Small Molecule Inhibitors and Substrate Analyses of Protein Arginine Methyltransferases

Amy Varney

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Trinity Term 2015

EPSRC and MRC Systems Approaches to Biomedical Science CDT

For Grandma, Irene, who always thought I was going to be a Doctor

For Mum, Chrissie, who taught me everything worth knowing

Abstract

Chapter 1 introduces the Protein Arginine Methyltransferases (PRMTs) as epigenetic regulators that decorate peptidic arginine with methyl groups. Evidence for PRMT involvement in cancer pathogenesis is reviewed and their plausibility as therapeutic targets is introduced. The methylation patterns conferred by the PRMTs is described, and the existence of novel patterns is considered. Techniques for assaying PRMT activity are compared and contrasted and a discussion of current PRMT inhibitors is presented. The chapter concludes by outlining the aims of the thesis.

Chapter 2 describes synthesis towards novel methylated arginine molecules that are fully protected for inclusion in peptides *via* solid phase peptide synthesis.

Chapter 3 outlines the total synthesis of protected δ-monomethylated arginine for peptide synthesis. This methylation pattern is known in yeast but has not yet been identified in humans.

Chapter 4 details a new mass spectrometry-based assay that can be applied for inhibitor and substrate analyses. Synthesis of a novel histone peptide containing the δ-monomethylated arginine, produced in Chapter 3, is also described and this is tested for relevance as a human epigenetic marker. Novel polymethylation patterns are also explored in a total of five histone peptides. This chapter concludes with discussion of possible methylation pattern rearrangements.

Chapter 5 describes the synthesis and testing of two series of putative PRMT inhibitors based on previously identified scaffolds within this research group. Data obtained from three different assays, including that outlined in Chapter 4, are analysed and suggestions as to the direction of PRMT assay design are offered.

Chapter 6 provides the experimental data to support Chapters 2-5, including all organic synthesis procedures, protein & peptide syntheses and assay methodology.

Declaration of Authorship

This thesis, and the work of which it is a record, were carried out by myself unless otherwise acknowledged.

Amy Varney

Oxford, Friday 9th October 2015

Acknowledgements

First and foremost, I must extend my warmest thanks to **Prof Angela Russell** for making this whole thing possible. For giving me freedom wherever possible and direction whenever necessary and for being an excellent mentor; it has been a pleasure to work with you. To my co-supervisors **Prof Chris Schofield** and **Dr Paul Brennan** – thank you both for your insight and perspectives and for helping guide this project with your respective expertise.

Dr Carole Bataille, I cannot thank you enough for your commitment to supervising me both in and out of the lab. Your diligence and dedication in proof reading this thesis were invaluable. Thank you for being an excellent mentor – without you this certainly wouldn't have been possible.

To one of the most unique people I have ever met - **Dr Gu**, you had me crying with laughter on more than one occasion! Thanks for making my time in the lab generally much sunnier – it has not been the same since you left.

Thank you **Nicky** for being my daily 7.30 am buddy and for making the long long days in the lab bearable. I very much enjoyed dancing the Macarena with you when it all got too much and enjoyed our spontaneous and outrageous conversations that made the Part II's a bit scared. Thank you for your proof-reading and occasional sanity-/intelligence testing me.

Jeggles. I can't remember if this is the second or third (or FOURTH?!!) time I've acknowledged you… hopefully you have a closer record than I do, but you know what they say… third time lucky. Good luck for your new career! It's been one heck of a journey and I am privileged that we got to make it together.

My little forest/lab-elf **Aini**. You've been my rock these last few years and given so many hilarious memories that I will treasure. I'd probably still be stranded in Dublin's Damnation without you. Or at least running around the airport. I'm sorry I always told you to fill up the dry ice. Thanks for putting up with me.

Caitlin, I'm very grateful that I've had someone who can send me 'interdisciplinary science' memes with absolutely no irony whatsoever, so thanks for that and for always been relied upon for a good laugh and a strong G&T.

A special mention must go to **Caitlin**, **Aini** and **Nicky** for being absolute heroes leading up to submission date. Thank you for all your help, encouragement and for being such wonderful friends.

Graham, in reality I have much to thank you for in terms of advice and for always being relied upon for an interesting tea-time debate. But mostly I am grateful for the pink wig you bought me in my third year – possibly the worst kept secret in the history of AJR Secret Santa.

Adam, thanks for introducing me to cheese and Nutella paninis during late-night Tuesdays in the lab. Thank you also for always allowing me to pick your brains and for being a reliable supplier of educational material, oh, and soil sieves.

Noelia and **Fernando**, you were an absolute pleasure to work with – your Buenos Nachos attitude and happy approach to life was ever inspirational and brightened up the days.

There are many people who cannot go without mention, in particular all current and past members of the AJR group for making G7 and G8 the unique and wonderful place that it is. **Beth** and **Lucia**, thank you for always making me smile and for some excellent banter – I'm so glad we had you in the group. Thank you also to the **AJRgrp whatsapp thread** for often keeping me entertained.

I am grateful to all at the DTC for making the first year of this DPhil such an inspiring, enlightening and educational experience while making it extremely fun along the way. In particular to my director **Prof Charlotte Deane** for your support both pastoral and academic, to **Prof Elspeth Garman** for some excellent advice in my second year and to **Prof David Gavaghan** for keeping the DTC ship shape. Also, all members of the admin and academic teams cannot go without mention for their absolute dedication to this fantastic programme and last but certainly not least, to my wonderfully intelligent friends and peers for being such an eclectic and fabulous bunch of people to work with. In particular, thanks to **James Mountain** and **Jonathan Ross** – I'm so glad we met over that fateful handshake.

It just so happens that **Dr Louise Walport** had the misfortune to help me out during two of the most stressful periods of my project. I cannot thank you enough for your help and wisdom and for directing me around Chris' many labs. To **Dr Richard Hopkinson** for being equally as accommodating and to all other members of the CJS group who always astounded me with their kindness and helpfulness, in particular **Gareth Langley** and **Becky Hancock**.

While I ran around Oxford like a headless chicken in desperate search of a spare incubator, it was my first ever research group who came to my rescue. Thank you **Dr Shilpa Bali** for your assistance, which

honestly felt like you were saving my life at the time, and to **Dr Despoina Mavridou** for a lovely catch up.

Thank you to my collaborators at the SGC, Toronto for providing me with some thought-provoking data. Also, to those members of the CRL who helped make much of my work possible, notably **Dr Barbara Odell**, **Prof Tim Claridge** and **Dr Tina Jackson** for their NMR expertise and to the mass spec team **Prof James McCullagh** and **Colin Sparrow**. **Dr James Wickens** deserves a special mention for his extensive assistance and endless patience for all things LC/MS-related.

Everything was made possible by my first Oxford mentor and the person who made the dream reality. **The Revd Prof Rob Gilbert**, thank you for giving me a place at **Magdalen** and for supporting and encouraging me throughout. Along with **Nicola Laurieri**, you were the greatest inspirer of my undergraduate studies.

I had the privilege to meet an amazing bunch of tremendously strong women as part of one of the hardest tests of my life. To all my friends at the Oxford University Women's Boat Club, thank you for everything. In particular, to my two boat race crews: Osiris 2012 and the Blue Boat 2013, to **Natasha Townsend** for being a truly inspirational role model, **Christine Wilson** for teaching me valuable lessons and to **Harriet Keane** for being such a reliable source of fantastic life advice and friendly support.

It remains for me to thank all of my family and friends for their continued support and for keeping me happy. There are no strong enough words to express my gratitude to my **Mum**, who amongst doing important things like birthing me and making this thesis possible, also acted as 'English Language Advisor' providing me with the elusive words on the tip of my tongue when all sense of my assumed intelligence had otherwise failed me.

Finally, to my wonderful **Alice**, thank you for being the loveliest person, for making me so happy and taking good care of me during write-up. Sorry you had to put up with terrible chemistry chat for so long and deal with my tantrums when 'The Magnet' broke.

Table of Contents

Abbreviations

AMI	Arginine Methyltransferase Inhibitor (series)
2-TY	Tryptone-Yeast
ABTS	2,2′-Azino-bis(3-ethylbenzothiazoline)-6-sulfonate
ACEC	α-Chloroethyl Chloroformate
ADMA	Asymmetrically Dimethylated Arginine
APCI	Atmospheric Pressure Chemical Ionisation
Arg, R	Arginine
BBN	Borabicyclo(3.3.1)nonane
Boc	*tert*-Butyloxycarbonyl
br	Broad Singlet
CARM1	Coactivator Associated Arginine Methyltransferase 1
CDK	Cyclin Dependent Kinase
CHCA	α-Cyano-4-hydroxycinnamic acid
ChIP	Chromatin Immunoprecipitation
cLogP	Calculated Partition Coefficient
CM	Cyclised Material
d	Doublet
DIPEA	Diisopropylethylamine
DLS	Dynamic Light Scattering
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNs	2,4-Dinitrobenzenesulfonamides
DPSO	Diphenylsulfoxide
EDTA	Ethylenediaminetetraacetic Acid
EGFR	Epidermal Growth Factor Receptor
EMT	Epithelial-Mesenchemal Transition
ER+	Estrogen Receptor Positive
ESI	Electrospray ionisation
EWG	Electron Withdrawing Group
Fmoc	Fluorenylmethyloxycarbonyl
GAR	Glycine, Arginine-Rich
GC/MS	Gas Chromatography Mass Spectrometry
Glu, E	Glutamic Acid
H4	Histone 4
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HMBC	Heteronuclear Multiple-Bond Correlation
HMT	Histone Methyltransferase

HRP	Horseradish Peroxidase
HTS	High Throughput Screen
IC_{50}	Half Maximal Inhibitory Concentration
IDMS	Isotope Dilution Mass Spectrometry
IHC	Immunohistochemistry
IPTG	β-D-Thiogalactopyranoside
KLH	Keyhole Limpet Hemocyanine
LC/MS	Liquid Chromatography Mass Spectrometry
Lys, K	Lysine
MALDI-TOF-MS	Matrix Assisted Laser Desorption Ionisation Time of Flight Mass Spectrometry
MCL	Mantle Cell Lymphoma
***m*CPBA**	*meta*-Chloroperoxybenzoic Acid
MITF	Microphthalmia-Associated Transcription Factor
MW	Molecular Weight
MLL	Mixed Lineage Leukaemia
MMA(δ)	δ-Monomethylated Arginine
MMA(ω)	ω-Monomethylated Arginine
MML	Monomethyl-Lys
MMO(δ)	δ-Monomethyl Orn
MS	Mass Spectrometry
MT	Methyltransferase
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MW	Microwave
NCBI	National Centre for Biotechnology Information
NCI	National Cancer Institute
NHL	Non-Hodgkin Lymphoma
NMR	Nuclear Magnetic Resonance
NOS	Nitric Oxide Synthase
Ns	Nitrobenzenesulfonyl
NSCLC	Non-Small Cell Lung Cancer
ORFs	Open Reading Frames
Orn	Ornithine
PABP1	Poly(A)-binding protein 1
PAD	Peptidylarginine Deiminase
PAINS	Pan-Assay Interference Compounds
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PGM	Proline, Glycine, Methionine-rich
pI	Isoelectric Point
PKMT	Protein Lysine Methyltransferase
PRMT	Protein Arginine Methyltransferase
RBL2	Retinoblastoma-Like 2
ROX	Ribosomal Oxygenase

s	Singlet
SAH	*S*-Adeonsyl-L-Homocysteine
SAHH	SAH hydroxylase
SAM	*S*-Adeonsyl-L-Methionine
SAR	Structure-Activity Relationship
SDM	Site-Directed Mutagenesis
SDMA	Symmetrically Dimethylated Arginine
SGC	Structural Genomics Consortium
siRNA	Small Interfering RNA
SM	Starting Material
SPM	Side Product Material
SPPS	Solid Phase Peptide Synthesis
t	Triplet
TB	Terrific Broth
TBAI	Tetrabutylammonium Iodide
TFA	Trifluroacetic Acid
TIM	Triosephosphateisomerase
TLC	Thin Layer Chromatography
TMO	Trimethylorthoformate
TRF	Time-Resolved Fluorescence
Tris	Trishydroxymethyl(aminomethane)
TSG	Tumour Suppressor Gene
UPLC	Ultra-High Performance Liquid Chromatography

Chapter 1: Introduction

Manipulation of the epigenome for therapy is a research-intensive field with particular hope for success in the fields of cancer, inflammatory and degenerative diseases[1,2]. As more of the complex biology involved in epigenetic regulation is uncovered, our ability to identify suitable targets for therapeutic intervention increases[3]. This thesis explores the possible existence of novel methylation patterns of the amino acid arginine **1** (Arg, R) under the working hypothesis that they could be epigenetic markers. Furthermore, this work describes the search for small molecule inhibitors of the Protein Arginine Methyltransferase (PRMT) family of enzymes as part of a cancer drug discovery project.

1.1 – Epigenetic Regulation

The prefix 'epi-' means 'above' or 'upon'; epigenetics is the layer of control above the genome. The concept of an 'epigenetic landscape' was first proposed by C.H. Waddington in 1957[4] although it wasn't until the explosion of data heralding the 21st century that the field really took off. A brief background to epigenetics is given in this section; covalent modifications, called epigenetic markers, are added to a huge variety of cellular substrates by epigenetic machinery. These covalent modifications bring about phenotypic changes in the cell.

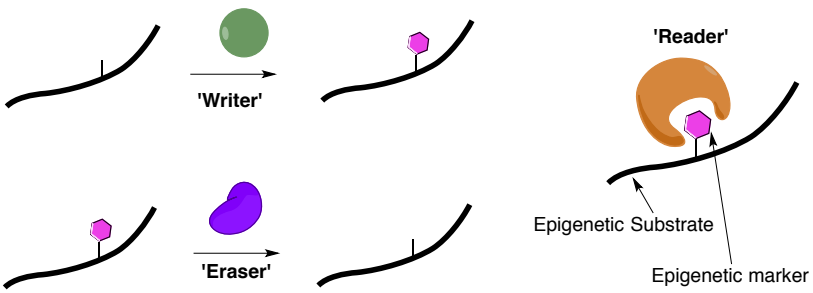


***Figure 1.1** – Epigenetic machinery can be thought of as 'writers', 'readers' or 'erasers'.*

1.1.1 –Epigenetic Machinery and Control

Epigenetic machinery can be broadly summarised into the categories of 'writer', 'reader' or 'eraser' (**Figure 1.1**). The 'writers' are enzymes that append a chemical group to the target, the 'readers' recognise and bind to markers in the context of the local environment and can be thought of as decipherers, which bring about phenotypic changes. The 'erasers' remove epigenetic markers. The exact combination and pattern of epigenetic markers on different substrates is thought to be integral to the cell phenotype and while some markers are transient, some are more permanent and are heritable. A variety of markers, 'readers', 'writers' and 'erasers' exist and the substrate scope upon which they act is substantial.

1.1.2 – Classical Epigenetics: Histones, DNA and Transcription

DNA is stored in cells wrapped around octamers of histone proteins that are stacked into chromosomes[5]. This octamer, called the nucleosome, incorporates two copies each of histone proteins H2A, H2B, H3 and H4 all of which are highly conserved[6]. Nucleosomes pack together *via* linker H1 regions into chromatin, which makes up the final unit of storage – the chromosome (**Figure 1.2**). Tightly bound regions of DNA (heterochromatin) are inaccessible to transcriptional machinery, whereas loosely bound regions of DNA (euchromatin) can be accessed and transcribed to mRNA.

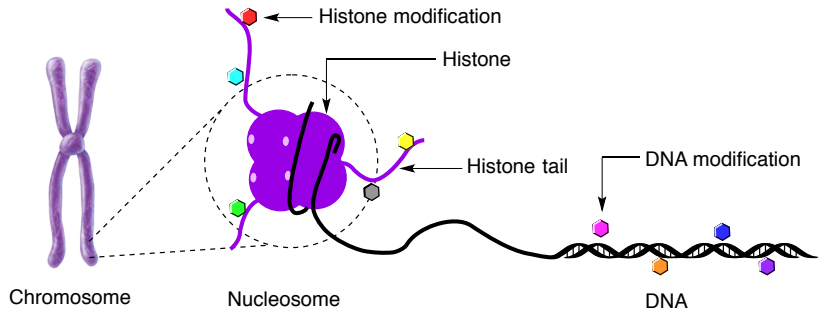


***Figure 1.2** – Organisation of DNA into chromosomes, showing an expanded view of the nucleosome and representation of epigenetic modifications by coloured hexagons, defined in **Figure 1.3**.*

Figure 1.3 exemplifies different epigenetic modifications. Histone proteins are major epigenetic substrates, with the possibility for modification by different 'writers' at various amino acid residues in their 'tail' regions. DNA is also a major epigenetic substrate with modifications occurring on the base cytosine. The nature of the chemical modifications influences whether chromatin exists as hetero- or euchromatin, by altering local charge and steric properties of both macromolecules. The associated structural changes of chromatin affect signalling *via* epigenetic 'readers', which goes some way to control the expression or silencing of individual genes.

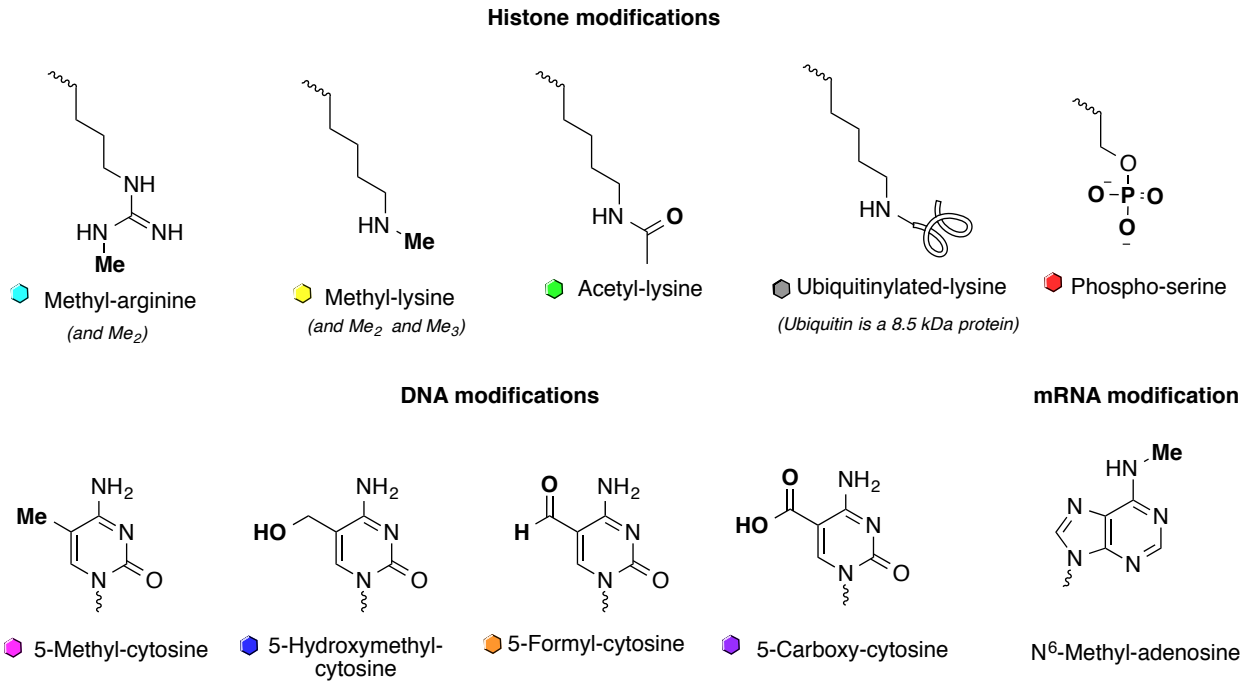


***Figure 1.3** – Epigenetic modifications of amino acids, the DNA base cytosine and the RNA base adenosine.*

1.1.3 – Emerging Epigenetic Concepts: RNA, the Ribosome, Translation and Trafficking

DNA and histone modifications are well studied, but the field of 'RNA epigenetics' is also coming into its own. A plethora of post-transcriptional RNA modifications have been identified although their purposes are less well defined[7]. One recent example is the epigenetic modification of mRNA; the mRNA marker N^6-methyladenosine (**Figure 1.3**) is shown to be important for stem cell differentiation in mouse models by altering mRNA translation[8]. RNA epigenetics also

involves modification of tRNA and rRNA. Modifications of the former are becoming better defined with roles emerging for specific tRNA chemical markers[9]. rRNA modifications can lead to alterations in ribosomal structure, which are postulated to affect the efficiency with which certain proteins are synthesised compared to others[10].

It has also been suggested that epigenetic machinery has a role to play in protein trafficking in the cell. The extent of acetylation on the microtubule subunit α-tubulin is partially under the control of histone deacetylase (HDAC) 6, an epigenetic 'eraser'. It is reported that transport of the Epidermal Growth Factor Receptor (EGFR) along microtubules is facilitated when HDAC6 is inhibited leading to α-tubulin acetylation, suggesting that epigenetic machinery has some control over protein trafficking[11].

1.1.4 – The Epigenome and Disease

Deciphering the 'epigenetic code' in order to understand the 'which?', 'when?', 'where?' and 'why?' of epigenetic modifications is an ongoing effort. Huge progress has been made in the last five years and in 2015 the first epigenomic database – 'The Epigenome Roadmap' – was made available online[12] which contained analyses of 111 reference human epigenomes[13].

Epigenetic dysregulation is prevalent in many diseases, notably those where gene expression or DNA replication/repair is aberrant such as cancer and inflammatory disease (reviewed by Chi)[14]. Drugging the epigenome is a continued aim for medicinal chemists fuelled by success stories such as the development of HDAC inhibitors and of bromodomain inhibitors[15] targeting epigenetic acetyl-Lys 'erasers' and 'readers' respectively. The first licenced example of a HDAC inhibitor is Vorinostat® (Merck) shown to have tangible clinical outcomes in treating Cutaneous T-Cell Lymphoma[16]. Undoubtedly HDAC inhibitors have good therapeutic value, but patients

receiving this treatment often suffer from side effects attributed to off-target interactions; the root cause is widely thought to be poor selectivity of the drugs for individual HDAC homologues[17]. Aided by the Epigenome Roadmap, epigenetic drug discovery continues to be a cutting-edge research field.

1.1.4.a –Methyltransferases

This work is concerned with the epigenetic 'writer' family of enzymes called the methyltransferases (MTs). Members of the overarching MT family have been divided into three classes (**Figure 1.4**), defined by mining genomic open reading frames (ORFs) using bioinformatic tools[18]. All family members transfer methyl groups from the universal cofactor *S*-adenosyl-L-methionine (SAM) **2** (**Figure 1.5**) to target substrates.

Class I MTs encompass the Protein Arginine Methyltransferases (PRMTs) and Class II the SET-domain-containing Protein Lysine Methyltransferases (PKMTs). These two classes catalyse methylation of Arg or lysine (Lys, K) residues respectively. Class III MTs are membrane-bound and are not further discussed herein. This work primarily focuses on the role of the Class I MTs, the PRMTs, in epigenetic regulation.

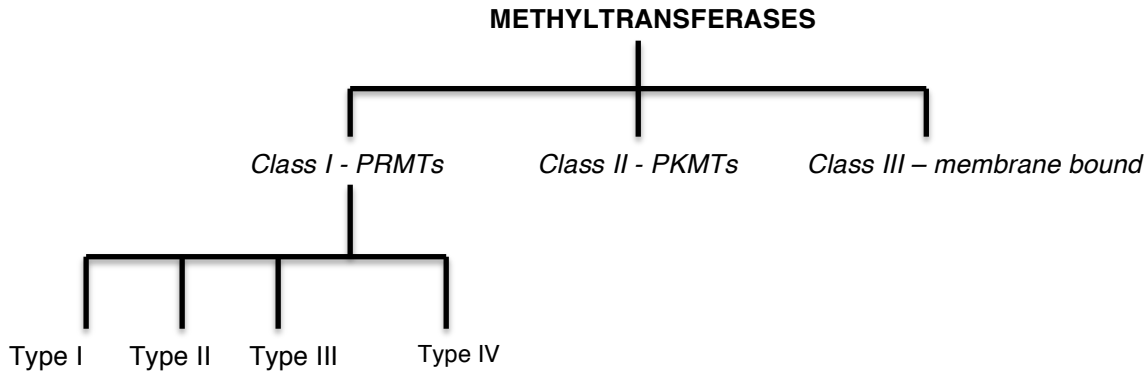


***Figure 1.4** – Classification of methyltransferases.*

1.2 – The Protein Arginine Methyltransferases

The PRMTs catalyse *N*-methylation of the guanidine side chain of Arg residues within various cellular proteins (**Figure 1.5**). This was first realised in the 1960's when radioactivity studies showed calf thymus nuclei capable of incorporating methyl groups from ^{14}C-labelled SAM **2** into non-Lys amino acids[19]. Earlier work[20], on what later became known as the PKMTs, meant that standards of methylated Lys homologues were available for chromatographic distinction from these newly identified[19] methylated amino acids. One of these new species was later defined[21-23] as ω-*N*-monomethylated Arg **3** (MMA(ω)) and subsequent work on human urine samples[24] suggested ω-*N, N*-(asymmetrically) dimethylated Arg **4** (ADMA) and ω-*N,N'*-(symmetrically) dimethylated Arg **5** (SDMA) accounted for the other species.

To date nine different putative human PRMTs have been discovered, split into subtypes I-IV[25]. PRMTs are named 1-9 but PRMT4 is more commonly referred to as CARM1 (Coactivator Associated Arginine Methyltransferase 1).

1.2.1 –Arginine Methylation

The guanidine side chain of Arg contains three nitrogen atoms, each of which is chemically amenable to methylation (**Figure 1.5**). Therefore in principle multiple methylation patterns are possible, of which four have so far been observed (**Figure 1.5**). The Lys side chain has one nitrogen atom and all of the mono-, di- and tri-methylated species have been defined[26].

The majority of PRMTs methylate Arg almost exclusively within a 'Glycine, Arginine-Rich' (GAR) sequence[27,28]. The exception is CARM1, which preferentially methylates at a 'Proline,

Glycine, Methionine-rich' (PGM) site[29], and PRMT5 which can methylate at both GAR and PGM sites[29,30].

1.2.1.a – Recognised Methylation Patterns

Different patterns of methylation are recognised by different 'readers' and have specific downstream effects in the cell. There are four types of reported methylation patterns on Arg, three of which have so far been shown to have physiological relevance in humans (**Figure 1.5**). Type I-III PRMTs generate the MMA(ω) pattern **3**. Type I PRMTs further methylate to the ADMA pattern **4** and type II to SDMA **5**, while type III is thought to catalyse no further reaction. A type IV PRMT has been discovered[31] in *Saccharomyces cerevisiae* where it catalyses formation of δ-*N*-monomethylarginine (MMA(δ)) **6**. No evidence for a human homologue has been reported to date.

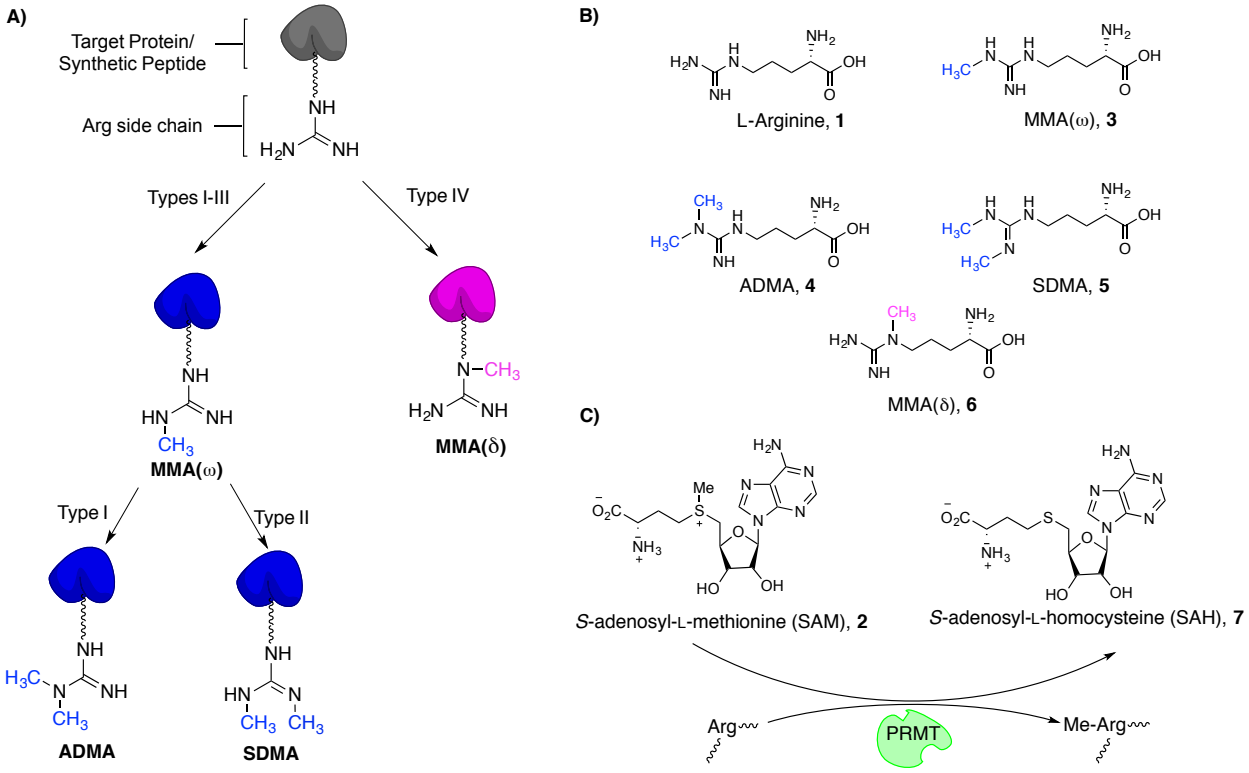


***Figure 1.5** – **A)** Methylation patterns conferred by types I-IV PRMTs **B)** Structure of known methylated Arg residues **C)** Schematic of PRMT-catalysed methylation of a protein-embedded Arg residue using SAM **2** as cofactor and producing breakdown product SAH **7**.*

1.2.1.b– Methylation Number and Regioisomerism

For the purpose of this thesis 'methylation number' is defined as the total number of methyl groups on any of the three nitrogen atoms in the guanidine group of Arg. Arginine residues with the same methylation number can be regioisomers of each other, for example: ADMA **4** and SDMA **5** both have a methylation number of two, but have different regiochemistry.

1.2.1.c – Methyl-Arginine Recognition

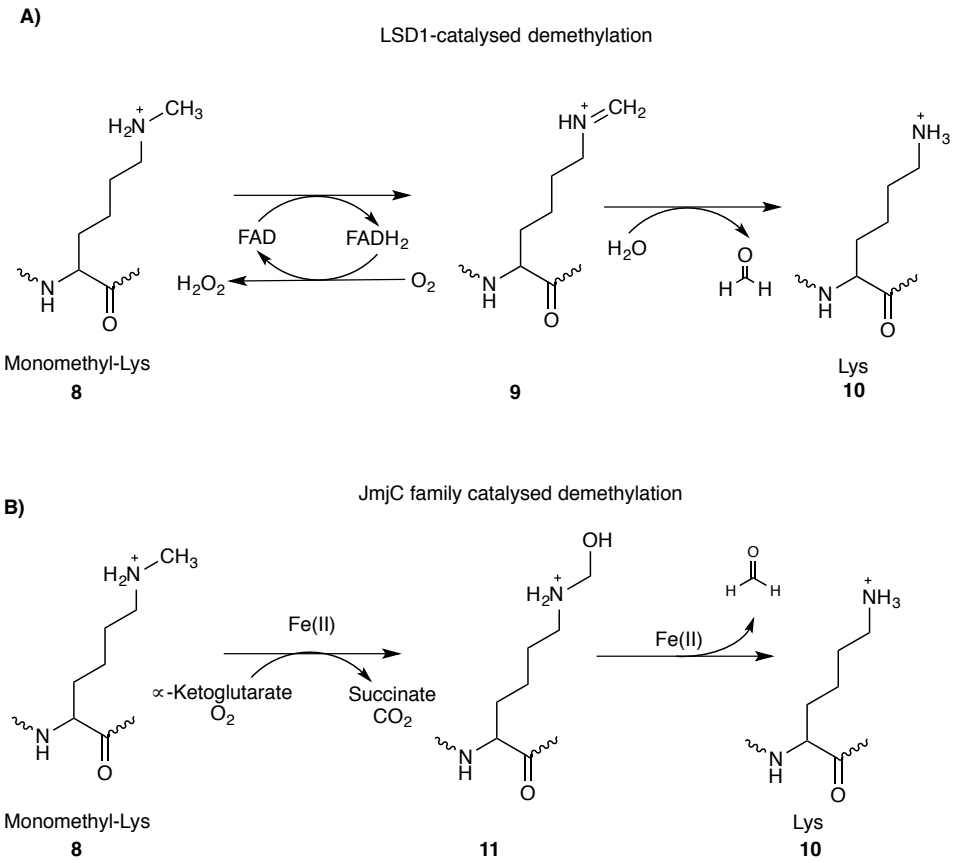
The major known 'readers' of methylated Arg are Tudor domain-, WD40 domain- and ADD domain- containing proteins[32]. The relevance of different marker-reader interactions is currently unclear but is an active area of research.

1.2.2 – Arginine Demethylation

Despite the early discovery of histone methyltransferases[19], it was decades later that any enzyme was discovered that could reverse the process[33-35]. In fact, it was initially thought that Arg/Lys methylation was either infrequently or never reversed based on work from the seventies showing low turnover of methyl-histones[36,37]; for a long time it was widely accepted that the high stability of the *N*-Me bond might act as a barrier to demethylation.

1.2.2.a – Oxygenation/Hydroxylation

The first true histone demethylase to be discovered was LSD1, which acts as an epigenetic 'eraser' on methyl-Lys (**Scheme 1.1A**)[38]. The second major class of Lys demethylases are the JmjC family of oxygenases, which act *via* a hydroxylation pathway **Scheme 1.1B**[39].

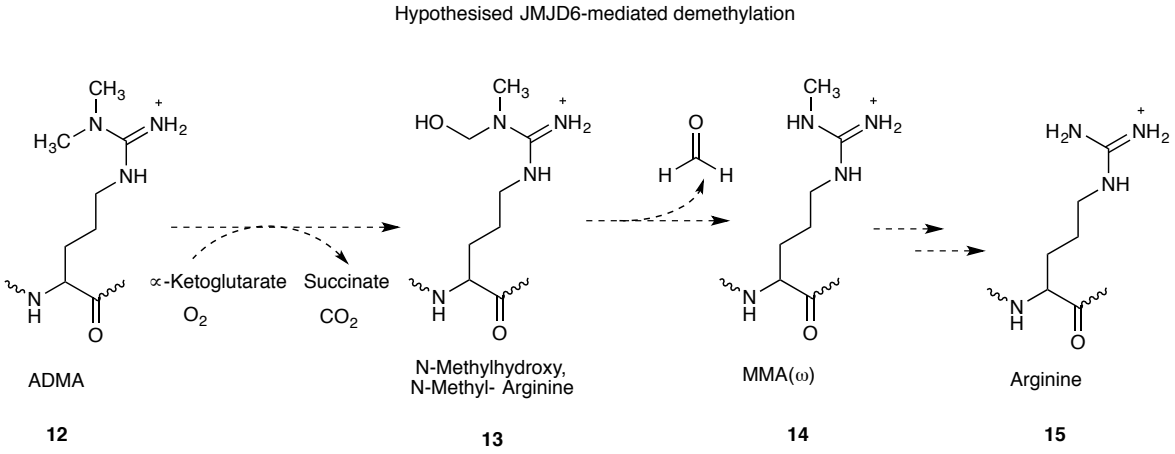


Scheme 1.1 – A) *LSD1-catalysed demethylation of monomethyl Lys* ***B)*** *JmjC family catalysed demethylation* via *hydroxylation.*

Demethylation of Arg however is much more controversial. In 2007 the oxygenase JMJD6 was reported to remove methyl groups at H3R2 and H4R3 *via* hydroxylation in an analogous manner to the Lys demethylases (**Scheme 1.2)**[40]. The authors used commercially available antibodies reportedly specific for mono- and di-methyl Arg to assess the extent of methylation on bulk histones *via* Western blotting[40]. However, the antibodies were generated against a synthetic peptidic immunogen (conjugated to the carrier protein keyhole limpet hemocyanine (KLH)) and in some cases the sequences are proprietary[a]. The discrepancy between immunogen (peptide) and experiment substrate (bulk histones) could undoubtedly lead to artefactual results. Indeed, in 2009 the first contradictory evidence for JMJD6's role as an Arg demethylase was published[41].

[a] Details of commercially available antibodies used are stated in the supplementary information for ref. 40. Details of immunogens can be found by cross-reference to the manufacturers' websites. Not every antibody was traceable.

This work, using mass spectrometry (MS) analyses and not relying upon antibody-based recognition, found that JMJD6 could not demethylate histone H3 and H4 fragments but that the enzyme was in fact a Lys hydroxylase. The search for a true Arg demethylase continues while de-convolution of JMJD6's role remains a point of discussion[42].



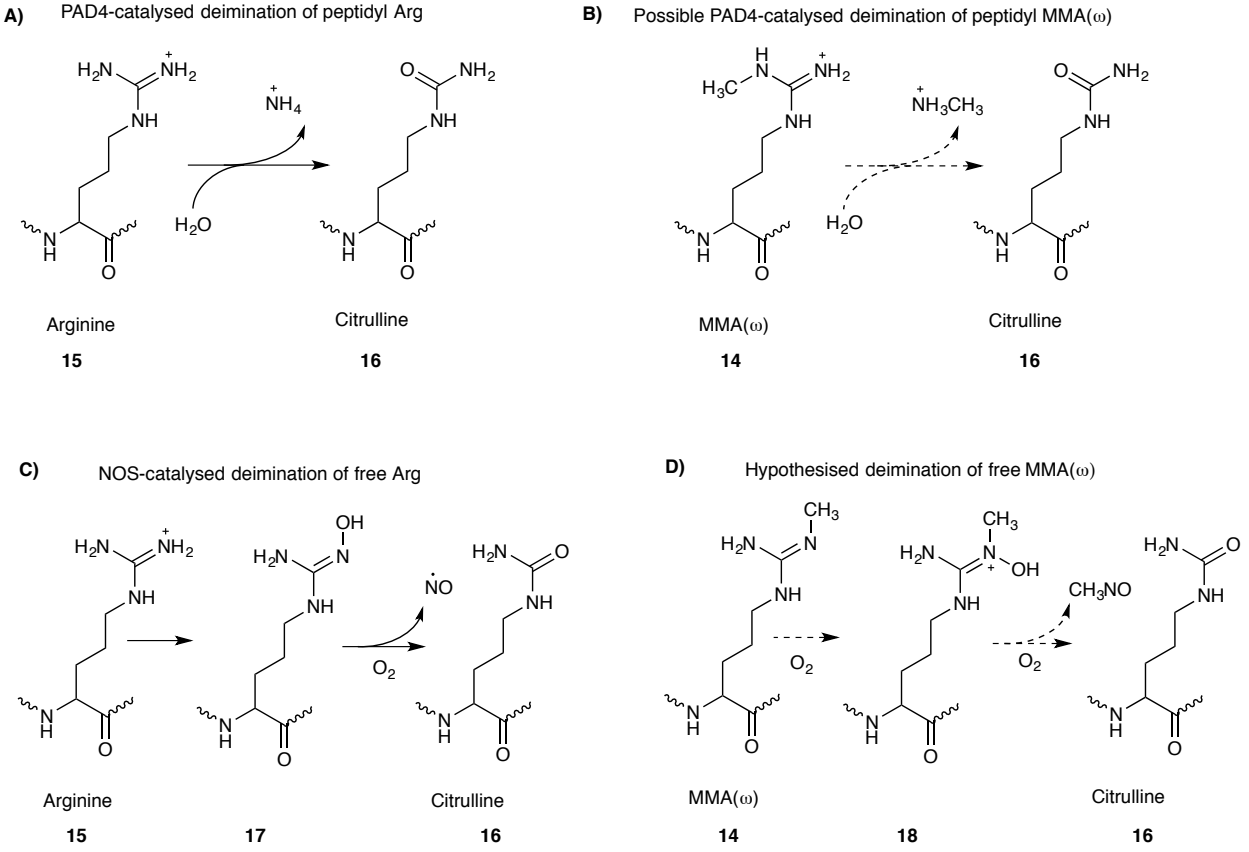
***Scheme 1.2** – Hypothesised mechanism of Arg demethylation by JMJD6 that has been opposed by subsequent characterisation of JMJD6 as a Lys hydroxylase.*

Arg hydroxylation has been described for ribosomal residues, catalysed by the ribosomal oxygenase (ROX) family of proteins[43]. However, so far structural evidence and substrate analyses point towards the ROX family as genuine Arg (and Lys) hydroxylases without a role as demethylases[44].

1.2.2.b - Deimination

The second proposed method of demethylation is *via* deimination although this does not restore unmethylated Arg, but instead produces citrulline. Peptidylarginine Deiminase 4 (PAD4) converts Arg to citrulline within proteins[45] (**Scheme 1.3A**) as evidenced by mutagenesis and crystallographic data[46]. There is much discussion as to whether PAD4 has any physiological relevance as a histone demethylase[47] acting *via* the mechanism outlined in **Scheme 1.3B**.

Certainly free mono- and di-methyl Arg are demethylated in this way by the dimethylargininse DDAH[48], but for peptidyl Arg opinion is moving towards the theory that (peptidyl-) citrulline itself is a post-translational marker mutually exclusive with methylation, and that the PAD family cannot demethylate histones *via* deimination[47]. Free Arg can also be converted to citrulline *via* nitric oxide synthase (NOS)[49] (**Scheme 1.3C**). By extension this principle might be applicable to free methyl-Arg (**Scheme 1.3D**), however it should be noted that this is speculation and that no such activity has been observed for peptidyl Arg to date.



Scheme 1.3 – A*) Mechanism of peptidyl Arg* ***15*** *deimination by PAD4* ***B)*** *Proposed mechanism for PAD4-catalysed deimination of peptidyl Me-Arg* ***15*** *that is now thought to have limited physiological relevance* ***C)*** *Known deimination of free Arg* ***15*** *to citrulline* ***16*** *catalysed by NOS* ***D)*** *Hypothesised deimination of free Me-Arg* ***14****.*

Despite the aforementioned controversies, it is generally accepted that arginine methylation/demethylation *in vivo* is a dynamic process[50]. This concept of methyl-marker

exchange sets the stage for a discussion of novel methylation patterns and their possible inter-conversion.

1.2.3– Novel Methylation Patterns

Since the guanidine side chain of Arg has three nitrogen atoms, there are nine theoretically possible permutations of methylation patterns. Four of these have been identified in cells as outlined in **Figure 1.5**, but there are no reports of investigations into the remaining five (**Figure 1.6**). This thesis hypothesises that Arg residues with a methylation number higher than two might exist and/or have physiological/biological relevance. This thesis also hypothesises that interconversion of regioisomers with the same methylation number could also be possible.

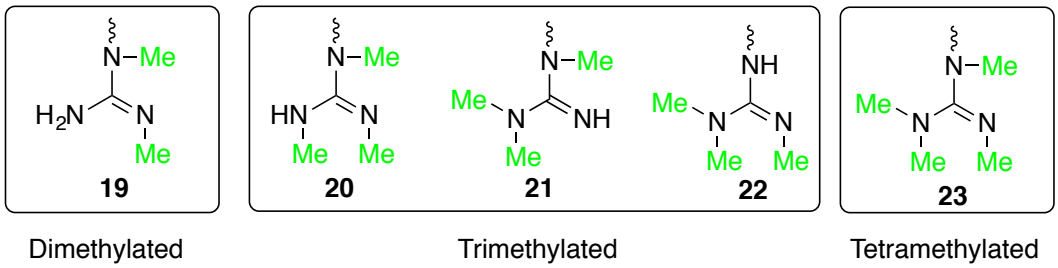


***Figure 1.6** – Novel methylation patterns.*

1.2.4 – Structure, Mechanism and Function of the PRMTs

The nine human PRMTs have two conserved domains – a Rossmann fold that binds SAM and a β-barrel that contributes to substrate binding (**Figure 1.7**). Individual PRMTs have smaller structural elements that influence their substrate specificity[51]. PRMT5 contains a large TIM-barrel that is important for recruitment of the cofactor MEP50 (also known as WD45) an essential component for activity[52]. Many of the PRMTs also harbour dimerisation arms within their β-barrel domains. In some PRMTs dimerisation is also assisted by a flexible conserved α-

helix; this motif has been hypothesised to have importance in cofactor binding by burying SAM in the β-barrel, in the process revealing the substrate binding site.

Crystal structures have assisted in determining the binding mode of substrates and inhibitors. Alongside site-directed mutagenesis (SDM) studies, they have also shed light on the mechanisms controlling the nature of methylation patterns (*vide infra*)[51].

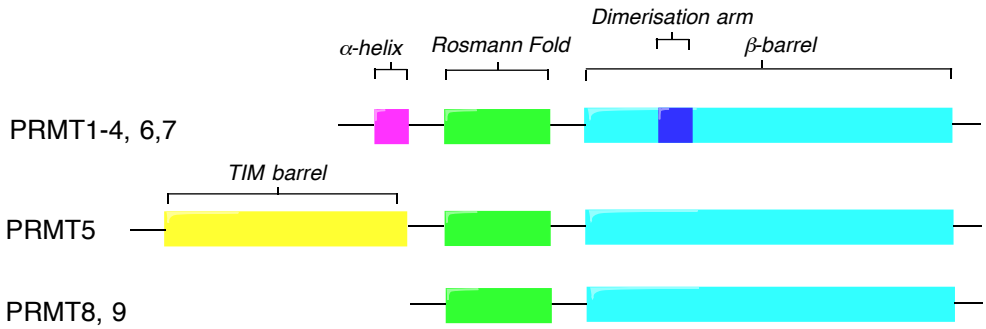
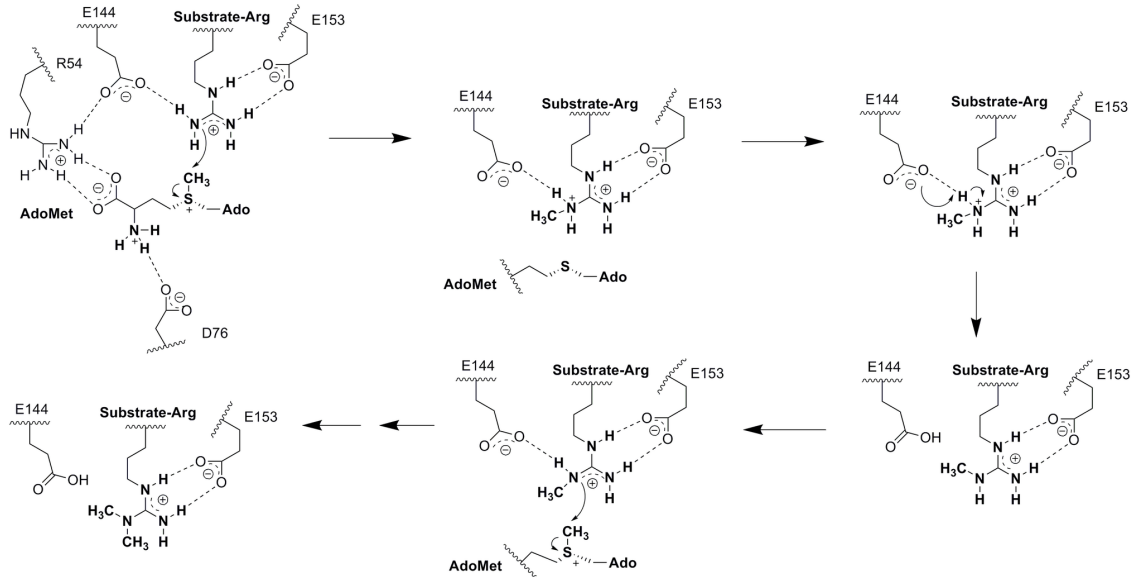


Figure 1.7 *– Core domains of the PRMTs, including conserved Rossmann fold and β-barrel domains. Adapted from Schapira* et al.[51].

1.2.4.a – Methylation Mechanism

All PRMTs have two conserved glutamic acid (Glu, E) residues – the 'double-E loop' – within the conserved 'Rosmann fold' domain that are essential to catalysis[53]. The mechanism for PRMT1-catalysed methylation is described below and depicted in **Scheme 1.4** as an archetype for the entire family.

The proposed mechanism relies on the conserved Glu residues to enhance nucleophilicity of the substrate guanidine and enable it to make an S_N2 attack on SAM[54]. In **Scheme 1.4**, the two Glu residues in the 'double-E loop' are shown localising and stabilising the positive charge developing on nascent methyl-Arg through hydrogen bonding[54,55].

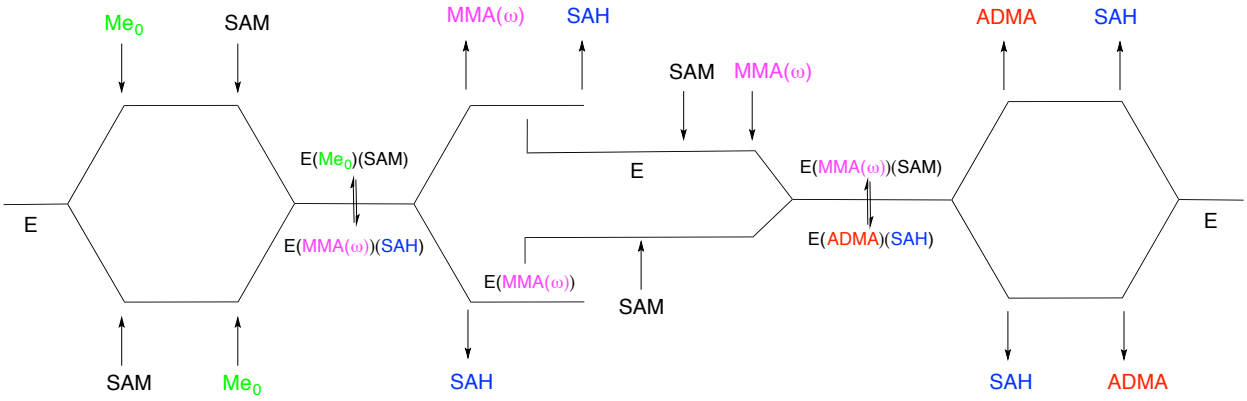


Scheme 1.4 – *Proposed mechanism for PRMT1-catalysed methylation of substrate Arg as an archetypal example for all PRMTs. Scheme reproduced from Zhang et al.*[54].

1.2.4.b - Kinetics

PRMT1 is reported to follow partially-processive enzyme kinetics, defined by Obianyo *et al.*[56] as follows: "The term partially processive was used to explain that PRMT1 catalyzes the production of MMA and ADMA containing peptides in approximately equimolar amounts, even in the presence of an excess of the unmethylated peptide substrate, ruling out a fully processive mechanism in which the production of the ADMA product is obligatory, i.e. the concentration of the intermediate does not rise above the concentration of the enzyme."

Importantly, dissociation of SAH must occur after the first methylation event and prior to the second to allow space for a second equivalent of SAM in the cofactor-binding pocket (**Scheme 1.5**). The term 'partially processive' in this context means that intermediate MMA(ω) may or may not dissociate before the second methylation event takes place.



***Scheme 1.5** – Proposed 'partially processive' mechanism for PRMT-catalysed methylation of Arg, adapted from Obianyo* et al.[56]. *Me_0 = substrate containing unmethylated Arg, MMA(ω) = intermediate containing ω-monomethylated Arg, ADMA = product containing asymmetrically dimethylated Arg, SAM = cofactor* S-*adenosyl-*L*-methionine, SAH = cofactor breakdown product* S-*adenosyl-*L*-homocysteine, E = enzyme.*

Scheme 1.5, adapted from Obianyo *et al.*[56], implies that order of binding is inconsequential for substrate and SAM. These conclusions were based on kinetic experiments and dead-end analogue inhibition studies, although the validity of these conclusions may be called into question given the complex and poorly understood nature of selective PRMT inhibition (further discussed in **Section 1.5**). In contrast to these findings, structural evidence suggests that SAM binding is necessary to stabilise a conformationally active state of the conserved α-helix that enables substrate recognition[51].

Molecular dynamic simulations (using rat PRMT1 in complex with SAH and peptide substrate[53]) predicts MMA(ω) turnover to be faster than unmethylated substrate turnover[54]. This was corroborated experimentally by kinetic experiments using un- and mono-methylated peptide pairs and comparing their k_{cat}/K_m values (**Table 1.1**)[57].

Peptide	Sequence	k_{cat}/K_m, ($x10^2$ M^{-1} s^{-1})
A	YIHRIG**R**GGR	13 ± 1.7
A-Me	YIHRIG**R(Me)**GGR	23 ± 5.5
B	KGGFGG**R**GGFGGKW	5.7 ± 0.3
B-Me	KGGFGG**R(Me)**GGFGGKW	7.6 ± 1.8
C	GG**R**GGFGGKGGFGGKW	17 ± 5.5
C-Me	GG**R(Me)**GGFGGKGGFGGKW	27 ± 8.9

Table 1.1 *– Data from Gui* et al.[57] *showing increased catalytic efficiency, as defined by an increased k_{cat}/K_m , for monomethylated peptides compared to their unmethylated counterparts.*

This latter piece of work postulated that degree of processivity (as defined above) and catalytic efficiency are strongly affected by the peptide sequence, length and the number of Args present in the substrate[57]. Additionally, subtle differences in the residues flanking the substrate GAR sequence are thought to have an impact on the overall methylation number of the product but further work will be necessary to unravel these intricacies.

1.2.4.c – Methylation Pattern is Conferred by Conserved Active Site Residues

As well as a role in catalysis, it is thought the two conserved Glu residues are involved in methylation pattern conferral. In type I PRMTs, exemplified by CARM1 in **Figure 1.8A**, one of these Glu's is predicted to hydrogen bond to the Arg substrate and hold it in position for attack on SAM[51]. H415 and E258 hydrogen bond to the ω-nitrogen and thus may act as a steric block to disfavour *N*-methylation at this position (**Figure 1.8A**). The cavity observed below Y518 however is thought to be large enough to accommodate two methyl groups on the same nitrogen of the substrate guanidine. Together these observations are consistent with an ADMA-producing phenotype.

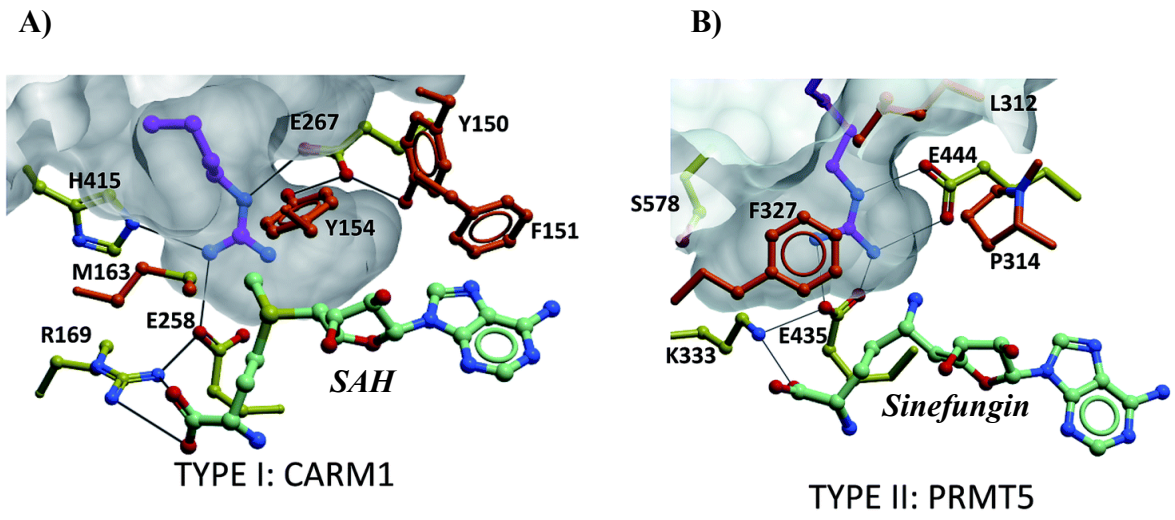


Figure 1.8 *– Structures of* ***A)*** *CARM1 with SAH (PDB:2Y1X)*[58] *as a representative of the type I PRMTs and* ***B)*** *PRMT5 with Sinefungin (PDB: 4GQB)*[59] *as a representative of the type II PRMTs. Structures adapted from Schapira* et al. *published by The Royal Society of Chemistry* [51].

For type II PRMTs, exemplified by PRMT5 in **Figure 1.8**, F327 is thought to be an essential contributor to the SDMA-producing phenotype[51,60]. SDM studies introduced an F327M mutation to PRMT5, which resulted in a mutant enzyme that could produce native SDMA and also unnaturally produce ADMA. This was determined by Western blot analysis using methylation-pattern specific antibodies (confirmed to be specific *vs.* controls) although no quantification of the ratio of products was obtained[60]. The authors hypothesise that this implicates F327 in conferring a type-II phenotype on PRMT5 by introducing conformational inflexibility. It was thought that the flexibility of Met – which is found in this position in type I enzymes e.g. M163 in CARM1 (**Figure 1.8A**) - plays some part in conferring an ADMA-phenotype. However, the authors comment that the reverse mutation (M to F) in PRMT1 did not confer a type II phenotype, although the data is not presented. It could however be argued that this explanation is not consistent with the observation that type I enzymes cannot produce SDMA. The relevance of this finding is therefore called into question, especially when one considers these results are corroborated by a crystal structure of *Caenorhabditis elegans* PRMT5 in complex with SAH rather than peptide (PDB: 1OR8)[60].

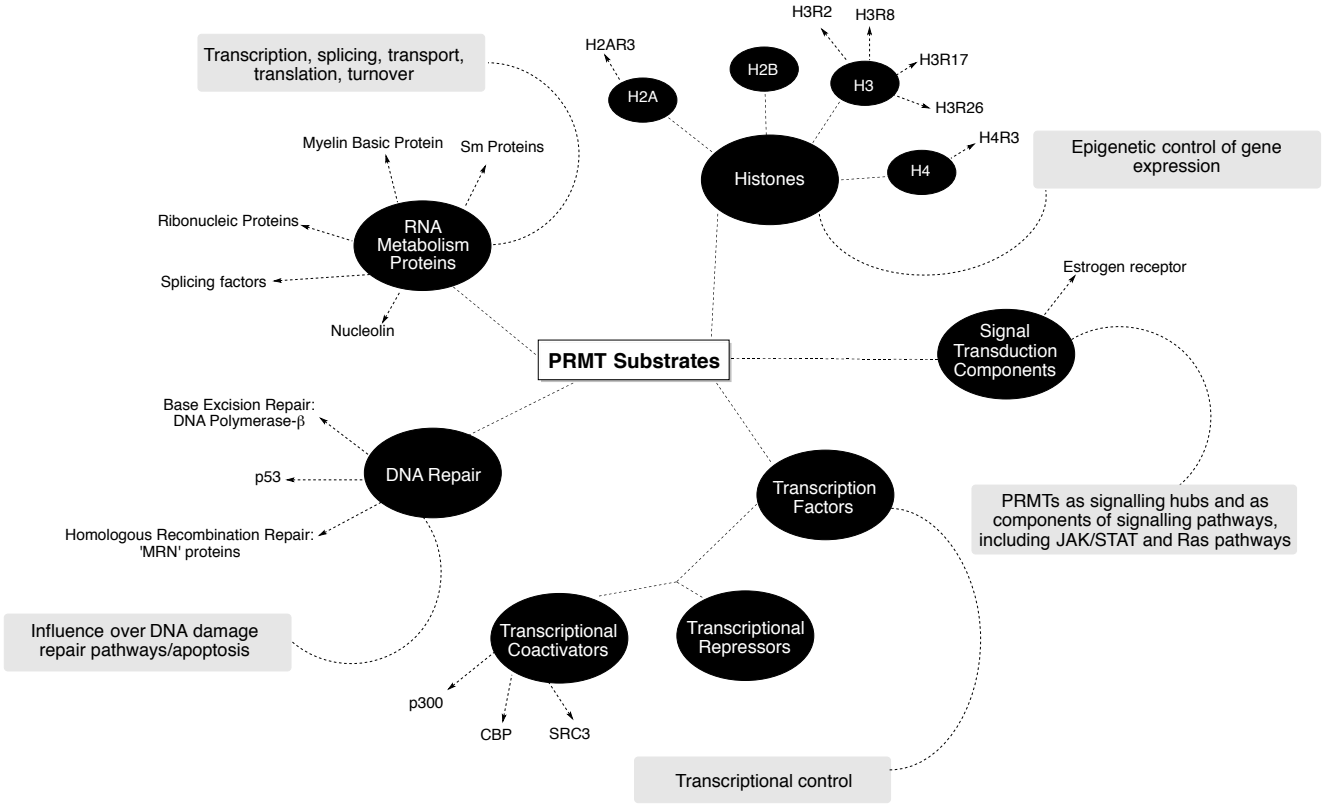
Human PRMT5 requires the coenzyme MEP50 for catalytic activity, and the subsequently reported structure of the human PRMT5:MEP50 complex (PDB: 3GQB) allowed a more detailed insight into possible structural features that may confer an SDMA-producing phenotype (**Figure 1.8B**)[59]. Here, the H415 in CARM1 is seen replaced by a serine (Ser, S); S578 is smaller than H415 suggesting an *N*-Me group may be accommodated in this cavity[51].

If F327 in PRMT5 truly does restrict repositioning of the Arg substrate[60] then it would be necessary for the MMA(ω) intermediate to dissociate fully from PRMT5 before re-binding ('non-processive kinetics') in a new orientation with *N*-Me occupying the cavity adjacent to S578 before a second methylation event could take place. Kinetic analyses[59] support this theory and until any further mutagenesis, structural and kinetic studies take place, it seems a reasonable working model.

1.2.4.d – The PRMT Family Has a Diverse Substrate Range and Functional Relevance

Figure 1.9 offers an insight into the type of processes PRMTs are implicated in and a broad overview of the classes of substrates upon which the PRMTs act, which include: histones, transcription factors, transcriptional coactivators, transcriptional repressors, RNA metabolic regulators, signal transducers, and DNA repair complexes and polymerases[61-64].

The network is expanding, and innovative methods are being employed to identify further substrates. Notably, allyl- and azido- SAM-analogues have been used to label PRMT substrates and subsequent 'click chemistry' reactions with an allyl/azido-reactive fluorophore have facilitated their identification[65].



PRMT	Function	Type	Primary Substrates
PRMT1	Transcriptional activation, signal transduction, RNA splicing, DNA repair	Type I	H4R3, MRE11, 53BP1, SAM68
PRMT2	Transcriptional regulation	Type I	H3R8
PRMT3	Ribosomal homeostasis	Type I	RPS2, P53
CARM1	Transcritional activation, RNA splicing, cell cycle progression, DNA repair	Type I	H3R17, AlB1, p300, CBP, RNA Pol II CTD
PRMT5	Transcriptional repression, signal transduction, piRNA pathway	Type II	H3R8, H4R3, E2F1, p53, EGFR, CRAF
PRMT6	Transcription regulation	Type I	H4R3, H2AR29
PRMT7	Male germline gene imprinting	Type II/III	H4R3, H2AR3, H3R2
PRMT8	Brain specific function	Type I	Unknown
PRMT9	Role in splicing	Type II	SAP145

***Figure 1.9** – Schematic depicts a broad overview of PRMT substrates and subsequent cellular effects; Black ovals: class of PRMT substrate; Grey boxes: processes that are affected; Unboxed writing: examples of substrates. Table gives examples of major substrate classes and is adapted from Yang & Bedford[28].*

1.2.5 – The PRMTs are Modern Therapeutic Targets

Evidently the PRMTs have a very broad substrate scope, therefore it is natural that dysregulation of PRMT expression, activity or control could cause a whole host of disease phenotypes. McBride and Silver's review points to possible PRMT involvement in cardiac disease, multiple

sclerosis and autoimmune disease[66] while a more updated account is provided by Bedford and Clarke who add spinal muscular atrophy and viral pathogenesis to the list[62]. The most evidence though has been gathered for PRMT involvement in cancer, which is the angle of interest for work presented in this thesis.

1.3 – The PRMTs Are Anti-Cancer Targets

1.3.1 –Cancer Pathogenesis

In their 2000 seminal review, Hanahan and Weinberg defined six 'hallmarks' that a healthy cell acquires to be cancerous[67]; these are defined by the authors as 'acquired functional capabilities that allow cancer cells to survive, proliferate and disseminate…'. Their updated model[68] includes two further hallmarks and two enabling characteristics. The hallmarks of cancer encompass properties that promote genome instability and dysfunctional DNA regulation, which facilitate mutation, and properties that enhance a cell's ability to survive, propagate and metastasise.

One of the hallmarks, 'deregulating cellular epigenetics', is of particular relevance when discussing PRMT involvement in cancer pathogenesis. Any process that exerts control over gene expression is prone to exploitation by cancer; as epigenetic 'writers', the PRMTs are heavily involved in epigenetic regulation.

The contributions of PRMT1, CARM1 and PRMT5 to cancer pathogenesis are well documented, with roles for the remaining family members being less well defined[28,64,69]; the following subsections discuss the evidence for this.

1.3.2 – PRMT1 and Cancer

One of the earliest indications of PRMT involvement in cancer was the role of PRMT1 in oncogenic transformation of mixed lineage leukaemia (MLL) cells. The *MLL* gene is commonly mutated in both acute myeloid and lymphoid leukaemias where it contributes to leukaemogenesis by forming one of a large number of different chromosomal fusions[70].

The first evidence of oncogenesis that is PRMT family-dependent was reported in 2007 where PRMT1 was found to be part of a three-protein complex (PRMT1-Sam68-CBP) that has oncogenic transcriptional activity in MLL cells[71]. Elegant *prmt1* truncation and mutation studies demonstrated that the asymmetric dimethylation pattern at H4R3 conferred by PRMT1 is critical to leukaemogenesis in MLL. Overall, H4R3 methylation by PRMT1 and H4 global acetylation by complex member CBP (a histone acetyltransferase (HAT)) were crucial for expression of key downstream MLL oncogenic transforming factors (**Figure 1.10**).

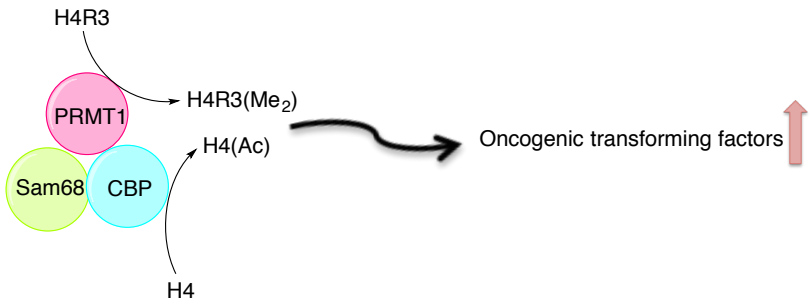


Figure 1.10 – *Upregulation of oncogenic transforming factors as described by Cheung* et al.[71].

In addition to its role in blood cancer, histological analyses have also showed PRMT1 (along with PRMT6) to be upregulated in lung, bladder and breast cancer[72]. Recently, it was found that PRMT1 might have some control over the epithelial-mesenchymal transition (EMT) in non-small cell lung cancer. EMT enables cell movement and therefore facilitates metastasis, one of the final hallmarks of cancer. E-Cadherin is frequently used as a biomarker in this context and its down-regulation heralds the EMT. Avasarala *et al.* altered PRMT1 levels in NSCLC cell lines by

overexpression and siRNA knockdown, and used immunoblotting to identify PRMT1 and E-Cadherin followed by densitometry to quantify them[73]. They showed that high PRMT1 expression is associated with low E-Cadherin levels, drawing the conclusion that PRMT1 can facilitate the EMT. Whilst densitometry data should always be treated with caution given the technique's limitations[74] (e.g. limited concentration range, non-equivalent staining etc.) these findings were augmented by microscopy work (wound-healing and invasiveness assays) showing a statistically significant correlation between PRMT1 upregulation and propensity for EMT. The authors suggest the E-Cadherin repressor Twist1 to be a PRMT1 substrate and the mediator of this effect.

1.3.3 – CARM1 and Cancer

Historically, there has been controversy surrounding the role of CARM1 in cancer development. CARM1 is most commonly associated with estrogen-receptor positive (ER+) breast cancer but there remains a conflict of opinion as to whether it accelerates or inhibits cancer development. For example, Frietze *et al.* propose that CARM1 upregulation is essential for proliferation of MCF-7 cells in ER+ breast cancer[75]. DNA fluorescence staining and flow-cytometry were used to quantify the proportion of cells in each phase of the cell cycle in the presence and absence of the hormone estradiol. Experiments where CARM1 was knocked down (*via* siRNA) gave a lower proportion of cells in S, G_2 and M phases (**Figure 1.11**) compared to control, interpreted by the authors to suggest a role for CARM1 in enabling cell-cycle progression. RT-PCR analyses suggest that CARM1-associated upregulation of the transcription factor E2F1 could bear responsibility for acquisition of this hallmark of cancer.

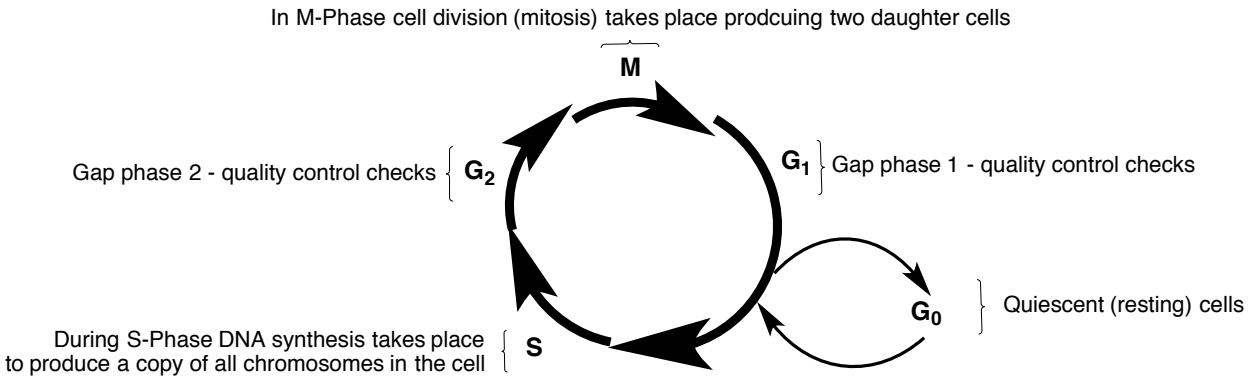
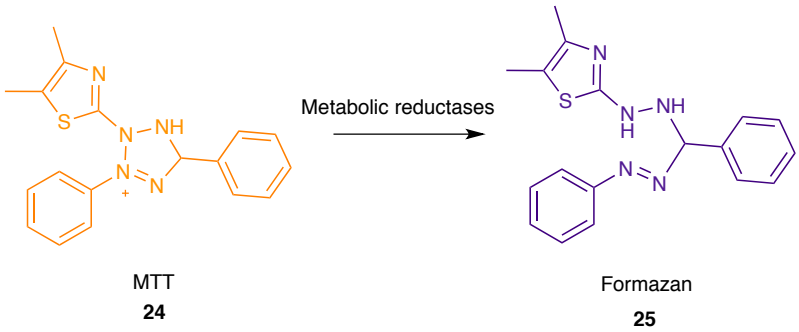


Figure 1.11 *– The cell cycle. During S phase DNA is synthesised and in M phase the cell divides by mitosis. G-phases are 'gap' phases whereby 'quality control' checks take place. Figure reproduced from www.nexcelom.com.*

In contrast to the evidence above, Kim *et al.* propose that CARM1 is not upregulated in breast cancer (and prostate) samples[76]. This study analysed tissue microarrays of real patient samples so arguably the data is of more clinical relevance than that of Frietze *et al.*[75] who used model cell lines. However, it should also be noted that immunohistochemical (IHC) analyses of these tissues were performed using polyclonal anti-CARM1 antibodies with no indication of how they were generated. Quality of antibody can have a profound impact on the sensitivity and/or specificity of an immunochemical experiment and there is a possibility that use of polyclonal antibodies, which are heterogeneous by nature, could have introduced error that skewed the results.

Going one step further than Kim *et al.*[76] is the study by Al-Dhaheri *et al.*[77] suggesting CARM1 can actually inhibit cell growth in ER+ cancer. This was monitored using a colorimetric cell proliferation assay that relies upon turnover of reagent 3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide **24** ('MTT') by metabolically active cells *via* dehydrogenase activity (**Scheme 1.6**). One of their proposals was that higher levels of CARM1 bring about an upregulation of the cyclin-dependent kinase inhibitors (CDKs) $p21^{cip1}$ and $p27^{Kip}$ therefore

inhibiting cell growth. The latter conclusion appears to have been based primarily on qualitative Western blot analysis, with no discussion of any attempt to quantify the two CDKs, so further work in this area is undoubtedly required.



Scheme 1.6 *– MTT-based cell proliferation assay relies on formation of purple formazan that can be quantified by spectrophotometry. MTT is turned over by metabolically active proliferating cells.*

Recently though, more evidence is accumulating to support the role of CARM1 as an oncogenic contributor in breast cancer. CARM1 is now thought to negatively regulate the tumour suppressor pRb, which controls cell-cycle restriction, by methylating it[78]. As part of this work pRb was identified as a CARM1 substrate by monitoring CT_3 transfer onto whole (and truncated) pRb followed by MS corroboration using synthetic peptides. Recently, CARM1 has also been shown to correlate with aggressiveness of tumour in ER+ breast cancer, although using polyclonal anti-CARM1 antibodies as part of an IHC study[79]. Perhaps most strikingly, in 2014 the chromatin remodelling factor BAF155 was shown to be methylated by CARM1 resulting in its activation and therefore enhancement of its downstream effects, which include tumour progression and metastasis in breast cancer cells[80]. Antibodies generated against asymmetrically dimethylated BAF155 peptide sequences were used in tandem with CARM1 knockout cell lines. Knockouts can often represent an unrealistic ablation of a protein's role, but the findings in this study were corroborated by shRNA knockdown. Each are not without their limitations and use of an accurate on-target chemical probe might give this work added value.

In addition to the plethora of reports documenting CARM1 involvement in breast cancer, it has also been implicated in androgen-resistant prostate cancer[81], colorectal cancer[76], melanoma[82] and early-stage hepatocarcinomas[83].

1.3.4 – PRMT5 Causes Malignancies by Influencing Cell-Cycle Control

A wealth of evidence supports the involvement of PRMT5 in malignant transformation of cells; the intimate relationship between PRMT5 and cell-cycle control proteins bear responsibility for this.

PRMT5 and the Guardian of the Genome:

PRMT5 interacts with p53, the master regulator of genomic stability; the former is recruited to the p53 complex when DNA has been damaged[84]. This work also showed that during DNA-damage, p53 influences the cell's decision to attempt repair *vs.* apoptosis. Using ectopic and endogenous PRMT5 and p53, as well as siRNA-knockdown, levels of PRMT5 in the cell were manipulated. It was concluded that PRMT5-mediated methylation of p53 influences the p53 response by altering its promoter specificity; p53 R to K mutants (R333K, R335K, R337K) were assayed *via* chromatin immunoprecipitation (ChIP) to support this claim. Notable was the observation that depletion of PRMT5 prompted p53-driven apoptosis.

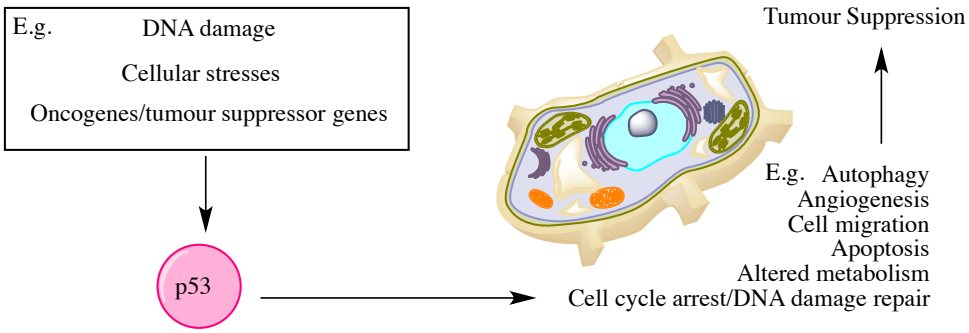


***Figure 1.12** – p53 is a transcription factor whose specificity for different promoters, and therefore downstream effects, is influenced by various different factors.*

Subsequent research using the *C. elegans* PRMT5 homologue, which has 48% sequence similarity to human PRMT5, suggested similar apoptotic-control in the nematode worm[85]. Of particular interest is the finding that nematode *prmt-5* inactivation increases germ cell apoptosis following exposure to ionising radiation. By extension, one might postulate that targeted radiotherapy could sensitise human cells in a similar manner; if a PRMT5 inhibitor is developed it could therefore offer a new therapeutic avenue in the treatment of apoptotic- or radiotherapy-resistant cancers.

PRMT5 and Lymphoma:

A direct link exists between PRMT5 and mantle cell lymphoma (MCL). Elevated levels of PRMT5 in MCL cells increases methylation at H4R3 and H3R8, which in turn represses expression of the tumour suppressor gene (TSG) 'ST7'[86]. Masking TSGs is a commonly employed tactic for nascent cancer cells whilst acquiring the hallmarks of cancer. Furthermore, this study showed that lentivirus-induced PRMT5-knockdown inhibited proliferation of MCL and Burkitt's lymphoma cells (JeKo and Raji cell lines respectively).

Chung *et al.* offer further proof of concept for PRMT5 as an anti-cancer therapeutic target in blood cancers[87]. Their work on Non-Hogkin Lymphoma (NHL) cells identified an inverse correlation between PRMT5 overexpression and amount of the pro-apoptotic factor RBL2. NHL cells typically exhibit overexpressed levels of PRMT5. Upon shRNA-knockdown of PRMT5 in a mouse lymphoma cell line, cancer cell growth was halted and apoptosis was induced. It is possible that this is a result of the concomitant increased expression of RBL2, as monitored by RT-PCR. The same relationship between PRMT5 and RBL2 levels was observed by Western blotting using polyclonal antibodies. Despite the associated flaws of a polyclonal antibody, in the context of the quantitative RT-PCR data this experiment substantiates their conclusions.

PRMT5's essentiality in lymphomagenesis, and therefore its plausibility as an oncogenic transforming agent in blood cancer, was only very recently demonstrated[88]. A dominant-negative *prmt5* construct was introduced to bone marrow cells and transplanted into mice, as was a shRNA construct. No effect on disease progression was observed but Western blot analysis suggested the mice overcame the shRNA-mediated knockdown and restored PRMT5 expression. This observation highlights the shortcomings of RNAi techniques to knockdown protein function. Despite this, the authors utilised their dominant-negative construct *in vitro* and showed that colony formation was inhibited, even when exposed to various different oncogenic drivers. This work therefore suggests, albeit tentatively, a necessity for PRMT5 in oncogenic transformation.

PRMT5 and Melanoma:

Depletion of PRMT5 in malignant melanoma cells by siRNA up-regulated expression of $p27^{Kip1}$ and down-regulated expression of the transcription factor MITF, although it should be noted these observations were from qualitative immunoblotting studies[89]. $p27^{Kip1}$ is a CDK inhibitor that inhibits cell cycle progression, whilst MITF is essential to melanocyte proliferation. Based on the observations, one would logically expect a decreased proliferation upon PRMT5 knockdown, but this was only observed for a subset of melanoma cell lines. In some cell lines proliferation was actually increased. It is possible that further quantitative data is required, e.g. by RT-PCR, in order to unravel the complexities described. This thesis postulates that an improved understanding of epigenetic markers, specifically the possible existence of novel Arg methylation patterns, could give some insight into the aforementioned oxymoron.

A second important observation was noted in this work by Nicholas *et al.*: histological analysis of athymic mice revealed diverse patterns of PRMT5 subcellular-localisation[89]. That same year, Shilo *et al.* showed that not only did different types of lung cancer have different PRMT5-localisation patterns, but the localisation pattern also correlated with tumour grade. An extension of this is that localisation patterns of PRMT5 could be used as predictors for lung cancer stage in diagnostic pathology[90]. A unified procedure using carefully developed and selected antibodies would be a prerequisite for this to be a successful diagnostic approach.

PRMT5 and Gastric, Lung and Bladder Cancers:

Indication of PRMT5 involvement in gastric cancer was initially suggested from cDNA microarray studies[91]. Later, expression of *prmt5* was screened for in a systematic genome-wide search of the NCBI 'GEO' database, revealing statistically significant upregulation in gastric, lung and bladder cancers as well as in metastatic colon cancer[92]. This was the first time PRMT5 was tested for its ability to transform cells; it is thought that PRMT5 upregulates various cyclins and CDKs to accelerate cell cycle progression through the checkpoint at G1-phase as analysed by flow cytometry[92].

It would seem that PRMT5 is also implicated in a variety of cancers other than the major examples discussed above, and as more reports surface we are able to paint a clearer picture of PRMT5's role in carcinogenesis. For example, reports of PRMT5 involvement in breast cancer[93,94] and neuroblastomas[95] are currently emerging.

1.3.5 – PRMT2, PRMT3, PRMT6, PRMT7, PRMT8 and PRMT9 in Cancer

The remaining family members – PRMTs 2, 3, 6, 7, 8 and 9– are certainly not as well studied as the founding members – PRMT1, CARM1, PRMT5 – in the context of cancer research. A small

amount of evidence suggests that a role in cancer pathogenesis might extend to all members of the PRMT family[28], but much of this data is tentative and some of these PRMTs are newly discovered and as yet poorly characterised.

1.3.6 – Inhibiting PRMTs Could be a New Therapeutic Avenue for Cancer

There are links between some PRMTs and some specific forms of cancer; certainly PRMT1, CARM1 and PRMT5 are now widely accepted as anti-cancer targets. At times the evidence for PRMT involvement in cancer can be contrasting, possibly because of inaccuracies resulting from the techniques used to draw such conclusions, as critiqued above.

Chemical probes that are able to selectively target PRMTs could be very useful tools for furthering our understanding of PRMT involvement in cancer and could certainly be used to augment some of the knockdown work implicating the family in cancers; indeed the search for selective inhibitors of PRMTs is ongoing. However, probe design and development should be a very careful process – in a recent commentary[96] a large cohort of leading chemical probe experts offered their opinions and advice on effective probe development. Irrelevant and inappropriate probes are in use every day, mostly as a result of poor specificity of these probes. In an effort to remedy this problem the authors have built an open source database: the 'Chemical Probes Portal'[97].

1.4 – Assaying For PRMT Activity

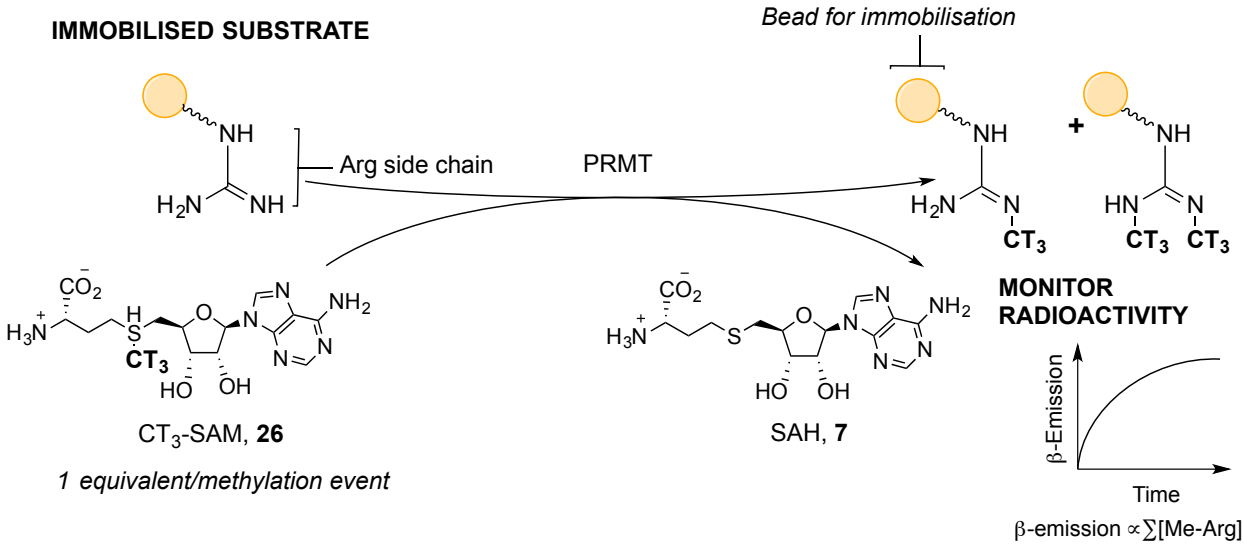
PRMT activity can be monitored directly or indirectly, though the former is more common. The assay techniques described below can also be applied to study of other MTs, such as the PKMTs, and can be adapted for most SAM-utilising enzymes.

1.4.1 – Direct Recognition of a Methylation Event

Direct recognition monitors the appearance of a methyl group either by radioactive-labelling of the methyl-donor, antibody-based recognition of the nascent methyl pattern or by mass spectrometry on the substrate/product.

1.4.1.a – Radioactivity Readout Using Tritiated or C^{14}-Labelled SAM

The radiometric assay using CT_3-labelled SAM has been widely considered the gold-standard PRMT assay. Substrates are immobilised and a percentage of hot SAM is added to the assay mixture. Unreacted cofactor is washed out when the assay is terminated and PRMT activity is monitored as a function of beta-radiation emission (**Scheme 1.7**). This technique is medium throughput but is not without its drawbacks. The assay monitors total radiation output so is vulnerable to noise from unreacted cofactor and requires specialist training and monitoring for radiation workers.



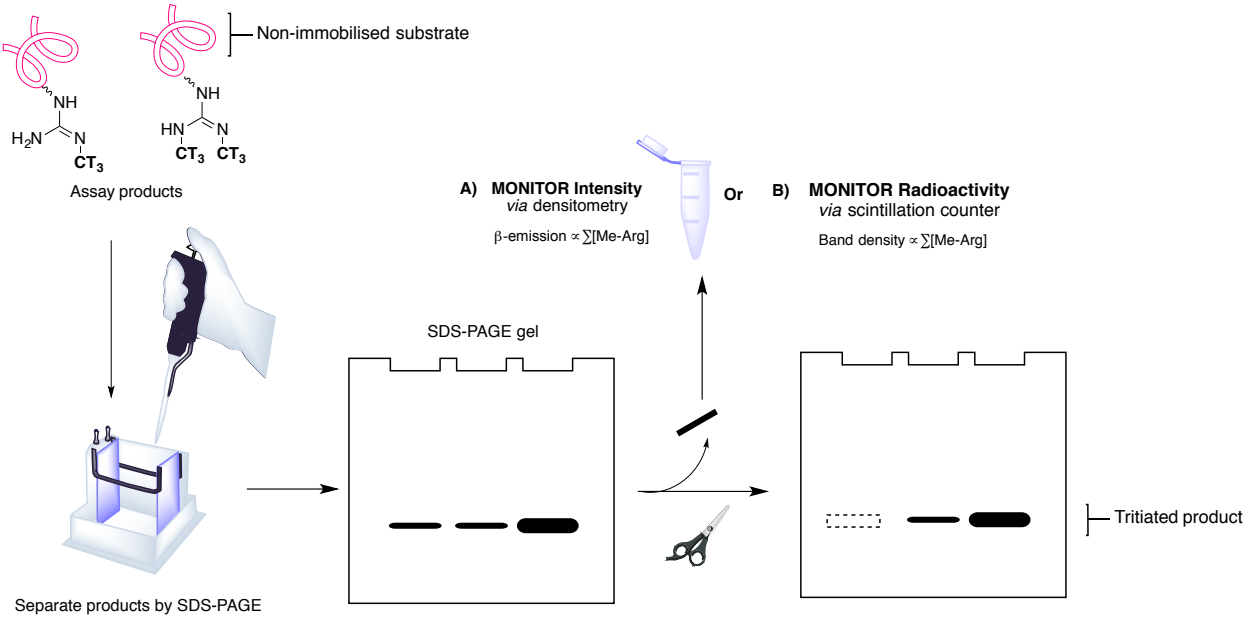
Scheme 1.7 *– The radiometric assay monitors transfer of a tritiated methyl group to substrate Arg.*

An analogous assay uses C^{14}-labelled SAM in a similar manner to above, with the same limitations.

A low-volume high throughput radiometric assay – HotSpot^{SM} – has recently been described for histone MTs[98]. HotSpot^{SM} negates some of the disadvantages of the original radiometric assay by reducing the amount of protein and radioactive SAM required.

Further variations on the radiometric assay include using SDS-PAGE to separate substrates, followed by densitometry or scintillation counting of excised radioactive bands or bands to elucidate activity (**Scheme 1.8**). This technique should be used cautiously because of the limitations of densitometry as discussed in **Section 1.3**[74].

The most pertinent disadvantage of radiometric assays is their inability to distinguish between mono- and di-methylated Arg. This thesis hypothesises that radiometric assays are at best insensitive to kinetic metadata, and at worst prone to inconsistency due to an inability to differentiate between mono- and di-methylated products.



