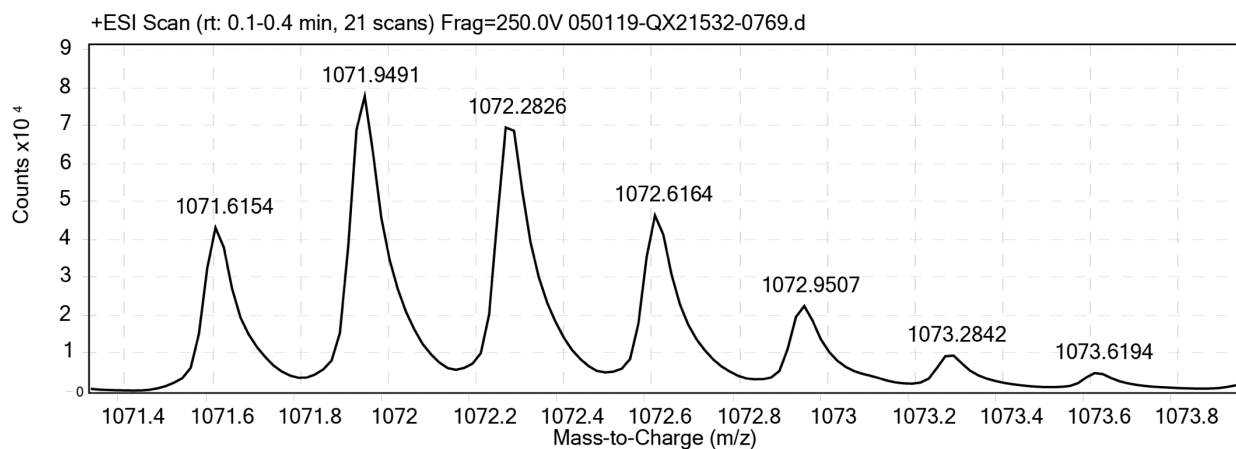


Figure S36. ESI-TOF MS data for **3a1-ER** (QX21532). Expected $[M+3H^+]/3 = 1071.602$.

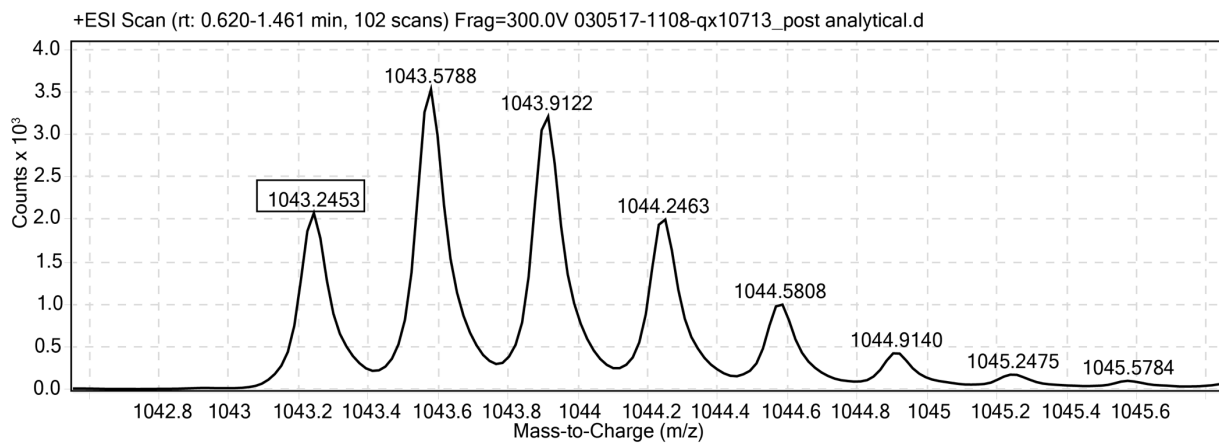


Figure S37. ESI-TOF MS data for **3a1-EA** (QX10713). Expected $[M+3H^+]/3 = 1043.247$.

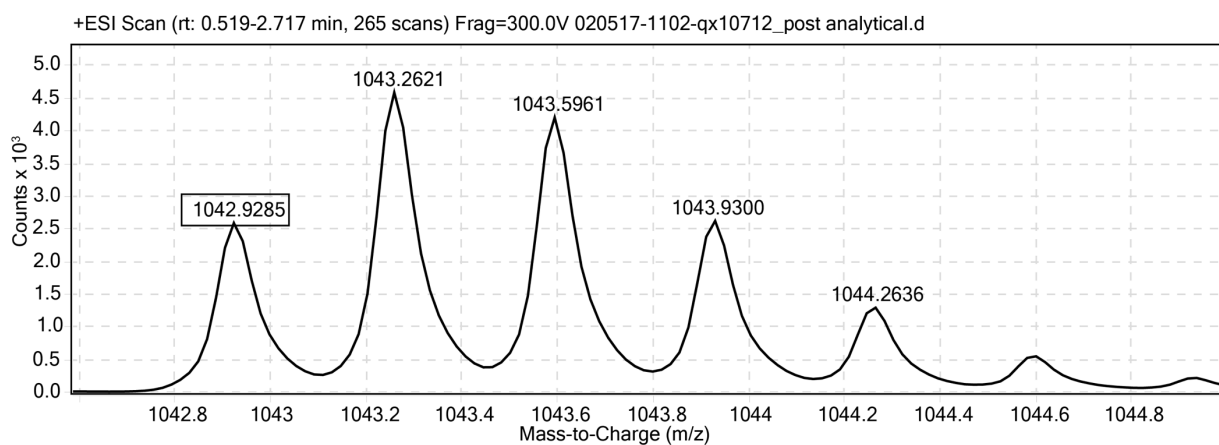


Figure S38. ESI-TOF MS data for **3a1-AK** (QX10712). Expected $[M+3H^+]/3 = 1042.931$.

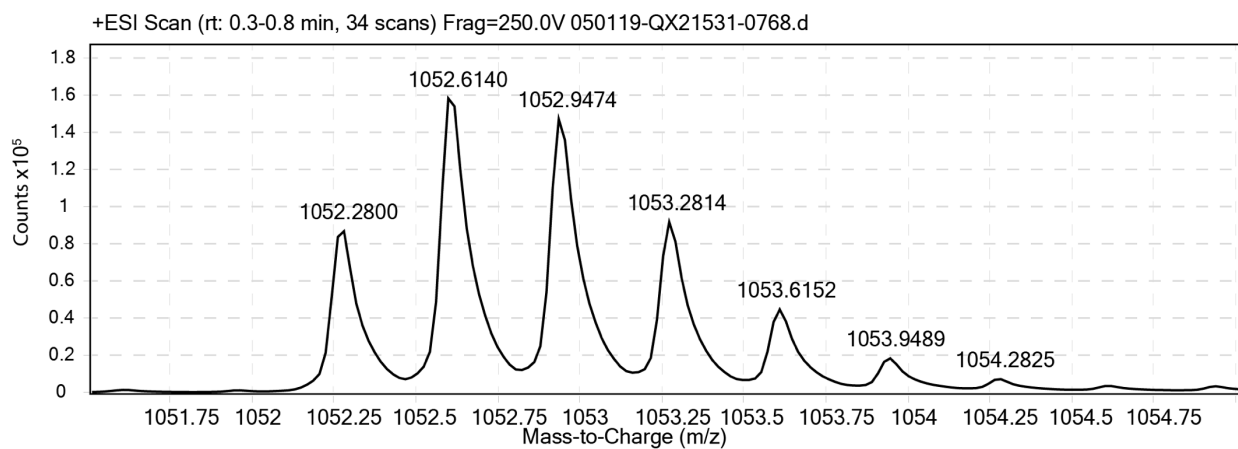


Figure S39. ESI-TOF MS data for **3a1-AR** (QX21531). Expected $[M+3H^+]/3 = 1052.267$.

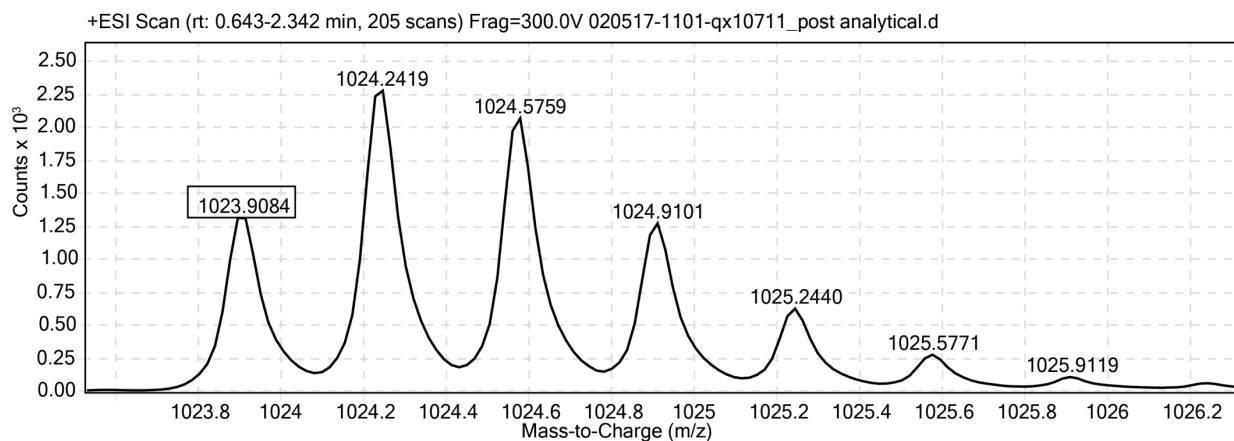


Figure S40. ESI-TOF MS data for **3a1-AA** (QX10711). Expected $[M+3H^+]/3 = 1023.912$.

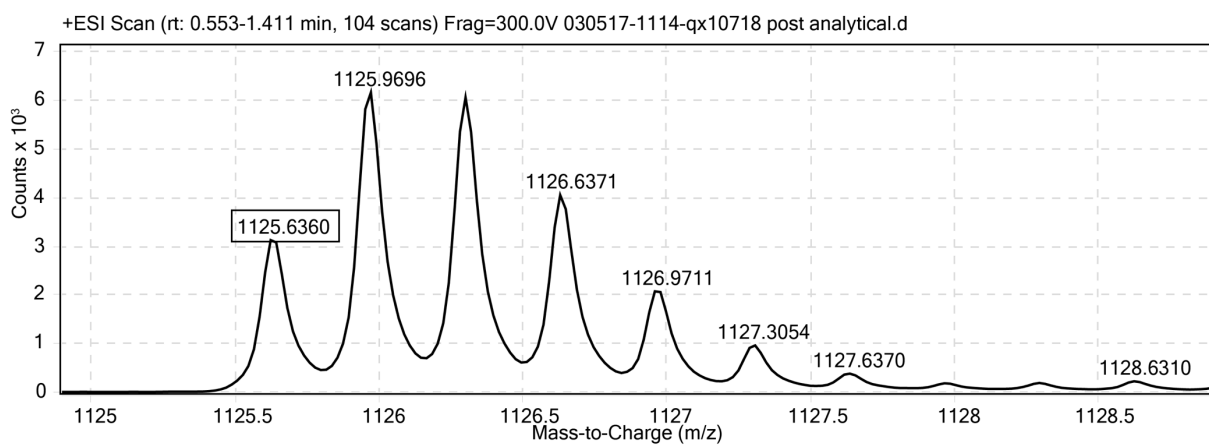


Figure S41. ESI-TOF MS data for **p3a1** (QX10718). Expected $[M+3H^+]/3 = 1125.640$.

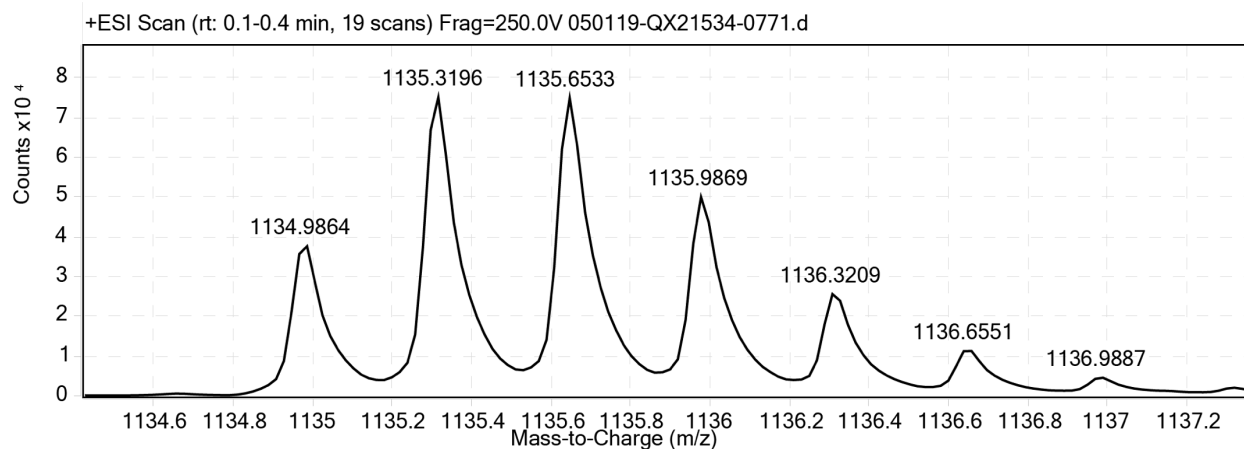


Figure S42. ESI-TOF MS data for **p3a1-ER** (QX21534). Expected $[M+3H^+]/3 = 1134.976$.

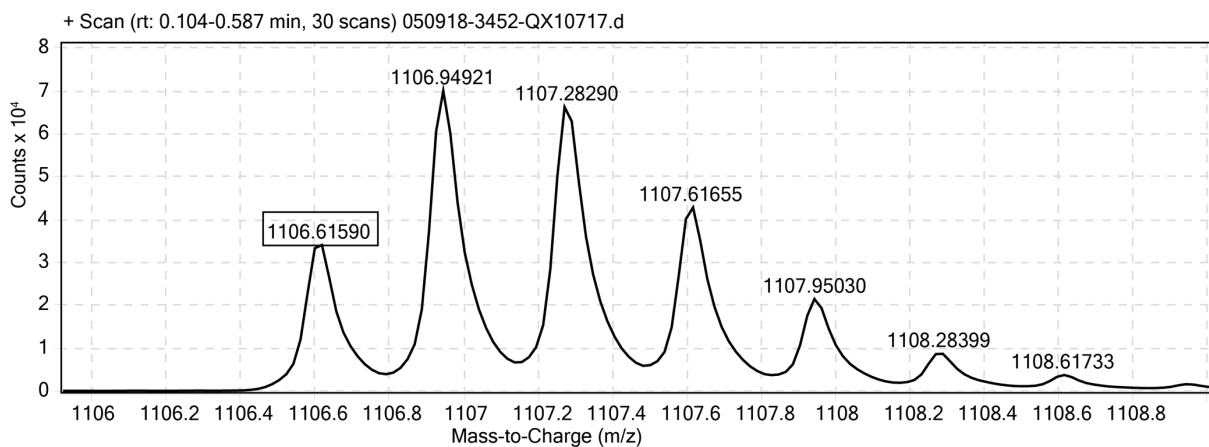


Figure S43. ESI-TOF MS data for **p3a1-EA** (QX10717). Expected $[M+3H^+]/3 = 1106.621$.

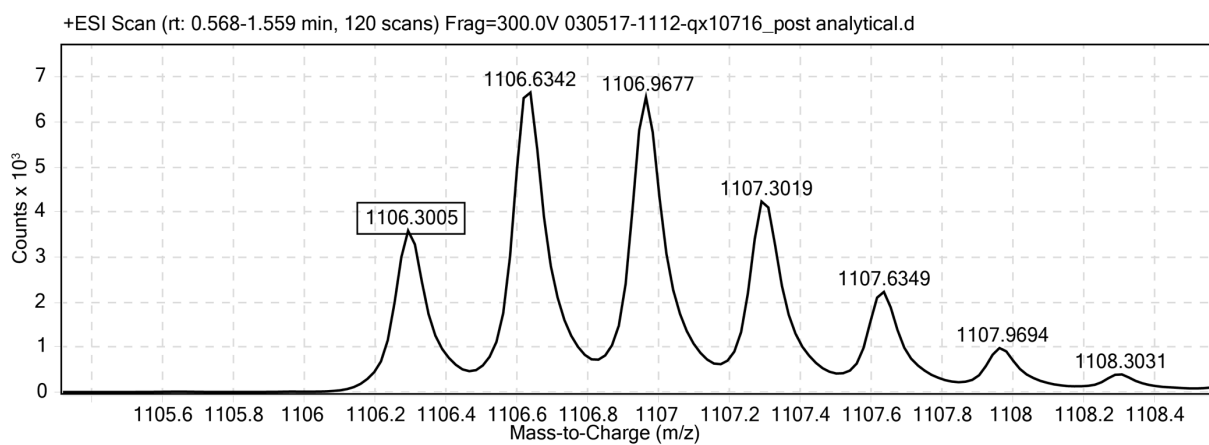


Figure S44. ESI-TOF MS data for **p3a1-AK** (QX10716). Expected $[M+3H^+]/3 = 1106.305$.

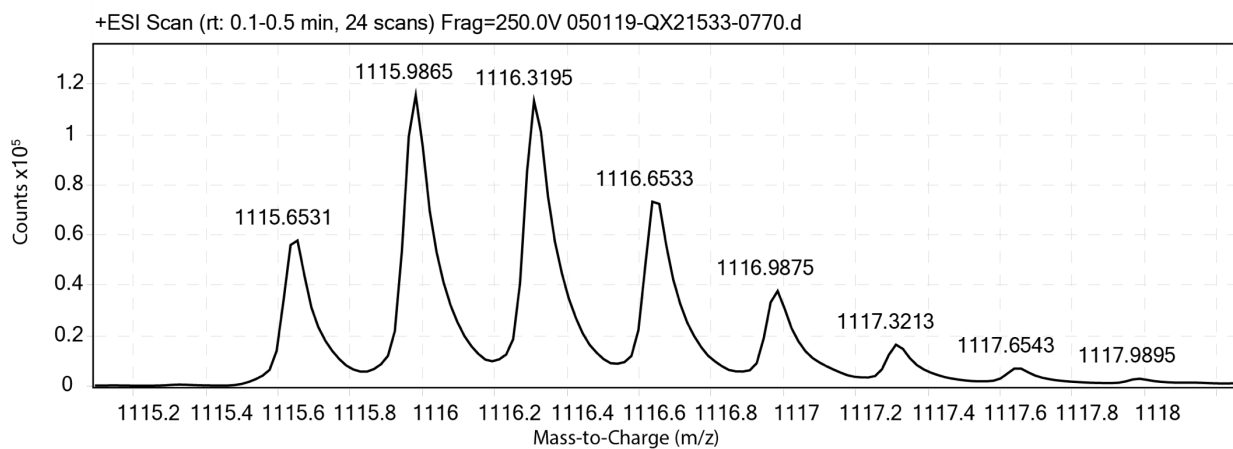


Figure S45. ESI-TOF MS data for **p3a1-AR** (QX21533). Expected $[M+3H^+]/3 = 1115.640$.

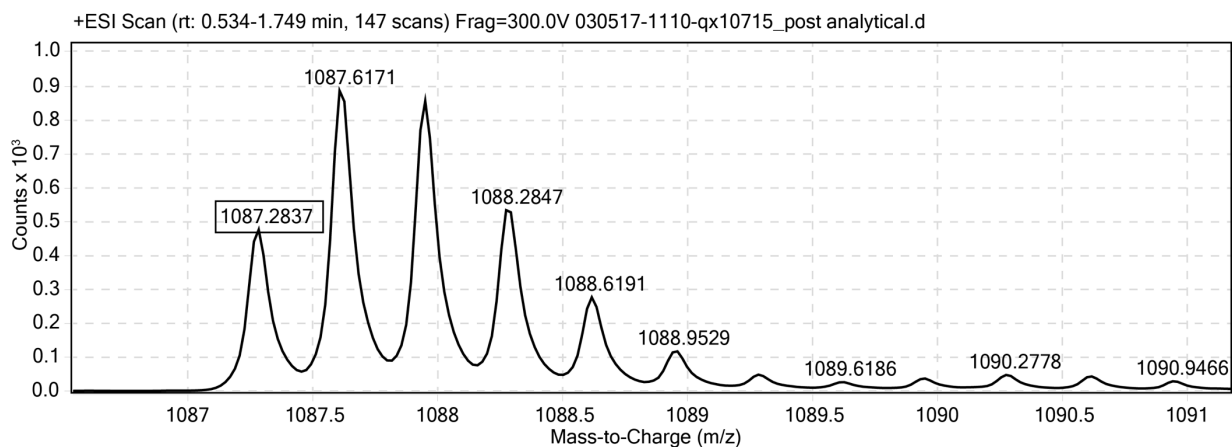


Figure S46. ESI-TOF MS data for **p3a1-AA** (QX10715). Expected $[M+3H^+]/3 = 1087.286$.

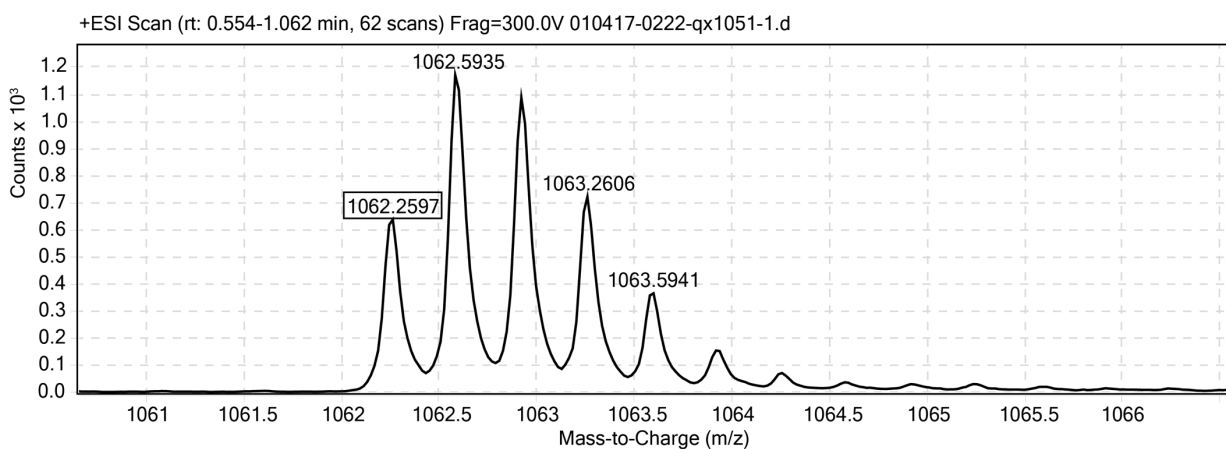


Figure S47. ESI-TOF MS data for **3a6** (QX10511). Expected $[M+3H^+]/3 = 1062.267$.

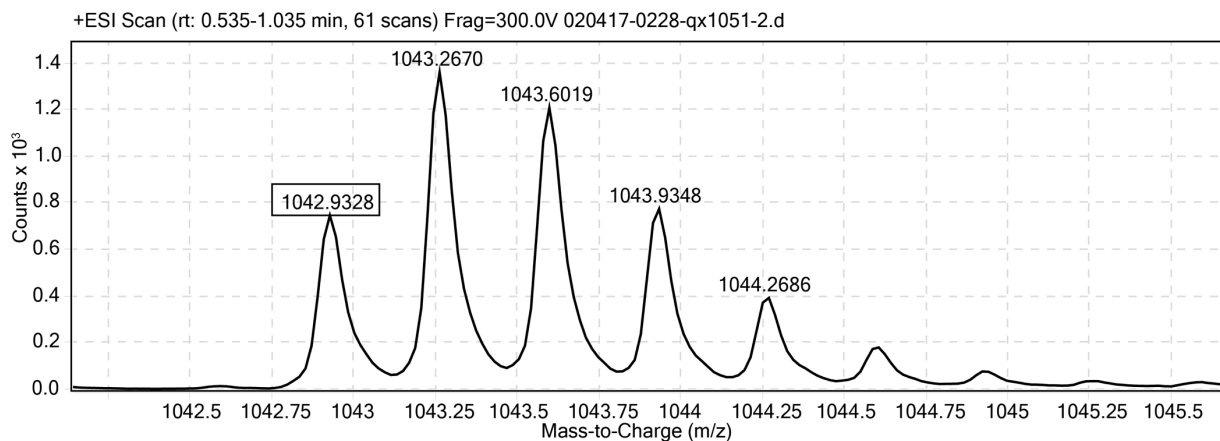


Figure S48. ESI-TOF MS data for **3a6-KA** (QX10512). Expected $[M+3H^+]/3 = 1042.931$.

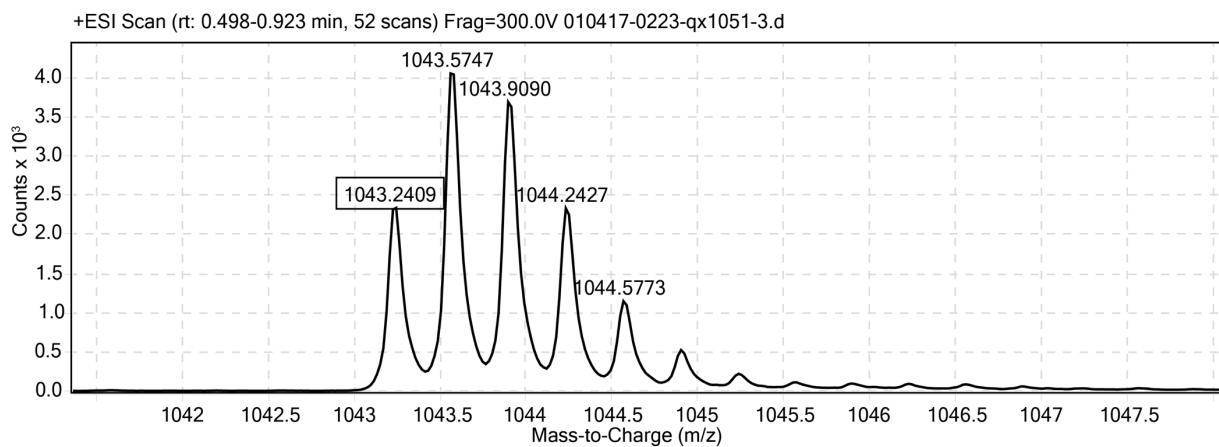


Figure S49. ESI-TOF MS data for **3a6-AE** (QX10513). Expected $[M+3H^+]/3 = 1043.247$.

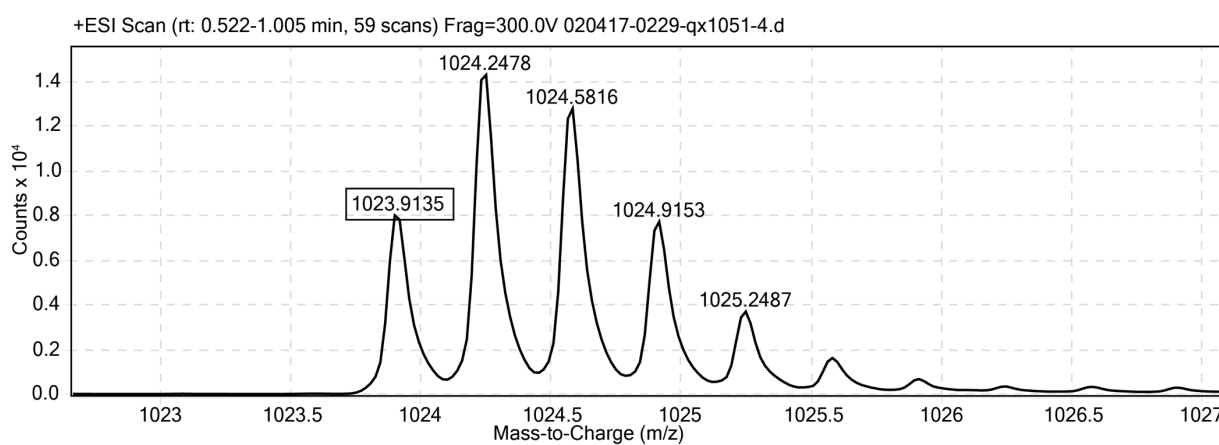


Figure S50. ESI-TOF MS data for **3a6-AA** (QX10514). Expected $[M+3H^+]/3 = 1023.912$.

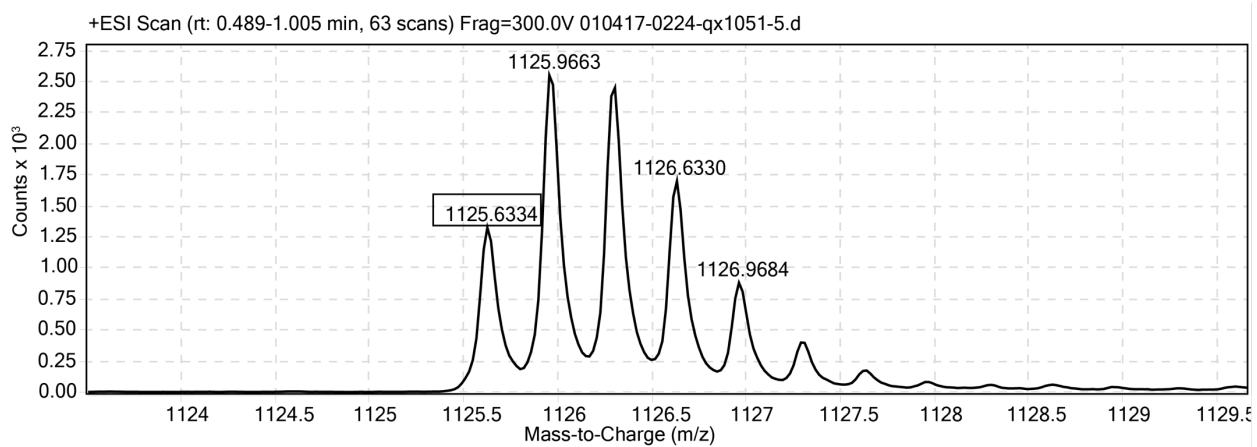


Figure S51. ESI-TOF MS data for **p3a6** (QX10515). Expected $[M+3H^+]/3 = 1125.640$.

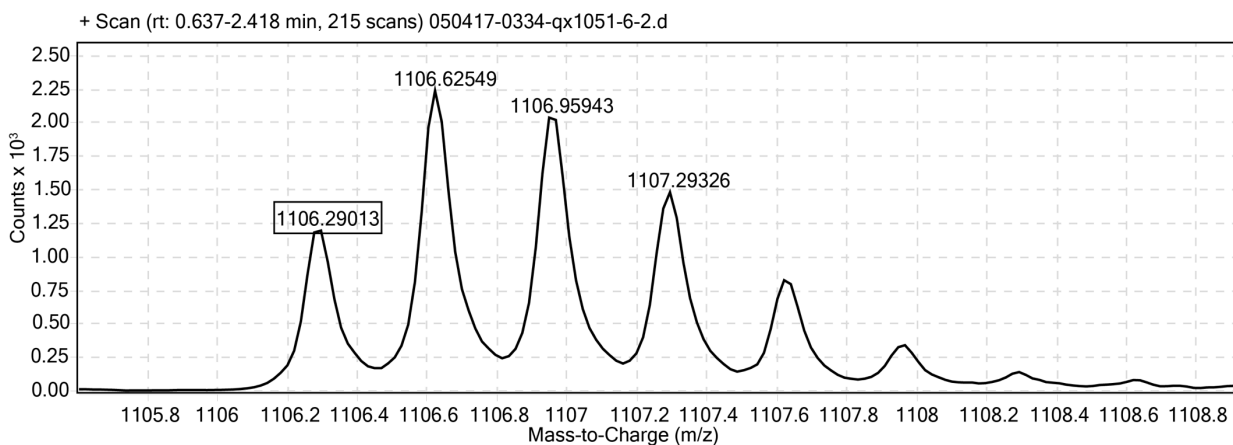


Figure S52. ESI-TOF MS data for **p3a6-KA** (QX10516). Expected $[M+3H^+]/3 = 1106.305$.

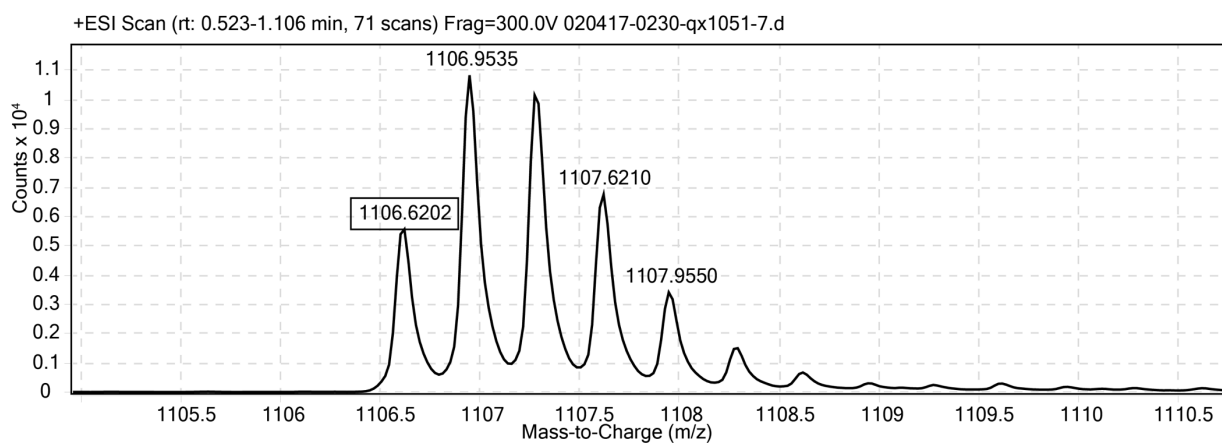


Figure S53. ESI-TOF MS data for **p3a6-AE** (QX10517). Expected $[M+3H^+]/3 = 1106.621$.

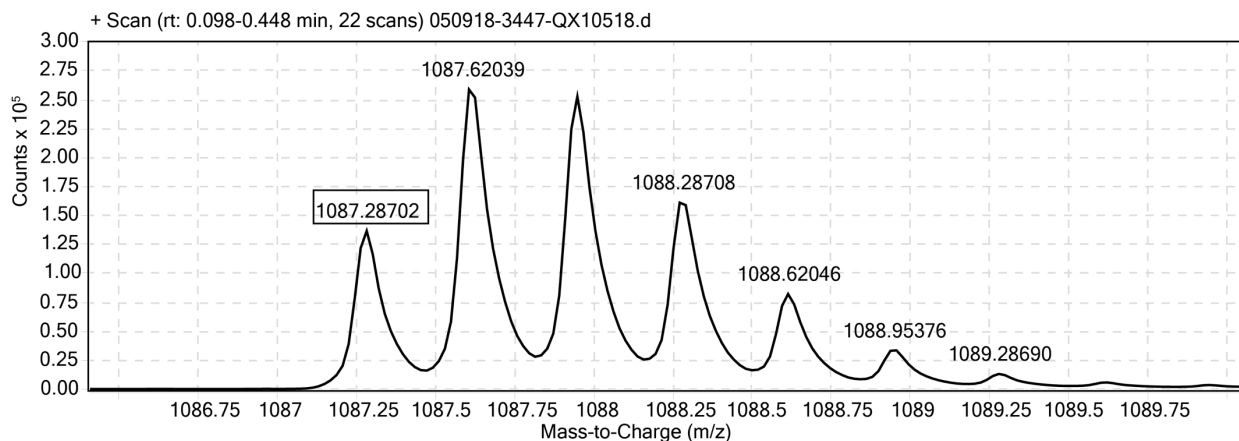


Figure S54. ESI-TOF MS data for **p3a6-AA** (QX10518). Expected $[M+3H^+]/3 = 1087.286$.

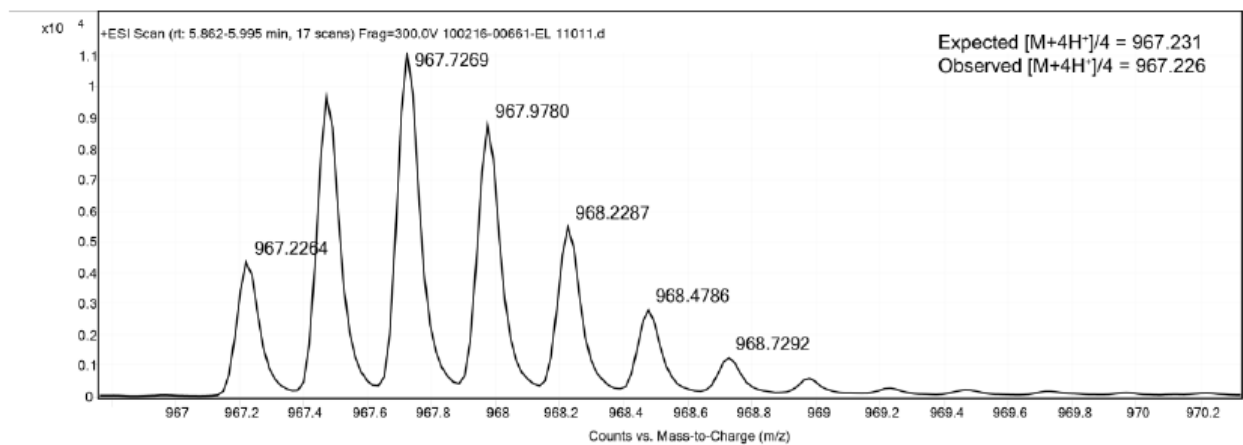


Figure S55. ESI-TOF MS data for $\beta 18$ (EL1101). Expected $[M+4H^+]/4 = 967.231$.

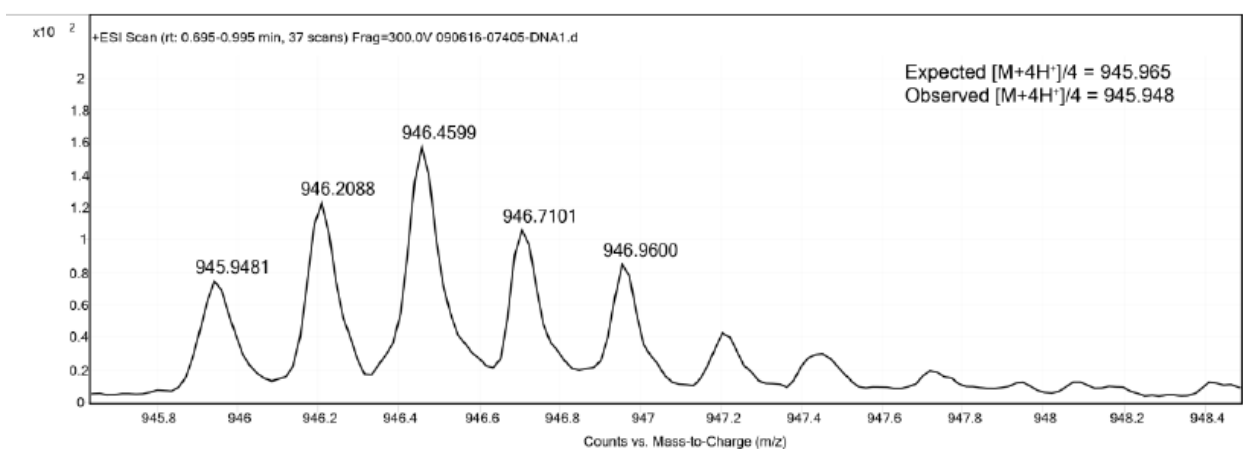


Figure S56. ESI-TOF MS data for $\beta 18$ -DA. Expected $[M+4H^+]/4 = 945.965$.

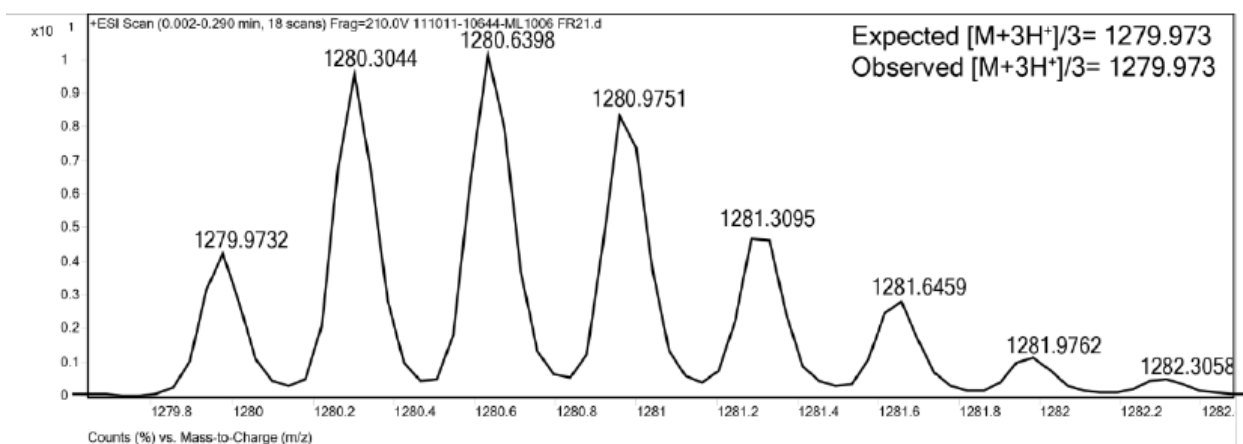


Figure S57. ESI-TOF MS data for $\beta 18$ -SR (ML1006). Expected $[M+4H^+]/3 = 1279.973$.

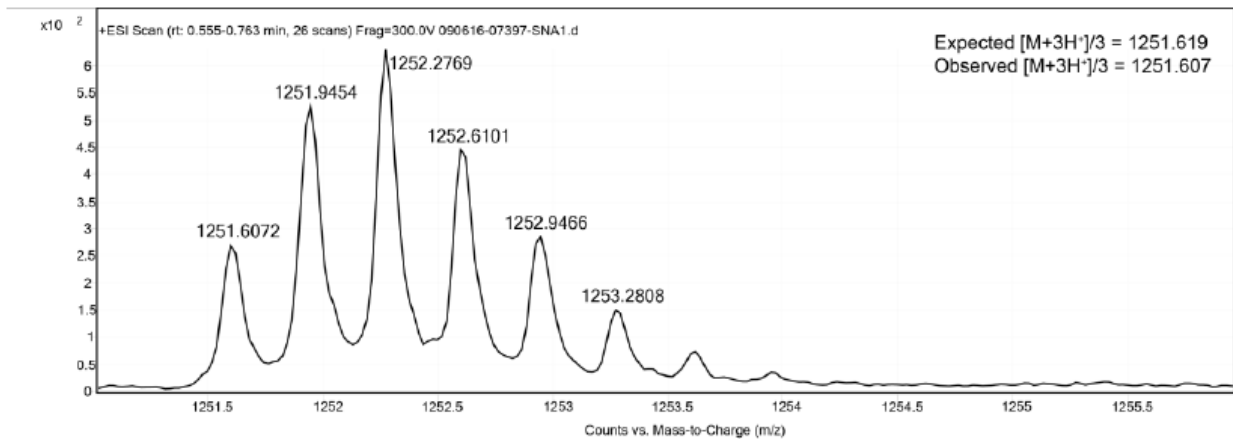


Figure S58. ESI-TOF MS data for β 18-SA. Expected $[M+4H^+]/3 = 1251.619$.

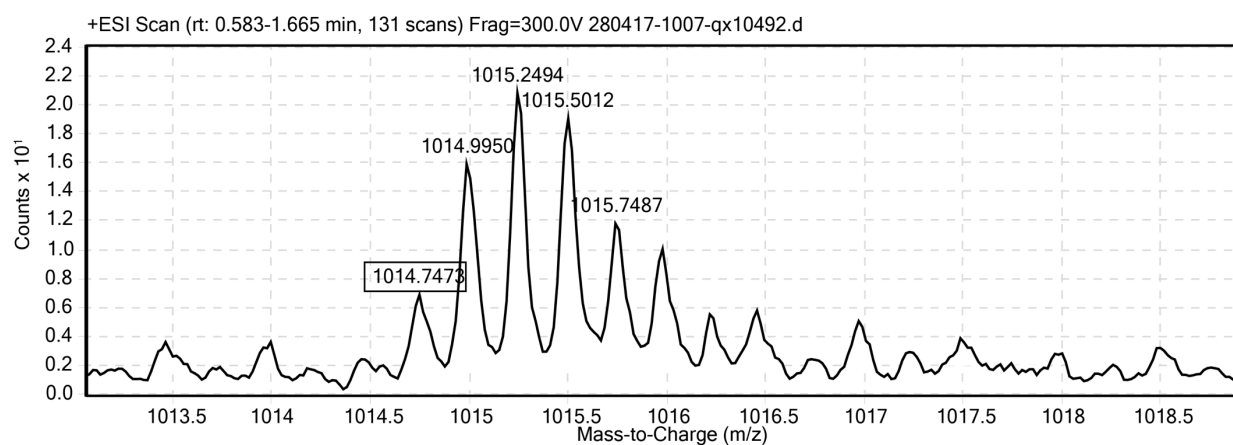


Figure S59. ESI-TOF MS data for $p\beta$ 18 (QX10492). Expected $[M+4H^+]/4 = 1014.747$.

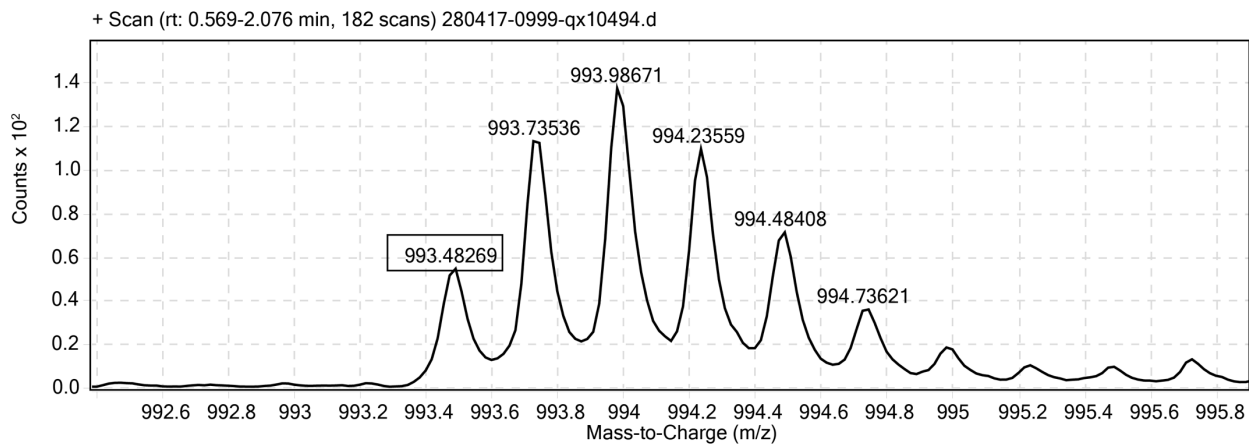


Figure S60. ESI-TOF MS data for $p\beta$ 18-DA (QX10494). Expected $[M+4H^+]/4 = 993.495$.

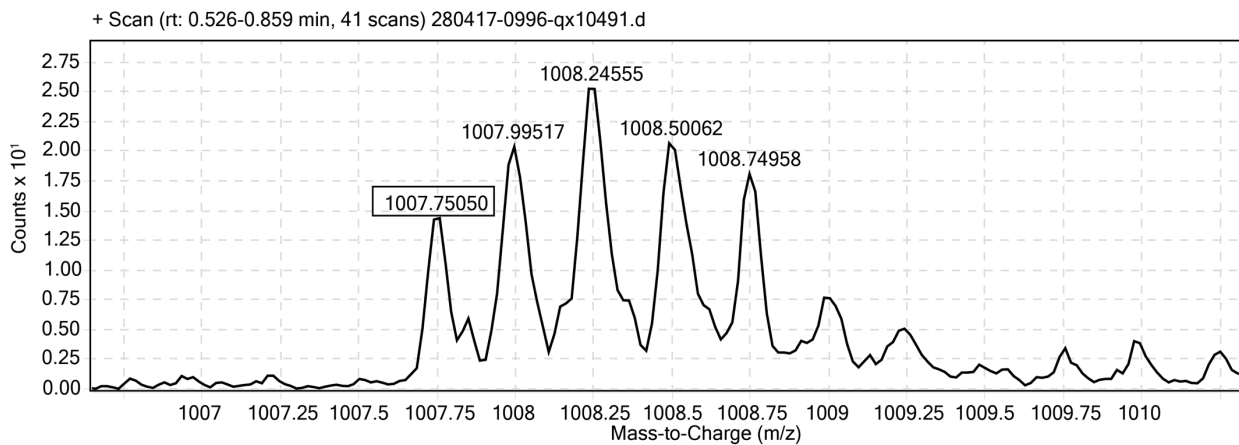


Figure S61. ESI-TOF MS data for **pβ18-SR** (QX10491). Expected $[M+4H^+]/4 = 1007.762$.

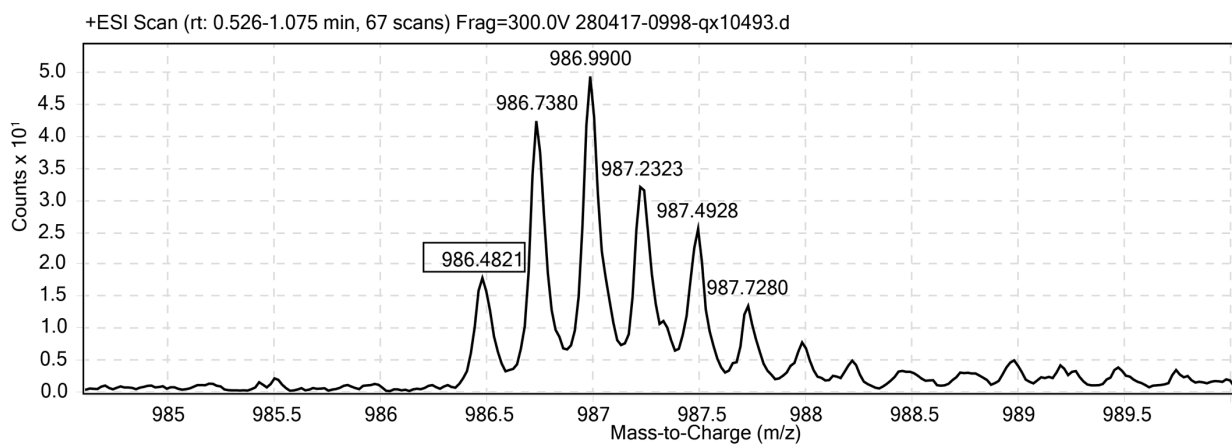


Figure S62. ESI-TOF MS data for **pβ18-SA** (QX10493). Expected $[M+4H^+]/4 = 986.496$.

