Development of fluorescence-based supramolecular tools for studying histone post-translational modifications

by

Sara Tabet
B.Sc., University of Victoria, 2011

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Chemistry

Supervisory Committee

Development of fluorescence-based supramolecular tools for studying histone post-translational modifications

by

Sara Tabet
B.Sc., University of Victoria, 2011

Supervisory Committee

Dr. Fraser Hof, Department of Chemistry
Supervisor

Dr. Natia Frank, Department of Chemistry
Departmental Member

Abstract

Supervisory Committee
Dr. Fraser Hof, Department of Chemistry
Supervisor
Dr. Natia Frank, Department of Chemistry
Departmental Member

A large variety of post-translational modifications can exist on the N-terminal tails of histone proteins H2A, H2B, H3 and H4. These have been of great interest as they have increasingly been shown to influence fundamental biological processes and human disease. Studying these modifications provides insight into their physiological functions and enables the search for potent small molecule inhibitors. In this thesis, new fluorescence-based supramolecular tools were developed and used to a) measure the binding of covalently modified peptide tails to a collection of synthetic receptors in neutral aqueous solution and b) monitor an enzyme that installs a post-translational modification (PTM) in real-time. Two different approaches were used to detect binding in these systems. The first was the optimization of a competitive dye-displacement method that relies on the ability of the cationic dye lucigenin. The second was the synthesis of novel conjugates that consist of calixarenes covalently appended with multiple different fluorescent dyes.

Table of Contents

List of Tables

List of Figures

List of Schemes

Abbreviations

Ac	Acetylated
AlphaScreen	Amplified luminescent proximity homogeneous assay screen Technology
CBX7	Chromobox protein homolog 7
CHD4	Chromodomain-helicase-DNA-binding protein 4
CN	Cyano
DACITC	7-Dimethylamino-4-methylcoumarin-3-isothiocyanate
DCM	Dichloromethane
DELFIA	Dissociation enhanced-lanthanide fluorescence immunoassay
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
EZH1/EZH2	Enhancer of zeste homologue 1/2
F	Emission at maximum wavelength upon treatment with a guest
F_0	Emission at maximum wavelength before treatment with a guest
FDH	Formaldehyde dehydrogenase
FITC	5(6)-Fluorescein Isothiocyanate
FITC-Ba	5(6)-Fluorescein Isothiocyanate Benzylamine
Fmoc	Fluorenylmethyloxycarbonyl chloride
G9a/EHMT2	Euchromatic histone methyltransferase 2
H_2O	Water
H3K27	Histone H3 lysine 27
H3K27me3	Histone H3 trimethylated at lysine 27
H3K36me2	Histone H3 dimethylated at lysine 36
H3K36me3	Histone H3 trimethylated at lysine 36

H3K4	Histone H3 lysine 4
H3K4me2	Histone H3 dimethylated at lysine 4
H3K4me3	Histone H3 trimethylated at lysine 4
H3K9	Histone H3 lysine 9
H3K9me2	Histone H3 dimethylated at lysine 9
H3K9me3	Histone H3 trimethylated at lysine 9
H4K5	Histone H4 lysine 5
H4K5ac	Histone H4 acetylated at lysine 5
HBTU	O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate
HDAC	Histone deacetylase
HDM	Histone demethylase
HMT	Histone methyltransferase
HP1	Heterochromatin protein 1
HRP	Horseradish peroxidase
ITC	Isothermal titration calorimetry
JMJD2A	Lysine-specific demethylase 4A
JMJD2C	Lysine-specific demethylase 4C
K	Lysine
K36	Lysine 36
K4	Lysine 4
K79	Lysine 79
Kac	Acetylated lysine
K_d	Association constant of the calixarene-guest complex
K_{ind}	Association constant for host-dye complex
Kme	Monomethyllysine
Kme2	Dimethyllysine
Kme3	Trimethyllysine
LANCE	Lanthanide chelate excite
LCG	Lucigenin
LSD1	Lysine-specific histone demethylase 1A
MALDI-TOF	Matrix-assisted laser desorption/ionization- Time of Flight

Me	Methylated
MeCN	Acetonitrile
MeO	Methoxy
$MgCl_2$	Magnesium chloride
MS	Mass spectrometry
NADH	Nicotinamide adenine dinucleotide
NMP	N-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
N-terminal	Amino terminal
PHD	Plant Homeodomain
PRC1	Polycomb repressive complex 1
PSC4	*p*-sulfonato calix[4]arene
PTM	Post-translational modification
R	Arginine
Rme	Monomethylarginine
Rme2a	Asymmetric dimethylated arginine
Rme2s	Symmetric dimethylated arginine
RP-HPLC	Reverse phase-High performance liquid chromatography
S	Serine
SAH	S-adenosylhomocysteine
SAHH	SAH-Hydrolase
SAM	S-adenosylmethionine
T	Threonine
TFA	Trifluoroacetic acid
TR-FRET	Time-resolved Fluorescence resonance energy transfer
TRITC	Tetramethylrhodamine Isothiocyanate
Y	Tyrosine

Acknowledgments

First, I would like to thank my supervisor Dr. Fraser Hof for giving me the opportunity to join his research group. Thank you for your guidance on this project and continued support. It was great having a supervisor who has a positive outlook and is always encouraging of new ideas. Thank you for allowing me to grow my research skills and learn lessons that will serve well in my future.

I would also like to thank the other group members of my research group. You have all been a great support and have helped me enjoy working on my project even through stressful times by creating a positive and enjoyable environment. A special thanks to Kevin Daze for all your assistance with my research, Kevin Allen for the kindest introduction anyone has done for me, and Chakravarthi Simhadri for being my big brother at work.

To Dr. Anna Patten, Geneviève Boice and Andrew Leung, thank you all for editing parts of this thesis. I appreciate your help and look forward to returning the favour in the form of baked-goods.

Finally I would like to thank my friends and family. To my friends: I doubt you will ever read this, but thank you so much for all your support. I am blessed and grateful to have friends like you. To my family:

شكراً ماما و بابا على كل دعمكم في دراستي و حياتي, بدوكن ما كنت وصلت. فراس و سحر, شكراً على كونكم احسن اخ و اخت و دعمكم بطريقتكم الخاصة.(SiYA)

To the love of my life, Mohamed, words cannot express how much I appreciate your support throughout the past two years that I spent in the lab and especially during the past couple of months as I wrote my thesis. You kept me calm when I got nervous, and helped me maintain my tranquility when I felt things were getting out of control.

Dedication

~Mama and Baba~

Chapter 1: Introduction

1.1 The Histone Code

In 1998, Klar made a statement that "We are more than the sum of our genes".[1] Numerous molecules play important structural and regulatory roles in the organization and regulation of DNA.[2] In the nuclei of all eukaryotic cells, the genetic information encoded in DNA is under strict control by specialized proteins that form a dynamic polymer called chromatin. Understanding how regulation and transduction of genetic information happens requires a closer look into this dynamic structure and its components.

1.1.1 Gene regulation and post-translational modifications

Chromatin proteins aid in packaging DNA within the cell's nucleus. The fundamental unit of chromatin is called a nucleosome which consists of 146 base pairs of DNA wrapped around proteins known as histones in a left-handed superhelix (Figure 1.1).[3-9] There are four types of core histone proteins known as H2A, H2B, H3 and H4, and two copies of each are present in each nucleosome.[5] The core histone proteins are predominantly globular and organized into a highly ordered structure except for their protruding unstructured N-terminal tails. Dynamic post-translational modifications (PTMs) that occur on these N-terminal tails have been the focus of much research in the last few decades and have been shown to be essential for regulation of gene expression.[10, 11]

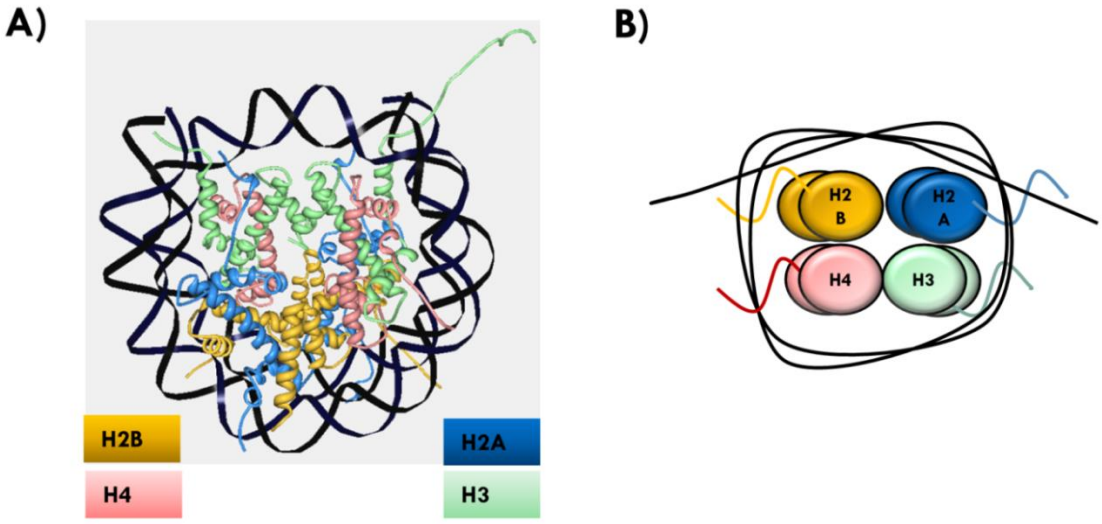


Figure 1.1 The basic structure of a nucleosome. A) Nucleosome unit Pdb1AOI showing the four core histone proteins H2A, H2B, H3 and H4 with 146 DNA base pairs wrapped around approximately 1.65 times. B) Cartoon depiction of a nucleosomal unit that will later be used when referring to histone post-translational modifications.

Chromatin is classified as having transcriptionally active or silent regions, which are referred to as euchromatin and heterochromatin, respectively. One central idea in gene regulation is that these states are controlled by PTMs. Figure 1.2 shows these two chromatin configurations where various modifications, such as acetylation, are done on the histone tails of nucleosomes. The nucleosomes in euchromatin are more accessible for transcription machinery than those in heterochromatin. Some general correlations between PTM state and the state of chromatin exist: While for example, silent heterochromatin is associated with hypoacetylated histones, active euchromatic regions of the chromosome are often associated with hyperacetylated histones.[12] This bulk effect is thought to arise from the fact that each histone acetylation neutralizes a cationic lysine side chain, and therefore weakens the compacting interactions between the anionic DNA phosphate backbone and cationic histone proteins.[13] Understanding the regulation of both the heterochromatic and euchromatic regions is crucial for understanding their role in gene transcription, heterochromatin formation, DNA replication, DNA repair, etc. [2, 14-16]

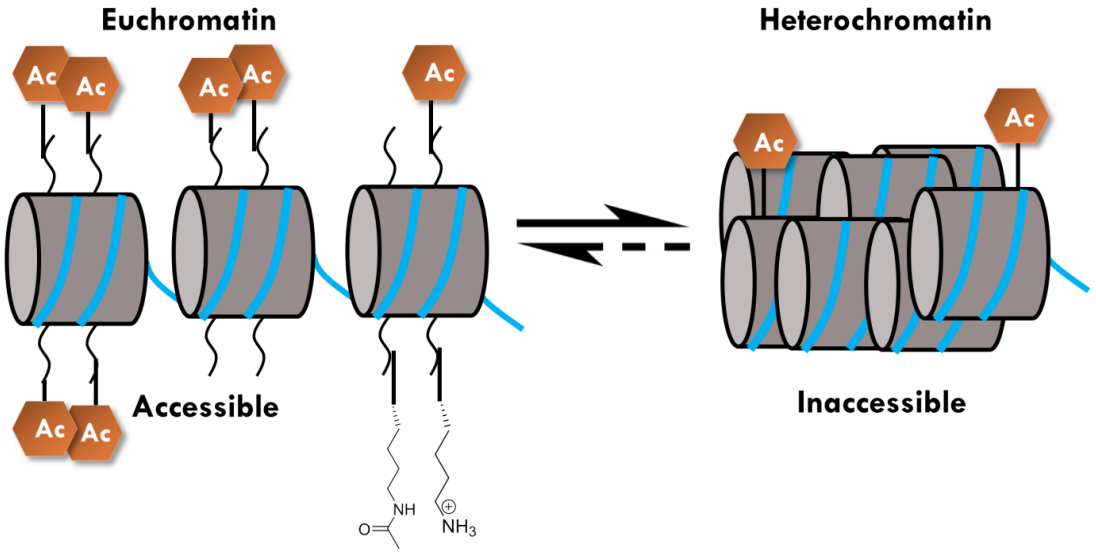


Figure 1.2 Euchromatin and heterochromatin with acetylated (Ac) modifications on histone tails. Hyperacetylated euchromatin (active chromatin) and hypoacetylated heterochromatin (silent chromatin) showing accessible and condensed nucleosomes. A cationic lysine side chain is shown alongside a neutralized acetylated lysine side chain in the active chromatin.

In addition to acetylation, there are numerous other post-translational modifications that occur on the N-terminal tails of the core histones as shown in Figure 1.3. Covalent modifications such as methylation of lysine (K) and arginine (R), acetylation of lysine, phosphorylation of serine (S) and threonine (T) as well as deimination of arginine to citrulline all occur.[17-19] Non-covalent modifications such as proline (P) *cis-trans* isomerization also exist.[20, 21] The "Histone code" therefore refers to these covalent and non-covalent modifications that occur on the charged N-termini of histones that cause downstream effects on gene transcription. Specific enzymes or complexes catalyze these modifications and these will be discussed in section 1.1.2. The molecular connections between each of these modifications and their impact on gene regulation are generally much more complicated than the above example given for lysine acetylation. Some examples will be discussed in more detail below. It is important to note that when referring to specific PTMs in this thesis, I will be using the conventional naming system as shown in Figure 1.3. For example, a trimethylation that occurs on lysine 9 of histone H3 will be referred to as H3K9me3.

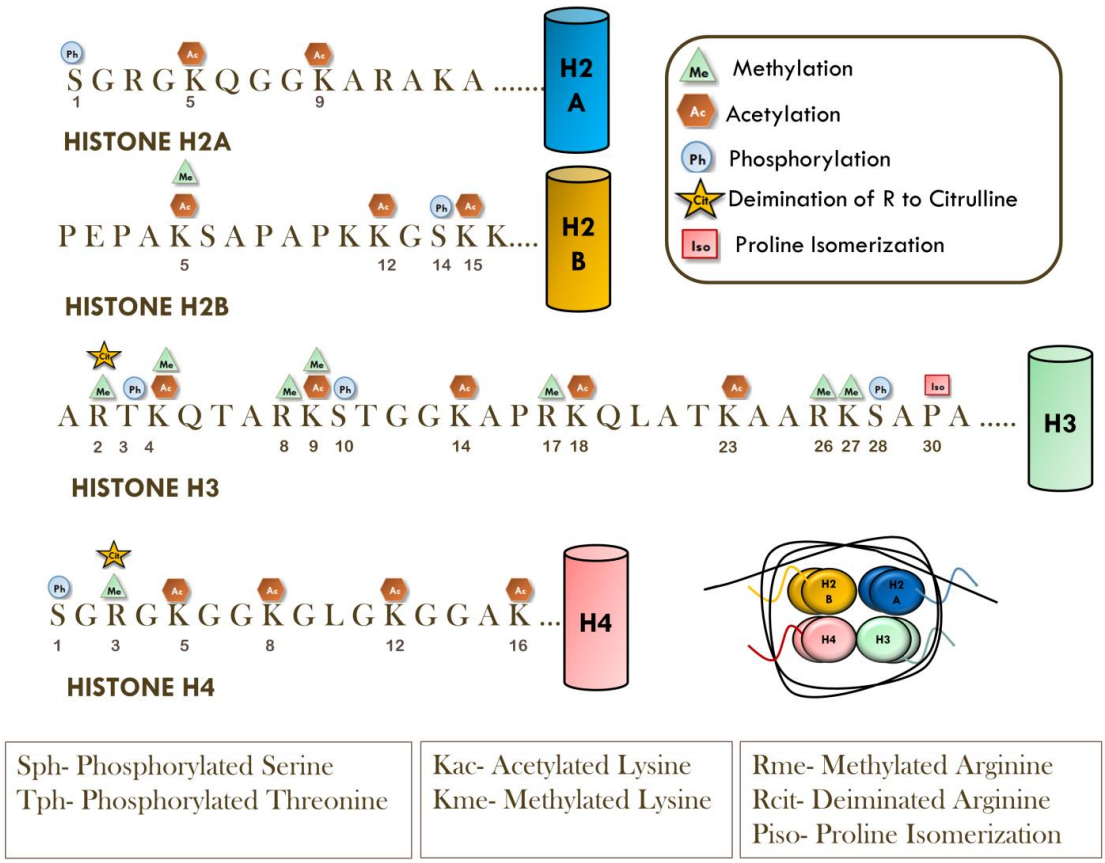
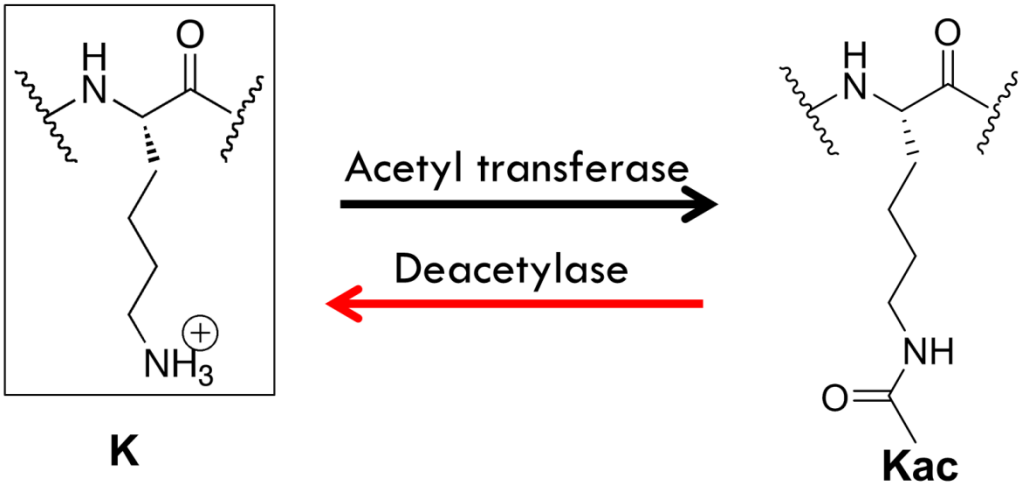


Figure 1.3 Post-translational modifications known to occur on histone tails.
One-letter codes for amino acid names are used for each histone tail sequence, and the core of each histone is represented as a coloured cylinder.

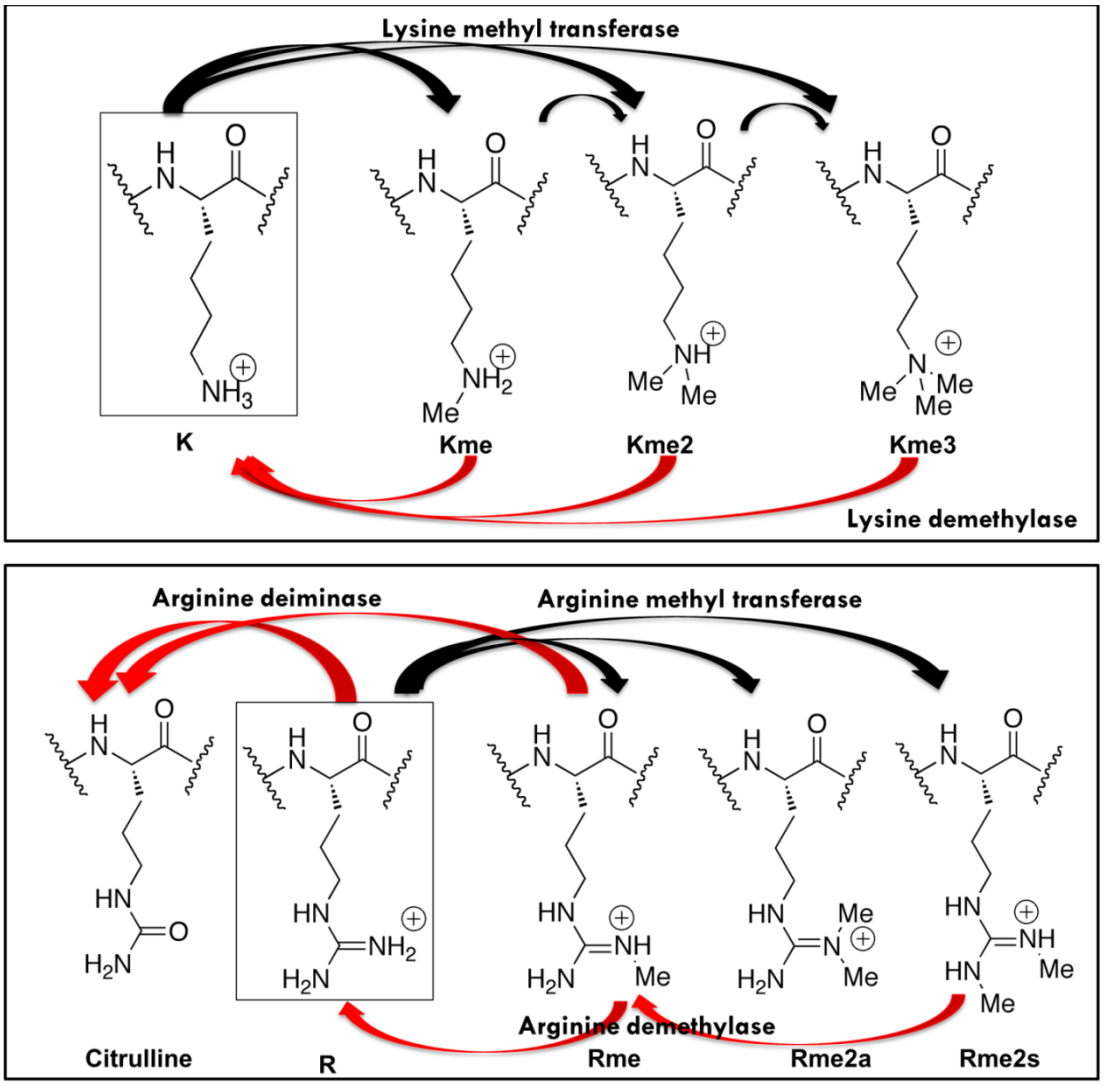
Acetylation (ac) occurs on lysine (**K**) residues of the four core histone proteins neutralizing the cationic side chain producing N-ε-acetyllysine (**Kac**) as shown in Scheme 1.1. An acetyl transferase catalyses the addition of an acetyl group which activates transcription,[22] while a deacetylase facilitates the removal of the acetyl group and correlates with transcriptional repression.[23] Most acetylations are found on the protruding N-terminal tails of histones, such as H4 (K5, K8,12,16)[12] as shown in Figure 1.3 but it is also possible to find such a modification on the globular folded core of histone 3, at H3K56.[24, 25]



Scheme 1.1 Lysine acetylation and deacetylation and the representative enzymes that regulate the modification.

Methylation (me) occurs on both arginine (R) and lysine (K) residues and is mediated by methyltransferases that have evolved specifically for each amino acid. Lysines can be monomethylated (Kme), dimethylated (Kme2) or trimethylated (Kme3) as shown in Scheme 1.2. Each of the lysine methylation states maintains the positive charge but increases the size and hydrophobicity of the residue while decreasing possible NH hydrogen bonding sites.[26] Depending on the modified state and its location on the histone tail, each modification can cause activation or repression of transcription. Generally, methylation of H3 (K4, K36 and K79) is known to cause activation of transcription while methylation of H3 (K9 and K27) causes repression.[27] These different outcomes do not arise directly from bulk changes in chromatin compaction, as with acetyllysine, but instead arise from the abilities of each kind of methylation mark to bind to and recruit different chromatin-modifying factors. These outcomes will be discussed in detail in section 1.1.2. Arginine residues can be monomethylated (Rme), dimethylated asymmetrically (Rme2a) or symmetrically (Rme2s). Like lysine methylation, these methylations do not change the overall charge of the side chain, but do change its size and hydrophobicity. Dimethylarginine modifications can also activate or repress transcription, depending on the exact location and depending on the isomeric form of the dimethylation that occurs. For example, methylation of H3R17me2a can activate transcription[28] whereas methylation of H4R3me2s acts as a repressor.[29] Both lysine and

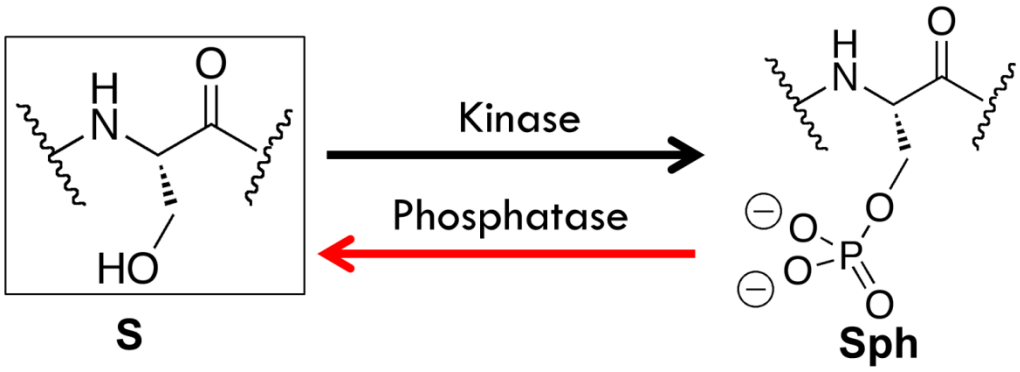
arginine methylations can be removed by their respective demethylases[30-33] as well as deiminases[34] in the case of arginine.



Scheme 1.2 Different post-translational states upon methylation of lysine (K) and arginine (R) and the respective enzymes that regulate them.

Phosphorylation (ph) occurs predominantly on serine (S), threonine (T) and tyrosine (Y) residues on histone tails. An example of phosphorylation of serine is shown in Scheme 1.3, where the addition and removal of a phosphate group is regulated by kinases or phosphatases, respectively. Unlike methylation, phosphorylation drastically alters the charge of the residue changing a neutral residue to an anionic one. Most

phosphorylation post-translational modifications are found on the N-terminal tails, such as H2AS1, H3T6, H4S1 but can also be found in the core regions, such as H3Y41.[35]



Scheme 1.3 An example of phosphorylation of a serine (S) residue and the respective enzymes that regulate the modification.

Why are histone post-translational modifications important to study? Epigenetics is defined as the study of "heritable changes in gene expression not encoded by the DNA sequence".[36] One of the major mechanisms by which this occurs involves gene regulation mediated by histone modifications, which are heritable because histones are passed along to daughter strands of DNA when DNA is replicated during cell division.[37] Histone PTMs have been linked to a variety of biological processes such as DNA damage, apoptosis, cell-cycle regulation and disease.[14] For example, apoptosis, or programmed cell death, is observed in osteosarcoma cells upon monomethylation of H3K27.[38] Other examples exist where combinations of modifications influence the overall chromatin structure and function. In order to better understand the modifications that occur and their role in biology, it is important to study the proteins that bring about these modifications.

1.1.2 Writers, Erasers and Readers

Proteins known as “writers”, “erasers” and “readers” allow addition, removal or detection of post-translational modifications respectively. In schemes 1.1, 1.2, and 1.3, we pointed out different “writer” and “eraser” proteins that regulate PTMs in a dynamic way. When these modifications are present, “reader” proteins are able to recognize and bind to a specific modification site and cause downstream effects by altering gene expression. For example, a writer protein known as EZH2,[39,40] allows the methylation of lysine 27 of histone 3 (H3K27) to make H3K27me3. Upon methylation of H3K27, a reader protein known as CBX7[41] recognizes and binds to the H3K27me3 modification site, and subsequently recruits other factors that cause repression of gene expression.[42] More about this particular modification will be discussed in Chapter 2.

Writer, eraser and reader proteins are required for the normal operation of the cell. As with many other kinds of gene regulation pathways, proteins associated with PTMs are very often co-opted by cancer cells in order to confer a growth advantage. Table 1.1 shows some examples of histone lysine methyltransferase (HKMT) and demethylase (HDM) enzymes and their known associations with cancer.[43] These associations are often highly specific to a certain tumour or tissue type. For example, the HDM eraser protein JMJD2C demethylates H3K9me3 and is associated with prostate and other cancers, whereas HKMT writer protein such as G9a (which will be the focus of Chapter 3) dimethylates H3K9 and is associated with gastric cancer and others (Table 1.1).

Table 1.1 Some examples of histone lysine methyltransferase (HMT) and demethylase (HDM) enzymes implicated in cancer.[43]

Enzyme Type	Gene Name	Substrate	Product	Associated Cancer
HMT	MLL1	H3K4	H3K4me1/2	Human lymphoid, myeloid leukemias
	EZH2	H3K27	H3K27me1/2/3	Prostate, breast, follicular, germinal center B cell lymphoma
	G9a	H3K9	H3K9me2	Hepatocellular carcinomas, gastric cancer
	Suv39H1	H3K9	H3K9me3	Colon
	SMYD3	H3K4	H3K4me2/3	Colon, breast hepatocellular carcinoma
	NSD1	H3K36	H3K36me2	Acute myeloid leukemia, neuroblastoma, glioma
	NSD2	H3K36	H3K36me2	Multiple myeloma
	NSD3	H3K36	H3K36me2	Leukemia, breast
	Ash2L	H3K4	H3K4me1/2	Squamous cell carcinomas of cervix and larynx, melanoma, rhabdomyosarcoma, breast and colon carcinomas, neuroendocrine carcinoma, pancreatic ductal adenocarcinomas and gastric carcinomas
HDM	LSD1	H3K4me1/2 H3K9me1/2	H3K4 H3K9	Prostate, neuroblastoma, breast
	JMJD2C	H3K9me3 H3K36me3	H3K9 H3K36	Prostate, esophageal squamous cell carcinoma, desmoplastic medulloblastoma, MALT lymphoma
	JMJD3	H3K27me2/3	H3K27	Prostate
	FBXL10	H3K4me3 H3K36me2	H3K4 H3K36	Lymphoma, brain, glioblastoma multiforme
	RBP2	H3K4me2/3	H3K4	Gastric cancer
	PLU-1	H3K4me1/2/3	H3K4	Breast, prostate, testis, ovary

1.2 Currently available methods for detection of post-translational modifications, and their application in assays for enzyme activity

1.2.1 The use of antibodies as post-translational modification recognition elements

Several biochemical assays have been developed that use antibody recognition to detect specific post-translational modifications. The methods used to develop antibodies have been described elsewhere.[44] While the use of antibodies has enabled investigators to create high-throughput assays, some of which are discussed in the subsections below, one general limitation to all of the following antibody-driven assays lies in the limited availability and low quality of specific antibodies for many PTMs. Egelhofer *et al.* showed in a study published in 2011 that out of 200 commercially available antibodies against post-translational modifications that were tested, 25% failed specificity tests and more than 20% failed in chromatin immunoprecipitation experiments.[45] Since there are more than 1000 different commercial antibodies raised against PTMs,[46] a database has been created[45] in order to allow other researchers to share their data. The view that antibodies against PTMs are lower quality than antibodies in general is commonly held in the literature.[44, 47, 48] Antibodies raised against PTMs tend to have difficulty distinguishing between closely related modifications (mono-, di- or tri-methyl) causing cross-reactivity, can be influenced in unpredictable ways by other nearby PTMs on the same substrate,[48] or can simply bind to non-targeted modifications. Nonetheless, antibodies have advanced the discovery and characterization of new PTMs, and remain the primary tools that allow the study of histone writers and erasers as well as their inhibition.

