Supporting Information

Discovery of an H3K36me3-derived peptidomimetic ligand with enhanced affinity for plant homeodomain finger protein 1 (PHF1)

Isabelle A. Engelberg,[†,#] Jiuyang Liu,[‡,#] Jacqueline L. Norris-Drouin,[†] Stephanie H. Cholensky,[†,§] Samantha A. Ottavi,[†] Stephen V. Frye,[†] Tatiana G. Kutateladze,[*,‡] Lindsey I. James[*,†]

[†]Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

[‡]Department of Pharmacology, University of Colorado School of Medicine, Aurora, CO, 80045, USA

# These authors contributed equally.

Table of Contents

Supplementary Figures

Figure S1, Engelberg and Liu et al.

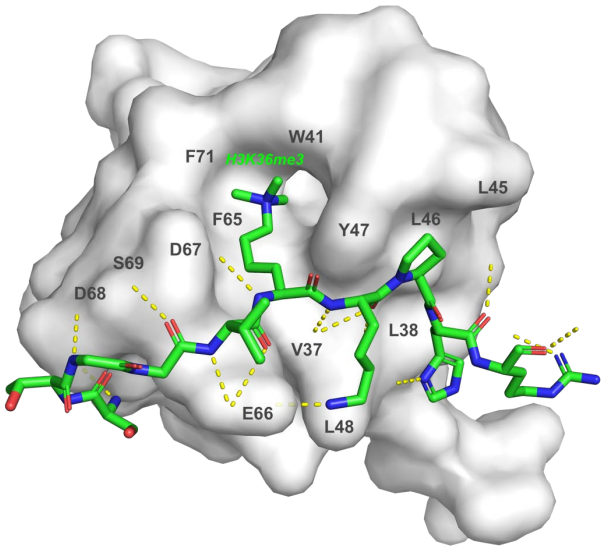


Figure S1. Structure of H3K36me3 bound to PHF1 Tudor.

H3K36me3 binds PHF1 Tudor in a surface-grove binding mode. Residues T32-R40 shown make polar contacts with the surface of the protein (PDB ID: 4HCZ), contributing to the micromolar binding affinity.

Figure S2, Engelberg and Liu et al.

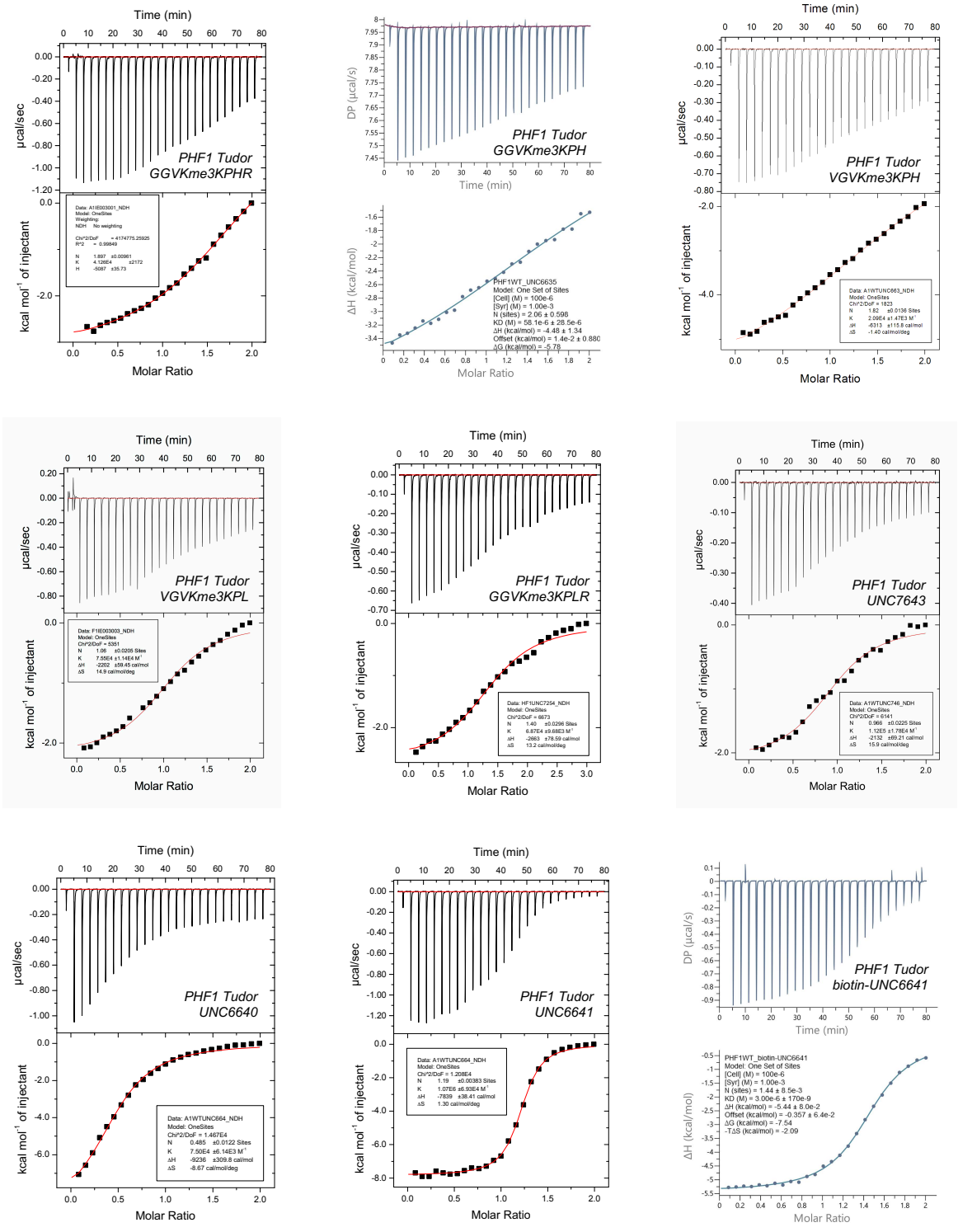


Figure S2. Representative ITC curves for PHF1 and PHF19 Tudor domain binding to H3K36me3 variants and select peptidomimetic ligands.

Figure S3, Engelberg and Liu et al.

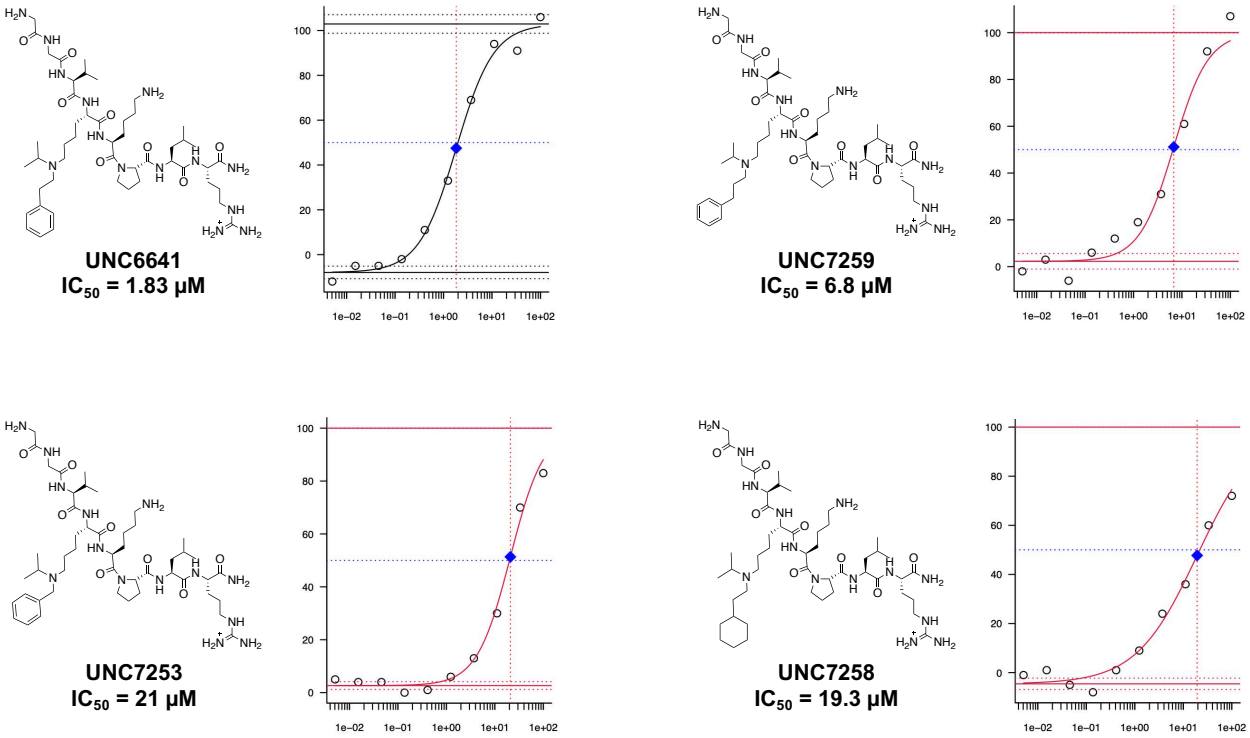


Figure S3. Representative TR-FRET curves for PHF1 Tudor domain binding to UNC6641 and closely related derivatives.

Figure S4, Engelberg and Liu et al.

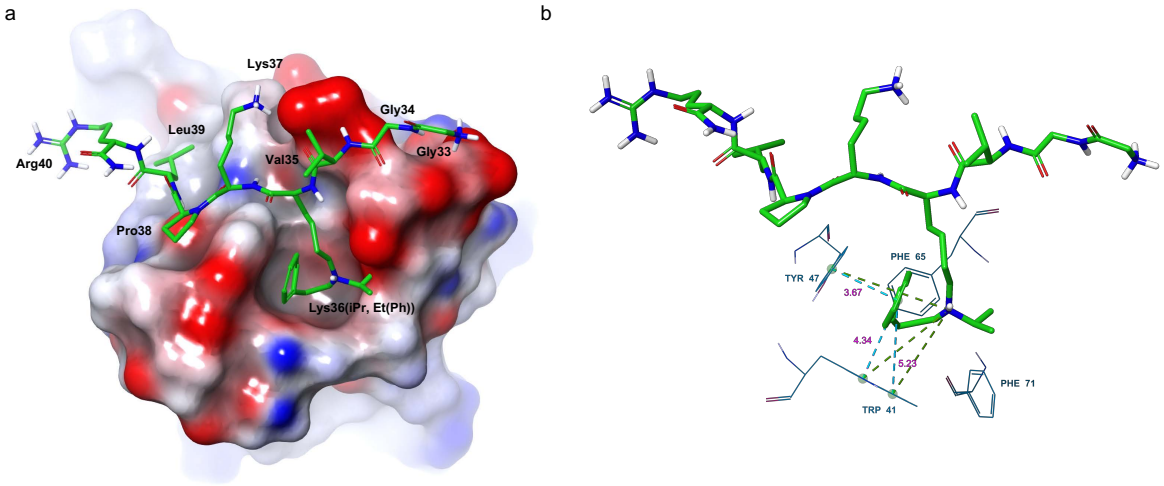


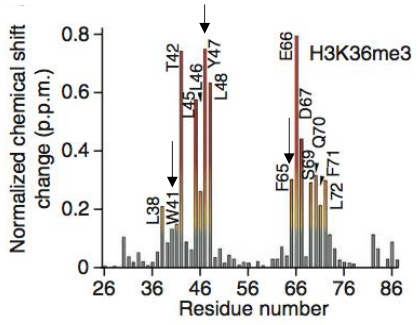
Figure S4. Molecular docking analysis of UNC6641 binding to PHF1 Tudor.

(A) Molecular docking of UNC6641 into PHF1 Tudor (PDB ID: 4HCZ) using H3K36me3 bound to PHF1 Tudor as backbone constraints.

(B) Using centroid-to-centroid distance and angle parameters defined by Maestro, PHF1 Y47 and W41 were predicted to engage the phenyl group of UNC6641 via π-π stacking (dashed blue lines). Our docking model suggests strong face-to-face π-π stacking with Y47 (< 4.4 Å centroid-to-centroid distance) and edge-to-face π-π stacking with W41 (< 5.5 Å centroid-to-centroid distance). All cation-π interactions with the ε-amino group (dashed green lines) remained within the suggested distance of 6.6 Å. As previously noted, the increased distance of F71 from the amino group may indicate a weaker role of cation-π interactions and a stronger contribution of hydrophobic interactions from this residue.

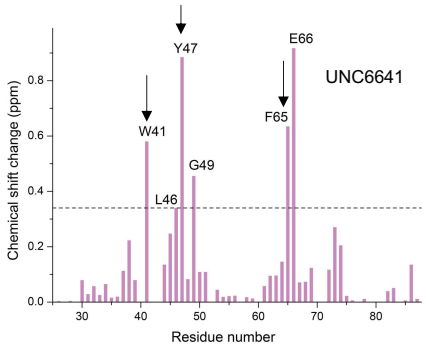
Figure S5, Engelberg and Liu et al.

a



PHF1 Residue	CSP (p.p.m.)
W41	0.15
Y47	0.75
F65	0.30
F71	0.20
E66	0.80

b



PHF1 Residue	CSP (p.p.m.)
W41	0.55
Y47	0.90
F65	0.65
F71	ND
E66	0.95

Figure S5. Comparison of chemical shift perturbations (CSPs) for H3K36me3 and UNC6641 upon binding to PHF1 Tudor.

(A) Histogram and quantification of CSPs for the H3K36me3 peptide upon binding to PHF1 Tudor. Histogram from: Musselman, C. A.; Avvakumov, N.; Watanabe, R.; Abraham, C. G.; Lalonde, M.-E.; Hong, Z.; Allen, C.; Roy, S.; Nuñez, J. K.; Nickoloff, J.; Kulesza, C. A.; Yasui, A.; Côté, J.; Kutateladze, T. G. Molecular Basis for H3K36me3 Recognition by the Tudor Domain of PHF1. *Nat. Struct. Mol. Biol.* **2012**, *19* (12), 1266–1272. https://doi.org/10.1038/nsmb.2435.

B) Histogram and quantification of CSPs for UNC6641 upon binding to PHF1 Tudor.

Figure S6, Engelberg and Liu et al.

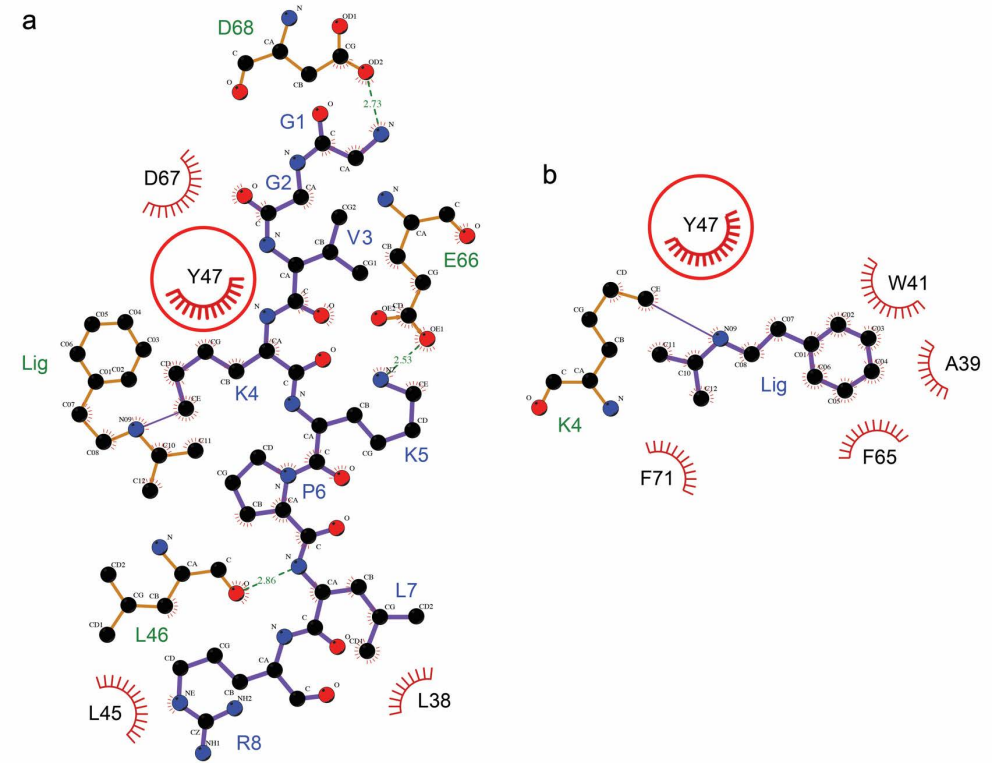


Figure S6. LigPlot interaction analysis of UNC6641 bound to PHF1 Tudor.

A) Analysis of interactions formed between UNC6641 (ball and stick) and residues of PHF1 (red circles) generated by LigPlot analysis software.

B) Closer view of the (isopropyl)phenethyl lysine substituent bound in the aromatic cage of PHF1.

Figure S7, Engelberg and Liu et al.