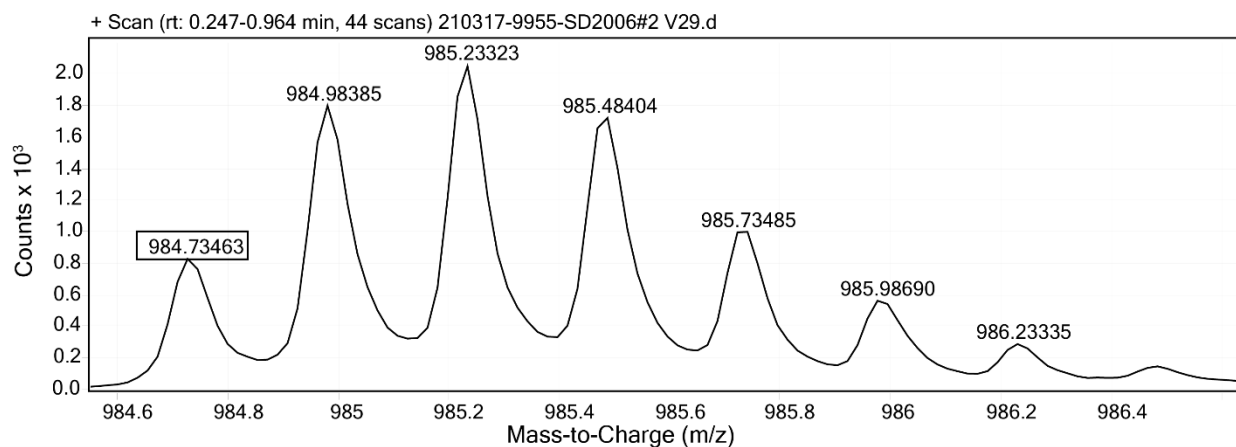


Figure S63. ESI-TOF MS data for **β23-EA** (SD2006#2). Expected $[M+4H^+]/4 = 984.741$.

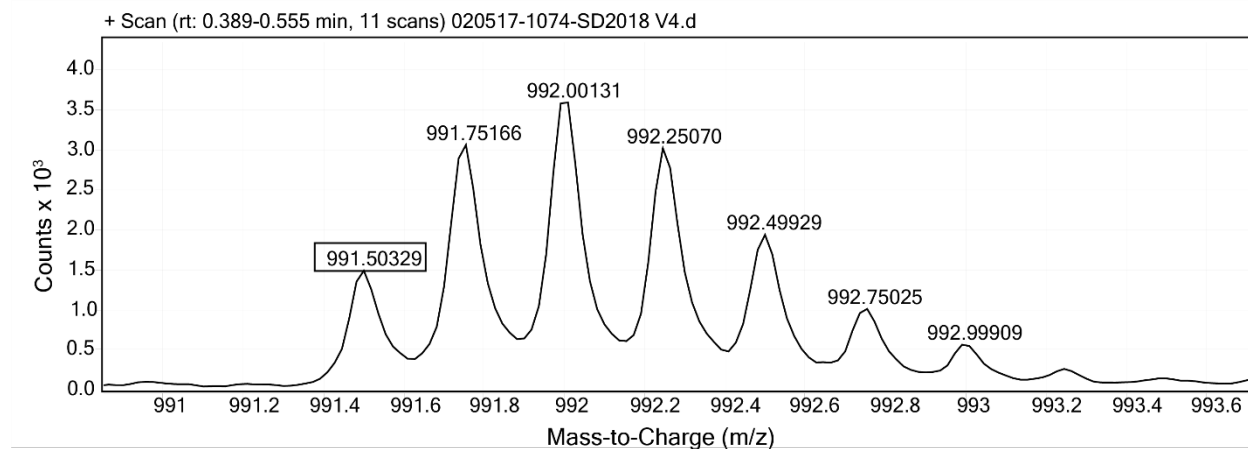


Figure S64. ESI-TOF MS data for **β 23-AR** (SD2018). Expected $[M+4H^+]/4 = 991.506$.

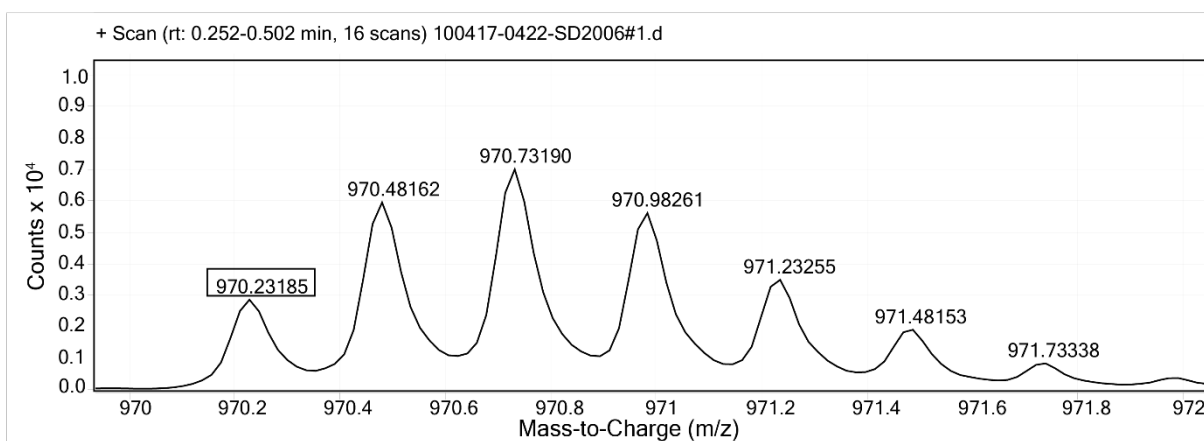


Figure S65. ESI-TOF MS data for **β 23-AA** (SD2006#1). Expected $[M+4H^+]/4 = 970.240$.

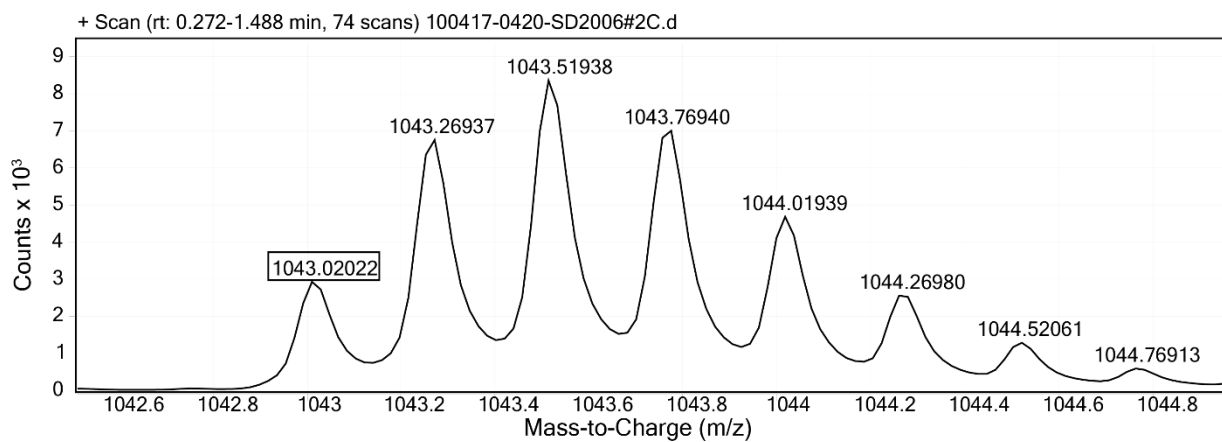


Figure S66. ESI-TOF MS data for **$p\beta$ 23-EA** (SD2006#2C). Expected $[M+4H^+]/4 = 1043.026$.

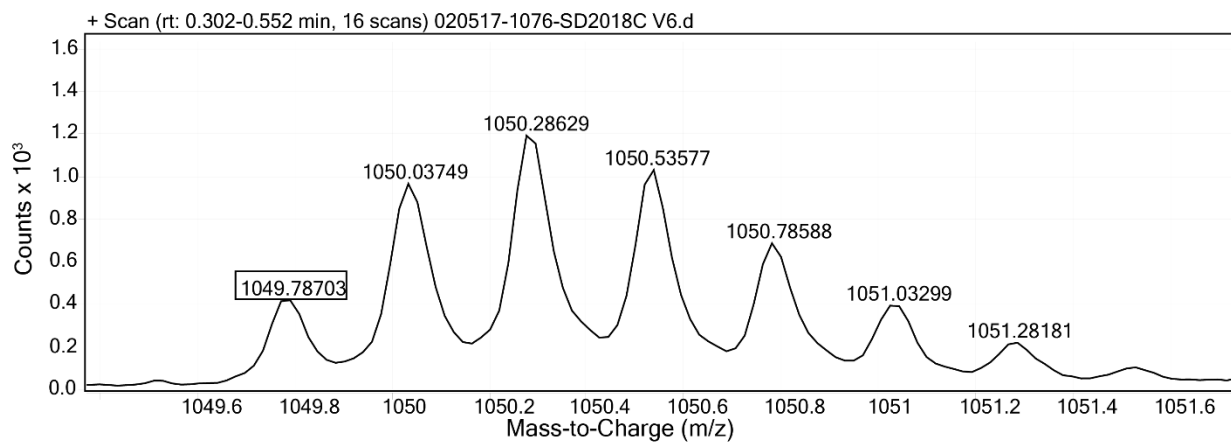


Figure S67. ESI-TOF MS data for **pβ23-AR** (SD2018C). Expected $[M+4H^+]/4 = 1049.790$.

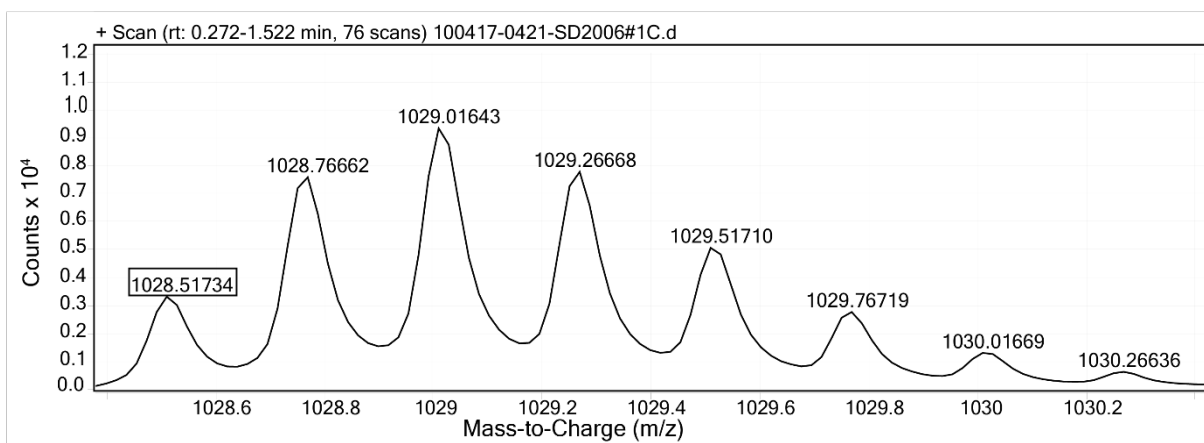


Figure S68. ESI-TOF MS data for **pβ23-AA** (SD2006#1C). Expected $[M+4H^+]/4 = 1028.524$.

Analytical HPLC data.

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.

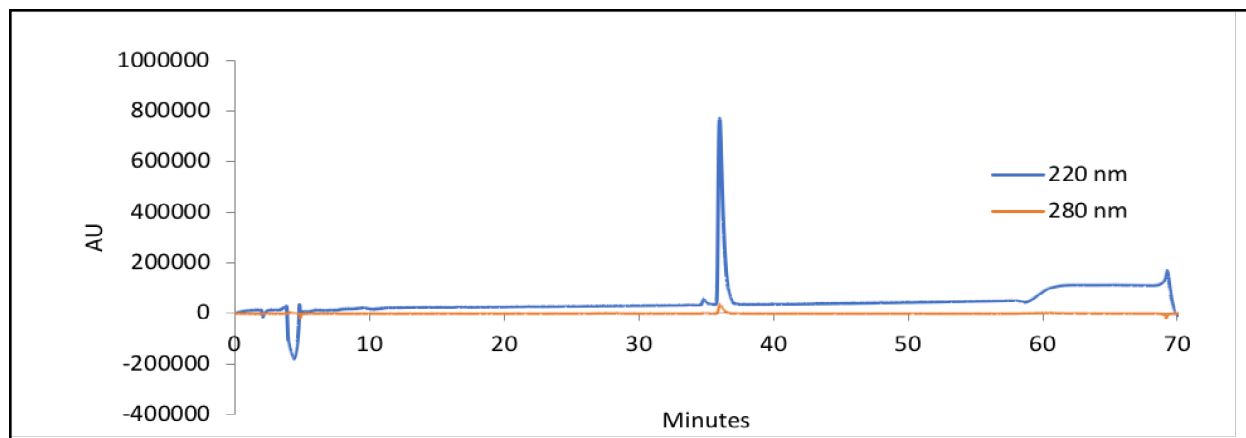


Figure S69. Analytical HPLC data for **2α1** (QX108111). Retention time = 36.042 minutes.

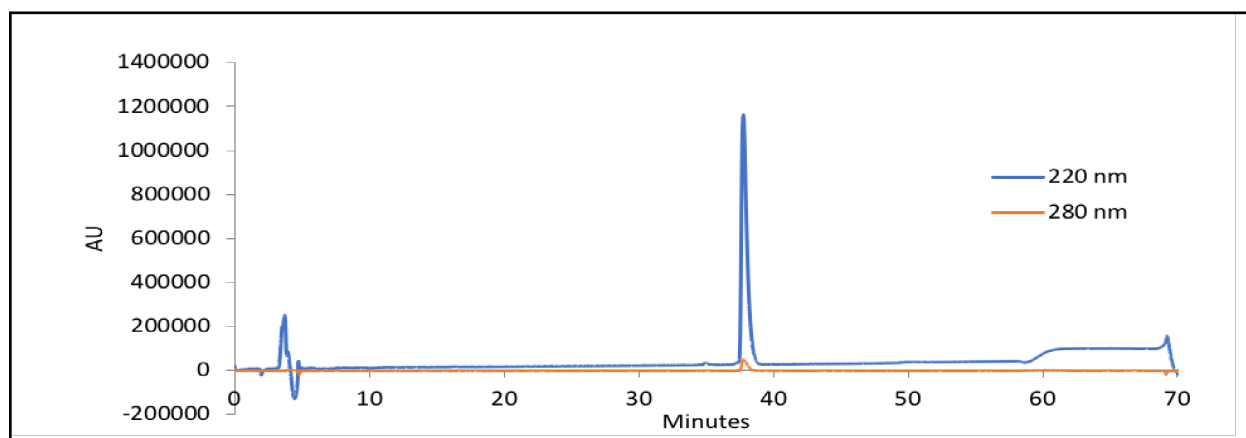


Figure S70. Analytical HPLC data for **p2α1** (QX108111p). Retention time = 37.750 minutes.

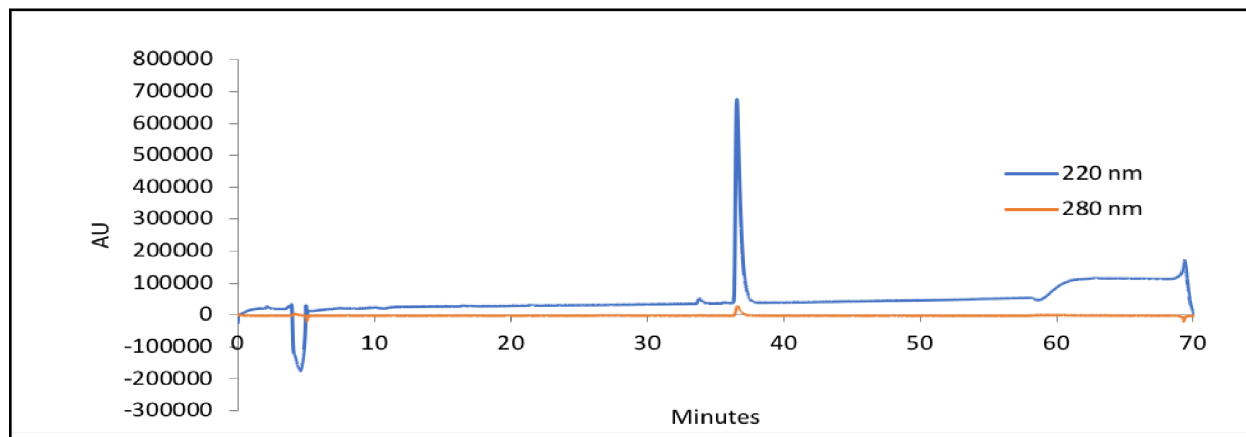


Figure S71. Analytical HPLC data for **2α3** (QX108110). Retention time = 36.625 minutes.

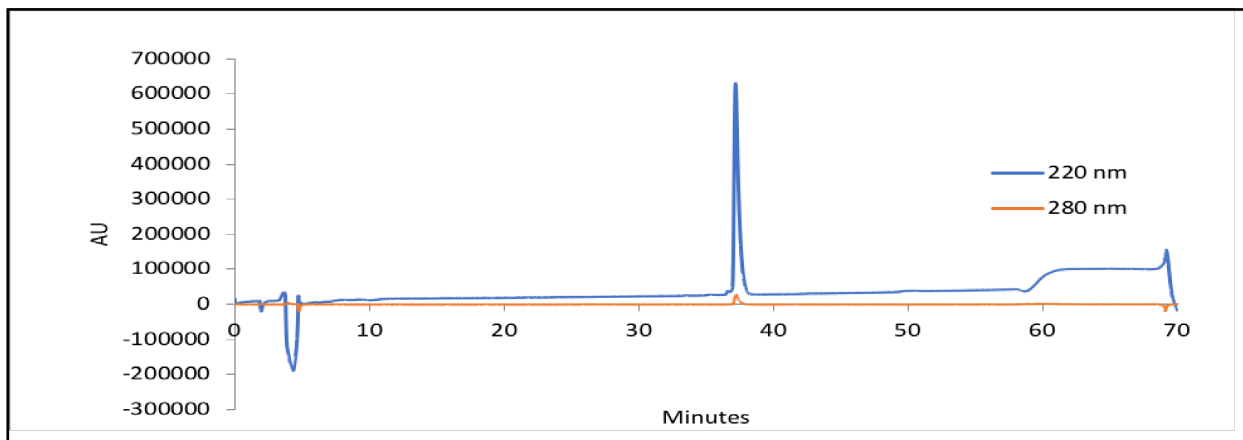


Figure S72. Analytical HPLC data for **p2 α 3** (QX108110p). Retention time = 37.233 minutes.

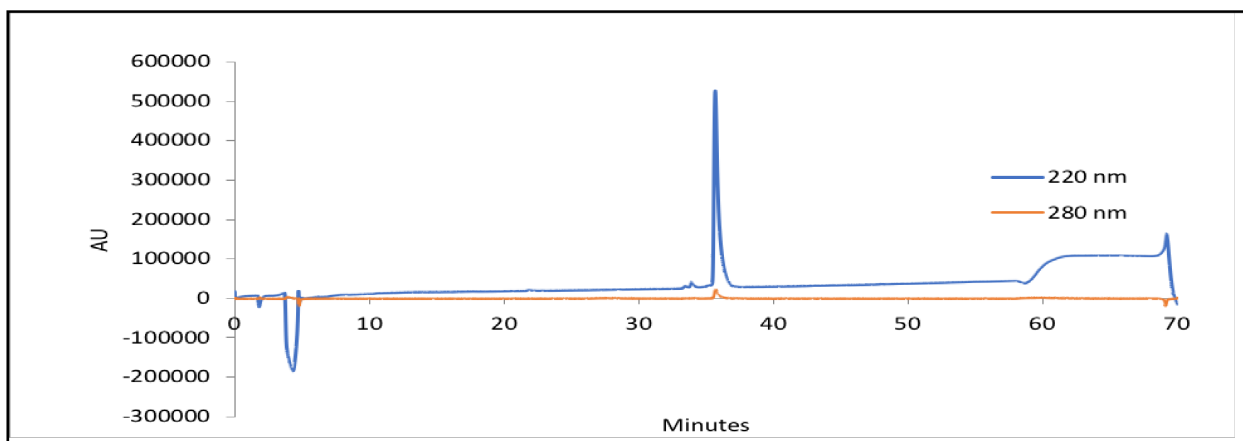


Figure S73. Analytical HPLC data for **2 α 4** (QX10819). Retention time = 35.683 minutes.

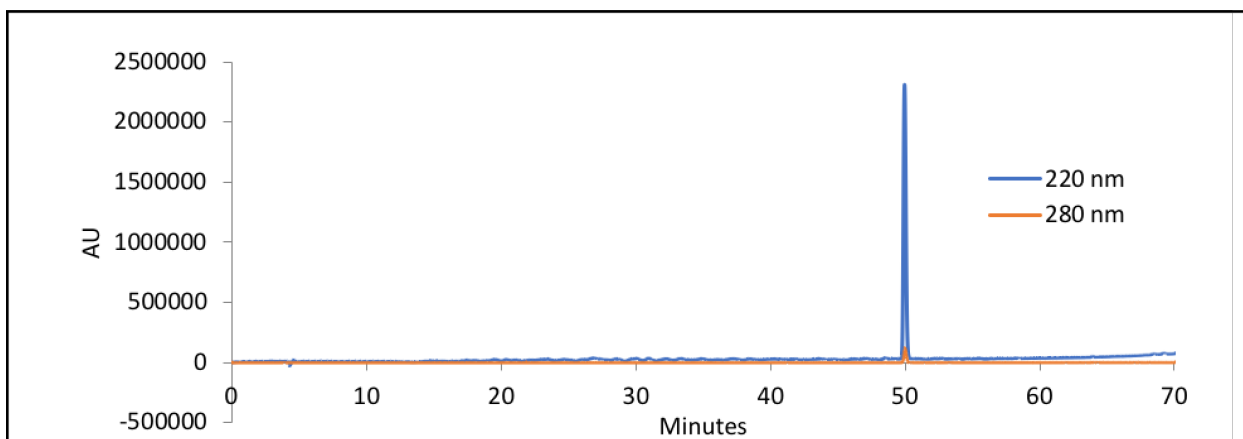


Figure S74. Analytical HPLC data for **2 α 4-KA** (QX11192). Retention time = 49.983 minutes.

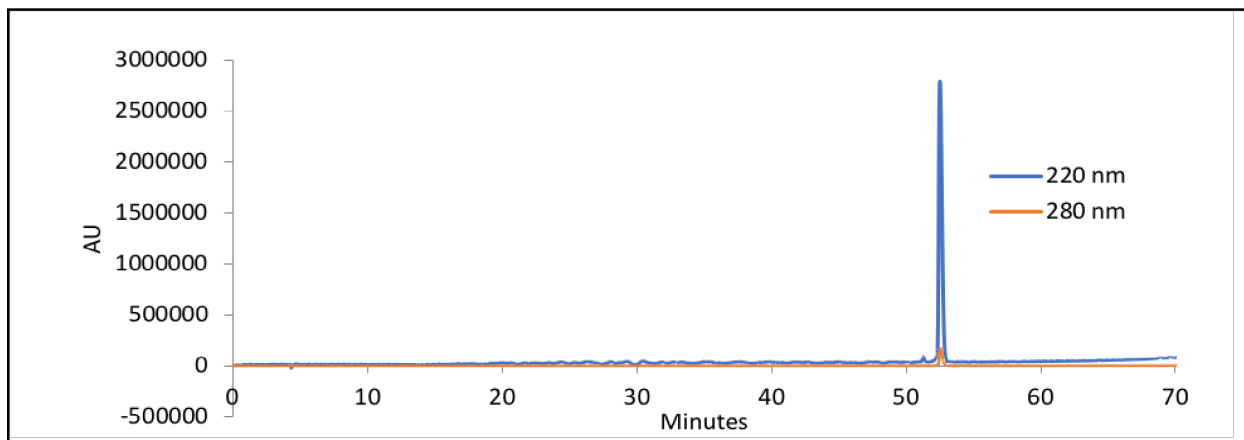


Figure S75. Analytical HPLC data for **2 α 4-AE** (QX11193). Retention time = 52.491 minutes.

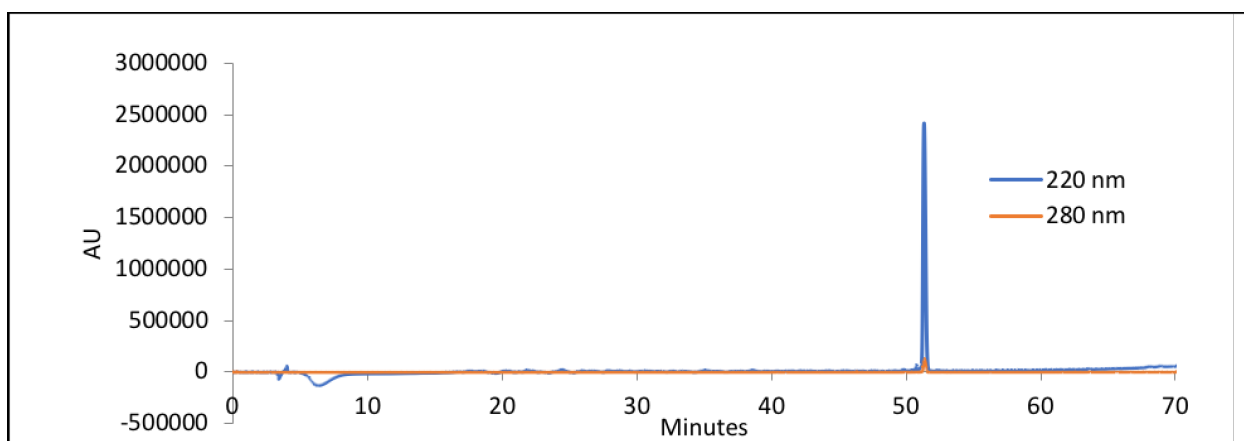


Figure S76. Analytical HPLC data for **2 α 4-AA** (QX11191). Retention time = 51.358 minutes.

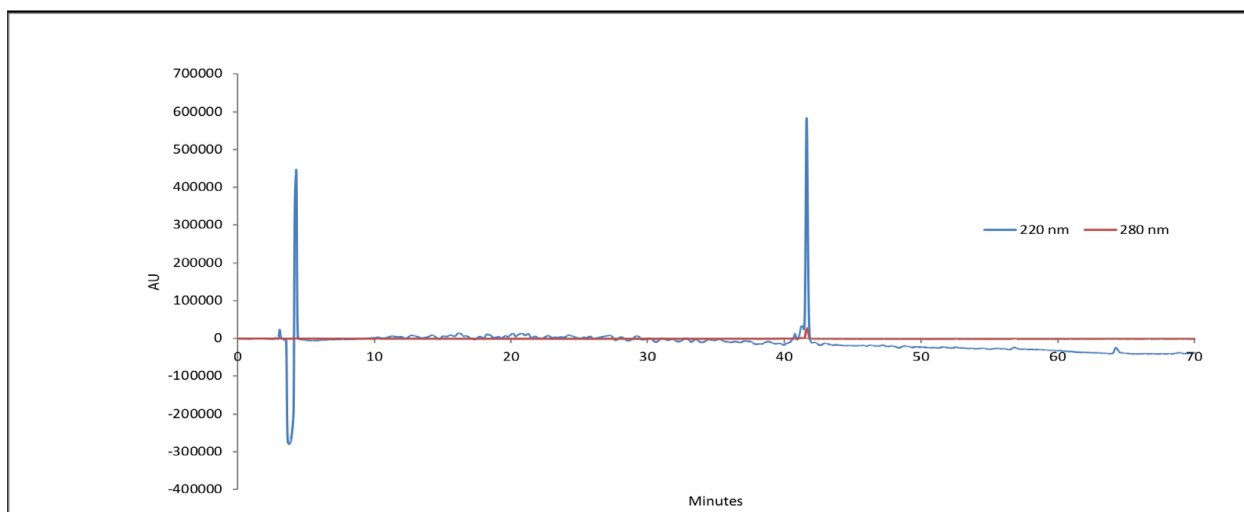


Figure S77. Analytical HPLC data for **p2 α 4** (QX10819p). Retention time = 41.625 minutes.

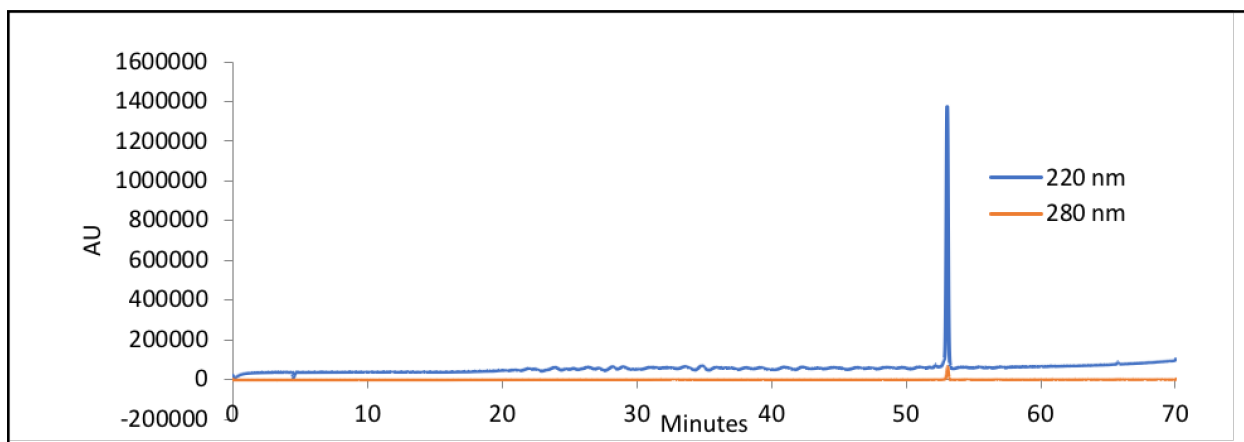


Figure S78. Analytical HPLC data for **p2 α 4-KA** (QX11195). Retention time = 53.067 minutes.

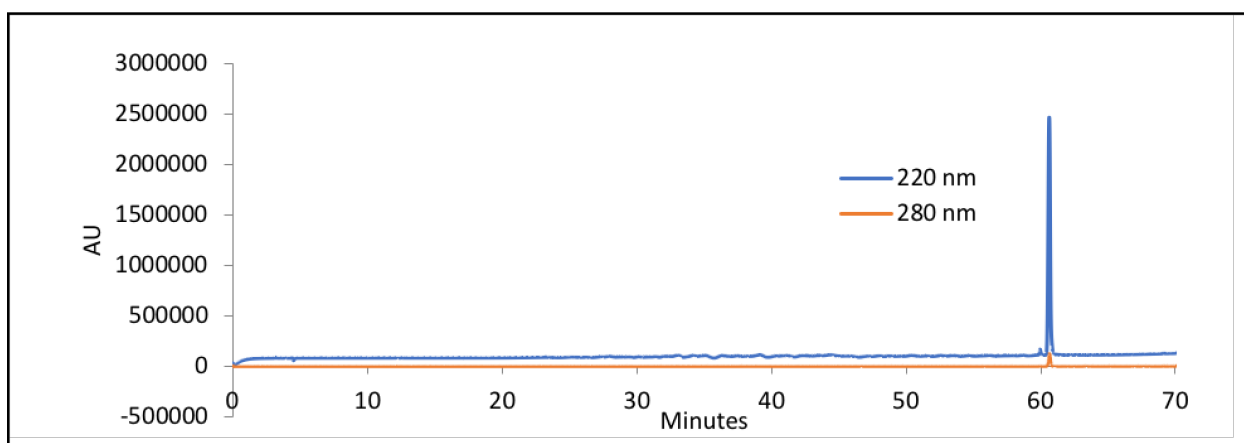


Figure S79. Analytical HPLC data for **p2 α 4-AE** (QX11196). Retention time = 60.617 minutes.

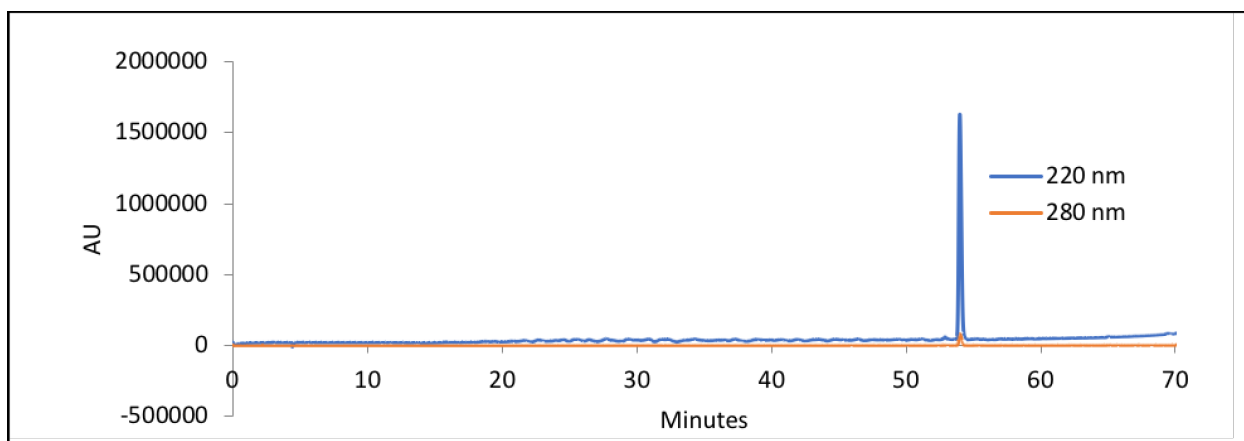


Figure S80. Analytical HPLC data for **p2 α 4-AA** (QX11194). Retention time = 54.008 minutes.

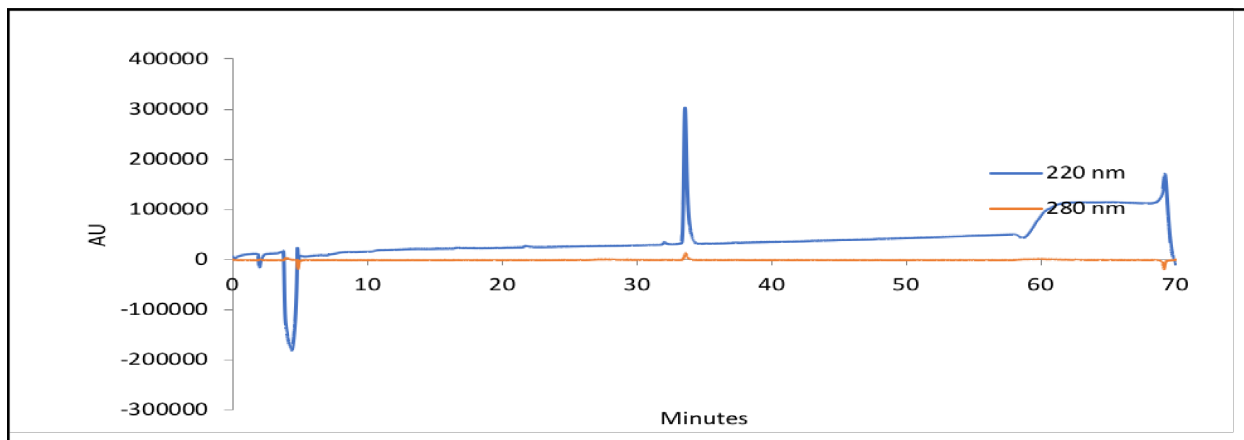


Figure S81. Analytical HPLC data for **2α6** (QX10818). Retention time = 33.617 minutes.

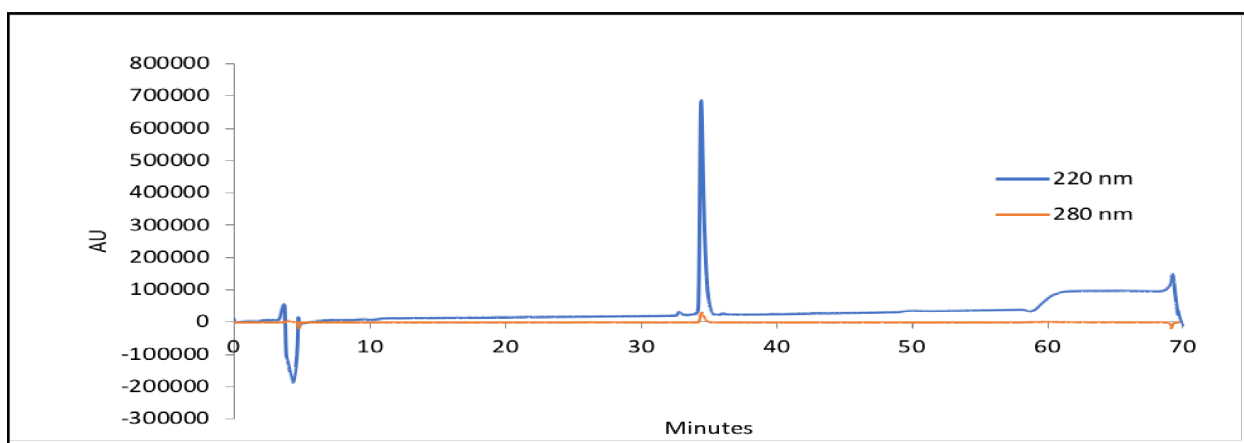


Figure S82. Analytical HPLC data for **p2α6** (QX10818p). Retention time = 34.450 minutes.

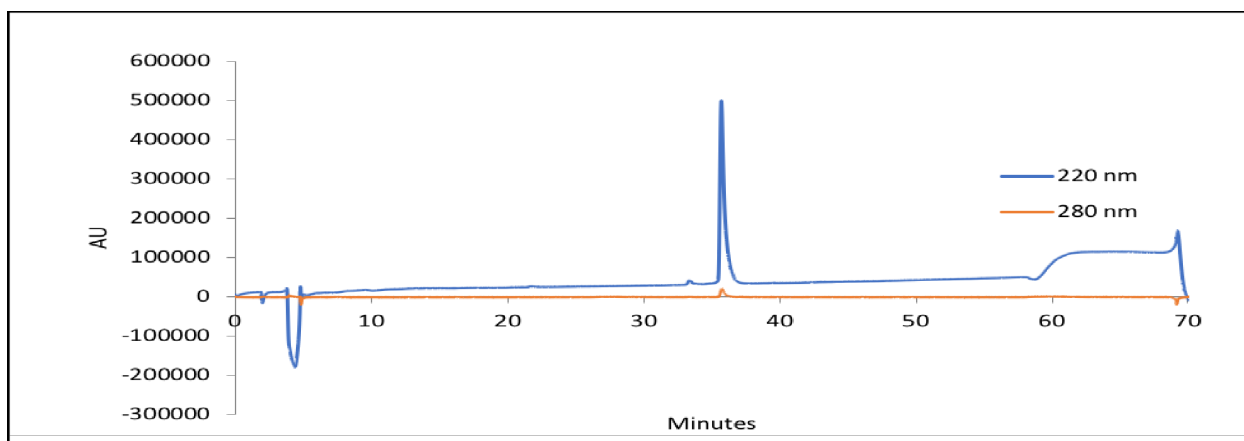


Figure S83. Analytical HPLC data for **2α7** (QX10817). Retention time = 35.750 minutes.

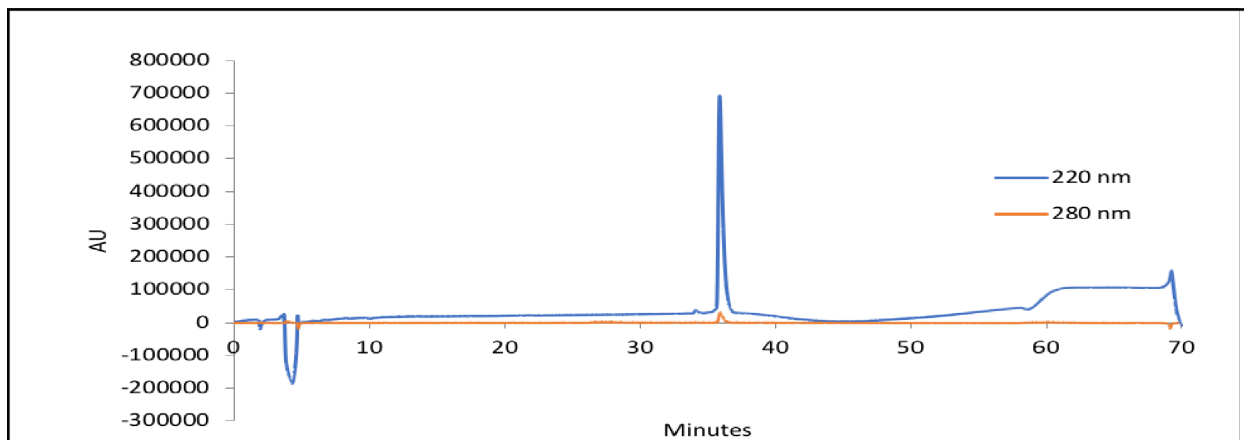


Figure S84. Analytical HPLC data for **p2 α 7** (QX10817p). Retention time = 35.900 minutes.

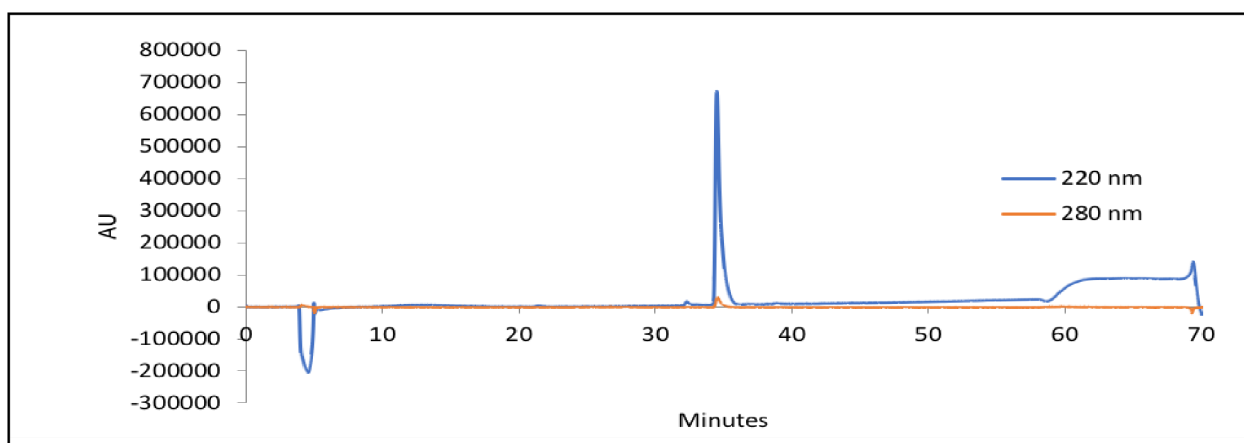


Figure S85. Analytical HPLC data for **2 α 10** (QX10816). Retention time = 34.525 minutes.

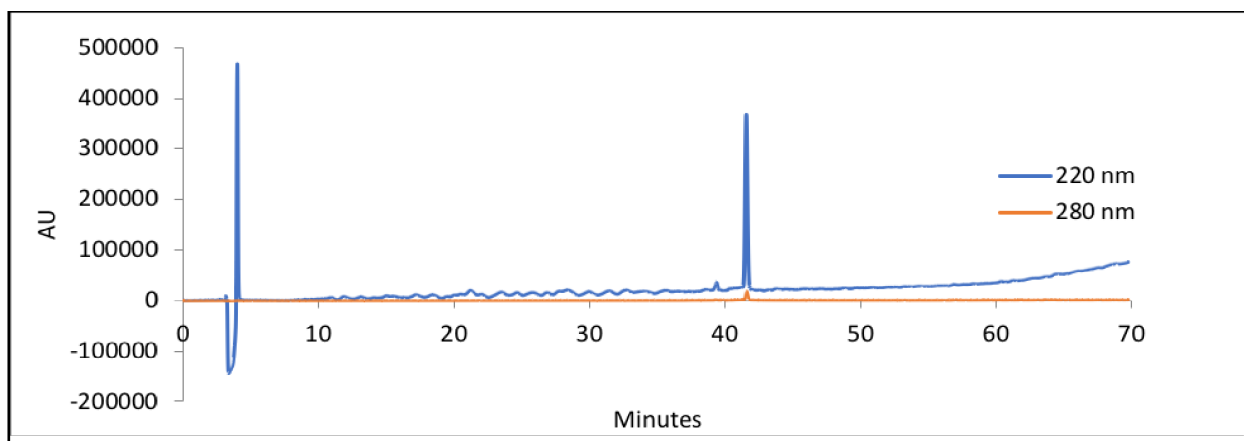


Figure S86. Analytical HPLC data for **p2 α 10** (QX10816p). Retention time = 41.600 minutes.

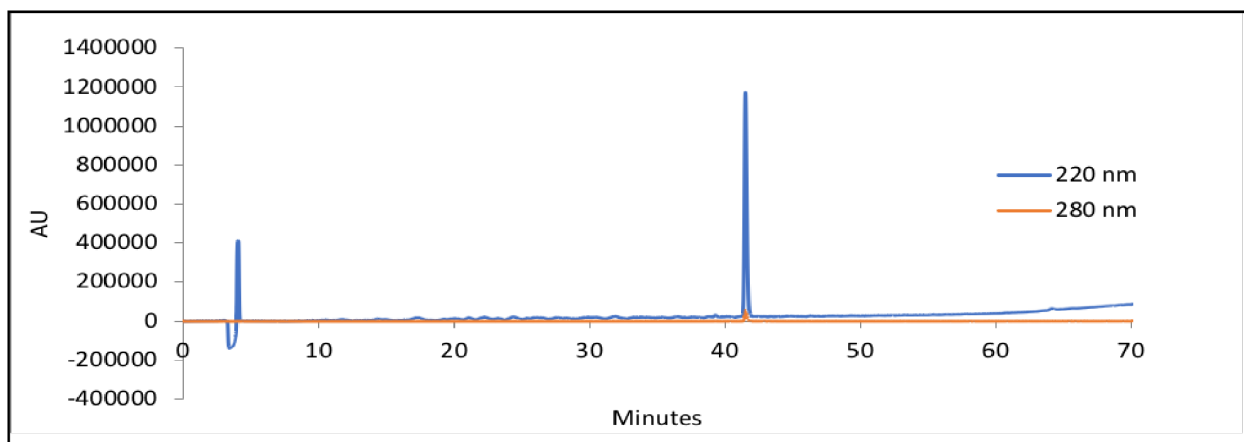


Figure S87. Analytical HPLC data for **2α14** (QX10815). Retention time = 41.525 minutes.

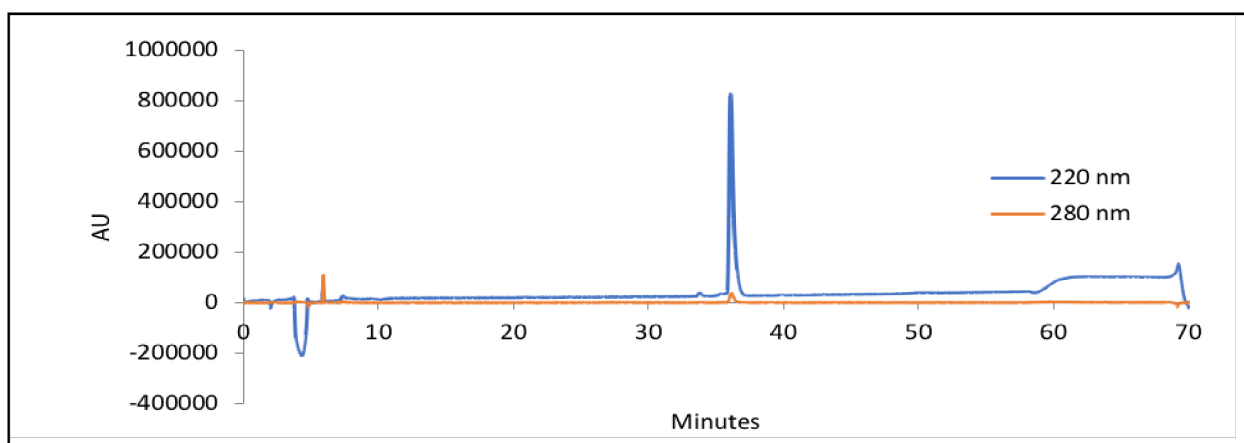


Figure S88. Analytical HPLC data for **p2α14** (QX10815p). Retention time = 36.100 minutes.