1.2.1.1 Enzyme-linked immunosorbent assay (Assay A)

Enzyme-linked immunosorbent assay (**ELISA**) has been used in the past to measure the activity of certain writer enzymes. Figure 1.4 shows a cartoon depiction of the principle of an ELISA assay. First, the substrate for a given enzymatic reaction is coated onto a multi-well plate followed by the addition of the enzyme and the cofactors necessary. The plate is incubated to allow for the PTM installation process catalyzed by the writer enzyme to occur. A primary antibody specific to the modification is added (1° antibody anti-modification), followed by a secondary antibody (2° antibody) that recognizes the primary antibody and is conjugated with the enzyme horseradish peroxidase (HRP). Rinsing after each treatment removes non-specifically bound antibodies. The enzyme HRP (the "E" in ELISA) allows for detection using a developing kit. Depending on the substrate used HRP catalyses the conversion of the chromogenic substrate into coloured products or produces light in the presence of a chemiluminescent substrate. For example, hydrazides such as luminol, undergo oxidation in the presence of hydrogen peroxide and the signal is enhanced in the presence of a chemiluminescent enhancer such as *p*-iodophenol to produce aminophthalate. The *p*-iodophenol enhancer acts as an electron-transfer mediator that increases the formation of the luminol radical. The luminol radical forms an α-hydroxy-peroxide intermediate which then forms the excited aminophthalate product.[49] The excited state product decays to a ground energy state by releasing photons with emission measured at 425 nm.[50] It is important to note that the enzymatic reaction is allowed to occur for a fixed period of time, prior to quenching, antibody treatment, and the detection step. Readout does not occur simultaneously to the PTM reaction, and so a continuous readout of reaction progress over time is not possible. A higher signal from HRP signifies that more substrate peptide became product peptide with modification. The first protein arginine methyltransferase (PRMT) enzyme inhibitor was identified using an ELISA-based assay.[51] ELISA assays are quantitative and sensitive which allow accurate detection of the modification being studied but the presence of antibodies at different dilutions in the assay, and also differences in incubation times, can affect the readout and run-to-run reproducibility.[44] It should be noted however, that other variations of ELISA assays such as sandwich

ELISA, direct ELISA, and competitive ELISA are also possible but are not described in detail here.

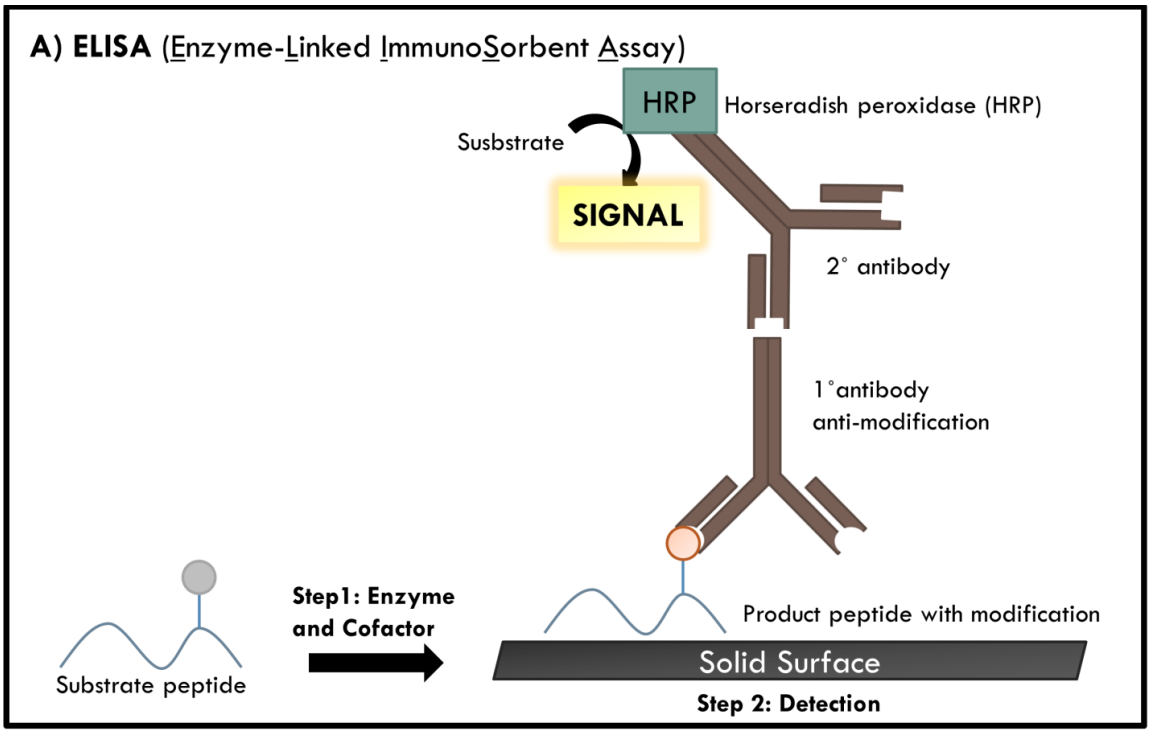


Figure 1.4 Cartoon depiction of the principles of an indirect ELISA assay.

Substrate peptide represents peptide before post-translational modification while product peptide with modification represents the peptide after the enzymatic reaction. ELISAs are run in two main steps: 1) The enzymatic reaction 2) The detection method, in this case horseradish peroxidase (HRP) enzyme catalyses a substrate producing an emission signal measured at 425 nm.

1.2.1.2 Dissociation enhanced-lanthanide fluorescence immunoassay (Assay B)

A more modern version of ELISA is known as dissociation enhanced-lanthanide fluorescence immunoassay (**DELFIA**) where the secondary antibody is instead labeled with a lanthanide chelate such as europium, terbium or samarium instead of HRP as shown in Figure 1.5. Similar to ELISA, product formation detection is based on the 2° antibody signal, which in this case is measured with time-resolved fluorescence of the long-lifetime lanthanide label. This technique was used by Kubicek and Spannhoff in 2007 for detection of both lysine methyl transferase inhibitors[52] and arginine methyltransferase inhibitors[53] respectively. Typically, europium is used as the fluorescent tag on the secondary antibody and emits at 620 nm for detection. This assay, like ELISA, requires the initial enzymatic reaction to occur for a fixed time, followed by quenching and antibody-based detection.

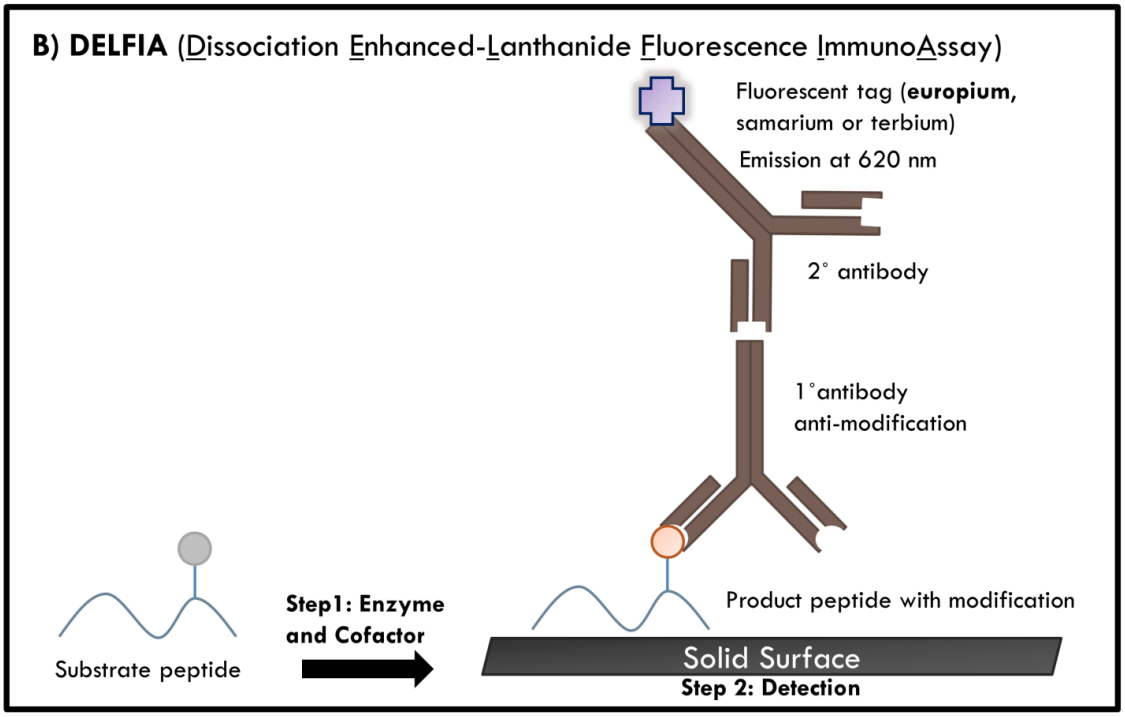


Figure 1.5 Cartoon depiction of principle of DELFIA assay. Substrate peptide represents peptide before post-translational modification while product peptide with modification represents the peptide after the enzymatic reaction. DELFIAs are run in two main steps: 1) The enzymatic reaction 2) The detection method, in this case a fluorescently (typically europium) tagged secondary antibody recognizes the primary anti-modification antibody and produces a signal at 620 nm.

1.2.1.3 Amplified luminescent proximity homogeneous assay screen technology (Assay C)

Another technology using antibodies is known as amplified luminescent proximity homogeneous assay screen (**AlphaScreen)**.[54, 55] This technique has been used to measure the activity of G9a histone methyltransferase writer protein[56] and of lysine demethylase eraser protein JMJD2E[57] *in vitro*. This homogenous assay does not require any washing steps unlike both ELISA and DELFIA but can also be done in multi-well plates like the previous two assays presented. As shown in Figure 1.6, this assay is a dual-bead based assay technology that utilizes an acceptor and donor bead. The excitation of the dyes embedded in the donor bead at 680 nm causes release of singlet oxygen, which can diffuse through solution during its short lifetime and generate a chemiluminescent signal at 520-620 nm when it reacts with the singlet oxygen-sensitized chromophores that are embedded in the acceptor bead. The chromophores embedded in the acceptor bead are thioxene, anthracene and rubrene. Upon excitation of the donor bead and the release of singlet oxygen, the singlet oxygen molecules can travel around 200 nm and react with thioxene. Thioxene produces light energy which is transferred to anthracene then rubrene.[58] The signal from rubrene is significant only when the beads are close enough together for the singlet oxygen to diffuse from one to another during its lifetime. The beads are brought together by programmed binding interactions as follows: the donor bead is usually coated with streptavidin and therefore able to recognize the biotin-modified substrate. The acceptor bead is coated with a secondary antibody that recognizes the primary anti-modification antibody as in the case of ELISA and DELFIA. Only when the PTM has been installed on the substrate are the beads brought together (see Figure 1.6). It is also possible with this assay to directly couple the primary anti-modification specific antibody to the acceptor bead instead of using a secondary antibody.

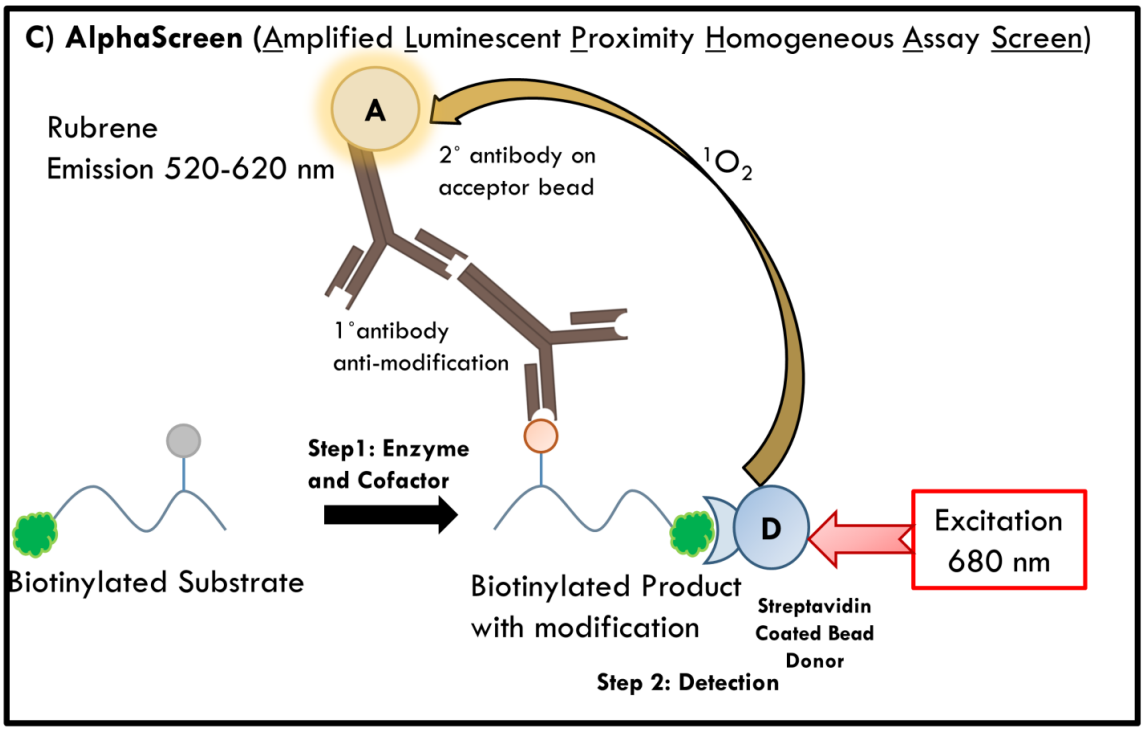


Figure 1.6 Cartoon depiction of the principles of the AlphaScreen assay which is a dual-bead based assay. Substrate peptide represents peptide before post-translational modification with the green cloud representing biotinylation of the peptide. Streptavidin coated bead (D) binds the biotinylated peptide and is excited at 680 nm. An anti-modification primary antibody specific to the post-translational modification is detected by an acceptor bead (A) coated with a secondary antibody. Excitation of the donor bead releases singlet oxygen molecules which are detected by the acceptor bead producing an emission at 580 nm.

This assay is not only used for tracking writer and eraser modifications *in vitro*, but also for detecting reader protein-PTM interactions. A study by the Wigle group published in 2010 showed that this technology can be used to detect malignant brain tumor (MBT) domain-containing proteins that bind certain methylation states of lysine residues on histone tail peptides.[59] Using the AlphaScreen principle, this group utilized the reader protein MBT in substitute of an antibody that would bind the methylated lysine on the peptide. The MBT protein was coated directly onto the acceptor bead. Upon excitation of the donor bead and release of the singlet oxygen, the acceptor bead/MBT protein would emit if bound to the methylation state of peptide and in close proximity.

This variant of the assay was set up to identify and characterize inhibitors of that protein-protein interaction.

1.2.1.4 Lanthanide chelate excite (Assay D)

Another technology similar to AlphaScreen, called lanthanide chelate excite (**LANCE**), utilizes a similar concept. The primary anti-modification specific antibody is labeled with europium and excited at 320 or 340 nm. Excitation of the europium labeled primary antibody allows time-resolved fluorescence energy transfer (TR-FRET) to a nearby *ULight* streptavidin bead attached to the biotinylated peptide.[60] The acceptor bead is labeled with a *ULight* dye that allows emission at 665 nm when in close proximity to the europium-tagged 1° antibody. This assay was used for the measurement of both SET7/9 methyltransferase writer protein and LSD1 demethylase eraser protein.[60] Figure 1.7 shows a cartoon of the principle of this technology.

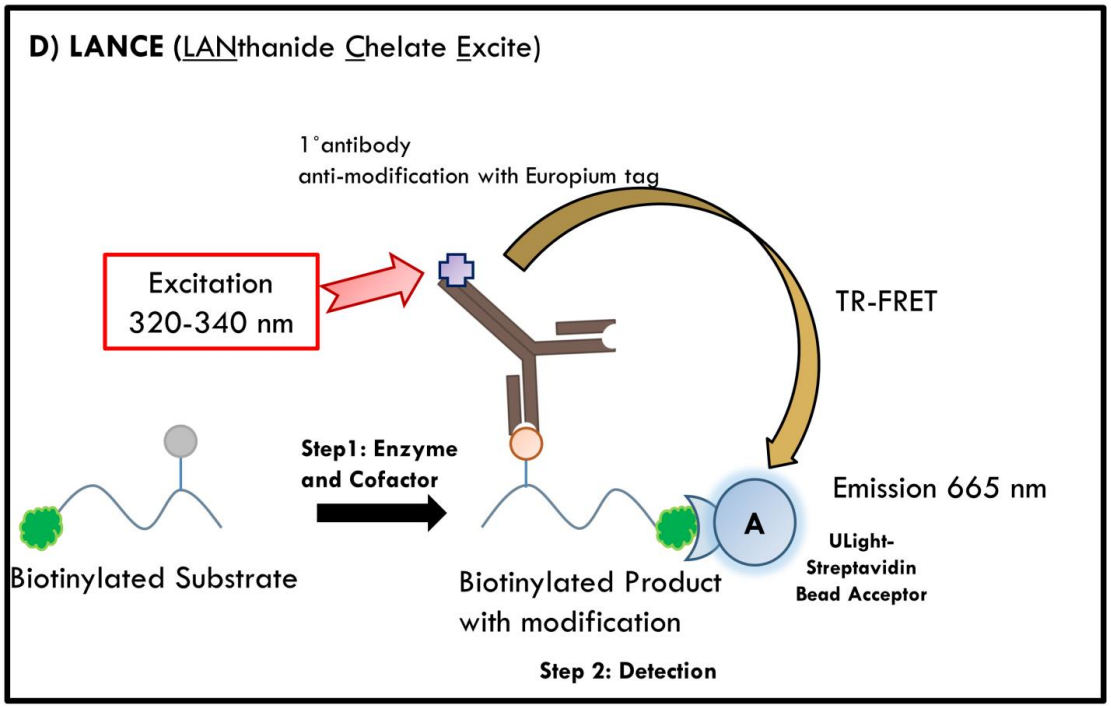


Figure 1.7 Cartoon depiction of the principle of the LANCE assay which is a single-bead based assay. Substrate peptide represents peptide before post-translational modification with the green cloud representing biotinylation of the peptide. Streptavidin coated bead (A) binds the biotinylated peptide and an anti-modification europium-tagged primary antibody specific to the post-translational modification is excited at 320-340 nm. Excitation of the europium-tagged

primary antibody causes emission of the *ULight* streptavidin bead acceptor at 665 nm by a time-resolved fluorescence resonance energy transfer (TR-FRET).

Table 1.2 shows a summary of antibody-based assays and describes the mode of detection, washing steps, high-throughput sensitivities, microplate size and analyte size of the different assays. The presence of a washing step indicates a heterogeneous assay (discontinuous) whereas the lack of one indicates a homogenous assay (continuous). The terms discontinuous and continuous will be described in a later section (section 1.2.5).

Table 1.2 Summary of antibody-based assays [51, 52, 54, 55]

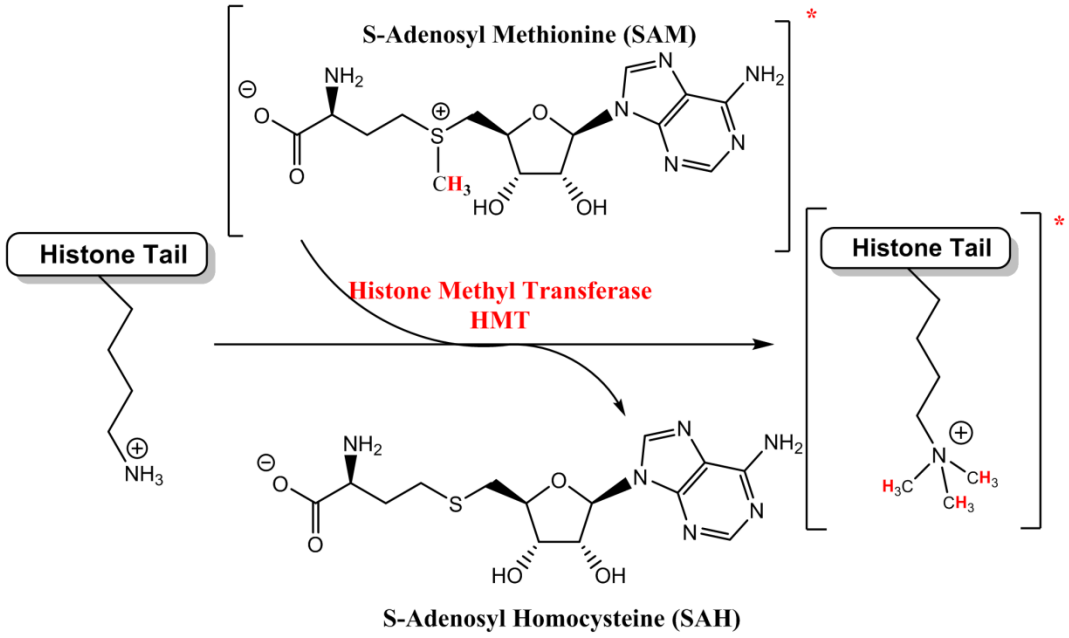
	A)ELISA	B)DELFIA	C)AlphaScreen	D)LANCE
Detection	Chemiluminescence	TRF	Luminescence Proximity	TR-FRET
Wash Step	Yes	Yes	No	No
Throughput	High	Medium	High	Ultra-high
Microplate	96, 384	96, 384	96, 384, 1536	96, 384, 1536
Analyte Size	Small molecules to whole cells	Small molecules to proteins	Small molecules to whole cells	Small molecules to peptides

1.2.2 Radiolabeling

Typically, for measurement of histone methyltransferase activity using a radiometric assay, the co-substrate S-adenosylmethionine (SAM) methyl group hydrogen atoms are replaced with radioactive tritium (^{3}H) atoms. Another option is ^{14}C labeling of the methyl group of SAM. Upon methylation with a specific histone methyltransferase, the radiolabeled methyl group on SAM will get transferred onto the product as shown in Scheme 1.4. Isolation of the product peptide from the reaction mixture (or of co-substrate SAM) allows for subsequent measurement of the amount of radiolabeled product, or the amount of radiolabelled SAM consumed, using a scintillation counter.[61, 62] While radiolabelling can be used for measuring enzyme activity in a quick a sensitive way, radiolabeled substances are costly for use and waste disposal, and also present

undesirable health risks. As with some of the antibody-driven methods, this kind of assay does not allow continuous monitoring of reaction progress.

In one radiochemistry-driven variant, the radioactive transfer of the label to a biotin-labeled product occurs in solution via the methyltransferase, the biotin-labeled product is then bound to a plate surface through avidin and washed to remove any free [^{3}H-labeled SAM]. The [^{3}H]-labeled product is then released and quantified by the scintillation counting method. The advantage is in better isolation of radioactive label from radioactive starting material, leading to a better overall signal. This method was used by Gowher *et al.* to detect the methyltransferase activity of Dim-5 enzyme to H3K9.[63] A new and improved radiolabeled-based method known as FlashPlate is described in section 1.2.2.1.



Scheme 1.4 Schematic of histone methyltransferase where ^{3}H-labeled SAM is transferred to the product peptide lysine group. S-adenosyl homocysteine (SAH) is formed as a by-product.

1.2.2.1 FlashPlate Assay (Assay E)

The FlashPlate assay is another variant that utilizes the same principle, but does not require any washing or filtering steps allowing the investigator to track enzymatic activity in real time.[64] The biotin-labeled peptide is bound to a streptavidin coated plate with wells coated with a thin layer of polystyrene-based scintillant. The radiolabeled co-substrate and enzyme are added to initiate the enzymatic reaction. The reaction progress is then monitored continuously by proximity scintillation counting. Free [^{3}H]-labeled SAM in solution only produces a strong signal when it gets to the bottom where the scintillant is found. Most free [^{3}H]-labeled SAM in solution is silent, except for the ones that get transferred to the substrate peptide (which is also bound to the bottom of the well). This technique was used by Dhayalan *et al.* to measure the activity of the writer protein G9a[65] and found it to be highly accurate and reproducible due to the decreased amounts of steps (pipetting, washing, filtering etc.).

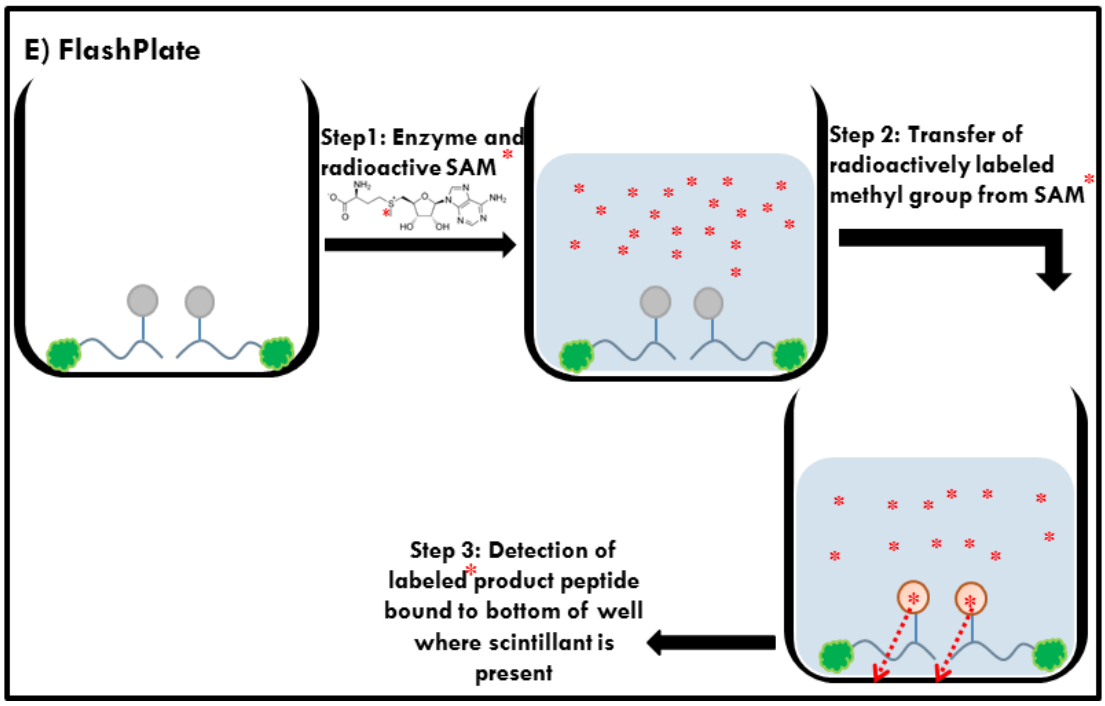


Figure 1.8 Cartoon depiction of the principle of the FlashPlate assay. Radiolabeled S-Adenosylmethionine (SAM) is added into a well coated with a biotinylated substrate peptide. Transfer of the tritiated ^{3}H to the product through a histone methyltransferase causes labeling of the product peptide as shown previously in scheme 1.4. Wells coated with a polystyrene-based scintillant allow measurement of high throughput signals during a course of time and subsequently measuring enzymatic activity.

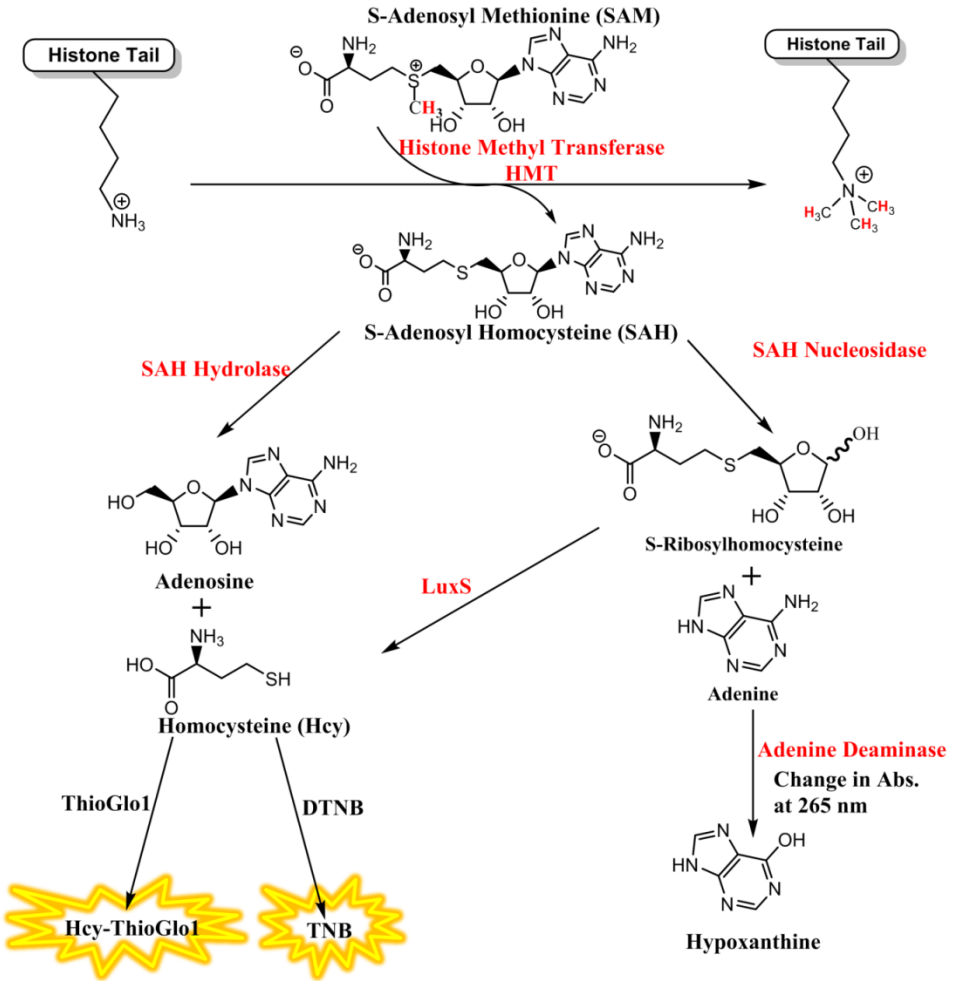
1.2.3 Mass spectrometry (Assay F)

Another technique used by researchers for the detection of PTMs is mass spectrometry (**MS**). Mass spectrometric assays for measuring enzyme activity are reliable and specific. They enable the investigator to distinguish between closely related methylation states but are relatively expensive in terms of instrumentation. MS allows the identification of different modifications on peptides as well as on intact proteins, as long as the modification has a unique mass and can be distinguished from starting material (and from other possible analytes in solution) on the basis of that mass. In one example, MS was used by Whetstine *et al.* to track enzymatic activity *in vitro*, specifically, a histone demethylase (JMJD2A). [66] This enzyme was shown to demethylate both H3K9me3 and H3K36me3 to H3K9me2 and H3K36me2 respectively corresponding to a loss of 14 Da (removal of CH_3 and addition of one H). This type of assay can be done in a continuous way in order to measure the activity of the enzyme in real time.[67] Sample preparation for a MS assay usually requires desalting to get rid of extra contamination but is fairly simple. Advances in MS instrumentation[68] and new experimental approaches[69, 70] open the door to explore more PTMs.

