



Virginia Commonwealth University  
VCU Scholars Compass

---

Theses and Dissertations

Graduate School


---

2017

## Chemical Probes for Protein $\alpha$ -N-Terminal Methylation

Brianna D. Mackie  
*Virginia Commonwealth University*

Follow this and additional works at: <https://scholarscompass.vcu.edu/etd>

 Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Enzymes and Coenzymes Commons](#), and the [Medicinal and Pharmaceutical Chemistry Commons](#)

© The Author

---

Downloaded from

<https://scholarscompass.vcu.edu/etd/4880>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).

CHEMICAL PROBES FOR PROTEIN  $\alpha$ -N-TERMINAL METHYLATION

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor  
of Philosophy at Virginia Commonwealth University

by

BRIANNA D. MACKIE, BS

Advisor:  
RONG HUANG, PHD  
ASSISTANT PROFESSOR, DEPARTMENT OF MEDICINAL CHEMISTRY

Virginia Commonwealth University  
Richmond, Virginia  
April, 2017

### Acknowledgements

I would like to thank Dr. Huang for training me and taking the time and effort to teach me and advocate for me. She pushed me to the limits of my capabilities and I am grateful for that. I would like to thank my former post-doc Gang Zhang, who I worked directly under for the first year of my training. He was uncompromisingly kind and taught me the basic principles of medicinal chemistry. I would like to thank my lab-mate Yunfei Mao for always encouraging me and making coming to work every day something I looked forward to. I would also like to thank all the other current and past members of the Rong Huang lab; Stacie Richardson, Guangping Dong, Kathryn Marshman, Dongxing Chen and Linjie Li. I also want to thank all my committee members, Drs. Martin Safo, Shijun Zhang, Ashton Cropp and Adam Hawkridge for guiding me toward the completion of this doctorate. Finally, I would like to thank my family and friends, specifically my husband Beau. Thanks to the support and encouragement from them, it has made this process possible.

## Table of Contents

List of Tables .....	vi
List of Figures .....	vii
List of Schemes .....	ix
List of Abbreviations .....	x
Abstract .....	xii
1. Introduction .....	1
1.1 $\alpha$ -N-terminal Acetylation .....	4
1.2 $\alpha$ -N-terminal Propionylation .....	11
1.3 $\alpha$ -N-terminal Myristoylation .....	12
1.4 $\alpha$ -N-terminal Palmitoylation .....	16
1.5 $\alpha$ -N-terminal Ubiquitination .....	20
1.6 $\alpha$ -N-terminal Formylation .....	24
1.7 $\alpha$ -N-terminal Methylation .....	26
2. Results and Discussion .....	35
2.1 NTMT1 Petidomimetic Inhibitors .....	35
2.1.1 Design .....	35
2.1.2 Docking Studies .....	37
2.1.3 Synthesis .....	38

2.1.4 MALDI-MS Studies .....	39
2.1.5 Inhibition Studies .....	41
2.1.6 Selectivity Studies .....	46
2.1.7 Inhibition Mechanism .....	48
2.2 Photoaffinity Probes .....	49
2.2.1 Design .....	49
2.2.2 Docking Studies.....	51
2.2.3 Synthesis.....	53
2.2.4 Recognition Studies.....	55
2.2.5 Photoaffinity Labeling .....	57
3. Conclusions.....	65
4. Future Direction.....	68
5. Experimental and Methods.....	70
5.1 Materials and Reagents .....	70
5.2 Instruments .....	70
5.3 Chemistry .....	70
5.3.1 Photoaffinity Probes .....	70
5.3.2 Peptide Inhibitors.....	71
5.4 Purification .....	72
5.4.1 Enzymes.....	72
5.4.2 Chemical Probes and Inhibitors.....	73
5.5 MALDI-MS Methylation Studies.....	73

5.5.1 Methylation of Probe 1 and 3.....	73
5.5.2 Methylation of Peptide Inhibitors.....	74
5.5.3 Methylation Progression after Incubation of Inhibitors.....	74
5.6 Docking Studies .....	75
5.7 SDS-PAGE Gel Studies .....	75
5.7.1 Concentration-dependence Studies .....	75
5.7.2 Time-dependence Studies.....	76
5.7.3 Competition Studies with Recombinant Protein.....	77
5.7.4 Selectivity Studies with Recombinant Protein.....	77
5.7.5 Cell Lysate Experiments .....	78
5.8 Kinetic Analysis of Compounds.....	78
5.8.1 Chemical Probes .....	78
5.8.2 Peptide Inhibitors.....	79
5.8.3 Inhibition Mechanism.....	79
5.9 Selectivity Studies .....	81
5.9.1 G9a.....	81
5.9.2 PRMT1 .....	82
References.....	83
Appendix .....	100
Vita .....	135

## List of Tables

1. Recognition motif of $\alpha$ -N-terminal writer proteins .....	3
2. Summary of the structural subunits, recognition motif and PDB ID of each NAT enzyme.....	5
3. Kinetic parameters of recombinant NMT from bovine cardiac muscle.....	14
4. Summary of modifications at 1 <sup>st</sup> position for tetramer inhibitors .....	36
5. Peptide inhibitors with 0 CH <sub>2</sub> groups between carbonyl and X position .....	42
6. Peptide inhibitors with 1-2 CH <sub>2</sub> groups between carbonyl and X position .....	43
7. Structure-Activity-Relationship of modifications at 2 <sup>nd</sup> position.....	44
8. Structure-Activity-Relationship of modifications at 3 <sup>rd</sup> position .....	44
9. Peptide inhibitors with modifications at 4 <sup>th</sup> position .....	45
10. Tetramer and hexamer peptide analogs.....	46
11 Summary of selectivity data for top peptide inhibitors .....	47
A1. Kinetic data for peptide inhibitors.....	120
A2. Michaelis-Menten constants for probes <b>1-3</b> and NTMT1 substrates.....	132
A3. Predicted and detected m/z of peptide inhibitors and probes.....	132

### List of Figures

1. Modifications made by writer proteins to the $\alpha$ -N-terminus.....	2
2. NAT inhibitors and their respective IC <sub>50</sub> s.....	10
3. NMT ordered bi-bi catalytic.....	13
4. NMT inhibitors and their respective IC <sub>50</sub> s .....	16
5. Hhat inhibitors and their respective .....	19
6. Ubiquitin hierarchal cascade.....	22
7. Formylation and deformylation cycle for protein synththesis .....	25
8. PDF inhibitors and their respective IC <sub>50</sub> s.....	26
9. NTMT2 homology model overlayed NTMT1 crystal structure.....	30
10. Bisubstrate strategy to inhibit methylation .....	32
11. NTMT1 inhibitors and their respective IC <sub>50</sub> s .....	33
12. Docking studies of peptide inhibitors .....	38
13. MALDI-MS methylation inhibition assay for BM-30.....	40
14. Inhibition mechanism of BM-30 .....	48
15. Structure of photoaffinity probes <b>1-6</b> .....	50
16. Docking studies of probes <b>1-3</b> .....	52



17. MALDI-MS of probe <b>1</b> before and after addition of NTMT1 and SAM.....	55
18. Steady-state kinetics of probes <b>1-3</b> and substrates .....	56
19. Concentration-dependence photolabeling of NTMT1 by probes.....	57
20. Time-dependence photolabeling of NTMT1 by probes.....	58
21. Competition studies with probe <b>1</b> .....	59
22. Competition studies with probe <b>2</b> .....	60
23. Competition studies with probes <b>4-6</b> .....	61
24. Selectivity studies with probe <b>1</b> and <b>3</b> .....	62
25. Photoaffinity labeling of probes <b>1-3</b> in HeLa cell lysates.....	64
A1. Mass spectrometry of peptide inhibitors .....	116
A2. MALDI-MS methylation inhibition assay for BM-47 .....	117
A3. MALDI-MS methylation inhibition assay for BM-11 .....	117
A4. MALDI-MS methylation inhibition assay for BM-46 .....	118
A5. MALDI-MS methylation inhibition assay for BM-34 .....	118
A6. IC <sub>50</sub> curves of select peptide inhibitors .....	128
A7. Mass spectrometry of probes <b>1-3</b> .....	129
A8. Mass spectrometry of probes <b>4-6</b> .....	130
A9. MALDI-MS of probe <b>3</b> before and after addition of NTMT1 and SAM .....	131

List of Schemes

1. Synthesis of photoaffinity probe <b>1</b> .....	53
A1. Synthesis of BM-45 .....	97
A2. Synthesis of BM-26 .....	97

### List of Abbreviations

CENP – Centromere protein

DDB2 – DNA Damage-binding protein

Eef1a – elongation factor 1A

Efm7 - Elongation factor methyltransferase 7

FMT – Formyltransferase

G9a – Euchromatic histone-Lys N-methyltransferase 2

HAT – Histone acetyltransferase

HHAT – Hedgehog acyltransferase

KAT – Lysine acetyltransferase

MALDI – Matrix assisted laser desorption ionization

MBOAT – Membrane bound O-acyltransferase

NAT – N-terminal acetyltransferase

NMT – N-terminal myristoyltransferase

NPT – N-terminal propionyltransferase

NTMT – N-terminal methyltransferase

PDF – Peptide deformylase

PRMT1 – Protein arginine methyltransferase

RB1 – Retinoblastoma 1

RCC1 – Regulator of chromosome condensation 1

SAH – S-adenosyl homocysteine

SAHH – S-adenosyl homocysteine hydrolase

SAM – S-adenosylmethionine

Shh – Sonic hedgehog

TFA – Trifluoroacetic acid

## Abstract

### CHEMICAL PROBES FOR PROTEIN $\alpha$ -N-TERMINAL METHYLATION

By Brianna D. Mackie, BS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2017

Advisor:  
RONG HUANG, PHD  
ASSISTANT PROFESSOR, DEPARTMENT OF MEDICINAL CHEMISTRY

While protein  $\alpha$ -N-terminal methylation has been known for nearly four decades since it was first uncovered on bacteria ribosomal proteins L33, the function of this modification is still not entirely understood. Recent discoveries have demonstrated  $\alpha$ -N-terminal methylation is essential to stabilize the interactions between regulator of chromosome condensation 1 (RCC1) and chromatin during mitosis, to localize and enhance the interaction of centromere proteins (CENPs) with chromatin, and to facilitate the recruitment of DNA damage-binding protein 2 (DDB2) to DNA damage foci. Identification of N-terminal methyltransferase 1 (NTMT1) unveiled the eukaryotic methylation writer for protein  $\alpha$ -N-termini. In addition, NTMT2 that shares over 50% sequence similarity, has been identified as another mammalian protein  $\alpha$ -N-terminal methylation writer. Knockdown of NTMT1 results in mitotic defects and sensitizes

chemotherapeutic agents in breast cancer cell lines, while NTMT1 knockout mice showed premature aging. Additionally, NTMT1 has been shown to be overexpressed in a colorectal and melanoma tumor tissues, and in lung and liver cancer cell lines.

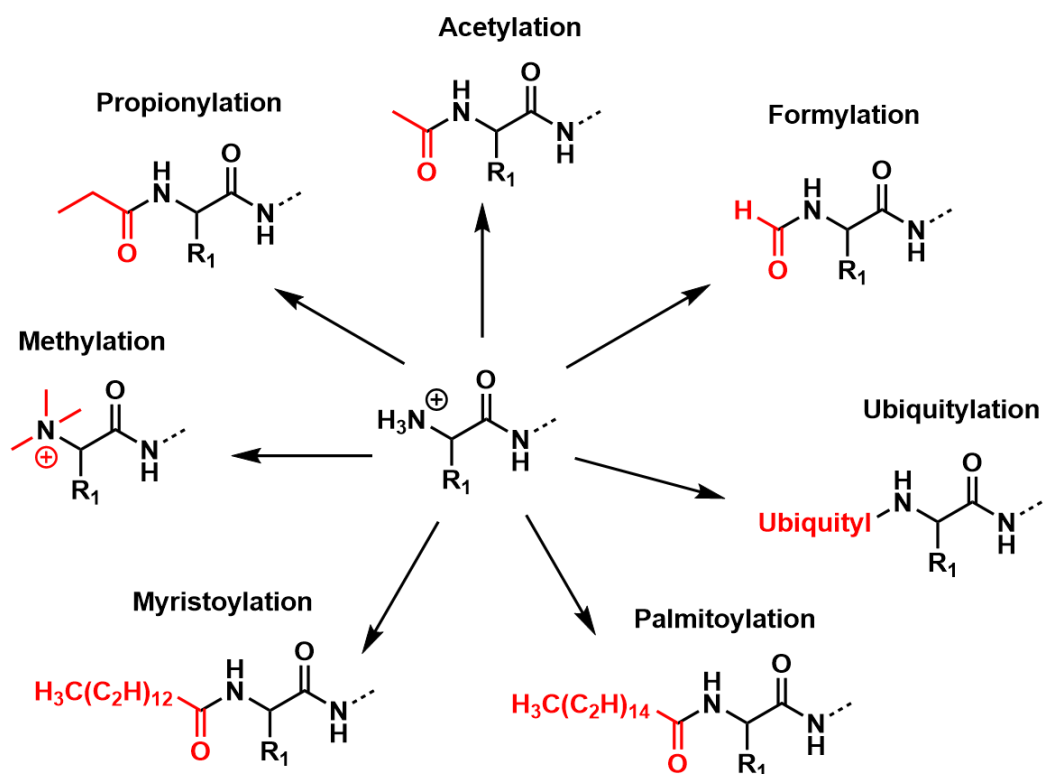
Given the vast array of clinical relevance, chemical probes and inhibitors for NTMT1 are vital to elucidate information about the function and downstream process of protein  $\alpha$ -N-terminal methylation. Therefore, 47 peptidomimetic compounds have been synthesized that target NTMT1. These peptide-based compounds range from three to six amino acids in length and the top 5 compounds have 3- to 300- fold selectivity for NTMT1 compared to other methyltransferases. An inhibition mechanism study has also been performed to verify the inhibitors are targeting the NTMT1 peptide binding site. Seven compounds have an  $IC_{50}$  of less than 5  $\mu$ M and our top inhibitor, BM-47, has an  $IC_{50}$  of  $0.32 \mu\text{M} \pm 0.06$  for NTMT1.

To further elucidate information about the NTMTs and their downstream effects, we utilized photoaffinity probes to target these enzymes. Our 6 photoaffinity probes exhibited in a dose- and time-dependent manner. Probe labeling has been shown to be driven by recognition and selectively and competitively label the NTMT writers in a complex cellular mixture. Our results also provided the first indication of substrate preferences among NTMT1/2. Methylated photoaffinity probes were also synthesized to identify novel proteins that recognize a methylated N-terminus and shed light on the function of  $\alpha$ -N-terminal methylation.

## 1. Introduction

Protein modifications may occur post- or co-translationally on the side chains, C-termini and backbone N-termini of proteins. These modifications are vital in downstream processes including gene expression, protein degradation, cell signaling and much more.<sup>1,2</sup> Modifications on the side chains have drawn extensive attention as sites for novel therapeutic targets. In addition, a variety of  $\alpha$ -N-terminal modifications including methylation, acetylation, propionylation, myristoylation, palmitoylation, ubiquitylation and formylation expose emerging interest. For most N-terminal modifications, the initial Met must first be cleaved by Met aminopeptidases and then subject to post-translational modifications by writers; however,  $\alpha$ -N-terminal acetylation can occur on the first Met and  $\alpha$ -N-terminal formylation is known to exclusively modify the first Met residue.<sup>2-5</sup>

Enzymes that catalyze the addition of a covalent modification onto its target are termed writers. Conversely, proteins that catalyze the removal of that modification are named erasers. Additionally, proteins that bind to these newly modified proteins are called readers. Until now, the only known eraser is for N-terminal formylation and there is sparse information about reader proteins that recognize  $\alpha$ -N-terminal modifications. Therefore, we will focus on the writers for  $\alpha$ -N-terminal modifications, which are summarized in Figure 1.



**Figure 1.** Modifications made by writers to the  $\alpha$ -N-terminus

The N-terminal acetyltransferase (NAT) family is the writer that catalyzes the addition of an acetyl group onto the  $\alpha$ -N-terminus.<sup>4</sup> The NAT family also acts as an N-terminal propionyltransferase (NPT) and catalyzes the addition of a propionyl group, although this modification is much less common.<sup>6</sup> Another newly discovered family is N-terminal methyltransferases (NTMTs), which adds a methyl group onto the  $\alpha$ -N-terminus.<sup>7</sup> Less common modifications include N-terminal palmitoylation catalyzed by Hedgehog acyltransferase (HHAT)<sup>8</sup> and N-terminal myristoylation carried out by N-terminal myristoyltransferases (NMTs).<sup>9</sup> N-terminal formylation occurs only on Met residues to indicate the initiation of protein synthesis and is catalyzed by methionyl-tRNA formyltransferase (FMT).<sup>5</sup> Lastly, N-terminal ubiquitylation has also been found to occur on Lys-deficient proteins by ubiquitin ligase.<sup>10</sup> For  $\alpha$ -N-terminal modifications, the initial amino acid sequence determines the types of modifications that can occur (Table 1). The



substrate preferences between the NAT family members varies greatly. NatA acetylates proteins beginning with smaller amino acids like Ala, Cys, Gly, Ser, Thr and Val<sup>11,12</sup> while the NatA catalytic subunit, Naa10, recognizes acidic amino acids like Asp and Glu.<sup>13</sup> NatB recognizes the initial Met residue followed by Asp, Glu, Asn, or Gln,<sup>11,14,15</sup> while NatC recognizes Met followed by a hydrophobic amino acid.<sup>11,15,16</sup> NatD has only been found to acetylate two substrates, H2A and H4 proteins, both starting with Ser-Gly-Arg-Gly.<sup>17,18</sup> NatE substrate preference is similar to NatC and will acetylate proteins beginning with Met followed by a hydrophobic residue.<sup>13</sup> NatF is only found in lower eukaryotes but has a substrate preference that overlaps with NatC and NatE.<sup>19</sup> NMTs add a myristic fatty acid chain only to proteins beginning with Gly<sup>20</sup> while HHAT adds a palmitic fatty acid chain to proteins starting with Cys.<sup>8</sup> N-terminal ubiquitination can occur on any protein with an unmodified N-terminus but often occurs on naturally Lys-deficient proteins.<sup>10</sup> As stated above, FMT enzyme only adds a formyl group to N-terminal Met residues.<sup>21</sup> Lastly, NTMT1 and 2 prefer an X-P-K/R (X=A, P, S, G) recognition motif<sup>22-24</sup> while a recently found NTMT3 enzyme has found to methylate one protein in yeast with initial sequence Gly-Lys.<sup>25</sup>

**Table 1.** Recognition motif of  $\alpha$ -N-terminal writer proteins<sup>3</sup>

Nt-acetylation							Nt-myristoylation	Nt-palmitoylation	Nt-ubiquitylation	Nt-formylation	Nt-methylation	
NatA	Naa10	NatB	NatC	NatD	NatE	NatF	NMT	HHAT	E2	FMT	NTMT1/2	NTMT3
A, C, G, S, T, V	D, E	ME, MD, MN, MQ	MI, ML, MF, MW	SGRG	MA, ML, MK, MM, MF, MS, MT, MY, MV	MA, MQ, MG, MI, ML, MK, MM, MS, MT, MY, MV	G	C	lysine-deficient proteins with unacetylated N-terminus	M	APK, PPK, SPK, GPK, GPR	GK

## 1.1 $\alpha$ -N-terminal Acetylation

$\alpha$ -N-terminal acetylation is a co- and posttranslational modification that occurs on 60% of yeast proteins and 80-90% of all human proteins.<sup>4,26</sup> This modification is catalyzed by NATs which are the enzymes that introduce an acetyl group on  $\alpha$ -N-termini of proteins. This modification neutralizes the positive charge on the free amino group.<sup>4</sup> NatA has been shown to undergo an ordered bi-bi kinetic mechanism where the peptide with the N-terminus to be modified binds to the NAT enzyme first, followed by the acetyl donor, Ac-CoA.<sup>27-29</sup> The NAT family shares a conserved Ac-CoA binding site that contains a conserved Glu24 residue, which acts as a general base to deprotonate the N-terminal amine of the protein substrate to facilitate the nucleophilic attack at the carbonyl position of the Ac-CoA cofactor.<sup>27</sup>

There are six eukaryotic NAT family enzymes NatA-NatF, which differs in substrate specificity and subunit composition. Most Nats contain one catalytic subunit and one or two auxiliary subunits that affect the substrate specificity. NatA, the most widely studied NAT enzyme, contains a catalytic subunit Naa10 and an auxiliary subunit Naa15. Naa10 also has the ability to acetylate protein substrates without the auxiliary subunit, but differs in substrate recognition motifs from the NatA complex.<sup>13</sup> NatB contains a catalytic subunit Naa20 and an auxiliary subunit Naa25. NatC is composed of one catalytic subunit Naa30 and two auxiliary subunits Naa35 and Naa38. NatE is composed of one catalytic domain Naa50 and one NatA complex; however, NatA and NatE have a different substrate specificity.<sup>30</sup> NatD and NatF only contain the catalytic subunits Naa40 and Naa60, respectively. The varying recognition motifs of the six NAT enzymes are shown in Table 2. Most NATs are associated with ribosomes, but NatF is associated with the cytosolic

side of Golgi and exhibits selectivity for membrane proteins. It is noted that the catalytic subunits of NatA and NatE have also been found to act as Lys acetyltransferases (KATs) and N-terminal propionyltransferases (NPTs) *in vitro*.<sup>6,31</sup>

**Table 2.** Summary of the structural subunits, recognition motif and crystal structures of NATs

NAT	NatA	NatB	NatC	NatD	NatE	NatF
Catalytic subunit	Naa10	Naa20	Naa30	Naa40	Naa50	Naa60
Auxiliary subunit	Naa15	Naa25	Naa35/ Naa38	--	Naa10/ Naa15	--
Recognition motif	A, C, G, S, T, V	ME, MD, MN, MQ	MI, ML, MF, MW	SGRG	MA, ML, MK, MM, MF, MS, MT, MY, MV	MA, MQ, MG, MI, ML, MK, MM, MS, MT, MY, MV
PDB ID	4KVX, 4KVO, 4KVM	--	--	4U9V, 4U9W	3TFY, 2OB0	5ICV, 5ICW

All the NAT family have a structurally conserved Ac-CoA binding site, which consists of a conserved fold containing four sequence motifs that is termed the N-acetyltransferase domain.<sup>32</sup> The first resolved crystal structure was the human Naa50 which validated the NAT preference for the  $\alpha$ -N-terminal amine over Lys side chains. The crystal structure identified that the Naa50 substrate peptide was bound to the protein through a series of backbone hydrogen bonds. Additionally, a hydrophobic pocket exists that forms van der Waals interactions with the initial Met residue, which prevents Naa50 from interacting with any other N-terminal amino acids.<sup>28</sup> Then, the NatA complex crystal

structure indicated the necessity of the auxiliary subunit in modulating catalytic activity and substrate preference. The auxiliary subunit of NatA, Naa15, contains 13 tetratricopeptide repeat motifs that wrap around and allosterically reconfigure the NatA catalytic subunit. Without the auxiliary subunit of NatA, key residues in the catalytic core including Leu22, Glu24, and Tyr26 are displaced and alters the Naa10 recognition motif to have preference for acidic amino acids like Asp and Glu. The crystal structure of Naa10 in complex with Naa15 shows a significant conformational change in the  $\alpha$ 1-loop- $\alpha$ 2 region of Naa10 compared to the crystal structure of Naa10 alone.<sup>13,27</sup> Superimposition of the crystal structures of NatA and NatE identified Glu35 and Val29 to be the key residues for their substrate specificity, respectively.<sup>33</sup> NatA can acetylate proteins that start with Ser, Ala, Thr, Gly and Val, but the peptide binding site cannot accommodate a Met residue. Structural information indicated that the Glu35 residue in the NatA peptide binding site would not accommodate a Met residue, due to steric hindrances.<sup>33</sup> Alternatively, the key residue in NatE is Val29, a smaller and more hydrophobic residue, which allows Met to bind in this site.<sup>28</sup> Mutation of NatA Glu35 to Ala or Val altered NatA substrate preference and enabled it to catalyze the acetylation of the N-terminal end of a NatE substrate, Met-Glu.<sup>33</sup> Recently, the NatD and NatF crystal structures have also been resolved.<sup>34,35</sup> NatD has the most specific recognition motif and is only known to acetylate histones H2A and H4, both beginning with a Ser-Gly-Ar-Gly-Lys sequence. The  $\alpha$ 1- $\alpha$ 2 and  $\beta$ 6- $\beta$ 7 loops of NatD orient the histone proteins N-terminus in a specific manner within the binding site. The peptide binding site of NatD is similar to NatA where it can only accommodate smaller residues; however, the NatD site is even more restricted than NatA. While the NatA binding site can accommodate Ser and Thr residues, NatD can

only accommodate Ser. Additionally, the 3<sup>rd</sup> position Arg of the histone substrates inserts into a pocket that is unique to the NatD enzyme. In addition, the N-terminus of NatD wraps around the catalytic core and stabilizes these interactions, indicating the N-terminus is also required for catalytic activity.<sup>34</sup> The crystal structure of NatF revealed that it is the most catalytically similar to NatE. Both NatE and NatF contain a hydrophobic pocket, which gives preference to Met at the 1<sup>st</sup> position. For the 2<sup>nd</sup> position of its substrate, NatE prefers a hydrophobic amino acid while NatF acetylates substrates with charged side chains at the 2<sup>nd</sup> position. This difference is due to the larger and more solvent exposed binding site of NatF. Another structural difference between NatE and NatF is NatF contains a longer  $\beta$ 6- $\beta$ 7 loop which mediates dimerization in the absence of a substrate peptide, which is unique has not been observed in any other member of the NAT family.<sup>35</sup>

$\alpha$ -N-terminal acetylation has demonstrated its significance in a variety of physiological processes like mediating protein complex formation, regulation of protein degradation, membrane attachment of small GTPases, and prevention of protein translocation from the cytosol to the endoplasmic reticulum.<sup>30,36,37</sup> For many years, it was hypothesized that proteins with acetylated N-termini are more stable and less susceptible to N-end rule pathway that governs the rate of protein degradation through recognition of the N-terminal residue of proteins. However, recent studies also suggest that N-terminal acetyl groups decrease the half-life for certain proteins through the Doa10 E3 ubiquitin ligase.<sup>4,38</sup> N-terminal acetylation also plays a role in subcellular localization, promotes proper association of Trm1p-II to the inner nuclear membrane, and prevents protein targeting to the endoplasmic reticulum secretory pathway.<sup>16,39</sup> N-terminal acetylation also

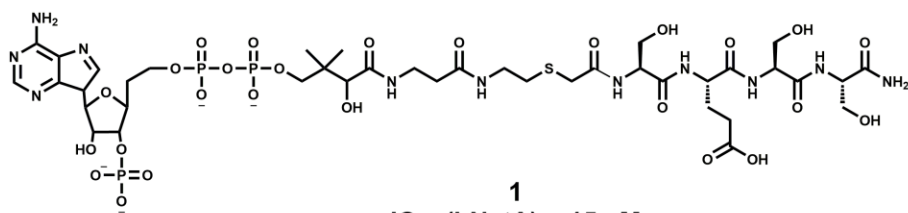
regulates protein-protein interactions; for example, the tropomyosin-actin complex formation is dependent upon the N-terminal acetylation of tropomyosin.<sup>40,41</sup> Finally, N-terminal acetylation plays a key role in protein folding. N-terminal acetylation of  $\alpha$ -synuclein, a key factor in Parkinson's disease, leads to an increase in helical folding and a resistance to aggregation.<sup>42</sup>

In addition, NATs have been implicated in several pathological conditions. Studies have shown a relationship between the catalytic subunit of NatA and many neurodegenerative disorders.<sup>4, 6, 16,43</sup> In 2011, a S37P mutation in the NatA catalytic domain impairs its catalytic activity and the complex formation. It is not known which specific proteins are not acetylated due to this mutation; however, given that N-terminal acetylation occurs on approximately 80% of all proteins and NatA is one of the major enzymes which catalyze this modification, many proteins might be affected. Such mutation causes lethal X-linked Ogden syndrome, which results in severe mental delays, fetal death, an aged appearance, craniofacial anomalies, and cardiac arrhythmias for boys.<sup>44</sup> NatA is also vital to prevent Htt aggregation since knockdown of N-terminal acetylation leads to an increase in aggregation of Huntingtin (Htt), a key protein in Huntington's disease.<sup>45</sup> NatA coexpression with  $\beta$ -amyloid precursor protein was found to suppress  $\beta$ -amyloid protein secretion, which generates the main component of amyloid plaques that are a signature biomarker of Alzheimer's disease.<sup>46</sup> As mentioned above, N-terminal acetylation of  $\alpha$ -synuclein has a significant impact on preventing its aggregation, which is one of the hallmarks of Parkinson's disease.<sup>42</sup> In addition to NATs involvement in neurodegenerative disorders, the NAT family is also linked to cancer. Naa15, the auxiliary subunit of NatA, has been found overexpressed in papillary thyroid carcinoma,

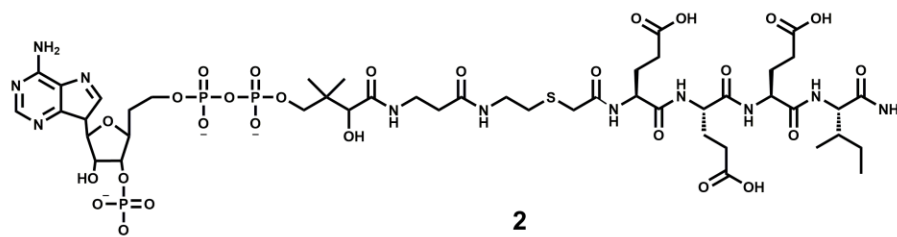
neuroblastomas and gastric cancer.<sup>47-49</sup> Elevated Naa10 expression has also occurred in breast cancer, colorectal cancer, lung cancer, and hepatocellular carcinoma.<sup>50-52</sup> Naa10 overexpression is implicated in increased cell proliferation by promoting cells to pass the cell cycle checkpoints.<sup>4,37</sup> In addition, NatB is overexpressed in hepatocellular carcinoma and plays a vital role in the cell cycle progression. Knockdown of the catalytic subunit of NatB led to the fraction of cells in the G<sub>0</sub>/G<sub>1</sub> phase to decrease while the cells in the G<sub>2</sub>/M phase to increase. This is an indicator that NatB is critical for cell-cycle progression.<sup>53</sup> Depletion of NatB in HeLa cells exhibited p21 upregulation, cell cycle arrest and p53 induction. Depletion of NatC in Hela cells reduced cell proliferation, induced p53 expression and p53-dependent cell death.<sup>16</sup>

The NAT family has shown to be a potential therapeutic target for a variety of diseases including cancer, Huntingtins disease and Ogden's syndrome. Therefore, NAT inhibitors would be very valuable chemical tools to elucidate its functions. Currently, there are three known NAT bisubstrate inhibitors (Figure 2). Each inhibitor contains the CoA portion of the Ac-CoA substrate linked to a peptide with amino acids specific for the NatA

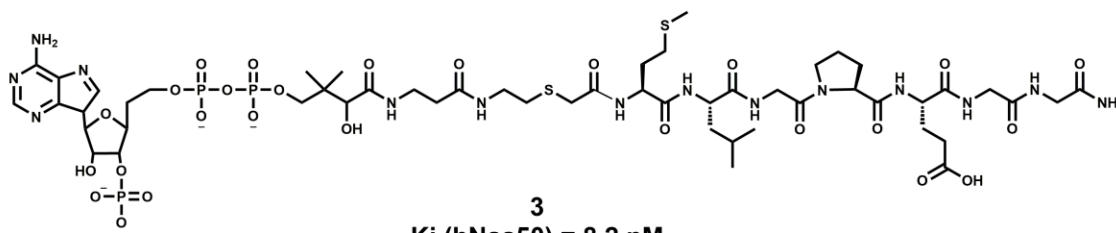
complex, NatA catalytic subunit or NatE catalytic subunit. Inhibitor **1** targets the NatA complex and contains a peptide portion derived from high mobility group protein A1.



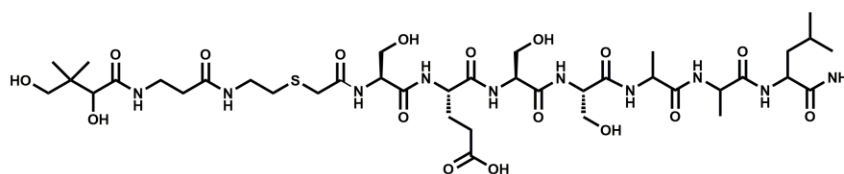
**1**  
 $IC_{50}$  (hNatA) = 15  $\mu$ M



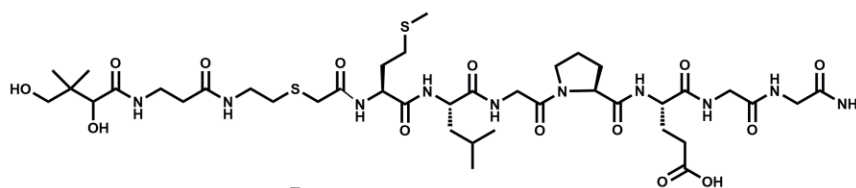
**2**  
 $K_i$  (hNaa10) = 1.6  $\mu$ M



**3**  
 $K_i$  (hNaa50) = 8.2 nM



**4**  
 $IC_{50}$  (hNaa10) = >1000  $\mu$ M



**5**  
 $IC_{50}$  (hNaa50) = 310  $\mu$ M

**Figure 2.** NAT inhibitors and their respective inhibitory activities<sup>54</sup>



Inhibitor **2** targets the catalytic subunit of NatA and incorporates a peptide derived from  $\gamma$ -actin while inhibitor **3** contains a portion of hnRNP F protein that targets the catalytic subunit of NatE. Inhibitor **1** has an  $IC_{50}$  of 15  $\mu$ M, while inhibitors **2** and **3** have  $K_i$  values of 1.6  $\mu$ M and 8.2 nM, respectively. Additionally, compounds **4** and **5** were synthesized without the adenosine ring and showed minimum inhibition, signifying that importance of the adenosine ring. These inhibitors have exhibited the different inhibition profiles for the NatA complex and the NatA catalytic subunit, which supports that the auxiliary subunit plays a role in substrate and inhibitor specificity.<sup>54</sup>

## 1.2 $\alpha$ -N-terminal Propionylation

There are approximately 18 proteins that have been identified and verified to be propionylated at their  $\alpha$ -N-termini *in vivo*.<sup>6,55,56</sup> Among them, 10 proteins carry the recognition motif that may be acetylated by either NatA, NatB or NatD based on the initial sequence. Additionally, both catalytic subunits of NatA and NatE have been found to catalyze the addition of a propionyl group from Prop-CoA onto the N-terminus of NatA and NatE substrates. Therefore, it is believed that the NAT family should be capable of propionylation.<sup>6</sup>

Compared with N-terminal acetylation that occurs on 80-90% of human proteins, propionylation on the N-terminus is less frequent, although the predicted amount of N-terminal propionylation is between 5 to 20%.<sup>6</sup> This severe disparity may result from the low availability of Prop-CoA, which is almost 20 times lower than the cellular concentration of Ac-CoA (3-30  $\mu$ M).<sup>57,58</sup> To investigate if the ratio of Ac-CoA/Prop-CoA is a key factor to dictate the predominant acetylation on the N-terminus, kinetic studies were carried out with three representative NatA substrate peptides. NatA introduced both acetylation and

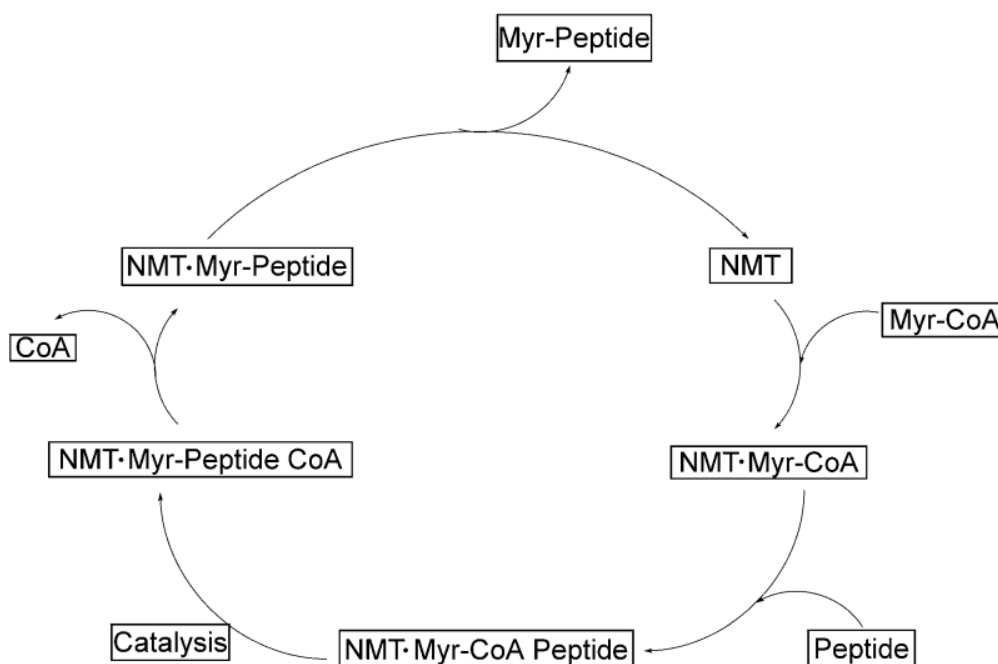
propionylation on  $\alpha$ -N termini of all three peptide substrates; however, the amount of acetylation ranged 3- to 20-fold higher compared to propionylation.<sup>6</sup> Although the  $k_{cat}$  for Ac-CoA and Prop-CoA were similar at  $0.78 \text{ min}^{-1}$  and  $0.70 \text{ min}^{-1}$ , respectively, the  $K_m$  values for Ac-CoA ( $K_m=18 \text{ }\mu\text{M}$ ) was about 3-fold lower than Prop-CoA ( $K_m=43 \text{ }\mu\text{M}$ ).<sup>6</sup> These kinetic differences validate that the scarcity of N-terminal propionylation is due to the low availability of Prop-CoA. Although the function of propionylation is not defined yet, propionylation does increase bulkiness and hydrophobicity with the addition of an extra methylene group as compared to acetylation. This modification may alter its ability to interact with proteins and vary downstream effects.<sup>59</sup>

### **1.3 $\alpha$ -N-terminal Myristoylation**

Although myristic acid is reported to be less than 1% of the total fatty acid in cells,<sup>60</sup> and only 0.5% of proteins in the cell are N-terminally myristoylated,<sup>61,62</sup> myristoylation has a wide array of relevance and function among posttranscriptional modifications for proteins. Proteins actin, gelsolin and p21-activated kinase 2 (PAK2) are N-terminally myristoylated post-translationally following cleavage of caspase-3, which leads to the up-regulation of apoptosis.<sup>9</sup> Additionally, the co-translational myristoylation of  $G_\alpha$ -protein assists in plasma membrane targeting and G-protein signaling.<sup>63</sup> N-Myristoyltransferase (NMT) is an enzyme that catalyzes the transfer of myristic acid from myristoyl-CoA, a 14-carbon fatty acid, to the N-terminal Gly residue of its substrates. NMT is a cytosolic monomeric enzyme that is present only in eukaryotes and not prokaryotes.<sup>64,65</sup> NMT plays a vital role in subcellular localization, protein folding, and enhancing hydrophobicity to increase protein-protein and protein-membrane interactions.<sup>66-68</sup> Studies have also confirmed that myristoylation is able to initiate the first step of the protein maturation

process and other conformational changes when in combination with either  $\text{Ca}^{2+}$  binding or phosphorylation.<sup>69,70</sup>

NMT is a part of the GCN5 acetyltransferase superfamily of proteins and is present as two forms, NMT-1 and NMT-2.<sup>71</sup> NMT-1 exists as four different isoforms, ranging in molecular weight from 49-69 kDa, while only one form of NMT-2 has been reported at 65 kDa.<sup>72</sup> NMT-1 has been shown to be vital for vegetative cell growth while knockdown of NMT-1 inhibits tumor progression.<sup>73,74</sup> NMT-2 has not been as extensively studied; however, elevated levels of NMT-2 are found in colorectal cancer and human brain tissues.<sup>75,76</sup> The NMT crystal structure (PDB 2NMT, 2.9Å) contains a characteristic symmetric two-fold saddle-shaped  $\beta$  sheets bordered by  $\alpha$  helices. The two binding sites of NMT are the myristoyl-CoA binding site, located in the N-terminal half of the enzyme, and the peptide binding site formed at the C-terminal half.<sup>77</sup>



**Figure 3.** NMT ordered bi-bi catalytic cycle<sup>78</sup>

The NMT family undergoes an ordered bi-bi kinetic mechanism. For proteins to undergo myristoylation, a two-step process must occur (Figure 3). The myristoyl-CoA substrate binds to NMT first, followed by the N-terminal Gly peptide substrate. Catalysis may then occur producing an N-myristoylated peptide.<sup>79-81</sup> There are many confirmed NMT substrates including various Tyr kinases, nitric oxide synthase, the  $\alpha$  subunits of heterotrimeric G-proteins, and many virus proteins including HIV-1.<sup>64,82</sup> A series of peptides originating from the N-terminal sequence of NMT substrates were synthesized and kinetically evaluated against NMT. The peptide substrates had  $K_m$  values which ranged from 16  $\mu$ M-100  $\mu$ M while myristoyl-CoA had a  $K_m$  value of 5.8  $\mu$ M.<sup>71</sup> A summary of the kinetic parameters of the NMT substrates are shown in Table 3.

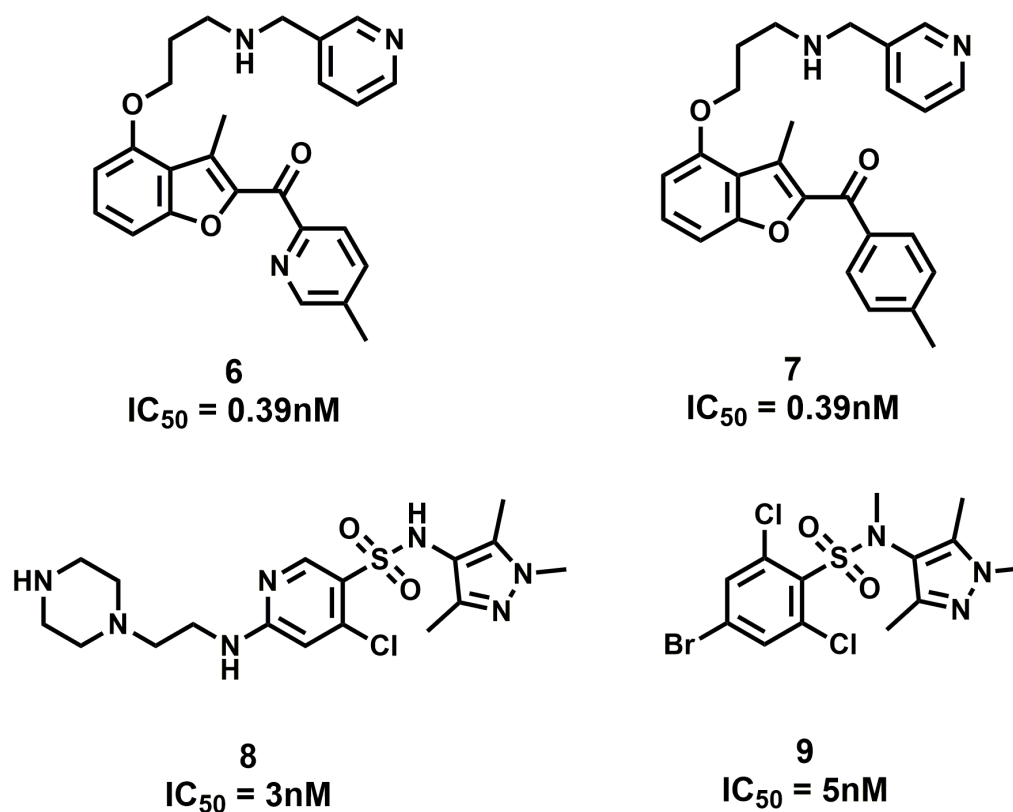
**Table 3.** Kinetic parameters of recombinant NMT from bovine cardiac muscle<sup>71</sup>

Substrate	Sequence	$K_m$ ( $\mu$ M)	Relative $V_{max}$
pp60 <sup>src</sup>	GSSKSKPKR	16	100
cAMP-dependent protein kinase	GNAAAKKRR	100	89
M2 gene segment	GNASSIKKK	50	27
Myristoyl-CoA	C14:0	5.8	100

NMT has been identified as a promising target in many therapeutic areas including anticancer, antifungal and antiparasitic. NMT-1 is a confirmed biomarker for the detection of breast, ovarian, lung and colon cancer.<sup>83,84</sup> Additionally, NMT has shown elevated activity in adenocarcinoma and gallbladder cancer.<sup>84-86</sup> NMT is also an antifungal target due to its vital role in the eukaryotic cell. Experiments show that many species of fungi cannot survive without NMT.<sup>87</sup> NMT is a single-copy gene in protozoan parasites and many NMTs from these parasites have been identified as valid drug targets.<sup>88,89</sup> NMT

exists in the human parasite *P. falciparum*, a primary agent in malaria, and inhibition of NMT lead to failure of the assembly of the inner membrane which is critical for the parasitic life cycle.<sup>89,90</sup> Given the wide clinical application potential of NMT, NMT inhibitors have been widely discovered and synthesized.

While over sixty NMT inhibitors have been reported for all three disease areas, there are many classes of NMT inhibitors for antifungal properties including benzofurans, peptidomimetics, and benzothiazoles. Two benzofuran derivatives, compounds **6** and **7**, showed IC<sub>50</sub> values of 0.39 nM against *C. albicans* NMT,<sup>91,92</sup> while peptidomimetic inhibitors have yet to show *in vivo* antifungal activity due to poor membrane permeability. The most potent inhibitor to date, compound **8**, is a pyrazole sulfonamide antiparasitic with an IC<sub>50</sub> value of 3 nM against human NMT.<sup>93</sup> This drug cured mouse models with stage 1 infection of *T. brucei* and *T. brucei rhodesiense*,<sup>93</sup> however, it was unable to cross the blood-brain barrier.<sup>94</sup> Recent studies infer the application of NMT inhibitors as novel anticancer drugs. A benzenesulfonamide inhibitor, compound **9**, had an IC<sub>50</sub> value of 5 nM against human NMT and inhibited the growth of breast and colon cancer cell lines at GI<sub>50</sub> values of less than 1 μM.<sup>95</sup> NMT is an ideal drug target for a variety of clinical fields and the amount of successful inhibitors discovered reflects that (Figure 4).



**Figure 4.** NMT inhibitors and their respective IC<sub>50</sub> values<sup>78</sup>

#### 1.4 $\alpha$ -N-terminal Palmitoylation

The addition of a palmitoyl group (C16) to proteins is important in cellular localization, protein-protein interactions, protein trafficking and protein stability.<sup>96,97</sup> This modification commonly occurs as S-palmitoylation by attaching a reversible thioester linkage to the side chain of Cys residues. However, N-palmitoylation occurs when a Cys residue is at the  $\alpha$ -N-terminus of the protein to undergo an S-palmitoylation intermediate and followed by a chemical rearrangement to yield a more stable N-palmitoylation at the  $\alpha$ -N-terminus. This N-palmitoylation was first discovered after analysis of the secreted morphogen Sonic Hedgehog.<sup>98</sup>

The Sonic Hedgehog (Shh) is a secreted signaling protein that plays significant roles in the brain during embryonic development.<sup>99</sup> Initially, Shh is synthesized as a 45-kDa precursor for the secretory pathway, but later undergoes autocleavage which results in a 19-kDa N-terminal signaling protein (ShhN). ShhN is the only known human protein that undergoes  $\alpha$ -N-palmitoylation.<sup>98,100</sup> Therefore, the enzyme that catalyzes this modification is named the Hedgehog acyltransferase (Hhat). Hhat is a member of the membrane-bound O-acyl-transferase (MBOAT) family which can be categorized into three groups based on function. The first group acylates a hydroxy group during neutral lipid biosynthesis, the second group acylates an amino acid within a protein, and the third acylates a lysophospholipid for phospholipid remodeling. Hhat falls in the second MBOAT family group which are enzymes involved in protein acylation.<sup>101,102</sup> The palmitoylation increases hydrophobicity of ShhN, which is essential in the strength of signaling and generation of a protein gradient in developing embryos.<sup>103</sup> Mice having palmitoylation-deficient ShhN, mutant Hhat or Hhat-knockout showed severe defects in limb development and neural tube.<sup>104</sup>

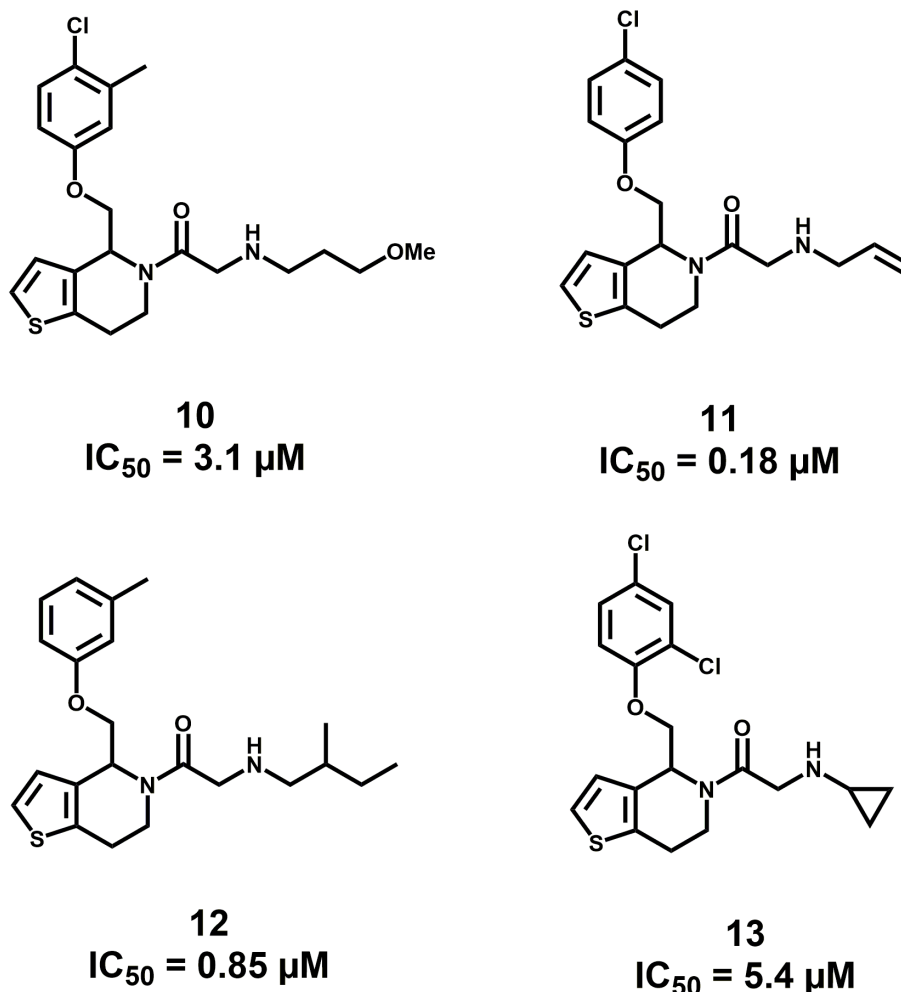
The first biochemical assay performed to verify Hhat specificity for ShhN was established in 2008. The mechanism of Hhat is dependent upon the presence of palmitoyl-CoA, which is 70-80  $\mu$ M in physiological conditions,<sup>105</sup> the substrate donor, and the availability of a free amino Cys residue. The  $\alpha$ -N-palmitoylation did not take place when the Cys residue was replaced with Ala or Ser residues. It was also found that a peptide containing an 11-amino acid sequence of the ShhN protein was sufficient for the palmitoylation to occur *in vitro*.<sup>100</sup> However, in 2012, studies revealed that the first six amino acids of mature ShhN (CGPGRG) are sufficient for palmitoylation by Hhat and the

N-terminal Cys amino acid is preferred but can be replaced by Ser. A basic amino acid at the fifth position is vital, but can be rescued if placed at position 7.<sup>106</sup> Lastly, it was identified that colocalization of Hhat and Shh occurred in the endoplasmic reticulum and Golgi.<sup>100</sup>

All acyltransferases within the MBOAT family have multiple hydrophobic transmembrane domains which has made the biochemical characterization of these enzymes difficult.<sup>102</sup> Multiple point mutations and truncations of Hhat were performed to analyze the effect of enzyme activity. Truncations at the N- and C-terminus resulted in reduced palmitoylation activity and protein stability.<sup>107</sup> Additionally, a conserved region of residues among MBOAT enzymes (residues 196-234) were found to be required for Hhat activity.<sup>107</sup> 11 point mutations and 10 deletion mutants were carried out within Hhat, and the majority had comparable stability as WT Hhat; however, they did cause a loss of catalytic activity. Two key residues were identified to be critical for catalytic activity or substrate recognition. Mutation of Asp339 to Ala resulted in a complete loss of WT activity, indicating this residue is vital for catalytic activity.<sup>107</sup> Residue H379 is conserved among the active site of all MBOAT enzymes and is verified to be critical for catalysis<sup>108-110</sup>; however, a H379A mutation in Hhat only led a 50% reduction in WT Hhat activity.<sup>107</sup> This suggests that this His residue may be more critical for recognition and binding of Shh and not catalytic activity.