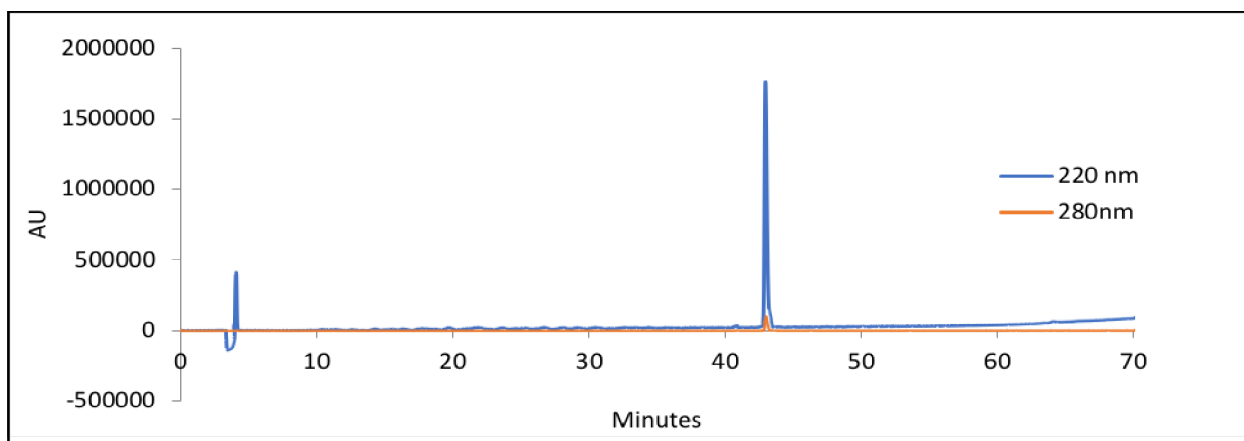


Figure S89. Analytical HPLC data for **2α18** (QX10814). Retention time = 42.958 minutes.

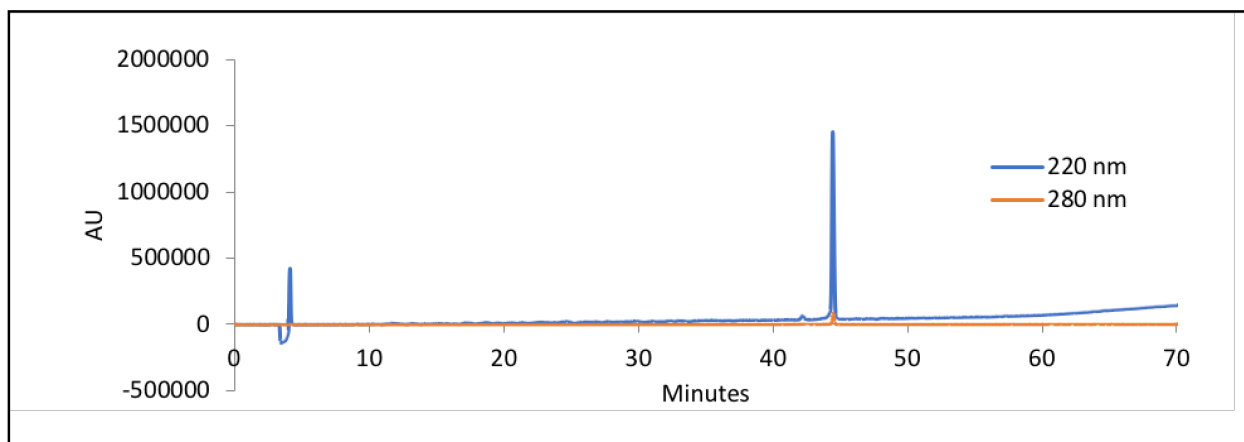


Figure S90. Analytical HPLC data for **2α18-EA** (NAB10212). Retention time = 44.467 minutes.

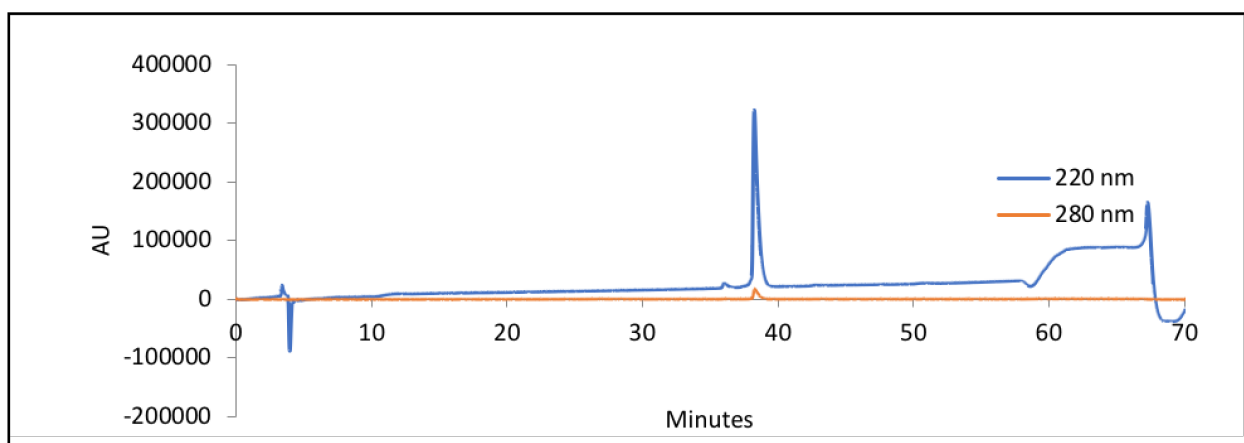


Figure S91. Analytical HPLC data for **2α18-AR** (NAB10211). Retention time = 38.275 minutes.

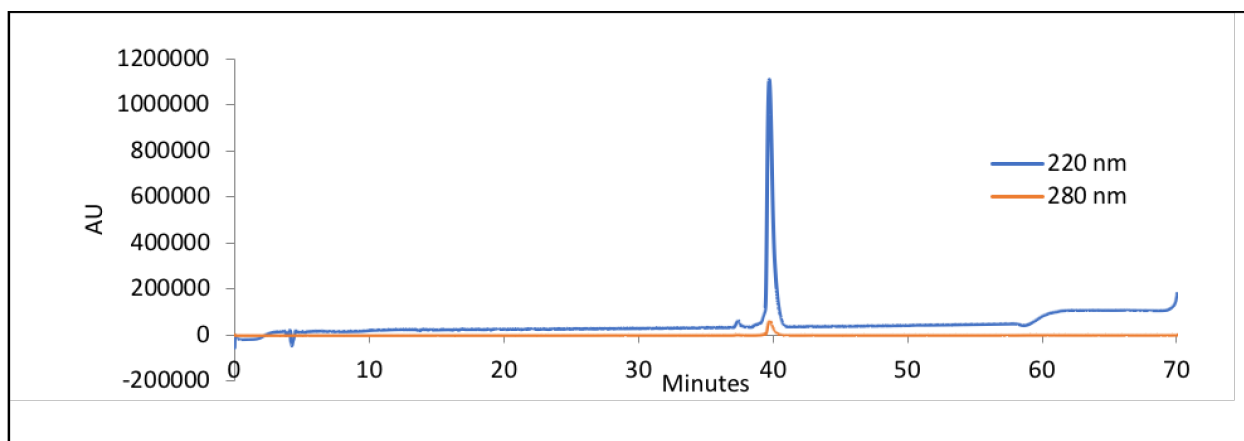


Figure S92. Analytical HPLC data for **2α18-AA** (NAB10213). Retention time = 39.775 minutes.

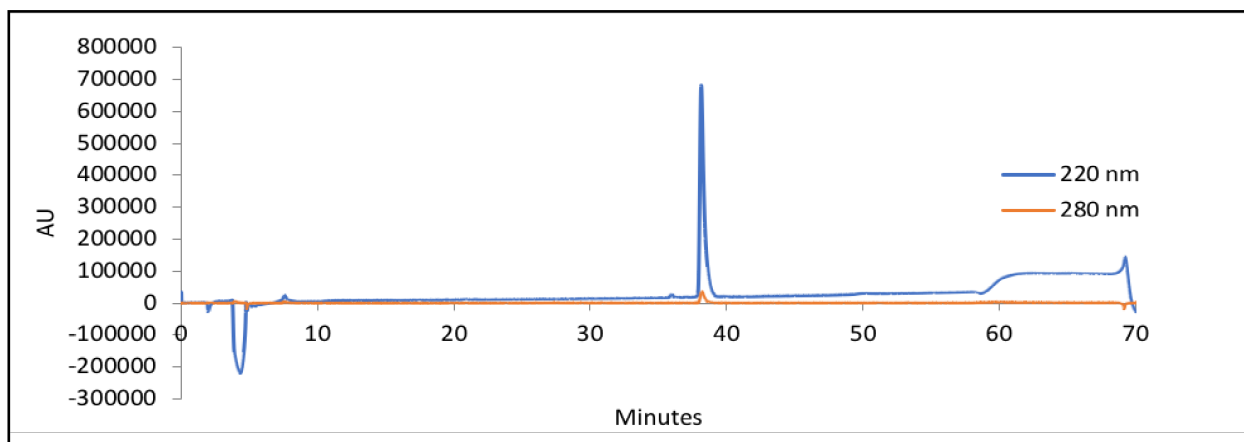


Figure S93. Analytical HPLC data for **p2 α 18** (QX10814p). Retention time = 38.117 minutes.

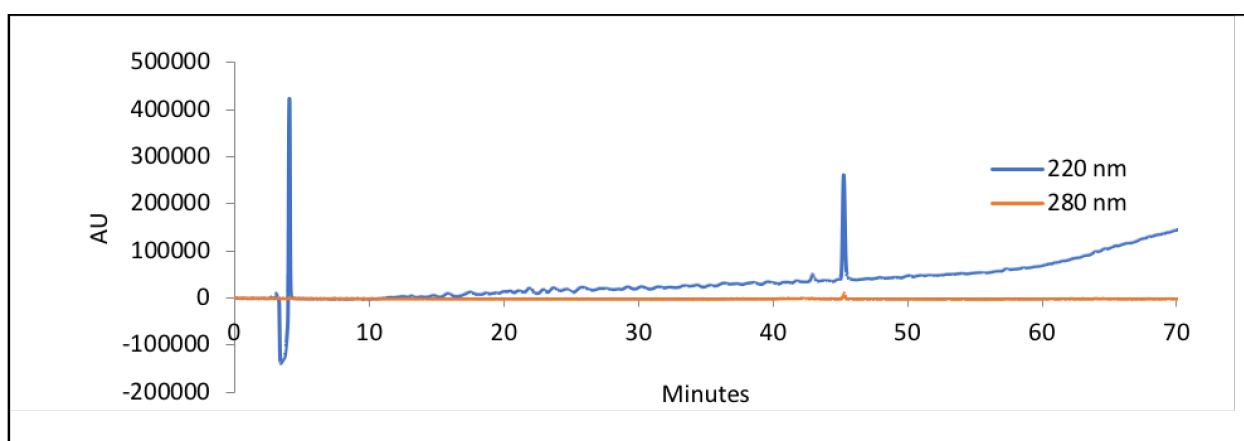


Figure S94. Analytical HPLC data for **p2 α 18-EA** (NAB10215). Retention time = 45.283 minutes.

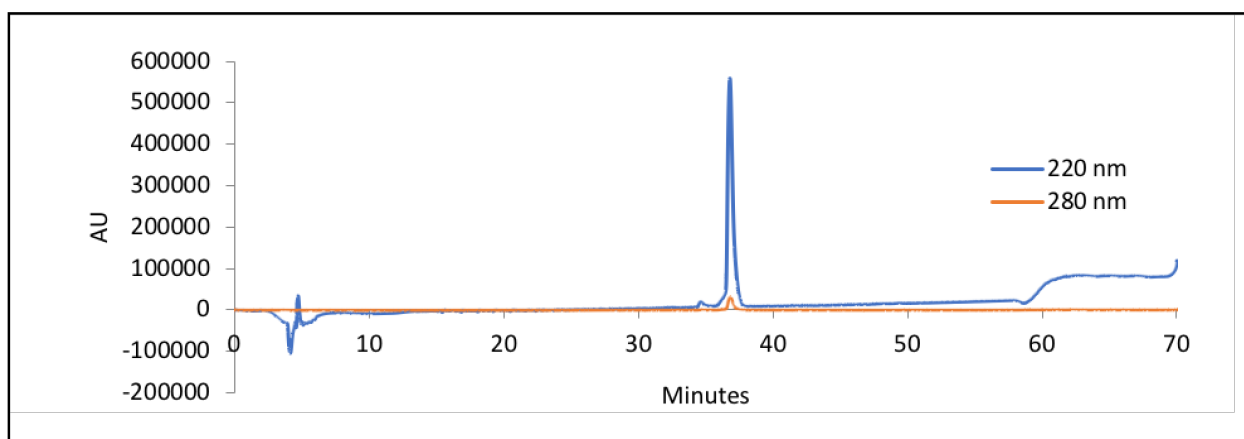


Figure S95. Analytical HPLC data for **p2 α 18-AR** (NAB10214). Retention time = 36.842 minutes.

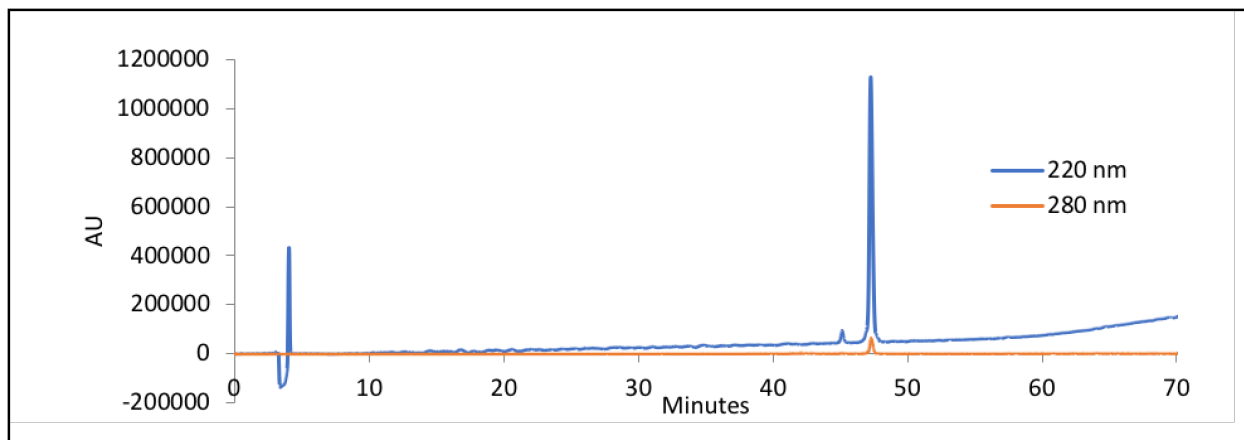


Figure S96. Analytical HPLC data for **p2 α 18-AA** (NAB10216). Retention time = 47.250 minutes.

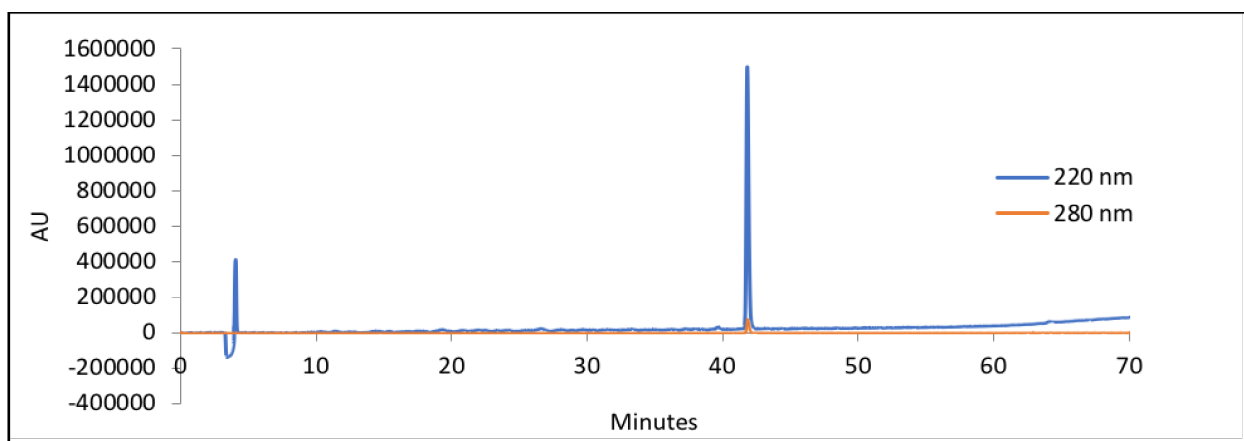


Figure S97. Analytical HPLC data for **2 α 21** (QX10813). Retention time = 41.825 minutes.

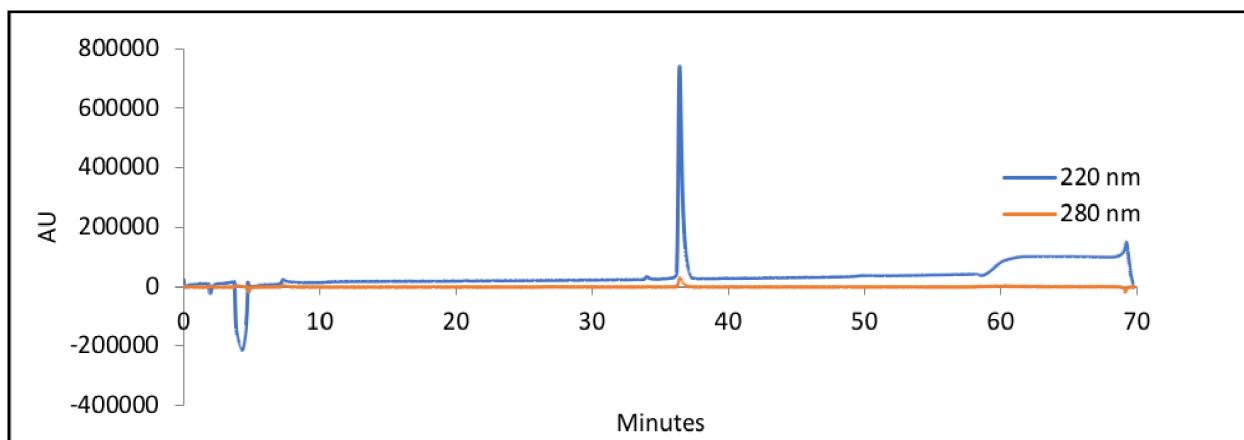


Figure S98. Analytical HPLC data for **p2 α 21** (QX10813p). Retention time = 36.450 minutes.

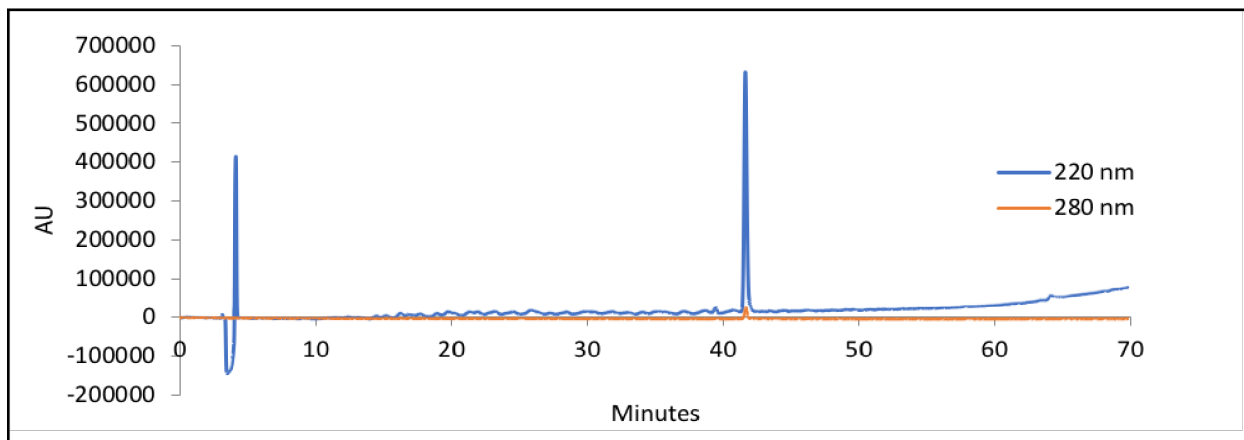


Figure S99. Analytical HPLC data for **2α25** (QX10812). Retention time = 41.667 minutes.

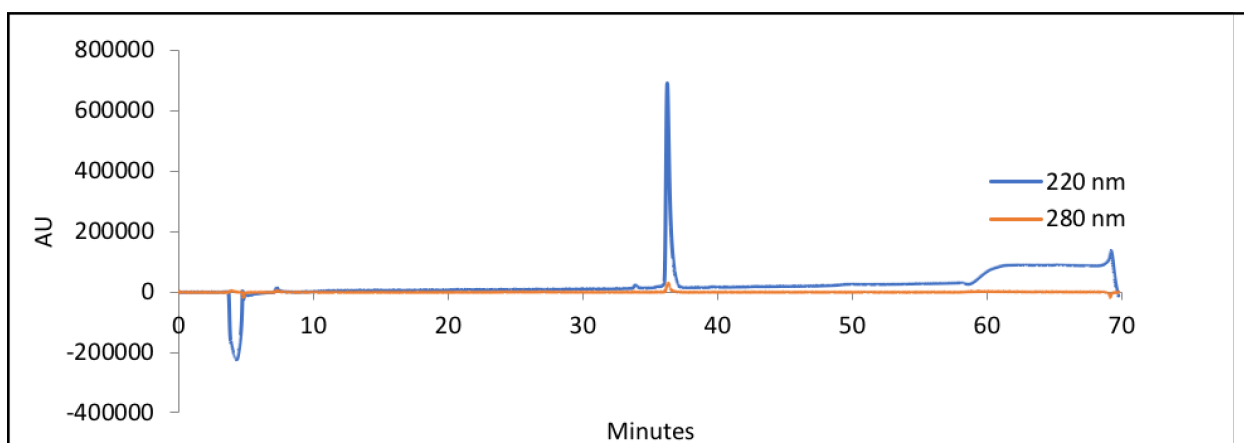


Figure S100. Analytical HPLC data for **p2α25** (QX10812p). Retention time = 36.300 minutes.

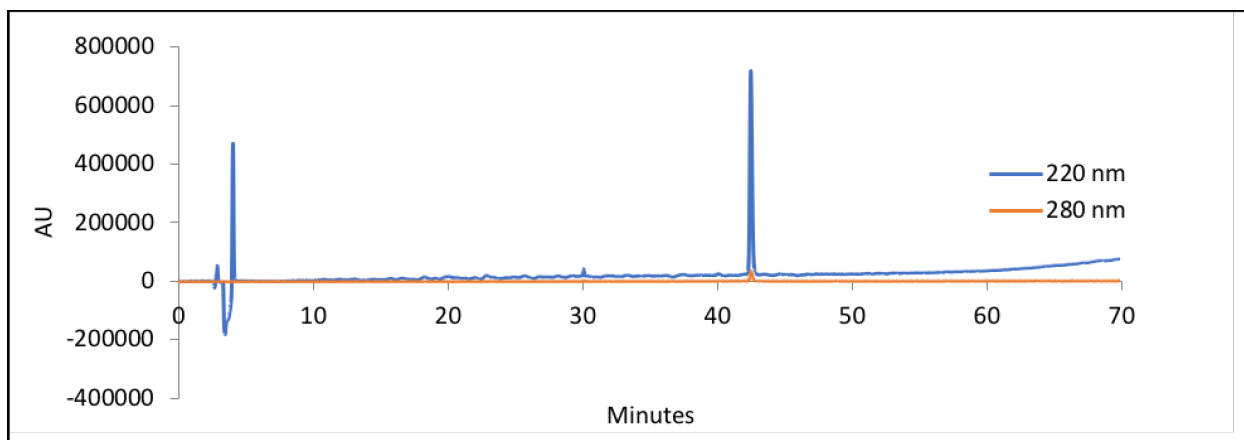


Figure S101. Analytical HPLC data for **2α28** (QX10811). Retention time = 42.492 minutes.

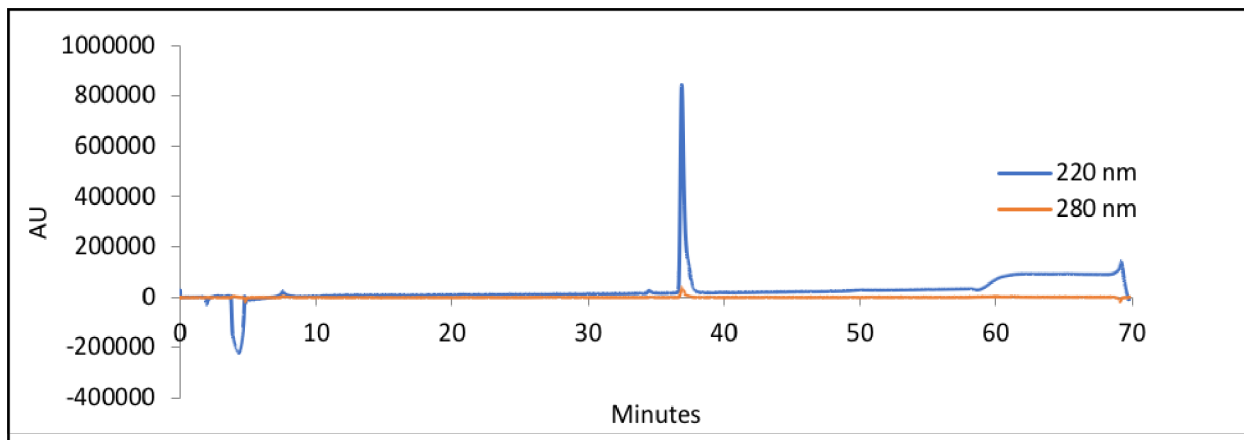


Figure S102. Analytical HPLC data for **p2a28** (QX10811p). Retention time = 36.917 minutes.

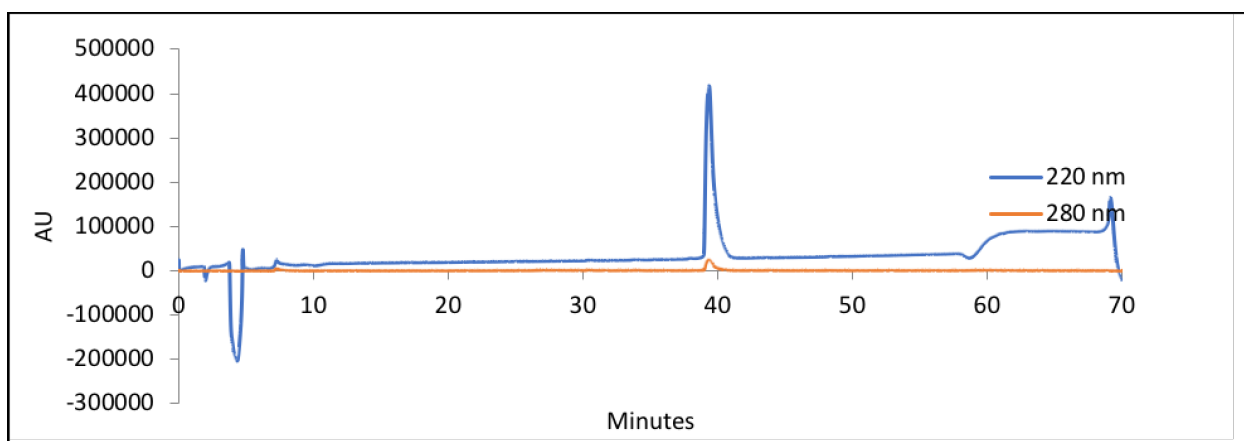


Figure S103. Analytical HPLC data for **3a1** (QX10714). Retention time = 39.442 minutes.

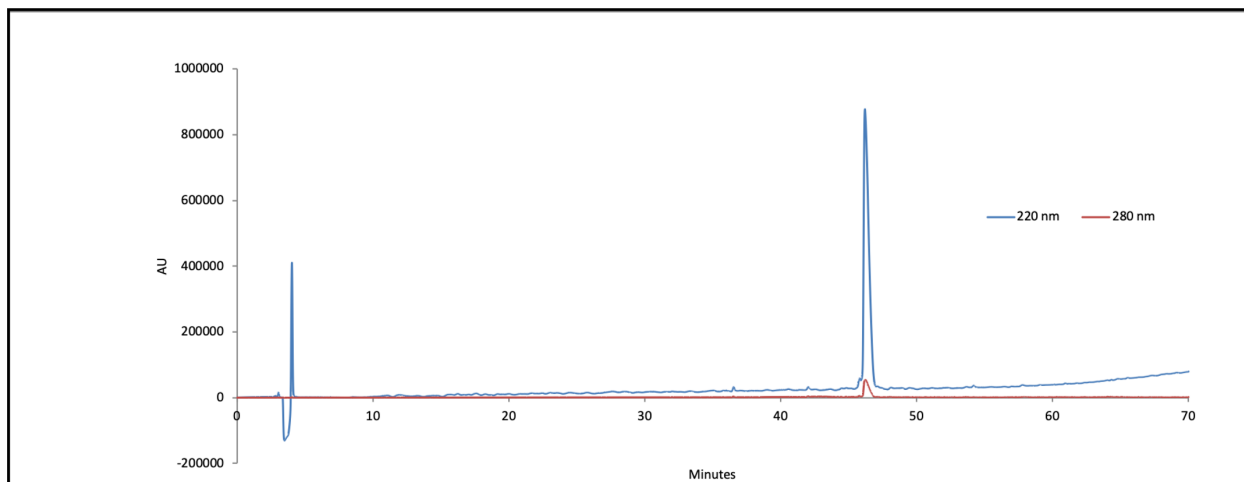


Figure S104. Analytical HPLC data for **3a1-ER** (QX21352). Retention time = 46.250 minutes.

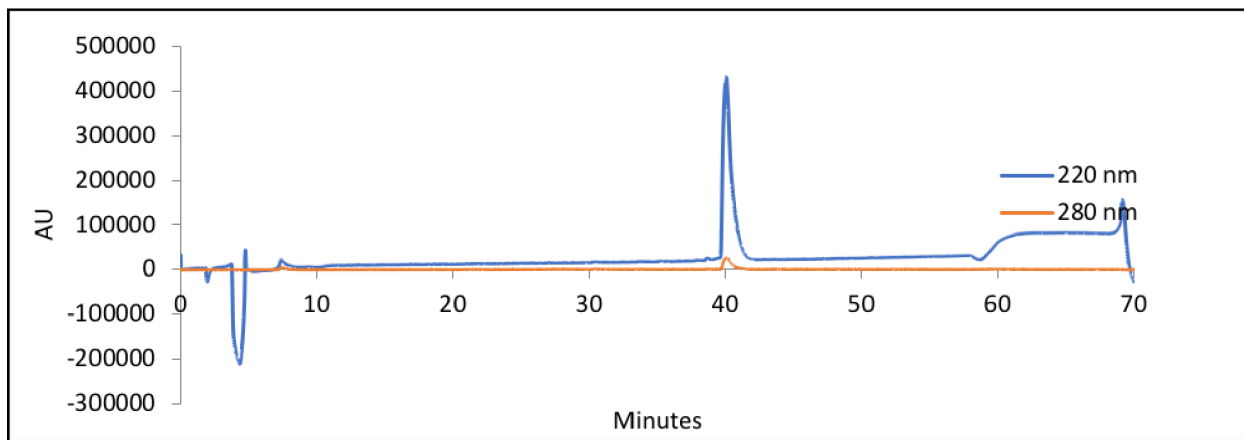


Figure S105. Analytical HPLC data for **3α1-EA** (QX10713). Retention time = 40.150 minutes.

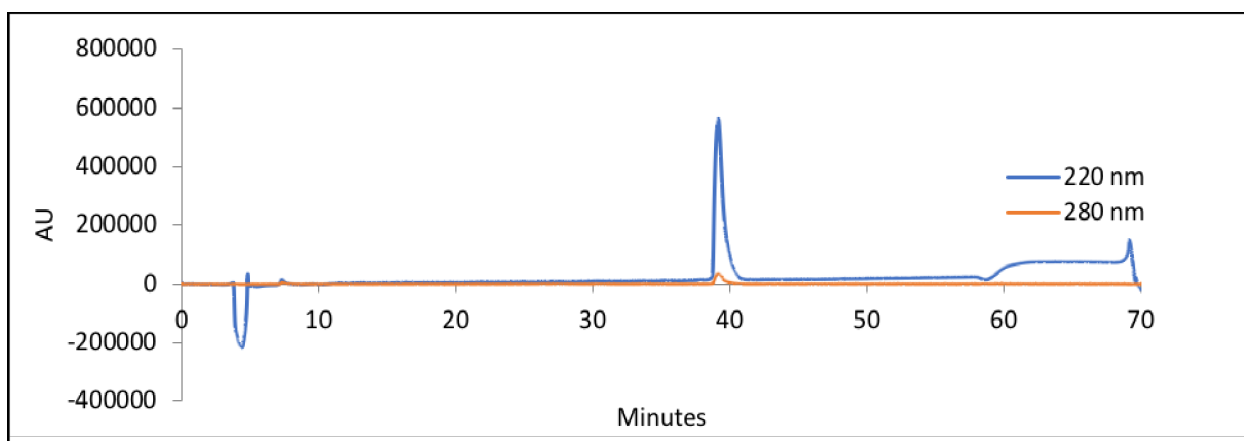


Figure S106. Analytical HPLC data for **3α1-AK** (QX10712). Retention time = 39.217 minutes.

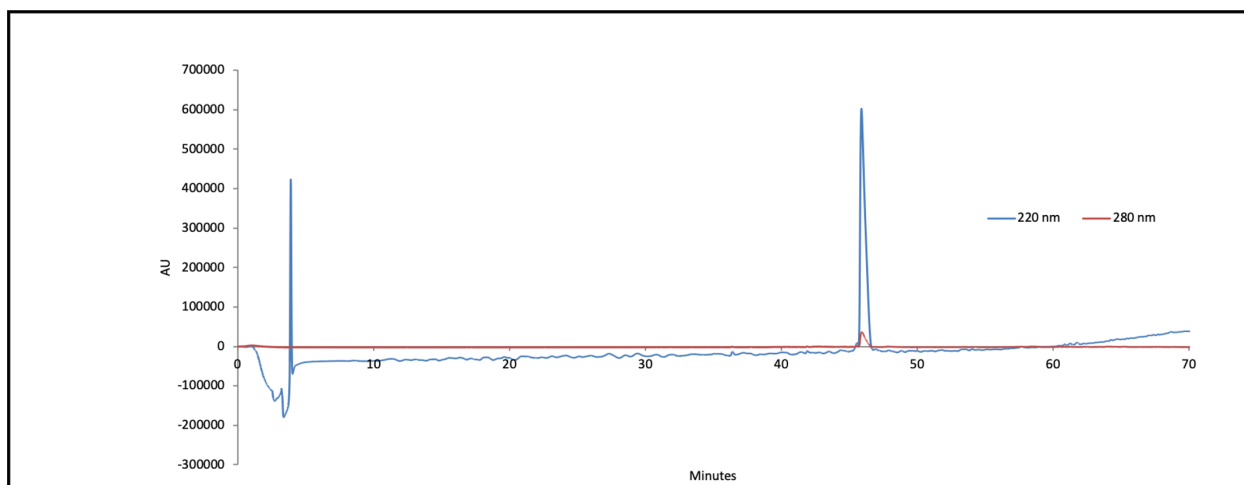


Figure S107. Analytical HPLC data for **3α1-AR** (QX21531). Retention time = 45.933 minutes.

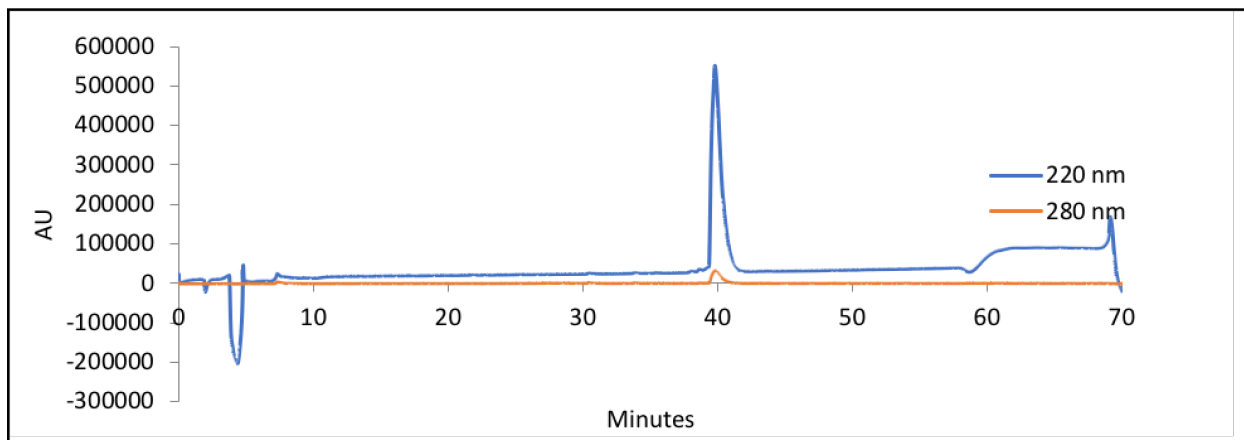


Figure S108. Analytical HPLC data for **3α1-AA** (QX10711). Retention time = 39.858 minutes.

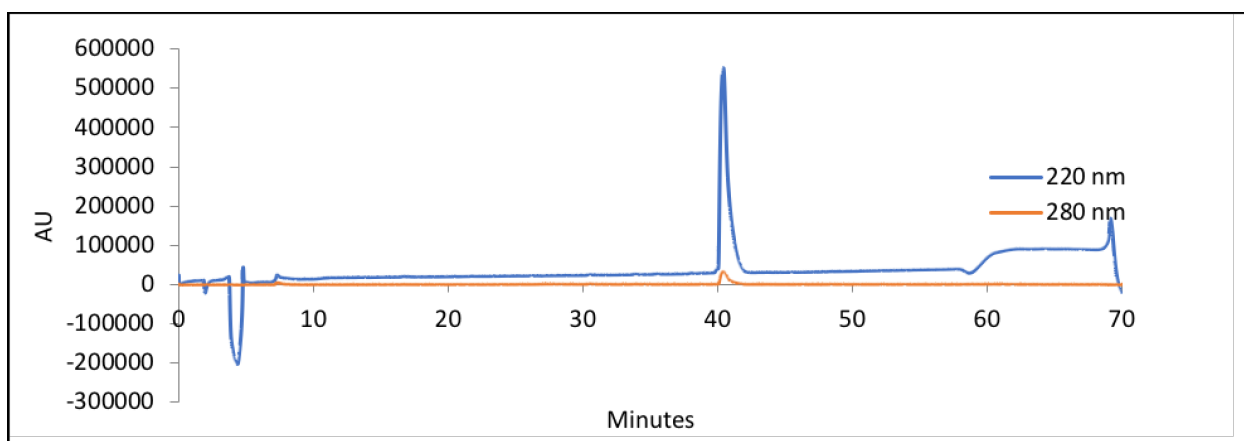


Figure S109. Analytical HPLC data for **p3α1** (QX10718). Retention time = 40.483 minutes.

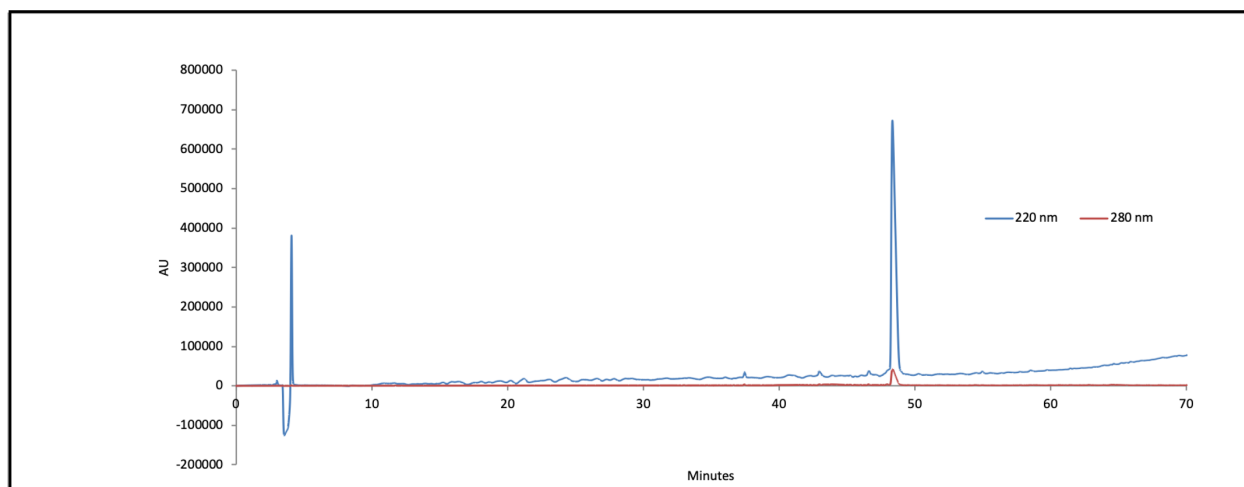


Figure S110. Analytical HPLC data for **p3α1-ER** (QX21354). Retention time = 48.367 minutes.

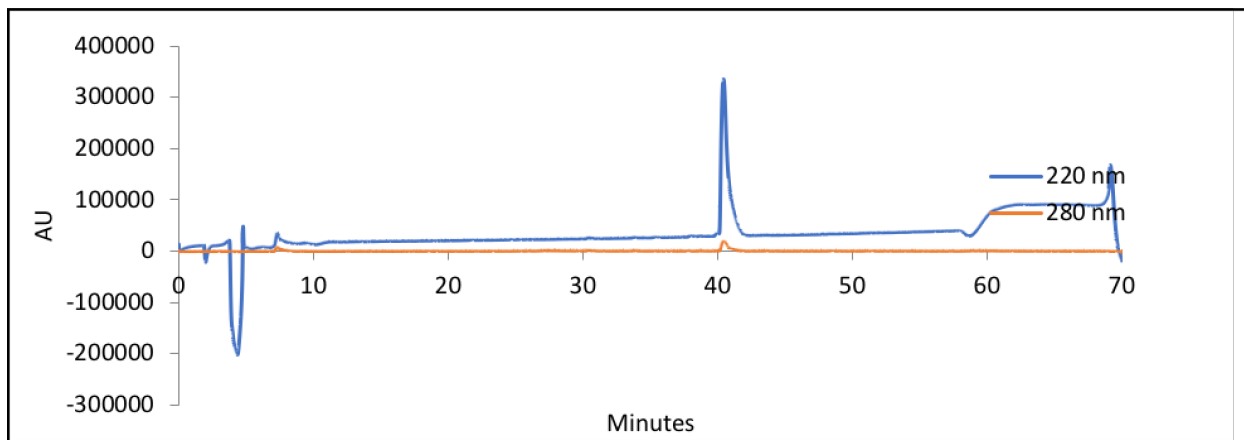


Figure S111. Analytical HPLC data for **p3α1-EA** (QX10717). Retention time = 40.492 minutes.

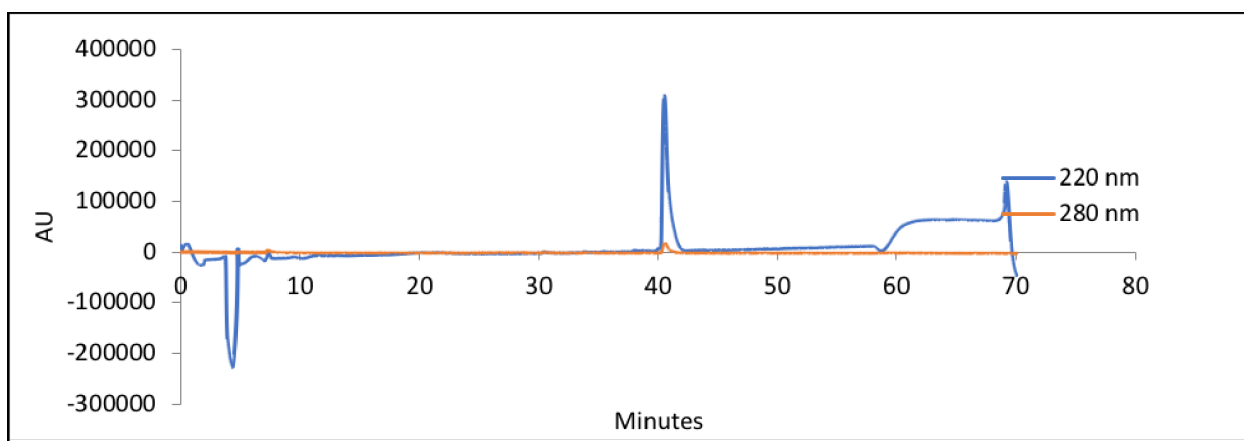


Figure S112. Analytical HPLC data for **p3α1-AK** (QX10716). Retention time = 40.575 minutes.

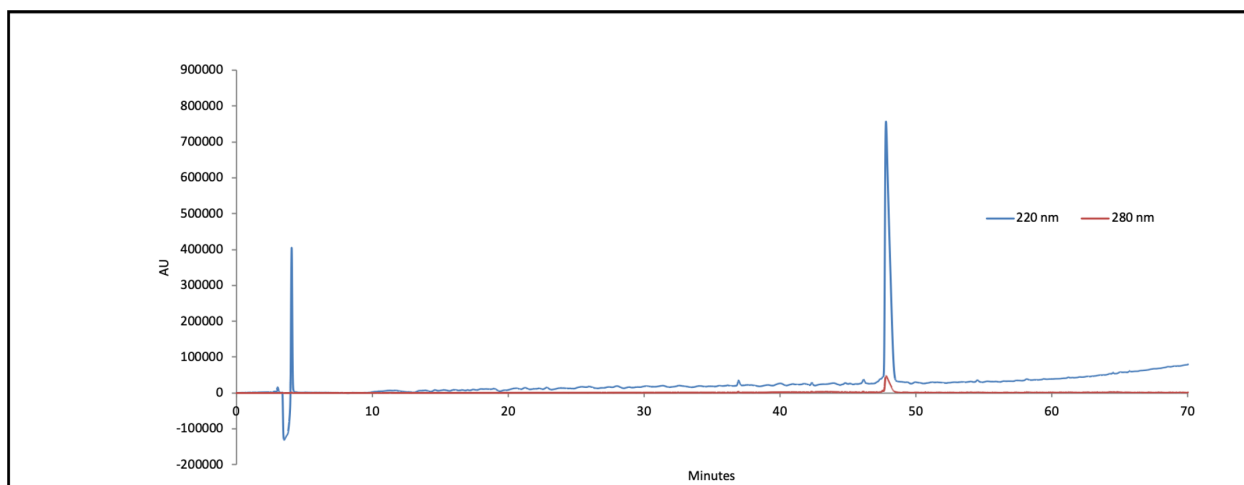


Figure S113. Analytical HPLC data for **p3α1-AR** (QX21353). Retention time = 47.808 minutes.

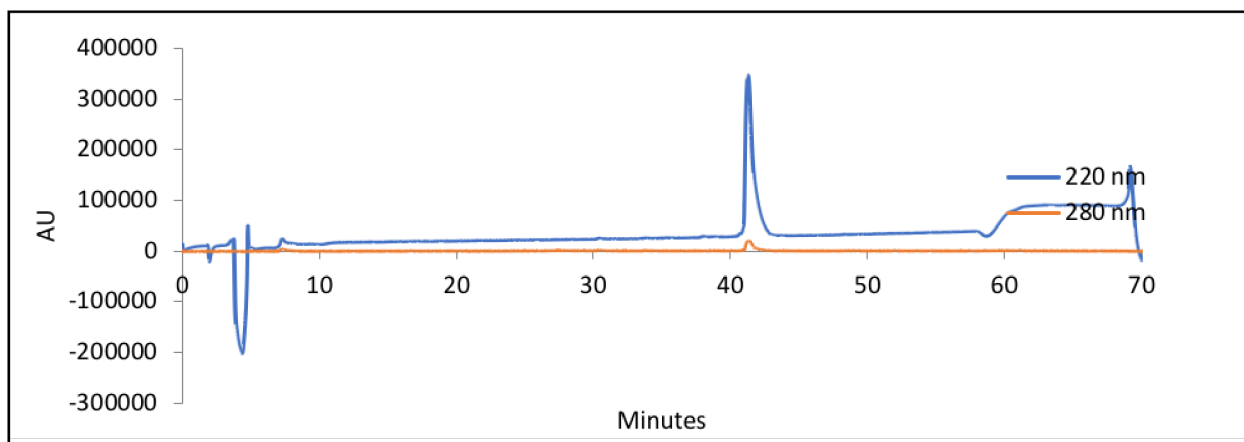


Figure S114. Analytical HPLC data for **p3a1-AA** (QX10715). Retention time = 41.375 minutes.

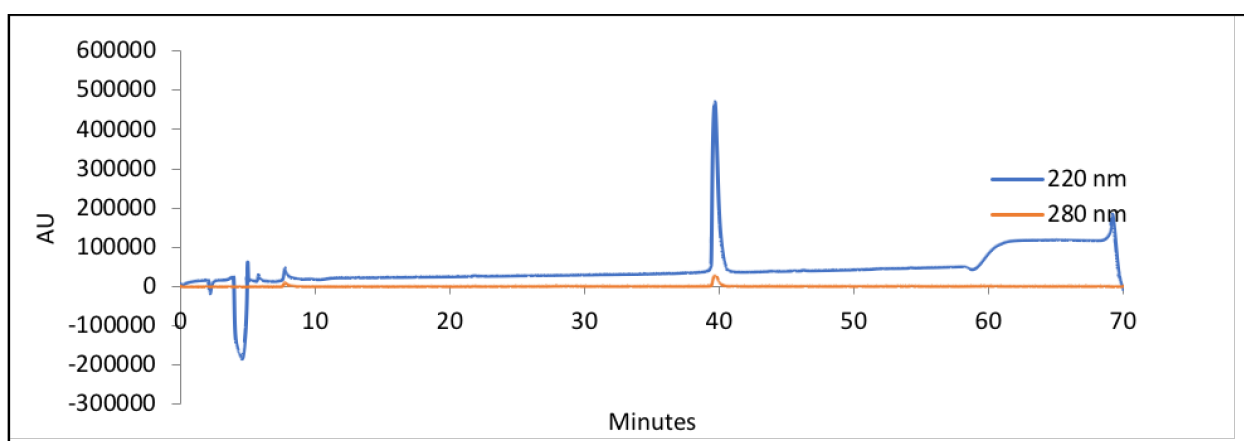


Figure S115. Analytical HPLC data for **3a6** (QX10511). Retention time = 39.725 minutes.