1.2.4 Coupled assays (Assay G)

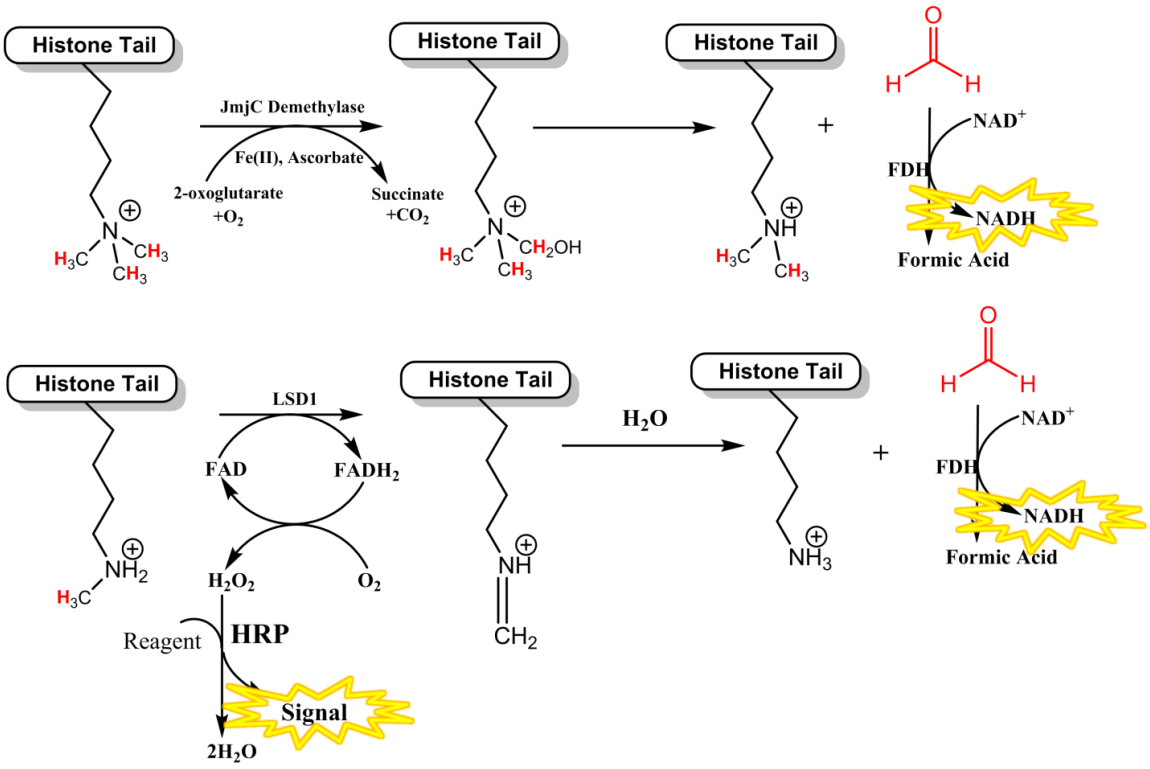
As shown in the radiometric assay above (section 1.2.2), methyltransferase enzymes require the transfer of the methyl group of SAM co-substrate to the lysine or arginine residues; upon the transfer, SAH is formed as a by-product. Due to some of the disadvantages of the radiometric assay, as previously mentioned, a new assay using this reaction process was developed by Collazo and co-workers.[71] A coupled fluorescence-based assay for SAM-dependent methyltransferases utilizes additional enzymes to quantify the generation of SAH as a direct proxy for the amount of methylated product formed (Scheme 1.5). The route from SAH to optical readout requires multiple additional reagents. SAH-Hydrolase (SAHH) catalyses the hydrolysis of SAH into adenosine and homocysteine (Hcy). Hcy free sulfhydryl group is subsequently reacted with the maleimido form of the fluorophore (ThioGlo1) which forms a highly fluorescent conjugate that emits at 515 nm.[72]. This assay allows the measurement of the enzymatic activity of histone methyltransferases without the use of radioactive compounds and

interruption of the procedure. Other coupled assay schemes exist, such as one developed by Hendricks and co-workers[73] using the same principle but using SAH Nucleosidase and LuxS enzymes to produce Hcy (Scheme 1.5), and Ellman's reagent (DTNB) as the agent used to react with Hcy in order to give an absorbance change at 412 nm. DTNB also reacts with the free sulhydryl group to form a mixed disulfide product and 3-thio-6-nitrobenzoate (TNB). TNB is coloured and can be measured using a spectrophometer.[74] SAH can also be utilized to form hypoxanthine to measure the enzymatic activity of methyltransferase proteins by observing the change in absorbance at 265 nm in real time.[75]



Scheme 1.5 Coupled enzyme assays used to measure SAM-dependent methyltransferase activity by transformation of SAH to products that can either be detected by fluorescence or UV-Vis spectroscopy.

Histone demethylase reactions (an example of eraser enzyme activity) generate different by-products, and can be measured using a different family of coupled assays that rely on the enzyme formaldehyde dehydrogenase (FDH). Demethylation of lysines with either JmjC domain demethylases or LSD1 yields one equivalent of formaldehyde that gets oxidized by FDH. FDH utilizes NAD^+ as a cofactor, which gets reduced to NADH during the reaction. The change in fluorescence[76] of NADH monitored at 465 nm provides a means to measure enzyme activity as it occurs *in vitro*.[77] Scheme 1.6 shows an example of this FDH-coupled assay for a Jmjc eraser enzyme and LSD1. Another option for measurement of an LSD1 demethylase is using a horseradish peroxidase (HRP)-coupled assay that uses the H_2O_2 produced during demethylation (see Scheme 1.6) to oxidize HRP substrates and produce a signal.[78] While these assays are useful in measuring demethylase activity, they do not provide specific identification of the product being formed.



Scheme 1.6 Coupled enzyme assays used to measure histone demethylase activity using an FDH-coupled assay and an HRP-coupled assay.

While coupled assays allow for a continuous measurement of methyltransferase or demethylase activity, careful selection of the components being used is necessary to minimize background fluorescence.[56] The addition of multiple components to allow for these assays to work means that further experimental optimization is required for each enzyme.[50] Although coupled assays have the advantage of being a radioisotope-free method, they too are not efficient assays for identification of the product being formed.

1.2.5 Summary of assays

In the above sections, a variety of assays were described that can be used to track enzyme activity *in vitro.* The assays can be described as either continuous or discontinuous. Table 1.3 shows a summary of the different types of assays presented. In general, a continuous assay is one that does not require any "time-lag" between steps—the reaction is initiated and the measurement of the enzymatic activity begins. These types of assays are desirable for high-throughput screen because they tend to require fewer, simpler steps, are easily automated, and provide continuous data about the modification occurring with the enzyme in question.[79] In discontinuous assays, the enzymatic reaction must be initiated, stopped, and then developed, thereby losing critical information about the particular enzyme in question. Although in the table below only ELISA and DELFIA assays are described as discontinuous, some radiometric assays (other than the FlashPlate assay) are also discontinuous in their nature due to the washing steps required. While having a continuous assay is important, not all continuous assays provide identification of the product being formed by the enzymatic reaction. Therefore, depending on the needs of one's experiment, certain assays may be chosen that provide continuity, product identification or both.

Table 1.3 Summary of assays presented indicating their ability to perform a continuous enzymatic assay or lack-thereof as well as specificity for product identification.

METHOD	CONTINUOUS?	SPECIFIC PRODUCT IDENTITY?
A) ELISA	**N**	**Y**
B) DELFIA	**N**	**Y**
C) ALPHASCREEN	**Y**	**Y**
D) LANCE	**Y**	**Y**
E) FLASHPLATE	**Y**	**N**
F) MASS SPECTROMETRY	**Y**	**Y**
G) COUPLED ASSAYS	**Y**	**N**

1.3 Motivating question: Can we develop a new method that allows the detection of post-translational modifications and use this method to track enzymatic modifications?

This chapter has introduced two main concepts: the variety and importance of post-translational modifications and the different assays developed to screen and measure, in real-time, the activity of the enzymes responsible for such modifications. As shown, current research relies on antibody-based assays, radiolabeling, mass spectrometry, and coupled assays. These methods have allowed for a better understanding of PTM enzymes. The fact that new variants are continually being developed points broadly to the importance of PTM enzymes in life sciences research, and also suggests that no one method has yet produced performance that is satisfactory to all end users. The aim of this thesis is to address this unmet research need with a supramolecular approach.

Prior to this research, the Hof group showed that the macrocyle *para*-sulfonatocalix[4]arene **PSC4** showed selectivity for methylated histone analytes. My motivation was to see if this macrocyle and its derivatives could be used to generate an optical signal for this fundamental binding event, and if such a set of supramolecular set of tools would allow me to detect post-translational modifications and to track the progress of enzymatic modifications.

1.4 Outline of the thesis

The following chapters describe the development of analytical tools based on this simple supramolecular binding motif. In Chapter 2, I report on a fluorescence-based method for the measurement of dissociation constants between a variety of supramolecular hosts and a particular histone tail modification. This serves as an important advance in our ability to measure such binding constants, and also as a forerunner to our efforts to track enzymatic activities. In Chapter 3, I report on my (ultimately unsuccessful) attempt to apply the fluorescence-based method developed in Chapter 2 to the continuous *in vitro* tracking of a histone methyltransferase reaction. I also report on the exploration of a second generation of calixarene-based supramolecular hosts, which ultimately allowed me to achieve this goal.

Chapter 2: Synthetic trimethyllysine receptors that bind histone 3, trimethyllysine 27 (H3K27me3): a dye displacement assay for convenient determination of equilibrium dissociation constants.

Parts of this Chapter have been previously published in a paper for which I was the first author.

Sara Tabet, Sarah F. Douglas, Kevin D. Daze, Graham A. E. Garnett, Kevin J. H. Allen, Emma M. M. Abrioux, Taylor T. H. Quon, Jeremy E. Wulff, Fraser Hof *Bioorg. Med. Chem*. (2013) *21* 7004–7010.

Link to paper: http://dx.doi.org/10.1016/j.bmc.2013.09.024

I conceived of the analytical method, planned binding experiments, collected and analyzed the data for direct titration and dye displacement assays, performed synthesis of peptides used as binding partners, and wrote the manuscript. The calixarenes used for assays in this Chapter were synthesized by Kevin D. Daze, Graham A.E. Garnett, and Kevin J. H. Allen. Complementary Fluorescence Polarization (FP) assays of these calixarenes were completed by Sarah F. Douglas and can be found in the published paper but are not included in this Chapter.

2.1 Introduction

The previous chapter illustrated the various covalent and non-covalent modifications that can occur on histone tails known as post-translational modifications (PTMs). Methylation (of Lys and Arg), acetylation (of Lys), phosphorylation (of Ser and Thr), lysine ubiquitilation and SUMOylation as well as non-covalent modifications such as *cis-trans* proline isomerization and others [17-21] are directly responsible for many downstream effects that include transcriptional repression or activation, chromatin remodelling, and DNA repair and recombination.[2, 14, 80] Proteins known as "writers" and "erasers" allow the addition and removal of such modifications respectively whereas "readers" allow the detection of these modifications.

Writer proteins EZH1/EZH2 are histone methyltransferases that enable the methylation of a lysine, in particular, Lys 27 of Histone H3. Trimethylated histone H3, lysine 27 (H3K27me3) is an epigenetic mark that is the focus of intense current interest in biomedical research due to its importance as a signalling element in multiple metastatic cancers.[40, 81-83] Histones bearing the H3K27me3 mark are generally associated with gene silencing by the downstream action of a multiprotein complex, called polycomb repressive complex 1 (PRC1).[84, 85] Recruitment of PRC1 occurs through the PRC1 component that is a H3K27me3-binding reader protein able to recognize the modification. Other proteins in PRC1 subsequently cause DNA methylation and stable silencing of the genetic information at that particular location in the genome.[84] The *Drosophila* parent of H3K27me3 reader module is called polycomb (the namesake of the entire pathway). In humans, there are five paralogs of polycomb called chromobox homolog (CBX), 2, 4, 6, 7, and 8 that can participate in different versions of the PRC1 complex. Despite their similar ability to bind H3K27me3 and to participate in polycomb-family gene silencing, each of these reader proteins is functionally distinct and operates at different areas in the genome.[84-88] The CBX7 reader protein is of particular interest, because CBX7 is specifically associated with the silencing of the gene for the tumour suppressors $p16^{INK4a}$ and $p14^{ARF}$ that are upstream controllers of Rb- and p53-mediated apoptosis respectively.[42, 89-91] In keeping with its role as a silencer of tumour suppressors, CBX7 expression is consistently shown to be strongly proliferative in castration resistant prostate cancer cell lines, embryonic and adult stem cells and in hematopoiesis and

lymphomagenisis.[42, 86, 88, 91-93] CBX7 is upregulated in prostate cancer upon progression from the androgen dependent state to the more aggressive androgen-independent state.[42] The molecular basis for targeting the H3K27me3-CBX7 complex is most clearly demonstrated by mutagenesis studies that show the complete blockage of proliferative signal when a single H3K27me3-binding residue of CBX7 is mutated.[42, 88, 92] The histone's trimethyllysine residue is a perfectly defined and potent hot spot for this protein–protein interaction, since CBX7 does not bind at all to unmethylated histone 3.[94]

In the natural protein–protein complex, the trimethyllysine residue is recognized and bound by an aromatic cage motif in CBX7 (Figure 2.1 A), which is a rigid pocket defined by Phe11, Trp32, and Trp35 (Figure 2.1 C).[92] Multiple cation–pi contacts between these pi-rich side chains and the methylated ammonium ion of Kme3 combine to drive complexation.[95] Mutation of Trp35 to an Alanine residue has been routinely used to shut down binding of H3K27me3 by CBX7 in various biological and biochemical studies, backing up the idea that these cation-pi interactions are critical for binding.[42, 88, 92] Previous work done by our group has shown how *para*-sulfonatocalix[4]arene (**PSC4**), which has a concave binding pocket made up of aromatic rings, is able to mimic in general CBX7's binding motif for methyllysines. This pocket-like macrocycle binds the methylated side chain of Kme3 via multiple charge–charge and cation–pi contacts.[96, 97] We have previously shown using a unimolecular, FRET-based biosensor assay that **PSC4** can disrupt the interaction of H3K27me3 with CBX7.[98]

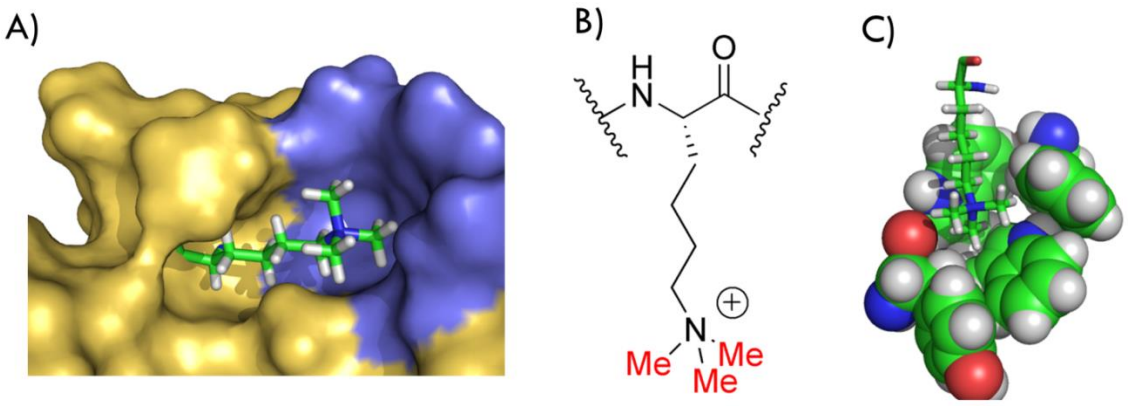


Figure 2.1 CBX7 and trimethyllysine. A) Trimethylysine binding in hydrophobic cavity (blue) of CBX7 protein (gold). B) Trimethyllysine free amino acid. C) Hydrophobic residues shown in blue in A) of CBX7 protein showing the binding of free amino acid trimethyllysine Pdb 2L1B.

Prior to my arrival, the Hof group showed that the **PSC4** macrocycle and its derivatives showed selectivity for methylated histone analytes and methylated peptides. Calculations of association constants were done using ITC and ^{1}H-NMR[96-99] and the dye displacement assay was used as a supramolecular sensor tool to generate distinct patterns of responses for different types of analytes. Possible applications for such methyllysine binding agents include both sensing applications and use as disruptors of various biological systems. Previous to this work, **PSC4** had been used as a component of an optical sensor by the Shinkai and Nau groups for sensing acetylcholine[100, 101] and histone methyltransferase activity.[102] These optical sensors operate using a dye displacement scheme as shown in Figure 2.2 where a fluorescent dye (such as lucigenin, **LCG**) emits when free in solution and is quenched upon titration of a host (modified *para*-sulfonatocalix[4]arene **PSC(X)**). Restoration of emission occurs upon addition of a guest molecule (H3K27 or H3K27me3 peptide tails) which acts as a competitor for binding within the host cavity.

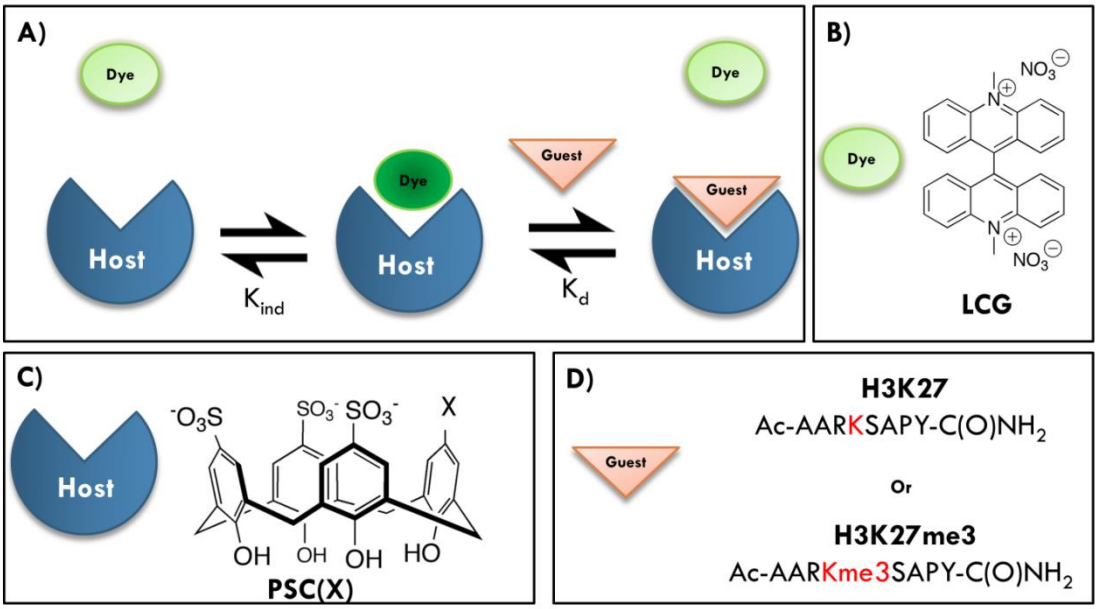


Figure 2.2 Indicator displacement assay for host-guest binding. A) Schematic of the indicator displacement assay showing first the quenching of the dye upon addition of the PSC(X) host (governed by the dissociation constant of the host-dye binding equilibrium (K_{ind})) followed by the competitive addition of an unmethylated **H3K27** or trimethylated peptide **H3K27me3** (guest) causing the release of the dye (governed by the host-guest dissociation constant (K_d)). B) Structure of lucigenin dye. C) Structure of host where **PSC(X)** represents hosts with different modifications on the upper rim of the calixarene. D) Unmethylated and trimethylated peptide sequences for representative guests.

Since we reported on **PSC4** as a methyllysine binder, other host-type macrocycles have been shown to bind Kme3 as the free amino acid[103] and within histone-tail peptide sequences[104, 105] representing an increasing interest in using structured macrocycles to target post-translational modifications. Three of these host-type macrocycles are shown in Figure 2.3. For example, the Macartney group [103] showed the binding of trimethylated lysine (Kme3) free amino acid with cucurbit[7]uril, **CB7** with a dissociation constant of 0.53 µM. Other studies by the Waters group showed that both polyanionic carboxylated cyclophane molecules **rac-A_2B**[104] and **A_2D**[105] bind trimethylated lysine within a peptide sequence with a dissociation constant of 25 µM and 3.9 µM respectively. But so far, those that operate on the more biologically relevant substrate (those within a peptide sequence) carry a major liability for biological studies in that they are composed of

disulfide-linked cycles that would not be stable in cells. Sulfonated calixarenes are chemically stable (e.g. they can survive heating in neat H_2SO_4), and have been shown to be stable in cells.[106-109] We aimed to make and study a set of calixarene derivatives that would retain these properties while also having tunable affinities for their methyllysine targets.

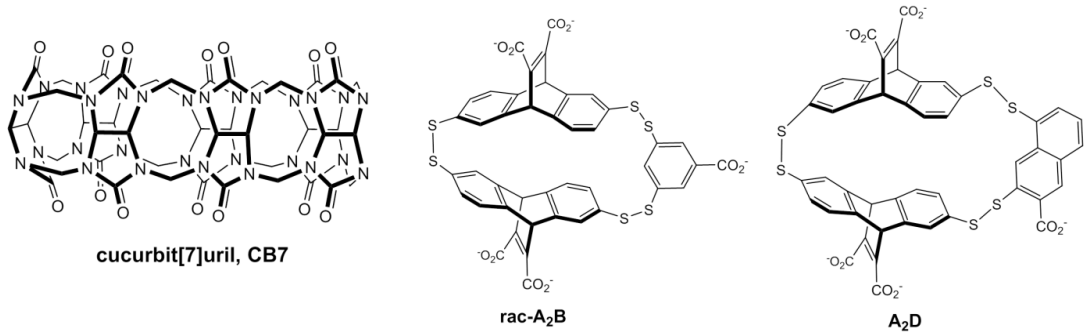


Figure 2.3 Macrocyles used for binding of post-translational modifications. CB7[103], rac-A_2B[104] and A_2D[105] macrocyles that bind methylated lysines.

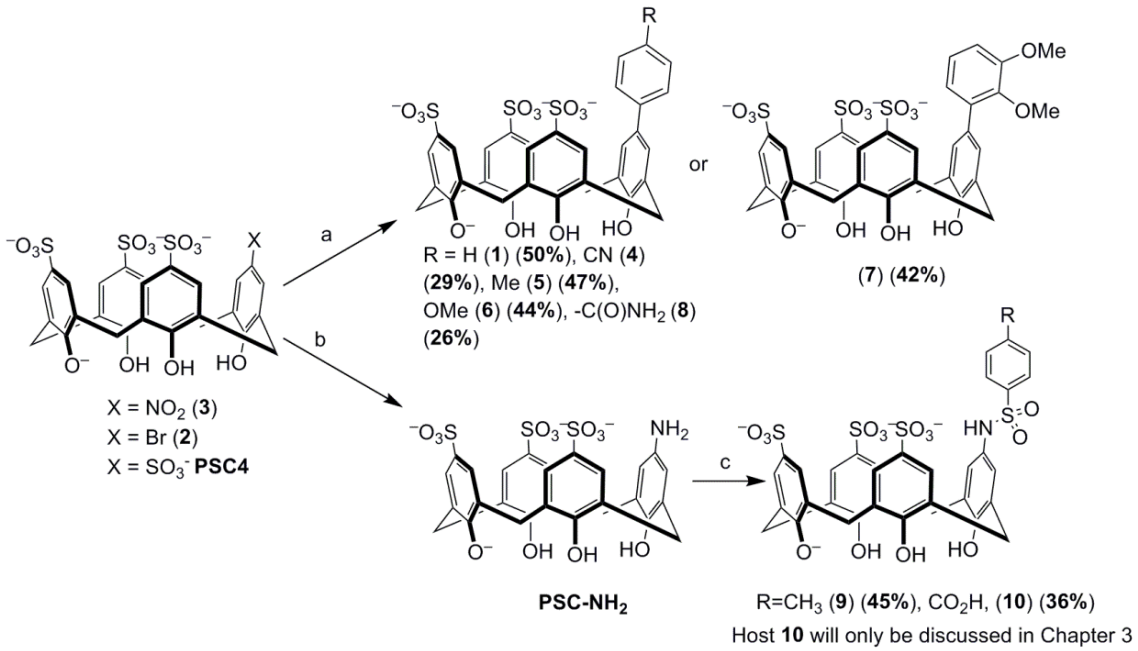
This Chapter reports the study of a set of macrocyclic compounds that constitute a new family of H3K27me3-targeting compounds that, unlike previously reported macrocycles, are easily modified to tune affinities and selectivities. The emphasis of this Chapter is on the development and use of a method for characterizing the direct binding of such compounds to peptidic partners using a competitive fluorescence-based dye displacement assay. Obtaining data on the binding affinities of these agents allows us to develop and understand structure-function relationships in a quick and easy way. The resulting structure–function relationships uncovered surprising aspects of molecular recognition for these macrocyclic agents.

2.2 Experimental methods

Synthesis — general

All reagents for synthesis were purchased from Aldrich and used as obtained. Lucigenin dye was purchased from Invitrogen or Sigma and stock solutions were prepared by sonicating the solid in distilled water and freezing at high concentrations. Dilutions for experiments were made as needed. The syntheses of compounds **1-4** and **8-10** have been previously published.[96, 98, 99] New hosts were made using **2** as starting material (see Scheme 2.1). All calixarenes and peptides were purified by HPLC (or HPLC-MS) on a preparative Apollo C18 column (Alltech, 5 μm, 22x250 mm) or preparative Luna C-18 column (Phenomenex, 5 μm, 21.2x250 mm), using a detection wavelength of 280 nm. Compounds were purified by running a gradient from 90:10 0.1% TFA in H_2O:0.1% TFA in MeCN to 10:90 0.1% TFA in H_2O:0.1% TFA in MeCN over 35 minutes. ESI-MS was performed on a Finnigan LCQ MS.

2.2.1 Synthesis of calixarenes[1]



Scheme 2.1 Synthesis of Sulfonato-calix[4]arenes 1-9. a) Ar-B(OH)$_2$, Na_2CO_3, TBAB, $Pd(OAc)_2$, µw, 2 hr, 150°C b) NaOH then RaNi, MeOH:H_2O (1:1), H_2, overnight c) TsCl, H_2O, 100 mM Na_2HPO_4, pH 8, overnight.

2.2.1.1 5-(4-methylphenyl)-25, 26, 27, 28-tetrahydroxy-11-17-23 trisulfonatocalix[4]arene (5).

Compound **2** (0.1011 g, 0.1360 mmol), 4-methylphenylboronic acid (0.0204 g, 1.1 equiv., 0.1496 mmol), $Pd(OAc)_2$ (0.0061 g, 20 mol%) and sodium carbonate (0.0548 g, 3.8 equiv., 0.517 mmol) were dissolved in 5 mL of deionized water inside a microwave vial, sealed, and heated to 150°C under microwave irradiation for 5 minutes with cooling air and stirring on. HPLC purification and evaporation of solvents *in vacuo* afforded a white powder in 47.5% yield (0.0489 g). Mp: 240°C (dec). IR (KBr pellet): 3350br, 1474s, 1457s, 1264w, 1211s, 1155s, 1113s, 1040s, 886w, 816w, 783m, 668m, 654m, 626m, 545m. ^{1}H NMR (500 MHz, D_2O): δ 7.83 (d, J=2.4 Hz, 2H), 7.75 (d, J=2.4 Hz, 2H), 7.60 (s, 2H), 7.19 (s, 2H), 6.80 (d, J=7.6 Hz, 2H), 5.70 (s, 2H), 4.06 (s, br, 8H), -0.68 (s, 3H). ^{13}C NMR (125 MHz, D_2O): δ 152.1, 150.8, 146.5, 136.8, 136.4, 134.3,

[1] Calixarenes with upper rim modifications of this sort will be referred to as 1st generation calixarenes in Chapter 2 whereas calixarenes that are covalently bound to a dye shown in Chapter 3 will be referred to as 2nd generation calixarenes.

133.9, 129.0, 128.8, 128.5, 128.4, 127.8 126.9, 126.8, 126.7, 126.5, 125.1, 30.7, 16.0. HR-ESI-MS: 753.07738 ([M-H]$^-$, $C_{35}H_{29}O_{13}S_3^-$; calcd 753.07758).

2.2.1.2 5-(4-methoxyphenyl)-25, 26, 27, 28-tetrahydroxy-11-17-23-trisulfonatocalix[4]arene (6).

Compound **2** (0.1080 g, 0.1454 mmol), 4-methoxyphenylboronic acid (0.0244 g, 1.1 equiv., 0.1606 mmol), $Pd(OAc)_2$ (0.0061 g, 20 mol%) and sodium carbonate (0.0551 g, 3.8 equiv., 0.519 mmol) were dissolved in 5 mL of deionized water inside a microwave vial and irradiated to 150°C for 5 minutes with cooling air and stirring. HPLC purification and evaporation of solvents *in vacuo* afforded a white powder in 43.9% yield (0.0491 g). Mp: 240°C (dec). IR (KBr pellet): 3245br, 1473s, 1457s, 1260w, 1239s, 1213s, 1180s, 1155s, 1114s, 1040s, 883w, 830w, 811w, 785m, 657m, 626m, 604m, 549m. ^{1}H NMR (500 MHz, D_2O): δ 7.78 (d, J=2.4 Hz, 2H), 7.71 (d, J=2.0 Hz, 2H), 7.27 (s, 2H), 7.06 (d, J=8.5 Hz, 2H), 6.03 (d, J=7.7Hz, 2H),), 4.06 (s, br, 8H), 1.60 (s, 3H). ^{13}C NMR (125 MHz, D_2O): δ 157.4, 151.8, 150.9, 146.9, 136.6, 136.2, 134.5, 131.6, 128.7, 128.5, 128.2, 128.0, 127.1, 127.0, 126.7, 126.5, 113.7, 52.8, 30.9, 30.6. HR-ESI-MS: 769.07107 ([M-H]$^-$, $C_{36}H_{31}O_{15}S_3^-$; calcd 769.07249).

2.2.1.3 5-(2,3-dimethoxyphenyl)-25, 26, 27, 28-tetrahydroxy-11-17-23-trisulfonatocalix[4]arene (7).

Compound **2** (0.041 g, 0.055 mmol), 2,3-dimethoxyphenylboronic acid (0.010 g, 1 equiv., 0.055 mmol), tetrabutylammonium bromide (0.0089 g, 0.5 equiv., 0.028 mmol), $Pd(OAc)_2$ (0.0025 g, 20 mol%) and sodium carbonate (0.026 g, 3.8 equiv., 0.209 mmol) were dissolved in 5 mL of deionized water inside a microwave vial and irradiated at 150°C for 5 minutes with cooling air and stirring. The aqueous solution was washed with CH_2Cl_2 (2x20 mL) then EtOAc (1x25 mL) and concentrated. HPLC purification and evaporation of solvents *in vacuo* afforded a white powder in 42% yield (0.018 g). Mp: 245°C (dec). IR (KBr pellet): 3366br, 1465s, 1261w, 1213s, 1160s, 1118s, 1042s, 889w, 784m, 661m, 625m, 559m. ^{1}H NMR (500 MHz, D_2O): δ 7.64 (d, J=2.1 Hz, 2H), 7.62 (d, J=2.0 Hz, 2H), 7.37 (s, 2H), 7.10 (s, 2H), 6.96 (t, J=7.8 Hz, 1H), 6.82 (d, J=7.8Hz, 1H), 6.76 (d, J=6.9Hz, 1H), 3.99 (m, br, 8H), 3.70 (s, 3H), 2.48 (s, 3H). ^{13}C NMR (125 MHz, D_2O): δ 153.3, 152.3, 150.9, 147.8, 145.2, 135.4, 135.3, 134.4, 131.4, 130.0, 129.1,

128.5, 128.2, 128.0, 126.6, 126.5, 126.4, 124.9, 122.3, 112.1, 59.6, 56.0, 30.7, 30.5. HR-ESI-MS: 799.07935 ([M-H]$^-$, $C_{36}H_{31}O_{15}S_3^-$; calcd 799.08303).

2.2.2 Peptide synthesis

All reagents used for peptides synthesis were purchased from ChemImpex or Sigma Aldrich except for Fmoc-Lys(Me3)-OH Chloride which was purchased from GL Biochem. Histone 3 peptides (H3K27 = Ac-AARKSAPY-C(O)NH_2, H3K27me3 = Ac-AARKme3SAPY-C(O)NH_2) were synthesized using the standard Fmoc solid-phase peptide synthesis protocol[110] as implemented on a CEM Liberty 1 microwave-based peptide synthesizer on Rink amide resin (ChemImpex). All sequences had a tyrosine introduced at the C-terminus to facilitate UV detection during HPLC purification.

Briefly: Peptides were synthesized on a 0.1 mmol scale on Rink amide resin. Alternating cycles of Fmoc deprotection and HBTU-mediated amino acid coupling were performed according to the default instrument protocols. Coupling solutions used by the peptide synthesizer included DIPEA in NMP (activator base solution) and HBTU in DMF (coupling reagent solution). Deprotection was performed with 20% piperidine in DMF. N-terminus of H3K27 and H3K27me3 peptides were acetylated off-line in a glass reaction vessel using 30:20:50 Pyridine:Acetic Anhydride: DCM for 1 hour at room temperature with occasional stirring. Peptides were cleaved off-line in a glass reaction vessel using 95% TFA/2.5% triisopropylsilane/2.5% H_2O for 2 hours at room temperature with occasional stirring. The cleaved mixture was rotovapped and precipitated in 45 mL of cold ether. Peptides were purified by preparative reversed-phase HPLC on a Apollo C18 column (Alltech, 5 µm, 22x250 mm) or a preparative Luna C-18 column (Phenomenex, 5 µm, 21.2x250 mm), using a gradient starting from 90:10 H_2O:MeCN (0.1% TFA) and running to 10:90 H_2O:MeCN (0.1% TFA) at a flow rate of 10 mL/min. Elution of the peptides was monitored at 280 nm and fractions were lyophilized to powder and characterised by ESI-MS.

