DESIGN AND APPLICATION OF SYNTHETIC RECEPTORS FOR RECOGNITION OF METHYLATED LYSINE AND SUPRAMOLECULAR AFFINITY LABELING

Isaiah Nathaniel Gober

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry (Organic) in the College of Arts & Sciences

Chapel Hill
2017

Approved by:

Marcey Waters

Eric Brustad

Jeffrey Johnson

David Lawrence

Frank Leibfarth

ABSTRACT

Isaiah Nathaniel Gober: Design and Application of Synthetic Receptors for Recognition of Methylated Lysine and Supramolecular Affinity Labeling
(Under the direction of Marcey Waters)

This dissertation involves the design and synthesis of new synthetic receptors and their application in the molecular recognition of methylated lysine and their use as tools for chemical biology. The dissertation is divided into four parts. The first section focuses on the development of a novel labeling method that is based on ligand-directed affinity labeling principles. In this labeling method, a synthetic receptor that binds to trimethyl lysine (Kme3) is attached through a linker to an electrophilic tag group that can react with a nucleophilic amine in a histone peptide. This affinity labeling probe, which we called CX_4-ONBD, is equipped with an electrophilic tag that allows for turn-on fluorescence labeling of Kme3 histone peitdes. We show that the probe gives a pronounced turn-on fluorescence response when it is incubated with a histone peptide that contains Kme3 and a nearby reactive lysine. This probe also displays >5-fold selectivity in covalent labeling over an unmethylated lysine peptide. This represents the first time a synthetic receptor has been used for affinity labeling purposes, and it also expands on the chemical toolkit that is available for sensing PTMs like lysine methylation.

In the second section, the supramolecular affinity labeling method that was optimized using CX_4-ONBD was applied to the development of a real-time assay for measuring enzymatic activity. More specifically, the probe was used to create a turn-on fluorescence assay for histone deacetylase (HDAC) activity and for inhibitor screening and IC_{50} determination. Most

commercial kits for HDAC activity have limited substrate scope, and other common methods used for characterizing enzymatic activity often require chromatographic separation and are therefore not high-throughput. This small molecule receptor-mediated affinity labeling strategy allowed for facile readout of HDAC activity and inhibition. Overall, this application of supramolecular affinity labeling expands on the possible ways for detecting PTMs and may find use in the development of new assays for enzymes that lack robust methods for measuring their activity.

The third section explores the development of new small molecule receptors capable of selectively binding hydrophilic guests in water, such as the lower methylation states of lysine. We identified a receptor, $\mathbf{A_2I}$, that has improved binding affinity and selectivity for dimethyllysine (Kme2). The receptor was discovered and synthesized by using dynamic combinatorial chemistry (DCC) to redesign a small molecule receptor ($\mathbf{A_2B}$) that preferentially binds trimethyllysine (Kme3). Incorporating a biphenyl monomer with ortho-di-substituted carboxylates into the receptor lead to the formation of a salt bridge interaction with Kme2. These favorable electrostatic and hydrogen bonding interactions produced a receptor with 32-fold tighter binding to Kme2, which is the highest affinity synthetic receptor for Kme2 in the context of a peptide that has been reported. This work provides insight into effective strategies for binding hydrophilic, cationic guests in water and is an encouraging result toward a synthetic receptor that selectively binds Kme2 over other methylation states of lysine.

In the final section, a small molecule receptor for Kme3 ($\mathbf{A_2B}$) was redesigned using DCC to incorporate either aromatic or acidic amino acids into the receptor. We proposed that the incorporation of amino acids could introduce additional non-covalent interactions (such as cation-π, electrostatic, and hydrogen bonding) with a guest bound inside the pocket of the

receptor. However, selective non-covalent interactions between the amino acid side chain on the modified receptor and the bound methylated lysine guest could not be achieved. This is most likely due to the conformational flexibility of the amino acid-functionalized receptors. Furthermore, attaching amino acids to the receptor seemed to increase non-specific electrostatic interactions, resulting in tighter binding to the unmethylated lysine peptide (compared to $\mathbf{A_2B}$). Ultimately, this highlights the importance of incorporating monomers with less conformational flexibility that can rigidly place functional groups into the binding pocket.

ACKNOWLEDGMENTS

First, I would like to thank Marcey Waters for being such an encouraging and supportive mentor. I greatly appreciate your openness to allow me to be creative and to pursue new ideas in my grad school career. I also want to thank those who served on my orals committee and defense committee: Jeff Johnson, Mike Gagné, Eric Brustad, David Lawrence, and Frank Leibfarth.

I would like to thank the different members of the Waters lab. I am grateful to those who helped teach and train me in various techniques and instruments and who were always available to answer any questions that I had. I also appreciate the many discussions about research that helped me to think critically and carefully about my work and to grow as a scientist.

I also want to thank my family for their love and support throughout the years. Your phone calls and visits always lifted my spirits and encouraged me to keep going strong.

I would like to say thanks to my friends at Triangle Church who have been an amazing support system for me for the past five years. I especially want to thank Dr. Onyi Okolie for his amazing friendship. You were always there when things we great and when times were tough. I will forever cherish the memories of the many times that you and Carla opened up both your home and your lives to me during my time at UNC.

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xi

LIST OF SCHEMES xxv

LIST OF ABBREVIATIONS xxvi

CHAPTER 1. Approaches for Site-Selective Labeling of Endogenous Peptides and Proteins 1

Introduction 1

Chemical Reactions for Modification of Endogenous Amino Acids 2

Activity Based Protein Profiling 6

Ligand-Directed Affinity Labeling 7

CHAPTER 2. Approaches for Detection of Histone Post-Translational Modifications 10

DNA Organization and Chromatin Structure 10

Covalent Modification of Histone Proteins 12

Aberrant PTMs and Disease 14

Current Methods to Detect Histone PTMs 15

Significance of This Work 19

CHAPTER 3. Supramolecular Affinity Labeling of Histone Peptides 21

Introduction and Significance 21

Design and Synthesis of a Probe for Labeling for Kme3 23

Characterization of Binding by Isothermal Titration Calorimetry 25

Optimization of Receptor-Mediated Labeling 27

Turn-On Fluorescence Labeling of Kme3 Peptides 29

Conclusions 32

Experimental 33

CHAPTER 4. A Supramolecular Affinity Labeling-Based Turn-On Fluorescence Assay for Histone Deacetylase Activity 75

Turn-On Fluorescence Assay for HDAC Activity 75

Conclusions 79

Experimental 80

CHAPTER 5. Design and Synthesis of Synthetic Receptors for Lower Methylation States of Lysine 90

Introduction and Significance 90

System Design 92

Synthesis of Monomers I and J 94

Dynamic Combinatorial Library Screening 96

NMR Characterization of Binding to Kme2 and Kme3 98

ITC Binding Studies 98

Influence of the Internal Carboxylate on Selectivity 99

Comparison of A_2I to A_2B with a biologically relevant peptide 99

Influence of Biphenyl Functional Group 102

Conclusions 102

Experimental 103

CHAPTER 6. Redesign of a Small Molecule Receptor for Methylated Lysine via Incorporation of Aromatic and Acidic Amino Acids 185

Introduction and Significance 185

System Design 186
Synthesis of Monomers 188
Dynamic Combinatorial Library Screening 189
Binding Studies 191
Influence of Attaching Aromatic Amino acids to Monomer B 192
Influence of Attaching an Acidic Amino Acid to Monomer B 192
Conclusions 194
Experimental 195
REFERENCES 235

LIST OF TABLES

Table 3.1. Peptides Used for ITC, Fluorescence Labeling, and Deacetylase Activity Screening 26

Table 3.3. Optimization of Labeling Reaction Conditions of CX4-ONBD Probe 5 with H3 Peptides 29

Table 3.4. Summary of ITC data from using a one-site binding model. 47

Table 3.5. Summary of peak integration data from HPLC trace of a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK. 69

Table 3.6. Summary of peak integration and labeling data from the HPLC trace of the reaction of 40 μM probe 5 with a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK. 70

Table 3.7. Summary of peak integration data HPLC trace of a mixture of 30 μM K9K14 and 30 μM RKme3GLLYK 71

Table 3.8. Summary of peak integration data from the HPLC trace of the reaction of 40 μM probe 5 with a mixture of 30 μM K9K14 and 30 μM RKme3GLLYK 72

Table 4.1. Summary of Fluorescence Intensities and Z Factor Calculation 89

Table 5.1. Change in chemical shift (Δδ) upon binding of Ac-Kme2-Gly-Gly-Tyr-NH_2 or Ac-Kme3-Gly-Gly-Tyr-NH_2 to A_2I-2 98

Table 5.2. Peptides used for ITC binding studies 99

Table 5.3. Thermodynamic binding data for the biding of A_2I and A_2J to peptides shown in Table 5.1 as measured by ITC. 101

Table 5.4. Summary of A_2I and A_2J ITC data from using a one-site binding model 141

Table 6.1. Peptides used for ITC. 192

Table 6.2. Thermodynamic binding data for the biding of A_2B-Phe, A_2B-Trp, A_2B-Tyr, A_2B-Asp-1, and A_2B-Asp-2 to peptides shown in Table 6.1 as measured by ITC. 193

Table 6.3. Thermodynamic binding data for the biding of A_2B-Phe, A_2B-Trp, A_2B-Tyr, A_2B-Asp-1, and A_2B-Asp-2 to peptides shown in Table 6.1 as measured by ITC. 219

LIST OF FIGURES

Figure 1.1. Classical reactions used in modifying lysine and cysteine residues..... 4

Figure 1.2. Ligand-directed ONBD chemistry. 9

Figure 2.1. Packaging of chromosomal DNA by winding around histone proteins..... 11

Figure 3.1. Turn-on fluorescence affinity labeling of a histone Kme3 peptide using a CX4-ONBD probe..... 23

Figure 3.2. Turn-on fluorescence labeling of K9me3K14 and K9K14 peptides with Probe 5 30

Figure 3.3. Analysis of distance dependence in the labeling of histone peptides by probe 5. 32

Figure 3.4. 1D ^{1}H-NMR (400 MHz) of Compound 1 in $CDCl_3$..... 34

Figure 3.5. High resolution mass spectrum of Compound 1. 35

Figure 3.6. Crude semi-preparative HPLC trace of the synthesis of 4 from 3 at 280 nm. 36

Figure 3.7. ^{1}H-NMR (600 MHz) of Compound 4 in D_2O. 37

Figure 3.8. High resolution mass spectrum of Compound 4. 38

Figure 3.9. Crude semi-preparative HPLC trace of 5 at 280 nm. 39

Figure 3.10. Semi-preparative HPLC trace prior to second purification of 5 at 280 nm. 39

Figure 3.11. ^{1}H-NMR (600 MHz) of Compound 5 in D_2O. 40

Figure 3.12. High resolution mass spectrum of compound 5 (ESI-). 40

Figure 3.13. High resolution mass spectrum of K9 K14 (ESI +). 42

Figure 3.14. High resolution mass spectrum of K9me_3 K14 (ESI +). 42

Figure 3.15. High resolution mass spectrum of K9me_3 R14 (ESI..... 43

Figure 3.16. High resolution mass spectrum of K4me_3 K9 (ESI +). 43

Figure 3.17. High resolution mass spectrum of K4me_3 K9ac (ESI +)..... 43

Figure 3.18. High resolution mass spectrum of K27 K36me_3 (ESI +). 44

Figure 3.19. High resolution mass spectrum of K27ac K36me$_3$ (ESI +)..... 44

Figure 3.20. High resolution mass spectrum of K23K27me3 (ESI 44

Figure 3.21. High resolution mass spectrum of K36me3K37 (ESI +)..... 45

Figure 3.22. High resolution mass spectrum of K4me3 N terminus (ESI +)..... 45

Figure 3.23. High resolution mass spectrum of K4me3 N terminus (ESI +)..... 45

Figure 3.24. Extinction coefficient determination of 4 from the slope of the line of regression..... 46

Figure 3.25. Heat of dilution for titrating K9me3 R14 (1.22 mM) into 10 mM potassium phosphate buffer pH 7.4 48

Figure 3.26. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.22 mM) titrated into 5 (82 μM) at 25 ^oC in 10 mM potassium phosphate buffer, pH 7.4 (Run 1). 49

Figure 3.27. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.22 mM) titrated into 5 (82 μM) at 25 ^oC in 10 mM potassium phosphate buffer, pH 7.4 (Run 2). 50

Figure 3.28. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.03 mM) titrated into 5 (81 μM) at 25 ^oC in 10 mM potassium phosphate buffer, pH 7.4 (Run 3). 51

Figure 3.29. Heat of dilution for titrating K9K14 (Ac-WGGQTARKSTGGK-NH_2) (1.43 mM) into 10 mM sodium borate buffer pH 8.6..... 52

Figure 3.30. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (1.43 mM) titrated into 4 (101 μM) at 25 ^oC in 10 mM sodium borate buffer, pH 8.6 (Run 1). 53

Figure 3.31. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (0.93 mM) titrated into 4 (85 μM) at 25 ^oC in 10 mM sodium borate buffer, pH 8.6 (Run 1). 54

Figure 3.32. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (0.93 mM) titrated into 4 (85 μM) at 25 ^oC in 10 mM sodium borate buffer, pH 8.6 (Run 2). 55

Figure 3.33. Heat of dilution for titrating K9me3K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.43 mM) into 10 mM sodium borate buffer pH 8.6 56

Figure 3.34. One trial of three of K9me3K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.43 mM) titrated into 4 (101 μM) at 25 ºC in 10 mM sodium borate buffer, pH 8.6 (Run 1). 57

Figure 3.35. One trial of three of K9me3 K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.11 mM) titrated into 4 (85 μM) at 25 ºC in 10 mM sodium borate buffer, pH 8.6 (Run 2). 58

Figure 3.36. One trial of three of K9me3 K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.11 mM) titrated into 4 (85 μM) at 25 ºC in 10 mM sodium borate buffer, pH 8.6 (Run 3). 59

Figure 3.37. LCMS characterization of CX_4-ONBD labeling of K9K14 and K9me3K14 60

Figure 3.38. MS spectra of the reaction of 100 μM 5 with 100 μM K9 K14 in 10 mM sodium borate buffer, pH 8.6 after 24 h. 61

Figure 3.39. HPLC trace of the reaction of 100 μM 5 with 100 μM K9 K14 (top) or 100 μM K9me3 K14 (bottom) in 10 mM potassium phosphate buffer, pH 7.4 after 24 h showing labeled and unlabeled peptide. 62

Figure 3.40. HPLC trace of the reaction of 100 μM 5 with 50 μM K9 K14 (top) or 50 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6 after 24 h. 63

Figure 3.41. HPLC trace of the reaction of 100 μM 5 with 50 μM K9 K14 (top) or 50 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 50 mM NaCl after 24 h. 64

Figure 3.42. HPLC trace of the reaction of 100 μM 5 with 30 μM K9 K14 (top) or 30 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h. 65

Figure 3.43. HPLC trace of the reaction of 100 μM 5 with 30 μM K9 K14 (top) or 30 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 150 mM NaCl after 24 h showing labeled and unlabeled peptide 66

Figure 3.44. HPLC trace of the reaction of 40 μM 5 with 20 μM K9 K14 (top) or 20 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h showing labeled and unlabeled peptide. 67

Figure 3.45. HPLC trace of the reaction of 40 μM 1 with 20 μM K9 K14 (blue) or 20 μM K9me3 K14 (black) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h demonstrating that the peptides do not react with a probe lacking the recognition motif. 68

Figure 3.46. HPLC trace of a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl to provide a concentration reference. 69

Figure 3.47. HPLC trace of the reaction of 40 μM probe 5 with a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl. 70

Figure 3.48. HPLC trace of a mixture of 30 μM K9K14 and 30 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl to provide a concentration reference. 71

Figure 3.49. HPLC trace of the reaction of 40 μM probe 5 with a mixture of 30 μM K9K14 and 30 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl. 72

Figure 3.50. Turn-on fluorescence labeling negative control experiment 73

Figure 3.51. Turn-on fluorescence labeling competition experiment. 74

Figure 4.1. Turn-on fluorescence HDAC1 activity assay. 77

Figure 4.2. Dose-response curve of the inhibition of the HDAC3/NCOR1 catalyzed deacetylation of K4me3K9ac by SAHA. Data are an average of duplicate measurements. 79

Figure 4.3. HPLC trace at 280 nm of a mixture of 20 μM K4me3 K9ac and 20 μM K4me_3 K9 (top) and the reaction of 0.007 μg/μL HDAC1 with 20 μM K4me3 K9ac after 30 min (bottom). 81

Figure 4.4. HPLC trace at 280 nm of a mixture of 20 μM K4me3 K9ac and 20 μM K4me_3 K9 (top) and the reaction of 0.1 μg/μL HDAC1 with 20 μM K4me3 K9ac after 30 min (bottom). 82

Figure 4.5. HPLC trace at 280 nm of a mixture of 20 μM K27ac K36me3 and 20 μM K27 K36me3 (top) and the reaction of 0.1 μg/μL HDAC1 with 20 μM K27ac K36me3 after 30 min (bottom). 82

Figure 4.6. HPLC trace at 280 nm of the reaction of 0.1 μg/μL HDAC1 with 20 μM K4me3 K9ac after 1 h. 83

Figure 4.7. HPLC trace at 280 nm of the reaction of 0.1 μg/μL HDAC1 with 20 μM K27ac K36me3 after 1 h. 83

Figure 4.8. Top: HPLC trace of a 1:1 mixture of 100 μM $K4me_3K9ac$ and 100 μM $K4me_3K9$. Bottom: HPLC trace at 214 of the reaction of 11.2 nM HDAC3/NCOR1 with 100 μM $K4me_3K9ac$ after 4 h. 83

Figure 4.9. Turn-on fluorescence HDAC1 activity assay. 84

Figure 4.10. Turn-on fluorescence labeling SAHA control experiment. 85

Figure 4.11. Turn-on fluorescence HDAC1 activity assay: fluorescence increase of deacetylated K4me3 K9ac after labeling for 6, 12, or 24 h 86

Figure 4.12. Turn-on fluorescence HDAC1 activity assay: fluorescence increase of deacetylated K27ac K36me3 after labeling for 6 ,12, or 24 h. 86

Figure 4.13. Turn-on fluorescence HDAC3/NCOR1 activity assay: fluorescence increase of deacetylated K4me3K9ac after labeling for 24 h 87

Figure 4.14. Plot of the dose response curve for the inhibition of HDAC3/NCOR1 activity by SAHA as monitored by the CX4-ONBD turn-on fluorescence HDAC assay. 88

Figure 4.15. Equation for calculating Z'. 89

Figure 4.16. Fluorescence intensities of positive control reaction and negative control reactions for eight wells for Z' determination. 89

Figure 5.1. Methylation states of lysine. 91

Figure 5.2. (a) Redesign of A_2B to A_2I. (b) Monomer I. (c) Gas-phase molecular model of one isomer of A_2I (blue) bound to butyldimethyl ammonium (cyan), a mimic of Kme2. 92

Figure 5.3. (a) BPTF PHD domain bound to Kme3. (b) BPTF PHD domain Y17E mutation bound to Kme2. 93

Figure 5.4. (a) DCLs of monomers A and I with different methylated butylamine guests (1 mM A, 1 mM I, 2 mM guest, 50 mM sodium borate buffer, pH 8.5). (b) Amplification of A_2I -1 and A_2I-2 in the presence of Kmex mimics. 97

Figure 5.5. 1D ^{1}H-NMR (600 MHz) of Compound 2 in $CDCl_3$. 105

Figure 5.6. 1D ^{13}C-NMR (600 MHz) of Compound 2 in $CDCl_3$. 106

Figure 5.7. High resolution mass spectrum of Compound 2. 106

Figure 5.8. 1D ^{1}H-NMR (600 MHz) of Compound 3 in $CDCl_3$. 107

Figure 5.9. 1D ^{13}C-NMR (600 MHz) of Compound 3 in $CDCl_3$. 108

Figure 5.10. High resolution mass spectrum of Compound 3. 108

Figure 5.11. 1D ^{1}H-NMR (600 MHz) of Monomer J in CD_3OD. 109

Figure 5.12. 1D ^{13}C-NMR (600 MHz) of Monomer J in CD_3OD. 110

Figure 5.13. High resolution mass spectrum of Monomer J. 110

Figure 5.14. 1D ^{1}H-NMR (600 MHz) of Compound 5 in $CDCl_3$. 112

Figure 5.15. 1D ^{13}C-NMR (600 MHz) of Compound 5 in $CDCl_3$. 113

Figure 5.16. High resolution mass spectrum of Compound 5. 113

Figure 5.17. 1D ^{1}H-NMR (600 MHz) of Compound 7 in $CDCl_3$. 114

Figure 5.18. 1D ^{13}C-NMR (600 MHz) of Compound 7 in $CDCl_3$. 115

Figure 5.19. High resolution mass spectrum of Compound 7. 115

Figure 5.20. 1D ^{1}H-NMR (600 MHz) of Monomer I in CD_3OD. 116

Figure 5.21. 1D ^{13}C-NMR (600 MHz) of Monomer I in CD_3OD. 117

Figure 5.22. High resolution mass spectrum of Monomer I. 117

Figure 5.23. Semi-preparative reversed-phase HPLC traces of the purification of A_2I from a preparative scale DCL. Absorbance measured at 214 nm. 118

Figure 5.24. High resolution mass spectrum of A_2I-1 (ESI 118

Figure 5.25. High resolution mass spectrum of A_2I-2 (ESI -). 119

Figure 5.26. (a) Low resolution mass spectrum of background (ESI -).
(b) High resolution mass spectrum of background (ESI -). 120

Figure 5.27. Semi-preparative reversed-phase HPLC traces of the
purification of A_2J from a preparative scale DCL.
Absorbance measured at 214 nm. 121

Figure 5.28. High resolution mass spectrum of A_2J-1 (ESI -). 121

Figure 5.29. High resolution mass spectrum of A_2J-1 (ESI -). 122

Figure 5.30. ^{1}H NMR spectrum of A_2I-1 in 50 mM sodium borate-d_3
D_2O, pH 8.5 at room temperature. 123

Figure 5.31. ^{1}H NMR spectrum of A_2I-2 in 50 mM sodium borate-d_3
D_2O, pH 8.5, 0.05 mM DSS (standard) at room temperature. 124

Figure 5.32. ^{1}H NMR spectrum of A_2J-1 in MeOD at 40 ^{o}C. 125

Figure 5.33. ^{1}H NMR spectrum of A_2J-2 in 50 mM sodium borate-d_3
D_2O, pH 8.5, 0.05 mM DSS (standard) at 40 ^{o}C 126

Figure 5.34. Overlaid HPLC traces of DCLs after 7 days containing
1 mM A, 1 mM J, and 2 mM guest (absorbance at 280 nm).
A_2J is amplified in the presence of Kme2 and Kme3. 127

Figure 5.35. High resolution mass spectrum of H3K9 (ESI +). 129

Figure 5.36. High resolution mass spectrum of H3K9me (ESI +). 129

Figure 5.37. High resolution mass spectrum of H3K9me2 (ESI +). 130

Figure 5.38. High resolution mass spectrum of H3K9me3 (ESI +). 130

Figure 5.39. High resolution mass spectrum of Kme2GGY (ESI 131

Figure 5.40. High resolution mass spectrum of Kme3GGY (ESI +). 131

Figure 5.41. Extinction coefficient determination of A_2I from the slope
of the line of regression. 132

Figure 5.42. Extinction coefficient determination of A_2J from the slope
of the line of regression. 133

Figure 5.43. ^{1}H NMR spectrum of Kme2GGY (5.6 mM) alone in 50 mM
sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard)
at room temperature. 134

Figure 5.44. ^{1}H NMR spectrum of A_2I-2 (0.75 mM) with Kme2GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at room temperature. 135

Figure 5.45. ^{1}H NMR spectrum of A_2I-2 (0.75 mM) with Kme2GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at 40 ^{o}C. 136

Figure 5.46. ^{1}H NMR spectrum of Kme3GGY (7.6 mM) alone in 50 mM sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard) at room temperature. 137

Figure 5.47. ^{1}H NMR spectrum of A_2I-2 (0.75 mM) with Kme3GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at room temperature. 138

Figure 5.48. ^{1}H NMR spectrum of A_2I-2 (0.75 mM) with Kme3GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at 40 ^{o}C. 139

Figure 5.49. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 1). 142

Figure 5.50. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 2). 143

Figure 5.51. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 3). 144

Figure 5.52. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 1). 145

Figure 5.53. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 2). 146

Figure 5.54. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into A_2I-2 (79 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 3). 147

Figure 5.55. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 μM) titrated into A_2I-2 (20 μM) at 25 ^{o}C in 10 mM sodium borate buffer pH 8.5 (Run 1). 148

Figure 5.56. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 μM) titrated into A_2I-2 (20 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 2)..... 149

Figure 5.57. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 μM) titrated into A_2I-2 (20 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 3)..... 150

Figure 5.58. One trial of three of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 μM) titrated into A_2I-2 (20 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 1)..... 151

Figure 5.59. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 μM) titrated into A_2I-2 (20 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 2)..... 152

Figure 5.60. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 μM) titrated into A_2I-2 (20 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 3)..... 153

Figure 5.61. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (225 μM) titrated into A_2I-2 (23 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 4)..... 154

Figure 5.62. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 μM) titrated into A_2I-1 (81 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 1)..... 155

Figure 5.63. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 μM) titrated into A_2I-1 (81 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 2)..... 156

Figure 5.64. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 μM) titrated into A_2I-1 (81 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 3)..... 157

Figure 5.65. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (900 μM) titrated into A_2I-1 (81 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 1)..... 158

Figure 5.66. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (700 μM) titrated into A_2I-1 (70 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 2)..... 159

Figure 5.67. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (700 μM) titrated into A_2I-1 (70 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5 (Run 3)..... 160

Figure 5.68. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (200 μM) titrated into A_2I-1 (20 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5 (Run 1)..... 161

Figure 5.69. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (200 μM) titrated into A_2I-1 (20 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5 (Run 2)..... 162

Figure 5.70. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 μM) titrated into A_2I-1 (20 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 163

Figure 5.71. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 μM) titrated into A_2I-1 (20 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 164

Figure 5.72. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (817 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 165

Figure 5.73. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (817 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 166

Figure 5.74. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (800 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 167

Figure 5.75. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (800 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 168

Figure 5.76. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (815 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 169

Figure 5.77. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (815 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 170

Figure 5.78. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (812 μM) titrated into A_2J-1 (76 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 171

Figure 5.79. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (812 μM) titrated into A_2J-1 (76 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 172

Figure 5.80. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (849 μM) titrated into A_2J-2 (84 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 173

Figure 5.81. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (849 μM) titrated into A_2J-2 (84 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 174

Figure 5.82. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (689 μM) titrated into A_2J-2 (84 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 175

Figure 5.83. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (689 μM) titrated into A_2J-2 (84 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 176

Figure 5.84. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (807 μM) titrated into A_2J-2 (86 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 177

Figure 5.85. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (807 μM) titrated into A_2J-2 (86 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 178

Figure 5.86. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (808 μM) titrated into A_2J-2 (86 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 179

Figure 5.87. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (808 μM) titrated into A_2J-2 (86 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 180

Figure 5.88. One trial of two of Kme2GGY (Ac-Kme2GGY-NH_2) (867 μM) titrated into A_2I-2 (91 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 181

Figure 5.89. One trial of two of Kme2GGY (Ac-Kme2GGY-NH_2) (867 μM) titrated into A_2I-2 (91 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2)..... 182

Figure 5.90. One trial of two of Kme3GGY (Ac-Kme3GGY-NH_2) (883 μM) titrated into A_2I-2 (91 μM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1)..... 183

Figure 5.91. One trial of two of Kme3GGY (Ac-Kme3GGY-NH_2) (883 μM) titrated into A_2I-2 (91 μM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 2). 184

Figure 6.1. Gas phase molecular models of one isomer of A_2B-Asp (left) and A_2B-Trp (right) bound to butyldimethyl ammonium and butyltrimethyl ammonium respectively. 188

Figure 6.2. Overlaid DCLs of monomer A, B-Phe, B-Tyr, and B-Trp with different methylated butylamine guests (1 mM A, 1 mM amino acid-functionalized monomer, and 4 mM guest, 50 mM sodium borate buffer, pH 8.5). 191

Figure 6.3. 1D ^{1}H-NMR (400 MHz) of Trt B in $CDCl_3$. 197

Figure 6.4. 1D ^{1}H-NMR (400 MHz) of Trt B Osu in $CDCl_3$ 198

Figure 6.5. 1D ^{1}H-NMR (400 MHz) of Trt B-Phe in $CDCl_3$ 199

Figure 6.6. 1D ^{1}H-NMR (400 MHz) of Trt B-Trp in $CDCl_3$. 200

Figure 6.7. 1D ^{1}H-NMR (400 MHz) of Trt B-Tyr in $CDCl_3$. 201

Figure 6.8. 1D ^{1}H-NMR (400 MHz) of Trt B-Asp in $CDCl_3$. 202

Figure 6.9. 1D ^{1}H-NMR (400 MHz) of B-Phe in MeOD. 203

Figure 6.10. High resolution mass spectrum of B-Phe. 204

Figure 6.11. 1D ^{1}H-NMR (400 MHz) of B-Tyr in MeOD. 205

Figure 6.12. High resolution mass spectrum of B-Tyr 206

Figure 6.13. 1D ^{1}H-NMR (400 MHz) of B-Trp in MeOD. 207

Figure 6.14. High resolution mass spectrum of B-Trp. 208

Figure 6.15. 1D ^{1}H-NMR (400 MHz) of B-Asp in MeOD. 209

Figure 6.16. High resolution mass spectrum of B-Asp (ESI -). 209

Figure 6.17. Semi-preparative reversed-phase HPLC traces of the purification of A_2B-Phe from a preparative scale DCL. Absorbance measured at 214 nm. 211

Figure 6.18. High resolution mass spectrum of A_2B-Phe (ESI -). 212

Figure 6.19. Semi-preparative reversed-phase HPLC traces of the purification of A_2B-Tyr from a preparative scale DCL. Absorbance measured at 214 nm. 212

Figure 6.20. High resolution mass spectrum of A_2B-Tyr (ESI -). 213

Figure 6.21. Semi-preparative reversed-phase HPLC traces of the purification of A_2B-Trp from a preparative scale DCL. Absorbance measured at 214 nm. 213

Figure 6.22. High resolution mass spectrum of A_2B-Trp (ESI -). 214

Figure 6.23. Semi-preparative reversed-phase HPLC traces of the purification of A_2B-Asp from a preparative scale DCL. Absorbance measured at 214 nm. 214

Figure 6.24. High resolution mass spectrum of A_2B-Asp-1 (ESI -). 215

Figure 6.25. High resolution mass spectrum of A_2B-Asp-2 (ESI -). 215

Figure 6.26. Extinction coefficient determination of A_2B-Phe from the slope of the line of regression. 216

Figure 6.27. Extinction coefficient determination of A_2B-Trp from the slope of the line of regression. 217

Figure 6.28. Extinction coefficient determination of A_2B-Tyr from the slope of the line of regression. 217

Figure 6.29. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into A_2B-Phe (82 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 220

Figure 6.30. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into A_2B-Phe (82 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 221

Figure 6.31. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into A_2B-Phe (82 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 222

Figure 6.32. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into A_2B-Tyr (98 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 223

Figure 6.33. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into A_2B-Tyr (98 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 224

Figure 6.34. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into A_2B-Tyr (98 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 225

Figure 6.35. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into A_2B-Trp (97 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 226

Figure 6.36. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into A_2B-Trp (97 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 227

Figure 6.37. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into A_2B-Trp (97 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 228

Figure 6.38. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.94 mM) titrated into A_2B-Asp-1 (85 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 229

Figure 6.39. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.04 mM) titrated into A_2B-Asp-1 (85 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 230

Figure 6.40. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.13 mM) titrated into A_2B-Asp-1 (85 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 231

Figure 6.41. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.94 mM) titrated into A_2B-Asp-2 (81 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 232

Figure 6.42. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.04 mM) titrated into A_2B-Asp-2 (81 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 233

Figure 6.43. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.13 mM) titrated into A_2B-Asp-2 (81 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. 234

LIST OF SCHEMES

Scheme 3.1: Synthesis of CX4-ONBD Probe..... 25

Scheme 3.2. Synthesis of CX_4-ONBD Probe 5. 33

Scheme 5.1: Synthesis of Monomers I and J. 95

Scheme 5.2. Synthesis of Monomer J 104

Scheme 5.3. Synthesis of Monomer I 111

Scheme 6.1. Synthesis of Trt B Osu. 188

Scheme 6.2. Synthesis of B-Phe, B-Tyr, B-Trp, and B-Asp. 189

Scheme 6.3. Synthesis of Trt B Osu 195

Scheme 6.4. Synthesis of B-Phe, B-Tyr, B-Trp, and B-Asp. 195

LIST OF ABBREVIATIONS

Ac	Acetyl
ACN	Acetonitrile
AcOH	Acetic Acid
Ala, A	Alanine
Arg, R	Arginine
Boc	Tert-butyloxycarbonyl
$CHCl_3$	Chloroform
COSY	Correlation Spectroscopy
CuAAC	Copper catalyzed azide-alkyne cycloaddition
CX4	Calix[4]arene
DABCO	N,N-diazabiocyclo[2.2.2]octane
DCC	Dynamic Combinatorial Chemistry
DCL	Dynamic Combinatorial Library
DCM	Dichloromethane
DIPEA	Diisopropylethylamine, Hünig's Base
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
ESI	Electrospray Ionization
EtOAc	Ethyl Acetate
Fmoc	Fluorenylmethoxycarbonyl
Gln, Q	Glutamine

Glu, E	Glutamic Acid
Gly, G	Glycine
H2A/B	Histone 2 A/B
H3	Histone 3
H4	Histone 4
HBTU	N,N,N’N’-Tetramethyl-O-(1H-benzotriazol-1-yl)uranium hexafluorophosphate
HDAC	Histone Deacetylase
HP1	Heterochromatin Protein 1
HPLC	High Performance Liquid Chromatography
HOBt	Hydroxybenzotriazole
HR ESI-MS	High Resolution Electrospray Ionization Mass Spectrometry
ITC	Isothermal Titration Calorimetry
K_d	Dissociation Constant
Kme	Monomethyllysine
Kme2	Dimethyllysine
Kme3	Trimethyllysine
LCMS	Liquid Chromatography Mass Spectrometry
Leu, L	Leucine
Lys/K	Lysine
m/z	Mass to Charge Ratio
MALDI	Matrix-Assisted Laser Desorption Ionization
MeOH	

MTBD	7-methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene
MS	Mass Spectrometry
NBD	Nitrobenzoxadiazole
NH_4OAc	Ammonium Acetate
NHS	N-hydroxysuccinimide
NMR	Nuclear Magnetic Resonance
ONBD	O-aryl-nitrobenzoxadiazole
PEG	Polyethylene Glycol
PTM	Post-translational Modification
RP-HPLC	Reversed-Phase High Performance Liquid Chromatography
SAHA	Suberoylanilide hydroxamic acid or Vorinostat
SAM	S-adenosyl-L-methionine
Ser, S	Serine
tBu	Tert-butyl
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic Acid
Thr, T	Threonine
TIPS	Triisopropylsilane
TLC	Thin Layer Chromatography
Trp, W	Tryptophan
Tyr	Tyrosine

CHAPTER 1. Approaches for Site-Selective Labeling of Endogenous Peptides and Proteins

Introduction

The selective modification of proteins is an important aspect of biology. It allows for extraordinary levels of chemical diversity to arise from a limited pool of proteinogenic amino acids. In nature, this diversity is achieved primarily through the use of post-translational modifications (PTMs): covalent chemical modifications made to residues in proteins after translation from the ribosome.[1] Common modifications include methylation,[2] acetylation,[3] phosphorylation,[4] ubiquitination,[5] and glycosylation.[6] These PTMs often allow for precise control of the structure and function of proteins and are crucial to normal biological function. In the field of chemical biology, site-specific labeling of amino acid residues in proteins is of extreme interest as it allows for interrogation of protein interactions and function and for manipulation of a protein's properties for novel and creative applications.

Several approaches have been developed for artificial chemical protein modification. One common method involves genetic manipulation to incorporate unnatural amino acids into proteins, a process pioneered by the Schultz lab.[7,8] This approach has been used to incorporate many unnatural and often biorthogonal functional groups into proteins including azides, alkynes, ketones, fluorophores, crosslinking groups, and many others. While unnatural amino acid mutagenesis is a powerful tool for site-selective modification, there are some downsides to this approach. It requires prior knowledge of the target for modification, which can be limiting depending on the desired application. It also does not allow for protein function to be studied

under natural protein expression levels and is often limited to model organisms. Native chemical ligation (NCL) is another approach that has been used for site-selective labeling of proteins, specifically at the N-terminus.[9] NCL allows for the synthesis of proteins from unprotected polypeptide fragments. One fragment contains an N-terminal cysteine, and the other contains a C-terminal α-thioester. Ligation occurs through a chemoselective trans-thioesterification followed by an S-to-N acyl shift to generate a native peptide bond.[10] A related approach, expressed chemical ligation, uses recombinant proteins containing α-thioesters to react with synthetic peptides containing an N-terminal Cys.[11] These recombinant thioesters are generated from inteins (polypeptide segments that are self-excised by formation of a thioester intermediate) that are incorporated into the protein sequence.[12] While NCL and expressed chemical ligation are powerful strategies for the synthesis of homogeneous, site-specifically modified proteins, they are not well suited for labeling endogenous proteins. The approaches that this chapter will focus on are those that involve chemical modifications of endogenous proteins containing natural amino acids. Predictably, artificial site-selective modification of endogenous amino acids is a challenging task. However, it is a worthwhile endeavor as it allows protein function to be interrogated under ideal, native conditions.

Chemical Reactions for Modification of Endogenous Amino Acids

The chemical reactions used for the artificial modification (or labeling) of endogenous proteins have several basic requirements, especially if the labeling is to occur in live cells.[13] First, reactions must occur in water or under predominantly aqueous conditions, as many organic solvents are toxic to cells. The reactions also should operate near physiological pH and under a physiological temperature range. The reaction must not use toxic reagents and should work at

low concentrations. The reaction kinetics should also be adapted to fit the desired application (many applications require rapid labeling rates).

Two of the most commonly modified amino acids are lysine and cysteine. Lysines are popular targets for modification as they are fairly abundant and often solvent exposed.[13,14] Cysteine is often a preferred target due to its exceptional nucleophilicity under physiological conditions. Several classical reactions (Figure 1.1) have been used to modify lysine and cysteine residues in proteins. N-hydroxysuccinimidyl esters, isocyanates, and isothiocyanates have been used to directly modify lysines through nucleophilic addition and elimination to give amide, urea, and thiourea conjugated products respectively.[13,15] Aldehydes have also been used to modify lysine residues though reductive amination to give amine linkages.[16] While electrophilic modification of lysines still finds common use, most of these classical reactions involve reagents that are non-selective and too reactive for use in labeling of proteins in cells. Traditionally, cysteine residues have also been modified using electrophilic reagents like maleimides and alpha-halocarbonyls.[17,18] Despite their widespread use, these chemical modifications of Cys are limited by formation of an unstable linkage and non-specific reactivity depending on the conditions.[19] Disulfides have also been used to target cysteine residues,[20] but the resulting mixed disulfide is still susceptible to oxidation in air and is easily reduced in a cellular environment.

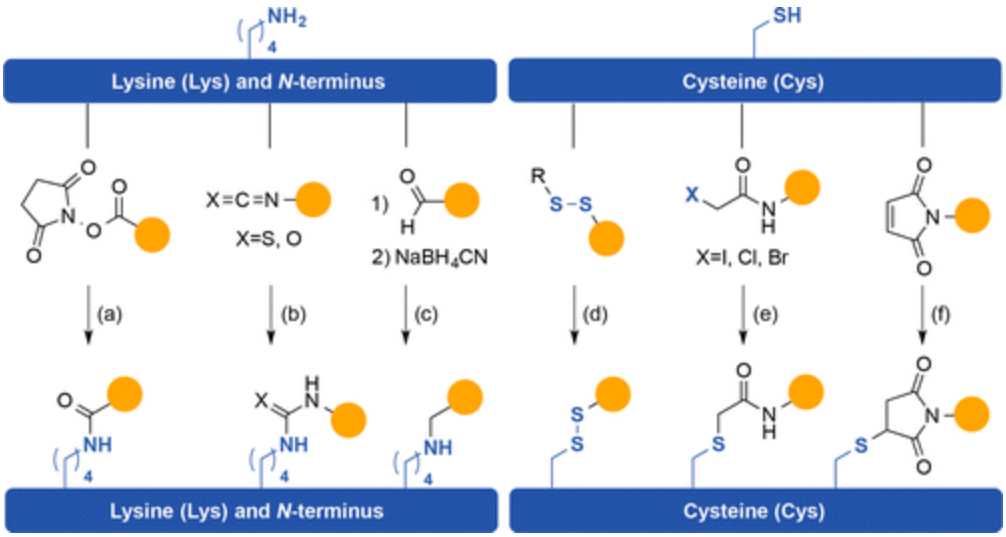


Figure 1.1. Classical reactions used in modifying lysine and cysteine residues. Reproduced with permission from the American Chemical Society: Chem. Rev. **2015**, 115, 2175.
DOI: 10.1021/cr500399p. This is an open access article published under a Creative Commons Attribution (CC-BY) License.

Several more recent chemistries have been developed that have vastly improved selectivities compared to the previously mentioned traditional reactions, making them far more amenable for use in chemical modification of proteins in live cells. One effective strategy for highly selective modification of lysine residues involves the reaction of lysine with an α,β,γ,δ-unsaturated aldehyde.[21] Imine formation followed by rapid aza-electrocyclization produces a stable, aromatic pyridinium conjugate. The reaction has been used to modify proteins *in vivo* with fluorescent dyes and PET imaging agents and shows remarkable selectivity for targeting solvent exposed lysines rather than interior lysines or the N-termini. Another selective reaction that has been used to modify lysines is by reaction with 2-formylphenyl boronic acids or 2-acetylphenyl boronic acids to form stable iminoboronate conjugates.[22,23] This reaction occurs readily under physiological conditions and is reversible by the addition of dopamine, fructose, or glutathione. For selective chemical modification of the N-terminus of proteins and peptides,

ketenes have been utilized.[24] At pH 6.3, this reaction allows for almost perfect selectivity in labeling the N-terminus versus lysine side chains.

Several selective cysteine modifications that produce relatively stable conjugates have been developed recently. One modification involves the use of 3-arylpropiolonitriles as Michael acceptors to rapidly produce conjugates with cysteine that exhibit remarkable stability to hydrolysis.[25] Hydroxymethyl naphthols have also been used to generate sulfhydryl-reactive Michael acceptors *in situ* upon irradiation with UV light.[26] The resulting 2-methylenecyclohexanone type intermediate reacts rapidly with cysteine to form a stable thioether bond. Radical initiated reactions such as the thiol-ene and thiol-yne have also bene used to modify cysteines.[27–29] In the case of thiol-ene chemistry, a stable thioether product is formed. For thiol-yne modifications a vinyl sulfide is formed initially, which can be reacted during a subsequent labeling step to produce a stable 1,2-dithioether.

Tyrosine and tryptophan residues, which are less abundant than lysine and are not involved in structure-stabilizing bonds like cysteine,[13] have also been modified selectively using a variety of approaches. A bench-stable formylbenzene aryl diazonium hexafluorophosphate has been used to directly modify tyrosine,[30] removing the need for *in situ* preparation from aniline under acidic conditions and allowing for a secondary functionalization step with the aldehyde to install a tag unit of interest. A transitional metal catalyzed approach has also been developed for site-selective modification of tyrosine: palladium catalyzed allylic O-alkylation using an allylic carbamate and the water-soluble phosphine TPPTS.[31] This modification occurs under mild conditions and has been used to prepare synthetic lipoproteins. Another reaction that has been used to label both tyrosine and tryptophan in proteins is the 3-component Mannich reaction.[32] This reaction requires the addition of an aniline and formaldehyde to give a stable amino-

alkylated product and has been used to prepare spin-labeled proteins for characterization of structural transitions. For targeting tryptophan residues, many of the existing strategies require toxic transition metal and harsh acidic conditions that limit their utility. However, an approach involving N-oxyl 9-azabicyclo[3.3.1]nonane organoradical reagents has been used to selectively modify tryptophan under mild aqueous conditions to produce a conjugates that are stable under acidic, basic, and reducing conditions.[33]

Activity Based Protein Profiling

While the previously mentioned chemical reactions can be applied broadly to any protein that contains a certain amino acid residue in sterically unhindered position, more sophisticated strategies have been developed for targeting and modifying specific proteins (generally enzymes) based on molecular recognition and catalytic activity. One approach that has seen frequent use for enzymes is activity-based protein profiling (ABPP).[34] ABPP uses carefully designed chemical probes to bind and covalently label at or near the active site of an enzyme. These probes typically contain several common components: a suicide inhibitor (frequently a substrate analog with a reactive electrophile or coordinating group), a linker, and a tag unit (alkynes or azides for conjugation in a secondary step,[35] fluorescent dye for visualization, biotin for pulldown, radiolabel, or NMR-active nuclei like fluorine).[36] Many ABPP probes also contain photo-crosslinking groups like benzophenones, phenyl azides, and diazirines to covalently tether the protein of interest to the probe.[37–40] Several classes of enzymes have been characterized using activity-based protein profiling. Metalloproteases have been targeted using ABPP probes[41] containing a hydroxamic acid inhibitor that coordinates to zinc in the active site, a photoaffinity group for covalent tethering, and an alkyne handle to click on a fluorophore. Several kinases have been profiled using chemical probes[42] that utilize ATP or ADP analogs conjugated to a tag

units, while phosphatases have been labeled and characterized using probes[43] with self-immolative phospho-tyrosine substrates that form electrophilic quinone methide reactive groups upon enzymatic cleavage of the phosphate group. In general, proteases have been subjected to ABPP by using probes with a reactive group (such as an epoxide or a phosphonate electrophile)[44,45] linked to a known peptide substrate, which is then linked to a tag unit.

Ligand-Directed Affinity Labeling

Another promising strategy for the labeling and modification of endogenous proteins is ligand-directed affinity labeling.[36,46] This strategy, pioneered by the Hamachi lab, involves the use of high affinity ligands to bring a connected reactive group in close proximity to a nucleophilic amino acid residue on the protein. Ligand directed affinity labeling typically employs two types of chemical probes and can covalently modify a protein through one of two types of labeling mechanisms. One type of probe has a high affinity ligand (typically for an enzyme) connected through a spacer to electrophilic reactive group which is connected to a tag unit (fluorophore, pulldown handle, etc.).[15,36,46–49] The second type of probe has the high affinity ligand connected through a linker to a catalyst that activates a particular amino acid functional group on the protein and requires the use of a secondary probe to label the protein with a tag unit.[50,51] In terms of labeling mechanisms, one possibility is to design probes where the entire probe remains tethered to the protein target after covalent reaction with the reactive group. However, this results in inhibition of the protein or enzyme. Another possibility is a traceless approach where reaction of a nucleophile on the protein with the probe installs the tag unit and displaces the high affinity ligand. The second labeling mechanism is often preferred as it allows an enzyme to be labeled with a unique tag unit without permanently inhibiting the activity of the enzyme. One of the most widely explored chemistries for ligand directed affinity labeling is

ligand-directed tosylate (LDT) chemistry where a ligand for the protein is linked to tag unit through a tosylate reactive group which reacts via an SN2 type mechanism.[46,52] LDT probes have been used label endogenous proteins in live cells and have successful been used to re-engineer proteins to serve as reporters of protein-protein interactions.[53–55] The Hamachi lab has also developed several other chemistries including ligand-directed acyl imidazole labeling,[47] affinity guided DMAP (AGD) catalysis,[50] and ligand-directed dibromophenyl benzoate (LDBB) labeling.[48] LDAI has been applied to the labeling of both intracellular and cell surface protein targets and has been used to engineer light-activatable caged enzymes.[56] AGD has been used to label cell surface protein receptors in a highly efficient manner, while LDBB has been applied to site-selective labeling of purified proteins and intracellular protein targets.

Others have also developed ligand-directed affinity labeling approaches that rely on an SN2 type mechanism. Fenical and co-workers developed ligand-directed acyl phenol probes in order to elucidate the intracellular target of a marine natural product.[49] Following a similar goal of validating a drug target, the Kakeya group has also prepared 5-sulfonyl tetrazole probes of cyclosporine A to label cyclophilin A intracellularly.[57]

Ligand-directed affinity labeling chemistries based on other reaction mechanisms have also been developed. The Sodeoka group developed a SNAr approach using probes with high affinity ligands attached to a bifunctional O-aryl nitrobenzoxadiazole (ONBD).[58] Since the ONBD probe is dark and the NHNBD probe is fluorescent, reaction of the probe with a nucleophilic lysine that is close proximity produces a turn-on fluorescence response (Figure 1.2). Another creative ligand-directed labeling approach was developed by the Nakamura lab where single electron transfer catalysts were used for chemical modification of proteins.[51] Probes containing a high affinity ligand were linked to a ruthenium tris(bipyridine) catalyst were used to

react with tyrosine residues to form a reactive tyrosyl radical, which can be trapped with a phenylenediamine tag unit to produce a carbon-carbon bonded conjugate.

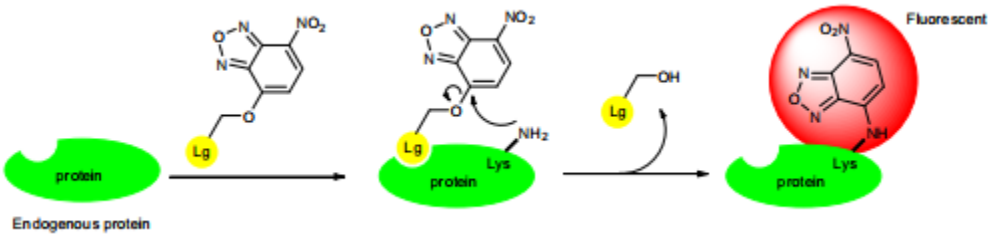


Figure 1.2. Ligand-directed ONBD chemistry. Reproduced with permission from Elsevier: Tetrahedron Lett., **2015**, 56, 2130. https://doi.org/10.1016/j.tetlet.2015.03.065. This article is published under the terms of the Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND).

CHAPTER 2. Approaches for Detection of Histone Post-Translational Modifications

DNA Organization and Chromatin Structure

DNA, commonly referred to as the blueprint of life, provides the genetic information necessary to construct every component of the cell and to regulate all functions necessary for life including cell growth, differentiation, and homeostasis. Over 223 million base pairs of DNA are contained within the largest human chromosome (chromosome 1).[59] If the DNA in chromosome 1 were extended out linearly, it would be close to 80 millimeters long.[60] In addition, each cell contains around 6 billion base pairs of DNA; if all the DNA in a single cell was measured from end to end, it would correspond to a length of 2 meters.[61] Since the average eukaryotic nucleus is only 6 micrometers in diameter, there must exist some way to package the DNA into smaller, higher order units so that it will fit inside the nucleus of the cell. This packaging and organization is accomplished through the use of chromatin. Each chromosome is made up of chromatin, and each unit of chromatin is made up of fundamental repeating units called nucleosomes. A nucleosome consists of 146 base pairs of DNA wrapped around eight proteins called histones.[62,63] Histones are a group of highly positively charged proteins that bind and coil negatively charged DNA around themselves. There are four core histone proteins: H2A, H2B, H3, and H4. Two units of each core histone protein dimerize to form an octamer around which 146 base pairs of DNA wrap.[64] Histone H1 serves as a linker between nucleosomes and stabilizes the higher order structure of chromatin.[65] The structure of each nucleosome can be thought of as a thread and spool, where the DNA is the thread that is tightly wound around the histone spool.

The nucleosomes fold on top of one another in a tight coil to form a fiber of chromatin that is 30 nm in diameter.[65] These packed chromatin fibers then form 300-nm long loops, which are folded even further to form a fiber that is 250 nm in width.[66] This 250-nm wide fiber is wound tightly to create a single chromatid of the chromosome. The chromosome itself is 1400 nm wide, which can be efficiently packaged within the eukaryotic nucleus.[66]

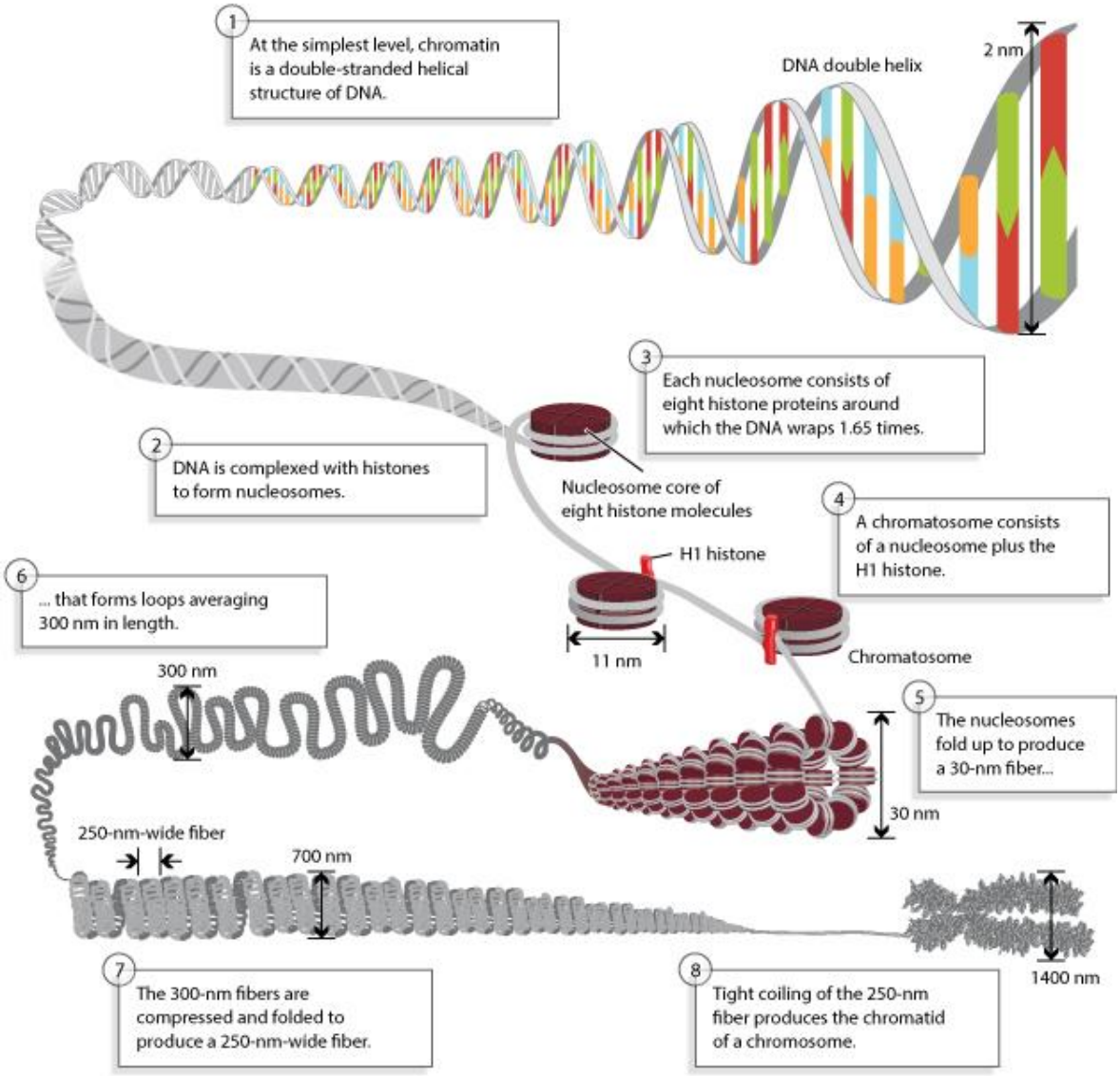


Figure 2.1. Packaging of chromosomal DNA by winding around histone proteins. Reproduced with permission from Nature Publishing Group: Nature Education, **2008**, 26, 1.

Covalent Modification of Histone Proteins

The higher order folding of nucleosomes into tightly coiled fibers leads to a form of chromatin called heterochromatin.[67] In this form of chromatin, there is tight association between DNA and the histone proteins which blocks enzymes and transcription factors from binding to the DNA and prevents transcription from taking place. One of the primary ways by which the DNA is made accessible for transcription is through covalent (most commonly enzymatic) modification of the histone proteins. These modifications, referred to as post-translational modifications or PTMs, occur after translation from the ribosome and provide a form of epigenetic control since it is independent of DNA sequence.[67] Each of the four core histones also contain unstructured peptide tails that extend from their globular domain. These unstructured tails can be chemically modified in order to alter the structure of chromatin and the accessibility of DNA within the gene complex. Some PTMs can lead to a state of chromatin called euchromatin where the DNA is partially dissociated from the histone complex allowing transcription factors to bind and transcription to occur.[67]

A wide variety of covalent modifications are made to amino acids on the tails of histone proteins including methylation,[2] acetylation,[3] phosphorylation,[4] ubiquitination,[5] and glycosylation.[6] The process of post-translational modification can be thought of in terms of writers, erasers, and readers. Writers are the enzymes (methyltransferases, acetyltransferases, kinases, ubiquitin ligases) responsible for installing the covalent mark, while erasers are the enzymes (demethylases, deacetylases, phosphatases, deubiquitinases) that remove the PTM. Readers refer to the proteins that contain a domain that allows the protein to bind to a PTM or a pattern of PTMs in order to manipulate the structure of chromatin and affect gene expression.

Histone PTMs are often categorized are being either activating or deactivating (or silencing). Histone H3 is the most extensively modified, and many of the marks have been well characterized. Expectedly, the majority of lysine acetylations and arginine citrullinations are activating, as the positive charge on the histone is reduced; this reduces the electrostatic interaction with the negatively charged DNA leading to dissociation from the gene complex. In addition to neutralizing charge, acetylated Lys also plays an active role in controlling gene expression via recognition by bromodomain reader proteins, which are part of larger protein complexes involved in transcription.[68] Methylation of lysine and arginine can be activating or repressive. Lysine has three methylation states: monomethyllysine (Kme), dimethyllysine (Kme2), and trimethyllysine (Kme3); arginine also has three methylation states: monomethylarginine, symmetric dimethylarginine (SDMA), and asymmetric dimethylarginine (ADMA). One of the hallmarks of active transcription is trimethylation of H3K4 (the lysine in the fourth position of histone H3), a mark that leads to the formation of euchromatin through the recruitment of chromosome remodeling proteins like CHD1 and BPTF.[69,70] Conversely, H3K9me3 is the trademark of repressive histone PTMs. This mark binds to heterochromatin protein 1 (HP1) and stabilizes the heterochromatin structure, which silences gene transcription.[71] For methylated arginine, active marks include H3R17me2a, H3R26me2a, H4R3me2a, and H2AR3me2a, while inactivating marks include H3R2me2a, H3R8me2s, H4R3me2s, and H2AR3me2s.[72,73]

Histone phosphorylation and ubiquitination also play important roles in controlling chromatin structure and regulating important nuclear processes. Several phosphorylation marks on histone H3 (H3S10, H3T11, and H3S28) have been found to coincide with histone acetylation and are thereby associated with activation of gene transcription.[4] Phosphorylation of H3S10,

H3S28, and H2AT120 are key PTMs for mitosis as they serve as the driving force for the compaction and condensation of chromatin during division of the nucleus.[74] Histone phosphorylation also plays a role in DNA repair as the H2AXS139ph mark is involved in recruiting DNA damage repair proteins to repair double-strand DNA breaks.[75] Similar to phosphorylation, histone ubiquitination (such as polyubiquitination of H2AK36) also plays a role in DNA damage repair.[5] Ubiquitination is also a mechanism by which gene transcription can be regulated. Monoubiquitination of K119 on H2A is involved in repressing transcription while monoubiquitination of K123 on H2B is transcriptionally activating.[5]

Aberrant PTMs and Disease

Histone post-translational modifications are crucial for the normal development and function of cells and tissues. When normal processes involving PTMs become dysregulated (such as overexpression or mutation of the enzymes involved in installing or removing a modification), the effects are often deleterious. Many cancers are characterized by the disruption of natural epigenetic regulatory mechanisms. For example, loss of acetylation at H4K16 and trimethylation at H4K20 and commonly observed across many cancer cell lines including colon, breast, and lung cancer cell lines.[76] Additionally, aberrant methylation of arginine residues is involved in blood cancers. More specifically, increased asymmetric dimethylation of H4R3 by the enzyme PRMT1 leads to overexpression of various oncogenes that are implicated in acute myeloid leukemia.[77] Several neurological disorders also arise from dysregulation of histone PTMs. Hypertrimethylation of H3K9 and histone hypoacetylation are observed in patients with Huntington's disease, Friedreich's ataxia, and Rubinstein-Taybi syndrome.[78]

Current Methods to Detect Histone PTMs

The importance of histone PTMs in regulating various biological functions (including gene transcription, modification of chromatin structure, and DNA repair) and their role in the progression of many diseases necessitates the development of robust and sensitive tools for their detection. Currently, antibodies are the workhorse for many of the techniques used in the detection of histone post-translational modifications, with over one thousand antibodies with specificity for histone PTMs available for commercial use.[79] The most commonly used methods for detection of histone PTMs are western blotting and mass spectrometry-based (MS-based) proteomics. In the analysis of PTMs from cellular samples, both methods typically require an enrichment step prior to analysis since many post-translational modifications are present in very low natural abundance. This enrichment is accomplished using techniques such as immunoprecipitation (IP), ionic interaction, or chemical derivatization.[80] For IP, antibodies that bind to a specific PTM are immobilized on a solid support such as agarose beads. The cell lysate is then incubated with the affinity beads. Any proteins that do not contain the PTM of interest are washed off, and afterwards the desired modified protein is isolated. An affinity enrichment strategy used specifically for phosphopeptides is immobilized metal affinity chromatography (IMAC), which is based on the ionic interaction of the phosphate group for iron ions.[81] Resin with immobilized Fe(III) is used to pulldown phosphopeptide by ion exchange and coordination to the iron. Chemical derivatization-based enrichment has been used for analysis of glycosylated peptides. Oxidation of the carbohydrate residues using sodium periodate generates aldehydes and allows the peptide or protein to be immobilized using hydrazide functionalized beads.[82] The glycopeptide or glycoprotein can finally be release from the resin via enzymatic reaction. After enrichment of the peptide or protein of interaction, analysis can be performed using Western

blotting or MS-based proteomics. For Western blotting, the sample is subsequently separated by SDS-PAGE and subjected to immunostaining to visualize proteins containing the PTM. For MS proteomics, a bottom-up workflow is generally used. After PTM-specific protein enrichment by IP, the protein is digested into smaller peptide fragments using enzymes that predictably cleave after certain residues. The peptide fragments are then separated using reversed-phase HPLC before analysis by MS/MS.

Both Western blotting and MS-based proteomics have several advantages for use in detecting PTMs. Western blotting has the benefit of being a very commonplace technique in biology that does not require any special equipment or expertise; MS-proteomics is a powerful detection method that offers the potential for unbiased, site-specific identification of histone PTMs.[83] In spite of these advantages, both methods have limitations. Some of these limitations are related to the use of antibodies in either the enrichment step (for proteomics) or in the final detection step (for Western blotting). Antibodies are often too sequence specific and may not recognize its target if there is a neighboring PTM which can lead to false negatives.[84] Batch-to-batch variability of antibody affinity and specificity is another common complication.[85] Additionally, the is no universally approved method for validating antibody, and an antibody that works well for one application may not work well in another application.[86] Western blotting suffers from the fact that it does not allow for the site of modification to be determined when pan-specific antibodies are used. For MS-based proteomics, the requirement of a mass spectrometer and expertise in operation of the spectrometer can be limiting factors. Also, extensive sample purification, enrichment, and preparation are essential for detecting low-abundant PTMs by MS.[87] Use of antibodies for enrichment often results in contamination of the sample with antibody fragments due to the harsh conditions needed to release the bound

peptides.[83] Furthermore, analytical challenges that arise from issues like poor ionization and solubility may hinder detection of a PTM by MS. Lastly, MS is low throughput.

As a complement to Western blotting and MS-based proteomics, several methods based on chemical labeling have been developed for detecting histone PTMs. Direct covalent modification of citrulline has been developed to characterize citrullination using a phenylglyoxal-based probe.[88] Incubation of peptides or proteins containing citrulline with the Rhodamine-conjugated-phenylglyoxal probe results in a chemoselective cyclization to form an imadazol-2-one adduct. The labeled targets can be visualized by in-gel fluorescence SDS-PAGE. For detecting methylated lysine by chemical means, a popular approach has been to engineer methyltransferase enzymes to accept unnatural S-adenosylmethionine (SAM) analogues in order to transfer alkynes or azides instead of methyl groups. Bertozzi and Gozani developed an alkyne-containing SAM cofactor capable of transferring an alkyne to substrates of a lysine methyltransferase.[89] After completion of the enzymatic reaction, click chemistry was label the alkyne-tagged protein with a FLAG tag. The Luo group improved on this approach by preparing a Se-Adenosyl-L-selenomethionine cofactor that has greater stability under physiological conditions.[90] Luo and colleagues also generated azide-containing SAM analogs for biorthogonal profiling of protein methylation.[91]

The development of small molecule receptors for the recognition of PTMs presents several advantages that could augment strategies using antibodies and mass spectrometry-based proteomics. Small molecules have much lower cost of production when compared to antibodies and can be synthesized reliably and reproducibly. Receptors of this kind also have the potential to bind to a PTM of interest in a non-sequence-specific manner due to its small size.

Several small molecule receptors that bind methylated lysine and arginine have been prepared using traditional rational design. One of the first reports[92] came from the Hof group who have developed several receptors based on a p-sulfonatocalix[4]arene receptor that bind Kme3 with low micromolar affinity.[93] More recently, tetracyanoresorcin[4]arene cyclophanes have been synthesized by Hamilton and co-workers that also bind to Kme3 with similar affinities as the calixarene receptors.[94] Cucurbiturils have also been shown to bind methylated lysine (CB-6) and methylated arginine (CB-7).[95]

The Waters lab has also developed small molecule receptors for methylated lysine and arginine by using a selection approach called dynamic combinatorial chemistry (DCC).[96–99] In developing small molecule receptors for molecular recognition of PTMs, a selection approach provides an advantage over traditional rational design as it allows evolution to determine the best receptor. DCC is an effective selection approach first reported by Otto and Sanders that can be used to select the best receptor for a guest from a library of receptors.[100,101] This technique allows macrocyles to be created from reversible exchange reactions of monomers under equilibrium conditions. If a guest with a PTM of interest is introduced in a DCC library (DCL), a host that forms favorable binding interactions with guest may be amplified. Rigid monomers with aromatic groups for molecular recognition through cation-π interactions, carboxylates for water solubility, and thiols for reversible disulfide exchange were used to generate mercaptophanes capable of binding methylated lysine and arginine with high affinity and impressive selectivity over unmethylated lysine.

Several of these receptors have been applied toward the development of new supramolecular methods for detecting PTMs. Hof, Waters, and Hooley have used synthetic receptors to create sensor array platforms for pattern recognition of combinations of PTMs on

histone peptides (known as the histone code).[102–104] The sensor array platforms are based on indicator displacement assays, where a small molecule receptor binds and quenches the fluorescence of a dye, and addition of a PTM-containing histone peptide displaces the dye which generates a turn-on fluorescence response. Small molecule receptors have also been used to create affinity reagents for enrichment applications. The Hof group has developed a supramolecular affinity chromatography approach where a calixarene receptor that binds methylated lysine is attached to a solid support. This approach generates "methyl affinity columns" that can enrich methylated peptides from a mixture of enzymatically digested proteins.[105] The Kanai group has also utilized small molecule receptors for affinity enrichment by synthesizing a multivalent calixarene-based affinity reagent that contains a biotin handle. The affinity reagent was able to pull-down methylated histone peptides, which could be detected through Western blotting.[106]

Significance of This Work

Significant progress has been made over the past decade in the detection of post-translational modifications as there has been substantial improvements in proteomic mass spectrometry and methods involving antibodies. However, the challenges and limitations that are still faced by antibodies and MS-based proteomics highlight the need for alternative approaches such as chemical derivatization and small molecule receptors. Chemical methods for selective detection of peptides and proteins containing PTMs are still lacking. The research detailed in this dissertation expands the toolkit available for detection of three important histone PTMs: trimethyllysine, acetylated lysine, and dimethyllysine. Chapter 3 describes the detection of trimethyllysine using an approach termed supramolecular affinity labeling that combines

molecular recognition by a small molecule receptor with chemical derivatization. Detection of acetylated lysine is presented in Chapter 4 with the detection of histone deacetylase (HDAC) activity using a supramolecular affinity labeling probe. Finally, the design and synthesis of new synthetic receptors for dimethyllysine and trimethyllysine using iterative redesign by dynamic combinatorial chemistry is discussed in Chapter 5 and Chapter 6.

CHAPTER 3. Supramolecular Affinity Labeling of Histone Peptides[1]

Introduction and Significance

Methylation of Lys and Arg in histone proteins is involved in many vital biological processes that regulate the dynamic structure of chromatin and alter the interactions of DNA with protein complexes. Accordingly, there is great interest in mapping out the epigenetic role of these post-translational modifications (PTMs) in development and disease.[87,107,108] Two of the most established methods to detect post-translational modifications of proteins include antibodies and mass spectrometry-based proteomics. While these methods are staples of PTM research, both face challenges and limitations. Antibodies often exhibit high sequence specificity, which complicates the detection of unknown PTMs.[86,109] Additionally, some antibodies suffer from irreproducibility from batch to batch and have been shown to have cross-reactivity with off-target antigens.[84] MS-based proteomics has been used to explore a number of PTMs including methylation of lysine and arginine[110–113], acetylation of lysine[114,115], and phosphorylation of serine and threonine[116]. However, MS analysis is expensive and time consuming and can be complicated by analytical challenges including low natural abundance of PTMs, the presence of complex mixtures, and difficulties in analyzing large ionized proteins.[117]

---

[1] Reprinted (adapted) with permission from Gober, I. N.; Waters, M. L. *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459. Copyright (2016) American Chemical Society.

Innovative tools for detecting PTMs would augment existing detection methods and expand the chemical tool kit available for epigenetics research. New chemical methods have the potential to further our understanding of how these modifications regulate gene expression and how the PTMs themselves are regulated. Several non-covalent, supramolecular approaches have been developed using small molecule receptors for recognition of PTMs. Calixarenes,[102] mercaptophans,[103] and cavitands[104] have been used to create sensor arrays that are based on an indicator displacement scheme. In this scheme, a small molecule receptor binds and quenches the fluorescence of a dye, and addition of a guest (i.e. PTM-containing analyte) displaces the dye which generates a turn-on fluorescence response. Others have developed chemical approaches based on site-selective covalent labeling of proteins or peptides containing PTMs[88,90] For example, direct covalent modification of citrulline has been developed to characterize citrullination.[88] For trimethyllysine, indirect labeling methods are needed since it cannot be directly covalently labeled. A popular approach has been to engineer methyltransferase enzymes to accept unnatural SAM analogs in order to transfer alkynes[89] or azides[91] instead of methyl groups, but this approach introduces additional challenges, such as differences in enzyme activity, cross-reactivity with native SAM, and validation of native substrate scope.

Taking inspiration from ligand-directed affinity labeling chemistries developed by Hamachi et al.[52,50,47] that demonstrate that high-affinity ligands with cleavable linkers can be used to promote traceless labeling of their target proteins, we looked to reverse the paradigm and use synthetic receptors[92,96,93,97] to direct affinity labeling of histone peptides containing the PTM trimethyl lysine (Kme3) (the ligand). Herein we transform a synthetic receptor for Kme3 into a chemical probe capable of selective turn-on fluorescence labeling of histone peptides containing Kme3, resulting in the first example, to our knowledge, of a synthetic receptor-mediated affinity

labeling reagent. In our labeling scheme, receptor binding to Kme3 places the electrophile in close proximity to a nucleophilic lysine on the target peptide, facilitating covalent labeling (Figure 1). We also demonstrate the utility of this labeling strategy in the development of a turn-on fluorescence assay for screening inhibitors of histone deacetylase activity (Chapter 4). This work represents the first example of synthetic-receptor directed affinity labeling and provides impetus for the use of synthetic receptors as agents for site-selective chemical modification of proteins mediated by PTMs.

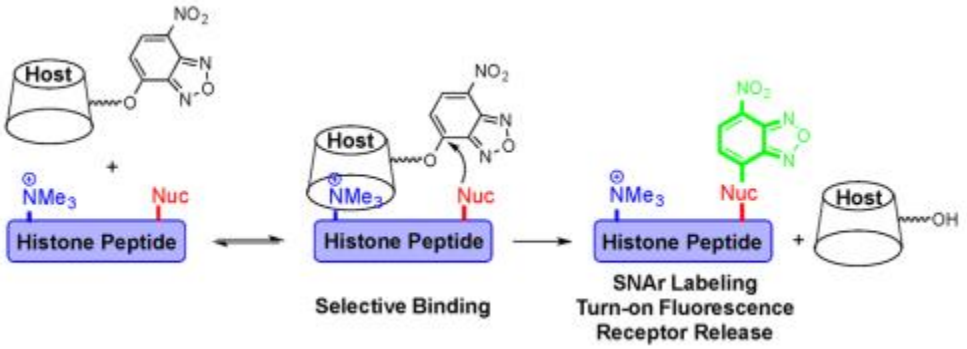


Figure 3.1. Turn-on fluorescence affinity labeling of a histone Kme3 peptide using a CX4-ONBD probe

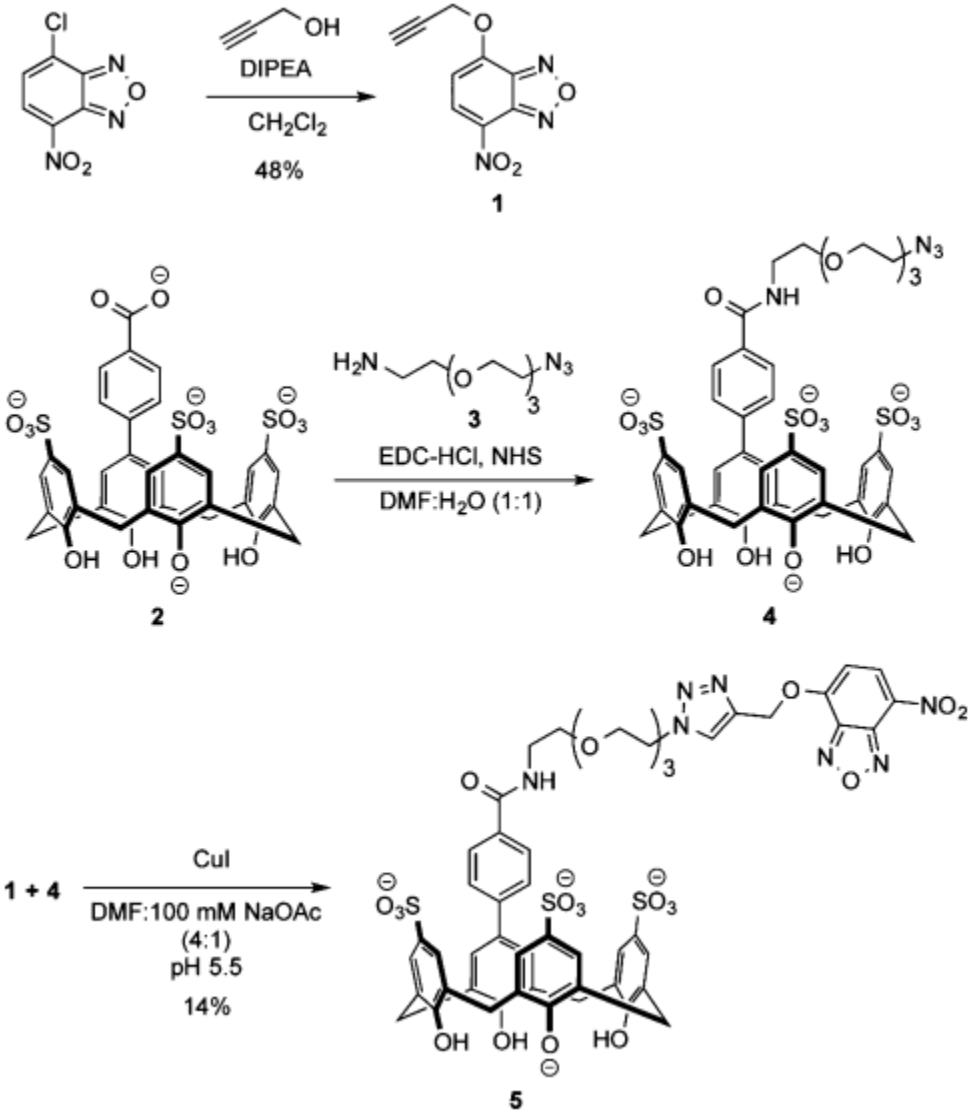
Design and Synthesis of a Probe for Labeling for Kme3

Analysis of histone sequences reveals that most known sites of Lys trimethylation contain another Lys 4-5 residues away, with a small number consisting of larger spacing between Kme3 and a neighboring Lys.[118] This pattern suggests that binding of a synthetic receptor to Kme3 brings it into close proximity of a nucleophilic residue, such that an affinity-labeling strategy is feasible. With the goal of preparing affinity probes capable of selective covalent labeling of histone peptides containing both Kme3 and a pendant unmethylated lysine, we designed a probe that features a synthetic receptor for Kme3 conjugated via a short flexible linker to a fluorogenic nitrobenzoxadiazole (NBD) group. For the synthetic receptor, we chose a tri-sulfonated calix-[4]-

arene (CX4-Ar) receptor first reported by the Hof group[93] that shows preferential binding to trimethyl lysine over lower methylation states of lysine. An O-aryl NBD group was incorporated as a tag unit to allow for turn-on fluorescence labeling. In this labeling chemistry[58] developed by the Sodeoka lab, the ONBD unit serves as both an electrophilic site and a tag unit that selectively reacts with amines (i.e. lysine) with release of the receptor, leading to traceless fluorescent labeling of the target protein or peptide.

We prepared the CX4-ONBD probe **5** by assembling three separate components: propargyl ONBD compound **1**, CX4-Ar receptor **2,** and amino-azide linker **3** (Scheme 1). First, the alkyne functionalized NBD compound **1** was prepared by treating NBD chloride with propargyl alcohol in the presence of tertiary amine base. The synthetic receptor **2** was prepared with minor alterations to published procedure.[93] Likewise, the flexible amino-azide linker **3** was prepared according to previously reported procedure[119] starting from tetraethylene glycol. After preparing each of the three components, tri-sulfonated CX4-Ar receptor **2** was functionalized with linker **3** using carbodiimide coupling to produce azide-functionalized **4**. Finally, copper catalyzed azide-alkyne cycloaddition (CuAAC)[120] was used to conjugate alkyne **1** to azide **4**, generating the CX4-ONBD probe **5** which was purified by reversed-phase HPLC.

Scheme 3.1. Synthesis of CX4-ONBD Probe



Characterization of Binding by Isothermal Titration Calorimetry

Previously reported binding studies[93] of **2** indicate that the receptor has favorable binding to Kme3 in the context of a histone peptide (H-ARTKQTARKme3STGY-NH_2), with a K_d of 0.19 μM in 10 mM phosphate buffer at pH 7.4 and 25 ^oC.[121] To determine whether covalent modification of this receptor influences binding, we measured the binding affinity of the CX4-ONBD probe **5** with a histone 3 (H3) peptide containing Kme3 (Table 1) using isothermal titration calorimetry (ITC). Peptides were synthesized with an N-terminal WGG unit to facilitate

calculation of peptide concentration by measuring absorbance at 280 nm. K9me3R14 was used in ITC experiments with **5**, where the lysine at position 14 was substituted for an arginine to prevent covalent labeling over the course of the ITC experiment without changing the overall charge of the peptide. Under phosphate buffer conditions at pH 7.4 with no added salt (Table 1 and Table 2, Entry 1), the CX4-ONBD probe had a K_d of 4.8 μM for K9me3R14. While most examples of ligand-directed affinity labeling typically take advantage of ligands with submicromolar or nanomolar binding to their targets, it has been demonstrated that micromolar affinity interactions can provide an increase in effective molarity that is satisfactory for proximity-driven labeling.[122] Thus, we proceeded with evaluation of probe **5** as an affinity labeling probe.

Table 3.1. Peptides Used for ITC, Fluorescence Labeling, and Deacetylase Activity Screening

Abbreviation	Peptide Sequence
K9me3R14	Ac-WGGGQTARKme3STGGR-NH_2
K9me3K14	Ac-WGGQTARKme3STGGK-NH_2
K9K14	Ac-WGGQTARKSTGGK-NH_2
K4me3K9ac	Ac-WGGTKme3QTARKacSTG-NH_2
K4me3K9	Ac-WGGTKme3QTARKSTG-NH_2
K27acK36me3	Ac-WGGRKacSAPATGGVKme3-NH_2
K27K36me3	Ac-WGGRKSAPATGGVKme3-NH_2

Table 3.2. Binding and Thermodynamic Data from ITC Experiments of 5 and 4 with K9me3K14 and K9K14

Entry	Host	Peptide	Charge	K_d	ΔG (kcal/mol)
1[a]	5	K9me3R14	+3	4.8 ± 1.5	−7.3 ± 0.2
2[b]	4	K9me3K14	+3	18.3 ± 7.7	−4.9 ± 0.6
3[b]	4	K9K14	+3	>390	

[a] 10 mM potassium phosphate, pH 7.4. Two trials using 82 μM probe 5 and 1.22 mM K9me3R14; One trial using 81 μM probe 5 and 1.03 mM K9me3R14. [b] 10 mM sodium borate, 100 mM NaCl, pH 8.6. Three trials using 85 μM probe 4 and 0.93 mM K9K14. Two trials using 85 μM probe 4 and 1.11 mM K9me3K14; One trial using101 μM probe 4 and 1.43 mM K9me3K14.

Optimization of Receptor-Mediated Labeling

To evaluate the reactivity of **5** toward histone peptides, the probe was incubated with either K9me3K14 or K9K14 in 10 mM potassium phosphate buffer (pH 7.4) at room temperature for 24 h (Table 3). The labeling reactions were analyzed by HPLC (See Experimental). Initially, concentrations well above the K_d were chosen for both probe and peptide to ensure complete host-guest complexation. Reaction yields after 24 h revealed slow reactivity of the CX4-ONBD probe toward the representative histone peptides with modest 2:1 selective labeling for K9me3K14 over K9K14 (29% and 12% labeling yield, respectively). Since the O-aryl NBD group is known to be more reactive toward amine nucleophiles at higher pH,[58] the slow reactivity for our system could be attributed in part to the extent of protonation at pH 7.4 of the □-amino group of the lysine residues in the histone peptides. Expectedly, increasing the pH of the buffer used in the labeling reaction to 8.6 resulted in a marked rate acceleration. Unfortunately, this rate acceleration was accompanied by a complete loss in labeling selectivity for K9me3K14 over the K9K14 peptide (75% and 70% labeling yield, respectively).

To circumvent the issue of unselective labeling, we screened conditions with different concentrations of NaCl, as the concentration of salt in aqueous solutions is known to influence

the strength of electrostatic interactions involved in host-guest complexation.[123] Varying NaCl concentration from 50 mM to 150 mM revealed two general trends: the overall labeling yield decreases slightly at higher concentrations of salt, while selectivity for K9me3K14 over K9K14 improves with increasing salt. Finally, we looked to improve selectivity further by lowering the concentration of both probe and peptide. Under conditions containing 40 μM probe and 20 μM peptide, the labeling selectivity improved to nearly 6-fold for K9me3K14 over K9K14. Also, despite micromolar binding and high salt concentration, the labeling rate for the reaction is appreciable even at lower concentrations of reagents, and the rate compares favorably to ligand-directed affinity labeling approaches.[52,47,48]

To determine the binding selectivity for Kme3 over unmethylated lysine under these optimized conditions, we used azide-functionalized receptor **4** as a surrogate for the CX4-ONBD probe as the azide is unreactive toward lysine-containing peptides during the ITC experiment. Additionally, **4** would not be expected to have binding interactions that are considerably different from **5** since the differences in functional groups between the two compounds are not close to the binding pocket. In 10 mM borate buffer with 100 mM NaCl (pH 8.6), receptor **4** exhibited about 20-fold selectivity for K9me3K14 over K9K14 with dissociation constants of 18.3 μM and 390 μM, respectively (Table 2, entries 2 and 3). This suggests that the affinity labeling is not achieving the maximum selectivity possible. Several factors likely contribute to this, including the flexibility of the linker that reduces the degree of pre-organization as well as the fact that the unmethylated peptide has twice the number of reactive lysines, which reduces the selectivity.

Table 3.3. Optimization of Labeling Reaction Conditions of CX4-ONBD Probe 5 with H3 Peptides

Peptide	Conditions [a,b]	Probe:peptide (μM)	Labeling yield (%)[c]
K9K14	pH 7.4[a]	100:100	12
K9me3K14	pH 7.4[a]	100:100	29
K9K14	pH 8.6[b]	100:50	70
K9me3K14	pH 8.6[b]	100:50	75
K9K14	pH 8.6, 50 mM NaCl[b]	100:50	59
K9me3K14	pH 8.6, 50 mM NaCl[b]	100:50	82
K9K14	pH 8.6, 100 mM NaCl[b]	100:30	28
K9me3K14	pH 8.6, 100 mM NaCl[b]	100:30	68
K9K14	pH 8.6, 150 mM NaCl[b]	100:30	22
K9me3K14	pH 8.6, 150 mM NaCl[b]	100:30	65
K9K14	pH 8.6, 100 mM NaCl[b]	40:20	11
K9me3K14	pH 8.6, 100 mM NaCl[b]	40:20	61

[a]10 mM potassium phosphate, pH 7.4. [b]10 mM sodium borate, pH 8.6. [c]Yield is based on integration of the peaks in the HPLC trace; error is estimated at ±10% of the listed value for each entry.

Turn-On Fluorescence Labeling of Kme3 Peptides

Next, we explored the turn-on fluorescent labeling of histone peptides by **5** by monitoring the fluorescence emission at 520 nm (excitation at 485 nm). Using optimized conditions of 40 μM probe **5** and 20 μM peptide in 10 mM phosphate buffer (100 mM NaCl, pH 8.6), measurements were taken every hour, and reactions were allowed to proceed under ambient temperature for 24 h (Figure 2). A pronounced turn-on fluorescence response was observed for labeling of K9me3K14 with **5** (blue diamonds). On the other hand, the turn-on response for labeling with the K9K14 peptide (green triangles) was significantly lower, even though the

peptide contains two reactive lysine residues. Trials with probe **5** alone appear to show only a slight increase in fluorescence signal over time, presumably due to slow hydrolysis of the O-NBD group.

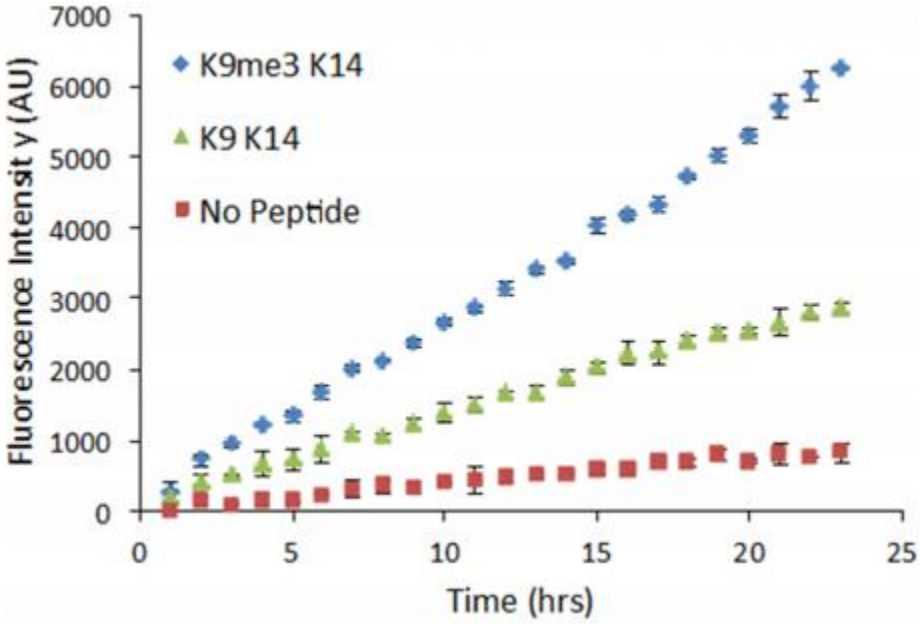


Figure 3.2. Turn-on fluorescence labeling of K9me3K14 and K9K14 peptides with Probe 5. Conditions: 40 μM 5 and 20 μM peptide in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl.

Fluorescence labeling experiments were also performed with Propargyl ONBD **1** as a negative control to examine the contribution of the synthetic receptor to probe design and function. Under the same optimized reaction conditions employed for labeling with **5**, negative control **1** was allowed to react with K9K14 and with K9me3K14 for 24 h. Unlike probe **5**, labeling with the negative control was slow and non-selective (see Experimental). Moreover, compound **1** reacted with K9K14 to a greater extent than with K9me3K14 as the former contains two reactive lysine residues rather than just one. These results are consistent with HPLC analysis of labeling and further confirm that preferential binding of the receptor to Kme3 over K directly translates into favored turn-on fluorescence labeling of Kme3 histone peptides. To further

confirm that the enhanced labeling rate of K9me3K14 by probe **5** is receptor mediated, a competition reaction was performed in which addition of K9me3R14, which can bind but not react with the probe, inhibited fluorescence labeling of K9me3K14 by probe **5** (see Experimental).

To probe the distance-dependence of labeling, we also investigated the reaction of probe **5** with biologically relevant histone peptides with varying numbers of amino acids separating the Kme3 binding site and the pendant reactive amine: K23K27me3 (three amino acids), K36me3K37 (zero amino acids), K4me3 N-terminus (two amino acids), K4me3K9 (four amino acids), and K27K36me3 (eight amino acids). Reactions were performed using 40 μM **5** and 20 μM peptide in 25 mM bicine buffer (137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, pH 8.02). Fluorescence measurements were taken after 24 h. Probe **5** reacted with K4me3K9 to the greatest extent, suggesting a preference for a spacing of four amino acids between Kme3 and the reactive amine (Figure 3). In contrast, K36me3K37 and K27K36me3, the two peptides with the shortest and longest spacing, respectively, showed considerably slower labeling. Interestingly, the peptide containing K4me3 and a free N-terminal amine (with a spacing of 2 amino acids) was the second most reactive in the series of peptides that were screened. This may be explained in part by the increased acidity of the alpha amino group relative to the Ɛ amine of the lysine sidechain, making it a better nucleophile at pH 8. Unexpectedly, probe **5** reacted with K23K27me3 (three amino acid spacing) the slowest, indicating that there may be subtleties in sequence and binding conformation that affect labeling rate in addition to amino acid spacing. While these results show some dependence on distance, all sequences tested gave a turn-on fluorescence response which demonstrates the generality of **5** for labeling relevant histone H3 peptides.