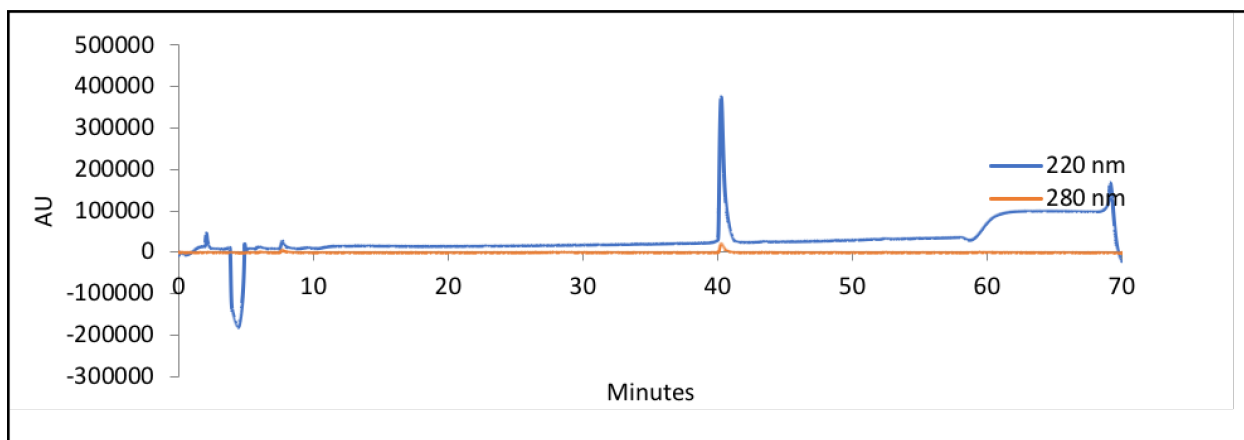


Figure S116. Analytical HPLC data for **3a6-KA** (QX10512). Retention time = 40.317 minutes.

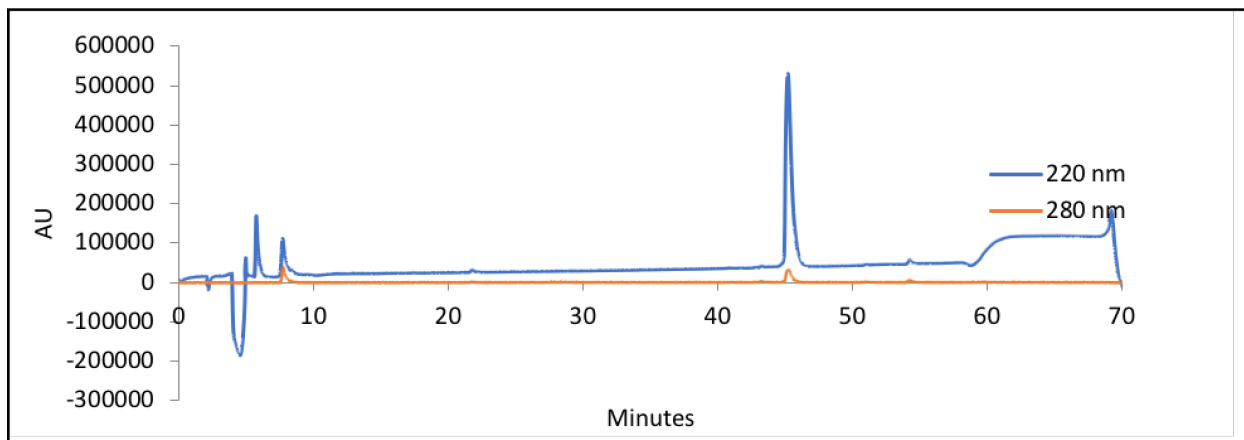


Figure S117. Analytical HPLC data for **3a6-AE** (QX10513). Retention time = 45.183 minutes.

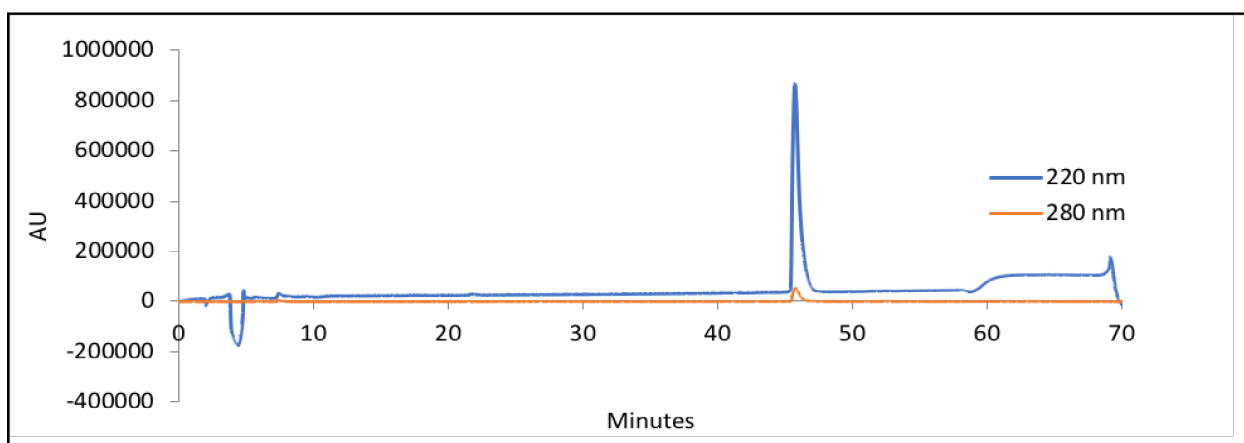


Figure S118. Analytical HPLC data for **3a6-AA** (QX10514). Retention time = 45.808 minutes.

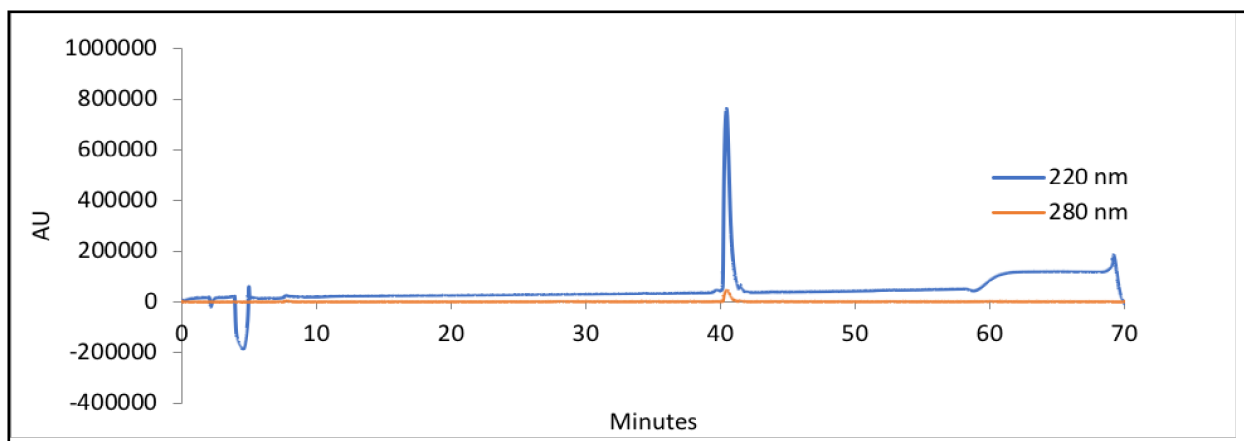


Figure S119. Analytical HPLC data for **p3a6** (QX10515). Retention time = 40.425 minutes.

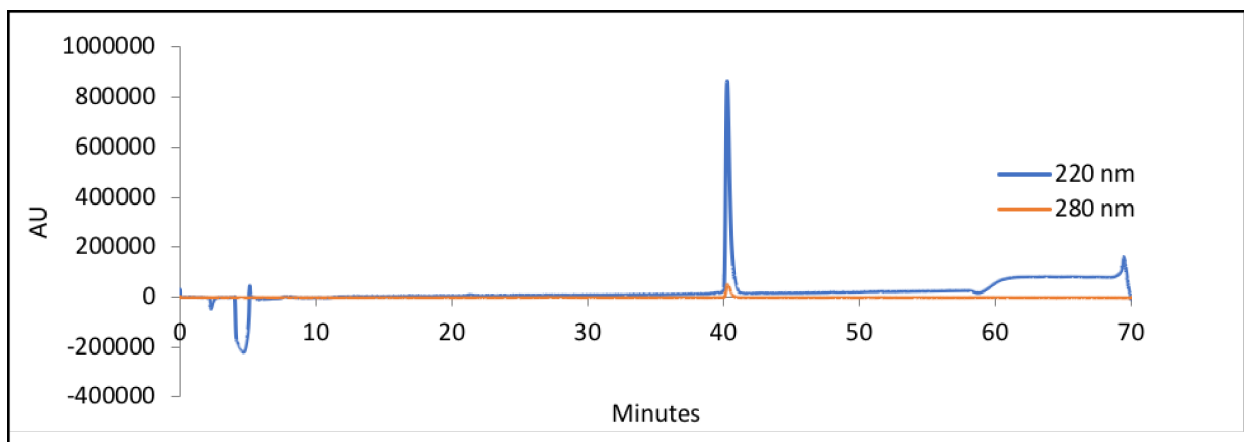


Figure S120. Analytical HPLC data for **p3 α 6-KA** (QX10516). Retention time = 40.267 minutes.

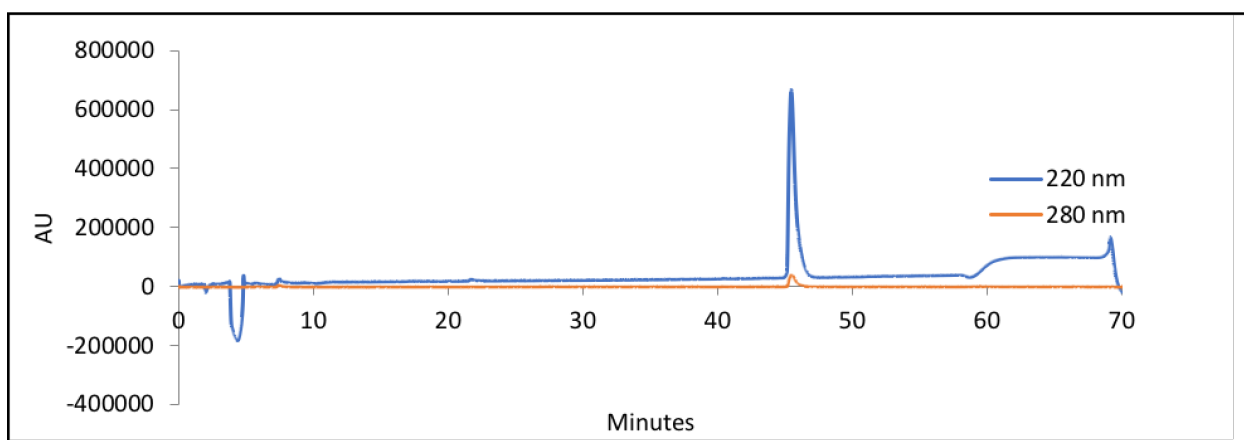


Figure S121. Analytical HPLC data for **p3 α 6-AE** (QX10517). Retention time = 45.517 minutes.

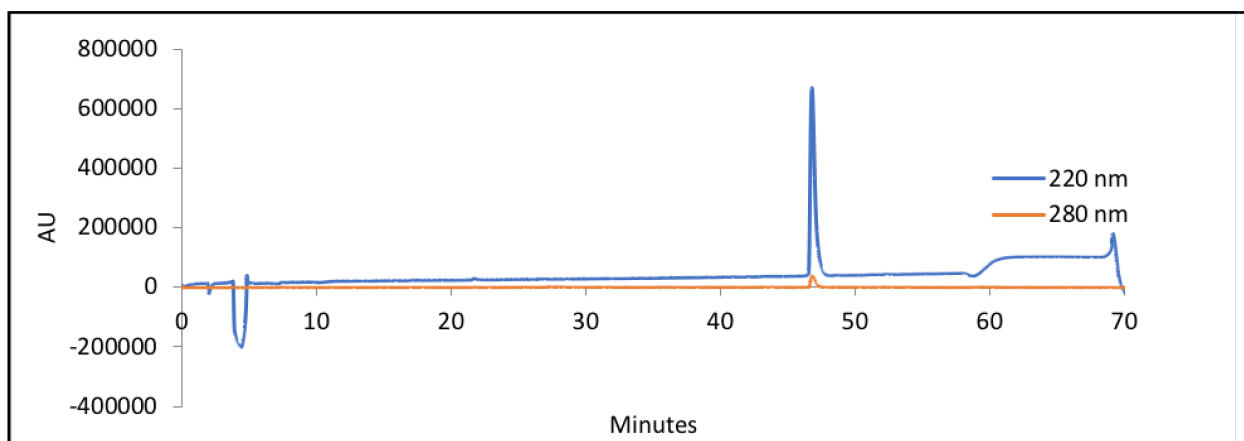


Figure S122. Analytical HPLC data for **p3 α 6-AA** (QX10518). Retention time = 46.850 minutes.

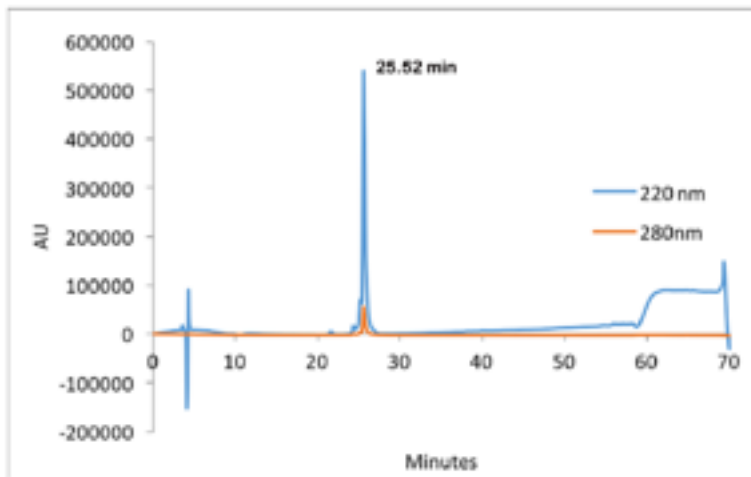


Figure S123. Analytical HPLC data for β 18 (EL1101). Retention time = 25.52 minutes.

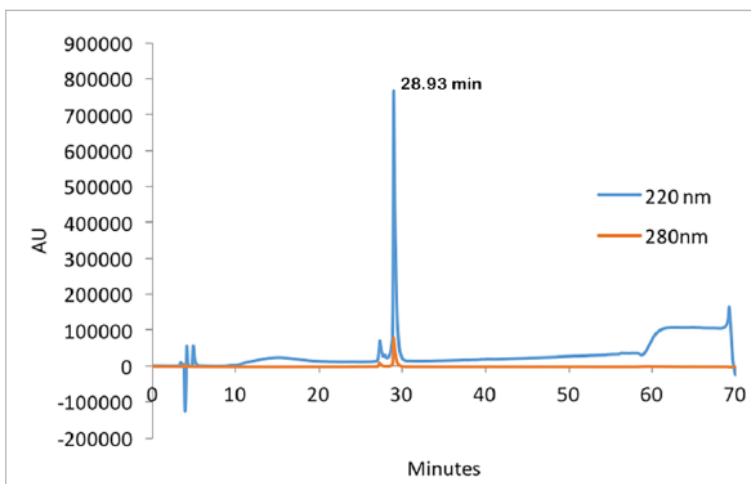


Figure S124. Analytical HPLC data for β 18-DA. Retention time = 28.93 minutes.

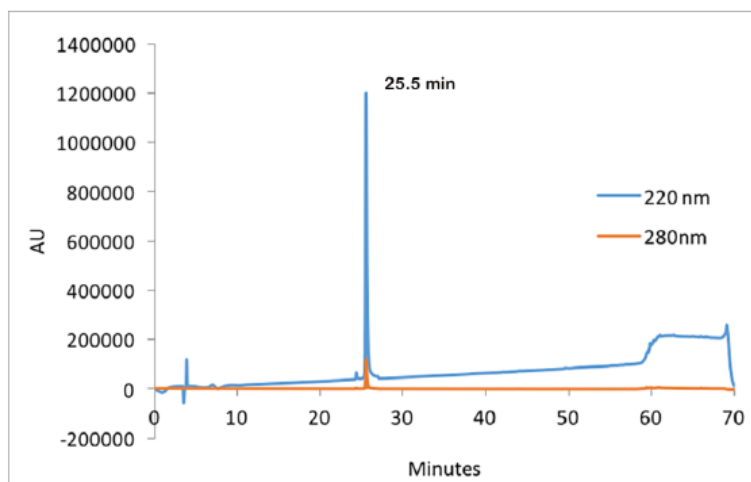


Figure S125. Analytical HPLC data for β 18-SR (ML1006). Retention time = 25.5 minutes.

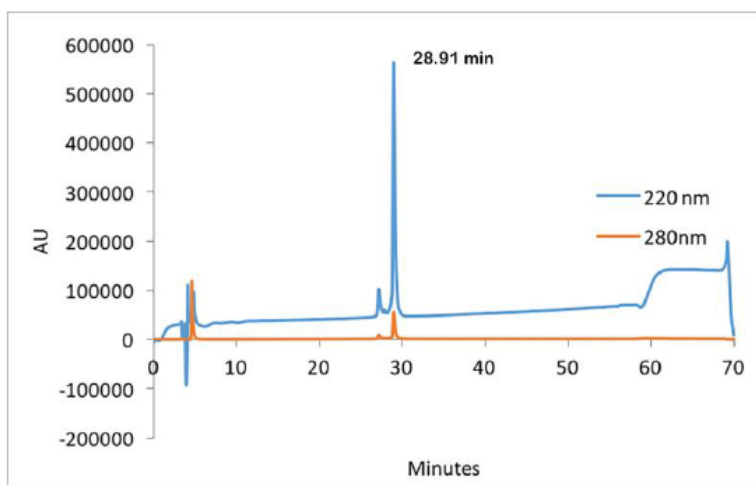


Figure S126. Analytical HPLC data for β 18-SA. Retention time = 28.91 minutes.

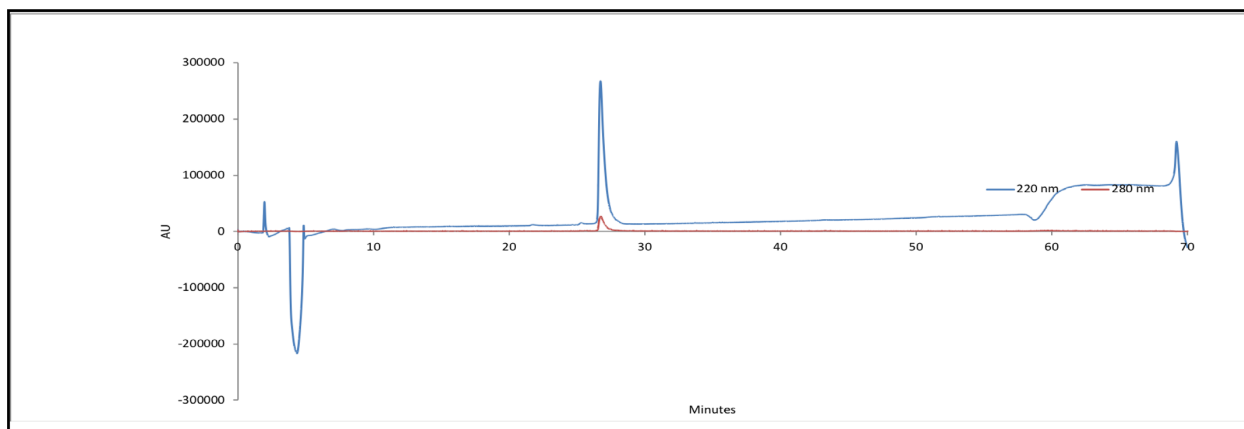


Figure S127. Analytical HPLC data for $p\beta$ 18 (QX10492). Retention time = 26.767 minutes.

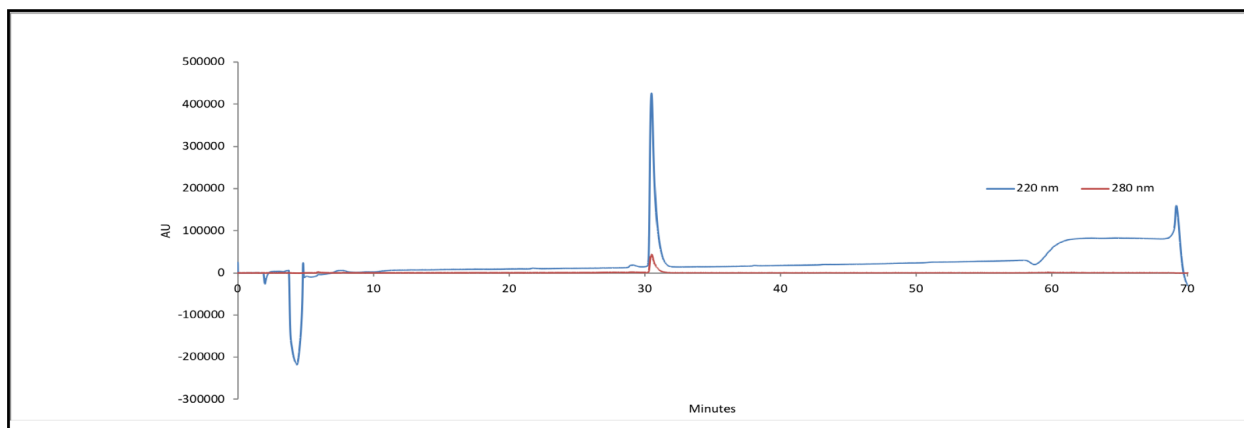


Figure S128. Analytical HPLC data for $p\beta$ 18-DA (QX10494). Retention time = 30.533 minutes.

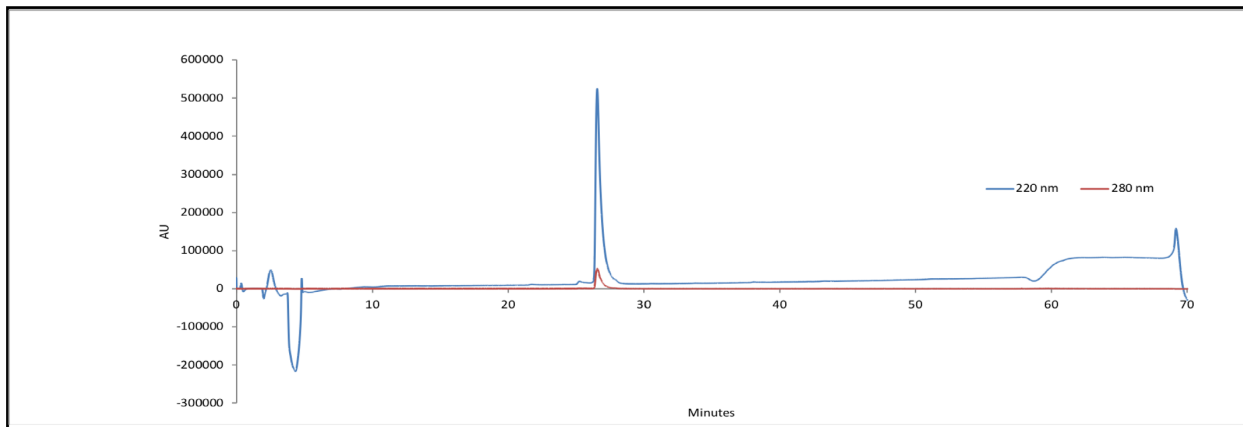


Figure S129. Analytical HPLC data for **pβ18-SA (QX10491)**. Retention time = 26.525 minutes.

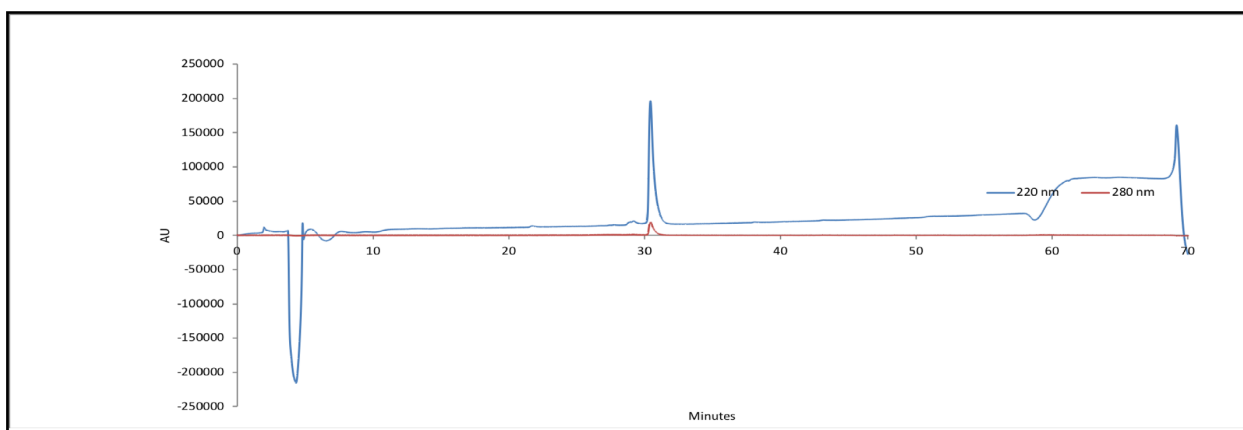


Figure S130. Analytical HPLC data for **pβ18-SA (QX10493)**. Retention time = 30.458 minutes.

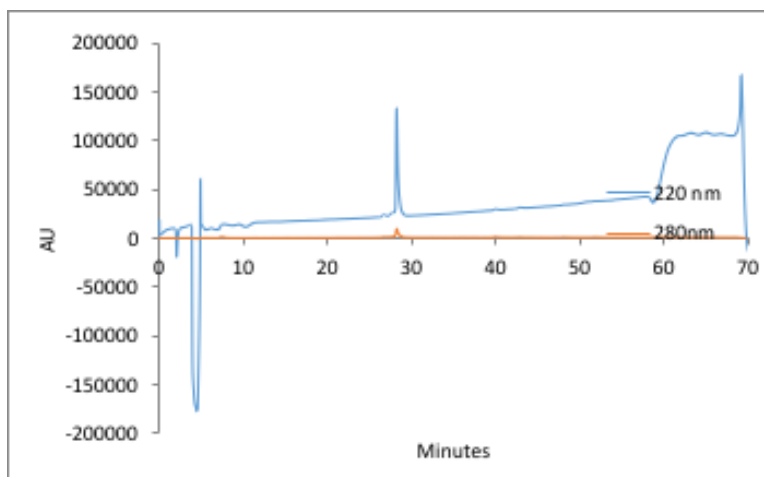


Figure S131. HPLC data for **β23-EA (SD2006#2)**.

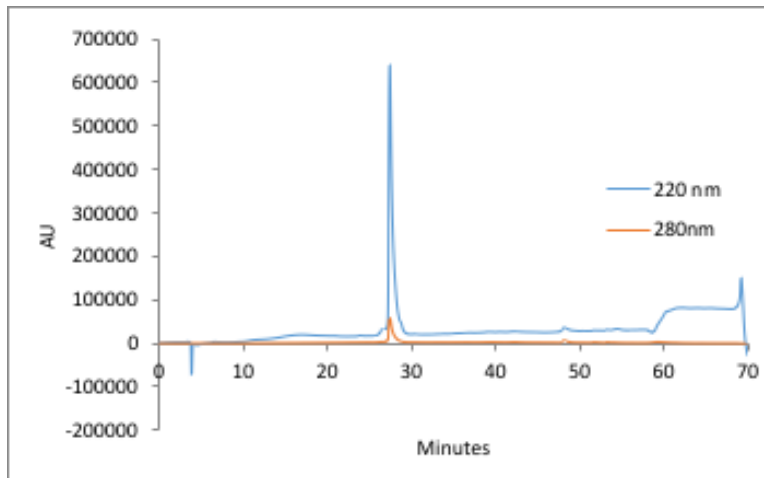


Figure S132. HPLC MS data for β 23-AR (SD2018).

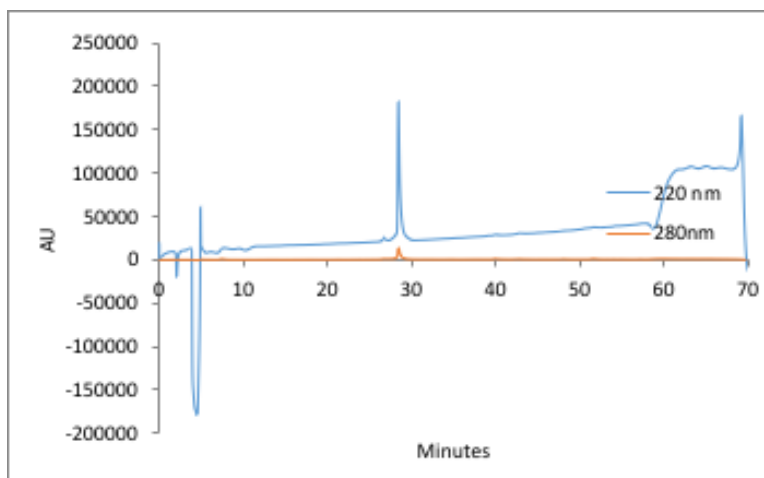


Figure S133. HPLC MS data for β 23-AA (SD2006#1).

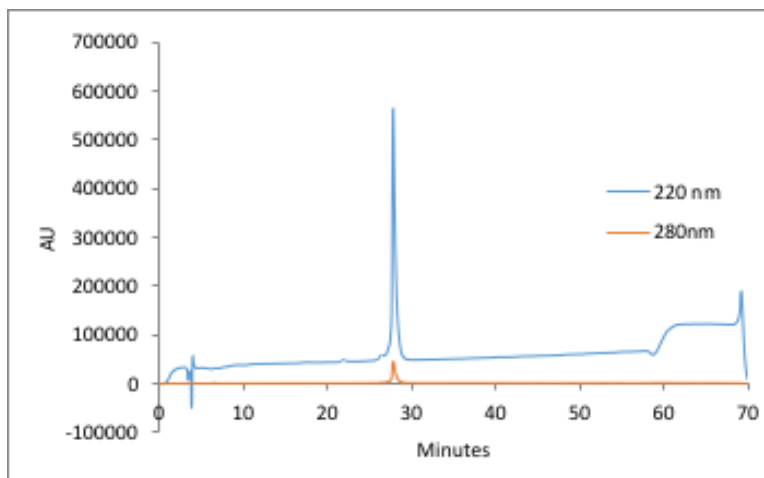


Figure S134. HPLC MS data for p β 23-EA (SD2006#2C).

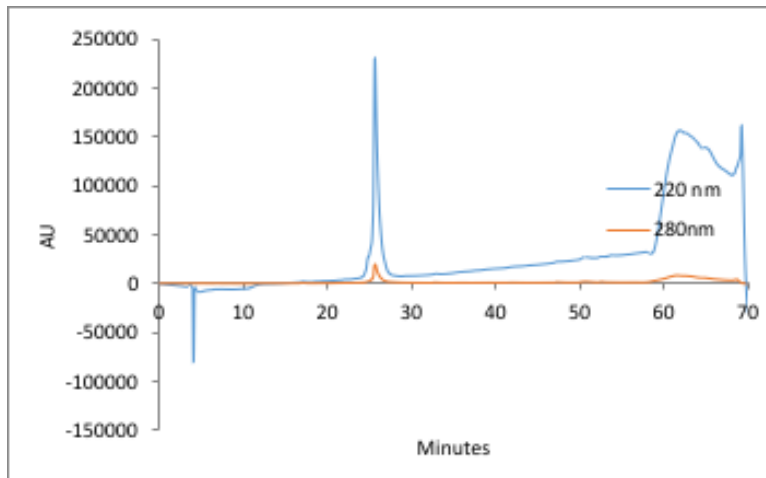


Figure S135. HPLC MS data for **pβ23-AR** (SD2018C).

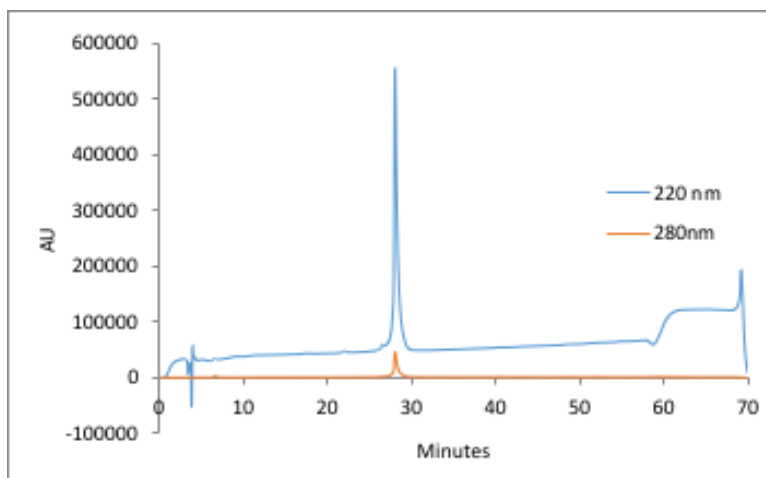


Figure S136. HPLC MS data for **pβ23-AA** (SD2006#1C).

2. Biophysical characterization of peptide variants

Self-association Properties of Peptides: Size Exclusion Chromatography: Previously characterized peptide **1CW** adopts a homotrimeric self-association state in solution, whereas **GCN4** adopts a homodimeric self-association state in solution. The peptides explored here (shown in Supplementary Table 1) precluded the use of time- and resource-intensive sedimentation equilibrium experiments to characterize their self-association properties. Consequently, we used the higher throughput size exclusion chromatography to characterize the self-association properties of peptides by comparing their retention times on a size-exclusion column to the retention times of homotrimeric **1CW**, homodimeric **GCN4** and monomeric α -helical **PSBD36**.⁵⁻⁷

Size exclusion chromatography (SEC) was done on a Shimadzu HPLC instrument using a Phenomenex Yarra 3u sec-3000 column (batches 1 and 2) or a Zenix-C SEC 100 column (batches 3 and 4). The columns were calibrated with internal **1CW**, **GCN4**, and **PSBD36** standards. Previous characterization of **1CW**, **GCN4**, and **PSBD36** by sedimentation equilibrium analytical ultracentrifugation under analogous buffer conditions demonstrates that **1CW** adopts a trimeric association state; that **GCN4** adopts a dimeric state; and that **PSBD36** is an α -helical monomer.

The retention times derived from SEC experiments on α_3 series of peptides are very close to that of trimeric **1CW**, suggesting that these variants likewise adopt a trimeric association state. Similarly, the retention times derived from SEC experiments on α_2 series of peptides are close to that of dimeric **GCN4**, suggesting that these variants likewise adopt a dimeric association state. Variants **α_210** , **$p\alpha_21$** and **$p3\alpha_1$ -AR** have retention time between dimer and monomer or dimer and trimer simply because they are in an equilibrium of those two states.

Table S4. Retention times of helical peptides on a Zenix-C SEC 100 column.

Batch 1		
Peptide	Retention Time	Inferred association state
PSBD36(monomer standard)	12.08	
GCN4/2 α 21(dimer standard)	10.28	
1CW(trimer standard)	9.41	
2 α 28	10.38	dimer
2 α 25	10.35	dimer
2 α 18	10.18	dimer
2 α 14	10.32	dimer
2 α 10	11.55	dimer/monomer
2 α 7	10.65	dimer
2 α 6	10.66	dimer
2 α 4	10.30	dimer
2 α 3	10.26	dimer
2 α 1	10.00	dimer
p2 α 28	10.26	dimer
p2 α 25	10.31	dimer
p2 α 21	10.14	dimer
p2 α 18	10.08	dimer
p2 α 14	10.08	dimer
p2 α 10	10.18	dimer
p2 α 7	10.53	dimer
p2 α 6	10.66	dimer
p2 α 4	11.18	dimer
p2 α 3	10.36	dimer
p2 α 1	9.98	dimer/trimer
2 α 4-AA	10.83	dimer
2 α 4-KA	10.65	dimer
2 α 4-AE	10.31	dimer
p2 α 4-AA	10.66	dimer
p2 α 4-KA	10.60	dimer
p2 α 4-AE	10.19	dimer
2 α 18-AR	11.29	dimer
2 α 18-EA	10.47	dimer
2 α 18-AA	10.88	dimer
p2 α 18-AR	11.04	dimer
p2 α 18-EA	10.48	dimer
p2 α 18-AA	10.91	dimer
3 α 6	9.65	trimer
3 α 6-KA	9.74	trimer
3 α 6-AE	9.54	trimer
3 α 6-AA	9.70	trimer
p3 α 6	9.71	trimer
p3 α 6-KA	9.81	trimer
p3 α 6-AE	9.48	trimer
p3 α 6-AA	9.68	trimer
3 α 1-EA	9.66	trimer
3 α 1	9.72	trimer
p3 α 1-EA	9.74	trimer
p3 α 1	9.87	trimer

Batch 2		
Peptide	Retention Time	Inferred association state
GCN4/2α21(dimer standard)	12.05	
p2α4	11.84	dimer
3α1-AR	10.74	trimer
3α1-ER	10.08	trimer
p3α1-AR	11.21	trimer/dimer
p3α1-ER	10.37	trimer
3α1-AA	10.26	trimer
3α1-AK	10.29	trimer
p3α1-AA	10.48	trimer
p3α1-AK	10.55	trimer

Circular Dichroism Spectropolarimetry: Measurements were made with an Aviv 420 Circular Dichroism Spectropolarimeter, using quartz cuvettes with a path length of 0.1 cm. Protein solutions were prepared in 20 mM sodium phosphate buffer, pH 7, and protein concentrations were determined spectroscopically based on tyrosine and tryptophan absorbance at 280 nm in 6 M guanidine hydrochloride + 20 mM sodium phosphate ($\epsilon_{\text{Trp}} = 5690 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon_{\text{Tyr}} = 1280 \text{ M}^{-1}\text{cm}^{-1}$).⁸ CD spectra of 30 μM solutions were obtained from 260 to 200 nm at 25°C. Variable temperature CD data were obtained at least in triplicate by monitoring the molar ellipticity $[\theta]$ at 222 nm of 30 μM solutions of each α -helical variant and at 227 nm of 50 μM solutions of each β -sheet variant in 20 mM sodium phosphate (pH 7) from 1 to 95 °C at 2 °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:

$$[\theta] = (u_0 + u_1 T)(1 - F_{\text{folded}}) + (f_0 + f_1 T)(F_{\text{folded}}) \quad (\text{S1})$$

In equation S1, T is the temperature in Kelvin; u_0 and u_1 are the intercept and slope of the pre-transition baseline (which represents the linear dependence of the unfolded ensemble CD signal $[\theta]$ on temperature); f_0 and f_1 are the intercept and slope of the pre-transition baseline (which represents the linear dependence of the folded state 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; its precise form depends on whether or not the associate state of the protein changes upon folding. Folding of the GCN4-p1 variants listed in Table S1 involves association of two unfolded monomers \mathbf{M} into a folded dimer \mathbf{D} with temperature-dependent equilibrium constant \mathbf{K} as defined below:



The constant total concentration of peptide in solution \mathbf{P} is defined by the equation S3:

$$\mathbf{P} = [\mathbf{M}] + 2[\mathbf{D}] = [\mathbf{M}] + 2\mathbf{K}[\mathbf{M}]^2 \quad (\text{S3})$$

The positive root of this quadratic equation provides an expression for $[\mathbf{M}]$ as a function of \mathbf{P} and \mathbf{K} :

$$[\mathbf{M}] = \frac{\sqrt{1+8\mathbf{K}\mathbf{P}}-1}{4\mathbf{K}} \quad (\text{S4})$$

Substitution of $[\mathbf{M}]$ into the definition of F_{folded} gives the following expression for the monomer-dimer equilibrium:

$$F_{\text{folded}} = \frac{2\mathbf{K}[\mathbf{M}]^2}{\mathbf{P}} = 1 + \frac{1}{4\mathbf{K}\mathbf{P}} - \sqrt{\frac{1}{16\mathbf{K}^2\mathbf{P}^2} - \frac{1}{2\mathbf{K}\mathbf{P}}} \quad (\text{S5})$$

Folding of the 1CW variants listed in Table S2 involves association of three unfolded monomers \mathbf{M} into a folded trimer \mathbf{Tri} with temperature-dependent equilibrium constant \mathbf{K} as defined below:



The constant total concentration of peptide in solution \mathbf{P} is defined by equation S7:

$$\mathbf{P} = [\mathbf{M}] + 3[\mathbf{D}] = [\mathbf{M}] + 3\mathbf{K}[\mathbf{M}]^3 \quad (\text{S7})$$

Rearranging equation S7 results in the following polynomial equation that is cubic in $[\mathbf{M}]$:

$$0 = [\mathbf{M}]^3 + \frac{[\mathbf{M}]}{3\mathbf{K}} - \frac{\mathbf{P}}{3\mathbf{K}} \quad (\text{S8})$$

Using Mathematica, we found the three roots of this polynomial, two of which are complex, whereas the third is real. The real root of equation S8 provides an expression for $[\mathbf{M}]$ as a function of \mathbf{P} and \mathbf{K} :

$$[\mathbf{M}] = \left(\frac{\mathbf{P}}{6\mathbf{K}} + \left(\frac{1}{729\mathbf{K}^3} + \frac{\mathbf{P}^2}{36\mathbf{K}^2} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}} - \frac{1}{9\mathbf{K} \left(\frac{\mathbf{P}}{6\mathbf{K}} + \left(\frac{1}{729\mathbf{K}^3} + \frac{\mathbf{P}^2}{36\mathbf{K}^2} \right)^{\frac{1}{2}} \right)^{\frac{1}{3}}} \quad (\text{S9})$$

Substitution of $[\mathbf{M}]$ into the definition of F_{folded} gives the following expression for the monomer-trimer equilibrium:

$$F_{\text{folded}} = \frac{3\mathbf{K}[\mathbf{M}]^3}{\mathbf{P}} \quad (\text{S10})$$

Folding of the WW variants listed in Table S3 involves an equilibrium between an unfolded monomer (**U**) and a folded monomer (**F**) temperature-dependent equilibrium constant **K** as defined below:

$$\mathbf{U} \rightleftharpoons \mathbf{F}; \quad \mathbf{K} = \frac{[\mathbf{F}]}{[\mathbf{U}]} \quad (\text{S11})$$

The constant total concentration of peptide in solution **P** is defined by equation S11:

$$\mathbf{P} = [\mathbf{U}] + [\mathbf{F}] = [\mathbf{U}] + \mathbf{K}[\mathbf{U}] \quad (\text{S12})$$

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

$$F_{\text{folded}} = \frac{\mathbf{K}}{1+\mathbf{K}} \quad (\text{S13})$$

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

$$\mathbf{K} = e^{-\frac{\Delta G_f}{RT}} \quad (\text{S14})$$

In turn, the temperature-dependence of ΔG_f for the dimer or trimer can be expressed as a second order polynomial:

$$\Delta G_f = \Delta G_0 + \Delta G_1(T - T_0) + \Delta G_2(T - T_0)^2 \quad (\text{S15})$$

where ΔG_0 , ΔG_1 and ΔG_2 are parameters to be determined via least-squares regression (though ΔG_2 is excluded when its standard error is too high), and T_0 is an arbitrary reference temperature that should be close to the melting temperature. We used this expression for the GCN4-p1 and 1CW variants, though many of the 1CW variants did not require the use of ΔG_2 . For the WW variants, we used the following expression for the temperature-dependence of ΔG_f , which is a function of folding enthalpy at the melting temperature $\Delta H_{(T_m)}$, folding heat capacity ΔC_p , and the melting temperature T_m (i.e., the temperature at which $F_{\text{folded}} = 0.5$ and $\Delta G_f = 0$ kcal/mol):

$$\Delta G_f = \frac{\Delta H_{(T_m)} \cdot (T_m - T)}{T_m} + \Delta C_p \cdot \left(T - T_m - T \cdot \ln \left[\frac{T}{T_m} \right] \right) \quad (\text{S16})$$

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 S137-S200, 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). For the monomer-dimer and monomer-trimer equilibria, the melting temperature T_m (defined as the temperature at which $F_{\text{folded}} = 0.5$) is not a parameter of the fit. In these cases, we used Mathematica to solve for T_m numerically.

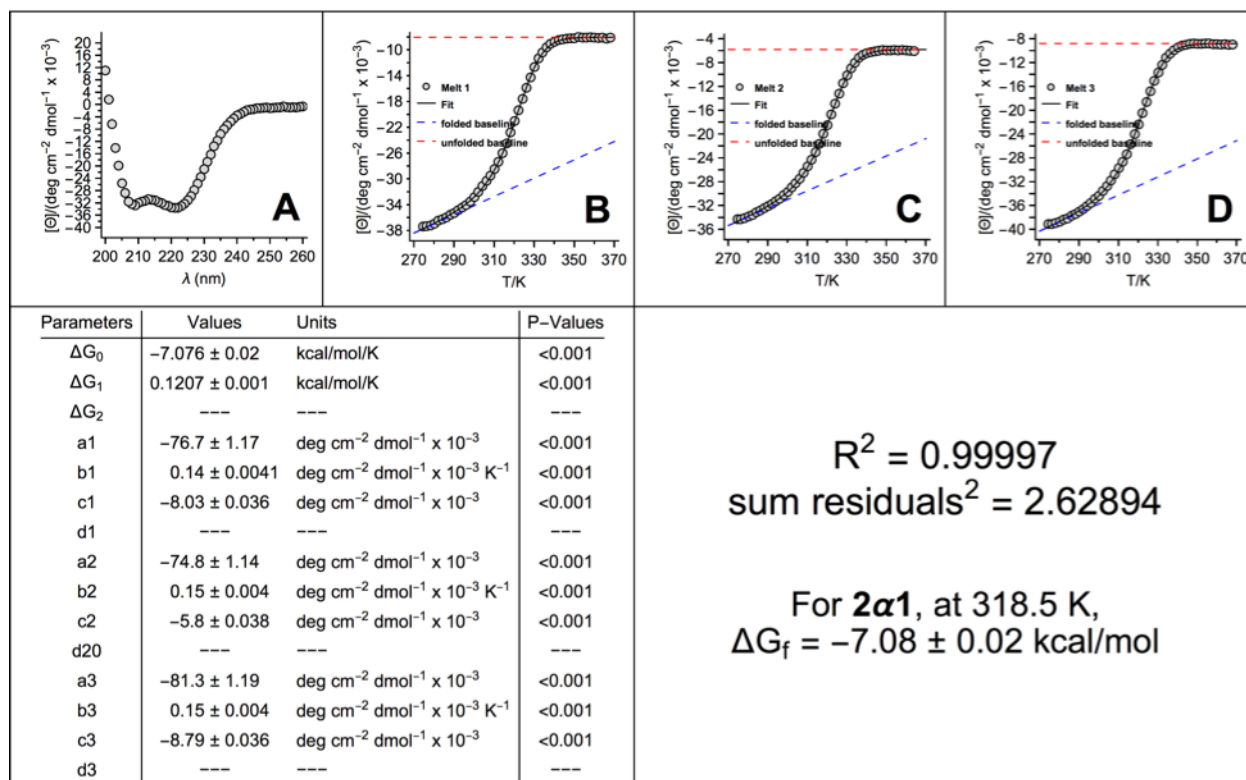


Figure S137. CD data spectra for **2α1** (QX108111).

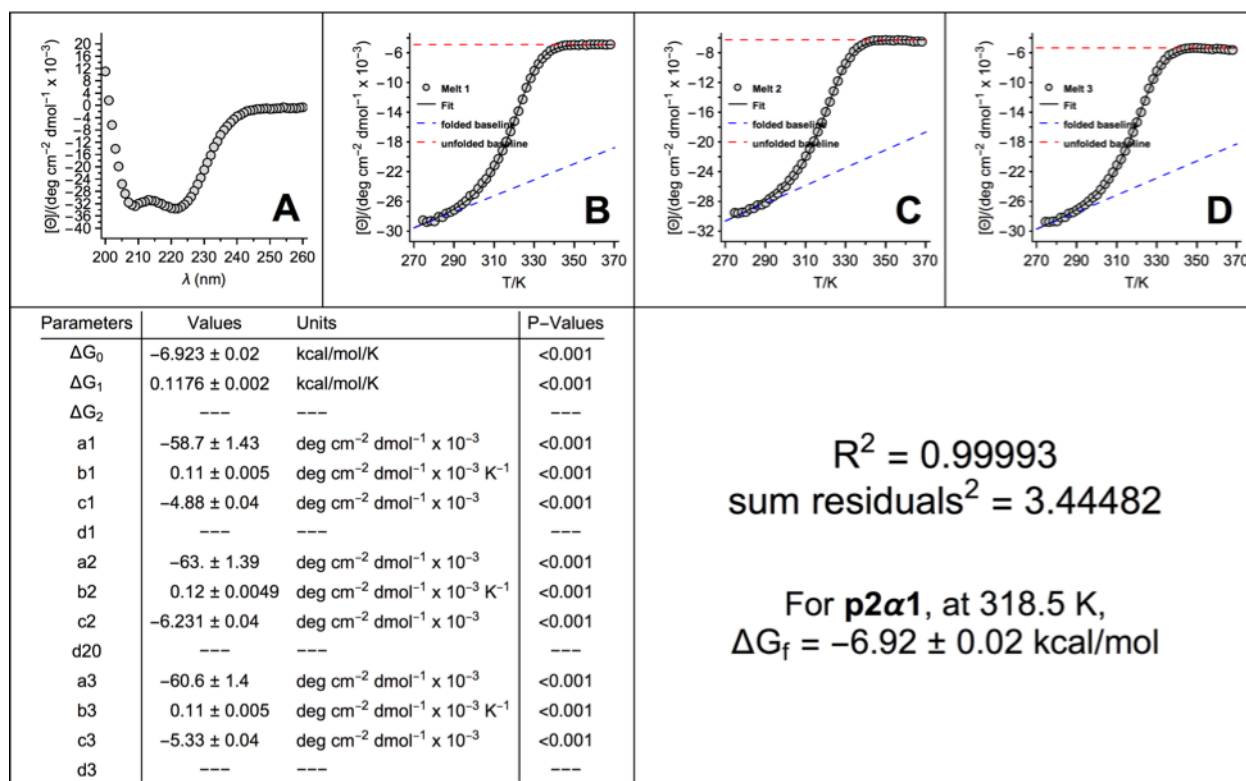


Figure S138. CD data for **p2α1** (QX108111p).