Peptides were used in assays without desalting. Stock solutions were made up using the cuvette reader accessory on the SpectraMax® M5 / M5e Microplate Reader with 700 µL quartz cuvettes. The extinction coefficient used was 1490 $M^{-1}cm^{-1}$ with a path length of 1 cm and absorption at 280 nm.

2.2.3 Protocols for binding constant determinations

2.2.3.1 K_{ind} determination — direct titrations for calixarene-dye affinities

Samples for the direct titration were prepared in NUNC 96 black-well plates with an optically clear bottom, and were composed of 0.01 M of phosphate buffer (Na_2HPO_4/NaH_2PO_4) at pH 7.4, 500 nM of lucigenin, and varying concentrations of hosts (0-5 µM) made up with distilled water to a final volume of 200 µL. Emission spectra from 445-645 nm using a SpectraMax® M5 / M5e Microplate Reader were collected at λ_{ex} 369 nm. All experiments were performed in duplicate. Calixarene-LCG K_{ind} values were determined by plotting emission intensity (dF_{obs}) as a function of calixarene concentration [Ht] and fitting the data to the following expression[111] using Origin:

Eq. 2.1 $dF_{obs} = (F_{max} - F_{min})*[(D + Ht + (1/K_{ind})) - sqrt ((D + Ht + (1/K_{ind}))^2 - (4D*Ht))/2D]$ Where; y equals the change in fluorescence ($dF_{obs}= F_{obs}- F_{min}$) and x equals the total host concentration Ht ([calixarene]t= 0-5 µM). Parameters, F_{min} and K_{ind} were adjustable where F_{min} equals the minimum fluorescence of dye when saturated with host. D and F_{max} were treated as constants at 0.5 µM and the maximum fluorescence of dye without presence of host respectively.

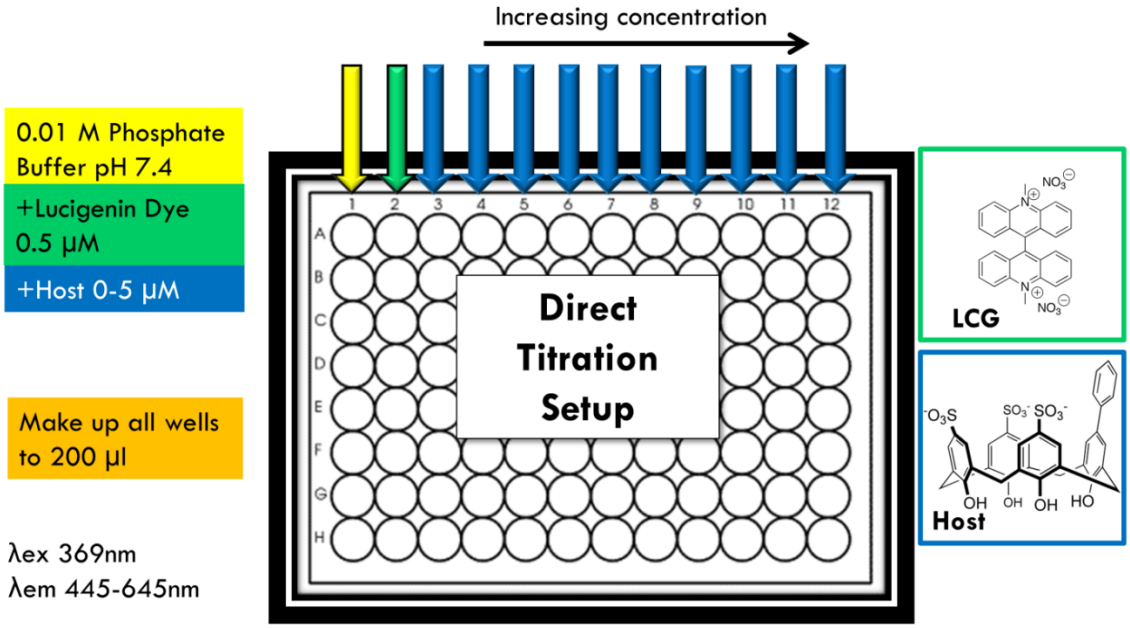


Figure 2.4 Direct titration plate setup using a 96-well Nunc plate with black sides and optically clear bottom. This setup shows the initial addition of phosphate buffer to the first blank well and the subsequent addition of dye then host in increasing concentrations. The dye used for all of these experiments was lucigenin (**LCG**) however numerous hosts (calixarene macrocycles) were used. The figure only shows example host **1** in the blue box. Plate was read using a SpectraMax® M5 / M5e Microplate Reader at λ_{ex} 369 nm and λ_{em} 445-645 nm at room temperature.

2.2.3.2 K_d determination — competition experiments to determine calixarene-peptide affinities

Samples for the dye displacement assay were prepared in NUNC-96 black well plates with optically clear bottom, and were composed of 0.01 M of phosphate buffer (Na_2HPO_4/NaH_2PO_4) at pH of 7.4, 500 nM of lucigenin, 1.25 µM of calixarene, and varying peptide concentrations (H3K27 or H3K27me3, 0-1.5 mM) made up with distilled water to a final volume of 200 µL. Emission spectra were collected as above. All experiments were performed in duplicate. Calixarene-peptide K_d values were determined by plotting emission intensity (dF_{obs}) as a function of peptide concentration $[G]_t$ and fitting the data with the program Origin using an expression and accompanying cubic equations initially derived by Nau and co-workers.[112-114] (These expressions were subsequently adapted and shown to be useful also for UV-Vis data.[115])

Equations used for fitting of dye displacement data:

Eq. 2.2 $dF_{obs} = Ei + ((Ehi-Ei) * ((K_{ind} * H)/(1 + (K_{ind} * H))))$

where dFobs = $F_{obs} - F_{min}$ (change in emission intensity), Ei = emission intensity of the uncomplexed dye and kept as an adjustable parameter, Ehi = the inherent emission intensity of the calixarene-dye complex and was set to 0, K_{ind} = association constant for host-dye complex was determined by previous 1:1 direct titration (see above **Section 2.2.3.1**) and held constant, and H= free concentration of host defined as below and used for iterative fitting with an initial guess of $H = H_t$. The term x is incorporated in the term H as shown below.

Fitting was achieved by iterative nonlinear least squares regression using the cubic step equation:

$(aH^3 + bH^2 + cH + d) / (3aH^2 + 2bH + c)$
for (H=Ht, step=1; abs(step)>1E-15; H=H-step)
$a = K_{ind} * K_d$;
$b = K_{ind} + K_d + K_{ind} * K_d * It + K_{ind} * K_d * x - K_{ind} * K_d * Ht$;
$c = 1 + K_{ind} * It + K_d * x - (K_{ind} + K_d) * Ht$;
$d = -Ht$;

where x = total guest concentration [G]t, which was varied from 0–1.5 mM for H3K27 and 0–0.5 mM for H3K27me3, It = total concentration of dye and held constant at [lucigenin]t = 0.5 µM, Ht = total concentration of host and held constant at [calixarene]t = 1.25 µM and K_d = association constant of the calixarene-guest complex and was kept as an adjustable parameter.

The binding constants from Origin software are calculated as *association values* but I will be referring to them in this Chapter as *dissociation values*.

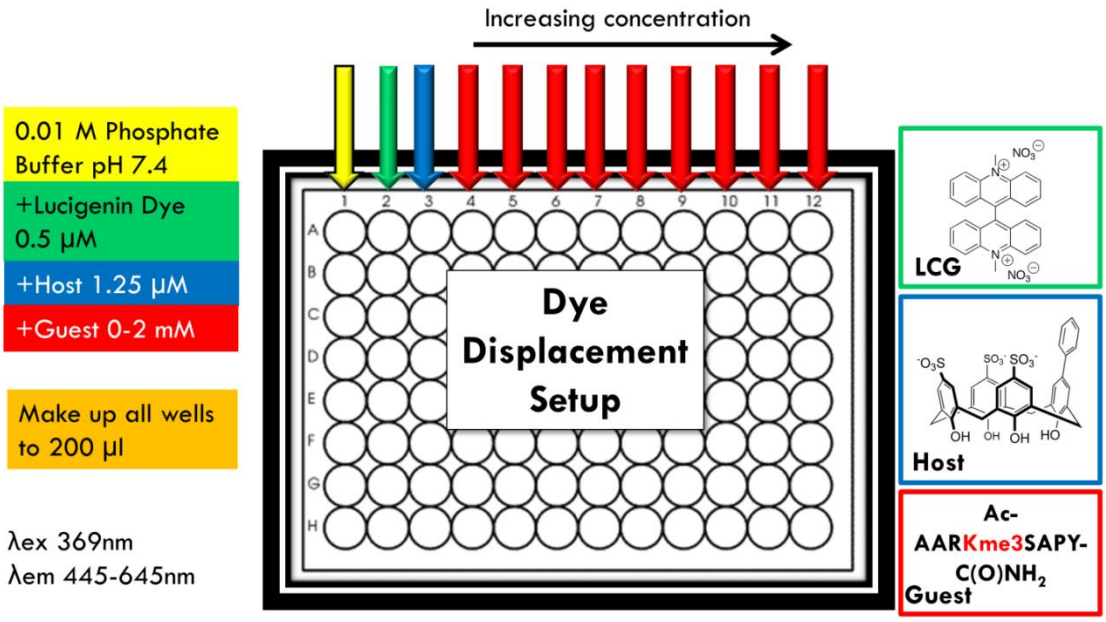


Figure 2.5 Dye displacement plate setup using a 96-well Nunc plate with black sides and optically clear bottom.This setup shows the initial addition of phosphate buffer to the first well and the subsequent addition of dye then host at constant concentration of 1.25 µM and increasing concentration of guest molecule. The dye used for all of these experiments was lucigenin (**LCG**) however numerous hosts (calixarene macrocycles) were used. The figure only shows host 1 in the blue box and the H3K27me3 guest as an example for setup. Plate was read using a SpectraMax® M5 / M5e Microplate Reader at λ_{ex} 369 nm and λ_{em} 445-645 nm at room temperature.

2.3 Results

2.3.1 Direct titration of lucigenin dye (LCG) and modified calixarene macrocyles

In order to carry out dye-displacement based studies of binding between calixarenes and different peptide binding partners, it was first necessary to determine the affinity of each calixarene for the indicator dye, lucigenin. As shown in Figure 2.4, a direct titration of each calixarene into lucigenin at a fixed concentration was conducted. Increasing host concentration caused reduced emission relative to the control well containing buffer and lucigenin only. Figure 2.6 A shows some exemplary data for the decrease in fluorescence upon increase in host concentration. The difference in fluorescence at the maximum emission wavelength between the buffer-dye well and the buffer-dye-host well is the dFobs in Eq. 2.1. These values were used in the software

Origin for the iterative fit that determined the value for K_{ind} for each host (see section 2.2.3.1). Figure 2.6 B shows a comparison of different binding curves for two different hosts, host **1** and host **9**, which shows qualitatively how the binding curve for a stronger binding host (**1**) and a weaker binding host (**9**) approach saturation at different rates. All curve fits were satisfactory, suggesting that each host binds this dye with the expected 1:1 host-guest ratio. K_{ind} values range from 0.0132 µM to 1.91 µM, and are reported in Table 2.1.

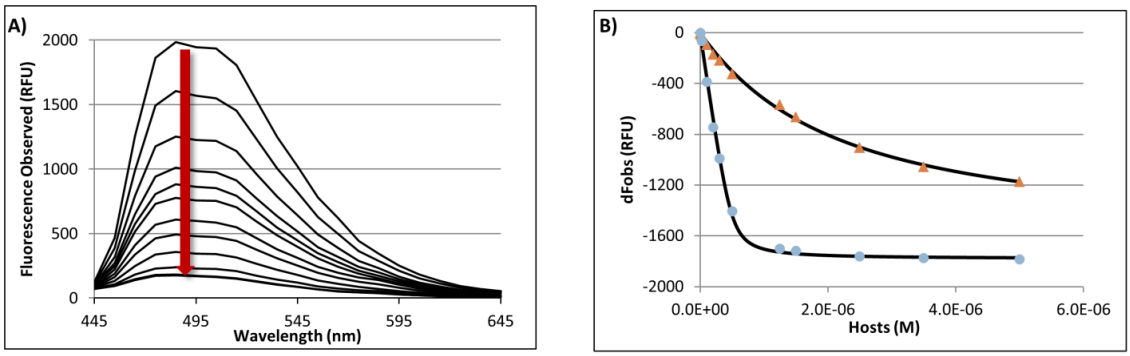


Figure 2.6 Direct titration assay for lucigenin-calixarene binding. A) Direct fluorescence titration of LCG (0.5 µM) with host **7** in 0.01 M phosphate buffer at pH 7.4, λ_{ex}= 369 nm. B) Direct titration curves for two different hosts, host **9** orange ▲ and host **1**, blue ● from high to low dissociation constant K_{ind} at 485 nm respectively.

2.3.2 Dye displacement for calixarene - guest molecules H3K27 or H3K27me3

The setup of the experiment was similar to that above. As shown in Figure 2.5, wells were made up to 200 µL with constant host and dye concentrations, and varying guest concentrations. The guests used in this Chapter were either **H3K27** or **H3K27me3**, but this method has been subsequently shown to work for a wide variety of different peptidic guests of different sequences and bearing different modifications (data not shown). Increasing guest concentration caused increased emission relative to the control well containing buffer, lucigenin and host. Figure 2.7 A shows the increase in fluorescence upon increase in guest concentration for host **5** and **H3K27me3**. The difference in fluorescence at the maximum emission wavelength between the buffer-dye-host well (now considered blank or F_{min}) and the buffer-dye-host-guest well is the dFobs in Eq.2.2, and these values are used in the software Origin which outputs data that allows the calculation of the association constant of the host-guest complex. Figure 2.7 B shows

a comparison of different binding curves for three different hosts, host **1, 5,** and **7** for both unmethylated **H3K27** and trimethylated **H3K27me3** peptides.

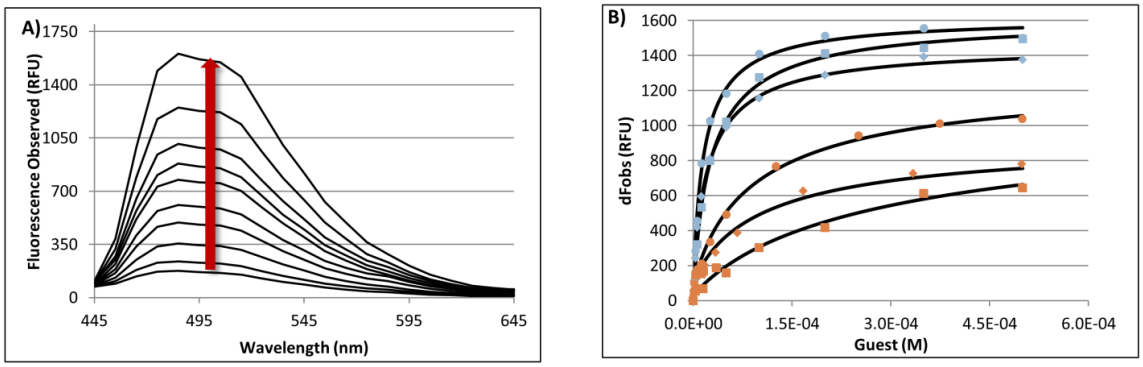


Figure 2.7 Dye displacement assay for calixarene-peptide binding. A) Exemplary dye-displacement assay curves done using 0.01 M phosphate buffer at pH 7.4, λ_{ex}= 369 nm, λ_{em}= 485 nm used for the calculation of the dissociation constant K_d at 500 nM lucigenin dye, 1.25 µM of host **5** and varying **H3K27me3** concentrations. B) Dye displacement curves for three different hosts, host **5** ●, host **1** ■, and host **7**♦ with varying concentrations of unmethylated (orange) or trimethylated peptide (blue) showing the different selectivities of each host to a particular peptide.

Two comparisons can be made from Figure 2.7 B. One is the difference in fluorescence for a single host with unmethylated and trimethylated peptides and therefore the strength of the guest at displacing the dye from the host cavity based on the presence or absence of a post-translational modification. The second is the difference in fluorescence for a single peptide with different hosts, therefore the effect of modifications done on the host's upper rim. Table 2.1 shows the dissociation constants for all nine hosts **1-9** and **PSC4** for **H3K27** and **H3K27me3** peptides. Selectivity of **H3K27me3/H3K27** was also calculated to determine which hosts would most selectively bind the trimethylated **H3K27me3** peptide.

Table 2.1 Activities of trimethyllysine-targeting compounds as determined by dye displacement assay.[a]

Calix[4]-arenes	Substituents	K_{ind} for LCG (μM)[b]	K_d for H3K27 (μM)[c]	K_d for H3K27me3 (μM)[c]	H3K27me3/ H3K27 selectivity
PSC4[d]	X=SO_3^-	n.d.	220±7[d]	5.4±0.1[d]	40
Host 1	X= Ph	0.055±0.011	19±4	0.75±0.18	25
Host 2	X= Br	0.085±0.025	13±8	0.70±0.17	19
Host 3	X= NO_2	0.460±0.04	30±60	10±13	3
Host 4	X= Ph(CN)	0.050±0.020	11±3	0.88±0.12	13
Host 5	X= Ph(Me)	0.013±0.013	2.7±1.0	0.34±0.03	8
Host 6	X= Ph(OMe)	0.120±0.01	20±9	0.78±0.22	26
Host 7	X= Ph(2,3-OMe)	0.103±0.017	25±14	2.13±0.5	12
Host 8	X= Ph($CONH_2$)	0.124±0.019	23±16	1.86±0.6	12
Host 9	X=$NHSO_2$Ph(Me)	1.91±0.3	> 100	> 100	n.d.

a) All K_d values are averages from duplicate determinations. b) Conditions for direct titration of calixarene into **LCG**: [LCG] =0.5 μM with host at varying concentrations between 0-5 μM in 0.01 M phosphate buffer at pH 7.4, λ_{ex} 369 nm , λ_{em} 485 nm done in duplicates. c) Conditions for the competitive titration of peptide into host/LCG solution: [**LCG**]=0.5 μM, [host]=1.25 μM with varying peptide concentrations of [**H3K27**]=0-1.5 mM and [**H3K27me3**]=0-0.5 mM in 0.01 M sodium phosphate buffer at pH 7.4, λ_{ex} 369 nm , λ_{em} 485 nm. d) Data for **PSC4** taken from reference.[98]

Although a similar assay had been previously used and validated to be accurate for **PSC4** and a variety of different cationic guests,[101] we further tested the accuracy of the K_d values for our guests determined by this method by using ITC.* Table 2.2 shows the comparison in dissociation constant for unmethylated **H3K27** and trimethylated **H3K27me3** with host **1** using both the dye-displacement method and ITC.

Table 2.2 Binding data for 1 determined by ITC and dye displacement methods.

Calixarene	Guest	N	K_d (μM)	ΔH (kJ mol^{-1})	–TΔS (kJ mol^{-1})
Host 1	H3K27	0.85	51±8	–9.73±1.52	–15.1
	H3K27[a]		19±4		
	H3K27me3	1.3	4±0.4	–27.1±0.18	–7.76
	H3K27me3[a]		0.75±0.18		

Averaged values determined by duplicate ITC titrations at 303 K in buffered H_2O (10 mM Na_2HPO_4/NaH_2PO_4, pH 7.7). Calixarene concentration in syringe was 0.750-1.5 mM titrated into buffer-matched guest solution at a concentration between 50-70 μM. Values determined using one-site binding model with supplied Origin software. a) Values determined by dye displacement assay are shown here for comparison.

2.4 Discussion

2.4.1 Choices and synthesis of compounds

Our earlier trimethyllysine-targeting compounds[98] were based on a synthetic pathway that created calixarene derivatives functionalized at the 'lower rim.' We found these compounds inherently poor at tuning affinities for trimethyllysine-containing peptides and proteins, simply because the lower rim is distal to the actual binding surface of the macrocycle. Others in the lab subsequently developed a new synthetic route to functionalization at the 'upper rim' and showed using simple amino acids that groups introduced at this position are in direct contact with the bound Kme3 residue, and therefore have a strong influence on binding affinities and selectivities.[96] The syntheses of compounds studied as H3K27me3 binders in this Chapter are shown briefly in Scheme 2.1 but will not be discussed in detail here. All of the compounds prepared and tested

* ITC data was collected by Kevin D. Daze.

here are characterized by having a desymmetrized upper rim bearing three sulfonates and a single, appended aromatic substituent that can modify interactions with the H3K27me3 target.

2.4.2 Dye displacement assay for unmethylated and trimethylated histone tail peptide

Previous studies have shown that the dye lucigenin binds to different, unadorned sulfonated calixarenes, including **PSC4** and its calix[5]arene and calix[6]arene analogues.[101, 116] Upon binding of lucigenin into the calixarene's electron-rich cavity, fluorescence decreases and the dissociation constant of the dye for host (K_{ind}) is determined by curve fitting of emission intensity as a function of concentration of host (Figure 2.6). This is thought to occur through a static quenching mechanism due to electron transfer from the electron-rich calixarene to the electron-accepting dye.[101] **LCG** has been studied by the Nau group and shown to bind in a 1:1 stoichiometry to the aforementioned simple calixarenes where one of the two acridine rings of **LCG** is included into the cavity of the calixarene.[101]

In this work, we found **LCG** to be generally well suited to use in a dye-displacement format with all of the upper-rim substituted calixarenes that we studied. In each case, addition of a calixarene caused a decrease in **LCG** emission intensity that could be fit to a 1:1 binding isotherm. Based on the properties of the host, the dissociation constant may increase or decrease yielding a weaker or stronger interaction of dye-host complex, respectively. We found that this indicator could be used, regardless of its own K_{ind} value, as long as the host-dye K_{ind} value is approximately 3 orders of magnitude of the host-guest K_d. This was true for all hosts studied here, with the exception of sulfonamide host **9** which will be discussed later. Based on these results, we proceeded to use **LCG** as an indicator for a quantitative assay for **H3K27me3** and **H3K27** binding to this wide variety of calixarenes, according to the schematic in Figure 2.2.

A constant concentration of the host is chosen from the direct titration and used for back titration of a methylated (**H3K27me3**) or unmethylated (**H3K27**) peptide. The host concentration selected for the experiments done in this chapter were all 0.5:1.25 dye:host where a strong fluorescent quenching is observed. As the peptide binds to the calixarene and displaces the dye, the fluorescence increases as shown in Figure 2.7 A.

Both unmethylated and methylated peptides are able to displace the dye and output a curve for the calculation of a dissociation constant K_d (Figure 3.7 B). Since **LCG** is known to be quenched by various halide anions, neutral amines, and other electron donating species in solution,[117] we avoided the use of such counterions as much as possible in our choice of binding partners and buffers. We found that some peptides at high concentrations can cause quenching of **LCG** themselves depending on the counterions present. Using our **H3K27** and **H3K27me3** peptides, as prepared, at relatively low concentrations allowed us to carry out the experiments without the need for desalting.

In order to verify the fluorescence-based dye displacement method for the calculation of dissociation constants for this system, Kevin Daze from our group conducted ITC titrations for host **1** and both the unmethylated **H3K27** and methylated **H3K27me3** peptides using the same 10 mM phosphate buffer at pH 7.7 which is almost identical to the buffer I used in the dye displacement assays (10 mM phosphate buffer at pH 7.4). As seen in Table 2.2, the dissociation constants for both **H3K27** and **H3K27me3** with host **1** were similar, but not identical, for both methods. This comparison showed a 3-fold difference for the H3K27-host **1** pair and a 5-fold difference for the H3K27me3-host **1** pair. The differences in dissociation constants can be understood due to differences in absolute concentration of calixarene in each method. A much higher calixarene concentration of 0.750-1.5 mM was needed for ITC measurement (compared to 1.25 μM in the dye displacement assay). At such higher concentrations, the contributions of additional, weaker binding equilibria such as host dimerization and/or the formation of 2:1 host-guest complexes might possibly occur. Similar discrepancies were seen in a related system done by Nau and co-workers.[101]

The calculation of K_d values for all other host-guest pairs for hosts **1-9** and guests **H3K27** and **H3K27me3** was completed. In general, the dissociation constants for hosts with **H3K27me3** were lower (stronger binding) than with **H3K27** unmethylated peptide, which was expected since previous work has shown selectivity for trimethyllysine over unmethylated lysine.[96-98] This trend is easily visible in the data shown in Figure 2.7 B, where the curves with the blue data (**H3K27me3**-host pairs) indicate a higher

fluorescence, therefore more dye being displaced than the curves with the orange data (**H3K27**-host pairs).

The affinities of each compound for **H3K27** and **H3K27me3** are reported in Table 2.1. Data were reproducible from replicate to replicate, and satisfactory fits for all curves were obtained except for those of compound **9**, which is a very weak binding compound (see below), and compound **3**, which itself contains a *p*-nitrophenol (PNP) chromophore in its structure. We were unable to obtain fits for compound **9** at all, whereas compound **3** fits were done but had a lower R^2 value than for other hosts. The K_d values for the target histone tail **H3K27me3** range from 0.34 μM to 10 μM. For comparison, the reported *in vitro* binding affinities of naturally evolved Kme3-reader proteins for trimethyllysine histone elements are in the range of 0.7-110 μM,[94, 95, 118] which indicates that these synthetic agents are generally as strong as the naturally evolved binding partners. All compounds showed a preference for the methylated peptide **H3K27me3** over unmethylated **H3K27**, with the magnitude of selectivity ranging from 3- to 26-fold. The most selective host was methoxy-substituted compound **6**.

While the dissociation constant of host **6** with **H3K27me3** is not the smallest value (strongest binder) out of the nine hosts tested, a decreased affinity for the unmethylated **H3K27** peptide allows for the best selectivity. Selectivity data for natural proteins is hard to obtain from the literature because affinities of reader proteins for unmethylated peptides are often not reported, or are too weak to measure under the conditions of a given biological assay. Some reported trimethylated/unmethylated selectivities exist, and seem to point to a general limit of >100-fold for trimethyllysine/lysine selectivities.[41, 95] One exception is the PHD fingers of reader protein CHD4, which binds trimethylated and unmethylated H3K9 partners with nearly equal affinities.[119] In any case, it seems that most of the naturally evolved proteins have better selectivities for methylated over unmethylated partners than do these synthetic compounds. The slight erosion of selectivity of the aryl-substituted compounds relative to **PSC4** seems to arise more generally from an *increased* affinity for unmethylated **H3K27** (relative to **PSC4**) than from any general change in affinities for the targeted **H3K27me3**. The compounds all make complexes with Kme3 in which the quaternary ammonium ion is buried deep within the macrocycle's binding pocket, and the appended aromatic rings

are in close contact with the trimethyllysine side chain's methylene groups, as previously confirmed for representative members of this class by ^{1}H NMR.[96] This structural model can explain the increased affinities for **H3K27** upon introduction of aryl groups, because the aryl groups make similar favourable contacts with the methylenes of *both* trimethyllysine and unmethylated lysine.

The appended aromatic rings have substituents with varying electron-donating or -withdrawing ability. We also looked at our data for the expected connections between ring electronics and strength of host-guest cation-pi interactions; i.e., that more electron-rich rings (e.g. methoxy-substituted **6**) would form stronger cation-pi interactions with the lysine/trimethyllysine side chains of the histone peptides than would a more electron-poor ring (e.g. cyano-substituted **4**). These trends are not present in the data for any aryl-appended host. For example, MeO-substituted **6** and CN-substituted **4** have equal affinities for **H3K27me3**, and similar affinities for **H3K27**. *But strong effects do exist for substituents on the main macrocycle*—Compound **3** (nitro-) differs from compound **2** (bromo-) by only a single group, but has ~15-fold lower affinity for **H3K27me3**. This suggests that a strong cation-pi interaction does exist for the aryl groups of the main macrocycle that does not exist for the aryl groups appended to the upper rim. Compound **7** (2,3-dimethoxy-) is significantly weaker than compound **6** (4-methoxy-), which is in this case best explained by the conformational differences induced by the presence of an *ortho*-substituted biaryl linkage in **7** that is not present in **6**. Compound **9**—the only member that bears a sulfonamide linked aryl group tested in this Chapter—has very low affinities for either methylated or unmethylated targets. One possibility is that the sulfonamide linker disrupts binding. But we also see evidence of self-inclusion in the ^{1}H NMR spectrum of **9** (a strongly upfield-shifted CH_3 resonance that indicates inclusion of the methyl group in the binding pocket of the calixarene), which offers an alternative explanation for the poor activity of **9** shown in Figure 2.8 A. Inclusion of the methyl group of host **9** into the binding pocket may cause difficulty for one of the acridine rings of **LCG**[101] to bind during direct titration and for the positively charged peptides during dye –displacement assay (See Figure 2.8 B).

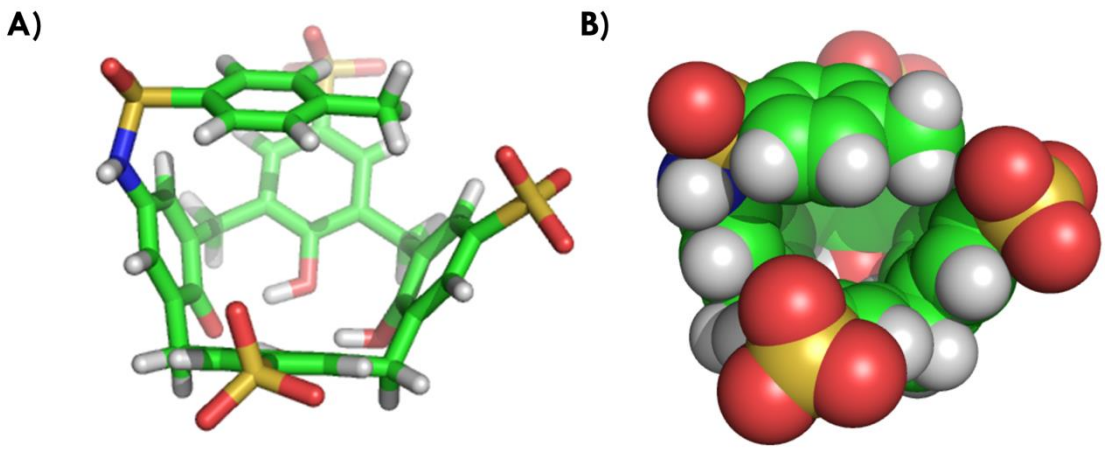


Figure 2.8 Collapsed molecular model of host 9. A) Shows a stick representation B) Shows a sphere representation of the collapsed host. Both were modeled using PyMol,

2.5 Conclusions and Future Work

The calixarenes that we report here have an unusual ability to bind to post-translational modifications (as opposed to binding within the ordered binding pockets of protein targets). This offers an effective route to the disruption of protein-protein interactions that are triggered by post-translational modifications. The structure-function relations identified here guide our understanding of the molecular recognition determinants for these synthetic receptors. Importantly, these compounds are created via a modular synthesis that suggests that a great number of them can be made and tested in order to create more potent and more selective analogs, as well as to identify agents with selectivities for other post-translational modification states.[105]

The studies reported in this chapter using the fluorescence-based dye displacement assay allow us to quickly and easily measure the specificity and selectivity of peptides with various methylation states to the synthesized hosts, and could easily be expanded to modifications of different types. The assay is operated in 96-well plates and consumes small amounts of sample and protein allowing for high-throughput readings. We are working to combine the high-throughput nature of this analytical method with the high-throughput synthesis of many analogues in order to identify agents with high affinities and selectivities for a variety of different post-translational modifications.

Although the compounds reported here are not drug-like, they are novel candidates for use as biochemical reagents for the *in vitro* characterization and disruption

of methylation signaling proteins. Given the extreme rarity of chemical agents of any kind that can disrupt the complexes of methyllysine-reader proteins, [120-124] we expect these agents to find diverse application in bioorganic and biochemical studies of methylation pathways such as their use as optical sensors for tracking enzymatic modifications. Some of our efforts in this direction will be reported in the next Chapter.