Abnormal Hhat signaling has shown to contribute to the growth of many cancers including lung squamous cell carcinoma, breast cancer, and the proliferation of pancreatic cancer cells.<sup>111–114</sup> Additionally, Hhat knockdown has also shown to reduce tumor growth in a pancreatic cancer mouse model supporting Hhat as a promising anticancer target.<sup>114</sup> Currently, there are only four known inhibitors for Hhat. High-throughput screening of a library of 63,885 compounds using an assay that monitors Hhat-mediated Shh palmitoylation was performed. An additional screen using a cell viability assay was then executed and led to 95 confirmed hits. The top four compounds named **10**, **11**, **12**, and **13** (Figure 5) were selected on their inhibitory activity and drug-like



**Figure 5.** Hhat inhibitors and their respective IC<sub>50</sub> values<sup>115</sup>

properties. Compound **12** was confirmed to demonstrate selective inhibition of Hhat that blocks Hhat-mediated Shh palmitoylation in cells.<sup>115</sup> Thus, inhibitor **12** was selected for further studies in animal studies, but **12** only had a short half-life of 17 min *in vivo*. Hhat inhibition offers a new anticancer target for cancers with Shh overexpression.<sup>107,115</sup>

### **1.5 $\alpha$ -N-terminal Ubiquitination**

Ubiquitination or ubiquitylation is another common posttranslational modification that adds a ubiquitin molecule onto substrate proteins. Ubiquitin is a 76 amino acid long, 8.5 kDa regulatory protein.<sup>116</sup> Initially, this modification was only known to occur on the side chains of Lys residues via an isopeptide bond. However, now it is understood that this modification also occurs on Ser/Thr side chains through an ester bond, Cys residues through a thioester bond, or on the N-terminus through a peptide bond.<sup>117–119</sup> The first verified protein to be ubiquitinated on the  $\alpha$ -N-terminus was in 1998 on the myogenic transcriptional switcher (MyoD) protein.<sup>120</sup> Currently, there are over twenty proteins, many which are naturally Lys-deficient, that undergo this modification solely on the N-terminus and not on the internal Lys.<sup>121–139</sup>

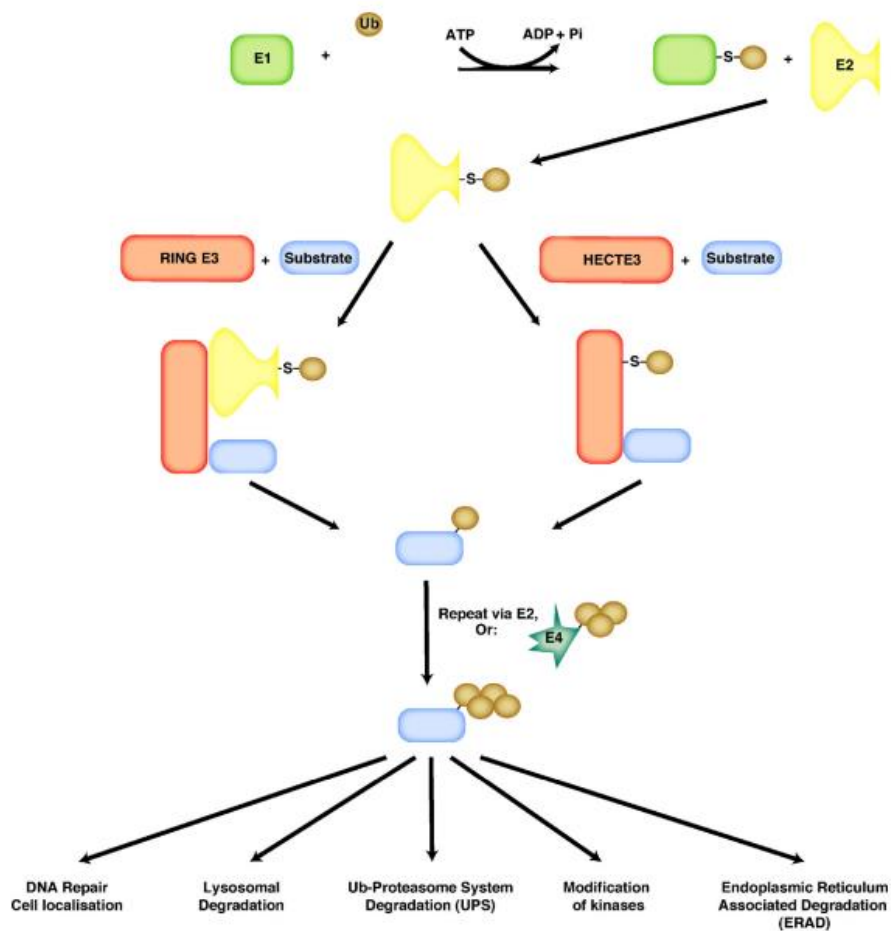
Ubiquitination has many functions including subcellular localization, subnuclear trafficking, DNA damage repair, activation or inactivation of proteins, modulating protein-protein interactions, and modification of kinases.<sup>118,140–142</sup> However, the main role associated with ubiquitination is as a signal for protease or lysosome degradation in eukaryotic cells.<sup>118,119</sup> The first study to indicate the role of N-terminal ubiquitination in protein degradation was with the MyoD protein. Nine internal Lys residues exist in the MyoD protein with the potential to be ubiquitinated and initiate the signal for degradation. All nine Lys residues were replaced; however, there was no significant decrease in the

degradation or conjugation of MyoD *in vivo* or *in vitro*. However, when all internal Lys residues were retained and the N-terminus was carbamylated, MyoD was stable *in vivo* and *in vitro*.<sup>120</sup> This was the first indicator that a protein can be prone to degradation through ubiquitination on the N-terminus and not through an internal Lys. Proteins that are susceptible to N-terminal ubiquitination are any proteins that have a free, unmodified N-terminus. Approximately 75% of all cellular proteins are N-terminally acetylated; therefore, the remaining 25% are potential substrates for N-terminal ubiquitination.<sup>10</sup>

There are many eukaryotic proteins that share the common ubiquitin fold and are often termed ubiquitin-like proteins. However, many of these proteins are structurally similar to ubiquitin but do not undergo conjugation to other proteins.<sup>143</sup> Ubiquitin itself is highly conserved among yeast and humans with only 3 of 76 amino acids varying.<sup>144</sup> The process of ubiquitination requires three different enzyme types. The first enzyme is the Ub-activating enzyme (E1) that adenylates the ubiquitin C-terminus using energy from ATP hydrolysis and forms an ubiquitin-adenylate intermediate.<sup>145</sup> The C-terminus of ubiquitin then binds to a Cys residue in E1 through a thioester linkage and releases AMP.<sup>10</sup> The next enzyme involved is an Ub-conjugating enzyme (E2). The ubiquitin molecule is transferred from E1 to a Cys residue on E2, where there are at least 50 in the mammalian genome.<sup>146,147</sup> Lastly, the E2 enzyme transfers ubiquitin to the protein substrates with the assistance of an Ub-ligase enzyme (E3). There are hundreds of E3 enzymes and they serve as the substrate recognition enzymes which have the ability to interact with both E2 and the protein substrates. This hierarchical cascade permits tight

regulation of the ubiquitination process (Figure 6).<sup>118</sup> This process can be repeated many times to add additional ubiquitin moieties and create a polyubiquitin chain.<sup>148</sup> In few cases, there are E4 enzymes that add a pre-formed polyubiquitin chain to the protein.<sup>149</sup> A target protein must contain a minimum of four linear ubiquitin moieties before it is recognized by proteases for degradation.<sup>150</sup> Although, it is not fully understood why certain proteins are targeted for degradation.

The last amino acid in the ubiquitin sequence, Gly76, will attach itself to its protein substrate at the N-terminus or Lys side chain. If the protein is then polyubiquitylated, further ubiquitin compounds will link their Gly76 to one of seven



**Figure 6.** Ubiquitin hierarchal cascade and varying downstream effects<sup>119</sup>

possible Lys residues or the N-terminus of the initial ubiquitin moiety.<sup>151</sup> The attachment of the ubiquitin at various Lys residues or N-terminus often indicates varying function for the target protein. Figure 6 shows the ubiquitination cascade and the varying downstream roles.<sup>119</sup> Once proteins are targeted for degradation, the proteins are swiftly degraded to smaller peptides and the ubiquitin subunits are cleaved and recycled for later use.<sup>152</sup>

The ubiquitin pathway is highly involved in a variety of clinical fields including neurodegenerative diseases, cancer, viral diseases, and cardiovascular diseases.<sup>153,154</sup> For example, the E3 Ub-ligase, Parkin, is associated with early-onset autosomal recessive forms of Parkinson's disease.<sup>155</sup> Additionally, another E3 ligase named HUWE1, is overexpressed in multiple cancer types including breast, lung and colorectal carcinomas.<sup>156</sup> Also, overexpression of MDM2 (HDM2), another E3 ligase, is also present in breast, lung, glioblastomas, esophageal carcinomas, and malignant melanomas and promotes the degradation of p53, the guardian of the genome.<sup>157</sup> Given the wide therapeutic potential of the ubiquitin pathway inhibitors, it is not surprising that a generous amount of inhibitors have already been discovered. In 2003, Bortezomib was the first drug approved by the FDA as a proteasome inhibitor for the treatment of refractory multiple myeloma.<sup>158</sup> It was also in clinical trials for diffuse large B cell lymphoma, non-Hodgkin's lymphoma and many other cancers.<sup>159</sup> Since then, another proteasome inhibitor, Carfilzomib, has been approved by the FDA and many other second-generation proteasome inhibitors are currently being developed.<sup>160</sup> Small-molecule inhibitors have been discovered and developed primarily through high-throughput screening techniques and have targeted each stage of the ubiquitination cascade.<sup>161</sup>

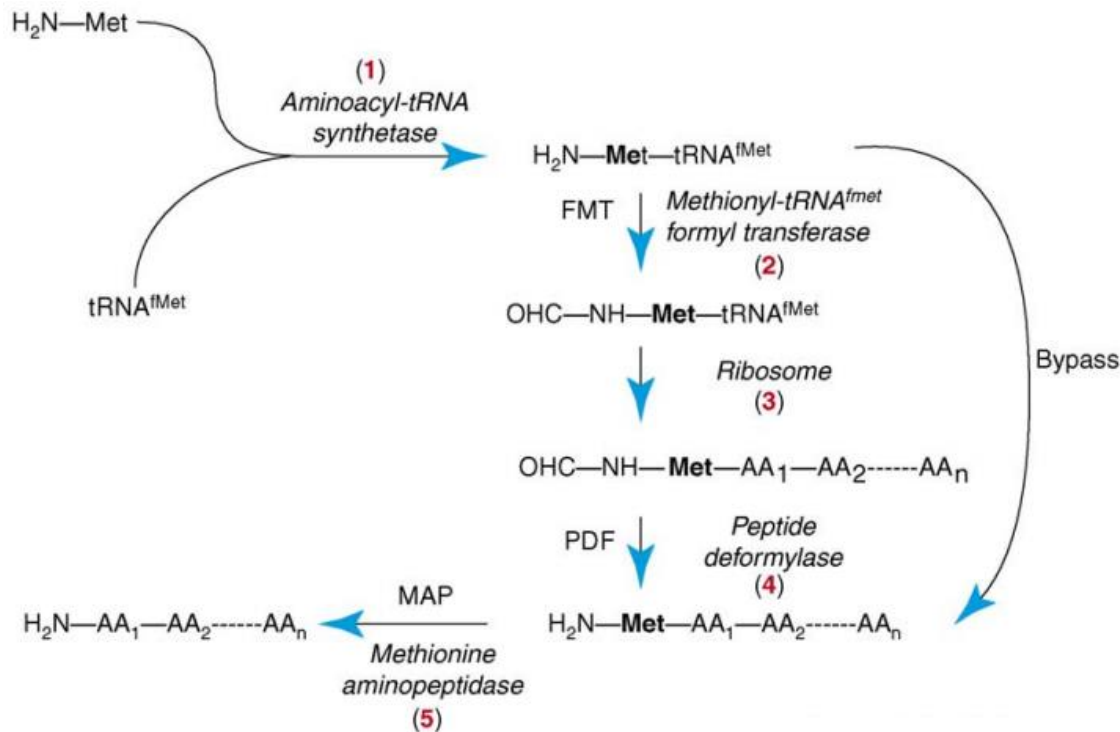
## 1.6 $\alpha$ -N-terminal Formylation

N-terminal formylation occurs during protein synthesis of bacteria and organelles.<sup>162</sup> The enzyme methionyl-tRNA formyltransferase (FMT) transfers a formyl group from 10-formyltetrahydrofolate to the N-terminus of Met residues.<sup>21</sup> The formylation of Met was first discovered in 1964 is now understood to signal the initiation of protein synthesis.<sup>162,163</sup> This reaction occurs in bacteria and organelles,<sup>21</sup> and formyl-Met (fMet) is the first residue that appears from bacterial ribosome. Although N-formylation is conserved among most bacteria, it is not understood which proteins are dependent upon this reaction.<sup>164,165</sup>

A proteomic study was performed to explore how mutations in FMT affects the metabolic capabilities of bacteria. The results indicated that the folic acid metabolism was altered after FMT mutation due to increased susceptibility to trimethoprim and sulfamethoxazole where the MICs were 3.5- and 3-fold lower, respectively.<sup>166</sup> The studies also verified that  $\alpha$ -N-terminally formylated peptides are critical for the oxidation of pyruvate and the FMT mutant consumed glucose less efficiently than FMT wild type. Finally, it was found that FMT mutant led to a lack of degradation of Arg through the Arg deiminase pathway,<sup>166</sup> which is needed as an ATP source by substrate level phosphorylation.<sup>167</sup> These results elucidated specific pathways that depend on N-terminal formylation; however, further studies to understand which enzymes rely on this reaction is required.

The protein synthesis cycle begins with aminoacyl-tRNA synthetase transferring a Met on to an initiating tRNA. This Met-tRNA product is now susceptible to formylation by FMT. After the fMet-tRNA complex is formed it is moved to the ribosome in order for the

protein synthesis to initiate. Next, the formyl group is removed from the initial Met by peptide deformylase (PDF) and then the Met can also be removed by Met aminopeptidases.<sup>168</sup> This protein synthesis cycle is shown in Figure 7.

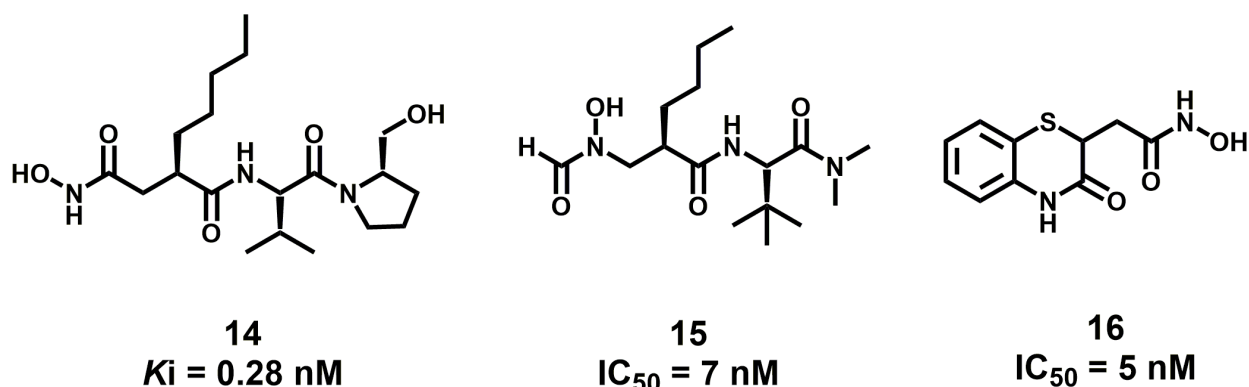


**Figure 7.** Formylation and deformylation cycle for protein synthesis<sup>168</sup>

While there are no inhibitors for FMT, designing inhibitors for PDF, the enzyme that catalyzes the removal of the formyl group, has been an attractive strategy. Given that PDF is present in prokaryotic cells and absent in mammalian cells, it is an ideal antibiotic target.<sup>169,170</sup> Inhibitor **14**, a naturally occurring tight-binding PDF inhibitor, has a  $K_i$  value of 0.28 nM.<sup>171</sup> Although **14** had poor bioavailability, the discovery of **14** led to the increase in rational design of inhibitors for PDF.

A variety of PDF inhibitors have been discovered including peptidic inhibitors and small molecule inhibitors. These inhibitors have been identified through high-throughput screening, ligand-based screening and ration drug design.<sup>172</sup> One example of a peptidic inhibitor is compound **15**, which was discovered through screening a library of metalloenzyme compounds. Compound **15**, which has a very similar scaffold as compound **14**, had an  $IC_{50}$  of 7 nM and high selectivity for PDF.<sup>173</sup> A small molecule bicyclic inhibitor, **16**, was also discovered through screening and had an  $IC_{50}$  of less than 5 nM and high selectivity for PDF, but had poor antibacterial activity.<sup>174</sup> PDF inhibitors **14**, **15** and **16** are shown in Figure 8. Inhibitors for PDF has been a verified as a validated effort given there are currently three compounds in Phase I or Phase II clinical trials.<sup>175–</sup>

177



**Figure 8.** PDF inhibitors and their respective  $IC_{50}$  values<sup>173,174</sup>

### 1.7 $\alpha$ -N-terminal Methylation

Protein  $\alpha$ -N-terminal methylation has been known for nearly four decades since it was first uncovered on bacteria ribosomal proteins L33.<sup>178</sup> Subsequent detection of this modification on histone H2B, cytochrome c-557, and myosin light chain proteins revealed that  $\alpha$ -N-terminal methylation normally occurs on an N-terminal X-P-K (X is A, P, S)



sequence.<sup>179–181</sup> Identification of the eukaryotic N-terminal methyltransferase 1 (NTMT1/NRMT1) unveiled the eukaryotic methylation writers for protein  $\alpha$ -N-termini.<sup>7, 22,182</sup> Since then, there have been an increasing number of reports of  $\alpha$ -N-terminal methylation detected on new protein substrates such as regulator of chromosome condensation 1 (RCC1), the tumor suppressor retinoblastoma 1 (RB1), oncoprotein SET (also known as I2PP2A, TAF1 $\alpha$ ), damaged DNA-binding protein 2 (DDB2), Poly [ADP-ribose] polymerase 3 (PARP3), and centromere proteins A and B (CENP-A and CENP-B).<sup>7,183–187</sup>

Protein  $\alpha$ -N-terminal methylation was originally proposed to regulate protein-protein interactions since the methylated proteins initially identified were involved in large macromolecular complexes.<sup>188</sup> Recent discoveries have demonstrated its relevance in protein-DNA interactions. The  $\alpha$ -N-terminal methylation is essential to stabilize the interactions between RCC1 and chromatin during mitosis, to localize and enhance the interaction of CENPs with chromatin, and to facilitate the recruitment of DDB2 to DNA damage foci.<sup>184–187</sup> In addition, the level of  $\alpha$ -N-terminal methylation increases in response to a variety of extracellular stimuli, including increased cell density, heat shock, and arsenite treatment.<sup>185,189</sup>

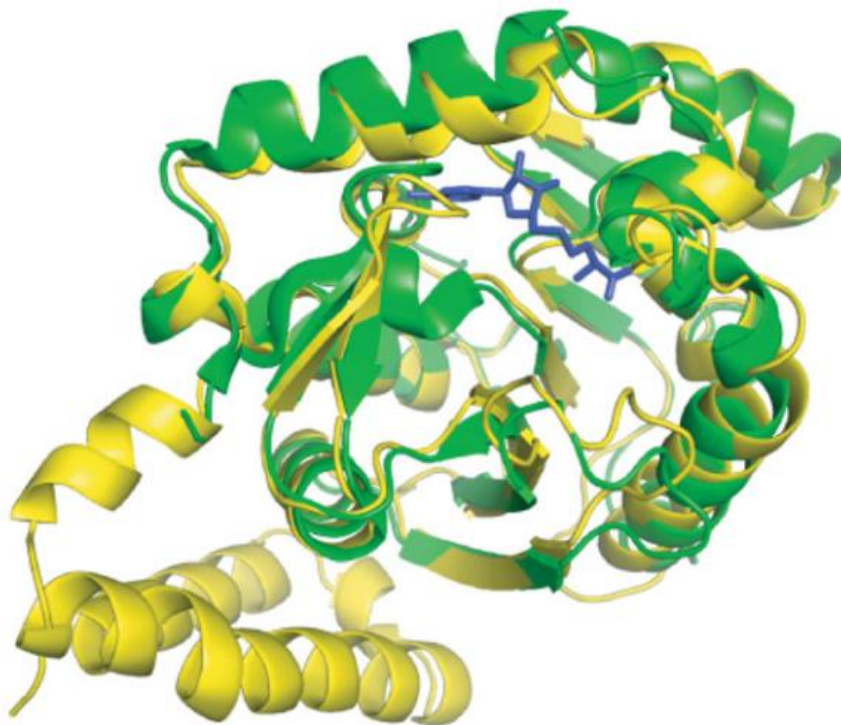
In addition to NTMT1, its close homolog NTMT2 that shares over 50% sequence similarity, has been identified as another mammalian protein  $\alpha$ -N-terminal methylation writer. Studies from Schanar-Tooley *et al.* indicated that NTMT1 and 2 share the same X-P-K recognition motif, but suggested that NTMT2 acts as a monomethylase to prime monomethylated substrates for NTMT1.<sup>22</sup> Recent structural studies from us and Li *et al.* substantiate that NTMT1 prefers an X-P-K/R (X can be any residue type other than D/E)

motif.<sup>23,24</sup> Although co-crystal structures of NTMT1 in a ternary complex with varying peptide substrates and SAH are now available, there is scarce knowledge of NTMT2 except the initial report in 2013.<sup>22</sup> One reason for the scarcity is the limited stability of recombinant full-length NTMT2; therefore, no structural information of NTMT2 is currently available. Recently, a third N-terminal methyltransferase has been discovered in yeast. Elongation factor methyltransferase 7 (Efm7) has shown to methylate eukaryotic elongation factor 1A (eEF1A), a protein that delivers aminoacyl-tRNAs to ribosome, *in vitro*. Additionally, studies verified that overexpression of Efm7 resulted in an increase in N-terminal methylation while knockdown resulted in a loss of methylation *in vivo*. Interestingly, Efm7 does not target the traditional X-P-K recognition motif and solely methylates an N-terminal Gly residue and the adjacent Lys. However, the adjacent Lys is methylated only after the N-terminal Gly is fully saturated. Nonetheless, Efm7 is the only known methyltransferase that methylates the N-terminus and an adjacent Lys residue of a single protein. Currently, eEF1A is the only known substrate of Efm7; however, there are 35 other yeast proteins with a beginning G-K sequence that have the potential to be substrates.<sup>25</sup>

Among the N-terminal methyltransferase family, only the structure of NTMT1 has been extensively studied. NTMT1 is highly conserved from yeast to humans<sup>182</sup> and many co-crystal structures of NTMT1 in complex with peptide substrates and the cofactor S-adenosyl homocysteine (SAH) have been determined.<sup>7,23,24</sup> The crystal structures show that NTMT1 folds as a single domain and include a Rossman fold that contains seven-strand  $\beta$  sheets and five  $\alpha$  helices which is common of methyltransferases. However, NTMT1 also comprises an N-terminal extension containing a  $3_{10}$  helix and two  $\alpha$  helices,

and a  $\beta$  hairpin inserted between strand  $\beta 5$  and helix  $\alpha 7$ , which is unique from other methyltransferases. The crystal structures also revealed that the peptide substrates all bind in a similar manner through a negatively charged channel connected to the cofactor S-adenosyl methionine (SAM) binding site. The N-terminal amine of the peptide substrate is pointing toward the SAM methyl donor.<sup>24</sup> This is distinct from other methyltransferase enzymes, which commonly have their peptide binding site on the surface and only insert the Lys or Arg side chain to be modified into a narrow channel pointing toward the SAM binding site.<sup>190</sup> The structural information gained from the crystal structure supported the selectivity NTMT1 has for the N-terminus over free amino side chains. Homology modeling of NTMT2 has also been performed and overlaid over the NTMT1 crystal structure (PDB: 2EX4). The NTMT1 crystal structure is shown in green with SAH shown in blue while the NTMT2 homology model is shown in yellow (Figure 9). The main distinction between the crystal structure and the model is the N-terminal 60-amino acid domain of NTMT2, which is not present in NTMT1.

Although the only endogenous substrates verified to be methylated *in vivo* begin with S/P/A/G, NTMT1 was found to methylate peptides beginning with F/Y/C/M/K/R/N/Q and H *in vitro*.<sup>24</sup> The tolerance for a variety of amino acids at the first position of the substrate can be explained by the spacious binding pocket surrounding the first position side chain. From the crystal structure, a hydrogen bond is shown between the S1 side chain and the main chain of Met30 from NTMT1.<sup>24</sup> This binding site is large enough to accommodate bulky aromatic residue Tyr and is substantiated by enzymatic binding data.<sup>24</sup> On the other hand, the negatively charged channel prevents negatively charged amino acids, like D and E, at the S1 position.<sup>7, 24,182</sup> Additionally, a vital hydrogen bond



**Figure 9.** Homology model of NTMT2 (yellow) having the same overall fold as NTMT1 crystal structure (green; PDB code 2EX4). SAH molecule in the NTMT1 crystal structure is shown in blue.<sup>22</sup>

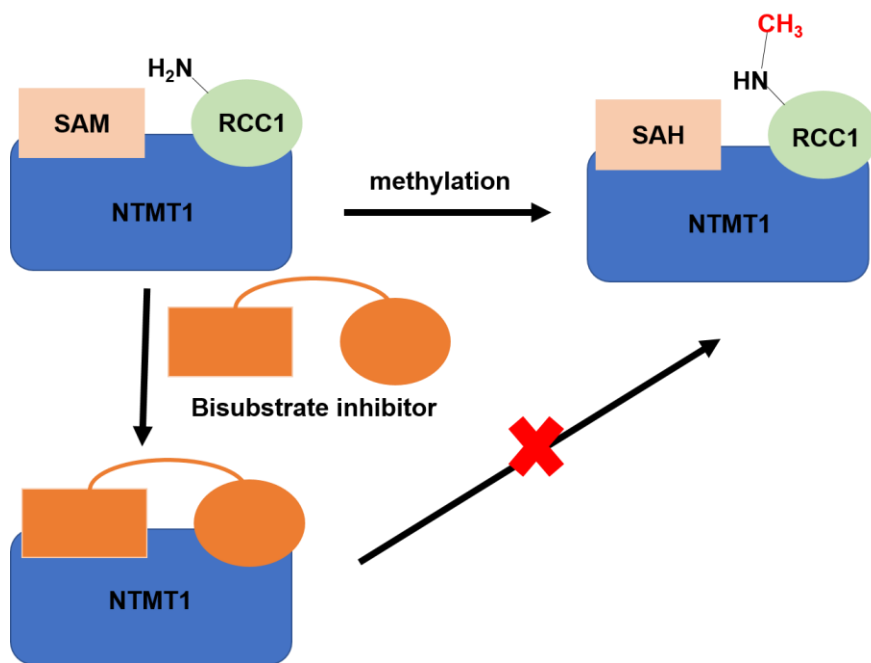
is found between the carbonyl oxygen of the first residue of the NTMT1 substrate and the carboxamide side chain of Asn168. Mutation of Asn168 to Lys resulted in a ~36-fold loss in  $K_m$ , indicating the importance of this interaction between NTMT1 and its peptide substrate. The second residue P2 is in a hydrophobic pocket formed by residues Leu31, Ile37, and Ile214.<sup>24</sup> It also forms a stacking interaction with the indole of Trp136. Mutagenesis of Trp136 or substitution of the Pro to any other amino acid resulted in a loss of enzymatic activity.<sup>24,182</sup> The third residue K3 forms two key hydrogen bonds with side chains Asp177 and Asp180 of NTMT1. NTMT1 can tolerate Lys or Arg residues at the third position of the peptide substrate.<sup>7, 24,191</sup> All known protein substrates contain the X-P-K initial sequence besides CENP-A, which contains an X-P-R motif. The amino acids found at the fourth position of NTMT1 substrates are Arg, Lys and Thr. The fourth residue

is adjacent to a negatively charged substrate binding channel.<sup>24</sup> The fifth and sixth residues, such as I5 and A6, only form a few backbone hydrogen bonds with NTMT1.<sup>24</sup>

NTMT1 catalytic transfer of methyl groups onto protein substrates takes place through a random-ordered bi-bi kinetic mechanism.<sup>192</sup> A ternary complex comprised of NTMT1, the cofactor SAM, and the protein substrate is formed with either substrate binding to NTMT1 first.<sup>192</sup> The key residues involved are the highly conserved Asp180 and His140 which act as general bases to enable the deprotonation of the  $\alpha$ -amino group of the N-terminus. The deprotonated amino group can then undergo a S<sub>N</sub>2 reaction and attack SAM to enable the transfer of the methyl group onto the peptide substrate. Additionally, crystal structures were also resolved with monomethylated peptide substrates and they had the same orientation as the unmethylated substrates.<sup>24</sup> These results support the understanding that NTMT1 follows a distributive mechanism and has no significant preference for unmethylated or monomethylated substrates.<sup>192</sup>

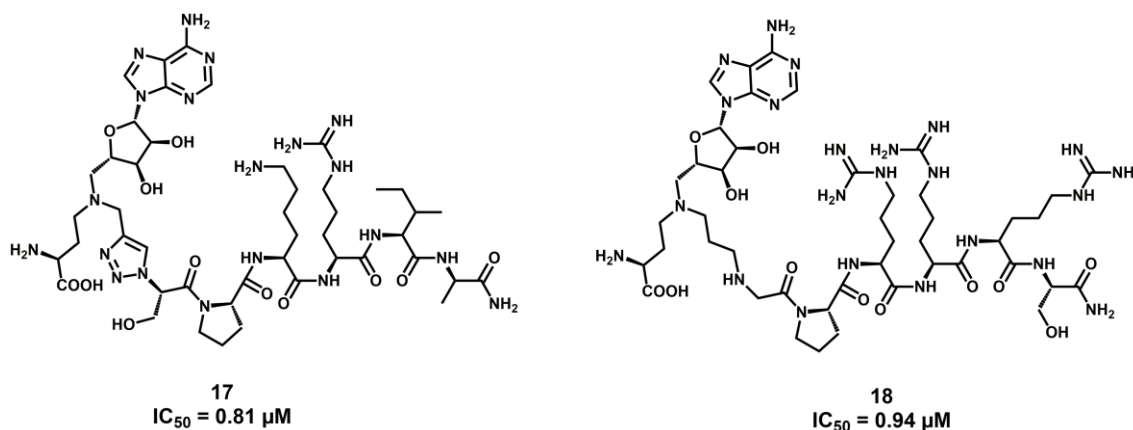
NTMT1 methylates damaged DNA-binding protein 2, which is crucial for DNA repair.<sup>184</sup> Knockdown of NTMT1 results in mitotic defects and sensitizes chemotherapeutic agents in breast cancer cell lines like MCF-7 and LCC9, while NTMT1 knockout mice showed premature aging.<sup>183,193</sup> Additionally, NTMT1 has been shown to be overexpressed in a variety of cancer cell lines and patients' tumor tissues including colorectal, melanoma, carcinoid, lung and liver according to data from ProteinAtlas.<sup>194</sup> Additionally, methylation of regulator of chromosome condensation 1 (RCC1) is essential to stabilize its interaction with chromatin throughout mitosis for proper division.<sup>185,186</sup>

Given that NTMT1 forms a ternary complex, bisubstrate inhibitors which target both binding sites have been proven to be an effective strategy to increase selectivity and potency (Figure 10). This technique has been applied in many enzymes with two binding sites including histone acetyltransferases (HATs) and protein Arg methyltransferases (PRMTs).<sup>54,195,196</sup> The first generation NTMT1 bisubstrate inhibitor is compound 17. The NAM portion of this bisubstrate inhibitor is a SAM analog; however, the sulfonium atom is replaced with a nitrogen to increase stability. The peptide portion of the inhibitor is the beginning six amino acids of RCC1. There is a triazole linker that was utilized due to its optimum size in coupling both substrate portions together. This optimum linker size was determined based on the distance found in docking studies between the sulfonium atom in SAM and the free amine of the peptide substrate. Through fluorescence-based assays and MALDI methylation assays, the  $IC_{50}$  of this compound was found to be 0.81  $\mu\text{M}$ . To verify specificity of this inhibitor, two other methyltransferases were also tested. Inhibitor



**Figure 10.** Bisubstrate strategy to inhibit methylation

**17** showed less than 15% inhibition against PRMT1 and less than 50% inhibition against protein Lys methyltransferase G9a at 50  $\mu\text{M}$ . Lineweaver Burke plots also illustrated that this inhibitor competes at both sides confirming the bisubstrate strategy was effective.<sup>197</sup> The second generation of bisubstrate inhibitors were also synthesized utilizing alkyl linkers in replace of the triazole in order to probe varying distances between the substrate portions. A series of inhibitors were synthesized and the most potent compound, **18**, had a propylene linker and an  $\text{IC}_{50}$  of 0.94  $\mu\text{M}$ . Selectivity studies against PRMT1 and G9a were also carried out and the bisubstrate inhibitor showed no significant inhibition at 30  $\mu\text{M}$  against either enzyme.<sup>198</sup> These inhibitors are the first bisubstrate inhibitors which target NTMT1 (Figure 11).



**Figure 11.** NTMT1 inhibitors and their respective  $\text{IC}_{50}$  values<sup>197,198</sup>

Bisubstrate inhibitors have been a successful strategy among methyltransferases given that they typically have higher selectivity and potency for their target. Nevertheless, bisubstrate inhibitors are often too large, not cell-permeable or drug-like. Therefore, there is a need for small molecule inhibitors to elucidate information about downstream process of protein  $\alpha$ -N-terminal methylation. While there are many reports that imply the importance of  $\alpha$ -N-terminal methylation, the function of protein  $\alpha$ -N-terminal

methylation is still in its infancy. Little progress has been made in characterizing the recognition motif and readers responsible for specifically recognizing this modification and translating it into a biological signal. Lack of such knowledge has greatly limited our understanding of its functions from a systematic viewpoint.

Given that the NTMT family has two binding sites, a small molecule inhibitor may target the SAM cofactor binding site or the peptide substrate binding site. Since SAM is a common cofactor for many methyltransferases, targeting the peptide binding site should achieve higher selectivity. Therefore, the intention of this project is to design potent and selective peptide inhibitors for NTMT1. Herein, we report the design and synthesis of the first series of selective photoaffinity probes for the NTMT family. Additionally, one approach to identify proteins which write a methyl group onto the N-terminus of proteins or recognize a methylated N-terminus of proteins is photoaffinity labeling. Photolabeling targets the active site of a specific protein where the covalent labeling is driven by specific recognition. A series of unmethylated photoaffinity probes which target the NTMT family and a series of methylated photoaffinity probes which target proteins downstream from the NTMTs have been synthesized. These photoaffinity probes have successfully labeled spiked-in NTMT1 in a complex cell mixture. Additionally, these probes have revealed novel information about substrate preferences among NTMT1 and NTMT2. These photoaffinity probes have the ability to identify information about the function of protein  $\alpha$ -N-terminal methylation.



## 2. Results and Discussion

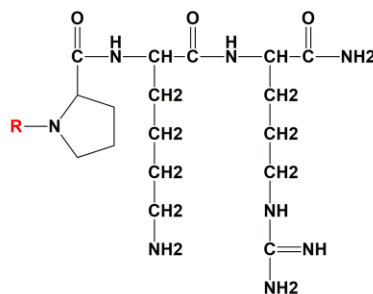
### 2.1 NTMT1 Peptidomimetic Inhibitors

#### 2.1.1 Design

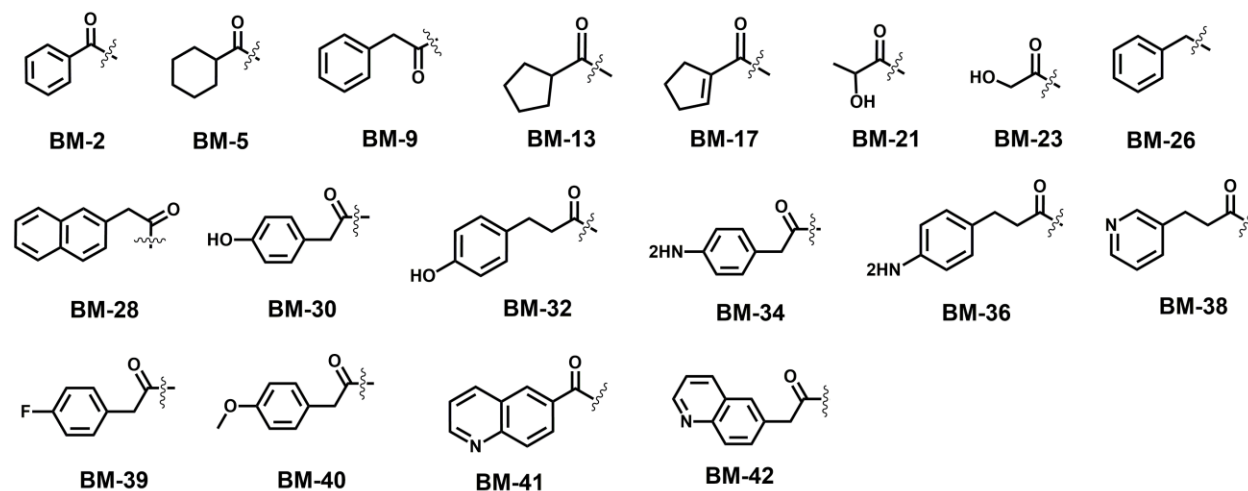
The ultimate goal of this project was to develop potent inhibitors that specifically inhibit NTMTs. Based on our crystal structures, the first amino acid at the N-terminus mainly interacts with N168 through the carbonyl amide bond and is tolerable towards NTMT1 recognition, which is substantiated by the wide range of amino acids which exist at the first position.<sup>24</sup> Given the mechanism of methylation, the  $\alpha$ -N-terminal amine is essential to serve as a nucleophile to attack SAM, but it is not critical for binding. Therefore, we hypothesized that NTMT1 peptidomimetic inhibitors can be developed through the removal of the N-terminal amine of peptide substrates. We initiated our efforts by incorporating carboxylic acids onto a tripeptide to mimic the Pro, Gly, Ser, or Tyr at the first position (Table 4). We also introduced a benzyl group at the first position to explore the importance of the first amide group.

In addition, we also explored the modifications at the second, third and fourth position in an attempt to increase inhibitory activity, as well as to enhance stability since peptide-based inhibitors are susceptible to degradation and peptidases. Pro is known to undergo *cis-trans* isomerization. From our crystal structures, Pro has only shown the *trans* conformation<sup>24</sup>; therefore, we hypothesized that the *trans* conformation is favorable to interact with NTMT1. Literature has stated that although  $\alpha$ -methyl-Pro has the ability

**Table 4.** Summary of modifications at 1<sup>st</sup> position for tetramer inhibitors



**R groups:**

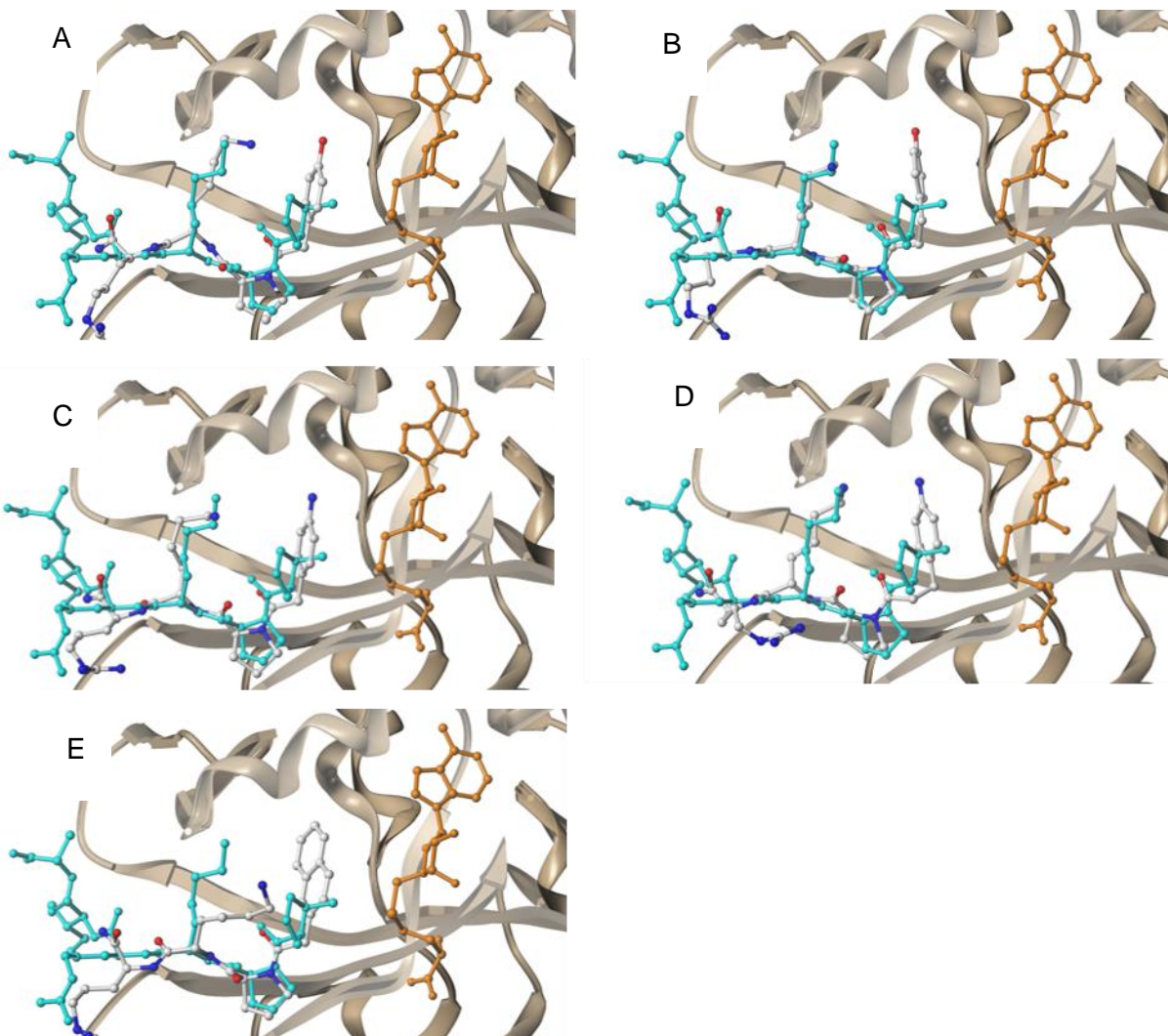


to exist as a trans or cis conformation, the cis conformation is the overwhelming majority.<sup>199</sup> Therefore, an alpha-methyl-Pro was designed at the 2<sup>nd</sup> position to lock its cis conformation and probe inhibition differences with changes at the P2 position. Arg has been detected at the 3<sup>rd</sup> or 4<sup>th</sup> position in NTMT1 substrates, so it was synthesized at either position to evaluate the inhibitory activity. Likewise, Thr was incorporated at the 4<sup>th</sup> position. Additionally, trimethylated Lys and D-Arg were designed to increase cell permeability and stability, respectively.<sup>200</sup> To investigate the optimal length for inhibitory activity, we also synthesized one peptidomimetic to mimic tripeptides and five to mimic hexapeptides.

### 2.1.2 Docking Studies

To support our hypothesis and validate the binding mode of our designed inhibitors, docking studies of the inhibitors were performed using the NTMT1 co-crystal structure (PDB id: 5E2B) using the program Gold53. The NTMT1 binding site was defined by a sphere of 6.0 Å where the substrate MePPKRIA was extracted and the inhibitors were docked into that site. The compounds for docking were prepared using SYBYL X2.1 and the energy was minimized using the external Tripos force field. The ChemPLP score from the Gold53 docking program was considered to determine the most likely docking conformations and top inhibitors.

The majority of the inhibitors overlaid well with the peptide substrate (MePPKRIA) in the binding pocket of NTMT1. The docking studies verified that there is a large site at the first position of the inhibitors which allowed for bulkier amino acid analogs at that position. The docking studies also indicated that electrostatic interactions could occur if an electronegative atom and hydrogen were introduced at the para position of the phenyl ring at the first position. This supported our design to incorporate a hydroxy phenyl and amino phenyl compound at the first position, which were substantiated by potent IC<sub>50</sub> values. The peptide inhibitors, BM-30, BM-32, BM-34, BM-36 and BM-28 overlaid the peptide substrate, MePPKRIA, are shown in Figure 12.



**Figure 12.** Superimposed structures of peptide inhibitors (atom color) with substrate peptide MePPKRIA (cyan) in the complex of NTMT1 (grey). (PDB: 5E2B) SAH is shown in orange (A) BM-30; (B) BM-32; (C) BM-34; (D) BM-36; (E) BM-28.

### 2.1.3 Synthesis

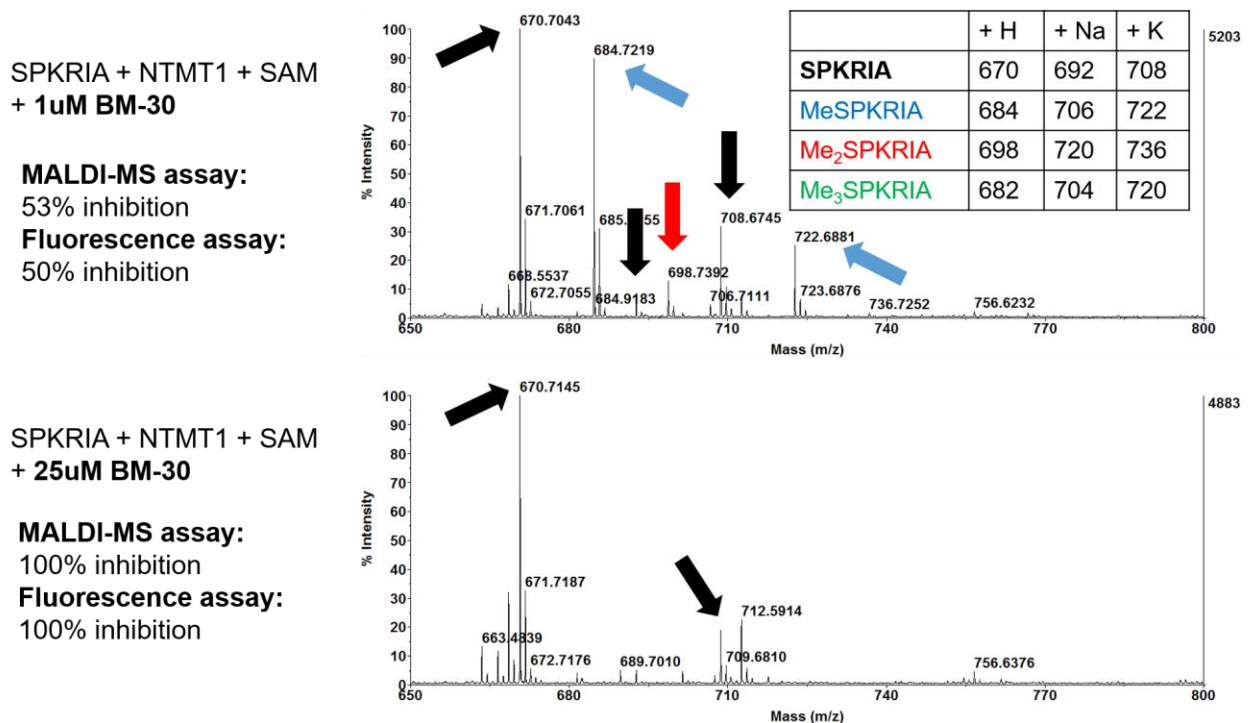
The peptides were synthesized following the standard Fmoc solid-phase synthesis on Rink amide resin using a CEM Liberty microwave peptide synthesizer. The peptidomimetics that have carboxylic acids at the first position were prepared through the standard amino acid coupling reaction except for BM-45 which has a trimethylated Lys at its 4<sup>th</sup> position. First, alloc protected Lys was manually coupled on to the Rink amide resin

following the standard Fmoc strategy. Next, orthogonal deprotection of the alloc group was performed using 3eq of I<sub>2</sub> in 4:1 ACCN:H<sub>2</sub>O and shaken for 48 hours.<sup>201</sup> followed by trimethylation on the Lys side chain amine<sup>202</sup> with an 80% yield after HPLC purification. Then the remaining amino acids were coupled following standard coupling procedure through the automatic peptide synthesizer (Scheme A1). For BM-26 and BM-27, the two peptidomimetics with a benzyl group at the first position, benzyl bromide was used as the starting material and added to the PKR and PKT tripeptide on resin through an S<sub>N</sub>2 reaction (Scheme A2). All compounds were cleaved from the solid support in a solution of trifluoroacetic acid/2,2'-(ethylenedioxy)diethanethiol/H<sub>2</sub>O/triisopropylsilane (94:2.5:2.5:1). All peptide inhibitors were purified by reverse-phase HPLC (Waters) and characterized by mass spectrometry (Figure A1).