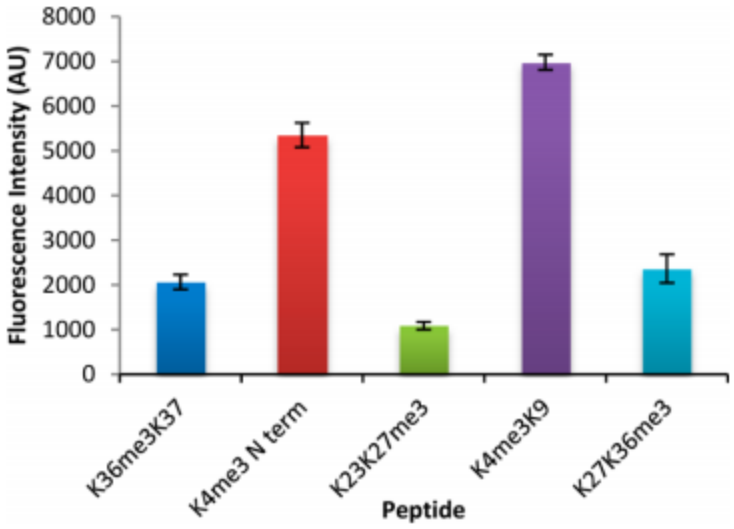


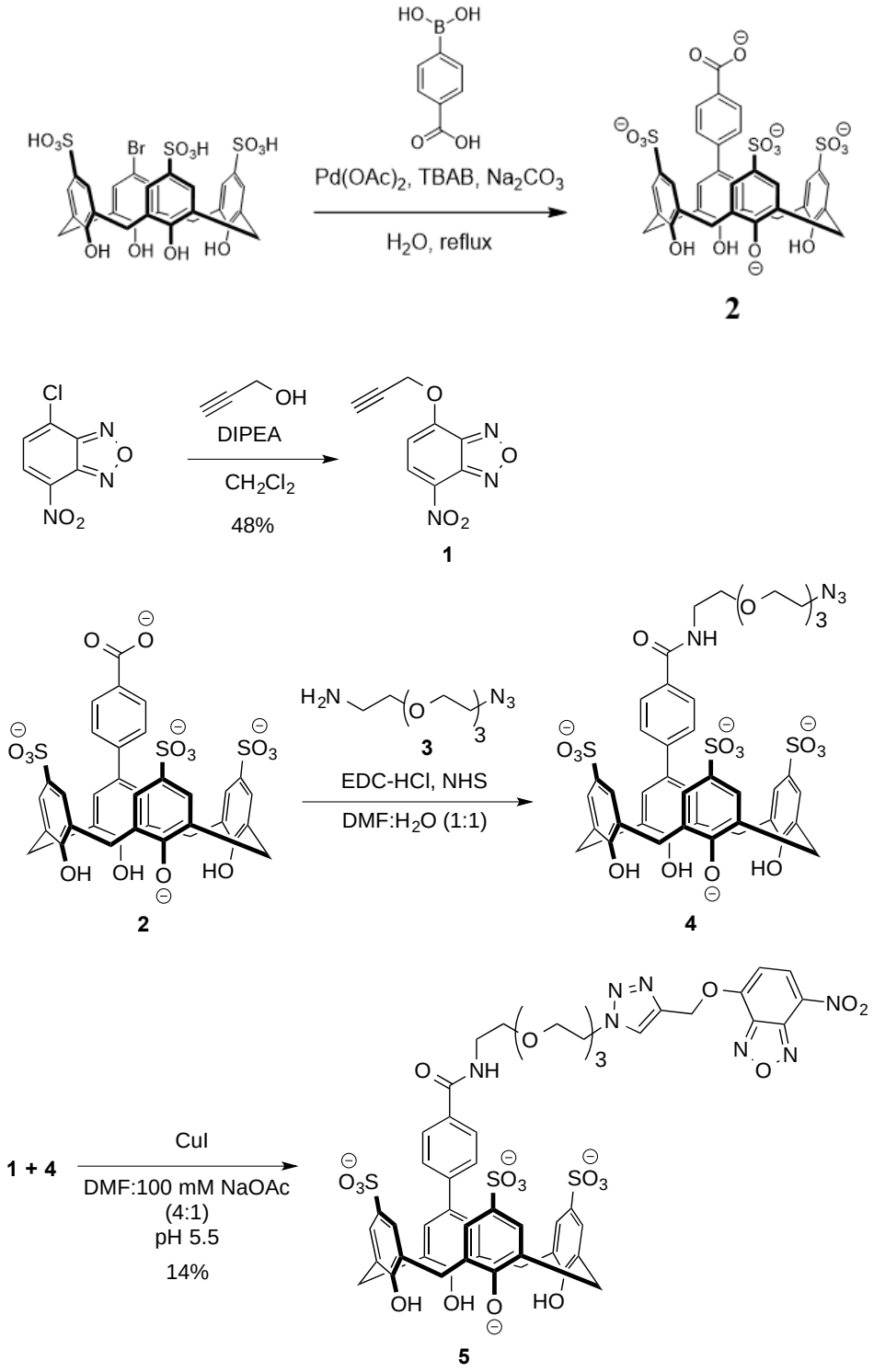
Figure 3.3. Analysis of distance dependence in the labeling of histone peptides by probe **5**. Conditons: 40 μM **5** and 20 μM peptide in 25 mM bicine, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, pH 8.02, 24 h.

Conclusions

Herein we have described the development of a synthetic receptor probe for affinity labeling of Kme3-containing histone peptides, providing a turn-on fluorescence signal for the presence of Kme3. To our knowledge, this is the first example of a labeling method using a synthetic receptor as an affinity labeling probe. We found that the rate and selectivity of the labeling reaction was considerably dependent on buffer pH, salt concentration, and reagent concentration. Using reaction buffer with slightly alkaline pH accelerated the labeling rate, while utilizing high salt concentration improved selectivity by minimizing non-specific electrostatic interactions between the probe and peptide. Labeling was further improved by selecting concentrations that maintained appreciable host-guest complexation while exploiting the weaker binding of the unmethylated histone peptide at lower concentrations. In the end, optimizing labeling conditions lead to selective covalent modification and turn-on fluorescence in the reaction of the probe with the target peptide.

Experimental

Scheme 3.2. Synthesis of CX_4-ONBD Probe **5**.



Synthesis of 1. 4-chloro-nitrobenzo[c][1,2,5]oxadiazole (0.100 g, 0.50 mmol, 1 equiv.) was dissolved in 5 mL CH_2Cl_2 under a nitrogen atmosphere. To this solution was added a mixture of propargyl alcohol (0.029 mL, 0.50 mmol, 1 equiv.) and DIPEA (0.087 mL, 0.50 mmol, 1 equiv.) in 5 mL CH_2Cl_2, and stirring was continued for 72 h. The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (EtOAc/Hexanes 1:1) to give the product as a yellow-brown solid (0.053 g, 48%). ^{1}H NMR ($CDCl_3$, 400 MHz): 8.558 (d, 1H, CH), 6.880 (d, 1H, CH), 5.129 (d, 2H, CH_2), 2.720 (t, 1H, CH). MS (calculated): 242.02 [M+Na^+]. HRMS (observed, ESI+): 242.01701 [M+Na^+].

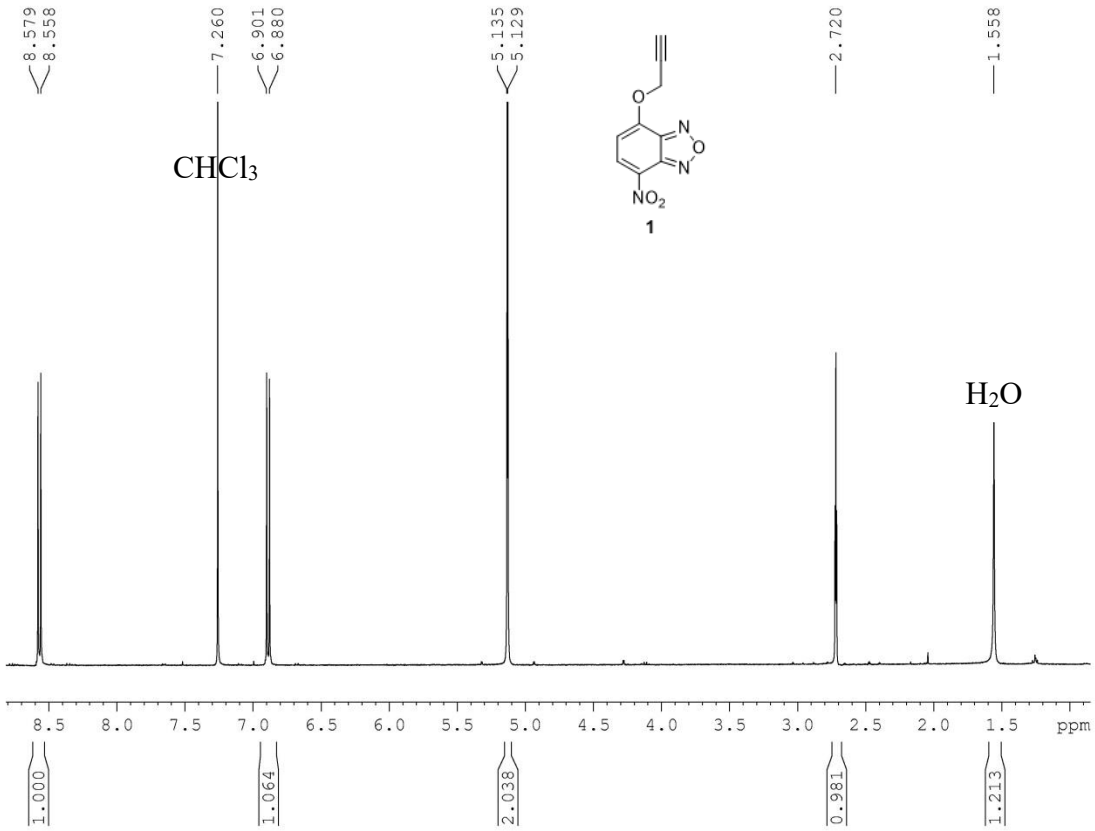


Figure 3.4. 1D ^{1}H-NMR (400 MHz) of Compound **1** in $CDCl_3$.

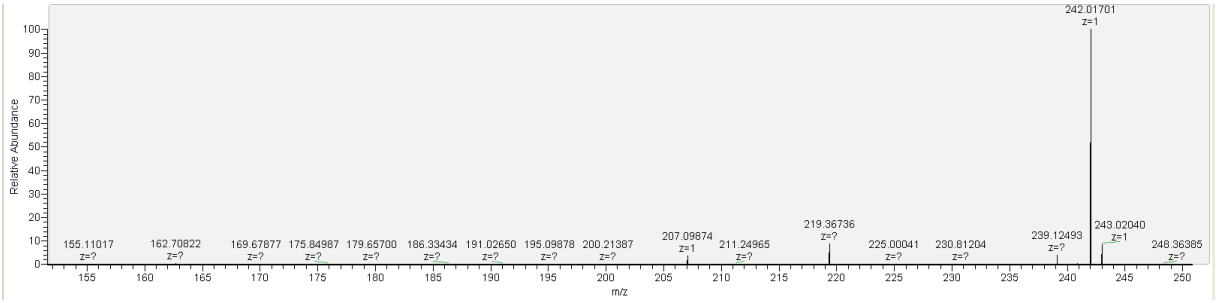


Figure 3.5. High resolution mass spectrum of Compound **1**. MS (calculated): 242.02 [M+Na^+]. HRMS (observed, ESI+): 242.01701 [M+Na^+].

Synthesis of 2. Compound **2** was synthesized according to published procedure[93] on a 0.5 g scale but with a modified purification procedure. 5-bromo-25,26,27,28-tetrahydroxy-11-17-23-trisulfonatocalix[4]arene (0.50 g, 0.067 mmol), $PdOAc_2$ (0.003 g, 0.013 mmol), tetrabutylammonium bromide (0.0108 g, 0.034 mmol), Na_2CO_3 (0.027 g, 0.25 mmol), and 4-carboxyphenylboronic acid (0.0112 g, 0.067 mmol) were dissolved in degassed H_2O under an atmosphere of nitrogen. The reaction mixture was heated to reflux using an oil bath for 1h. After completion of the reaction, the aqueous solution was washed with DCM (3 x 20 mL) and EtOAc (3 x 25 mL). The aqueous solvent was removed under reduced pressure. The residue was taken up in MeOH/H_2O (1:1), and palladium was removed by QuadraPure® TU resin. The resin (0.5 g) was added to the solution and heated to reflux under nitrogen for 24 h. The mixture was filtered, and the solvent was removed under reduced pressure. The crude product was used without further purification.

Synthesis of 3. Compound **3** was prepared according to published procedure[119].

Synthesis of 4. EDC-HCl (0.087 g, 0.45 mmol, 2 equiv.) and N-hydroxysuccinimide (0.050 g, 0.43 mmol, 2 equiv.) were added to a solution of **2** (0.176 g, 0.22 mmol, 1 equiv.) in 5 mL DMF/H_2O (1:1) under a nitrogen atmosphere, and the mixture was stirred for 1 h. **3** (0.100 g,

0.46 mmol, 2.1 equiv.) was added, and the solution was allowed to stir at room temperature overnight. The solvent was removed under reduced pressure, and the resulting residue was used in the next step without further purification. ^{1}H NMR (D_2O, 600 MHz): 7.718 (d, 2H), 7.660 (d, 2H), 7.587 (d, 2H),7.498 (d, 2H), 7.478 (s, 2H), 7.464 (s, 2H), 4.280 (d, 2H), 3.509-3.633 (m, 12H), 3.321 (s, 2H), 3.052 (s, 2H), 2.765 (s, 2H), 2.263 (s, 2H), 1.787 (t, 2H). MS (calculated): 983.18 [M-H], 1005.16 [M-2H+Na^+], 491.09 [M-2H], 502.58 [M-3H+Na^+]. HRMS (observed, ESI-) 983.17713 [M-H], 1005.15929 [M-2H+Na^+], 491.08576 [M-2H], 502.57840 [M-3H+Na^+].

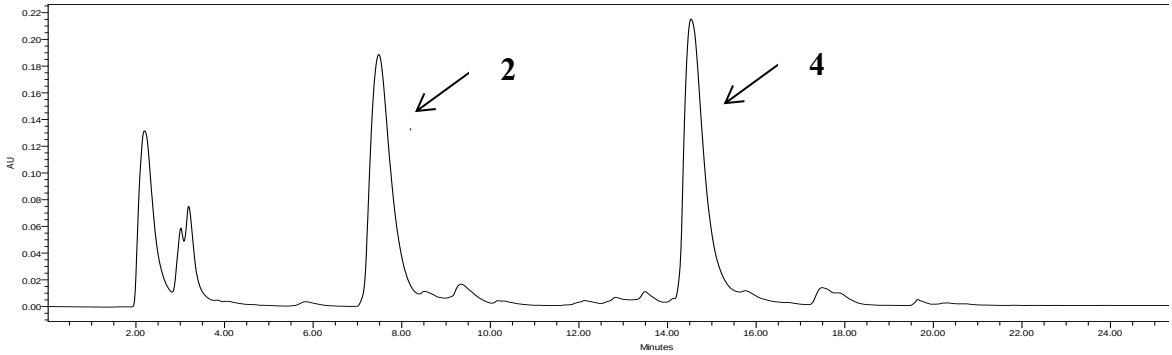


Figure 3.6. Crude semi-preparative HPLC trace of the synthesis of **4** from **3** at 280 nm.

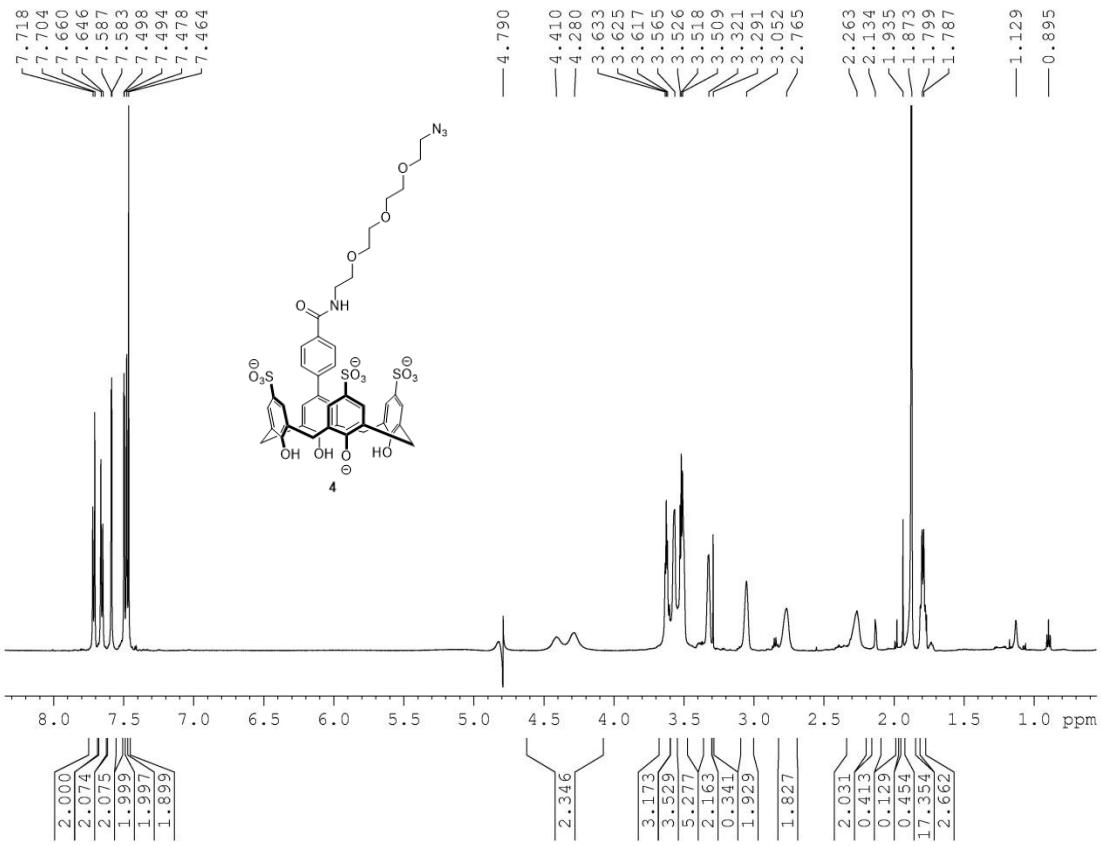


Figure 3.7. ^{1}H-NMR (600 MHz) of Compound **4** in D_2O.

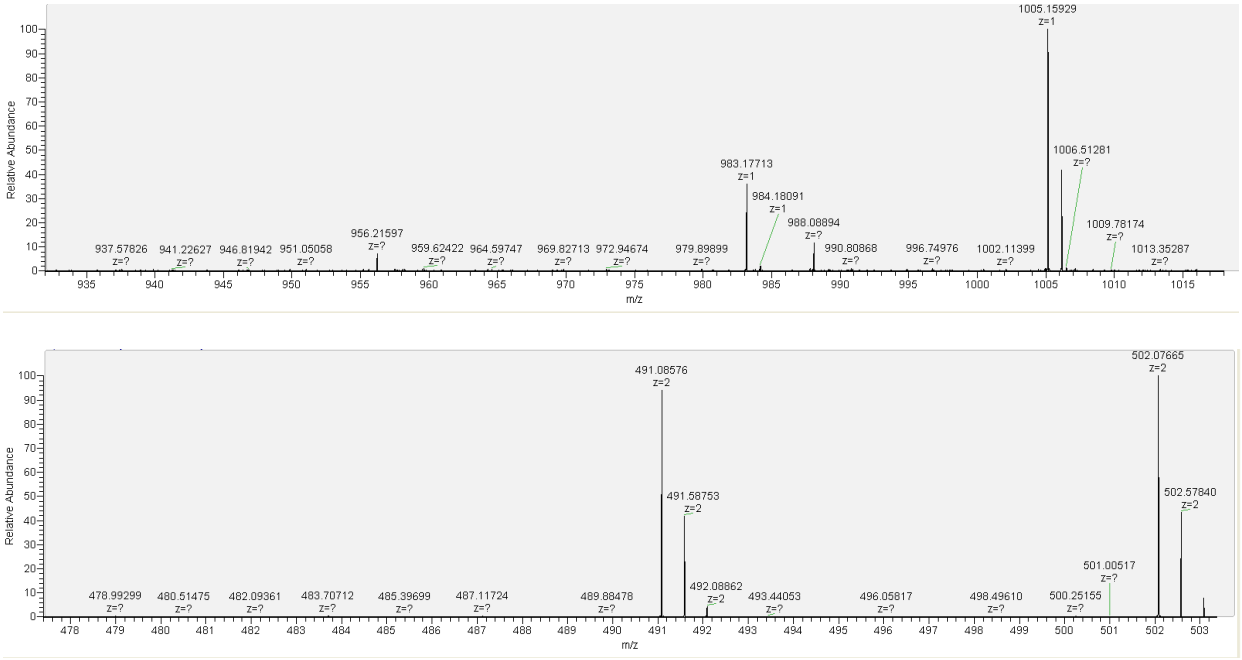


Figure 3.8. High resolution mass spectrum of Compound **4**. MS (calculated): 983.18 [M-H], 1005.16 [M-2H+Na$^+$], 491.09 [M-2H], 502.58 [M-3H+Na$^+$]. HRMS (observed, ESI-) 983.17713 [M-H], 1005.15929 [M-2H+Na$^+$], 491.08576 [M-2H], 502.57840 [M-3H+Na$^+$].

Synthesis of 5: The crude residue containing **4** (0.037 g, 0.038 mmol, 1 equiv.) and **1** (0.049 g, 0.22 mmol, 5.8 equiv.) were dissolved in 5 mL DMF/100 mM sodium acetate, pH 5.5 (4:1). Copper iodide (0.001 g, 0.005 mmol, 0.13 equiv.) was added, and the solution was allowed to stir at room temperature for 48 h. The solvent was removed under reduced pressure, and the resulting residue was purified by semi-preparative reversed-phase HPLC (solvent A: 10 mM NH_4OAc in 9:1 H_2O:CH_3CN; solvent B: 10 mM NH_4OAc in 9:1 CH_3CN:H_2O) using the gradient: 0-100% B from 0-60 min with a flow rate of 4.0 mL/min to produce the desired compound in 14% yield. ^{1}H NMR (D_2O, 600 MHz): 7.996 (s, 1H), 7.662 (s, 2H), 7.639 (s, 2H), 7.264-7.295 (m, 4H), 7.104-7.128 (m, 5H), 6.630 (d, 1H), 5.167 (s, 2H), 4.502 (t, 2H), 3.841 (t, 2H), 3.506-3.657 (m, 12H). MS (calculated) 1202.21 [M-H], 1224.18 [M-2H+Na$^+$], 600.60 [M-2H], 611.59 [M-3H+Na$^+$]. HRMS (observed, ESI-) 1202.20580 [M-H], 1224.18764 [M-2H+Na$^+$], 600.59975 [M-2H], 611.59092 [M-3H+Na$^+$].

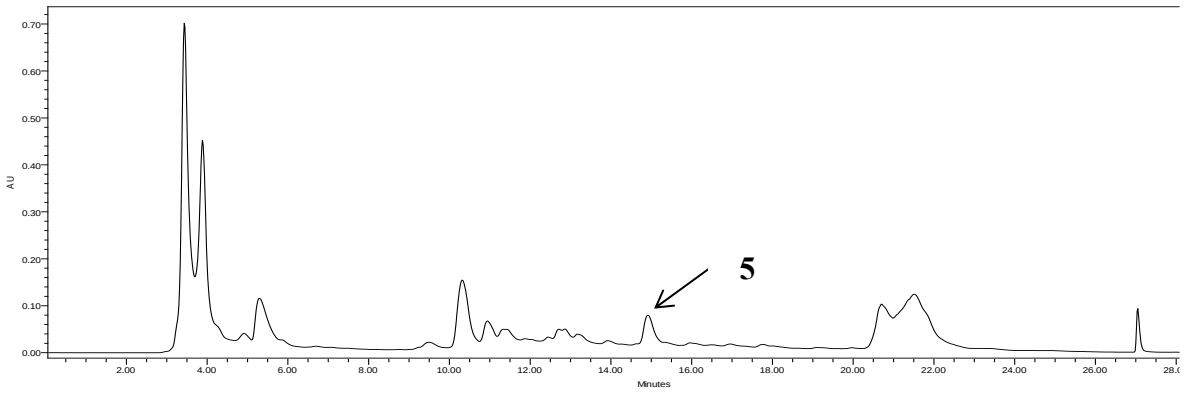


Figure 3.9. Crude semi-preparative HPLC trace of **5** at 280 nm.

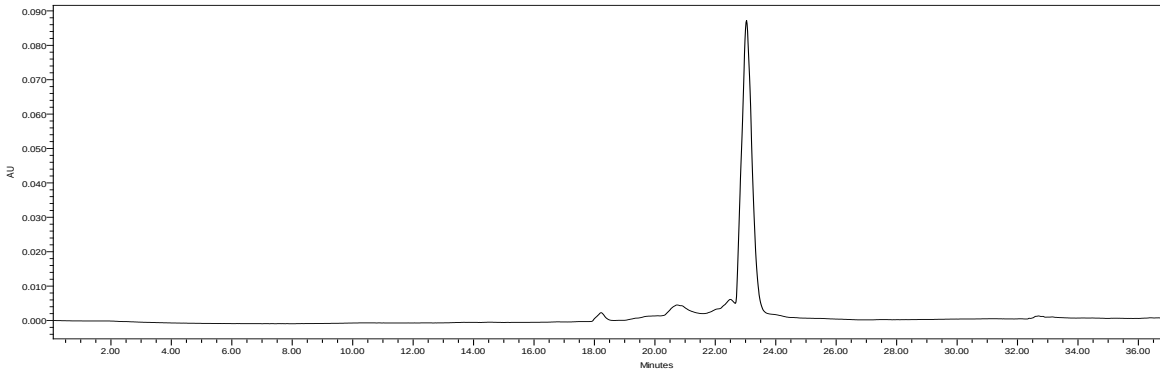


Figure 3.10. Semi-preparative HPLC trace prior to second purification of **5** at 280 nm.

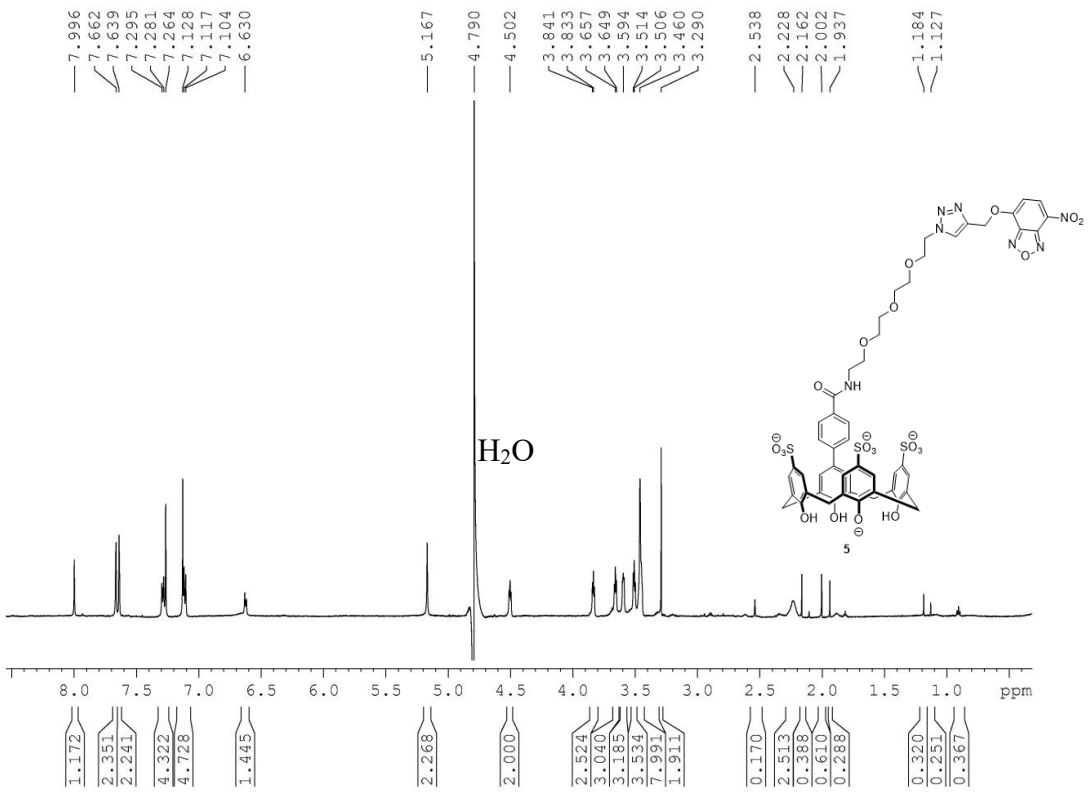


Figure 3.11. ^{1}H-NMR (600 MHz) of Compound **5** in D_2O.

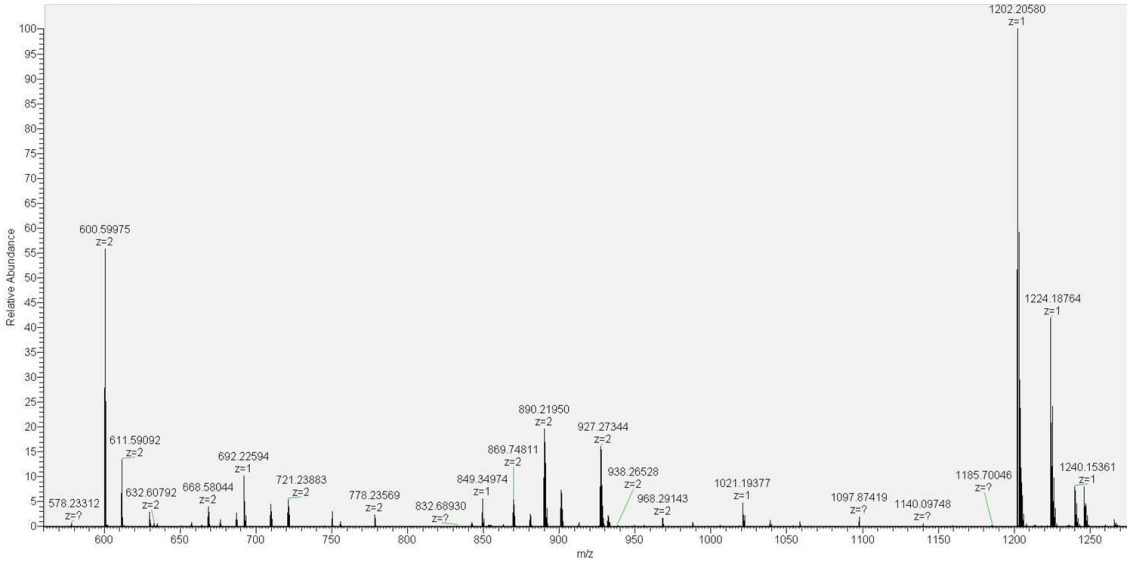


Figure 3.12. High resolution mass spectrum of compound **5** (ESI-). MS (calculated) 1202.21 [M-H], 1224.18 [M-2H+Na^+], 600.60 [M-2H], 611.59 [M-3H+Na^+]. HRMS (observed, ESI-) 1202.20580 [M-H], 1224.18764 [M-2H+Na^+], 600.59975 [M-2H], 611.59092 [M-3H+Na^+].

Peptide Synthesis

Peptides were synthesized on a 0.6 mmol scale via solid-phase peptide synthesis in a peptide synthesis flask using Fmoc protected amino acids. Coupling reagents were HOBt/HBTU in DMF (4 equiv each), with coupling times of 60 min, followed by rinsing with DMF, MeOH, and CH_2Cl_2. Deprotection was performed with 20% piperidine in DMF. All peptides were acylated at the N-terminus with a solution of 5% acetic anhydride and 6% 2,6-lutidine in DMF. Cleavage was performed using a cocktail of 95% TFA/2.5% triisopropylsilane/2.5% H_2O for 3 h. Peptides were purified by semi-preparative reverse-phase HPLC on a C18 column at a flow rate of 4 mL/min. Peptides were purified with a linear gradient of A and B (A: 95% H_2O/5% CH_3CN with 0.1% TFA, B: 95% CH_3CN/5% H_2O with 0.1 % TFA).

Methylated peptides were synthesized with 2 equivalents of Fmoc-Lys$(Me)_2$-OH HCl purchased from EMP Millipore and coupled for 5 h. The trimethyllysine containing peptides were synthesized by reacting corresponding dimethylated peptides (0.6 mmol scale) prior to cleavage from the resin with MTBD (7-Methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene, 10 equiv.) and methyl iodide (10 equiv.) in DMF (5 mL) for 5 h with bubbling nitrogen in a peptide flask. After washing the resin with DMF and CH_2Cl_2 and drying, the peptide was cleaved as normal. Acetylated peptides were prepared from 4 equivalents of Fmoc-Lys(ivDde)-OH purchased from Chem Impex with 1 h couplings. The ivDde group was deprotected using 5% hydrazine hydrate in DMF (3 x 10 mL) for 5 minutes. After washing resin with DMF, MeOH, and CH_2Cl_2, peptides were acetylated using a solution of 5% acetic anhydride and 6% 2,6-lutidine in DMF. Peptides were cleaved from the resin as normal after washing with DMF, MeOH, and CH_2Cl_2.

Peptides:

K9K14: Ac-WGGQTARKSTGGK-NH_2

K9me3K14: Ac-WGGQTARKme3STGGK -NH_2

K9me3R14: Ac-WGGGQTARKme3STGGR-NH_2

K4me3K9ac: Ac-WGGTKme3QTARKacSTG-NH_2

K4me3K9: Ac-WGGTKme3QTARKSTG-NH_2

K27acK36me3: Ac-WGGRKacSAPATGGVKme3-NH2

K27K36me3: Ac-WGGRKSAPATGGVKme3-NH_2

RKme3GLLYK: Ac-RKme3GLLYK-NH_2

K9K14

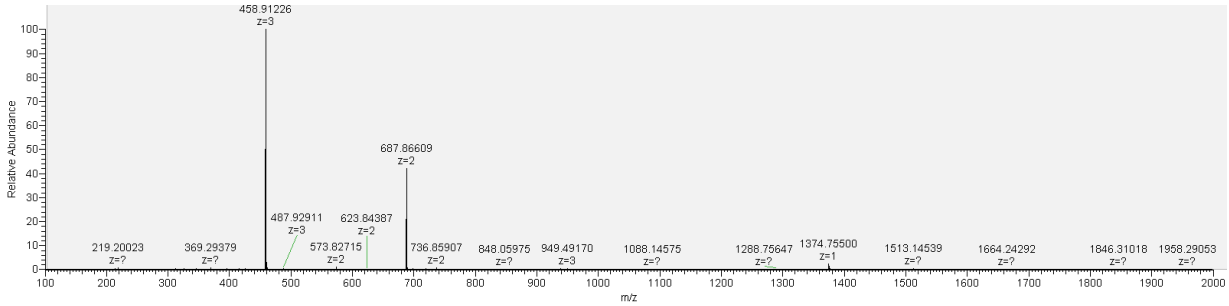


Figure 3.13. High resolution mass spectrum of K9 K14 (ESI +). MS calculated: 1374.73 [M+H], 687.87 [M+2H], 458.91 [M+3H]. MS observed: 1374.75500 [M+H], 687.86609 [M+2H], 458.91226 [M+3H].

K9me$_3$K14

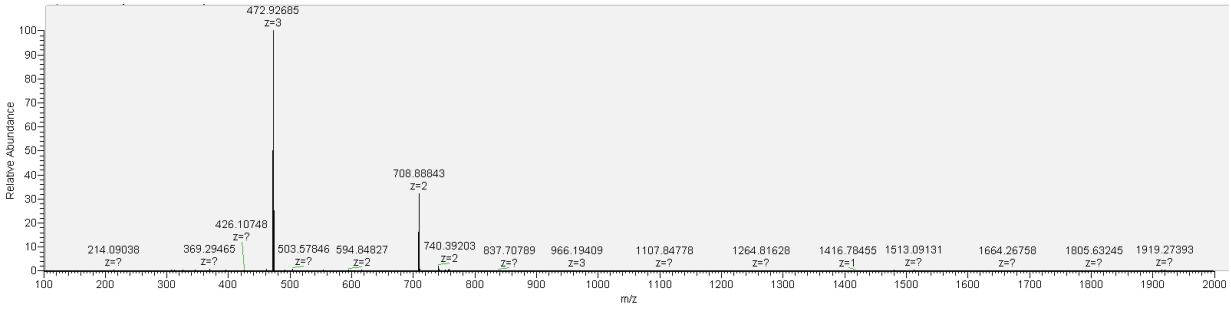


Figure 3.14. High resolution mass spectrum of K9me$_3$ K14 (ESI +). MS calculated: 1416.77 [M], 708.89 [M+H], 472.93 [M+2H]. MS observed: 1416.78455 [M], 708.88843 [M+H], 472.92685 [M+2H].

K9me$_3$R14

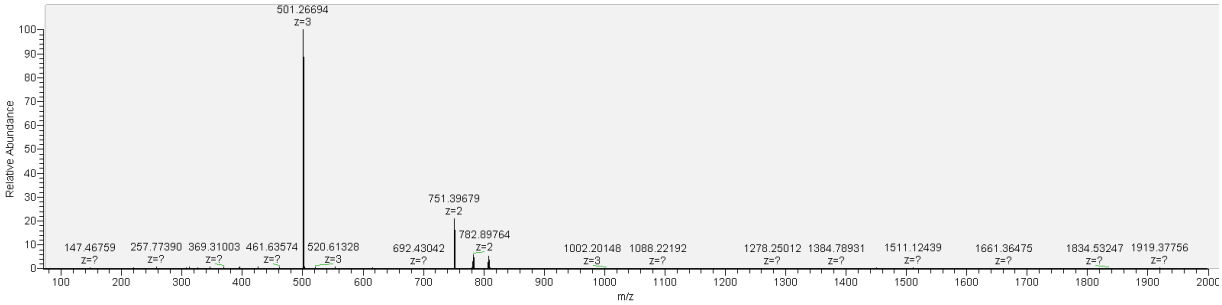


Figure 3.15. High resolution mass spectrum of K9me$_3$ R14 (ESI +). MS calculated: 751.40 [M+1H], 501.27 [M+2H]. Mass observed: 751.39679 [M+1H], 501.26694 [M+2H].

K4me$_3$K9

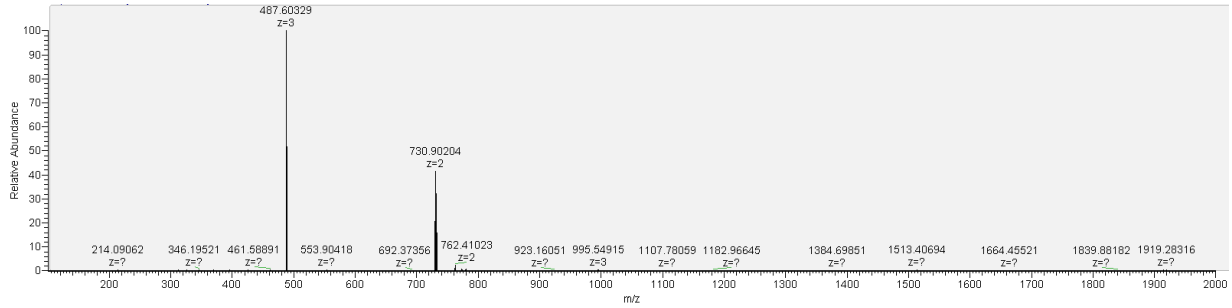


Figure 3.16. High resolution mass spectrum of K4me$_3$ K9 (ESI +). MS calculated: 730.90 [M+H], 487.61 [M+2H]. MS observed: 730.90204 [M+H], 487.60329 [M+2H].

K4me$_3$K9ac

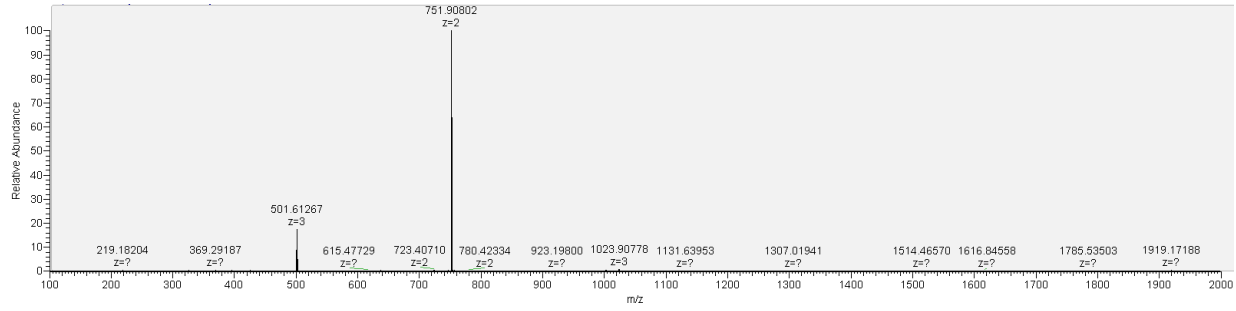


Figure 3.17. High resolution mass spectrum of K4me$_3$ K9ac (ESI +). MS calculated: 751.91 [M+H], 501.61 [M+2H]. MS observed: 751.90802 [M+H], 501.61267 [M+2H].

K27K36me$_3$

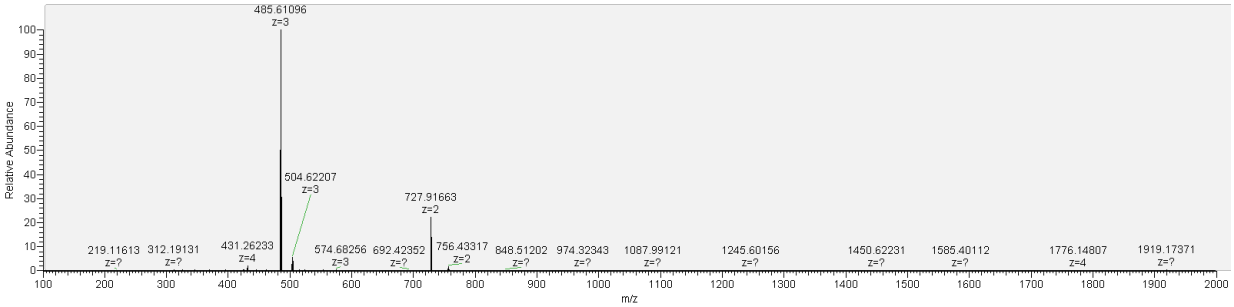


Figure 3.18. High resolution mass spectrum of K27 K36me$_3$ (ESI +). MS calculated: 727.91 [M+H], 485.61 [M+2H]. MS observed: 727.91663 [M+H], 485.61096 [M+2H].

K27acK36me$_3$

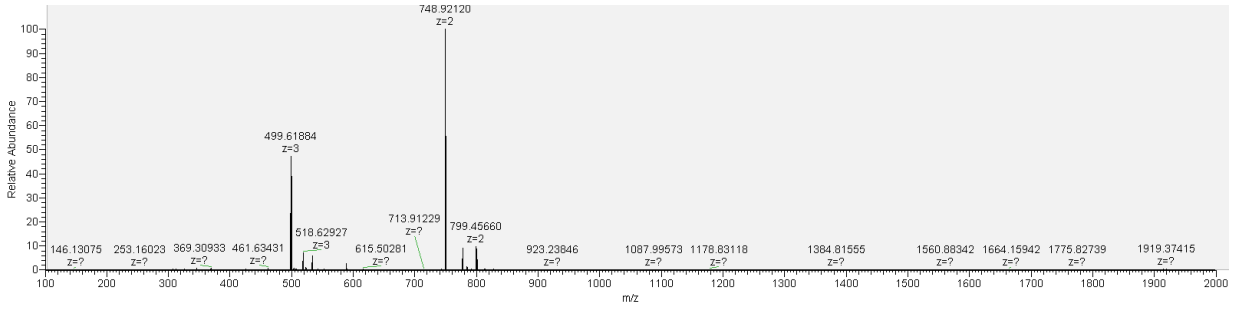


Figure 3.19. High resolution mass spectrum of K27ac K36me$_3$ (ESI +). MS calculated: 748.92 [M+H], 499.62 [M+2H]. MS observed: 748.92120 [M+H], 499.61884 [M+2H].

K23K27me$_3$

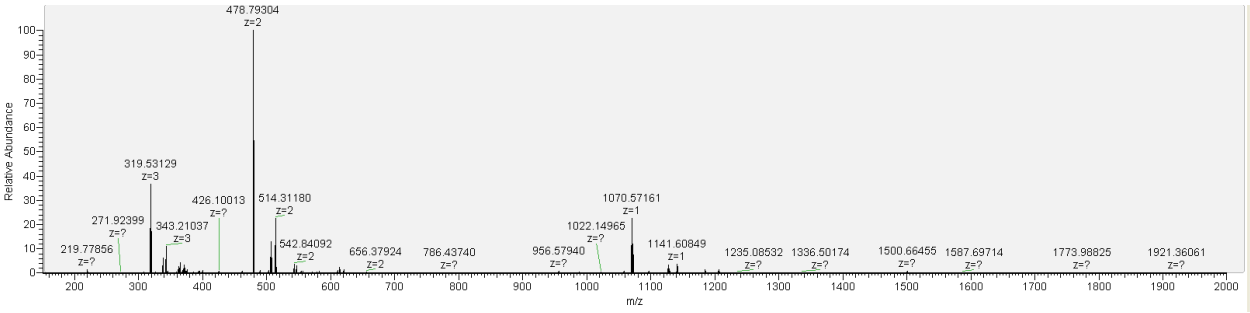


Figure 3.20. High resolution mass spectrum of K23K27me3 (ESI +). MS calculated: 956.58 [M], 1070.57 [M+TFA+H], 478.79 [M+H], 319.53 [M+2H]. MS observed: 956.57940 [M], 1070.57161 [M+TFA+H], 478.79304 [M+H], 319.53129 [M+2H].

K36me$_3$K37

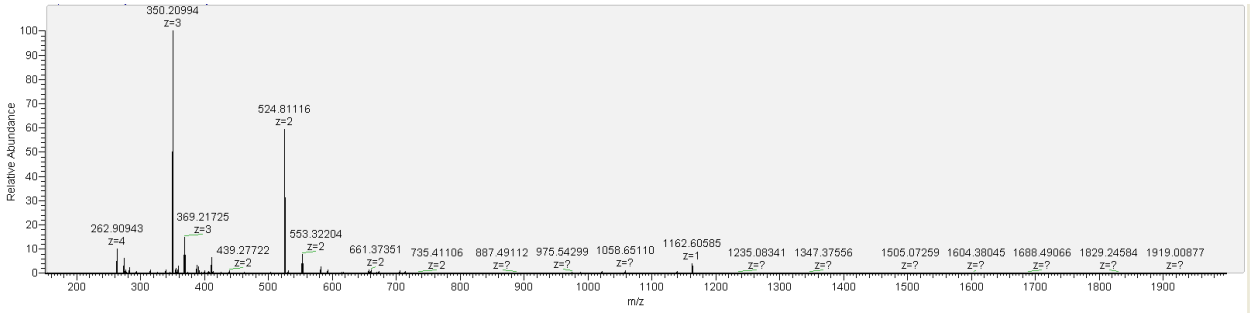


Figure 3.21. High resolution mass spectrum of K36me3K37 (ESI +). MS calculated: 1162.61 [M+TFA+H], 524.81 [M+H], 350.21 [M+2H]. MS observed: 1162.60585 [M+TFA+H], 524.81116 [M+H], 350.20994 [M+2H].

K4me$_3$ N terminus

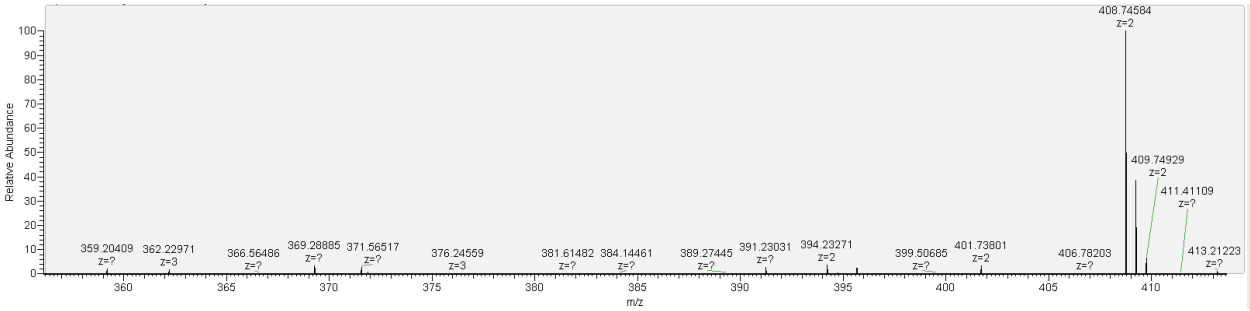


Figure 3.22. High resolution mass spectrum of K4me3 N terminus (ESI +). MS calculated: 408.74 [M+H]. MS observed: 408.74584 [M+H].

RKme3GLLYK

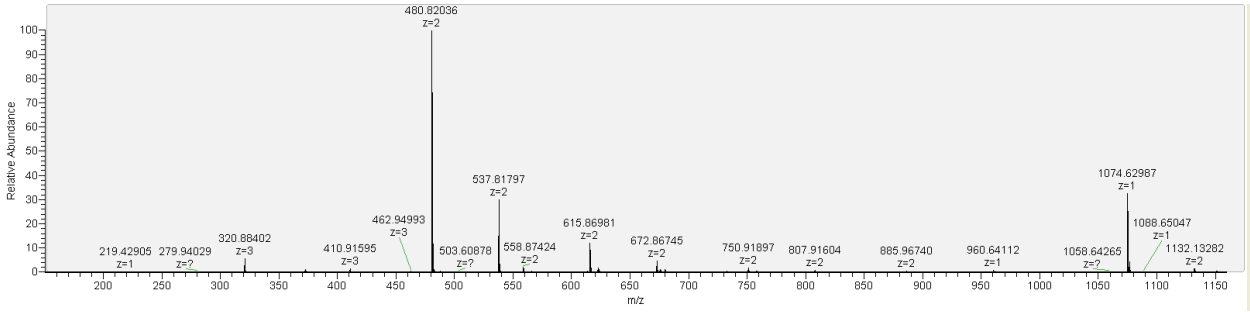


Figure 3.23. High resolution mass spectrum of K4me3 N terminus (ESI +). MS calculated: 960.64 [M], 1074.63 [M+TFA+H], 480.82 [M+H]. MS observed: 960.64112 [M], 1074.62987 [M+TFA+H], 480.82036 [M+H].

Extinction Coefficient Determination for 4.

Compound **4** was purified using reversed phase-HPLC with 10 mM NH_4OAc as the mobile phase additive. After purification, **4** was lyophilized for five days to remove volatile NH_4OAc salts. The dried compound was then dissolved in anhydrous methanol and filtered using a 0.33 μm filter to remove any remaining salts. Methanol was evaporated and **4** was further dried under vacuum. The mass was accurately determined, and a 1 mM stock solution of **4** in 10 mM sodium borate buffer, 100 mM NaCl, pH 8.6 was prepared. The stock solution was serially diluted (80:20, **4**:buffer) to give 12 concentrations. The absorbance at 315 nm was measured for each concentration, which was then plotted against the concentration. The extinction coefficient of **4** was determined from the slope of the line of regression using Beer's law and was found to be 900 $M^{-1}cm^{-1}$.

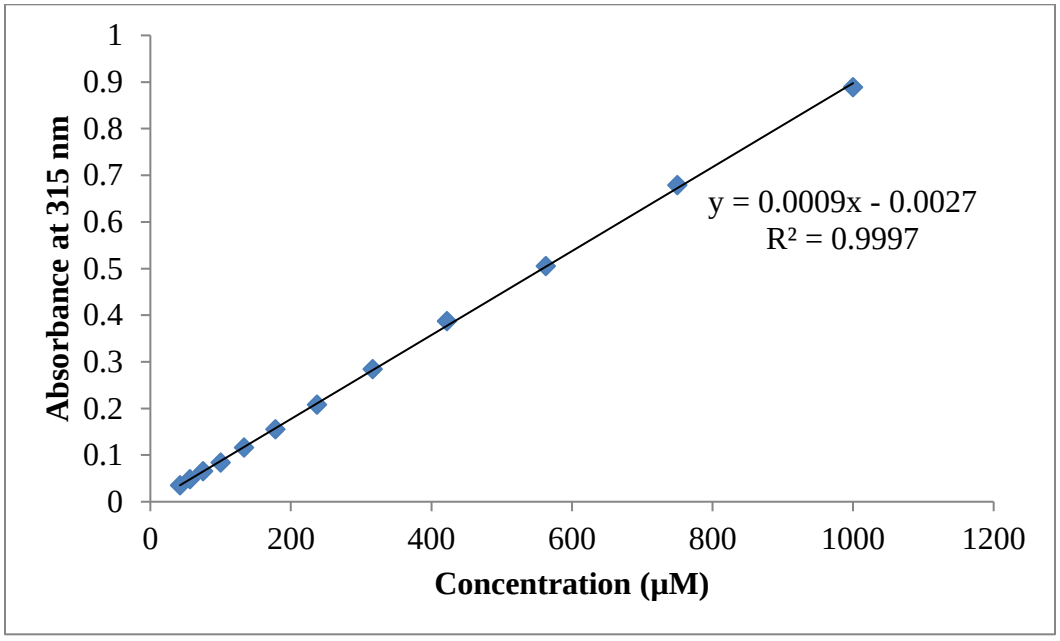


Figure 3.24. Extinction coefficient determination of **4** from the slope of the line of regression.

Isothermal Titration Calorimetry

ITC binding experiments were conducted using a Microcal AutoITC200. Titrations were performed at 25 ^{o}C in 10 mM potassium phosphate, pH 7.4 or 10 mM sodium borate, 100 mM NaCl, pH 8.6. The concentration of **5** was determined by measuring the UV–vis absorbance at 373 nm, using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. The concentration of **4** was determined from UV-vis absorbance at 315 nm. A solution of 0.9 – 1.2 mM of peptide was titrated into a 81-85 μM solution of **4** or **5**, using 2.0 μL increments every 3 min. Heats of dilution (determined from separate runs where the peptide was titrated into the ITC buffer) were subtracted prior to fitting. Binding curves were produced using the supplied Origin software and fit using one-site binding models.

Table 3.4. Summary of ITC data from using a one-site binding model. Values represent an average of 3 measurements. Errors are from standard deviation from three separate runs.

Host	Peptide	Net Charge	Kd (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	N	ΔG (Kcal/mol)
5[a]	K9me3R14	+3	4.8 ± 1.5	-6.6 ± 0.4	6.6±2.9	1.03± 0.04	-7.3 ± 0.2
4[b]	K9K14	+3	390 ±440	-	-	0.05 ± 0.05	-
4[b]	K9me3K14	+3	18.3 ± 7.7	-4.9 ± 0.7	1.62 ± 0.9	0.53 ± 0.006	-6.5 ± 0.2

(a) 10 mM potassium phosphate, pH 7.4; (b) 10 mM sodium borate, 100 mM NaCl, pH 8.6
N values that deviate from 1 are likely due to the error in concentration determination for **4**, but does not influence the measurement of K_d.

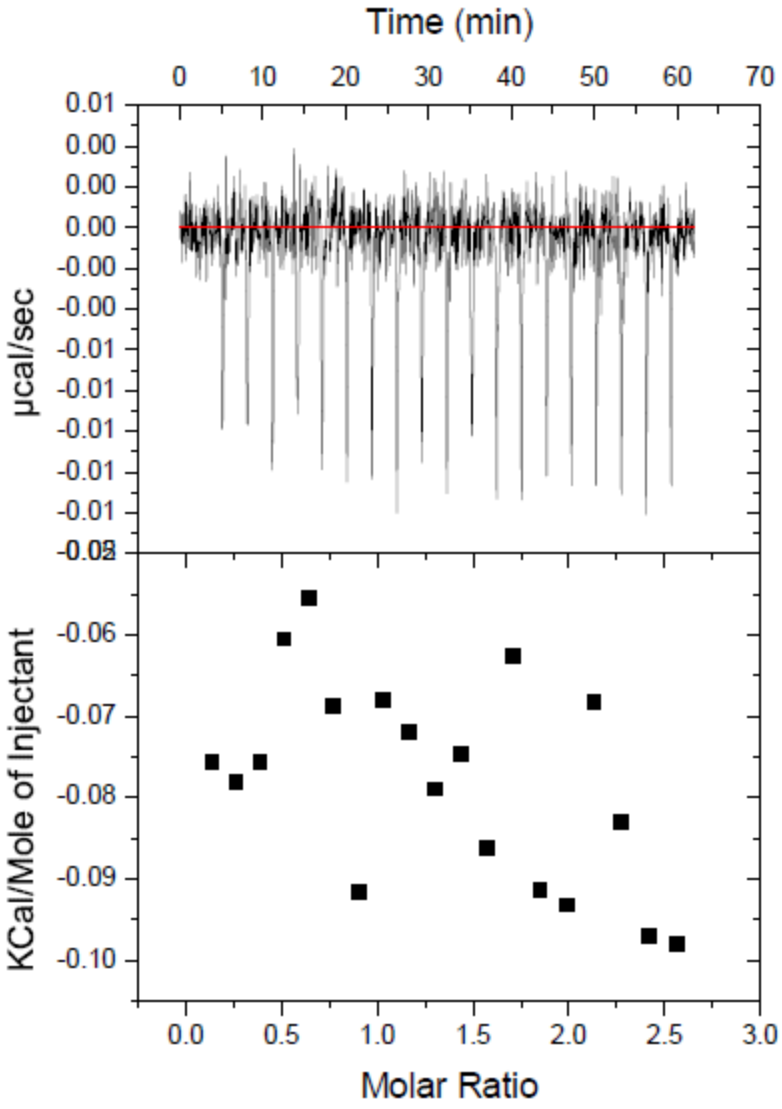


Figure 3.25. Heat of dilution for titrating K9me3 R14 (1.22 mM) into 10 mM potassium phosphate buffer pH 7.4.

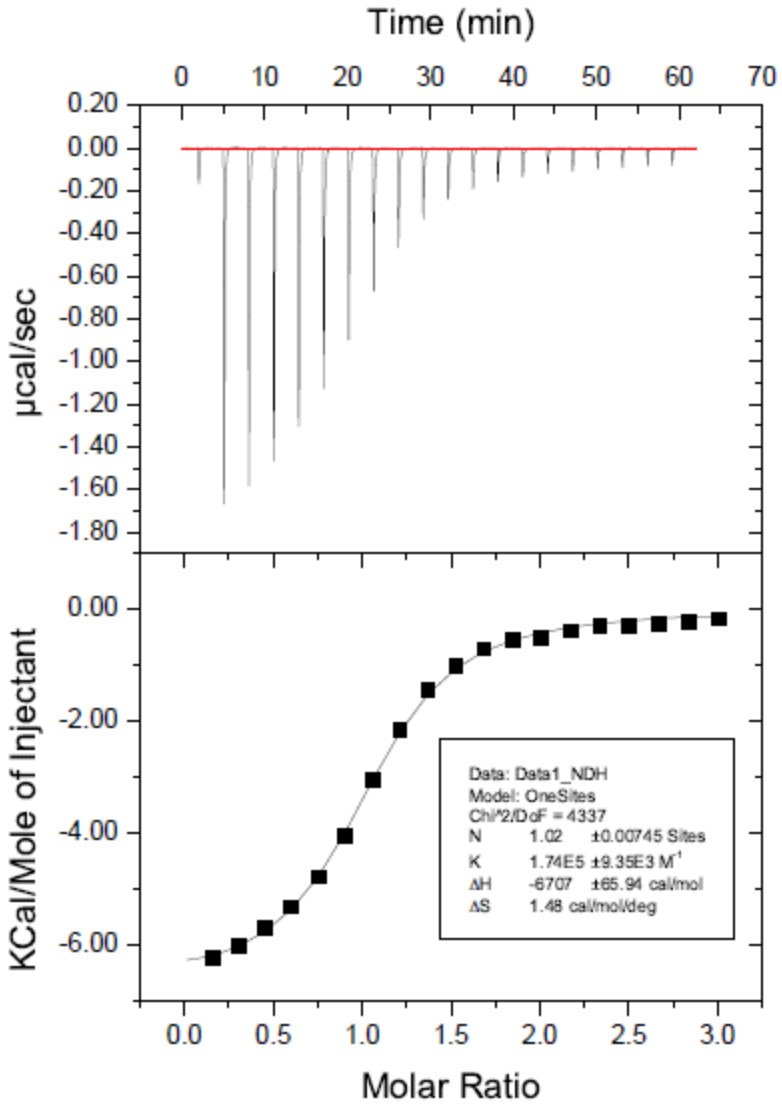


Figure 3.26. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.22 mM) titrated into **5** (82 µM) at 25 ºC in 10 mM potassium phosphate buffer, pH 7.4 (Run 1).

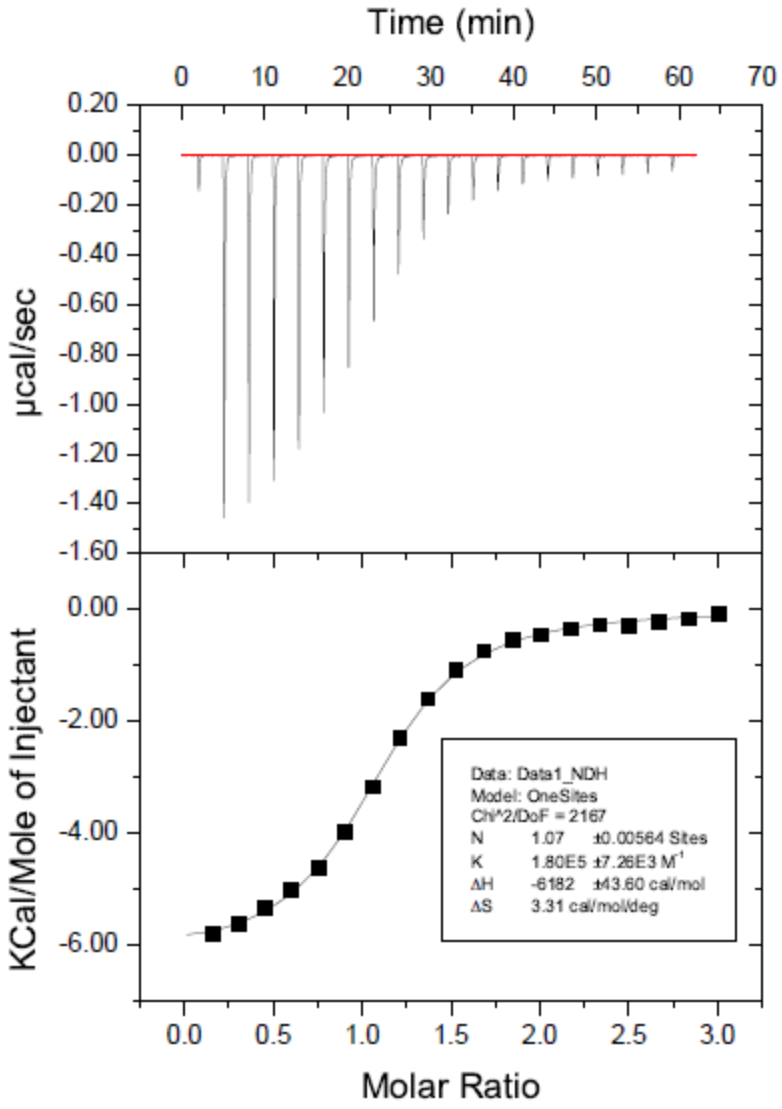


Figure 3.27. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.22 mM) titrated into **5** (82 µM) at 25 °C in 10 mM potassium phosphate buffer, pH 7.4 (Run 2).

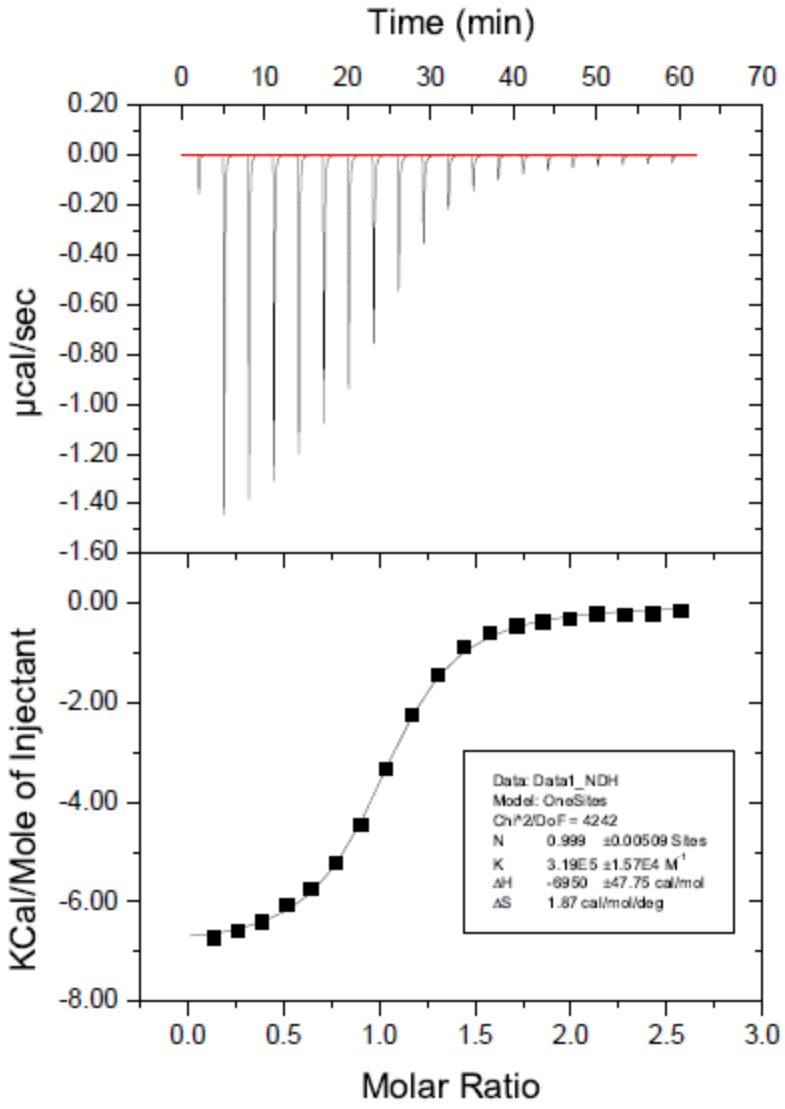


Figure 3.28. One trial of three of K9me3 R14 (Ac-WGGQTARKme3STGGR-NH_2) (1.03 mM) titrated into **5** (81 µM) at 25 °C in 10 mM potassium phosphate buffer, pH 7.4 (Run 3).

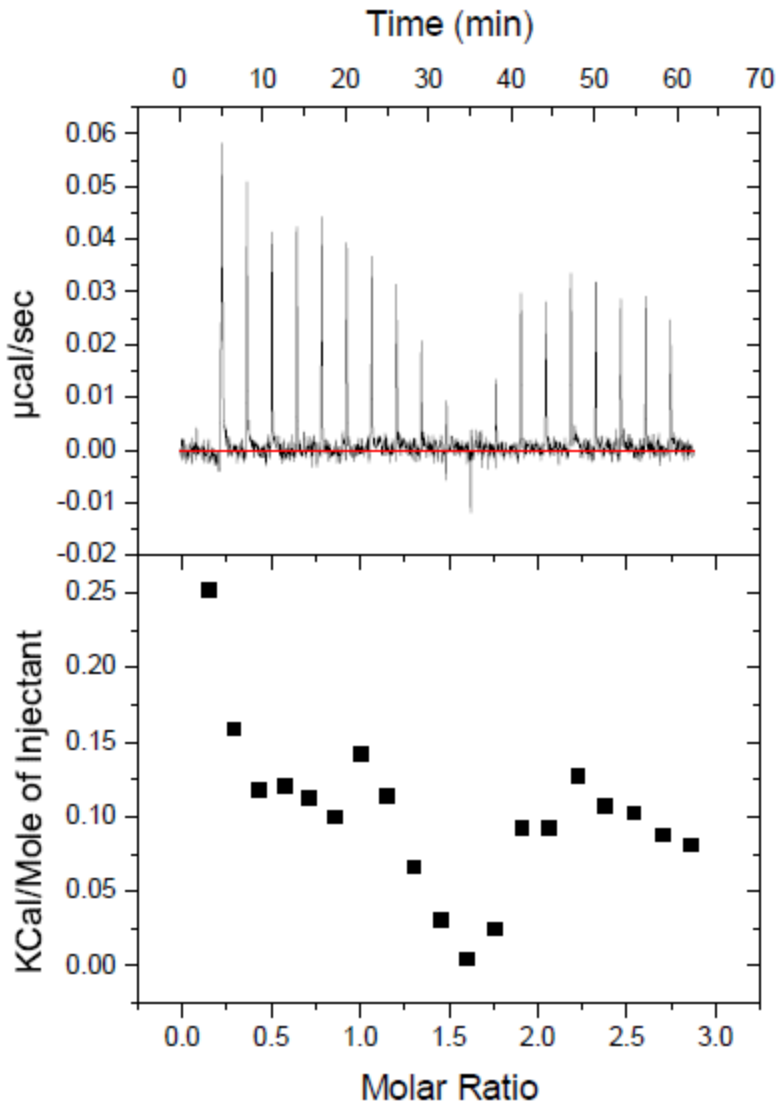


Figure 3.29. Heat of dilution for titrating K9K14 (Ac-WGGQTARKSTGGK-NH_2) (1.43 mM) into 10 mM sodium borate buffer pH 8.6.

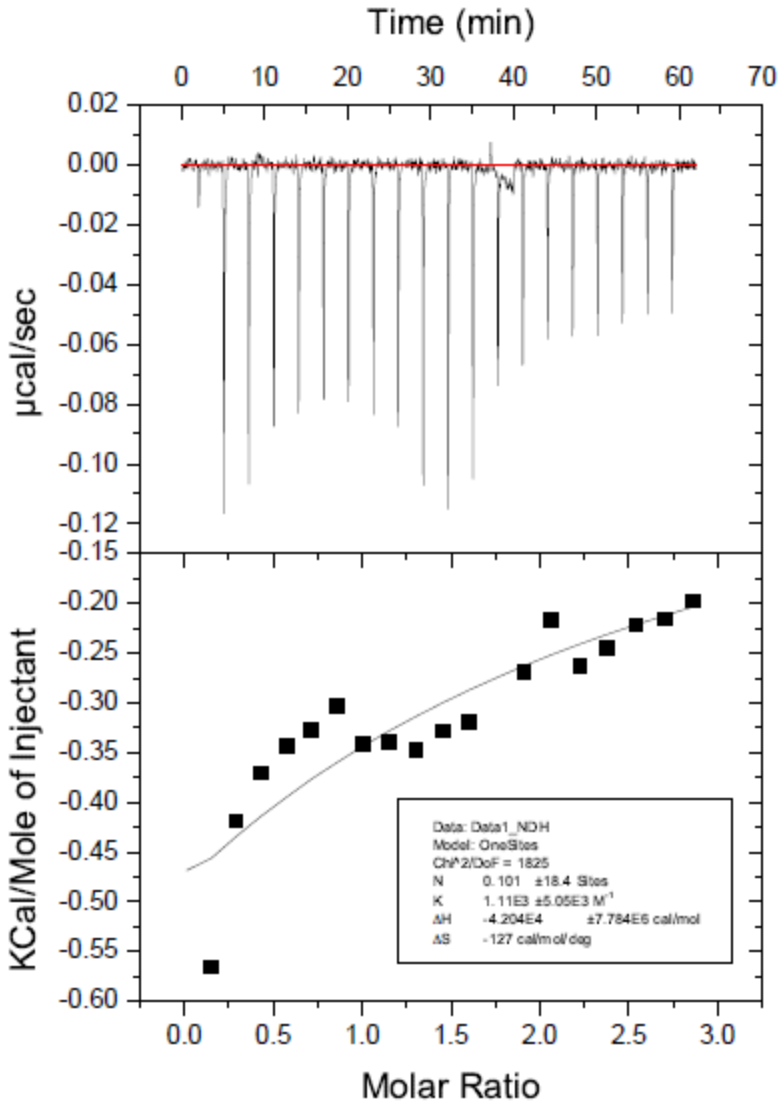


Figure 3.30. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (1.43 mM) titrated into **4** (101 µM) at 25 ºC in 10 mM sodium borate buffer, pH 8.6 (Run 1).

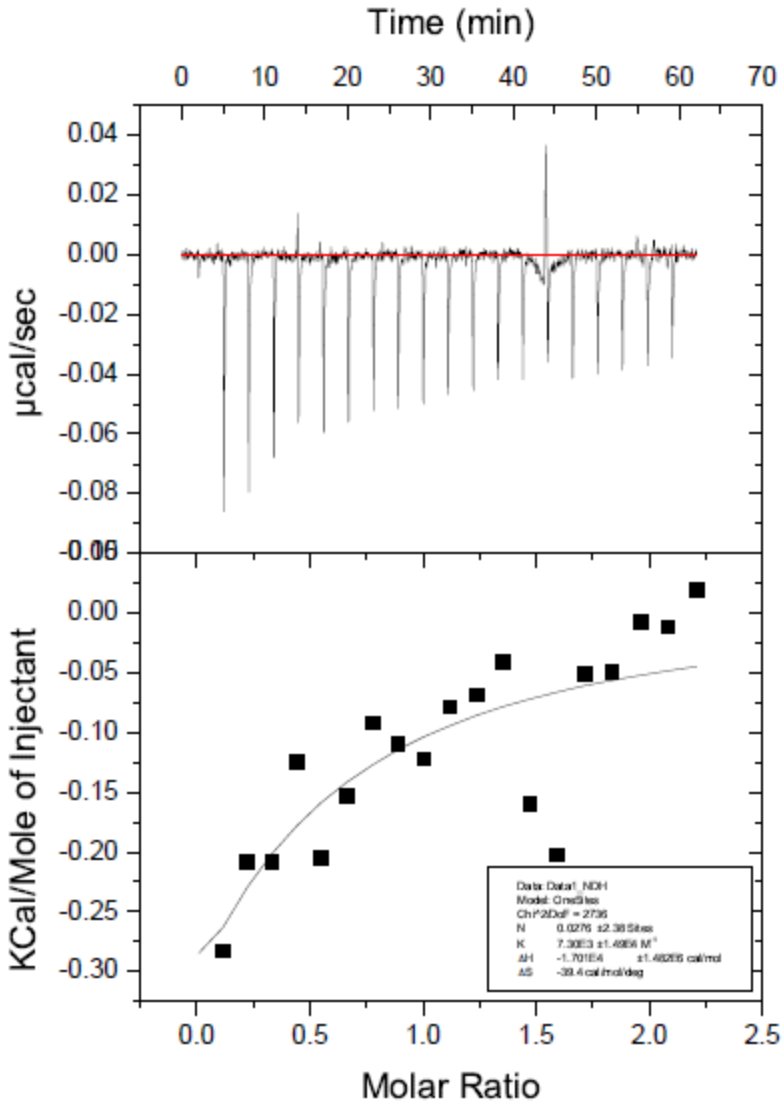


Figure 3.31. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (0.93 mM) titrated into **4** (85 µM) at 25 ^oC in 10 mM sodium borate buffer, pH 8.6 (Run 1).

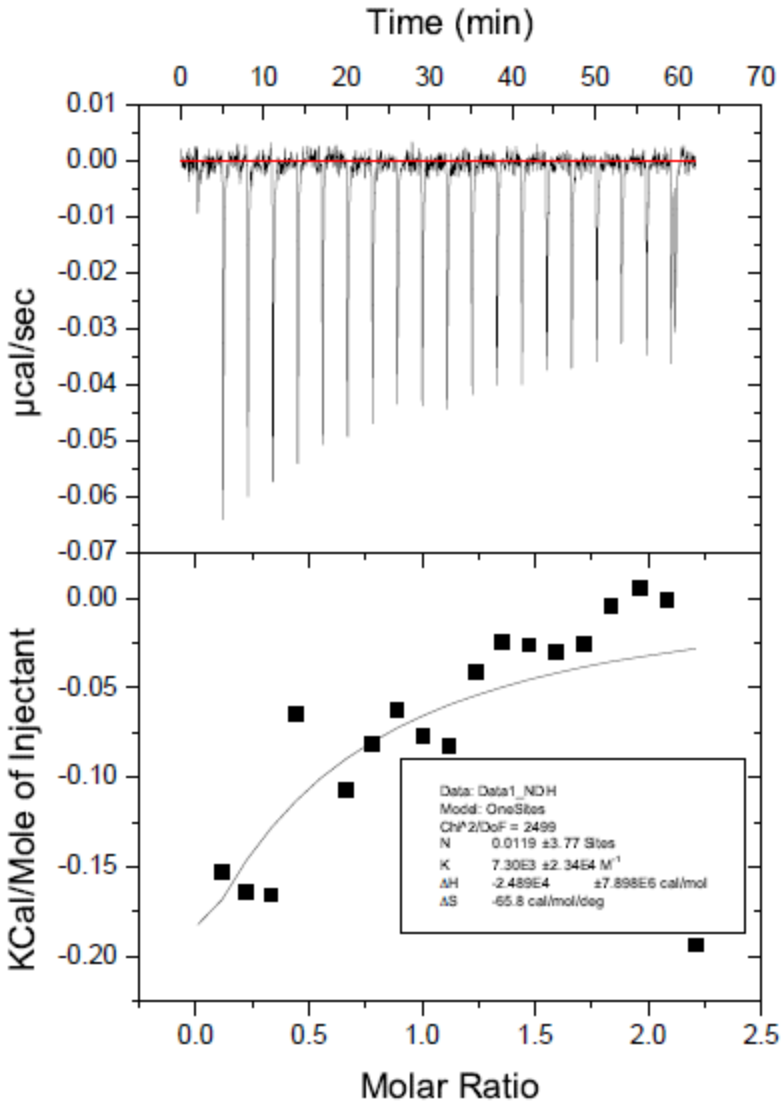


Figure 3.32. One trial of three of K9 K14 (Ac-WGGQTARKSTGGK-NH_2) (0.93 mM) titrated into **4** (85 µM) at 25 ºC in 10 mM sodium borate buffer, pH 8.6 (Run 2).

.

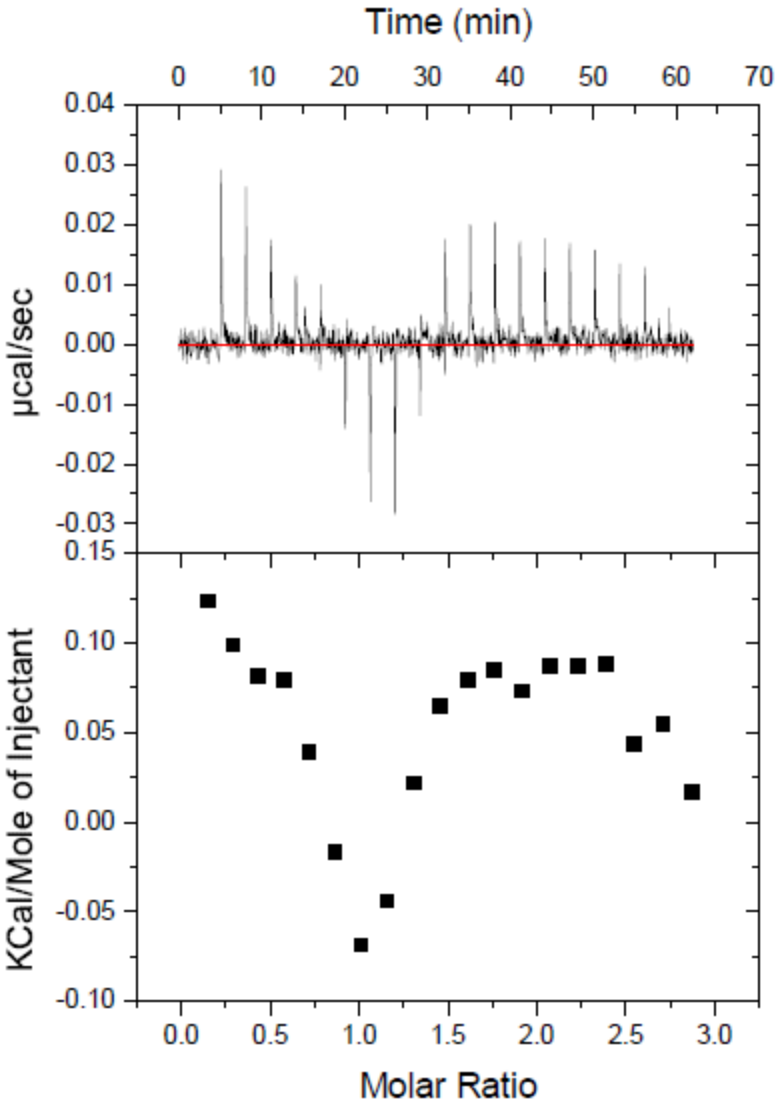


Figure 3.33. Heat of dilution for titrating K9me3K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.43 mM) into 10 mM sodium borate buffer pH 8.6.

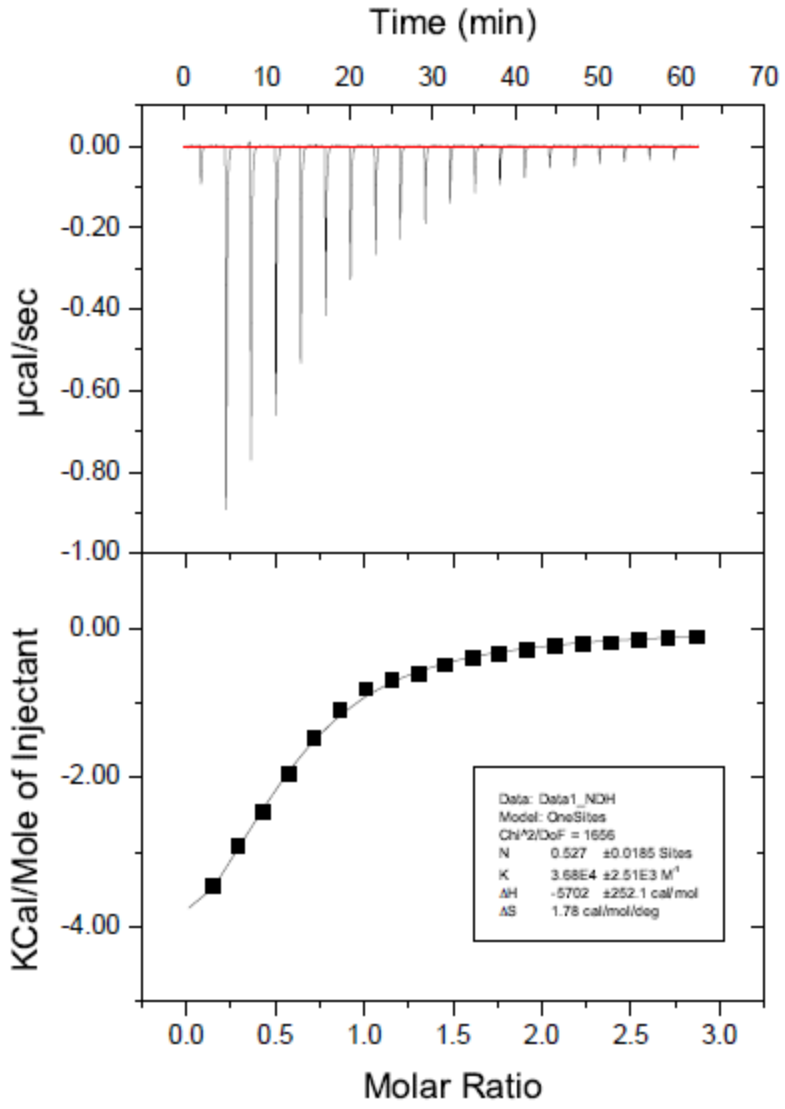


Figure 3.34. One trial of three of K9me$_3$K14 (Ac-WGGQTARKme3STGGK-NH$_2$) (1.43 mM) titrated into **4** (101 µM) at 25 ^{o}C in 10 mM sodium borate buffer, pH 8.6 (Run 1).

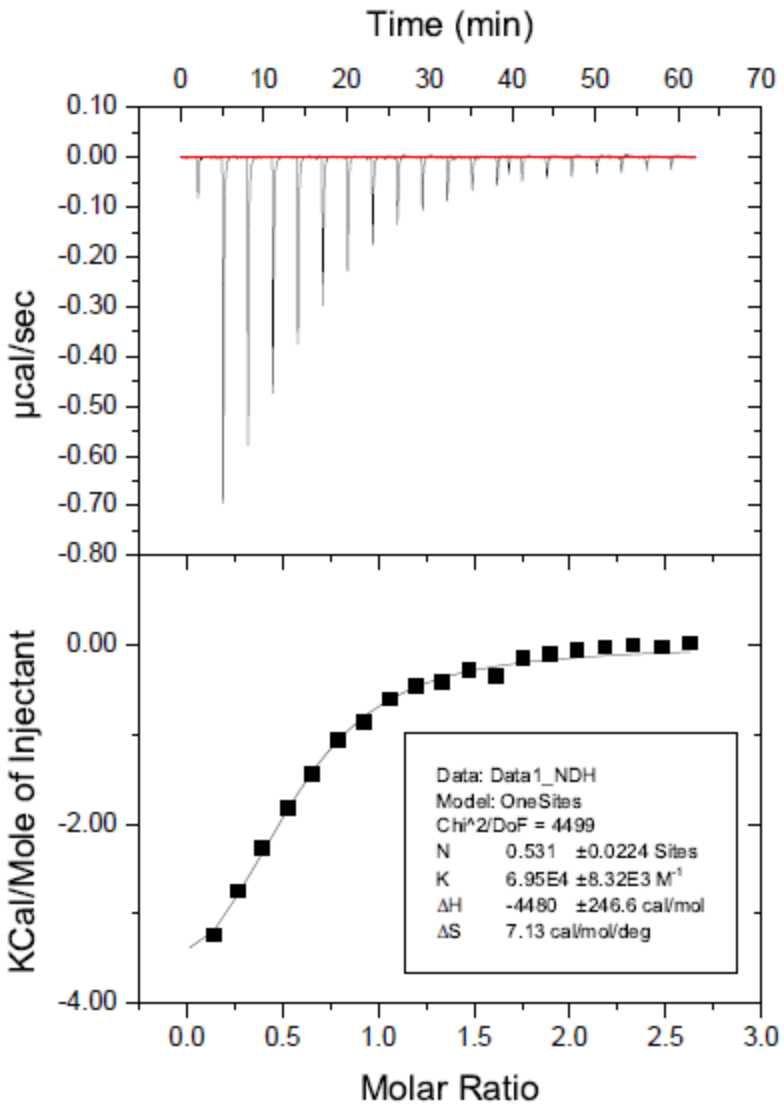


Figure 3.35. One trial of three of K9me$_3$ K14 (Ac-WGGQTARKme3STGGK-NH$_2$) (1.11 mM) titrated into **4** (85 µM) at 25 ^{o}C in 10 mM sodium borate buffer, pH 8.6 (Run 2).

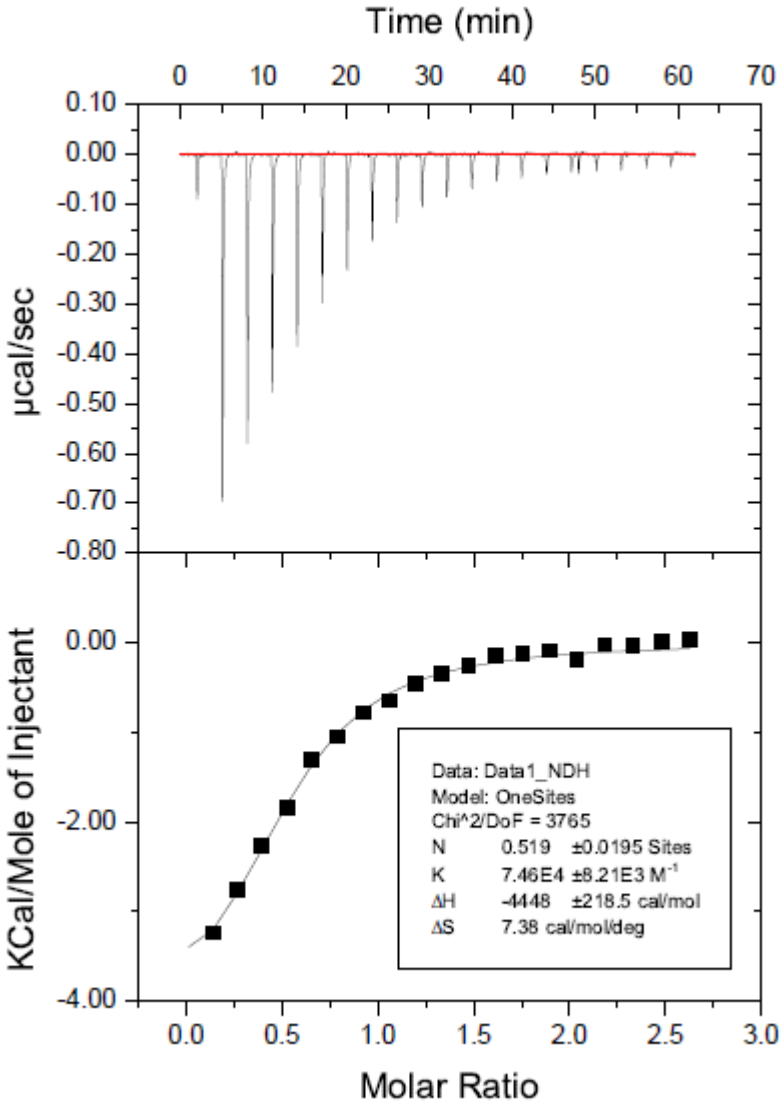


Figure 3.36. One trial of three of K9me3 K14 (Ac-WGGQTARKme3STGGK-NH_2) (1.11 mM) titrated into **4** (85 µM) at 25 ^oC in 10 mM sodium borate buffer, pH 8.6 (Run 3).

LCMS characterization of CX_4-ONBD labeling of K9K14 and K9me$_3$K14

Reactions were analyzed by LCMS using a C18 column at a flow rate of 0.2 mL/min. Chromatographic separation was carried out using a 25-min linear gradient of solvent A (95% H_2O, 5% CH_3CN, and 0.1% formic acid) and solvent B (95% CH_3CN, 5% H_2O, and 0.1% formic acid).

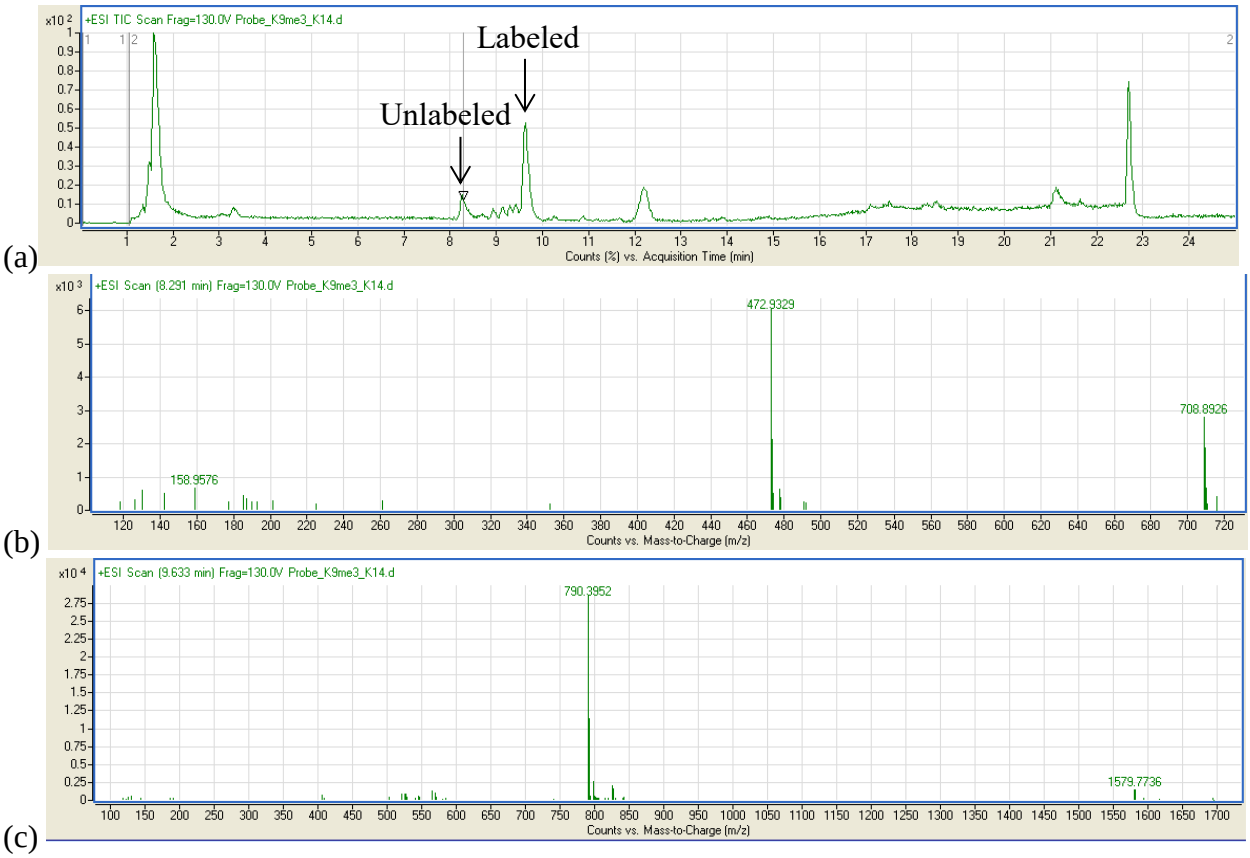


Figure 3.37. LCMS characterization of CX_4-ONBD labeling of K9K14 and K9me3K14 (a) LCMS absorbance trace of the reaction of 100 µM **5** with 100 µM K9me3 K14 in 10 mM sodium borate buffer, pH 8.6 after 48 h. (b) MS analysis of the peak at ~8.25 min (unlabeled peptide). MS (calculated): 708.89 (M+2) [Unlabeled+1H], 472.93 (M+3) [Unlabeled+2H]. MS (observed, ESI+): 708.8926 (M+2) [Unlabeled+1H], 472.9329 (M+3) [Unlabeled+2H]. (c) MS analysis of the peak at ~9.75 min (labeled peptide). MS (calculated): 1579.77 (M+1) [Labeled], 790.39 (M+2) [Labeled+1H]. MS (observed, ESI+):1579.7736 (M+1) [Labeled], 790.3952 (M+2) [Labeled+1H].