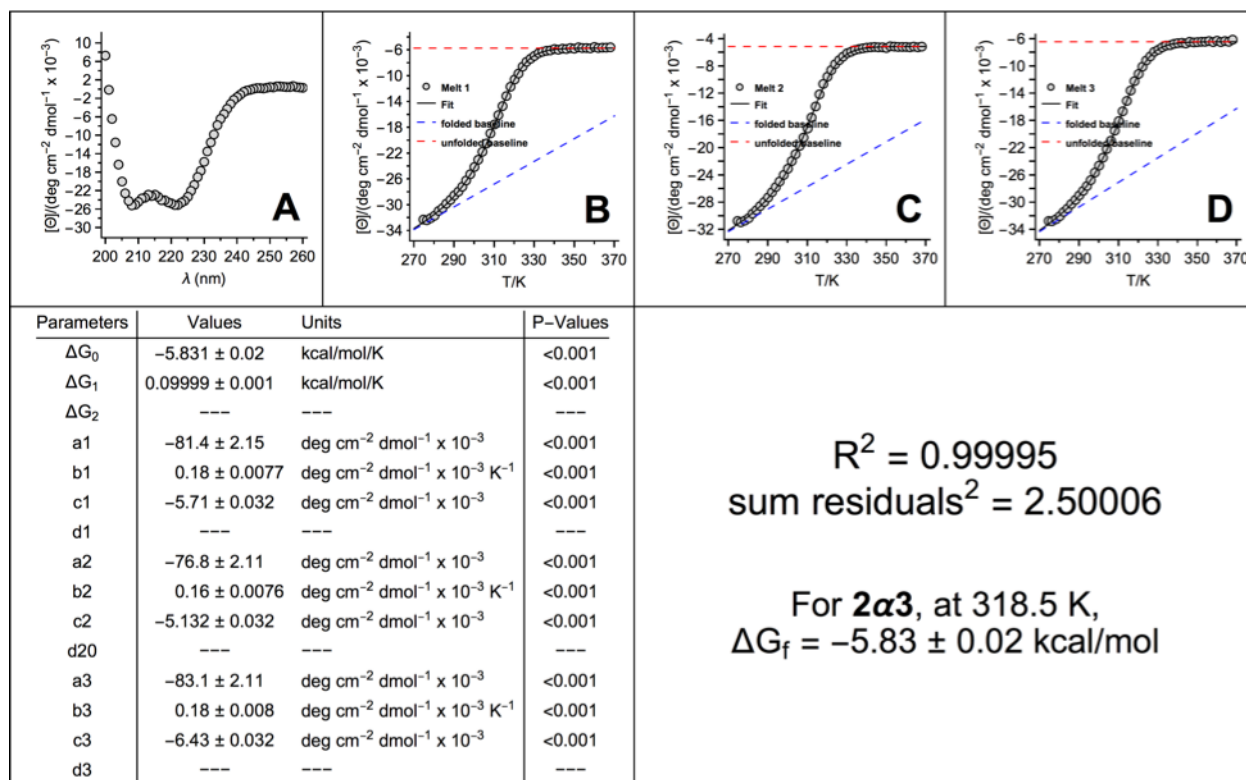


Figure S139. CD data for **2α3** (QX108110).

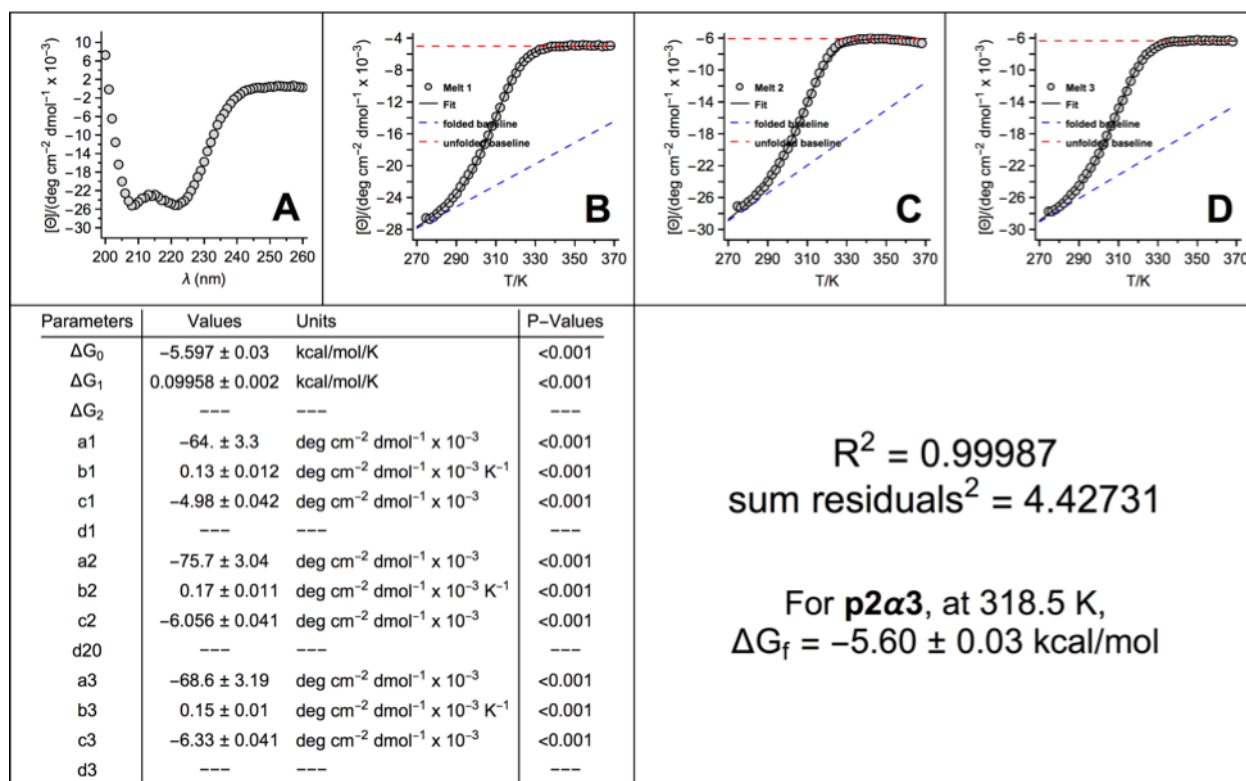


Figure S140. CD data for **p2α3** (QX108110p).

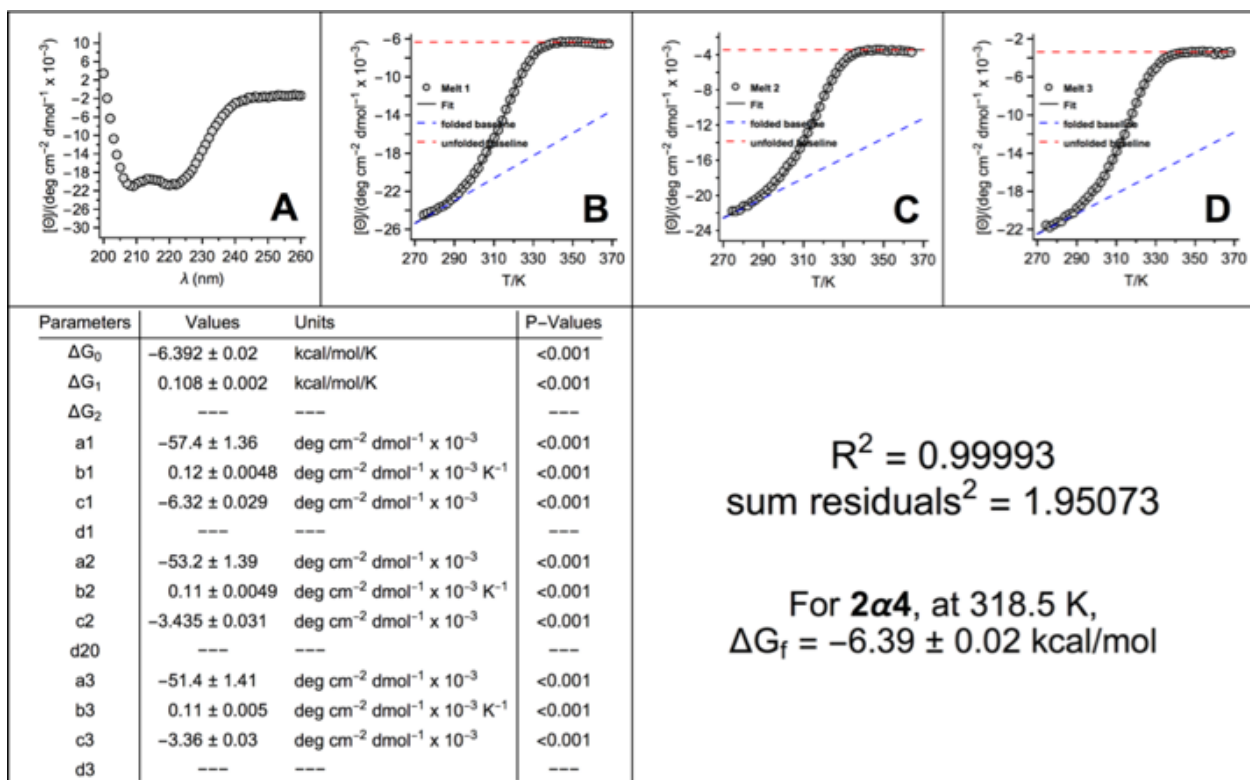


Figure S141. CD data for **2α4** (QX10819).

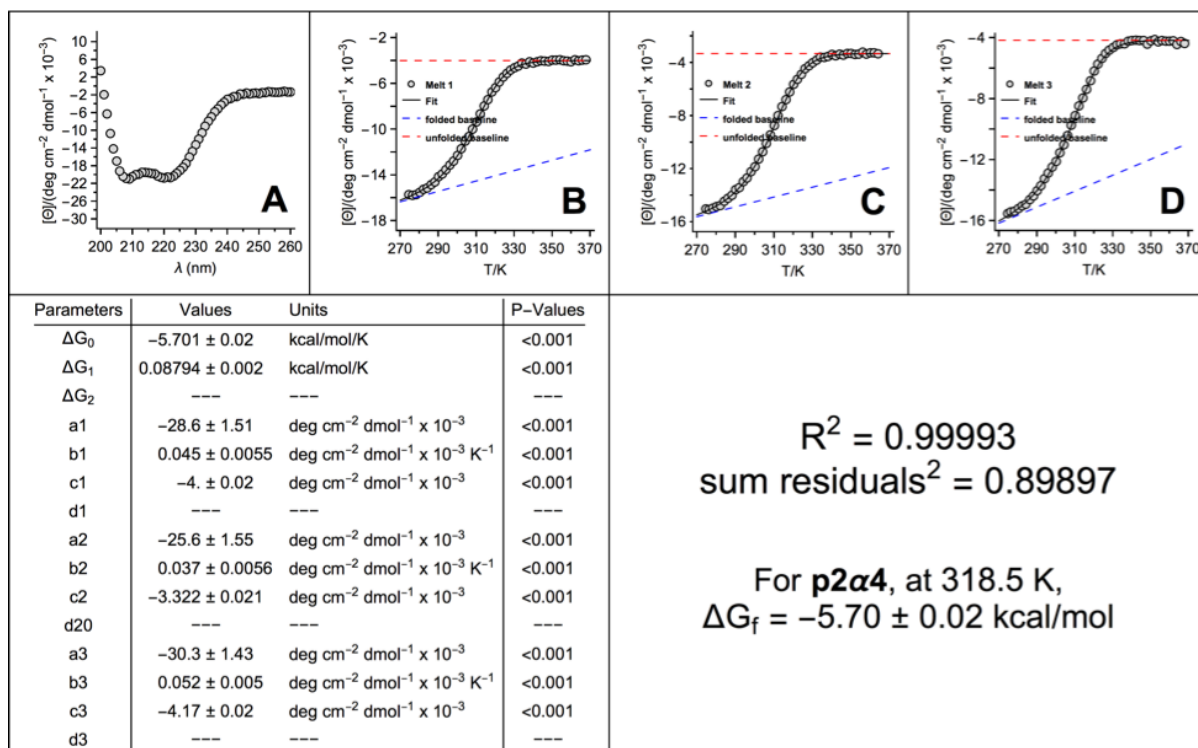


Figure S142. CD data for **p2α4** (QX10819p).

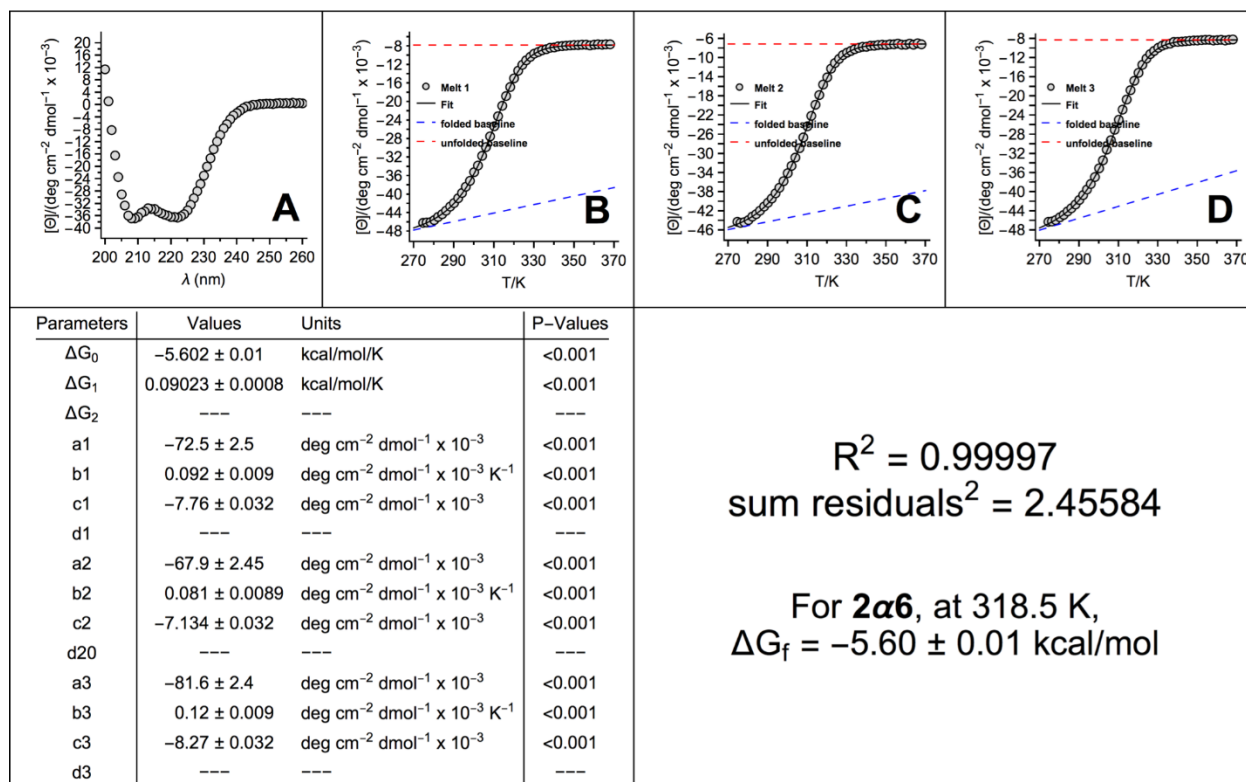


Figure S143. CD data for **2α6** (QX10818).

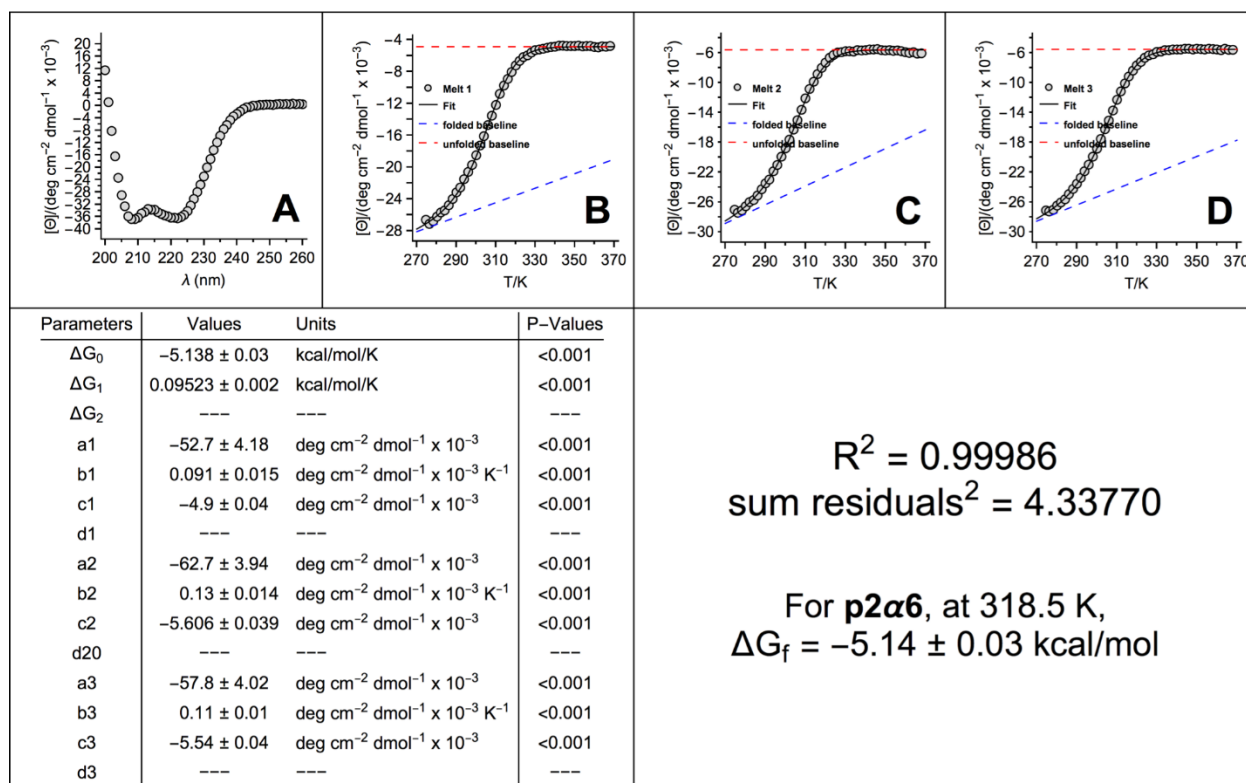


Figure S144. CD data for **p2α6** (QX10818p).

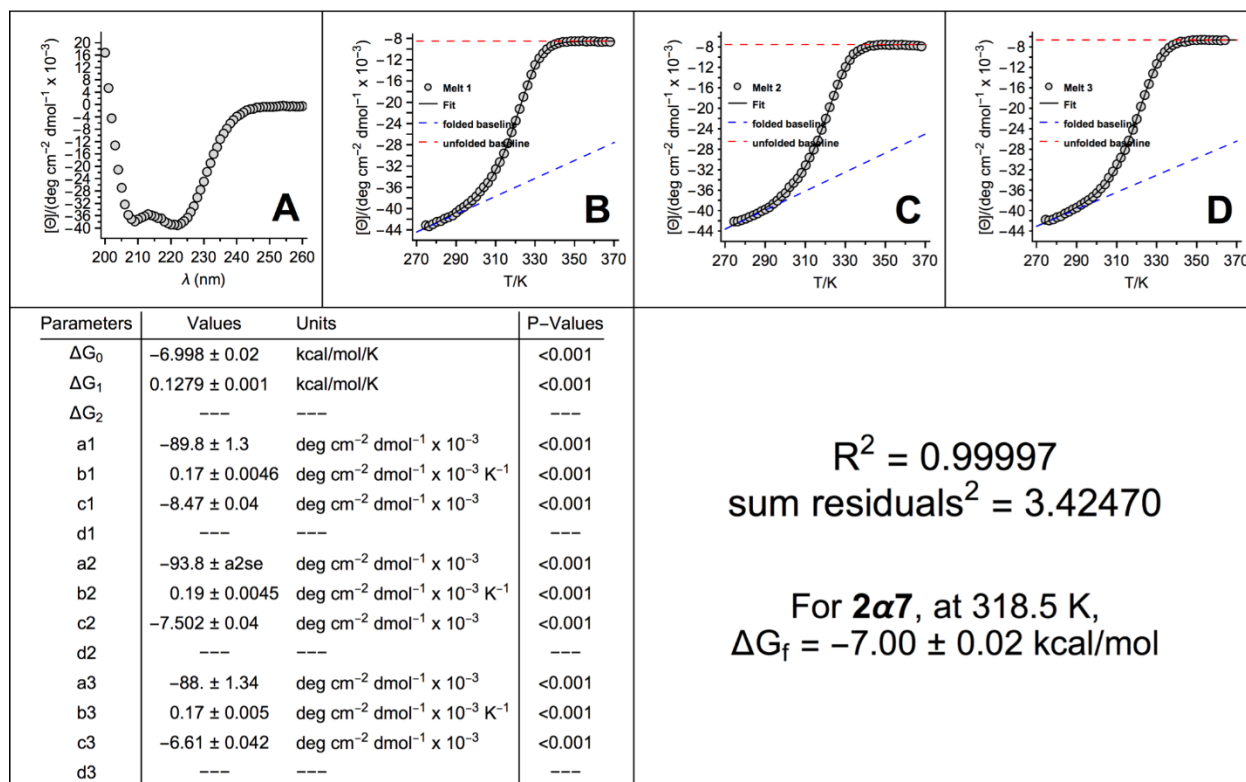


Figure S145. CD data for **2α7** (QX10817).

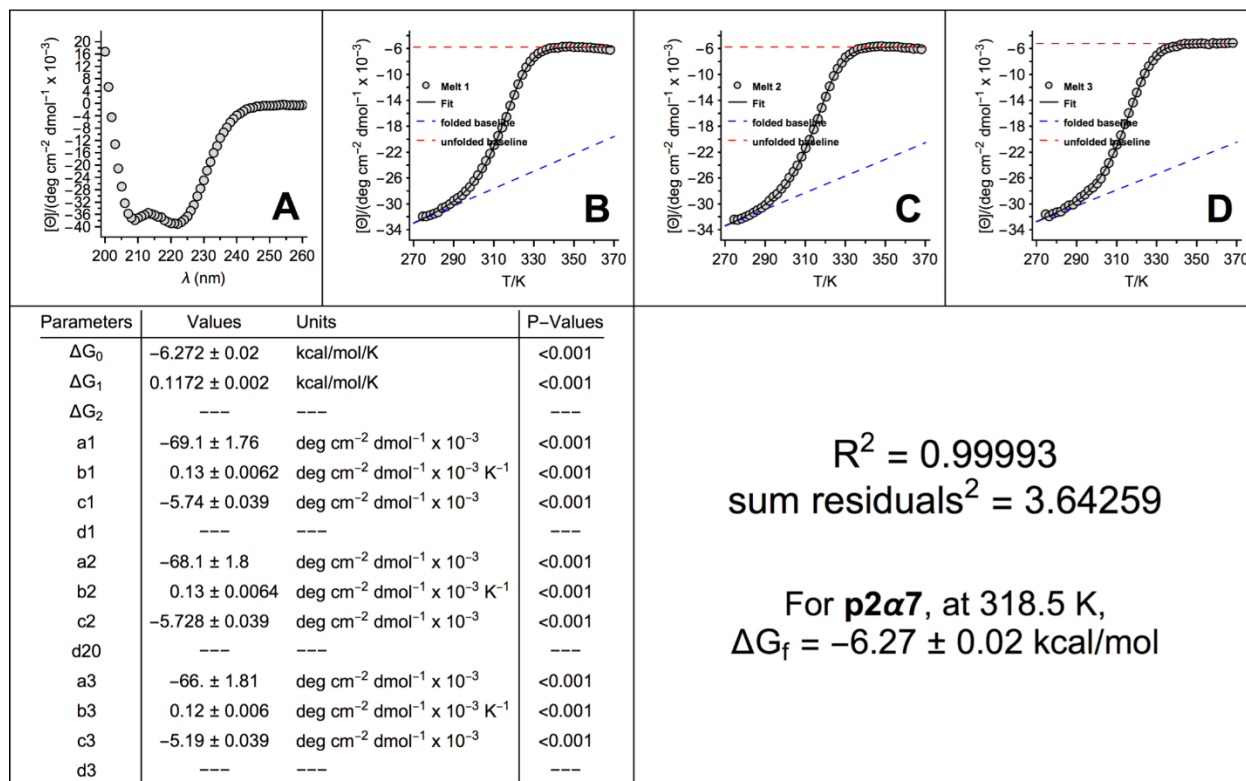


Figure S146. CD data for **p2α7** (QX10817p).

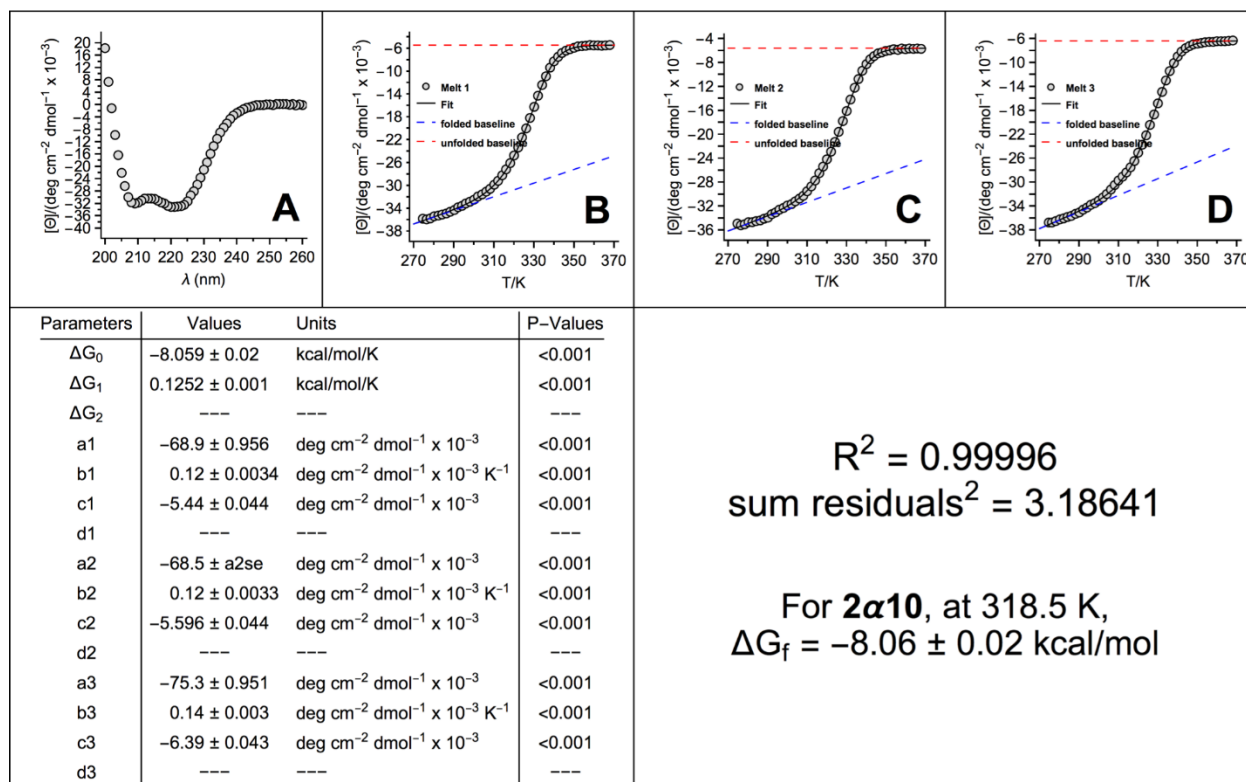


Figure S147. CD data for **2 α 10** (QX10816).

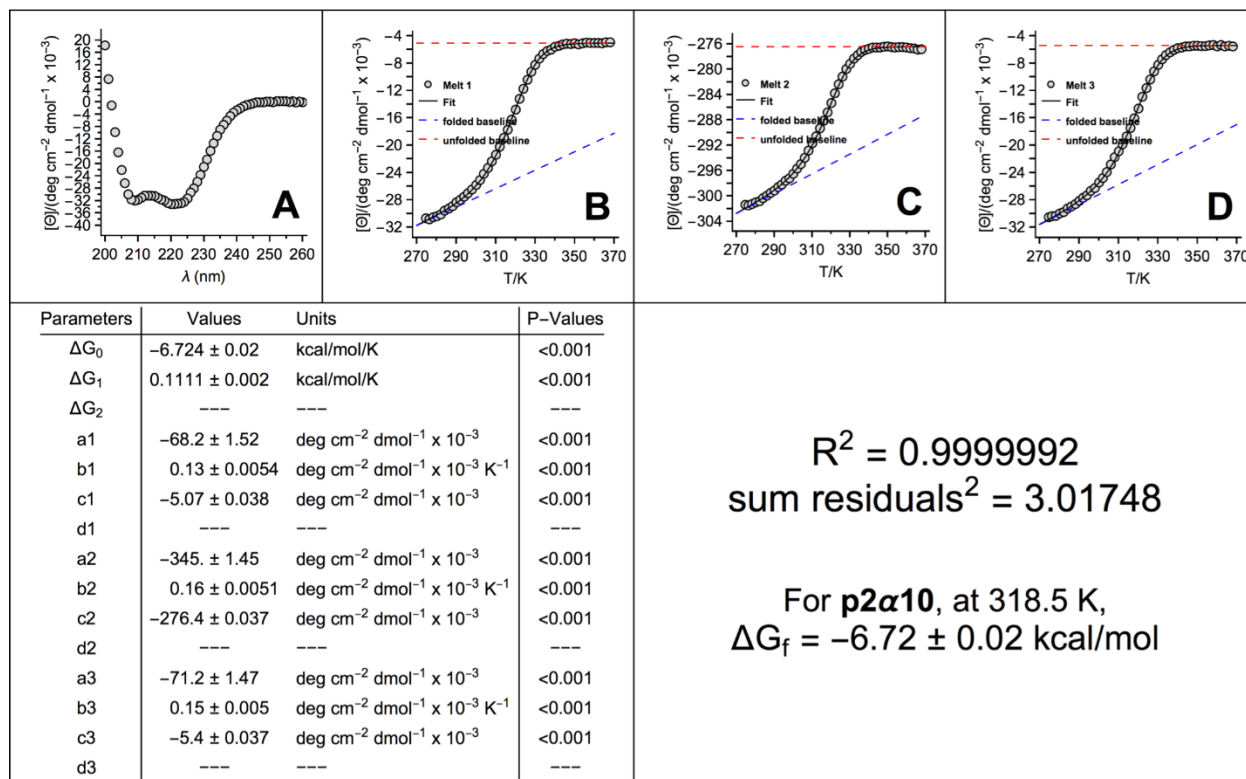


Figure S148. CD data for **p2 α 10** (QX10816p).

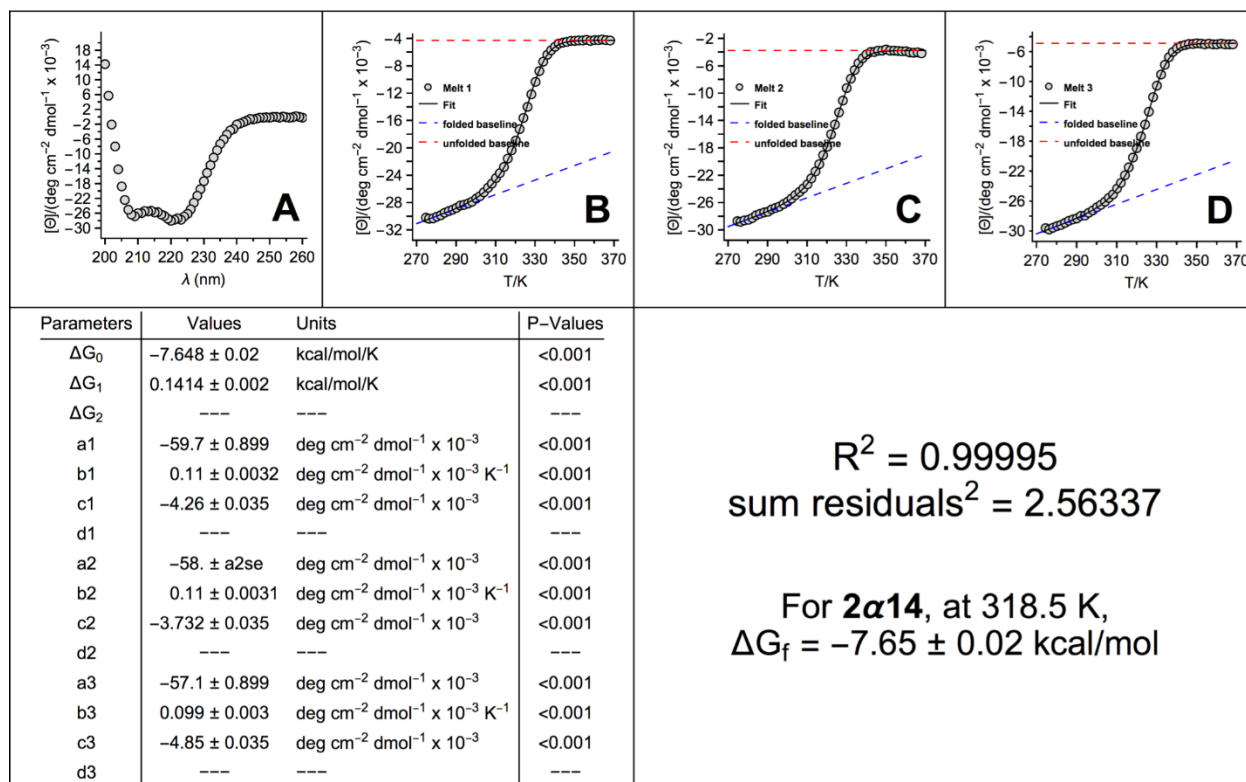


Figure S149. CD data for **2α14** (QX10815).

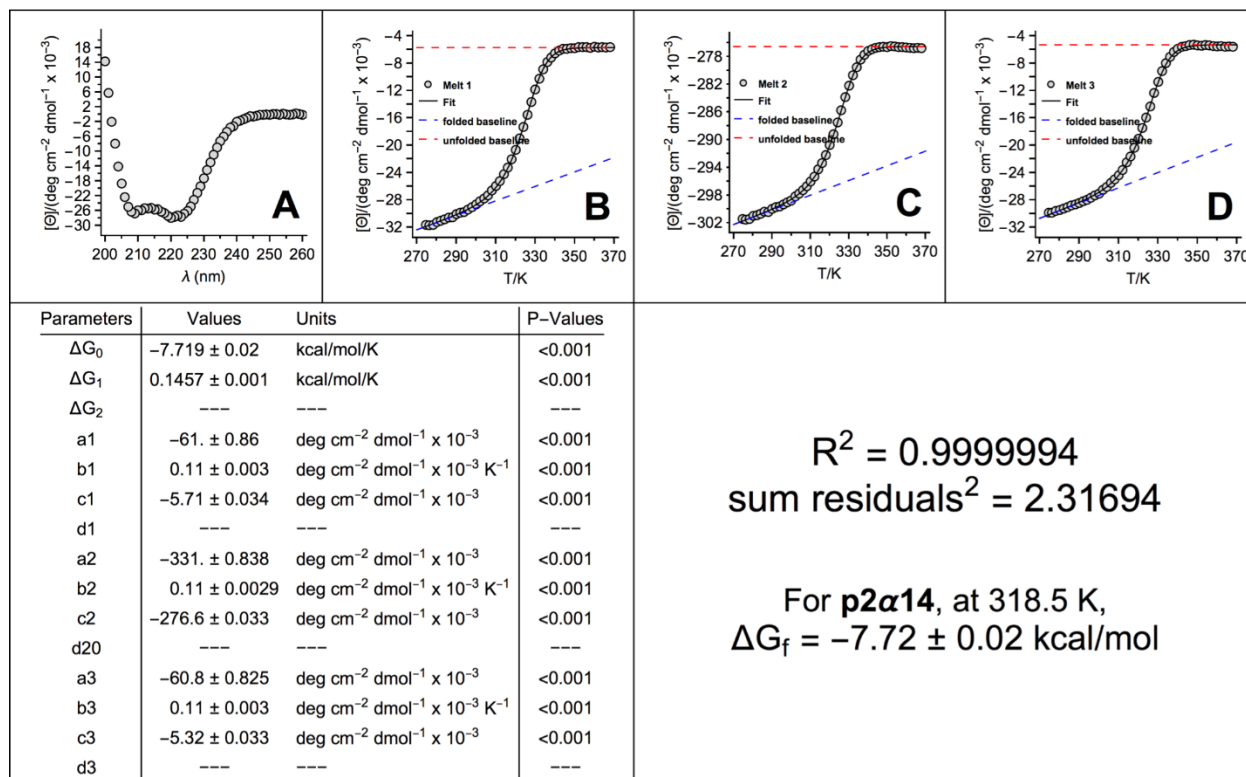


Figure S150. CD data for **p2α14** (QX10815p).

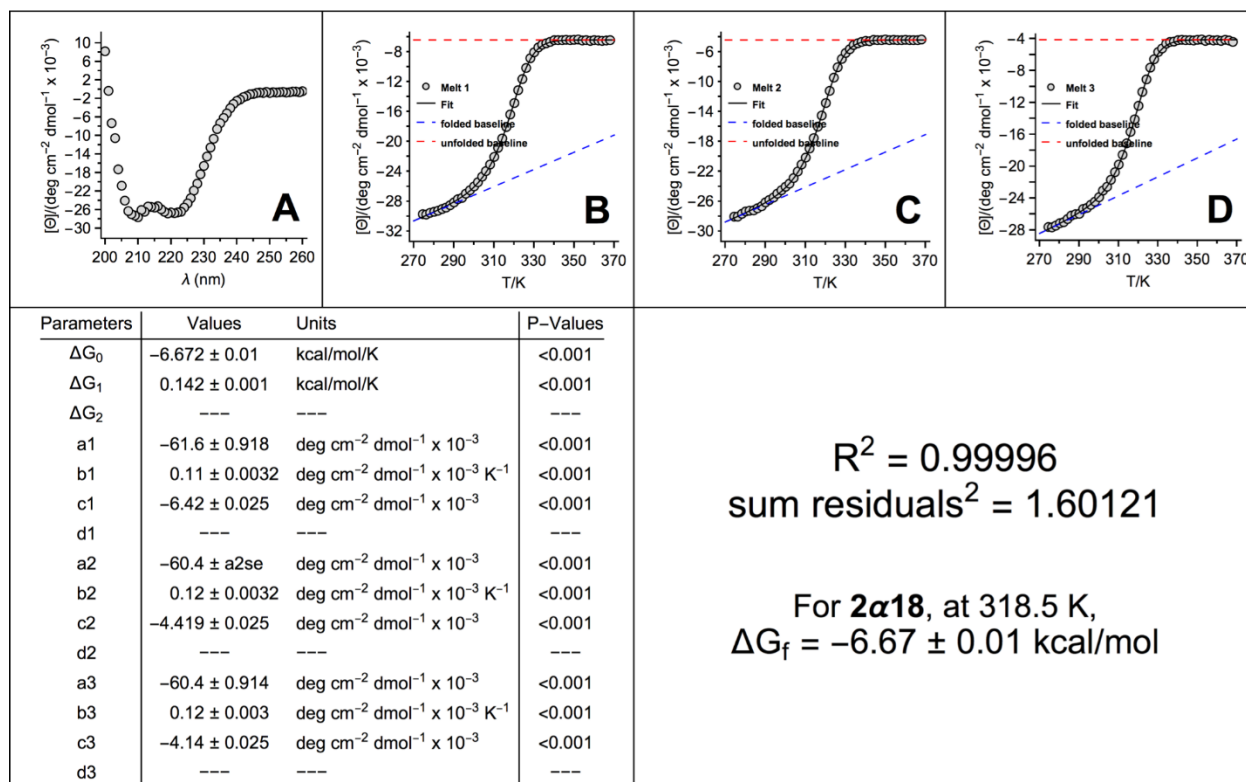


Figure S151. CD data for **2α18** (QX10814).

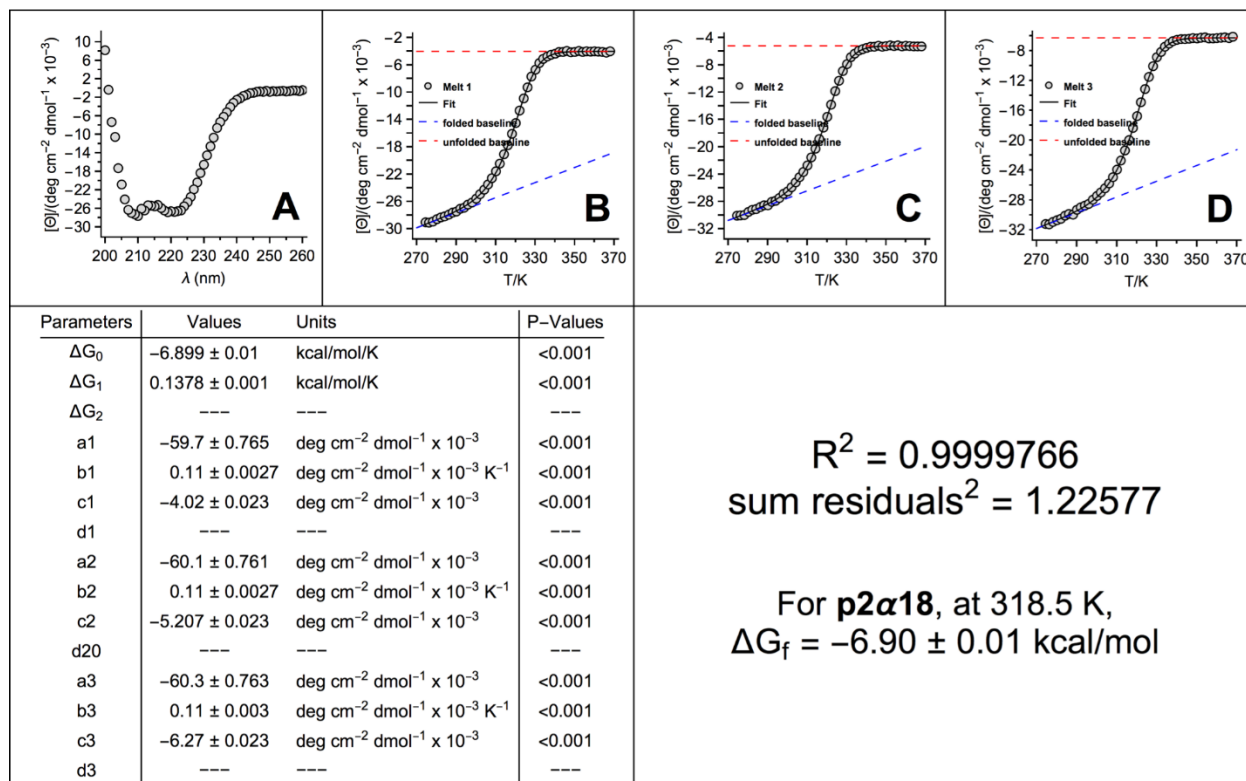


Figure S152. CD data for **p2α18** (QX10814p).

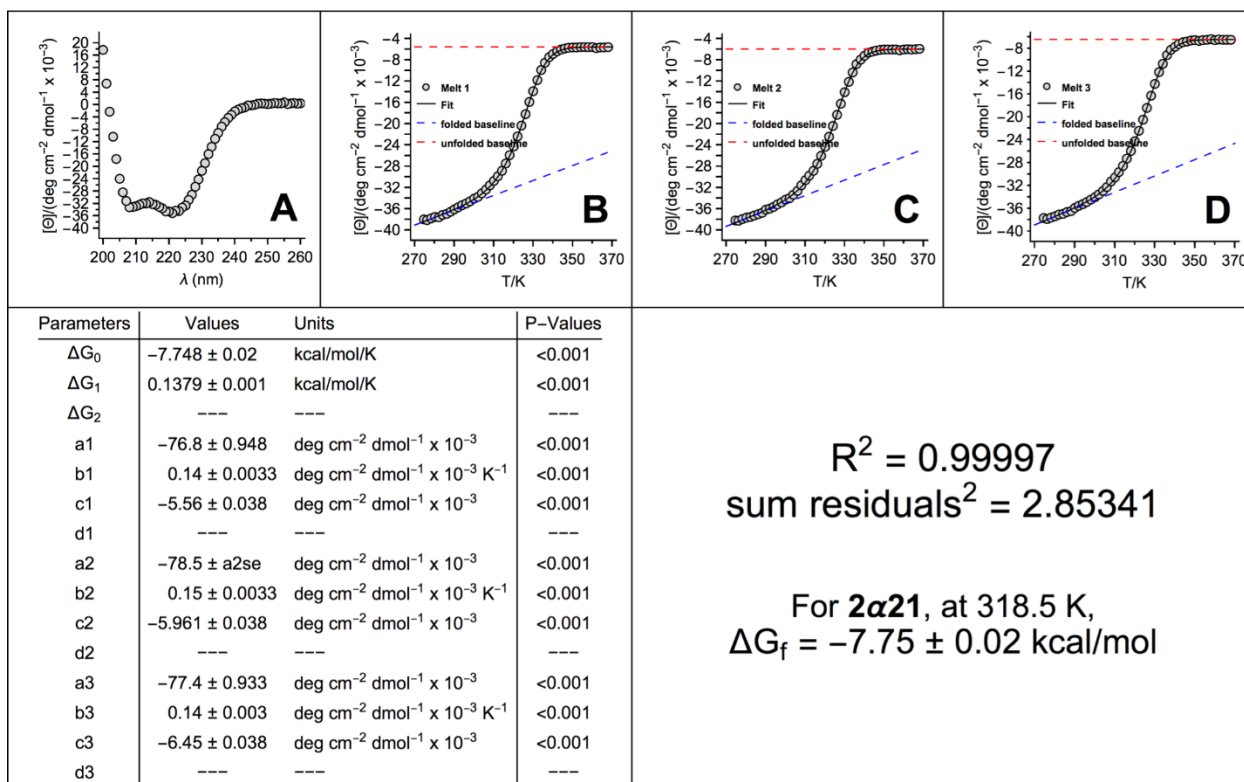


Figure S153. CD data for **2α21** (QX10813).

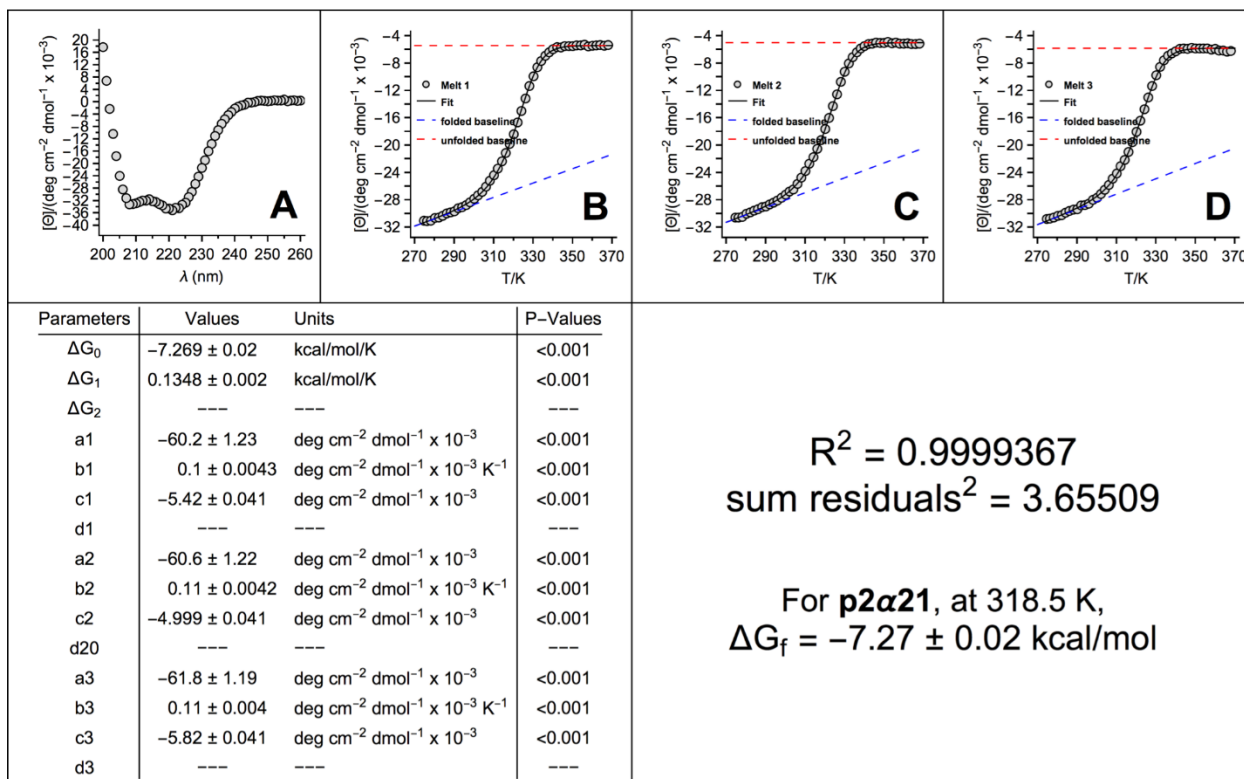


Figure S154. CD data for **p2α21** (QX10813p).

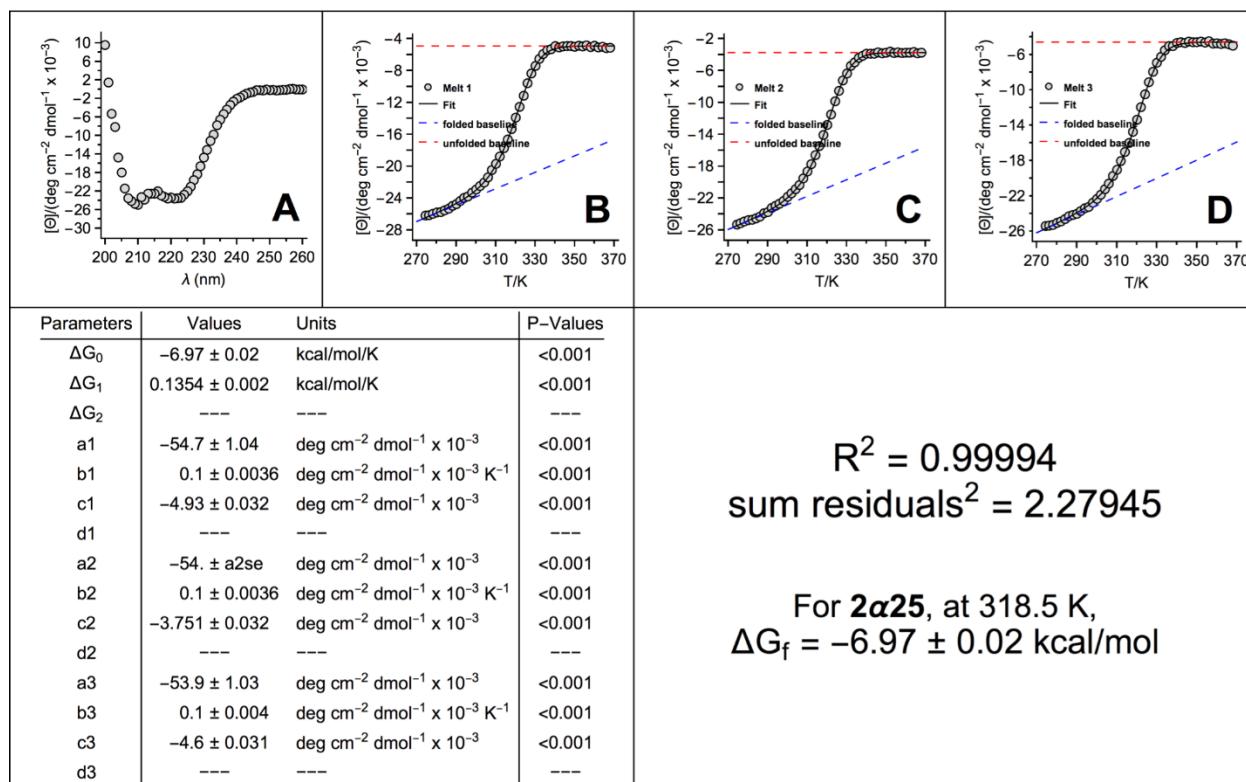


Figure S155. Cd data for **2α25** (QX10812).

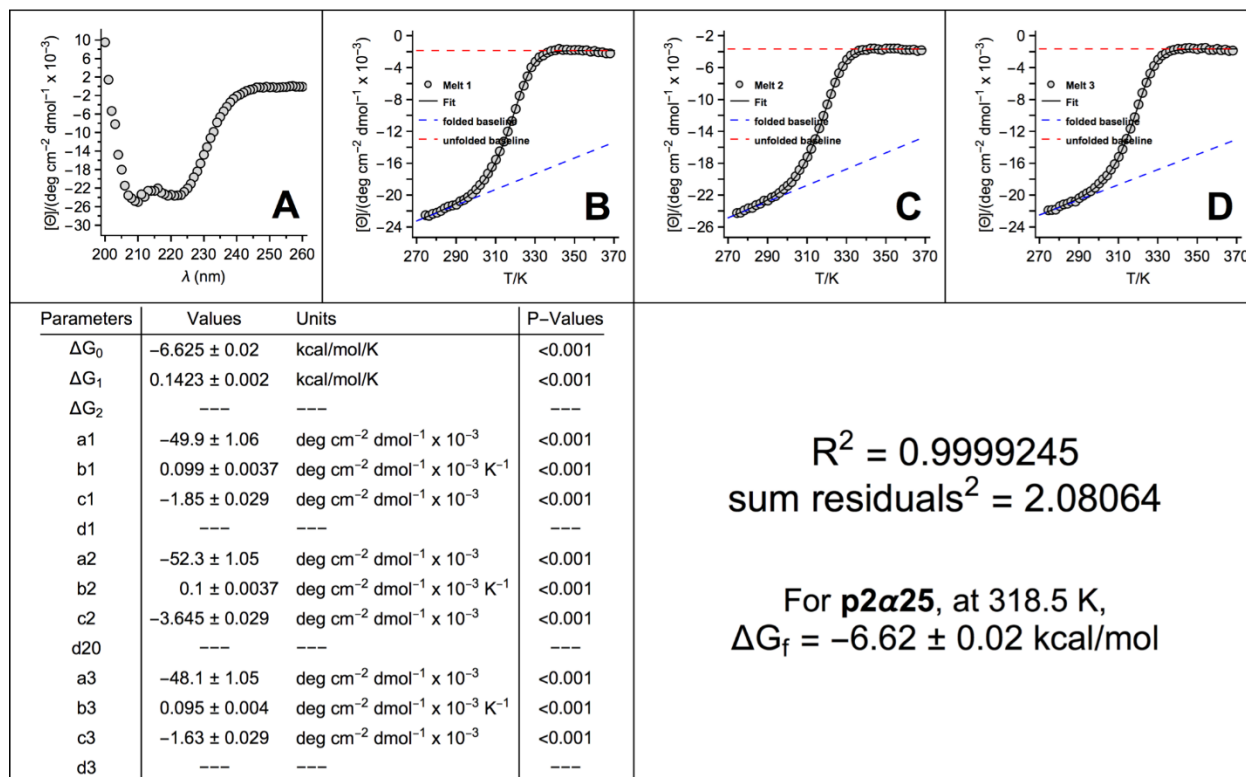


Figure S156. CD data for **p2α25** (QX10812p).

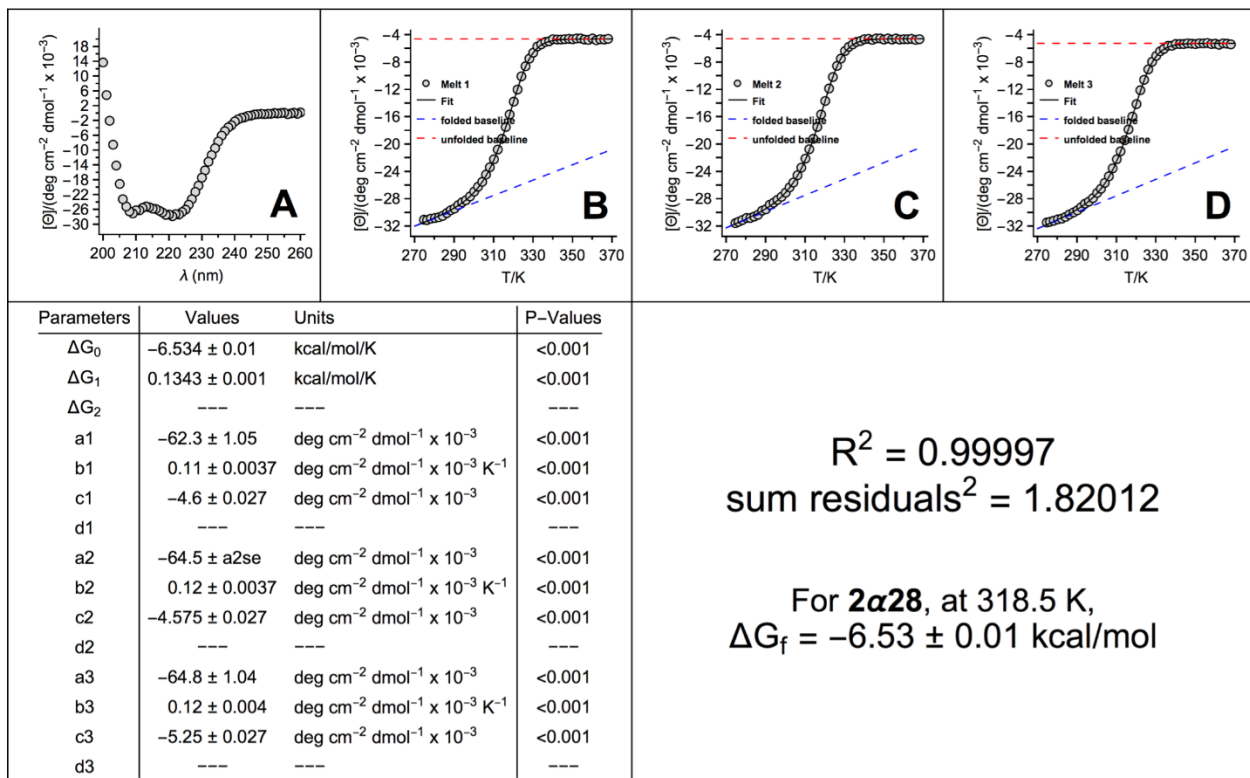


Figure S157. CD data for **2α28** (QX10811).

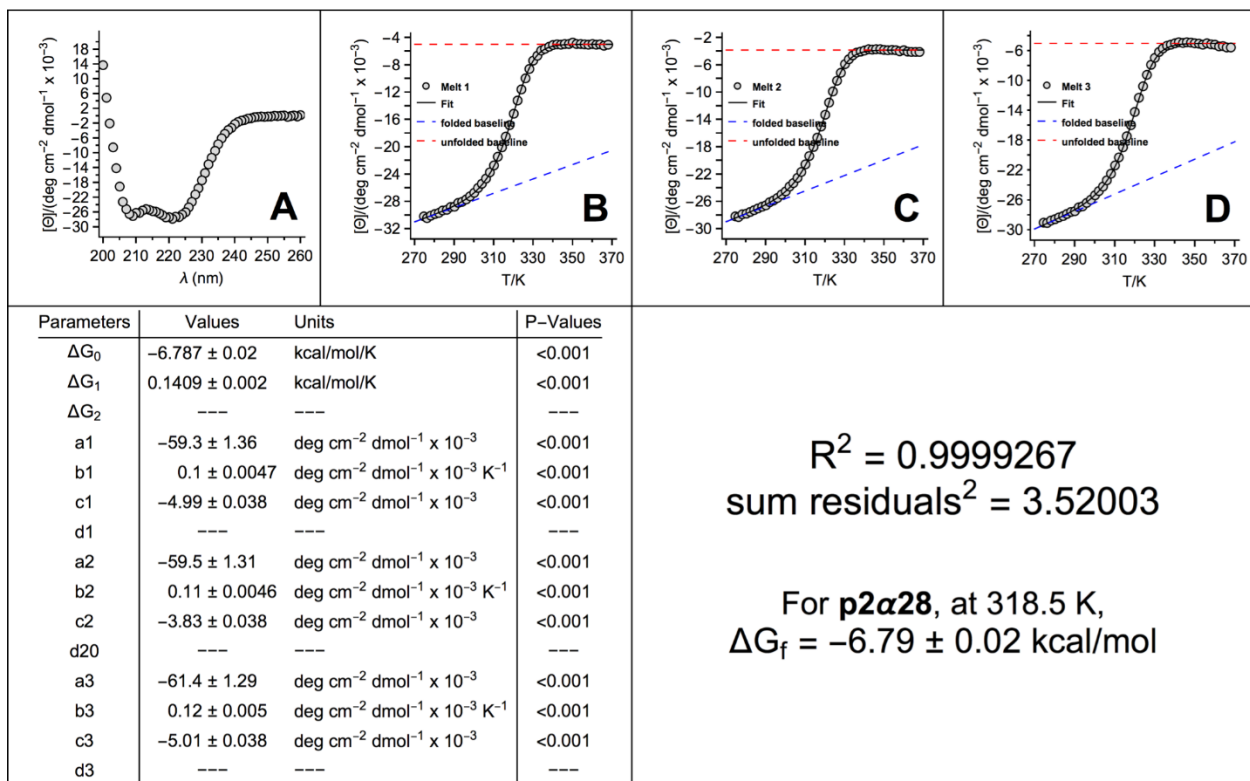


Figure S158. CD data for **p2α28** (QX10811p).

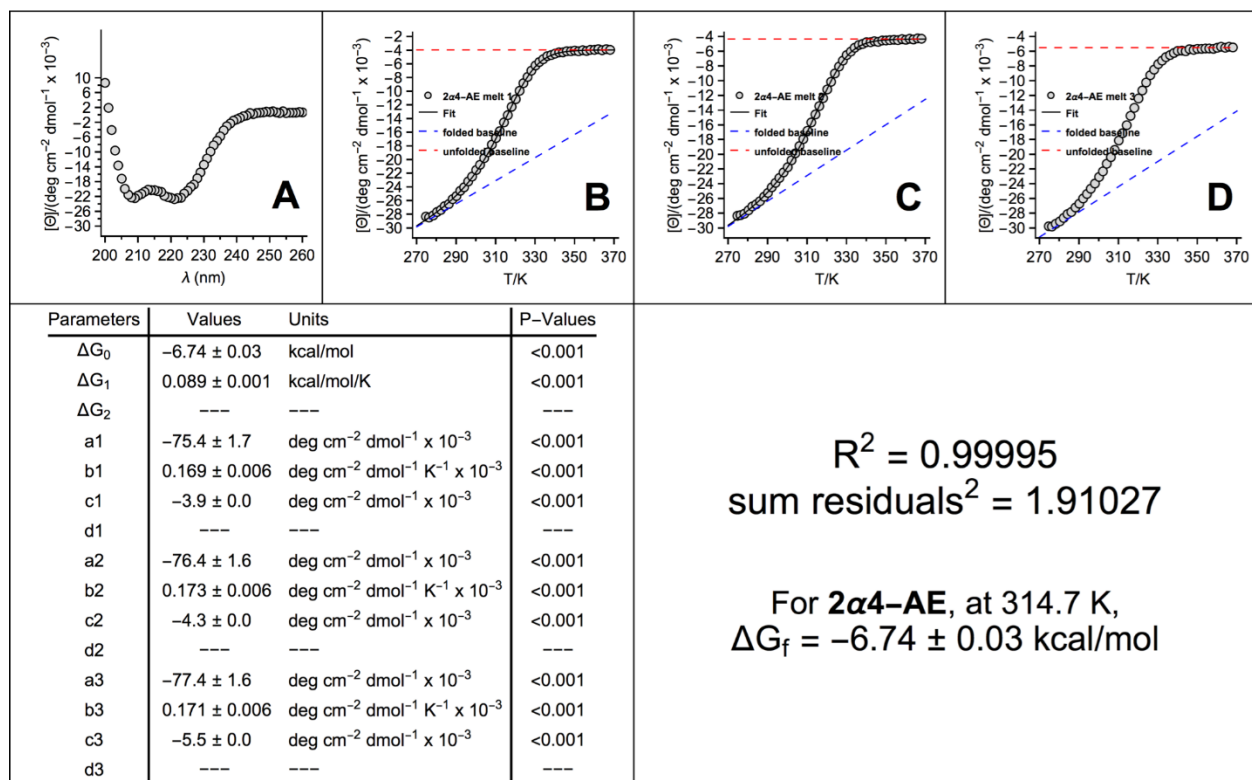


Figure S159. CD data for **2 α 4-AE** (QX11193).

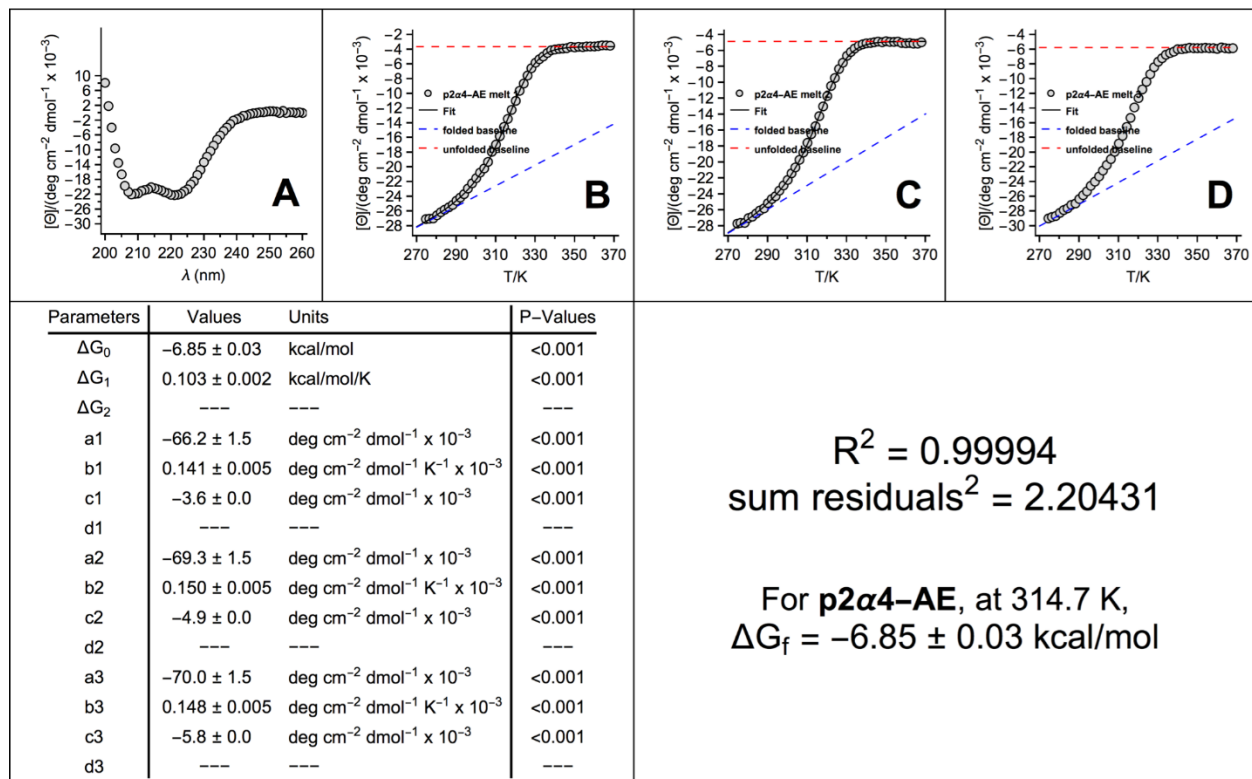


Figure S160. CD data for **p2 α 4-AE** (QX11196).

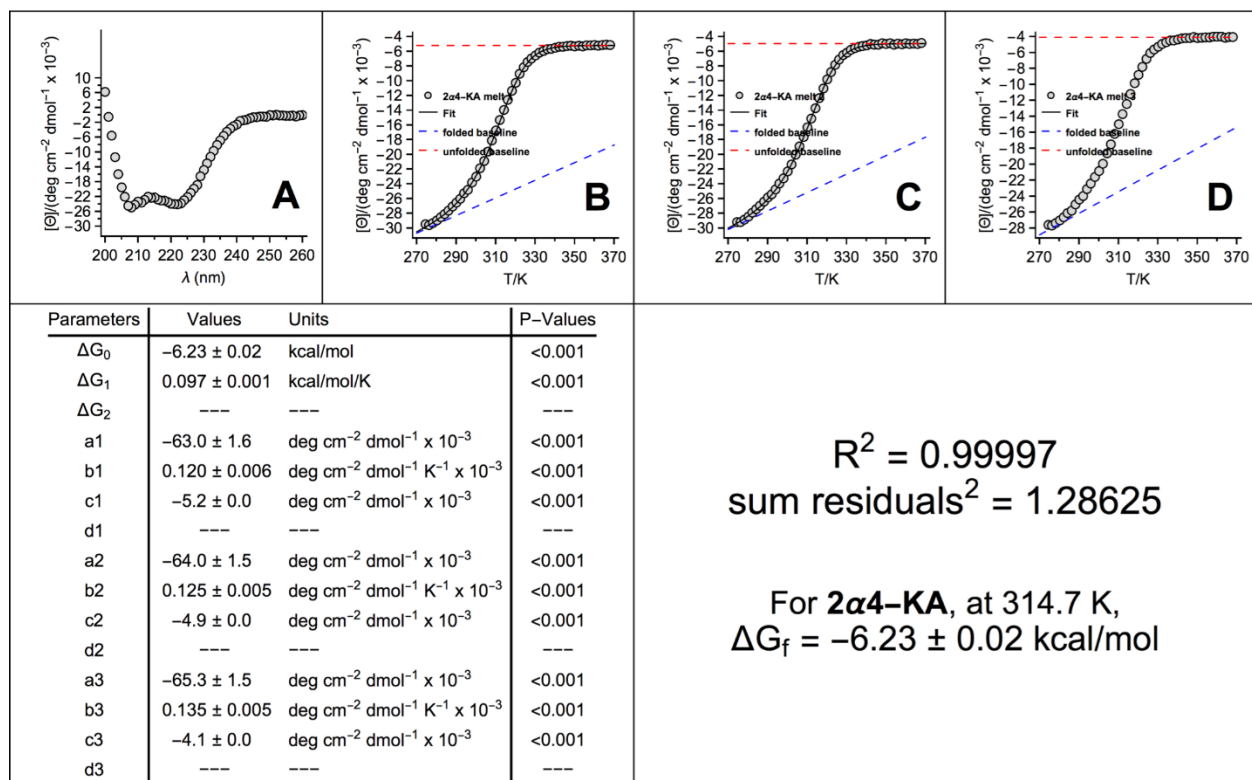


Figure S161. CD data for **2α4-KA** (QX11192).

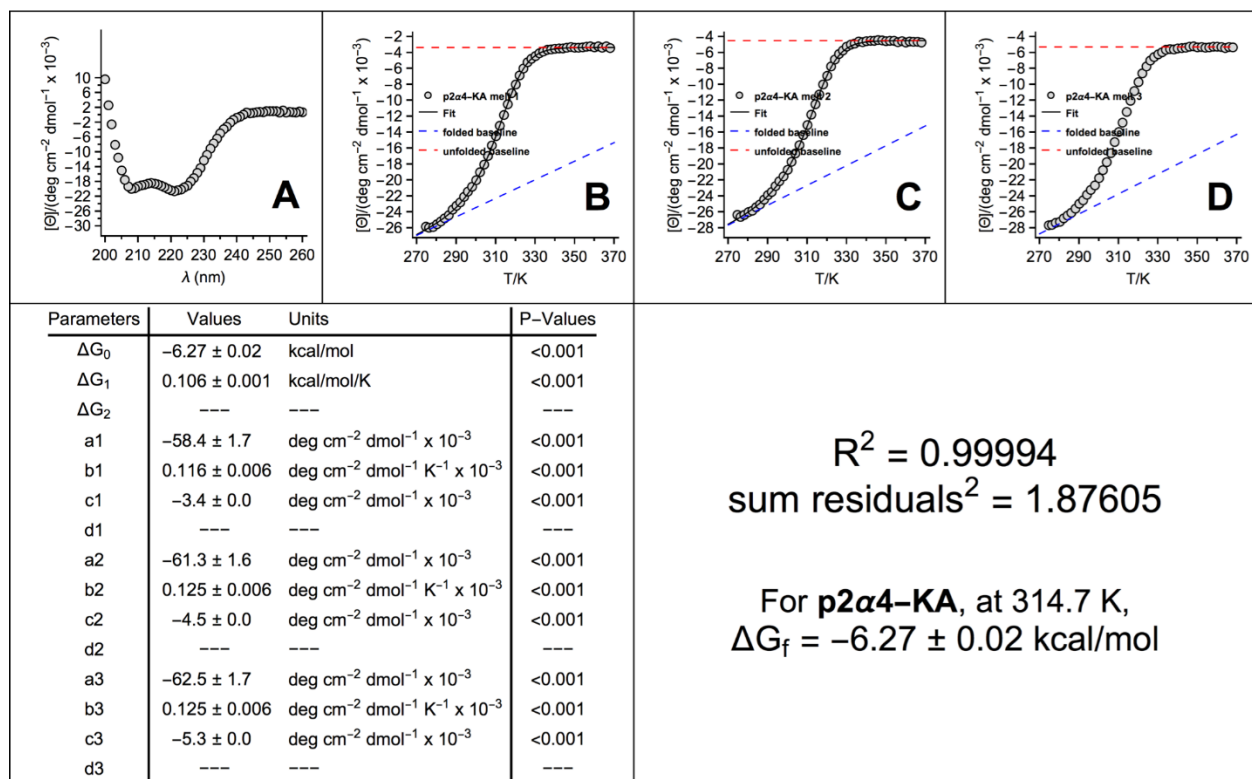


Figure S162. CD data for **p2α4-KA** (QX11195).

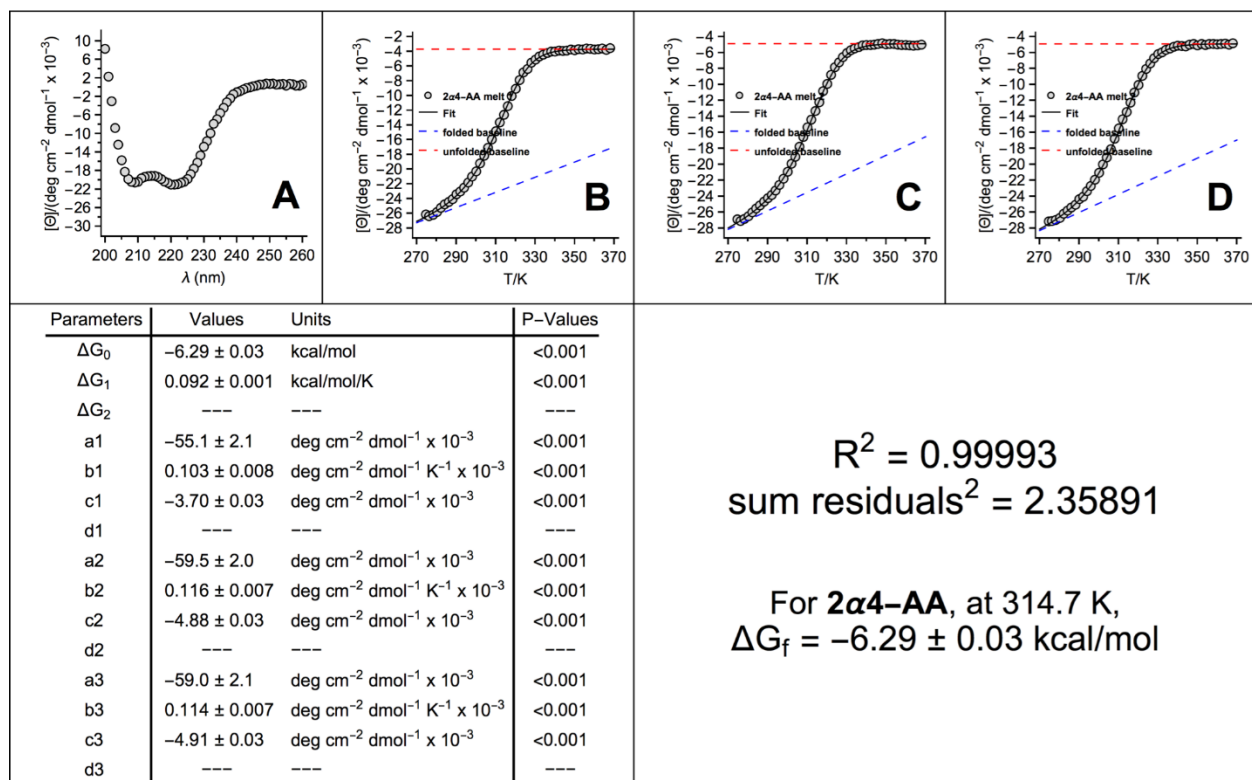


Figure S163. CD data for **2 α 4-AA** (QX11191).

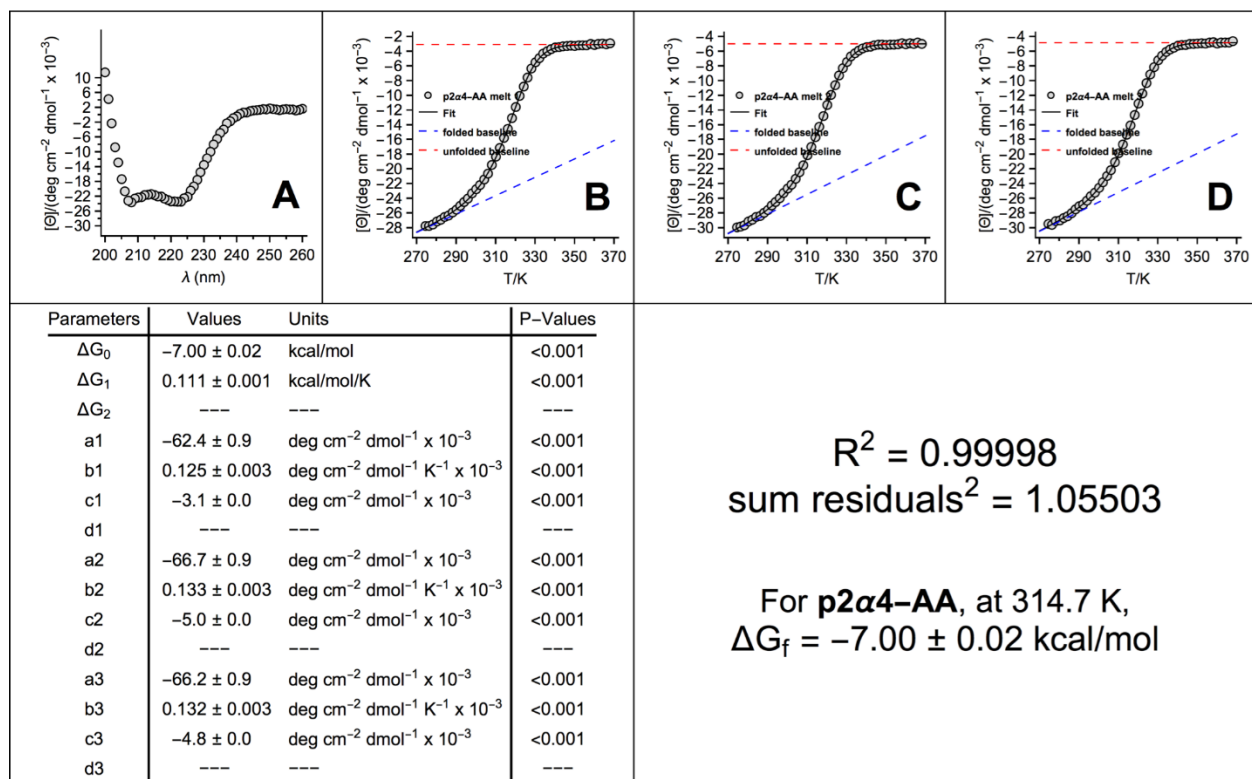


Figure S164. CD data **p2 α 4-AA** (QX11194).

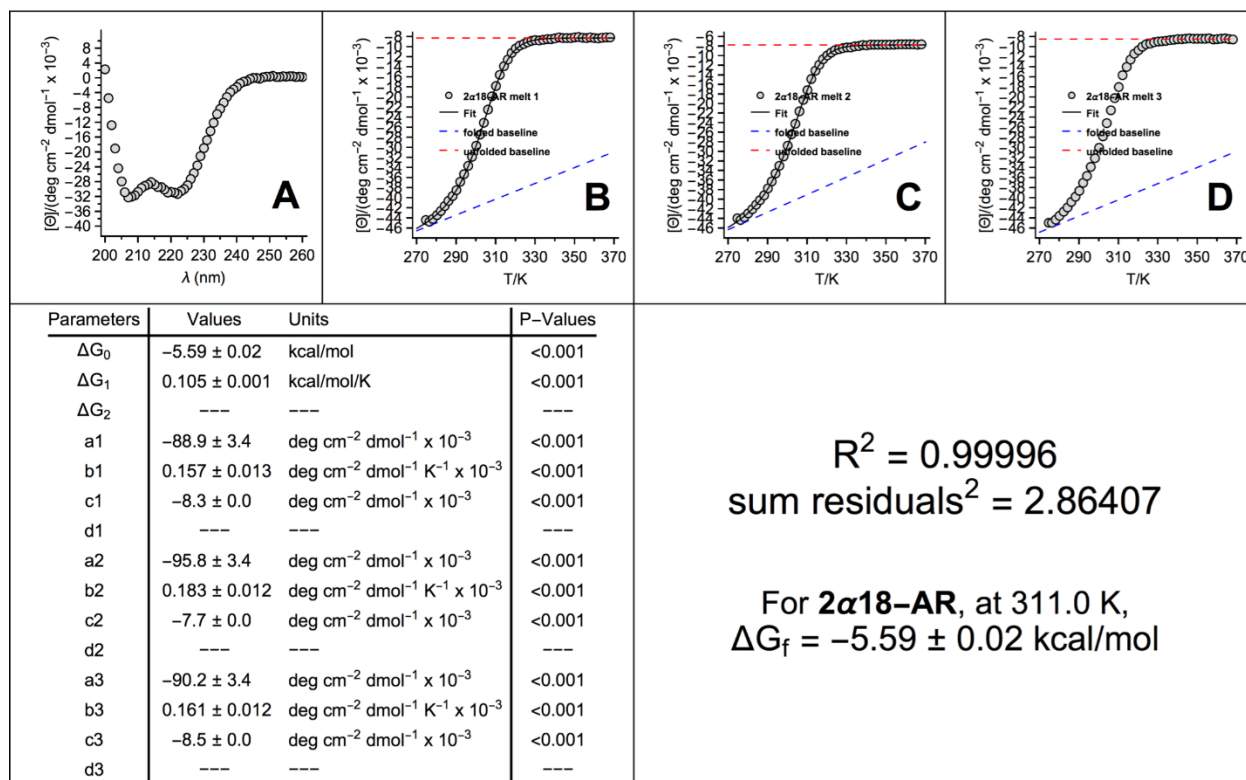


Figure S165. CD data for **2 α 18-AR** (NAB10211).

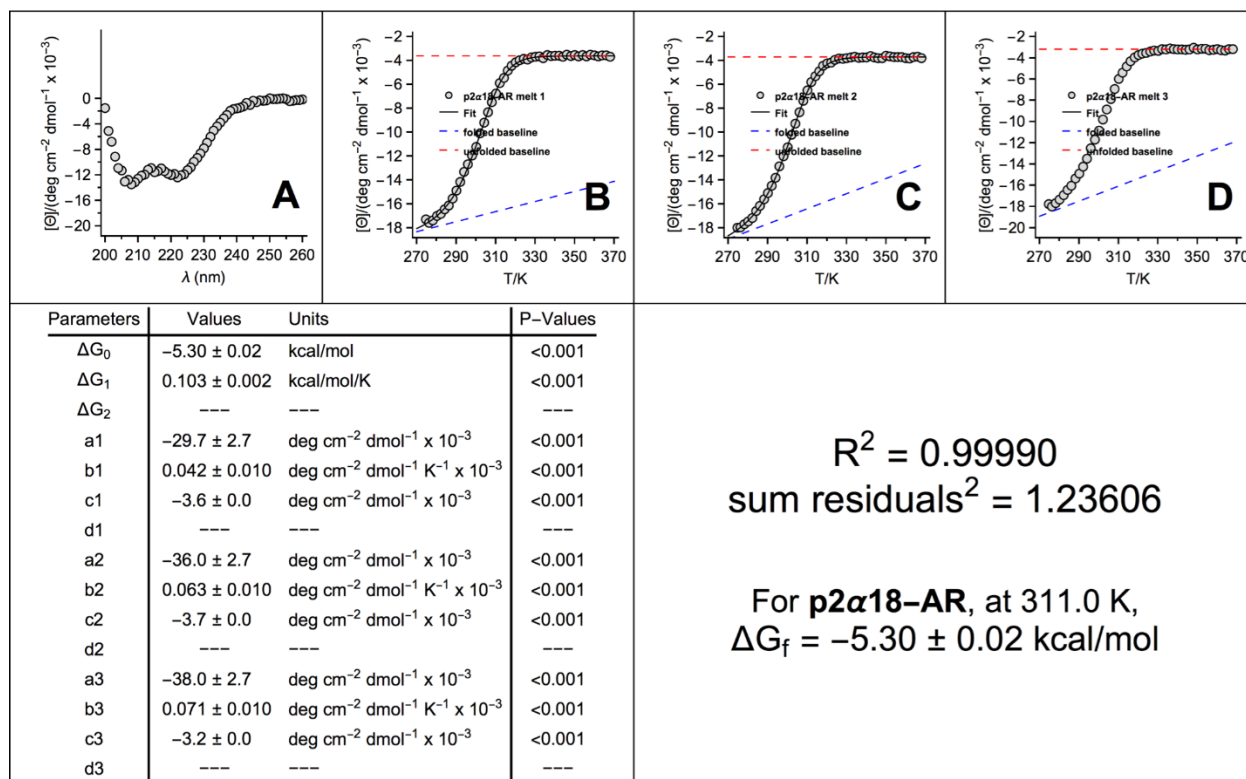


Figure S166. CD data for **p2 α 18-AR** (NAB10214).

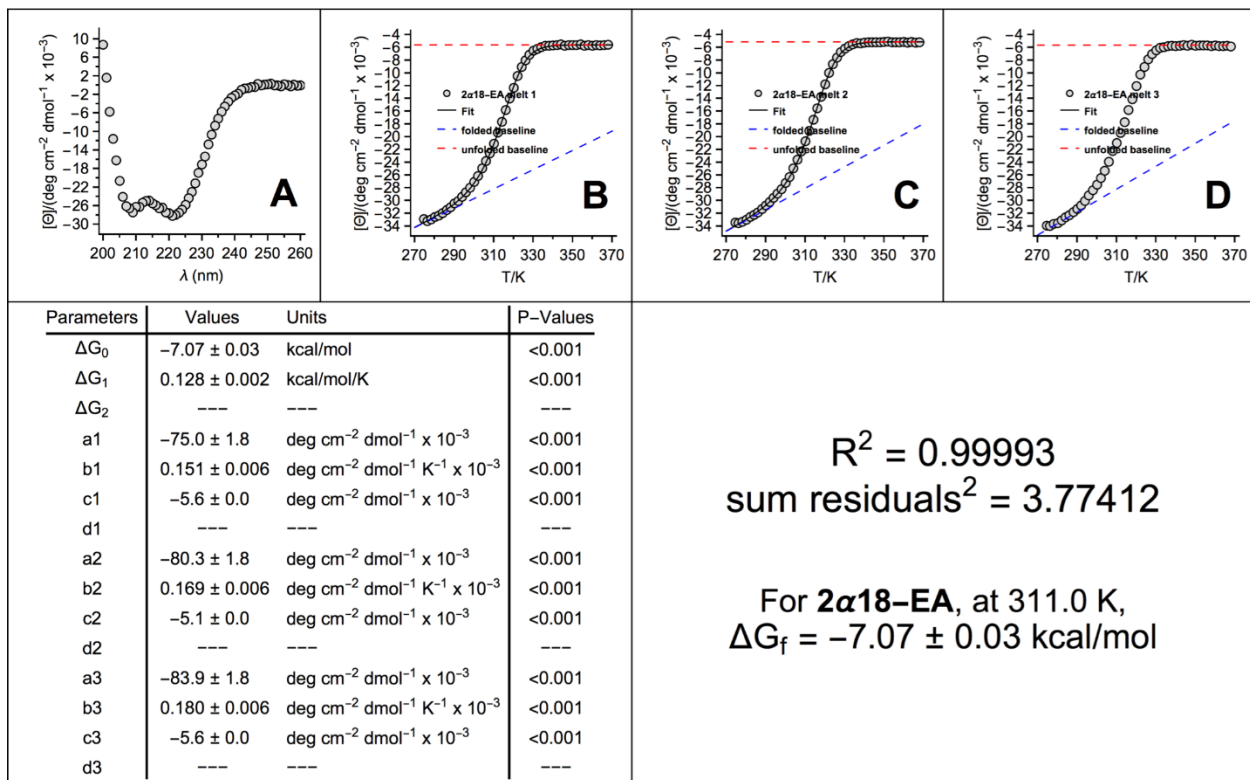


Figure S167. CD data **2 α 18-EA** (NAB10212).

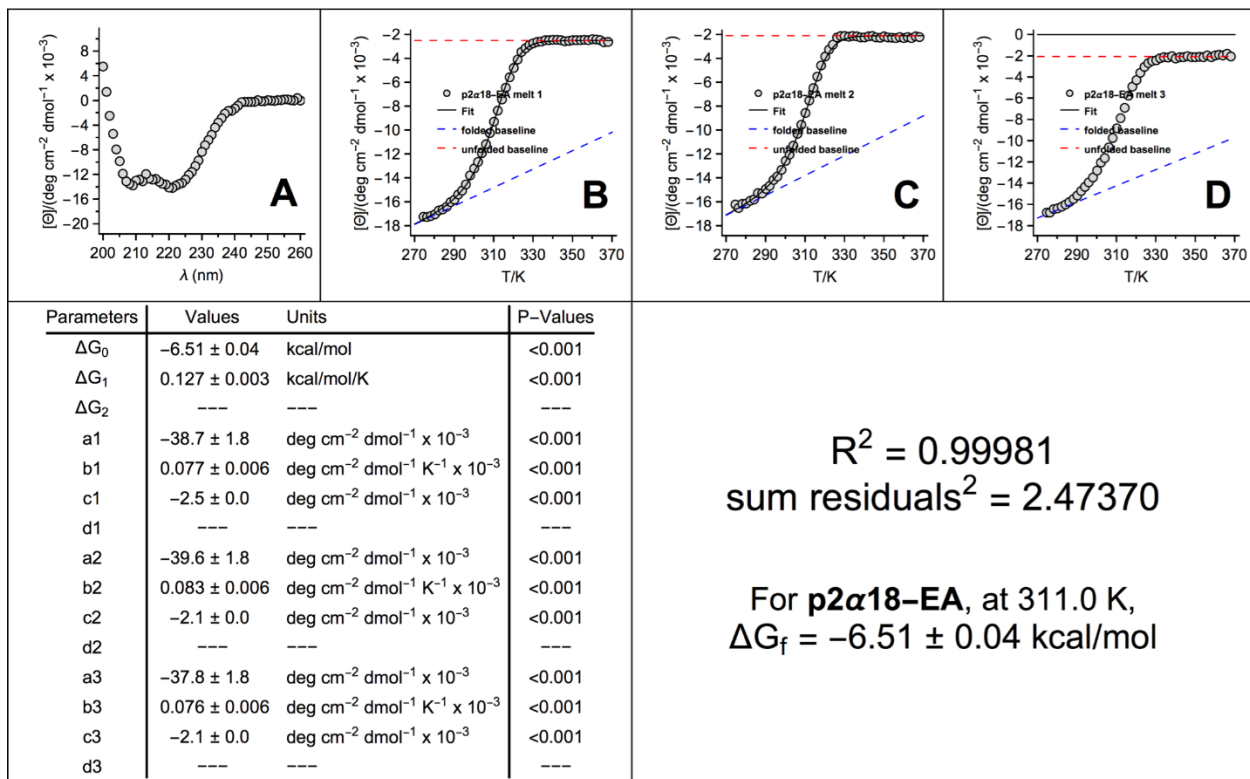


Figure S168. CD data for **p2 α 18-EA** (NAB10215).

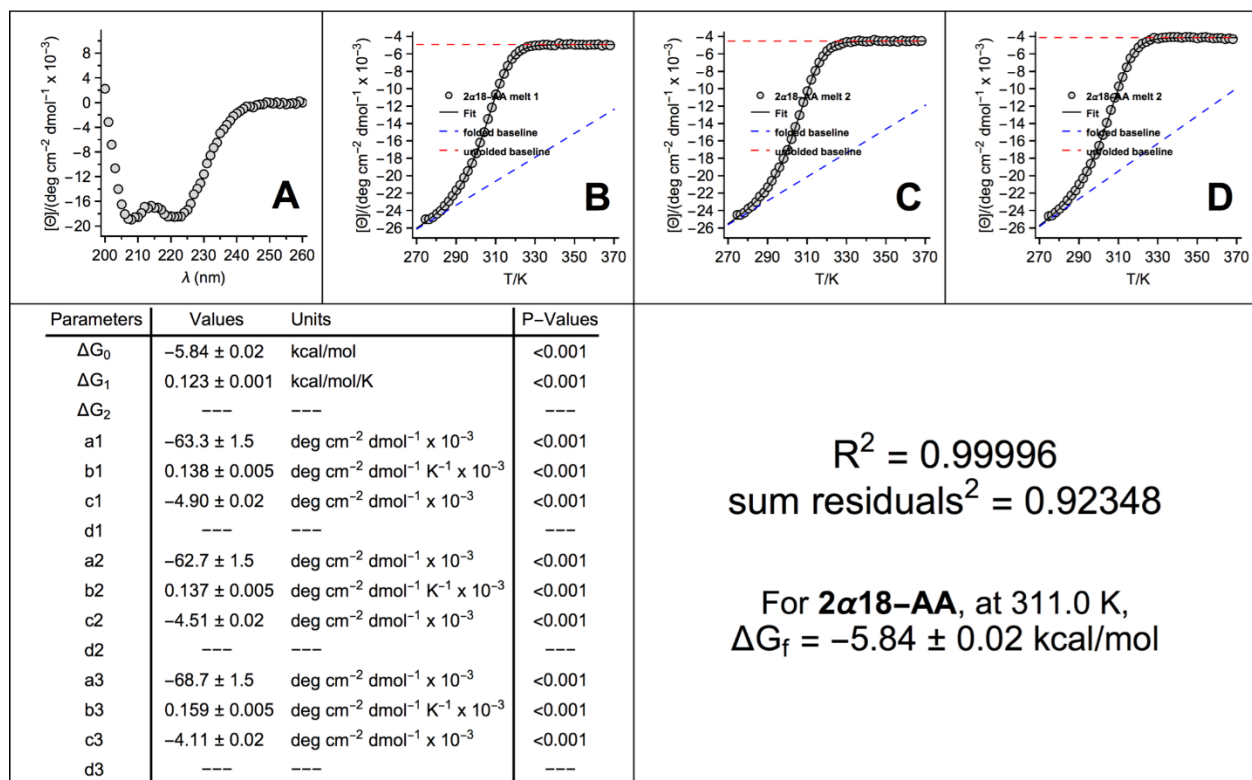


Figure S169. CD data for **2 α 18-AA** (NAB10213).

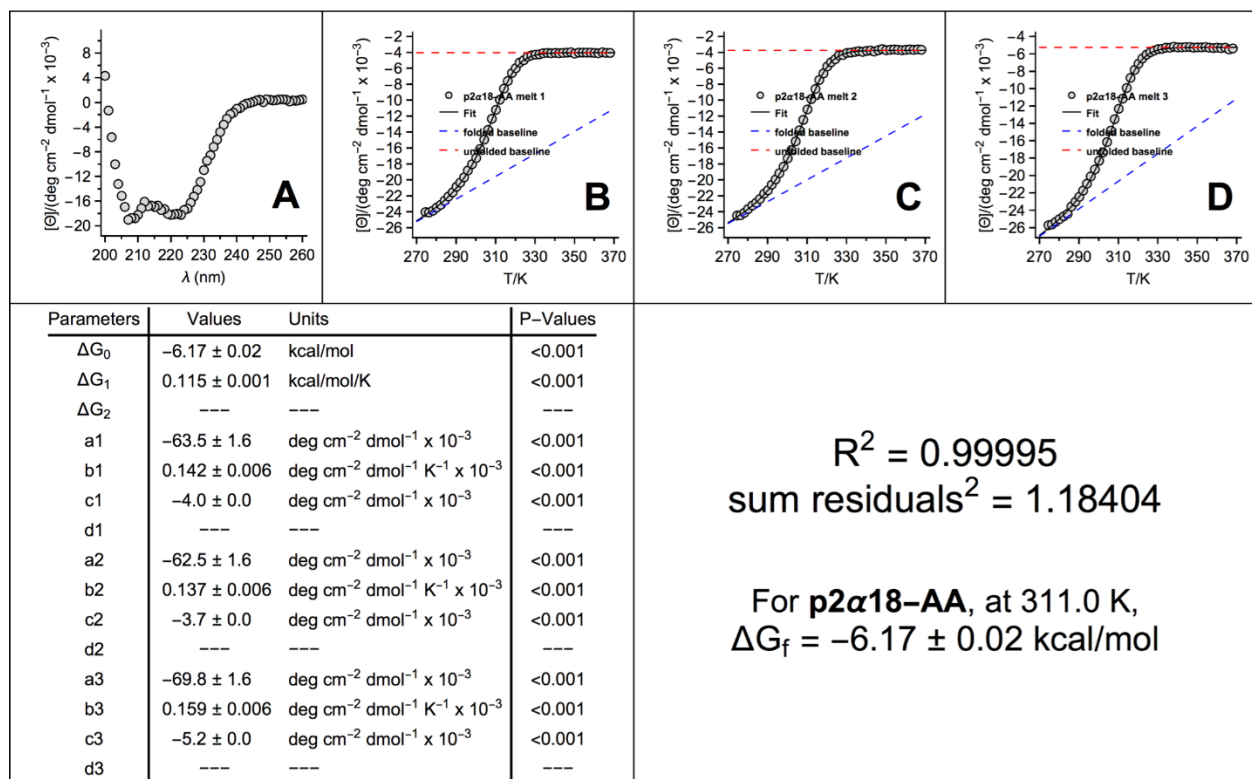


Figure S170. CD data for **p2 α 18-AA** (NAB10216).