#### **2.1.4 MALDI-MS Studies**

To determine if the peptide inhibitors were possible substrates, the inhibitors underwent a methylation progression study monitored via MALDI-MS as described before.<sup>203</sup> NTMT1 was incubated with the peptide inhibitors, and after 5 min, SAM was added to initiate methylation. The reaction was quenched after one hour and the compounds were analyzed through MALDI-MS and processed in Data Explorer. The peptide inhibitors BM-11, BM-30, BM-32, BM-34, BM-39, and BM-42 were subjected to NTMT1 methylation and none showed any methylation, indicating that they are not NTMT1 substrates.

To validate our inhibition values obtained from our fluorescence-based methyltransferase assay, a MALDI-MS methylation progression assay was performed following a similar procedure as described above. NTMT1 was incubated with RCC1-6 and the peptide inhibitors BM-11, BM-30, BM-34, BM-46 and BM-47 at two concentrations (1  $\mu$ M and 25  $\mu$ M). After 5 min, SAM was added to initiate methylation of RCC1-6. The reaction was quenched after one hour.<sup>203</sup> The methylation progression was monitored via MALDI-MS and data was analyzed in Data Explorer by quantifying the area under the curve to determine the percentage of inhibition.<sup>203</sup> The peptide inhibitors tested showed similar inhibition values in this MALDI-MS methylation inhibition assay compared to the fluorescence-based methyltransferase assay. Figure 13 shows the MALDI-MS methylation progression after the addition of the two concentrations of BM-30. The inhibition values found in this assay are the same or comparable to the inhibition values



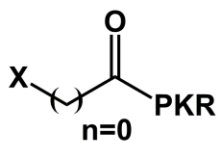
**Figure 13.** MALDI-MS methylation inhibition assay for peptide inhibitor BM-30

found in the fluorescence assay. The four other peptide inhibitors tested are shown in Figure A2-A5.

### 2.1.5 Inhibition Studies

All synthesized peptidomimetics were evaluated in our established SAH-coupled fluorescence assay.<sup>192</sup> The  $K_m$  values for both peptide substrate and SAM were determined as described before. The assay was performed at 100  $\mu\text{M}$  of SAM and at the  $K_m$  value of peptide substrate. Initial screening was carried out at the concentration of 5  $\mu\text{M}$  and 30  $\mu\text{M}$  of each compound. Those compounds that exhibited more than 50% of inhibition at 30  $\mu\text{M}$  were subjected to  $\text{IC}_{50}$  studies. Percentage of inhibition is calculated by subtracting the slope of [SAH] production vs. time for each concentration of inhibitor from the slope of [SAH] production vs time of the positive control with no inhibitor present.  $\text{IC}_{50}$  studies were carried out by ranging the inhibitor concentration from 0.14  $\mu\text{M}$  – 100  $\mu\text{M}$  in a three-fold dilution. They were incubated with 0.1  $\mu\text{M}$  NTMT1 and the reaction was initiated with the peptide substrate, RCC1-6 at  $K_m$  value. Fluorescence intensity was monitored for 15 min and the rates were fit to the log[inhibitor] vs response model using least squares nonlinear regression. A summary of the kinetic data for the peptide inhibitors is shown in Table A1 and  $\text{IC}_{50}$  curves are shown in Figure A6.

Table 5 shows all modifications made at the first position followed immediately by a P-K-R sequence. The X groups consisted of many Pro mimics including cyclopentane, cyclopentane, cyclohexane, phenyl and a benzyl moiety. Smaller groups to mimic Gly and Ser were also incorporated in the X position. Finally, a larger aromatic group, a quinolone, was also added to probe steric hindrances at the X position. All of these

**Table 5.** Peptide inhibitors with 0 CH<sub>2</sub> groups between carbonyl and X position

Compound	Structure	IC <sub>50</sub> (μM) ± SD	Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-2		>100	BM-5		>100
BM-13		>100	BM-17		>100
BM-21		>100	BM-24		>100
BM-26		>100	BM-41		>100

inhibitors also contained 0 CH<sub>2</sub> groups between the carbonyl and X position. Interestingly, none of these inhibitors showed inhibition toward NTMT1 at 100 μM. To explore the importance of the carbonyl at the first position, it was removed in compounds BM-26 and BM-27, and neither inhibitor showed any inhibitory activity at 100 μM. This lack of inhibition may be an indicator of the significance of the hydrogen bond with the carbonyl in that position or may be due to the distance of the X group.

On the other hand, introducing methylene or ethylene between the carbonyl and X position led to a significant increase in inhibition (Table 6). While BM-2 showed no inhibition at 100 μM, the addition of one more CH<sub>2</sub> group in BM-9 led to an IC<sub>50</sub> value of 33 μM. This was also observed in BM-41 and BM-42 which increased in inhibition after the addition of a methylene group. The increase in inhibition from an extra CH<sub>2</sub> group

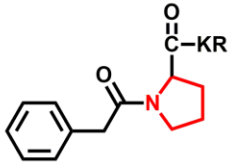
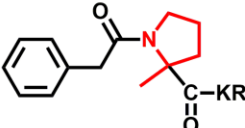


**Table 6.** Peptide inhibitors with 1-2 CH<sub>2</sub> groups between carbonyl and X position

Compound	Structure	IC <sub>50</sub> (μM) ± SD	Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-9		33 ± 1.5	BM-32		4.5 ± 0.47
BM-28		14 ± 2.5	BM-36		12 ± 1.4
BM-30		1.0 ± 0.13	BM-38		>100
BM--34		4.3 ± 0.28			
BM-39		12 ± 2.9			
BM-40		~80			
BM-42		~80			

may indicate the need for more flexibility or distance at the first position. Alternatively, an ethylene group led to a decrease in inhibition compared to the methylene analogs. BM-32 had a 4.5-fold loss in inhibition after the incorporation of two CH<sub>2</sub> groups compared to BM-30. Similarly, BM-36 had a 3-fold loss in inhibition compared to BM-34. Therefore, these results suggest that the optimum inhibitor for the NTMT1 binding site should have a methylene group. Inhibitors which contain no CH<sub>2</sub> groups may not have the distance or flexibility to attain high inhibitory activity while two CH<sub>2</sub> groups may cause steric hindrances.

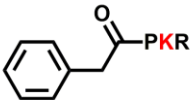
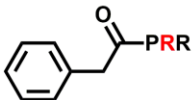
**Table 7.** Structure-Activity-Relationship of modifications at 2<sup>nd</sup> position

Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-9		33 ± 1.5
BM-43		>100

Additionally, a variety of substituents were incorporated at the para position on the X position, including a hydroxy group, an amine, a fluoro group and a methoxy group. The inhibitors which contained a hydrogen donor at this para position, had the greatest inhibitory effects. An 80-fold decrease in inhibition was found in BM-40, which had a methoxy group, compared to BM-30, which had a hydroxy group. This supports the necessity of a hydrogen donor at this first position.

Only one modification was made at the second position of the peptide inhibitors and is shown in Table 7. Adding an alpha-methyl-Pro at the second position forces the P2 to exist in a cis configuration,<sup>199</sup> whereas Pro more only exists in the trans configuration in our co-crystal structures.<sup>24</sup> This modification led to a complete loss of

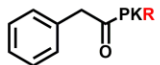
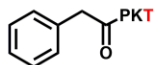
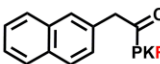
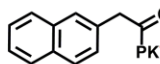
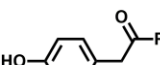
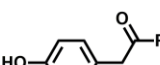
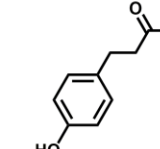
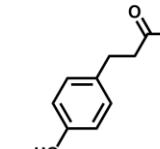
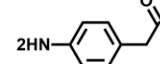
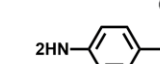
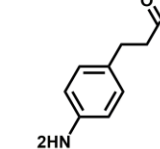
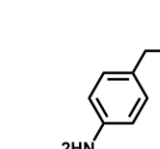
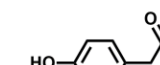
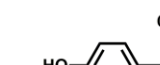

**Table 8.** Structure-Activity-Relationship of modifications at 3<sup>rd</sup> position

Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-9		33 ± 1.5
BM-10		>100

inhibition for BM-43. These results substantiate previous findings that a Pro at the second position prefers the trans configuration; however, the additional methyl group may also be contributed to the loss in inhibition.

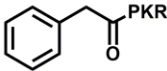
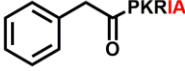
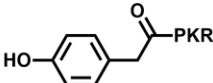
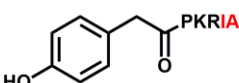
Many inhibitors were synthesized that had the sequence X-P-K-R and X-P-R-R given that Lys and Arg exist in the third position of NTMT1 substrates. All inhibitors that ended in the P-R-R sequence showed less inhibitory activity in the double dose screenings compared to the P-K-R sequence. None of the inhibitors with an Arg at the third position were evaluated in an IC<sub>50</sub> assay due to low initial screenings. Table 8 indicates the preference for Lys over Arg at the third position.

**Table 9.** Peptide inhibitors with modifications at 4<sup>th</sup> position

Compound	Structure	IC <sub>50</sub> (μM) ± SD	Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-9		33 ± 1.5	BM-8		>100
BM-28		14 ± 2.5	BM-29		>100
BM-30		1.0 ± 0.13	BM-31		4.8 ± 0.31
BM-32		4.5 ± 0.47	BM-33		>100
BM-34		4.3 ± 0.28	BM-35		>100
BM-36		12 ± 1.4	BM-37		48 ± 2.0
BM-44		~90	BM-45		~35
BM-46		4.2 ± 1.2			

NTMT1 substrates have Arg and Thr at the fourth position; therefore, inhibitors were synthesized with both amino acids at this position and are shown in Table 9. Every inhibitor that ended in an Arg residue had greater inhibitory activity compared to its Thr analog. The inhibitors ending in Arg had 4- to 25-fold greater inhibition compared to their Thr counterparts. These results indicate the significance of the Arg residue at the fourth position. Eliminating the fourth amino acid entirely and converting the inhibitor to a trimer also significantly diminished inhibition, which is shown by BM-44 compared to BM-30. Two modifications were also made at the fourth position in attempts to increase permeability or stability. Adding trimethylated Lys at the fourth position (BM-45) was performed to increase permeability by adding three additional methyl groups, yet a significant loss in inhibitory activity occurred. Converting L-Arg to the non-natural amino acid D-Arg (BM-46) was carried out to increase stability against peptidases; however, resulted in a four-fold loss in inhibition. Although there was a decrease in inhibition due to this modification, there may be an increase in stability that is worth the loss in inhibition. Lastly, the hexamer peptide compounds showed greater inhibition than their tetramer analogs. BM-11 had a 12-fold increase in inhibition compared to BM-9, while BM-47 has a 3-fold increase compared to BM-30 (Table 10). These results suggest that increasing peptide inhibitor length increases inhibitory activity due to additional binding interactions.

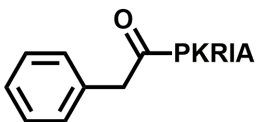
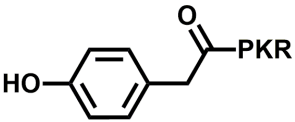
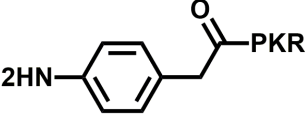
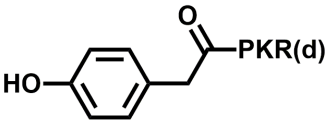
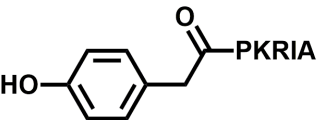
**Table 10.** Tetramer and hexamer peptide analogs

Compound	Structure	IC <sub>50</sub> (μM) ± SD	Compound	Structure	IC <sub>50</sub> (μM) ± SD
BM-9		33 ± 1.5	BM-11		2.7 ± 0.7
BM-30		1.0 ± 0.13	BM-47		0.32 ± 0.06

### 2.3.6 Selectivity Studies

Given that there are many methyltransferase enzymes, selectivity among these inhibitors is critical. For example, G9a is a methyltransferase that transfers a methyl to Lys side chains while PRMT1 is a methyltransferase that transfers methyl groups to Arg side chains. Both enzymes also share the conserved SAM binding site with NTMT1. Therefore, selectivity of these compounds for NTMT1 is vital. The top five inhibitors with IC<sub>50</sub> values lower than 5 μM were tested against G9a and PRMT1 to ensure selectivity. All peptide inhibitors except BM-46 showed >70 μM IC<sub>50</sub> values against G9a and all the inhibitors showed greater than >100 μM IC<sub>50</sub>'s against PRMT1. The results indicate that

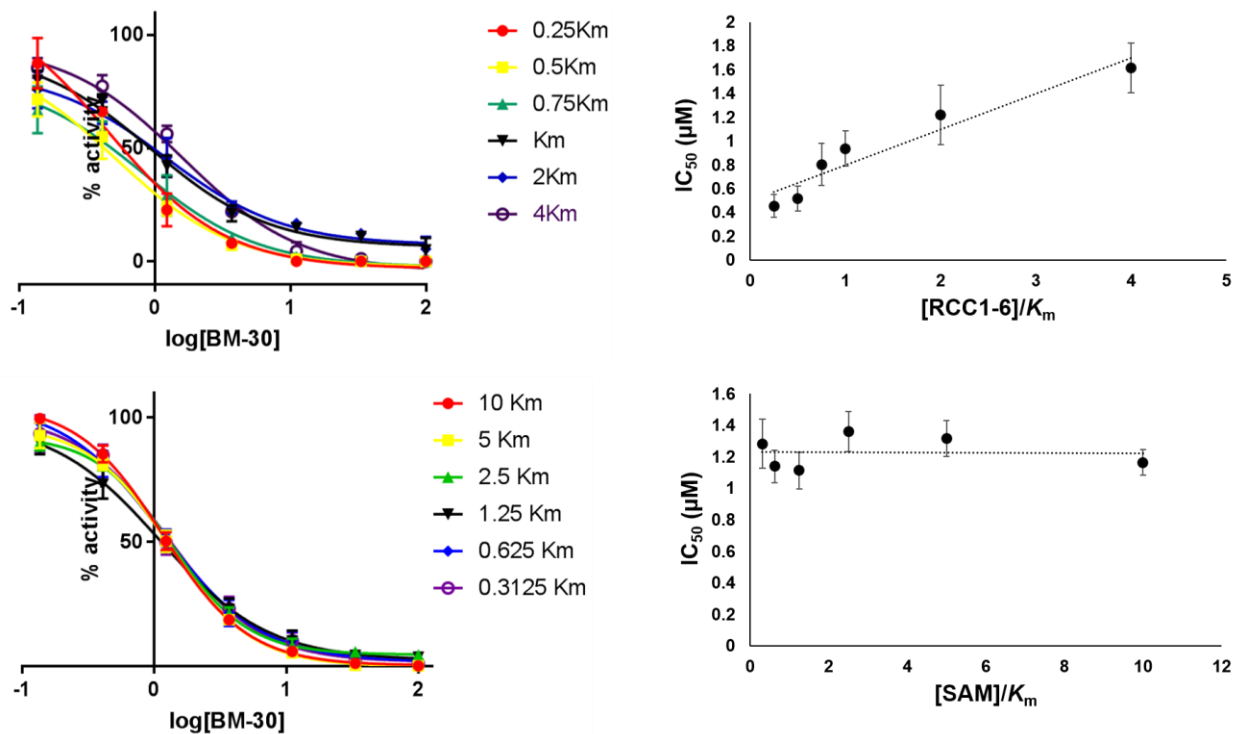
**Table 11.** Summary of selectivity data for top peptide inhibitors

Name	Structure	IC <sub>50</sub> (μM) ± SD		
		NTMT1	G9a	PRMT1
BM-11		2.7 ± 0.7	~70	>100
BM-30		1.0 ± 0.13	>100	>100
BM-34		4.3 ± 0.28	>100	>100
BM-46		4.2 ± 1.2	13 ± 4.1	>100
BM-47		0.32 ± 0.06	>100	>100

the inhibitors have 3- to 300-fold selectivity for NTMT1 over these other two methyltransferases. The selectivity data is summarized in Table 11.

### 2.3.7 Inhibition Mechanism

An inhibition mechanism study of one of the top inhibitors, BM-30, was performed to verify which binding site the peptide inhibitor competes. Figure 14 shows the inhibitor is acting through competitive inhibition with respect to the peptide substrate. This is demonstrated by the linear dependence of the  $IC_{50}$  value relative to the increase in peptide substrate concentration. Alternatively, BM-30 was noncompetitive with respect to SAM, indicated by the flat line that minimally fluctuates based on SAM concentration.



**Figure 14.** BM-30 is a SAM-noncompetitive, peptide-competitive inhibitor of NTMT1. BM-30 is competitive with the RCC1-6 peptide substrate, as the  $IC_{50}$  increases linearly with with compound concentration. BM-30 does not compete with cofactor SAM, as the  $IC_{50}$  was not affected by the compound.

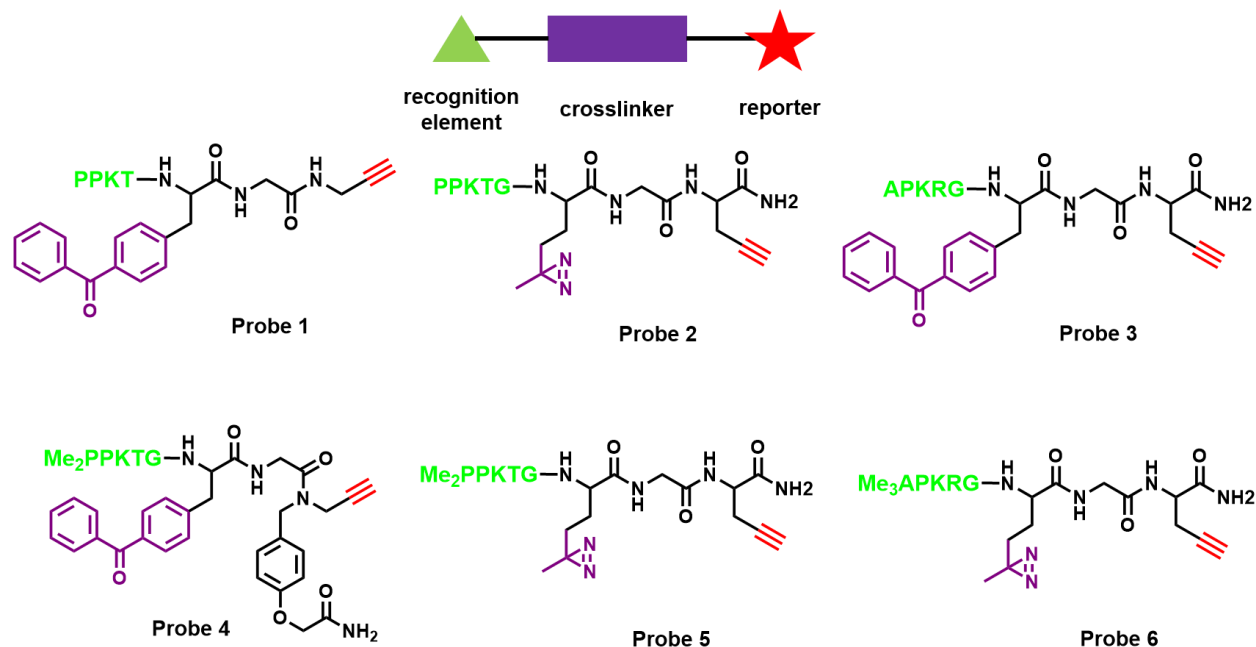
## 2.2 Photoaffinity Probes

### 2.2.1 Design

The goal is to identify the reader for  $\alpha$ -N-terminal methylation to provide insight into the NTMT1-regulated downstream pathway. Until now, there is no known interacting partner that has been identified to recognize  $\alpha$ -N-terminal methylation, except the NTMT enzymes that install this methylation. Photoaffinity labeling has been used as an attractive approach to identify binding partners, delineate drug-protein binding interactions, and isolate cellular targets.<sup>204–207</sup> Hence, we sought to apply a photoaffinity approach to catch novel readers for  $\alpha$ -N-terminal methylation. A typical photoaffinity probe contains three main components: a recognition element, a crosslinker, and a clickable handle (Figure 15). The recognition element conveys selectivity, the photocrosslinker covalently tethers the probe to the targets, and the clickable handle allows for derivations to link a fluorescent reporter or biotinylated tag.

We initiated our efforts by optimizing our photoaffinity probes using NTMT1 which is the only known interacting partner as the model system. Probes **1-3** contain an unmethylated N-terminus and have the capability of being enzymatically recognized by the NTMTs. Probes **4-6** have a methylated N-terminus and mimic the product of the NTMT family and can be recognized by NTMT downstream reader proteins. Additionally, NTMT still recognizes its methylated product; therefore, probes **4-6** can still be optimized using NTMT1 as a model. To achieve specific photolabeling, we aim to position the crosslinker close enough to the recognition motif while not interfering with the recognition. Based on our co-crystal structures of NTMT1 in complex with substrates, the first four

residues of the peptide substrate contribute significantly to the binding while the fifth and sixth residues only form backbone hydrogen bonds with NTMT1 and are solvent-exposed.<sup>24</sup> Hence, we decided to only incorporate the first four residues from selected protein substrates as the specific recognition element and attached the crosslinker at either the fifth or sixth position on our designed probes **1-6** (Figure 15). Among all tested peptide substrates, substrate peptides starting with P-P-K have the highest binding affinity for NTMT1 (1.5 nM) and can be efficiently methylated to the fully methylated state by recombinant NTMT1.<sup>24, 191,203</sup> Such substrate proteins include tumor suppressor RB1, mouse RCC1, ribosomal subunits S25 and L12, and histone H2B; therefore, we initiated our efforts by incorporating a P-P-K-T peptide that is derived from RB1 protein into probes **1, 2, 4, and 5** as the recognition element. In addition, we incorporated A-P-K-R that is derived from oncoprotein SET as an alternative recognition element for probe **3 and 6**



**Figure 15.** Structure of photoaffinity probes **1-6**



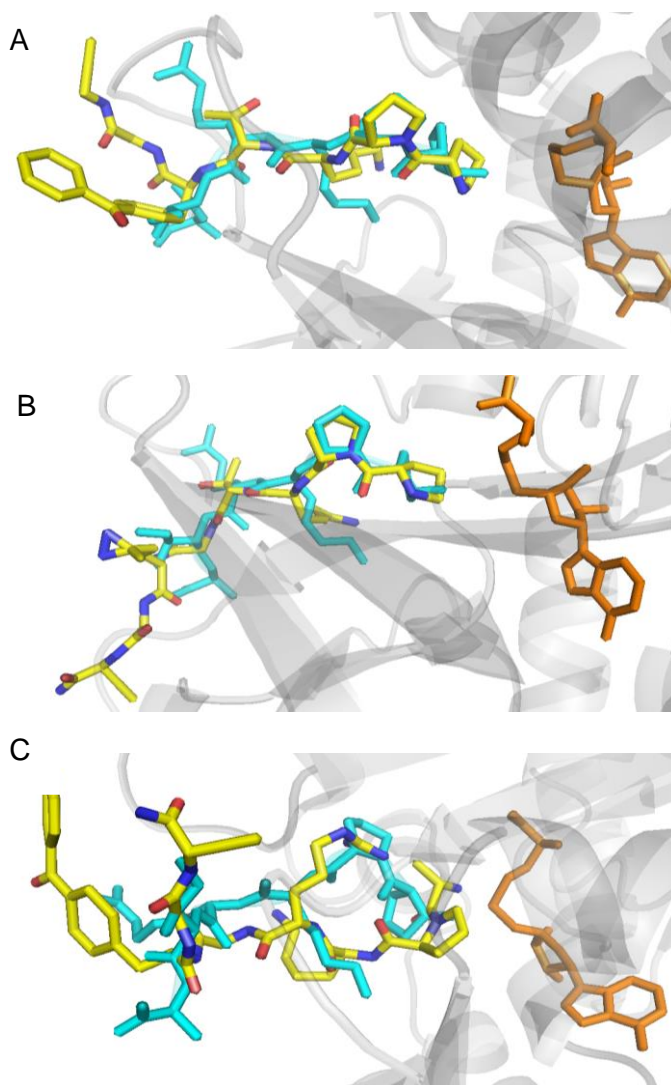
since validated substrates like oncoprotein SET, DDB2, and PARP3 contain the sequence APK at their N-termini.

We initially selected a benzophenone as our crosslinker considering its commercial availability, high stability and its preferential reaction with unreactive C-H bonds.<sup>208</sup> However, the benzophenone bulky size may also cause low labeling efficiency and lead to nonspecific labeling. Therefore, we also synthesized photo-Met to investigate the effect of a diazirine crosslinker,<sup>209</sup> which is smaller in size. The diazirine has high labeling efficiency and consequently increases the sensitivity of the probe, although the reactive diazirine intermediate could also lead to nonspecific labeling.<sup>210-212</sup> Both crosslinkers were placed at either the fifth or sixth amino acid position of the probe, which is after the recognition portion. Again, this design was to minimize interference with the affinity of the recognition element while still positioned in close enough proximity that it would specifically bind to the target. One Gly linker was introduced right after the crosslinker (Probe **1**) or one on either side of the crosslinker (Probes **2-6**) to allow conformational flexibility and possibly prevent possible steric hindrance from interactions.<sup>213</sup> A propargyl group was used as a clickable handle and added to the C-terminal of the peptide through a Backbone Amide Linker (BAL) or propargylglycine.

### **2.2.2 Docking Studies**

To support our design and further validate that the crosslinkers did not affect binding within the active site, docking studies of the probes was performed using the NTMT1 co-crystal structure (PDB id: 5E2B). The NTMT1 binding site was defined by a

sphere of 6.0 Å where the substrate MePPKRIA was extracted and probes **1-3** were docked into that site. The compounds for docking were prepared using SYBYL X2.1 and the energy was minimized using the external Tripos force field. The ChemPLP score from the Gold53 docking program was considered to determine the most likely docking conformations and top inhibitors. Probe **1** overlaid well with the peptide substrate (MePPKRIA) in the binding pocket of NTMT1. The benzophenone crosslinker and the alkyne portion pointed away from the surface of the protein. Probe **2** bound in a similar

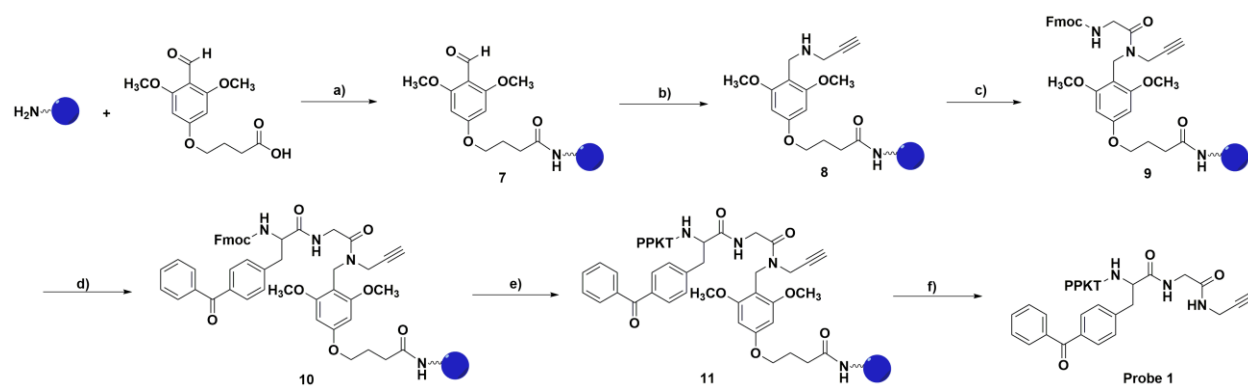


**Figure 16.** Superimposed structures of probe **1-3** (yellow) with substrate peptide MePPKRIA (cyan) in the complex of NTMT1 (grey). (PDB: 5E2B) SAH is shown in orange (A) Probe **1**; (B) Probe **2**; (C) Probe **3**.

manner as probe **1** and the diazirine crosslinker still pointed to the surface of the protein. Surprisingly, probe **3** displayed slightly different binding mode from the MePPKRIA peptide, with the N-terminus in closer proximity to SAH. Nevertheless, the crosslinker moieties of all three probes imposed minimal interference on binding to NTMT1 (Figure 16).

## 2.2.3 Synthesis

All probes were synthesized on solid phase via standard Fmoc peptide coupling procedure. Briefly, the synthesis of photoaffinity Probe **1** is illustrated in Scheme 1. The commercially available BAL handle was loaded onto Rink amide resin first to yield **7**.<sup>214</sup> Compound **7** was converted to the alkyne, **8**, after reductive amination with propargylamine and sodium cyanoborohydride.<sup>214</sup> Next, a Gly linker was manually coupled to yield **9** followed immediately with a manual coupling of Fmoc-4-benzoyl-L-phenylalanine to yield **10**. The N-terminal amino acids of PPKT



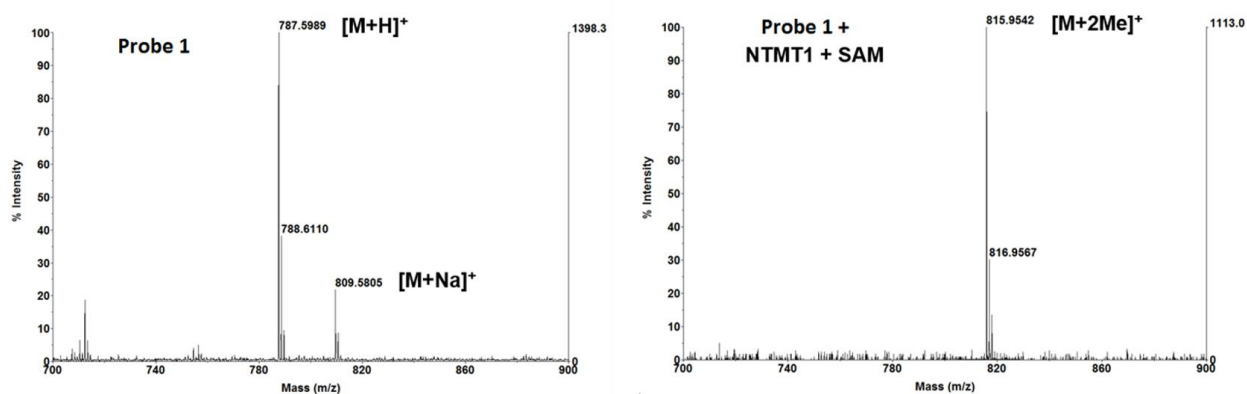
**Scheme 1.** Synthesis of photoaffinity probe **1** a) piperidine, HOBT, DMF, rink amide resin, 2x10 min; HOBT, HBTU, DIPEA, NMP, 12 hr; Ac<sub>2</sub>O, DCM, 2x5 min; b) propargylamine, NaBH<sub>3</sub>CN, AcOH, MeOH, 2x12 hr; c) Fmoc-Gly-OH, DIC, NMP, DCM, 2x90 min; Ac<sub>2</sub>O, DCM, 2x5 min; d) piperidine, HOBT, DMF, 2x10 min; Fmoc-4-Bpa-L-phenylalanine, HOBT, DIC, DMF, 12 hr; e) automated iterative synthesis,  $\mu$ W, Gly, Thr, Lys, Pro, Pro; f) trifluoroacetic acid, 2,2'-(ethylenedioxy)diethanethiol, Triisopropylsilane, H<sub>2</sub>O, 4 hr.

were added via a CEM Liberty microwave automatic peptide synthesizer to produce **11** and followed by cleavage from resin to yield probe **1**. Cleavage from resin using trifluoroacetic acid (TFA) also cleaves the BAL linker from the probe. The probe was verified by mass spectrometry with a detected mass of 787.5989 (predicted mass 787.4137). The synthesis of photoaffinity probe **4** is very similar as the synthesis of probe **1**, except the BAL linker is not cleaved after the addition of TFA. Additionally, probe **4** was dimethylated using methyl iodide and potassium carbonate.<sup>201</sup> Probe **4** was cleaved and was verified by mass spectrometry with a detected mass of 1035.5136 (predicted mass 1035.5304).

The syntheses of probes **2**, **3**, **5**, and **6** were completed in a similar fashion except Fmoc-propargylglycine was purchased from ChemImpex and manually coupled onto the Rink amide resin. This step eliminated the incorporation of the Bal linker and the reductive amination reaction. After successful coupling of the Fmoc-propargylglycine, manual coupling of Fmoc-glycine and Fmoc-benzophenone or Fmoc-Photo-Met followed as described above. The Fmoc-Photo-Met was synthesized as described in Yang, *et al.*<sup>209</sup> The remaining amino acids were coupled using the automatic peptide synthesizer. Probes **5** and **6** were then dimethylated or trimethylated and were also verified by mass spectrometry with a detected mass of 817.4768 (predicted mass 817.4685) and 860.6093 (predicted mass 860.5219), respectively. Probes **2** and **3** were also verified by mass spectrometry with a detected mass of 789.4373 (predicted mass 789.4366) and 930.4953 (predicted mass 930.4945), respectively. All six photoaffinity probes were purified by reverse-phase HPLC (Waters) and characterized by mass spectrometry (Figure A7-A8).

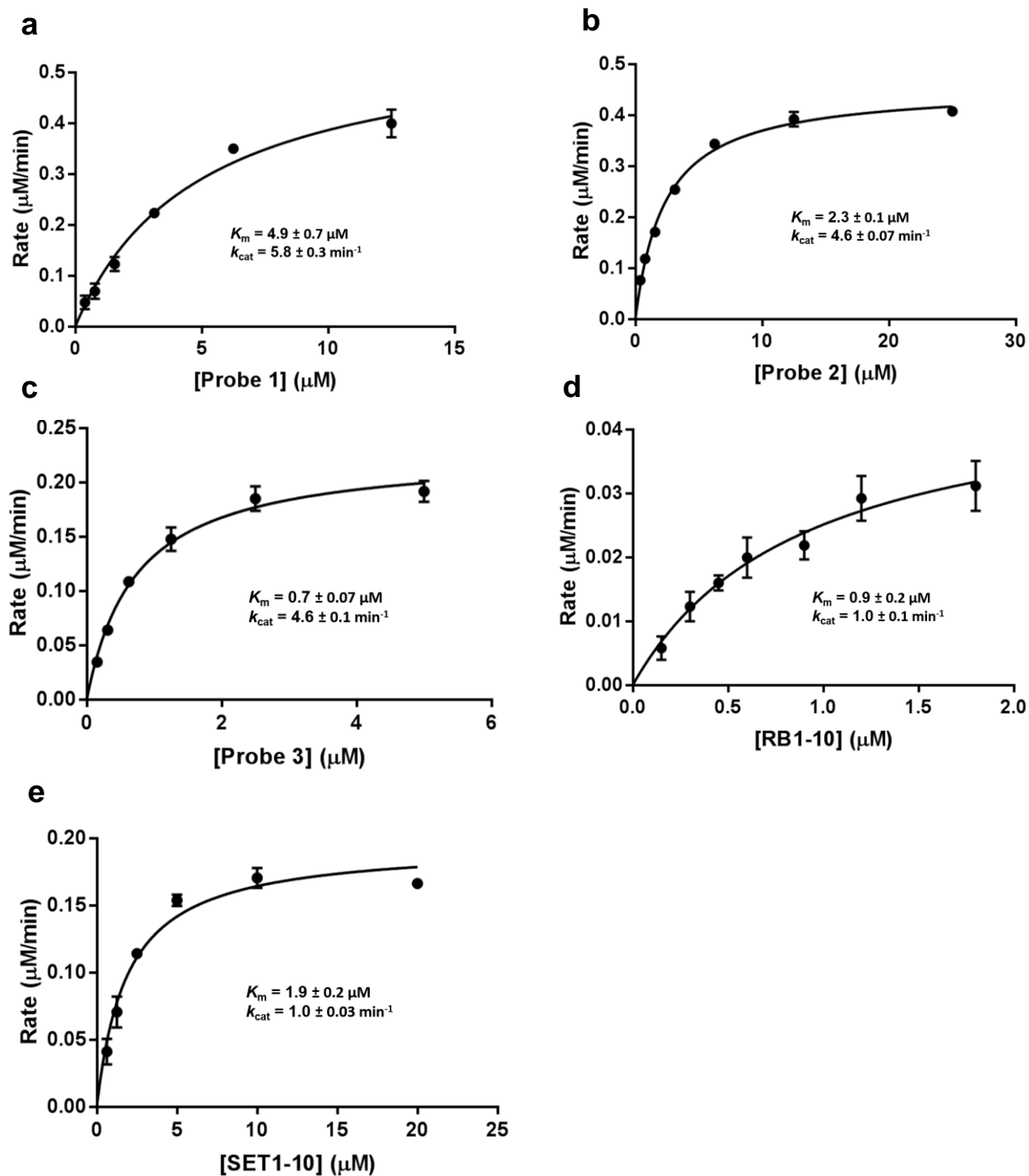
## 2.2.4 Recognition Studies

To validate the specific recognition of the probes, we first examined if our unmethylated probes could be recognized and methylated by NTMT1 through a MS-based study. Briefly, photoaffinity probe **1** (5  $\mu\text{M}$ ) was incubated with NTMT1 (0.2  $\mu\text{M}$ ) and SAM (50  $\mu\text{M}$ ) as reported previously.<sup>192,203</sup> The reaction was monitored at various time points via MALDI-MS. As shown in Figure 17, the photoaffinity probe was fully methylated within five minutes in the presence of SAM as only dimethylated peak ( $[\text{M}+2\text{Me}]^+ = 815.9542$ ) was detected. This phenomenon indicates minimal interference of the benzophenone with NTMT1 recognition and catalysis. This experiment was also repeated with probe **3**; however, after 15 min, there was a mixture of un-, mono-, and dimethylated product (Figure A9).



**Figure 17.** MALDI-MS of probe **1** (theory  $m/z$ : 787.4143, detected  $m/z$ : 787.5989) (left). MALDI-MS confirming photoaffinity probe **1** is fully dimethylated by NTMT1 in the presence of S-adenosyl methionine (SAM) after 5 minutes (theory  $m/z$ : 815.4456, detected  $m/z$ : 815.9542) (right).

In addition, we characterized the steady-state kinetic parameters of probes **1-3** for methylation using an SAH hydrolase-coupled fluorescence assay, previously described by Richardson, *et al.*<sup>192</sup> The  $K_m$  values of probes **1**, **2** and **3** were found to be  $4.9 \pm 0.7$   $\mu\text{M}$ ,  $2.3 \pm 0.1$   $\mu\text{M}$  and  $0.7 \pm 0.07$   $\mu\text{M}$ , respectively (Figure 18a-c, Table A2). To understand



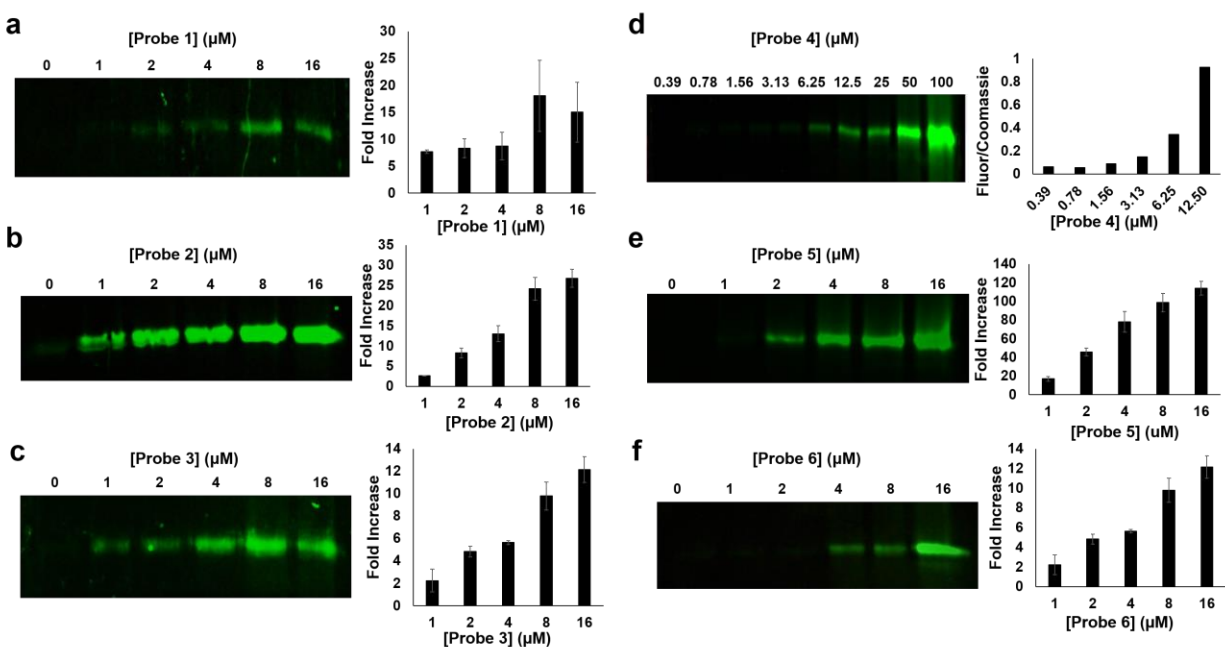
**Figure 18.** Steady-state kinetics of His-NTMT1 activity as detected by fluorescence assay. Varied concentration of (a) probe 1, (b) 2, (c) 3, (d) RB1-10, and (e) SET1-10 peptide with SAM at 100  $\mu\text{M}$ .

how the crosslinker of the probes affects NTMT1 methylation, we also determined the  $K_m$  values of the RB1-10 (PPKTPRKTA A) and SET1-10 (APKRQSPLPP) peptides to be 0.9

$\pm 0.2 \mu\text{M}$  and  $1.9 \pm 0.2 \mu\text{M}$ , respectively (Figure 18d-e). Hence, the comparable  $K_m$  values of our probes and peptide substrates validate our design strategy.

### 2.2.5 Photoaffinity Labeling

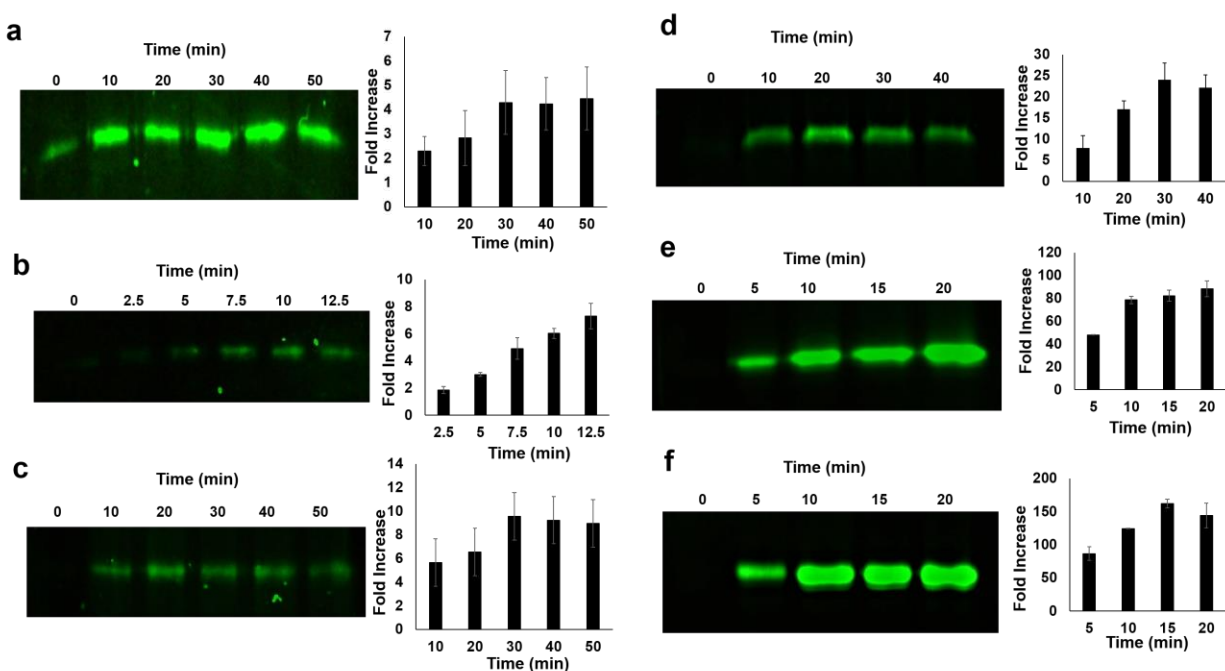
To investigate the effect of probe concentration on labeling efficiency, NTMT1 (4  $\mu\text{M}$ ) was directly incubated with varying concentrations of the probes (0  $\mu\text{M}$  to 16  $\mu\text{M}$ ) on ice for 10 min and followed by irradiation for 30 min at 312 nm.<sup>215</sup> Tetramethylrhodamine (TAMRA) azide and the “click” reagents were then added and incubated in the dark for one hour. Then the samples were quenched with SDS loading dye and analyzed after electrophoresis. In order to remove the excess fluorophore and eliminate background noise from unreacted probes, the SDS-PAGE gels were washed with destaining buffer<sup>216</sup> before the samples were quantified through fluorescence imaging analysis (Figure 19).



**Figure 19.** Concentration-dependence photolabeling of NTMT1 by (a) probe 1, (b) probe 2, (c) probe 3, (d) probe 4, (e) probe 5 and (f) probe 6.

The results showed that all six probes display a concentration-dependent manner given that the amount of fluorescent labeling increases as the concentration of the probes increases. The optimal concentration based on the signal-to-noise ratio and background fluorescence was determined to be 8  $\mu\text{M}$  for the probes which contain a benzophenone (probes 1, 3 and 4) and 4  $\mu\text{M}$  for probes which have a diazirine crosslinker (probes 2, 5 and 6). These concentrations of photoaffinity probes are used for all following studies.

A time-dependent study was performed to determine the optimal UV dosage for efficient photocrosslinking. NTMT1 (4  $\mu\text{M}$ ) was incubated with benzophenone photoaffinity probes 1, 2 and 4 (8  $\mu\text{M}$ ) on ice for 10 min then irradiated for 0, 10, 20, 30, 40 or 50 min at 312 nm.<sup>215</sup> The diazirine crosslinker has been reported to have a much faster crosslinking time; therefore, the time used for probes 2, 5 and 6 was 0, 2.5, 5, 7.5, 10 and 12.5 min or 0, 5, 10, 15 and 20 min.<sup>209</sup> The results indicated that an optimal

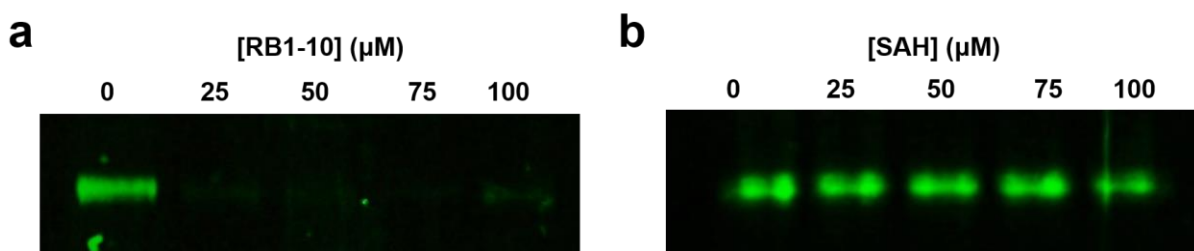


**Figure 20.** Time-dependence photolabeling of NTMT1 by (a) probe 1, (b) probe 2, (c) probe 3, (d) probe 4, (e) probe 5 and (f) probe 6.



labeling was reached at approximately 30 min for the benzophenone probes **1**, **3** and **4** (Figures 20a,c,d) and 10 min for the diazirine probes **2**, **5** and **6** (Figure 20b,e,f).