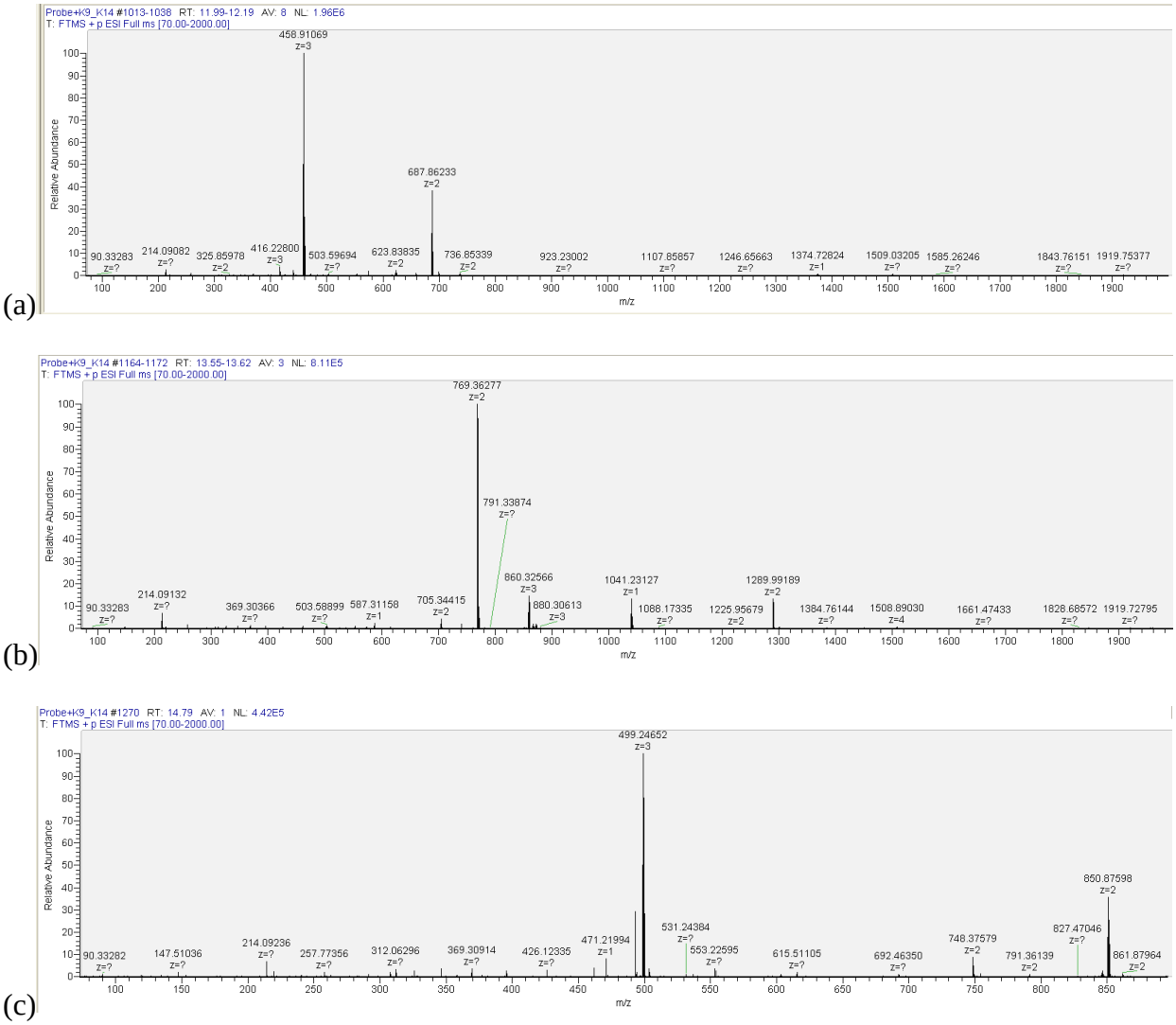


Figure 3.38. MS spectra of the reaction of 100 μM **5** with 100 μM K9 K14 in 10 mM sodium borate buffer, pH 8.6 after 24 h. (a) MS analysis of the peak at ~12.09 min (unlabeled peptide). MS (calculated): 687.87 (M+2) [Unlabeled+1H], 458.91 (M+3) [Unlabeled+2H]. MS (observed, ESI+): 687.86233 (M+2) [Unlabeled+1H], 458.91069 (M+3) [Unlabeled+2H]. (b) MS analysis of the peak at ~13.37 min (singly labeled). MS (calculated): 769.37 (M+2) [Mono-Labeled+2H]. MS (observed, ESI+): 769.36277 (M+2) [Mono-Labeled+2H]. (c) MS analysis of the peak at ~14.79 min (doubly labeled). MS (calculated): 850.88 (M+2) [Di-Labeled+2H]. MS (observed, ESI+): 850.87598 (M+2) [Di-Labeled+2H].

HPLC Analysis of Labeling

Reactions were analyzed by analytical HPLC using a C18 column at a flow rate of 1 mL/min. Chromatographic separation was carried out using a 45 min linear gradient of solvent A (95% H_2O, 5% CH_3CN, and 0.1% TFA) and solvent B (95% CH_3CN, 5% H_2O, and 0.1% TFA). Absorbance was monitored at 214 nm.

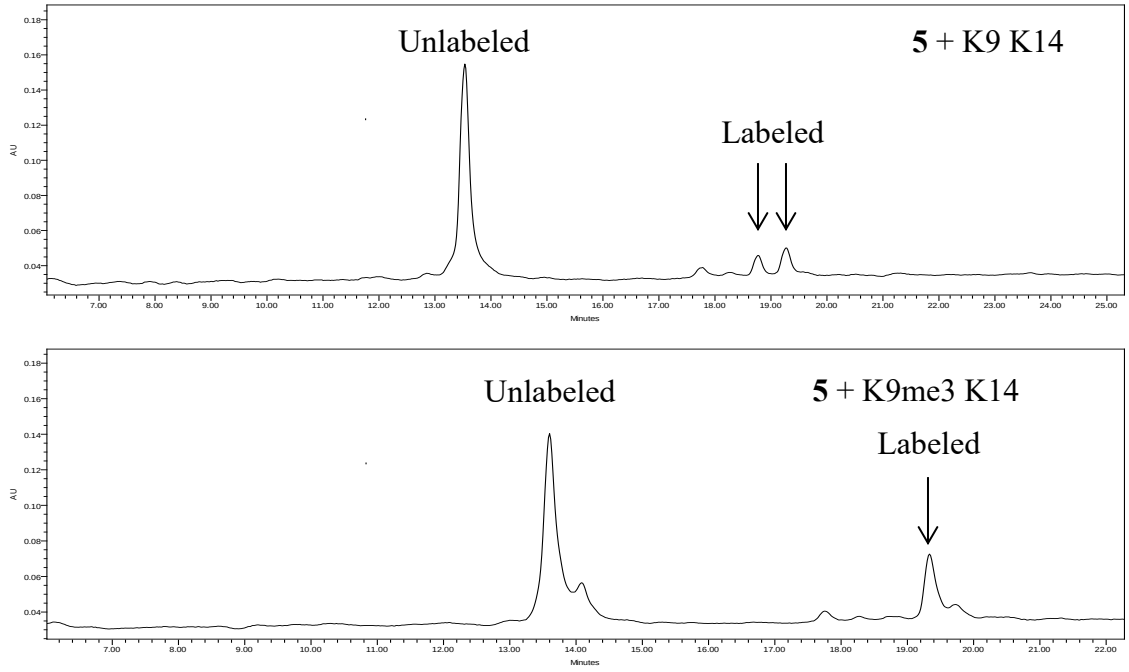


Figure 3.39. HPLC trace of the reaction of 100 μM **5** with 100 μM K9 K14 (top) or 100 μM K9me3 K14 (bottom) in 10 mM potassium phosphate buffer, pH 7.4 after 24 h showing labeled and unlabeled peptide. .

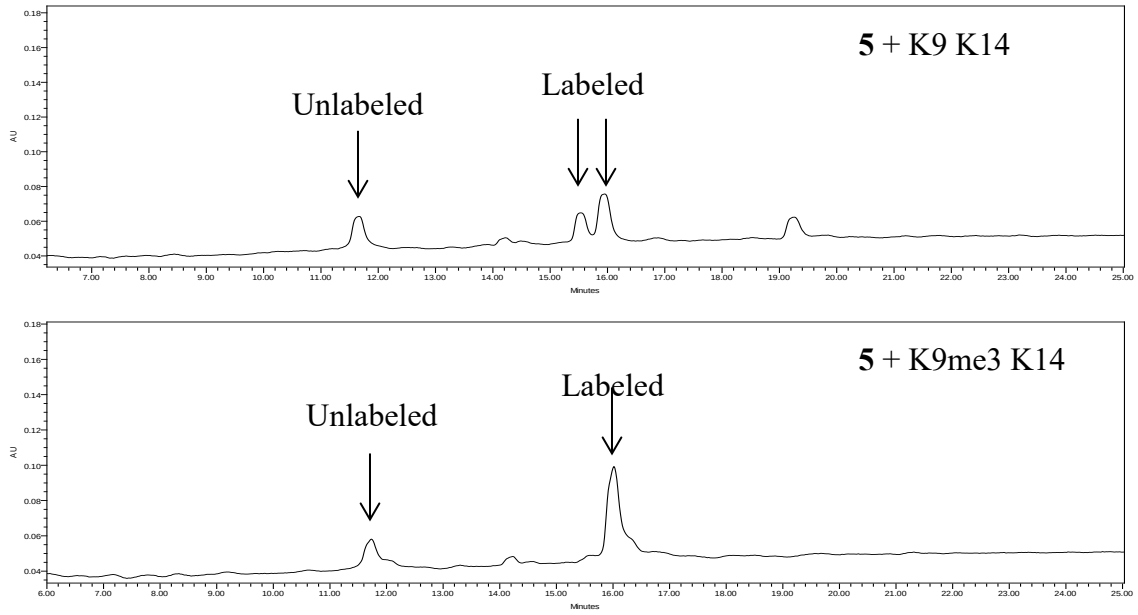


Figure 3.40. HPLC trace of the reaction of 100 μM **5** with 50 μM K9 K14 (top) or 50 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6 after 24 h.

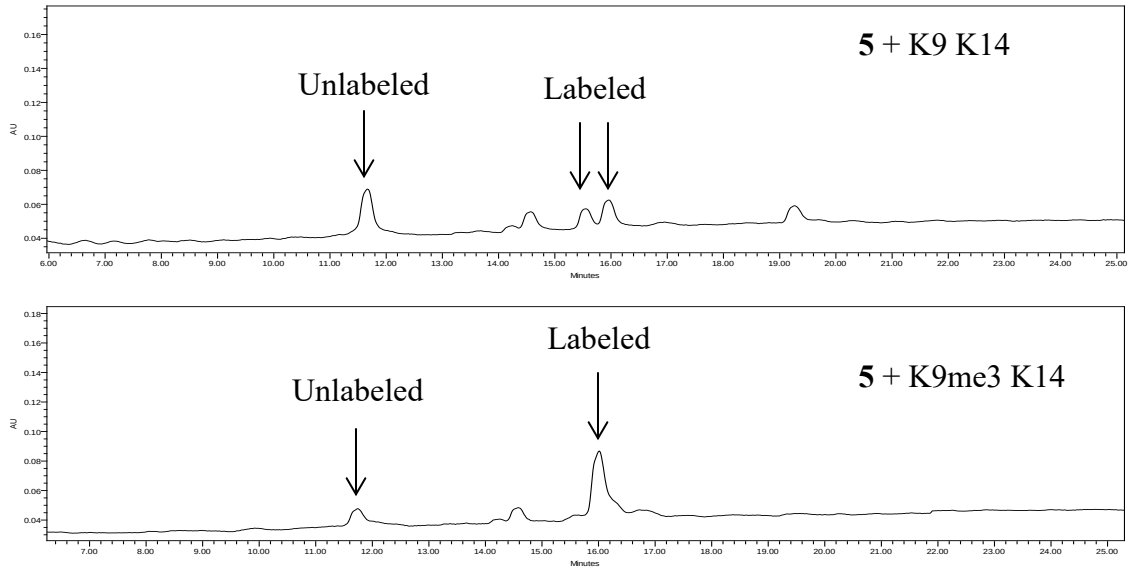


Figure 3.41. HPLC trace of the reaction of 100 μM **5** with 50 μM K9 K14 (top) or 50 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 50 mM NaCl after 24 h.

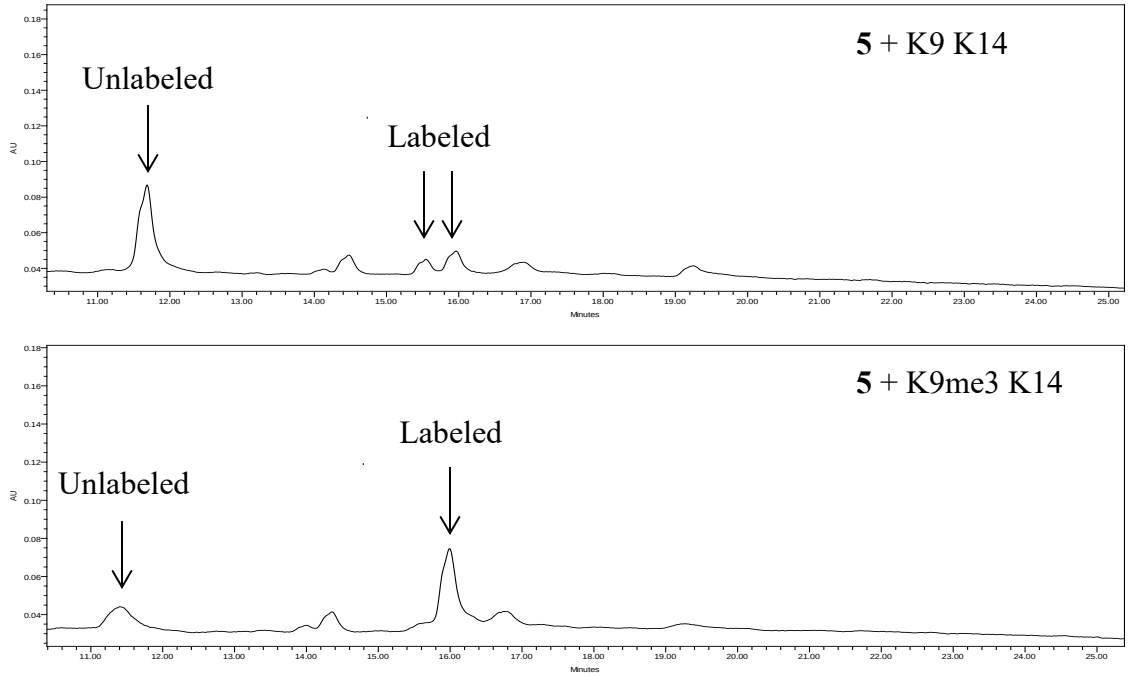


Figure 3.42. HPLC trace of the reaction of 100 μM 5 with 30 μM K9 K14 (top) or 30 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h.

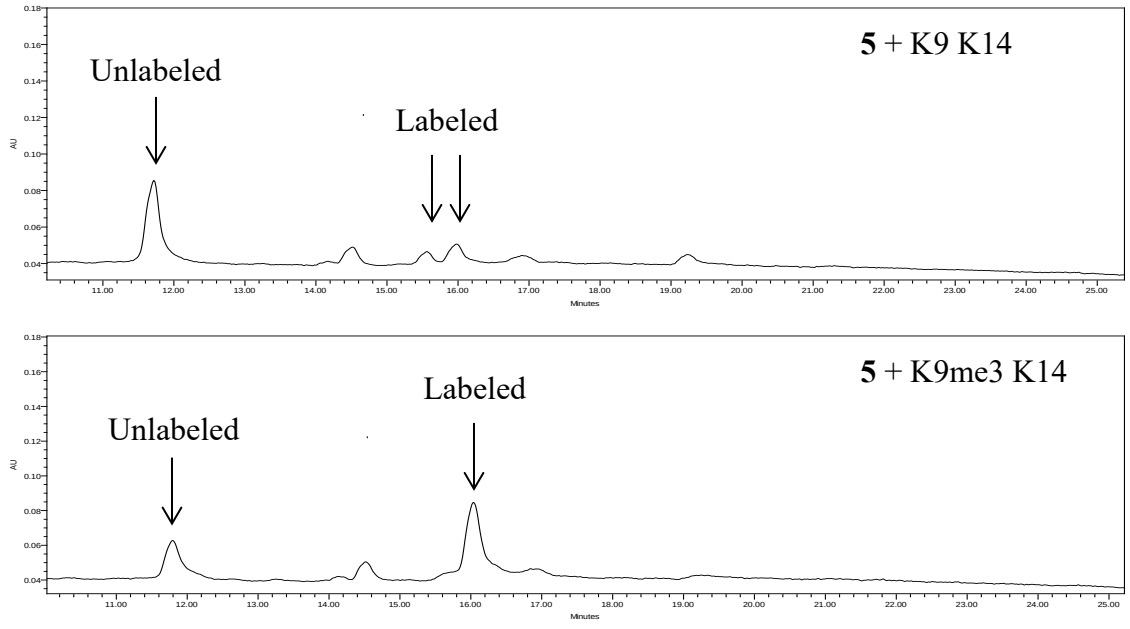


Figure 3.43. HPLC trace of the reaction of 100 μM **5** with 30 μM K9 K14 (top) or 30 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 150 mM NaCl after 24 h showing labeled and unlabeled peptide.

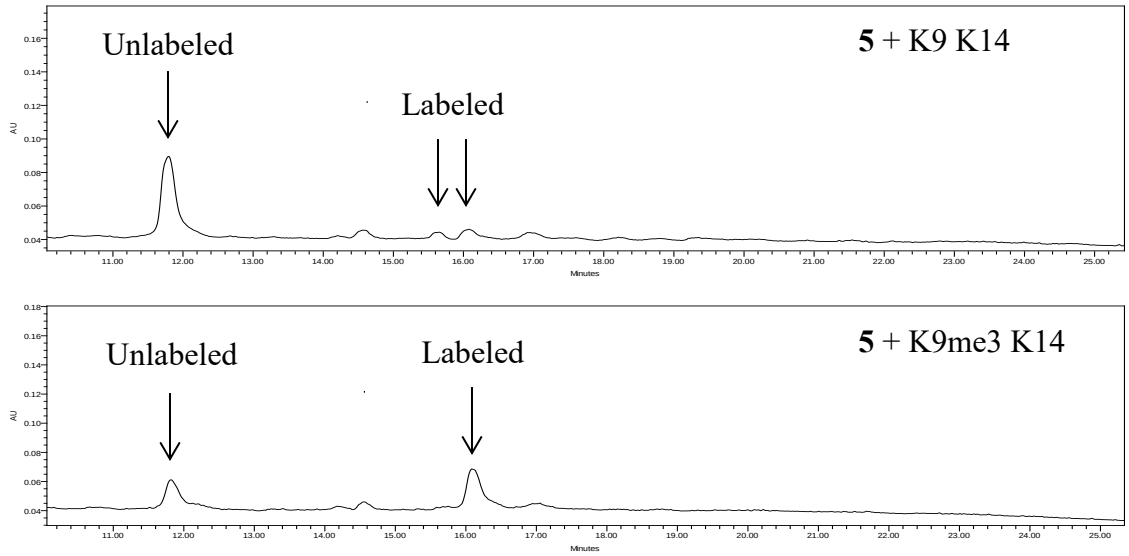


Figure 3.44. HPLC trace of the reaction of 40 μM **5** with 20 μM K9 K14 (top) or 20 μM K9me3 K14 (bottom) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h showing labeled and unlabeled peptide.

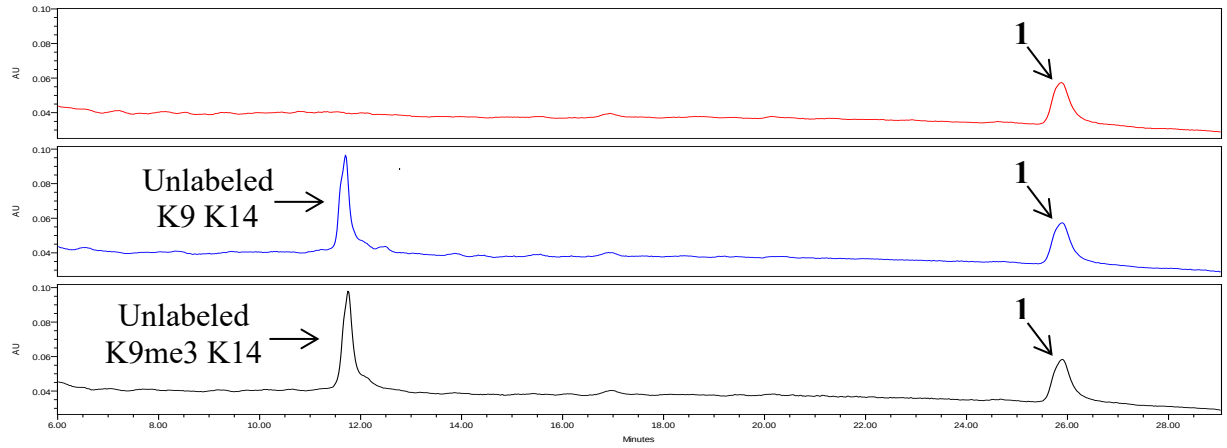


Figure 3.45. HPLC trace of the reaction of 40 μM **1** with 20 μM K9 K14 (blue) or 20 μM K9me3 K14 (black) in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl after 24 h demonstrating that the peptides do not react with a probe lacking the recognition motif.

Kme3/K Competition Labeling Experiment

To assess the ability of probe **5** to selectively label trimethylated peptides in the presence of a lysine-containing peptide, competition labeling experiments were conducted. RKme3GLLYK was used for the trimethylated peptide, and K9K14 was used for the unmodified lysine peptide. Two conditions were tested: one reacting 40 μM of **5** with a mixture of 20 μM RKme3GLLYK and 40 μM K9K14 and another reacting 40 μM of **5** with 30 μM RKme3GLLYK and 30 μM K9K14. Reactions were analyzed by analytical HPLC using a C18 column at a flow rate of 1 mL/min. Chromatographic separation was carried out using a 100 min linear gradient of solvent A (95% H_2O, 5% CH_3CN, and 0.1% TFA) and solvent B (95% CH_3CN, 5% H_2O, and 0.1% TFA). Absorbance was monitored at 214 nm.

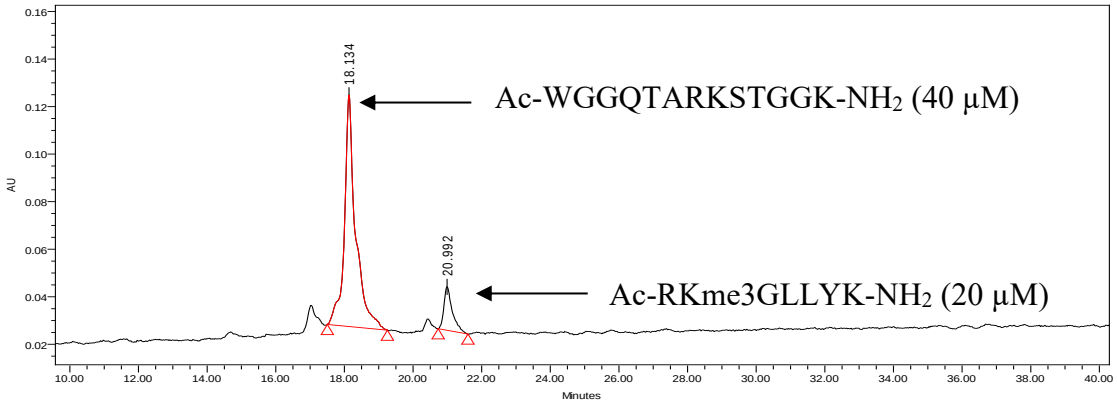


Figure 3.46. HPLC trace of a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl to provide a concentration reference.

Table 3.5. Summary of peak integration data from HPLC trace of a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK. Error is estimated to be near 10%.

	Retention Time	Area	% Area	Height
1	18.134	2210624	87.60	97325
2	20.992	312987	12.40	18352

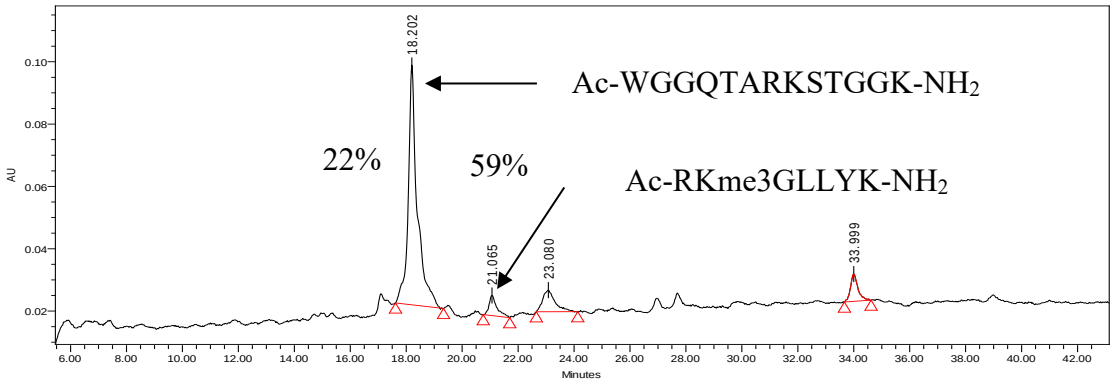


Figure 3.47. HPLC trace of the reaction of 40 μM probe **5** with a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl. After 24 h, 22% of K9K14 reacted with the probe while 59% of RKme3GLLYK was labeled by the probe.

Table 3.6. Summary of peak integration and labeling data from the HPLC trace of the reaction of 40 μM probe **5** with a mixture of 40 μM K9K14 and 20 μM RKme3GLLYK.

	Retention Time	Area	% Area	Height	Percent of Peptide Labeled
1	18.202	1725267	77.34	76903	22%
2	21.065	127580	5.72	6488	59%
3	23.080	215293	9.65	6974	-
4	33.999	162552	7.29	8771	-

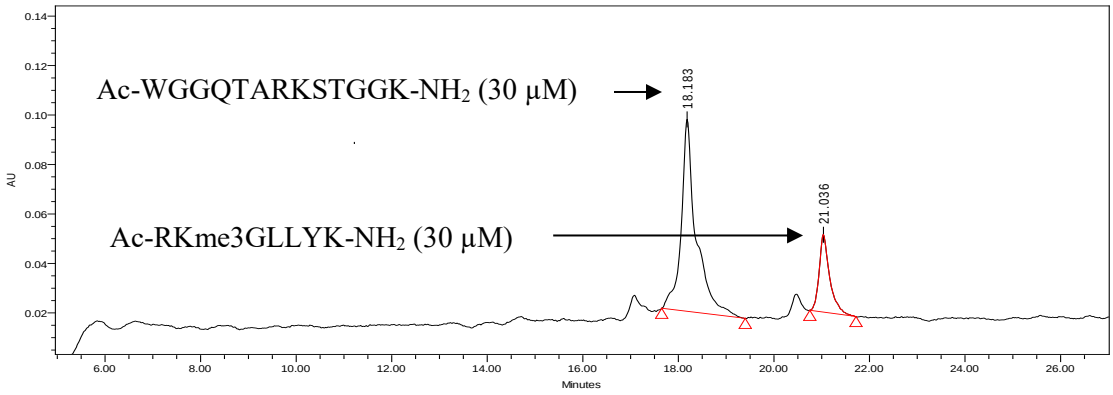


Figure 3.48. HPLC trace of a mixture of 30 µM K9K14 and 30 µM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl to provide a concentration reference.

Table 3.7. Summary of peak integration data HPLC trace of a mixture of 30 µM K9K14 and 30 µM RKme3GLLYK. Error is estimated to be near 10%.

	Retention Time	Area	% Area	Height
1	18.183	1689794	77.88	77563
2	21.036	479927	22.12	31370

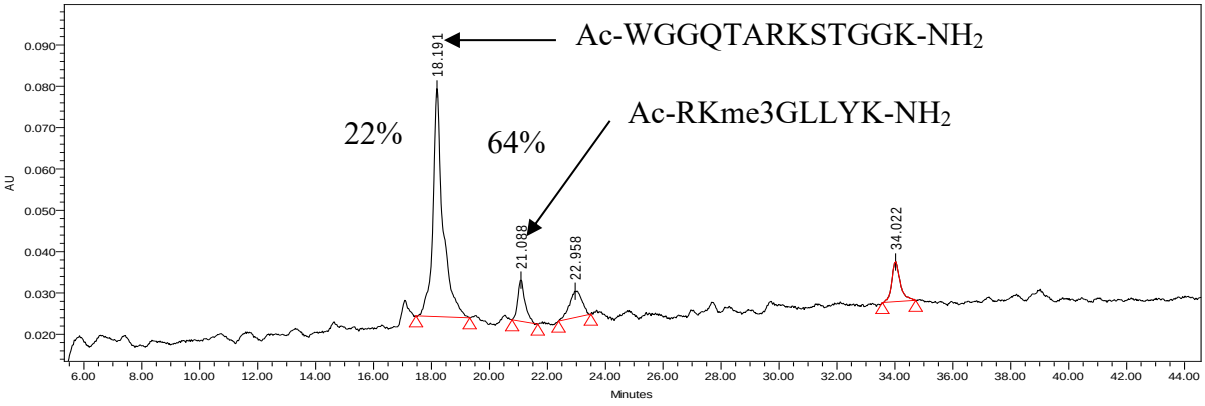


Figure 3.49. HPLC trace of the reaction of 40 µM probe **5** with a mixture of 30 µM K9K14 and 30 µM RKme3GLLYK in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl. After 24 h, 22% of K9K14 reacted with the probe while 64% of RKme3GLLYK was labeled by the probe.

Table 3.8. Summary of peak integration data from the HPLC trace of the reaction of 40 µM probe **5** with a mixture of 30 µM K9K14 and 30 µM RKme3GLLYK. Error is estimated to be near 10%.

	Retention Time	Area	% Area	Height	Percent of Peptide Labeled
1	18.191	1324422	70.17	55268	22%
2	21.088	174149	9.23	10016	64%
3	22.958	192180	10.18	6424	-
4	34.022	196655	10.42	9641	-

Turn-On Fluorescence Labeling Control Experiments

Fluorescence labeling experiments were performed in black 96 well optical bottom plates (non-treated, non-sterile, polystyrene) using a POLARstar Omega (BMG Labtech, Inc.) plate reader. A fluorescence optic with a 485 nm excitation filter and a 520 nm emission filter was used. Labeling reactions were carried out at ambient temperature, and Titer Tops® sealing film was used to prevent evaporation. Fluorescence intensity was measured every hour for 24 h.

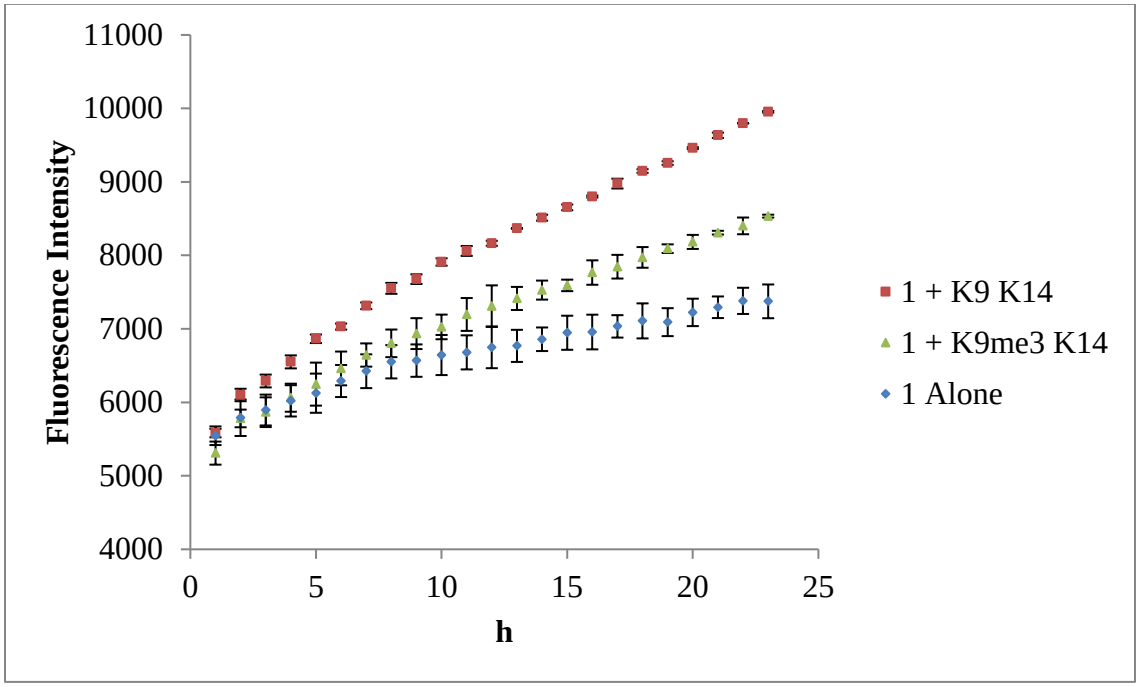


Figure 3.50. Turn-on fluorescence labeling negative control experiment. Labeling of 40 μM of **1** alone (blue) or in the presence of 100 μM K9 K14 (red) or 100 μM K9me3 K14 (green) in 10 mM potassium phosphate buffer (pH 7.4). Each data point represents an average of three runs. Error bars are standard deviations.

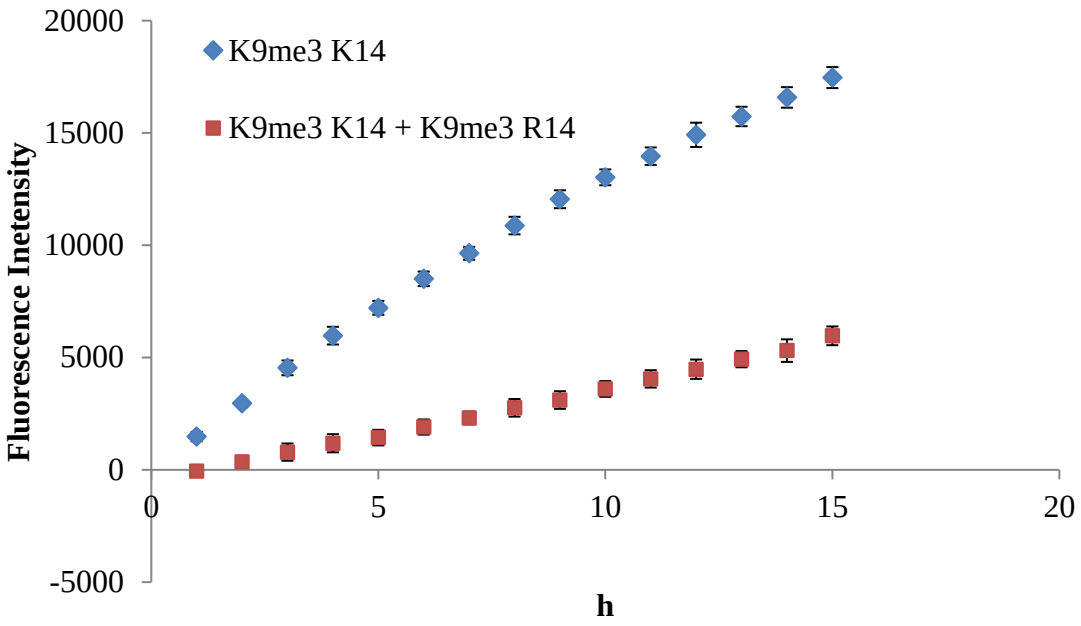


Figure 3.51. Turn-on fluorescence labeling competition experiment. 40 μM of **5** was reacted with 20 μM K9me3 K14 in the presence (red) or absence of 120 μM K9me3 R14 (blue) in 10 mM sodium borate buffer (pH 8.6), 100 mM NaCl. Each data point represents an average of three runs. Error bars are standard deviations.

CHAPTER 4. A Supramolecular Affinity Labeling-Based Turn-On Fluorescence Assay for Histone Deacetylase Activity[1]

Turn-On Fluorescence Assay for HDAC Activity

After establishing selective turn-on fluorescence labeling of a Kme3 histone peptide, we applied this labeling methodology toward the development of a chemical assay for histone deacetylase (HDAC) activity. HDACs play a crucial role in regulating chromatin structure and ultimately gene expression via removal of acetyl groups from the sidechains of acetylated lysine residues on histone tails.[124,125] Generally, deacetylation of histones is associated with a condensed chromatin state and repression of gene transcription. Many HDACs function as part of multi-unit protein complexes and are implicated in processes ranging from cell differentiation to cell development.[126,127] Expectedly, dysregulation of HDACs often has detrimental biological effects, in some cases promoting oncogenesis or the development of inflammatory disease.[128,129]

Existing commercial kits for measuring HDAC activity predominantly rely on the use of a peptide or amino acid substrate with a C-terminal fluorogenic 7-amino-4-methylcoumarin tag directly linked to an acetylated lysine residue.[130] Removal of the acetyl group by an HDAC allows trypsin to be used in a subsequent development step to cleave the C-terminal amide, releasing a fluorophore. While this method is convenient and widely used and HDACs generally lack significant *in vitro* sequence selectivity, it is nonetheless severely limited in substrate scope

[1] Reprinted (adapted) with permission from Gober, I. N.; Waters, M. L. *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459. Copyright (2016) American Chemical Society.

since the coumarin tag must be immediately adjacent to the acetyl lysine. Furthermore, when screening for novel HDAC inhibitors, the required secondary trypsin reaction presents the possibility of false positives.[131,132] We envisioned designing a HDAC inhibitor assay in which probe **5** could be used to determine the extent of deacetylation in the context of a peptide containing an acetylated lysine residue adjacent to trimethyl lysine. Trimethylation and acetylation of lysine are known to co-exist and act cooperatively to influence the function of chromatin-modifying proteins.[133]

To demonstrate this assay concept, we selected histone deacetylase 1 (HDAC1) as a model histone deacetylase. HDAC1 is a class I HDAC with a zinc dependent active site whose function is disrupted by inhibitors containing hydroxamate functionality, including vorinostat (SAHA).[134] For our assay, we prepared two H3 peptides representative of biologically relevant HDAC substrates (K4me3K9ac and K27acK36me3). Histone H3 acetylation at K9 and K27 as well as methylation at K4 are histone PTMs known to be associated with active gene transcription.[135,136] Conversely, methylation at K36 has been linked to heterochromatin and may signal the recruitment of HDACs.[137,138] Furthermore, it has been shown that all four PTMs (K4me3, K9ac, K27ac, and K36me3) co-occur on a single H3 histone in cells treated with the HDAC inhibitor sodium butyrate.[139] K4me3K9ac has the same spacing between the two Lys residues as our initial model peptide (K9me3K14), whereas the spacing is larger for K27acK36me3. Thus, comparison of these two peptides provides information about the distance-dependence of labeling.

Reactions between acetylated Kme3 peptides (200 μM) and HDAC1 (0.1 μg/μL) were carried out in 25 mM bicine buffer, 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl2 at pH 8.02 and 37 °C in the presence or absence of SAHA (5 μM). After 1 h, reactions were diluted with the

bicine reaction buffer to quench the reaction before addition to a 96-well plate containing **5**. Fluorescence labeling was performed with a final concentration of 40 µM **5** and 20 µM peptide at ambient temperature for 24 h. Samples containing only **5** and HDAC1 were used as negative controls and showed a slow increase in fluorescence over time, likely due to the presence of the primary amine Tris in the stock buffer of the commercial enzyme. H3 peptides K4me3K9 and K27K36me3 were used as positive controls for determining the maximum extent of fluorescence labeling possible for peptides deacetylated by HDAC1. Additionally, the CX4-ONBD showed no reactivity toward the hydroxamate group of the SAHA inhibitor (see Experimental), which further displays the probe's selectivity toward amine nucleophiles.

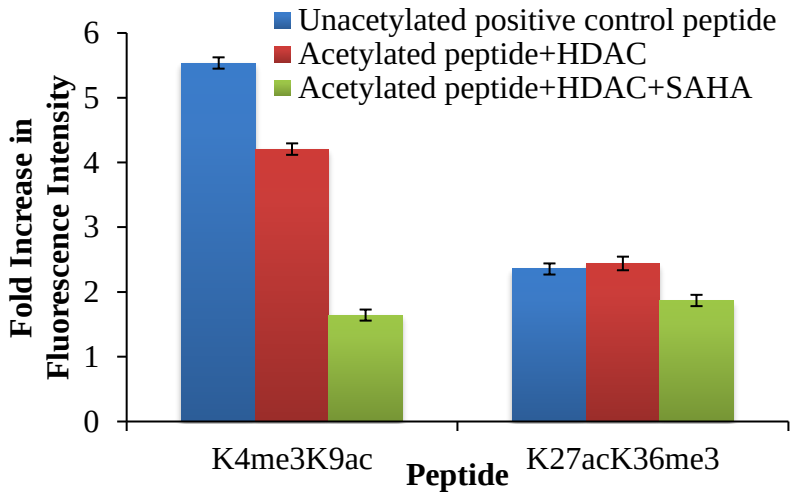


Figure 4.1. Turn-on fluorescence HDAC1 activity assay. Fluorescence monitoring of the deacetylation of histone H3 peptides K4me3 K9ac and K27ac K36me3 with probe **5** after reaction with HDAC1 in the absence or presence of 5 µM SAHA for 1 h at 37 °C. Enzymatic conditions are described in the text. Fluorescence response is shown after labeling for 24 h. The fold increase in fluorescence intensity for each reaction was normalized to the reaction containing HDAC1 and **5**. Blue bars: unacetylated control peptides K4me3K9 and K27K36me3; Red bars: acetylated peptides K4me3K9ac or K27acK36me3 + HDAC1; Green bars: acetylated peptides K4me3K9ac or K27acK36me3 + HDAC1 + SAHA.

As seen in Figure 4, this assay clearly distinguishes between active and inhibited enzyme in the case of K4me3K9ac. The fluorescence signal for the sample containing HDAC1 does not fully recapitulate the positive control because the enzymatic reaction does not reach completion

under these conditions, as analyzed by HPLC (see Experimental). Likewise, SAHA did not completely inhibit the reaction because conditions were not optimized to maximize inhibition. The fact that differences in fluorescence are clearly distinguishable despite the unoptimized conditions bodes well for application to inhibitor screens. Moreover, analysis of the fluorescence output after 6 h exhibits identical results as the data at 24 h, despite low conversion, indicating that shorter assay times are feasible (see Experimental).

To illustrate the distance-dependence on labeling in the context of our HDAC assay, we also investigated K27acK36me3, which has eight amino acids separating Kme3 and K compared to only four amino acids for K4me3K9. Not surprisingly, labeling of K4me3K9 was faster than that of K27K36me3, as indicated by the difference in fluorescence intensity of the positive controls (Figure 4). This distance dependence has a negative impact on the ability of the probe to report on HDAC activity for the K27acK36me3 peptide, as there is little difference in the degree of labeling in the absence or presence of SAHA (Figure 4). This suggests that the labeling may not be significantly receptor-mediated in this case. Importantly, in histone 3, all but two Lys residues have another Lys within 5 residues or less (K56 and K79 are the exceptions), meaning that this first-generation probe is quite general already.

In order to further validate the utility of our turn-on fluorescence assay for HDAC inhibitor screening, we carried out a dose-response experiment for the inhibition of HDAC3/NCOR1 by SAHA. For these experiments, inhibitor was serially diluted into 50 mM bicine buffer (137 mM NaCl, 2.7 mM KCl, 1 mM MgCl2, 0.5 mg/mL BSA, pH 8.0), and 11.2 nM HDAC3/NCOR1 was allowed to react with 100 μM K4me3K9ac for 4 h. After quenching each reaction by the addition of 10 μM SAHA, samples were added to a 96 well plate containing probe **5**, giving a final concentration of 40 μM **5**, 75 μM peptide, and 8.4 nM HDAC3/NCOR1.

Fluorescence measurements were made after 24 h and were used to generate a dose-response curve. An IC_{50} value of 1.5 µM was determined for the inhibition of HDAC3/NCOR1 by SAHA (Figure 5), which is in close agreement with a previously reported measurement for inhibition of the decrotonylase activity of HDAC3/NCOR1 by SAHA using a 7-amino-4-methylcoumarin tagged substrate.[140] Overall these results demonstrate that probe **5** is able to efficiently detect deacetylase activity via a sensitive fluorogenic response and allows for accurate determination of IC_{50} values, making it amenable to HDAC inhibitor screening.

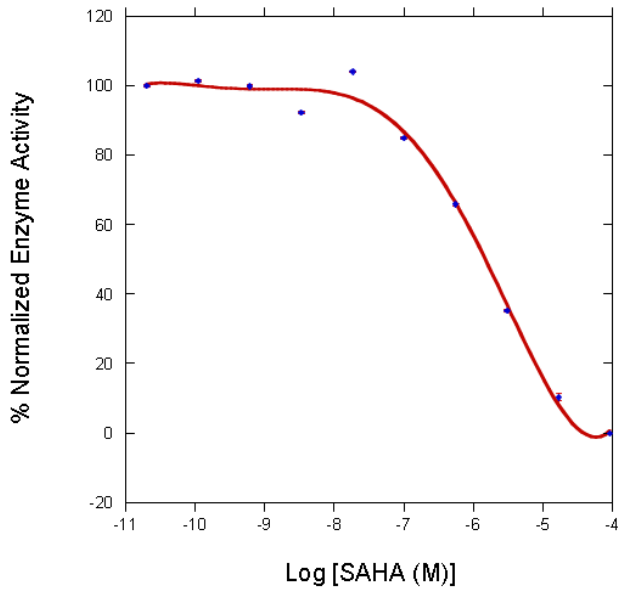


Figure 4.2. Dose-response curve of the inhibition of the HDAC3/NCOR1 catalyzed deacetylation of K4me3K9ac by SAHA. Data are an average of duplicate measurements. Conditons: 40 µM **5**, 75 µM peptide, serial dilution of SAHA (starting from 92 µM) in 50 mM bicine, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, pH 8.02, 24 h.

Conclusions

We also illustrated the utility of our labeling concept in the development of a turn-on fluorescence HDAC assay. Inhibition of HDAC1 and HDAC3/NCOR1 activity by a known inhibitor, SAHA, could be monitored without the need for chromatographic separation of reaction components or a secondary trypsin development step. While the assay was demonstrated

for monitoring deacetylase activity, this probe could theoretically be applied to other histone modifying enzymes such as acetyltransferases, demethylases, or methyltransferases. Moving forward with this labeling methodology, the largest challenge will be gaining the selectivity necessary for many of the possible applications. Ultimately, we believe that the design of synthetic receptors with improved binding properties and the incorporation of alternative reactive group chemistries will be great assets in helping synthetic receptor-directed approaches to realize these potential applications.

Experimental

In Vitro HDAC Assay

For HDAC1, enzymatic reactions were performed in 25 mM Bicine buffer, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$ (pH 8.02) with 0.1 μg/μL HDAC1 (purchased from Active Motif) and 200 μM acetylated peptide at 37 °C. After 1 h, the reactions were diluted with buffer before adding to a 96-well plate containing probe **5** to give final concentrations of 20 μM peptide, 40 μM probe **5**, and 10 ng/μL HDAC1 (reactions containing SAHA were diluted to contain 5 μM SAHA). Fluorescence labeling was monitored every hour for 24 h.

For in vitro inhibition experiments with HDAC3/NCOR1, a 5.5 dilution series (starting with 92 μM SAHA) was prepared in 50 mM Bicine buffer, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, 0.5 mg/mL BSA (pH 8.02). Enzymatic reactions were performed using 11.2 nM HDAC3/NCOR1 (purchased from Enzo Life Sciences) and 100 μM acetylated peptide at 37 °C. After 4 h, the reactions were quenched with 10 μM SAHA before adding to a 96-well plate containing probe **5** to give final concentrations of 77 μM peptide, 40 μM probe **5**, and 8.6 nM HDAC3/NCOR1. Fluorescence labeling was monitored after 24 h. The data were analyzed using KaleidaGraph version 3.6 to generate and fit a dose-response curve and to calculate IC_{50} values.

HPLC Analysis of HDAC1 and HDAC3/NCOR1 Reactions

Reactions were analyzed by analytical HPLC using a C18 column at a flow rate of 1 mL/min. Chromatographic separation was carried out using a 100-min linear gradient of solvent A (95% H_2O, 5% CH_3CN, and 0.1% TFA) and solvent B (95% CH_3CN, 5% H_2O, and 0.1% TFA). Absorbance was monitored at 280 nm.

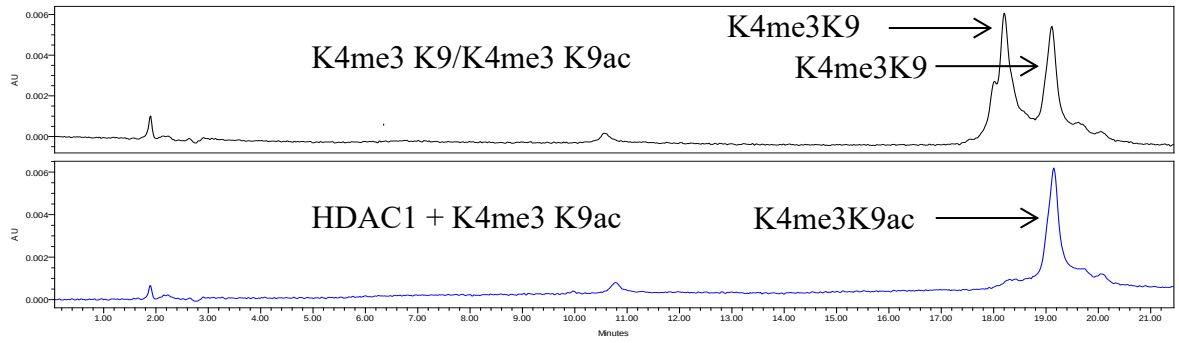


Figure 4.3. HPLC trace at 280 nm of a mixture of 20 µM K4me3 K9ac and 20 µM $K4me_3$ K9 (top) and the reaction of 0.007 µg/µL HDAC1 with 20 µM K4me3 K9ac after 30 min (bottom).

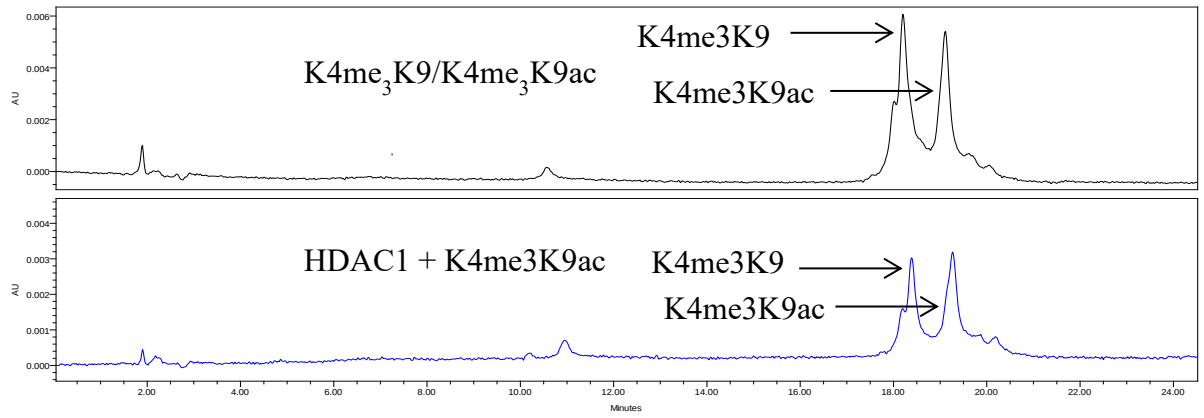


Figure 4.4. HPLC trace at 280 nm of a mixture of 20 µM K4me3 K9ac and 20 µM K4me$_3$ K9 (top) and the reaction of 0.1 µg/µL HDAC1 with 20 µM K4me3 K9ac after 30 min (bottom).

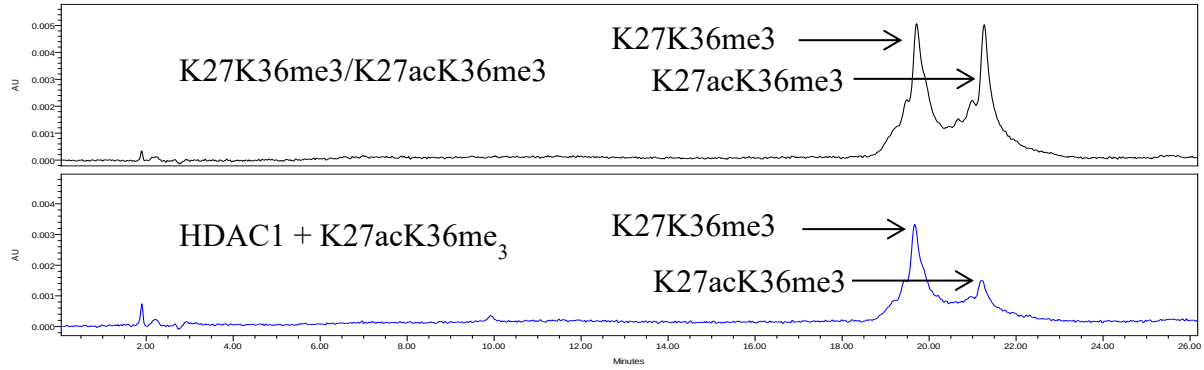


Figure 4.5. HPLC trace at 280 nm of a mixture of 20 µM K27ac K36me3 and 20 µM K27 K36me$_3$ (top) and the reaction of 0.1 µg/µL HDAC1 with 20 µM K27ac K36me3 after 30 min (bottom).

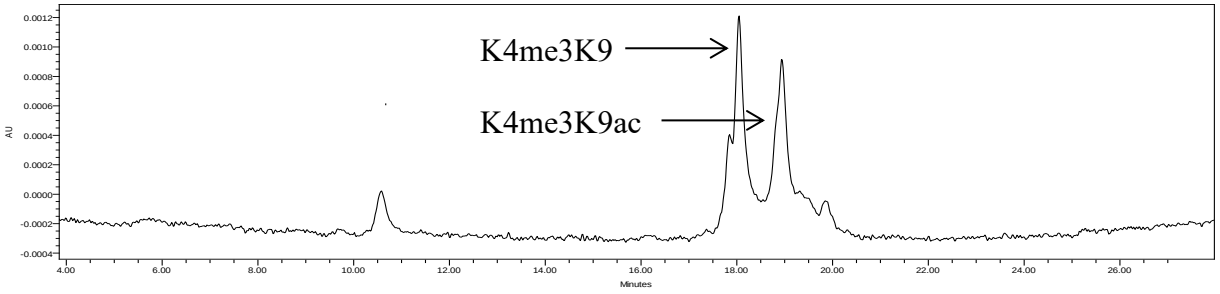


Figure 4.6. HPLC trace at 280 nm of the reaction of 0.1 µg/µL HDAC1 with 20 µM K4me3 K9ac after 1 h.

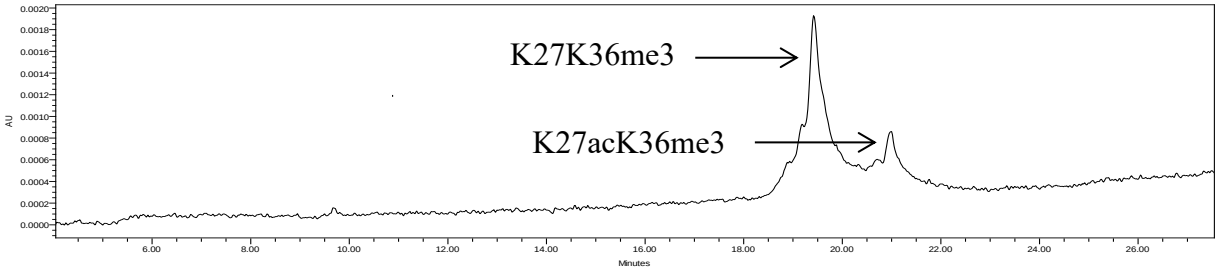


Figure 4.7. HPLC trace at 280 nm of the reaction of 0.1 µg/µL HDAC1 with 20 µM K27ac K36me3 after 1 h.

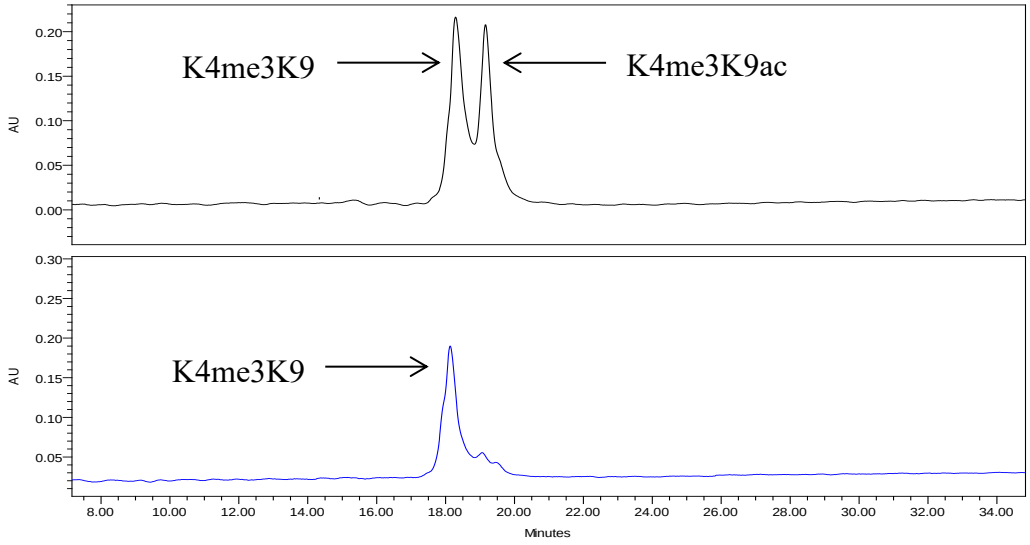


Figure 4.8. Top: HPLC trace of a 1:1 mixture of 100 µM K4me$_3$K9ac and 100 µM K4me$_3$K9. Bottom: HPLC trace at 214 of the reaction of 11.2 nM HDAC3/NCOR1 with 100 µM K4me$_3$K9ac after 4 h.

Turn-on Fluorescence HDAC1 Activity Assay

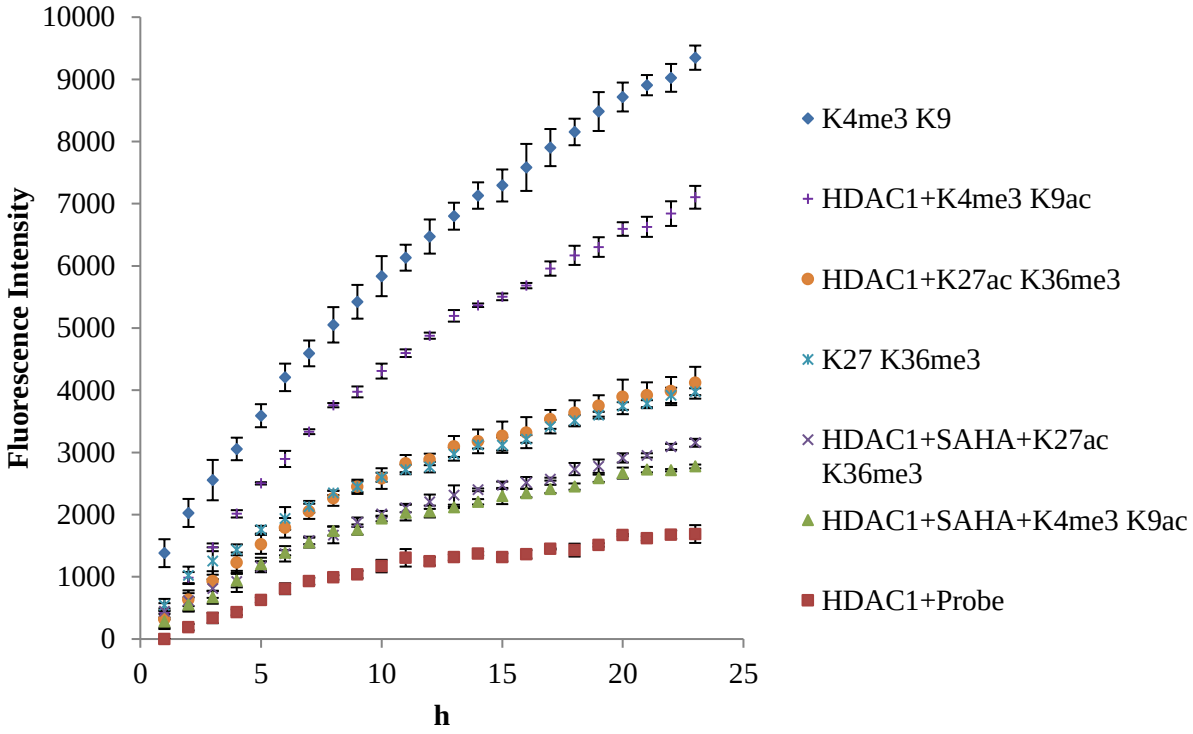


Figure 4.9. Turn-on fluorescence HDAC1 activity assay. Fluorescence monitoring of the deacetylation of histone H3 peptides K4me3 K9ac and K27ac K36me3 with probe **5** after reaction with HDAC1 in the absence or presence of 5 μM SAHA for 1 h at 37 ºC. Enzymatic reactions were performed in 25 mM Bicine buffer, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$ (pH 8.02) with 0.1 μg/μL HDAC1 and 200 μM acetylated peptide. After 1 h, the reactions were diluted with buffer before adding to a 96-well plate containing probe **5** to give final concentrations of 20 μM peptide, 40 μM probe **5**, and 10 ng/μL HDAC1 (reactions containing SAHA were diluted to contain 5 μM SAHA). Fluorescence labeling was monitored every hour for 24 h. Each data point represents an average of three runs. Error bars are standard deviations.

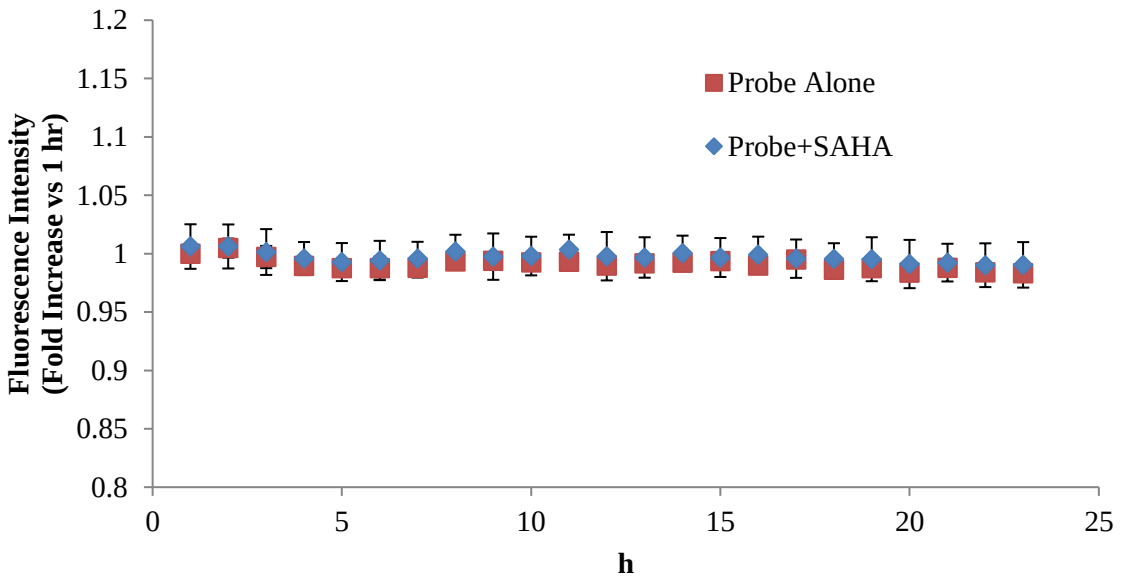


Figure 4.10. Turn-on fluorescence labeling SAHA control experiment. Labeling of 40 μM **5** alone (red) or in the presence of 5 μM SAHA (blue) in 25 mM Bicine buffer, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$ (pH 8.02). Each data point represents an average of three runs. Error bars are standard deviations.

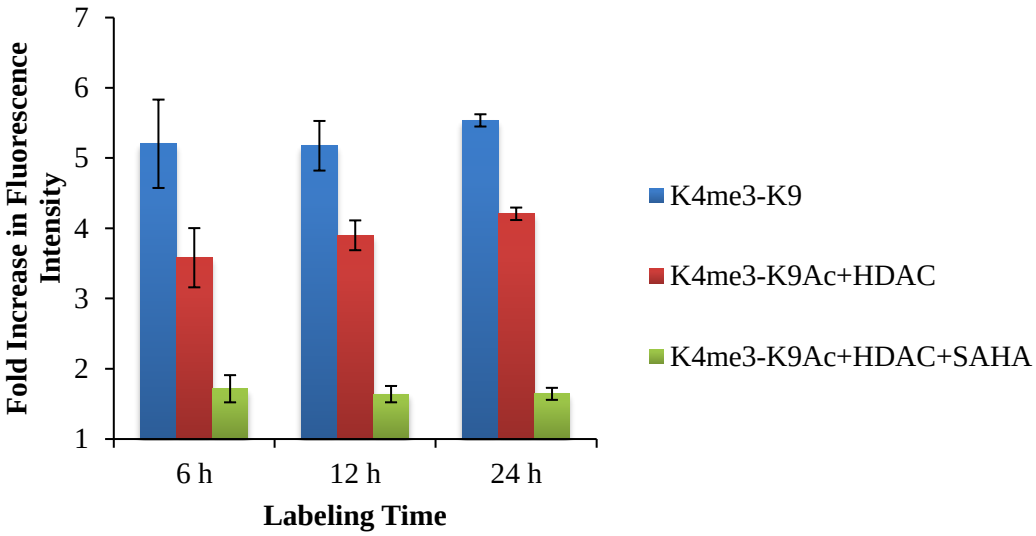


Figure 4.11. Turn-on fluorescence HDAC1 activity assay: fluorescence increase of deacetylated K4me3 K9ac after labeling for 6, 12, or 24 h. Fold increase in fluorescence intensity was normalized to the negative control containing HDAC1 and **5** alone after 6, 12, or 24 h. Each data point represents an average of three runs. Error bars are standard deviations.

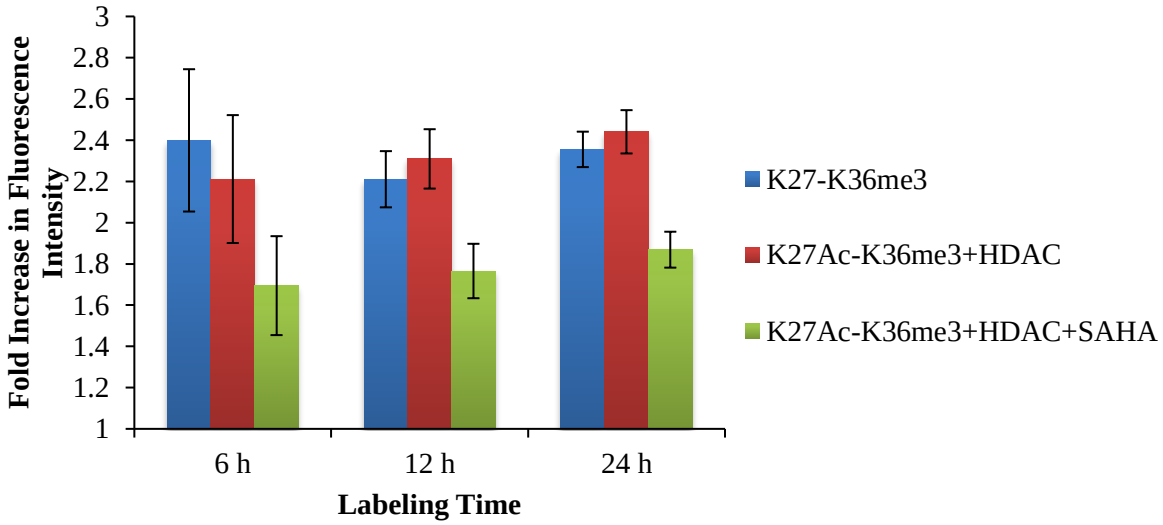


Figure 4.12. Turn-on fluorescence HDAC1 activity assay: fluorescence increase of deacetylated K27ac K36me3 after labeling for 6 ,12, or 24 h. Fold increase in fluorescence intensity was normalized to the negative control containing HDAC1 and **5** alone after 6, 12, or 24 h. Each data point represents an average of three runs. Error bars are standard deviations.

Turn-on Fluorescence HDAC3/NCOR1 Activity Assay

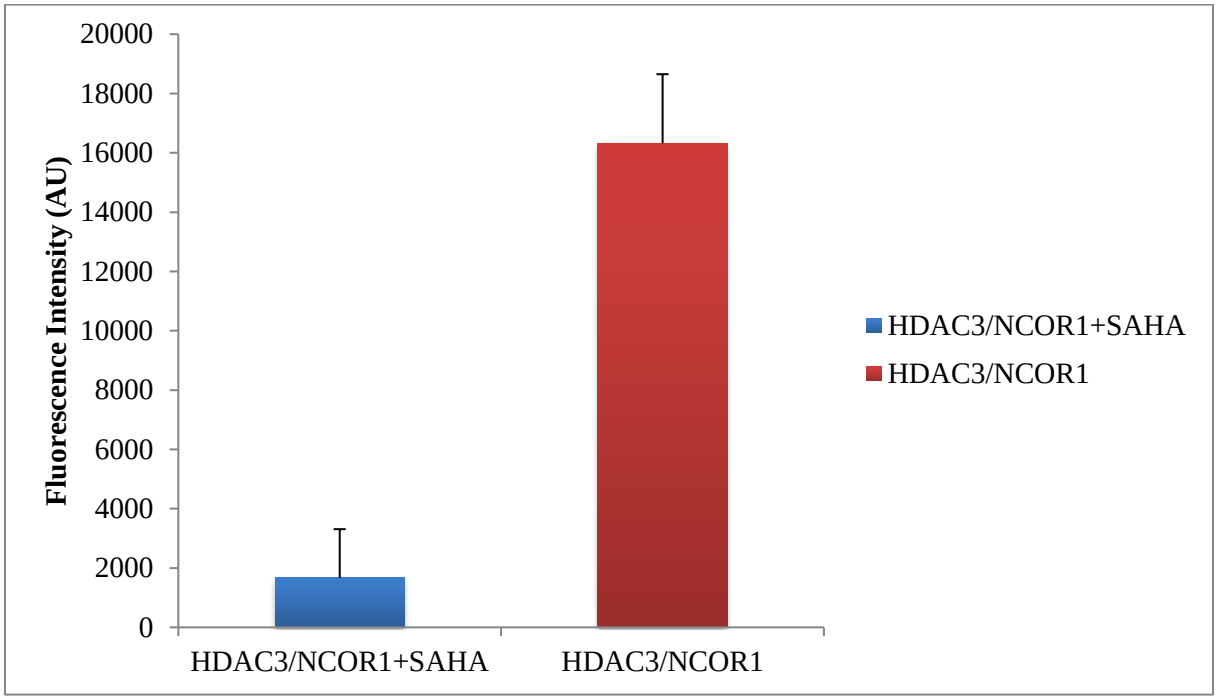


Figure 4.13. Turn-on fluorescence HDAC3/NCOR1 activity assay: fluorescence increase of deacetylated K4me3K9ac after labeling for 24 h. Enzymatic reactions were performed using 11.2 nM HDAC3/NCOR1, 100 μM acetylated peptide, and with or without 17 μM SAHA at 37 ^{o}C. After 4 h, the reactions were quenched with 10 μM SAHA before adding to a 96-well plate containing probe **5** to give final concentrations of 77 μM peptide, 40 μM probe **5**, and 8.6 nM HDAC3/NCOR1. Each data point represents an average of two runs. Error bars are standard deviations.

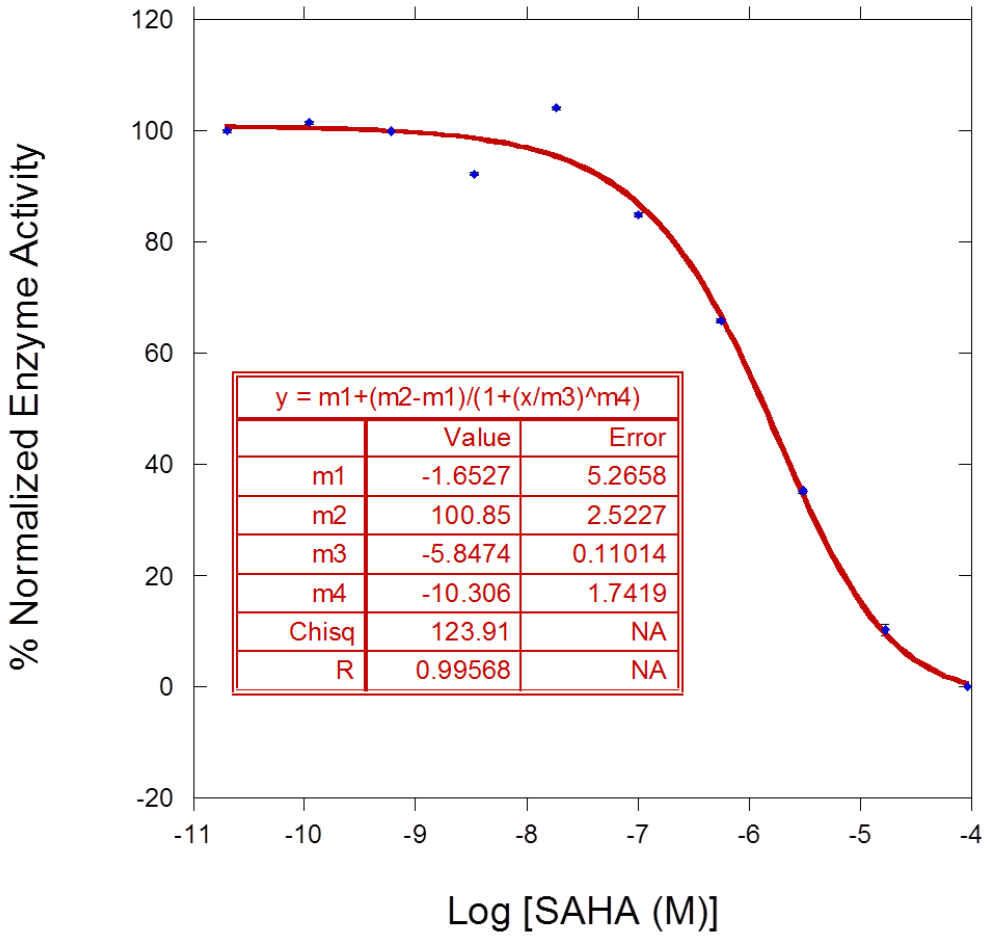


Figure 4.14. Plot of the dose response curve for the inhibition of HDAC3/NCOR1 activity by SAHA as monitored by the CX4-ONBD turn-on fluorescence HDAC assay. SAHA. Data are an average of duplicate measurements. Conditions: 40 µM 5, 75 µM peptide, serial dilution of SAHA (starting from 92 µM) in 50 mM bicine, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, pH 8.02, 24 h. Error bars represent the standard deviation of each data point. In the equation provided for the IC_{50} fit, m1 and m2 correspond to the minimum and maximum Y values, respectively. The value m3 is the IC_{50} value (the X value at a Y value of 50), and m4 is the slope at the midpoint of the curve.

Z Factor Determination

For determining the suitability of the fluorogenic assay for HTS screening, the Z factor[141] was calculated. Eight wells of positive control reaction and eight wells of negative control reaction were used in 96-well plate format. K4me3K9ac was used as a negative control, and K4me3K9 was used as a positive control. Reactions were performed with 40 μM probe 5 and 20 μM peptide in 10 mM sodium borate buffer, pH 8.6, 100 mM NaCl. After 24 h, the fluorescence intensity was measured.

$$Z' = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

Figure 4.15. Equation for calculating Z'. The values σ_p and σ_n are the standard deviation of the positive control and negative control respectively. The values μ_p and μ_n refer to the mean of the fluorescence intensity for the positive control and negative control.

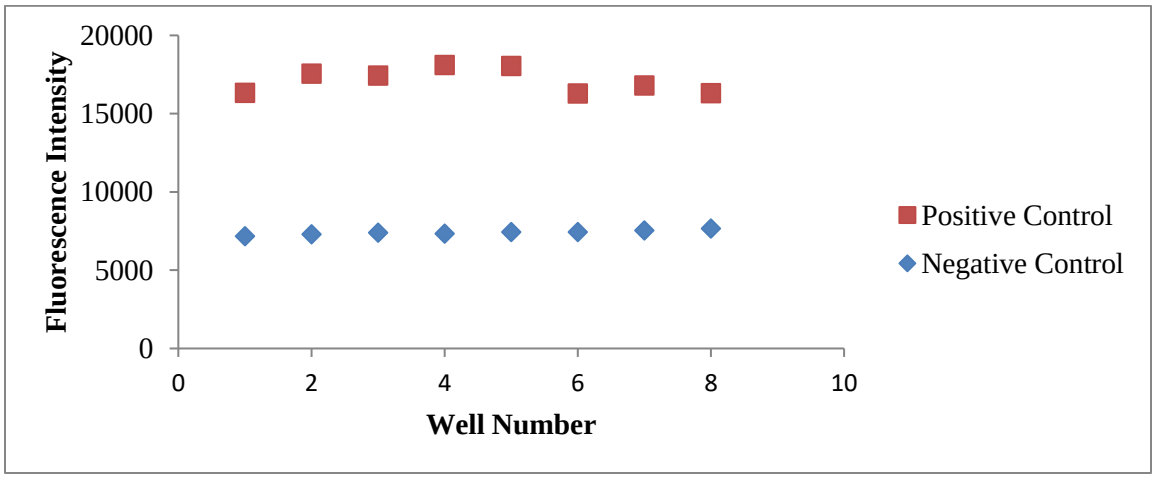


Figure 4.16. Fluorescence intensities of positive control reaction and negative control reactions for eight wells for Z' determination.

Table 4.1. Summary of Fluorescence Intensities and Z Factor Calculation

	1	2	3	4	5	6	7	8	Avg	Stdev	Z Factor
Positive Control	16337	17561	17439	18120	18052	16282	16802	16319	17114	776.68	0.71
Negative Control	7172	7299	7396	7345	7437	7446	7541	7673	7413.625	151.20	

CHAPTER 5. Design and Synthesis of Synthetic Receptors for Lower Methylation States of Lysine

Introduction and Significance

Over the past decade, supramolecular approaches for the recognition of lysine methylation (Figure 1), a protein post-translational modification that regulates gene expression, have expanded greatly. To date, at least three different classes of molecular receptor have been reported to bind to trimethyllysine (Kme3): cyclophanes (Waters and Hamilton)[96,98,99], calixarenes (Hof)[142], and cucurbiturils(Macartney)[95]. Additionally, several of these receptors have been applied to the field of chemical biology to develop sensors (Hof ,Waters, Nau, Hooley, and Zhong)[102–104,143–145] and affinity reagents (Hof and Kanai)[105,106]. Nonetheless, little progress has been made in developing synthetic receptors for the lower methylation states of lysine. This is likely due to the greater hydrophilic nature of mono- and dimethyl lysine, which are protonated at physiological pH and hydrogen bond with water, resulting in a much higher desolvation penalty than Kme3. In general, binding hydrophilic guests in aqueous environments is a significant challenge in the field of molecular recognition due to the necessity of overcoming interactions between the guest and water molecules.[146–151] In spite of this thermodynamic hurdle, proteins overcome this challenge by incorporating buried polar interactions to bind a variety of hydrophilic molecules including sugars[152], DNA[153,154], and polar amino acids.[155] The methyllysine "reader" proteins are another such example that utilize this strategy.[156,157] While the reader proteins for Kme3 contain an aromatic cage which contributes to binding via cation-π

interactions, binding to lower methylation states is accomplished by directing an additional polar functional group, typically a carboxylate from Asp or Glu, into the pocket, which forms a salt bridge to the NH group of Kme or Kme2.

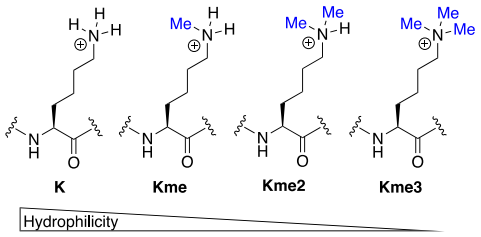


Figure 5.1. Methylation states of lysine.

Herein, we use dynamic combinatorial chemistry (DCC)[100,158] to develop a new receptor, **A_2I**, which directs a carboxylate into its binding pocket to form a salt bridge with bound Kme2, with the goal of mimicking reader proteins for Kme2 (Figure 2). This receptor displays a 0.2 μM affinity for Kme2 in the context of a histone peptide, which represents a 32-fold enhancement over **A_2B**. This is the highest affinity receptor for Kme2 to date and is also the first to bind Kme2 as well or better than Kme3. This work highlights the power of DCC for the discovery of receptors with new binding properties by simple redesign of monomers and represents a step forward in the development of synthetic receptors that preferentially bind lower methylation states of lysine.

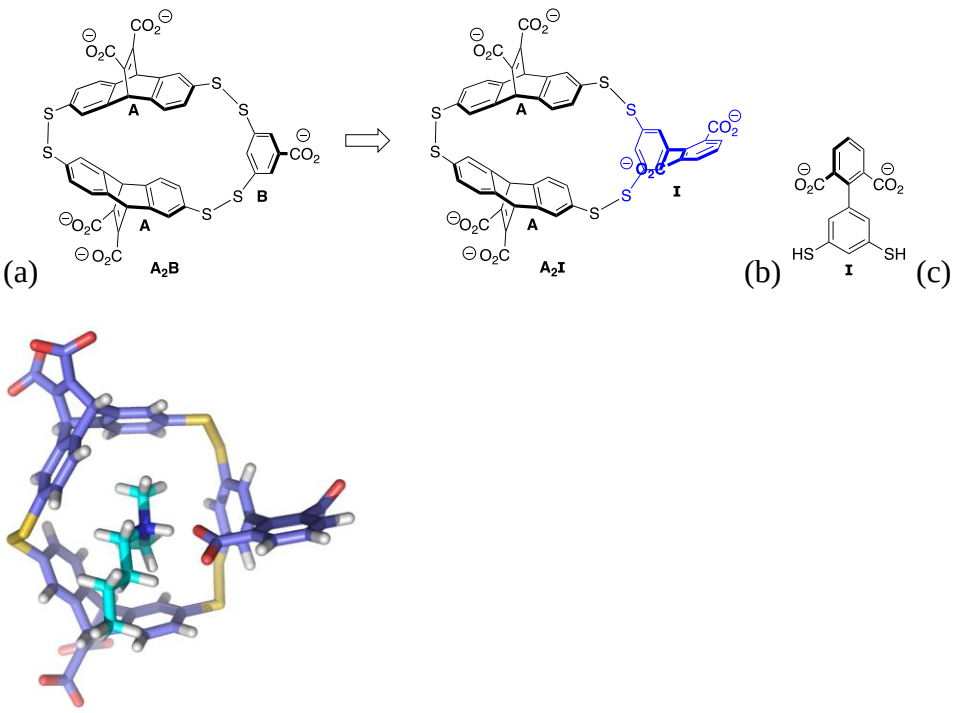


Figure 5.2. (a) Redesign of **A$_2$B** to **A$_2$I**. (b) Monomer **I**. (c) Gas-phase molecular model of one isomer of **A$_2$I** (blue) bound to butyldimethyl ammonium (cyan), a mimic of Kme2.

System Design

We previously reported[97–99] the development of several small molecule receptors for methylated lysine and arginine that mimic the binding motif of reader proteins. These receptors were prepared using iterative redesign by DCC. DCC[100] has proven to be a powerful tool for identifying small molecule receptors for methylated lysine, as it allows for both facile synthesis of macrocyclic receptors from monomer building blocks and for rapid analysis of potential binders. In the redesign process, an individual monomer from a known receptor is synthetically modified or varied to introduce new structural elements and functional groups that are expected to influence guest binding. DCC is then used to rapidly screen for receptors that incorporate the new monomer and bind tightly or selectively to a particular guest. To date, receptors for methylated lysine have consistently shown a stronger preference for binding Kme3 over lower

methylation states of lysine. This is consistent with the fact that the lower methylation states are less hydrophobic and can hydrogen bond with water, which generally makes binding inside a hydrophobic aromatic pocket less favorable compared to Kme3.

In designing new receptors for lower methylation states of lysine that mimic nature's reader proteins, we took clues from protein engineering work on methyl lysine reader proteins. Li et. al investigated lysine methylation state specificity by mutating the residues near the aromatic binding pocket of the BPTF PHD Finger, a binding motif which specifically binds H3K4me3 in the context of a histone peptide (Figure 3a).[159] By mutating a tyrosine to a glutamic acid, the selectivity in binding was reversed to favor binding Kme2 over Kme3 by a factor of two (Figure 3b). This change in selectivity was attributed to a salt-bridge interaction between the glutamic acid carboxylate and the N-H group of Kme2. By following an analogous redesign strategy, we looked to alter the methylation state specificity of a known receptor for Kme3 by introducing a favorably positioned carboxylate within the cavity to form a receptor with greater selectivity for Kme2.

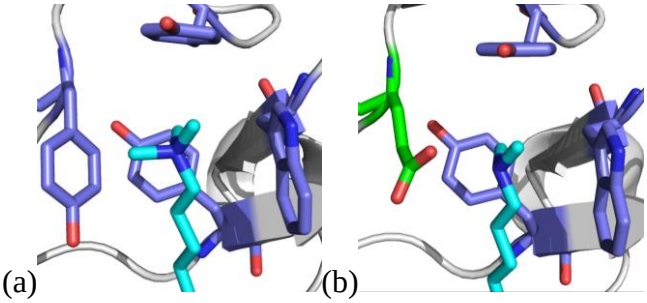


Figure 5.3. (a) BPTF PHD domain bound to Kme3. (b) BPTF PHD domain Y17E mutation bound to Kme2.

The synthetic receptor **A_2B** has been shown to selectively bind Kme3 by forming an aromatic cage that is able to engage the methyl groups of Kme3 in cation-π interactions.[96]

Additionally, **A_2B** was shown to bind Kme3 with two-fold tighter affinity compared to Kme2, presumably due to Kme2 having fewer methyl groups to engage in cation- π interactions and having a higher energetic cost of desolvation. We reasoned that selectivity for lower methylation states of lysine like dimethyllysine could be achieved by redesigning monomer **B** into a biphenyl monomer, **I**, with ortho-disubstituted carboxylic acids. A biphenyl monomer with two ortho carboxylates would force one of the carboxylic acids to be inside the cavity of the receptor. Theoretically, this would allow for a favorable salt bridge with Kme2, which is the combination of both electrostatic and hydrogen bonding interactions. Since Kme3 is not able to engage in a salt bridge due to its lack of a hydrogen bond donor, we expected that incorporation of monomer **I** into the receptor would reduce the preference for binding Kme3 over Kme2.

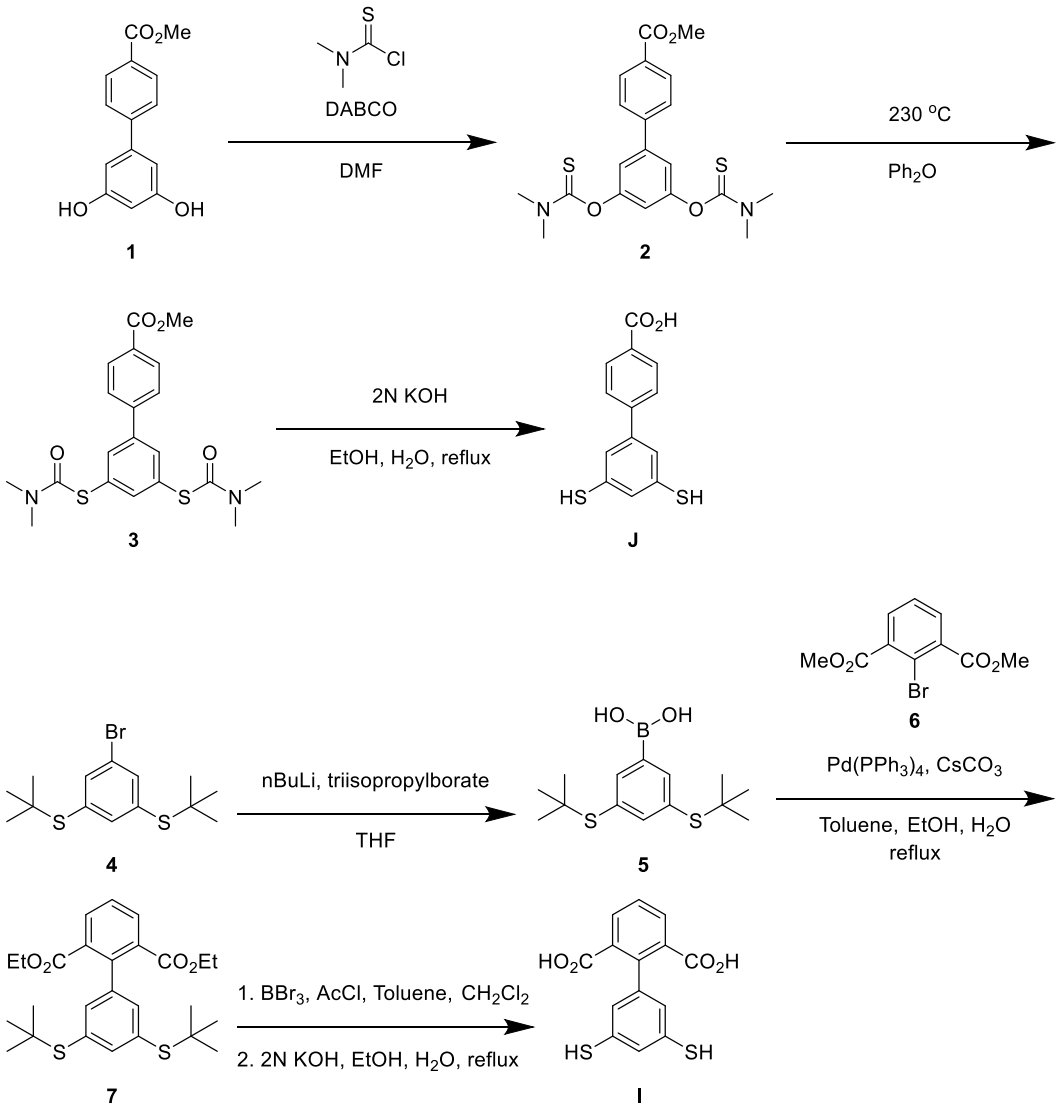
Synthesis of Monomers I and J

The control monomer **J** was synthesized following a similar synthesis reported for preparing monomer **B** (Scheme 1).[101] Compound **1** was prepared according to literature procedure.[160] The O-thiocarbamate intermediate **2** is prepared by reacting the hydroxyl groups of **1** with DMTC. A Newman-Kwart rearrangement provides the S-thiocarbamate **3**, and subsequent base hydrolysis affords the monomer **J**.

Initial attempts to prepare monomer **I** via an O-thiocarbamate intermediate were unsuccessful, possibly due to the steric hindrance of the ortho-disubstituted biphenyl. An alternative synthetic approach that proved to be effective involving the incorporation of tert-butyl protected thiols prior to cross coupling. First, the di-S-t-butyl aryl bromide **4** was prepared according to literature procedure via an S_NAr reaction between 1-bromo-3,5-difluorobenzene and tert-butyl thiol.[161] Compound **4** was then converted to a boronic acid (**5**) through lithiation followed by borylation with triisopropylborate. The boronic acid **5** was used in a Suzuki-Miyaura

reation with compound **6** (prepared according to previously reported procedure[162]) to yield biphenyl precursor **7**. The S-di-tbutyl groups of **7** were converted to thioacetate groups using boron tribromide and acetyl chloride. Finally, base hydrolysis of the ester and thioacetate groups produced the monomer **I** in good yield.

Scheme 5.1: Synthesis of Monomers **I** and **J**.



Dynamic Combinatorial Library Screening

Following similar conditions reported for the preparation of **A_2B**, five unbiased dynamic combinatorial libraries (DCLs) were prepared using 1 mM **A**, 1 mM biphenyl monomer, and 2 mM of guest (or no guest for the untemplated library) in 50 mM sodium borate buffer at pH 8.5. Methylated butylamine guests were used as convenient Kme_x side chain mimics: n-butylamine for K, N-butylmethylamine for Kme, N,N-dimethylbutylamine for Kme2, or butyltrimethylammonium bromide for Kme3. These Kme_x mimics were used as they are inexpensive and commercially available or easily prepared in the case of butyltrimethylammonium bromide.[163]

The DCLs were allowed to undergo disulfide exchange and equilibration for five days at room temperature in air before analysis by LCMS. For the libraries containing monomers **A** and **I**, two major species were amplified (Figure 4a) both of which corresponded to the mass of **A_2I,** corresponding to rac- and meso-stereoisomers, since monomer **A** was used as a racemic mixture. Significant amplification of the receptor is seen even in the presence of the K mimic (3 to 4 fold above the untemplated); however the higher methylation state guests still amplify the receptor to the greatest extent (nearly 7-fold greater than the untemplated library) (Figure 4b). Interestingly, the two stereoisomers are amplified to different degrees depending on the guest. In the presence of the Kme2 mimic, the peak that elutes second by HPLC is preferentially amplified, whereas the isomer that elutes first is amplified to a greater extent in the presence of the Kme3 mimic.

