Supporting Information

for

Nanomolar Binding of Peptides Containing Noncanonical Amino Acids by a Synthetic Receptor

Leigh A. Logsdon, Christopher L. Schardon, Vijayakumar Ramalingam, Sharon K. Kwee, and Adam R. Urbach*

Department of Chemistry, Trinity University, 1 Trinity Place, San Antonio, TX 78212

* to whom correspondence should be addressed: aurbach@trinity.edu

Pages	**Contents**
S2-S3	Isothermal titration calorimetry data
S4-S11	1D ^{1}H NMR spectra of guests and their Q7 complexes
S12-S13	2D ROESY spectra of tBuPhe-Gly-Gly (**2**) in the presence and absence of Q7
S14-S25	Mass spectra
S26	Table of limiting chemical shifts for modified phenylalanine sidechains
S27-S29	Derivation of the competitive equilibrium analysis
S30-S34	Experimental details

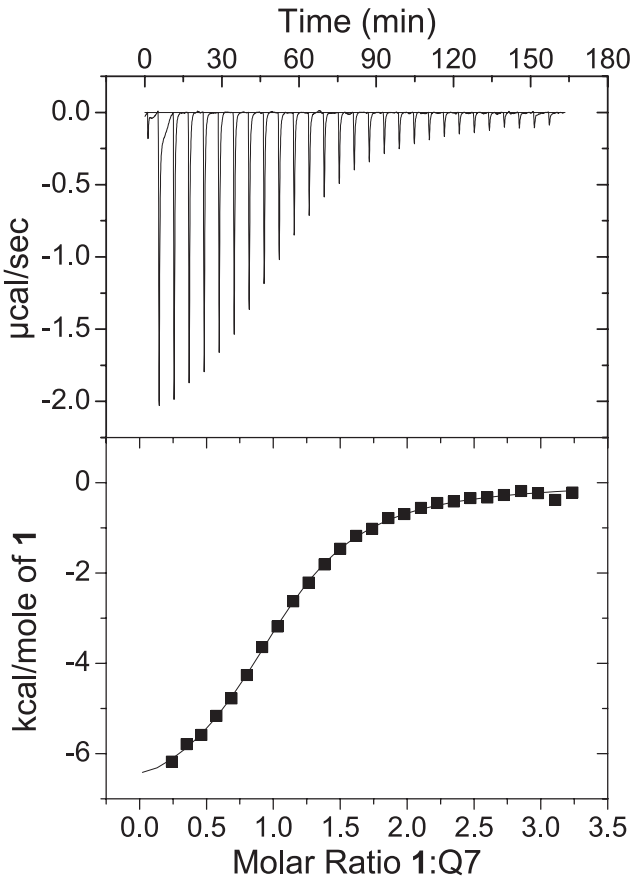


Figure S1. ITC of **1** titrated into Q7.

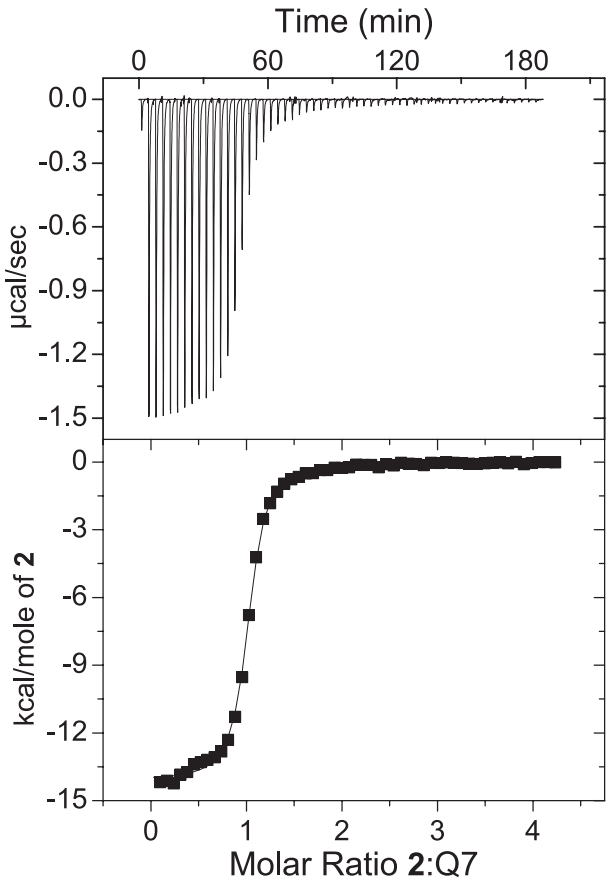


Figure S2. ITC of **2** titrated into Q7.

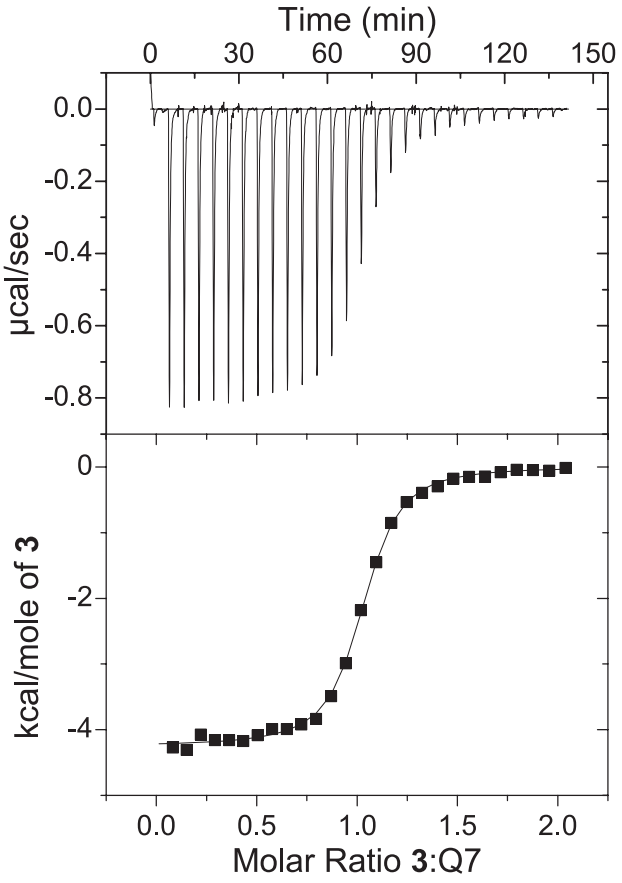


Figure S3. ITC of **3** titrated into Q7.

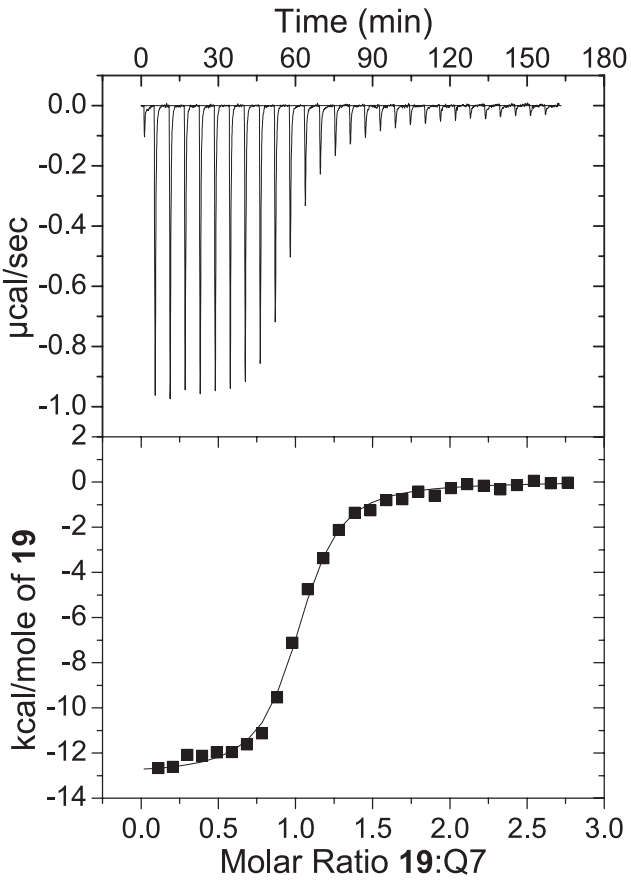


Figure S4. ITC of **19** titrated into Q7.

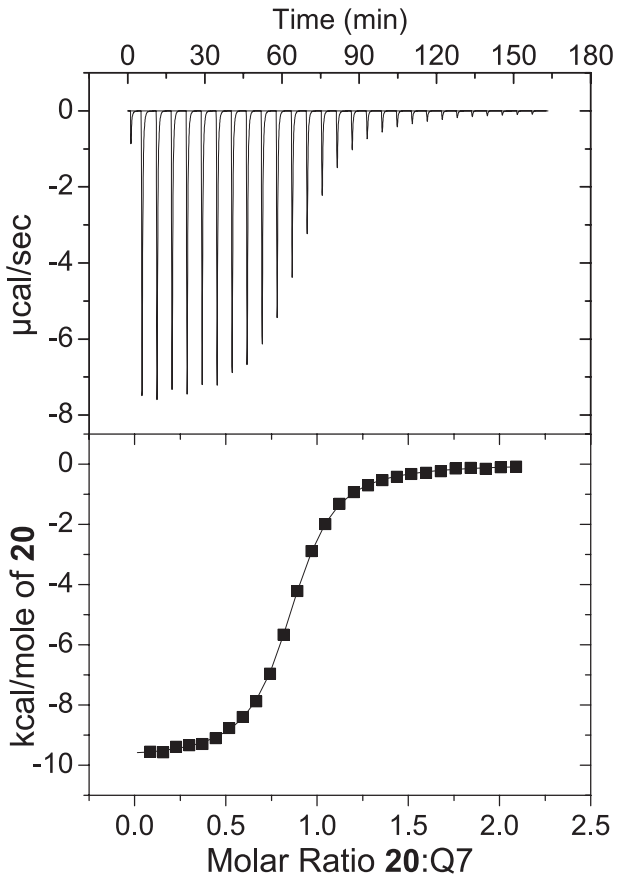


Figure S5. ITC of **20** titrated into Q7.

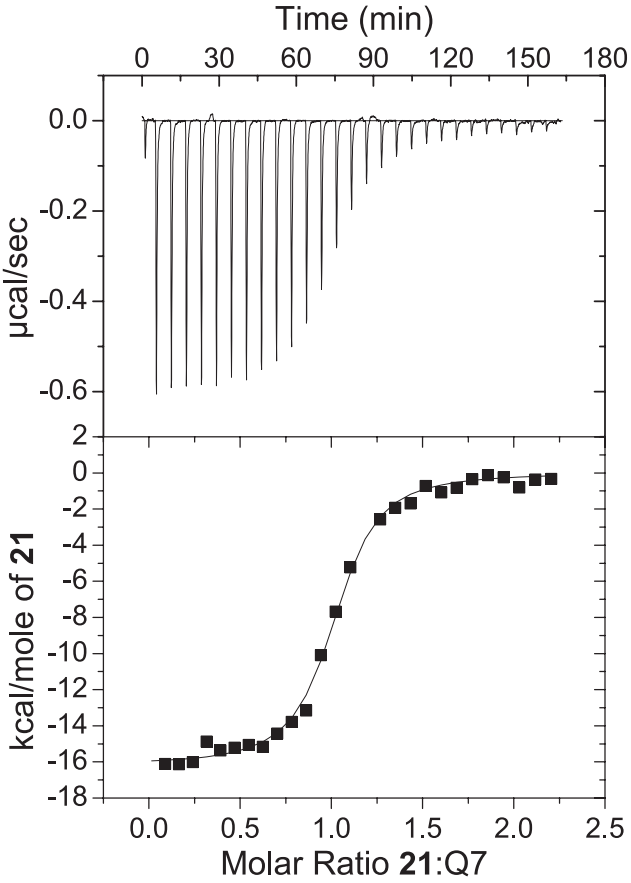


Figure S6. ITC of **21** titrated into Q7.

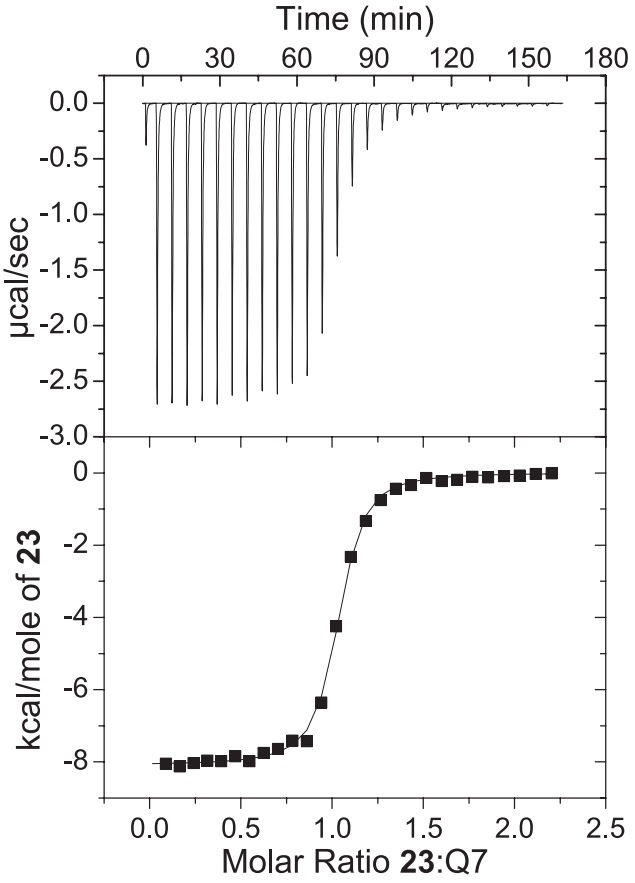


Figure S7. ITC of **23** titrated into Q7.