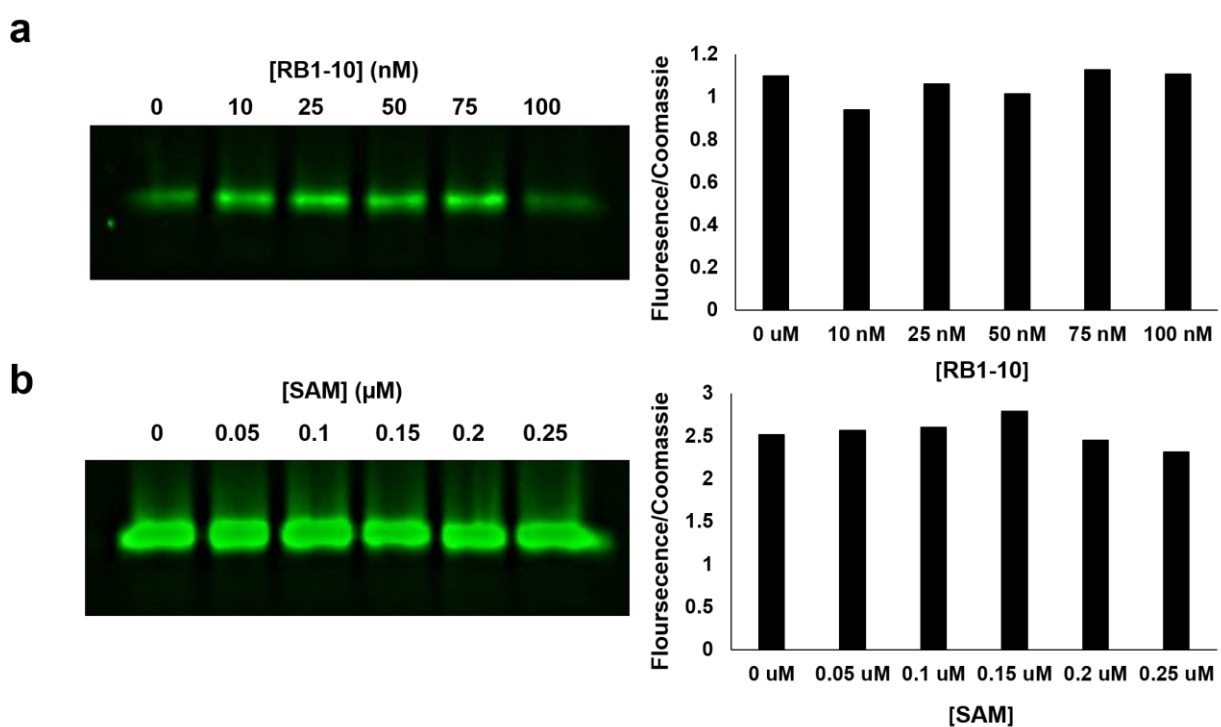
To investigate if the labeling is specifically directed by recognition, we performed NTMT1 labeling in the presence of an RB1-10 substrate peptide, which served as a competitor to bind to NTMT1. NTMT1 (4  $\mu\text{M}$ ) was incubated with either competitor peptide RB1-10 (0, 25, 50, 75 and 100  $\mu\text{M}$ ) or S-adenosylhomocysteine (SAH) (0, 25, 50, 75 and 100  $\mu\text{M}$ ) on ice for 10 min to allow for any possible interaction. Next, photoaffinity probe **1** (8  $\mu\text{M}$ ) was added and incubated on ice for an additional 10 min to allow the probe to compete with any of the possible interactions with NTMT1. Compared to the control, total fluorescent labeling for probe **1** decreased over three-fold in the presence of 25  $\mu\text{M}$  RB1-10, but did not significantly change in the presence of SAH (Figure 21). Since SAH binds at the SAM binding site of NTMT1, our results validate that our designed photoaffinity probes bind in the peptide binding pocket, revealing that the labeling is driven by specific recognition.



**Figure 21.** Competition studies. (a) Cross-linking reactions of probe **1** with recombinant NTMT1 in the presence of RB1-10 peptide at 0, 25, 50, 75, and 100  $\mu\text{M}$ . (b) Cross-linking reactions of probe **1** with recombinant NTMT1 in the presence of SAH at 0, 25, 50, 75, and 100  $\mu\text{M}$ .

Experiments were also carried out with RB1-10 and SAM at physiological concentrations to evaluate if the photoaffinity probes still effectively label NTMT1. Physiological concentrations of RB1 protein and SAM is reported to be 45 nM and 0.15

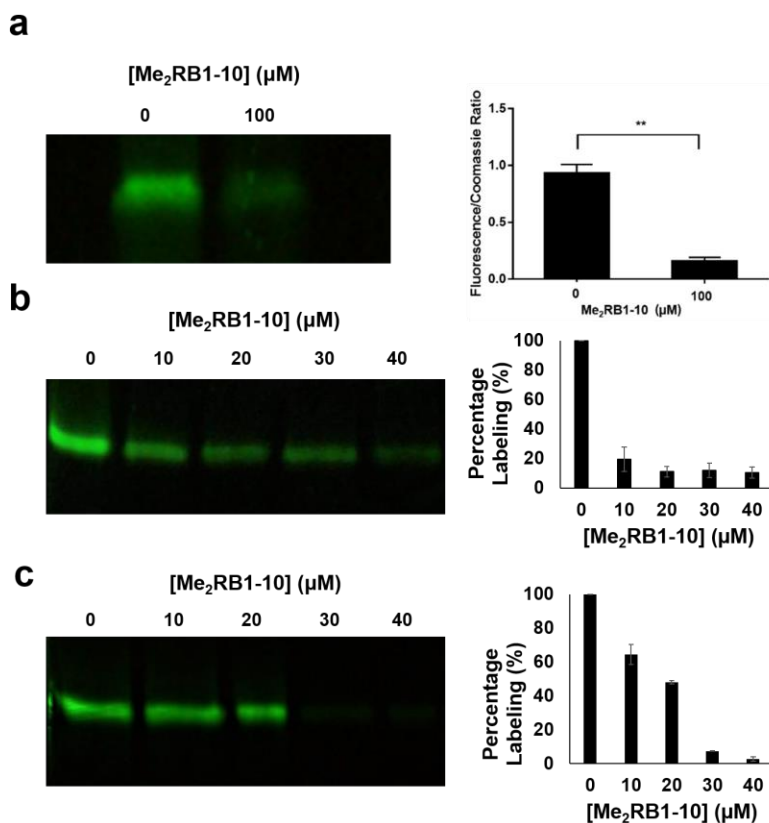
$\mu\text{M}$ , respectively. Therefore, NTMT1 ( $4\ \mu\text{M}$ ) was incubated with RB1-10 (0, 10, 25, 50, 75 and 100 nM) or cofactor SAM (0, 0.05, 0.1, 0.15, 0.2, and 0.25  $\mu\text{M}$ ) on ice for 10 min to allow for any competition. Next, photoaffinity probe **2** ( $4\ \mu\text{M}$ ) was added and incubated on ice for an additional 10 min to allow the probe to label NTMT1. Compared to the control, total fluorescent labeling for probe **2** did not decrease in the presence of RB1 or SAM (Figure 22). These results validate that these photoaffinity probes would be effective in labeling NTMTs against physiological concentrations of competitors.



**Figure 22.** NTMT1 labeling with probe **2** at physiological concentrations of competitors (a) Probe **2** with recombinant NTMT1 in the presence of RB1-10 peptide at 0, 10, 25, 50, 75, and 100 nM. (b) Probe **2** with recombinant NTMT1 in the presence of SAM at 0, 0.05, 0.1, 0.15, 0.2 and 0.25  $\mu\text{M}$ .

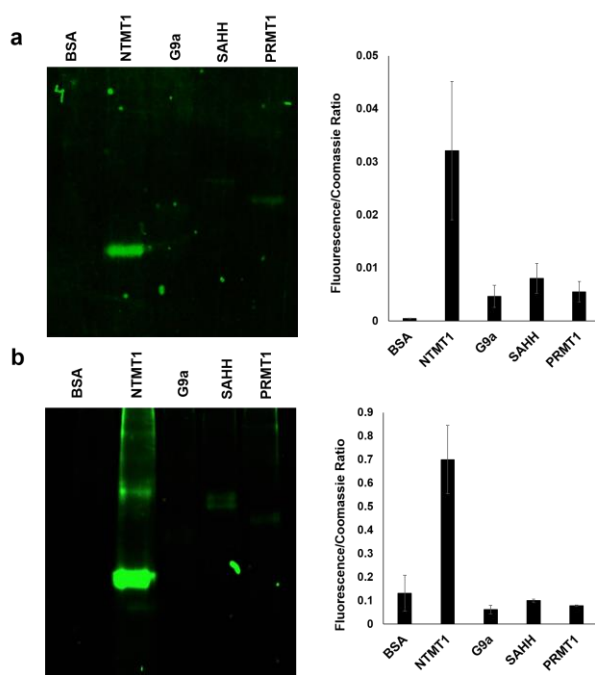
Competition experiments were also performed with the probes **4-6** and a methylated RB1-10 competitor. NTMT1 ( $4\ \mu\text{M}$ ) was incubated with either competitor peptide Me<sub>2</sub>RB1-10 (0, 10, 20, 30, 40 or 100  $\mu\text{M}$ ) on ice for 10 min to allow for any possible interaction. Next, the photoaffinity probes (8 or 4  $\mu\text{M}$ ) were added and incubated on ice

for an additional 10 min to allow the probe to compete with any of the possible interactions with NTMT1. Compared to the control, total fluorescent labeling for probe **4** decreased over five-fold in the presence of 100  $\mu\text{M}$  Me<sub>2</sub>RB1-10 (Figure 23a). Total percentage of labeling of NTMT1 by photoaffinity probe **5** also decreased to 20% after introducing 10  $\mu\text{M}$  of the Me<sub>2</sub>RB1-10 competitor (Figure 23b). Finally, total percentage of labeling by probe **6** only decreased to 60% after introducing 10  $\mu\text{M}$  of Me<sub>2</sub>RB1-10 (Figure 23c). These results validate that our reader photoaffinity probes bind in the peptide binding pocket, revealing that the labeling is driven by specific recognition.



**Figure 23.** Competition studies. (a) Cross-linking reactions of probe **4** with recombinant NTMT1 in the presence of Me<sub>2</sub>RB1-10 peptide at 0 and 100  $\mu\text{M}$ . (b) Cross-linking reactions of probe **5** with recombinant NTMT1 in the presence of Me<sub>2</sub>RB1-10 peptide at 0, 10, 2, 30 and 40  $\mu\text{M}$ . (c) Cross-linking reactions of probe **6** with recombinant NTMT1 in the presence of Me<sub>2</sub>RB1-10 peptide at 0, 10, 2, 30 and 40  $\mu\text{M}$ .

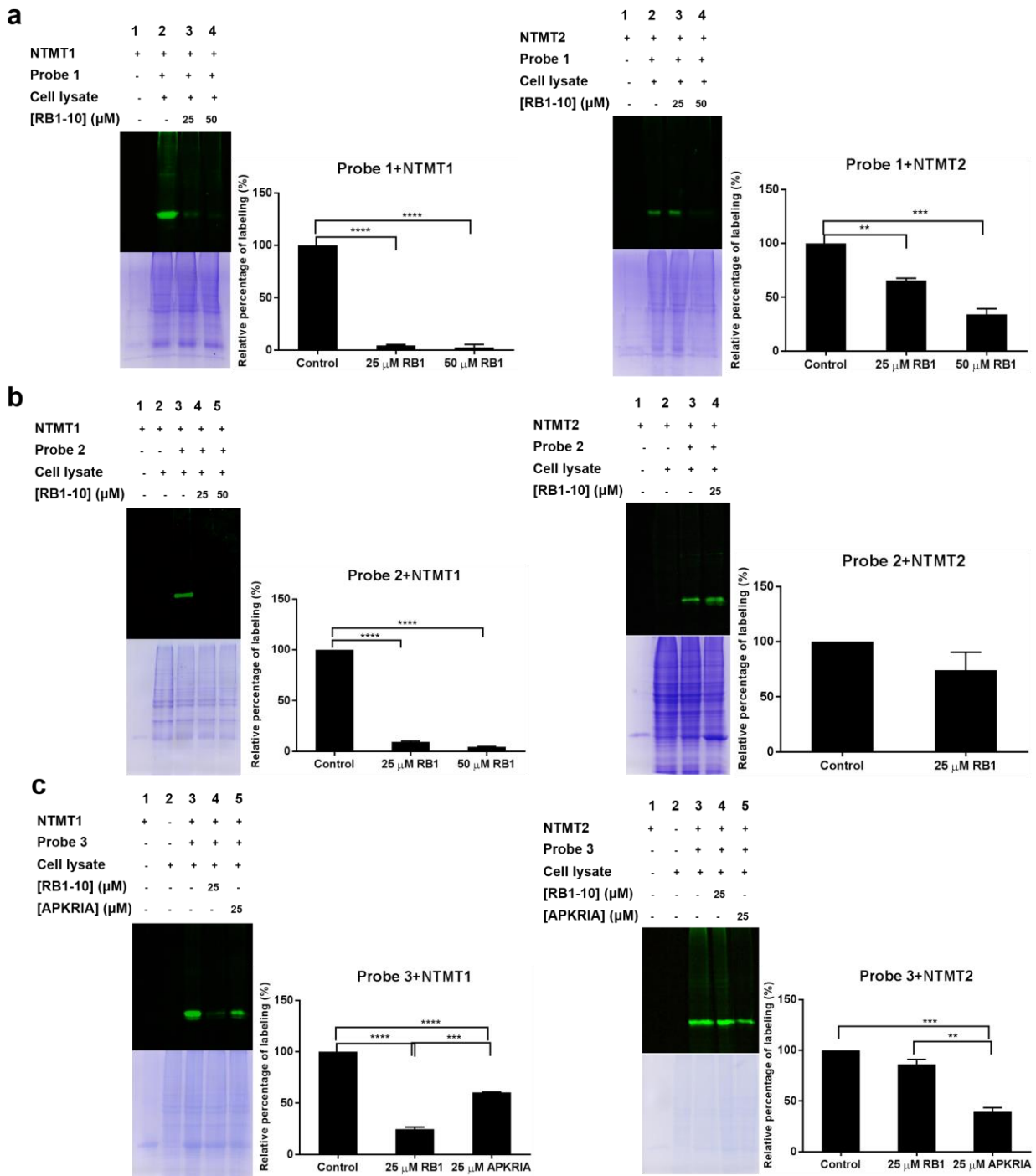
Selectivity of the probes is critical for their applications in labeling specific interactions. We assembled a pilot five-member panel of proteins including NTMT1, bovine serum albumin (BSA), SAH hydrolase (SAHH), and two other protein methyltransferases including euchromatic histone-Lys N-methyltransferase 2 (G9a) and protein Arg methyltransferase 1 (PRMT1). We performed the photoaffinity labeling experiment for each member with our probes. Figure 24 show that probe 1 and probe 3 efficiently labeled NTMT1 more than 4- and 5-fold, respectively, over the other tested proteins.



**Figure 24.** Selectivity studies. (a) Cross-linking reactions of probe 1 for a panel of five recombinant enzymes. (b) Cross-linking reactions of probe 3 for a panel of five recombinant enzymes.

Although the photoaffinity probes inferred selectivity among a panel of methyltransferases, we wanted to ensure the photoaffinity probes were selective in a complexed context like the cell lysate. Here, we incubated probes 1-3 (8  $\mu$ M) with HeLa cell lysates (1 mg/mL) that were spiked in with either NTMT1 or NTMT2 (4  $\mu$ M) in the

presence of a competitor peptide at various concentrations (0-50  $\mu$ M). Within the whole cell lysate, the photoaffinity probes specifically labeled NTMT1 and NTMT2 and did not show labeling for any other proteins. For probe **1**, the total fluorescent labeling for NTMT1 decreased to 4% and 2% in the presence of RB1-10 at 25  $\mu$ M and 50  $\mu$ M, respectively (Figure 25a left). However, probe **1** labeling of NTMT2 only decreased to 66% and 34% in the presence of 25  $\mu$ M and 50  $\mu$ M RB1-10, respectively (Figure 25a right). More effective competition of RB1-10 peptide against probe **1** for NTMT1 over NTMT2 infers that RB1-10 binds better to NTMT1 over NTMT2. The same substrate preference was observed in the competition experiments with probe **2**. Total labeling of NTMT1 decreased to 9% when incubated with 25  $\mu$ M RB-10 and 4% when incubated with 50  $\mu$ M RB1-10 (Figure 25b left), while labeling of NTMT2 only decreased to 74% when incubated with 25  $\mu$ M RB-10 (Figure 25b right). Similar experiments were carried out with probe **3** with the addition of incorporating an APKRIA substrate competitor. After incubating 25  $\mu$ M of RB1-10 or APKRIA, photoaffinity labeling of NTMT1 decreased to 25% and 60%, respectively (Figure 25c left). Interestingly, labeling of NTMT2 decreased to 86% and 40%, respectively (Figure 25c right). These results imply that NTMT1 has a substrate preference for PPKT over APKR which is substantiated by the kinetic analysis of these two substrates shown in Figure 18 and Table A2. Furthermore, these results suggest that NTMT2 prefers APKR substrate over PPKT.



**Figure 25.** Photoaffinity labeling of probes 1-3 in HeLa cell lysates. (a) Probe 1 labeling in spiked cell lysates with NTMT1 (left) and NTMT2 (right) in presence of competitor RB-10 peptide at 0, 25, and 50  $\mu$ M. (b) Probe 2 labeling in spiked cell lysates with NTMT1 (left) and NTMT2 (right) in presence of competitor RB-10 peptide at 0, 25, and 50  $\mu$ M (c) Probe 3 labeling in spiked cell lysates with NTMT1 (left) and NTMT2 (right) without or with 25  $\mu$ M RB-10 peptide or 25  $\mu$ M APKRIA peptide.

### 3. Conclusions

The first series of peptide-based inhibitors have been synthesized that target NTMT1. All 47 inhibitors were synthesized via solid-phase and an extensive structure-activity-relationship (SAR) has been developed. The compounds with the greatest inhibitory activity contain a hydrogen donor, like a hydroxy or amino group, at the first amino acid in the para position. The optimum length between the first position and the carbonyl of the second position Pro is one CH<sub>2</sub> group. Pro at the P2 position is preferred to alpha-methyl-Pro given its preference for the trans conformation. The inhibitors with a Lys at the third position and an Arg at the fourth position are preferred for highest inhibition. Among all synthesized tetrapeptidomimetics, our top inhibitor displays an IC<sub>50</sub> of 1.02 μM. Finally, extending the compound length from a tetramer to a hexamer, also led to a 3- to 10-fold increase in inhibition. Thusly, seven inhibitors have IC<sub>50</sub> values of less than 5 μM and our top inhibitor has an IC<sub>50</sub> value of 0.32 μM. Many of our top peptide inhibitors were subjected to a MALDI-MS methylation inhibition assay and the results substantiated the inhibitory values found from the fluorescent-based methyltransferase assay. The inhibitor BM-30, also underwent an inhibition mechanism assay to validate which NTMT1 binding site it is targeting. BM-30 was found to target the peptide substrate binding site and not the SAM binding site. Lastly, the top five inhibitors underwent selectivity studies with Lys methyltransferase, G9a, and Arg methyltransferase, PRMT1.

All five inhibitors exhibited 3- to 300-fold selectivity for NTMT1 over the other methyltransferases.

We have also designed, synthesized and characterized the first series of photoaffinity probes that selectively label protein  $\alpha$ -N-terminal methylation writers. A critical issue in the design of photoaffinity probes is the positioning of the photoreactive group and the selection of the recognition element, since it is imperative that the labeling is regulated by specific recognition. Guided by our co-crystal structures and canonical X-P-K motif, we have focused on incorporating an initial four amino acids; hence, we varied two different recognition elements that were derived from protein substrate RB1 and SET. Our photoaffinity probes have exhibited in a dose- and time-dependent manner to label NTMT1/2. Additionally, labeling has proven to be driven by specific recognition given that the photoaffinity probes can be enzymatically methylated by NTMT1 in a MS-based methylation assay, and competitively inhibited in the presence of substrate peptide, RB1-10. Alternatively, probe labeling did not decrease in the presence of physiological concentrations of RB1 or SAM, validating that the probes can be effective in a cellular context. The designed probes also selectively label NTMT1 among a panel of enzymes including other methyltransferase enzymes, G9a and PRMT1. Different labeling of NTMT1 and NTMT2 by photoaffinity probes **1-3** in the presence of two substrate competitors provides the first information to shed light on possible structural differences or recognition preferences between NTMT1/2, which could aid in the design of selective inhibitors for either NTMT family member. Furthermore, probes **1-3** have demonstrated selectivity for NTMT1/2 in HeLa cell lysates, suggesting immense potential of utilizing photoaffinity probes with a methylated  $\alpha$ -N-terminus in the characterization of the protein



$\alpha$ -N-terminal modification network. We anticipate using this strategy as a guide to design the first probes to map the interacting protein partners of  $\alpha$ -N-terminal methylation.

#### 4. Future Directions

Thus far, the peptidomimetic inhibitors have shown a lot of promise. The next step is to determine stability and permeability of these inhibitors given that peptide-based compounds are susceptible to peptidases and degradation *in vivo*. An LC-MS study can be performed to verify stability of the compounds and a Caco-2 study can be carried out to test permeability. Based on the results of these studies, new inhibitors can be synthesized with modifications to increase stability and permeability. Although eliminating the amide bond between the first and second position resulted in no inhibitory activity, elimination of the amide bond at the other positions should be attempted. Alternatively, the amide bonds can also be methylated to increase stability. Incorporating more non-natural amino acids like D-Arg, D-Lys,  $\beta$ -Arg, or  $\beta$ -Lys would also make these inhibitors less prone to degradation. In order to increase permeability, lipophilic groups like can be added to side chains of polar residues. Finally, the inhibitors need to be tested in cell-based and animal studies and inhibition of protein  $\alpha$ -N-terminal methylation needs to be evaluated at a cellular level.

The photoaffinity probes have been optimized for an NTMT1 model and have demonstrated competitive labeling for NTMT1. Probes **1-3** have exhibited selective labeling of spiked-in NTMT1 in a complex cell mixture. However, the next step is to successfully pull-down NTMTs from a complex cell mixture without spiking-in the NTMTs.

Identifying the NTMT writer proteins within the cell is critical to verify that the photolabeling technique is effective at a cellular level. After successful optimization of pulling-down the NTMT family using probes **1-3**, the methylated photoaffinity probes (probes **4-6**) can be utilized to identify novel downstream proteins. Therefore, optimization of this approach in a cellular context is critical.

## **5. Methods**

### **5.1 Materials and Reagents**

All chemicals and reagents were used as purchased without further purification. Most chemicals were purchased from Aldrich, Fisher, VWR, EMD, Calbiochem and Chem-Impex.

### **5.2 Instruments**

The peptide-based inhibitors and probes were synthesized on a CEM Liberty microwave automatic peptide synthesizer. The compounds were purified by a C18 reverse-phase HPLC column (Waters). The compounds were characterized by an Applied Biosystems Voyager MALDI time-of-flight mass spectrometer in reflector mode or by a Perkins Elmer ESI time-of-flight mass spectrometer in positive mode. Visualization of the SDS-PAGE gel experiments was carried out by an Amersham gel imager 600 at wavelength 520 and 630 nm.

### **5.3 Chemistry**

#### **5.3.1 Photoaffinity Probes**

Probes **1-3** were prepared following the standard Fmoc strategy by solid-phase synthesis on Rink amide resin. Fmoc removal was performed with 20% (v/v) piperidine in DMF and 0.1M HOBt (2 x 10 min). After deprotection and coupling, the resin was washed

with DMF (2x5 min), NMP (2x5 min), CH<sub>2</sub>Cl<sub>2</sub> (2x5 min) and MeOH (2x5 min). Coupling of a BAL linker (4 eq) was performed with HOBt (4 eq), HBTU (4 eq) and DIPEA (10 eq).<sup>214</sup> Reductive amination was performed with propargylamine (20 eq) and sodium cyanoborohydride (20 eq).<sup>214</sup> For the probes without a BAL linker, coupling of Fmoc-propargylglycine (2 eq) was performed with HOBt (2 eq) and DIC (2 eq). Coupling of Fmoc-glycine (10 eq) was performed with DIC (5 eq) and HOBt (5 eq) and the coupling of Fmoc-Bz or Fmoc-Dz<sup>209</sup> (2 eq) was performed with HOBt (2 eq) and DIC (2 eq) in DMF. Coupling of the sequence PPKT, APKRG or PPKTG was then performed using a CEM Liberty microwave peptide synthesizer. The compounds were cleaved from the solid support in a solution of trifluoroacetic acid/2,2'-(ethylenedioxy)diethanethiol/H<sub>2</sub>O/triisopropylsilane (94:2.5:2.5:1) and confirmed by mass spectrometry.

Probes **4-6** followed the same procedure as above. The methylation of the N-terminal amine occurred through the addition of iodomethane (5 eq) and K<sub>2</sub>CO<sub>3</sub> (6 eq) in DMF.<sup>202</sup> The compounds were cleaved from the solid support in a solution of trifluoroacetic acid/2,2'-(ethylenedioxy)diethanethiol/H<sub>2</sub>O/triisopropylsilane (94:2.5:2.5:1) and confirmed by mass spectrometry.

### 5.3.2 Peptide Inhibitors

The peptides and peptide inhibitors were prepared following the standard Fmoc strategy by solid-phase synthesis on Rink amide resin. All amino acid derivatives at the first position were carboxylic acids to ensure that the standard amino acid coupling reaction would still occur. Fmoc removal was performed with 20% (v/v) piperidine in DMF and 0.1M HOBt (2 x 10 min). After deprotection and coupling, the resin was washed with

DMF (2x5 min), NMP (2x5 min), CH<sub>2</sub>Cl<sub>2</sub> (2x5 min) and MeOH (2x5 min). Synthesis of the peptides and peptide inhibitors was performed using a CEM Liberty microwave peptide synthesizer.

Three peptide inhibitors had a varying protocol from the one described above. The first is inhibitor BM-45 had a trimethylated Lys residue at the fourth position. Alloc protected Lys was coupled on to the Rink amide resin following the standard Fmoc strategy. Next, orthogonal deprotection of the alloc group was performed using I<sub>2</sub> (3 eq) in 4:1 ACCN:H<sub>2</sub>O, shaken for 48 hours.<sup>201</sup> Next, trimethylation of the Lys side chain amine occurred through the addition of iodomethane (5 eq), K<sub>2</sub>CO<sub>3</sub> (6 eq), 18-crown-6 (0.1 eq) in DMF.<sup>202</sup> Then, standard coupling of the remaining amino acids occurred through the automatic peptide synthesizer.

The other two peptide inhibitors synthesized with varying protocol from the standard synthesizer were BM-26 and BM-27. Benzyl chloride (2 eq) and Cs<sub>2</sub>CO<sub>3</sub> (2 eq) in DMF were added to the PKR and PKT resin through an S<sub>N</sub>2 reaction. All compounds were cleaved from the solid support in a solution of trifluoroacetic acid/2,2'-(ethylenedioxy)diethanethiol/H<sub>2</sub>O/triisopropylsilane (94:2.5:2.5:1) and confirmed by mass spectrometry.

## 5.4 Purification

### 5.4.1 Enzymes

His-NTMT1 was purified as previously described by Richardson, *et al.*<sup>192</sup> The gene of NTMT2 (58-278) was cloned into pET28-MKH8SUMO vector, NTMT2 was overexpressed in *Escherichia coli* BL21 (DE3) cells by induction with 0.2mM IPTG at

16°C overnight. The cells were harvested and purified by Ni-NTA, The SUMO tag was removed by TEV protease, and the protein was further purified by ion-exchange column and Superdex 200 gel filtration column. The gel filtration buffer contains 20mM Tris-HCl, pH 7.5, 150mM NaCl and 1mM DTT. Enzymes G9a and PRMT1 were expressed in *E. Coli* BL21 (DE3) codon plus RIL cells in LB medium in the presence of 50 µg/ml of kanamycin, respectively.<sup>217,218</sup>

#### **5.4.2 Chemical Probes and Inhibitors**

All peptides, peptide inhibitors, and peptide-based photoaffinity probes were purified by a C18 reverse-phase HPLC column (Waters). The compounds were characterized by an Applied Biosystems Voyager MALDI time-of-flight mass spectrometer in reflector mode or by a Perkins Elmer ESI time-of-flight mass spectrometer in positive mode. A table of the predicted and detected masses can be found in Table A3.

### **5.5 MALDI-MS Methylation Studies**

#### **5.5.1 Methylation of Probe 1 and 3**

NTMT1 (1 µL of 4 µM) was incubated with the photoaffinity probe (2 µL of 50 µM) in Tris (2 µL of pH 7.5, 250 mM), KCl (2 µL of 500 mM) and 12 µL of H<sub>2</sub>O at 37°C. After 5 min, SAM (1 µL of 1 mM) was added to initiate methylation. The reaction was quenched in a 1:1 ratio with quenching solution (50% MeCN:20 mM ammonium phosphate:0.4% TFA) at 5, 10, and 15 min.<sup>203</sup> α-Cyano-3-hydroxycinnamic acid (CHCA) was recrystallized and dissolved to a final concentration of 2 mg/mL in matrix solution (10mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub>, 0.2% (v/v) TFA in 1:1 acetonitrile/water). Samples (1 µL) were directly spotted with 1 µL of CHCA matrix solution. An average of five acquisitions were

performed for each well. The methylation progression was monitored via MALDI-MS and data was processed in Data Explorer as described before.<sup>203</sup> Final concentrations are NTMT1 (0.2  $\mu$ M), probe (5  $\mu$ M), Tris (pH 7.5, 25 mM), KCl (50 mM) and SAM (50  $\mu$ M).

### **5.5.2 Methylation of Peptide Inhibitors**

NTMT1 (1  $\mu$ L of 4  $\mu$ M) was incubated with the peptide inhibitors (2  $\mu$ L of 50  $\mu$ M) in Tris (2  $\mu$ L of pH 7.5, 250 mM), KCl (2  $\mu$ L of 500 mM) and 12  $\mu$ L of H<sub>2</sub>O at 37°C. After 5 min, SAM (1  $\mu$ L of 1 mM) was added to initiate methylation. The reaction was quenched with quenching solution (50% MeCN:20 mM ammonium phosphate:0.4% TFA) at 60 min.<sup>203</sup> The methylation progression was monitored via MALDI-MS and data was processed in Data Explorer as described before.<sup>203</sup> Final concentrations are NTMT1 (0.2  $\mu$ M), inhibitors (5  $\mu$ M), Tris (pH 7.5, 25 mM), KCl (50 mM) and SAM (50  $\mu$ M).

### **5.5.3 Methylation Progression of Substrate when Incubated with Inhibitors**

NTMT1 (1  $\mu$ L of 4  $\mu$ M) was incubated with RCC1-6 (2  $\mu$ L of 50  $\mu$ M) in Tris (2  $\mu$ L of pH 7.5, 250 mM), KCl (2  $\mu$ L of 500 mM) and 10  $\mu$ L of H<sub>2</sub>O at 37°C. The peptide inhibitors BM-11, BM-30, BM-34, BM-44, and BM-46 were then added at two concentrations (2  $\mu$ L of 10  $\mu$ M or 250  $\mu$ M). After 5 min, SAM (1  $\mu$ L of 1 mM) was added to initiate methylation of RCC1-6. The reaction was quenched with quenching solution (50% MeCN:20 mM ammonium phosphate:0.4% TFA) at 60 min.<sup>203</sup> The methylation progression was monitored via MALDI-MS and data was analyzed in Data Explorer by quantifying the area under the curve to determine the percentage of inhibition for each time point.<sup>203</sup> Final concentrations are NTMT1 (0.2  $\mu$ M), RCC1-6 (5  $\mu$ M), Tris (pH 7.5, 25 mM), KCl (50 mM), SAM (50  $\mu$ M) and inhibitors (1  $\mu$ M and 25  $\mu$ M).



## 5.6 Docking Studies

All chemical probes and peptide inhibitors were docked into the NTMT1 co-crystal structure (PDB id: 5E2B) using the program Gold53. The NTMT1 binding site was defined by a sphere of 6.0 Å where the substrate MePPKRIA was extracted and the probes and inhibitors were docked into that site. The compounds for docking were prepared using SYBYL X2.1 and the energy was minimized using the external Tripos force field. The ChemPLP score from the Gold53 docking program was considered to determine most likely docking conformations and top inhibitors.

## 5.7 SDS-PAGE Gel Studies

### 5.7.1 Concentration-dependence

NTMT1 (2 µL of 1 mg/mL) was incubated in 96-well plates with the photoaffinity probes at varying concentrations (2 µL of 0 µM, 10 µM, 20 µM, 40 µM, 80 µM, and 160 µM) in Tris (2 µL pH 7.5, 250 mM) and KCl (2 µL of 500 mM) and 12 µL of H<sub>2</sub>O on ice for 10 min for a total volume of 20 µL. The final concentrations are NTMT1 (0.1 mg/mL), probes (0 µM, 1 µM, 2 µM, 4 µM, 8 µM, and 16 µM), Tris (pH 7.5, 25 mM) and KCl (50 mM). This solution was then irradiated for 10 min (for probes containing a diazirine crosslinker) or 30 min (for probes containing a benzophenone crosslinker) at 312 nm and 4 °C. TAMRA azide (1 µL of 4.2 mM), aminoguanidine (9 µL of 12 µM), THPTA (4.5 µL of 0.5 µM), CuSO<sub>4</sub> (0.9 µL of 12 µM) and 35.6 µL of H<sub>2</sub>O were then premixed and 8.5 µL of the mixed solution was aliquoted into individual Eppendorf tubes. The 96-well plate solutions were then added to the premixed solution in the Eppendorf tubes, followed by the addition of sodium ascorbate (1.5 µL of 100 µM) for a total volume of 30 µL and reacted for one hour in the dark. The final concentrations are TAMRA azide (200 µM),

aminoguanidine (5  $\mu$ M), THPTA (0.1  $\mu$ M), CuSO<sub>4</sub> (0.5  $\mu$ M) and sodium ascorbate (5  $\mu$ M). The reaction was quenched by the addition of 7.5  $\mu$ L of 5x SDS loading dye. 20  $\mu$ L of each sample was loaded into an SDS-PAGE gel well and the reaction was analyzed on SDS-PAGE gels (12%) at 180V for 50 min. The SDS-PAGE gels were washed with destaining buffer (45% H<sub>2</sub>O:45% MeOH:10% AcOH) for one hour to remove excess fluorescent reagents, followed by washing with water (3 x 1 min). Fluorescence was detected with Amersham gel imager 600 at wavelength 520 and 630 nm. Next, the gel was stained with Coomassie Blue R-250 for 1 hour, followed by destaining with 100% destaining buffer for 1 hour and then 50:50 destaining buffer:H<sub>2</sub>O overnight. The gels were quantified using BioRad software by taking the ratio of the fluorescent band density over the Coomassie band density to account for any inconsistencies in protein loading on the gel. Studies were performed in duplicate.

### 5.7.2 Time-dependence

NTMT1 (2  $\mu$ L of 1 mg/mL) was incubated in 96-well with the photoaffinity probes (2  $\mu$ L of 80  $\mu$ M for probes with benzophenones, 2  $\mu$ L of 40  $\mu$ M for probes with diazirines) in Tris (2  $\mu$ L pH 7.5, 250 mM) and KCl (2  $\mu$ L of 500 mM) and 12  $\mu$ L of H<sub>2</sub>O on ice for 10 min for a total volume of 20  $\mu$ L. The final concentrations are NTMT1 (0.1 mg/mL), probes (8 or 4  $\mu$ M), Tris (pH 7.5, 25 mM) and KCl (50 mM). Photoaffinity probes **1** and **3** were then irradiated for 0 min, 10 min, 20 min, 30 min, 40 min or 50 min at a wavelength of 312 nm at 4 °C. Probe **2** was irradiated for 0 min, 2.5 min, 5 min, 7.5 min, 10 min or 12.5 min. Probe **4** was irradiated for 0, 10, 20, 30 and 40 min while probes **5** and **6** were irradiated for 0, 5, 10, 15, or 20 min. The click chemistry reaction and gel analysis was performed as described above. Studies were performed in duplicate.

### 5.7.3 Competition with Recombinant Protein

NTMT1 (2  $\mu$ L of 1 mg/mL) was incubated in 96-well plates with the competitor RB1-10 (2  $\mu$ L of 0, 250, 500, 750 or 1000  $\mu$ M), RB1-10 (2  $\mu$ L of 0, 100, 250, 500, 750, 1000 nM), Me<sub>2</sub>RB1-10 (2  $\mu$ L of 0, 100, 200, 300 or 400  $\mu$ M), SAH (2  $\mu$ L of 0, 250, 500, 750 or 1000  $\mu$ M) or SAM (2  $\mu$ L of 0, 0.5, 1, 1.5, 2, 2.5  $\mu$ M) in Tris (2  $\mu$ L pH 7.5, 250 mM), KCl (2  $\mu$ L 500 mM) and 8  $\mu$ L of H<sub>2</sub>O for a total volume of 18  $\mu$ L on ice for 10 min. Next, photoaffinity probes with benzophenone crosslinkers (2  $\mu$ L of 80  $\mu$ M) or diazirine crosslinkers (2  $\mu$ L of 40  $\mu$ M) were added for a final volume of 20  $\mu$ L. The final concentrations are NTMT1 (0.1 mg/mL), photoaffinity probes with benzophenone crosslinkers (8  $\mu$ M) or diazirine crosslinkers (4  $\mu$ M), RB1-10 (0, 25, 50, 75 or 100  $\mu$ M), RB1-10 (0, 10, 25, 50, 75, 100 nM), Me<sub>2</sub>RB1-10 (0, 10, 20, 30 or 40  $\mu$ M), SAH (0, 25, 50, 75 or 100  $\mu$ M), SAM (0, 0.05, 0.1, 0.15, 0.2, 0.25  $\mu$ M), Tris (pH 7.5, 25 mM) and KCl (50 mM). The photoaffinity probes were then irradiated for 30 min (probes with benzophenone crosslinkers) or 10 min (probes with diazirine crosslinkers) at 312 nm and 4 °C. The following click reaction and gel analysis was performed as described previously in the time- and concentration-dependent studies. Studies were performed in duplicate.

### 5.7.4 Selectivity with Recombinant Protein

Each enzyme (BSA, NTMT1, G9a, SAHH PRMT1; 2  $\mu$ L of 1 mg/mL) was incubated in 96-well plates with photoaffinity probes **1** and **3** (2  $\mu$ L of 80  $\mu$ M) in Tris (2  $\mu$ L of pH 7.5, 250 mM) and KCl (2  $\mu$ L of 500 mM) and 12  $\mu$ L of H<sub>2</sub>O for a total volume of 20  $\mu$ L on ice for 10 min then irradiated for 30 min at 312 nm and 4 °C. Final concentrations are enzymes (0.1 mg/mL), photoaffinity probes **1** and **3** (8  $\mu$ M), Tris (pH 7.5, 25 mM) and KCl

(50 mM). The click chemistry reaction and gel analysis was performed as described above. Studies were performed in duplicate.

### 5.7.5 Cell Lysate Studies

NTMT1 or NTMT2 (2  $\mu$ L of 1 mg/mL) was spiked in to HeLa cell lysates (2  $\mu$ L of 10 mg/mL) and incubated in 96-well plates with RB1-10 (2  $\mu$ L of 0, 250, 500  $\mu$ M) or APKRIA (2  $\mu$ L of 0, 250  $\mu$ M) in Tris (2  $\mu$ L of pH 7.5, 250 mM) and KCl (2  $\mu$ L of 50 mM) and  $\mu$ L of H<sub>2</sub>O for a total of 18  $\mu$ L on ice for 10 min. Photoaffinity probes **1** and **3** (2  $\mu$ L of 80  $\mu$ M) were then added on ice for 10 min and irradiated for 30 min at 312 nm and 4 °C, while photoaffinity probe **2** (2  $\mu$ L of 40  $\mu$ M) was added on ice for 10 min and irradiated for 10 min at 312 nm and 4 °C. Final concentrations are NTMT1/2 (0.1 mg/mL), HeLa cell lysate (1 mg/mL), photoaffinity probes **1** and **3** (8  $\mu$ M) or photoaffinity probe **2** (4  $\mu$ M), RB1-10 (0, 25, 50  $\mu$ M) or APKRIA (0, 25  $\mu$ M) in Tris (pH 7.5, 25 mM) and KCl (50 mM). The following click reaction and gel analysis was performed as described previously. Studies were performed in duplicate.

## 5.8 Kinetic Analysis of Compounds

### 5.8.1 Photoaffinity Probes

Kinetic characterization of the probes were determined using an SAH hydrolase-coupled fluorescence assay, which assess the production of SAH.<sup>192</sup> Probes **1** and **3** were incubated with NTMT1 (0.5  $\mu$ L of 20  $\mu$ M), while probe **2** was incubated with NTMT1 (0.5  $\mu$ L of 10  $\mu$ M). The reagents of the well-solution were added in the following order; H<sub>2</sub>O, buffer (10  $\mu$ L of 250 mM Tris, pH 7.4 and 500 mM KCl), SAM (1  $\mu$ L of 10 mM), SAH hydrolase (5  $\mu$ L of 10 mg/mL), NTMT1 and ThioGlo1 (1  $\mu$ L of 1.5 mM). After 10 min of

incubation at 37 °C, the reaction was initiated with varying concentrations of each photoaffinity probe in a two-fold dilution series (10 µL of 0 µM – 32 µM) for a total volume of 100 µL. The final concentrations are buffer (25 mM Tris, pH 7.4 and 50 mM KCl), SAM (100 µM), SAH hydrolase (10 µM), NTMT1 (0.1 or 0.05 µM), probe (0 – 32 µM) and ThioGlo1 (15 µM). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37 °C for 15 min. The rates were fit to the Michaelis-Menten model using least squares nonlinear regression through GraphPad Prism 7 software. All experiments were performed in triplicate.

### **5.8.2 Peptide Inhibitors**

Kinetic characterization of the peptide inhibitors were determined using a similar protocol as described above.<sup>192</sup> The inhibitors ranging in concentration (1 µL of 0.014 mM – 10 mM) and following a three-fold dilution were incubated in the well-solution that was added in the following order: H<sub>2</sub>O, buffer (10 µL of 250 mM Tris, pH 7.4 and 500 mM KCl), SAM (1 µL of 10 mM), SAH hydrolase (5 µL of 10 mg/mL), NTMT1 (0.5 µL of 40 µM) and ThioGlo1 (1 µL of 1.5 mM). After 10 min of incubation at 37 °C, the reaction was initiated with 10 µL of 50 µM RCC1-6 for a total volume of 100 µL. The final concentrations are buffer (25 mM Tris, pH 7.4 and 50 mM KCl), SAM (100 µM), SAH hydrolase (10 µM), NTMT1 (0.2 µM), inhibitors (0 – 100 µM), RCC1-6 (5 µM) and ThioGlo1 (15 µM). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37 °C for 15 min. The rates were fit to the log[inhibitor] vs response model using least squares nonlinear regression through GraphPad Prism 7 software. All experiments were performed in triplicate.

### **5.8.3 Inhibition Mechanism**

Kinetic analysis of one of the top inhibitors, BM-30, was performed to verify the inhibition mechanism using the fluorescent-based assay described above.<sup>192,219</sup> Six independent IC<sub>50</sub> studies of BM-30 were performed in triplicate with varying concentrations of substrate peptide, RCC1-6 and SAM at its  $K_m$  value. The inhibitors ranging in concentration (1  $\mu$ L of 0.14  $\mu$ M – 10 mM) and following a three-fold dilution were incubated in the well-solution that was added in the following order: H<sub>2</sub>O, buffer (10  $\mu$ L of 250 mM Tris, pH 7.4 and 500 mM KCl), SAM (1  $\mu$ L of 10  $\mu$ M), SAH hydrolase (5  $\mu$ L of 10 mg/mL), NTMT1 (0.5  $\mu$ L of 40  $\mu$ M) and ThioGlo1 (1  $\mu$ L of 1.5 mM). After 10 min of incubation at 37 °C, the reaction was initiated with each concentration of RCC1-6 (10  $\mu$ L of 5, 10, 15, 20, 40, 80  $\mu$ M) for a total volume of 100  $\mu$ L. The final concentrations are buffer (25 mM Tris, pH 7.4 and 50 mM KCl), SAM (1  $\mu$ M), SAH hydrolase (10  $\mu$ M), NTMT1 (0.2  $\mu$ M), inhibitors (0 – 100  $\mu$ M), RCC1-6 ( $0.25K_m$ ,  $0.5K_m$ ,  $0.75K_m$ ,  $K_m$ ,  $2K_m$ , and  $4K_m$ , 0.5, 1, 1.5, 2, 4, 8  $\mu$ M, respectively) and ThioGlo1 (15  $\mu$ M). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37 °C for 15 min. The rates were fit to the log[inhibitor] vs response model using least squares nonlinear regression through GraphPad Prism 7 software. The average IC<sub>50</sub> value of each independent triplicate study was then plotted against the concentration of the [RCC1-6]/ $K_m$ .

Next, the same experiment was repeated with six IC<sub>50</sub> studies of BM-30 in triplicate at varying concentrations of SAM and RCC1-6 at its  $K_m$  value. The inhibitors ranging in concentration (1  $\mu$ L of 0.14  $\mu$ M – 10 mM) and following a three-fold dilution were incubated in the well-solution that was added in the following order: H<sub>2</sub>O, buffer (10  $\mu$ L of 250 mM Tris, pH 7.4 and 500 mM KCl), SAM (1  $\mu$ L of 31.25, 62.5, 125, 250, 500, and

1000  $\mu\text{M}$ ), SAH hydrolase (5  $\mu\text{L}$  of 10 mg/mL), NTMT1 (0.5  $\mu\text{L}$  of 40  $\mu\text{M}$ ) and ThioGlo1 (1  $\mu\text{L}$  of 1.5 mM). After 10 min of incubation at 37  $^{\circ}\text{C}$ , the reaction was initiated with each concentration of RCC1-6 (10  $\mu\text{L}$  of 20  $\mu\text{M}$ ) for a total volume of 100  $\mu\text{L}$ . The final concentrations are buffer (25 mM Tris, pH 7.4 and 50 mM KCl), SAM (0.3125 $K_m$ , 0.625 $K_m$ , 1.25 $K_m$ , 2.5 $K_m$ , 5 $K_m$ , and 10 $K_m$ , 0.3125, 0.625, 1.25, 2.5, 5, 10  $\mu\text{M}$ , respectively), SAH hydrolase (10  $\mu\text{M}$ ), NTMT1 (0.2  $\mu\text{M}$ ), inhibitors (0 – 100  $\mu\text{M}$ ), RCC1-6 (2  $\mu\text{M}$ ) and Thioglo1 (15  $\mu\text{M}$ ). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37  $^{\circ}\text{C}$  for 15 min. The rates were fit to the log[inhibitor] vs response model using least squares nonlinear regression through GraphPad Prism 7 software. The average  $\text{IC}_{50}$  value of each independent triplicate study were plotted against the concentration of the  $[\text{SAM}]/K_m$ .

## 5.9 Selectivity studies

### 5.9.1 G9a

To determine selectivity, kinetic analysis of the peptide inhibitors were also carried out using the SAH hydrolase-coupled fluorescence assay against G9a. The inhibitors ranging in concentration (1  $\mu\text{L}$  of 0  $\mu\text{M}$  – 10 mM) and following a three-fold dilution were incubated with a reaction mixture containing H<sub>2</sub>O, PBS buffer (10  $\mu\text{L}$  of 10x buffer), SAM (1  $\mu\text{L}$  of 10 mM), SAH hydrolase (5  $\mu\text{L}$  of 10 mg/mL), G9a (1  $\mu\text{L}$  of 2.5  $\mu\text{M}$ ) and ThioGlo1 (1  $\mu\text{L}$  of 1.5 mM). After 10 min of incubation at 37  $^{\circ}\text{C}$ , the reaction was initiated with 10  $\mu\text{L}$  of 200  $\mu\text{M}$  of H3-15.<sup>217</sup> The final concentrations are buffer (1x PBS), SAM (100  $\mu\text{M}$ ), SAH hydrolase (10  $\mu\text{M}$ ), G9a (25 nM), inhibitors (0 – 100  $\mu\text{M}$ ), H3-15 (20  $\mu\text{M}$ ) and Thioglo1 (15  $\mu\text{M}$ ). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37  $^{\circ}\text{C}$  for 15 min. The rates were fit to the

log[inhibitor] vs response model using least squares nonlinear regression through GraphPad Prism 7 software. All experiments were performed in triplicate.

### 5.9.2 PRMT1

To determine selectivity, kinetic analysis of the peptide inhibitors were also carried out using the SAH hydrolase-coupled fluorescence assay against PRMT1. The inhibitors ranging in concentration (1  $\mu$ L of 0  $\mu$ M – 10 mM) and following a three-fold dilution were incubated with a reaction mixture containing H<sub>2</sub>O, buffer (10  $\mu$ L of 25 mM HEPES, 250 mM NaCl, 250  $\mu$ M EDTA, 500  $\mu$ M TCEP), SAM (1  $\mu$ L of 10 mM), SAH hydrolase (5  $\mu$ L of 10 mg/mL), PRMT1 (2  $\mu$ L of 10  $\mu$ M) and ThioGlo1 (1  $\mu$ L of 1.5 mM). After 10 min of incubation at 37 °C, the reaction was initiated with H4-12 (10  $\mu$ L of 1 mM).<sup>218</sup> The final concentrations are buffer (2.5 mM HEPES, 25 mM NaCl, 25  $\mu$ M EDTA, 50  $\mu$ M TCEP), SAM (100  $\mu$ M), SAH hydrolase (10  $\mu$ M), PRMT1 (0.2  $\mu$ M), inhibitors (0 – 100  $\mu$ M), H4-12 (100  $\mu$ M) and Thioglo1 (15  $\mu$ M). Fluorescence intensity was monitored using a ClarioStar microplate reader (Ex = 370 nm, Em=500 nm) at 37 °C for 15 min. The rates were fit to the log[inhibitor] vs. response model using least squares nonlinear regression through GraphPad Prism 7 software. All experiments were performed in triplicate.



### List of References

1. Gillette, T. G.; Hill, J. A. Readers, writers, and erasers: Chromatin as the whiteboard of heart disease. *Circ. Res.* **2015**, *116*, 1245–1253.
2. Lai, Z. W.; Petrera, A.; Schilling, O. Protein amino-terminal modifications and proteomic approaches for N-terminal profiling. *Curr. Opin. Chem. Biol.* **2015**, *24*, 71–79.
3. Varland, S.; Osberg, C.; Arnesen, T. N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics* **2015**, *15*, 2385–2401.
4. Aksnes, H.; Hole, K.; Arnesen, T. *Molecular, Cellular, and Physiological Significance of N-Terminal Acetylation*; Elsevier Ltd, 2015; Vol. 316.
5. Mader, D.; Liebeke, M.; Winstel, V.; Methling, K.; Leibig, M.; Götz, F.; Lalk, M.; Peschel, A. Role of N-terminal protein formylation in central metabolic processes in *Staphylococcus aureus*. *BMC Microbiol.* **2013**, *13*, 1–9.
6. Foyen, H.; Van Damme, P.; Stove, S. I.; Glomnes, N.; Evjenth, R.; Gevaert, K.; Arnesen, T. Protein N-Terminal acetyltransferases act as N-Terminal propionyltransferases in vitro and in vivo. *Mol. Cell. Proteomics* **2012**, 42–54.
7. Tooley, C. E. S.; Petkowski, J. J.; Muratore-Schroeder, T. L.; Balsbaugh, J. L.; Shabanowitz, J.; Sabat, M.; Minor, W.; Hunt, D. F.; Macara, I. G. NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. *Nature* **2010**, *466*, 1125–1128.
8. Buglino, J. A.; Resh, M. D. *Palmitoylation of Hedgehog Proteins*, 1st ed.; Elsevier Inc., 2012; Vol. 88.
9. Martin, D. D.; Beauchamp, E.; Berthiaume, L. G. Post-translational myristoylation: Fat matters in cellular life and death. *Biochimie* **2011**, *93*, 18–31.
10. Ciechanover, A.; Ben-Saadon, R. N-terminal ubiquitination: More protein substrates join in. *Trends Cell Biol.* **2004**, *14*, 103–106.
11. Polevoda, B.; Norbeck, J.; Takakura, H.; Blomberg, A.; Sherman, F. Identification and specificities of N-terminal acetyltransferases from *Saccharomyces cerevisiae*. *EMBO J.* **1999**, *18*, 6155–6168.
12. Arnesen, T.; Van Damme, P.; Polevoda, B.; Helsens, K.; Evjenth, R.; Colaert, N.; Varhaug, J. E.; Vandekerckhove, J.; Lillehaug, J. R.; Sherman, F.; Gevaert, K. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8157–8162.