After characterizing amplification, receptors **A_2I** and **A_2J** were synthesized on a preparative scale using butyltrimethylammonium bromide as a guest. Each preparative library contained 2.5 mM monomer **A**, 2.5 mM of the respective biphenyl monomer, and 10 mM guest

dissolved in 10 mM sodium borate buffer at pH 8.5. Prep scale DCLs were allowed to equilibrate for a week before purification by semi-preparative reversed-phase HPLC.

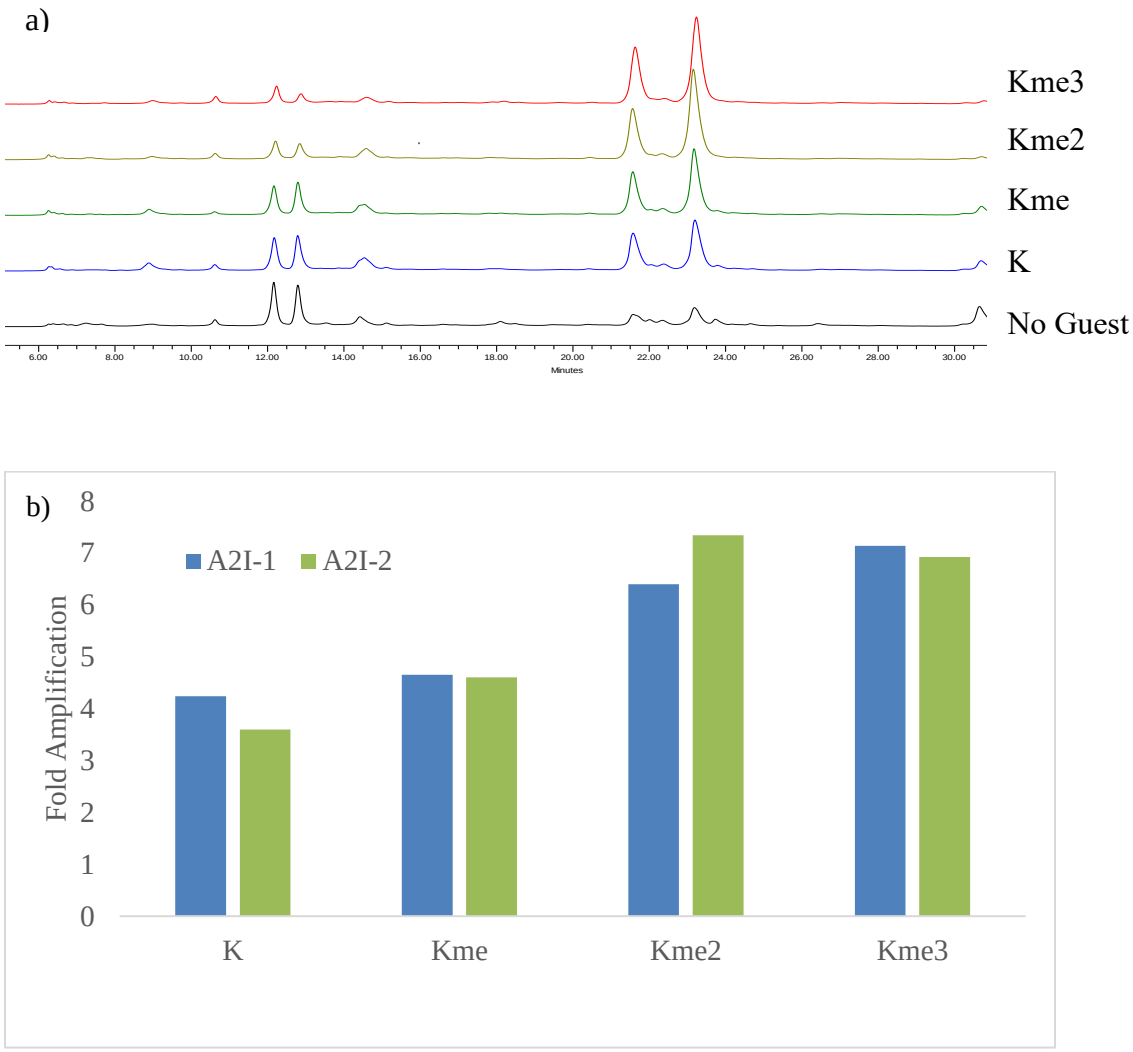


Figure 5.4. (a) DCLs of monomers **A** and **I** with different methylated butylamine guests (1 mM **A**, 1 mM **I**, 2 mM guest, 50 mM sodium borate buffer, pH 8.5). (b) Amplification of **A_2I -1** and **A_2I-2** in the presence of Kmex mimics. Fold amplification is defined as the peak area of the guest-templated library divided by the peak area of the untemplated library determined from the HPLC trace at 280 nm.

NMR Characterization of Binding to Kme2 and Kme3

In order to investigate the mode of binding of **A_2I-2** to methylated lysine, we observed the ^{1}H NMR spectrum of the receptor in the presence of Ac-Kme2GGY-NH_2 or Ac-Kme3GGY-NH_2 under concentrations where all the peptide is bound. Protons on functional groups that are bound within the pocket of the receptor would be expected to experience shielding from being in close proximity to aromatic rings. At room temperature, the chemical shifts of several of the protons from the lysine side chain of the peptide are perturbed, and the NMe_2 or NMe_3 protons in particular experience significant broadening. When the ^{1}H NMR spectrum is measured at 40 ^{o}C, the NMe_2/NMe_3 protons become sharper and are shifted upfield by over 1 ppm. This indicates that **A_2I-2** binds to Kme2 or Kme3 by making cation-π interactions with the quaternary ammonium group.

Table 5.1. Change in chemical shift (Δδ) upon binding of Ac-Kme2-Gly-Gly-Tyr-NH_2 or Ac-Kme3-Gly-Gly-Tyr-NH_2 to **A_2I-2**

Peptide protons	A_2I-2 Δδ (ppm)[a]	A_2B[b,c]
NMe_2	-1.15	-
NMe_3	-0.93	-1.59

[a] Conditions: 50 mM sodium borate-d3 in D_2O, pH 8.5, 0.05 mM DSS at 40 ^{o}C. [b] ^{1}H NMR carried out at room temperature. [c] Data reported by Pinkin and Waters.

ITC Binding Studies

Isothermal titration calorimetry (ITC) was used to characterize the binding of **A_2I** and **A_2J** to a short model peptide consisting of methylated Lys (Kme_x) appended with a C-terminal GGY tag for measuring peptide concentration by UV-Vis and/or a representative histone H3 peptide containing different methylation states of lysine with an N-terminal WGGG tag, for comparison to previous host-guest studies[98,99] (see Table 1). ITC experiments were conducted at

25 °C in 10 mM borate buffer at pH 8.5. The thermodynamic binding data collected from ITC binding studies of **A_2I** and **A_2J** are summarized in Table 2.

Table 5.2. Peptides used for ITC binding studies

Peptide	Sequence
Kme_xGGY (x = 2,3)	Ac-Kme_xGGY-NH_2
H3 K9me_x (x = 0-3)	Ac-WGGG-QTARKme_xSTG-NH_2

Influence of the Internal Carboxylate on Selectivity

Preliminary binding studies were performed using short model peptide KmexGGY (x = 2 or 3) to determine the influence of the internal carboxylate on selectivity. The second isomer of **A_2I** (**A_2I-2**) was used as it exhibited slightly greater amplification in the presence of Kme2 than Kme3. Both the Kme2 and Kme3 peptides exhibited low micromolar affinity with a 1.3:1 preference for Kme2 vs Kme3, consistent with the amplification data. Thus, incorporation of an internal carboxylate did indeed result in a small reversal of selectivity, presumably due to a favorable salt bridge with Kme2 that that compensates for the loss of one cation-π interaction relative to Kme3.

Comparison of A_2I to A_2B with a biologically relevant peptide

Next, we investigated binding of **A_2I-1** and **A_2I-2** to a more biologically relevant peptide: the histone 3 tail containing Kme_x (x = 0, 1, 2, or 3) at position 9. We have used this same peptide to evaluate **A_2B**, such that a direct comparison can be made.[98] Incorporation of monomer **I** had a pronounced influence on binding affinity to methylated lysine relative to **A_2B**. Overall when compared to **A_2B**, both isomers of receptor **A_2I** bind to Kme2 with 32-fold higher binding affinity and represents the highest affinity small molecule receptor for Kme2 to date.

Additionally, receptor stereochemistry affected the affinity and selectivity for the different methylation states. ITC revealed that **A_2I-1** had 20-fold higher affinity for Kme3 (0.13 μM ± 0.005 μM) than the parent **A_2B** (2.6 μM ± 0.1) receptor and is roughly 2-fold more selective for binding Kme3 over K relative to **A_2B** (16-fold versus 8-fold respectively) (Table 2). **A_2I-1** maintains slight selectivity for Kme3 over Kme2 in the context of the H3 peptide, with a selectivity of 1.4 compared to 2.4 for **A_2B**.

Interestingly, receptor **A_2I-2** displays essentially identical binding to Kme2 and Kme3 in the context of the H3 peptide sequence while maintaining the same selectivity for binding Kme2/3 over Kme and K as **A_2B**. The tighter overall binding and nearly equal binding to Kme3 and Kme2 with this histone peptide relative to the KmexGGY peptide suggests that additional interactions are contributing to binding. Earlier studies[164] from the Waters lab have demonstrated that the Arg in position 8 that neighbors Lys 9 in the histone H3 peptide sequence influences the binding affinity and selectivity of **A_2B** and similar synthetic receptors for the different methylation states of lysine. While the presence of the Arg slightly diminishes the Kme2/Kme3 selectivity by about 0.2 kcal/mol, it has a much more substantial influence on overall binding affinity, amounting to about 1.6-1.8 kcal/mol, presumably due to additional nonspecific electrostatic interactions, giving rise to the sub-micromolar binding.

Table 5.3. Thermodynamic binding data for the biding of **A_2I** and **A_2J** to peptides shown in **Table 5.1** as measured by ITC.

Entry	*Receptor*	*Peptide*	*Charge*	*K_d [b] (μM)*	*Selectivity Factor[d]*	*ΔG^c*
1	A_2B	H3 K9me3	+2	2.6 ± 0.1	-	-7.63 ± 0.03
2	A_2B	H3 K9me2	+2	6.3 ± 0.3	2.4	-7.10 ± 0.07
3	A_2B	H3 K9me	+2	13.9 ± 0.1	5.4	-6.64 ± 0.01
4	A_2B	H3 K9	+2	22 ± 1	8.3	-6.38 ± 0.02
5	A_2I-2	H3 K9me3	+2	0.22 ± 0.03 [c]	-	-9.11 ± 0.09
6	A_2I-2	H3 K9me2	+2	0.20 ± 0.01	0.9	-9.11 ± 0.09
7	A_2I-2	H3 K9me	+2	1.00 ± 0.04	4.5	-8.18 ± 0.04
8	A_2I-2	H3 K9	+2	1.79 ± 0.1 [c]	8	-7.88 ± 0.1
9	A_2I-1	H3 K9me3	+2	0.13 ± 0.005	-	-9.39 ± 0.04
10	A_2I-1	H3 K9me2	+2	0.18 ± 0.01	1.4	-9.19 ± 0.2
11	A_2I-1	H3 K9me	+2	1.23 ± 0.04 [c]	9.5	-8.03 ± 0.06
12	A_2I-1	H3 K9	+2	2.09 ± 0.15 [c]	16	-7.78 ± 0.2
13	A_2I-2	Kme3GGY	+1	4.30 ± 0.33	-	-7.32 ± 0.02
14	A_2I-2	Kme2GGY	+1	3.32 ± 0.20	0.77	-7.48 ± 0.01
15	A_2J-1	H3 K9me3	+2	2.61 ± 0.23	-	-7.62 ± 0.10
16	A_2J-1	H3 K9me2	+2	5.21 ± 0.44	2.0	-7.21 ± 0.10
17	A_2J-1	H3 K9me	+2	6.36 ± 0.64	2.4	-7.09 ± 0.12
18	A_2J-1	H3 K9	+2	7.02 ± 0.75	2.7	-7.03 ± 0.13
19	A_2J-2	H3 K9me3	+2	2.25 ± 0.18	-	-7.70 ± 0.09
20	A_2J-2	H3 K9me2	+2	6.74 ± 0.62	3.0	-7.05 ± 0.11
21	A_2J-2	H3 K9me	+2	8.37 ± 1.0	3.7	-6.93 ± 0.14
22	A_2J-2	H3 K9	+2	10.60 ± 1.1	4.7	-6.79 ± 0.12

[a] Conditions: 25° C in 10 mM sodium borate buffer, pH 8.5. [b] Errors are averages of the error in the fit as determined by the Origin 7 software. [c] Errors are from averages of the runs. [d] The selectivity factor is calculated as the K_d of Kmex (where x = 0-2) divided by the K_d of Kme3.

Influence of Biphenyl Functional Group

Monomer **J** was used to assess whether or not the differences in affinity and selectivity exhibited by **A_2I** relative to **A_2B** were the result of carboxylate position or the effect of incorporating a phenyl group into the structure of monomer **B**. Measurement of binding by ITC showed that either isomer of **A_2J** shows about a 2-fold to 3-fold preferential binding to Kme3 over Kme2. Furthermore, the binding affinity of the receptor for Kme3 and Kme2 is nearly identical to the parent **A_2B** receptor, indicating that incorporation of the biphenyl scaffold alone does not contribute favorably to binding for the higher methylation states of lysine.

Conclusions

In summary, we have successfully employed DCC to redesign the receptor **A_2B**, which has a 2.4-fold preference for Kme3 over Kme2, into **A_2I**, which places a carboxylate directed into the pocket and provides the first receptor with a preference for Kme2. This work illustrates the value of rationally designing monomers for DCC with unique architectures (i.e. ortho-disubstituted biphenyls) for controlling interactions within a binding pocket. It also provides insight into effective strategies for the binding and selective molecular recognition of hydrophilic guests in water. Placing anionic functionality near or inside a hydrophobic aromatic binding pocket likely minimizes competing interactions with water, lending selectivity for guests capable of forming hydrogen bonding and cation-π interactions inside an aromatic cavity. The result of this work was a synthetic receptor, **A_2I**, with 32-fold greater affinity than **A_2B** for Kme2 in the context of a H3 peptide. When controlling for the influence of neighboring positive charge, **A_2I - 2** displays a marginal preference for binding Kme2 over Kme3. The high binding affinities and the changes in selectivities can be attributed to the formation of a salt-bridge interaction with Kme2 that arises from directing a carboxylic acid into the aromatic binding pocket. This

represents the highest affinity receptor for Kme2 in the context of a peptide that has been reported and is a promising step toward a small receptor with appreciable selectivity in binding Kme2 over other methylation states of lysine. Looking forward, strategies to more rigidly direct a carboxylate into the binding pocket should significantly improve biding to Kme2 over Kme3. Additionally, the use of receptors for methylated lysine with small binding pockets should sterically occlude Kme3 leading to greater selectivities for Kme2.

Experimental

Monomer **A** was prepared according to a previously reported procedure.[101]

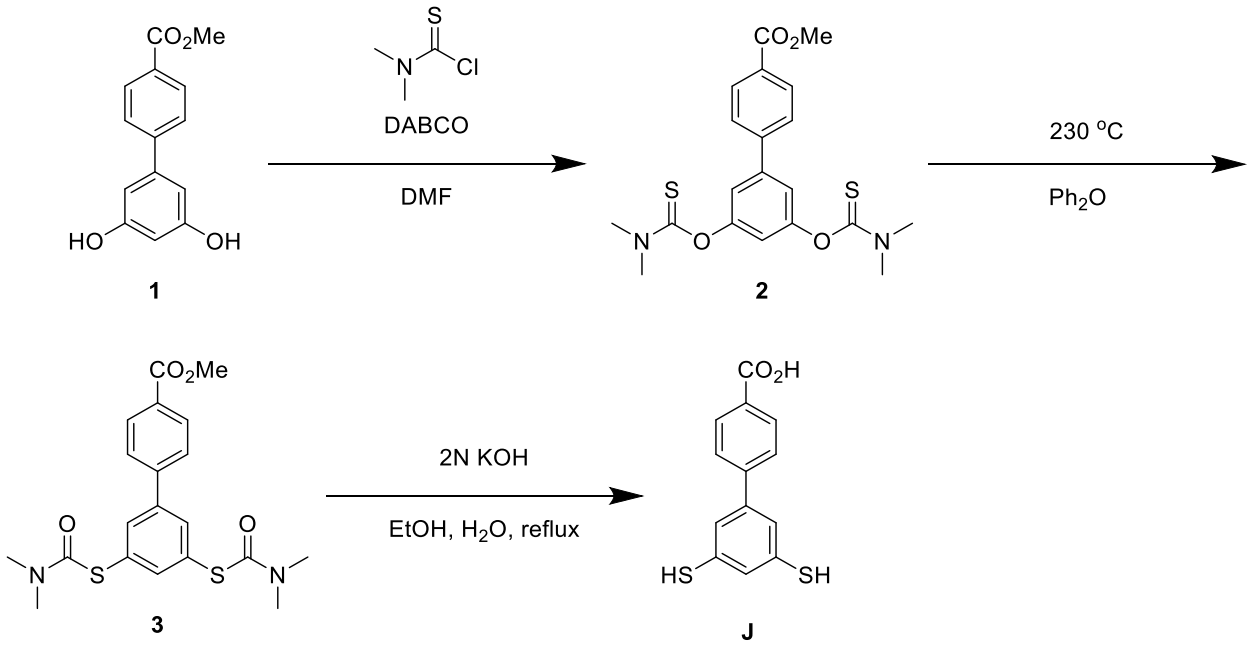
DCLs were made from stock solutions of the monomers dissolved in 50 mM sodium borate at pH 8.5. Components of the library were mixed to give final monomer concentrations of 1 mM at a total volume of 500 μL. The concentration of methylated butylamine guest was equivalent to the total of concentration of monomer in the library. DCLs were allowed to undergo oxidation and disulfide exchange equilibration for a week. After equilibration, the library was filtered using a 0.22 μm PVDF syringe filter and was analyzed by analytical reversed phase HPCL (Atlantis T3 4.7 x 150 mm 5 μm C-18 column) using by monitoring UV absorbance at 280. An optimized gradient of A and B (A: 100% H_2O, 10 mM NH_4OAc; B: 90% CH_3CN, 10% H_2O, 10 mM NH_4OAc) was used for DCL analysis.

Analysis of DCLs by analytical HPLC was carried out using an Agilent Rapid Resolution LC-MSD instrument situated with an online degasser, binary pump, autosampler, heated column compartment, and diode array detector. Separations were achieved using a mobile phase of H_2O (5 mM NH_4OAc) and CH3CN (95% CH_3CN, 5% H_2O, 5 mM NH_4OAc) at pH 5.5 on a Zorbax Extend C18 column (4.6 Å 2.1 x 50 mm, 1.8 micron). Negative ion electrospray ionization MS

analysis was performed with a single quad mass spectrometer. Data were analyzed using the Agilent ChemStation software.

ITC binding experiments were conducted using a Microcal AutoITC200. Titrations were performed at 25 °C in 10 mM sodium borate buffer at pH 8.5. The concentration of receptor was determined by measuring the UV–vis absorbance at 300 nm or 315 nm, using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. Solutions of 0.2–1.3 mM of peptide were titrated into a 20–91 μM solution of receptor, using 2.0 μL increments every 3 min. Heats of dilution were taken as the last three or four points and subtracted prior to fitting. Binding curves were produced using the supplied Origin software and fit using one-site binding models.

Scheme 5.2. Synthesis of Monomer J



Synthesis of 1. Compound **1** was prepared according to published procedure.[160]

Synthesis of 2. Compound **1** (0.23 g, 0.94 mmol, 1 equiv.) was dissolved in 2 mL of anhydrous DMF under a nitrogen atmosphere, and the solution was cooled to 0°C. To this solution was

added 1,4-diazabicyclo[2.2.2]octane (0.42 g, 3.7 mmol, 4 equiv.) in small portions. A solution of dimethylthiocarbamoyl chloride (0.47 g, 3.7 mmol, 4 equiv.) in 2 mL anhydrous DMF was added dropwise to the resulting suspension. The reaction mixture was allowed to warm to room temperature and stirred for 24 h. Afterward, 40 mL of water was added to the reaction mixture to precipitate out the crude solid, which was collected by filtration to yield a white solid (0.37 g, 95%). ^{1}H NMR ($CDCl_3$, 600 MHz): 8.087,8.073 (d, 2H, CH), 7.664,7.650 (d, 2H, CH), 7.249,7.245 (d, 2H, CH), 6.893 (t, 1H, CH), 3.935 (s, 3H, CH_3), 3.466 (s, 6H, N-CH_3), 3.361 (s, 6H, N-CH_3). ^{13}C NMR ($CDCl_3$, 600 MHz): 187.336, 167.000, 154.586, 143.833, 141.448, 130.241, 127.288, 119.516, 117.731, 52.324, 43.498, 39.025. MS (calculated): 419.11 [M+H], 441.09 [M+Na^+], 837.23 [2M+H], 859.21 [2M+Na^+]. HRMS (observed, ESI+): 419.10986 [M+H], 441.09189 [M+Na^+], 837.21382 [2M+H], 859.19383 [2M+Na^+].

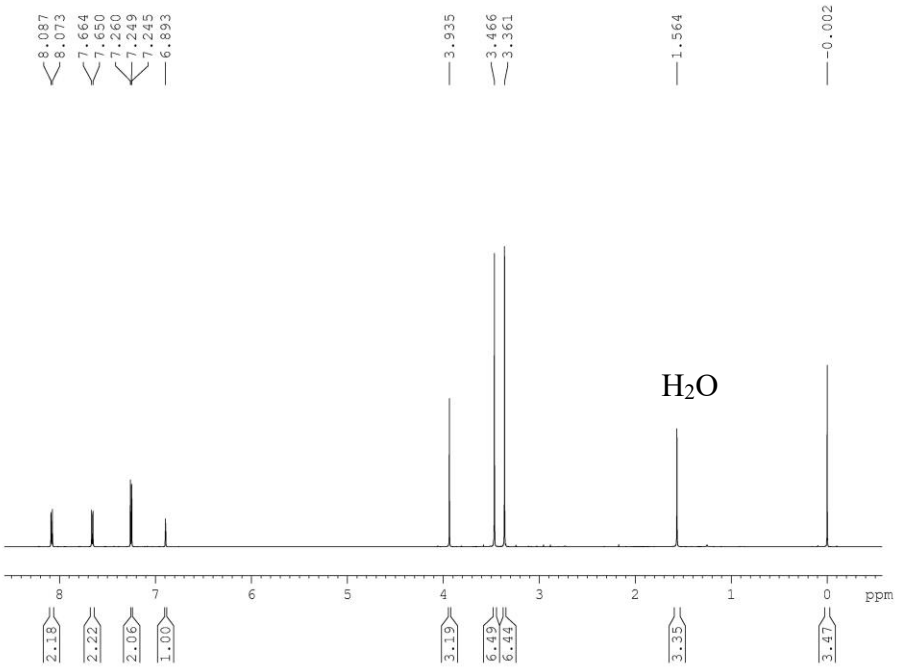


Figure 5.5. 1D ^{1}H-NMR (600 MHz) of Compound **2** in $CDCl_3$.

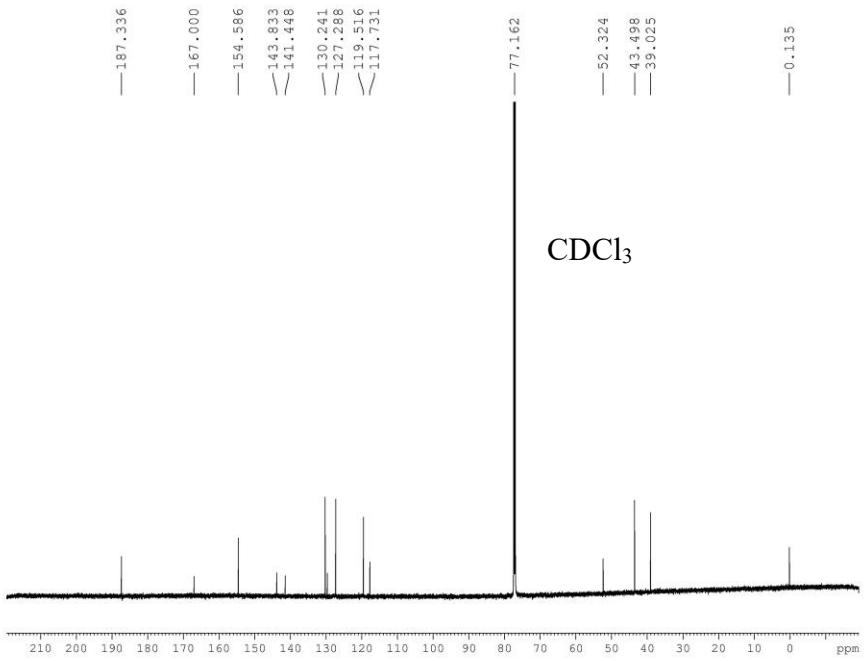


Figure 5.6. 1D ^{13}C-NMR (600 MHz) of Compound **2** in $CDCl_3$.

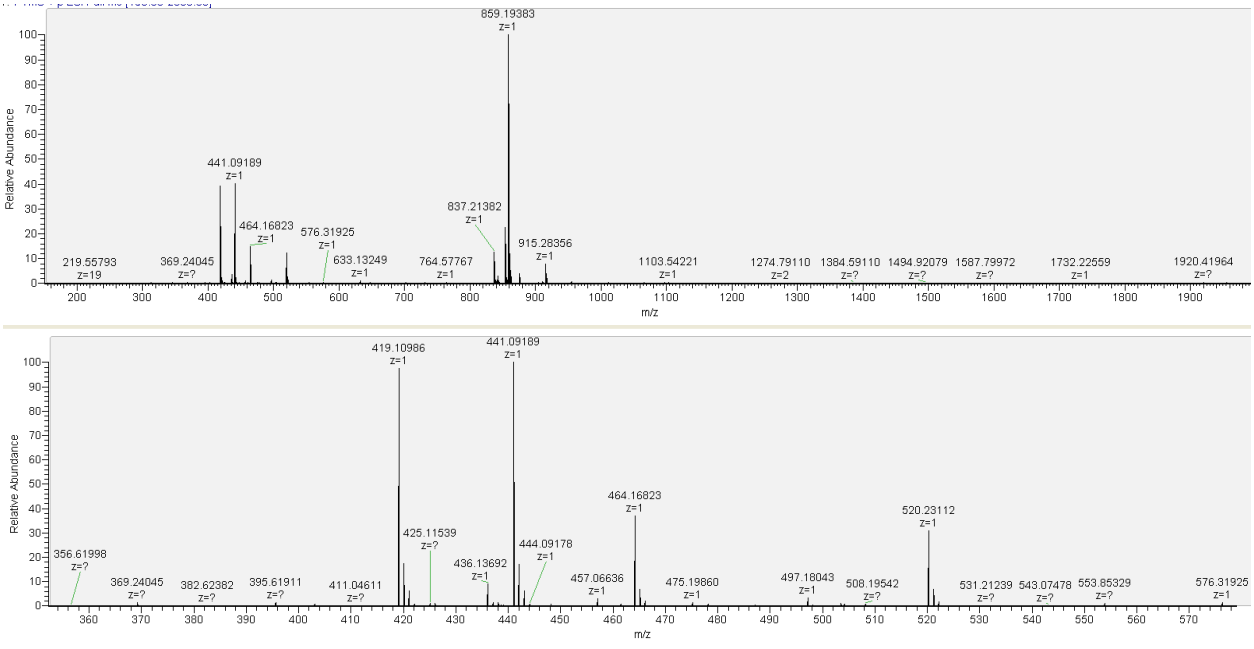


Figure 5.7. High resolution mass spectrum of Compound **2**.

Synthesis of 3. Compound **2** (0.30 g, 0.72 mmol, 1 equiv.) was dissolved in 5 mL of warm diphenyl ether and heated to 230ºC for 3 h. The reaction mixture was cooled to room temperature and diluted with 15 mL hexanes. The solution was purified by column chromatography (ethyl acetate/hexanes gradient from 0 to 50%) to obtain a white solid (0.25 g, 83%). ^{1}H NMR ($CDCl_3$, 600 MHz): 8.084, 8.070 (d, 2H, CH), 7.768, 7.765 (d, 2H, CH), 7.663, 7.660 (d, 2H, CH), 7.650 (s, 1H, CH), 3.934 (s, 3H, CH_3), 3.013 (s, 6H, N-CH_3), 3.035 (s, 6H, N-CH_3). ^{13}C NMR ($CDCl_3$, 600 MHz): 167.056, 166.371, 144.104, 141.550, 141.390, 135.328, 130.276, 130.210, 129.459, 127.460, 52.294, 37.113. MS (calculated): 419.11 [M+H], 441.09 [M+Na^+], 837.23 [2M+H], 859.21 [2M+Na^+]. HRMS (observed, ESI+): 419.11038 [M+H], 441.09259 [M+Na^+], 837.21406 [2M+H], 859.19353 [2M+Na^+].

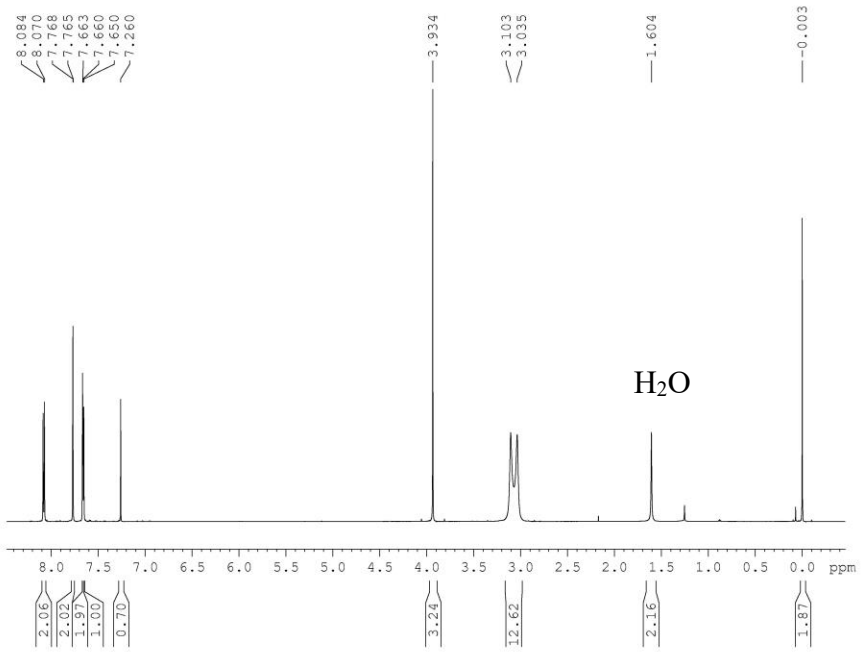


Figure 5.8. 1D ^{1}H-NMR (600 MHz) of Compound **3** in $CDCl_3$.

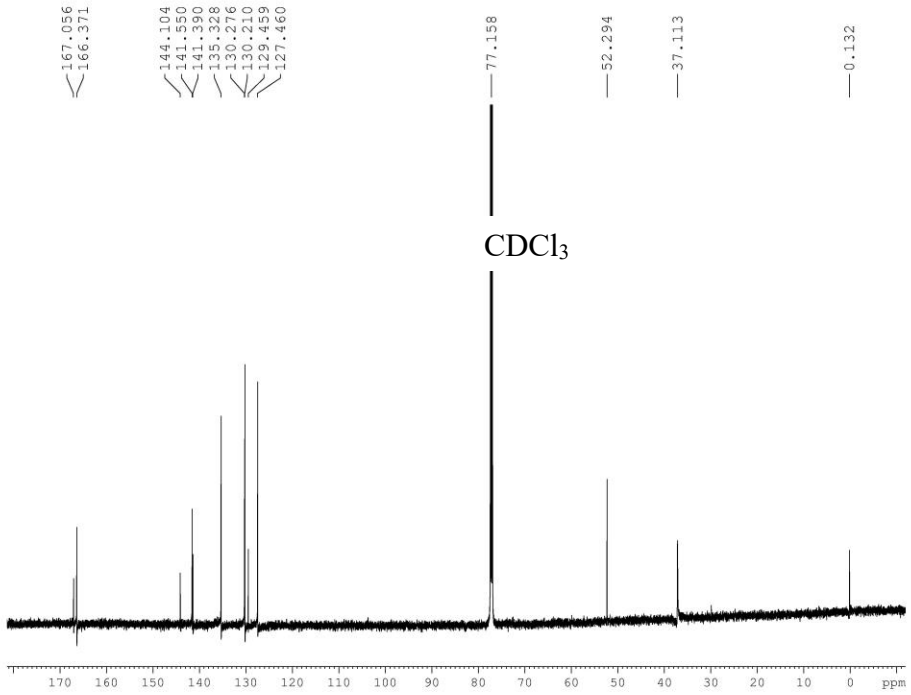


Figure 5.9. 1D ^{13}C-NMR (600 MHz) of Compound **3** in $CDCl_3$.

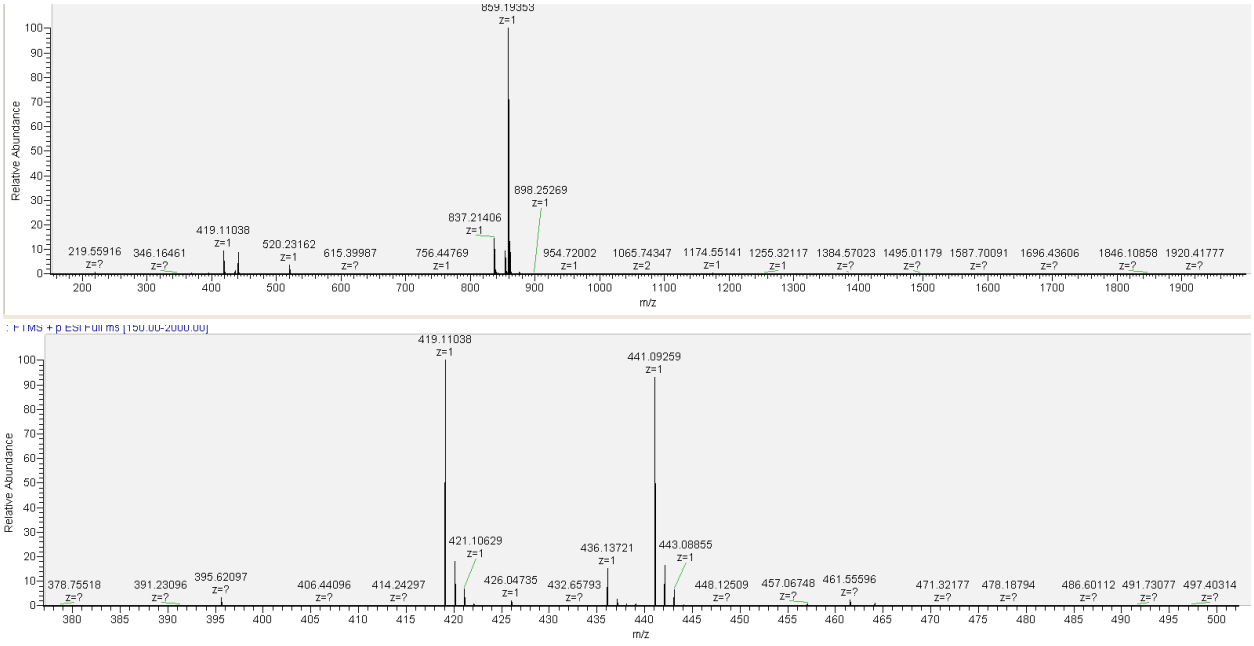


Figure 5.10. High resolution mass spectrum of Compound **3**.

Synthesis of J. Potassium hydroxide (0.80 g, 14 mmol, 33 equiv.) was dissolved in a solution of ethanol (2 mL) and water (4.5 mL), and the solution was degassed for 2 h. To this solution was added compound **3** (0.18 g, 0.43 mmol, 1 equiv.). The reaction was refluxed for 3 h and then cooled to room temperature. Ethanol was removed under reduced pressure, and 10% HCl (10 mL) was added. The precipitate was collected by vacuum filtration and dried under vacuum to give a white solid (0.102 g, 92%). ^{1}H NMR (CD_3OD 600 MHz): 8.093, 8.079 (d, 2H, CH), 7.684, 7.670 (d, 2H, CH), 7.343, 7.340 (d, 2H, CH), 7.234 (s, 1H, CH). ^{13}C NMR ($CDCl_3$, 600 MHz): 169.560, 145.557, 142.835, 135.383, 131.365, 128.616, 128.040, 125.076. MS (calculated): 261.00 [M-H], HRMS (observed, ESI-): 261.00495.

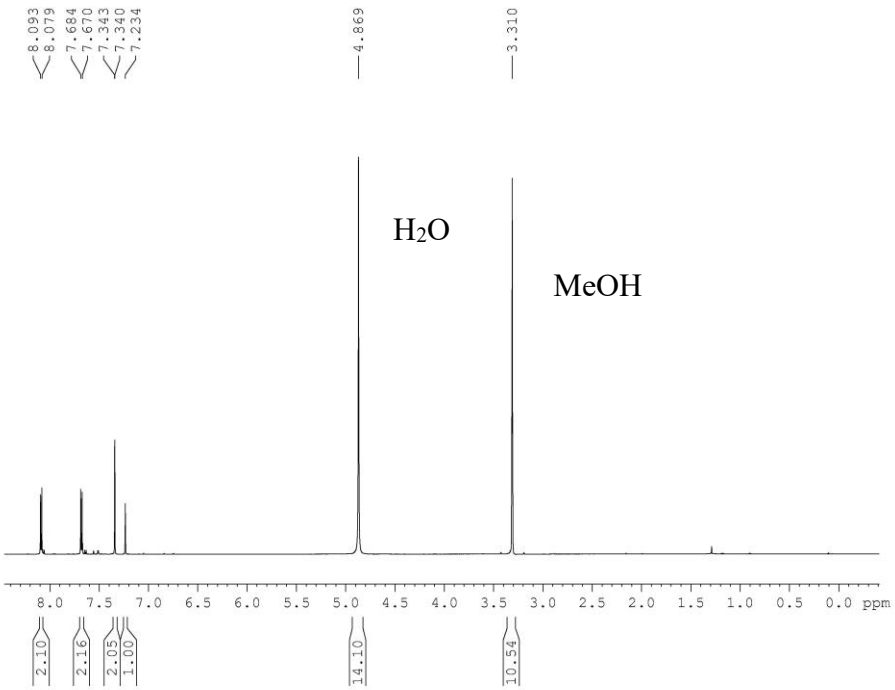


Figure 5.11. 1D ^{1}H-NMR (600 MHz) of Monomer **J** in CD_3OD.

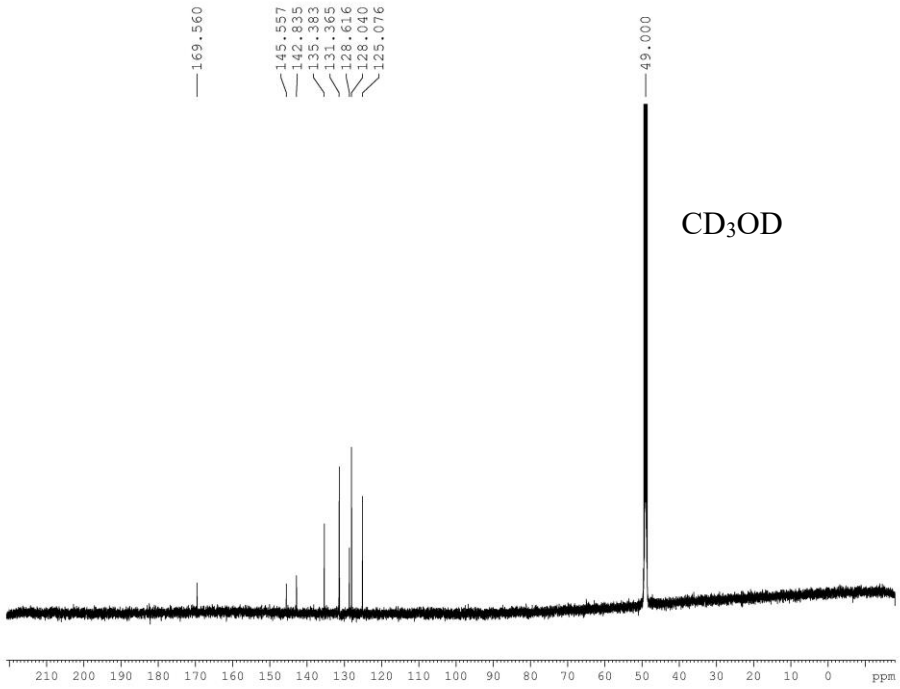


Figure 5.12. 1D ^{13}C-NMR (600 MHz) of Monomer **J** in CD_3OD.

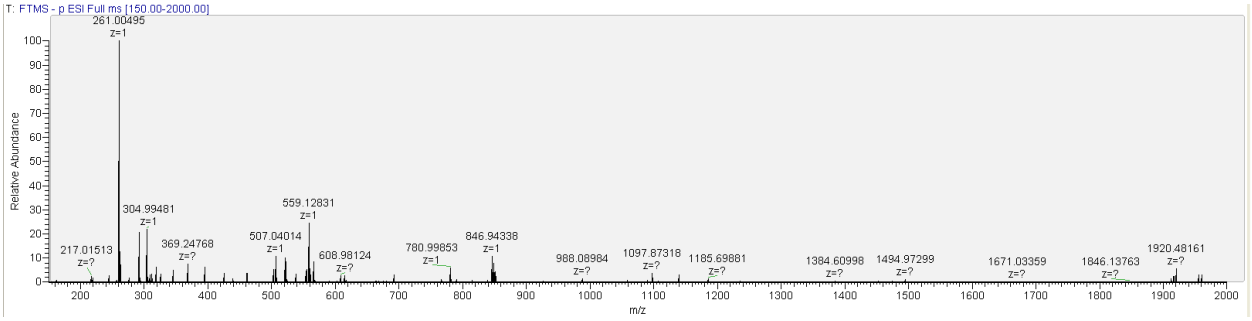
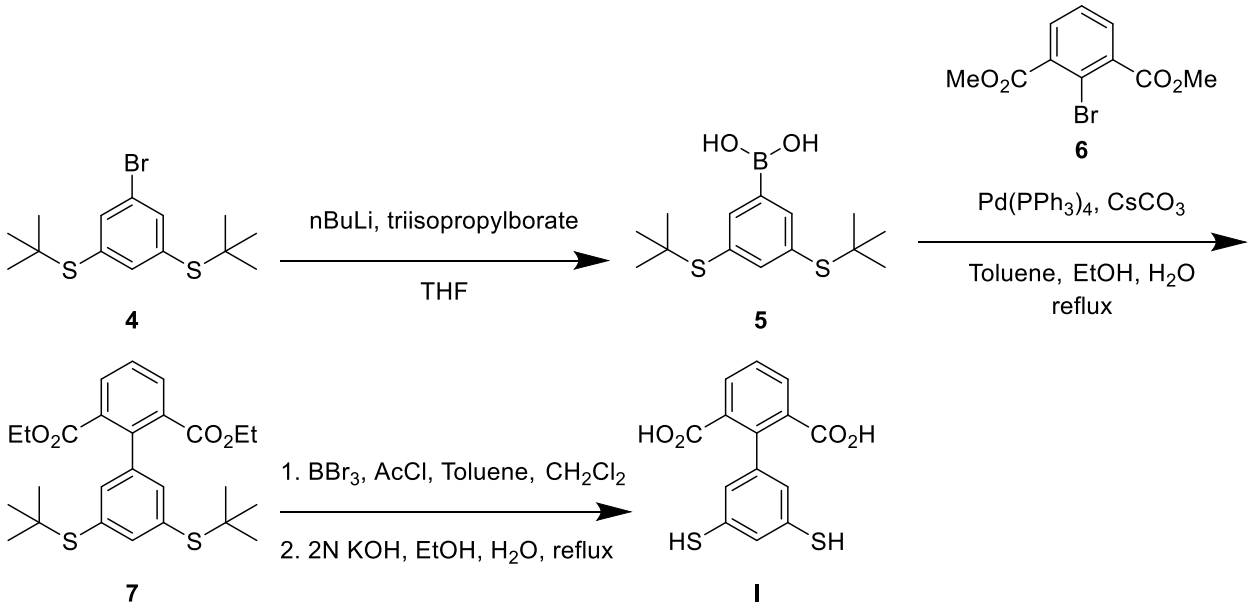


Figure 5.13. High resolution mass spectrum of Monomer **J**.

Scheme 5.3. Synthesis of Monomer **I**



Synthesis of 4. Compound **4** was synthesized according to published procedure.[161]

Synthesis of 5. Compound **4** (0.50 g, 1.51 mmol) was dissolved in anhydrous THF (6 mL) at -78 ^{o}C, and n-butyl lithium (2.5 M in hexanes, 0.75 mL, 1.9 mmol) was added dropwise. The solution was stirred at -78 ^{o}C for 1 h, and then triisopropyl borate (1.2 mL, 6.4 mmol) was added dropwise. The reaction was stirred at -78 ^{o}C for 3 h and then allowed to warm to room temperature overnight. Water (1 mL) was added to the reaction, and the solvent was removed under reduced pressure. The crude product was dissolved in dichloromethane, and 1N HCl (3 mL) was added. After stirring for 1 h, the organic fraction was separated, and the product was purified by column chromatography (ethyl acetate/hexanes gradient from 10% to 35%) to give a white solid (0.404 g, 90% yield). ^{1}H NMR ($CDCl_3$, 600 MHz): 7.857,7.854 (d, 1H, CH), 7.752-7.857 (m, 2H, CH), 1.275 (s, 18H, CH_3). ^{13}C NMR ($CDCl_3$, 600 MHz): 148.130, 147.303, 142.906, 142.777, 132.515, 46.284, 31.088. MS (calculated): 577.23 [cyclic boronic ester dimer + OH], 875.35 [open boronic ester trimer + OH], 1137.45 [cyclic boronic ester + OH], 1417.57 [cyclic boronic ester pentamer + OH]. HRMS (observed, ESI-): 577.23339 [cyclic boronic ester

dimer + OH], 875.35906 [open boronic ester trimer + OH], 1137.46613 [cyclic boronic ester + OH], 1417.58859 [cyclic boronic ester pentamer + OH].

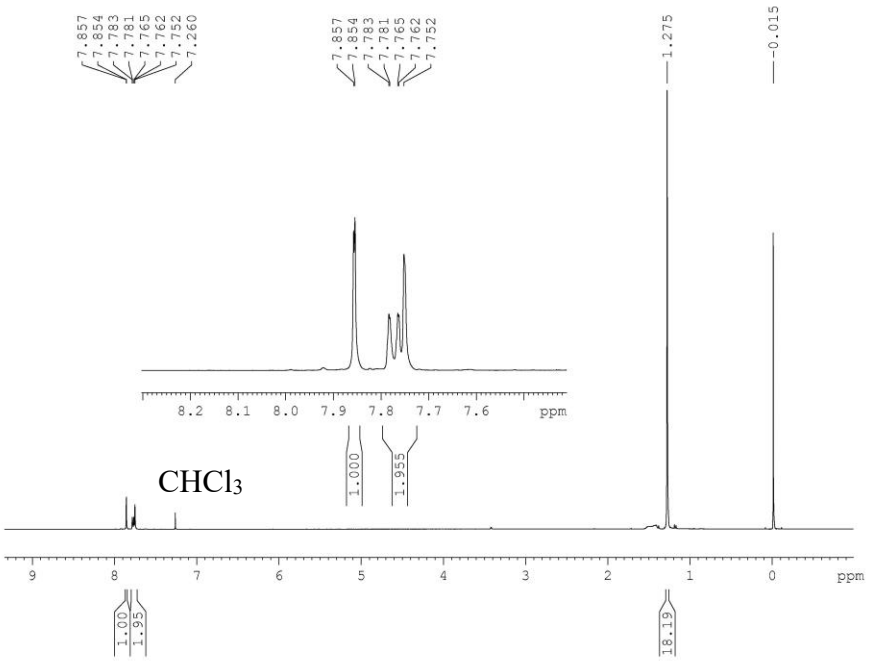


Figure 5.14. 1D ^{1}H-NMR (600 MHz) of Compound **5** in $CDCl_3$.

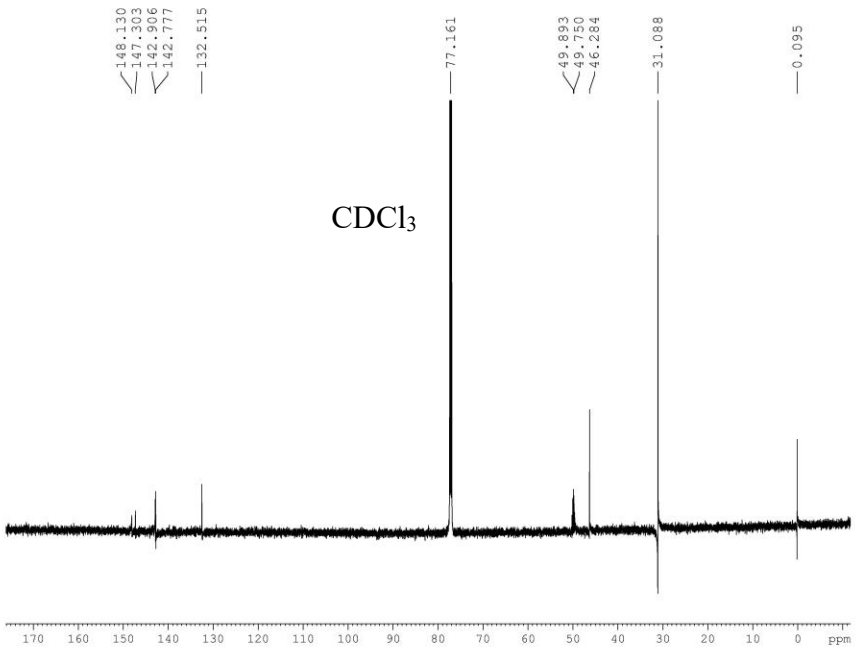


Figure 5.15. 1D ^{13}C-NMR (600 MHz) of Compound **5** in $CDCl_3$.

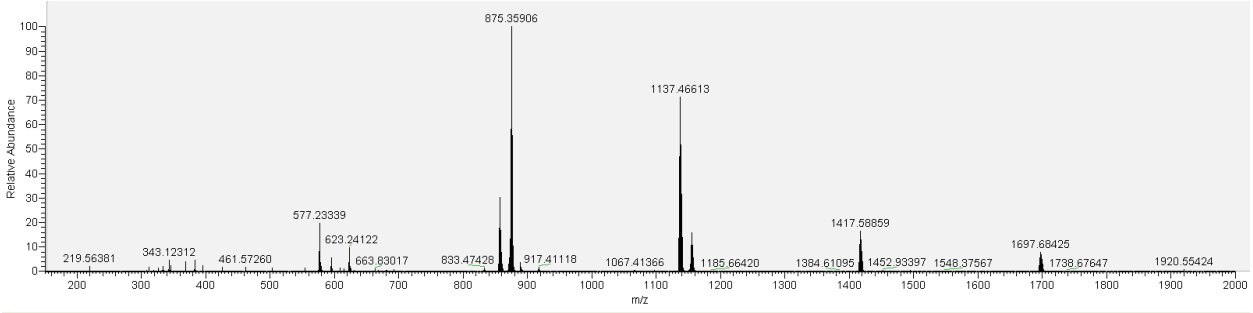


Figure 5.16. High resolution mass spectrum of Compound **5**.

Synthesis of 6. Compound **6** was synthesized according to published procedure.[162]

Synthesis of 7. A solution of $CsCO_3$ (3.60 g, 11 mmol) in H_2O (4.8 mL), toluene (21 mL), and EtOH (21 mL) was degassed by bubbling with argon for 2 h. To this flask under an atmosphere of argon were added **6** (0.69 g, 2.5 mmol), **5** (0.78 g, 2.6 mmol), and $Pd(PPh_3)_4$ (0.30 g, mmol). The solution was heated to reflux and vigorously stirred overnight. After cooling to room temperature, the solvent was removed under reduced pressure. The resulting crude solid was

purified by column chromatography (ethyl acetate/hexanes gradient from 15% to 35%) to give a white solid (0.59 g, 56% yield). ^{1}H NMR ($CDCl_3$, 600 MHz): 7.935, 7.922 (d, 2H, CH), 7.722 (s, 1H, CH), 7.506 (t, 1H, CH), 7.400, 7.398 (d, 2H), 4.024, 4.012 (q, 4H, CH_2), 1.303 (s, 18H, CH_3), 1.010 (t, 6H, CH_3). ^{13}C NMR ($CDCl_3$, 600 MHz): 167.580, 144.815, 139.949, 137.854, 133.349, 132.343, 127.801, 61.324, 46.339 31.069, 14.004. MS (calculated): 475.20 [M+H], 949.39 [2M+H], 971.37 [2M+Na^+]. HRMS (observed, ESI+): 475.19894 [M+H], 949.39195 [2M+H], 971.37069 [2M+Na^+].

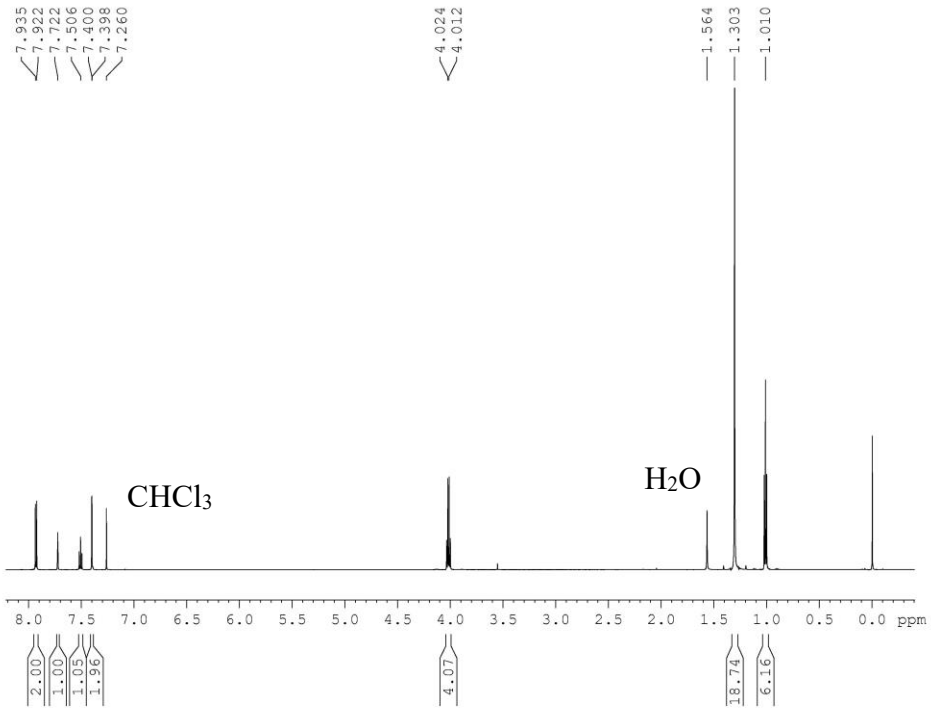


Figure 5.17. 1D ^{1}H-NMR (600 MHz) of Compound **7** in $CDCl_3$.

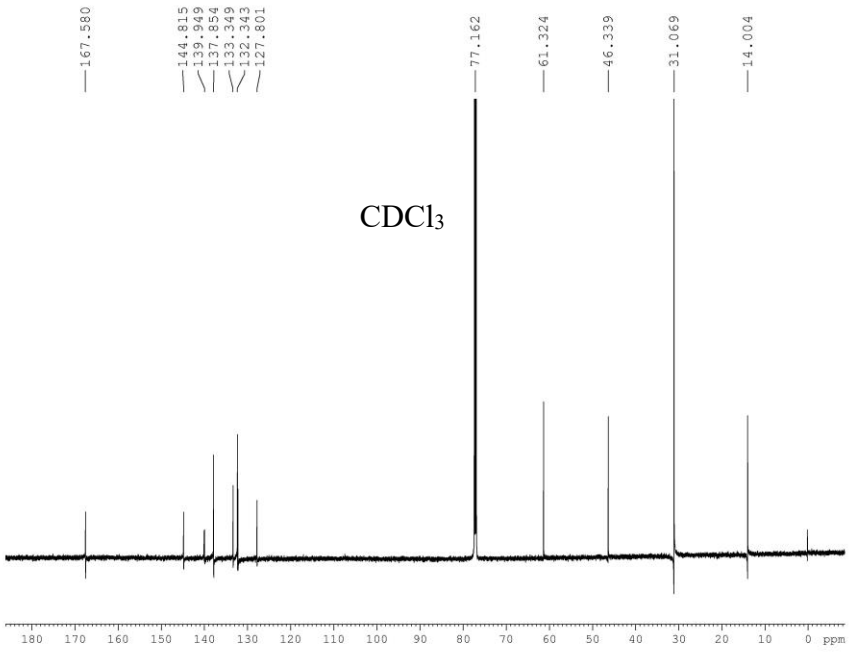


Figure 5.18. 1D ^{13}C-NMR (600 MHz) of Compound **7** in $CDCl_3$.

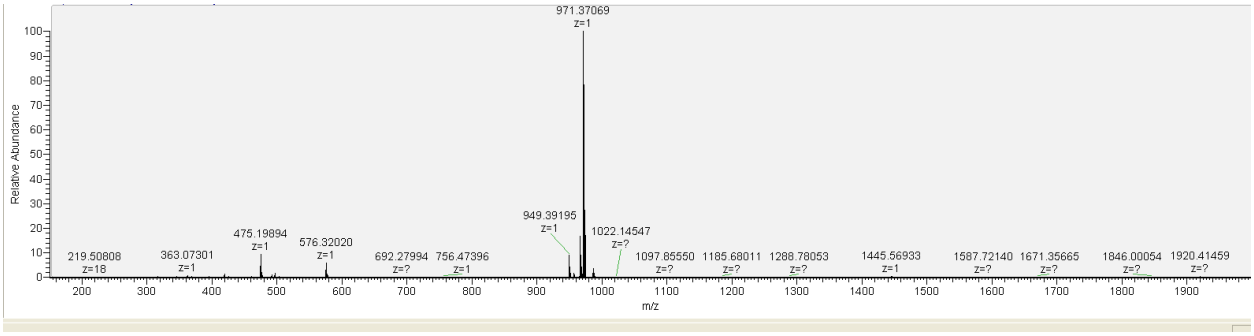


Figure 5.19. High resolution mass spectrum of Compound **7**.

Synthesis of I. Compound **7** (0.50 g, 1.1 mmol) was dissolved in a solution of toluene (2.2 mL) and acetyl chloride (1.12 mL, 15.7 mmol) under an atmosphere of nitrogen. The solution was cooled on an ice bath, and a 1.0 M solution of BBr_3 in dichloromethane (2.47 mL, 2.5 mmol) was added dropwise. Afterward, the ice bath was removed and the reaction was allowed to warm to room temperature. After 2h, the reaction was quenched by slow, dropwise addition of water

(10 mL). The organic phase was separated, washed with water (3 x 10 mL), and dried over magnesium sulfate. The solvent was removed under reduce pressure to provide a crude thioacetate intermediate. A solution of potassium hydroxide (0.80 g, 14.3 mmol) in water (4.5 mL) and ethanol (2 mL) was degassed by bubbling with nitrogen for 2 h. To this solution was the thioacetate intermediate (0.38 g, 0.91 mmol) under an atmosphere of nitrogen. The reaction was heated at reflux for 3 h. After cooling to room temperature, the solution was concentrated under reduced pressure. The product was precipitated out by addition of 3N HCl. The precipitate was collected by vacuum filtration and dried under vacuum overnight to give a white solid (0.24 g, 73% yield over 2 steps). ^{1}H NMR (CD_3OD 600 MHz): 7.892, 7.879 (d, 2H, CH), 7.537 (t, 1H, CH), 7.170 (s, 1H, CH), 6.905, 6.902 (s, 2H, CH). ^{13}C NMR ($CDCl_3$, 600 MHz): 171.071, 142.068, 140.533, 135.252, 133.333, 132.375, 128.777, 128.331, 127.169. MS (calculated): 304.99 [M-H], 610.99 [2M-H]. HRMS (observed, ESI+): 304.99499 [M-H], 610.99692 [2M-H].

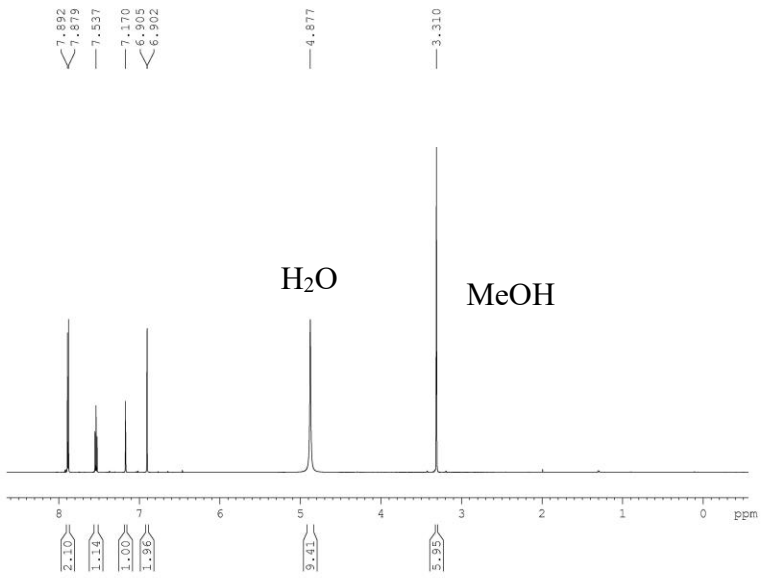


Figure 5.20. 1D ^{1}H-NMR (600 MHz) of Monomer **I** in CD_3OD.

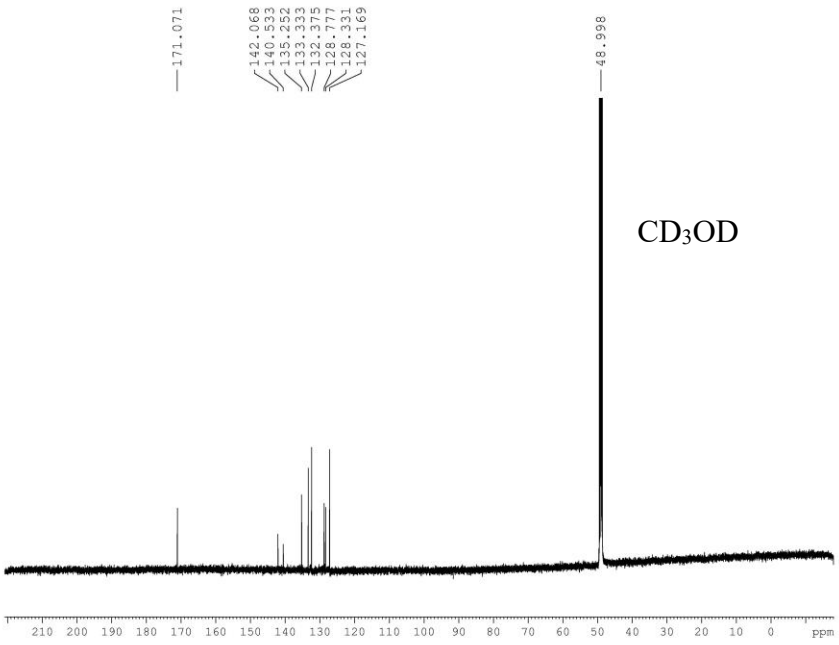


Figure 5.21. 1D ^{13}C-NMR (600 MHz) of Monomer **I** in CD_3OD.

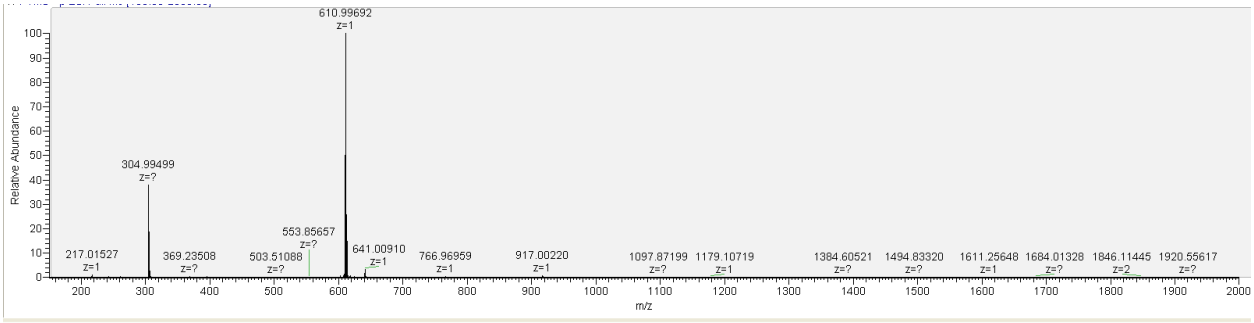


Figure 5.22. High resolution mass spectrum of Monomer **I**.

Synthesis of A_2I and A_2J

DCLs were prepared on a preparative scale using butyltrimethylammonium bromide as a guest for templating the formation of the desired receptor. Libraries were prepared using 2.5 mM **A**, 2.5 mM biphenyl monomer (**I** or **J**), and 10 mM guest in 50 mM sodium borate buffer at pH 8.5. Equilibration was allowed to occur for at least 5 days, after which the desired receptor was

purified by semi-preparative reversed-phase HPLC (solvent A: 10 mM NH_4OAc in H_2O; solvent B: 10 mM NH_4OAc in 9:1 CH_3CN:H_2O). For **A_2I**, a gradient from 10% B to 100% B in 1 h was used for purification. For **A_2J**, the following gradient was used: 0% B to 35% B over 1 min, 35% B to 40% B over 5 min, 40% B to 55% B over 30 min, and 55% B to 100% B over 30 min. Collected fractions were lyophilized to give off-white solids.

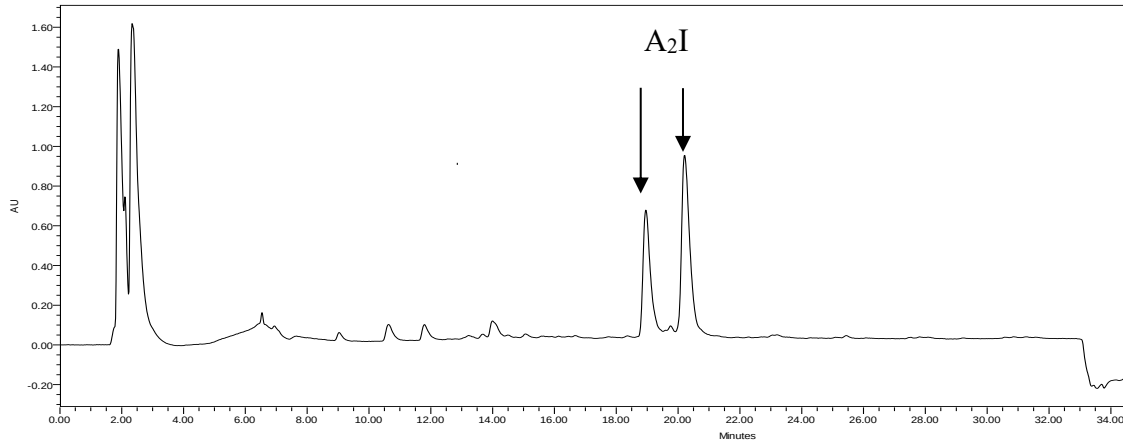


Figure 5.23. Semi-preparative reversed-phase HPLC traces of the purification of **A_2I** from a preparative scale DCL. Absorbance measured at 214 nm.

A_2I-1

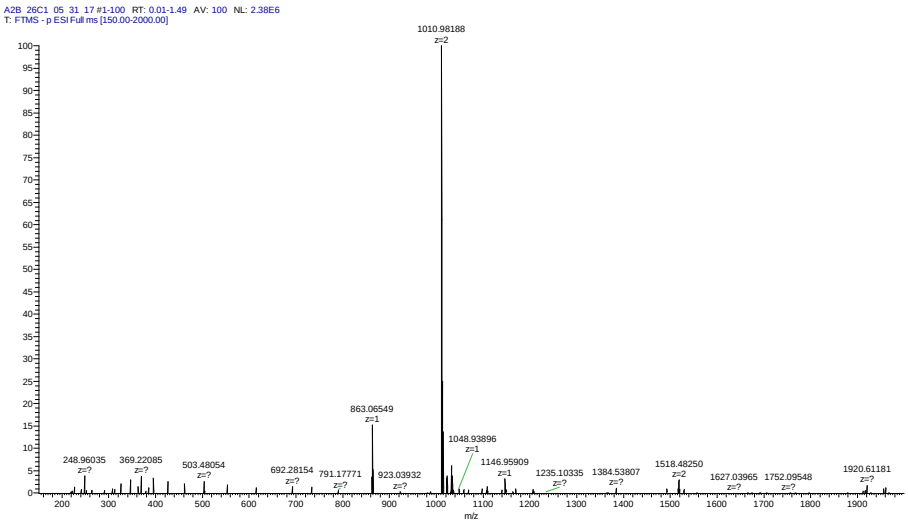


Figure 5.24. High resolution mass spectrum of **A_2I-1** (ESI -). MS calculated: 1010.98 [M-H]. MS observed: 1010.98188 [M-H].

A_2I-2

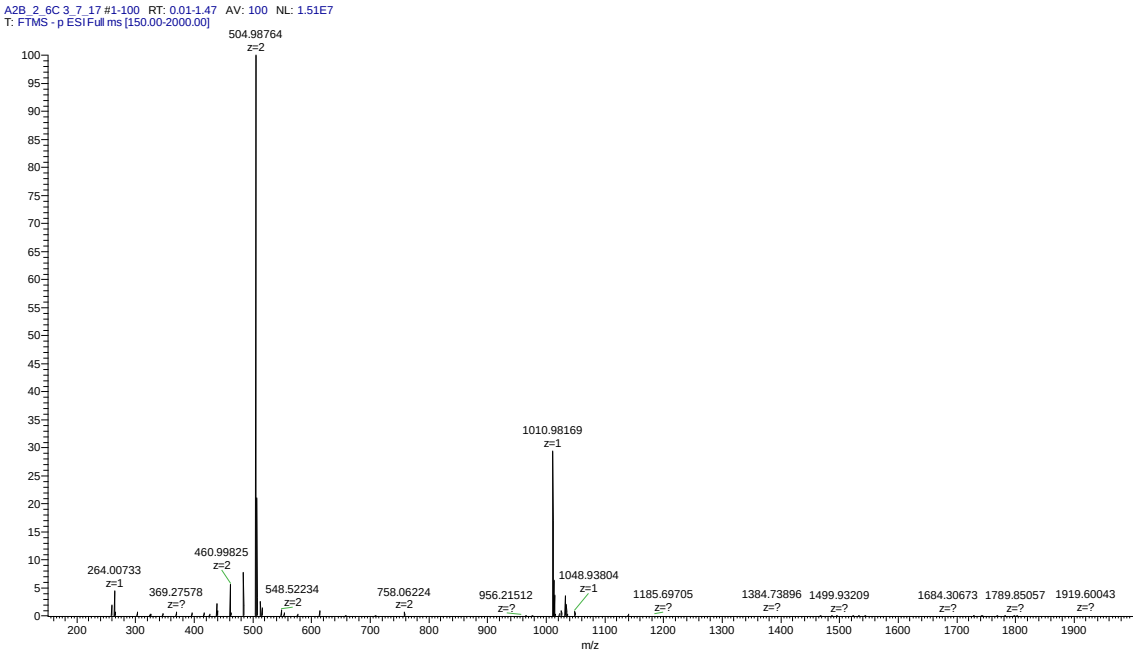
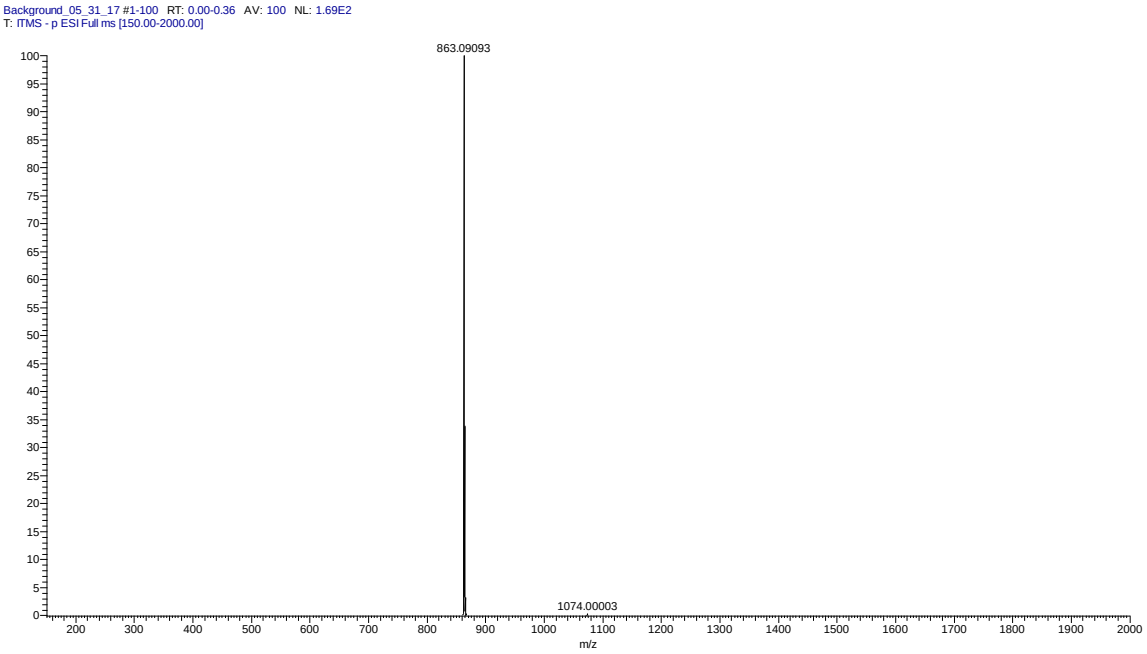


Figure 5.25. High resolution mass spectrum of **A_2I-2** (ESI -). MS calculated: 1010.98 [M-H], 504.99 [M-2H]. MS observed: 1010.98169 [M-H], 504.98764 [M-2H].

Background

(a)



(b)

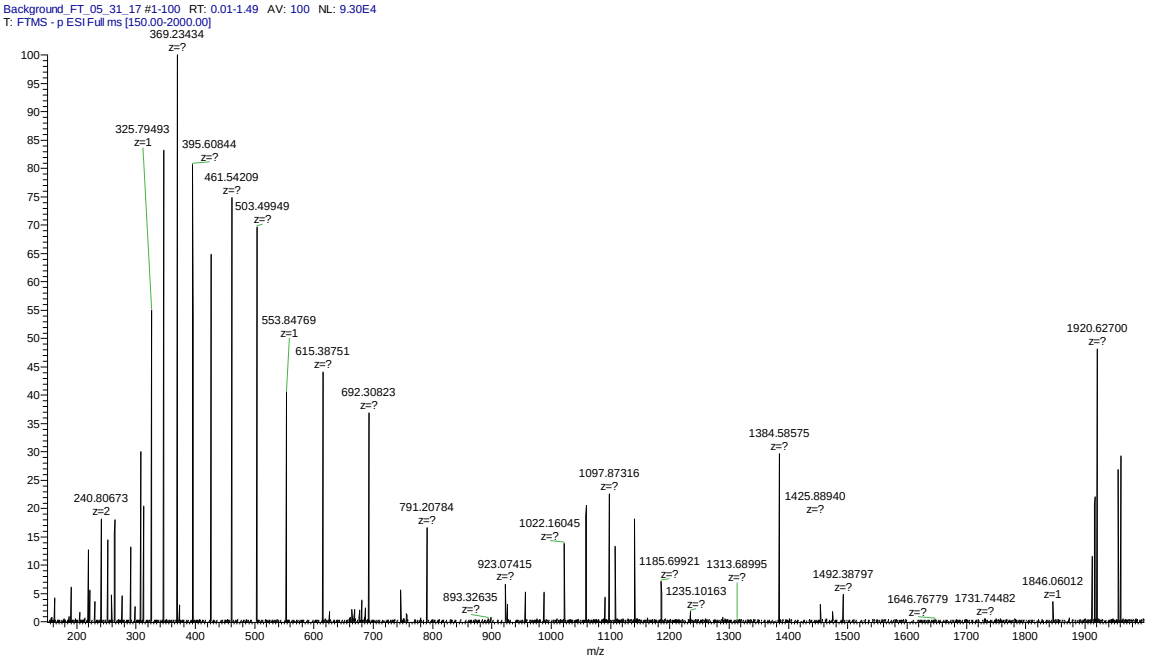


Figure 5.26. (a) Low resolution mass spectrum of background (ESI -). (b) High resolution mass spectrum of background (ESI -).

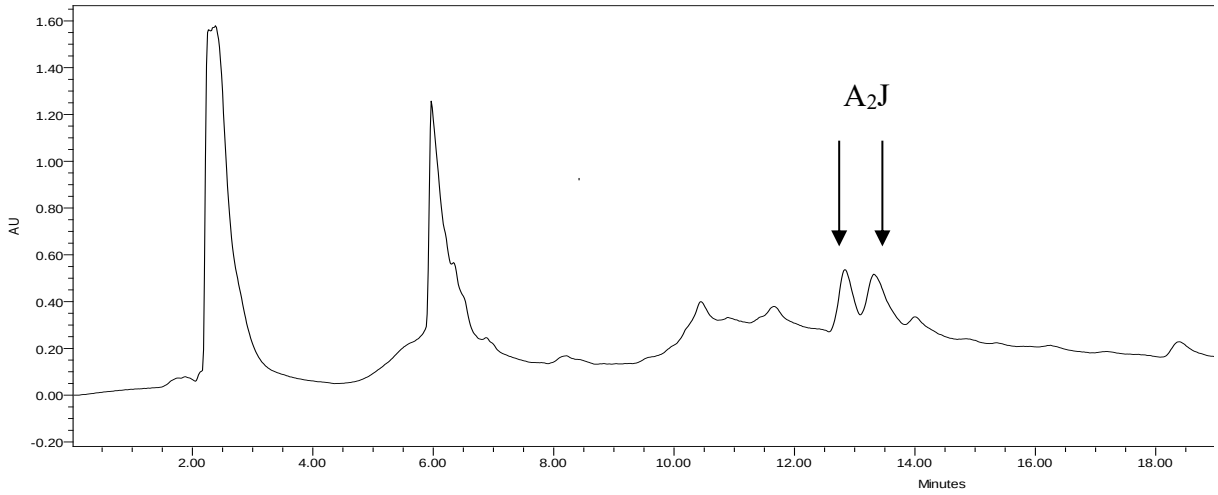


Figure 5.27. Semi-preparative reversed-phase HPLC traces of the purification of **A2J** from a preparative scale DCL. Absorbance measured at 214 nm.

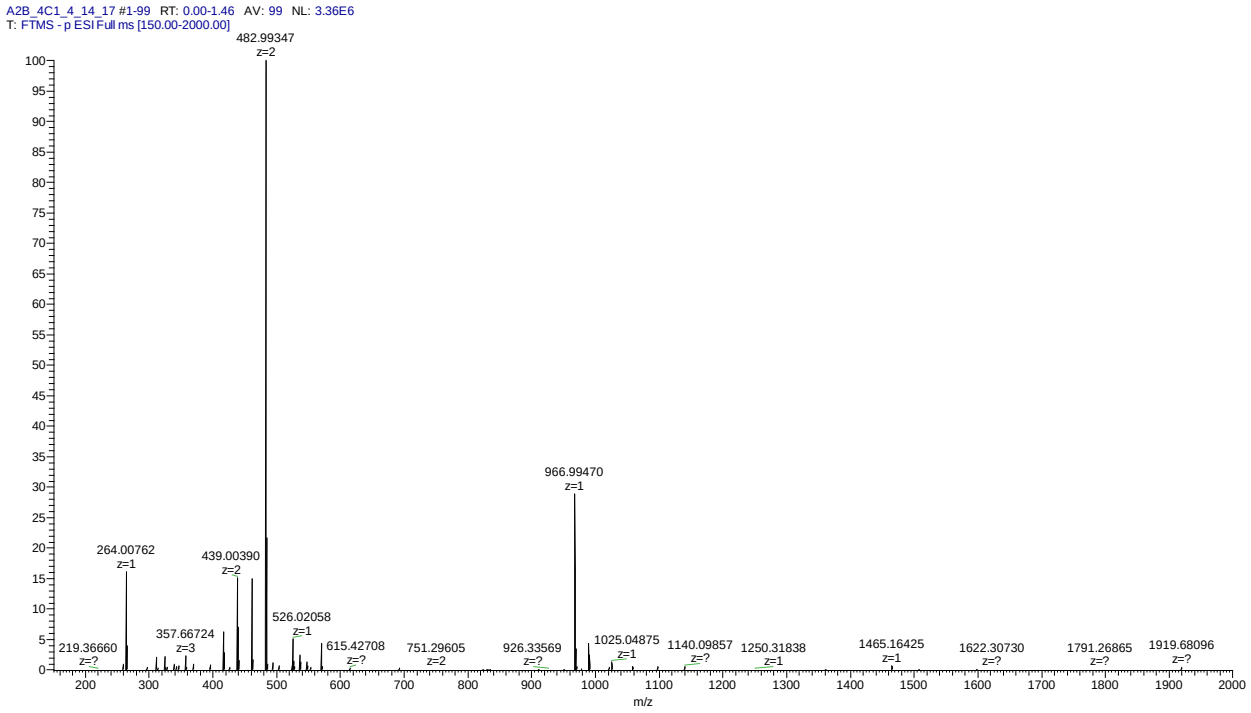


Figure 5.28. High resolution mass spectrum of **A2J-1** (ESI -). MS calculated: 966.99 [M-H], 482.99 [M-2H]. MS observed: 966.99470 [M-H], 482.99347 [M-2H].

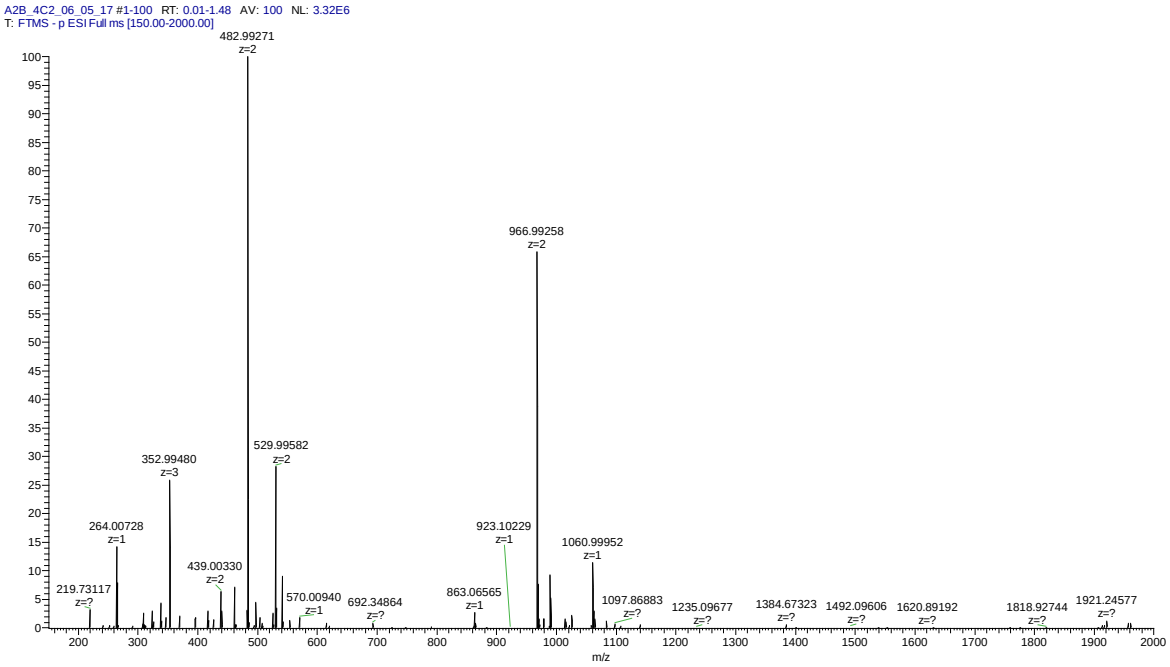


Figure 5.29. High resolution mass spectrum of **A_2J-1** (ESI -). MS calculated: 966.99 [M-H], 482.99 [M-2H]. MS observed: 966.99258 [M-H], 482.99271 [M-2H].

NMR Characterization of A_2I and A_2J

Structural characterization of **A_2I-1**, **A_2I-2**, and **A_2J-2** was conducted in 50 mM sodium borate-d3 D_2O, pH 8.5 with 0.05 mM DSS (as a NMR standard) at room temperature using a Bruker 500 MHz instrument (40 ^oC, 400 scans each, 3s pre-saturation and water suppression). The ^1H NMR spectrum of **A_2J-1** was taken in MeOD (40 ^oC, 900 scans, 3s pre-saturation and water suppression).

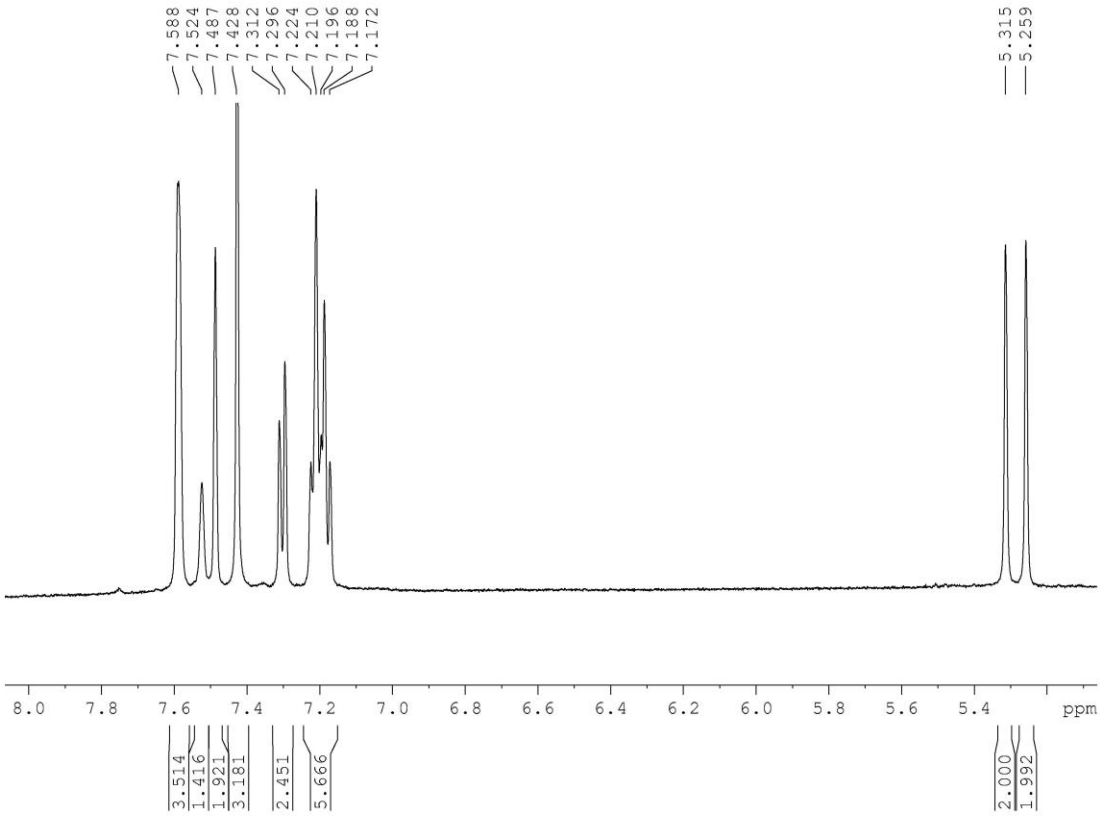


Figure 5.30. ^1H NMR spectrum of **A_2I-1** in 50 mM sodium borate-d_3 D_2O, pH 8.5 at room temperature. 7.588 (d, 4H, Ar-H), 7.524 (s, 1H, Ar-H), 7.487 (s, 2H, Ar-H), 7.428 (m, 3H, Ar-H), 7.294-7.312 (d, 2H, Ar-H), 7.172-7.224 (m, 6H, Ar-H), 5.315 (s, 2H, C-H), 5.259 (s, 2H, C-H).

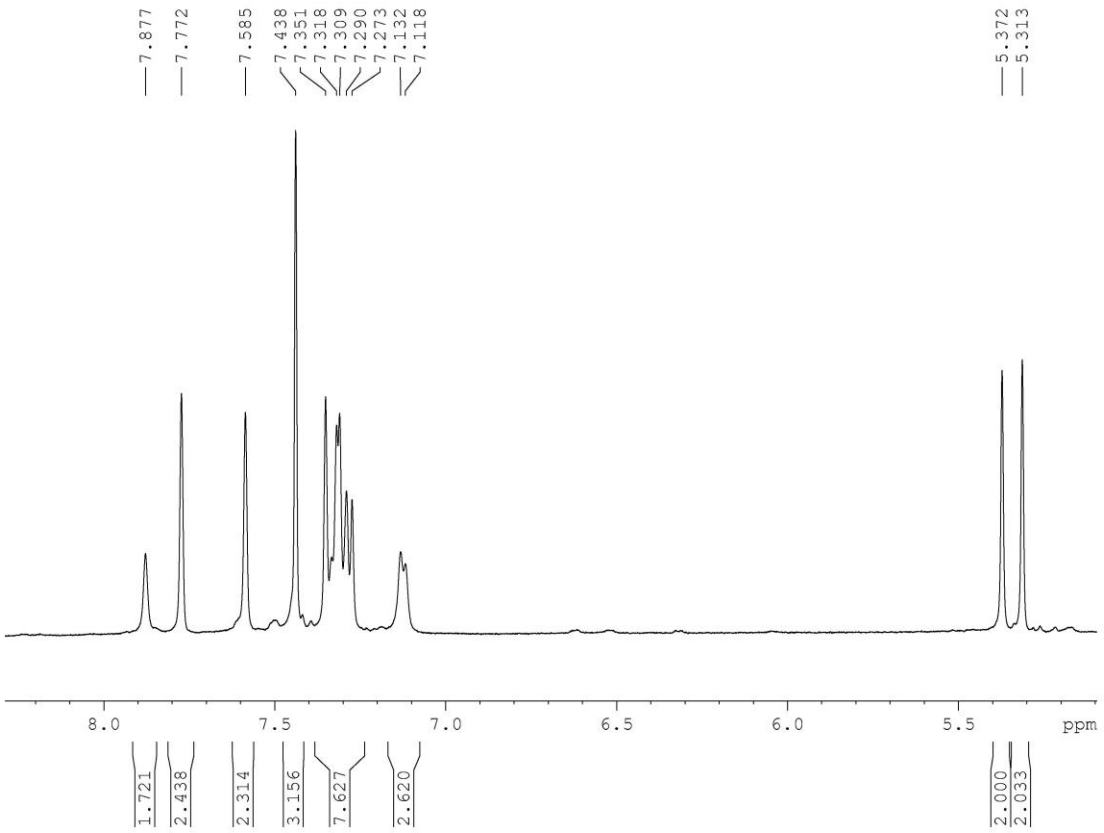


Figure 5.31. ^{1}H NMR spectrum of **A_2I-2** in 50 mM sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard) at room temperature. 7.877 (s, 2H, Ar-H), 7.772 (s, 2H, Ar-H), 7.585 (s, 2H, Ar-H), 7.438 (m, 3H, Ar-H), 7.273-7.7.351 (m, 7H, Ar-H), 7.118-7.132 (d, 2H, Ar-H), 5.372 (s, 2H, C-H), 5.313 (s, 2H, C-H).