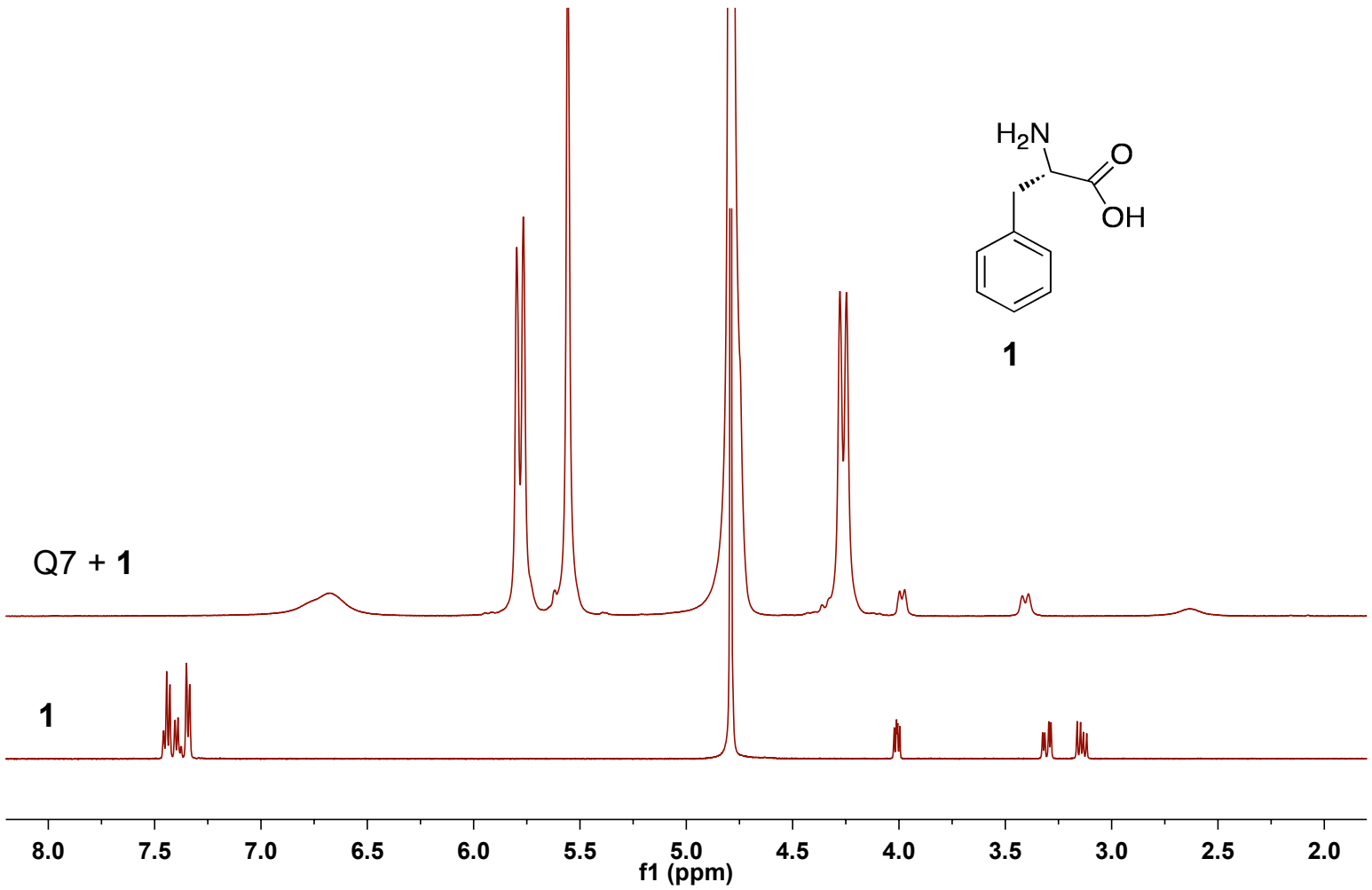


Figure S8: Overlay of ^{1}H NMR spectra for (**1,** 2 mM + Q7, 2 mM)(top) with **1,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

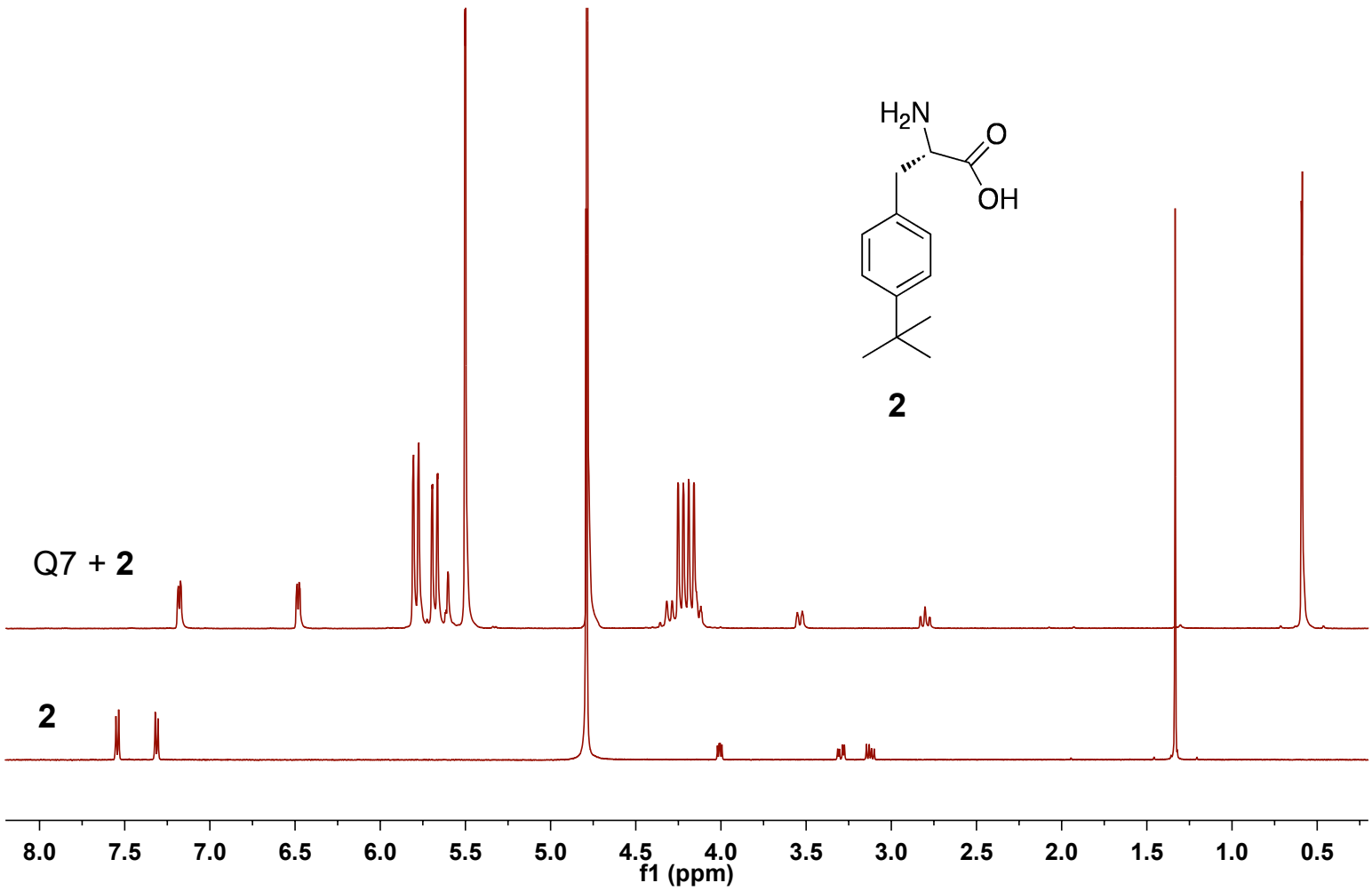


Figure S9: Overlay of ^{1}H NMR spectra for (**2,** 2 mM + Q7, 2 mM)(top) with **2,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

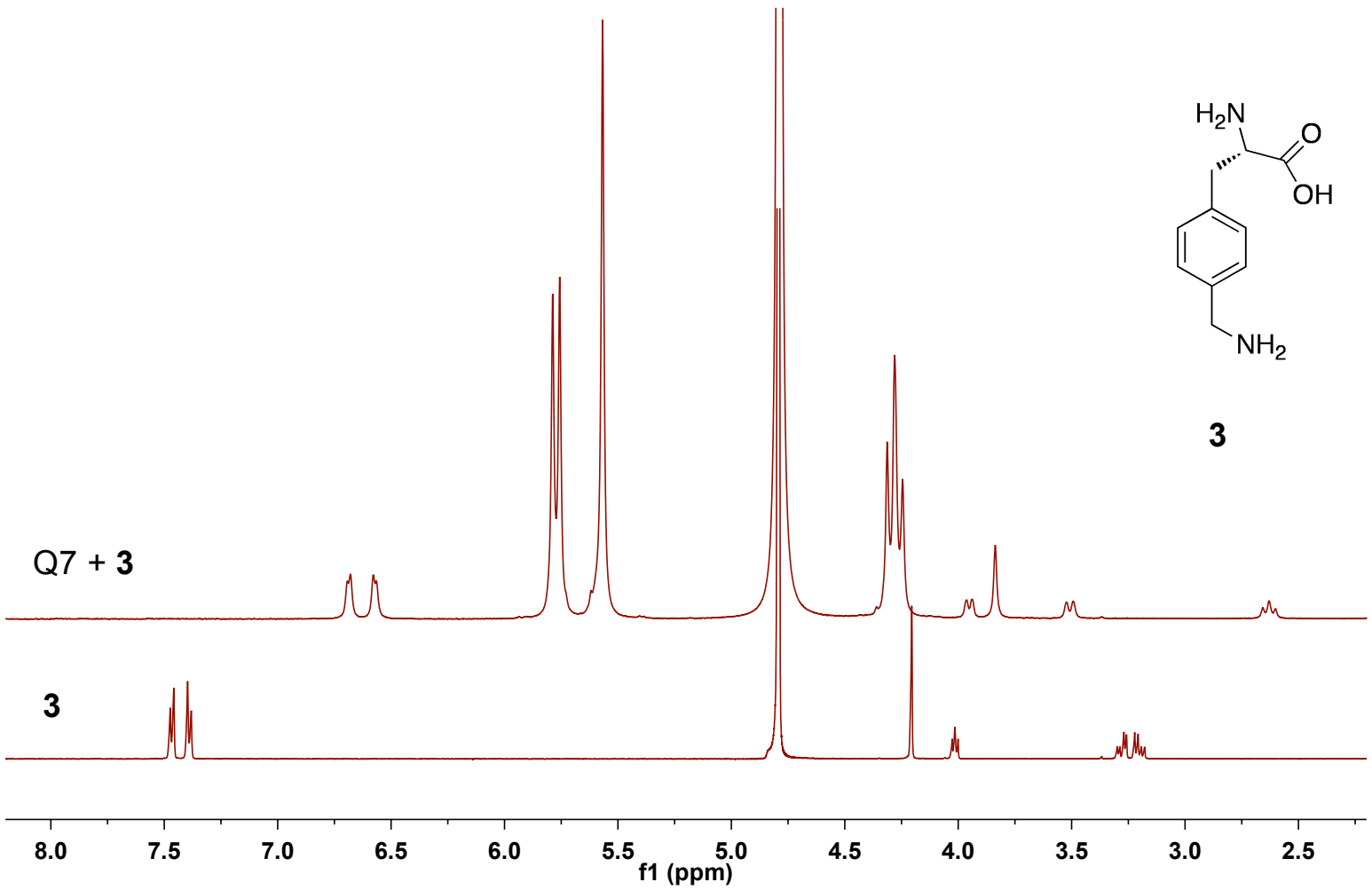


Figure S10: Overlay of ^{1}H NMR spectra for (**3,** 2 mM + Q7, 2 mM)(top) with **3,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

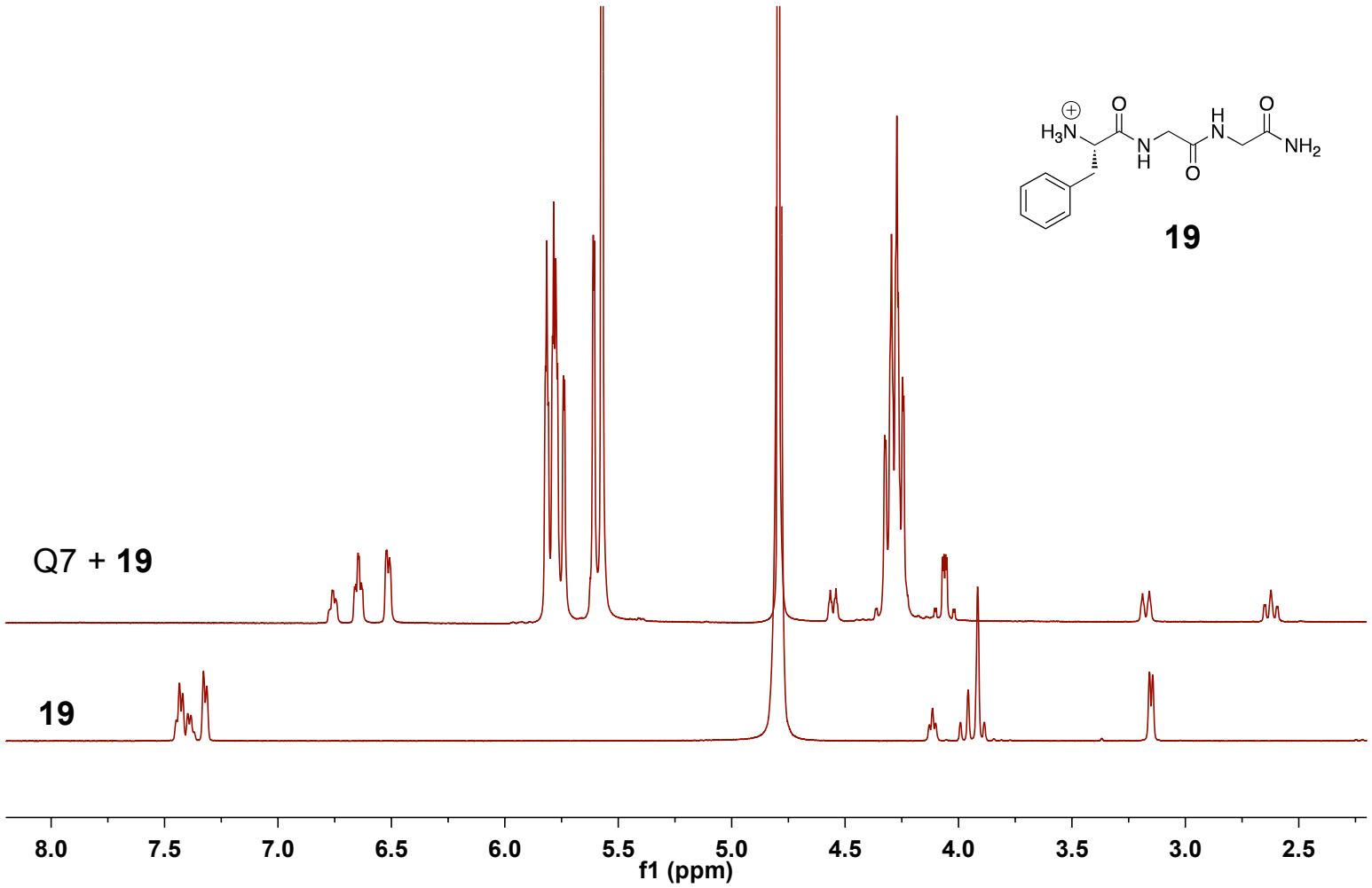


Figure S11: Overlay of ^{1}H NMR spectra for (**19,** 2 mM + Q7, 2 mM)(top) with **19,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

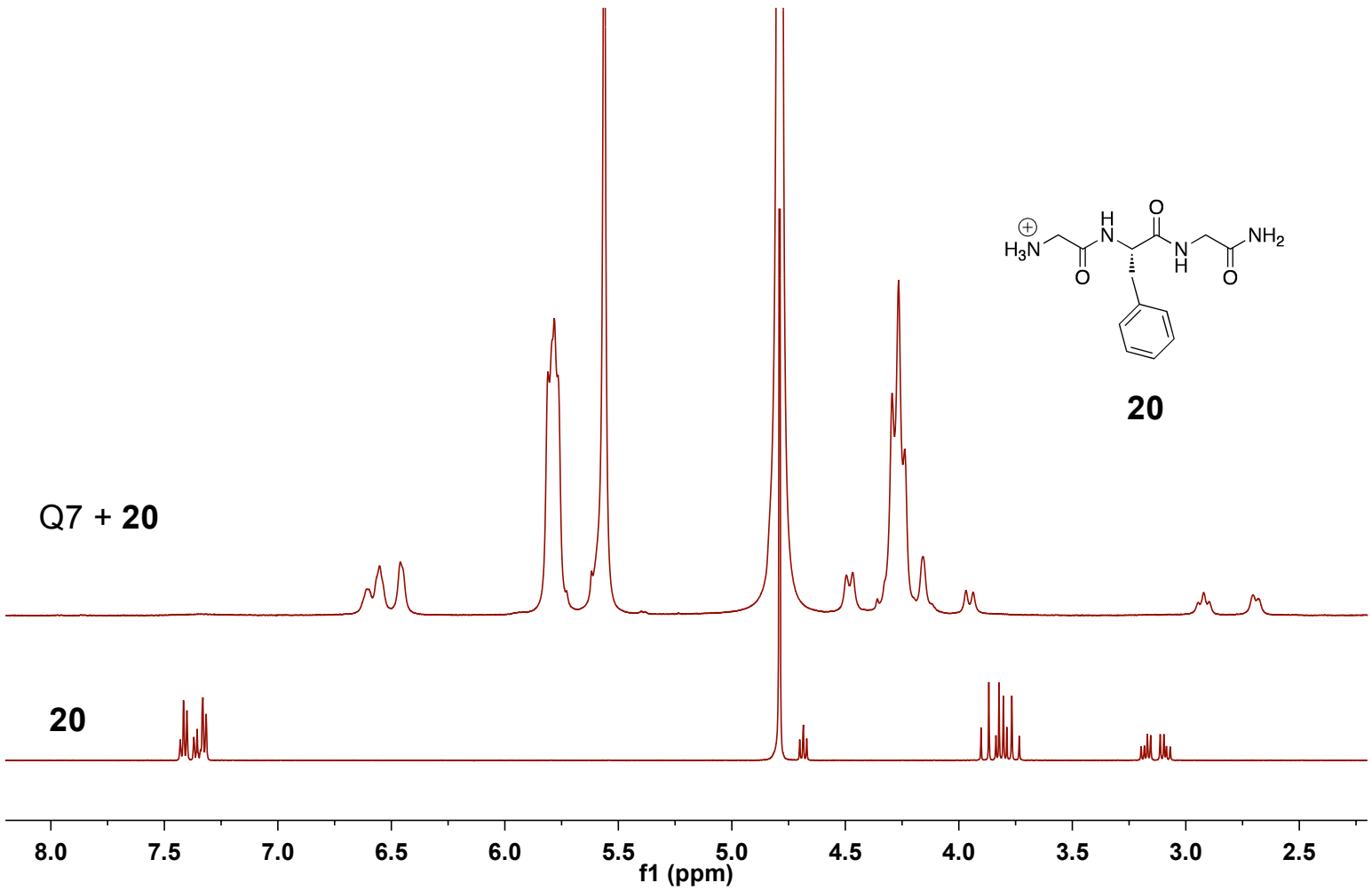


Figure S12: Overlay of ^{1}H NMR spectra for (**20,** 2 mM + Q7, 2 mM)(top) with **20,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

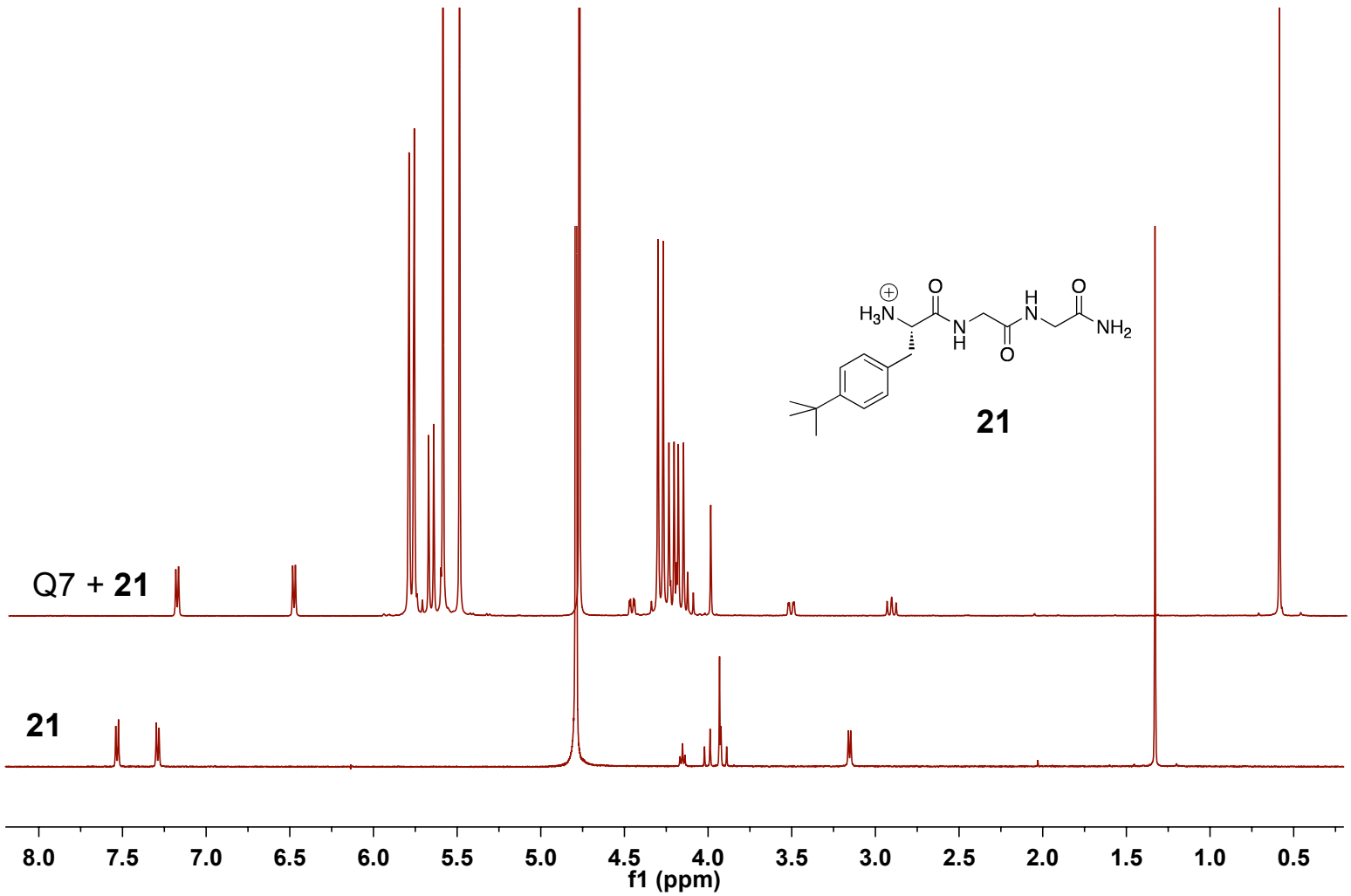


Figure S13: Overlay of ^{1}H NMR spectra for (**21,** 2 mM + Q7, 2 mM)(top) with **21,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

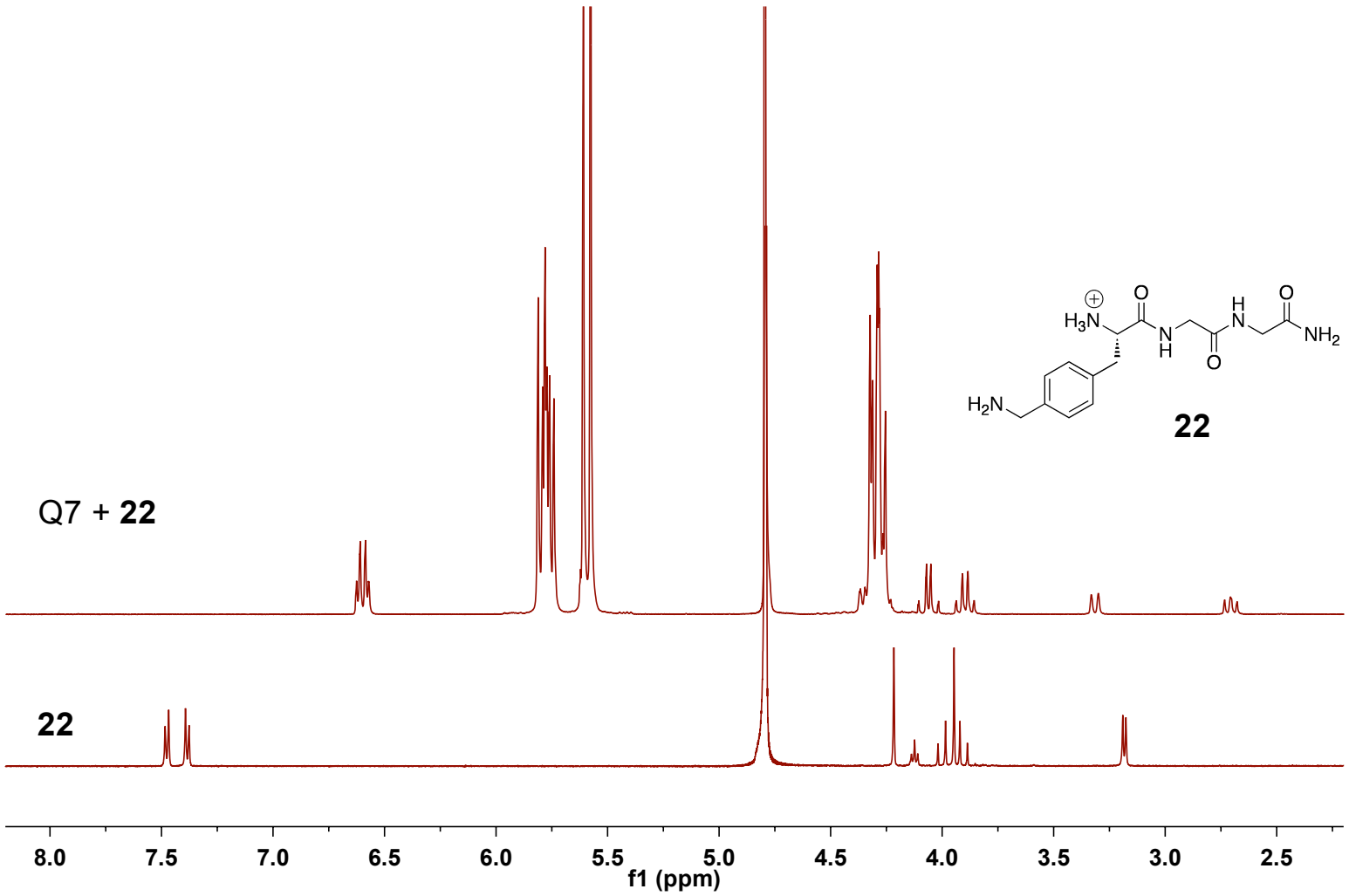


Figure S14: Overlay of ^{1}H NMR spectra for (**22,** 2 mM + Q7, 2 mM)(top) with **22,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

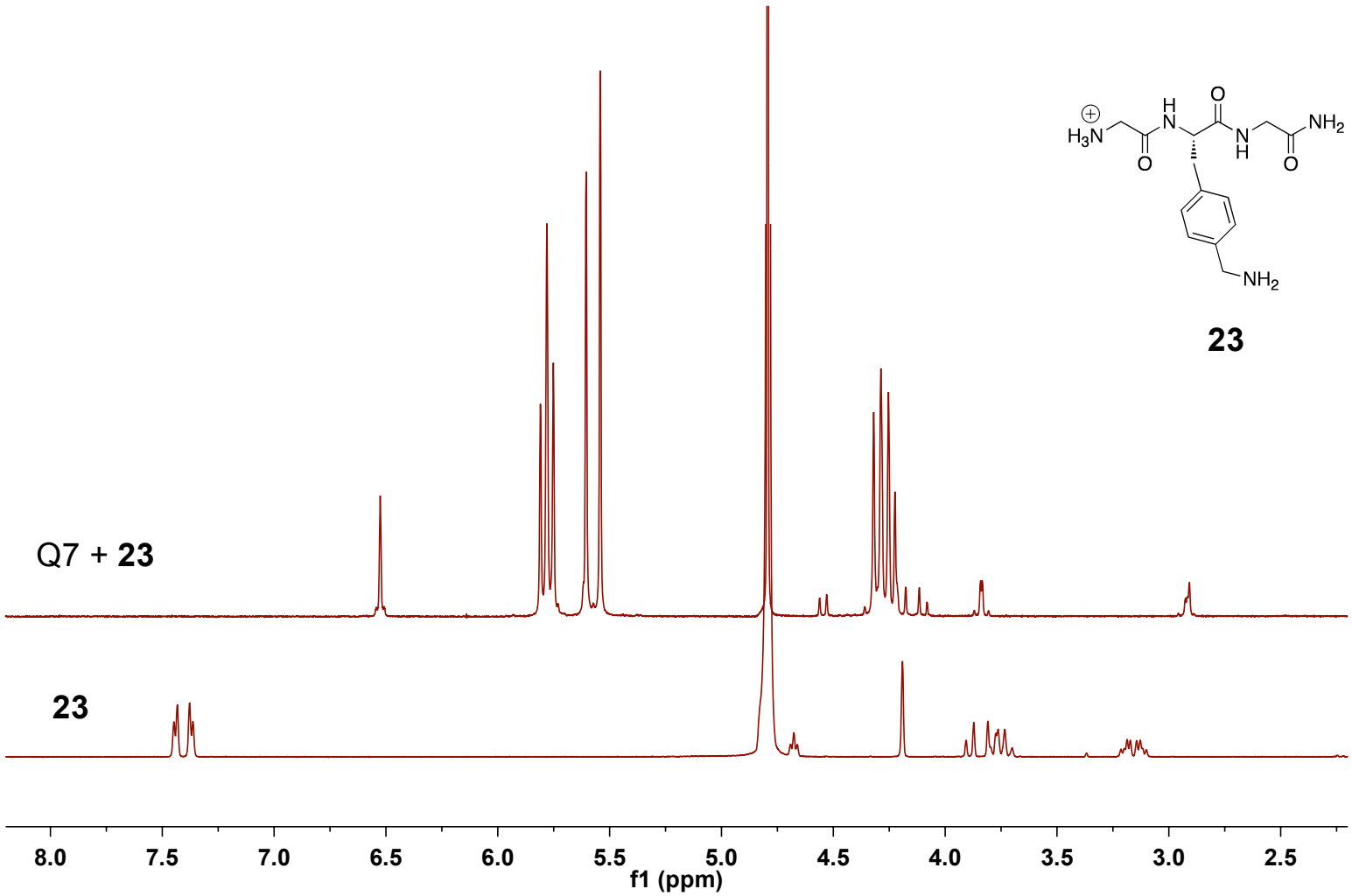
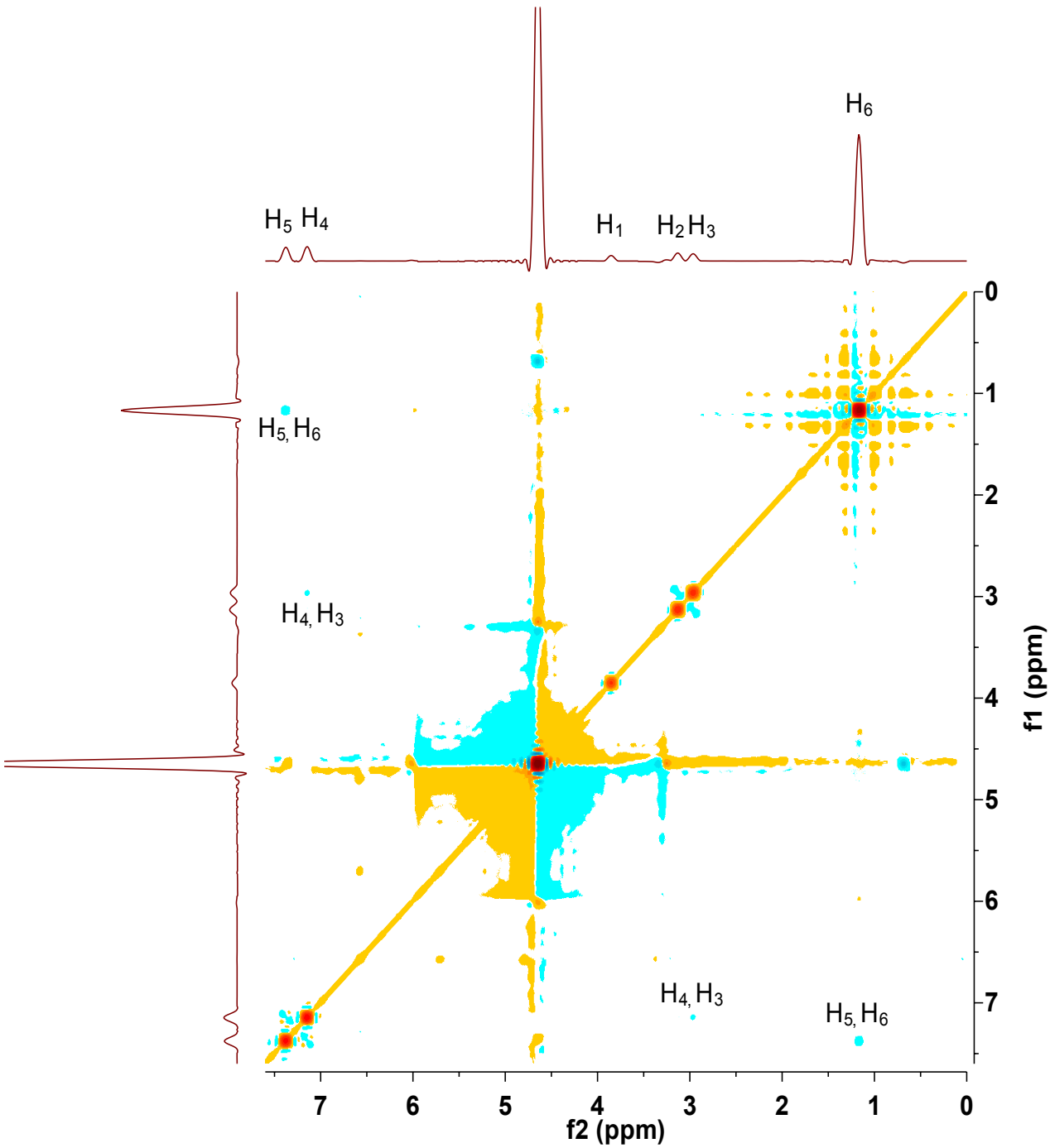
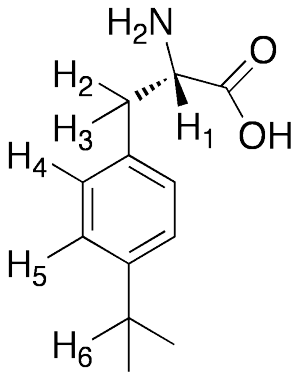


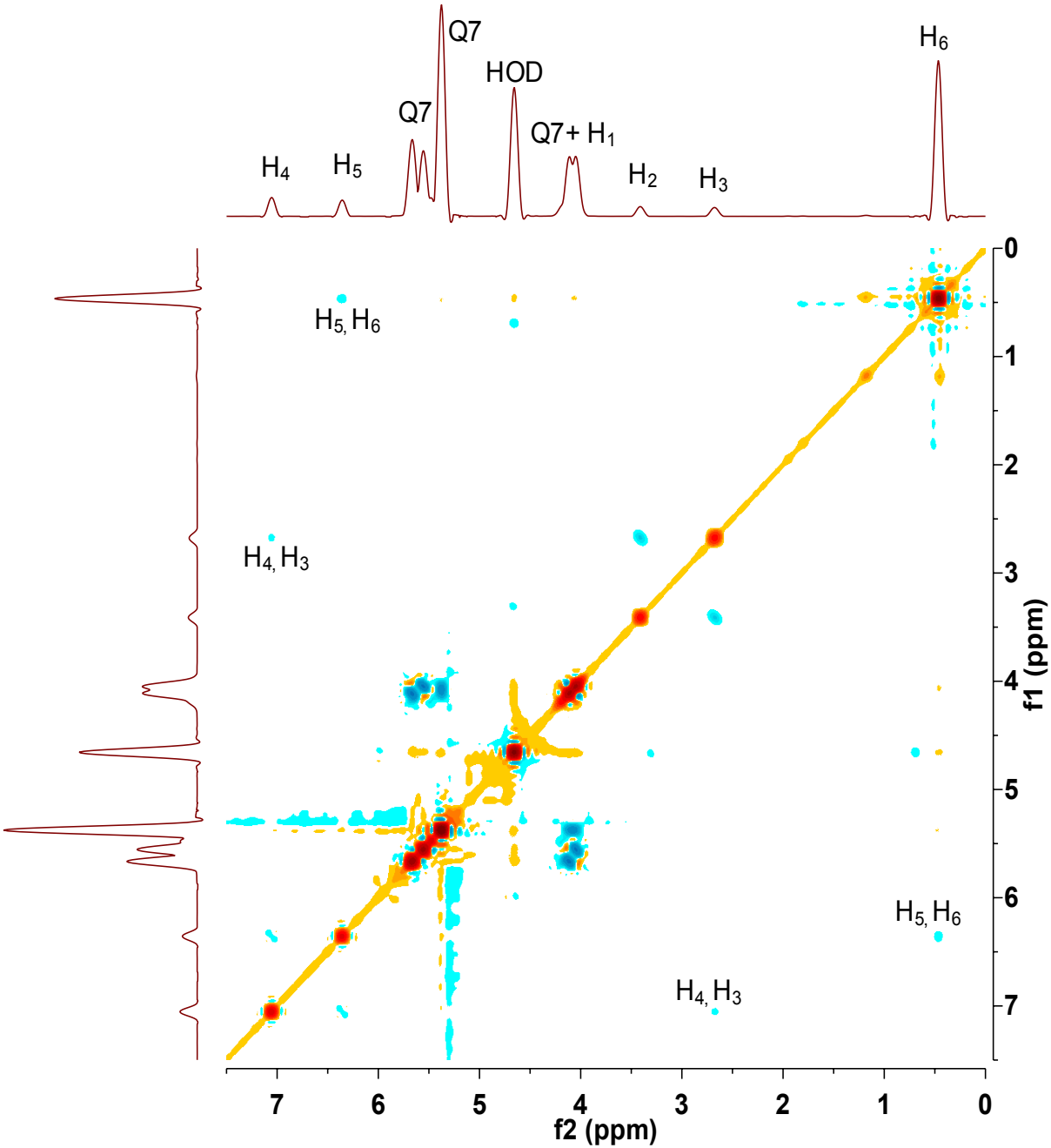
Figure S15: Overlay of ^{1}H NMR spectra for (**23,** 2 mM + Q7, 2 mM)(top) with **23,** 2 mM (bottom) in D_20 with 10 mM sodium phosphate, pH 7.4.

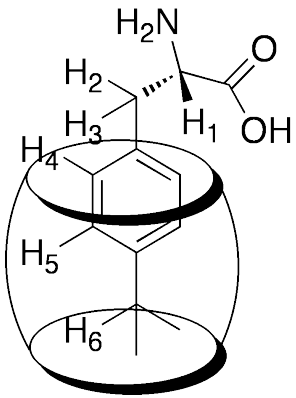




2

Figure S16: 500 MHz ROESY spectrum tBu-Phe (**2**) at 25 °C in deuterium oxide with 10 mM sodium phosphate, pH 7.4.





2•Q7

Figure S17: 500 MHz ROESY spectrum of an equimolar mixture of Q7 and tBu-Phe (**2**) at 25 °C in deuterium oxide with 10 mM sodium phosphate, pH 7.4.

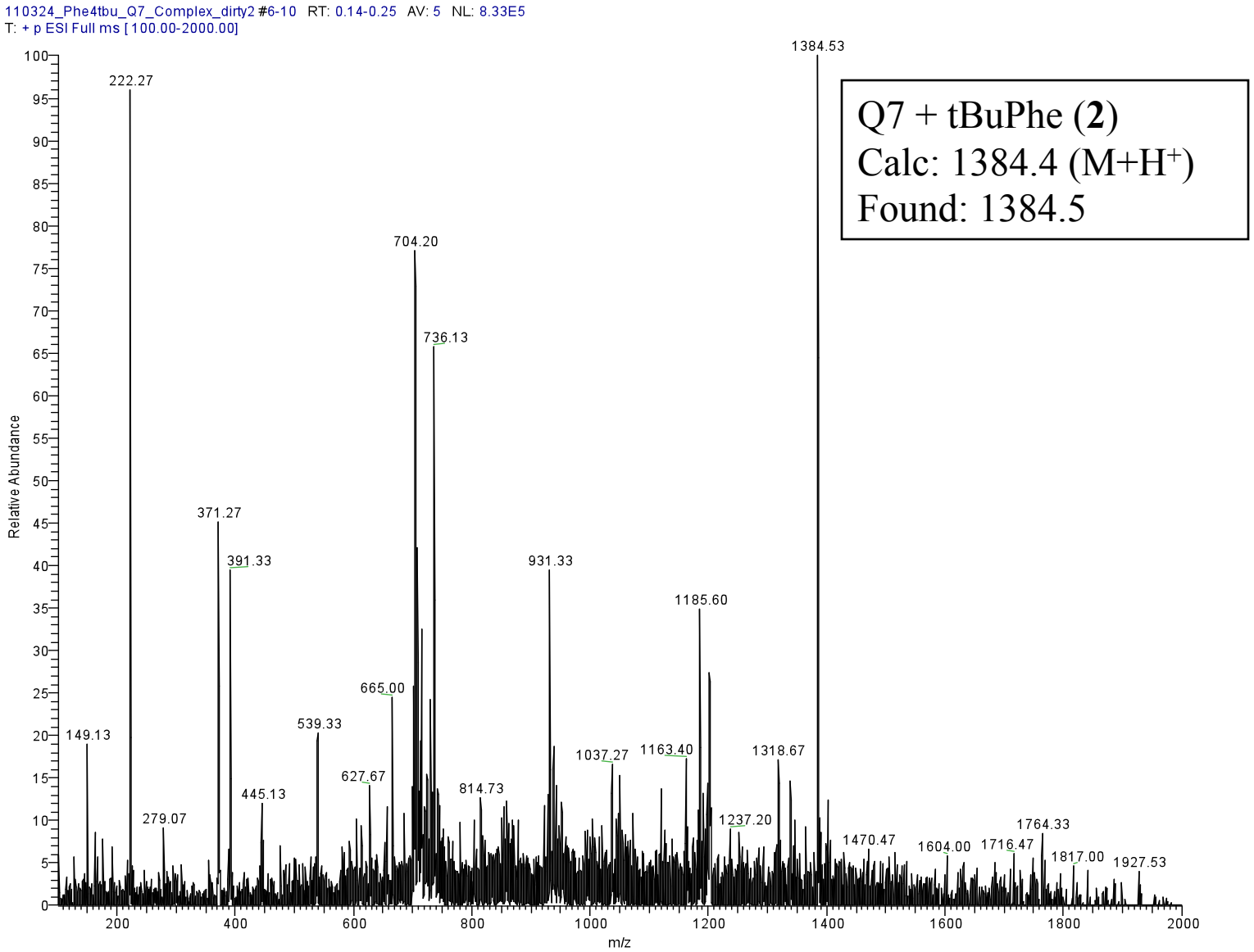
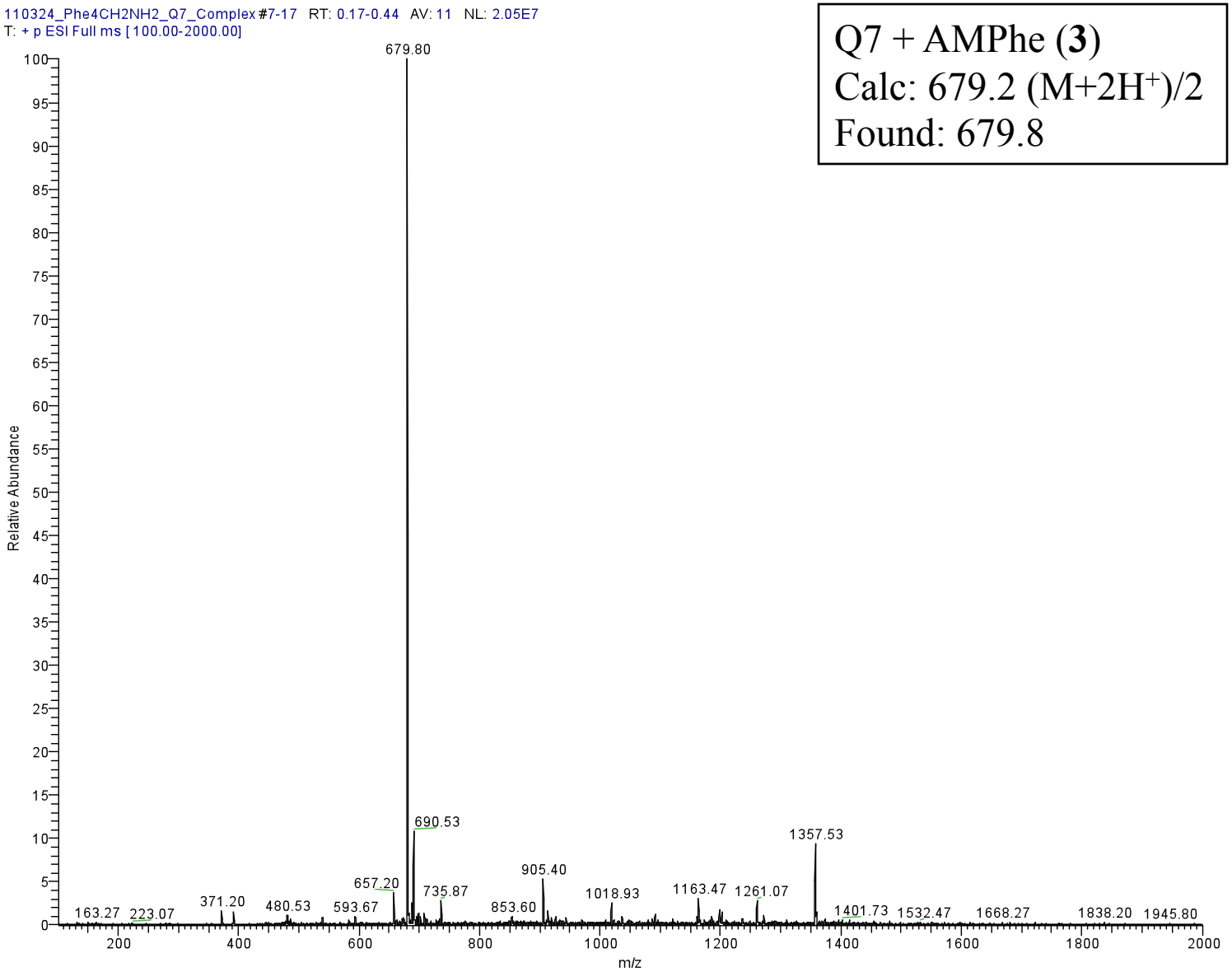
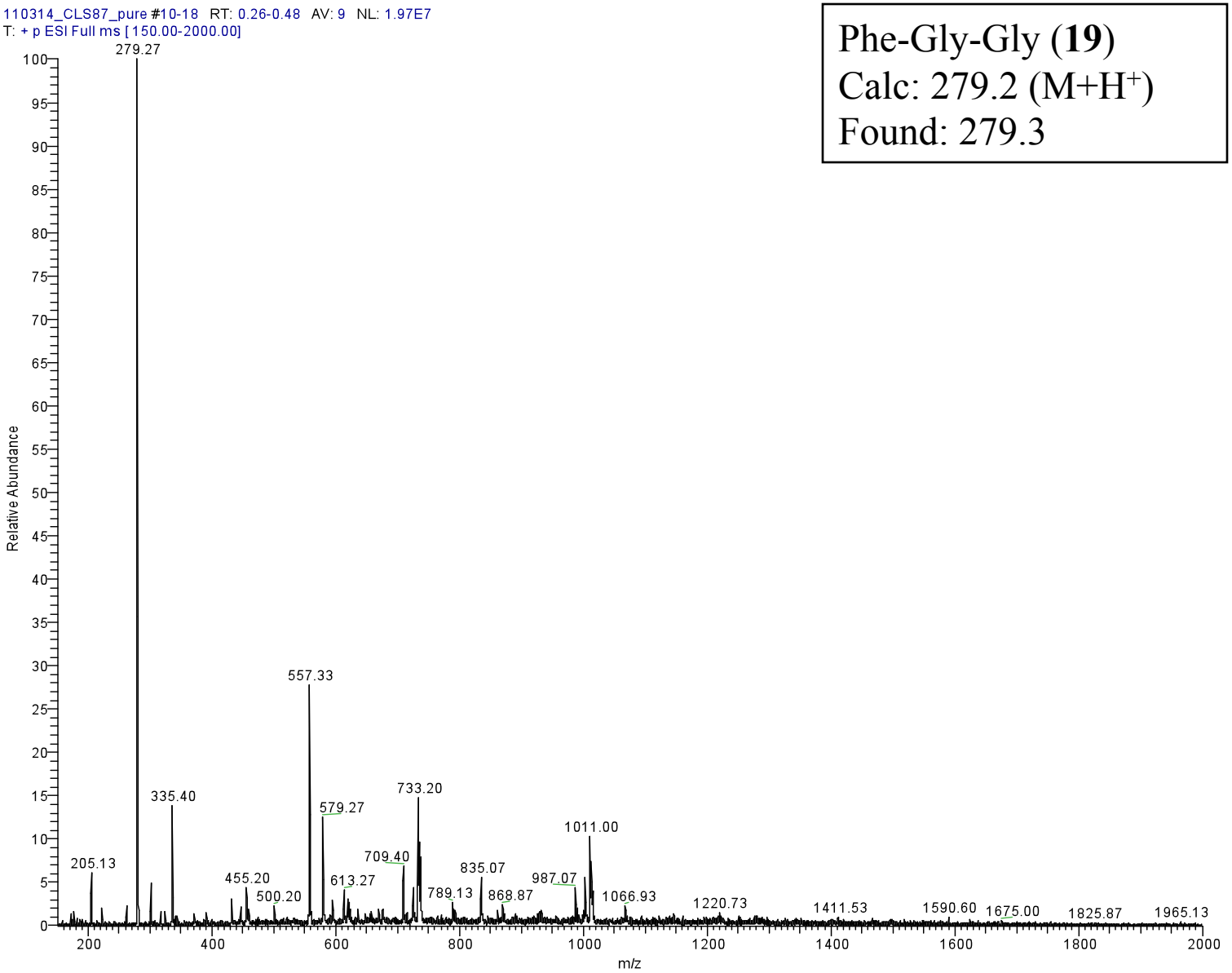


Figure S18. ESI mass spectrum of Q7 + **2** obtained in positive ion mode.



Q7 + AMPhe (**3**)
Calc: 679.2 $(M+2H^+)/2$
Found: 679.8

Figure S19. ESI mass spectrum of Q7 + **3** obtained in positive ion mode.



Phe-Gly-Gly (**19**)
Calc: 279.2 ($M+H^+$)
Found: 279.3

Figure S20. ESI mass spectrum of **19** obtained in positive ion mode.

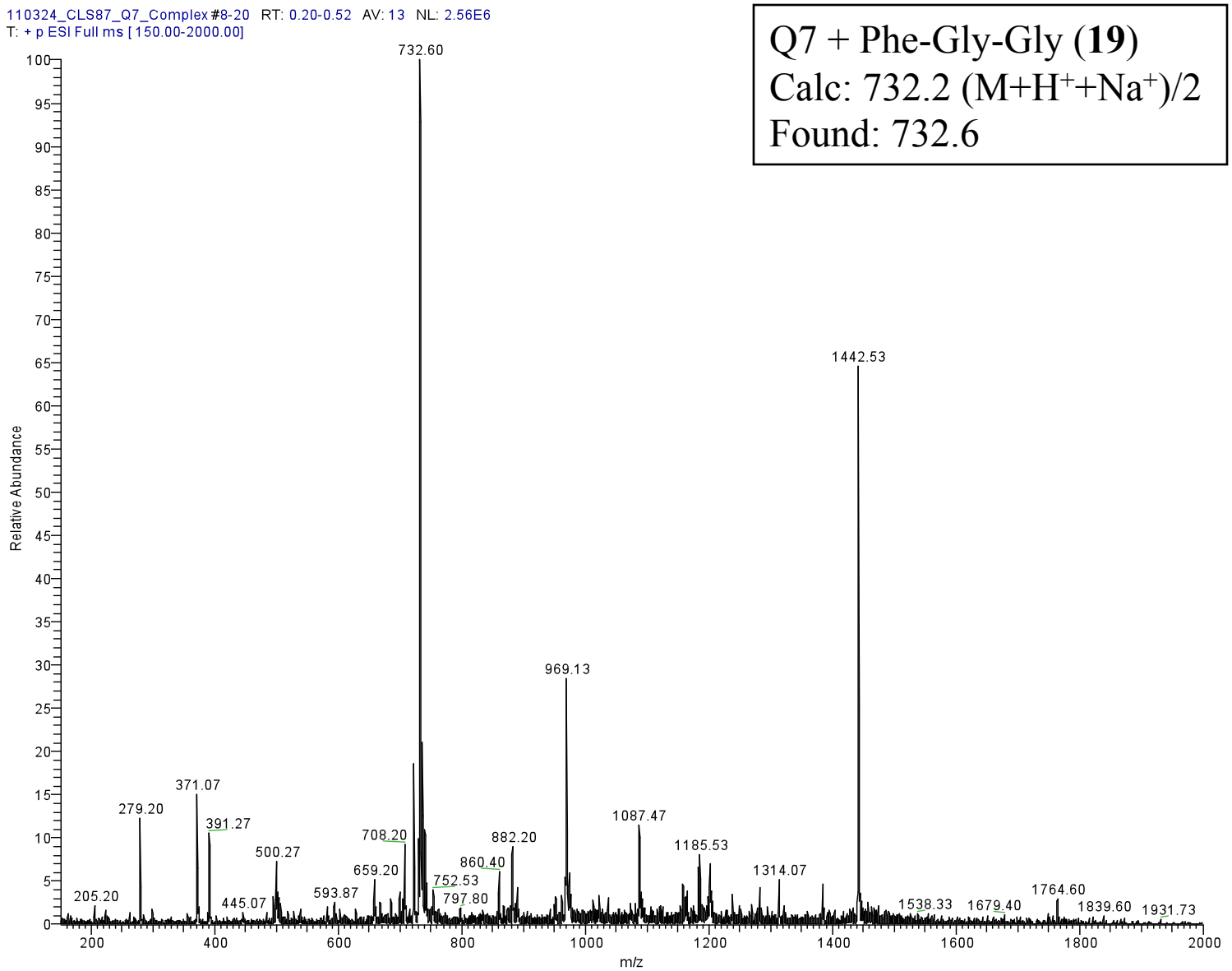
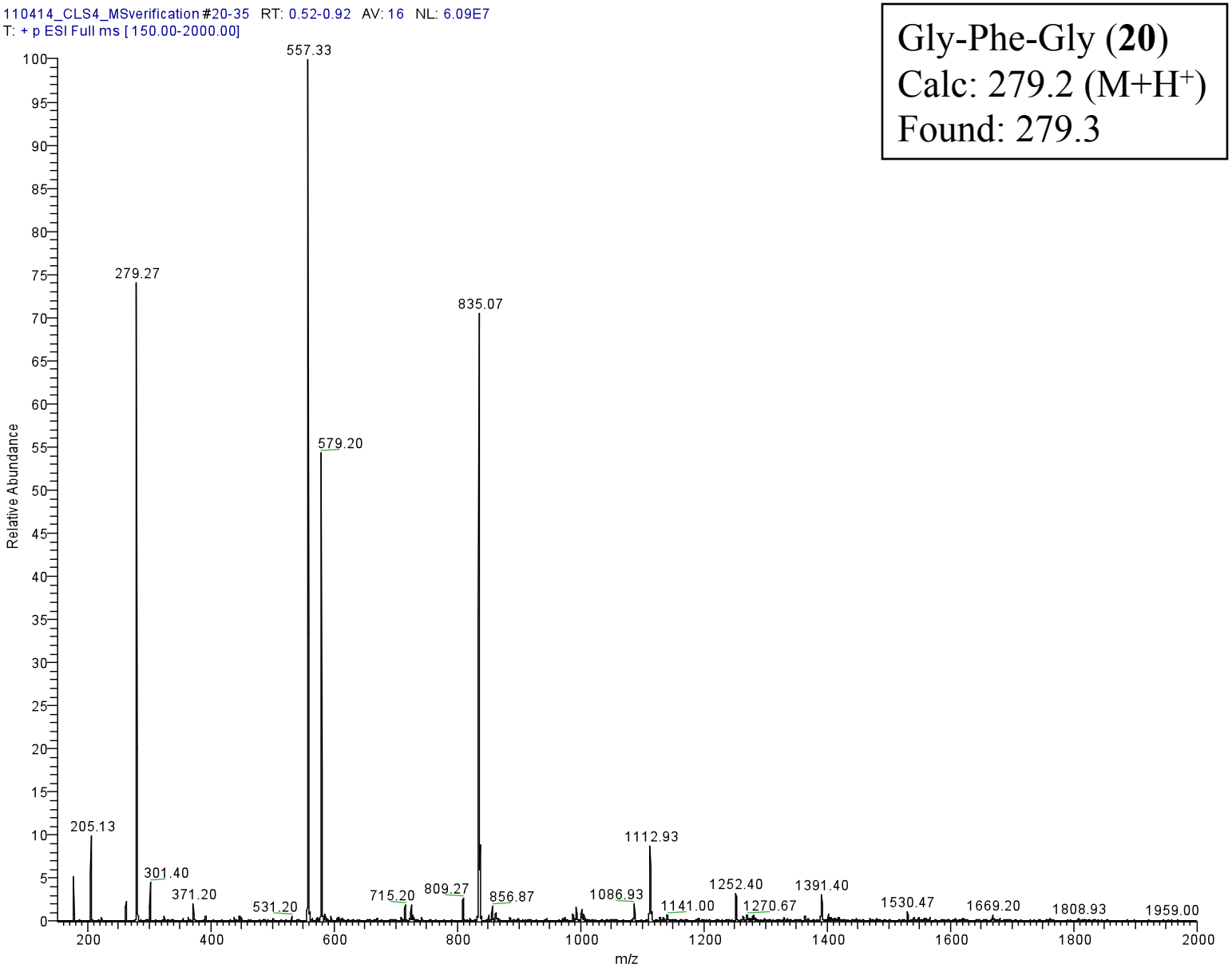


Figure S21. ESI mass spectrum of Q7 + **19** obtained in positive ion mode.



Gly-Phe-Gly (**20**)
Calc: 279.2 ($M+H^+$)
Found: 279.3

Figure S22. ESI mass spectrum of **20** obtained in positive ion mode.

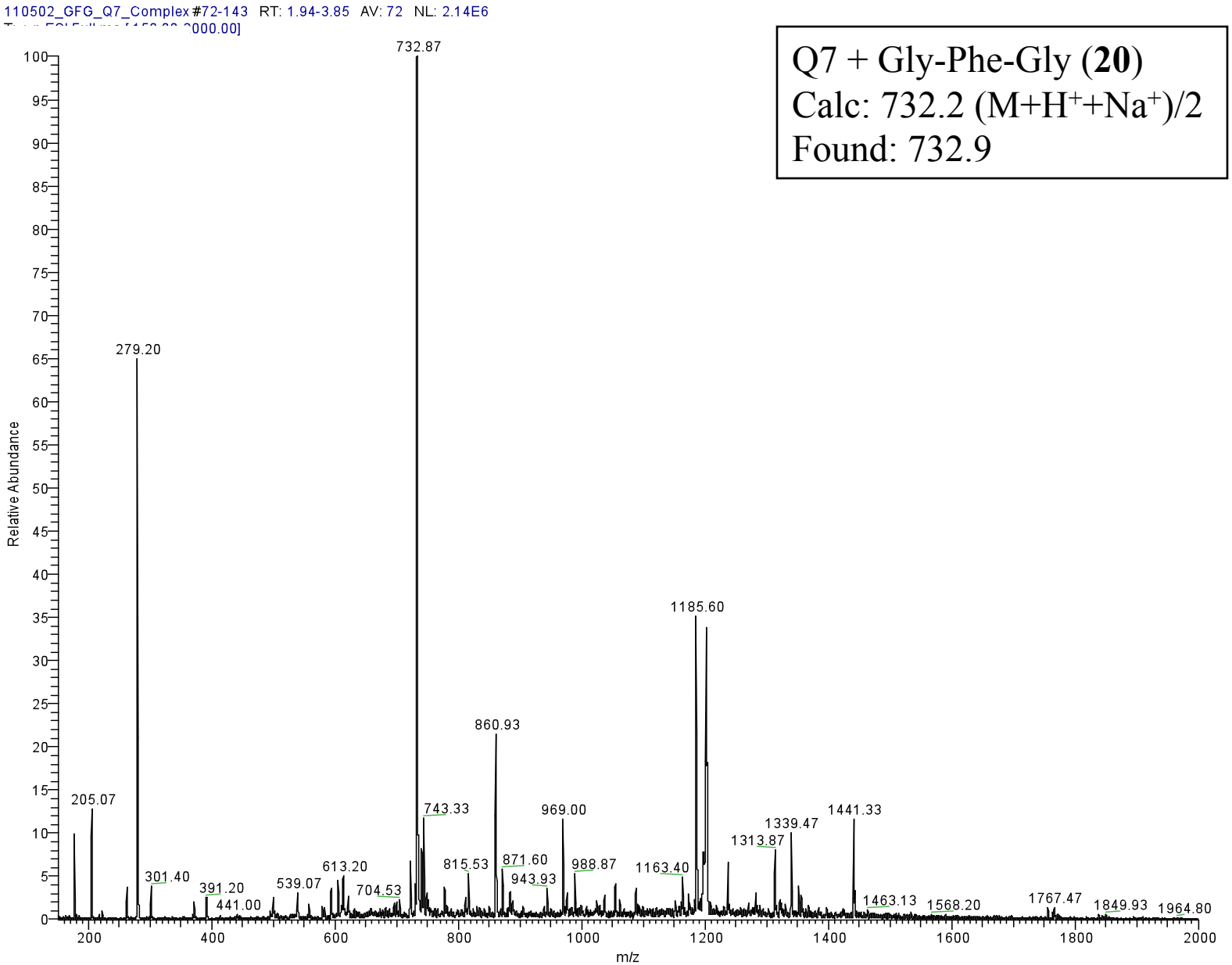
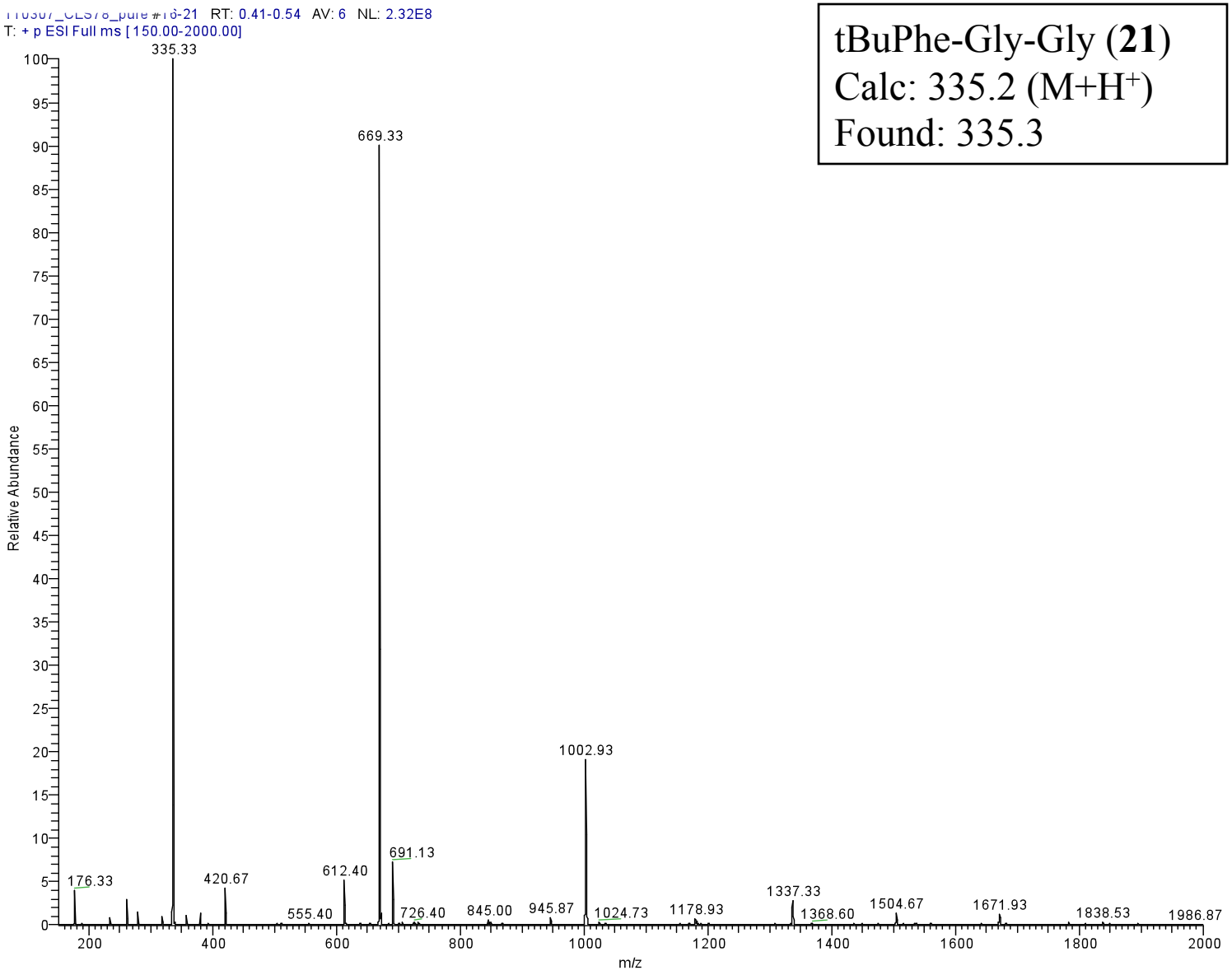


Figure S23. ESI mass spectrum of Q7 + **20** obtained in positive ion mode.



tBuPhe-Gly-Gly (**21**)
Calc: 335.2 ($M+H^+$)
Found: 335.3

Figure S24. ESI mass spectrum of **21** obtained in positive ion mode.

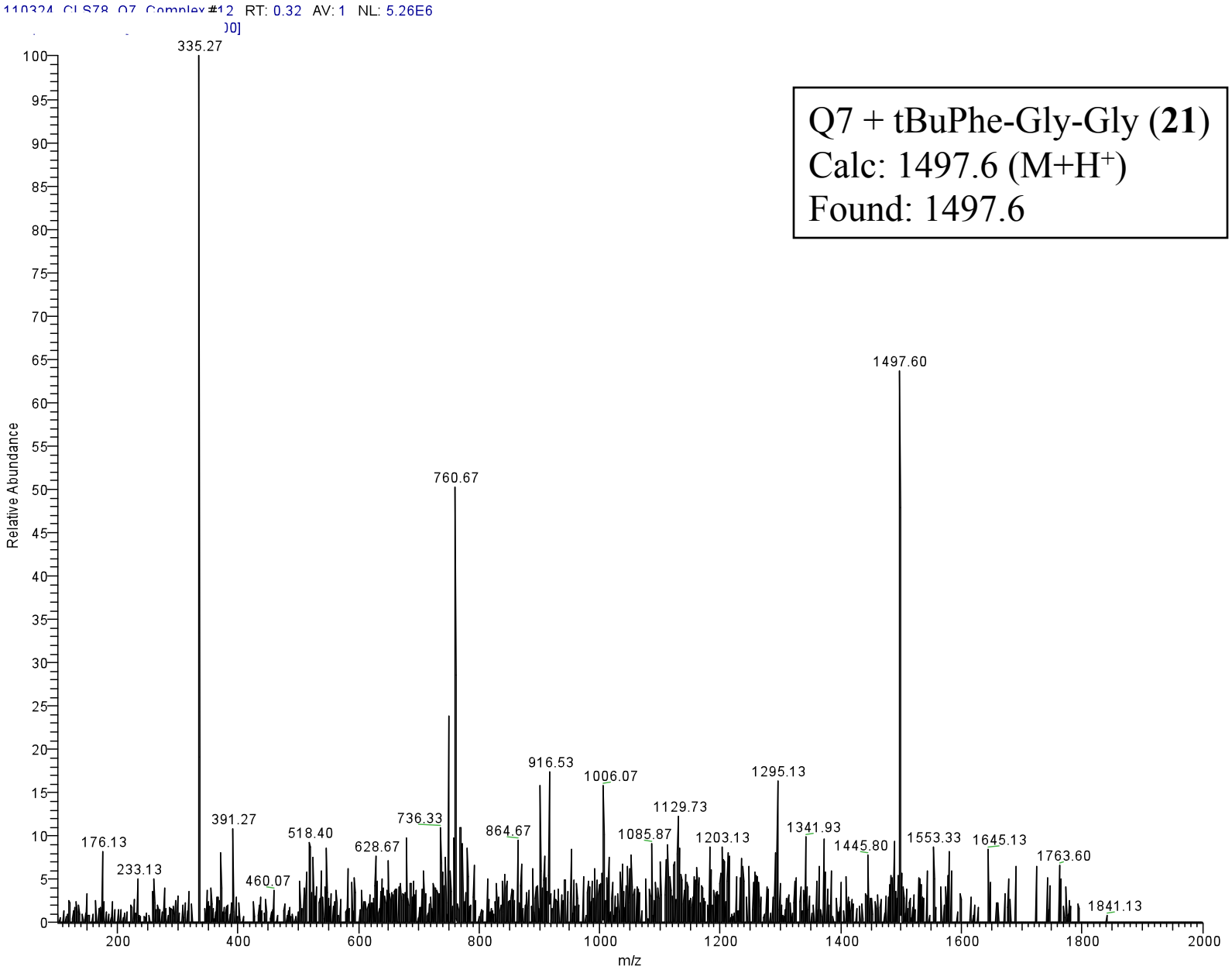
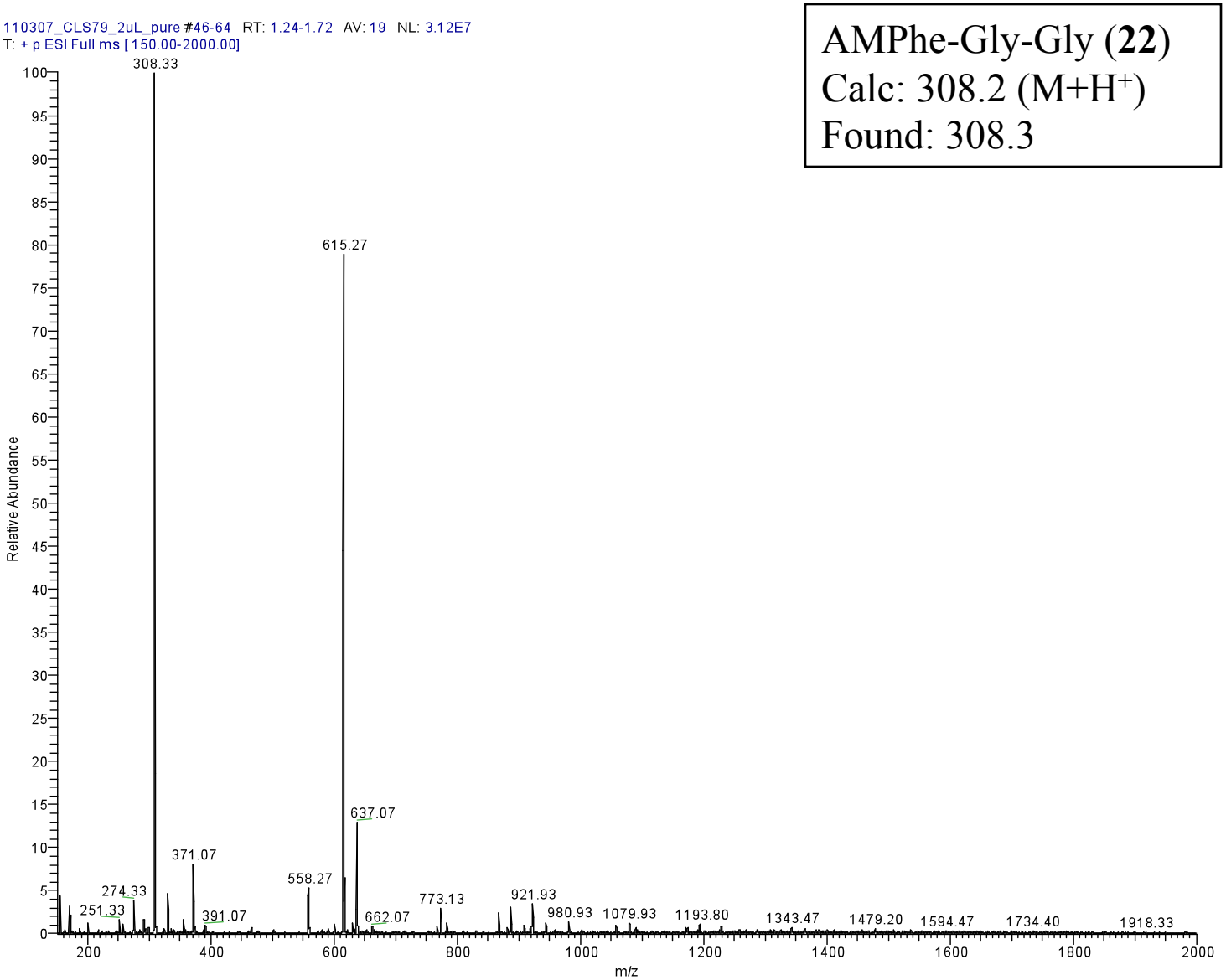
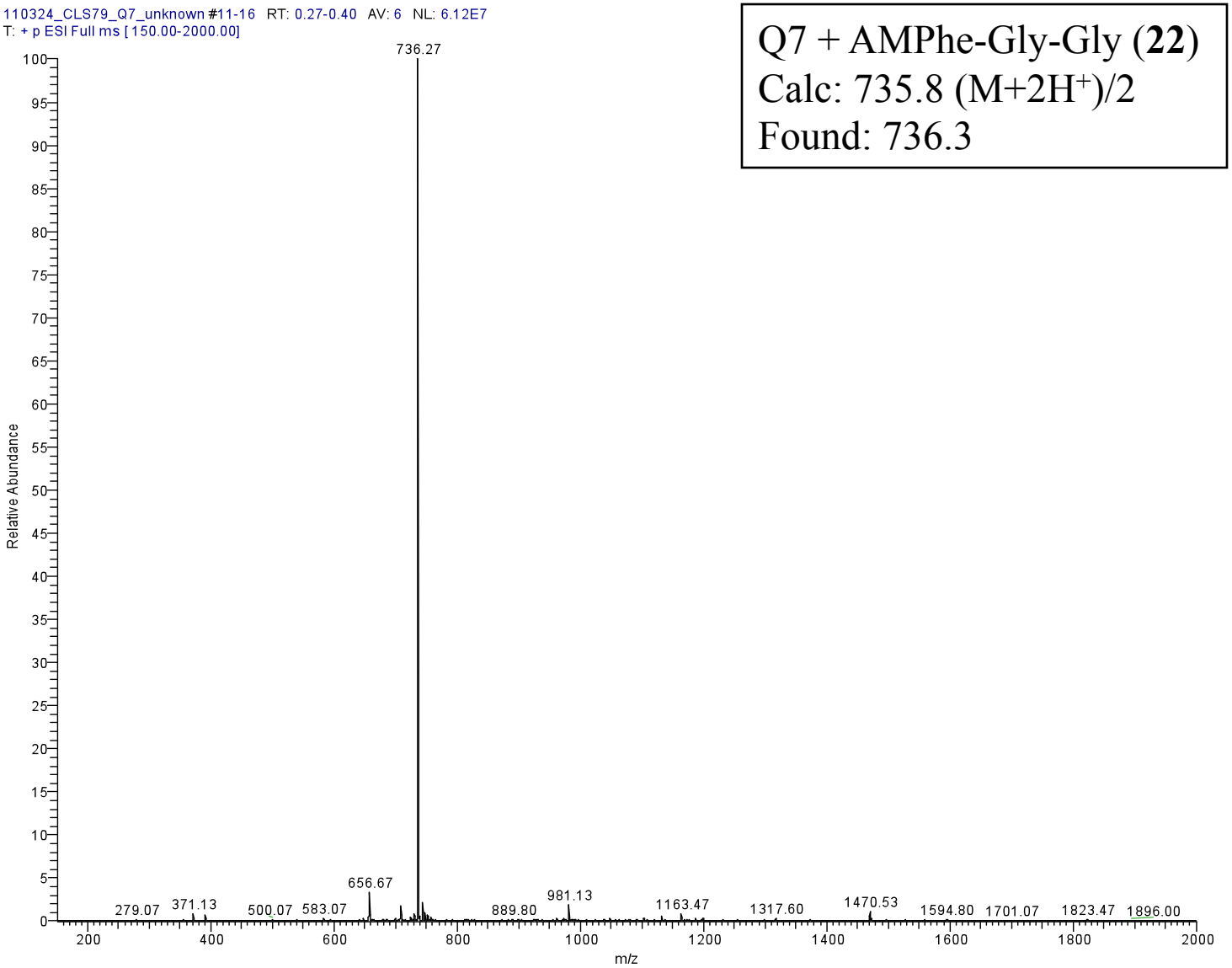


Figure S25. ESI mass spectrum of Q7 + **21** obtained in positive ion mode.



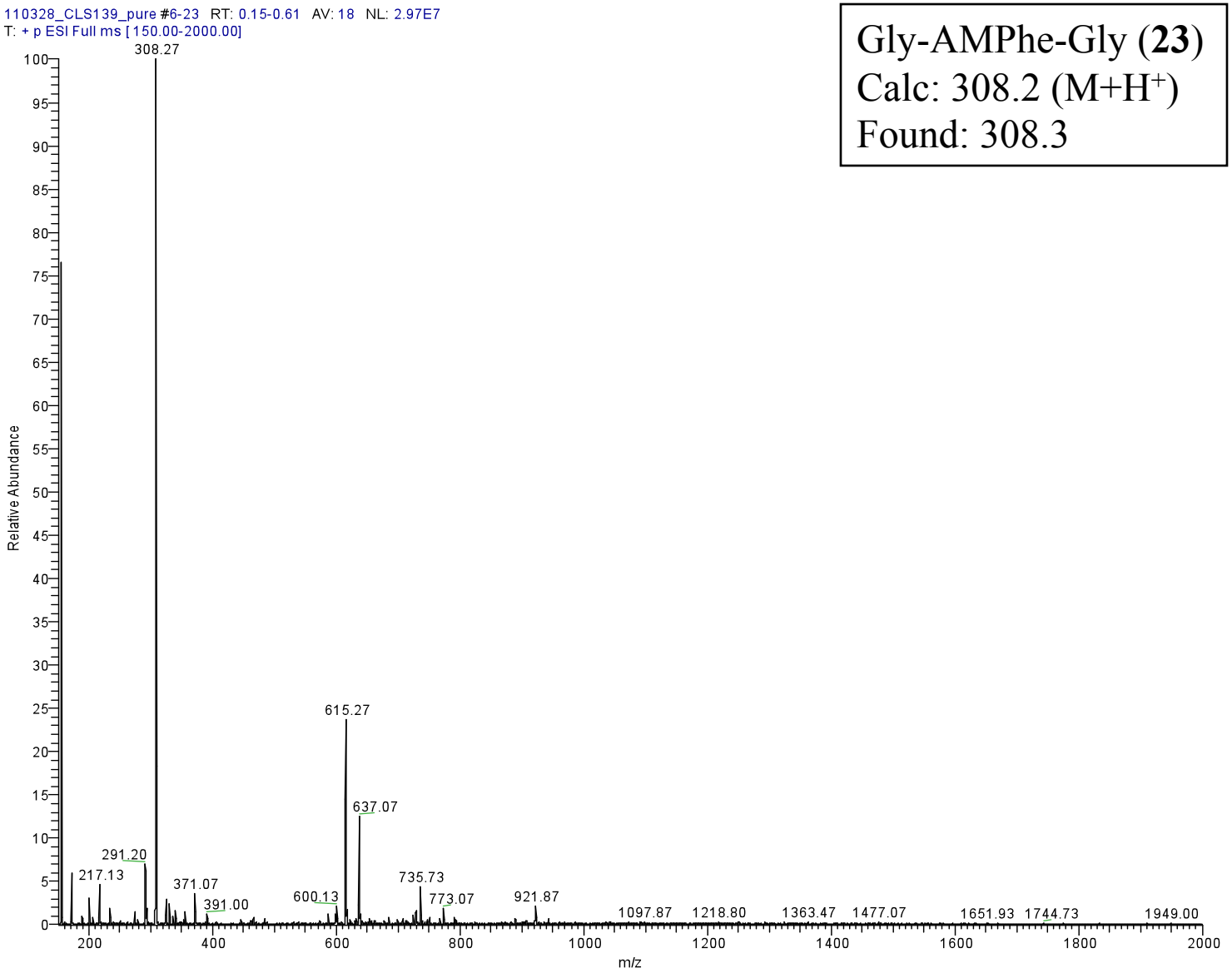
AMPhe-Gly-Gly (**22**)
Calc: 308.2 ($M+H^+$)
Found: 308.3

Figure S26. ESI mass spectrum of **22** obtained in positive ion mode.



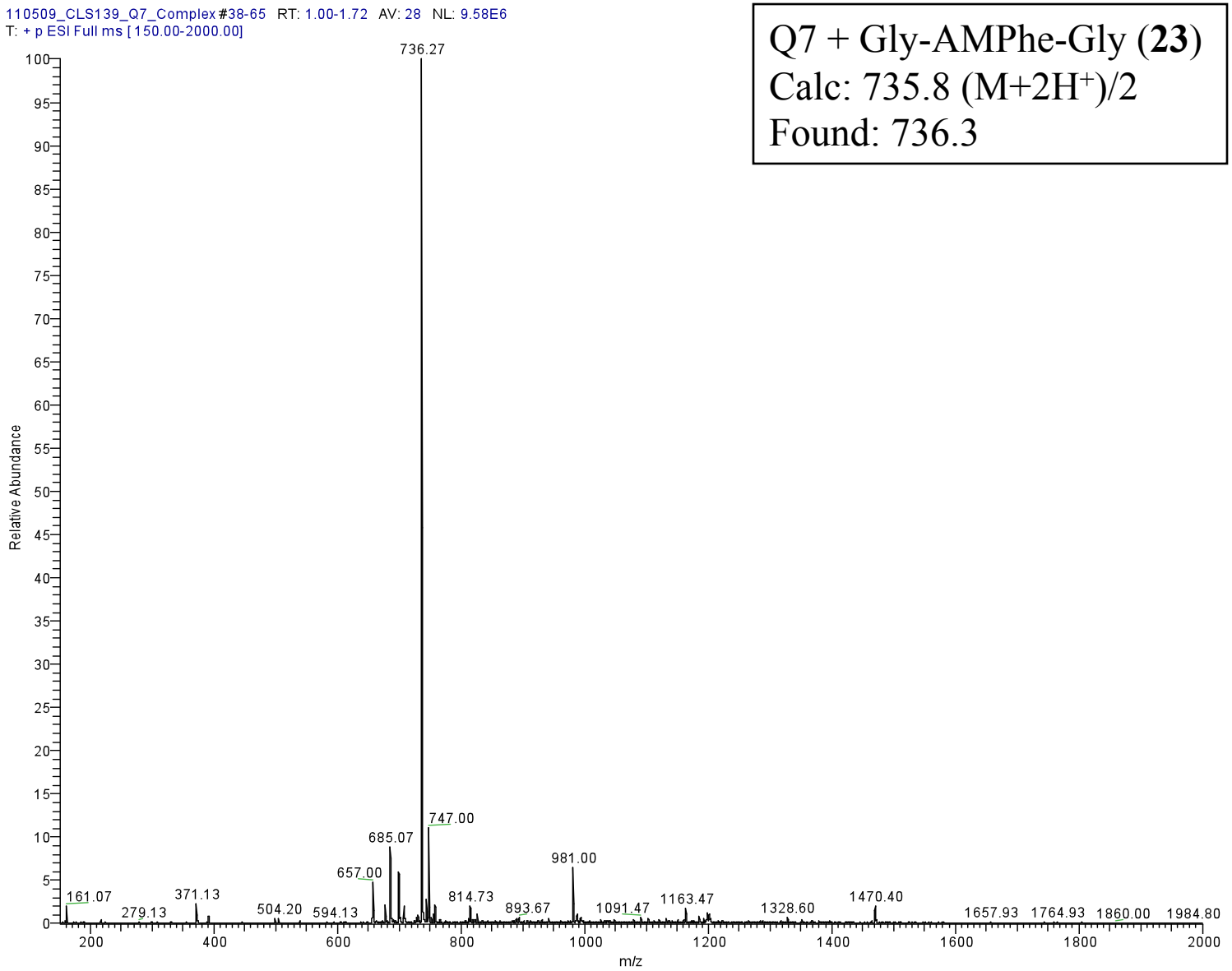
Q7 + AMPhe-Gly-Gly (**22**)
Calc: 735.8 $(M+2H^+)/2$
Found: 736.3

Figure S27. ESI mass spectrum of Q7 + **22** obtained in positive ion mode.



Gly-AMPhe-Gly (**23**)
Calc: 308.2 ($M+H^+$)
Found: 308.3

Figure S28. ESI mass spectrum of **23** obtained in positive ion mode.



Q7 + Gly-AMPhe-Gly (**23**)
Calc: 735.8 $(M+2H^+)/2$
Found: 736.3

Figure S29. ESI mass spectrum of Q7 + **23** obtained in positive ion mode.

Table S1. Chemical Shifts (δ) for Derivatized Phenylalanine Sidechains.

Guest	δ aromatic no Q7 (ppm)[a]	δ aromatic with Q7 (ppm)[a]	δ substituent[b] no Q7 (ppm)[a]	δ substituent[b] with Q7 (ppm)[a]
tBuPhe (**2**)	7.54, 7.32	7.19, 6.49	1.33	0.59
AMPhe (**3**)	7.47, 7.39	6.69, 6.58	4.20	3.84
tBuPhe-Gly-Gly (**21**)	7.53, 7.29	7.19, 6.50	1.32	0.60
AMPhe-Gly-Gly (**22**)	7.47, 7.38	6.61, 6.59	4.21	3.89
Gly-AMPhe-Gly (**23**)	7.40, 7.37	6.52	4.19	3.83

[a] Limiting chemical shifts of unbound ("no Q7") and bound ("with Q7") hydrogens on the modified phenylalanine sidechains. All species were at 2 mM concentration in a buffer of 10 mM sodium phosphate, pH 7.4. Spectra were acquired at 25 °C on a 500 MHz instrument. [b] t-butyl or aminomethyl substituent.

Competitive Equilibrium Analysis

The competitive equilibrium analysis used for AMFGG (**22**) is derived here. The competitive equilibrium is described by two binary equilibria

$$H{\bullet}G \underset{}{\overset{K_G}{\rightleftharpoons}} H + G + C \overset{K_C}{\rightleftharpoons} H{\bullet}C$$

where the symbols are defined as

H = host

G = low-affinity guest

C = high-affinity competitor

and the binary equilibrium association constants are defined as

$$K_G = \frac{[H \bullet G]}{[H][G]}$$

$$K_C = \frac{[H \bullet C]}{[H][C]}$$

The following analysis requires that the experimental conditions satisfy two assumptions:

1. There is negligible unbound host.
2. The free guest concentration is approximately equal to the total guest.

Both of these assumptions can be satisfied under the condition:

$[G]_O >> [H]_O$ (we use 100-fold excess)

and therefore we can assume

$[G] \approx [G]_O$

The dissociation of the weak-binding guest can be combined with the association of strong-binding competitor to obtain

$$\frac{K_C}{K_G} = \frac{[H \bullet C][G]}{[H \bullet G][C]}$$

which, based on our assumptions, can be rewritten as

$$\frac{K_C}{K_G} = \frac{[H \bullet C][G]_O}{[H \bullet G][C]}$$

Rearranging this expression, we define the observed equilibrium constant, K_{obs}

$$K_{obs} = \frac{[H \bullet C]}{[H \bullet G][C]} = \frac{K_C}{K_G[G]_O}$$

This expression is analogous to a simple binary equilibrium association reaction based on the assumption of negligible free host. Therefore, the association constant for the high-affinity competitor

$$K_C = K_{obs} K_G [G]_O$$

can be derived from knowledge of the total guest concentration, the affinity of the weak-binding guest (determined by prior experimentation), and measurement of the observed competitive equilibrium constant.

Here we discuss further the experimental conditions in the case of isothermal titration calorimetry (ITC). Fitting ITC data to a simple binary equilibrium expression is ideal when

$$10 < K_{obs}[cell] < 400$$

where [*cell*] is the concentration of interacting species in the sample cell (e.g., the host). We recommend placing a mixture of the weak-binding guest and the host in the sample cell in a 100:1 ($[G]_O$:$[H]_O$) ratio. Our ITC model (VP-ITC from Microcal) has a ratio of the volume of cell:syringe of 5:1, and thus in a typical binary equilibrium titration we use a syringe concentration that is 10-12 times that of the cell.

To satisfy all of these conditions, the weak-binding guest must be chosen carefully so that K_{obs}, i.e., the ratio

$$\frac{K_C}{K_G[G]_O}$$

will satisfy the condition

$$10 < K_{obs}[H]_O < 400$$

with $[H]_O$ that gives adequate signal:noise in an ITC experiment (at least 5 μcal/injection), and a $[G]_O$ that is sufficiently soluble when $[G]_O = 100[H]_O$. In the experiment described in this manuscript, we placed in the sample cell [Q7] = 50 μM and [Phe] = 5 mM, and in the syringe [AMF-G-G] = 600 μM. This gave adequate signal:noise, and $K_{obs}[Q7]_O = 91$, which yields an ideal curvature for fitting the data to an equilibrium binding model, as seen in Figure 3 of the manuscript.

Enthalpy. The enthalpy of binding for Q7 with AMFGG (**22**) was determined from the same ITC experiment as described above. The experimental ITC conditions used to determine K_{obs} also give an observed binding enthalpy (ΔH_{obs}), which is the sum of the dissociation enthalpy of the weak-binding guest ($-\Delta H_G$) and the association enthalpy of the strong-binding competitor (ΔH_C) according to the sum of those reactions:

$$\begin{array}{rcll} H{\bullet}G & \rightleftharpoons & H + G & -\Delta H_G \\ H + C & \rightleftharpoons & H{\bullet}C & \Delta H_C \\ \hline H{\bullet}G + C & \rightleftharpoons & H{\bullet}C + G & \Delta H_{obs} \end{array}$$

and thus if the value of ΔH_G has been determined by prior experiment with the weak-binding guest, then we can obtain the desired value as

$$\Delta H_C = \Delta H_{obs} + \Delta H_G$$

Experimental Details

Materials. The following compounds of analytical purity grade unless otherwise indicated were used without purification: phenylalanine (**1**) and tyrosine (**13**) (Sigma); H-Phe(4-tBu)-OH, H-Phe(4-CH_2NH_2)-OH (**2**), H-Phe(4-CH_3)-OH (**3**), H-Phe(4-CF_3)-OH (**5**), H-H-Phe(4-NH_2)-OH (**6**), H-Phe(4-F)-OH (**7**), H-Phe(4-Cl)-OH (**8**), H-Phe(4-Br)-OH (**9**), H-Phe(4-I)-OH (**10**), H-Phe(4-NO_2)-OH (**11**), H-Tyr(Me)-OH (**12**), H-Tyr(Ac)-OH (**14**), H-Phe(4-N_3)-OH (**15**), H-Phe(4-CN)-OH (**16**), H-Tyr(H_2PO_3)-OH (**17**), H-Phe(4-CO_2H)-OH (**18**), fmoc-Phe(4-tBu)-OH, and fmoc-Phe(4-CH_2NHBoc)-OH (Chem Impex International); Rink amide MBHA resin, HBTU, Fmoc-Phe-OH, and Fmoc-Gly-OH (Peptides International); biotech-grade dimethyl formamide (DMF), diisopropylethylamine (DIEA), triisopropyl silane (TIS), trifluoroacetic acid (TFA), piperidine, anhydrous dichloromethane, and HPLC-grade acetonitrile (Aldrich). All amino acids were the L enantiomer. Cucurbit[7]uril was synthesized by the group of Dr. Anthony Day and purchased from Unisearch. Nuclease-free ultra-pure water (US Biochemicals) was used for all analytical experiments.

A stock solution of 1.0 M sodium phosphate buffer was adjusted to pH 7.0 and sterile filtered. All ITC and fluorescence experiments were performed in 10 mM sodium phosphate buffer at pH 7.0, which was made by diluting the 1.0 M stock and adjusting to pH 7.0. The pH was checked periodically. All analytes were massed to ±0.02 mg with an accuracy of at least three significant digits. The concentration of Q7, accounting for salts and waters of recrystallization, was determined by titration (ITC) with a known sample of methyl viologen, which was determined spectroscopically (ε_{257} = 20,400 $M^{-1}cm^{-1}$). The

NMR solvent was prepared separately by dissolving a premixed PBS buffer packet (Sigma) in deuterium oxide.

Synthesis. The synthesis of tripeptides **19-23** was achieved by standard fmoc solid-phase methods using an automatic peptide synthesizer (CS Bio) on Rink amide MBHA resin. Each synthesis was carried out on the scale of 0.30 g resin (0.156 meq/g). Fmoc deprotection was accomplished by shaking the resin in a solution of 20% piperidine in DMF at room temp for 20 min, followed by extensive rinsing. For each peptide coupling, a mixture of fmoc-protected amino acid, HBTU and DIEA (4 equiv. each) was dissolved in 10 mL DMF, mixed for 15 min, and added to the resin. Couplings proceeded at room temp for 60 min followed by extensive rinsing. After the final fmoc deprotection, the resin was transferred to a fritted flask and rinsed with dichloromethane, and then the peptide was cleaved from the resin by treatment with two portions of 10 mL of a mixture of 95% TFA, 2.5% TIS, and 2.5% water for one hour each at room temperature with shaking. The mixture was then filtered, rinsed with 5 mL of the TFA mixture, and the filtrate collected and concentrated *in vacuo* to dryness. The crude peptide was then dissolved in 30 mL 0.1% aqueous TFA and purified in three batches by reverse phase HPLC (Waters Delta 600 Semi-Prep system with a Waters Nova-Pak C18 column) using a 1%/min gradient of acetonitrile in 0.1% aqueous TFA. Fractions were checked by analytical HPLC (Agilent 1100 system with Waters Nova-Pak C18 column), and pure fractions were combined and lyophilized to dryness.

H-Phe-Gly-Gly-NH_2 (19). ^{1}H NMR (500 MHz, 10 mM phosphate buffer D_2O, pH 7.4): δ = 7.43 (2H, t, *J* = 6.5 Hz), 7.38 (1H, t, *J* = 7.0 Hz), 7.32 (2H, d, *J* = 7.0 Hz), 4.11 (1H, t, *J* = 7.0 Hz), 3.90 (4H, m), 3.15 (2H, m).

H-Gly-Phe-Gly-NH_2 (20). ^{1}H NMR (500 MHz, 10 mM phosphate buffer D_2O, pH 7.4): δ = 7.41 (2H, t, *J* = 7.5 Hz), 7.35 (1H, t, *J* = 7.5 Hz), 7.30 (2H, d, *J* = 7.0 Hz), 4.68 (1H, t, *J* = 8.0 Hz), 3.80 (4H, m), 3.15 (2H, m).

H-Phe(4-tBu)-Gly-Gly-NH_2 (21). ^{1}H NMR (500 MHz, 10 mM phosphate buffer D_2O, pH 7.4): δ = 7.53 (2H, d, *J* = 8.0 Hz), 7.29 (2H, d, *J* = 8.5 Hz), 4.15 (1H, t, *J* = 7.5 Hz), 3.90 (4H, m), 3.15 (2H, d, *J* = 7.0 Hz), 1.32 (9H, s).

H-Phe(4-CH_2NH_2)-Gly-Gly-NH_2 (22). ^{1}H NMR (500 MHz, 10 mM phosphate buffer D_2O, pH 7.4): δ = 7.47 (2H, d, *J* = 8.0 Hz), 7.38 (2H, d, *J* = 8.5 Hz), 4.21 (2H, s), 4.12 (1H, t, *J* = 7.5 Hz), 3.98 (4H, m), 3.18 (2H, d, *J* = 7.0 Hz).

H-Gly-Phe(4-CH_2NH_2)-Gly-NH_2 (23). ^{1}H NMR (500 MHz, 10 mM phosphate buffer D_2O, pH 7.4): δ = 7.40 (2H, d, *J* = 8.0 Hz), 7.37 (2H, d, *J* = 7.5 Hz), 4.67 (1H, t, *J* = 7.5 Hz), 4.19 (2H, s), 3.77 (4H, m), 3.14 (2H, m).

Fluorescence Spectroscopy. Fluorescence emission intensities were measured at room temperature using a Tecan Infinite 200 Pro fluorescence plate reader with an excitation wavelength 485 nm (bandwidth 9 nm) and emission wavelength of 515 nm

(bandwidth 20 nm) with a constant amount of gain, signal averaging, integration time, and focal depth. All samples were dissolved in 10 mM sodium phosphate, pH 7.0. Analyte samples ($I_{analyte}$) contained 20 μM analyte (e.g., phenylalanine), 2 μM Q7, and 2 μM acridine orange (AO). The control for maximum fluorescence (I_{max}) contained 2 μM Q7 and 2 μM AO. The control for minimum fluorescence (I_{min}) contained 2 μM AO. Percent quenching was calculated as:

$$\%Quenching = \frac{I_{analyte} - I_{\min}}{I_{\max} - I_{\min}}$$

Experimental uncertainties were plotted (in Figure 2) as the standard deviation from three experiments.

Isothermal Titration Calorimetry (ITC). Titration experiments were carried out in 10 mM sodium phosphate buffer (pH 7.0) at 27 ºC on a VP-ITC calorimeter from Microcal Inc. In a typical experiment, Q7 was in the sample cell at a concentration of 0.01 – 0.1 mM, and the guest was in the injection pipet at a concentration of 0.1 – 1.0 mM. The titration schedule consisted of 28-40 consecutive injection of 2-10 μL with at least a 200 s interval between injections. Heats of dilution were subtracted from each data set. All solutions were degassed prior to titration. The data were analyzed using Origin software and fit well to the "one-set-of-sites" binding model supplied with the software. This is a standard binary equilibrium (1+1) model that assumes all sites are independent and non-interacting.

NMR Spectroscopy. ^{1}H and ^{13}C NMR spectra were collected in deuterium oxide solution containing 10 mM sodium phosphate, pH 7.4, on a Varian 500 MHz spectrometer at 25 °C using a presaturation pulse to suppress the signal from residual protiated solvent.

Mass Spectrometry. Data were acquired using a Thermo Deca XP Plus spectrometer with an electrospray source in positive ion mode. All samples were dissolved in pure water at a concentration of 10-100 μM.