Heterochiral Knottin Protein: Folding and Solution Structure

Electronic Supplementary Information

Surin K. Mong[a], Frank V. Cochran[b], Hongtao Yu[c], Zachary Graziano[c], Yu-Shan Lin[c], Jennifer R. Cochran[b], Bradley L. Pentelute*[a]

[a] S. K. Mong, Prof. Dr. B. L. Pentelute
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA 02139, U.S.A.
E-mail: blp@mit.edu

[b] Dr. F. V. Cochran, Prof. Dr. J. R. Cochran
Department of Bioengineering
Stanford University
450 Serra Mall, Stanford, CA 94305, U.S.A.

[c] Dr. H. Yu, Z. Graziano, Prof. Dr. Y.-S. Lin
Department of Chemistry
Tufts University
62 Talbot Ave., Medford, MA 02155, U.S.A.

1. Table of Contents

2. List of Supplementary Tables and Figures

3. Materials and Instrumentation

3.1. Reagents

Polyethylene glycol resin functionalized with Rink-amide linker (H-Rink Amide-ChemMatrix® HYR) was sourced primarily from PCAS BioMatrix; (L) and (D)-α-Fmoc protected amino acids were sourced primarily from CreoSalus and possessed the following side-chain protection schemes: Arg(Pbf), Asp(OtBu), Asn(Trt), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc), and Tyr(tBu); (L) and (D)-β-Fmoc protected amino acids were sourced primarily from Chem-Impex International with identical side chain protection; 2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) was sourced from Chem-Impex International; N,N-dimethylformamide (DMF), dichloromethane (DCM), acetonitrile (ACN), and diethyl ether (ether) were sourced primarily from EMD Millipore. Diisopropylethylamine (DIEA), piperidine, HPLC-grade trifluoroacetic acid (TFA), LC-MS grade formic acid (FA), 1,2-ethanedithiol (EDT), and triisopropylsilane (TIPS) were sourced primarily from Sigma-Aldrich. Deionized water (H_2O) was filtered and obtained using a Milli-Q® system. Tris(hydroxymethyl)aminomethane (TRIS), sodium phosphate, sodium chloride (NaCl), guanidine hydrochloride (Guan.), calcium chloride ($CaCl_2$) tris(2-carboxyethyl)phosphine (TCEP), cysteine, cystine, reduced and oxidized glutathione, dimethyl sulfoxide (DMSO), sodium azide, sodium hydroxide (NaOH), hydrochloric acid (HCl), trypsin, and proteinase K were sourced from a variety of vendors including Sigma-Aldrich, Amresco, and Hampton Research. Deuterium oxide (D_2O) was sourced from Cambridge Isotope Laboratories.

3.2. Liquid Chromatography and Mass Spectrometry (LC-MS)

Solvent mixtures used throughout both analytical as well as preparative chromatography include the following: [A] = H_2O with 0.1% (v/v) TFA; [B] = ACN with 0.1% (v/v) TFA; [A'] = H_2O with 0.1% (v/v) FA; [B'] = ACN with 0.1% (v/v) FA. Two different liquid chromatography-mass spectrometry (LC-MS) systems were used throughout this work. The two LC-MS setups are as

follows: “System 1” = Agilent 1260 Infinity LC System paired with an Agilent 6520 Accurate-Mass Quadruple Time-of-Flight (Q-TOF) mass analyzer; “System 2” = Agilent 1290 Infinity LC System paired with an Agilent 6550 Accurate-Mass Q-TOF mass analyzer. LC-MS solvents were degassed and typical autosampler injection quantities for peptides were on the order of 250 ng and 25 ng for System 1 and System 2, respectively. LC-MS data was analyzed using Agilent MassHunter software. With these instruments, several different analytical chromatography methods were employed (Table S1-S4) as described below:

Table S1: LC-MS “Method A”. Chromatography setup, source-related parameters, and instrument time table.

“Method A” on LC-MS System 1 using A’/B’ solvents			
LC: Agilent Zorbax® 300SB-C3 (300 Å, 5 µm, 2.1 x 150 mm), 40 °C, 0.8 mL/min			
MS: Positive Polarity, Gas Temp. = 350 °C, Drying Gas = 11 l/min, Nebulizer = 60 psig, Vcap = 4000 V, Fragmentor = 175 V, Skimmer = 65 V, OCT 1 RF Vpp = 750 V			
Timetable, linear gradients:			
LC Time (min)	B’ (%)	MS Time (min)	State (ON/OFF)
0	5	0	OFF
2	5	4	ON
14	37		
15.5	65		

Table S2: LC-MS “Method B”. Chromatography setup, source-related parameters, and instrument time table.

“Method B” on LC-MS System 1 using A’/B’ solvents			
LC: Agilent Zorbax® 300SB-C3 (300 Å, 5 µm, 2.1 x 150 mm), 40 °C, 0.8 mL/min			
MS: Positive Polarity, Gas Temp. = 350 °C, Drying Gas = 11 l/min, Nebulizer = 60 psig, Vcap = 4000 V, Fragmentor = 175 V, Skimmer = 65 V, OCT 1 RF Vpp = 750 V			
Timetable, linear gradients:			
LC Time (min)	B’ (%)	MS Time (min)	State (ON/OFF)
0	1	0	OFF
2	1	4	ON
11	61	12	OFF
12	61		

Table S3: LC-MS "Method C". Chromatography setup, source-related parameters, and instrument time table.

"Method C" on LC-MS System 1 using A'/B' solvents			
LC: Agilent Zorbax® 300SB-C3 (300 Å, 5 µm, 2.1 x 150 mm), 40 °C, 0.8 mL/min			
MS: Positive Polarity, Gas Temp. = 350 °C, Drying Gas = 11 l/min, Nebulizer = 60 psig, Vcap = 4000 V, Fragmentor = 175 V, Skimmer = 65 V, OCT 1 RF Vpp = 750 V			
Timetable, linear gradients:			
LC Time (min)	B' (%)	MS Time (min)	State (ON/OFF)
0	5	0	OFF
4	5	4	ON
17.5	35	17.5	OFF
19.5	70		

Table S4: LC-MS "Method D". Chromatography setup, source-related parameters, and instrument time table.

"Method D" on LC-MS System 2 using A'/B' solvents			
LC: Jupiter® C4 column (300 Å, 5 µm, 1.0 x 150 mm), 40 °C, 0.1 mL/min			
MS: Positive Polarity, Gas Temp. = 250°C, Drying Gas 15 l/min, Nebulizer = 55 psig, VCap = 5000 V, Nozzle Voltage (Expt) = 2000 V, Fragmentor = 365 V, Skimmer = 0 V, OCT 1 RF Vpp = 750 V			
Timetable, linear gradients:			
LC Time (min)	B' (%)	MS Time (min)	State (ON/OFF)
0	1	0	OFF
2	1	4	ON
12	61	12	OFF
16	91		

Peptides were purified using a Waters 600 preparative High Performance Liquid Chromatography (HPLC) system and its associated controller. Peptides were detected using a Varian 320 ProStar Ultraviolet-Visible (UV) Detector and fractionated with a Gilson FC 203B Fraction Collection. Fractions containing the purified and desired compound, as determined by LC-MS, were pooled and lyophilized. Isolated yields were calculated by mass. Different preparative HPLC methods (Table S5-S6) that were employed and are described below:

Table S5: Preparative HPLC "Method E". Chromatography setup and instrument timetable.

"Method E" using A/B solvents		
LC: Agilent Zorbax® 300SB-C3 (300 Å, 5 µm, 9.4 x 250 mm), 19 °C, 5.0 mL/min		
Timetable, linear gradients:		
LC Time (min)	B (%)	
0	5	
5	5	
10	10	
90	50	

Table S6: Preparative HPLC "Method F". Chromatography setup and instrument timetable.

"Method F" using A/B solvents		
LC: Agilent Zorbax® 300SB-C8 (300 Å, 5 µm, 9.4 x 250 mm), 19 °C, 5.0 mL/min		
Timetable, linear gradients:		
LC Time (min)	B (%)	
0	5	
5	5	
10	10	
90	50	

3.3. Nuclear Magnetic Resonance (NMR)

NMR spectra were acquired at the National Magnetic Resonance Facility at Madison using a Bruker Avance III 600 MHz spectrometer equipped with a triple resonance cryogenic probe at 25 °C. Samples were dissolved in buffered H_2O containing 5% (v/v) D_2O for instrument lock.

4. Experimental Methodology

4.1. Chemical Synthesis of Peptides

The sequences for wild type Ecballium elaterium trypsin inhibitor II (WT-EETI-II SP), as well as a couple of integrin-binding EETI-II variants evolved via yeast cell-surface display (EETI-II 2.5F and EETI-II 2.5D) are shown below (Table S7).[1,2] The "constant" core region of the scaffold protein are underlined and the "binding loops" are highlighted in red. 'SP' was appended to the C-terminus of WT EETI-II in order to facilitate efficient folding of the WT EETI-II scaffold.[3]

Table S7: Trypsin (wild type) and integrin binding EETI-II primary sequences. Binding loops underlined.

Protein	Sequence
WT EETI-II SP	GCPRILMRCK QDSDCLAGCV CGPNGFCGSP
EETI-II 2.5F	GCPRPRGDNP PLTCKQDSDC LAGCVCGPNG FCG
EETI-II 2.5D	GCPQGRGDWA PTSCKQDSDC LAGCVCGPNG FCG

Unless otherwise noted, all peptides were synthesized using Fmoc-based solid-phase peptide synthesis (SPPS) protocols detailed below. Both (L) and (D)-EETI-II constant regions for eventual EETI-II 2.5F and 2.5D analogs were synthesized in batch on 4.0 g of 0.45 mmol/g loading H-Rink Amide-ChemMatrix® HYR using ChemGlass GL-25 100 mL peptide synthesis vessels. Resins were pre-swelled in DCM and then DMF. For every constant region coupling step, 7.2 mmol (4 equiv.) of each Fmoc-protected amino acid was dissolved in 18 mL of 0.38 M HATU in DMF stock (3.8 equiv.); to this mixture, 1.88 mL (10.8 mmol, 1.5 equiv.) of DIEA was added and the resulting solution was mixed vigorously and subsequently added to a drained resin bed. After allowing a coupling step to proceed for 20 min, the coupling solution was washed from the resin bed using a gravity-driven flow wash of several 20 mL DMF additions totaling approximately 100-150 mL of DMF. Deprotection was then achieved by first washing the resin, in gravity flow, with approximately 20 mL of 20% (v/v) piperidine in DMF, and then treating

the resin with 20 mL of the same deprotection solution for 2 x 5 min. The deprotection solution was washed from the resin as previously described for the post-coupling wash step before proceeding to couple the next amino acid.

(L)-WT EETI-II SP constant region was synthesized using rapid-flow based peptide synthesis instrumentation and protocols previously reported.[4,5] Briefly, 0.160 g of H-Rink Amide-ChemMatrix® HYR were loaded into the flow reaction vessel. Synthesis was achieved using a synthetic cycle of 40 seconds nominal coupling, 20 seconds DMF wash, 20 seconds deprotection with 20% (v/v) piperidine in DMF, 60 seconds DMF wash. Each coupling step involved delivery of 1 mmol of Fmoc-protected amino acid dissolved in 2.5 mL of 0.38 M HATU in DMF (0.95 equiv.) and activated with 190µL or 500 µL of DIEA for Cys or all other amino acids, respectively.

After assembly of EETI-II constant regions, synthesis of the binding loop regions was then accomplished using batch synthesis over a syringe manifold. Each EETI-II analog was synthesized on an amount of peptidyl-loaded constant region resin that approximately equated 60 mg (0.027 mmol) of starting resin. Coupling and deprotection conditions were similar but scaled down relative to those described above for batch synthesis of the constant regions; for each coupling, 0.5 mmol of Fmoc-protected amino acid was used (18.5 equiv.).

After completing synthesis of the peptides, resins were washed extensively with DCM and then dried completely using the syringe manifold vacuum. Each peptide was then transferred to an individual Falcon 50 mL tube for global deprotection and cleavage from the solid support. This was achieved for each peptide by using 5 mL of a cleavage cocktail comprised of the following: 94% (v/v) TFA / 2.5% (v/v) H_2O / 2.5% (v/v) EDT / 1.0% (v/v) TIPS. Cleavage was allowed to proceed for 2 hrs at room temperature. After completion, the resin-TFA slurries were evaporated to dryness using nitrogen and then the peptides were triturated using approximately

20 mL ether. The peptides and resin were sedimented in a table-top centrifuge and the ether decanted. Residual solids were again thoroughly washed 2 more times using ether and then allowed to dry in a fume hood after the final decantation. Once dry, the peptides were taken up in approximately 10 mL of 50% (v/v) / 50% (v/v) A/B, filtered using a 0.22 μm syringe filter, frozen using liquid nitrogen, and then lyophilized to yield the crude peptide as a white powder. Crude peptides were analyzed and confirmed by LC-MS using Method B and yields were quantified by mass.

4.2. Purification of Peptides

Crude peptides were initially dissolved in 2.5 mL of 6 M Guan. containing 30 mM DTT. Peptides were left for 20 mins and then filtered using a 0.22 μm syringe. The filtered solution was then loaded on to a preparative HPLC column for purification and fractionation. Fractions were analyzed by LC-MS using Method B, pooled, and then lyophilized to a powder. Purified LC-MS chromatograms and purification yields for all EETI-II analog are summarized below (Table S8) as well as individually each peptide (Figure S1-S28) in the subsequent section.

4.3. Oxidative Folding of EETI-II Analogs

A number of oxidative folding solutions were investigated in order to identify one that would enable “efficient” oxidative folding of (L)-EETI-II 2.5F. Efficient folding was defined as nearly quantitative conversion of reduced (L)-EETI-II 2.5F to a distinct LC-MS peak demonstrating both loss of 6 Da (formation of 3 disulfide bonds) and a dramatic loss in retention time by RP-HPLC (burial of hydrophobic patches), relative to the starting material. The oxidative folding solutions that were initially investigated included solutions containing reduced and oxidized glutathione, reduced and oxidized cysteine, DMSO, and DMSO with reduced glutathione. Of these preliminary experiments, the use of a cysteine/cystine redox pair yielded the most efficient folding results. Creating a matrix of pH versus cysteine concentration resulted in the optimized folding conditions described further below. Non-optimal folding conditions often

resulted in a heterogenous product (presumably misfolded) as well as peptides that exhibited thiolate-ligand adducts (i.e. addition of 1 or 2 glutathione or cysteine moieties). (L)-EETI-II 2.5F that was oxidatively folded under optimal conditions was purified and exhibited binding activity towards U87MG cells in agreement with what has been previously described in literature (data not shown).[2]

The following buffers were prepared fresh and used in the optimized and analytical scale oxidative folding of all EETI-II analogs: “initialization buffer” (50 mM TRIS buffer, pH 7.8, 5.7 M Guan., 5 mM TCEP), “redox buffer” (50 mM TRIS buffer, pH 7.8, saturated or approximately 0.47 mM cystine, 1 mM cysteine). For a given EETI-II analog, oxidative folding was performed by first dissolving the purified and lyophilized peptide to 5.0 mg/mL in initialization buffer; the initialization buffer was used in order to ensure that all peptides were being oxidatively folded from a reduced and denatured state. After approximately 20 mins, a portion of the peptide in initialization buffer was pipetted directly into and rapidly mixed into redox buffer in order to initiate oxidative folding. The final concentration of EETI-II analog in redox buffer was 0.25 mg/mL (approximately 0.075 mM). Oxidative folding was monitored by using LC-MS to analyze samples quenched at various time points. Quenching was achieved by removing an aliquot of the sample and diluting it 3-fold into H_2O containing 0.2% (v/v) TFA. Quenched samples were analyzed by LC-MS using Method C to inject and analyze 4 µL of sample (330 ng peptide). Time zero samples were quenched directly from peptide dissolved in initialization buffer to achieve the same final peptide concentration for LC-MS analysis. Oxidative folding results are summarized below (Table S8) and LC-MS chromatograms and for each individual EETI-II analog are shown in the subsequent section (Figure S1-S28).

Table S8: Summary of EETI-II analogs synthesized, purified, and analytically folded. Disulfide (-SS-) Retention times (r.t.) correspond to the major products in the LC-MS chromatograms at time 0 min, 2 min, and 16 hrs. Abbreviated naming convention for EETI-II analogs: $^{[\text{Core Chirality}],[\text{Loop Chirality}]}_{(\text{Loop Chirality stretch})}$[Base Sequence](Substitutions).

Summary of analytical oxidative folding experiments, EETI-II analogs

Purpose	EETI-II Analog	Crude Peptide Yield (mg)	Crude Purification Yield (%) [mg isolated/mg loaded]	Analytical Folding, t = 0 min, Principle LCMS Product Monoisotopic Mass (Da)	Analytical Folding, t = 2 min, Principle LCMS Product Monoisotopic Mass (Da)	Analytical Folding, t = 16 hrs, Principle LCMS Product Monoisotopic Mass (Da)	0 -SS- r.t. (min)	2 -SS- r.t. (min)	3 -SS- r.t. (min)	Δ[3 - 0] r.t. (min)	Δ[2 - 0] r.t. (min)
wild type EETI-II	^{L,L}WT-SP	33.5	12.1 [1.3/10.7]	3085.34	3081.31	3079.29	13.28	11.20	10.02	-3.26	-2.08
	^{L,L}WT-SP(P3β_A,R8β_A)	32.8	21.4 [3.7/17.3]	2974.27	2970.23	2968.21	13.57	12.74	11.45	-2.12	-0.83
	$^{L,L(3\text{-}8)}$WT-SP	32.1	17.6 [2.8/15.9]	3085.34	3081.31	3319.32	13.00	12.03	x	x	-0.97
	$^{L,D(3\text{-}8)}$WT-SP(P3β_A,R8β_A)	28.2	23.6 [3.9/16.5]	2974.25	2970.22	2968.21	13.48	12.64	11.89	-1.59	-0.84
different, longer loop	^{L,L}2.5D	45.7	25.4 [5.3/20.9]	3287.32	3283.29	3281.28	13.10	12.00	10.46	-2.64	-1.10
	^{L,L}2.5D(P3β_A,S13β_A)	38.5	26.5 [6.1/23.0]	3245.31	3241.28	3239.27	12.97	12.50	11.13	-1.84	-0.47
	$^{L,D(3\text{-}13)}$2.5D	45.5	51.6 [11.1/21.5]	3287.32	3283.29	3281.29	13.25	12.82	12.43	-0.82	-0.43
	$^{L,D(3\text{-}13)}$2.5D(P3β_A,S13β_A)	42.8	35.8 [8.2/22.9]	3245.31	3241.28	3239.27	12.99	11.74	10.17	-2.82	-1.25
different, longer loop	^{L,L}2.5F	54.8	35.4 [19.4/54.8]	3335.43	3331.38	3329.43	12.32	11.56	9.59	-2.73	-0.75
	^{L,L}2.5F(P3β_A,T13β_A)	59.2	39.9 [23.6/59.2]	3279.42	3275.46	3273.39	12.20	11.75	10.24	-1.96	-0.44
	$^{L,D(3\text{-}13)}$2.5F	46.1	34.3 [15.8/46.1]	3335.37	3331.44	3329.31	12.39	11.86	12.45	0.06	-0.53
	$^{L,D(3\text{-}13)}$2.5F(P3β_A,T13β_A)	61.4	37.1 [22.8/61.4]	3279.48	3275.37	3273.39	12.19	11.70	10.38	-1.81	-0.49
positioning of β_A	^{L,L}2.5F(P3β_A,P10β_A)	56.2	23.5 [3.9/16.6]	3283.35	3279.33	3277.32	12.29	11.28	9.50	-2.79	-1.01
	$^{L,D(3\text{-}10)}$2.5F	47.4	32.6 [7.8/23.9]	3335.44	3331.41	3329.39	12.53	11.44	10.87	-1.66	-1.09
	$^{L,D(3\text{-}10)}$2.5F(P3β_A,P10β_A)	48.7	30.5 [7.4/24.3]	3283.41	3279.38	3277.37	12.32	11.32	9.67	-2.65	-1.00
enantiomers	^{D,D}2.5F	58.9	44.1 [26.0/58.9]	3335.44	3331.40	3329.39	12.41	11.70	9.67	-2.75	-0.71
	$^{D,L(3\text{-}13)}$2.5F(p3β_A,t13β_A)	62.0	44.2 [27.4/62.0]	3279.41	3275.38	3273.37	12.26	11.82	10.46	-1.80	-0.44
	$^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A)	55.8	21.5 [3.7/17.2]	3283.45	3279.34	3277.33	12.28	11.29	9.66	-2.62	-0.99
alternative transitions, scaffold to loop	$^{L,D(3\text{-}13)}$2.5F(P3β_A)	43.0	36.6 [6.8/18.6]	3309.43	3305.39	3303.38	12.25	11.83	10.92	-1.32	-0.42
	$^{L,D(3\text{-}13)}$2.5F(T13β_A)	55.0	40.4 [22.2/55.0]	3305.43	3301.40	3299.39	12.39	11.93	11.21	-1.18	-0.46
	$^{L,D(3\text{-}13)}$2.5F(P3G,T13G)	48.8	31.9 [4.4/13.8]	3251.36	3247.31	3245.31	12.23	11.73	10.13	-2.10	-0.50
	$^{L,D(3\text{-}13)}$2.5F(P3β_A,T13$^{D}\beta_{hP}$)	48.8	42.3 [10.5/24.8]	3319.50	3315.38	3313.40	12.67	12.18	10.90	-1.77	-0.49
	$^{L,D(3\text{-}12)}$2.5F(P3β_A,T13$^{L}\beta_{hE}$)	47.0	27.4 [4.5/16.4]	3351.22	3347.37	3345.35	12.14	11.90	10.36	-1.78	-0.24
	$^{L,D(3\text{-}10)}$2.5F(P3β_A)	52.4	27.2 [4.6/16.9]	3309.37	3305.35	3303.34	12.45	11.27	10.27	-2.18	-1.18
	$^{L,D(3\text{-}10)}$2.5F(P10β_A)	57.7	16.6 [2.9/17.5]	3309.38	3305.35	3303.34	12.42	11.41	10.30	-2.12	-1.01
	$^{L,D(3\text{-}10)}$2.5F(P3G,P10G)	86.8	21.8 [3.8/17.4]	3255.34	3251.31	3249.30	12.28	11.12	9.32	-2.96	-1.16
	$^{L,D(3\text{-}10)}$2.5F(P3β_A,P10$^{D}\beta_{hP}$)	41.4	21.3 [3.8/17.8]	3323.40	3319.37	3317.36	12.43	11.57	9.84	-2.59	-0.86
	$^{L,D(3\text{-}9)}$2.5F(P3β_A,P10$^{L}\beta_{hF}$)	55.9	19.6 [3.2/16.3]	3373.42	3369.38	3367.37	13.46	13.01	11.02	-2.44	-0.45

4.4. NMR Sample Preparation, Data Acquisition, and Analysis

The folded heterochiral EETI-II analog used in protein NMR spectroscopy experiments possessed the following primary sequence (uppercase denotes an achiral or (L)-α-amino acid, lowercase denotes a (D)-α-amino acid, and X denotes β-alanine):

GcXRPRGDNX pltcsqdsdc laGcvcGpnG fcG

Compared to native EETI-II 2.5F, Position 15 was changed from Lys to Ser in order to enable direct comparison to an undisclosed solution structure. The reduced polypeptide precursor was synthesized as a C-terminal amide using H-Rink Amide-ChemMatrix® HYR using rapid-flow based peptide synthesis and protocols described above. Crude peptide was subjected to optimized folding conditions, acid quenched after approximately 24 hours, and purified in order to yield the folded heterochiral EETI-II 2.5F analog (Figure S29). This purified analog was dissolved in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2% (w/v) sodium azide to create 600 µL of an approximately 3 mM protein NMR sample. This sample was filtered through a 0.22 µm membrane and dynamic light scattering was employed to ensure that there was no protein aggregation. The sample was then sent to NMRFAM for analysis using a 600 MHz spectrometer at 25 °C. A number of spectra were acquired in order to assist proton chemical shift assignments (Figure S30-S36).