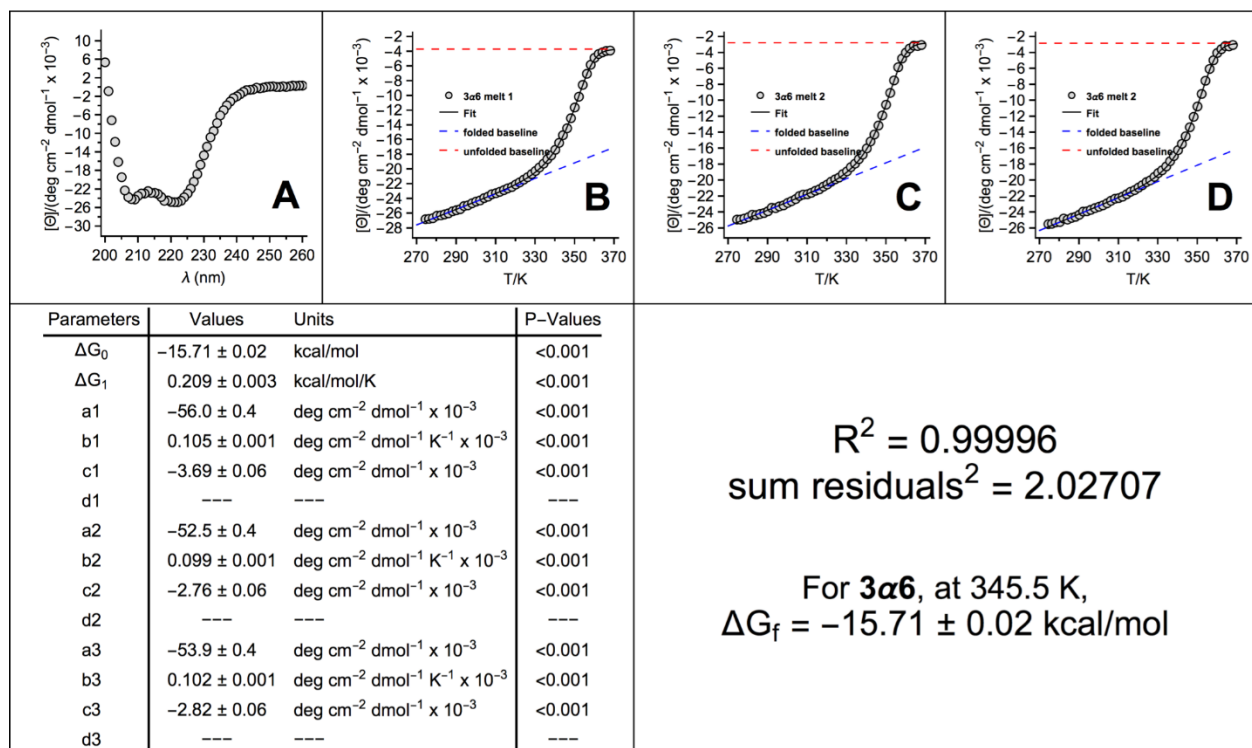


Figure S171. CD data for **3α6** (QX10511).

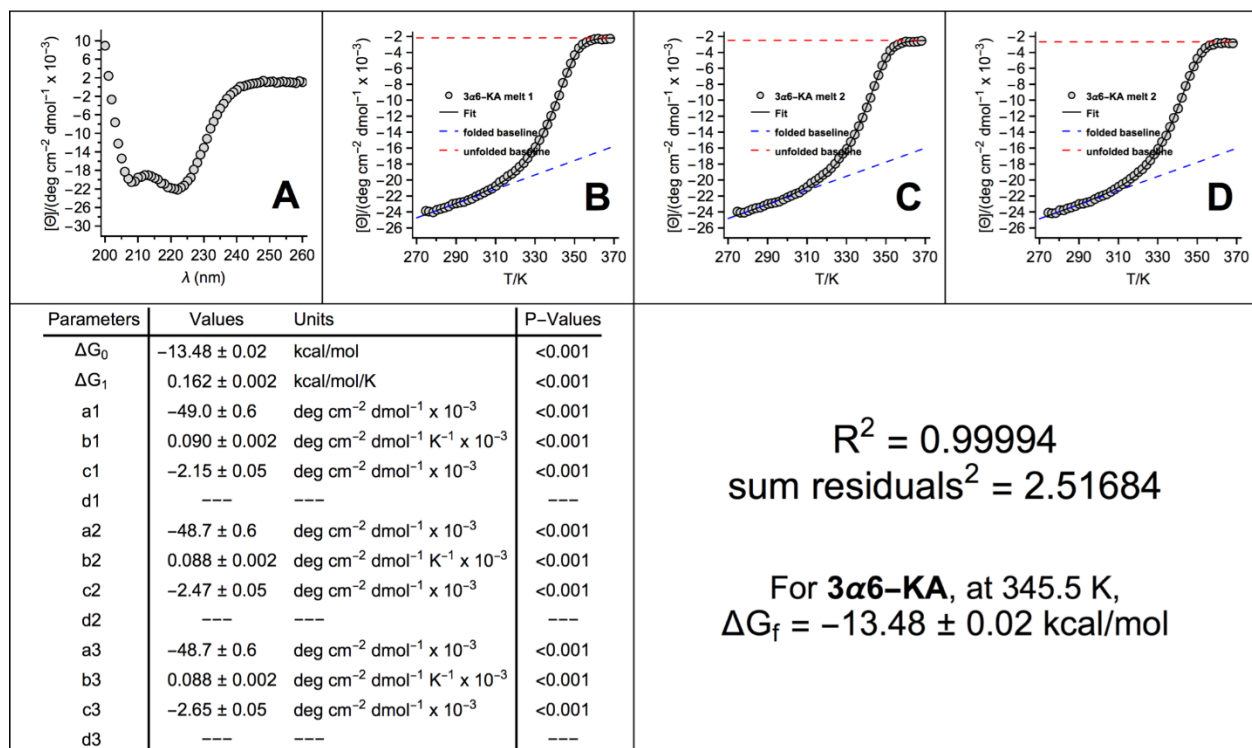


Figure S172. CD data for **3α6-KA** (QX10512).

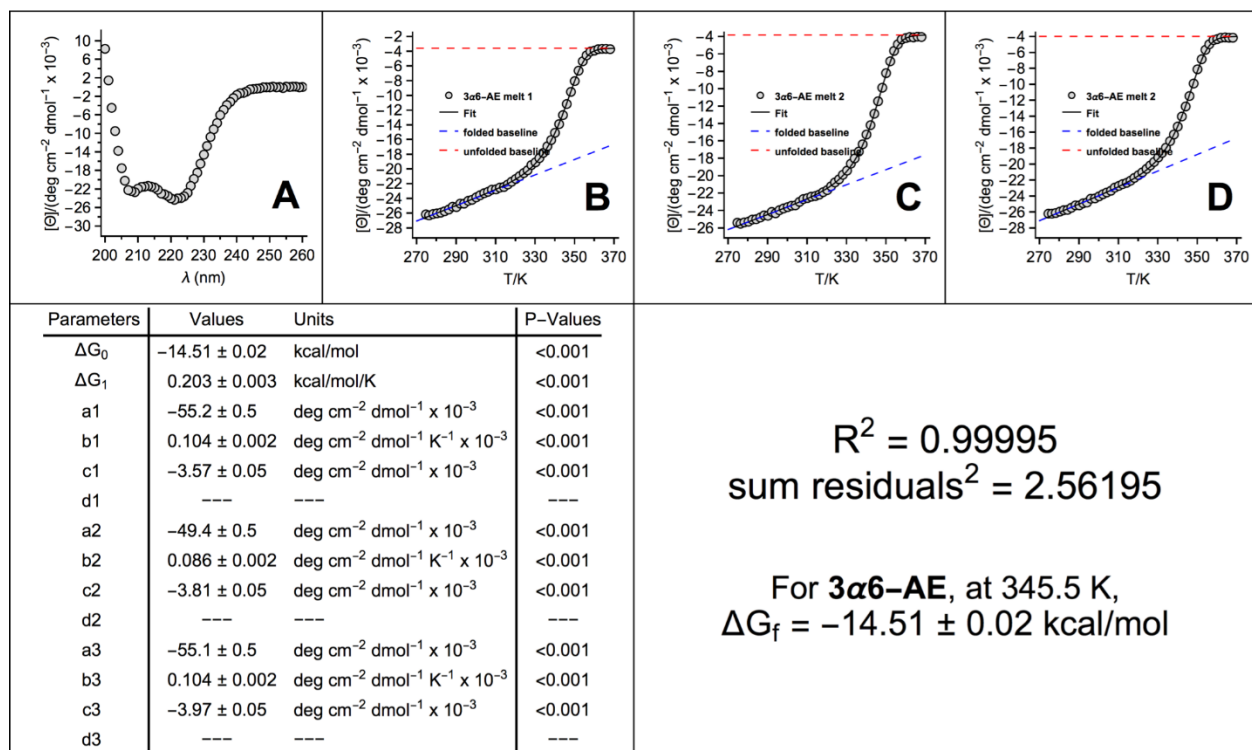


Figure S173. CD data for 3 α 6-AE (QX10513).

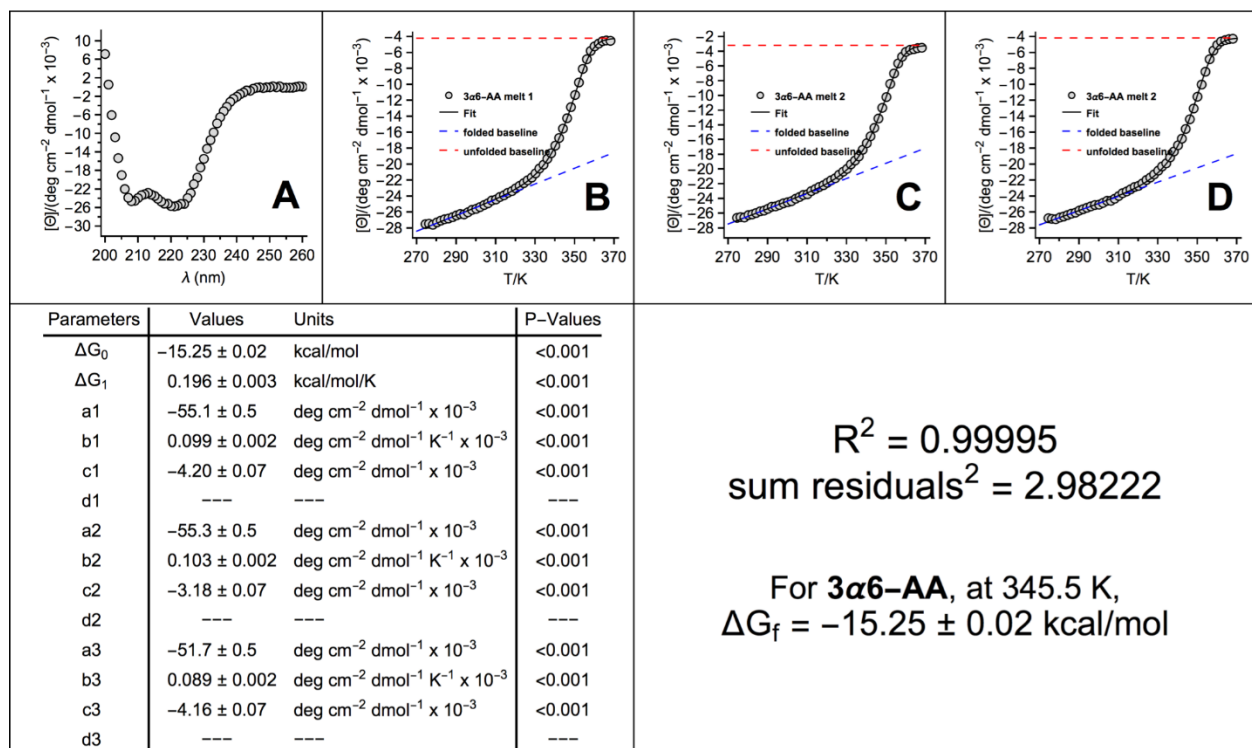


Figure S174. CD data for 3 α 6-AA (QX10514).

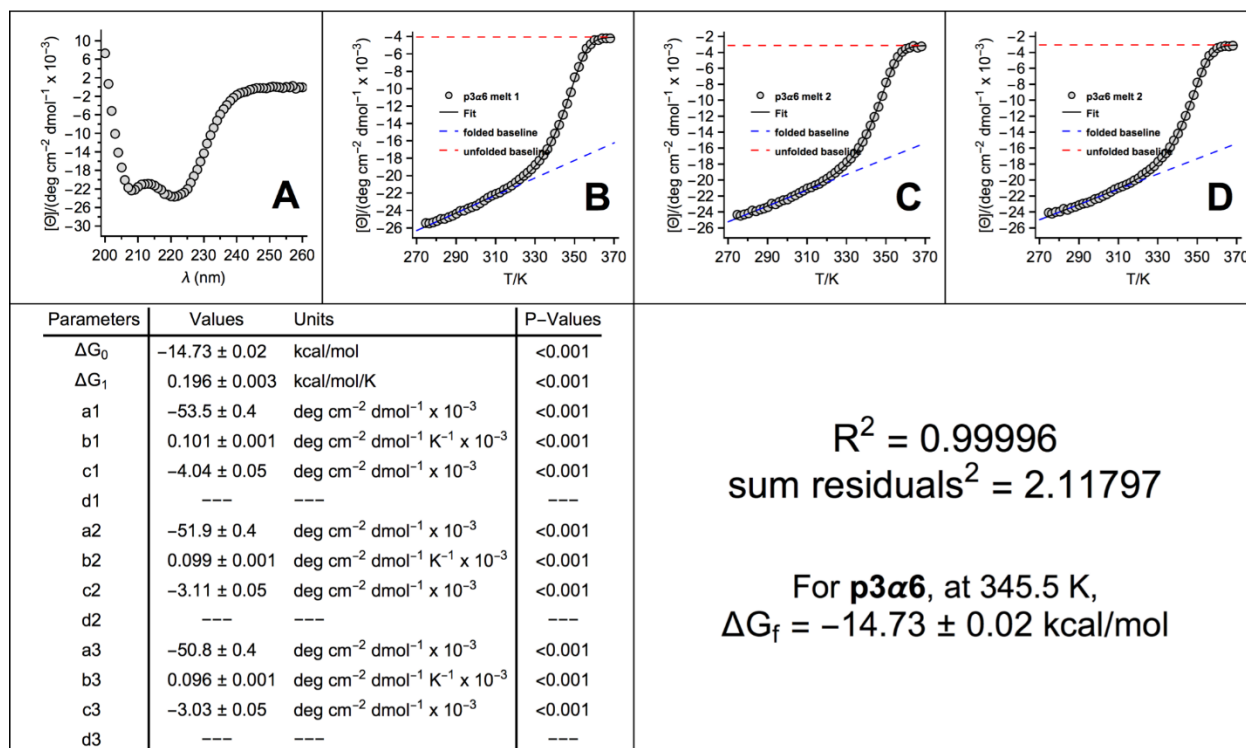


Figure S175. CD data for p3α6 (QX10515).

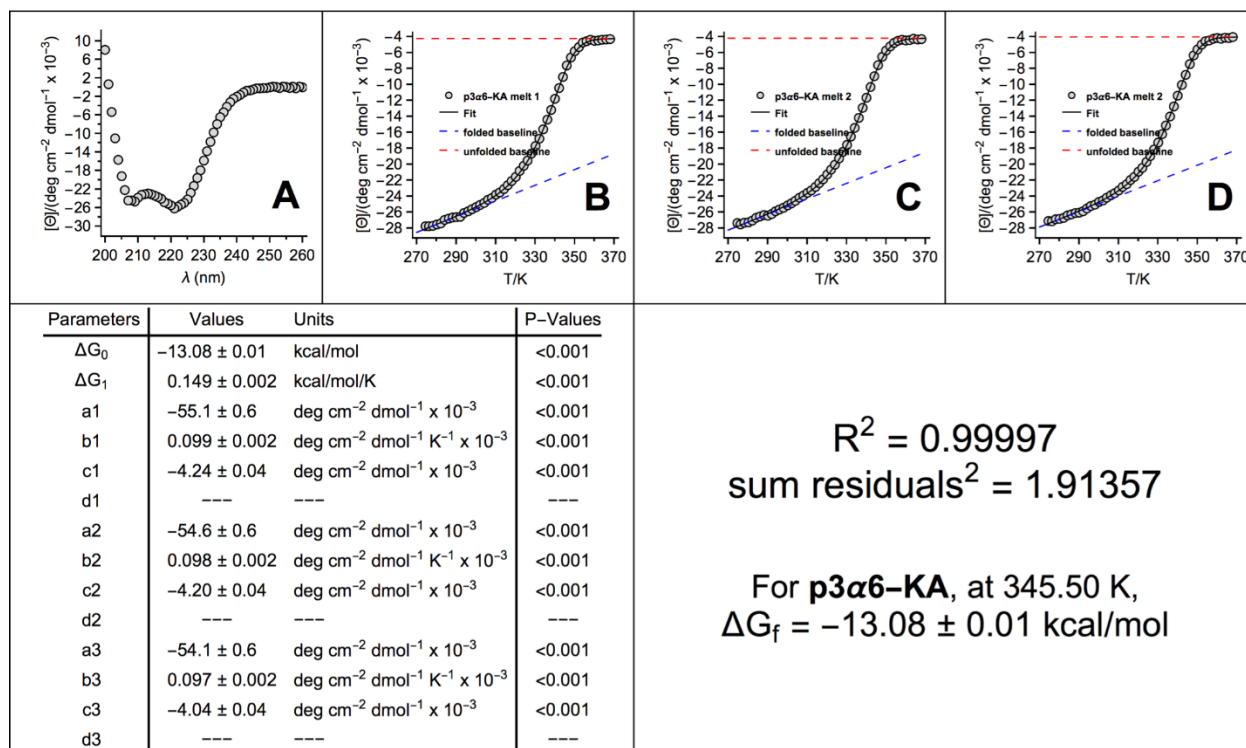


Figure S176. CD data for p3α6-KA (QX10516).

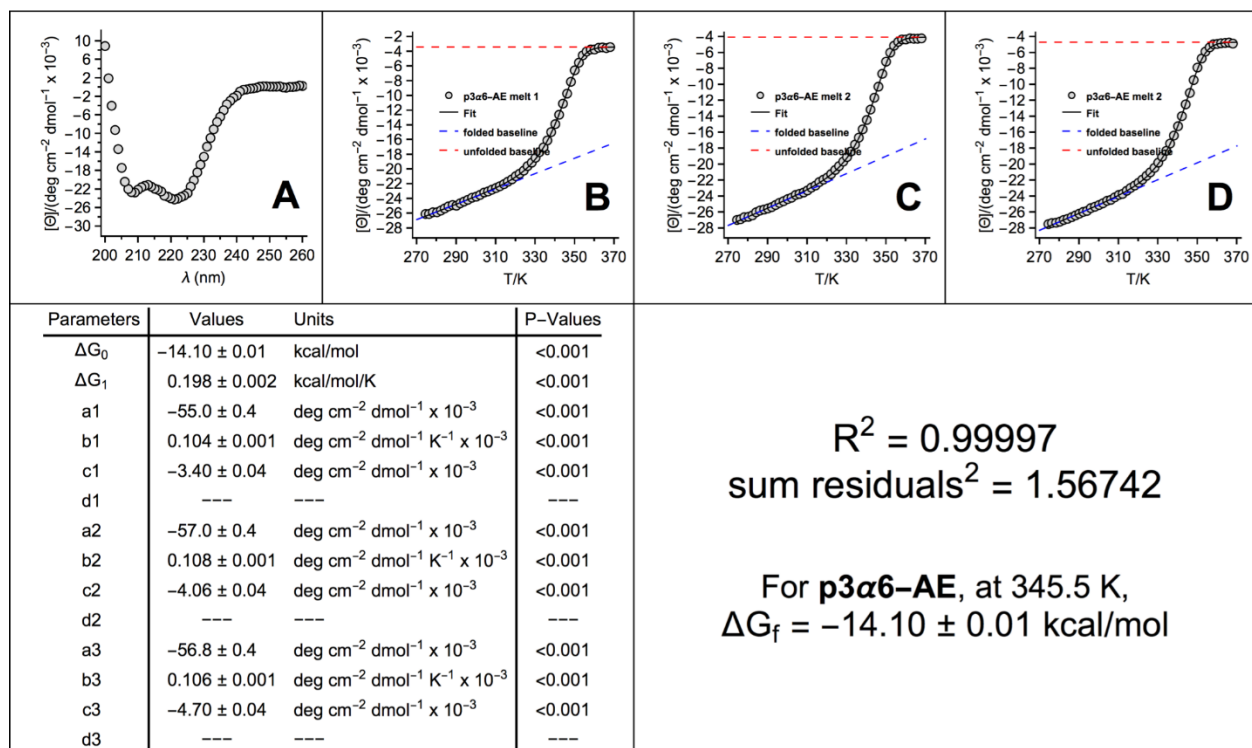


Figure S177. CD data for p3a6-AE (QX10517).

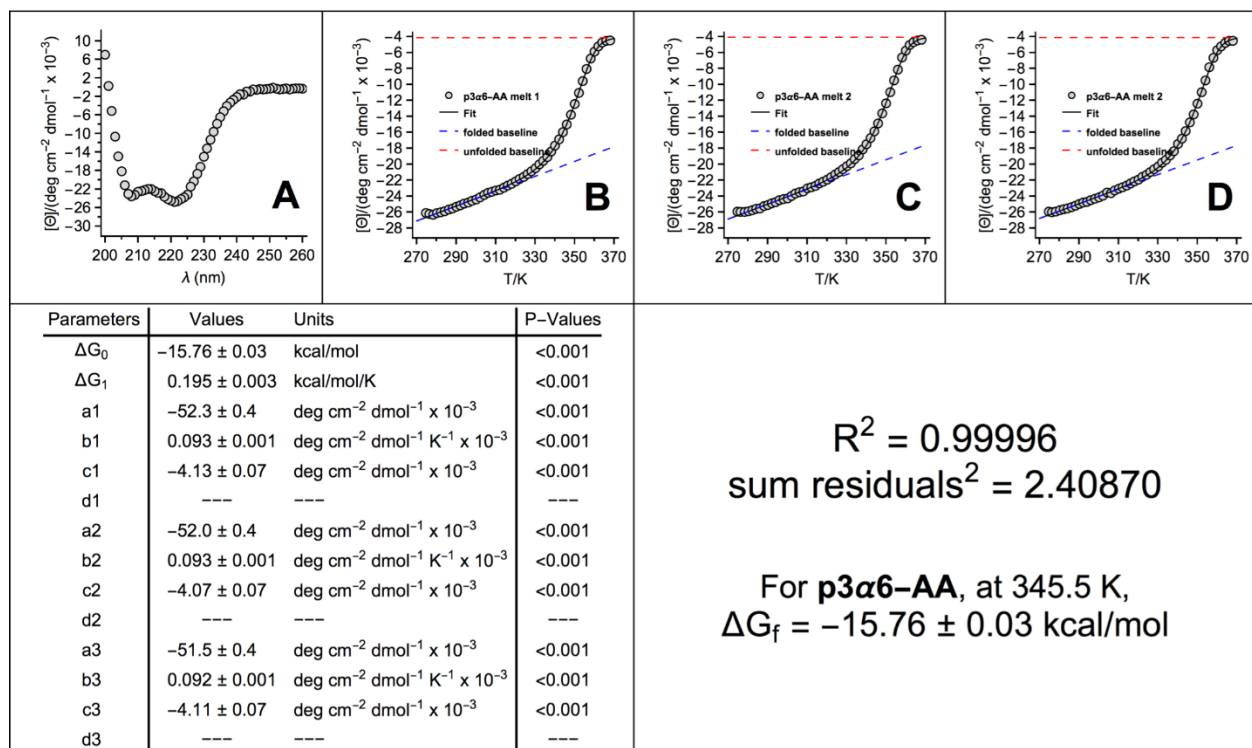


Figure S178. CD data for p3a6-AA (QX10518).

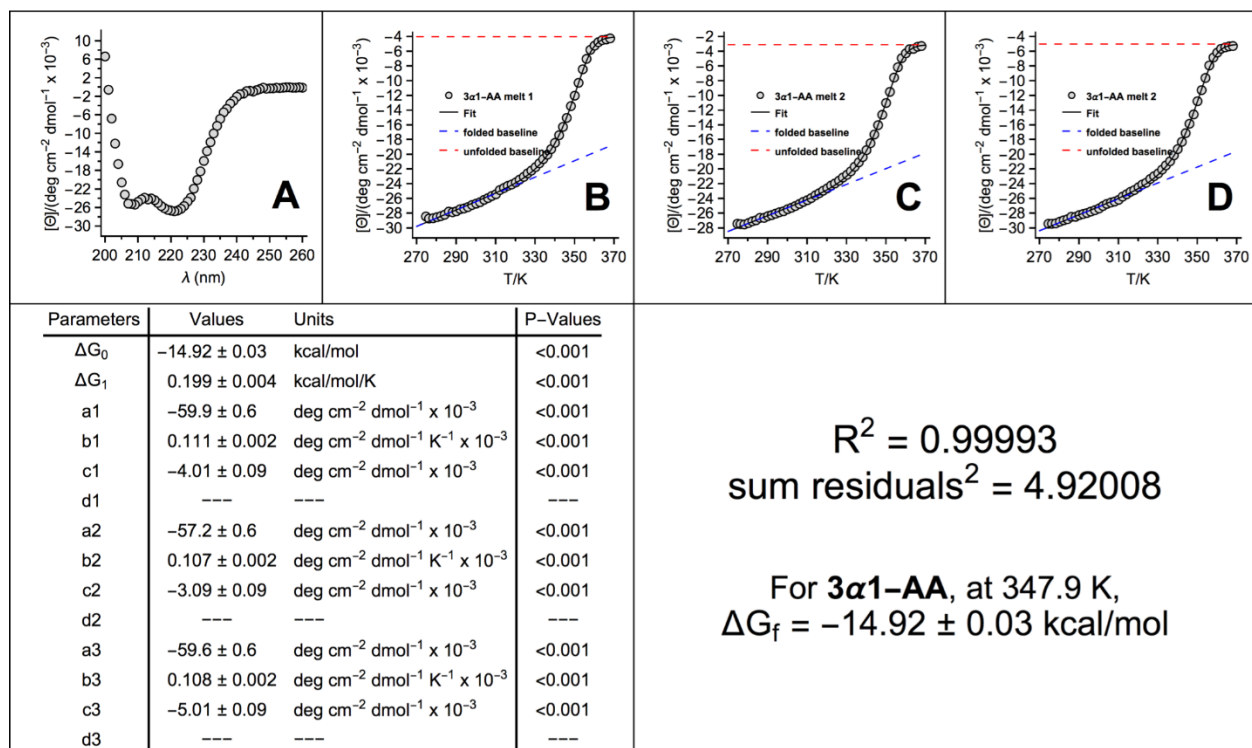


Figure S179. CD spectra 3 α 1-AA (QX10711).

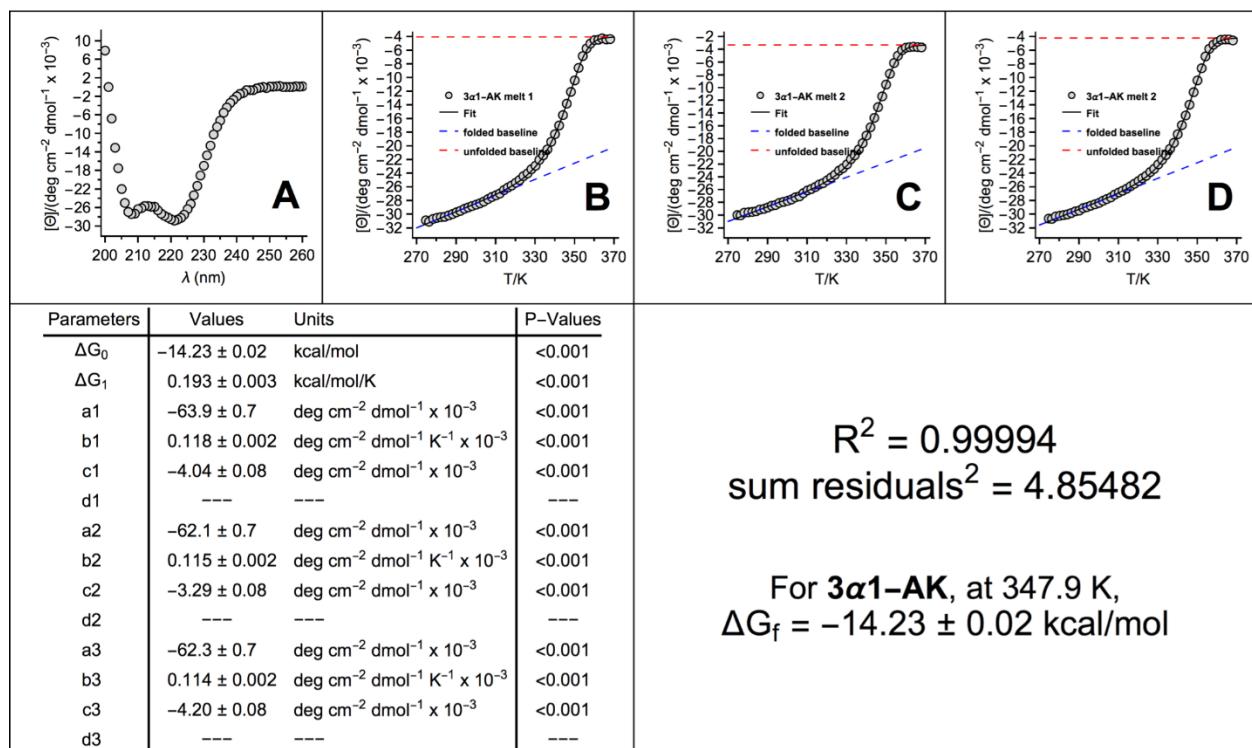


Figure S180. CD spectra for 3 α 1-AK (QX10712).

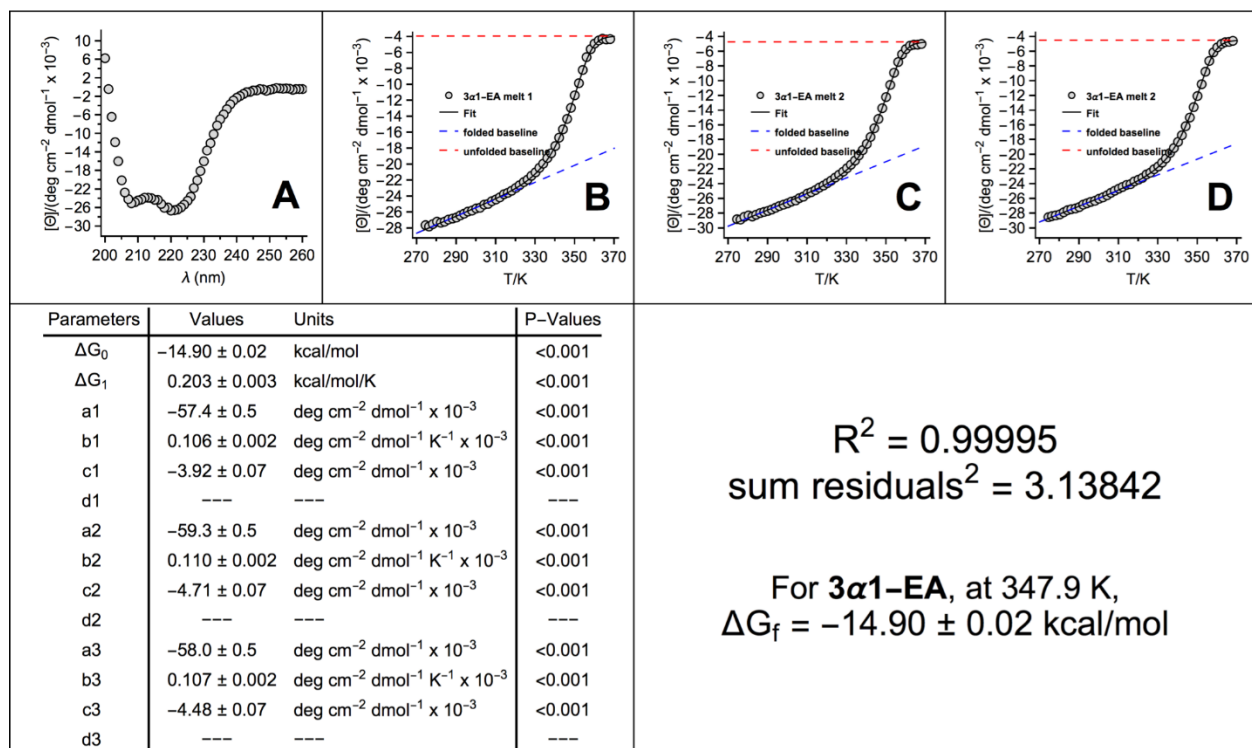


Figure S181. CD data for 3 α 1-EA (QX10713).

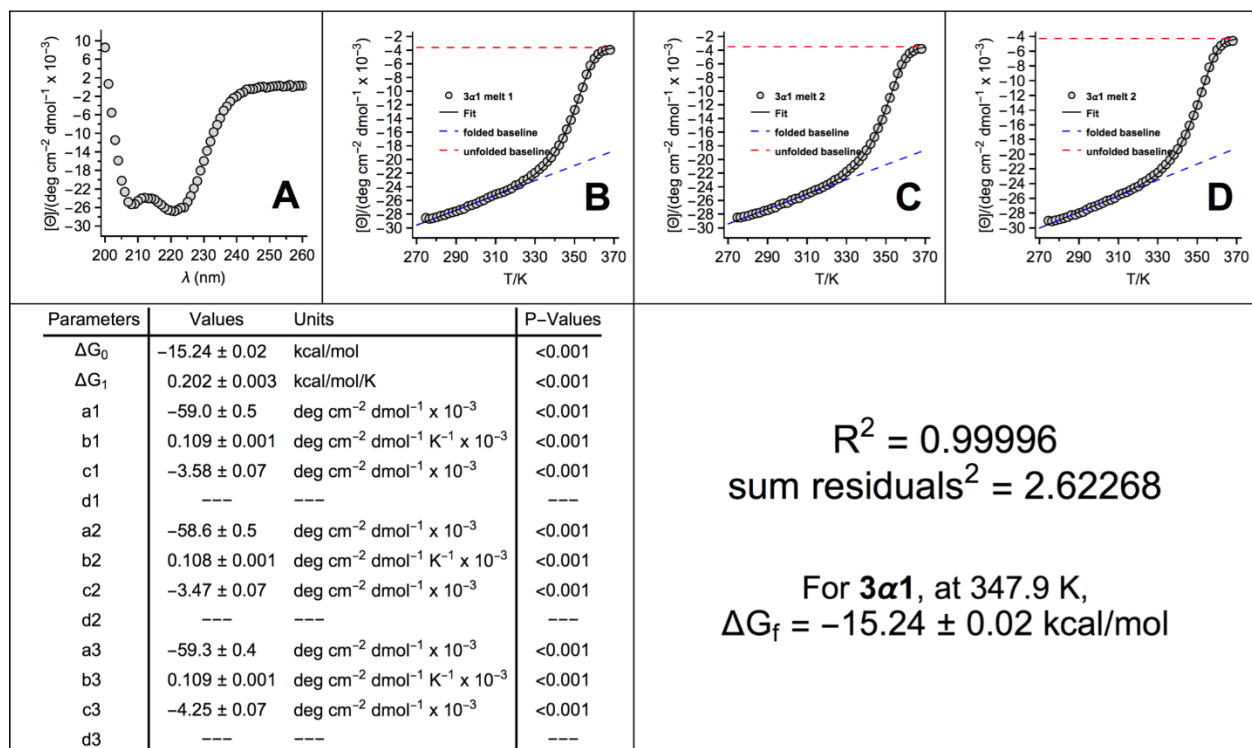


Figure S182. CD data for 3 α 1 (QX10714).

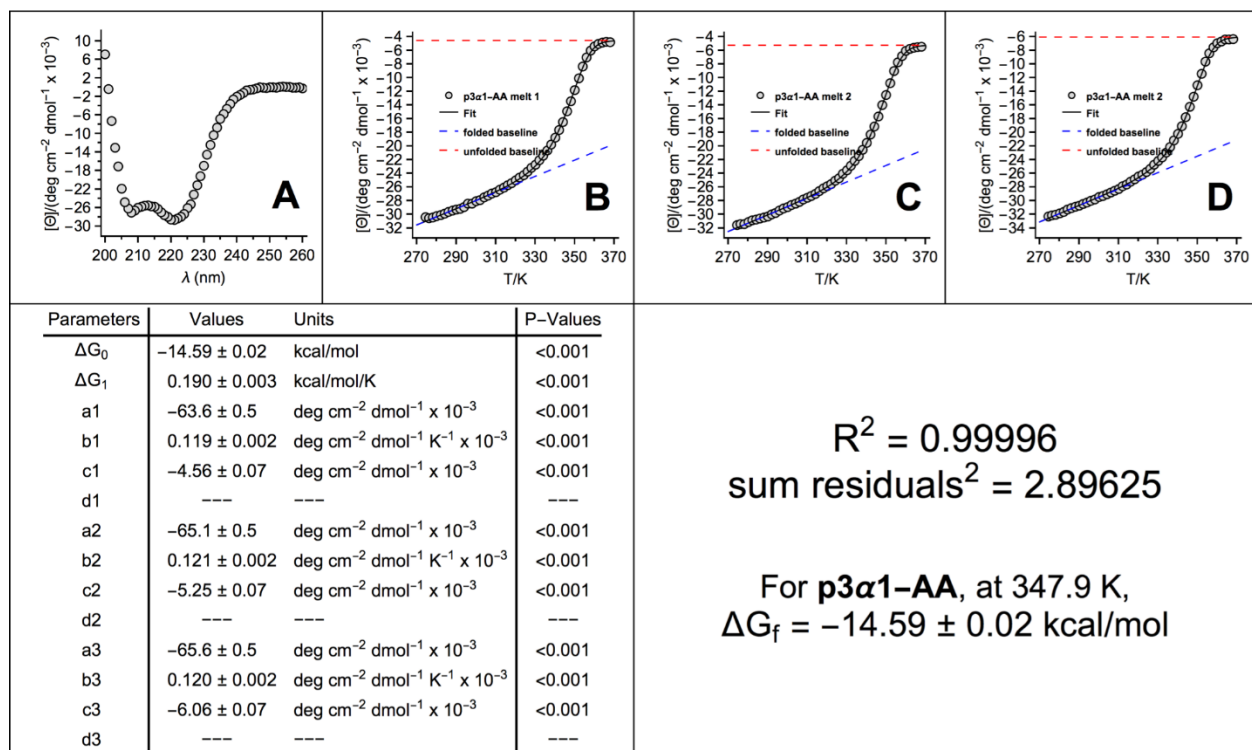


Figure S183. CD data for p3α1-AA (QX10715).

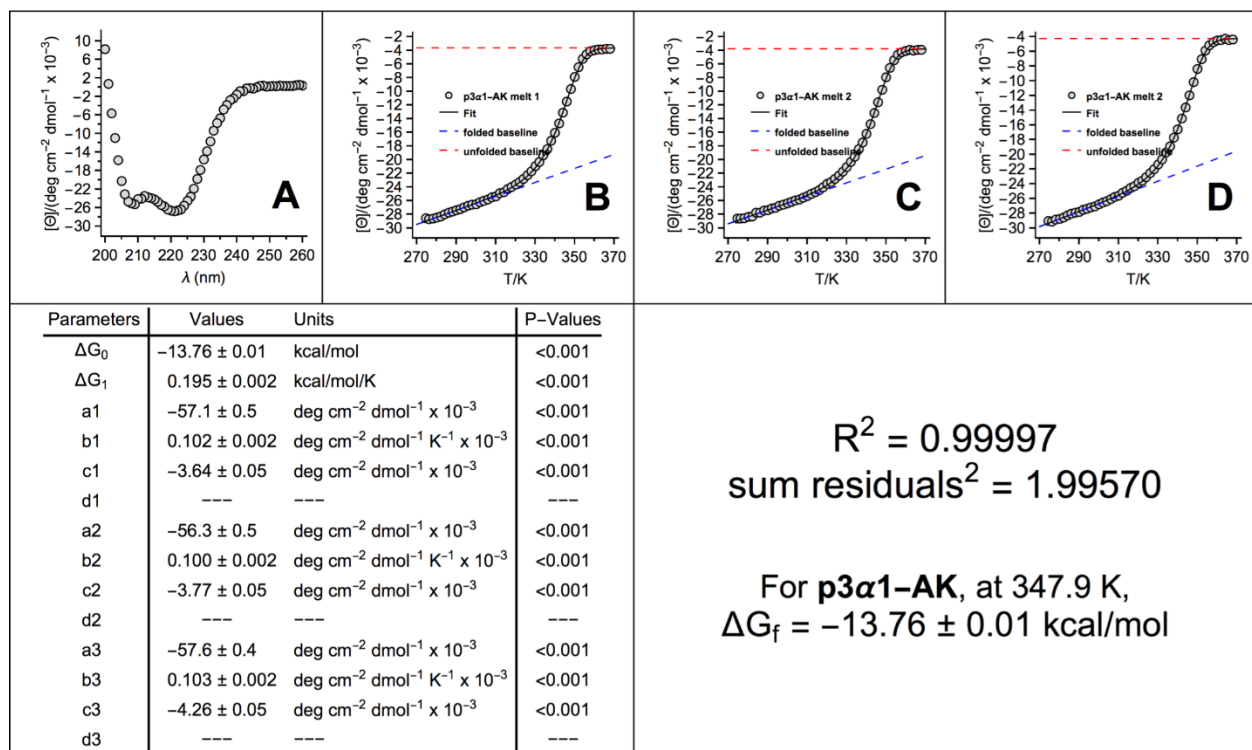


Figure S184. CD data for p3α1-AK (QX10716).

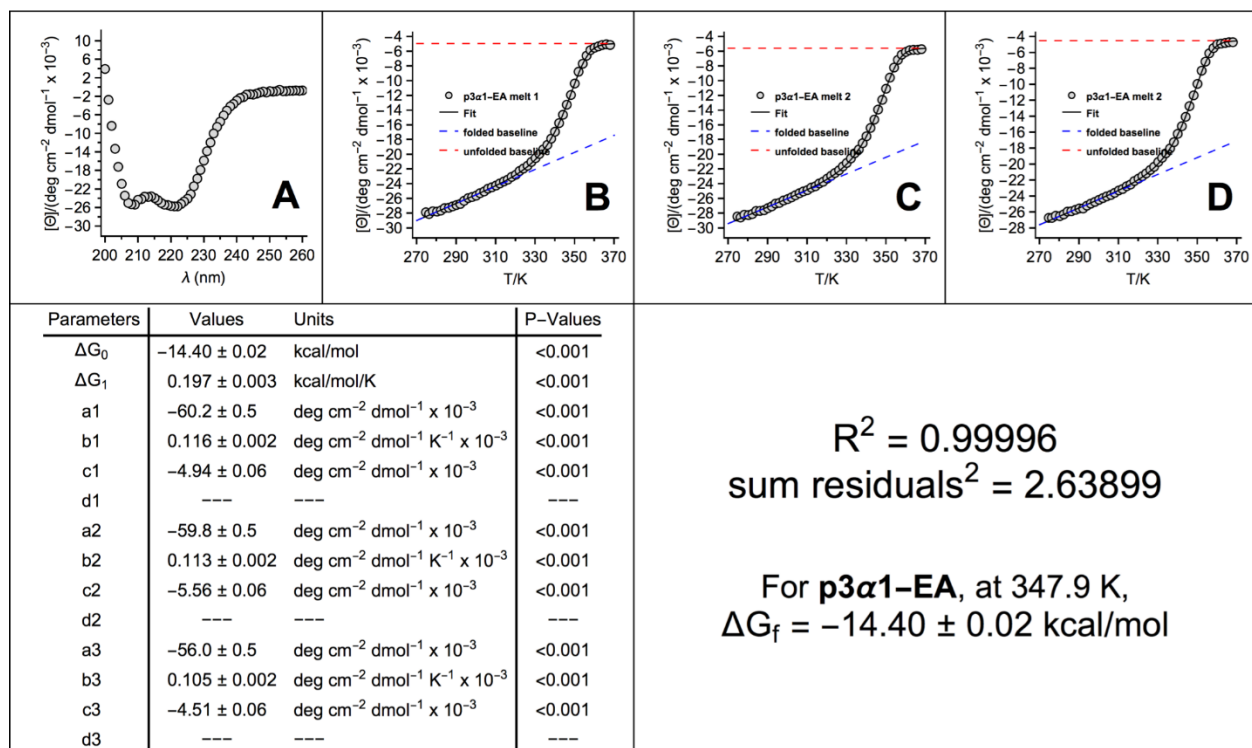


Figure S185. CD data for p3α1-EA (QX10717).

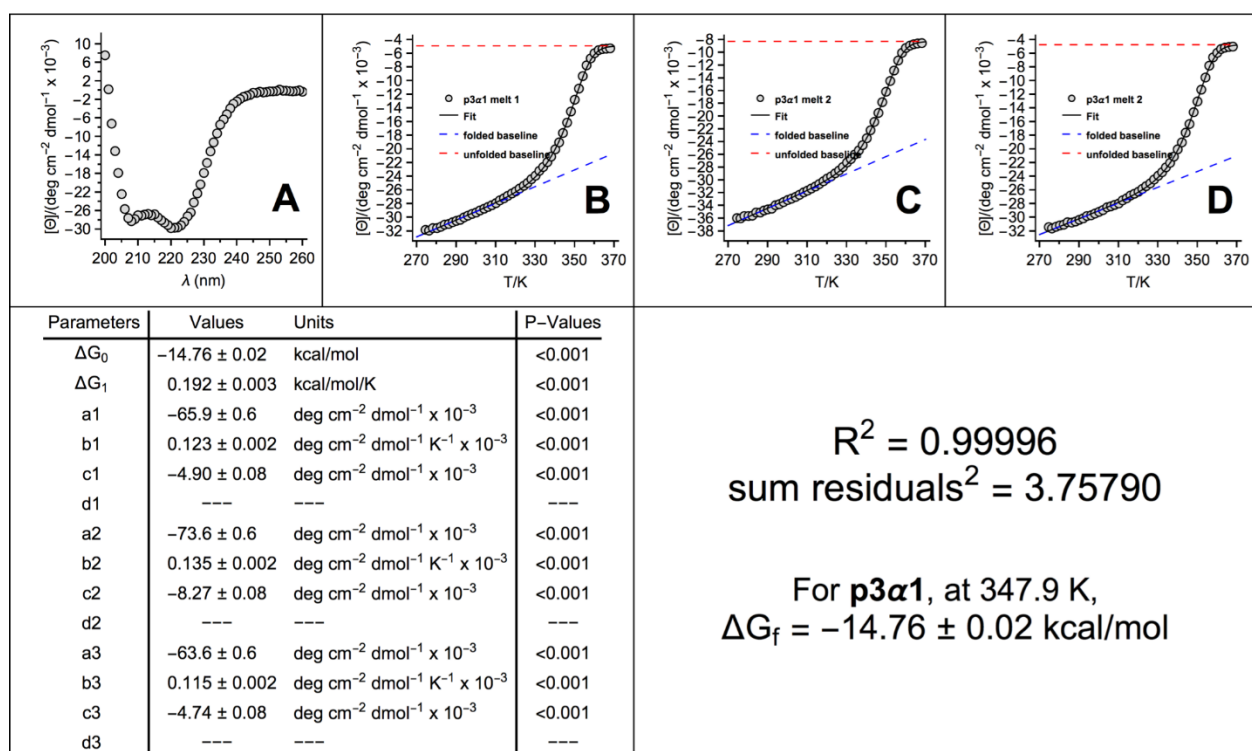


Figure S186. CD data for p3α1 (QX10718).

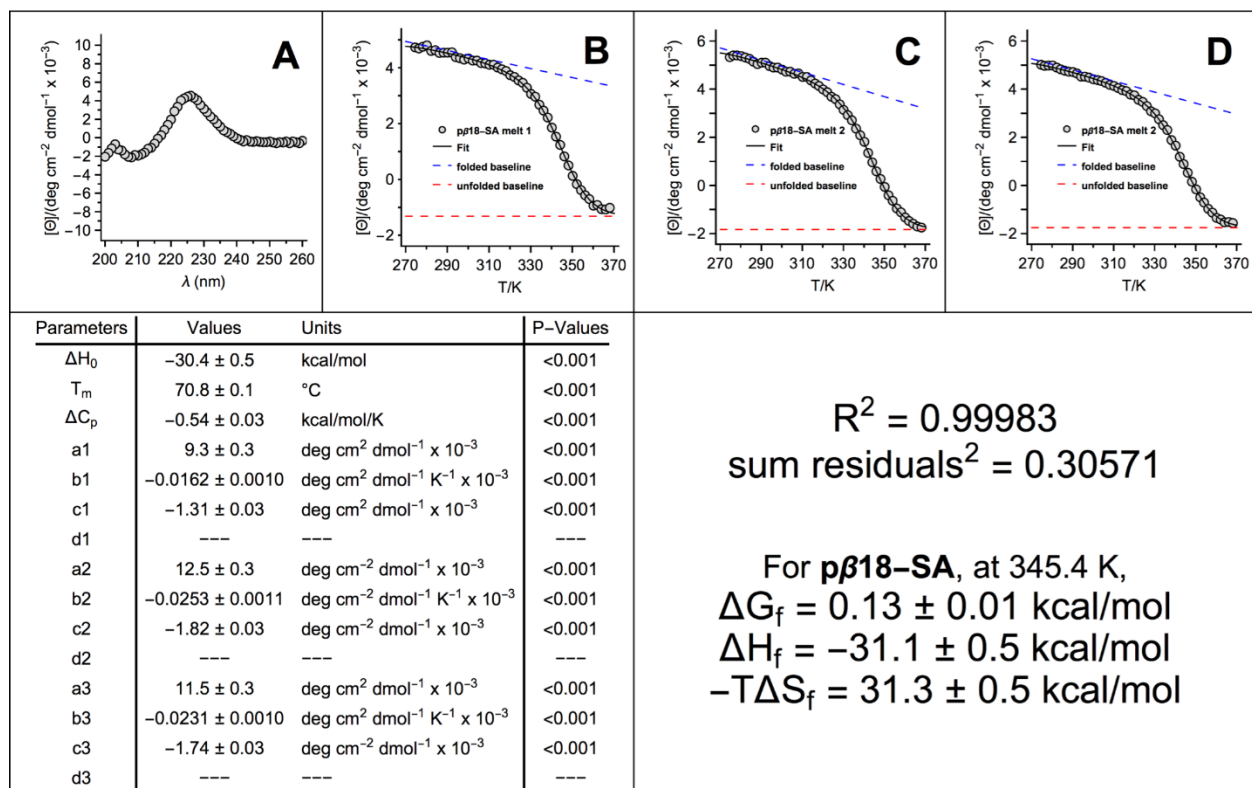


Figure S187. CD data for pβ18-SA (QX10491).