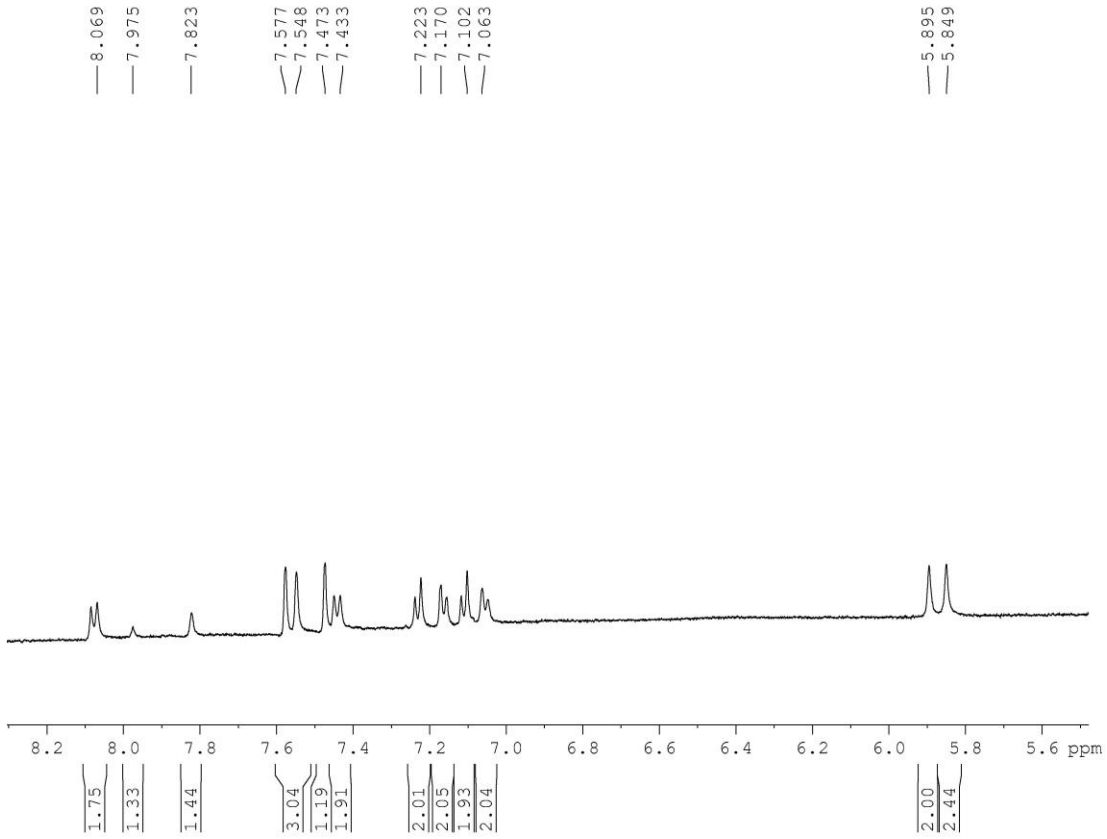


Figure 5.32. ^{1}H NMR spectrum of **A$_2$J-1** in MeOD at 40 ^{o}C. 8.069 (d, 2H, Ar-H), 7.975 (s, 1H, Ar-H), 7.823 (s, 2H, Ar-H), 7.548-7.577 (m, 3H, Ar-H), 7.473 (s, 1H, Ar-H), 7.433 (d, 2H, Ar-H), 7.223 (d, 2H, Ar-H), 7.170 (d, 2H, Ar-H), 7.102 (d, 2H, Ar-H), 7.063 (d, 2H, Ar-H), 5.895 (s, 2H, C-H), 5.849 (s, 2H, C-H).

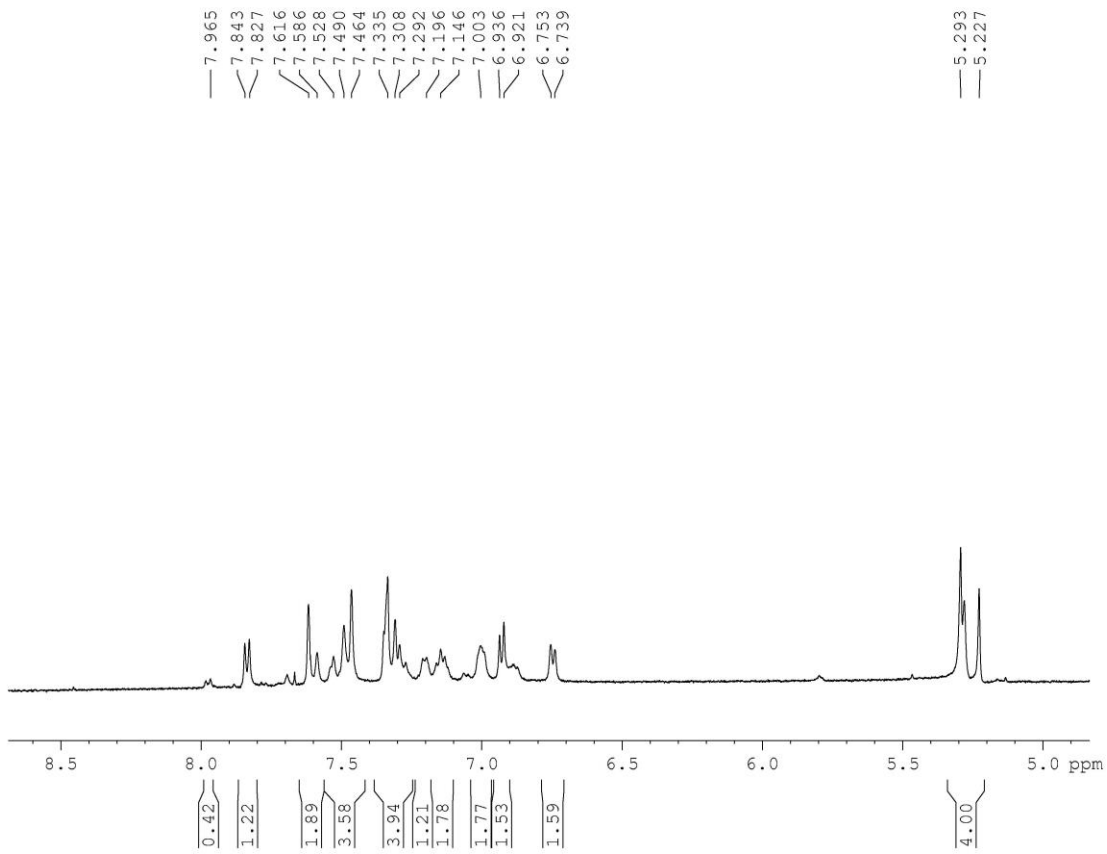


Figure 5.33. ^{1}H NMR spectrum of **A_2J-2** in 50 mM sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard) at 40 ºC. Due to difficulty in purification, the second isomer could only be obtained in 83% purity, with the impurity presumably arising from HPLC peak overlap with the first isomer. 7.827-7.843 (d, 1H, Ar-H), 7.568-7.616 (m, 2H, Ar-H), 7.464-7.490 (m, 3H, Ar-H), 7.292-7.335 (m, 4 H, Ar-H), 7.196 (d, 1H, Ar-H), 7.146 (d, 2H, Ar-H), 7.003 (d, 2H, Ar-H), 6.921-6.936 (d, 2H, Ar-H), 6.739-6.753 (d, 2H, Ar-H), 5.227-5.293 (m, 4H, C-H).

Dynamic Combinatorial Libraries for A_2J

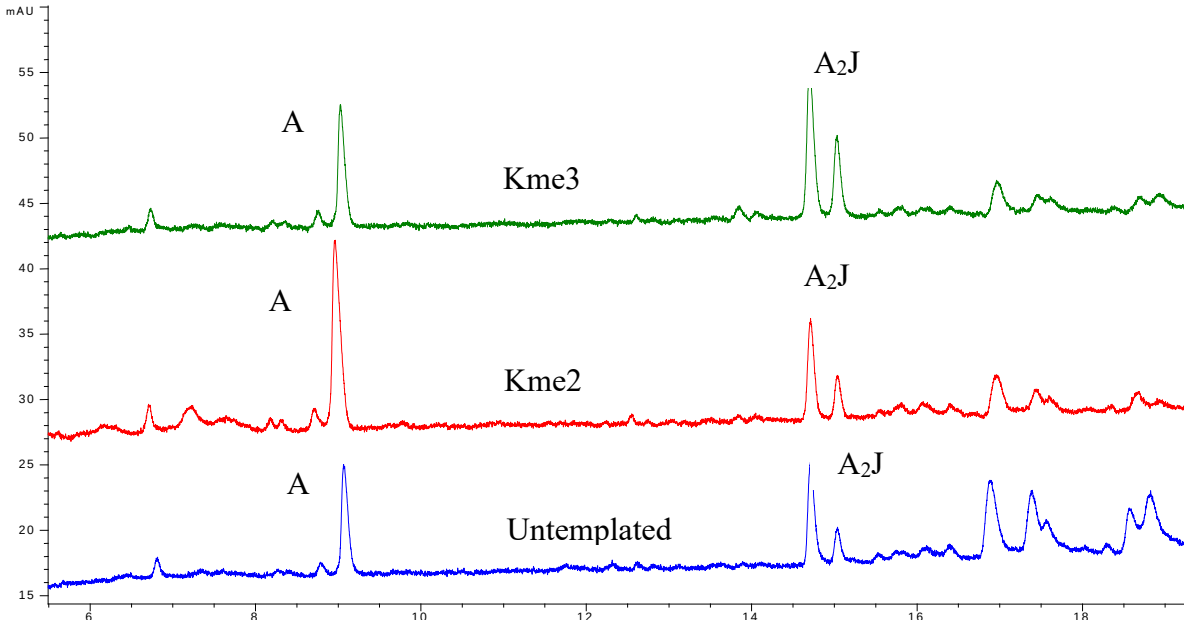


Figure 5.34. Overlaid HPLC traces of DCLs after 7 days containing 1 mM **A**, 1 mM **J**, and 2 mM guest (absorbance at 280 nm). **A_2J** is amplified in the presence of Kme2 and Kme3.

Peptide Synthesis

Peptides were synthesized on a 0.6 mmol scale via solid-phase peptide synthesis in a peptide synthesis flask using Fmoc protected amino acids. Coupling reagents were HOBt/HBTU in DMF (4 equiv each), with coupling times of 60 min, followed by rinsing with DMF, MeOH, and CH_2Cl_2. Deprotection was performed with 20% piperidine in DMF. All peptides were acylated at the N-terminus with a solution of 5% acetic anhydride and 6% 2,6-lutidine in DMF. Cleavage was performed using a cocktail of 95% TFA/2.5% triisopropylsilane/2.5% H_2O for 3 h. Peptides were purified by semi-preparative reverse-phase HPLC on a C18 column at a flow rate of 4 mL/min. Peptides were purified with a linear gradient of A and B (A: 95% H_2O/5% CH_3CN with 0.1% TFA, B: 95% CH_3CN/5% H_2O with 0.1 % TFA).

Methylated peptides were synthesized with 2 equivalents of Fmoc-Lys$(Me)_2$-OH HCl or Fmoc-Lys(Me,Boc)-OH purchased from EMP Millipore and coupled for 5 h. The trimethyllysine containing peptides were synthesized by reacting corresponding dimethylated peptides (0.06 mmol scale) prior to cleavage from the resin with MTBD (7-Methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene, 10 equiv.) and methyl iodide (10 equiv.) in DMF (5 mL) for 5 h with bubbling nitrogen in a peptide flask. After washing the resin with DMF and CH_2Cl_2 and drying, the peptide was cleaved as normal.

H3K9: Ac-WGGGQTARKSTG-NH_2

H3K9me: Ac-WGGGQTARKmeSTG -NH_2

H3K9me_2: Ac-WGGGQTARKme2STG-NH_2

H3K9me3: Ac-WGGGQTARKme_3STG-NH_2

Kme2GGY: Ac-Kme2GGY-NH_2

Kme3GGY: Ac-Kme3GGY-NH_2

H3K9

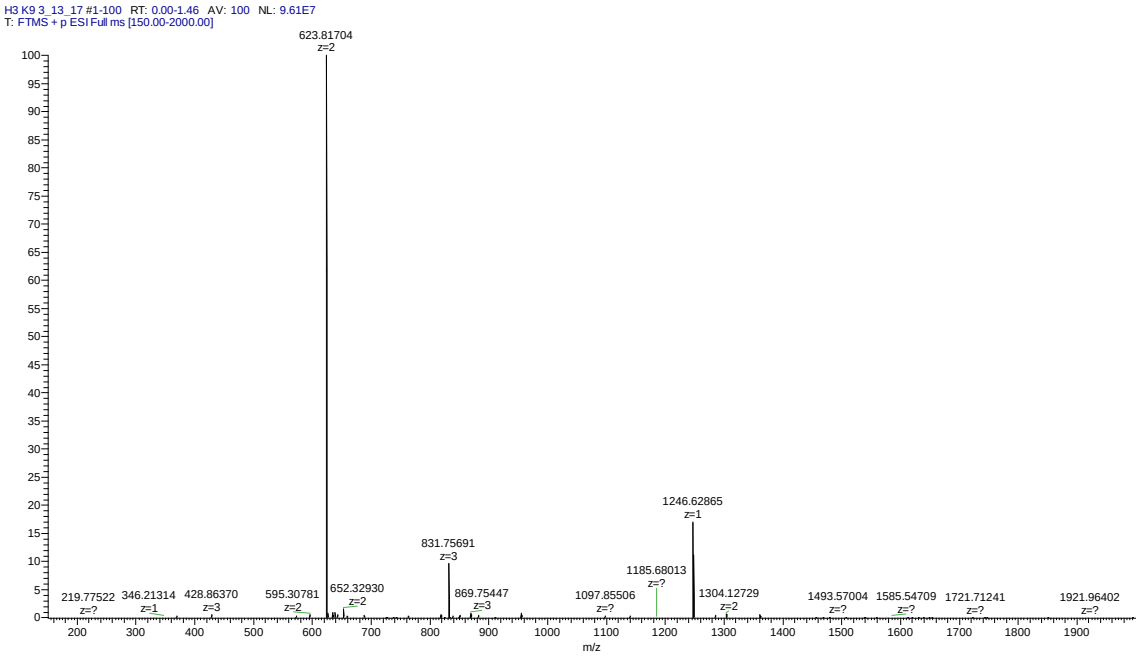


Figure 5.35. High resolution mass spectrum of H3K9 (ESI +). MS calculated: 1246.63 [M+H], 623.82 [M+2H]. MS observed: 1246.62865 [M+H], 623.81704 [M+2H].

H3K9me

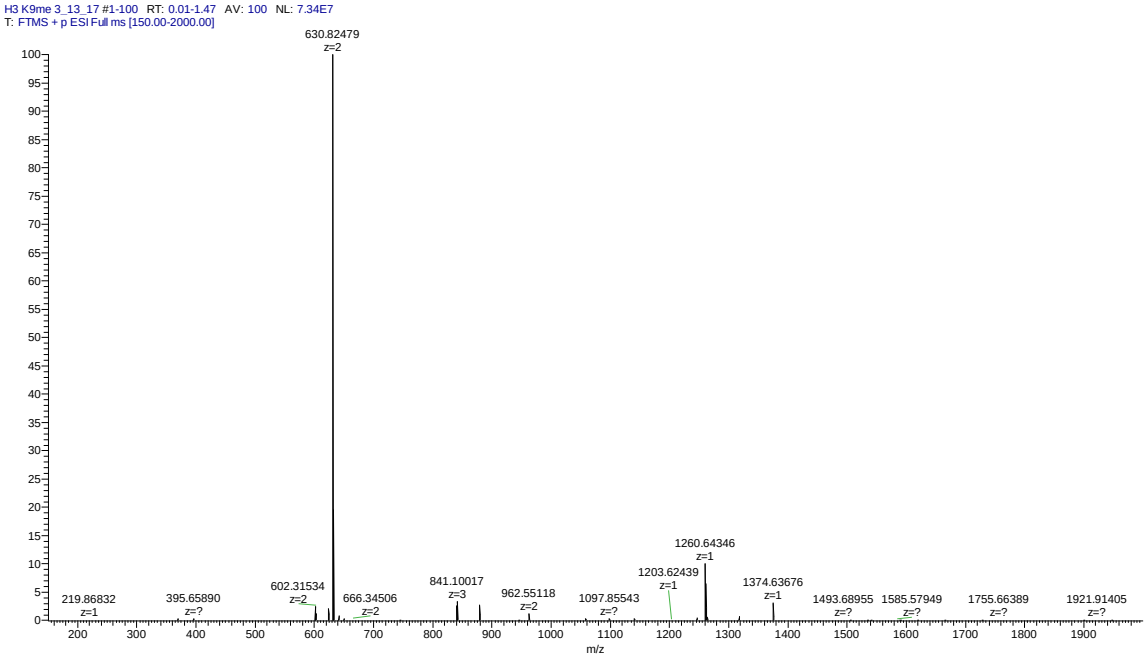


Figure 5.36. High resolution mass spectrum of H3K9me (ESI +). MS calculated: 1260.64 [M+H], 630.83 [M+2H]. MS observed: 1260.64346 [M+H], 630.82479 [M+2H].

H3K9me2

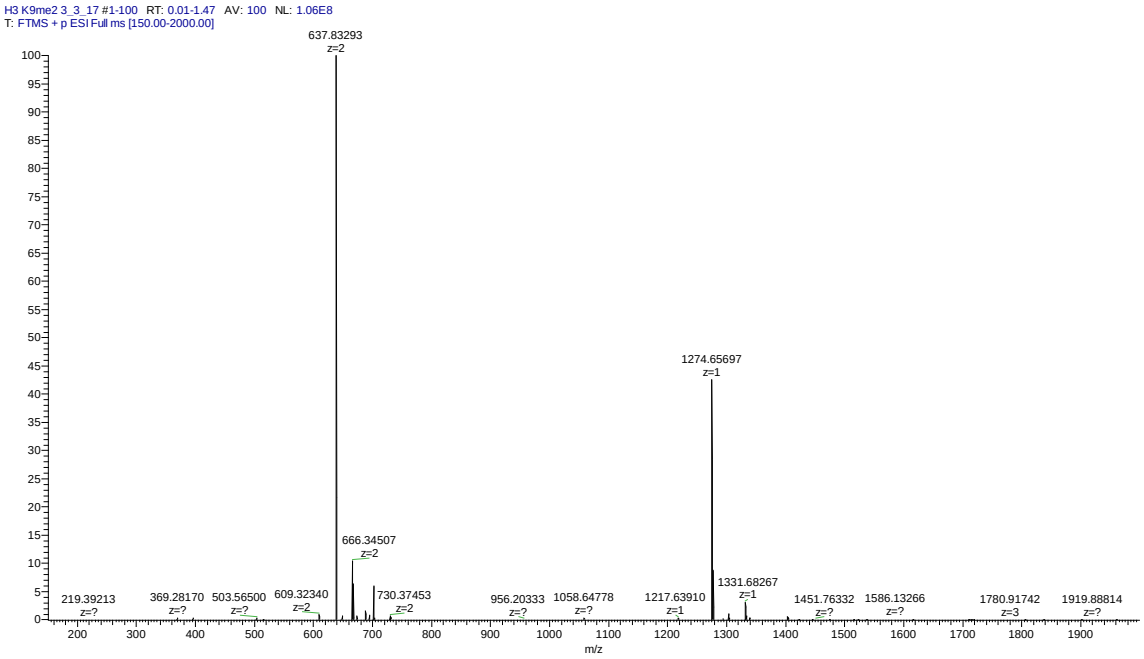


Figure 5.37. High resolution mass spectrum of H3K9me2 (ESI +). MS calculated: 1274.66 [M+H], 637.83 [M+2H]. MS observed: 1274.65697 [M+H], 637.83293 [M+2H].

H3K9me3

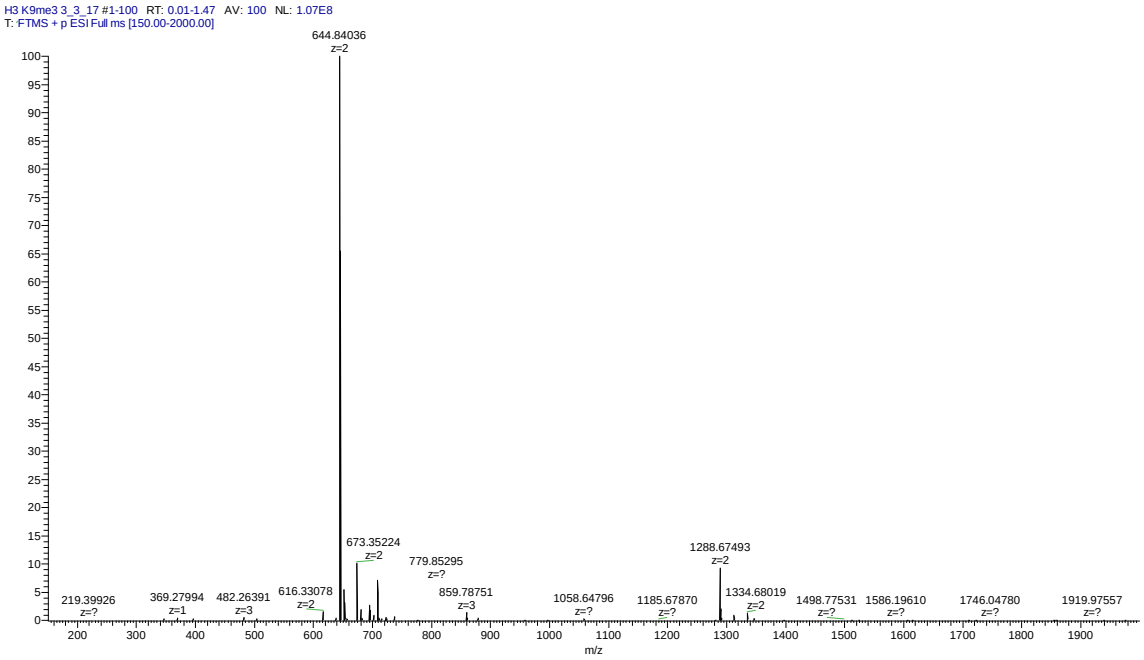


Figure 5.38. High resolution mass spectrum of H3K9me3 (ESI +). MS calculated: 1288.68 [M+H], 644.84036 [M+2H]. MS observed: 1288.67493 [M+H], 644.84 [M+2H].

Kme2GGY

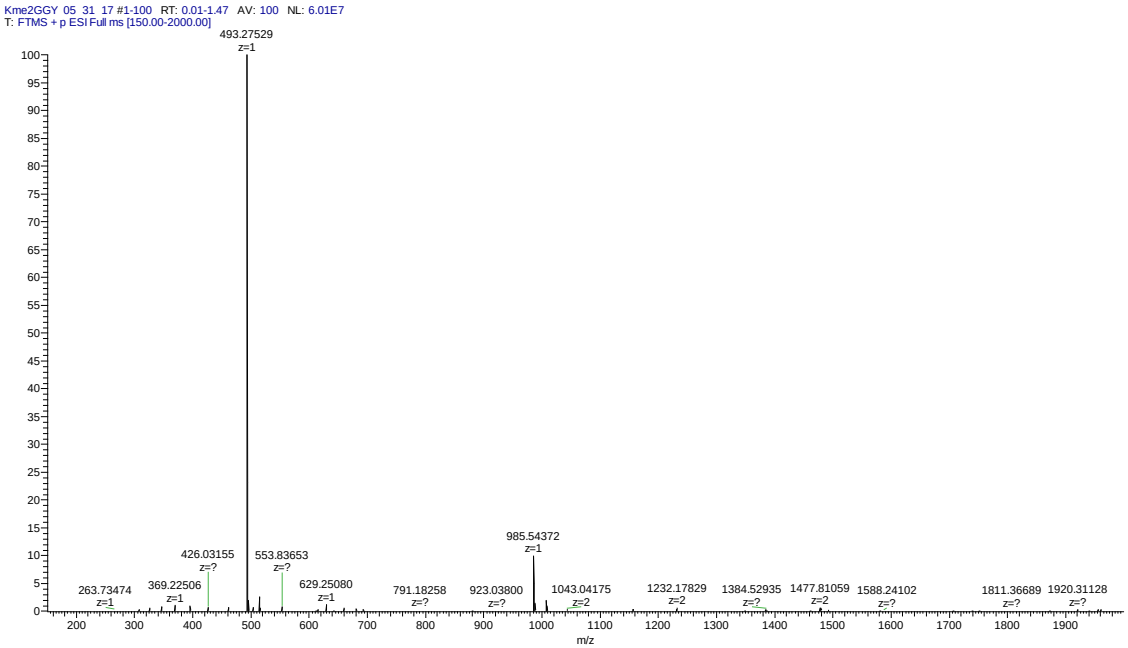


Figure 5.39. High resolution mass spectrum of Kme2GGY (ESI +). MS calculated: 492.28 [M+H], 985.55 [2M+H]. MS observed: 493.27529 [M+H], 985.54372 [2M+H].

Kme3GGY

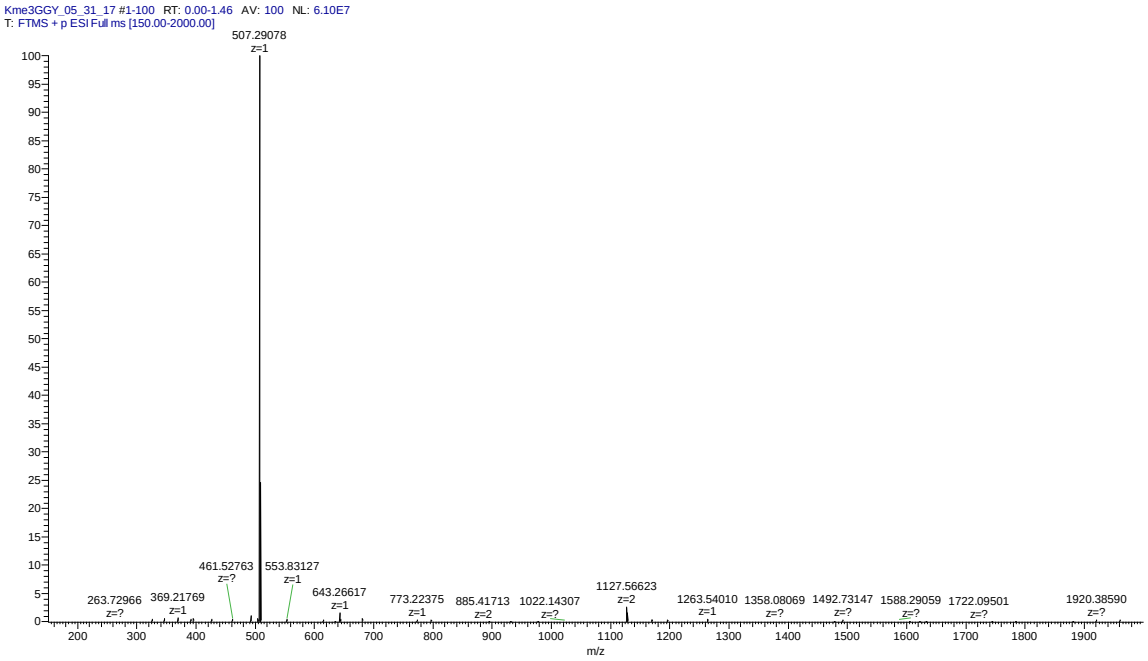


Figure 5.40. High resolution mass spectrum of Kme3GGY (ESI +). MS calculated: 507.29 [M+H]. MS observed: 507.29078 [M+H].

Extinction Coefficient Determination for A_2I and A_2J.

Receptors were purified using reversed phase-HPLC with 10 mM NH_4OAc as the mobile phase additive. After purification, they were lyophilized for five days to remove volatile NH_4OAc salts. The dried compound was then dissolved in anhydrous methanol and filtered using a 0.33 μm filter to remove any remaining salts. Methanol was evaporated and the receptors were further dried under vacuum. The mass was accurately determined, and a stock solution of receptor (1.07 mM for **A_2I** and 0.607 mM for **A_2J**) in 10 mM sodium borate buffer, pH 8.5 was prepared. The stock solution was serially diluted (80:20, receptor: buffer) to give 10 to 12 concentrations. The absorbance at 315 nm or 300 nm was measured for each concentration, which was then plotted against the concentration. The extinction coefficient of the receptors was determined from the slope of the line of regression using Beer's law and was found to be 401.5 $M^{-1}cm^{-1}$ for **A_2I** and 1.4 $M^{-1}cm^{-1}$ for **A_2J**.

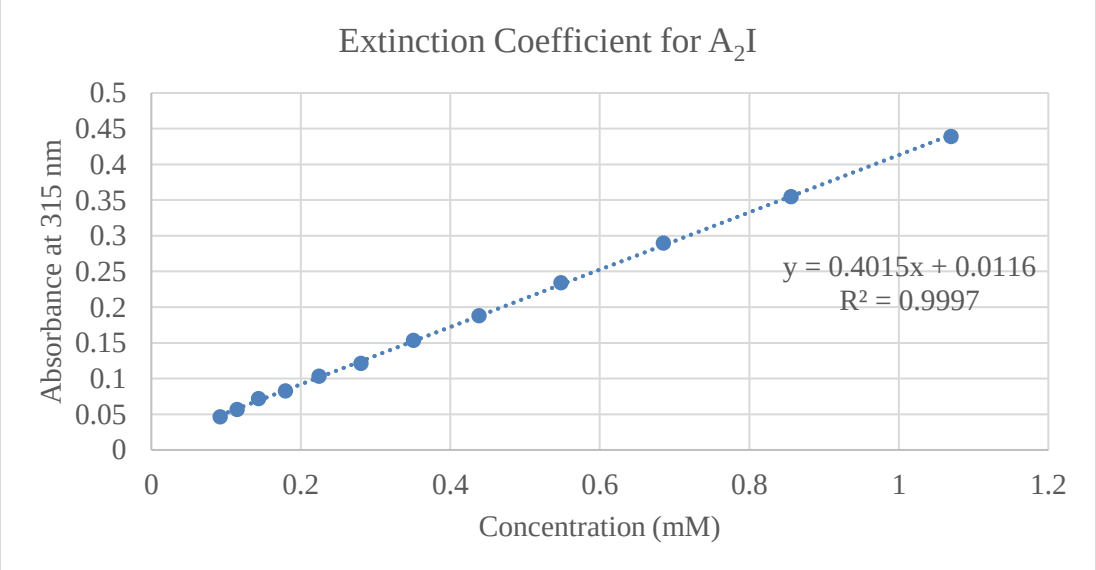


Figure 5.41. Extinction coefficient determination of **A_2I** from the slope of the line of regression.

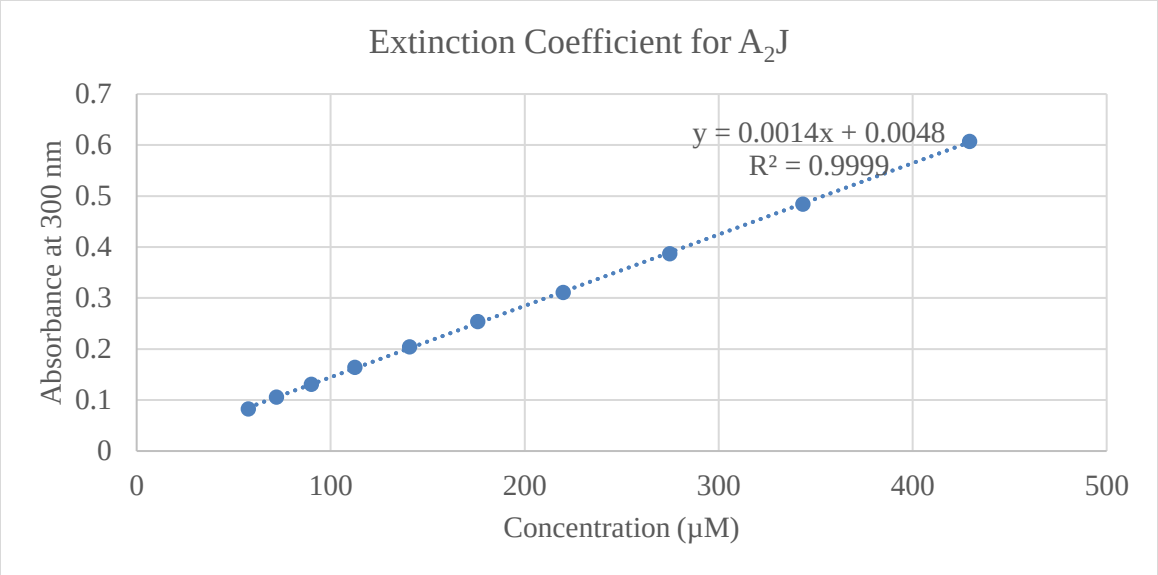


Figure 5.42. Extinction coefficient determination of **A_2J** from the slope of the line of regression.

NMR Analysis of Binding to Kme2 and Kme3

NMR binding experiments were conducted to characterize the mode of binding of **A_2I-2** to a Kme2GGY and Kme3GGY peptide. Experiments were carried out in 50 mM sodium borate-d3 D_2O, pH 8.5 with 0.05 mM DSS (as an internal standard) at room temperature and 40 ^oC using a Bruker 500 MHz instrument (400 scans each, 3s pre-saturation and water suppression). suppression). Stock solutions of receptor and peptide were prepared in the NMR buffer (the concentrations of each were determined using a NanoDrop^{TM}). Samples were prepared by measuring out aliquots from the stock and diluting with buffer to the appropriate concentration.

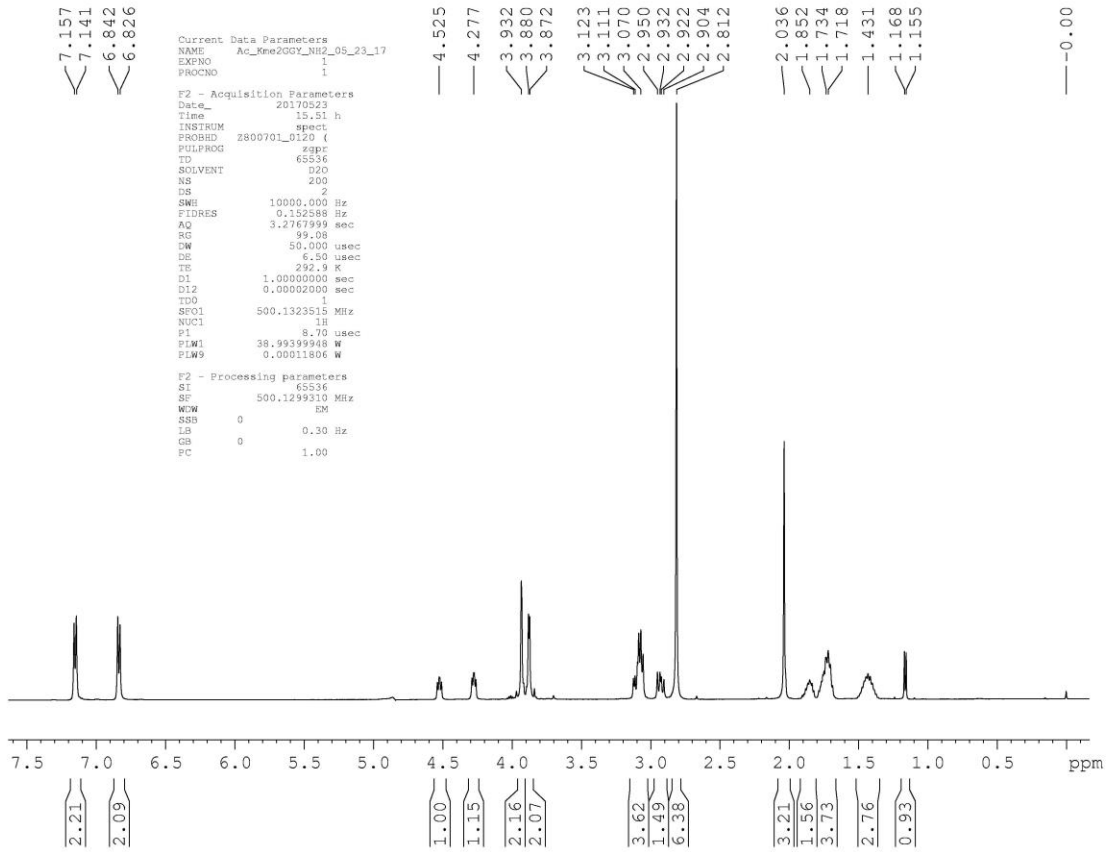


Figure 5.43. ^{1}H NMR spectrum of **Kme2GGY** (5.6 mM) alone in 50 mM sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard) at room temperature.

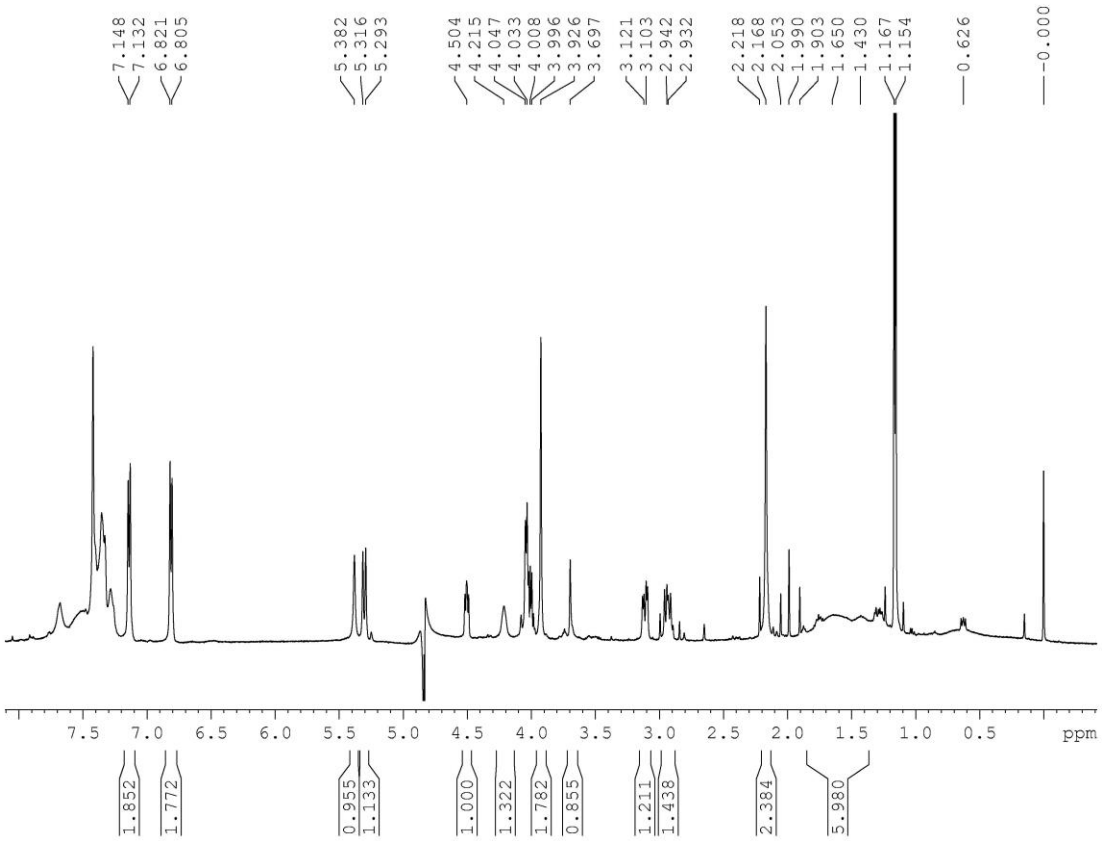


Figure 5.44. ^{1}H NMR spectrum of **A_2I-2** (0.75 mM) with Kme2GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at room temperature.

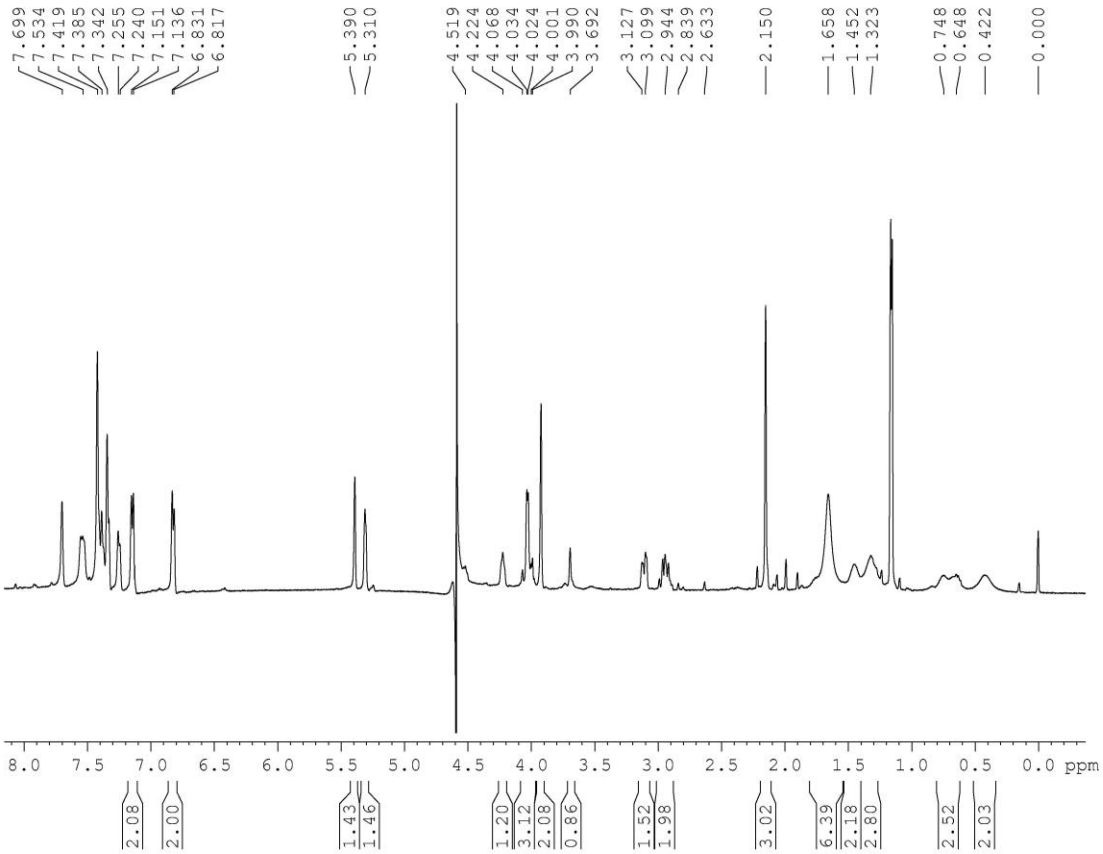


Figure 5.45. ^{1}H NMR spectrum of **A_2I-2** (0.75 mM) with Kme2GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at 40 ºC.

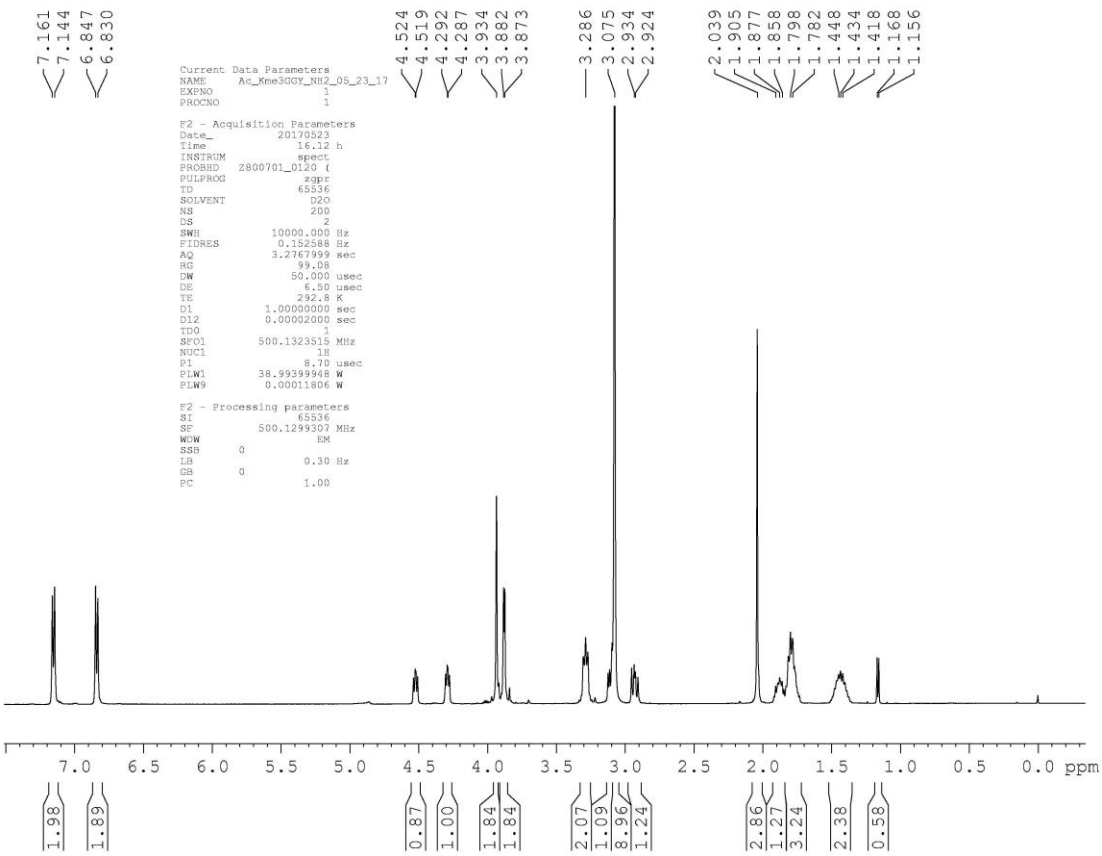


Figure 5.46. ^{1}H NMR spectrum of **Kme3GGY** (7.6 mM) alone in 50 mM sodium borate-d_3 D_2O, pH 8.5, 0.05 mM DSS (standard) at room temperature.

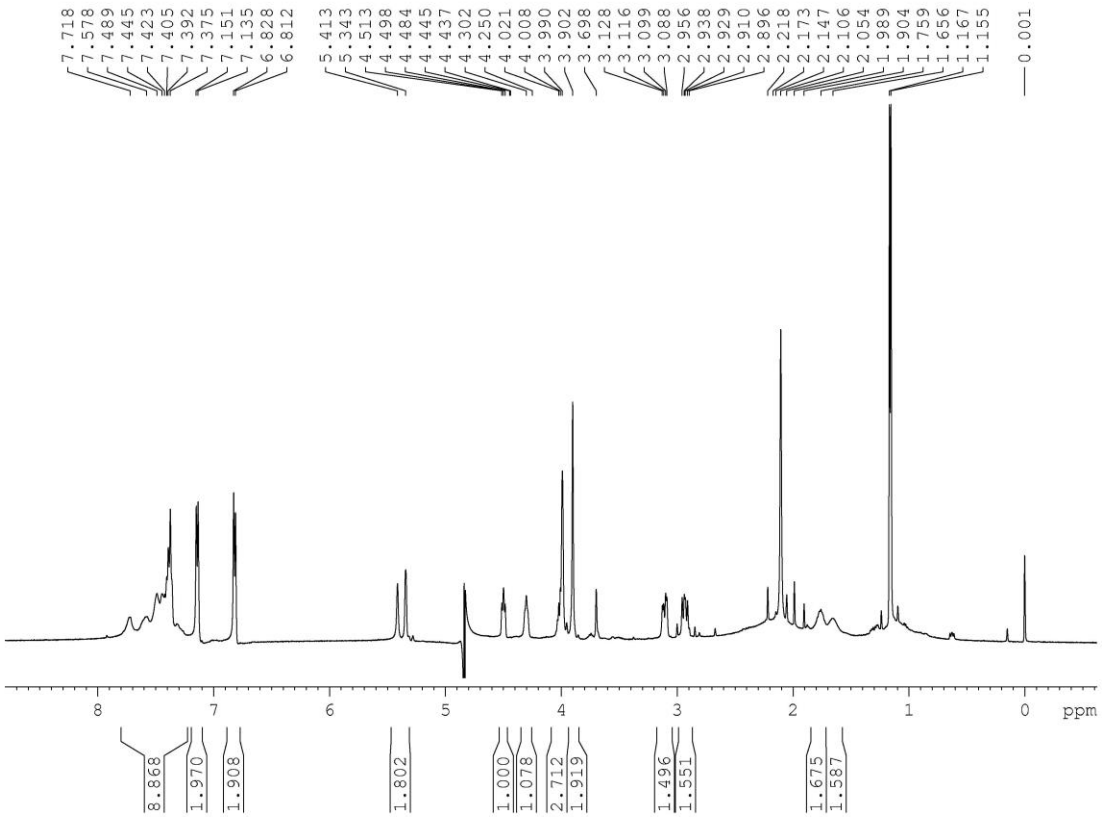


Figure 5.47. ^{1}H NMR spectrum of **A_2I-2** (0.75 mM) with Kme3GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at room temperature.

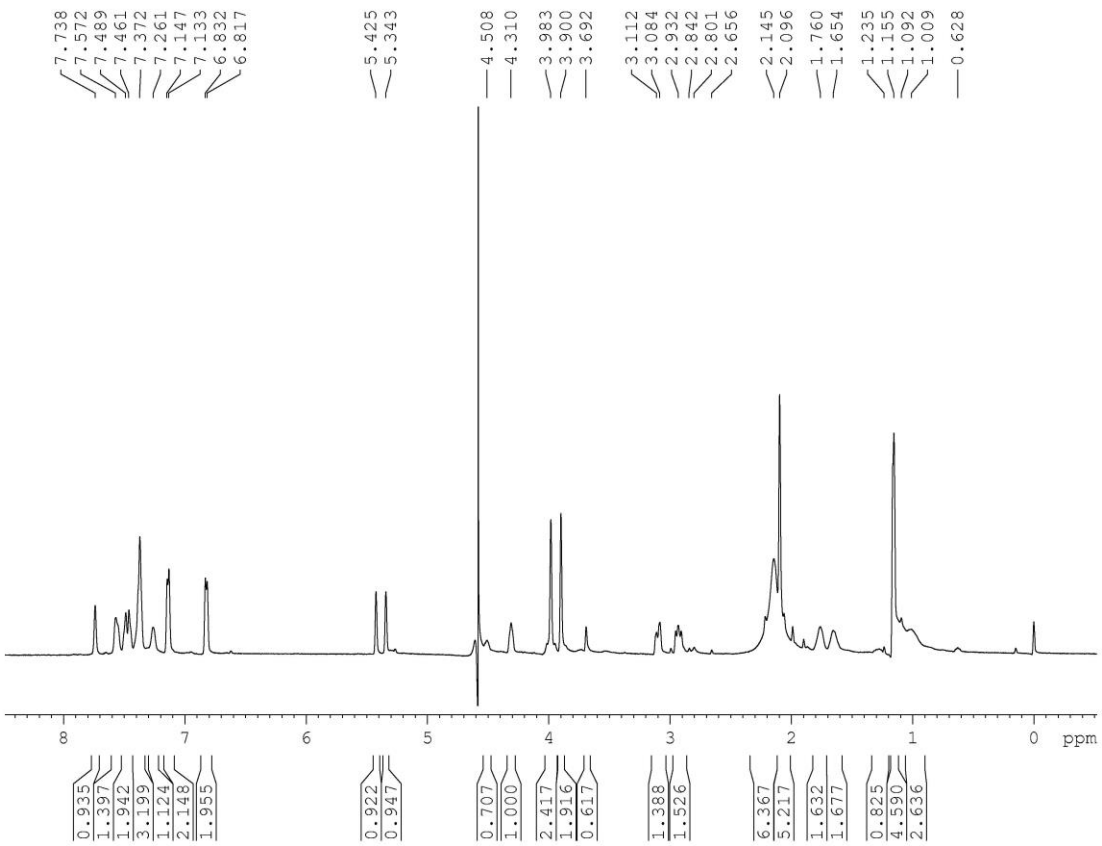


Figure 5.48. ^{1}H NMR spectrum of **A_2I-2** (0.75 mM) with Kme3GGY (0.70 mM) in 50 mM sodium borate-d_3 D_2O, pH 8.5 at 40 ºC.

Isothermal Titration Calorimetry

ITC binding experiments were conducted using a Microcal AutoITC200. Titrations were performed at 25 °C in 10 mM sodium borate buffer at pH 8.5. The concentration of receptor was determined by measuring the UV–vis absorbance at 300 nm or 315 nm, using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. Solutions of 0.2–1.3 mM of peptide were titrated into a 20–91 μM solution of receptor, using 2.0 μL increments every 3 min. Heats of dilution were taken as the last three or four points and subtracted prior to fitting. Binding curves were produced using the supplied Origin software and fit using one-site binding models.

Table 5.4. Summary of **A_2I** and **A_2J** ITC data from using a one-site binding model. Values represent an average of 3 measurements. Errors are from standard deviation from three separate runs.

Entry[a]	Receptor	Peptide	Charge	K_d[b]	SF[d]	ΔG[c]	ΔH[c]	TΔS[c]
1	A_2B	H3 K9me3	+2	2.6 ± 0.1	-	-7.63 ± 0.03		
2	A_2B	H3 K9me2	+2	6.3 ± 0.3	2.4	-7.10 ± 0.07		
3	A_2B	H3 K9me	+2	13.9 ± 0.1	5.4	-6.64 ± 0.01		
4	A_2B	H3 K9	+2	22 ± 1	8.3	-6.38 ± 0.02		
5	A_2I-2	H3 K9me3	+2	0.22 ± 0.03 [c]	-	-9.11 ± 0.09	-11.11 ± 0.05	-2.0 ± 0.07
6	A_2I-2	H3 K9me2	+2	0.20 ± 0.01	0.9	-9.11 ± 0.09	-11.11 ± 0.05	-2.0 ± 0.02
7	A_2I-2	H3 K9me	+2	1.00 ± 0.04	4.5	-8.18 ± 0.04	-10.18 ± 0.04	-2.0 ± 0.01
8	A_2I-2	H3 K9	+2	1.79 ± 0.1 [c]	8	-7.88 ± 0.1	-9.88 ± 0.04	-2.0 ± 0.1
9	A_2I-1	H3 K9me3	+2	0.13 ± 0.005	-	-9.39 ± 0.04	-12.29 ± 0.04	-2.9 ± 0.02
10	A_2I-1	H3 K9me2	+2	0.18 ± 0.04	1.4	-9.19 ± 0.2	-11.39 ± 0.08	-2.2 ± 0.2
11	A_2I-1	H3 K9me	+2	1.23 ± 0.04 [c]	9.5	-8.03 ± 0.06	-8.83 ± 0.03	-0.8 ± 0.05
12	A_2I-1	H3 K9	+2	2.09 ± 0.15 [c]	16	-7.78 ± 0.2	-8.48 ± 0.06	-0.7 ± 0.2
13	A_2I-2	Kme3GGY	+1	4.30 ± 0.33	-	-7.32 ± 0.1	-8.72 ± 0.1	-1.40 ± 0.05
14	A_2I-2	Kme2GGY	+1	3.32 ± 0.20	0.8	-7.48 ± 0.1	-8.96 ± 0.1	-1.48 ± 0.08
15	A_2J-1	H3 K9me3	+2	2.61 ± 0.23	-	-7.62 ± 0.8	-	-
16	A_2J-1	H3 K9me2	+2	5.21 ± 0.44	2.0	-7.21 ± 0.7	-	-
17	A_2J-1	H3 K9me	+2	6.36 ± 0.64	2.4	-7.09 ± 0.7	-	-
18	A_2J-1	H3 K9	+2	7.02 ± 0.75	2.7	-7.03 ± 0.7	-	-
19	A_2J-2	H3 K9me3	+2	2.25 ± 0.18	-	-7.70 ± 0.7	-	-
20	A_2J-2	H3 K9me2	+2	6.74 ± 0.62	3.0	-7.05 ± 0.7	-	-
21	A_2J-2	H3 K9me	+2	8.37 ± 1.0	3.7	-6.93 ± 0.7	-	-
22	A_2J-2	H3 K9	+2	10.60 ± 1.1	4.7	-6.79 ± 0.7	-	-

[a] Conditions: 25 °C in 10 mM sodium borate buffer, pH 8.5. [b] Errors are averages of the error in the fit as determined by the Origin 7 software. [c] Errors are from averages of the runs. [d] The selectivity factor is calculated as the K_d of Kmex (where x = 0-2) divided by the K_d of Kme3.

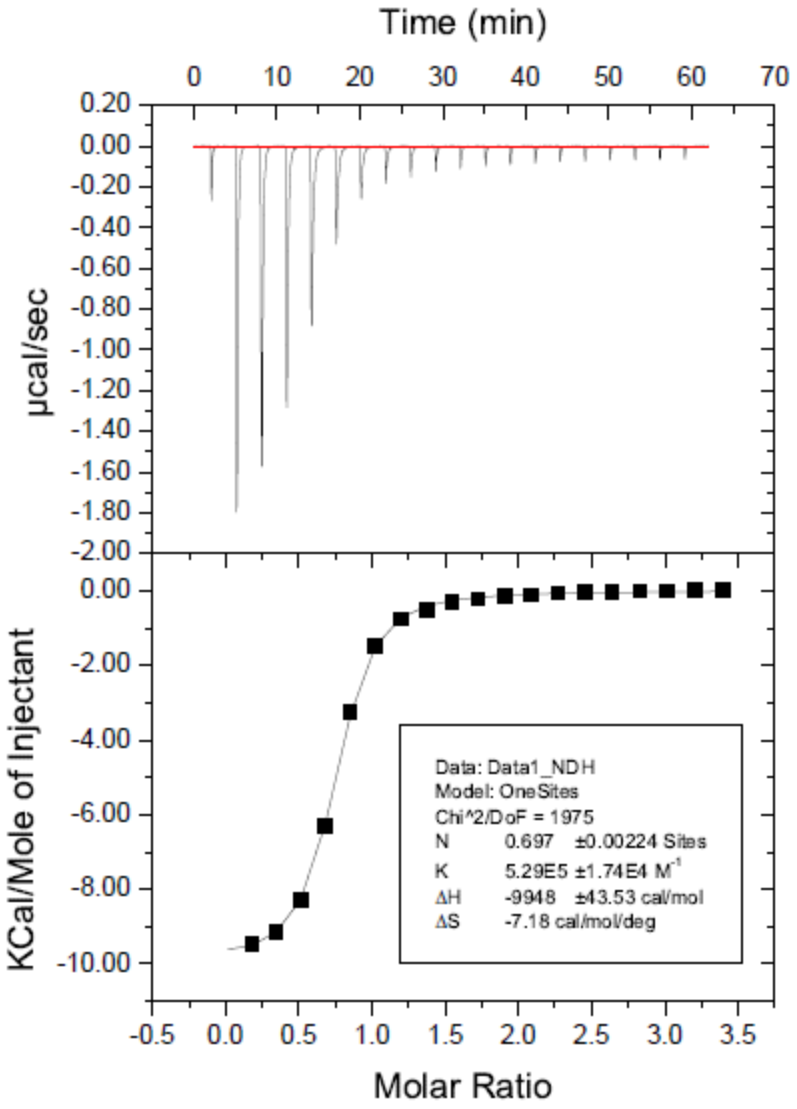


Figure 5.49. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into **A_2I-2** (79 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 1).

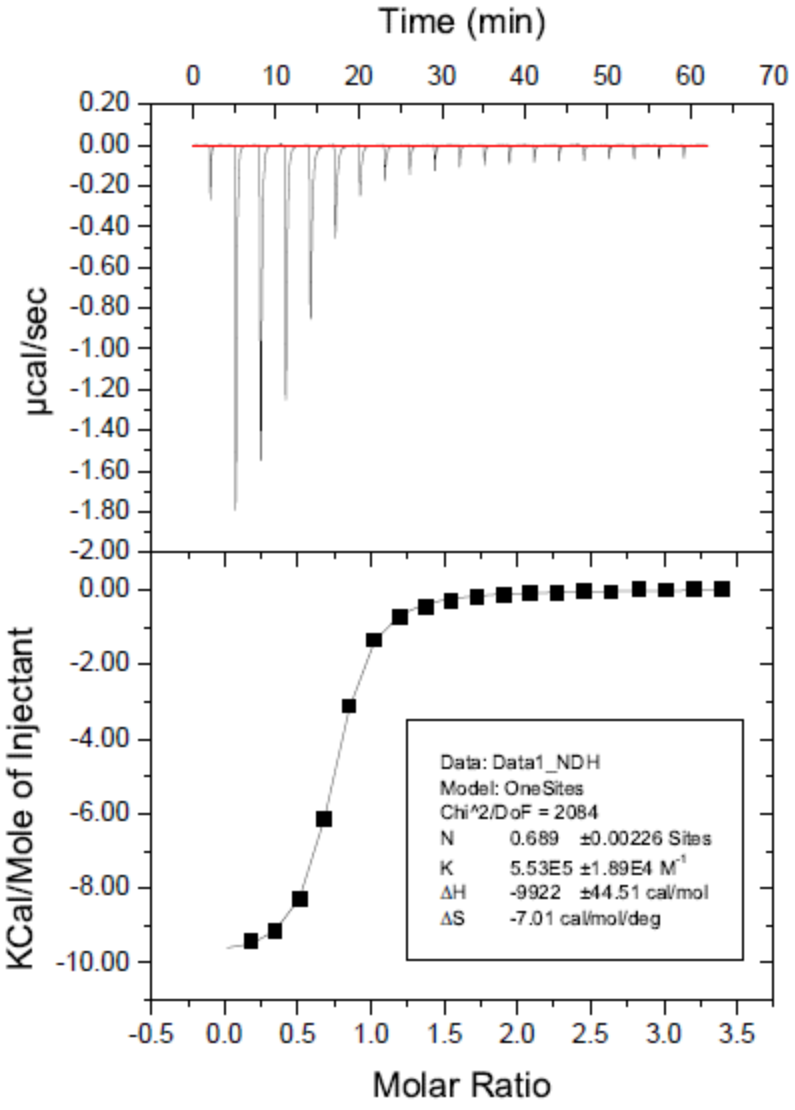


Figure 5.50. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into **A_2I-2** (79 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 2).

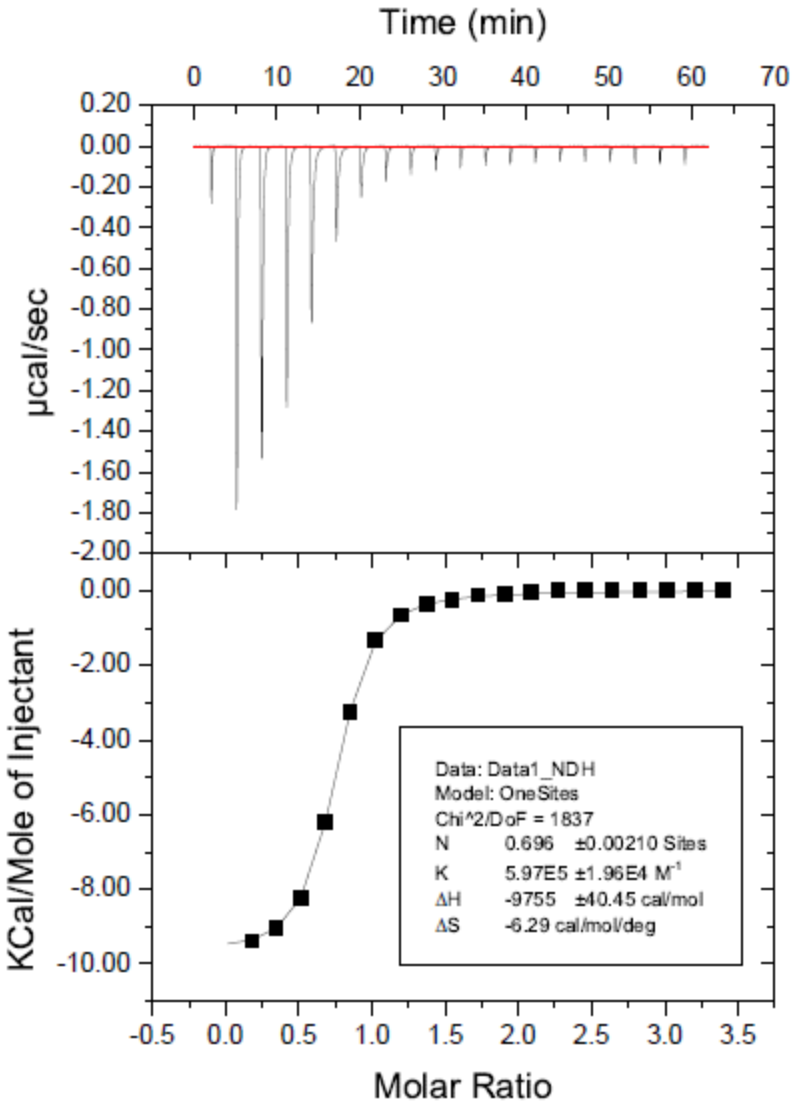


Figure 5.51. One trial of three of H3K9 (Ac-WGGGQTARKSTG-NH_2) (1.33 mM) titrated into **A_2I-2** (79 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 3).

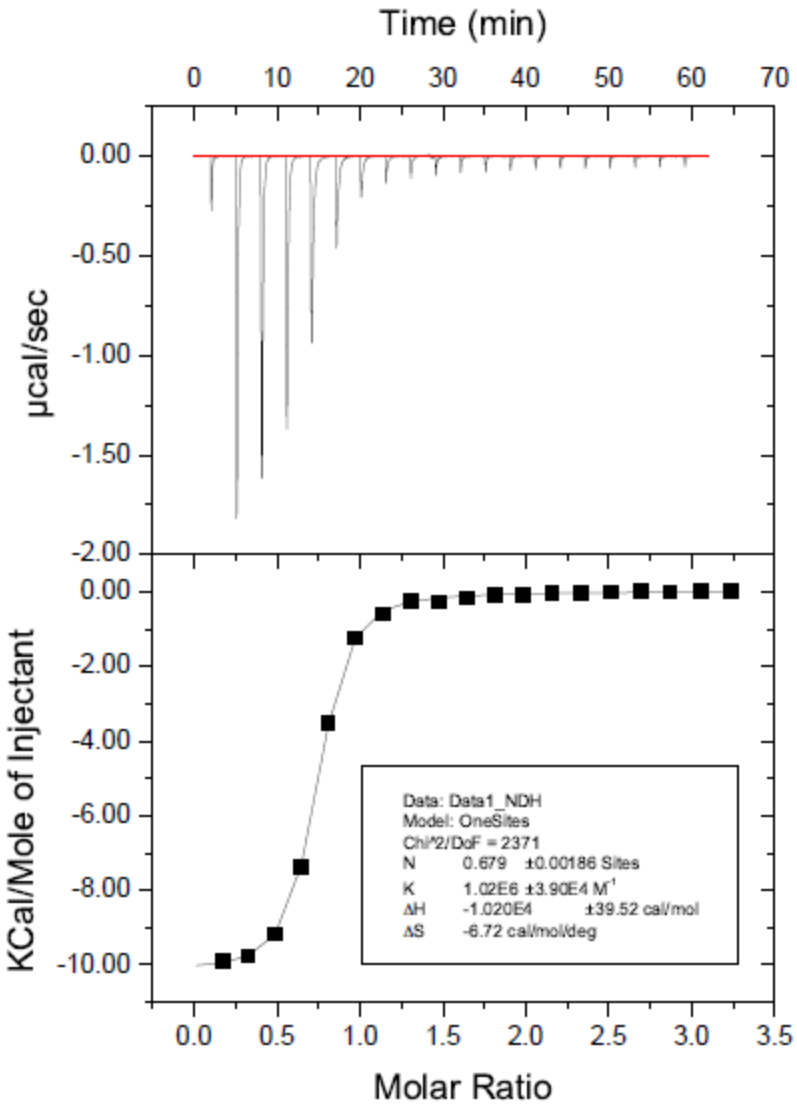


Figure 5.52. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into **A_2I-2** (79 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 1).

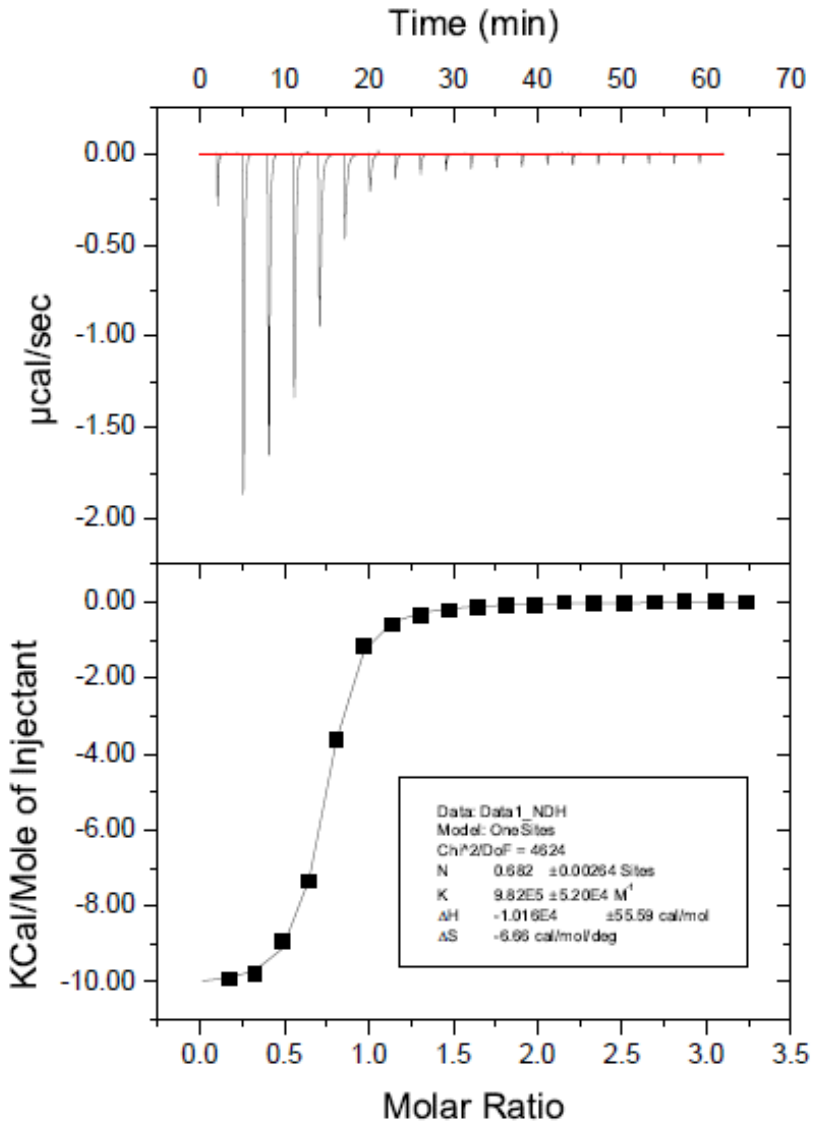


Figure 5.53. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into **A_2I-2** (79 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5 (Run 2).

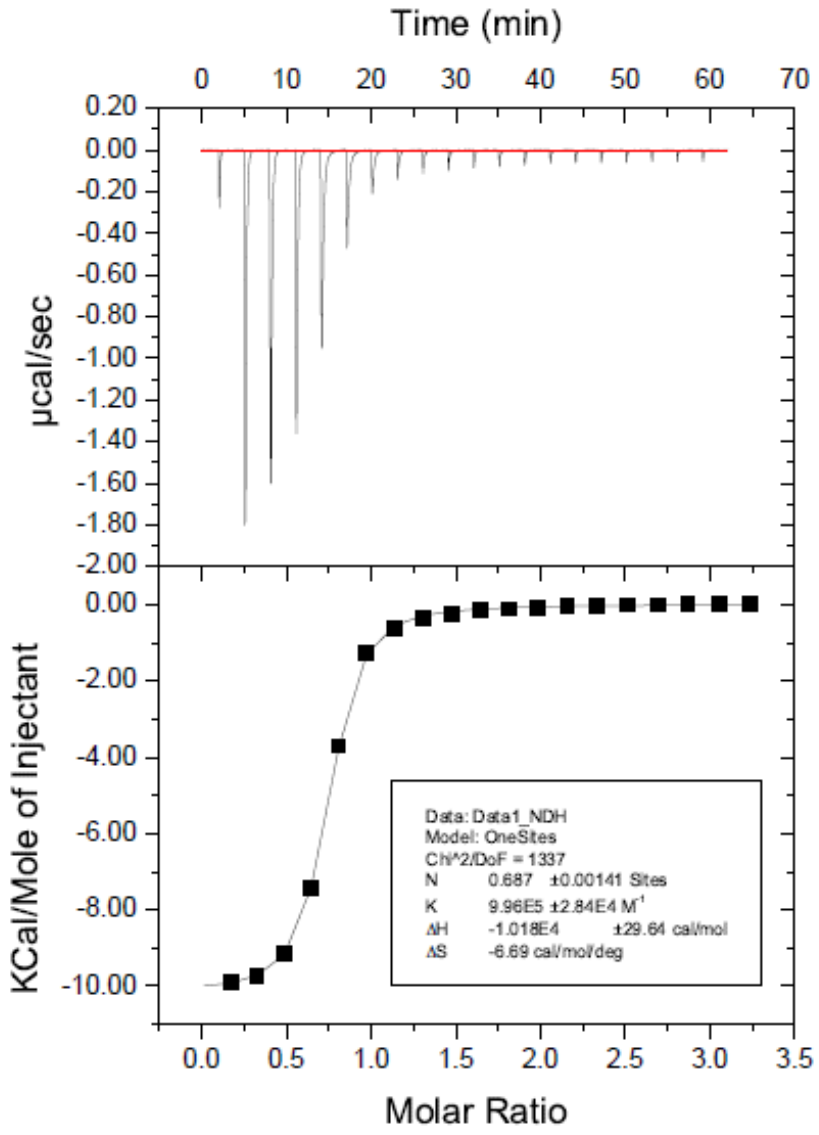


Figure 5.54. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (1.26 mM) titrated into **A_2I-2** (79 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 3).

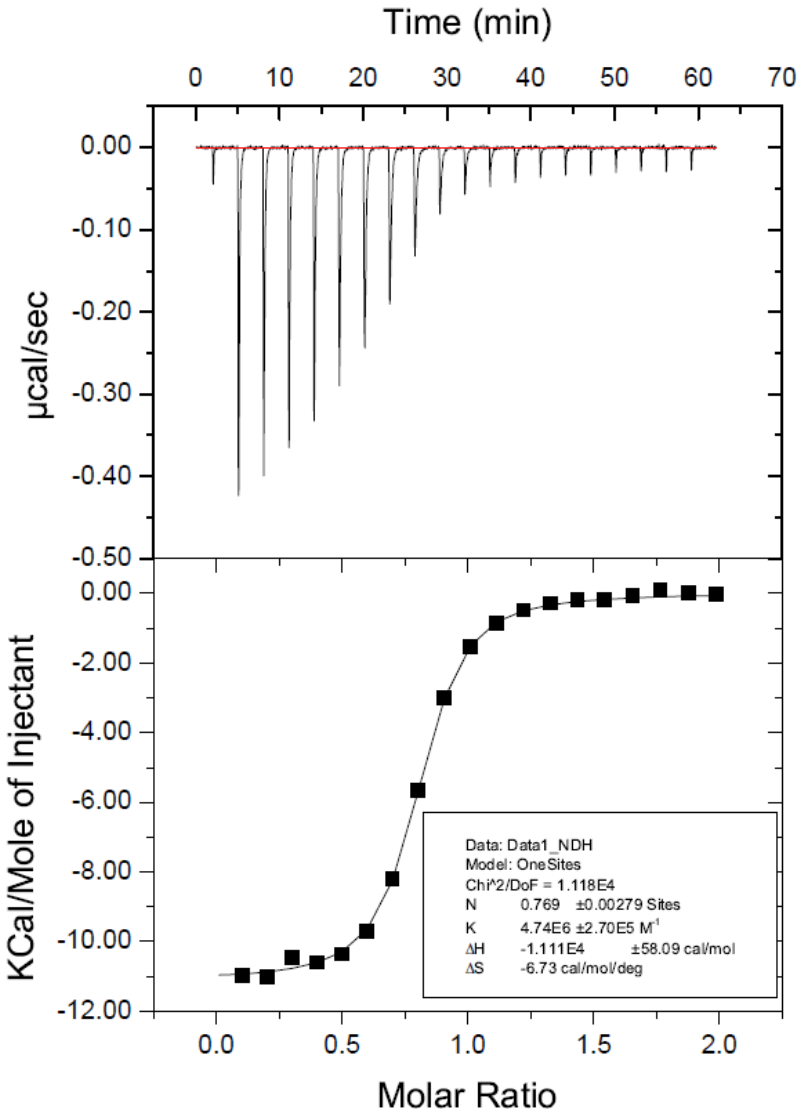


Figure 5.55. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 µM) titrated into **A_2I-2** (20 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 1).

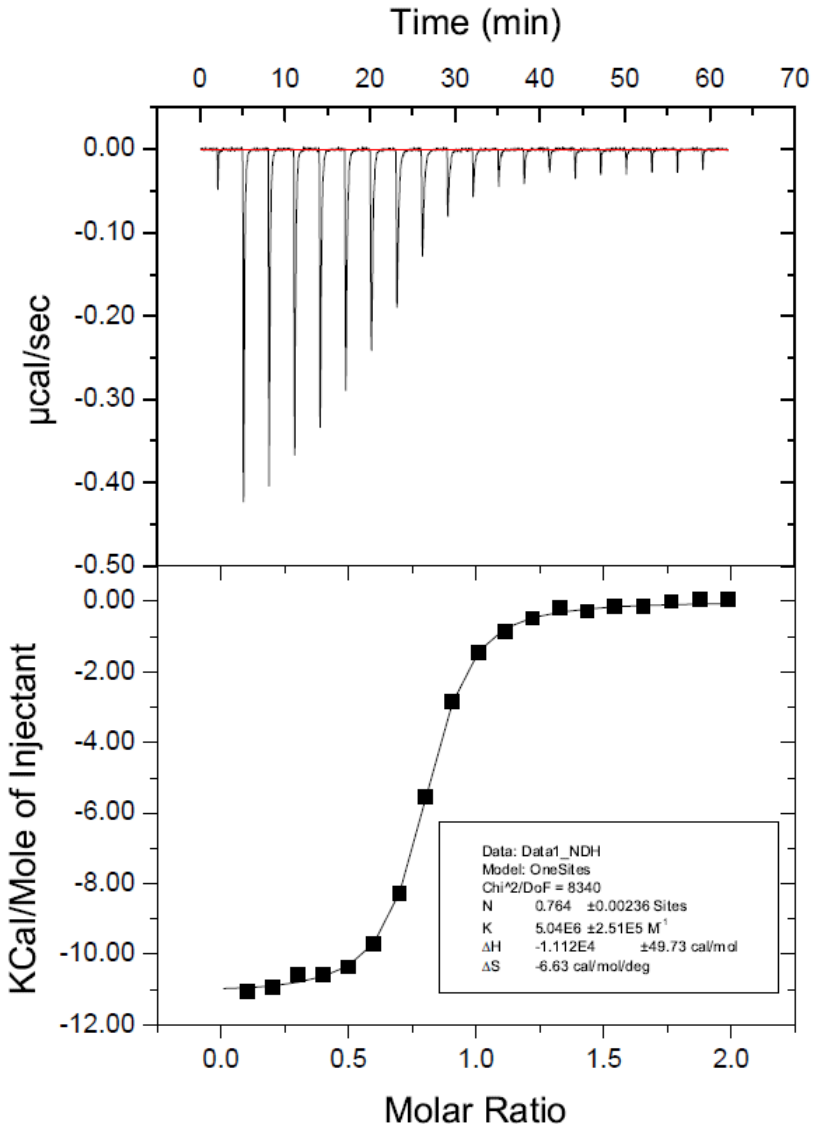


Figure 5.56. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 µM) titrated into **A_2I-2** (20 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 2).

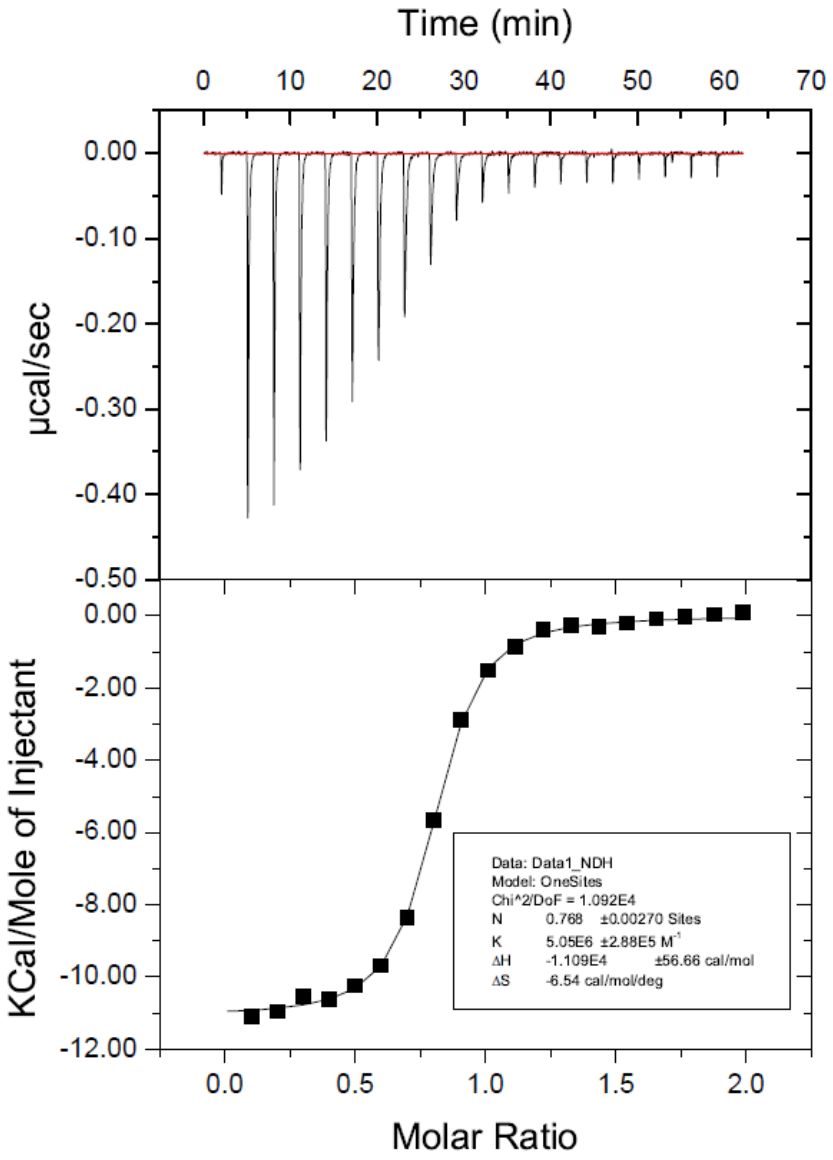


Figure 5.57. One trial of three of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (197 µM) titrated into **A_2I-2** (20 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 3).

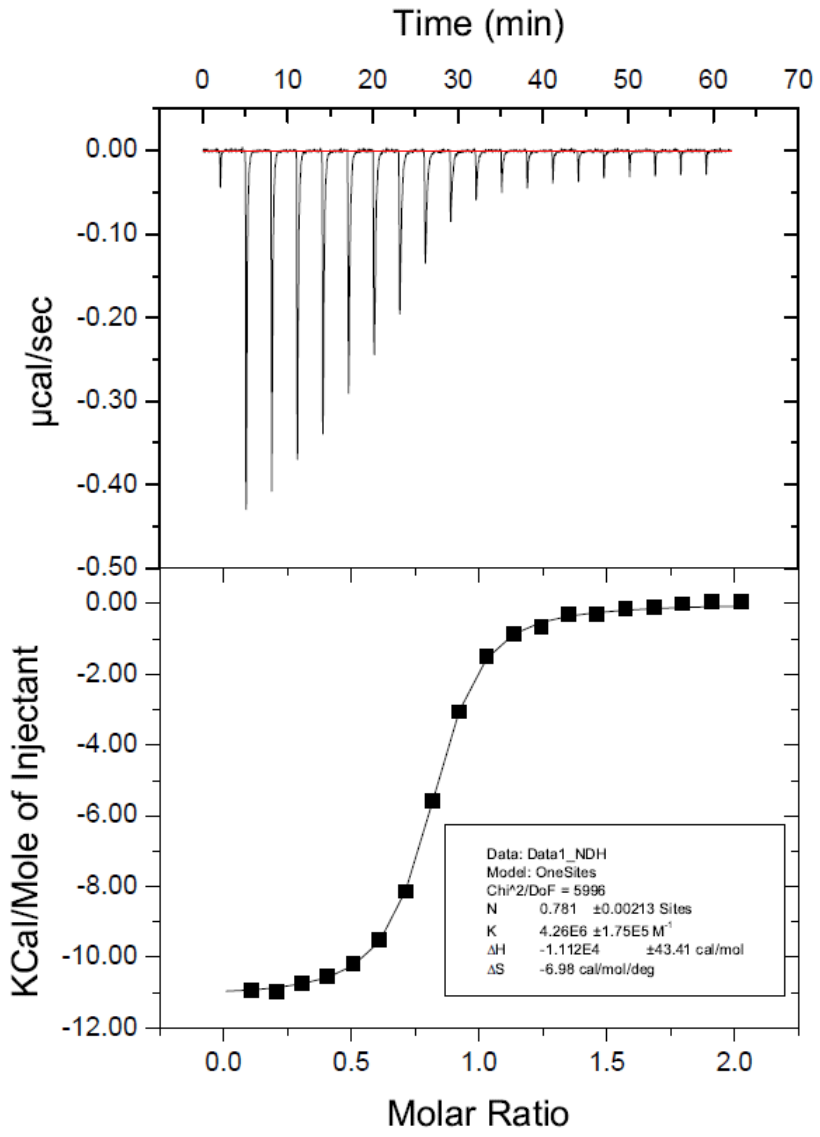


Figure 5.58. One trial of three of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 µM) titrated into **A_2I-2** (20 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 1).

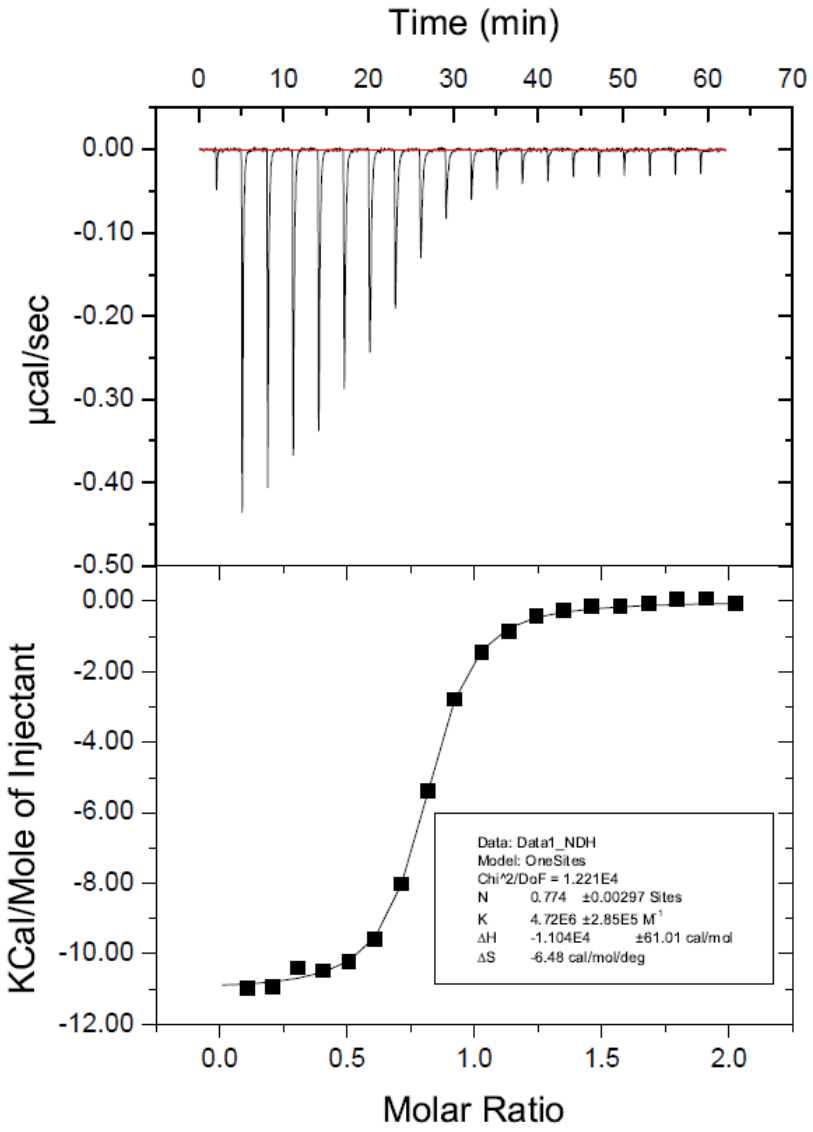


Figure 5.59. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 µM) titrated into **A_2I-2** (20 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 2).

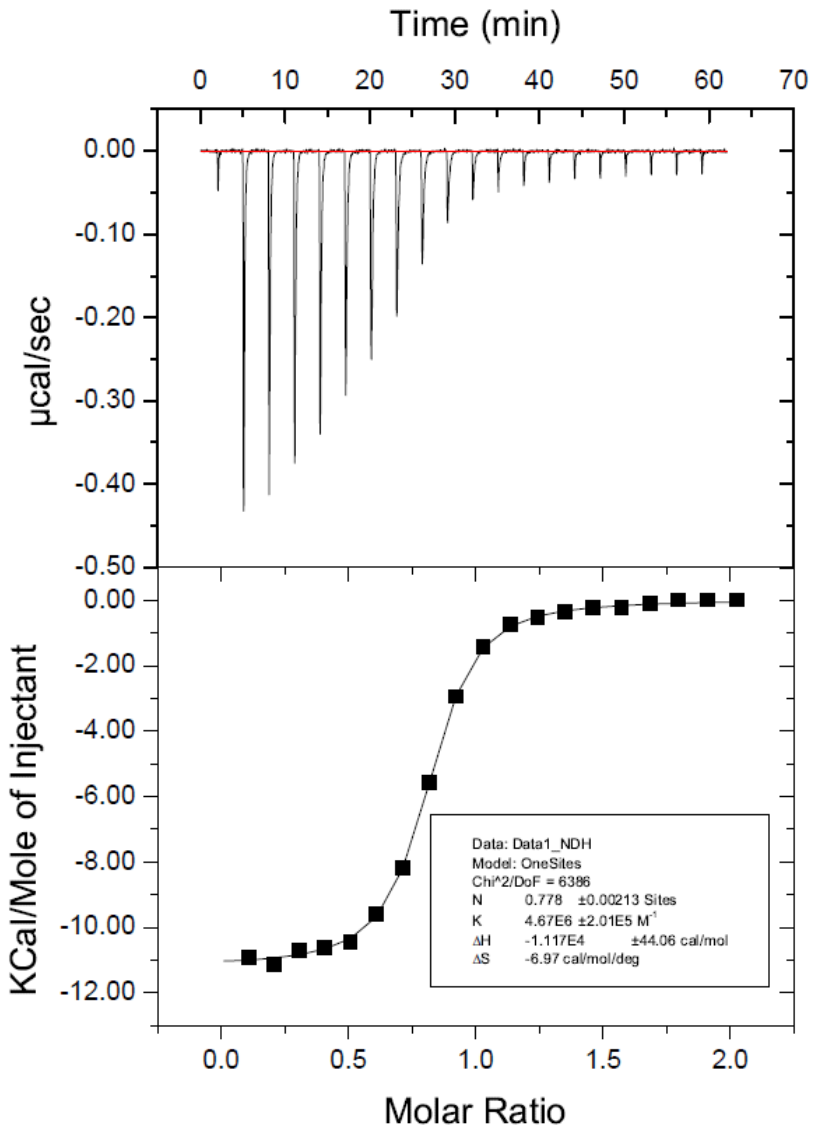


Figure 5.60. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 µM) titrated into **A_2I-2** (20 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 3).

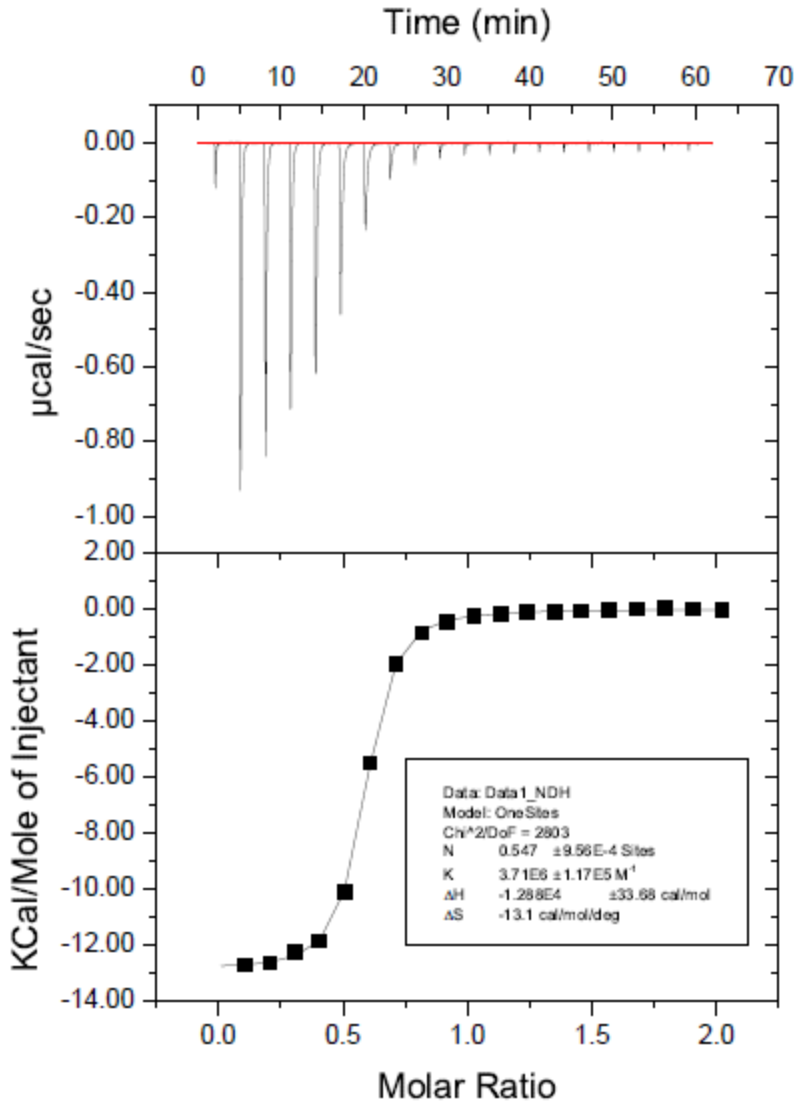


Figure 5.61. One trial of four of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (225 µM) titrated into **A_2I-2** (23 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 4).

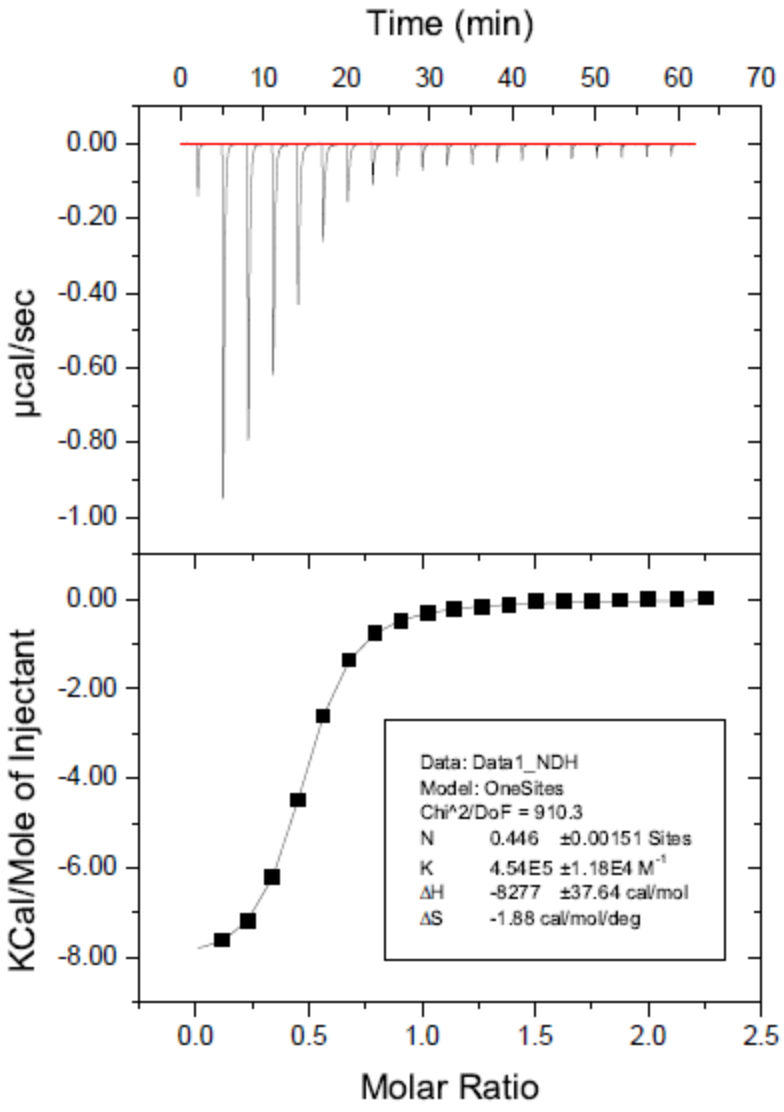


Figure 5.62. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 µM) titrated into **A_2I-1** (81 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 1).

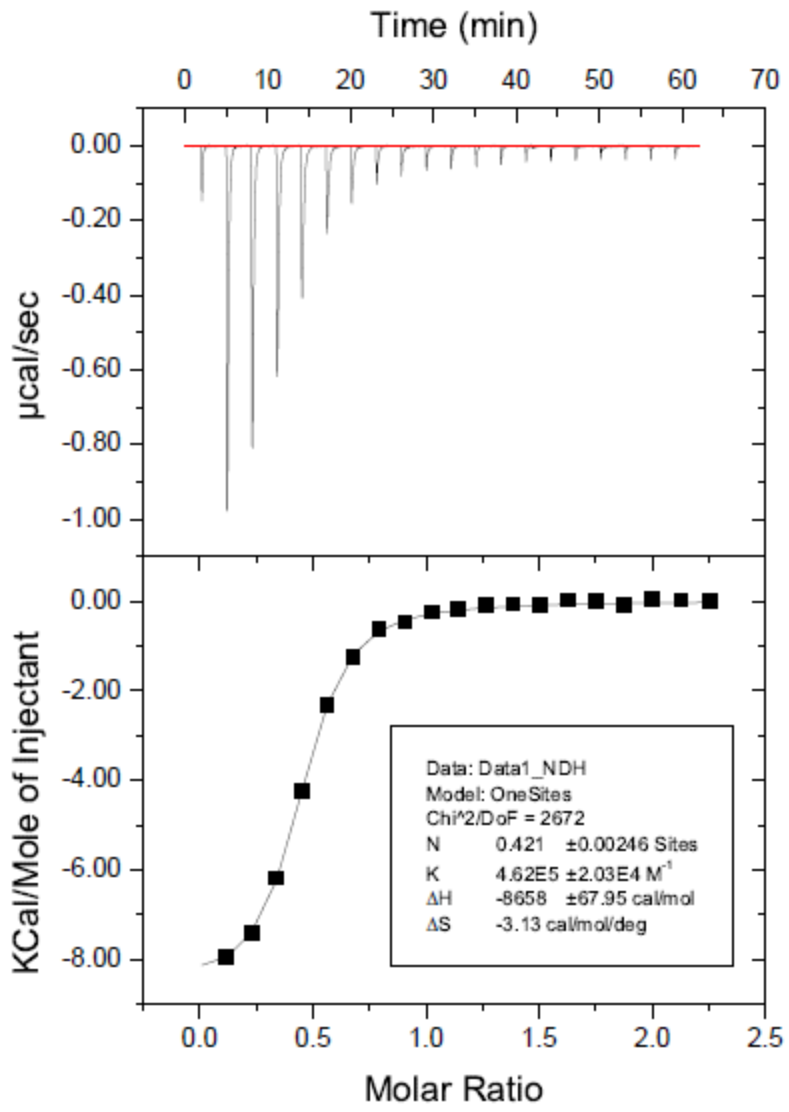


Figure 5.63. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 µM) titrated into **A_2I-1** (81 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 2).

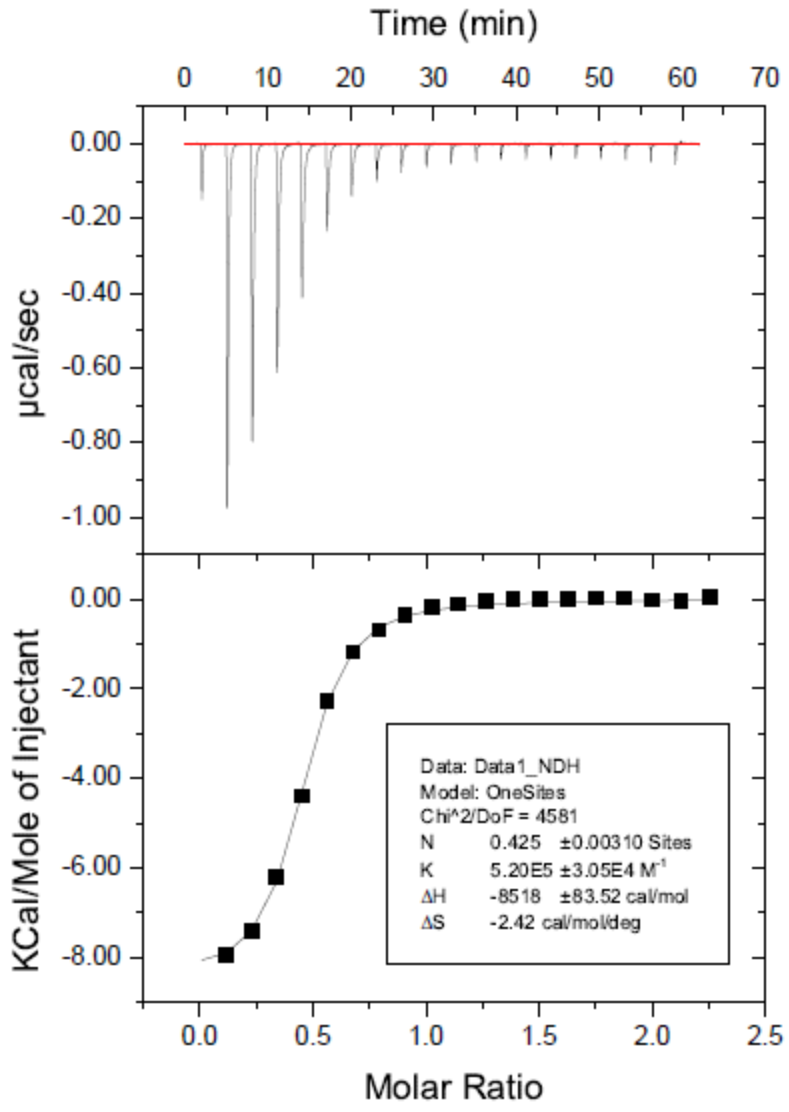


Figure 5.64. One trial of four of H3K9 (Ac-WGGGQTARKSTG-NH_2) (900 µM) titrated into **A_2I-1** (81 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 3).

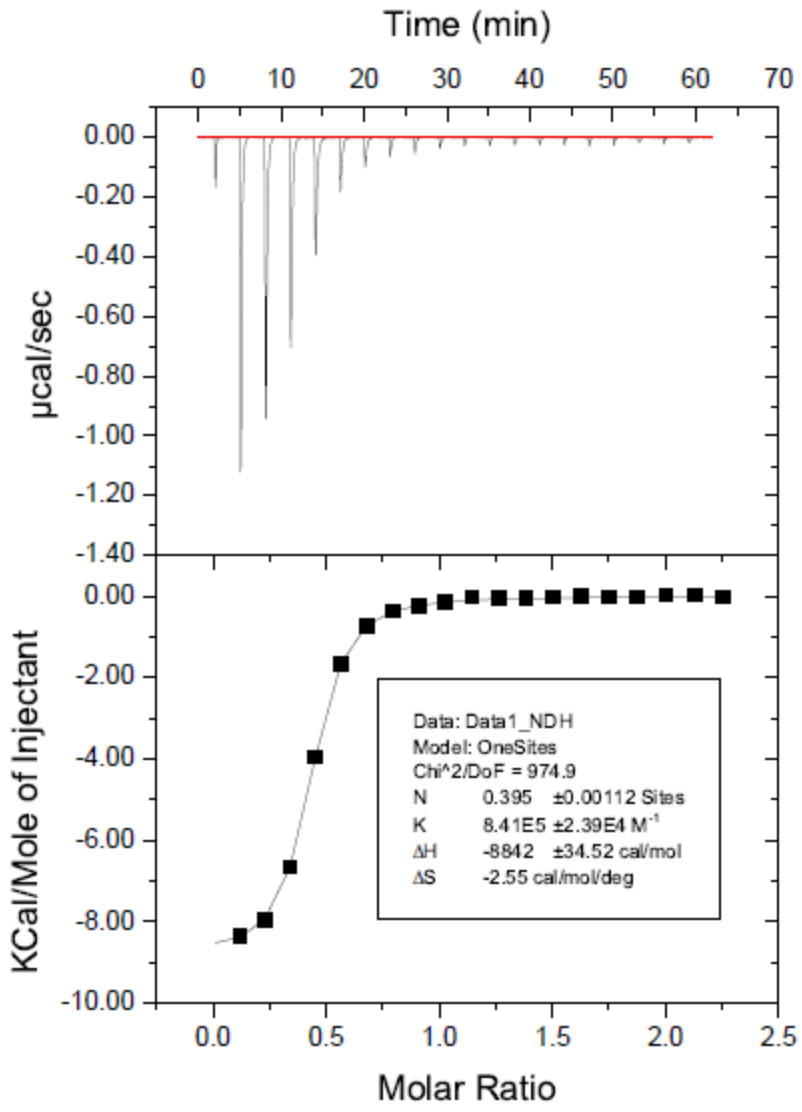


Figure 5.65. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (900 µM) titrated into **A_2I-1** (81 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 1).

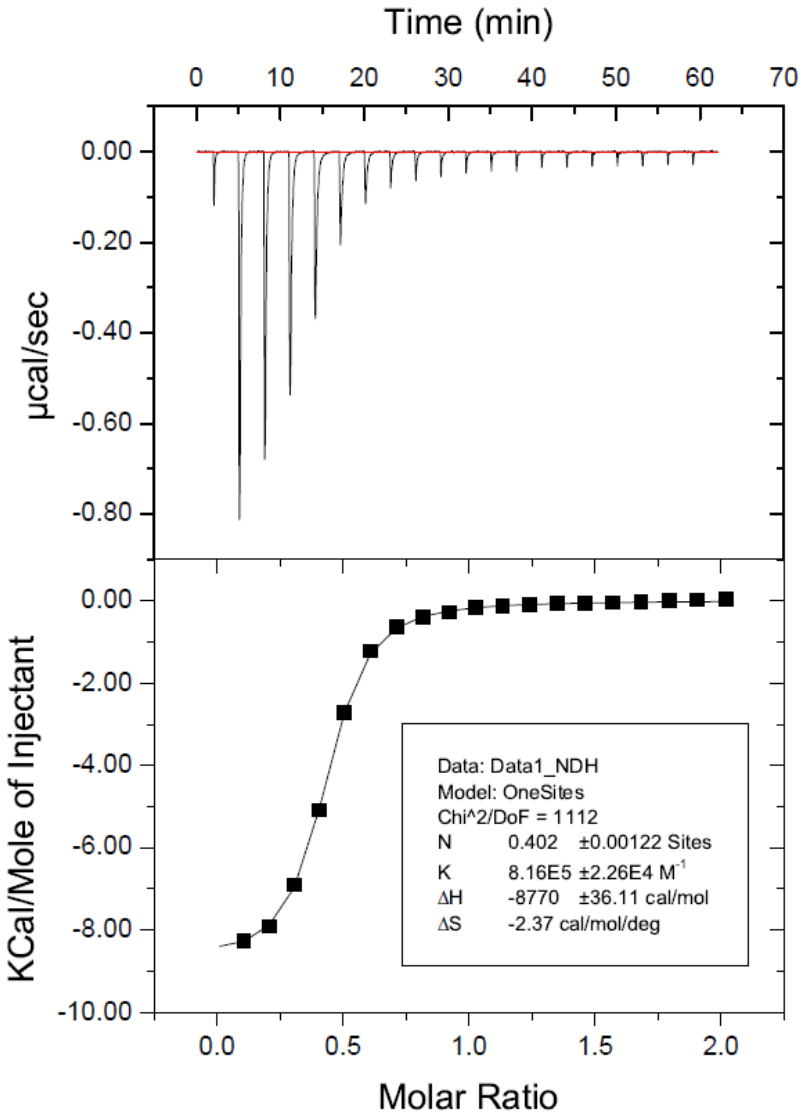


Figure 5.66. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (700 µM) titrated into **A_2I-1** (70 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 2).

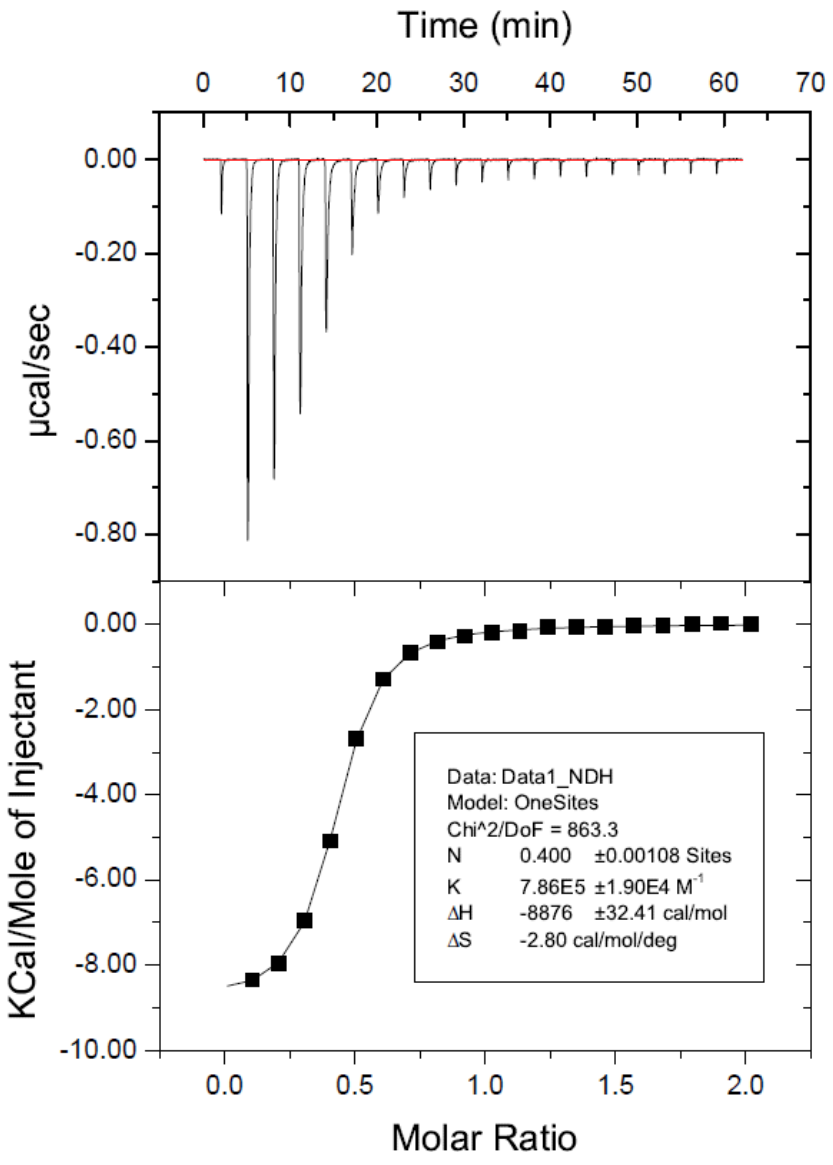


Figure 5.67. One trial of three of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (700 µM) titrated into **A_2I-1** (70 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5 (Run 3).

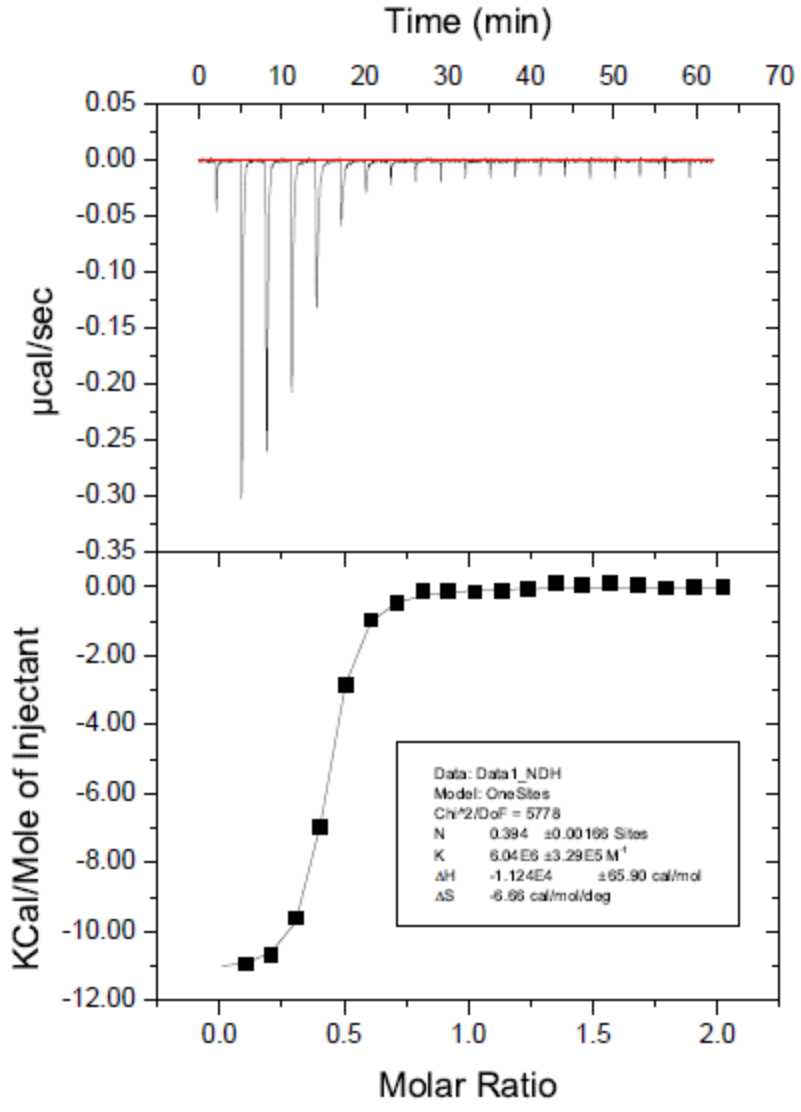


Figure 5.68. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (200 µM) titrated into **A_2I-1** (20 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5 (Run 1).

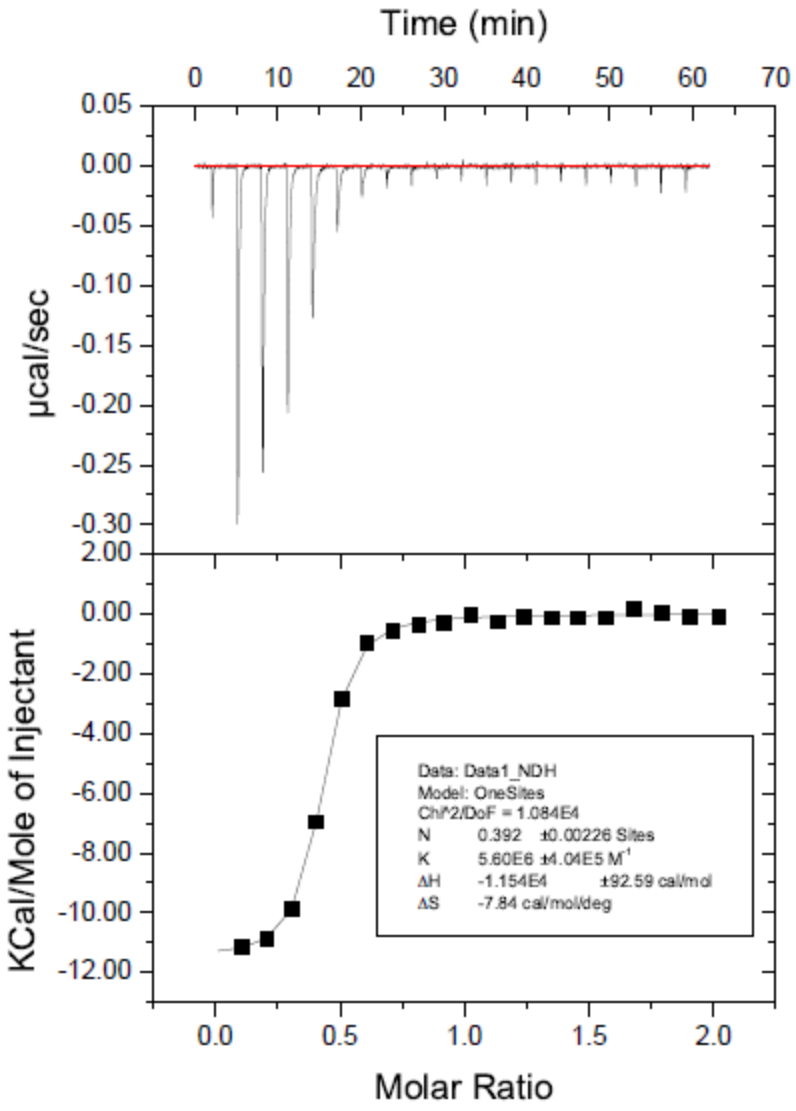


Figure 5.69. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (200 µM) titrated into **A_2I-1** (20 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5 (Run 2).

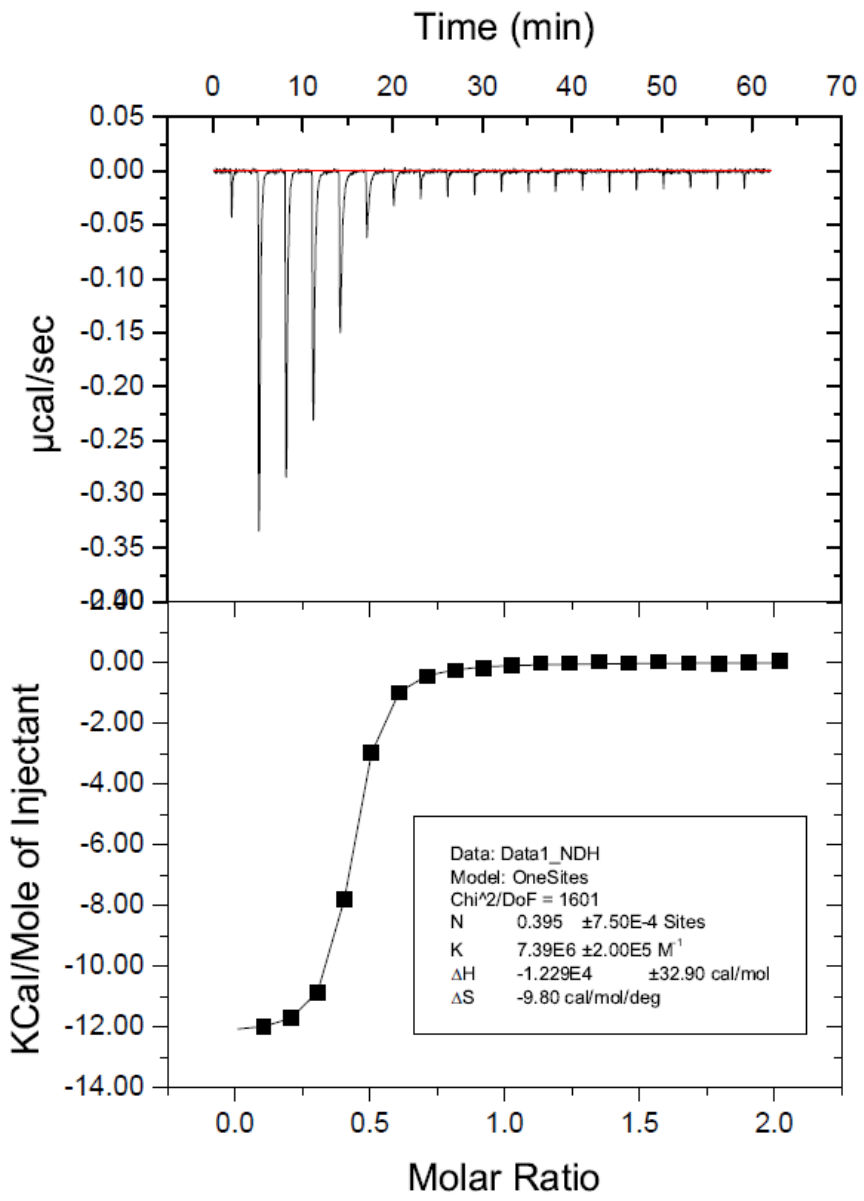


Figure 5.70. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 µM) titrated into **A_2I-1** (20 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1).

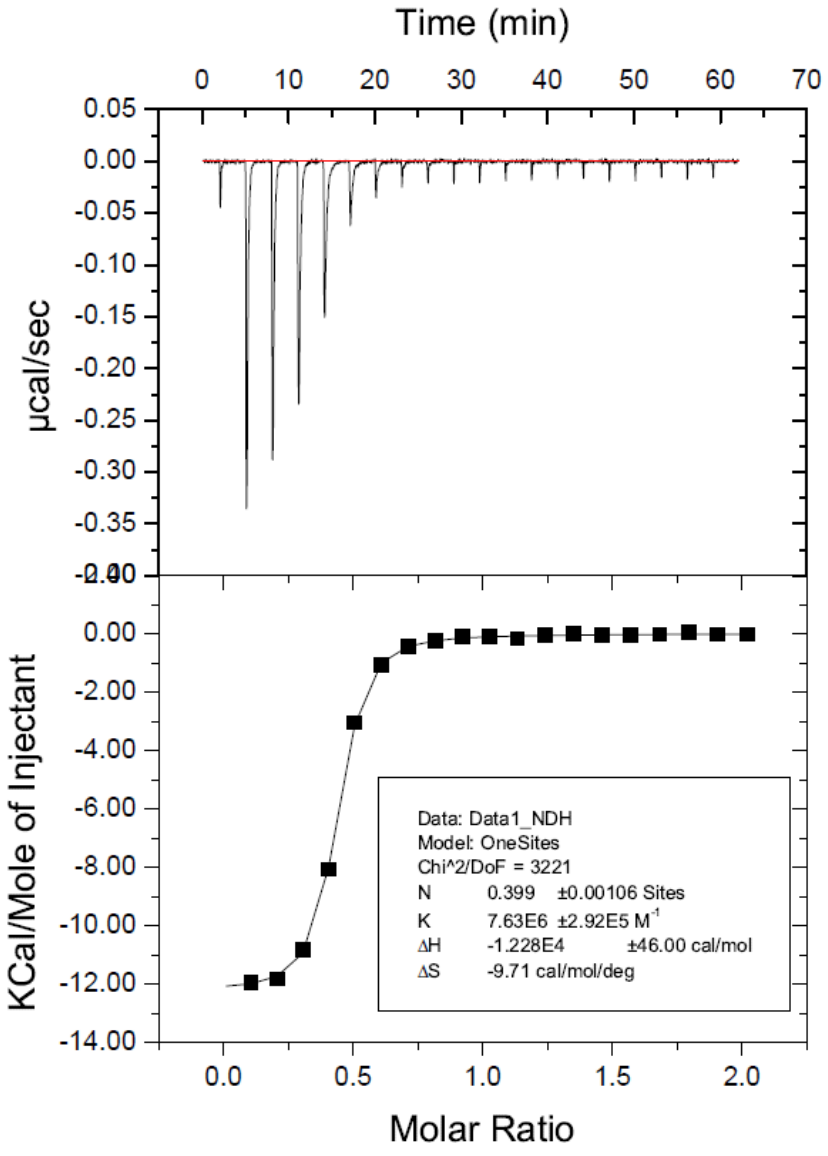


Figure 5.71. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (200 µM) titrated into **A_2I-1** (20 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2).

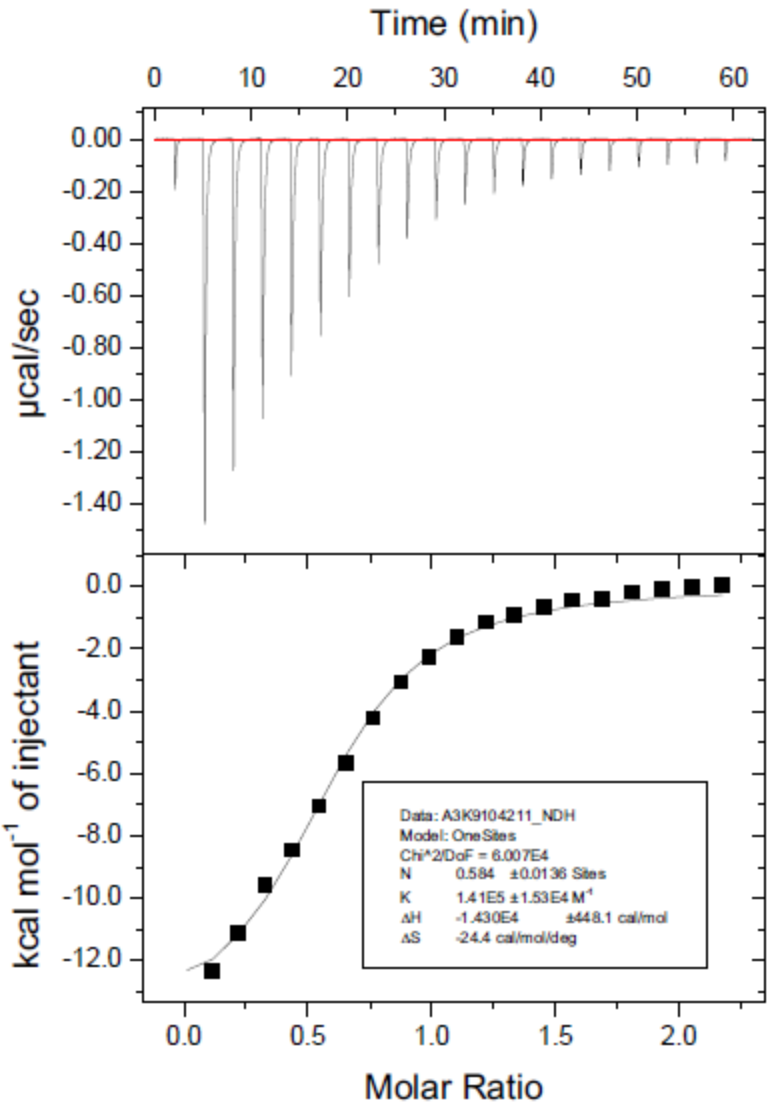


Figure 5.72. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (817 µM) titrated into **A_2J-1** (76 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1).

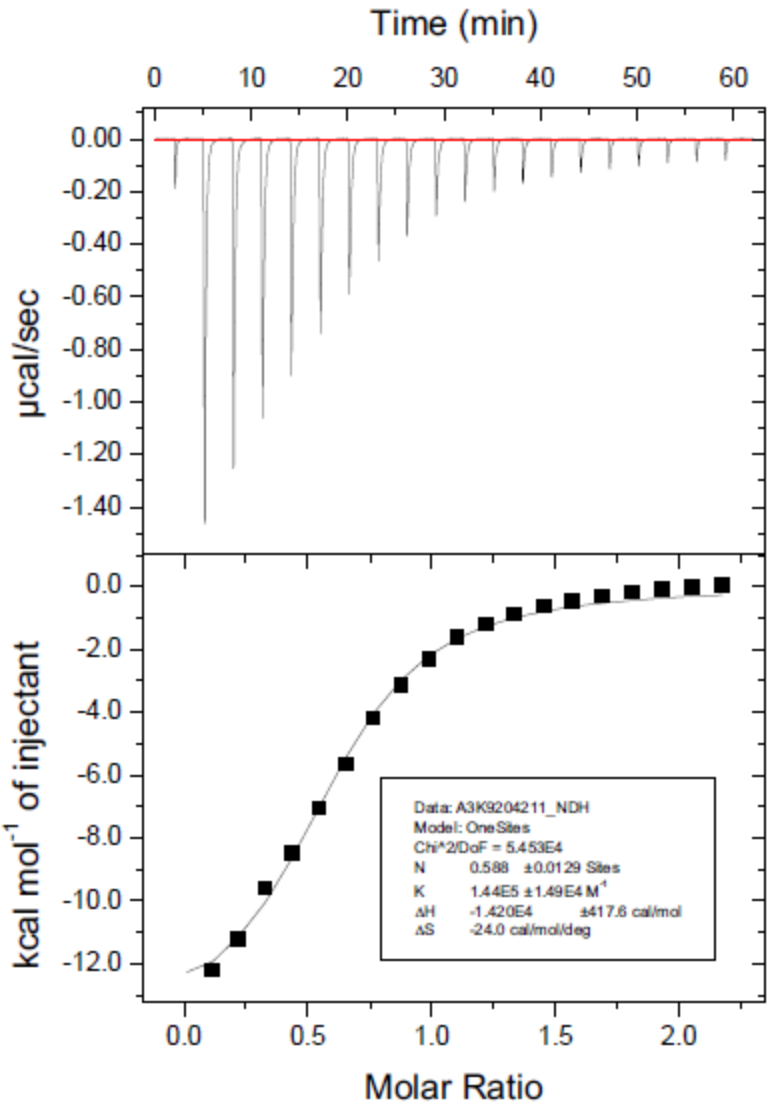


Figure 5.73. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (817 µM) titrated into **A_2J-1** (76 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2).

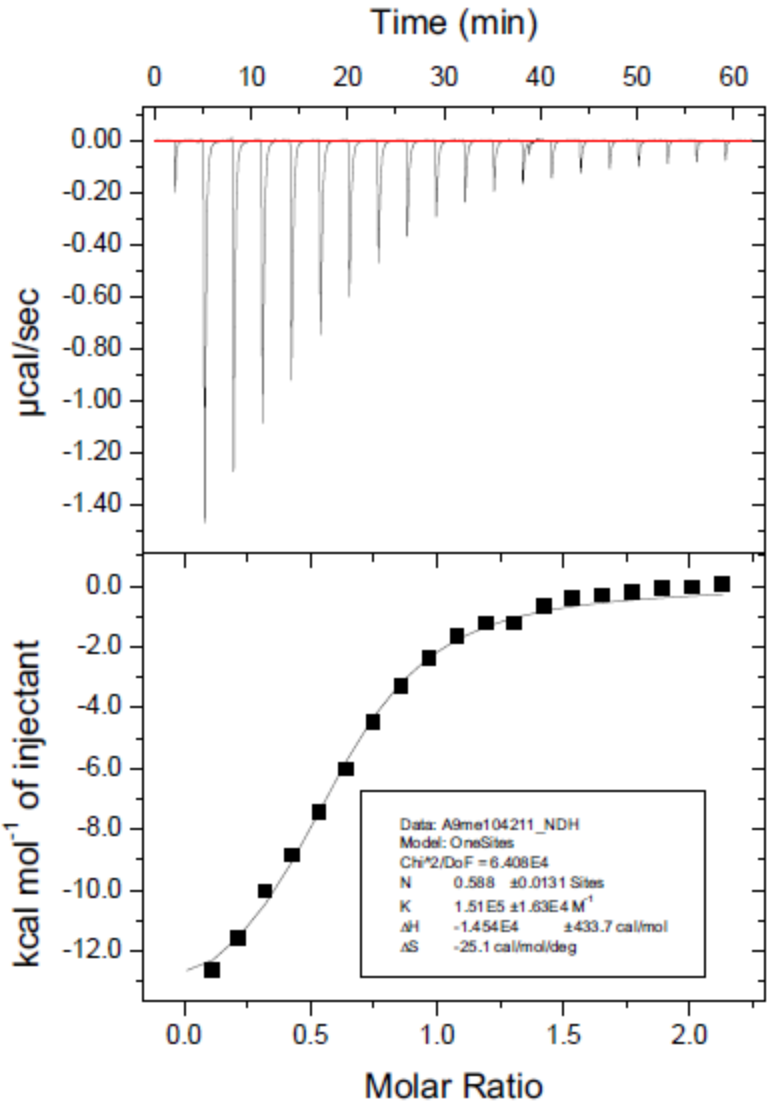


Figure 5.74. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (800 µM) titrated into **A_2J-1** (76 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1).

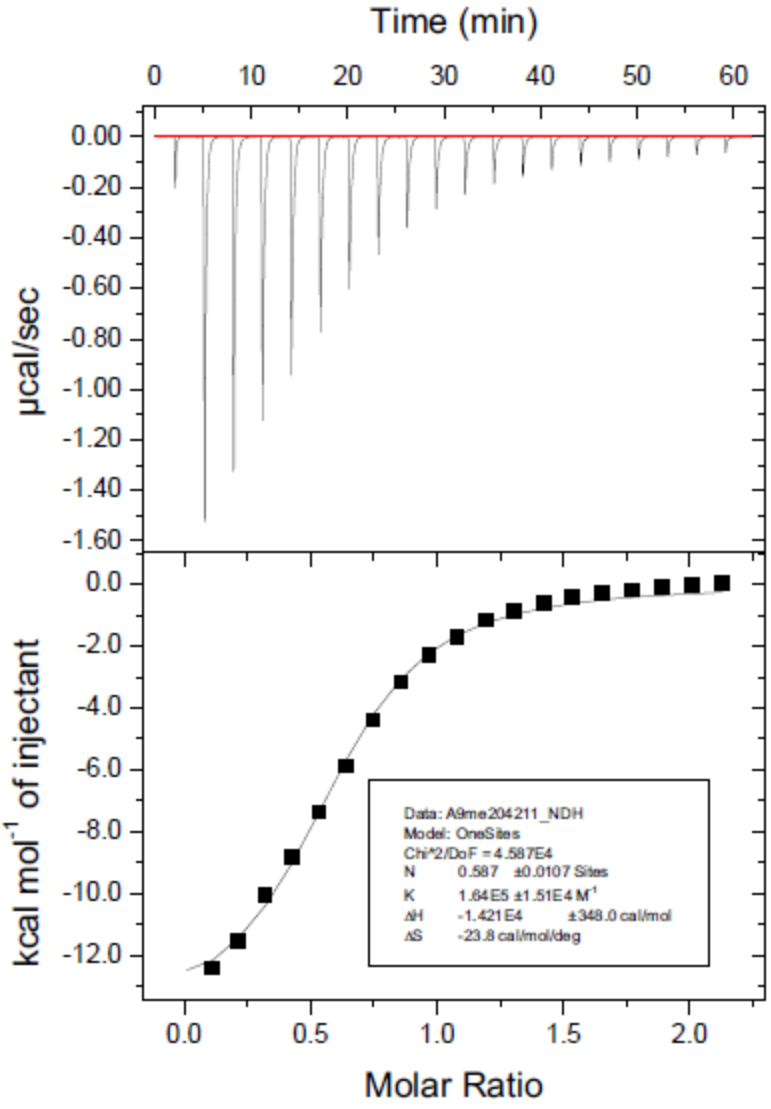


Figure 5.75. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (800 µM) titrated into **A_2J-1** (76 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2).

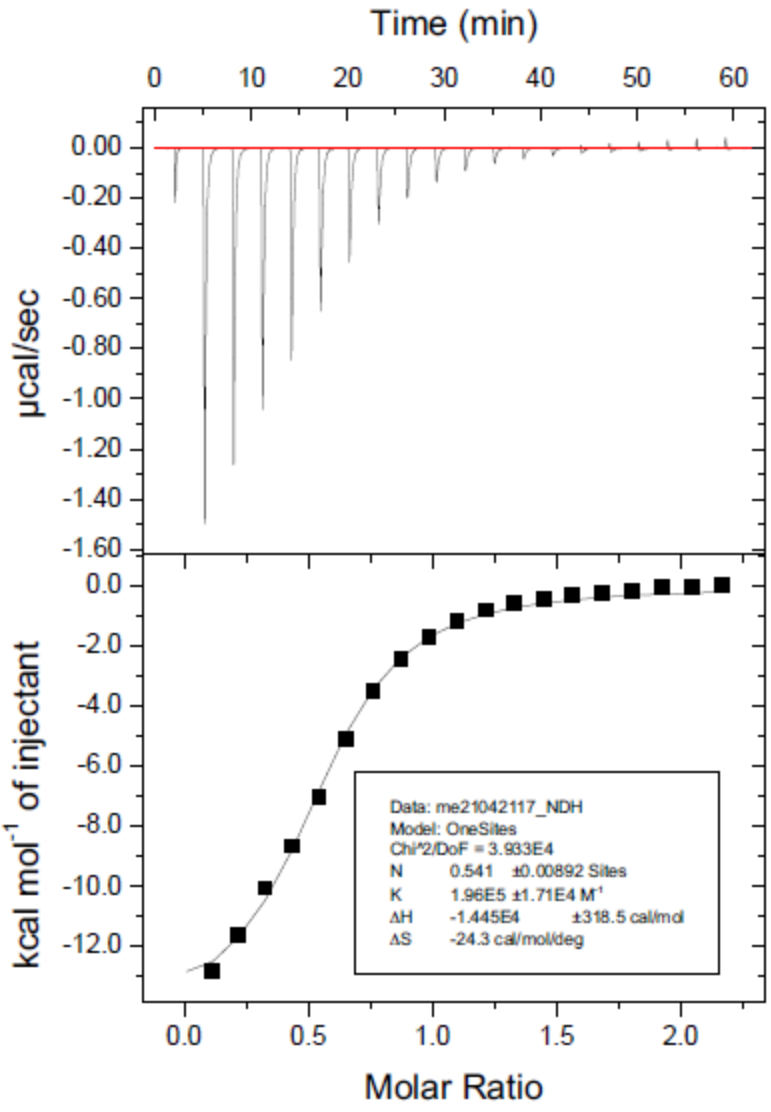


Figure 5.76. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (815 µM) titrated into **A_2J-1** (76 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 1).

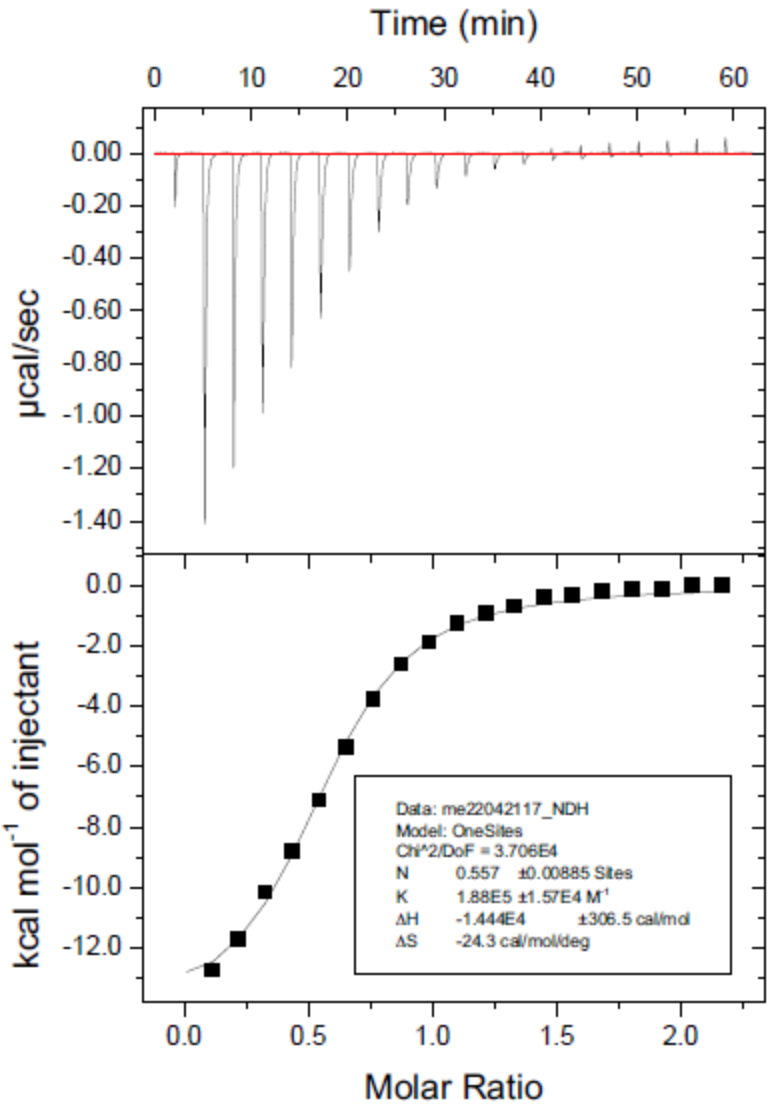


Figure 5.77. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (815 µM) titrated into **A_2J-1** (76 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2).

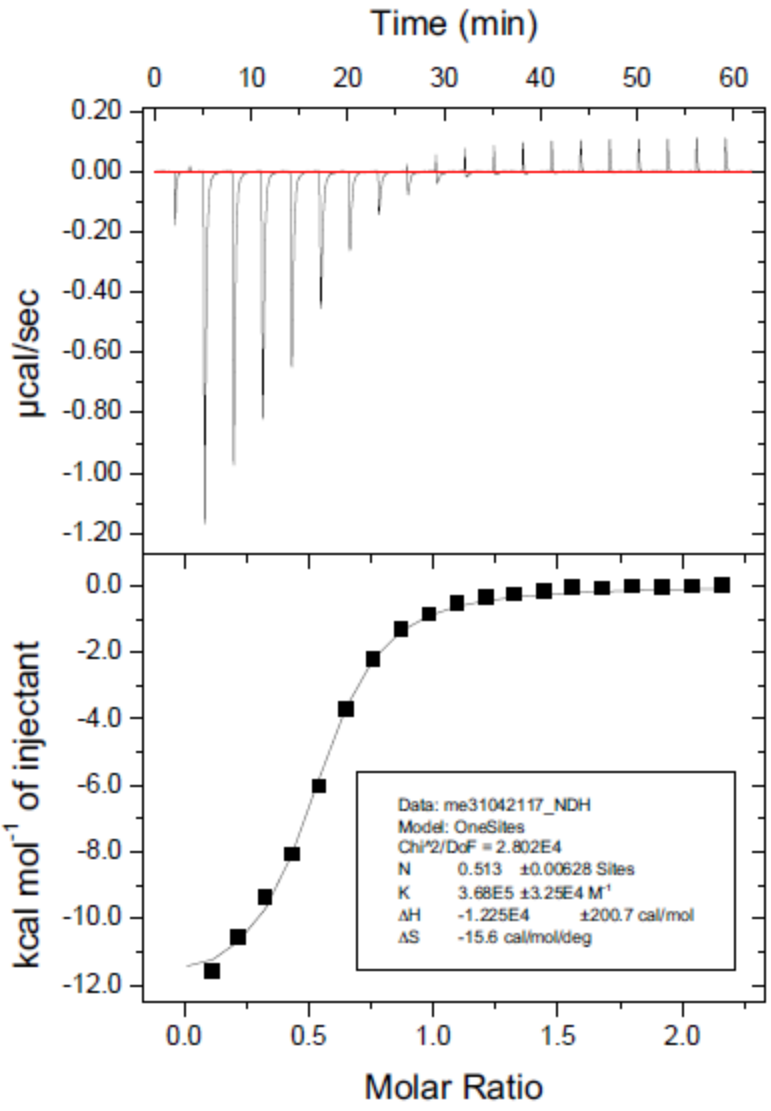


Figure 5.78. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (812 µM) titrated into **A_2J-1** (76 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 1).

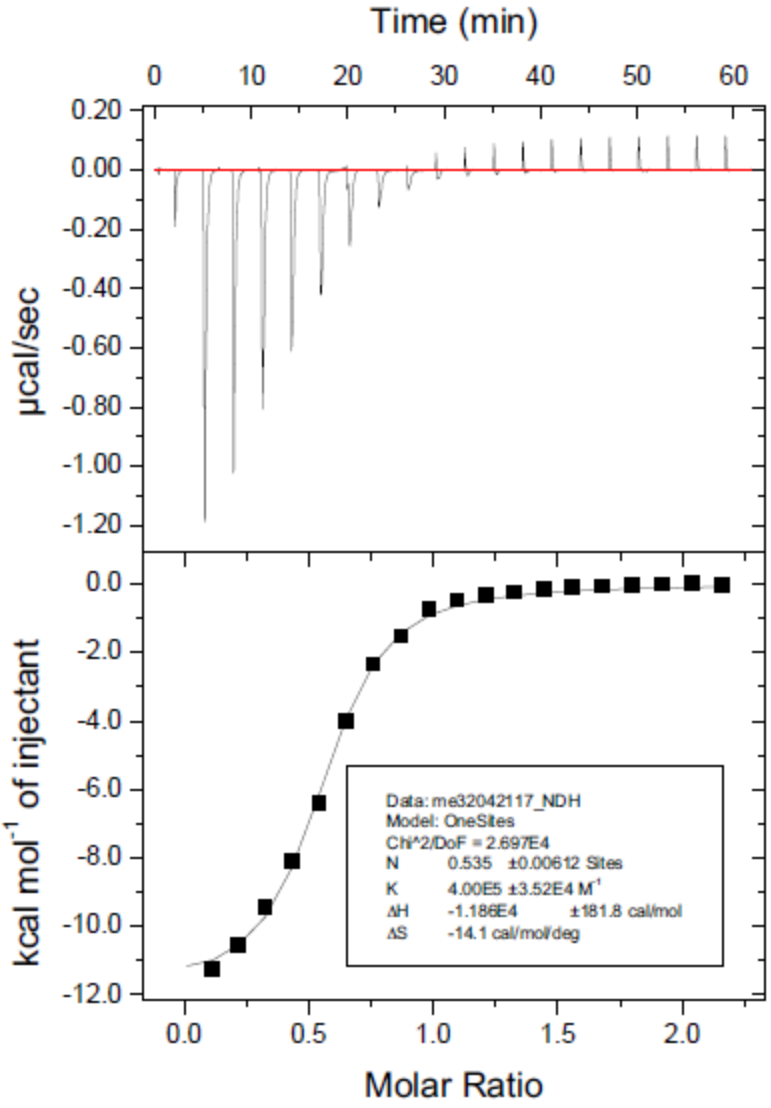


Figure 5.79. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (812 µM) titrated into **A_2J-1** (76 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 2).

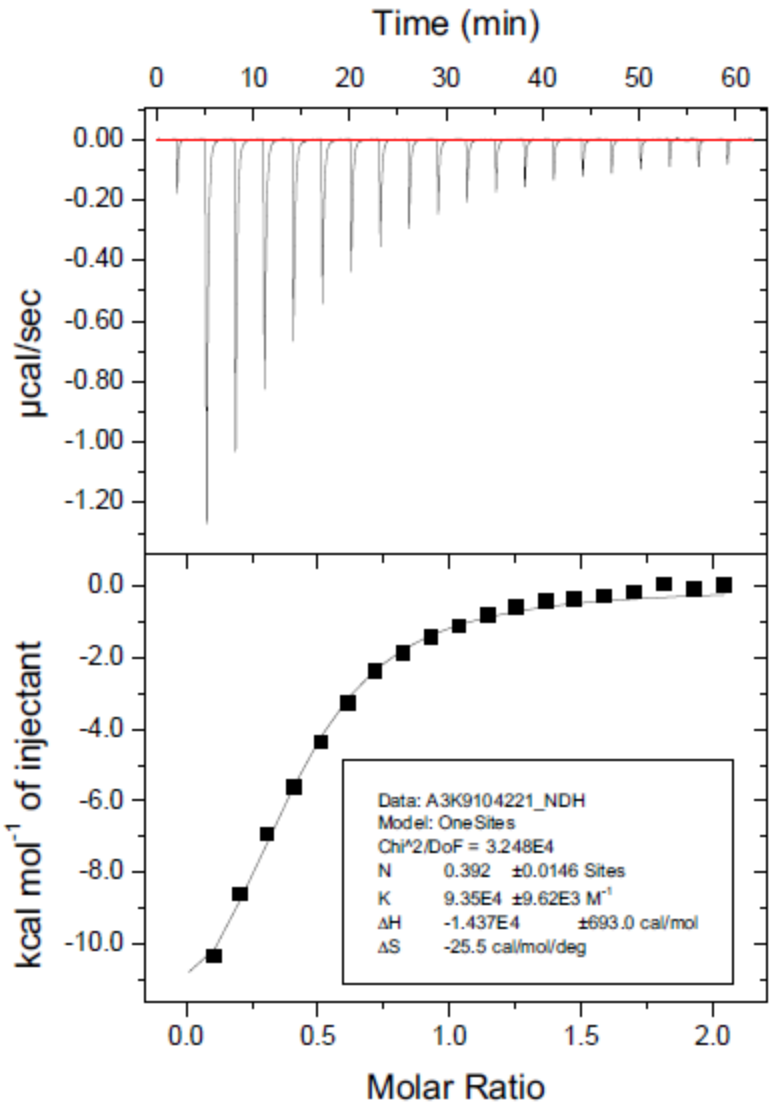


Figure 5.80. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (849 µM) titrated into **A_2J-2** (84 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1).

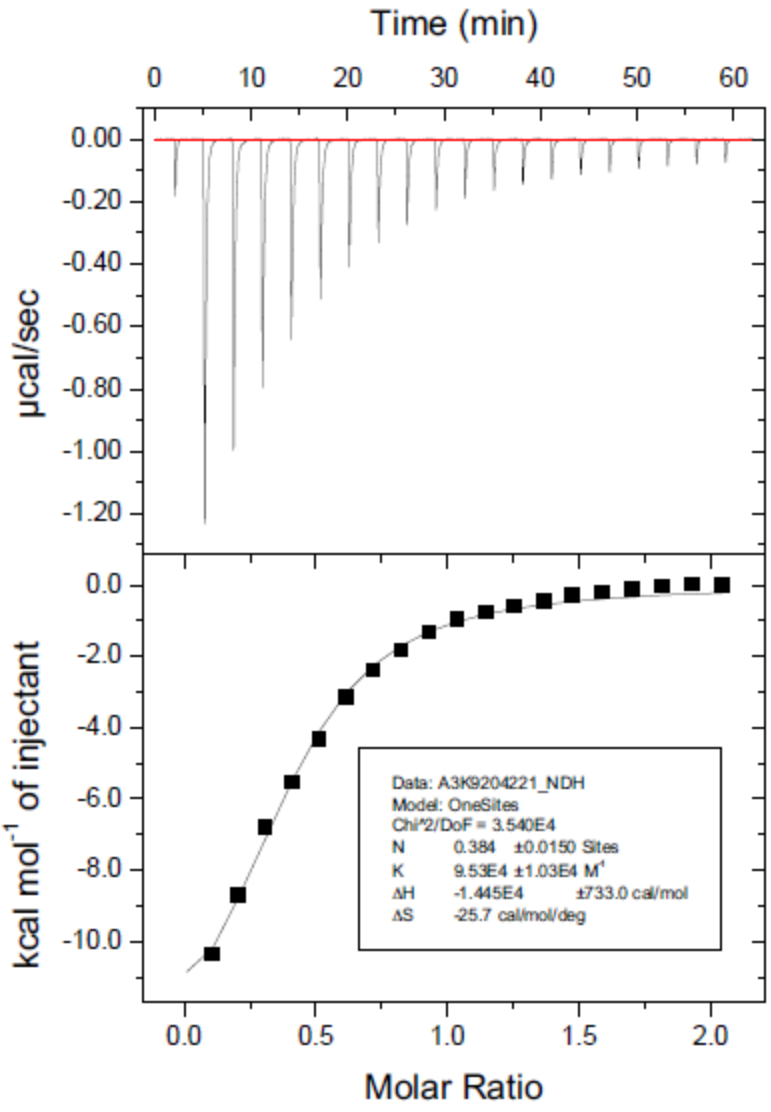


Figure 5.81. One trial of two of H3K9 (Ac-WGGGQTARKSTG-NH_2) (849 µM) titrated into **A_2J-2** (84 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 2).

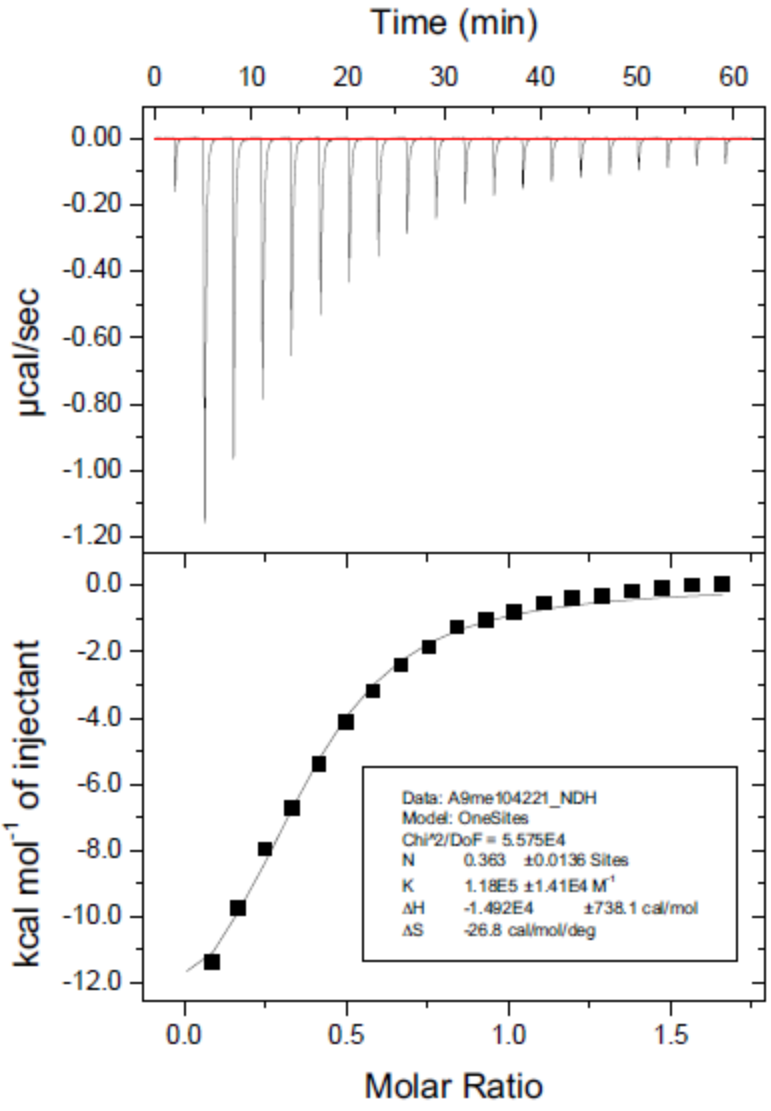


Figure 5.82. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (689 µM) titrated into **A_2J-2** (84 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1).

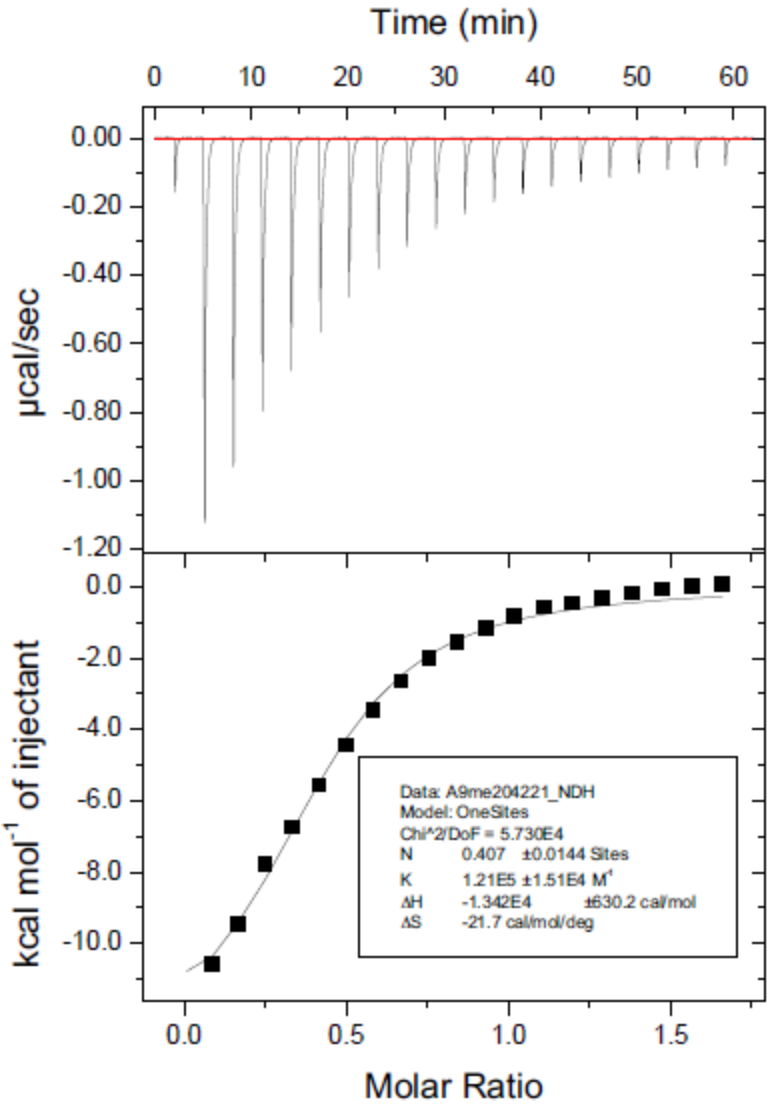


Figure 5.83. One trial of two of H3K9me (Ac-WGGGQTARKmeSTG-NH_2) (689 µM) titrated into **A_2J-2** (84 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 2).

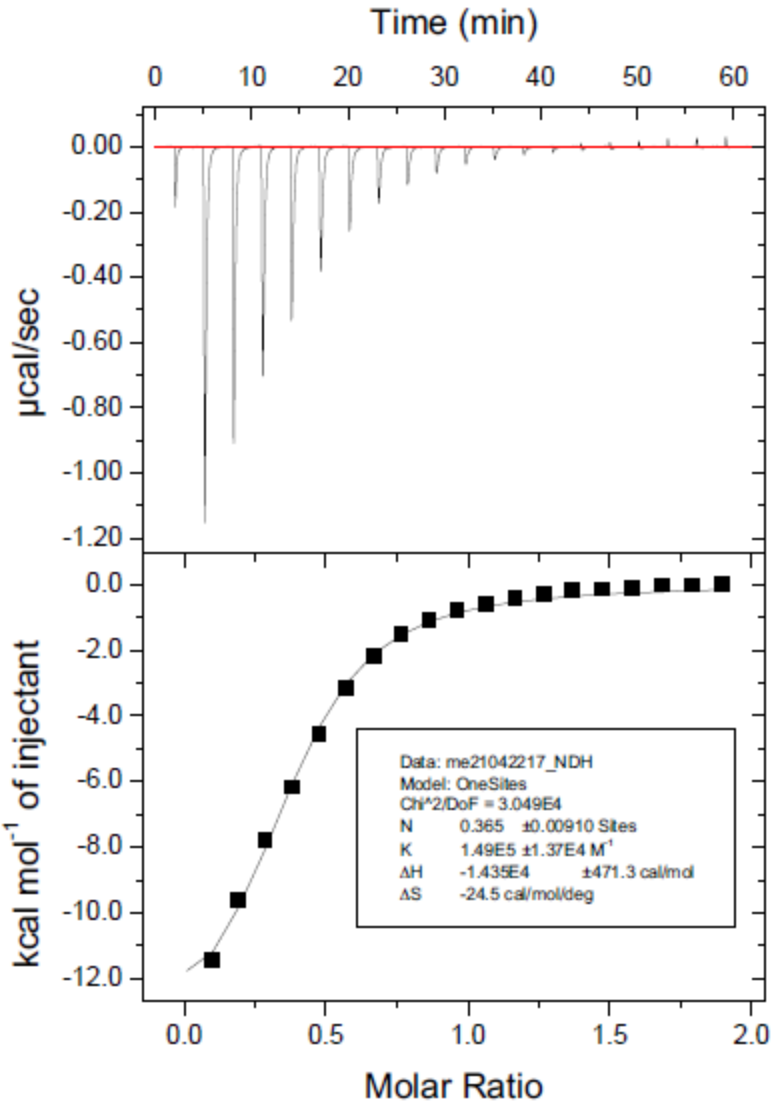


Figure 5.84. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (807 µM) titrated into **A_2J-2** (86 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 1).

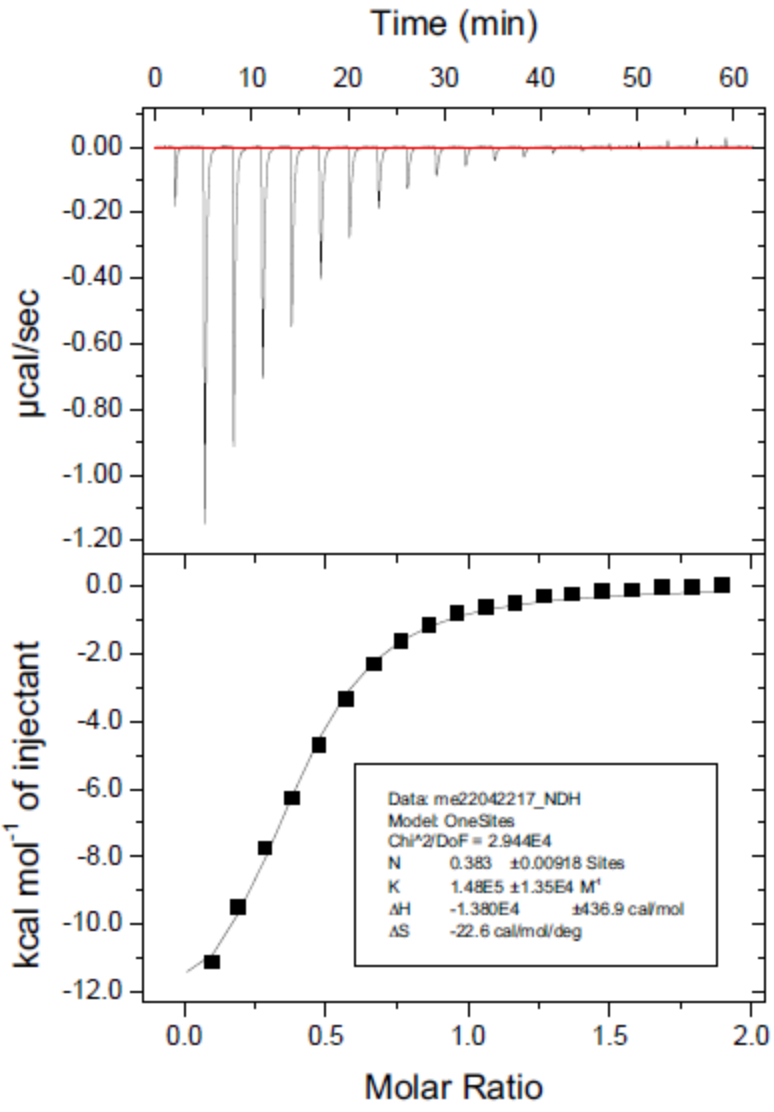


Figure 5.85. One trial of two of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (807 µM) titrated into **A_2J-2** (86 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2).

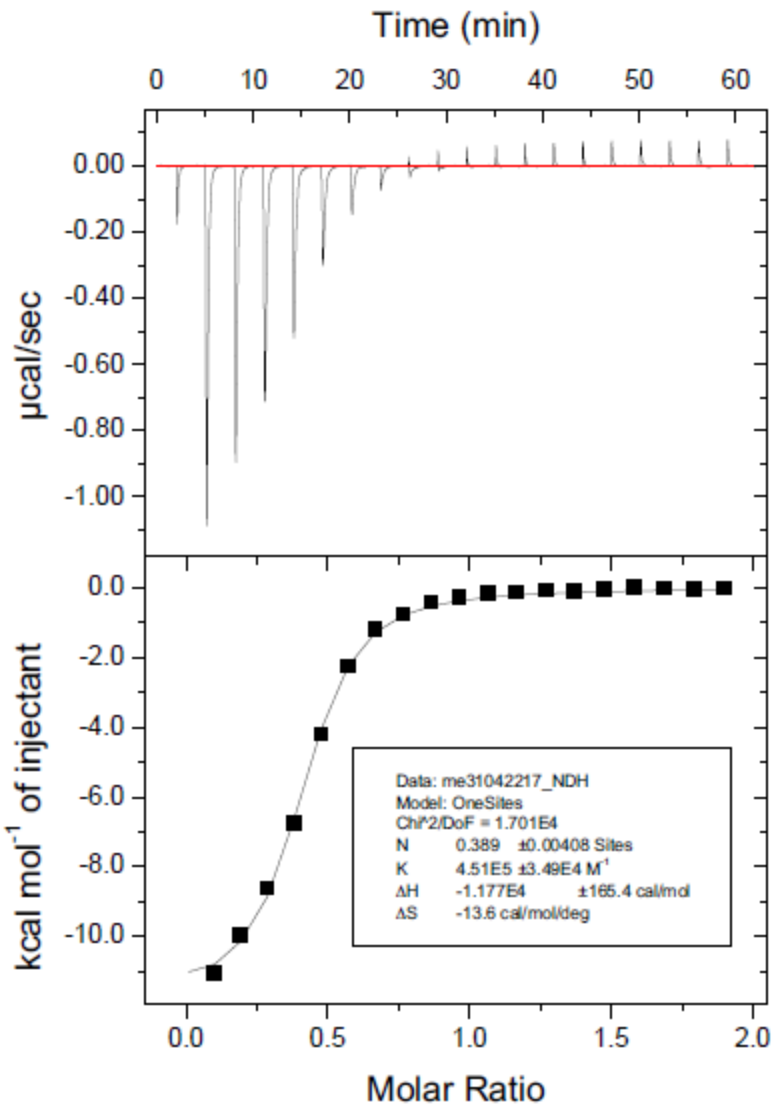


Figure 5.86. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (808 µM) titrated into **A_2J-2** (86 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 1).

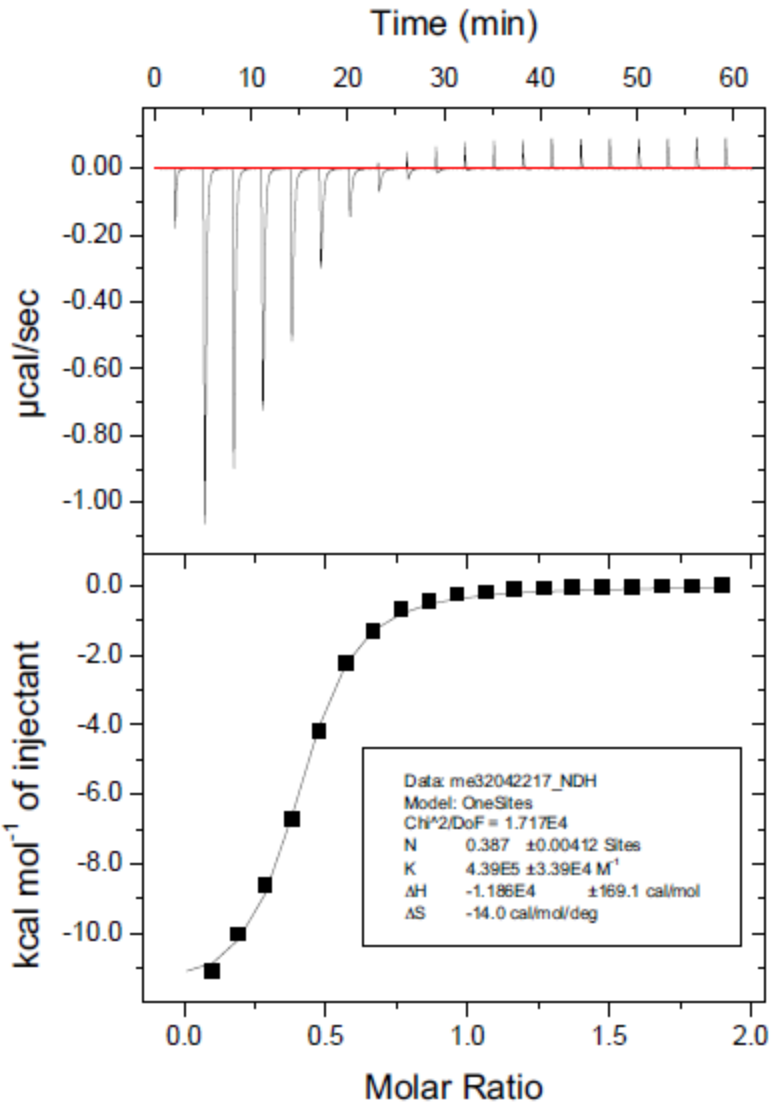


Figure 5.87. One trial of two of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (808 µM) titrated into **A_2J-2** (86 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 2).

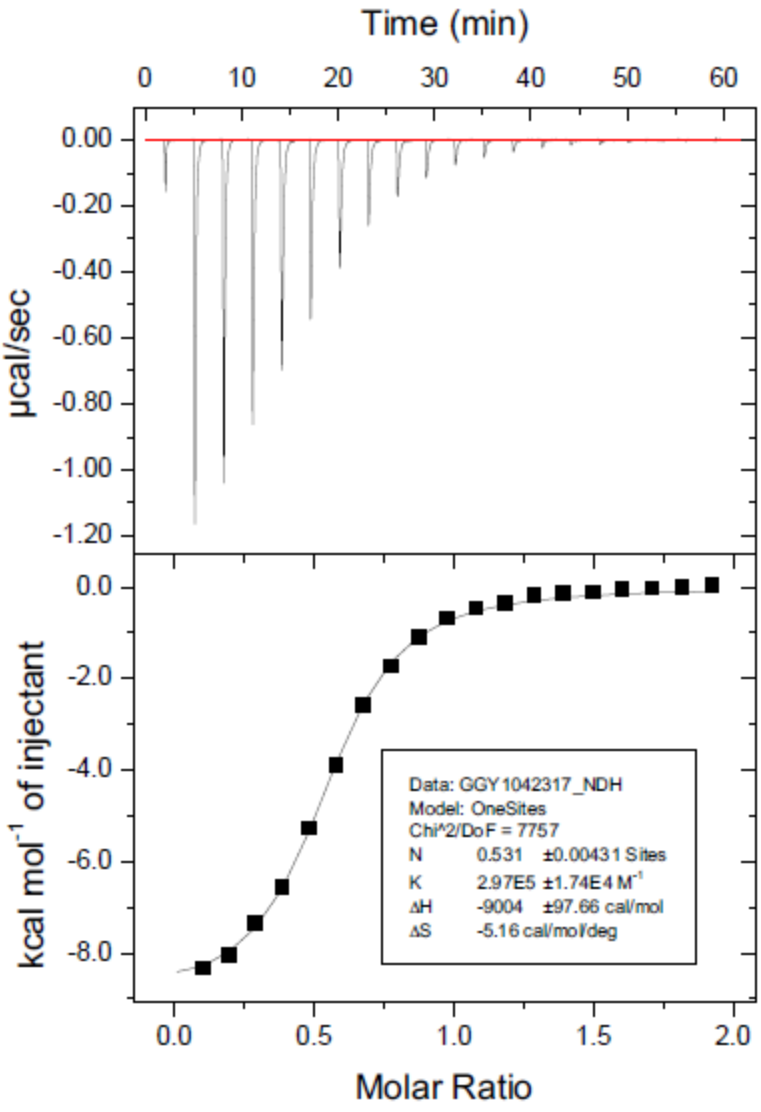


Figure 5.88. One trial of two of Kme2GGY (Ac-Kme2GGY-NH_2) (867 µM) titrated into **A_2I-2** (91 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5. (Run 1).

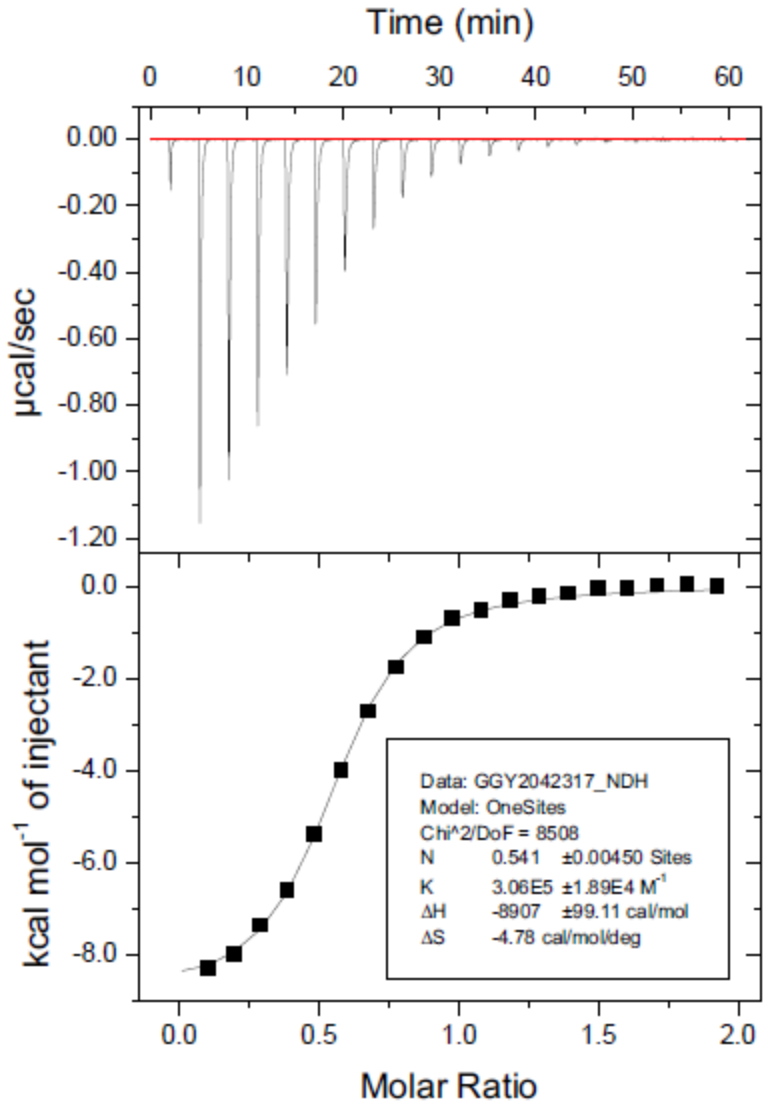


Figure 5.89. One trial of two of Kme2GGY (Ac-Kme2GGY-NH_2) (867 µM) titrated into **A_2I-2** (91 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 2).

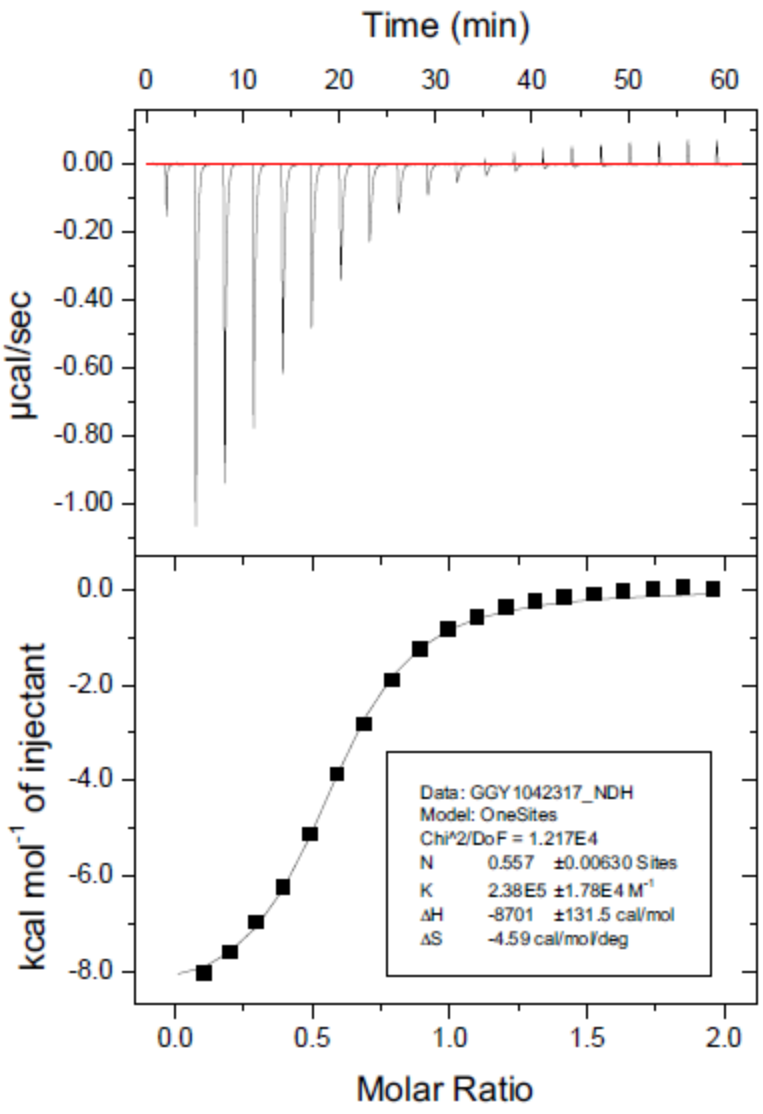


Figure 5.90. One trial of two of Kme3GGY (Ac-Kme3GGY-NH_2) (883 µM) titrated into **A_2I-2** (91 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5. (Run 1).

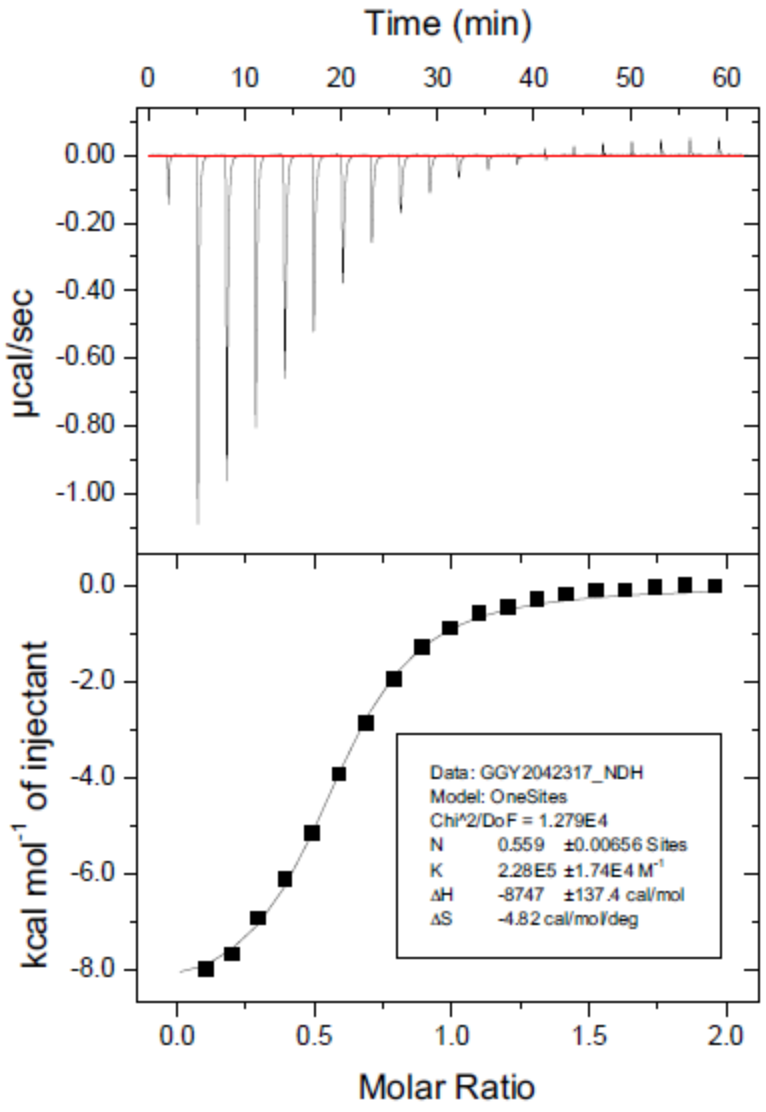


Figure 5.91. One trial of two of Kme3GGY (Ac-Kme3GGY-NH_2) (883 µM) titrated into **A_2I-2** (91 µM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5. (Run 2).

CHAPTER 6. Redesign of a Small Molecule Receptor for Methylated Lysine via Incorporation of Aromatic and Acidic Amino Acids

Introduction and Significance

Current methods to detect post-translational modifications of proteins (such as lysine methylation) include the use of antibodies and mass spectrometry-based proteomics. Even though these methods have been used successfully to identify PTMs, they both face challenges and limitations. While antibodies have been widely used with Western blotting for the identification of PTMs, they are sequence specific and generally require the protein or peptide sequence of interest to be known.[86,109] This makes the detection of unknown post-translational modifications difficult. Their sequence specificity also hinders the detection of the patterns of PTMs that make up the histone code. Additionally, antibodies suffer from high cost of production and irreproducibility of quality from batch to batch.[85] One major challenge facing mass spectrometry-based proteomics is the low natural abundance of post-translational modifications.[165] This significantly complicates detection of proteins with PTMs as the protein of interest is present only in small amounts in a complex mixture of other proteins. Analysis is further complicated by the fact that purification, enrichment, and sample preparation for proteins with a specific PTM are challenging and involve multiple steps.[87,117] Complications with both antibodies and MS-proteomics highlight the need for new tools for detection of PTMs.

The development of small molecule receptors for the recognition of PTMs presents several advantages that would augment strategies using antibodies and mass spectrometry-based

proteomics. Small molecules have much lower cost of production when compared to antibodies and can be synthesized reliably and reproducibly. Receptors of this kind also have the potential to bind to a PTM of interest in a non-sequence-specific manner due to their smaller size.

The Waters lab has developed a small molecule receptor, **A_2B**, that binds trimethyllysine (Kme3) using dynamic combinatorial chemistry.[96] Rigid monomers with aromatic groups for molecular recognition through cation-π interactions, carboxylates for water solubility and electrostatic interactions, and thiols for reversible disulfide exchange were used in dynamic combinatorial libraries (DCLs) to rapidly screen for macrocycles that bind Kme3. The identified receptor bound to histone peptide Ac-WGGGQTARKme3STG-NH_2 with a K_d of 2.6 μM for racemic **A_2B**, which is comparable to the binding affinity of the native Kme3 reader chromodomain, Heterochromatin Protein 1 (HP1).[98] Additionally, the receptor displayed a modest 2-fold selectivity over the Kme2 variant of the peptide.

We sought to redesign receptor **A_2B** with two goals in mind: to improve the affinity and selectivity of the receptor for Kme3 and to reverse the selectivity of the receptor to favor binding to dimethyllysine (Kme2). Here, we use dynamic combinatorial chemistry (DCC) to expolore the effect of covalently attaching amino acids with either aromatic or acidic functionality to a known small molecule receptor for Kme3, **A_2B**.

System Design

A biomimetic approach derived from PTM-binding effector modules was used in an effort to identify synthetic receptors with improved selectivity for Kme3 and Kme2. These modules are known as reader proteins, and they achieve specific molecular recognition of histone PTMs through a balance of cation-π, electrostatic, and hydrogen bonding interactions in the protein's binding pocket.[156] Li et al. demonstrated that the selectivity of a nucleosome

remodeling protein BPTF can be switched from Kme3 to Kme2 by changing the balance of non-covalent interactions in the binding pocket of the protein.[159] The wild-type BPTF binds Kme3 through an aromatic cage consisting of three tyrosines and a trytophan. By simply mutating the tyrosine at position 17 to a glutamic acid, the BPTF protein was engineered to selectively recognize the lower methylation state Kme2 over Kme3.

We looked to mimic this strategy by coupling amino acids to monomer **B** to introduce new functionality that can participate in non-covalent interactions with a guest bound to receptor **A_2B**. We posited that the binding affinity of **A_2B** for Kme3 could be improved by attaching aromatic amino acids like phenylalanine, tyrosine, and tryptophan to monomer **B**. Modeling suggests that this could position an aromatic group over the top of the receptor that should add more cation-π contact area and increase binding to Kme3 peptides (Figure 6.1). Previous studies have shown that the carboxylate of monomer **B** can be converted to an amide without disrupting the function of **A_2B**, suggesting that the proposed modifications will not negatively affect binding.[166] To re-design **A_2B** to be selective for Kme2, we looked to functionalize monomer **B** with an amino acid like apartic acid that is capable of making hydrogen bonding and electrostatic interactions with bound Kme2. Modeling suggests that the side chains of the derivatized monomers can adopt an orientation that allows favorable hydrogen bonding to Kme_2 (Figure 6.1).

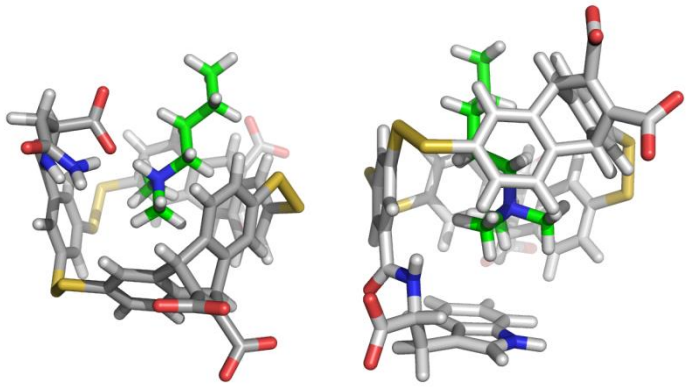
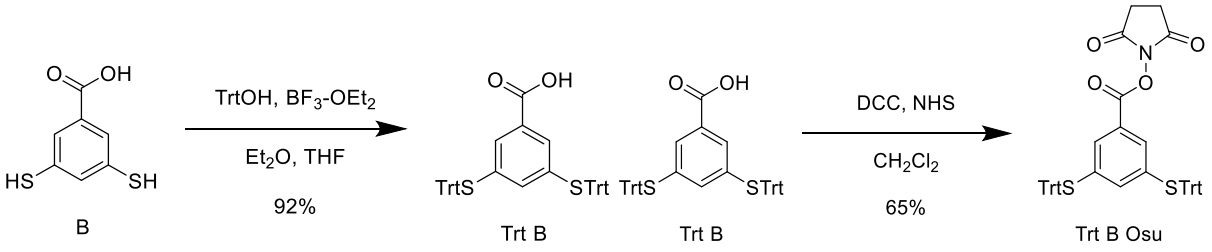


Figure 6.1. Gas phase molecular models of one isomer of **A$_2$B-Asp** (left) and **A$_2$B-Trp** (right) bound to butyldimethyl ammonium and butyltrimethyl ammonium respectively.

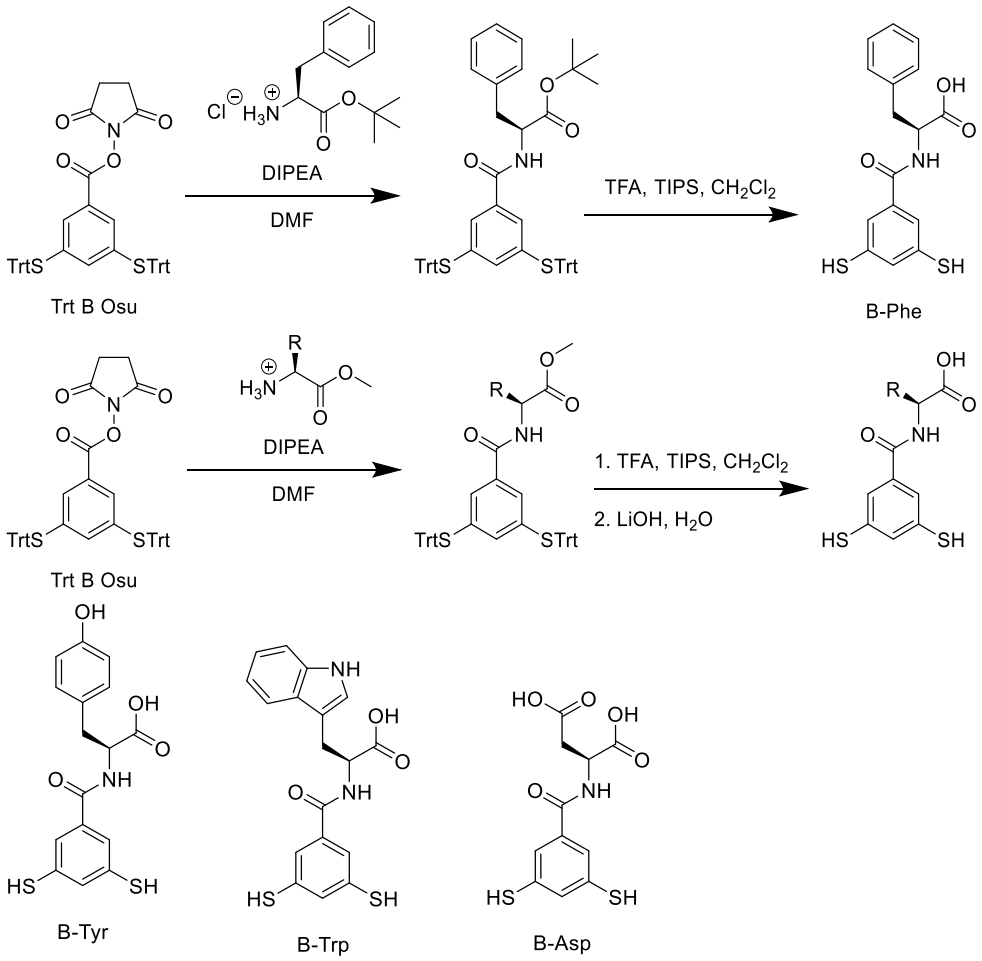
Synthesis of Monomers

The amino acid-functionalized monomers were synthesized starting from monomer **B**. First, the thiol groups were protected with trityl groups by reacting the monomer with triphenylmethanol and boron trifluoride. The trityl protected monomer (**Trt B**) was then converted to the N-hydroxysuccinimidyl ester (**Trt B Osu**) using carbodiimide coupling chemistry. **Trt B Osu** was then used to prepare each of the amino acid-functionalized monomers by reacting with either phenylalanine tertbutyl ester, tyrosine methyl ester, tryptophan methyl ester, or aspartic acid dimethyl ester. Deprotection with TFA gives monomer **B-Phe**, while TFA deprotection and a subsequent lithium hydroxide (or potassium hydroxide for **B-Asp**) hydrolysis is required to obtain **B-Try**, **B-Trp**, and **B-Asp**.

Scheme 6.1. Synthesis of **Trt B Osu**.



Scheme 6.2. Synthesis of **B-Phe**, **B-Tyr**, **B-Trp**, and **B-Asp**.



Dynamic Combinatorial Library Screening

DCC was used for rapid screening of receptors that incorporate the new monomers and that bind to the desired guest. To evaluate the effect of incorporating aromatic amino acids, five unbiased dynamic combinatorial libraries (DCLs) were prepared containing 1 mM **A**, 1 mM **B-Phe**, 1 mM **B-Tyr**, 1-mM **B-Trp** and 4 mM of guest (or no guest for the untemplated library) in 50 mM sodium borate buffer at pH 8.5. Methylated butylamine guests were used as convenient Kme_x side chain mimics: n-butylamine for K, N-butylmethylamine for Kme1, N,N-dimethylbutylamine for Kme2, or butyltrimethylammonium bromide for Kme3. These Kme_x mimics were used as they are inexpensive and commercially available or easily prepared in the

case of butyltrimethylammonium bromide.[163] For monomer **B-Asp**, a discovery library was not prepared. A prep library (described below) was prepared to isolate receptor for characterization of binding by ITC.

The DCLs were allowed to undergo disulfide exchange and equilibration for four days at room temperature in air before analysis by LCMS. For the library containing **A** and the monomers functionalized with aromatic amino acids, two sets of peaks were amplified according to the HPLC trace. One set of peaks highlighted in red in Figure 6.2 corresponded to the mass of **A_2B-Tyr**. The other set of peaks highlighted in blue corresponding to both **A_2B-Phe** and **A_2B-Trp** (which indicates that the two receptors have similar retention times). Due to peak overlap and the complexity of the DCL, values for the extent of amplification compared to the untemplated library could not be accurately determined.

Receptors **A_2B-Phe**, **A_2B-Tyr**, **A_2B-Trp**, and **A_2B-Asp** were synthesized individually on a preparative scale using butyltrimethylammonium bromide as a guest. For the aromatic amino monomers, each preparative library contained 2.5 mM monomer **A**, 2.5 mM of the amino acid-functionalized monomer, and 10 mM guest dissolved in 50 mM sodium borate buffer at pH 8.5. For the preparation of **A_2B-Asp**, the library contained 5 mM monomer **A**, 5 mM of the amino acid-functionalized monomer, and 10 mM guest dissolved in 50 mM sodium borate buffer at pH 8.5. Prep scale DCLs were allowed to equilibrate for a week before purification by semi-preparative reversed-phase HPLC. For **A_2B-Phe**, **A_2B-Tyr**, and **A_2B-Trp**, each receptor was purified as a mixture of racemic and meso isomers due to peak overlap of the isomers. For **A_2B-Asp**, two isomers were purified and isolated.

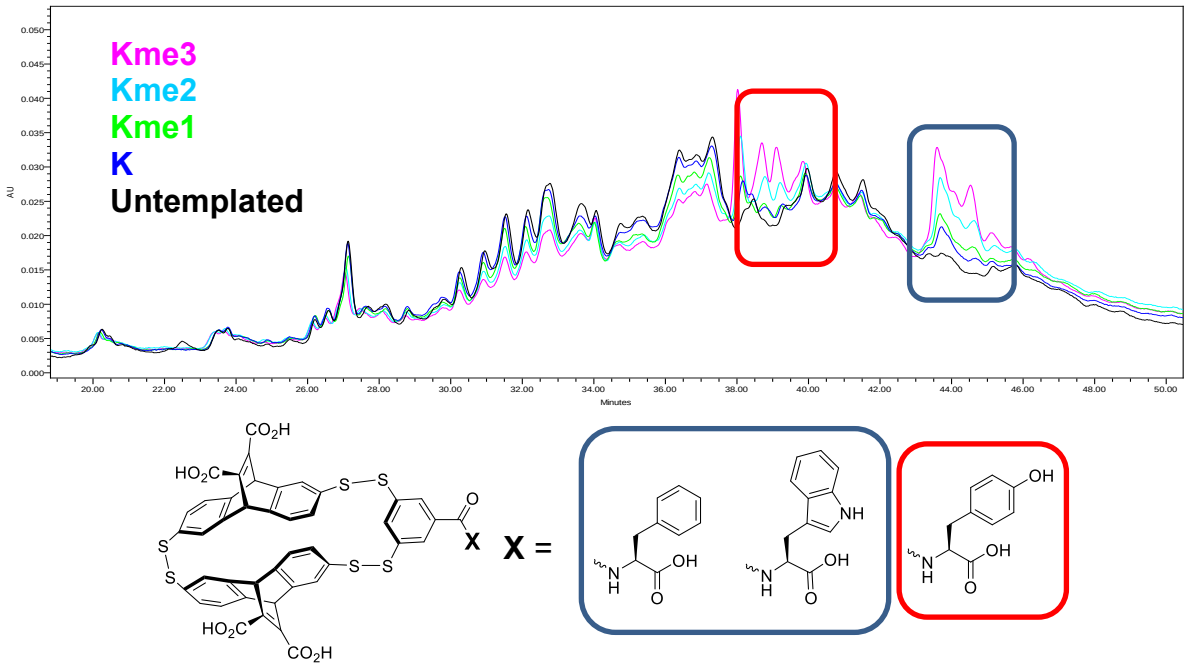


Figure 6.2. Overlaid DCLs of monomer **A**, **B-Phe**, **B-Tyr**, and **B-Trp** with different methylated butylamine guests (1 mM A, 1 mM amino acid-functionalized monomer, and 4 mM guest, 50 mM sodium borate buffer, pH 8.5).

Binding Studies

Binding of the receptors to model H3 histone peptides was characterized using ITC. The H3 peptides (corresponding to the different methylation states of lysine) contained a N-terminal WGGG tag for measuring peptide concentration by UV-Vis. The same peptide sequence that was used for characterizing binding of **A_2B** was used so that a direct comparison could be made[98,99] (see Table 6.1). ITC experiments were conducted at 25 °C in 10 mM borate buffer at pH 8.5. The thermodynamic binding data collected from ITC binding studies of **A_2Phe**, **A_2B-Tyr**, **A_2B-Trp**, and **A_2B-Asp** are summarized in Table 6.2.

Table 6.1. Peptides used for ITC.

Peptide	Sequence
H3 K9me$_x$ (x = 0, 2, or 3)	Ac-WGGG-QTARKme$_x$STG-NH_2

Influence of Attaching Aromatic Amino acids to Monomer B

Measurement of binding by ITC revealed that **A_2B-Phe** and **A_2B-Trp** bind Kme3 to the same extent (4.4 μM and 4.9 μM respectively), while **A_2B-Tyr** binds slightly tighter with a K_d of 1.6 μM. None of the receptors show a significant improvement in affinity for Kme3 when compared to the parent **A_2B** receptor. Additionally, selectivity for binding Kme3 over the lower methylation states of lysine is reduced relative to **A_2B**. It is likely that the side chains of the attached aromatic amino acids have too much conformational flexibility such that the arene spends a significant amount of time in conformations that do not contribute favorably to binding. The loss in selectivity over the lower methylation states may be due to non-specific electrostatic interactions with the peptide. These interactions may arise from the change in position of carboxylic acid on the **B** monomer which may bring it closer proximity to the neighboring Arg.

Influence of Attaching an Acidic Amino Acid to Monomer B

Analysis of binding of **A_2B-Asp** to the model histone peptides showed that either isomer binds to Kme3 with similar affinity to the parent receptor **A_2B**, with **A_2B-Asp-1** (the isomer that eluted from the HPLC first) having a K_d of 2.1 μM and **A_2B-Asp-2** having a K_d of 3.7 μM. Both isomers bound Kme3 and Kme2 to the same extent, indicating an improvement in binding to Kme2. However, the selectivity over unmethylated lysine was significantly decreased. The increase in affinity for Kme2 and K is likely due to non-specific electrostatic interactions, which may arise from repositioning of the carboxylate on **B** to bring it in closer proximity to a neighboring Arg. A more specific salt bridge interaction with the bound Kme2 guest is not

observed presumably due to the considerable conformational flexibility around the amide bond connecting **B** to Asp in addition to rotation around the side chain of the amino acid.

Table 6.2. Thermodynamic binding data for the biding of **A_2B-Phe**, **A_2B-Trp**, **A_2B-Tyr**, **A_2B-Asp-1**, and **A_2B-Asp-2** to peptides shown in **Table 6.1** as measured by ITC.

Entry	***Receptor***	***Peptide***	***Charge***	***K_d [b] (μM)***	***SF[d]***	***ΔG[c] (kcal/mol)***
1	A_2B[a]	H3K9	+2	22 ± 1	8.3	-6.38 ± 0.02
2	A_2B[a]	H3 K9me2	+2	6.3 ± 0.3	2.4	-7.10 ± 0.07
3	A_2B[a]	H3 K9me3	+2	2.6 ± 0.1	-	-7.63 ± 0.03
4	A_2B-Phe	H3K9	+2	7.9 ± 1.1	1.8	-6.96 ± 0.97
5	A_2B-Phe	H3K9me2	+2	6.5 ± 0.4	1.5	-7.07 ± 0.44
6	A_2B-Phe	H3 K9me3	+2	4.4 ± 0.2	-	-7.31 ± 0.33
7	A_2B-Trp	H3K9	+2	>16	3.2	>-6.56
8	A_2B-Trp	H3 K9me2	+2	7.2 ± 0.5	1.5	-7.02 ± 0.49
9	A_2B-Trp	H3 K9me3	+2	4.9 ± 0.2	-	-7.24 ± 0.30
10	A_2B-Tyr	H3K9	+2	8.6 ± 0.7	5.3	-6.91 ± 0.56
11	A_2B-Tyr	H3 K9me2	+2	3.0 ± 0.2	1.9	-7.53 ± 0.50
12	A_2B-Tyr	H3K9me3	+2	1.6 ± 0.1	-	-7.91 ± 0.49
13	A_2B-Asp-1	H3K9	+2	6.3 ± 0.5	3.0	-7.10 ± 0.56
14	A_2B-Asp-1	H3 K9me2	+2	2.3 ± 0.1	1.1	-7.69 ± 0.33
15	A_2B-Asp-1	H3 K9me3	+2	2.1 ± 0.1	-	-7.76 ± 0.37
16	A_2B-Asp-2	H3K9	+2	8.5 ± 0.8	2.3	-6.91 ± 0.65
17	A_2B-Asp-2	H3K9me2	+2	3.5 ± 0.2	0.9	-7.44 ± 0.43
18	A_2B-Asp-2	H3K9me3	+2	3.7 ± 0.3	-	-7.42 ± 0.60

[a] Reported by Pinkin and Waters.[98] [b] Errors are from the error in the fit as determined by the Origin 7 software. [c] ΔG was calculated from the K_d. Errors are from the error in the fit as determined by the Origin 7 software [d] The selectivity factor is calculated as the K_d of Kmex (where x = 0, 2, or 3) divided by the K_d of Kme3.

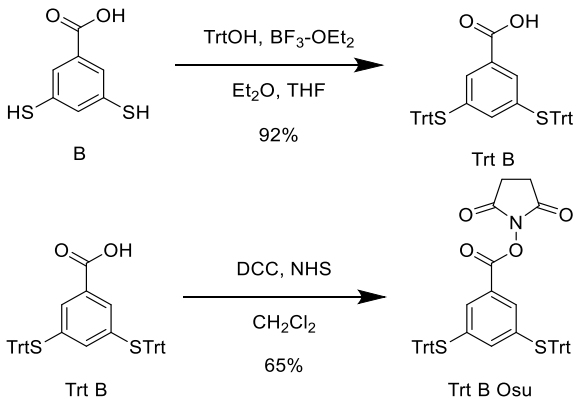
Conclusions

In summary, we have utilized DCC to explore the effect of covalently attaching various aromatic and acidic amino acids to a previously reported small molecule receptor for Kme3 (**A_2B**). Attaching Phe, Tyr, or Trp did not lead to an improvement in binding to Kme3. Furthermore, selectivity over the lower methylation states was diminished. Attaching Asp to **A_2B** lead to a receptor with equal affinity for Kme3 and Kme2 but with weakened selectivity over K. Overall, selective non-covalent interactions between the amino acid side chain and the bound methylated lysine guest could not be achieved, presumably due to the high conformational flexibility of the modified receptors. Additionally, we believe that non-specific electrostatic interactions are responsible for the tighter binding to the unmethylated lysine peptide (compared to **A_2B**). Since attaching the aromatic amino acids produced receptors with the same overall charge as the parent **A_2B** receptor, the non-specific electrostatic interactions cannot be due simply to an increase in overall charge. The non-specific interactions may be due to placing a carboxylate (on the amino acid) in closer proximity to the positively charge arginine on the peptide. Ultimately, monomers that more rigidly direct functional groups into the binding pocket of the receptor may hold more promise for either increasing cation-π interactions or introducing favorable electrostatic and hydrogen bonding interactions within the receptor.

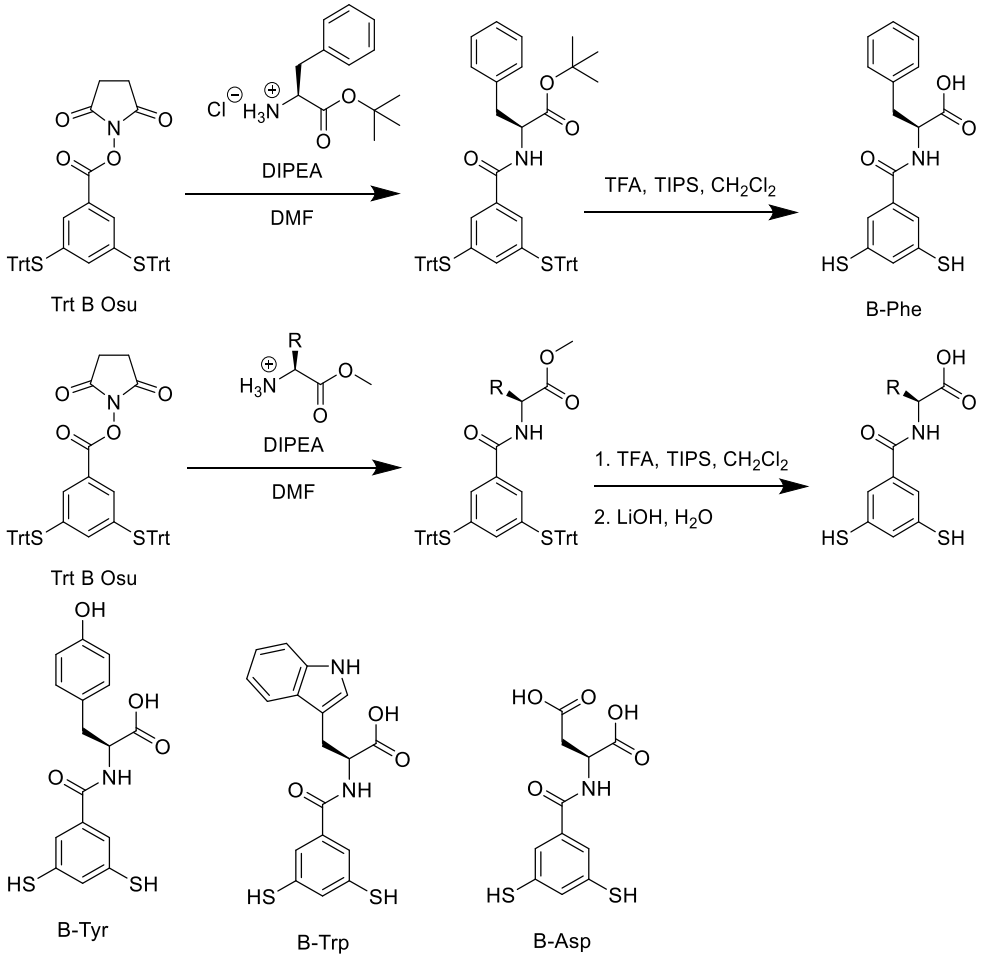
Experimental

Monomer **A** was prepared according to a previously reported procedure.[101]

Scheme 6.3. Synthesis of **Trt B Osu**



Scheme 6.4. Synthesis of **B-Phe**, **B-Tyr**, **B-Trp**, and **B-Asp**.



Synthesis of Trt B

Monomer **B** (1.0 g, 5.37 mmol) and triphenylmethanol (2.91 g, 11.18 mmol) were added to a round bottom flask and dissolved in 15 mL diethyl ether under an atmosphere of nitrogen to form an off-white slurry. Boron trifluoride diethyl etherate (48% in diethyl ether, 3.0 mL, 11.67 mmol) was added dropwise to the flask (a yellow color develops upon addition). The slurry disappears as the reaction progresses, and a white precipitate precipitates out of solution. To create a homogenous solution, 10 mL of THF was added. After completion of the reaction (monitored by TLC, 100% CH_2Cl_2), the solvent was removed under vacuum. The crude solid was taken up in diethyl ether and washed with 1N NaOH (25 mL), H_2O (25 mL), and brine (25 mL). The organic fraction was dried over $MgSO_4$, and the solvent was removed under reduced pressure to yield a white solid (3.3 g, 92%). ^1H NMR ($CDCl_3$, 400 MHz): 6.973-7.339 (m, 33H, CH).

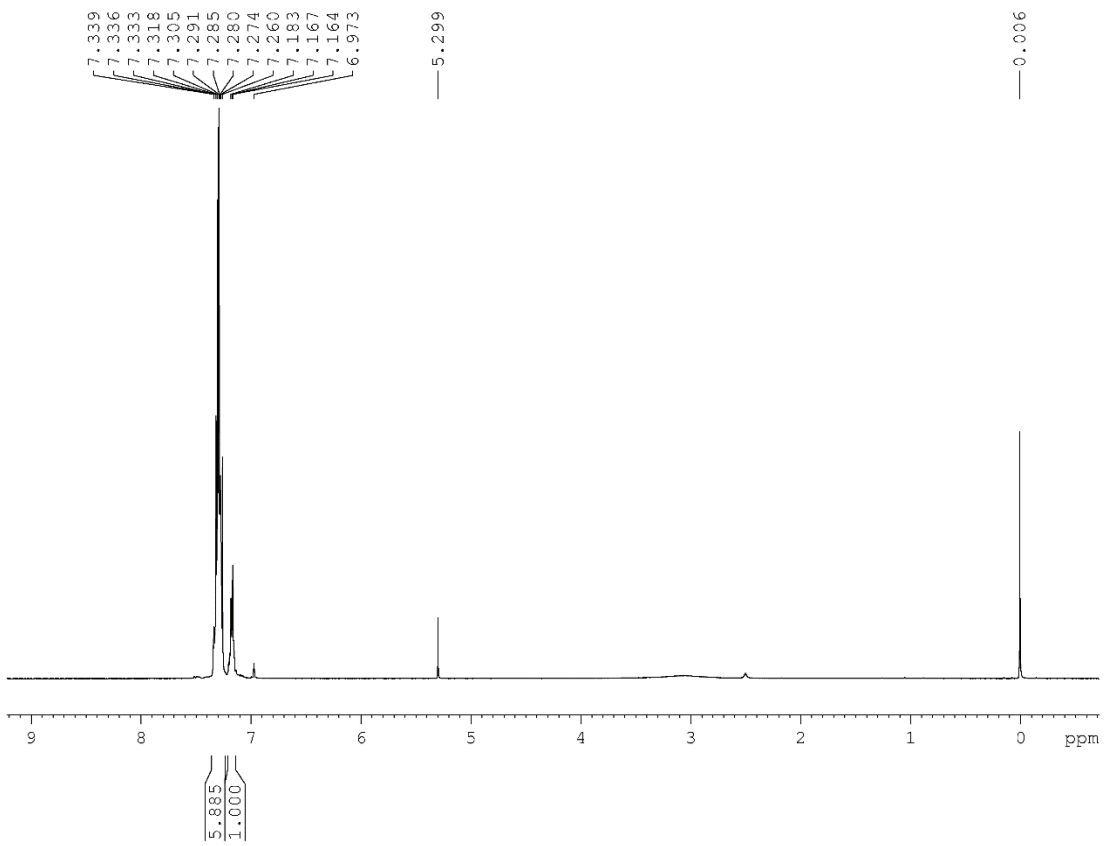


Figure 6.3. 1D ^{1}H-NMR (400 MHz) of **Trt B** in $CDCl_3$.

Synthesis of Trt B Osu

Trt B (1.5 g, 2.24 mmol), N-hydroxysuccinimide (0.27 g, 2.35 mmol), and N,N'-dicyclohexylcarbodiimide (0.48 g, 2.32 mmol) were dissolved in 15 mL CH_2Cl_2 under an atmosphere of nitrogen. The reaction was allowed to stir overnight. After completion of the reaction (monitored by TLC, 100% CH_2Cl_2), the solvent was removed under reduced pressure. The crude was purified by silica column chromatography (100% CH_2Cl_2) to give a crystalline white solid (1.7 g, 65%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.161-7.308 (m, 33H, CH); 2.833 (s, 4H, CH).

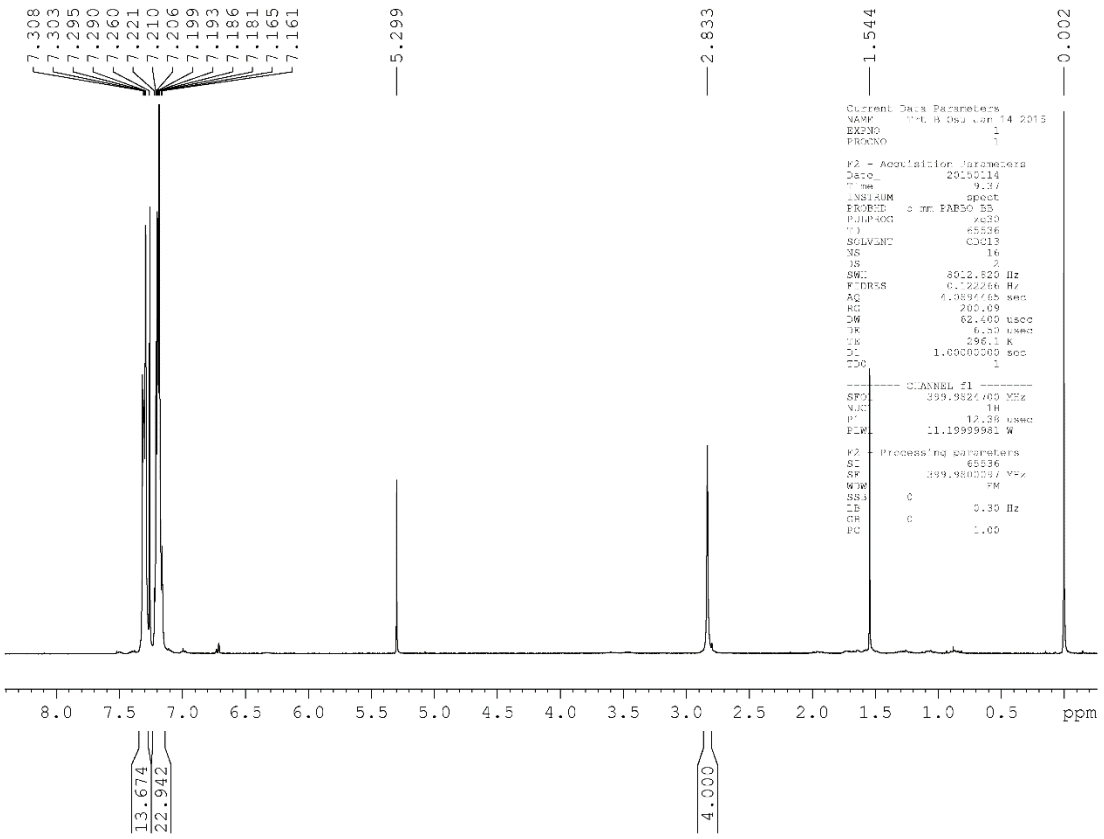


Figure 6.4. 1D ^{1}H-NMR (400 MHz) of **Trt B Osu** in $CDCl_3$.

Synthesis of Trt B-Phe

Trt B Osu (0.100 g, 0.13 mmol) and phenylalanine tert-butyl ester hydrochloride (0.134 g, 0.52 mmol) were dissolved in 1.0 mL DMF in a flask under a nitrogen atmosphere. Triethylamine (0.15 mL, 1.08 mmol) was added, and the reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, it was diluted with toluene (20 mL). Solvent was removed under vacuum, and the crude was purified by flash column chromatography (1:2 EtOAc:hexanes) to give a whitish-gray solid (0.108 g, 95%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.056-7.310 (m, 36 H, CH); 6.840 (s, 2H, CH); 5.777, 5.758 (d, 1H, NH); 4.739, 4.719 (m, 1H, CH); 4.131, 4.113 (m, 2H, CH); 1.420 (s, 9H, CH).

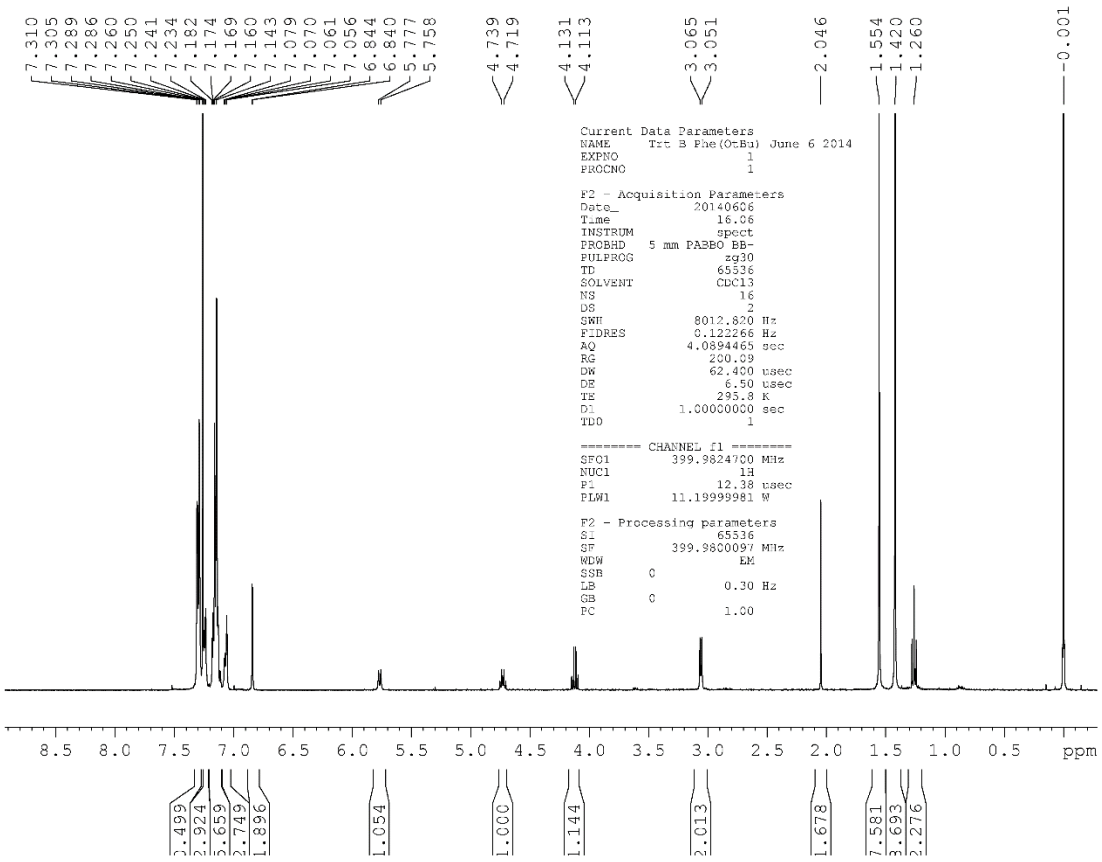


Figure 6.5. 1D ^{1}H-NMR (400 MHz) of **Trt B-Phe** in $CDCl_3$.

Synthesis of Trt B-Trp

Trt B Osu (0.50 g, 0.65 mmol) and tryptophan methyl ester hydrochloride (0.66 g, 2.59 mmol) were dissolved in DMF (5.0 mL) in a flask under a nitrogen atmosphere. DIPEA (0.91 mL, 5.22 mmol) was added, and the reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, it was diluted with water (50 mL) to precipitate out a white solid, which was collected by vacuum filtration. The crude material was purified using silica column chromatography (0-1% MeOH in CH_2Cl_2) to give a yellow solid (0.53 g, 93%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.107-7.772 (m, 34H, CH); 7.002 (s, 1H, CH); 6.777 (s, 2H, CH);

6.749 (s, 1H, CH); 5.686, 5.665 (d, 1H, NH); 4.930 (m, 1H, CH); 3.724 (s, 3H, CH); 3.297, 3.288 (m, 2H, CH).

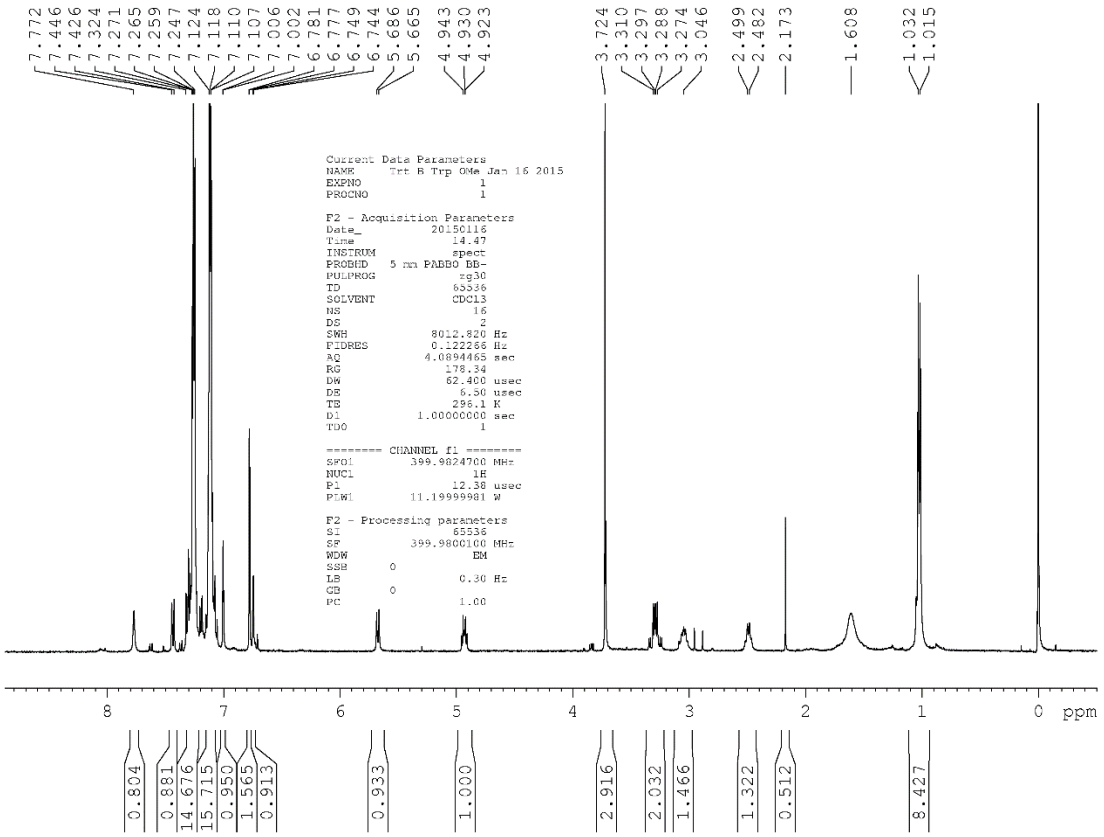


Figure 6.6. 1D ^{1}H-NMR (400 MHz) of **Trt B-Trp** in $CDCl_3$.

Synthesis of Trt B-Tyr

Trt B Osu (0.50 g, 0.65 mmol) and tyrosine methyl ester hydrochloride (0.60 g, 2.59 mmol) were dissolved in DMF (5.0 mL) in a flask under a nitrogen atmosphere. DIPEA (0.91 mL, 5.22 mmol) was added, and the reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, it was diluted with water (50 mL) to precipitate out a white solid, which was collected by vacuum filtration. The crude material was purified using silica column chromatography (0-1% MeOH in CH_2Cl_2) to give a white solid (0.50 g, 91%). ^{1}H

NMR ($CDCl_3$, 400 MHz): 7.082-7.303 (m, 31H, CH), 6.808 (s, 2H, CH); 6.837, 6.858 (d, 2H, CH); 6.672, 6.693 (d, 2H, CH); 5.688, 5.668 (d, 1H, NH); 4.848, 4.828 (m, 1H, CH); 3.734 (s, 3H, CH); 2.998 (m, 2H, CH).

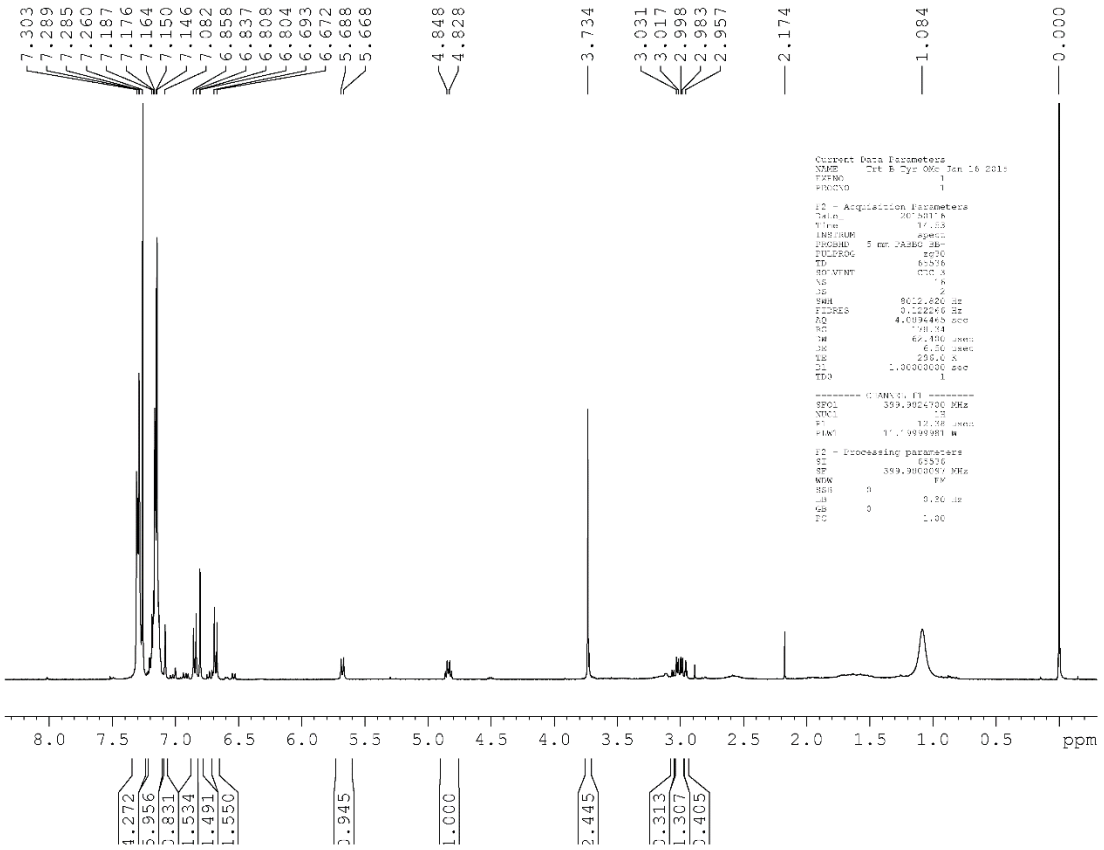


Figure 6.7. 1D ^{1}H-NMR (400 MHz) of **Trt B-Tyr** in $CDCl_3$.

Synthesis of Trt B-Asp

Trt B Osu (0.50 g, 0.69 mmol) and aspartic acid dimethyl ester hydrochloride (0.206 g, 1.04 mmol) were dissolved in DMF (4.0 mL) in a flask under a nitrogen atmosphere. DIPEA (0.36 mL, 2.09 mmol) was added, and the reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, it was diluted with water (50 mL) to

precipitate out a white solid, which was collected by vacuum filtration. The crude material was purified using silica column chromatography (1:2 EtOAc:hexanes) to give a white solid (0.49 g, 92%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.177-7.325 (m, 30H, CH); 7.083 (s, 1H, CH); 6.862, 6.858 (s, 2H, CH); 6.281, 6.261 (d, 1H, NH); 4.845 (m, 1H, CH); 3.763 (s, 3H, CH); 3.690 (s, 3H, CH); 2.977, 2.829 (m, 2H, CH).

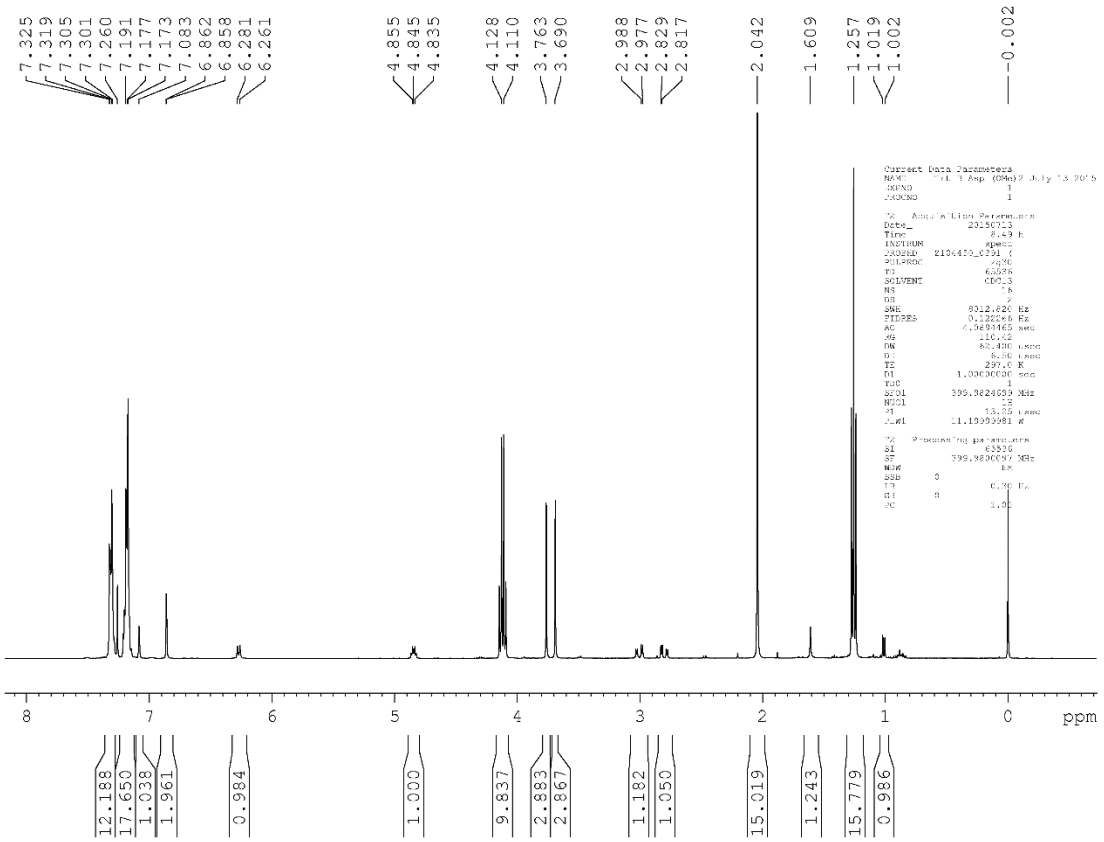


Figure 6.8. 1D ^{1}H-NMR (400 MHz) of **Trt B-Asp** in $CDCl_3$.

Synthesis of B-Phe

A cocktail of TFA (10 mL), TIPS (0.5 mL), and CH_2Cl_2 (2 mL) was added to Trt B-Phe (0.50 g, 0.57 mmol) in a flask under a nitrogen atmosphere. The reaction progress was monitored

by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, solvent was removed by bubbling with nitrogen. The crude was triturated in hexanes (3 X 5 mL) and dried under vacuum. The crude product was stored under nitrogen and was used without further purification (0.167 g, 87%). ^{1}H NMR (MeOD, 400 MHz): 7.825 (s, 2H, CH); 7.683 (s, 1H, CH); 7.141-7.250 (m, 5H, CH); 4.694 (m, 1H, CH); 3.099 (m, 2H, CH). MS (calculated): 332.04 [M-H], 665.09 [2M-H]. MS (observed, ESI-): 332.12 [M-H], 665.12 [2M-H].

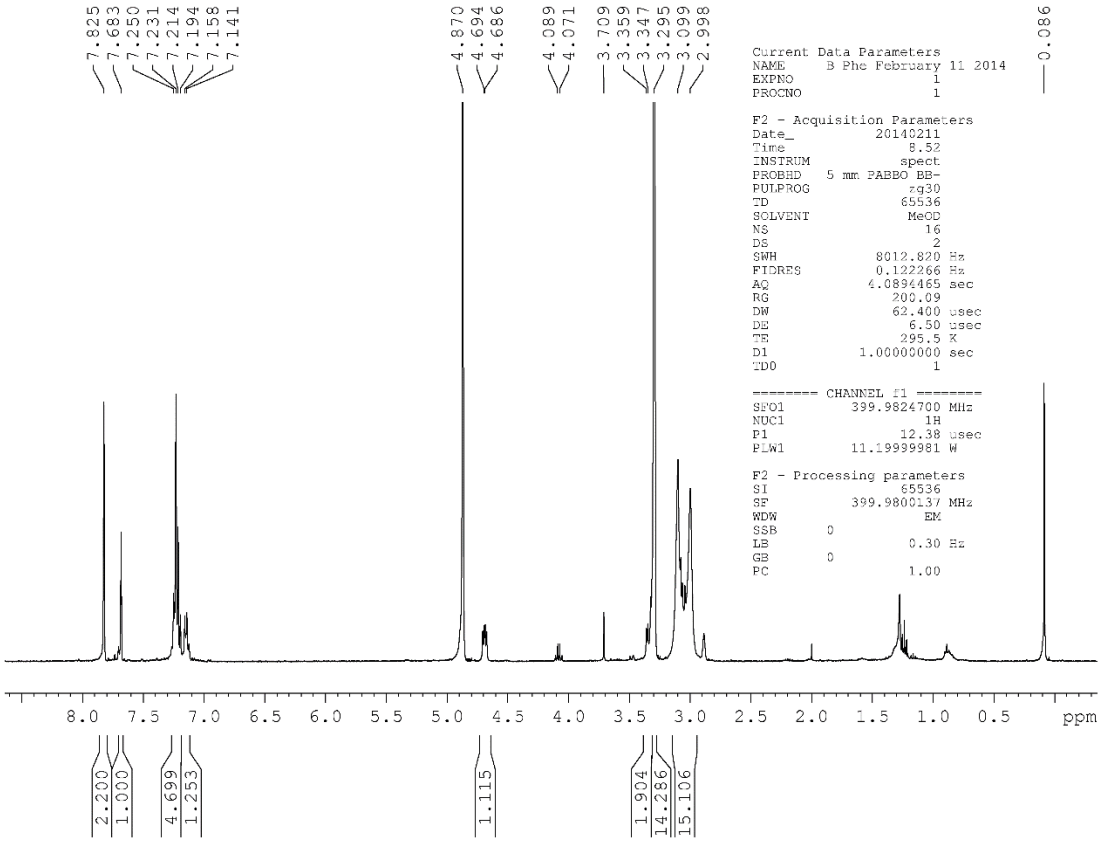


Figure 6.9. 1D ^{1}H-NMR (400 MHz) of **B-Phe** in MeOD.

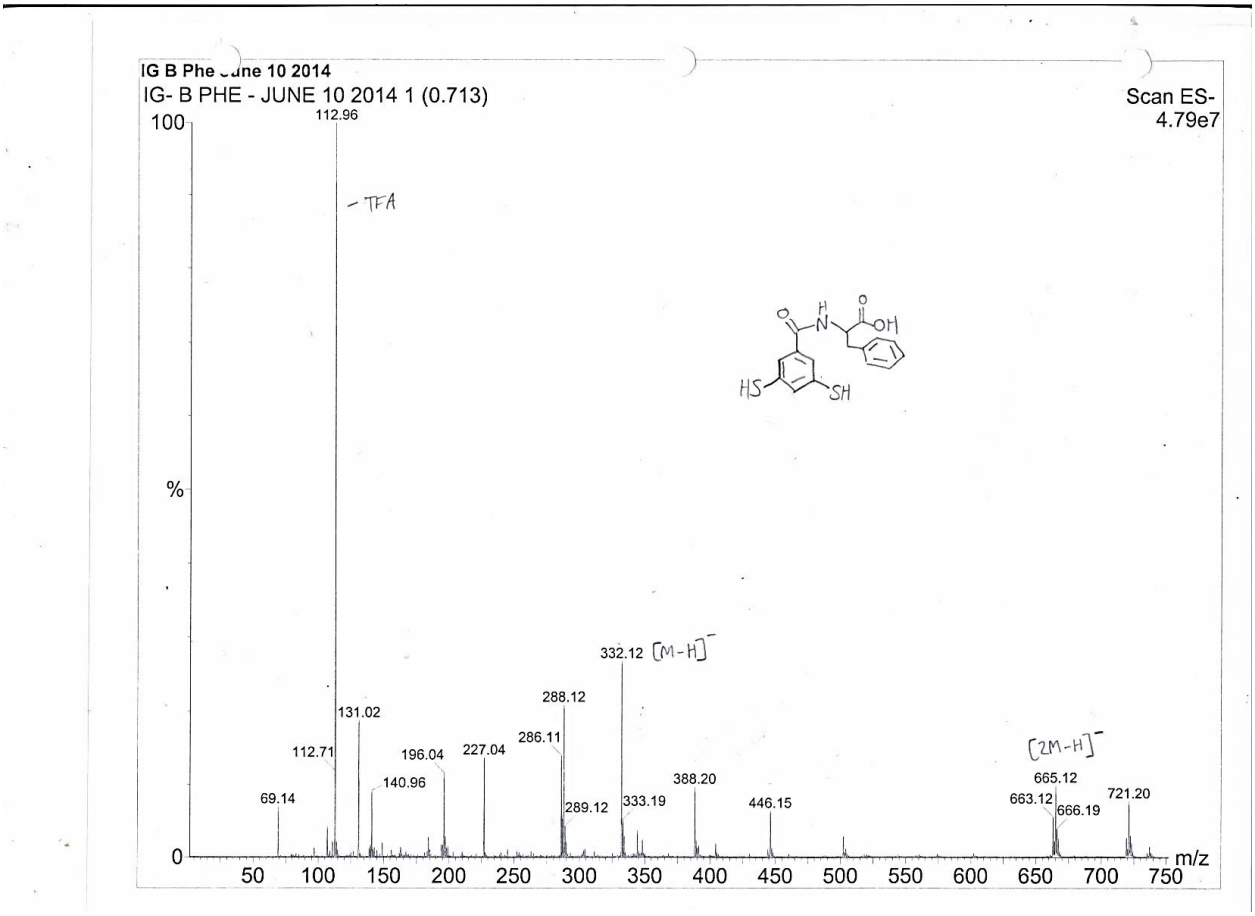


Figure 6.10. High resolution mass spectrum of **B-Phe**. MS (calculated): 332.04 [M-H], 665.09 [2M-H]. MS (observed, ESI-): 332.12 [M-H], 665.12 [2M-H].

Synthesis of B-Tyr

A cocktail of TFA (10 mL), TIPS (0.5 mL), and CH_2Cl_2 (2 mL) was added to Trt B-Tyr (0.40 g, 0.47 mmol) in a flask under a nitrogen atmosphere. The reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, solvent was removed by bubbling with nitrogen. The crude was triturated in hexanes (3 X 5 mL) and dried under vacuum. A solution of LiOH monohydrate (0.02 g, 4.7 mmol) in H_2O was degassed by bubbling with nitrogen for 2 h. The crude solid was dissolved in 5 mL of the LiOH solution, and the reaction was allowed to stir overnight. The product was precipitated by addition of 5 mL of 1N HCl and collected by vacuum filtration. The crude product was used without further purification

(0.117 g, 71% over two steps). ^{1}H NMR (MeOD, 400 MHz): 7.318 (s, 3H, CH); 7.057, 7.036 (d, 2H, CH); 6.688, 6.667 (d, 2H, CH); 4.714 (m, 1H, CH); 2.926-3.214 (m, 2H, CH). MS (calculated): 349.04 [M-2H+D]. MS (observed, ESI-): 349.04 [M-2H+D]. Sample was run in MeOD, which lead to H/D exchange.

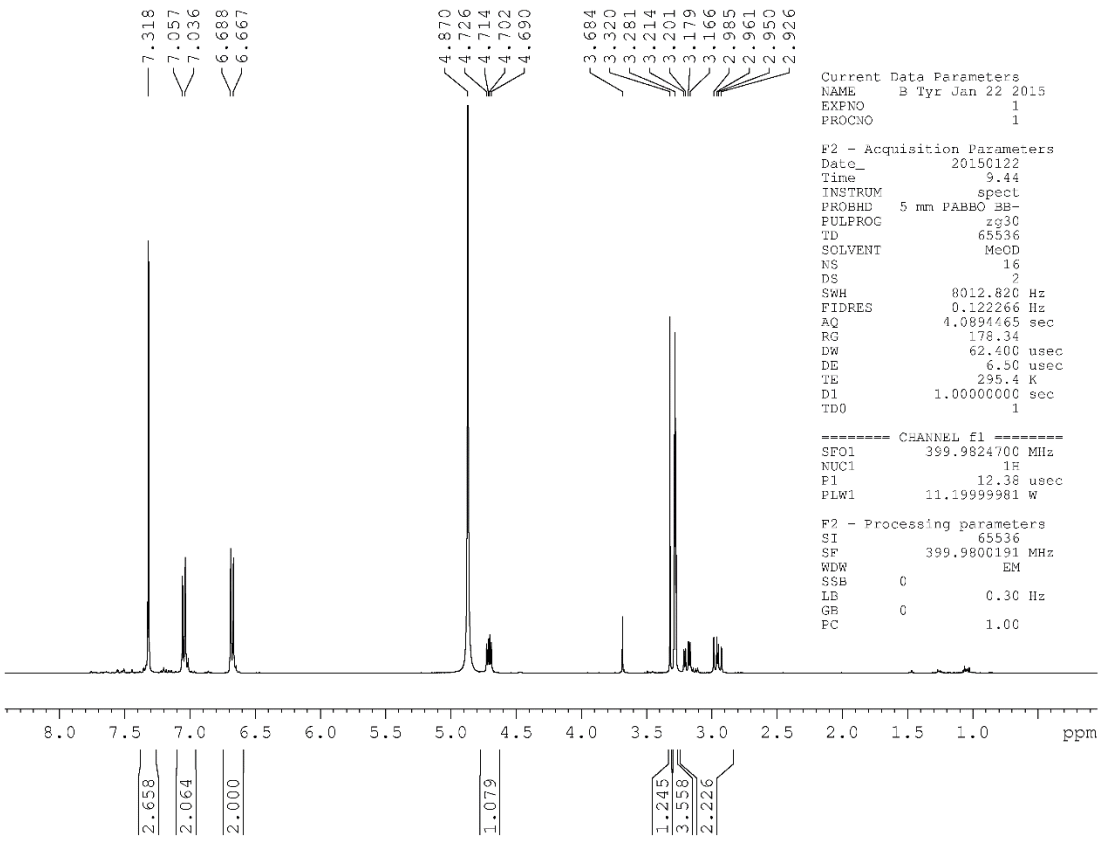


Figure 6.11. 1D ^{1}H-NMR (400 MHz) of **B-Tyr** in MeOD.

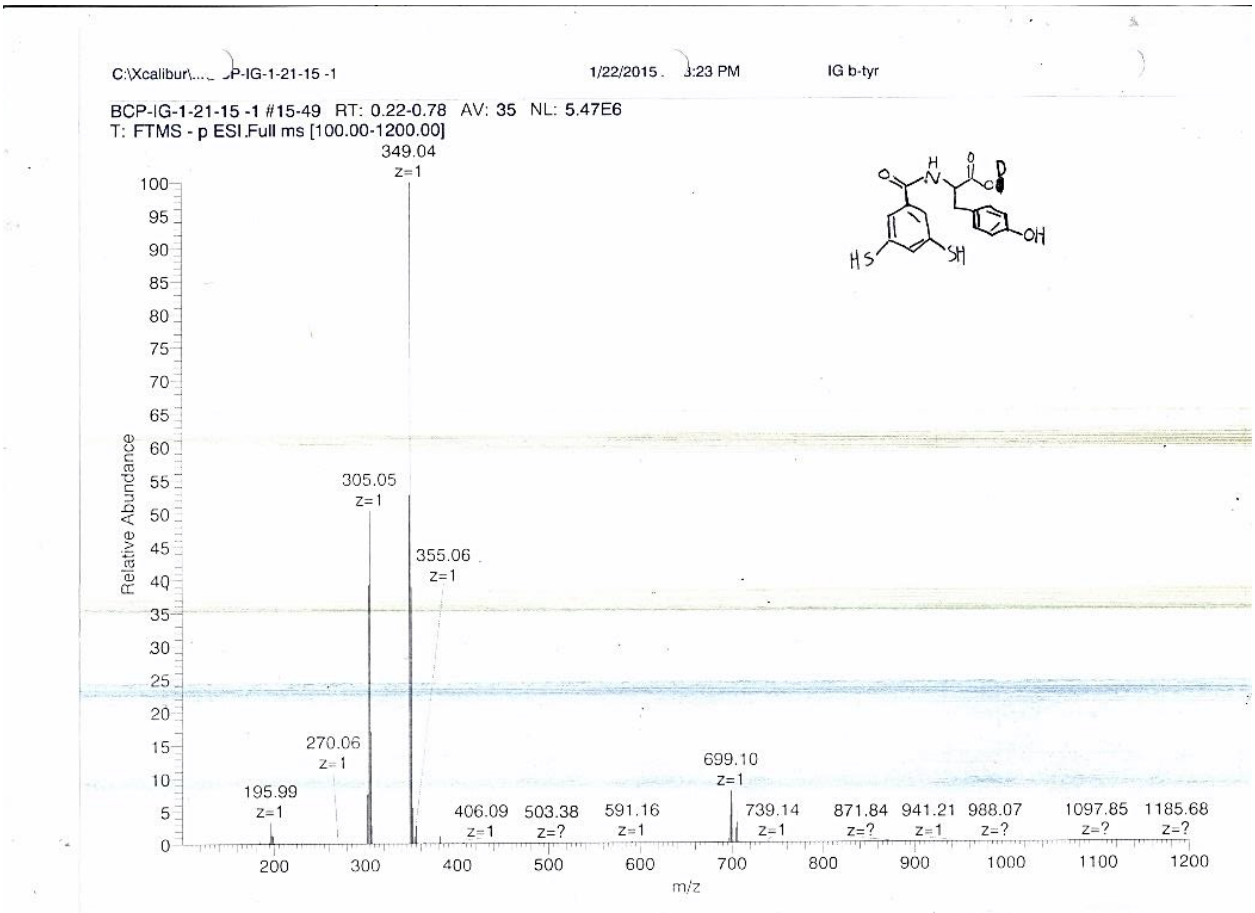


Figure 6.12. High resolution mass spectrum of **B-Tyr**. MS (calculated): 349.04 [M-2H+D]. MS (observed, ESI-): 349.04 [M-2H+D]. Sample was run in MeOD, which lead to H/D exchange.

Synthesis of B-Trp

A cocktail of TFA (10 mL), TIPS (0.5 mL), and CH_2Cl_2 (2 mL) was added to Trt B-Phe (0.40 g, 0.46 mmol) in a flask under a nitrogen atmosphere. The reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, solvent was removed by bubbling with nitrogen. The crude was triturated in hexanes (3 X 5 mL) and dried under vacuum to give a crude solid. A solution of LiOH monohydrate (0.02 g, 4.7 mmol) in H_2O was degassed by bubbling with nitrogen for 2 h. The crude solid was dissolved in 5 mL of the LiOH solution, and the reaction was allowed to stir overnight. The product was precipitated by addition of 5 mL of 1N HCl and collected by vacuum filtration. The crude product was used without

further purification (0.149 g, 87% over two steps). ^{1}H NMR (MeOD, 400 MHz): 7.001-7.573 (m, 6H, CH); 6.966, 6.983 (s, 2H, CH); 6.810, 6.806 (s, 1H, NH); 4.762 (m, 1H, CH); 3.452, 3.428 (m, 2H, CH). MS (calculated): 373.06 [M-3H+2D]. MS (observed, ESI-): 373.07 [M-3H+2D]. Sample was run in MeOD, which lead to H/D exchange.

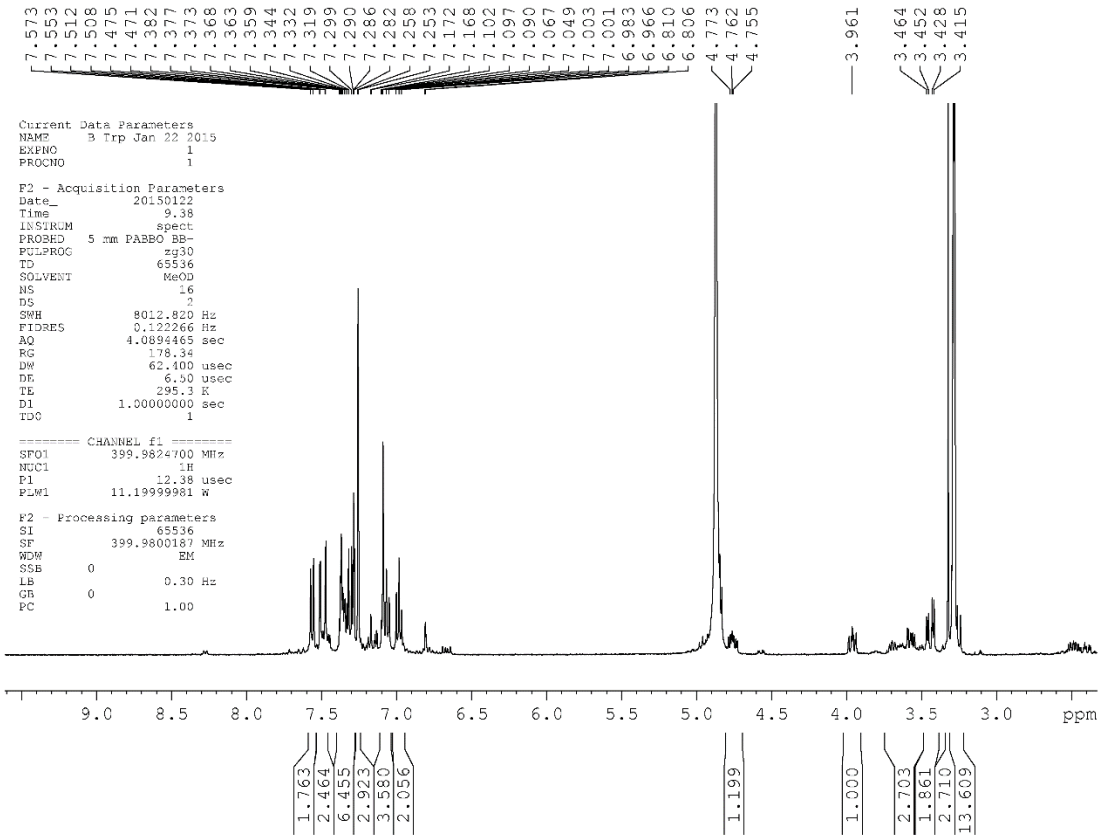


Figure 6.13. 1D ^{1}H-NMR (400 MHz) of **B-Trp** in MeOD.

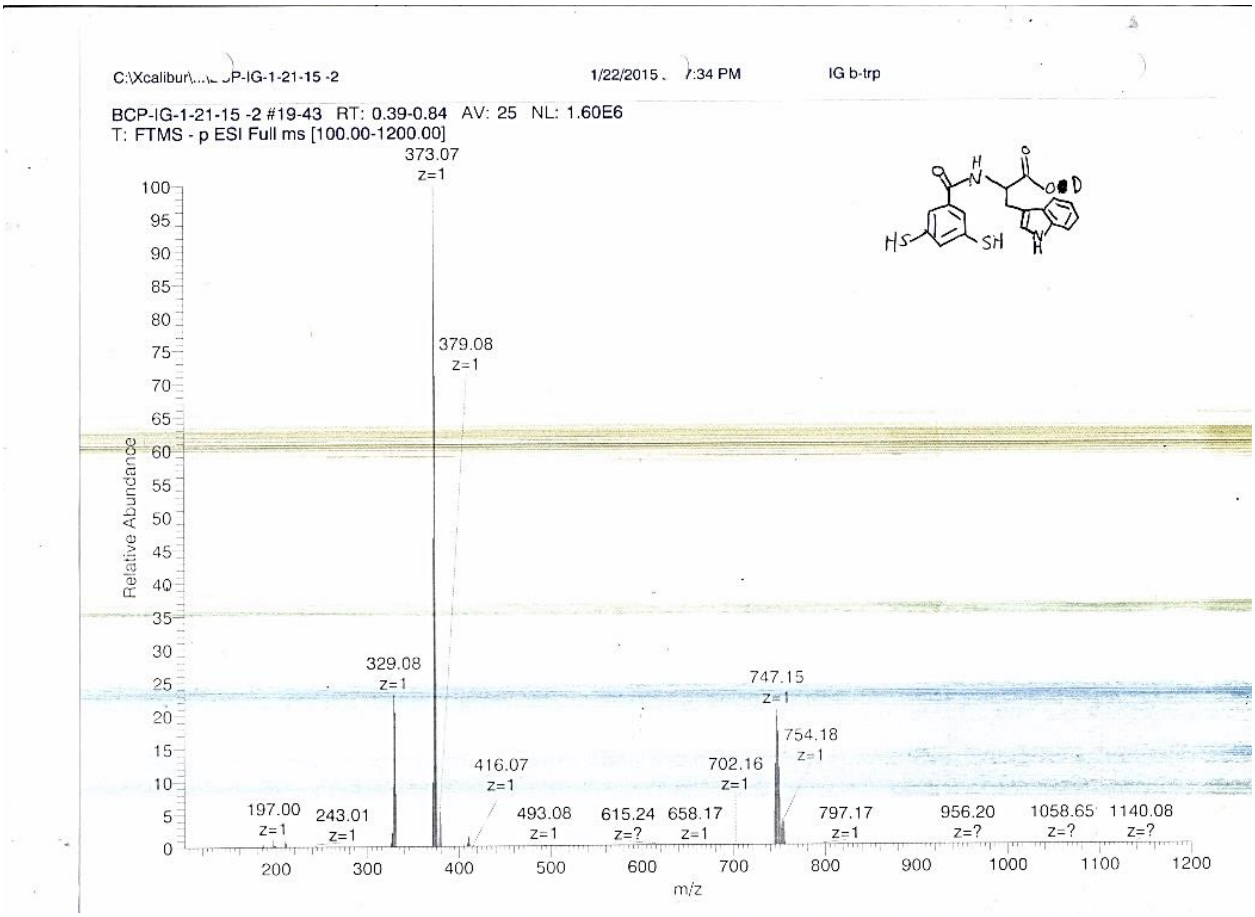


Figure 6.14. High resolution mass spectrum of **B-Trp**. MS (calculated): 373.06 [M-3H+2D]. MS (observed, ESI-): 373.07 [M-3H+2D]. Sample was run in MeOD, which lead to H/D exchange.

Synthesis of B-Asp

A cocktail of TFA (9.65 mL), TIPS (0.25 mL), and CH_2Cl_2 (2.0 mL) was added to Trt B-Asp (0.49 g, 0.61 mmol) in a flask under a nitrogen atmosphere. The reaction progress was monitored by TLC (35% Ethyl acetate in hexanes). After the reaction was complete, solvent was removed by bubbling with nitrogen. The crude was triturated in hexanes (3 X 5 mL) and dried under vacuum. The crude solid was dissolved in 10 mL of a solution of 10% KOH in 10:1 H_2O:MeOH (degassed by bubbling N_2 for 2 h) and allowed to stir overnight for 24 h. The product was used without further purification (0.082 g, 45% over two steps). ^{1}H NMR (MeOD,

400 MHz): 7.368 (s, 2H, CH); 7.272 (s, 1H, CH); 3.831 (m, 1H, CH), 2.830, 2.871 (m, 2H, CH).

MS calculated: 300.00 [M-H]. HRMS observed: 300.00037 [M-H].

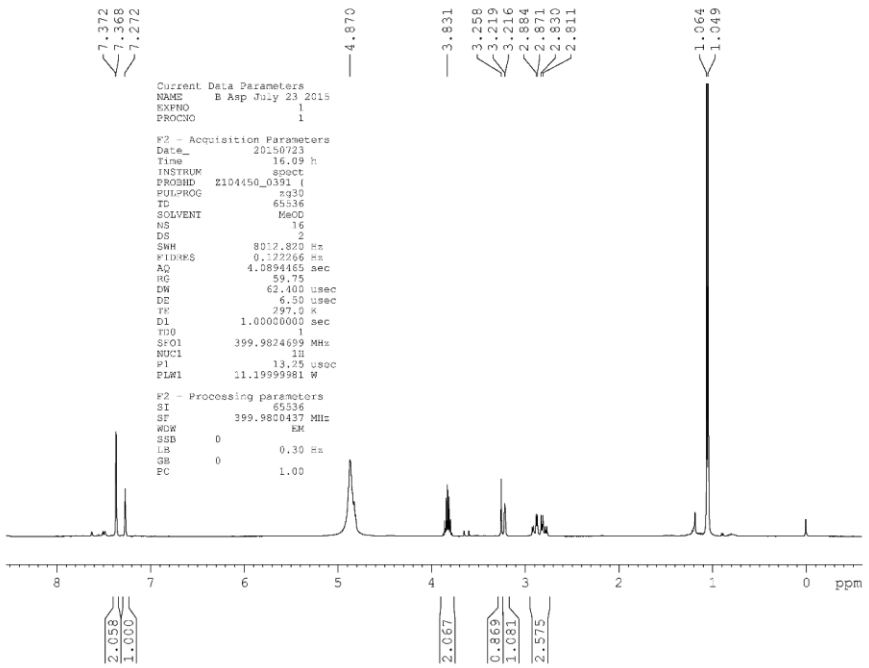


Figure 6.15. 1D ^{1}H-NMR (400 MHz) of **B-Asp** in MeOD.

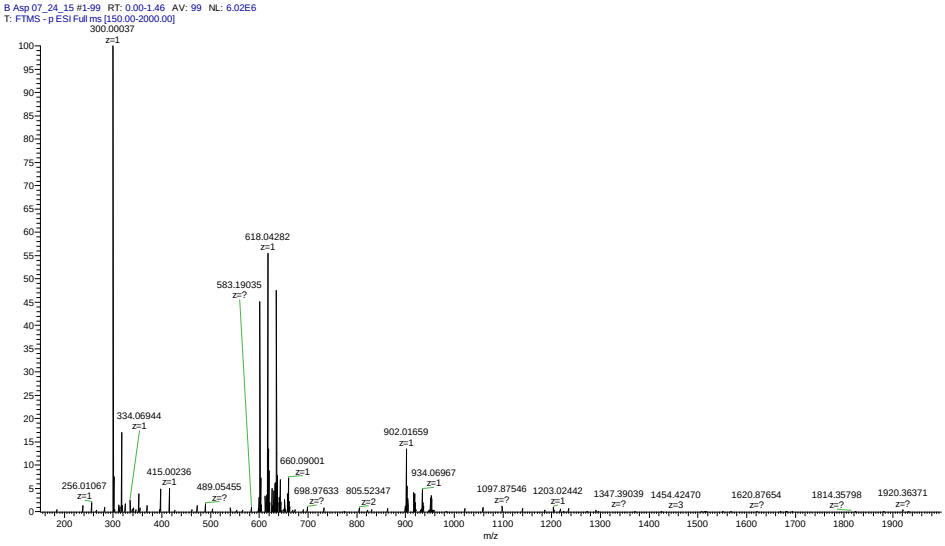


Figure 6.16. High resolution mass spectrum of **B-Asp** (ESI -). MS calculated: 300.00 [M-H]. MS observed: 300.00037 [M-H].

Peptide Synthesis

Peptides were synthesized on a 0.6 mmol scale via solid-phase peptide synthesis in a peptide synthesis flask using Fmoc protected amino acids. Coupling reagents were HOBt/HBTU in DMF (4 equiv each), with coupling times of 60 min, followed by rinsing with DMF, MeOH, and CH_2Cl_2. Deprotection was performed with 20% piperidine in DMF. All peptides were acylated at the N-terminus with a solution of 5% acetic anhydride and 6% 2,6-lutidine in DMF. Cleavage was performed using a cocktail of 95% TFA/2.5% triisopropylsilane/2.5% H_2O for 3 h. Peptides were purified by semi-preparative reverse-phase HPLC on a C18 column at a flow rate of 4 mL/min. Peptides were purified with a linear gradient of A and B (A: 95% H_2O/5% CH_3CN with 0.1% TFA, B: 95% CH_3CN/5% H_2O with 0.1 % TFA).

Methylated peptides were synthesized with 2 equivalents of Fmoc-Lys$(Me)_2$-OH HCl purchased from EMP Millipore and coupled for 5 h. The trimethyllysine containing peptides were synthesized by reacting corresponding dimethylated peptides (0.06 mmol scale) prior to cleavage from the resin with MTBD (7-Methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene, 10 equiv.) and methyl iodide (10 equiv.) in DMF (5 mL) for 5 h with bubbling nitrogen in a peptide flask. After washing the resin with DMF and CH_2Cl_2 and drying, the peptide was cleaved as normal.

Peptides:

H3K9: Ac-WGGGQTARKSTG-NH_2

H3K9me2: Ac-WGGGQTARKme2STG-NH_2

H3K9me3: Ac-WGGGQTARKme3STG-NH_2

Synthesis of A_2B-Phe, A_2B-Tyr, A_2B-Trp, and A_2B-Asp

DCLs were prepared on a preparative scale using butyltrimethylammonium bromide as a guest for templating the formation of the desired receptor. For the aromatic amino monomers, each preparative library contained 2.5 mM monomer **A**, 2.5 mM of the amino acid-functionalized monomer, and 10 mM guest dissolved in 50 mM sodium borate buffer at pH 8.5. For the preparation of **A_2B-Asp**, the library contained 5 mM monomer **A**, 5 mM of the amino acid-functionalized monomer, and 50 mM guest dissolved in 50 mM sodium borate buffer at pH 8.5. Prep scale DCLs were allowed to equilibrate for a week before purification by semi-preparative reversed-phase HPLC (solvent A: 10 mM NH_4OAc in H_2O; solvent B: 10 mM NH_4OAc in 9:1 CH_3CN:H_2O). For **A_2B-Phe**, **A_2B-Tyr**, and **A_2B-Trp**, each receptor was purified as a mixture of racemic and meso isomers due to peak overlap of the isomers. For **A_2B-Asp**, two isomers were purified and isolated. All receptors were purified using the following gradient: 0% B to 35% B over 1 min, 35% B to 40% B over 5 min, 40% B to 55% B over 30 min, and 55% B to 100% B over 30 min. Collected fractions were lyophilized to give off-white solids.

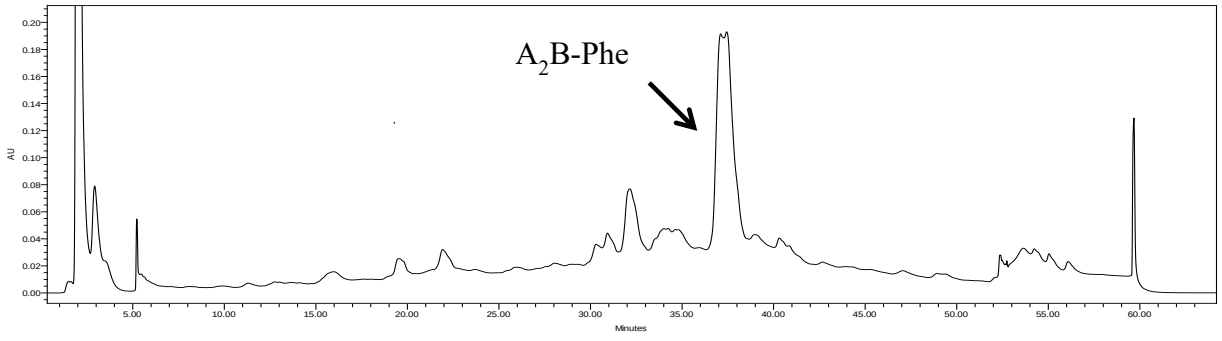


Figure 6.17. Semi-preparative reversed-phase HPLC traces of the purification of **A_2B-Phe** from a preparative scale DCL. Absorbance measured at 214 nm.

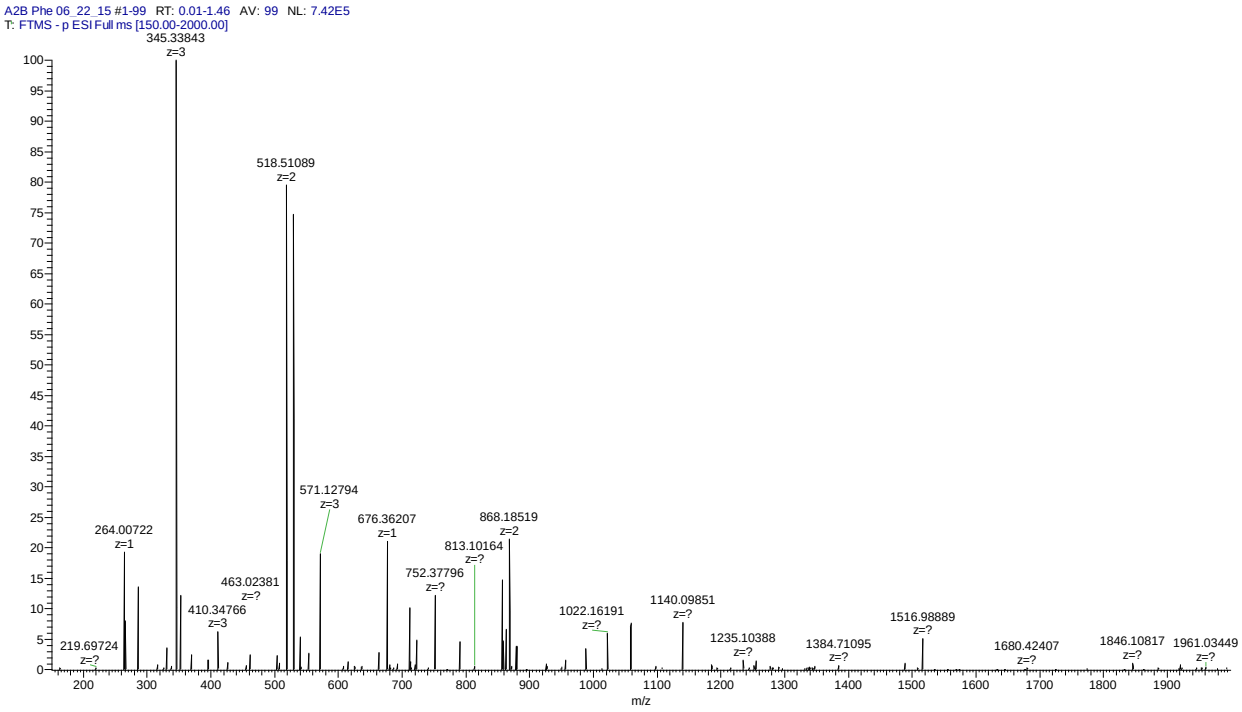


Figure 6.18. High resolution mass spectrum of **A$_2$B-Phe** (ESI -). MS calculated: 518.51 [M-2H], 345.34 [M-3H]. MS observed: 518.51089 [M-2H], 345.33843 [M-3H].

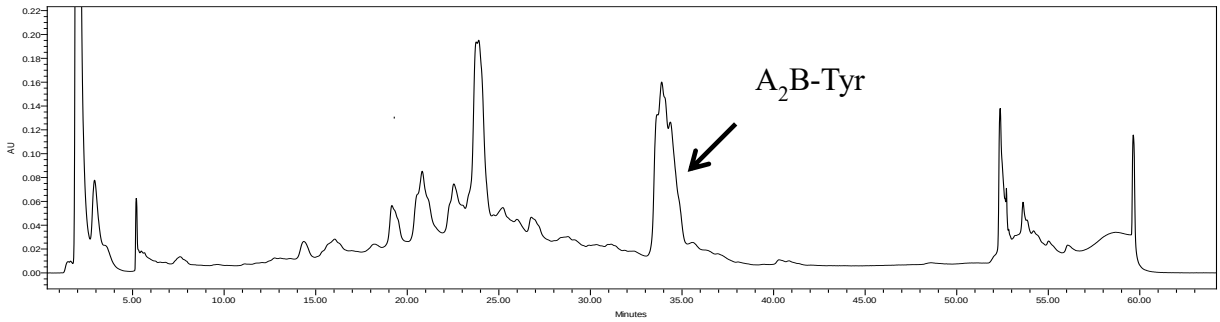


Figure 6.19. Semi-preparative reversed-phase HPLC traces of the purification of **A$_2$B-Tyr** from a preparative scale DCL. Absorbance measured at 214 nm.

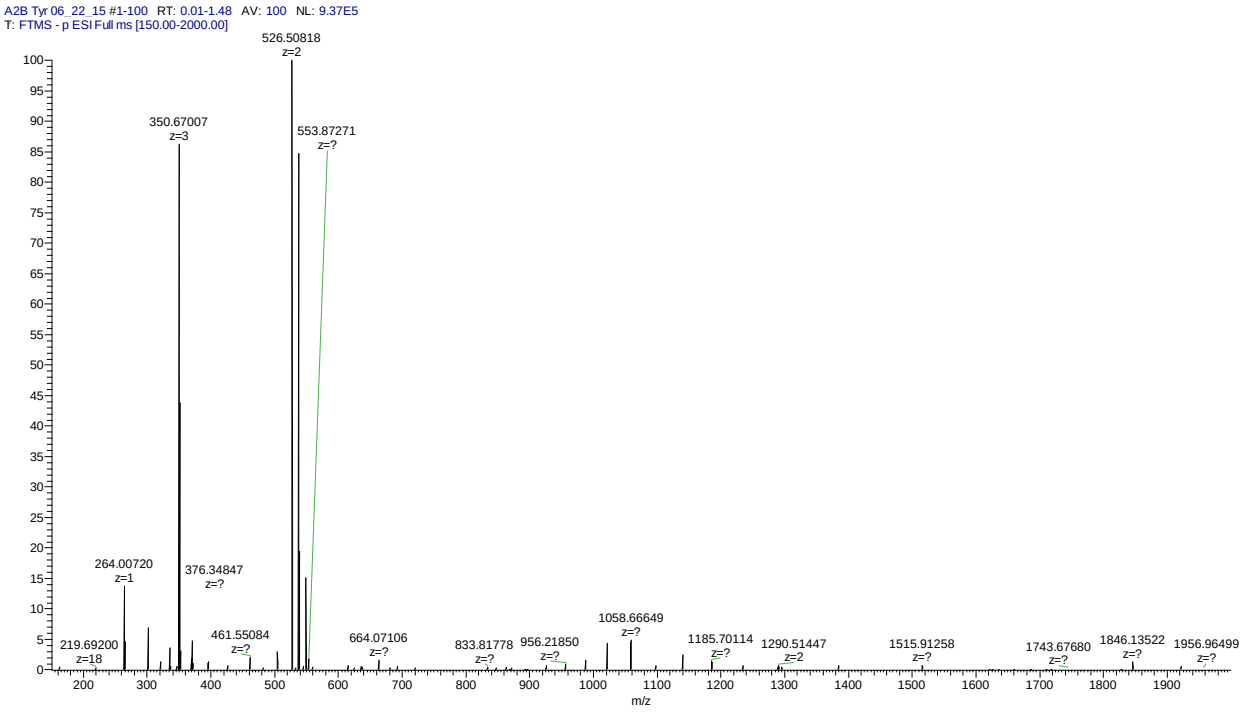


Figure 6.20. High resolution mass spectrum of **A_2B-Tyr** (ESI -). MS calculated: 526.51 [M-2H], 350.67 [M-3H]. MS observed: 526.50818 [M-2H], 350.67007 [M-3H].

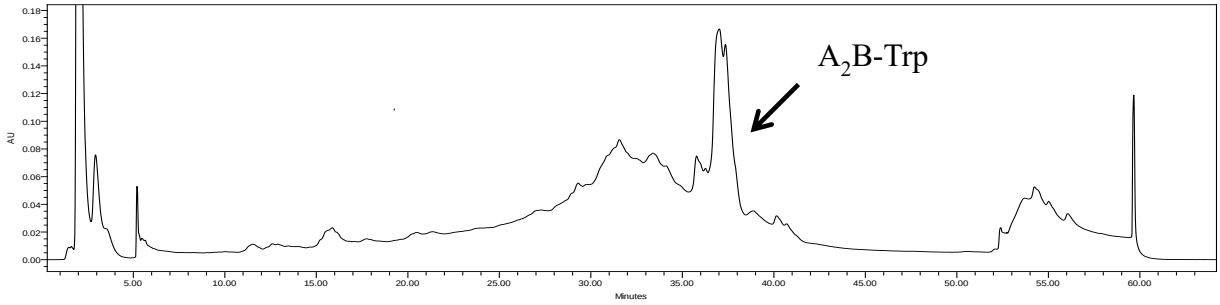


Figure 6.21. Semi-preparative reversed-phase HPLC traces of the purification of **A_2B-Trp** from a preparative scale DCL. Absorbance measured at 214 nm.

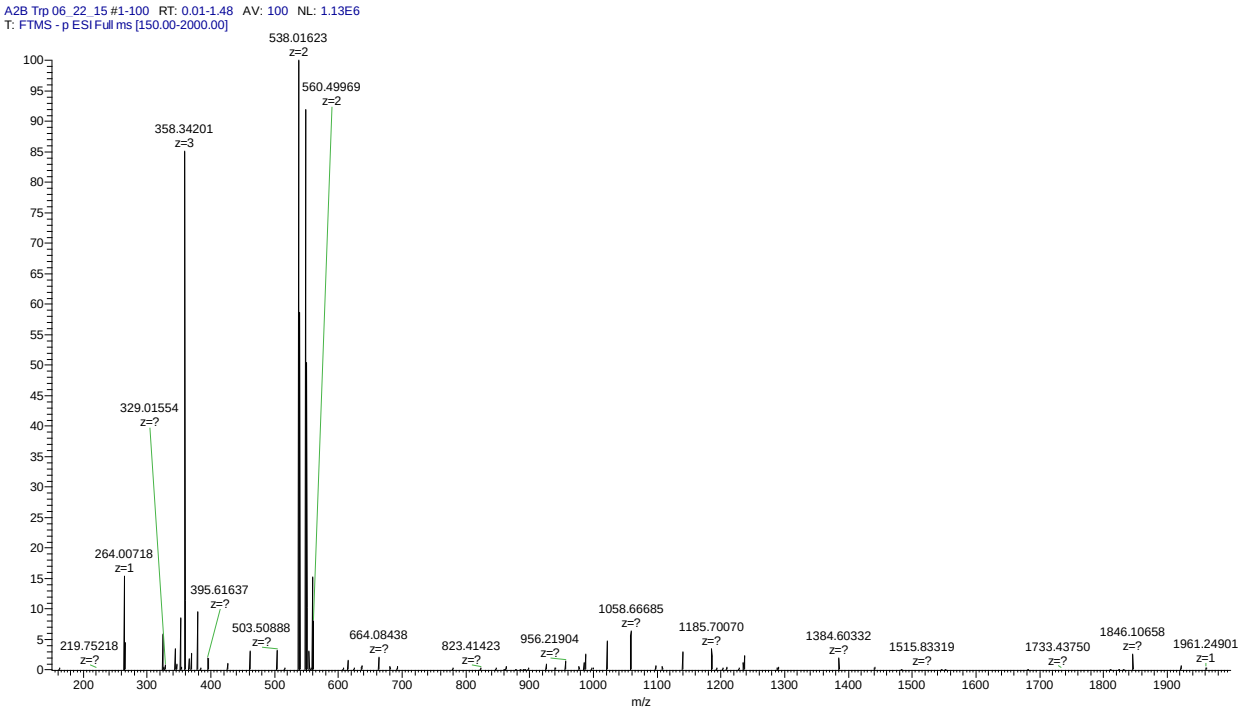


Figure 6.22. High resolution mass spectrum of **A_2B-Trp** (ESI -). MS calculated: 538.02 [M-2H], 358.34 [M-3H]. MS observed: 538.01623 [M-2H], 358.34201 [M-3H].

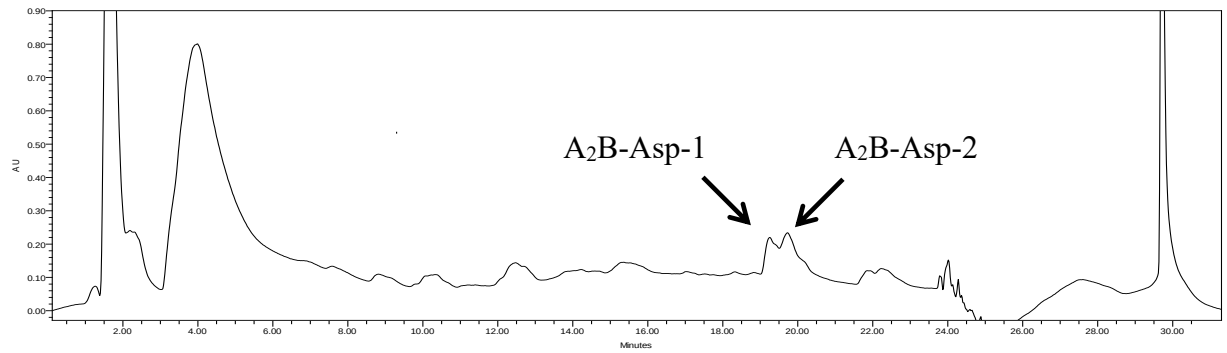


Figure 6.23. Semi-preparative reversed-phase HPLC traces of the purification of **A_2B-Asp** from a preparative scale DCL. Absorbance measured at 214 nm.

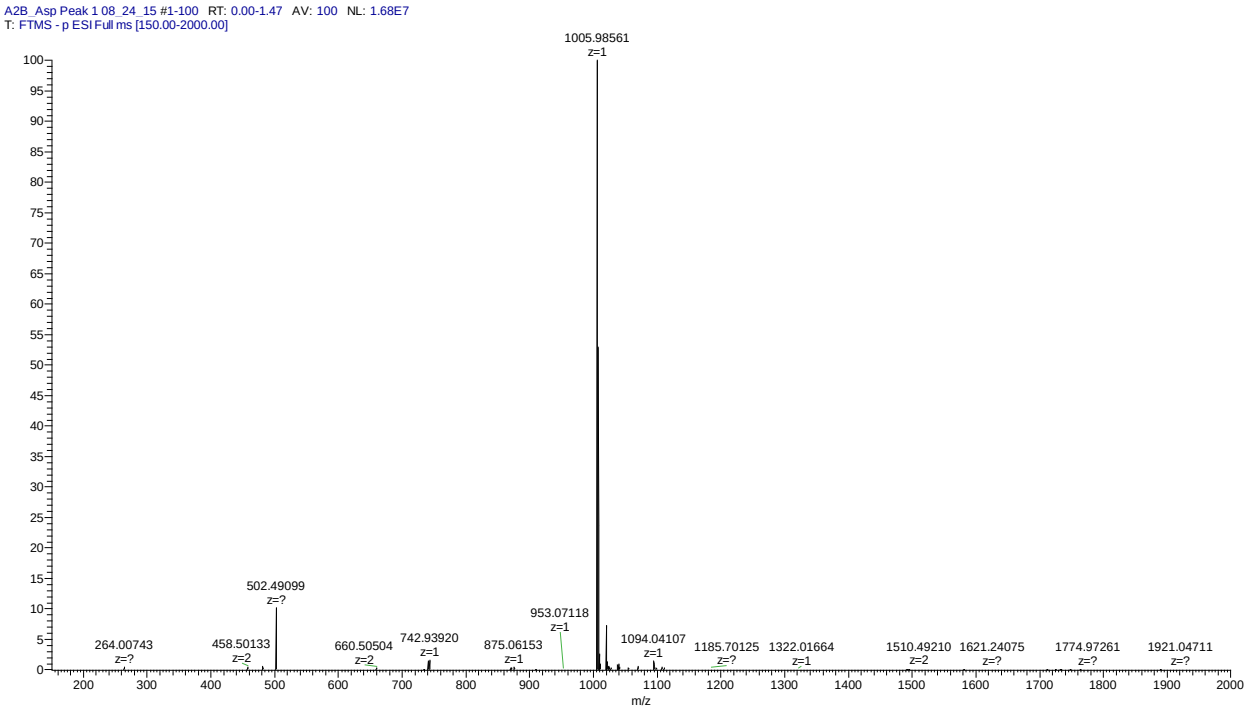


Figure 6.24. High resolution mass spectrum of **A$_2$B-Asp-1** (ESI -). MS calculated: 1005.99 [M-H], 502.49206 [M-2H]. MS observed: 1005.98561 [M-H], 502.49099 [M-2H].