13. Van Damme, P.; Evjenth, R.; Foyn, H.; Demeyer, K.; De Bock, P.; Lillehaug, J. R.; Vandekerckhove, J.; Arnesen, T.; Gevaert, K. Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of N(alpha)-acetyltransferases and point to hNaa10p as the post-translational actin N(alpha)-acetyltransferase. *Mol. Cell. proteomics MCP* **2011**, *10*, 1–12.
14. Van Damme, P.; Lasa, M.; Polevoda, B.; Gazquez, C.; Elosegui-Artola, A.; Kim, D. S.; De Juan-Pardo, E.; Demeyer, K.; Hole, K.; Larrea, E.; Timmerman, E.; Prieto, J.; Arnesen, T.; Sherman, F.; Gevaert, K.; Aldabe, R. N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB. *Proc. Natl. Acad. Sci.* **2012**, *109*, 12449–12454.
15. Polevoda, B.; Sherman, F. N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. *J. Mol. Biol.* **2003**, *325*, 595–622.
16. Starheim, K. K.; Gromyko, D.; Evjenth, R.; Rynningen, A.; Varhaug, J. E.; Lillehaug, J. R.; Arnesen, T. Knockdown of Human N-Terminal Acetyltransferase Complex C Leads to p53-Dependent Apoptosis and Aberrant Human Arl8b Localization. *Mol. Cell. Biol.* **2009**, *29*, 3569–3581.
17. Hole, K.; van Damme, P.; Dalva, M.; Aksnes, H.; Glomnes, N.; Varhaug, J. E.; Lillehaug, J. R.; Gevaert, K.; Arnesen, T. The human N-Alpha-acetyltransferase 40 (hNaa40p/hNatD) is conserved from yeast and N-terminally acetylates histones H2A and H4. *PLoS One* **2011**, *6*, 9–15.
18. Song, O. K.; Wang, X.; Waterborg, J. H.; Sternglanz, R. An N-alpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. *J. Biol. Chem.* **2003**, *278*, 38109–38112.
19. Van Damme, P.; Hole, K.; Pimenta-Marques, A.; Helsens, K.; Vandekerckhove, J.; Martinho, R. G.; Gevaert, K.; Arnesen, T. NatF contributes to an evolutionary shift in protein N-terminal acetylation and is important for normal chromosome segregation. *PLoS Genet.* **2011**, *7*, 7–16.
20. Towler, D. A.; Eubanks, S. R.; Towery, D. S.; Adams, S. P.; Glaser, L. Amino-terminal processing of proteins by N-myristoylation: Substrate specificity of N-myristoyl transferase. *J. Biol. Chem.* **1987**, *262*, 1030–1036.
21. Kozak, M. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **1983**, *47*, 1–45.
22. Petkowski, J. J.; Bonsignore, L. A.; Tooley, J. G.; Wilkey, D. W.; Merchant, M. L.; Macara, I. G.; Schaner Tooley, C. E. NRMT2 is an N-terminal monomethylase that primes for its homologue NRMT1. *Biochem. J.* **2013**, *456*, 453–462.
23. Wu, R.; Yue, Y.; Zheng, X.; Li, H. Molecular basis for histone N-terminal methylation by NRMT1. *Genes Dev.* **2015**, *29*, 2337–2342.
24. Dong, C.; Mao, Y.; Tempel, W.; Qin, S.; Li, L.; Loppnau, P.; Huang, R.; Min, J. Structural basis for substrate recognition by the human N-terminal

- methyltransferase 1. *Genes Dev.* **2015**, *29*, 2343–2349.
25. Hamey, J. J.; Winter, D. L.; Yagoub, D.; Overall, C. M.; Hart-Smith, G.; Wilkins, M. R. Novel N-terminal and Lys methyltransferases that target translation elongation factor 1A in yeast and human. *Mol. Cell. Proteomics* **2016**, *2*, 164–176.
  26. Lee, K.-E.; Ahn, J.-Y.; Kim, J.-M.; Hwang, C.-S. Synthetic lethal screen of NAA20, a catalytic subunit gene of NatB N-terminal acetylase in *Saccharomyces cerevisiae*. *J. Microbiol.* **2014**, *52*, 842–848.
  27. Liszczak, G.; Goldberg, J. M.; Foyn, H.; Petersson, E. J.; Arnesen, T.; Marmorstein, R. Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1098–1105.
  28. Liszczak, G.; Arnesen, T.; Marmorsteins, R. Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. *J. Biol. Chem.* **2011**, *286*, 37002–37010.
  29. Evjenth, R. H.; Brenner, A. K.; Thompson, P. R.; Arnesen, T.; Frøystein, N. Å.; Lillehaug, J. R. Human protein N-terminal acetyltransferase hNaa50p (hNAT5/hSAN) follows ordered sequential catalytic mechanism: Combined kinetic and NMR study. *J. Biol. Chem.* **2012**, *287*, 10081–10088.
  30. Starheim, K. K.; Gevaert, K.; Arnesen, T. Protein N-terminal acetyltransferases: when the start matters. *Trends Biochem. Sci.* **2012**, *37*, 152–161.
  31. Evjenth, R.; Hole, K.; Karlsen, O. A.; Ziegler, M.; Arnesen, T.; Lillehaug, J. R. Human Naa50p (Nat5/San) displays both protein N $\alpha$ - and N $\epsilon$ -acetyltransferase activity. *J. Biol. Chem.* **2009**, *284*, 31122–31129.
  32. Dyda, F.; Klein, D. C.; Hickman, A. B. GCN5-related N-acetyltransferases: A structural overview. *Annu Rev Biophys Biomol Struct* **2000**, *29*, 81–103.
  33. Chang, Y.-Y.; Hsu, C.-H. Structural basis for substrate-specific acetylation of N $\alpha$ -acetyltransferase Ard1 from *Sulfolobus solfataricus*. *Sci. Rep.* **2015**, *5*, 8673.
  34. Magin, R. S.; Liszczak, G. P.; Marmorstein, R. The molecular basis for Histone H4- and H2A-specific amino-terminal acetylation by NatD. *Structure* **2015**, *23*, 332–341.
  35. Stove, S. I.; Magin, R. S.; Foyn, H.; Haug, B. E.; Marmorstein, R.; Arnesen, T. Crystal structure of the golgi-associated human Na-acetyltransferase 60 reveals the molecular determinants for substrate-specific acetylation. *Structure* **2016**, *24*, 1044–1056.
  36. Van Damme, P.; Lasa, M.; Plevoda, B.; Gazquez, C.; Elosegui-Artola, A.; Kim, D. S.; De Juan-Pardo, E.; Demeyer, K.; Hole, K.; Larrea, E.; Timmerman, E.; Prieto, J.; Arnesen, T.; Sherman, F.; Gevaert, K.; Aldabe, R. N-terminal acetylome analyses and functional insights of the N-terminal acetyltransferase NatB. *Proc. Natl. Acad. Sci.* **2012**, *109*, 12449–12454.
  37. Kalvik, T. V.; Arnesen, T. Protein N-terminal acetyltransferases in cancer. *Oncogene* **2012**, *32*, 269–276.

38. Hwang, C.-S.; Shemorry, A.; Varshavsky, A. N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* **2010**, *327*, 973–977.
39. Hofmann, I. An N-terminally acetylated Arf-like GTPase is localised to lysosomes and affects their motility. *J. Cell Sci.* **2006**, *119*, 1494–1503.
40. Singer, J. M.; Shaw, J. M. Mdm20 protein functions with Nat3 protein to acetylate Tpm1 protein and regulate tropomyosin-actin interactions in budding yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7644–7649.
41. Plevoda, B.; Sherman, F. Composition and function of the eukaryotic N-terminal acetyltransferase subunits. *Biochem. Biophys. Res. Commun.* **2003**, *308*, 1–11.
42. Bartels, T.; Kim, N. C.; Luth, E. S.; Selkoe, D. J. N-alpha-acetylation of  $\alpha$ -synuclein increases its helical folding propensity, GM1 binding specificity and resistance to aggregation. *PLoS One* **2014**, *9*, 1–10.
43. Forte, G. M.; Pool, M. R.; Stirling, C. J. N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum. *PLoS Biol.* **2011**, *9*, e1001073.
44. Rope, A. F.; Wang, K.; Evjenth, R.; Xing, J.; Johnston, J. J.; Swensen, J. J.; Johnson, W. E.; Moore, B.; Huff, C. D.; Bird, L. M.; Carey, J. C.; Opitz, J. M.; Stevens, C. A.; Jiang, T.; Schank, C.; Fain, H. D.; Robison, R.; Dalley, B.; Chin, S.; South, S. T.; Pysker, T. J.; Jorde, L. B.; Hakonarson, H.; Lillehaug, J. R.; Biesecker, L. G.; Yandell, M.; Arnesen, T. Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *Am. J. Hum. Genet.* **2011**, *89*, 28–43.
45. Arnesen, T.; Starheim, K. K.; Damme, P. Van; Evjenth, R.; Dinh, H.; Betts, M. J.; Rynningen, A.; Gevaert, K.; Anderson, D. The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of huntingtin aggregation. *Mol. Cell. Biol.* **2010**, *30*, 1898–1909.
46. Asaumi, M.; Iijima, K.; Sumioka, A.; Iijima-Ando, K. Interaction of N-terminal acetyltransferase with the cytoplasmic domain of  $\beta$ -amyloid precursor protein and its effect on A $\beta$  secretion. **2005**, *137*, 147–155.
47. Fluge, Ø.; Bruland, O.; Akslén, L. A.; Varhaug, J. E.; Lillehaug, J. R. NATH, a novel gene overexpressed in papillary thyroid carcinomas. *Oncogene* **2002**, *21*, 5056–5068.
48. Martin, D. T.; Gendron, R. L.; Jarzembowski, J. A.; Perry, A.; Collins, M. H.; Pushpanathan, C.; Miskiewicz, E.; Castle, V. P.; Paradis, H. Tubedown expression correlates with the differentiation status and aggressiveness of neuroblastic tumors. *Clin. Cancer Res.* **2007**, *13*, 1480–1487.
49. Linē, A.; Stengrēvics, A.; Slucka, Z.; Li, G.; Jankevics, E.; Rees, R. C. Serological identification and expression analysis of gastric cancer-associated genes. *Br. J. Cancer* **2002**, *86*, 1824–1830.
50. Yu, M.; Gong, J.; Ma, M.; Yang, H.; Lai, J.; Wu, H.; Li, L.; Li, L.; Tan, D. Y. Immunohistochemical analysis of human arrest-defective-1 expressed in cancers

- in vivo. *Oncol. Rep.* **2009**, *21*, 909–915.
51. Midorikawa, Y.; Tsutsumi, S.; Taniguchi, H.; Ishii, M.; Kobune, Y.; Kodama, T.; Makuuchi, M.; Aburatani, H. Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling analysis. *Jpn. J. Cancer Res.* **2002**, *93*, 636–643.
  52. Lee, C. F.; Ou, D. S. C.; Lee, S. B.; Chang, L. H.; Lin, R. K.; Li, Y. S.; Upadhyay, A. K.; Cheng, X.; Wang, Y. C.; Hsu, H. S.; Hsiao, M.; Wu, C. W.; Juan, L. J. HNa10p contributes to tumorigenesis by facilitating DNMT1-mediated tumor suppressor gene silencing. *J. Clin. Invest.* **2010**, *120*, 2920–2930.
  53. Ametzazurra, A.; Larrea, E.; Civeira, M. P.; Prieto, J.; Aldabe, R. Implication of human N- $\alpha$ -acetyltransferase 5 in cellular proliferation and carcinogenesis. *Oncogene* **2008**, *27*, 7296–7306.
  54. Foyn, H.; Jones, J. E.; Lewallen, D.; Narawane, R.; Varhaug, J. E.; Thompson, P. R.; Arnesen, T. Design, synthesis, and kinetic characterization of protein N-terminal acetyltransferase inhibitors. *ACS Chem. Biol.* **2013**, *8*, 1121–1127.
  55. Dormeyer, W.; Mohammed, S.; Van Breukelen, B.; Krijgsveld, J.; Heck, A. J. R. Targeted analysis of protein termini. *J. Proteome Res.* **2007**, *6*, 4634–4645.
  56. Zhang, X.; Ye, J.; Højrup, P. A proteomics approach to study in vivo protein N $\alpha$ -modifications. *J. Proteomics* **2009**, *73*, 240–251.
  57. Cai, L.; Sutter, B. M.; Li, B.; Tu, B. P. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Mol. Cell* **2011**, *42*, 426–437.
  58. Hosokawa, Y.; Shimomura, Y.; Harris, R. A.; Ozawa, T. Determination of short-chain acyl-coenzyme A esters by high-performance liquid chromatography. *Anal. Biochem.* **1986**, *153*, 45–49.
  59. Tooley, J. G.; Schaner Tooley, C. E. New roles for old modifications: Emerging roles of N-terminal post-translational modifications in development and disease. *Protein Sci.* **2014**, *23*, 1641–1649.
  60. Khandwala, A. S.; Kasper, C. B. The Fatty Acid Composition of Individual Phospholipids from Rat Liver Nuclear Membrane and Nuclei. *J. Biol. Chem.* **1971**, *246*, 6242–6246.
  61. Sugii, M.; Okada, R.; Matsuno, H.; Miyano, S. Performance improvement in protein N-Myristoyl classification by bonsai with insignificant indexing symbol. *Genome Informatics* **2007**, *18*, 277–286.
  62. Maurer-Stroh, S.; Eisenhaber, B.; Eisenhaber, F. N-terminal N-Myristoylation of Proteins: Prediction of Substrate Proteins from Amino Acid Sequence. *J. Mol. Biol.* **2002**, *317*, 541–557.
  63. Jennings, B. C.; Linder, M. E. Regulation of G proteins by covalent modification. *Handb. Cell Signaling*, *2/e* **2010**, *2*, 1629–1633.

64. Boutin, J. A. Myristoylation. *Cell. Signal.* **1997**, *9*, 15–35.
65. Muszbek, L.; Laposata, M. Myristoylation of proteins in platelets occurs predominantly through thioester linkages. *J. Biol. Chem.* **1993**, *268*, 8251–8255.
66. Robbins, S. M.; Quintrell, N. A.; Bishop, J. M. Myristoylation and differential palmitoylation of the HCK protein-tyrosine kinases govern their attachment to membranes and association with caveolae. *Mol. Cell. Biol.* **1995**, *15*, 3507–3515.
67. Yonemoto, W.; McGlone, M. L.; Taylor, S. S. N-myristylation of the catalytic subunit of cAMP-dependent protein kinase conveys structural stability. *J. Biol. Chem.* **1993**, *268*, 2348–2352.
68. Raju, R. V; Moyana, T. N.; Sharma, R. K. N-Myristoyltransferase overexpression in human colorectal adenocarcinomas. *Exp. Cell Res.* **1997**, *235*, 145–154.
69. Cox, A. D.; Der, C. J. Protein prenylation: more than just glue? *Curr. Opin. Cell Biol.* **1992**, *4*, 1008–1016.
70. Kataoka, M.; Mihars, K.; Tokunaga, F. Recoverin alters its surface properties depending on both calcium-binding and N-terminal myristoylation. *J. Biochem.* **1993**, *114*, 535–540.
71. Raju, R. V; Kakkar, R.; Datla, R. S.; Radhi, J.; Sharma, R. K. Myristoyl-coA:protein N-myristoyltransferase from bovine cardiac muscle: molecular cloning, kinetic analysis, and in vitro proteolytic cleavage by m-calpain. *Exp. Cell Res.* **1998**, *241*, 23–35.
72. Giang, D. K.; Cravatt, B. F. A second mammalian N-myristoyltransferase. *J. Biol. Chem.* **1998**, *273*, 6595–6598.
73. Gelb, M. H.; Van Voorhis, W. C.; Buckner, F. S.; Yokoyama, K.; Eastman, R.; Carpenter, E. P.; Panethymitaki, C.; Brown, K. A.; Smith, D. F. Protein farnesyl and N-myristoyl transferases: Piggy-back medicinal chemistry targets for the development of antitrypanosomatid and antimalarial therapeutics. *Mol. Biochem. Parasitol.* **2003**, *126*, 155–163.
74. Ducker, C. E. Two N-myristoyltransferase ssozymes play unique roles in protein myristoylation, proliferation, and apoptosis. *Mol. Cancer Res.* **2005**, *3*, 463–476.
75. Lu, Y.; Selvakumar, P.; Ali, K.; Shrivastav, A.; Bajaj, G.; Resch, L.; Griebel, R.; Fourny, D.; Meguro, K.; Sharma, R. K. Expression of N-myristoyltransferase in human brain tumors. *Neurochem. Res.* **2005**, *30*, 9–13.
76. Selvakumar, P.; Lakshmikuttyamma, A.; Sharma, R. K. Biochemical characterization of bovine brain myristoyl-CoA:protein N-myristoyltransferase type 2. *J. Biomed. Biotechnol.* **2009**, *2009*, 907614.
77. Bhatnagar, R.; Fütterer, K.; Waksman, G.; Gordon, J. The structure of myristoyl-CoA:protein N-myristoyltransferase. *Biochim. Biophys. Acta* **1999**, *1441*, 162–167.
78. Zhao, C.; Ma, S. Recent advances in the discovery of N-myristoyltransferase

- inhibitors. *ChemMedChem* **2014**, *9*, 2425–2437.
79. Rocque, W. J.; McWherter, C. A.; Wood, D. C.; Gordon, J. I. A comparative analysis of the kinetic mechanism and peptide substrate specificity of human and *Saccharomyces cerevisiae* myristoyl-CoA:protein N-myristoyltransferase. *J. Biol. Chem.* **1993**, *268*, 9964–9971.
  80. Rudnick, D. A.; Rocque, W. J.; McWherter, C. A.; Toth, M. V.; Jackson-Machelski, E.; Gordon, J. I. Use of photoactivatable peptide substrates of *Saccharomyces cerevisiae* myristoyl-CoA:protein N-myristoyltransferase (Nmt1p) to characterize a myristoyl-CoA-Nmt1p-peptide ternary complex and to provide evidence for an ordered reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 1087–1091.
  81. Yamauchi, S.; Fusada, N.; Hayashi, H.; Utsumi, T.; Uozumi, N.; Endo, Y.; Tozawa, Y. The consensus motif for N-myristoylation of plant proteins in a wheat germ cell-free translation system. *FEBS J.* **2010**, *277*, 3596–3607.
  82. Bhatnagar, R. S.; Gordon, J. I. Understanding covalent modifications of proteins by lipids: Where cell biology and biophysics mingle. *Trends Cell Biol.* **1997**, *7*, 14–20.
  83. Chen, L.; Ling, B.; Alcorn, J.; Yang, J. Quantitative analysis of the expression of human N-myristoyltransferase 1 (hNMT-1) in cancers. *Open Biomark. J.* **2009**, *1*, 6–10.
  84. Shrivastav, A.; Varma, S.; Saxena, A.; DeCoteau, J.; Sharma, R. K. N-myristoyltransferase: a potential novel diagnostic marker for colon cancer. *J. Transl. Med.* **2007**, *5*, 58–63.
  85. Rajala, R. V. S.; Radhi, J. M.; Kakkar, R.; Datla, R. S. S.; Sharma, R. K. Increased expression of N-myristoyltransferase in gallbladder carcinomas. *Cancer* **2000**, *88*, 1992–1999.
  86. Shrivastav, A.; Suri, S. S.; Mohr, R.; Janardhan, K. S.; Sharma, R. K.; Singh, B. Expression and activity of N-myristoyltransferase in lung inflammation of cattle and its role in neutrophil apoptosis. *Vet. Res.* **2010**, *41*, 1–12.
  87. Georgopapadakou, N. H. Antifungals targeted to protein modification: focus on protein N-myristoyltransferase. *Expert Opin. Investig. Drugs* **2002**, *11*, 1117–1125.
  88. Gunaratne, R. S.; Sajid, M.; Ling, I. T.; Tripathi, R.; Pachebat, J. A.; Holder, A. A. Characterization of N-myristoyltransferase from *Plasmodium falciparum*. *Biochem. J.* **2000**, *348 Pt 2*, 459–463.
  89. Price, H. P.; Menon, M. R.; Panethymitaki, C.; Goulding, D.; Mckean, P. G.; Smith, D. F.; Leishmania, I. Myristoyl-CoA: Protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J. Biol. Chem.* **2003**, *278*, 7206–7214.
  90. Wright, M. H.; Clough, B.; Rackham, M. D.; Rangachari, K.; Brannigan, J. A.; Grainger, M.; Moss, D. K.; Bottrill, A. R.; Heal, W. P.; Broncel, M.; Serwa, R. A.; Brady, D.; Mann, D. J.; Leatherbarrow, R. J.; Tewari, R.; Wilkinson, A. J.; Holder, A. A.; Tate, E. W. Validation of N-myristoyltransferase as an antimalarial drug target

- using an integrated chemical biology approach. *Nat. Chem.* **2014**, *6*, 112–121.
91. Kawasaki, K. ichi; Masubuchi, M.; Morikami, K.; Sogabe, S.; Aoyama, T.; Ebiike, H.; Niizuma, S.; Hayase, M.; Fujii, T.; Sakata, K.; Shindoh, H.; Shiratori, Y.; Aoki, Y.; Ohtsuka, T.; Shimma, N. Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 3. *Bioorganic Med. Chem. Lett.* **2003**, *13*, 87–91.
  92. Taha, M. O.; Qandil, A. M.; Al-Haraznah, T.; Khalaf, R. A.; Zalloum, H.; Al-Bakri, A. G. Discovery of new antifungal leads via pharmacophore modeling and QSAR analysis of fungal N-myristoyl transferase inhibitors followed by In Silico screening. *Chem. Biol. Drug Des.* **2011**, *78*, 391–407.
  93. Gilbert, I. H. Drug discovery for neglected diseases: Molecular target-based and phenotypic approaches. *J. Med. Chem.* **2013**, *56*, 7719–7726.
  94. Tate, E. W.; Bell, A. S.; Rackham, M. D.; Wright, M. H. N-Myristoyltransferase as a potential drug target in malaria and leishmaniasis. *Parasitology* **2014**, *141*, 37–49.
  95. Das, U.; Kumar, S.; Dimmock, J. R.; Sharma, R. K. Inhibition of protein N-myristoylation: A therapeutic protocol in developing anticancer agents. *Curr. Cancer Drug Targets* **2012**, *12*, 667–692.
  96. Linder, M. E.; Deschenes, R. J. Palmitoylation: policing protein stability and traffic. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 74–84.
  97. Aicart-Ramos, C.; Valero, R. A.; Rodriguez-Crespo, I. Protein palmitoylation and subcellular trafficking. *Biochim. Biophys. Acta* **2011**, *1808*, 2981–2994.
  98. Pepinsky, R. B.; Zeng, C.; Rayhorn, P.; Baker, D. P.; Williams, P.; Bixler, S. A.; Christine, M.; Garber, E. A.; Taylor, F. R.; Elizabeth, A.; Galdes, A.; Wen, D.; Williams, K. P.; Ambrose, C. M.; Miatkowski, K.; Wang, E. a. Identification of a palmitic acid-modified form of human sonic hedgehog. *J. Biol. Chem.* **1998**, *273*, 14037–14045.
  99. Fuccillo, M.; Joyner, A. L.; Fishell, G. Morphogen to mitogen: the multiple roles of hedgehog signalling in vertebrate neural development. *Nat. Rev. Neurosci.* **2006**, *7*, 772–783.
  100. Buglino, J. A.; Resh, M. D. Hhat is a palmitoylacyltransferase with specificity for N-palmitoylation of Sonic Hedgehog. *J. Biol. Chem.* **2008**, *283*, 22076–22088.
  101. Chang, C. C. Y.; Sun, J.; Chang, T.-Y. Membrane-bound O-acyltransferases (MBOATs). *Front. Biol. (Beijing)*. **2011**, *6*, 177–182.
  102. Hoffman, K. A superfamily of membrane-bound O-acetyltransferases with implications for Wnt signalling. *Trends Biochem. Sci.* **2000**, *25*, 111–112.
  103. Taylor, F. R.; Wen, D.; Garber, E. A.; Carmillo, A. N.; Baker, D. P.; Arduini, R. M.; Williams, K. P.; Weinreb, P. H.; Rayhorn, P.; Hronowski, X.; Whitty, A.; Day, E. S.; Boriack-Sjodin, A.; Shapiro, R. I.; Galdes, A.; Pepinsky, R. B. Enhanced potency of human Sonic hedgehog by hydrophobic modification. *Biochemistry* **2001**, *40*,



- 4359–4371.
104. Chen, M. H.; Li, Y. J.; Kawakami, T.; Xu, S. M.; Chuang, P. T. Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. *Genes Dev.* **2004**, *18*, 641–659.
  105. Constantinides, P. P.; Steim, J. M. Physical properties of fatty acyl-CoA. *J. Biol. Chem.* **1985**, *260*, 7573–7580.
  106. Hardy, R. Y.; Resh, M. D. Identification of N-terminal residues of sonic hedgehog important for palmitoylation by Hedgehog acyltransferase. *J. Biol. Chem.* **2012**, *287*, 42881–42889.
  107. Buglino, J. A.; Resh, M. D. Identification of conserved regions and residues within hedgehog acyltransferase critical for palmitoylation of Sonic hedgehog. *PLoS One* **2010**, *5*, 17–20.
  108. Yang, J.; Brown, M. S.; Liang, G.; Grishin, N. V.; Goldstein, J. L. Identification of the acyltransferase that octanoylates ghrelin, an appetite-stimulating peptide hormone. *Cell* **2008**, *132*, 387–396.
  109. Bosson, R.; Jaquenoud, M.; Conzelmann, A. GUP1 of *Saccharomyces cerevisiae* encodes an O-acyltransferase involved in remodeling of the GPI anchor. *Mol. Biol. Cell* **2006**, *17*, 2636–2645.
  110. Guo, Z. Y.; Lin, S.; Heinen, J. A.; Chang, C. C. Y.; Chang, T. Y. The active site His-460 of human acyl-coenzyme A:cholesterol acyltransferase 1 resides in a hitherto undisclosed transmembrane domain. *J. Biol. Chem.* **2005**, *280*, 37814–37826.
  111. Konitsiotis, A. D.; Chang, S. C.; Jovanovic, B.; Ciepla, P.; Masumoto, N.; Palmer, C. P.; Tate, E. W.; Couchman, J. R.; Magee, A. I. Attenuation of hedgehog acyltransferase-catalyzed sonic hedgehog palmitoylation causes reduced signaling, proliferation and invasiveness of human carcinoma cells. *PLoS One* **2014**, *9*, e89899.
  112. Justilien, V.; Walsh, M. P.; Ali, S. A.; Thompson, E. A.; Murray, N. R.; Fields, A. P. The PRKCI and SOX2 oncogenes are co-amplified and cooperate to activate Hedgehog signaling in lung squamous cell carcinoma. *Cancer Cell* **2014**, *25*, 139–151.
  113. Matevossian, A.; Resh, M. D. Hedgehog Acyltransferase as a target in estrogen receptor positive, HER2 amplified, and tamoxifen resistant breast cancer cells. *Mol. Cancer* **2015**, *14*, 72.
  114. Petrova, E.; Matevossian, A.; Resh, M. D. Hedgehog acyltransferase as a target in pancreatic ductal adenocarcinoma. *Oncogene* **2015**, *34*, 263–268.
  115. Petrova, E.; Rios-Esteves, J.; Ouerfelli, O.; Glickman, J. F.; Resh, M. D. Inhibitors of Hedgehog acyltransferase block Sonic Hedgehog signaling. *Nat. Chem. Biol.* **2013**, *9*, 247–249.
  116. Goldstein, G.; Scheid, M.; Hammerling, U.; Schlesinger, D. H.; Niall, H. D.; Boyse,

- E. A. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 11–15.
117. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **2001**, *70*, 503–533.
  118. Pickart, C. M.; Eddins, M. J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta* **2004**, *1695*, 55–72.
  119. McDowell, G. S.; Philpott, A. Non-canonical ubiquitylation: Mechanisms and consequences. *Int. J. Biochem. Cell Biol.* **2013**, *45*, 1833–1842.
  120. Breitschopf, K.; Bengal, E.; Ziv, T.; Admon, A.; Ciechanover, A. A novel site for ubiquitination: The N-terminal residue, and not internal Lys of MyoD, is essential for conjugation and degradation of the protein. *EMBO J.* **1998**, *17*, 5964–5973.
  121. Bloom, J.; Amador, V.; Bartolini, F.; DeMartino, G.; Pagano, M. Proteasome-mediated degradation of p21 via N-terminal ubiquitylation. *Cell* **2003**, *115*, 71–82.
  122. Scaglione, K. M.; Basrur, V.; Ashraf, N. S.; Konen, J. R.; Elenitoba-Johnson, K. S. J.; Todi, S. V.; Paulson, H. L. The ubiquitin-conjugating enzyme (E2) ube2w ubiquitinates the N terminus of substrates. *J. Biol. Chem.* **2013**, *288*, 18784–18788.
  123. Sadeh, R.; Breitschopf, K.; Bercovich, B.; Zoabi, M.; Kravtsova-Ivantsiv, Y.; Kornitzer, D.; Schwartz, A.; Ciechanover, A. The N-terminal domain of MyoD is necessary and sufficient for its nuclear localization-dependent degradation by the ubiquitin system. *Proc. Natl. Acad. Sci.* **2008**, *105*, 15690–15695.
  124. Coulombe, P.; Rodier, G.; Bonneil, E.; Thibault, P. N-terminal ubiquitination of extracellular signal-regulated Kinase 3 and p21 directs their degradation by the proteasome. *Mol. Cell. Biol.* **2004**, *24*, 6140–6150.
  125. Kuo, M. L.; Den Besten, W.; Bertwistle, D.; Roussel, M. F.; Sherr, C. J. N-terminal polyubiquitination and degradation of the Arf tumor suppressor. *Genes Dev.* **2004**, *18*, 1862–1874.
  126. Ben-Saadon, R.; Fajerman, I.; Ziv, T.; Hellman, U.; Schwartz, A. L.; Ciechanover, A. The tumor suppressor protein p16INK4a and the human papillomavirus oncoprotein-58 E7 are naturally occurring Lys-less proteins that are degraded by the ubiquitin system: Direct evidence for ubiquitination at the N-terminal residue. *J. Biol. Chem.* **2004**, *279*, 41414–41421.
  127. Reinstein, E.; Scheffner, M.; Oren, M.; Ciechanover, A.; Schwartz, A. Degradation of the E7 human papillomavirus oncoprotein by the ubiquitin-proteasome system: targeting via ubiquitination of the N-terminal residue. *Oncogene* **2000**, *19*, 5944–5950.
  128. Aviel, S.; Winberg, G.; Massucci, M.; Ciechanover, A. Degradation of the Epstein-Barr virus latent membrane protein 1 (LMP1) by the ubiquitin-proteasome pathway: Targeting via ubiquitination of the N-terminal residue. *J. Biol. Chem.* **2000**, *275*, 23491–23499.

129. Ikeda, M.; Ikeda, A.; Longnecker, R. Lys-independent ubiquitination of Epstein–Barr virus LMP2A. *Virology* **2002**, *300*, 153–159.
130. Li, H.; Okamoto, K.; Peart, M. J.; Prives, C. Lys-independent turnover of cyclin G1 can be stabilized by B'alpha subunits of protein phosphatase 2A. *Mol. Cell. Biol.* **2009**, *29*, 919–928.
131. Wang, Y.; Shao, Q.; Yu, X.; Kong, W.; Hildreth, J. E. K.; Liu, B. N-terminal hemagglutinin tag renders Lys-deficient APOBEC3G resistant to HIV-1 Vif-induced degradation by reduced polyubiquitination. *J. Virol.* **2011**, *85*, 4510–4519.
132. Traasch-Azar, J. S.; Lingbeck, J.; Ciechanover, A.; Schwartz, A. L. Ubiquitin-proteasome-mediated degradation of Id1 is modulated by MyoD. *J. Biol. Chem.* **2004**, *279*, 32614–32619.
133. Trausch-Azar, J.; Leone, T. C.; Kelly, D. P.; Schwartz, A. L. Ubiquitin proteasome-dependent degradation of the transcriptional coactivator PGC-1a via the N-terminal pathway. *J. Biol. Chem.* **2010**, *285*, 40192–40200.
134. Vosper, J. M. D.; McDowell, G. S.; Hindley, C. J.; Fiore-Herliche, C. S.; Kucerova, R.; Horan, I.; Philpott, A. Ubiquitylation on canonical and non-canonical sites targets the transcription factor neurogenin for ubiquitin-mediated proteolysis. *J. Biol. Chem.* **2009**, *284*, 15458–15468.
135. Tatham, M. H.; Plechanovová, A.; Jaffray, E. G.; Salmen, H.; Hay, R. T. Ube2W conjugates ubiquitin to  $\alpha$ -amino groups of protein N-termini. *Biochem. J.* **2013**, *453*, 137–145.
136. Yang, J.; Hong, Y.; Wang, W.; Wu, W.; Chi, Y.; Zong, H.; Kong, X.; Wei, Y.; Yun, X.; Cheng, C.; Chen, K.; Gu, J. HSP70 protects BCL2L12 and BCL2L12A from N-terminal ubiquitination-mediated proteasomal degradation. *FEBS Lett.* **2009**, *583*, 1409–1414.
137. Fajerman, I.; Schwartz, A. L.; Ciechanover, A. Degradation of the Id2 developmental regulator: Targeting via N-terminal ubiquitination. *Biochem. Biophys. Res. Commun.* **2004**, *314*, 505–512.
138. McDowell, G. S.; Kucerova, R.; Philpott, A. Non-canonical ubiquitylation of the proneural protein Ngn2 occurs in both *Xenopus* embryos and mammalian cells. *Biochem. Biophys. Res. Commun.* **2010**, *400*, 655–660.
139. Vittal, V.; Shi, L.; Wenzel, D. M.; Scaglione, K. M.; Duncan, E. D.; Basrur, V.; Elenitoba-Johnson, K. S. J.; Baker, D.; Paulson, H. L.; Brzovic, P. S.; Klevit, R. E. Intrinsic disorder drives N-terminal ubiquitination by Ube2w. *Nat Chem Biol* **2015**, *11*, 83–89.
140. Mukhopadhyay, D.; Riezman, H. Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science (80- )*. **2007**, *315*, 201–205.
141. Schnell, J. D.; Hicke, L. Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *J. Biol. Chem.* **2003**, *278*, 35857–35860.

142. Garcia-Higuera, I.; Taniguchi, T.; Ganesan, S.; Meyn, M. S.; Timmers, C.; Hejna, J.; Grompe, M.; D'Andrea, A. D. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol. Cell* **2001**, *7*, 249–262.
143. Hochstrasser, M. Evolution and function of ubiquitin-like protein-conjugation systems. *Nat. Cell Biol.* **2000**, *2*, E153-157.
144. Varshavsky, A. The ubiquitin system. *Annu. Rev. Biochem.* **1998**, *22*, 383–387.
145. Hershko, A.; Ciechanover, A.; Rose, I. A. Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown. *J. Biol. Chem.* **1981**, *256*, 1525–1528.
146. Wu, P. Y.; Hanlon, M.; Eddins, M.; Tsui, C.; Rogers, R. S.; Jensen, J. P.; Matunis, M. J.; Weissman, A. M.; Wolberger, C. P.; Pickart, C. M. A conserved catalytic residue in the ubiquitin-conjugating enzyme family. *EMBO J.* **2003**, *22*, 5241–5250.
147. Hershko, A.; Heller, H.; Elias, S.; Ciechanover, A. Components of ubiquitin-protein ligase system. *J. Biol. Chem.* **1983**, *258*, 8206–8214.
148. Koegl, M.; Hoppe, T.; Schlenker, S.; Ulrich, H. D.; Mayer, T. U.; Jentsch, S. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **1999**, *96*, 635–644.
149. Grossman, S. R.; Deato, M. E.; Brignone, C.; Chan, H. M. Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science (80- )*. **2003**, *300*, 342–344.
150. Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M.; Amerik, A.; Swaminathan, S.; Krantz, B.; Wilkinson, K.; Hochstrasser, M.; Pickart, C.; et al. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **2000**, *19*, 94–102.
151. Komander, D.; Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **2012**, *81*, 203–229.
152. Lecker, S. H. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *J. Am. Soc. Nephrol.* **2006**, *17*, 1807–1819.
153. Bedford, L.; Lowe, J.; Dick, L. R.; Mayer, R. J.; Brownell, J. E. Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. *Nat. Rev. Drug Discov.* **2011**, *10*, 29–46.
154. Kessler, B. M. Ubiquitin - omics reveals novel networks and associations with human disease. *Curr. Opin. Chem. Biol.* **2013**, *17*, 59–65.
155. Corti, O.; Lesage, S.; Brice, A. What genetics tells us about the causes and mechanisms of Parkinson's disease. *Physiol Rev.* **2011**, *91*, 1161–1218.
156. Adhikary, S.; Marinoni, F.; Hock, A.; Hulleman, E.; Popov, N.; Beier, R.; Bernard, S.; Quarto, M.; Capra, M.; Goettig, S.; Kogel, U.; Scheffner, M.; Helin, K.; Eilers, M. The ubiquitin ligase HectH9 regulates transcriptional activation by Myc and is essential for tumor cell proliferation. *Cell* **2005**, *123*, 409–421.
157. Li, M.; Brooks, C. L.; Wu-baer, F.; Chen, D.; Baer, R.; Li, M.; Brooks, C. L.; Wu-

- baer, F.; Chen, D. Mono- versus polyubiquitination: Differential control of p53 fate by Mdm2. *Science* (80-. ). **2003**, *302*, 1972–1975.
158. Richardson, P. G.; Barlogie, B.; Berenson, J.; Singhal, S.; Jagannath, S.; Irwin, D.; Rajkumar, S. V.; Srkalovic, G.; Alsina, M.; Alexanian, R.; Siegel, D.; Orłowski, R. Z.; Kuter, D.; Limentani, S. a; Lee, S.; Hideshima, T.; Esseltine, D.-L.; Kauffman, M.; Adams, J.; Schenkein, D. P.; Anderson, K. C. A phase 2 study of bortezomib in relapsed, refractory myeloma. *N. Engl. J. Med.* **2003**, *348*, 2609–2617.
  159. Cohen, P.; Tcherpakov, M. Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell* **2010**, *143*, 686–693.
  160. Rentsch, A.; Landsberg, D.; Brodmann, T.; B??low, L.; Girbig, A. K.; Kalesse, M. Synthesis and pharmacology of proteasome inhibitors. *Angew. Chemie - Int. Ed.* **2013**, *52*, 5450–5488.
  161. Zhang, W.; Sidhu, S. S. Development of inhibitors in the ubiquitination cascade. *FEBS Lett.* **2014**, *588*, 356–367.
  162. Adams, J. M.; Capecchi, M. N-Formylmethionyl-sRNA as the initiator of protein synthesis. *Proc. Natl. Acad. Sci.* **1966**, *55*, 147–155.
  163. Marcker, K.; Sanger, F. N-Formyl-methionyl-S-RNA. *J. Mol. Biol.* **1964**, *8*, 835–IN8.
  164. Newton, D. T.; Creuzenet, C.; Mangroo, D. Formylation is not essential for initiation of protein synthesis in all eubacteria. *J. Biol. Chem.* **1999**, *274*, 22143–22146.
  165. Margolis, P. S.; Hackbarth, C. J.; Young, D. C.; Wang, W.; Chen, D.; Yuan, Z.; White, R.; Trias, J. Peptide deformylase in *Staphylococcus aureus*: Resistance to inhibition is mediated by mutations in the formyltransferase gene. *Antimicrob. Agents Chemother.* **2000**, *44*, 1825–1831.
  166. Mader, D.; Liebeke, M.; Winstel, V.; Methling, K.; Leibig, M.; Götz, F.; Lalk, M.; Peschel, A. Role of N-terminal protein formylation in central metabolic processes in *Staphylococcus aureus*. *BMC Microbiol.* **2013**, *13*, 1–9.
  167. Galkin, A.; Kulakova, L.; Sarikaya, E.; Lim, K.; Howard, A.; Herzberg, O. Structural insight into Arg degradation by Arg deiminase, an antibacterial and parasite drug target. *J. Biol. Chem.* **2004**, *279*, 14001–14008.
  168. Leeds, J. A.; Dean, C. R. Peptide deformylase as an antibacterial target: a critical assessment. *Curr. Opin. Pharmacol.* **2006**, *6*, 445–452.
  169. Mazel, D.; Pochet, S.; Marliere, P. Genetic characterization of polypeptide deformylase, a distinctive enzyme of eubacterial translation. *EMBO J.* **1994**, *13*, 914–923.
  170. Giglione, C.; Pierre, M.; Meinnel, T. Peptide deformylase as a target for new generation, broad spectrum antimicrobial agents. *Mol. Microbiol.* **2000**, *36*, 1197–1205.
  171. Chen, D. Z.; Patel, D. V.; Hackbarth, C. J.; Wang, W.; Dreyer, G.; Young, D. C.;

- Margolis, P. S.; Wu, C.; Ni, Z. J.; Trias, J.; White, R. J.; Yuan, Z. Actinonin, a naturally occurring antibacterial agent, is a potent deformylase inhibitor. *Biochemistry* **2000**, *39*, 1256–1262.
172. Jain, R.; Chen, D.; White, R. J.; Patel, D. V.; Yuan, Z. Bacterial Peptide deformylase inhibitors: a new class of antibacterial agents. *Curr. Med. Chem.* **2005**, *12*, 1607–1621.
173. Clements, J. M.; Beckett, R. P.; Brown, A.; Catlin, G.; Lobell, M.; Palan, S.; Thomas, W.; Whittaker, M.; Wood, S.; Salama, S.; Baker, P. J.; Rodgers, H. F.; Barynin, V.; Rice, D. W.; Hunter, M. G. Antibiotic activity and characterization of BB-3497, a novel peptide deformylase inhibitor. *Antimicrob. Agents Chemother.* **2001**, *45*, 563–570.
174. Molteni, V.; He, X.; Nabakka, J.; Yang, K.; Kreuzsch, A.; Gordon, P.; Bursulaya, B.; Warner, I.; Shin, T.; Biorac, T.; Ryder, N. S.; Goldberg, R.; Doughty, J.; He, Y. Identification of novel potent bicyclic peptide deformylase inhibitors. *Bioorganic Med. Chem. Lett.* **2004**, *14*, 1477–1481.
175. Verma, S. K.; Jat, R. K.; Nagar, N. L.; Saharan, R.; Sharma, V.; Pandey, S.; Bansal, K. A novel antibacterial target: Peptidyl deformylase. *Pharmacophore* **2011**, *2*, 114–123.
176. Azoulay-Dupuis, E.; Mohler, J.; Bédos, J. P. Efficacy of BB-83698, a novel peptide deformylase inhibitor, in a mouse model of Pneumococcal Pneumonia. *Antimicrob. Agents Chemother.* **2004**, *48*, 80–85.
177. Naderer, O. J.; Jones, L. S.; Zhu, J.; Kurtinecz, M.; Dumont, E. Safety, tolerability, and pharmacokinetics of oral and intravenous administration of GSK1322322, a peptide deformylase inhibitor. *J. Clin. Pharmacol.* **2013**, *53*, 1168–1176.
178. Wittmann-Liebold, B.; Pannenbecker, R. Primary structure of protein L33 from the large subunit of the Escherichia Coli ribosome. *FEBS Lett.* **1976**, *68*, 115–118.
179. Martinage, A.; Briand, G.; Van Dorsselaer, A.; Turner, C. H.; Sautiere, P. Primary structure of histone H2B from gonads of the starfish *Asterias rubens*. *Eur. J. Biochem.* **1985**, *147*, 351–359.
180. Pettigrew, G. W.; Smith, G. M. Novel N-terminal protein blocking group identified as dimethylproline. *Nature* **1977**, *265*, 661–662.
181. Henry, G. D.; Winstanley, M. A.; Dalgarno, D. C.; Scott, G. M. M.; Levine, B. A.; Trayer, I. P. Characterization of the actin-binding site on the alkali light chain of myosin. *Biochim. Biophys. Acta* **1985**, *830*, 223–243.
182. Webb, K. J.; Lipson, R. S.; Al-hadid, Q.; Whitelegge, J. P.; Clarke, S. G. Identification of protein N-terminal methyltransferases in yeast and humans. *Biochemistry* **2010**, *49*, 5225–5235.
183. Bonsignore, L. A.; Tooley, J. G.; Van Hoose, P. M.; Wang, E.; Cheng, A.; Cole, M. P.; Schaner Tooley, C. E. NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. *Mech. Ageing Dev.* **2015**, *146*–148,

42–52.

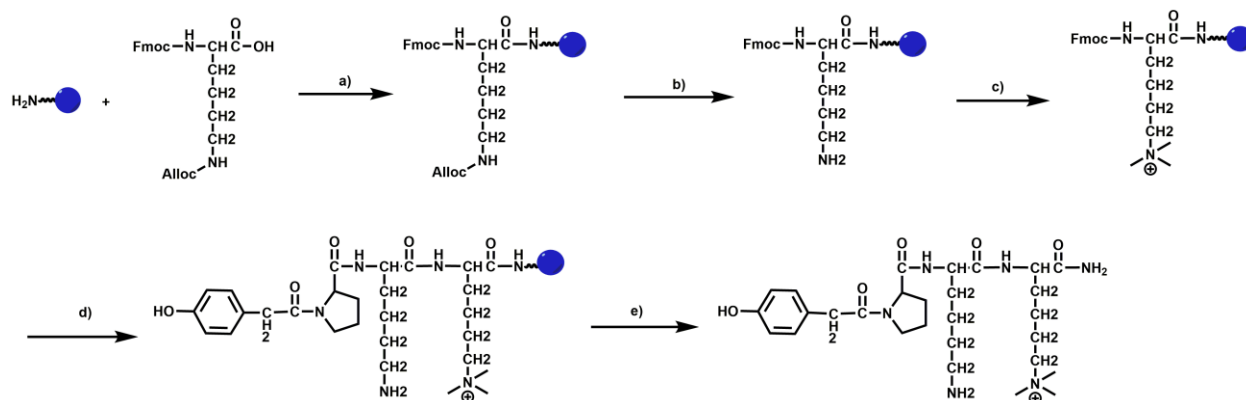
184. Cai, Q.; Fu, L.; Wang, Z.; Gan, N.; Dai, X.; Wang, Y. A-N-methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair. *J. Biol. Chem.* **2014**, *289*, 16046–16056.
185. Dai, X.; Otake, K.; You, C.; Cai, Q.; Wang, Z.; Masumoto, H.; Wang, Y. Identification of novel  $\alpha$ -N-methylation of CENP-B that regulates its binding to the centromeric DNA. *J. Proteome Res.* **2013**, *12*, 4167–4175.
186. Bailey, A. O.; Panchenko, T.; Sathyan, K. M.; Petkowski, J. J.; Pai, P.-J.; Bai, D. L.; Russell, D. H.; Macara, I. G.; Shabanowitz, J.; Hunt, D. F.; Black, B. E.; Foltz, D. R. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 11827–11832.
187. Chen, T.; Muratore, T. L.; Schaner-Tooley, C. E.; Shabanowitz, J.; Hunt, D. F.; Macara, I. G. N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. *Nat. Cell Biol.* **2007**, *9*, 596–603.
188. Stock, A.; Clarke, S.; Stock, J. N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett.* **1987**, *220*, 8–14.
189. Villar-Garea, A.; Forne, I.; Vetter, I.; Kremmer, E.; Thomae, A.; Imhof, A. Developmental regulation of N-terminal H2B methylation in *Drosophila melanogaster*. *Nucleic Acids Res.* **2012**, *40*, 1536–1549.
190. Antonysamy, S.; Bonday, Z.; Campbell, R. M.; Doyle, B.; Druzina, Z.; Gheyi, T. Crystal structure of the human PRMT5:MEP50 complex. *Proc. Natl. Acad. Sci.* **2012**, *109*, 1–6.
191. Petkowski, J. J.; Tooley, C. E. S.; Anderson, L. C.; Shumilin, I. A.; Balsbaugh, J. L.; Hunt, D. F.; Minor, W.; Macara, I. G. Substrate specificity of mammalian N-terminal  $\alpha$ -amino methyltransferase NRMT. *Biochemistry* **2012**, *51*, 5942–5950.
192. Richardson, S. L.; Mao, Y.; Zhang, G.; Hanjra, P.; Peterson, D. L.; Huang, R. Kinetic mechanism of protein N-terminal methyltransferase 1. *J. Biol. Chem.* **2015**, *290*, 11601–11610.
193. Bonsignore, L. A.; Butler, J. S.; Klinge, C. M.; Tooley, C. E. S.; Schaner Tooley, C. E. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. *Oncotarget* **2015**, *6*, 12248–12263.
194. Uhlen, M.; Björk, E.; Agaton, C.; Szigartyo, C. A.; Amini, B.; Andersen, E.; Andersson, A.; Angelidou, P.; Asplund, A.; Ponte, F.; et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **2005**, *4*, 1920–1932.
195. Dowden, J.; Hong, W.; Parry, R. V.; Pike, R. A.; Ward, S. G. Toward the development of potent and selective bisubstrate inhibitors of protein Arg methyltransferases. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2103–2105.
196. Osborne, T.; Roska, R. L. W.; Rajski, S. R.; Thompson, P. R. In situ generation of

- a bisubstrate analogue for protein Arg methyltransferase 1. *J. Am. Chem. Soc.* **2008**, *130*, 4574–4575.
197. Zhang, G.; Richardson, S. L.; Mao, Y.; Huang, R. Design, synthesis, and kinetic analysis of potent protein N-terminal methyltransferase 1 inhibitors. *Org. Biomol. Chem.* **2015**, *13*, 4149–4154.
  198. Zhang, G.; Huang, R. Facile synthesis of SAM–peptide conjugates through alkyl linkers targeting protein N-terminal methyltransferase 1. *RSC Adv.* **2016**, *6*, 6768–6771.
  199. De Poli, M.; Moretto, A.; Crisma, M.; Peggion, C.; Formaggio, F.; Kaptein, B.; Broxterman, Q. B.; Toniolo, C. Is the backbone conformation of C $\alpha$ -methyl proline restricted to a single region? *Chem. - A Eur. J.* **2009**, *15*, 8015–8025.
  200. Stuckey, J. I.; Simpson, C.; Norris-Drouin, J. L.; Cholensky, S. H.; Lee, J.; Pasca, R.; Cheng, N.; Dickson, B. M.; Pearce, K. H.; Frye, S. V.; James, L. I. Structure-activity relationships and kinetic studies of peptidic antagonists of CBX chromodomains. *J. Med. Chem.* **2016**, *59*, 8913–8923.
  201. Szumigala, R. H.; Onofiok, E.; Karady, S.; Armstrong, J. D.; Miller, R. A. Mild non-transition metal catalyzed deprotection of N-allyloxycarbonyl amines. *Tetrahedron Lett.* **2005**, *46*, 4403–4405.
  202. Santhiya, D.; Dias, R. S.; Shome, A.; Das, P. K.; Miguel, M. G.; Lindman, B.; Maiti, S. Role of linker groups between hydrophilic and hydrophobic moieties of cationic surfactants on oligonucleotide-surfactant interactions. *Langmuir* **2009**, *25*, 13770–13775.
  203. Richardson, S. L.; Hanjra, P.; Zhang, G.; Mackie, B. D.; Peterson, D. L.; Huang, R. A direct, ratiometric, and quantitative MALDI–MS assay for protein methyltransferases and acetyltransferases. *Anal. Biochem.* **2015**, *478*, 59–64.
  204. Schmidinger, H.; Hermetter, A.; Birner-Gruenberger, R. Activity-based proteomics: Enzymatic activity profiling in complex proteomes. *Amino Acids* **2006**, *30*, 333–350.
  205. Willems, L. I.; Overkleeft, H. S.; Van Kasteren, S. I. Current developments in activity-based protein profiling. *Bioconjug. Chem.* **2014**, *25*, 1181–1191.
  206. Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383–414.
  207. Serim, S.; Haedke, U.; Verhelst, S. H. L. Activity-based probes for the study of proteases: Recent advances and developments. *ChemMedChem* **2012**, *7*, 1146–1159.
  208. Dormán, G.; Prestwich, G. D. Benzophenone photophores in biochemistry. *Biochemistry* **1994**, *33*, 5661–5673.
  209. Yang, T.; Liu, Z.; Li, X. D. Developing diazirine-based chemical probes to identify histone modification “readers” and “erasers.” *Chem. Sci.* **2015**, *6*, 1011–1017.