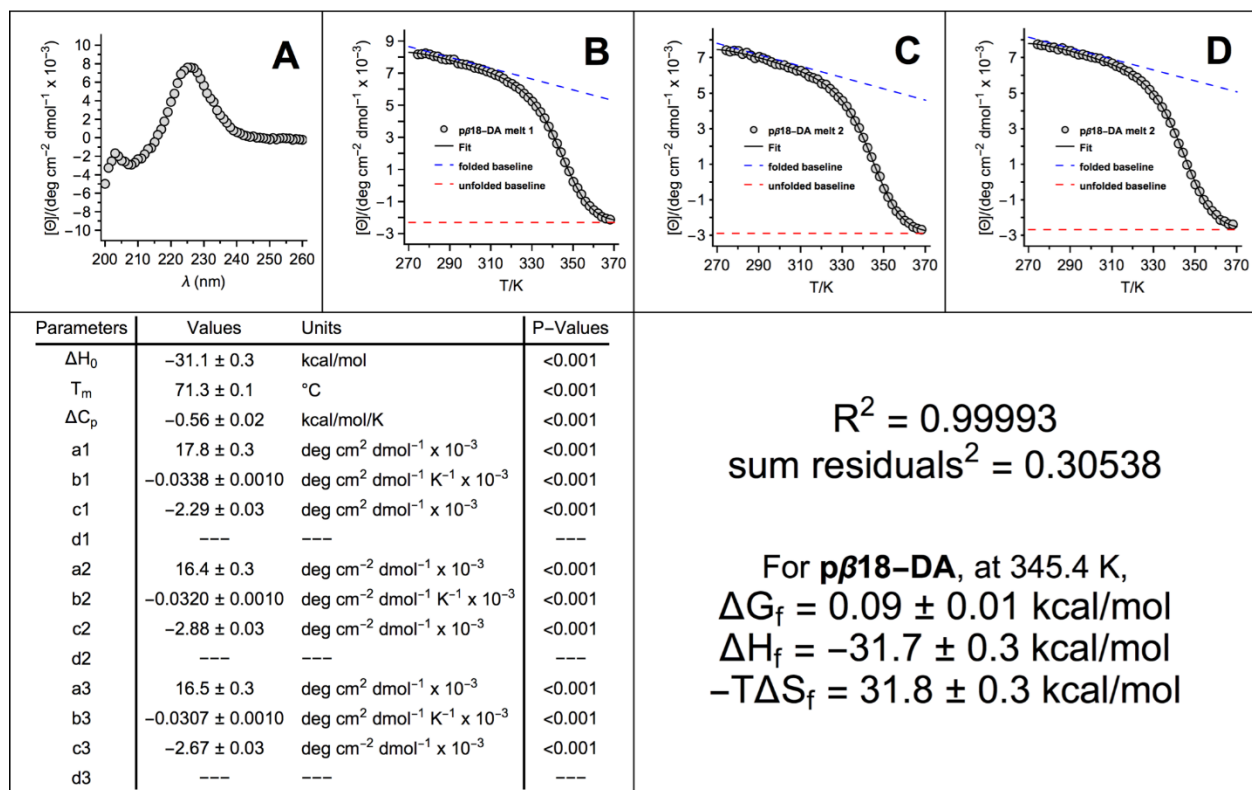


Figure S188. CD data for pβ18-DA (QX10492).

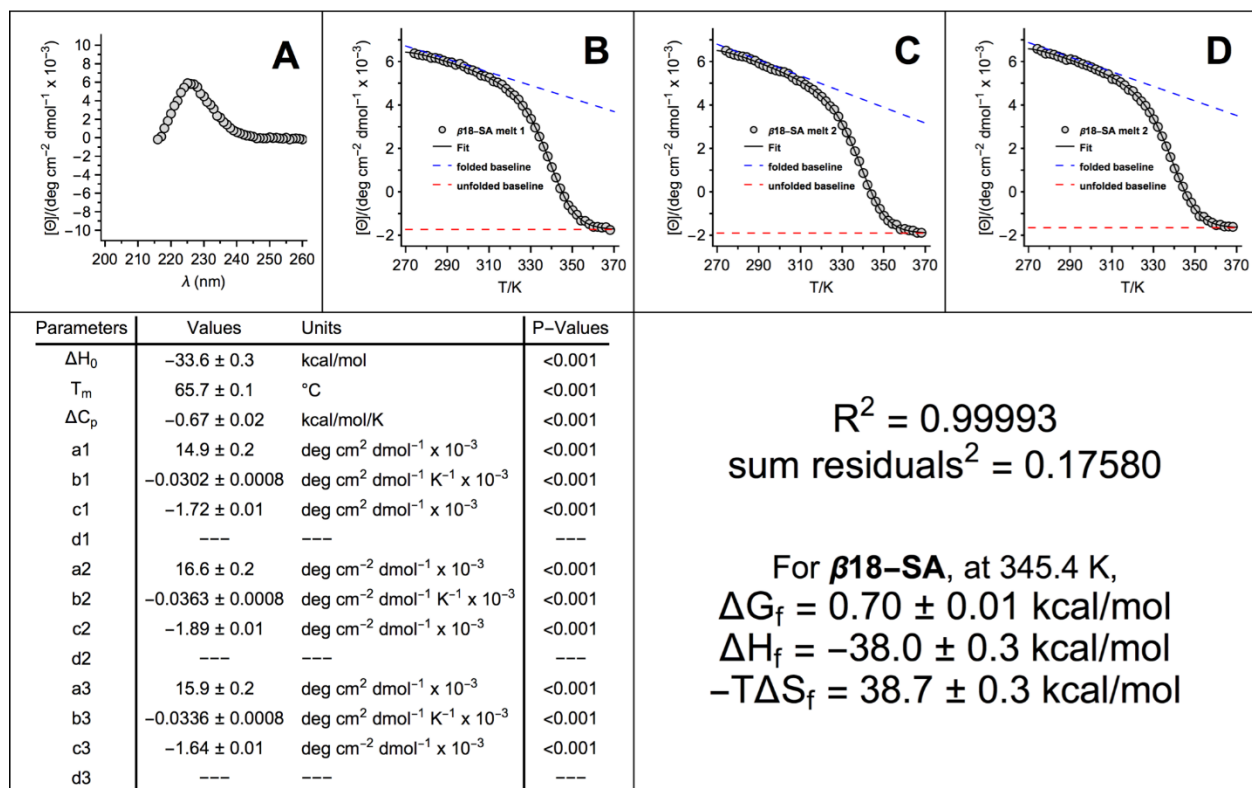


Figure S189. CD data for $\beta 18\text{-SA}$ (QX10493).

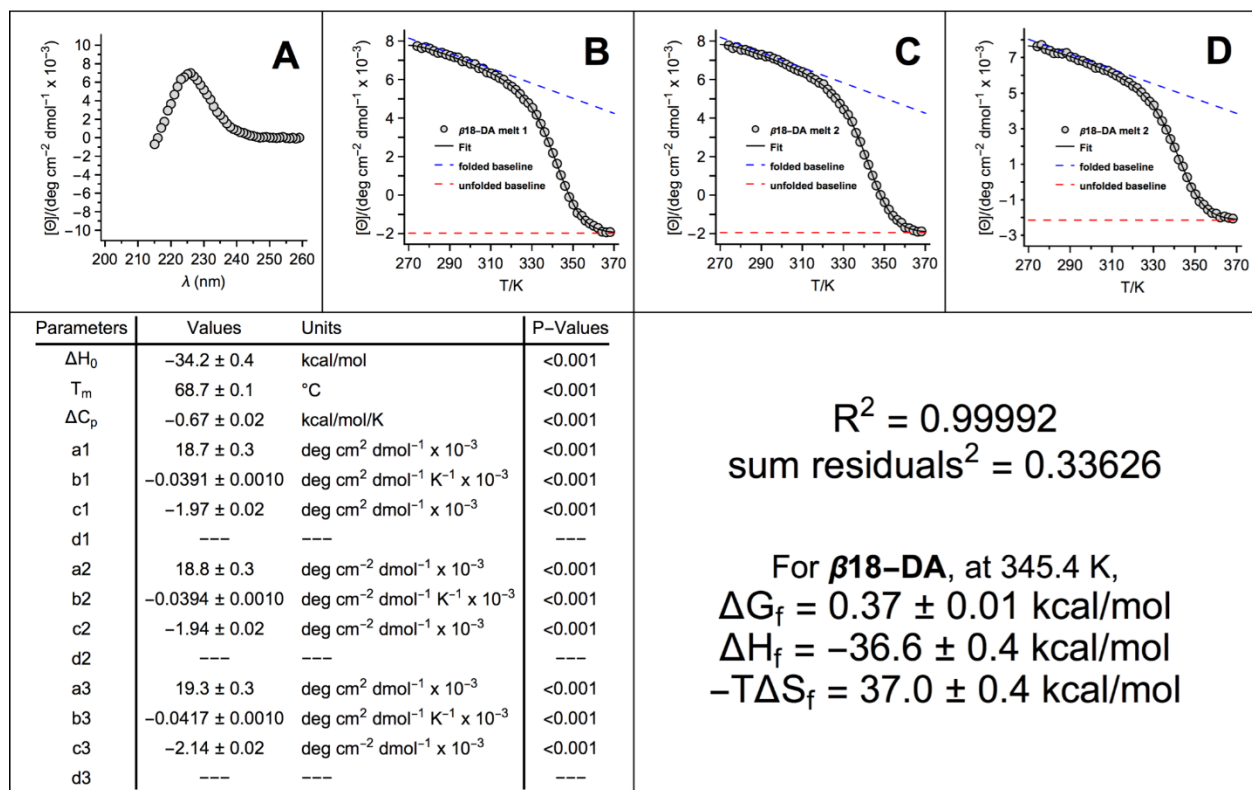


Figure S190. CD data for $\beta 18\text{-DA}$ (QX10494).

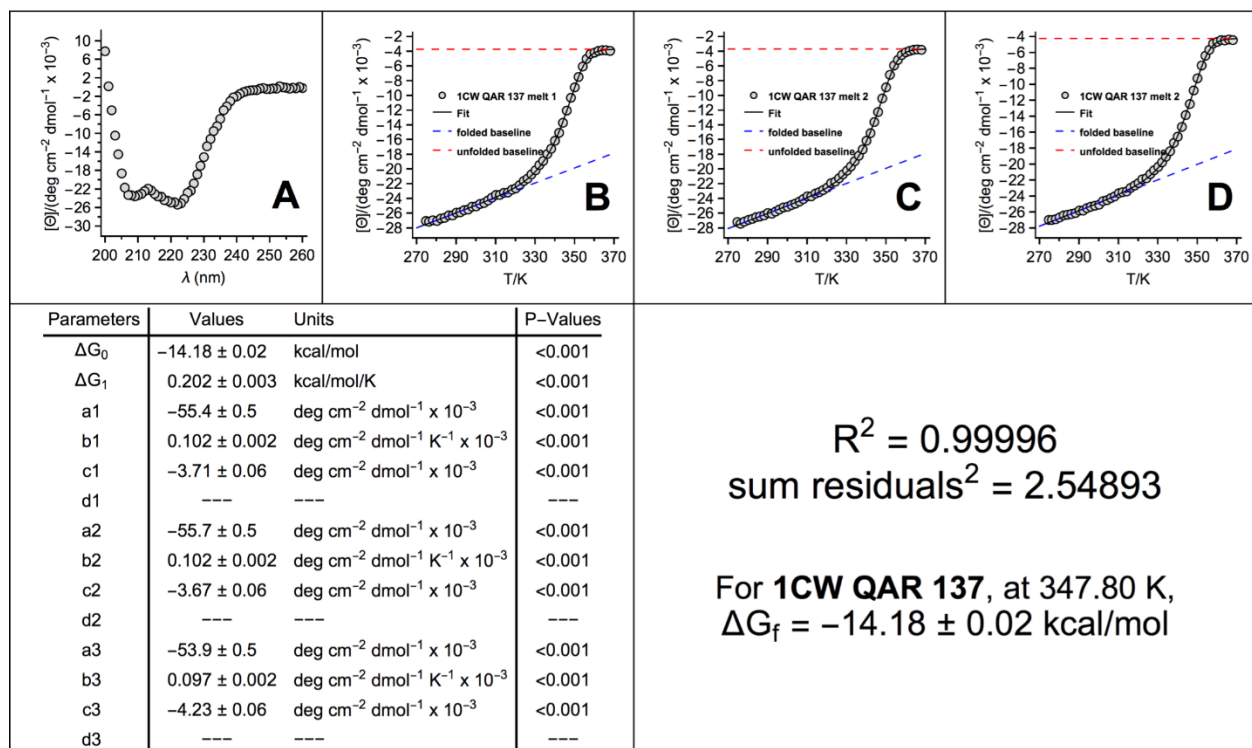


Figure S191. CD data for 3a1-AR (QX21531).

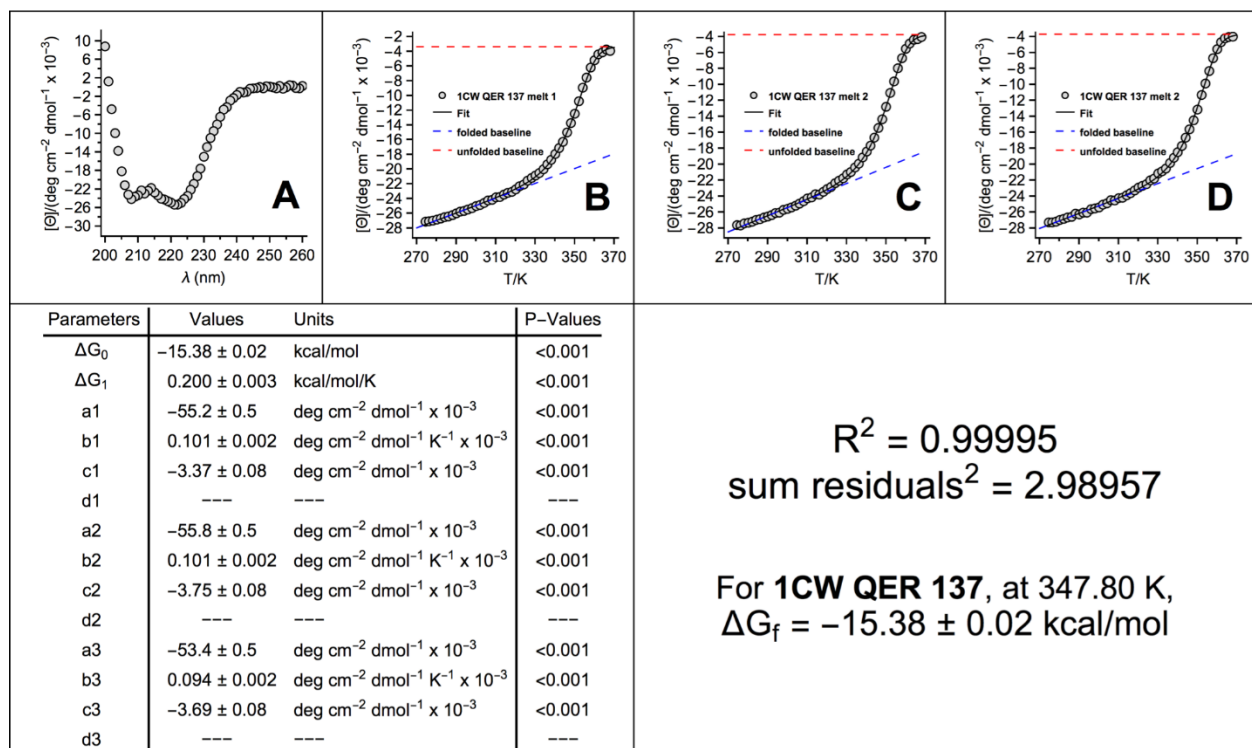


Figure S192. CD data for 3a1-ER (QX21532).

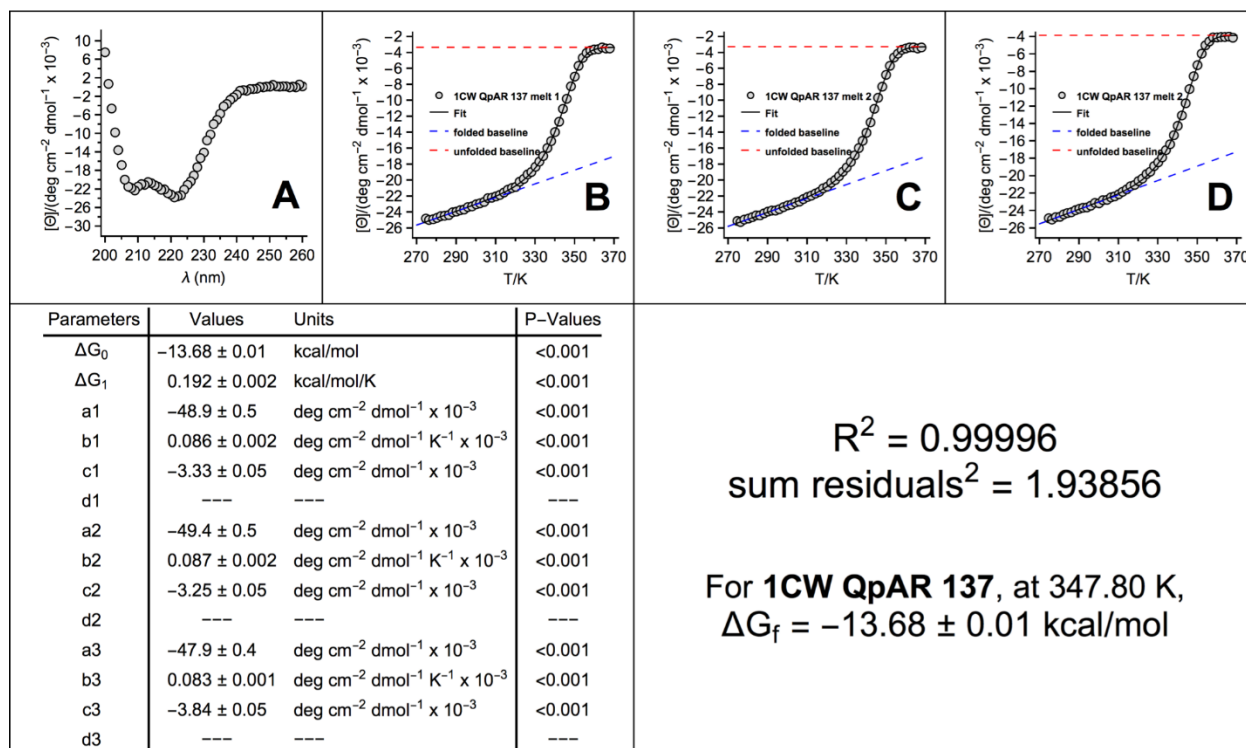


Figure S193. CD data for p3a1-AR (QX21533).

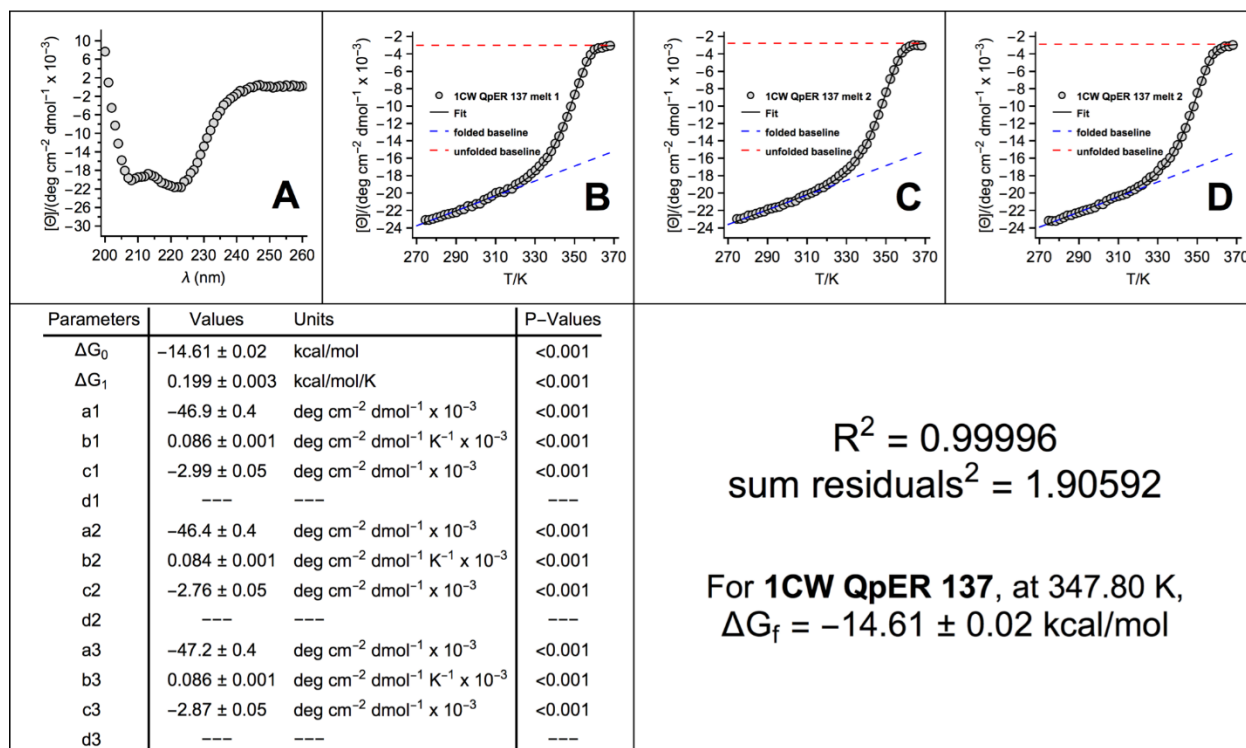


Figure S194. CD data for p3a1-ER (QX21534).

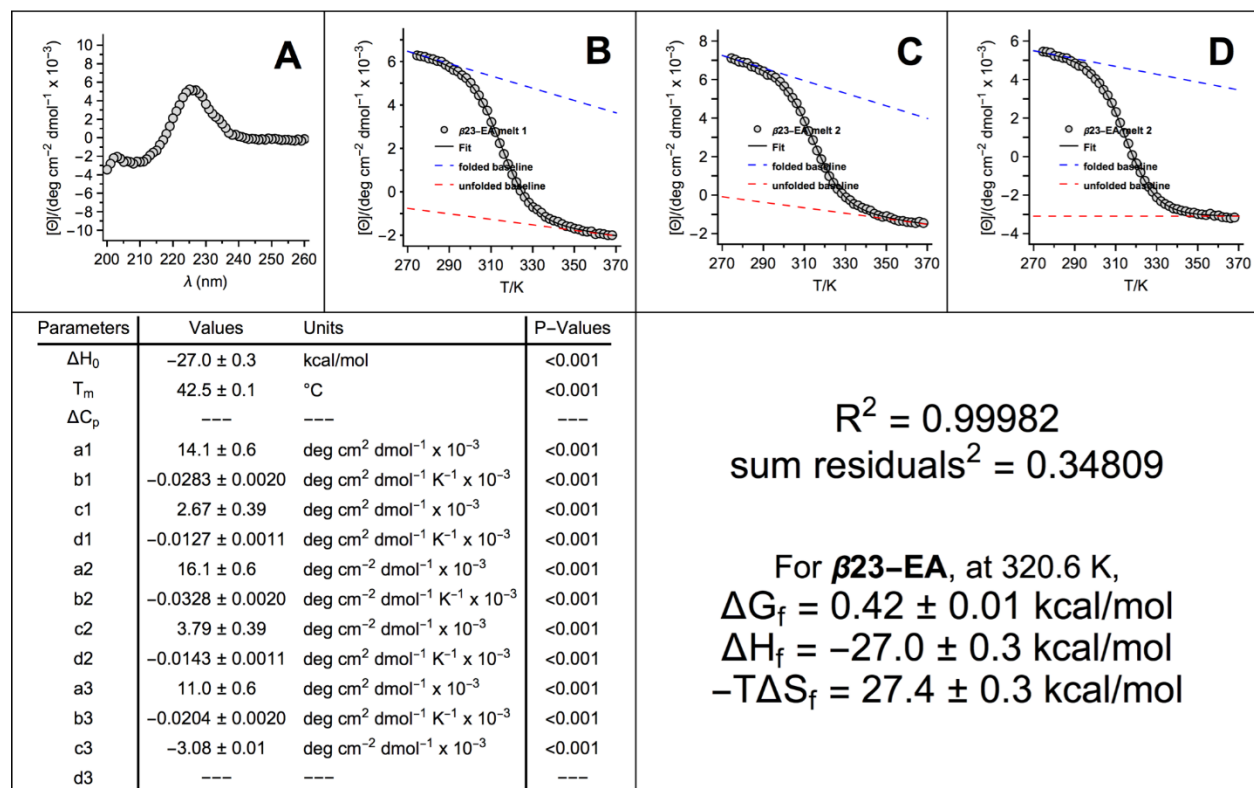


Figure S195. CD data for $\beta 23\text{-EA}$ (SD2006#2).

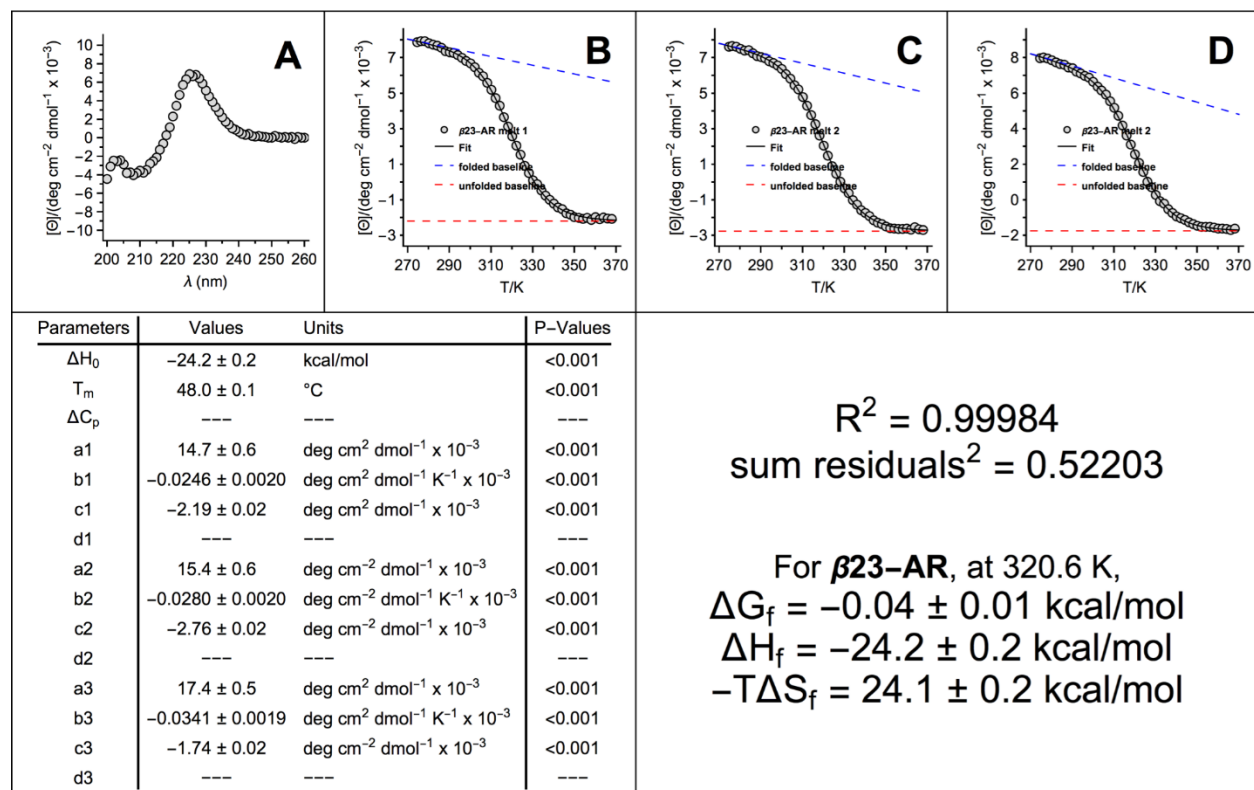


Figure S196. CD data for $\beta 23\text{-AR}$ (SD2018).

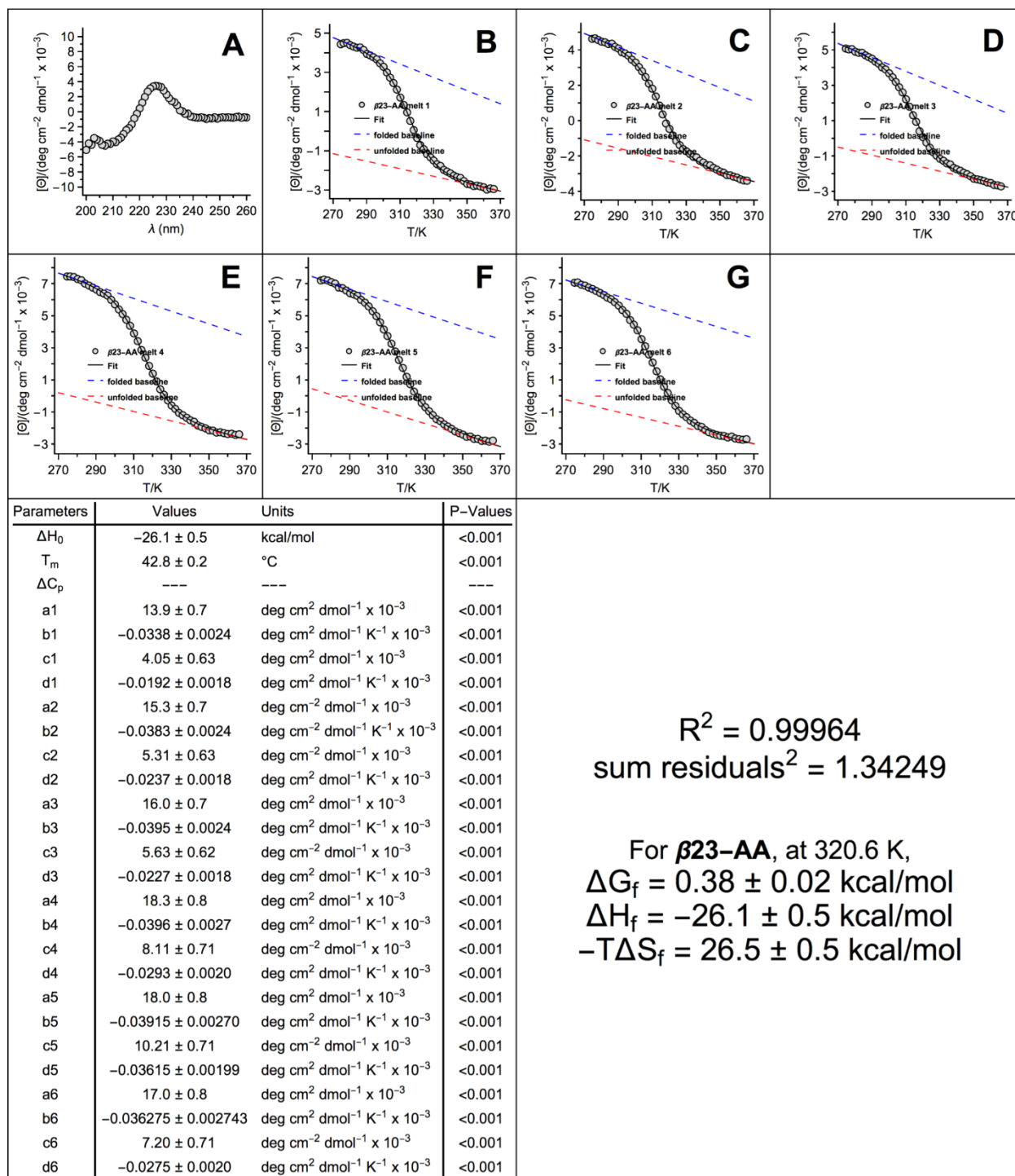


Figure S197. CD data for $\beta 23\text{-AA}$ (SD2006#1).

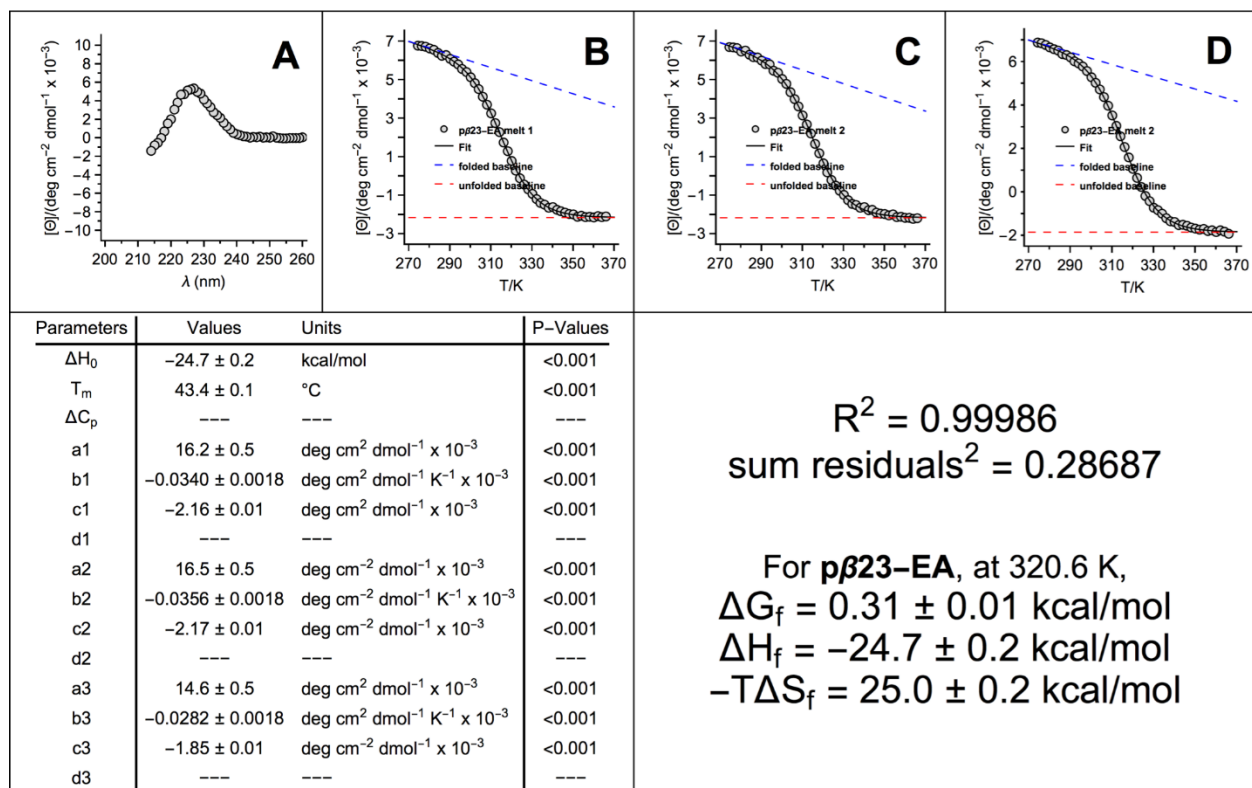


Figure S198. CD data for pβ23-EA (SD2006#2C).

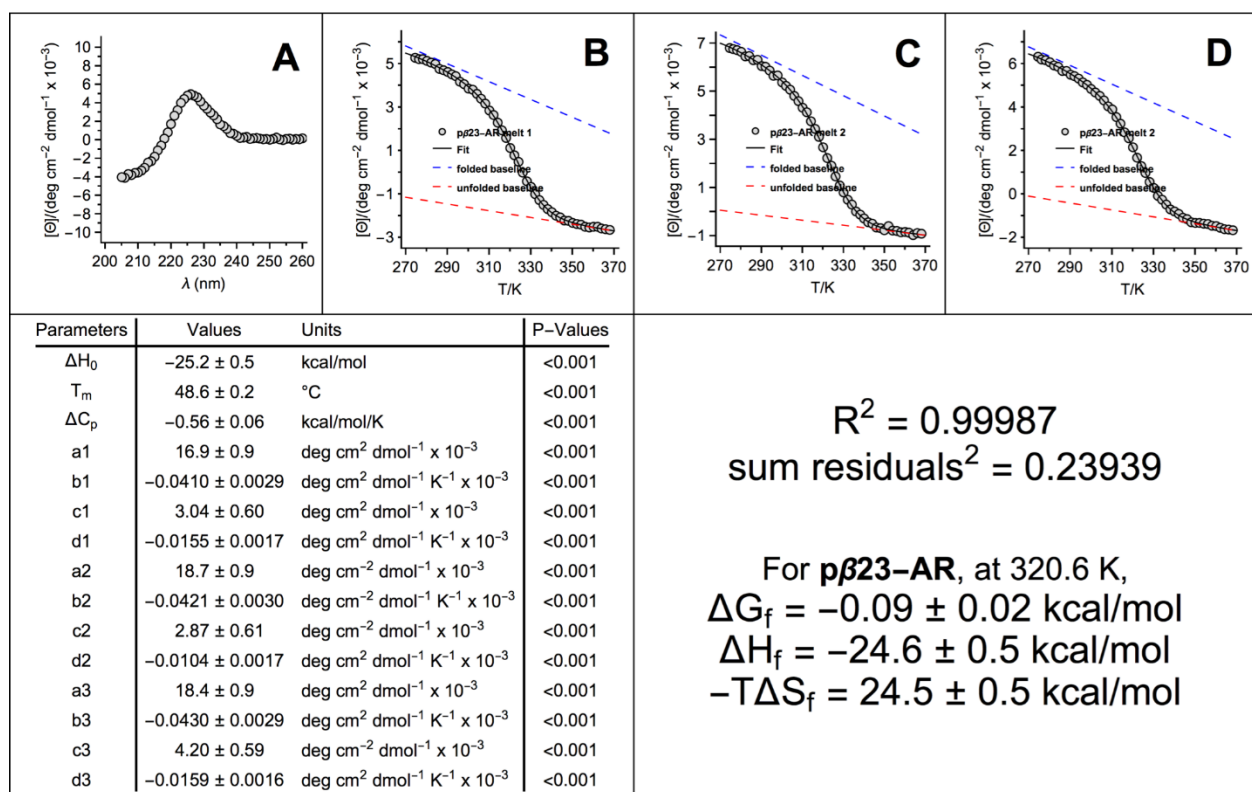


Figure S199. CD data for pβ23-AR (SD2018C).

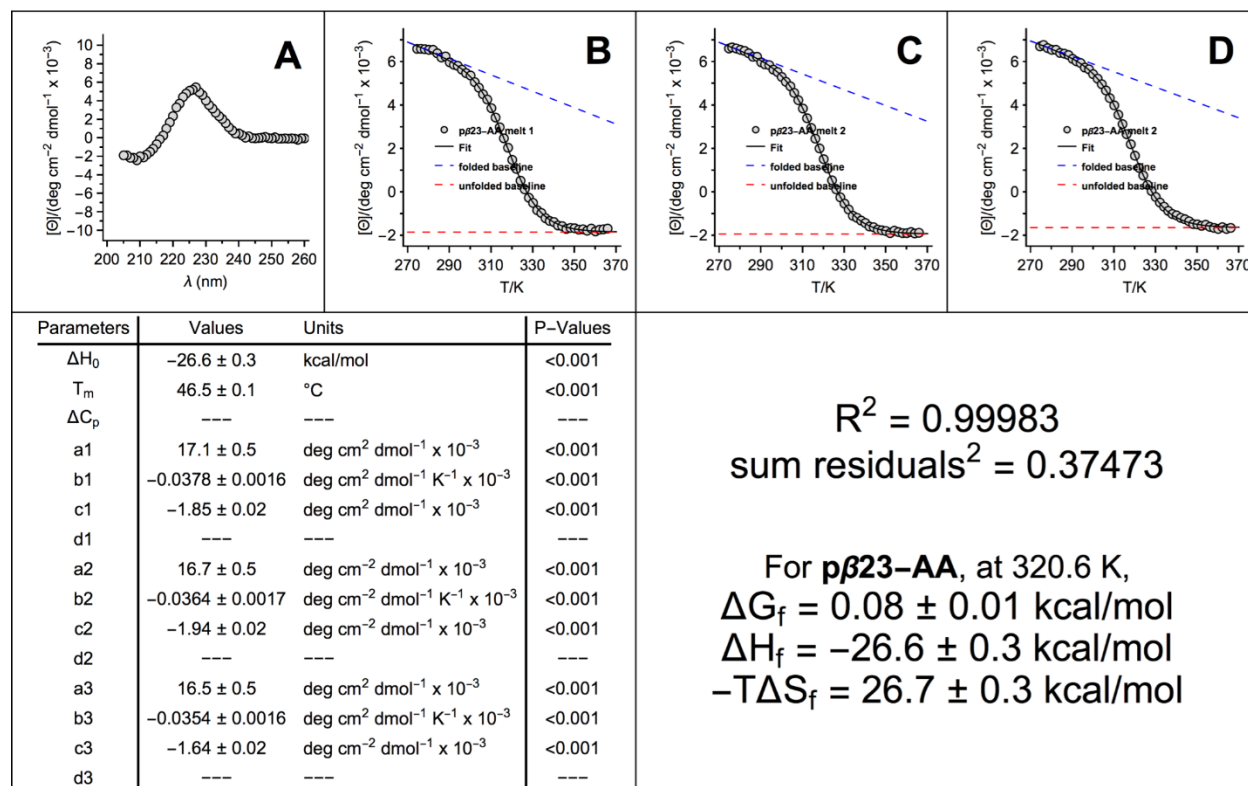


Figure S200. CD spectra for pβ23-AA (SD2006#1C).

Table S5. Triple mutant box analysis of the impact of PEGylation on salt-bridge strength within peptide p3a1-ER.^a

Peptide	Sequence	T_m (°C)	ΔG_f (kcal/mol)	$\Delta\Delta G_f^b$ (kcal/mol)	$\Delta\Delta\Delta G_f^c$ (kcal/mol)	$\Delta\Delta\Delta\Delta G_f^d$ (kcal/mol)
3a1-ER	QV E ALE R KVEALESKVQKLEKKVEALEHGWDGR	77.9	-15.38 ± 0.02		-0.60 ± 0.02	
3a1-EA	QV E ALE A	76.0	-14.90 ± 0.02			
3a1-AR	QV A ALE R	73.0	-14.18 ± 0.02			
3a1-AA	QV A ALE A	76.1	-14.92 ± 0.03			
p3a1-ER	QV E ALE R	74.7	-14.61 ± 0.02	0.38 ± 0.01	-0.56 ± 0.02	0.05 ± 0.03
p3a1-EA	QV E ALE A	73.9	-14.40 ± 0.02	0.17 ± 0.01		
p3a1-AR	QV A ALE R	70.7	-13.68 ± 0.01	0.25 ± 0.01		
p3a1-AA	QV A ALE A	74.7	-14.59 ± 0.02	0.11 ± 0.01		

^aQ = GlnPEG. Variable temperature CD experiments were performed in 20 mM phosphate buffer (pH 7) at 30 μM protein concentration. ΔG_f values are given \pm standard error at the average melting temperature for these variants (347.8 K). ^bImpact of PEGylation on peptide/protein conformational stability. ^cStrength of salt-bridge interaction. ^dImpact of PEGylation on salt-bridge strength.

3. Crystallographic characterization of 2 α 18 and p2 α 18

2 α 18: 2 α 18 was crystalized by vapor diffusion in sitting drops where the well solution contained 0.1 M PCTP (sodium propionate, sodium cacodylate trihydrate, and bis-Tris Propane) and 25% w/v PEG 1500 at pH 4. Each drop contained 0.3 μ L well solution and 0.3 μ L peptide (10 mg/ml in water). Crystals were looped and cryocooled by plunging them into liquid nitrogen prior to data collection. Data were collected at 100 K with a copper rotating anode X-ray source (Bruker FR-591 Dual Source Low Temperature X-ray Diffractometer with CCD Detector).

p2 α 18: 2 α 18 was crystalized by vapor diffusion in sitting drops where the well solution contained 0.1 M sodium malonate dibasic monohydrate and 25% w/v PEG 1500 at pH 4. Each drop contained 0.3 μ L well solution and 0.3 μ L peptide (10 mg/ml in water). Crystals were looped and cryocooled by plunging them into liquid nitrogen prior to data collection. Data were collected at 100 K with a copper rotating anode X-ray source (Bruker FR-591 Dual Source Low Temperature X-ray Diffractometer with CCD Detector).

The data were integrated and scaled using Protium. The molecular replacement and refinement were done in Phenix. Model building was carried out in winCOOT.

Table S6. Crystallographic statistics

Data collection 2α18 unPEGylated	PDB ID: 6O2E	Data collection p2α18 PEGylated	PDB ID: 6O2F
Space Group	I 21 21 21	Space Group	I 21 21 21
Unit cell dimensions (Å)	19.3, 30.0, 107.1; 90, 90, 90	Unit cell dimensions (Å)	19.2, 30.0, 106.8; 90, 90, 90
Resolution (Å)	28.9-1.90	Resolution (Å)	28.9-1.80
Total Observations	15,787	Total Observations	33,482
Unique observations	2,679	Unique observations	3,133
Redundancy	5.9 (1.9)	Redundancy	9.8 (1.75)
Completeness (%)	99.04 (93.02)	Completeness (%)	99.74 (99.66)
I/ σ	9.5 (0.76)	<I/ σ I>	11.9 (0.76)
Rpim	0.067	Rpim	0.047
Refinement		Refinement	
Resolution (Å)	28.9-1.90	Resolution (Å)	28.9-1.80
Rcryst ^b	0.214 (0.298)	Rcryst ^b	0.180 (0.248)
Rfree ^c	0.248 (0.311)	Rfree ^c	0.235 (0.310)
Average B-factor	26.21	Average B-factor	22.48
RMSD: bonds (Å) / angles (°)	0.008 / 1.04	RMSD: bonds (Å) / angles (°)	0.012 / 1.26

4. References:

1. Lawrence, P. B., Billings, W. M., Miller, M. B., Pandey, B. K., Stephens, A. R., Langlois, M. I., and Price, J. L. (2016) Conjugation Strategy Strongly Impacts the Conformational Stability of a PEG-Protein Conjugate, *ACS Chem. Biol.* **11**, 1805–1809.
2. Smith, M. S., Lawrence, E. E. K., Billings, W. M., Larsen, K. S., Bécar, N. A., and Price, J. L. (2017) An Anion- π Interaction Strongly Stabilizes the β -Sheet Protein WW, *ACS Chemical Biology*.
3. Deiters, A., Cropp, T. A., Mukherji, M., Chin, J. W., Anderson, J. C., and Schultz, P. G. (2003) Adding Amino Acids with Novel Reactivity to the Genetic Code of *Saccharomyces Cerevisiae*, *J. Am. Chem. Soc.* **125**, 11782-11783.
4. Lawrence, P. B., Gavrilov, Y., Matthews, S. S., Langlois, M. I., Shental-Bechor, D., Greenblatt, H. M., Pandey, B. K., Smith, M. S., Paxman, R., Torgerson, C. D., Merrell, J. P., Ritz, C. C., Prigozhin, M. B., Levy, Y., and Price, J. L. (2014) Criteria for Selecting PEGylation Sites on Proteins for Higher Thermodynamic and Proteolytic Stability, *J. Am. Chem. Soc.* **136**, 17547-17560.
5. Pandey, B. K., Smith, M. S., Torgerson, C., Lawrence, P. B., Matthews, S. S., Watkins, E., Groves, M. L., Prigozhin, M. B., and Price, J. L. (2013) Impact of Site-Specific PEGylation on the Conformational Stability and Folding Rate of the Pin WW Domain Depends Strongly on PEG Oligomer Length, *Bioconjugate Chem.* **24**, 796-802.
6. Ciani, B., Bjelić, S., Honnappa, S., Jawhari, H., Jaussi, R., Payapilly, A., Jowitt, T., Steinmetz, M. O., and Kammerer, R. A. J. P. o. t. N. A. o. S. (2010) Molecular basis of coiled-coil oligomerization-state specificity, **107**, 19850-19855.
7. Spector, S., Young, P., and Raleigh, D. P. J. B. (1999) Nativelike structure and stability in a truncation mutant of a protein minidomain: the peripheral subunit-binding domain, **38**, 4128-4136.
8. Edelhoch, H. (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins, *Biochemistry* **6**, 1948–1954.