Chapter 3: Use of supramolecular hosts as a continuous tracking method for monitoring enzymatic activity

3.1 Introduction

The previous chapter described the ability to measure the dissociation constants of particular histone peptide tails with various calixarenes. We wanted to understand the strength of binding between parts of histone peptide tails and modified calixarenes and use these results as a guide for our efforts to disrupt the interaction between CBX7 and the H3K27me3. These studies made it clear that we could detect the difference between unmethylated and methylated peptides. We were interested in whether these fluorescence-based measurements could be used for real-time monitoring of enzymatic reactions that install or remove methyl groups. One precedent for such an assay exists in the literature. Using **PSC4** and **LCG**, Nau[102] was able to monitor the progress of Dim-5 methyltransferase which trimethylates H3K9 peptide. This reaction produces H3K9me3 which binds more strongly to **PSC4** and displaces more **LCG**. This turn-on sensor was optimized to conditions of pH 10.8 in glycine buffer that the enzyme preferred, and allowed for the real-time monitoring of the methylation by an increase of fluorescence. This kind of approach, which we'll call **first generation approach,** shown in Figure 3.1, could potentially be used to report on the activities of a wide variety of different enzymes that install or remove different post-translational modifications, including histone deacetylases (HDACs), histone methyltransferases (HMTs) and histone demethylases (HDMs).

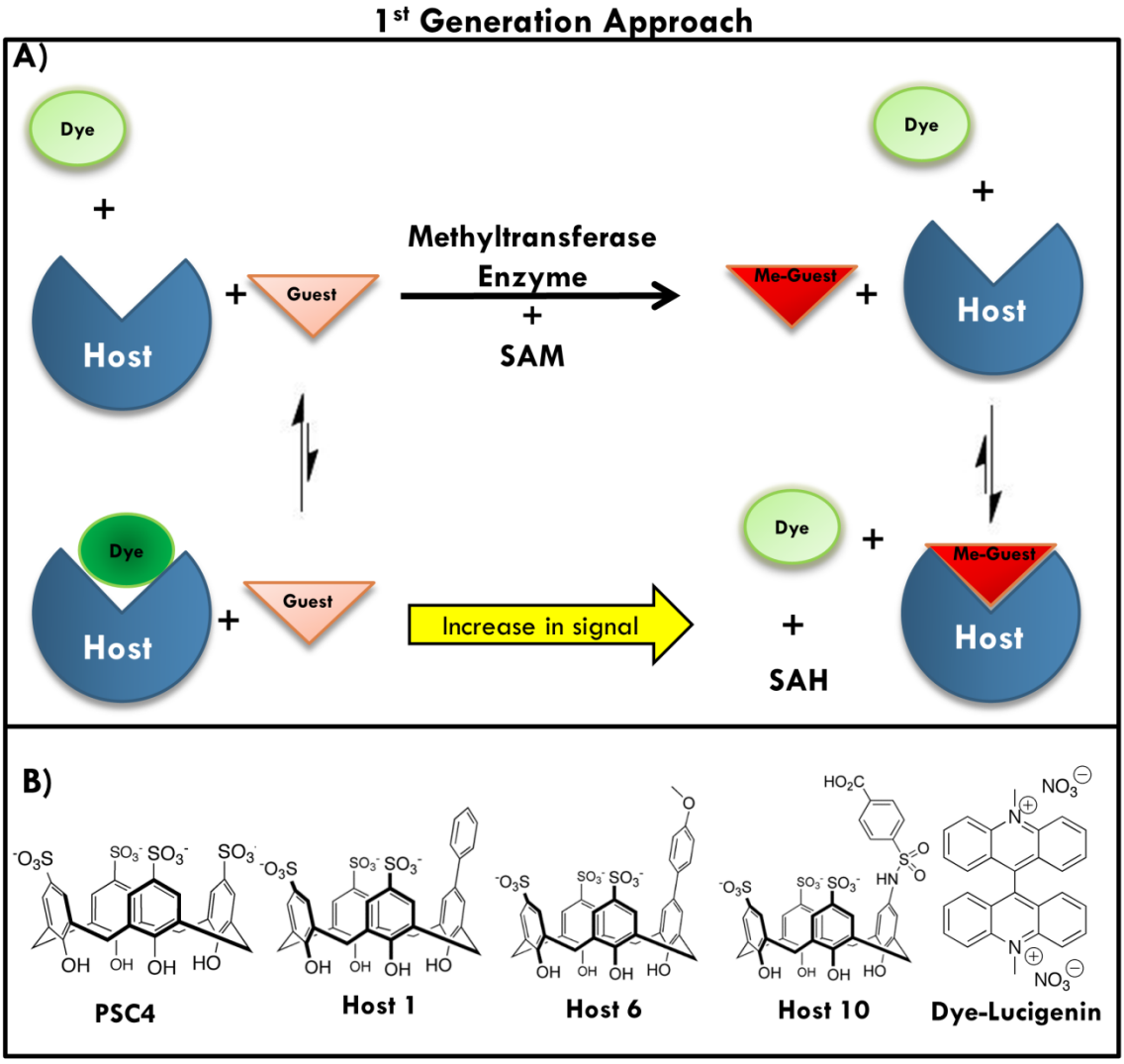


Figure 3.1 First generation approach. A) Dye displacement assay showing the typical direct titration equilibrium followed by the addition of a methyltransferase enzyme with S-adenosylmethionine (SAM) co-substrate that can methylate the substrate guest and yield a methylated guest (Me-Guest) and S-adenosylhomocysteine (SAH) by-product. The methylated guest can now displace more of the dye and allow for real-time measurement of methylation. Both unmethylated and methylated guests bind to the host cavity as the reaction progresses but with a stronger affinity when the guest is methylated, therefore increasing the signal. B) Hosts used for 1st generation approach are shown (**PSC4**, host **1**, **6** and **10**) and LCG dye.

Chapter 1 introduced the importance of having methods that allow us to monitor post-translational modifications. In general these methods were divided into two categories, **continuous** (mass spectrometry[125], coupled assays[71] etc.) and **discontinuous**

(ELISA[51], DELFIA[52] etc.). The ability to monitor modifications in a continuous, sensitive and convenient way is important to further our understanding about these writer and eraser proteins. The approaches that monitor enzymatic activity in real-time are very promising for the study of future enzymes because they meet many of the criteria necessary for a successful assay. They are "one-pot" assays that allow the direct measurement of the product of the enzymatic reaction, therefore avoiding complicated coupled assays while maintaining the continuous ability to observe progress of reaction. It is also a fluorescence-based assay which promises ease of use with common equipment such as plate readers, good sensitivity to low concentrations of reactants, and overall safe operation due to lack of radioisotopes.

This chapter reports on my efforts to use **LCG**-based dye displacement methods (**1st generation approach**, Figure 3.1), analogous to those reported in Chapter 2, as well as the development and usage of a novel family of supramolecular sensors that are composed of supramolecular hosts covalently linked to fluorescent reporter dyes (**2nd generation approach,** Figure 3.2). In addition, my efforts at developing assays for G9a, LSD1 and HDAC1 are reported. The data presented will mainly focus on the difficult, but eventually successful development of a readout of G9a histone methyltransferase activity. I will also attempt to provide the reader with a good understanding of the barriers encountered for the development of these kinds of assays in general, and for the other above-mentioned enzymes in particular.

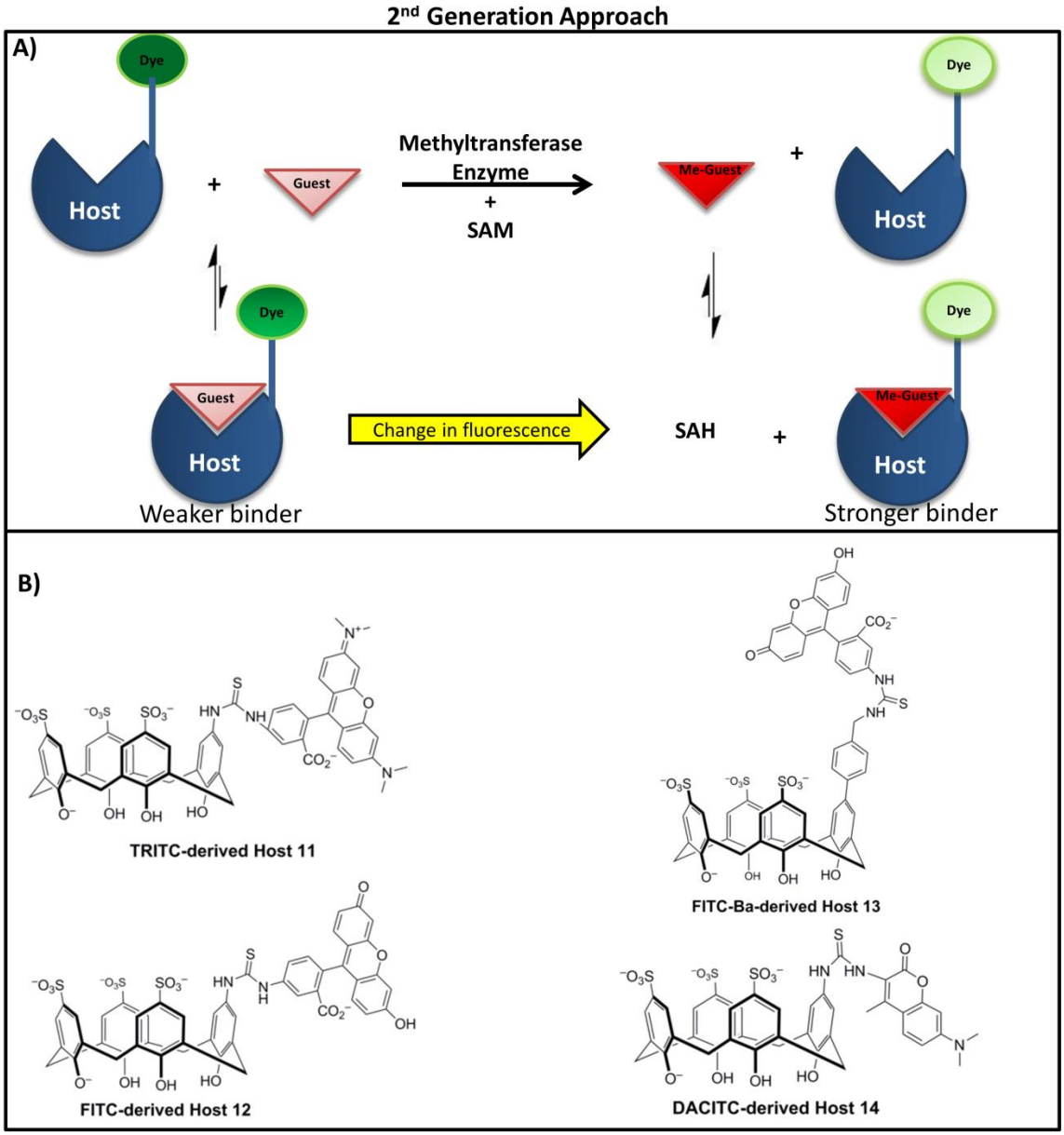


Figure 3.2 Second generation approach. A) Fluorescence-based assay showing a dye-derived host with the addition of a methyltransferase enzyme and S-adenosylmethionine (SAM) co-substrate that can methylate the substrate guest and produce a methylated guest (Me-Guest) and S-adenosylhomocysteine (SAH) by- product. The methylated guest can now bind the dye-derived host with a stronger affinity and change its inherent fluorescence. Both the unmethylated and methylated guests are able to bind in the host cavity as the reaction progresses. B) Dye-derived hosts used for 2nd generation approach are shown (hosts **11**, **12**, **13** and **14**).

Histone demethylases such as LSD1[126-128] were briefly introduced in Chapter 1 as an example in the section on coupled assays (Scheme 1.6). LSD1 specifically demethylates histone H3 lysine 4 (H3K4me2) through an oxidative reaction that generates formaldehyde and the unmethylated product. Using either the first or second generation approach, we would expect the LSD1 substrate peptide (H3K4me2) to bind better to the hosts than the product unmethylated (H3K4) peptide.

Histone deacetylases (HDACs) remove acetyl groups from lysine side chains, with different subtypes having selectivity for different acetylated substrates. Unlike LSD1, where the methylated substrate binds better to the host than does the unmethylated product, histone deacetylases such as HDAC1 generates a charged product that binds better to the host cavity than does the neutral, acetylated substrate. HDAC1 removes an acetyl group from a lysine residue such as H4K5ac and generates an unacetylated product (H4K5).

Methyltransferases are represented in this Chapter by G9a or Euchromatic Histone Methyltransferase 2 (EHMT2). It is a member of the Suvar (3-9) family of SET domain histone methyltransferases (HMTases). G9a can methylate both lysine 9 and lysine 27 of histone H3. According to other studies, it has also been shown to methylate non-histone targets.[129] Part of the N-terminal catalytic domain has been found to automethylate one of its own lysine residues, located in a core motif that strongly resembles the H3K9 methylation site.[130] G9a plays a vital role in the euchromatic histone H3K9 methylation by allowing association with effector proteins. *In vitro*, G9a has been observed to both di- and trimethylate H3K9[130] creating a binding site for chromodomain (CD)-containing proteins of the heterochromatin protein 1 (HP1). Inactivation of G9a has been seen to cause severe growth factor retardation and early lethality.[131] Methylation of a histone peptide tail such as H3K9 using G9a HMT and S-adenosylmethionine (SAM) co-substrate as shown in Scheme 1.4 produces a methylated peptide that is able to bind more strongly to one of our hosts than the unmethylated substrate.

For any given enzyme, the **first aim** would be to optimize a supramolecular sensing system that allows a good, fluorescence-based discrimination between substrate and product peptides; these experiments will be referred to as **substrate-product discrimination experiments.** Both sensor and enzymes are sensitive to buffer conditions

in different ways, and so for any given optimization task, buffer conditions must be identified that work for sensing and that are also likely to support the activity of the enzyme. Following the achievement of the first aim, my **second aim** would then be to monitor the progress of an enzymatic reaction in real-time; these will be referred to as **kinetic runs**.

3.2 Experimental

3.2.1 Materials and methods

All chemicals were purchased from Chemimpex, VWR Canada or Sigma-Aldrich unless otherwise noted. Lucigenin stock solutions were prepared by sonicating the solid in distilled water and freezing at high concentrations. Dilutions for experiments were made as needed. Emission spectra and kinetic reads were done using SpectraMax® M5 / M5e Microplate Reader. The spectra were obtained at room temperature and the kinetic reads were obtained at both room temperature and 30°C where indicated. Wells were prepared by addition of stock solutions through single or multichannel pipettes to Nunc 96-well optical plates (black sides, clear bottom) and were diluted to their final concentrations at 200 µL total volume. Peptides were characterized using mass spectrometry on a Thermo Electron Corporation, Finnigan LCQ Classic ion trap mass spectrometer with electrospray ionization (ESI). Peptide solution concentrations were quantified by UV-Vis using 700 µL quartz cuvette with the cuvette reader accessory for the above plate reader. The extinction coefficient used was 1490 $M^{-1}cm^{-1}$ with a path length of 1 cm and absorption at 280 nm. All enzymes used were supplied by Active Motif.

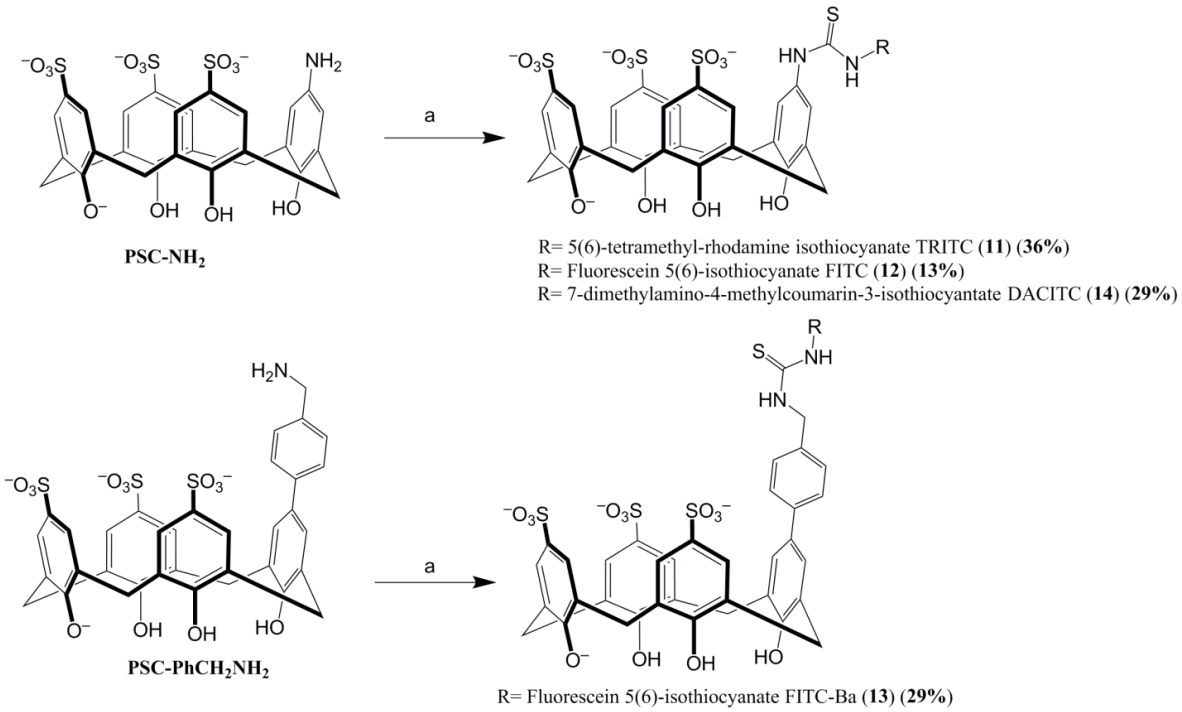
3.2.2 Peptide synthesis

All sequences had a tyrosine introduced at the C-terminus to facilitate UV detection during HPLC purification and subsequent quantification of stock solution concentrations. Histone 3 **H3K9me2 (1-16)** NH_2-ARTKQTARK**me2**STGGKAPY-$C(O)NH_2$ was purchased from Cedarlane. For peptides synthesized in house, all Fmoc-protected amino acids were purchased from ChemImpex except for Fmoc-Lys(Me3)-OH chloride which was purchased from GL Biochem. All other reagents used for peptide synthesis were purchased from ChemImpex or Sigma-Aldrich. Histone 3 **H3K9 (1-12)**

and **H3K4 (1-12)** NH_2-ARTKQTAR**K**STGY-C(O)NH_2 , **H3K9me3 (1-12)** NH_2-ARTKQTAR**Kme3**STGY-C(O)NH_2 , **H3K4me3 (1-12)** NH_2-ART**Kme3**QTARKSTGY-C(O)NH_2 were synthesized by Kevin Daze, **H3K4 (1-8)** NH_2-ART**K**QTARY-C(O)NH_2 , **H3K4ac (1-8)** NH_2- ART**Kac**QTARY -C(O)NH_2 were synthesized by Janessa Li and **H3K9 (1-16)** NH_2-ARTKQTAR**K**STGGKAPY-C(O)NH_2 , **H3K4 (1-21)** NH_2-ART**K**QTARKSTGGKAPRKQLAY-C(O)NH_2, **H3K4me2 (1-21)** NH_2-ARTK**me2**QTARKSTGGKAPRKQLAY-C(O)NH_2, **H4K5 (3-5)** Ac-RG**K**Y-C(O)NH_2 **H4K5ac (3-5)** Ac-RG**Kac**Y-C(O)NH_2 were synthesized by me using the standard Fmoc solid-phase peptide synthesis protocol[110] as implemented on a CEM Liberty 1 microwave-based peptide synthesizer. A more detailed procedure on peptide synthesis is shown in **Section 2.2.2**.

3.2.3 Calixarene synthesis

All calixarenes discussed in this chapter were synthesized by Kevin Daze. First generation calixarene synthesis was described in Chapter 2 and second generation calixarene synthesis is described in Scheme 3.1. All calixarenes were purified by HPLC on a preparative Luna C-18 column (Phenomenex, 5 μm, 21.2x250 mm) using a gradient starting from 90:10 H_2O:MeCN (0.1% TFA) and running to 10:90 H_2O:MeCN (0.1% TFA) at a flow rate of 10 mL/min over 30 minutes. Analytical HPLC traces for compounds **11, 12, 13** and **14** are found in appendix A.



Scheme 3.1 Synthesis of Sulfonato-calix [4]arenes 11-14. a) Treatment with R-N=C=S in 1:1 DMF:pyridine, overnight at room temperature.

3.2.3.1 Tetramethylrhodamine Isothiocyanate (TRITC) - derived host 11

PSC-NH_2 Calixarene and isothiocyanate dye (1.1 eq) are dissolved in a 2 mL mixture of pyridine and DMF (1:1) and stirred overnight at room temperature in the dark. The reaction mixture is poured into 15 mL of H_2O and extracted with 2x20 mL DCM, 1x15 mL EtOAc and the aqueous layer is lyophilized to dryness. The calixarene is purified by Reverse Phase –High Performance Liquid Chromatography (RP-HPLC). Yield: 36% ESI-MS: calc for $C_{53}H_{47}N_4O_{16}S_4^+$: 1123.2 found: 1124 $[M+H]^+$

3.2.3.2 5(6)-Fluorescein Isothiocyanate (FITC) - derived host 12

PSC-NH_2 Calixarene and isothiocyanate dye (1.1 eq) are dissolved in a 2 mL mixture of pyridine and DMF (1:1) and stirred overnight at room temperature in the dark. The reaction mixture is poured into 15 mL of H_2O and extracted with 2x20 mL DCM, 1x15 mL EtOAc and the aqueous layer is lyophilized to dryness. The calixarene is purified by Reverse Phase –High Performance Liquid Chromatography (RP-HPLC). Yield: 13% HR-ESI-MS: calc for $C_{49}H_{36}N_2O_{18}S_4$: 1067.0768 found: 1067.0769 $[M-H]^-$. Purity determined to be 97% by HPLC (see appendix A).

3.2.3.3 5(6)-Fluorescein Isothiocyanate Benzylamine (FITC-Ba) - derived host 13

PSC-$PhCH_2NH_2$ Calixarene and isothiocyanate dye (1.1 eq) are dissolved in a 2 mL mixture of pyridine and DMF (1:1) and stirred overnight at room temperature in the dark. The reaction mixture is poured into 15 mL of H_2O and extracted with 2x20 mL DCM, 1x15 mL EtOAc and the aqueous layer is lyophilized to dryness. The calixarene is purified by Reverse Phase –High Performance Liquid Chromatography (RP-HPLC). Yield: 29% ESI-MS: calc for $C_{56}H_{40}N_2O_{18}S_4^{2-}$: 1158.1 found: 1157.5 $[M-H]^-$. Purity determined to be 98% by HPLC (see appendix A)

3.2.3.4 Dimethylamino - 4 - methylcoumarin - 3 -isothiocyanate *(*DACITC)- derived host 14

PSC-NH_2 calixarene and dansyl chloride are dissolved in 1 M sodium phosphate buffer (pH 8) and stirred overnight at room temperature in the dark. The solution is extracted with 2x20 mL DCM and 1x15 mL EtOAc and the aqueous layer is lyophilized to dryness and purified by Reverse Phase –High Performance Liquid Chromatography (RP-HPLC). Yield: 29% ESI-MS: calc for $C_{41}H_{37}N_3O_{15}S_4$: 939.1 found: 938.2 $[M-H]^-$. Purity determined to be 98% by HPLC (see appendix A).

3.2.4 Sample Protocol

All buffers used in this Chapter are shown in the table and will be referred to by their one letter code throughout this Chapter.

Table 3.1 Buffers A-H used in this Chapter.

Buffer	Components
A	20 mM Tris NO_3^- at pH 8.5, 5 mM $Mg(NO_3)_2$ and 4 mM DTT
B	0.01 M phosphate buffer at pH 7.5
C	20 mM Tris NO_3^- at pH 8.5
D	20 mM Tris/Cl^- pH 9.12, 5 mM $MgCl_2$, and 4 mM DTT
E	0.2 M of HEPES
F	50 mM phosphate, 25% glycerol
G	50 mM Tris/Cl^- pH 8.5, 5 mM $MgCl_2$, 50 mM KCl
H	50 mM Tris/Cl^- at pH 8.0, 137 mM NaCl, 2.7 mM KCl and 1 mM $MgCl_2$

3.2.4.1 First generation approach, substrate-product discrimination setup

For setting up a plate, four different controls were done as shown in the table below.

Table 3.2 96-well black optical bottom plate setup for substrate-product discrimination experiment using 1st generation calixarenes. Plate is excited at 369 nm and emission is observed between 445-615 nm or at the λ_{max} using the SpectraMax® M5 / M5e Microplate Reader.

	Volume (µL)	Well #1	Well #2	Well #3	Well #4	Well #5	Well #6	Well #7
10X Buffer (A,B, or C)	20							
5 µM LCG	20							
12.5 µM Host	20							
500 µM H3K9 (1-12)	20							
500 µM H3K9me3 (1-12)	20							
1.5 mM SAM	20							
dH_2O	**Make up all wells to 200 µL**							

Legend of the names of each well:
Well #1: Buffer blank
Well #2: LCG only
Well #3: LCG and host
Well #4: Positive control (product peptide)
Well #5: Negative control (substrate peptide)
Well #6: SAM control
Well #7: Substrate peptide and SAM control

3.2.4.2 Second generation approach, kinetic run setup

For setting up a plate, four different controls were done as shown in the table below.

Table 3.3 Exemplary 96-well black optical bottom plate setup for substrate-product discrimination experiments using 2nd generation calixarenes and kinetic run of enzymatic activity monitoring in real-time. Plate is excited at each dye's λ_{exc} (found in Table 3.4) and emission is observed at an emission range for spectrum reads and at the maximum emission λ_{max} for kinetic run using the SpectraMax® M5 / M5e Microplate Reader.

	Volume (µL)	Well #1	Well #2	Well #3	Well #4	Well #5	Well #6	Well #7
10X Buffer (D,E or F)	20	■	■	■	■	■	■	■
5 µM dye-derived host 11	20		■	■	■	■	■	■
125 µM H3K9 (1-16)	20			■			■	■
125 µM H3K9me2 (1-16)	20				■			
1.25 mM SAM	20					■	■	■
2 µg/µL G9a	0.125							■
dH_2O	**Make up all wells to 200 µL**							

Legend of the names of each well:
Well #1: Buffer blank
Well #2: Dye-derived host
Well #3: Negative control –substrate peptide
Well #4: Positive control – product peptide
Well #5: SAM control
Well #6: SAM and substrate control
Well #7: Reaction

3.3 Results and Discussion

The first enzyme studied was G9a histone lysine methyltransferase. It is a high value target for new assays because it plays a major role in epigenetic regulation of gene transcription. Preliminary studies for other enzymes such as HDAC1 histone deacetylase and LSD1 lysine demethylases were initiated and some barriers to success on those enzyme assays were discovered and will be discussed. The majority of the results will be focused on the development of an assay for G9a methyltransferase.

Earlier studies done in Chapter 2 showed the ability of calixarenes to distinguish between unmethylated and methylated peptides.[132] Here we explored two possible ways of measuring methylation by G9a methyltransferase in real-time. The first is using a dye-displacement assay, derived from those shown in Chapter 2, which relies on displacing a dye, lucigenin, from a calixarene cavity upon guest binding. The use of lucigenin as a dye-displacement indicator for use with sulfonated calixarene hosts is an idea that was adapted from Nau.[102, 133]

3.3.1 Optimizations using first generation approach

3.3.1.1 First aim: Dye displacement sensing of unmethylated H3K9 (1-12) and trimethylated H3K9me3 (1-12) peptides with first generation calixarenes

Initial substrate-product discrimination experiments using **H3K9 (1-12)** and **H3K9me3 (1-12)** were done using hosts **PSC4**, and host **1**,**6** and **10** with Lucigenin in order to discriminate between unmethylated and trimethylated peptides. The dye displacement sensing protocol was adapted from Chapter 2 and optimized to work in 96-well plates with a total volume of 200 µL. Four types of samples were prepared for each dye-displacement assay: the free dye **LCG** , the dye with host (**PSC4, 1, 6 or 10** shown in Figure 3.1), the dye with host and guest molecule (either the **H3K9** or **H3K9me3**) as shown in Table 3.2. The **H3K9(me3) (1-12)** shorter peptide sequences were used because they were available and this experiment was for proof of principle. If a particular host is able to differentiate between, the 12-mer undecorated and trimethylated states, it should be able to differentiate between the G9a methyltransferase susbtrate and product.[2]

[2] The first generation calixarenes were synthesized before the second generation calixarenes and at the time, I only had access to H3K9 (1-12) and H3K9me3 (1-12) synthesized by Kevin Daze. For tracking G9a methylation, H3K9 (1-16) and H3K9me2 (1-16) were used instead since shorter peptides do not act as substrates for G9a writer protein.

The experiment was run in three different buffer conditions in order to determine the conditions most able to discriminate between substrate and product. Buffer components for this assay had to be modified due to Lucigenin's inherent sensitivity to quenching by anions (and especially halides) in the solvent environment. Therefore all chloride containing salts were substituted with oxyanions like nitrate and phosphate that cause less **LCG** quenching. Buffer **A** was adapted from Pradhan,[134] buffer **B** was taken from Chapter 2, and buffer **C** was adapted from Weiss[135] with final concentrations shown in Table 3.1 In each well, the final dye, host and peptide concentrations used were 0.5 µM, 1.25 µM and 50 µM respectively. The free dye well shows a strong emission of free **LCG** when excited at 369 nm. Upon addition of host, the observed fluorescence is decreased compared to the free dye well. Restoration of the fluorescence is observed upon addition of a guest, either the unmethylated or trimethylated peptide in this case. A good **sensor** is one that shows good substrate-product discrimination for the unmethylated and trimethylated peptides. Figure 3.3 A, B summarizes the percent change between emission intensities at λ_{max} of 485 nm for samples containing the unmethylated (blue bars) and trimethylated (red bars) peptides with two **PSC4** and host **10**. **PSC4** is shown in Figure 3.3 A and host **10** in Figure 3.3 B in buffers **A, B** and **C.**

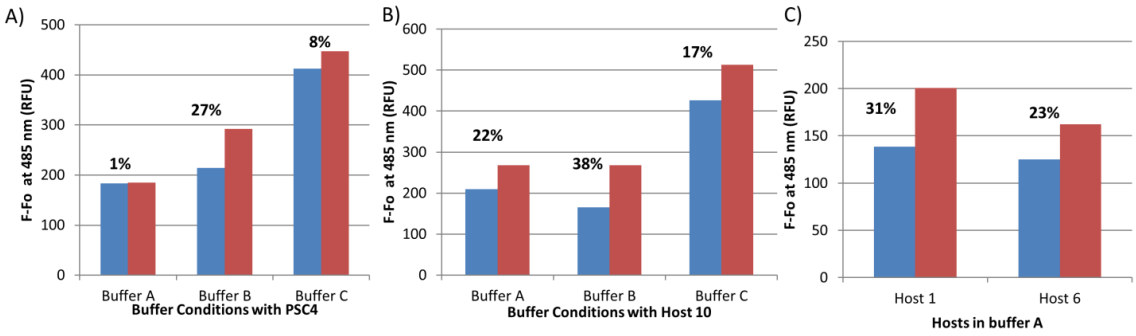


Figure 3.3 Dye displacement substrate-product discrimination between samples containing 50 µM H3K9 (1-12) (blue bars) or 50 µM H3K9me3 (1-12) (red bars) in buffers A, B , and C. A) F-F_0 values for guests with PSC4 at 1.25 µM and B) host **10** at 1.25 µM. F-F_0 values calculated by measuring the observed fluorescence intensity at λ_{max} subtracted from the blank (host and LCG well). Percent change calculated by the following equation: % = $(((F\text{-}F_0)_{H3K9me3} - (F\text{-}F_0)_{H3K9}) / (F\text{-}F_0)_{H3K9me3}) \times 100$.

As seen in Figure 3.3 A and B, buffer **B** shows the best conditions with the highest percent change for sensing between both peptides for both **PSC4** and host **10**, which is expected since buffer **B** is the same buffer used in Chapter 2 assays and has proven to be an efficient buffer for this sort of dye displacement assay. Buffer **A** was adapted from a chloride-containing buffer which has been used in literature for *in vitro* G9a methyltransferase activity[134] and therefore was selected as the buffer of choice. Substitution of chloride for nitrate in the buffer solution still allowed for a good distinction between the peptides when using host **10** and was chosen as the buffer for further assays using 1st generation calixarenes and G9a methyltransferase. Host **1** and **6** were also tested using buffer **A** showing a percent change of 31 % and 23% respectively (Figure 3.3 C). Due to the availability of host **1**, and the high percent discrimination of 31 % in buffer **A**, it was chosen as the host for further experiments with the G9a methyltransferase in the 1st generation approach.