NMR spectra were analyzed and chemical shifts were assigned using NMRFAM-SPARKY software (Table S8-S10).[6] After identifying and assembling Nuclear Overhauser Effect (NOE) cross-peaks, CYANA software was used to assemble initial ensemble of structures based using an unassigned list of NOE distance restraints as well as disulfide-bond distance restraints (observed from LC-MS experiments).[7] Custom library entries were created for (D)-amino acids as well as β-alanine. CYANA calculations were performed for all possible permutations of disulfide pairing and the arrangement of Cys2-Cys24, Cys14-Cys26, Cys20-32

was in greatest agreement with the experimentally observed NOE cross-peaks. Out of the 100 initial structures calculated for this disulfide arrangement, the 20 with the lowest CYANA target function scores were carried forward for explicit solvent refinement using YASARA software.[8] Inspection of the YASARA-refined ensemble revealed several potential hydrogen bonds that were present in greater than 60% of the structures. 5 potential hydrogen bond restraints that had been previously been observed for an EETI-II variant were used as restraints in a final structure calculation and refinement.[3] The finalized structure ensemble was then evaluated (Table S11).

Table S9: Assigned ^{15}N and ^{13}C resonance chemical shifts (ppm). Acquired on a 600 MHz spectrometer at 25 °C. CYANA atomic nomenclature.

			Resonance chemical shift (ppm), CYANA nomenclature												
Position	Chirality	Amino Acid	N	ND2	NE2	CA	CB	CG	CG1	CG2	CD	CD1	CD2	CE1	CZ
1	-	Gly	-	-	-	40.71	-	-	-	-	-	-	-	-	-
2	D	Cys	122.9	-	-	51.51	39.26	-	-	-	-	-	-	-	-
3	-	$^{\beta}$Ala	122.5	-	-	34.75	36.19	-	-	-	-	-	-	-	-
4	L	Arg	126.6	-	-	50.84	27.52	24.31	-	-	40.70	-	-	-	-
5	L	Pro	-	-	-	60.59	29.42	24.59	-	-	47.95	-	-	-	-
6	L	Arg	120.8	-	-	53.89	28.07	24.32	-	-	40.66	-	-	-	-
7	-	Gly	110.0	-	-	42.65	-	-	-	-	-	-	-	-	-
8	L	Asp	118.4	-	-	53.26	36.34	-	-	-	-	-	-	-	-
9	L	Asn	106.7	113.00	-	51.84	38.46	-	-	-	-	-	-	-	-
10	-	$^{\beta}$Ala	117.8	-	-	33.31	35.57	-	-	-	-	-	-	-	-
11	D	Pro	-	-	-	59.94	29.83	24.38	-	-	47.94	-	-	-	-
12	D	Leu	123.1	-	-	51.85	40.55	24.39	-	-	-	21.29	22.32	-	-
13	D	Thr	114.4	-	-	58.51	67.75	-	-	19.46	-	-	-	-	-
14	D	Cys	115.3	-	-	51.31	45.06	-	-	-	-	-	-	-	-
15	D	Ser	114.1	-	-	56.23	62.44	-	-	-	-	-	-	-	-
16	D	Gln	114.1	-	112.50	51.25	29.10	28.64	-	-	-	-	-	-	-
17	D	Asp	124.1	-	-	56.33	36.87	-	-	-	-	-	-	-	-
18	D	Ser	112.1	-	-	57.57	59.56	-	-	-	-	-	-	-	-
19	D	Asp	120.6	-	-	53.27	39.26	-	-	-	-	-	-	-	-
20	D	Cys	117.4	-	-	49.92	37.48	-	-	-	-	-	-	-	-
21	D	Leu	120.5	-	-	52.39	40.02	24.35	-	-	-	20.25	22.61	-	-
22	D	Ala	121.4	-	-	51.16	15.37	-	-	-	-	-	-	-	-
23	-	Gly	110.1	-	-	42.33	-	-	-	-	-	-	-	-	-
24	D	Cys	119.4	-	-	51.94	41.37	-	-	-	-	-	-	-	-
25	D	Val	115.0	-	-	57.12	31.68	-	16.03	18.82	-	-	-	-	-
26	D	Cys	123.9	-	-	52.63	35.85	-	-	-	-	-	-	-	-
27	-	Gly	117.4	-	-	42.55	-	-	-	-	-	-	-	-	-
28	D	Pro	-	-	-	61.80	28.80	-	-	-	47.23	-	-	-	-
29	D	Asn	118.4	111.80	-	49.72	35.06	-	-	-	-	-	-	-	-
30	-	Gly	-	-	-	43.94	-	-	-	-	-	-	-	-	-
31	D	Phe	114.5	-	-	53.23	39.40	-	-	-	-	129.30	-	128.60	127.00
32	D	Cys	123.7	-	-	53.33	38.89	-	-	-	-	-	-	-	-
33	-	Gly	111.6	-	-	42.94	-	-	-	-	-	-	-	-	-

Table S10: Assigned ^{1}H resonance chemical shift (ppm). Acquired on a 600 MHz spectrometer at 25 °C. CYANA atomic nomenclature.

			Resonance chemical shift (ppm), CYANA nomenclature																							
Position	Chirality	Amino Acid	H	HA	HA2	HA3	QA	HB	HB2	HB3	QB	HG	QG	QG1	QG2	HD2	HD3	HD21	HD22	QD	QD1	QD2	HE	HE21	HE22	QE	HZ
1	-	Gly	-	-	-	-	3.77	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	D	Cys	8.71	4.50	-	-	-	-	-	-	2.93	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	$^{\beta}$Ala	8.34	-	-	-	2.44	-	3.31	3.45	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	L	Arg	8.21	4.51	-	-	-	-	1.60	1.76	-	-	1.59	-	-	-	-	-	-	3.12	-	-	7.18	-	-	-	-
5	L	Pro	-	4.33	-	-	-	-	1.83	2.23	-	-	1.95	-	-	3.56	3.75	-	-	-	-	-	-	-	-	-	-
6	L	Arg	8.38	4.19	-	-	-	-	1.72	1.80	-	-	1.58	-	-	-	-	-	-	3.12	-	-	7.23	-	-	-	-
7	-	Gly	8.37	-	3.86	3.91	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	L	Asp	8.25	4.51	-	-	-	-	2.64	2.72	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	L	Asn	8.12	4.49	-	-	-	-	2.59	2.52	-	-	-	-	-	-	-	6.81	7.50	-	-	-	-	-	-	-	-
10	-	$^{\beta}$Ala	7.88	-	2.48	2.59	-	-	-	-	3.34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11	D	Pro	-	4.31	-	-	-	-	1.78	2.11	-	-	1.74	-	-	-	-	-	-	3.45	-	-	-	-	-	-	-
12	D	Leu	8.52	4.49	-	-	-	-	1.58	1.62	-	1.72	-	-	-	-	-	-	-	-	0.91	0.94	-	-	-	-	-
13	D	Thr	7.87	4.56	-	-	-	4.10	-	-	-	-	-	-	1.01	-	-	-	-	-	-	-	-	-	-	-	-
14	D	Cys	8.15	4.84	-	-	-	-	3.10	3.14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15	D	Ser	9.48	4.50	-	-	-	-	3.75	3.78	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16	D	Gln	8.00	4.67	-	-	-	-	1.95	2.12	-	-	2.19	-	-	-	-	-	-	-	-	-	-	6.64	7.55	-	-
17	D	Asp	9.10	4.08	-	-	-	-	2.60	2.73	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
18	D	Ser	8.28	4.17	-	-	-	-	3.74	4.02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
19	D	Asp	7.64	4.53	-	-	-	-	-	-	2.93	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20	D	Cys	7.83	4.78	-	-	-	-	2.79	3.05	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21	D	Leu	8.11	4.18	-	-	-	-	1.49	1.60	-	1.68	-	-	-	-	-	-	-	-	0.83	0.85	-	-	-	-	-
22	D	Ala	8.09	4.00	-	-	-	-	-	-	1.29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
23	-	Gly	8.56	-	3.53	4.26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24	D	Cys	8.19	4.78	-	-	-	-	2.99	3.30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	D	Val	8.51	4.25	-	-	-	2.06	-	-	-	-	-	0.70	0.75	-	-	-	-	-	-	-	-	-	-	-	-
26	D	Cys	9.31	4.75	-	-	-	-	2.45	2.76	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27	-	Gly	8.23	-	4.02	4.33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	D	Pro	-	4.32	-	-	-	-	1.95	2.27	-	-	2.00	-	-	3.63	3.68	-	-	-	-	-	-	-	-	-	-
29	D	Asn	8.00	4.59	-	-	-	-	2.91	3.02	-	-	-	-	-	-	-	6.93	7.64	-	-	-	-	-	-	-	-
30	-	Gly	8.13	-	3.59	3.80	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
31	D	Phe	6.97	5.28	-	-	-	-	2.61	2.94	-	-	-	-	-	-	-	-	-	6.94	-	-	-	-	-	7.18	7.18
32	D	Cys	8.97	4.96	-	-	-	-	-	-	2.91	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
33	-	Gly	9.25	-	4.02	4.16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table S11: CYANA input data and structural statistics.

CYANA input data and structural statistics	
Distance Constraints	
Total NOEs:	333
with short-range assignment |i-j|<=1:	274
with medium-range assignment 1<|i-j|<5:	25
with long-range assignment |i-j|>=5:	34
Hydrogen bonds:	5
Disulfide bonds:	3
Structural Statistics	
Total number of structures calculated:	100
Total number of structures in final ensemble:	20
Final ensemble average CYANA function value ($Å^2$)	0.35 ± 0.31
Violations observed:	2
Overall RMSD	
average backbone RMSD (Å):	1.82 ± 0.26
average heavy atom RMSD (Å):	2.50 ± 0.33
Loop (residue 1-10) RMSD	
average backbone RMSD (Å):	1.99 ± 0.34
average heavy atom RMSD (Å):	2.99 ± 0.44
Core (residue 11-33) RMSD	
average backbone RMSD (Å):	0.69 ± 0.15
average heavy atom RMSD (Å):	1.08 ± 0.15

4.5. Molecular Dynamics

The molecular dynamics (MD) simulations were performed using Gromacs 4.6.7[9] with PLUMED 2.1 plugin.[10] The proteins were modeled using a recently developed residue-specific force field (RSFF1).[11,12] Because the RSFF1 parameters for *β*-Ala are currently unavailable, the OPLS-AA/L parameters were used for this amino acid.[13] The solvent was represented by the TIP4P-Ew water model.[14] The L-scaffold was built from the EETI-II PDB structure (PDB ID: 1H9H).[15] The D-loops were added by using the UCSF Chimera modeling package.[16] For the simulation of the $^{D,L(3\text{-}10)}$2.5F(P3β,P10β,K15S) protein (manuscript Figure 6, [5]), the NMR structure was used. The NH_3^+ group was used for the N-terminus and the C-terminus was capped by an amide. The disulfide bond between Cys2 and Cys24 was removed. For each system, the initial structure was first minimized for 1000 steps and then solved in a cubic water box. The size of the box was chosen such that the minimum distance between the protein and the box edges is 15 Å. For the L-scaffold/D-loop systems, one chloride ion was also added to neutralize the net charge. The solvated system was further minimized for 5000 steps using the steepest descent algorithm. With the heavy atoms of residues 14–33 restrained, each system was heated from 5 K to 600 K within 50 ps and relaxed for 200 ps. The system was then cooled down to 300 K with a cooling rate of 6 K/ps and equilibrated for 100 ps at 300 K. The equilibrated system was used for subsequent steered molecular dynamics (SMD) simulations,[17] the distance between the two S_γ atoms of Cys2 and Cys24, *r*, was used as a reaction coordinate in the SMD simulations. For each system, 100 independent SMD runs starting from different initial velocities were performed. In each run, the system was first pulled to the initial state (*r* = 20 Å) within 100 ps using a force constant of 10 kcal/mol/Å^2. After an additional 100 ps equilibration at the initial state, the production SMD was implemented to pull the system from the initial state (*r* = 20 Å) to the final state (*r* = 3 Å) with a pulling speed of 10 Å/ns. The potential of mean force (PMF) as a function of *r* was evaluated by using the Jarzynski's equality.[18]

All the SMD simulations were performed in the NPT (isothermal-isobaric) ensemble at 300 K/1 bar. The temperature was controlled using the Nosé-Hoover thermostat[19,20] with a time constant of 1.0 ps. To alleviate the "hot-solvent/cold-solute" artifact,[21,22] two separate thermostats were applied to both the protein and the solvent molecules. The pressure of the system was regulated using an isotropic Parrinello-Rahman algorithm[23] with a coupling time of 2.0 ps and a compressibility of 4.5×10^{-5} bar^{-1}. All bonds were constrained with the LINCS algorithm[24] to enable the use of a 2 fs time step with the leap-frog algorithm.[25] The non-bonded interactions (Lennard-Jones and Columbic) were truncated at 8 Å. Long-range Columbic interactions beyond the cut-off distance were treated using the Particle Mesh Ewald (PME) summation method.[26] A long-range analytic dispersion correction was applied to both the energy and pressure to account for the truncation of Lennard-Jones interactions.[27]

4.6. Proteolysis Stability Assays

The four EETI-II analogs used in proteolysis experiments were initially folded and purified from crude starting material as described earlier in NMR sample preparation. These purified peptides were dissolved into neat H_2O in order to create 5.0 mg/mL stock solutions. A "proteolysis buffer" (50 mM TRIS, pH 7.4, 150 mM NaCl, 10 mM $CaCl_2$) was used throughout both trypsin and proteinase K experiments.

For proteolysis experiments involving trypsin, lyophilized trypsin was dissolved to 1.0 mg/mL in neat H_2O. Using a multi-channel pipette, 6 µL of stock solution for each EETI-II analog and was added to a strip of polymerase chain tubes containing 94 µL of proteolysis buffer. To these, 0.67 µL of the trypsin stock solutions were added and the resulting solutions mixed and incubated at 37 °C in order to initiate degradation before. Various trypsin experimental time points were obtained by quenching 10 µL of the mixture in 20 µL of solvent A/B, 30/70. The experimental setup for measuring the extent of degradation by proteinase K was similar in setup for trypsin with the exception that 2 µL of a 10 mg/mL proteinase K stock was used to initiate proteolysis.

Additionally, proteinase K samples were quenched by adding 10 µL the proteolysis mixture to 10 µL of solvent B. Trypsin and proteinase K samples were analyzed by LC-MS (Figure S37-S40). Proteinase K data workup involved measuring intact peptide remaining by quantifying the extracted ion current for the starting peptide. A standard curve was created to ensure the linearity of measuring the total ion current and extracted ion current over the range experimentally relevant range of peptide analyzed (data not shown).

5. LC-MS Chromatograms and NMR Spectra

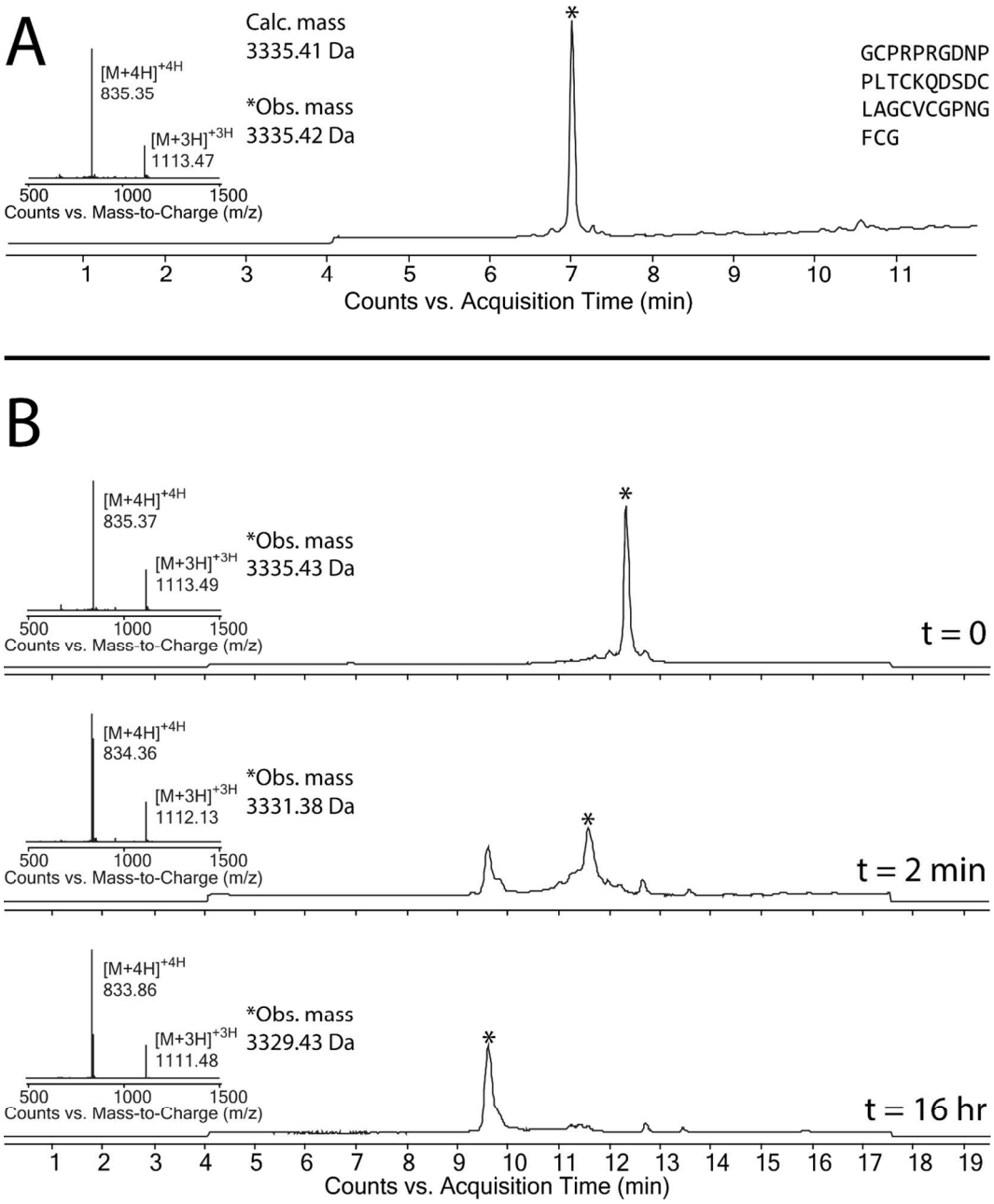

Figure S1: Purification and oxidative folding of ^{L,L}2.5F. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (35.4 % yield, 19.4 mg isolated, 54.8 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

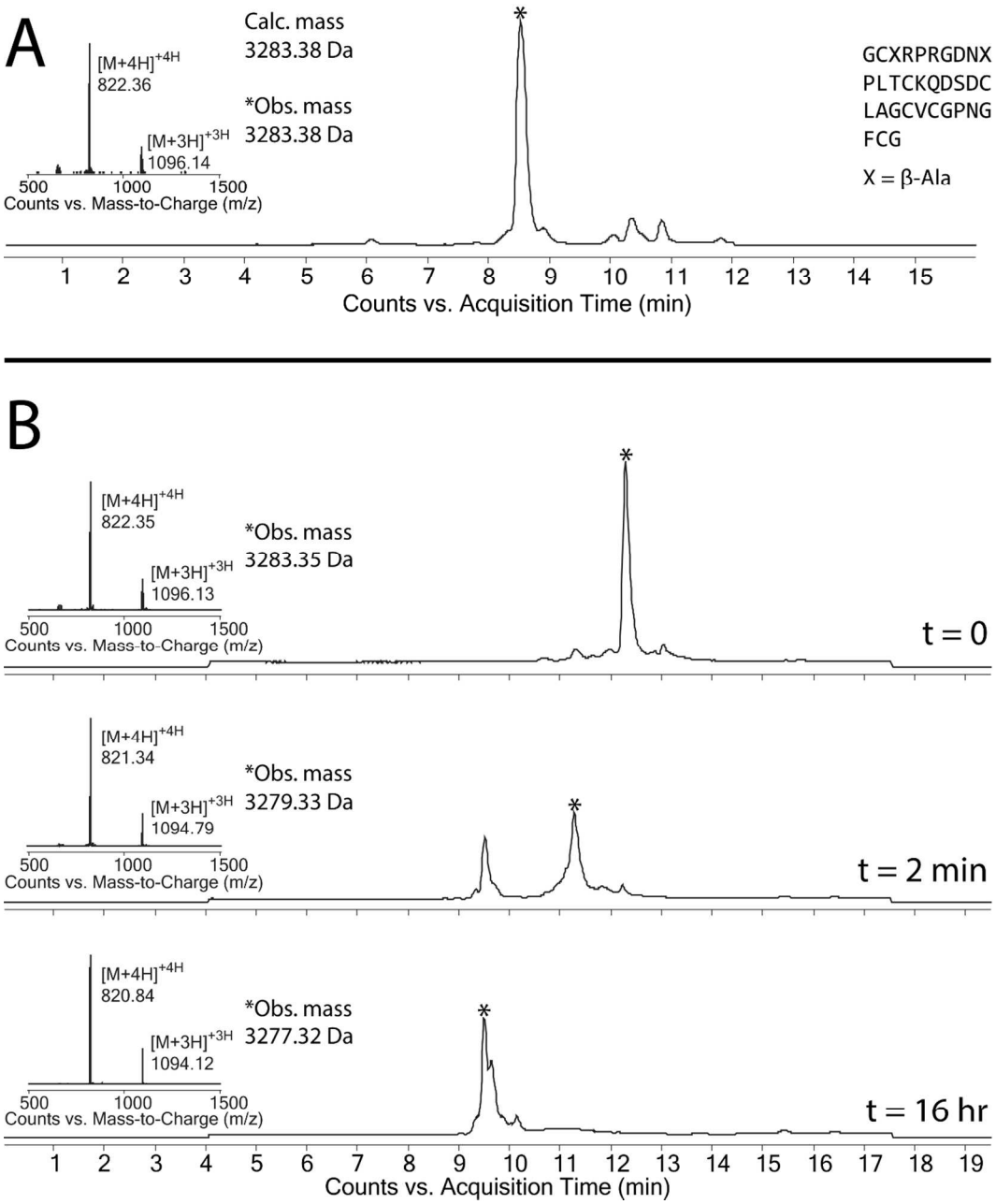

Figure S2: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P3β_A,P10β_A). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (23.5 % yield, 3.9 mg isolated, 16.6 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

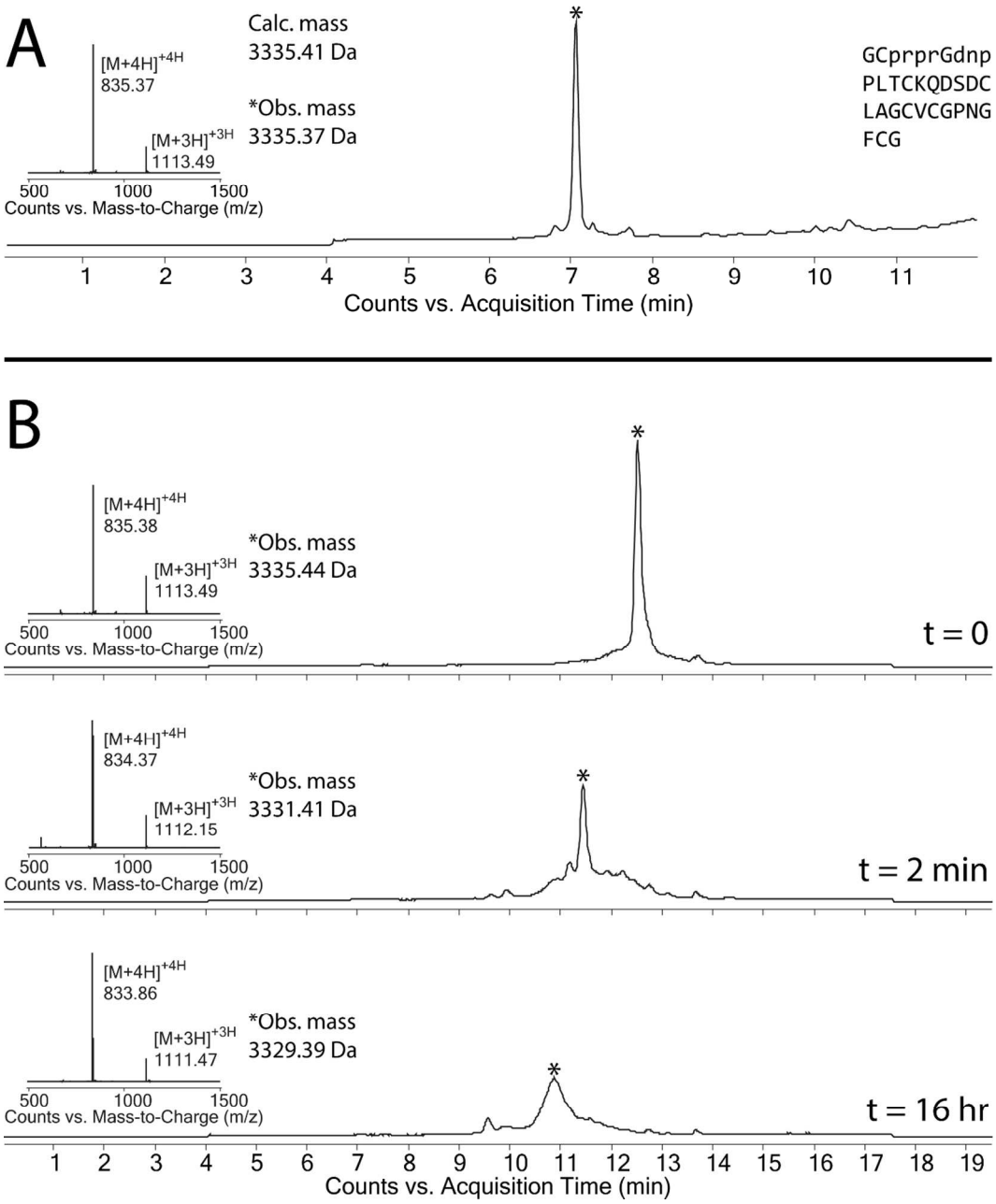

Figure S3: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (32.6 % yield, 7.8 mg isolated, 23.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

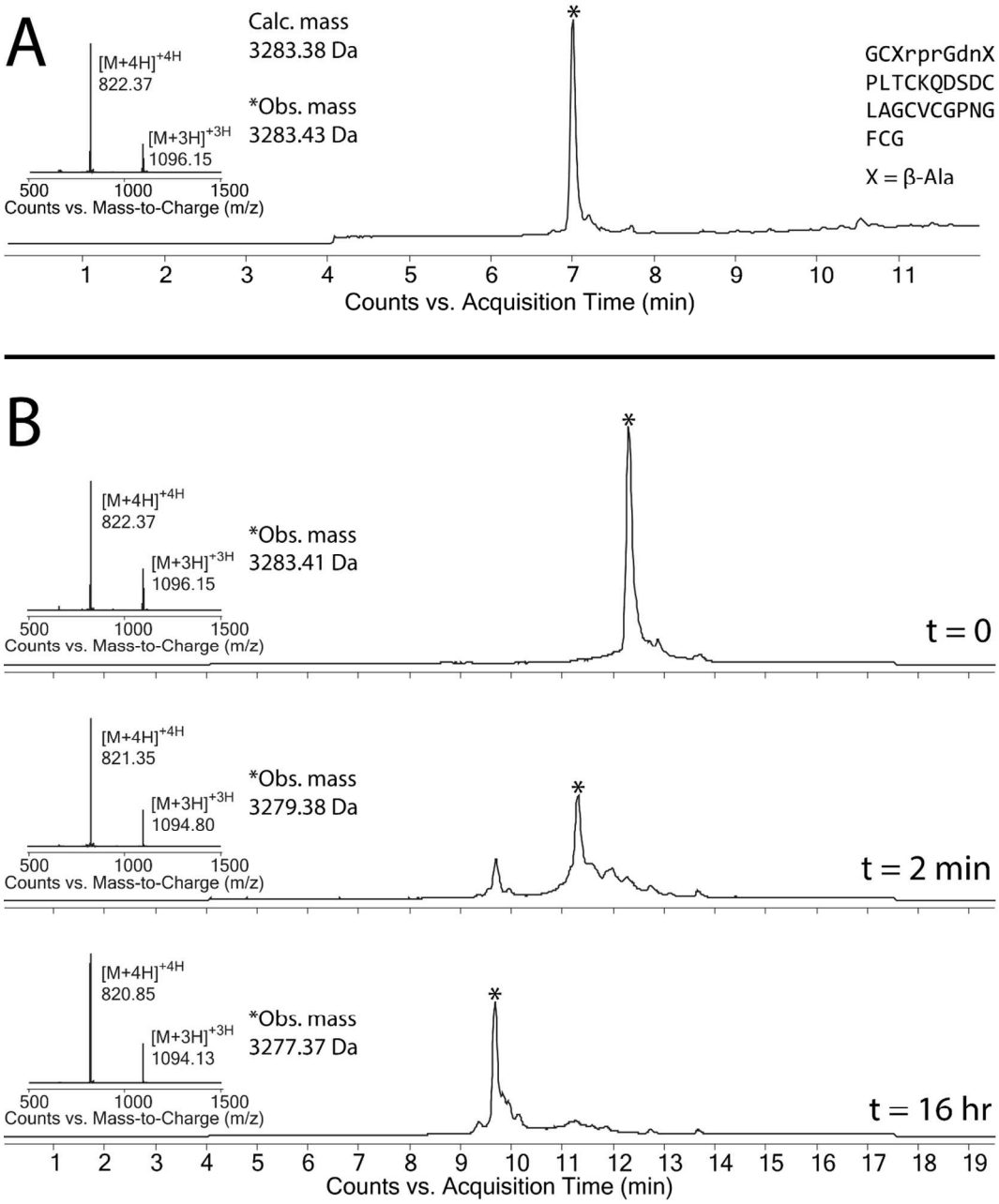

Figure S4: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P3β_A,P10β_A). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (30.5 % yield, 7.4 mg isolated, 24.3 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

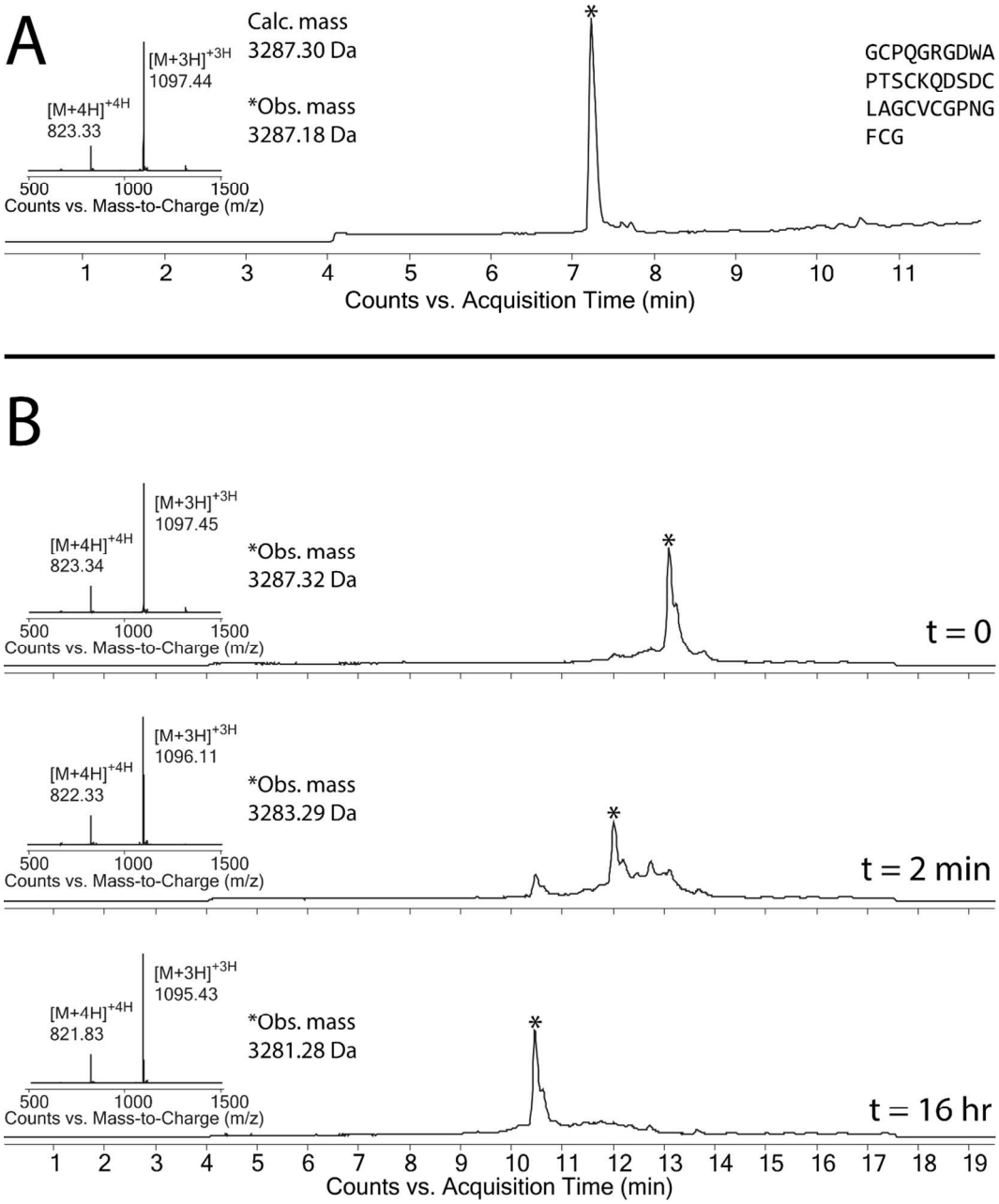

Figure S5: Purification and oxidative folding of ^{L,L}2.5D. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method F (25.4 % yield, 5.3 mg isolated, 20.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

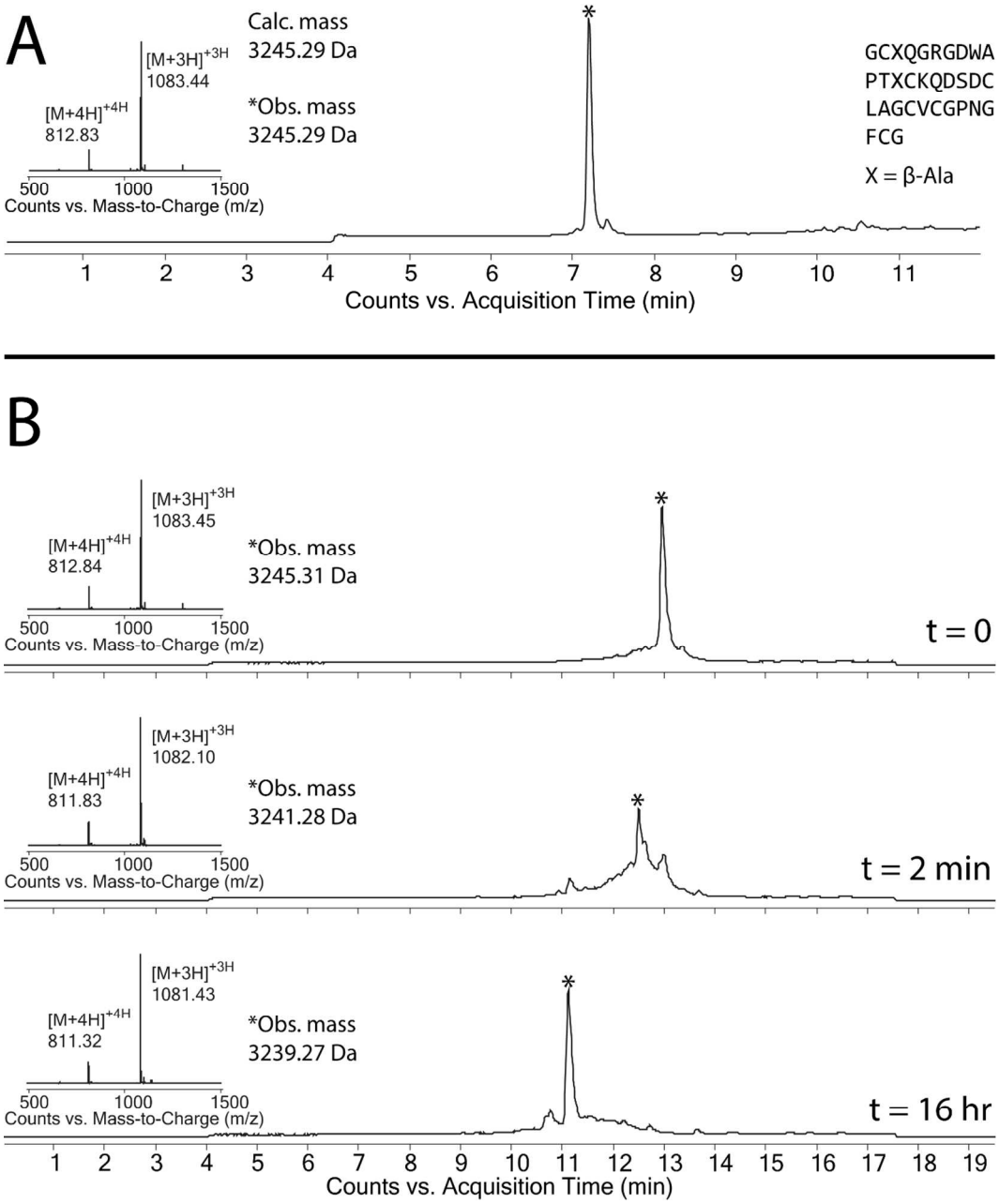

Figure S6: Purification and oxidative folding of ^{L,L}2.5D(P3β_A,S13β_A). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method F (26.5 % yield, 6.1 mg isolated, 23.0 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

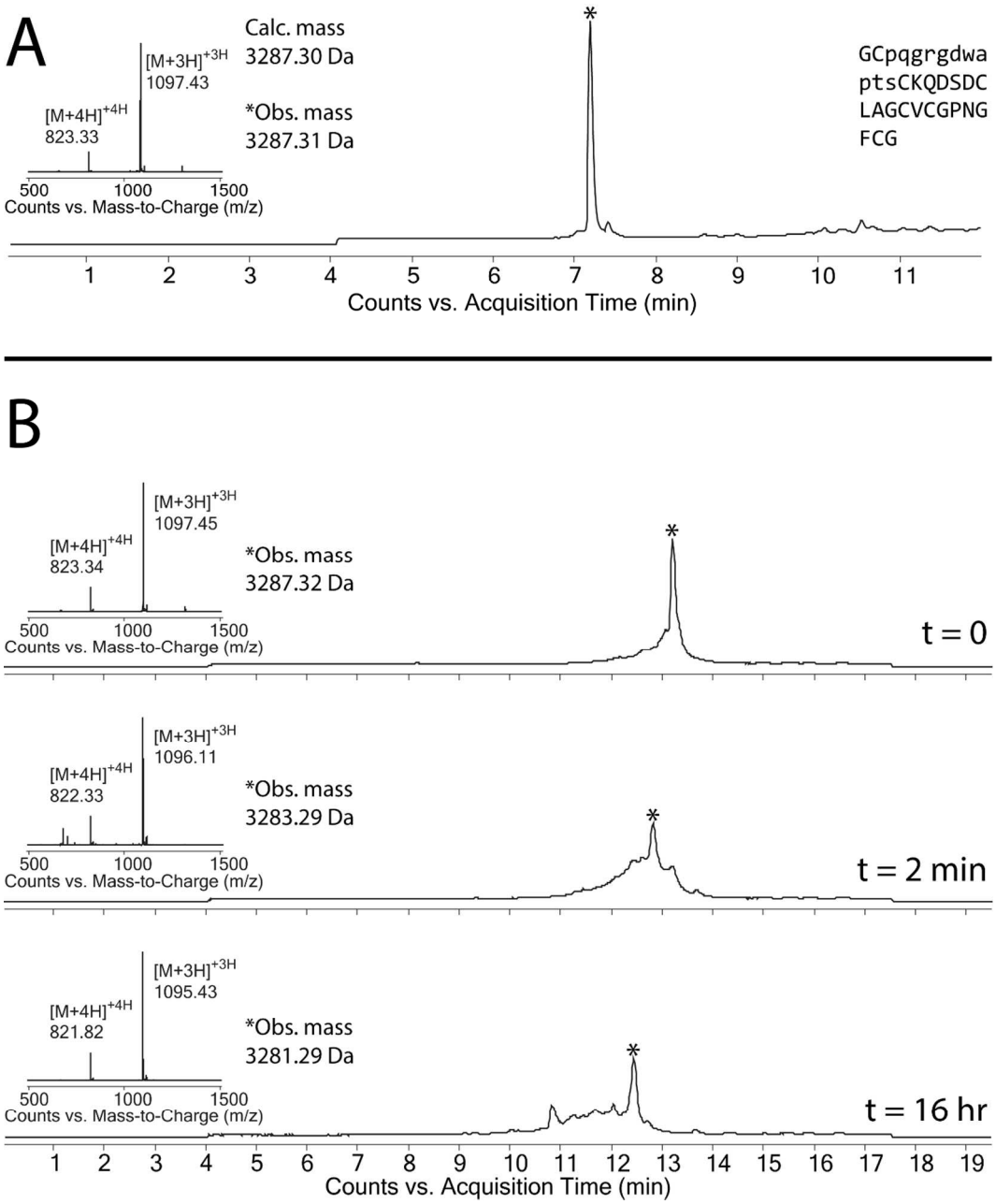

Figure S7: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5D. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method F (51.6 % yield, 11.1 mg isolated, 21.5 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

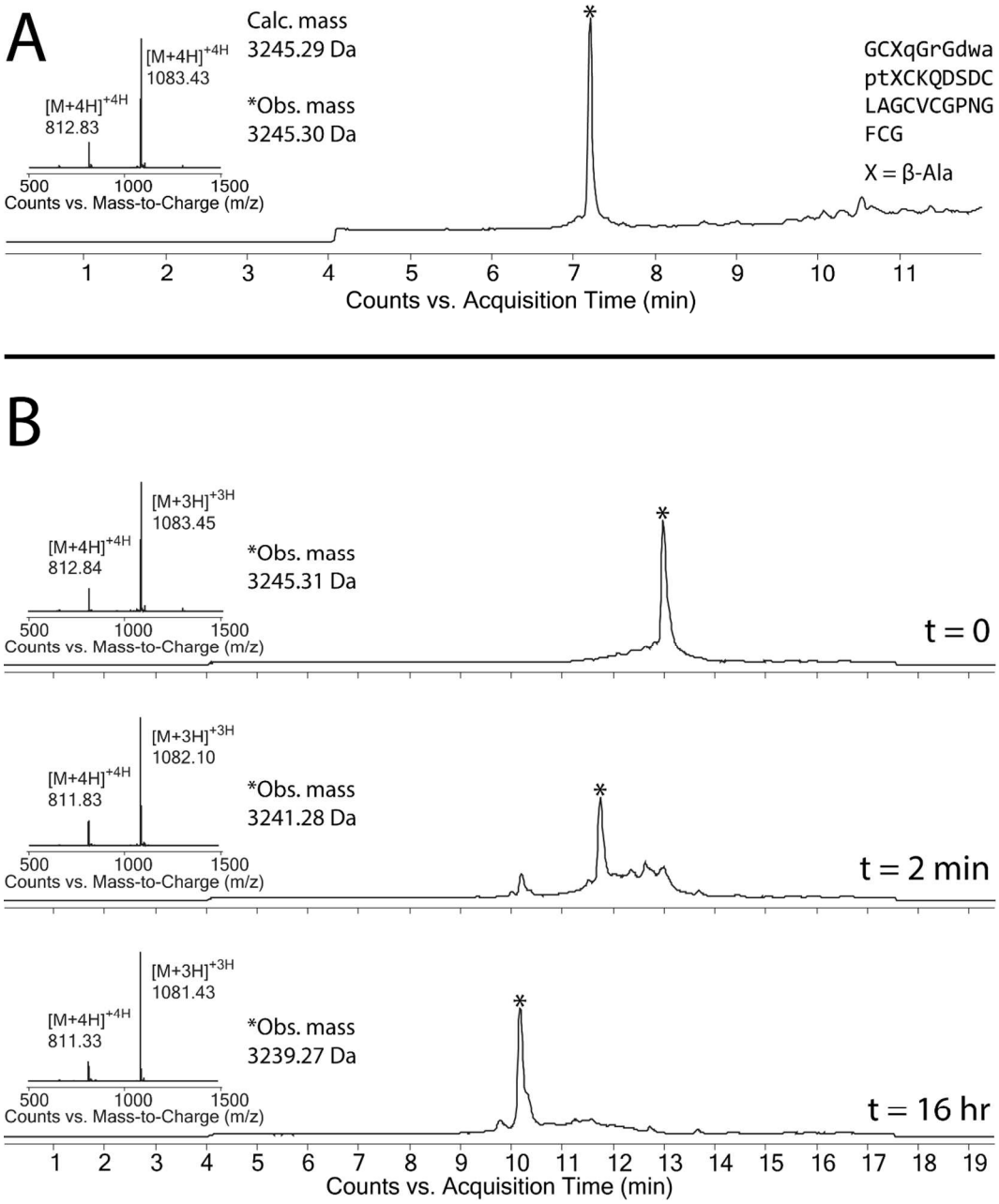

Figure S8: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5D(P3β$_A$,S13β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method F (35.8 % yield, 8.2 mg isolated, 22.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

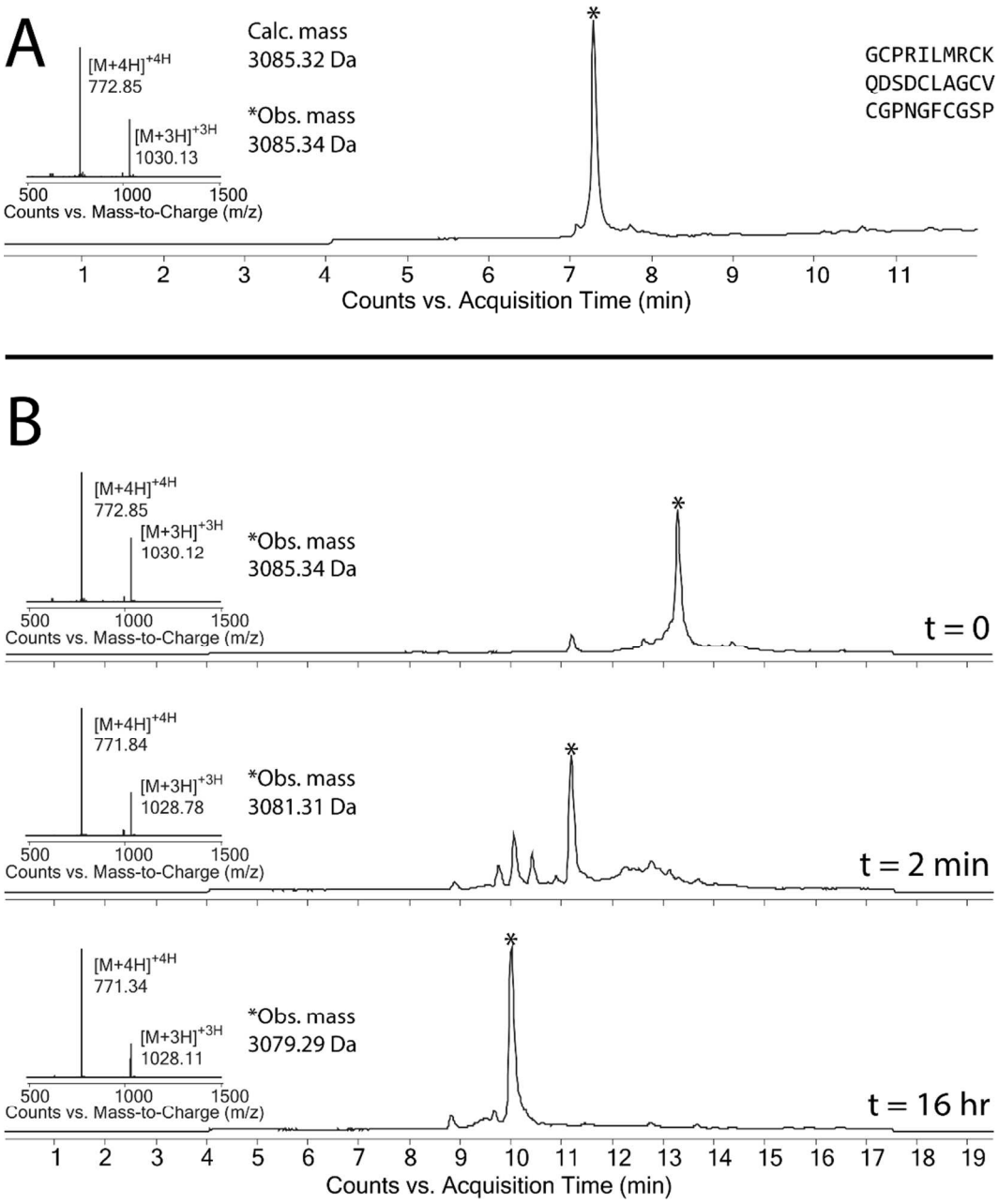

Figure S9: Purification and oxidative folding of ^{L,L}WT-SP. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (12.1 % yield, 1.3 mg isolated, 10.7 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

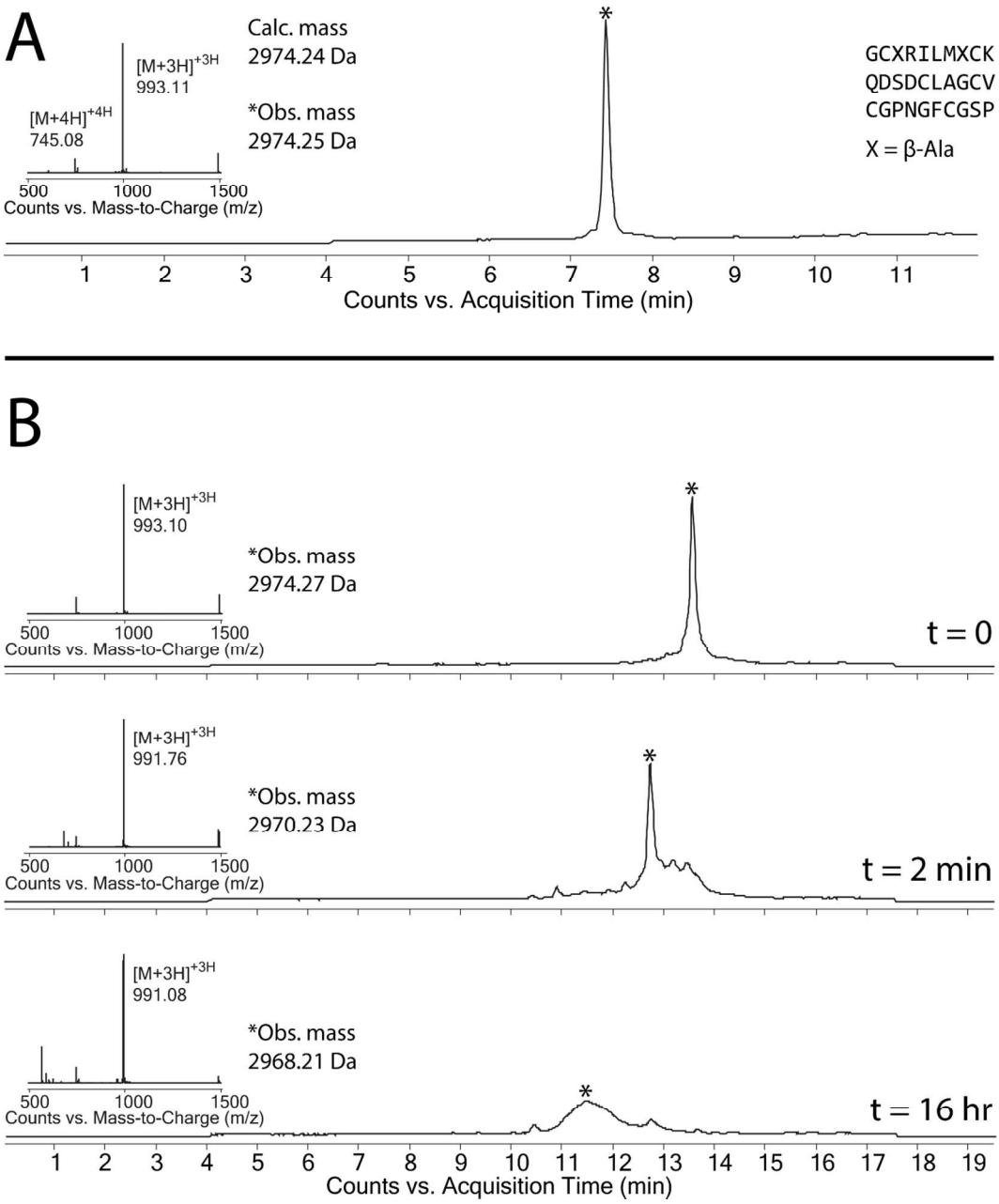

Figure S10: Purification and oxidative folding of ^{L,L}WT-SP(P3β$_A$,R8β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (21.4 % yield, 3.7 mg isolated, 17.3 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

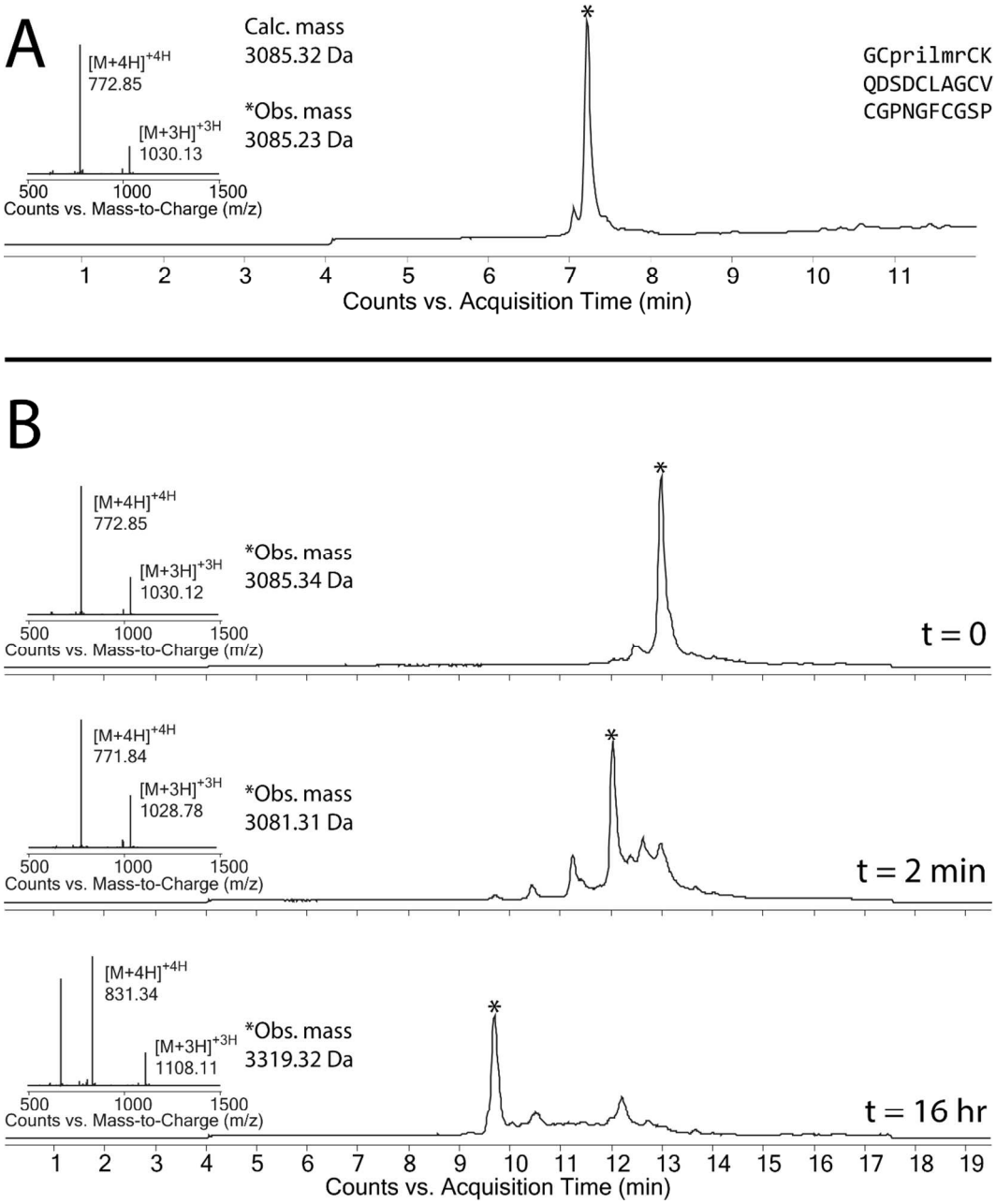

Figure S11: Purification and oxidative folding of $^{L,D(3-8)}$WT-SP. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (17.6 % yield, 2.8 mg isolated, 15.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. The mass of the "t = 16hr" major product corresponds to a 2 disulfide EETI-II analog + 2 cysteine adduct. Observed and calculated masses are monoisotopic.

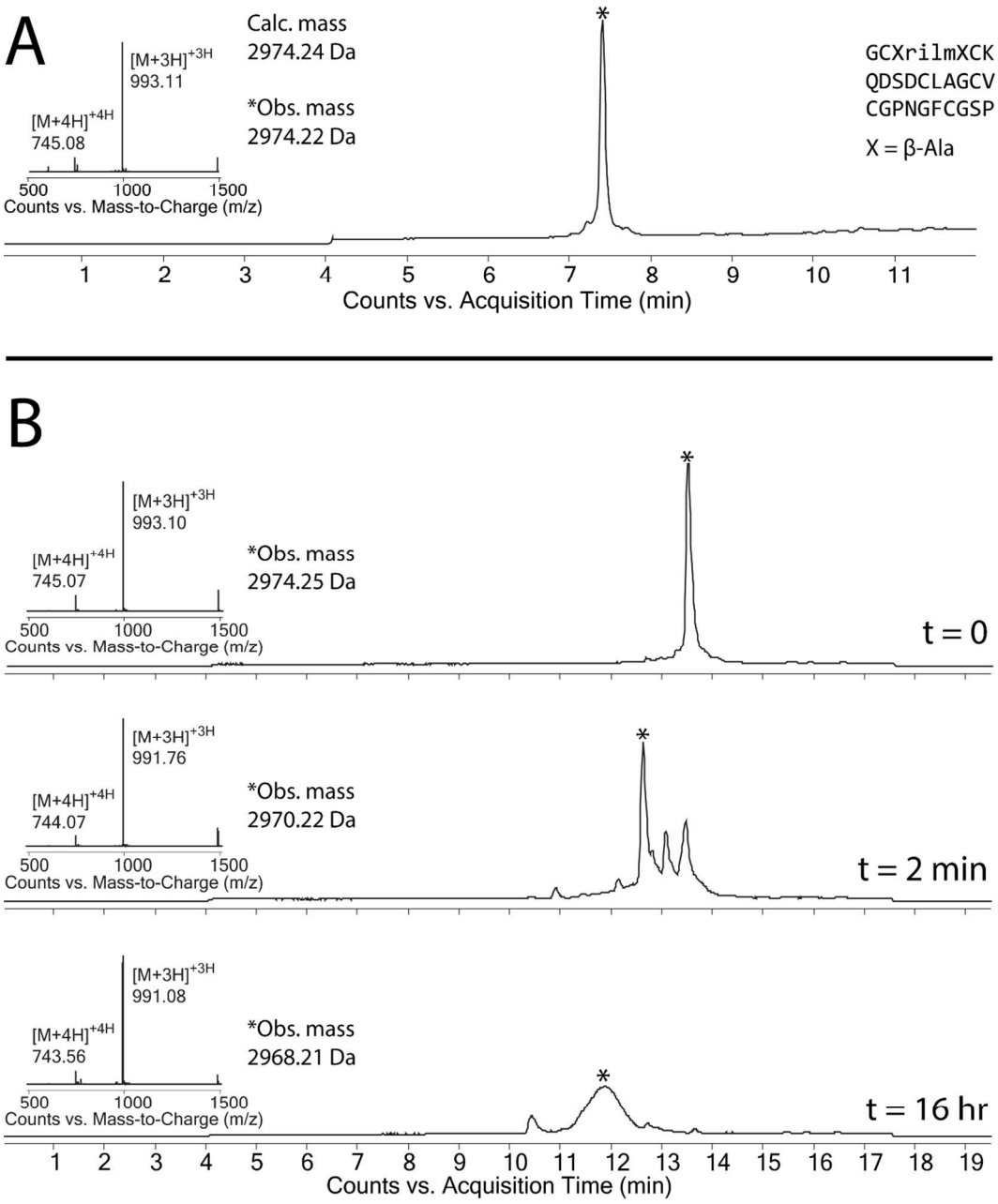

Figure S12: Purification and oxidative folding of $^{L,D(3-8)}$WT-SP(P3β$_A$,R8β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (23.6 % yield, 3.9 mg isolated, 16.5 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

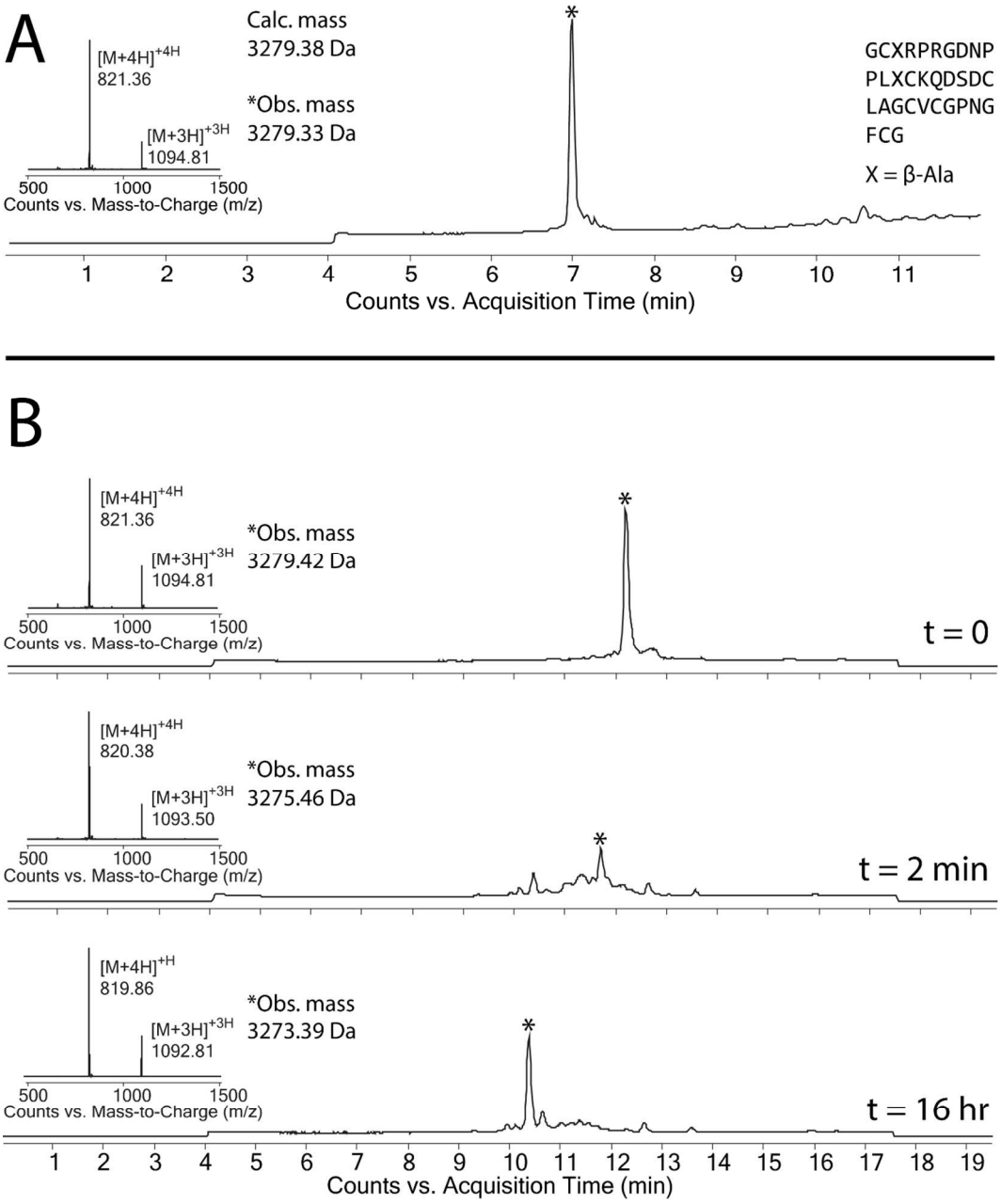

Figure S13: Purification and oxidative folding of ^{L,L}2.5F(P3β$_A$,T13β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (39.9 % yield, 23.6 mg isolated, 59.2 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

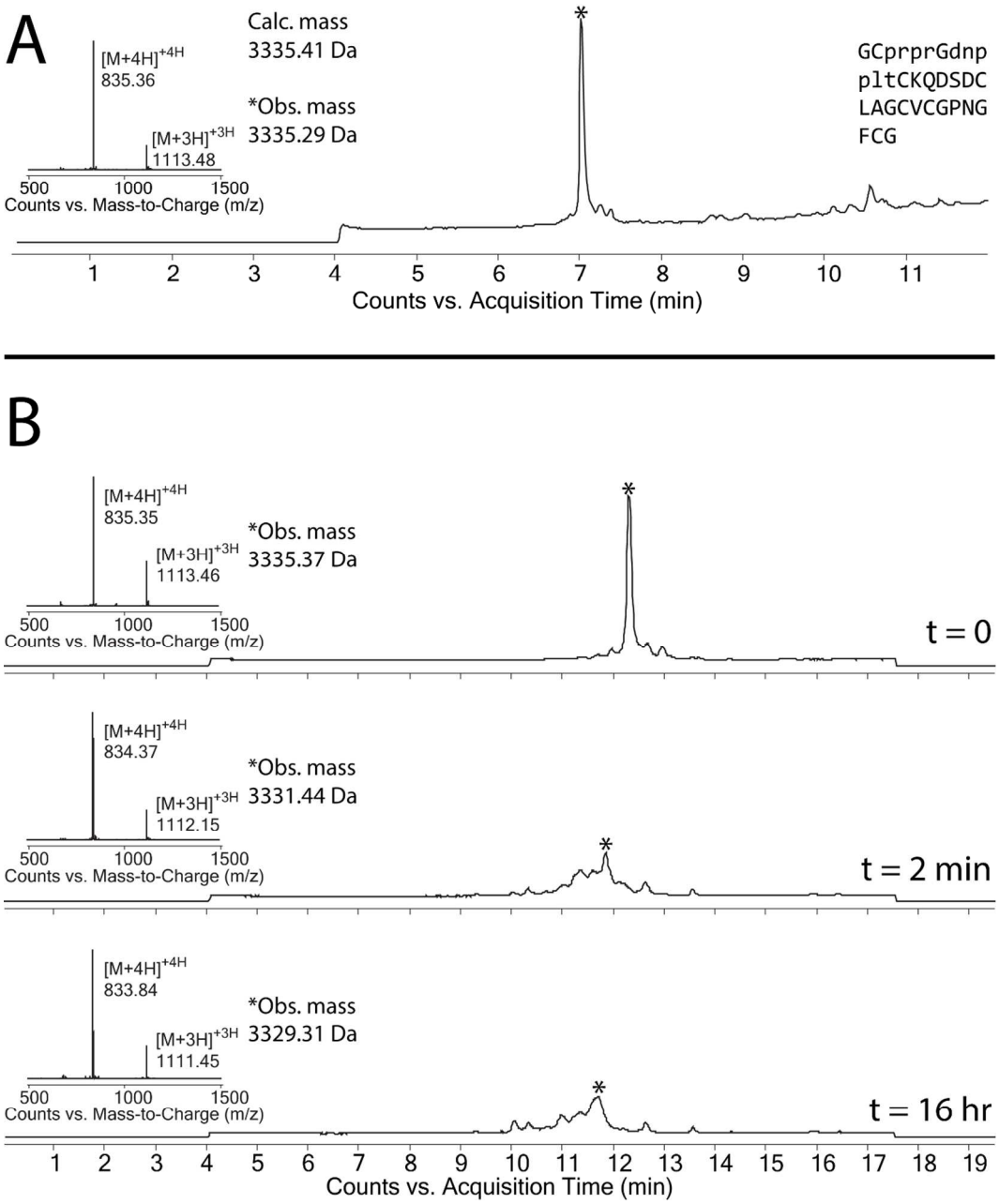

Figure S14: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (34.3 % yield, 15.8 mg isolated, 46.1 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

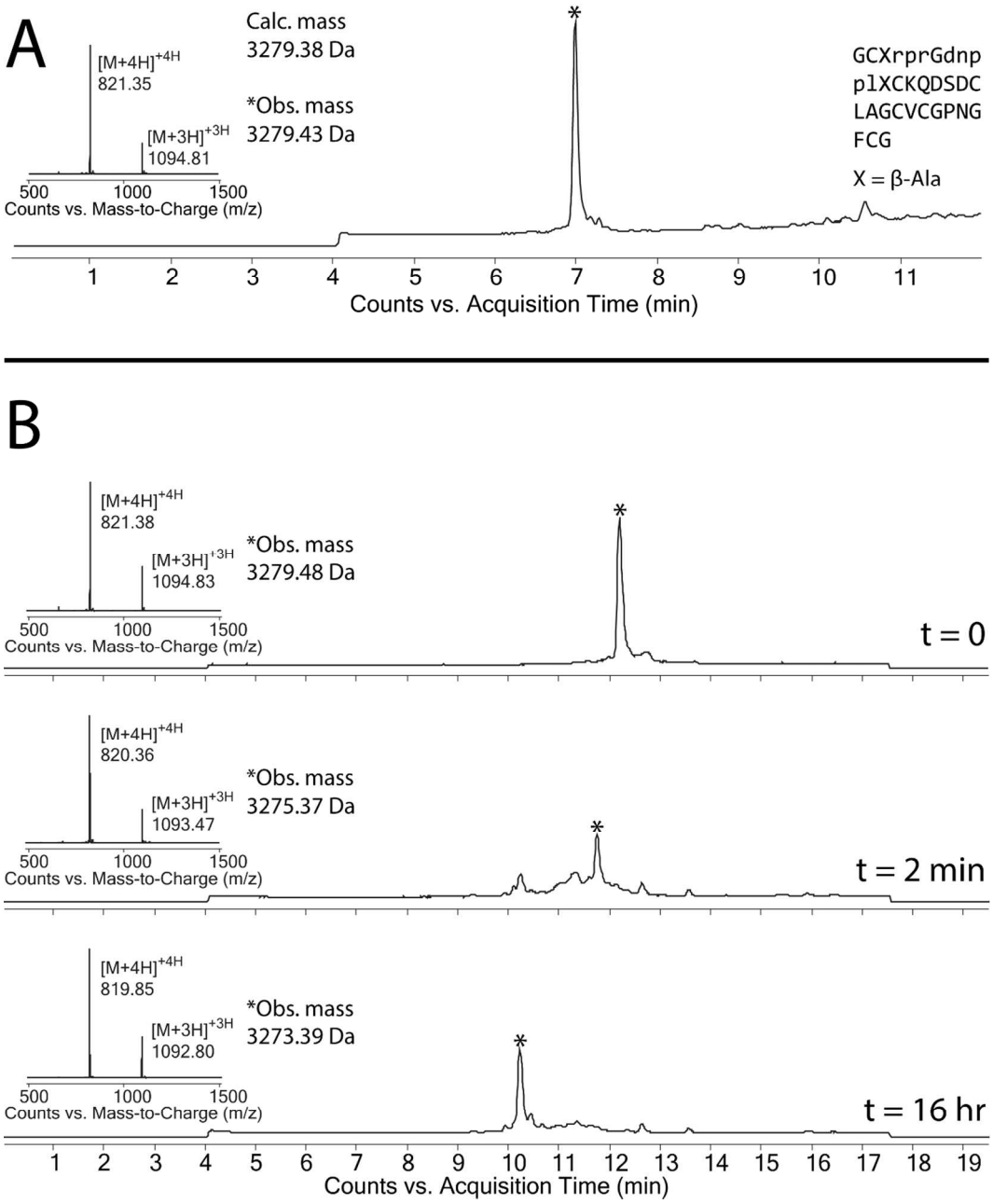

Figure S15: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F(P3β$_A$,T13β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (37.1 % yield, 22.8 mg isolated, 61.4 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