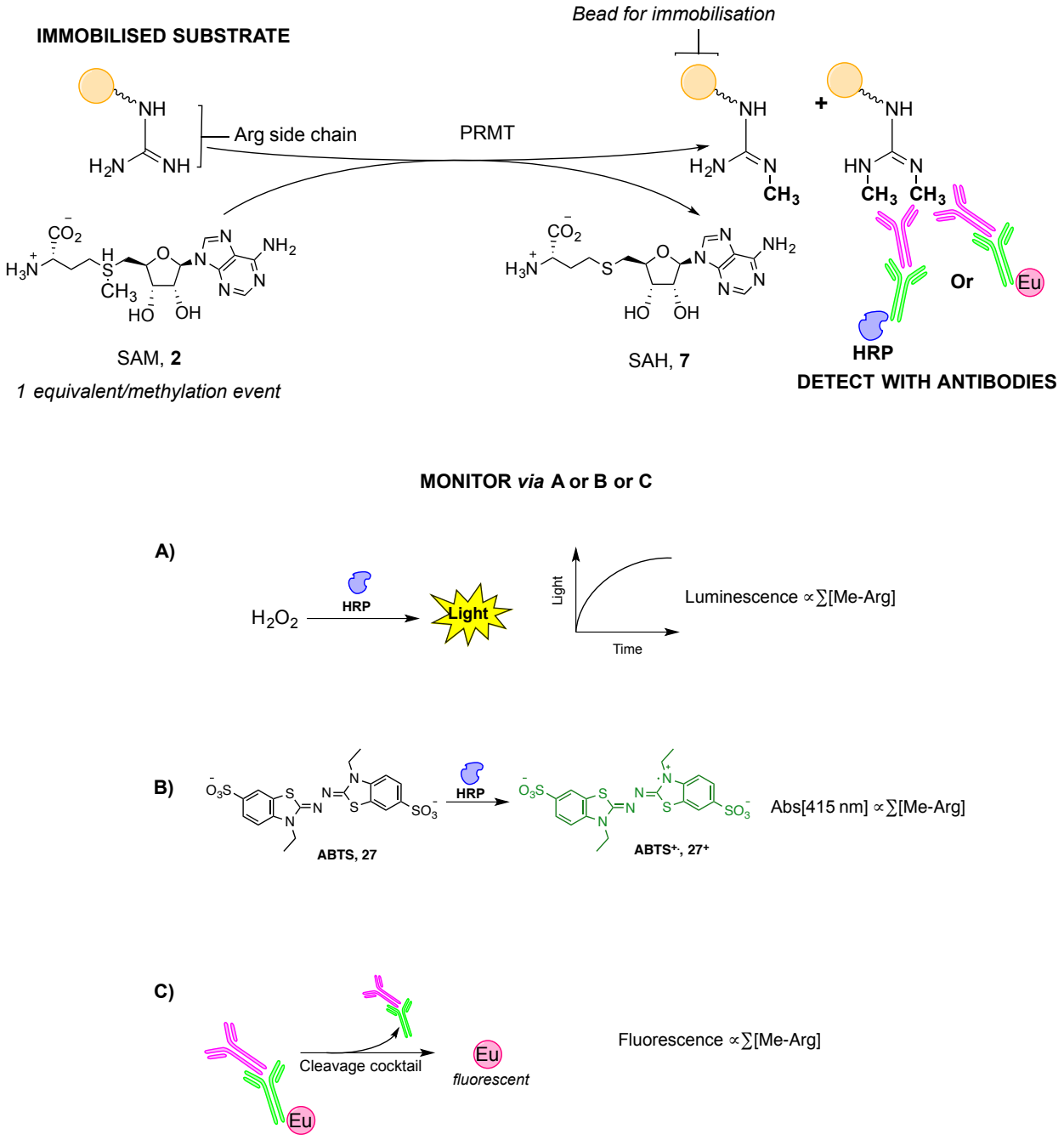
Scheme 1.8 *– Use of SDS-PADE to separate non-immobilised radioactive assay products. Following excision of product bands from the gel, PRMT activity is supposedly quantified by monitoring* ***A)*** *density of individual bands or* ***B)*** *radioactivity of individual bands.*

1.4.1.b – Antibody-Based Recognition of Methyl Markers

Antibodies can also be used to assay for PRMT activity using immobilised substrate. A primary antibody that recognises methylated Args is first used, followed by a secondary antibody that is typically conjugated to a readout-effector and facilitates quantification of PRMT activity. Conjugation to horseradish peroxidase (HRP) allows quantification by monitoring chemiluminescence upon addition of the HRP substrate H_2O_2 (**Scheme 1.9A**).

Conjugation to α-mouse peroxidase has also been used to quantify PRMT activity *via* a colorimetric assay[99] where the substrate 2,2′-azino-bis(3-ethylbenzothiazoline)-6-sulfonate (ABTS) is used to generate a green radical cation which absorbs at 415 nm (**Scheme 1.9B**). HRP is also able to catalyse the colorimetric reaction.

Some assays use time-resolved fluoresence (TRF) to monitor PRMT activity after detection of product with antibodies. Europium conjugated to the secondary antibody is often used as the fluorophore[100]. Generally, after a washing step, Eu is cleaved from the secondary antibody and exposed to flash of light. Fluorescence decay is monitored to quantify extent of methylation (**Scheme 1.9C**).



Scheme 1.9 *– An antibody-based assay using primary and secondary antibodies to detect Arg methylation patterns. Secondary antibodies conjugated to horseradish peroxidase (HRP) can be used to monitor* ***A)*** *chemiluminescence using H_2O_2 as a substrate or* ***B)*** *fluorescence using, for example, ABTS as a substrate. Secondary antibodies conjugated to* ***C)*** *Eu can be used in a time-resolved fluorescence (TRF) assay.*

Antibodies are highly specific for their antigen, but their ability to distinguish un-, mono- and di-methylated Arg has been called into question[101]. Egelhofer *et al.*[101] have created an online database to which researchers may submit pass/fail assessments of different antibodies based on pre-defined metrics[102]. To date, this database indicates a 0-50% pass rate for 18 different supposedly specific antibodies that target histone arginine methylation patterns. This brings in for stark examination the suitability of using antibodies in assessing PRMT-conferred methylation patterns. For antibody-based PRMT assays, even if epitope-specific antibodies were to be used, it would remain impossible to distinguish whether luminescent/colorimetric readout originates from mono- or di-methyl Arg and kinetic information is lost this way, leading to a similar disadvantage observed for radiometric assays. Another consideration is that in the assays described above, a secondary antibody is used to amplify signal, which can be another source of error.

When using chemiluminescence a final further disadvantage arises; luminescent readout has a narrow working range, meaning signal saturation can happen very easily and reduce assay sensitivity. However, chemiluminescence and colorimetric assays enable higher throughput than standard radiometric assays since their readout can be monitored rapidly and automatically on a plate-reader. However, the validity of such data, for the reasons alluded to above, could be questionable. Furthermore, a similar problem exists for fluorescence-based readout where quenching and phosphorescence at orthogonal wavelengths can be problematic (**Figure 1.13**).

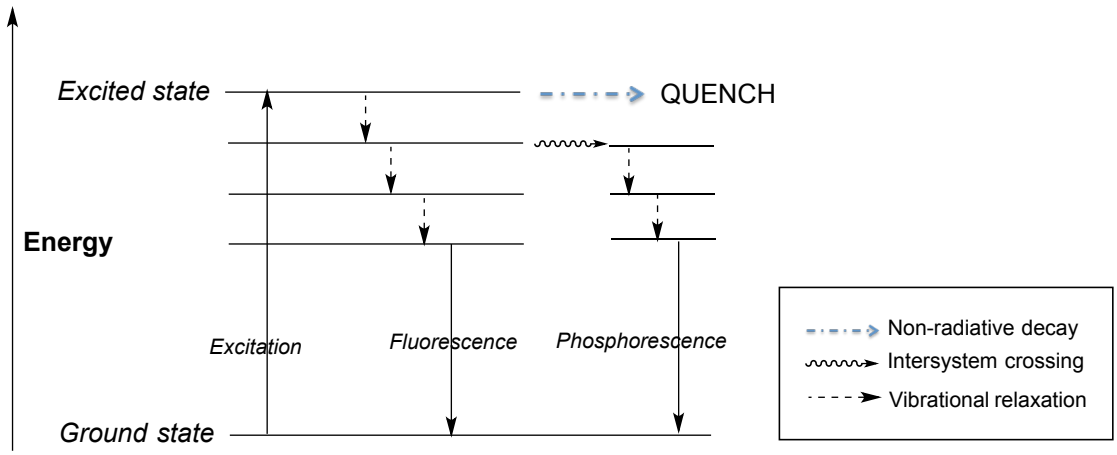


Figure 1.13 *– Fluorescence energy diagram.*

1.4.1.c – Monitoring Methylation by Mass Spectrometry

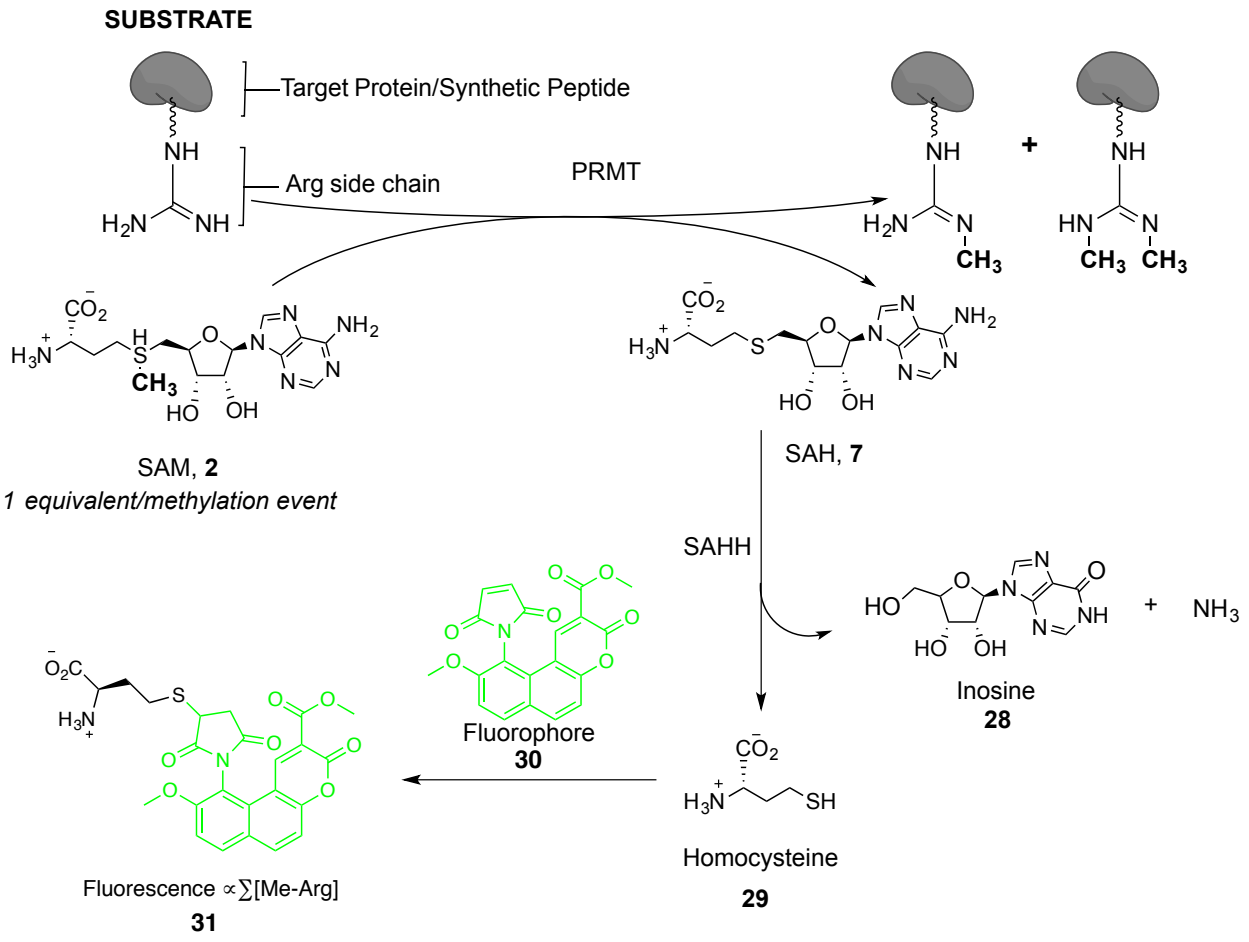
Addition of a methyl group to a peptide substrate can be detected by observing the concomitant mass increase using mass spectrometry (MS). Published accounts include those using ultra-high performance liquid chromatography (UPLC) followed by MS/MS analysis to monitor PRMT-mediated methylation of peptides containing fluorinated and ethylated Args by following their fragmentation patterns[103]. Furthermore, a MS 'fingerprint' exists for the fragmentation of different methylated Args; Gehrig *et al.* used electrospray triple quadruple tandem MS to distinguish between different methylated Arg standards based on their ionic fragmentation patterns[104]. If this technique were applied to an assay protocol in the future, then a powerful MS assay might result which could inform us of the methylation regiochemistry.

Matrix-assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF-MS, or just MALDI) has been used in kinetic studies[57] to distinguish between methylated peptides with different methylation number, but not regioisomers of the same methylation number. Both MS/MS and MALDI techniques are highly superior to other assays for monitoring distribution of

un-, mono- and di-methyl Args. However, these assays are low to medium throughput, which might go some way to explaining why they are under-represented in the PRMT inhibition literature. The data gathered from a MS assay is enriched with kinetic metadata compared to radiometric and antibody-based assays. MALDI assays are also applicable to monitoring other MTs, demethylases and other epigenetic modifications. MS techniques are advancing all the time and with high throughput technology such as rapid fire MS[105]an assay could be developed that delivers enriched data.

1.4.2 – Indirect Recognition of a Methylation Event by Coupled Assays

Assays that monitor the appearance of SAH **7** offer an alternative to those described above. SAH hydroxylase (SAHH) can be used to couple PRMT activity to fluorescent readout *via* SAHH-mediated breakdown of SAH to the thiol homocysteine **29**. Collazo *et al.* use this assay technique on both the PRMTs and PKMTs[106]. Addition of a thiol-sensitive fluorophore (e.g. **30**) allows quantification of homocysteine **29** (**Scheme 1.10**). One advantage is that fluorescence can be quantified in a high-throughput manner and can be automated using a suitable plate reader. A similar assay has been described for another SAM-dependent methyltransferase (salicylic acid carboxyl methyltransferase) which uses 5,5'-dithiobis(2-nitrobenzene) to quantify based on absorption rather than fluorescence emission[107] but the principle is also applicable to the PRMTs.



Scheme 1.10 *– The SAHH-coupled assay shown using fluorescence to quantify PRMT activity.*

The main disadvantage of any secondary enzyme-coupled assay arises when testing compounds for inhibitory properties; it must be confirmed that compounds are inhibiting the target (i.e. PRMT) rather than the secondary enzyme (e.g. SAHH) to avoid unwanted false positives. The additional limitation of using fluorescence-based readout also applies here (*vide supra*).

1.4.3 – Assay Substrates

As discussed in **Section 1.2.2.c**, the PRMT family acts upon a wealth of different substrates and different family members have different substrate specificities. Full-length or truncated human histone proteins, as well as histone peptides, are very commonly used as assay substrates and use of chicken erythrocyte histones has also been reported[108]. For MALDI assays, a peptide

representing a truncated histone is the most convenient choice for detection; MALDI machines have varying lower and upper limits of mass detection and this should be accounted for when selecting appropriate peptides.

Arguably, the most important consideration is the known/prospective kinetics of the selected substrate. The structural features that control PRMT substrate selectivity and processivity, reviewed in **Section 1.2.2**, must be given particular consideration in PRMT assay design as the researcher considers the information they wish to obtain from their experiment.

1.4.4 – PRMT Sources

The PRMT source is partly dependent on the type of experiment being run and the level of protein purity required. Impure cell lysates can be used, but purified recombinant PRMTs are necessary for early drug discovery in order to establish structure-activity relationship (SAR) data and enzyme specificity. Where recombinant PRMTs are used it is important to avoid expressing the protein from a FLAG-tagged construct since this has been shown to enrich for PRMT5 activity during affinity purification[109]. *E. coli* can be used as the expression system for the majority of PRMTs, the notable exception being PRMT5 which requires co-expression in mammalian cells alongside its coenzyme MEP50. Where mammalian cell expression is used, PRMTs can be isolated by immunoprecipitation. Example protocols for PRMT expression and purification have been published[110] and live versions are available online from the lab of Mark Bedford at the MD Anderson Cancer Centre, University of Texas[111].

1.5 – Drug Discovery of PRMT Inhibitors

1.5.1 – Early PRMT Inhibition

The first identified PRMT inhibitor is the natural product and cofactor mimetic Sinefungin **32** (**Figure 1.14**), which non-specifically inhibits SAM-utilising enzymes[112]. The first example of selectivity for the PRMTs over the PKMTs was reported by Bedford and co-workers[99] from screening a 9000-member Chembridge library. This work identified the AMI (Arginine Methyltransferase Inhibitor) series – unsurprisingly compounds with similarity to Sinefungin **32** were common in this series, and such compounds were found to also inhibit the SET-domain containing PKMTs tested. The original screen (*vs*. 9000 compounds) was carried out using an antibody-based detection method. A monoclonal antibody was raised against methylated Npl3p – a yeast RNA-binding protein. Using Western-blotting, the authors report that treatment of Npl3p with human PRMT1 can also generate an epitope recognisable by the same antibody. It is noted however that this work is lacking a negative control that incubates PRMT1 with antibody in the absence of Npl3p substrate, which arguably gives reason to doubt the accuracy of this screen. Furthermore, the assay readout is monitored by chemiluminescence using a secondary antibody conjugated to a peroxidase as described in **Section 1.4.1.b**.

Hit molecules from the initial screen were assessed for concentration-dependent inhibition *via* IC_{50}. A different assay was used for this work compared to the original screen. The same antibody-based detection was used but instead of chemiluminescence as the readout, the substrate ABTS was added to generate a colorimetric output (described in **Section 1.4.1.b**). AMI-1 **33**, a symmetrical bis-aryl sulfonated urea, gave an IC_{50} of 8.8 μM *vs*. PRMT1 using the colorimetric assay (**Figure 1.14**). To test for methyltransferase specificity, AMI-1 was tested *vs*. a panel of different PRMTs and PKMTs using a radiometric assay with CT_3-SAM; the products of the reaction were separated by SDS-PAGE and the degree of inhibition quantified using

densitometry (**Section 1.4.1.a**). As previously discussed (*vide supra*), densitometry is notoriously imprecise and no other corroborative inhibition assay was reported. The authors' conclusion is that AMI-1 displays no inhibition of SET-domain containing PKMTs.

This inhibitor was tested in MCF7 cells and was shown to suppress downstream roles of PRMT1 and CARM1 *in vitro* and was also shown to not compete for the cofactor-binding pocket by SDS-PAGE. Despite having an IC_{50} over five times less potent than Sinefungin **32** (1.63 μM *Vs.* PRMT1 in this study) AMI-1 **33** was touted as a promising lead compound for PRMT-specificity.

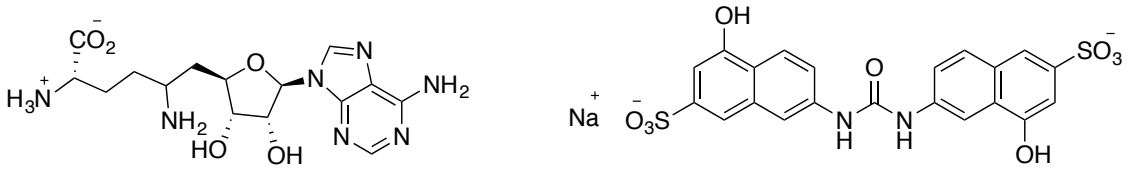


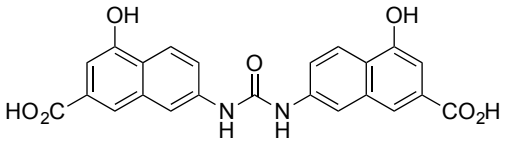
Figure 1.14 *– Pan-MT inhibitor Sinefungin **32** and the first PRMT-specific inhibitor AMI-1, **33**.*

1.5.2 – Small Molecule Inhibitors

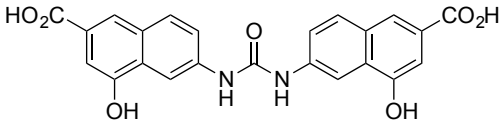
1.5.2.a – Improving Upon AMI-1

Following the seminal paper reporting AMI-1 **33** as the first PRMT-specific inhibitor, a series of target-based studies were conducted in an attempt to find an improvement. AMI-1 was predicted to have low bioavailability and is analogous to suramin-type sulfonated ureas that are thought to have pleiotropic interactions with other proteins, so a cross-collaborative effort began to optimise the compound[113]. Compounds **34** and **35** (**Figure 1.15**) were identified as promising new leads against both recombinant and cellular PRMT1 using a combination of colorimetric and

radiometric assays (**Section 1.4.1**). The authors report selectivity for the PRMT family over the PKMTs but not between different PRMTs (PRMT1, PRMT3, CARM1, PRMT6). It is noted however that these selectivity assessments were carried out using a similar method described in the original work by Bedford[99] involving CT_3 SAM, SDS-PADE separation and relying on quantification by densitometry.



34



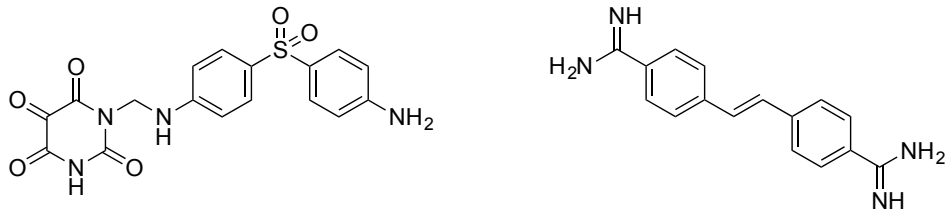
35

Figure 1.15 *– Carboxy- derivatives of the first hit small molecule PRMT inhibitor AMI-1* ***33****.* ***34*** *and* ***35*** *showed selectivity for PRMTs over PKMTs but not between PRMT family members.*

Symmetry was a recurring characteristic of many of the early PRMT inhibitors but Jung and co-workers identified the asymmetric substituted bis-aryl sulfonated urea Allantodapsone **36** as a new lead[100]. Using a homology model of PRMT1 (built using data from co-crystal structures of the rat proteins PRMT1 and PRMT3) *in silico* docking was performed against almost 2000 compounds from the NCI diversity set to identify possible hits.

Despite the limitations of docking experiments based on a homology model, the authors validated potential hits using recombinant human PRMT1 and a TRF Europium-conjugated antibody-based detection assay (**Section 1.4.1.b**). SAR data was suggested and a common feature of the hit molecules showed interaction of a basic moiety with Glu152 in the homology model of human PRMT1. Inhibitors were tested in cellular assays and the top hit Allantodapsone **36**, along with the symmetrical hit Stilbamidine **37** (**Figure 1.16**) were found to be most effective. Kinetic analysis of both **36** and **37** revealed that neither were SAM-competitive but both were substrate-competitive. Interestingly, docking of AMI-1 **33** in this model suggested that

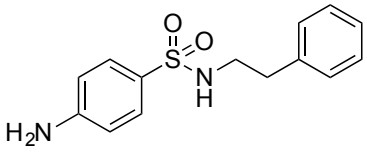
favourable binding conformations were only possible when the cofactor SAH **7** was removed. The presence or absence of the cofactor has a significant influence on inhibitor binding – this correlates well with PRMT structural data reviewed in **Section 1.2.2**.



Allantodapsone, **36** Stilbamidine, **37**

***Figure 1.16** – Allantodapsone **36** and Stilbamadine **37** inhibit PRMT1 in cell assays.*

Jung and co-workers developed this work using a more sophisticated virtual screening tool to assess a 328,000-strong Chembridge library for novel inhibitors[114]. This entailed screening against the previously ascertained human PRMT1 homology model[100] but with the added constraint that hit molecules must pass a 'phamacophore model' built from data and structural information obtained from their earlier work and from the seminal AMI-1 study. Four top hits were pulled from clusters of similar molecules; notably three of these contained sulfonamide bridges, although despite this each hit was classified as a unique chemotype according to their Tanimoto coefficients. The sulfonamide **38** was identified as the top hit in this screen from the Europium TRF assay (**Figure 1.17**).



38

***Figure 1.17** – Top hit from the Jung group, **38**, passed a pharmacophore model-based virtual screen.*

Two further studies sought to improve upon the AMI series, both generating inhibitors with broadly similar potency to the initial report. Bonham *et al.* set out to develop a less polar analogue of AMI-1 **33** by 'melding' different elements of the previously identified AMIs leading to the identification of **39**, a cell-permeable PRMT1 inhibitor (**Figure 1.18**)[115]. Fontán *et al.* identified three symmetrical AMI-based urea-linked scaffolds, **40-42**, that gave poor potency and were ruled out of further investigation (**Figure 1.18**)[116].

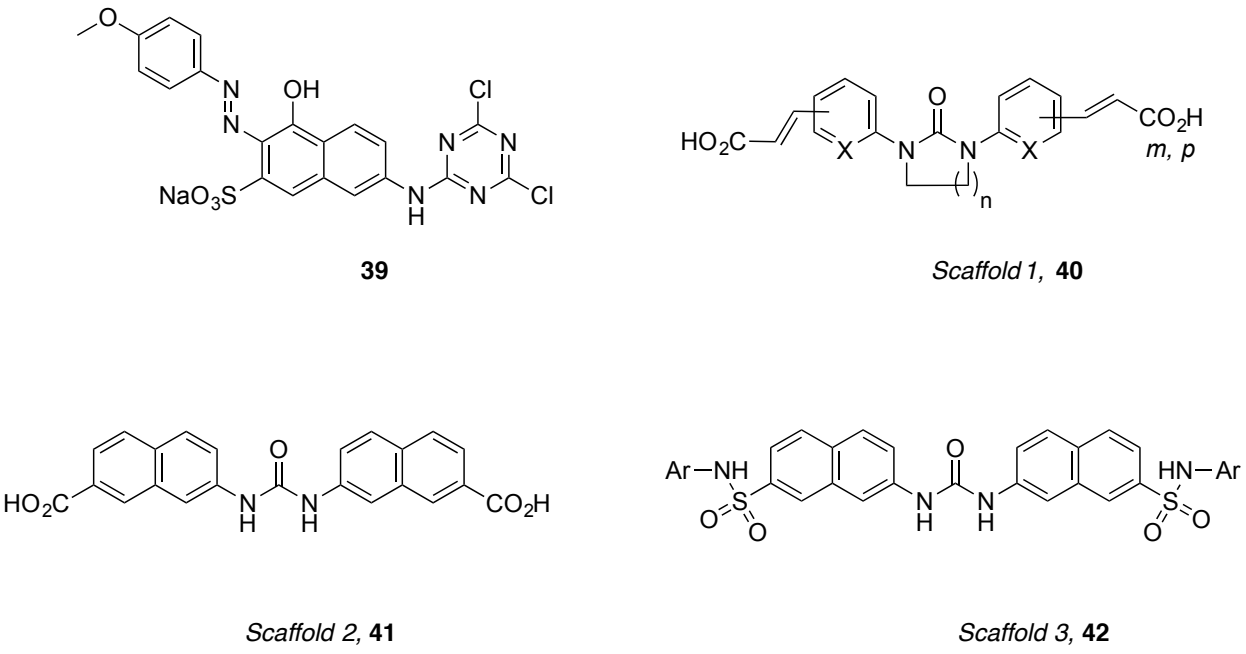
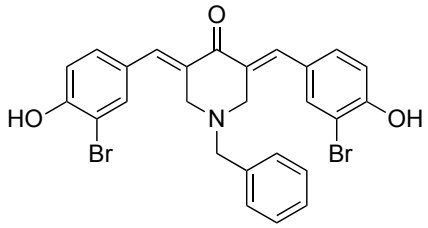


***Figure 1.18 – 39** is a cell-permeable PRMT1 inhibitor based on the structures of the AMI series of compounds. Scaffolds **40-42** are poor PRMT inhibitors, despite being based on the same AMI series.*

Meanwhile, work was underway to selectively inhibit CARM1. High throughput screening and hit-to-lead SAR development identified a series of pyrazole and indole-based inhibitors[117-120]. Valuable lessons, applicable to the wider PRMT field, were learnt from the follow-up work[58]; co-crystal structures of CARM1 in complex with cofactor analogues (SAH **7** or Sinefungin **32**) and hit pyrazole or indole inhibitors revealed the importance of cofactor binding for bringing about the necessary structural changes to facilitate inhibitor binding, hence corroborating aforementioned reports.

Despite demonstrating good inhibitory activity against recombinant CARM1, no compound from the initial screens had demonstrated good potency in a cellular context. Moving back towards AMI-like compounds, Bedford and co-workers addressed this problem and identified **43** (**Figure 1.19**), which demonstrated efficacy against the prostate cancer cell line LNCaP. This result was attributed to the concentration-dependent inhibition of CARM1[121]. Radiometric assays using recombinant PRMT sources suggested **43** has specificity for inhibiting CARM1. Using an inducible CARM1 substrate (PABP1) *in vitro* cellular assays were performed that assessed presence or absence of methylation on PABP1 using an antibody developed by the authors. The antibody reportedly recognises methyl-PABP1 but exact details of how the antibody was raised or evidence of its selectivity were absent, so no comment can be made on the relevance of the finding that **43** is able to prevent methylation of PABP1.



43

***Figure 1.19** – Within this series, position and extent of bromination is thought to be key to the selectivity between the PRMTs. **43** is a highly potent and selective CARM1 inhibitor.*

43, a symmetrical brominated bis-aryl molecule, was also tested against other PRMT family members and it is thought that the position and extent of bromination on the aromatic rings is integral to the selectivity between different PRMTs. The observations and SAR inferences were based on use of a radiometric assay using CT_3-SAM that relies on separation of products by SDS-PAGE[121], followed by excision of protein bands and quantification of methylation by counting the number of decays per minute (**Section 1.4.1.a**). The reliability of this technique is

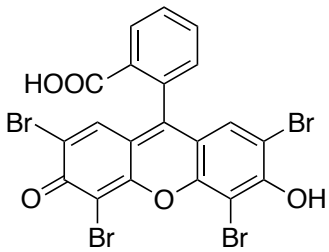
likely to be insufficient for making SAR observations that could have a profound impact on future generations of inhibitors.

1.5.1.b – Dye-Like Small Molecule Development

Noticing that the AMI series were dye-like derivatives, a small focused library of dye-like small molecules was assembled and screened, using a radiometric assay, against human PRMT1 and the putative model protein RmtA, a fungal Arg methyltransferase from *A. nidulans* sharing 56% sequence identity with human PRMT1[108]. The top hit from this focused screen was **44.**

Inhibition data correlated well between the two enzymes, therefore a homology structure of RmtA was composed using three crystal structures of PRMT1-homologous proteins (**Figure 1.20**) for docking work. AutoDock was used for *in silico* binding studies; top hits from this study and molecules from the AMI series were docked into a hypothesised binding pocket based on the reported[99] non-competitive binding mode of AMI-1 **33** (*vs.* PRMT1).

PDB code	Protein	Cofactor	Sequence identity[b] with human PRMT1
1OR8	rat PRMT1	SAH	100%
1G6Q	yeast RMT1/Hmt1	none	52%
1F3L	rat PRMT3	SAH	50%



44

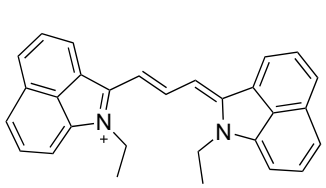
***Figure 1.20** – Table of homology between PRMT1 and model proteins; Dye-like inhibitor **44** identified from one of the first targeted screening approaches using a PRMT1 homology model.*

[b] Sequence identity calculated using a BLAST algorithm accessed at http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins

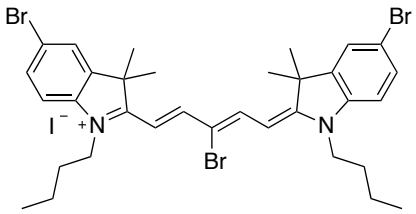
Analysis of the docking results from this work prompted classification of inhibitors into three categories based on their calculated binding mode: docked in the Arg pocket, docked in the SAM pocket or docked in both pockets. The majority of small molecule PRMT inhibitors are thought to bind in one of the two pockets; **Section 1.5.3** further discusses those compounds that are designed to bind to both pockets at once. The validity of such a homology model however is likely to be limited and the only way to unequivocally verify these claims would be to obtain a co-crystal structure of the purported inhibitor with PRMT1.

In an attempt to exploit the dye-like nature of the predominant PRMT inhibitors, work to identify a fluorescent dye that could be used as an imaging agent generated carbocyanine **45** as a potential lead (**Figure 1.21**)[122]. The molecule was docked into a human PRMT1 homology model (modelled on rat PRMT3) using AutoDock and a possible allosteric mode of binding was suggested. Amino acid residues thought to be key for protein interaction with the inhibitors were identified as Y148, T81, T292 and Y152. SDM was used to generate four mutant proteins where each of these residues in turn was altered to an alanine and the top inhibitors tested against the mutants. From this study, all residues except Y152 are thought to be important for the mechanism of inhibitor binding. Despite being initially based on docking work, the molecular biology experiments using mutant proteins brings more authenticity to the results.

Very recently this work was developed and **46** was reported as a more selective agent for inhibiting PRMT1 over CARM1, PRMT5 and PRMT8. This work was carried out using the radiometric assay[123].



45

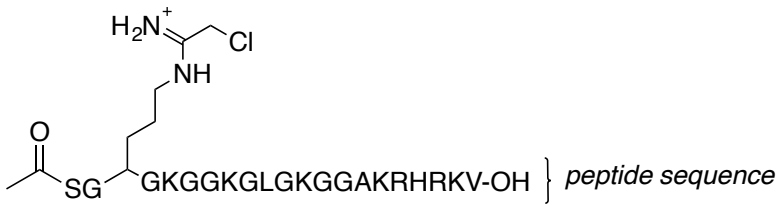


46

Figure 1.21 – *Cyanine dye-like PRMT inhibitors.*

1.5.3 – Peptidic and Bisubstrate Inhibitors

In an attempt to develop inhibitors that were selective to individual PRMT family members, a series of 'bisubstrate' compounds that made use of both the SAM and substrate binding sites were synthesised[124]. Follow-up work on this series identified the peptidic chloroacetamide **47** (**Figure 1.22**) as a PRMT1 inhibitor showing selectivity for PRMT1 (IC_{50} = 1.8 ± 0.1 μM) and PRMT6 (IC_{50} = 8.8 ± 0.5 μM) over PRMT3 and CARM1 (both IC_{50} = >500 μM)[125]. These results were generated using an SDS-PAGE assay (**Section 1.4.1.a**) with ^{14}C-radiolabelled SAM. The bands were quantified by 'Phosphorimager' analysis (Molecular Dynamica Inc.) – a type of densitometry that relies on laser-optical imaging; the IC_{50} value had relatively small associated error, although it is noted that experiments were usually carried out in duplicate.



47

Figure 1.22 – 47 *is a PRMT1 and PRMT6-selective peptide inhibitor.*

Inhibitor **47** was subsequently used as a tool to probe PRMT1 substrate recognition in a combinatorial approach[126].

Four novel ethylated and fluorinated Args, **48-51** (**Figure 1.23**), were incorporated into short (12-mer) peptides for testing as PRMT substrates and inhibitors[103]. Using the UPLC MS assay described in **Section 1.4.1.c**, it was found that **48** was methylated by PRMT1 and that all four peptides inhibited PRMT1 and PRMT6, plus CARM1 to a lesser extent. Building on these promising results, the authors developed a second series of peptidic bisubstrate compounds with the intention of incorporating a cofactor-analogue to increase potency[127]. Due to the perceived difficulty in appending a bulky adenosine group, it was instead decided to add an ornithine- (Orn) or Lys-like tail to the guanidine of the peptidic Arg (**52** and **53**). The length of these Orn- and Lys-like fragments mimic the size of the linker region that would be found if true SAM/SAH analogues were used. The Orn analogue **52** did not inhibit PRMT1 or CARM1 but did inhibit PRMT6. The Lys analogue **53** inhibited all three tested enzymes and was most efficacious *vs.* PRMT1.

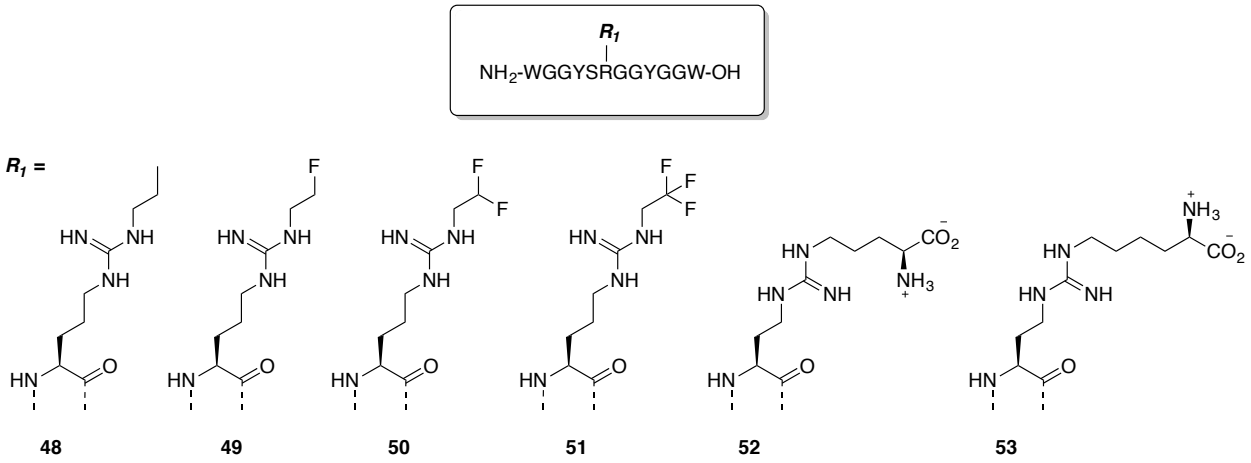


***Figure 1.23 – 48-51** inhibit PRMT1 and PRMT6[103]. **52** and **53** have appended Orn- and Lys-like chains respectively, which mimic the linker region of cofactor SAM[127]*

As part of a targeted effort to discriminately inhibit different PRMTs using bisubstrate compounds, Dowden *et al.*[128] synthesised non-peptidic true adenosine-containing cofactor analogues and appended substrate Arg-like fragments (**Figure 1.24**). Using molecular modelling

for assistance, this study found that varying the number of methylene linkers in **54** between the SAM-like component and the Arg-like component were crucial for inhibitor selectivity between PRMT1 and CARM1[128]. These results seem plausible, but it should be noted that the assay used to make these SAR observations was radiometric and relied upon excision and scintillation counting of SDS-PAGE bands (**Section 1.4.1.a**).

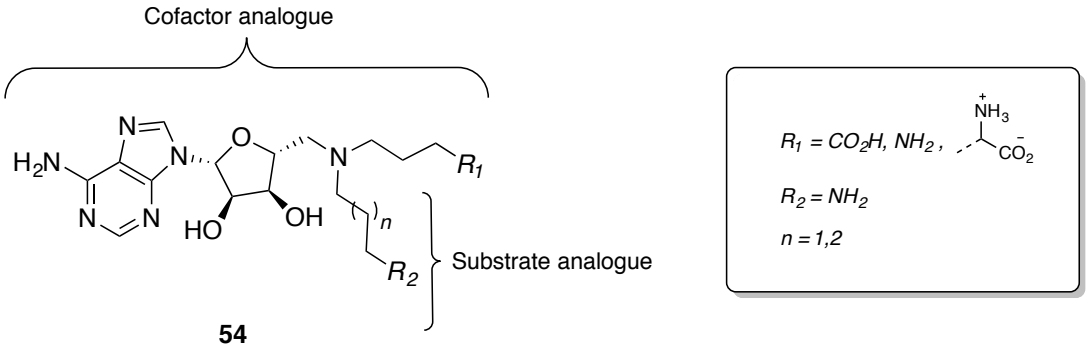


***Figure 1.24** – Inhibitors produced by Dowden* et al.[128].

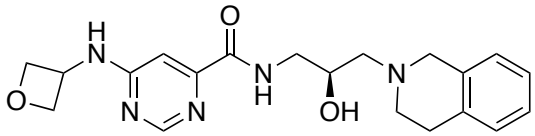
With this proof-of-concept for small molecule bisubstrates as inhibitors, and building on the promising results in **Figure 1.23**, the Martin research group turned their attention to small molecule bisubstrate inhibitors. They found success in developing potent PRMT inhibitors and also generating SAR data that points to mechanisms of inter-PRMT selectivity[129].

1.5.4 – Inhibition of Type II and III PRMTs

The majority of PRMT inhibitors reported are active against type I enzymes. A highly potent (IC_{50} = 22 nM) PRMT5 inhibitor **55** (**Figure 1.25**) was recently described – the best type II PRMT inhibitor to date[130]. **55** is reported to be highly selective over other PRMTs, have excellent oral bioavailability and have anti-proliferative effects *in vitro* and *in vivo* against MCL phenotypes[130]. Inhibition data was generated using antibody-based detection of monomethyl-Arg and a TRF assay. The antibody was purchased from Abcam although the exact product was not

defined in the original report; at the time of printing, the online database cataloguing antibody reliability[102] contained one entry that could correspond to the antibody used in this study *vs* MMA(ω) and that had been passed as a reliable tool by dot-blot analysis.

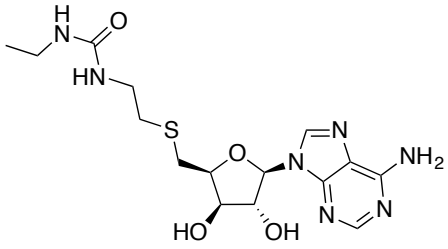
In any case, the authors use an orthogonal assay (radiometric) for hit compounds in an effort to rule out any false positives. Furthermore they tested the top hit in cell-based assays that showed good correlation with the recombinant PRMT results. Surface-plasmon resonance studies strongly supported a true binding event between compound **55** and PRMT5:MEP50. Finally, an X-ray co-crystal structure was obtained that added further proof for this interaction. Overall, the study was extremely thorough and there is little reasonable doubt that **55** is a true PRMT5 inhibitor and shows promise for future development.



55

***Figure 1.25** – The most potent PRMT5 inhibitor that has been reported to date.*

A second recent discovery, from the Structural Genomics Consortium, is of the dual PRMT5-PRMT7 inhibitor **56** (**Figure 1.26**)[131]. PRMT7, a type III enzyme, can only ω-monomethylate Arg and one widely discussed theory is that PRMT5 and PRMT7 could act in tandem in the cell to generate an SDMA methylation pattern. With this in mind, it is of particular interest that **56** selectively inhibits both PRMT5 and PRMT7 but not other PRMTs or PKMTs in the 27-member panel that was screened using a radiometric assay.



56

Figure 1.26 – *The first reported dual PRMT5-PRMT7 inhibitor.*

1.5.5 – Allosteric PRMT Inhibition

All hitherto mentioned inhibitors are thought to occupy either the cofactor binding site, the substrate binding site or both at once. The first proposition of an allosteric binding mode for **45** in PRMT3 came from homology modelling[122].

The first example of a confirmed allosteric inhibitor was reported in 2012; Compound **57** was shown to inhibit PRMT3 in a non-competitive manner with respect to SAM **2** or substrate (**Figure 1.27**)[132]. Barring very weak inhibition of PRMT5, **57** is reported completely selective for PRMT3 over a panel of four PKMTs, PRMT1, CARM1, PRMT5 and PRMT8 (determined by radiometric assay). The co-crystal structure of PRMT3 with **57** clearly revealed binding at the dimer interface of the protein. This work provides proof-of-concept that allosteric inhibition could be a successful way to selectively inhibit PRMTs; allosteric sites could allow researchers to work in non-conserved PRMT chemical space and exploit differences between family members. Indeed, potential allosteric binding sites have been identified by the aforementioned study in PRMT1 and CARM1. **57** had a reported IC_{50} of 2.5 ± 0.1 μM; building on this work, the authors gathered SAR data using a radiometric assay leading to the apparently more potent PRMT3 inhibitor **58 (**IC_{50} = 0.23 ± 0.03 μM)[133]. The crystal structural information reported for this work (PDB: 3SMQ[132]) lends weight to the argument that **57** and **58** are true inhibitors. **58**

had the added benefit of being completely inactive against PRMT5, unlike its predecessor **57**. A brief summary of the SAR indications for this series is given below in **Figure 1.27**.

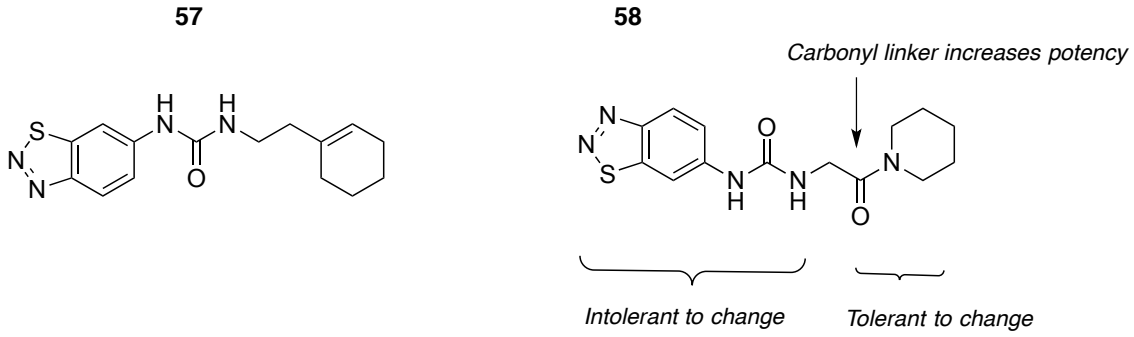


***Figure 1.27** – Initial top hit[132] **57** vs. PRMT3 with an IC_{50} of 2.5 ± 0.1 μM and top hit **58** from the follow-up study[133] with an IC_{50} of 0.23 ± 0.03 μM shown with SAR indications realised from the series analysis.*

1.5.6 – Summary of Currently Proposed Inhibition Modes

The inhibitors reviewed in this section can be broadly categorised as competitive with cofactor and/or substrate, uncompetitive with cofactor and/or substrate, or allosteric. Prominent inhibitors reviewed above and their proposed mechanisms of action are summarised in **Appendix A**.

1.6 – Aims of This Research

PRMTs are major epigenetic regulators and are implicated in a large number of disease phenotypes. Targeting the epigenome for therapeutic purposes is a research-intensive field; success in developing epigenetic drugs is aided by good understanding of the underlying biology and reliable screening and testing methods for putative drugs.

This thesis describes a multi-disciplinary project towards the following broad aims:

- Widen biological understanding of arginine methylation patterns
- Develop a potent PRMT inhibitor

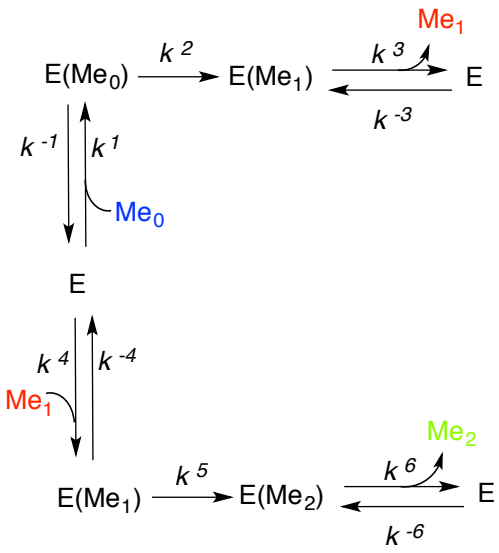
The following major hypotheses are proposed:

1.6.1 – Hypothesis One – Concept of Alternative Methylation Patterns

Arg residues with a methylation number higher than two could exist and could act as inhibitors and/or substrates when incorporated into peptides. Currently, the only way Arg methylation patterns have been identified is from isolating proteins from cells and hydrolysing them to release the component amino acids, discussed in **Section 1.2.1**. Arg residues can then be assessed for methylation using fragmentation and HPLC analyses. However, in the absence of any standards for novel methylation patterns, it is not possible to conclusively state that no such higher order methylation might exist. Furthermore, this thesis begins to look at the possibility that known (and by extension any unknown) methylation patterns might be capable of inter-converting; this could challenge our perception of the histone code.

1.6.2 – Hypothesis Two – Current Assay Techniques are Mostly Inadequate

Following the current model of mono- and di-methylation by PRMTs (**Figure 1.5**) a distribution of methylated products is to be expected in any PRMT-catalysed methylation reaction, the kinetics of which are depicted in **Scheme 1.11**.



Scheme 1.11 *– Kinetics of a type I or II PRMT-catalysed reaction.*

It is hypothesised herein that assays that do not allow analysis of this methylation distribution can generate inconsistent inhibitory data. As an extension of this hypothesis, it is postulated herein that novel multi-methylated Args could be produced by PRMTs but that this information is being lost in the absence of metadata that describes substrate/product methylation number. Using an assay that enables distinction between different methylated species, such as a MS-based assay, may be required for consistently accurate and precise inhibition data.

1.6.3 –Specific Aims of this Thesis

With the intention of addressing the two hypotheses presented above, this thesis describes work towards the following specific aims:

1) Exploration of synthetic routes towards novel ω-methylated Arg residues (Chapter Two).
2) Identification of the most appropriate route to protecting Arg amino acids that facilitates their use in peptide synthesis (Chapter Two).

3) Total synthesis of the amino acid MMA(δ) with suitable protection for peptide synthesis (Chapter Three).
4) Peptide synthesis incorporating protected MMA(δ) into a histone sequence (Chapter Three).
5) Design and optimisation of a new MALDI-based assay (Chapter Four).
6) Testing of the novel MMA(δ) peptide for activity *vs*. PRMT1 (Chapter Four).
7) Synthesis of a library of putative PRMT inhibitors based on previous scaffolds in our laboratory (Chapter Five).
8) Implementation of chemiluminescence-, radiometric- and MALDI- based assays to assess inhibitory potency of the aforementioned library (Chapter Five).

Chapter 2: Towards the Synthesis of Bespoke Methylated Arginines

This chapter describes work towards the synthesis of selective ω-methylated Arg for use as building blocks in peptide synthesis.

2.1 – Introduction

In the field of systems biology, 'top-down' and 'bottom-up' are terms to describe the type of approach used to probe a biological system[134]. Broadly speaking, top-down approaches begin with an observation of a system and aim to work back down to understand the components, while bottom-up approaches do the reverse and start with the components of a system that can be used to elucidate higher-order functional relevance.

2.1.1 – Novel Methylated Arginines

Top-down approaches have identified the Arg methylation patterns shown in **Figure 1.5**; human PRMTs are reported to produce MMA(ω) **3**, ADMA **4**, or SDMA **5**[19]. In yeast MMA(δ) **6** is produced by the PRMT family homologue RMT1[31,135,136] but both enzyme and marker are poorly understood, with the downstream effects of MMA(δ) **6** in yeast remaining unreported and no directly homologous protein to RMT1 yet identified in humans.

MMA(δ) **6** has not been identified in humans, but little work has been undertaken to investigate this possible marker; in particular the stability of MMA(δ) **6** is unknown. Furthermore, specific detection methods for methylated Args are limited, as discussed in **Section 1.4**, and it is possible that inadequate assays have failed to identify such a marker.

The concept of 'methylation number' was introduced in the previous chapter. No Arg has been identified with a methylation number greater than two, although there are three possible *N*-methylation positions on guanidine - two ω and one δ (**Figure 2.1A**). The α-position is also shown but this is implicated in peptide bond formation and is not available for methylation.

This work aims to explore the idea that novel methylated Args could be substrates for, or inhibitors of, human PRMTs. **Figure 2.1B** depicts the methylation patterns currently thought to have physiological relevance. The novel methylation patterns under investigation are shown in **Figure 2.1C** and include polymethylated Args **19-23** and the yeast marker MMA(δ) **6**. Using a bottom-up approach, the aim was to synthesise each of these selectively methylated analogues for testing with PRMTs. However, the PRMTs will not catalyse the methylation of free Arg, requiring the amino acid to be part of a peptide/protein to be accepted as a substrate[137]. To test methylated Args for substrate/inhibitory prowess, the amino acids must be incorporated into peptides.

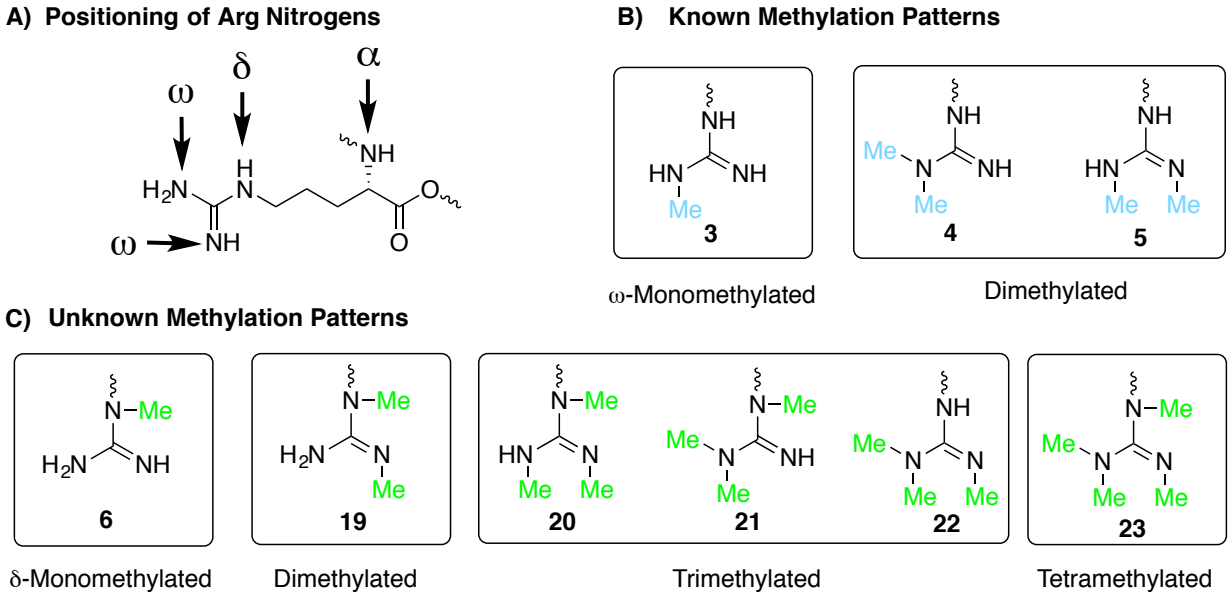


***Figure 2.1** – **A)** Possible methylation positions include the ω and δ - nitrogens, while the α-NH_2 is occupied in peptide bonding. **B)** Known and **C)** novel methylation patterns on Arg termini.*

2.1.3 – Background to Peptide Synthesis

Solid phase peptide synthesis[138] (SPPS) relies on use of a solid support to which the carboxylate group of the *C*-terminal residue is attached. The solid support is typically a polystyrene-based resin with a linker attached[138]. A series of washing-deprotection-washing-coupling steps are performed to polymerise amino acid residues into a peptide (**Figure 2.2**). Each new residue is introduced to the reaction system with its α-amino group Fmoc-protected such that the only coupling that can occur is between the solid-supported nascent chain amino group and the carboxylate group of the newly introduced amino acid; this cycle is repeated until the peptide synthesis is complete, building up from *C*-terminus to *N*-terminus. Once the peptide backbone has been constructed, the *C*-terminus is cleaved from the linker and any remaining side-chain protecting groups are removed to furnish the peptide, typically with acidic conditions.

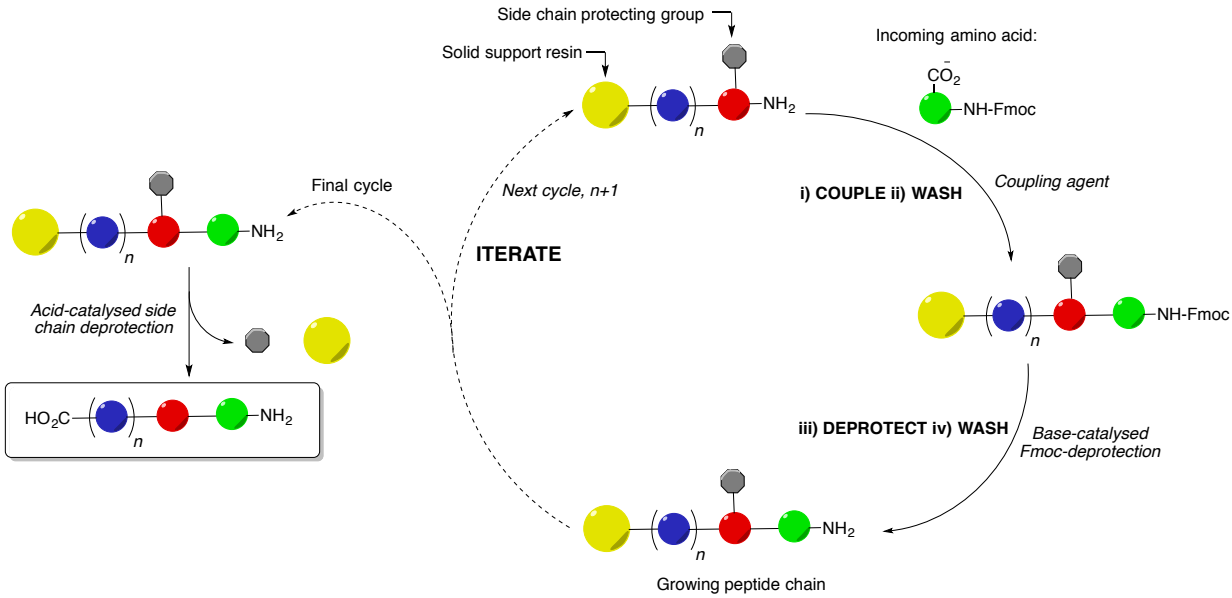
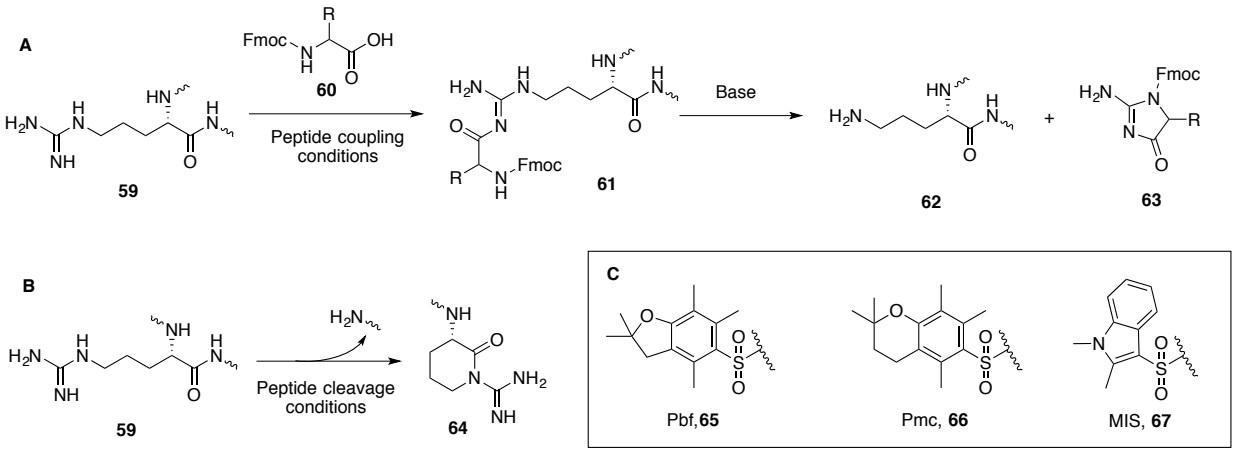


Figure 2.2 *– Solid phase peptide synthesis sequence iteration. Reagents and conditions: i) coupling; activating agent e.g. HOBt/HATU and DIC ii) washing; e.g. DMF iii) deprotection; piperidine iv) washing e.g. DMF.*

Arginine is a special case, requiring a protecting group that masks the strongly basic nature of the side-chain guanidine group (pK_A = 12.5 in H_2O)[139]. Under normal Fmoc-based SPPS conditions, peptidyl-Arg **59** is prone to deguanylation to form peptidyl-Orn **62** *via* acylation of

the guanidine from an incoming amino acid **60** (**Scheme 2.1A**)[140]. Alternatively, under basic conditions such as those experienced when deprotecting Fmoc-groups, the guanidine side-chain of peptidyl-Arg **59** readily cyclises to form the δ-lactam **64** (**Scheme 2.1B**)[141]. Much research has gone into developing a good protecting group for the guanidine of Arg, and currently the 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) **65** group is favoured for its commercial availability and resilience to SPPS conditions and moderately facile deprotection under acidic conditions (**Scheme 2.1C**)[141]. Other sulfonyl protecting groups are available but are either not acid-labile enough (e.g. 'Pmc' **66**)[142] for the final de-protection step or not easily synthesised or not commercially available (e.g. 'MIS' **67**)[143].



Scheme 2.1 – A) *Acylation and subsequent de-guanylation of unprotected peptidyl-Arg* ***59*** *to peptidyl-Orn* ***62*** *during peptide synthesis, adapted from Isidro-Llobet*[141] ***B)*** *Cyclisation of Arg* ***59*** *to δ-lactam* ***64*** ***C)*** *Pbf* ***65****, Pmc* ***66****, MIS* ***67*** *protecting groups for Arg-OH* ***1****.*

The greatest challenge with SPPS is reaction yield. For a peptide of length *n* residues, the overall yield is x^n where *x* is the percentage yield of each coupling step. In order to promote the highest possible overall yielding reaction, each amino acid to be coupled is added in great excess with respect to the nascent peptide. For commercially available amino acids this is often inconsequential, however SPPS efficiency can be project critical for novel bespoke amino acids. In theory it is possible to recover un-reacted amino acids following each individual coupling

step, but in practice this is not a viable or attractive option given the need to remove accompanying contaminants and solvent DMF.

2.2 – Chapter Aims

This chapter and the next discuss work towards the overall aim of synthesising bespoke methylated Arg molecules with α-NH_2 Fmoc protection and suitable ω-guanidine Pbf protection on a preparative scale (~1 mmol) which facilitates their incorporation into peptides *via* SPPS. Reported in the current chapter is work towards the following specific aims:

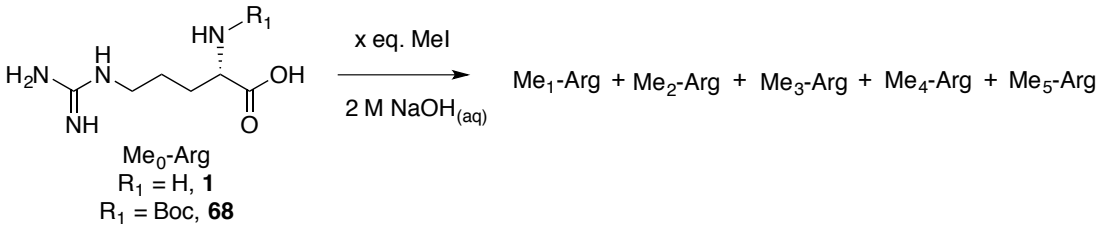
i. Identification of suitable and high-yielding protecting group strategies for Orn **69** and Arg **1**
ii. Development of an optimised method for converting Orn **69** to Arg **1** *via* guanylation to facilitate incorporation of different ω-*N*-methylation patterns

Synthesis of MMA(δ) **6** is discussed separately in Chapter Three.

2.3 – Strategy for Synthesis of Protected Arginines for SPPS

2.3.1 – Towards Methylation of Unprotected Arginine

It was anticipated that treatment of readily available Boc-protected and unprotected Arg derivatives with a methylating agent, would generate a mixture of different products, which while inefficient with respect to yield, may prove separable. Investigations were made as to the regioselectivity of the reaction between L-Arg **1** and the methylating agent iodomethane (**Scheme 2.2**).



			Apparent distribution methylated species (LC/MS analysis)[c]					
Entry	**R_1**	**Eq. MeI**	**Me_0-Arg**	**Me_1-Arg**	**Me_2-Arg**	**Me_3-Arg**	**Me_4-Arg**	**Me_5-Arg**
Arg-OH, **1**	H	1	32%,	29%	17%	19%	3%	0.1%
Arg(Boc)-OH, **68**	Boc	1	The different methylated compounds could not be resolved by LC/MS					
Arg-OH, **1**	H	0.1	60%,	29%	8%	3%	<0.1%	negligible
Arg(Boc)-OH, **68**	Boc	0.1	The different methylated compounds could not be resolved by LC/MS					

Scheme 2.2 *– Reaction between free Arg-OH **1** or Arg(Boc)-OH **68** and iodomethane generates a mixture of methylated products. Reagents and conditions: as stated.*

^{1}H and 2D NMR studies suggest that complex mixtures of *N*-methylated and *O*-methylated products are present. Accurate mass data from LC/MS chromatographic peaks correspond to molecular formulae consistent with un- (Me_0-Arg), mono- (Me_1-Arg), di- (Me_2-Arg), tri- (Me_3-Arg), tetra- (Me_4-Arg) and penta- (Me_5-Arg) methylated Arg. Regioisomeric Me_1-Arg (MMA(δ) **6** and MMA(ω) **3**) and regioisomeric Me_2-Arg (ADMA **4** and SDMA **5**) are indistinguishable by this method.

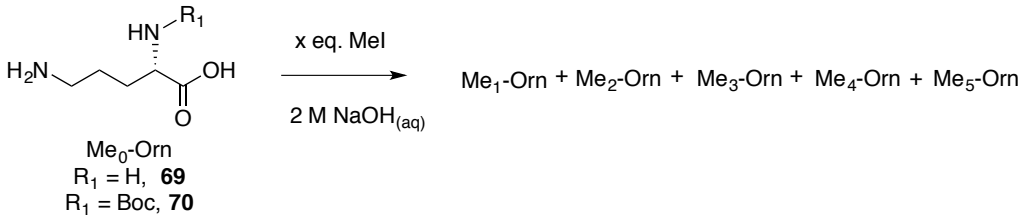
The distribution of methylated products was calculated by totalling the integrated 'area under curve' for each extracted ion mass chromatogram corresponding to Me_0-Me_5 and expressing each methylated Arg as a percentage of the total. The percentage distributions are given in **Scheme 2.2** and the LC/MS traces are given in **Appendix B**. These traces show poor separation

[c] LC/MS analyses were carried out using a 100 mm x 2.1 mm Cogent Diamond Hydride™ column (particle size: 4 μM, pore size: 100 Å) with aqueous-normal-phase elution. Percentage distribution was evaluated from the extracted ion mass chromatogram.

of compounds at analytical scale and complete overlap at preparative scale, suggesting attempts to isolate individual methylated Arg derivatives from a mixture would be an infeasible approach to synthesising bespoke Arg molecules on preparative scale.

It is noted that for experiments using 0.1 eq. methylating agent, the percentage distribution represents an impossible outcome – in theory no more than 10% of the total compounds could truly be methylated. This discrepancy is explainable – the calculations rely upon integration of extracted ion chromatogram peaks and the assumption was made that each methylated Arg would have similar chromatographic properties. However, the data presented herein suggests this to be untrue, meaning the calculated quantities are a misrepresentation of the true distribution of methylated products. In the absence of standard compounds, it is not possible to make an assessment of the actual percentage distribution. However, for the purpose of this experiment, the data undoubtedly shows that a mixture of methylated products exists, and the LC/MS traces in **Appendix B** demonstrate that these are not separable.

The same reactions were carried out using Orn derivatives with the same limitations and the same conclusions (**Scheme 2.3**). Thus, alternative procedures allowing for chemo- and regio-selective methylation and more ready separation of isomers were explored.

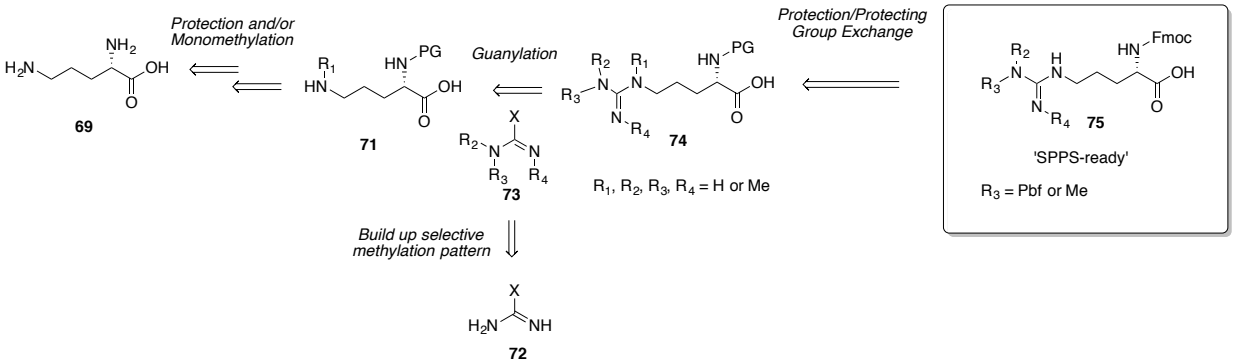


			Distribution methylated species (LC/MS analysis)					
Entry	**R_1**	**Eq. MeI**	**Me_0-Orn**	**Me_1-Orn**	**Me_2-Orn**	**Me_3-Orn**	**Me_4-Orn**	**Me_5-Orn**
Orn-OH **69**	H	1	19%	29%	38%	8%	4%	2%
Orn(Boc)-OH **70**	Boc	1	53%	24%	11%	11%	0%	0%
Orn-OH **69**	H	0.1	51%	39%	8%	1%	0.5%	0.1%
Orn(Boc)-OH **70**	Boc	0.1	41%	22%	14%	23%	0%	0%

***Scheme 2.3** – Reaction between free Orn-OH **69** or Orn(Boc)-OH **70** and iodomethane generates a mixture of methylated products. Reagents and conditions: as stated.*

2.3.2 – Ornithine as a Building Block for Methylated Arginine

Orthogonal protecting groups must first be introduced to enable selective methylation.



***Scheme 2.4** – Planned retrosynthesis of selectively methylated Arg **74** from Orn-OH **69**; PG = protecting group; X = leaving group.*

A disconnection was made at the guanidine terminus of Arg **74** to allow incorporation of methyl groups at a late stage, which results in the most efficient convergent synthesis. Conversion of Orn-OH **69** to protected Arg **74** would be achieved *via* a guanylating agent (**Scheme 2.4)**. To be

suitable for SPPS, the α-NH_2 requires Fmoc-protection and the guanidine terminus must be Pbf-protected **75** unless all of R_2, R_3 and R_4 are methyl groups[d].

In all cases, this chapter discusses reactions where R_1 = H. Reactions using R_1 = Me for the synthesis of MMA(δ) **6** are dealt with in Chapter Three.

2.4 – Purification Strategies

One of the biggest challenges working with amino acids on the scale required for SPPS is the associated purification difficulty due to their high polarity; certainly the LC/MS data presented in **Section 2.3** confirms this. Normal-phase chromatographic purification (SiO_2) that can separate non-polar molecules was useful but only at the limit of operation for fully-protected Orn and Arg, and often with poor separation and yield and occasionally with product degradation. Poor results were also obtained using reverse-phase HPLC techniques and aqueous normal-phase HPLC techniques (*vide supra*).

Use of a cation-exchange resin (a Dowex® column[e]) enables purification of a single amine from a mixture of organics but is not typically used to resolve two different amines; all basic groups are strongly bound to the resin and then eluted under gravity with $NH_4OH_{(aq)}$. As part of this work an ion-exchange purification procedure was developed that is able to separate very polar molecules by using different ionic strength eluents:

[d] It is anticipated that no Pbf protection is required when all ω-positions are methylated since de-guanylation is thought to occur *via* acylation at these positions. (140) Rink, H.; Sieber, P.; Raschdorf, F. *Tetrahedron Lett.* **1984**, *25*, 621..

[e] All references to Dowex® in this thesis refer to the nuclear sulfonic acid strongly cationic ion exchange resin (Dowex-50-H) with mesh size 200-400 and 8% cross-linkage from which products can be eluted in aqueous and methanolic basic solutions. Dowex® was purchased from Sigma-Aldrich.

A report from the 1950's used dilute aqueous ammonium acetate and formate solutions at different buffered pH values to elute individual amino acids on Dowex® resin by exploiting their isoelectric points (pI)[144]. Good separation could be achieved for Lys and Arg; by extension the principle should be applicable to Orn/Arg separation insofar as Orn and Lys are congeneric with similar pI's compared to Arg (**Table 2.1**). The published method was adapted for separation of a 50:50 Orn:Arg standard mixture (detailed in the Experimental Section). It was anticipated that this purification procedure, successful on a scale conducive for SPPS, could also be applied for unprotected methylated Arg derivatives.

Amino Acid	pK_A α-COOH	pK_A α-NH_2	pK_A side chain	pI
Arg	2.02	9.00	12.50	7.84
Lys	2.18	8.95	10.53	7.22
Orn	1.71	8.69	10.76	7.05

Conditions for Optimised Separation:

Resin: 12.5 cm x 2.5 cm

Flow-rate: ~3 mL/min

Orn eluent: 0.2 M $NH_4C_2H_3O_2$, pH 5.5

Arg eluent: 0.5 M $NH_4C_2H_3O_2$, pH 6.8

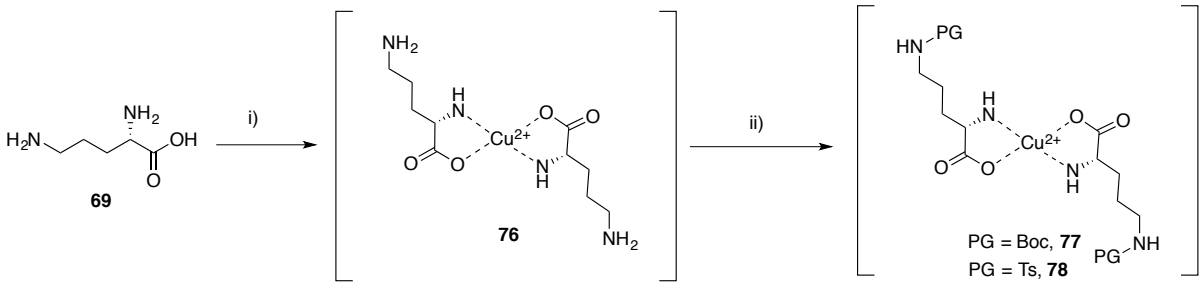
***Table 2.1** – Isoelectric properties of Arg, Lys and Orn and details of ion-exchange purification.*

2.5 – Protecting Group Strategies for Ornithine

2.5.1 – Copper Complexation to Allow Selective δ-Protection of Orn

The use of copper complexation to mask the α-amino and α-carboxylate groups is reported for Lys and Orn[145,146] (**Scheme 2.5**). Complex **76** was formed by addition of a Cu^{2+} source to a solution of Orn-OH **69** in $NaOH_{(aq)}$ (2 M), upon which an instant colour change from teal to dark blue was observed; $CuSO_4$ and $Cu(OAc)_2$ were both tested as the Cu^{2+} source but the latter was found to be more soluble under the reaction conditions, giving a more homogenous solution.

Addition of bulky protecting groups at the accessible δ-nitrogen afforded a light blue precipitate (**77, 78)** that could be isolated by *vacuum* filtration. Any unreacted starting materials were removed in the aqueous flow-through after washing the precipitate with MeOH and an acetone:water solution as reported by Wiejak[146]. It is reasoned that this step is important to avoid low yields and/or complicated mixtures of products because it facilitates removal of unreacted protecting group reagent by washing. Precipitates **77** and **78** were analysed by MS and TLC to confirm successful protection had occurred; the presence of copper made it impossible to gather NMR data.



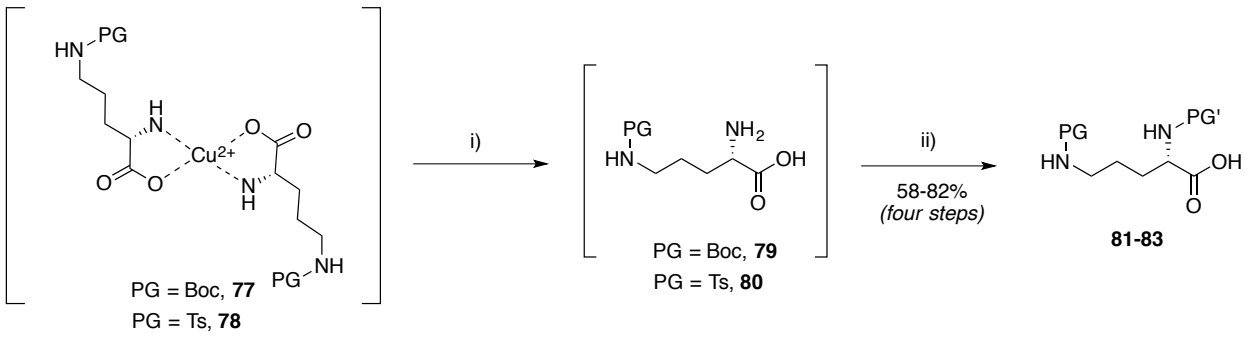
Scheme 2.5: *Reagents and conditions: all reactions at room temperature i) $Cu(OAc)_2$, 2 M $NaOH_{(aq)}$ ii) Boc_2O or Fmoc-Cl & N-hydroxysuccinimide or Ts-Cl & Et_3N. Procedure adapted from Wiejak et. al*[146].

2.5.2 – Copper Chelation and α-Protection

Removal of copper from the reaction mixture was achieved using a chelating agent. Sodium ethylenediaminetetraacetic acid (EDTA)[147] or H_2S[148] can be used for this purpose, but 8-quinolinol was found to be preferential[146]; test reactions using **77** suggested 8-quinolinol was more facile to remove than EDTA *via* acid-base extraction during work-up.

A one-pot synthesis was employed that bypasses isolation of mono-δ-protected derivatives of Orn (**79**, **80**) and achieves chelation and α-protection (**Scheme 2.6**). The copper precipitate (**77** or **78**) was treated with 8-quinolinol for 1 h and then the second protecting group introduced:

Fmoc protection was achieved using Fmoc-Cl and *N*-hydroxysuccinimide for 16 h, Cbz protection was achieved using benzyl chloroformate and *N*-hydroxysuccinimide for 16 h and Boc protection achieved using di-*tert*-butyl dicarbonate for 72 h. After this time, the copper:8-quinolinol chelate was removed by *vacuum* filtration and the filtrate concentrated to remove acetone. Dual α-,δ-protected Orn derivatives **81-83** could be isolated cleanly from the residual aqueous solution *via* standard acid-base extraction methods. High yields (58-82%) and multi-gram scale could be achieved using this route, without the need for further purification.



Compound	α-Protection (PG')	δ-Protection (PG)	Amount, g	Yield, %
Boc-Orn(Fmoc)-OH, **81**	Fmoc	Boc	23.5	58
Boc-Orn(Cbz)-OH, **82**	Cbz	Boc	30.1	82
Ts-Orn(Boc)-OH, **83**	Boc	Ts	24.6	64

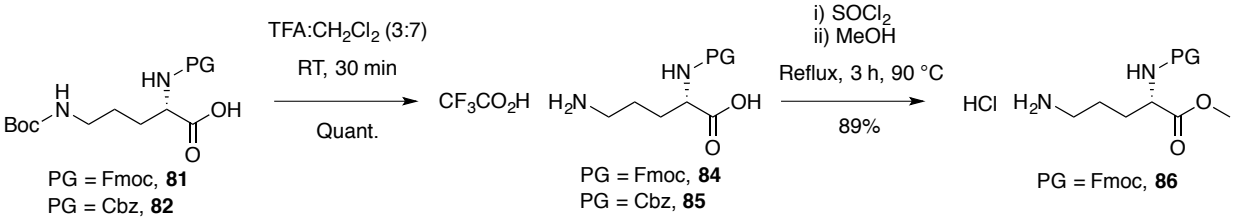
Scheme 2.6 *– Route towards α-protection from the δ-protected Orn derivatives **79**, **80**. Reagents and conditions: All reactions at room temperature unless otherwise stated: i) 8-quinolinol, 1 h ii) Boc_2O or Fmoc-Cl &* N-*hydroxysuccinimide or benzyl chloroformate &* N-*hydroxysuccinimide 16-72 h.*

81 and **82** were both synthesised on approximately 2.2 times the scale previously reported; **81** was isolated in lower yield (58% *vs.* reported[146] 80%) and **82** in identical[146] yield (82%). The novel *N*-tosylated derivative **83** was isolated in good yield (64%) and subsequent reactions of this derivative are discussed in Chapter Three.

Although α-Fmoc-protection was the least efficient transformation, it is a prerequisite for SPPS. Therefore it was decided to protect with Fmoc early in the synthetic route to minimise loss of novel high value derivatives that would otherwise require a step inefficient protecting group exchange for SPPS.

2.5.3 - δ-Deprotection

Orthogonally protected **81** and **82** were δ-Boc-deprotected in quantitative yield by treatment with trifluoroacetic acid (TFA), as required, to expose the free δ-amino group for guanylation (**Scheme 2.7**).



***Scheme 2.7** – δ-Boc deprotection and methyl esterification; Reagents and conditions: i) TFA:CH_2Cl_2 (3:7) RT, 30 min ii) $SOCl_2$ 30 min, RT iii) MeOH, 90 °C, 3 h.*

To improve solubility in organic solvents and thus ease of handling, Orn(Fmoc)-OH **84** was converted to the methyl ester Orn(Fmoc)-OMe **86.**

2.6 – Guanylating Agent Investigations

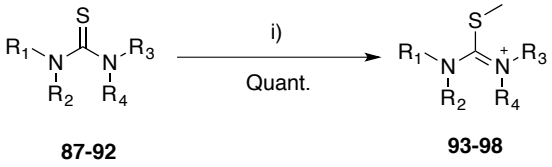
Lots of different guanylating agents have been reported[149] in the literature, however for the most part they have been employed in less polar systems. Towards the goal of this project, it would be preferable if a series of methylated guanylating agent could be employed as part of an optimised general procedure towards accessing all known and novel methylated Arg molecules, as this

would allow the requisite ω-methylation number and regiochemistry to be incorporated in a single step.

2.6.1 – Use of Isothioureas

Thioureas are reported to guanylate primary[150], secondary[150] and tertiary[151] amines in good yields but using polar non-protic solvents, rendering them potentially unsuitable for this work. However, selectively methylated thioureas are commercially available, which could assist towards the goal of this project and so alternative procedures utilising these guanylating reagents were researched. Their derivatives, the *S*-methylisothioureas, are reported to be applicable in more polar systems[149] and could be synthesised in one step from the precursor thiourea using iodomethane. Moreover, a precedent exists for their use in guanylating the methyl ester of α-*N*-protected Lys[152]. All commercially available *S*-methylisothioureas were synthesised in quantitative yield from the corresponding thioureas (**Scheme 2.8**), except expensive **89**, which was not attempted prior to subsequent proof of concept for downstream reactions.

Entry	SM	R_1	R_2	R_3	R_4	Product	Yield, %
1	**87**	H	H	H	H	**93**	Quant.
2	**88**	Me	H	H	H	**94**	Quant.
3	**89**	Me	Me	H	H	**95**	-
4	**90**	Me	H	Me	H	**96**	Quant.
5	**91**	Me	Me	H	Me	**97**	Quant.
6	**92**	Me	Me	Me	Me	**98**	Quant.



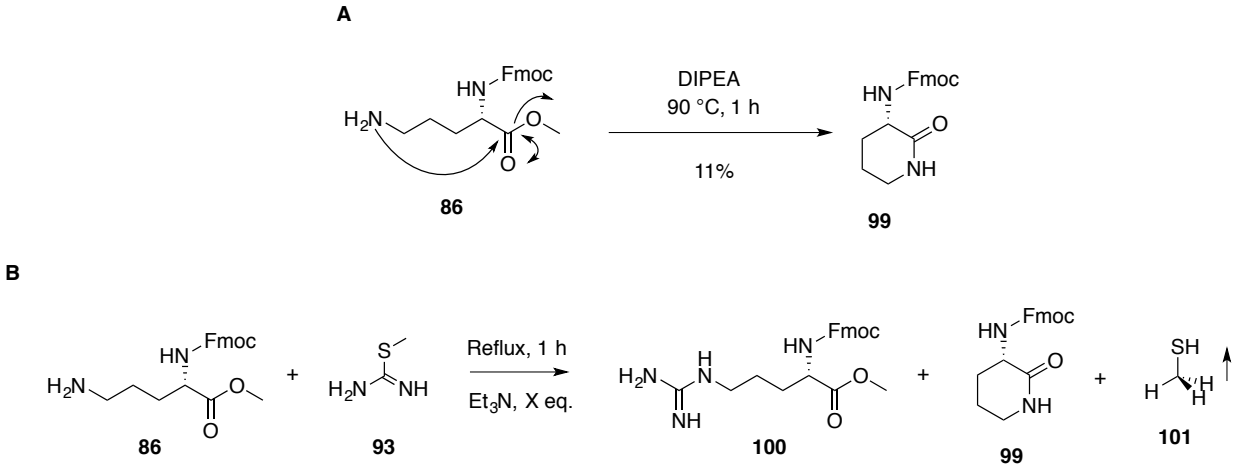
87-92 **93-98**

Scheme 2.8 *– Thioureas* ***87-92*** *and conversion to* S*-methylisothioureas* ***93-98****. Reagents and conditions: i) MeI, 120 °C, 40 min, acetone. SM = starting material.*

The *S*-methylisothioureas were confirmed to be the *S*-methylated and not the *N*-methylated products (where relevant) by comparison of ^{1}H NMR spectra with literature standards where

possible. Furthermore, comparison of the ^{1}H and ^{13}C NMR spectra of *N*-methylthiourea **88** and putative *S*-methylisothiourea **93** indicated diagnostic *N*- and *S*- methyl peaks; this information was extrapolated to identify distinctive chemical shifts in the remaining methylated products of these reactions.

Attempts were made to guanylate Orn **86** using *S*-methyliosthourea **93** (**Scheme 2.9B**), but the Orn methyl esters were vulnerable to cyclisation under the tested conditions forming the δ-lactam **99** - a phenomenon reported for some Orn derivatives[153]. A standard of δ-lactam **99** was therefore synthesised (**Scheme 2.9A**) by heating **86** with DIPEA (11% yield); the extent of cyclisation in the attempted guanylations could therefore be quantified by ^{1}H NMR.



Entry	Eq. 86	Eq. 93	Base	Eq. Base	Temp, °C	SM (86), %	PM (100), %	CP (99), %	CP (MS evidence)
1	1	2	Et_3N	1	140	Deg.	Deg.	Deg.	Y
2	1	2	Et_3N	1	90	Deg.	Deg.	Deg.	Y
3	1	2	none	0	90	100	0	0	N
4	1	2	none	0	140	45	25	25	Y

***Scheme 2.9** – **A)** Synthesis of δ-lactam **99 B)** Attempted guanylation of **86** to **100** using* S*-methylisothiourea **93**. Reagents and conditions: i) MeOH, DIPEA, 90 °C, 1 h ii) All reactions in MeOH - conditions stated in table. Reaction outcome assessed by ^{1}H NMR analyses of crude mixtures and expressed as percentages of starting material (SM) **86**, product material (PM) **100** and cyclised by-product (CP) **99**. Reactions where degradation occurred are denoted by 'Deg.'.*

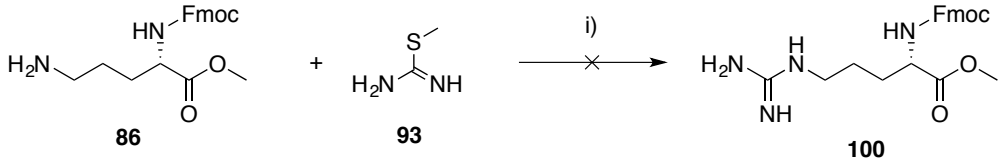
Entry one (**Scheme 2.9B**) shows an attempt to form Arg(Fmoc)-OMe **100** by heating at 140 °C with base. Unfortunately a complex mixture was obtained that contained no distinguishable protons in the ^{1}H NMR spectrum of the crude reaction mixture. However, evidence for cyclisation was observed by TLC and MS analyses. It was postulated that the high temperature might be causing degradation of one or more of the compounds in the mixture, which could explain the complex ^{1}H NMR spectrum. It was postulated that these complex mixtures are a result of Fmoc cleavage occurring under the reaction conditions; this suggestion is tentatively supported by MS analyses.

The temperature was lowered to 90 °C in an effort to reduce this supposed degradation but to no effect – the exact same conclusions were drawn for entry two as for entry one (**Scheme 2.9B**). Maintaining this lower temperature, the base was removed from the reaction (entry three) but ^{1}H NMR analysis of the crude mixture showed no reaction had taken place and 100% starting material remained. In an attempt to push this reaction, the temperature was raised back to 140 °C in the absence of base (entry four). Here, a mixture of starting material **86**, product **100** and cyclised material **99** could be distinguished by ^{1}H NMR analysis of the crude reaction mixture. Synthesis of the cyclised material **99** under these conditions was an unexpected result; it was thought that base would be essential for formation of the δ-lactam **99** but the result presented in entry four suggests that high temperatures are sufficient for cyclisation in the absence of base.

These unsatisfactory results led to investigation of alternative procedures that utilise the *S*-methylisothioureas.

2.6.1.a – Attempts to Accelerate Reaction With $AgNO_3$

Mercuric salts (notably $HgCl_2$ and $Hg(ClO_4)_2$) can be used in stoichiometric quantities to significantly improve guanylation yields and avoid heating by promoting de-sulfurisation[149], but have undesirable safety properties. Successful use of $AgNO_3$ as an alternative Lewis acid is reported for room temperature reactions[154]. Attempts to implement this methodology with Orn(Fmoc)-OMe **86** had limited success (**Scheme 2.10**) affording a complex mixture of products that could not be separated using normal-phase or ion-exchange chromatography.



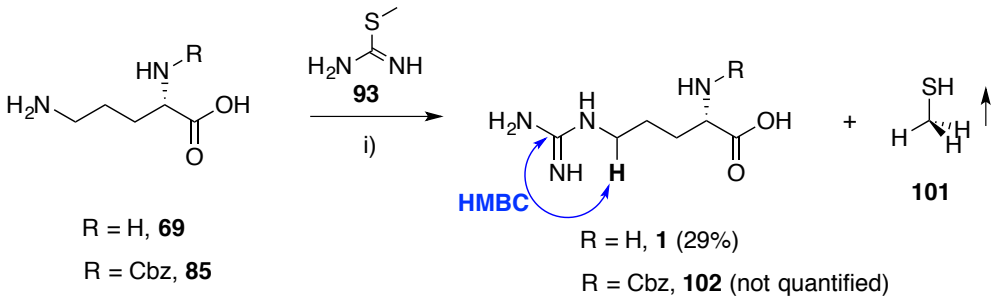
***Scheme 2.10** – Attempted guanylation of **86** to **100** using $AgNO_3$ as Lewis acid to facilitate desulfurisation. Reagents and conditions: i) $AgNO_3$, Et_3N, $MeCN/CH_2Cl_2$, RT, 4 h in dark.*

2.6.1.b – Alternative α-Protection

It was thought that preventing the cyclisation reaction and avoiding postulated Fmoc deprotection (*vide supra*) would ameliorate the analysis of crude reaction mixtures, alongside the obvious advantages of encouraging the desired reaction. To prevent the observed cyclisation (**Scheme 2.9**) carboxylic acid derivatives of Orn were used rather than the methyl ester derivative **86**. The postulated Fmoc lability was addressed by exploring two new approaches – use of base-stable α-Cbz protection **85** and use of free amine **69**; both options also allowed use of a stronger base. It was planned that any resultant products would be subsequently converted to the Fmoc-protected derivative for SPPS.

A report from the early 1990's suggested that guanylation with *S*-methylisothiourea **93** could be achieved regioselectively at the δ-nitrogen of unprotected Orn-OH **69**[155]. However, data from

Section 2.3.1 suggests methylation of unprotected Orn-OH **69** is not regioselective and it was therefore anticipated that guanylation might follow the same pattern.



Entry	R	Yield, %	Purity, %	Side products (assessed by ^{1}H NMR)
1, **1**	H	29	91	Orn-OH **69**, 4%. Thiourea **93**, 4%
2, **102**	Cbz	n/a	n/a	Could not be purified. Quantification not possible in the absence of standards

Scheme 2.11 *– Reagents and conditions: i) 2 M $NaOH_{(aq.)}$, 100 °C, 5 h.*

Orn-OH **69** and Orn(Cbz)-OH **85** were treated with *S*-methylisothiourea **93** in $NaOH_{(aq)}$ (2 M) with heating at 100 °C (**Scheme 2.11**).