3.3.1.2 Second aim: Monitoring methylation in real-time using first generation calixarenes

Having accomplished the **first aim** with the 1st generation calixarenes and determined buffer conditions that allow host **1** to discriminate between H3K9 (1-12) substrate and H3K9me3 (1-12) product, the **second aim** was to see if the optimized conditions would allow the methylation of a particular substrate with G9a methyltransferase. In order to monitor the enzymatic activity of G9a, the reading mode used was that allowed by the instrument's "**kinetic run**" settings, which allows emission intensity at a single wavelength to be observed continually over time. This mode allows multiple wells to be observed simultaneously, so in order to correlate emission intensities with the conversion of substrate to product, multiple control wells were also observed at the same time as the well containing the enzymatic reaction. A black 96-well optical bottom plate was setup as seen in Table 3.2 with an additional well containing the G9a enzyme. The reaction well contained buffer **A**, host **1** at 1.25 µM, **LCG** at 0.5 µM, 50 µM of **H3K9 (1-16)** substrate, 150 µM of S-adenosylmethionine (SAM), and 2 µg of G9a enzyme. The plate was excited at 369 nm, and emission was measured at 485 nm over 90 minutes at 30°C. As expected, the addition of **LCG** increased the fluorescence

signal (green line in Figure 3.4), while the addition of host **1** decreased it (dark red line). Addition of the susbstrate H3K9 (1-16) (light blue line) to the next well allowed the displacement of some of the **LCG** and restoration of the fluorescence. Since SAM is a co-substrate for the G9a methyltransferase, it was important to see whether the addition of SAM to the substrate would cause a shift in fluorescence.

We found that adding SAM co-substrate without peptide increased the fluorescence by displacing some of the **LCG** by an average of 41% (black line). However, the addition of SAM co-substrate to the substrate peptide only decreased the fluorescence by an average of 3% (dark blue line). Since the addition of SAM to the **H3K9 (1-16)** peptide only decreased the fluorescence slightly, this result would be acceptable to use for monitoring G9a methylation. Finally, 2 µg of the G9a methyltransferase was added to the well to monitor methylation. As seen in Figure 3.4, the well containing G9a methytransferase (bright red line) slightly decreased in fluorescence compared to the well without the enzyme (substrate and SAM only dark blue line). An increased fluorescence signal, closer to the green line, was expected but was not observed for the G9a methyltransferase reaction well. At this stage of the experiment, I did not have access to a positive control such as H3K9me2 (1-16) but regardless, no enzymatic activity was observed.

The measurement of an enzymatic reaction is much more complex than simply discriminating substrates and products, because the sensing system must function in the presence of additional components, such as co-substrates and the enzyme itself. It also must function in the presence of a buffer in which the enzyme actually works. Substitution of buffer components containing chloride with nitrate may have been the reason for the lack of methylation in our system. In support of this, a study done by Pradhan *et al.*[134] showed that G9a methyltransferase is highly dependent on reaction conditions. For example, at pH 7, G9a displays poor activity and shows increasing activity as the pH increases up to pH 9. Another set of experiments done by same group showed that an insignificant increase in methylation is observed when the temperature is shifted between 25-42°C. While surveying these and other literature references to determine what buffer concentration to use, it became clear that there are many different buffers possible to use and temperatures where the enzyme would be active. However,

finding the correct buffer conditions that can a) act as a good buffer for our 1st generation approach calixarenes and accommodate **LCG**'s inherent sensitivity and b) allow for enzyme activity is a very difficult task that we could not achieve.

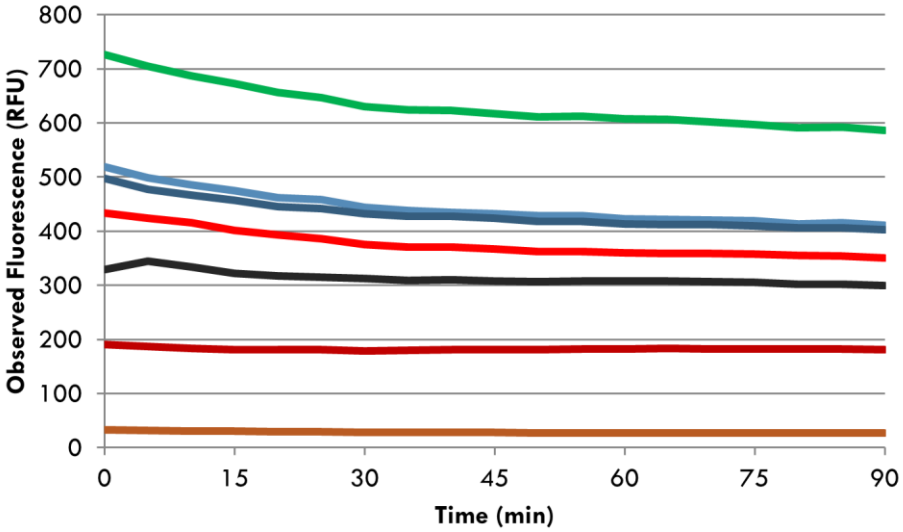


Figure 3.4 G9a methylation using 1st generation approach. 96-well black optical plate run in a SpectraMax® M5 / M5e Microplate Reader at 30°C over 90 minutes. Buffer **A** well (brown line), host **1** used at 1.25 µM (dark red line), with substrate **H3K9 (1-16)** at 50 µM (light blue line) with S- adenosylmethionine (**SAM**) at 150 µM (dark blue line). All wells except buffer well contained 0.5 µM of **LCG**. 1 µL of G9a was added to the final well indicated in bright red (total concentration 2 µg/µL). Plate was excited at 369 nm and emission at 485 nm was detected at 30 second intervals. An additional well containing buffer, host **1**, **LCG** and 150 µM of SAM is shown in black.

The inability to use buffers containing $MgCl_2$ or other components with chloride due to the **LCG** sensitivity, forced us to move onto a new system that would allow for us to use fluorescent dyes compatible with a wider range of buffer conditions. The 1st generation, LCG-binding calixarenes were no longer used and 2nd generation calixarenes that would hopefully eliminate buffer problems were the focus of all further efforts.

3.3.2 Optimizations using second generation approach.

3.3.2.1 First aim: Discrimination of unmethylated H3K9 (1-16) and dimethylated H3K9me2 (1-16) peptides with second generation calixarenes

The 2nd generation approach involves the calixarenes that are directly affixed with fluorescent dyes shown in Figure 3.2. These dyes are all known to be compatible with a wide variety of buffer conditions. The first task — substrate-product discrimination tests,

were performed using dye-labeled hosts (**11**, **12, 13** or **14** shown in Figure 3.2) with guest molecule (**H3K9 (1-16)** or **H3K9me2 (1-16)**). The buffer used for these experiments, buffer **D**, shown in Table 3.1 was the same as buffer **A** used in the first generation approach, but with all of the chloride-containing components that had originally been reported in the literature. The free dye-derived host well #2 in Table 3.3 was used at either 0.5 μM (for **11, 14**) or 0.1 μM (for **12, 13**). Depending on the dye-derived host being used, different excitation and emission wavelengths as well as emission cut offs were used as shown in Table 3.4.

Unlike the fluorescence trend seen in the previous section, the mechanism for the fluorescence responses of the second-generation hosts are not as straightforward as those of the dye-displacement mechanisms in the first-generation systems, in which the addition of substrate or product guest molecules always produced an increase in fluorescence output due to displacement of the **LCG** into free solution (where it is more emissive). The dye-derived hosts display different trends depending on the identity of the dye attached. Upon addition of a guest to the dye-derived hosts, the environment of the fluorophore is changed affecting the emission. We presume that the overall anionic nature of these dye-derived hosts drives the formation of complexes with the peptides – all of which are positively charged. FITC-derived host **12**, FITC-benzylamine-derived host **13** as well as DACITC-derived host **14** show a decrease in intensity upon binding of the peptide **H3K9me2(1-16)** as shown in Figure 3.5. Host **12** and **13** contain a fluorescein moiety that is known to be quenched by some electron-rich amino acids such as tryptophan and tyrosine and thought to occur through a static and dynamic quenching process.[136, 137] This quenching behaviour is explainable with all the peptides used here as they all contain a tyrosine at the C-terminus.

Table 3.4 Dye-derived hosts excitation and emission wavelengths and cut offs used for assays using the SpectraMax® M5 / M5e Microplate Reader.

	Excitation wavelength λ_{exc} (nm)	Emission wavelength range λ_{em} (nm)	Emission maximum λ_{max} (nm)	Emission cut off λ_{CutOff} (nm)
Host 11	532	560-650	585	550
Host 12	450	480-600	515	470
Host 13	450	480-600	515	470
Host 14	390	440-510	470	430

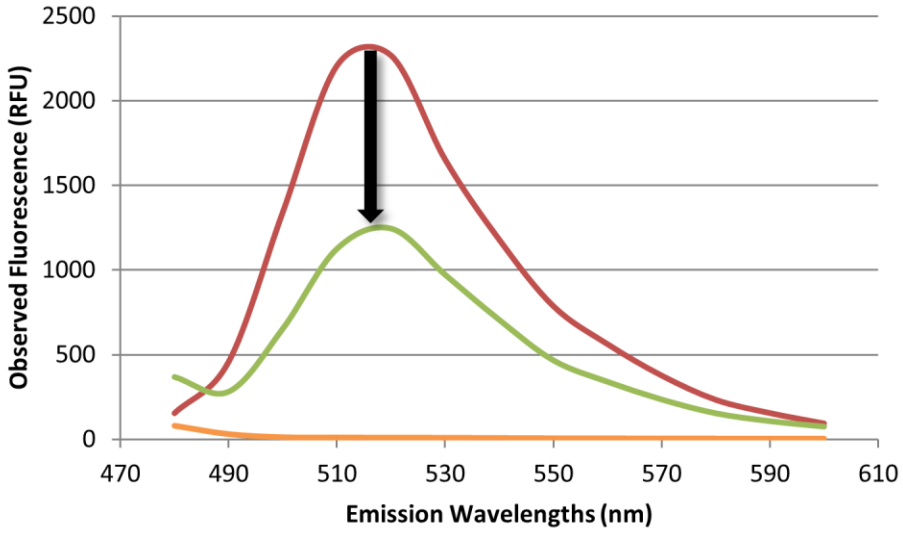


Figure 3.5 Representative spectrum showing a decrease in fluorescence from FITC-derived host 12 at 1 µM final concentration (red line) and decreased fluorescence upon H3K9me2 (1-16) guest addition (green line) at 27.5 µM final concentration. Orange line represents blank buffer **D** well. Samples were excited at 450 nm and emission was observed at 480-600 nm with an emission cut off of 470 nm. Similar trends are observed for hosts **13** and **14** (observed using wavelengths reported in Table 3.4).

TRITC-derived host **11** however shows similar trends as the 1st generation sensor in that upon addition of a guest, the fluorescence increases as shown in Figure 3.6. Emission increases could be achieved upon addition of peptides due to the decrease of self-quenching by host **11**. Host **11** contains a tetramethylrhodamine moiety which is known to homodimerize and self-quench in solution by a non-radiative process,[138-140] and addition of a competing guest molecule may disrupt the dimerization and allow an

increase in fluorescence. These suggestions warrant further experimentation in order to determine photochemical mechanisms of quenching and emission enhancement, in the hopes that such knowledge would allow the further improvement of the performance of each of these different photochemical sensors.

Initial optimization experiments were done in order to assess the sensitivity (substrate-product discrimination) of each of the dye-derived compounds. Having removed the troublesome Lucigenin component from the sensor scheme, I was able to now use the buffer conditions that are optimal for G9a methyltransferase activity as previously shown by Pradhan *et al.*.[134]

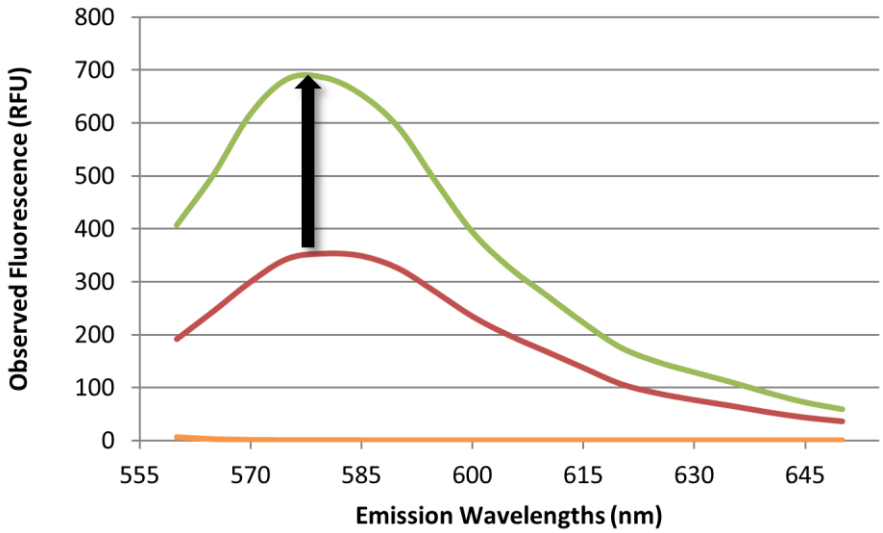


Figure 3.6 Representative spectrum showing an increase in fluorescence from TRITC-derived host 11 at 0.5 µM final concentration (red line) upon H3K9me2 (1-16) guest addition (green line) at 12.5 µM final concentration. Orange line represents blank buffer **D** well. Plate was excited at 532 nm and emission observed at 560-650 nm with an emission cut off of 550 nm.

Figure 3.7 shows the percent change for the four different dye-sensors in the chloride containing buffer **D** with both unmethylated **H3K9 (1-16)** and dimethylated **H3K9me2 (1-16)** (the longer, dimethylated product peptide had become available in the intervening time since the abandonment of my first-generation approach to performing enzyme assays). Figure 3.7 A represents data similar to those reported in Figure 3.3 where the biggest percent change in fluorescence between the unmethylated substrate

(blue bars) and dimethylated product (red bars) is with host **11** at 70%. Dye-derived hosts **12** and **13** show poor sensitivity to product (i.e. low percent change, 13% and 0.4% respectively) while host **14** shows a percent change of 32%. I had become increasingly aware of the possibility that the methyl-donor enzyme co-substrate *S*-adenosylmethionine (SAM) could itself bind to calixarenes and induce a fluorescence response. I determined the size of the effect of SAM by preparing wells of substrate peptide **H3K9 (1-16)** with and without SAM, and calculating the percent change in intensity induced by the addition of SAM. As seen in Figure 3.7 B, wells containing substrate peptide and SAM (red bars) caused a relatively large percent change for FITC-derived host **12** and FITC-Ba-derived host **13** at 22% and 14%, respectively while causing a smaller change of 8% of emission intensity for the TRITC-derived host **11.** The experiment using SAM was not performed with DACITC-derived host **14** and therefore was not added to the graph in Figure 3.7 B.

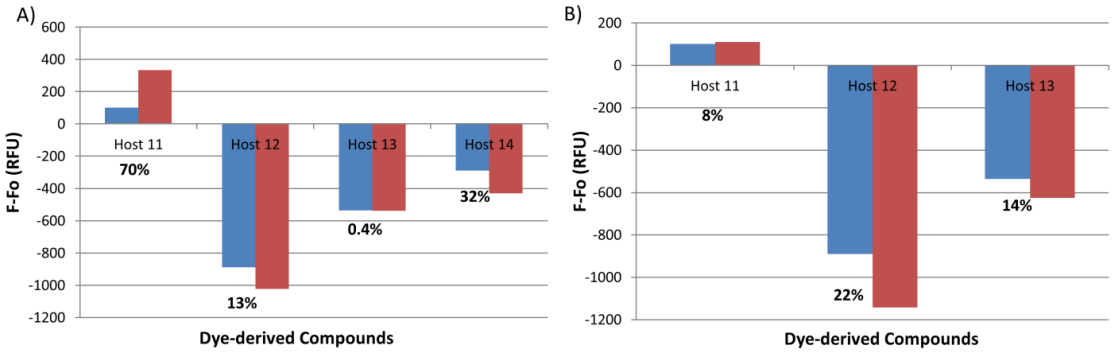


Figure 3.7 Substrate-product discrimination for four different dye-derived calixarenes. A) F-Fo values for unmethylated **H3K9 (1-16)** substrate (blue bars) and dimethylated **H3K9me2 (1-16)** product (red bars) at λ_{max} of each host. B) F-Fo values for substrate **H3K9 (1-16)** (blue bars) and the substrate upon addition of S-adenosylmethionine (SAM) (red bars). F-Fo values calculated by measuring the observed fluorescence intensity at λ_{max} subtracted from the blank (dye-derived calixarene). Percent changes were calculated by the following equations: % for A = $(((F\text{-}F_0)_{H3K9me2} - (F\text{-}F_0)_{H3K9}) / (F\text{-}F_0)_{H3K9me2}) \times 100$ and % for B $(((F\text{-}F_0)_{H3K9\text{-}SAM} - (F\text{-}F_0)_{H3K9}) / (F\text{-}F_0)_{H3K9\text{-}SAM}) \times 100$.

SAM is in fact a common co-substrate for hundreds of enzymes (e.g. DNA methyltransferases, protein methyltransferases, and catechol-*O*-methyltransferases). The sensitivity of the other hosts to SAM suggests that they might be developed in the future

as assay reagents that detect changes in SAM concentration during the course of enzymatic reactions. But since the concentrations of SAM in an enzymatic reaction are normally high compared to those of substrate, and since the SAM concentrations are decreasing during the course of a methylation reaction, the sensitivity to SAM (host **12** and **13**) is another reason that these particular dye-derived hosts would not be suitable for this G9a assay. TRITC-derived host **11** shows a desirable small change of 8% between the substrate with and without the SAM co-substrate and a big percent change of 70% between the substrate and product supporting its use as our sensor for measuring real-time methylation. Full emission spectra of each sample in the experiments on the best host, **11**, are shown in Figure 3.8.

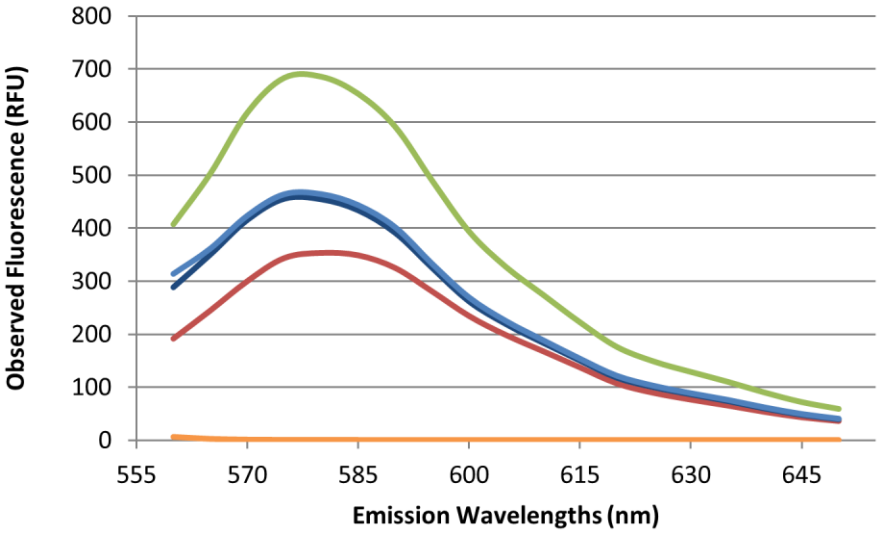


Figure 3.8 Representative spectra showing the substrate-product discrimination experiments. Orange line represents blank buffer D and dH_2O well. The other spectra show an increase in fluorescence from TRITC-derived host **11** at 0.5 μM (red line) upon being treated with product **H3K9me2 (1-16)** at 12.5 μM (green line). Dark blue line represents the **H3K9 (1-16)** substrate at 12.5 μM, whereas the lighter blue shows result of adding both **H3K9 (1-16)** substrate at 12.5 μM and S-adenosylmethionine (**SAM**) co-substrate at 125 μM. Samples were excited at 532 nm and emission observed at 560-650 nm with an emission cut off of 550 nm.

3.3.2.2 Second aim: Reading in real-time methylation of Histone 3 lysine 9 (1-16) peptide tail using tetramethylrhodamine isothiocyanate-derived host 11 (TRITC-derived host 11)

The second generation calixarenes with a covalently attached dye were not sensitive to an environment containing chloride and therefore were able to be tested with enzymes under conventional buffer conditions that are better able to support activity of the enzyme. The assay was done in a 96-well black optical bottom plate in buffer **D**. A **kinetic run** mode was used here again to determine the change in fluorescence over time. Details for the setup of each well are given in the Experimental Section (Table 3.3). Figure 3.9 A shows the increase in fluorescence over 90 minutes from the level of the substrate control well (blue line) up to the level of the product control well (green line). This suggests a) that the enzymatic reaction is working, and b) that the sensor is able to track the progress of the reaction with time.

Samples from the experiment shown in Figure 3.9 A were injected onto the LC/MS by Kevin Daze to see if we would be able to detect a mass change after the methylation process. Although it was hard to correlate the mass of the substrate and product peptide to a specific peak, we were able to see changes in the SAM-related peaks. As seen in the appendix B, both SAM and SAH masses were detected. The well containing buffer **D**, host **11** and SAM co-factor showed a big peak with the mass of SAM and a smaller peak, with the mass of SAH. However, the reaction well containing buffer **D**, host **11**, SAM co-factor, initial substrate peptide and G9a methyltransferase showed a bigger peak correlating to the mass of SAH. Although this does not directly prove that the substrate peptide was methylated, it is a good indication of the transfer of the methyl group from SAM to the substrate peptide, therefore showing an increased SAH peak in the reaction well mixture. Another kinetic run was done over a period of 60 minutes at room temperature to ensure enzymatic activity. The sample reaction was given to Dr. Ori Granot and a MALDI experiment was conducted. The reaction sample was filtered through a zip-tip and spotted onto a MALDI plate. Appendix C shows the MALDI-trace where some of the initial substrate peptide mass remains, as well as the mono-, di-, and tri-methylated product masses which confirm the transfer of a methyl group, once, twice or three times, showing that G9a methylation occurs.

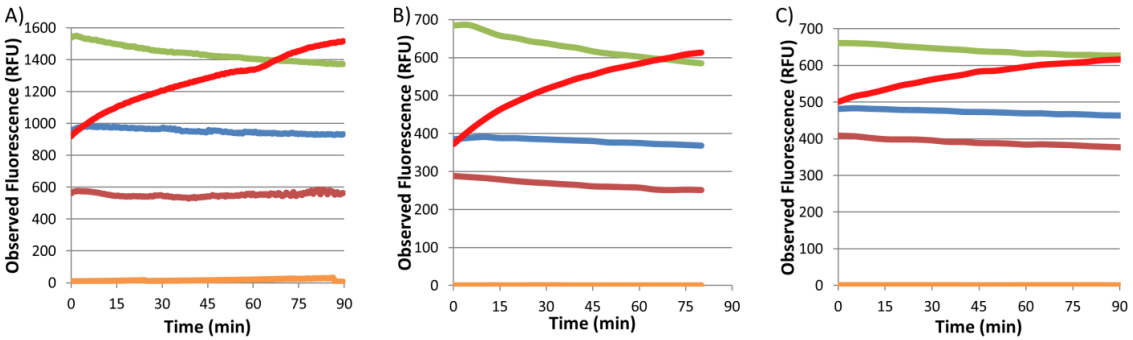


Figure 3.9 Kinetic runs showing the methylation of H3K9 (1-16) substrate over 90 minutes shown at 5 minute intervals. Orange line indicates buffer D well, red line indicates TRITC-sensor well at 0.5 µM, blue line indicates **H3K9 (1-16)** with S-adenosylmethionine (SAM), green line indicates positive control containing **H3K9me2 (1-16)** and the bright red line indicating the reaction well upon addition of the enzyme. A) Run at 30°C with final concentrations of 27.5 µM for peptides, 275 µM for SAM, and 2 µg of G9a. B) Run at 30°C with final concentrations of 25 µM for peptides, 250 µM for SAM, and 0.5 µg of G9a. This experiment was run only up to 80 min. C) Run at room temperature with final concentrations of 12.5 µM for peptides, 125 µM for SAM, and 0.25 µg of G9a. Plate was excited at 532 nm and emission observed at 585 nm.

Two additional kinetic runs were made with progressively lower concentrations of reagents, and each suggested that both the enzymatic reaction and sensor continued to operate under each set of conditions Figure 3.9 B and C. It is important to note that in each of the sets of kinetic run data, there is an obvious decrease in intensity for all the curves over time. A similar decrease in fluorescence was observed for the 1st generation host **1** kinetic run in Figure 3.4. Kinetic runs in Figure 3.9 A and B were done at 30°C whereas C was run at room temperature. Since samples were prepared at room temperature prior to insertion in the plate reader, it is possible that slow thermal equilibration up to the set point of 30°C inside of the plate reader is a reason for these readings. This potential problem arises because no direct measurement of temperature is reported by the plate reader, and so the effect of temperature on emission intensity and the rate of thermal equilibration are unclear from data collected so far. The other, simpler possibility is that significant photobleaching occurs during the course of the experiments.

In order to understand the reasons behind the decrease in intensity, four experiments were done in duplicate. The first involved pre-equilibrating samples inside the plate reader for 15 minutes at 30°C or at room temperature, and recording emission spectra for samples of **11** with and without product **H3K9me2 (1-16)**. Figure 3.10 A reports on these data, and shows that there is in fact a larger change in fluorescence between starting material and product at 30°C (dashed lines) than at room temperature (solid lines). These results are not consistent with the kinetic run data in Figure 3.9, in which signal decreases slowly over time. Figure 3.10 B shows kinetic runs over 60 minutes for samples that were pre-equilibrated in the plate reader for 15 minutes in the dark at either room temperature or 30°C prior to the start of the kinetic run. All of these samples show decreases in intensities with time that can only be attributed to photobleaching over time that is independent of reaction temperature. These results explain the decrease in intensities observed for all 'control' wells in Figure 3.9. Additionally, the fact that the intensity of the reaction wells ends up slightly higher than those of the wells containing pure product (seen most clearly in Figure 3.9 A), might be explained by a difference in photobleaching rates in the reaction wells and the product control wells, which contain the same dye-derived host but slightly different other components. Since our assays are run in 200 µL of total volume, we can imagine that the initial increase in fluorescence at less than 15 minutes seen in Figure 3.10 B in particular is due to the increase in temperature and equilibration of the sample solution while the decrease over time is due to photobleaching.

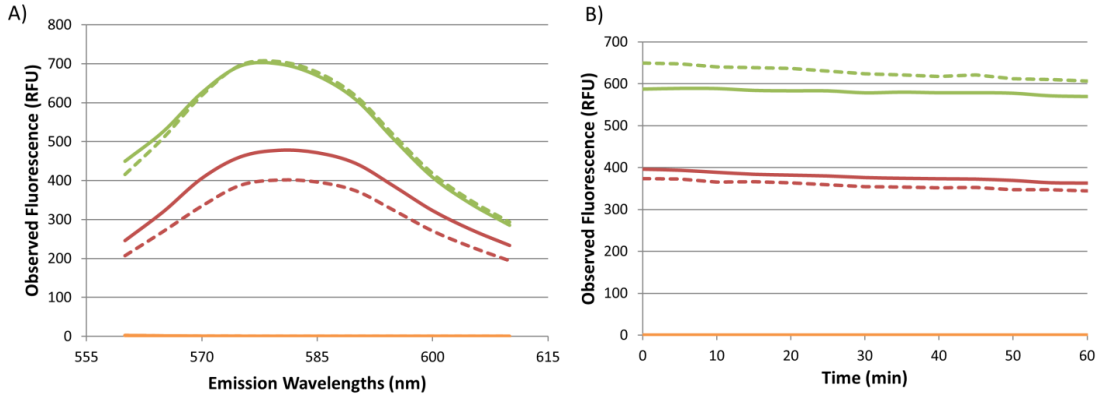


Figure 3.10 Temperature and photobleaching effects. Orange line represents blank buffer **D** and dH_2O well. TRITC-derived host **11** 0.5 µM (red line) increases in fluorescence upon binding of product **H3K9me2 (1-16)** 12.5 µM (green line). A) Plate was excited at 532 nm and emission observed at 560-650 nm with an emission cut off of 550 nm at room temperature (solid lines) and 30°C (dashed lines) after 15 minutes of equilibration in the plate reader. B) Plate was excited at 532 nm and emission observed at 580 nm with an emission cut off of 550 nm at room temperature (solid lines) and 30°C (dashed lines) after 15 minutes of equilibration in the plate reader.

3.3.3 Reading in real-time inhibition of G9a methyltransferase

One major reason for inventing such an enzyme assay is the potential to use it for the identification and characterization of enzyme inhibitors. One well known and readily accessible inhibitor of G9a is S-adenosylhomocysteine (SAH), [65, 134, 141] the by-product arising after transfer of the methyl group from the SAM co-substrate. To test the ability of SAH to act as an inhibitor, a real-time monitoring of methylation with the addition of SAH at different concentrations was done. A 96-well black optical bottom plate was used and a similar setup was done as explained previously for the monitoring of the methylation over a course of time. Three more wells were added, all containing the same components as the reaction mixture well with G9a and **H3K9 (1-16)** substrate with SAM co-substrate (red line in Figure 3.11) as well as increasing concentrations of SAH. Concentrations of SAH were chosen to achieve ratios of 0.25:1, 0.5:1 and 1:1 for SAH:SAM concentrations. Figure 3.11 for both A and B shows the kinetic run of the inhibition studies over 90 minutes at room temperature.

The reaction mixture well containing G9a (red line) increased over time towards the positive control **H3K9me2 (1-16)** well (green line) in Figure 3.11 A. Addition of SAH decreased the methylation but was not dose-dependent in a simple way. In other

words, the biggest inhibition was seen with the 125 µM SAH well (1:1 SAH: SAM) followed by the 31.25 µM SAH well, (0.25:1 SAH:SAM) then the 61.5 µM SAH well (0.5:1 SAH:SAM). Figure 3.11 B shows a repetition of the inhibition study where the G9a reaction does get inhibited in a dose-dependent manner but another obstacle is encountered. In this trial, the uninhibited G9a reaction well (red line) does not move as far as expected toward the positive control well **H3K9me2 (1-16)** (green line). But the inhibition in this set of averaged duplicates is more cleanly dose dependent. These enzymatic reactions are quite sensitive to the environment and this could have been due to a few factors, one of which may have been the stability of SAM co-substrate. SAM co-substrate is most stable when stored in H_2SO_4 at -80°C and should not undergo thaw/freeze cycles. If the co-substrate, substrate or enzyme were not stored properly before use, the enzymatic reaction would not be successful.

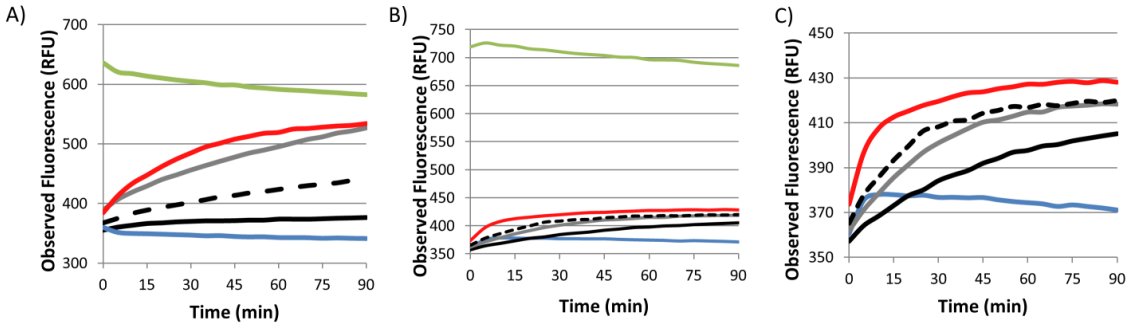


Figure 3.11 SAH inhibition studies. In real-time methylation measured using a SpectraMax® M5 / M5e Microplate Reader at room temperature over 90 minutes shown in 5 minute intervals. Positive control 12.5 µM **H3K9me2 (1-16)** indicated in green, G9a reaction well indicated in bright red, 12.5 µM substrate peptide **H3K9 (1-16)** indicated in blue and different SAH concentrations 31.25 µM (black dashes), 61.5 µM (grey solid line) and 125 µM (black solid line).

3.3.4 Preliminary studies and barriers

Preliminary studies done for HDAC1 and LSD1 enzymes did not lead to a successful readout of enzyme activity. Selected data from these efforts are shown below in order to highlight some of the barriers to success and general lessons learned when using 2nd generation calixarenes.

3.3.4.1 Buffer matters

Numerous buffer studies for every new enzyme need to be done in order to first achieve good host sensitivity towards substrate and product peptides and second monitor the enzymatic activity. Data collected for similar peptides as the substrate and product peptides for the enzyme LSD1 nicely illustrate this issue. A 96-well black optical bottom plate was used with FITC-derived host **12,** four different buffers **A**, **B**, **E**, and **F** and unmethylated **H3K4 (1-12)** (product) and trimethylated **H3K4me3 (1-12)** (substrate). The sample was excited at 450 nm and emission intensity was measured at 520 nm. Figure 3.12 shows the different emission intensities of FITC-derived host **12** in the presence of substrate or product in each of the four different buffers varying from 53% to 86% difference in substrate-product discrimination.

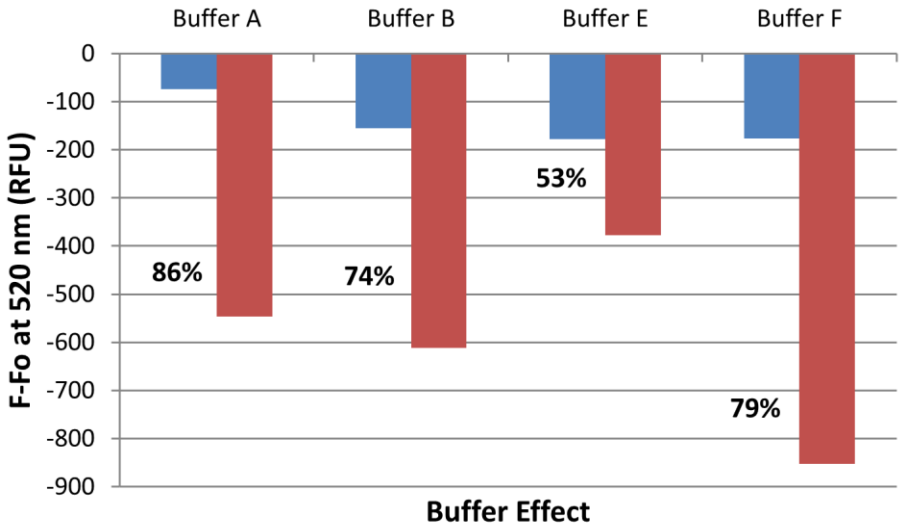


Figure 3.12 Buffer effect with FITC-derived host 12. Buffers **A**, **B**, **E** and **F** with 10 µM of either **H3K4 (1-12)** (blue) or **H3K4me3 (1-12)** (red). Plate was made up to 200 µL with dH_2O and was excited at 450 nm and emission at 520 nm is plotted as $F\text{-}F_0$, where F is the fluorescence observed upon guest binding and F_0 is the fluorescence of the dye-derived calixarene. Percent change calculated as shown in caption of Figure 3.3.

3.3.4.2 Substrate length matters

Showing that a particular host in certain buffer conditions is a good sensor for one peptide does not automatically mean that this sensor will work for other, longer peptides. LSD1 demethylase has been shown *in vitro* to demethylate substrates with a minimum length of 20 amino acids, placing a requirement to use such long peptides in any assay for LSD1 activity.[142] We have previously shown (Figure 3.7) that DACITC-derived host **14** acts as a good sensor in buffer **D** for the 16-mer G9a substrate **H3K9 (1-16)** and product **H3K9me2 (1-16)**, and the same was as evident for **H3K4 (1-21)** LSD1 product and **H3K4me2 (1-21**) LSD1 substrate (Figure 3.13). Although these two peptides do not have the dimethylation modification on the same position, they give an indirect example, measured in the same buffer, of the difference that peptide length may have on the host's ability to discriminate among modification states. The unmethylated peptides **H3K4 (1-21)** and **H3K9 (1-16)** shown in blue showed a smaller change in fluorescence than do their dimethylated counterparts shown in red. The shorter 16-mer peptide showed the same discrimination as the 21-mer peptide of 32% but the later peptide showed smaller raw fluorescence emission intensity.

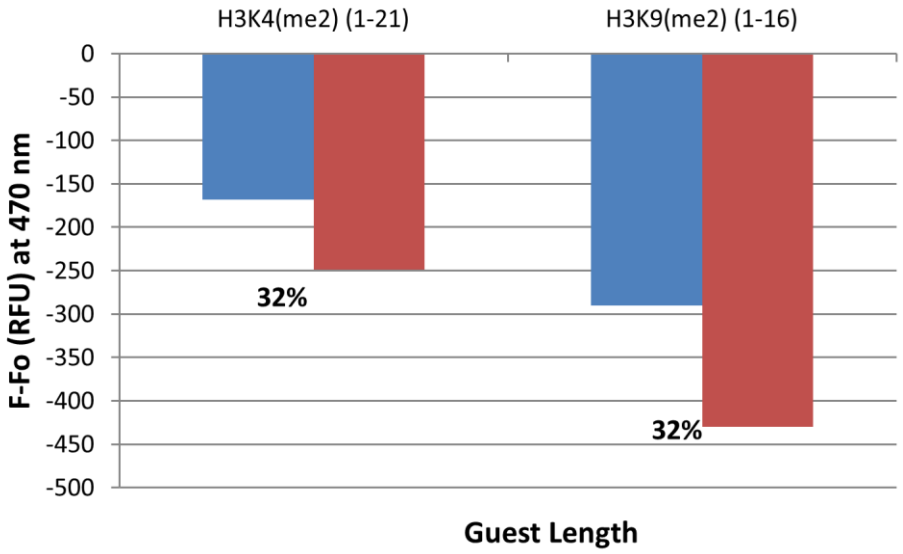


Figure 3.13 Substrate length effect with DACITC-derived host 14. 10 µM of either 21-mer or 16-mer unmethylated and dimethylated peptide was used. **H3K4 (1-21)** (NH_2-ARTKQTARKSTGGKAPRKQLAY-C(O)NH_2) and **H3K9 (1-16)** NH_2-ARTKQTARKSTGGKAPY-C(O)NH_2) shown in blue, whereas **H3K4me2 (1-21)** (NH_2-ARTK**me2**QTARKSTGGKAPRKQLAY-C(O)NH_2) and **H3K9me2 (1-16)** NH_2-ARTKQTARK**me2**STGGKAPY-C(O)NH_2) are shown in red. Plate was made up to 200 µL with buffer **D**. The plate was excited at 380 nm and emission at 470 nm is plotted as F-F_0, where F is the fluorescence observed upon guest binding and F_0 is the fluorescence of the dye-derived calixarene. Measurements were taken at 37°C. Percent change calculated as shown in caption of Figure 3.3

3.3.4.3 Sealer effect

Some enzymes have optimal enzymatic activity at higher temperature. For example, both HDAC1 deacetylase[143] and LSD1 demethylase[126] have been shown to function at 37°C. When running these assays in a 96-well microplate with a final volume of 200 µL, higher temperatures present the problem that evaporation of sample is significant during the time course of the experiments. Plate sealers (VIEWseal, Greiner) prevent evaporation of sample for longer kinetic reads and are sold as being highly transparent for optical measurements. Figure 3.14 shows an example of a substrate-product discrimination example with FITC-benzylamine derived host **13** and unmethylated **H3K4 (1-21)** and dimethylated **H3K4me2 (1-21)**. Since **H3K4me2 (1-21)** is the substrate for LSD1, and a temperature of 37°C is required for enzymatic activity,

an assay was run in buffer **G** with and without a plate sealer. The sample was excited at 480 nm and emission at 520 nm was recorded. F-F_0 values showed a smaller discrimination between substrate and product peptide of 12% for the sample covered with a plate sealer compared to 33% without the sealer, showing that the film is not completely innocent, and placing a strict limitation on the temperatures at which assays can be run. Varying the host, buffer, and type of plate sealer used are all options that need to be considered to achieve better substrate-product discrimination for enzymes that are active only at higher temperatures.

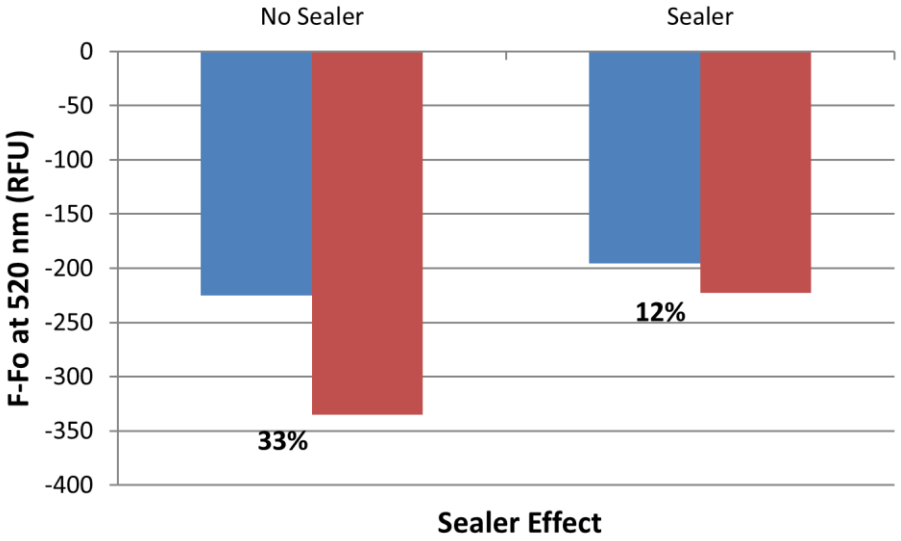


Figure 3.14 Sealer effect with FITC-Ba-derived host 13. 10 µM of **H3K4 (1-21)** (NH_2-ARTKQTARKSTGGKAPRKQLAY-C(O)NH_2) (blue) or **H3K4me2 (1-21)** (NH_2-ARTK**me2**QTARKSTGGKAPRKQLAY-C(O)NH_2) (red) plated in presence or absence of plate sealer. Plate was made up to 200 µL in buffer **G** and dH_2O. The plate was excited at 520 nm and emission at 480 nm is plotted as F-F_0, where F is the fluorescence observed upon guest binding and F_0 is the fluorescence of the dye-derived calixarene. Measurements were taken at 37°C. Percent change calculated as shown in caption of Figure 3.3.

3.3.4.4 Enzyme effect

Another obstacle encountered during the preliminary studies was the effect of the enzyme itself on the dye-derived host. Figure 3.15 shows the change in fluorescence for DACITC-derived host **14** in buffer **H** with the addition of either, **H4K5 (3-5)** (Ac-RGKY-C(O)NH_2) , **H4K5ac (3-5)** (Ac-RGKacY-C(O)NH_2) or HDAC1 enzyme. The HDAC1 product **H4K5 (3-5)** showed a bigger change in fluorescence than the acetylated

substrate **H4K5ac (3-5)** of 27%, however, the addition of HDAC1 enzyme only to **14** showed an even bigger change in fluorescence of 39%. The increased change in fluorescence caused by the enzyme alone causes problems for monitoring of the enzymatic reaction. Addition of an enzyme that causes a bigger change in fluorescence than the addition of product restricts this dye-derived host to act as a sensor for monitoring the deacetylation of **H4K5ac (3-5)** (Ac-RGKacY-C(O)NH_2).

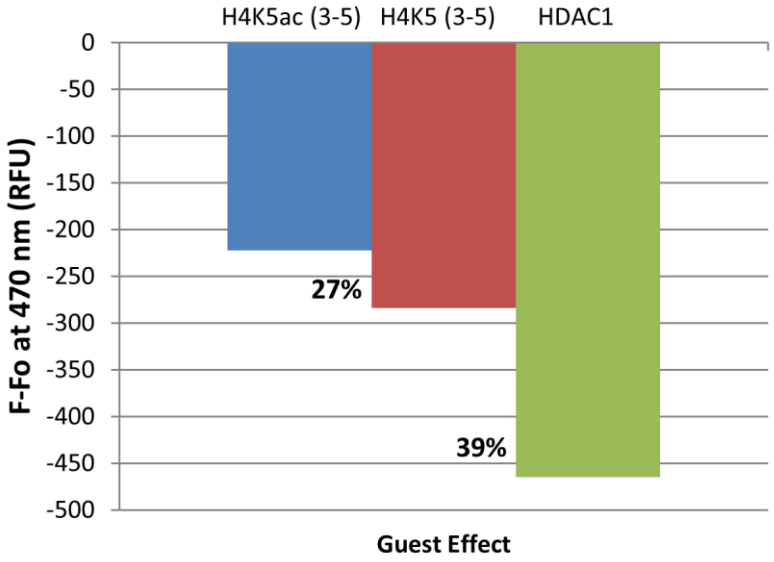


Figure 3.15 Guest effect on DACITC-derived host 14. 10 µM of either unmodified (Ac-RGKY-C(O)NH_2) or acetylated peptides (Ac-RG**Kac**Y-C(O)NH_2) and 0.17 µg of HDAC1 enzyme alone. Plate was made up to 200 µL in buffer **H** and dH_2O. The plate was excited at 380 nm and emission at 470 nm is plotted as F-F_0, where F is the fluorescence observed upon guest binding and F_0 is the fluorescence of the dye-derived calixarene. Percent change calculated as shown in caption of Figure 3.3.

3.4 Conclusions and Future Work

The very difficult task of developing a fluorescence-based assay for tracking enzymatic activity in real-time was achieved by Nau *et al*. as previously mentioned for the Dim-5 methyltransferase enzyme.[102] Using **PSC4** and the **LCG**, Nau's group showed the ability of this host-dye pair to act as a sensor for the trimethylation of lysine 9 of histone 3. While I was unable to optimize the 1st generation approach as a fluorescence-based supramolecular tool to measure PTM enzyme activity in real-time, I believe that it can still be used for more diverse enzyme assays requiring simpler buffers.

If I were to further explore the real-time assay using the 1st generation calixarenes, I might approach this particular project from a different angle. Instead of attempting to optimize my assay within the scope of my developing assay, I would begin with an already functioning assay from literature. After selecting a buffer that accommodates the host-dye pair for a methyltransferase, I would use an enzyme-coupled assay to identify the proper substrate and co-substrate concentrations needed for the reaction to occur. Choosing a secondary assay that monitors enzymatic activity in a homogenous solution would be critical to verify a) that the enzyme itself is active b) that the peptide length of the substrate is suitable for the enzyme and c) that the substrate and co-substrate are used in the correct concentrations.

In this Chapter, second generation calixarenes, in particular TRITC-derived host **11** showed the ability to act as a sensor that can discriminate unmethylated and methylated peptides. The sensor's sensitivity allowed for real-time measurement of methylation of a 16-mer peptide H3K9 (1-16) using S-adenosylmethionine co-substrate and G9a methyltransferase.

The development of these novel unimolecular, second generation dye-derived calixarenes is an advantage over using a bi-molecular first generation approach system like that of Nau's because it enables the researcher to use conditions optimal for enzymatic activity. Unlike Nau's system, which can really only be applied in systems where enzymes can function in buffers that do not contain any lucigenin quenchers, this homogeneous "one-pot" assay may be used for numerous enzymes in different buffer conditions. So far, this new second generation approach system worked on G9a lysine methyltransferase, and lysine demethylases JMJ2DA which demethylates H3K36me3[144]

(work done by Kevin Daze) but can be expected to work on other methyltransferases such as Set2[145] which methylates H3K36, arginine methyltransferases such as CARM1[28] that dimethylates H3R17 asymmetrically, arginine demethylase JMJD6[31] as well as other types of enzymes like deacetylases and acetyltransferases. The simplicity of this experimental setup is motivating for the optimization of several other sensors like this one.

The development of this new second-generation approach was not focused on the extraction of enzyme kinetics but rather the development of a convenient tool for the rapid, continuous, high-throughput monitoring of enzymatic activity. Although this assay type still needs optimization in order to allow for universality for other enzymes and systems, it has the potential to be a much easier setup than current assays such as ELISA-based detection of methylated substrates,[52] AlphaScreen,[56] microfluidic capillary electrophoresis[146] and enzyme-coupled measurements of SAH production.[71] The studies reported here are simple and easy to carry out and will allow further experiments to be done to gain more understanding about the sensitivity and limits of these dye-derived calixarene sensors.

Since the end of the G9a project, Kevin Daze has synthesized at least five new dye-derived calixarenes that may be useful for future study for the optimization of other sensors for different enzymes. Further detailed studies regarding the photochemical mechanisms of quenching and fluorescence for these dye-derived hosts will remain as work for a future student as well as MALDI-TOF or LC/MS experiments that can accompany the enzymatic studies to allow for quantitative analysis of enzymatic activity.

Bibliography

2. Jenuwein, T. & Allis, C.D. Translating the histone code. *Science* **293**, 1074-1080 (2001).

3. Kornberg, R.D. Chromatin Structure: A repeating unit of histones and DNA. *Science* **184**, 868-871 (1974).

4. Kornberg, R.D. & Thomas, J.O. Chromatin Structure: Oligomers of the histones. *Science* **184**, 865-868 (1974).

5. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8A resolution. *Nature* **389**, 251-260 (1997).

6. Olins, A.L. & Olins, D.E. Spheroid chromatin units (ν Bodies). *Science* **183**, 330-332 (1974).

7. Olins, A.L., Carlson, R.D., Wright, E.B. & Olins, D.E. Chromatin υ bodies: isolation, subfractionation and physical characterization. *Nucl Acids Res* **3**, 3271-3292 (1976).

8. Oudet, P., Gross-Bellard, M. & Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**, 281-300 (1975).

9. Richmond, T.J. & Davey, C.A. The structure of DNA in the nucleosome core. *Nature* **423**, 145-150 (2003).

10. Durrin, L.K., Mann, R.K., Kayne, P.S. & Grunstein, M. Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65**, 1023-1031 (1991).

11. Wallis, J.W., Hereford, L. & Grunstein, M. Histone H2B genes of yeast encode two different proteins. *Cell* **22**, 799-805 (1980).

12. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349-352 (1997).

13. Garcia-Ramirez, M., Rocchini, C. & Ausio, J. Modulation of chromatin folding by histone acetylation. *J Biol Chem* **270**, 17923-17928 (1995).

14. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).

15. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**, 41-5 (2000).

16. Richards, E.J. & Elgin, S.C.R. Epigenetic codes for heterochromatin formation and silencing: Rounding up the usual suspects. *Cell* **108**, 489-500 (2002).

17. Musselman, C.A., Lalonde, M.-E., Cote, J. & Kutateladze, T.G. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* **19**, 1218-1227 (2012).

18. Chi, P., Allis, C.D. & Wang, G.G. Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer* **10**, 457-469 (2010).

19. Ruthenburg, A.J., Li, H., Patel, D.J. & David Allis, C. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**, 983-994 (2007).

20. Dilworth, D., Gudavicius, G., Leung, A. & Nelson, C.J. The roles of peptidyl-proline isomerases in gene regulation. *Biochem Cell Biol* **90**, 55-69 (2012).

21. Nelson, C.J., Santos-Rosa, H. & Kouzarides, T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126**, 905-916 (2006).

22. Sterner, D.E. & Berger, S.L. Acetylation of histones and transcription-related factors. *Microbiol and Mol Biol Rev* **64**, 435-459 (2000).

23. Kurdistani, S.K. & Grunstein, M. Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* **4**, 276-284 (2003).

24. Xu, F., Zhang, K. & Grunstein, M. Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* **121**, 375-385 (2005).

25. Xie, W. et al. Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell* **33**, 417 - 427 (2009).

26. Scharf, A.N.D. & Imhof, A. Every methyl counts – Epigenetic calculus. *FEBS Lett* **585**, 2001-2007 (2011).

27. Pekowska, A. et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *The EMBO journal* **30**, 4198-4210 (2011).

28. Bauer, U.-M., Daujat, S., Nielsen, S.J., Nightingale, K. & Kouzarides, T. Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Reports* **3**, 39-44 (2002).

29. Fabbrizio, E. et al. Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO Reports* **3**, 641-645 (2002).

30. Anand, R. & Marmorstein, R. Structure and Mechanism of Lysine-specific demethylase enzymes. *J Biol Chem* **282**, 35425-35429 (2007).

31. Chang, B., Chen, Y., Zhao, Y. & Bruick, R.K. JMJD6 is a histone arginine demethylase. *Science* **318**, 444-447 (2007).

32. Smith, B.C. & Denu, J.M. Chemical mechanisms of histone lysine and arginine modifications.*BBA - Gene Reg Mech* **1789**, 45-57 (2009).

33. Klose, R.J. & Zhang, Y. Regulation of histone methylation by demethylimination and demethylation. *Nat Rev Mol Cell Biol* **8**, 307-318 (2007).

34. Vossenaar, E.R., Zendman, A.J.W., van Venrooij, W.J. & Pruijn, G.J.M. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *BioEssays* **25**, 1106-1118 (2003).

35. Dawson, M.A. et al. JAK2 phosphorylates histone H3Y41 and excludes HP1 alpha from chromatin. *Nature* **461**, 819-822 (2009).

36. Jin, B., Li, Y. & Robertson, K.D. DNA Methylation: Superior or subordinate in the epigenetic hierarchy? *Gen & Cancer* **2**, 607-617 (2011).

37. Abmayr, Susan M. & Workman, Jerry L. Holding on through DNA replication: Histone modification or modifier? *Cell* **150**, 875-877 (2012).

38. Fullgrabe, J., Hajji, N. & Joseph, B. Cracking the death code: apoptosis-related histone modifications. *Cell Death Differ* **17**, 1238-1243 (2010).

39. Yoo, K.H. & Hennighausen, L. EZH2 methyltransferase and H3K27 methylation in breast cancer. *Int J Biol Sci* **8**, 59-65 (2012).

40. Knutson, S.K. et al. A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat Chem Biol* **8**, 890-896 (2012).

41. Bernstein, E. et al. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* **26**, 2560-2569 (2006).

42. Bernard, D. et al. CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the Ink4a/Arf locus. *Oncogene* **24**, 5543-5551 (2005).

43. Varier, R.A. & Timmers, H.T.M. Histone lysine methylation and demethylation pathways in cancer. *BBA - Reviews on Cancer* **1815**, 75-89 (2011).

44. Sarma, K., Nishioka, K. & Reinberg, D. in Methods in enzymology 255-269 (Academic Press, 2003).

45. Egelhofer, T.A. et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* **18**, 91-93 (2011).

46. Fuchs, S.M. & Strahl, B.D. Antibody recognition of histone post-translational modifications: emerging issues and future prospects. *Epigenomics* **3**, 247-249 (2011).

47. Nishikori, S. et al. Broad ranges of affinity and specificity of anti-histone antibodies revealed by a quantitative peptide immunoprecipitation assay. *J Mol Biol* **424**, 391-399 (2012).

48. Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L. & Strahl, B.D. Influence of combinatorial histone modifications on antibody and effector protein recognition. *Current Biol* **21**, 53-58 (2011).

49. Rongen, H.A.H., Hoetelmans, R.M.W., Bult, A. & Van Bennekom, W.P. Chemiluminescence and immunoassays. *J Pharm Biomed Anal* **12**, 433-462 (1994).

50. Li, K.K., Luo, C., Wang, D., Jiang, H. & Zheng, Y.G. Chemical and biochemical approaches in the study of histone methylation and demethylation. *Med Res Rev* **32**, 815-867 (2012).

51. Cheng, D. et al. Small molecule regulators of protein arginine methyltransferases. *J Biol Chem* **279**, 23892-23899 (2004).

52. Kubicek, S. et al. Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol Cell* **25**, 473-481 (2007).

53. Spannhoff, A. et al. Target-based approach to inhibitors of histone arginine methyltransferases. *J Med Chem* **50**, 2319-2325 (2007).

54. Bosse, R. Principles of Alphascreen. (PerkinElmer Application Note, 2001).

55. Guenat, S. et al. Homogeneous and nonradioactive high-throughput screening platform for the characterization of kinase inhibitors in cell lysates. *J Biomol Scr* **11**, 1015-1026 (2006).

56. Quinn, A.M., Allali-Hassani, A., Vedadi, M. & Simeonov, A.Y. A chemiluminescence-based method for identification of histone lysine methyltransferase inhibitors. *Mol BioSyst* 782 (2010).

57. Kawamura, A. et al. Development of homogeneous luminescence assays for histone demethylase catalysis and binding. *Anal Biochem* **404**, 86-93 (2010).

58. Eglen, R.M., T, R., P., R. & Rouleau, N. The use of alphaScreen technology in HTS: Current status. *Curr Chem Genom* **1**, 2-10 (2008).

59. Wigle, T.J. et al. Screening for inhibitors of low-affinity epigenetic peptide-protein interactions: An AlphaScreen™-based assay for antagonists of methyl-lysine binding proteins. *J Biomol Scr* **15**, 62-71 (2010).

60. Gauthier, N. et al. Development of homogeneous nonradioactive methyltransferase and demethylase assays targeting histone H3 lysine 4. *J Biomol Scr* **17**, 49-58 (2012).

61. Suh-Lailam, B.B. & Hevel, J.M. A fast and efficient method for quantitative measurement of S-adenosyl-l-methionine-dependent methyltransferase activity with protein substrates. *Anal Biochem* **398**, 218-224 (2010).

62. Greiner, D., Bonaldi, T., Eskeland, R., Roemer, E. & Imhof, A. Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. *Nat Chem Biol* **1**, 143-145 (2005).

63. Gowher, H., Zhang, X., Cheng, X. & Jeltsch, A. Avidin plate assay system for enzymatic characterization of a histone lysine methyltransferase. *Anal Biochem* **342**, 287-291 (2005).

64. Rathert, P. et al. Protein lysine methyltransferase G9a acts on non-histone targets. *Nat Chem Biol* **4**, 344 - 346 (2008).

65. Dhayalan, A., Dimitrova, E., Rathert, P. & Jeltsch, A. A continuous protein methyltransferase (G9a) assay for enzyme activity measurement and inhibitor screening. *J Biomol Scr* **14**, 1129-1133 (2009).

66. Whetstine, J.R. et al. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125**, 467-481 (2006).

67. Young, N.L. et al. High throughput characterization of combinatorial histone codes. *Mol & Cell Proteomics* **8**, 2266-2284 (2009).

68. Silva, A.M.N., Vitorino, R., Domingues, M.R.M., Spickett, C.M. & Domingues, P. Post-translational modifications and mass spectrometry detection. *Free Rad Biol Med* **65**, 925-941 (2013).

69. Phanstiel, D. et al. Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. *Proc Natl Acad Sci* **105**, 4093-4098 (2008).

70. Nicklay, J.J. et al. Analysis of Histones in Xenopus laevis: II. Mass spectrometry reveals an index of cell type-specific mofications on H3 and H4. *J Biol Chem* **284**, 1075-1085 (2009).

71. Collazo, E., Couture, J.F., Bulfer, S. & Trievel, R.C. A coupled fluorescent assay for histone methyltransferases. *Anal Biochem* **342**, 86-92 (2005).

72. Palmer, N.A., Sattler, S.E., Saathoff, A.J. & Sarath, G. A Continuous, quantitative fluorescent assay for plant caffeic acid O-methyltransferases. *J Agr Food Chem* **58**, 5220-5226 (2010).

73. Hendricks, C.L., Ross, J.R., Pichersky, E., Noel, J.P. & Zhou, Z.S. An enzyme-coupled colorimetric assay for S-adenosylmethionine-dependent methyltransferases. *Anal Biochem* **326**, 100-105 (2004).

74. Ellman, G.L. Tissue sulfhydryl groups. *Arch. Biochem.Biophys.* **82**, 70-77 (1959).

75. Dorgan, K.M. et al. An enzyme-coupled continuous spectrophotometric assay for S-adenosylmethionine-dependent methyltransferases. *AnalBiochem* **350**, 249-255 (2006).

76. Sakurai, M. et al. A miniaturized screen for inhibitors of Jumonji histone demethylases. *Mol BioSyst* **6**, 357-364 (2010).

77. Couture, J.-F., Collazo, E., Ortiz-Tello, P.A., Brunzelle, J.S. & Trievel, R.C. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nat Struct Mol Biol* **14**, 689-695 (2007).

78. Ueda, R. et al. Identification of cell-active lysine specific demethylase 1-selective inhibitors. *J Am Chem Soc* **131**, 17536-17537 (2009).

79. Williams, K. & Scott, J. Enzyme assay design for hight-throughput screening. *High T Scr,* 107-126 (2009).

80. Sawan, C. & Herceg, Z. Histone modifications and cancer. *Adv Genet*, 57-85 (2010).

81. Qi, W. et al. Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc Natl Acad Sci* **109**, 21360-21365 (2012).

82. Verma, S.K. et al. Identification of potent, selective, cell-active inhibitors of the histone lysine methyltransferase EZH2. *ACS Med Chem Lett* **3**, 1091-1096 (2012).

83. McCabe, M.T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108-112 (2012).

84. Gieni, R.S. & Hendzel, M.J. Polycomb group protein gene silencing, non-coding RNA, stem cells, and cancer. *Biochem Cell Biol* **87**, 711-746 (2009).

85. Sauvageau, M. & Sauvageau, G. Polycomb Group Proteins: Multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* **7**, 299-313 (2010).

86. O'Loghlen, A. et al. MicroRNA regulation of cbx7 mediates a switch of polycomb orthologs during ESC differentiation. *Cell Stem Cell* **10**, 33-46 (2012).

87. Morey, L. et al. Nonoverlapping functions of the polycomb group cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10**, 47-62 (2012).

88. Klauke, K. et al. Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nat Cell Biol* **15**, 353-362 (2013).

89. Aguilo, F., Zhou, M.-M. & Walsh, M.J. Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. *Cancer Res.* **71**, 5365-5369 (2011).

90. Maertens, G.N. et al. Several distinct polycomb complexes regulate and co-localize on the *INK4a* tumor suppressor locus. *PLoS One* **4**, e6380 (2009).

91. Gil, J., Bernard, D., Martinez, D. & Beach, D. Polycomb CBX7 has a unifying role in cellular lifespan. *Nat Cell Biol* **6**, 67-72 (2004).

92. Yap, K.L. et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**, 662-674 (2010).

93. Bracken, A.P. & Helin, K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nat Rev Cancer* **9**, 773-784 (2009).

94. Kaustov, L. et al. Recognition and specificity determinants of the human cbx chromodomains. *J Biol Chem* **286**, 521-529 (2011).

95. Hughes, R.M., Wiggins, K.R., Khorasanizadeh, S. & Waters, M.L. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proc Natl Acad Sci* **104**, 11184-11188 (2007).

96. Daze, K.D., Ma, M.C., Pineux, F. & Hof, F. Synthesis of new trisulfonated calix[4]arenes functionalized at the upper rim, and their complexation with the trimethyllysine epigenetic mark. *Org Lett* **14**, 1512-5 (2012).

97. Beshara, C.S., Jones, C.E., Daze, K.D., Lilgert, B.J. & Hof, F. A simple calixarene recognizes post-translationally methylated lysine. *ChemBioChem* **11**, 63-66 (2010).

98. Daze, K.D. et al. Supramolecular hosts that recognize methyllysines and disrupt the interaction between a modified histone tail and its epigenetic reader protein. *Chem Sci* **3**, 2695-2699 (2012).

99. Daze, K.D. & Hof, F. The cation-pi interaction at protein-protein interaction interfaces: Developing and learning from synthetic mimics of proteins that bind methylated lysines. *Acc Chem Res* **46**, 937-945 (2013).

100. Koh, K.N., Araki, K., Ikeda, A., Otsuka, H. & Shinkai, S. Reinvestigation of calixarene-based artificial-signaling acetylcholine receptors useful in neutral aqueous (water/methanol) solution. *J Am Chem Soc* **118**, 755-758 (1996).

101. Guo, D.S., Uzunova, V.D., Su, X., Liu, Y. & Nau, W.M. Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. *Chem Sci* **2**, 1722-1734 (2011).