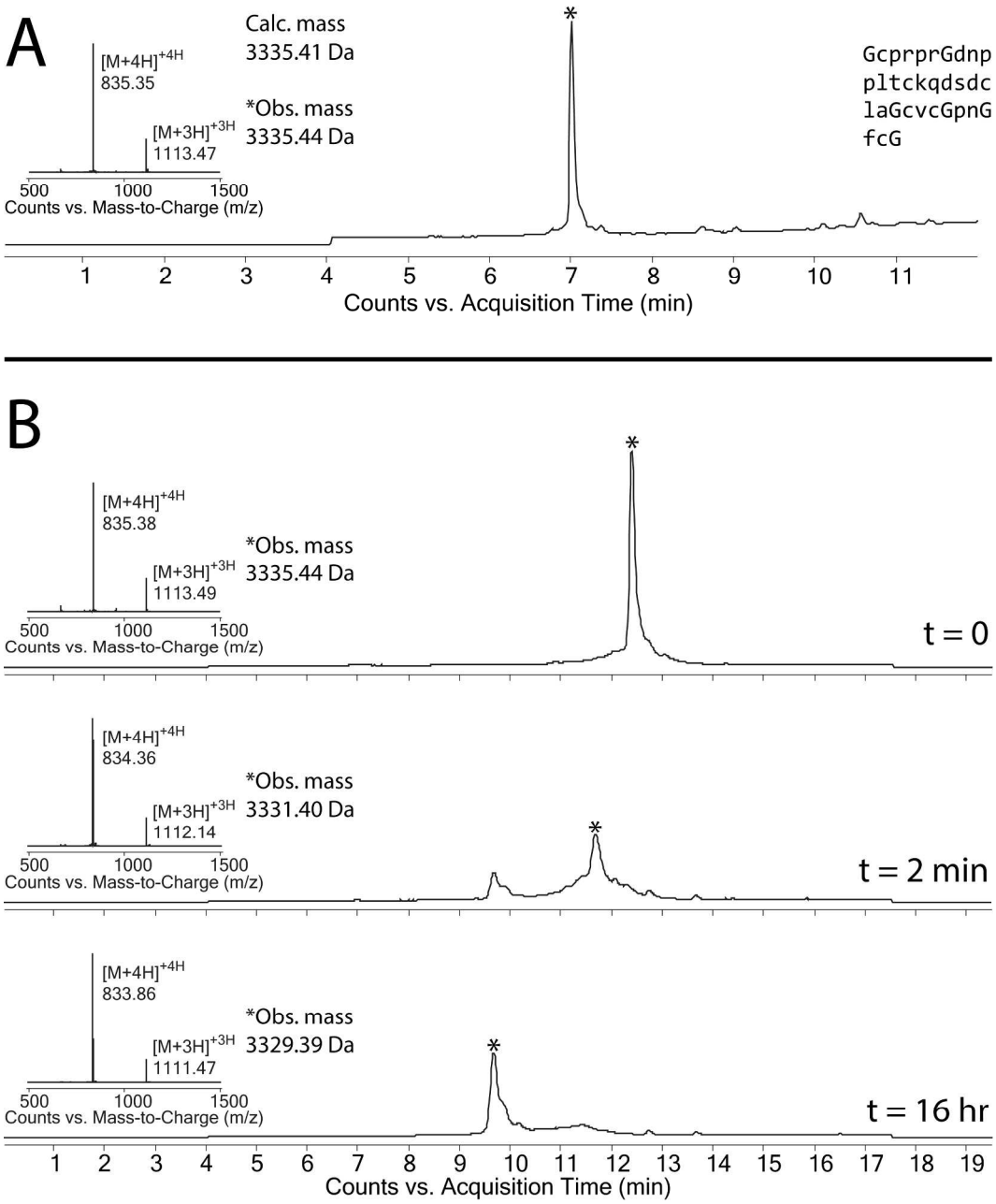

Figure S16: Purification and oxidative folding of ^{D,D}2.5F. LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (44.1 % yield, 26.0 mg isolated, 58.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

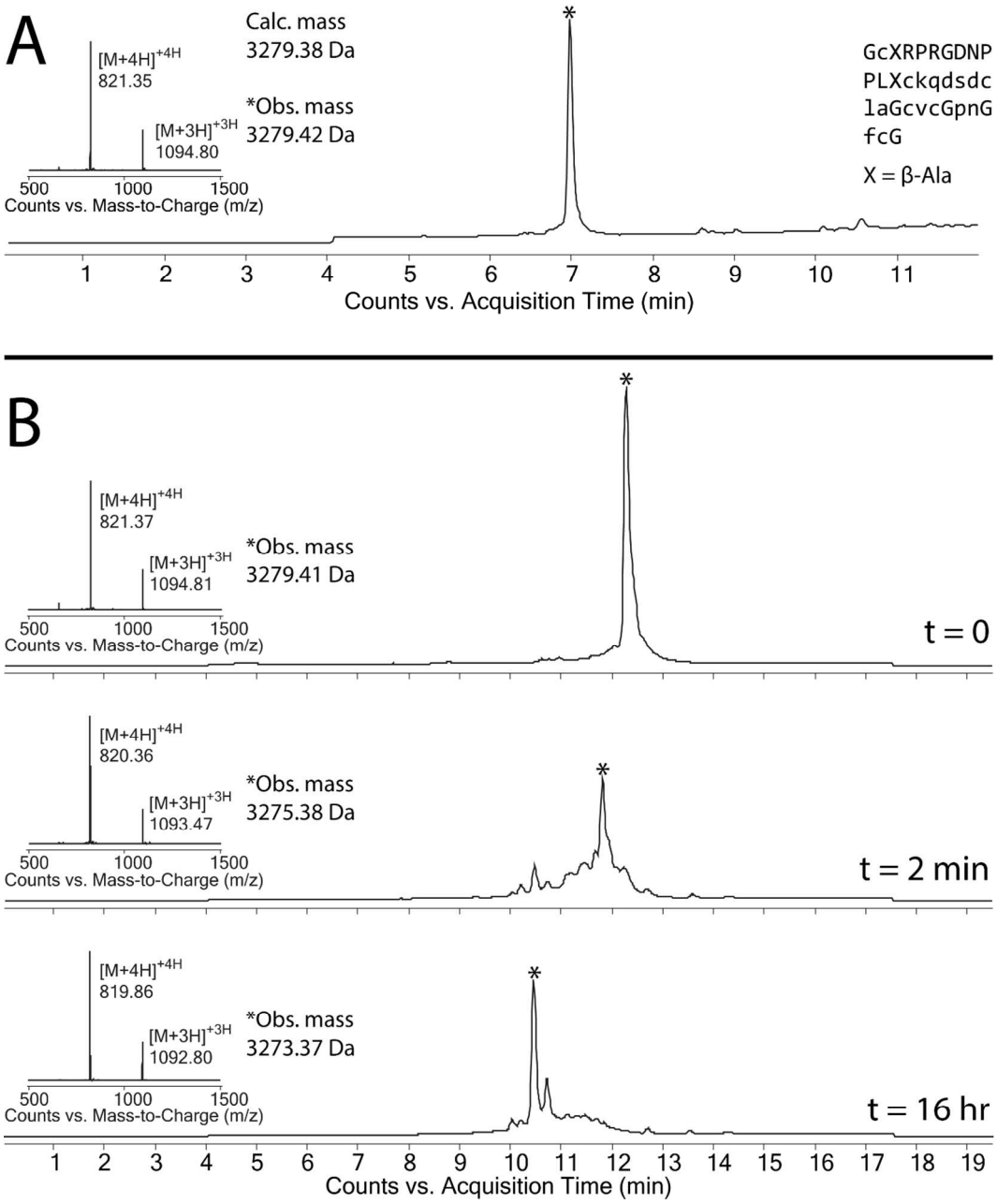

Figure S17: Purification and oxidative folding of $^{D,L(3\text{-}13)}$2.5F(p3β_A,p10β_A). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (44.2 % yield, 27.4 mg isolated, 62.0 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

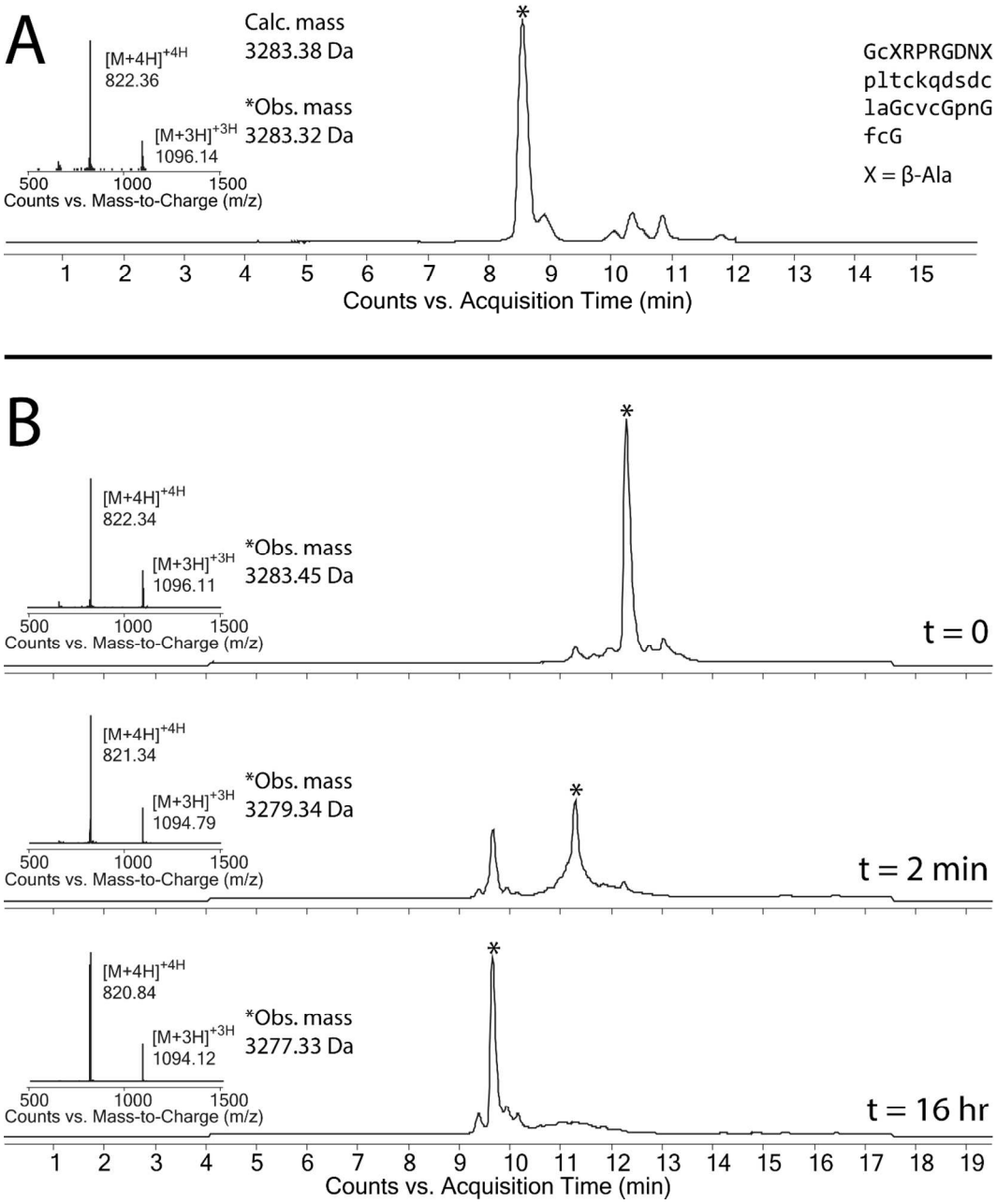

Figure S18: Purification and oxidative folding of $^{D,L(3\text{-}10)}$2.5F(p3β$_A$,p10β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (21.5 % yield, 3.7 mg isolated, 17.2 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

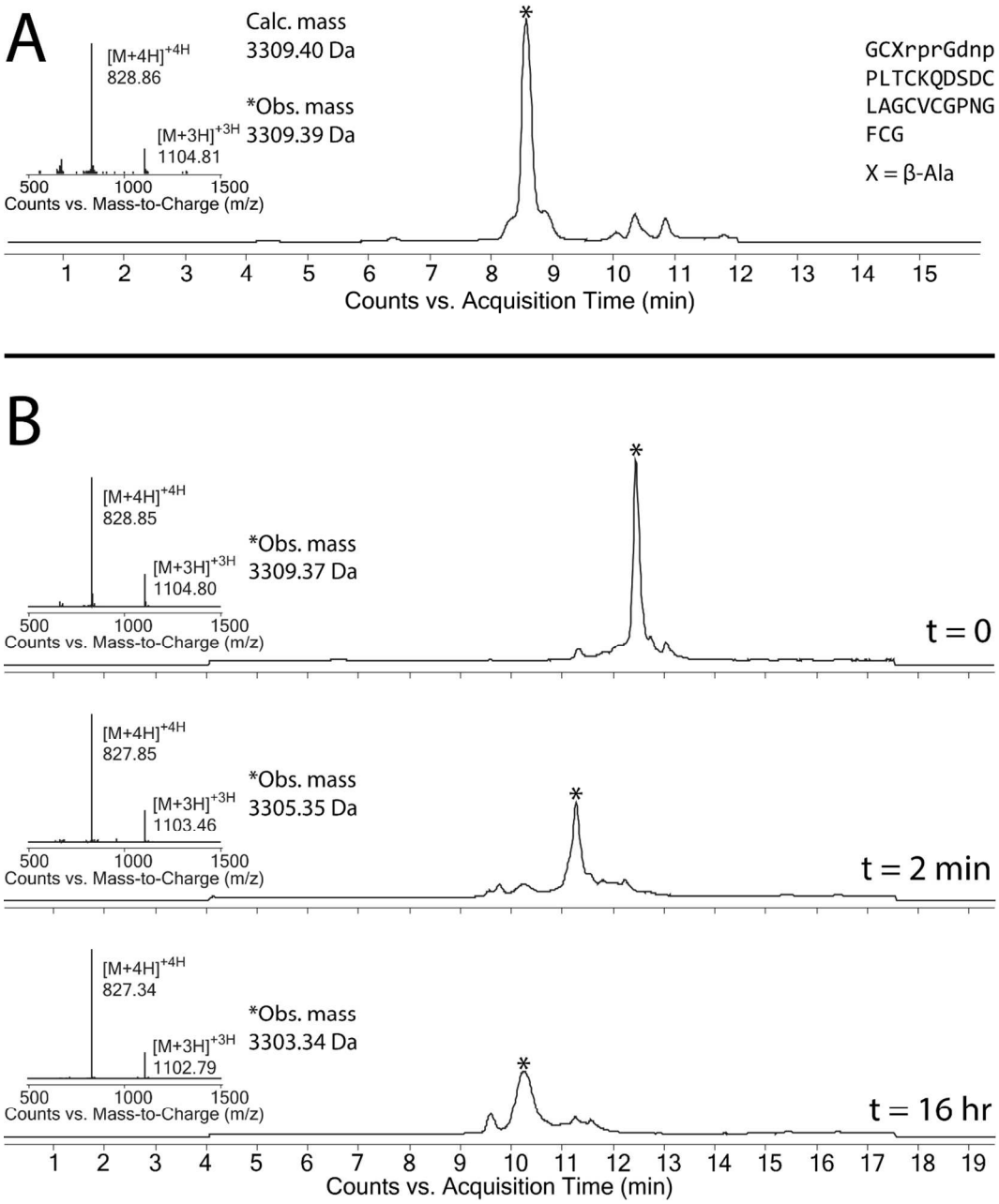

Figure S19: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P3β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (27.2 % yield, 4.6 mg isolated, 16.9 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

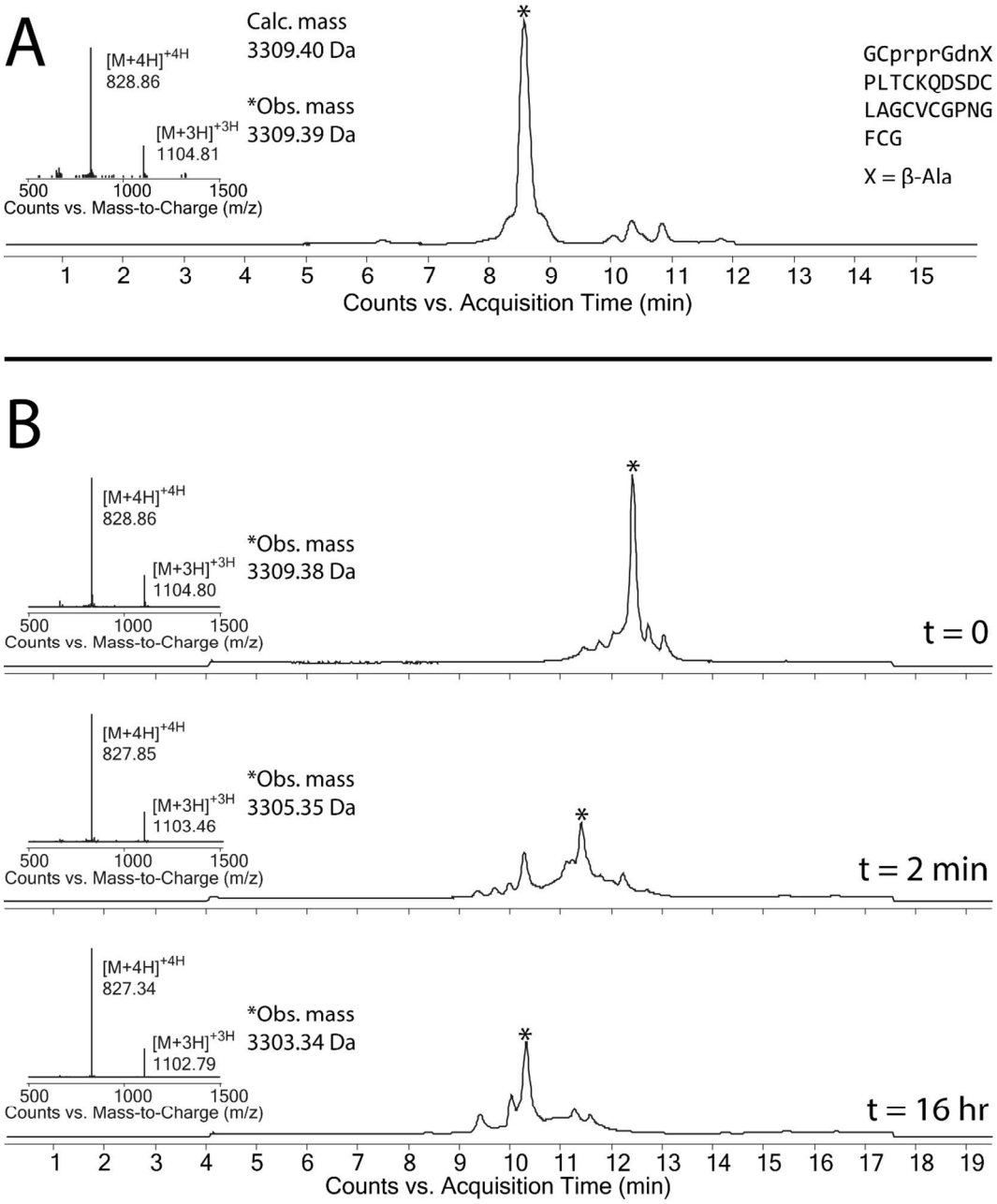

Figure S20: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P10β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (16.6 % yield, 2.9 mg isolated, 17.5 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

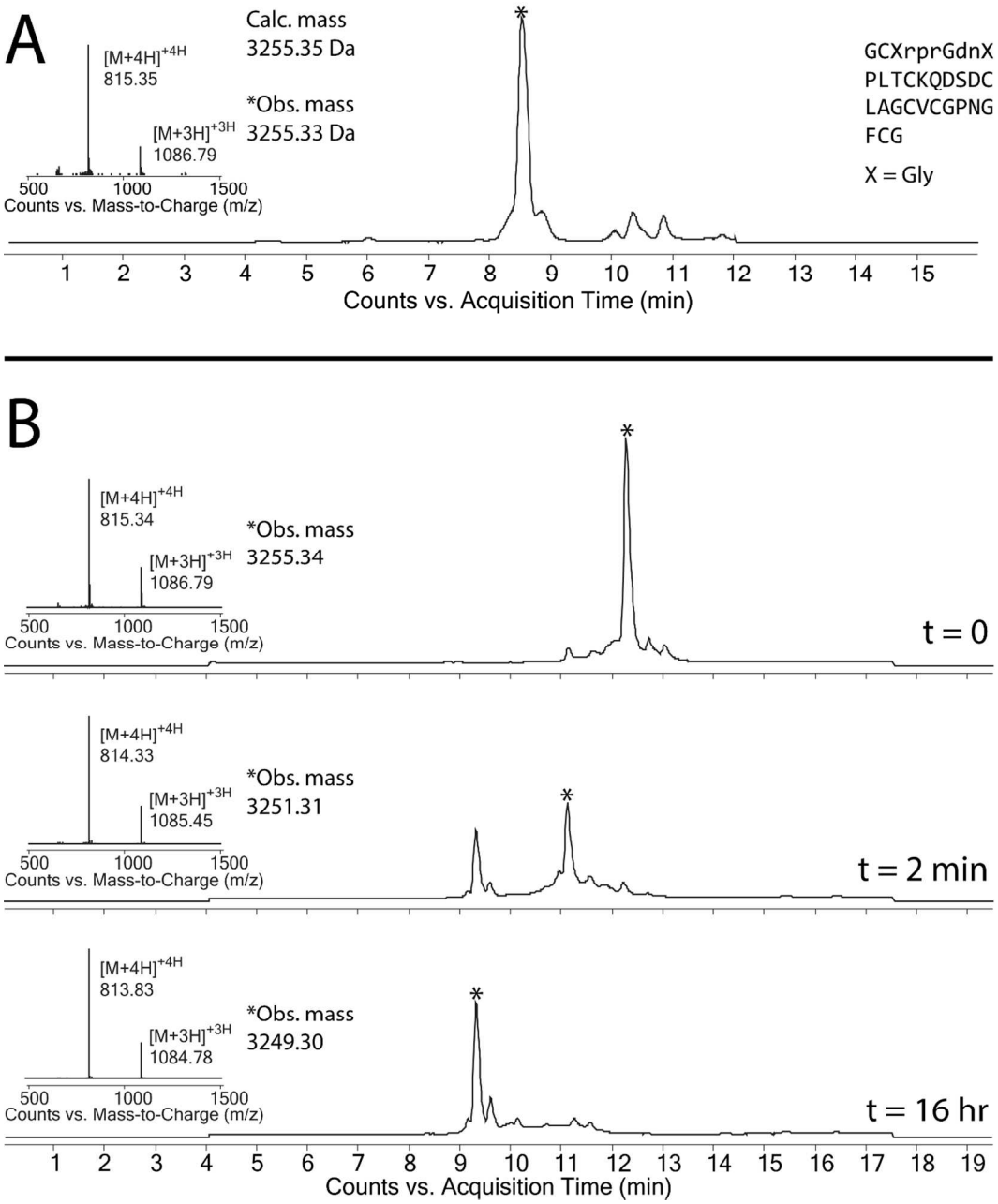

Figure S21: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P3G,P10G). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (21.8 % yield, 3.8 mg isolated, 17.4 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

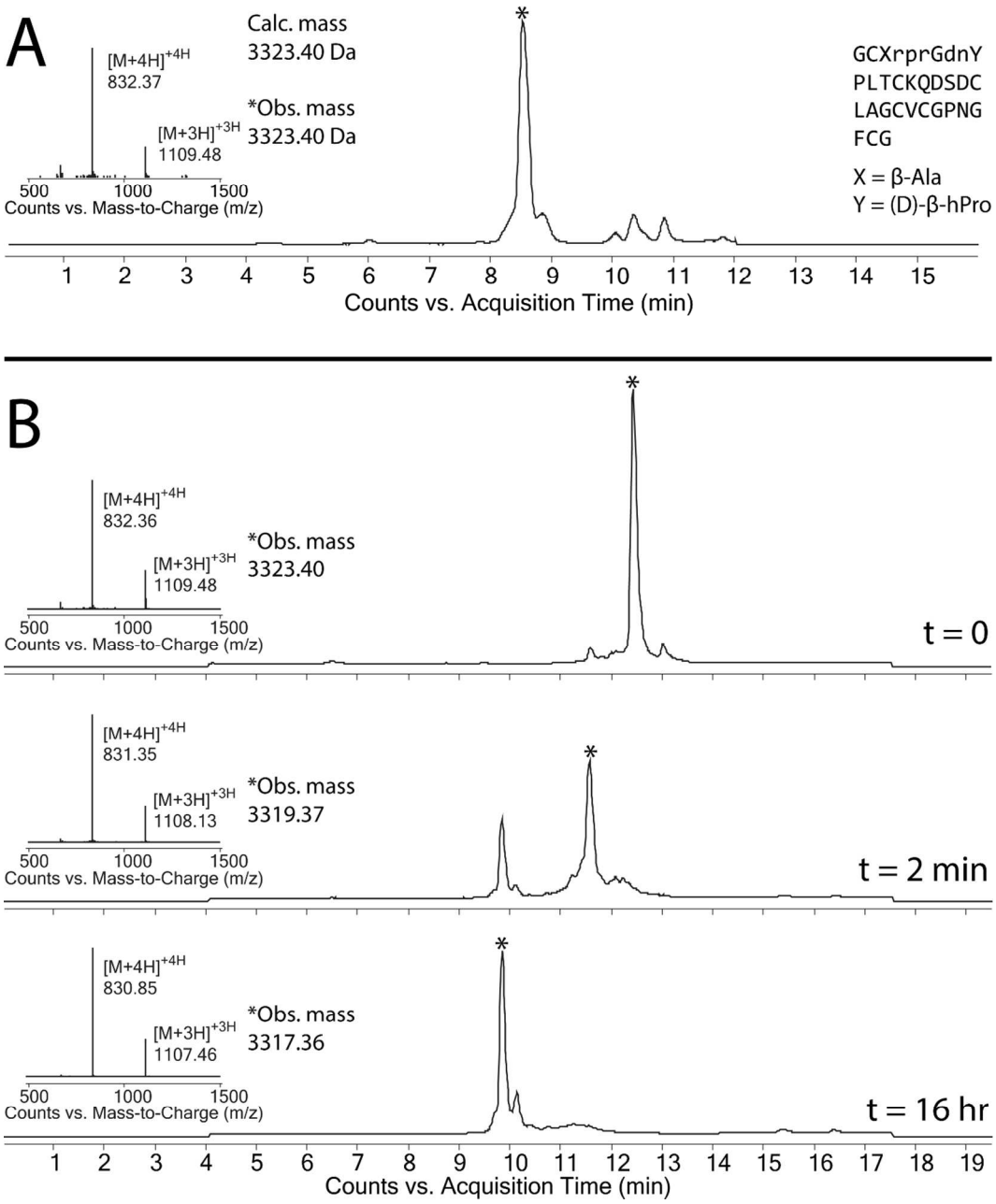

Figure S22: Purification and oxidative folding of $^{L,D(3\text{-}10)}$2.5F(P3β$_A$,P10^{D}β$_{hP}$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (21.3 % yield, 3.8 mg isolated, 17.8 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