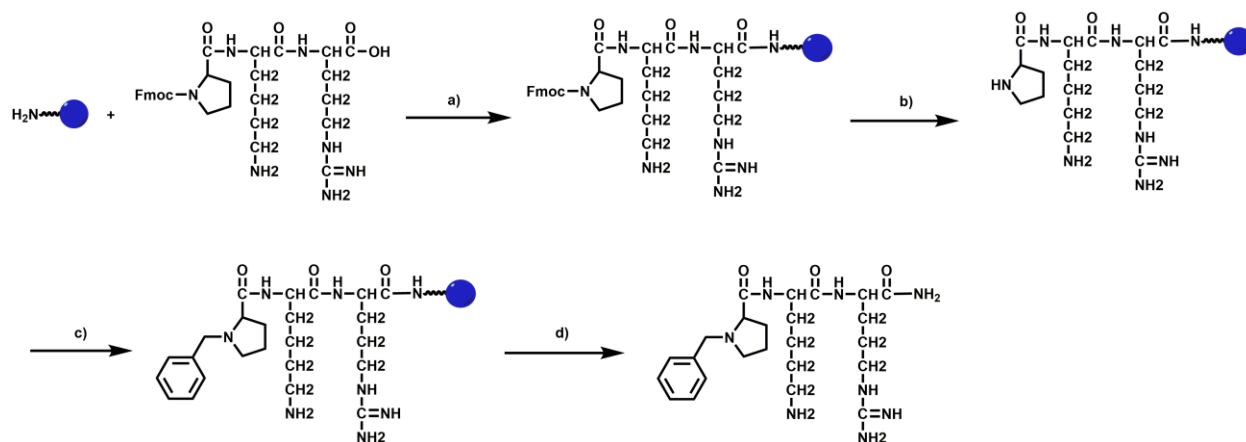


210. Dubinsky, L.; Krom, B. P.; Meijler, M. M. Diazirine based photoaffinity labeling. *Bioorganic Med. Chem.* **2012**, *20*, 554–570.
211. Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J. Am. Chem. Soc.* **2010**, *132*, 14819–14824.
212. Meng Zhang, P. C. A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat. Chem. Biol.* **2011**, *7*, 671–677.
213. Kawamura, A.; Hindi, S.; Mihai, D. M.; James, L.; Aminova, O. Binding is not enough: Flexibility is needed for photocrosslinking of Lck kinase by benzophenone photoligands. *Bioorganic Med. Chem.* **2008**, *16*, 8824–8829.
214. Brandt, M.; Gammeltoft, S.; Jensen, K. J. Microwave heating for solid-phase peptide synthesis: General evaluation and application to 15-mer phosphopeptides. *Int. J. Pept. Res. Ther.* **2006**, *12*, 349–357.
215. Morozov, Y. I.; Agaronyan, K.; Cheung, A. C. M.; Anikin, M.; Cramer, P.; Temiakov, D. A novel intermediate in transcription initiation by human mitochondrial RNA polymerase. *Nucleic Acids Res.* **2014**, *42*, 3884–3893.
216. Montgomery, D. C.; Sorum, A. W.; Meier, J. L. Chemoproteomic profiling of Lys acetyltransferases highlights an expanded landscape of catalytic acetylation. *J. Am. Chem. Soc.* **2014**, *136*, 8669–8676.
217. Liu, F.; Chen, X.; Allali-hassani, A.; Quinn, A. M.; Wigle, T. J.; Wasney, G. A.; Dong, A.; Senisterra, G.; Chau, I.; Siarheyeva, A.; Norris, J. L.; Kireev, D. B.; Jadhav, A.; Herold, J. M.; Janzen, W. P.; Arrowsmith, C. H.; Frye, S. V.; Brown, P. J.; Simeonov, A. Protein Lys methyltransferase G9a inhibitors: Design, synthesis, and structure activity relationships of 2,4-diamino-7-aminoalkoxy-quinazolines. *J. Med. Chem.* **2010**, *373*, 5844–5857.
218. Feng, Y.; Xie, N.; Jin, M.; Stahley, M. R.; Stivers, J. T.; Zheng, Y. G. A transient kinetic analysis of PRMT1 catalysis. *Biochemistry* **2011**, *50*, 7033–7044.
219. Chan-Penebre, E.; Kuplast, K. G.; Majer, C. R.; Boriack-Sjodin, P. A.; Wigle, T. J.; Johnston, L. D.; Rioux, N.; Munchhof, M. J.; Jin, L.; Jacques, S. L.; West, K. A.; Lingaraj, T.; Stickland, K.; Ribich, S. A.; Raimondi, A.; Scott, M. P.; Waters, N. J.; Pollock, R. M.; Smith, J. J.; Barbash, O.; Pappalardi, M.; Ho, T. F.; Nurse, K.; Oza, K. P.; Gallagher, K. T.; Kruger, R.; Moyer, M. P.; Copeland, R. A.; Chesworth, R.; Duncan, K. W. A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. *Nat. Chem. Biol.* **2015**, *11*, 432–437.

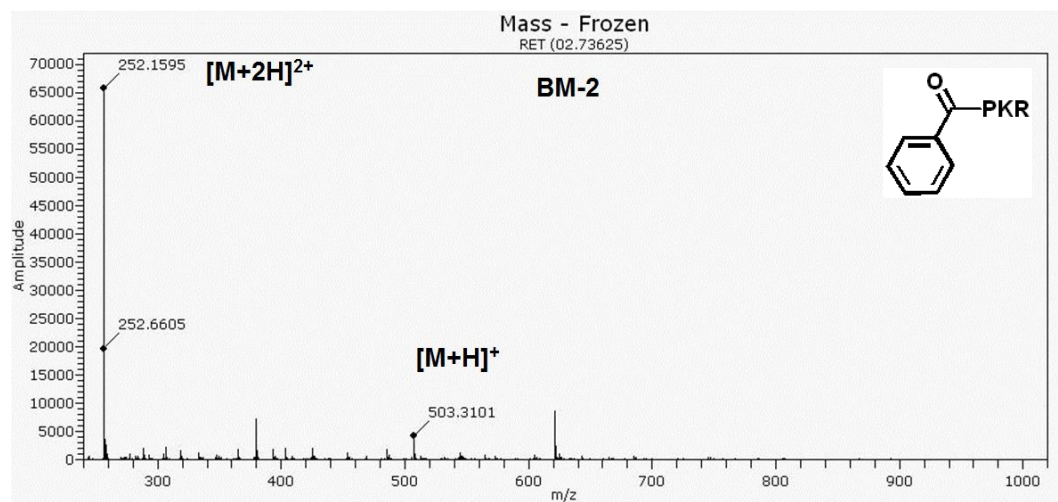
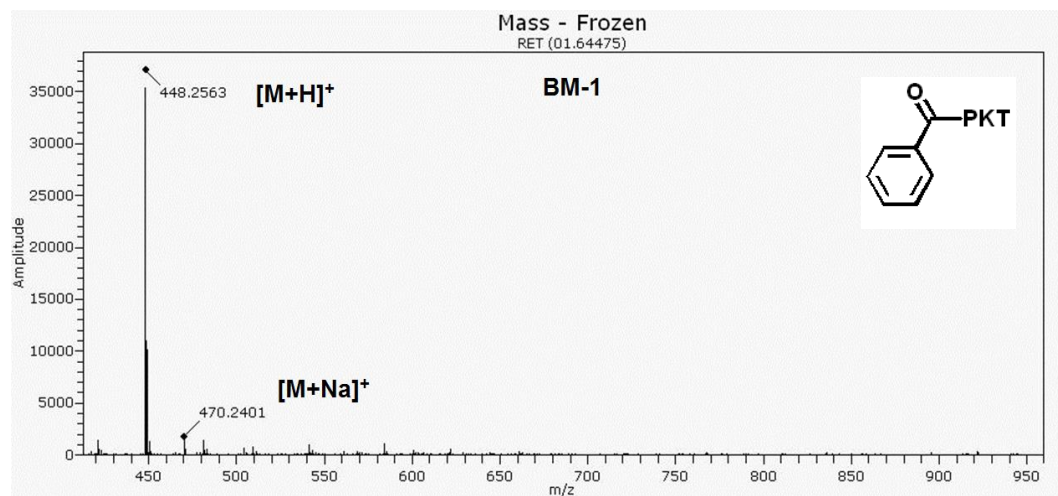
## Appendix

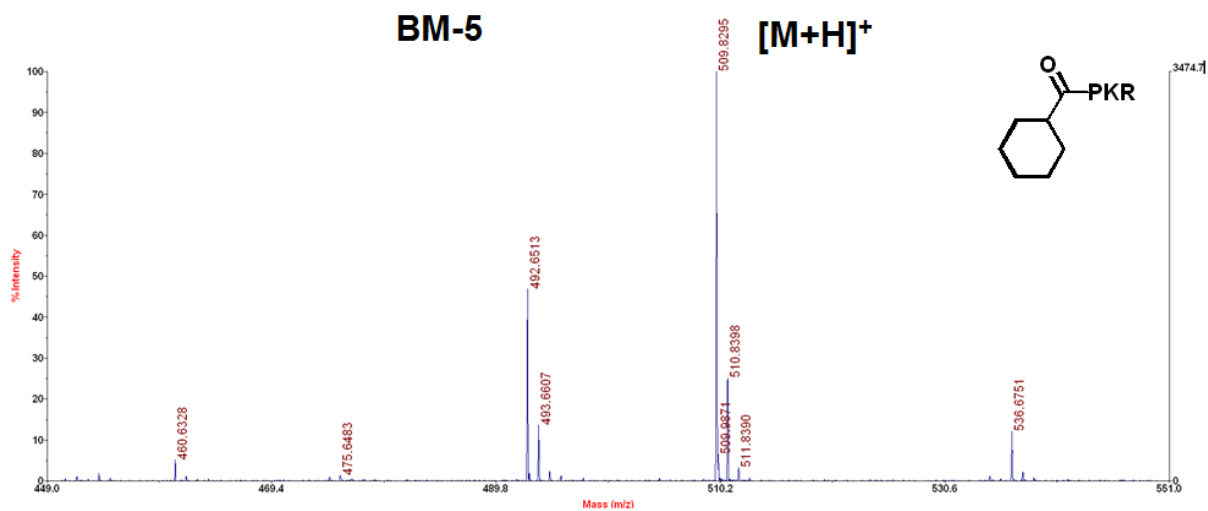
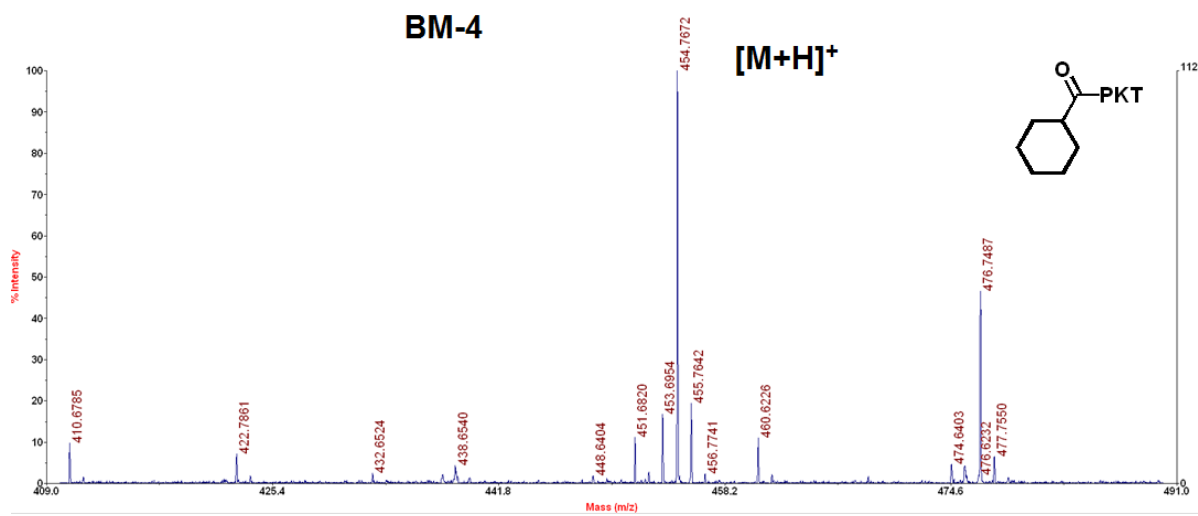
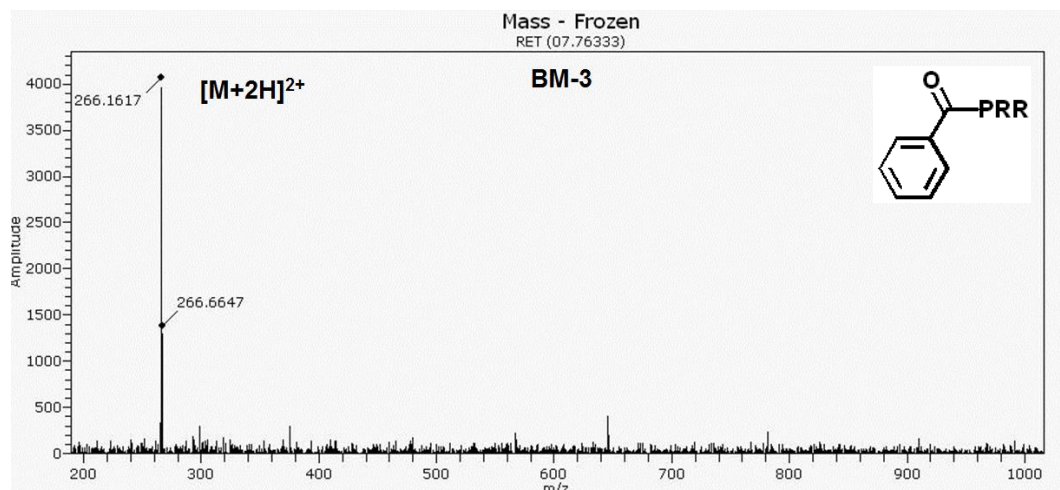


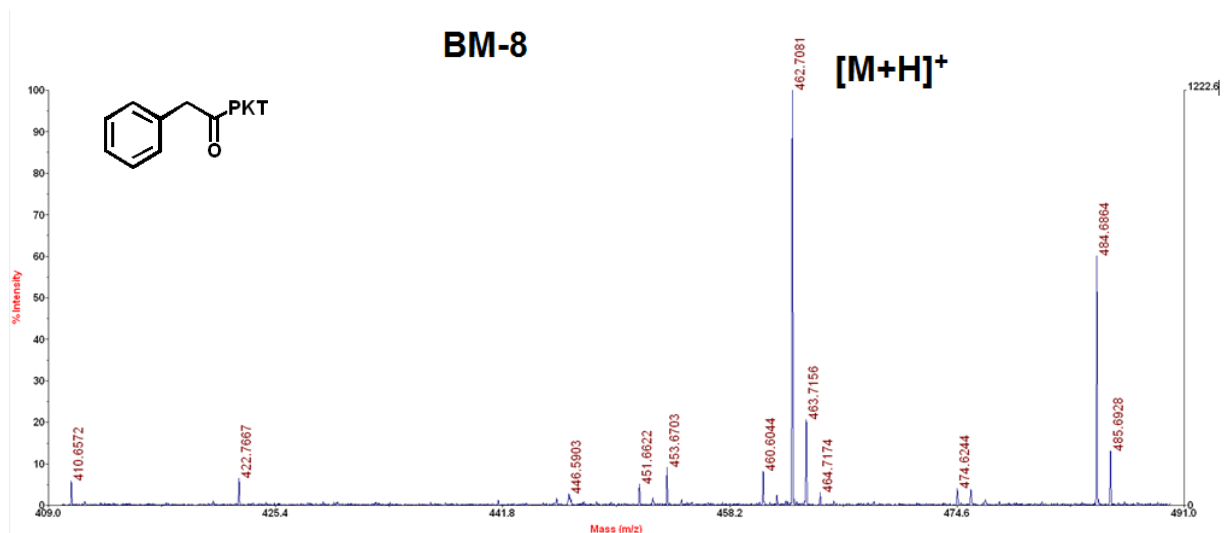
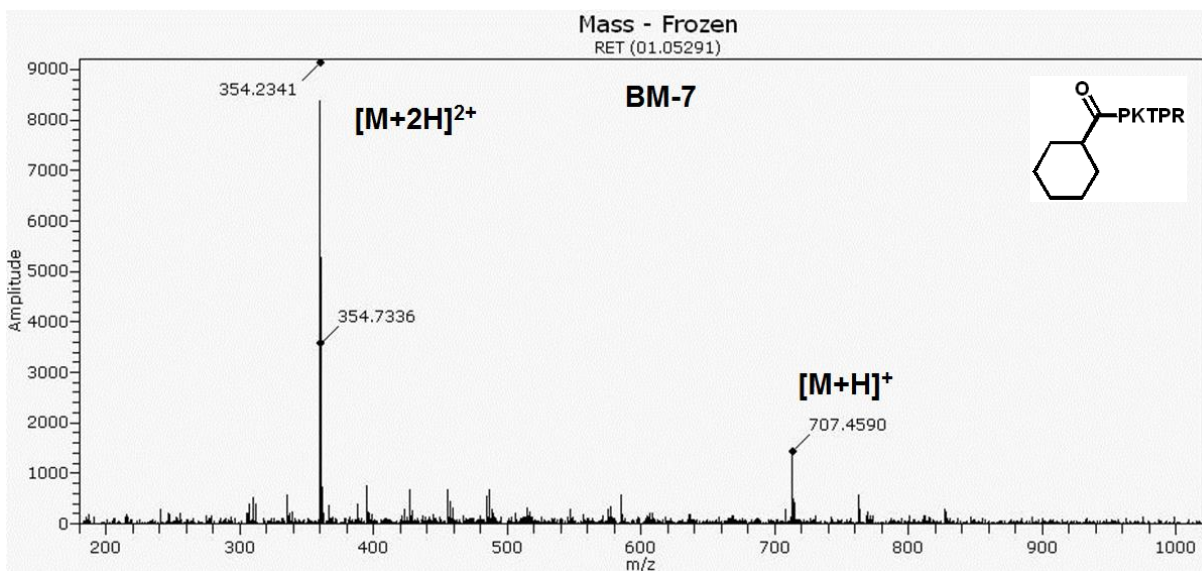
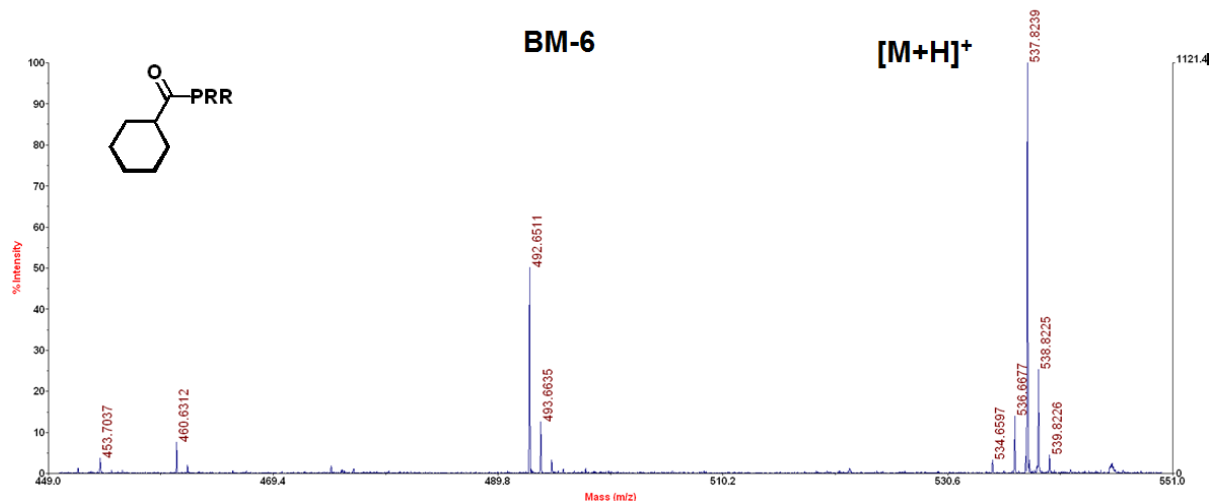
**Scheme A1.** Synthesis of BM-45 **1** a) piperidine, HOBt, DMF, rink amide resin, 2x10 min; Fmoc-Lys(alloc)-OH, HOBt, DIC, DMF, 12 hr; b)  $I_2$ , ACCN,  $H_2O$ , 48hr; c) MeI,  $K_2CO_3$ , 18-crown-6, DMF, 12 hr; d) automated iterative synthesis,  $\mu W$ , Lys, Pro, 4-hydroxyphenylacetic acid; e) trifluoroacetic acid, 2,2'-(ethylenedioxy)diethanethiol, Triisopropylsilane,  $H_2O$ , 4 hr.

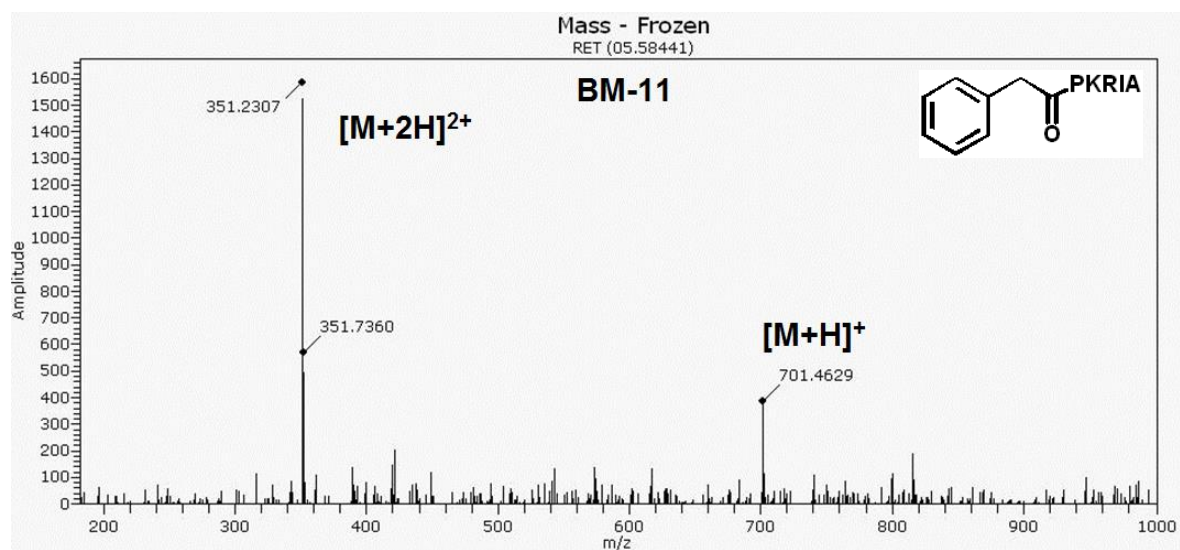
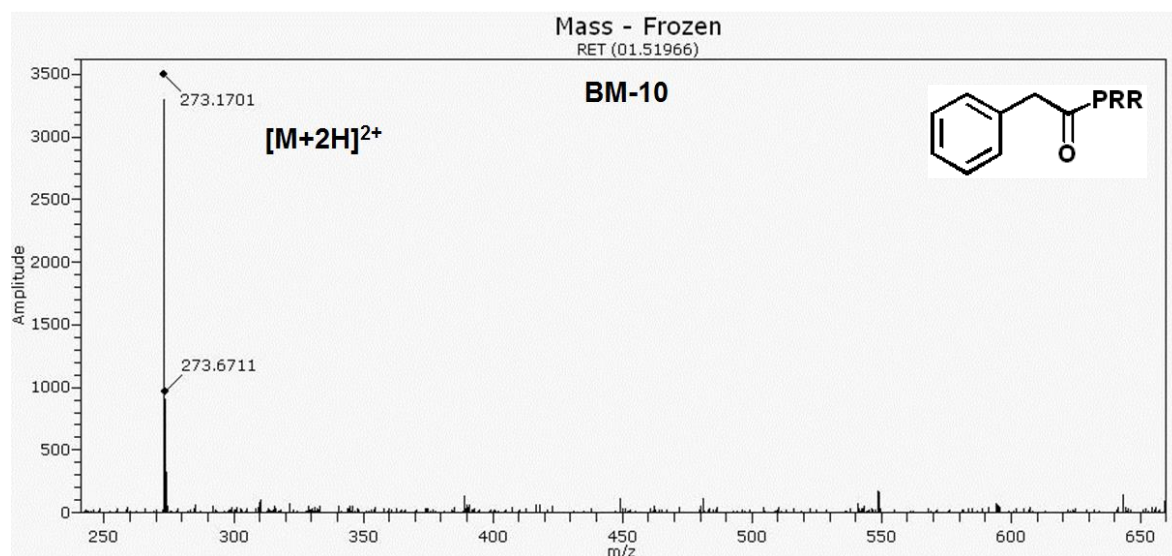
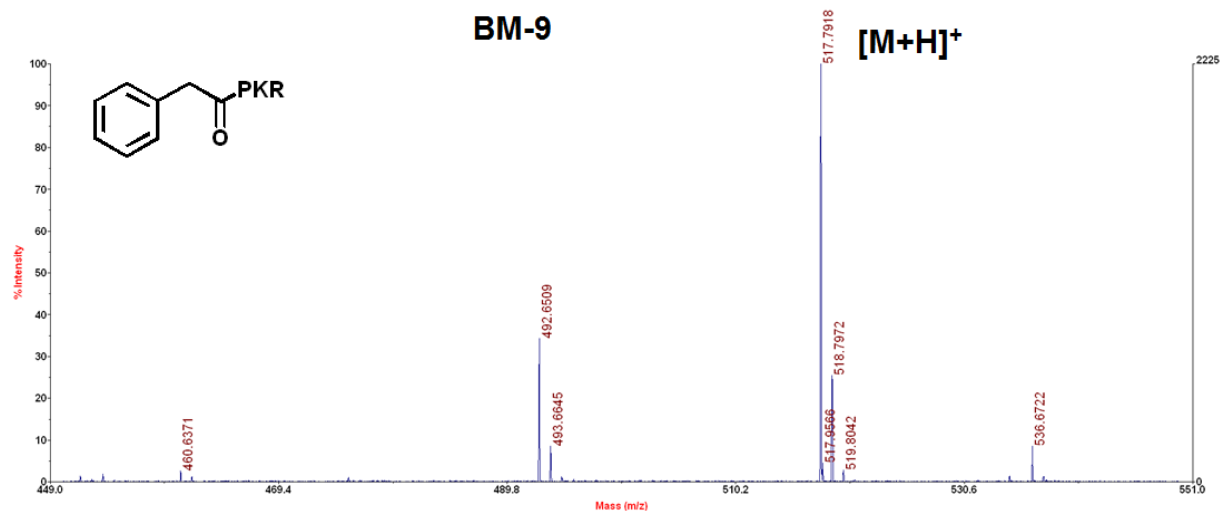


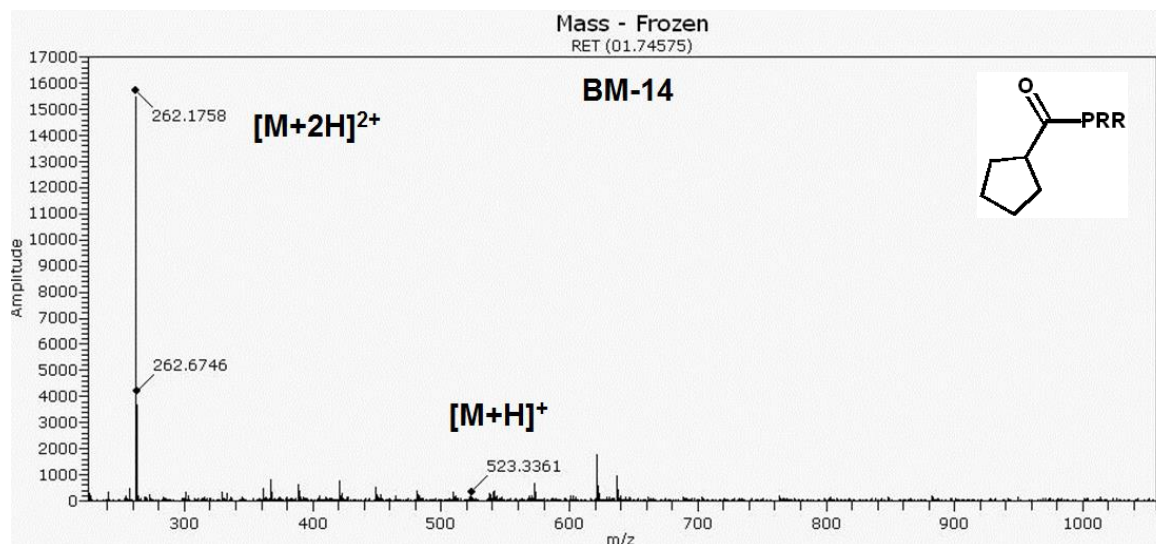
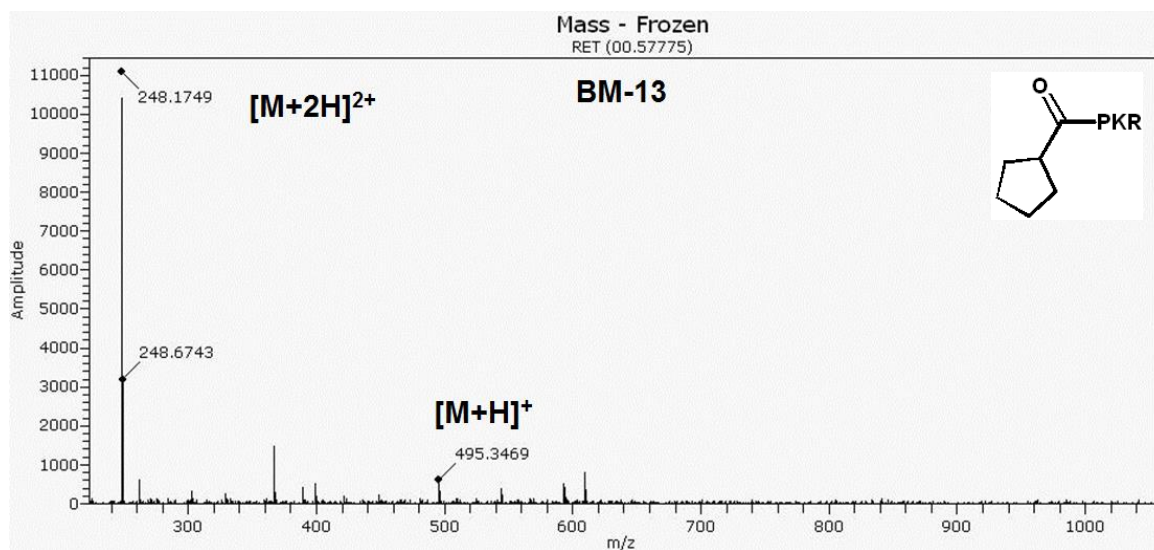
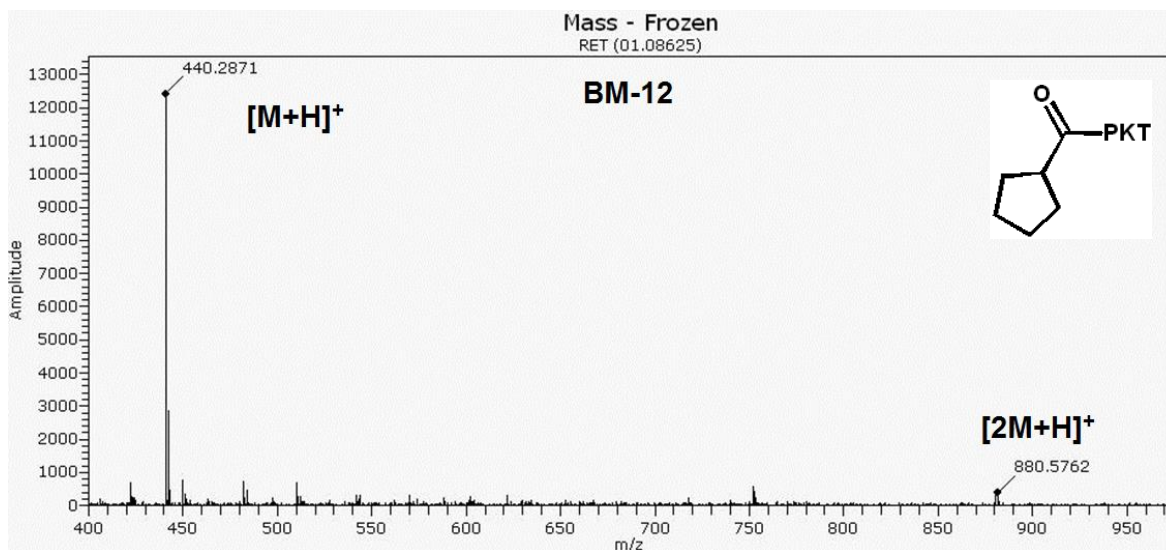
**Scheme A2.** Synthesis of BM-26 **1** a) automated iterative synthesis,  $\mu W$ , Arg, Lys, Pro; b) piperidine, HOBt, DMF, rink amide resin, 2x10 min; c) benzyl chloride,  $CS_2CO_3$ , DMF, 12 hr d) trifluoroacetic acid, 2,2'-(ethylenedioxy)diethanethiol, Triisopropylsilane,  $H_2O$ , 4 hr.

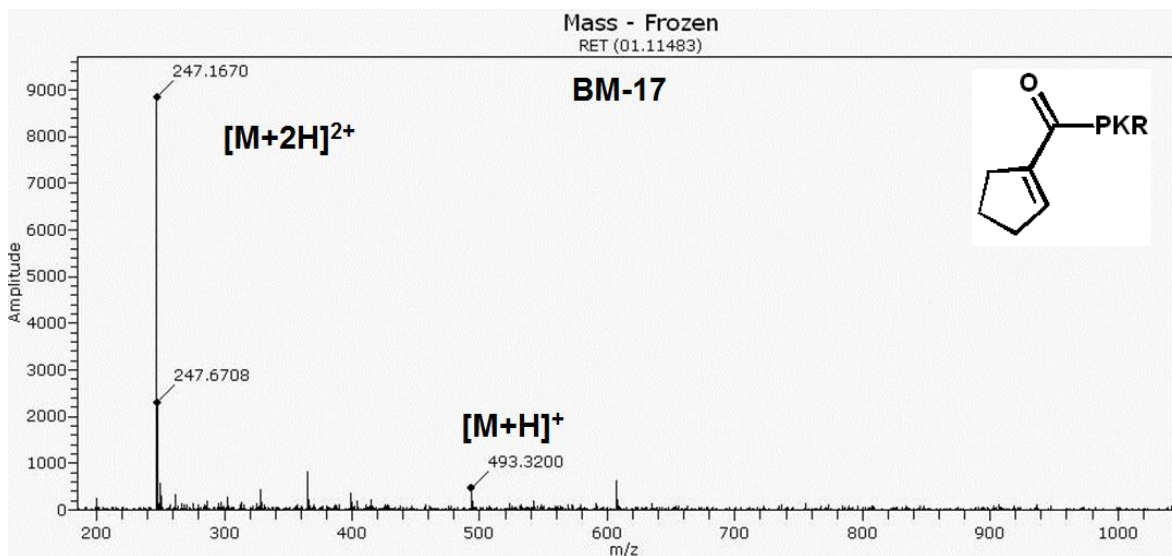
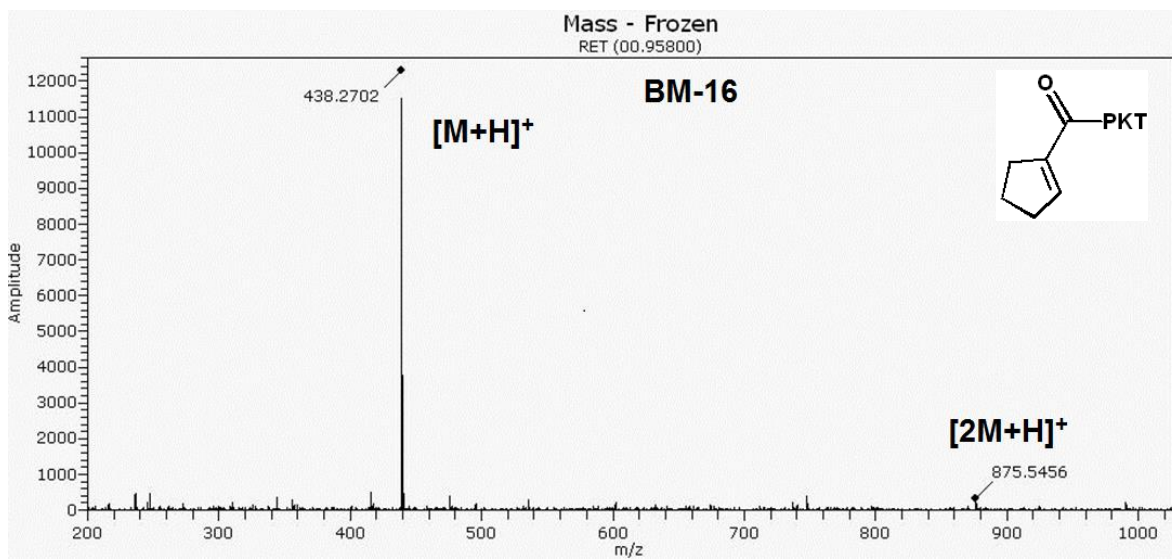
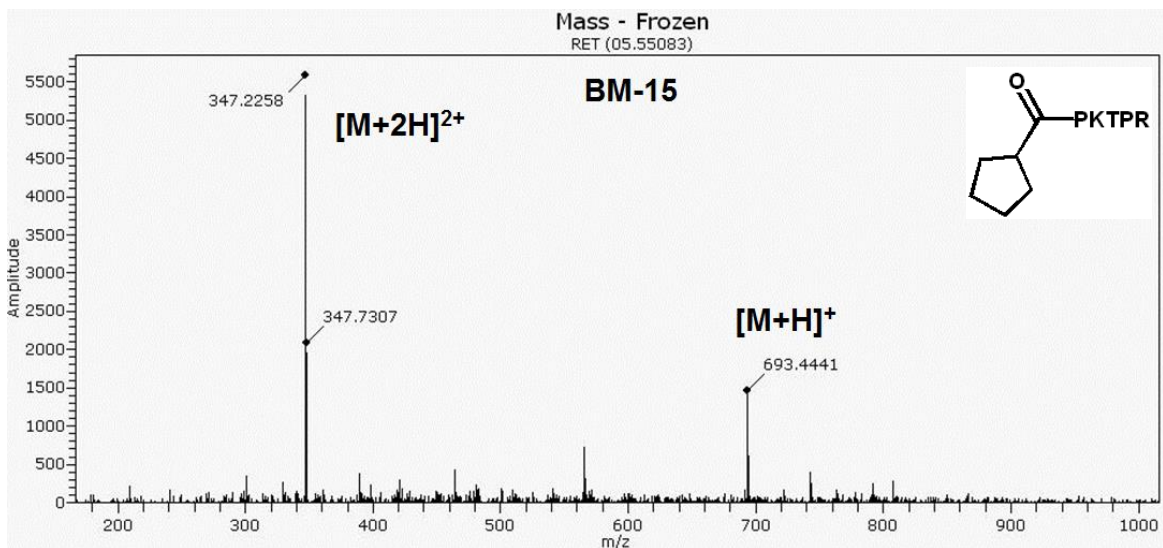




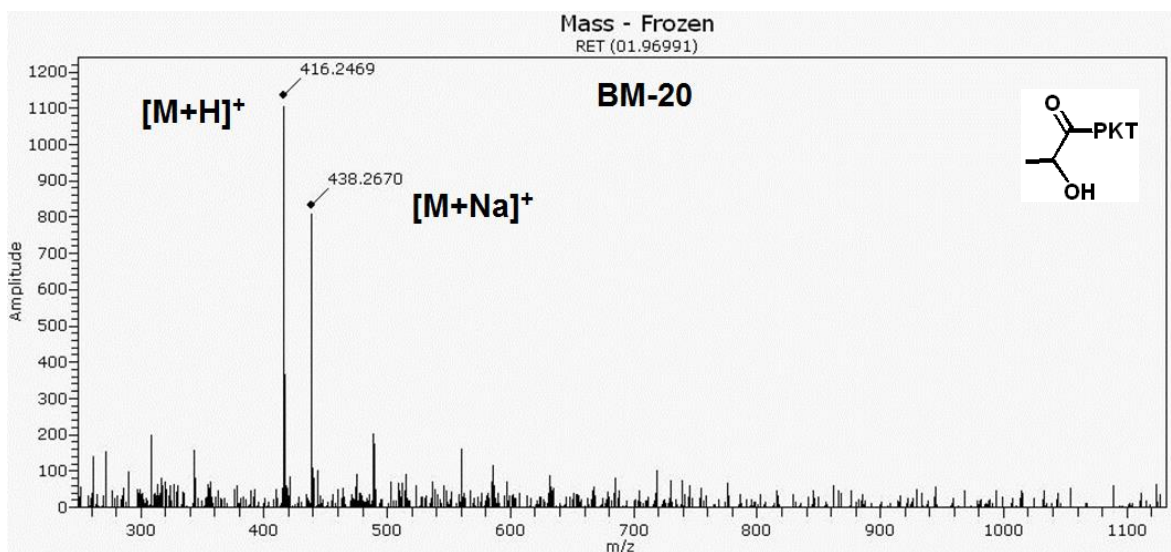
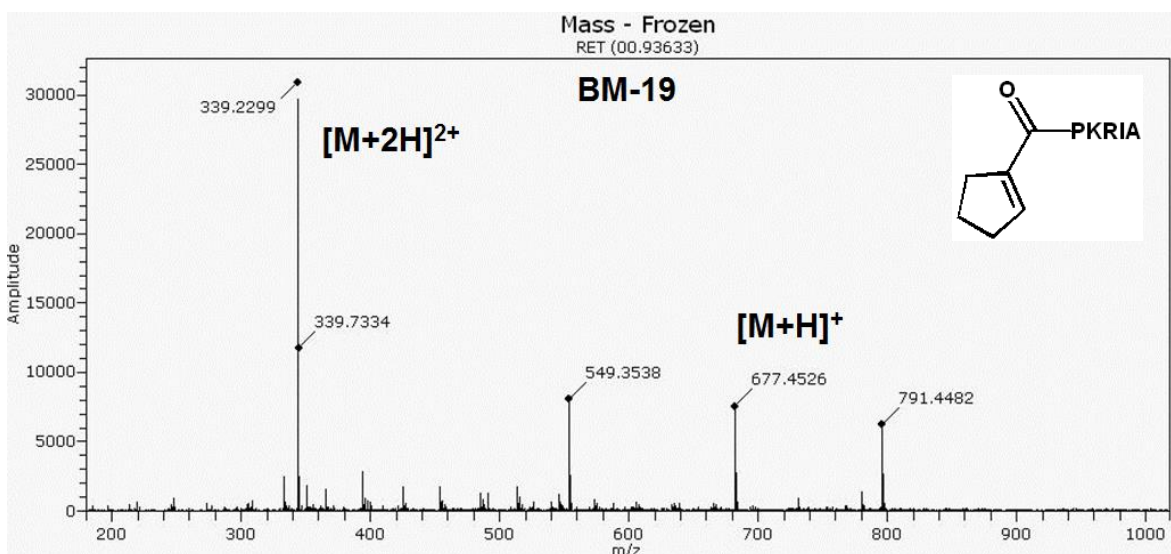
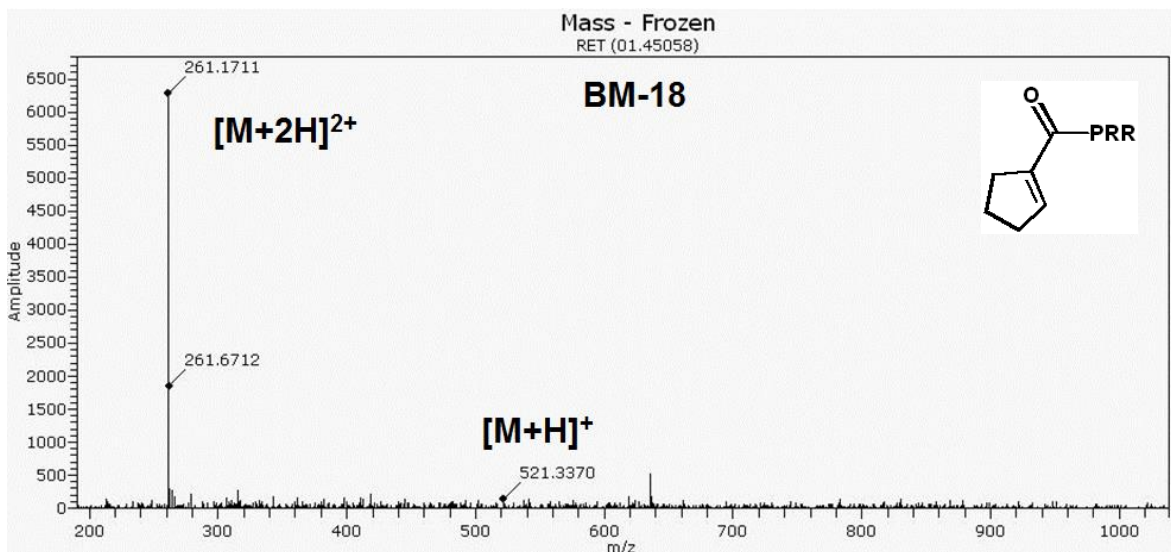