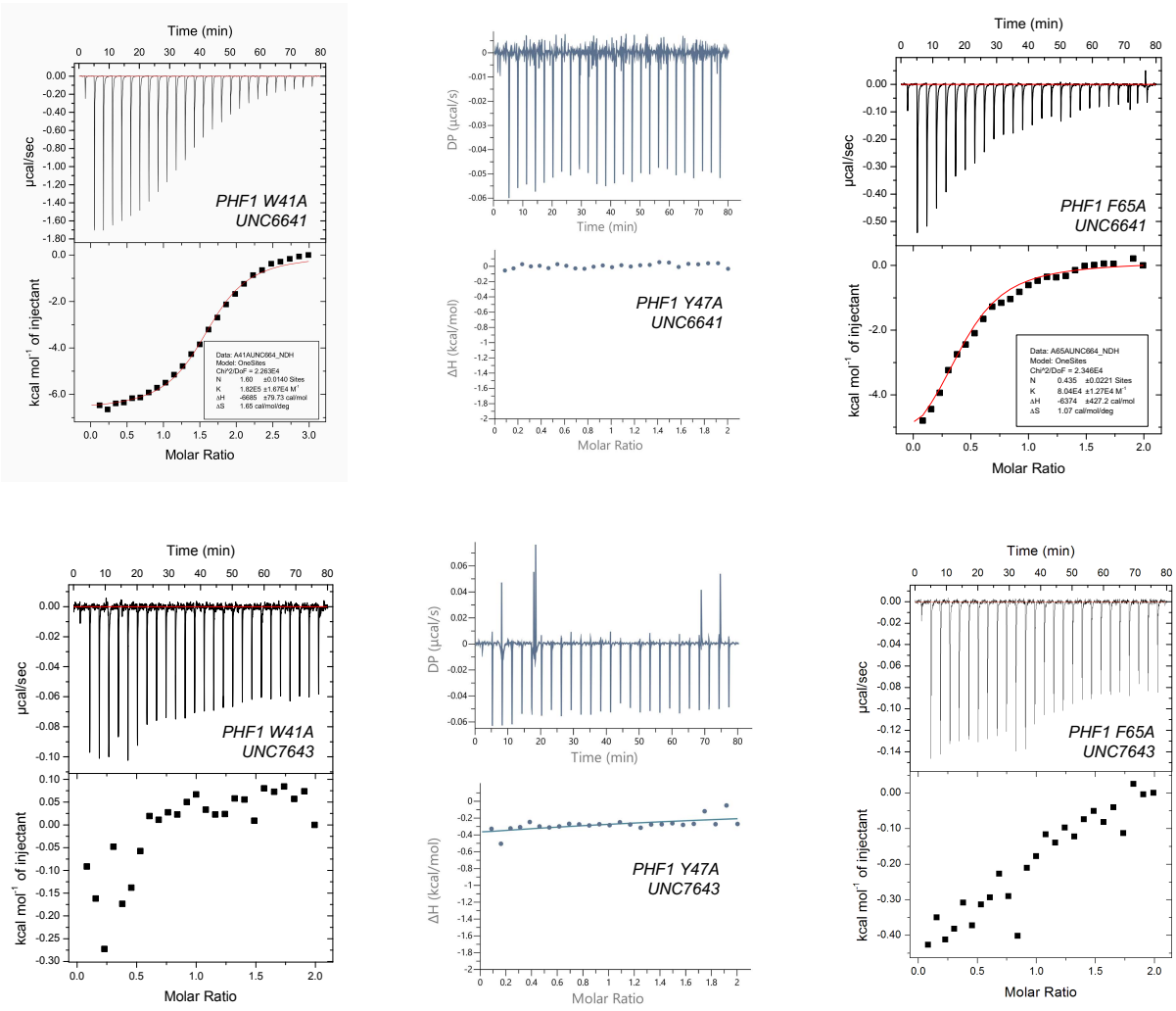


Figure S7. Representative ITC curves for binding of PHF1 aromatic cage mutants to UNC6641.

Figure S8, Engelberg and Liu et al.

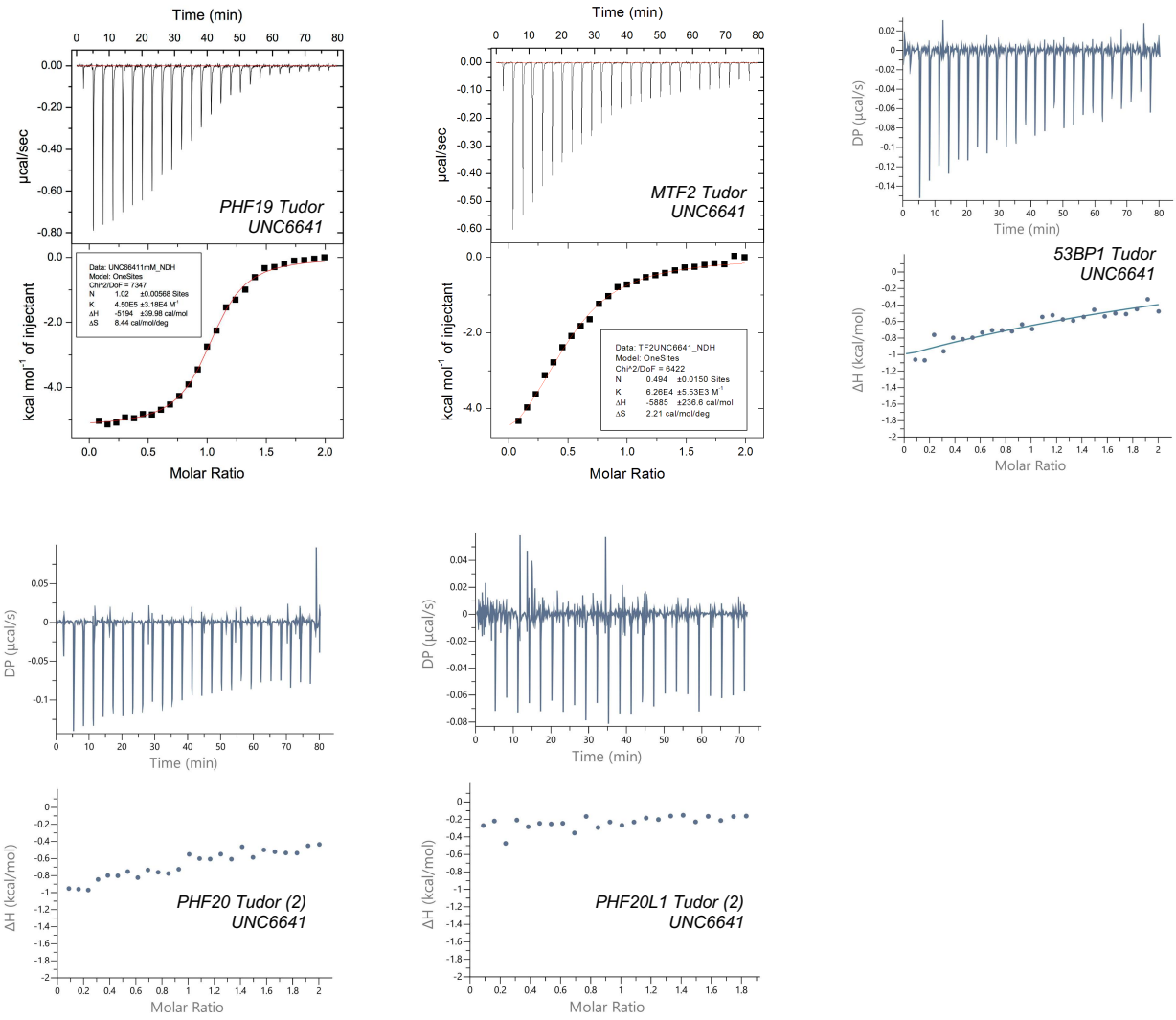


Figure S8. Representative ITC curves for binding of related Tudor domains to UNC6641.

Figure S9, Engelberg and Liu et al.

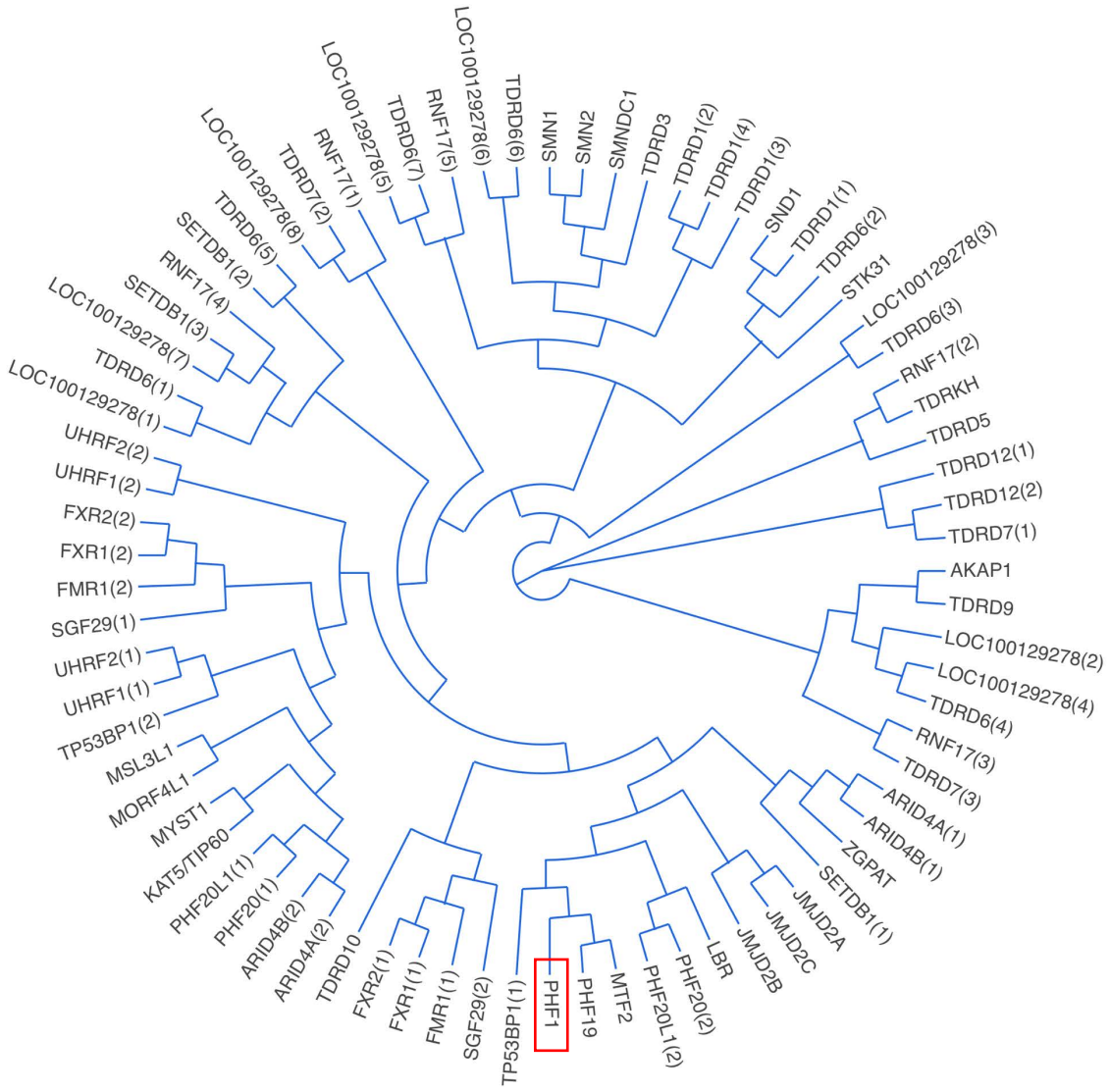


Figure S9. Phylogenetic tree of human Tudor domains.

Complete phylogenetic tree of Tudor domains from Liu L, Zhen XT, Denton E, Marsden BD, Schapira M., Bioinformatics (2012): **ChromoHub: a data hub for navigators of chromatin-mediated signaling**. PHF1 is outlined in red.

Figure S10, Engelberg and Liu et al.

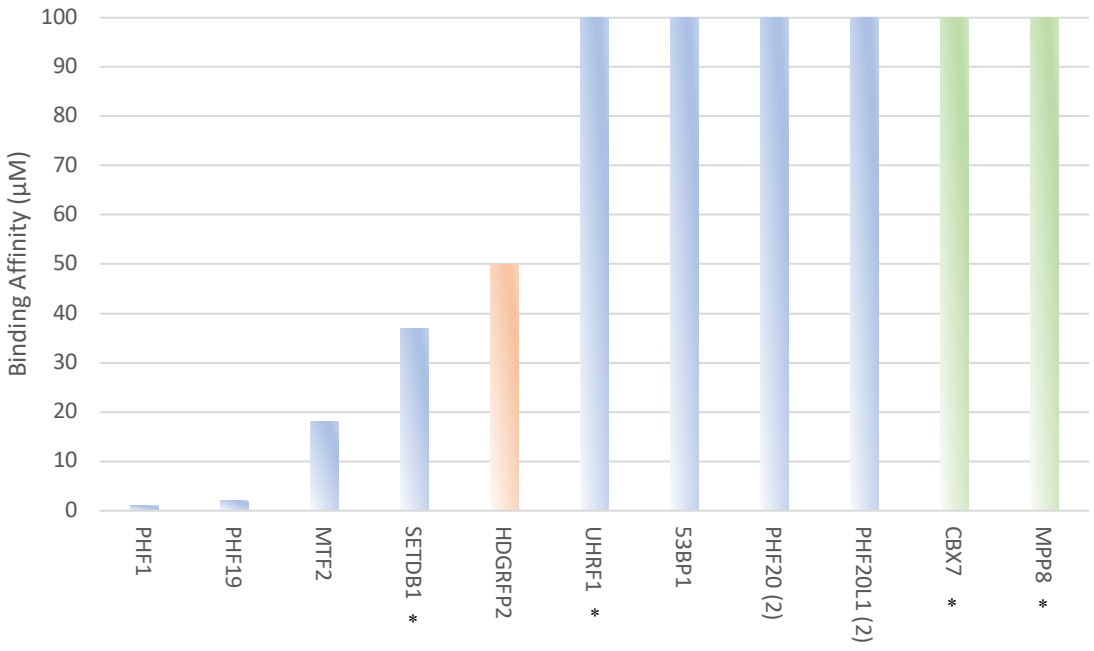


Figure S10. UNC6641 selectively binds PHF1 and PHF19 within the panel of Kme readers tested.

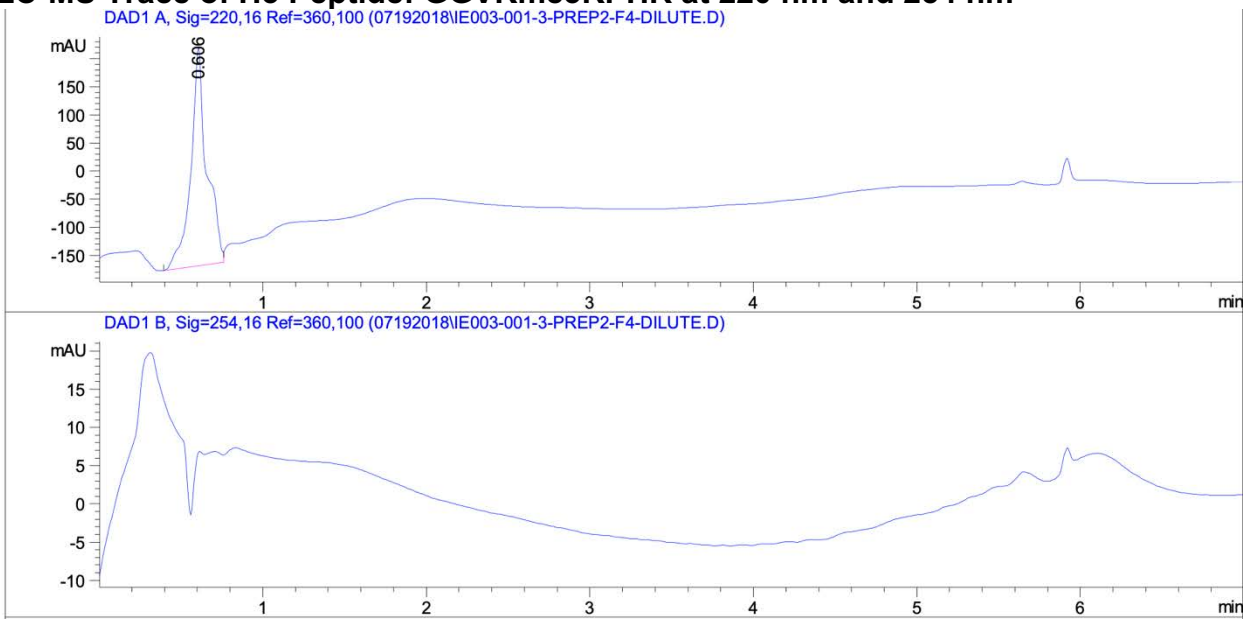
Binding affinities measured by ITC (K_d) or TR-FRET (IC_{50}, denoted by *) were calculated as an average of at least two independent experiments. Blue = Tudor domain. Orange = PWWP domain. Green = Chromodomain.