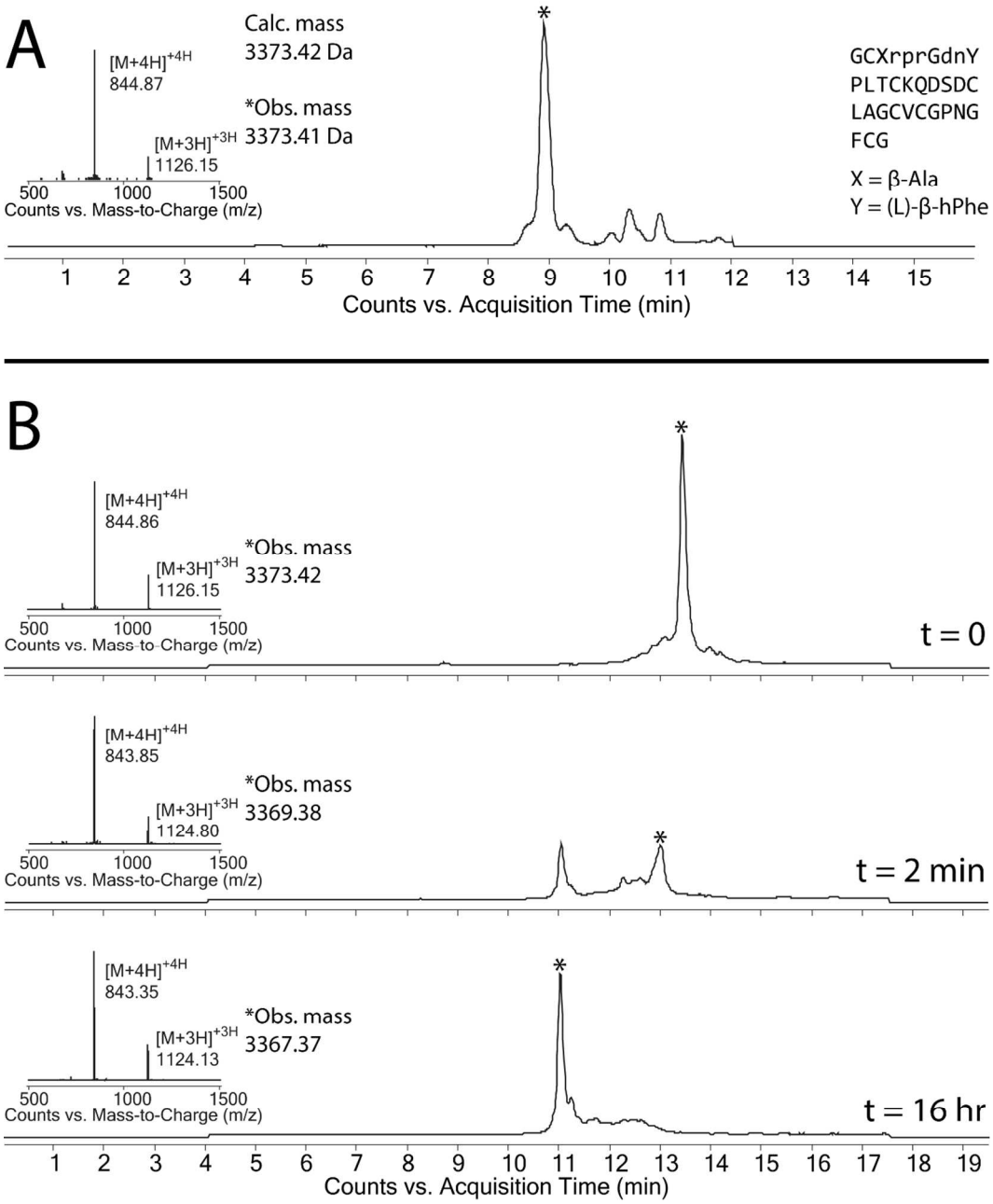

Figure S23: Purification and oxidative folding of $^{L,D(3-9)}$2.5F(P3β$_A$,P10^{L}β$_{hF}$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (19.6 % yield, 3.2 mg isolated, 16.3 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

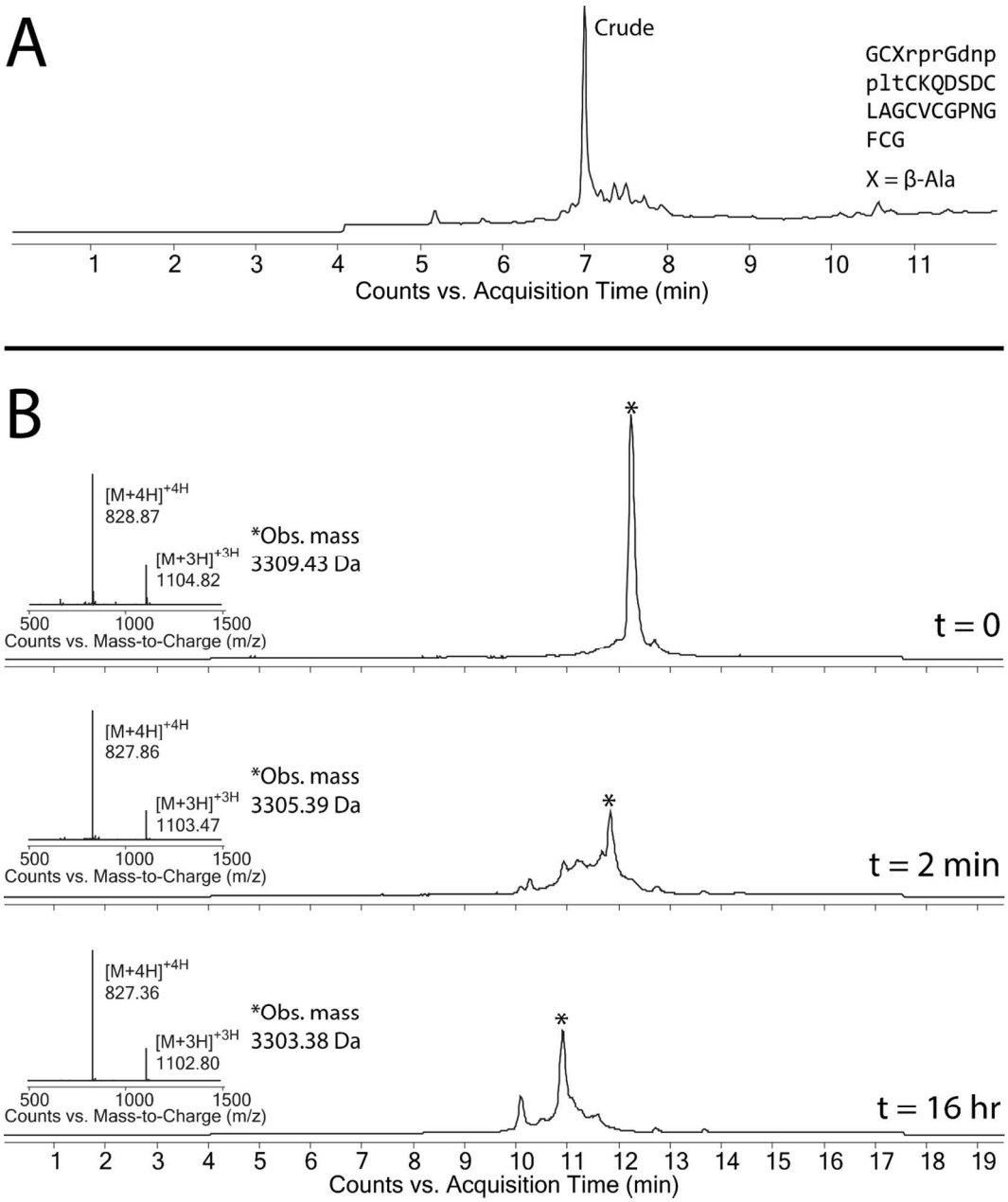

Figure S24: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F(P3β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (36.6 % yield, 6.8 mg isolated, 18.6 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

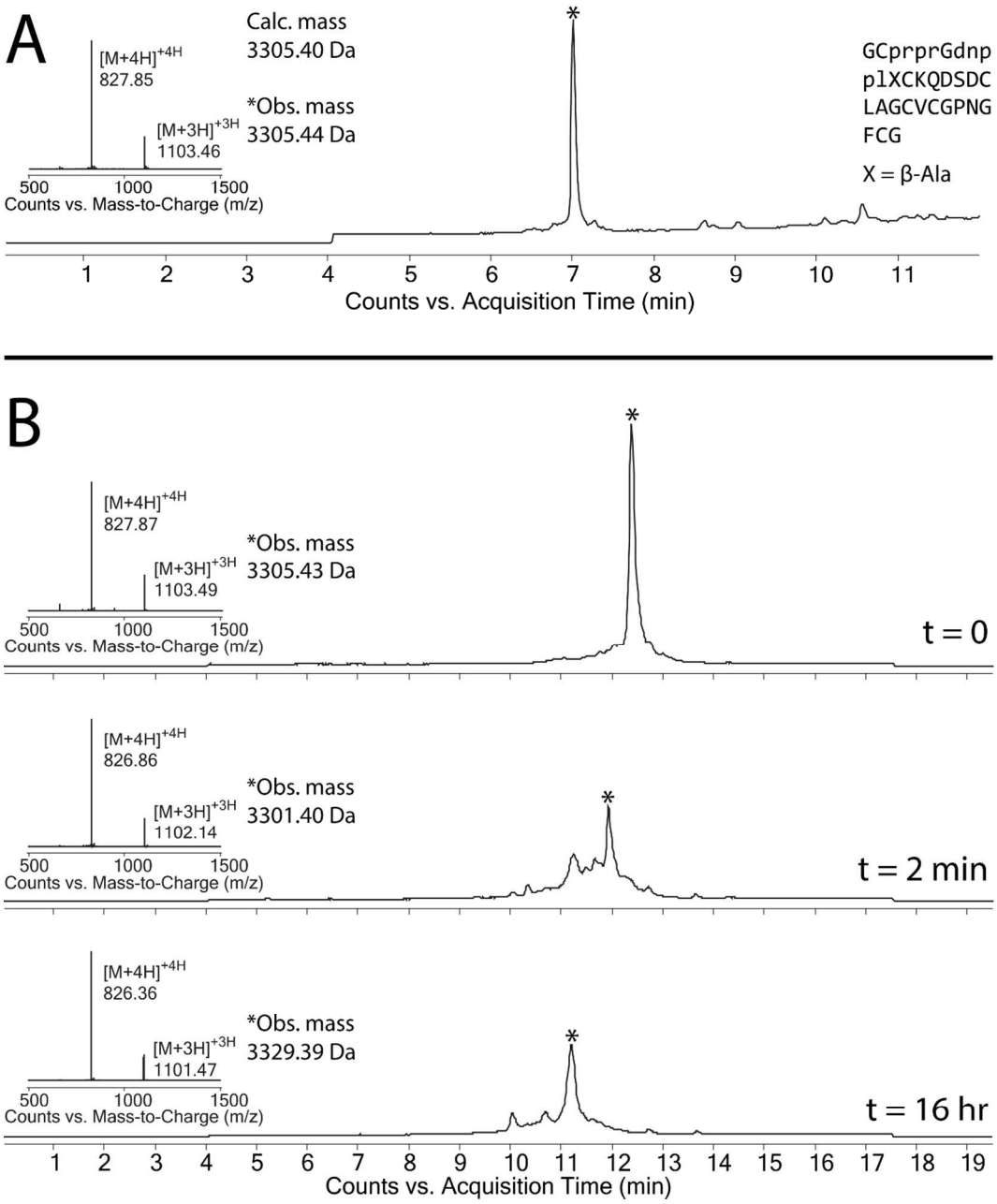

Figure S25: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F(T13β$_A$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (40.4 % yield, 22.2 mg isolated, 55.0 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

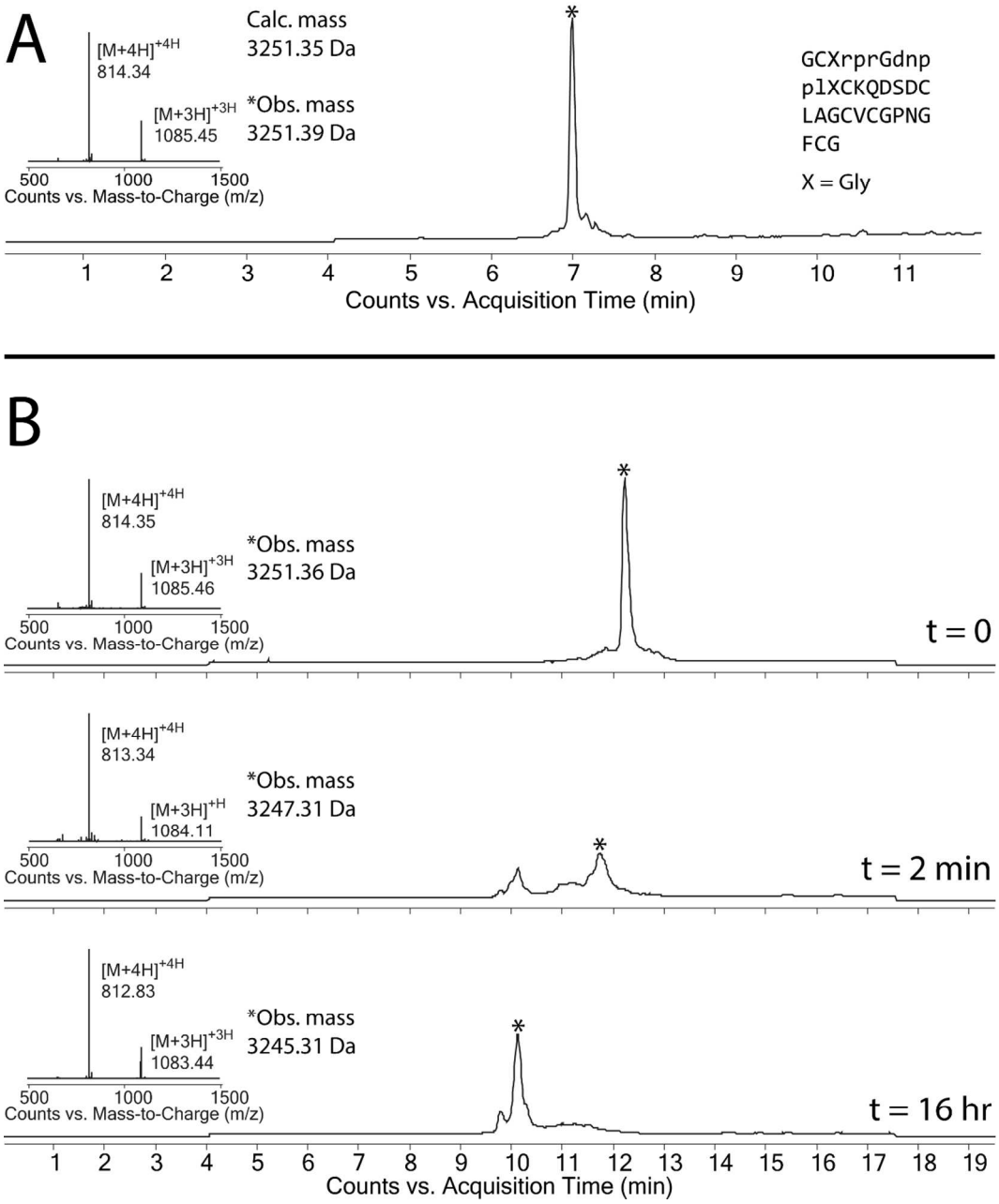

Figure S26: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F(P3G,T13G). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (31.9 % yield, 4.4 mg isolated, 13.8 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

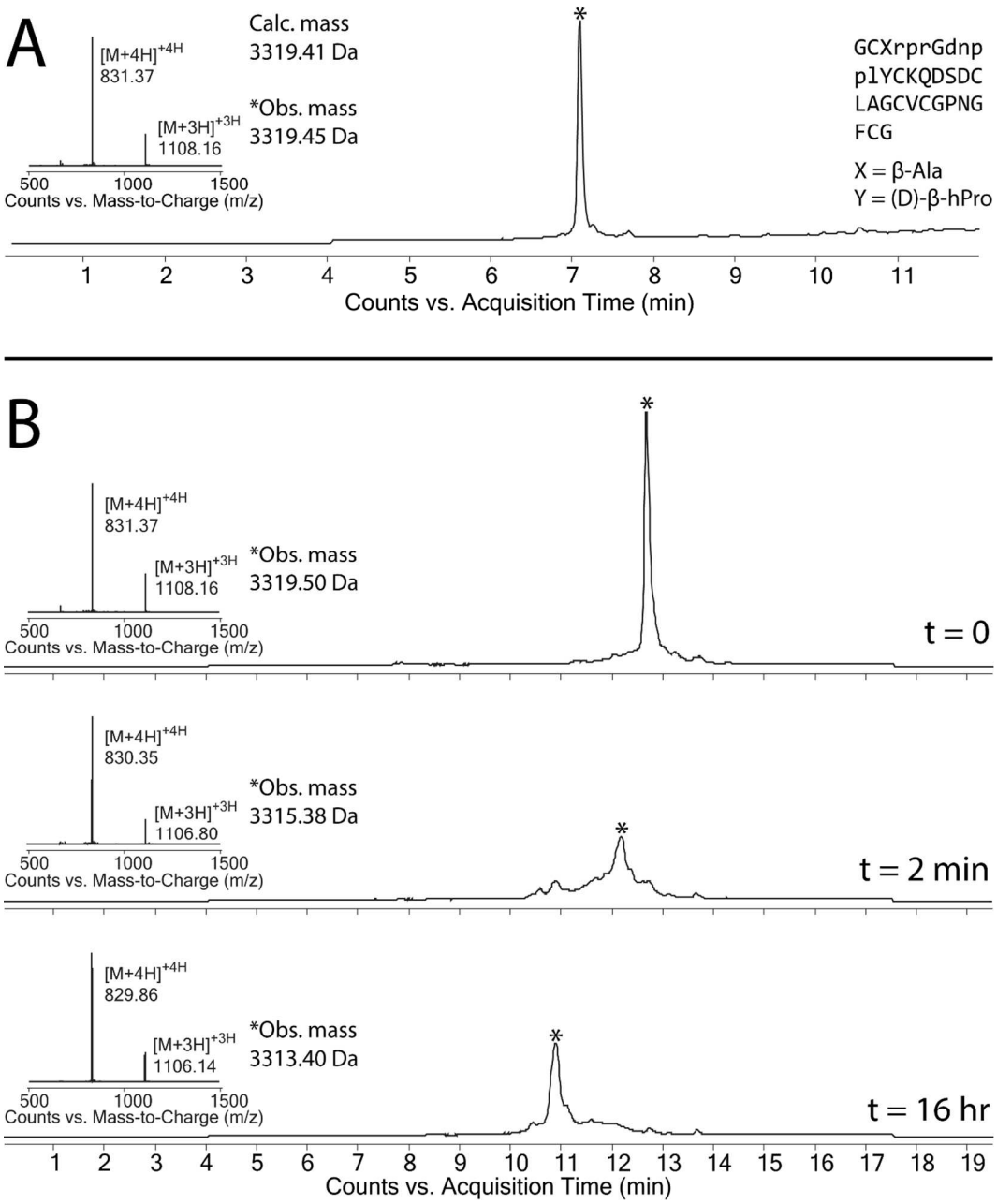

Figure S27: Purification and oxidative folding of $^{L,D(3\text{-}13)}$2.5F(P3β$_A$,T13^{D}β$_{hP}$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (42.3 % yield, 10.5 mg isolated, 24.8 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

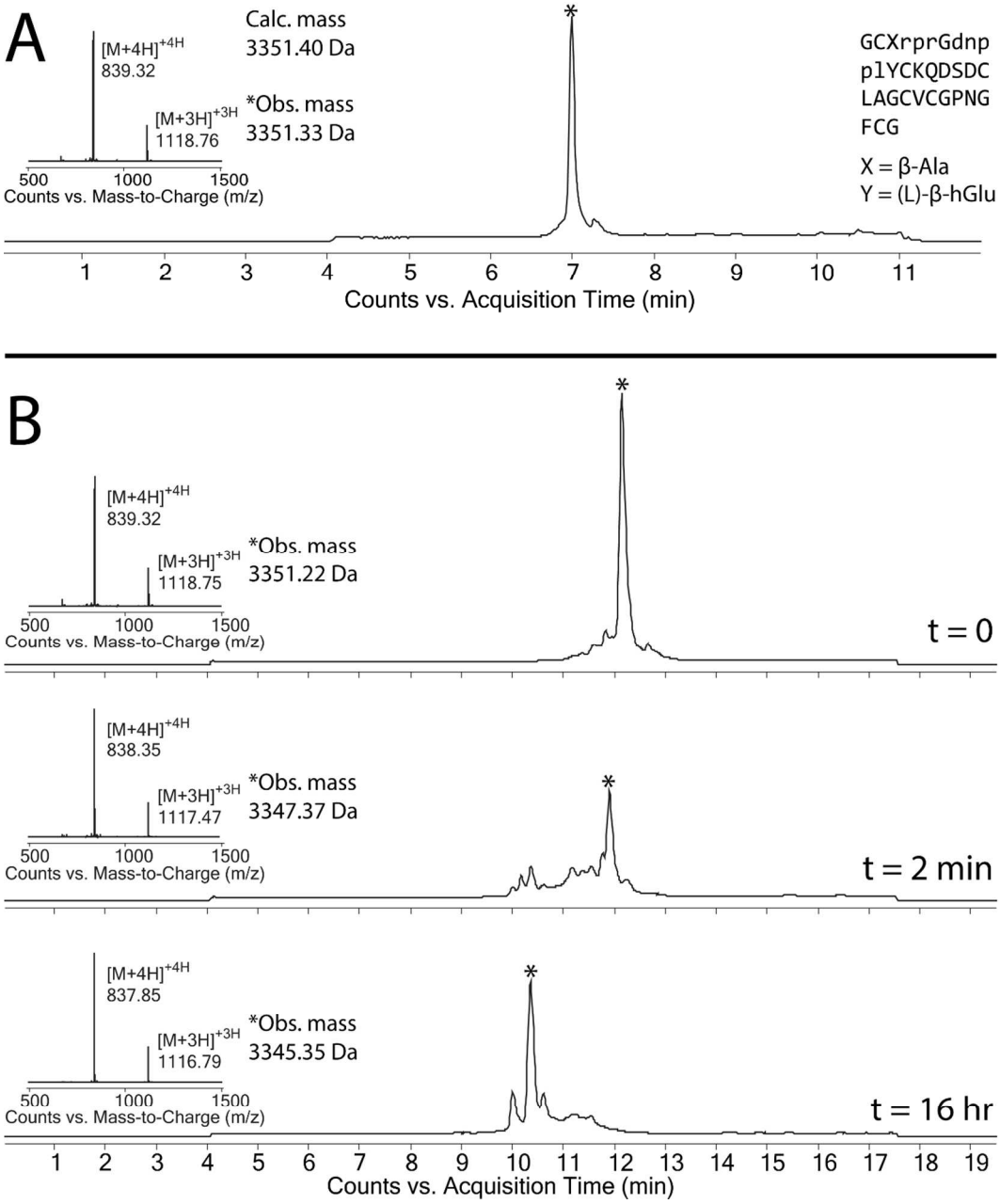

Figure S28: Purification and oxidative folding of $^{L,D(3\text{-}12)}$2.5F(P3β$_A$,T13^{L}β$_{hE}$). LC-MS data (total ion current versus time) of: **A)** reduced polypeptide purified by Method E (27.4 % yield, 4.5 mg isolated, 16.4 mg loaded) and analyzed by Method D; **B)** the oxidative folding of the peptide in A), as monitored by Method C, using an optimized cysteine/cystine containing redox buffer as described earlier. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

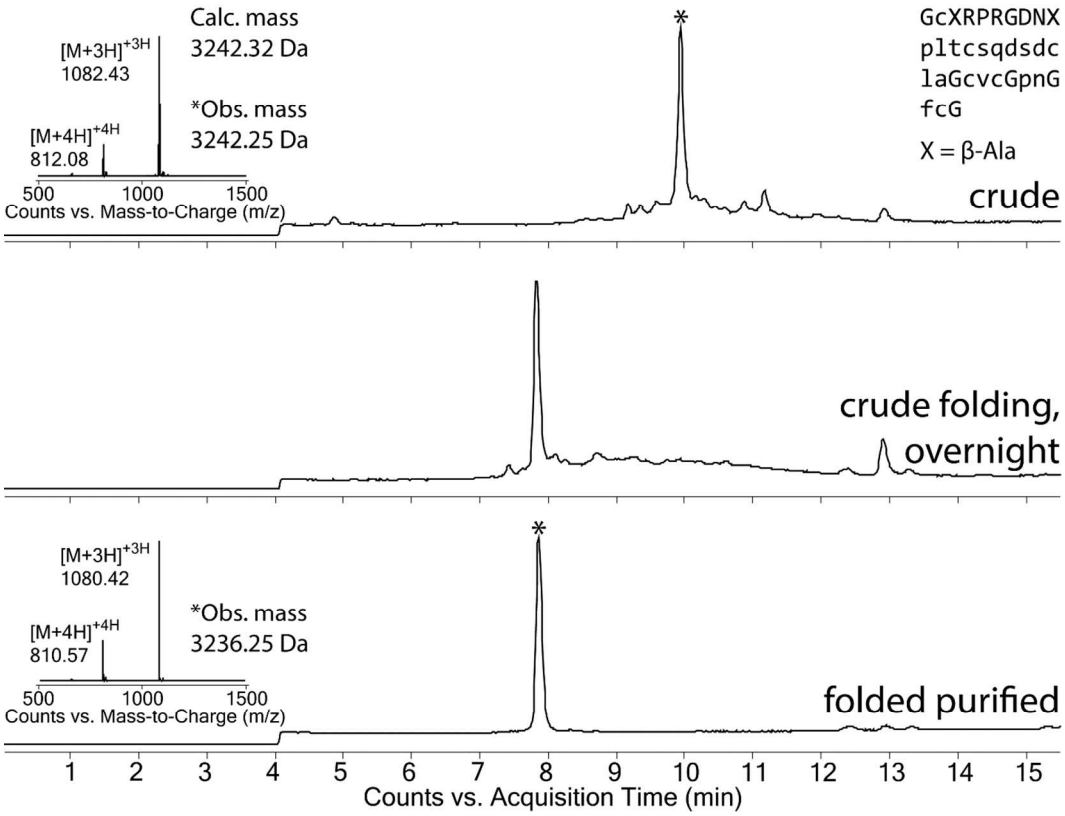

Figure S29: Crude peptide oxidative folding of $^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A,k15s) for NMR. LC-MS data (total ion current versus time) of reduced crude, crude oxidative folding, purified folded peptide for NMR. Folding was achieved using an optimized cysteine/cystine containing redox buffer as described earlier. Purification was achieved using Method E (16.7 % yield, 10.2 mg isolated, 61.2 mg loaded). The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic.

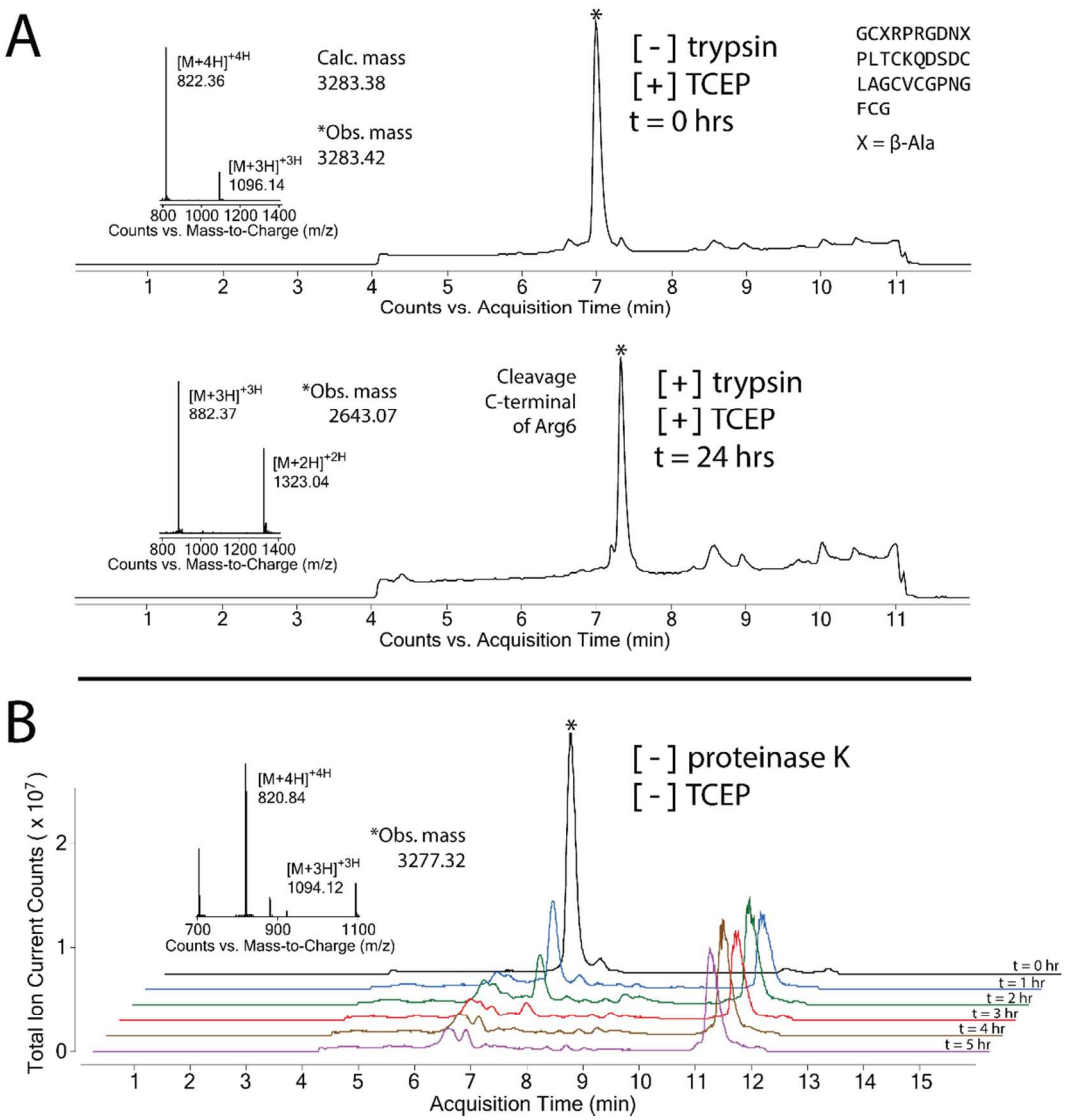

Figure S30: Proteolytic degradation of ^{L,L}2.5F(P3β$_A$,P10β$_A$). LC-MS data (total ion current versus time) of folded, purified peptide subjected to degradation by **A)** trypsin or **B)** proteinase K over time, as monitored by Method B or D, respectively. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic. A) demonstrates cleavage of the loop region by trypsin and B) demonstrates non-specific cleavage of the entire protein by proteinase K.

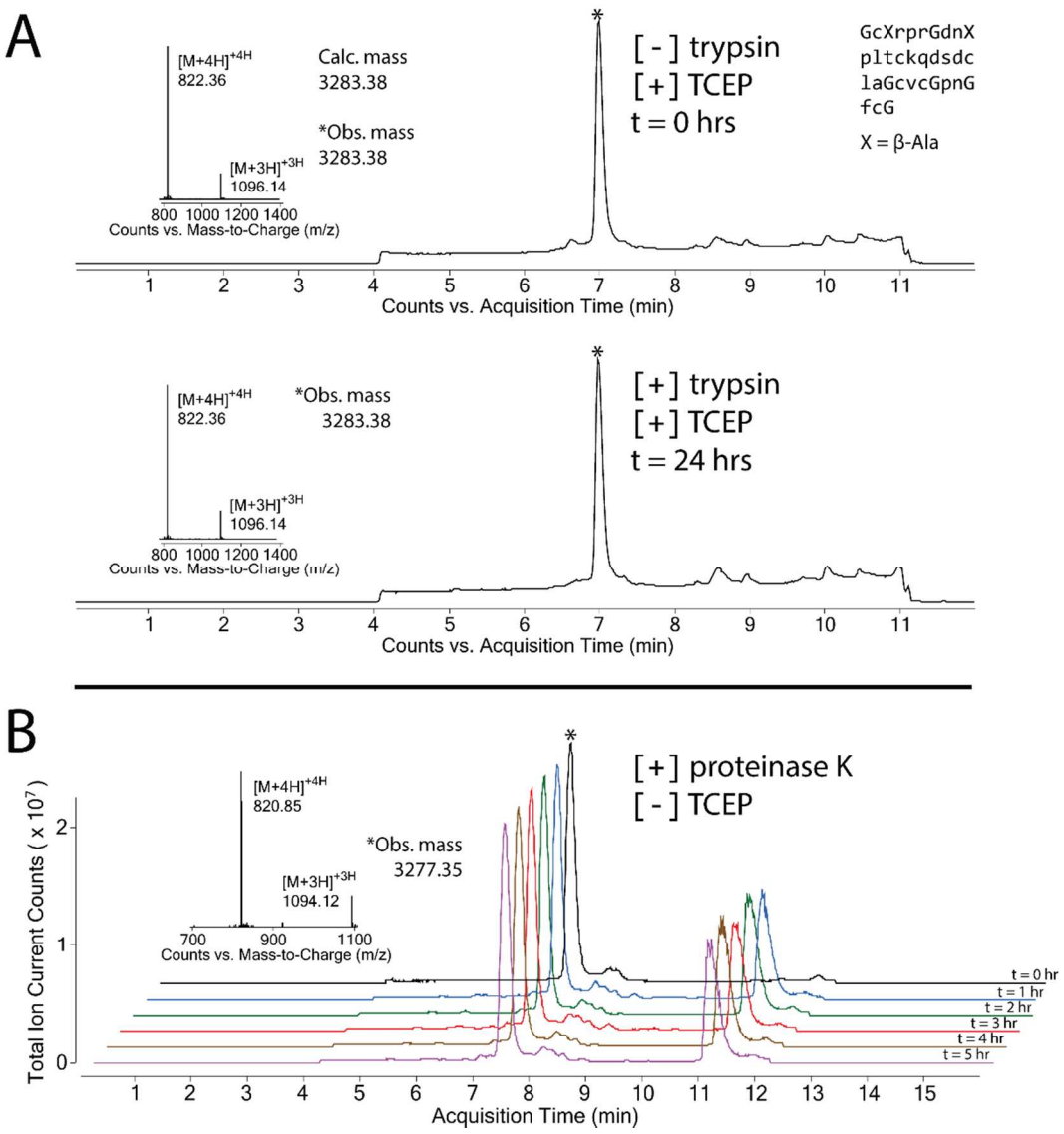

Figure S31: Proteolytic degradation of ^{D,D}2.5F(p3β$_A$,p10β$_A$). LC-MS data (total ion current versus time) of folded, purified peptide subjected to degradation by **A)** trypsin or **B)** proteinase K over time, as monitored by Method B or D, respectively. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic. A) and B) demonstrate that this analog is resistant to proteolysis through the duration of the assay.

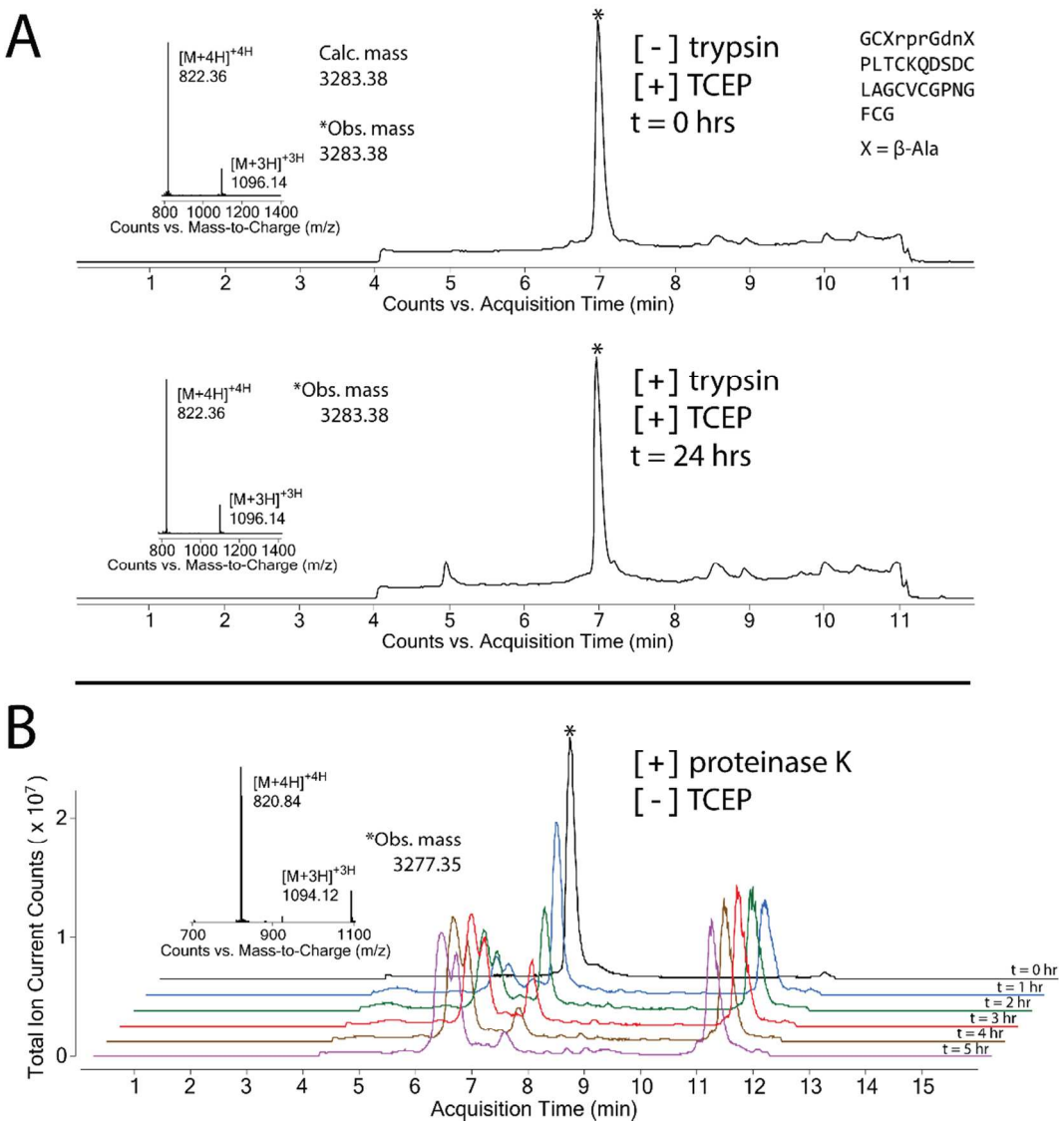

Figure S32: Proteolytic degradation of $^{L,D(3\text{-}10)}$2.5F(P3β$_A$,P10β$_A$). LC-MS data (total ion current versus time) of folded, purified peptide subjected to degradation by **A)** trypsin or **B)** proteinase K over time, as monitored by Method B or D, respectively. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic. A) demonstrates resistance of the loop region to proteolysis by trypsin and B) demonstrates non-specific cleavage of the entire protein by proteinase K, albeit at a slower rate than if the entire protein were to be comprised of (L)-amino acids.

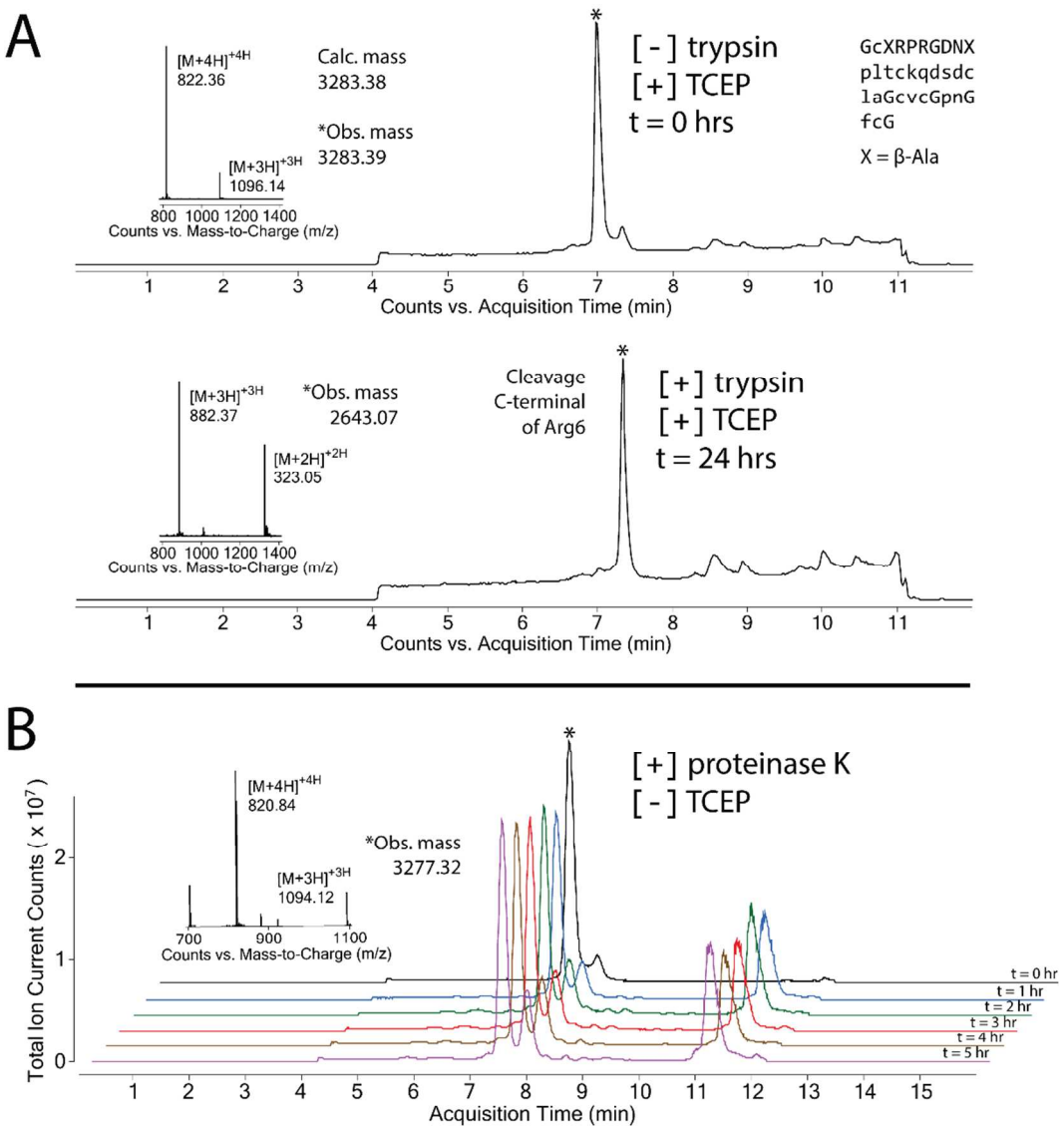

Figure S33: Proteolytic degradation of $^{D,L(3\text{-}10)}$2.5F(p3β$_A$,p10β$_A$). LC-MS data (total ion current versus time) of folded, purified peptide subjected to degradation by **A)** trypsin or **B)** proteinase K over time, as monitored by Method B or D, respectively. The insets containing the charge state series and observed mass correspond to the major product (*) of each individual chromatogram. Observed and calculated masses are monoisotopic. A) demonstrates susceptibility of the loop region to proteolysis by trypsin and B) demonstrates non-specific cleavage of the entire protein by proteinase K, albeit at a slower rate than if the entire protein were to be comprised of (L)-amino acids. The major product in B) is a cleavage product with near identical retention identical time as the starting material. Extracted ion current analysis reveals the starting material is slowly consumed (data not shown).

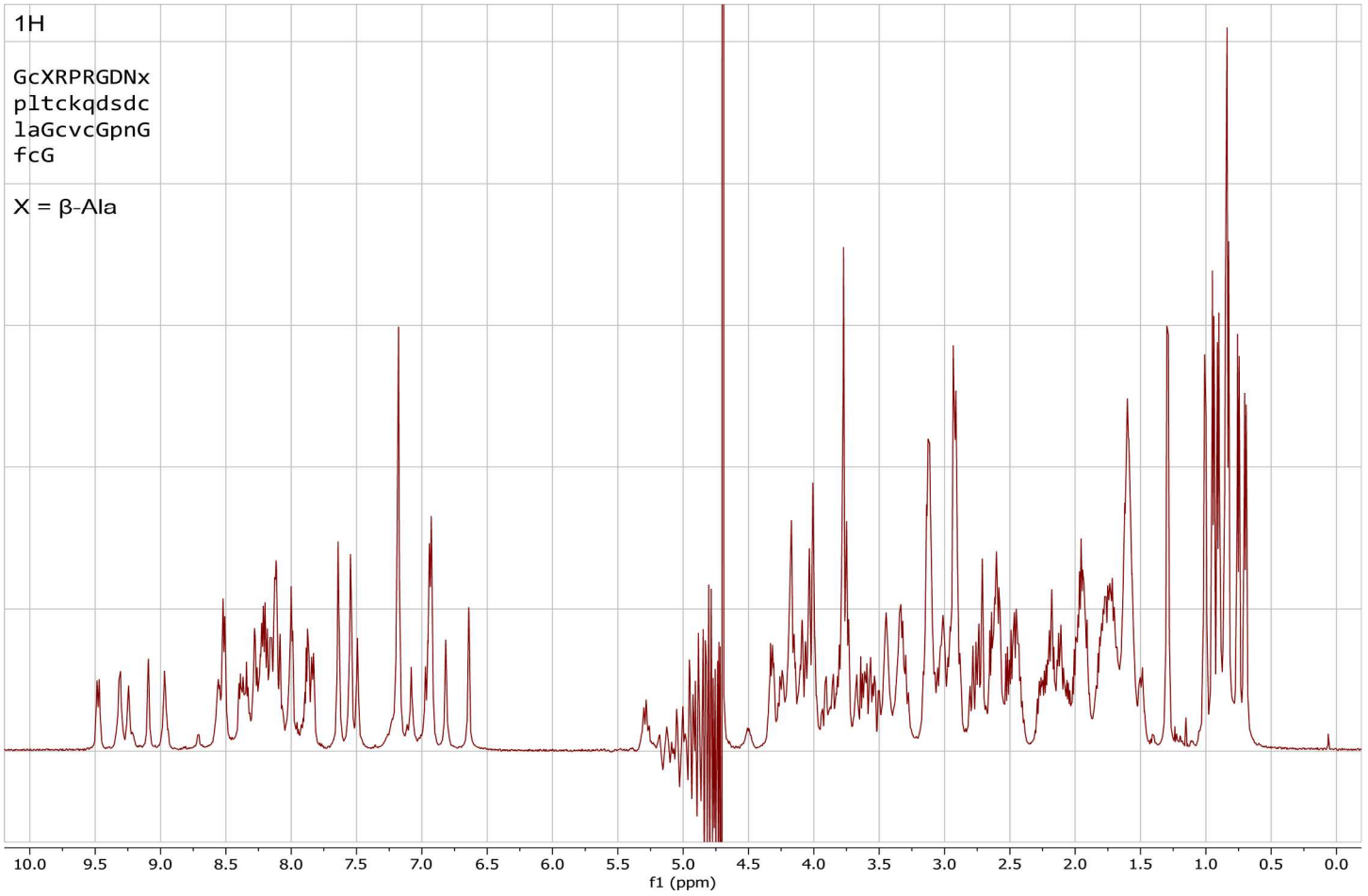

Figure S34: 1D ^{1}H spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β$_A$,p10β$_A$,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

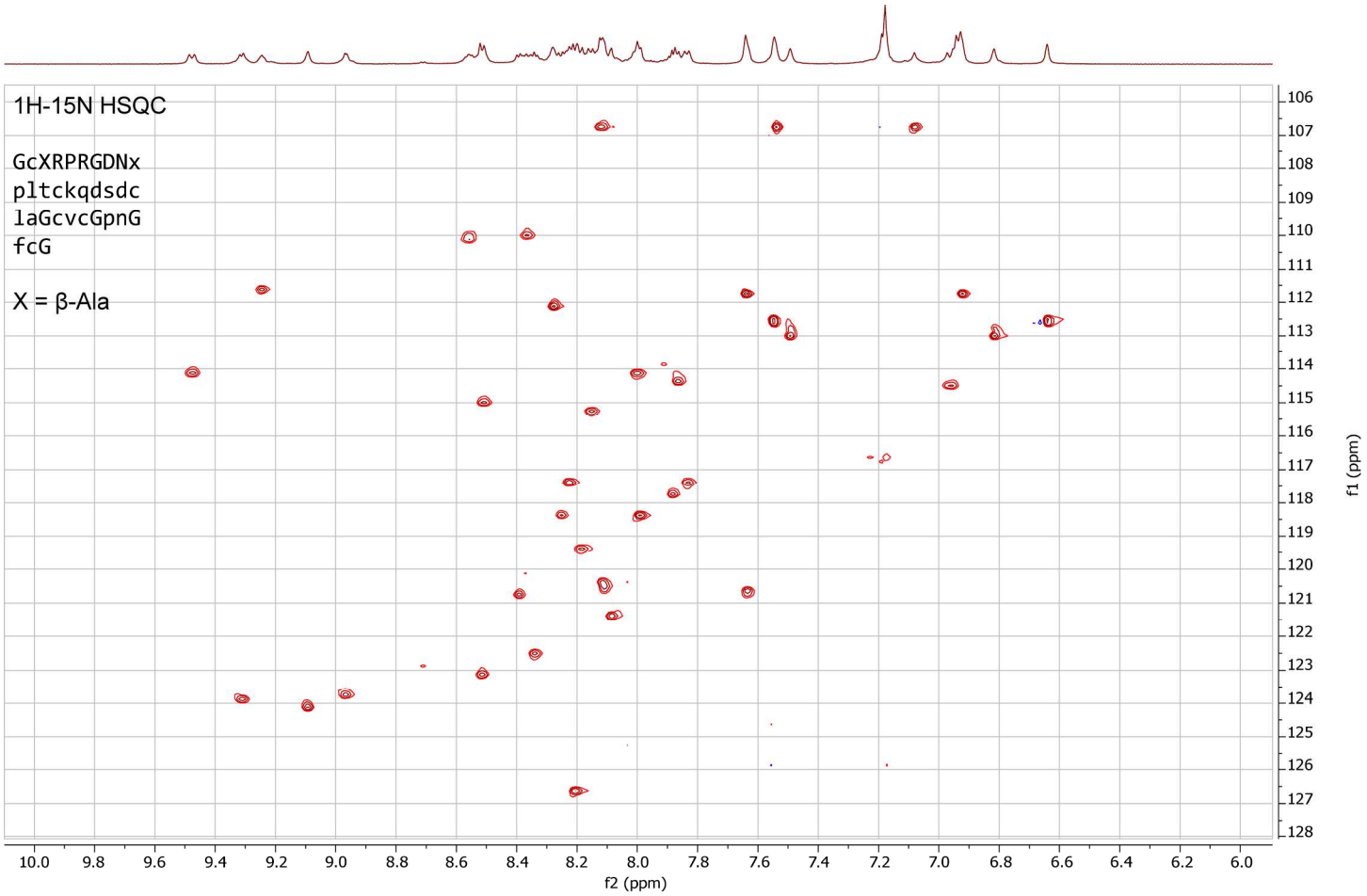

Figure S35: 2D ^{1}H-^{15}N HSQC spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β$_A$,p10β$_A$,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

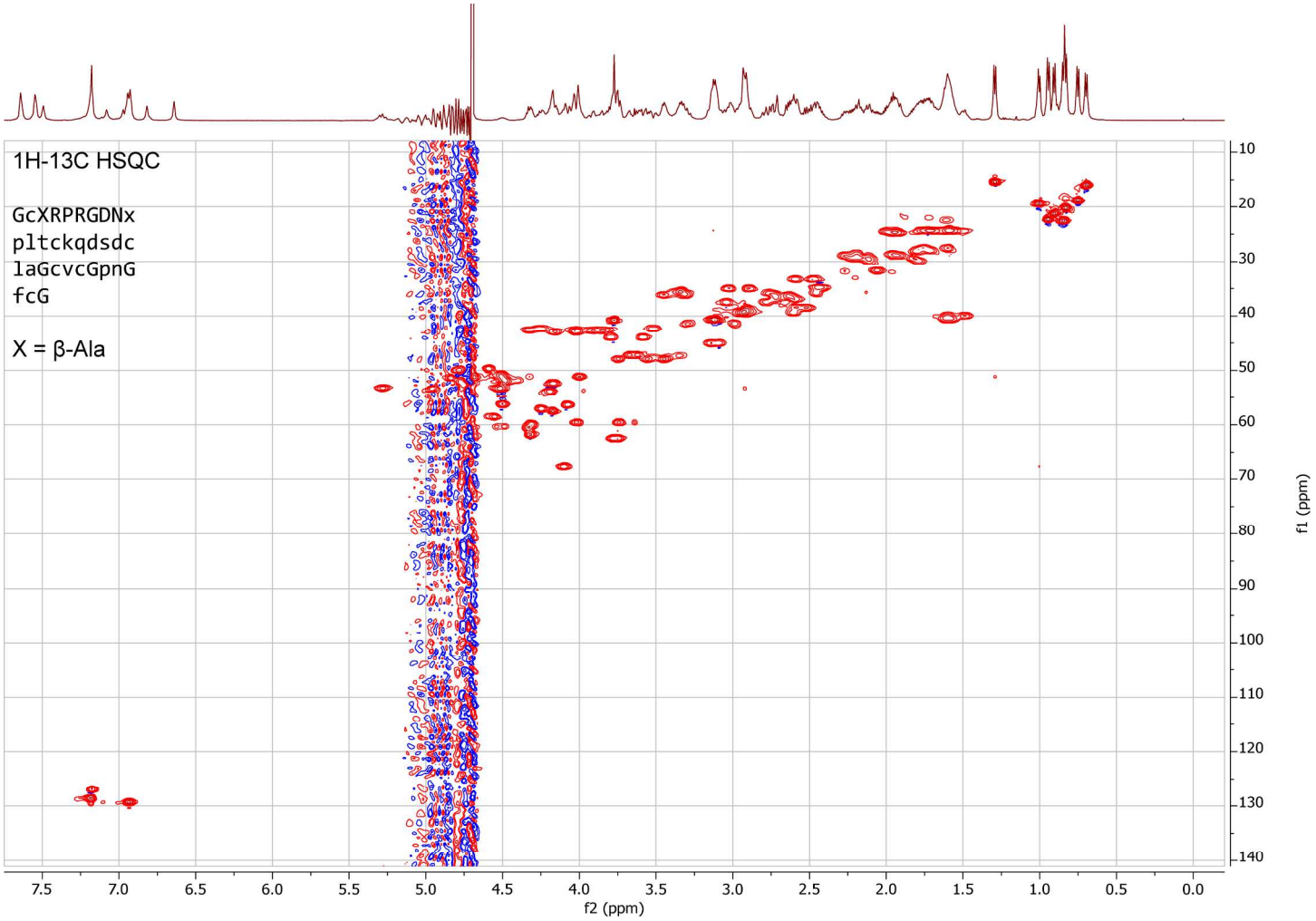

Figure S36: 2D ^{1}H-^{13}C HSQC spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

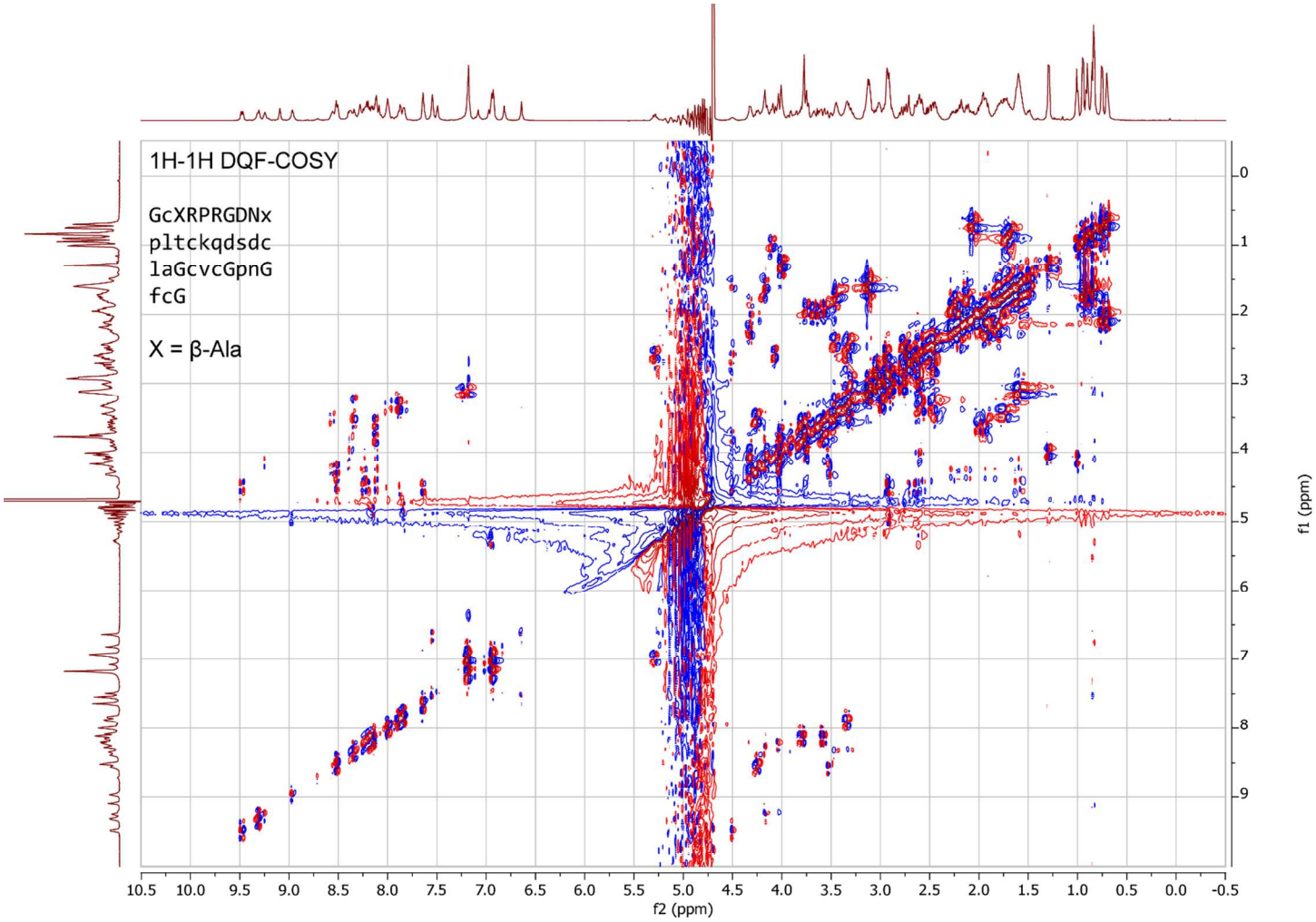

Figure S37: 2D ^{1}H-^{1}H DQF-COSY spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

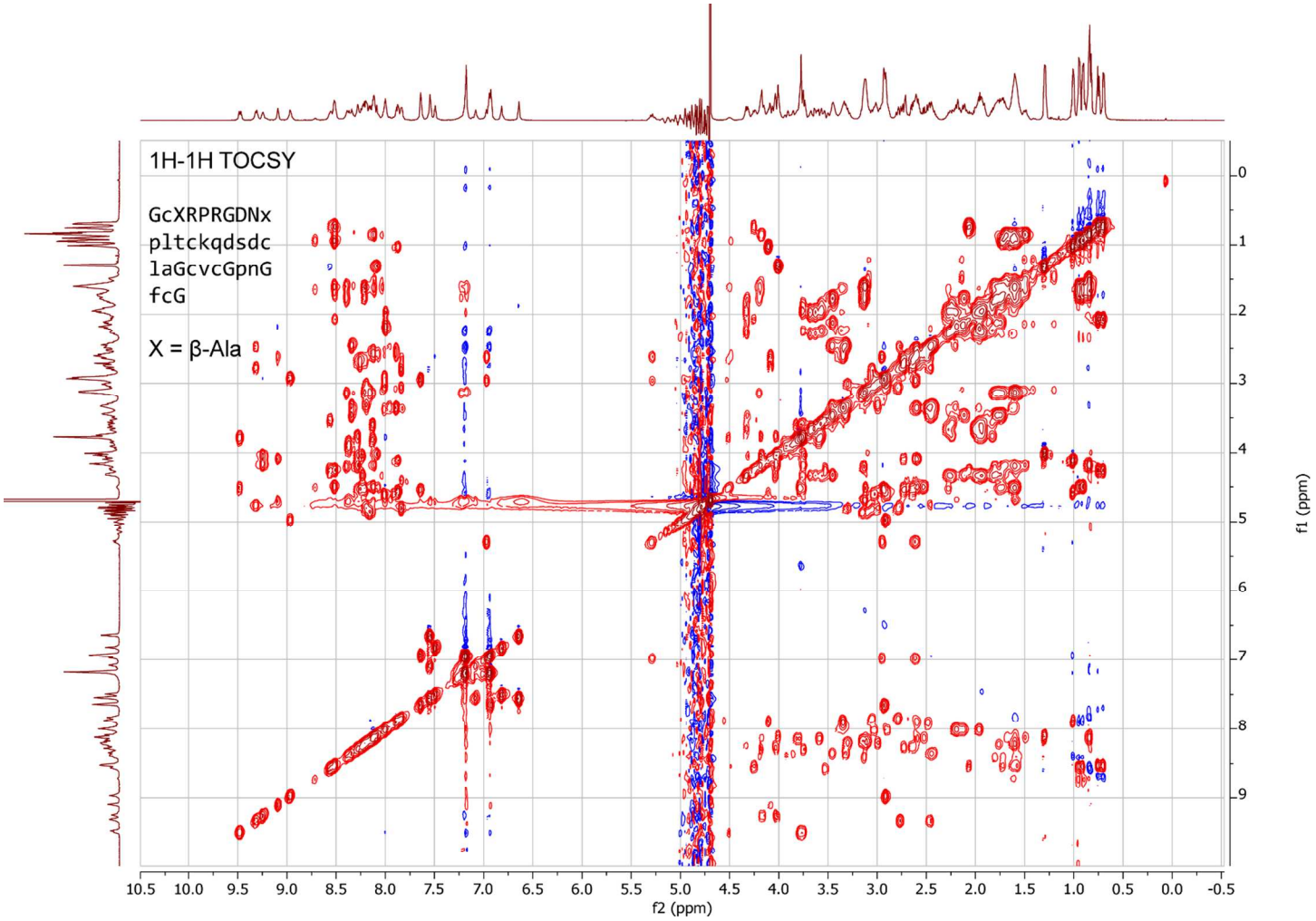

Figure S38: 2D ^{1}H-^{1}H TOCSY spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

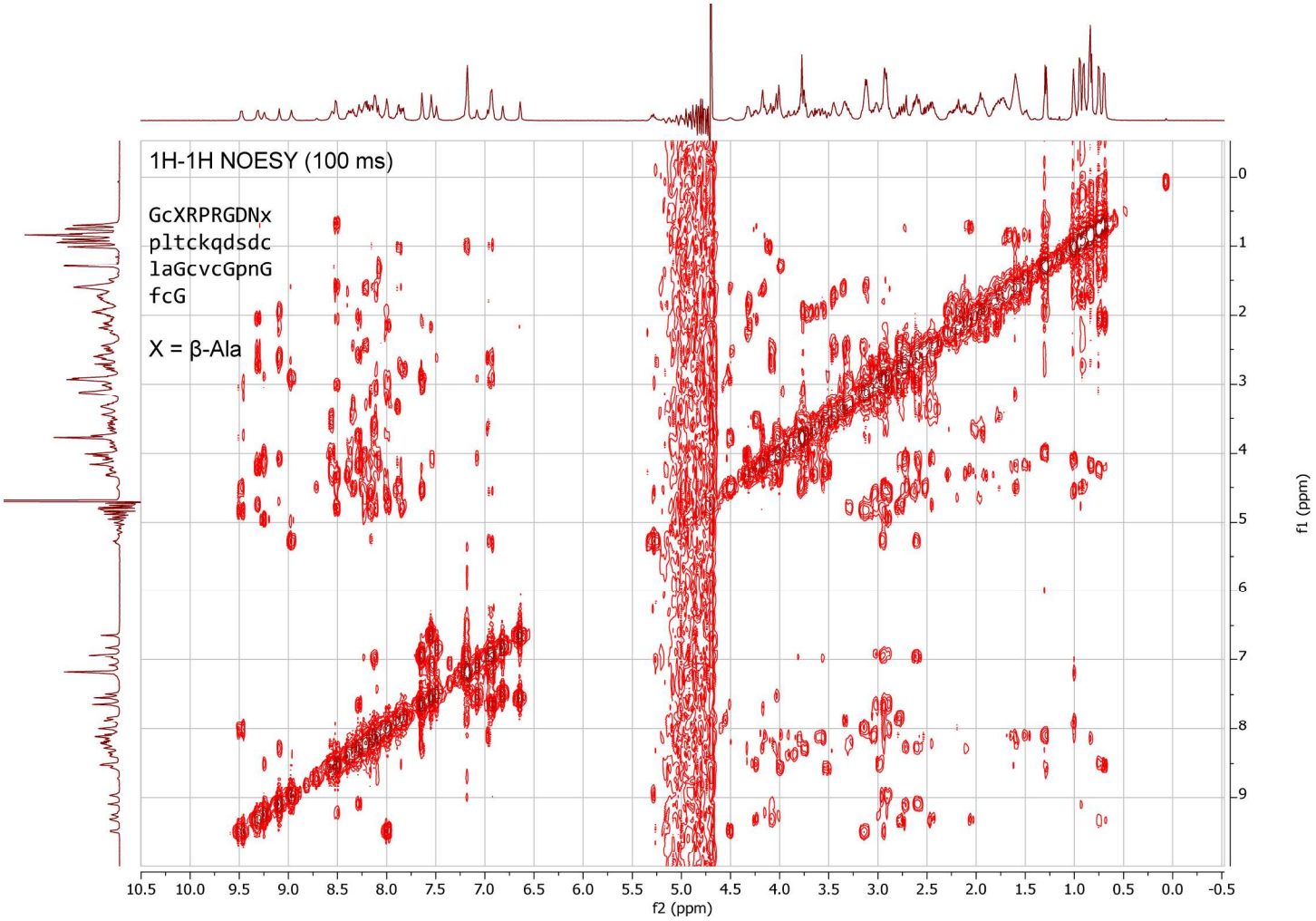

Figure S39: 2D ^{1}H-^{1}H NOESY (100 ms) spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β$_A$,p10β$_A$,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

Figure S40: 2D ^{1}H-^{1}H NOESY (300 ms) spectrum of $^{D,L(3\text{-}10)}$2.5F(p3β_A,p10β_A,k15s). Acquired on a 600 MHz spectrometer at 25 °C. Sample was dissolved to 3 mM in 95/5 H_2O/D_2O containing 20 mM sodium phosphate, pH 6, and 0.2 % (w/v) sodium azide.

6. References

(1) Favel, A., Mattras, H., Coletti-Previero, M. A., Zwilling, R., Robinson, E. A., and Castro, B. (1989) Protease inhibitors from Ecballium elaterium seeds. *Int. J. Pept. Protein Res. 33*, 202–208.

(2) Kimura, R. H., Levin, A. M., Cochran, F. V., and Cochran, J. R. (2009) Engineered cystine knot peptides that bind αvβ3, αvβ5, and α5β1 integrins with low-nanomolar affinity. *Proteins Struct. Funct. Bioinforma. 77*, 359–369.

(3) Nielsen, K. J., Alewood, D., Andrews, J., Kent, S. B. H., and Craik, D. J. (1994) An 1H NMR determination of the three-dimensional structures of mirror-image forms of a Leu-5 variant of the trypsin inhibitor form Ecballium elaterium (EETI-II). *Protein Sci. 3*, 291–302.

(4) Simon, M. D., Heider, P. L., Adamo, A., Vinogradov, A. A., Mong, S. K., Li, X., Berger, T., Policarpo, R. L., Zhang, C., Zou, Y., Liao, X., Spokoyny, A. M., Jensen, K. F., and Pentelute, B. L. (2014) Rapid flow-based peptide synthesis. *ChemBioChem 15*, 713–720.

(5) Mong, S. K., Vinogradov, A. A., Simon, M. D., and Pentelute, B. L. (2014) Rapid total synthesis of DARPin pE59 and barnase. *ChemBioChem 15*, 721–733.

(6) Lee, W., Tonelli, M., and Markley, J. L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics 31*, 1325–1327.

(7) Güntert, P., and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR 62*, 453–471.

(8) Krieger, E., and Vriend, G. (2015) New ways to boost molecular dynamics simulations. *J. Comput. Chem. 36*, 996–1007.

(9) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for

highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435–447.

(10) Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014) PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun. 185*, 604–613.

(11) Jiang, F., Zhou, C. Y., and Wu, Y. D. (2014) Residue-specific force field based on the protein coil library. RSFF1: modification of OPLS-AA/L. *J Phys Chem B 118*, 6983–6998.

(12) Jiang, F., and Wu, Y.-D. (2014) Folding of fourteen small proteins with a residue-specific force field and replica-exchange molecular dynamics. *J. Am. Chem. Soc. 136*, 9536–9539.

(13) Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B 105*, 6474–6487.

(14) Horn, H. W., Swope, W. C., Pitera, J. W., Madura, J. D., Dick, T. J., Hura, G. L., and Head-Gordon, T. (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys. 120*, 9665–9678.

(15) Kratzner, R., Debreczeni, J. E., Pape, T., Schneider, T. R., Wentzel, A., Kolmar, H., Sheldrick, G. M., and Uson, I. (2005) Structure of Ecballium Elaterium Trypsin Inhibitor II (EETI-II): A Rigid Molecular Scaffold. *Acta Crystallogr. Sect. D 61*, 1255–1262.

(16) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

(17) Park, S., Khalili-Araghi, F., Tajkhorshid, E., and Schulten, K. (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys. 119*,

3559–3566.

(18) Jarzynski, C. (1997) Nonequilibrium equality for free energy differences. *Phys. Rev. Lett. 78*, 2690–2693.

(19) Nosé, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys. 52*, 255–268.

(20) Hoover, W. G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A 31*, 1695–1697.

(21) Cheng, A., and Merz, K. M. (1996) Application of the Nosé−Hoover chain algorithm to the study of protein dynamics. *J. Phys. Chem. 100*, 1927–1937.

(22) Lingenheil, M., Denschlag, R., Reichold, R., and Tavan, P. (2008) The “hot-solvent/cold-solute” problem revisited. *J. Chem. Theory Comput. 4*, 1293–1306.

(23) Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys. 52*, 7182–7190.

(24) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463–1472.

(25) Hockney, R. W., Goel, S. P., and Eastwood, J. W. (1974) Quiet high-resolution computer models of a plasma. *J. Comput. Phys. 14*, 148–158.

(26) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577–8593.

(27) Allen, M. P., and Tildesley, D. J. (1989) Computer simulation of liquids. Oxford University Press.