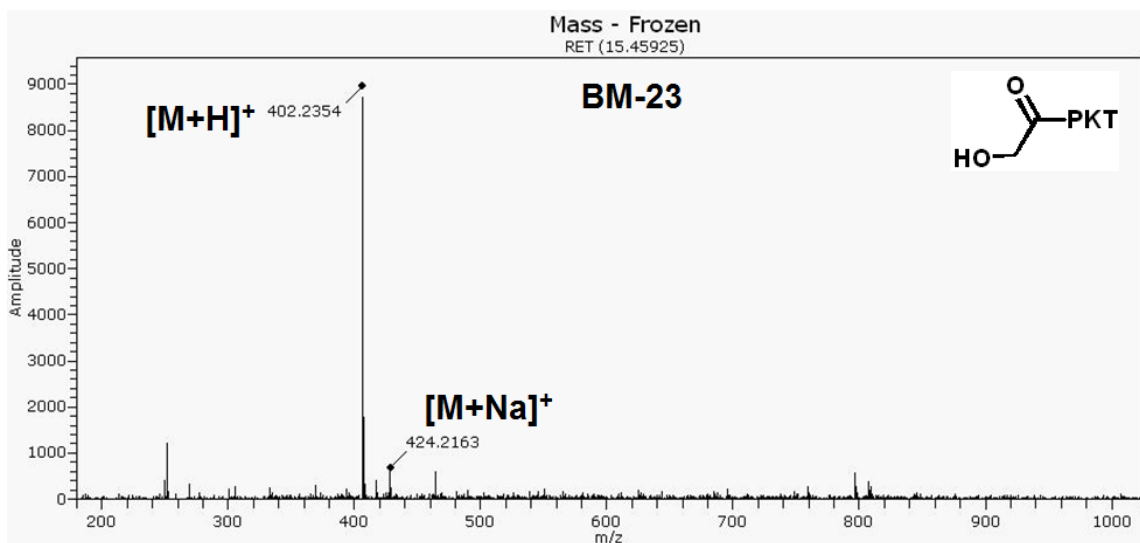
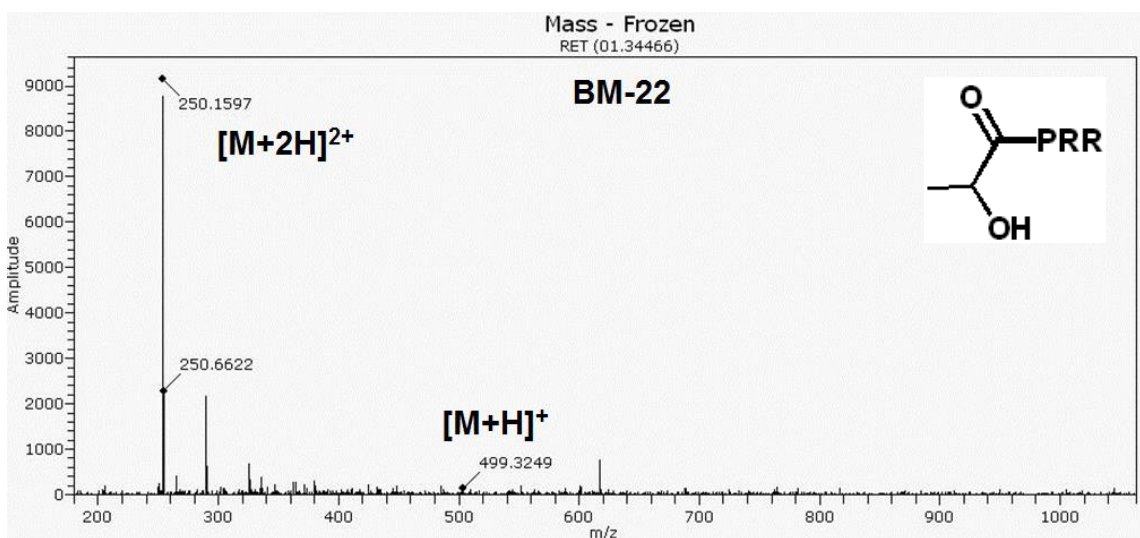
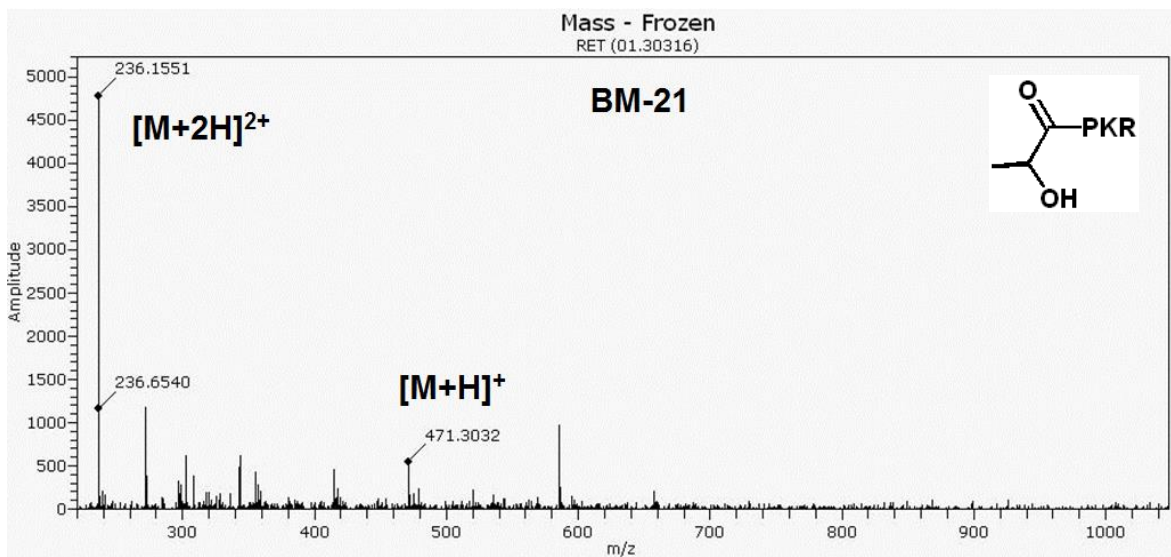


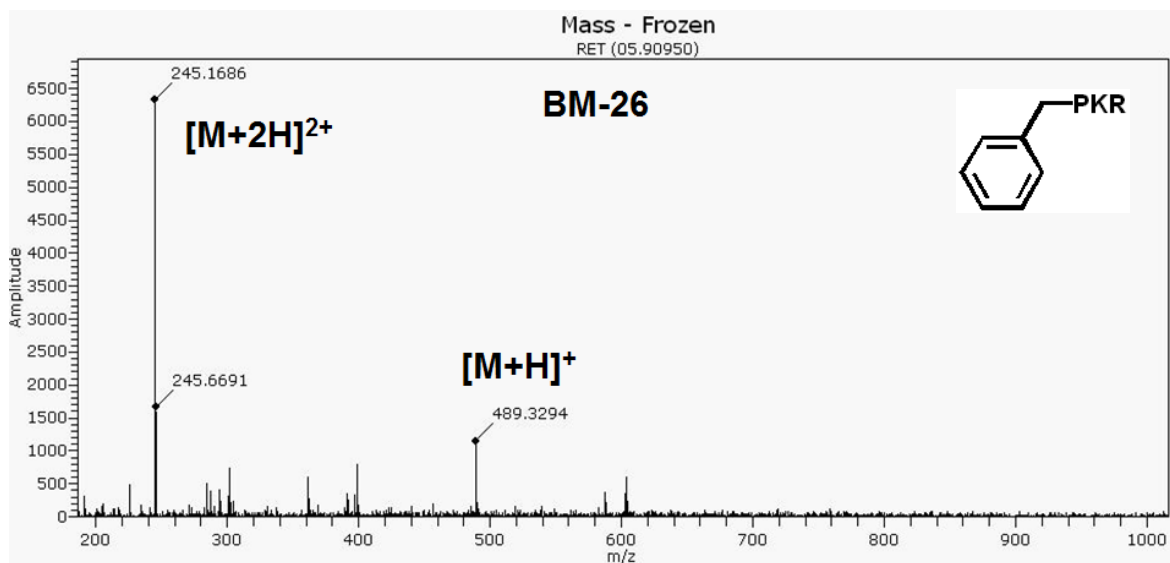
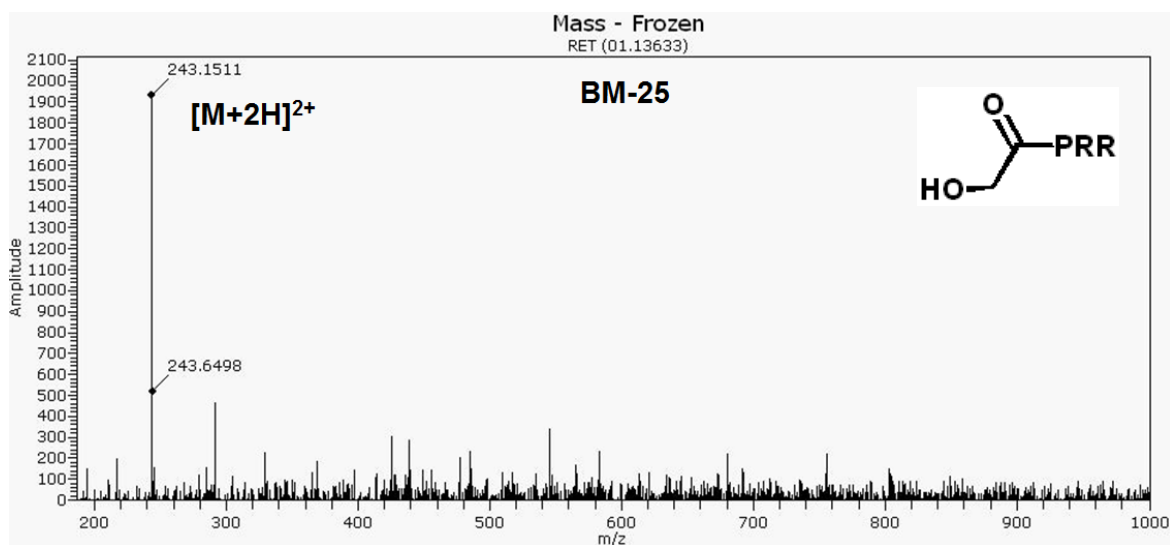
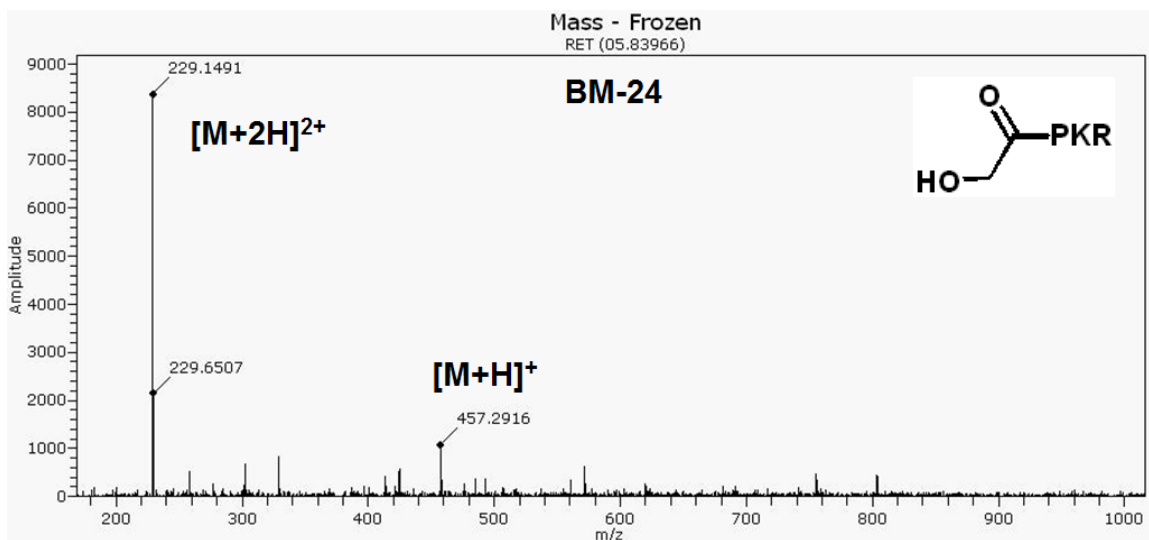


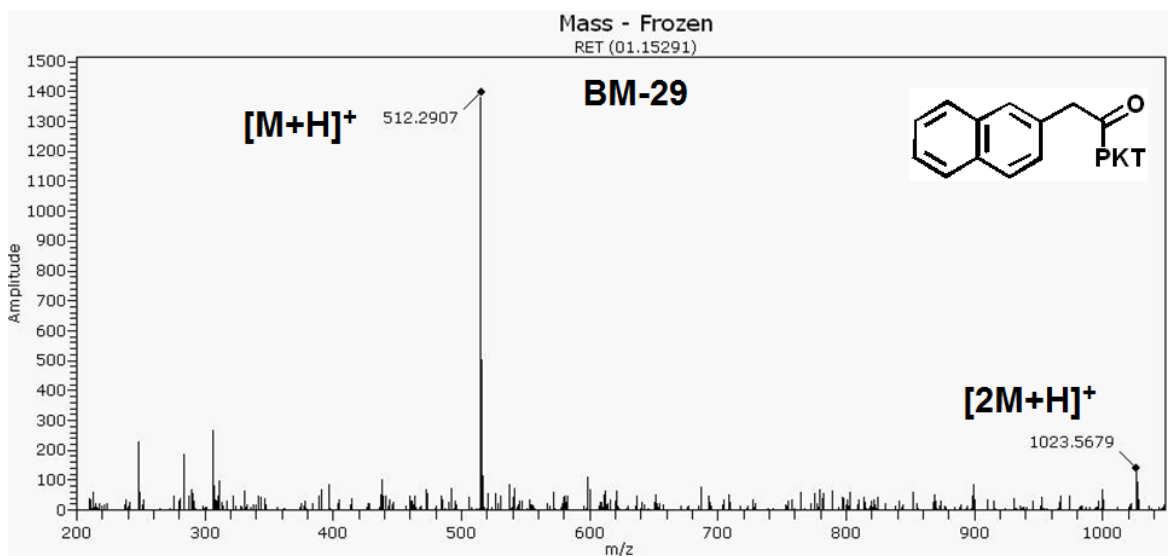
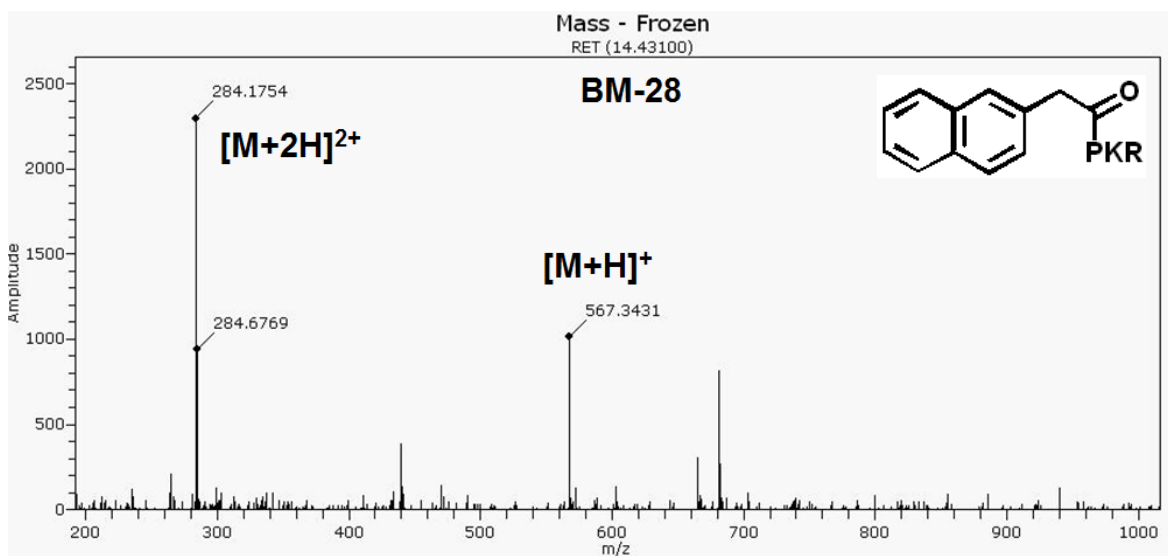
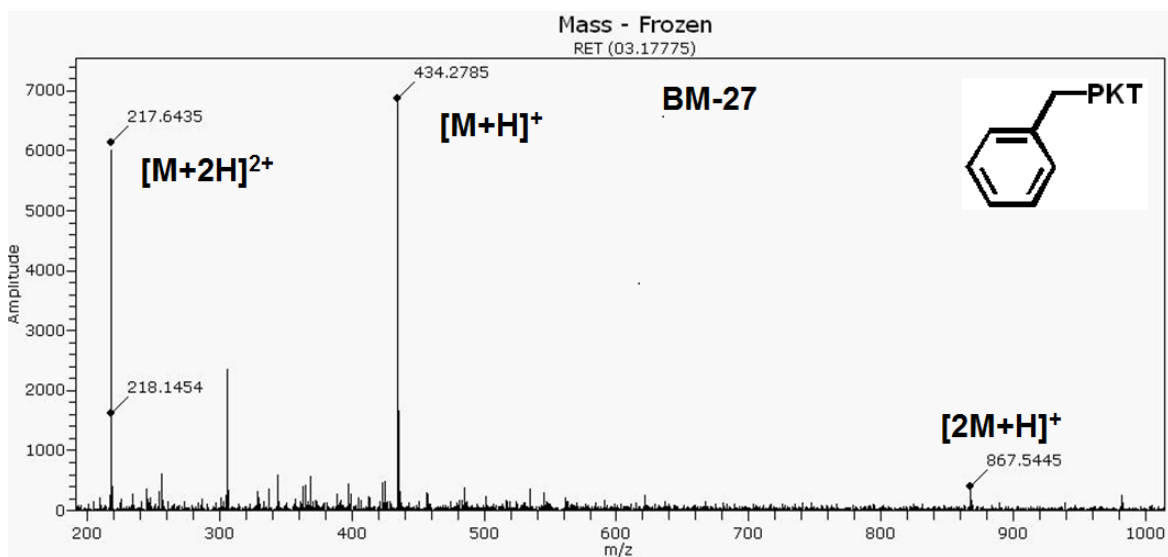


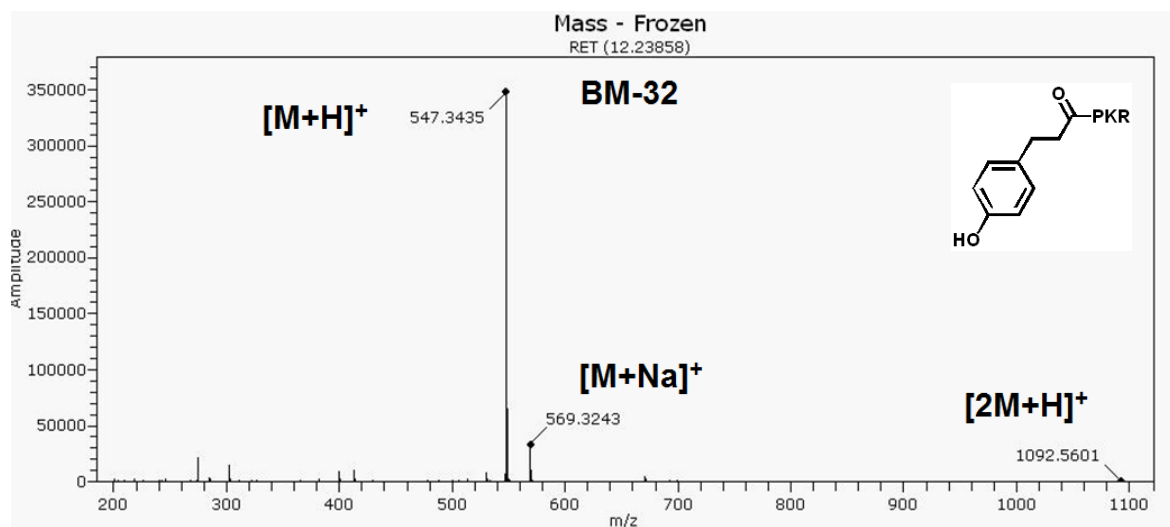
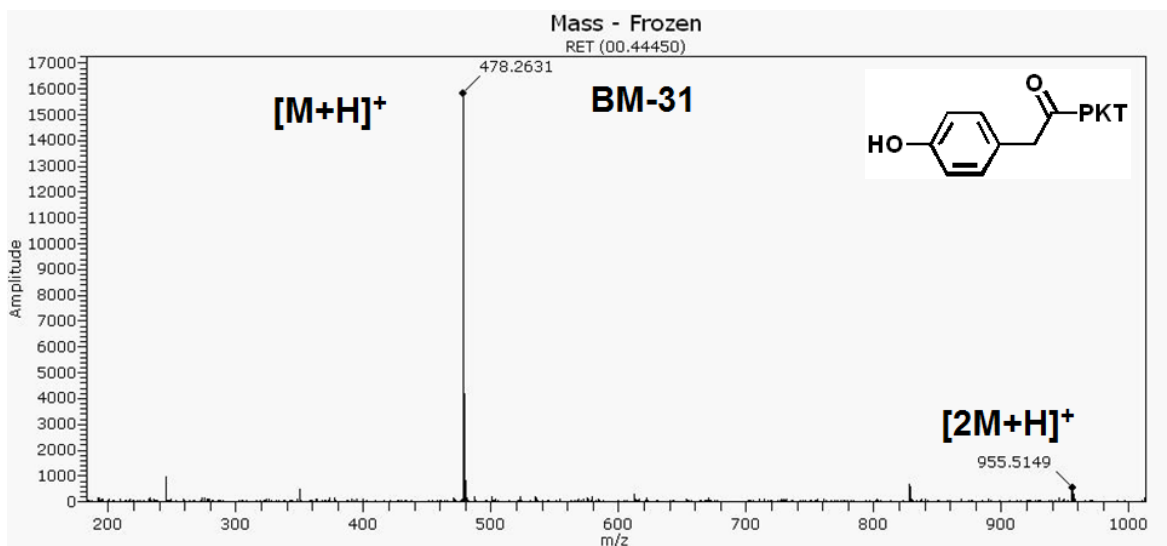
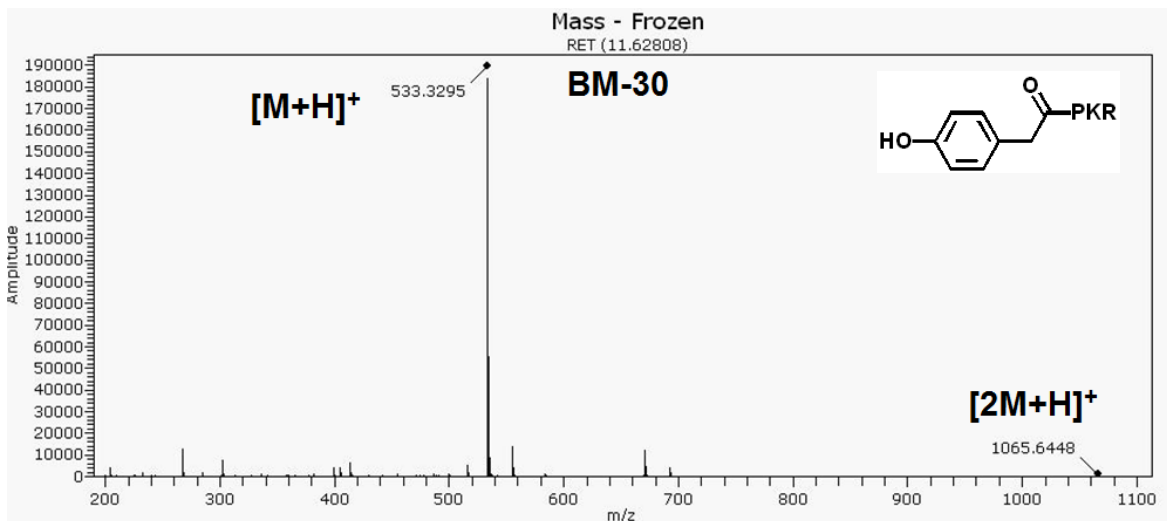


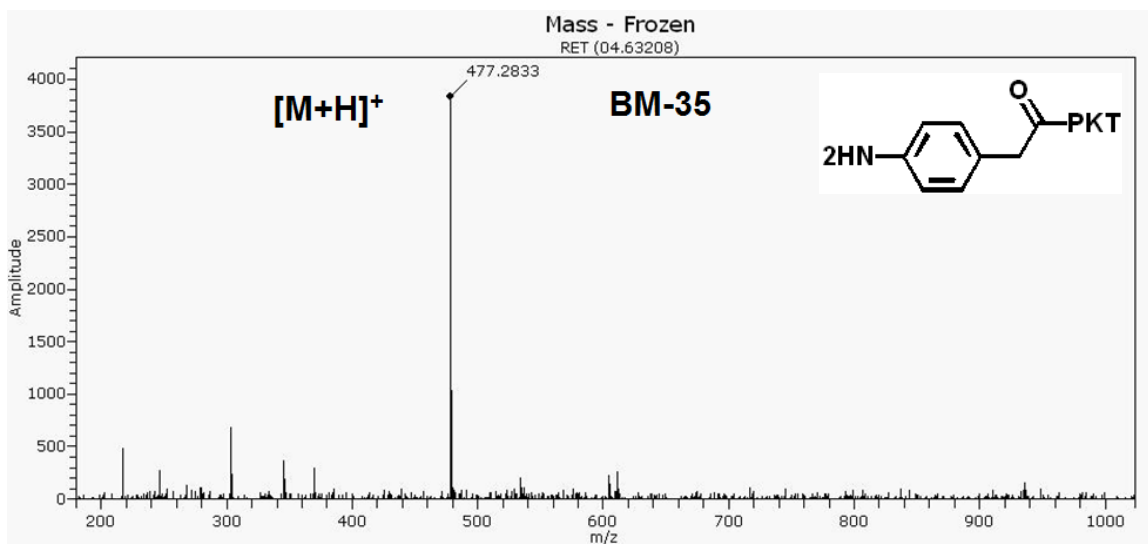
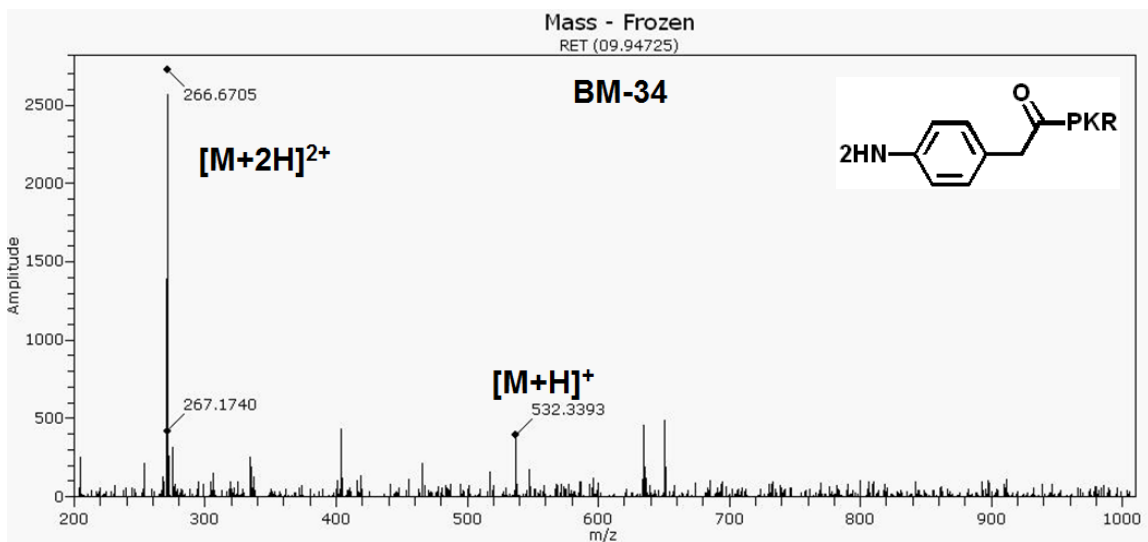
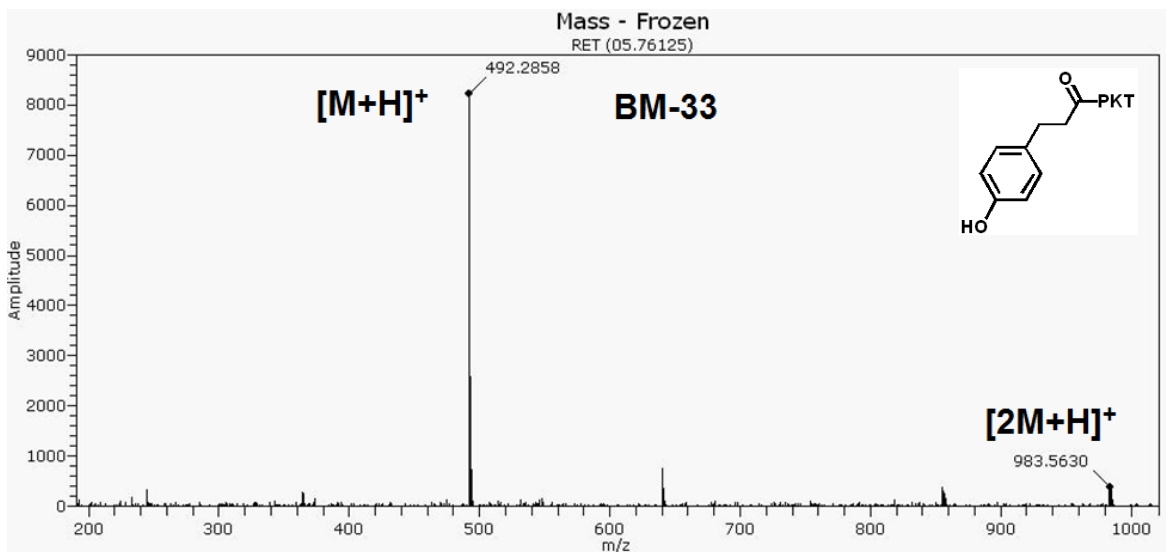


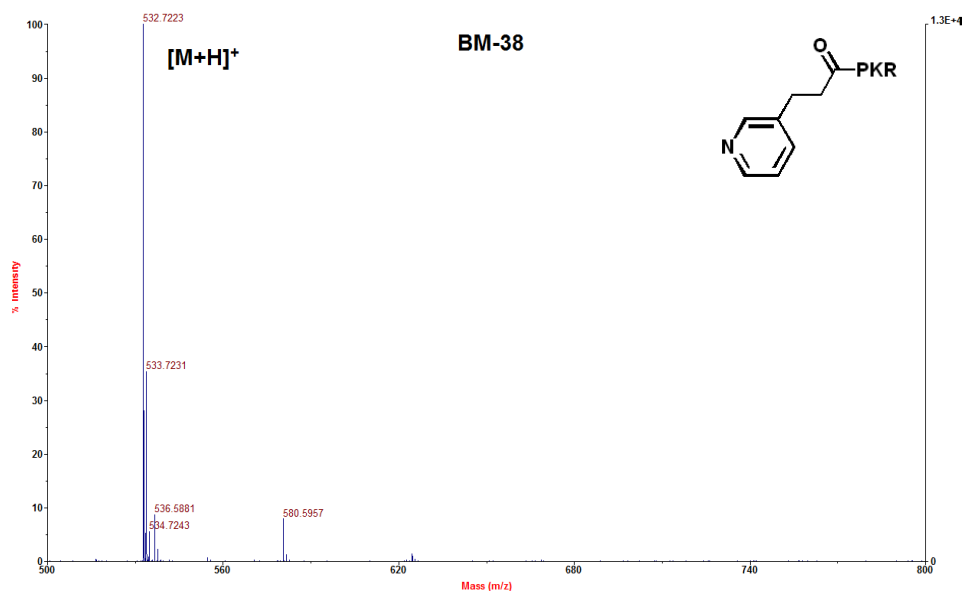
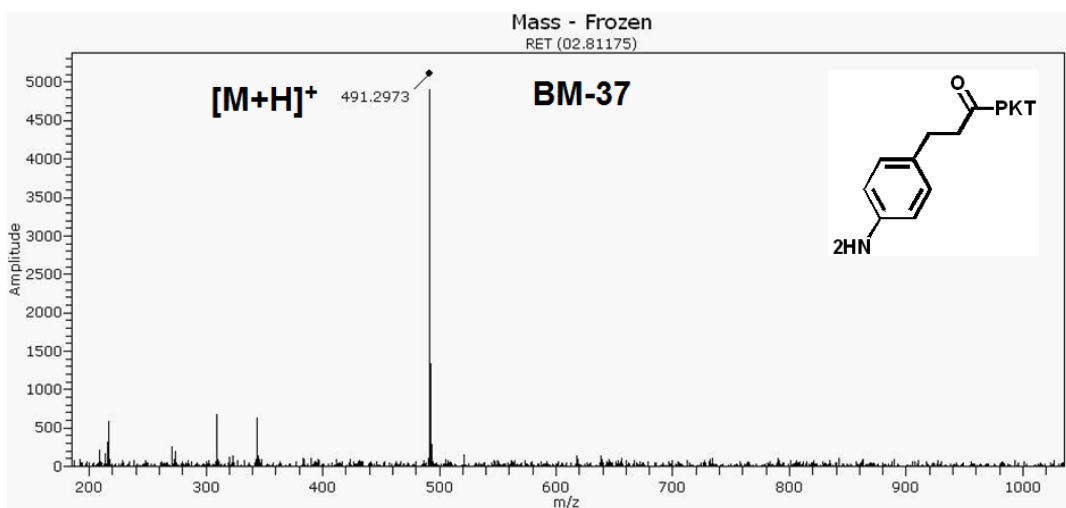
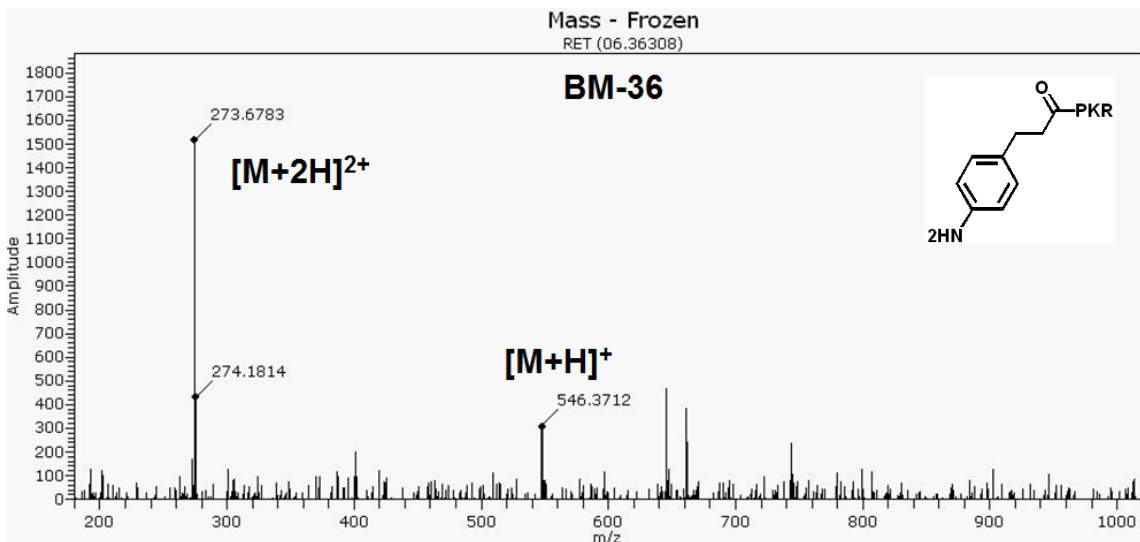


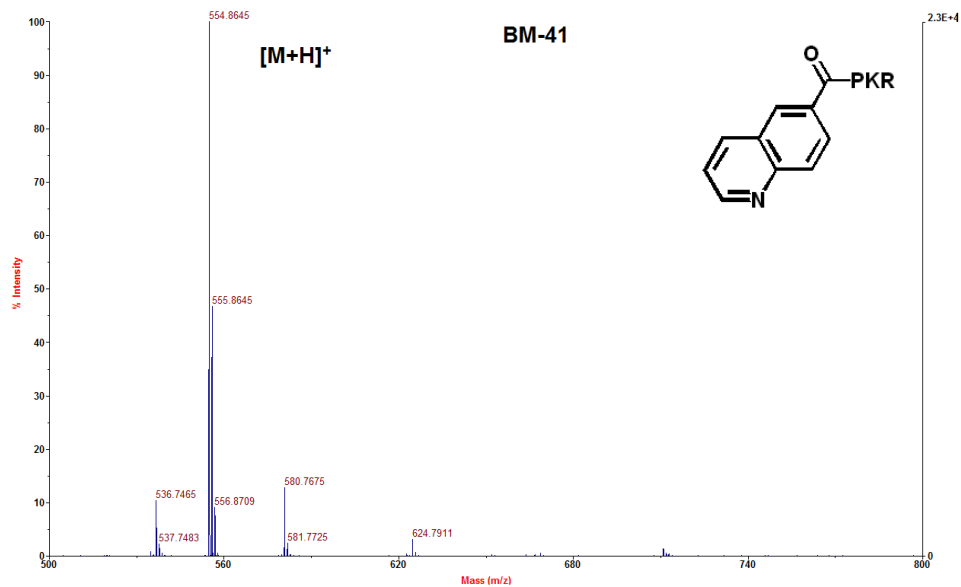
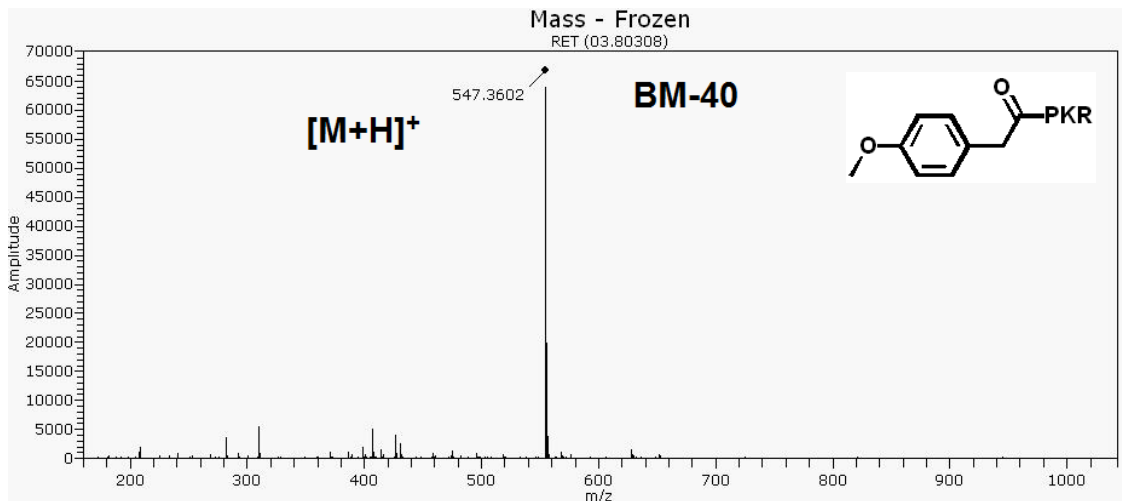
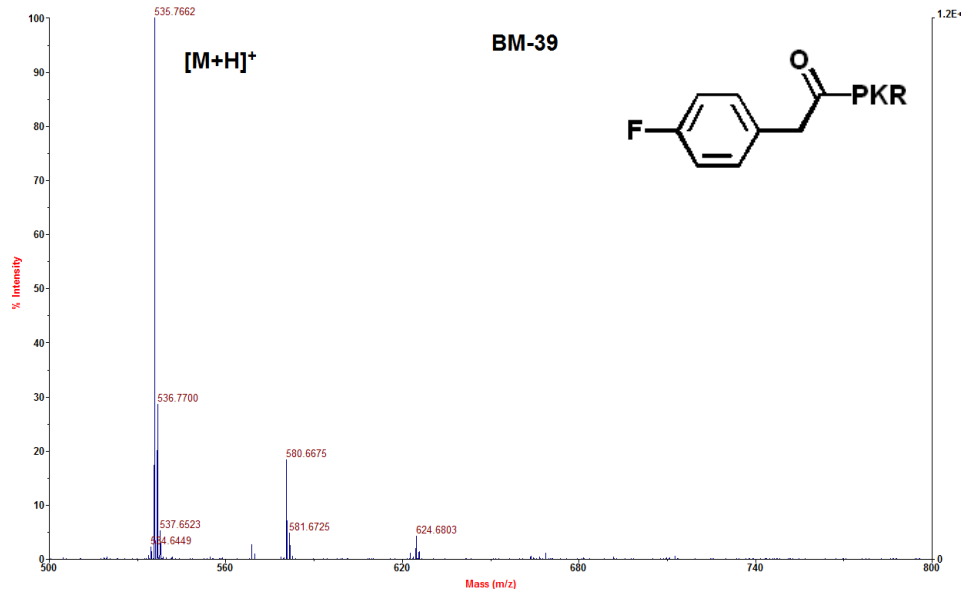




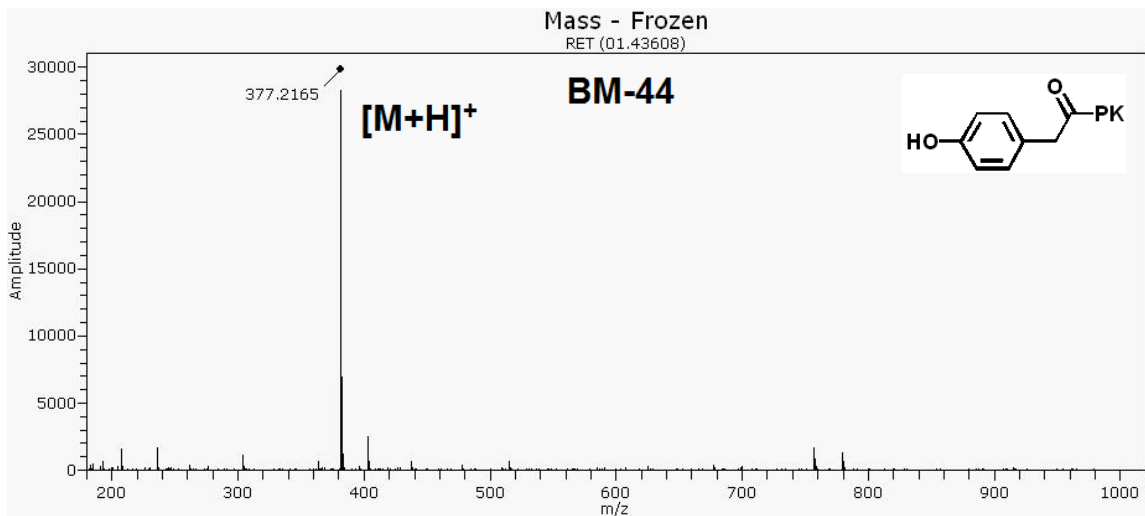
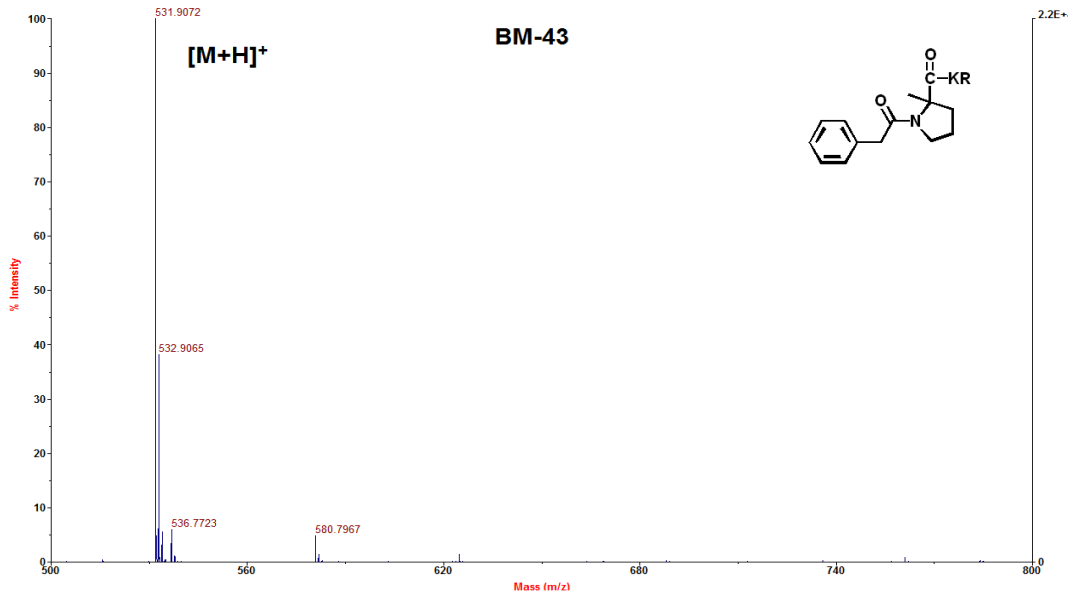
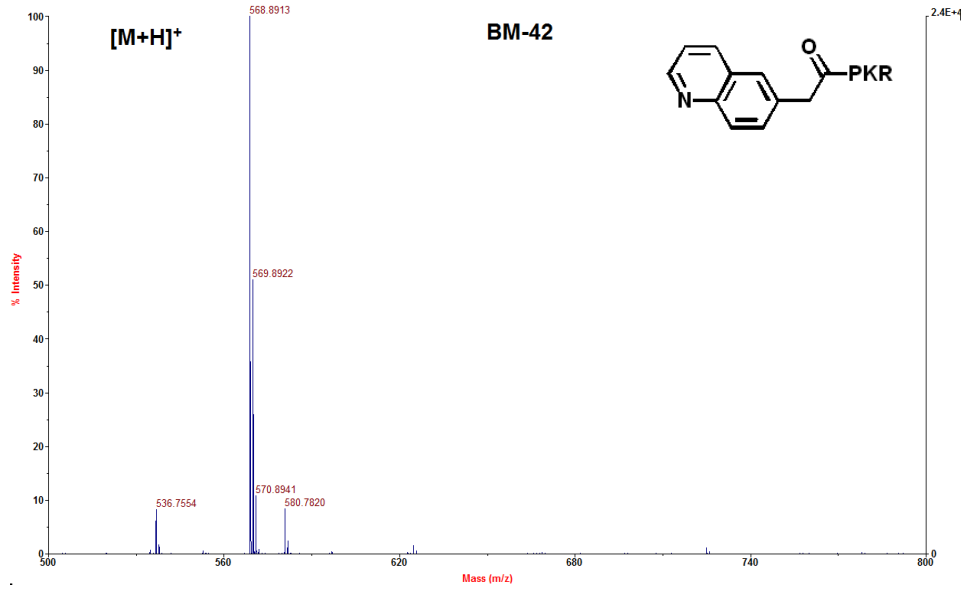












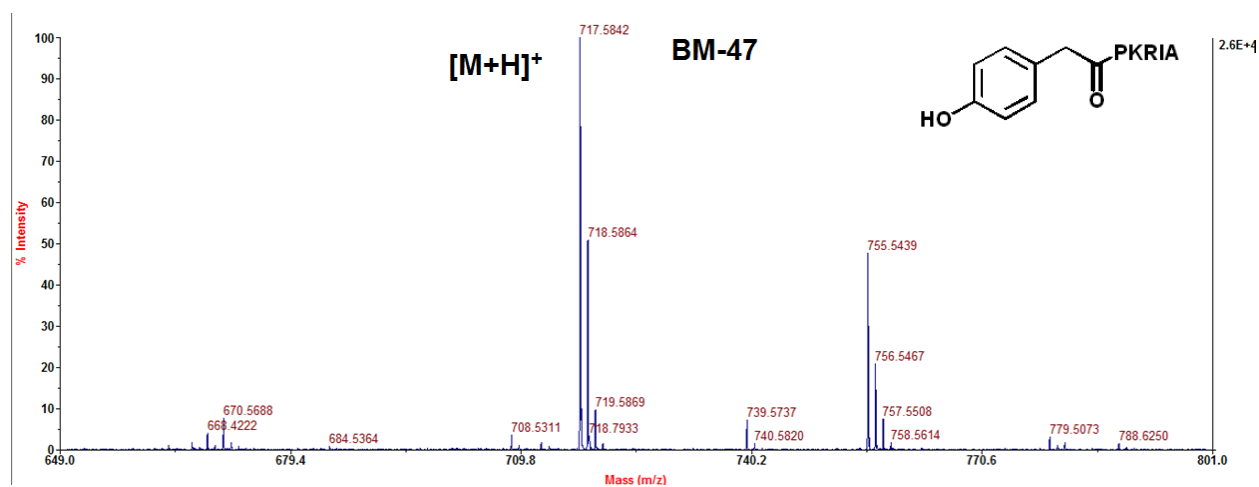
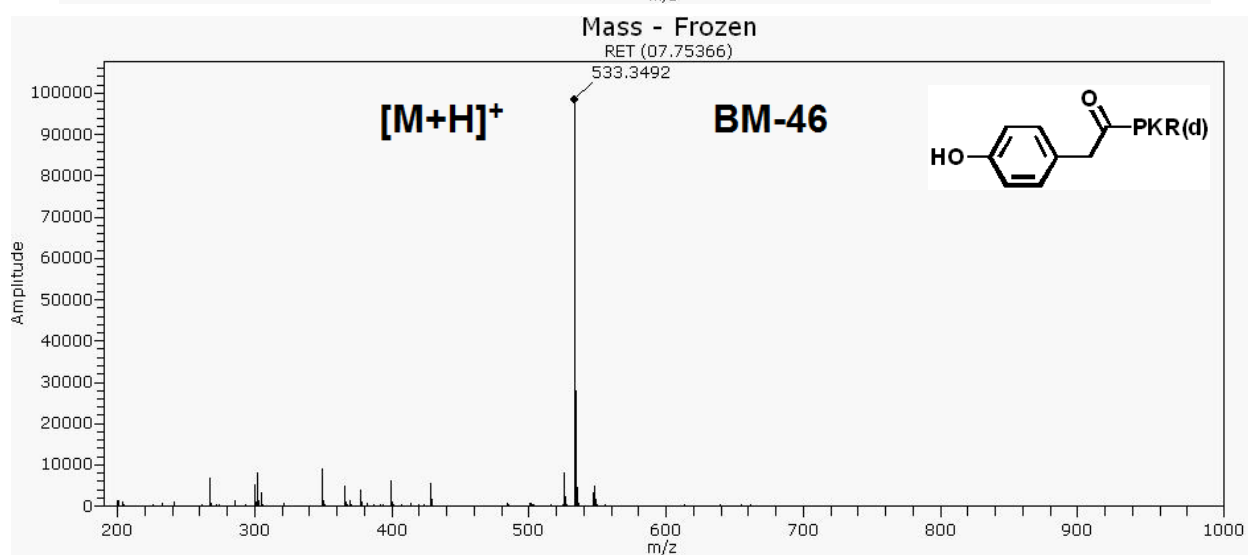
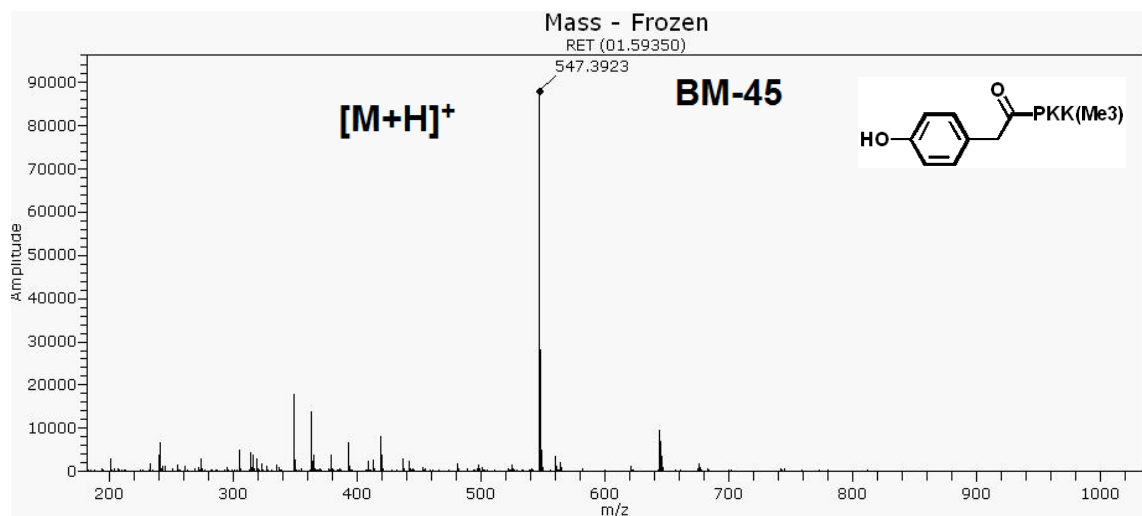


Figure A1. MALDI-MS and ESI-TOF-MS characterization of all peptide inhibitors

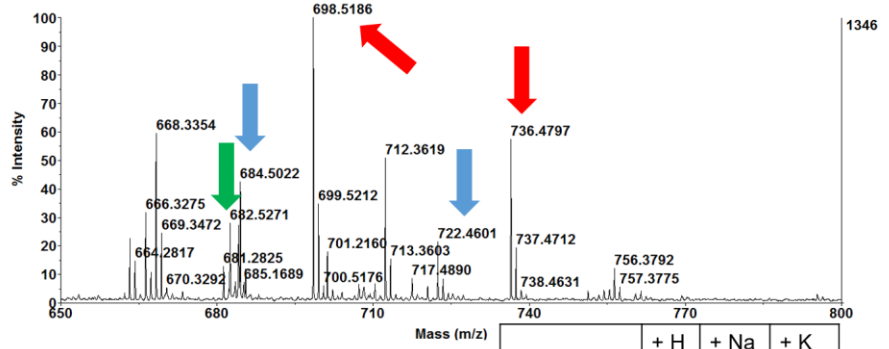
SPKRIA + NTMT1 + SAM  
+ 1uM BM-47

**MALDI-MS assay:**

6% inhibition

**Fluorescence assay:**

40% inhibition

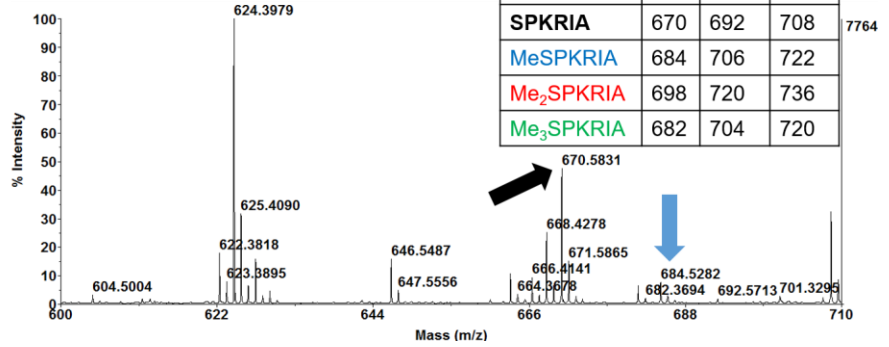


SPKRIA + NTMT1 + SAM  
+ 25uM BM-47

**MALDI-MS assay:** 91%  
inhibition

**Fluorescence assay:**

88% inhibition



**Figure A2.** MALDI-MS methylation inhibition assay for peptide inhibitor BM-47

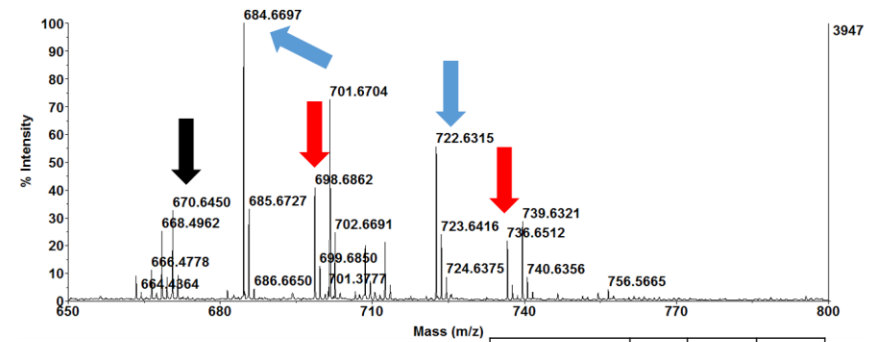
SPKRIA + NTMT1 + SAM  
+ 1uM BM-11

**MALDI-MS assay:**

19% inhibition

**Fluorescence assay:**

40% inhibition



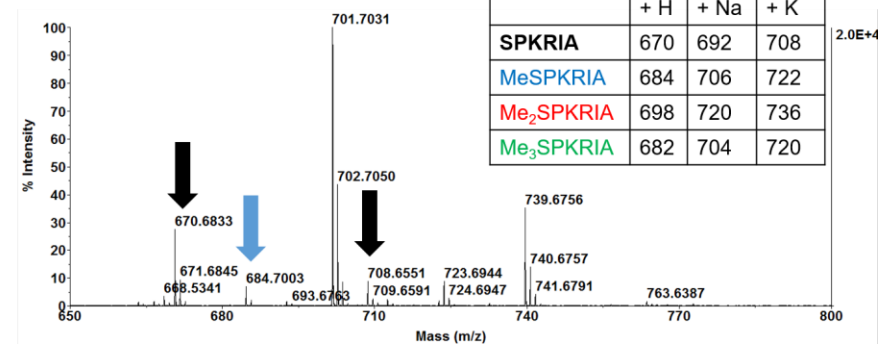
SPKRIA + NTMT1 + SAM  
+ 25uM BM-11

**MALDI-MS assay:**

82% inhibition

**Fluorescence assay:**

75% inhibition



**Figure A3.** MALDI-MS methylation inhibition assay for peptide inhibitor BM-11

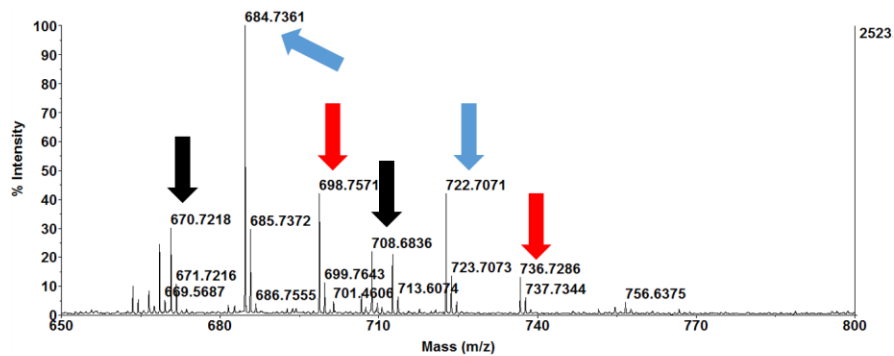
SPKRIA + NTMT1 + SAM  
+ 1uM BM-46

**MALDI-MS assay:**

22% inhibition

**Fluorescence assay:**

12% inhibition



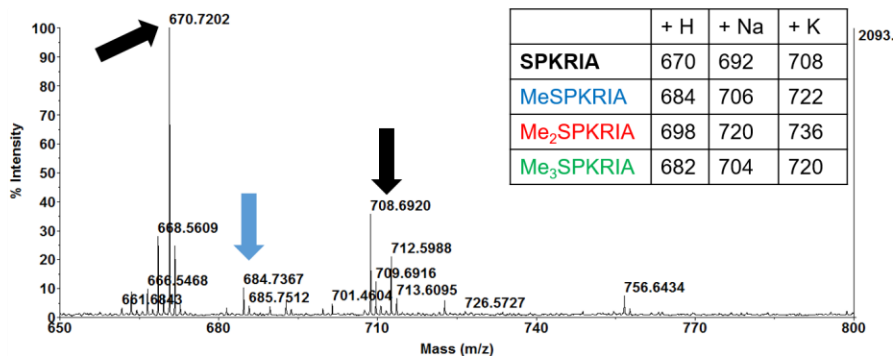
SPKRIA + NTMT1 + SAM  
+ 25uM BM-46

**MALDI-MS assay:**

90% inhibition

**Fluorescence assay:**

85% inhibition



**Figure A4.** MALDI-MS methylation inhibition assay for peptide inhibitor BM-46

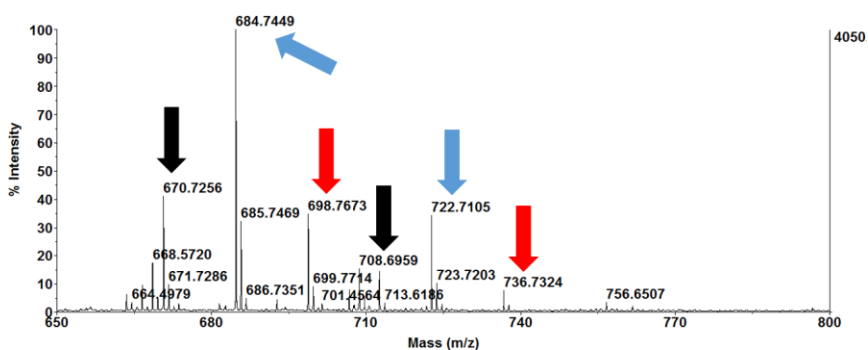
SPKRIA + NTMT1 + SAM  
+ 1uM BM-34

**MALDI-MS assay:**

26% inhibition

**Fluorescence assay:**

18% inhibition



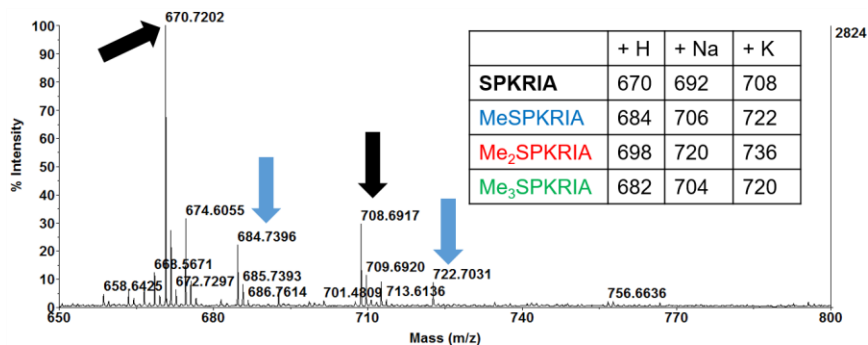
SPKRIA + NTMT1 + SAM  
+ 25uM BM-34

**MALDI-MS assay:**

81% inhibition

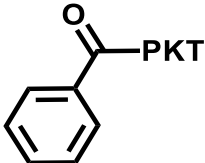
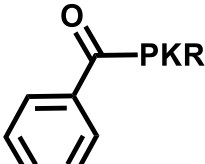
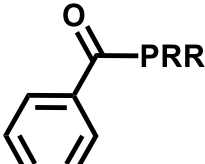
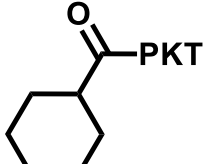
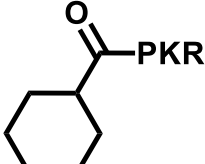
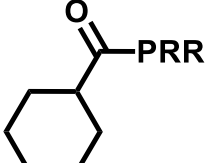
**Fluorescence assay:**

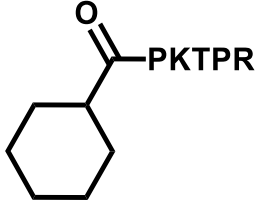
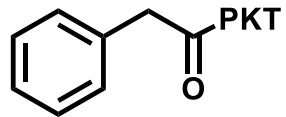
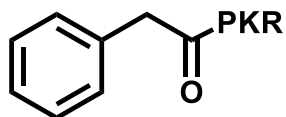
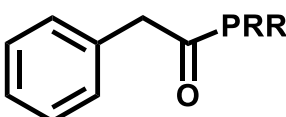
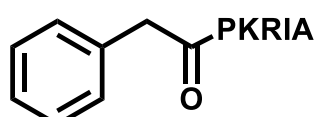
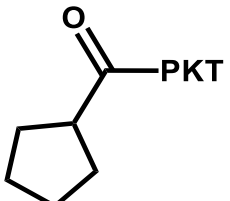
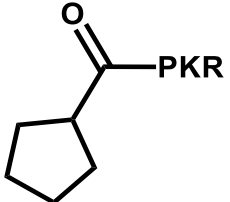
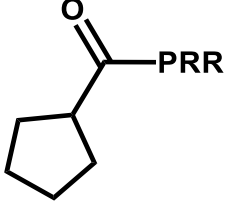
78% inhibition

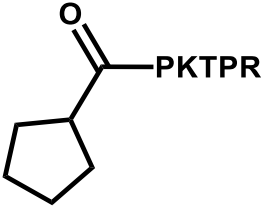
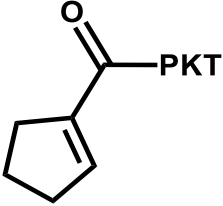
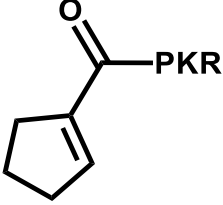
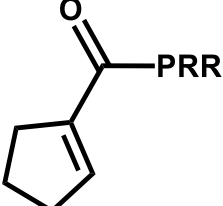
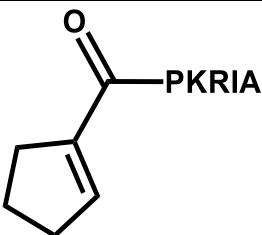
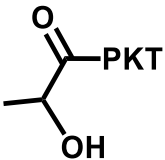
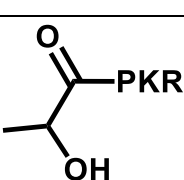
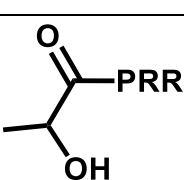


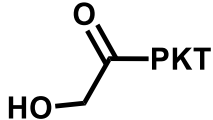
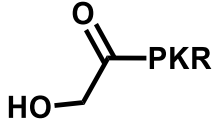
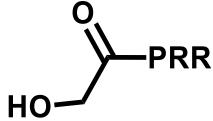
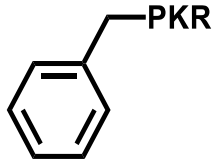
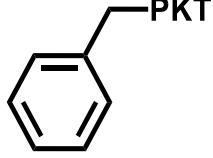
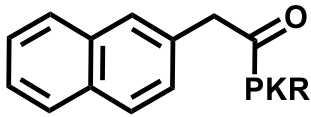
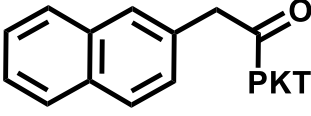
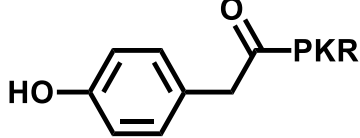
**Figure A5.** MALDI-MS methylation inhibition assay for peptide inhibitor BM-34

**Table A1.** Kinetic data for peptide inhibitors

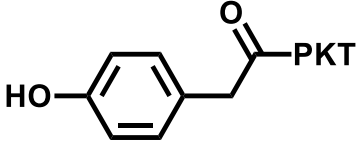
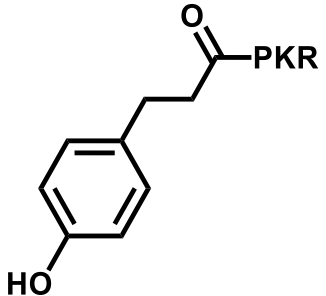
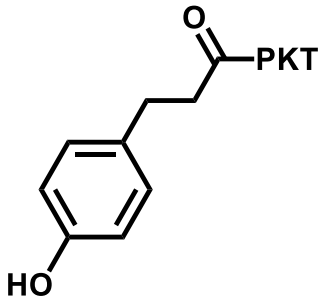
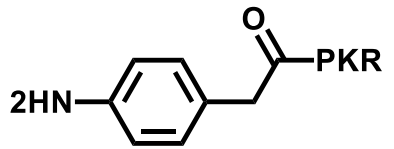
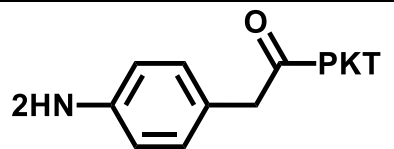
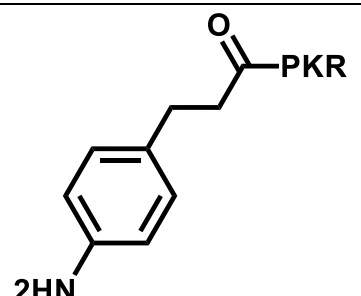
Name	Structure	IC <sub>50</sub> (μM)		
		NTMT1	G9a	PRMT1
BM-1		>100		
BM-2		>100		
BM-3		>100		
BM-4		>100		
BM-5		>100		
BM-6		>100		

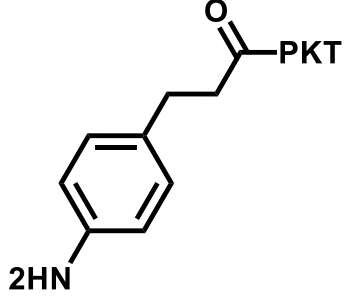
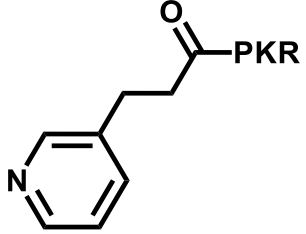
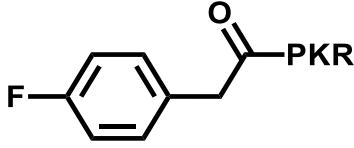
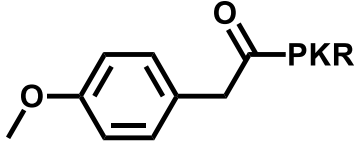
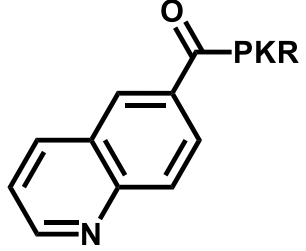
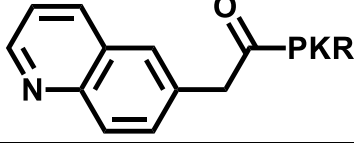
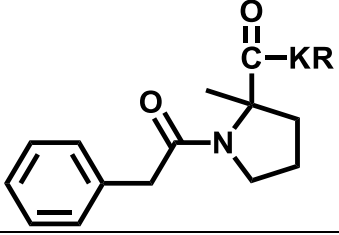
BM-7		>100		
BM-8		>100		
BM-9		$33 \pm 1.5$		
BM-10		>100		
BM-11		$2.7 \pm 0.7$	~70	>100
BM-12		>100		
BM-13		>100		
BM-14		>100		

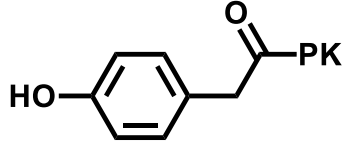
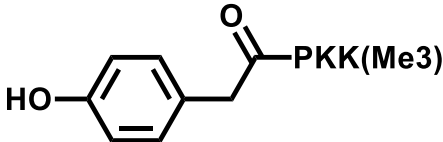
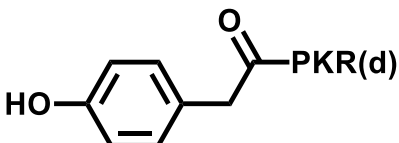
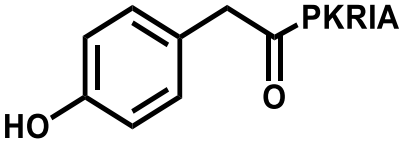
BM-15		>100		
BM-16		>100		
BM-17		>100		
BM-18		>100		
BM-19		>100		
BM-20		>100		
BM-21		>100		
BM-22		>100		

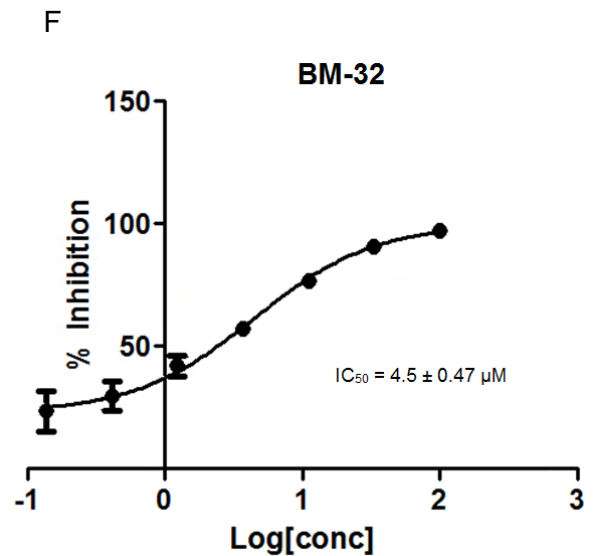
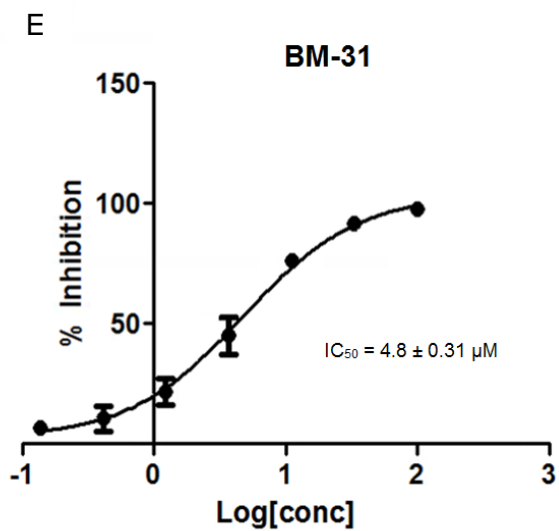
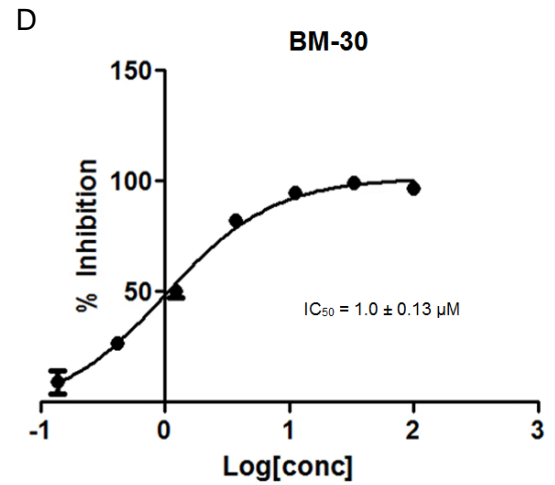
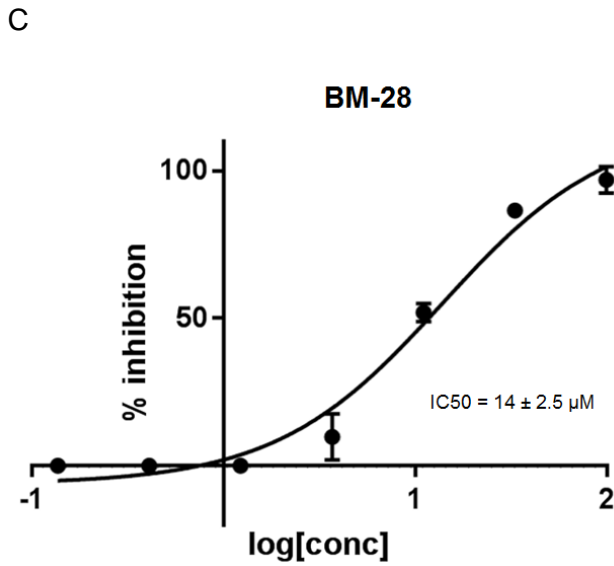
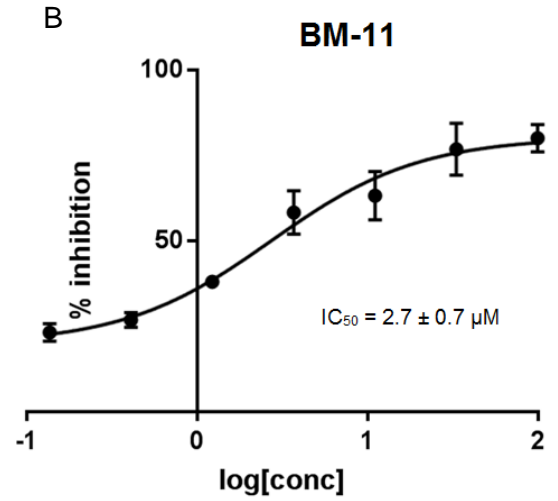
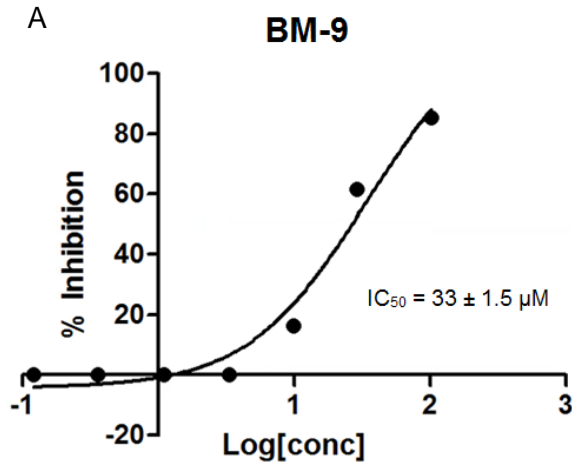
BM-23		>100		
BM-24		>100		
BM-25		>100		
BM-26		>100		
BM-27		>100		
BM-28		$14 \pm 2.5$		
BM-29		>100		
BM-30		$1.0 \pm 0.13$	>100	>100

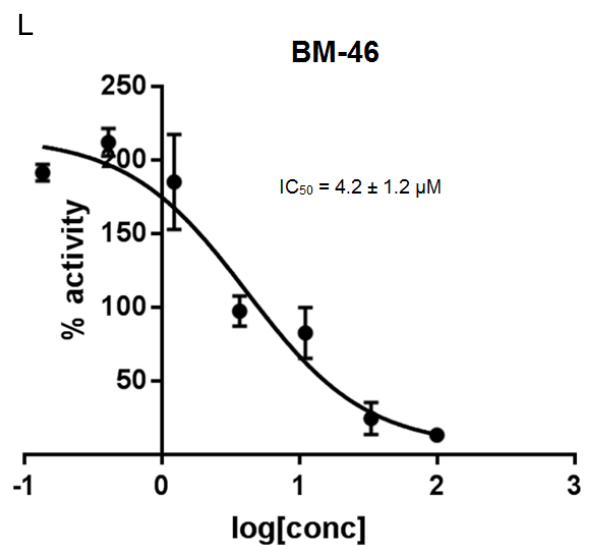
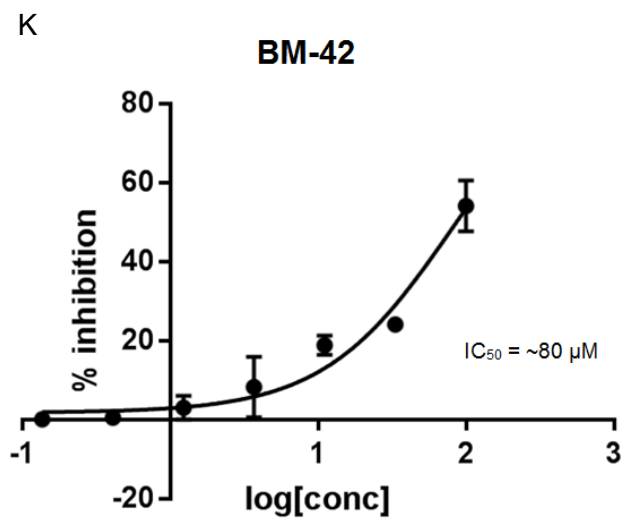
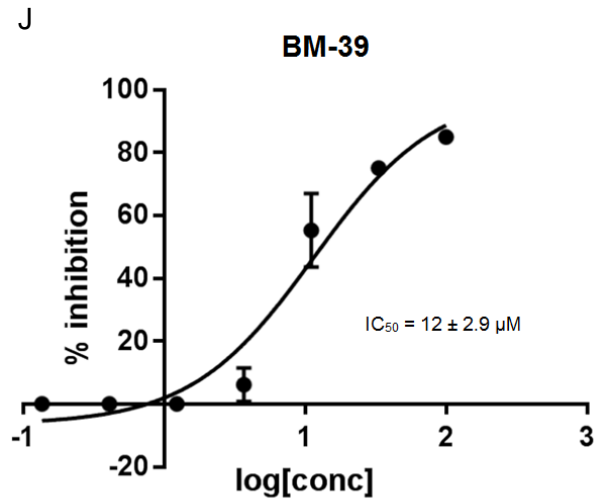
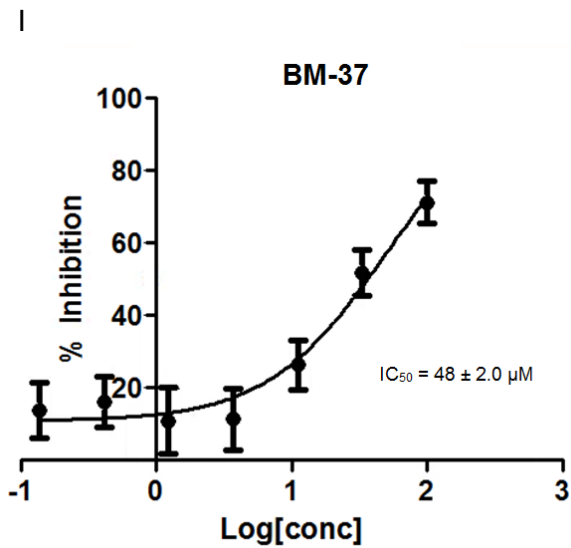
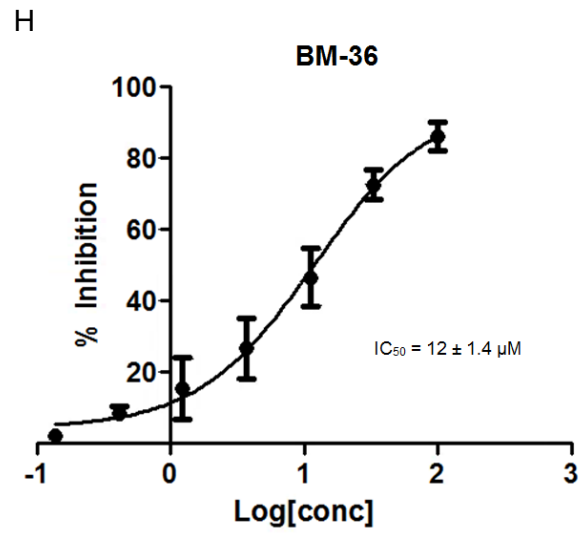
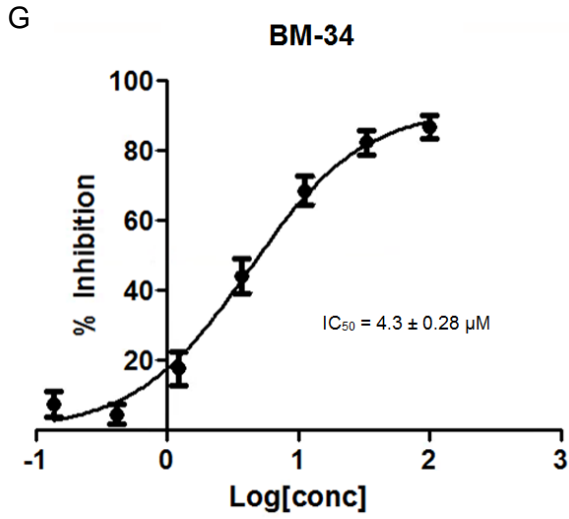


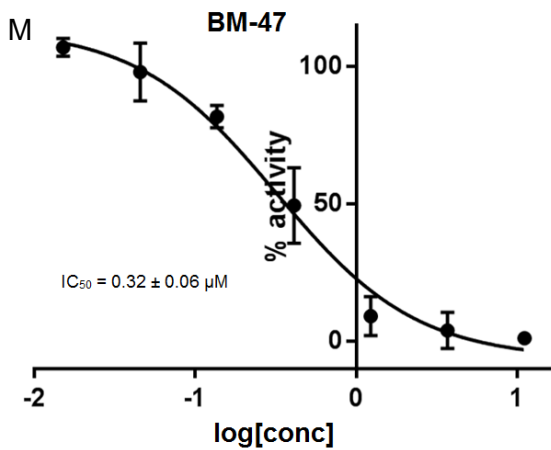
<b>BM-31</b>		<b>4.8 ± 0.31</b>		
<b>BM-32</b>		<b>4.5 ± 0.47</b>		
<b>BM-33</b>		<b>&gt;100</b>		
<b>BM-34</b>		<b>4.3 ± 0.28</b>	<b>&gt;100</b>	<b>&gt;100</b>
<b>BM-35</b>		<b>&gt;100</b>		
<b>BM-36</b>		<b>12 ± 1.4</b>		

BM-37		48 ± 2.0		
BM-38		>100		
BM-39		12 ± 2.9		
BM-40		~80		
BM-41		>100		
BM-42		~80		
BM-43		>100		

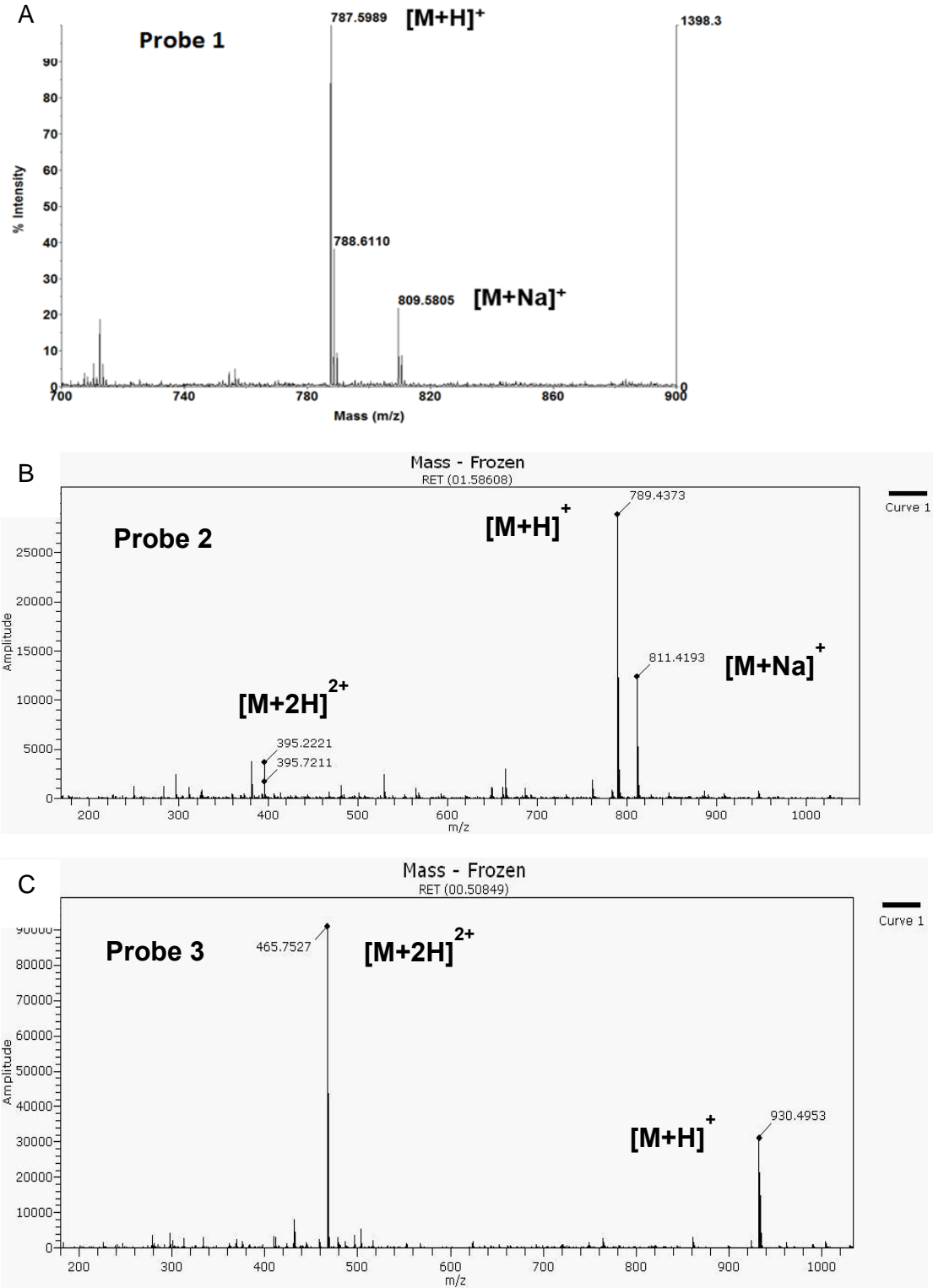
<b>BM-44</b>		<b>~90</b>		
<b>BM-45</b>		<b>~35</b>		
<b>BM-46</b>		<b>4.2 ± 1.2</b>	<b>13 ± 4.1</b>	<b>&gt;100</b>
<b>BM-47</b>		<b>0.32 ± 0.06</b>	<b>&gt;100</b>	<b>&gt;100</b>



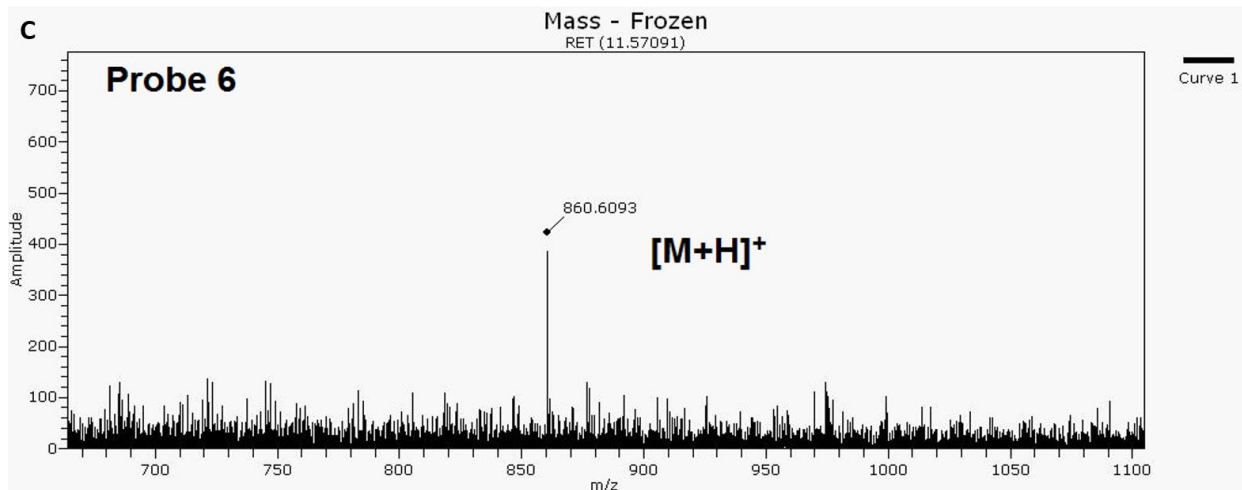
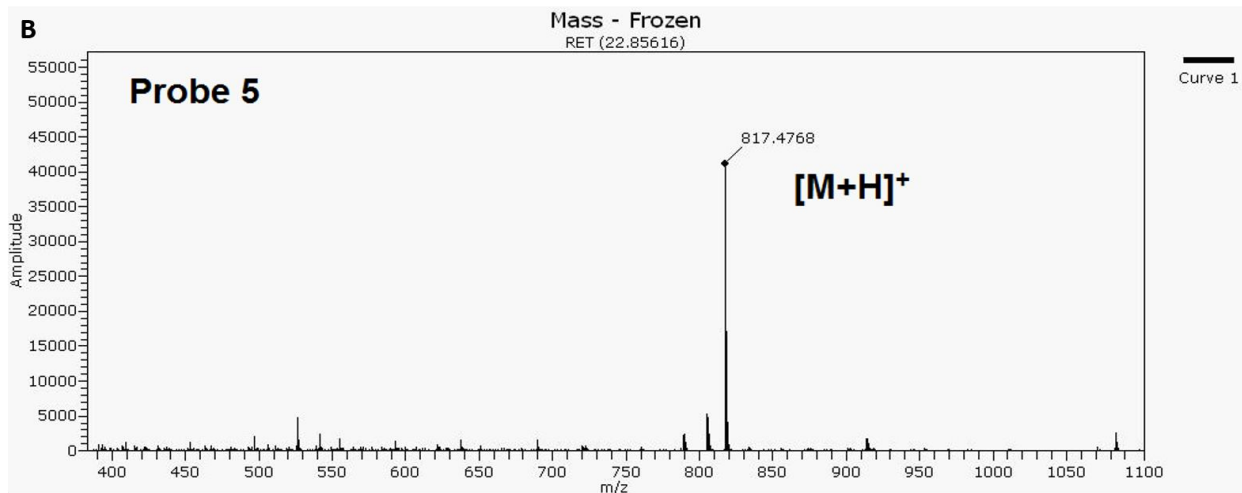
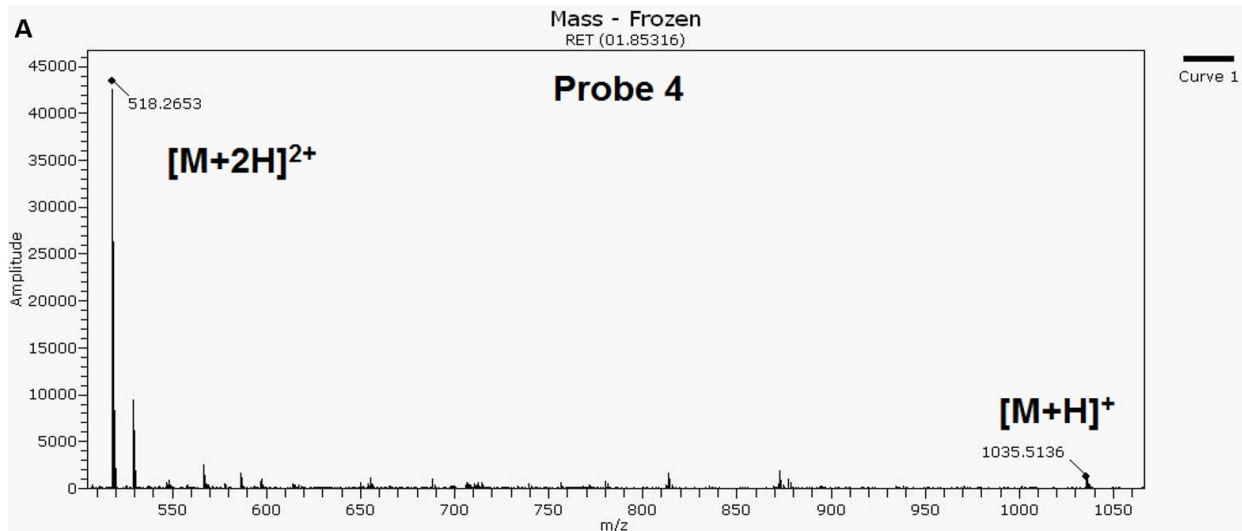




**Figure A6.** IC<sub>50</sub> curves of (A) BM-9; (B) BM-11, (C) BM-28; (D) BM-30; (E) BM-31; (F) BM-32; (G) BM-34; (H) BM-36; (I) BM-37; (J) BM-39; (K) BM-42; (L) BM-46; (M) BM-47

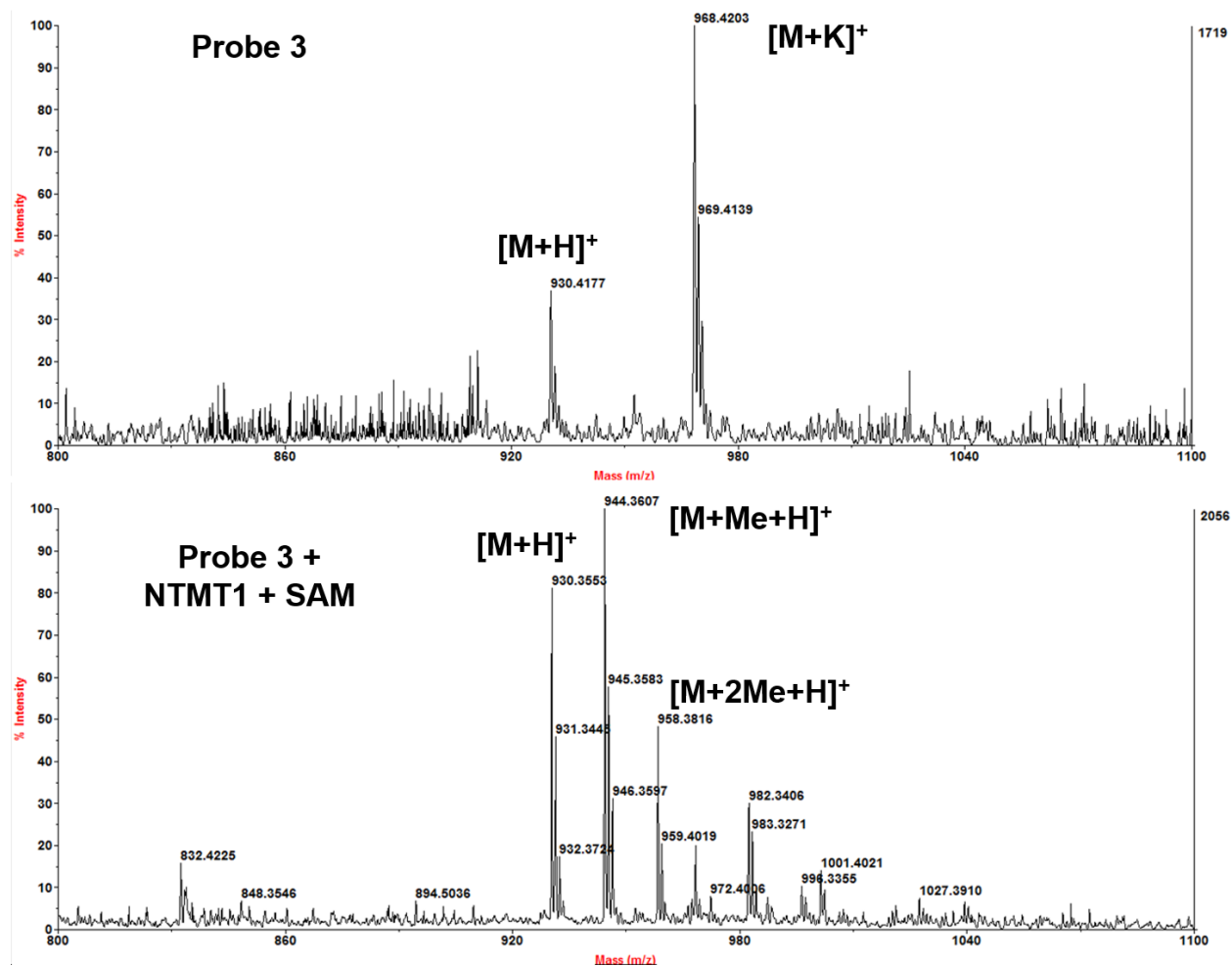


**Figure A7.** Mass spectrum of (A) probes **1** (predicted m/z: 787.4137, detected m/z: 787.5989), (B) **2** (predicted m/z: 789.4366, detected m/z: 789.4373) and (C) **3** (predicted m/z: 930.4945, detected m/z: 930.4953).



**Figure A8.** Mass spectrum of (A) probes 4 (predicted m/z: 1035.5304, detected m/z: 1035.5136), (B) 5 (predicted m/z: 817.4685, detected m/z: 817.4768) and (C) 6 (predicted m/z: 860.5219, detected m/z: 860.6093).





**Figure A9.** MALDI-MS of probe **3** (predicted m/z: 930.4903, detected m/z: 930.4177) (top). MALDI-MS of mixture of unmethylated (predicted m/z: 930.4903, detected m/z: 930.3553), mono- (predicted m/z: 944.5112, detected m/z: 944.3607), and dimethylated (predicted m/z: 958.4503, detected m/z: 958.3816) photoaffinity probe **3** by NTMT1 in the presence of S-adenosyl methionine (SAM) after 15 minutes (bottom).

**Table A2.** Michaelis-Menten constants for probes **1-3** and substrates RB1-10 and SET1-10.

<b>Compound</b>	$K_m$	$k_{cat}$	$k_{cat}/K_m$
<b>Probe 1</b>	4.9 ± 0.7 μM	5.8 ± 0.3 min <sup>-1</sup>	1.2 μM <sup>-1</sup> min <sup>-1</sup>
<b>Probe 2</b>	2.3 ± 0.1 μM	4.6 ± 0.07 min <sup>-1</sup>	2.0 μM <sup>-1</sup> min <sup>-1</sup>
<b>Probe 3</b>	0.7 ± 0.07 μM	4.6 ± 0.1 min <sup>-1</sup>	6.3 μM <sup>-1</sup> min <sup>-1</sup>
<b>RB1-10</b>	0.9 ± 0.2 μM	1.0 ± 0.1 min <sup>-1</sup>	1.1 μM <sup>-1</sup> min <sup>-1</sup>
<b>SET1-10</b>	1.9 ± 0.2 μM	1.0 ± 0.03 min <sup>-1</sup>	0.5 μM <sup>-1</sup> min <sup>-1</sup>

**Table A3.** Predicted and detected masses of peptide inhibitors and photoaffinity probes

<b>Compound</b>	<b>Predicted m/z</b>	<b>Detected m/z</b>
BM-1	448.2555	448.2563
BM-2	503.3088	503.3101
BM-3	531.3150	531.3234
BM-4	454.3024	454.7672
BM-5	509.3558	509.8295
BM-6	537.3620	537.8239
BM-7	707.4563	707.4590
BM-8	462.2711	462.7081
BM-9	517.3245	517.7918
BM-10	545.3207	545.3402

BM-11	701.4457	701.4629
BM-12	440.2868	440.2871
BM-13	495.3402	495.3469
BM-14	523.3463	523.3361
BM-15	693.4406	693.4441
BM-16	438.2711	438.2702
BM-17	493.3245	493.3200
BM-18	521.3307	521.3370
BM-19	677.4457	677.4526
BM-20	416.2404	416.2469
BM-21	471.3038	471.3032
BM-22	499.3099	499.3249
BM-23	402.2347	402.2354
BM-24	457.2881	457.2916
BM-25	485.2943	485.3022
BM-26	489.3296	489.3294
BM-27	434.2762	434.2785
BM-28	567.3402	567.3431
BM-29	512.2868	512.2907
BM-30	533.3194	533.3295
BM-31	478.2660	478.2631
BM-32	547.3351	547.3435
BM-33	492.2817	492.2858

BM-34	532.3354	532.3393
BM-35	477.2820	477.2833
BM-36	546.3511	546.3712
BM-37	491.2976	491.2973
BM-38	532.3301	532.7223
BM-39	535.3193	535.7662
BM-40	547.3387	547.3602
BM-41	554.3128	554.8645
BM-42	568.3392	568.8913
BM-43	531.3365	531.9072
BM-44	377.2183	377.2165
BM-45	547.3608	547.3923
BM-46	533.3194	533.3492
BM-47	717.4406	717.5842
Probe 1	787.4137	757.5989
Probe 2	789.4366	789.4373
Probe 3	930.4945	940.4953
Probe 4	1035.5305	1035.5136
Probe 5	817.4685	817.4768
Probe 6	860.5219	860.6093

### Vita

Brianna Danielle Mackie was born on September 14<sup>th</sup>, 1991 in Bitburg, Germany and is an American citizen. She graduated from Grafton High School in Yorktown, Virginia in 2009. She received her Bachelor of Science in Chemical Engineering from Virginia Polytechnic Institute and State University in Blacksburg, Virginia in 2013.