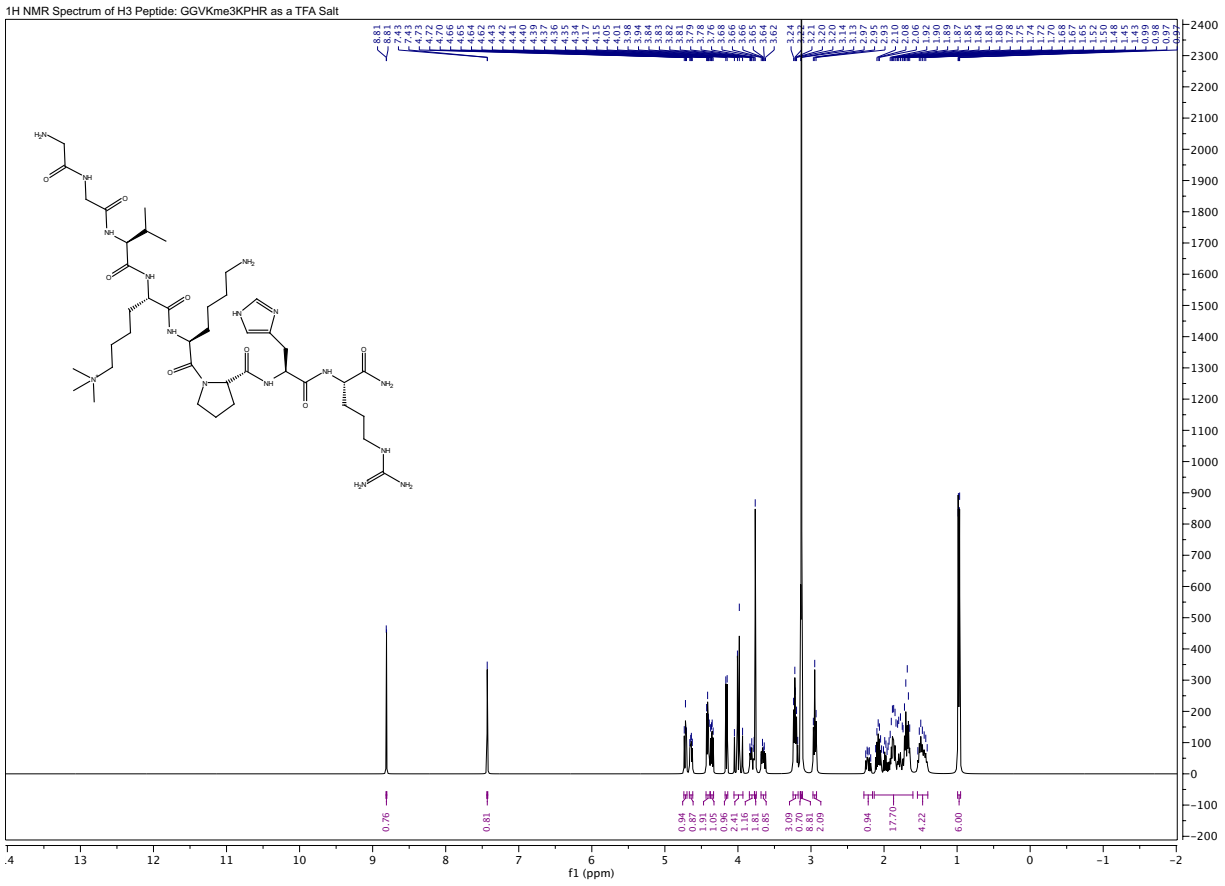
Supplementary Table 1. Data collection and refinement statistics.

Tudor of PHF1 in complex with UNC6641	
PDB ID	**7LKY**
Data Collection	
Space group	*P* 1
Cell dimensions	
a, b, c (Å)	43.39, 45.35, 64.44
α, β, γ (˚)	90.02, 95.19, 90.07
Wavelength(Å)	1.28
Resolution* (Å)	43.22-1.85 (1.916 -1.85)
Completeness (%)	90.87 (63.60)
Multiplicity	1.9 (1.9)
$CC_{1/2}$	0.997 (0.839)
CC*	0.999 (0.955)
R_{sym} or R_{merge} (%)	5.8 (33.4)
$I/\sigma I$	8.85 (1.85)
Refinement	
Resolution (Å)	43.22-1.85
Unique reflections	37960
R_{work}/R_{free}	23.3/28.1
Root-mean-square deviation	
Bondslengths (Å)	0.003
Bond angles (˚)	0.667
B-factors (Å^2)	
Protein	19.48
Ligand	24.56
Water	17.56
No. atoms	
Protein	3600
Ligand	483
Water	258
Ramachandran plot	
Favored/ allowed/ outlier(%)	95.95/4.05/0.00

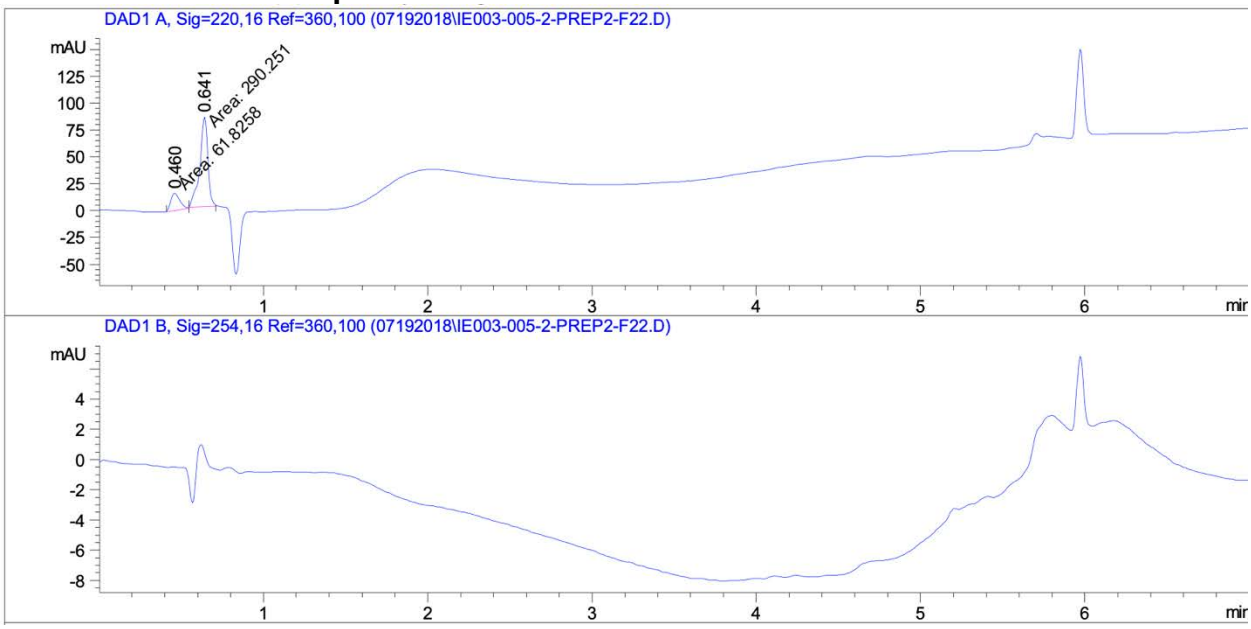
* Values in parentheses are for highest-resolution shell.

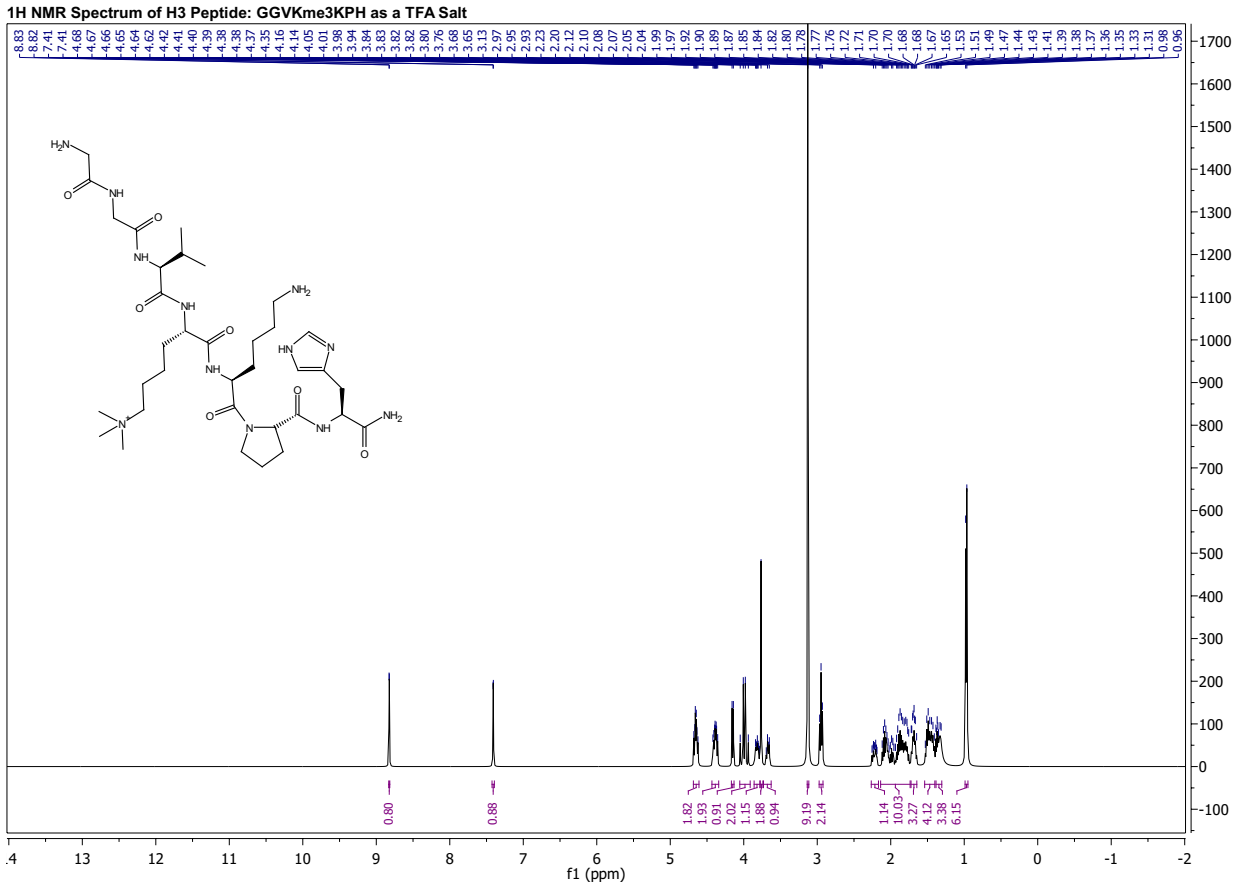
LC-MS Trace of H3 Peptide: GGVKme3KPHR at 220 nm and 254 nm



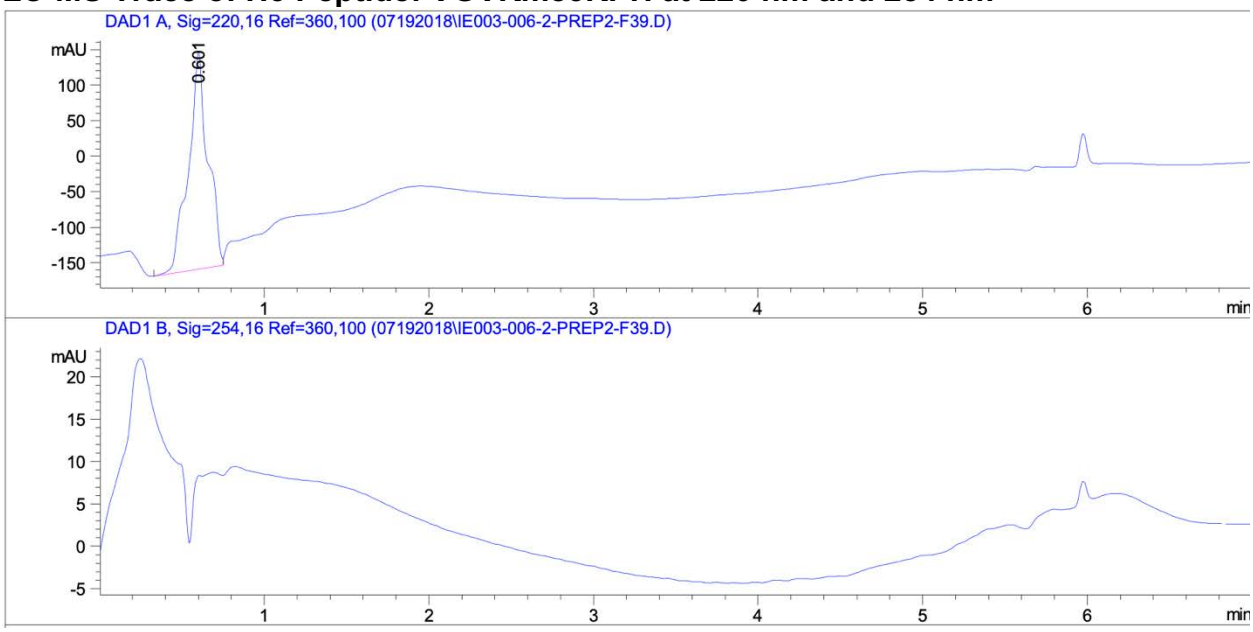


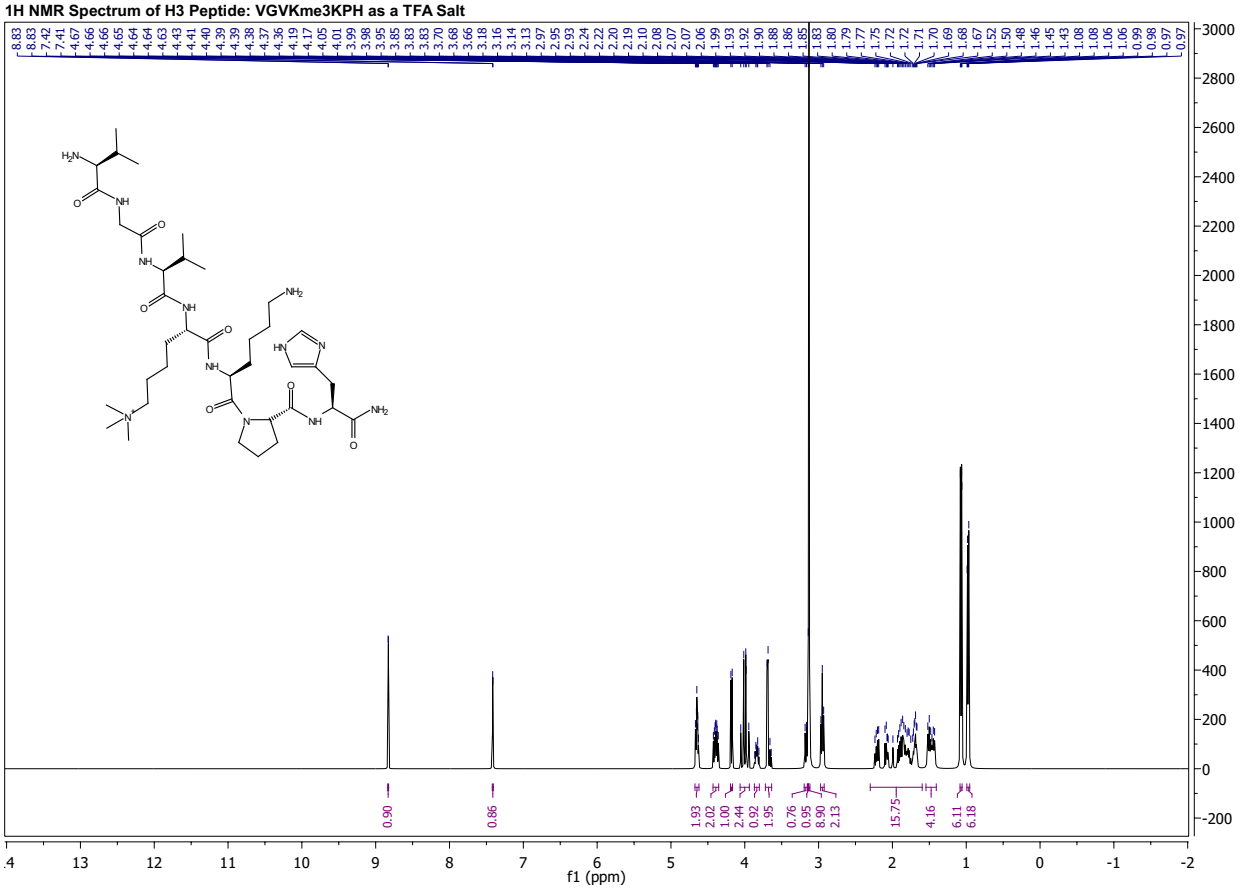
LC-MS Trace of H3 Peptide: GGVKme3KPH at 220 nm and 254 nm