For the reaction with Orn(Cbz)-OH **85**, ^{13}C NMR analysis of the crude reaction mixture showed that guanylation had been successful; a quaternary centre correlating to the guanidine carbon was observed. It was anticipated that conversion could be estimated from the ^{13}C NMR spectrum of the crude mixture. It was reasoned that the relaxation properties of the β-CH_2 carbon atoms of each of Orn and Arg would be very similar and that the integration of these peaks could therefore reasonably be compared. Unfortunately, in practice this was not possible and individual chemical shifts could not be resolved, meaning no estimation of conversion was possible. Purification was attempted *via* SiO_2 chromatography but neither product **102** nor starting material **85** could be isolated. This approach, using Orn(Cbz)-OH **85**, was not further pursued.

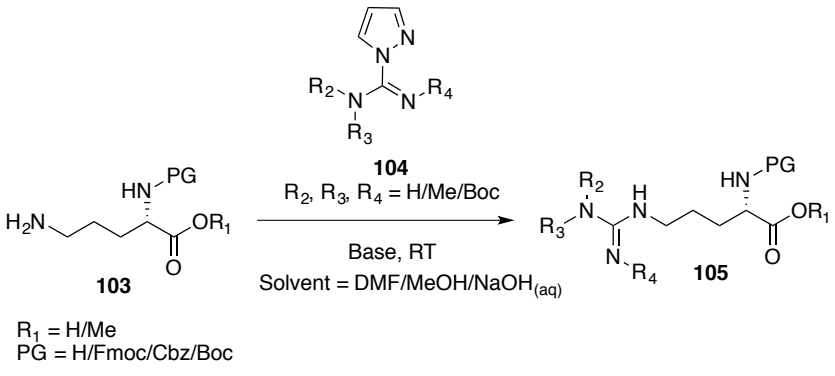
The analogous reaction using Orn-OH **69** proved to be more fruitful. Gratifyingly, the guanylation of unprotected Orn-OH **69** was found to be completely regioselective for the δ-NH_2 over the α-NH_2 (as observed previously[155]) producing Arg-OH **1** and confirmed to be the desired regioisomer by HMBC NMR (see Experimental Chapter). Purification was possible *via* selective Dowex® elution, yielding 29% Arg-OH **1** (91% purity) from the reaction mixture. Small impurities of **93** (4%) and starting material **69** (4%) were observed.

A higher-yielding approach was desired, and the procedures described above offered little room for optimisation since strong base and high temperatures were already in place; unprotected amino acids are susceptible to racemisation at high temperatures[156] so harsher conditions were ruled out.

Attention was therefore turned to a different class of guanylating agent, the pyrazole-1*H*-carboxamidines, which are reputedly successful reagents for guanylation under milder conditions[157].

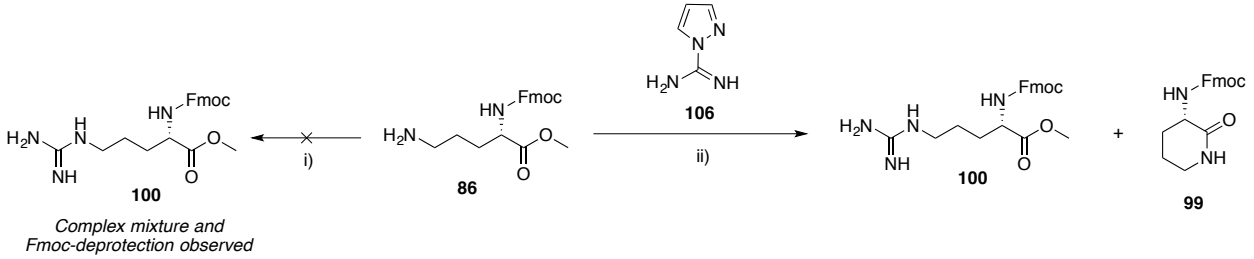
2.6.3 – Use of Pyrazole-1*H*-Carboxamidines

The following general scheme was envisaged towards synthesis of protected and methylated Arg analogues for SPPS using pyrazole-1*H*-carboxamidines (**Scheme 2.12**).



***Scheme 2.12** – General procedure towards synthesis of Arg derivatives **105** from Orn derivatives **103**.*

LiOH was tested as the base, but this was quickly ruled out following preliminary investigations showing this readily removed the Fmoc group (**Scheme 2.13**). A range of conditions were then tested for the guanylation of Orn(Fmoc)-OMe **86** with commercially available pyrazole **106** (**Scheme 2.13**). Unlike for the *S*-methylisothiourea test reactions (**Scheme 2.9**), degradation was not observed, which was attributed to the milder conditions used.



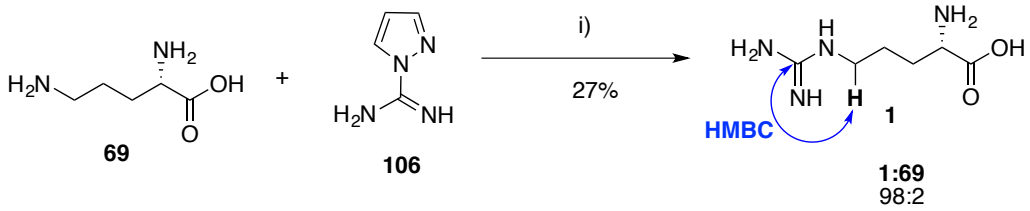
Entry	Orn(Fmoc)-OMe, 86	Eq. 106	Solvent	Base	Eq. base	Temp, °C	Time, h	SM 86, %	PM 100, %	CP 99, %
1	1	1.1	DMF	DIPEA	2.2	25	16	75	0	25
2	1	1.1	DMF	none	0	25	16	75	25	0
3	1	1.1	DMF	DIPEA	2.2	50	16	5	30	65
4	1	1.1	DMF	none	0	50	16	60	40	0
5	1	1.1	MeOH	DIPEA	2.2	25	16	0	65	35

Scheme 2.13 – *Attempted guanylations with pyrazole **106**. Reagents and conditions: i) LiOH, H_2O, 80 h, RT ii) as stated in table. Reaction outcome assessed by ^1H NMR analyses of crude mixtures, using standards for comparison, as percentages of starting material (SM) **86**, product material (PM)**100** and cyclised by-product (CP)**99**.*

Comparison of entries one and two suggests that use of base has a negative impact on the desired reaction while favouring the cyclisation reaction, as expected. The same conclusion can be drawn by comparing entries three and four; these reactions at the elevated temperature of 50 °C generally have a better outcome with respect to desired product **100**, but absence of base is preferable (entry four) because it disfavours cyclisation. Comparison of entries three and five suggests that use of MeOH is preferable over DMF as a solvent, and that even in the presence of base, a 65% conversion to product **100** can be obtained (entry five).

Attempts were made to purify using Dowex® resin and SiO_2 chromatography but both failed to separate the compounds of interest in all cases, presumably because of their amphipathic nature. Samples of crude reaction mixtures were subjected to LC/MS analysis in an attempt to corroborate conversion rates and identify suitable conditions for separation. Unfortunately, co-elution of starting material **86**, product **100** and cyclised product **99** was observed on analytical scale using aqueous normal-phase separation. This explains the failure to separate these compounds *via* SiO_2 or Dowex® purification and suggests that further optimisation of this reaction would be necessary to facilitate a near 100% conversion rate to avoid purification difficulties.

To overcome this purification challenge, attempts were made to apply the regioselective guanylation procedure (**Scheme 2.11**) to unprotected Orn-OH using the milder reagent pyrazole-1*H*-carboxamidine **106** (**Scheme 2.14**). It was anticipated that this way, the ion-exchange protocol outlined in **Section 2.4** could be used to separate Orn-OH **69** from Arg-OH **1** and circumvent the requirement for a near perfect conversion rate. This purification had been successful in separating Orn-OH **69** and Arg-OH **1** for the analogous procedure using *S*-methylisothiourea as the guanylating agent (**Section 2.6.1.b**).



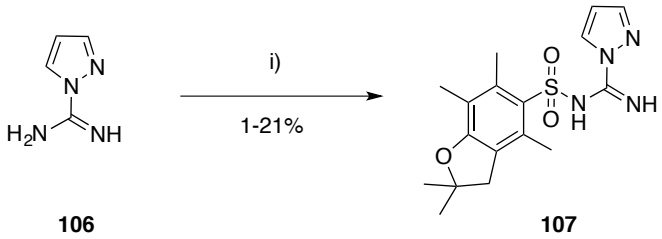
Scheme 2.14 *– Regioselective guanylation of Orn-OH* ***69*** *to Arg-OH* ***1****. Reagents and conditions: i) 8 M* $NaOH_{(aq)}$*, RT, 5 d. Purity assessed by NMR analysis.*

Orn-OH **69** was treated with pyrazole-1*H*-carboxamidine **106** in $NaOH_{(aq)}$ (8 M) to give Arg-OH **1** in 27% yield after successful purification *via* selective Dowex® elution. Attempts were made

to estimate reaction conversion from ^{13}C NMR analysis of the crude mixture, but this was not successful due to overlapping chemical shifts.

Product **1** was confirmed to be the desired regioisomer by HMBC analyses (see Experimental Section). This outcome is a comparable result with the analogous *S*-methylisothiourea reaction that gave a 29% yield (**Scheme 2.11**), although pyrazole **106** offers the advantages of milder conditions and improved purity (**Scheme 2.14**). This class of guanylating agent was therefore selected for further investigation.

Previous reports suggest carbamate-protected guanylating reagents are more reactive than their unprotected parents[158]. Despite not being applicable for accessing all methylation patterns, since ω-protection blocks prospective methylation sites, a protected guanylating reagent could prove extremely useful for producing some of the target Args in higher yield. Boc and Cbz protection are the two reported derivatives, however it would be more useful if Pbf-protection were used since this protecting group is required for SPPS. It was theorised that incorporation of both Pbf and guanidine groups might be achievable using one pyrazole-based reagent. The novel prospective guanylating agent *N*-Pbf-pyrazole-1*H*-carboxamidine **107** was therefore synthesised to test this hypothesis (**Scheme 2.15**). Pyrazole-1*H*-carboxamidine **106** was treated with Pbf-Cl and DIPEA to afford the Pbf-protected pyrazole **107** in 1-21% yield.

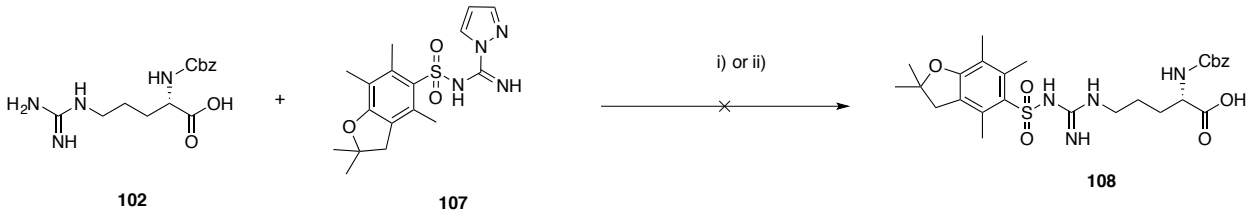


106 **107**

Entry	Solvent	Temperature, °C	Time, h	Eq. Base	Yield, %
1	Acetone	25	16	1.3	1
2	DMF	100	16	5	18
3	DMF	100, MW	1	5	21

Scheme 2.15 – *Synthesis of **107** as a prospective guanylaing agent. Reagents and conditions: i) Pbf-Cl, DIPEA, temperature and reaction time as stated in table. MW = microwave.*

With **107** now available, albeit in low yield, test reactions investigating its utility as a guanylating agent were carried out. Arg(Cbz)-OH **102** was treated with **107** using DIPEA and then $NaOH_{(aq)}$ but regrettably no reaction was observed; ^{1}H NMR analysis of the crude reaction mixture showed only starting material **102** was present (**Scheme 2.16**). It was thought that the decision to avoid this procedure would be the most time-efficient option; optimisation of two poorly yielding reactions (**Scheme 2.15** and **2.16**) was not an attractive prospect for a synthetic route that demanded a high overall yield. Unfortunately attempts to estimate conversion of the reaction *via* ^{1}H NMR analyses of the crude reaction mixtures were unsuccessful; chemical shifts of the starting material **107** and product **108** were overlapping.

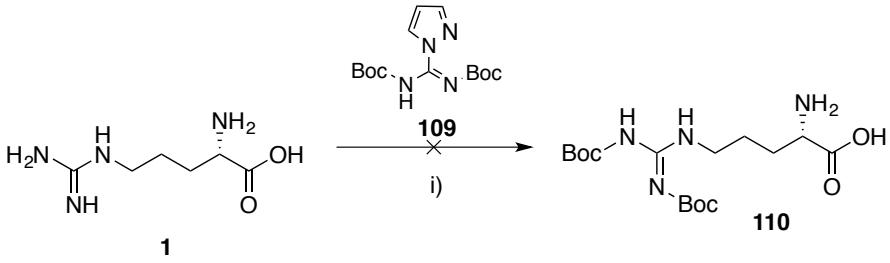


102 **107** **108**

Scheme 2.16 – N-*Pbf-pyrazole-1*H*-carboxamidine **107** did not guanylate Orn(Cbz)-OH **102**. Reagents and conditions: i) MeOH/CH_2Cl_2, DIPEA, RT, 16 h ii) $NaOH_{(aq.)}$ 2 M, RT, 16 h.*

Therefore, the natural progression was to use di-Boc-protected pyrazole **109**, which is reportedly[158] more reactive than the unprotected version.

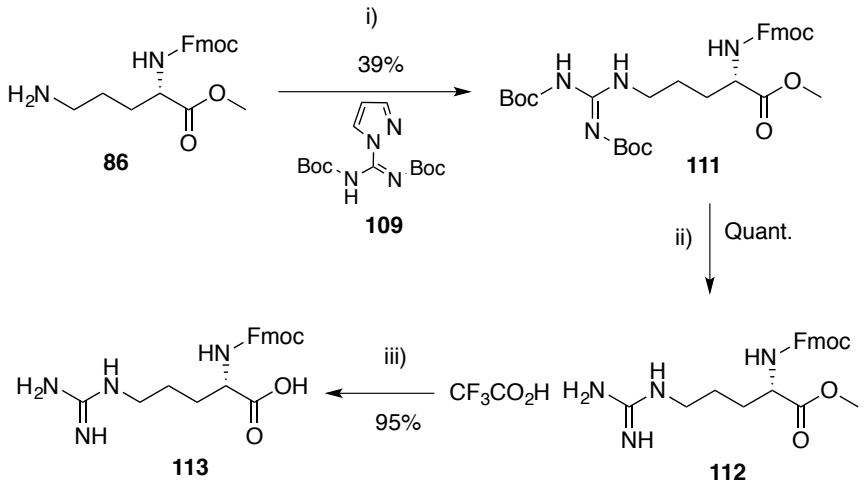
To circumvent the need for protection/deprotection steps, regioselective guanylation was attempted using *N,N'-bis*-Boc-pyrazole1*H*-carboxamidine **109** and unprotected Arg-OH **1** (**Scheme 2.17**).



Scheme 2.17 *– Reagents and conditions: i) DIPEA, MeOH 48 h, RT.*

Unfortunately this resulted in a complex mixture of products from which the desired product **110** could not be isolated by SiO_2 or Dowex® chromatography. Furthermore, the crude complex mixture was not amenable to analysis by ^{13}C or ^{1}H NMR, meaning no estimation of the reaction conversion could be made. Based on these inconclusive results, it was decided not to further explore this procedure.

Attempts were instead made to utilise protected Orn(Fmoc)-OMe **86** as the starting material, in the hope that the product would be more amenable to analysis and/or purification (**Scheme 2.18**).

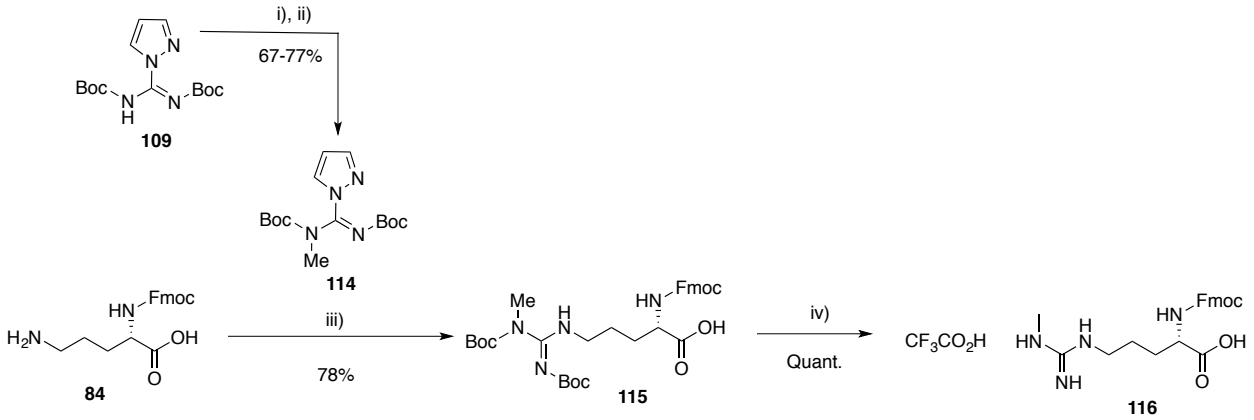


Scheme 2.18 *– Reagents and conditions: i) DIPEA, MeOH 48 h, RT ii) TFA:CH_2Cl_2 (3:7), RT, 2 h iii) 0.5 M $HCl_{(aq.)}$, 100 °C, 14 h.*

Orn(Fmoc)-OMe **86** was treated with the protected pyrazole **109** and DIPEA at room temperature for 48 h. It was possible to partially purify the product *N,N'-bis*-Boc-Arg(Fmoc)-OMe **111** *via* SiO_2 chromatography to give a 39% yield, but an approximate 50 mol% contaminant of DIPEA was found repeatedly difficult to remove. It is postulated that either the gelatinous nature of the product traps the base or a complex forms between them making it impossible to remove by high *vacuum* techniques or by purification. Nevertheless, partially-pure **111** was used in the next step; Boc-deprotection of **111** proceeded successfully by treatment with TFA:CH_2Cl_2. Arg(Fmoc)-OMe **112** was afforded in quantitative yield and the DIPEA contaminant from the previous step was eliminated during work-up (concentration under *vacuum*), making this route the current best option for guanylation. Subsequent ester hydrolysis was also achieved in excellent yield (95%) by refluxing in 0.5 M $HCl_{(aq.)}$ affording Arg(Fmoc)-OH **113** (**Scheme 2.18**). Following Pbf-protection this compound would be ready for inclusion into a peptide using SPPS.

2.7 – Synthesis of MMA(ω)

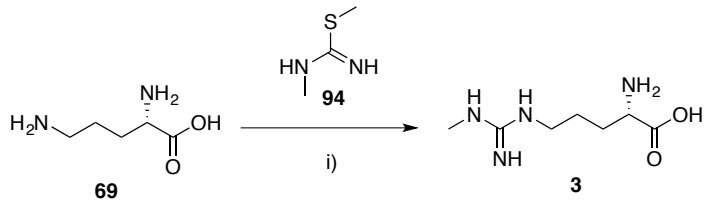
Given the successful guanylation of Arg(Fmoc)-OH **113** from Orn(Fmoc)-OMe **86** using protected pyrazole **109**, synthesis and application of an ω-methylated version was attempted using Orn(Cbz)-OH **85** as starting material. A patented procedure[159] exists for the synthesis of methylated pyrazole **114** by treatment of *N,N'-bis*-Boc-pyrazole1*H*-carboxamidine **109** with iodomethane (**Scheme 2.19**). This method was found to be reproducible; 0.7 g of said guanylating agent **114** was obtained in 67% yield and in up to 77% yield for smaller quantities.



Scheme 2.19 *– Synthesis of* ***116****; Reagents and conditions: i) NaH, N,N'-bis-boc-pyrazole-1H-carboxamidine* ***109****, DMF, RT, 5 min ii) MeI, 24-48 h iii) DIPEA, MeOH 48 h, RT iv) TFA:CH_2Cl_2 (3:7), 2 h.*

Orn(Cbz)-OH **85** was then treated with pyrazole **114** to furnish *N(ω)*-Me-Arg(Cbz)-OH **116** in 78% yield over two steps.

In an attempt towards generalising the guanylation procedure for methylated Arg derivatives, the monomethylated thiourea **94** was used with unprotected Orn-OH **69** under the premise that the reaction would be regioselective and that ion-exchange could be used to separate Orn-OH **69** from *N(ω)*-Me-Arg-OH **3** (**Scheme 2.20**).



Scheme 2.20 *– Attempted regioselective guanylation of Orn-OH* ***69*** *to* N(ω)*-Me-Arg-OH* ***3****. Reagents and conditions: i) 2 M* $NaOH_{(aq.)}$*, 100 °C, 5 h. Evidence of guanylation was observed by* ^{13}C *NMR but starting material and product material could not be separated for quantitative analysis.*

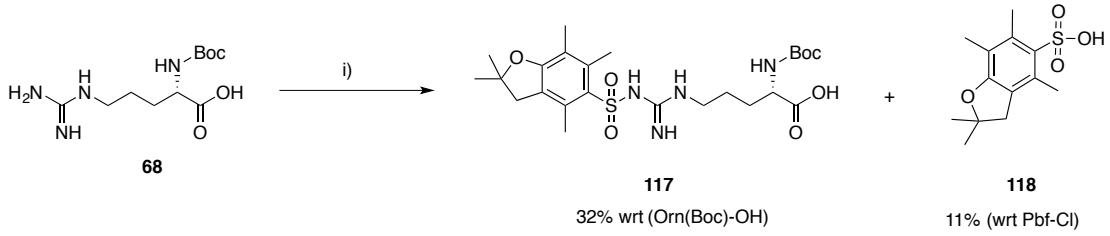
The same conditions were used as reported in **Section 2.6.1.b**, and it was possible to observe guanylation by ^{13}C NMR but not to estimate conversion from the spectrum of the crude mixture. Unfortunately the product and starting material co-eluted during attempts to purify *via* Dowex® ion-exchange chromatography; this observation was backed up by analytical scale LC/MS analyses, which showed co-elution of product **3** and starting material **69** under aqueous-normal phase conditions.

2.8 – Incorporating the Pbf Group

The difficulties encountered in both synthesising Pbf-pyrazole **107** and attempting to react it with an Orn derivative (**Section 2.7**) suggested harsh conditions, notably strongly basic conditions, might be necessary to incorporate the Pbf group onto Arg. Therefore, base-stable α-NH_2 protection of Orn was used to prevent cleavage of the Fmoc group under the anticipated reaction conditions. This would necessitate a protecting group exchange later in the synthetic route to install α-Fmoc protection on the Arg derivative to enable SPPS.

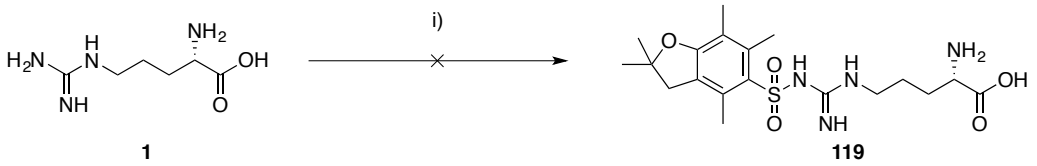
Attempts were made to introduce the Pbf group to the guanidine terminus of Arg(Boc)-OH **68**, by treatment with Pbf-Cl in $NaOH_{(aq)}$ (3 M) at room temperature, but these were low yielding

(**Scheme 2.21**). The reactions proceeded sluggishly, attributed to hydrolysis of Pbf-Cl to the sulfonic acid (Pbf-OH, **118**), which could be isolated *via* SiO_2 chromatography.



***Scheme 2.21** – Reagents and conditions: i) Pbf-Cl, $NaOH_{(aq.)}$ (3 M), RT, 18 h.*

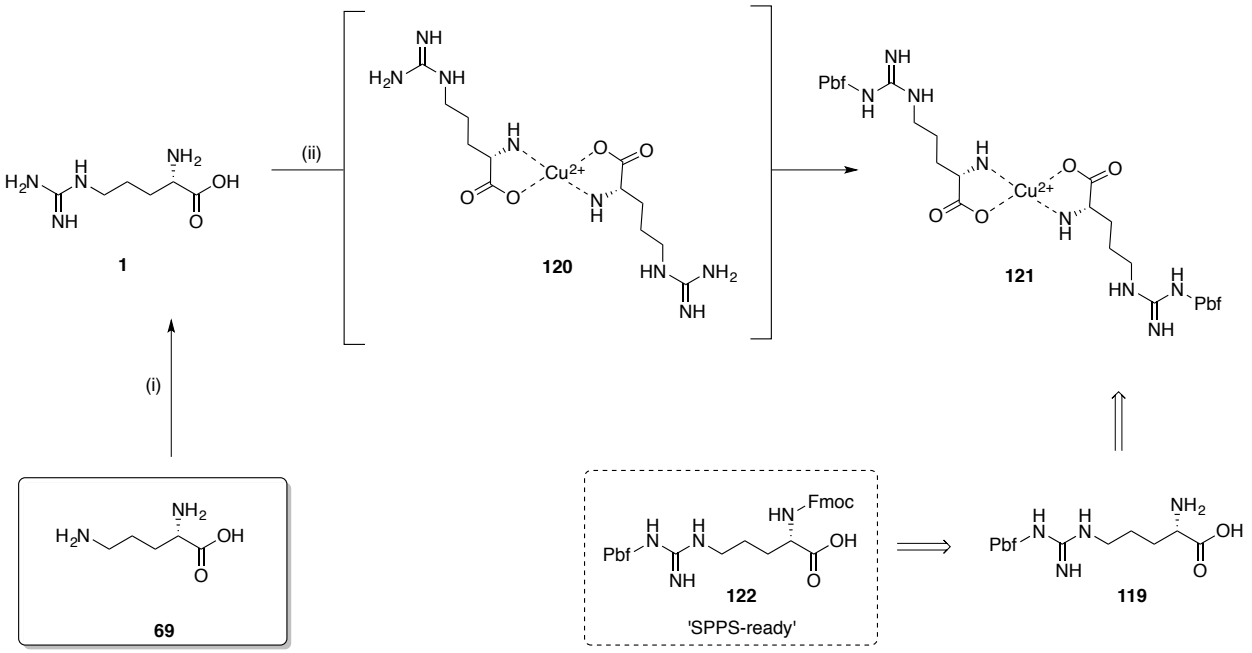
Given the regioselectivity observed for the reactions of pyrazole-1*H*-carboxamidines or *S*-methylisothioureas with unprotected Orn-OH **69**, it was hypothesised that a similar result could be obtained for the reaction of Pbf-Cl with unprotected Arg-OH **1** in aqueous base (**Scheme 2.22**).



***Scheme 2.22** – Attempted regioselective Pbf-protection. Reagents and conditions: i) Pbf-Cl, $NaOH_{(aq)}$ (4 M), RT, 18 h.*

This could have been an advantageous route for synthesising methylated Args from unprotected Orn, but in practice there was no regioselectivity and a mixture of mono- and di-substituted compounds were observed so this route was not further pursued.

Having now established that α-protection is clearly necessary to enforce selective ω-Pbf protection, the following retrosynthesis was envisaged (**Scheme 2.23**). This route aimed to build on the findings that complexing the α-groups with copper is the most practical approach to isolating high yields of δ-protected Orn:



***Scheme 2.23** – Reagents and conditions: (i)* S*-methylisothioureas (2 M NaOH, 100 °C, 5 h) or pyrazole-1*H*-carboxamidines (8 M $NaOH_{(aq)}$, RT, 5 d) or protected versions thereof and subsequent quantitative deprotection (ii) $Cu(OAc)_2$, 2 M $NaOH_{(aq)}$, Pbf-Cl.*

Based on ^{1}H NMR and MS analyses, the Pbf-protection itself was a success (**Scheme 2.23** step i), but unfortunately the [Arg-Pbf]$_2$Cu **121** complex was water soluble and did not precipitate. Having previously established in **Section 2.5** that isolation of precipitate is essential for achieving good purity and high yield, this planned synthesis would require further optimisation to reach the desired product **122.** It remains to be seen if this approach could prove useful for accessing some or all of the novel ω-methylated Arg's in the future and further work should focus on exploring this.

2.8 - Conclusions and Future Work

By adapting literature techniques, a high-yielding synthesis afforded protected Orn derivatives (**81-83**) on a multi-gram scale. These derivatives were used to assess reactivity and suitability of two major classes of guanylating reagent, the *S*-methylisothioureas and the pyrazole-1*H*-carboxamidines. It was hoped that a unified guanylation procedure using *S*-methylisothiourea **93**

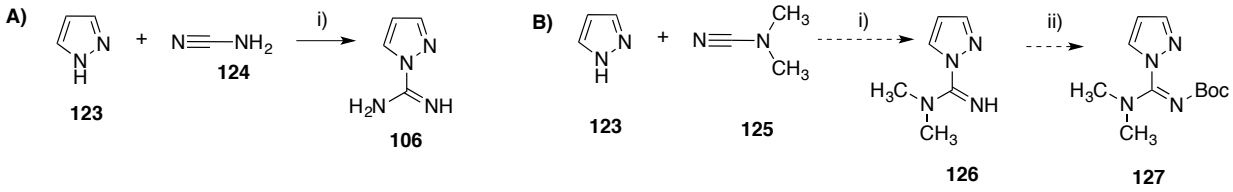
and its methylated analogues could allow access to the six novel methylated Args depicted in **Figure 2.1**. Unfortunately though, this class of reagent was found to be suboptimal; reaction conversion was poor and product purification ranged from low yielding to prohibitively difficult. The best outcome, using unprotected Orn-OH **69**, afforded unprotected Arg-OH **1** in 29% yield (91% purity). However, regioselective introduction of protecting groups was found to be much less achievable for Arg **1** than for Orn **69**; deriving a fully-protected version of the nascent Arg for SPPS was thus condemned as an unviable synthetic route.

The pyrazole-1*H*-carboximidines offered a slight improvement. A comparative yield (27%) and marginally improved purity (98%) was recorded for guanylation of naked Orn **69** to Arg **1**, achieved under much milder reaction conditions using pyrazole **106**. Use of the pyrazole-1*H*-carboxamidines as guanylating reagents was substantiated by the improved yield of Arg(Fmoc)-OMe **100** from Orn(Fmoc)-OMe **86** (39% over two steps) using the *bis*-Boc-protected pyrazole analogue **109**. Separation of starting material **86** from product **111** was facilitated by the fully-protected nature of **111,** enabling use of SiO_2 chromatography.

Promisingly, the *N*-methylated *bis*-Boc-protected pyrazole **114** also proved to be successful, guanylating Orn(Fmoc)-OH **84** to produce *N'*-Me-Arg(Fmoc)-OH **116** (78% over two steps). Purification problems prevented use of unprotected *N*-methylated *S*-methylisothiourea **94** towards regioselective guanylation of Orn **69**. All the guanylation data together suggests that methylated Args will require synthesis on a case-by-case basis and that no generalised procedure is applicable without using preparative scale reverse-phase HPLC, which might hold some promise for assisting delineation of some of the non-quantifiable reactions described herein. Such generalised syntheses should be the basis of future studies.

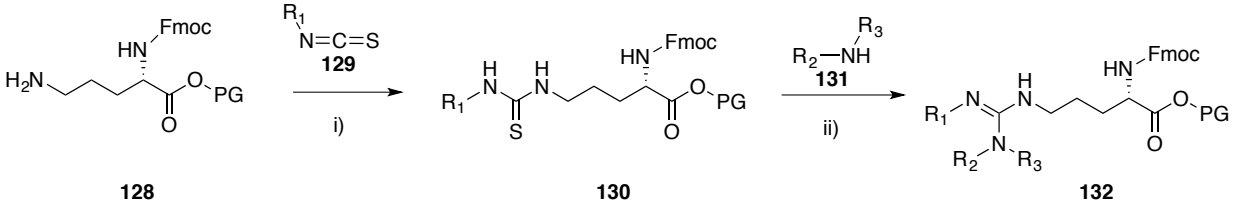
Testing the biological relevance of the multi-methylated Args outlined in **Figure 2.1** remains a valid and interesting line of research. It is recommended that future work should aim to synthesise each methylated Arg on a case-by-case basis using a combination of different guanylating reagents. Certainly the methylated and protected pyrazole **114** could be used to access the novel dimethylated Arg **19.** Also, identification of a strategy for orthogonally protecting Arg **1**, rather than Orn **69**, could allow utilisation of the methylated *S*-methylisothioureas.

Should optimisation of the *S*–methylisothiourea route prove unsuccessful, there is scope for synthesis of an asymmetrically dimethylated pyrazole-1*H*-carboxamidine **126** (**Scheme 2.24B**) which could prove useful for synthesis of **21**. By adapting a similar procedure (**Scheme 2.24A**)[160] it might be possible to access **126** in the following fashion (**Scheme 2.24B)**:



Scheme 2.24 – A) *Reported[160] reaction for synthesis of pyrazole-1*H*-carboxamidine. Reagents and conditions: i) Δ* ***106 B)*** *Proposed synthesis of the* N,N*-dimethyl derivative* ***126*** *and subsequent Boc-protection to give* ***127****. Reagents and conditions: i) Δ ii) Boc_2O.*

Additionally the following route, adapted from Martin *et al*[161], was identified as a possibility for accessing the remaining methyl patterns:



Scheme 2.25 – *Reagents and conditions: i) CH_2Cl_2, Et_3N ii) EDCl, CH_2Cl_2.*

This route requires a non-protic solvent and therefore manipulation of the carboxylate group would be required for solubility. The associated optimisations regarding potential cyclisation are predicted to require substantial future work if this route is to be used.

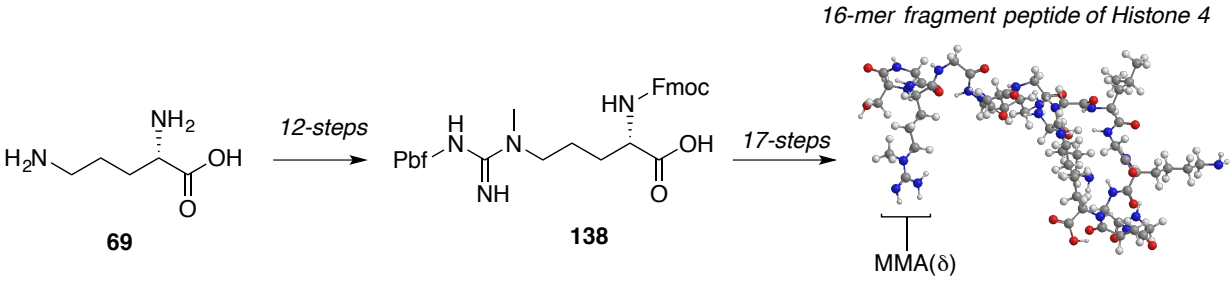
Given the success introducing unmethylated guanidine *via* application of *N,N'-bis*-Boc-pyrazole-1*H*-carboxamidine **109**, the decision was made to focus attention solely on synthesis of MMA(δ) **6.** Subsequent investigations as to whether this yeast-relevant epigenetic marker could have importance in humans are of particular interest. Chapter Three builds on the foundations of protection and guanylation outlined in the present chapter and applies them to the total synthesis of protected MMA(δ) **6** for SPPS.

Chapter 3: Total Synthesis of *N(δ)*-Me, *N*-Pbf-Arg(Fmoc)-OH

This chapter discusses development of a route for the total synthesis of *N(δ)*-Me, *N*-Pbf-Arg(Fmoc)-OH **138** from L-Orn on a scale that allowed its use in SPPS.

3.1 – Introduction and Aims

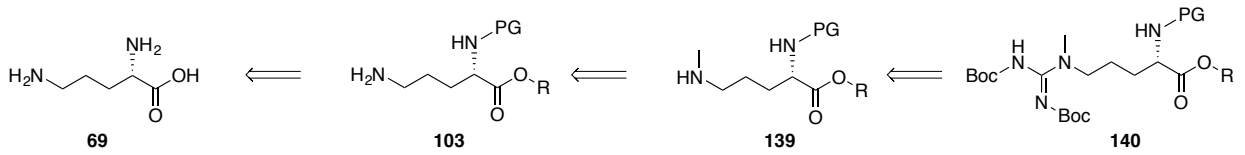
MMA(δ) **6** is a yeast epigenetic marker but has not previously been observed in humans or other higher organisms. In order to test this marke/r as a human PRMT substrate/product/inhibitor MMA(δ) **6** needed to be synthesised and incorporated into peptides for testing. The previous chapter focused on developing routes towards protected arginines for use in SPPS. Namely, the guanidine group of Arg requires Pbf protection and the α-NH_2 requires Fmoc protection. Building on the work in Chapter Two, the present chapter describes the synthesis of a protected MMA(δ) derivative **138** from the starting material L-Orn **69** (**Scheme 3.1**).



Scheme 3.1 *– Synthesis of protected MMA(δ)**138** and subsequent incorporation into a 16-mer fragment peptide of histone 4.*

It was concluded in Chapter Two that the most effective approach towards the synthesis of target methylated Args was to build up a methylation pattern *via* guanylation of Orn (**Scheme 3.2**). Application of this route for synthesis of protected MMA(δ) would first require synthesis of a δ-

monomethyl Orn (MMO(δ)) derivative. It was expected MMO(δ) **6** could be converted to a protected MMA(δ) derivative using *bis*-Boc-pyrazole-1*H*-carboxamidine **109** - identified in Chapter Two as the most appropriate reagent for such a conversion.



***Scheme 3.2** – Planned retrosynthesis of protected MMA(δ)* via *protected MMO(δ).*

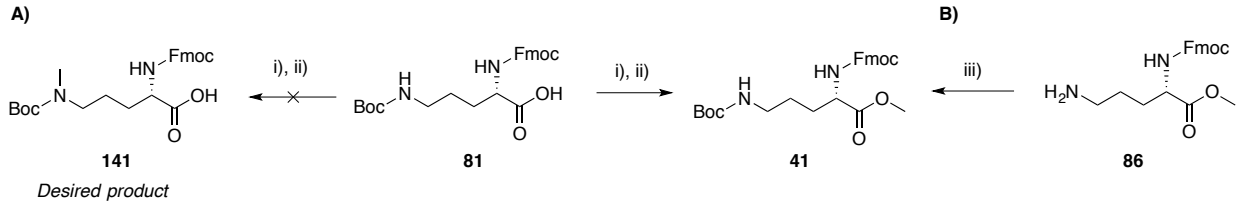
It was anticipated that monomethylation of Orn would be challenging given installation of one methyl unit would render the δ-nitrogen more nucleophilic and therefore prone to di-methylation[31].

In general there are limited examples of MMO(δ) synthesis[31,162,163], with no reported cases of protected derivatives that might facilitate synthesis of protected MMA(δ) **138** for SPPS.

3.2 – Attempted Derivatisation of MMO(δ) from *N*-Boc-Orn

3.2.1 – Towards Direct δ-Methylation of *N*-Boc-Orn(Fmoc)-OH

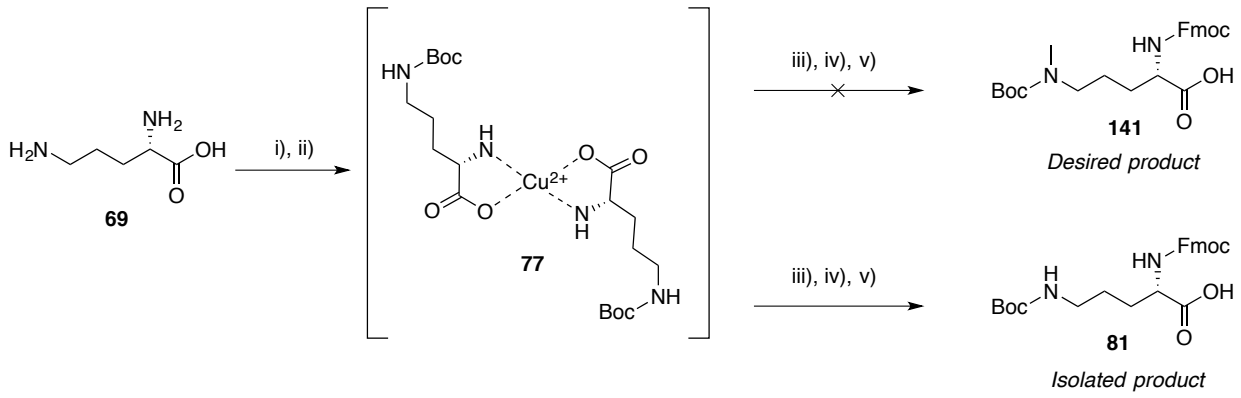
At first, options for directly installing a methyl group were explored. Iodomethane can be used to install a monomethyl group on *bis*-Boc-protected pyrazole **109** to form monomethylated *bis*-Boc-pyrazole **114** (**Section 2.7** and Asami[159]). It was asked whether this methodology could be applied to derive and isolate *N*-Me,*N*-Boc-Orn(Fmoc)-OH **141** from reaction of *N*-Boc-Orn(Fmoc)-OH **81** with NaH and iodomethane in DMF (**Scheme 3.3**).



Scheme 3.3 – A) *Attempted methylation on Boc-protected nitrogen resulted in methyl esterification and* ***B)*** *synthesis of a standard of* ***41****. Reagents and conditions: i) MeI, NaH (1 or 2 eq.), DMF, RT, 30 min ii) 60 °C, 2 h iii)* Boc_2O*,* $NaHCO_3$*, RT, 16 h.*

Unfortunately, the reaction did not generate the desired product **141** but instead afforded the methyl ester **41** as determined by HMBC analysis of the crude mixture, which also matched a standard of **41** synthesised by treating Orn(Fmoc)-OMe **86** with di-*tert*-butyl dicarbonate (**Scheme 3.3B**).

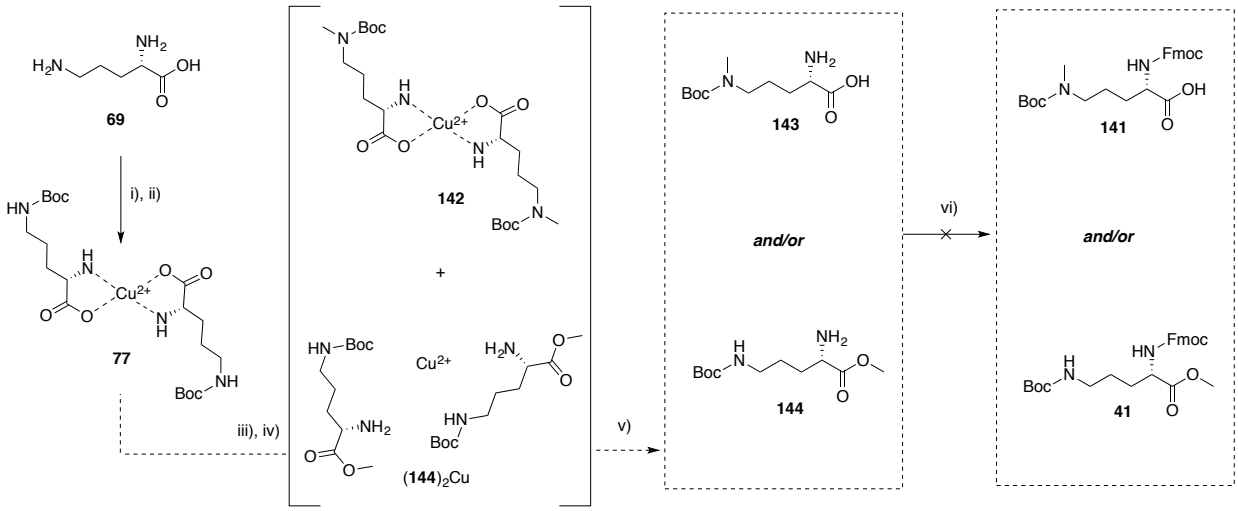
To prevent the observed methyl esterification and favour the desired reaction, attempts were made to shield the α-NH_2 and α-CO_2H of Orn *via* complexation with Cu^{2+}, before treatment with di-*tert*-butyl dicarbonate and then iodomethane (**Scheme 3.4**). 8-Quinolinol was used to remove copper and free the α-NH_2 for reaction with Fmoc-Cl. This approach had previously allowed isolation of a dual-protected derivative by acid-base extraction (**Section 2.5.2**).



Scheme 3.4 *– Attempted methylation on Boc-protected nitrogen* via *copper complexation i)* $Cu(OAc)_2$, $NaOH_{(aq)}$ (2 M), RT, 2 h ii) Boc_2O iii) MeI, NaH, DMF, 2 h, RT iv) 8-quinolinol, RT, 1 h v) Fmoc-Cl, N-hydroxysuccinimide, RT, 2 h.

Unfortunately, ^{1}H NMR analysis of the crude reaction mixture revealed that no methylation had taken place and that the unmethylated dual protected **81** had been made, a standard of which had already been synthesised (**Section 2.5.2**).

The procedure was reattempted with heating to 60 °C for 2 h upon treatment with iodomethane (**Scheme 3.5** step iv).

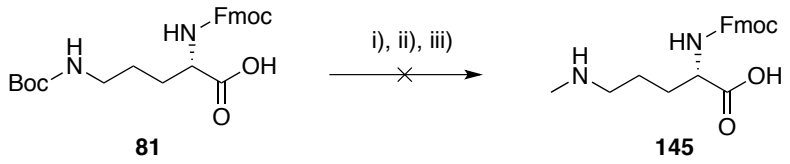


Scheme 3.5 – *Attempted methylation on Boc-protected δ-nitrogen. Reagents and conditions: i)* $Cu(OAc)_2$, NaOH(aq) (2 M), RT, 2 h ii) Boc_2O, iii) MeI, NaH, DMF, RT, 2 h iv) 60 °C, 2 h v) 8-quinolinol, RT, 2 h vi) Fmoc-Cl, N-hydroxysuccinimide, RT, 2 h.

MS analyses suggested a monomethylated species might be present. However, a colour change from blue to grey was noted, concomitant with heating the reaction during step iv, suggesting that the copper-mediated complex may have broken down under the conditions described. It was speculated this putative monomethylation might have occurred at the carboxylate group, as previously observed (**Scheme 3.3**), which could have been liberated from the copper complex during heating. Attempts were made to derivatise to the α-Fmoc-protected Orn **141** and isolate the desired product *via* acid-base extraction (as above and **Section 2.5.2**). However, no products could be isolated.

3.2.2 – Towards Reduction of *N*-Boc to *N*-Me

An alternative procedure was sought; according to Phuan *et al*.[164] $LiAlH_4$ can be used to reduce Boc groups to methyl groups. This methodology was applied in an attempt to reduce *N(δ)*-Boc to *N(δ)*-Me by refluxing with $LiAlH_4$ for 16 h in THF (**Scheme 3.6**).



81 **145**

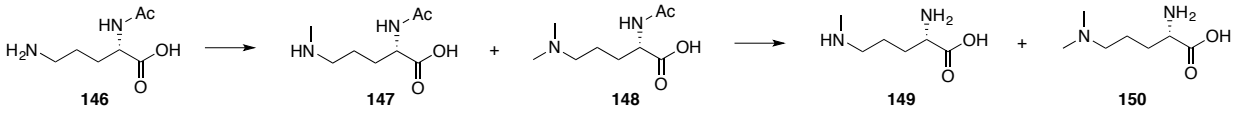
Scheme 3.6 *– Attempted reduction of* N-*Boc to* N(δ)-*Me. Reagents and conditions: i) $LiAlH_4$, THF, 0 °C, ii) 70 °C, 16 h iii) saturated aqueous solution of $KNaC_4H_4O_6{\cdot}4H_2O$ to quench reaction, 25 °C, 6 h.*

The crude reaction mixture was poorly soluble in all organic solvents attempted, except in acidic aqueous solution (1 M $HCl_{(aq)}$, pH 1). Analysis of this crude reaction mixture by ^{1}H NMR in D_2O/DCl revealed a complex mixture of products within which the desired product **145** could not be detected.

3.3 –Reductive Amination Procedure

The current project aims to produce a protected derivative of MMA(δ) on a scale close to 1 mmol to facilitate SPPS. Zobel-Thropp *et al.*[31] describe use of a reductive amination procedure to synthesise MMO(δ) **172** on a 0.012 mmol scale *via* the α-protected *N*-acetyl derivative **146**. Using 1 eq. formaldehyde in $NaOH_{(aq)}$ and reduction with $NaBH_4$, they observed a mixture of starting material **146**, MMO(δ) **149** and dimethylated Orn **150** (**Scheme 3.7**). Their crude mixture was purified by HPLC and the authors alluded to problematic co-elution of the mono- and di-methylated species but offered no indication of reaction conversion or yield. This

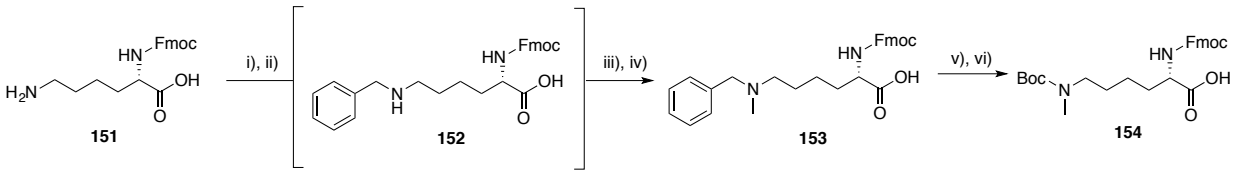
procedure was therefore considered an unsuitable starting point for synthesis of preparative scale protected MMO(δ) **149** for conversion to protected MMA(δ) **138**.



***Scheme 3.7** – Synthesis of analytical quantities of MMO(δ) **149**. Reagents and conditions: i) 1 M $NaOH_{(aq)}$, formaldehyde ii) $NaBH_4$ iii) 6 M $HCl_{(aq)}$.*

Huang *et al.*[162] described a sequential reductive amination procedure for synthesis of α-NH_2 Fmoc-protected monomethyl-Lys (MML) **154**, which is congeneric with Orn, in very good yield (85%).

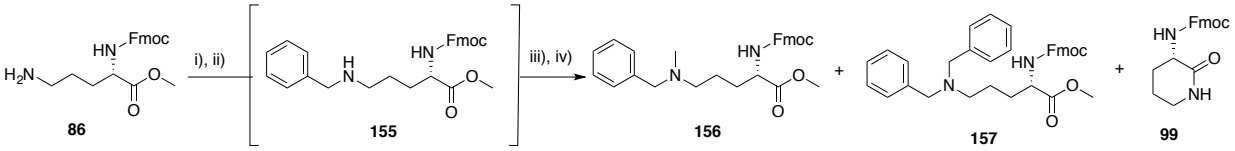
According to Huang *et al.*[162], treatment of Lys(Fmoc)-OH **151** with benzaldehyde and 3 Å molecular sieves followed by reduction with $NaCNBH_3$ afforded the monobenzylated derivative **152** (**Scheme 3.8**). In a one-pot procedure, aqueous formaldehyde was subsequently added followed by reduction with $NaCNBH_3$ to form *N*-benzyl, *N*-Me-Lys(Fmoc)-OH **153** in 85% yield following purification on SiO_2. All reactions took place at room temperature. The authors replaced the benzyl group with Boc *via* catalytic hydrogenolysis (Pd/C, H_2) and treatment with Boc_2O in a one pot procedure. They report a yield of 93% for the product affording *N*-Boc, *N*-Me-Lys(Fmoc)-OH **154.**



***Scheme 3.8** – Procedure by Huang* et al.[162] *for selective reductive amination of Lys(Fmoc)-OH **151** to MML derivative **154**. Reagents and conditions: i) benzaldehyde, 3 Å molecular sieves ii) $NaCNBH_3$ iii) formaldehyde iv) $NaCNBH_3$ v) 10% Pd/C, H_2 vi) Boc_2O.*

3.3.1 – Reductive Amination to Selectively Mono-methylate Ornithine

The method described by Huang *et al.*[162], was therefore adapted for the synthesis of Orn(Fmoc)-OMe **86** (**Scheme 3.9**). It was elected to use the methyl ester rather than the carboxylic acid because of the ease of handling (**Section 2.5.3**).



Scheme 3.9 *– Synthesis of* N-*Me,* N-*benzyl-Orn(Fmoc)-OMe* ***156****. Reagents and conditions: i) benzaldehyde (1.2 eq.), 3 Å molecular sieves ii)* $NaCNBH_3$ *iii) formaldehyde iv)* $NaCNBH_3$

Early attempts using a range of concentrations, reaction times or equivalents of $NaCNBH_3$ generated complex mixtures within which the desired product could not be distinguished by ^1H NMR for estimation of conversion. Attempts to separate the products on SiO_2 were partially successful, affording small quantities of partially purified product **156** and the putative dibenzylated side product **157** as determined by MS and ^1H NMR of the crude mixture.

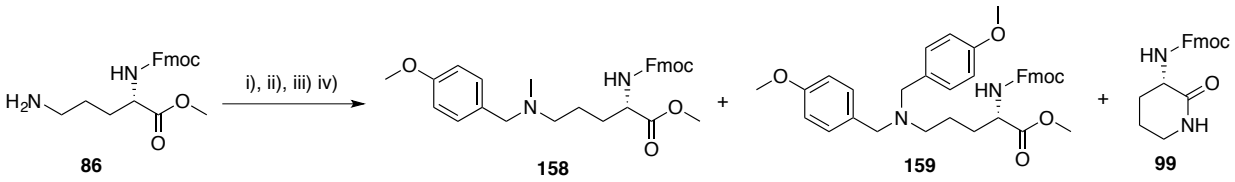
In an effort to encourage formation of the desired product **156**, attempts were made to isolate the mono-benzylated intermediate **155** using similar conditions, but a mixture of putative dibenzylated derivative **157** and partially pure δ-lactam **99** (*vide infra*) were the only isolated products.

3.3.2 – Attempts to Improve Imine Formation

In the original report using Lys[162], the authors report no dibenzylation and propose that the terminal nitrogen of the mono-benzylated speices is sterically hindered by the bulky aromatic group such that a second equivalent of benzaldehyde is unable to approach to give the undesired iminium ion intermediate.

However, in this case low overall yields contaminated with putative dibenzyl-side product supported that there was poor initial conversion to the benzaldehyde-imine intermediate.

Locke *et al*[165] report that benzaldehydes substituted with *para*-electron donating groups react more readily with primary amines to form the requisite imine. Reactions using *p*-anisaldehyde instead of benzaldehyde were therefore attempted under a series of similar conditions to those reported by Huang *et al.*[162] in an effort to improve overall conversion (**Scheme 3.10**).



Entry	Reducing Agent	pH	Conditions	SM 86, %	PM 158, %	SPM 159, %	CM 99, %
1	$NaCNBH_3$	7-8	16 h, 25 °C then 2.5 h, 25 °C	34	10	42	0
2	$NaCNBH_3$	7-8	2 h, 80 °C then 48 h, 25 °C	0	0	22	14
3	$Na(CH_3CO_2)_3BH_3$	4-5	16 h, 25 °C then 2.5 h, 25 °C	51	3	0	0

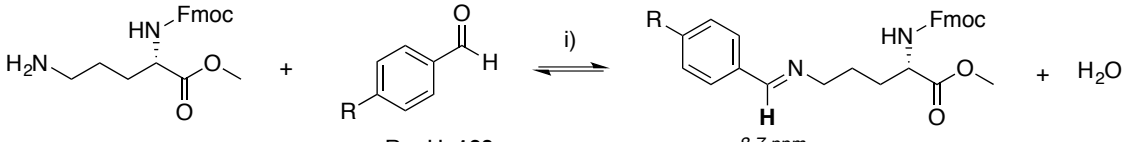
Scheme 3.10 *– Towards synthesis of protected derivatives of MMO(δ)* via *selective reductive amination. Reagents and conditions: i)* p*-anisaldehyde, 3 Å molecular sieves ii) $NaCNBH_3$ iii) formaldehyde iv) $NaCNBH_3$.*

Following Si_2O chromatography, an assessment of reaction success was made based on recovery of starting material (SM) **86**, desired product material (PM) **158**, di-substituted side product material (SPM) **159** and the δ-lactam cyclised material (CM) **99**. Entry one shows that under conditions similar to those reported by Huang *et al*.[162] a disappointing yield of desired product **158** was obtained with 42% disubstituted **159.** Entry two describes the use of elevated temperature in an attempt to encourage imine formation. Unfortunately this encouraged cyclisation to the δ-lactam **99** with 14% yield isolated from purification *via* SiO_2. Furthermore, production of disubstituted product **159** remained a problem, with a 22% yield being afforded

under these conditions. It is noted that an apparent loss of material was observed with respect to percentage yield. This is explained as possible degradation on SiO_2 or during reaction heating.

The effect of pH on the reaction was considered – generally the optimum pH for imine formation is somewhere close to five[166] and a more acidic pH might discourage cyclisation to **99**. A pH of four was established using AcOH (entry three). For this test reaction, an alternative reducing agent was used ($Na(CH_3CO_2)_3BH_3$) because of the associated hazards of handling $NaCNBH_3$ in acidic conditions. Regrettably this reaction was gave poor conversion, retuning 51% starting material **86** from purification on SiO_2 and only 3% of desired product **158**, although gratifyingly no disubstituted side-product **159**. A similar problem with loss of material was recorded for entry three as had been observed for entry two.

In light of the disappointing results presented in **Scheme 3.10**, alternative drying agents were investigated in an attempt to improve imine conversion. To this end, trimethylorthoformate (TMO) (10 eq.) was used as the drying agent in the reaction of 1 eq. of *p*-anisaldehyde **161** with Orn(Fmoc)-OMe **86** (**Scheme 3.11**).



86

R = H, **160**
R = OMe, **161**

R = H, **162**
R = OMe, **163**

Entry	Aldehyde	R	Conditions	pH	Drying agent	Imine, %
1	*p*-Anisaldehyde, **161**	OMe	6 h, 80 °C	4	TMO, 10 eq.	8
2	Benzaldehyde, **160**	H	6 h, 80 °C	4	TMO, 10 eq.	0
3	*p*-Anisaldehyde, **161**	OMe	6 h, 80 °C	4	TMO, 10-70 eq.	2-8
4	*p*-Anisaldehyde, **161**	OMe	6 h, 80 °C	4	3 Å Sieves	25
5	Benzaldehyde, **160**	H	6 h, 80 °C	4	3 Å Sieves	4

***Scheme 3.11** – Attempted synthesis of imine intermediate **162** or **163**. Reagents and condition: i) AcOH added dropwise until pH 4 was reached. All other conditions as stated in table.*

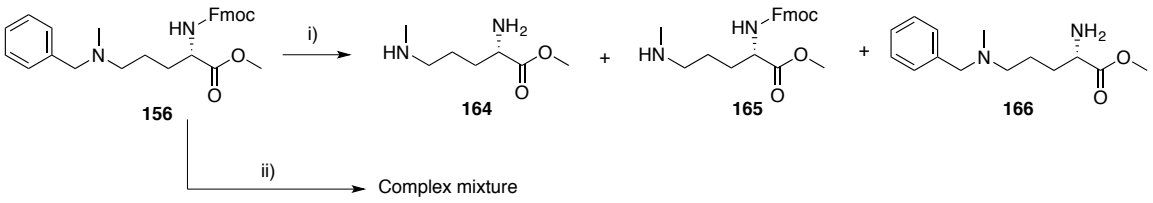
The crude reaction mixture was analysed by ^{1}H NMR and percentage conversion assessed by quantifying the ratio of imine (~8.7 ppm, C*H*=N) **162**/**163** to starting material **86** (~7.8 ppm, Fmoc-C*H*) (**Scheme 3.11**).

TLC suggested no imine was formed after 2 h at room temperature so the mixture was heated to 80 °C for 6 h to encourage imine formation and subsequent analysis of the crude by ^{1}H NMR revealed 8% imine had formed (entry one). The same reaction was carried out using benzaldehyde **160** as a point of reference to the original work[162] and this time ^{1}H NMR analysis of the crude revealed production of 0% imine (entry two), confirming the preferential use of *p*-anisaldehyde **161**.

It was hypothesised that increasing the quantity of drying agent in the reaction would improve conversion to imine. However, increasing the equivalents of TMO gave no improvement on imine conversion (entry three). For all reactions described above, the pH was adjusted to four by dropwise addition of AcOH. The best conversion that could be obtained was using 3 Å molecular sieves (entry four, 25% imine) but overall it was concluded that imine conversion was generally poor.

3.3.2 – Debenzylation of *N*-Me, *N*-Benzyl-Orn(Fmoc)-OMe to form a MMO(δ) Derivative

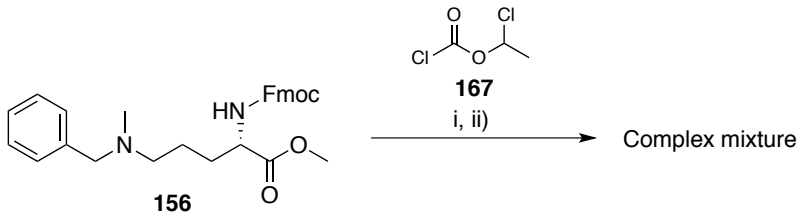
Using a small sample of partially pure *N*-Me, *N*-benzyl-Orn(Fmoc)-OMe **156** synthesised in the previous section, attempts were made to selectively remove the benzyl group to give *N*-Me-Orn(Fmoc)-OMe **165** (**Scheme 3.12**).



Scheme 3.12 *– Attempted selective debenzylation of* N-*Me,* N-*benzyl-Orn(Fmoc)-OMe* ***156****. Reagents and conditions: i) 20% Pd(OH)$_2$/C (50% w/w), H$_2$ ii) 10% Pd/C, CO$_2$HNH$_4$, 80 °C, 30 min.*

At first, attempts were made to selectively remove the benzyl group using Pd(OH)$_2$/C *via* catalytic hydrogenolysis, but unfortunately a complex mixture of products was afforded that could not be interpreted by ^{1}H NMR. MS and TLC analyses suggested the presence of three putative products (**164-166**); attempts were made to separate these using SiO_2 but were unsuccessful.

An alternative procedure was attempted using Pd/C following a procedure optimised in our laboratory utilising ammonium formate as the H_2 source. Unfortunately the resultant complex mixture of products could neither be separated *via* SiO_2 nor interpreted by ^{1}H NMR. MS and analyses suggested the Fmoc group may be labile under these conditions; a precedent exists for Fmoc deprotection using heterogeneous catalytic hydrogenation conditions[167].



Scheme 3.13 *– Attempted selective debenzylation of* N-*Me,* N-*benzyl-Orn(Fmoc)-OMe* ***156****. Reagents and conditions: i) 0 °C, 15 min, CH$_2$Cl$_2$ iv) 50 °C, 1 h ii) concentrate then reflux in MeOH, 50 °C, 45 min.*

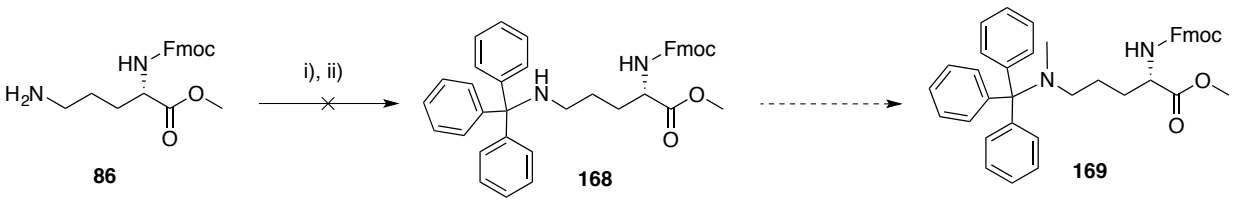
According to Olofson *et al.*[168], α-chloroethyl chloroformate (ACEC) **167** followed by methanolysis can be used as a dealkylating method and is reported to preferentially remove

benzyl groups over methyl groups. Compound **156** was treated with ACEC **167** but unfortunately none of the desired product **165** was obtained and only a complex mixture of products was recovered (**Scheme 3.13**).

Based on the series of poor reaction outcomes observed in both the reductive amination reactions and the subsequent attempted debenzylations, the decision was taken to abandon attempts to access a protected MMO(δ) *via* this route.

3.3.4 – Attempted Use of the Trityl Protecting Group for the δ-Amino Substituent

Through their reductive amination procedure, Huang *et al.*[162] had suggested the incorporation of a bulky substituent at the δ-amino group lends some control over poly-substitution. Orn(Fmoc)-OMe **86** was therefore treated with trityl chloride[169] (**Scheme 3.14**) to try and introduce a bulky substituent at the δ-amino group without resorting to a reductive amination procedure. It was anticipated that a methylation reaction might be possible on *N*-trityl-Orn(Fmoc)-OMe **168** to produce the monomethylated product **169** that could be subsequently trityl-deprotected.



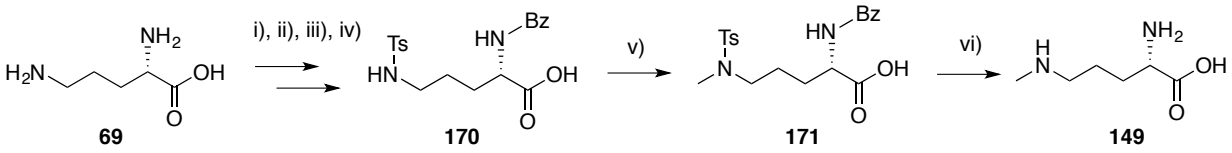
***Scheme 3.14** – Failed attempt to sterically hinder the δ-nitrogen to facilitate a monomethylation reaction. Reagents and conditions: i) Et_3N, Ph_3CCl, EtOH, 0 °C to RT, 2 h ii) 40 °C, 3 h.*

This reaction was not successful, producing a complex mixture that could not be distinguished by ^{1}H NMR; TLC analysis suggested that the majority of the reaction mixture contained starting material.

Alternative procedures were therefore investigated for introduction of a single methyl group onto the δ-nitrogen of ornithine towards the synthesis of protected MMA(δ) **138**.

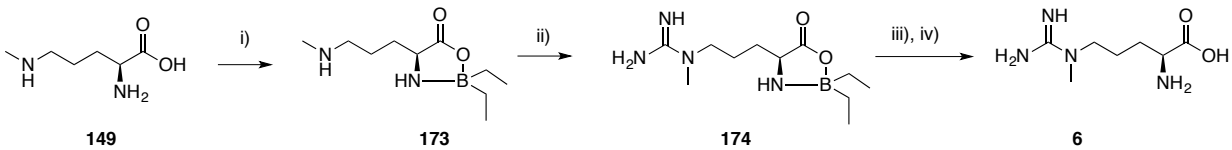
3.4 –Synthesis of MMO(δ) *via N*-Sulfonyl-Protected Orn

According to Luzzi & Marletta[170], MMO(δ) **149** can be synthesised by methylation of a tosyl-protected δ-nitrogen of Orn using iodomethane (**Scheme 3.15**). Global de-protection in refluxing conc. $HBr_{(aq)}$ generated the unprotected amino acid **149**.



Scheme 3.15 *– Synthesis of MMO(δ)* ***149*** *from L-Orn* ***69****. Reagents and conditions: i) $CuCO_3$ $Cu(OH)_2$, reflux ii)* p-*toluenesulfonyl chloride, 2 M $NaOH_{(aq)}$ iii) H_2S, 2 M, HCl iv) benzoyl chloride, 1 M $NaOH_{(aq)}$ v) MeI, 2 M NaOH vi) HBr 48%, reflux.*

The authors went on to convert MMO(δ) **149** to MMA(δ) **6** by shielding the α-groups *via in situ* generation of a triethylborane complex (**Scheme 3.16**). After guanylation, which in this case was achieved using $HgCl_2$ and Boc-protected thiourea, the borane complex was broken by refluxing in $HCl_{(aq)}$ followed by treatment with TFA at room temperature to afford MMA(δ) **6**.



Scheme 3.16 *– Conversion of MMO(δ)* ***149*** *to MMA(δ)* ***6*** via *complexation through borane. Reagents and conditions: i) 1 M E_3N, THF, dimethoxyethane ii)* N,N'-*bis*-tert-*butoxycarbonylthiourea, $HgCl_2$, Et_3N, DMF iii) 1.5 M, $HCl_{(aq)}$, 100 °C iv) TFA.*

Selective orthogonal protection of Arg was found to be less facile compared to Orn (**Section 2.8**), meaning the method reported by Luzzi & Marletta[170] would require significant adaptation to render a product suitably protected for SPPS. Additionally, the authors' synthetic route had an overall yield of only 6% after six steps, so it was decided their approach was not wholly suitable for producing a sufficient quantity of the fully-protected derivative of MMA(δ) for SPPS.

However, a number of solutions were envisaged that might allow exploitation of Luzzi & Marletta's[170] approach to mono-methylation of Orn to eventually afford the desired protected MMA(δ) for SPPS.

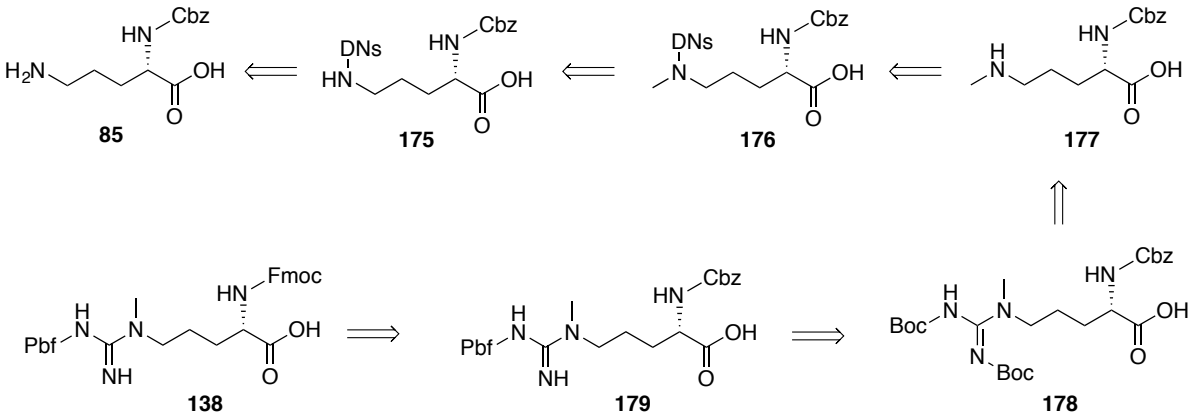
3.4.1 – Exploration of More Labile Sulfonyl-Protecting Groups

Ultimately, the biggest drawback to Luzzi & Marletta's[170] synthetic route is the deprotection step requiring reflux in $HBr_{(aq)}$, which necessitates subsequent orthogonal protection of MMO(δ) if a protected MMA(δ) is to be synthesised. There are two reasons this is disadvantageous:

1) Harsh conditions may encourage amino acid racemisation
2) Compete de-protection of MMO(δ) is step-inefficient since subsequent conversion of MMO(δ) is necessary towards protected MMA(δ) suitable for SPPS.

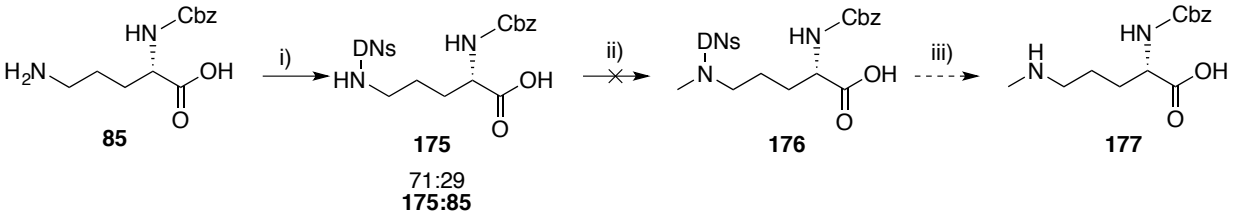
The need to remove the tosyl group demands these harsh conditions since this protecting group is very robust and non-labile. Tosyl protection of the δ-nitrogen does not preclude its subsequent methylation and so an alternative sulfonyl protecting group was sought that might be more facile to remove and avoid use of $HBr_{(aq)}$ for deprotection. Fukuyama *et al.*[171] describe the use of 2- and 4-nitrobenzenesulfonamides (Ns) and 2,4-dinitrobenzenesulfonamides (DNs) as protecting groups for primary and secondary amines. Moreover, they are extremely facile to deprotect and

in particular the DNs group can be removed by *ipso* substitution using *n*-propylamine; these mild conditions could enable a protecting group exchange at the DNs-protected δ-nitrogen while leaving orthogonal α-NH_2 protecting groups intact. The following retrosynthesis was therefore envisaged (**Scheme 3.17**):



***Scheme 3.17** – Retrosynthetic approach towards protected MMA(δ) **138** for SPPS* via *DNs-protected and monomethylated **176**.*

Treatment of Orn(Cbz)-OH **85** with 2,4-dinitrobenzenesulfonyl chloride in $NaOH_{(aq)}$ (1 M) afforded a 71:29 mixture of dual protected **175**:starting material **85** after three days; the crude mixture containing novel **175** was characterised as fully as possible (**Scheme 3.18**). Unfortunately, this compound appeared to be unstable on SiO_2 and attempts to purify *via* normal-phase chromatography were unsuccessful, returning a complex mixture that was less interpretable by ^{1}H NMR than the crude reaction mixture. **175** was used in subsequent steps as a mixture with **85**.



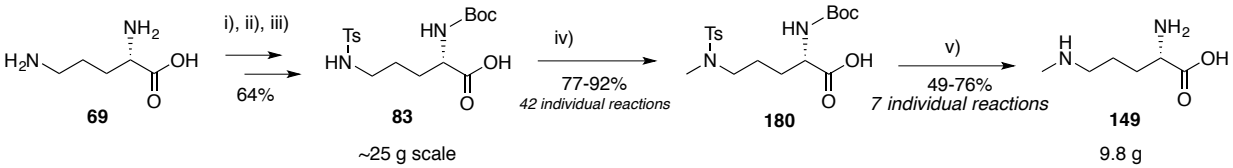
***Scheme 3.18** – Synthesis of **175** and attempted synthesis of methylated derivative **176**. Reagents and conditions: i) 2,4-dinitrobenzenesulfonyl chloride, $NaOH_{(aq)}$ (1 M), RT, 3 d ii) MeI, K_2CO_3, DMF, 2 h. iii)* n*-propylamine, CH_2Cl_2, 23 °C, 10 min.*

Crude *N*-DNs-Orn(Cbz)-OH **175** was treated with iodomethane and five eq. K_2CO_3. Unfortunately, the method described by Fukuyama *et al.*[171] could not be successfully applied to this reaction system and the conditions led to removal of the DNs group, as confirmed by ^{1}H NMR. Furthermore, ^{1}H NMR and MS analyses of the crude mixture showed no indication that methylation had been successful.

It is anticipated that with optimisation this route might hold further promise, for example, use of the reportedly less labile Ns protecting group(s) might be a more appropriate course of action and prevent the observed deprotection during the methylation step.

3.4.2 – Improved Synthesis of Unprotected MMO(δ)

The first step in the reported procedure described by Luzzi & Marletta[170] involved synthesis of a benzoyl-protected α-NH_2 *via* copper-complexation. As part of the current work, the Boc group had been successfully and selectively installed on the α-NH_2 of Orn in high yield (**Section 2.5.2**) and so was used instead of the reported benzoylation. This modified procedure led to a more productive synthesis of MMO(δ) **149** with a better overall yield (45%) compared to that reported (12%).



Scheme 3.19 *– Synthesis of MMO(δ)* ***149*** *using a procedure adapted from Luzzi & Marletta[170]. Reagents and conditions: i)* $Cu(OAc)_2$, $NaOH_{(aq)}$ (2 M), p-toluenesulfonyl chloride, Et_3N, 2 d ii) 8-quinolinol iii) Boc_2O, 3 d iv) *MeI, 2 M NaOH v) HBr 48%, reflux.*

Synthesis of the first intermediate *N*-Ts-Orn(Boc)-OH **83** proceeded in good yield (64% compared to reported[170] 17%) and on large scale (24.6 g) *via* copper complexation (**Section**

2.5.2) (**Scheme 3.19**). Orn-OH **69** was incubated with $Cu(OAc)_2$ in $NaOH_{(aq)}$ (2 M) and treated with *p*-toluenesulfonyl chloride and Et_3N for two days. After removal of copper by chelation with 8-quinolinol, di-*tert*-butyl dicarbonate was added and the reaction stirred for three days. *N*-Ts-Orn(Boc)-OH **83** was isolated cleanly by acid-base extraction (64% yield).

Methylation of the tosylated nitrogen proceeded smoothly and in good yield (77-92% compared to reported[170] 86%) using three sequential equivalents of iodomethane at approx. 1 h spacing and refluxing in $NaOH_{(aq)}$ as previously described (**Scheme 3.19**). It was necessary to perform this step under pressure, achieved using sealed vials, and this restricted the scale to ~2.5 mmol per reaction. Approximately 40-fold this scale was desirable for subsequent reactions with MMO(δ) **149** towards synthesis of protected MMA(δ) **138** for SPPS. Attempts were made to methylate **83** to **180** using reflux apparatus, which would facilitate a larger scale reaction system and avoid the need for pressurised vials, but unfortunately the yield for this step was found to be significantly worse (55% compared to 77-92% for the pressurised method). Presumably volatile iodomethane more readily escapes an open reaction system compared to a closed one, which may explain the difference in yield. Overall, 42 individual methylation reactions were carried out as part of the scale-up synthesis in sealed vials and the products pooled for larger scale global de-protection towards **138**. In the future, optimisation of this step towards a more time-efficient procedure could be beneficial, for instance, use of large scale pressurised reaction vessels.

To remove the tosyl group, it was necessary to implement the aforementioned global deprotection step using conc. $HBr_{(aq)}$. Similar work by Schade *et al.*[163] made two stipulations that are purportedly necessary to avoid racemisation at the α-carbon of MMO(δ) **149** during this step. Firstly, it is advised that a reaction time no longer than 1.5 h is enforced and secondly, removal of acid under high *vacuum* must take place at temperatures no higher than 50-60 °C.

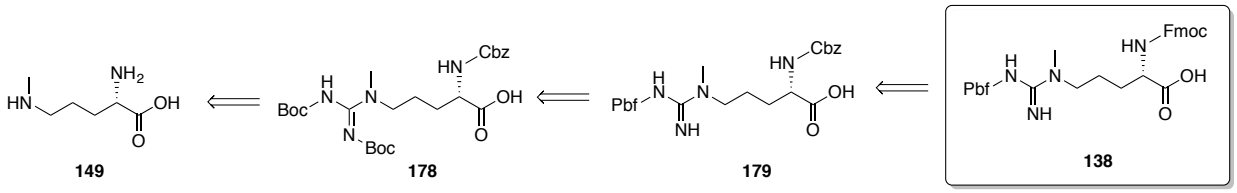
Dual-protected MMO(δ) **180** was refluxed on no larger than ~6 g (~15 mmol) scale in approximately 60 mL conc. $HBr_{(aq)}$ at 125 °C. Incomplete deprotection of the tosyl group was observed after 1.5 h on this scale, so the reaction time was extended to 2 h. Importantly, analysis of optical rotation of MMO(δ) **149** obtained as part of this work was found to have a very similar specific rotation compared to that reported by Schade *et al.*[163] who assessed MMO(δ) **149** to have an *e.e.* of 98.2% by chiral HPLC (**Table 3.1**).

Compound	Specific Rotation	Solvent	Concentration
MMO(δ), **149**	$[\alpha]_D^{22}$ + 19.8	6 M $HCl_{(aq)}$	1.4 g/100 mL
	$[\alpha]_D^{20}$ + 23.3 (*ee = 98.2%)*[163]	6 M $HCl_{(aq)}$	2 g/100 mL

***Table 3.1** – Comparisons of specific rotation properties suggesting MMO(δ) **149** has not racemised during synthesis.*

Initial attempts to remove the aqueous acid from the crude mixture under *vacuum* were met with difficulty and resulted in moderate yield (49%). To avoid increasing the temperature above the recommended[163] maximum of 50-60 °C, a purification procedure was developed that bypassed the need to concentrate the mixture. This procedure facilitated direct application of the entire reaction volume to the column and involved an incubation step with Dowex® resin and use of three tandem columns to trap the entire product, which is eluted in $NH_4OH_{(aq)}$ (9 M) . The procedure, described in more detail in the Experimental Section, was implemented to afford MMO(δ) **149** in good yield (76%). A total of 3.7 g (30 mmol) was synthesised (six reactions).

With moderate quantities of MMO(δ) **149** now at hand from this newly optimised procedure, a retrosynthetic scheme towards deriving protected MMA(δ) was planned (**Scheme 3.20**).



***Scheme 3.20** – Retrosynthesis of MMO(δ) towards protected MMA(δ) **138** for SPPS.*

In Chapter Two, it was determined that the conditions required for Pbf-protection of the guanidine group are strongly basic, so the base-labile Fmoc-group must be installed on the α-NH_2 as the final step in the synthesis. Use of α-Cbz protection throughout the route was therefore planned according to the retrosynthesis shown in **Scheme 3.20** and a protecting group exchange planned as the last step.

Two different approaches were attempted towards synthesising protected MMA(δ) **138** from unprotected MMO(δ) **149**.

3.4.3 – Use of Borane Complexation to Shield α-NH_2 and α-CO_2H

The first approach again builds on Luzzi & Marletta's work[170] describing the use of triethylborane to complex the α-NH_2 and α-CO_2H (**Scheme 3.16**). Instead of introducing an unprotected guanidine to the δ-nitrogen at this stage, reaction with a protected guanylating reagent was envisaged (**Scheme 3.21**).

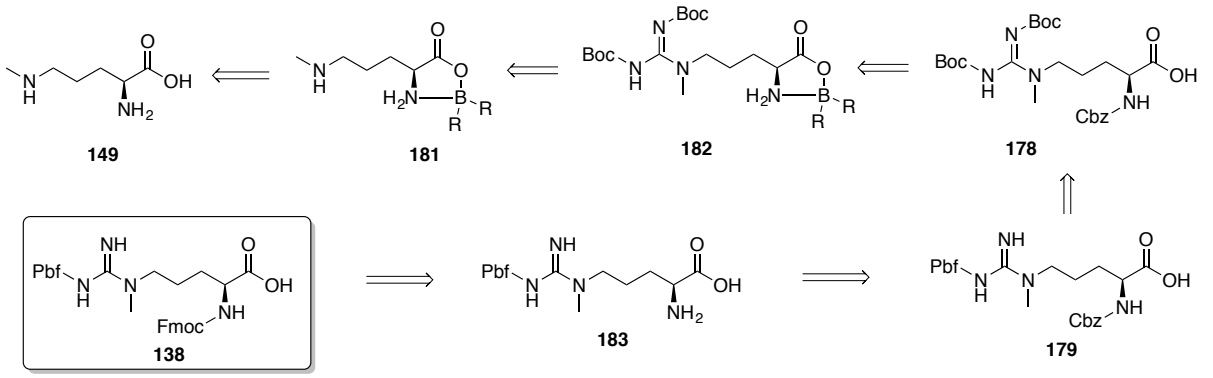
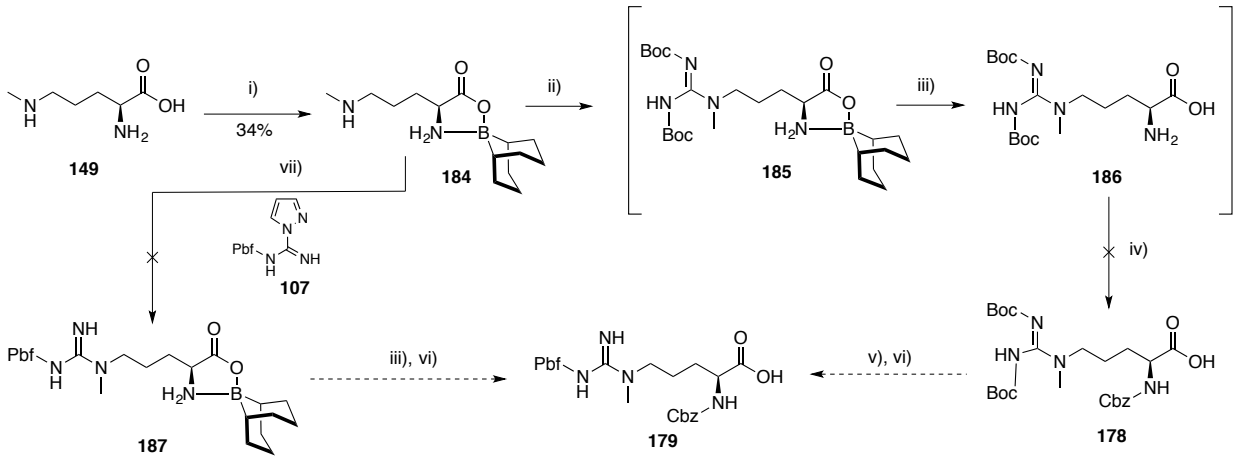


***Figure 3.21** – Retrosynthetic analysis towards protected MMA(δ) **138** via complexation with boron.*

A subsequent report by Schade *et al.*[163] that uses boron to similar effect showed use of 9-borabicyclo[3.3.1]nonane (9-BBN) to be preferential over triethylborane because of its facile deprotection in weak base. Certainly in the envisaged synthesis towards protected MMA(δ) **138**, decomplexation of this borane complex would be beneficially achieved in base rather than acid to retain the planned guanidine Boc-protection (**Scheme 3.20**).



***Scheme 3.22** – Attempted synthesis of protected derivatives of MMA(δ) via a 9-BBN complex. Reagents and conditions: i) 9-BBN, MeOH, 65 °C, 1.5 h ii)* bis*-Boc-pyrazole-1*H*-carboxamidine,DIPEA, MeOH iii) $NaOH_{(aq)}$ (2 M) iv) Cbz-Cl,* N*-hydrozysuccinimide v) TFA:CH_2Cl_2 vi) Pbf-Cl, $NaOH_{(aq)}$ vii) DIPEA.*

Synthesis of the 9-BBN complex **184** proceeded in 34% yield by treatment of MMO(δ) **149** with 9-BBN at 65 °C (**Scheme 3.22**). Despite the poor results observed in **Section 2.6.3** using the Pbf-

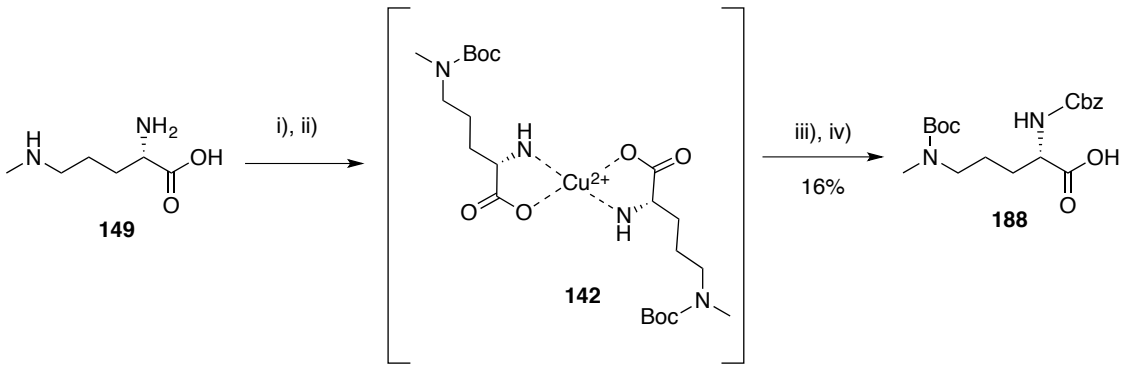
protected pyrazole **107** as a guanylating agent, a test reaction was attempted using *N*-Me-Orn(9-BBN) **184** as the starting material. It was thought that a monomethylated amine might be more nucleophilic towards the putative guanylating reagent but unfortunately this small scale reaction, carried out at room temperature and with DIPEA, was not successful and only starting material was observed in ^{1}H NMR analysis of the crude mixture. This was in part attributed to poor solubility properties of the pyrazole **107** leading to a heterogeneous reaction mixture. Optimisation was not attempted in the context of the poor success previously observed for Orn in **Section 2.6.3**.

Guanylation of **184** to the protected Arg-derivative **185** was achieved using *bis*-Boc-pyrazole-1*H*-carboxamidine **109** and DIPEA. The resultant crude product mixture was used without purification and was treated with $NaOH_{(aq)}$ (2 M) which was successful in breaking the borane complex to form the free amino acid **186**. The majority of contaminating organics were removed *via* Dowex® purification and **178** eluted from the column with $NH_4OH_{(aq)}$ (1 M) but the cleaved by-product, 9-BBN-OH, was not separable from the desired product **178** under these conditions as assessed by ^{1}H NMR. Attempts were made to derivatise **178** to α-Cbz-protected **179** in the hope that acid-base extraction, which had been successful in isolating orthogonally protected Orn derivatives (**Section 2.5.2**), might have some application here. The amino acid **178** was treated with *N*-hydroxysucinimide and Cbz-Cl in 10% aqueous Na_2CO_3, emulating the conditions that had previously been successful for similar transformations with Orn (**Section 2.5.2**). Unfortunately, the procedure was unsuccessful; ^{1}H NMR of the crude organic extract showed a large amount of unreacted Cbz-Cl alongside a complex mixture of decomposition products, within which starting material **178** and product **179** could not be detected.

Rather than attempt optimisation of this procedure, it was anticipated that the highly successful orthogonal protection method used for Orn-OH **69** (**Section 2.5.2** and **Section 3.4.2**) could instead be applied here. The remaining quantity of MMO(δ) **149** was therefore used in an orthogonal dual-protection reaction *via* copper complexation, described below.

3.4.4 – Copper Complexation to Achieve Protected MMO(δ)

The method described in **Section 2.5.2** was utilised for the orthogonal protection of MMO(δ) **149**. MMO(δ) **149** was treated with $Cu(OAc)_2$ to form the copper complex that shields the α-groups, and Boc_2O was added towards production of **142**, which was anticipated to be a chalky blue precipitate isolatable *via* filtration under *vacuum* (**Scheme 3.23**). Unfortunately no precipitate was observable and any attempts to isolate a precipitate produced a negligible amount of material, meaning an aqueous work-up procedure was required to remove residual Boc_2O.



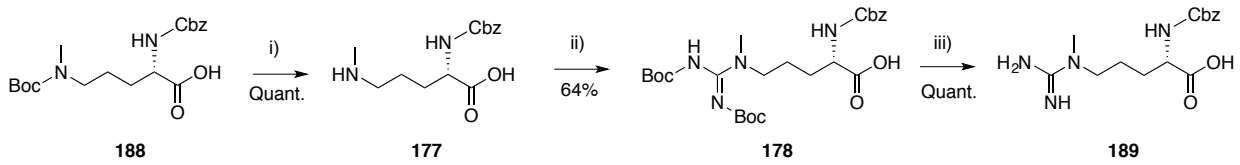
Scheme 3.23 *– Synthesis of dual protected* N-*Boc,*N-*Me-Orn(Cbz)-OH* ***188*** *from MMO(δ)* ***149*** *as part of the total synthesis of protected MMA(δ)* ***138*** *for SPPS. Reagents and conditions: i)* $Cu(OAc)_2$, NaOH(aq) (2 M) ii) Boc_2O iii) 8-quinolinol iv) N-hydroxysuccinimide, Cbz-Cl, $Na_2CO_{3(aq)}$, 16 h, RT.

8-Quinolinol was added to the reaction mixture and reacted for 2 h to liberate the copper-complexed amino acids and contaminating organics were extracted into EtOAc and discarded. A

sample of the crude aqueous phase was dried and analysed by ^{1}H NMR experiments, which revealed that the reaction mixture contained the desired product δ-*N*-Boc, *N*-Me-Orn-OH **188**. Gratifyingly, HMBC analysis confirmed that Boc protection had been selective for the δ-position.

The crude mixture was treated with Cbz-Cl and *N*-hydroxysuccinimide and reacted for 16 h. The reaction mixture was concentrated and then extracted with EtOAc; following extraction the desired product was not sufficiently clean, unlike the other derivatives of amino acid synthesised using this method (**Section 2.5.2**). Crude **188** was purified on normal-phase SiO_2 doped with 1% Et_3N. The eluted product required washing in 1 M $HCl_{(aq)}$ to remove Et_3N; subsequent re-extraction into EtOAc afforded pure *N*-Boc, *N*-Me-Orn(Cbz)-OH **188** in an overall yield of 16% (2.92 mmol) (**Scheme 3.23**).

3.5 – δ-Monomethyl Ornithine to δ-Monomethyl Arginine



***Scheme 3.24** – Synthesis of Cbz-protected MMA(δ) **189** from dual-protected MMO(δ) **188**. Reagents and conditions: i) TFA:CH_2Cl_2 (3:7), RT, 30 min ii)* bis-*Boc-pyrazole-1*H*-carboxamidine, DIPEA, MeOH, RT, 16 h iii) TFA:CH_2Cl_2 (3:7), RT, 2 h.*

The dual-protected **188** was Boc-deprotected by reaction with TFA:CH_2Cl_2 at room temperature in quantitative yield to afford **177** (**Scheme 3.24**). This was treated with *bis*-Boc-pyrazole-1*H*-carboxamidine **109** to afford *N,N'-bis*-Boc-, *N*-Me-Arg(Cbz)-OH **178** in 64% yield following purification on Et_3N-doped SiO_2. The tri-protected MMA(δ) **178** was selectively deprotected *via* reaction with TFA:CH_2Cl_2 to give *N*-Me-Arg(Cbz)-OH **189** in quantitative yield.

Attempts were made to assess enantiopurity at this stage in the synthetic route using chiral solvating/complexation agents. The specific rotation of MMO(δ) **149** had already been measured and shown to be comparable to a literature standard (**Section 3.4.2**). None of the reaction steps after the production of MMO(δ) **149** were anticipated to scramble the stereocentre of the amino acid, but assessment of enantiopurity at this later stage in the synthetic route was nevertheless attempted.

It was hoped that *N*-Me-Arg(Cbz)-OH **189** would complex with a racemic sample of chiral solvating agent to form diastereomers that could be resolved by NMR and allow estimation of *e.e.* by integration. Despite not being the final product, *N*-Me-Arg(Cbz)-OH **102** was chosen as the candidate for assessing enantiopurity. Firstly, the unmethylated analogue Arg(Cbz)-OH **102** is commercially available and test experiments could be carried out using this model compound to establish if chiral resolution with individual solvating agents is possible as a predictor for success with the much more precious *N*-Me-Arg(Cbz)-OH **189**.

Chiral solvating/complexation agents with a primary amine functionality are reported to give the best results for resolution of carboxylic acids[172,173]. Attempts were made to complex **102** with racemic (±)α-methylbenzylamine **190** but this approach failed to resolve two diastereomeric salts by ^{1}H NMR (**Figure 3.1**).

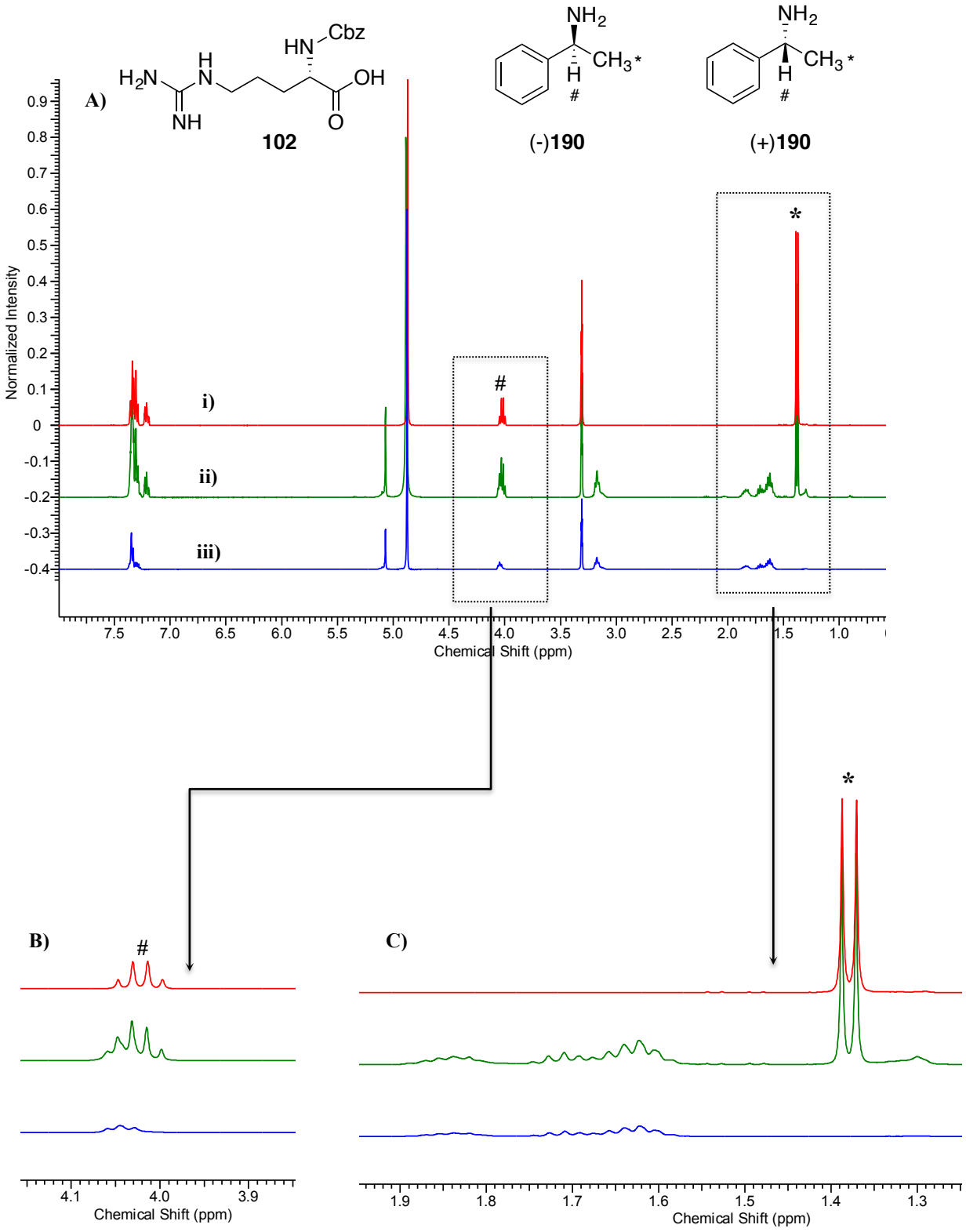
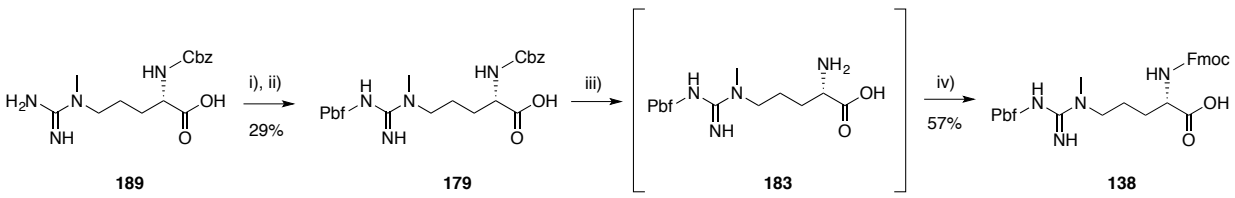


Figure 3.1 *– Attempt to use the chiral solvating agent α-methylbenzylamine* ***190*** *to estimate the enantiopurity of model compound L-Arg(Cbz)-OH* ***102****. **A)** Overlay of ^{1}H NMR spectra of **i)** racemic (+)**190** & (-)**190** (red) **ii)** experiment containing 1:1 L-Arg(Cbz)-OH **102** and racemic (±)**190** (green) and **iii)** enantiopure L-Arg(Cbz)-OH **102** (blue). Zoomed in regions of **B)** multiplet # and **C)** doublet * assigned to the indicated protons showing no diastereomeric chemical shifts.*

In order to be used in SPPS, **189** required Pbf protection of the guanidine functionality, and a protecting group exchange for Fmoc at the α-amino group.



***Scheme 3.25** – Successful synthesis of fully-protected MMA(δ) **138** for SPPS. Reagents and conditions: i)Pbf-Cl, acetone, 0 °C ii) 1 h RT (six cycles) iii) 10% Pd/C (20% w/w), H_2, MeOH iv) Fmoc-Cl,* N*-hydroxysuccinimide, 10% aqueous Na_2CO_3, 16 h.*

N-Me-Arg(Cbz) **189** was treated with one equivalent Pbf-Cl in acetone at 0 °C and stirred for 1 h (**Scheme 3.25**). A second portion of Pbf-Cl was added (~0.5 eq.) and stirred for 90 min at 0 °C and a third portion (also ~0.5 eq.) was added and stirred for 30 min at 0 °C. The reaction was allowed to stir for 16 h at room temperature following which it was filtered to remove the side-product Pbf-OH – a phenomenon described in **Section 2.8**. Four further portions of Pbf-Cl (each ~0.5 eq.) were added at 1 h intervals, all at room temperature until TLC and MS analyses suggested the starting amino acid **189** had been consumed. In order to purify the product, it was necessary to include AcOH in the eluent to facilitate chromatography on SiO_2. Since the Pbf group is labile in strong acid conditions, column fractions containing product were pooled and mixed with H_2O prior to evaporation of the volatile eluent solvents (CH_2Cl_2:MeOH) under *vacuum*. Once the solution had been concentrated and only an aqueous phase remained (containing dilute AcOH) a small amount of $HCl_{(aq)}$ was added and the product immediately extracted into CH_2Cl_2. After concentration and drying, this process afforded *N*-Pbf-, *N*-Me-Arg(Cbz) **179** in 29% yield. The closest analogous reaction described in the literature reports a yield of 20% for this step using two equivalents of a different guanidine sulfonyl protecting group 'MIS'[143] (**Section 2.1.1**) added in one portion.

The final stage in this synthesis was to exchange the Cbz-protecting group for α-Fmoc-protection to enable SPPS. *N*-Pbf-, *N*-Me-Arg(Cbz)-OH **179** was treated with H_2, and hydrogenolysis of the Cbz group catalysed by 10% Pd/C. The crude was treated with *N*-hydroxysuccinimide and Fmoc-Cl in a 10% aqueous solution of Na_2CO_3; the final product **138** was isolated *via* purification on SiO_2 in 57% yield over two steps.

Overall, the total synthesis of protected MMA(δ) **138** for SPPS took place over 12 steps and in a yield of 0.76%. The total amount of end product synthesised was 145 mg (0.219 mmol); the SPPS procedure that would be used to produce the peptide containing MMA(δ) **138** was to be conducted on a 0.1 mmol scale (final peptide). Usually each component amino acid is added in 10-times excess (1 mmol each) to encourage high peptide productivity (**Section 2.1.3**). Although 0.219 mmol was not the optimum quantity for SPPS, it could be used in a roughly two-fold excess and so was used in the synthesis of 16mer peptide H4R3-MMA(δ), described in Chapter Four.

Specific rotation was assessed at each step throughout the 12-step total synthesis of **138** described, and at no point were any of the materials racemic; the final compound exhibited an $[\alpha]_D^{25}$ of +5.3 (*c* = 1.0 in MeOH:CH_2Cl_2 55:45) although there was no literature standard for comparison. The exact *e.e.* could not be quantified by chiral gas chromatography mass spectrometry (GC/MS) or chiral HPLC; these techniques require a standard racemic sample. Attempts to synthesise the latter from racemic Orn-OH **69** following **Scheme 3.26** were unsuccessful, failing at step one in the synthesis – the racemic starting material would not readily form a complex with Cu^{2+} and no precipitate could be isolated. Synthesis of an authentic sample of the D-enantiomer proved more successful, but the full 12-step route could not be completed

within the confines of this project. Only one step in the synthesis was of particular concern with respect to racemisation: the global deprotection using refluxing $HBr_{(aq)}$. Following this step, the optical rotation of the product MMO(δ) **149** was assessed and compared to a literature standard; the data were consistent in sign and similar in magnitude.

3.6 – Conclusions and Future Work

This chapter had been concerned with the synthesis of protected derivatives of MMO(δ) that could be manipulated to produce MMA(δ) suitably protected for incorporation into a peptide *via* SPPS. Very few examples of MMO(δ) **149** syntheses are reported in the literature and the few identified procedures had significant shortcomings. Zobel-Thropp *et al.*[31] described the synthesis of an α-*N*-acetyl protected MMO(δ) **147** *via* reductive amination using only formaldehyde. This procedure was not attempted; the authors observe the *N,N*-dimethylated side-product **148** and also supply incomplete characterisation data.

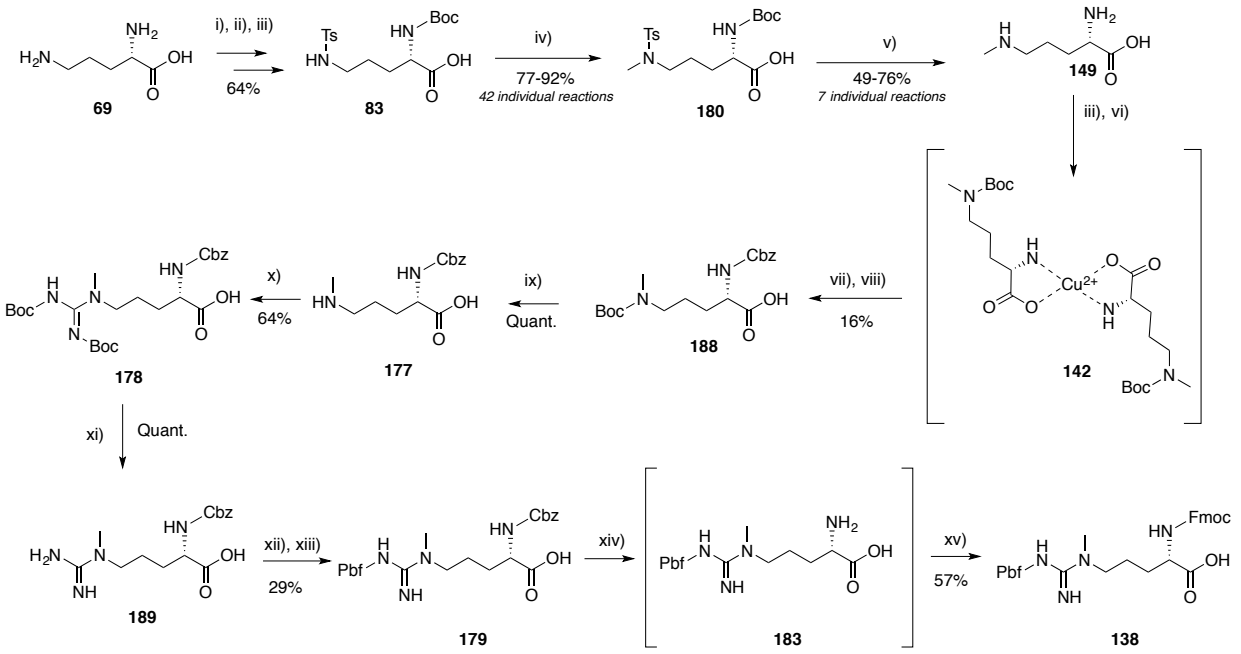
Attempts were made to generate MMO(δ) derivatives *via* δ-*N*-Boc-protected Orn derivatives. Direct *N*-methylation using the alkylating agent iodomethane was unsuccessful, instead generating the corresponding methyl ester as the major product. According to a reported procedure by Phuan *et al.*[164], reduction of the Boc group was attempted using $LiAlH_4$ but this was also unsuccessful, generating a mixture of insoluble aluminium salts and causing unwanted reduction of the carboxylate group.

Huang *et al.*[162] had reported a procedure for synthesising MML *via* sequential reductive aminations that reportedly relied upon steric hindrance at this position to control the number of substitutions. Reductive amination was reportedly used to achieve a mono-benzylation at the δ-

nitrogen. The authors used a one-pot procedure to then mono-methylate the same nitrogen *via* reductive amination with formaldehyde. This methodology was applied to Orn(Fmoc)-OMe **86** but with poor success; putative dibenzylation was observed at the δ-nitrogen (**157**). It was thought that this problem was symptomatic of a poor conversion to the benzylimine. Attempts were made to control this problem by altering concentration, temperature, drying agent and by using a reportedly[165] more reactive aldehyde (*p*-anisaldehyde **161**). Unfortunately though, the best imine conversion that could be obtained was 25% (as assessed by ^{1}H NMR), which was considered to be insufficient.

By adapting the sequential reductive amination method described by Huang *et al*[162], a limited amount of desired product *N*-Me, *N*-benzyl-Orn(Fmoc)-OMe **156** was obtained. Subsequently attempts were made to selectively remove the benzyl group to afford *N*-Me-Orn(Fmoc)-OMe **165**. Three different methods using $Pd(OH)_2$, Pd/C and ACEC respectively as debenzylating agents were not successful; the Fmoc group was found to be labile under all conditions tested.

Instead, an approach reported by Luzzi & Marletta[170] was implemented that generates unprotected MMO(δ) **149**. The overall route towards protected MMA(δ) **138** for SPPS is given below (**Scheme 3.26**):



Scheme 3.26 – *Total synthesis of protected MMA(δ)* ***138*** *for SPPS. Reagents and conditions: i)* $Cu(OAc)_2$, $NaOH_{(aq)}$ (2 M), p-toluenesulfonyl chloride, Et_3N ii) 8-quinolinol iii) Boc_2O iv) *MeI, 2 M NaOH v) HBr 48%, reflux vi)* $Cu(OAc)_2$, NaOH(aq) (2 M) vii) 8-quinolinol viii) N-hydroxysuccinimide, Cbz-Cl, $Na_2CO_{3(aq)}$ ix) TFA:CH_2Cl_2 (3:7), RT, 30 min x) bis-*Boc-pyrazole-1*H-*carboxamidine, DIPEA, MeOH, RT, 16 h xi) TFA:*CH_2Cl_2 *(3:7), RT, 2 h xii) Pbf-Cl, acetone, 0 °C xiii) 1 h RT (six cycles: xii-xiii) xiv) 10% Pd/C (20% w/w),* H_2*, MeOH xv) Fmoc-Cl,* N-*hydroxysuccinimide, 10% aqueous* Na_2CO_3*, 16 h.*

Where previous attempts had failed to install a methyl group on a Boc-protected nitrogen (**Section 3.3**) the key to success in this procedure was methylation of a sulfonyl-protected Orn derivative. Building on the work reported in Chapter Two for selective orthogonal protection of Orn, the intermediate *N*-Ts-Orn(Boc)-OH **83** was synthesised on large scale (24.6 g) and good yield (64%). This represented a significant improvement on Luzzi & Marletta's 16% yield for the analogous *N*-Ts-Orn(Bz)-OH **170**. Methylation of **83** proceeded in good yield using iodmethane as the methylating agent, but the disadvantage to this step is the requirement for a sealed reaction system that confers a pressurised environment that is presumed to prevent escape of iodomethane. In order to produce *N*-Me, *N*-Ts-Orn(Boc)-OH **180** on a sufficiently large scale for subsequent reactions, it was necessary to carry out 42 individual reactions in sealed tubes. This is clearly a very time-inefficient step but attempts to scale-up this procedure required use of

an open-reaction system, which had a detrimental impact on reaction yield (55% compared to 77-92%). Optimisation of this step, perhaps requiring investigation into alternative pieces of equipment, should be a priority for subsequent iterations of this synthesis.

Deprotection of *N*-Me, *N*-Ts-Orn(Boc)-OH **180** required refluxing in conc. $HBr_{(aq)}$; this was necessitated by the non-labile tosyl group. This step represents a significant bottleneck in the reaction efficiency because it generates naked MMO(δ) that then requires a series of orthogonal protection/deprotection steps, which is costly for both yield and time. Attempts were made to circumvent this bottleneck by using a more labile sulfonyl-protecting group in place of the tosyl group. The 2,4-dinitobenzenesulfonyl (DNs) protecting group was reported to be labile under mild conditions[171] and so was introduced to the δ-nitrogen of Orn. Unfortunately attempts to install a methyl group on the same nitrogen caused cleavage of the DNs group. It is recommended that future work should investigate use of a reportedly[171] intermediary labile 2- or 4- (mono)nitrobenzenesulfonyl (Ns) protecting group, which could make for a more efficient overall synthetic route, bypassing the need to selectively reprotect/deprotect MMO(δ) which was inefficient with respect to yield. Nevertheless, the reprotection of MMO(δ) to *N*-Me, *N*-Boc-Orn(Cbz)-OH **188** proceeded successfully, although in much poorer than yield (16%) than expected. *N*-Me, *N*-Boc-Orn(Cbz)-OH **188** required purification, but was successfully taken forward for guanylation and protecting group exchanges that eventually yielded 0.219 mmol of the final product *N*-Me *N*-Pbf-Arg(Fmoc)-OH **138** (**Scheme 3.25**). The second bottleneck in this overall 12-step procedure occurred during Pbf-protection of the guanidine group. An overall 29% yield was obtained from treatment of precursor *N*-Me-Arg(Cbz)-OH **189** with multiple quantities of Pbf-Cl followed by purification on SiO_2. It is postulated that Pbf-Cl breaks down to the sulfonic acid Pbf-OH **118** in the aqueous conditions necessitated by this reaction.

Optimisation of this step is much-needed for synthesis of any Arg for SPPS on large scale and future work might wish to make investigations here.

After a 12-step synthesis (**Scheme 3.26**) the overall yield of the route was 0.75%. Areas for improvement have been identified and discussed, and should be viewed with excitement for the future, since optimisation of this route could allow access to novel methylation patterns incorporating both δ-Me modification and ω-Me modifications if combined with use of the prospective guanylating agents identified in Chapter Two.

This amino acid *N*-Me-, *N*-Pbf-Arg(Cbz)-OH **138** was used in the SPPS of an H4R3 peptide, the synthesis of which is described in the following chapter.

Chapter 4 – Assay Development and Biological Investigations of Novel Epigenetic Markers

This chapter describes the development of a MS-based assay to quantify arginine-residue methylation status, and its implementation in investigations concerning novel methylated substrates for PRMT1

4.1 – Introduction

In **Section 1.4**, the various advantages and disadvantages of different non-MS based literature assay techniques for assessing PRMT activity were discussed. The major shortcomings of many of these assays can be summarised as follows:

1. Poor or unknown selectivity in distinguishing Args with different methylation number
2. Insensitivity of the assay readout methods
3. Possible variability between readouts from the same assay

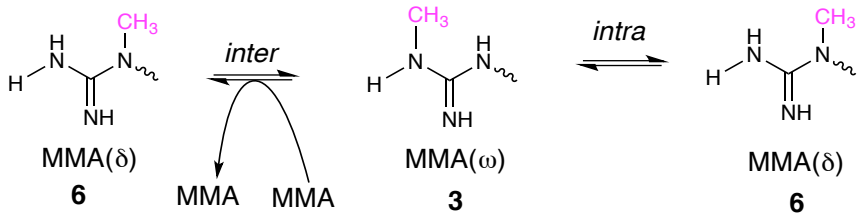
The most notable advantage of MS is its ability to detect and enable quantification of un-, mono- and di- methylated arginine-containing substrates/products as well as possible higher-order methylation. Polymethylation of arginine-residues has not previously been observed; the vast majority of currently employed assay techniques are not capable of distinguishing higher order methylation number. Furthermore, methylation patterns have previously been defined by isolating methylated products of PRMT-catalysed reactions and comparing to known standards of un-, mono- and di-methylation patterns[19,21-23]. The absence of any standards of polymethylated Arg residues suggests polymethylation may have been overlooked. An extension

of this postulation is that MMA(δ), an Arg modification found in yeast conferred by the PRMT homologue RmtA[31], could be a substrate for or inhibitor of human PRMTs.

4.2 – Chapter Aims

This chapter describes work directed towards the following specific aims:

1. Design and optimisation of a new MALDI-TOF-MS-based PRMT methylation assay.
2. Investigate possible polymethylation of arginine by 'forcing' the PRMT1-catalysed reactions.
3. Test a peptide incorporating MMA(δ) for substrate activity *vs.* PRMT1.
4. Investigate the possible rearrangement of MMA(δ) to give MMA(ω) (**Scheme 4.1**)



***Scheme 4.1** – Possible inter- or intra-molecular rearrangement between MMA(ω) **3** and MMA(δ) **6**.*

4.3 – Assay Development

4.3.1 – Protein

It was decided to use a model PRMT for experiments described in this chapter; human PRMT1 was selected because it is well-characterised and is the founding member of the PRMT family[62] with established procedures for its recombinant expression and purification[174].

Recombinant human PRMT1 with a hexa-histidine tag at its *N*-terminus was produced in *Escherichia coli* cells using the pNIC28-BSA4 vector conferring kanamycin-resistance as previously reported[174]. The vector was introduced to ultra-competent BL21 *E. coli* cells *via* heat-shock transformation, which were then incubated on kanamycin-containing agar plates at 37 °C overnight. An individual colony of cells was picked and grown as part of a large-scale expression using the media outlined in **Table 4.1**; a preferential yield was obtained for expression in Terrific Broth (TB) over tryptone-yeast (2-TY) media.

Entry	Media	Culture Size, L	Total PRMT1, mg	Expression Efficiency, μg/L
1	2-TY	5	0.4	80
2	TB	12	20.3	1,700

Table 4.1 *– PRMT1 recovered from two different expression systems. Protein concentration was calculated using the Beer-Lambert law: a NanoDrop spectrometer was used to measure* A_{280} *and the molar extinction coefficient of PRMT1 was obtained using the ExPASy ProtParam tool*[f].

Following induction of protein expression using isopropyl β-D-1-thiogalactopyranoside (IPTG) cells were harvested and lysed to isolate a crude lysate containing recombinant PRMT1. The protein was purified by affinity chromatography *via* the *N*-terminal hexa-His tag to give 0.4-20.3 mg human PRMT1 as part of two different expression procedures (**Table 4.1**).

N-terminal His-PRMT1 was found to be active using the MALDI-MS assay (*vide infra*) so was used without removal of the hexa-His tag. Expression and purification protocols are detailed in the Experimental Section.

[f] The ExPASy ProtParam tool can be accessed at http://web.expasy.org/cgi-bin/protparam/protparam

4.3.1 – Peptide Synthesis

It was decided that the substrate for these MALDI assays should be based on histone 4 (H4) because it is well established as a major 'epigenetic' substrate of PRMT1[175]. The sequence of the whole H4 protein is given in **Figure 4.1**.

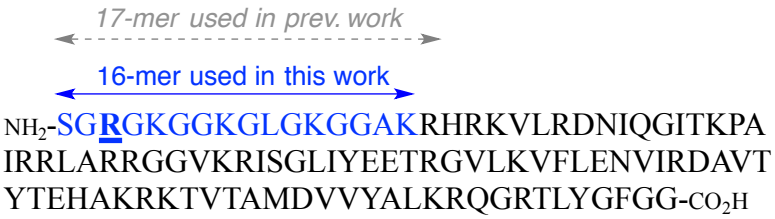


***Figure 4.1** – Amino acid sequence of human H4, showing the 16-mer fragment used for assays reported herein.*

A MALDI-based PRMT assay has been reported for comparing catalytic turnover of different peptide sequences[57]. As part of this study, the authors used MALDI-MS to investigate the time-dependent modifications of peptides. In addition to this, the authors used LC-MS/MS analysis to explore the effect of peptide sequence and length on PRMT1-catalysed methylation; peptide sequences containing more than one Arg residue were tested. Of particular interest to the current project is the finding that a 17-mer fragment of H4, containing Arg residues at positions three and 17 (**Figure 4.1**), could be methylated at both Arg residues. By analysing the LC-MS/MS fragmentation patterns of this peptide, the authors demonstrated that PRMT1 will preferentially methylate *N*-terminal Arg but will also methylate more *C*-terminal positions in multi-Arg containing peptides[57].

Based on these observations, a 16-mer fragment of histone 4 (H4) was chosen as the most appropriate substrate for MALDI-MS assays described in this chapter, because it is the longest *N*-terminal H4 canonical sequence that contains only one Arg residue. It was thought that a

peptide containing multiple Arg residues would unnecessarily complicate the planned investigations into possible polymethylation of H4R3 fragment peptides.

The peptides used for methylation investigations are detailed in **Table 4.2**. Those peptides synthesised in-house were made using automated Fmoc-based solid phase chemistry on a 0.1 mmol peptide scale.

Entry	Peptide	Sequence	Source	Purity, %
1	H4R3	SGRGKGGKGLGKGGAK	Synthesised	100
2	H4R3-MMA(ω)	SGR(ω-Me)GKGGKGLGKGGAK	Thermo Fisher	100
3	H4R3-MMA(δ)	SGR(δ-Me)GKGGKGLGKGGAK	Synthesised	ND
4	H4R3-ADMA	SGR(*N,N*-Me_2)GKGGKGLGKGGAK	Thermo Fisher	96
5	H4R3-SDMA	SGR(*N,N'*-Me_2)GKGGKGLGKGGAK	Thermo Fisher	96
6	H4R3(Δ)	SG-GKGGKGLGKGGAK	Synthesised	100

Table 4.2 *– Peptide sequences used as part of this study. Purities were assessed by analytical LC/MS. ND = not determined due to co-elution with major contaminant of the 15-mer MMA(δ) deletion product (H4R3(Δ)).*

α-Protected Fmoc-amino acids were activated with coupling agents (diisopropylcarbodiimide and hydroxybenzotriazole) and mixed with a rink-amide solid-support resin in 10 times excess (1 mmol/amino acid). Couplings took place over 3 hours with MW irradiation. Arg however cannot be subjected to MW irradiation because it is prone to deguanylation and/or cyclisation as outlined in **Scheme 2.1**. Coupling times for introduction of Arg residues were doubled to 6 hours at room temperature. Inefficient coupling at any stage in the peptide synthesis can lead to production of 'deletion products', which are shorter peptides where an amino acid has been missed from the chain due to an unsuccessful coupling.

Synthesis of the H4R3 standard from component protected amino acids afforded ~15 mg in 100% purity. A trace impurity of trishydroxymethyl(aminomethane) (Tris) buffer was also found to be present in the HPLC trace that was also found in the blank runs preceding and following the sample of interest. The HPLC traces and associated MALDI-MS data as part of these peptide syntheses are given in the **Appendix C**.

The synthesis of the MMA(δ)-containing peptide was less forthcoming. Only 145 mg (0.22 mmol) of the bespoke protected amino acid MMA(δ) could be synthesised as part of this work, as described in Chapter Three, representing only two-fold excess compared to the desired 10-fold excess. To encourage complete consumption of available MMA(δ) the coupling time for this residue was increased to 16 h at room temperature. Attempted purification of the resulting peptide by preparative HPLC afforded a small quantity of desired product (<0.1 mg) alongside a significant contaminant of the 15-mer deletion product. Attempts to estimate the purity of the desired MMA(δ)-containing peptide by analytical LC/MS were unsuccessful because of co-elution of desired product and the 15-mer deletion peptide from the HPLC column.

The poor results were attributed to poor reaction conversion resulting from the low quantity of the MMA(δ) amino acid available. This was diagnosed by presence of a significant impurity that corresponded to the 15-mer deletion product in the HPLC trace, with MMA(δ) absent (see **Appendix C**).

4.3.3 – Assay Optimisation

4.3.3.a – Assay Parameters and Conditions

The conditions defined in **Table 4.3** were used for the PRMT1 assays and mimicked conditions reported by Siarheyeva *et al.*[132] and Smil *et al.*[131].

Assay Component	Volume	Final Assay Conditions
SAM	10 µL	At approximated K_m value: 10 µM (*vide infra*)
Peptide substrate	10 µL	Close to reported[132] K_m value: 15 µM
Enzyme	20 µL	Dependent on specific activity
Buffer	59 µL	20 mM Tris, 10 mM DTT
Inhibitor in DMSO	1 µL	1% DMSO
TOTAL volume	100 µL	

***Table 4.3** – Assay components used the MALDI-MS-based assay described herein.*

Typically SAM-dependent assays are carried out below 37 °C because SAM is unstable even at room temperature[110,176]. As part of the current work, PRMT1-catalysed decline of unmethylated substrate peptide was monitored over a 120 min period at 37 °C and 30 °C. A plateau in PRMT activity was observed after 105 min for both assays, so data up to and including this time point were used to compare reaction progression at 30 °C and 37 °C (**Figure 4.2**). These results demonstrated that the PRMT1-catalysed reaction proceeded to a similar extent under both conditions. Since there were no associated disadvantages to using a lower temperature, which might in fact be favourable for SAM stability, 30 °C was selected as the working temperature for all future assays.

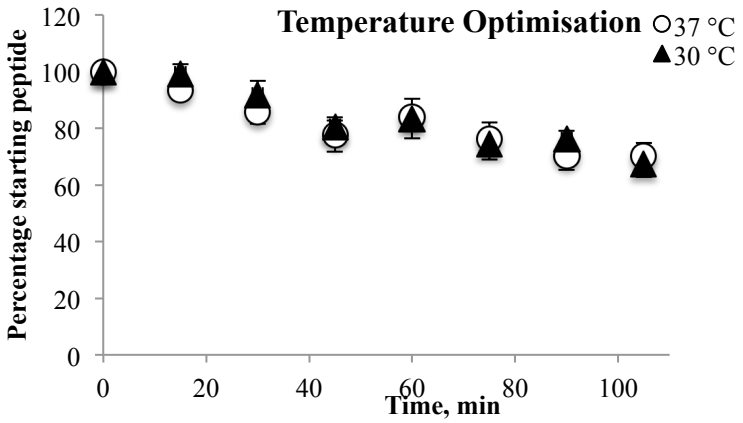


Figure 4.2 *– Temperature optimisation for the PRMT-catalysed reaction. Assay carried out as described in* ***Table 4.3****. Conditions: [PRMT1] = 80 nM, [H4R3] = 15 μM, [SAM] = 10 μM.*

4.3.3.b - Quenching

Various different quenching agents for the PRMT1 reaction were tested – chaotropic salts, acid and organic solvent; use of 7.5 M guanidine HCl interfered with the MALDI readout, as did use of a 10% aqueous TFA solution. Adding MeOH to a final concentration of 50% was found to be effective at quenching the reaction whilst not affecting the MALDI readout.

4.3.3.c - Quantification

MALDI ion peak intensities are not guaranteed to give a linear response with increasing concentration, meaning quantification of assay readout cannot always be achieved directly from the raw data; this represents the biggest limitation of using a MALDI-based assay. A number of methods have been reported to quantify readout but can be time-consuming and/or costly[57,177,178].

Preliminary investigations were made as to the utility of incorporating a single generic standard, in this case the 10-mer peptide Angiotensin I. It was thought it might be possible to quantify histone peptides by constructing a standard curve comparing the ratios of the predominant histone and angiotensin isotopes. Unfortunately this idea was abandoned given the non-linearity observed and the high degree of error encountered (**Figure 4.3**).

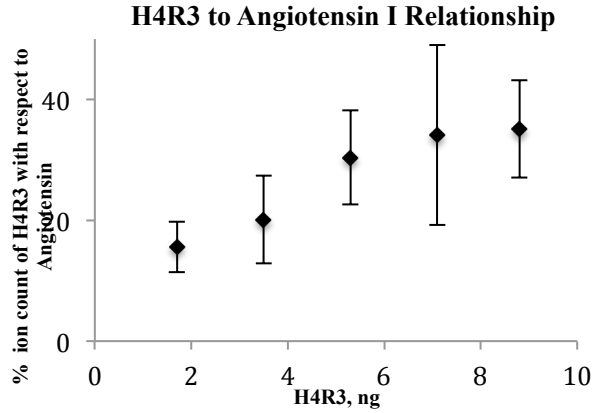


Figure 4.3 – *Percentage ion count of H4R3 compared to a constant 5.3 ng of Angiotensin I where n = 10. Error given as ±standard deviation.*

It was decided to attempt quantification by applying a weighting to the intensity of each ion peak based on the corresponding peptide's propensity to crystallise, ionise and fly. A similar method had been applied in work reported by Gui *et al.*[57]. For peptides with equivalent weighting, the intensities for all peptides of interest could be totalled and each peptide then expressed as a percentage of this (**Figure 4.4**).

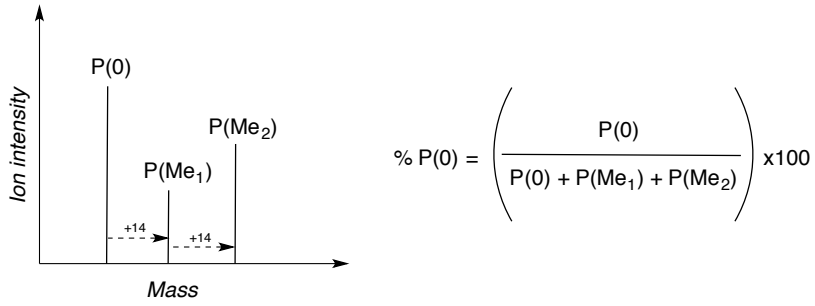


Figure 4.4- *Peak-picking and percentage intensities. P(0) = unmethylated peptide, $P(Me_1)$ = monomethylated peptide and $P(Me_2)$ = dimethylated peptide.*

First, the tendencies of un-, mono- and di-methylated peptides to crystallise, ionise and fly must be assessed. Two homogenous solutions replicating assay conditions were prepared containing absolute and known quantities of the three standard peptides. Samples for each repeat (n = 10)

were individually crystallised with the matrix α-cyano-4-hydroxycinnamic acid (CHCA) and the percentage distribution of their ion intensities was noted following subjection to MALDI-MS.

When ADMA is included (**Figure 4.5A**) each of the means are within two standard deviations of each other, which for the purpose of this work suggests no appreciable difference in intensity is present for each different peptide. This means the assay can be used for quantification of un-, mono-, and asymmetrically dimethylated peptides by giving equal weighting to the peak intensities observed for each peptide and calculating distribution according to **Figure 4.4**. Since PRMT1 is reported to generate ADMA, this assay could be used for accurately monitoring activity for this PRMT homologue.

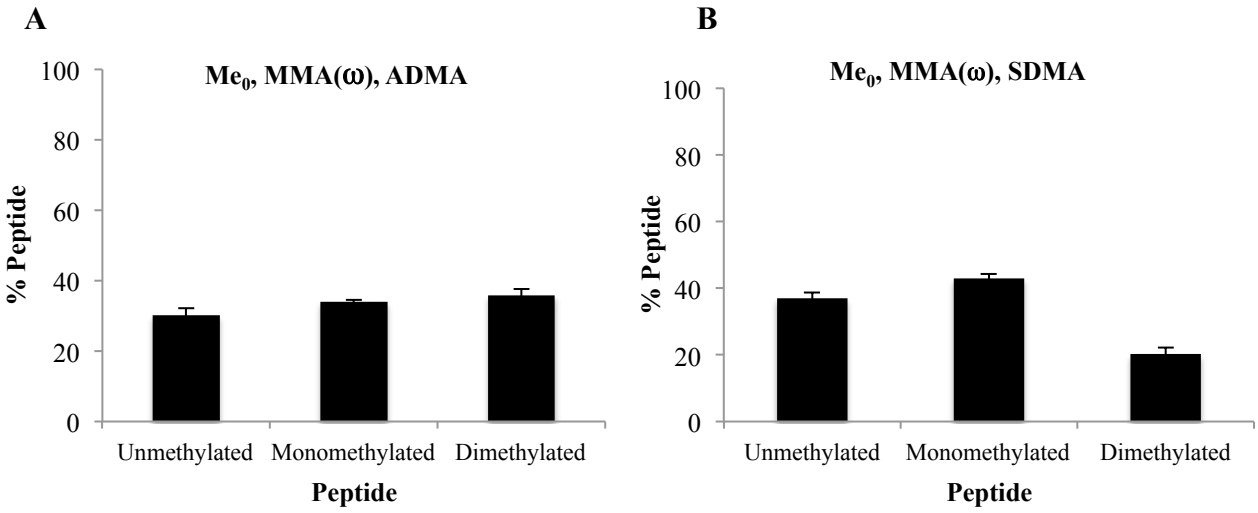


Figure 4.5 – *Distribution of ion peak intensities from a solution of exact and equivalent quantities of peptides with different masses. Unmethylated H4R3 and MMA(ω) were added to each experiment alongside* ***A)*** *ADMA and* ***B)*** *SDMA. Data is given as the mean ± standard deviation and n = 10 for each sample.*

However, a marked difference was recorded for the observed and null distributions when SDMA was used (**Figure 4.5B**). Further tests were carried out to assess the purity of the SDMA peptide, which had been supplied with a certificate of analysis stating 96% purity. Samples of SDMA were analysed by MALDI-MS and showed a 17% (±2%) impurity corresponding to the mass of

the unmethylated peptide. Analytical LC/MS on the SDMA peptide was re-run in house and a theoretical purity of 100% was obtained. Together, these three pieces of data suggest any of the following:

1) An impurity of unmethylated peptide is present in the SDMA peptide sample and the two peptides cannot be distinguished under the HPLC conditions used both in-house and by the supplier
2) The SDMA peptide spontaneously degrades under MALDI-MS conditions and no contamination is present in the sample
3) The apparent contamination of unmethylated peptide in the SDMA sample is simply a fragmentation product produced under MALDI-MS conditions

For assays described below where SDMA is used, this artefact and the non-equivalent peptide distribution is taken into account.

It should be noted that in some of the samples tested as part of this work a low overall ion count was recorded, possibly caused by insufficiently pure or concentrated samples or instrumental artefact. For those experiments showing low overall ion count, it is speculated that this method of calculation might be less reliable because the effect of noise may skew the data. Individual cases where this arose are discussed in subsequent sections.

4.3.3.d – Optimisation for Linearity

A time-course assay was carried out over a 120 min period to identify initial linear velocity conditions. Assays were initiated at time = 0, 15, 30, 45, 60, 75, 90, 105, 120 min and quenched at t = 120 min. Decline of unmethylated starting peptide was plotted as a function of time

(**Figure 4.6**). Linear reaction velocity was observed for the 0-90 min period and so all future assays were carried out for 90 min unless otherwise stated.

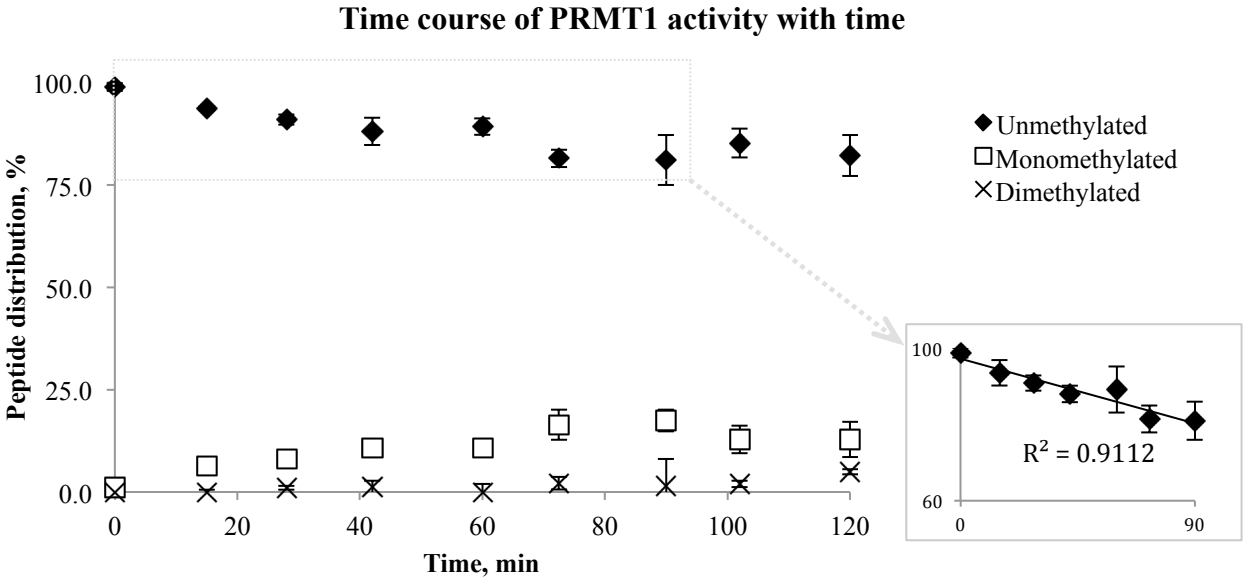


***Figure 4.6** – Time-course assay of PRMT1 at 30 °C to elucidate initial linear velocity. Inset graph shows linear PRMT1 activity in the 0-90 min period. Assay carried out as described in **Table 4.3**. Conditions: [PRMT1] = 80 nM, [H4R3] = 15 μM, [SAM] = 10 μM.*

4.3.3.e – Verifications

The H4R3 peptide K_m was experimentally determined to be 25.8 ± 1.2 μM (two standard deviations) (**Figure 4.7**) using a 20-fold excess of SAM, which is close to the literature estimate of 15 μM[132]. It was not possible to experimentally determine the K_m of SAM because the required excess of peptide overwhelmed the MALDI signal and was outside the limit of sensitivity for the MALDI reader, so the literature value (10 μM) was used[132].

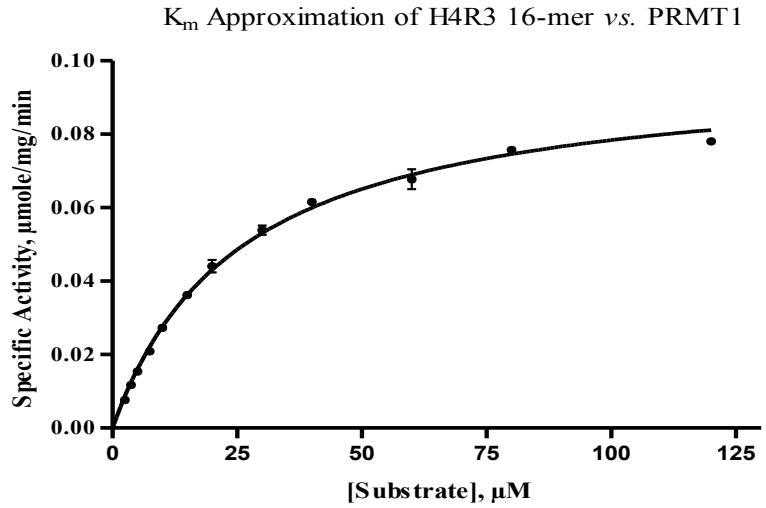


Figure 4.7 *- K_m approximation of H4R3 16-mer and other kinetic parameters calculated using GraphPad Prism software and Michaelis-Menten kinetic analysis. K_m = 25.8 ± 1.2 µM, r^2 = 0.99 (2 sf), k_{cat} = 0.035 ± 0.007 s^{-1}, r^2 = 0.99 (2 sf) and V_{max}= 5.61 nM s^{-1}. Data points are given as the mean ± standard deviation and n = 3 for each. Assay carried out as described in* ***Table 4.3****. Conditions: [PRMT1] = 80 nM. 13 concentration points of peptide H4R3 (0-120 µM) were tested and SAM was added in 20-fold excess for each of these (final [SAM] = 0-2.4 mM).*

The V_{max} was calculated to be 5.61 nM s^{-1} and the k_{cat} was calculated as 0.035 ± 0.007 s^{-1}. **Figure 4.6** above, showing optimisation of assay conditions for linearity, shows low PRMT1-catlaysed turnover of substrate. Insufficient data exists in the literature to allow comparison of PRMT1 specific activity using an H4R3 substrate. However, k_{cat} data presented in **Table 4.4** shows similar reaction rates for PRMT1-catalysed methylation of H4R3 substrates in the literature and in the present study. It is emphasised that k_{cat} values cannot be directly compared, but that k_{cat} can act as a useful gauge for estimating the rate of an enzyme-catalysed reaction[179].

Substrate	**k_{cat} s^{-1}**
Histone H4, full length[180]	0.001
Histone Ac-H4, amino acids 2-21[137]	0.008
Histone Ac-H4, amino acids 2-21[126]	0.009
Histone H4, 16-mer *(this study)*	0.035

Table 4.4 *– Comparison of k_{cat} values from literature reports and the present study.*

Assay parameters and conditions were found to faithfully replicate the literature IC_{50} value[99] for the potent pan-PRMT inhibitor Sinefungin **32** (**Figure 4.8**), which acts as a cofactor mimetic. This assay and the applied parameters and quantification procedures were therefore deemed applicable for the study of two series of inhibitors, outlined in Chapter Five, and for substrate analyses, outlined below.

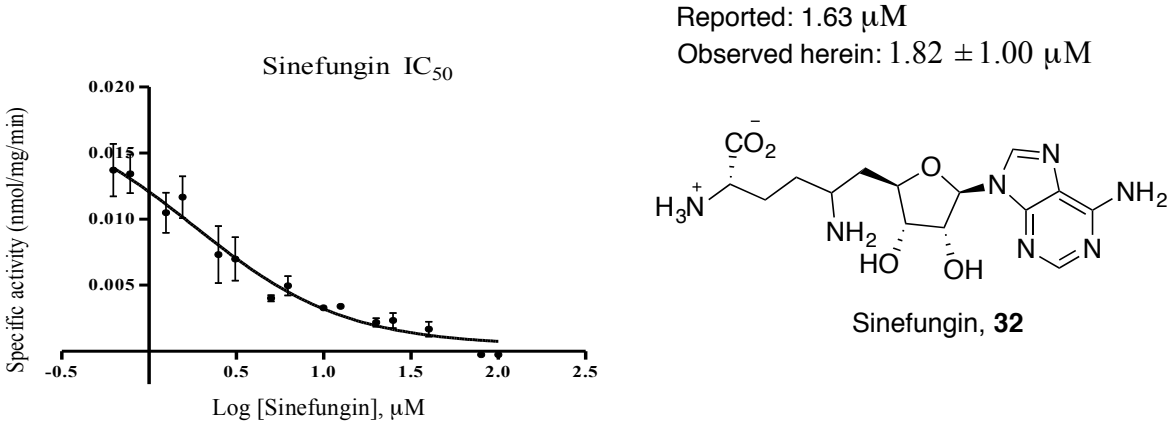


Figure 4.8 *– IC_{50} approximation of Sinefungin **32*** vs. *PRMT1 calculated using GraphPad Prism software. IC_{50} = 1.82 ± 1.00 μM, r^2 = 0.86 (2 sf). Data points are given as mean ± standard deviation and n = 3 for each. Reported IC_{50} from Cheng* et al[99]*. Assay conditions: [PRMT1] = 80 nM, [H4R3] = 15 μM, [SAM] = 10 μM.*

Exemplar readouts showing ion count output from control and test experiments from the MALDI assay are detailed in **Appendix D**.

4.3.3.f – Miniaturisation Assay

For substrate investigations a smaller total assay volume of 20 μL was used in order to conserve the precious methylated peptides. Conditions were identical to the larger scale assay and specific activity was found to be unaffected (<1 standard deviation of each other's means) compared to the larger assay volume that had already been optimised for linearity (**Figure 4.6**).

Assay Component	Volume	Final Assay Conditions
SAM	5 µL	At approximated K_m value: 10 µM
Peptide substrate	5 µL	Close to K_m value: 15 µM
Enzyme	5 µL	Dependent on specific activity
Buffer	5 µL	20 mM Tris, 10 mM DTT, 1% DMSO
TOTAL volume	20 µL	

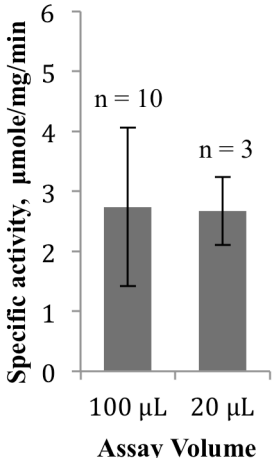


Figure 4.9 – *Conditions for small-scale assay set up and comparison of specific activities observed for small and large scale assays showing faithful replication of optimised assay conditions. Data given as mean ± standard deviation and carried out in at least triplicate. Assay conditions stated in table.*

4.4 – Investigating Novel Methylation Patterns

4.4.1 – Exploring Polymethylation Using Known Human Methylated Peptides

Initially, all peptides shown in **Table 4.2**, bar the MMA(δ) peptide, were tested as PRMT1 substrates under the established small scale assay conditions but with 20-times excess of SAM (with respect to peptide) to encourage reaction progression. Individual MALDI traces were searched for evidence of polymethylation. Ions corresponding to tri- and tetra-methylated Arg were not observed under these standard conditions; the resulting peptide distributions from this assay are given in **Figure 4.10A**. Controls with no enzyme are given in **Figure 4.10B**, which account for the impurity/artefact found in the SDMA peptide data described in **Section 4.3.3.b** and also shows an artefact for the MMA(ω) peptide that is was analysed to be 100% pure by HPLC.

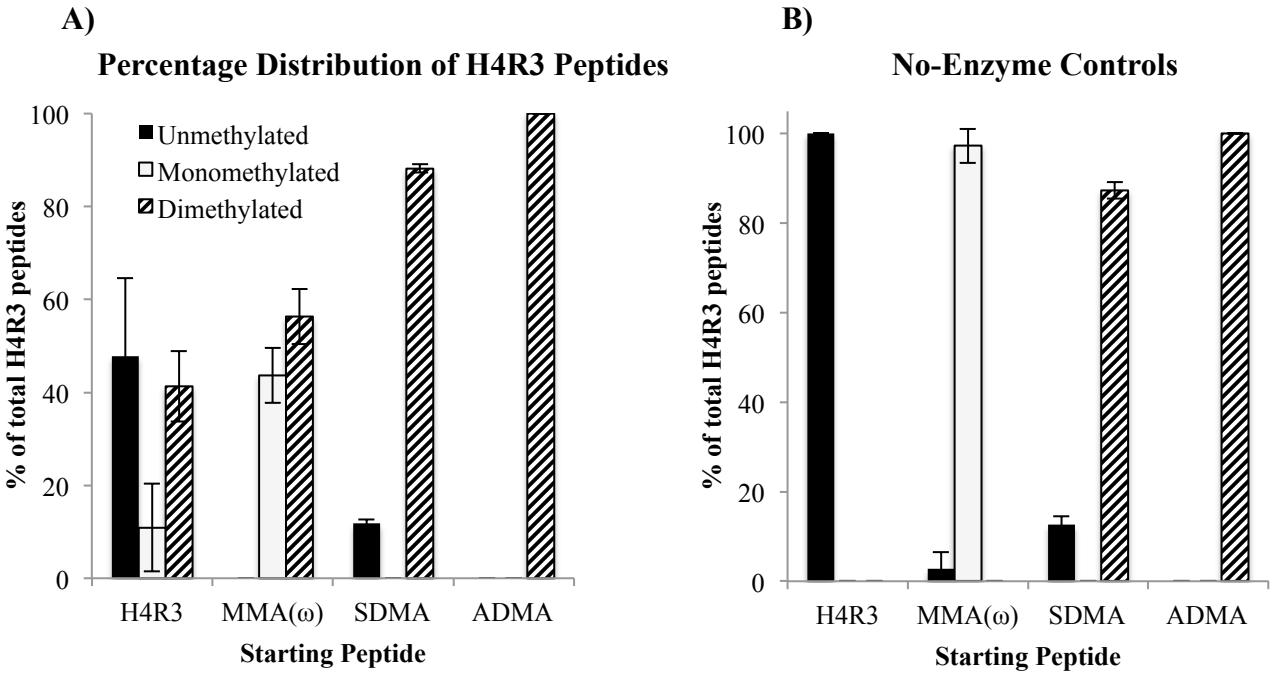


Figure 4.10 – *Percentage distribution of peptides following incubation with PRMT1 and a 10-fold excess of SAM for 90 min. Data given as mean ± standard deviation and n = 3 for each. Assay conditions: [PRMT1] = 640 nM, [peptide] = 15 μM, [SAM] = 300 μM.*

No detectable tri- or tetra- methylation was observed under the initial attempted conditions; attempts were next made to push PRMT1 to polymethylate the different peptides. Experiments were again carried out with a 20-times excess of SAM (with respect to peptide) but this time with a 1:1 ratio of peptide to PRMT1 (usually 23:1) over an extended reaction time of 22 h. All samples were initiated at t = 0 h and analysed at time points 1.5, 3, 4.5, 6 and 21 h (**Figure 4.11**). Suitable controls were included to ensure the enzyme had not lost catalytic ability and that SAM had not degraded over the 6 h time-course although it is possible that PRMT1 and/or SAM could have degraded in the 6 h – 21 h period.

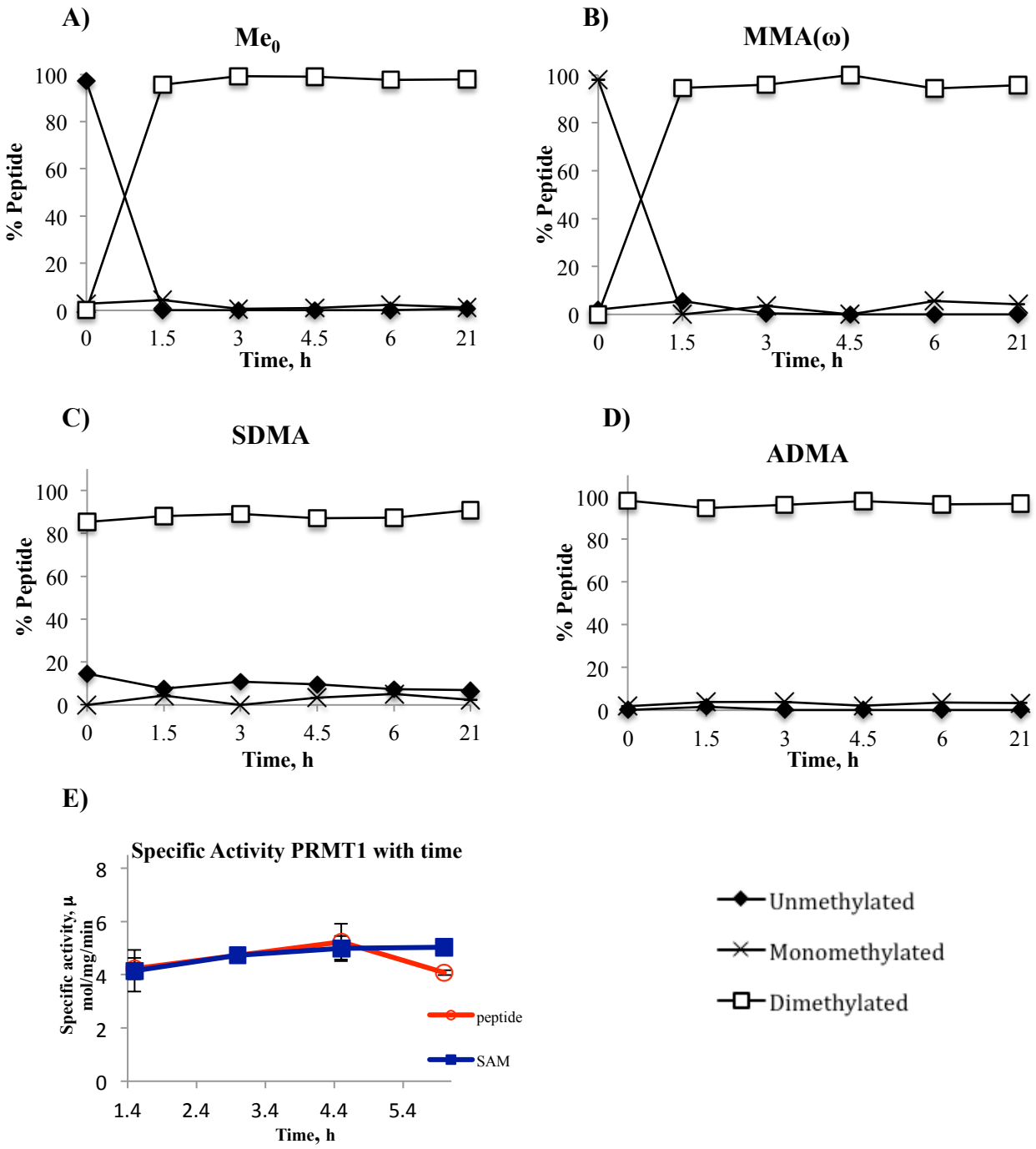


Figure 4.11 *– Peptide distribution with time when* ***A)*** *Unmethylated,* ***B)*** *ω-monomethylated* ***C)*** *asymmetrically dimethylated and* ***D)*** *symmetrically dimethylated H4R3 peptides were used as substrates for PRMT1-catalysed methylation reaction using a 1:1 ratio of peptide to enzyme.* ***E)*** *Shows PRMT1 remains active throughout the experiment. All data points given as mean ± standard deviation and n = 3 for each. Assay conditions: [PRMT1] = 640 nM, [peptide] = 640 nM, [SAM] = 12.8 μM.*

As expected based on literature reports[57], those reactions using known PRMT1 substrates H4R3 (**Figure 4.11A**) and H4R3-MMA(ω) (**Figure 4.11B**) quickly showed conversion to the

dimethylated peptide. As before however, no peaks with masses corresponding to putative tri- or tetra-methylation patterns were observed.

A final attempt was made to push the reaction even harder by using an excess of protein (1:4 peptide:protein) under otherwise identical conditions (**Figure 4.12**).

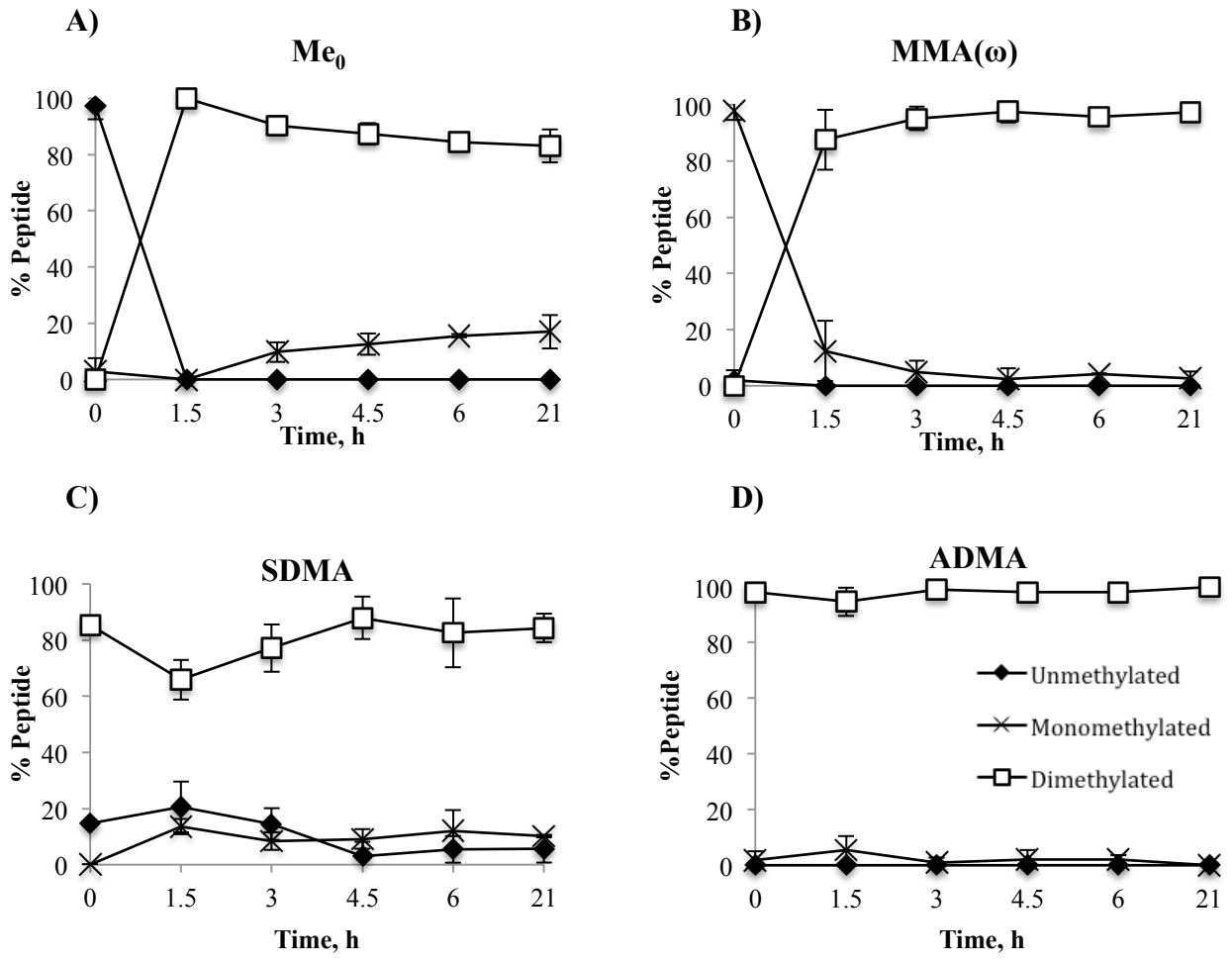


Figure 4.12 – *Peptide distribution with time when **A)** Unmethylated, **B)** ω-monomethylated **C)** asymmetrically dimethylated and **D)** symmetrically dimethylated H4R3 peptides were used as substrates for PRMT1-catalysed methylation reaction using a 1:4 ratio of peptide to enzyme. All data points given as mean ± standard deviation and n = 3 for each. Protein activity data identical to **Figure 4.11** above. Assay conditions: [PRMT1] = 640 nM, [peptide] = 160 nM, [SAM] = 3.2 μM.*

Again no higher order methylation could be observed in any case. Using unmethylated H4R3 as a substrate, **Figure 4.12A** shows a rapid production of dimethylated peptide followed by a

gradual and small decline alongside an associated rise in monomethylated peptide. It was speculated that the apparent decline in dimethylated peptide between t = 1.5 h and t = 21 h might not represent a real decrease of this peptide, so a two-tailed Mann-Whitney U-test[181] was used in an attempt to assess the significance of the difference in SDMA levels at t = 1.5 h and t = 21 h. A significance value of p = 0.064 was calculated (using GraphPad Prism software); where p>0.05, the two data sets are considered not significantly different at the 95% confidence interval. It is emphasised that the low sample size (n = 3 for each data set) could mean the calculated value of p = 0.064 is skewed, but this result gives a tentative indication that the measured decline is not statistically significant.

It seems unlikely that spontaneous non-enzymatic demethylation is occurring given the stability of the *N*-Me bond[36,37]. The most likely explanation for the observation described above is that the assay is suffering from an artefact – in particular, the total ion counts obtained for the MALDI readings as part of this series of reactions were low (**Appendix E**). Since calculation of peptide distribution is dependent on totalling the ion count for each experiment, those experiments with weak overall signal (i.e. low ion count) are naturally more prone to error and artefacts. Future work should look into circumventing this problem as it represents one of the major drawbacks to this MALDI-MS assay.

4.4.2 – Testing MMA(δ)

The yeast methylation pattern MMA(δ) was then tested as a substrate for human PRMT1, first under standard assay conditions with a 10-times SAM excess and secondly under the 'pushed' reaction conditions described above for non-novel substrates.

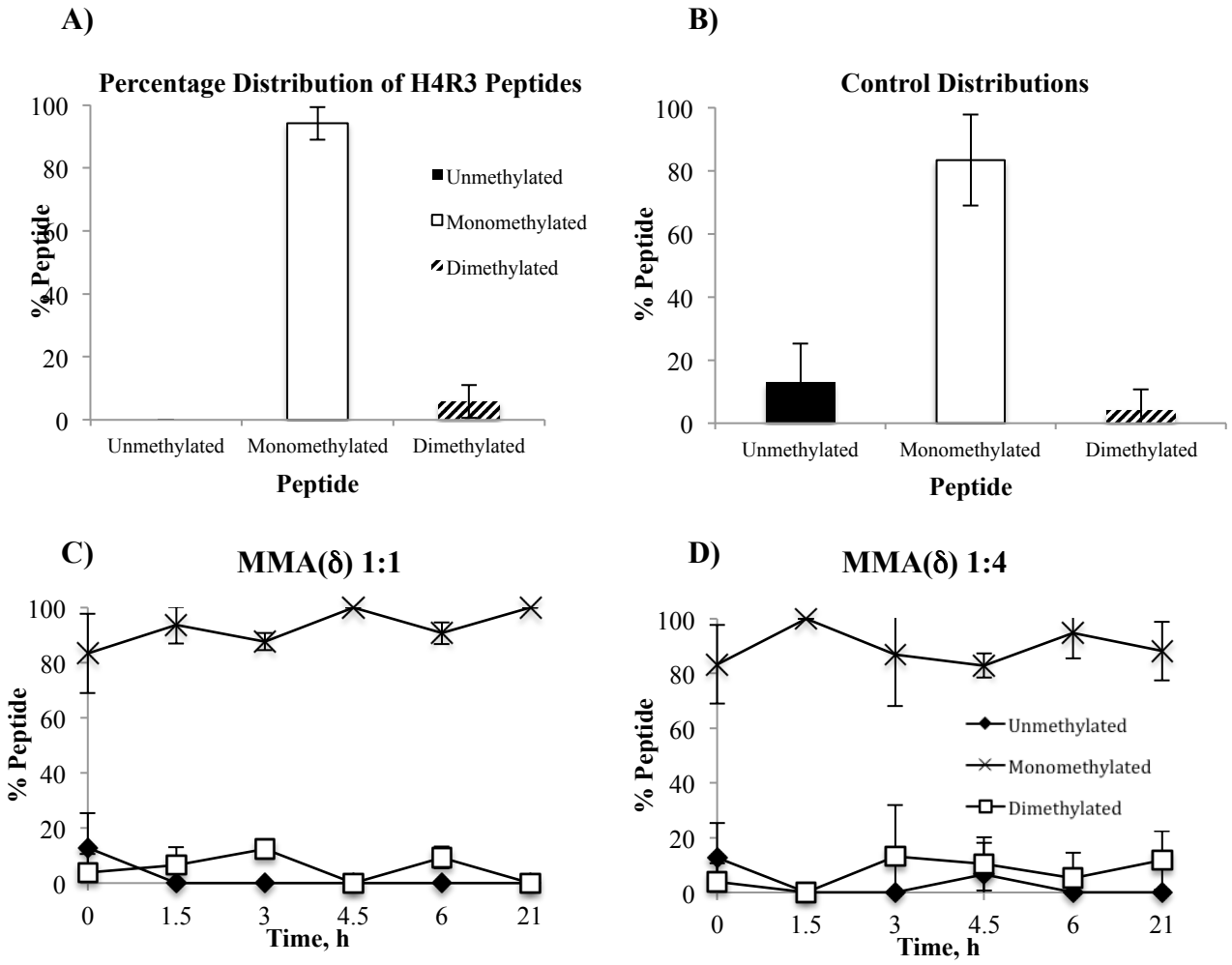


Figure 4.13 – *Peptide distribution with time when δ-monomethylated peptide is used as substrate for the PRMT1-catalysed methylation reaction under **A)** standard conditions (23:1 peptide to enzyme, 1.5 h): [PRMT1] = 640 nM, [MMA(δ)] = 15 μM, [SAM] = 300 μM. and **B)** the respective control distribution. **C)** 1:1 (peptide to enzyme) conditions: [PRMT1] = 640 nM, [MMA(δ)] = 640 nM, [SAM] = 12.8 μM and **D)** 1:4 (peptide to enzyme) conditions: [PRMT1] = 640 nM, [MMA(δ)] = 160 nM, [SAM] = 3.2 μM. All data points given as mean ± standard deviation and n = 3 for each.*

The spectra in **Figure 4.13** suggest that little to no activity was observed. It is noted however that this series of experiments also suffered from a low total ion count (**Appendix E**). Two recommendations are made: firstly, more concentrated MALDI-MS samples should be used to try and overcome the problem of low ion count. Secondly, this assay should be repeated on a larger scale such that any putative products can be isolated and characterised by accurate mass

LC/MS. Unfortunately, given the limited availability of the MMA(δ) amino acid and peptide, neither of these options were possible as part of this project.

The only difference between the H4R3 and MMA(ω) 16-mer fragment peptides of H4, that were shown to be PRMT1 substrates in this work, is the methylation pattern on the δ-nitrogen of Arg. One explanation for lack of MMA(δ) turnover is that the peptide is not a substrate and that the MMA(δ) methylation pattern is either unable to fit in the PRMT1 active site or does not bind in the correct conformation for catalysis. Schematics from mechanistic studies of the type I PRMTs suggest that interactions between the δ-nitrogen of the Arg substrate and conserved active site Glu residues are important for orienting the substrate for nucleophilic attack on SAM (**Section 1.2.4**)[51,54]. There is a possibility that a δ-monomethyl pattern abrogates this postulated interaction, which could mean that a MMA(δ) is not able to act as a PRMT1 substrate.

It is however possible that the observed inactivity of the MMA(δ) peptide is a false negative result. The MMA(δ) peptide used was of unknown purity and contained a significant contamination of the 15-mer deletion product as described in **Section 4.3.1**. To ascertain whether or not the deletion product might be acting as a PRMT1 inhibitor, and therefore masking any possible MMA(δ) turnover, an IC_{50} experiment was run with a standard of the 15-mer (**Figure 4.14**).

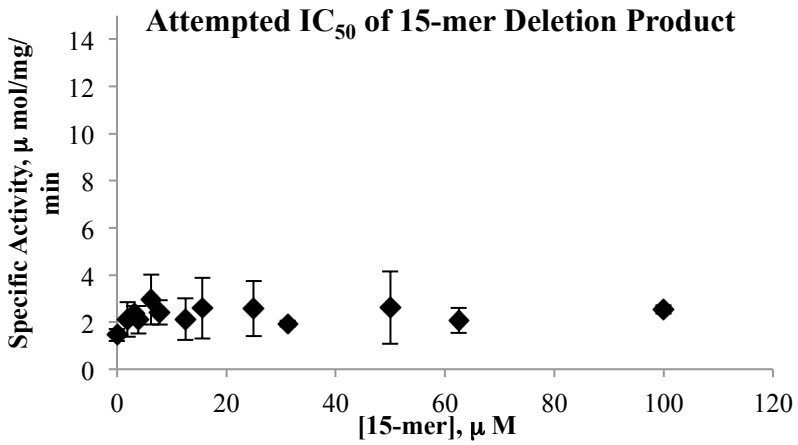
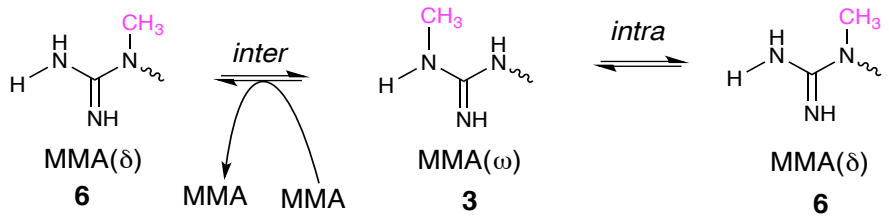


Figure 4.14 *– Specific activity varying with concentration of 15-mer deletion peptide. All data points given as mean ± standard deviation and n = 3 for each. Assay conditions: [PRMT1] = 80 nM, [H4R3] = 15 μM, [SAM] = 10 μM.*

The 15-mer did not exhibit concentration-dependent inhibition or in fact have any notable impact on PRMT1 activity. Lack of inhibition does not however preclude the 15-mer from binding to PRMT1 or even from interacting with the MMA(δ) peptide, and further work should focus on producing a cleaner MMA(δ) peptide sample and re-evaluating this methyl marker as a PRMT1 substrate.

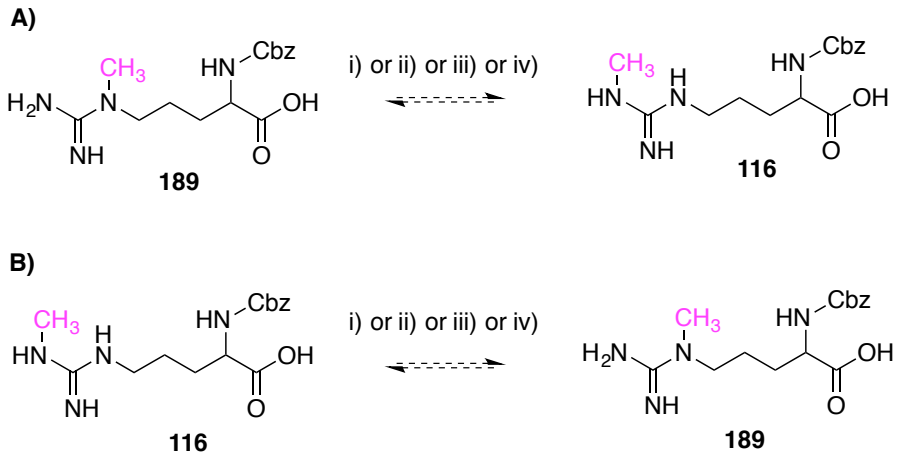
4.4.3 – Rearrangement Studies of MMA(δ) and MMA(ω)

To test whether or not the MMA(δ) and MMA(ω) methylation patterns are interchangeable (**Scheme 4.1**) samples of the amino acids were analysed by ^{1}H NMR; a precedent exists for intermolecular alkyl rearrangements *via* a 1,3-shift[182].



Scheme 4.1 *– Possible inter- or intra-molecular rearrangement between MMA(ω)* ***3*** *and MMA(δ)* ***6****.*

The α-Cbz protected derivatives, **116** and **189**, were both tested at room temperature (12 h period) and with heating (6-7 h period) in different solvents (**Scheme 4.2**). Freshly prepared samples of *N*-Me(δ)-Arg(Cbz)-OH **189** and *N*-Me(ω)-Arg(Cbz)-OH **116** were dissolved in CD_3OD and showed distinctive chemical shifts for the ω- and δ- *N*-Me peaks allowing monitoring of their possible inter-conversion (**Figure 4.15**).

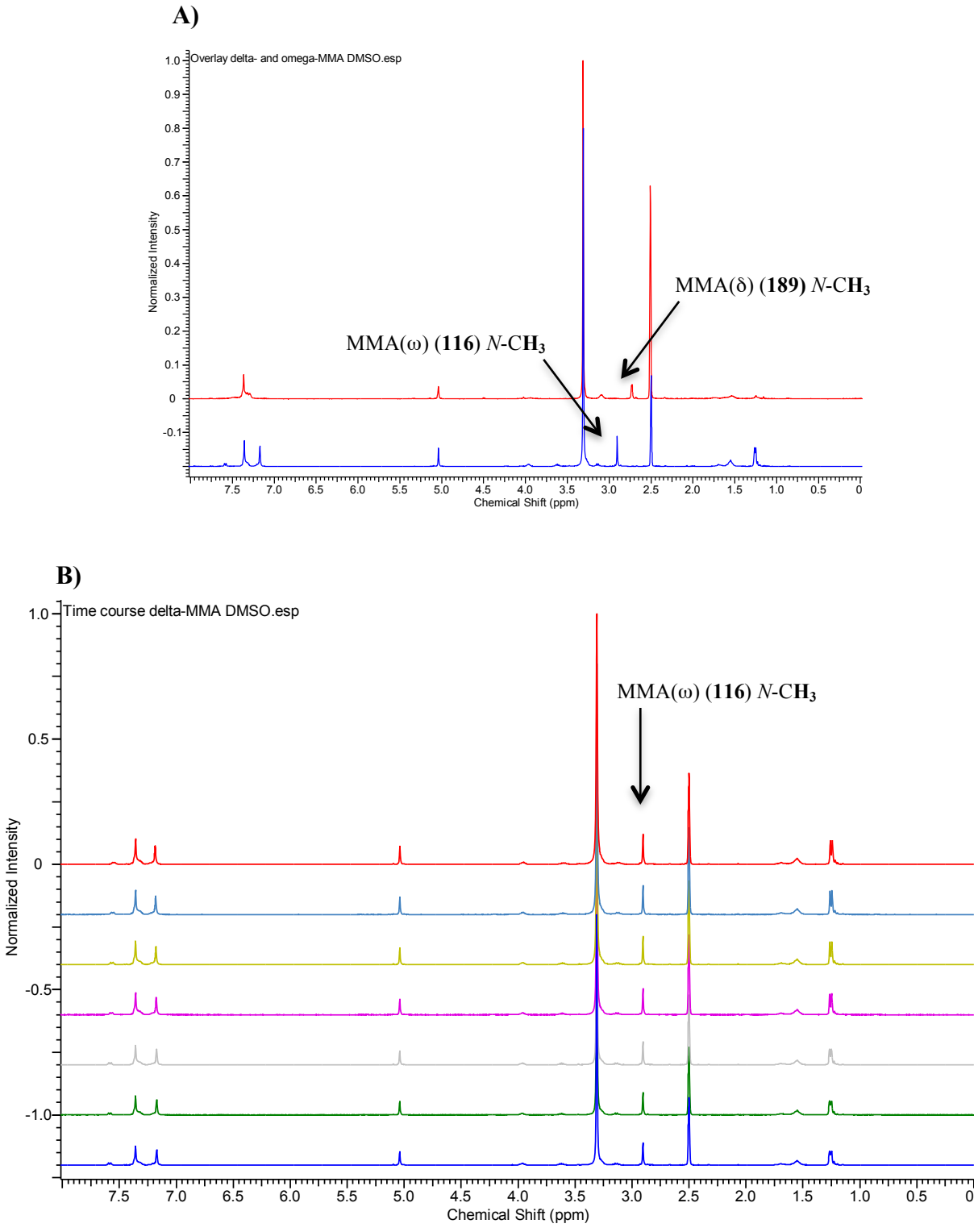


***Scheme 4.2** – Investigating possible rearrangement of **A)** MMA(δ) derivative **189** and **B)** MMA(ω) derivative **116** with each other. Reagents and conditions: i) CD_3OD, RT, 12 h ii)CD_3OD, 60 °C, 7h iii) DMSO-d_6, RT, 12h iv) DMSO-6_6, 100 °C, 6 h.*

Samples were monitored every hour for 12 hours at room temperature, showing no rearrangement. Following this they were heated to 60 °C, just below the boiling point of methanol, and analysed every hour for seven hours but again showed no rearrangement. Data for these experiments are given in **Appendix F**.

Attempts to observe rearrangement were made at a higher temperature, requiring use of a higher boiling point solvent. The samples had poor solubility in D_2O but were soluble in DMSO-d_6 and in similar fashion to their CD_3OD solutions, gave distinctive spectra but no rearrangement over a 12 hour room temperature period. Samples were heated to 100 °C and analysed every hour for

six hours, and again showed no rearrangement. Data for the DMSO-based rearrangements are given in **Figure 4.15** below.



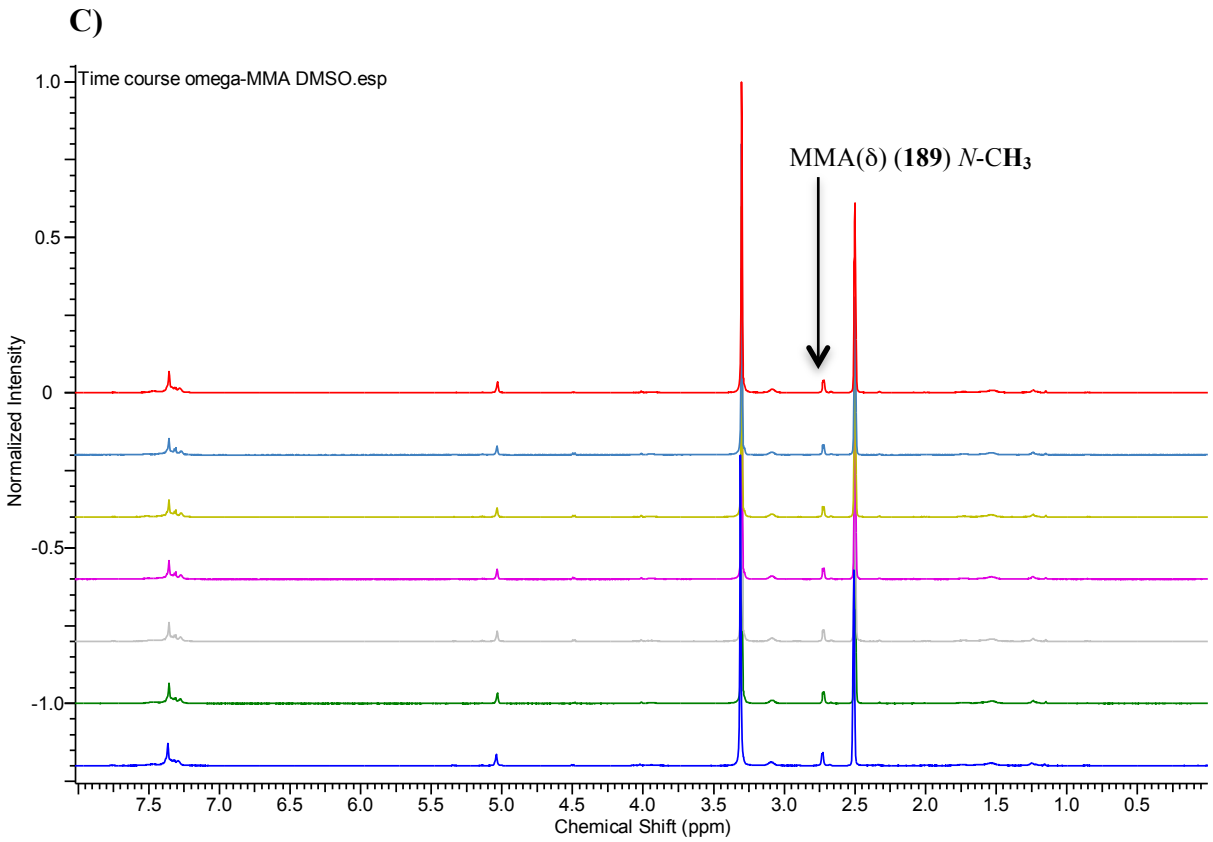


Figure 4.15 – A) *Overlapping ^{1}H NMR spectra of MMA(ω) derivative* ***116*** *(bottom, blue) and MMA(δ) derivative* ***189*** *(top, red) showing different chemical shifts of the ω- and δ-N-Me. Time course at 100 °C showing no change in the N-Me chemical shift for* ***B)*** *MMA(ω) derivative* ***116*** *and* ***C)*** *MMA(δ) derivative* ***189****.*

Unfortunately this study could not be further pursued due to limited material availability, but future work should focus on investigating the effect of acid and base as well as a broader range of temperatures.

4.5 – Conclusions and Future Work

This chapter describes the development of a MALDI-MS assay that is able to semi-quantitatively distinguish un-, mono- and di-methylated peptides. The assay is implemented here to search for evidence of PRMT1-catalysed polymethylation of known histone peptides, but found no evidence that any higher-order methylation was accrued. It is noted that these findings are not irrevocable and the idea that polymethylation patterns might exist, as hypothesised in this thesis,

is very much still a possibility. Certainly in the absence of polymethylated Arg-containing peptide standards, it is difficult to be certain that this negative result is in fact indicative of PRMT1's inability to catalyse higher-order methylation patterns.

It seems unlikely that polymethylated peptides would not crystallise with the MALDI matrix used and/or fly in the MS chamber given the known standard peptides all do, but this possibility cannot be excluded. This is especially true for tetra-methylated Arg, which could have a drastically different propensity to 'fly' in the mass spectrometer compared to its less substituted cousins, given the altered physicochemical properties (especially polarity) of the guanidine terminus when four methyl groups are present. Future work might wish to focus on synthesis of all possible methylated Args, building on the synthetic advances made in Chapters Two and Three, which might be able to offer a more definitive conclusion to this work.

Similarly for experiments testing turnover of the MMA(δ) peptide it would be inappropriate to conclude that this peptide is not a substrate of PRMT1 given the impurity of the sample tested. It is recommended that the MMA(δ) amino acid protected for SPPS (**138**) is synthesised on a greater scale to facilitate improved downstream results. The procedure outlined in Chapter Three, and the possible improvements discussed therein, can be utilised to this effect.

One of the biggest limitations to this piece of work was the low ion count experienced during some of the polymethylation investigations. The miniaturisation assay had demonstrated comparable PRMT1 activity to the large-scale optimised assay, but it is possible that some optimisation of MALDI sample preparation is necessary to circumvent the problem of low ion count. Alternatively, as part of future work it might be necessary to investigate different methods

of quantifying the MALDI readout. Below are suggested possible techniques that might offer an improved quantification method.

Isotope Dilution Mass Spectrometry (IDMS) can be used to generate quantitative data. The sample of interest is 'spiked' with a known quantity of an 'enriched isotope' internal standard[177]. Concentration of the sample of interest can be calculated by comparing the isotopes A and B (**Figure 4.16**).

R_m = Isotope ration of A to B

$$R_m = \frac{A_X C_X W_X + A_Y C_Y W_Y}{B_X C_X W_X + B_Y C_Y W_Y} \qquad C_X = \left(\frac{C_Y W_Y}{W_X}\right)\left(\frac{A_Y - R_m B_Y}{R_m B_X - A_X}\right)$$

A_X = peptide fraction in sample A_Y = peptide fraction in spike
B_X = peptide fraction in sample B_Y = peptide fraction in spike

C_X = Concentration peptide in sample W_X = Weight (sample)
C_Y = Concentration peptide in spike W_Y = Weight (spike)

Figure 4.16 *– Use of Isotope Dilution Mass Spectrometry to calculate concentration of a sample. Adapted from Fassett & Paulsen[177].*

IDMS is a powerful technique for quantification of individual peptides and would be hugely beneficial for investigating novel methylation patterns. Unfortunately, synthesis of isotope-enriched standards was outside the scope of this project but could be used to great effect as part of future work.

The alternative method of 'standard addition' can also be used if isotopically-enriched standards prove non-facile to make. This technique involves spiking the sample with a known amount of identical sample and calculation of original concentration based on the augmented signal[178]. This method is likely to be slightly more readily applicable for extension of the current work than IDMS; however, for an analysis of polymethylation standards of the novel methylated peptides would need to be synthesised.

For quantifying the non-novel methylation patterns (where standards already exist) the technique would be readily applicable. However, it is noted that this quantification step would render the assay extremely low throughput.

Following the suggested improvements for peptide testing and assessment of higher-order methylation, it is recommended that all peptide assays described above are carried out against the full panel of PRMTs. Furthermore, these novel peptides could be of high interest and value across the wider epigenetic field, particularly in the controversial field of putative Arg demethylases where they might provide some proof of concept for proposed demethylation reactions.

In conclusion a MALDI-MS assay and semi-quantification method described herein to monitor PRMT activity has been developed. This assay was utilised to great effect in the testing of compounds for PRMT inhibition - the results of this work are presented in the following chapter.

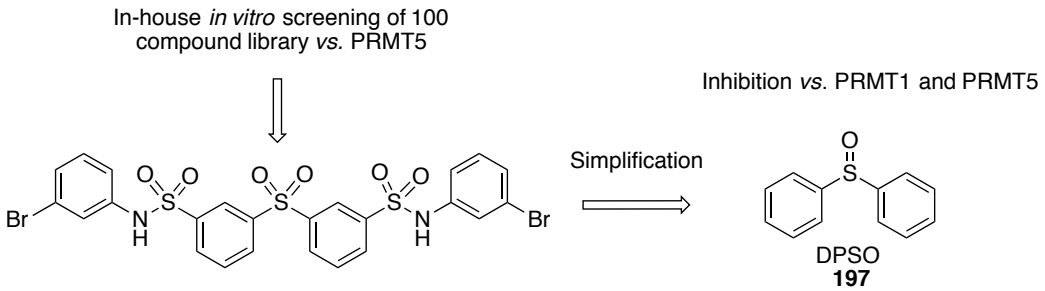
Chapter 5: Inhibitor Design, Synthesis and Testing

5.1 – Introduction

Previously in this laboratory 100 compounds were cherry-picked from a library of 20,000 for their similarity to published[99,100] PRMT inhibitors. Using radiometric- and antibody-based assays, a HTS against PRMT1 and PRMT5 were carried out respectively by Dr Alan Jones and a collaboration with Prof Steve Jane's laboratory at Monash University, Australia, that led to the identification of hit **196** as a PRMT1 and PRMT5 inhibitor with specificity against the SET7 and SET9 PKMTs (**Scheme 5.1**).

Lipinski's 'rule of five' states, amongst other guidelines, that a cLogP greater than five and a molecular weight (MW) greater than 500 generally act as predictors for poor oral bioavailability of a drug[183]. Oprea *et al.* suggest that a lead compound should have a MW no higher than 450 and a cLogP below 4.5 (but greater than -3.5) to allow for SAR development that will likely add molecular weight and lipophilicity while keeping the final drug-like compound within Lipinksi's 'rule of five'[184]. Based on these guidelines hit compound **196** had too large a cLogP value[g] (6.36) and too high a MW (686) and so represented an undesirable starting point for a drug discovery project. The structure was simplified to DPSO **197**, which had more desirable properties (cLogP = 2.09 and MW = 202) and appeared to retain inhibitory prowess *vs*. PRMT1 and PRMT5 (**Scheme 5.1**).

[g] cLogP values in this thesis were calculated using ChemDraw Professional version 15.0



Hit molecule
196

cLogP: 6.36, MW = 686

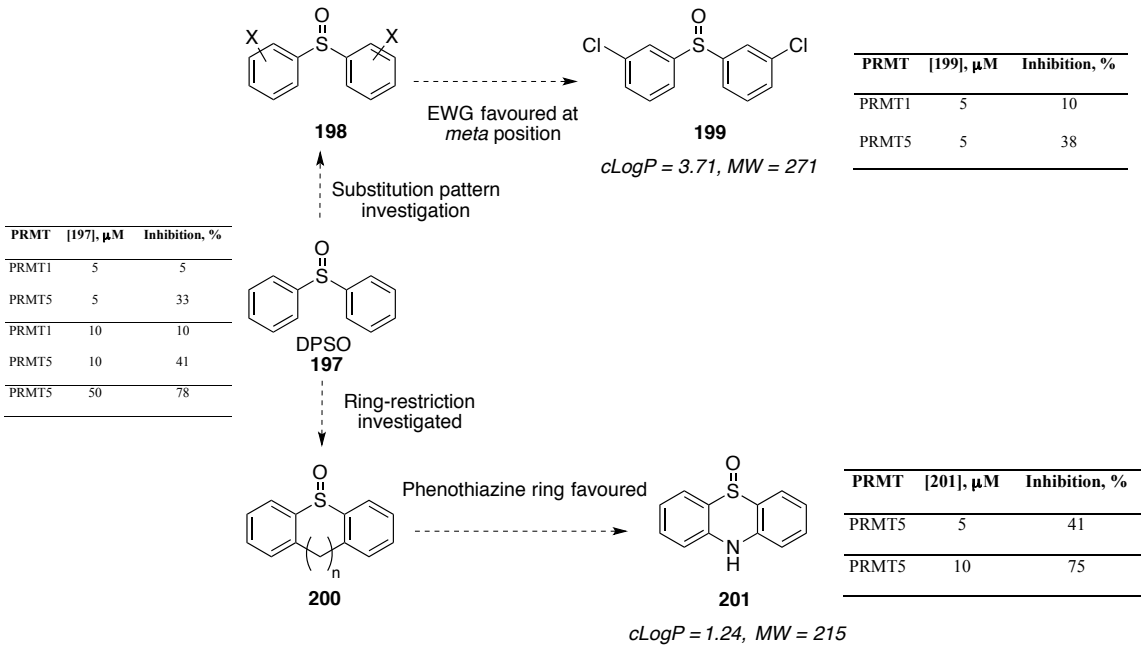
PRMT	[196], μM	Inhibition, %
PRMT1	25	65
PRMT5	25	36
PRMT1	50	90

cLogP: 2.09, MW = 202

PRMT	[197], μM	Inhibition, %
PRMT1	5	5
PRMT5	5	33
PRMT1	10	10
PRMT5	10	41
PRMT5	50	78

Scheme 5.1 *– Early progression towards identifying a potent and selective PRMT5 inhibitor. PRMT1 was tested using a commercially available radiometric assay*[185]*. PRMT5 was tested using an antibody-based assay* via *collaboration with Prof. S. Jane's laboratory at Monash University, Australia.*

Following this, SAR development led to the observation that placing electron-withdrawing groups (EWGs) *meta* to the sulfoxide **199** gave similarly potent inhibition of PRMT5 and PRMT1 (**Scheme 5.2**). It is important to note that comparisons between these compounds must be cautiously made, since no dose-response or IC_{50} data was determined. Whilst at first glance the meta-substituted **199** and the phenothiazine **201** seem to be slightly more potent than the starting scaffold **197**, this cannot be firmly concluded from these individual data points. The real attractiveness of these compounds is that the former (**199**) suggests tolerance to substitution and the latter (**201**) has an improved cLogP value (1.24).



PRMT	[199], μM	Inhibition, %
PRMT1	5	10
PRMT5	5	38

PRMT	[197], μM	Inhibition, %
PRMT1	5	5
PRMT5	5	33
PRMT1	10	10
PRMT5	10	41
PRMT5	50	78

PRMT	[201], μM	Inhibition, %
PRMT5	5	41
PRMT5	10	75

Scheme 5.2 *– Early progression towards identifying a potent and selective PRMT inhibitor. PRMT1 was tested using a commercially available radiometric assay*[185]*. PRMT5 was tested using an antibody-based assay* via *collaboration with Prof. S. Jane's laboratory at Monash University, Australia.*

This chapter describes two approaches that attempt to facilitate an improved PRMT inhibitor:

1. Improve upon the inhibitory properties of the aforementioned scaffolds
2. Make an assessment of and improve upon the techniques used for assaying PRMTs

Section 1.4 presented a critical discussion of commonly used PRMT assay techniques and their shortcomings were highlighted. It is hypothesised that readouts from different assays can have a marked effect on inhibition data and give rise to artefacts. The following specific aims were envisaged towards identification of a more potent PRMT inhibitor building on early SAR data described above and utilising the MS-based assay described in the previous chapter:

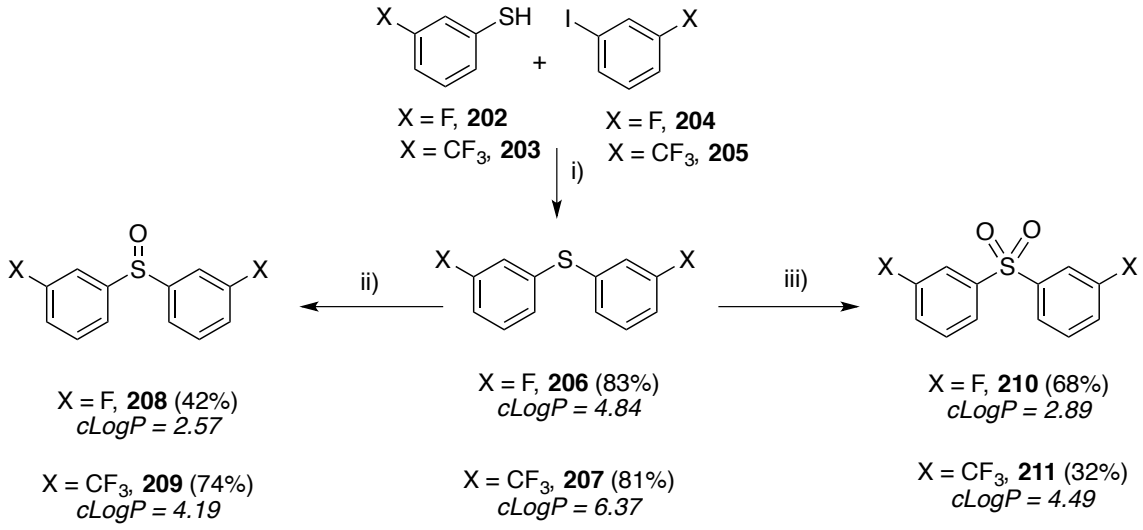
5.2 – Chapter Aims

1. Expand library of *bis*-aryl and tricylic analogues based on hit molecule DPSO **197**
2. Test compounds for PRMT inhibitory activity using three different assay approaches

5.3 – Inhibitor Synthesis

5.3.1 – Bis-aryl Series

Based on scaffold **197**, a small array of *bis*-aryls were synthesised with *meta* EWGs. Using CuI as a catalyst, the sulfides were obtained *via* an Ullman-type coupling[186] from the corresponding thiophenols **202**, **203** and aryl-iodides **204**, **205** in 81-83% yield. Additionally, in a systematic investigation of sulfur oxidation state on PRMT activity and selectivity, the sulfides **206** and **207** were oxidised *via* a Gif-type[187] reaction using $FeCl_3$ and H_5IO_6 to the sulfoxides **208** and **209** in 42-74% yield. The sulfones **210** and **211** were synthesised *via* a Prizelhaev-type[188] reaction using H_2O_2 and AcOH, giving 32-68% yield (**Scheme 5.3**).

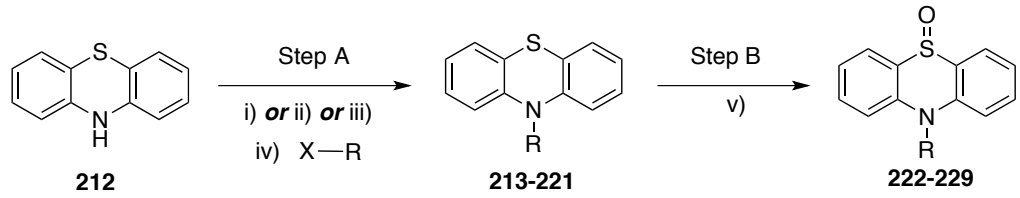


***Scheme 5.3** – Synthesis of bis-aryl array. Reagents and conditions: i) CuI, K_2CO_3, iPrOH, ethylene glycol, 80 °C, 24 h ii) $FeCl_3$, H_5IO_6, RT, 16 h iii) $H_2O_{2(aq)}$ (33%), AcOH, 2 h, 125 °C.*

5.3.2 – *N*-Substituted Tricyclic Series

Based on scaffold **201**, an array of *N*-substituted tricyclic phenothiazines (**213-221**) were synthesised using alkyl halides and base (NaH or Cs_2CO_3 or KHMDS) giving yields of 10-68%.

These sulfides were oxidised to sulfoxides **222-229** using *m*CPBA in 37-69% yield (**Scheme 5.4**).

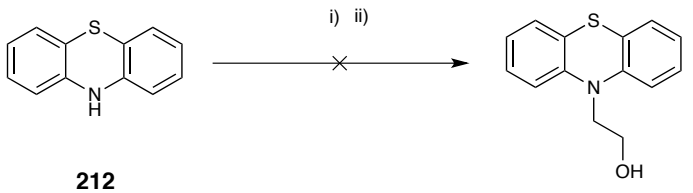


Product, Step A *(cLogP)*	R	Conditions (Step A)	Yield, % (Step A)	X	Product, Step B *(cLogP)*	Yield, % (Step B)
213 *(4.13)*		i)	68	I	**222** *(1.48)*	37
214 *(4.66)*		i)	65	Br	**223** *(2.01)*	65
215 *(5.19)*		iii)	53	Br	**224** *(2.54)*	54
216 *(5.72)*		ii)	10	Br	**225** *(3.07)*	69
217 *(5.11)*		i)	20	Br	**226** *(2.54)*	65
218 *(5.66)*		iii)	44	Br	*n/a*	*n/a*
219 *(6.22)*		iii)	17	I	**227** *(3.57)*	38
220 *(6.78)*		iii)	52	Br	**228** *(4.13)*	66
221 *(6.40)*		i)	37	Br	**229** *(4.45)*	62

Scheme 5.4 – *Synthesis of phenothiazine analogues. Reagents and conditions: i) NaH, 30 min, RT, THF ii) TBAI, Cs_2CO_3, 30 min, RT, DMF iii) KHMDS, 0 °C, 15 min, THF iv) requisite alkyl halide, 16 h, RT v)* m*CPBA, CH_2Cl_2, RT, 16 h.*

Synthesis of some of the sulfone derivatives was attempted *via* a Prizelhaev-type[188] reaction using H_2O_2 and AcOH. However, crude ^{1}H NMR data suggested the sulfones had poor solubility in DMSO, meaning they were not suitable candidates for inhibitor testing and were therefore not purified or further pursued.

The sulfides had high cLogP values (4.13-6.78) while the sulfoxides were less hydrophobic (cLogP = 1.48-4.45) and could make for better lead-like molecules[184]. In an attempt to improve the cLogP properties of the sulfides in accordance with Oprea *et al.*'s[184] guidelines for lead-like molecules and to improve solubility, it was decided to introduce more polar substituents at the ring nitrogen starting *via* S_N2 reaction with halogenated alcohols (**Scheme 5.5**).



212 **230**

cLogP = 3.14

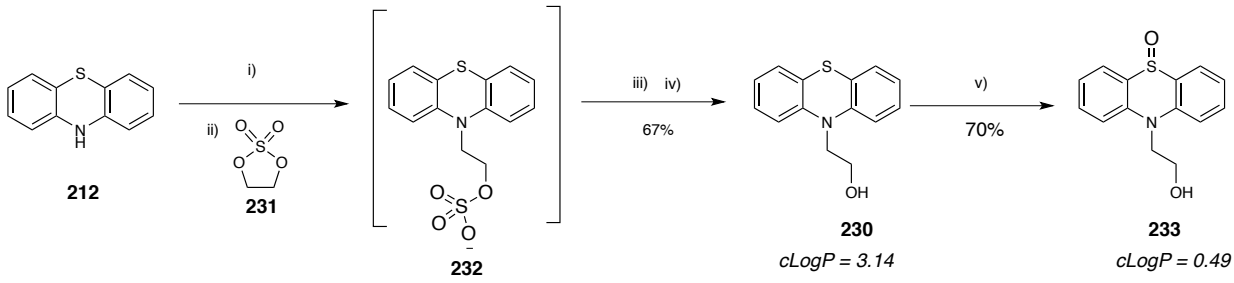
Entry	Reagents	Base	Time, h	Temperature, °C	SM 212, %	PM 230, %
1	Cl-EtOH	NaH	20	0 → 25	100	0
2	Cl-EtOH	K_2CO_3	43	60 → 95→ 110	100	0
3	Br-EtOH	NaH	20	0 → 25	100	0
4	Br-EtOH + TBAI	Cs_2CO_3	20	0 → 25	100	0

Scheme 5.5 – *Attempted **s**ynthesis of **230**, reagents and conditions: as stated in table, i) requisite halogenated alcohol ii) Base.*

The table above shows attempts to afford **230** *via* this route using chlorinated and brominated alcohol but this was not forthcoming. Entry four shows an attempt to generate the more electrophilic iodo- species *in situ* using the transfer agent tetrabutylammonium iodide (TBAI),

but to no avail. It is postulated that the halogenated alcohol readily formed the more stable alcoholate then formed the epoxide *via* intramolecular attack, which the phenothiazine -NH was subsequently unable to react with.

Instead, according to Kulig *et al.*[189], 1,3,2-dioxathiolane 2,2-dioxide **231** was used to afford **230** in 67% yield (**Scheme 5.6**). Subsequent oxidation using *m*CPBA afforded the sulfoxide **233** in 70% yield; oxidation was confirmed to have taken place at sulfur rather than at the amine or alcohol groups by X-ray crystallography (**Figure 5.1**).



***Scheme 5.6** – Successful synthesis of **230** and **233**. Reagents and conditions: i) THF, KHMDS, 0 °C, 1 h ii) 16 h iii) conc. H_2SO_4 iv) H_2O v)* m*CPBA, 2 h, RT.*

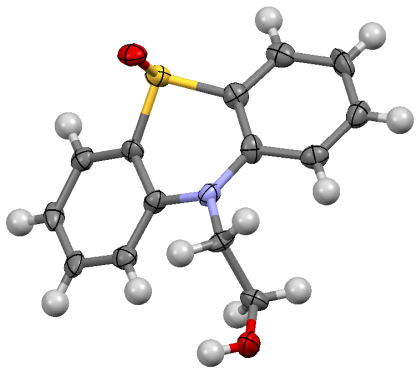
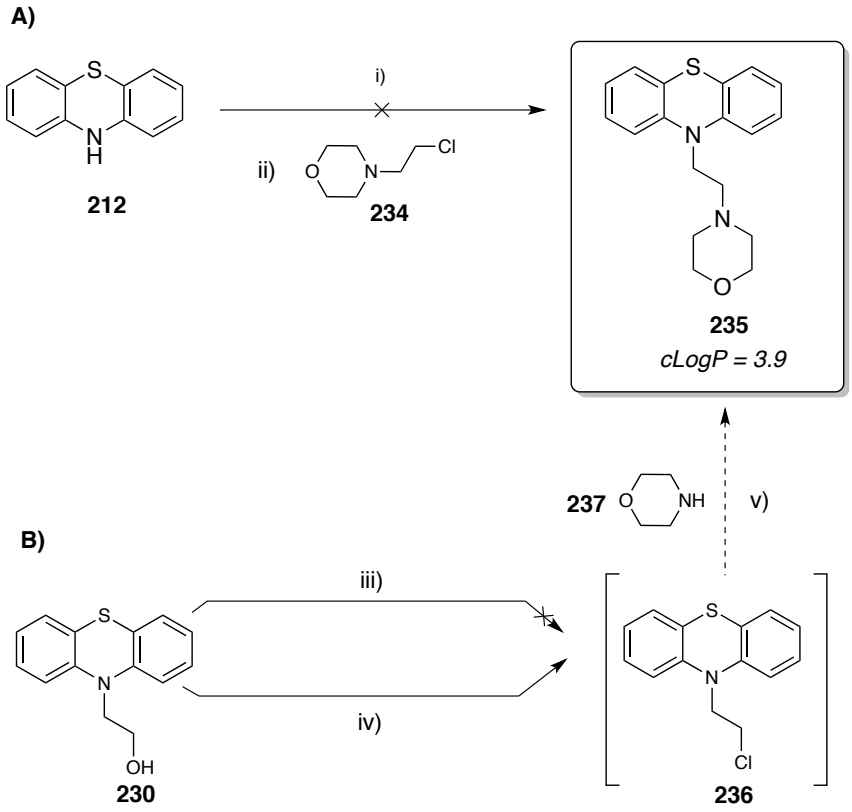


***Figure 5.1** – Crystal structure of **233** demonstrating oxidation has occurred at sulfur, not at the –OH. Data collection and processing carried out by Dr R. Cooper. See **Appendix G** for crystallographic information.*

Further endeavours were made to alleviate hydrophobicity in this series by synthesising an ethylmorpholine derivative of phenothiazine (**Scheme 5.7A**).



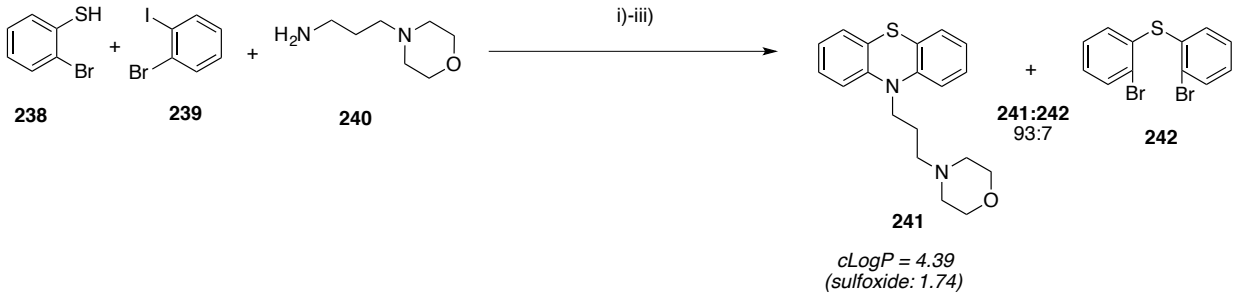
***Scheme 5.7** – Attempted **s**yntheses of morpholine-substituted phenothiazine. Reagents and conditions:**A)** i) KHMDS, 0 °C, 15 min ii) RT, 16 h **B)** iii) Me_3SiCl, DMSO, RT, 15 min, adapted from Snyder[190] iv) PPh_3, CCl_4, RT, 2 h - Appel reaction[191] v) DIPEA, CH_2Cl_2 , RT, 16 h.*

At first, a substitution reaction with chloroethylmorpholine **234** was attempted, returning a complex mixture containing both starting material **212** and an unidentified degradation product after 16 h as assessed by ^{1}H NMR.

A different approach was attempted *via* chlorination of newly synthesised alcohol-substituted phenothiazine **230** in an attempt to facilitate a substitution reaction with morpholine **237** itself (**Scheme 5.7B**). Chlorination of **230** towards **236** using Me_3SiCl and catalysed by DMSO was not successful. Column chromatography on SiO_2 returned starting material as determined by MS, alongside an unidentified degradation product.

Instead, an Appel reaction[191] was attempted using PPh_3 and CCl_4, which showed disappearance of the diagnostic –O*H* triplet in the ^{1}H NMR spectrum of the crude reaction mixture. In order to avoid the hypothesised degradation observed previously, the crude reaction mixture was treated with morpholine without purification but a complex mixture was afforded and no conclusions could be drawn.

According to Dahl *et al*.[192] Pd_2dba_3, dppf and NaOtBu could be used to couple three components *via* click-chemistry. This methodology was employed with heating at 60 °C for synthesis of propylmorpholino derivative **241** (**Scheme 5.8**).

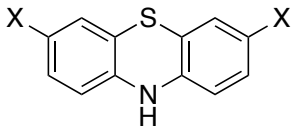


Scheme 5.8 *– Click chemistry synthesis of morpholine-substituted phenothiazine. Reagents and conditions: i) Pd_2dba_3, dppf, NaOtBu ii) 60 °C, MW, 20 min iii) 160 °C, MW, 2 h. Purity assessed by NMR.*

Purification of the reaction mixture was attempted using SiO_2 chromatography but this did not resolve all the products. Ion-exchange chromatography was used to remove unconsumed starting amine **240** followed by a second attempt at SiO_2 chromatography, which was unable to separate the two products **241** and **242**. Preparative-scale TLC was also attempted, but this left only an analytical quantity of desired product **241** containing a 7% contamination of the non ring-closed product **242**. This impurity is a recognised side-product of analogous reactions from the original report and unfortunately precluded **241** from being tested as a PRMT inhibitor. Had **241** been isolated in 100% purity, attempts to synthesise the sulfoxide derivative would have been made, since this compound has a favourable cLogP of 1.74.

5.3.3 – Ring-Substituted Tricycles

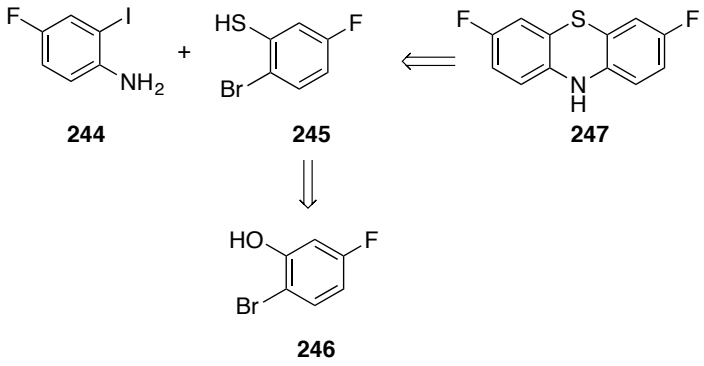
Given the SAR data presented in the introduction to this chapter, the next logical step was an attempt to combine the positive structural elements reported for the *bis*-aryl compounds (namely *meta*-substituted EWGs) with those reported for the ring-restricted compounds (phenothiazine-like core). Attempts were therefore made to synthesise phenothiazines *meta*-substituted with electron-withdrawing halogens such as **243** (**Figure 5.2**) and then explore oxidation state at sulfur.



243

Figure 5.2 *– Generic structure of symmetrical ring-substituted phenothiazine.*

Direct functionalization of the ring was quickly ruled out as a viable tactic due to lack of regio-control in electrophilic substitutions of phenothiazines[193-195]. It was therefore decided to build up substitution patterns from component arenes. **Scheme 5.9** outlines a possible retrosynthesis towards the di-fluorinated derivative:

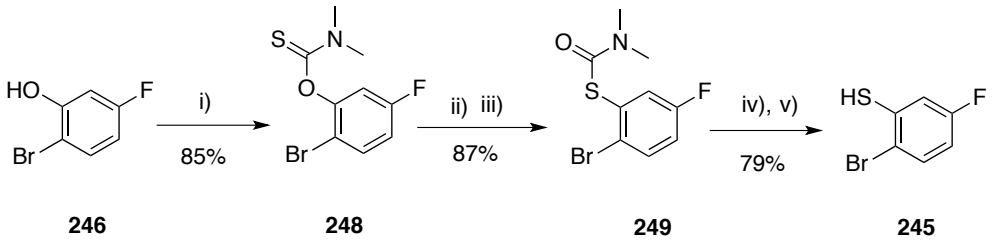


244 **245** **247**

246

Scheme 5.9 *– Retrosynthetic scheme for synthesis of* meta-*substituted fluorinated phenothiazines.*

The fluorinated thiol **245** is not commercially available and was synthesised *via* the Newmann-Kwart rearrangement in good overall yield (58%, 3 steps) (**Scheme 5.10**). The *O*-carbamothioate **248** was afforded in 85% yield by reaction of phenol **246** with dimethylthiocarbamoyl chloride at room temperature. The rearrangement to *S*-carbamothioate **249** (87% yield) was achieved by heating under harsh conditions (220 °C with MW irradiation for 2 h) followed by treatment with 4 M $HCl_{(aq)}$. Subsequent cleavage of the dimethylthiocarbamoyl group was achieved by heating in $NaOH_{(aq)}$ at 100 °C to yield thiol **245** in 79% yield.

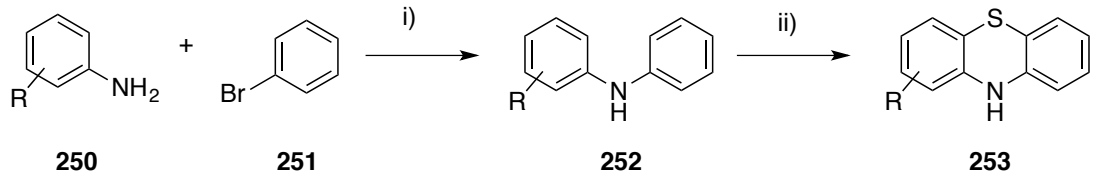


Scheme 5.10 *– Newmann-Kwart rearrangement to form fluorothiophenol **245**. Reagents and conditions:i) Dimethylthiocarbamoyl chloride, K_2CO_3, DMF, 25 ° C, 17 h ii) 220 °C, 2 h MW,* N,N-*diethylaniline iii) 4 M $HCl_{(aq.)}$ iv) MeOH, $NaOH_{(aq)}$ 1 M, 100 °C, 2 h v) $HCl_{(aq)}$ 1 M to pH 1.*

Synthesis of *N*-propylmorpholino phenothiazine **241**, described in **Scheme 5.8** above, found formation of the ring-closing C-N-C bond to be problematic. It was anticipated that a similar problem might also arise for synthesis of ring-substituted phenothiazines. Procedures that proceed *via* synthesis of the C-N-C bond as the first step were investigated with the aim of overcoming this problem.

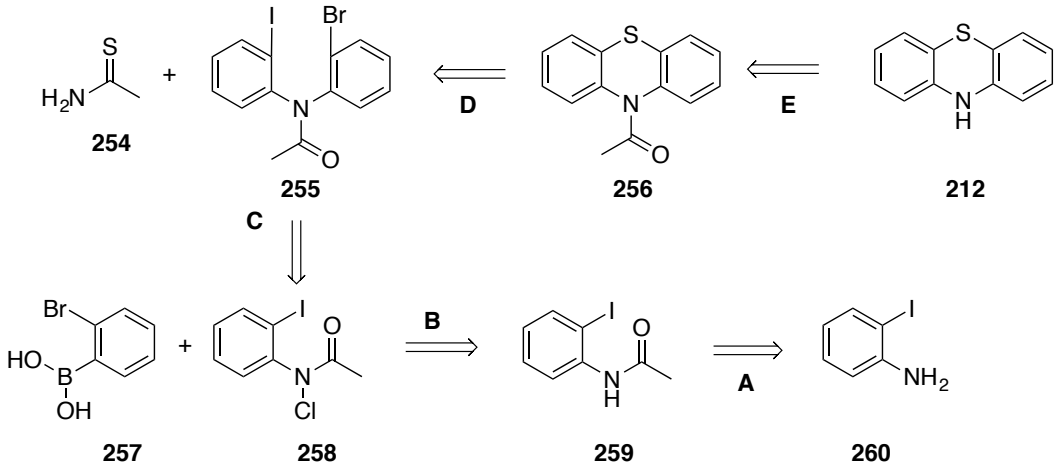
A precedent does exist for initial C-N-C bond formation but failure to ring close *via* sulfurization was observed[196], and other reports state that conversion to intermediate **252** is not forthcoming[197]. **Scheme 5.11** describes one of the more successful procedures but even here yields of only 10-50% are obtained, with starting materials being recovered[196]. The authors used Pd_2dba_3 and X-Phos to catalyse coupling of starting anilines with bromobenzenes to form the C-

N-C bond, followed by I_2-catalysed sulfurization using elemental sulfur and refluxing at 190 °C to effect ring closure.



***Scheme 5.11** – Bis-aryl coupling followed by sulfurization as described by Madrid* et al.*[196]. Reagents and conditions: i) Pd_2dba_3, X-Phos, K_3PO_4, reflux 20 h ii) Sulfur, I_2, MW 190 °C, 20 min, H_2O.*

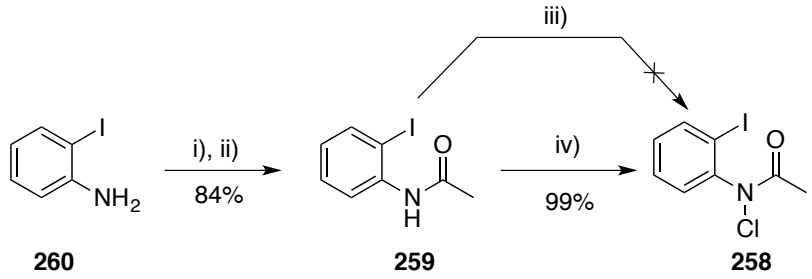
Given the generally poor success reported for such reactions, a different route was sought. One approach that held some promise was the reportedly mild Cu(I)Cl-catalysed coupling of arylboronic acids with *N*-chloroamides[198], adapted in **Scheme 5.12C** for the model synthetic route towards phenothiazine **212**. According to Tao *et al.*[199], thioacetamide can be used to access *bis*-aryl thioethers **256** from **255** (**Scheme 5.12D**).



***Scheme 5.12** – Retrosynthesis of phenothiazine **212** as a model for substituted phenothiazines.*

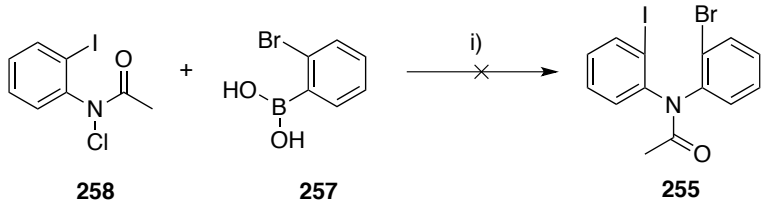
It was anticipated this procedure could be used for synthesis of ring-substituted phenothiazines but first, to conserve more precious and limited substituted starting materials, the retrosynthetic route was tested for its ability to produce phenothiazine itself **212**. **Scheme 5.13** shows the

successful generation of *N*-chloroamide **258** by following a procedure adapted from He *et al.*[200] using calcium hypochlorite in a H_2O:Et_2O biphasic reaction system after a failed attempt to use milder conditions utilising *N*-chlorosuccinimide as the chlorinating agent[201].



Scheme 5.13 *– Reagents and conditions: i) Ac_2O, Et_3N, CH_2Cl_2, 0 °C ii) RT, 2 h iii) N-chlorosuccinimide, t-BuOK, THF iv) $Ca(ClO)_2$, $NaHCO_3$, Et_2O, H_2O.*

No success was encountered however for the coupling of *N*-chloroamide **258** with commercially available boronic acid **257** using CuCl as the catalyst at room temperature (**Scheme 5.14**).

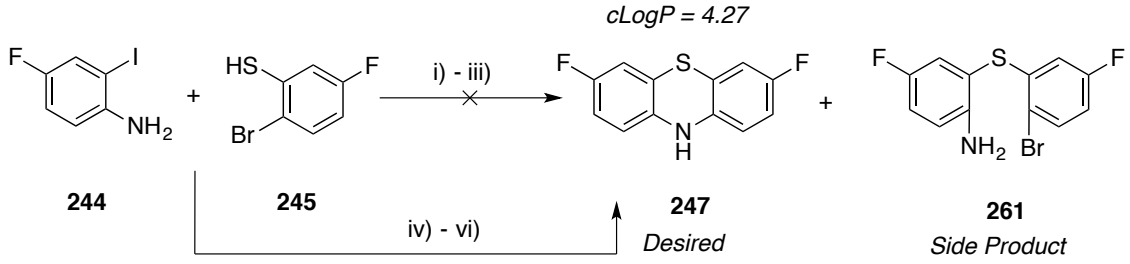


Scheme 5.14 *– Attempted **s**ynthesis of **255** towards phenothiazine **212**. Reagents and conditions: i) Cu(I)Cl, Na_2CO_3, THF, RT, 36 h.*

Reaction conversion was poor – starting material was identified by TLC analysis but the crude mixture was complex and not amenable to ^{1}H NMR analysis. Two reasons are thought to explain the observed poor conversion. Firstly, CuCl is prone to spontaneous oxidation to $CuCl_2$. Following an initial failed attempt at this reaction, fresh CuCl was used and anoxic conditions maintained using Schlenk line techniques but the reaction still afforded a complex mixture with no detectable desired product. The second reason for poor reaction outcome is proposed to be the instability of *N*-chloroamide **258**; ^{1}H NMR analysis revealed degradation back to the starting phenylacetamide **259** after 1 d at room temperature. Although attempts were made using freshly

prepared reagent, the reaction still offered disappointing results and identification and isolation of **255** from the complex reaction mixture remained prohibitively challenging.

At this stage it seemed that use of a 'click' coupling procedure was unavoidable, despite the anticipated difficulty of forming the C-N-C bond and closing the tricylic ring system. According to Ma *et al.*[202], CuI and L-Proline-catalysed coupling could be used to produce substituted phenothiazines from the corresponding substituted 2-iodoanilines and 2-bromothiophenols. This methodology was implemented using 4-fluoro-2-iodoaniline **244** and 4-fluoro-2-bromothiophenol **245** towards di-fluorinated phenothiazine **247** (**Scheme 5.15,** i-iii).



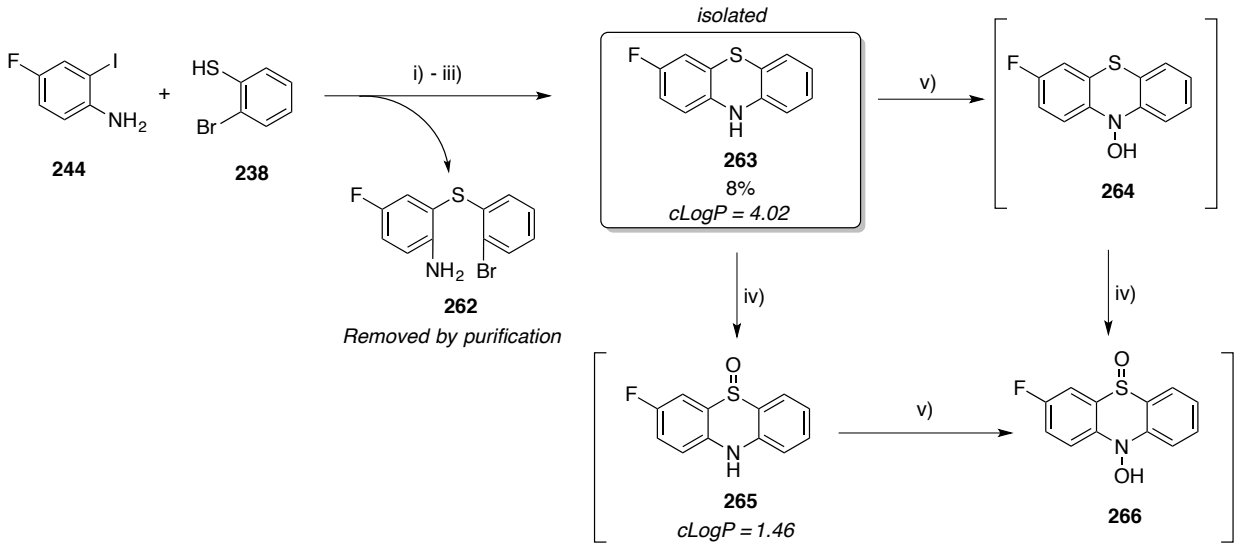
Scheme 5.15 *– Coupling to form di-fluorinated phenothiazine **247**. Reagents and conditions: i)CuI, L-proline , K_2CO_3, 2-methoxyethanol, 90 °C, 48 h ii) 110 °C 72 h iii) 130 °C 72 h. iv) Pd_2dba_3, dppf, NaOtBu v) 60 °C, MW, 20 min vi) 160 °C, MW, 2 h.*

Ma *et al.*[202] report that presence of EWG *meta* to the sulfide can slow the aromatic substitution reaction that brings about ring closure and these observations were corroborated in this work. ^{1}H NMR analyses of the crude product showed a complex mixture was present and it was predicted that desired product **247** and side-product **261** were present after 5 days of reaction. ^{19}F NMR analyses of the crude reaction mixture confirmed the presence of four different fluorinated compounds. The reaction was pushed, as suggested in the original paper, for a further 3 days at 130 °C but did not reach full conversion as assessed by ^{19}F NMR of the crude reaction mixture. Similar purification difficulties were encountered here as for the *N*-propylmorpholino derivative

241 with co-elution of the ring-closed **247** and ring-open **261** products occurring using SiO_2 chromatography, so different coupling conditions were sought.

According to Dahl *et al.*[192] 2-iodobromobenzene, 2-bromothiophenol and a substituted primary amine can be coupled using Pd_2dba_3, dppf and NaOtBu and heating with MW irradiation. This method was found to have some success producing *N*-propylmorpholino-phenothiazine **241** and was adapted for a two-component coupling using 4-fluoro-2-iodoaniline **244** and 4-fluoro-2-bromothiophenol **245** rather than the reported three-component click reaction (**Scheme 5.15** iv-vi).

In this case a complex mixture of products was produced that could not be distinguished by ^1H NMR. ^{19}F NMR of the crude mixture suggested at least three distinct fluorinated species were present, thought to be the two products and starting aniline (supported by TLC analysis). Silica chromatography could not separate the aniline from the product but ion-exchange chromatography of the crude mixture was able to remove the aniline from the mix leaving behind two fluorinated entities as confirmed by ^{19}F NMR. Attempts to separate these products by silica chromatography were partially successful; separation was possible but almost immediately upon elution the desired product appeared to be oxidising to an unknown entity (+ 16 in the mass spectrum) and the ^1H NMR spectrum showed a highly complex mix of aromatic protons. It was considered that the strongly electron-withdrawing properties of the two fluorine groups might have some effect on the stability of **247**. It was decided to attempt synthesis of mono-fluorophenothiazine in the hope that this might be more stable (**Scheme 5.16**). 4-fluoro-2-iodoaniline and 2-bromothiophenol were heated with Pd_2dba_3, dppf and NaOtBu at 60 °C with MW irradiation for 20 min, followed by 2 h at 160 °C.



Scheme 5.16 – *Reagents and conditions: i) Pd_2dba_3, dppf, NaO*t*Bu ii) 60 °C, MW, 20 min iii) 160 °C, MW, 2 h iv)* m*CPBA, CH_2Cl_2 v) Spontaneous oxidation at RT of fluorinated phenothiazine over an indeterminate period of up to five days.*

Desired product **263** was isolated in poor yield (8%) following the same dual purification procedure described above – removal of aniline *via* ion-exchange and separation of **263** and **262** by silica chromatography. However, spontaneous oxidation of **263** was also observed as had been noted for the di-fluoro analogue. To probe whether oxidation was occurring on sulfur, attempts were made to push the transformation to completion using *m*CPBA to form **265**. However, it quickly became apparent that the oxidation was happening at either the nitrogen or on the arene ring, since a mixture of four fluorinated products could be observed by ^{19}F NMR following oxidation with *m*CPBA. LC-MS/MS identified **264** and **266** *via* fragmentation analyses. Atmospheric pressure chemical ionisation (APCI) can be used in tandem with electrospray ionisation (ESI) to make chemical structure inferences. Notably, presence of a –O ion peak (-16 mass) in the APCI MS/MS spectrum is diagnostic for an *N*-oxide[203]. Samples of mono-fluorophenothiazine **263** and its *S*-oxide **265** were tested: **263** contained a mixture of the desired product and the *N*-oxide, as determined by APCI. The sample of intended **265** contained almost entirely **266**.

With these new findings in mind, the ring-substituted phenothiazines were dropped from further investigation as part of this project; unquestionably it would not be productive to submit mixtures of unknown purity, stability and composition to the biological assay. All other bis-aryl and tricyclic structures synthesised herein, barring *N*-propylmorpholino-phenothiazine that contained a non-removable impurity, were taken forward for testing against PRMTs for inhibitory potency.

5.4 – Testing

Three different assays were used to test the aforementioned compounds for inhibitory potency: a chemiluminescence assay, a radiometric assay and the MS-based assay described in Chapter 4. An overview of these assay types was given in **Section 1.4**. Results from each assay are presented on a case-by-case basis in this section and a summary of results is tabulated in **Figure 5.11** on pages 178-180.

Of particular interest was the ability of these new compounds to inhibit PRMT5. These compounds were synthesised as part of the current project building on previous work within our laboratory (**Scheme 5.1** and **Scheme 5.2**) that had identified the scaffold DPSO **197** and subsequent *meta*-substituted **199** and ring-restricted **201** derivatives as PRMT inhibitors.

5.4.1 – Chemiluminescence Assay

This commercially available assay purchased from Amsbio uses a primary antibody purported to recognise methylated Arg and a secondary antibody conjugated to HRP to generate chemiluminescence using H_2O_2 as a substrate (**Scheme 1.9B**). This protocol was broadly outlined in **Section 1.4.1.b** and is explained in detail in the Experimental Section. Immobilised

histone substrate was incubated with PRMT1/5 (supplied with the assay kit) and inhibitor for 1-1.5 h; following a washing step to remove all assay components from the immobilised histones, the samples were blocked to prevent non-specific binding. Primary antibodies (anti Me-Arg) were used to detect methylated Arg. The exact methylation pattern recognised by the antibodies was proprietary information – it is unknown whether they bind to mono-or dimethyl- Arg or both.

Unbound antibody was washed off and the samples blocked again. A secondary antibody conjugated to HRP was then applied. After a final washing step, the plate was treated with H_2O_2, a substrate for HRP, and the resultant chemiluminescence monitored using an automated plate-reader device.

With limited access to these assay components due to their prohibitive expense, only a small sample of *bis*-aryl and tricyclic compounds were tested at 50 μM against PRMT5 (**Figure 5.3**). Those chosen for testing were some of the *meta*-substituted electron-withdrawing *bis*-aryl motifs and a selection of tri-cyclic phenothiazine-like cores. A limited amount of PRMT1 was available for these assays and some of the selected compounds were also tested against this enzyme in an attempt to monitor any selectivity between PRMT1 and PRMT5 (**Figure 5.3**).

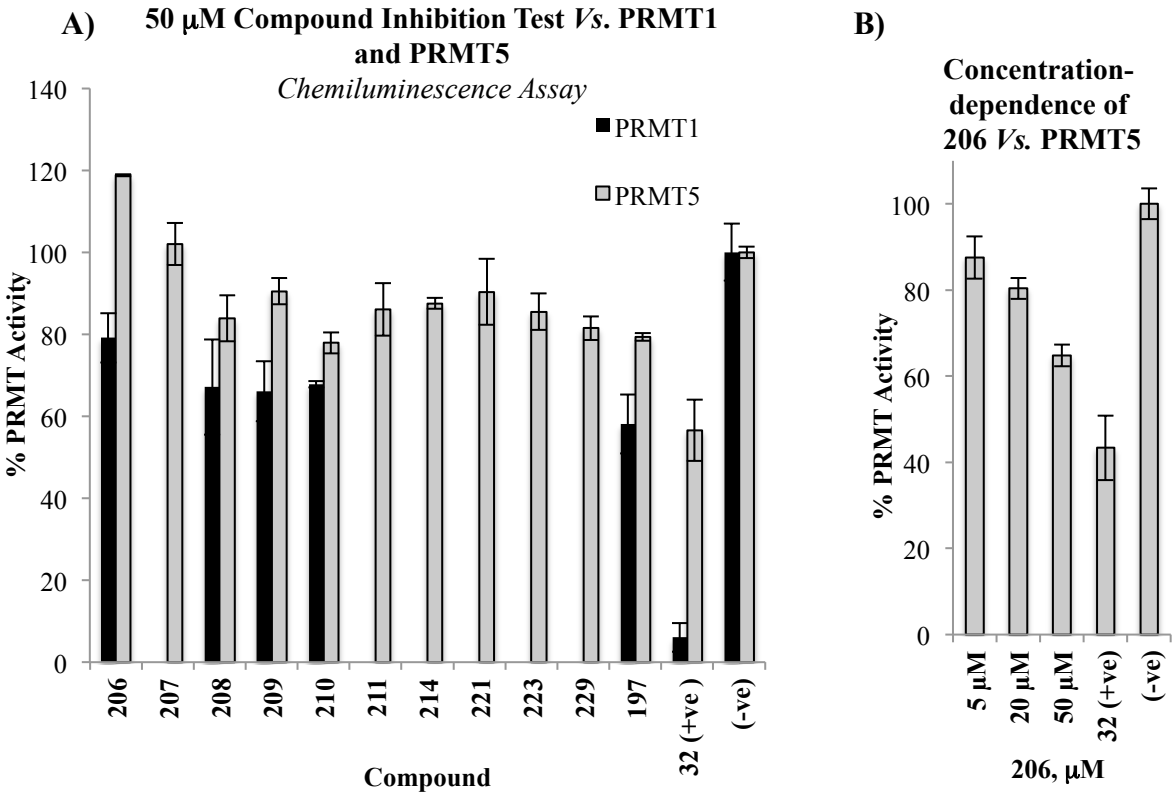


Figure 5.3 – *Inhibitory tests of a select group of compounds using the commercially available chemiluminescence assay kit purchased from AmsBio.* ***A)*** *Compounds tested at 50 μM* vs. *PRMT1 and PRMT5 and* ***B)*** *compounds tested at 5, 10 and 50 μM* vs. *PRMT5 in an attempt to observe any dose-dependent response. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin* ***32*** *and/or SAH* ***7****. Data is presented as the mean ± standard deviation and n = 3 for each data point.*

Positive and negative controls were carried out in the presence of Sinefungin and absence of inhibitor respectively. The positive control using Sinefungin **32** at 50 μM *vs.* PRMT5 suggested the dataset should be treated with caution. Sinefungin **32**, a universal and potent inhibitor of all SAM-utilising enzymes, only gives 44% inhibition of PRMT5 as calculated by this assay; it is expected that the observed inhibition should be much greater at 50 μM given Sinefungin has a reported IC_{50} of 8.6 μM *vs*. PRMT5:MEP50[59]. Possibly the exact batch of enzyme used was extraordinarily active, or perhaps a degree of human and/or instrumental error was involved, since these measurements were made only once (although in triplicate). However, the notion that antibody-based detection of methylated Arg is often unreliable was raised and discussed in **Section 1.4.1.b**. The results reported herein go some way to supporting this theory, although

further repeats would be necessary. It is advised though that data from this assay is treated with extreme caution.

The results obtained were somewhat disappointing with little inhibition being observed against both enzymes. DPSO **197** was found to inhibit PRMT5 by 21% at 50 μM compared to the previously observed value of 78%. Previously DPSO **197** was tested at 5 and 10 μM *vs.* PRMT1 (5% and 10% inhibition respectively). The analogous 50 μM test was not carried out in the previous study and enzyme availability limited tests herein such that the lower concentrations could not be tested. **Table 5.1** documents these data comparisons for DPSO **197**.

PRMT	**[197], μM**	**Inhibition, %**	
		Previous study	*Current study*
PRMT1	5	5	-
PRMT5	5	33	-
PRMT1	10	10	-
PRMT5	10	41	-
PRMT1	50	-	42
PRMT5	50	78	21

Table 5.1 – *Comparison of previously obtained inhibition data and inhibition data from the current study.*

Table 5.1 shows an inconsistency between the results observed in the present study and in the previous study for the single concentration test of DPSO (50 μM) *vs.* PRMT5. Parham *et al.*[204] discuss the importance of using concentration-dependent data to allow inter-compound comparisons of inhibitory prowess; it is emphasised that the data presented here is at the preliminary single concentration point stage. It is more probable that the inconsistencies observed are a result of some form of inherent error, supported by the observed poor inhibition for the positive control of 50 μM Sinefungin **32** *vs.* PRMT5.

Further inconsistencies using this assay were noted. In a separate run of the assay, concentration dependency was tested for **206** against PRMT5 (**Figure 5.3B**). At first glance, a mild concentration-dependency appears to exist for **206**. However, comparing the percentage inhibitions (for **206** at 50 μM) obtained for the two assays shown in **Figure 5.3** shows profound inconsistency: **Figure 5.3A** shows 119% PRMT5 activity whereas **Figure 5.3B** shows 65% PRMT5 activity. This suggests an assay-based inconsistency is affecting results.

Not wanting to abandon the series of compounds in case it transpired any of the results above were true positives, it was deemed prudent to re-test in an alternative assay. Described below is use of a radiometric assay *vs.* a panel of PRMTs and use of the MALDI-MS assay described in the preceding chapter.

5.4.2 – Radiometric Assay

Access to a panel of PRMTs (1,3 and 5-8) was facilitated through a collaboration with the SGC, Toronto. Compounds were tested by Dr M. Eram using a radiometric assay against all six of the available PRMTs, which were purified by Dr T. Hajian and Dr E. Gibson. PRMT was incubated with peptide substrate, CT_3-SAM and inhibitor for 30 min at 23 °C. After being quenched with 7.5 M guanidine HCl, samples were washed and analysed using a scintillation counter. Raw data processing was carried out by Dr M. Eram but all subsequent analyses and conclusions are the author's own.

PRMT1 and PRMT5 are of particular interest given the previous work in our laboratory so data for these two enzymes are examined first (**Figure 5.4**).

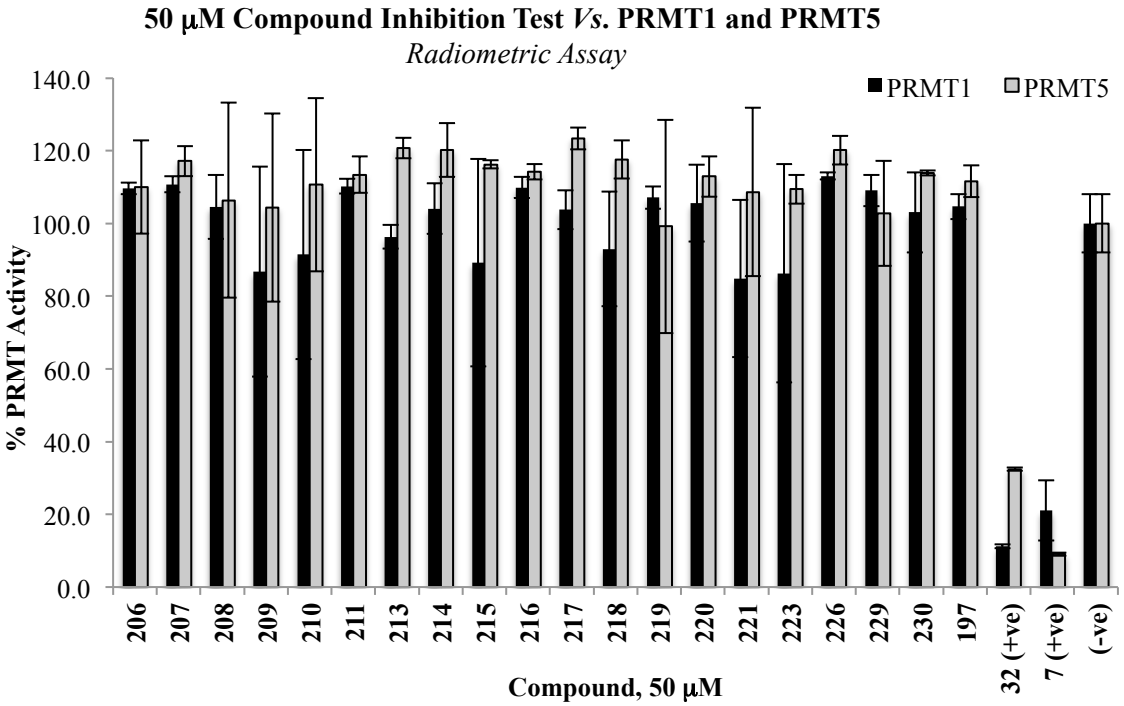


Figure 5.4 – *Inhibitory tests of a select group of compounds at 50 μM* vs. *PRMT1 and PRMT5 using the radiometric assay. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin **32** and/or SAH 7. Data is presented as the mean ± standard deviation and n = 3 for each data point. Assays were performed by Dr M. Eram of the SGC, Toronto.*

The strong inhibition observed for positive controls SAH **7** and Sinefungin **32** demonstrate that this set of results is most likely not artefactual (**Figure 5.4**). Very little inhibitory activity was observed for all tested compounds against PRMT1. Without exception, those compounds that exhibited some activity (*i.e.* more than 10%) had large associated standard deviations, so no conclusions as to their possible inhibitory prowess were made.

The situation was arguably worse for PRMT5, with no inhibition recorded at all. In fact, it appears as though most of the compounds are having a mild activating effect on the protein. The most disappointing result was that DPSO **197** did not show any inhibition of either PRMT1 or PRMT5.

The compound library was also screened at 50 μM against the other accessible PRMTs, but very similar conclusions were drawn for these compounds *vs.* PRMT3 and PRMT6-8 (**Figure 5.5**). During hit identification, the DPSO **197** scaffold showed little to no inhibitory activity towards PRMTs 1,3,4 and 6 so it could be argued that this result was to be expected.

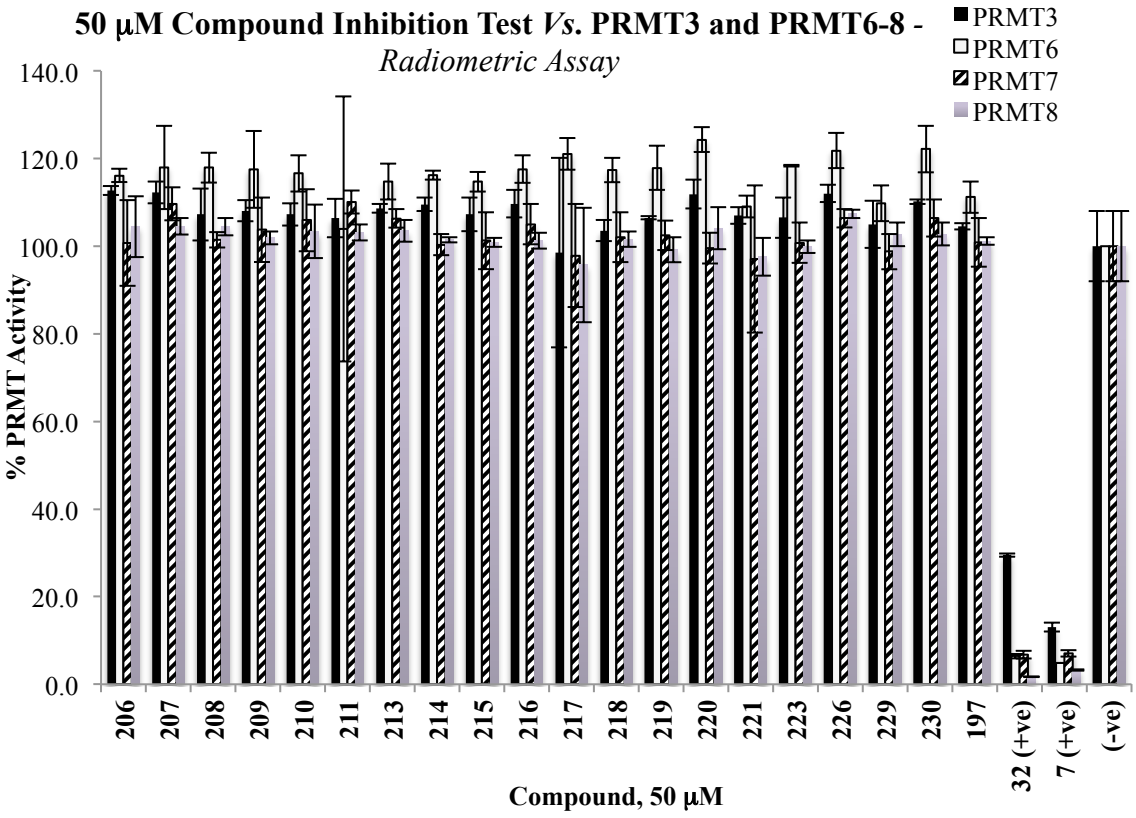


Figure 5.5 – *Inhibitory tests of a select group of compounds at 50 μM* vs. *PRMT3,6-8 using the radiometric assay. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin* ***32*** *and/or SAH 7. Data is presented as the mean ± standard deviation and n = 3 for each data point. Assays were performed by Dr M. Eram of the SGC, Toronto.*

With these results from the radiometric assays in mind, it was considered that the original hits identified in **Scheme 5.1** and **Scheme 5.2** might have been false positives i.e. the apparent decline in PRMT activity observed in previous work was not a result of true PRMT inhibition but instead was of a result of an assay artefact/inconsistency. One point to consider here is that the majority of the tested compounds had relatively high cLogP values (summarised in **Figure**

5.11) meaning they are poorly soluble and likely prone to aggregation under aqueous assay conditions, which has been shown to cause false positive effects in HTS assays[205]. Amongst other factors, the final percentage of DMSO (vehicle) in the assay would likely have an effect on the propensity for aggregation exhibited by this series of compounds. The radiometric assay described above[131] was carried out at a final assay concentration of 0.25% DMSO for the 50 μM inhibition tests. The chemiluminescence assay had been carried out at a final concentration of 1% DMSO as per manufacturer's instructions. Subsequent analyses using the MALDI-MS assay were also carried out at a final concentration of 1% DMSO (*vide infra*). Original screening that identified DPSO **197** as a lead was carried out at 1-2% DMSO. A DMSO-dependency test should be carried out to establish the importance of DMSO concentration in assays as part of future work. As a cursory test to investigate the effect of assay DMSO concentration on the inhibitory prowess of the current series of compounds, the radiometric assay was carried at by Dr M. Eram at a final DMSO concentration of 1% and using three concentration points of test compounds (1, 10 and 50 μM) (**Figure 5.6**). Positive and negative controls behaved as expected.

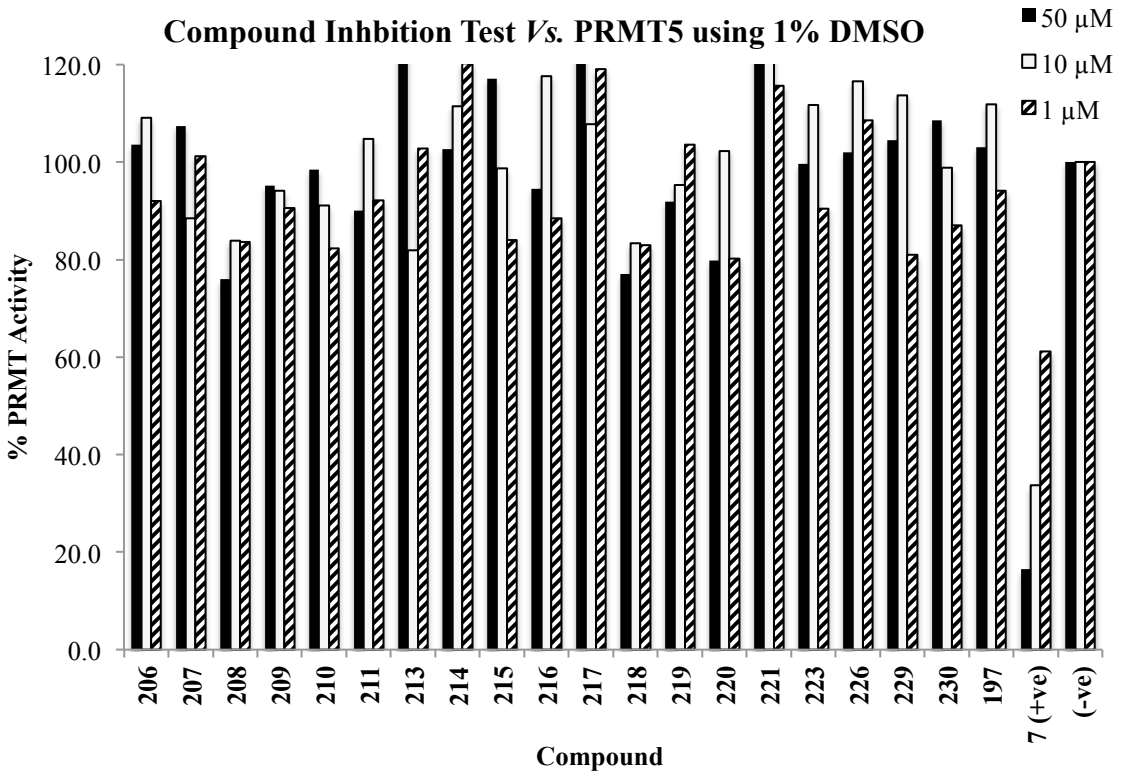


Figure 5.6 - *Inhibitory tests of a select group of compounds at 1, 10 and 50 µM* vs. *PRMT5 using the radiometric assay with a final DMSO concentration of 1%. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin **32** and/or SAH **7**. Data for this graph was collected in singlicate only. Assays were performed by Dr M. Eram of the SGC, Toronto.*

It was anticipated that some inhibition might have been observed with this increased DMSO concentration but these results do not support this. However, it is heavily emphasised that this data set was collected in singlicate only and so no error can be given and the data should therefore be treated with extreme caution.

The MALDI-MS assay described in Chapter Four avoids many of the problems associated with the radiometric and antibody-based assay methods and is able to deliver enriched data that distinguishes between mono- and di-methylated Arg. The original series of compounds (**216-211, 214, 221, 223**) synthesized as part of this work and additional compounds, (**213, 215-220, 226, 229**) were tested using this newly developed protocol; the assay is reliable and robust, as

discussed in the previous chapter, and it was thought to have the power to confirm or refute the identity of this series as true PRMT inhibitors.

5.4.3 – MALDI-MS Assay

The compounds were tested at 50 μM *vs.* PRMT1 using the MALDI-MS assay following the exact protocol outlined in the Experimental Section. It is important to note here the two different methods of calculating inhibition (**Figure 5.7**).

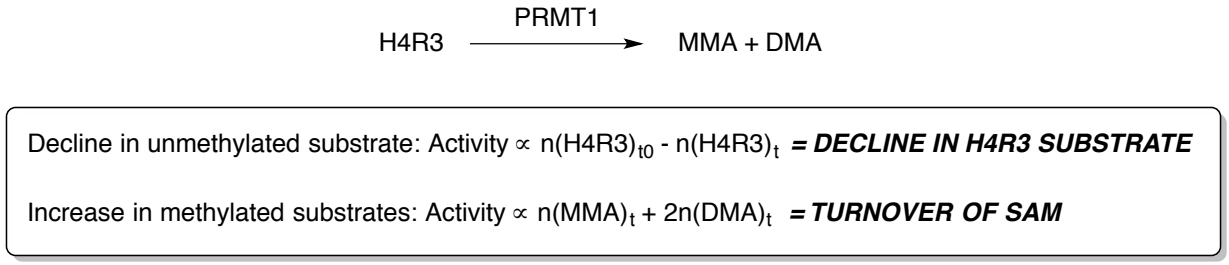


***Figure 5.7** – Calculating inhibition of PRMT1 using the MALDI-MS assay.*

For the former technique, the decline in starting substrate (H4R3) was monitored. For the latter technique the increase in products were calculated - the di-methylated peptides were given a double weighting during calculation because they represent a double turnover with respect to unmethylated substrate. This latter technique is analogous to monitoring the turnover of SAM.

The following inhibitory values were obtained when using the first method that monitors decline of unmethylated substrate (**Figure 5.8**). Positive and negative controls were as expected, although quite high error was observed for the negative control.

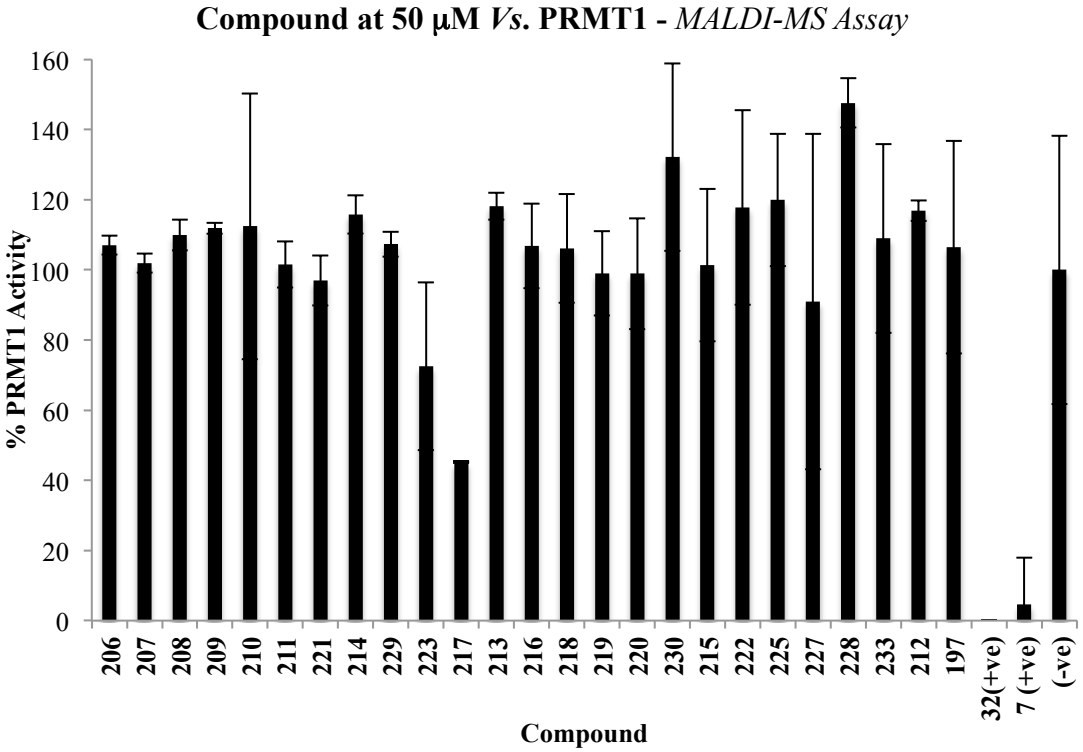


Figure 5.8 – *Inhibitory tests of compounds at 50 μM* vs. *PRMT1 using the MALDI-MS assay. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin* ***32*** *and/or SAH* ***7****. Data is presented as the mean ± standard deviation (n = 3) and were processed according to decline in starting substrate.*

Within error, zero inhibition was observed for almost all of these compounds at 50 μM *vs.* PRMT1; **217** was the exception showing 55% inhibition. It is noted here that these data represent single concentration points, and for a comparison of their inhibitory prowess concentration-dependency would need to be tested for[204].

The raw data was also processed using method two, which is the most similar analysis to the radiometric assay but avoids the problem of erroneous signal that could arise from non-specific radioactive readout. **Figure 5.9** below depicts the difference in inhibition observed when using each of these methods; in both cases the positive and negative controls gave the expected outcomes.

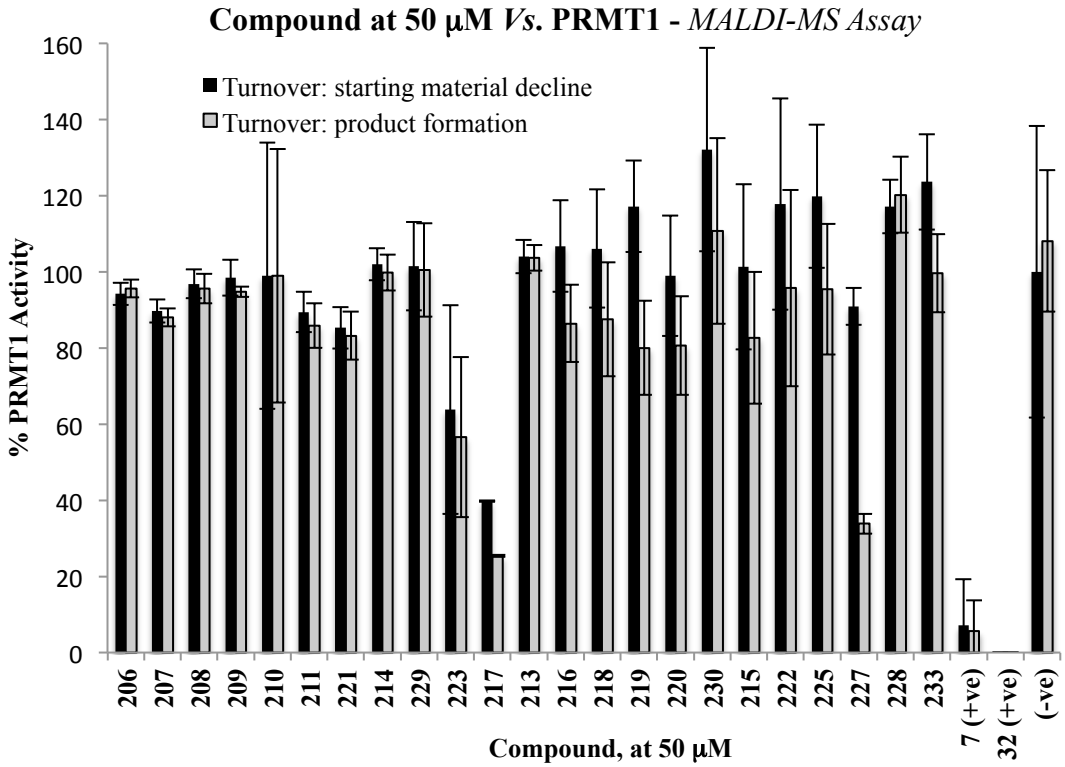


Figure 5.9 – *Inhibitory tests of compounds at 50 μM* vs. *PRMT1 using the MALDI-MS assay. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin* ***32*** *and/or SAH 7. Data is presented as the mean ± standard deviation (n = 3) and were processed according to either decline in starting substrate or increase in methylated products – a comparison of which is offered.*

For the majority of cases there is no apparent difference between inhibition calculated from the two different methods of processing – *i.e.* each mean for method one is within ≤1 standard deviation of the mean for method two. For the purposes of this work, it was considered that these methods of analysis proffer equivalent outcomes. The notable exceptions are for **217** and **227** where a marked difference is recorded.

It had in fact been hypothesised there would be a much greater difference in the outcomes of processing data according to methods one and two. The preliminary evidence above could be interpreted as suggesting it is unnecessary to collect data that indicates the distribution of mono

and dimethylated products. This is a promising result for the reliability of the radiometric assay, which is unable to distinguish such products and whose analysis would be expected to give broadly similar results to the second method of processing used for the MALDI-MS data. This comment is made with some caution though and it is anticipated that further data and in-depth statistical analysis will be required to make a firm conclusion as to whether use of different PRMT assays, with different monitoring protocols, has an effect on phenotypic-relevance. The next section gives a brief comparison of results from testing against the three aforementioned assays.

Overall, results from this section using the MALDI-assay to test inhibitory potency suggest that this series of compounds are not PRMT1 inhibitors, except **217** that gave modest inhibition at high concentration.

5.4.4 – Inter-Assay Comparisons

Figure 5.10 depicts the differences in inhibitory values recorded for each compound tested at 50 μM *vs.* PRMT1 using the three different assays. When comparing results from the radiometric assay and the MALDI-MS assay, the vast majority of inhibitory values fall within 1-2 standard deviations of the mean from each assay. The chemiluminescence data is included for those compounds tested but is excluded from comparison given the shortcomings of the assay (*vide supra*) and the likelihood that it is therefore artefactual.

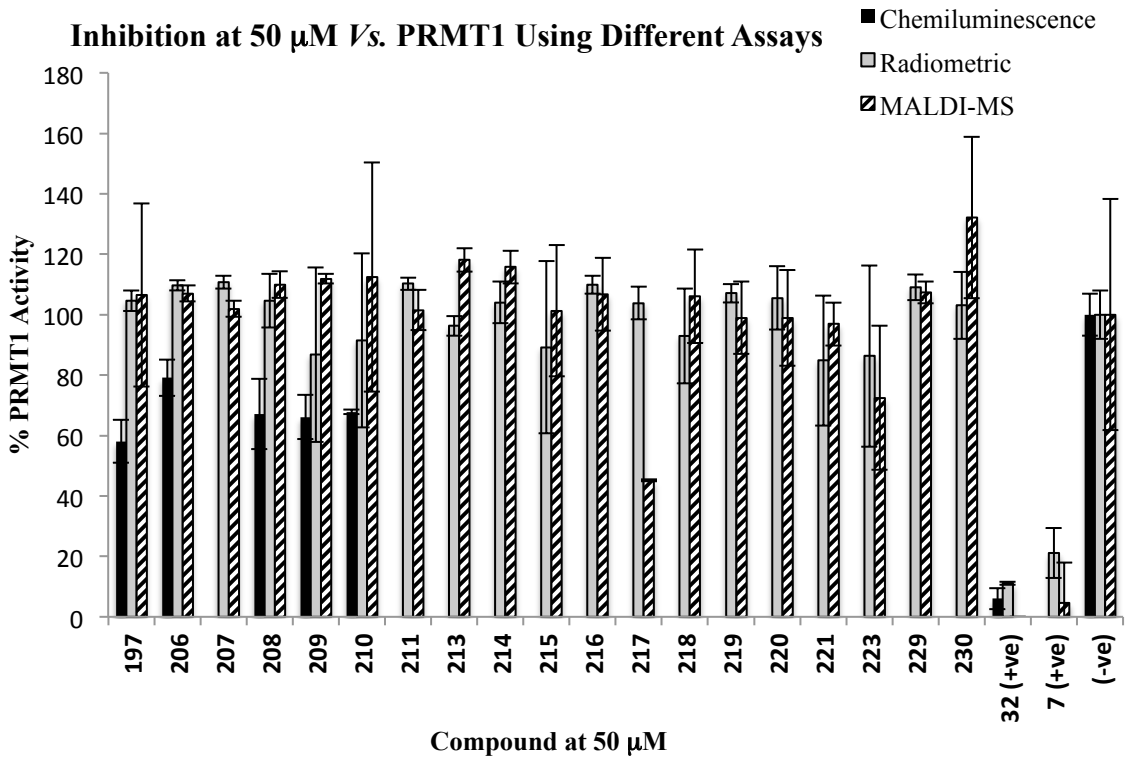
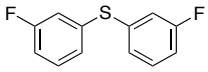
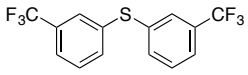
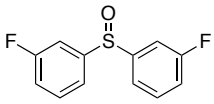
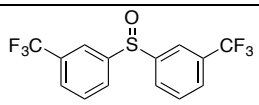
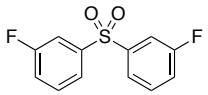
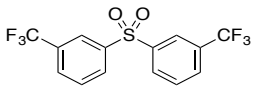
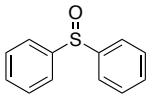
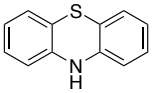
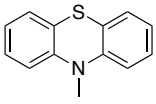
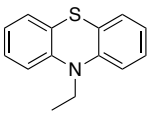
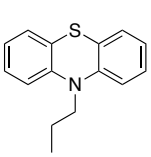
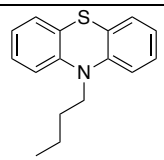
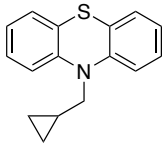
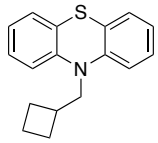
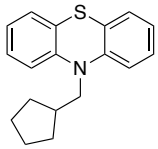
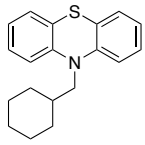
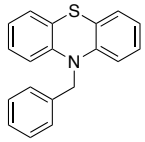
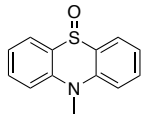
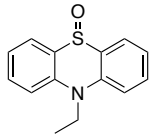
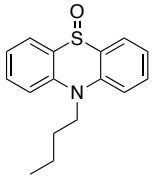


Figure 5.10 – *Inhibitory tests of compounds at 50 μM* vs. *PRMT1 using the chemiluminescence assay, the radiometric assay and the MALDI-MS assay. All assays were carried out in triplicate: % activity is pinned to a negative control containing no inhibitor. Controls containing no enzyme and containing no SAM were carried out, as were positive controls using Sinefungin* ***32*** *and/or SAH* ***7****. Data is presented as the mean ± standard deviation (n = 3). MALDI assay data is processed according to decline in starting peptide.*

Compound	Structure	cLogP	Inhibition, %		
			Chemiluminescence assay	*Radiometric assay*	*MALDI-MS assay*
206		4.85	21 ± 6	-10 ± 2	-7 ± 3
207		6.37	-	-11 ± 2	-2 ± 3
208		2.57	33 ± 12	-5 ± 9	-10 ± 4
209		4.19	34 ± 7	13 ± 29	-12 ± 2
210		2.89	32 ± 1	9 ± 29	-12 ± 38
211		4.49	-	-10 ± 2	-2 ± 7
DPSO, 197		2.09	42 ± 7	-5 ± 3	-7 ± 30
212		3.56	-	-	-17 ± 3
213		4.13	-	4 ± 3	-18 ± 4
214		4.66	-	-4 ± 7	-16 ± 5
215		5.19	-	11± 29	-1 ± 22

216		5.72	-	-10 ± 3	-7 ±12
217		5.11	-	-4 ± 5	55 ± 0
218		5.66	-	7 ± 16	-6 ± 16
219		6.22	-	-7 ± 3	1 ± 12
220		6.78	-	-6 ± 11	1 ± 16
221		6.40	-	15 ± 22	3 ± 7
222		1.48	-	-	-18 ± 28
223		2.01	-	14 ± 30	27 ± 24
225		3.07	-	-	-20 ± 19

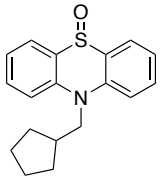
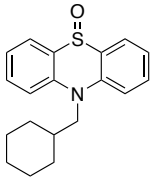
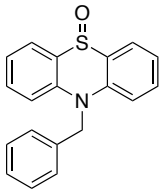
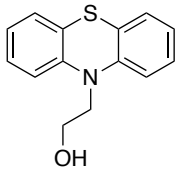
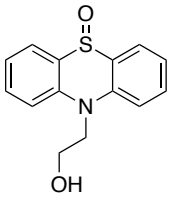
227		3.57	-	-	9 ± 48
228		4.13	-	-	-48 ± 7
229		4.45	-	-9 ± 4	-7 ± 4
230		3.14	-	-3 ± 11	-32 ± 27
233		0.49	-	-	-9 ± 27

Figure 5.11 – *Cross-assay comparisons of inhibition data* vs *PRMT1.* ***217*** *is highlighted as the only example of moderate inhibition (at 50 μM).*

When both radiometric and MALDI-MS assay data are considered it is concluded that all these compounds, including **217** and **227**, are unlikely to be true PRMT1 inhibitors and that in general no appreciable difference in assay readout is observed between these two techniques. It might prove interesting to submit **217** to further testing (in particular concentration-dependent testing) given the good level of inhibition suggested by MALDI-MS data, but it is anticipated that this result may be anomalous.

DPSO **197** was originally identified as a possible PRMT1 and PRMT5 inhibitor using a combination of radiometric and chemiluminescence assays. Subsequent work - reported herein - showed a reduction in PRMT1 and PRMT5 activity in the presence of DPSO **197** using a chemiluminescence assay (**Section 5.4.1**). However, this assay relied upon antibody-based detection of methylated end-products, which has it's associated disadvantages[101] and the assay returned inconsistent results upon repeated testing, so this data was treated with extreme caution and was effectively disregarded. The radiometric assay was subsequently used to test this series of compounds and DPSO **197** was found not to inhibit PRMT5 or PRMT1. Results from the MALDI-MS assay also provided support that DPSO was not an inhibitor of PRMT1.

Many of the compounds synthesised as part of this study had high cLogP values, which could go some way to reconciling the differing observations discussed herein. Compounds with high cLogP values are prone to aggregation[183] in aqueous systems (i.e. the assay environments used herein) and aggregates of compounds are notorious for bringing about false positives during HTS[205]. It is postulated that the original work suffered from this aggregation-based effect and that DPSO **197** was erroneously identified as an inhibitor of PRMT1 and PRMT5, when in fact may have been a false positive.

In light of all the evidence discussed in this chapter, it is recommended that this series is not further pursued. It would be interesting to perform a similar analysis of all other reported PRMT inhibitors given their structural similarity to the sulfone/phenothiazine derivatives (**Section 1.5**).

5.4.5 – Aggregation Studies

In an attempt to observe aggregation, ^{1}H NMR was employed. For compounds that aggregate, it is often possible to observe a 'line-broadening' effect[h] whereby the resolution of spectrum multiplets become less defined and flattened (shorter, broader peaks – for instance see Foster *et al.*[206]). Preliminary investigations tested DPSO **197** at different concentrations in D_2O up to and including 50 μM – the concentration used for the majority of assays described herein. The intention had been to carry out analogous tests using deuterated buffers to replicate assay conditions but unfortunately the test experiments in D_2O showed very weak spectra that were not amenable to 'line-broadening' analyses, so this work was not further pursued.

Dynamic light scattering (DLS) could be used in future work to make quantitative assessments of the aggregation properties of this series of compounds. A precedent exists for using DLS to assess the aggregation of phenothiazine-based drugs[207].

5.5 – Cheminformatics

5.5.1 – PAINS Filter

Pan assay interference compounds (PAINS) is a recognised concept first described by Baell[208] and refers to chemical substructures that are likely to give false-positive or artefactual results during high throughput screening for hit identifications. It is possible to use computational techniques to identify such compounds, and in practice a PAINS filter would ideally be used at the pre-screening stage. As part of the current work, an open-access PAINS filter was adapted

[h] Increasing the molecular weight of a compound results in slower tumbling and altered relaxation time that leads to a 'line-broadening' effect. Aggregation can be considered a form of apparent molecular weight increase.

and utilised for assessing whether any of the aforementioned series of compounds are classified as PAINS.

Knime[i] is an open-source data analysis platform that operates on a workflow basis using connections of nodes that perform specific tasks and have specific data outputs. Users can code their own nodes and/or construct their own workflows as well as use or adapt existing ones created by the Knime community. Saubern *et al.* published their PAINS filter workflow[209], which was adapted as part of this project (**Figure 5.12**).

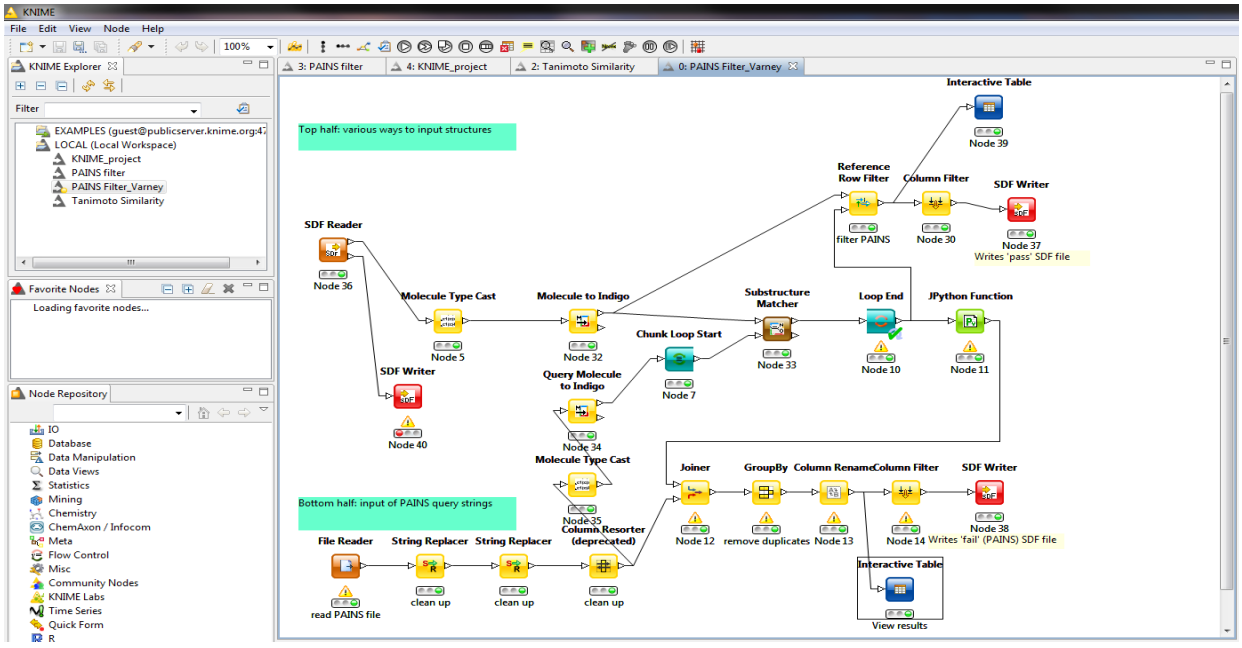


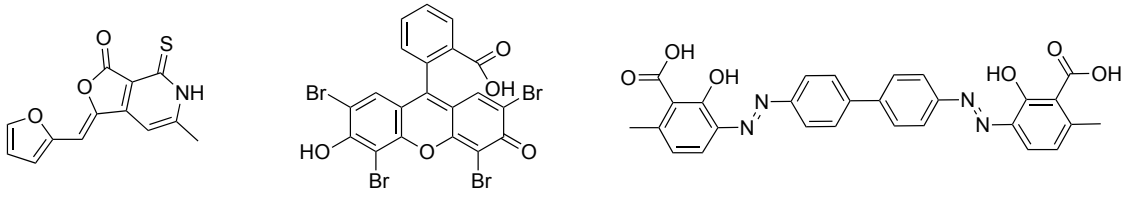
Figure 5.12 *– Knime workflow adapted from Saubern* et al.[209] *that scans libraries of compounds for PAINS (pan-assay interference compounds), which are postulated to give false-positive results in high throughput screening programmes. The bottom left node (orange: 'read PAINS file') can be altered so that the user can change the parameters that define a compound as a PAINS or not. The parameters used for this work are the same as those described in the original work*[209]*. The output of this workflow is a list of 'pass' and 'fail' compounds. The top right of the workflow outputs the 'pass' compounds and the bottom right the 'fail' compounds: the blue nodes allow this data to be viewed as interactive tables and the red nodes generate a chemical structure-data file (SDF).*

Every compound synthesised as part of this project passed the PAINS filter, as did DPSO **197** and phenothiazine **197**. This result is an indicator that the compounds are unlikely to have caused

[i] Knime software can be accessed *via* https://www.knime.org/

during HTS, although this does not completely preclude them from causing off-target effects since the PAINS filter is only as good as its defined parameters.

The original HTS in our laboratory had been carried out on a focused library set of 100 molecules, each of which had been cherry-picked for structural similarity to compounds reported in Bedford[99] and Jung's[100] work. As part of the current work, a database of all reported PRMT inhibitors was constructed and also run through the PAINS filter. Interestingly three of the inhibitors from Bedford's seminal paper failed owing to presence of azo-, enone-, halogenated quinone and unsaturated five-membered heterocycle functionalities (**Figure 5.13**). Possibly the presence of PAINS early on in the PRMT inhibition field, and also in the compound library that was screened as the starting point for this work, meant that our initial HTS was carried out against a library containing possible PAINS. Such discrepancies early on in the field may have manifested itself downstream as inconsistent results and/or false inhibition in later generations of inhibitor series. This theory may go some way to explaining the results described herein for the series of compounds under investigation.



AMI-2 **267** **AMI-5** **268** **AMI-9** **269**

Figure 5.13 *– Molecules reported as PRMT inhibitors in the seminal work by Cheng* et al.[99] *that were classified as PAINS (pan-assay interference compounds) by the Knime-based PAINS filter depicted in* ***Figure 5.12****.*

5.6 – Conclusions and Future Work

This chapter presents the synthesis and testing of a series of *bis*-aryl and tricyclic compounds based on early SAR development that led to hit DPSO **197** and subsequently to its substituted and ring-restricted derivatives. An Ullman-type coupling was used to access *meta*-EWG substituted *bis*-aryl sulfides in very good yields from the requisite thiophenols and aryl iodides. These were oxidised to sulfoxides *via* a Gif-type reaction and to the sulfones *via* a Prizelhaev-type reaction, both in reasonably good yields, to furnish a small *bis*-aryl library.

Similarly, a library of ring-restricted phenothiazines was prepared; *N*-substituted phenothiazines were synthesised from requisite alkyl halides under varying substitution conditions in poor to good yields. Stoichiometric *m*CPBA was used to generate the sulfoxide derivatives in moderate to good yields but owing to solubility problems, synthesis of the corresponding sulfones was not attempted. More polar substituents were not amenable to a direct substitution reaction but were introduced using individually developed protocols. The 2-hydroxyethyl substituent was incorporated by an elegant reaction with cyclic sulphate **231** in good yield and the sulfoxide derivative was produced using *m*CPBA also in good yield. Oxidation was confirmed to have selectively occurred at sulfur rather than at the amine or alcohol by X-ray crystallography.

Other methods explored for synthesis of more polar functionalities focused on coupling. A three-component click chemistry reaction utilising 2-bromothiophenol, 2-iodobromobenzene and a requisite substituted primary amine was reported to generate *N*-substituted phenothiazines[192]. This palladium and dppf-catalysed reaction was used to synthesise *N*-propylmorpholino-phenothiazine **241** as the major product, although the minor product (the non-ring closed intermediate) was challenging to remove from the crude mixture despite multiple purification attempts. This unfortunately excluded **241** from testing against PRMTs given the 7% ring-open

impurity. Attempts were made to modify this procedure to a two-component coupling utilising 2-iodoanilines in place of 2-iodobromobenzene and a primary amine. This procedure was used to synthesis mono- and di-fluorinated phenothiazines in an attempt to investigate the effect of *meta*-EWG substituted tricyclic structures on PRMT activity. Unfortunately the resultant compounds were unstable, undergoing spontaneous oxidation to the *N*-oxides as assessed by ESI and APCI LC-MS/MS.

A library of 25 viable compounds was taken forward for testing *vs.* PRMTs. Three different assay techniques were used: a chemiluminescence, a radiometric and a MALDI-MS assay. The original hit DPSO **197** was originally shown to inhibit PRMT1 and PRMT5 during hit identification and early SAR. However, this result could not be replicated here and it was concluded, following inter-assay data comparisons and a consideration of the limitations of each assay technique, that DPSO **197** was not a true PRMT inhibitor. It is possible that DPSO **197** had been a false positive hit from the screen against a cherry-picked library in the original work. This idea is substantiated by two observations:

1. Firstly, hydrophobic compounds with high cLogP values have been shown to self-aggregate and cause false positive readouts under aqueous assay conditions[205]. The original hit **196** had a cLogP value of 6.89 and could conceivably have represented a false positive. Hit **196** was the precursor to DPSO **197** and although the latter had a lower cLogP value (2.02) it is plausible that a similar aggregation effect was being observed, especially given the planarity of the molecule which facilitates π-π stacking.

2. Secondly, the majority of early SAR development using DPSO **197** had been carried out with an assay using antibody-based detection. Throughout this thesis, the suitability of antibodies

against methylated arginines has been called into question. In 2011 it was reported that more than 25% of anti-histone antibodies described in the scientific literature are not fit for purpose[101].

A further hypothesis of this thesis had been that the inability to distinguish between mono- and dimethylated- arginine was a disadvantage for PRMT assays. The MALDI-MS assay, development of which is described in the preceding chapter, is able to provide enriched data that shows the distribution of un-, mono- and di-methylated arginine-containing peptides. This was exploited in the current chapter and used to compare the difference in inhibitory readout from processing the data in two different ways: turnover was monitored firstly by calculating decline in unmethylated substrate and second by monitoring increase of the two methylated products. For the latter technique the di-methylated peptides were given a double weighting during calculation because they represent a double turnover with respect to unmethylated substrate. No appreciable difference for most compounds was noted between these two methods of processing, which is evidence lodged against the theory that this peptide distribution metadata is important for inhibitory potency. This observation also tentatively suggests that the radiometric assay, the current literature gold-standard, is able to generate accurate data, which is promising for the wider PRMT field that so relies upon this assay. It is heavily emphasised though that a great deal of further work is required to make a firm and final conclusion on this matter. Certainly a much larger sample size will be required and further manipulation of the MALDI-MS data is necessary. In particular, it is thought that kinetic parameters could be applied to MALDI-MS data outputs to make some sort of assessment as to whether an inhibitor might be competing with any of the three relevant peptides and if inhibitors might have a different effect on different phases of catalysis. This will require in-depth and kinetic analyses outside the scope of this project.

Regrettably it was concluded that the series described herein are most likely not true inhibitors of PRMT1,3,5,6,7 or 8. **217** showed some inhibition in the MALDI-MS assay, but not in others, and may be worth following up but should be rigorously tested as a true inhibitor prior to any subsequent attempted development. Of particular relevance would be to carry out an aggregation assay that investigates detergent-dependency[205] (e.g. Triton/Tween), and a concentration-dependency assay to establish if the compound is suffering from single-concentration point artefacts[204].

A PAINS filter was implemented to test whether or not the compounds might be interfering with assay readout but all compounds passed. However, some of the early inhibitors from the seminal work[99] were found to fail the PAINS test. The original screening library that identified DPSO **197** was constructed of cherry-picked compounds based upon these scaffolds; it is not inconceivable that the error of identifying DPSO **197** as a PRMT inhibitor originated here, especially since the paper was published in a pre-PAINS-aware era. However, it is emphasised that just because a PAINs filter flags a compound as a possible pan-assay interferer, does not necessarily mean that the compound is a false positive. It is reiterated that re-testing of early inhibitors/hits reported in the literature, in particular those that might now be considered PAINS, could prove to be a worthwhile and interesting venture.

In **Section 1.5**, the development of peptide-based and bisubstrate PRMT inhibitors were discussed. In light of the disappointing results obtained from the small molecule symmetrical inhibitors reported herein, peptide-based inhibitors might be a productive new avenue towards inhibiting PRMTs.

In particular, it may be possible to exploit the research presented in Chapters Two and Three to explore new chemical space that might allow development of highly specific PRMT probes. Reproduced below is **Figure 1.24** showing the scaffold for a PRMT inhibitor/probe described by Dowden *et al.*[128]. Addition of an MMA(δ) residue synthesised as part of this work (Chapter Three) at R_1 and/or R_2 could make for an interesting avenue of research.

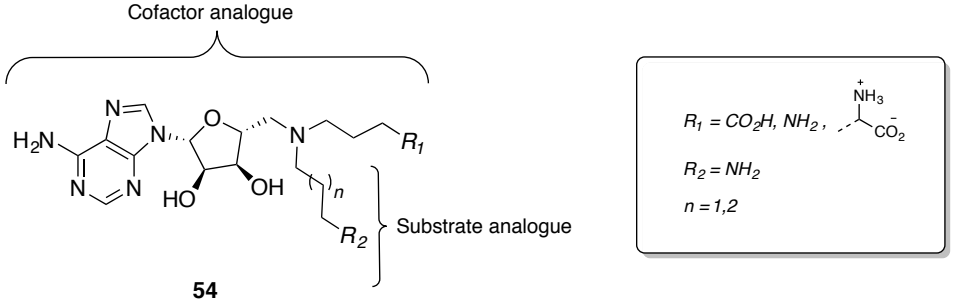
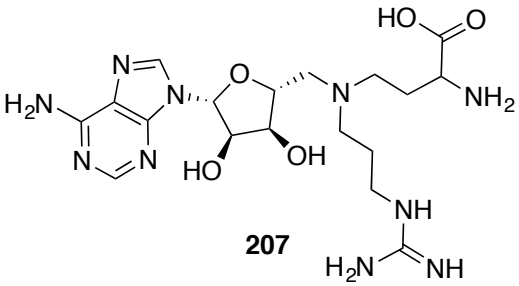


***Figure 1.24** – The first true adenosine-containing bisubstrate PRMT inhibitors.*

Furthermore, one of the aims for the future should be to build on the work presented in Chapters Two and Three and synthesise the full range of possible methylated arginines. Testing of differentially methylated arginines as part of histone peptides could be carried out using the MALDI-MS assay, developed as part of this work, against the whole panel of PRMTs. Any possible reactivity with and/or selectivity between PRMTs could be exploited in the design of specific PRMT probes. Future work may wish to explore this potentially exciting option. Some preliminary docking work was carried out on the hypothetical SAM-analogue **270** shown in **Figure 5.14A**. The guanidine terminus could be substituted with specific methylation patterns. **Figure 5.14B** shows cofactor-mimetic Sinefungin **32** (magenta) and a peptide substrate (grey) in the PRMT5 active site (PDB code: 4GQB)[59]. The guanidine terminus of the Arg residue in the substrate can be seen bridging towards the $–NH_2$ of Sinefungin, which would be replaced with the $–CH_3$ to be transferred in the cofactor SAM. The cofactor-binding pocket is shown in **Figure 5.14C** with sinefungin (magenta) and hypothesised structure **270** (green) superimposed. Both

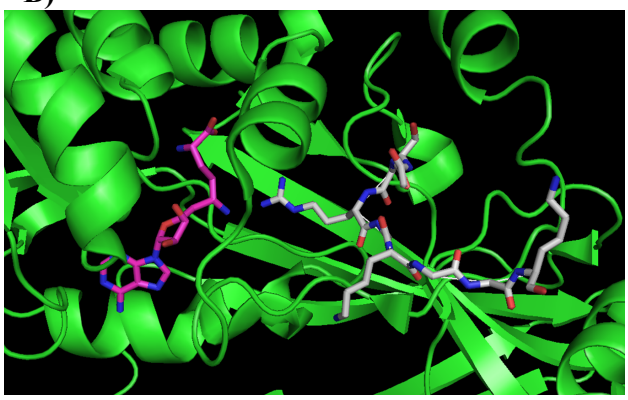
compounds were docked using Mol® software and represent predicted energy minima. The guanidine terminus of structure **270** can be seen protruding out of the plane of the page, towards the reader. The aim would be to fill the space in which the true substrate Arg would usually fit. Different methylation patterns at this position may allow for inter-PRMT specificity.

A)



207

B)



C)

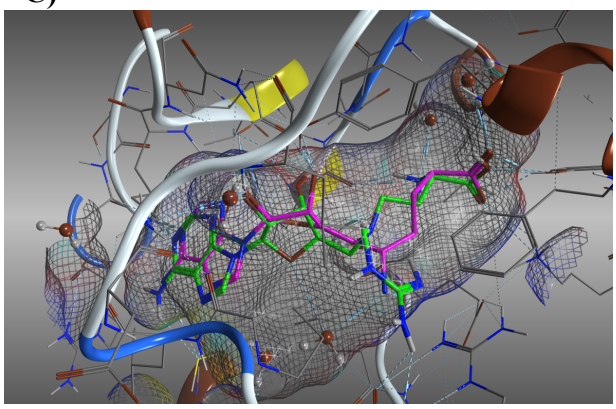


Figure 5.14 – A) *Hypothesised structure of a next generation PRMT inhibitor/probe based on bisubstrate inhibitors reported by Dowden* et al.[128] ***B)*** *Positioning of the cofactor mimetic Sinefungin (magenta) and peptide (grey) in the PRMT5 active site (PDB code: 4GQB)*[59] ***C)*** *Overlay of Sinefungin (magenta) and structure* ***270*** *docked into the PRMT1 active site using Mol® software.*

Thesis Summary

This thesis describes a multi-disciplinary project that had the following broad aims:

- Widen biological understanding of arginine methylation patterns in the context of the PRMT family of enzymes
- Develop a potent PRMT inhibitor

One of the major hypotheses of this work was that alternative Arg methylation patterns might exist and have physiological and/or epigenetic relevance. The assays used routinely in the PRMT field were reviewed; it was speculated that their associated disadvantages, namely their lack of selectivity for different methylated Arg substrates and products, could have prevented identification of novel methylated derivatives.

Work presented in this thesis was partly directed towards synthesis of standards of novel methylated Arg derivatives that could assist in testing such a hypothesis. One of the major successes of this work is the development of a 12-step route for large-scale synthesis of protected Me(δ)-containing derivatives that can be used in SPPS; the full route, including subsequent incorporation into a 16-mer peptide, was successfully implemented for synthesis of peptidyl-MMA(δ) methylation pattern.

Testing of this peptide for relevance as a possible PRMT1 substrate was made possible by development of a MALDI-MS assay that enables distinction between peptides with different methylation numbers. This assay was used to investigate possible PRMT1-catalysed polymethylation, but no evidence of higher order methyl patterns were observed using this technique.

Since PRMTs are epigenetic drug targets, a library of *bis*-aryl and tricyclic small molecules were prepared that were tested for inhibitory activity against a panel of PRMTs. The MALDI-MS assay developed as part of this work was utilised, alongside a radiometric and a chemiluminescence assay. No strong inhibitors were identified but the variability in results between the three assays was discussed and could serve as a platform for future work investigating the applicability of assays for monitoring PRMT activity.

Future work should aim to synthesise the full range of novel methylated Arg derivatives and use the MALDI-MS assay to make investigations as to their relevance as substrates and/or inhibitors of not just the PRMTs, but also other epigenetic enzymes.

Chapter 6 – Experimental Section

6.1 – Chemistry

6.1.1 – General Experimental

All reactions involving water-sensitive reagents were carried out under a nitrogen or argon atmosphere using standard *vacuum* line techniques and glassware that was flame-dried before use. Solvents were dried following the procedure outlined by Grubbs *et al*.[210] Water was purified by an Elix® UV-10 system. All other solvents and reagents were obtained from Sigma Aldrich, Fluorochem, Alfa Aesar, TCI Fine Chemicals UK or Acros and were used as supplied (analytical or HPLC grade) without prior purification. Organic layers were dried over Na_2SO_4. Brine refers to a saturated (sat.) aqueous solution of NaCl. Under *vacuum* refers to the use of a rotary evaporator attached to a diaphragm pump. Pet. ether refers to the fraction of petroleum spirit boiling between 30 and 40 °C. Thin layer chromatography was performed on aluminium plates coated with 60 F_{254} silica (Merck). Plates were visualised under UV light (254 nm) or by staining with either 1% (w/v) aq. $KMnO_4$ or 0.3% (w/v) ninhydrin in a solution of 3% (v/v) AcOH in *n*-butanol. Flash column chromatography was performed on Kieselgel 60 silica in a glass column. Melting points were recorded on a Gallenkamp Hot Stage apparatus and are uncorrected. Infrared spectra were recorded on a Bruker Tensor 27 FT-IR spectrometer, as neat or thin film samples. Selected characteristic peaks are reported in wavenumbers (cm^{-1}). Optical rotations were recorded on a Perkin-Elmer 241 polarimeter with a water-jacketed 10 cm cell at 22 °C or on a Schmidt Haensch Unipol polarimeter with a water-jacketed cell at 25 °C. Specific rotations are reported in 10^{-1} decomp cm^2 g^{-1} and concentrations in g/100 mL. NMR spectra were recorded on Bruker Avance spectrometers (DPX200, DQX400, AVIII400, AVII 500 or DRX500) in the deuterated solvent stated. The field was locked by external referencing to the

relevant deuteron resonance. Chemical shifts (δ) are reported in parts per million (ppm) relative to tetramethylsilane (TMS) where δ_H (TMS) = 0.00 and δ_C (TMS) = 0.00. Coupling constants (*J*) are quoted in Hz and are reported to the nearest 0.1 Hz. The coupling constants were determined by analysis using ACD Labs software. Low-resolution mass spectra were recorded on either a VG MassLab 20-250 or a Micromass Platform 1 spectrometer, operating in positive or negative mode, from solutions of MeOH. Accurate mass measurements were run on either a Bruker MicroTOF internally calibrated with polyalanine, or a Micromass GCT instrument fitted with a Scientific Glass Instruments BPX5 column (15 m × 0.25 mm) using amyl acetate as a lock mass, by the mass spectrometry department of the Chemistry Research Laboratory, University of Oxford, UK. *m/z* values are reported in Daltons. IUPAC names of all compounds are given except in the case of amino acid derivatives, which are referred to by their recognised names. Additionally amino acids are abbreviated for clarity according to the following rules: stereochemistry about the α-carbon is denoted by L- or D- prefixes. Standard three letter abbreviations are used for naturally occurring amino acids and additionally ornithine is abbreviated to 'Orn'. Protection of the α-amino group is denoted in brackets immediately following this code *e.g.* Orn(Fmoc). Any side-chain protection is denoted before the code *e.g. N*-Fmoc-Orn. The α-carboxylic acid group is denoted as either the acid (*e.g.* Orn-OH) or the relevant ester (*e.g*. Orn-OMe). An example: L-Ornithine with α-amino Fmoc protection, terminal amino Boc protection and α-carboxylate ethyl esterification would be denoted as follows: *N*-Boc-Orn(Fmoc)-OEt. In ^{1}H NMR spectra where two protons are attached to the same carbon atom but have different chemical shits (e.g. diastereotopic protons), the protons are arbitrarily defined as H_A and H_B (**Figure 6.1**) and are assigned accordingly.

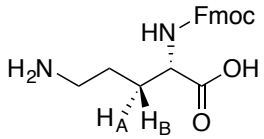


***Figure 6.1** – Example of diastereotopic protons defined as H_A and H_B for the purpose of ^1H NMR assignments.*

6.1.2 - General Procedure for Ion-Exchange Purification using Dowex® Resin

Dowex® (50WX8) was prepared according to the following washing procedure: 1 M $HCl_{(aq)}$ (2 column volumes, cv), H_2O (until eluent runs neutral), *x* M $NH_4OH_{(aq)}$ (2 cv), H_2O (2 cv), 1 M $HCl_{(aq)}$ (until eluent runs neutral). The crude mixture was dissolved in the minimum volume of water or $HCl_{(aq)}$ (up to 2 M) and added drop-wise to the stationary phase. In the case of an acidic mobile phase, the column was washed with H_2O until the eluent ran neutral. The column was washed with MeOH (2 cv) then H_2O (2-4 cv) and then the product eluted with *x* M $NH_4OH_{(aq)}$ (2-4 cv) under gravity.

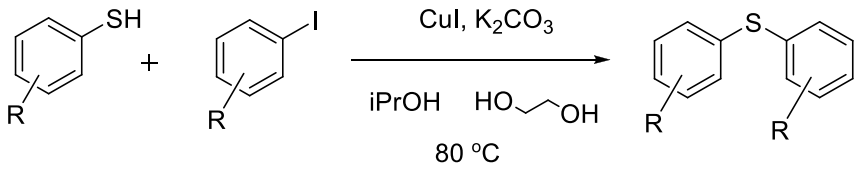
6.1.3 - Procedure for Separation of Orn and Arg using Ion-Exchange

A column of 2.5 cm diameter was filled with Dowex® resin, and prepared as described above, to a height of 12.5 cm. Following application of the crude mixture to the stationary phase, as described above, the column was washed with H_2O (2 cv). Aqueous 0.2 M $NH_4C_2H_3O_2$, pH 5.5 (approx. 1 L) was used to elute Orn. Gentle pressure could be applied manually *via* bellows up to a flow rate of ~3 mL/min. Following complete elution of Orn, analysed by TLC (5:1 MeOH:NH_4OH), the second eluent 0.5 M $NH_4C_2H_3O_2$, pH 6.8 (approx. 600 mL) could be applied to the column for elution of Arg under the same flow rate as before. Fractions containing product were pooled and concentrated under *vacuum* until a slurry formed. The general procedure for ion-exchange chromatography (**Section 6.1.2**) was employed to remove the salts from the product.

6.1.4 - General Experimental Procedures

General Procedure 1:

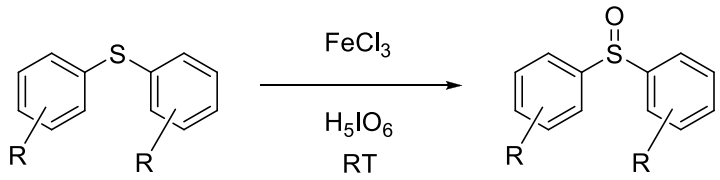
Ullman-type coupling for *S,S-bis*(aryl)sulfide coupling



CuI (0.05 eq.), K_2CO_3 (2 eq.) and the requisite aryl iodide (1 eq.) were added sequentially to a microwave vial before sequential addition of ethylene glycol (2 eq.), isopropyl alcohol (1 mL/mmol) and the requisite aryl thiol (1 eq.). The reaction mixture was heated at 80 °C for 24 h, cooled down, filtered and the filtrate concentrated under *vacuum*. The residue was purified on SiO_2.

General Procedure 2:

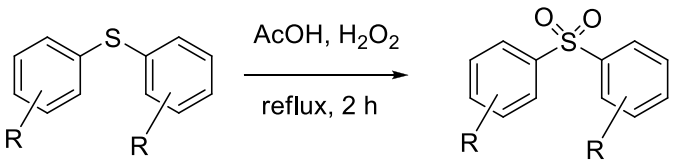
Oxidation of *S,S-bis*(aryl)sulfide to *S,S*-bis(aryl)sulfoxide



A solution of *S,S-bis*(aryl)sulfide (1 eq.) and Fe(III)Cl_3 (0.03 eq.) in MeCN (10 mL/mmol) was stirred for 5 min at room temperature before addition of H_5IO_6 (1.1 eq.), and the resulting mixture stirred for 16 h. The reaction was quenched with saturated $Na_2S_2O_{3\,(aq)}$ (25 mL/mmol) and extracted with CH_2Cl_2 (3x25 mL). The combined organic layer was dried, filtered, concentrated under *vacuum* and the residue was purified on SiO_2.

General Procedure 3:

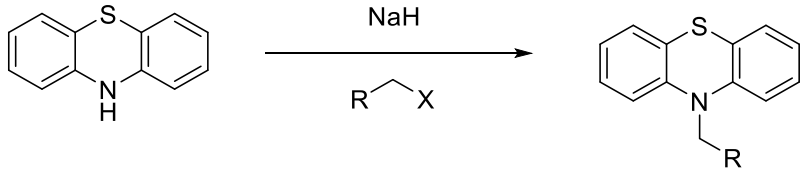
Oxidation of *S,S-bis*(aryl)sulfide to *S,S*-bis(aryl)sulfone



A solution of *S,S-bis*(aryl)sulfide (1 eq.) in glacial AcOH (5 mL/mmol) and $H_2O_{2(aq)}$ (33%, 5 mL/mmol) was refluxed for 2 h, cooled to room temperature and the solvent was removed under *vacuum*. The residue was dissolved in CH_2Cl_2 (50 mL/mmol) and washed with water (100 mL/mmol). The organic layer was dried, filtered, concentrated under *vacuum* and the residue was purified on SiO_2.

General Procedure 4a:

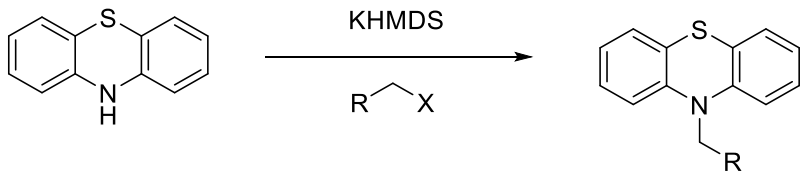
Alkylation of phenothiazine



Phenothiazine (1 eq.) was added to a suspension of NaH (1 eq.) in THF (10 mL/mmol) and the resulting mixture was stirred at room temperature for 30 min before drop-wise addition of the requisite alkyl halide (1 eq.), and the reaction stirred for 16 h. The mixture was treated with water (30 mL) and EtOAc (30 mL), the organic layer was washed with brine (30 mL) and the aqueous layer extracted further with EtOAc (3x50 mL). The combined organic layer was dried, filtered, concentrated under *vacuum* and the residue was purified on SiO_2.

General Procedure 4b:

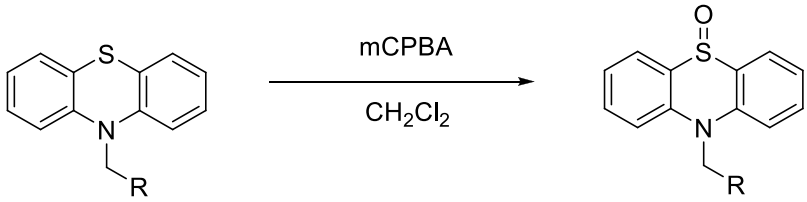
Alkylation of phenothiazine



A solution of phenothiazine (1 eq.) in THF (10 mL/mmol) was cooled to 0 °C before drop-wise addition of KHMDS (1.05 eq., 0.5 M in toluene), and the resulting mixture brought to room temperature and stirred for 15 min. The requisite alkyl halide (2.2 eq.) was then added drop-wise and the resulting reaction was stirred for 16 h. The mixture was treated with water, the aqueous layer extracted with Et_2O (30 mL) and EtOAc (2x30 mL). The organic layers were combined, washed with brine, dried, filtered, concentrated under *vacuum* and the residue was purified was purified on SiO_2.

General Procedure 5:

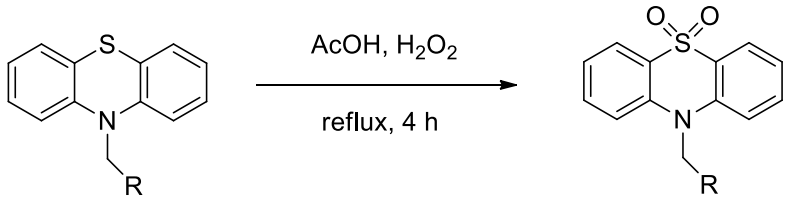
Oxidation of substituted phenothiazine to *S*-phenothiazine-sulfoxide



*m*CPBA (1 eq.) was added to a solution of phenothiazine (1 eq.) in CH_2Cl_2 (12 mL/mmol) and the reaction stirred for 16 h. The mixture was quenched with a saturated aqueous solution of $NaHCO_3$ and washed with water and the aqueous layer extracted with CH_2Cl_2 (3 x 50 mL). The combined organic layer was dried, filtered, concentrated under *vacuum* and the residue purified by silica column chromatography.

General Procedure 6:

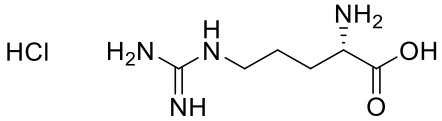
Oxidation of substituted phenothiazine to *S,S*-phenothiazine-sulfone



$H_2O_{2(aq)}$ (33%, 5 mL/mmol) was added to a solution of phenothiazine (1 eq.) in glacial AcOH (5 mL/mmol) and the resulting solution refluxed for 4 h. The reaction mixture was cooled to room temperature and the solvent removed under *vacuum*. The residue was dissolved in CH_2Cl_2 (50 mL/mmol) and washed with water (100 mL/mmol). The organic layer was dried, filtered, concentrated under *vacuum* and the residue purified by recrystallisation.

L-Arginine HCl[211], 1

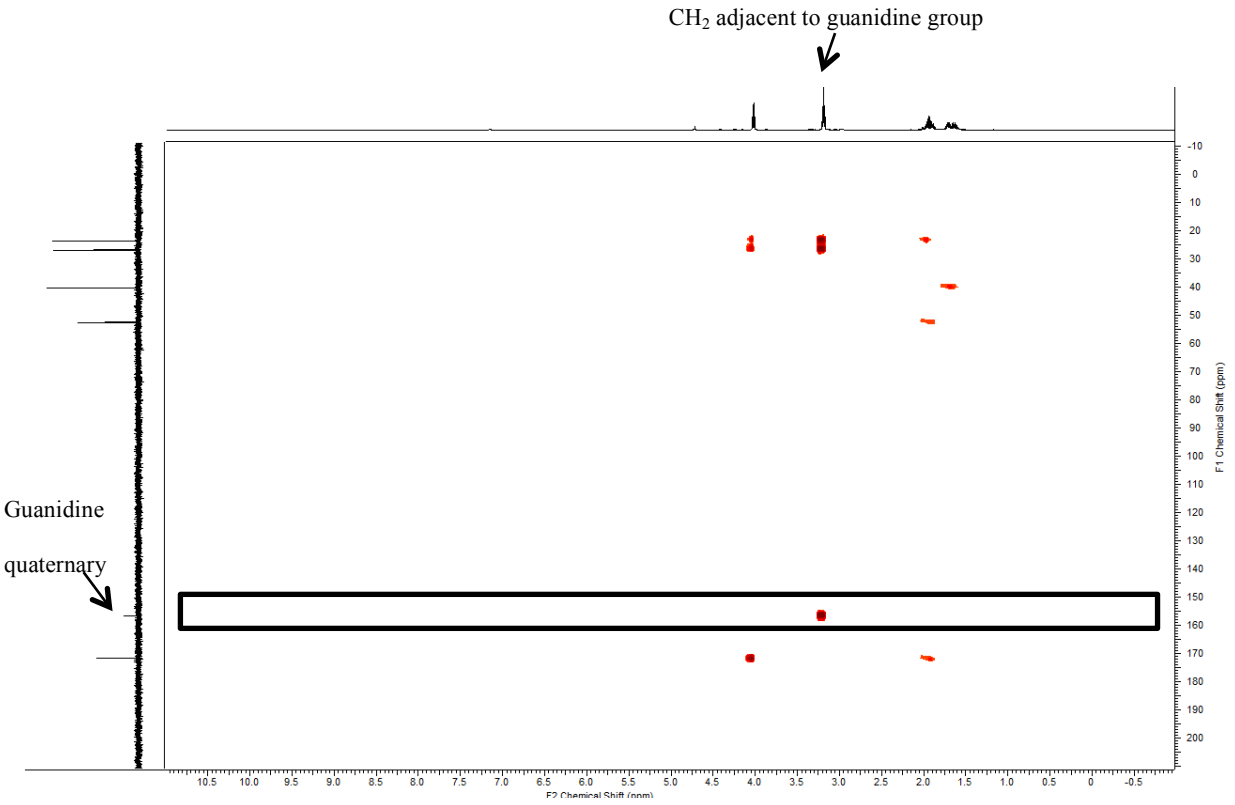
Arg-OH



Method A:

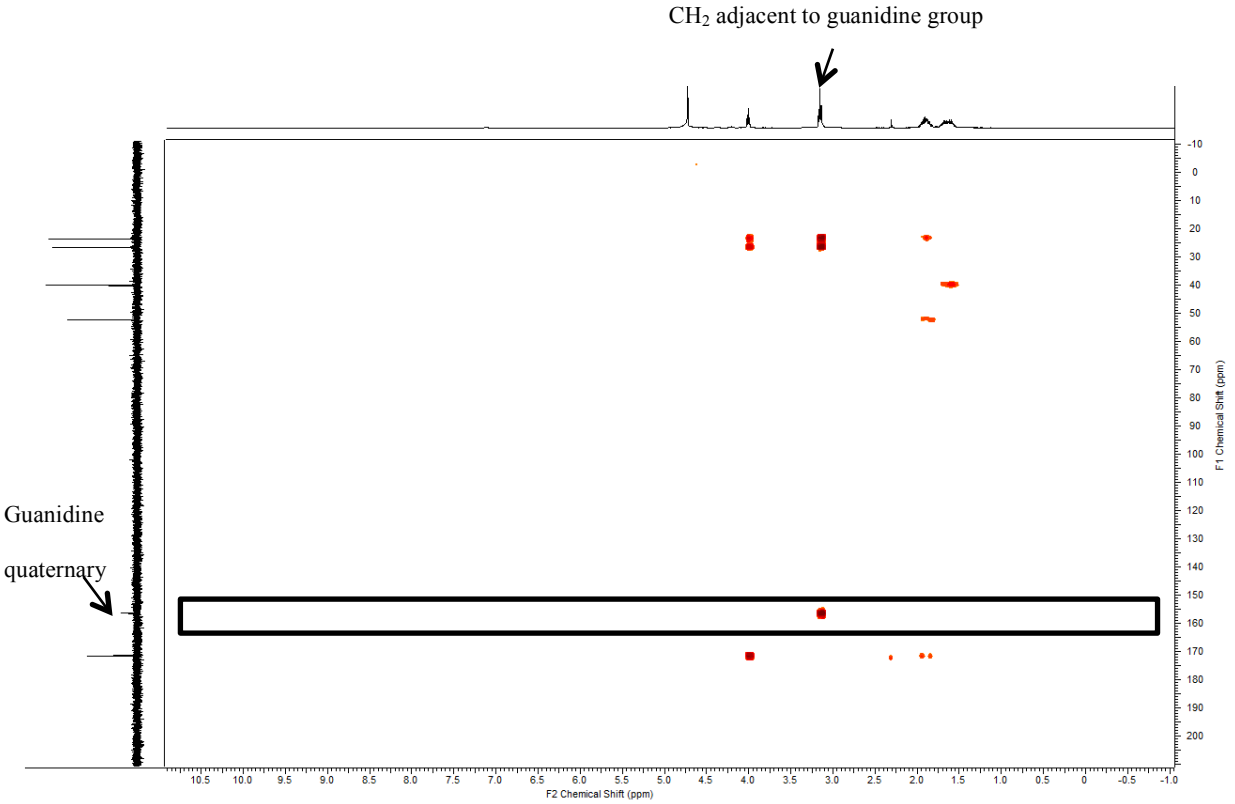
1*H*-Pyrazole-1-carboxamidine hydrochloride (1.13 g, 7.71 mmol, 1.3 eq.) was added to a solution of L-Orn·HCl (1.00 g, 5.93 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (8 M, 2 mL), and the reaction stirred for 5 d at room temperature. The mixture was acidified to pH 6 with $HCl_{(aq)}$ and evaporated under *vacuum* to leave ~1 mL of solution. Ion exchange chromatography (**Section 6.1.2**) was used to afford the title compound as a brown oil (330 mg, 27%) after washing with 1 M $HCl_{(aq)}$; $[\alpha]_D^{22}$ +12.7 (*c* = 1.0 in 6 M $HCl_{(aq)}$) (lit.[212] $[\alpha]_D^{20}$ +22 (*c* = 8 in 6 M $HCl_{(aq)}$)); ^{1}H NMR δ (400 MHz, D_2O) 1.55-1.76 (2H, m, $RNHCH_2C\mathit{H}_2CH_2$), 1.84-2.01 (2H, m,

RNHCH$_2$CH$_2$C*H$_2$*), 3.18 (2H, t, *J* = 6.9 Hz, RNHC*H$_2$*CH$_2$CH$_2$), 4.01 (1H, t, *J* = 6.3 Hz, CO$_2$C*H*); LRMS *m/z* (ESI$^+$) 175 [M+H]$^+$. HMBC data proving guanylation has been regioselective at the δ-NH$_2$ of Orn:



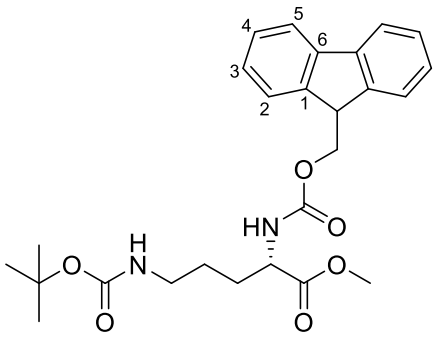
Method B:

2-Methylisothiouronium iodide **93** (1.23 g, 5.64 mmol, 1.9 eq.) was added to a solution of L-Orn·HCl (500 mg, 2.97 mmol, 1.0 eq.) in NaOH$_{(aq)}$ (1 M, 3 mL) and the reaction refluxed at 100 °C for 5 h. The reaction was cooled to room temperature and acidified to pH 4 with 1 M HCl$_{(aq)}$ and purified on Dowex® as described in **Section 6.1.3**, to afford the title compound as a brown oil (149 mg, 29%). The starting material L-Orn was observed by NMR as a 13% impurity; selected data confirming synthesis of the title compound: ^{1}H NMR δ (400 MHz, D$_2$O) 1.55-1.76 (2H, m, RNHCH$_2$C*H$_2$*CH$_2$), 1.84-2.01 (2H, m, RNHCH$_2$CH$_2$C*H$_2$*), 3.18 (2H, t, *J* = 6.9 Hz, RNHC*H$_2$*CH$_2$CH$_2$), 4.01 (1H, t, *J* = 6.3 Hz, CO$_2$C*H*); LRMS *m/z* (ESI$^+$) 175 [M+H]$^+$. HMBC data proving guanylation has been regioselective at the δ-NH$_2$ of Orn:



<u>*N*α-(9-Fluorenyl)methoxycarbonyl-*N*δ-tert-Butoxycarbonyl-L-ornithine methyl ester,[213] 41</u>

***N*-Boc-Orn(Fmoc)-OMe**

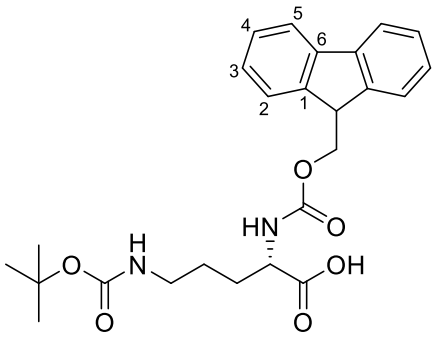


To a solution of Orn(Fmoc)-OMe **86** (100 mg, 0.247 mmol, 1 eq.) in MeOH (1 mL) was added di-*tert*-butyl dicarbonate (70 mg, 0.321 mmol, 1.3 eq.) in acetone (1 mL) and $NaHCO_3$ (63 mg, 0.750 mmol, 3 eq.). The reaction was stirred for 16 h at room temperature and concentrated under *vacuum* to afford the title compound as a colourless gel (19 mg, 16% yield) after purification on SiO_2 (1:9 MeOH:CH_2Cl_2); ^{1}H NMR δ (400 MHz, CD_3OD) 1.44 (9H, s, *t*-Bu)

1.47-1.71 (3H, m, $NHCH_2CH_2CH_2(H_A)$) 1.78-1.89 (1H, m, $NHCH_2CH_2CH_2(H_B)$) 3.02-3.09 (2H, m, $NHCH_2CH_2CH_2$) 3.71 (3H, s, OCH_3) 4.15-4.25 (2H, m, ArC*H*, CO_2CH) 4.33-4.39 (2H, m, $ArCH_2$) 7.31 (2H, dd, J = 6.9, 7.3 Hz, C(3)*H*) 7.39 (2H, dd, J = 7.3, 7.5, C(4)*H*) 7.67 (2H, dd, J = 6.4, 6.6 Hz, C(5)*H*) 7.80 (2H, d, J = 7.5 Hz, C(2)*H*); LRMS *m/z* (ESI^+) 409 $[M+Na]^+$.

N^{α}-(9-Fluorenyl)methoxycarbonyl-N^{δ}-*tert*-Butoxycarbonyl-L-ornithine[146], 81

***N*-Boc-Orn(Fmoc)-OH**

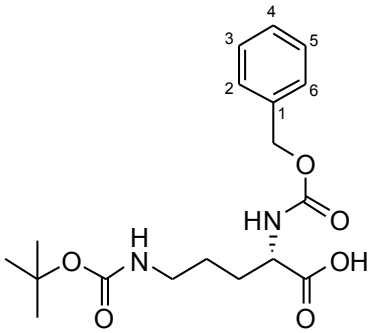


Reaction mixture 1 was prepared as follows: [*N*-Boc-Orn-OH]$_2$Cu **77** (23.2 g, 44.3 mmol, 1.0 eq.) was suspended in acetone (100 mL) and stirred for ~15 min before addition of water (100 mL). The resulting mixture was stirred for an additional 10 min before sequential addition of 10% aqueous Na_2CO_3 (200 mL) and 8-quinolinol (16.7 g, 115 mmol, 2.6 eq.) and the reaction stirred for 1 h. Reaction mixture 2 was prepared as follows: $Na_2CO_{3(s)}$ (4.95 g, 46.7 mmol, 1.1 eq.) and Fmoc-Cl (23.0 g, 88.9 mmol, 2.0 eq.) in acetone (100 mL) were added sequentially to a solution of *N*-hydroxysuccinimide (10.7 g, 93.0 mmol, 2.1 eq.) in water (50 mL). The reaction was stirred for 30 min, following which it was poured into reaction mixture 1. The resulting reaction was stirred for 1 h after which the copper-quinolate precipitate was filtered and washed with water. The filtrate and washings were concentrated under *vacuum*. The residual aqueous solution was poured into a separating funnel and extracted with toluene (3x30 mL, discarded).

The aqueous phase was extracted with EtOAc (3x30 mL). The combined organic layers were washed with 0.5 M $HCl_{(aq)}$ (1x30 mL), 0.25 M $HCl_{(aq)}$ (2x30 mL) and brine (1x30 mL), dried and concentrated under *vacuum* to give a foam. The foam was dissolved in EtOAc and the resulting solution refluxed at 90 °C for 15 min before cooling to RT to afford the title compound as a white solid (23.5 g, 58%) after recrystallization from hexane-EtOAc; mp (from EtOAc-hexane) 120.0-124.5 ˚C (lit.[146] (from EtOAc-hexane) 112.5 ˚C); $[\alpha]_D^{22}$ +7.0 (*c* = 1.0 in MeOH) (lit.[146] $[\alpha]_D^{20}$ −7.95 (*c* = 1.0 in DMF)); ^{1}H NMR δ (400 MHz, DMSO-*d*$_6$) 1.37 (9H, s, *t*-Bu), 1.40-1.77 (4H, m, $NHCH_2C\mathit{H}_2C\mathit{H}_2$), 2.85-2.97 (2H, m, $NHC\mathit{H}_2$), 3.86-3.96 (1H, m, $CO_2HC\mathit{H}$), 4.18-4.32 (3H, m, $ArC\mathit{H}C\mathit{H}_2$), 7.33 (2H, td, *J* = 7.5, 1.3 Hz, 2xC(3)*H*), 7.39-7.44 (2H, m, 2xC(4)*H*), 7.65 (2H, d, *J* = 8.1 Hz, amide-N*H*), 7.73 (2H, d, *J* = 7.3, 2xC(5)*H*), 7.89 (2H, dd, *J* = 7.5, 0.6 Hz, 2xC(2)*H*); LRMS *m/z* (ESI^+) 477 $[M+Na]^+$.

<u>*N*$^{\alpha}$-(Benzyloxy)carbonyl- , *N*$^{\delta}$-*tert*-butoxycarbonyl L-ornithine, 82</u>

***N*-Boc-Orn(Cbz)-OH**

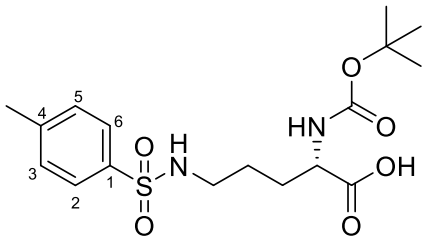


$Cu(CH_3CO_2)_2{\cdot}H_2O$ (9.98 g, 50.0 mmol, 0.5 eq.) in H_2O (50 mL) was added to a solution of L-Orn·HCl (16.9 g, 100 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (2 M, 100 mL), and the reaction mixture stirred for 10 min. A solution of di-*tert*-butyl dicarbonate (28.7 g, 132 mmol, 1.3 eq.) in acetone (200 mL) was added and the reaction stirred at room temperature for 24 h. A pale blue suspension began to form and additional acetone (100 mL) was added and the reaction left stirring for a

further 24 h. The pale blue chalky precipitate was filtered under *vacuum*, washed with acetone:water (2:1, 200 mL) and re-suspended in acetone (90 mL) and a 10% aqueous solution of Na_2CO_3 (10%, 180 mL) with vigorous stirring. 8-Quinolinol (13.1 g, 90.3 mmol, 0.9 eq.) was added and stirred for 1.5 h affording a green suspension. In a separate flask, benzyl chloroformate (12.9 mL, 90.4 mmol, 0.9 eq.) in acetone (45 mL) was added to *N*-hydroxysuccinimide (10.4 g, 90.4 mmol, 0.9 eq.) in water (45 mL) and the reaction stirred for 30 min at 0 °C. The contents of the two flasks were mixed and the reaction was stirred for 2 h at room temperature. The green precipitate was removed by filtration under *vacuum* and washed with H_2O (3x45 mL). Acetone was evaporated from the filtrate and the remaining pH 7-8 aqueous phase extracted with CH_2Cl_2 (3x50 mL, discarded) to remove residual 8-quinolinol. The aqueous phase was adjusted to pH 3 with $HCl_{(aq)}$ and extracted with EtOAc (3x50 mL). The organic layers were combined and dried to afford the title compound as a yellow gel (30.1 g, 82%) without the need for further purification; $[\alpha]_D^{25}$ –2.7 (*c* = 1.0 in MeOH) (lit. $[\alpha]_D^{25}$ 4.0-–2.0 (*c* = 1 in MeOH)); ^{1}H NMR δ (400 MHz, CD_3OD) 1.43 (9H, s, *t*-Bu), 1.50-[214]1.73 (3H, m, $NHCH_2CH_2CH_2(H_A)$), 1.80-1.92 (1H, m, $NHCH_2CH_2CH_2(H_B)$), 3.05 (2H, dd, *J* = 6.7, 6.9 Hz, $NHCH_2CH_2CH_2$), 4.13-4.18 (1H, m, CO_2CH), 5.09 (2H, s, CH_2-Ar), 7.26-7.39 (5H, m, Ar); ^{1}H NMR (ref[215]) δ (400 MHz, $CDCl_3$) 1.44 (9H, s, *t*-Bu), 1.49-1.76 (3H, m, $NHCH_2CH_2CH_2(H_A)$), 1.83-1.97 (1H, m, $NHCH_2CH_2CH_2(H_B)$), 3.03-3.24 (2H, m, $NHCH_2CH_2CH_2$), 4.35-4.47 (1H, m, CO_2CH), 5.11 (2H, s, CH_2-Ar), 7.29-7.38 (5H, m, Ar); LRMS *m/z* (ESI^-) 365.

N^{α}-(*tert*-butoxycarbonyl)-, N^{δ}-*p*-toluenesulfonyl L-ornithine, 83

***N*-Ts-Orn(Boc)-OH**

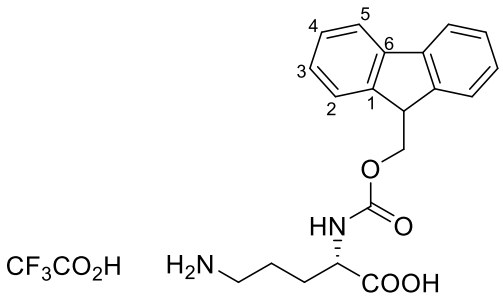


A solution of $Cu(CH_3CO_2)_2{\cdot}H_2O$ (9.98 g, 50.0 mmol, 0.5 eq.) in water (50 mL) was added to a solution of L-Orn·HCl (16.9 g, 100 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (2 M, 100 mL), and the resulting mixture stirred for 10 min. *p*-Toluenesulfonyl chloride (24.8 g, 130 mmol, 1.3 eq.) in acetone (200 mL) was added followed by Et_3N (18.1 mL, 130 mmol, 1.3 eq.) and the reaction stirred at room temperature for 24 h. A pale blue suspension began to form and additional acetone (100 mL) was added and the reaction left stirring for a further 48 h. The pale blue chalky precipitate was filtered under *vacuum*, washed with acetone:water (2:1, 200 mL) and re-suspended in acetone (90 mL) and $NaOH_{(aq)}$ (2 M, 100 mL) with vigorous stirring. 8-Quinolinol (13.1 g, 90.3 mmol, 0.9 eq.) was added and the resulting reaction stirred for 1.5 h affording a green suspension, following which di-*tert*-butyl dicarbonate (28.7 g, 132 mmol, 1.3 eq.) in acetone (45 mL) was added. The reaction was stirred for 72 h, the green precipitate removed by filtration under *vacuum* and washed with H_2O (3x45 mL). Acetone was evaporated from the filtrate under *vacuum* and the remaining pH 7-8 aqueous phase extracted with CH_2Cl_2 (3x50 mL, discarded) to remove 8-quinolinol. The aqueous phase was adjusted to pH 3 with $HCl_{(aq)}$ (1 M) and extracted with EtOAc (3x50 mL). The organic layers were combined and dried to afford the title compound as a brown gel (24.6 g, 64%) without the need for further purification, which slowly solidifies to a pale brown solid over a period of days; mp (from CH_2Cl_2) 77.1-101.9 °C; $[\alpha]_D^{22}$ +3.7 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 3500-3000 (broad), 2976 (broad), 1687 (medium), 1595

(medium), 1159 (sharp); ^{1}H NMR δ (400 MHz, CD_3OD) 1.42 (9H, s, *t*-Bu), 1.47-1.66 (3H, m, $ArNHCH_2CH_2CH_2(H_A)$), 1.68-1.84 (1H, m, $ArNHCH_2CH_2CH_2(H_B)$), 2.42 (3H, s, Ar-CH_3), 2.81-2.86 (2H, m, $ArNHCH_2$), 3.86-3.94 (1H, m, CO_2CH), 7.37 (2H, d, J = 8.1 Hz, C(3)*H*, C5)*H*), 7.72 (2H, d, J = 8.4 Hz, C(2)*H*, C(6)*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 20.1 (C, Ar-CH_3), 25.5 ($ArNHCH_2CH_2CH_2$), 27.4 (3x*t*-BuC(CH_3)), 30.2 ($ArNHCH_2CH_2CH_2$), 42.5 ($ArNHCH_2$), 55.2 (CO_2CH), 78.6 (*t*-Bu$C(CH_3)$), 126.7 (*C*(2), *C*(6)), 129.3 (*C*(3), *C*(5)), 137.5 (*C*(1)), 143.1 (*C*(4)), 156.2 (*t*-BuO*C*=O), 178.0 (CO_2H); LRMS *m/z* (ESI^-) 385 $[M-H]^-$; HRMS (ESI^-) $C_{17}H_{25}N_2O_6S^-$ $[M-H]^-$ requires 385.1439; found 385.1436.

<u>N^{α}-(9-Fluorenyl)methoxycarbonyl- L-ornithine, 84</u>

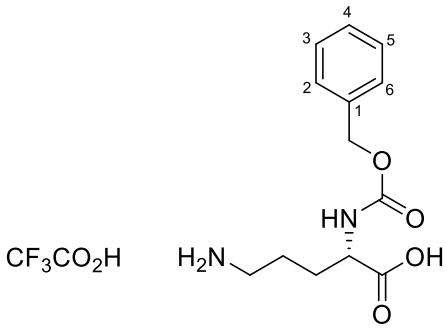
Orn(Fmoc)-OH·CF_3CO_2H



N-Boc-Orn(Fmoc)-OH **81** (2.00 g, 4.40 mmol) was dissolved in TFA:CH_2Cl_2 (3:7, 10 mL) and stirred for 30 min, the crude mixture was then concentrated under *vacuum*. The product was re-crystallised in EtOAc, washed with Et_2O and dried to afford the title compound as an off-white powder (2.20 g, quant.); mp (from Et_2O) 143.9-147 ˚C (lit.[216] (HCl salt, from EtOAc) 97.9 ˚C); $[\alpha]_D^{22}$ +24.2 (*c* = 0.8 in MeOH); ^{1}H NMR δ (400 MHz, DMSO-d_6) 1.54-1.71 (3H, m, $NH_2CH_2CH_2CH_2(H_A)$), 1.73-1.85 (1H, m, $NH_2CH_2CH_2CH_2(H_B)$), 2.72-2.83 (2H, m, NH_2CH_2), 3.92-4.01 (1H, m, CO_2HCH), 4.20-4.36 (3H, m, $ArCHCH_2$), 7.32-7.37 (2H, m, 2xC(3)*H*) 7.42 (2H, td, J = 7.5, 0.8 Hz, 2xC(4)*H*), 7.67-7.79 (3H, m, 2xC(5)*H*, amide-N*H*), 7.91 (2H, d, J = 7.5 Hz, 2xC(2)*H*); LRMS *m/z* (ESI^+) 355 $[M+H]^+$.

<u>N^{α}-(Benzyloxy)carbonyl- L-ornithine, 85</u>

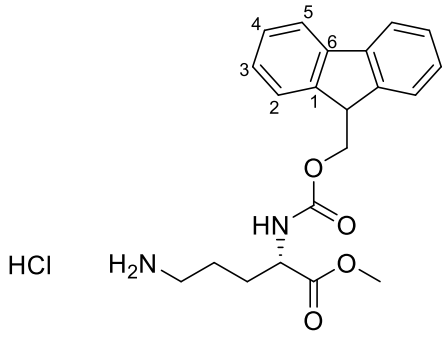
Orn(Cbz)-OH



N-Boc-Orn(Cbz)-OH **82** (1.31 g, 3.58 mmol) was dissolved in TFA:CH_2Cl_2 (3:7, 10 mL) and stirred for 30 min. The solvent was evaporated under *vacuum* to afford the title compound as a pale brown oil (1.14 g, 99%) without the need for further purification; $[\alpha]_D^{25}$ −2.7 (*c* = 1.0 in MeOH); ^{1}H NMR δ (400 MHz, CD_3OD) 1.71-1.82 (3H, m, $NH_2CH_2CH_2CH_2(H_A)$), 1.89-2.03 (1H, m, $NH_2CH_2CH_2CH_2(H_B)$), 2.91-2.98 (2H, m, $NH_2CH_2CH_2CH_2$), 4.18-4.23 (1H, m, CO_2CH), 5.10 (2H, s, CH_2-Ar), 7.29-7.38 (5H, m, Ar); LRMS *m/z* (ESI^+) 267 $[M+H]^+$.

<u>N^{α}-(9-Fluorenyl)methoxycarbonyl- L-ornithine methyl ester, 86</u>

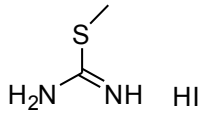
Orn(Fmoc)-OMe·HCl



$SOCl_2$ (839 µL, 11.5 mmol, 1.8 eq.) was added drop-wise to Orn(Fmoc)-OH·CF_3CO_2H **84** (3.04 g , 6.49 mmol, 1.0 eq.) in MeOH (5 mL) and the reaction mixture stirred for 30 min, refluxed for

3 h and stirred for 16 h at room temperature. The solvent was removed under *vacuum*; a yellow solid was obtained, which was triturated with EtOAc and filtered to afford the title compound as a white powder (3.08 g, 89%); mp 153.7-156.5 °C; $[\alpha]_D^{22}$ −3.5 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 3320 (broad), 2952 (broad), 1692 (sharp), 1531 (sharp); ^{1}H NMR δ (400 MHz, CD_3OD) 1.67-2.00 (4H, m, $NH_2CH_2C\mathit{H_2}C\mathit{H_2}$), 2.92-2.98 (2H, m, $NH_2C\mathit{H_2}$), 3.74 (3H, s, $CO_2\mathit{CH_3}$), 4.19-4.28 (2H, m, NHC*H*, Ar*CH*), 4.31-4.38 (1H, m, ArCHC$\mathit{H_2}(\mathit{H_A})$), 4.41-4.49 (1H, m, ArCH$\mathit{CH_2}(\mathit{H_B})$), 7.33 (2H, app dd, *J* = 7.5, 7.3 Hz 2xC(3)*H*), 7.41 (2H, app dd, *J* = 7.5, 7.3 Hz, 2xC(4)*H*), 7.68 (2H, t, *J* = 7.3 Hz, 2xC(5)*H*), 7.82 (2H, d, *J* = 7.3 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 23.7 ($NH_2CH_2CH_2\mathit{C}H_2$), 28.1 ($NH_2CH_2\mathit{C}H_2CH_2$), 38.8 ($NH_2\mathit{C}H_2CH_2CH_2$), 47.0 (NH*C*H) 51.5 (O*C*H_3), 53.4 (Ar*C*H), 66.6 (ArCH*C*H_2), 119.6 (2x*C*(2)), 124.8 (2x*C*(5)), 126.8 (2x*C*(3)), 127.4 (2x*C*(4)), 140.1 (2x*C*(6)), 143.8 (2x*C*(1)), 157.4 (carbamate *C*=O), 172.3 (carboxylate *C*=O); LRMS *m/z* (ESI^+) 369 $[M+H]^+$; HRMS (ESI^+) $C_{21}H_{25}N_2O_4^+$ $[M+H]^+$ requires 369.1809; found 369.1808.

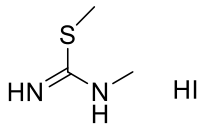
2-Methylthiouronium iodide[217], 93



Iodomethane (2.8 mL, 45.0 mmol, 1.4 eq.) was added drop-wise to a solution of thiourea (2.47 g, 32.5 mmol 1.0 eq.) in EtOH (25 mL) at 0 °C. The reaction was brought to room temperature and then heated to reflux at 90 °C for 90 min after which it was cooled to room temperature. The solvent was removed under *vacuum* affording the title compound as a yellow-brown salt (6.99 g, quant.) without the need for further purification; mp (from EtOH) 116.8-118.8 °C (lit.[217] (from

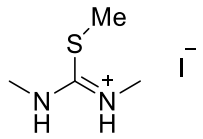
Et_2O) 115.3-117.6 °C); ^{1}H NMR δ (400 MHz, DMSO-d_6) 2.56 (3H, s, S-CH_3), 8.90 (4H, br. s, NH_2CNH_2); HRMS (EI) $C_2H_6N_2S$ [M]$^{\cdot}$ requires 90.0252; found 90.0254.

1,2-Dimethylisothiouronium iodide[218], 94



Iodomethane (953 μL, 15.3 mmol, 1.4 eq.) was added drop-wise to a solution of *N*-methylthiourea (1.00 g, 11.1 mmol 1.0 eq.) in EtOH (15 mL) at 0 °C. The reaction was brought to room temperature and then heated to reflux at 90 °C for 90 min after which it was cooled to room temperature. The solvent was removed under *vacuum* affording the title compound as a yellow-brown salt (2.54 g, 99%) without the need for further purification; mp (from EtOH) 134.0-135.4 °C (lit.[219] (from acetone) 135 °C); ^{1}H NMR δ (400 MHz, D_2O) 2.58 (3H, s, S-CH_3), 2.99 (3H, s, NHCH_3); LRMS *m/z* (ESI$^+$) 105 [M+H]$^+$.

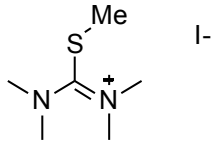
1,2,3-Trimethylthiouronium iodide, 96



Iodomethane (1.3 mL, 21.2 mmol, 1.1 eq.) was added dropwise to a solution of *N,N'*-dimethylthiourea (2 g, 19.2 mmol, 1 eq.) in acetone (80 mL) and the reaction heated to reflux for 40 min after which it was cooled to room temperature. The solvent was removed under *vacuum* affording the title compound as an off white salt (4.5 g, quant. yield) without the need for purification; mp 205.9-210.8 °C; NMR δ_H ppm (400 MHz, DMSO-d_6) 2.63 (3H, s, S-CH_3)

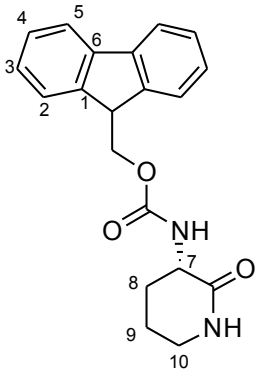
2.94 (6H, dd, *J* = 15.8, 4.5 Hz, 2xN-C*H*$_3$) 8.71 (1H, br.s., N*H* or =N*H*) 9.04 (1H, br.s., N*H* or =N*H*); LRMS *m/*z (ESI$^+$) 119 $[M+H]^+$.

1,1,2,3,3-Pentamethylthiouronium iodide, 98



Iodomethane (1 mL, 16.6 mmol, 1.1 eq.) was added dropwise to a solution of tetramethylthiourea (2 g, 15.1 mmol, 1 eq.) in acetone (80 mL) and the reaction heated to reflux for 40 min after which it was cooled to room temperature. The solvent was removed under *vacuum* affording the title compound as a bright yellow salt (quant. yield) without the need for purification; mp 179.8-180.4 °C lit[220] (180-185 °C (dec)); NMR δ_H ppm (400 MHz, DMSO-d_6) 2.08 (1H, s, N*H*) 2.55 (3H, s, SC*H*$_3$) 3.21 (12H, s, =NC*H*$_3$C*H*$_3$, NC*H*$_3$C*H*$_3$); LRMS *m/z* (ESI$^+$) 147 $[M+H]^+$.

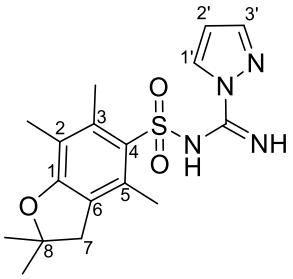
(9H-Fluoren-9-yl)methyl (2-oxopiperidin-3-yl)carbamate[221], 99



DIPEA (96 µL, 0.550 mmol, 2.2 eq.) was added to a solution of Orn(Fmoc)-OMe·HCl **86** (100 mg, 0.250 mmol, 1.0 eq.) in MeOH (2 mL) and the reaction heated to reflux at 90 °C for 1 h. The solvent was removed under *vacuum* to afford the title compound as a white gel (20 mg, 11%)

after purification on SiO_2 (80:20 EtOAc:pet. ether); $[\alpha]_D^{25}$ −5.7 (*c* = 0.9 in MeOH) (lit.[221] not provided); IR ν_{max} cm^{-1} 2948 (broad), 2360 (broad), 2161 (broad), 2031 (broad), 1708 (broad), 1651 (broad); ^{1}H NMR δ (400 MHz, DMSO-d_6) 1.55-1.88 (3H, m, C(8)H_A, C(9)H_2), 1.88-2.07 (1H, m, C(8)H_B), 3.12 (2H, s, C(10)H_2), 3.81-3.99 (1H, m, C(7)*H*), 4.16-4.39 (3H, m, ArC*H*CH_2), 7.34 (2H, t, *J* = 7.2 Hz, 2xC(4)*H* or C(5)*H*), 7.43 (2H, t, *J* = 7.3 Hz, 2xC(4)*H* or C(5)*H*), 7.49 (1H, d, *J* = 8.6 Hz, Fmoc-N*H*), 7.57 (1H, s, lactam N*H*), 7.74 (2H, d, *J* = 7.1 Hz, 2xC(3)*H*), 7.90 (2H, d, *J* = 7.6 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, DMSO-d_6) 21.7 (C7)), 28.3 (*C*(8)), 41.6 (*C*(10)), 47.1 (Ar*C*H)), 51.1 (*C*(7)), 66.0 (*C*(Ar*C*H_2)), 120.6 (2x*C*(2)), 125.8 (2x*C*(3)), 127.5 (2x*C*(5)), 128.1 (2x*C*(4)), 141.2 (2x*C*(6)), 144.4 (2x*C*(1)), 156.6 (carbamate *C*=O)), 170.3 (lactam *C*=O)); LRMS *m/z* (ESI^+) 359 $[M+Na]^+$; HRMS $(ESI)^+$ $C_{20}H_{20}N_2NaO_3^+$ requires 359.1366, found 359.1368.

<u>*N*-((2,2,4,6,7-Pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)-*1H*-pyrazole-1-carboximidamide, 107</u>

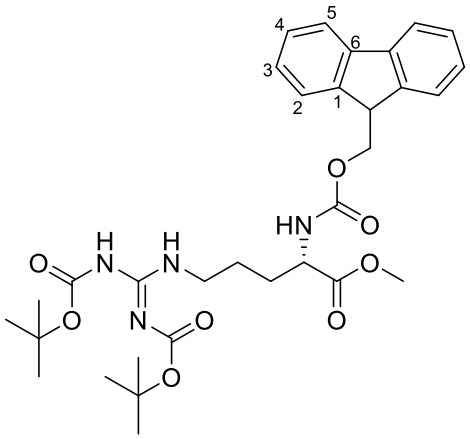


DIPEA (5.9 mL, 33.9 mmol, 5.0 eq.) was added to a solution of 1*H*-Pyrazole-1-carboxamidine (1.00 g, 6.82 mmol, 1.0 eq.) and Pbf chloride (751 mg, 9.00 mmol, 1.3 eq.) in DMF (3 mL), and the reaction heated to 100 °C for 2 h with MW irradiation. The solvent was removed under *vacuum* and the residue suspended in water (20 mL) and extracted with EtOAc (3x20 mL). The combined organic layers were washed with water (4x50 mL), dried and concentrated under *vacuum* to afford the title compound as a white solid (511 mg, 21%) after purification on SiO_2 (pet. ether:EtOAc 90:10 to 80:20); mp (from EtOAc) 182.4-183.3 °C; IR ν_{max} cm^{-1} 3460 (weak,

sharp), 3352 (weak, sharp), 1722 (moderate, sharp), 1554 (moderate, sharp), 1149 (strong, sharp); ^{1}H NMR δ (400 MHz, $CDCl_3$) 1.48 (6H, s, C(8)(CH_3)$_2$), 2.12 (3H, s, C(2)CH_3), 2.57 (3H, s, C(5)CH_3), 2.63 (3H, s, C(3)CH_3), 2.99 (2H, s, C(7)H_2), 6.41 (1H, dd, J = 1.7, 2.9 Hz, C(2')H), 7.68 (1H, dd, J = 0.7, 1.7 Hz, C(3')H), 8.22 (1H, dd, J = 0.7, 2.8 Hz, C(1')H); ^{13}C NMR δ (100 MHz, $CDCl_3$) 12.4 (C(2)CH$_3$), 17.9 (C(3)CH$_3$), 19.2 (C(5)CH$_3$), 28.6 (C(8)(CH$_3$)$_2$), 43.1 (C(7)), 86.8 (C(8)), 109.6 (C(2')), 117.9 (C(3)), 124.9 (C(6)), 129.0 (C(1')), 131.1 (C(4)), 133.1 (C(5)), 139.3 (C(2)), 143.6 (C(3')), 148.8 (RHNC=NH), 159.6 (C(1)); HRMS (FI) $C_{17}H_{22}N_4O_3$ $[M]^{\cdot}$ requires 362.1410; found 362.1413.

N^{α}-(9-Fluorenyl)methoxycarbonyl- N^{ω},N^{ω}-*bis*(*tert*-butoxycarbonyl) L-arginine methyl ester, 111

***N,N'*-Di-Boc-Arg(Fmoc)-OMe**

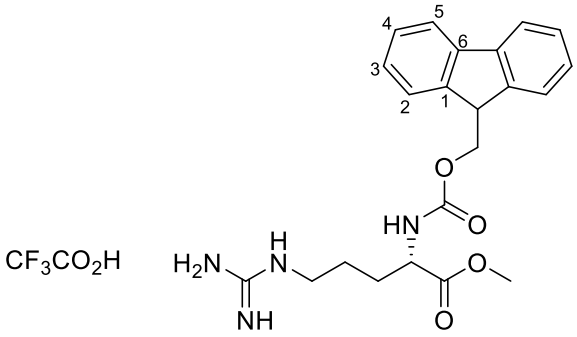


N-N'-di-Boc-1*H*-pyrazole-1-carboxamidine **109** (279 mg, 0.90 mmol, 1.2 eq.) and DIPEA (312 µL, 1.79 mmol, 2.4 eq.) were added sequentially to Orn(Fmoc)-OMe·HCl **86**(300 mg, 0.740 mmol, 1.0 eq.) in MeOH (5 mL) and the reaction stirred for 16 h. The solvent was removed under *vacuum* to afford a white solid (crude, 175 mg) and a small sample taken for analysis. Attempts to purify the crude mixture by silica column chromatography (20:80 EtOAc:Pet. ether) or alumina column chromatography had limited success – degradation to the unprotected guanidine **112** and the cyclised δ-lactam **99** was observed. This compound was therefore used

immediately in the next step as a crude mixture. Selected data supporting synthesis of the title compound is presented below: ^{1}H NMR δ (400 MHz, DMSO-d_6) 1.38 (9H, s, *t*-Bu), 1.47 (9H, s, *t*-Bu), 1.34-1.74 (4H, m, $NH_2CH_2CH_2CH_2$), 3.24-3.32 (2H, s, NH_2CH_2), 3.62 (3H, s, CO_2CH_3), 3.99-4.08 (1H, m, NHC*H*), 4.17-4.33 (3H, m, ArC*H*CH_2), 7.33 (2H, t, *J* = 7.4 Hz, C(4)*H* or C(5)*H*), 7.41 (2H, t, *J* = 7.4 Hz, C(4)*H* or C(5)*H*), 7.71 (2H, d, *J* = 7.2 Hz, C(3)*H*), 7.83 (1H, d, *J* = 7.9 Hz, CH_2CH_2NH), 7.89 (2H, d, *J* = 7.5 Hz, C(2)*H*); LRMS *m/z* (ESI$^+$) 611 $[M+H]^+$; HRMS (ESI$^+$) $C_{32}H_{43}N_4O_8^+$ $[M+H]^+$ requires 611.3075; found 611.3071.

<u>N^{α}-(9-Fluorenyl)methoxycarbonyl- L-arginine methyl ester, 112</u>

Arg(Fmoc)-OMe·CF_3CO_2H

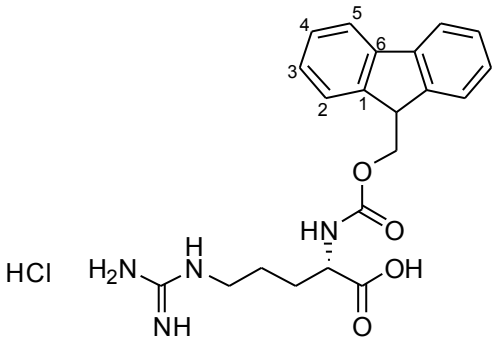


N,N'-Di-Boc-Arg(Fmoc)-OMe **111**(72 mg, 0.118 mmol) was added to TFA:CH_2Cl_2 (3:7, 5 mL) and the reaction stirred for 2 h at room temperature. The solvent was removed under *vacuum* to afford the title compound as pale yellow salt (46 mg, 73%) without the need for purification; mp 156.8-160.2 °C; $[\alpha]_D^{25}$ –4.3 (*c* = 0.6 in MeOH); IR ν_{max} cm^{-1} 3351 (broad), 3162 (broad), 2933 (broad), 1676 (strong), 1433 (moderate), 1365 (weak), 1203 (strong), 1176 (strong), 1135 (sharp, strong); ^{1}H NMR δ (400 MHz,CD_3OD) 1.59-1.75 (3H, m, $NHCH_2CH_2CH_2(H_A)$), 1.85-1.94 (1H, m, $NHCH_2CH_2CH_2(H_B)$), 3.17-3.21 (2H, m, $NHCH_2CH_2CH_2$), 3.72 (3H, s, OCH_3), 4.19-4.24 (2H, m, NHC*H*, Ar*CH*), 4.33-4.43 (ArCHCH_2), 7.29-7.33 (2H, m, 2xC(3)*H*), 7.37-7.41 (2H, m, 2xC(4)*H*), 7.66 (2H, dd, *J* = 7.5, 7.6 Hz, 2xC(2)*H*), 7.79 (2H, d, *J* = 7.6 Hz, 2xC(5)*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 25.0 ($NHCH_2CH_2CH_2$), 28.3 ($NHCH_2CH_2CH_2$), 40.4 ($NHCH_2CH_2CH_2$),

47.0 (Ar*C*H), 51.4 (O*C*H$_3$), 53.5 (NH*C*H), 66.5 (ArCH*C*H$_2$), 119.6 (2x*C*(5)), 124.8 (d, *J* = 4.8 Hz, 2x*C*(1)), 126.8 (2x*C*(3)), 127.4 (2x*C*(4)), 141.2 (2x*C*(6)), 143.8 (d, *J* = 19.1 Hz 2x*C*(1)), 157.3 (d, *J* = 7.6 Hz, carbamate *C*=O), 161.7 (q, *J* = 34.3 Hz, guanidine *C*), 172.8 (carboxylate *C*=O); LRMS *m/z* (ESI$^+$) 411 [M+H]$^+$; HRMS (ESI$^+$) $C_{22}H_{27}N_4O_4^+$ [M+H]$^+$ requires 411.2027; found 411.2037.

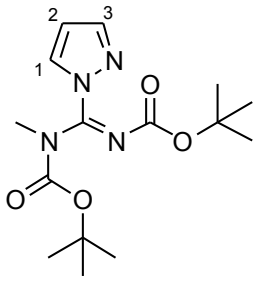
<u>*N*$^{\alpha}$-(9-Fluorenyl)methoxycarbonyl- L-arginine hydrochloride[222], 113</u>

Arg(Fmoc)-OH·HCl



L-Arg(Fmoc)-OMe·HCl **112** (92 mg, 0.206 mmol) was added to 0.5 M HCl$_{(aq)}$ (10 mL) and heated to reflux at 100 °C for 16 h. The reaction was cooled and the water removed under *vacuum* to yield the title compound as a white salt (92 mg, quant.); mp (from 0.5 M HCl$_{(aq)}$) 160.2-162.0 °C (lit.[222] (from CH_2Cl_2:*n*-hexane) 120-124 °C); $[\alpha]_D^{22}$ +10.9 (*c* = 0.9 in DMF) (lit.[222] $[\alpha]_D^{20}$ +8.0 (in DMF)); ^{1}H NMR δ (400 MHz, DMSO-*d*$_6$) 1.45-1.67 (3H, m, NHCH$_2$C*H*$_2$CH(*H*$_A$)), 1.70-1.82 (1H, m, NHCH$_2$C*H*$_2$CH(*H*$_B$)), 3.06-3.13 (2H, m, NHC*H*$_2$), 3.95 (1H, td, *J* = 8.6, 4.8 Hz, NHC*H*), 4.18-4.33 (3H, m, ArC*H*C*H*$_2$), 7.33 (2H, td, *J* = 7.5, 1.0 Hz, 2xC(3)*H*), 7.42 (2H, app. t, *J* = 7.3 Hz, 2xC(4)*H*), 7.67-7.77 (3H, m, 2xC(5)*H*, CH$_2$N*H*), 7.90 (2H, d, *J* = 7.6 Hz, 2xArC(2)*H*); LRMS *m/z* (ESI$^+$) 397 [M+H]$^+$.

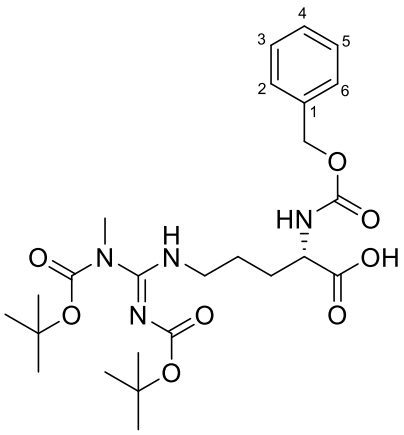
(Z)-*tert*-Butyl(((*tert*-butoxycarbonyl)imino)(1H-pyrazol-1yl)methyl)(methyl)carbamate[223], 114



NaH (60% in oil, 48 mg, 1.20 mmol, 1.2 eq.) was added to a solution of *N-N'*-di-boc-1*H*-pyrazole-1*H*-carboxamidine **109** (310 mg, 1.00 mmol, 1.0 eq.) in DMF (5 mL). The reaction was stirred for 5 min before drop-wise addition of iodomethane (68 µL, 1.09 mmol, 1.1 eq.) in DMF (2 mL),and was stirred for 48 h at room temperature. Water was then added and the aqueous layer extracted with EtOAc (3x30 mL), the combined organic layers were dried and concentrated under *vacuum* to afford the title compound as a colourless oil (252 mg, 77%) after purification on SiO_2 (10:90 EtOAc:pet. ether); IR ν_{max} cm^{-1} 2979 (broad), 1730 (broad), 1688 (broad), 1139 (strong); ^{1}H NMR δ (400 MHz, $CDCl_3$) 1.23 (9H, s, *t*-Bu), 1.46 (9H, s, *t*-Bu), 3.18 (3H, s, NCH_3), 6.36 (1H, dd, *J* = 2.7, 1.6 Hz, C(2)*H*), 7.63 (1H, dd, *J* = 1.6, 0.6 Hz, C(3)*H*), 7.93 (1H, br.s., C(1)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 27.8 (*t*-Bu(3x*C*H_3)), 28.0 (*t*-Bu(3x*C*H_3), 35.7 (N*C*H_3), 82.5 (*t*-Bu(*C*CH_3)), 82.8 (*t*-Bu(*C*CH_3)), 109.2 (*C*(2)), 129.9 (*C*(3), *C*(1)), 143.3(N=*C*-N-CH_3), 152.5 (C=N*C*=O), 157.5 (C-N(CH_3)*C*=O); LRMS *m/z* (ESI^+) 325 $[M+H]^+$; HRMS (ESI^+) $C_{15}H_{24}N_4NaO_4^+$ $[M+Na]^+$ requires 347.1690; found 347.1696.

N^{α}-(Benzyloxy)carbonyl- N^{ω} -methyl, N^{ω}, N^{ω}-*bis*(*tert*-butoxycarbonyl) L-arginine, 115

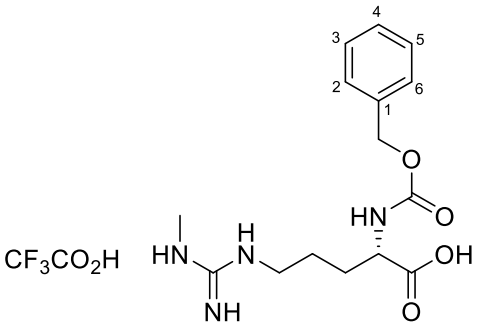
***N*-Me(ω)-, *N,N'*-di-Boc-Arg(Cbz)-OH**



Pyrazole **109** (101 mg, 0.311 mmol, 1.0 eq.) and DIPEA (111 μL, 0.637 mmol, 2.0 eq.) were added sequentially to a solution of Orn(Cbz)-OH **85** (86 mg, 0.323 mmol, 1.0 eq.) in MeOH (2 mL), and the reaction stirred for 16 h at room temperature. The mixture was concentrated under *vacuum* to afford the title compound as white crystals (112 mg, 57%) after purification on SiO_2 (CH_2Cl_2:MeOH 96:4 to 88:12); mp (from MeOH) 122.4-126.5 °C; $[\alpha]_D^{25}$ +4.4 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 1026 (strong), 1701 (medium), 2360 (broad, weak), 2977 (broad, weak); ^{1}H NMR δ (400 MHz, CD_3OD) 1.46 (9H, s, 2x*t*-Bu), 1.59-1.72 (3H, m, $NHCH_2CH_2CH_2(H_A)$), 1.82-1.92 (1H, m, $NHCH_2CH_2CH_2(H_B)$), 3.01 (3H, s, $NHCH_3$), 3.18-3.27 (2H, $NHCH_2CH_2CH_2$), 4.04-4.10 (1H, CO_2CH), 5.06 (2H, td, *J* = 9.2, 12.4, $ArCH_2$), 7.26-7.37 (5H, m, phenyl); ^{13}C NMR δ (100 MHz, CD_3OD) 24.8 ($NHCH_2CH_2CH_2$), 27.2 ($2x(CH_3)_3$), 29.7 ($NHCH_2CH_2CH_2$), 34.1 (NCH_3), 42.4 ($NHCH_2CH_2CH_2$), 55.5 (br. s, CO_2CH), 66.3 ($ArCH_2$), 79.2 ($C(CH_3)_3$), 81.9 ($C(CH_3)_3$.), 127.6 (*C*(3), *C*(5)), 127.7 (*C*(4)), 128.1 (*C*(2), *C*(6)), 136.8 (*C*(1)), 153.2 (*C*=NH), 157.1 ($ArCH_2$), 179.0 (CO_2H); LRMS *m/z* (ESI^+) 523 $[M+H]^+$; HRMS (ESI^+) $C_{25}H_{39}N_4O_8^+$ $[M+H]^+$ requires 523.2762; found 523.2761.

N^{α}-(Benzyloxy)carbonyl- N^{ω} –methyl L-arginine, 116

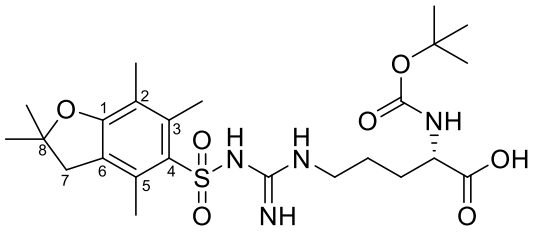
***N*-Me(ω)-Arg(Cbz)-OH**



Nω -Me-, *N,N'*-di-Boc-Arg(Cbz)-OH (112 mg, 0.214 mmol) was dissolved in TFA:CH_2Cl_2 (3:7, 10 mL) and stirred for 5 h. The solvent was evaporated under *vacuum* to afford the title compound as a pale brown oil (94 mg, quant.) without the need for further purification; $[\alpha]_D^{25}$ – 4.4 (*c* = 1.0 in MeOH, 0.15% DIPEA); IR ν_{max} cm^{-1} 3231 (weak, broad), 1635 (strong), 1184 (strong); ^{1}H NMR δ (400 MHz, CD_3OD) 1.62-1.78 (3H, m, $NMeCH_2C\mathit{H_2}CH_2(\mathit{H_A})$), 1.85-1.97 (1H, m, $NMeCH_2CH_2CH_2(\mathit{H_B})$), 2.82 (3H, s, $NC\mathit{H_3}$), 3.16-3.23 (2H, m, $NMeC\mathit{H_2}CH_2CH_2$), 4.16-4.22 (1H, m, $CO_2C\mathit{H}$), 5.09 (2H, s, $ArC\mathit{H_2}$); ^{13}C NMR δ (100 MHz, CD_3OD) 25.0 ($NMeCH_2\mathit{C}H_2CH_2$), 26.7 ($N\mathit{C}H_3$), 28.5 ($NMeCH_2CH_2\mathit{C}H_2$), 40.5 ($NMe\mathit{C}H_2CH_2CH_2$), 53.4 ($CO_2\mathit{C}H$), 66.3 ($Ar\mathit{C}H_2$), 127.4 (*C*(3), *C*(4)), 127.6 (*C*(4)), 128.1 (*C*(2), *C*(6)), 136.7 (*C*(1)), 157.1 (d, *J* = 58.7 Hz, *C*=NH), 160.6 (q, *J* = 36.0 Hz, NH*C*=O), 174.1 (CO_2H); LRMS *m/z* (ESI^+) 323 $[M+H]^+$; HRMS (ESI^+) $C_{15}H_{23}N_4O_4^+$ $[M+H]^+$ requires 323.1714; found 323.1714.

N^{α}-(*tert*-Butoxycarbonyl)-N^{ω}-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl) L-arginine[224], 117

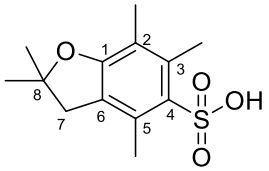
***N*-Pbf-Orn(Boc)-OH**



A solution of Pbf-Cl (1.02 g, 3.53 mmol, 2.2 eq.) in acetone (3.2 mL) was added drop-wise to a ice-cold solution of L-Arg(Boc)·HCl (500 mg, 1.61 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (1 M, 1.6 mL), and the reaction was stirred at 0 °C for 2 h before being brought to room temperature for 16 h. A white precipitate was removed by filtration, which was characterised as the side-product Pbf-OH **118**. The filtrate was extracted with Et_2O to remove unreacted Pbf-Cl and the residual aqueous phase was slowly acidified to pH 6-7 with $HCl_{(aq)}$. The resulting precipitate was isolated by *vacuum* filtration and washed with H_2O and evaporated under high *vacuum* to afford the title compound as a white solid (166 mg, 32%) without the need for further purification; mp (from H_2O) 150.6-156.6 °C (lit.[225] 110-120 °C (solvent unspecified)); $[\alpha]_D^{25}$ +16.1 (c = 1.0 in MeOH) (lit.[225] $[\alpha]_D^{20}$ +4.0 (*c* = 4 in MeOH)); ^{1}H NMR δ (400 MHz, CD_3OD) 1.43 (9H, s, *t*-Bu), 1.45 (6H, s, C(8)(CH_3)$_2$), 1.51-1.70 (3H, m, RNHCH$_2$CH_2CH$_2$(H_A)), 1.74-1.85 (1H, m, RNHCH$_2$CH$_2$CH$_2$(H_B), 2.08 (3H, s, C(2)CH_3), 2.51 (3H, s, C(5)CH_3), 2.57 (3H, s, C(3)CH_3), 3.00 (2H, s, C(7)H_2), 3.14-3.21 (2H, m, RNHCH_2CH$_2$CH$_2$), 4.00-4.05 (1H, m, CO$_2$C*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 11.2 (C(2)*C*H$_3$), 17.0 (C(3)*C*H$_3$), 18.2 (C(5)*C*H$_3$), 27.4 (*t*-BuC(*C*H$_3$)$_3$, C(8)(*C*H$_3$)$_2$), 28.8 (RNHCH$_2$CH$_2$*C*H$_2$), 40.2 (RNH*C*H$_2$CH$_2$CH$_2$), 42.6 (*C*(7)), 53.4 (CO$_2$*C*H), 79.2 (t-Bu*C*CH$_3$), 86.4 (*C*(8)), 117.1 (*C*(3)), 124.7 (*C*(6)), 132.1 (*C*(4)), 132.8 (*C*(5)), 138.0 (*C*(2)), 156.8 (*C*(1)), 158.5 (*C*O$_2$); LRMS *m/z* (ESI$^+$) 527 [M+H]$^+$.

2,2,4,6,7-Pentamethyl-2,3-dihydrobenzofuran-5-sulfonic acid, 118

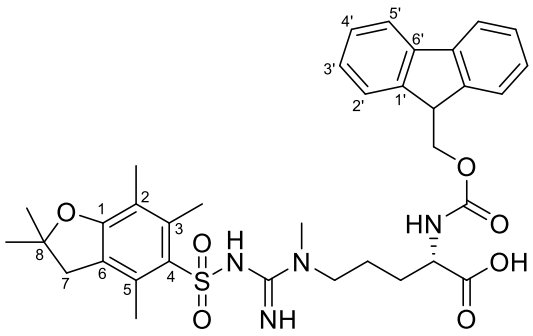
Pbf-OH



The title compound **118** was isolated as a white solid (104 mg, 11%); mp (from acetone) decomp. past 210 °C; IR ν_{max} cm^{-1} 3418 (broad), 2988-2929 (weak), 2360-2341 (weak), 1209 (sharp, strong), 1180 (sharp, strong), 1010 (strong); ^{1}H NMR δ (400 MHz, CD_3OD) 1.43 (6H, s, C(8)(CH_3)$_2$), 2.07 (3H, s, C(5)H_3), 2.49 (3H, s, C(3)H_3), 2.56 (3H, s, C(2)H_3), 2.97 (2H, s, C(7)H_2); ^{13}C NMR δ (100 MHz, CD_3OD) 11.1 (C(5)CH$_3$), 16.9 (C(2)CH$_3$), 18.2 (C(3)CH$_3$), 27.3 (C(6)(CH$_3$)$_2$), 42.7 (C(7)), 85.7 (C(8)), 116.2 (C(3)), 123.9 (C(6)), 131.1 (C(4)), 134.7 (C(5)), 136.8 (C(2)), 157.3 (C(1)); LRMS *m/z* (ESI^-) 269 $[M-H]^-$; HRMS (ESI^-) $C_{13}H_{17}O_4S^-$ $[M-H]^-$ requires 269.0853; found 269.0859.

N^{α}-(9-Fluorenyl)methoxycarbonyl-N^{ω}-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)-N^{δ} -methyl-L-arginine, 138

***N*-Pbf-, *N*-Me(δ)-Arg(Fmoc)-OH**

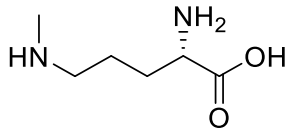


10% Pd/C(50 mg, ~20% w/w) was added to a solution of *N*-Pbf-, *N*-Me-Arg(Cbz)-OH **179** (236 mg, 0.411 mmol, 1.0 eq.) in MeOH (6 mL). The reaction was sealed, degassed and back-flushed

with H_2 (three cycles). The reaction was stirred for 5 h, filtered through Celite®, eluted with MeOH and concentrated under *vacuum*. The crude residue was dissolved in 10% aqueous Na_2CO_3 (1 mL). In a separate flask, a solution of Fmoc-Cl (116 mg, 0.448 mmol, 1.1 eq.) in acetone (1 mL) was added to 10% aqueous Na_2CO_3 (4 mL), followed by *N*-hydroxysuccinimide (58 mg, 0.504 mmol, 1.2 eq.), and the reaction stirred for 30 min. The two solutions were combined and the resulting reaction stirred for 16 h at room temperature after which the solvent was removed under *vacuum*. The residue was dissolved in EtOAc and washed with 1 M $HCl_{(aq)}$ (3x30 mL). The organic layers were combined and dried to afford the title compound as white crystals (145 mg, 57%) after purification on SiO_2 (CH_2Cl_2:MeOH 95:5 to 90:10 to 85:15); mp (from MeOH) 165.3-169.9 °C; $[\alpha]_D^{25}$+5.3 (*c* = 1.0 in MeOH:CH_2Cl_2 55:45); IR ν_{max} cm^{-1} 2971 (weak, broad), 1553 (strong); ^{1}H NMR δ (400 MHz, CD_3OD) 1.39 (6H, s, C(8)H_3), 1.50-1.64 (3H, m, NMeCH$_2$CH_2CH$_2$(H_A)), 1.71-1.82 (1H, m, NMeCH$_2$CH$_2$CH$_2$(H_B)), 2.05 (3H, s, C(2)H_3), 2.50 (3H, s, C(5)H_3) 2.58 (3H, C(3)H_3), 2.88 (3H, s, NCH_3), 2.92 (2H, s, C(7)H_2), 3.96-4.02 (1H, m, CO_2C*H*), 4.16 (1H, t, *J* = 6.4, ArC*H*), 4.26-4.33 (1H, ArCH$_2$(H_A)), 4.36-4.44 (1H, ArCH$_2$(H_B)), 7.22-7.28 (2H, m, C(3')*H*), 7.31-7.37 (2H, m, C(4')*H*), 7.57-7.64 (2H, m, C(5')*H*), 7.72-7.79 (2H, m, C(2')*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 11.1 (C(2)CH_3), 17.1(C(3)CH_3), 18.2 (C(5)CH_3), 23.5 (NMeCH$_2$CH$_2$CH$_2$), 27.3 (C(8)(CH_3)$_2$), 29.5 (NMeCH$_2$CH$_2$CH$_2$), 34.2 (NCH_3), 42.6 (C(7)CH_2), 47.2 (ArCH), 48.9 (NMeCH$_2$CH$_2$CH$_2$), 56.1 (CO_2CH), 66.3 (ArCH_2), 86.2 (*C*(8)), 117.2 (*C*(3)), 119.5 (*C*(2')), 124.6 (*C*(6)), 124.8 (*C*(5')), 126.7 (*C*(3')), 127.4 (*C*(4')), 131.9 (*C*(4)), 133.2 (*C*(5)), 137.9 (*C*(2)), 141.2 (*C*(6')), 144.0 (*C*(1')), 155.9 (NH=*C*), 157.0 (Ar*C*=O), 158.4 (*C*(1)); LRMS *m/z* (ESI$^+$) 663 $[M+H]^+$; HRMS (ESI$^+$) $C_{35}H_{42}N_4NaO_7S^+$ $[M+Na]^+$ requires 685.2666; found 685.2662.

N^{δ}-Methyl L-ornithine[170], 149

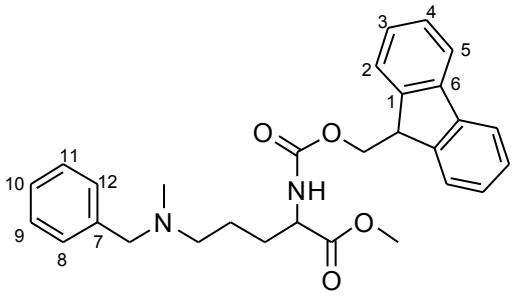
***N*-Me(δ)-Orn-OH**



N-Ts-*N*-Me-Orn(Boc)-OH **180** (232 mg, 0.579 mmol) was dissolved in 48% $HBr_{(aq)}$ (2 mL) and refluxed at 122 °C for 2 h. The reaction was cooled to room temperature and solid contaminants removed by gravity filtration. The filtrate was purified *via* ion-exchange chromatography by adapting the general procedure described in **Section 6.1.2**. In this instance, the acidic crude mixture was stirred for 30 min with Dowex® resin, prepared according to the general procedure, before being added to a column. The flow-through was collected and stirred with a second batch of Dowex® for 5 min before application to a second column. This process was repeated for a second time and the three resulting columns washed and eluted in tandem, requiring $NH_4OH_{(aq)}$ (9 M) as eluent. The product fractions were pooled to afford the title compound as a brown oil (64 mg, 76%); $[\alpha]_D^{22}$ +19.8 (*c* = 1.4 in 6 M $HCl_{(aq)}$) (lit.[163] $[\alpha]_D^{20}$ +23.3 (*c* = 2 in 6 M $HCl_{(aq)}$); ^{1}H NMR δ (400 MHz, D_2O) 1.64-1.90 (4H, m, $MeNHCH_2C\mathit{H_2}C\mathit{H_2}$), 2.65 (3H, s, $C\mathit{H_3}$), 3.02 (2H, dd, *J* = 7.8, 7.2, $MeNHC\mathit{H_2}CH_2CH_2$), 3.71 (1H, dd, *J* = 6.1, 5.3 $CO_2C\mathit{H}$); LRMS *m/z* (ESI^+) 147 $[M+H]^+$.

Methyl 2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-(benzyl(methyl)amino)pentanoate, 156

***N*-Methyl, *N*-Benzyl-Orn(Fmoc)-OMe·HCl**

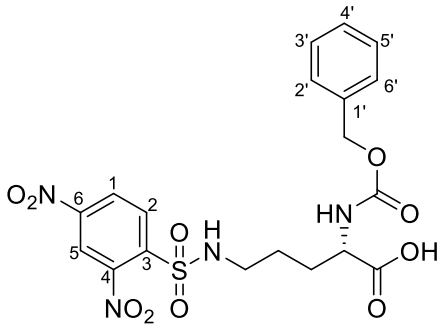


Benzaldehyde (75.5 µL, 0.74 mmol, 1.2 eq.) was added to a solution of Orn(Fmoc)-OMe·HCl **86** (250 mg, 0.62 mmol, 1 eq.) in MeOH (2 mL) and 3Å molecular sieves (crushed). The flask was sealed and purged with argon and stirred for 1 h before addition of $NaCNBH_3$ (46.5 mg, 0.74 mmol, 1.2 eq.). The reaction was stirred for 30 min and then 37% formaldehyde (46 µL, 1.24 mmol, 2 eq.) was added followed immediately by $NaCNBH_3$ (54.7 mg, 0.87 mmol, 1.4 eq.). The reaction was stirred for 30 min and subsequently quenched by addition of a few drops of $HCl_{(aq)}$ (1 M) and filtered through Celite®, using MeOH and CH_2Cl_2 as eluent. The solvent was removed under *vacuum* to afford the title compound as a partially pure off-white gel (158 mg) containing putative dibenzylated species **157** after attempted purification on SiO_2 (EtOAc:Pet. ether, 2:8 to 8:2); selected data: IR ν_{max} cm^{-1} 2917 (weak, broad), 2330 (weak, broad), 1733 (weak, sharp), 1449 (medium, sharp), 732 (strong, sharp); ^1H NMR δ (400 MHz, CD_3OD) 1.52-1.84 (4H, m, (Bn)(Me)$NCH_2CH_2CH_2$), 2.14 (3H, s, NCH_3), 2.37 (2H, t, J = 6.7 Hz, (Bn)(Me)NCH_2), 3.46 (2H, s, NCH_2Ph), 3.70 (3H, s, CO_2CH_3), 4.13-4.23 (2H, m, NHC*H*, Ar*CH*), 4.37 (2H, m, $ArCHCH_2$), 7.20-7.43 (9H, m, NCH_2Ph, (Ar)H_4H_5), 7.65 (2H, t, J = 7.0 Hz,(Ar)H_3), 7.77 (2H, d, J = 7.3 Hz, (Ar)H_2); ^{13}C NMR δ (100 MHz, CD_3OD) 23.0 ($BnNMeCH_2CH_2$), 32.5 ($BnNMeCH_2CH_2CH_2$), 41.3 ($PhCH_2NMeCH_2$), 51.5 ($NHCO_2CH_2$), 53.9 ($NHCO_2CH_2CAr$), 56.9 ($MeCO_2CHNH$, NCH_3), 62.0 (OCH_3) 107.3 (2x*C*(2)), 119.8 (2x*C*(3)), 121.1 (2x*C*(5)), 127.2 (2x*C*(4)), 127.4 (2x*C*(1)), 128.3 (*C*(10)), 128.9 (*C*(9), *C*(11)), 129.6 (*C*(8),

C(12)), 138.2 (*C*(7)), 140.4 (2x*C*(6)), 143.9 (NH*C*O_2), 176.1 (*C*O_2Me); LRMS *m/z* (ESI^+) 473; HRMS (ESI^+) $C_{29}H_{33}N_2O_4^+$ $[M+H]^+$ requires 473.2435; found 473.2414.

<u>N^{α}-(Benzyloxy)carbonyl-, N^{δ}-2,4-dinitrobenzylsulfonyl L-ornithine, 175</u>

***N*-DNs-Orn(Cbz)-OH**

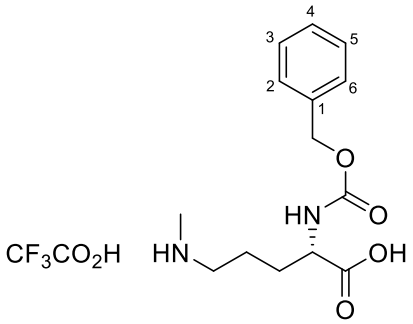


2,4-Dinitrobenzenesulfonyl chloride (4.8 g, 18.0 mmol, 1.1 eq.) was dissolved in $NaOH_{(aq)}$ (1 M, 15 mL), slowly to avoid precipitation and added to a solution of Orn(Cbz)-OH·CF_3CO_2H **85** (6.22 g, 16.4 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (1 M, 10 mL), and the resulting mixture stirred at room temperature for 72 h. The mixture was washed with Et_2O (2x30 mL, discarded), the aqueous phase cooled to 0 °C and acidified to pH 2 with vigorous stirring. The aqueous phase was extracted with EtOAc (3x30 mL) and the organic layers combined and dried to afford a pale yellow solid (71% title compound, 29% starting amino acid). Attempts to purify the compound resulted in degradation (notably on SiO_2) and so it was used for subsequent reactions without purification. Selected data: mp (from EtOAc) 106.0-113.7 °C; IR ν_{max} cm^{-1} 3321 (broad), 1696 (sharp, medium), 1541 (sharp, strong), 1243 (medium); $[\alpha]_D^{25}$ −6.2 (*c* = 1.0 in MeOH); ^{1}H NMR δ (400 MHz, CD_3OD) 1.71-1.83 (3H, m, RNHCH$_2$C*H*$_2$CH$_2$(*H*$_A$)), 1.90-1.98 (1H, m, RNHCH$_2$CH$_2$CH$_2$(*H*$_B$), 2.92-2.99 (2H, m, RNHC*H*$_2$CH$_2$CH$_2$), 4.17-4.22 (1H, m, CO_2C*H*), 5.10 (2H, s, C*H*$_2$Ar), 7.27-7.38 (5H, m, benzyl), 8.27 (1H, d, *J* = 8.7 Hz, C(2)*H*), 8.45-8.49 (1H, m, C(1)*H*), 8.52 (1H, d, *J* = 2.2, C(5)*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 23.8 (NHCH$_2$*C*H$_2$CH$_2$),

28.3 ($NHCH_2CH_2\mathit{C}H_2$), 38.9 ($NH\mathit{C}H_2CH_2CH_2$), 53.3 ($NH\mathit{C}H$), 66.3 ($Ar\mathit{C}H_2$), 118.6 (*C*(5)), 125.5 (*C*(1)), 127.4 (*C*(2'), *C*(6')), 127.7 (*C*(4')), 128.1 (*C*(3'), *C*(5')), 130.9 (*C*(2)), 142.7 (*C*(6)), 157.3 (carbamate *C*=O). Carbons 3, 4 and 1' were not seen. HMBC data suggests a coupling to the 136-137 ppm region which could correspond to *C*(1') and a coupling to the 147-149 region which could correspond to *C*(3); LRMS *m/z* (ESI^+) 519 $[M+Na]^+$; HRMS could not be found.

<u>N^{α}-(Benzyloxy)carbonyl-N^{δ}-methyl-L-ornithine, 177</u>

***N*-Me(δ)-Orn(Cbz)-OH**

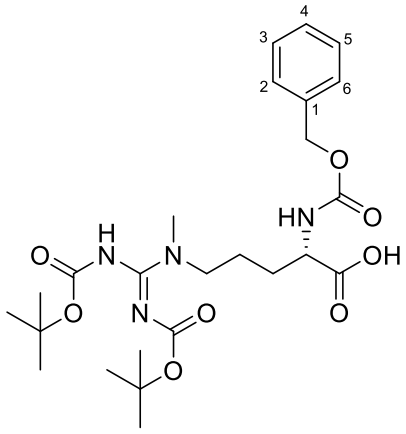


N-Me, *N*-Boc-Orn(Cbz)-OH **188** (990 mg, 2.60 mmol) was dissolved in TFA:CH_2Cl_2 (3:7, 10 mL) and stirred for 30 min. The solvent was evaporated under *vacuum* to afford the title compound as a pale brown oil (1.01 g, quant.) without the need for further purification; $[\alpha]_D^{25}$ -5.0 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 3035 (broad, weak), 1671 (strong), 1198 (strong); ^1H NMR δ (400 MHz, CD_3OD) 1.70-1.83 (3H, m, $NHCH_2C\mathit{H}_2CH_2(\mathit{H}_A)$), 1.96 (1H, m, $NHCH_2CH_2CH_2(\mathit{H}_B)$, 2.67 (3H, s, $NC\mathit{H}_3$), 2.97-3.04 (2H, m, $NHC\mathit{H}_2CH_2\mathit{C}H_2$), 4.18-4.24 (1H, m, $CO_2C\mathit{H}$), 5.10 (2H, s, $ArC\mathit{H}_2$), 7.24-7.40 (5H, m, phenyl); ^{13}C NMR δ (100 MHz, CD_3OD) 22.3 ($NHCH_2\mathit{C}H_2CH_2$), 28.3 ($NHCH_2CH_2\mathit{C}H_2$), 32.1 ($N\mathit{C}H_3$), 48.4 ($NH\mathit{C}H_2CH_2CH_2$), 53.2 ($CO_2\mathit{C}H$), 66.3 (Ar-$\mathit{C}H_2$), 127.5 (*C*(3), *C*(5)), 127.7 (*C*(4)), 128.1 (*C*(2), *C*(6)), 136.9 (*C*(1)),

159.3; LRMS *m/z* (ESI^+) 281 $[M+H]^+$; HRMS $(ESI)^+$ $C_{14}H_{21}N_2O_4^+$ requires 281.1496, found 281.1497.

<u>N^{α}-(Benzyloxy)carbonyl- N^{δ} -methyl, N^{ω}, N^{ω}-*bis*(*tert*-butoxycarbonyl)-L-arginine, 178</u>

***N*-Me(δ)-, *N,N'*-Di-Boc-Arg(Cbz)-OH**

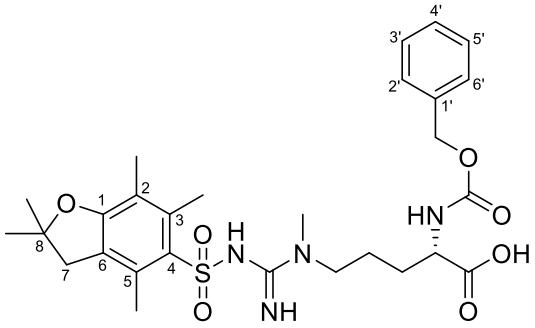


N-Me-Orn(Cbz)-OH, **177 (**999 mg, 2.53 mmol, 1.0 eq.) was dissolved in MeOH (5 mL) and *N,N'-bis*-boc-pyrazole-1*H*-carboxamidine (865 mg, 2.79 mmol, 1.1 eq.) added. The reaction was stirred for 24 h and then concentrated under *vacuum* to afford the title compound as a white solid (846 mg, 64%) after purification on SiO_2 (CH_2Cl_2:MeOH:Et_3N 97:2:1); mp (from 91:9 MeOH:DIPEA) 69.0-75.3 °C; $[\alpha]_D^{25}$+1.8 (*c* = 1.0 in MeOH, 1% DIPEA); IR ν_{max} cm^{-1} 2978 (broad, weak), 1601 (moderate), 1147 (strong); ^1H NMR δ (400 MHz, CD_3OD) 1.46 (18H, s, 2x*t*-Bu), 1.58-1.73 (3H, m, $NMeCH_2C\mathit{H_2}CH_2(\mathit{H_A})$), 1.75-1.88 (1H, m, $NMeCH_2CH_2CH_2(\mathit{H_B})$), 2.95 (3H, s, $NC\mathit{H_3}$), 3.38-3.46 (2H, m, $NMeC\mathit{H_2}CH_2CH_2$), 4.03-4.09 (1H, m, $CO_2C\mathit{H}$), 5.02-5.11 (2H, m, $ArC\mathit{H_2}$), 7.25-7.38 (5H, m, phenyl); ^{13}C NMR δ (100 MHz, CD_3OD) 23.2 ($NMeCH_2\mathit{C}H_2CH_2$), 27.2 (2x(*t*-$BuC(\mathit{C}H_3)_3$)), 29.4 ($NMeCH_2CH_2\mathit{C}H_2$), 35.3 ($H\mathit{C}H_3$), 49.9 ($NMe\mathit{C}H_2CH_2CH_2$), 55.2 ($CO_2\mathit{C}H$), 66.1 ($\mathit{C}H_2Ar$), 79.9 (2x(*t*-$Bu\mathit{C}(CH_3)_3$)), 127.5 (*C*(2), *C*(6)), 127.6 (*C*(4)), 128.1 (*C*(3), *C*(5)), 136.9 (*C*(1)), 153.9 (*C*=NH), 156.9 (Ar-*C*=O), 177.1 ($\mathit{C}O_2H$);

LRMS *m/z* (ESI^+) 523 $[M+H]^+$; HRMS (ESI^+) $C_{25}H_{39}N_4O_8^+$ $[M+H]^+$ requires 523.2762; found 523.2763.

N^{α}-(Benzyloxy)carbonyl-N^{ω}-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)-N^{δ}-methyl-L-arginine, 179

***N*-Pbf-, *N*-Me(δ)-Arg(Cbz)-OH**

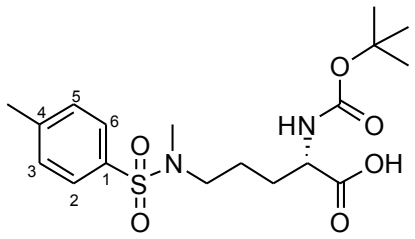


A solution of Pbf-Cl (402 mg, 1.39 mmol, 1.0 eq.) in acetone (9 mL) was added over 10 min to an ice-cold solution of *N*-Me-Arg(Cbz)-OH **188** (620 mg, 1.42 mmol, 1.0 eq.) in acetone (5 mL) and 3 M $NaOH_{(aq)}$ (5 mL), and the reaction stirred for 1 h at 0 °C. A second portion of Pbf-Cl (197 mg, 0.682 mmol, 0.5 eq.) in acetone (4.5 mL) was added and the resulting mixture stirred at 0 °C for 90 min. A third portion of Pbf-Cl (197 mg, 0.682 mmol, 0.5 eq.) in acetone (4.5 mL) was added and the reaction stirred at 0 °C for 30 min, brought to room temperature and stirred for 16 h after which it was filtered and concentrated. The crude mixture was re-suspended in acetone (3 mL) and 3 M $NaOH_{(aq)}$ and a further four portions of Pbf-Cl (each 197 mg, 0.682 mmol, 0.5 eq.) added at 1 h intervals at room temperature. The mixture was filtered, concentrated and purified on SiO_2 (CH_2Cl_2:MeOH:AcOH 100:0:0 to 96:3:1 to 93:6:1). Fractions containing product were pooled, mixed with H_2O and the organics removed under *vacuum*. The residual aqueous solution was acidified to ~pH 2 using 1 M $HCl_{(aq)}$ and extracted with CH_2Cl_2 (3x 30 mL). The organic layers were combined, washed with 10% aqueous citric acid, dried

(Na_2SO_4) and concentrated under *vacuum* to afford the title compound as a pale yellow oil (236 mg, 29%); $[\alpha]_D^{25}$ –1.2 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 3339 (weak, broad), 2930 (weak, broad), 1553 (moderate), 1089 (strong); ^{1}H NMR δ (400 MHz, CD_3OD) 1.44 (6H, s, C(8)(CH_3)$_2$), 1.53-1.65 (3H, m, NMeCH$_2$CH_2CH$_2$(H_A)), 1.71-1.81 (1H, m, NMeCH$_2$CH$_2$CH$_2$(H_B)), 2.07 (3H, C(2)H_3), 2.50 (3H, C(5)H_3), 2.57 (3H, s, C(3)H_3), 2.89 (3H, s, NCH_3), 2.98 (2H, s, C(7)H_2), 3.34-3.39 (2H, m, NMeCH_2CH$_2$CH$_2$), 4.09-4.14 (1H, m, CO_2C*H*), 5.08 (2H, td, *J* = 4.6, 12.4, ArCH_2), 7.26-7.38 (5H, m, phenyl); ^{13}C NMR δ (100 MHz, CD_3OD), 11.2 (*C*(2)), 17.1 (*C*(3)), 18.3 (*C*(5)), 23.6 (NMeCH$_2$*C*H$_2$CH$_2$), 27.4 (C(8)(*C*H$_3$)$_2$), 28.5 (NMeCH$_2$CH$_2$*C*H$_2$), 34.0 (N*C*H$_3$), 42.6 (*C(7)*), 48.6 (NMe*C*H$_2$CH$_2$CH$_2$), 53.6 (CO_2*C*H), 66.3 (Ar*C*H$_2$), 86.3 (*C*(8)), 117.1 (*C*(3)), 124.7 (*C*(6)), 127.4 (*C*(3'), *C*(5')), 127.6 (*C*(4')), 128.1 (*C*(2'), *C*(6')), 131.9 (*C*(4)), 133.2 (*C*(5)), 136.8 (*C*(1')), 137.8 (C(2)), 155.9 (*C*=NH), 157.2 (NH*C*=O), 158.4 (*C*(1)), 174.1 (CO_2H); LRMS *m/z* (ESI^+) 575 $[M+H]^+$; HRMS (ESI^+) $C_{28}H_{39}N_4O_7S^+$ $[M+H]^+$ requires 575.2534; found 575.2535.

<u>*N*$^{\alpha}$-(*tert*-Butoxycarbonyl)- *N*$^{\delta}$-methyl, *N*$^{\delta}$-*p*-toluenesulfonyl L-ornithine, 180</u>

***N*-Ts-*N*-Me(δ)-Orn(Boc)-OH**

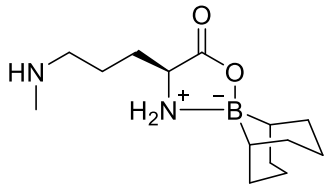


A solution of *N*-Ts-Orn(Boc)-OH **83** (1.00 g, 2.59 mmol, 1.0 eq.) in $NaOH_{(aq)}$ (2 M, 10 mL) was sealed in a microwave vial. Iodomethane (161 μL, 2.59 mmol, 1.0 eq.) was injected to the vial and the reaction heated to 70 °C for 1 h. The reaction was cooled to below 40 °C and a second equivalent of iodomethane (161 μL, 2.59 mmol, 1.0 eq.) added before heating to 70 °C for 1 h.

This cycle was repeated once more with one further equivalent of iodomethane (161 μL, 2.59 mmol, 1.0 eq.) and the reaction cooled to room temperature and then to 0 °C. The mixture was acidified to pH 2 with $HCl_{(aq)}$ (1 M) to precipitate the product as a brown gel, which was extracted from the aqueous phase with CH_2Cl_2 (3x25 mL). The organic phases were combined and dried to afford the title compound as a brown oil (800 mg, 77%) without the need for further purification; $[\alpha]_D^{22}$ –5.6 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 2977 (broad), 1712 (broad), 1338 (weak), 1160 (strong); ^{1}H NMR δ (400 MHz, CD_3OD) 1.44 (9H, m, *t*-Bu)), 1.56-1.72 (3H, m, $ArNHCH_2C\mathit{H}_2CH_2(\mathit{H}_A)$), 1.78-1.91 (1H, m, $ArNHCH_2CH_2CH_2(\mathit{H}_B)$), 2.43 (3H, s, Ar-$C\mathit{H}_3$), 2.68 (3H, s, *N*-$C\mathit{H}_3$), 2.94-3.07 (2H, m $ArNHC\mathit{H}_2CH_2CH_2$), 4.07-4.12 (1H, s, $CO_2C\mathit{H}$), 7.40 (2H, d, *J* = 8.1 Hz, C(3)*H*, C(5)*H*), 7.67 (2H, d, *J* = 8.2 Hz, C(2)*H*, C(6)*H*); ^{13}C NMR δ (100 MHz, CD_3OD) 20.1 (Ts-$C\mathit{H}_3$), 23.7 ($ArNHCH_2C\mathit{H}_2CH_2$), 27.3 (3x*t*-BuC($C\mathit{H}_3$)), 28.4 ($ArNHCH_2CH_2C\mathit{H}_2$), 33.7 (N-$C\mathit{H}_3$), 49.4 ($ArNHC\mathit{H}_2$), 53.1 ($CO_2C\mathit{H}$), 79.1 (*t*-Bu$C$($CH_3$)), 127.2 ($C$(2), C(6)), 129.5 (C(3), C(5)), 134.3 (C(1)), 144.3 (C(4)), 156.7 (*t*-BuOC=O), 174.6 ($C$$O_2$H); LRMS *m/z* ($ESI^+$) 423 $[M+Na]^+$; HRMS (ESI^+) $C_{18}H_{28}N_2NaO_6S^+$ $[M+Na]^+$ requires 423.1560; found 423.1558.

9-Borabicyclo[3.3.1]non-9-yl[*N*-methyl-ornithinato-*O,N*]-boron, 184

***N*-Me-Orn(9-BBN)**

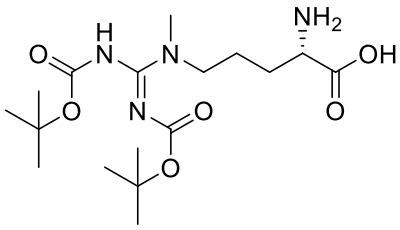


To a solution of *N*-Me-Orn-OH **149** (205 mg, 1.40 mmol, 1.4 eq.) in MeOH (0.75 mL) was added 9-BBN in THF (3.4 mL, 1.70 mmol, 0.5 M) and refluxed for 1.5 h. The resulting suspension was filtered, and the residue washed with *n*-hexane (1x20 mL) and Et_2O (1x20 mL) to give the title compound as a crude brown gel. Selected data of the crude mixture supporting

synthesis of title compound: ^{1}H NMR δ (400 MHz, CD_3OD) 1.57-2.19 (18H, m, 9-BBN, $NHCH_2CH_2CH_2$), 2.93 (3H, s, NCH_3), 3.33-3.40 (2H, m, $NHCH_2$), 3.55-3.59 (1H, m, CO_2CH); LRMS *m/z* (ESI^+) 267 $[M+H]^+$.

<u>*N*δ -methyl, *N*ω, *N*ω-*bis*(*tert*-butoxycarbonyl) L-arginine, 186</u>

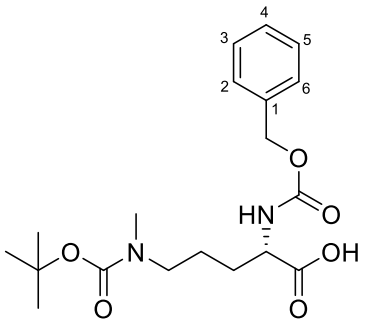
***N*-Me(δ)-, *N*-, *N'-bis*-Boc-Arg-OH**



To a solution of crude *N*-Me-Orn(9-BBN)-OH **184** (127 mg, 0.477 mmol, 1 eq.) in MeOH (1 mL) was added *bis*-Boc-pyrazole-1*H*-carboxamidine **109** (180 mg, 0.580 mmol, 1.2 eq.) and DIPEA (100 μL, 0.574 mmol, 1.2 eq.). The mixture was concentrated, re-dissolved in $NaOH_{(aq)}$ (2 M) and stirred for 1 h. Purification *via* Dowex® resin afforded the title compound as a partially pure brown residue. Selected data supporting synthesis of the title compound: ^{1}H NMR δ (400 MHz, D_2O) 1.37 (18H, s, *t*-Bu), 1.53-1.79 (4H, m, $NH_2CH_2CH_2CH_2$), 2.93 (3H, s, OCH_3), 3.34-3.44 (2H, m, $NH_2CH_2CH_2$), 3.64-3.70 (1H, CO_2CH); LRMS *m/z* (ESI^+) 389 $[M+H]^+$.

N^{α}-(Benzyloxy)carbonyl-, N^{δ}-methyl, N^{δ}-*tert*(butoxycarbonyl)-L-ornithine, 188

***N*-Boc, *N*-Me(δ)-Orn(Cbz)-OH**

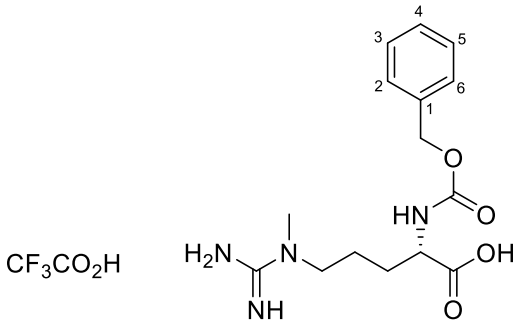


$Cu(CH_3CO_2)_2 \cdot H_2O$ (1.76 g, 8.82 mmol, 0.5 eq) in H_2O (9 mL) was added to a solution of *N*-Me-Orn-OH, **149** (2.58 g, 17.7 mmol, 1.0 eq.) in 2 M $NaOH_{(aq)}$ (17 mL), and the reaction mixture stirred for 10 min. Di-*tert*-butyl dicarbonate (3.9 g, 17.9 mmol, 1.0 eq.) in acetone (34 mL) was added and the reaction stirred at room temperature for 17 h. The solution was filtered under *vacuum*, and the filtrate was treated with 8-quinolinol (2.30 g, 15.8 mmol, 0.9 eq.) and stirred for 2 h after which the copper:quinolinate complex was removed by *vacuum* filtration. The filtrate (pH 10) was extracted with CH_2Cl_2 (3x20 mL, discarded) and the aqueous phase slowly acidified to pH 4 using conc. $HCl_{(aq)}$, extracted with EtOAc (3x20 mL, discarded) and concentrated before being re-dissolved in 10% $Na_2CO_{3(aq)}$ (32 mL). In a separate flask *N*-hydroxysuccinimide (1.80 g, 15.6 mmol, 0.9 eq.) was added to a 10% aqueous solution of Na_2CO_3 (8 mL) and benzyl chloroformate (2.3 mL, 16.1 mmol, 0.9 eq.) added in acetone (8 mL) and the mixture stirred on ice for 30 min. The two solutions were combined and stirred for 24 h after which the mixture was acidified to pH 4 with 1 M $HCl_{(aq)}$ and extracted with EtOAc (3x20 mL). The organic layers were combined, dried and concentrated to afford the title compound as a colourless oil (1.11 g, 16%) after purification on SiO_2 (CH_2Cl_2:MeOH:Et_3N 98:1:1); $[\alpha]_D^{25}$ –2.59 (*c* = 1.0 in MeOH); IR ν_{max} cm^{-1} 2974 (broad, weak), 1692 (broad, strong), 1159 (sharp, strong); ^1H NMR δ (400 MHz, CD_3OD) 1.45 (9H, s, *t*-Bu), 1.58-1.70 (3H, m, $NCH_2CH_2CH_2(H_A)$), 1.82 (1H, m,

$NCH_2CH_2CH_2(H_A)$), 2.82 (3H, s, NCH_3), 3.20-3.27 (2H, m, $NCH_2CH_2CH_2$), 4.16-4.22 (1H, m, CO_2CH), 5.06-5.14 (2H, m, CH_2Ar), 7.26-7.39 (5H, m, phenyl); ^{13}C NMR δ (100 MHz, CD_3OD), 23.6 ($NHCH_2CH_2CH_2$), 27.3 (*t*-Bu$C(CH_3)_3$), 28.5 ($NHCH_2CH_2CH_2$), 33.0 (NCH_3), 53.6 (CO_2CH), 66.2 (Ar-CH_2), 79.5 (*t*-Bu$C(CH_3)_3$), 127.4 (*C*(3), *C*(5)), 127.6 (*C*(4)), 128.1 (*C*(2), *C*(6)), 136.9 (*C*(1)), 156.3 (MeN*C*=O), 157.3 ($ArCH_2C$=O), 174.3 (CO_2); LRMS *m/z* (ESI^-) 379; HRMS (ESI^+) $C_{19}H_{28}N_2O_6$ $[M+H]^+$ requires: 381.2020; found 381.2013.

<u>N^{α}-(benzyloxy)carbonyl- N^{δ} -methyl-L-arginine, 189</u>

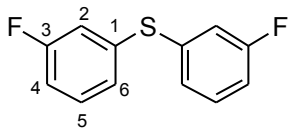
***N*-Me(δ)-Arg(Cbz)-OH**



N-Me-,*N,N'*-di-Boc-Arg(Cbz)-OH, **178** (906 mg, 1.73 mmol) was dissolved in TFA:CH_2Cl_2 (3:7, 20 mL) and stirred for 2 h. The solvent was evaporated under *vacuum* to afford the title compound as a pale brown oil (760 mg, quant.) without the need for further purification; $[\alpha]_D^{25}$ – 4.4 (*c* = 1.0 in MeOH, 0.15% DIPEA); IR ν_{max} cm^{-1} 3370 (weak, broad), 1628 (moderate), 1145 (strong); ^1H NMR δ (400 MHz, CD_3OD) 1.62-1.78 (3H, m, $NMeCH_2CH_2CH_2(H_A)$), 1.84-1.94 (1H, m, $NMeCH_2CH_2CH_2(H_B)$), 2.99 (3H, s, NCH_3), 3.34-3.42 (2H, m, $NMeCH_2CH_2CH_2$), 4.19-4.25 (1H, m, CO_2CH), 5.06-5.15 (2H, m, $ArCH_2$); ^{13}C NMR δ (100 MHz, CD_3OD) 23.0 ($NMeCH_2CH_2CH_2$), 28.3 ($NMeCH_2CH_2CH_2$), 35.1 (NCH_3), 49.5 ($NMeCH_2CH_2CH_2$), 54.5 (CO_2CH), 66.3 ($ArCH_2$), 127.4 (*C*(3), *C*(4)), 127.6 (*C*(4)), 128.1 (*C*(2), *C*(6)), 136.8 (*C*(1)) 156.8

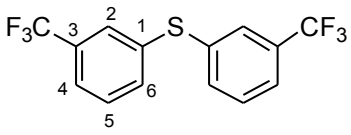
(*C*=NH), 157.4 (NH*C*=O), 173.8 (CO_2H); LRMS *m/z* (ESI^+) 323 $[M+H]^+$; HRMS (ESI^+) $C_{15}H_{23}N_4O_4^+$ $[M+H]^+$ requires 323.1714; found 323.1714.

<u>*S,S-bis*(3-fluorophenyl)sulfide, 206</u>



Following general procedure 1, 3-fluoroiodobenzene (275 μL, 2.34 mmol), 3-fluorothiophenol (198 μL, 2.34 mmol) and ethylene glycol (260 μL, 4.68 mmol) afforded the title compound as a pale yellow oil (432 mg, 83%) after purification on SiO_2 (pet. ether); IR ν_{max} cm^{-1} 1578 (sharp), 1472 (sharp), 1216 (sharp); ^1H NMR δ (400 MHz, $CDCl_3$) 6.92 (2H, tdd, *J* = 8.4, 2.6, 0.9 Hz, 2xC(4)*H*), 6.97-7.01 (2H, m, 2xC(2)*H*), 7.08 (2H, ddd, *J* = 7.8, 1.6, 1.0 Hz, 2xC(6)*H*), 7.20-7.26 (2H, m, 2xC(5)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 114.5 (d, *J* = 21.3 Hz, 2x*C*(4)), 117.9 (d, *J* = 23.5 Hz, 2x*C*(2)), 126.7 (d, *J* = 2.9 Hz, 2x*C*(6)) 130.6 (d, *J* = 8.1 Hz, 2x*C*(5)), 137.2 (d, *J* = 7.3 Hz, 2x*C*(1)), 163.0 (d, *J* = 249.4 Hz, 2x*C*(3)); ^{19}F NMR δ (377 MHz, $CDCl_3$) −111.5 (F, m, 2xC*F*); LRMS not available due to compound being undetectable by ESI; HRMS (EI^+) $C_{12}H_8F_2S^+$ $[M]^+$ requires 222.0315; found 222.0314.

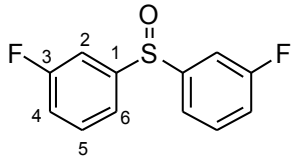
<u>*S,S-Bis*(3-(trifluoromethyl)phenyl)sulfide[226], 207</u>



Following general procedure 2, 3-iodobenzotrifluoride (340 μL, 2.34 mmol), 3-(trifluoromethyl)thiophenol (200 μL, 2.34 mmol) and ethylene glycol (260 μL, 4.68 mmol) afforded the title compound as a pale yellow oil (614 mg, 81%) after purification on SiO_2 (pet. ether); IR ν_{max} cm^{-1} 1318 (sharp), 902 (weak), 792 (strong); ^1H NMR δ (400 MHz, $CDCl_3$) 7.43-

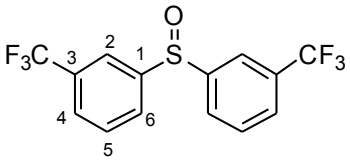
7.53 (4H, m, 2xC(5)*H*, 2xC(6)*H*), 7.55 (2H, d, *J* = 7.0 Hz, 2xC(4)*H*), 7.63 (2H, s, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 123.6 (q, *J* = 272.9 Hz, 2x*C*F_3), 124.4 (q, *J* = 3.7 Hz, 2x*C*(4)), 127.7 (q, *J* = 3.7 Hz, 2x*C*(2)), 129.9 (2x*C*(5)), 131.9 (q, *J* = 33.0 Hz, 2x*C*(3)), 134.2 (2x*C*(6)), 136.2 (2x*C*(1)); ^{19}F NMR δ (377 MHz, $CDCl_3$) −62.9 (F, 2xC*F*$_3$); LRMS not available due to compound being undetectable by ESI; HRMS *m/z* (FI) $[M]^{\cdot}$ $C_{14}H_8F_6S^{\cdot}$ requires 322.0251; found 322.0244.

3,3'-Sulfinyl-*bis*(fluorobenzene), 208



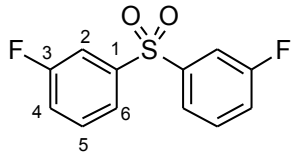
Following general procedure 2, *S,S-bis*(3-fluorophenyl)sulfide **206** (220 mg, 1.00 mmol) afforded the title compound as an off-white crystalline solid (160 mg, 42%) after purification on SiO_2 (10:90 EtOAc:pet. ether); mp (from EtOAc) 72.9-73.7 °C; IR ν_{max} cm^{-1} 1216 (sharp), 1038-1063 (broad); ^{1}H NMR δ (400 MHz, $CDCl_3$) 7.13-7.18 (2H, m, 2xC(4)*H*), 7.33-7.52 (6H, m, 2xC(2)*H*, 2xC(5)*H*, 2xC(6)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 111.7 (d, *J* = 23.8 Hz, 2x*C*(2)), 118.6 (d, *J* = 21.5 Hz, 2x*C*(4)), 120.3 (d, *J* = 3.2 Hz, 2x*C*(6)), 131.2 (d, *J* = 7.9 Hz, 2x*C*(5)), 147.7 (d, *J* = 5.6 Hz, 2x*C*(1)), 163.0 (d, *J* = 252.7 Hz, 2x*C*(3)); ^{19}F NMR δ (377 MHz, $CDCl_3$) −109.0 - −108.9 (F, m, 2xC*F*); LRMS *m/z* (ESI^+) 261 $[M+Na]^+$; HRMS (ESI^+) $C_{12}H_8F_2NaOS^+$ $[M+Na]^+$ requires 261.0156; found 261.0157.

3,3'-Sulfinyl-*bis*((trifluoromethyl)benzene)[227], 209



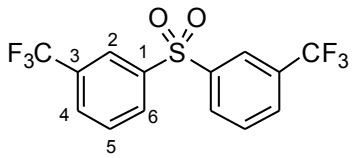
Following general procedure 1, *S,S-bis*(3-(trifluoromethyl)phenyl)sulfide **207** (320 mg, 1.00 mmol) afforded the title compound as a yellow oil (251 mg, 74%) after purification on SiO_2 (20:80 EtOAc:pet. ether); IR ν_{max} cm^{-1} 1057 (strong); ^1H NMR δ (400 MHz, $CDCl_3$) 7.65 (2H, t, *J* = 7.8 Hz, 2xC(5)*H*), 7.75 (2H, d, *J* = 8.0 Hz, 2xC(4)*H*), 7.85 (2H, d, *J* = 7.9 Hz, 2xC(6)*H*), 7.98 (2H, s, 2xC(2)*H*); ^{19}F NMR δ (377 MHz, $CDCl_3$) −62.8 (F, 2xCF_3); LRMS *m/z* (ESI^+) 361 $[M+Na]^+$.

3,3'-Sulfonyl*bis*(fluorobenzene), 210



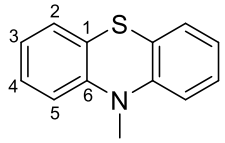
Following general procedure 3, *S,S-bis*(3-fluorophenyl)sulfide **206** (100 mg, 0.450 mmol) afforded the title compound as an off-white powder (77.5 mg, 68%) after purification on SiO_2 (20:80 EtOAc:pet. ether); mp (from EtOAc) 93.4-94.7 °C; IR ν_{max} cm^{-1} 3048 (weak), 1218 (sharp), 1148 (strong); ^1H NMR δ (400 MHz, $CDCl_3$) 7.30 (2H, tdd, *J* = 8.3, 2.5, 0.9 Hz, 2xC(4)*H*), 7.49-7.56 (2H, m, 2xC(5)*H*), 7.65 (2H, ddd, *J* = 7.8, 2.1 Hz, 2xC(2)*H*), 7.75 (2H, d, *J* = 7.8 Hz, 2xC(6)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 115.1 (d, *J* = 24.2 Hz, 2x*C*(2)), 120.9 (d, *J* = 21.3 Hz, 2x*C*(4)), 123.6 (d, *J* = 3.7 Hz, 2x*C*(6)), 131.3 (d, *J* = 7.3 Hz, 2x*C*(5)), 143.1 (d, *J* = 5.9 Hz, 2x*C*(1)), 162.5 (d, *J* = 253.1 Hz, 2x*C*(3)); ^{19}F NMR δ (377 MHz, $CDCl_3$) −108.8 (F, 2xC*F*); LRMS *m/z* (ESI^+) 277 $[M+Na]^+$; HRMS (ESI^+) $C_{12}H_8F_2NaO_2S^+$ $[M+Na]^+$ requires 277.0105; found 277.0102.

3,3'-Sulfonyl*bis*((trifluoromethyl)benzene), 211



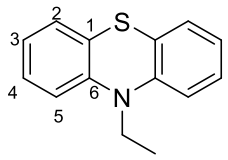
Following general procedure 3, *S,S-bis*(3-(trifluoromethyl)phenyl)sulfide **207** (100 mg, 0.310 mmol) afforded the title compound as a white solid (34 mg, 32%) after purification on SiO_2 (20:80 EtOAc:pet. ether); mp (from EtOAc) 93.2-93.4 °C (lit.[228] (from pet. ether) 87-89 °C); IR ν_{max} cm^{-1} 1325 (sharp), 1130 (strong); ^1H NMR δ (400 MHz, $CDCl_3$) 7.72 (2H, dd, *J* = 7.8, 8.0 Hz, 2xC(5)*H*), 7.88 (2H, d, *J* = 7.8 Hz, 2xC(6)*H*), 8.16 (2H, d, *J* = 7.8 Hz, 2xC(4)*H*), 8.25 (2H, s, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 124.9 (q, *J* = 3.7 Hz, 2x*C*(2)), 130.4 (2x*C*(5), 2x*C*(6)), 131.1 (2x*C*(4)), 142.1 (2x*C*(1)) note: *C*(3) and *C*F_3 were not resolved);^{19}F NMR δ (377 MHz, $CDCl_3$) −62.9 (F, 2xC*F*$_3$); LRMS *m/z* (ESI^+) 377 $[M+Na]^+$; HRMS (ESI^+) $C_{12}H_8F_2NaOS^+$ $[M+Na]^+$ requires 377.0041; found 377.0043.

10-Butyl-10*H*-phenothiazine[229], 213



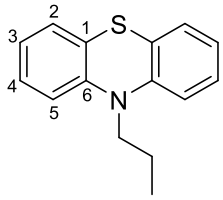
Following general procedure 4b, methyl iodide (312 µl, 5.01 mmol) afforded the title compound as a white solid (725 mg, 68%) after purification on SiO_2 (100% pet. ether); mp 100.5-101.6 °C (from EtOAc) (lit.[229] (from EtOH) 101.7-104.1°C); ^1H NMR δ (400 MHz, C_6D_6) 2.67 (3H, s, NC*H*$_3$), 6.36 (2H, d, *J* = 8.1 Hz, 2xC(5)*H*), 6.70 (2H, td, *J* = 7.5, 1.1 Hz, 2xC(3)*H*), 6.90 (2H, td, *J* = 7.8, 1.5 Hz, 2xC(4)*H*), 7.05 (2H, dd, *J* = 7.6, 1.4 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 35.2 (N*C*H$_3$), 114.7 (2x*C*(5)), 123.0 (2x*C*(3)), 124.4 (2x*C*(1)), 127.8 (2x*C*(2)), 127.8 (2x*C*(4)), 146.5 (2x*C*(6)); LRMS *m/z* (ESI^+) 214 $[M+H]^+$.

10-Ethyl-10*H*-phenothiazine[230], 214



Following general procedure 4a, phenothiazine (1.00 g, 5.03 mmol) and ethyl bromide (370 µl, 4.96 mmol) afforded the title compound as a white powder (740 mg, 65%) after purification on SiO_2 (10:90 EtOAc:pet. ether); mp (from EtOAc) 103.4-104.6 °C (lit.[230] (from MeOH) 103-104 °C); ^{1}H NMR ppm (400 MHz, $CDCl_3$) 1.43 (3H, t, *J* = 7.0 Hz, C*H*$_3$), 3.94 (2H, q, *J* = 7.0 Hz, NC*H*$_2$), 6.82-6.95 (4H, m, 2xC(3)*H*, 2xC(2)*H*), 7.08-7.19 (4H, m, 2xC(4)*H*, 2xC(5)*H*); LRMS *m/z* (ESI$^+$) 228 $[M+H]^+$.

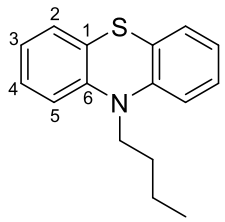
10-propyl-10*H*-phenothiazine[231], 215



Following general procedure 4b, 1-bromopropane (500 µL, 5.50 mmol) afforded the title compound as a white crystalline solid (321 mg, 53%) after purification on SiO_2 (100% pet. ether); mp (from pet. ether) 50.1-50.8 °C (lit.[231] (from EtOH) 48-49 °C); IR ν_{max} cm^{-1} 1455 (moderate), 752 (sharp); ^{1}H NMR δ (400 MHz, C_6D_6) 0.72 (3H, t, *J* = 7.4 Hz, $CH_2CH_2CH_3$), 1.51 (2H, sxt, *J* = 7.4 $CH_2CH_2CH_3$), 3.36 (2H, t, *J* = 7.4 Hz, $CH_2CH_2CH_3$), 6.54 (2H, d, *J* = 8.2 Hz, 2xC(5)*H*), 6.67-6.73 (2H, m, 2xC(3)*H*), 6.89-6.95 (2H, m, 2xC(4)*H*), 7.07 (2H, dd, *J*=7.6, 1.2 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 10.9 ($CH_2CH_2CH_3$), 19.9 ($CH_2CH_2CH_3$), 48.7

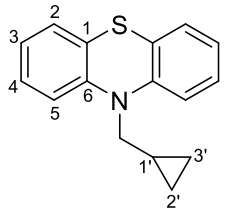
($CH_2CH_2CH_3$), 115.56 (2x*C*(2)), 122.3 (2x*C*(4)), 125.5 (2x*C*(1)), 126.9 (2x*C*(3)), 127.5 (2x*C*(5)), 145.5 (2x*C*(6)); LRMS *m/z* (ESI^+) 242 $[M + H]^+$; HRMS (ESI^+) $C_{15}H_{16}NS^+$ $[M+H]^+$ requires 242.0998; found 242.0999.

10-Butyl-10*H*-phenothiazine[232], 216



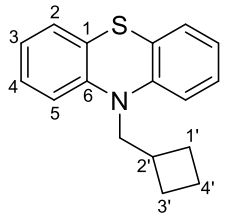
Cs_2CO_3 (815 mg, 2.50 mmol, 1.0 eq.) was added to phenothiazine (500 mg, 2.51 mmol, 1.0 eq.) in DMF (10 mL) before addition of TBAI (925 mg, 2.50 mmol, 1 eq.) and the mixture stirred for 1 h before cooling to 0 °C. 1-Bromobutane (591 µL, 5.50 mmol, 2.2 eq.) was added drop-wise, the reaction brought to room temperature and stirred for 16 h. The reaction was treated with water and EtOAc (3x30 mL). The organic layers were combined, washed with water (2x30 mL) and brine (1x30 mL), dried, filtered and concentrated under *vacuum* to afford the title compound as a colourless oil (140 mg, 10%) after purification on SiO_2 (100% pet. ether; ^1H NMR δ (400 MHz, C_6D_6) 0.70 (3H, t, *J* = 7.5 Hz, $CH_2CH_2CH_2CH_3$), 1.18 (2H, dq, *J* = 15.0, 7.5 Hz, $CH_2CH_2CH_2CH_3$), 1.45-1.54 (1H, m, $CH_2CH_2CH_2CH_3$), 3.42 (1H, t, *J* = 7.0 Hz, $CH_2CH_2CH_2CH_3$), 6.58 (2H, d, *J* = 8.1 Hz, 2xC5*H*), 6.70 (1H, td, *J* = 7.5, 1.2 Hz, 2xC(3)*H*), 6.91-6.96 (1H, m, 2xC(4)*H*), 7.08 (1H, dd, *J* = 7.6, 1.5 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 13.5 ($CH_2CH_2CH_2CH_3$), 19.9 ($CH_2CH_2CH_2CH_3$), 28.8 ($CH_2CH_2CH_2CH_3$), 46.7 ($CH_2CH_2CH_2CH_3$), 115.4 (2x*C*(2)), 122.3 (2x*C*(3)), 125.5 (2x*C*(1)), 126.9 (2x*C*(4)), 127.5 (2x*C*(5)), 145.5 (2x*C*6); LRMS *m/z* (ESI^+) 256 $[M+H]^+$.

10-(Cyclopropylmethyl)-10*H*-phenothiazine, 217



Following general procedure 4a, phenothiazine (1.00 g, 5.03 mmol) and (bromomethyl)cyclopropane (485 μL, 5.00 mmol) afforded the title compound as an off-white solid (247 mg, 20%) after purification on SiO_2 (pet. ether); mp (from pet. ether) 122-126 °C; IR ν_{max} cm^{-1} 3000 (weak), 1456 (moderate), 750 (strong); ^1H NMR δ (400 MHz, acetone-d_6); 0.37-0.42 (2H, m, C(2')H_A and C(3')H_A), 0.57-0.63 (2H, m, C(2')H_B and C(3')H_B), 1.19-1.30 (1H, m, C(1')*H*), 3.80 (2H, d, *J* = 5.8 Hz, NCH_2), 6.91-6.96 (2H, m, 2xC(3)*H*), 7.10 (2H, dd, *J* = 8.2, 1.2 Hz, 2xC(5)*H*), 7.14 (2H, dd, *J* = 7.7, 1.4 Hz, 2xC(2)*H*), 7.17-7.22 (2H, m, 2xC(4)*H*); ^{13}C NMR δ (100 MHz, acetone-d_6) 5.4 (*C*(2'), *C*(3')), 10.0 (*C*(1')), 52.3 (N*C*H_2), 116.3 (2x*C*(5)), 122.8 (2x*C*(3)), 124.8 (2x*C*(1)), 127.4 (2x*C*(2)), 127.8 (2x*C*(4)), 146.0 (2x*C*(6)); LRMS not available due to compound being undetectable by ESI; HRMS (FI) $C_{16}H_{15}NS$ $[M]^{\cdot}$ requires 253.0925; found 253.0926.

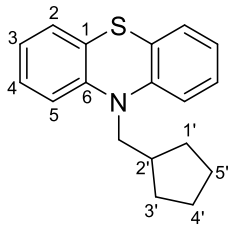
10-(Cyclobutylmethyl)-10*H*-phenothiazine, 218



Following general procedure 4b, bromomethylcyclobutane (309 μL, 2.75 mmol) afforded the title compound as a white crystalline solid (148 mg, 44%) after purification on SiO_2 (100% pet. ether); mp (from pet. ether) 80.3-84.1 °C; IR ν_{max} cm^{-1} 2931 (weak); ^1H NMR δ (400 MHz,

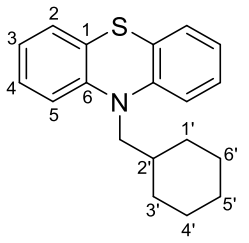
C_6D_6) 1.44-1.65 (4H, m, C(1')H_2, C(3')H_2), 1.82-1.94 (2H, m, C(4')H_2), 2.58-2.72 (1H, m, C(2')*H*), 3.51 (2H, d, *J* = 6.9 Hz, NCH_2), 6.56 (2H, d, *J* = 8.1 Hz, 2xC(5)*H*), 6.71 (2H, dd, *J* = 7.5, 1.2 Hz, 2xC(3)*H*), 6.92-6.97 (2H, m, 2xC(4)*H*), 7.08 (2H, dd, *J* = 7.6, 1.4 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 18.1 (*C*(4')), 26.6 (*C*(1'), *C*(3')), 32.9 (*C*(2')), 53.2 (N*C*H_2), 115.8 (2x*C*(5)), 122.4 (2x*C*(3), 125.9 (2x*C*(1)), 126.9 (2x*C*(4)), 127.5 (2x*C*(2)), 145.8 (2x*C*(6)); LRMS *m/z* (ESI$^+$) 268 $[M+H]^+$; HRMS (ESI$^+$) $C_{17}H_{18}NS^+$ $[M+H]^+$ requires 268.1155; found 268.1151.

10-(Cyclopentylmethyl)-10*H*-phenothiazine, 219



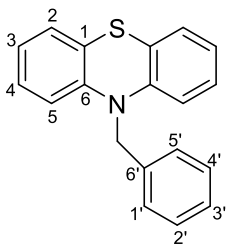
Following general procedure 4b, iodomethylcyclopentane (360 μL, 2.75 mmol) afforded the title compound as a white solid (49 mg, 17%) after purification on SiO_2 (100% pet. ether); mp (from pet. ether) 108.7-109.2 °C; IR ν_{max} cm^{-1} 2950 (weak), 1454 (moderate), 748 (strong, sharp); ^{1}H NMR δ (400 MHz, C_6D_6) 1.13 (2H, dd, *J* = 12.1, 7.2 Hz, C(1')H_A, C(3')H_A), 1.23 - 1.34 (2H, m, C(4')H_A, C(5')H_A), 1.34 - 1.45 (2H, m, C(4')H_B, C(5')H_B), 1.61 (2H, dd, *J* = 12.7, 4.9 Hz, C(1')H_B, C(3')H_B), 2.40 (1H, spt, *J* = 7.5 Hz, C(2')*H*), 3.42 (2H, d, *J* = 7.5 Hz, NCH_2), 6.63 (2H, d, *J* = 8.1 Hz, 2xC(2)*H*), 6.71 (2H, td, *J* = 7.5, 1.1 Hz, 2xC(4)*H*), 6.95 (2H, ddd, *J* = 7.7, 1.5 Hz, 2xC(3)*H*), 7.11 (2H, dd, *J* = 7.6, 1.4 Hz, 2xC(5)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 25.0 (*C*(4'), *C*(5')), 30.2 (*C*(1'), *C*(3')) 36.8 (*C*(2')), 52.4 (N*C*H_2), 115.9 (2x*C*(5)), 122.4 (2x*C*(3)), 126.2 (2x*C*(1)), 126.9 (2x*C*(4)), 127.7 (2x*C*(2)), 145.9 (2x*C*(6)); LRMS *m/z* (ESI$^+$) 282 $[M + H]^+$; HRMS (ESI$^+$) $C_{18}H_{20}NS^+$ $[M+H]^+$ Requires 282.1311; found 282.1310.

10-(Cyclohexylmethyl)-10*H*-phenothiazine, 220



Following general procedure 4b, bromomethylcyclohexane (384 μL, 2.75 mmol) afforded the title compound as a white solid (192 mg, 52%) after purification on SiO_2 (100% pet. ether); mp (from pet. ether) 84.8-85.1 °C; IR ν_{max} cm^{-1} 2927 (weak), 1451 (sharp), 750 (sharp); ^1H NMR δ (400 MHz, C_6D_6) 0.66-0.81 (2H, m, C(1')H_A, C(3')H_A), 0.89-1.03 (3H, m, C(4')H_A, C(5')H_A, C(6')H_A), 1.41-1.54 (3H, m, C(4')H_B, C(5')H_B, C(6')H_B), 1.80-1.91 (2H, m, C(1')H_B, C(2')*H*, C(3')H_B), 3.40 (2H, d, *J* = 6.6 Hz, NCH_2), 6.65 (2H, d, *J* = 8.1 Hz, 2xC(5)*H*) 6.70 (2H, td, *J* = 7.5, 1.1 Hz, 2xC(3)*H*), 6.93-6.98 (2H, m, 2xC(4)*H*), 7.09 (2H, dd, *J* = 7.6, 1.5 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 25.7 (*C*(4'), *C*(6')), 26.5 (*C*(5')), 30.8 (*C*(1'), *C*(3')), 34.6 (*C*(2')), 53.7 (N*C*H_2), 115.8 (*C*(5')), 122.4 (*C*(3')), 126.1 (*C*(1')), 126.9 (*C*(4)), 127.6 (*C*(2)), 146.0 (*C*(3)); LRMS *m/z* (ESI^+) 296 $[M+H]^+$; HRMS (ESI^+) $C_{19}H_{21}NS^+$ $[M+H]^+$ requires 296.1468; found 296.1465.

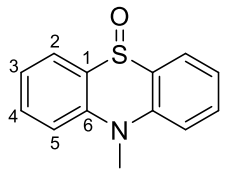
10-Benzyl-10*H*-phenothiazine[192], 221



Following general procedure 4a, phenothiazine (1.00 g, 5.00 mmol) and benzyl bromide (590 μl, 5.00 mmol) afforded the title compound as a white crystalline solid (535 mg, 37%) after

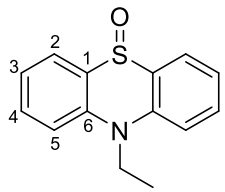
purification on SiO_2 (10:90 CH_2Cl_2:pet. ether); mp (from CH_2Cl_2:pet. ether) 92.7-94.4 °C (lit.[192] (from heptane) 90-91 °C); ^{1}H NMR δ (400 MHz, $CDCl_3$) 5.10 (2H, s, NC*H*$_2$), 6.64 (2H, dd, *J* = 8.2, 1.0 Hz, 2xC(2)*H*), 6.87 (2H, td, *J* = 7.5, 1.4 Hz, 2xC(3)*H*), 6.96-7.00 (2H, 2xC(4)*H*), 7.09 (2H, dd, *J* = 7.5, 1.7 Hz, 2xC(5)*H*), 7.27-7.38 (5H, m, C(1'-5')*H*); LRMS *m/z* (ESI$^+$) 290 $[M+H]^+$.

10-Methyl-10*H*-phenothiazine 5-oxide[233], 222



Following general procedure 5, 10-methyl-10H-phenothiazine **213** (150 mg, 0.703 mmol) afforded the title compound as a white solid (60 mg, 37%) after purification on SiO_2 (acetone:pet. ether 30:70); mp (from acetone) 188.6-195.3 °C (lit.[233] (from EtOH) 187-189 °C); ^{1}H NMR δ (400 MHz, C_6D_6) 2.66 (3H, s, C*H*$_3$), 6.59 (2H, d, *J* = 8.4 Hz, 2xC(5)*H*), 6.81 (2H, td, *J* = 7.4, 0.9 Hz, 2xC(3)*H*), 7.07 (2H, ddd, *J* = 8.4, 7.2, 1.7 Hz, 2xC(4)*H*), 7.74 (2H, dd, *J* = 7.6, 1.5 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 34.0 (*C*H$_3$), 114.9 (2x*C*(5)), 121.4 (2x*C*(3)), 126.6 (2x*C*(1)), 130.3 (2x*C*(2)), 131.4 (2x*C*(4)), 139.6 (2x*C*(6)); HRMS (FI$^+$) $C_{13}H_{11}NOS$ $[FI]^{\cdot}$ requires 229.0561; found 229.0562.

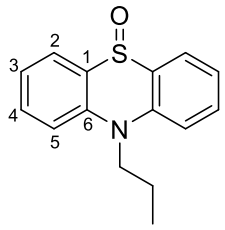
10-Ethyl-10*H*-phenothiazine 5-oxide, 223



Following general procedure 5, 10-ethyl-10H-phenothiazine **214** (114 mg, 0.501 mmol) afforded the title compound as an off-white powder (79 mg, 65%) after purification on SiO_2 (20:80

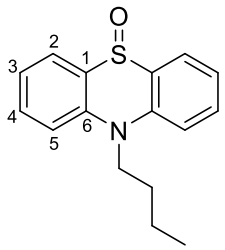
EtOAc:CH_2Cl_2); mp (from EtOAc) 163.2-165.6 °C (lit.[231] (from EtOH) 162-164 °C); IR ν_{max} cm^{-1} 1051 (moderate), 751 (strong); ^{1}H NMR δ (400 MHz, $CDCl_3$) 1.60 (3H, t, *J* = 7.1 Hz, C*H*$_3$), 4.38 (2H, q, *J* = 7.1 Hz, NC*H*$_2$), 7.24-7.29 (2H, m, 2xC(3)*H*), 7.48 (2H, d, *J* = 8.6 Hz, 2xC(5)*H*), 7.65 (2H, ddd, *J* = 8.7, 7.1, 1.6 Hz, 2xC(4)*H*), 7.95-7.98 (2H, m, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 12.1 (*C*H$_3$), 42.7 (N*C*H$_2$), 115.5 (2x*C*(5)), 121.7 (2x*C*(3)), 124.0 (2x*C*(1)), 131.9 (2x*C*(2)), 132.9 (2x*C*(4)), 138.1 (2x*C*(6)); LRMS *m/z* (ESI^+) 266 $[M+Na]^+$; HRMS (ESI^+) $C_{14}H_{13}NNaOS^+$ $[M+Na]^+$ requires 244.0791; found 244.0787.

10-Propyl-10*H*-phenothiazine 5-oxide, 224



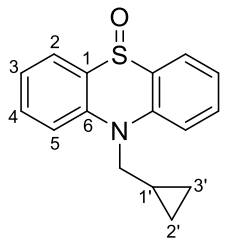
Following general procedure 5, 10-propyl-10H-phenothiazine **215** (100 mg, 0.414 mmol) afforded the title compound as a white solid (57 mg, 54%) after purification on SiO_2 (pet. ether:acetone 80:20); mp (from acetone) 140.2-140.9 °C (lit.[234] (from EtOH) 138-140 °C); ^{1}H NMR δ (400 MHz, C_6D_6); 0.62 (3H, t, *J* = 7.4 Hz, $CH_2CH_2CH_3$), 1.39 (2H, dq, *J*=15.2, 7.5 Hz, $CH_2CH_2CH_3$), 3.38-3.45 (2H, m, $CH_2CH_2CH_3$), 6.79-6.90 (4H, m, 2xC(3)*H*, 2xC(5)*H*), 7.12 (2H, ddd, *J* = 8.6, 7.1, 1.7 Hz, 2xC(4)*H*), 7.76 (2H, dd, *J* = 7.6, 1.7 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6)10.5 ($CH_2CH_2CH_3$), 19.4 ($CH_2CH_2CH_3$), 48.6 ($CH_2CH_2CH_3$), 115.6 (2x*C*(5)), 121.3 (2x*C*(3)), 126.4 (2x*C*(1)), 131.1 (2x*C*(2)), 131.7 (2x*C*(4)), 138.5 (2x*C*6)); LRMS *m/z* (ESI^+) 537 $[2M+Na]^+$.

10-Butyl-10*H*-phenothiazine 5-oxide, 225



Following general procedure 5, 10-butyl-10H-phenothiazine **216** (60 mg, 1.24 mmol) afforded the title compound as a white solid after purification on SiO_2 (pet. ether:acetone 80:20); mp (from acetone) 128.4-129.4 °C; IR ν_{max} cm^{-1} 2962 (weak), 1572 (weak), 1480 (moderate), 1027 (moderate); ^{1}H NMR δ (400 MHz, C_6D_6) 0.69 (3H, t, *J* = 7.3 Hz, $CH_2CH_2CH_2CH_3$), 1.8 (2H, tq, *J* = 7.5, 7.3 Hz, $CH_2CH_2CH_2CH_3$), 1.36-1.45 (2H, m, $CH_2CH_2CH_2CH_3$), 3.52 (2H, dd, *J* = 7.8, 7.3 Hz, $CH_2CH_2CH_2CH_3$), 6.81-6.86 (2H, m, 2xC(3)*H*), 6.93 (2H, d, *J* = 8.6 Hz, 2xC(5)*H*), 7.11-7.17 (2H, m, 2xC(4)*H*), 7.76 (2H, dd, *J* = 7.7, 1.6 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 13.4 ($CH_2CH_2CH_2CH_3$), 19.6 ($CH_2CH_2CH_2CH_3$), 28.1 ($CH_2CH_2CH_2CH_3$), 46.8 ($CH_2CH_2CH_2CH_3$), 115.5 (2x*C*(5)), 121.3 (2x*C*(3)), 126.5 (2x*C*(1)), 131.0 (2x*C*(2)), 131.6 (2x*C*(4)), 138.5 (2x*C*(6)); LRMS *m/z* (ESI^+) 272 $[M+H]^+$; HRMS (ESI^+) $C_{16}H_{18}NOS^+$ $[M+H]^+$ requires 272.1104; found 272.1103.

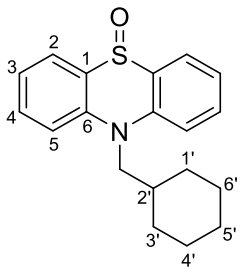
10-(Cyclopropylmethyl)-10*H*-phenothiazine 5-oxide, 226



Following general procedure 5, 10-(cyclopropylmethyl)-10H-phenothiazine, **217** (114 mg, 0.450 mmol) afforded the title compound as an off-white powder (79 mg, 65%) after purification on

SiO_2 (EtOAc:CH_2Cl_2 10:90); mp (from EtOAc) decomp. past 160 °C; IR ν_{max} cm^{-1} 2922 (weak) 1053 (medium); ^1H NMR δ (400 MHz, $CDCl_3$) 0.55-0.60 (2H, m, C(2')H_A and C(3')H_A), 0.81-0.90 (2H, m, C(2')H_B and C(3')H_B), 1.16-1.23 (1H, m, C(1')*H*), 4.17 (2H, d, *J* = 5.0 Hz, NCH_2), 7.23-7.29 (2H, m, 2xC(5)*H*), 7.59-7.66 (4H, m, 2xC(3)*H*, 2xC(4)*H*), 7.93-7.97 (2H, m, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 6.5 (*C*(2'), *C*(3')), 9.7 (*C*(1')), 52.3 (N*C*H_2), 116.5 (2x*C*(5)), 121.8 (2x*C*(3)), 124.6 (2x*C*(1)), 131.4 (2x*C*(2)), 132.7 (2x*C*(4)), 139.0 (2x*C*(6)); LRMS *m/z* (ESI^+) 270 $[M+H]^+$ 290 $[M+Na]^+$; HRMS (ESI^+) $C_{16}H_{15}NOSNa^+$ $[M+Na]^+$ requires 292.0767; found 292.0756.

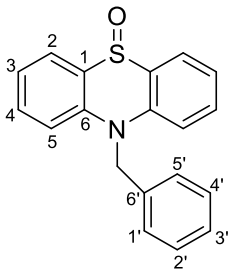
10-(Cyclohexylmethyl)-10*H*-phenothiazine 5-oxide, 228



Following general procedure 5, 10-(cyclohexylmethyl)-10H-phenothiazine **220** (50 mg, 0.169 mmol) afforded the title compound as a white solid (35 mg, 66%) after purification on SiO_2 (pet. ether:acetone 80:20); mp (from acetone) 165.1-165.8 °C; IR ν_{max} cm^{-1} 2918 (broad), 1457 (sharp), 1023 (sharp, strong), 748 (sharp, strong); ^1H NMR δ (400 MHz,C_6D_6) 0.58-0.73 (2H, m, C(1')H_A, C(3')H_A), 0.82-0.97 (3H, m, C(4')H_A, C(5')H_A, C(6')H_A), 1.36-1.49 (3H, m, C(4')H_B, C(5')H_B, C(6')H_B), 1.72 (2H, d, *J* = 13.3 Hz, C(1')H_B, C(3')H_B), 1.82 (1H, ttt, *J* = 11.0, 7.3, 3.7 Hz, C(2')*H*), 3.62 (2H, d, *J* = 7.0 Hz, NCH_2), 6.78-6.84 (2H, m, 2xC(3)*H*) 7.02 (2H, d, *J* = 8.4 Hz, 2xC(5)*H*), 7.11-7.18 (2H, m, 2xC(4)*H*), 7.71 (2H, dd, *J* = 7.6, 1.5 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 25.5 (*C*(4'), *C*(6')), 26.3 (*C*(5')), 30.5 (*C*(1'), *C*(3')), 35.9 (*C*(2')), 52.2

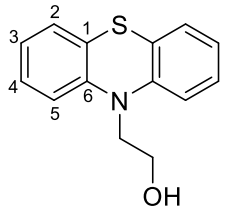
(N*C*H$_2$), 116.3 (2x*C*(5)), 121.6 (2x*C*(3)), 127.8 (2x*C*(1)), 130.1 (2x*C*(2)), 131.3 (2x*C*(4)), 139.4 (2x*C*(6)); LRMS *m/z* (ESI$^+$) 312 [M+H]$^+$; HRMS (ESI$^+$) $C_{19}H_{22}NOS^+$ [M+H]$^+$ requires 312.1417; found 312.1427.

10-Benzyl-10*H*-phenothiazine 5-oxide, 229



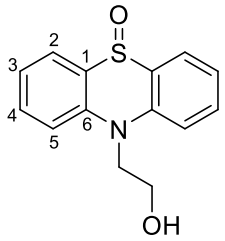
Following general procedure 5, 10-benzyl-10H-phenothiazine **221** (145 mg, 0.501 mmol) afforded the title compound as a pale pink solid (95 mg, 62%) after purification on SiO_2 (10:90 EtOAc:CH_2Cl_2); mp (from EtOAc) 215-220 °C (lit.[231] (solvent unspecified) 226.5-228 °C); IR ν_{max} cm^{-1} 1461 (moderate), 1014 (sharp); ^{1}H NMR δ (400 MHz $CDCl_3$) 5.52 (2H, s, NC*H*$_2$), 7.14-7.18 (2H, m, C(1')*H*, C(5')*H*), 7.21 (2H, d, *J* = 8.6 Hz, 2xC(5)*H*), 7.24-7.30 (2H, m, 2xC(3)*H*), 7.30-7.40 (3H, m, C(2'-4')*H*), 7.51 (2H, ddd, *J* = 8.6, 7.2, 1.6 Hz, 2xC(4)*H*), 8.00 (2H, dd, *J* = 7.7, 1.6 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 52.9 (N*C*H$_2$), 116.7 (2x*C*(5)), 122.3 (2x*C*(3)), 123.6 (2x*C*(1)), 125.8 (*C*(1'), *C*(5')), 127.6 (*C*(3')), 129.2 (*C*(2'), *C*(4')), 131.5 (2x*C*(2)), 133.2 (2x*C*(4)), 135.2 (2x*C*(6')), 139.1 (2x*C*(6)); LRMS *m/z* (ESI$^+$) 328 [M+Na]$^+$; HRMS (ESI$^+$) $C_{19}H_{15}NNaOS^+$ [M+Na]$^+$ requires 306.0947, found 306.0940.

2-(10*H*-Phenothiazin-10-yl)ethan-1-ol[235], 230



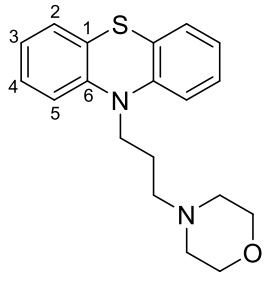
Phenothiazine (2 g, 10.0 mmol, 1.0 eq.) was dissolved in THF (20 mL), potassium bis(trimethylsilyl)amide (as 0.5 M solution in toluene, 31.2 mL, 15.6 mmol, 1.5 eq.) was added drop-wise and the resulting mixture brought to room temperature and stirred for 15 min. 1,3,2-Dioxathiolane 2,2-dioxide (1.49 g, 12.0 mmol, 1.2 eq.) was added in one portion and the reaction stirred for 18 h. $H_2SO_{4(aq)}$ (conc.) was added drop-wise until the solution reached pH 1. Water (5 mL) was added drop-wise and the reaction stirred for 3 h. The reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (20 mL) and extracted with Et_2O (1x30 mL) and EtOAc (2x30 mL). The organic layers were combined, washed with water (3x30 mL) and dried to give the title compound as a dark purple oil (1.64 g, 67%) after purification on SiO_2 (30:70 Et_2O:pet. ether); ^{1}H NMR δ (400 MHz, DMSO-d_6) 3.71 (2H, td, J = 6.1, 5.6 Hz, CH_2CH_2OH), 3.90-3.98 (2H, m, CH_2CH_2OH), 4.90 (1H, t, J = 5.4 Hz, O*H*), 6.93 (2H, td, J=7.5, 1.2 Hz, 2xC(3)*H*), 7.03 (2H, dd, J = 8.3, 1.1 Hz, 2xC(5)*H*), 7.12 (2H, dd, J = 7.6, 1.5 Hz, 2xC(2)*H*), 7.16-7.22 (2H, m, 2xC(4)*H*); ^{13}C NMR δ (100 MHz, DMSO-d_6) 50.0 (CH_2CH_2OH), 58.2 (CH_2CH_2OH), 116.1 (2x*C*(5)), 123.0 (2x*C*(3)), 123.6 (2x*C*(1)), 127.5 (2x*C*(2)), 128.1 (2x*C*(4)), 145.1 (2x*C*(6)); LRMS *m/z* (ESI^+) 244 $[M + H]^+$, 266 $[M+Na]^+$, 282 [M+K]+.

10-(2-Hydroxyethyl)-10*H*-phenothiazine 5-oxide, 233



Following general procedure 5, 10-(2-hydroxyethyl)-10H-phenothiazine **230** (108 mg, 0.444 mmol) afforded the title compound as a pink crystalline solid (79 mg, 70%) after purification on SiO_2 (CH_2Cl_2:MeOH 98:2 to 97:3); mp (from MeOH) 174.2-176.8 °C; IR ν_{max} cm^{-1} 3507 (very broad), 2981 (broad), 2362 (weak), 1589 (moderate), 1466 (strong, sharp); ^{1}H NMR δ (400 MHz, DMSO-d_6) 3.86 (2H, dt, J = 6.4, 6.3 Hz, CH_2CH_2OH), 4.42 (2H, t, J = 6.5 Hz, CH_2CH_2OH), 5.01 (1H, t, J = 5.7 Hz, CH_2CH_2OH), 7.3 (2H, ddd, J = 7.8, 6.6, 1.5 Hz, 2xC(3)*H*), 7.68-7.77 (4H, m, 2xC(4)*H*, 2xC(5)*H*), 7.95 (2H, dd, J = 7.6, 1.4 Hz, 2xC(2)*H*); ^{13}C NMR δ (100 MHz, DMSO-d_6) 50.3 (CH_2CH_2OH), 58.2 (CH_2CH_2OH), 117.3 (2x*C*(5)), 122.2 (2x*C*(3)), 124.6 (2x*C*(1)), 131.5 (2x*C*(2)), 133.4 (2x*C*(4)), 138.7 (2x*C*(6)); LRMS *m/z* (ESI^+) 260 $[M+H]^+$; HRMS (ESI^+) $C_{14}H_{14}NO_2S^+$ $[M+H]^+$ requires 260.0740; found 260.0740.

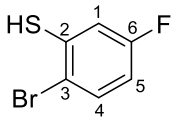
4-(3-(10*H*-phenothiazin-10-yl)propyl)morpholine, 241



To a solution of 2-bromothiophenol (361 μL, 3.00 mmol, 1 eq.), 2-bromoiodobenzene (386 μL, 3.00 mmol, 1 eq.) and 3-morpholinopropylamine (438 μL, 3.00 mmol, 1 eq.) in toluene (13 mL) was added NaO*t*Bu (1.15 g, 12.0 mmol, 4 eq.), Pd_2dba_3 (69 mg, 0.075 mmol, 0.025 eq.) and 1,1'-

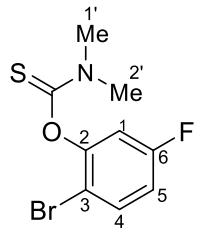
bis(diphenylphosphino)ferrocene (56 mg, 0.119 mmol, 0.040 eq.), and the mixture heated at 60 °C for 20 min with MW irradiation and then to 120 °C for 2 h also with MW irradiation. The mixture was cooled to room temperature and filtered twice through filter paper. The filtrate was concentrated, treated with water (30 mL) and extracted into Et_2O (30 mL) and EtOAc (2x 30 mL). The organic layers were dried and concentrated under *vacuum*. Two attempts were made to purify the residue *via* SiO_2 chromatography (99:1 CH_2Cl_2:MeOH) then (99.8:0.2 to 99.5:0.5 to 99:1 CH_2Cl_2:MeOH) and a subsequent attempt was made to increase purity by preparative TLC (98:2 CH_2Cl_2:MeOH). Selected data of the crude mixture supporting synthesis of the title compound: ^{1}H NMR δ (400 MHz, C_6D_6) 1.57-1.65 (2H, m), 2.02-2.12 (6H, m), 3.47-3.52 (4H, m), 3.58 (2H, t, J = 6.9 Hz), 6.63-6.72 (4H, m), 6.70 (2H, J = 7.5, 0.8 Hz), 7.07 (2H, J = 7.6, 1.4 Hz); ^{13}C NMR δ (100 MHz, C_6D_6) 24.0, 44.7, 53.7, 55.5, 66.8, 115.5, 122.4, 125.4, 127.0, 145.4; LRMS *m/z* (ESI^+) 327 $[M+H]^+$.

2-Bromo-5-fluorobenzenethiol[236], 245



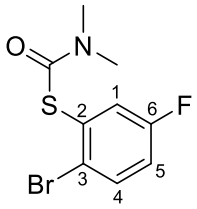
S-(2-Bromo-5-fluorophenyl) dimethylcarbamothioate **249** (1.82 g, 6.59 mmol) was added to MeOH (60 mL) and $NaOH_{(aq)}$ (1N, 60 mL) and heated to reflux (100 °C) for 2 h. The mixture was cooled to room temperature and concentrated under *vacuum*. The residue was diluted with water and extracted with CH_2Cl_2 (3x30 mL, discarded) then acidified to pH 1 with $HCl_{(aq)}$ (1 M) and extracted with EtOAc (3x30 mL) to afford the title compound as a colourless oil (1.07 g, 79%) which was used without further purification. ^{1}H NMR δ (400 MHz, $CDCl_3$) 6.74 (1H, ddd, J = 8.9, 7.9, 2.9 Hz, C(5)*H*), 7.10 (1H, dd, J = 8.7, 2.9 Hz, C(4)*H*), 7.48 (1H, dd, J = 8.9, 5.5 Hz, C(1)*H*); ^{19}F NMR δ (377 MHz, $CDCl_3$) −114.2 (F, C*F*); LRMS *m/z* (ESI^-) 204 $[M-H]^-$.

***O*-(2-Bromo-5-fluorophenyl) dimethylcarbamothioate, 248**



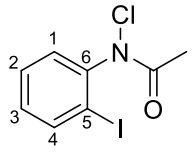
2-Bromo-5-fluorophenol (1.16 mL, 10.4 mmol, 1.0 eq.) was added to a suspension of K_2CO_3 (2.20 g, 15.9 mmol, 1.5 eq.) in DMF (10.5 mL). Dimethylthiocarbamoyl chloride (1.55 g, 12.5 mmol, 1.2 eq.) was added and the reaction stirred at room temperature for 17 h. The mixture was poured slowly into water (8 mL) containing 2 g of ice and the resulting slurry was stirred for 2 h. The precipitate was filtered off by *vacuum* filtration, dissolved in CH_2Cl_2 and washed with water (3x100 mL). The organic layer was dried, filtered, and concentrated under *vacuum* to afford the title compound as a bright yellow solid (2.47 g, 85%), which did not require further purification; mp (from CH_2Cl_2) 68.7-70.1 °C (lit.[237] (from DMF) 63.0 °C); IR ν_{max} cm^{-1} 1147 (strong); ^1H NMR δ (400 MHz, $CDCl_3$) 3.33 (3H, s, C*H*$_3$(2')), 3.40 (3H, s, C*H*$_3$(1')), 6.81-6.85 (1H, m, C(5)*H*), 6.87 (1H, dd, *J* = 8.7, 2.9 Hz, C(1)*H*), 7.48 (1H, dd, *J* = 8.8, 5.7 Hz, C(4)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 39.0 (*C*H$_3$(2')), 43.5 (*C*H$_3$(1')), 111.9 (d, *J* = 4.8 Hz, *C*(3)), 113.5 (d, *J* = 24.8 Hz, *C*(1)), 114.7 (d, *J* = 21.9 Hz, *C*(5)), 133.4 (d, *J* = 8.6 Hz, *C*(4)), 151.7 (d, *J* = 11.4 Hz, *C*(2)), 161.8 (d, *J* = 248.9 Hz, *C*(6)), 185.6 (*C*=S); ^{19}F NMR δ (377 MHz, $CDCl_3$) −112.8 (F, C*F*); HRMS (FI) $C_9H_9BrFNOS^{\cdot}$ $[M]^{\cdot}$ requires 276.9572; found 276.9567.

_S_-(2-Bromo-5-fluorophenyl) dimethylcarbamothioate, 249



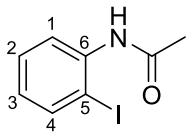
O-(2-Bromo-5-fluorophenyl) dimethylcarbamothioate **248** (2.40 g, 8.63 mmol, 1.0 eq.) was added to *N,N*-diethylaniline (8 mL) and heated for 2 h with MW irradiation at 220 °C. The reaction mixture was cooled to room temperature and poured slowly onto 20 g of ice in 30 mL $HCl_{(aq.)}$ (4N). The slurry was stirred for 30 min and left to stand for 17 h, it was then stirred for a further 15 min. The precipitate was filtered off by *vacuum* filtration, dissolved in CH_2Cl_2 (30 mL) and washed with water (3x100 mL). The organic layer was dried, filtered, and concentrated under *vacuum* to afford the title compound as a brown solid (2.08 g, 87%) after purification on SiO_2 (pet. ether:EtOAc 95:5). The product was freshly prepared as required since it spontaneously rearranges from the carbonyl **249** back to the thiocarbonyl **248** after a few days; mp (from CH_2Cl_2) 69.1-69.6 °C (lit.[237] (from *N,N*-diethylaniline) 69.1-69.6 °C); IR ν_{max} cm^{-1} 680 (strong); ^{1}H NMR δ (400 MHz, $CDCl_3$) 3.00-3.20 (6H, m, 2xNC*H*$_3$), 6.96-7.04 (1H, m, C(5)*H*), 7.38-7.42 (1H, m, C(1)*H*), 7.64 (1H, dd, *J* = 8.9, 5.3 Hz, C(4)*H*); ^{13}C NMR δ (100 MHz, $CDCl_3$) 37.1 (2x*C*H$_3$), 118.2 (d, *J* = 21.9 Hz, *C*(5)), 124.9 (d, *J* = 22.9 Hz, *C*(1)), 132.4 (d, *J* = 8.6 Hz, *C*(2)), 134.1 (d, *J* = 7.6 Hz, *C*(4)), 161.3 (d, *J* = 249.9 Hz, *C*6), 164.5 (*C*=O), C(3) not resolved; ^{19}F NMR δ (377 MHz, $CDCl_3$) −114.1 (F, C*F*); LRMS *m/z* (ESI^+) 278 $[M + H]^+$; HRMS (ESI^+) $C_9H_9BrFNNaOS^+$ $[M+Na]^+$ requires 299.9465; found 299.9456.

***N*-chloro-*N*-(2-iodophenyl)acetamide, 258**



To a solution of *N*-(2-Iodophenyl)acetamide, **259** (100 mg, 0.383 mmol, 1 eq.) in Et_2O (5 mL) was slowly added a solution of $Ca(ClO)_2$ (163 mg, 1.14 mmol, 3 eq.) and $NaHCO_3$ (96 mg, 1.14 mmol, 3 eq.) in water (5 mL) at 0 °C. The biphasic reaction was stirred for 30 min and the aqueous phase separated and extracted with Et_2O (3x20 mL). The organic layers were combined, dried and concentrated under *vacuum* to afford the title product, which was used immediately without the need for purification. Selected data of the crude mixture supporting synthesis of title compound: ^{1}H NMR δ (400 MHz, C_6D_6) 1.66 (3H, s, C*H*$_3$), 6.39 (1H, *J* = 7.5, 7.8 Hz, C(2)*H*), 6.69-6.76 (1H, m, C(1)*H*), 6.82 (1H, br. s, C(4)*H*), 7.42 (1H, d, *J* = 8.0 Hz); ^{13}C NMR δ (100 MHz, C_6D_6) 21.6 (*C*H$_3$), 100.2 (*C*(5)), 129.5 (*C*(2)), 130.2 (*C*(3)) 130.7 (*C*(1)), 140.1 (*C*(4)), 145.6 (*C*(6)); LRMS *m/z* (ESI$^+$) 295 $[M+H]^+$; HRMS not found.

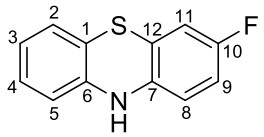
***N*-(2-Iodophenyl)acetamide, 259**



2-Iodoaniline (5 g, 22.85 mmol, 1 eq.) and triethylamine (9.55 mL, 68.5 mmol, 3.0 eq.) were dissolved in CH_2Cl_2 (40 mL). Acetic anhydride (6.45 mL, 68.2 mmol, 3.0 eq.) was added drop-wise at 0 °C and the reaction was brought to room temperature and stirred for 2 h. The mixture was diluted with CH_2Cl_2 and washed with water and 1 M $NaOH_{(aq.)}$ to afford the title compound as a white solid (4.99 g mg, 84%) after purification on SiO_2 (pet. ether:EtOAc 80:20); mp (from

EtOAc) 109.2-109.6 °C (lit.[238] (from MeOH) 109-110 °C); ^{1}H NMR δ (400 MHz, C_6D_6) 1.51 (3H, s, C*H*$_3$), 6.35 (1H, td, *J* = 7.6, 1.5 Hz, C(3)*H*), 6.96 (1H, t, *J* = 7.6 Hz, C(2)*H*), 7.46 (1H, dd, *J* = 7.9, 1.5 Hz, C(4)*H*), 8.63 (1H, d, *J* = 7.3 Hz, C(1)*H*); ^{13}C NMR δ (100 MHz, C_6D_6) 23.6 (*C*H$_3$), 89.3 (*C*(5)), 121.7 (*C*(1)), 125.3 (*C*(3)), 129.1 (*C*(2)), 138.5 (*C*(4)), 138.9 (*C*(6)), 167.1 (C, C=O); LRMS *m/z* (ESI$^-$) 260 [M−H]$^-$.

3-fluoro-10*H*-phenothiazine, 263



To a solution of 2-bromothiophenol (361 μL, 3.00 mol, 1 eq.), and 4-fluoro-2-iodoaniline (354 μL, 3.00 mmol, 1 eq.) in toluene (13 mL) was added NaO*t*Bu (384 mg, 4.00 mmol, 1.3 eq.), Pd_2dba_3 (69 mg, 0.075 mmol, 0.025 eq.) and 1,1'-*bis*(diphenylphosphino)ferrocene (166 mg, 0.299 mmol, 0.10 eq.) and the mixture heated at 60 °C for 20 min with MW irradiation and then to 160 °C for 2 h also with MW irradiation. The mixture was cooled to room temperature and filtered twice through filter paper then treated with water (20 mL) and extracted into Et_2O (3x20 mL). The organic layers were dried and concentrated under *vacuum* and the residue purified *via* Dowex® chromatography, followed by SiO_2 chromatography (1:9 EtOAc:Pet. ether) to afford the title compound as a blue solid (52 mg, 8%). Selected data supporting the synthesis of the title compound: ^{1}H NMR δ (400 MHz, C_6D_6) 4.81 (1H, br. s, N*H*), 5.61 (1H, dd, *J* = 4.7 Hz), 5.91 (1H, d, *J* = 7.9 Hz), 6.40 (1H, td, *J* = 8.4, 2.8 Hz), 6.51-6.59 (2H, m) 6.71-6.80 (2H, m); ^{19}F NMR δ (377 MHz, C_6D_6) −122.3 (1F, C(10)*F*); LRMS not found.

6.2 – Biochemistry and Molecular Biology

6.2.1 - General

Chemicals were purchased Apollo Scientific, Sigma Aldrich, Alfa Aesar or Fisher Scientific unless otherwise stated. Standard sterile techniques were practised where necessary. All equipment and media requiring sterilisation were autoclaved at 121 °C for 20 min. For heat-sensitive solutions, filtration through 0.22 μm filters was carried out to ensure aseptic conditions. Water was purified using a Millipore Elix® 10 system and then sterilised through a Millipore Milli-Q purification system fitted with a 0.22 μm filter.

6.2.2 - Measurement of pH

The probe used to monitor pH of buffer solutions was a Hanna Instruments HI 9321 microprocessor pH meter fitted with a 5 mm diameter electrode. Solutions were stirred with a magnetic bar and pH adjusted by drop-wise addition of either $HCl_{(aq)}$ (conc., 2 M or 1 M) or $NaOH_{(aq)}$ (8 M, 2 M or 1 M). The instrument was calibrated before use with standard calibration solutions (Fischer Scientific).

6.2.3 - Waterbaths and Static Incubators

For 'heat shock' bacterial transformation, a Grant W14 waterbath was used at 42 °C.

For plated cell growth and assays, a Hereaus incubator was used at either 30 °C or 37 °C as stated.

6.2.4 - Media Recipes

2-Tryptone-Yeast (2-TY):

A 1 L solution contains 16 g bacto-tryptone, 10 g yeast extract and 5 g NaCl.

Terrific Broth (TB):

A 900 mL solution containing 12 g bacto-tryptone, 24 g yeast extract and 4 mL glycerol was autoclaved. Separately, 100 mL KH_2PO_4 (0.17 M), K_2HPO_4 (0.72 M) was autoclaved. Once the two solutions had cooled to room temperature they were mixed to a total volume of 1 L.

6.2.5 - SDS-PAGE Gel Electrophoresis

Precast 4-12% NuPAGE® Bis-Tris Mini Gels (Novex) were used with 1X NuPAGE® MES or MOPS SDS running buffer prepared by diluting 20X standards purchased from Novex. 500 µL NuPAGE® antioxidant was added to the upper buffer chamber prior to running the gel. Samples were prepared by mixing 6.5 µL of protein sample with 2.5 µL NuPAGE® LDS sample buffer (4X) and 1 µL NuPAGE® reducing agent (10X) and heating to 70 °C for 10 min. 7 µL of the mixture was loaded onto the gel and 4 µL of standard ladder PageRuler™ (Thermo Fisher) added to an adjacent lane.

Gels were run at 200 V constant for 35 min (MES buffer systems) or 50 min (MOPS buffer systems) using Bio-Rad apparatus. Gels were stained for 1 h in a solution of 45% EtOH, 45% H_2O, 10% AcOH containing 0.25% (w/v) Coomassie™ Blue stain. Gels were de-stained for 3 h in the same solution but without the addition of Coomassie™ Blue.

The gel from showing purified His-PRMT1 from the two different expressions is shown below (**Figure 6.2**).

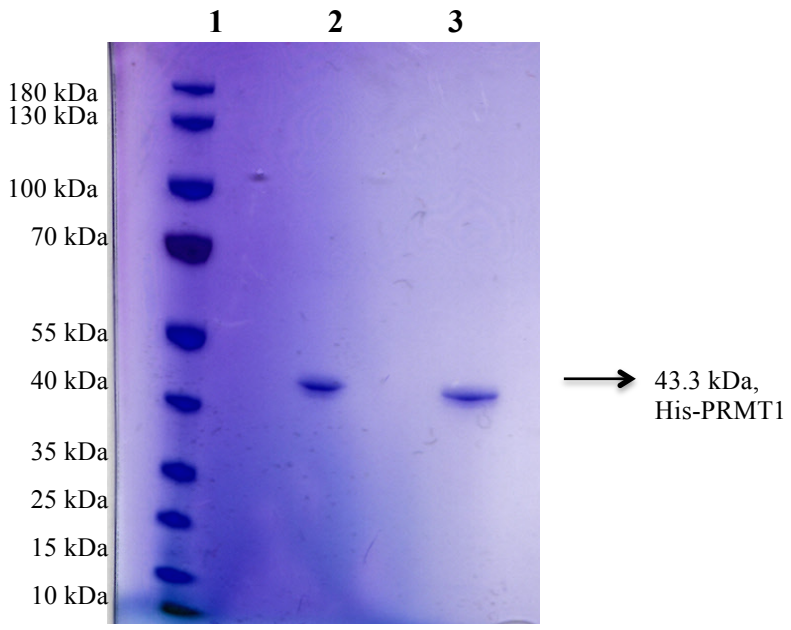


Figure 6.2 – *SDS-PAGE of His-PRMT purified* via *affinity chromatography. Lane 1 = ladder (PageRuler™), Lane 2 = Pure recombinant His-PRMT1 from expression using 2-TY media, Lane 3 = Pure recombinant His-PRMT1 from expression using TB media.*

6.2.6 - Gene Amplification and Vectors

PRMT1 was expressed from a pNIC28-BSA4 vector encoding an *N*-terminal hexa-His tag and kanamycin resistance gene. This vector was gifted from Dr L. Walport.

6.2.7 - Transformation

25 µL BL21 (DE3) *E. coli* competent cells, gifted from Dr L. Walport, were thawed on ice and 1µL plasmid DNA (100 ng/µL) added. The cells were left on ice for 20 min before being heat-shocked at 42 ºC for 30 s. Cells were immediately cooled on ice for a further 2 min before addition of 400 µL 2-TY media. The suspension was incubated for 1 h at 37 ºC (without shaking) before plating of 250 µL onto a 2-TY agar plate containing kanamycin (30 µg/mL). Plates were incubated at 37 ºC for 24 h.

6.2.8 - Protein Expression

Starter Cultures:

Using a sterile pipette tip, a single colony was picked from the agar plates containing transformed cells and added to a 500 mL conical flask containing 100 mL 2-TY media (prepared by Mr H. M. Jubier) and 30 µg/mL kanamycin. The flask was incubated for 14 h at 37 ºC with 250 rpm shaking in a New Brunswick Scientific G25 environmental shaker.

Large-Scale Growth:

6 mL of starter culture was used to inoculate 600 mL of media in a 2 L Tunair® flask containing 30 µg/mL kanamycin. Five 2 L flasks were used with 2-TY media for the first PRMT1 expression and twelve 2 L flasks were used with TB media for the second expression. Flasks were incubated at 37 ºC with 180 rpm shaking until OD_{600} reached ~0.6, upon which cells were induced with IPTG to a final concentration of 100 µM. Induced cells were incubated at 18 ºC with 180 rpm shaking for 16 h before being harvested by centrifugation (8000 rpm, 8 min, 20 °C) and transferred to re-sealable plastic bags for storage at -80 ºC.

6.2.9 - Protein Purification

Buffers required for protein purification:

The following buffers were all pH adjusted to 7.5 and contained 5% glycerol:

Binding buffer: 50 mM HEPES, 500 mM NaCl, 20 mM imidazole, 1 mM TCEP.

Wash buffer: 50 mM HEPES, 500 mM NaCl, 40 mM imidazole, 1 mM TCEP.

Elution buffers: 50 mM HEPES, 500 mM NaCl, 250/500 mM imidazole, 1 mM TCEP.

Protein storage buffer: 10 mM HEPES, 200 mM NaCl.

Isolation and purification:

Frozen cell pellets from large-scale expressions were manually smashed into small pieces and then defrosted at room temperature in 200 mL of binding buffer. Once the suspension was homogenous, 10 µg/mL DNase I (Roche) and 2 EDTA-free protease inhibitor tablets (Roche) were added. Cells were lysed on ice by sonication (3 min 30 sec of 3 s on, 4 s off) using a Vibra Cell VCX 500 with a 13 mm probe. The resultant suspension was centrifuged (22,000 rpm, 25 min, 6 ºC) and the supernatant filtered through a 0.4 µm filter before being loaded onto a His-trap purification column (GE Healthcare).

PRMT1 was purified by Ni-affinity chromatography using an Akta purification system. The crude mixture was loaded onto the column in binding buffer, washed with wash buffer and then PRMT1 eluted in 250 mM and 500 mM elution buffers using a step gradient.

Concentration and storage:

Fractions containing PRMT1 were pooled and concentrated using a 30 kDa-exclusion filter tube (Millipore) and centrifuging for 5-20 min intervals (2,000 rpm, 4 °C). Buffer-exchange was achieved by resuspension-concentration-resuspension using protein storage buffer until imidazole concentration was below 40 mM. Protein was aliquoted at 4°C and flash-frozen in liquid nitrogen before immediate transfer to a -80 °C freezer for storage for up to three months.

6.3 – Peptide Synthesis

6.3.1 - Synthesis

Standard amino acids were purchased from Novabiochem or CS Bio. SPPS was carried out using a Liberty Blue (CEM Corporation) automated peptide synthesiser and peptides were synthesised

from *C*- to *N*-terminus using a 100-200 mesh rink amide resin (4-(2',4'-dimethoxyphenyl-fmoc-aminomethyl)-phenoxyacetamidonorleucyl-MBHA) (AGTC Bioproducts) as the solid support. For each cycle, protected amino acids were activated with DIC and HoBt and added to the reactor in 10 times excess (except MMA(δ), which was added in two-times excess) as a 0.2 M solution in DMF. Each coupling reaction was carried out using standard Liberty Blue procedures: reactions were heated with MW irradiation for 3 h with stirring, except in the case of Arg, which was reacted for 6 h at room temperature. For the coupling of MMA(δ) the reaction was stirred for 8 h at room temperature. After each coupling, the α-Fmoc group was removed by treatment with 20% piperidine (Merck) in DMF (Rathburn Chemicals).

6.3.2 – Cleavage and Isolation

After the final coupling reaction, the resin was filtered from the solvent under *vacuum*, washed with copious amounts of CH_2Cl_2 and dried under high *vacuum* for 10 min. The dry resin was added to a solution containing 4 mL TFA, 100 μL triisopropylsilane, 100 μL water and 100 μL 1,3-dimethoxybenzene. The solution was stirred gently for 3.5 h at room temperature and then the resin removed by filtration under *vacuum*. The peptide was precipitated from the filtrate by addition of 40 mL ice-cold Et_2O. Centrifugation (2000 rpm, 10 min) pelleted the peptide – the supernatant was removed and the pellet re-suspended in the minimum volume of water. The solution was lyophilized to afford the crude peptide.

6.3.3 – Preparative HPLC Purification

The crude peptide was dissolved in 2 mL water and filtered through a 0.22 μm filter before injection onto a Vydac C_{18} reverse-phase HPLC column. A gradient of eluent A (0.1% TFA in water) to eluent B (0.1% TFA in MeCN) was applied from 100:0 to 80:20 over 45 min. Fractions

containing peptide (assessed by MALDI-MS) were pooled and lyopholised to afford the pure product.

6.3.4 – Analytical LC/MS

Low resolution positive ion ESI mass spectrometry data were recorded using an Agilent 1200 series LC/MS system (6120 Quadrupole MS) with a Waters Sunfire C_{18} column.

6.4 – Assay Protocols

All assay procedures were run in triplicate and with suitable positive and negative controls.

6.4.1 – Chemiluminescence Assay

PRMT1 and PRMT5 chemiluminescence assay kits were purchased from Amsbio (product codes 52004L and 52002L respectively). Substrate histones were immobilised to a 96-well plate. Micro-wells were rehydrated by addition of 150 µL TBST buffer (50 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20, pH 8.0) and incubation at RT for 15 min. 40 µM SAM and 1 µL inhibitor (final conditions of 1% DMSO) were added per well and the volume adjusted to 50 µL using supplied HMT Tris-based assay buffer. Enzyme was added to initiate the reaction (6 ng PRMT1/well or 700 ng PRMT5/well) and the plate incubated at 37 °C for 1 h (PRMT1) or 2 h (PRMT5). Wells were washed three times with 200 µL TBST buffer. Primary antibody was diluted as per manufacturer's instructions and 100 µL added to the wells for incubation with gentle shaking (RT, 50 rpm, 1 h). Plates were washed as before and then incubated in 100 µL of a supplied BSA-based blocking solution (RT, 50 rpm, 10 min). HRP-conjugated secondary antibody was diluted as per manufacturer's instruction and incubated with gentle shaking (RT, 50 rpm, 30 min). Wells were washed and blocked as before following which 100 µL of supplied HRP substrate was added and chemiluminescence immediately detected using an Omega

FLUOstar plate reader (BMG Labtech). Results were blank-corrected and anomalies were excluded from analysis.

6.4.2 – Radiometric Assay

The scintillation proximity assay was used as previously reported[239]. The assays were set up as follows:

PRMT1:

Assay conditions: 20 mM Tris-HCl, pH 8.0 containing 0.01% Triton X-100 and 5 mM DTT; 20 nM PRMT1, 130 nM peptide (Biotin-H4-1-24) and 4.3 μM SAM.

PRMT3:

Assay conditions: 20 mM Tris-HCl, pH 7.5 containing 0.01% Tween-20 and 5 mM DTT; 20 nM PRMT3, 0.6 μM peptide (Biotin-H4-1-24) and 28 μM SAM.

PRMT5:

Assay conditions: 20 mM Tris-HCl, pH 8.5 containing 0.01% Tween-20 and 10 mM TCEP; 15 nM PRMT5-MEP50 complex, 120 nM peptide (Biotin-H4-1-24) and 2 μM SAM.

PRMT6:

Assay conditions: 20 mM BTP, pH 7.5; 0.01% Tween-20 and 10 mM DTT; 50 nM PRMT6; 0.6 μM peptide (H4 1-24) and 2.3 μM SAM

PRMT7:

20 mM Tris-HCl, pH 8.5 containing 0.01% Tween-20 and 5 mM DTT, 25 nM PRMT7, 0.3 μM peptide (H2B 23-37) and 1.1 μM SAM

PRMT8:

20 mM Tris-HCl, pH 8.0 containing 0.01% Triton X-100 and 5 mM DTT; 25 nM PRMT8, 0.7 μM peptide (Biotin-H4-1-24) and 2.2 μM SAM.

All reactions were incubated at 23 °C for 30 min and were quenched by addition of equal volumes of 7.5 M guanidine hydrochloride and the volume finalised to 200 μL by addition of 20 mM Tris-HCl (pH 8.0). Samples were transferred to scintillation proximity assay plates (FlashPlate® PLUS, PerkinElmer Life Sciences) and were incubated for 1 h. The signals were measured using a TopCount NXT™ Microplate Scintillation and Luminescence Counter (PerkinElmer Life Sciences).

All assays were carried out in triplicate unless otherwise stated and expressed as averages ±standard deviation. The final concentration of DMSO differed depending on the concentration of compound tested; 50 μM tests (0.25% DMSO), 25 μM tests (0.05% DMSO) or 5 μM (0.01% DMSO). Percentage activity of each individual test compound was expressed with respect to the negative control, which was pinned to 100% activity.

6.4.3 – Large Scale MALDI-MS Assay

Inhibitor in DMSO was added to each test well in a 96-well plate. Buffer was added followed by peptide and then SAM. Reactions were initiated by addition of enzyme with aspiration. Controls containing DMSO in place of inhibitor, water in place of SAM and buffer in place of enzyme

were also initiated. Assays were incubated for 90 min at 30 °C and then quenched by pipetting 100 µL MeOH into each well. 1 mg α-cyano-4-hydroxycinnamic acid (CHCA) was dissolved in 100 µL matrix solution (1:1:0.01 MeCN:H_2O:TFA (v/v/v)). 1 µL of the matrix solution was mixed with 1 µL of sample on the MALDI plate and allowed to dry at room temperature. Data was shot on a Waters® Micromass® MALDI micro MX™ using 12,000 Volt laser energy and 10-12 scans per well.

The following quantities were added:

Assay Component	**Volume**	**Final Assay Conditions**
SAM	10 µL	At approximated K_m value: 10 µM
Peptide substrate	10 µL	Close to K_m value: 15 µM
Enzyme	20 µL	Dependent on specific activity
Buffer	59 µL	20 mM Tris, 10 mM DTT
Inhibitor in DMSO	1 µL	1% DMSO
TOTAL volume	100 µL	

6.4.4 - Small Scale MALDI-MS Assay

The small scale assay was set up identically to the large scale assay except using the following reduced volumes and 384 well plates.

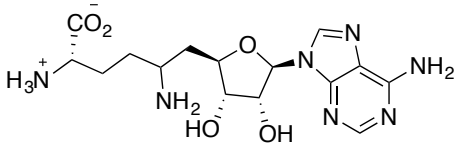
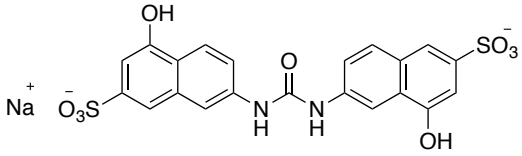
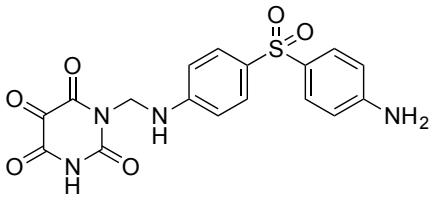
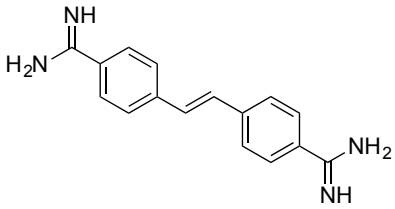
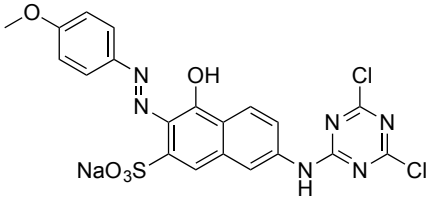
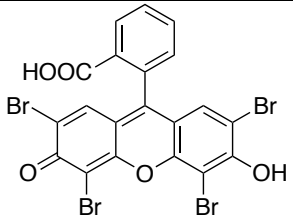
Assay Component	**Volume**	**Final Assay Conditions**
SAM	5 µL	At approximated K_m value: 10 µM
Peptide substrate	5 µL	Close to K_m value: 15 µM
Enzyme	5 µL	Dependent on specific activity
Buffer	5 µL	20 mM Tris, 10 mM DTT, 1% DMSO
TOTAL volume	20 µL	

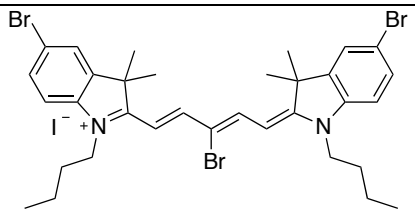
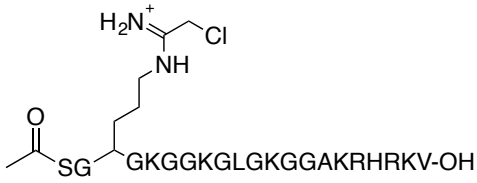
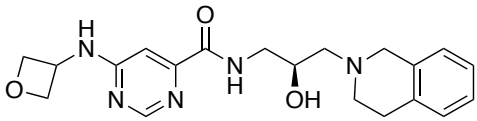
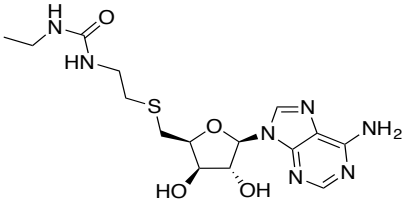
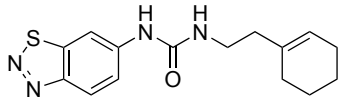
For time-course assays used to monitor altering distributions of peptides, 1 μL samples were removed from the wells at each time point for crystallisation with 1 μL matrix solution without quenching.

Data was collected in an identical manner as for the large scale assay.

Appendix A – PRMT Literature Inhibitors

Summary of prominent inhibitors, mechanism of action and potency as reviewed in **Section 1.5**.

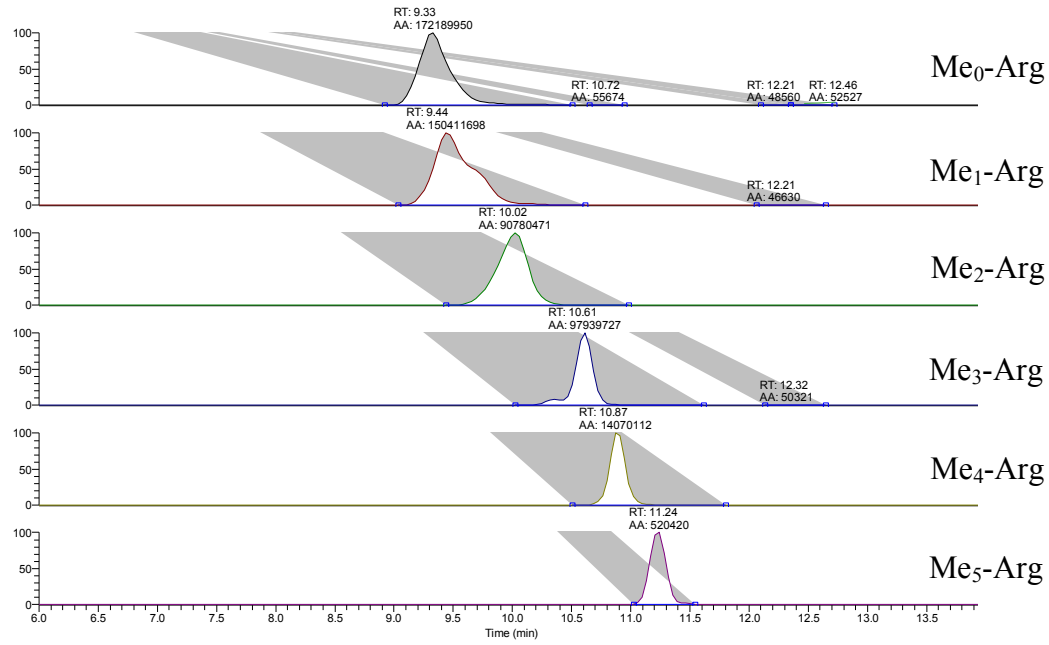
Compound	Proposed mechanism	IC_{50}	Structure
Sinefungin, **32**	Cofactor competitive	1.63 μM *vs* PRMT1	 Sinefungin, **32**
AMI-1, **33**	Cofactor uncompetitive	8.81 μM *vs* PRMT1	 AMI-1, **33**
Allantodapsone, **36**	Cofactor uncompetitive, substrate competitive	1.7 μM *vs* PRMT1	 Allantodapsone, **36**
Stilbamidine, **37**	Cofactor uncompetitive, substrate competitive	56.9 μM *vs* PRMT1	 Stilbamidine, **37**
AMI-composite, **39**	Cofactor uncompetitive	4.15 μM *vs* PRMT1 2.65 μM *vs* CARM1	 **39**
AMI-5 **44**	Thought to bind to SAM pocket (*in silico* docking studies)	1.41 μM *vs* PRMT1	 **44**

46	Allosteric (proposed)	4.1 μM *vs* PRMT1	 **46**
47	SAM and substrate pocket binding	1.8 μM *vs* PRMT1	 **47**
55	Cofactor uncompetitive	22 nM *vs* PRMT5	 **55**
56	Cofactor competitive, substrate uncompetitive	3.38 μM *vs* PRMT1	 **56**
57	Allosteric (confirmed)	2.5 μM *vs* PRMT3	 **57**

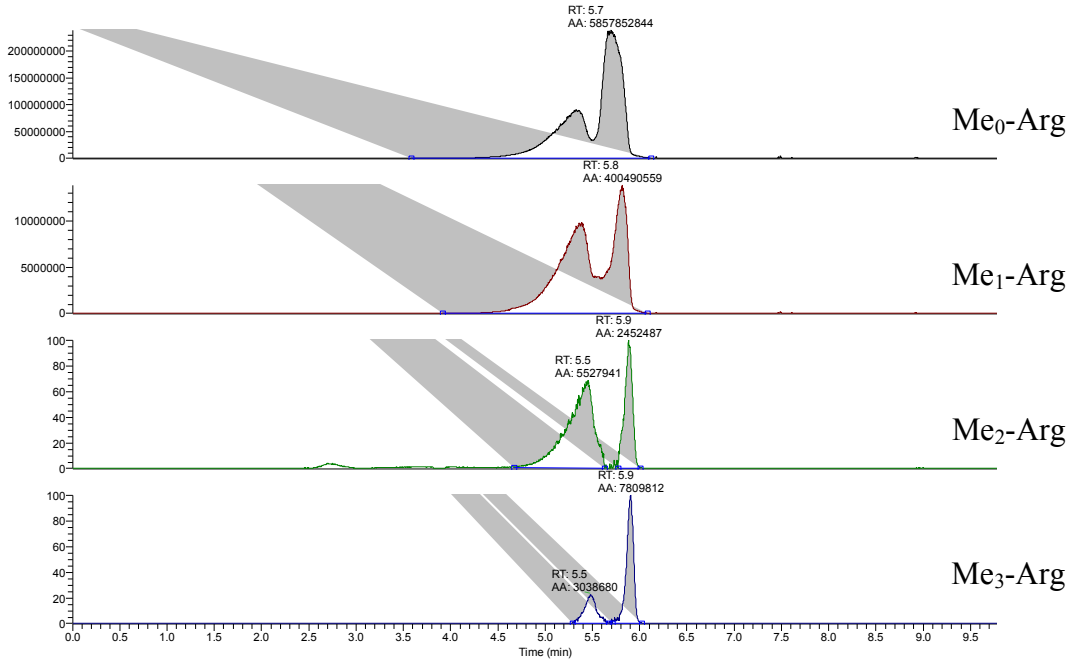
Appendix B – LC/MS traces for Arg/Orn Methylations

These LC/MS traces accompany the data presented in Section 2.3.

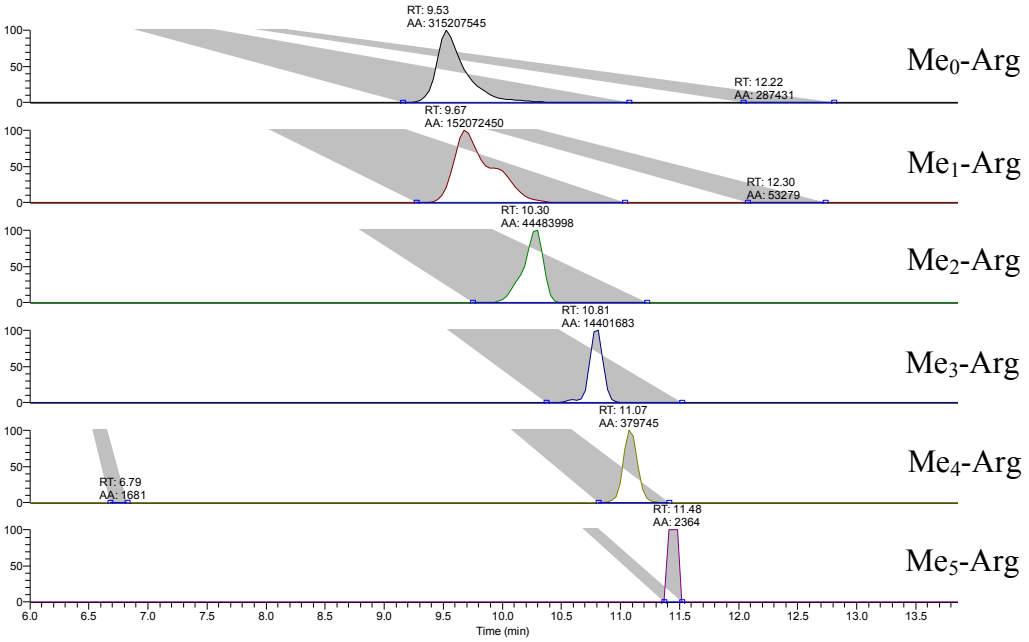
Arg-OH with 1 eq. MeI:



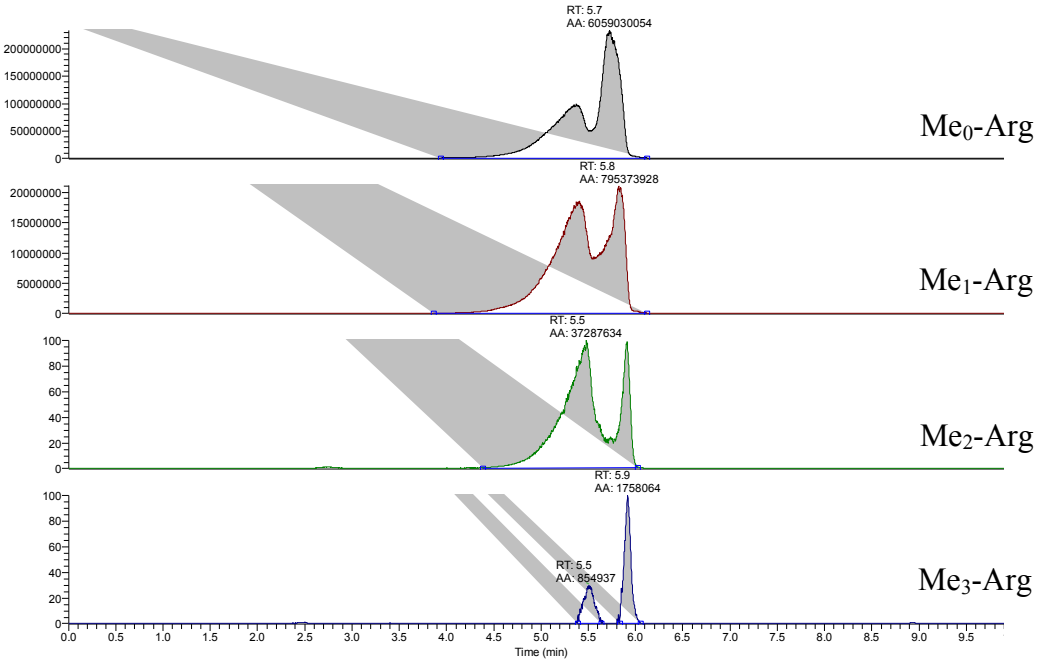
Arg(Boc)-OH with 1 eq. MeI:



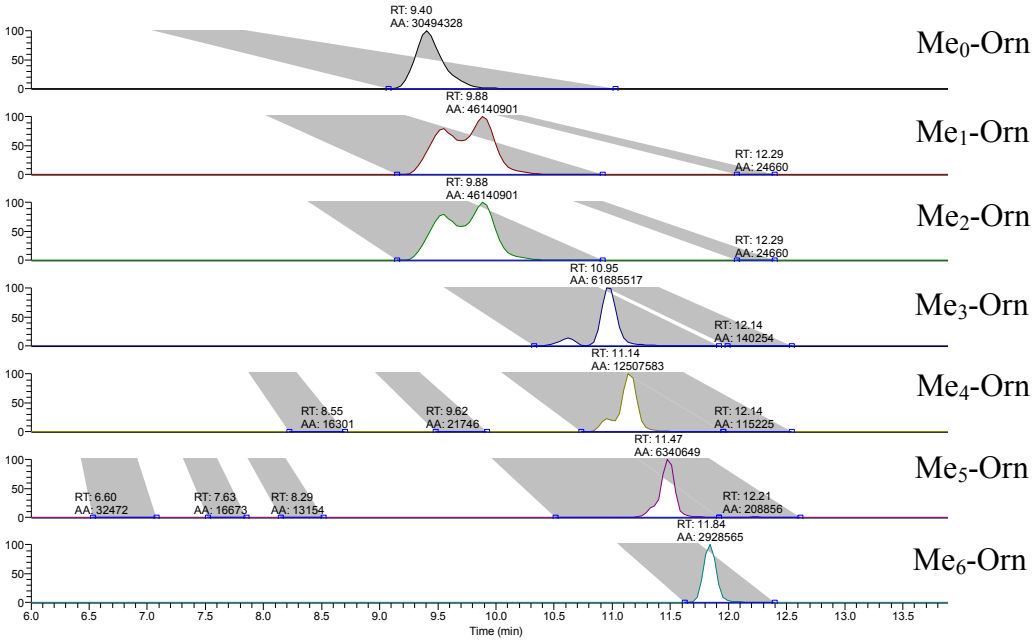
Arg-OH with 0.1 eq. MeI:



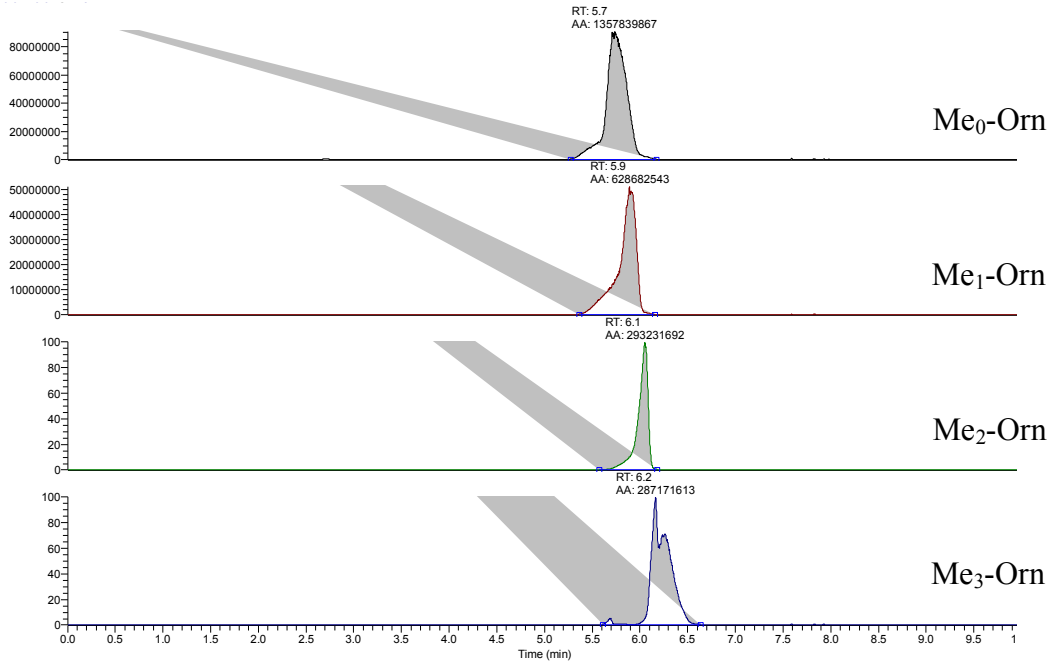
Arg(Boc)-OH with 0.1 eq. MeI:



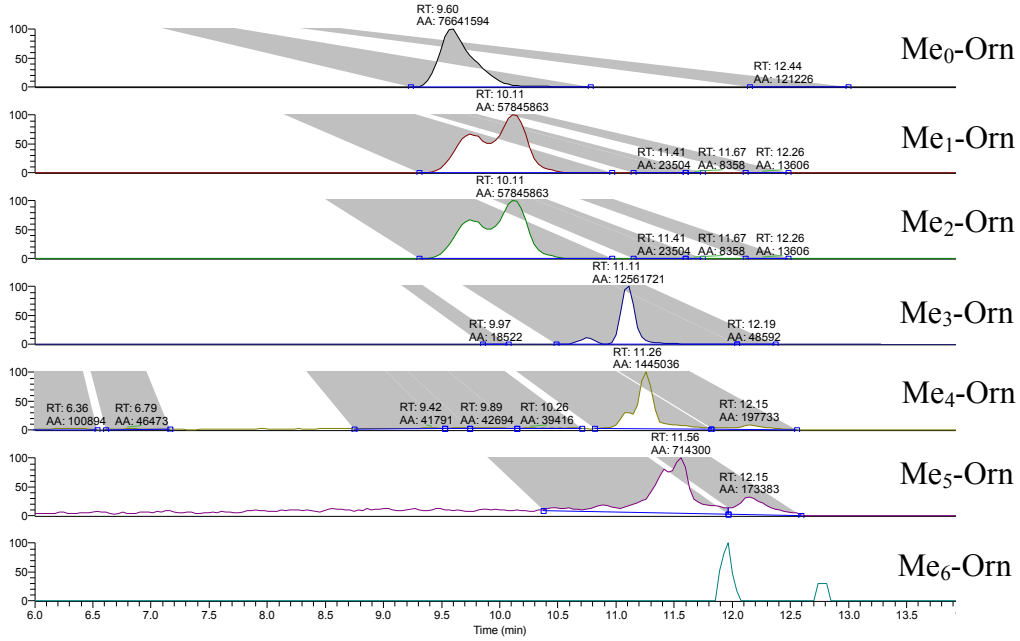
Orn-OH with 1 eq. MeI:



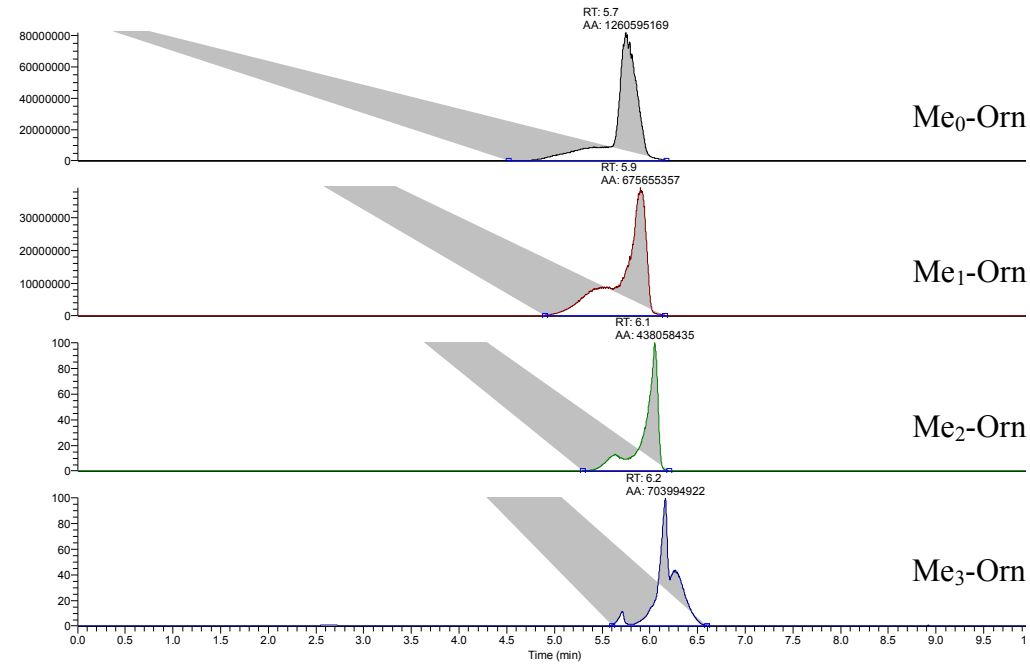
Orn(Boc)-OH with 1 eq. MeI:



Orn-OH with 0.1 eq MeI:

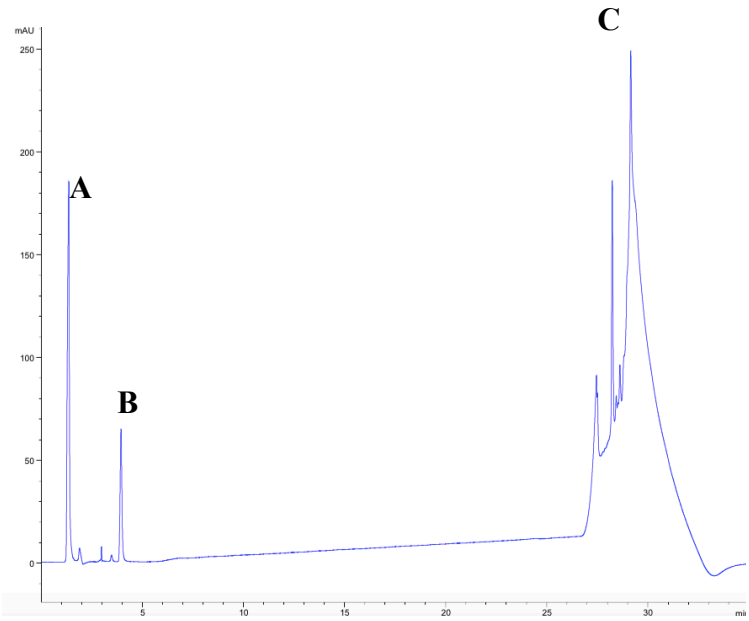


Orn(Boc)-OH with 0.1 ew. MeI:



Appendix C – Peptide HPLC Traces

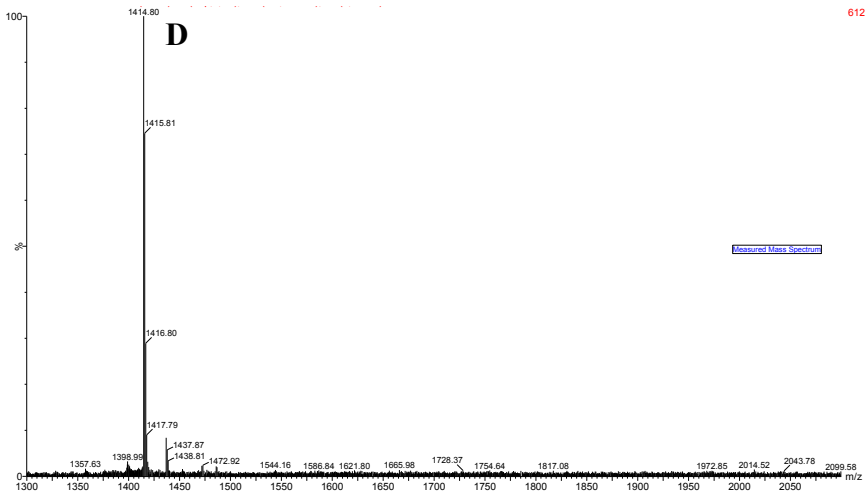
HPLC trace for H4R3 unmethylated peptide



Peptide sequence: NH_2-SGRGKGGKGLGKGGAK-OH

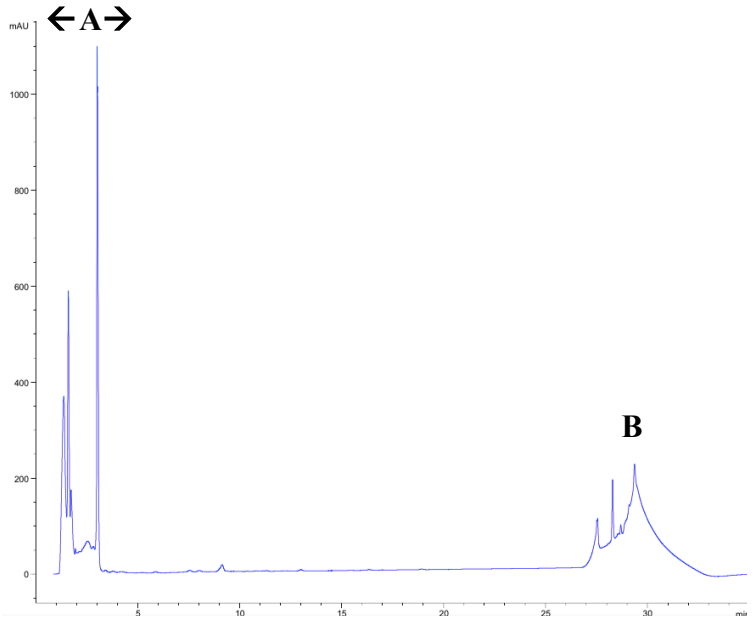
A: LRMS *m/z* 472 (H4R3 16-mer)
B: LRMS *m/z* 121 (Tris base)
C: Column flush

MALDI-TOF-MS trace of this sample



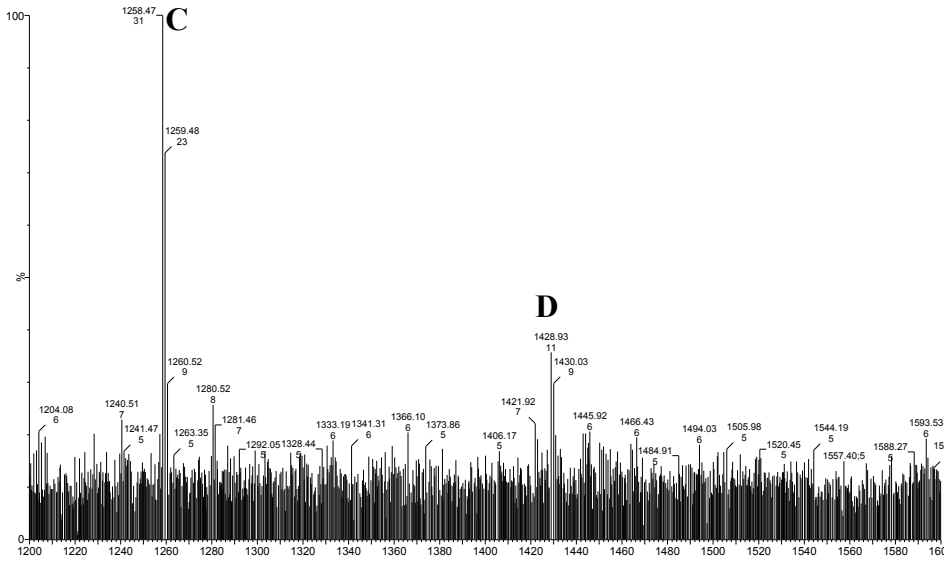
D: 16-mer desired product (1415, M+H)

HPLC trace for MMA(δ)-containing 16-mer fragment peptide of H4



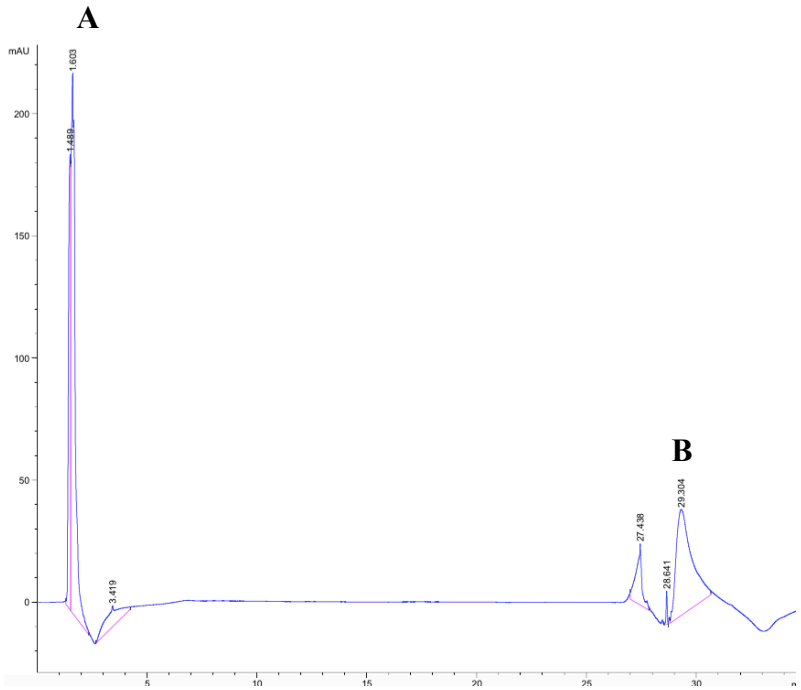
Peptide sequence: NH_2-SGR(δMe)GKGGKGLGKGGAK-OH
A: LRMS *m/z* $(ESI)^+$ 420 (15-mer deletion product)
B: Column flush

MALDI-TOF-MS trace of this sample



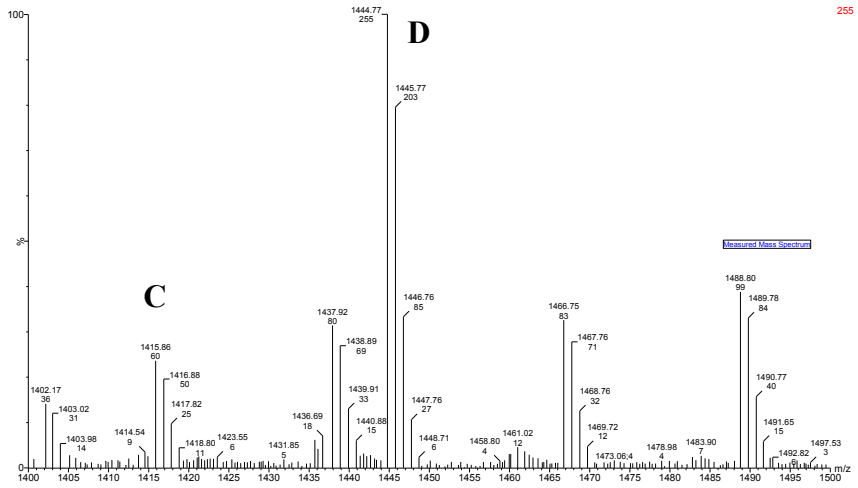
C: 15-mer deletion product (1258, M+H)
D: 16-mer desired product (1429, M+H)

HPLC trace for SDMA-containing 16-mer fragment peptide of H4



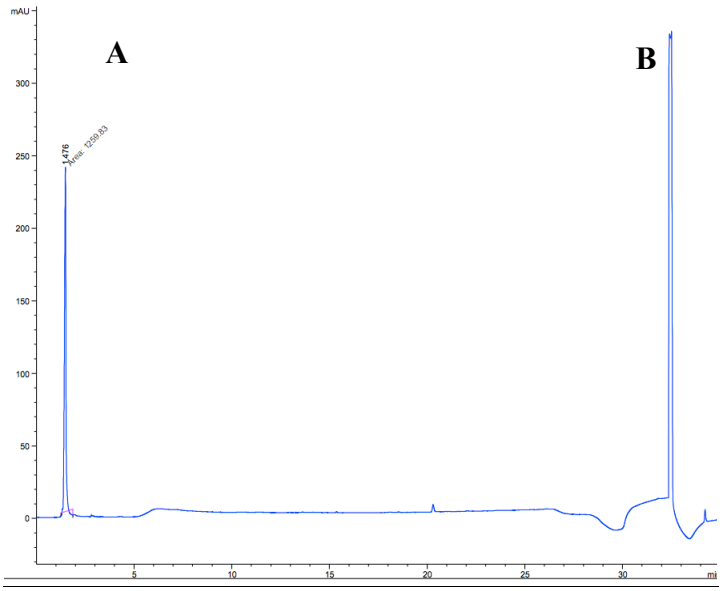
Peptide sequence: NH_2-SGR(*Sym*-Me_2)GKGGKGLGKGGAK-OH
A: LRMS *m/z* 481 (H4R3(SDMA) 16-mer)
B: Column flush

MALDI-TOF-MS trace of this sample



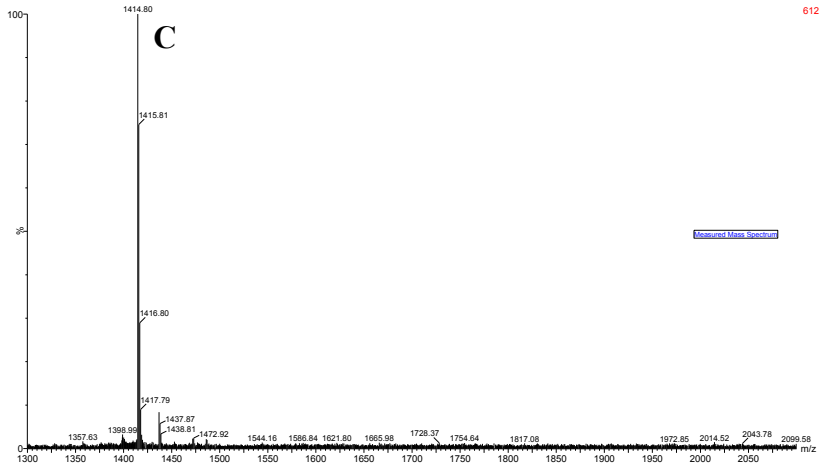
C: 16-mer of putative unmethylated peptide (1415, M+H)
D: 16-mer desired product (1444, M+H)

HPLC trace for the 15-mer deletion product generated as part of the synthesis of the 16-mer peptide containing MMA(δ)



Peptide sequence: NH_2-SGGKGGKGLGKGGAK-OH
A: LRMS *m/z* 629 (15-mer deletion product)
B: Column flush

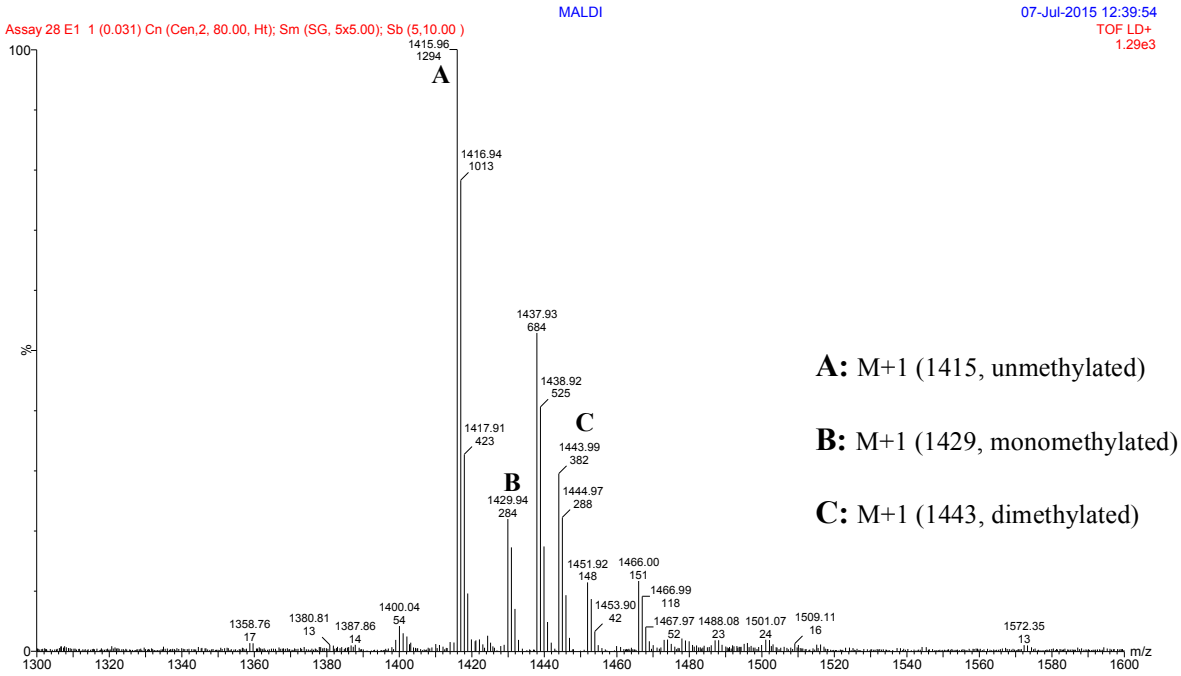
MALDI-TOF-MS trace of this sample



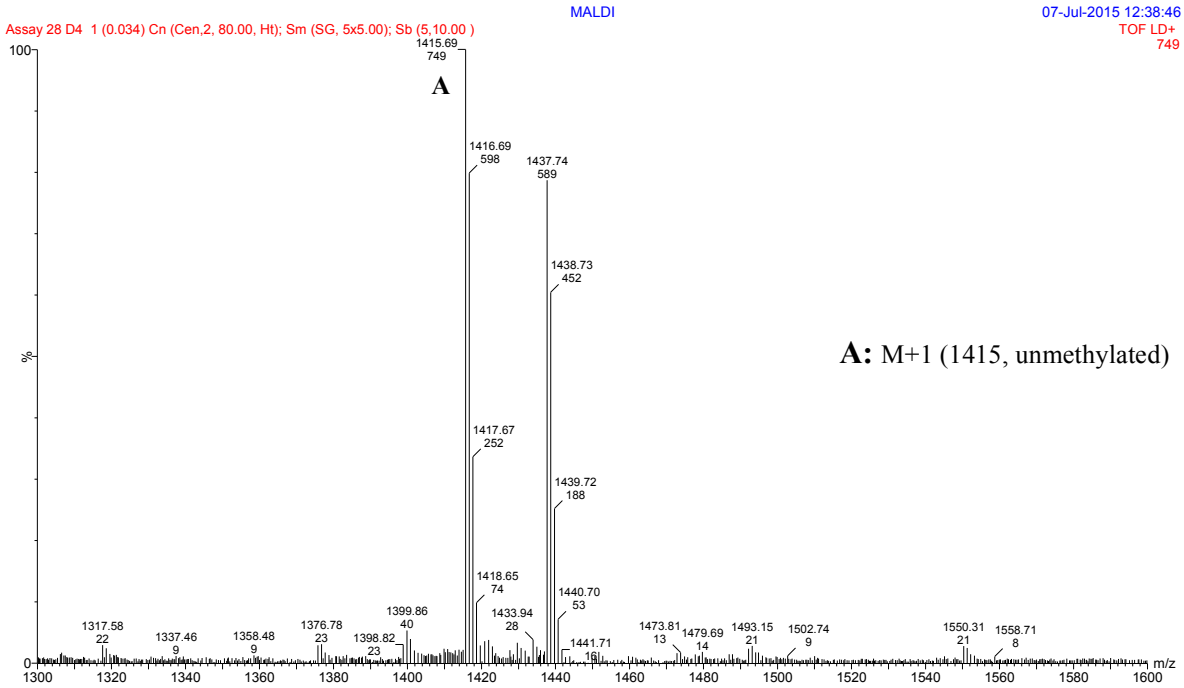
C: 15-mer deletion product peptide (1258, M+H)

Appendix D – Exemplar MALDI Assay Traces

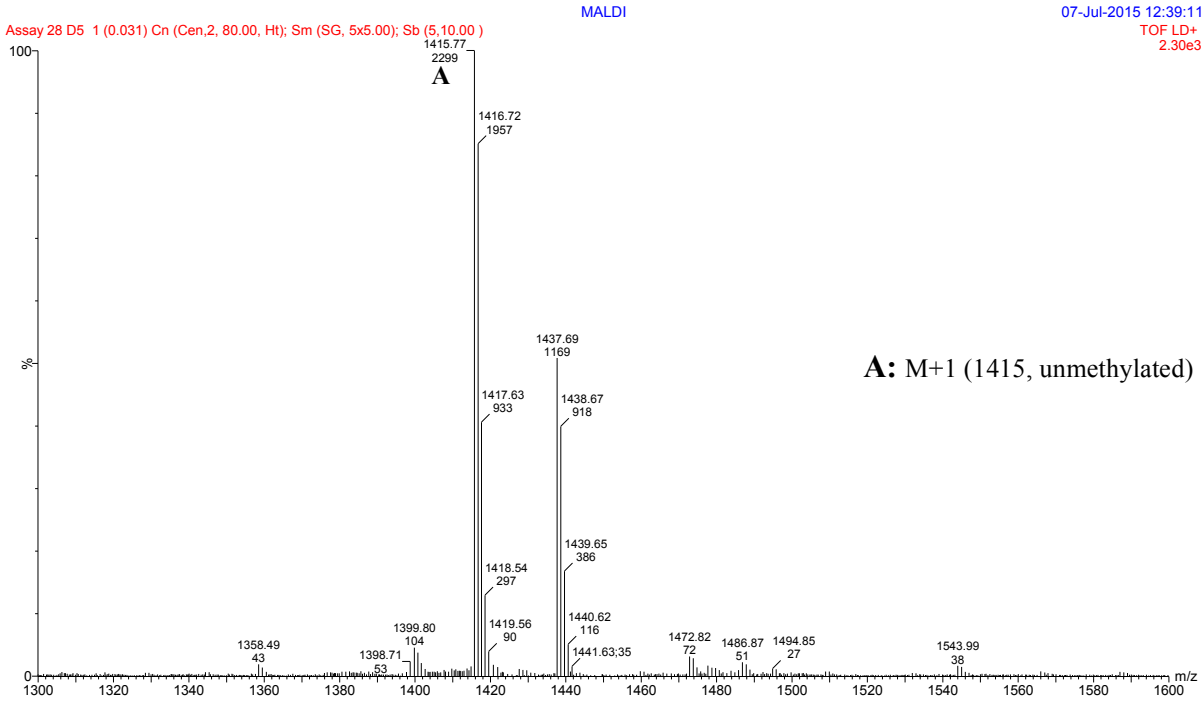
Negative control using no inhibitor (expected full enzyme activity):



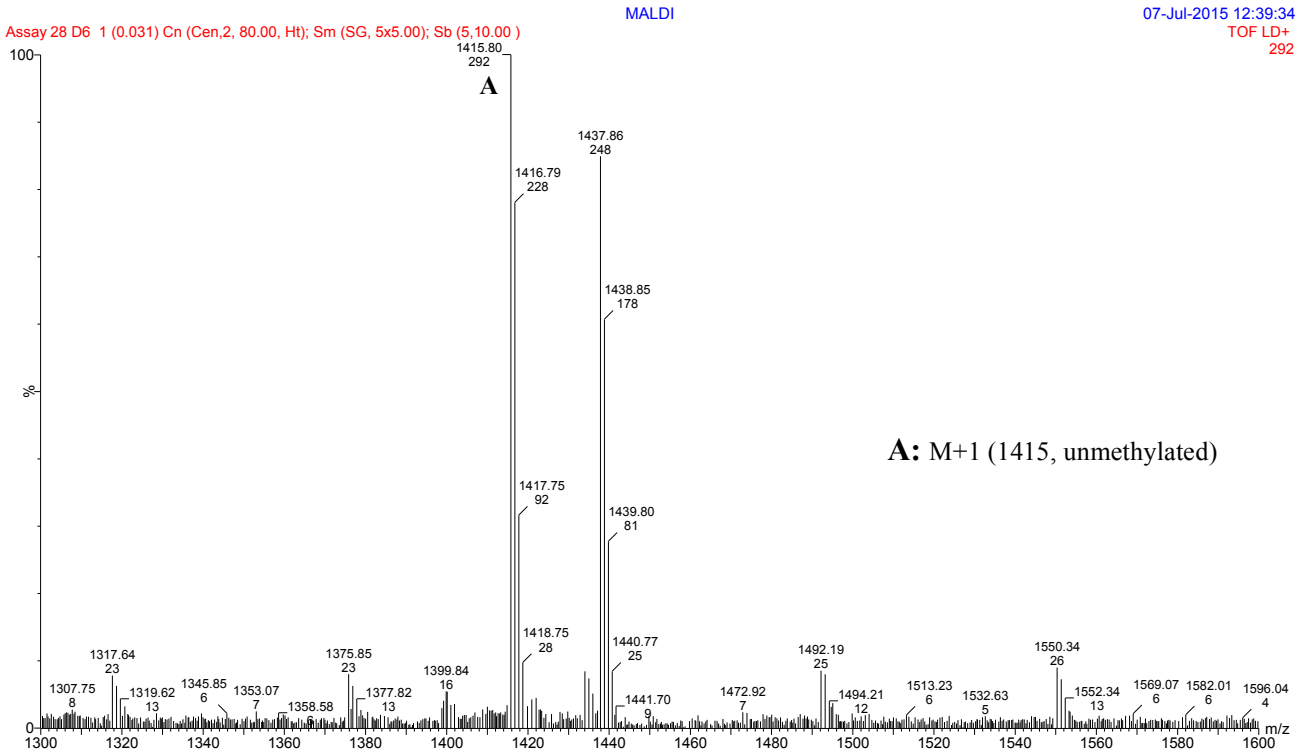
Positive control using Sinefungin **32** as an inhibitor:



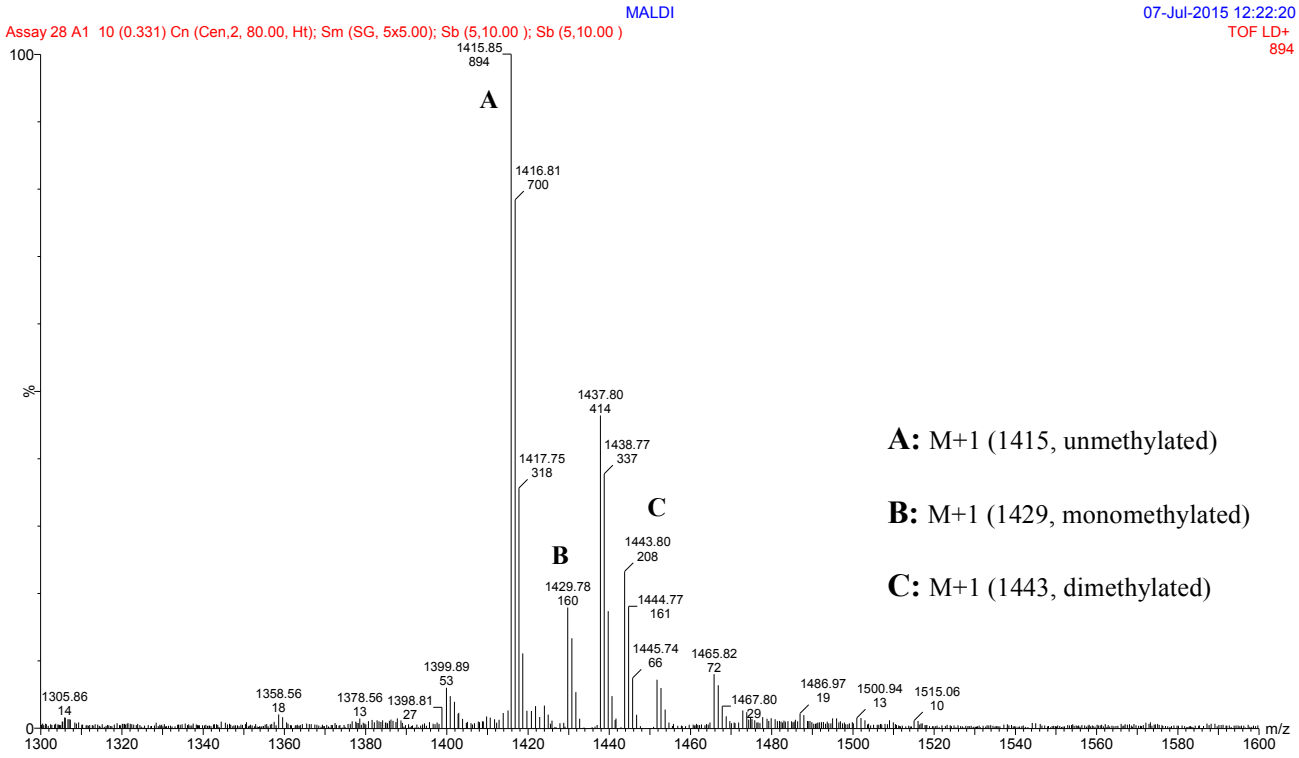
Control using no SAM **2**:



Control using no enzyme:



Example trace testing compound **206** as an inhibitor of PRMT:



Appendix E – Total MALDI Ion Counts

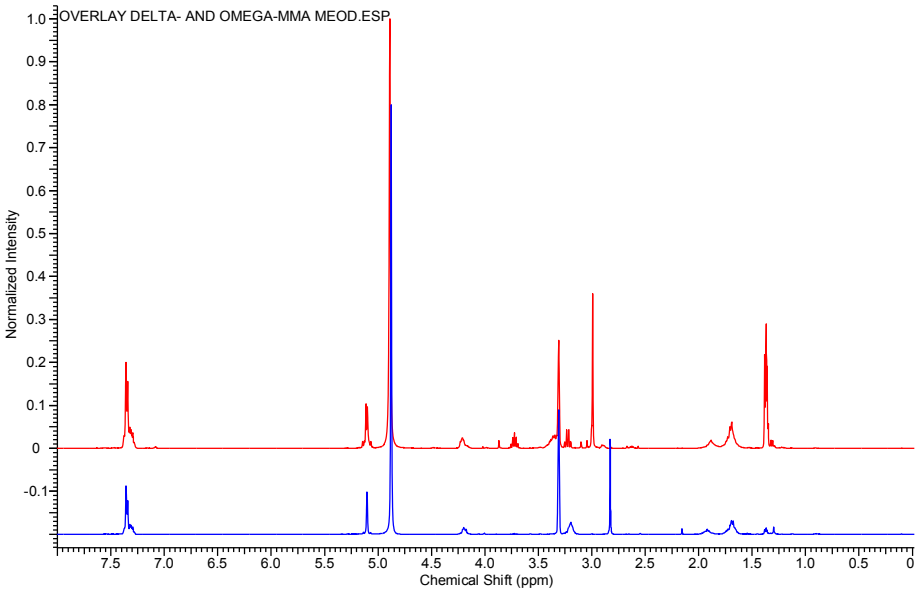
Total ion counts for each experiment recorded in **Figure 4.11, Figure 4.12** and **Figure 4.13**

		1:1 Me0	1:1 MMA(ω)	1:1 SDMA	1:1 ADMA	1:1 MMA(δ)
t = 0 h	Repeat 1	246	139	0	285	60
	Repeat 2	144	81	75	169	24
	Repeat 3	158	307	0	225	61
t = 1.5 h	Repeat 1	276	85	216	373	120
	Repeat 2	225	140	275	140	101
	Repeat 3	310	55	104	115	91
t = 3 h	Repeat 1	315	180	239	490	90
	Repeat 2	759	467	29	131	58
	Repeat 3	490	363	139	45	62
t = 4.5 h	Repeat 1	228	66	81	46	39
	Repeat 2	225	126	382	230	207
	Repeat 3	683	357	278	144	42
t = 6 h	Repeat 1	113	149	130	545	73
	Repeat 2	115	592	260	1874	183
	Repeat 3	774	419	585	247	122
t = 21 h	Repeat 1	408	245	300	163	25
	Repeat 2	213	194	378	487	61
	Repeat 3	315	578	499	229	41

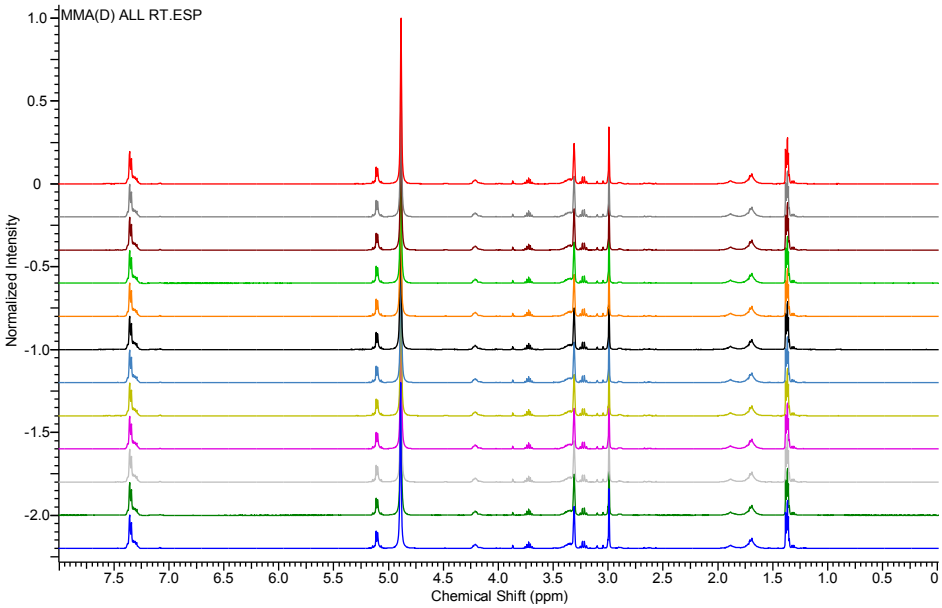
		1:4 Me0	1:4 MMA(ω)	1:4 SDMA	1:4 ADMA	1:4 MMA(δ)
t = 0 h	Repeat 1	61	30	80	152	38
	Repeat 2	67	9	87	38	26
	Repeat 3	72	47	31	35	38
t = 1.5 h	Repeat 1	143	187	189	214	15
	Repeat 2	61	120	56	99	14
	Repeat 3	51	208	107	99	10
t = 3 h	Repeat 1	125	331	230	201	0
	Repeat 2	157	401	187	183	44
	Repeat 3	102	328	101	312	34
t = 4.5 h	Repeat 1	84	66	100	553	30
	Repeat 2	41	209	106	854	33
	Repeat 3	233	189	258	282	31
t = 6 h	Repeat 1	134	224	77	719	44
	Repeat 2	84	205	96	914	100
	Repeat 3	68	371	536	760	22
t = 21 h	Repeat 1	50	234	85	402	32
	Repeat 2	43	1490	71	23	28
	Repeat 3	78	448	92	361	38

Appendix F – NMR Rearrangement Studies

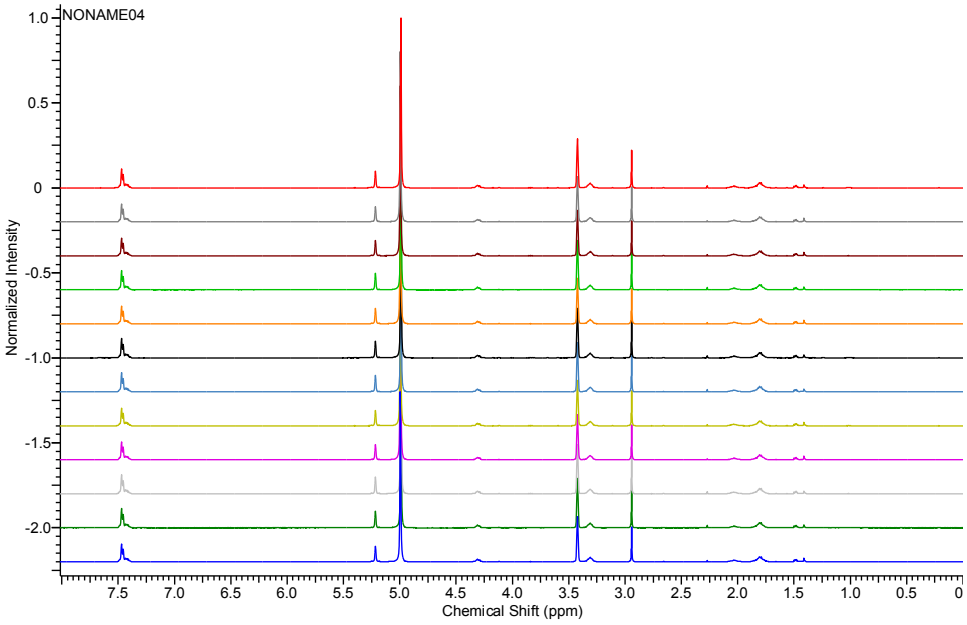
Overlay of *N*-Me(δ)-Arg(Cbz)-OH **189** and *N*-Me(ω)-Arg(Cbz)-OH **116** in CD_3OD at RT



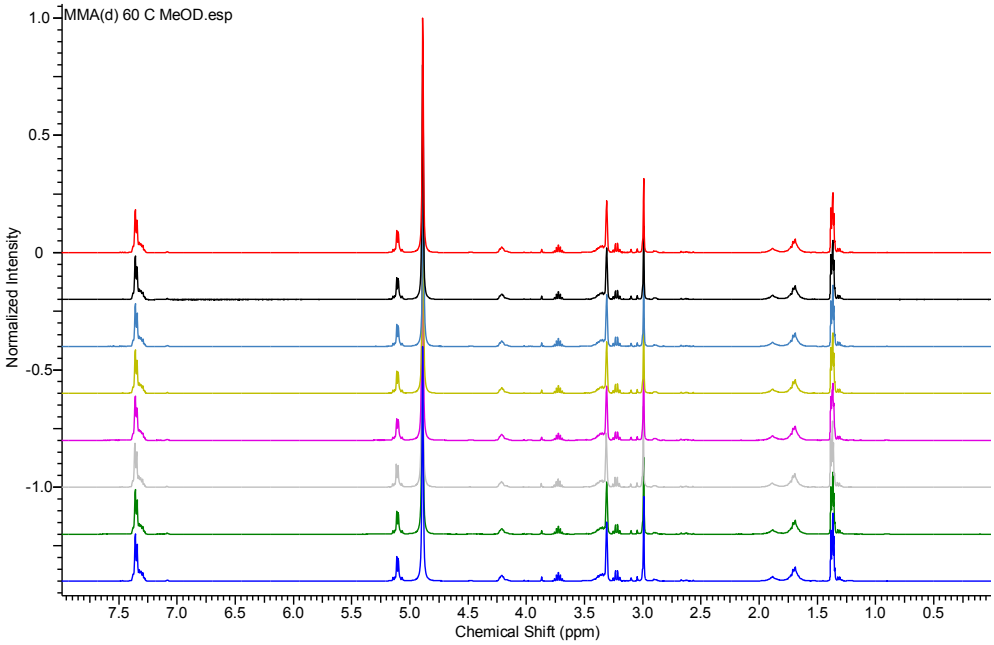
12-hour time course NMR of *N*-Me(δ)-Arg(Cbz)-OH **189** in CD_3OD at RT



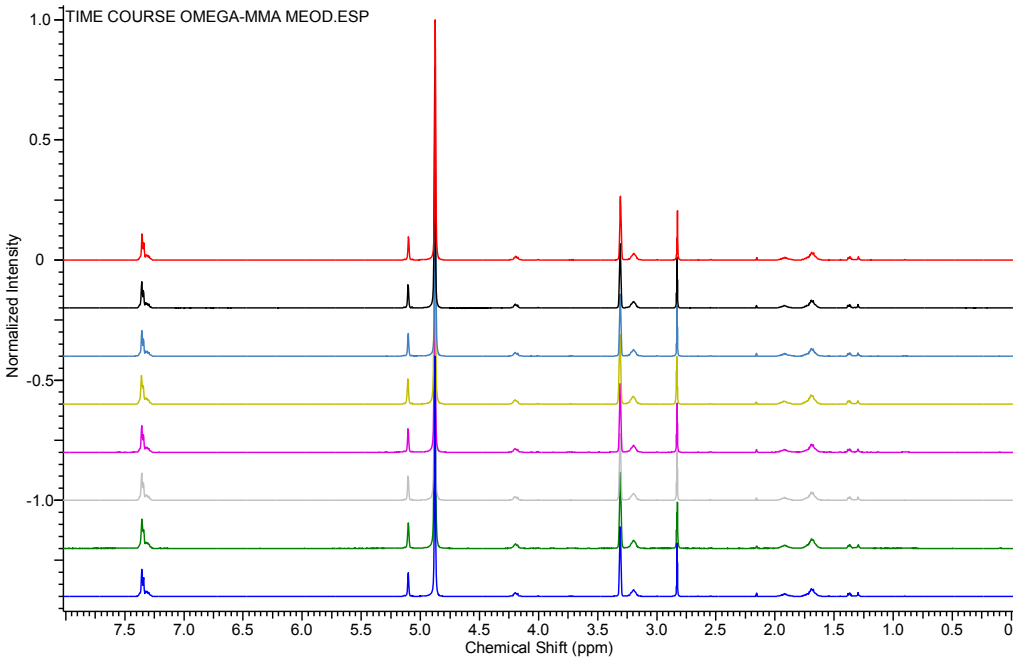
12-hour time course NMR of *N*-Me(ω)-Arg(Cbz)-OH **116** in CD_3OD at RT



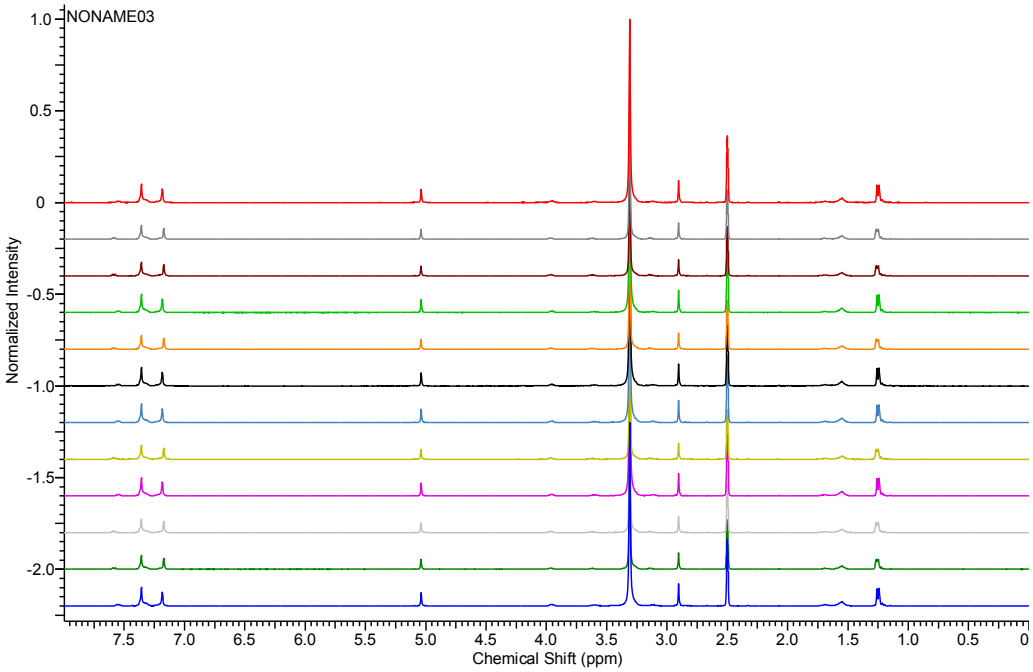
7-hour time course NMR of *N*-Me(δ)-Arg(Cbz)-OH **189** in CD_3OD at 60 °C



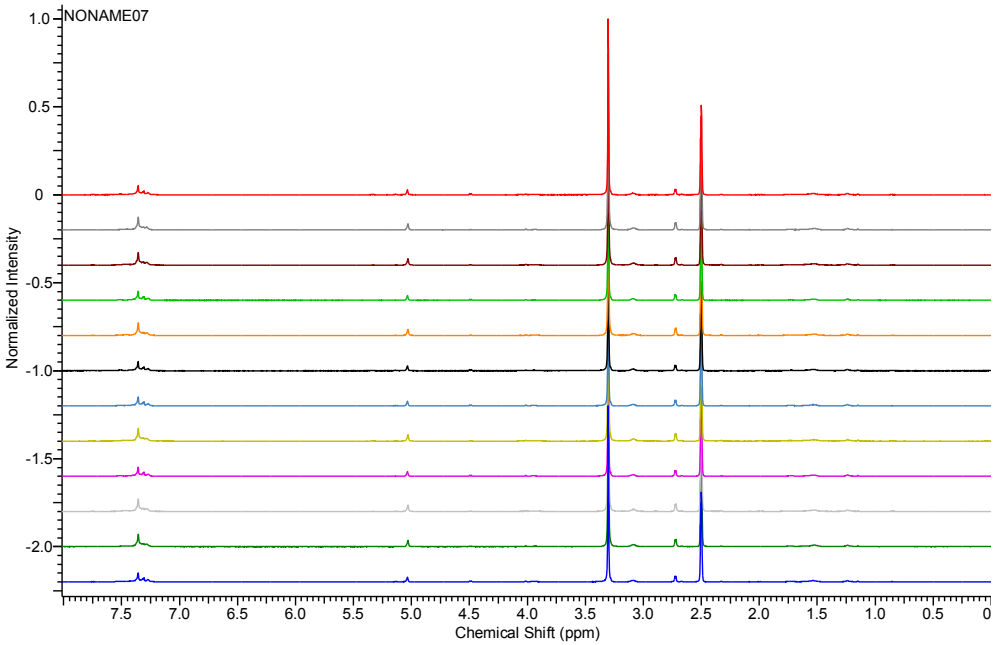
7-hour time course NMR of *N*-Me(ω)-Arg(Cbz)-OH **116** in CD_3OD at 60 °C



12-hour time course NMR of *N*-Me(δ)-Arg(Cbz)-OH **189** in DMSO-d_6 at RT



12-hour time course NMR of *N*-Me(ω)-Arg(Cbz)-OH **116** in DMSO-d_6 at RT



Appendix G – Crystallographic Information for 233

collect in F:\Data\av2203\impor At 15: 9 on 7/ 9

Crystal Data

a = 8.2239(4) Å	α = 100.730(2)°		
b = 8.2328(4) Å	β = 101.721(2)°		
c = 10.2779(5) Å	γ = 116.240(2)°		
Volume	580.19(5) Å^3	Crystal Class	triclinic
Space group	P -1	Z =	2
Formula	C_{14} H_{13} N_1 O_2 S_1	M_r	259.33
Cell determined from	1934 reflections	Cell θ range =	5 - 27°
Temperature	150K		
Pressure	101 kPa	Shape	block
Colour	colourless	Size	0.10 × 0.10 × 0.20 mm
D_x	1.48 Mg m^{-3}	F000	272.000
μ	0.271 mm^{-1}		
Absorption correction	multi-scan		
T_{min}	0.97	T_{max}	0.97

Data Collection

Diffractometer	multi-scan
Scan type	ω scans
Reflections measured	4043
Independent reflections	2471
Rint	0.0337
θ_{max}	27.5781
h =	-10 → 10
k =	-10 → 10
l =	-13 → 13

Refinement

$\Delta\varrho_{min}$ =	-0.42 e Å^{-3}
$\Delta\varrho_{max}$ =	0.37 e Å^{-3}
Reflections used	2471
Cutoff: I >	-10.00σ(I)
Parameters refined	166
S =	0.94
R-factor	0.064
weighted R-factor	0.107
Δ/σ_{max}	0.0006
Refinement on	F^2
w =	$1/[\sigma^2(F_{obs}^2) + (0.025 \times P)^2 + 0.356 \times P + 0.000 + 0.000 \times \sin\theta]$, $P = 0.333 \times \max(F_{obs}^2,0) + 0.667 \times F_{calc}^2$

Parameters

Label	*x*	*y*	*z*	$U_{iso/equiv}$	Occupancy
S1	0.87587(7)	0.19442(8)	0.34938(5)	0.0298	1.0000
O2	0.9250(2)	0.1090(2)	0.23056(14)	0.0339	1.0000
C3	0.7190(3)	0.2696(3)	0.2777(2)	0.0268	1.0000
C4	0.5282(3)	0.1373(3)	0.1944(2)	0.0254	1.0000
N5	0.4449(2)	-0.0538(2)	0.18828(17)	0.0272	1.0000
C6	0.2694(3)	-0.1953(3)	0.0715(2)	0.0286	1.0000
C7	0.0806(3)	-0.2210(3)	0.0887(2)	0.0322	1.0000
O8	-0.0658(2)	-0.3206(2)	0.04431(16)	0.0385	1.0000
C9	0.5144(3)	-0.1087(3)	0.2972(2)	0.0273	1.0000
C10	0.7047(3)	0.0054(3)	0.3886(2)	0.0284	1.0000
C11	0.7676(3)	-0.0388(3)	0.5073(2)	0.0321	1.0000
C12	0.6464(3)	-0.1988(3)	0.5341(2)	0.0356	1.0000
C13	0.4619(3)	-0.3177(3)	0.4402(2)	0.0349	1.0000
C14	0.3956(3)	-0.2760(3)	0.3232(2)	0.0312	1.0000
C15	0.4242(3)	0.2067(3)	0.1193(2)	0.0302	1.0000
C16	0.5038(3)	0.3982(3)	0.1351(2)	0.0333	1.0000
C17	0.6901(3)	0.5297(3)	0.2219(2)	0.0333	1.0000
C18	0.7977(3)	0.4639(3)	0.2912(2)	0.0308	1.0000
H61	0.2661(3)	-0.3183(3)	0.0564(2)	0.0339	1.0000
H62	0.2806(3)	-0.1503(3)	-0.0097(2)	0.0337	1.0000
H71	0.0449(3)	-0.2983(3)	0.1490(2)	0.0373	1.0000
H72	0.0949(3)	-0.0959(3)	0.1289(2)	0.0387	1.0000
H111	0.8945(3)	0.0430(3)	0.5683(2)	0.0380	1.0000
H121	0.6898(3)	-0.2265(3)	0.6149(2)	0.0425	1.0000
H131	0.3779(3)	-0.4310(3)	0.4571(2)	0.0421	1.0000
H141	0.2698(3)	-0.3575(3)	0.2609(2)	0.0363	1.0000
H161	0.4291(3)	0.4424(3)	0.0860(2)	0.0407	1.0000
H171	0.7399(3)	0.6610(3)	0.2318(2)	0.0408	1.0000
H181	0.9272(3)	0.5483(3)	0.3478(2)	0.0360	1.0000
H151	0.2956(3)	0.1189(3)	0.0578(2)	0.0355	1.0000
H81	-0.028(4)	-0.256(4)	-0.094(3)	0.0581	1.0000

Thermal Parameters

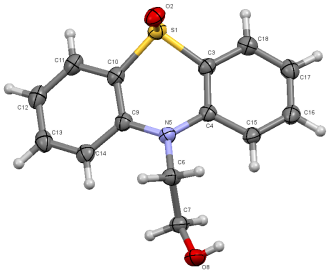
Label	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
S1	0.0249(3)	0.0341(3)	0.0278(3)	0.0089(2)	0.0048(2)	0.0145(2)
O2	0.0320(8)	0.0446(9)	0.0323(8)	0.0130(7)	0.0129(6)	0.0235(7)
C3	0.0269(10)	0.0309(11)	0.0233(9)	0.0083(9)	0.0073(8)	0.0154(9)
C4	0.0260(10)	0.0286(11)	0.0250(9)	0.0093(9)	0.0111(8)	0.0149(9)
N5	0.0231(8)	0.0289(9)	0.0274(8)	0.0108(7)	0.0059(7)	0.0113(7)
C6	0.0244(10)	0.0316(11)	0.0265(10)	0.0078(9)	0.0075(8)	0.0120(9)
C7	0.0245(10)	0.0377(12)	0.0323(11)	0.0132(10)	0.0087(9)	0.0132(9)
O8	0.0237(8)	0.0414(10)	0.0400(9)	0.0152(7)	0.0039(7)	0.0095(7)
C9	0.0297(10)	0.0313(11)	0.0271(10)	0.0109(9)	0.0119(9)	0.0184(9)
C10	0.0303(10)	0.0317(11)	0.0270(10)	0.0094(9)	0.0091(9)	0.0185(9)
C11	0.0364(11)	0.0383(12)	0.0276(10)	0.0095(9)	0.0084(9)	0.0247(10)
C12	0.0468(13)	0.0460(14)	0.0314(11)	0.0184(10)	0.0172(10)	0.0331(12)
C13	0.0449(13)	0.0351(12)	0.0410(12)	0.0196(10)	0.0236(11)	0.0264(11)
C14	0.0326(11)	0.0324(12)	0.0348(11)	0.0115(9)	0.0137(9)	0.0198(10)
C15	0.0244(10)	0.0385(12)	0.0321(11)	0.0166(10)	0.0109(9)	0.0165(9)
C16	0.0344(11)	0.0422(13)	0.0387(12)	0.0220(10)	0.0198(10)	0.0248(10)
C17	0.0378(12)	0.0319(12)	0.0382(11)	0.0144(10)	0.0194(10)	0.0195(10)
C18	0.0302(11)	0.0302(11)	0.0295(10)	0.0072(9)	0.0105(9)	0.0134(9)

Distances

S1	O2	1.5104(15)Å		S1	C3	1.7550(19)Å
S1	C10	1.759(2)Å		C3	C4	1.406(3)Å
C3	C18	1.401(3)Å		C4	N5	1.393(2)Å
C4	C15	1.409(3)Å		N5	C6	1.473(2)Å
N5	C9	1.393(2)Å		C6	C7	1.524(3)Å
C6	H61	0.983Å		C6	H62	0.979Å
C7	O8	1.425(2)Å		C7	H71	0.965Å
C7	H72	0.980Å		O8	H81	0.814(16)Å
C9	C10	1.410(3)Å		C9	C14	1.406(3)Å
C10	C11	1.402(3)Å		C11	C12	1.375(3)Å
C11	H111	0.944Å		C12	C13	1.390(3)Å
C12	H121	0.945Å		C13	C14	1.383(3)Å
C13	H131	0.957Å		C14	H141	0.942Å
C15	C16	1.374(3)Å		C15	H151	0.960Å
C16	C17	1.391(3)Å		C16	H161	0.950Å
C17	C18	1.376(3)Å		C17	H171	0.947Å
C18	H181	0.948Å				

Angles

O2	S1	C3	107.20(9)°		O2	S1	C10	107.74(9)°
C3	S1	C10	95.80(9)°		S1	C3	C4	121.20(15)°
S1	C3	C18	117.24(15)°		C4	C3	C18	121.04(18)°
C3	C4	N5	121.35(17)°		C3	C4	C15	117.18(18)°
N5	C4	C15	121.46(18)°		C4	N5	C6	118.83(15)°
C4	N5	C9	120.85(16)°		C6	N5	C9	120.04(16)°
N5	C6	C7	116.58(17)°		N5	C6	H61	107.388°
C7	C6	H61	109.096°		N5	C6	H62	106.128°
C7	C6	H62	106.263°		H61	C6	H62	111.395°
C6	C7	O8	109.10(17)°		C6	C7	H71	109.303°
O8	C7	H71	107.552°		C6	C7	H72	109.530°
O8	C7	H72	111.050°		H71	C7	H72	110.265°
C7	O8	H81	105(2)°		N5	C9	C10	121.14(17)°
N5	C9	C14	121.29(18)°		C10	C9	C14	117.53(17)°
S1	C10	C9	121.14(14)°		S1	C10	C11	117.73(16)°
C9	C10	C11	120.90(18)°		C10	C11	C12	120.4(2)°
C10	C11	H111	118.853°		C12	C11	H111	120.700°
C11	C12	C13	118.93(19)°		C11	C12	H121	119.738°
C13	C12	H121	121.326°		C12	C13	C14	121.7(2)°
C12	C13	H131	119.262°		C14	C13	H131	119.028°
C9	C14	C13	120.3(2)°		C9	C14	H141	118.974°
C13	C14	H141	120.725°		C4	C15	C16	120.64(19)°
C4	C15	H151	119.241°		C16	C15	H151	120.108°
C15	C16	C17	121.88(19)°		C15	C16	H161	119.148°
C17	C16	H161	118.971°		C16	C17	C18	118.5(2)°
C16	C17	H171	119.726°		C18	C17	H171	121.800°
C3	C18	C17	120.6(2)°		C3	C18	H181	118.390°
C17	C18	H181	120.962°					



References

(1) Heightman, T. D. *Expert Opinion on Therapeutic Targets* **2011**, *15*, 729.

(2) DeWoskin, V. A.; Million, R. P. *Nature Reviews Drug Discovery* **2013**, *12*, 661.

(3) Skipper, M.; Eccleston, A.; Gray, N.; Heemels, T.; Le Bot, N.; Marte, B.; Weiss, U. *Nature* **2015**, *518*, 313.

(4) Waddington, C. H. *The Strategy of the Genes: a Discussion of Some Aspects of Theoretical Biology*; Allen & Unwin, London, 1957.

(5) Kornberg, R. D. *Science* **1974**, *184*, 868.

(6) Kornberg, R. D.; Lorch, Y. *Cell* **1999**, *98*, 285.

(7) He, C. *Nature Chemical Biology* **2010**, *6*, 863.

(8) Geula, S.; Moshitch-Moshkovitz, S.; Dominissini, D.; Mansour, A. A.; Kol, N.; Salmon-Divon, M.; Hershkovitz, V.; Peer, E.; Mor, N.; Manor, Y. S.; Ben-Haim, M. S.; Eyal, E.; Yunger, S.; Pinto, Y.; Jaitin, D. A.; Viukov, S.; Rais, Y.; Krupalnik, V.; Chomsky, E.; Zerbib, M.; Maza, I.; Rechavi, Y.; Massarwa, R.; Hanna, S.; Amit, I.; Levanon, E. Y.; Amariglio, N.; Stern-Ginossar, N.; Novershtern, N.; Rechavi, G.; Hanna, J. H. *Science* **2015**, *347*, 1002.

(9) Phizicky, E. M.; Hopper, A. K. *Rna* **2015**, *21*, 483.

(10) Gilbert, W. V. *Trends in Biochemical Sciences* **2011**, *36*, 127.

(11) Liu, W.; Fan, L. X.; Zhou, X.; Sweeney, W. E., Jr.; Avner, E. D.; Li, X. *PLoS One* **2012**, *7*, e49418.

(12) Nature; Nature Publishing Group: http://www.nature.com/collections/vbqgtr/?; Vol. 2015, p Describes the main findings of the NIH Roadmap Epigenomics Program.

(13) Roadmap Epigenomics, C.; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; Ziller, M. J.; Amin, V.; Whitaker, J. W.; Schultz, M. D.; Ward, L. D.; Sarkar, A.; Quon, G.; Sandstrom, R. S.; Eaton, M. L.; Wu, Y.-C.; Pfenning, A. R.; Wang, X.; Claussnitzer, M.; Liu, Y.; Coarfa, C.; Harris, R. A.; Shoresh, N.; Epstein, C. B.; Gjoneska, E.; Leung, D.; Xie, W.; Hawkins, R. D.; Lister, R.; Hong, C.; Gascard, P.; Mungall, A. J.; Moore, R.; Chuah, E.; Tam, A.; Canfield, T. K.; Hansen, R. S.; Kaul, R.; