

## Electronic Supplementary Information for “Long-range PEG Stapling: Macrocyclization for Increased Protein Conformational Stability and Resistance to Proteolysis”

Qiang Xiao, Dallin S. Ashton, Zachary B. Jones, Katie P. Thompson and Joshua L. Price

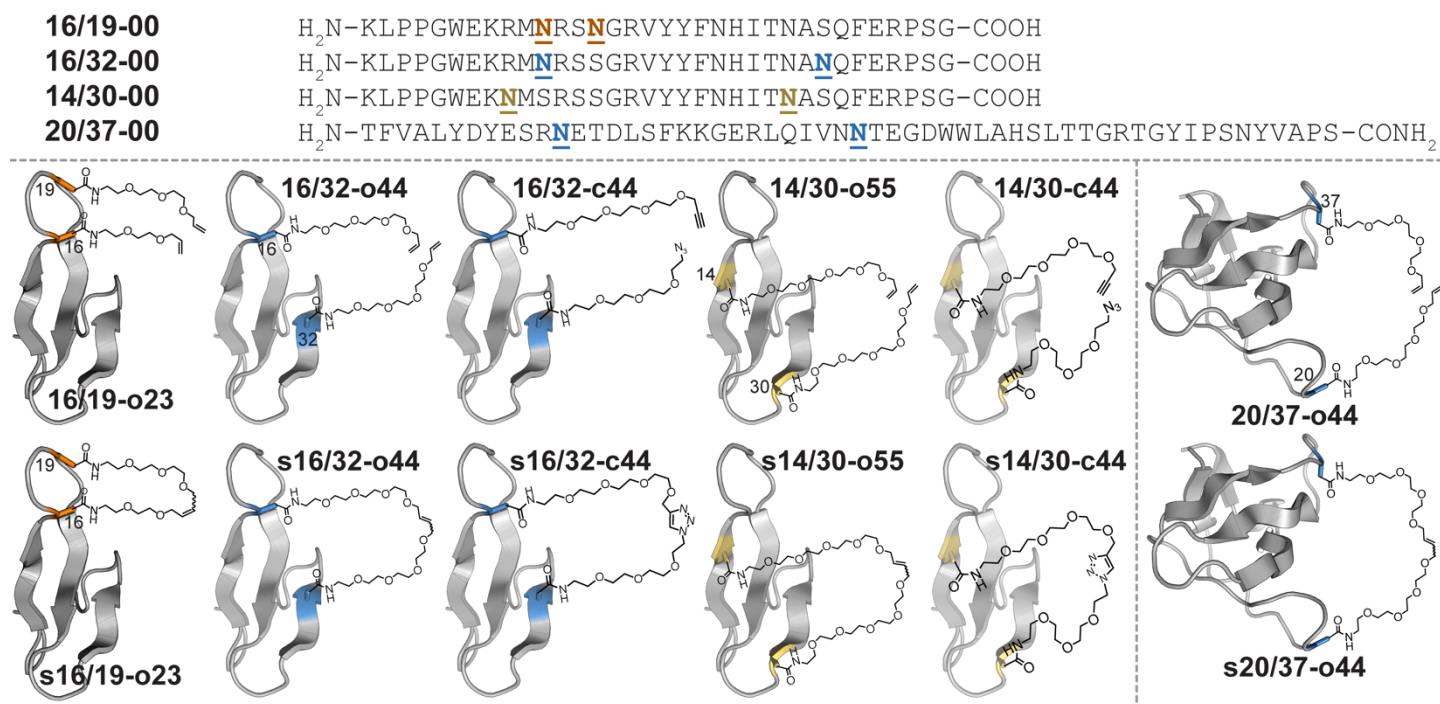
Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, United States

### Safety and Hazards:

The work described below was carried out using standard laboratory PPE, including a lab coat, nitrile gloves, and goggles. Work with trifluoroacetic acid or with thiols was conducted in a fume hood.

### 1. Structures and sequences of WW, SH3, and GCN4 variants

The structures and sequences of newly synthesized WW and SH3 variants are shown in Figure S1. WW variants **16/19-00**, **16/19-o23**, **s16/19-o23**, **16/32-00**, **16/32-o44**, **s16/32-o44** and SH3 variant **SH3 20/37-00** mentioned in the main text, those structures and sequences were reported before by our lab.<sup>1,2</sup>



**Figure S1.** Sequences and structures of olefin-stapled WW variants **s16/19-o23**, **s16/32-o44** and **s14/30-o55**; triazole-stapled WW variants **s16/32-c44** and **s14/30-c44**; and olefin-stapled SH3 variant **s20/37-o44** and their non-stapled and non-PEGylated counterparts **N** represents a PEG-modified Asn residue; the PEG oligomer(s) within each variant have the number of ethylene oxide units and the olefin, azide, alkyne, or triazole functional groups as indicated in the structural drawings.



**Figure S2.** Acidic GCN4 monomer **27A-c4**, basic GCN4 monomer **29A'-c0**, noncovalent GCN4 heterodimer **27/29'-c40**, and its triazole-stapled counterpart **27/29'-c40**. Also shown are cysteine-containing acidic GCN4 monomer **27-c4**, basic GCN4 monomer **29'-c0**, disulfide-bound GCN4 heterodimer **d27/29'-c40**, and its triazole-stapled counterpart **sd27/29'-c40**. **X** represents propargyl glycine; **N** represents a PEG-modified Asn residue; the PEG oligomer(s) within each variant have the number of ethylene oxide units indicated in the structural drawings

## 2. Synthesis of WW, SH3 and GCN4 variants

**Peptide Synthesis:** WW variants **16/32-00**, **16/32-o44**, **s16/32-o44**, and SH3 variant **20/37-00** were prepared previously. WW variants **16/32-c44**, **14/30-00**, **14/30-o55**, and **14/30-c44** were synthesized as C-terminal acids on Fmoc-Gly-Wang LL resin (EMD Biosciences) and SH3 variant **20/37-o44** and GCN4 variants **27-c4**, **29'-c0**, and their Cys33Ala mutants **27A-c4**, **29A'-c0** were prepared as C-terminal amides on and Rink amide MBHA LL resin (EMD Biosciences), by microwave-assisted solid phase peptide synthesis using a standard Fmoc N $\alpha$  protection strategy as described previously.<sup>2</sup> Fmoc-protected amino acids were purchased from Advanced Chem Tech, except for the PEGylated asparagine derivatives, which were synthesized as described in section 7 below. Unstapled WW, SH3, and GCN4 variants were cleaved from resin and 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 trifluoroacetic acid. Peptide identity was confirmed by electrospray ionization time-of-flight mass spectrometry. We used an analogous approach to prepare the stapled WW variants and SH3 variant, except we reduced the resin loading (i.e. the number of Fmoc-protected amino groups on resin) by approximately 50% using a 1:1 mixture of Fmoc- and Boc-protected amino acids during the first coupling reaction.

**Stapling via olefin metathesis:** We prepared olefin-stapled variants **s14/30-o55** and **s20/37-o44** from their resin-bound non-stapled precursors **14/30-o55** and **20/37-o44** via on-resin olefin metathesis, as we have done previously, except we heated the reaction mixture to 60 °C. We monitored reaction completeness by

microcleavage and ESI-TOF MS. Following the reaction, we cleaved and purified **s14/30-o55** and **s20/37-o44** as described above.

**Disulfide formation in the GCN4 heterodimer:** We prepared the disulfide-bonded GCN4 heterodimer **d27/29'-c40** (notebook number **QX22031**) by mixing its purified cysteine-containing precursors **27-c4** (notebook number **QX21971**) and **29'-c0** (notebook number **QX21973**) in a 1:1 ratio (total peptide concentration of 8 mg/mL) in an aqueous solution of ammonium bicarbonate with exposure to air for 3 h. Reaction completeness was monitored by analytical HPLC; GCN4 variant **d27/29'-c40** was then purified by preparative HPLC and characterized by ESI-TOF MS.

**Stapling via copper-catalyzed azide/alkyne cycloaddition (CuAAC):** We prepared triazole-stapled variants **s16/32-c44**, **s14/30-c44**, **sd27/29'-c40**, and **s27/29'-c40** (i.e., the non-disulfide-bound counterpart of **sd27/29'-c40**) by stirring their purified non-stapled precursors **16/32-c44**, **14/30-c44**, **d27/29'-c40**, and **27/29'-c40** (i.e., the non-covalent GCN4 heterodimer formed by a 1:1 mixture of **27A-c4** and **29A'-c0**) in 2:1 (v/v) water/*tert*-butanol with 10 eq. copper sulfate pentahydrate and 10 eq. sodium ascorbate at room temperature for 2 hours. The reaction was monitored by analytical HPLC, where we observed small changes in retention time upon stapling. The triazole-stapled variants were purified via preparative HPLC.

The successful conversion of non-covalent heterodimer **27/29'-c40** into triazole-stapled **s27/29'-c40** was readily confirmed by ESI-TOF MS; the non-covalent heterodimer **27/29'-c40** separates into its component monomers **27A-c4** (notebook number **QX22091**) and **29A'-c0** (notebook number **QX22092**) in the mass spectrometer, whereas triazole-stapled **s27/29'-c40** does not. However, the CuAAC reaction does not change the mass of the monomeric triazole-stapled variants **s16/32-c44**, **s14/30-c44**, **sd27/29'-c40** (notebook number **QX2204**) relative to their non-stapled azide/alkyne precursors **16/32-c44**, **14/30-c44**, and **d27/29'-c40** (notebook number **QX22031**). To confirm the completion of the CuAAC reaction, we subjected each azide-containing variant (**16/32-c44**, **14/30-c44**, and **d27/29'-c40**) and its triazole-stapled counterpart (**s16/32-c44**, **s14/30-c44**, and **sd27/29'-c40**) separately to a solution of dithiothreitol (DTT, 31 mg) and diisopropylethylamine (DIEA, 17  $\mu$ L) in 100  $\mu$ L water, followed by stirring for 1 h. Presence of the azide groups in non-stapled **16/32-c44** (notebook number **QX21852**), **14/30-c44**, and **d27/29'-c40** was confirmed by their conversion to amino groups under reducing conditions, as detected by ESI-TOF MS (i.e., a loss of N<sub>2</sub> and the addition of two hydrogens). Presence of the triazole group in **s16/32-c44** (notebook number **QX2188**), **s14/30-c44**, and **sd27/29'-c40** were confirmed by the absence of this mass change under reducing conditions. The results of this analysis are shown below in Figures S3–S8.

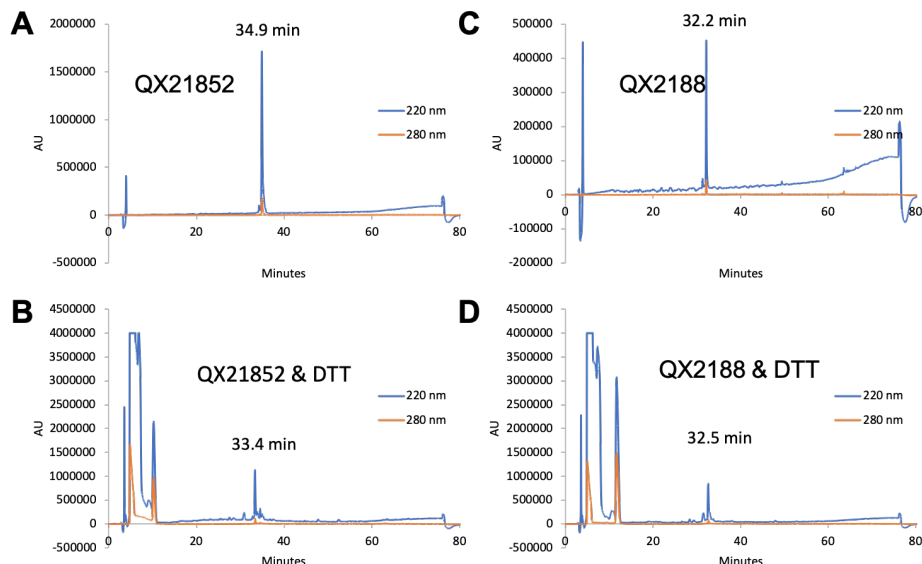
**Choice of PEG length in variants s16/19-o23, s16/32-o44, s16/32-c44, s14/30-o55, and s14/30-c44.** We previously found that the PEG linker between positions 16 and 19 in WW variant **s16/19-o23** provided optimal stabilization relative to linkers of other lengths, including one with two four-unit PEGs.<sup>1</sup> In the same previous

publication, we used two four-unit PEG linkers to connect positions 16 and 32 owing to the increased distance between positions 16 and 32 relative to that between positions 16 and 19 in the crystal structure of the parent WW domain from which these variants were derived. We did not attempt any further optimization of linker length between positions 16 and 32. Here, we have continued to use two four-unit PEG linkers to facilitate direct comparison between the impact of click-stapling on variant **s16/32-c44** vs. olefin-stapling on variant **s16/32-o44**.

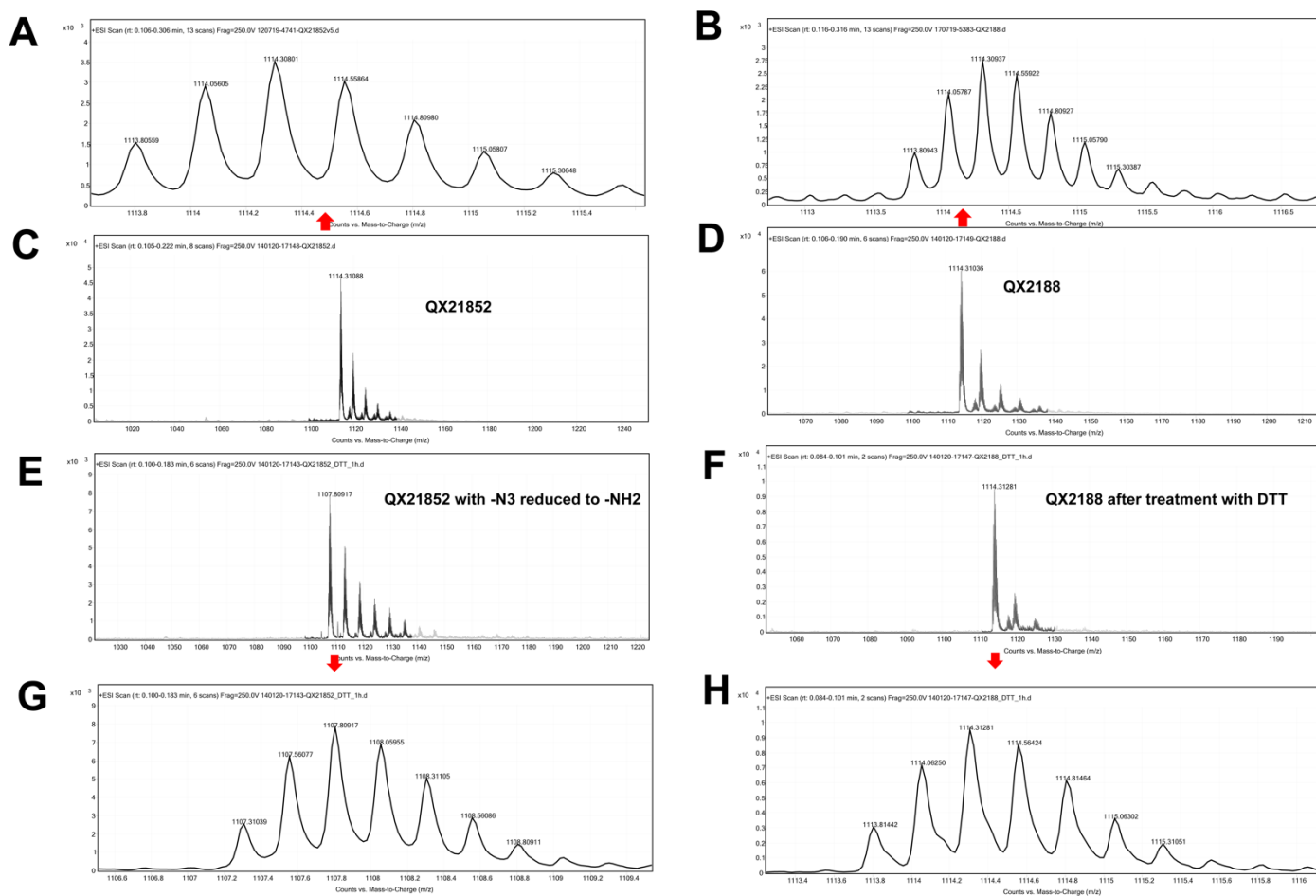
In preparing the current manuscript, we initially attempted to use two four-unit PEG linkers to staple positions 14 and 30. We successfully prepared PEGylated variant **14/30-44** and its non-PEGylated counterpart **14/30-00**. However, several attempts at on-resin olefin metathesis failed to provide sufficient amounts of **s14/30-o44** for variable temperature CD experiments, precluding us from assessment of the impact of this particular staple on WW conformational stability. We hypothesized that the on-resin olefin metathesis reaction might have failed because the two four-unit PEGs were not sufficiently long to bridge the distance between positions 14 and 30 in the ensemble of conformations adopted by the protected resin-bound protein. Accordingly, we prepared a new olefin-terminated Asn-PEG residue with five ethylene oxide units instead of four. We incorporated this residue at positions 14 and 30 to generate WW variant **14/30-o55**. This time, on-resin olefin-metathesis provided sufficient amounts of stapled WW variant **s14/30-o55** for variable temperature CD experiments.

Click-stapling efforts at these positions occurred before we knew that our efforts to obtain **s14/30-o44** would be unsuccessful. Accordingly, we prepared WW variant **14/30-c44** (in which an Asn-linked propargyl-terminated four-unit PEG occupies position 14, whereas an Asn-linked azide-terminated four-unit PEG occupies position 30). The click reaction was carried out in a 2:1 solution of water/*tert*-butanol and successfully converted the purified non-stapled precursor **14/30-c44** into its stapled counterpart **s14/30-c44**. We speculate that the conformation adopted by deprotected purified **14/30-c44** in this water/*tert*-butanol solution allowed positions 14 and 30 to be close enough to each other that the click reaction could succeed. Unfortunately, our efforts to carry out olefin metathesis on purified deprotected **14/30-o44** in aqueous solution have not yet been successful.

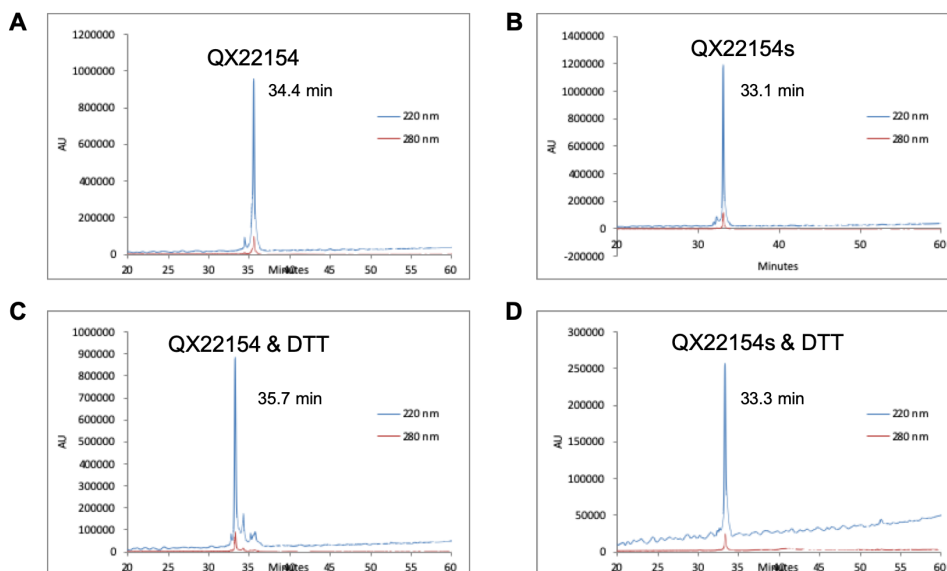
We realize that comparing the staple-based stabilization of variant **s14/30-c44** vs. **s14/30-o55** conflates the impact of staple chemistry with the impact of different PEG lengths. Ideally, we would have prepared variants **14/30-c55** and **s14/30-c55** to facilitate a more direct comparison. However, at this point comparison between variants **s16/32-o44** and **s16/32-c44** (where linker length stayed constant) had already revealed that click-stapling provide similar levels of stabilization as does olefin-stapling (see Table 1 in the main text). The comparison between variants **s14/30-c44** vs. **s14/30-o55** did not contradict this finding, especially in light of previous observations where changes in PEG-length beyond three or four-units has a minimal effect.<sup>2</sup> We concluded that the limited insights to be gained by generating variants **14/30-c55** and **s14/30-c55** did not sufficiently justify the extensive effort that would have been required to prepare the two new Asn-PEGs comprised of propargyl or azide-terminated five-unit PEGs.



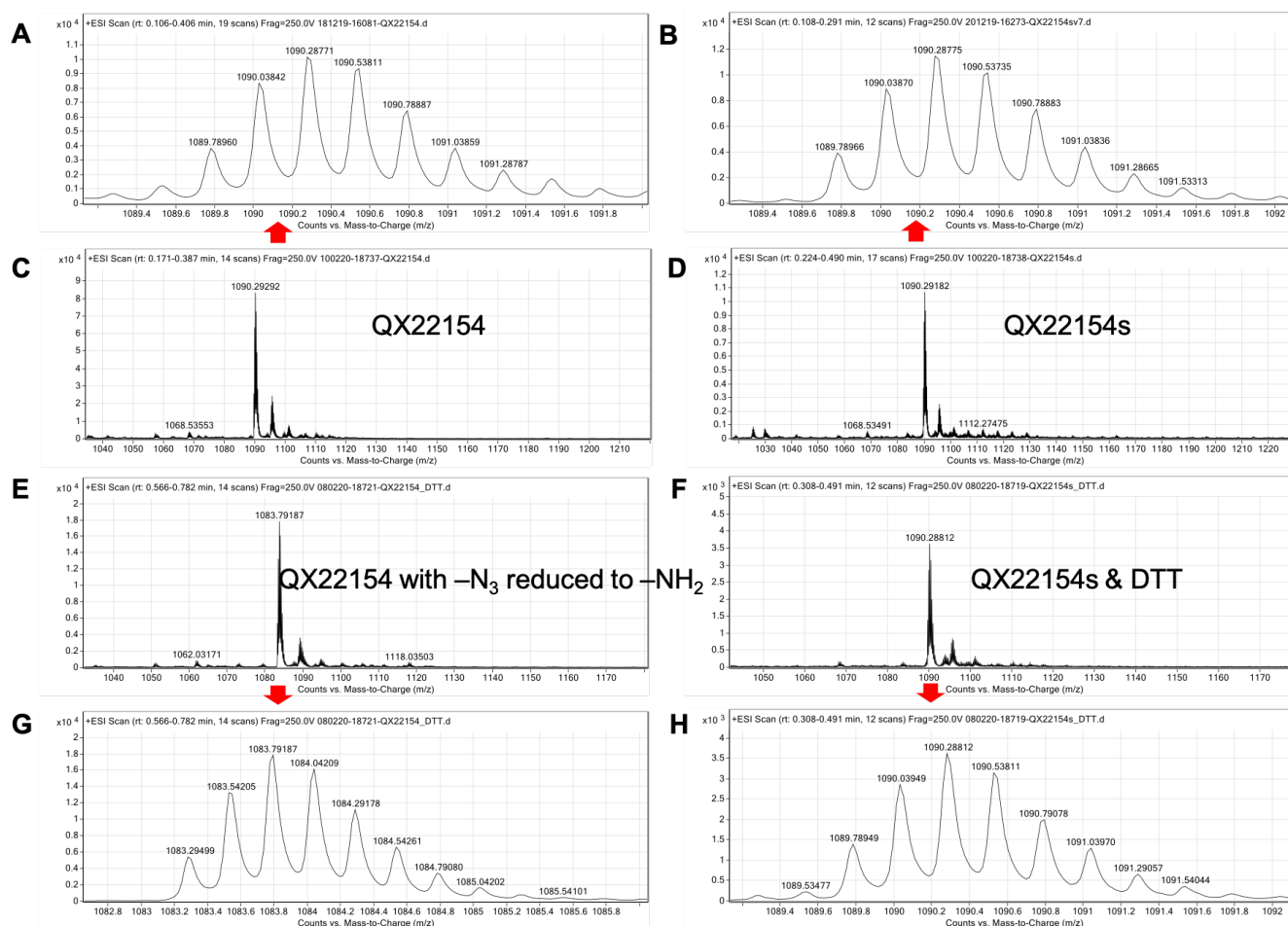
**Figure S3.** Analytical HPLC retention time for WW variant **16/32-c44** (notebook number **QX21852**) before (A) and after (B) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h; and for its triazole-stapled counterpart **s16/32-c44** (notebook number **QX2188**) before (C) and after (D) exposure to reducing conditions for 1 h. Peptide solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



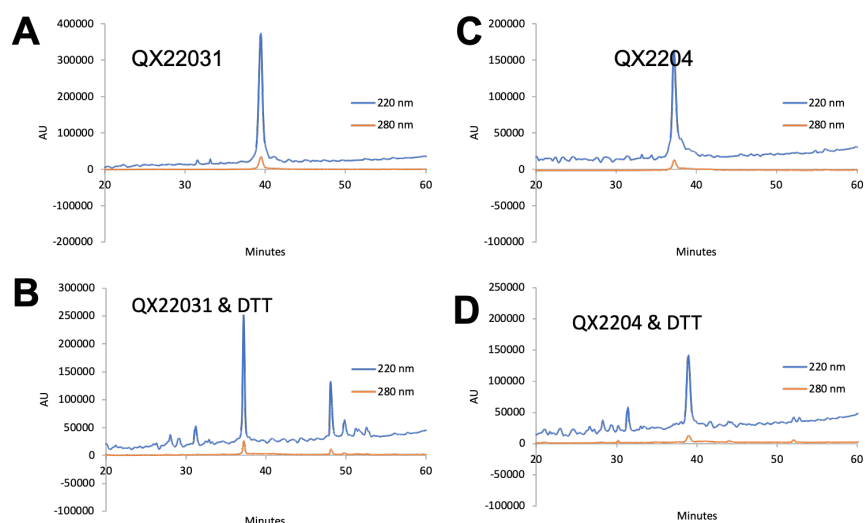
**Figure S4.** ESI-TOF MS data for azide-containing WW variant **16/32-c44** (notebook number **QX21852**) before (A,C) and after (E,G) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h. Expected  $[M+4H^+]/4 = 1113.82$  Da for **16/32-c44** prior to reduction; expected  $[M+4H^+]/4 = 1107.32$  Da after reduction of the azide at position 32 to the corresponding amine. Also shown are ESI-TOF MS data for triazole-stapled **s16/32-c44** (notebook number **QX2188**) before (B,D) and after (F,H) exposure to reduction conditions for 1 h. Expected  $[M+4H^+]/4 = 1113.82$  Da for triazole-stapled **s16/32-c44** both before and after reduction.



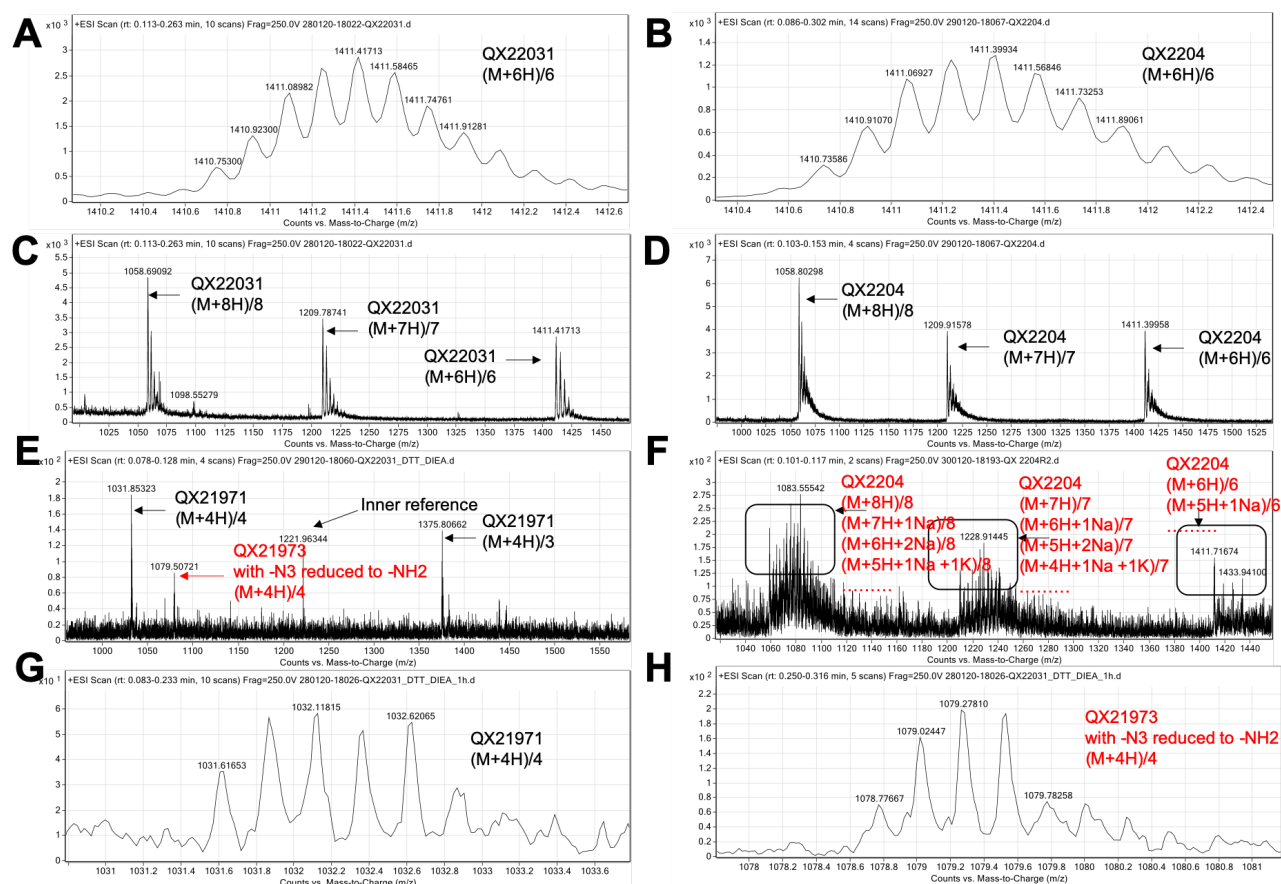
**Figure S5.** Analytical HPLC retention time for WW variant **14/30-c44** (notebook number **QX22154**) before (A) and after (B) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h; and for its triazole-stapled counterpart **s14/30-c44** (notebook number **QX22154s**) before (C) and after (D) exposure to reducing conditions for 1 h. Peptide solution was injected onto a C18 analytical column and eluted using a linear gradient of 10–60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



**Figure S6.** ESI-TOF MS data for azide-containing WW variant **14/30-c44** (notebook number **QX22154**) before (A,C) and after (E,G) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h. Expected  $[M+4H^+]/4 = 1089.80$  Da for **14/30-c44** prior to reduction; expected  $[M+4H^+]/4 = 1083.30$  Da after reduction of the azide at position 30 to the corresponding amine. Also shown are ESI-TOF MS data for triazole-stapled **s14/30-c44** (notebook number **QX22154s**) before (B,D) and after (F,H) exposure to reducing conditions for 1 h. Expected  $[M+4H^+]/4 = 1089.80$  Da for triazole-stapled **s14/30-c44** both before and after reduction.



**Figure S7.** Analytical HPLC retention time for disulfide-bound GCN4 heterodimer **d27/29'-c40** (notebook number **QX22031**) before (A) and after (B) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h; and for its triazole-stapled counterpart **sd27/29'-c40** (notebook number **QX2204**) before (C) and after (D) exposure to reducing conditions for 1 h. Peptide solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B = MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



**Figure S8.** ESI-TOF MS data for disulfide-bound GCN4 heterodimer **d27/29'-c40** (notebook number **QX22031**) before (A,C) and after (E,G,H) exposure to reducing conditions (aqueous DTT and DIEA) for 1 h. Expected  $[M+6H]^+/6 = 1410.59$  Da for **d27/29'-c40** prior to reduction. Following reduction, **d27/29'-c40** splits into its component peptides: alkyne-containing **27-c4** (notebook number **QX21971**); expected  $[M+4H]^+/4 = 1031.62$  Da and **29'-c0** (notebook number **QX21973**) in which the azide at position 29' has been reduced to an amine (expected  $[M+4H]^+/4 = 1078.78$  Da). Also shown are ESI-TOF MS data for triazole-stapled GCN4 variant **sd27/29'-c40** before (B,D) and after (F) exposure to reducing conditions. Expected  $[M+6H]^+/6 = 1410.59$  Da for **sd27/29'-c40** prior to reduction. After reduction, the triazole staple should be intact, but the disulfide bond should not; expected  $[M+6H]^+/6 = 1410.92$  Da.

### 3. Global Fitting of Variable Temperature CD Data

Following purification and characterization (as described above), the conformational stability of stapled and unstapled WW, SH3, and GCN4 variants was assessed by variable-temperature circular dichroism spectropolarimetry. Data from three replicate variable temperature CD experiments were each fit globally to a model for a two-state thermally induced unfolding transition using the program Mathematica (Wolfram Research).

We used a model based on two-state monomer folding for WW variants **16/19-00**, **16/19-o23**, **s16/19-o23**, **16/32-c44**, **s16/32-c44**, **14/30-00**, **14/30-o55**, **s14/30-o55**, **14/30-c44** and **s14/30-c44**; SH3 variant **s20/37-o44**; and GCN4 variants **d27/29'-c40**, **sd27/29'-c40**, and **s27/29'-c40**. This model is described in equation S1:

$$[\theta] = \frac{[K \cdot (a_n + b_n \cdot T) + (c_n + d_n \cdot T)]}{1 + K}, \quad (\text{S1})$$

where  $T$  is temperature in Kelvin,  $a_n$  is the  $y$ -intercept and  $b_n$  is the slope of the pre-transition baseline for melt  $n$  ( $a_1$  and  $b_1$  for replicate 1,  $a_2$  and  $b_2$  for replicate 2,  $a_3$  and  $b_3$  for replicate 3, etc.);  $c_n$  is the  $y$ -intercept and  $d_n$  is the slope of the post-transition baseline for replicate  $n$  ( $c_1$  and  $d_1$  for replicate 1,  $c_2$  and  $d_2$  for replicate 2,  $c_3$  and  $d_3$  for replicate 3, etc.); and  $K$  is the temperature-dependent folding equilibrium constant.  $K$  is related to the temperature-dependent free energy of folding  $\Delta G$  according to the following equation:

$$K = \exp\left[-\frac{\Delta G}{RT}\right], \quad (\text{S2})$$

where  $R$  is the universal gas constant (0.0019872 kcal/mol/K).  $\Delta G$  is a function of temperature, as shown in the following equation:

$$\Delta G = \frac{\Delta H_0 \cdot (T_m - T)}{T_m} + \Delta C_p \cdot (T - T_m - T \cdot \ln\left[\frac{T}{T_m}\right]), \quad (\text{S3})$$

where  $T_m$  is the midpoint of the unfolding transition and the temperature at which  $\Delta G_f = 0$ ;  $\Delta H_0$  is the change in enthalpy upon folding at  $T = T_m$ ; and  $\Delta C_p$  is the change in heat capacity upon folding.

In contrast, we used a model based on two-state dimer folding for noncovalent GCN4 heterodimer **27/29'-c40**, as described in equation S4:

$$[\theta] = F \cdot (a_n + b_n \cdot T) + (1 - F) \cdot (c_n + d_n \cdot T), \quad (\text{S4})$$

where  $T$ ,  $a_n$ ,  $b_n$ ,  $c_n$ , are defined as described above, and  $F$  is the temperature-dependent fraction of the total peptide concentration that is folded.  $F$  is related to the temperature-dependent equilibrium constant  $K$  and the total peptide concentration  $P$  as shown in equation S5:

$$F = 1 + \frac{1}{4KP} - \sqrt{\frac{1}{16K^2P^2} + \frac{1}{2KP}} \quad (\text{S5})$$



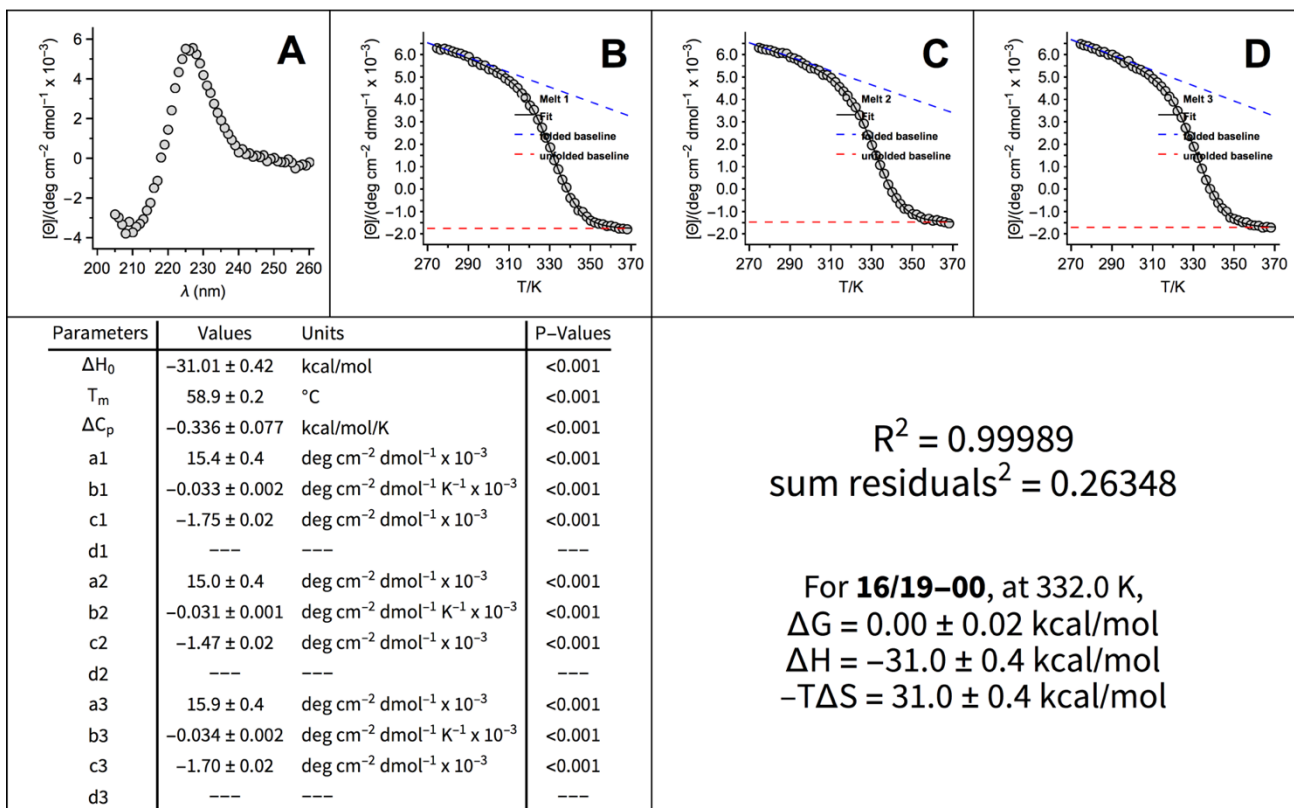
K is related to the folding free energy  $\Delta G$  as described above in equation S2. Because equation S3 is only appropriate for monomer folding equilibria, we used the following polynomial expression to describe the temperature-dependence of  $\Delta G$  for noncovalent GCN4 heterodimer **27/29'-c40**:

$$\Delta G = \Delta G_0 + \Delta G_1 \cdot (T - T_0) + \Delta G_2 \cdot (T - T_0)^2, \quad (\text{S6})$$

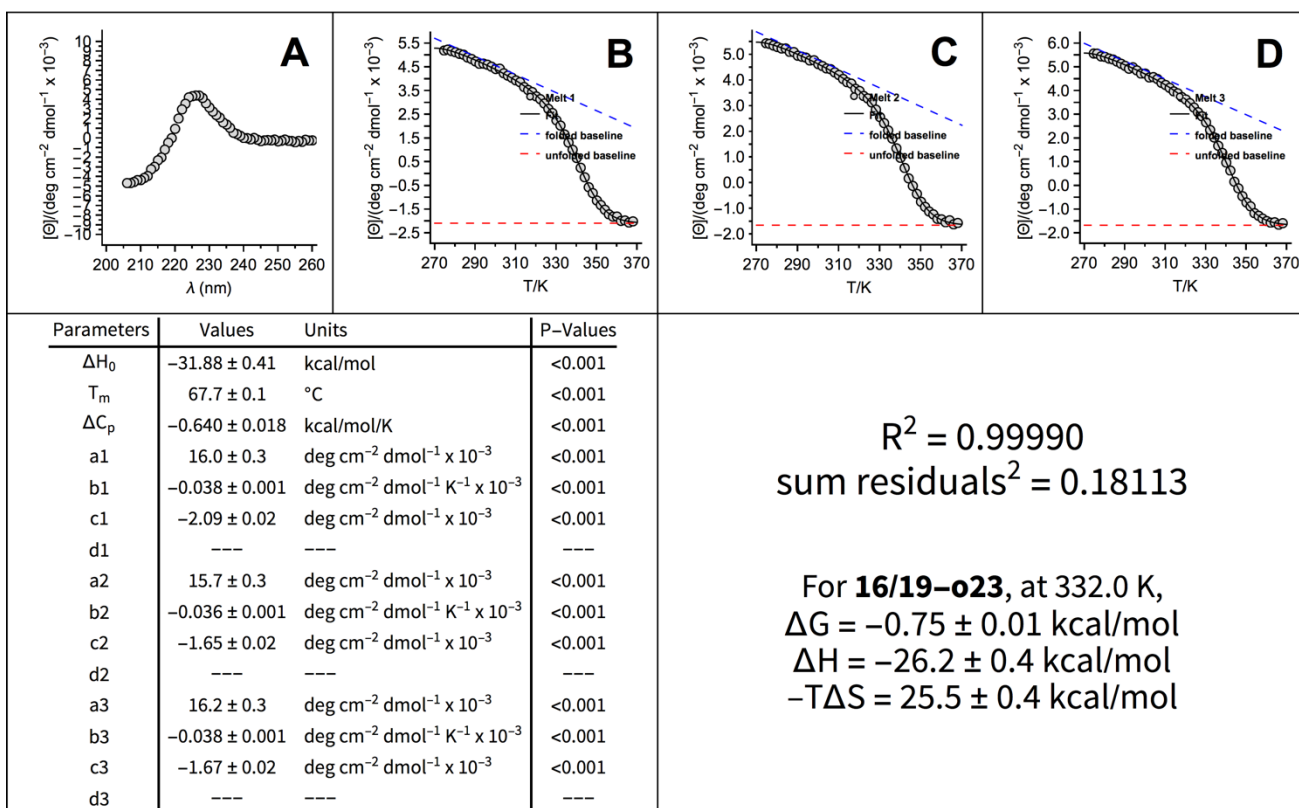
where  $\Delta G_0$ ,  $\Delta G_1$ , and  $\Delta G_2$  are parameters of the fit and  $T_0$  is an arbitrary reference temperature, usually chosen as  $T_m$ , the temperature at which  $F = 0.5$  (determined numerically for monomer-dimer equilibria).

In some cases, we found that some fit parameters had sufficiently high standard errors as to render them indistinguishable from zero and therefore not essential to the fit (as judged by their p-values). When this occurred, we repeated the fitting process without the non-essential parameters. We used the fit parameters for each variant to calculate the  $\Delta G$  values given in the main text; we calculated the uncertainty for each  $\Delta G$  value by propagation of error using the standard errors of the fit parameters.

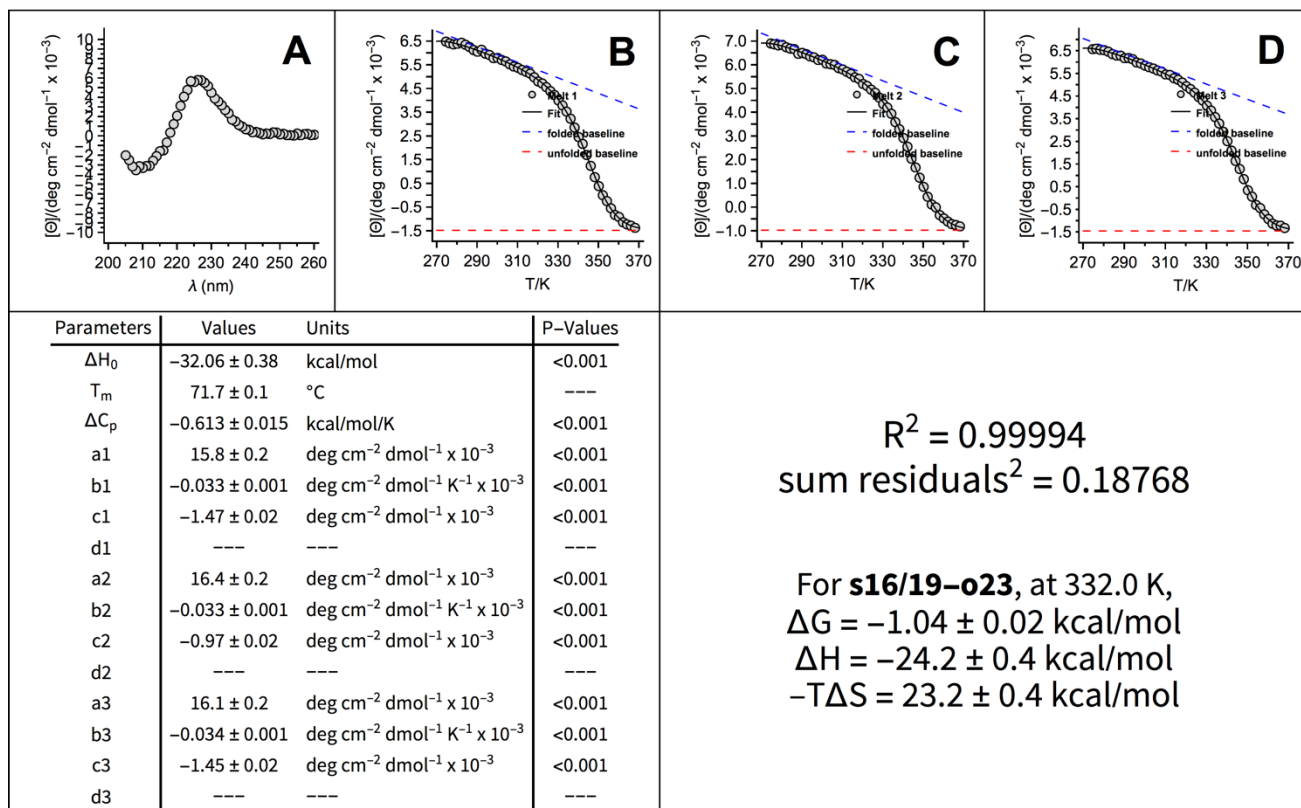
CD spectra and variable temperature CD data for WW variants **16/19-00**, **16/19-o23**, **s16/19-o23**, **16/32-c44**, **s16/32-c44**, **14/30-00**, **14/30-o55**, **s14/30-o55**, **14/30-c44** and **s14/30-c44**; SH3 variant **s20/37-o44**; and GCN4 variants **d27/29'-c40**, **sd27/29'-c40**, **27/29'-c40**, and **s27/29'-c40** are shown in Figures S9–S27, along with the fit parameters  $\pm$  standard error. Standard parameter errors were used to estimate the uncertainty in the thermodynamic values given in the main text by propagation of error.



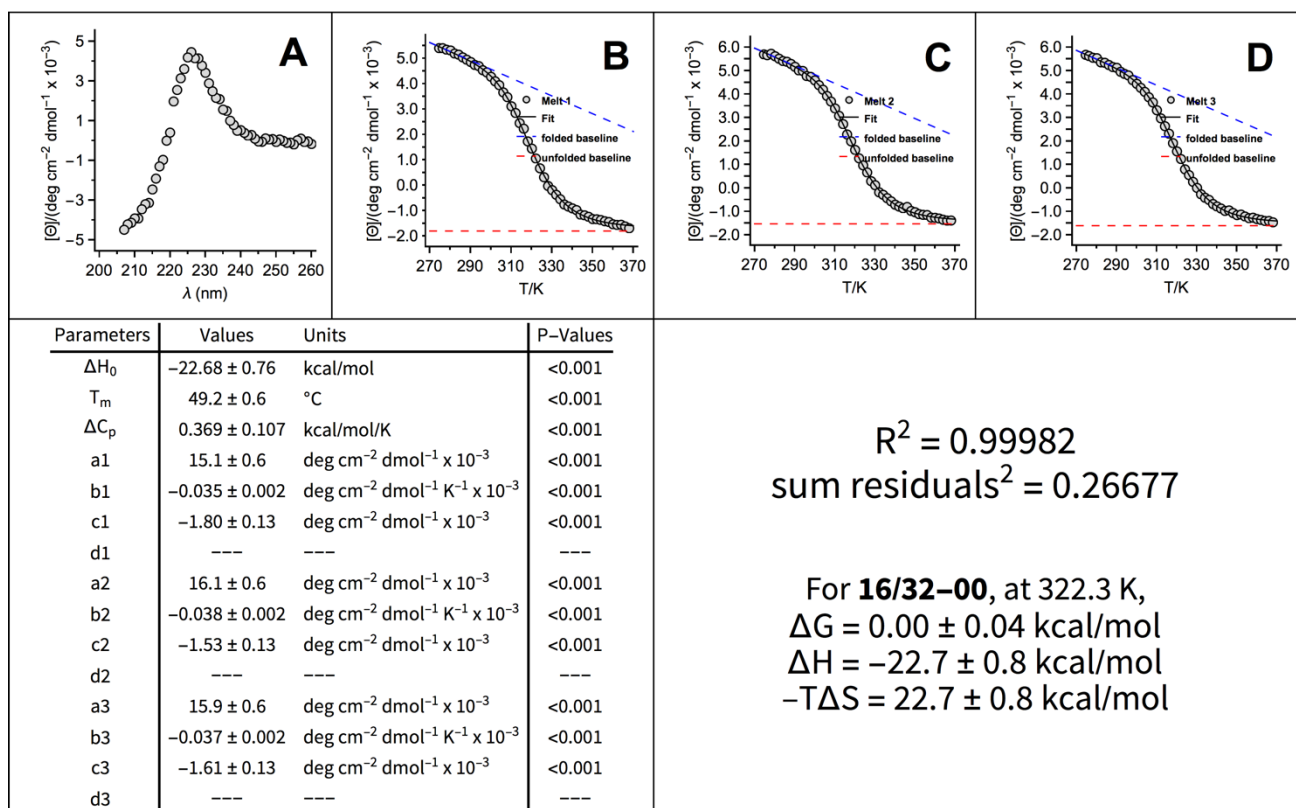
**Figure S9.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **16/19-00** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 332.0 K (the melting temperature of **16/19-00**), with the indicated standard errors.



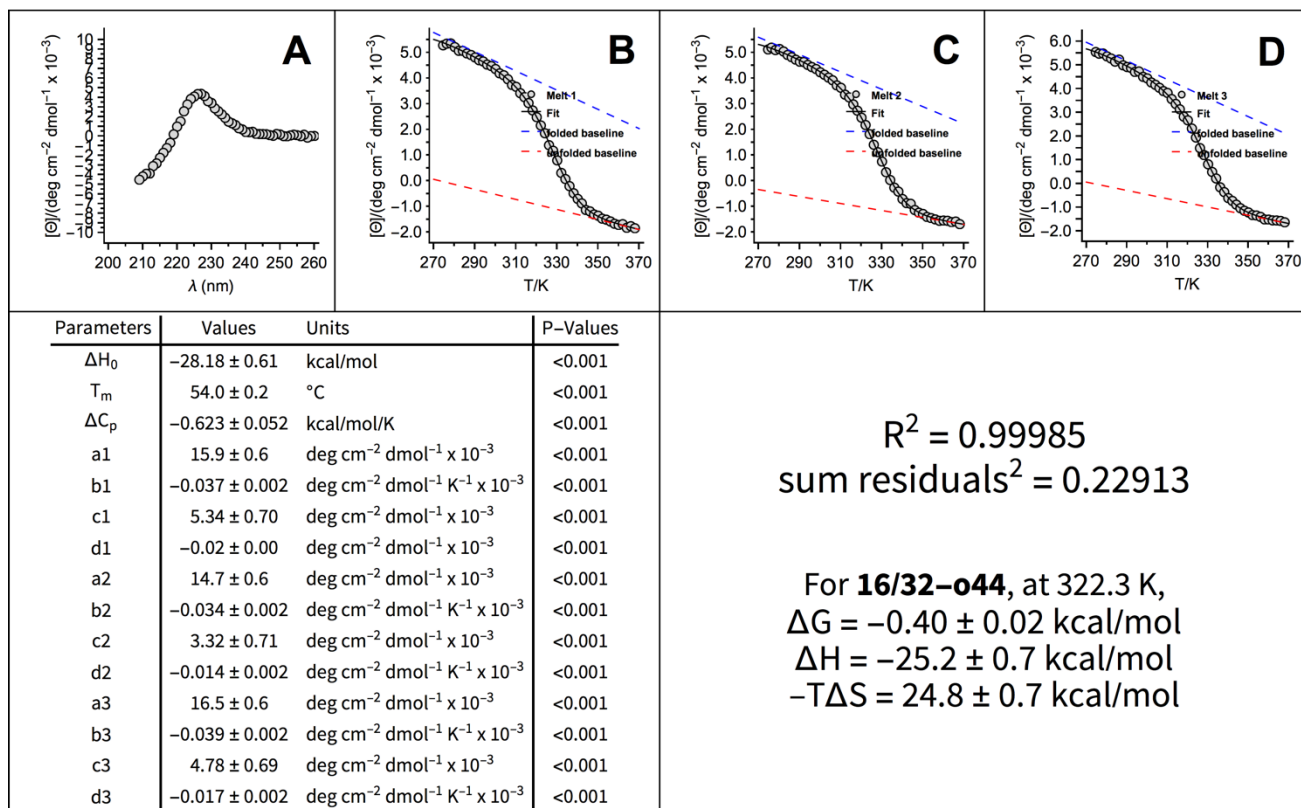
**Figure S10.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **16/19-o23** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 332.0 K (the melting temperature of **16/19-00**), with the indicated standard errors.



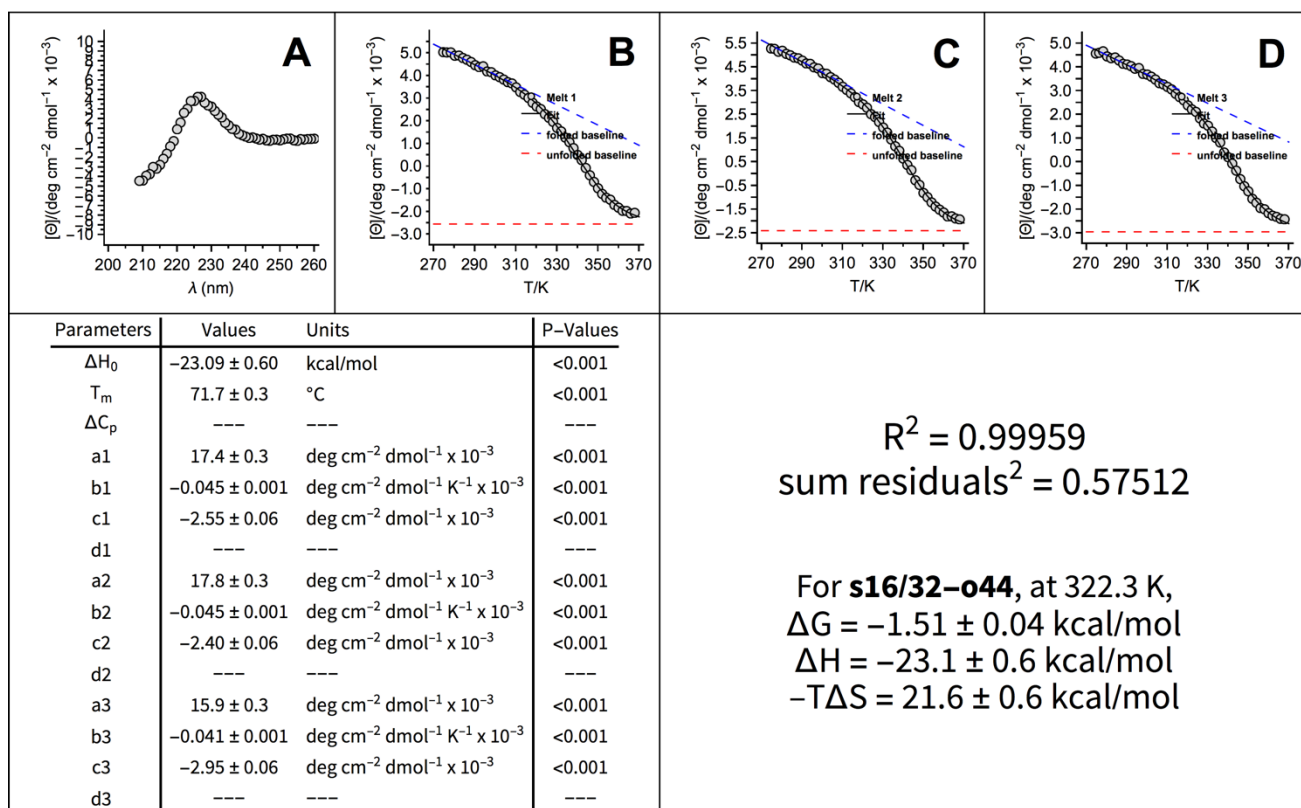
**Figure S11.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **s16/19-o23** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 332.0 K (the melting temperature of **16/19-00**), with the indicated standard errors.



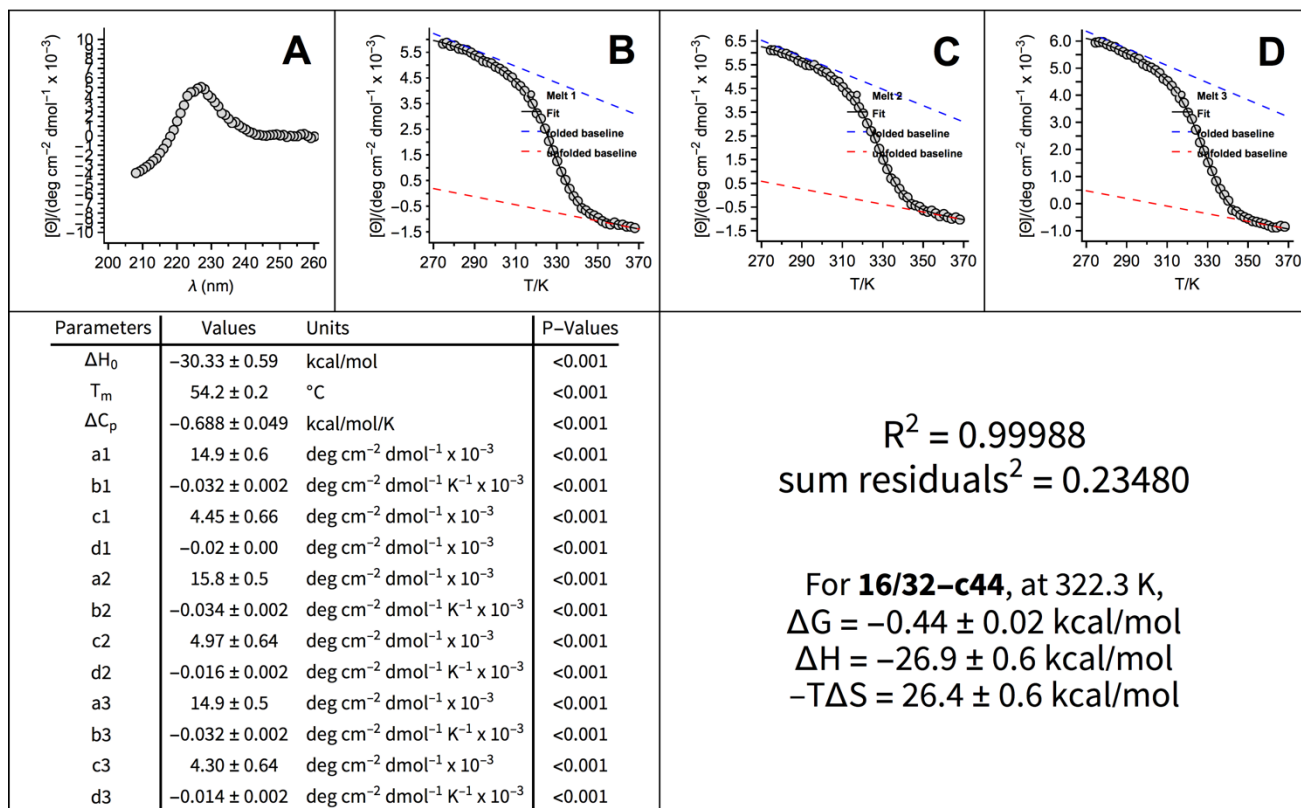
**Figure S12.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **16/32-00** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 322.3 K (the melting temperature of **16/32-00**), with the indicated standard errors.



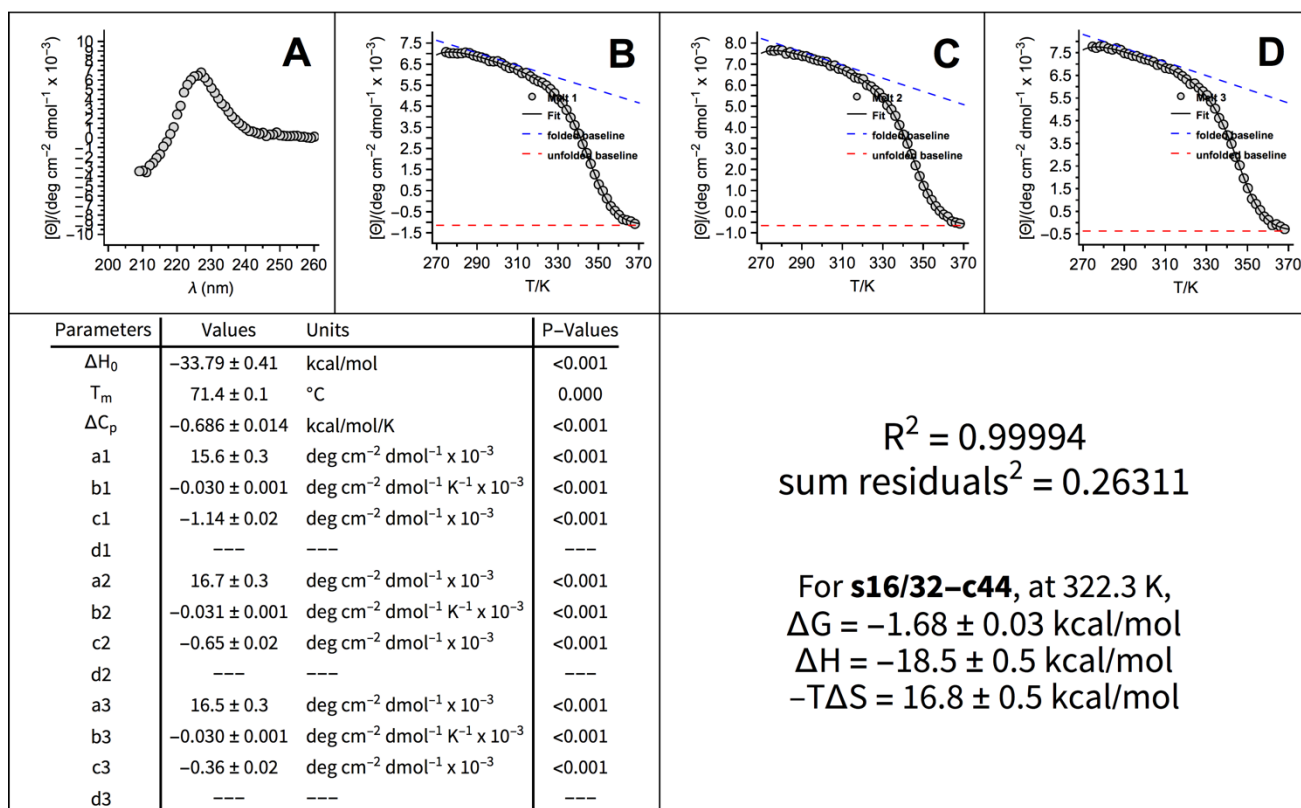
**Figure S13.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu$ M WW variant **16/32-o44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 322.3 K (the melting temperature of **16/32-00**), with the indicated standard errors.



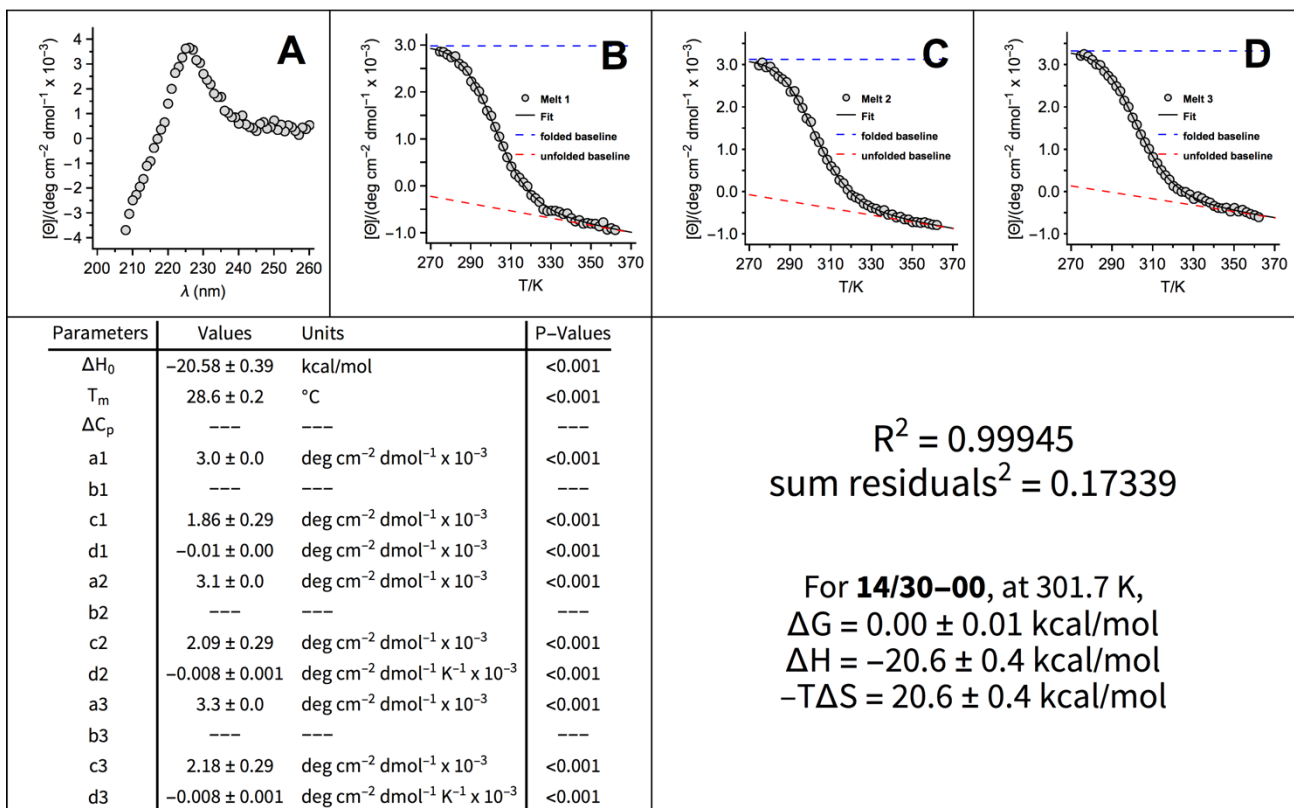
**Figure S14.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu$ M WW variant **s16/32-o44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 322.3 K (the melting temperature of **16/32-00**), with the indicated standard errors.



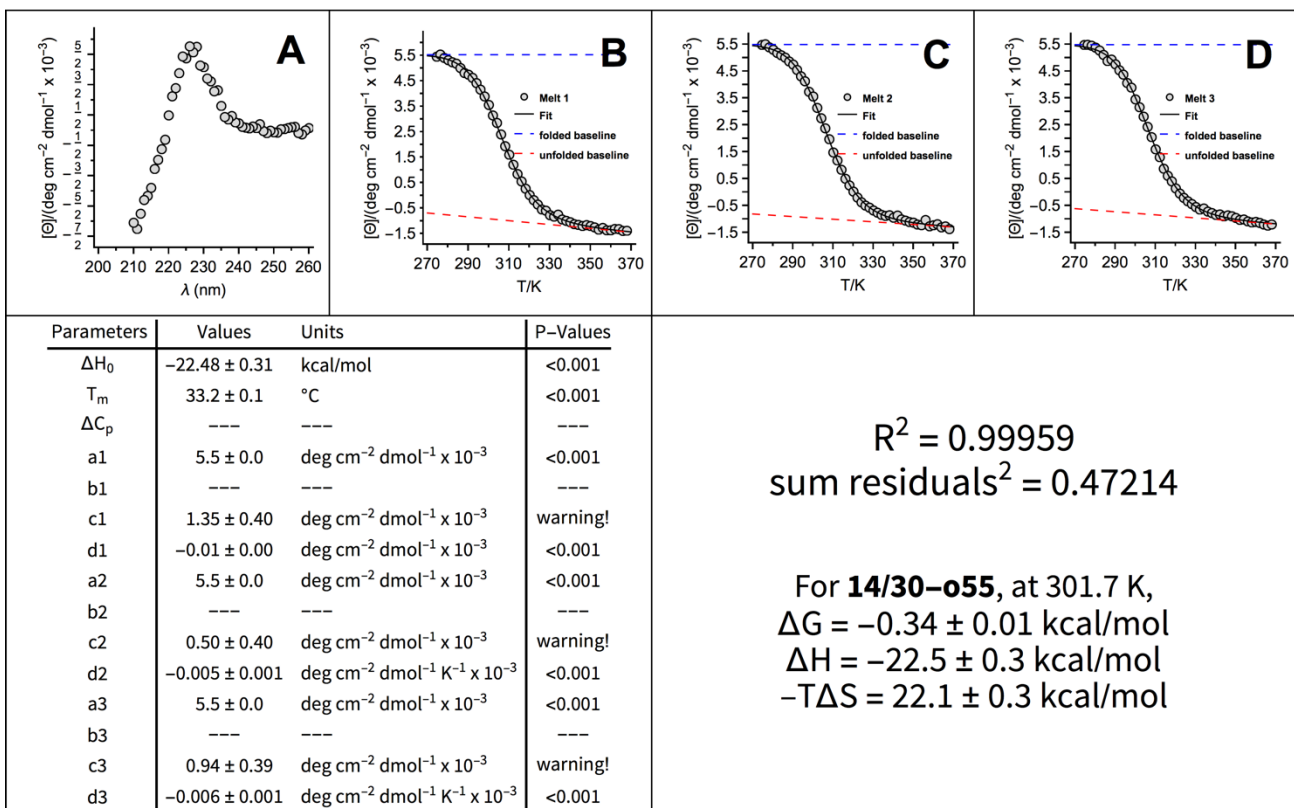
**Figure S15.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **16/32-c44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 322.3 K (the melting temperature of **16/32-00**), with the indicated standard errors.



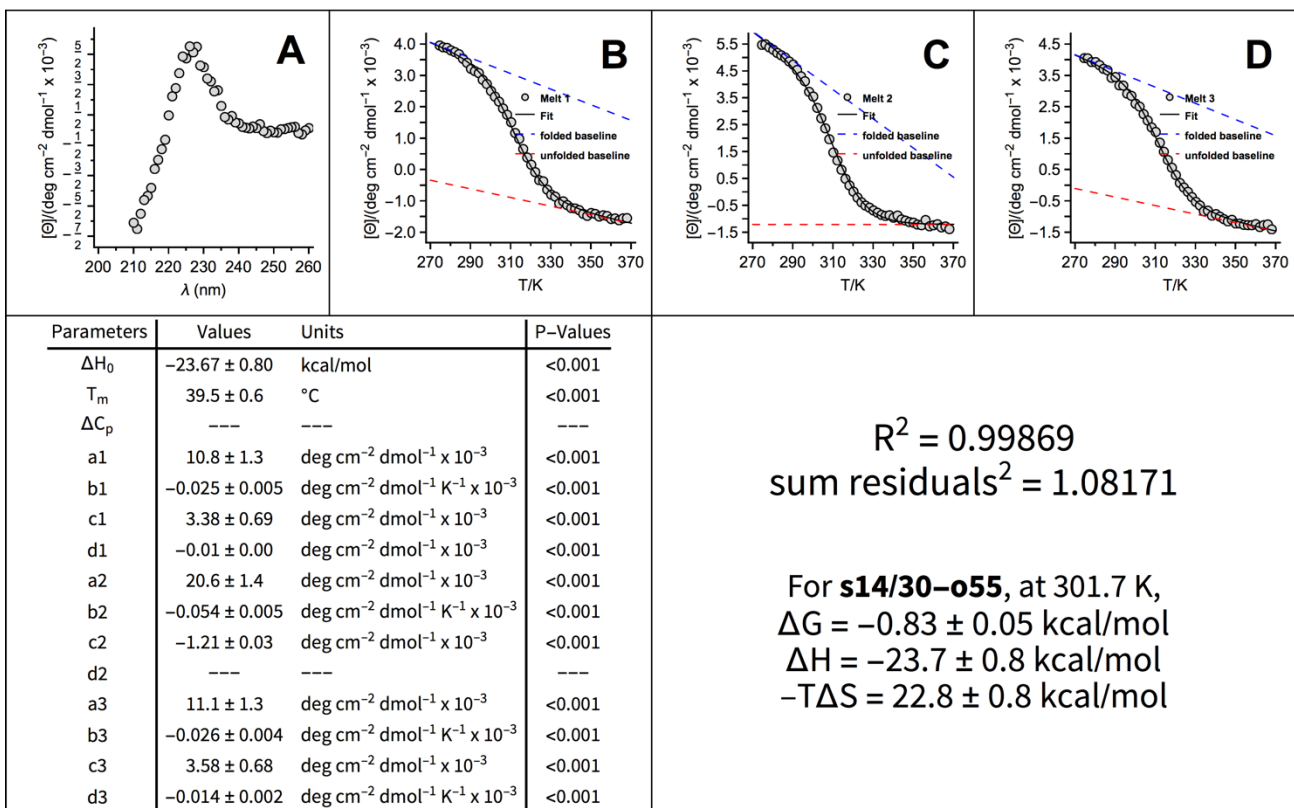
**Figure S16.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **s16/32-c44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 322.3 K (the melting temperature of **16/32-00**), with the indicated standard errors.



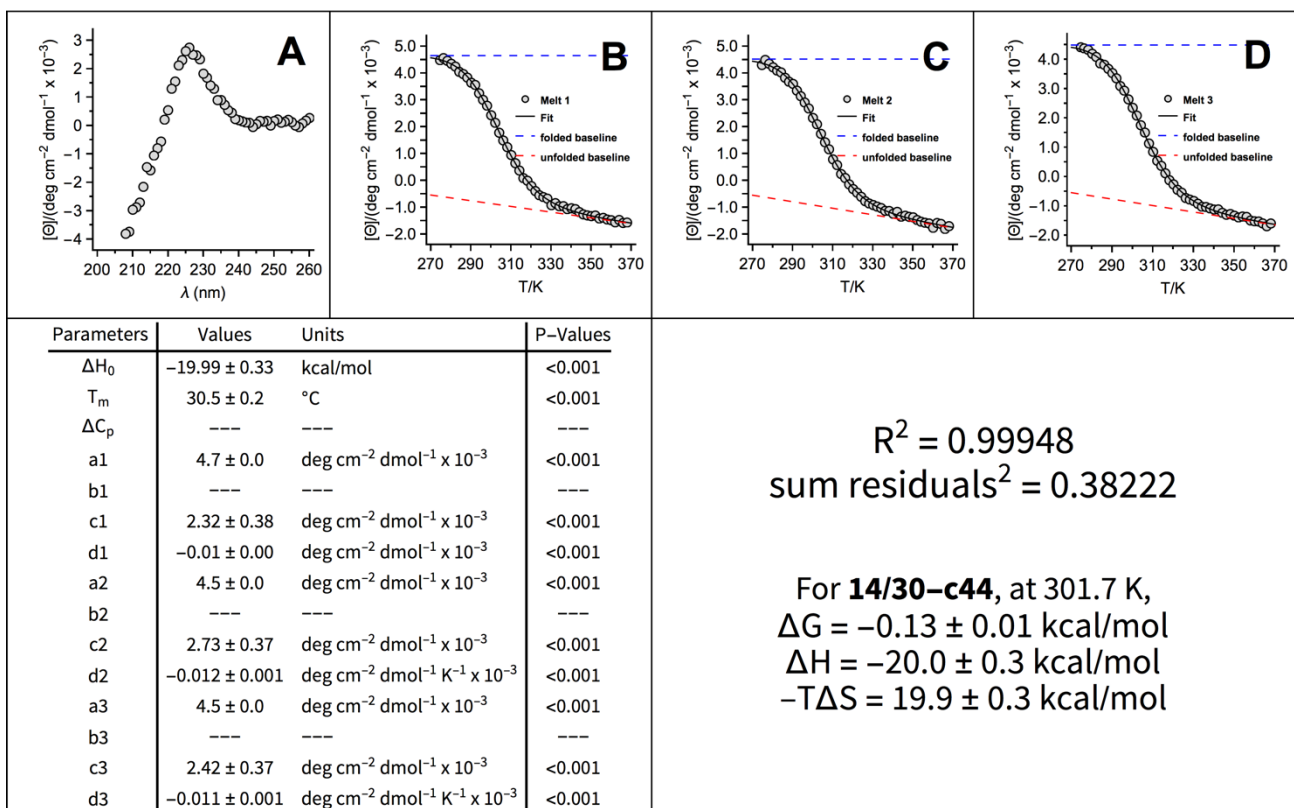
**Figure S17.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **14/30-00** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 301.7 K (the melting temperature of **14/30-00**), with the indicated standard errors.



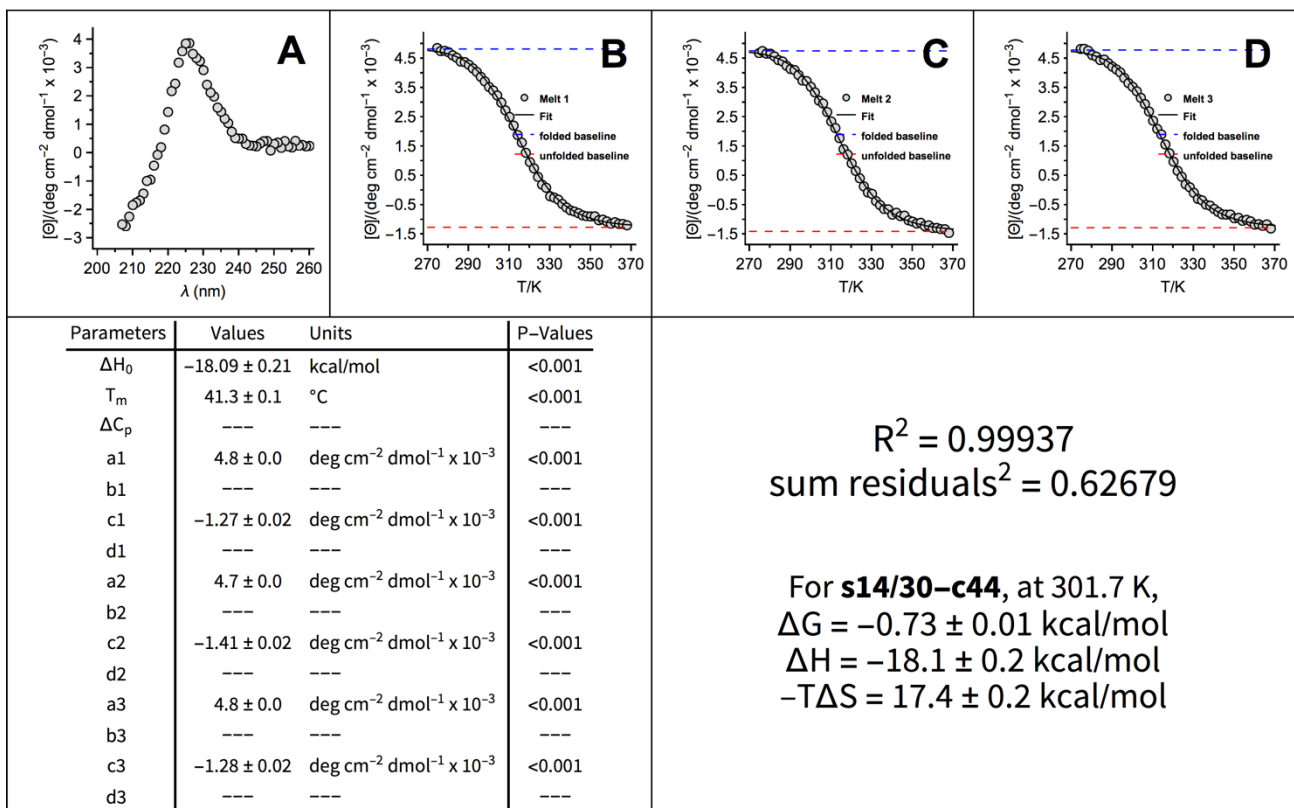
**Figure S18.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **14/30-o55** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 301.7 K (the melting temperature of **14/30-00**), with the indicated standard errors.



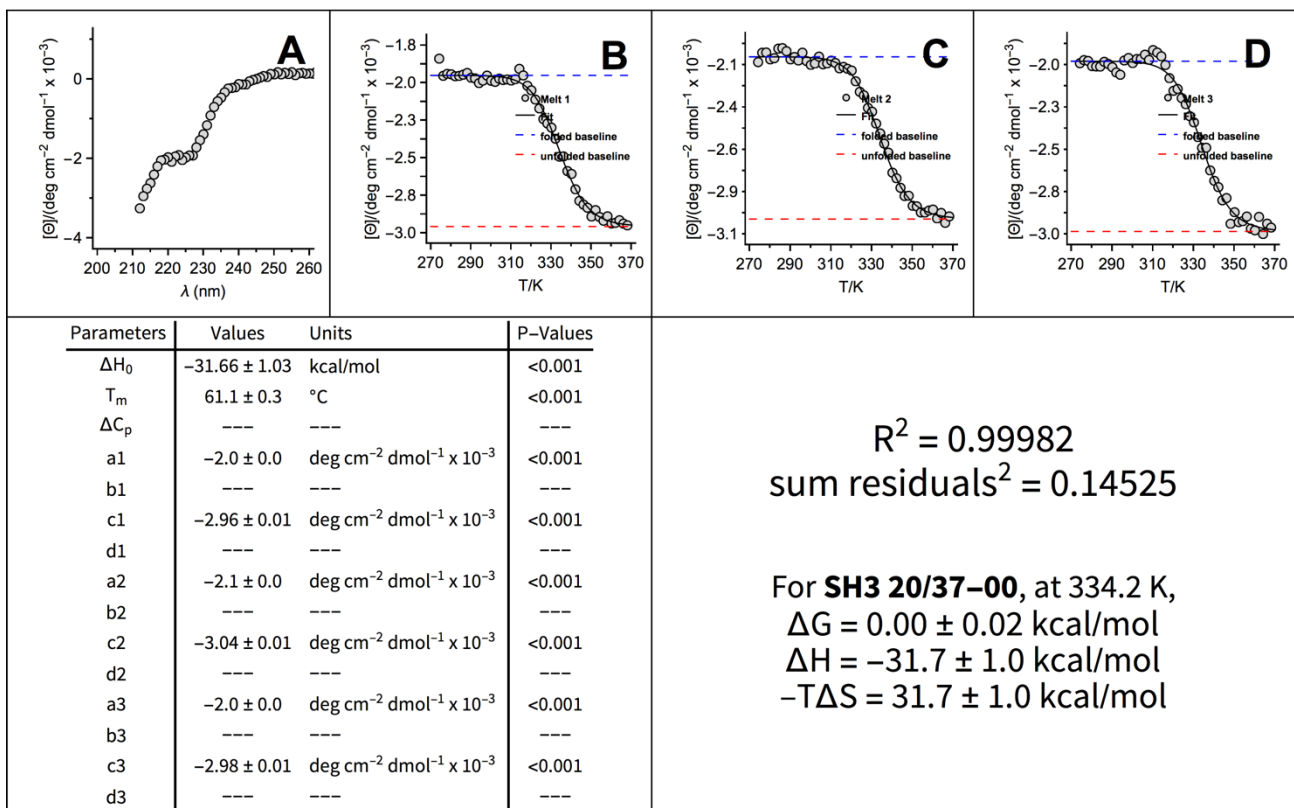
**Figure S19.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **s14/30-o55** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 301.7 K (the melting temperature of **14/30-00**), with the indicated standard errors.



**Figure S20.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **14/30-c44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 301.7 K (the melting temperature of **14/30-00**), with the indicated standard errors.

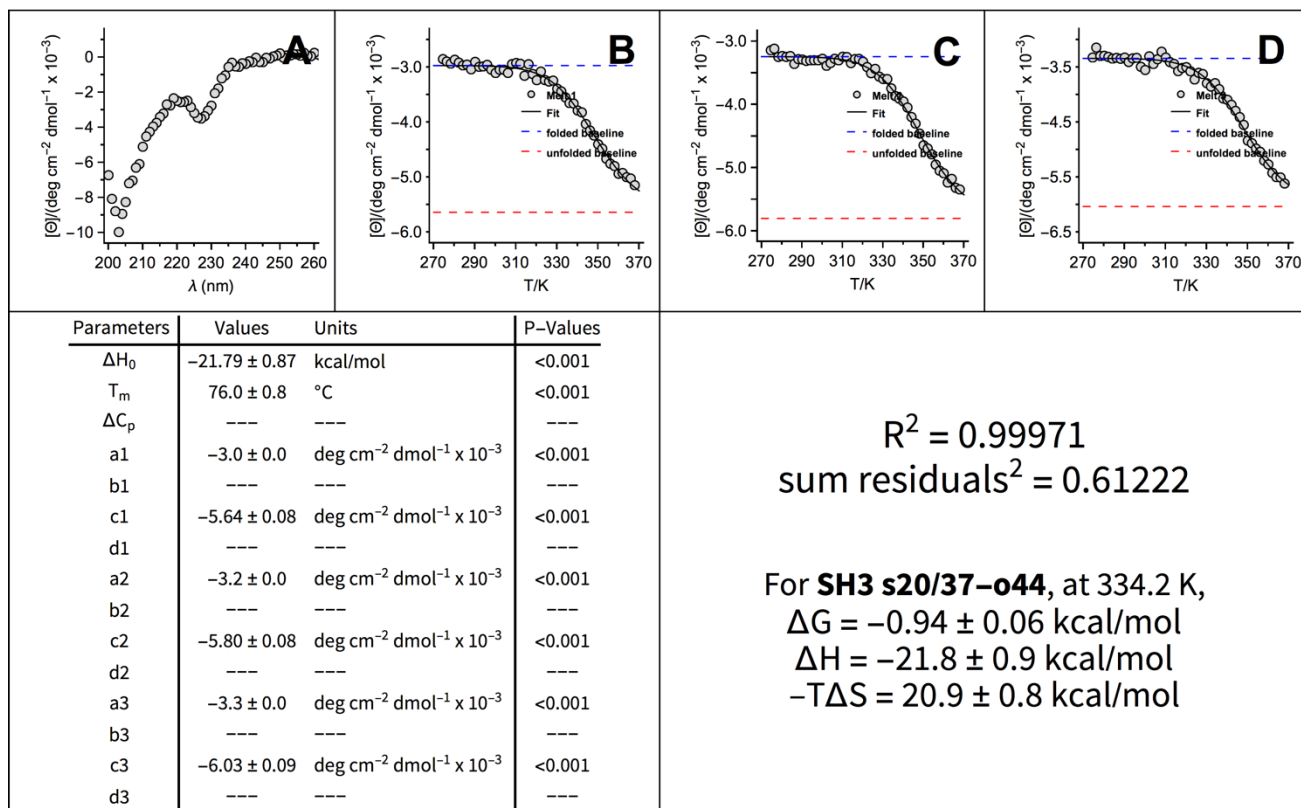


**Figure S21.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  WW variant **s14/30-c44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 301.7 K (the melting temperature of **14/30-00**), with the indicated standard errors.

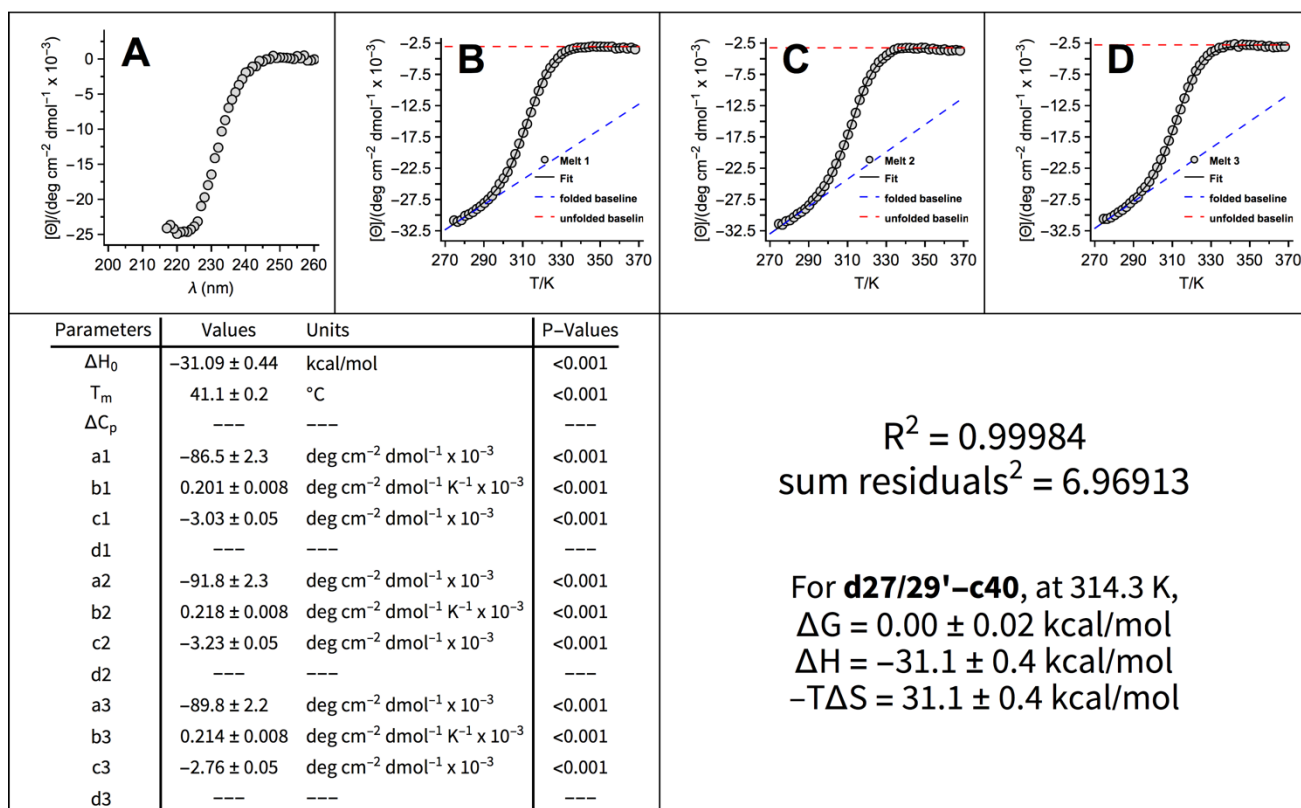


**Figure S22.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  SH3 variant **20/37-00** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 334.2 K (the melting temperature of **20/37-00**), with the indicated standard errors.

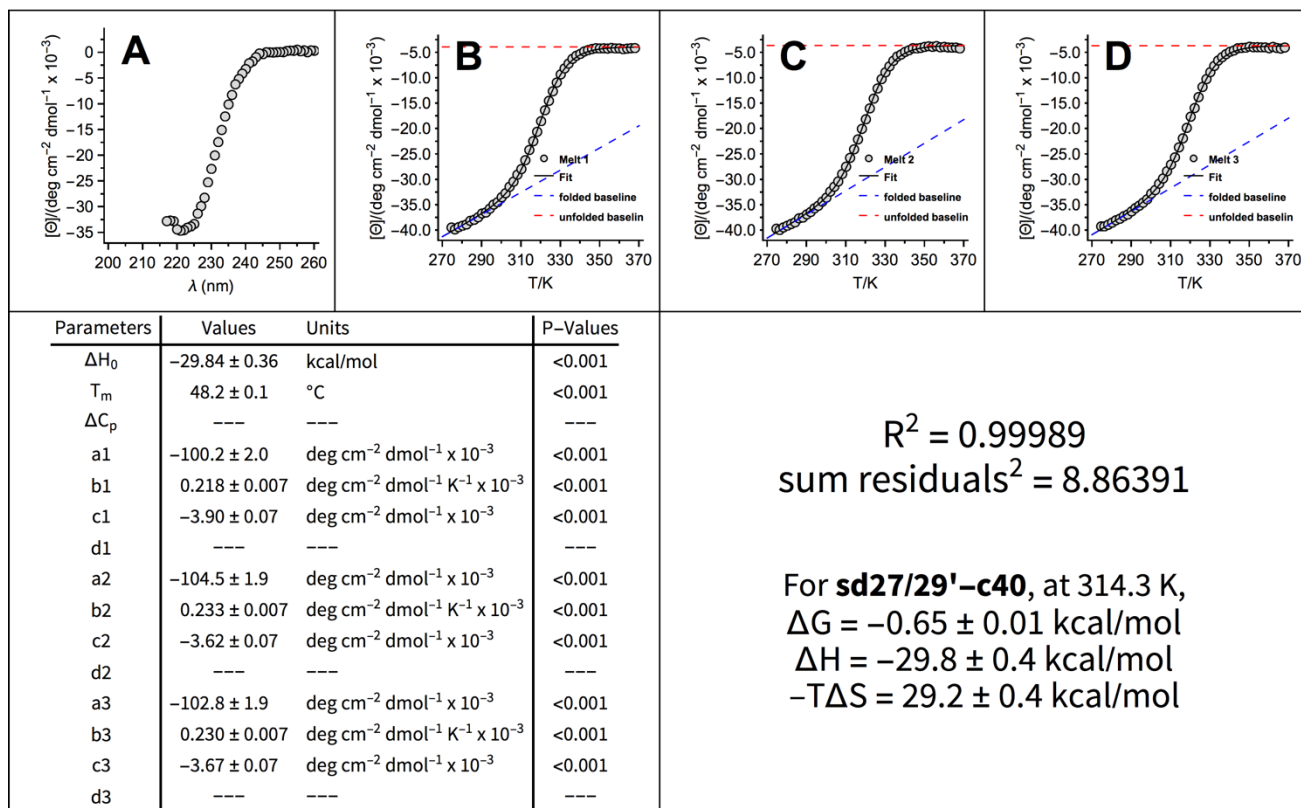




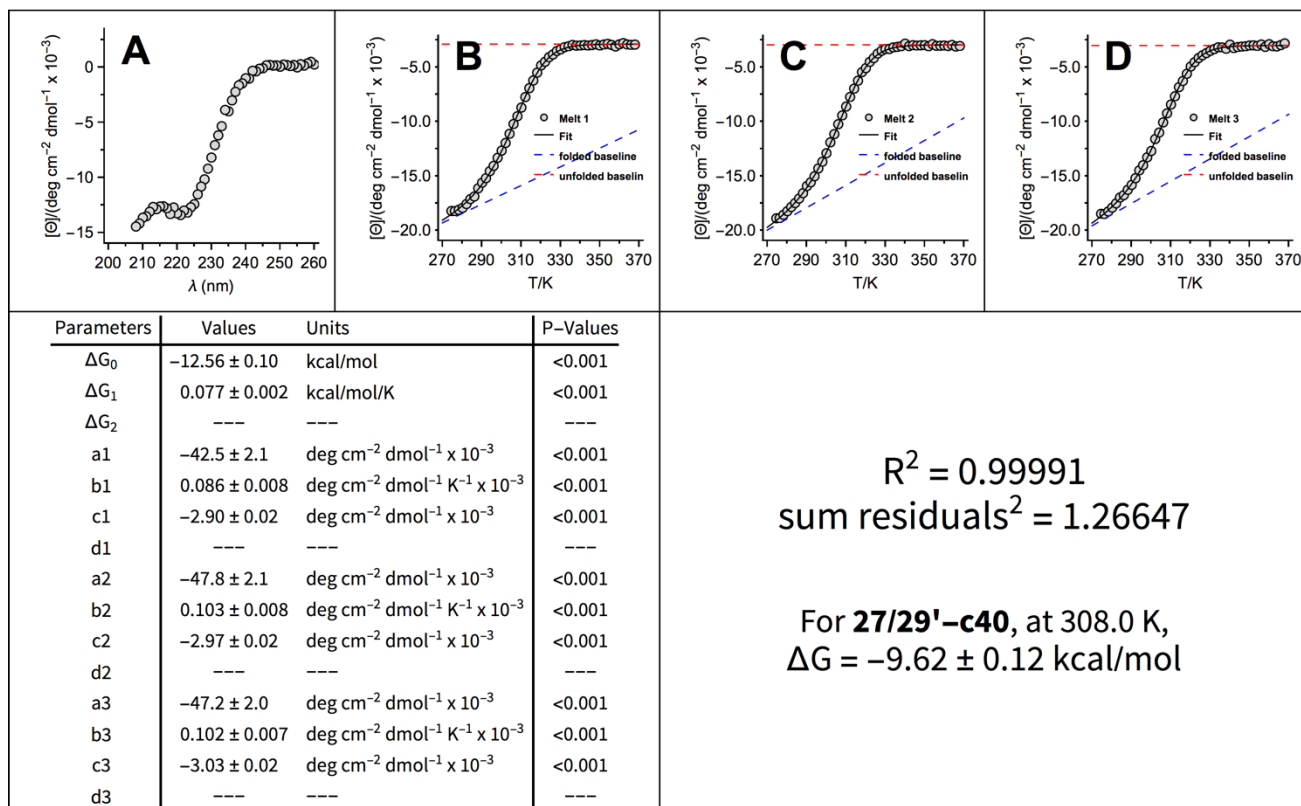
**Figure S23.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 50  $\mu\text{M}$  SH3 variant **s20/37-o44** in 20 mM sodium phosphate (pH 7). Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-\Delta S_f$  at 334.2 K (the melting temperature of **20/37-00**), with the indicated standard errors.



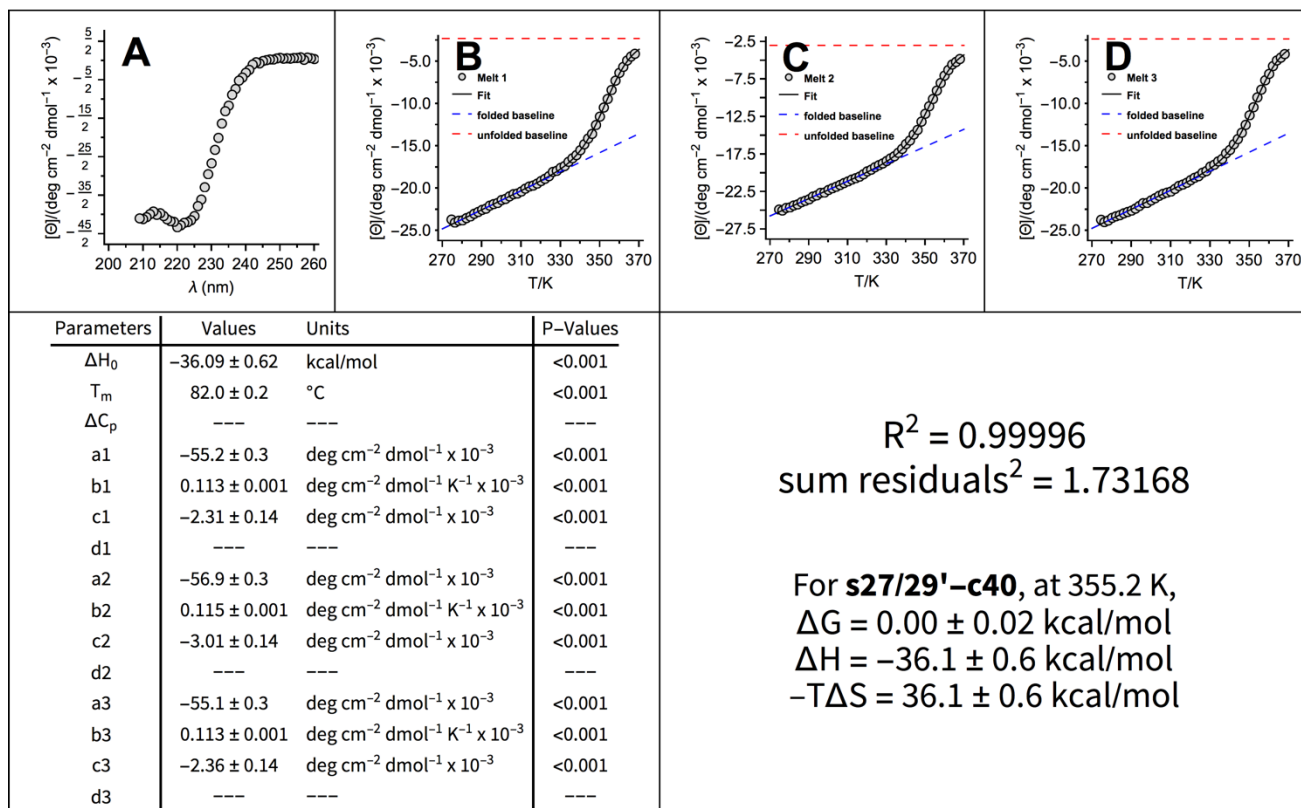
**Figure S24.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 15  $\mu\text{M}$  disulfide bound GCN4 heterodimer **d27/29'-c40** in 20 mM sodium phosphate (pH 7) with 4 M GdnHCl. Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-\Delta S_f$  at 334.2 K (the melting temperature of **d27/29'-c40**), with the indicated standard errors.



**Figure S25.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 15  $\mu\text{M}$  disulfide bound triazole-stapled GCN4 variant **sd27/29'-c40** in 20 mM sodium phosphate (pH 7) with 4 M GdnHCl. Fit parameters from equations S1–S3 appear in the table, as do calculated values for  $\Delta G_f$ ,  $\Delta H_f$ , and  $-T\Delta S_f$  at 334.2 K (the melting temperature of **d27/29'-c40**), with the indicated standard errors.



**Figure S26.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 15  $\mu\text{M}$  non-covalent GCN4 heterodimer **27/29'-c40** in 20 mM sodium phosphate (pH 7) with 0.5 M GdnHCl. Fit parameters from equations S4–S6 appear in the table, as does the calculated value for  $\Delta G_f$  at 308.0 K (the melting temperature of **27/29'-c40**), with the indicated standard errors.



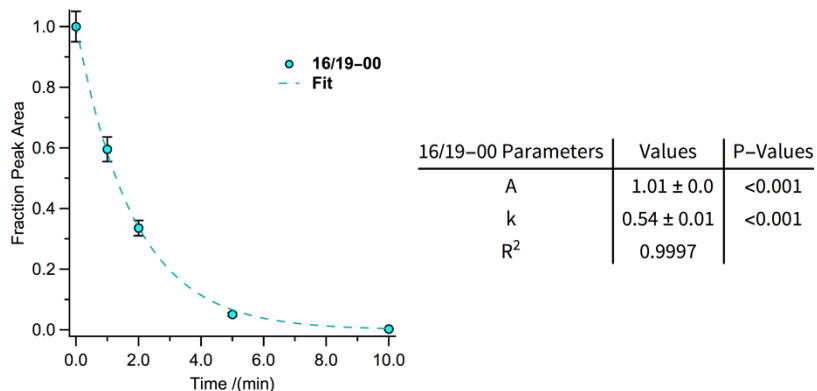
**Figure S27.** (A) CD spectra and (B–D) variable temperature CD data (triplicate) for 15  $\mu\text{M}$  triazole-stapled GCN4 **s27/29'-c04** in 20 mM sodium phosphate (pH 7) with 0.5 M GdnHCl. Fit parameters from equations S1–S3 appear in the table, as does the calculated value for  $\Delta G$ ,  $\Delta H$ , and  $-\Delta S$  at 355.2 K (the melting temperature of **s27/29'-c04**), with the indicated standard errors.

#### 4. Proteolysis of WW variants

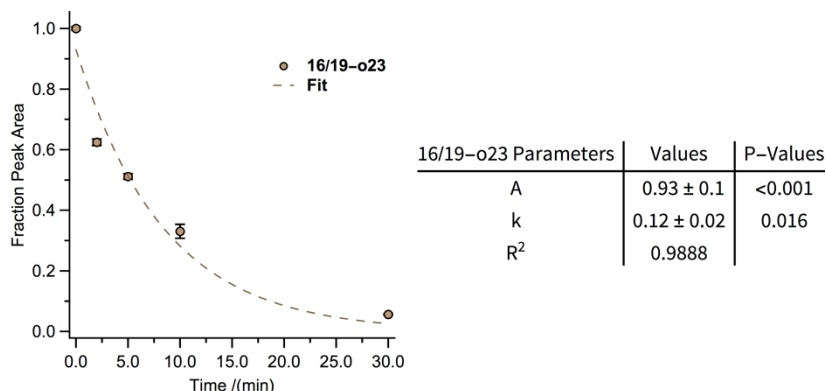
50  $\mu\text{M}$  protein solutions in 20 mM sodium phosphate buffer (pH 7) were incubated at ambient temperature with 17  $\mu\text{g/mL}$  proteinase K respectively for up to 1 hour. At each of the several time points, the proteolysis reaction was quenched by adding 100  $\mu\text{L}$  of aqueous trifluoroacetic acid (1% v/v) to 50  $\mu\text{L}$  of the reaction mixture. The quenched mixture was then analyzed in triplicate by reverse phase HPLC analytical column, monitored by a UV-Vis detector at 220 nm. The degradation of the proteins was assessed using the integrated HPLC peak area to account for how much of the full-length protein remained at each time point. The protein half-lives were calculated by fitting the integrated peak areas as a function of time to a mono exponential decay equation:

$$\text{Area}(t) = A \cdot \exp[-kt],$$

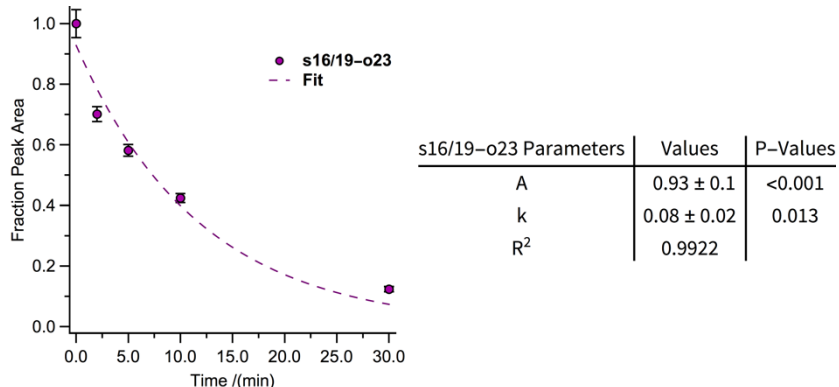
where  $t$  is time in minutes,  $A$  is a constant corresponding to relative integrated peak area at  $t = 0$ , and  $\tau$  is the decay time, which is related to the protein half-life  $t_{1/2}$  ( $t_{1/2} = \tau \ln 2$ ). Decay traces for proteins WW variants, **16/19-00**, **16/19-o23**, **s16/19-o23**, **16/32-00**, **16/32-o44**, **s16/32-o44**, are shown in Figures S28– S33.



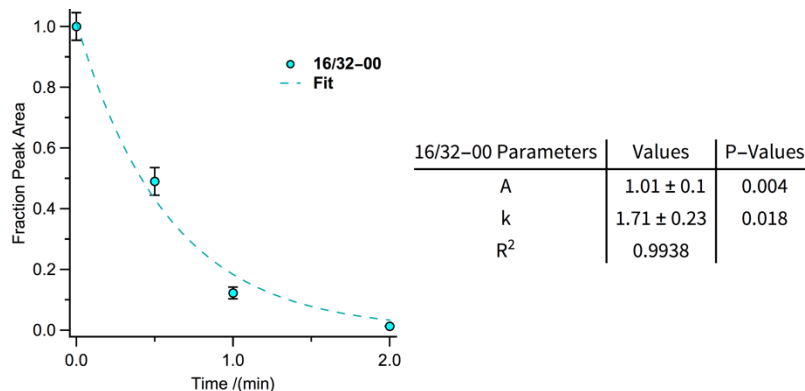
**Figure S28.** Proteolysis of **16/19-00** (50  $\mu\text{M}$  protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu\text{g}/\text{mL}$ ) as monitored by HPLC. Data points for **16/19-00** are shown as blue circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.



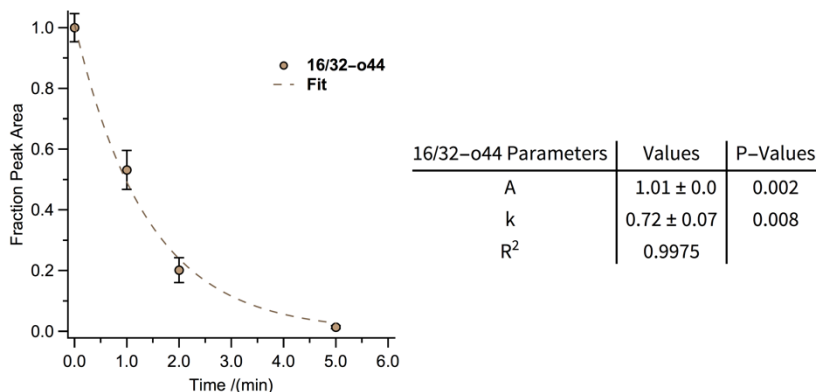
**Figure S29.** Proteolysis of **16/19-o23** (50  $\mu\text{M}$  protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu\text{g}/\text{mL}$ ) as monitored by HPLC. Data points for **16/19-o23** are shown as orange circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.



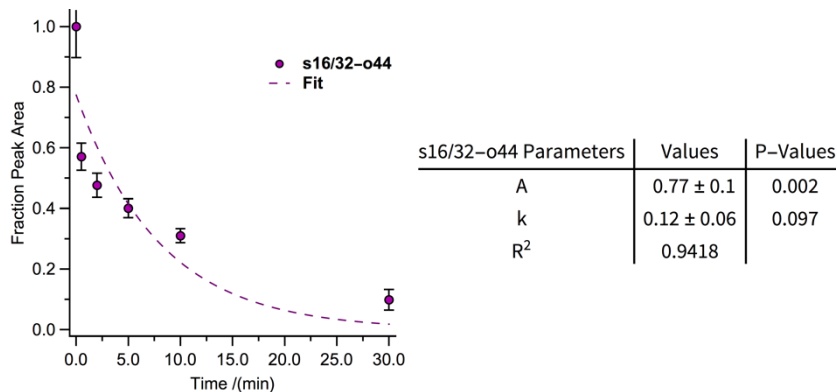
**Figure S30.** Proteolysis of **16/19-o23** (50  $\mu\text{M}$  protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu\text{g}/\text{mL}$ ) as monitored by HPLC. Data points for **16/19-o23** are shown as pink circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.



**Figure S31.** Proteolysis of **16/32-00** (50  $\mu$ M protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu$ g/mL) as monitored by HPLC. Data points for **16/32-00** are shown as blue circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.



**Figure S32.** Proteolysis of **16/32-o44** (50  $\mu$ M protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu$ g/mL) as monitored by HPLC. Data points for **16/32-o44** are shown as orange circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.



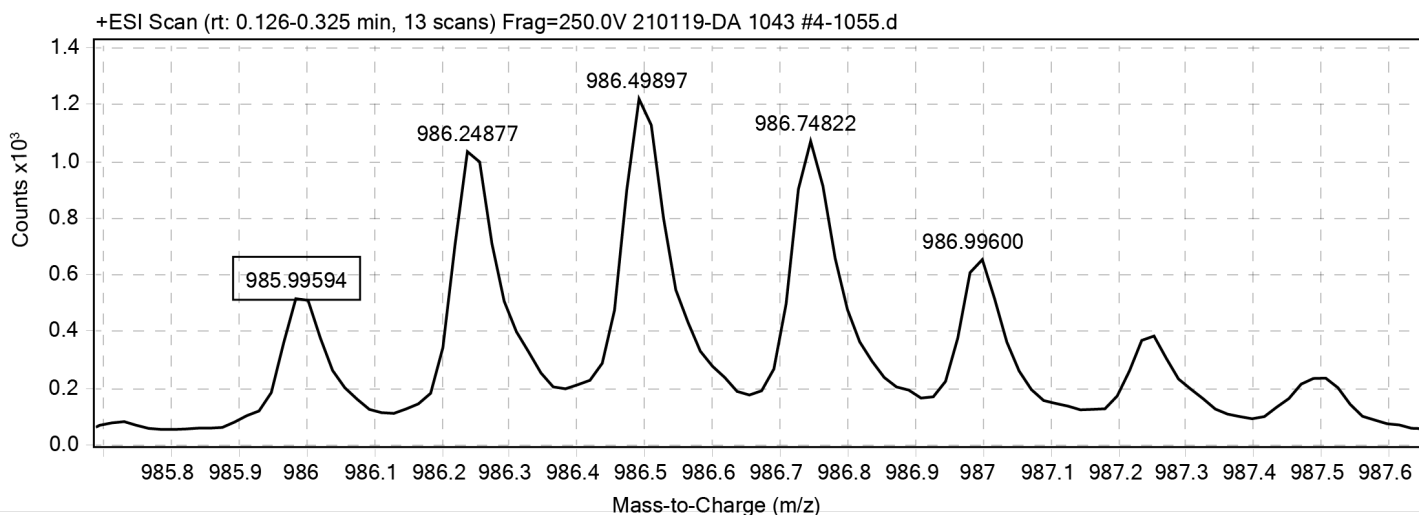
**Figure S33.** Proteolysis of **s16/32-o44** (50  $\mu$ M protein concentration in 20 mM sodium phosphate buffer, pH 7) by proteinase K (17  $\mu$ g/mL) as monitored by HPLC. Data points for **s16/32-o44** are shown as orange circles, and each represents the average of three replicate experiments. Solid lines represent fits of the data to a mono exponential decay function, which was used to calculate the indicated half-lives.

## 5. ESI-TOF MS data for WW, SH3, and GCN4 variants

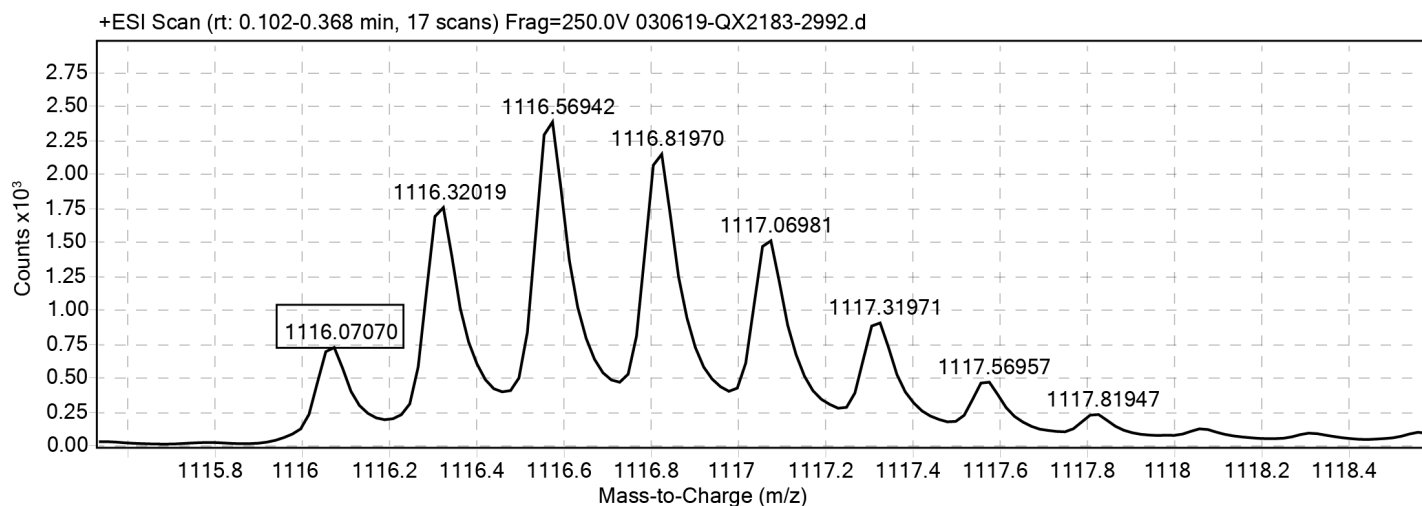
**Table S1. Summary of the mass spectrum data for the new WW, SH3, and GCN4 variants described here.**

Name	Notebook Number	Molecular Formula	z	Expected [M+z·H]/z	Observed [M+z·H]/z
14/30-00	DA10434	C <sub>175</sub> H <sub>261</sub> N <sub>52</sub> O <sub>51</sub> S	4	985.989	985.996
14/30-55	QX2183	C <sub>201</sub> H <sub>310</sub> N <sub>52</sub> O <sub>61</sub> S	4	1116.070	1116.071
s14/30-55	QX2183s	C <sub>199</sub> H <sub>306</sub> N <sub>52</sub> O <sub>61</sub> S	4	1109.062	1109.058
14/30-c44	QX22154	C <sub>194</sub> H <sub>294</sub> N <sub>55</sub> O <sub>57</sub> S	4	1089.797	1089.790
s14/30-c44	QX22154s	C <sub>194</sub> H <sub>294</sub> N <sub>55</sub> O <sub>57</sub> S	4	1089.797	1089.790
16/32-c44	QX21852	C <sub>198</sub> H <sub>303</sub> N <sub>59</sub> O <sub>57</sub> S	4	1113.817	1113.806
s16/32-c44	QX2188	C <sub>198</sub> H <sub>303</sub> N <sub>59</sub> O <sub>57</sub> S	4	1113.817	1113.810
20/37-44	KT1027	C <sub>310</sub> H <sub>471</sub> N <sub>77</sub> O <sub>98</sub>	4	1711.114	1171.092
s20/37-44	QX2217	C <sub>308</sub> H <sub>467</sub> N <sub>77</sub> O <sub>98</sub>	4	1704.106	1704.108
27-c4	QX21971	C <sub>187</sub> H <sub>324</sub> N <sub>56</sub> O <sub>46</sub> S	4	1031.620	1031.618
29'-c0	QX21973	C <sub>184</sub> H <sub>293</sub> N <sub>51</sub> O <sub>68</sub> S	4	1085.277	1085.271
d27/29'-c40	QX22031	C <sub>371</sub> H <sub>611</sub> N <sub>105</sub> O <sub>114</sub> S <sub>2</sub>	6	1410.592	1410.753
sd27/29'-c40	QX2204	C <sub>371</sub> H <sub>611</sub> N <sub>105</sub> O <sub>114</sub> S <sub>2</sub>	6	1410.592	1410.736
27A-c4	QX22091	C <sub>187</sub> H <sub>324</sub> N <sub>56</sub> O <sub>46</sub>	4	1023.627	1023.620
29A'-c0	QX22092	C <sub>184</sub> H <sub>293</sub> N <sub>51</sub> O <sub>68</sub>	4	1077.284	1077.276
s27/29'-c40	QX2214	C <sub>371</sub> H <sub>613</sub> N <sub>105</sub> O <sub>114</sub> S <sub>2</sub>	6	1400.271	1400.583

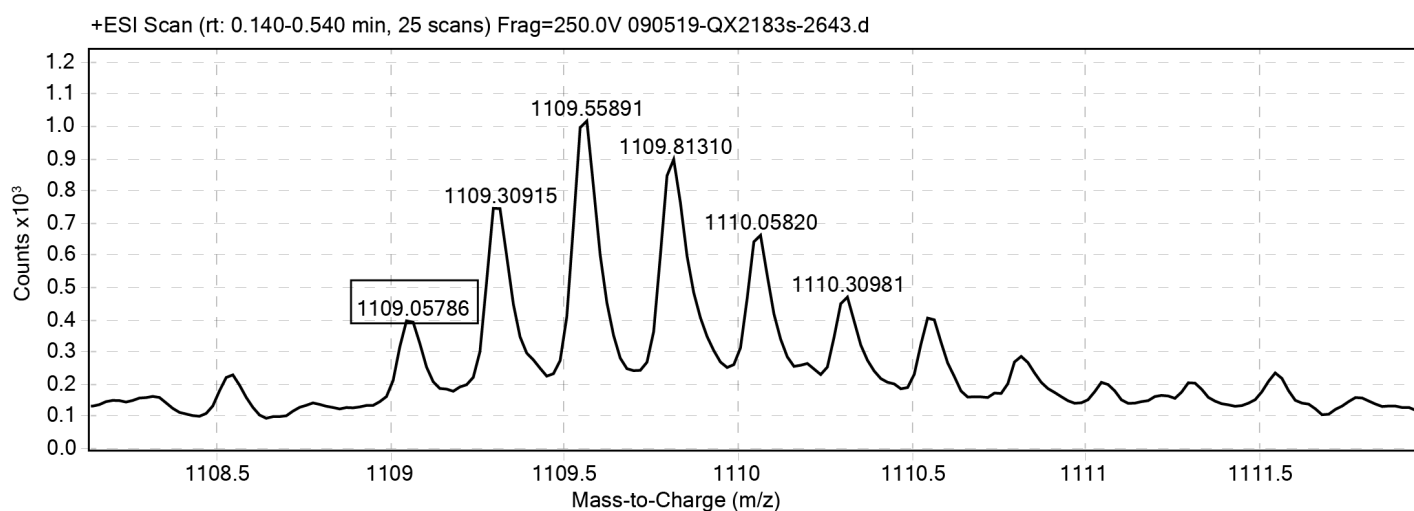
ESI-TOF MS data for WW variants **14/30-00**, **14/30-o55**, **s14/30-o55**, **14/30-c44**, **s14/30-c44**, **16/32-c44** and **s16/32-c44**; SH3 variants **20/37-o44** and **s20/37-o44**; and GCN4 variants **27-c4**, **29'-c0**, **d27/29'-c40**, **sd27/29'-c40**, **27A-c4**, **29A'-c0**, and **s27/29'-40** are shown in **Figures S34–S49**.



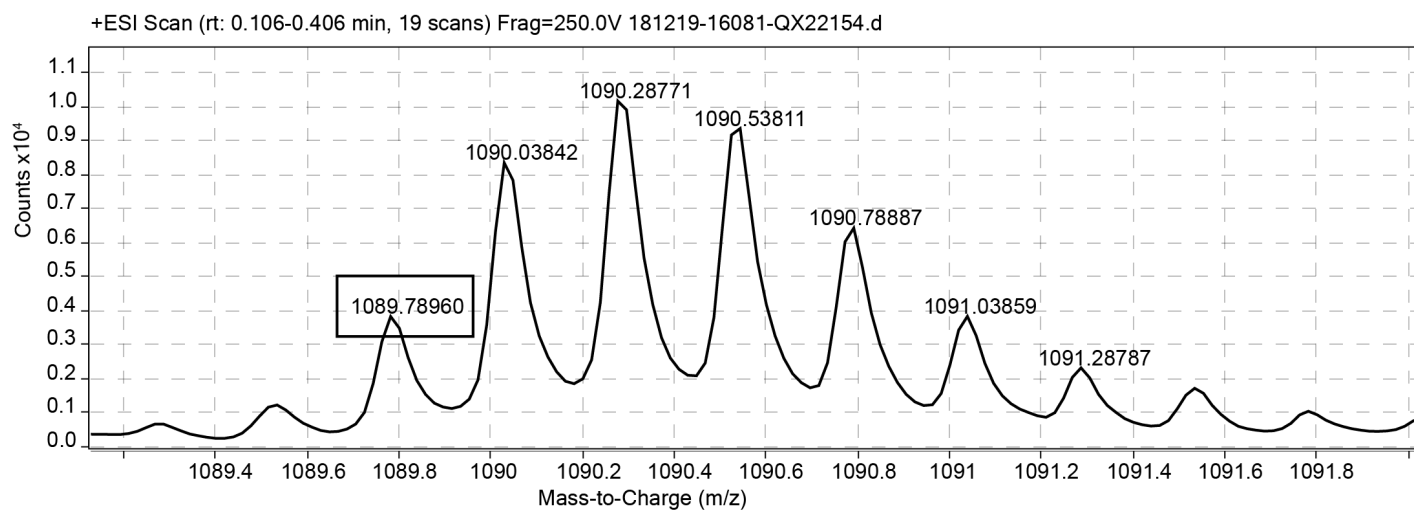
**Figure S34.** ESI-TOF spectrum for WW variant **14/30-00**. Expected  $[M+4H^+]/4 = 985.989$  Da. Observed  $[M+4H^+]/4 = 985.996$  Da.



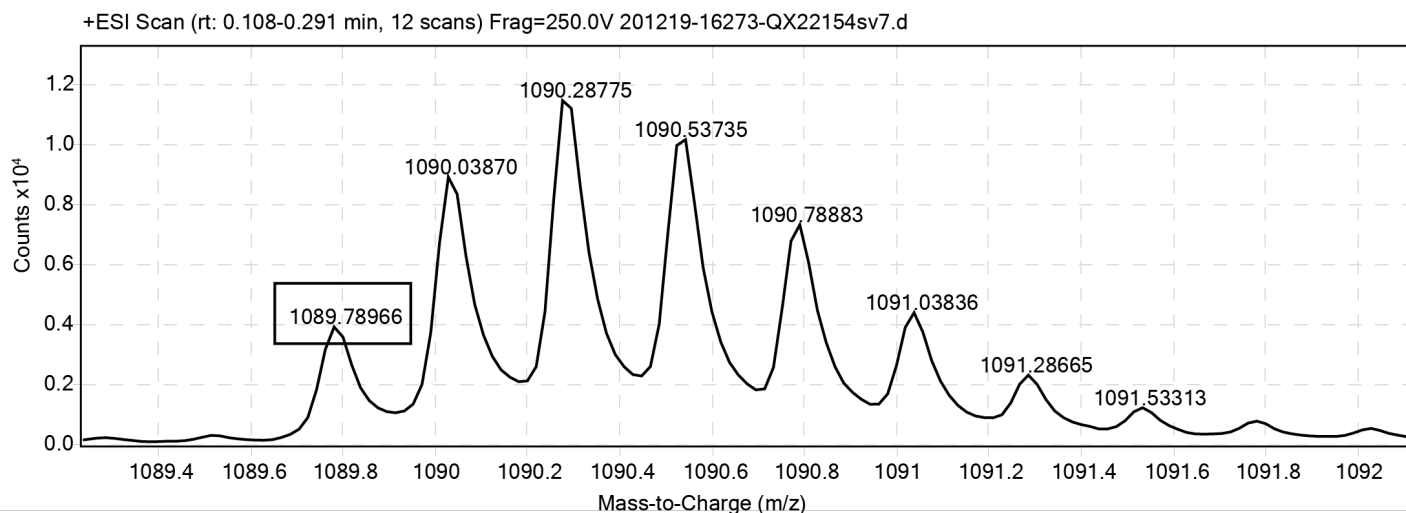
**Figure S35.** ESI-TOF spectrum for WW variant **14/30-o55**. Expected  $[M+4H^+]/4 = 1116.070$  Da. Observed  $[M+4H^+]/4 = 1116.071$  Da.



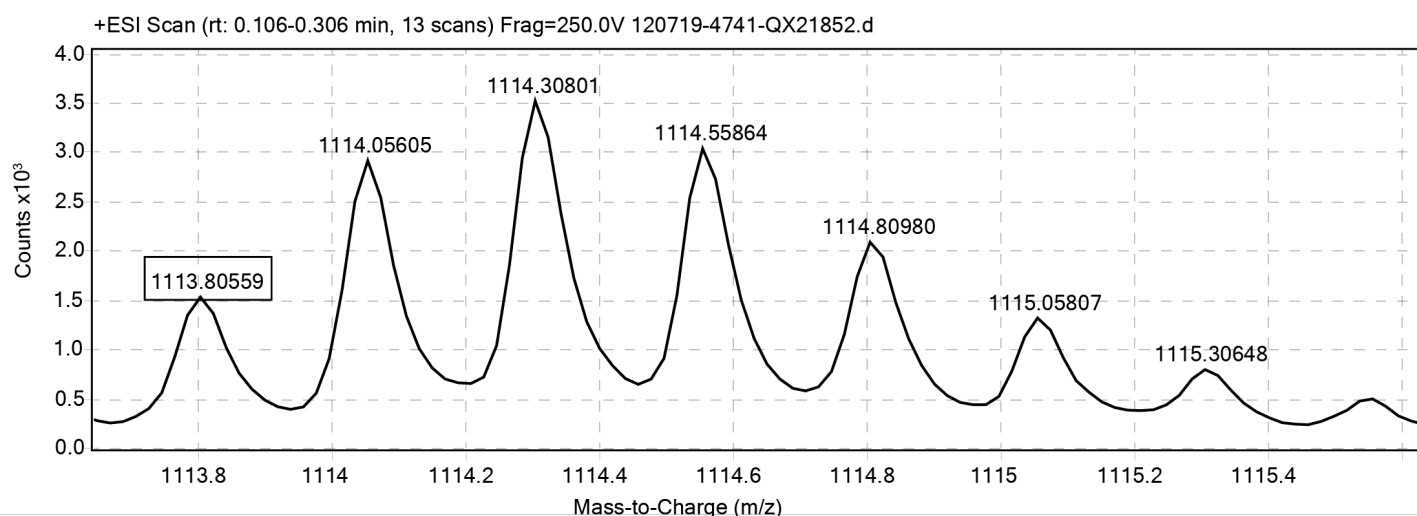
**Figure S36.** ESI-TOF spectrum for WW variant **s14/30-o55**. Expected  $[M+4H^+]/4 = 1109.062$  Da. Obs.  $[M+4H^+]/4 = 1109.058$  Da.



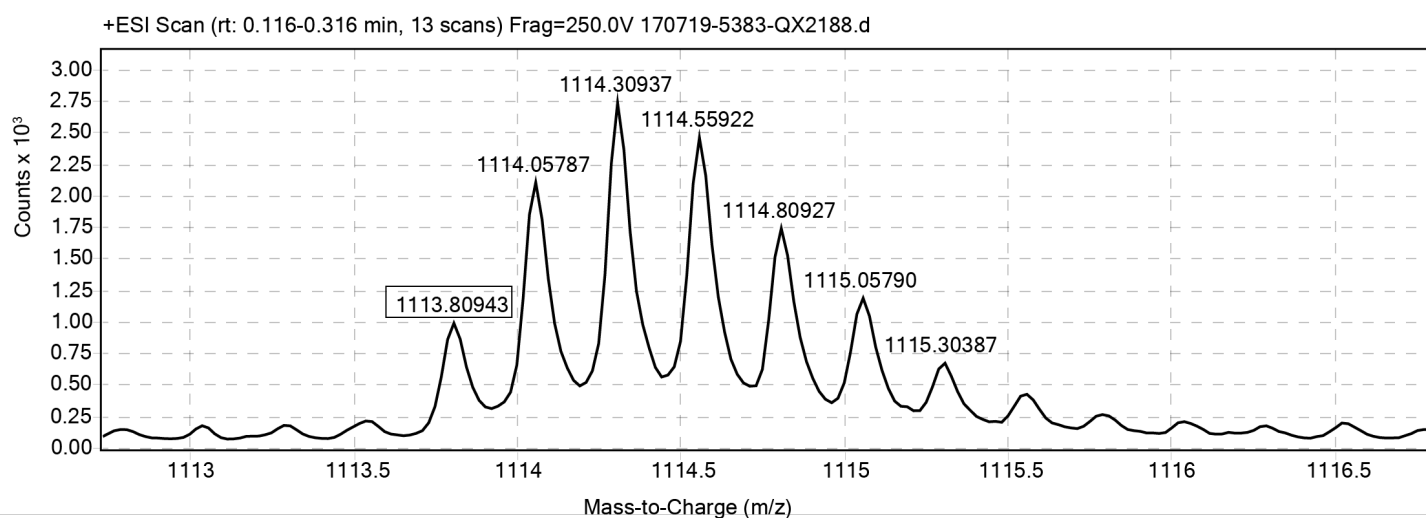
**Figure S37.** ESI-TOF spectrum for WW variant **14/30-c44**. Expected  $[M+4H^+]/4 = 1089.797$  Da. Observed  $[M+4H^+]/4 = 1089.790$  Da.



**Figure S38.** ESI-TOF spectrum for WW variant **s14/30-c44**. Expected  $[M+4H^+]/4 = 1089.797$  Da. Observed  $[M+4H^+]/4 = 1089.790$  Da.



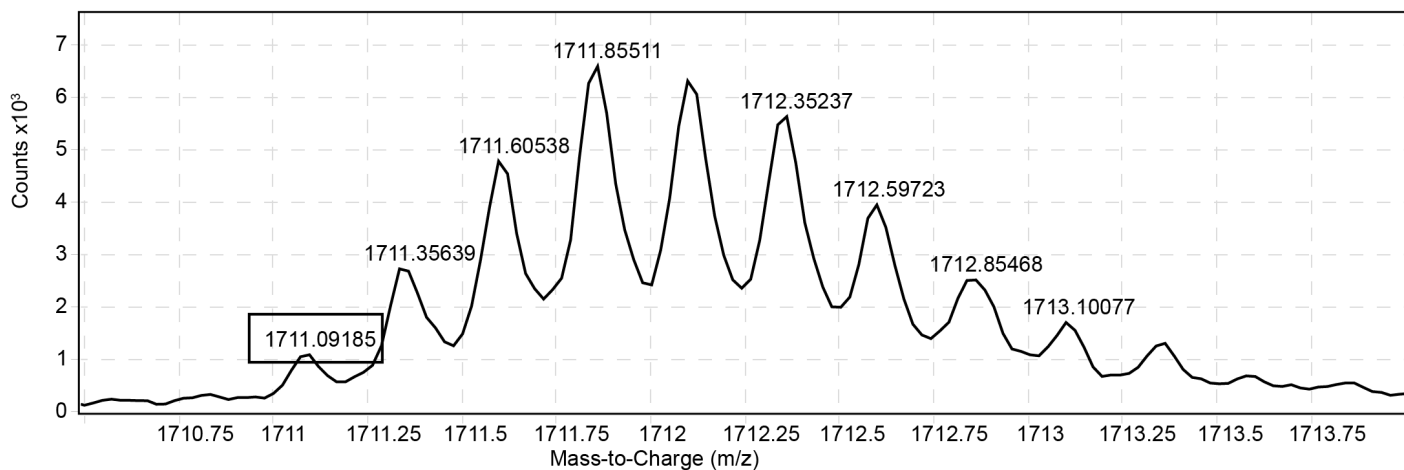
**Figure S39.** ESI-TOF spectrum for WW variant **16/32-c44**. Expected  $[M+4H^+]/4 = 1113.817$  Da. Observed  $[M+4H^+]/4 = 1113.806$  Da.



**Figure S40.** ESI-TOF spectrum for WW variant **s16/32-c44**. Expected  $[M+4H^+]/4 = 1113.817$  Da. Observed  $[M+4H^+]/4 = 1113.809$  Da.

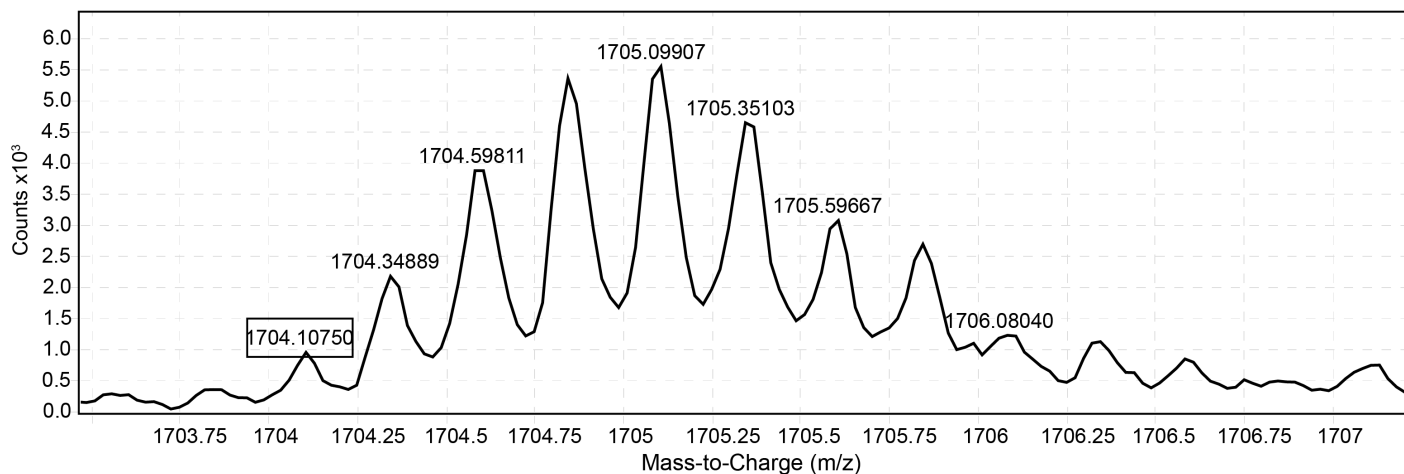


+ESI Scan (rt: 0.099-0.282 min, 12 scans) Frag=250.0V 271018-4452-KT1027v12.d



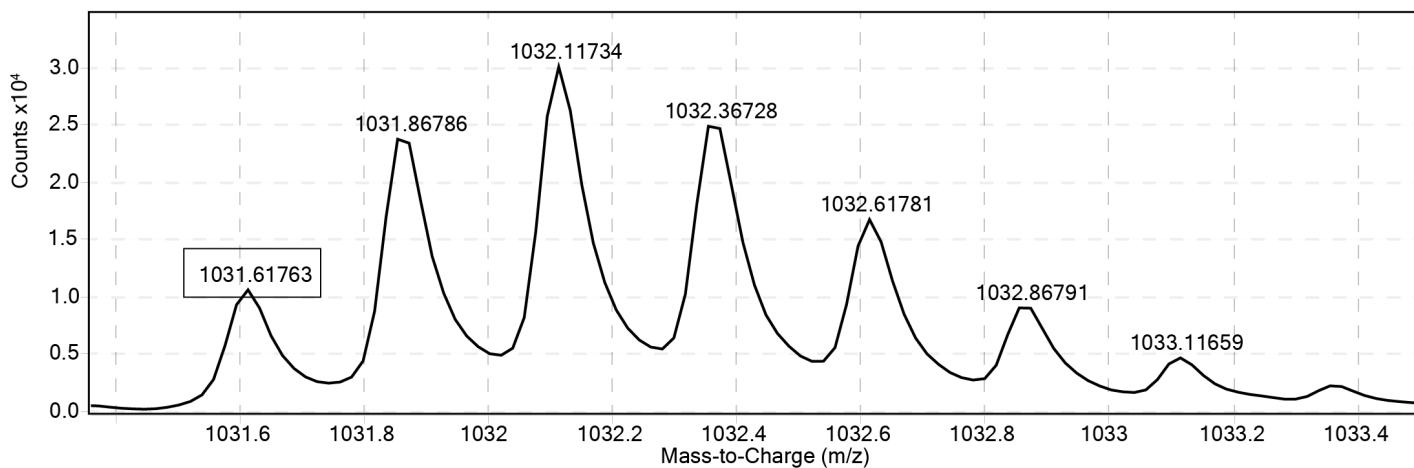
**Figure S41.** ESI-TOF spectrum for SH3 variant **20/37-o44**. Expected  $[M+4H^+]/4 = 1711.114$  Da. Observed  $[M+4H^+]/4 = 1711.092$  Da.

+ Scan (rt: 0.107-0.173 min, 5 scans) 100918-3523-QX2117.d



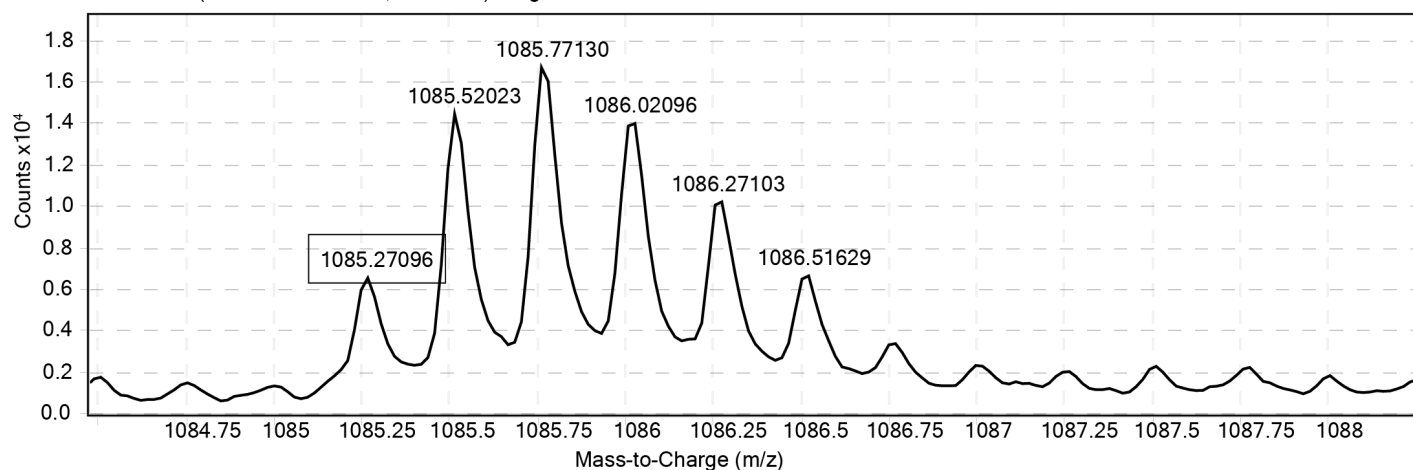
**Figure S42.** ESI-TOF spectrum for SH3 variant **s20/37-o44**. Expected  $[M+4H^+]/4 = 1704.106$  Da. Observed  $[M+4H^+]/4 = 1704.108$  Da.

+ESI Scan (rt: 0.131-0.364 min, 15 scans) Frag=250.0V 081119-13783-QX21971.d



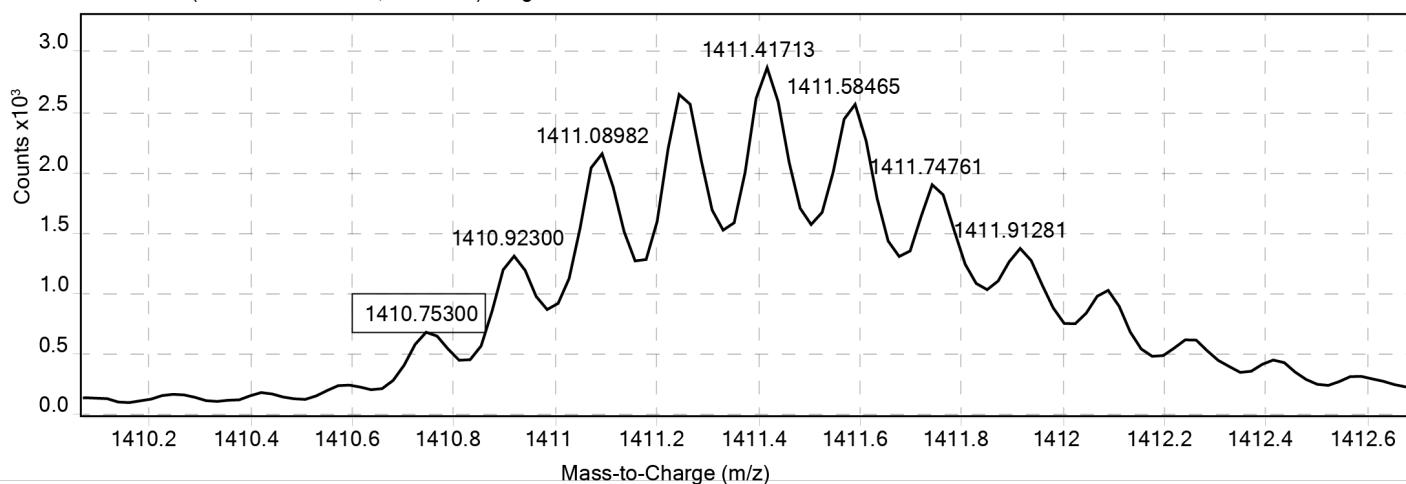
**Figure S43.** ESI-TOF spectrum for GCN4 monomer **27-c4**. Expected  $[M+4H^+]/4 = 1031.620$  Da. Observed  $[M+4H^+]/4 = 1031.618$  Da.

+ESI Scan (rt: 0.118-0.285 min, 11 scans) Frag=250.0V 101119-13946-QX21973v3.d



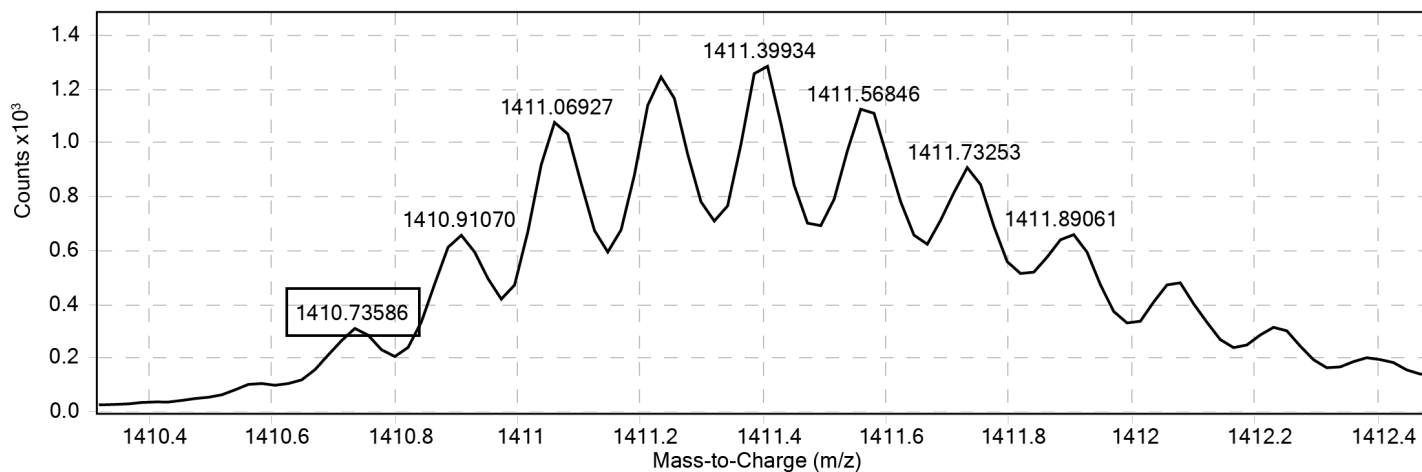
**Figure S44.** ESI-TOF spectrum for GCN4 monomer **29'-c0**. Expected  $[M+4H^+]/4 = 1085.277$  Da. Obs.  $[M+4H^+]/4 = 1085.271$  Da.

+ESI Scan (rt: 0.113-0.263 min, 10 scans) Frag=250.0V 280120-18022-QX22031.d

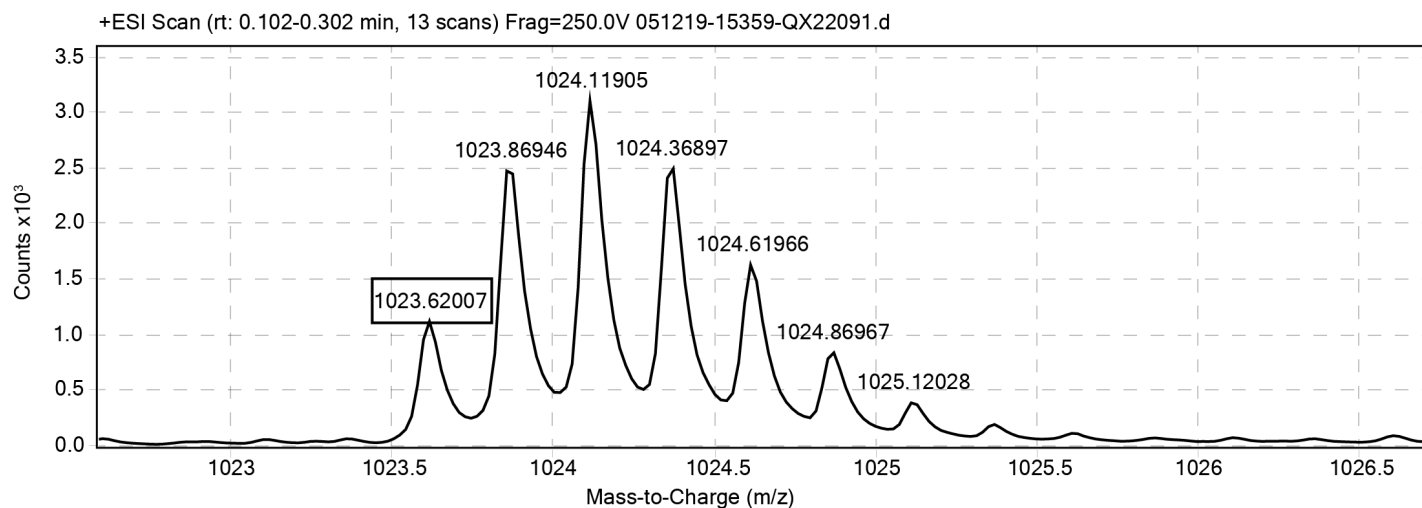


**Figure S45.** ESI-TOF spectrum for disulfide-bound GCN4 heterodimer **d27/29'-c40**. Expected  $[M+6H^+]/6 = 1410.592$  Da. Observed  $[M+6H^+]/6 = 1410.753$  Da.

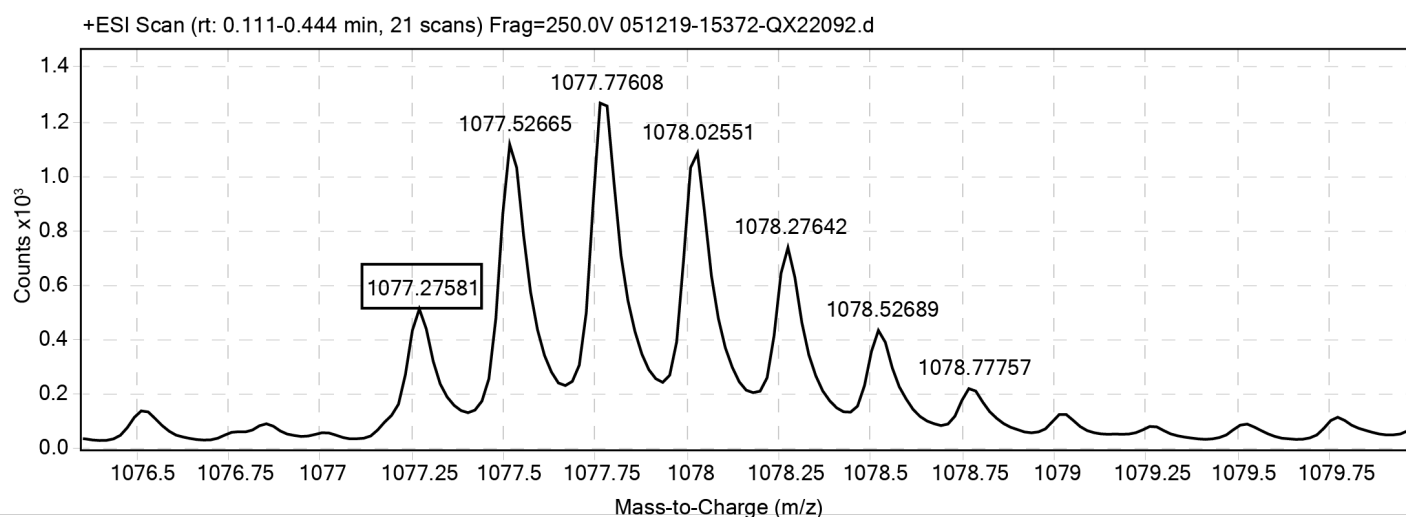
+ESI Scan (rt: 0.086-0.302 min, 14 scans) Frag=250.0V 290120-18067-QX2204.d



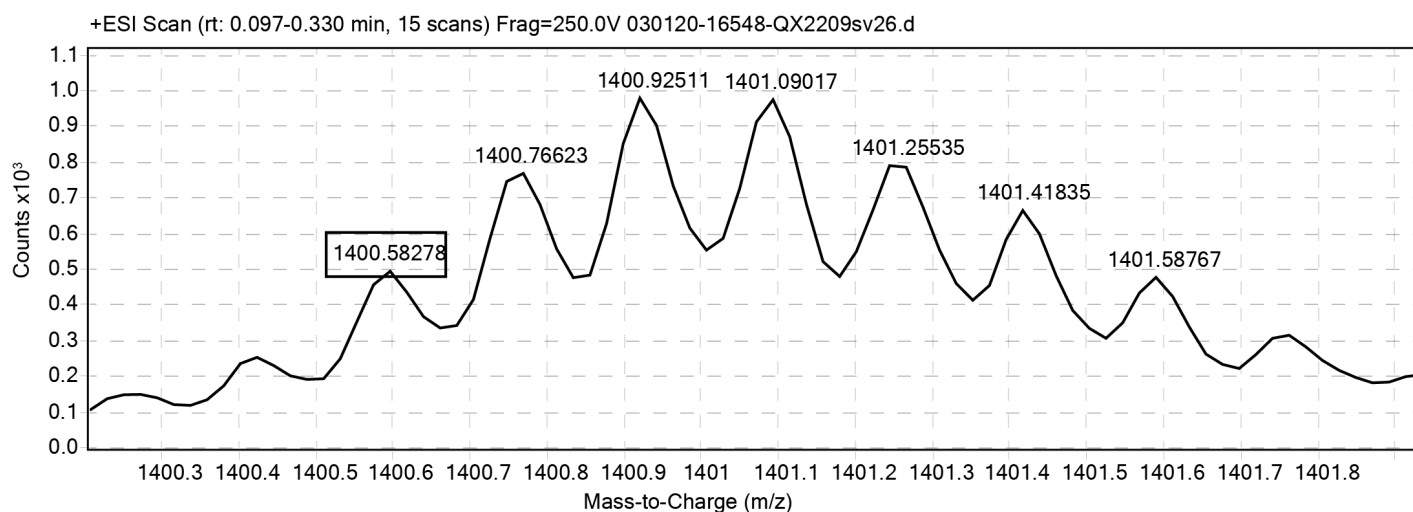
**Figure S46.** ESI-TOF spectrum for triazole-stapled disulfide-bound GCN4 heterodimer **sd27/29'-c40**. Expected  $[M+6H^+]/6 = 1410.592$  Da. Observed  $[M+6H^+]/6 = 1410.736$  Da.



**Figure S47.** ESI-TOF spectrum for GCN4 monomer 27A-c4. Expected  $[M+4H^+]/4 = 1023.627$  Da. Obs.  $[M+4H^+]/4 = 1023.620$  Da.



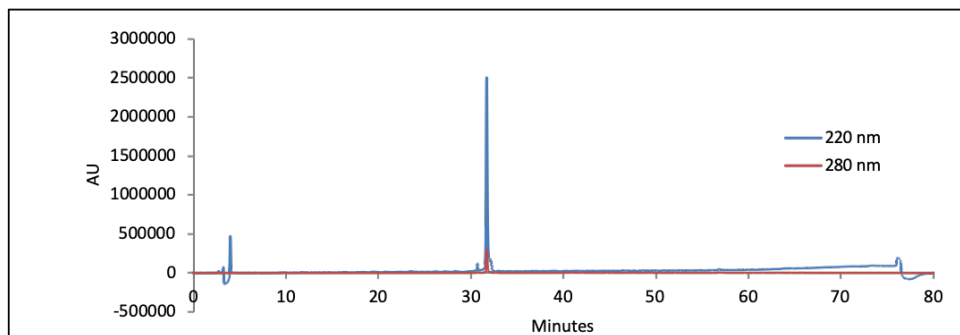
**Figure S48.** ESI-TOF spectrum for GCN4 monomer 29A'-c0. Expected  $[M+4H^+]/4 = 1077.284$  Da. Obs.  $[M+4H^+]/4 = 1077.276$  Da.



**Figure S49.** ESI-TOF spectrum for triazole-stapled GCN4 heterodimer s27/29'-c40. Expected  $[M+6H^+]/6 = 1400.271$  Da. Observed  $[M+6H^+]/6 = 1400.583$  Da.

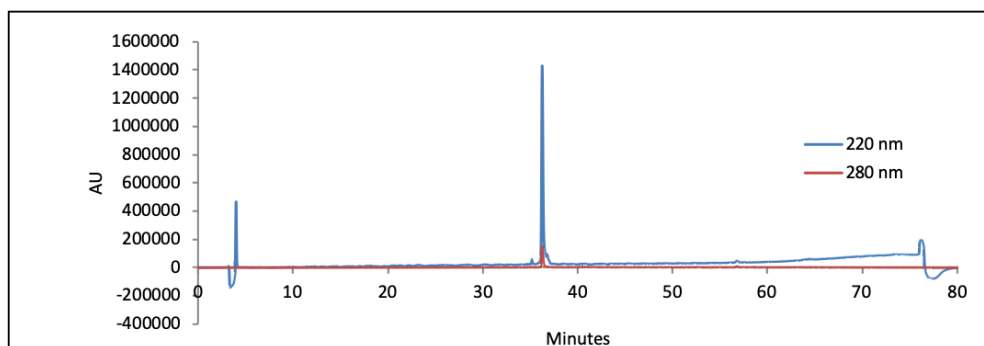
## 6. Analytical HPLC data

Analytical HPLC data WW variants **14/30-00**, **14/30-o55**, **s14/30-o55**, **14/30-c44**, **s14/30-c44**, **16/32-c44** and **s16/32-c44**; SH3 variants **20/37-o44** and **s20/37-o44**; and GCN4 variants **27-c4**, **29'-c0**, **d27/29'-c40**, **sd27/29'-c40**, **27A-c4**, **29A'-c0**, and **s27/29'-40** are shown in Figures S50–S65.



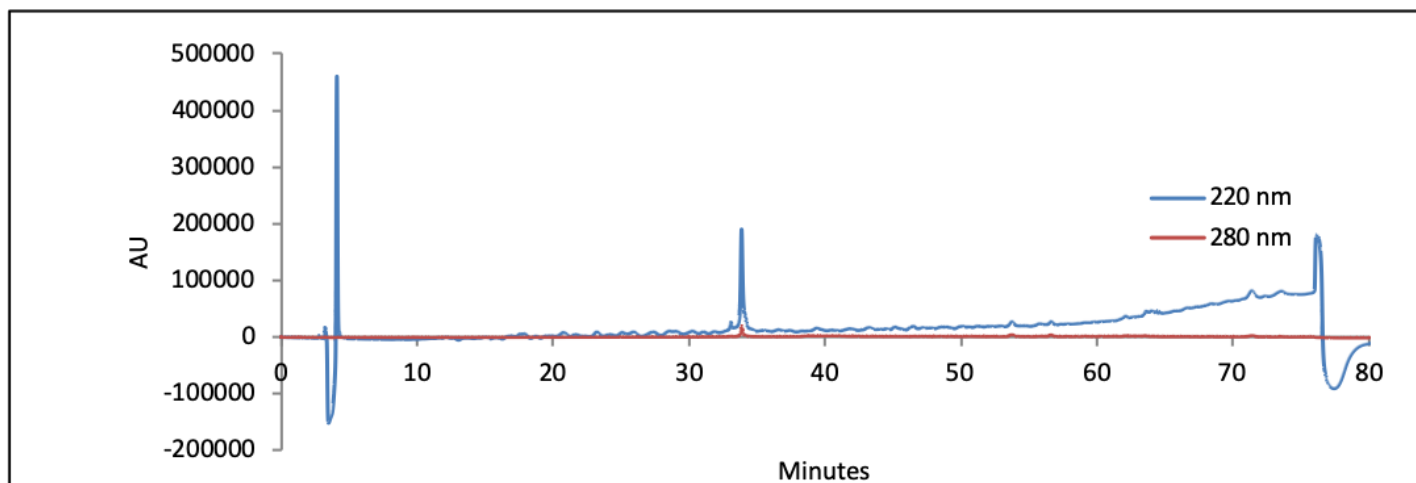
**Figure S50.** Analytical HPLC Data for WW variant **14/30-00** (notebook number **DA10434**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B = MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.

---



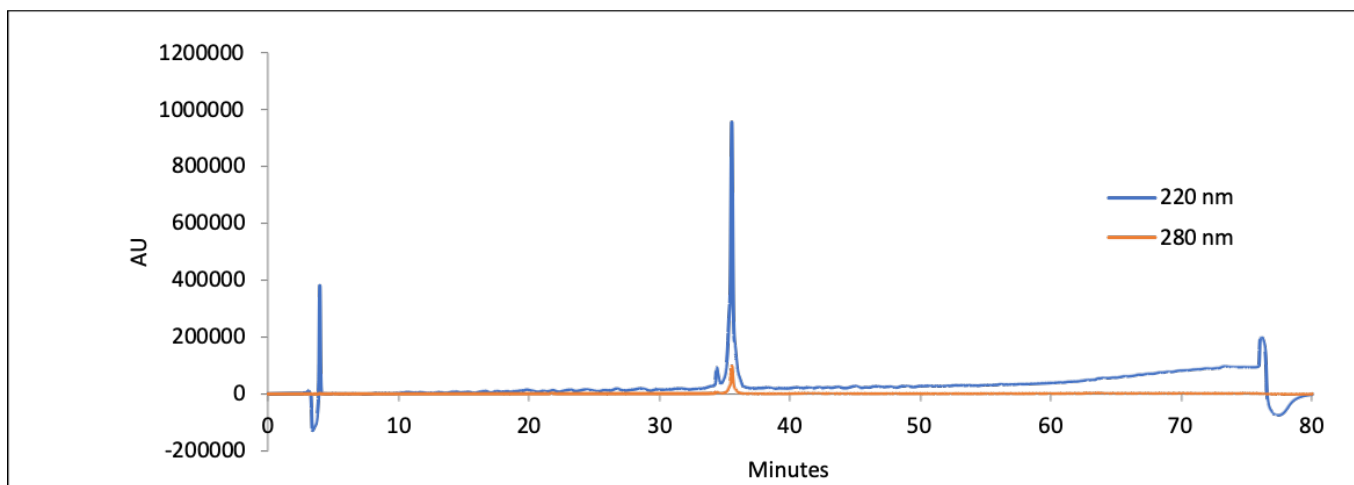
**Figure S51.** Analytical HPLC Data for WW variant **14/30-o55** (notebook number **QX2183**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B = MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.

---

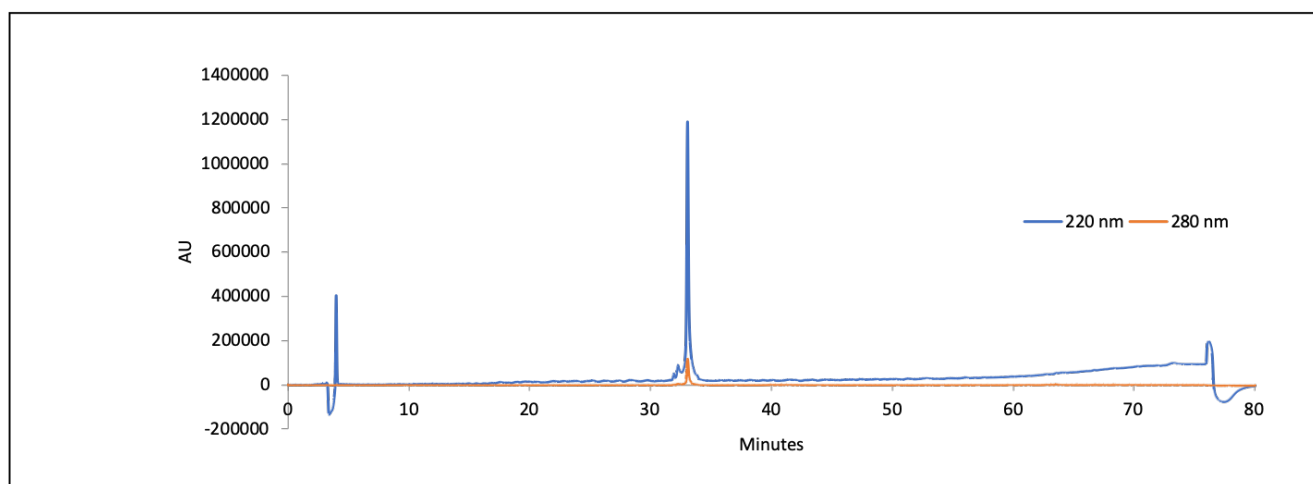


**Figure S52.** Analytical HPLC Data for WW variant **s14/30-o55** (notebook number **QX2183S**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B = MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.

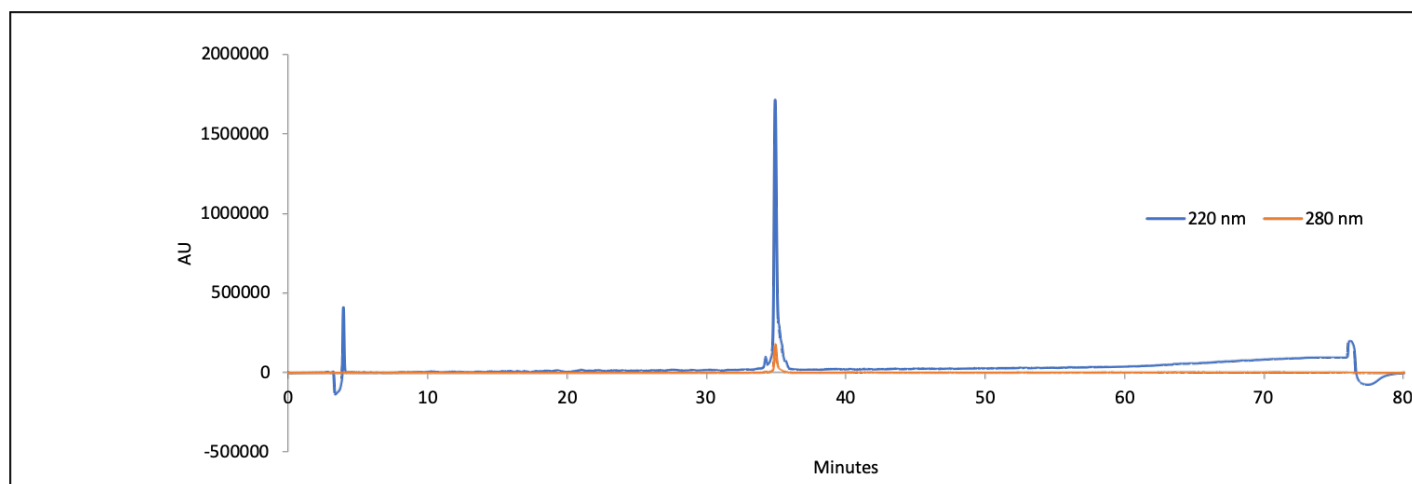
---



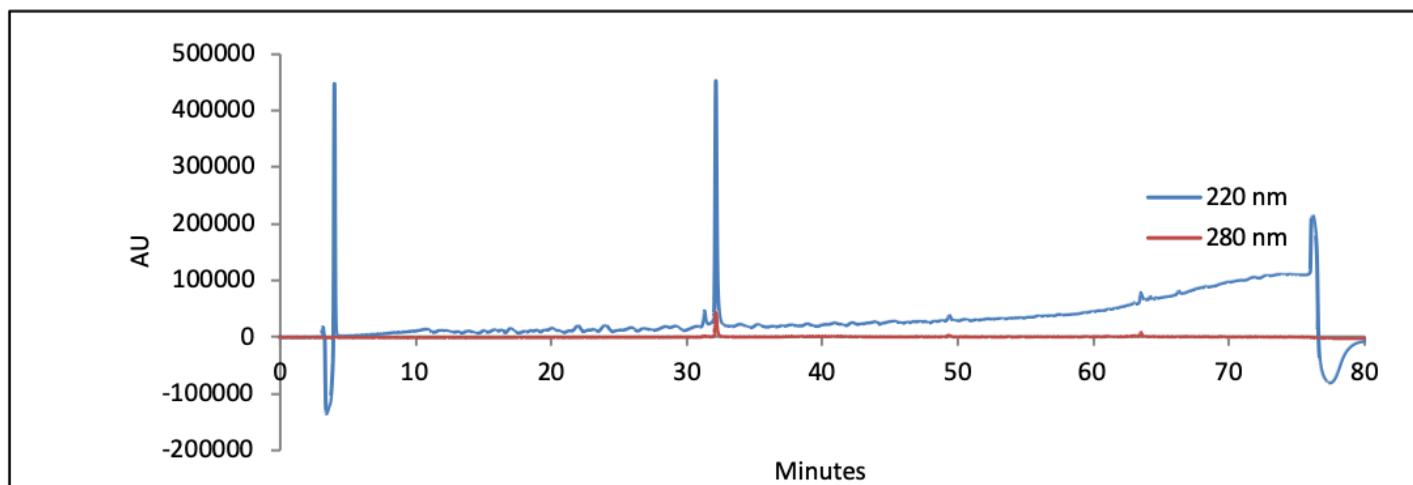
**Figure S53.** Analytical HPLC Data for WW variant **14/30-c44** (notebook number **QX22154**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



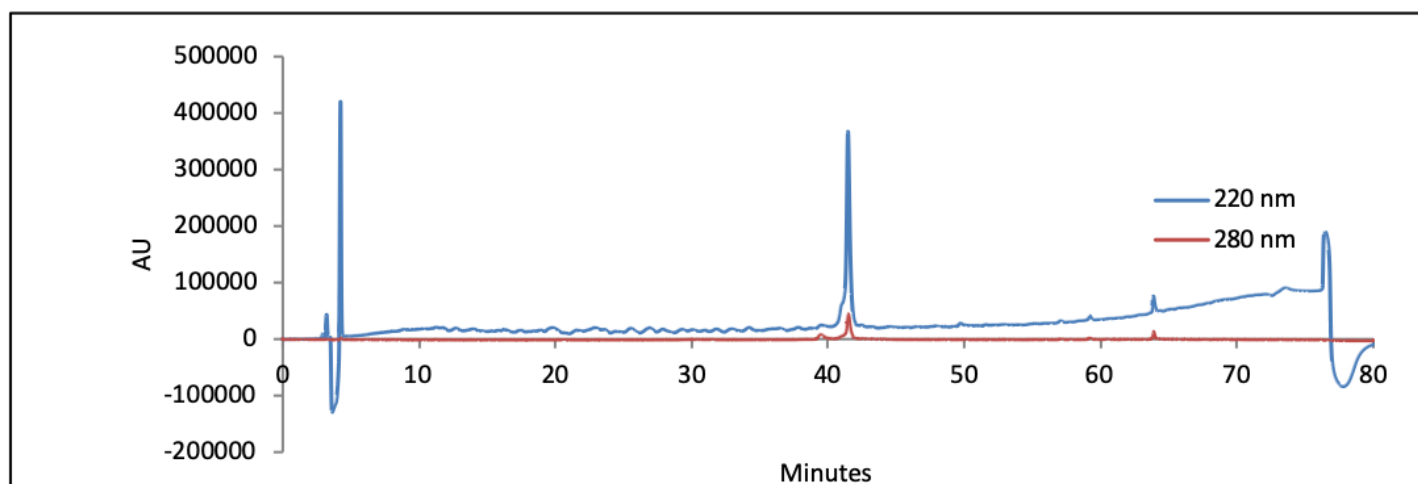
**Figure S54.** Analytical HPLC Data for WW variant **s14/30-c44** (notebook number **QX22154s**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



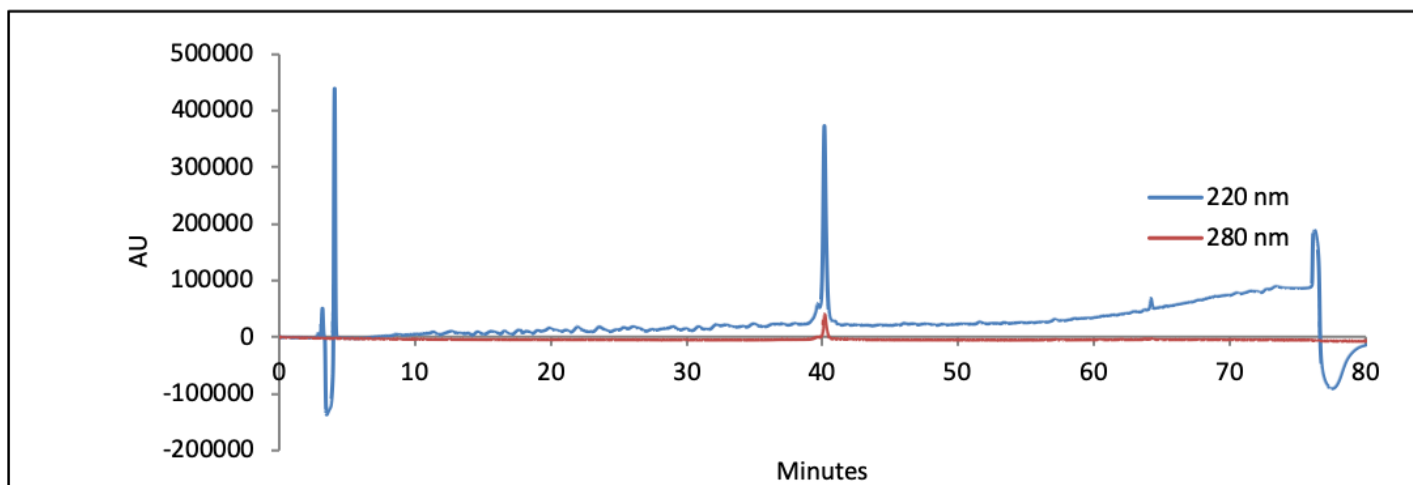
**Figure S55.** Analytical HPLC Data for WW variant **16/32-c44** (notebook number **QX21852**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



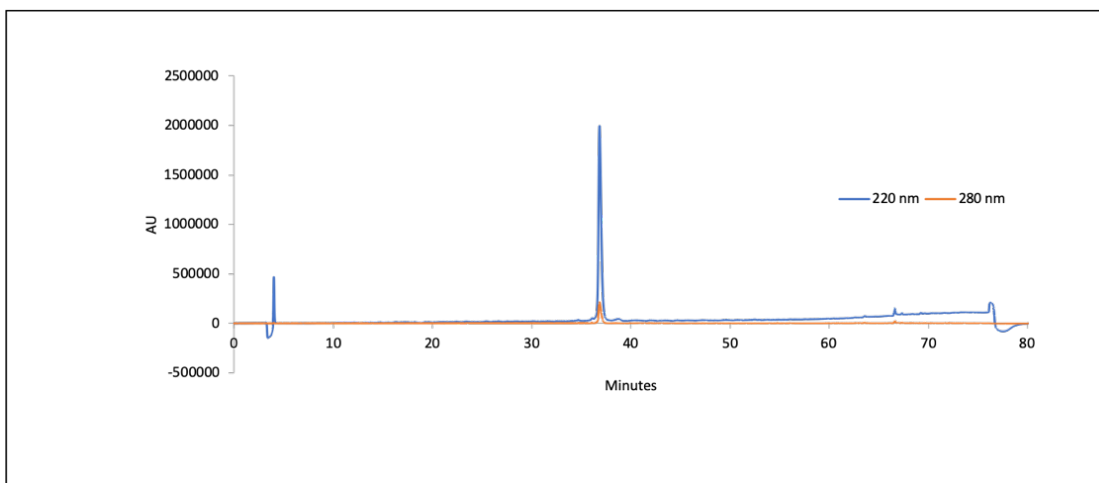
**Figure S56.** Analytical HPLC Data for WW variant **s16/32-c44** (notebook number **QX2188**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



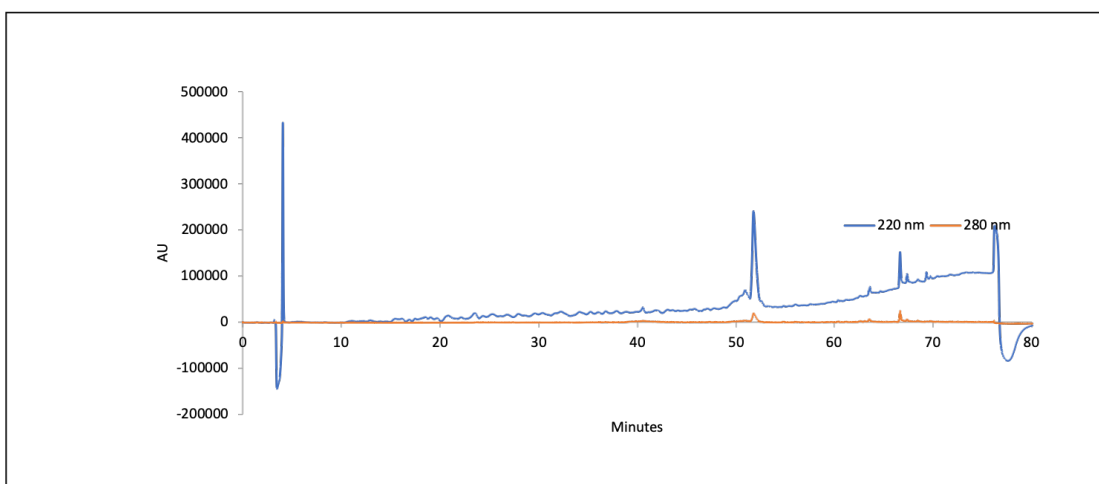
**Figure S57.** Analytical HPLC Data for SH3 variant **20/37-o44** (notebook number **KT1027**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



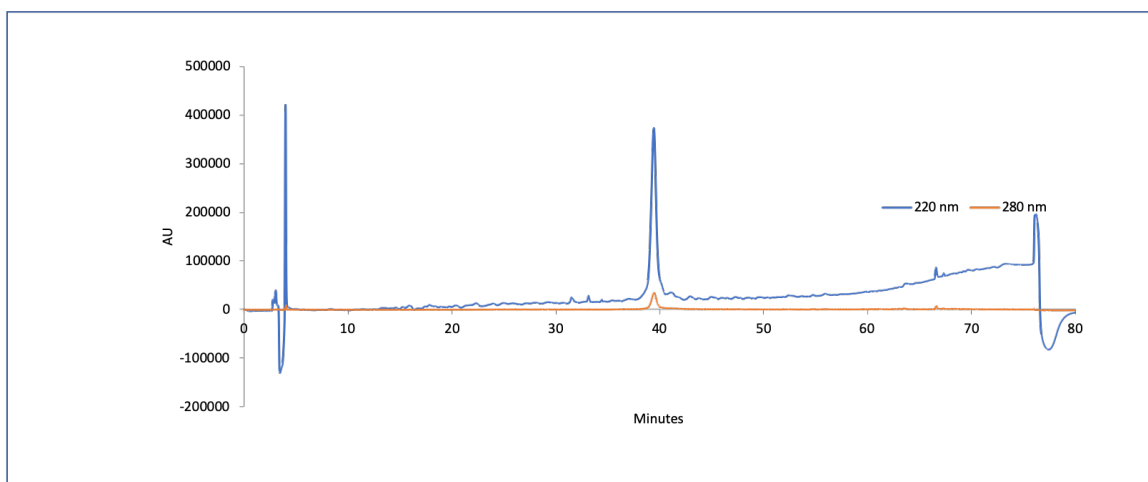
**Figure S58.** Analytical HPLC Data for SH3 variant **s20/37-44** (notebook number **QX2117**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



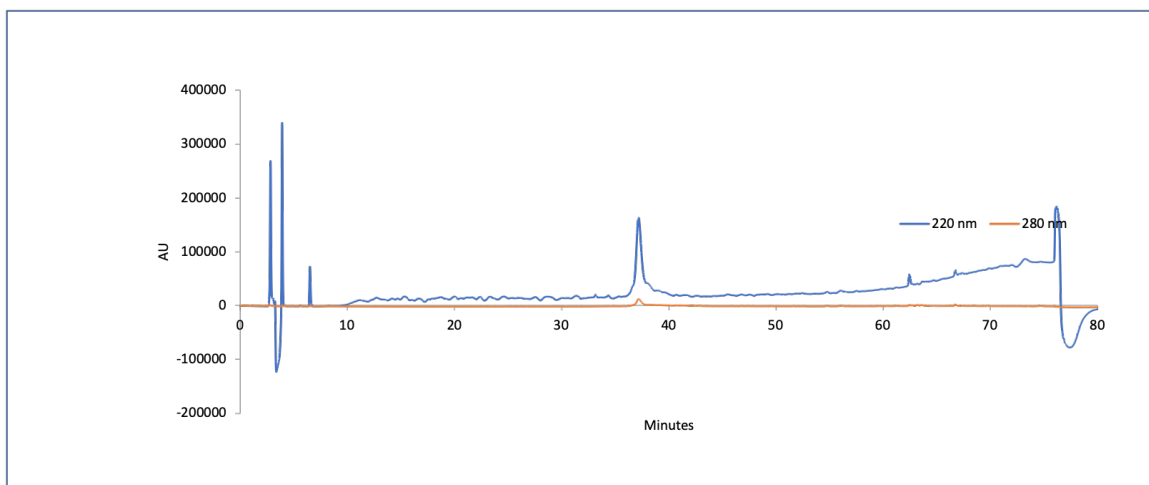
**Figure S59.** Analytical HPLC Data for GCN4 monomer **27-c4** (notebook number **QX21971**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



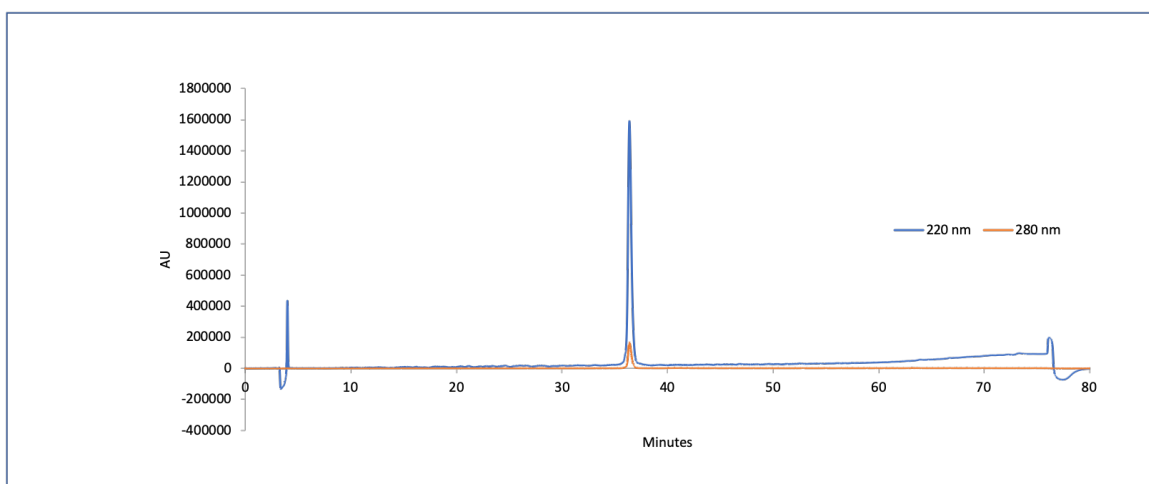
**Figure S60.** Analytical HPLC Data for GCN4 monomer **29'-c0** (notebook number **QX21972**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



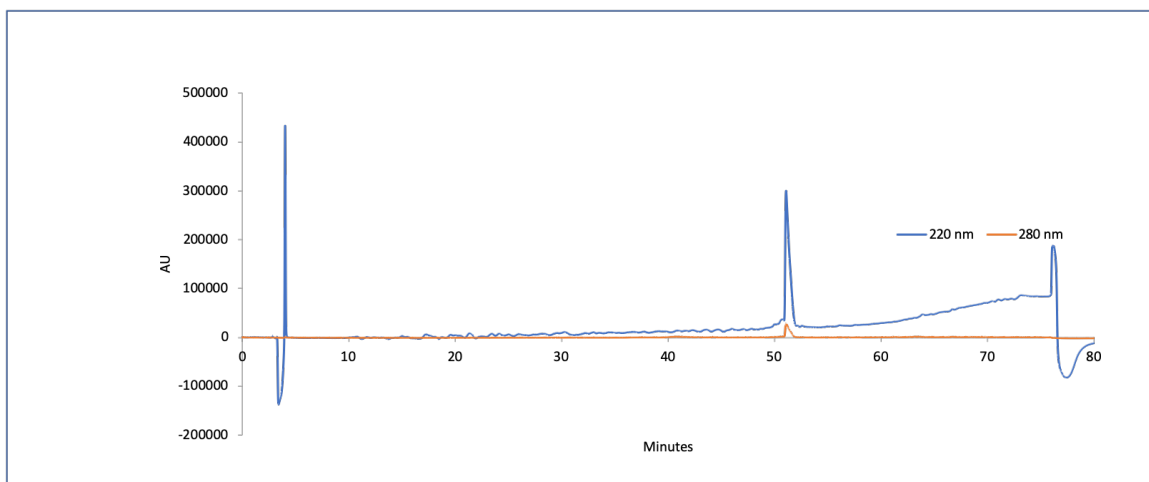
**Figure S61.** Analytical HPLC Data for disulfide-bound GCN4 heterodimer **d27/29'-c40** (notebook number **QX22031**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



**Figure S62.** Analytical HPLC Data for triazole-stapled disulfide-bound GCN4 variant **sd27/29'-c40** (notebook number **QX2204**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.

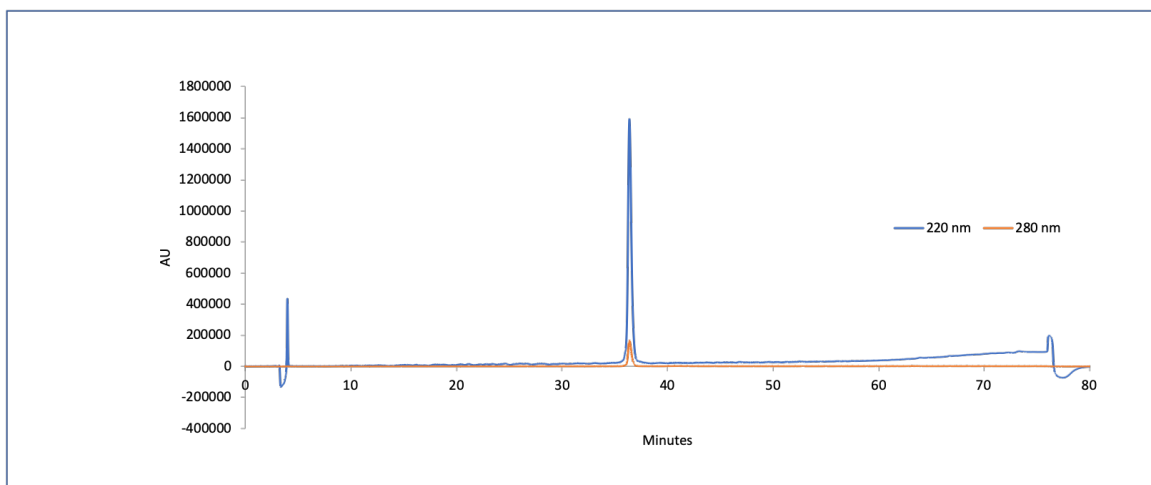


**Figure S63.** Analytical HPLC Data for GCN4 monomer **27A'-c4** (notebook number **QX22091**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



**Figure S64.** Analytical HPLC Data for GCN4 monomer **29A'-c0** (notebook number **QX22092**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.



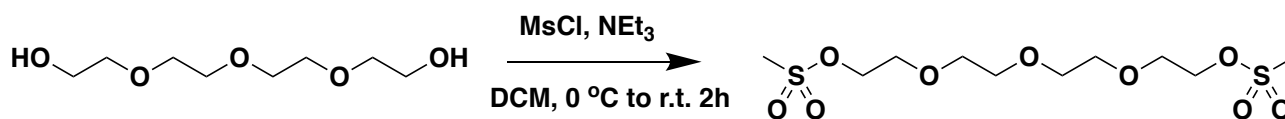


**Figure S65.** Analytical HPLC Data for triazole-stapled GCN4 heterodimer **s27/29'-c40** (notebook number **QX2214**). Protein solution was injected onto a C18 analytical column and eluted using a linear gradient of 10-60% B (A = H<sub>2</sub>O, 0.1% TFA; B= MeCN, 0.1% TFA) over 50 minutes, followed by a 10-minute rinse (95% B), and a 10-minute column re-equilibration.

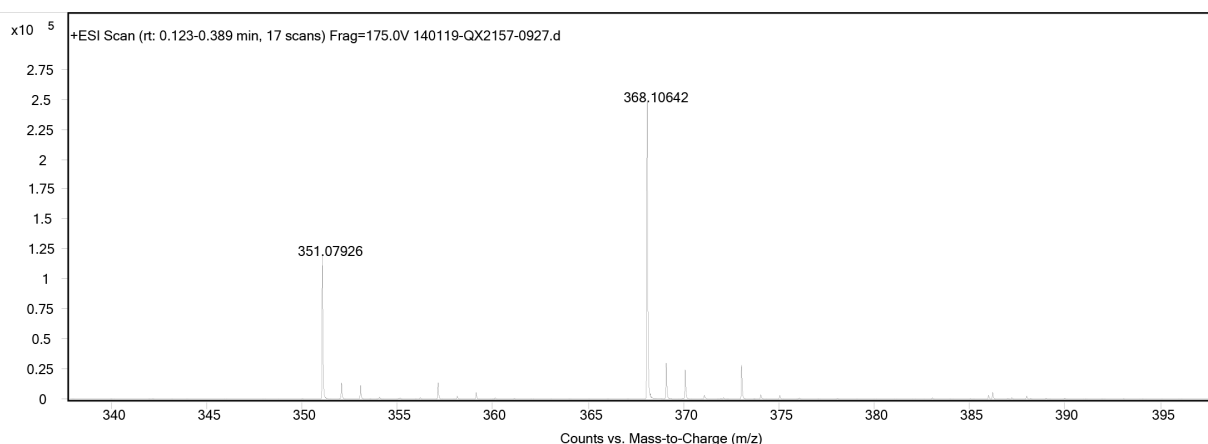
---

## 7. Synthesis and characterization of PEGylated Asn derivatives

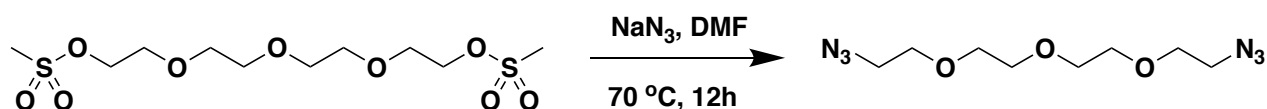
### ((Oxybis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl) dimethanesulfonate (QX2157)



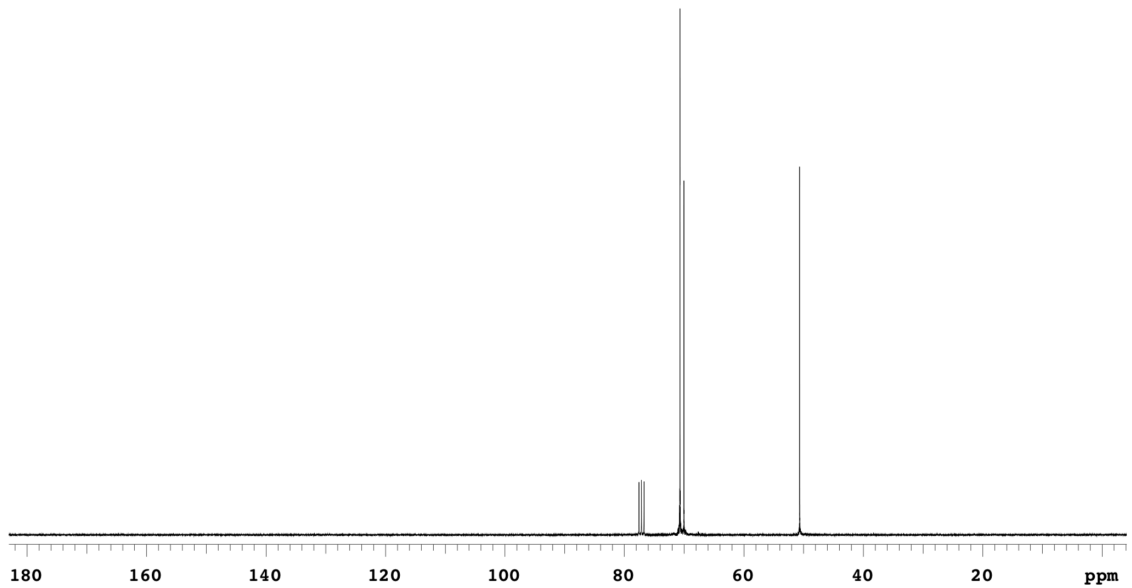
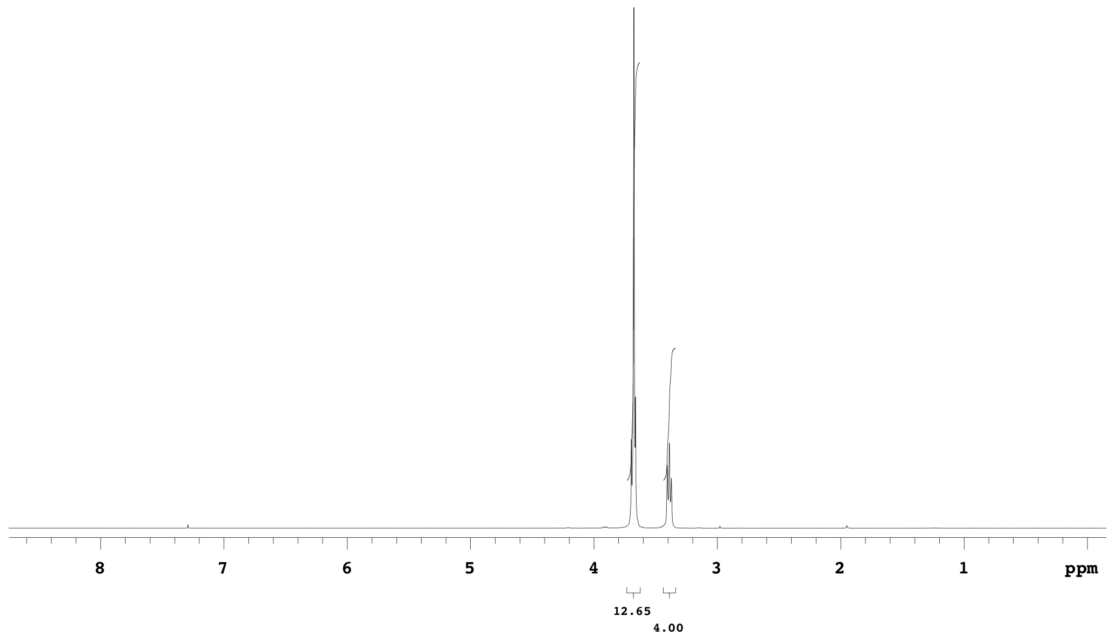
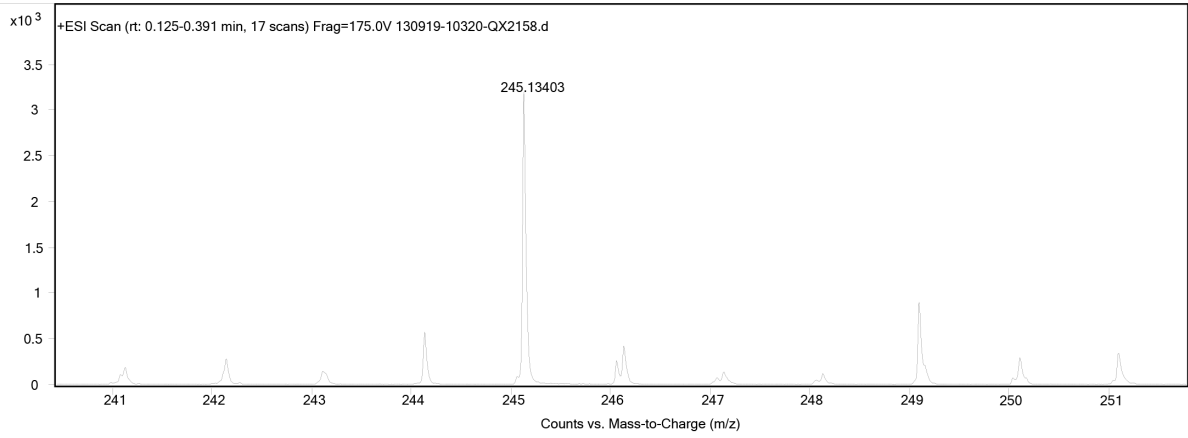
Methanesulfonyl chloride (23 g, 200 mmol) was added dropwise to a stirred solution of tetra ethylene glycol (10 g, 50 mmol) and TEA (20 g, 200 mmol) in dichloromethane (250 mL) at 0 °C. After the addition was complete, the resulting mixture was stirred at r.t. for 15 h. Water was added to quench the reaction. The organic phase was separated, and the aqueous phase was extracted with dichloromethane (2 × 100 mL). The combined organic layers were washed with brine (3 × 100 mL), dried with anhydrous sodium sulfate, filtered, and the solvent was removed by rotary evaporation to afford 17.5 g of colorless oil, which was used in the next step without purification. Yield quantitative. MS(ESI-TOF)  $m/z$  calc. for  $C_{10}H_{23}O_9S_2^+$  351.08, found 351.08  $[M+H^+]$ ; calc. for  $C_{10}H_{26}NO_9S_2^+$  368.10, found 368.11  $[M+NH_4^+]$ .

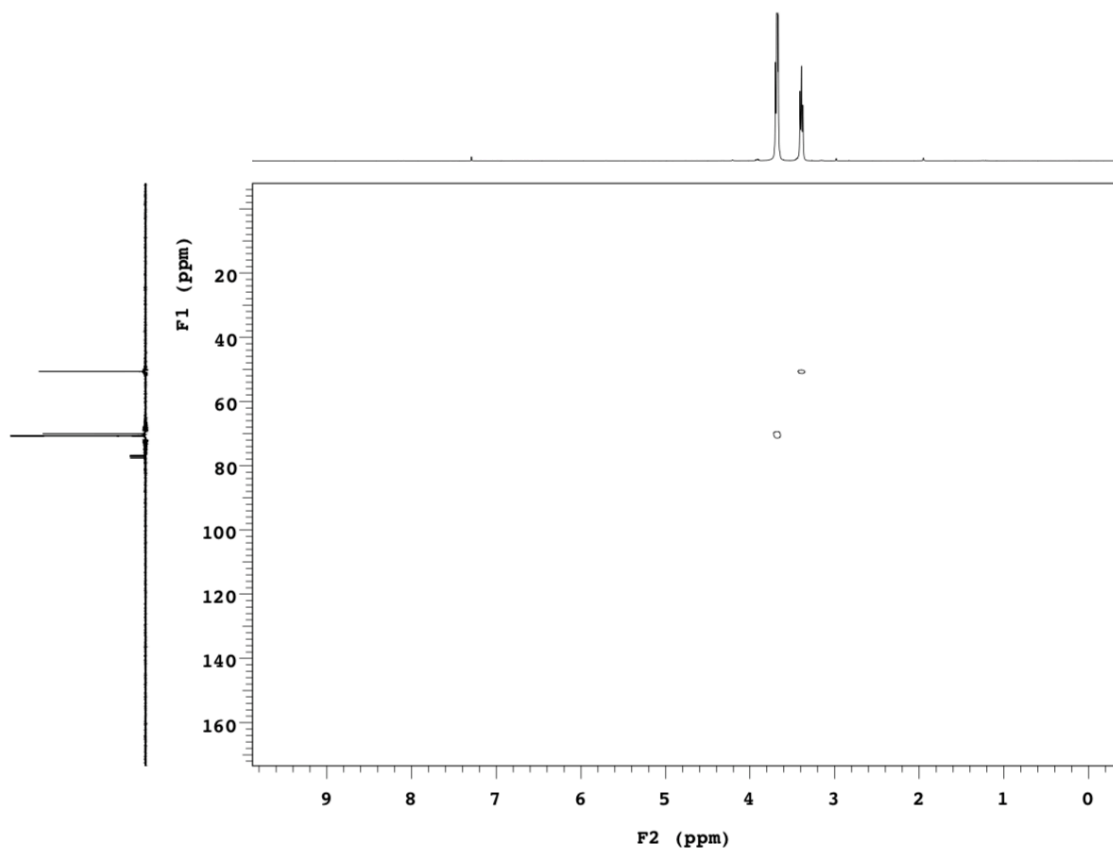


### 1-Azido-2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethane (QX2158)

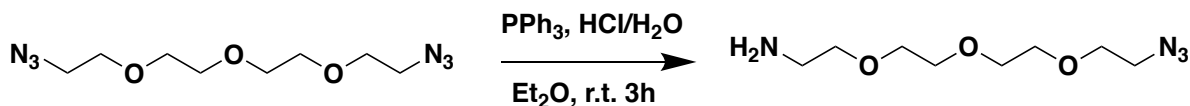


To a solution of  $NaN_3$  (13 g, 200 mmol) in DMF (200 mL) was added **QX2157** (17.5 g, 50 mmol) at room temperature. The reaction mixture was heated to 70 °C and stirred for 12 hours. The crude mixture was extracted with DCM for 3 times after the reaction was completed. The combined organic phases were washed with saturated brine and dried over anhydrous sodium sulfate, evaporated and purified by flash column (Hexane/EA 1:1 to 0:1) to obtain a colorless oil. (10 g, 82% yield). MS(ESI-TOF)  $m/z$  calc. for  $C_8H_{17}N_6O_3^+$  245.14, found 245.13  $[M+H^+]$ .  $^1H$  NMR (300 MHz, Chloroform-*d*)  $\delta$  3.70 – 3.66 (m, 12H), 3.39 (t,  $J$  = 5.5 Hz, 4H).  $^{13}C$  NMR (126 MHz, Chloroform-*d*)  $\delta$  70.74, 70.73, 70.71, 70.70, 70.66, 70.04, 50.67.

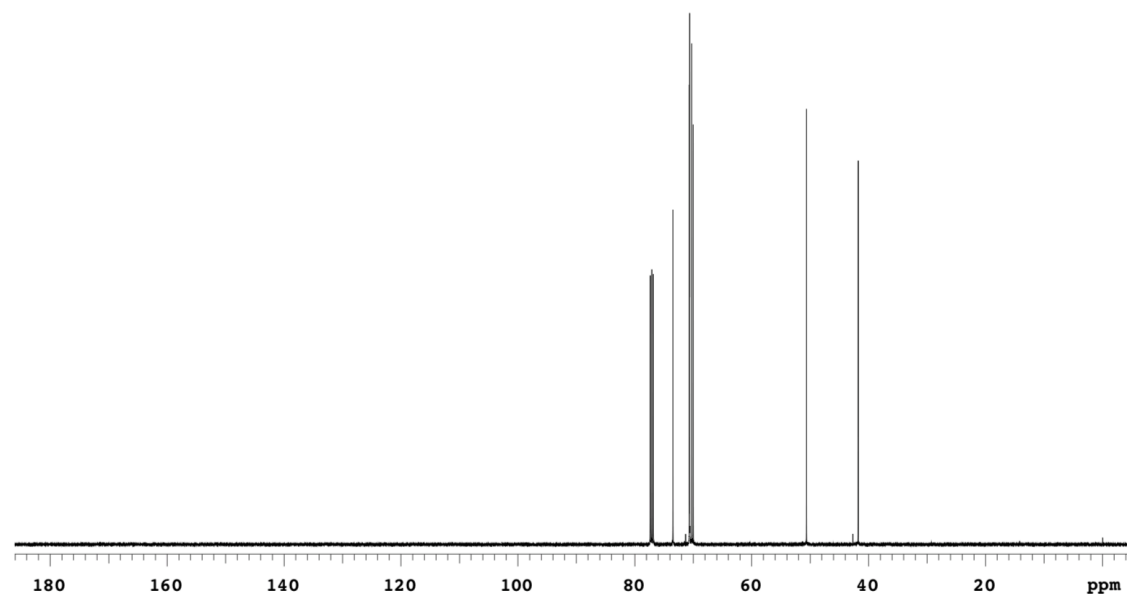
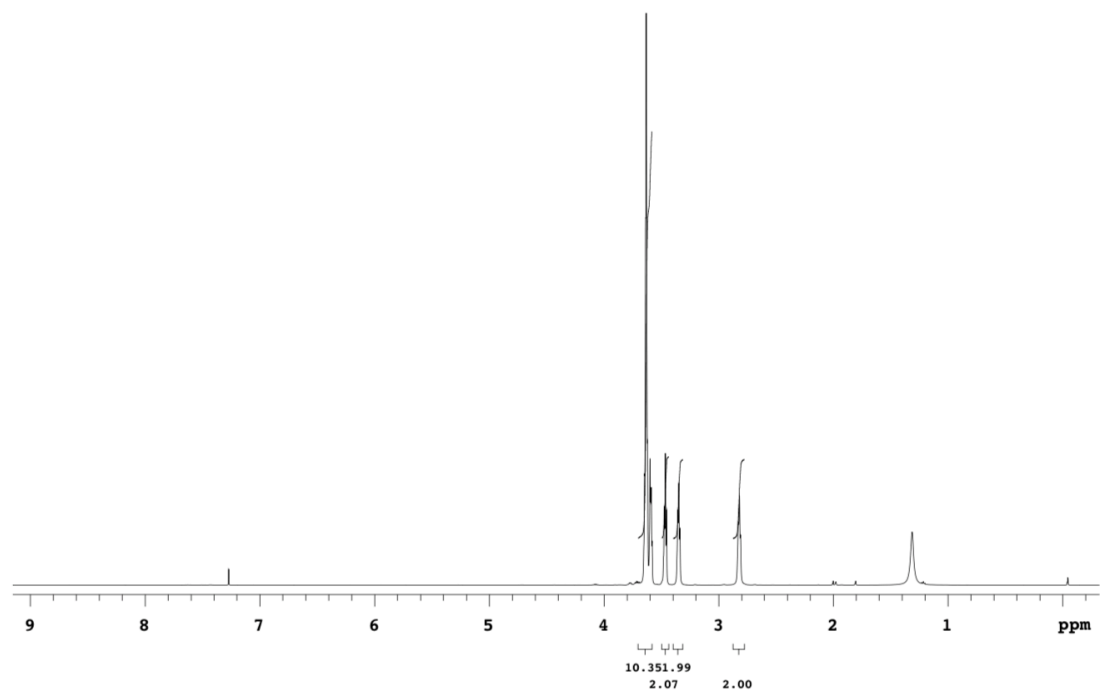
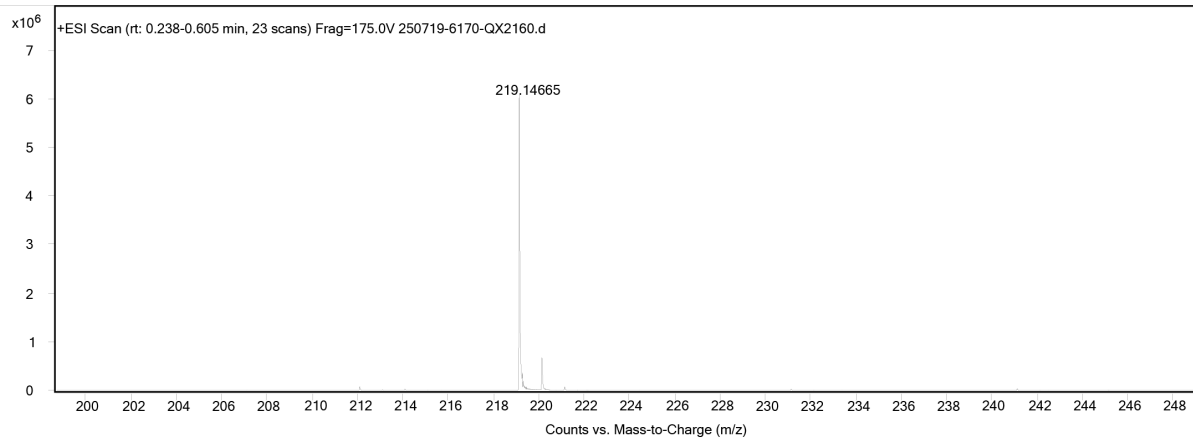


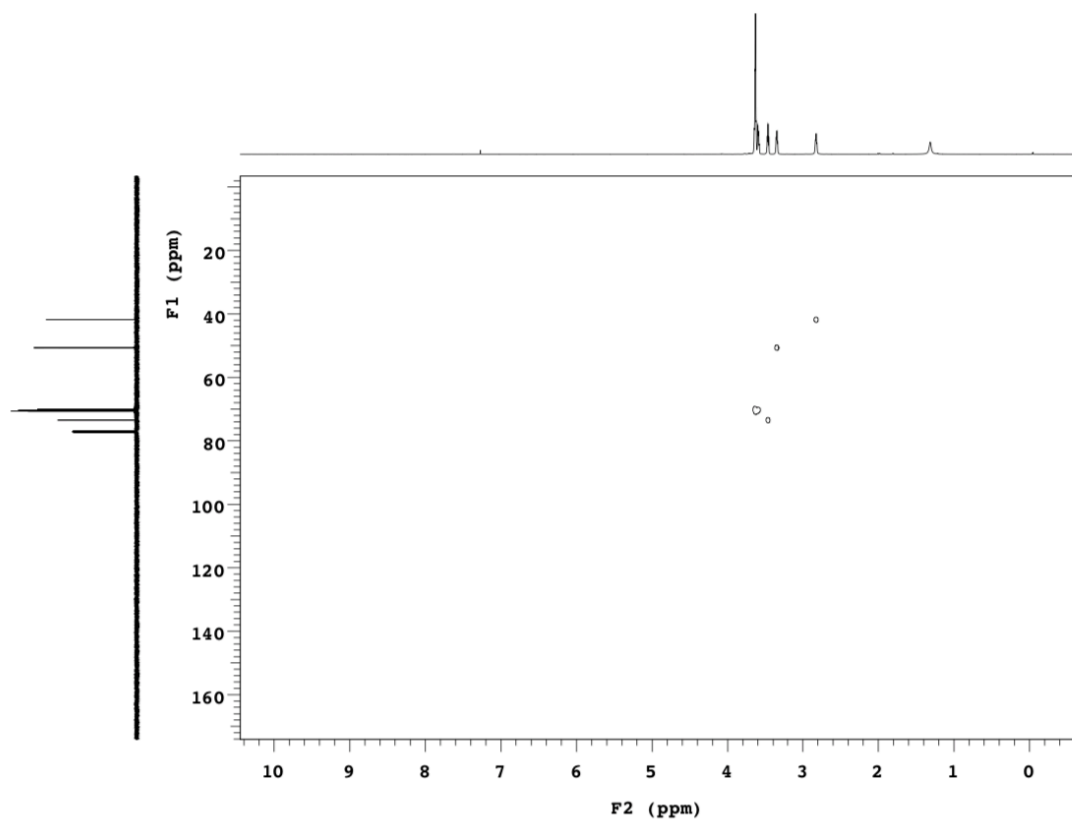


**2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethan-1-amine (QX2160)**

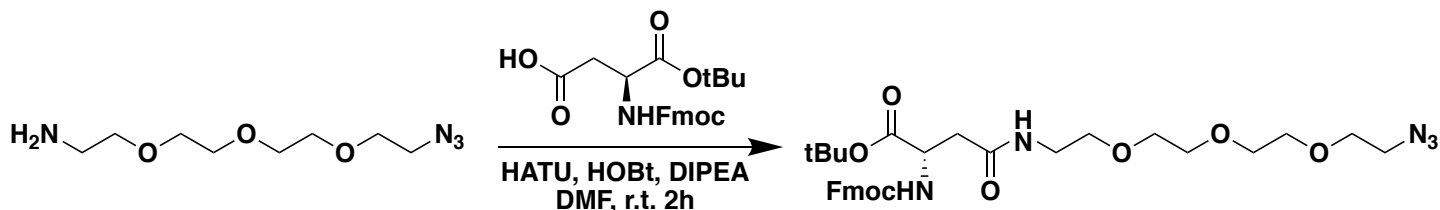


Triphenylphosphine (4 g, 15 mmol, 0.9 eq.) dissolved in ether (75 mL) was added to a solution of **QX2158** (4 g, 16.5 mmol) in 5% aqueous HCl (50 mL). Addition was performed in 30 minutes at room temperature and the reaction was stirred for additional 2.5 hours. Phases were separated by a separation funnel and the aqueous layer was washed with DCM (3 × 50 mL). The aqueous layer was adjusted to pH 10 using sodium hydroxide powder. Product was then extracted with DCM (3 × 75 mL). Combined organic layer was dried over anhydrous sodium sulfate and filtered. After removal of the solvent under reduced pressure, a yellow oil was afforded (1.5 g, yield 46%) MS(ESI-TOF) *m/z* calc. for C<sub>8</sub>H<sub>19</sub>N<sub>4</sub>O<sub>3</sub><sup>+</sup> 219.15, found 219.15 [M+H<sup>+</sup>]. <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.65 – 3.62 (m, 10H), 3.60 – 3.58 (m, 2H), 3.47 (t, *J* = 5.5 Hz, 2H), 3.35 (t, *J* = 5.5 Hz, 2H), 2.8 (t, *J* = 5.5 Hz, 2H). <sup>13</sup>C NMR (500 MHz, Chloroform-*d*) δ 73.49, 70.69, 70.63, 70.61, 70.26, 70.02, 50.64, 41.7.

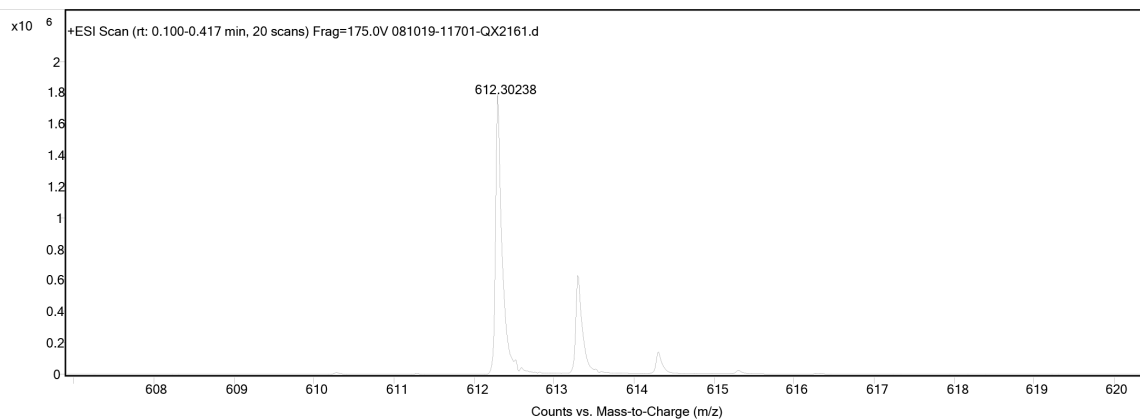




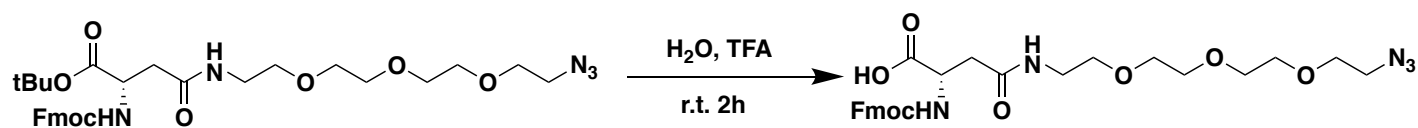
**tert-Butyl (S)-15-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-1-azido-13-oxo-3,6,9-trioxa-12-azahexadecan-16-oate (QX2161)**



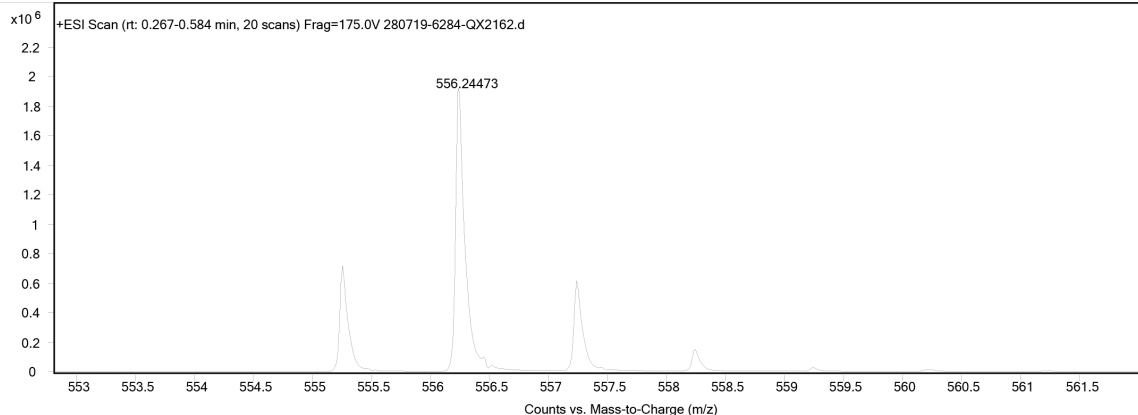
To Fmoc-Asp-tBu (1.15 g, 2.8 mmol) dissolved in dry DMF (20 ml) was added HATU (1.6 g, 4.2 mmol), HOBT (567.5 mg, 4.2 mmol), DIPEA (1.46 ml, 8.4 mmol). Then the mixture was stirred for 15 minutes at room temperature. Then compound **QX2160** (730 mg, 3.25 mmol, dissolved in 5 ml DMF) was added to the mixture and the system was stirred for another 2 hours at room temperature. Upon completion of the reaction monitored by TLC, 30 ml water was added to the flask and extracted 3 times with ethyl acetate. The organic phases were combined, washed with saturated brine, dried over anhydrous sodium sulfate and evaporated to dryness. The crude was used in the next step without purification; MS(ESI-TOF)  $m/z$  calc. for  $C_{31}H_{41}N_5O_8H^+$  612.30, found 612.30  $[M+H^+]$ ;

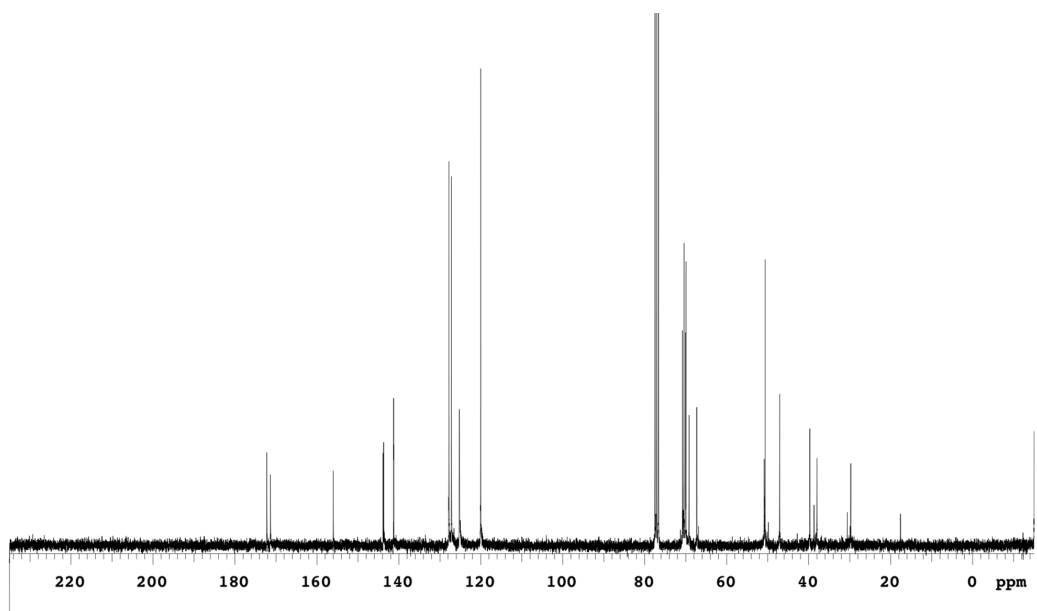
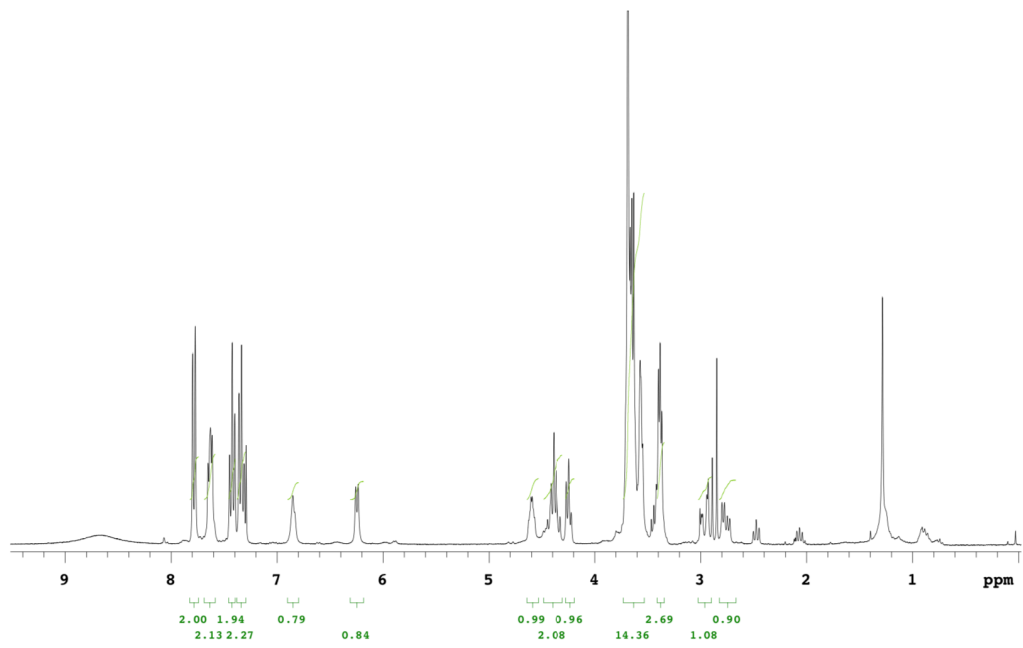


**(S)-15-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-1-azido-13-oxo-3,6,9-trioxa-12-azahexadecan-16-oic acid (QX2162)**

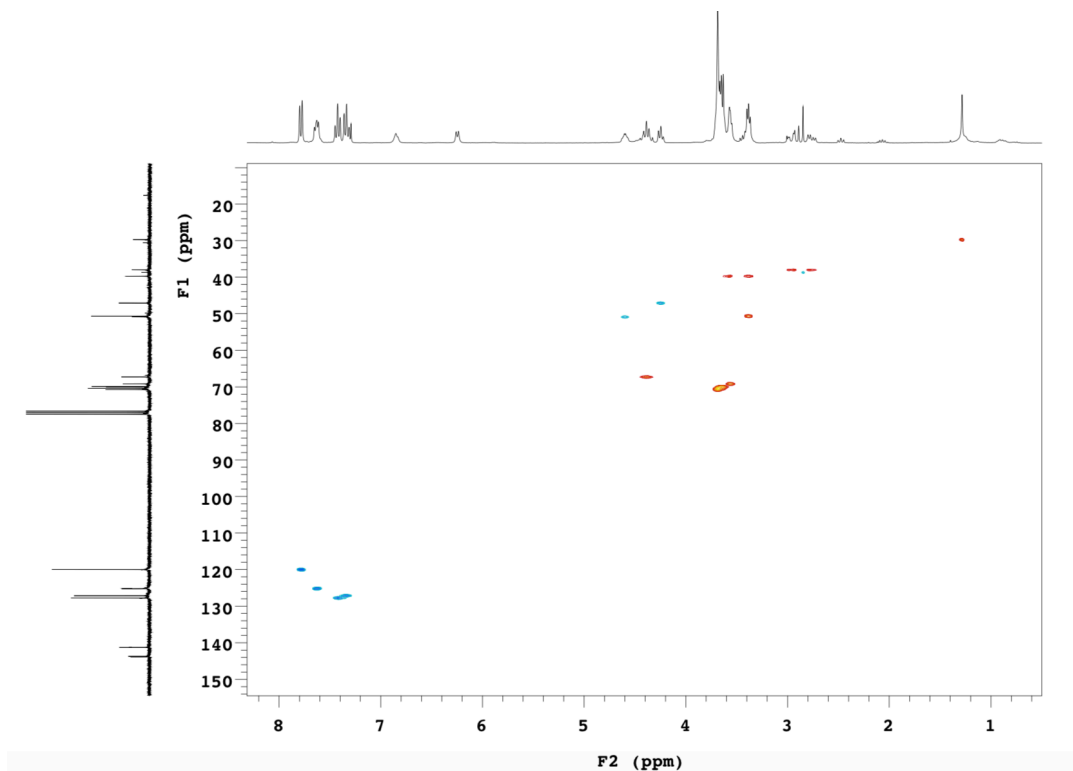


Compound **QX2161** (800 mg, 1.31 mmol) was dissolved in a mixture of TFA (3 ml) and water (150  $\mu$ L). Then the reaction mixture was stirred at room temperature for 2 hours. After completion of the reaction, TFA and water were removed by rotary evaporation. Yield quantitative; MS(ESI-TOF)  $m/z$  calc. for  $C_{27}H_{34}N_5O_8H^+$  556.24, found 556.24  $[M+H^+]$ ;  $^1H$  NMR (300 MHz, Chloroform- $d$ )  $\delta$  7.78 (d,  $J = 6.5$  Hz, 2H), 7.63 (t,  $J = 6.5$  Hz, 2H), 7.42 (t,  $J = 6.5$  Hz, 2H), 7.34 (t,  $J = 6.5$  Hz, 2H), 6.85 (br s., 1H), 6.25 (d,  $J = 6.9$  Hz, 1H), 4.60 (m., 1H), 4.48 – 4.33 (m, 2H), 4.25 (t,  $J = 6.9$  Hz, 1H), 3.68 – 3.63 (m, 12H), 3.57 – 3.55 (m, 2H), 3.42 – 3.37 (m, 2H), 2.97 (dd,  $J = 15, 7.2$  Hz, 1H), 2.76 (dd,  $J = 15, 7.2$  Hz, 1H).  $^{13}C$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  172.22, 171.34, 156.03, 143.88, 141.29, 127.76, 127.14, 125.25, 125.19, 120.00, 70.75, 70.44, 70.40, 70.02, 69.94, 50.85, 50.63, 47.07, 39.71, 38.69, 37.98.

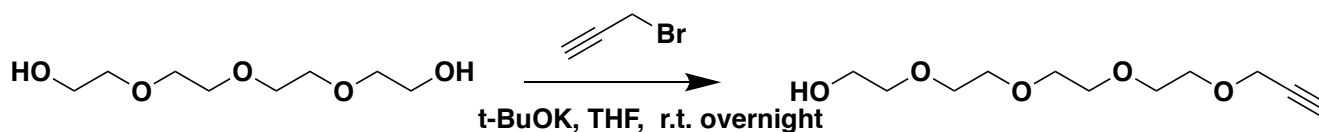




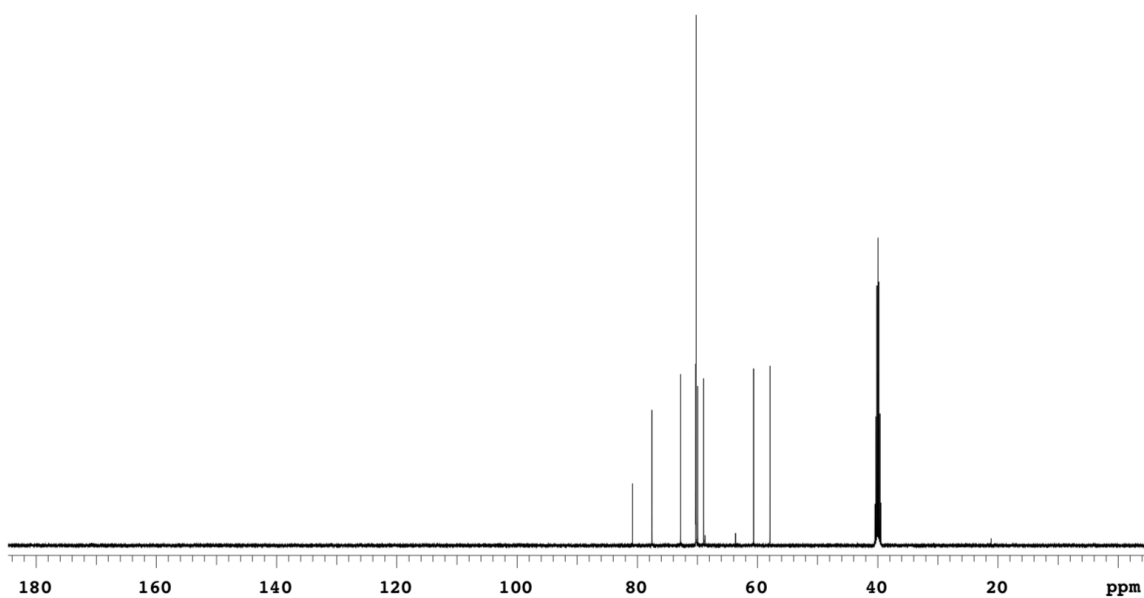
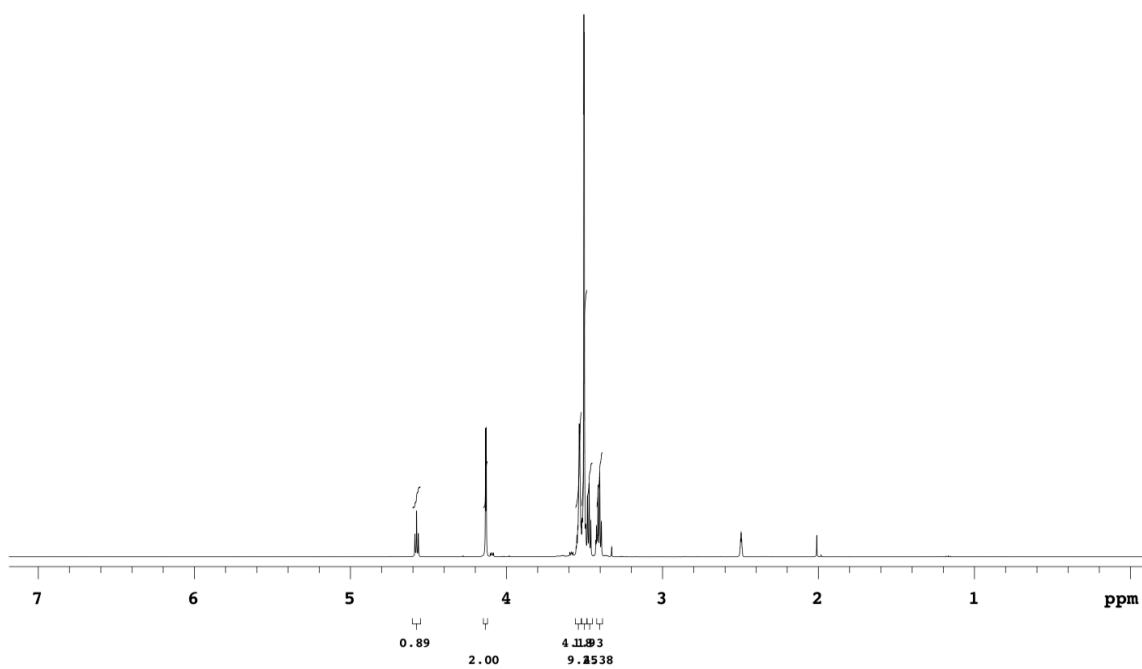
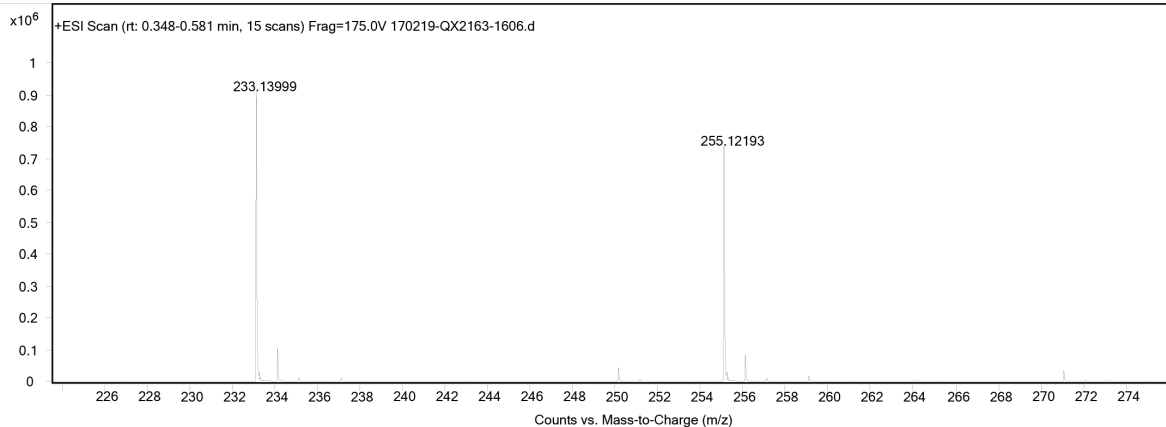




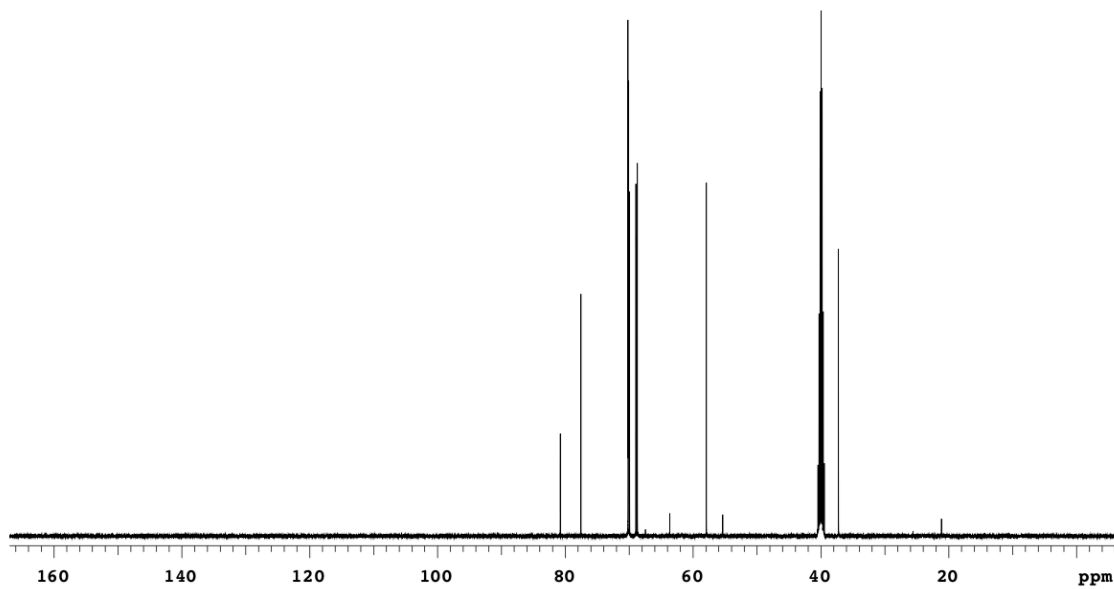
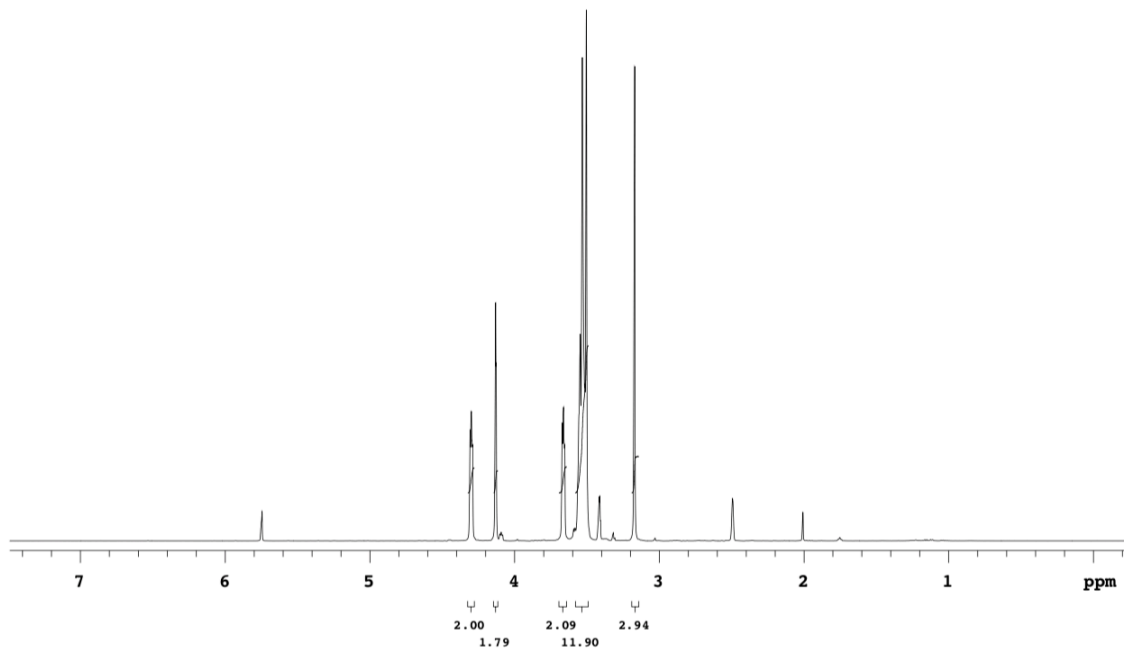
### **3,6,9,12-Tetraoxapentadec-14-yn-1-ol (QX2163)**

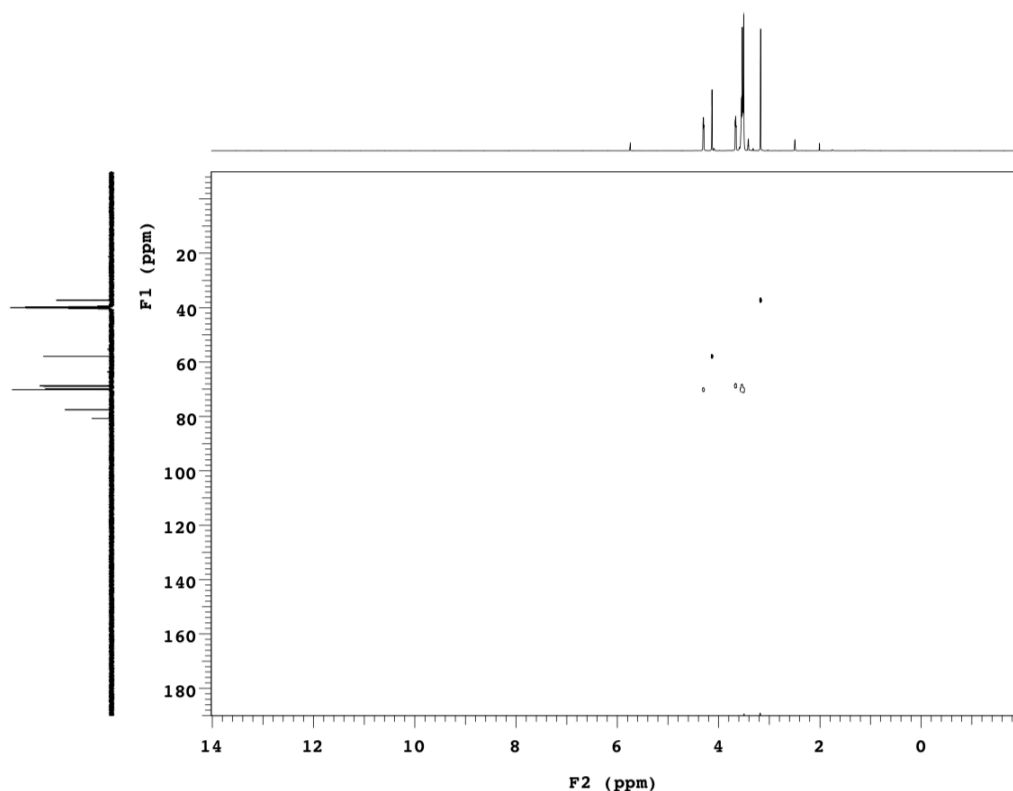


To a solution of tetra ethylene glycol (3.88 g, 20 mmol) in 80 mL of freshly distilled THF was added t-BuOK (1.12 g, 10 mmol) at 0 °C, the mixture was stirred at r.t. for 30 mins, then cooled to 0 °C, and propargyl bromide (1.2 g, 10 mmol) dissolved in 5 mL of THF was added dropwise. After that the reaction mixture was stirred at 0 °C for 10 mins then warmed to r.t. and stirred at r.t. overnight. Then the reaction was filtered, washed with THF, and the filtrate was concentrated and purified by flash chromatography (Hexane/EA : 1/2 to 0/1), the desired intermediate was obtained as a yellow oil. (2.0 g Yield 86%) MS(ESI-TOF)  $m/z$  calc. for  $C_{11}H_{21}O_5^+$  233.14, found 233.14  $[M+H^+]$ ; calc. for  $C_{11}H_{20}NaO_5^+$  255.12 found 255.12  $[M+Na^+]$ .  $^1H$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  4.58 (t,  $J = 5.5$  Hz, 1H), 4.13 (d,  $J = 2.5$  Hz, 2H), 3.55 – 3.52 (m, 4H), 3.52 – 3.50 (m, 8H), 3.47 (t,  $J = 5.5$  Hz, 2H), 3.42 (t,  $J = 2.5$  Hz, 1H), 3.40 (t,  $J = 5.5$  Hz, 2H).  $^{13}C$  NMR (500 MHz, DMSO- $d_6$ )  $\delta$  80.80, 77.57, 72.80, 70.69, 70.27, 70.23, 69.50, 68.97, 60.66, 57.94.

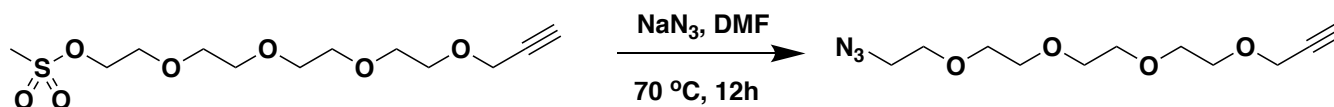




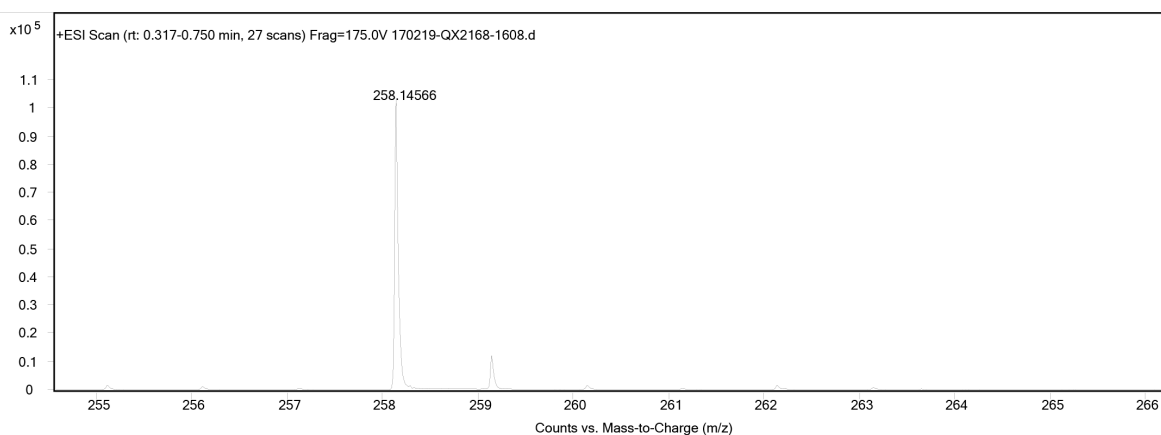


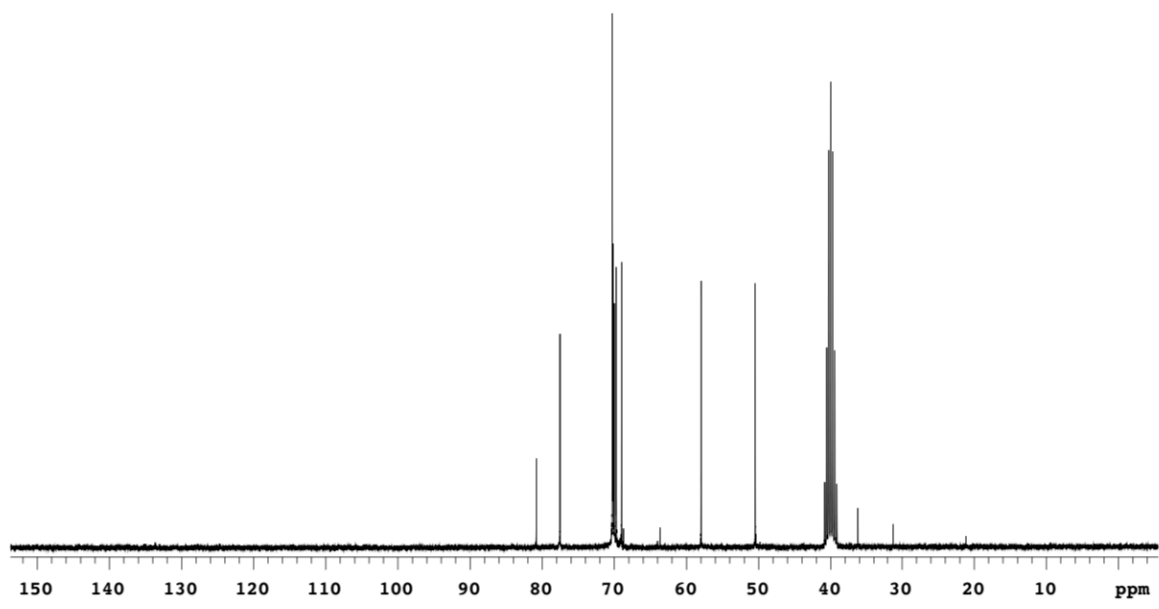
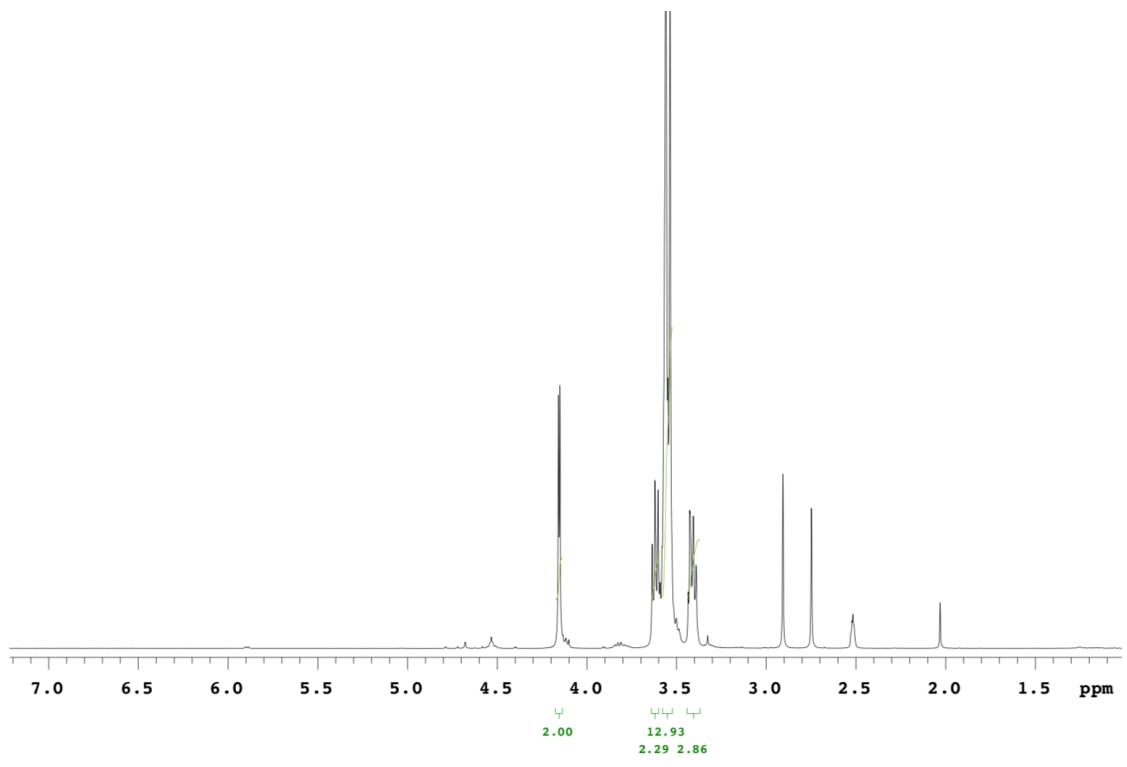


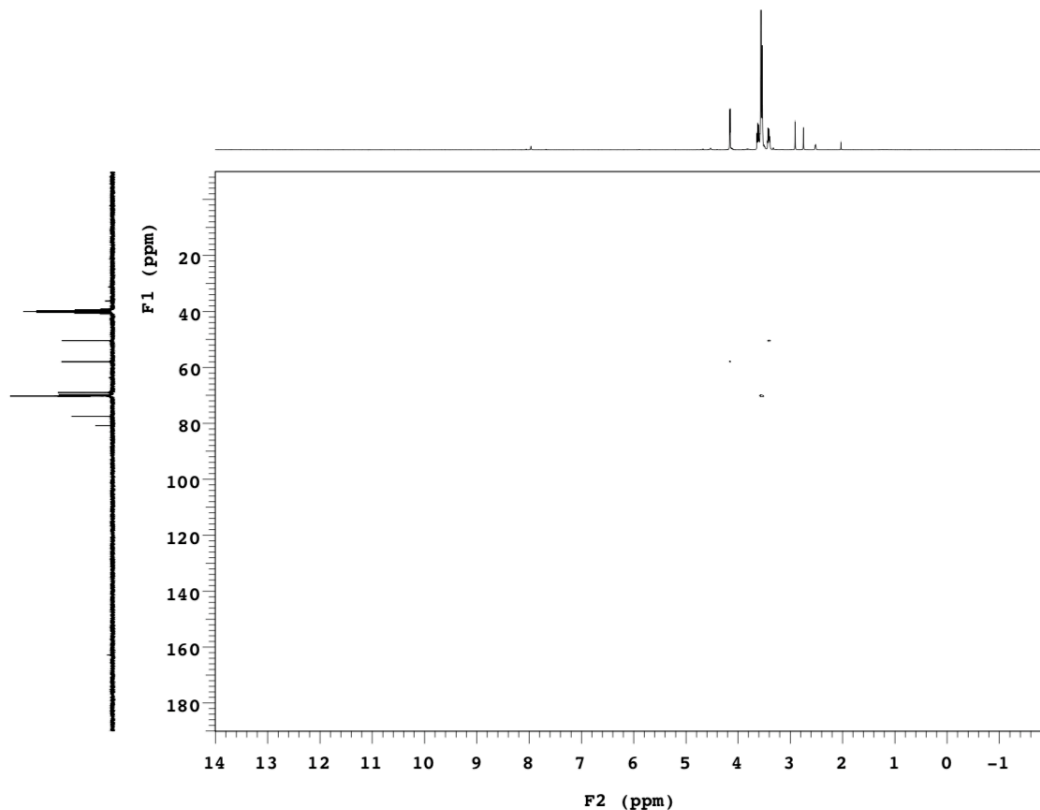
### 1-Azido-3,6,9,12-tetraoxapentadec-14-yne (QX2168)



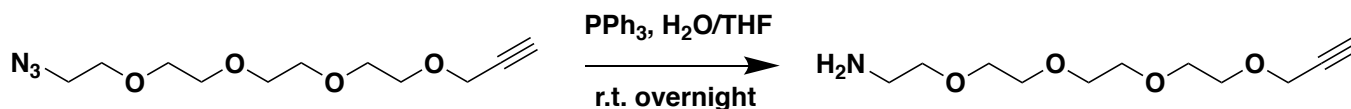
Procedure was performed as described above for compound **QX2158** but using **QX2167** as starting material. Product was obtained as yellow oil (1.68 g Yield: 97%) MS(ESI-TOF)  $m/z$  calc. for  $C_{11}H_{20}N_3O_4^+$  258.14, found 258.15  $[M+H]^+$ .  $^1H$  NMR (300 MHz,  $DMSO-d_6$ ) 4.16 (d,  $J = 2.4$  Hz, 2H), 3.64 – 3.59 (m, 2H), 3.58 – 3.54 (m, 12H), 3.43 – 3.39 (m, 3H).  $^{13}C$  NMR (300 MHz,  $DMSO-d_6$ )  $\delta$  80.80, 77.54, 70.29, 70.24, 70.21, 70.17, 69.99, 69.74, 69.00, 57.97, 50.48.



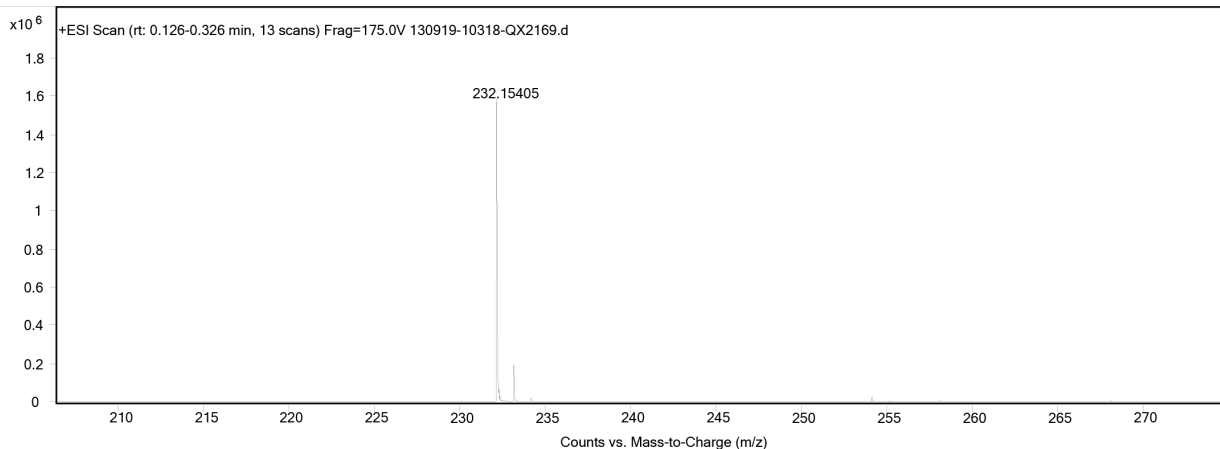


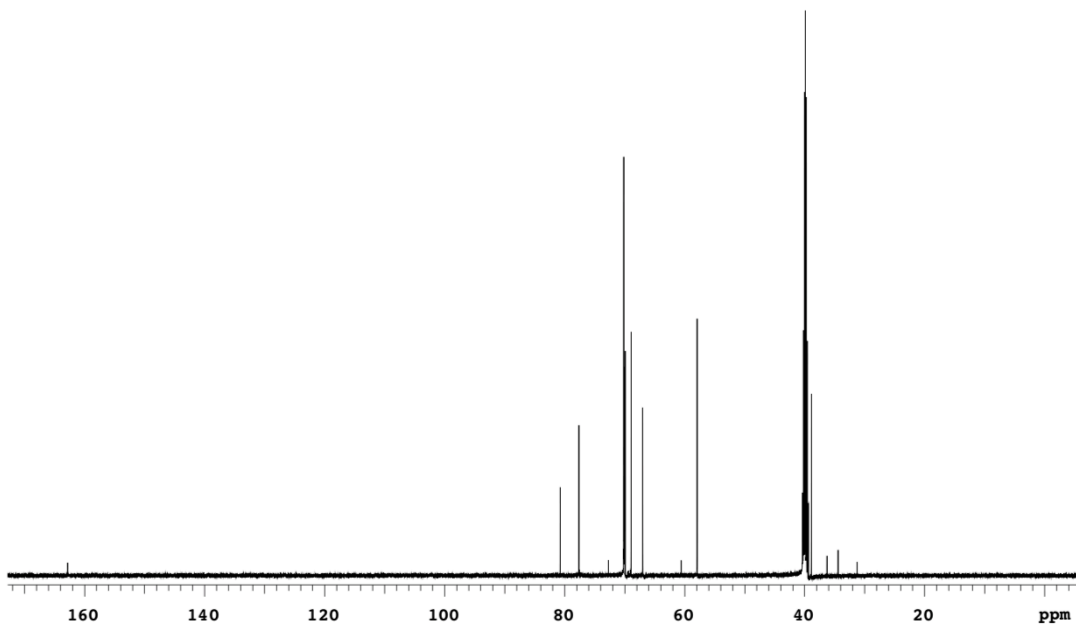
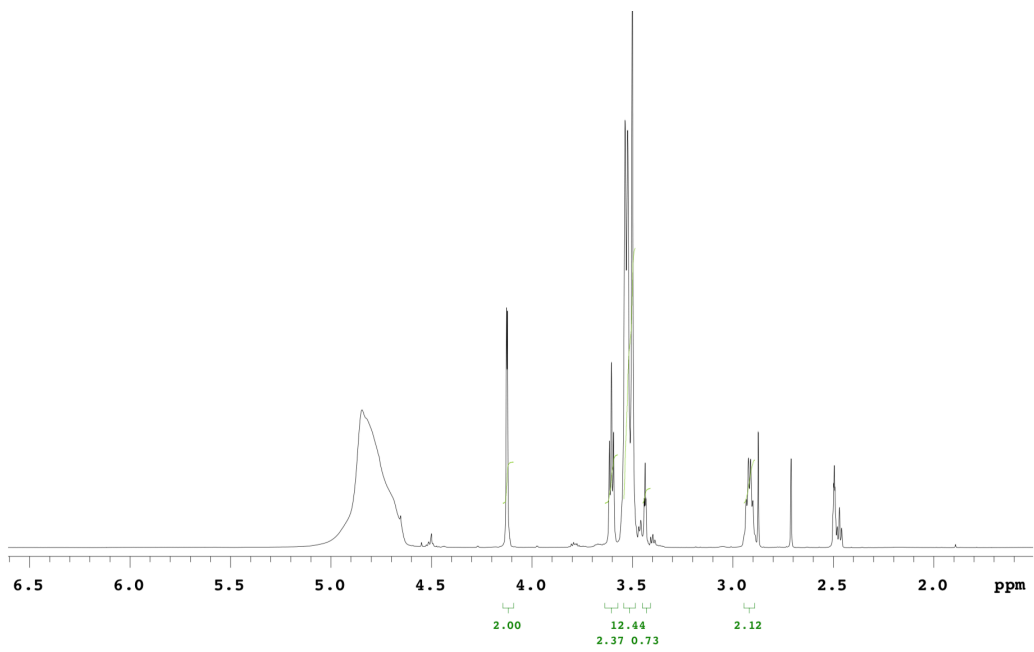


### 3,6,9,12-Tetraoxapentadec-14-yn-1-amine (QX2169)

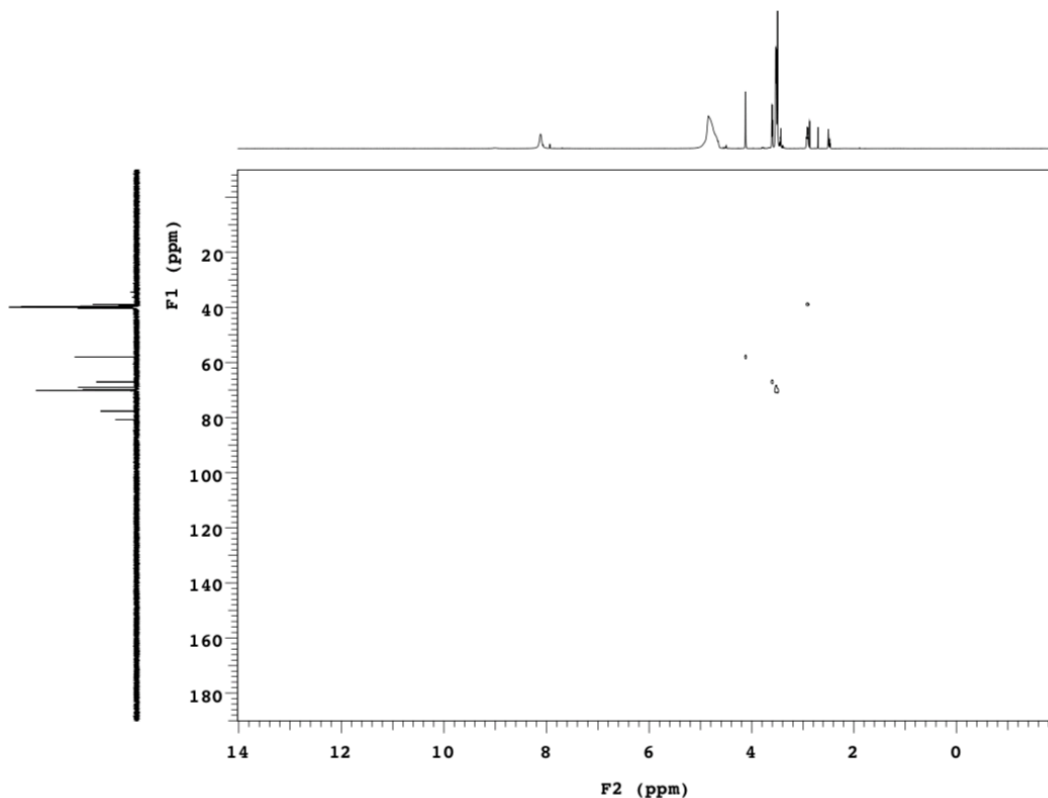


Compound **QX2168** (1.68 g, 6.5 mmol) and PPh<sub>3</sub> (3.4 g, 13.0 mmol) in a mixture of THF (30 mL) and water (1.6 mL) were stirred at room temperature overnight. 1M HCl (60 mL) was added and extracted 3 times with diethyl ether. The aqueous phase was lyophilized to obtain the product as colorless oil without further purification. Yield quantitative; MS(ESI-TOF) *m/z* calc. for C<sub>11</sub>H<sub>22</sub>NO<sub>4</sub>H<sup>+</sup> 232.15, found 232.15 [M+H<sup>+</sup>]. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 4.13 (d, *J* = 2.0 Hz, 2H), 3.60 (t, *J* = 5.5 Hz, 2H), 3.54 – 3.50 (m, 12H), 3.43 (t, *J* = 2.5 Hz, 1H), 2.93 – 2.90 (m, 2H). <sup>13</sup>C NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 80.78, 77.64, 70.17, 70.09, 70.06, 69.91, 68.95, 67.02, 57.94, 38.87.

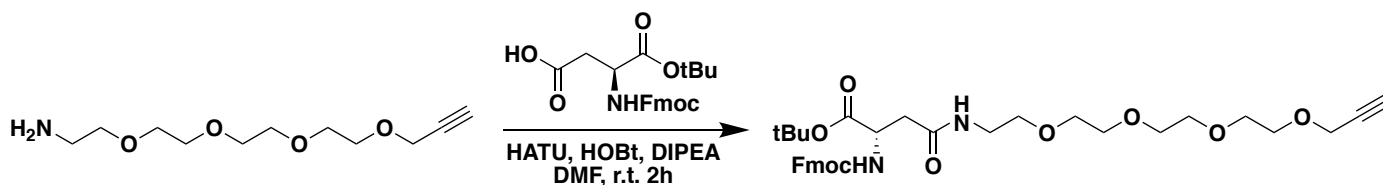




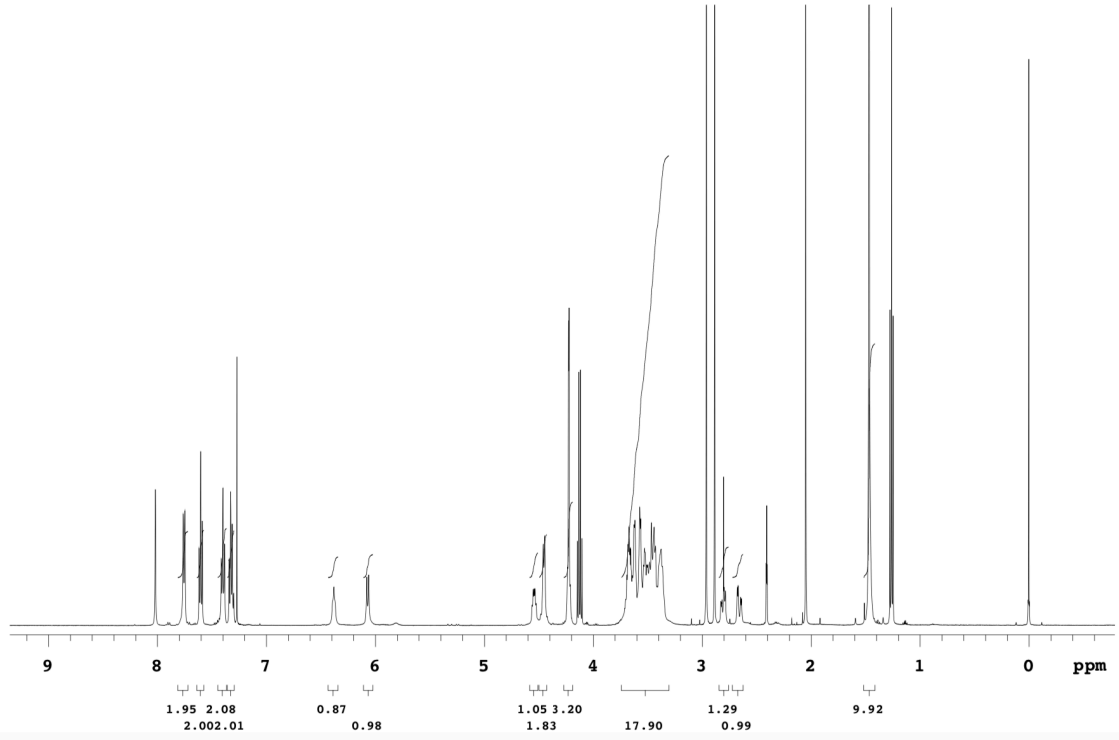
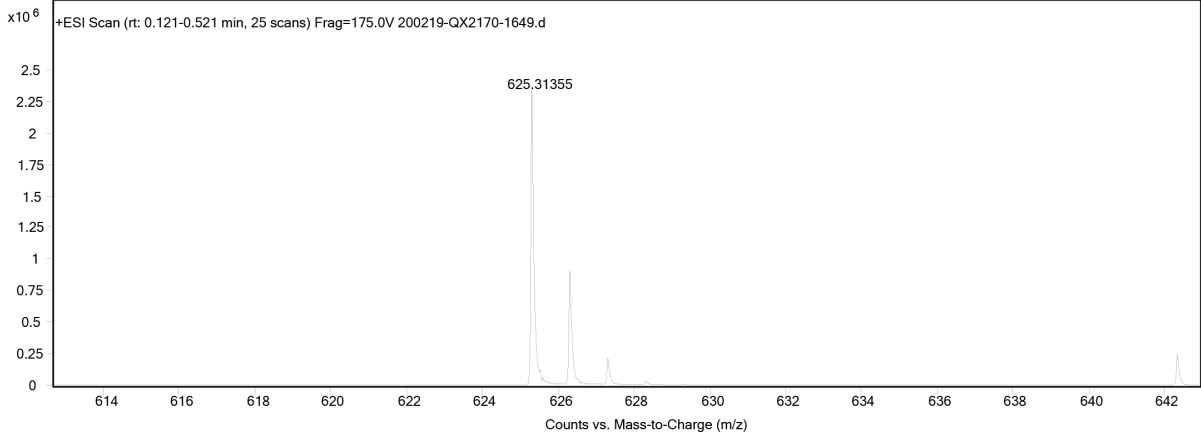


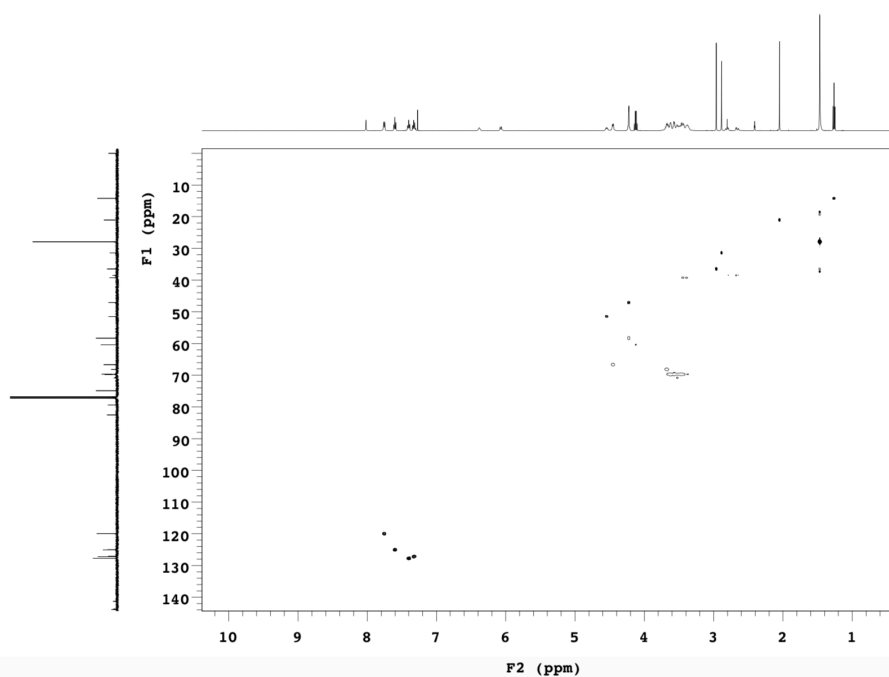
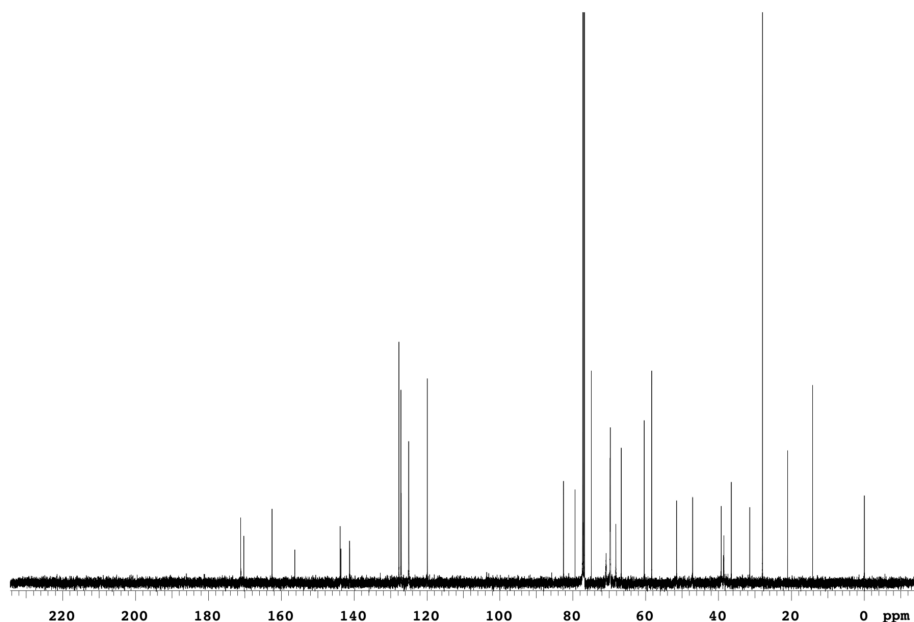


**tert-Butyl (S)-19-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-17-oxo-4,7,10,13-tetraoxa-16-azaicos-1-yn-20-oate (QX2170)**

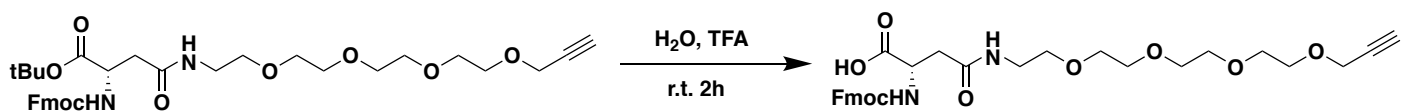


Procedure was performed as described above for compound **QX2161** but using **QX2169** as starting material. Product was obtained as colorless oil (2.4 g Yield: 59%) MS(ESI-TOF)  $m/z$  calc. for  $C_{34}H_{45}N_2O_9^+$  625.31, found 625.31  $[M+H]^+$ .  $^1H$  NMR (500 MHz, Chloroform-*d*)  $\delta$  7.78 (d,  $J = 7.5$  Hz, 2H), 7.60 (dd,  $J = 7.5, 5.5$  Hz, 2H), 7.40 (t,  $J = 7.5$  Hz, 2H), 7.33 (t,  $J = 7.5$  Hz, 2H), 6.38 (br.s., 1H), 6.07 (d,  $J = 8.5$  Hz, 1H), 4.54 (dt,  $J = 8.5, 5.0$  Hz, 1H,  $\alpha$ H), 4.45 (d,  $J = 7.5$  Hz, 2H), 4.24 – 4.21 (m, 3H, 1H of Fmoc, 2H of propargyl group), 3.71 – 3.37 (m, 16H), 2.80 (dd,  $J = 15.5, 5.0$  Hz, 1H), 2.66 (dd,  $J = 15.5, 5.0$  Hz, 1H), 1.49 (s, 9H).  $^{13}C$  NMR (500 MHz, Chloroform-*d*)  $\delta$  171.17, 170.30, 156.31, 143.87, 143.68, 141.31, 141.23, 127.76, 127.20, 127.11, 125.08, 125.02, 119.99, 82.53, 79.38, 74.89, 69.76, 69.67, 68.17, 66.69, 58.33, 51.49, 47.11, 39.27, 38.53, 36.50, 31.43, 27.91.



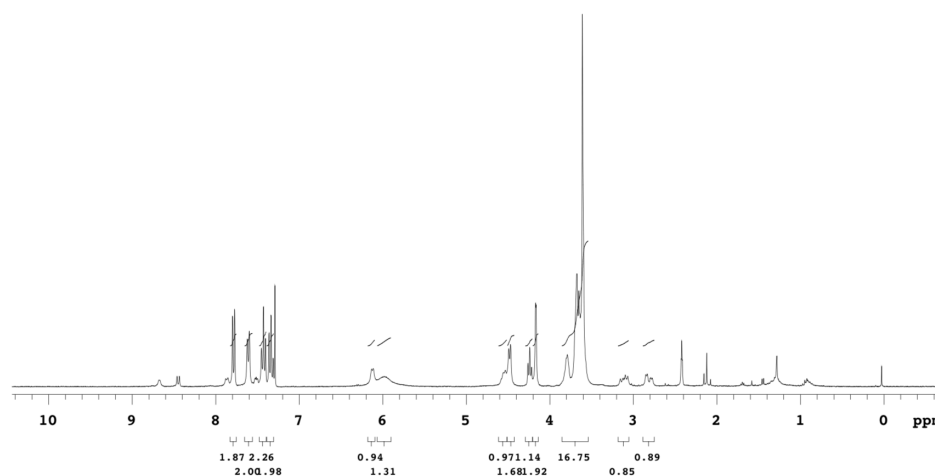
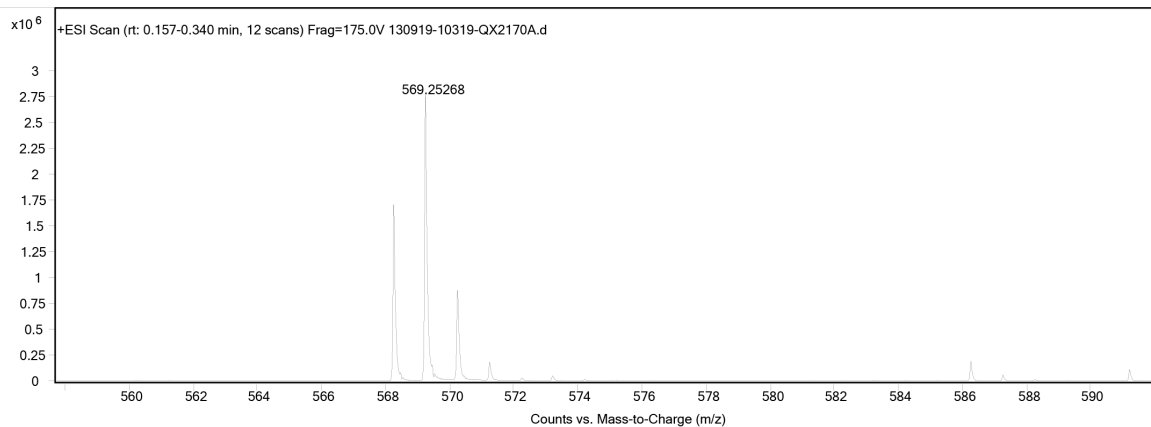


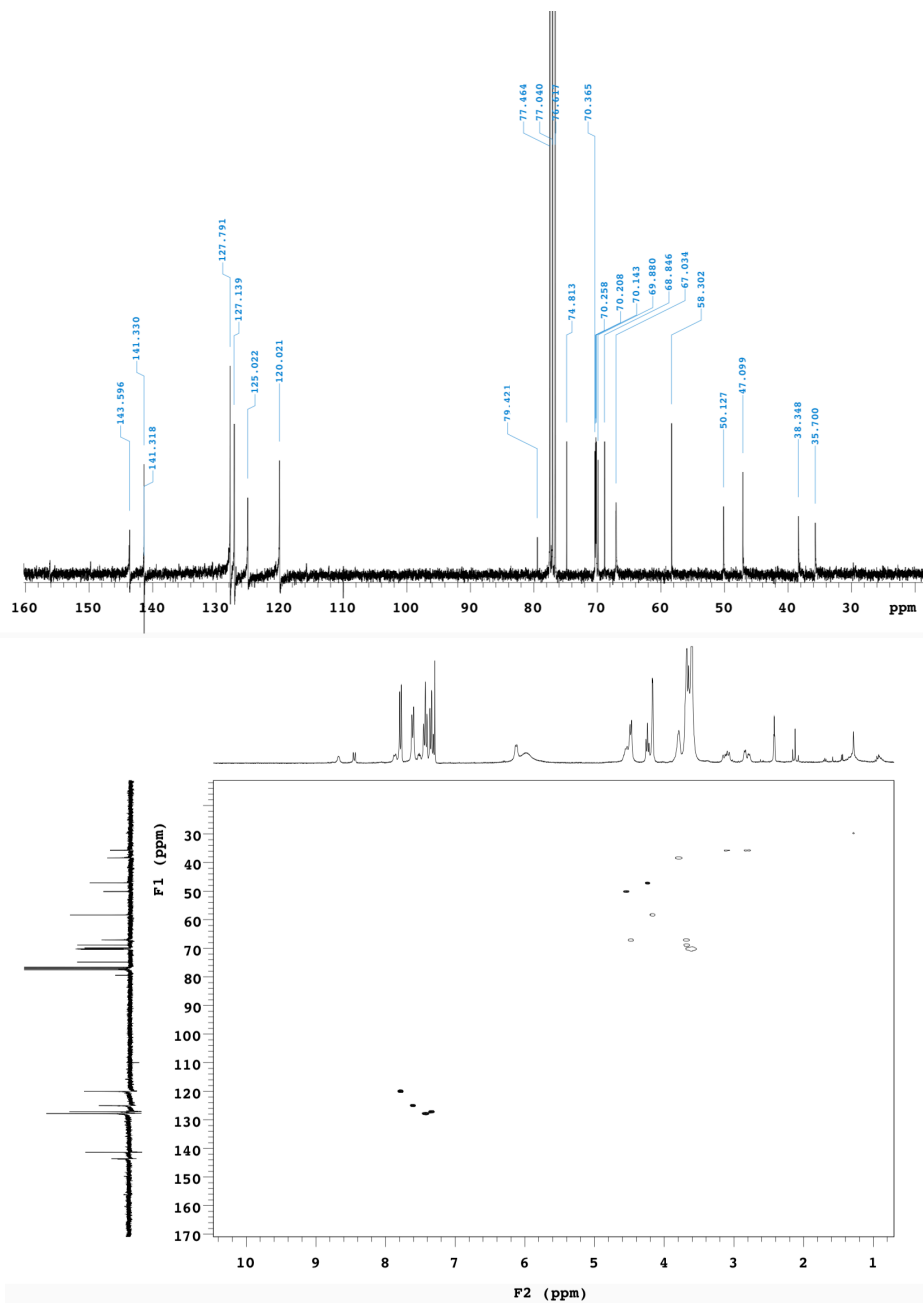
**(S)-19-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-17-oxo-4,7,10,13-tetraoxa-16-azaicos-1-yn-20-oic acid (QX2170A)**



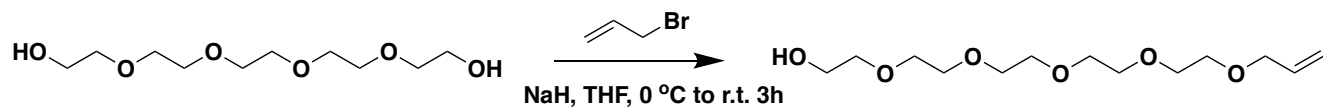
Procedure was performed as described above for compound **QX2162** but using **QX2170** as starting material. Product was obtained as yellow oil. Yield quantitative. MS(ESI-TOF)  $m/z$  calc. for  $C_{30}H_{37}N_2O_9^+$  569.25, found 569.25  $[M+H^+]$ .  $^1H$  NMR (300 MHz, Chloroform- $d$ )  $\delta$  7.79 (d,  $J = 7.5$  Hz, 2H), 7.61 (d,  $J = 7.2$  Hz, 2H), 7.43 (t,  $J = 7.2$  Hz, 2H), 7.34 (t,  $J = 7.5$  Hz, 2H), 6.12 (d,  $J = 6$  Hz, 1H), 5.98 (br.s., 1H), 4.58 – 4.50 (m, 1H), 4.48 (d,  $J$

= 6.6 Hz, 2H), 4.24 (t,  $J = 6.6$  Hz, 1H, from Fmoc), 4.16 (d,  $J = 1.8$  Hz, 2H), 3.79 – 3.61 (m, 16H), 3.11 (dd,  $J = 17.7, 9$  Hz, 1H), 2.82 (dd,  $J = 17.4, 5.1$  Hz, 1H).  $^{13}\text{C}$  NMR (300 MHz, Chloroform- $d$ )  $\delta$  143.61, 143.35, 141.33, 127.81, 127.15, 125.04, 120.03, 79.43, 74.82, 70.37, 70.26, 70.21, 70.14, 69.88, 69.85, 67.03, 58.30, 50.13, 47.10, 38.35, 35.70.



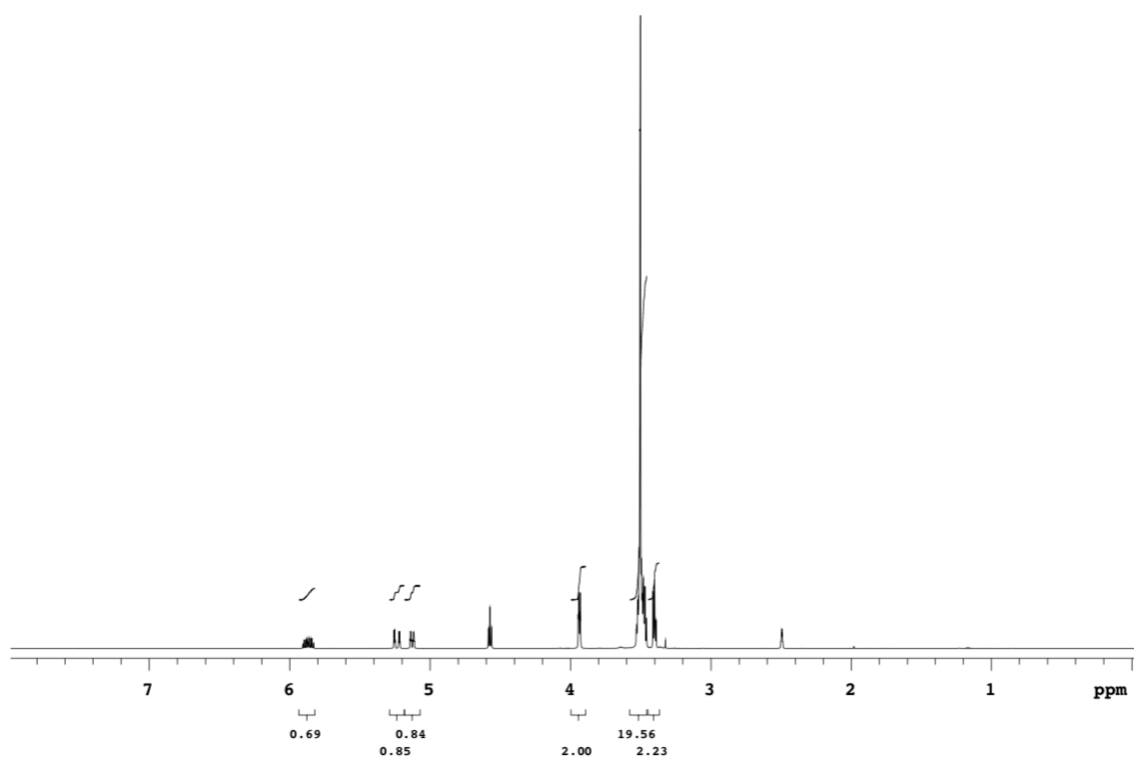
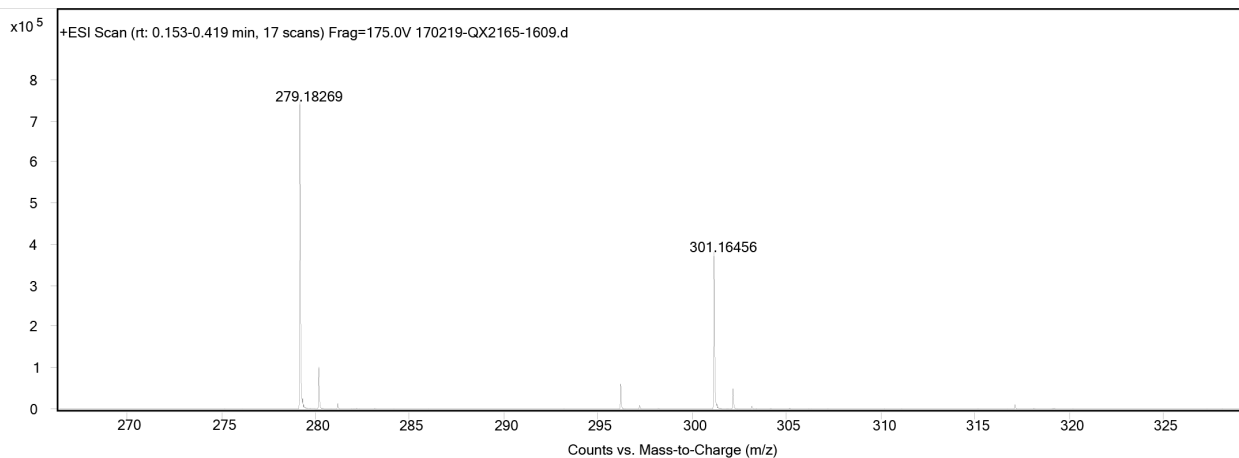


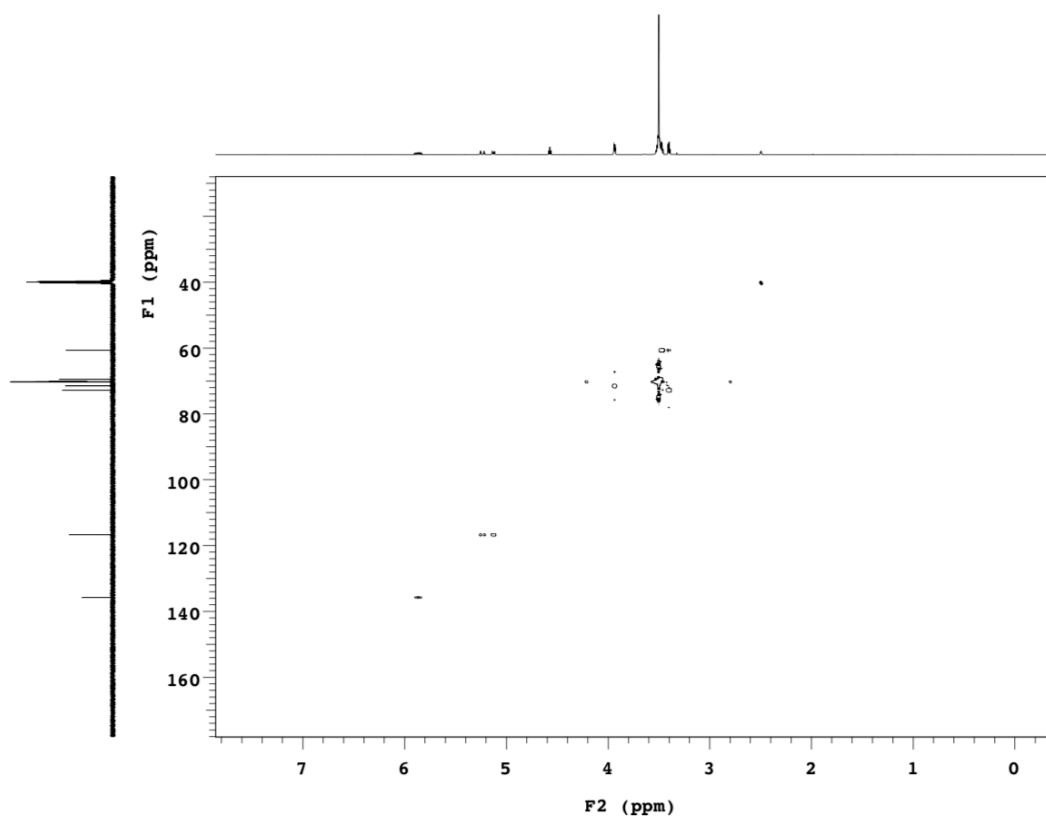
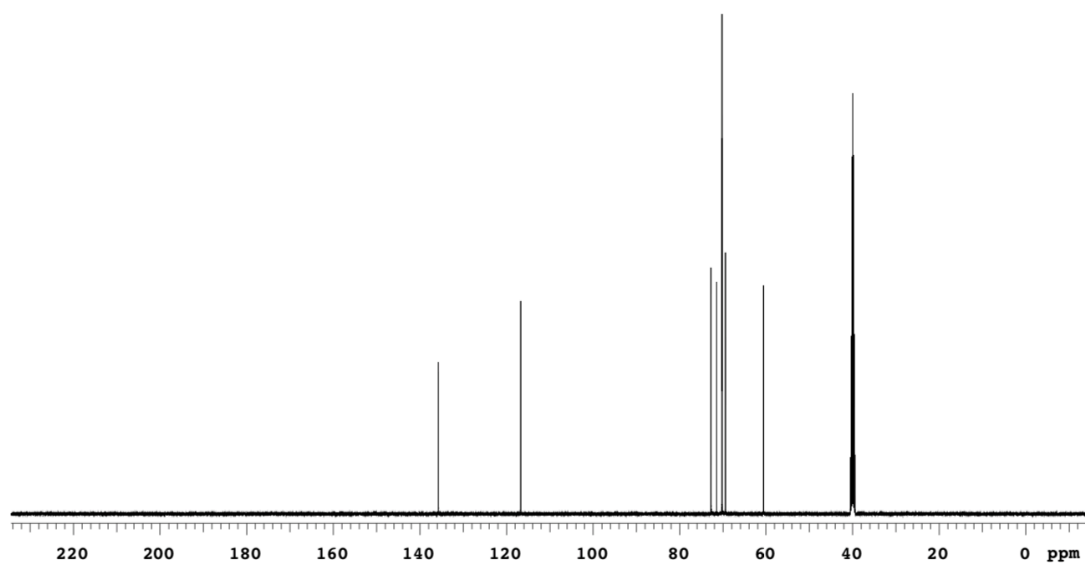
### 3,6,9,12,15-Pentaoxaoctadec-17-en-1-ol (QX2165)



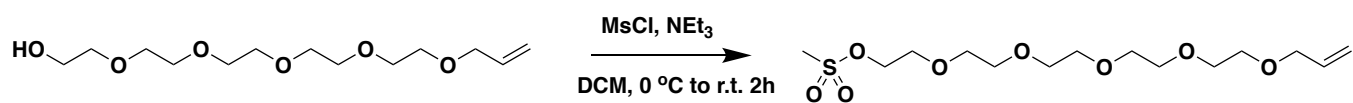
To NaH (60 % in mineral oil, 360 mg, 9 mmol, washed with dry hexane) in THF (30 mL) was added pentaerythritol (4.76 g, 20 mmol) at 0°C under argon. The reaction mixture was stirred at 0°C for 1 hour, allyl bromide (846  $\mu\text{L}$ , 10 mmol) in THF (10 mL) was added slowly at 0°C and stirring was continued for another 2 hours at room temperature. The reaction mixture was poured into cold saturated ammonium chloride solution (50 mL) and extracted with ethyl acetate. The combined organic phases were washed with saturated brine and dried over anhydrous sodium sulfate, evaporated to dryness. The products were separated by flash column chromatography to obtain the desired product QX2165 as yellow oil. (1.2 g, Yield 43%); MS(ESI-TOF)  $m/z$  calc. for  $\text{C}_{13}\text{H}_{27}\text{O}_6^+$

279.18, found 279.18 [M+NH<sub>4</sub><sup>+</sup>]; calc. for C<sub>13</sub>H<sub>26</sub>NaO<sub>6</sub><sup>+</sup> 301.16, found 301.16 [M+Na<sup>+</sup>]. <sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 5.91 – 5.83 (m, 1H), 5.24 (dd, *J* = 17.5, 2.0 Hz, 1H), 5.13 (dd, *J* = 10.5, 1.5 Hz, 1H), 3.94 (dt, *J* = 5.5, 1.5 Hz, 2H), 3.53 – 3.46 (m, 18H), 3.40 (t, *J* = 5.0 Hz, 2H). <sup>13</sup>C NMR (500 MHz, DMSO-*d*<sub>6</sub>) δ 135.74, 116.74, 72.80, 71.50, 70.28, 70.25, 70.23, 70.22, 69.46, 60.66.



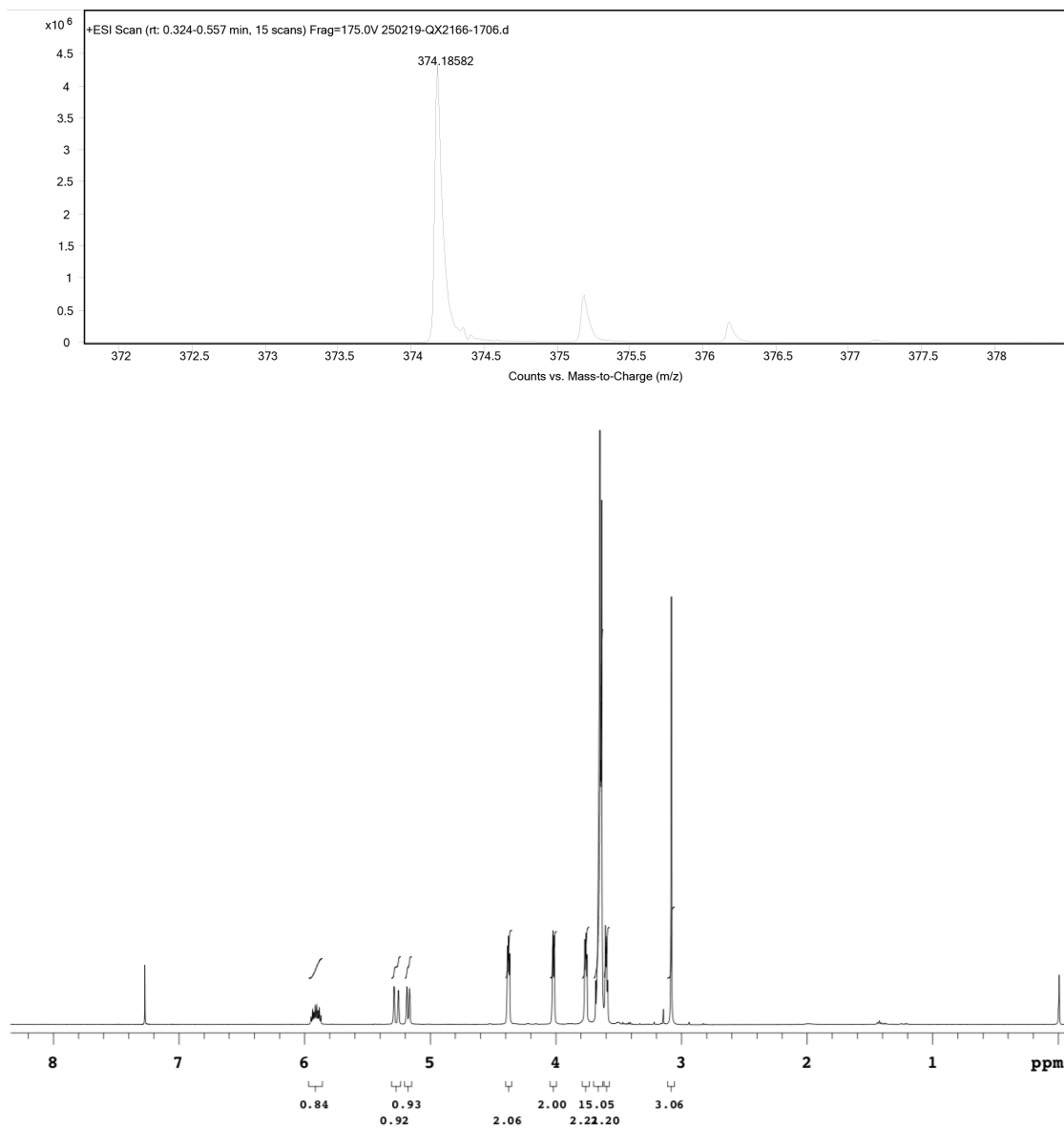


### 3,6,9,12,15-Pentaoxaoctadec-17-en-1-yl methanesulfonate (QX2166)

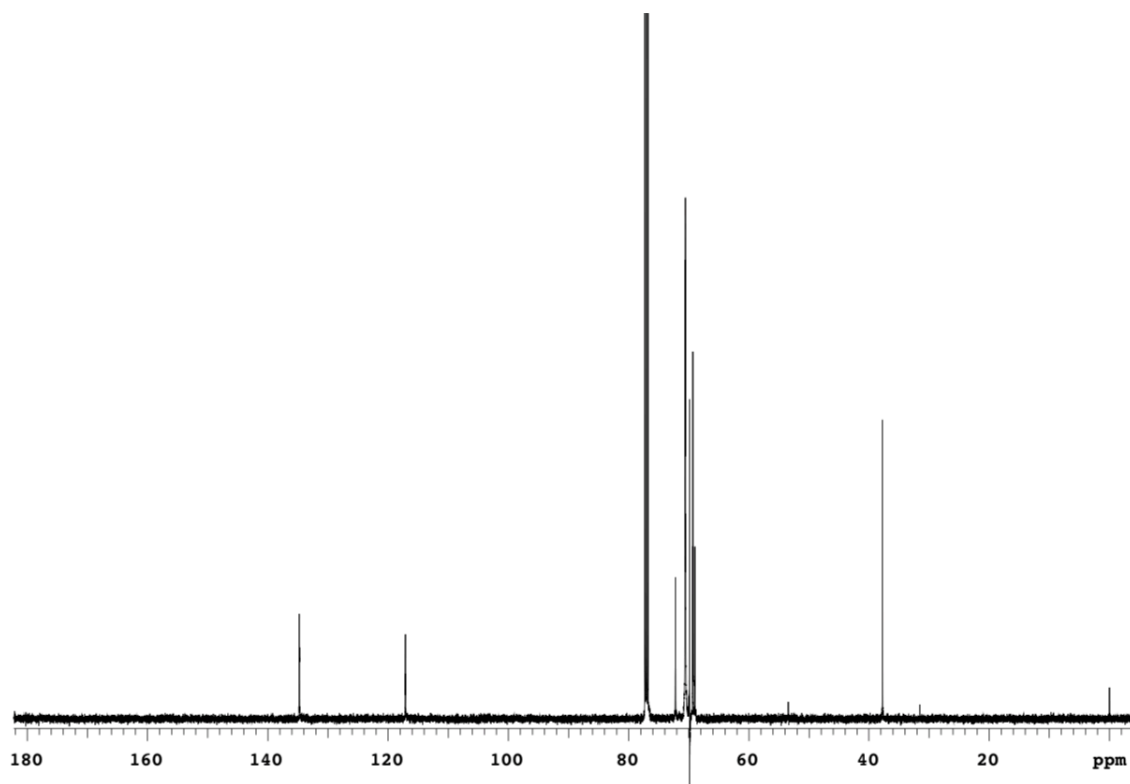


Procedure was performed as described above for compound **QX2157** but using **QX2165** as starting material. Product was obtained as yellow oil (1.45 g Yield: 95%) MS(ESI-TOF) *m/z* calc. for C<sub>14</sub>H<sub>32</sub>NO<sub>8</sub>S<sup>+</sup> 374.18, found 374.19 [M+NH<sub>4</sub><sup>+</sup>]. <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 5.90 (ddt, *J* = 17.5, 10.5, 5.5 Hz, 1H), 5.27 (dd, *J* =

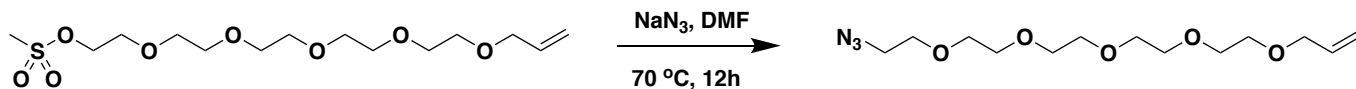
17.5, 1.5 Hz, 1H), 5.18 (d,  $J = 10.5$  Hz, 1H), 4.39 – 4.37 (m, 2H), 4.02 (d,  $J = 6.0$  Hz, 2H), 3.77 – 3.75 (m, 2H), 3.65 – 3.64 (m, 14H), 3.61 – 3.59 (m, 2H), 3.08 (s, 3H).  $^{13}\text{C}$  NMR (500 MHz, Chloroform- $d$ )  $\delta$  134.75, 117.10, 72.23, 70.63, 70.58, 70.53, 70.52, 69.89, 69.87, 69.41, 69.34, 69.01, 37.74.



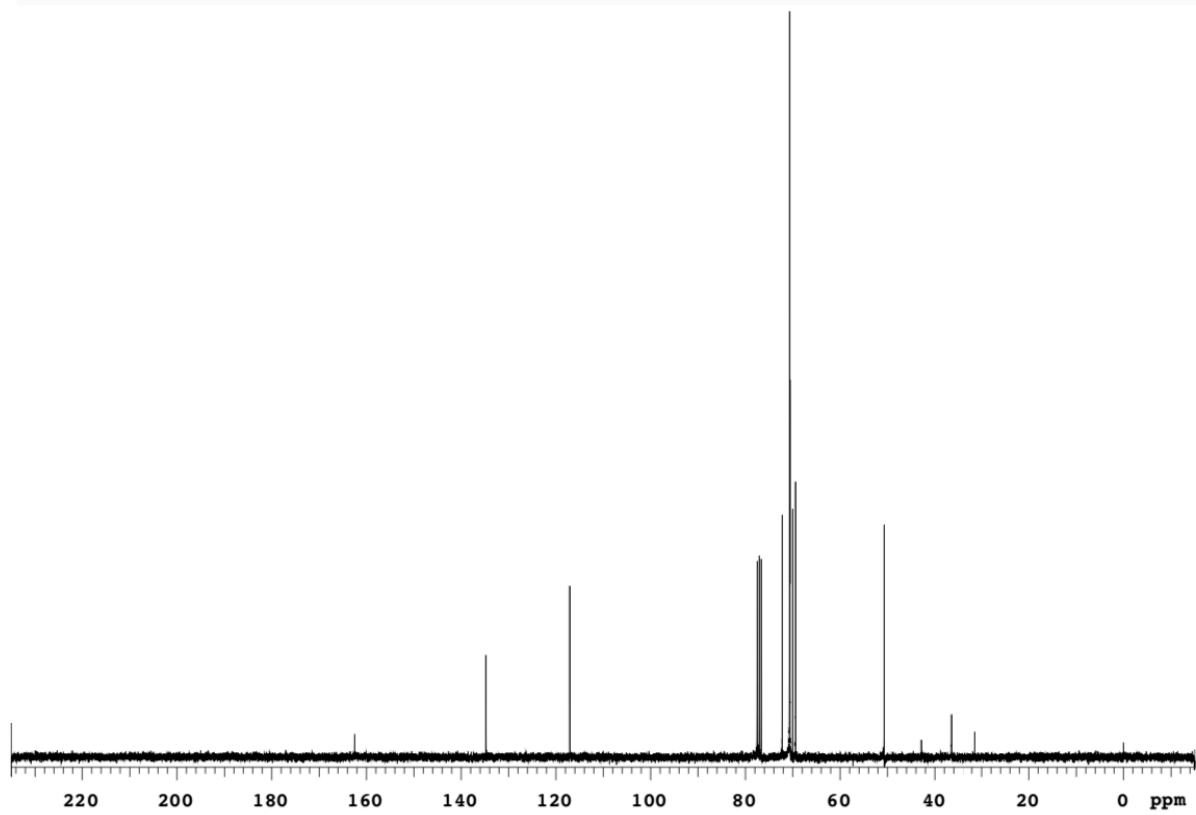
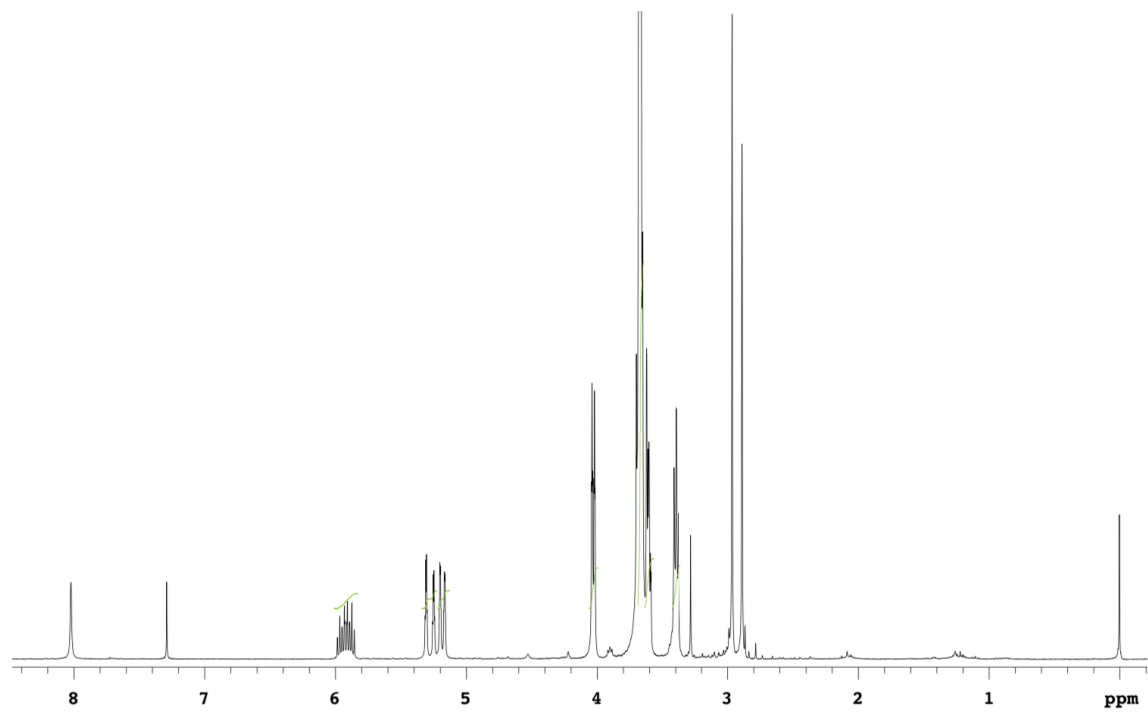




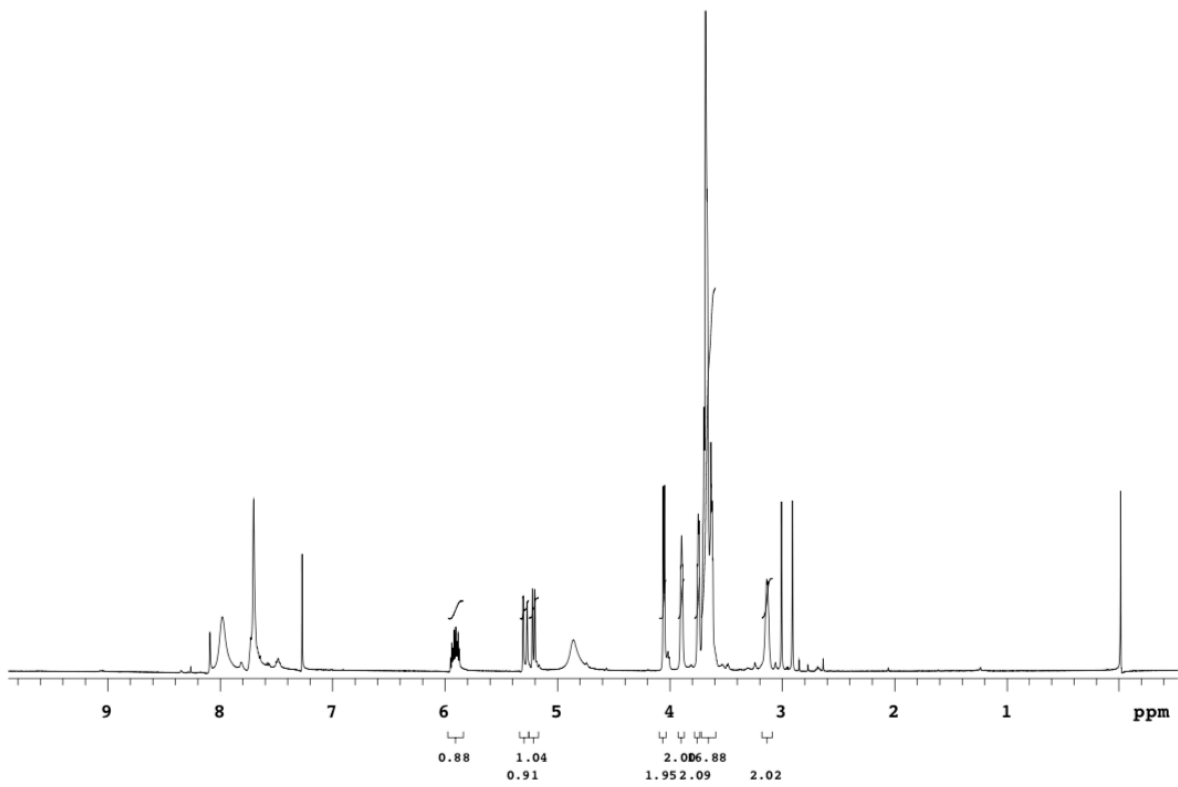
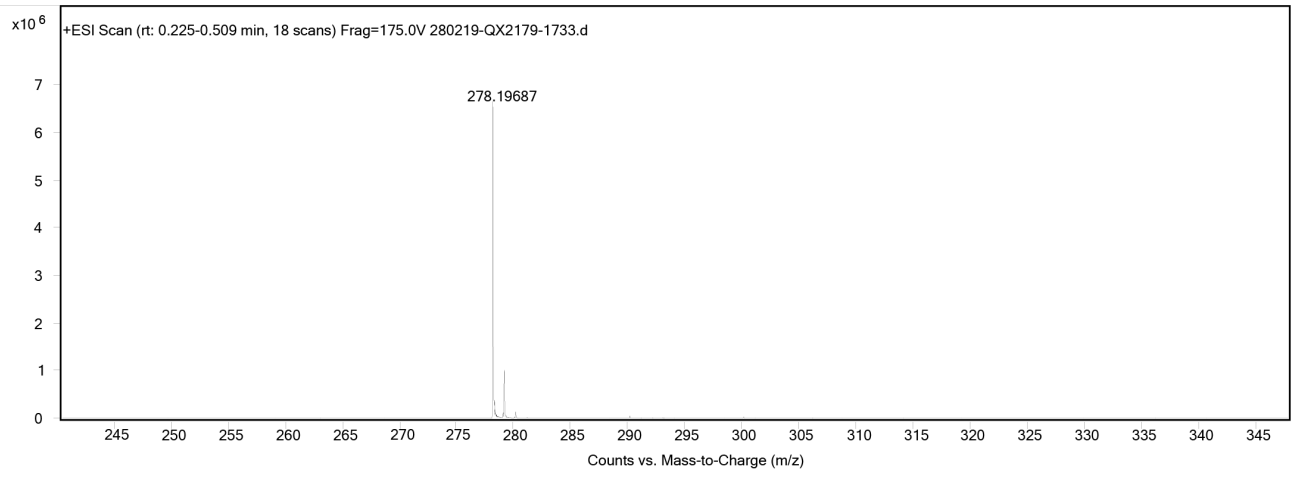
**1-Azido-3,6,9,12,15-pentaoxaoctadec-17-ene(QX2178)**

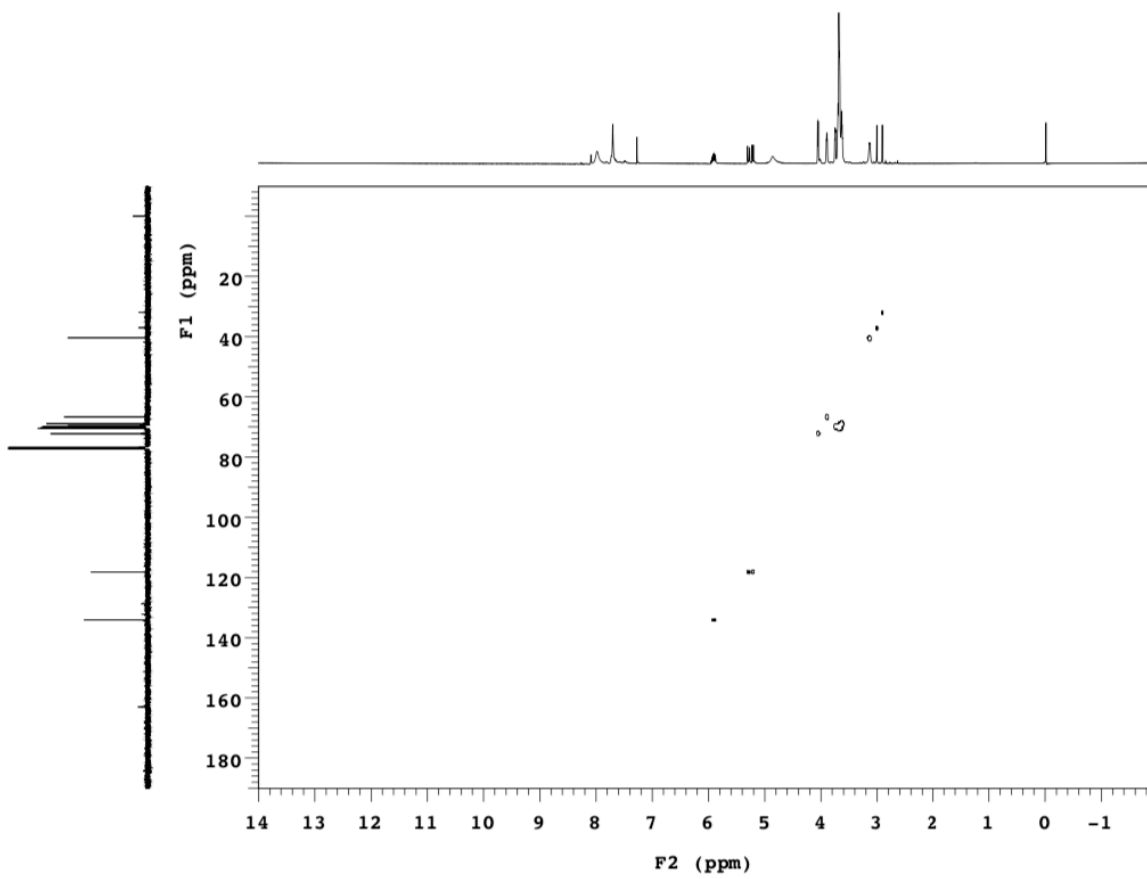
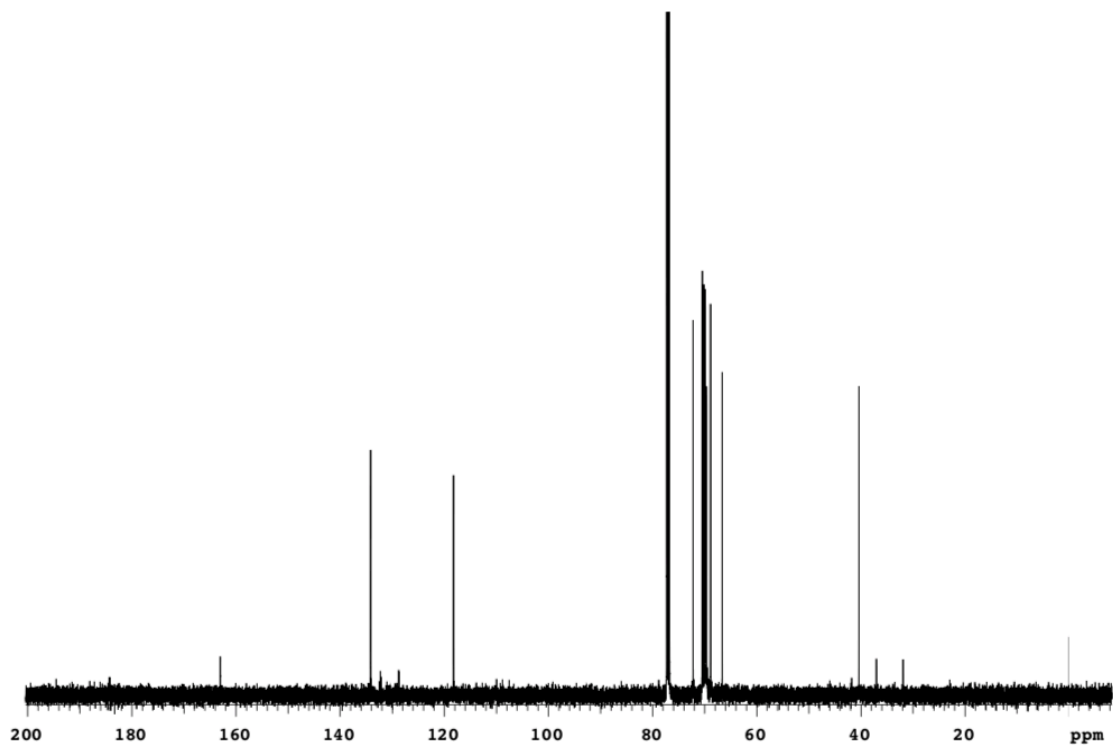


Procedure was performed as described above for compound **QX2158** but using **QX2166** as starting material. Product was obtained as colorless oil. Yield quantitative. (The mass spectrum information was not detected since it is hard to be charged.)  $^1\text{H}$  NMR (300 MHz, Chloroform-*d*)  $\delta$  5.92 (ddt,  $J = 17.5, 10.5, 5.5$  Hz, 1H), 5.27 (dd,  $J = 17.4, 1.8$  Hz, 1H), 5.18 (dd,  $J = 10.5, 1.5$  Hz, 1H), 4.03 (d,  $J = 5.4$  Hz, 2H), 3.70 – 3.65 (m, 16H), 3.62 – 3.60 (m, 2H), 3.40 (t,  $J = 5.1$  Hz, 2H).  $^{13}\text{C}$  NMR (300 MHz, Chloroform-*d*)  $\delta$  134.78, 117.10, 72.24, 70.70, 70.68, 70.63, 70.64, 70.61, 70.04, 69.43, 50.69.

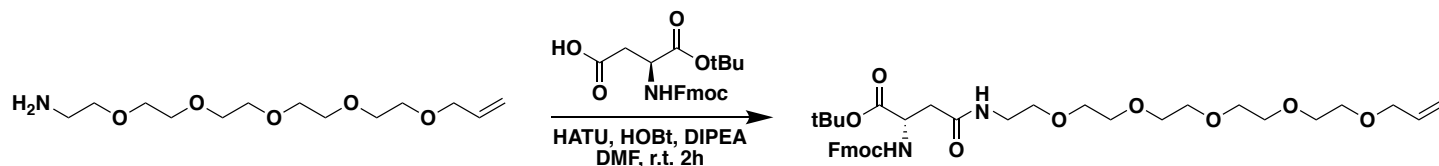




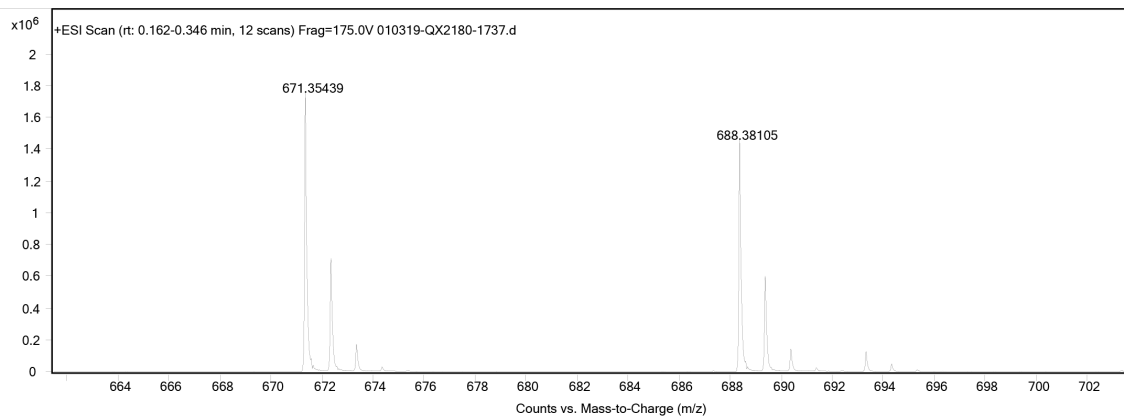


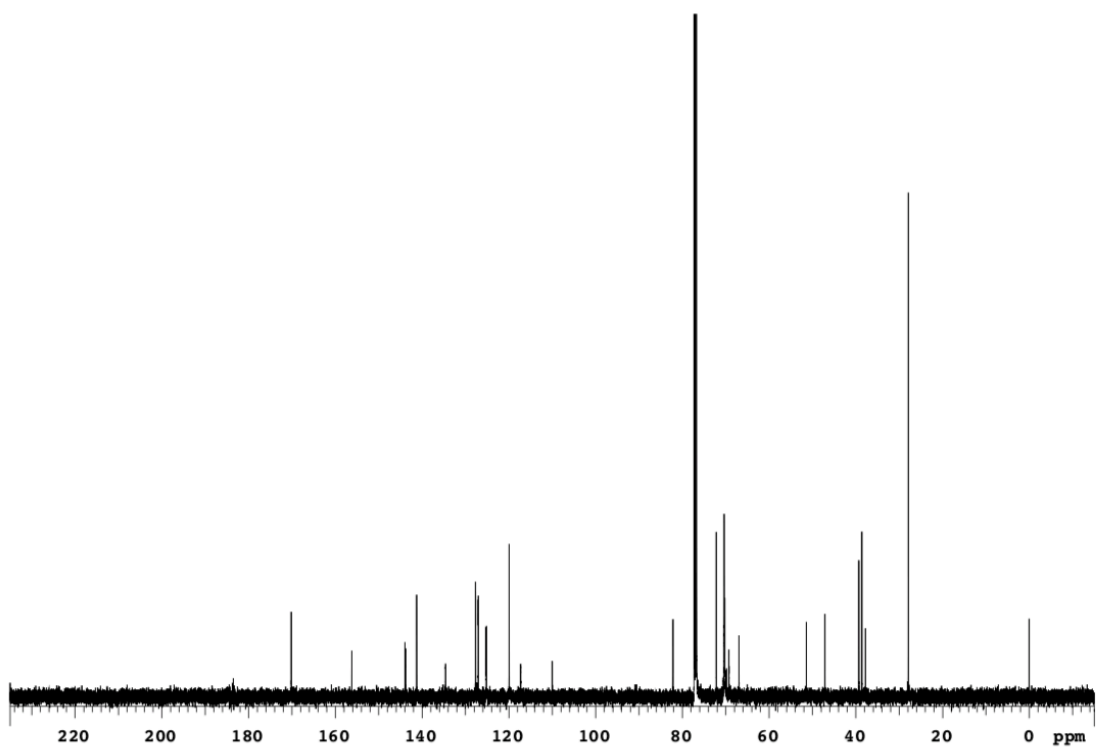
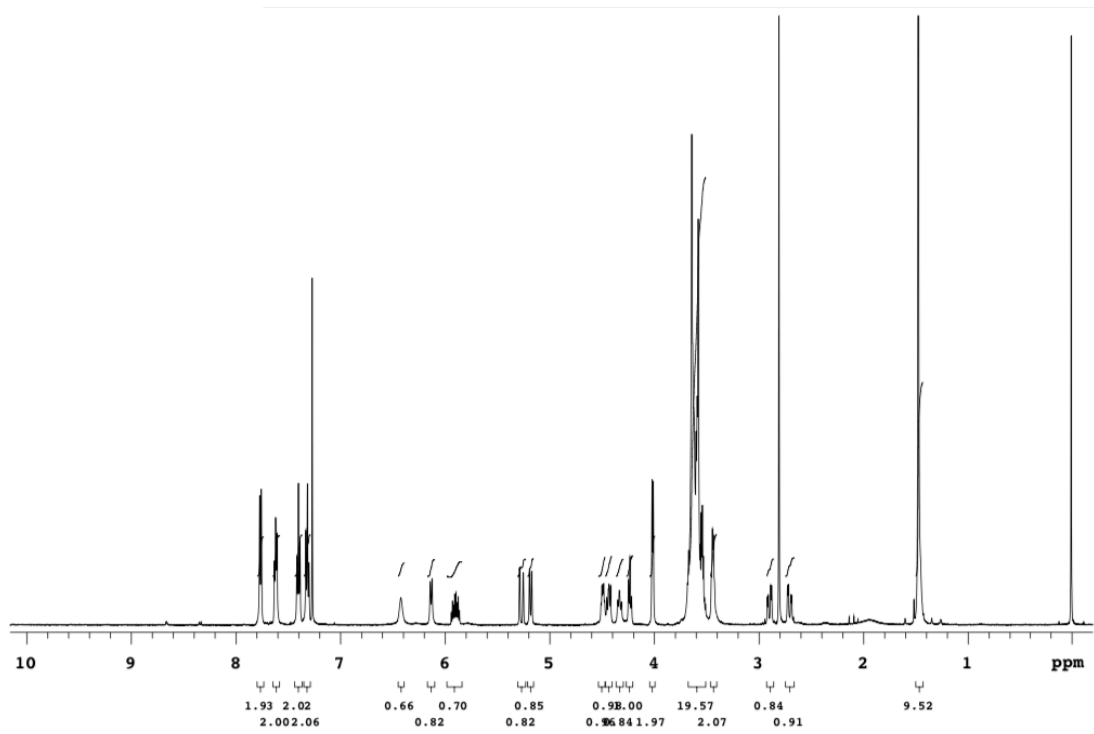


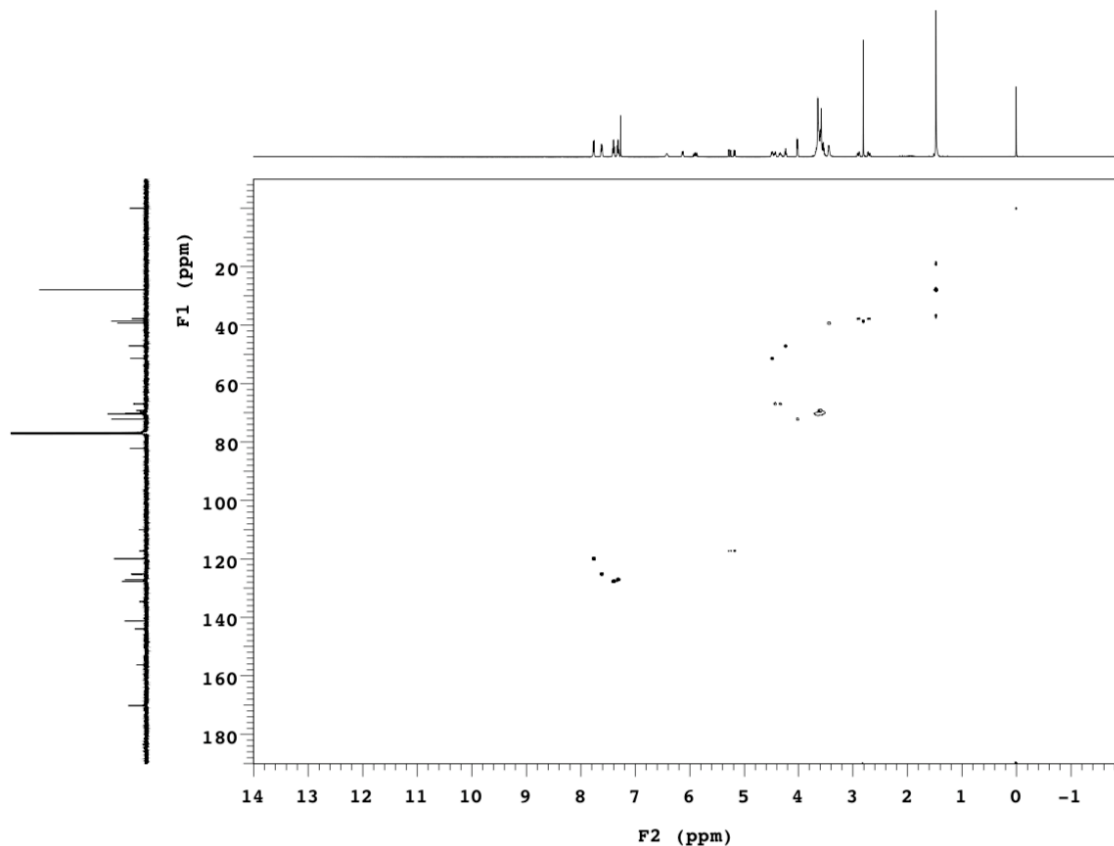
**tert-Butyl (S)-22-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-20-oxo-4,7,10,13,16-pentaoxa-19-azatricos-1-en-23-oate (QX2180)**



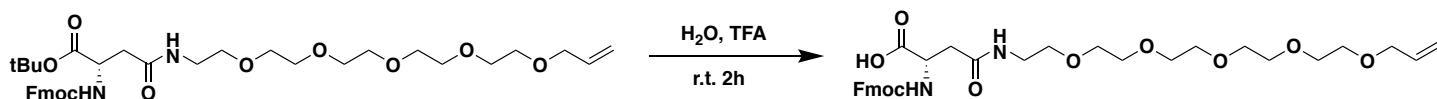
Procedure was performed as described above for compound **QX2161** but using **QX2179** as starting material. Product was obtained as colorless oil. (0.94 g Yield: 41%) MS(ESI-TOF)  $m/z$  calc. for  $C_{36}H_{51}N_2O_{10}^+$  671.35, found 671.35  $[M+H^+]$ ; calc. for  $C_{36}H_{54}N_3O_{10}^+$  688.38, found 688.38  $[M+NH_4^+]$ .  $^1H$  NMR (500 MHz, Chloroform-*d*)  $\delta$  7.77 (d,  $J = 7.5$  Hz, 2H), 7.62 (dd,  $J = 7.5, 5.5$  Hz, 2H), 7.40 (t,  $J = 7.5$  Hz, 2H), 7.32 (t,  $J = 7.5$  Hz, 2H), 6.43 (br.s., 1H), 6.14 (d,  $J = 8.5$  Hz, 1H), 5.89 (ddt,  $J = 17.5, 10.5, 5.5$  Hz, 1H), 5.27 (dd,  $J = 17.5, 1.5$  Hz, 1H), 5.18 (d,  $J = 10.5$  Hz, 1H), 4.49 (dt,  $J = 8.5, 4.5$  Hz, 1H), 4.44 (dd,  $J = 10.5, 7.5$  Hz, 1H), 4.34 (dd,  $J = 10.5, 7.5$  Hz, 1H), 4.23 (t,  $J = 7.5$  Hz, 1H), 4.02 (d,  $J = 5.5$  Hz, 2H), 3.68 – 3.51 (m, 18H), 3.44 (dt,  $J = 4.5, 5.0$  Hz, 2H), 2.90 (dd,  $J = 15.5, 5.0$  Hz, 1H), 2.71 (dd,  $J = 15.5, 5.0$  Hz, 1H), 1.47 (s, 9H).  $^{13}C$  NMR (500 MHz, Chloroform-*d*)  $\delta$  170.19, 156.24, 143.98, 143.83, 141.27, 134.62, 127.68, 127.10, 127.08, 125.25, 125.18, 119.94, 117.28, 110.00, 82.16, 72.21, 70.38, 70.21, 69.86, 69.27, 66.96, 51.42, 47.13, 39.32, 38.62, 37.82, 27.93.





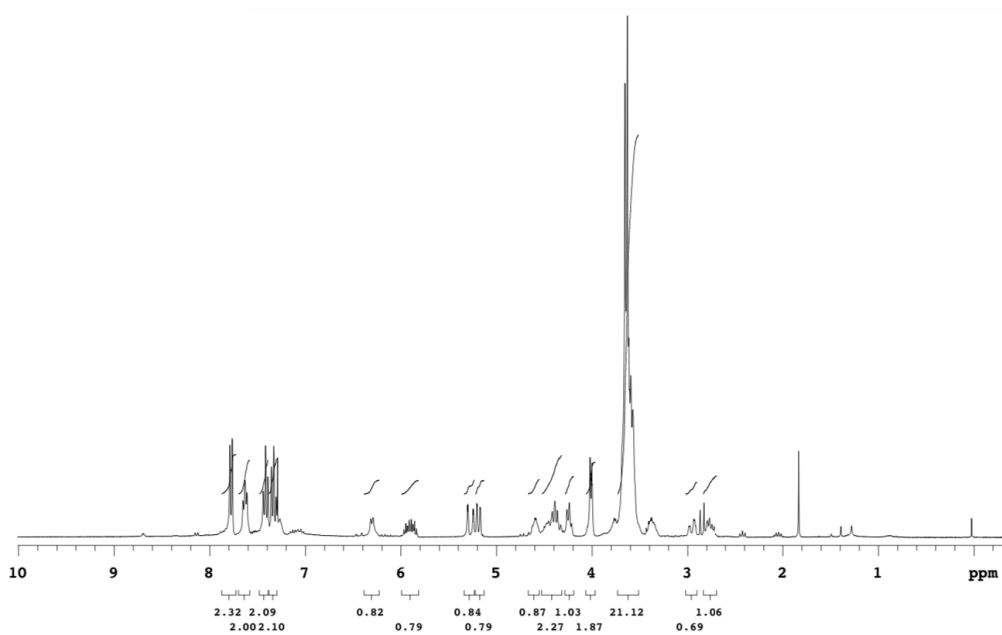
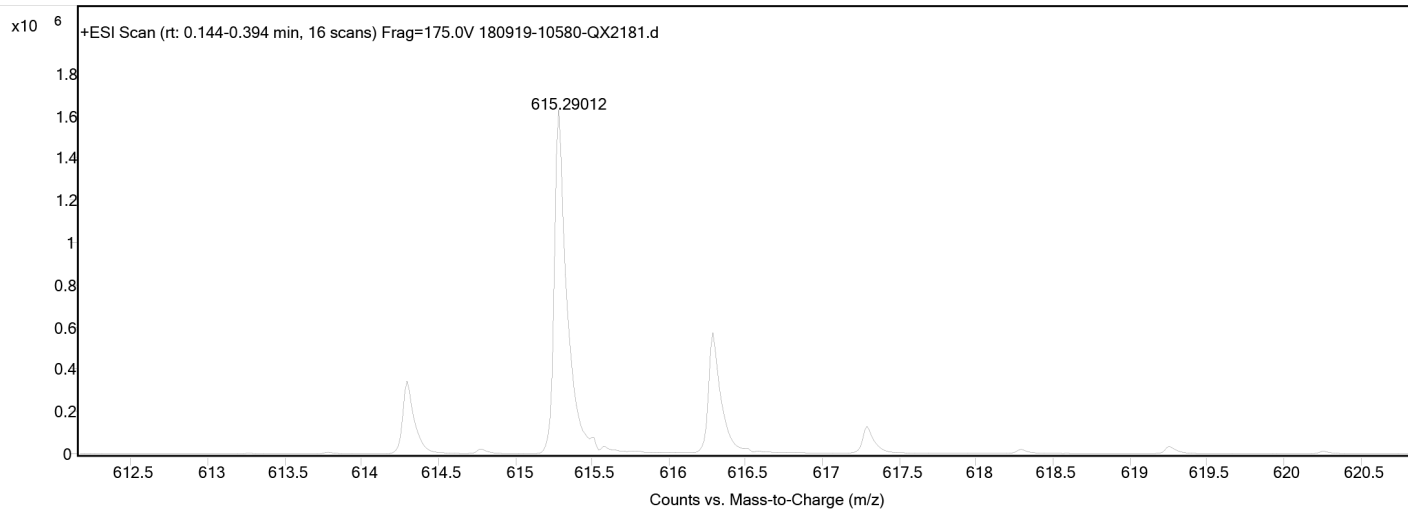


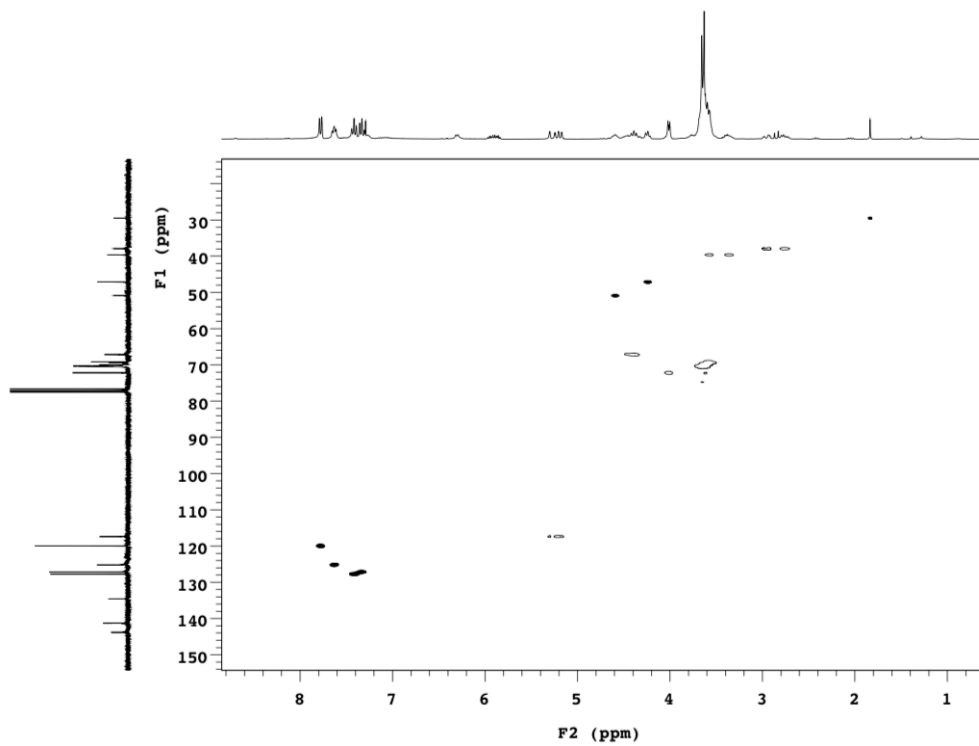
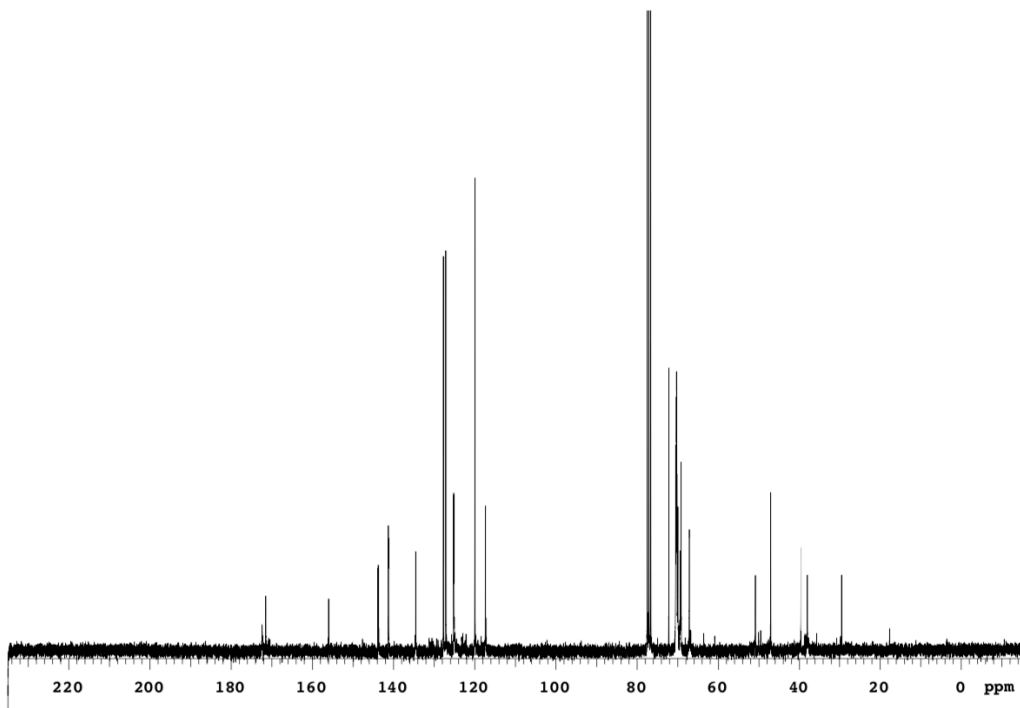
**(S)-22-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-20-oxo-4,7,10,13,16-pentaoxa-19-azatricos-1-en-23-oic acid (OX2181)**



Procedure was performed as described above for compound **QX2162** but using **QX2180** as starting material. Product was obtained as yellow oil. Yield quantitative. MS(ESI-TOF)  $m/z$  calc. for  $C_{32}H_{43}N_2O_{10}^+$  615.29, found 615.29  $[M+H]^+$ .  $^1H$  NMR (300 MHz, Chloroform-*d*)  $\delta$  7.78 (d,  $J = 7.5$  Hz, 2H), 7.63 (dd,  $J = 7.5$ , 5.5 Hz, 2H), 7.42 (t,  $J = 7.5$  Hz, 2H), 7.33 (t,  $J = 7.5$  Hz, 2H), 6.30 (d,  $J = 7.2$  Hz, 1H), 5.90 (ddt,  $J = 17.5$ , 10.5, 5.5 Hz, 1H), 5.27 (dd,  $J = 17.5$ , 1.5 Hz, 1H), 5.18 (d,  $J = 10.5$  Hz, 1H), 4.60 (dt,  $J = 8.5$ , 4.5 Hz, 1H), 4.47 – 4.42 (m, 1H), 4.39 (dd,  $J = 9.9$ , 7.5 Hz, 1H), 4.24 (t,  $J = 8.1$  Hz, 1H), 4.01 (d,  $J = 5.7$  Hz, 2H), 3.66 – 3.38 (m, 20H), 2.96 (dd,  $J = 15.0$ , 2.7 Hz, 1H), 2.76 (dd,  $J = 15.0$ , 7.2 Hz, 1H).  $^{13}C$  NMR (300 MHz, Chloroform-*d*)  $\delta$  171.49, 156.00, 143.93, 143.78, 141.28, 134.54, 127.73, 127.14, 125.25, 125.18, 119.97, 117.40, 72.18, 70.51, 70.34, 70.30, 70.18, 69.95, 69.37, 69.17, 67.14, 50.87, 47.09, 39.63, 37.91.







## SI References

1. Q. Xiao, N. A. Becar, N. P. Brown, M. S. Smith, K. L. Stern, S. R. E. Draper, K. P. Thompson and J. L. Price, *Org Biomol Chem*, 2018, **16**, 8933-8939.
2. P. B. Lawrence, Y. Gavrilov, S. S. Matthews, M. I. Langlois, D. Shental-Bechor, H. M. Greenblatt, B. K. Pandey, M. S. Smith, R. Paxman, C. D. Torgerson, J. P. Merrell, C. C. Ritz, M. B. Prigozhin, Y. Levy and J. L. Price, *J Am Chem Soc*, 2014, **136**, 17547-17560.