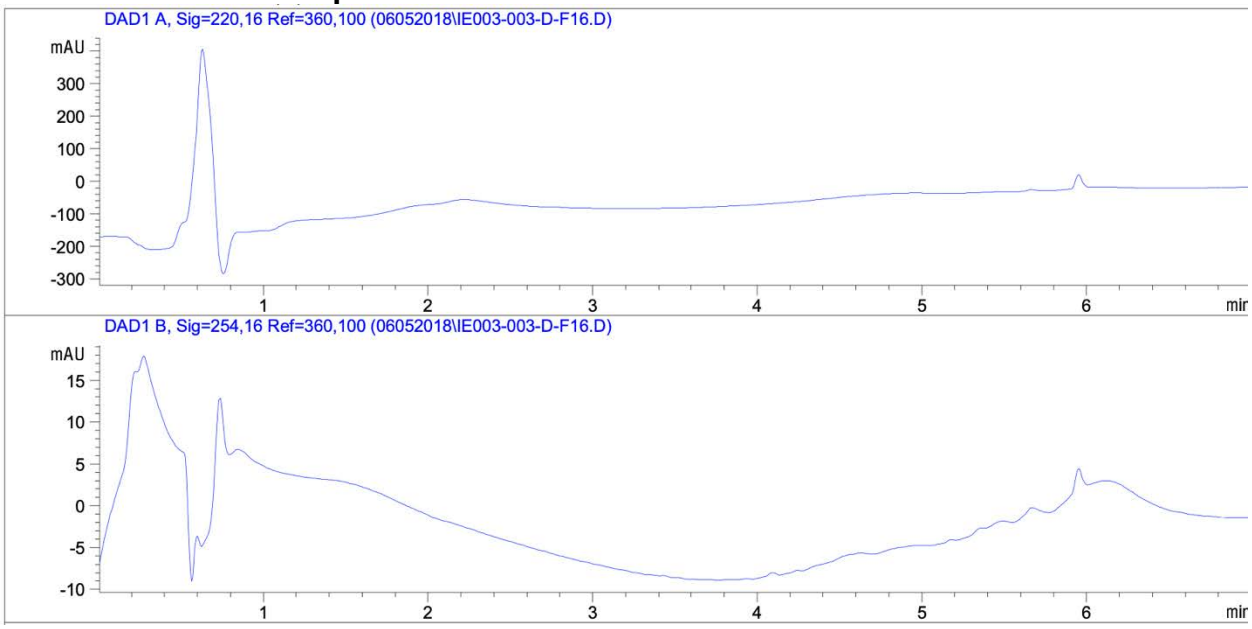


LC-MS Trace of H3 Peptide: VGVKme3KPH at 220 nm and 254 nm

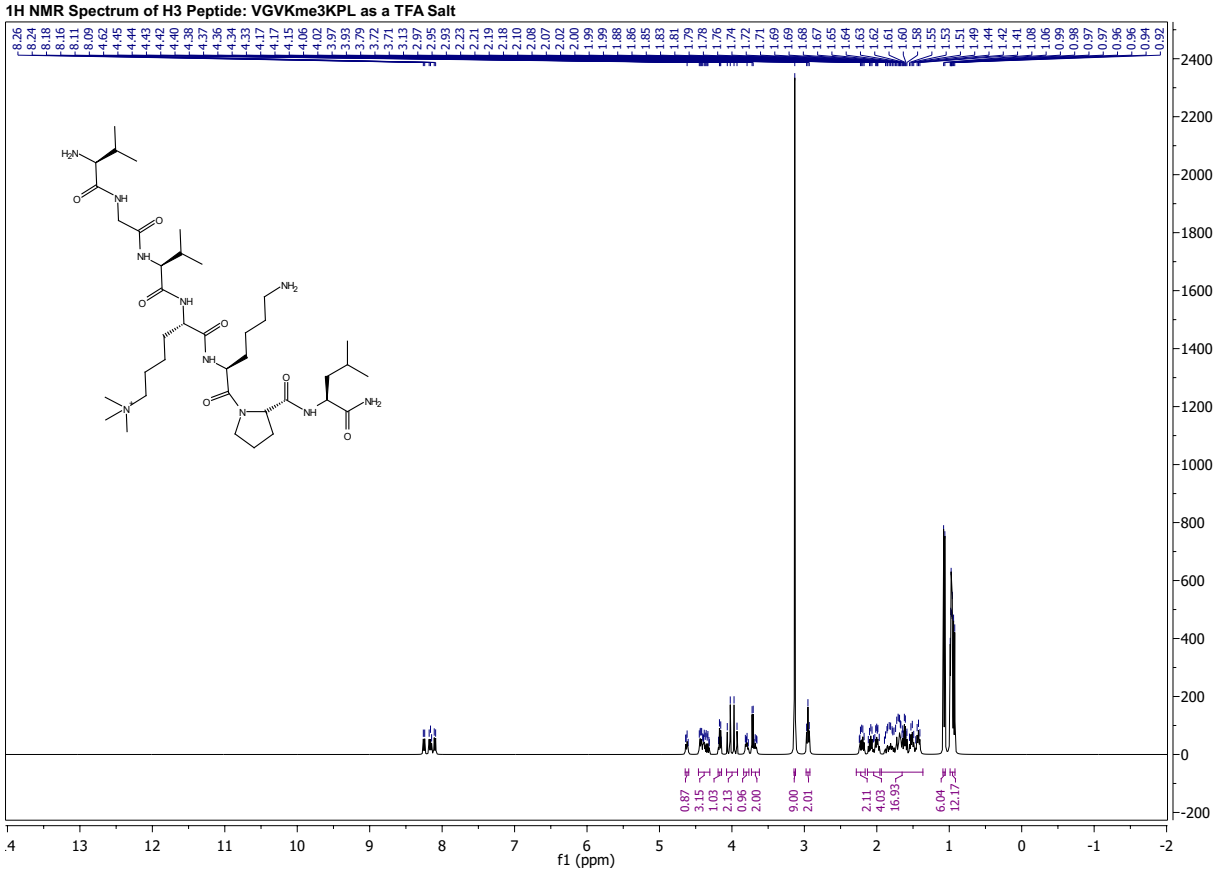




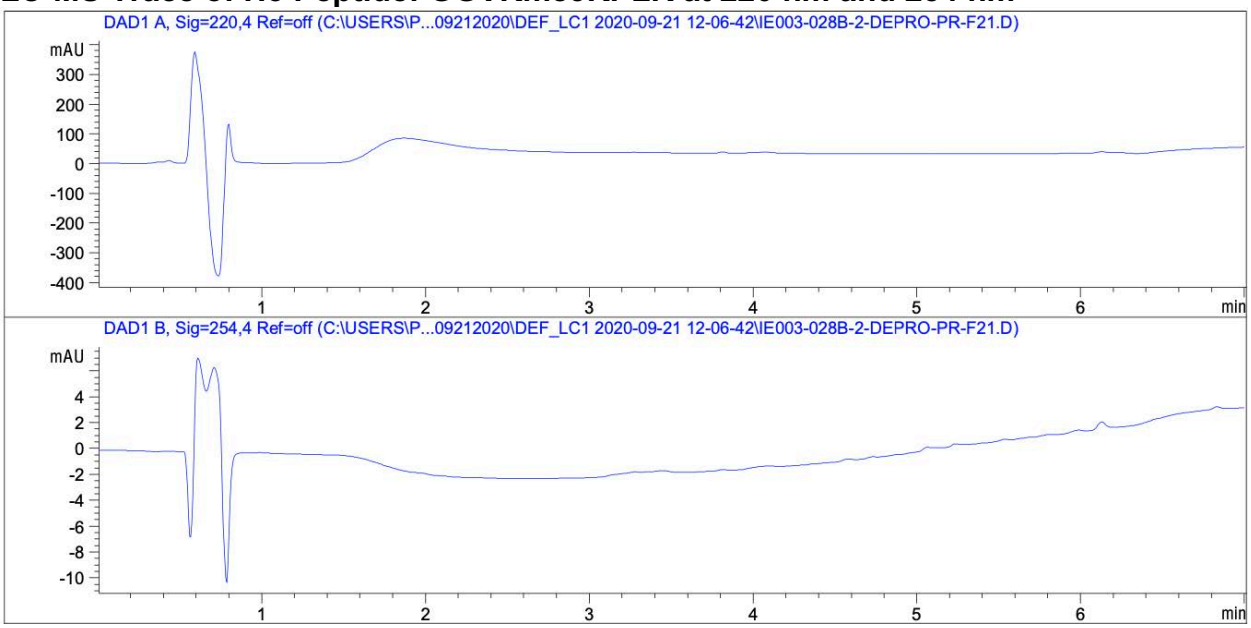
LC-MS Trace of H3 Peptide: VGVKme3KPL at 220 nm and 254 nm

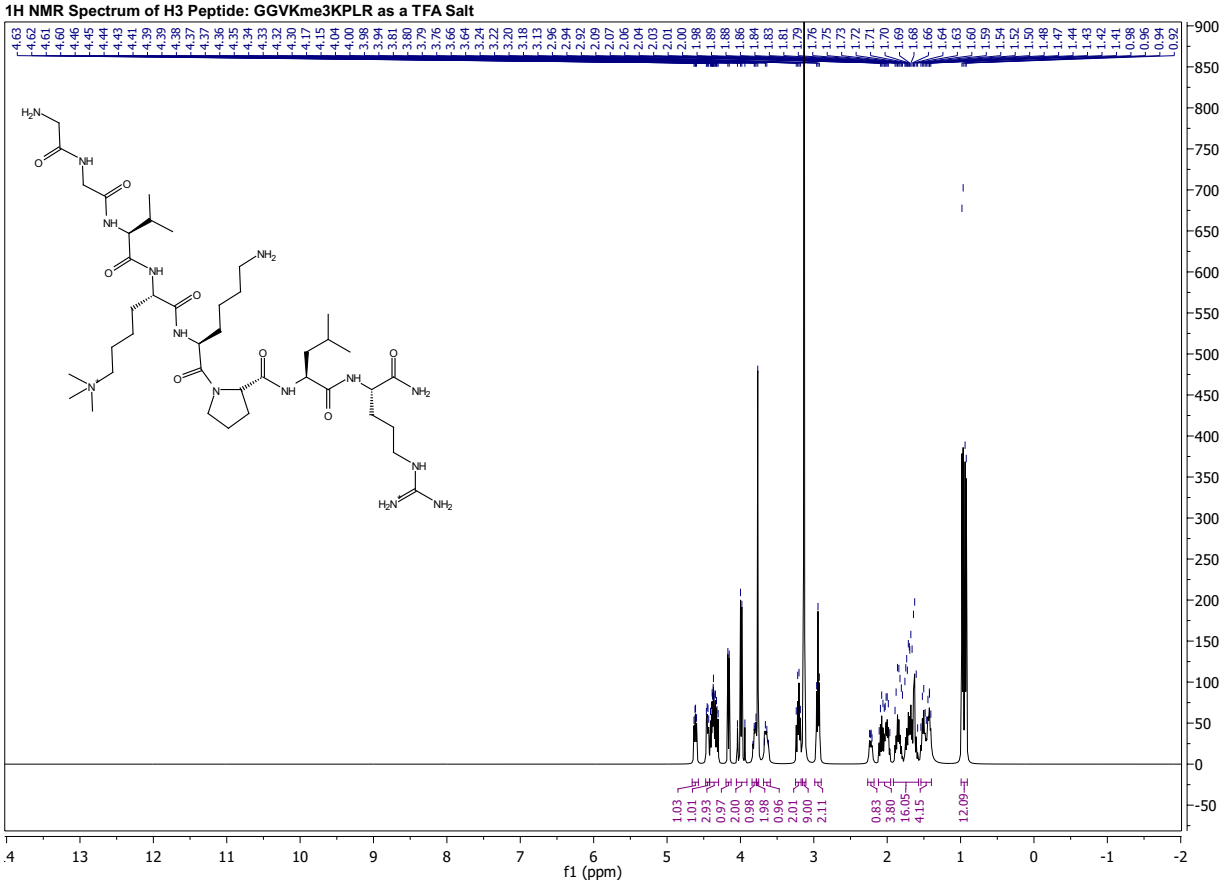


1H NMR Spectrum of H3 Peptide: VGVKme3KPL as a TFA Salt

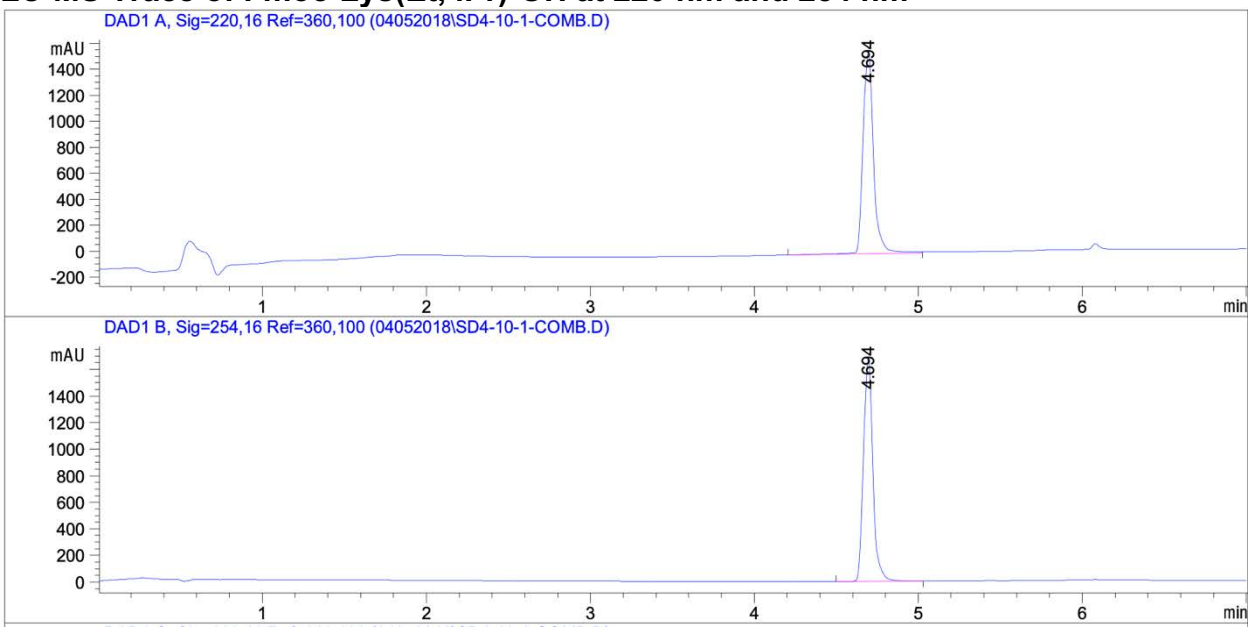


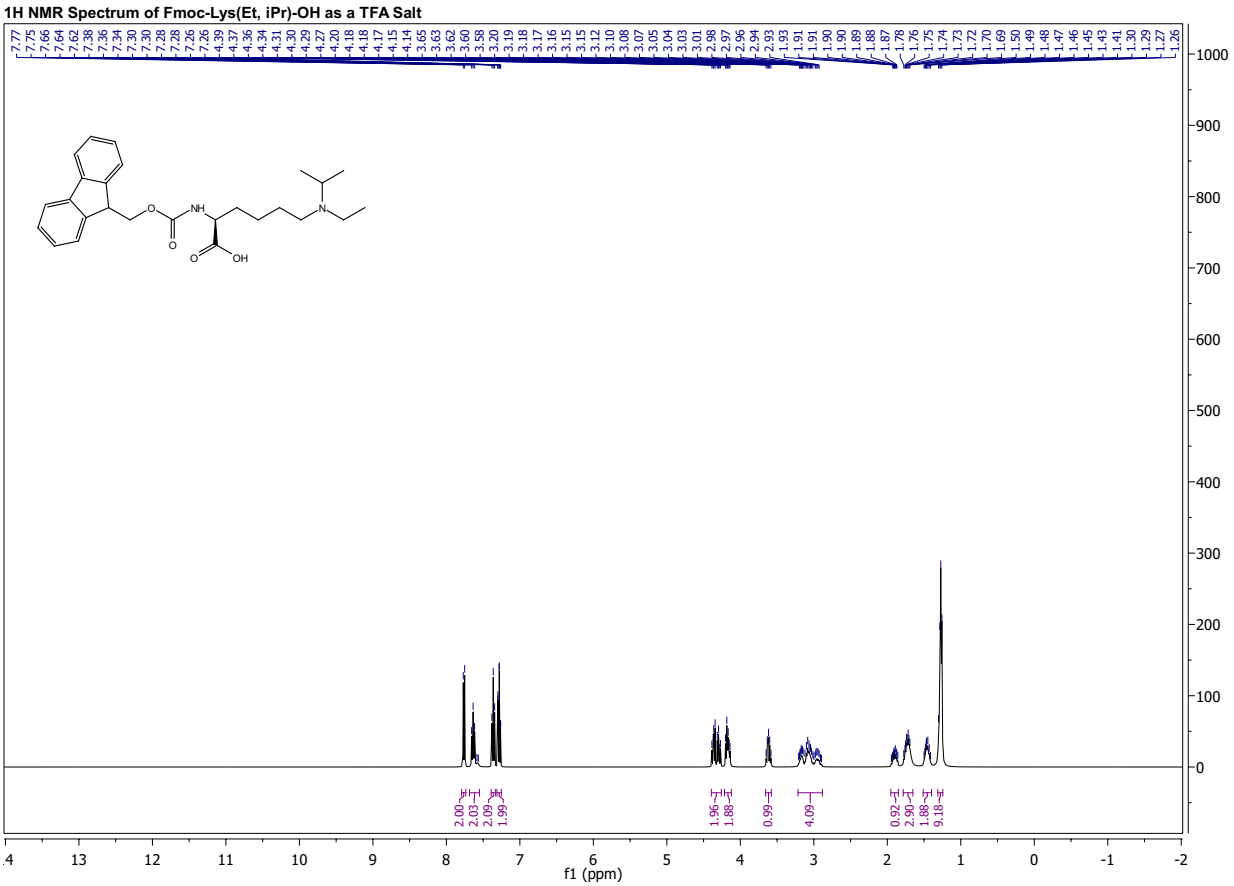
LC-MS Trace of H3 Peptide: GGVKme3KPLR at 220 nm and 254 nm



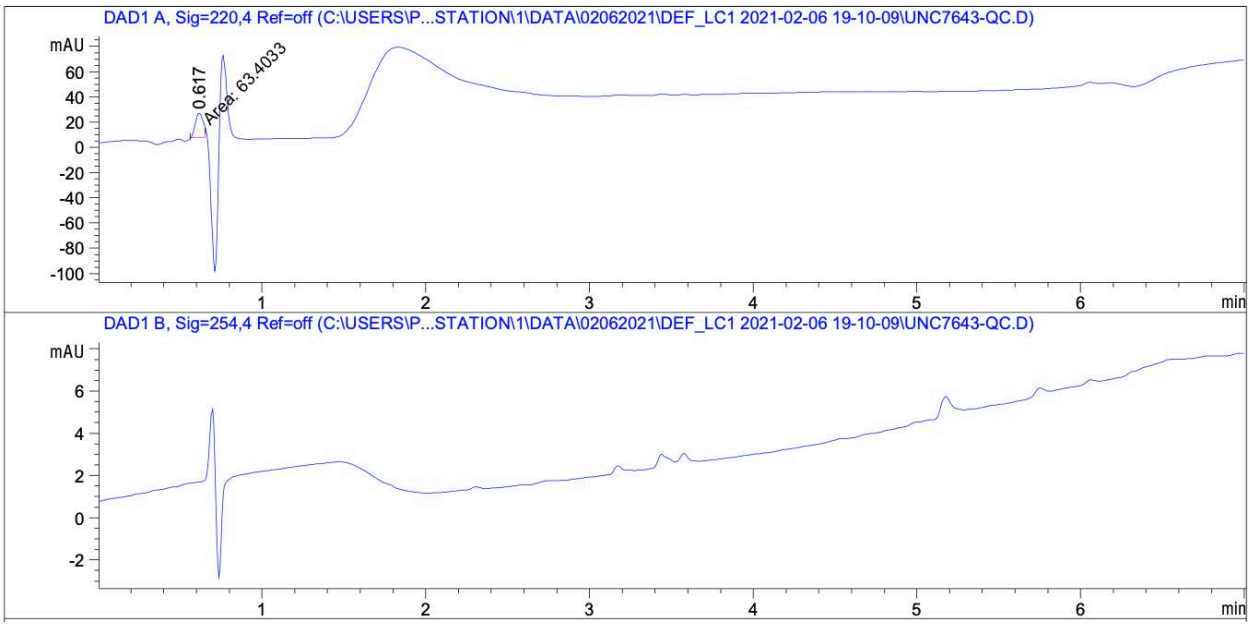


LC-MS Trace of Fmoc-Lys(Et, iPr)-OH at 220 nm and 254 nm

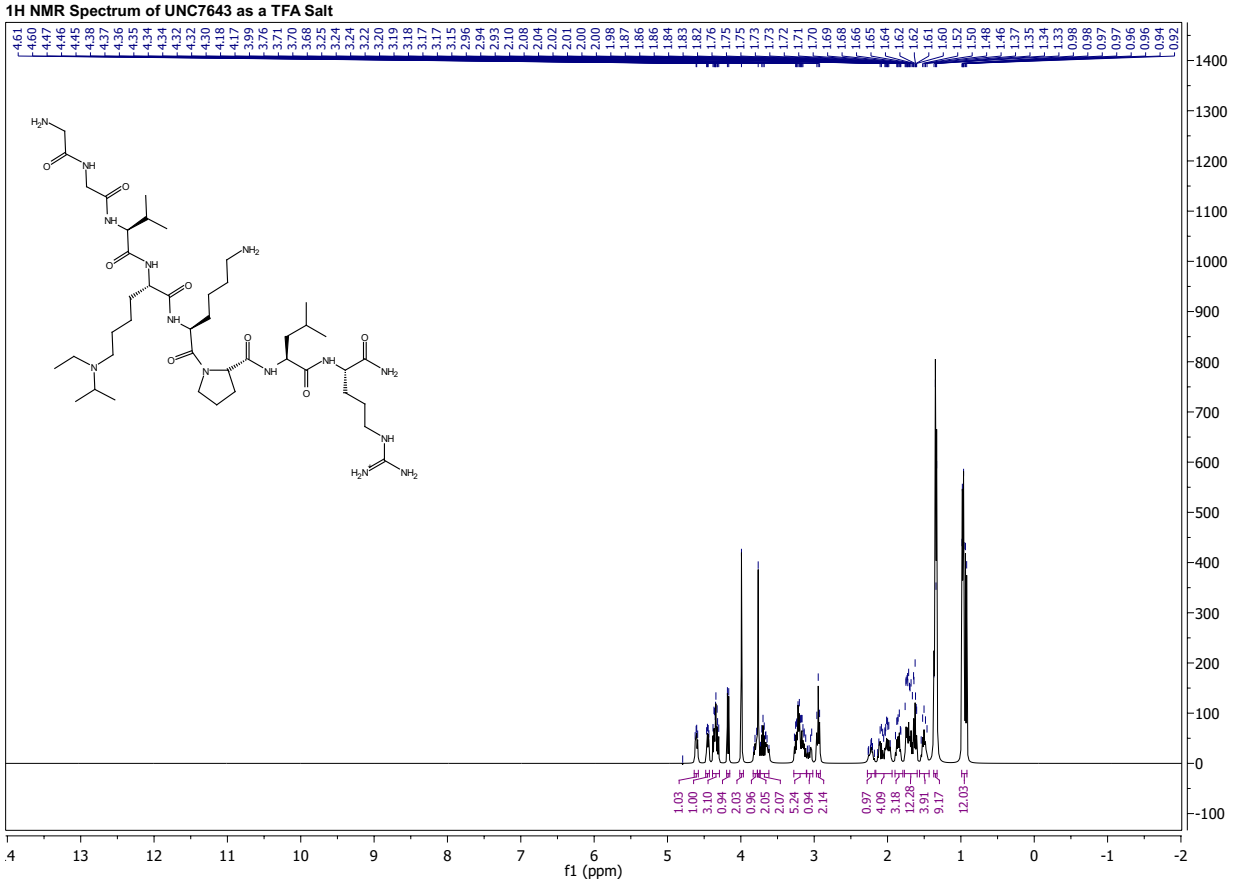




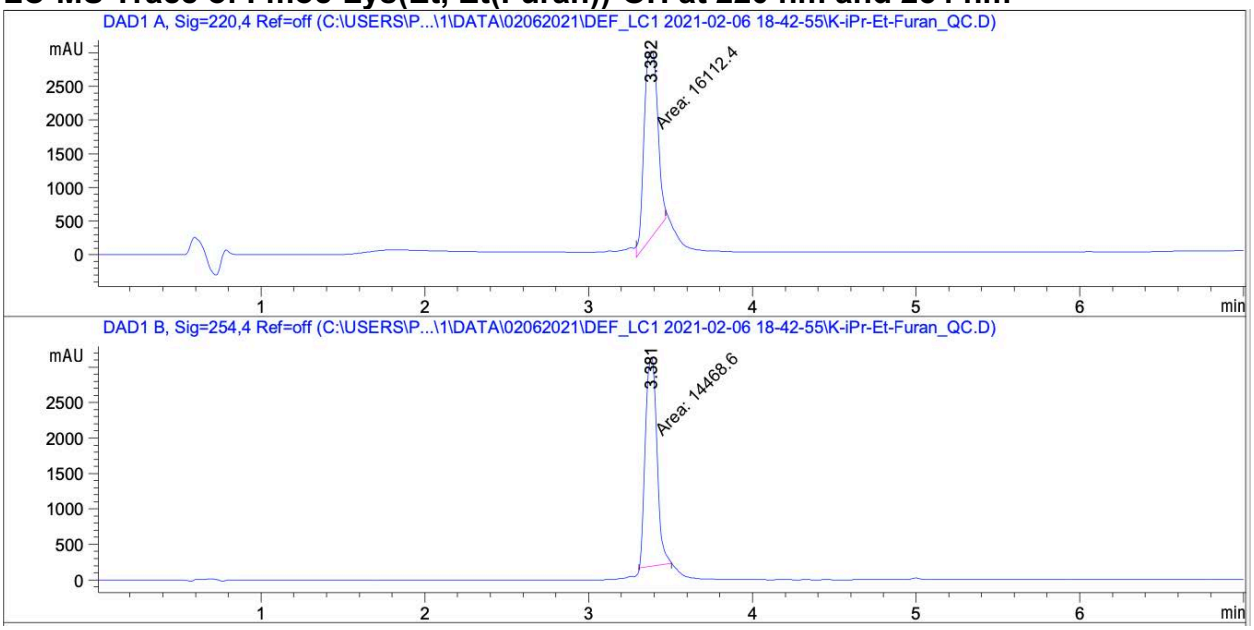
LC-MS Trace of UNC7643 at 220 nm and 254 nm

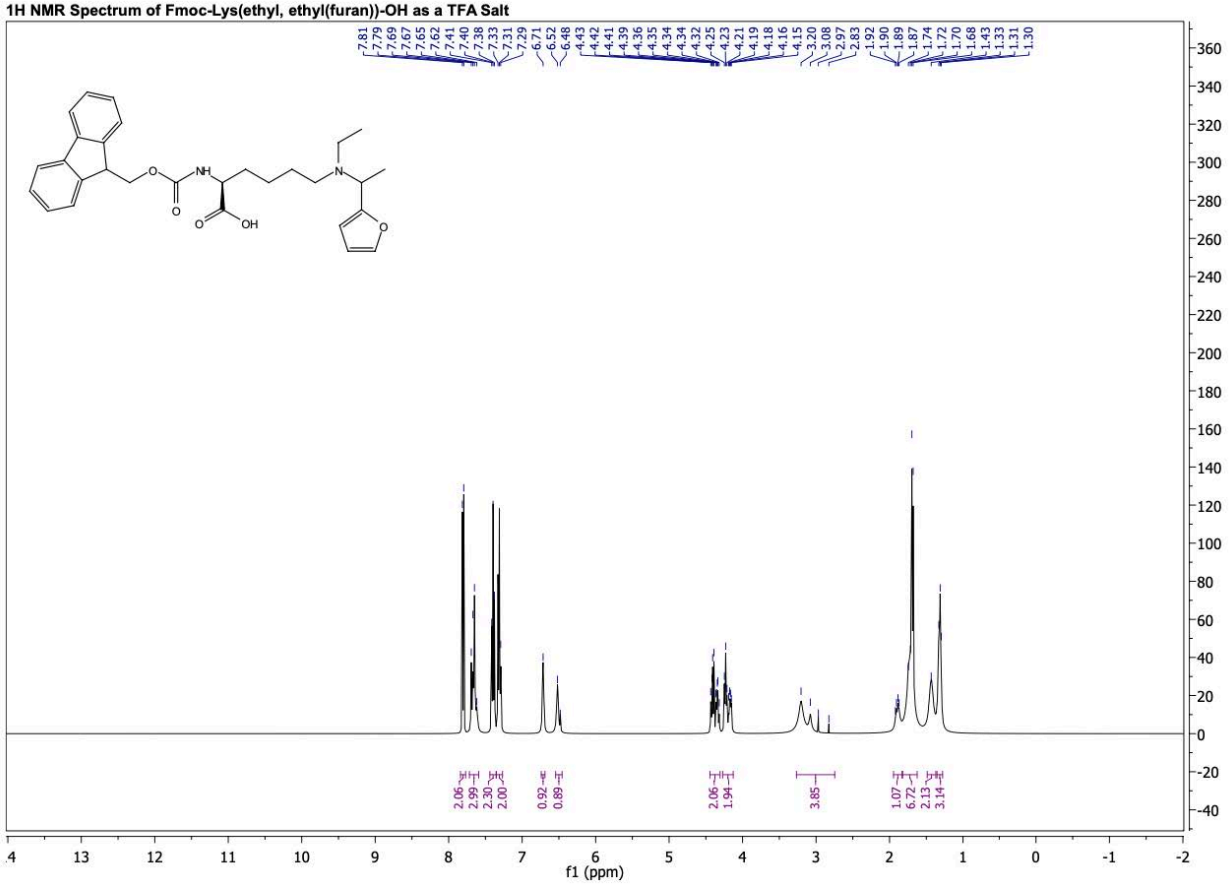


1H NMR Spectrum of UNC7643 as a TFA Salt

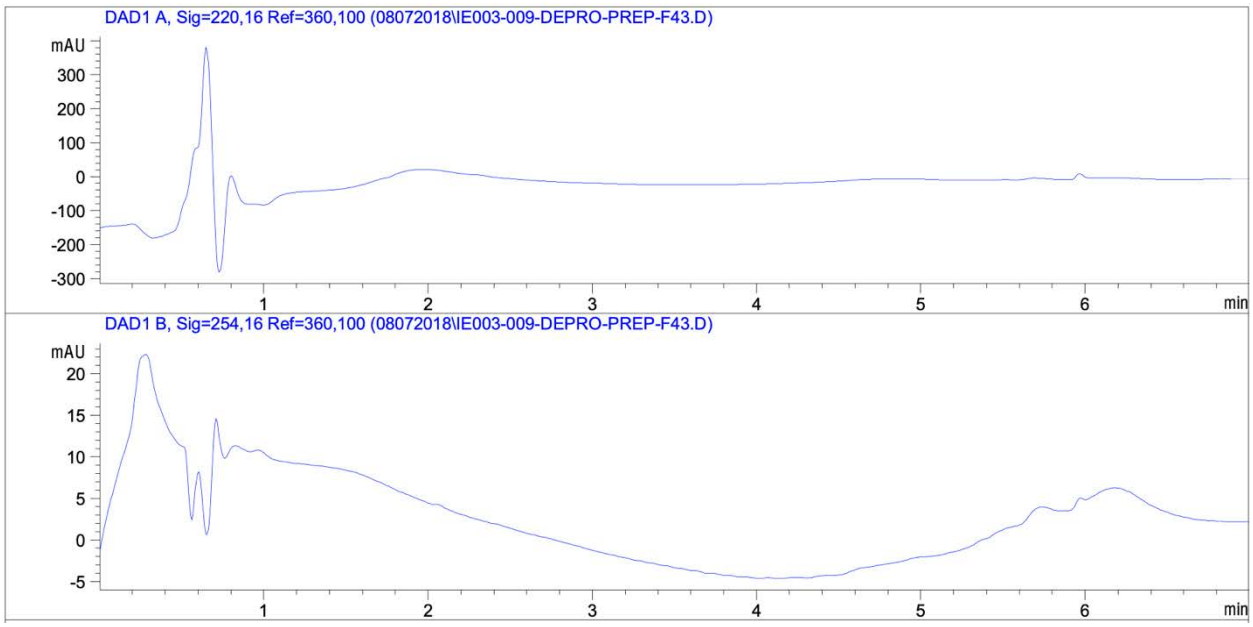


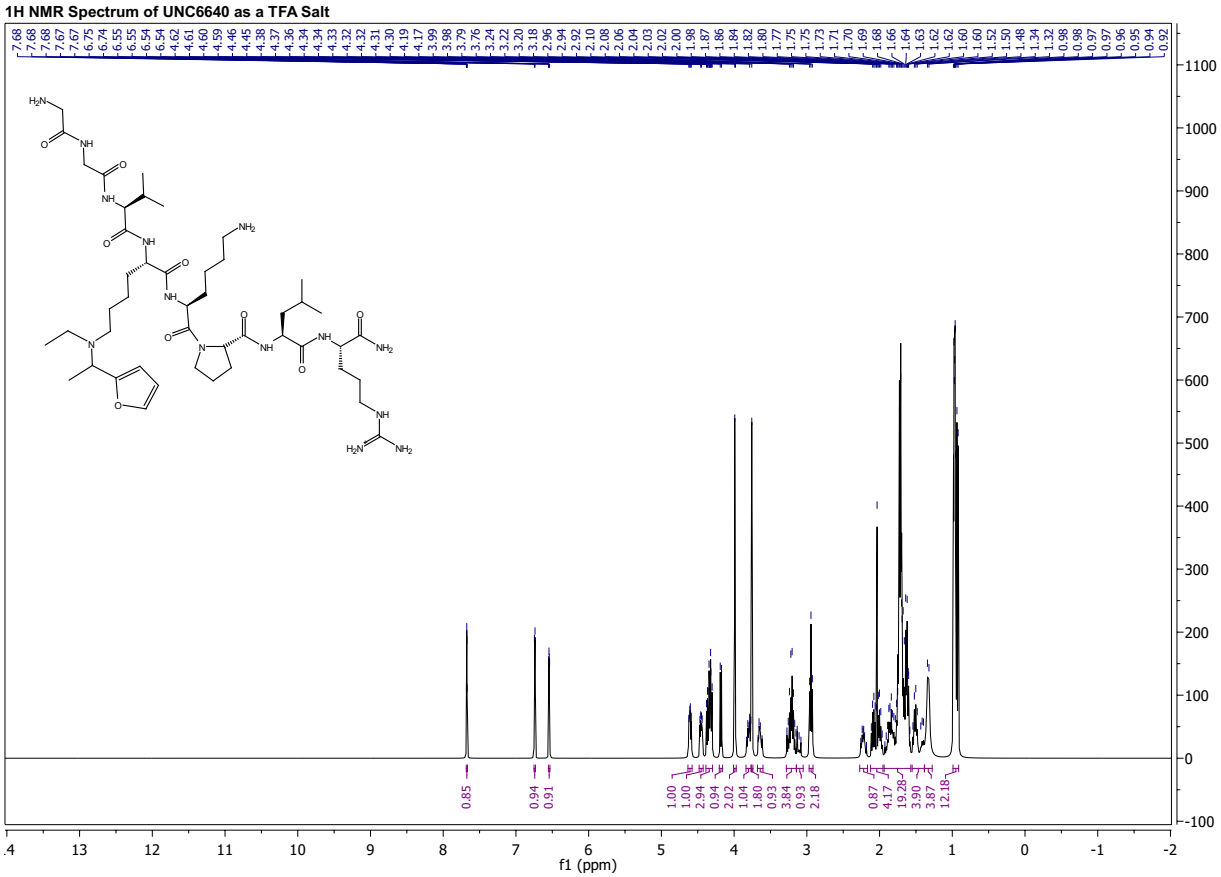
LC-MS Trace of Fmoc-Lys(Et, Et(Furan))-OH at 220 nm and 254 nm



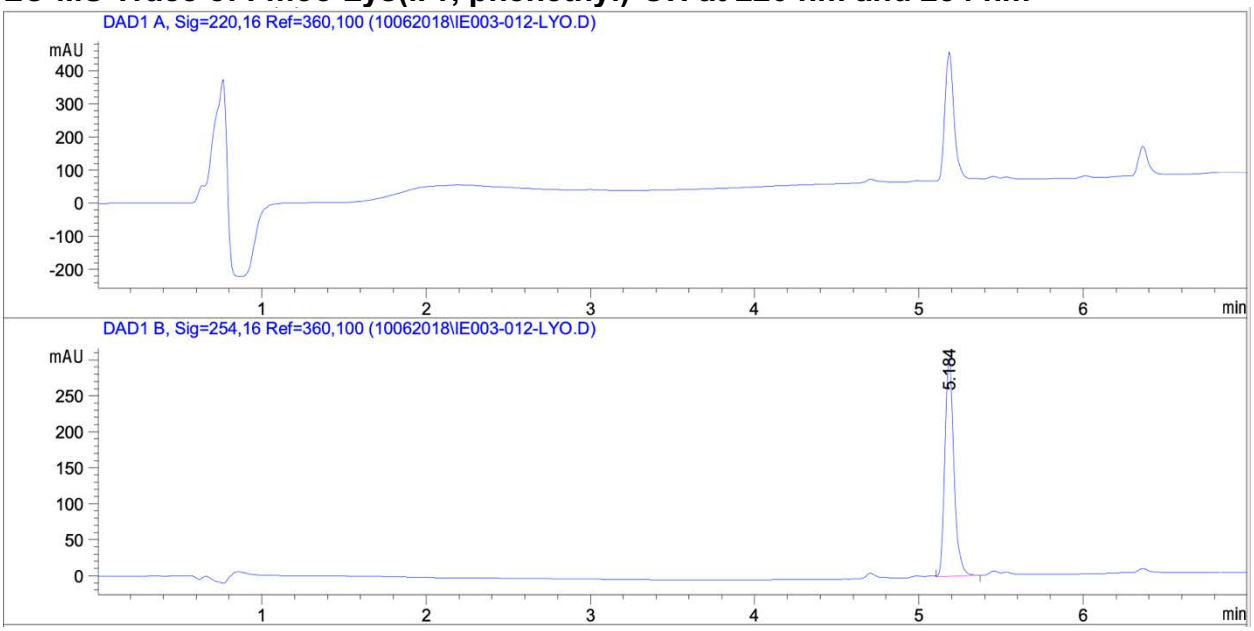


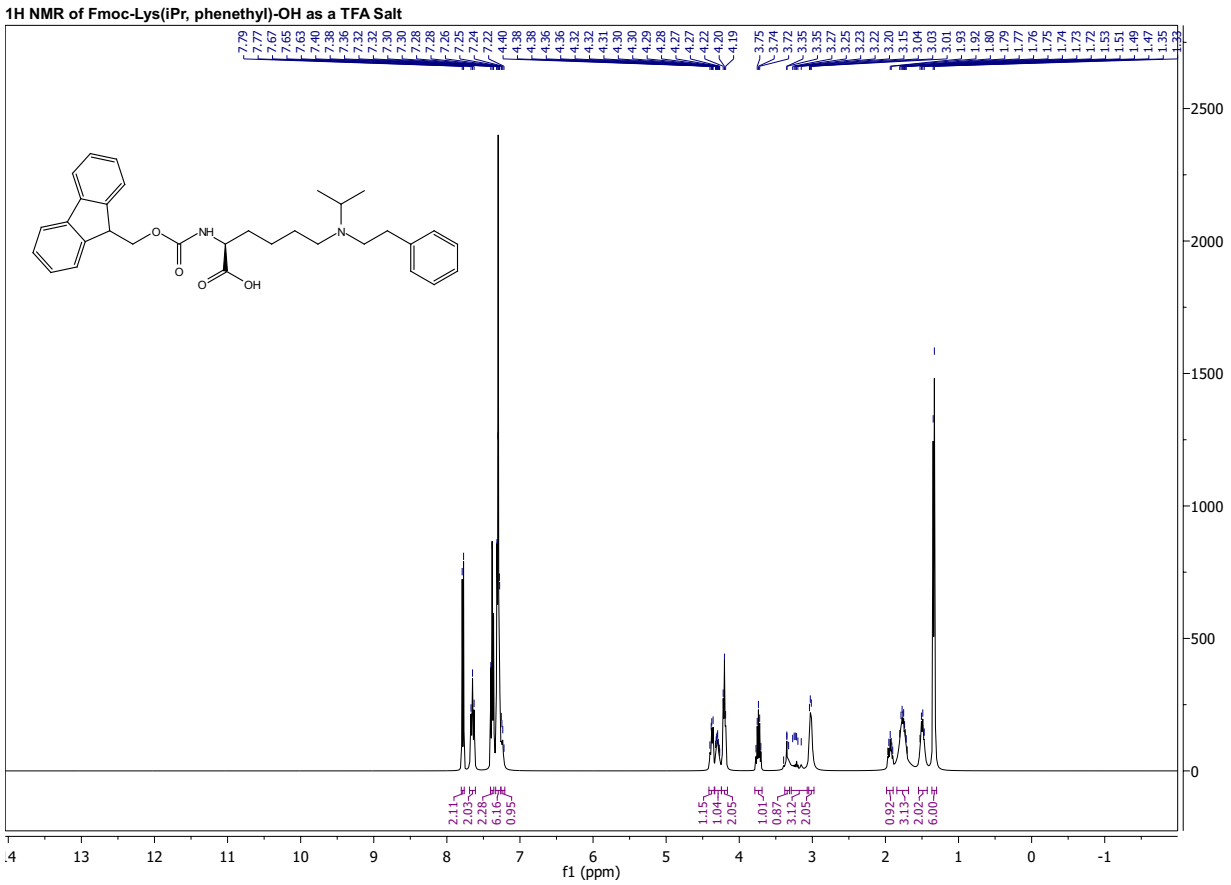
LC-MS Trace of UNC6640 at 220 nm and 254 nm



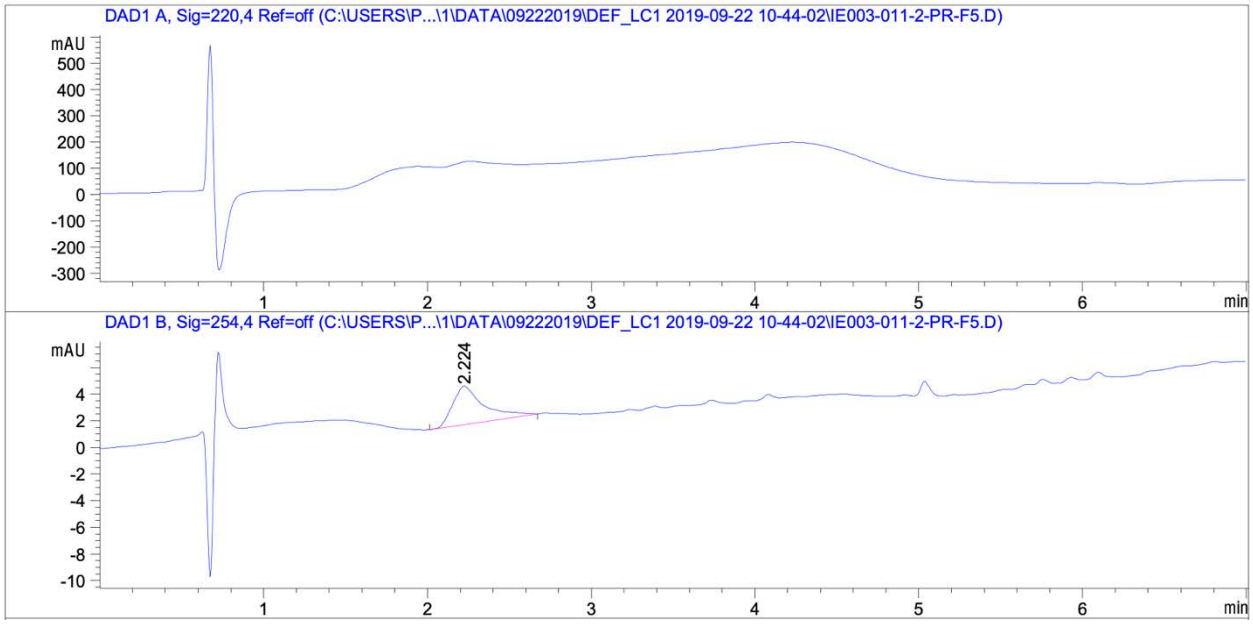


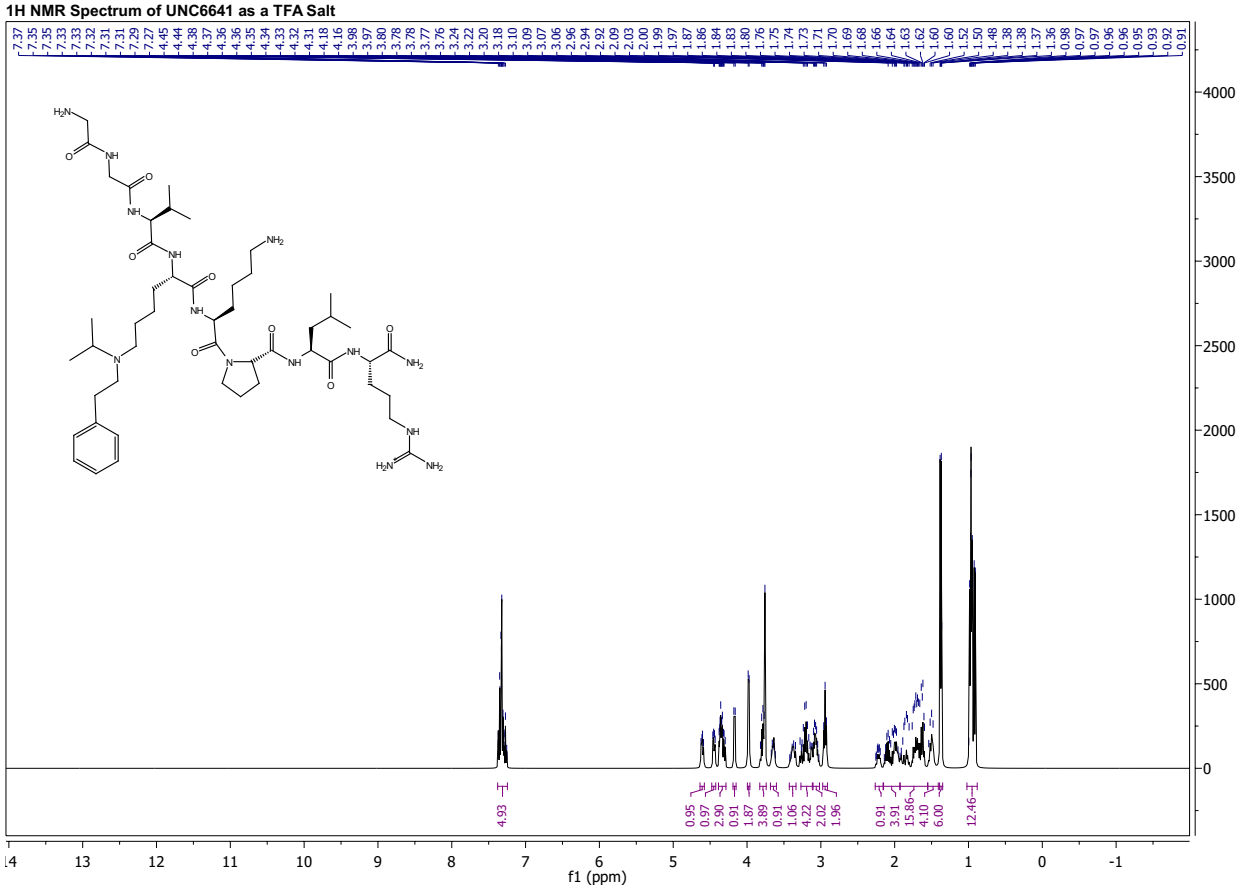
LC-MS Trace of Fmoc-Lys(iPr, phenethyl)-OH at 220 nm and 254 nm



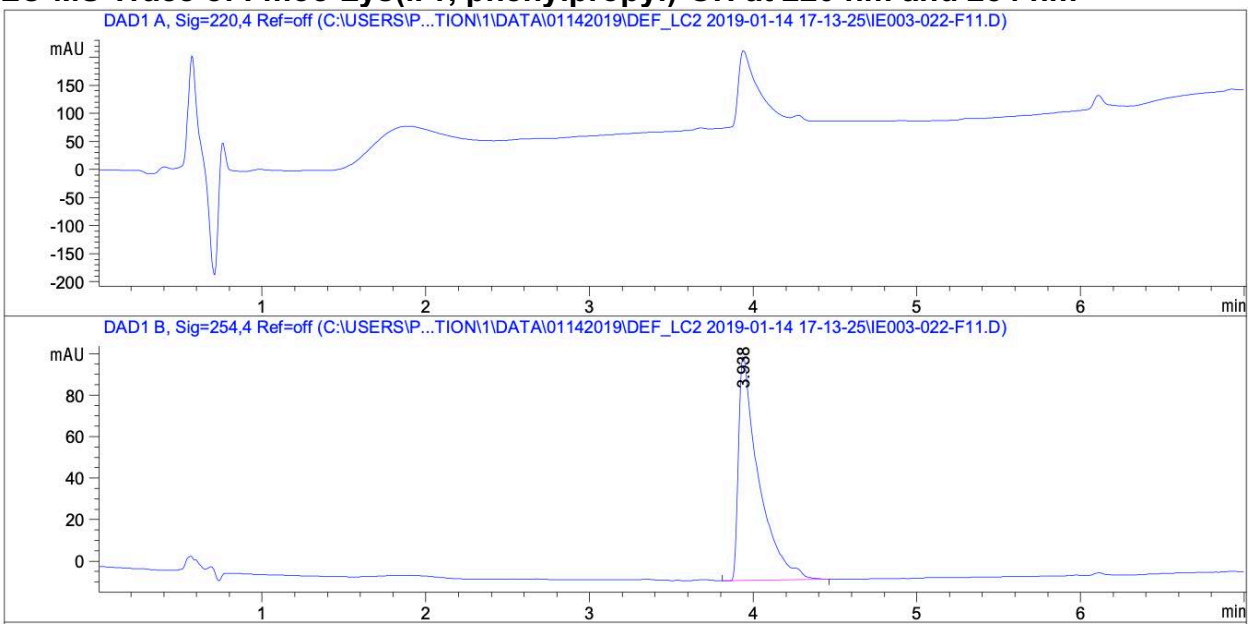


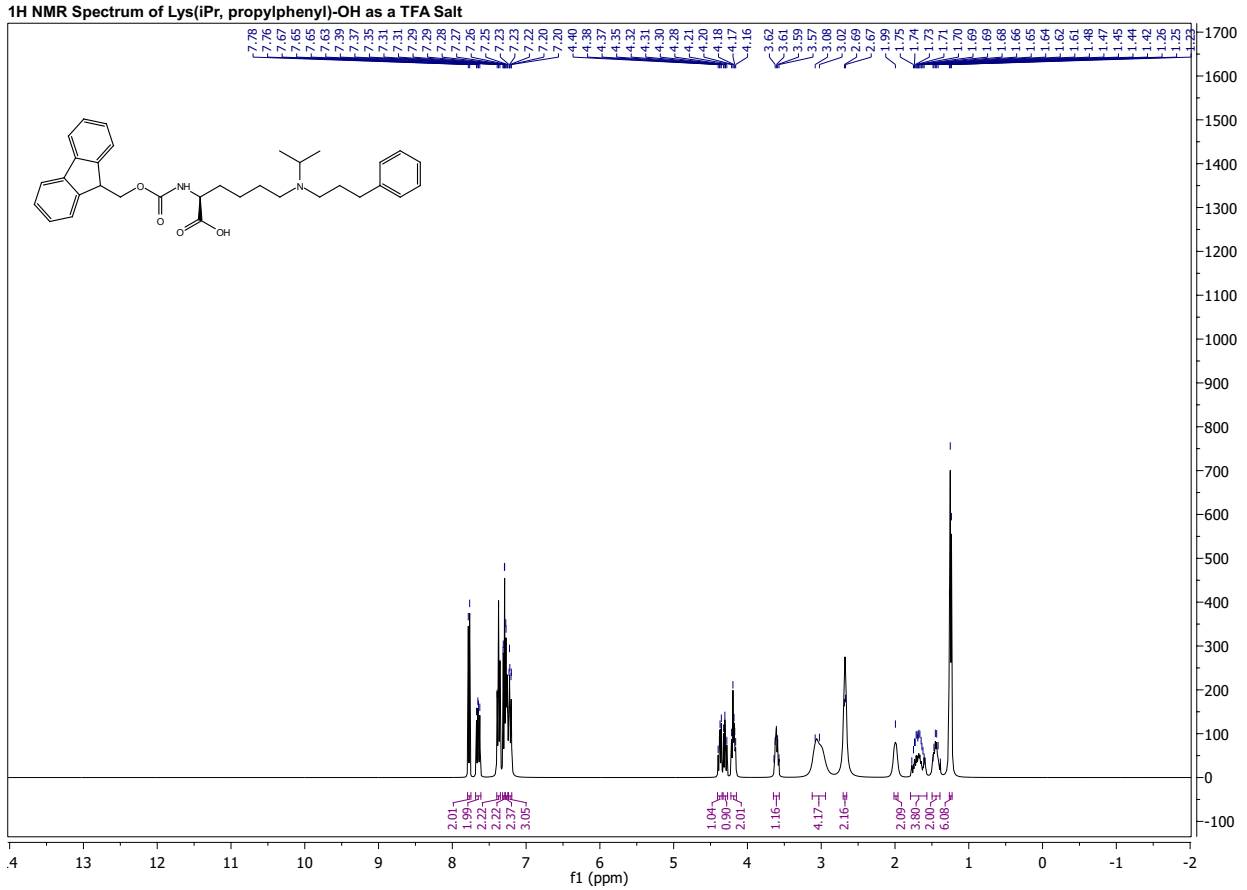
LC-MS Trace of UNC6641 at 220 nm and 254 nm



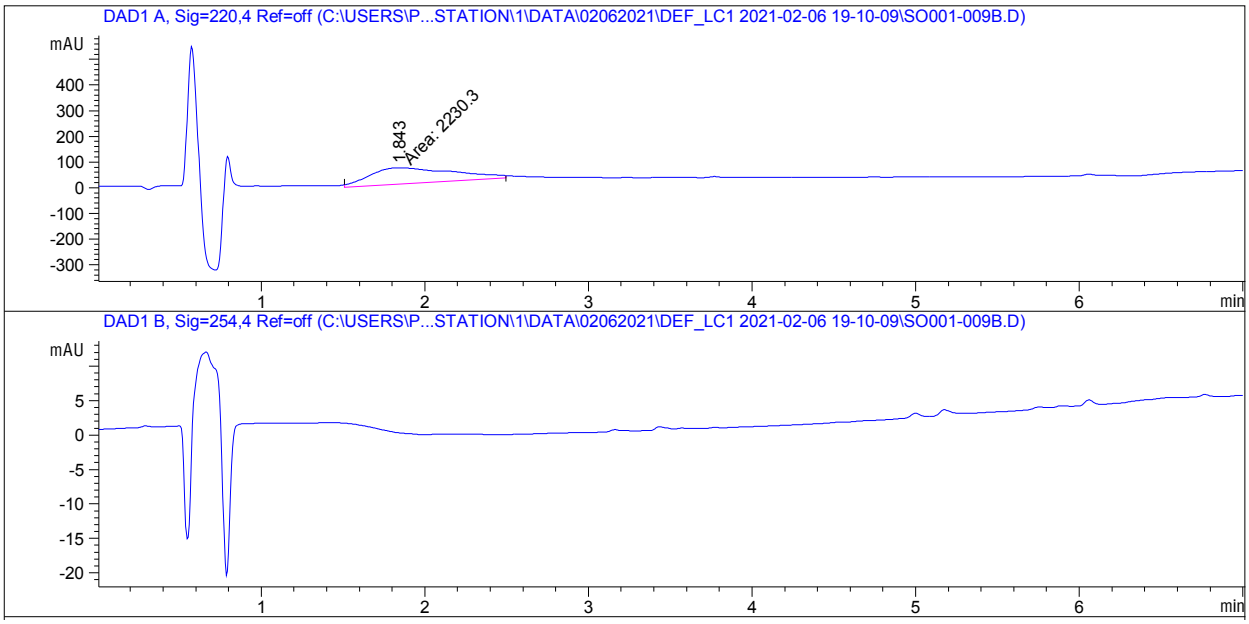


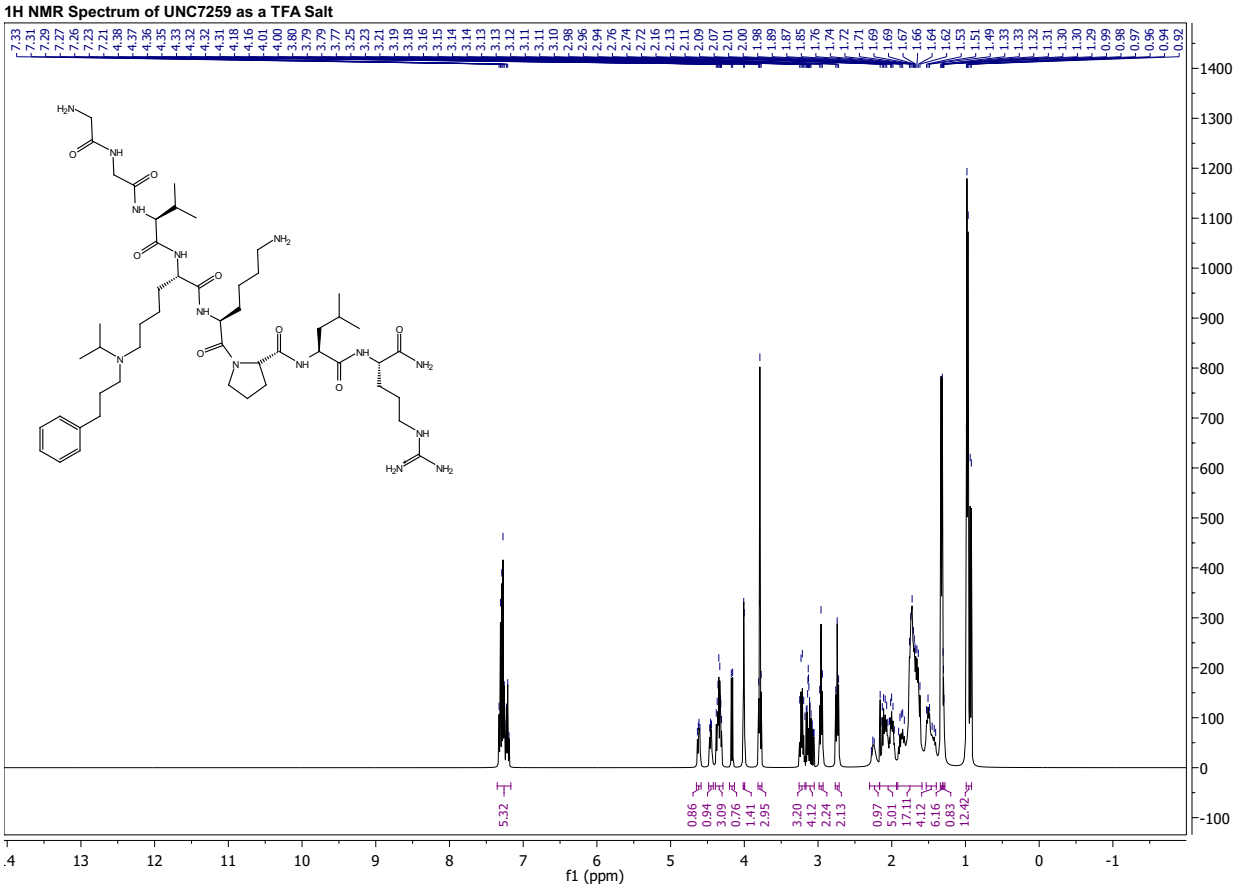
LC-MS Trace of Fmoc-Lys(iPr, phenylpropyl)-OH at 220 nm and 254 nm



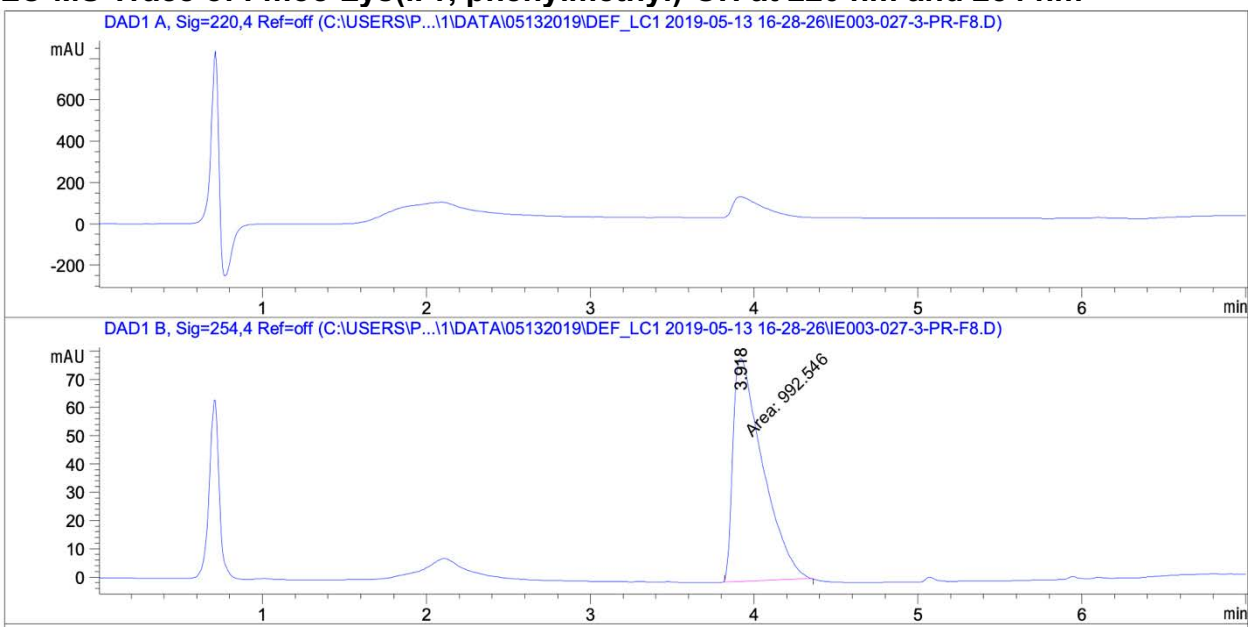


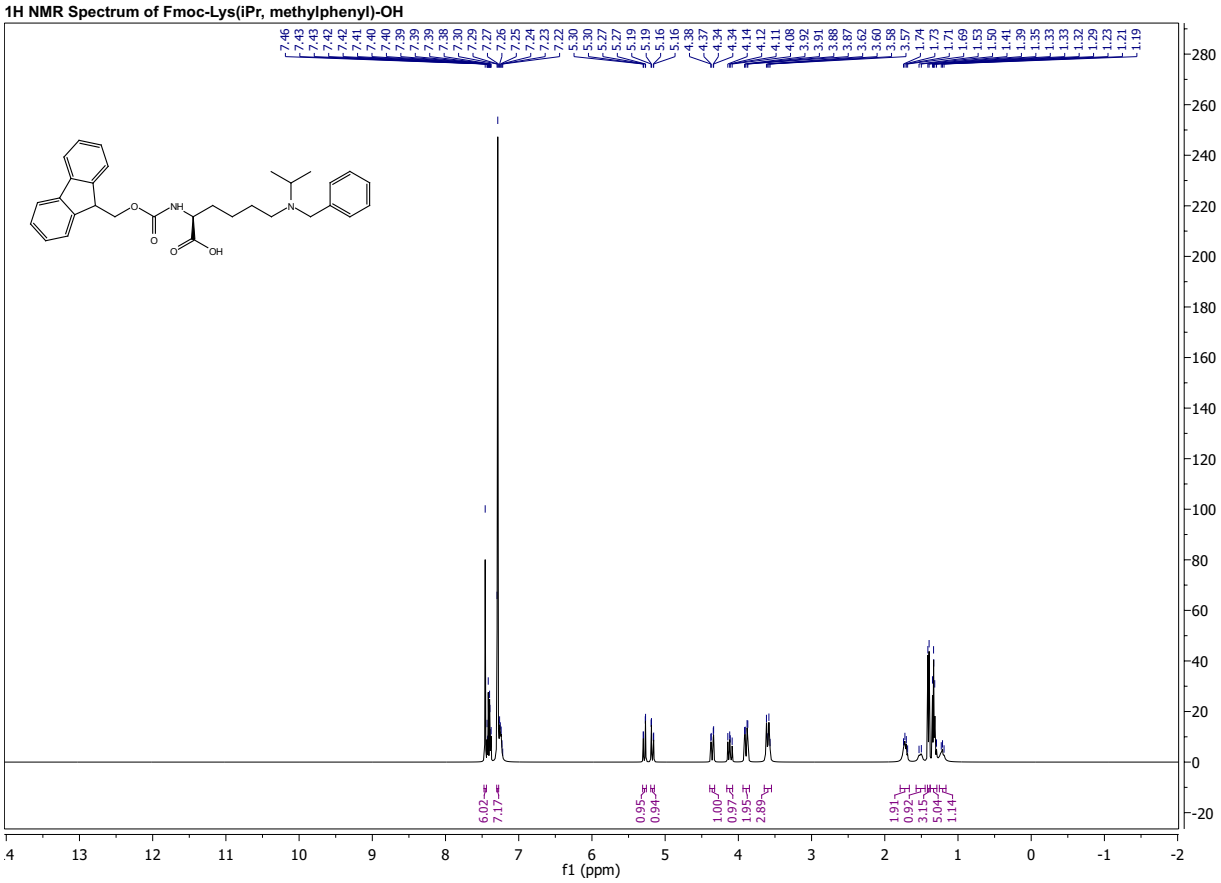
LC-MS Trace of UNC7259 at 220 nm and 254 nm



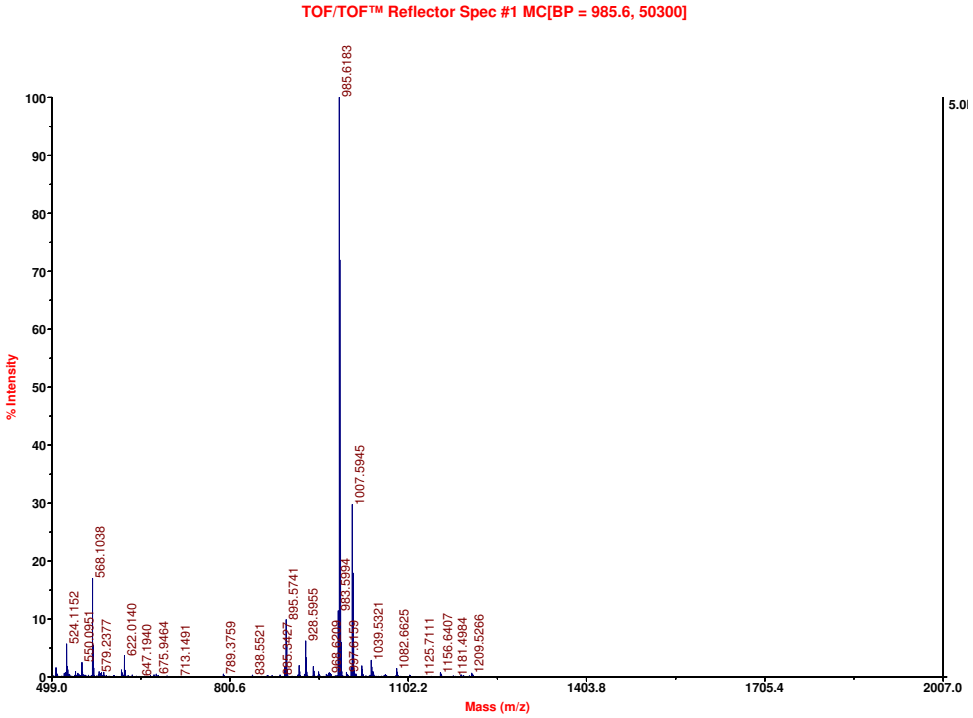


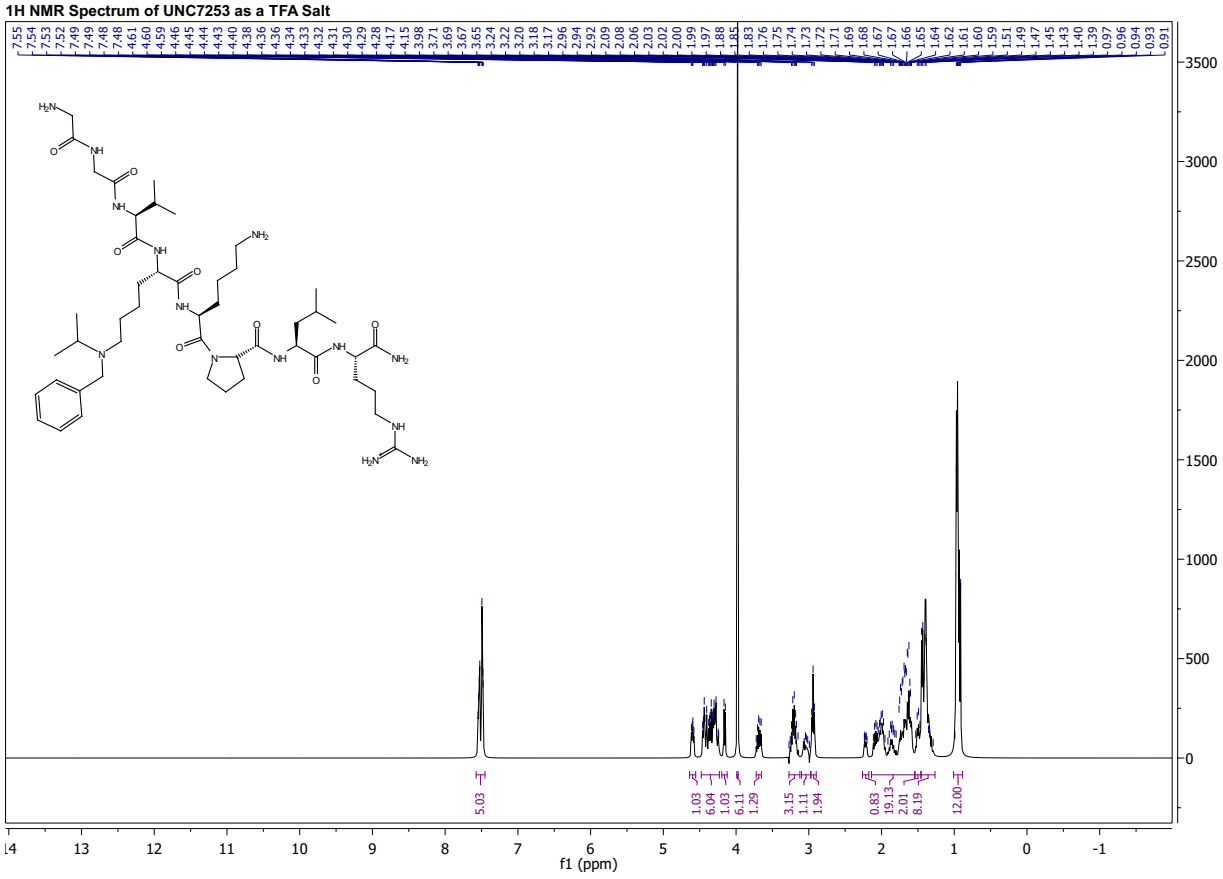
LC-MS Trace of Fmoc-Lys(iPr, phenylmethyl)-OH at 220 nm and 254 nm



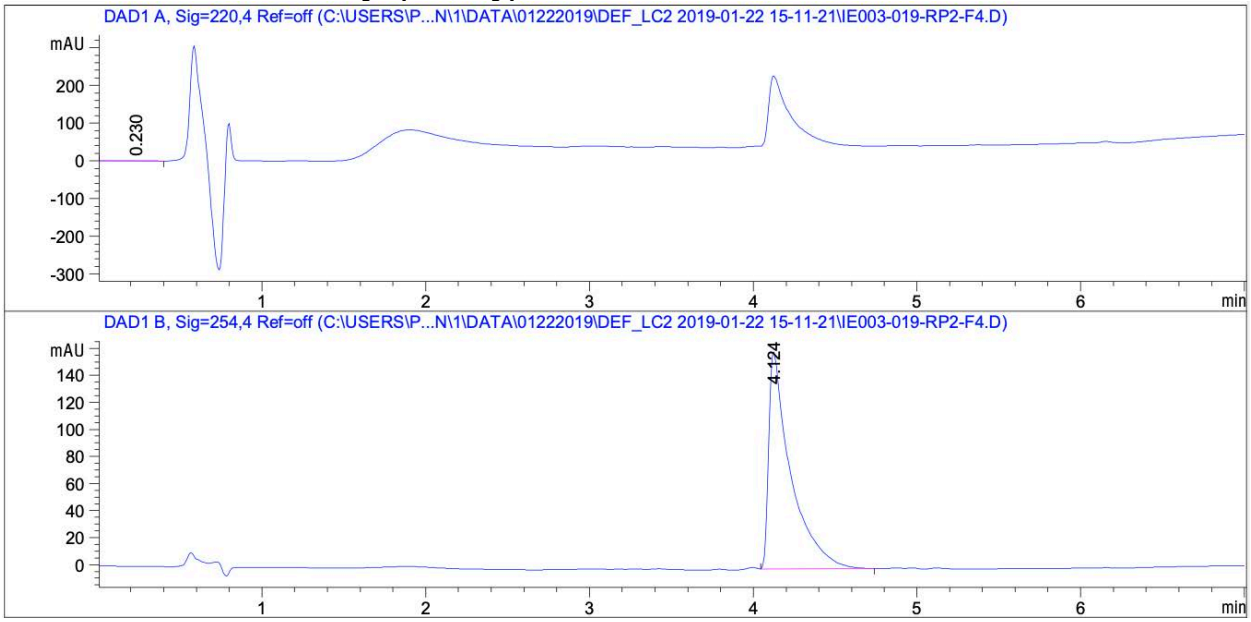


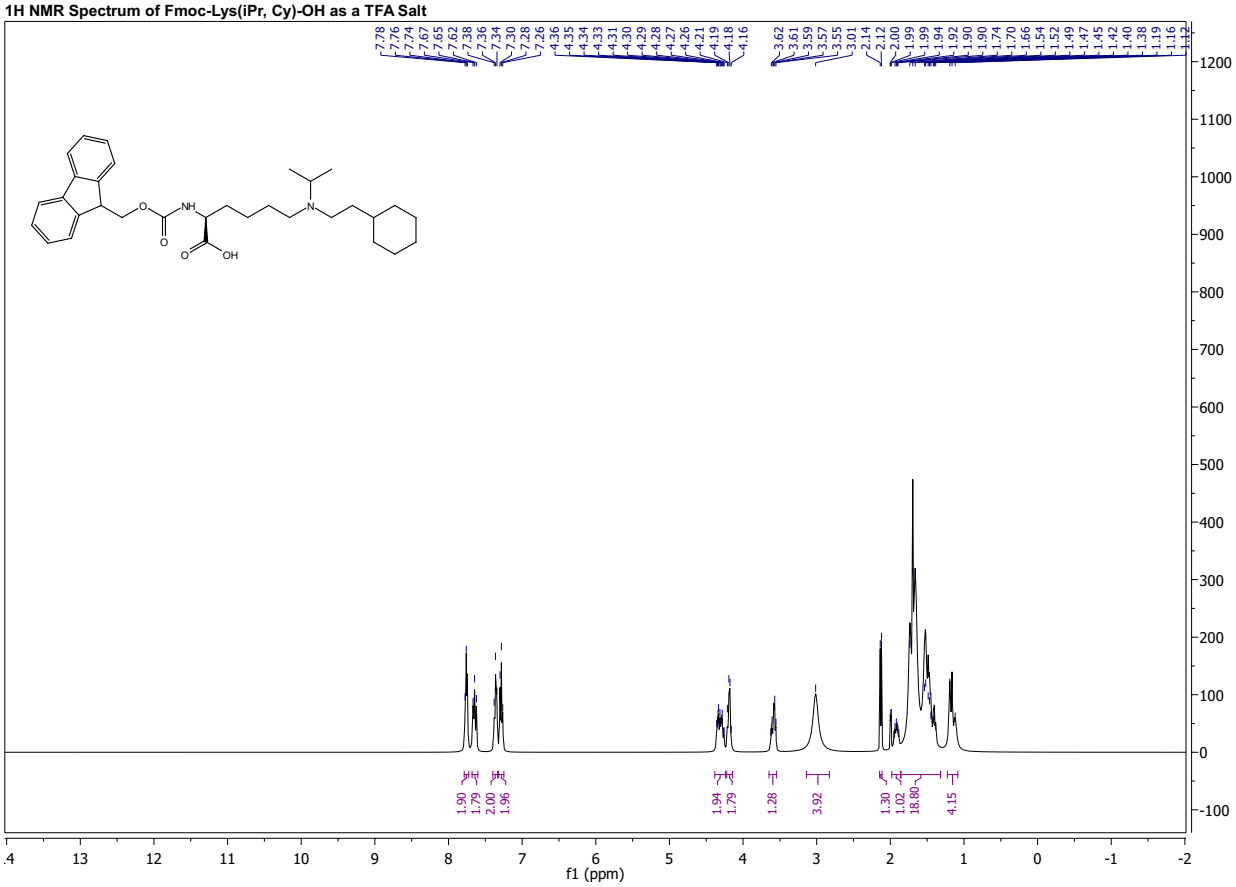
MALDI TOF/TOF Trace of UNC7253



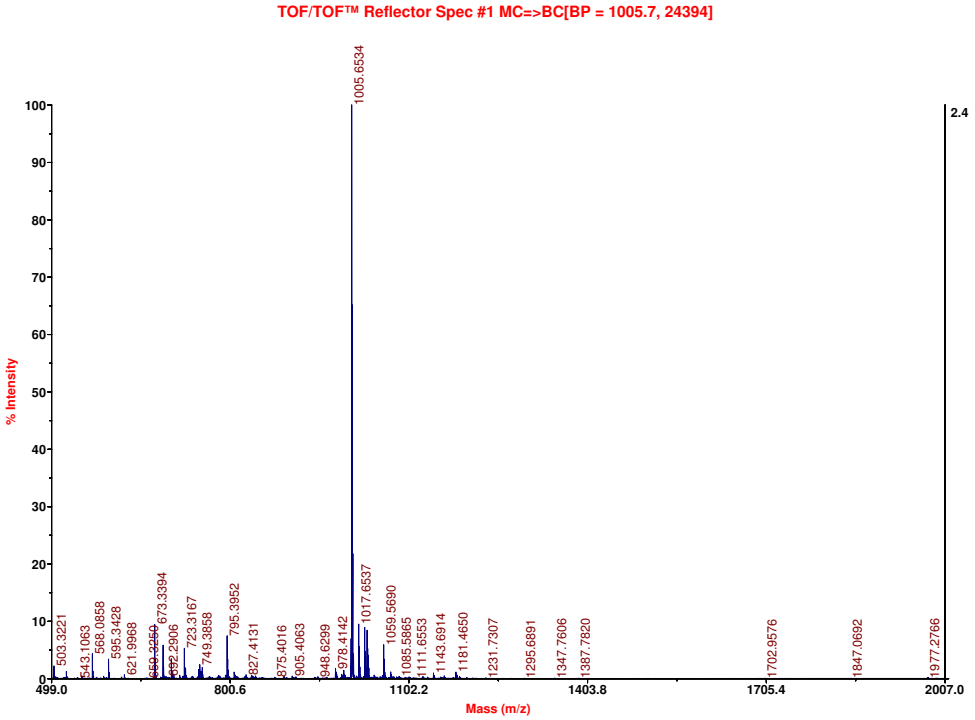


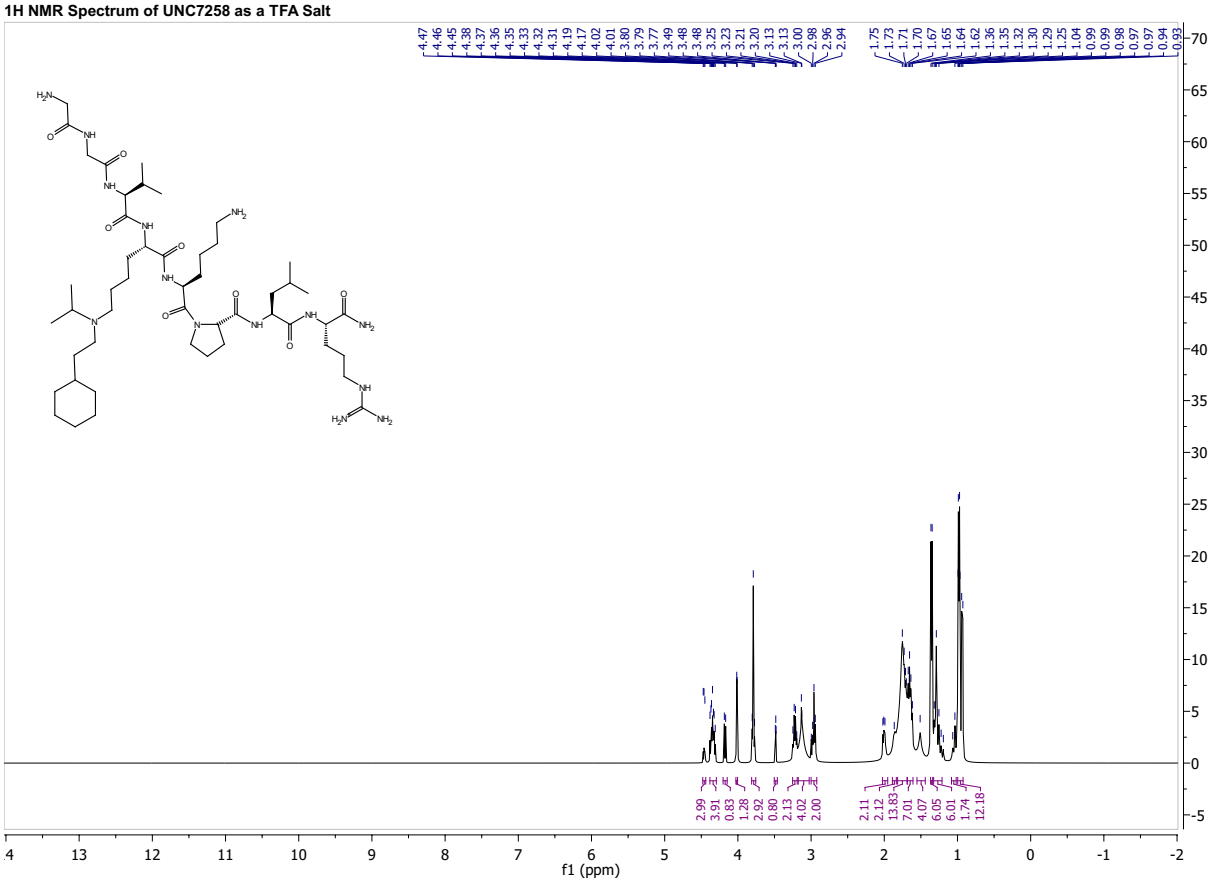
LC-MS Trace of Fmoc-Lys(iPr, Cy)-OH at 220 nm and 254 nm



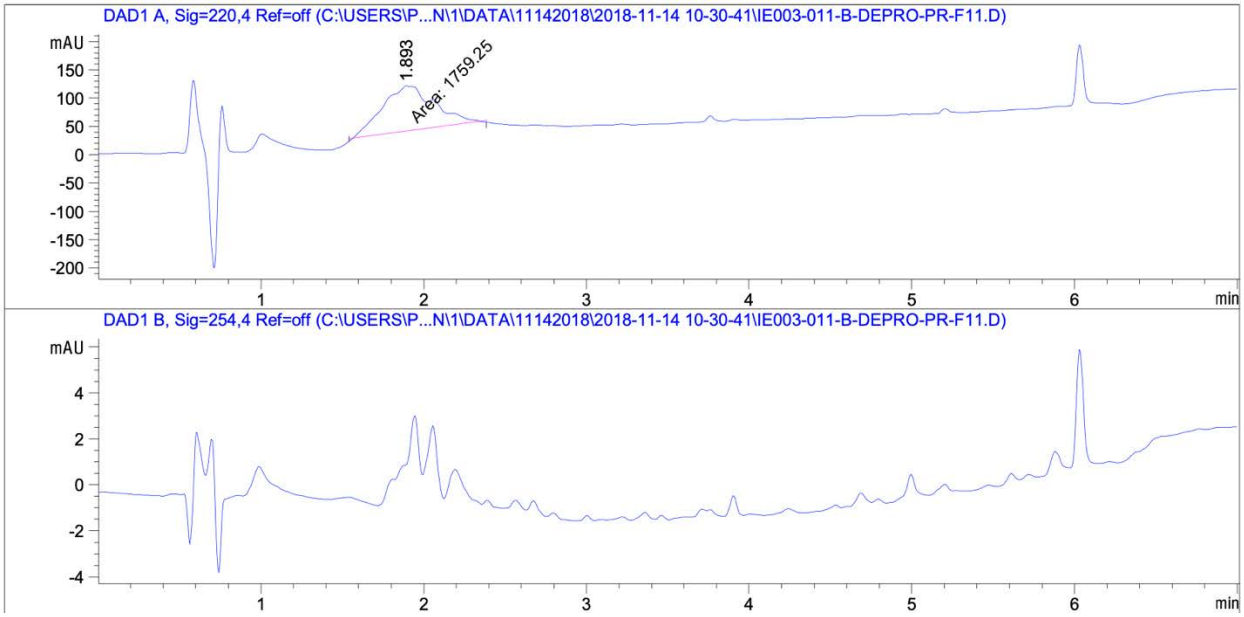


MALDI TOF/TOF Trace of UNC7258





LC-MS Trace of alkynyl-UNC6641 at 220 nm and 254 nm



LC-MS Trace of biotin-UNC6641 at 220 nm and 254 nm