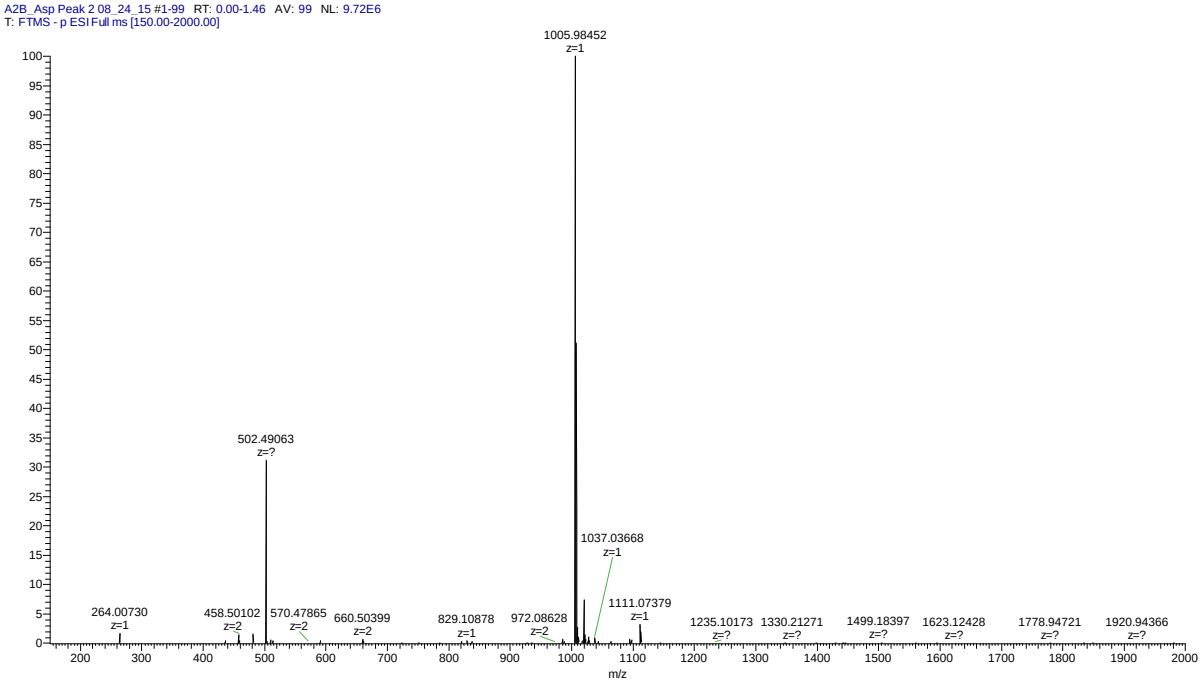


Figure 6.25. High resolution mass spectrum of **A$_2$B-Asp-2** (ESI -). MS calculated: 1005.99 [M-H], 502.49206 [M-2H]. MS observed: 1005.98452 [M-H], 502.49063 [M-2H].

Extinction Coefficient Determination for A_2B-Phe, A_2B-Trp, and A_2B-Tyr

Receptors were purified using reversed phase-HPLC with 10 mM NH_4OAc as the mobile phase additive. After purification, they were lyophilized for five days to remove volatile NH_4OAc salts. The dried compound was then dissolved in anhydrous methanol and filtered using a 0.33 μm filter to remove any remaining salts. Methanol was evaporated and the receptors were further dried under vacuum. The mass was accurately determined, and a stock solution of receptor (1.0 mM) in 10 mM sodium borate buffer, pH 8.5 was prepared. The stock solution was serially diluted (80:20, receptor: buffer) to give 8 concentrations. The absorbance at 280 nm was measured for each concentration, which was then plotted against the concentration. The extinction coefficient of the receptors was determined from the slope of the line of regression using Beer's law and was found to be 1300 $M^{-1}cm^{-1}$ for **A_2B-Phe,** 1600 $M^{-1}cm^{-1}$ for **A_2B-Trp**, and 800 $M^{-1}cm^{-1}$ for **A_2B-Tyr**. For **A_2B-Asp-1** and **A_2B-Asp-2**, the previously reported extinction coefficient for **A_2B** was used.[98]

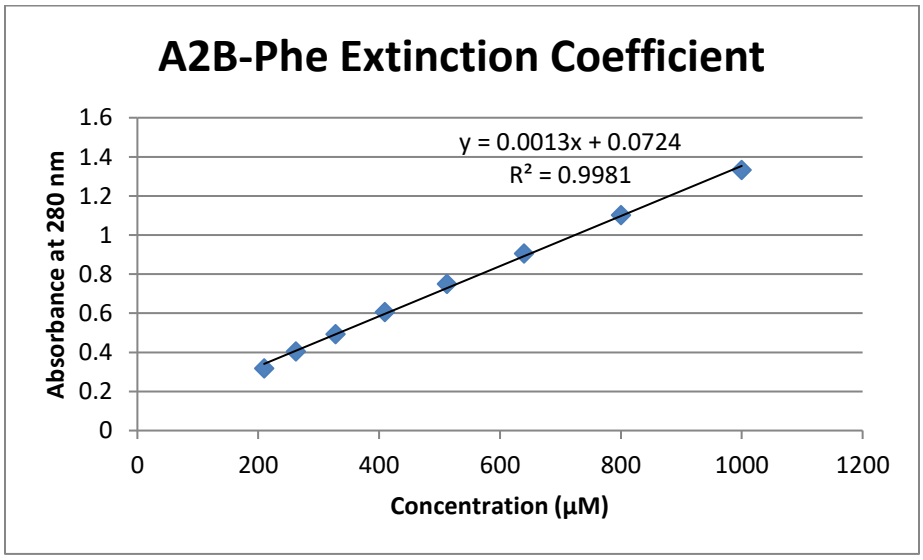


Figure 6.26. Extinction coefficient determination of **A_2B-Phe** from the slope of the line of regression.

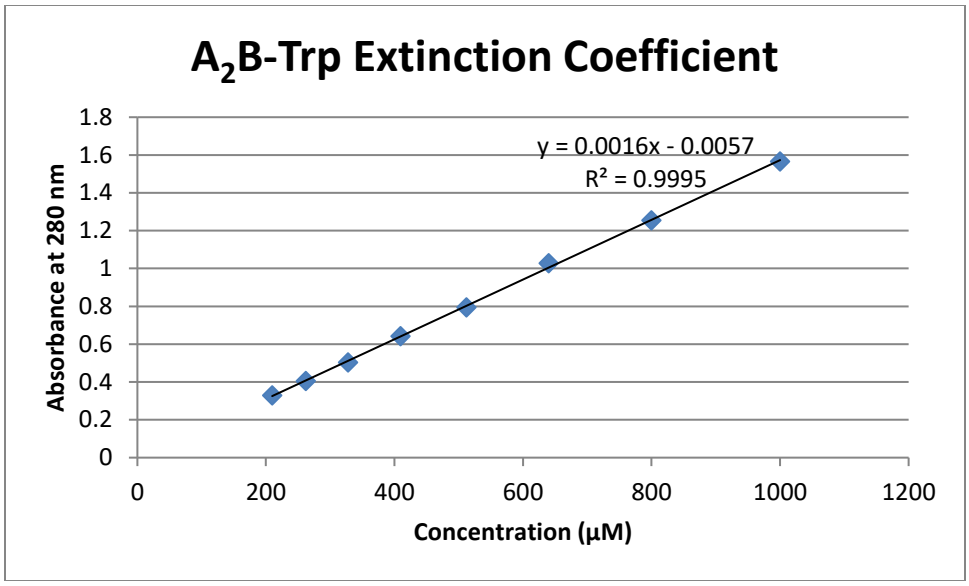


Figure 6.27. Extinction coefficient determination of **A_2B-Trp** from the slope of the line of regression.

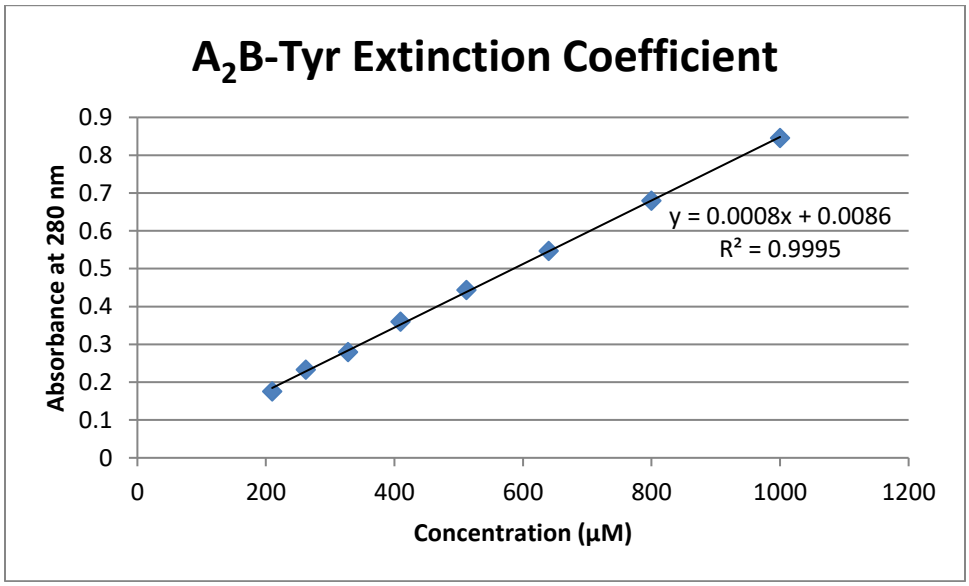


Figure 6.28. Extinction coefficient determination of **A_2B-Tyr** from the slope of the line of regression.

Isothermal Titration Calorimetry

ITC binding experiments were conducted using a Microcal AutoITC200. Titrations were performed at 25 °C in 10 mM sodium borate buffer at pH 8.5. The concentration of receptor was determined by measuring the UV–vis absorbance at 300 nm or 315 nm, using a NanoDrop2000 with a xenon flash lamp, 2048 element linear silicon CCD array detector, and 1 mm path length. Solutions of 0.94–1.14 mM of peptide were titrated into an 81–98 μM solution of receptor, using 2.0 μL increments every 3 min. Heats of dilution were taken as an average of the last three points and subtracted prior to fitting. Binding curves were produced using the supplied Origin software and fit using one-site binding models.

Table 6.3. Thermodynamic binding data for the biding of **A_2B-Phe**, **A_2B-Trp**, **A_2B-Tyr**, **A_2B-Asp-1**, and **A_2B-Asp-2** to peptides shown in **Table 6.1** as measured by ITC.

Entry	*Receptor*	*Peptide*	*Charge*	*K_d [b] (μM)*	*SF[d]*	*ΔG[c] (kcal/mol)*	*ΔH[b] (kcal/mol)*	*ΔS (cal/mol/deg)*	*N*
1	A_2B[a]	H3K9	+2	22 ± 1	8.3	-6.38 ± 0.02	-	-	-
2	A_2B[a]	H3 K9me2	+2	6.3 ± 0.3	2.4	-7.10 ± 0.07	-	-	-
3	A_2B[a]	H3 K9me3	+2	2.6 ± 0.1	-	-7.63 ± 0.03	-	-	-
4	A_2B-Phe	H3K9	+2	7.9 ± 1.1	1.8	-6.96 ± 0.97	-5.53 ± 0.17	4.80	0.920
5	A_2B-Phe	H3K9me2	+2	6.5 ± 0.4	1.5	-7.07 ± 0.44	-10.14 ± 0.16	-10.3	0.726
6	A_2B-Phe	H3 K9me3	+2	4.4 ± 0.2	-	-7.31 ± 0.33	-10.72 ± 0.07	-11.4	0.808
7	A_2B-Trp	H3K9	+2	>16	3.2	>-6.56	-5.28 ± 0.47	4.30	0.624
8	A_2B-Trp	H3 K9me2	+2	7.2 ± 0.5	1.5	-7.02 ± 0.49	-8.99 ± 0.19	-6.61	0.486
9	A_2B-Trp	H3 K9me3	+2	4.9 ± 0.2	-	-7.24 ± 0.30	-10.75 ± 0.11	-11.8	0.547
10	A_2B-Tyr	H3K9	+2	8.6 ± 0.7	5.3	-6.91 ± 0.56	-7.93 ± 0.21	-3.42	0.430
11	A_2B-Tyr	H3 K9me2	+2	3.0 ± 0.2	1.9	-7.53 ± 0.50	-11.99 ± 0.20	-15.0	0.320
12	A_2B-Tyr	H3K9me3	+2	1.6 ± 0.1	-	-7.91 ± 0.49	-11.75 ± 0.07	-12.9	0.358
13	A_2B-Asp-1	H3K9	+2	6.3 ± 0.5	3.0	-7.10 ± 0.56	-8.23 ± 0.17	-3.80	0.516
14	A_2B-Asp-1	H3 K9me2	+2	2.3 ± 0.1	1.1	-7.69 ± 0.33	-11.88 ± 0.11	-14.1	0.413
15	A_2B-Asp-1	H3 K9me3	+2	2.1 ± 0.1	-	-7.76 ± 0.37	-11.05 ± 0.07	-11.0	0.484
16	A_2B-Asp-2	H3K9	+2	8.5 ± 0.8	2.3	-6.91 ± 0.65	-8.56 ± 0.31	-5.51	0.453
17	A_2B-Asp-2	H3K9me2	+2	3.5 ± 0.2	0.9	-7.44 ± 0.43	-11.88 ± 0.19	-14.9	0.345
18	A_2B-Asp-2	H3K9me3	+2	3.7 ± 0.3	-	-7.42 ± 0.60	-12.38 ± 0.23	-16.6	0.384

[a] Reported by Pinkin and Waters.[98] [b] Errors are from the error in the fit as determined by the Origin 7 software. [c] ΔG was calculated from the K_d. Errors are from the error in the fit as determined by the Origin 7 software [d] The selectivity factor is calculated as the K_d of Kmex (where x = 0, 2, or 3) divided by the K_d of Kme3.

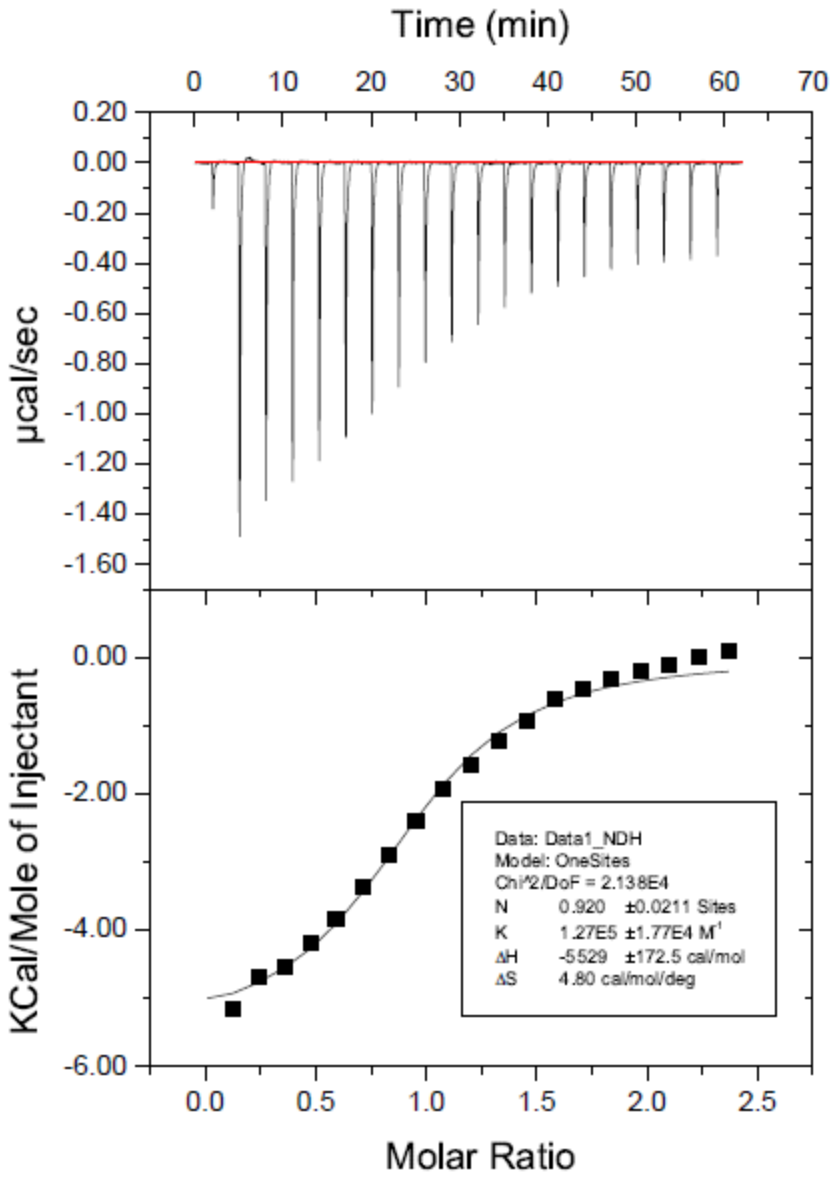


Figure 6.29. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into **A_2B-Phe** (82 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5.

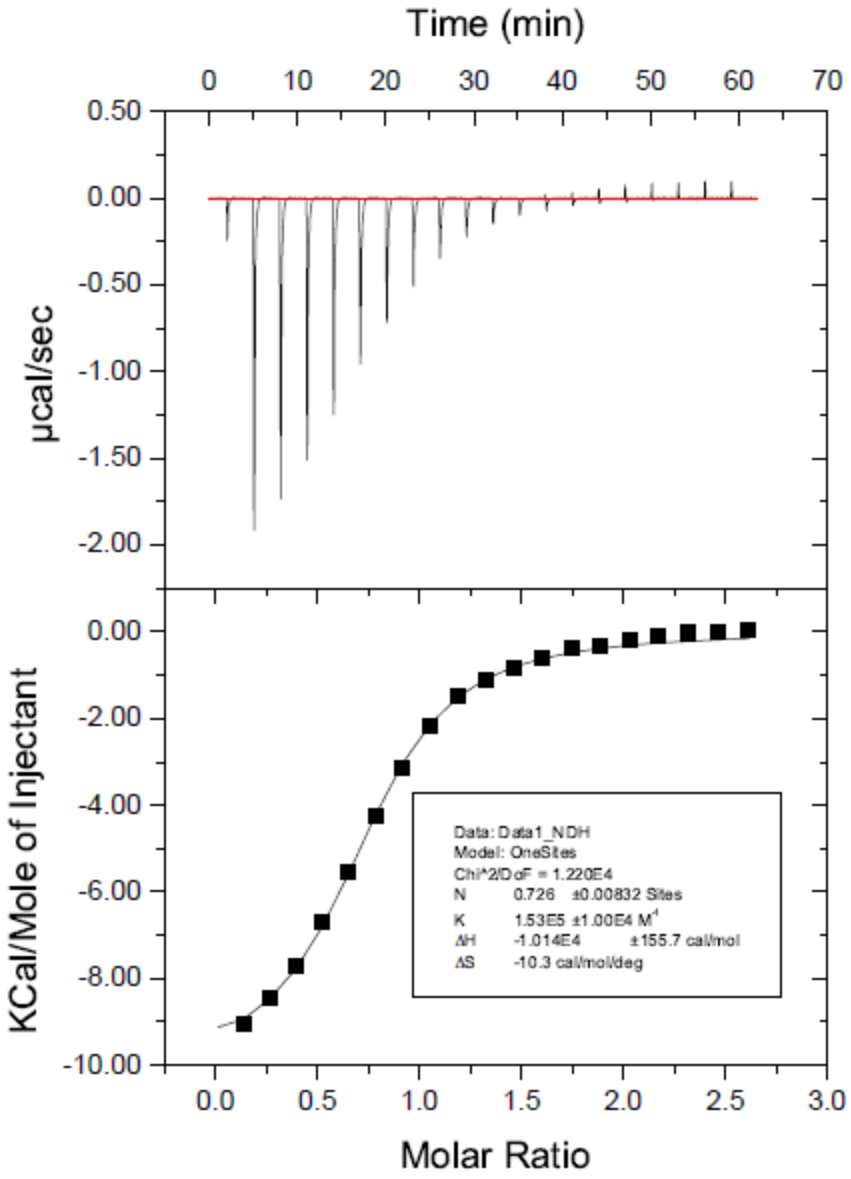


Figure 6.30. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into **A_2B-Phe** (82 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

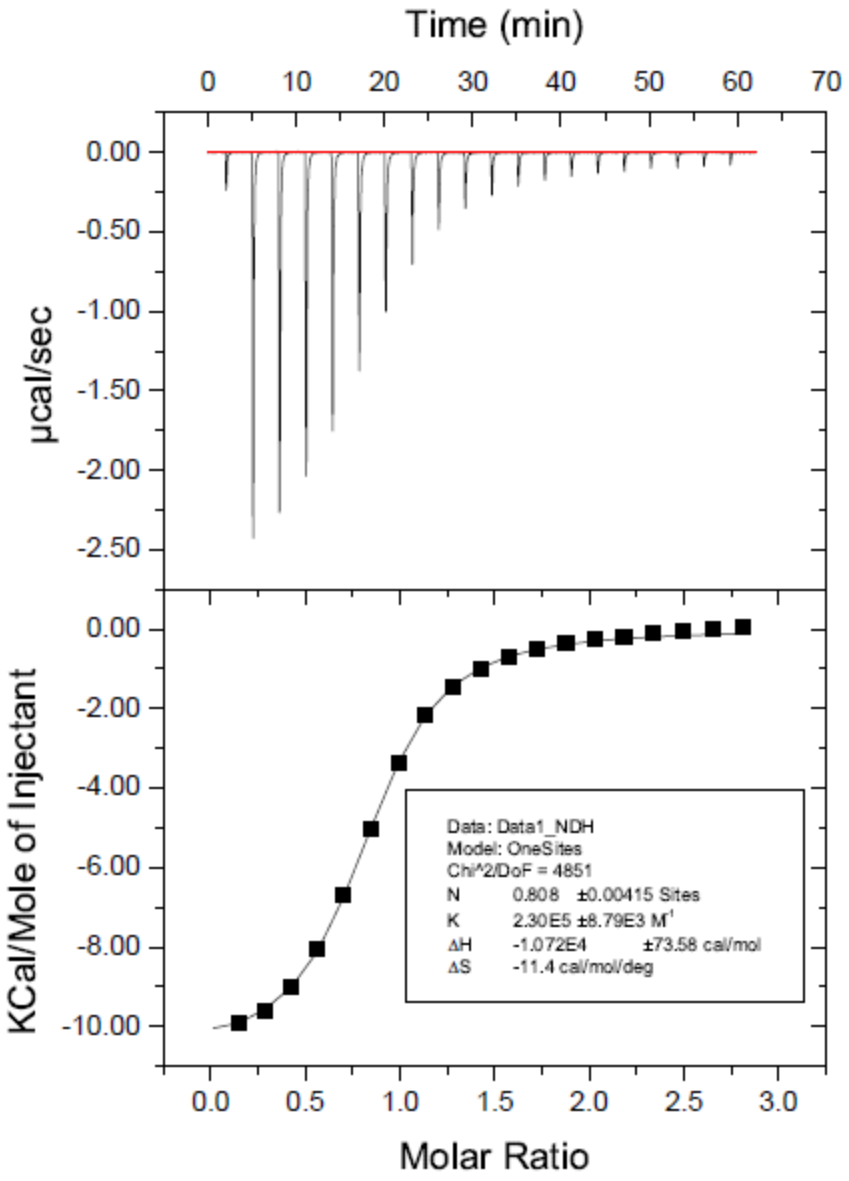


Figure 6.31. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into **A_2B-Phe** (82 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5.

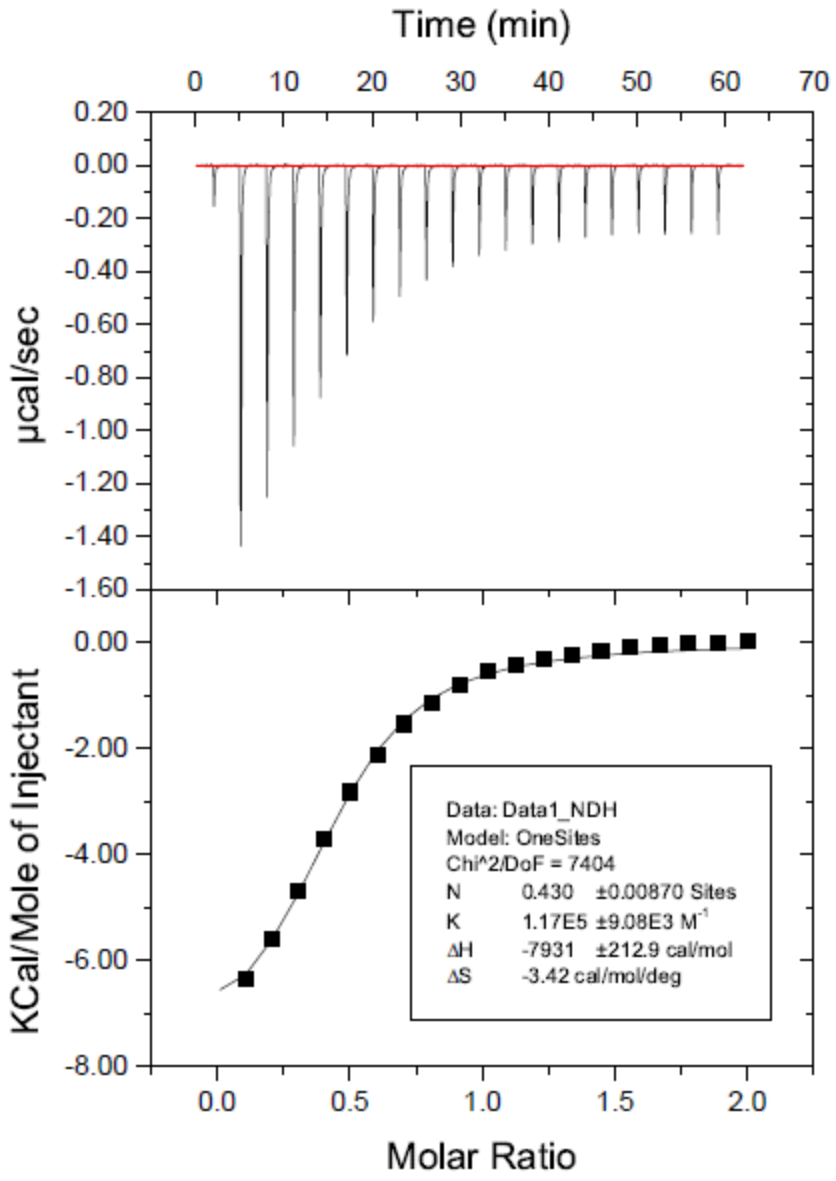


Figure 6.32. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into **A_2B-Tyr** (98 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

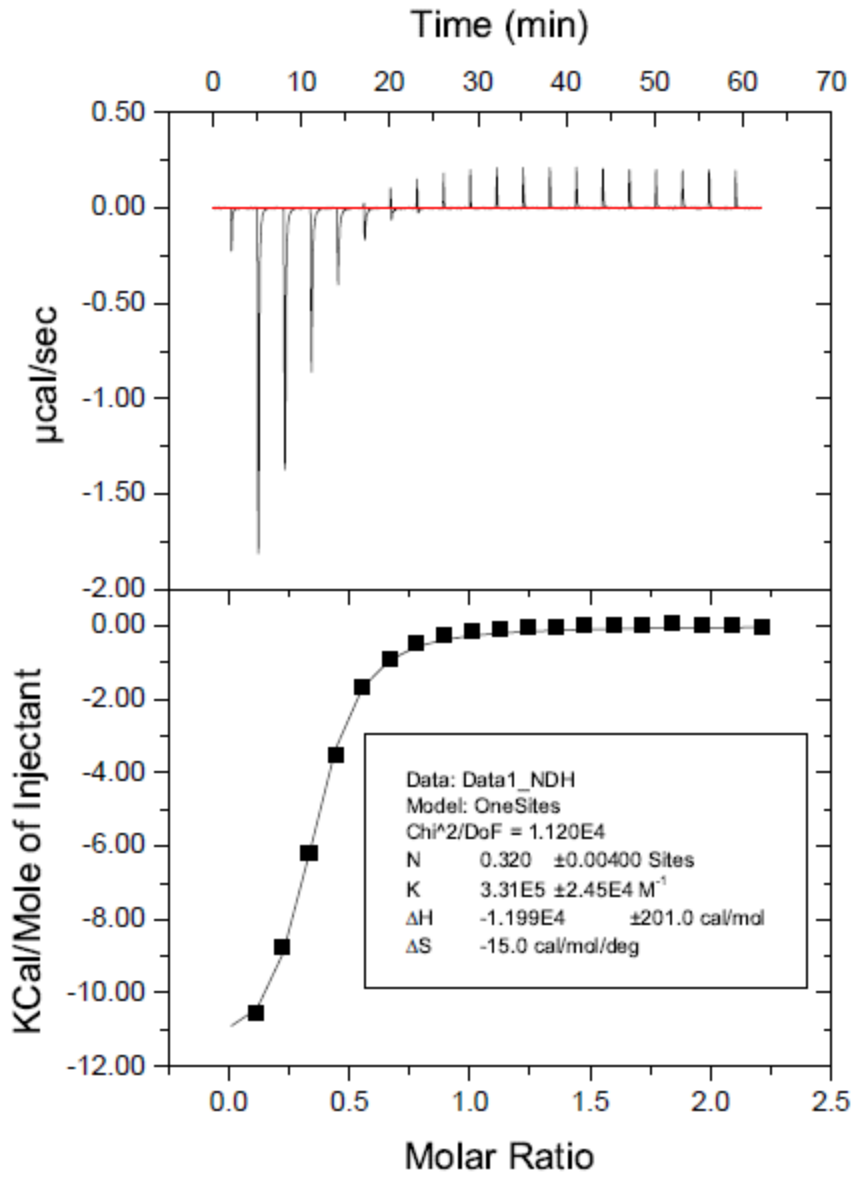


Figure 6.33. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into **A_2B-Tyr** (98 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5.

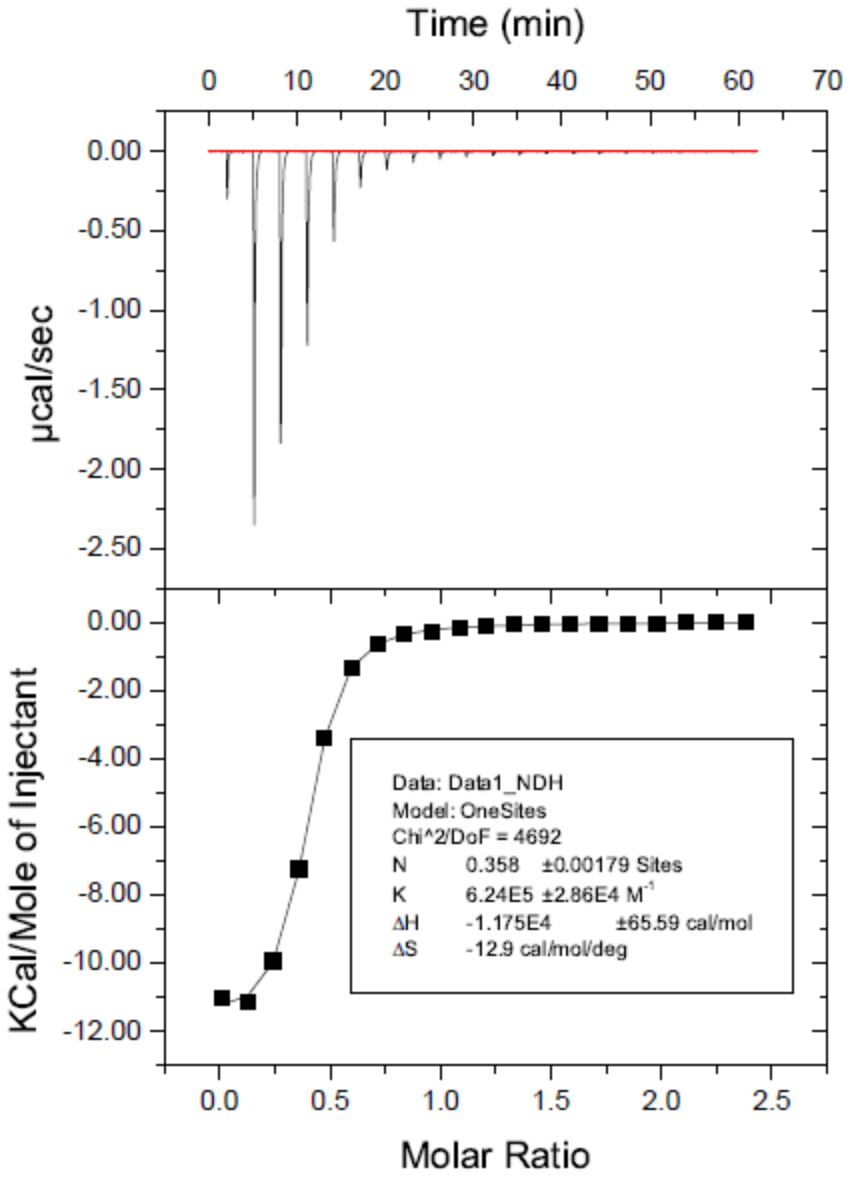


Figure 6.34. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into **A_2B-Tyr** (98 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

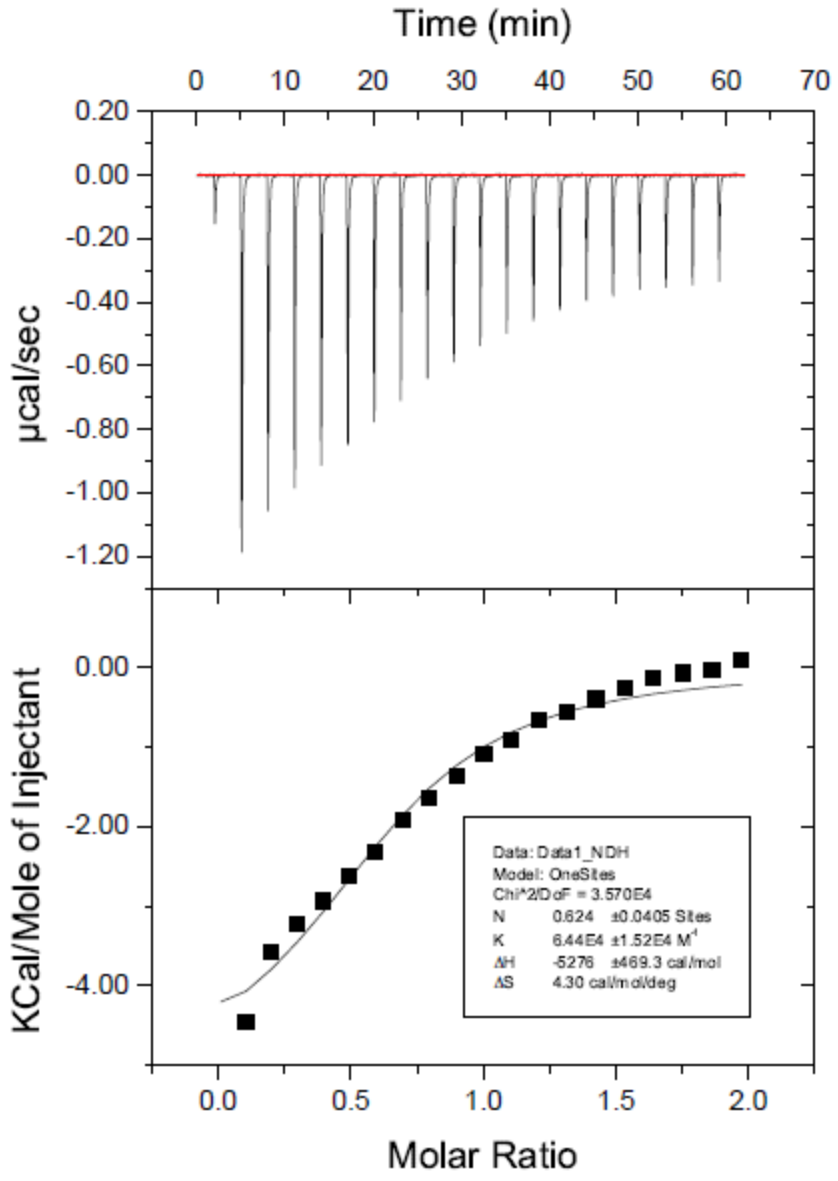


Figure 6.35. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.96 mM) titrated into **A_2B-Trp** (97 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5.

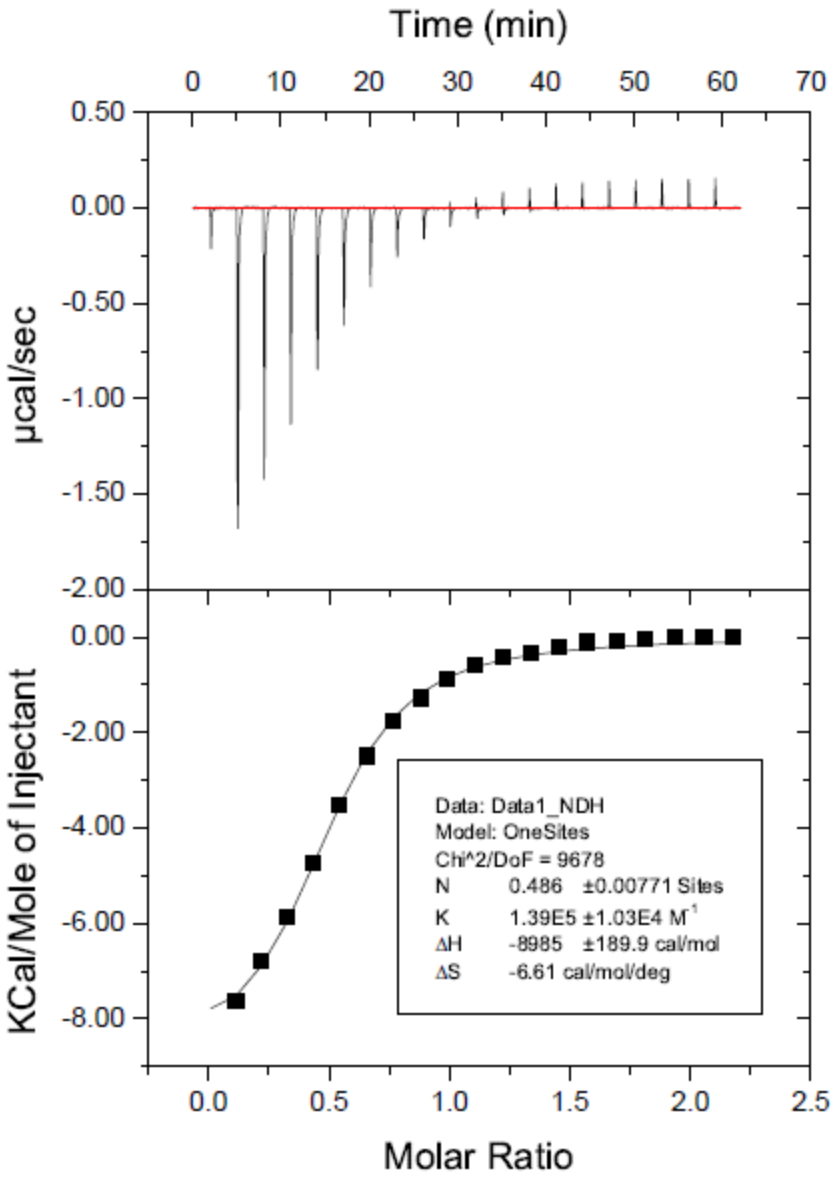


Figure 6.36. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.06 mM) titrated into **A_2B-Trp** (97 µM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

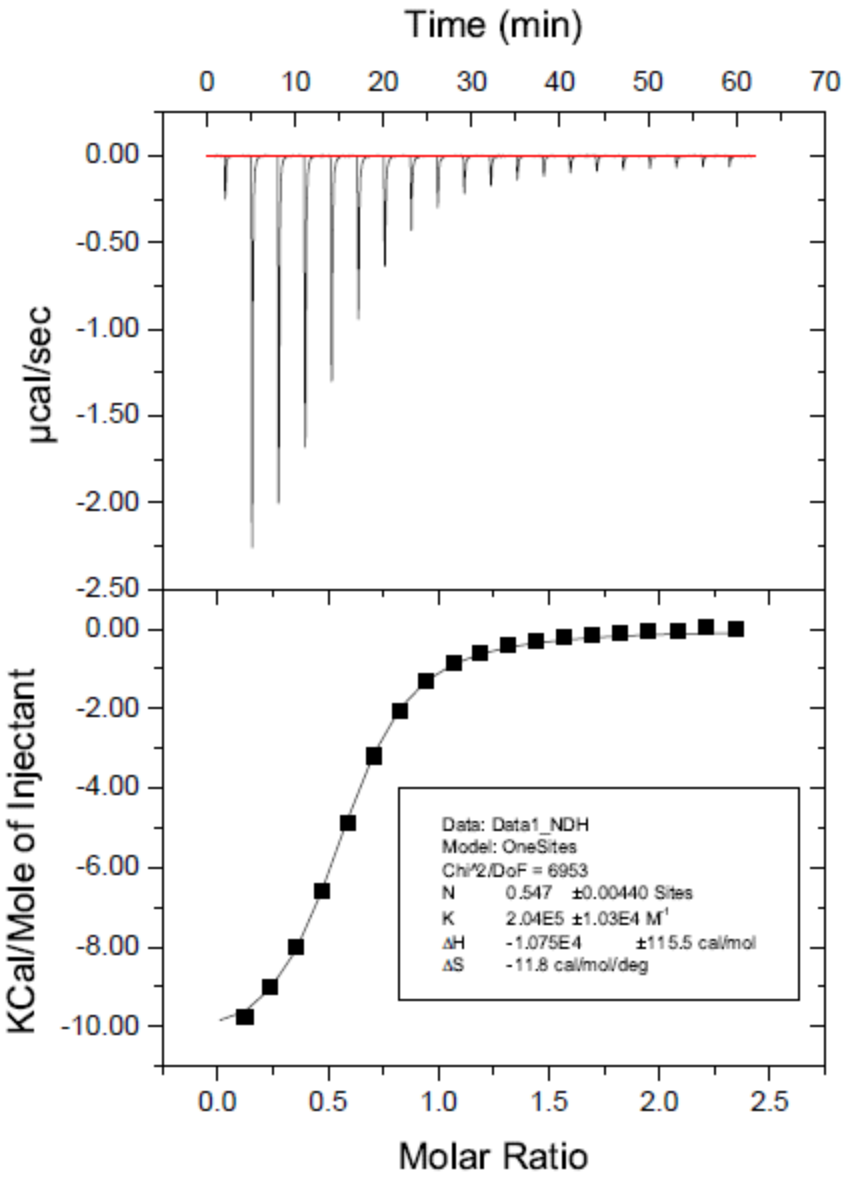


Figure 6.37. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.14 mM) titrated into **A_2B-Trp** (97 µM) at 25 ºC in 10 mM sodium borate buffer pH 8.5.

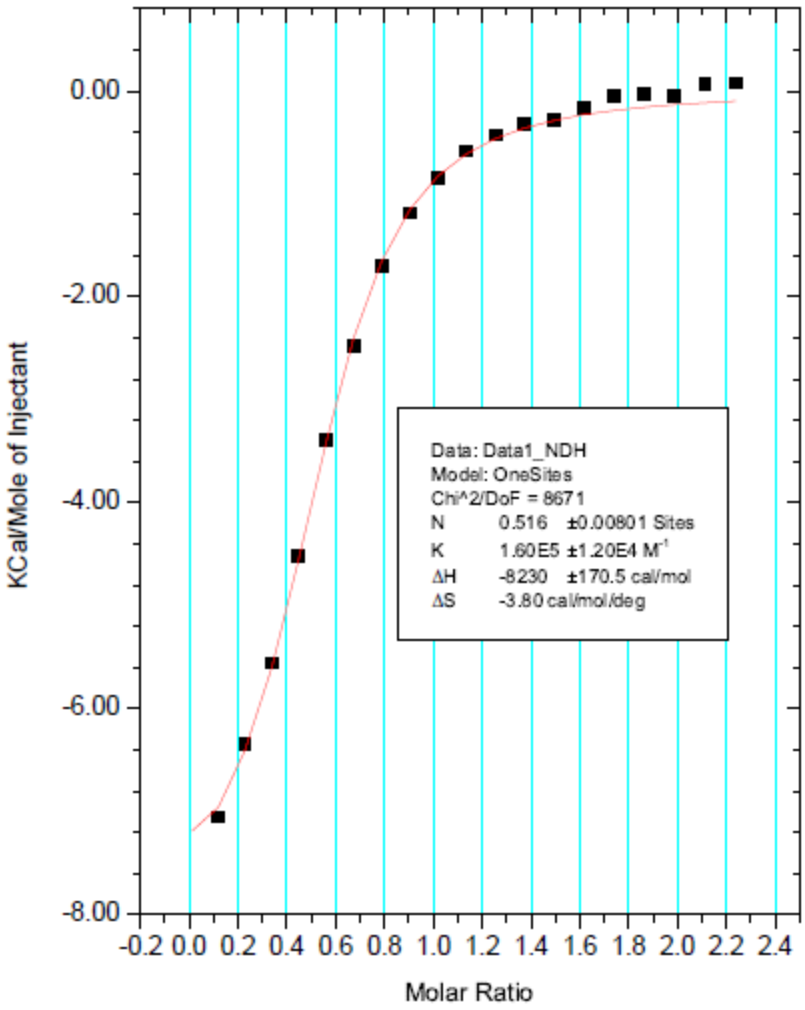


Figure 6.38. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.94 mM) titrated into **A_2B-Asp-1** (85 μM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

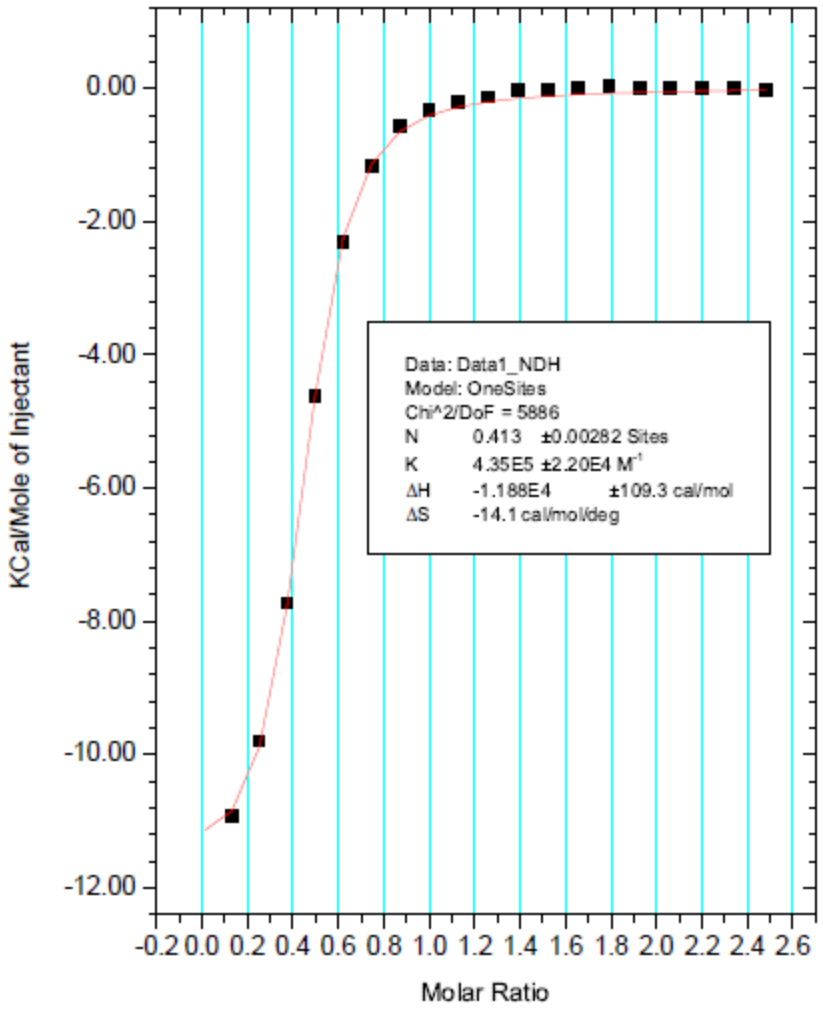


Figure 6.39. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.04 mM) titrated into **A_2B-Asp-1** (85 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5.

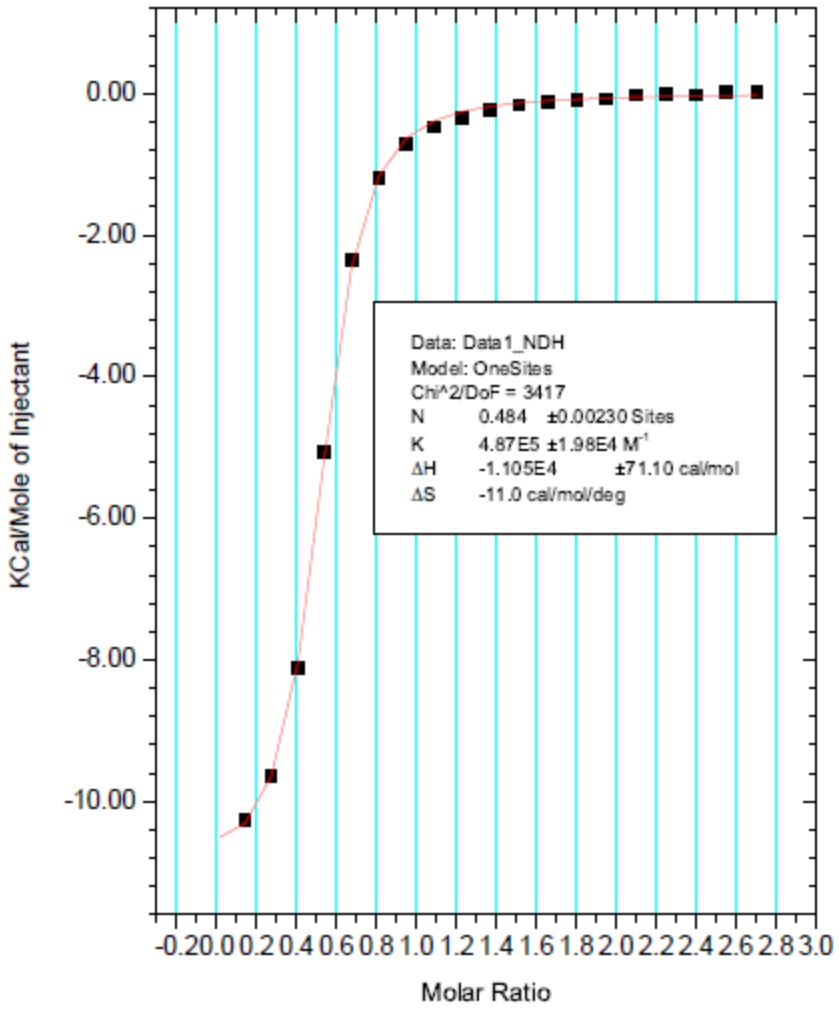


Figure 6.40. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.13 mM) titrated into **A_2B-Asp-1** (85 μM) at 25 ^oC in 10 mM sodium borate buffer pH 8.5.

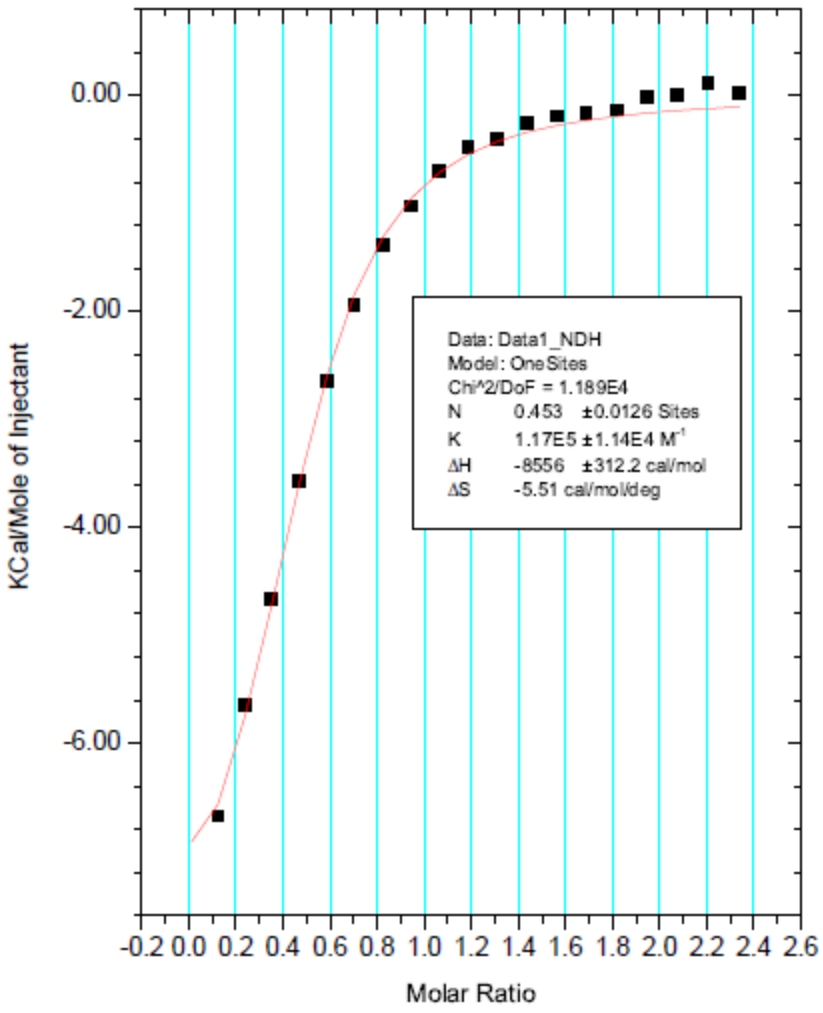


Figure 6.41. One trial of H3K9 (Ac-WGGGQTARKSTG-NH_2) (0.94 mM) titrated into **A_2B-Asp-2** (81 μM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

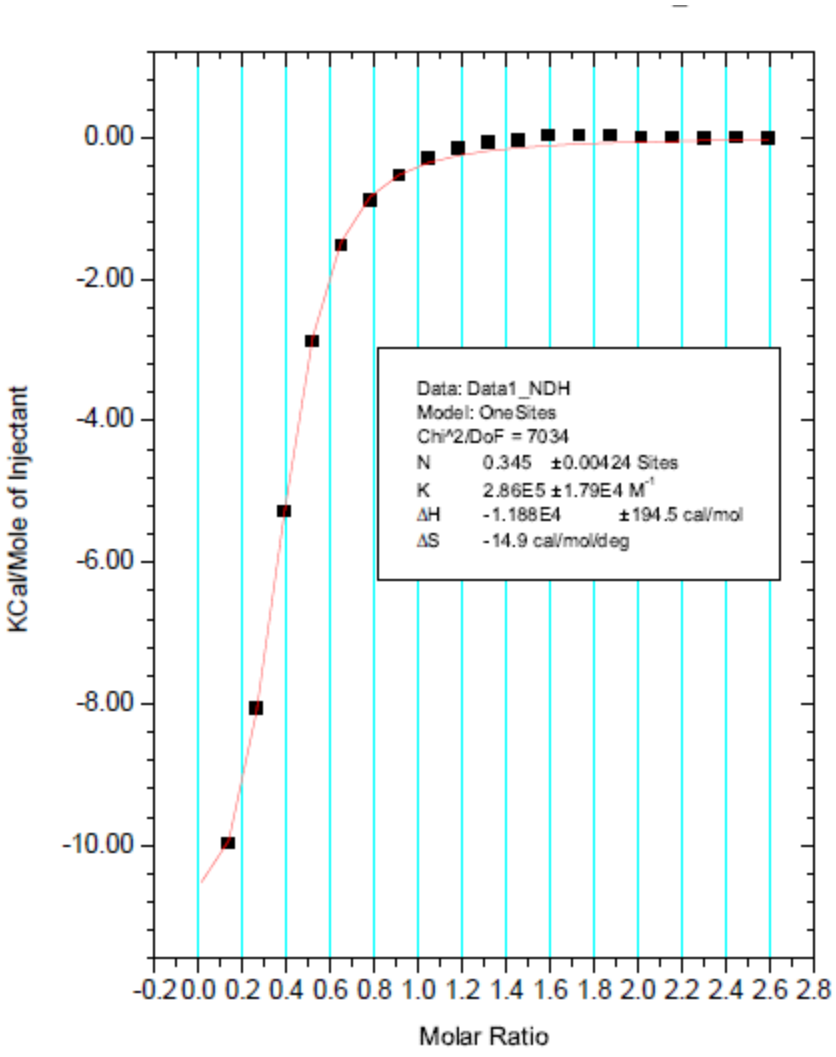


Figure 6.42. One trial of H3K9me2 (Ac-WGGGQTARKme2STG-NH_2) (1.04 mM) titrated into **A_2B-Asp-2** (81 μM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

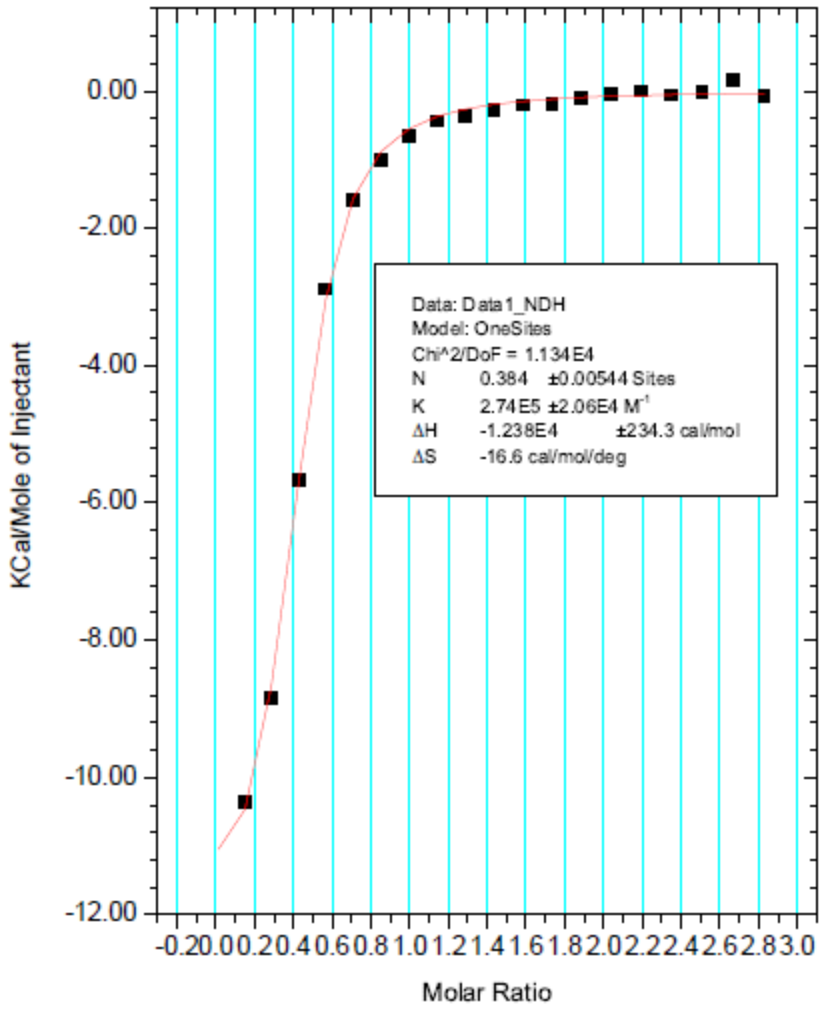


Figure 6.43. One trial of H3K9me3 (Ac-WGGGQTARKme3STG-NH_2) (1.13 mM) titrated into **A_2B-Asp-2** (81 μM) at 25 °C in 10 mM sodium borate buffer pH 8.5.

REFERENCES

(1) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J. *Angew. Chemie Int. Ed.* **2005**, *44*, 7342–7372.

(2) Greer, E. L.; Shi, Y. *Nat. Rev. Genet.* **2012**, *13*, 343–357.

(3) Verdin, E.; Ott, M. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 258–264.

(4) Rossetto, D.; Avvakumov, N.; Côté, J. *Epigenetics* **2012**, *7*, 1098–1108.

(5) Cao, J.; Yan, Q. *Front. Oncol.* **2012**, *2*, 26.

(6) Dehennaut, V.; Leprince, D.; Lefebvre, T. *Front. Endocrinol. (Lausanne).* **2014**, *5*, 155.

(7) Alexander Deiters; T. Ashton Cropp; Mridul Mukherji; Jason W. Chin; J. Christopher Anderson, and; Schultz*, P. G. **2003**.

(8) Wang, Q.; Parrish, A. R.; Wang, L. *Chem. Biol.* **2009**, *16*, 323–336.

(9) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. *Science (80-.).* **1994**, *266*.

(10) Malins, L. R.; Payne, R. J. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70–78.

(11) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6705–6710.

(12) Berrade, L.; Camarero, J. A. *Cell. Mol. Life Sci.* **2009**, *66*, 3909–3922.

(13) Boutureira, O.; Bernardes, G. J. L. *Chem. Rev.* **2015**, *115*, 150220140113005.

(14) Siemion, I. Z. *Biosystems* **1994**, *32*, 163–170.

(15) Gong, Y.; Pan, L. *Tetrahedron Lett.* **2015**, *56*, 2123–2132.

(16) Jentoft, N.; Dearborn, D. G. *J. Biol. Chem.* **1979**, *254*, 4359–4365.

(17) Massa, S.; Xavier, C.; De Vos, J.; Caveliers, V.; Lahoutte, T.; Muyldermans, S.; Devoogdt, N. *Bioconjug. Chem.* **2014**, *25*, 979–988.

(18) Hemantha, H. P.; Bavikar, S. N.; Herman-Bachinsky, Y.; Haj-Yahya, N.; Bondalapati, S.; Ciechanover, A.; Brik, A. *J. Am. Chem. Soc.* **2014**, *136*, 2665–2673.

(19) Cal, P. M. S. D.; Bernardes, G. J. L.; Gois, P. M. P. *Angew. Chemie Int. Ed.* **2014**, *53*, 10585–10587.

(20) Rademann, J. *Angew. Chemie Int. Ed.* **2004**, *43*, 4554–4556.

(21) Tanaka, K.; Fukase, K.; Katsumura, S. *Synlett* **2011**, *2011*, 2115–2139.

(22) Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.

(23) Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P.; Huang, Y.; Liu, L.; Krall, N.; Cruz, F. P. da; Boutureira, O.; Bernardes, G. J. L.; Shannon, D. A.; Weerapana, E.; Boutureira, O.; Bernardes, G. J. L.; Koniev, O.; Wagner, A.; Takaoka, Y.; Ojida, A.; Hamachi, I.; Stephanopoulos, N.; Francis, M. B.; Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G.; Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; González, M. F.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofielda, C. J.; Davis, B. G.; Ratner, V.; Kahana, E.; Eichler, M.; Haas, E.; Moody, P.; Chudasama, V.; Nathani, R. I.; Maruani, A.; Martin, S.; Smith, M. E. B.; Caddick, S.; Cal, P. M. S. D.; Bernardes, G. J. L.; Gois, P. M. P.; Rashidian, M.; Kumarapperuma, S. C.; Gabrielse, K.; Fegan, A.; Wagner, C. R.; Distefano, M. D.; Maruani, A.; Smith, M. E. B.; Miranda, E.; Chester, K. A.; Chudasama, V.; Caddick, S.; Wang, T.; Riegger, A.; Lamla, M.; Wiese, S.; Oeckl, P.; Otto, M.; Wu, Y.; Fischer, S.; Barth, H.; Kuan, S. L.; Weil, T.; Hu, Q. Y.; Berti, F.; Adamo, R.; Patterson, D. M.; Nazarova, L. A.; Prescher, J. A.; Ngo, J. T.; Tirrel, D. A.; Xie, J.; Schultz, P. G.; Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Voorhis, T. V.; Pentelute, B. L.; Sachdeva, A.; Wang, K.; Elliott, T.; Chin, J. W.; Zhang, D.; Devarie-Baez, N. O.; Jack, Q. L.; Jr, R. L.; Xian, M.; Mühlberg, M.; Hoesl, M. G.; Kuehne, C.; Dernedde, J.; Budisa, N.; Hackenberger, C. P. R.; Kim, J.; Seo, M.-H.; Lee, S.; Cho, K.; Yang, A.; Woo, K.; Kim, H.-S.; Park, H.-S.; Grünewald, J.; Jones, D. H.; Brock, A.; Chiu, H.-P.; Bursulaya, B.; Ng, K.; Vo, T.; Patterson, P.; Uno, T.; Hunt, J.; Spraggon, G.; Geierstanger, B. H.; Nathani, R. I.; Moody, P.; Chudasama, V.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S.; Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S.; Ren, H.; Xiao, F.; Zhan, K.; Kim, Y.-P.; Xie, H.; Xia, Z.; Rao, J.; Yuan, Y.; Liang, G.; Zhang, L.; Tam, J. P.; Casi, G.; Dezot, N. H.; Zuberbühler, K.; Scheuermann, J.; Neri, D.; Bernardes, G. J. L.; Steiner, M.; Hartmann, I.; Neri, D.; Casi, G.; Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P.; Cal, P. M. S. D.; Frade, R. F. M.; Chudasama, V.; Cordeiro, C.; Caddick, S.; Gois, P. M. P.; Cal, P. M. S. D.; Frade, R. F. M.; Cordeiro, C.; Gois, P. M. P.; Schmidt, P.; Stress, C.; Gillingham, D.; Bandyopadhyay, A.; Gao, J.; Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S.; Draganov, A. B.; Wang, K.; Holmes, J.; Damera, K.; Wang, D.; Dai, C.; Wang, B.; Bandyopadhyay, A.; Cambray, S.; Gao, J.; Koniev, O.; Leriche, G.; Nothisen, M.; Remy, J.-S.; Strub, J.-M.; Reiss, C. S.; Dorsselaer, A. V.; Baati, R.; Wagner, A. *Chem. Sci.* **2016**, *8*, 101.

(24) Chan, A. O.-Y.; Ho, C.-M.; Chong, H.-C.; Leung, Y.-C.; Huang, J.-S.; Wong, M.-K.; Che, C.-M. *J. Am. Chem. Soc.* **2012**, *134*, 2589–2598.

(25) Koniev, O.; Leriche, G.; Nothisen, M.; Remy, J.-S.; Strub, J.-M.; Schaeffer-Reiss, C.; Van Dorsselaer, A.; Baati, R.; Wagner, A. *Bioconjug. Chem.* **2014**, *25*, 202–206.

(26) Arumugam, S.; Guo, J.; Mbua, N. E.; Friscourt, F.; Lin, N.; Nekongo, E.; Boons, G.-J.; Popik, V. V. *Chem. Sci.* **2014**, *5*, 1591–1598.

(27) Valkevich, E. M.; Guenette, R. G.; Sanchez, N. A.; Chen, Y.; Ge, Y.; Strieter, E. R. *J. Am. Chem. Soc.* **2012**, *134*, 6916–6919.

(28) Li, F.; Allahverdi, A.; Yang, R.; Lua, G. B. J.; Zhang, X.; Cao, Y.; Korolev, N.; Nordenskiöld, L.; Liu, C.-F. *Angew. Chemie Int. Ed.* **2011**, *50*, 9611–9614.

(29) Conte, M. Lo; Staderini, S.; Marra, A.; Sanchez-Navarro, M.; Davis, B. G.; Dondoni, A.; Lo, A.; Bertozzi, C. R. *Chem. Commun.* **2011**, *47*, 11086.

(30) Gavrilyuk, J.; Ban, H.; Nagano, M.; Hakamata, W.; Barbas, C. F. *Bioconjug. Chem.* **2012**, *23*, 2321–2328.

(31) Tilley, S. D.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 1080–1081.

(32) McFarland, J. M.; Joshi, N. S.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 7639–7644.

(33) Seki, Y.; Ishiyama, T.; Sasaki, D.; Abe, J.; Sohma, Y.; Oisaki, K.; Kanai, M. *J. Am. Chem. Soc.* **2016**, *138*, 10798–10801.

(34) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem.* **2008**, *77*, 383–414.

(35) Kolb, H. C.; Sharpless, K. B. *Drug Discov. Today* **2003**, *8*, 1128–1137.

(36) Takaoka, Y.; Ojida, A.; Hamachi, I. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 4088–4106.

(37) Brunner, J. *Annu. Rev. Biochem.* **1993**, *62*, 483–514.

(38) Knowles, J. R. *Acc. Chem. Res.* **1972**, *5*, 155–160.

(39) Singh, A.; Thornton, E. R.; Westheimer, F. H. *J. Biol. Chem.* **1962**, *237*, 3006–3008.

(40) Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol. Biosyst.* **2008**, *4*, 473.

(41) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. *Proc. Natl. Acad. Sci.* **2004**, *101*, 10000–10005.

(42) Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W. *Biochemistry* **2007**, *46*, 350–358.

(43) Lee-Chiang, L.; Te-Ling, P.; Chi-Hsien, K.; Ying-Ling, C.; Hsin-Yi, W.; Jing-Jer, L. *J. Proteome Res.* **2002**, *1*, 35–40.

(44) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. *Proc. Natl. Acad. Sci.* **1999**, *96*, 14694–14699.

(45) Serim, S.; Haedke, U.; Verhelst, S. H. L. *ChemMedChem* **2012**, *7*, 1146–1159.

(46) Tsukiji, S.; Hamachi, I. *Curr. Opin. Chem. Biol.* **2014**, *21*, 136–143.

(47) Fujishima, S.; Yasui, R.; Miki, T.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2012**, *134*, 3961–3964.

(48) Takaoka, Y.; Nishikawa, Y.; Hashimoto, Y.; Sasaki, K.; Hamachi, I. *Chem. Sci.* **2015**, *6*,

3217–3224.

(49) Hughes, C. C.; Yang, Y.-L.; Liu, W.-T.; Dorrestein, P. C.; La Clair, J. J.; Fenical, W. *J. Am. Chem. Soc.* **2009**, *131*, 12094–12096.

(50) Koshi, Y.; Nakata, E.; Miyagawa, M.; Tsukiji, S.; Ogawa, T.; Hamachi, I. *J. Am. Chem. Soc.* **2008**, *130*, 245–251.

(51) Sato, S.; Nakamura, H. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 8681–8684.

(52) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. *Nat. Chem. Biol.* **2009**, *5*, 341–343.

(53) Tamura, T.; Kioi, Y.; Miki, T.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2013**, *135*, 6782–6785.

(54) Tsukiji, S.; Wang, H.; Miyagawa, M.; Tamura, T.; Takaoka, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2009**, *131*, 9046–9054.

(55) Tamura, T.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2012**, *134*, 2216–2226.

(56) Matsuo, K.; Kioi, Y.; Yasui, R.; Takaoka, Y.; Miki, T.; Fujishima, S.; Hamachi, I.; Tochio, H.; Shirakawa, M.; Hamachi, I.; Pingoud, A.; Tsien, R. Y. *Chem. Sci.* **2013**, *4*, 2573.

(57) Otsuki, S.; Nishimura, S.; Takabatake, H.; Nakajima, K.; Takasu, Y.; Yagura, T.; Sakai, Y.; Hattori, A.; Kakeya, H. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1608–1611.

(58) Yamaguchi, T.; Asanuma, M.; Nakanishi, S.; Saito, Y.; Okazaki, M.; Dodo, K.; Sodeoka, M. *Chem. Sci.* **2014**, *5*, 1021.

(59) Gregory, S. G.; Barlow, K. F.; McLay, K. E.; Kaul, R.; Swarbreck, D.; Dunham, A.; Scott, C. E.; Howe, K. L.; Woodfine, K.; Spencer, C. C. A.; Jones, M. C.; Gillson, C.; Searle, S.; Zhou, Y.; Kokocinski, F.; McDonald, L.; Evans, R.; Phillips, K.; Atkinson, A.; Cooper, R.; Jones, C.; Hall, R. E.; Andrews, T. D.; Lloyd, C.; Ainscough, R.; Almeida, J. P.; Ambrose, K. D.; Anderson, F.; Andrew, R. W.; Ashwell, R. I. S.; Aubin, K.; Babbage, A. K.; Bagguley, C. L.; Bailey, J.; Beasley, H.; Bethel, G.; Bird, C. P.; Bray-Allen, S.; Brown, J. Y.; Brown, A. J.; Buckley, D.; Burton, J.; Bye, J.; Carder, C.; Chapman, J. C.; Clark, S. Y.; Clarke, G.; Clee, C.; Cobley, V.; Collier, R. E.; Corby, N.; Coville, G. J.; Davies, J.; Deadman, R.; Dunn, M.; Earthrowl, M.; Ellington, A. G.; Errington, H.; Frankish, A.; Frankland, J.; French, L.; Garner, P.; Garnett, J.; Gay, L.; Ghori, M. R. J.; Gibson, R.; Gilby, L. M.; Gillett, W.; Glithero, R. J.; Grafham, D. V.; Griffiths, C.; Griffiths-Jones, S.; Grocock, R.; Hammond, S.; Harrison, E. S. I.; Hart, E.; Haugen, E.; Heath, P. D.; Holmes, S.; Holt, K.; Howden, P. J.; Hunt, A. R.; Hunt, S. E.; Hunter, G.; Isherwood, J.; James, R.; Johnson, C.; Johnson, D.; Joy, A.; Kay, M.; Kershaw, J. K.; Kibukawa, M.; Kimberley, A. M.; King, A.; Knights, A. J.; Lad, H.; Laird, G.; Lawlor, S.; Leongamornlert, D. A.; Lloyd, D. M.; Loveland, J.; Lovell, J.; Lush, M. J.; Lyne, R.; Martin, S.; Mashreghi-Mohammadi, M.; Matthews, L.; Matthews, N. S. W.; McLaren, S.;

Milne, S.; Mistry, S.; M oore, M. J. F.; Nickerson, T.; O'Dell, C. N.; Oliver, K.; Palmeiri, A.; Palmer, S. A.; Parker, A.; Patel, D.; Pearce, A. V.; Peck, A. I.; Pelan, S.; Phelps, K.; Phillimore, B. J.; Plumb, R.; Rajan, J.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Sehra, H. K.; Sheridan, E.; Shownkeen, R.; Sims, S.; Skuce, C. D.; Smith, M.; Steward, C.; Subramanian, S.; Sycamore, N.; Tracey, A.; Tromans, A.; Van Helmond, Z.; Wall, M.; Wallis, J. M.; White, S.; Whitehead, S. L.; Wilkinson, J. E.; Willey, D. L.; Williams, H.; Wilming, L.; Wray, P. W.; Wu, Z.; Coulson, A.; Vaudin, M.; Sulston, J. E.; Durbin, R.; Hubbard, T.; Wooster, R.; Dunham, I.; Carter, N. P.; McVean, G.; Ross, M. T.; Harrow, J.; Olson, M. V.; Beck, S.; Rogers, J.; Bentley, D. R. *Nature* **2006**, *441*, 315–321.

(60) Mandelkern, M.; Elias, J. G.; Eden, D.; Crothers, D. M. *J. Mol. Biol.* **1981**, *152*, 153–161.

(61) Annunziato, A. T. *Nat. Educ.* **2008**, *1*, 26.

(62) Van Holde, K. E. *Chromatin: Springer Series in Molecular Biology*; Springer-Verlag: New York, 1988.

(63) Wolffe, A. P. *Chromatin: Structure and Function*; 3rd ed.; Academic: San Diego, 1999.

(64) Thomas, J. O.; Kornberg, R. D. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 2626–2630.

(65) Bednar, J.; Horowitz, R. A.; Grigoryev, S. A.; Carruthers, L. M.; Hansen, J. C.; Koster, A. J.; Woodcock, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 14173–14178.

(66) Pierce, B. *Genetics: A Conceptual Approach*; 2nd ed.; W H Freeman & Co: New York, 2005.

(67) Kubicek, S. Epigenetics - A Primer http://www.the-scientist.com/?articles.view/articleNo/30737/title/Epigenetics-A-Primer/ (accessed Jun 30, 2017).

(68) Filippakopoulos, P.; Knapp, S. *Nat. Rev. Drug Discov.* **2014**, *13*, 337–356.

(69) Flanagan, J. F.; Mi, L.-Z.; Chruszcz, M.; Cymborowski, M.; Clines, K. L.; Kim, Y.; Minor, W.; Rastinejad, F.; Khorasanizadeh, S. *Nature* **2005**, *438*, 1181–1185.

(70) Li, H.; Ilin, S.; Wang, W.; Duncan, E. M.; Wysocka, J.; Allis, C. D.; Patel, D. J. *Nature* **2006**, *442*, 91–95.

(71) Lehnertz, B.; Ueda, Y.; Derijck, A. A. H. A.; Braunschweig, U.; Perez-Burgos, L.; Kubicek, S.; Chen, T.; Li, E.; Jenuwein, T.; Peters, A. H. F. M. *Curr. Biol.* **2003**, *13*, 1192–1200.

(72) Di Lorenzo, A.; Bedford, M. T. *FEBS Lett.* **2011**, *585*, 2024–2031.

(73) Gayatri, S.; Bedford, M. T. *Biochim. Biophys. Acta* **2014**, *1839*, 702–710.

(74) Kschonsak, M.; Haering, C. H. *Bioessays* **2015**, *37*, 755–766.

(75) Pinto, D. M. S.; Flaus, A. In *Sub-cellular biochemistry*; 2010; Vol. 50, pp. 55–78.

(76) Fraga, M. F.; Ballestar, E.; Villar-Garea, A.; Boix-Chornet, M.; Espada, J.; Schotta, G.; Bonaldi, T.; Haydon, C.; Ropero, S.; Petrie, K.; Iyer, N. G.; P?rez-Rosado, A.; Calvo, E.; Lopez, J. A.; Cano, A.; Calasanz, M. J.; Colomer, D.; Piris, M. ?ngel; Ahn, N.; Imhof, A.; Caldas, C.; Jenuwein, T.; Esteller, M. *Nat. Genet.* **2005**, *37*, 391–400.

(77) Cheung, N.; Chan, L. C.; Thompson, A.; Cleary, M. L.; So, C. W. E. *Nat. Cell Biol.* **2007**, *9*, 1208–1215.

(78) Urdinguio, R. G.; Sanchez-Mut, J. V; Esteller, M. *Lancet Neurol.* **2009**, *8*, 1056–1072.

(79) Perez-Burgos, L.; Peters, A. H. F. .; Opravil, S.; Kauer, M.; Mechtler, K.; Jenuwein, T. 2003; pp. 234–254.

(80) Zhao, Y.; Jensen, O. N. *Proteomics* **2009**, *9*, 4632–4641.

(81) Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y. *Anal. Chem.* **2005**, *77*, 5144–5154.

(82) Tian, Y.; Zhou, Y.; Elliott, S.; Aebersold, R.; Zhang, H. *Nat. Protoc.* **2007**, *2*, 334–339.

(83) Cytoskeleton. Post-translational Modification Detection Techniques. www.cytoskeleton.com/about-signal-seeker-ptm-detection (accessed Jun 30, 2017).

(84) Rothbart, S. B.; Dickson, B. M.; Raab, J. R.; Grzybowski, A. T.; Krajewski, K.; Guo, A. H.; Shanle, E. K.; Josefowicz, S. Z.; Fuchs, S. M.; Allis, C. D.; Magnuson, T. R.; Ruthenburg, A. J.; Strahl, B. D. *Mol. Cell* **2015**, *59*, 502–511.

(85) Nishikori, S.; Hattori, T.; Fuchs, S. M.; Yasui, N.; Wojcik, J.; Koide, A.; Strahl, B. D.; Koide, S. *J. Mol. Biol.* **2012**, *424*, 391–399.

(86) Blow, N. *Nature* **2007**, *447*, 741–744.

(87) Sidoli, S.; Cheng, L.; Jensen, O. N. *J. Proteomics* **2012**, *75*, 3419–3433.

(88) Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R. *J. Am. Chem. Soc.* **2012**, *134*, 17015–17018.

(89) Binda, O.; Boyce, M.; Rush, J. S.; Palaniappan, K. K.; Bertozzi, C. R.; Gozani, O. *Chembiochem* **2011**, *12*, 330–334.

(90) Bothwell, I. R.; Islam, K.; Chen, Y.; Zheng, W.; Blum, G.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2012**, *134*, 14905–14912.

(91) Islam, K.; Bothwell, I.; Chen, Y.; Sengelaub, C.; Wang, R.; Deng, H.; Luo, M. *J. Am. Chem. Soc.* **2012**, *134*, 5909–5915.

(92) Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. *Chembiochem* **2010**, *11*,

63–66.

(93) Daze, K. D.; Ma, M. C. F.; Pineux, F.; Hof, F. *Org. Lett.* **2012**, *14*, 1512–1515.

(94) Peacock, H.; Thinnes, C. C.; Kawamura, A.; Hamilton, A. D. *Supramol. Chem.* **2016**, *28*, 575–581.

(95) Gamal-Eldin, M. A.; Macartney, D. H.; Zhou, M.-M.; Kim, S.-Y.; Lee, J. W.; Hakamoto, S.; Yamaguchi, K.; Kim, K.; Kim, J.; Selvapalam, N.; Ryu, S. H.; Kim, K.; Gilson, M. K.; Kim, K.; Inoue, Y.; Mapelli, M.; Guccione, E. *Org. Biomol. Chem.* **2013**, *11*, 488–495.

(96) Ingerman, L.; Cuellar, M.; Waters, M. L. *Chem. Commun.* **2010**, *46*, 1839.

(97) James, L.; Beaver, J. E.; Rice, N.; Waters, M. L. *J. Am. Chem. Soc.* **2013**, *135*, 6450–6455.

(98) Pinkin, N. K.; Waters, M. L. *Org. Biomol. Chem.* **2014**.

(99) Beaver, J. E.; Peacor, B. C.; Bain, J. V.; James, L. I.; Waters, M. L. *Org. Biomol. Chem.* **2015**, *13*, 3220–3226.

(100) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590–593.

(101) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chemistry* **2008**, *14*, 2153–2166.

(102) Minaker, S. A.; Daze, K. D.; Ma, M. C. F.; Hof, F. *J. Am. Chem. Soc.* **2012**, *134*, 11674–11680.

(103) Peacor, B. C.; Ramsay, C. M.; Waters, M. L. *Chem. Sci.* **2017**, *8*, 1422–1428.

(104) Liu, Y.; Perez, L.; Mettry, M.; Gill, A. D.; Byers, S. R.; Easley, C. J.; Bardeen, C. J.; Zhong, W.; Hooley, R. J.; Cheng, Q.; Hooley, R. J.; Moghaddam, S.; Gilson, M. K.; Kim, K. M.; Inoue, Y. *Chem. Sci.* **2017**, *8*, 3960–3970.

(105) Garnett, G. A. E.; Starke, M. J.; Shaurya, A.; Li, J.; Hof, F. *Anal. Chem.* **2016**, *88*, 3697–3703.

(106) Kimura, Y.; Saito, N.; Hanada, K.; Liu, J.; Okabe, T.; Kawashima, S. A.; Yamatsugu, K.; Kanai, M. *ChemBioChem* **2015**, *16*, 2599–2604.

(107) Carlson, S. M.; Gozani, O. *J. Mol. Biol.* **2014**, *426*, 3350–3362.

(108) Jones, P. A.; Baylin, S. B. *Cell* **2007**, *128*, 683–692.

(109) Kazanecki, C. C.; Kowalski, A. J.; Ding, T.; Rittling, S. R.; Denhardt, D. T. *J. Cell. Biochem.* **2007**, *102*, 925–935.

(110) Carlson, S. M.; Moore, K. E.; Green, E. M.; Martín, G. M.; Gozani, O. *Nat. Protoc.* **2014**, *9*, 37–50.

(111) Bremang, M.; Cuomo, A.; Agresta, A. M.; Stugiewicz, M.; Spadotto, V.; Bonaldi, T. *Mol. Biosyst.* **2013**, *9*, 2231–2247.

(112) Guo, A.; Gu, H.; Zhou, J.; Mulhern, D.; Wang, Y.; Lee, K. A.; Yang, V.; Aguiar, M.; Kornhauser, J.; Jia, X.; Ren, J.; Beausoleil, S. A.; Silva, J. C.; Vemulapalli, V.; Bedford, M. T.; Comb, M. J. *Mol. Cell. Proteomics* **2014**, *13*, 372–387.

(113) Geoghegan, V.; Guo, A.; Trudgian, D.; Thomas, B.; Acuto, O. *Nat. Commun.* **2015**, *6*, 6758.

(114) Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V; Mann, M. *Science* **2009**, *325*, 834–840.

(115) Choudhary, C.; Weinert, B. T.; Nishida, Y.; Verdin, E.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 536–550.

(116) Gronborg, M. *Mol. Cell. Proteomics* **2002**, *1*, 517–527.

(117) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G. *Nat. Methods* **2007**, *4*, 798–806.

(118) Ohe, Y.; Iwai, K. *J. Biochem.* **1981**, *90*, 1205–1211.

(119) Bonger, K. M.; van den Berg, R. J. B. H. N.; Heitman, L. H.; IJzerman, A. P.; Oosterom, J.; Timmers, C. M.; Overkleeft, H. S.; van der Marel, G. A. *Bioorg. Med. Chem.* **2007**, *15*, 4841–4856.

(120) Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2002**, *41*, 2596–2599.

(121) Florea, M.; Kudithipudi, S.; Rei, A.; González-Álvarez, M. J.; Jeltsch, A.; Nau, W. M. *Chem. - A Eur. J.* **2012**, *18*, 3521–3528.

(122) Huang, F.; Nie, Y.; Ye, F.; Zhang, M.; Xia, J. *Bioconjug. Chem.* **2015**.

(123) Daze, K. D.; Jones, C. E.; Lilgert, B. J.; Beshara, C. S.; Hof, F. *Can. J. Chem.* **2013**, *91*, 1072–1076.

(124) Cheung, W. L.; Briggs, S. D.; Allis, C. D. *Curr. Opin. Cell Biol.* **2000**, *12*, 326–333.

(125) Strahl, B. D.; Allis, C. D. *Nature* **2000**, *403*, 41–45.

(126) Lai, A. Y.; Wade, P. A. *Nat. Rev. Cancer* **2011**, *11*, 588–596.

(127) Denslow, S. A.; Wade, P. A. *Oncogene* **2007**, *26*, 5433–5438.

(128) Huang, L. *J. Cell. Physiol.* **2006**, *209*, 611–616.

(129) Lin, H.-Y.; Chen, C.-S.; Lin, S.-P.; Weng, J.-R.; Chen, C.-S. *Med. Res. Rev.* **2006**, *26*, 397–413.

(130) Hoffmann, K.; Jung, M.; Brosch, G.; Loidl, P. *Nucleic Acids Res.* **1999**, *27*, 2057–2058.

(131) Acker, M. G.; Auld, D. S. *Perspect. Sci.* **2014**, *1*, 56–73.

(132) Sodji, Q. H.; Kornacki, J. R.; Mrksich, M.; Oyelere, A. K. *Epigenetic Technological Applications*; Elsevier, 2015.

(133) Schwämmle, V.; Aspalter, C.-M.; Sidoli, S.; Jensen, O. N. *Mol. Cell. Proteomics* **2014**, *13*, 1855–1865.

(134) Marks, P. A.; Breslow, R. *Nat. Biotechnol.* **2007**, *25*, 84–90.

(135) Martin, C.; Zhang, Y. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838–849.

(136) Eberharter, A.; Becker, P. B. *EMBO Rep.* **2002**, *3*, 224–229.

(137) Chantalat, S.; Depaux, A.; Héry, P.; Barral, S.; Thuret, J.-Y.; Dimitrov, S.; Gérard, M. *Genome Res.* **2011**, *21*, 1426–1437.

(138) Carrozza, M. J.; Li, B.; Florens, L.; Suganuma, T.; Swanson, S. K.; Lee, K. K.; Shia, W.-J.; Anderson, S.; Yates, J.; Washburn, M. P.; Workman, J. L. *Cell* **2005**, *123*, 581–592.

(139) Young, N. L.; DiMaggio, P. A.; Plazas-Mayorca, M. D.; Baliban, R. C.; Floudas, C. A.; Garcia, B. A. *Mol. Cell. Proteomics* **2009**, *8*, 2266–2284.

(140) Madsen, A. S.; Olsen, C. A. *Angew. Chemie Int. Ed.* **2012**, *51*, 9083–9087.

(141) Zhang, J.; Chung, T.; Oldenburg, K. *J. Biomol. Screen.* **1999**, *4*, 67–73.

(142) Shaurya, A.; Dubicki, K. I.; Hof, F. *Supramol. Chem.* **2014**, *26*, 583–590.

(143) Gober, I. N.; Waters, M. L. *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459.

(144) Norouzy, A.; Azizi, Z.; Nau, W. M. *Angew. Chemie Int. Ed.* **2015**, *54*, 792–795.

(145) Liu, Y.; Perez, L.; Mettry, M.; Easley, C. J.; Hooley, R. J.; Zhong, W. *J. Am. Chem. Soc.* **2016**, *138*, 10746–10749.

(146) Kataev, E. A.; Müller, C. *Tetrahedron* **2014**, *70*, 137–167.

(147) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. *Angew. Chemie Int. Ed.* **2007**, *46*, 2366–2393.

(148) Schneider, H.-J.; Yatsimirsky, A. K. *Chem. Soc. Rev.* **2008**, *37*, 263–277.

(149) Butterfield, S. M.; Julius Rebek, J. *J. Am. Chem. Soc.* **2006**, *128*, 15366–15367.

(150) Schrader, T.; Bitan, G.; Klärner, F.-G.; Mori, Y.; Sajiki, H.; Tremaine, L. M.; Gales, L.; Saraiva, M. J.; Almeida, M. R.; Gallo, G.; Buxbaum, J.; Shimada, K.; Takahashi, K.;

Yamamura, K.; Maeda, S.; Pham, E.; Masliah, E.; Gage, F. H.; Riek, R.; Wettig, B.; Heid, C.; Kirchhoff, F.; Weil, T.; Klärner, F. G.; Schrader, T.; Bitan, G.; Sanchez-Garcia, E.; Winter, R.; Shorter, J.; Münch, J. *Chem. Commun.* **2016**, *52*, 11318–11334.

(151) Huang, G.-B.; Wang, S.-H.; Ke, H.; Yang, L.-P.; Jiang, W. *J. Am. Chem. Soc.* **2016**, *138*, 14550–14553.

(152) Berg, J. M.; Tymoczko, J. L.; Stryer, L. In *Biochemistry. 5th edition*; W H Freeman: New York, 2002.

(153) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1984**, *53*, 293–321.

(154) Tainer, J. A.; Cunningham, R. P. *Curr. Opin. Biotechnol.* **1993**, *4*, 474–483.

(155) Yun, M.; Wu, J.; Workman, J. L.; Li, B. *Cell Res.* **2011**, *21*, 564–578.

(156) Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.

(157) Adams-Cioaba, M. A.; Min, J. *Biochem. Cell Biol.* **2009**, *87*, 93–105.

(158) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.

(159) Li, H.; Fischle, W.; Wang, W.; Duncan, E. M.; Liang, L.; Murakami-Ishibe, S.; Allis, C. D.; Patel, D. J. *Mol. Cell* **2007**, *28*, 677–691.

(160) Percec, V.; Holerca, M. N.; Nummelin, S.; Morrison, J. J.; Glodde, M.; Smidrkal, J.; Peterca, M.; Rosen, B. M.; Uchida, S.; Balagurusamy, V. S. K.; Sienkowska, M. J.; Heiney, P. A. *Chem. - A Eur. J.* **2006**, *12*, 6216–6241.

(161) Umemoto, T.; Garrick, L. M.; Saito, N. *Beilstein J. Org. Chem.* **2012**, *8*, 461–471.

(162) Dahl, B. J.; Branchaud, B. P. *Org. Lett.* **2006**, *8*, 5841–5844.

(163) Meng, L.-B.; Zhang, W.; Li, D.; Li, Y.; Hu, X.-Y.; Wang, L.; Li, G.; Zhang, D.; Wang, L. *Chem. Commun.* **2015**, *51*, 14381–14384.

(164) Pinkin, N. K.; Liu, I.; Abron, J. D.; Waters, M. L. *Chem. - A Eur. J.* **2015**, *21*, 17981–17986.

(165) Parker, C. E.; Mocanu, V.; Mocanu, M.; Dicheva, N.; Warren, M. R. *Mass Spectrometry for Post-Translational Modifications*; CRC Press/Taylor & Francis, 2010.

(166) Pinkin, N. K.; N. Power, A.; Waters, M. L. *Org. Biomol. Chem.* **2015**, *13*, 10939–10945.