102. Florea, M. et al. A Fluorescence-based supramolecular tandem assay for monitoring lysine methyltransferase activity in homogeneous solution. *Chem Eur J* **18**, 3521-3528 (2012).

103. Gamal-Eldin, M.A. & Macartney, D.H. Selective molecular recognition of methylated lysines and arginines by cucurbit[6]uril and cucurbit[7]uril in aqueous solution. *Org & Biomol Chem* **11**, 488-495 (2013).

104. Ingerman, L.A., Cuellar, M.E. & Waters, M.L. A small molecule receptor that selectively recognizes trimethyl lysine in a histone peptide with native protein-like affinity. *Chem Commun* **46**, 1839-1841 (2010).

105. James, L.I., Beaver, J.E., Rice, N.W. & Waters, M.L. A synthetic receptor for asymmetric dimethyl arginine. *J Am Chem Soc* **135**, 6450-5 (2013).

106. Lalor, R., Baillie-Johnson, H., Redshaw, C., Matthews, S.E. & Mueller, A. cellular uptake of a fluorescent calix[4]arene derivative. *J Am Chem Soc* **130**, 2892-2893 (2008).

107. Mueller, A., Lalor, R., Cardaba, C.M. & Matthews, S.E. Stable and sensitive probes for lysosomes: cell-penetrating fluorescent calix[4]arenes accumulate in acidic vesicles. *J Int Soc Anal Cyt* **79**, 126-136 (2011).

108. Wang, K. et al. Highly effective binding of viologens by p-sulfonatocalixarenes for the treatment of viologen poisoning. *J Med Chem* **52**, 6402-6412 (2009).

109. Coleman, A.W. et al. Toxicity and biodistribution of para-sulfonato-calix[4]arene in mice. *New J Chem* **32**, 780-782 (2008).

110. Amblard, M., Fehrentz, J.-A., Martinez, J. & Subra, G. Methods and protocols of modern solid phase peptide synthesis. *Mol Biotechnol* **33**, 239-254 (2006).

111. Hirose, K. A practical guide for the determination of binding constants. *J Incl Phenom Macro Chem* **39**, 193-209 (2001).

112. Bakirci, H. & Nau, W.M. Fluorescence regeneration as a signaling principle for choline and carnitine binding: A refined supramolecular sensor system based on a fluorescent azoalkane. *Adv Funct Mat* **16**, 237-242 (2006).

113. Hennig, A., Bakirci, H. & Nau, W.M. Label-free continuous enzyme assays with macrocycle-fluorescent dye complexes. *Nat Meth* **4**, 629-632 (2007).

114. Bailey, D.M., Hennig, A., Uzunova, V.D. & Nau, W.M. Supramolecular tandem enzyme assays for multiparameter sensor arrays and enantiomeric excess determination of amino acids. *Chem Eur J* **14**, 6069-6077 (2008).

115. Hargrove, A.E., Zhong, Z., Sessler, J.L. & Anslyn, E.V. Algorithms for the determination of binding constants and enantiomeric excess in complex host : guest equilibria using optical measurements. *New J Chem* **34**, 348-354 (2010).

116. Minaker, S.A., Daze, K.D., Ma, M.C.F. & Hof, F. Antibody-free reading of the histone code using a simple chemical sensor array. *J Am Chem Soc* **134**, 11674-11680 (2012).

117. Legg, K.D. & Hercules, D.M. Quenching of lucigenin fluorescence. *J Phys Chem* **74**, 2114-2118 (1970).

118. Nielsen, P.R. et al. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103-107 (2002).

119. Mansfield, R.E. et al. Plant homeodomain (PHD) fingers of CHD4 are Histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. *J Biol Chem* **286**, 11779-11791 (2011).

120. Gao, C. et al. Biophysical probes reveal a “compromise” nature of the methyl-lysine binding oocket in L3MBTL1. *J Am Chem Soc* **133**, 5357-5362 (2011).

121. Herold, J.M. et al. Small-molecule ligands of methyl-lysine binding proteins. *J Med Chem* **54**, 2504-2511 (2011).

122. James, L.I. et al. Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat Chem Biol* **9**, 184-91 (2013).

123. Kireev, D. et al. Identification of non-peptide malignant brain tumor (MBT) repeat antagonists by virtual screening of commercially available compounds. *J Med Chem* **53**, 7625-7631 (2010).

124. Wagner, E.K., Nath, N., Flemming, R., Feltenberger, J.B. & Denu, J.M. Identification and characterization of small molecule inhibitors of a plant homeodomain finger. *Biochem* **51**, 8293-8306 (2012).

125. Unnikrishnan, A., Gafken, P.R. & Tsukiyama, T. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nat Struct Mol Biol* **17**, 430-437 (2010).

126. Shi, Y. et al. Histone Demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953 (2004).

127. Shi, Y.J. et al. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* **19**, 857-864 (2005).

128. Hino, S. et al. FAD-dependent lysine-specific demethylase-1 regulates cellular energy expenditure. *Nat Commun* **3**, 758 (2012).

129. Chin, H.G. et al. Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. *Nucl Acids Res* **35**, 7313-7323 (2007).

130. Sampath, S.C. et al. Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol Cell* **27**, 596-608 (2007).

131. Tachibana, M. et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**, 1779-1791 (2002).

132. Tabet, S. et al. Synthetic trimethyllysine receptors that bind histone 3, trimethyllysine 27 (H3K27me3) and disrupt its interaction with the epigenetic reader protein CBX7. *Bioorg Med Chem* **21**, 7004-7010 (2013).

133. Dsouza, R.N., Hennig, A. & Nau, W.M. - Supramolecular tandem enzyme assays. *Chem Eur J* **12,**3444 (2012).

134. Patnaik, D. et al. Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase. *J Biol Chem* **279**, 53248 - 53258 (2004).

135. Weiss, T. et al. Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigen Chromatin* **3**, 7 (2010).

136. Watt, R.M. & Voss Jr, E.W. Mechanism of quenching of fluorescein by anti-fluorescein IgG antibodies. *Immunochem* **14**, 533-541 (1977).

137. Togashi, D.M., Szczupak, B., Ryder, A.G., Calvet, A. & O'Loughlin, M. Investigating tryptophan quenching of fluorescein fluorescence under protolytic equilibrium. *J Phys Chem A* **113**, 2757-2767 (2009).

138. Arbeloa, F.L., Ojeda, P.R. & Arbeloa, I.L. Flourescence self-quenching of the molecular forms of Rhodamine B in aqueous and ethanolic solutions. *J Lum* **44**, 105-112 (1989).

139. Lakowicz, J.R. Principles of fluorescence spectroscopy (Springer, New York, 2006).

140. Setiawan, D., Kazaryan, A., Martoprawiro, M.A. & Filatov, M. A first principles study of fluorescence quenching in rhodamine B dimers: how can quenching occur in dimeric species? *Phys Chem* **12**, 11238-11244 (2010).

141. Lin, Y. et al. Detecting S-adenosyl-l-methionine-induced conformational change of a histone methyltransferase using a homogeneous time-resolved fluorescence-based binding assay. *Anal Biochem* **423**, 171-177 (2012).

142. Forneris, F., Binda, C., Vanoni, M.A., Mattevi, A. & Battaglioli, E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett* **579**, 2203-2207 (2005).

143. Hassig, C.A. et al. A role for histone deacetylase activity in HDAC1-mediated transcriptional repression. *Proc Natl Acad Sci* **95**, 3519-3524 (1998).

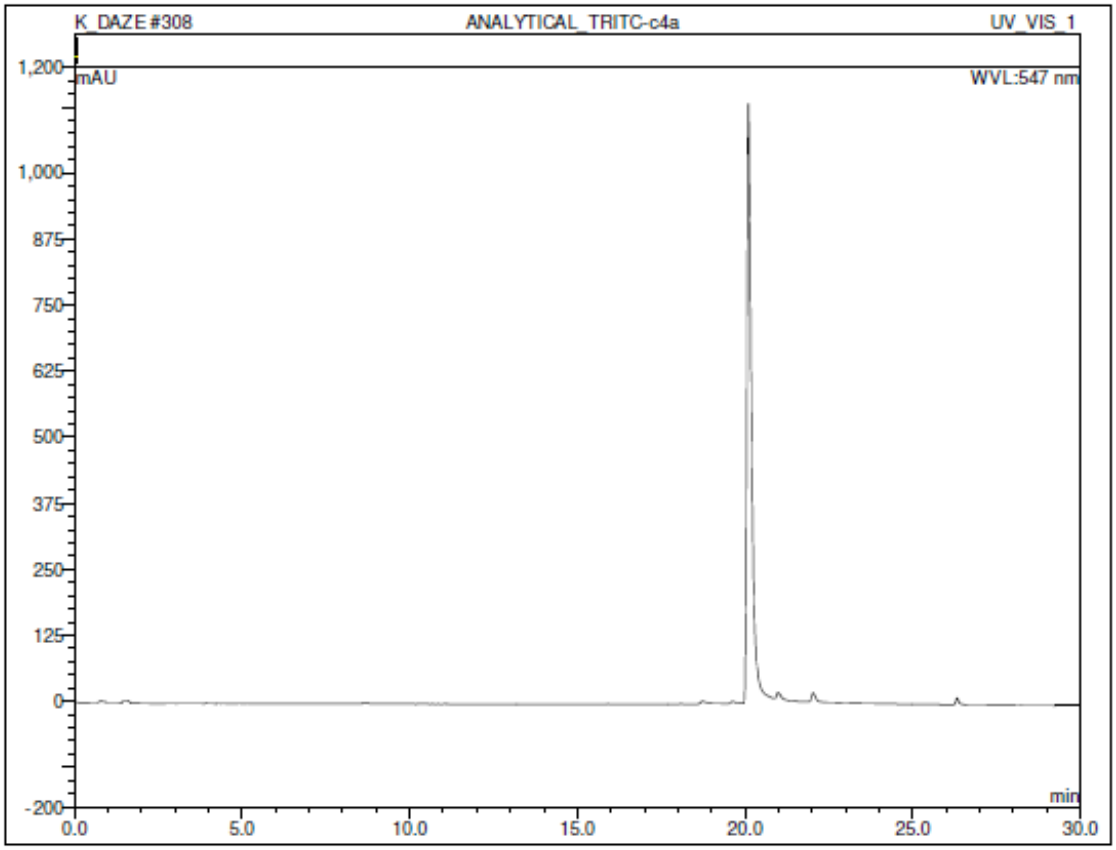
144. Ng, S.S. et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87-91 (2007).

145. Strahl, B.D. et al. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* **22**, 1298-1306 (2002).

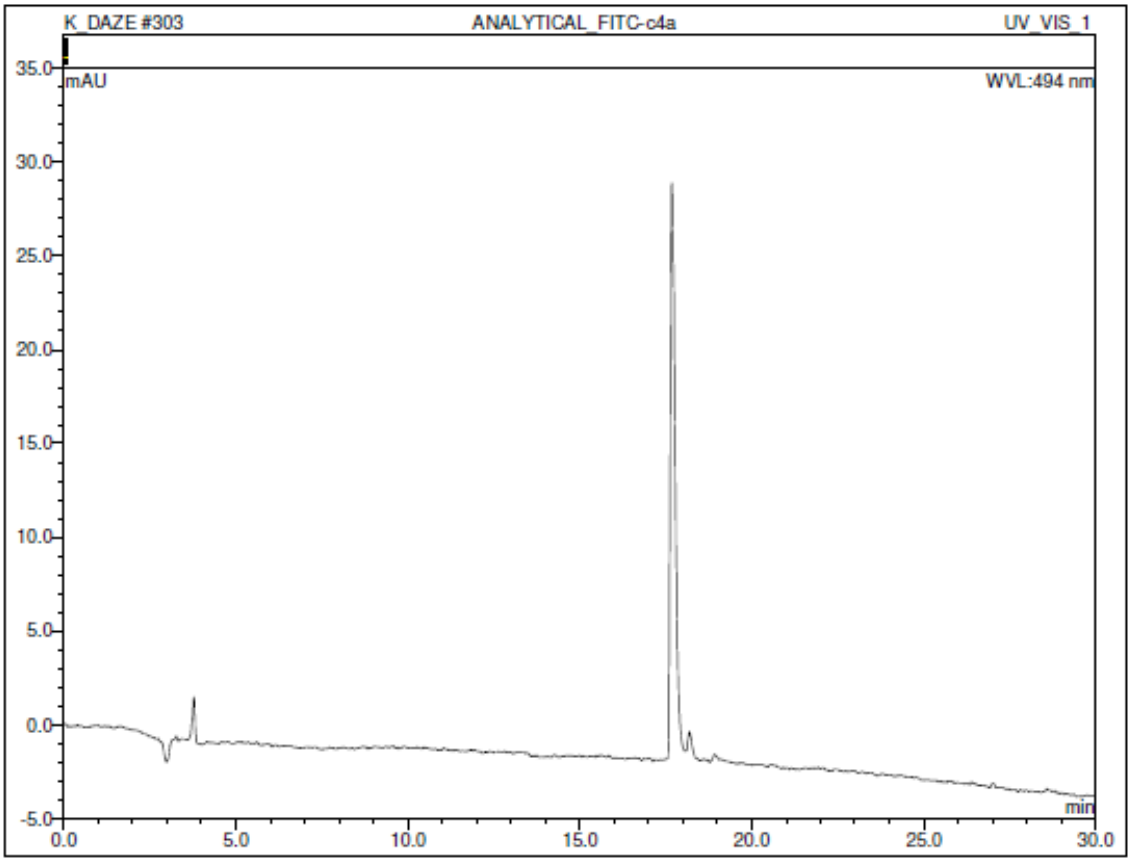
146. Wigle, T.J. et al. Accessing protein methyltransferase and demethylase enzymology using microfluidic capillary electrophoresis. *Chem Biol* **17**, 695-704 (2010).

Appendices

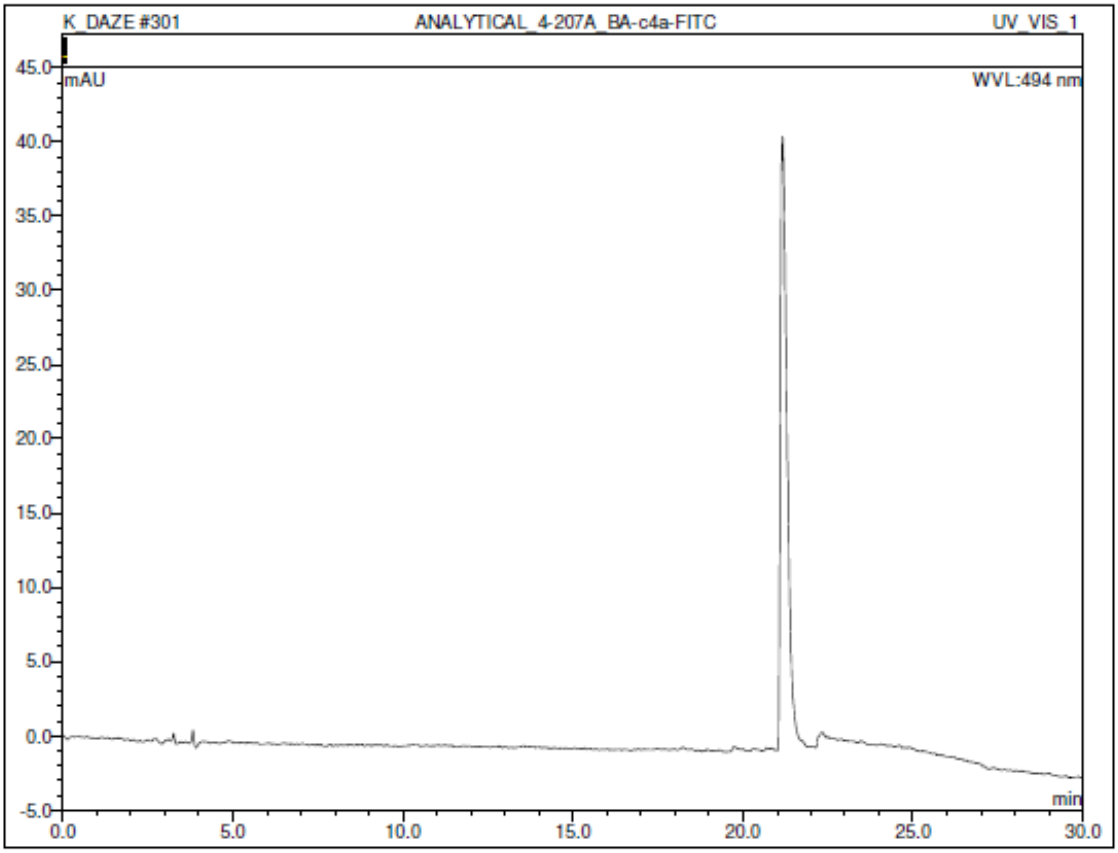
Appendix A- LC/MS traces for compounds 11-14



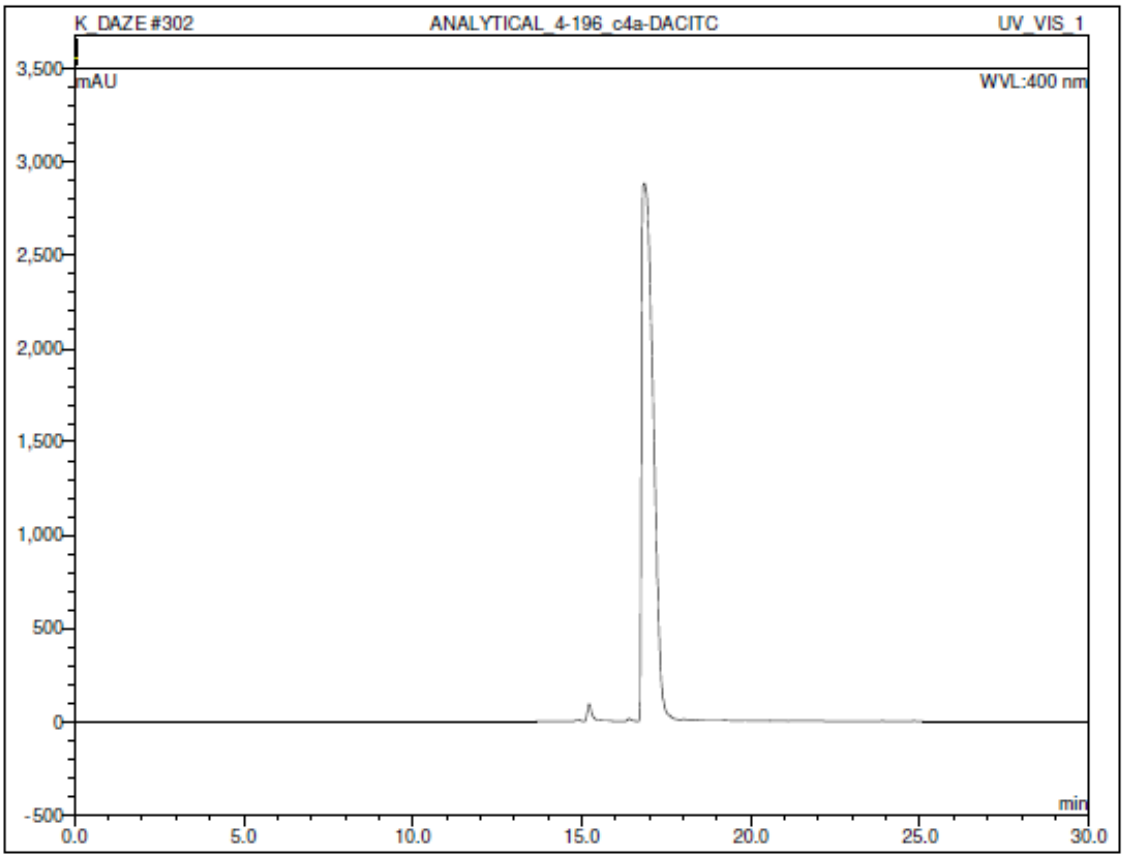
LC/MS Trace 1. Tetramethylrhodamine Isothiocyanate (TRITC)-derived host 11 on a preparative Luna C-18 column (Phenomenex, 5 µm, 21.2x250 mm), at a flow rate of 10 mL/min. Compound was purified by running a gradient from 90:10 0.1% TFA in H_2O:0.1% TFA in MeCN to 10:90 0.1% TFA in H_2O:0.1% TFA in MeCN over 30 minutes.



LC/MS Trace 2. 5(6)-Fluorescein Isothiocyanate (FITC) - derived host 12 on a preparative Luna C-18 column (Phenomenex, 5 µm, 21.2x250 mm), at a flow rate of 10 mL/min. Compound was purified by running a gradient from 90:10 0.1% TFA in H_2O:0.1% TFA in MeCN to 10:90 0.1% TFA in H_2O:0.1% TFA in MeCN over 30 minutes.

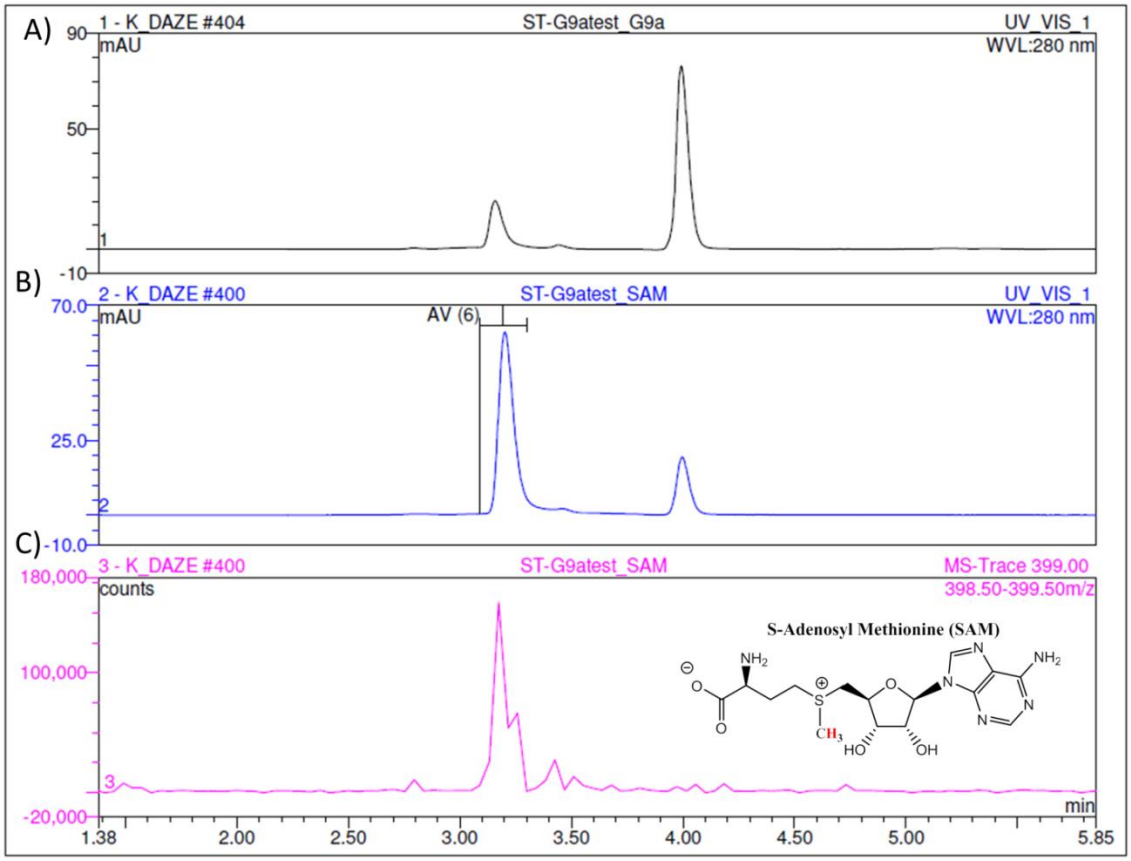


LC/MS Trace 3. 5(6)-Fluorescein Isothiocyanate Benzylamine (FITC-Ba) - derived host 13 on a preparative Luna C-18 column (Phenomenex, 5 µm, 21.2x250 mm), at a flow rate of 10 mL/min. Compound was purified by running a gradient from 90:10 0.1% TFA in H_2O:0.1% TFA in MeCN to 10:90 0.1% TFA in H_2O:0.1% TFA in MeCN over 30 minutes.

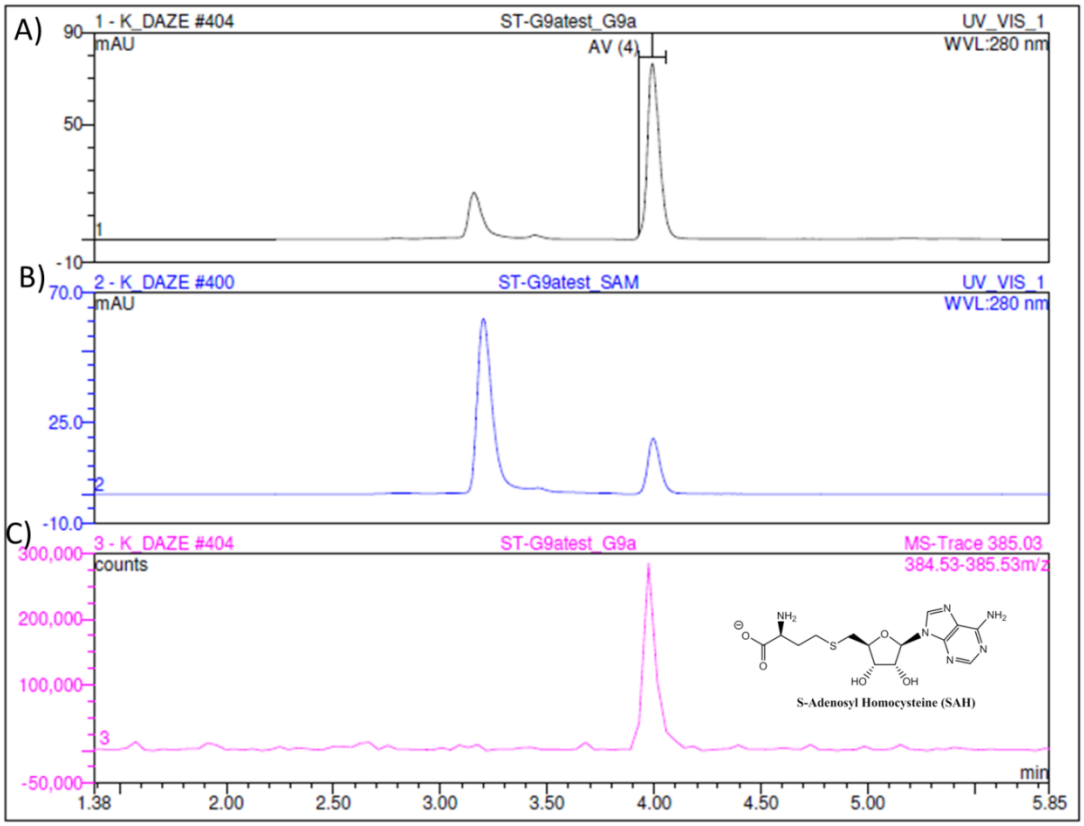


LC/MS Trace 4. 7-Dimethylamino-4-methylcoumarin-3-isothiocyanate *(*DACITC) - derived host 14 on a preparative Luna C-18 column (Phenomenex, 5 µm, 21.2x250 mm), at a flow rate of 10 mL/min. Compound was purified by running a gradient from 90:10 0.1% TFA in H_2O:0.1% TFA in MeCN to 10:90 0.1% TFA in H_2O:0.1% TFA in MeCN over 30 minutes.

Appendix B- LC/MS traces for G9a enzymatic activity monitoring

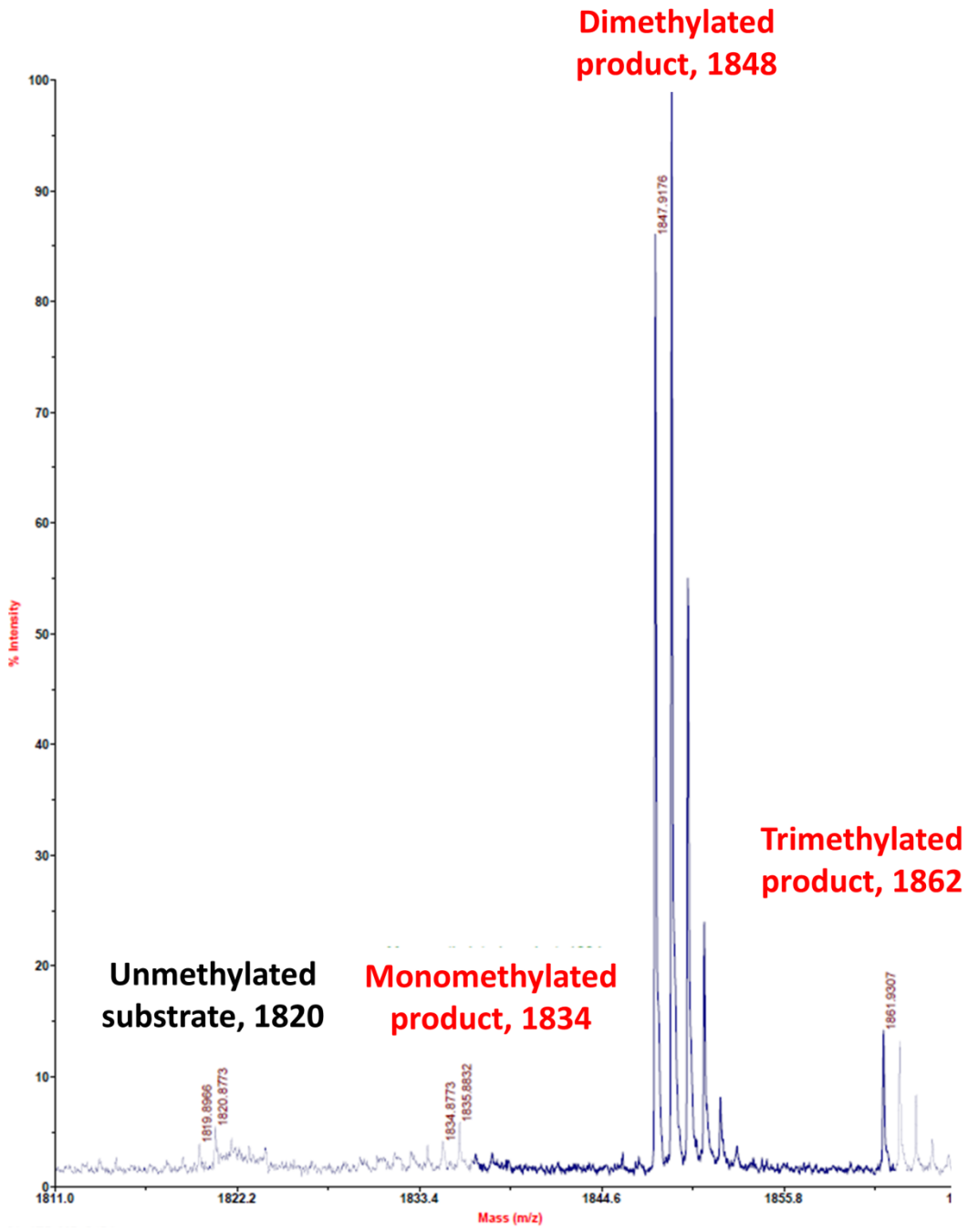


LC/MS Trace 5. A) LC/MS trace of sample containing buffer D, host 11, peptide, SAM cofactor and G9a methyltransferase. Sample injected after 90 minute kinetic run at 30°C. B) Control well containing buffer D, host 11 and SAM-cofactor also incubated for 90 minutes at 30°C. C) Mass of the biggest peak in B), indicating the mass of SAM.



LC/MS Trace 6. A) LC/MS trace of sample containing buffer D, host 11, peptide, SAM cofactor and G9a methyltransferase. Sample injected after 90 minute kinetic run at 30°C. B) Control well containing buffer D, host 11 and SAM-cofactor also incubated for 90 minutes at 30°C. C) Mass of the biggest peak in A), indicating the mass of SAH.

Appendix C- MALDI trace for G9a enzymatic activity monitoring



MALDI Trace 7. Sample containing buffer D, host 11, peptide, SAM cofactor and G9a methyltransferase incubated at room temperature over 60 minutes. Sample was filtered through a zip-tip and spotted on a MALDI-plate. Spectrum was zoomed in to region of interest.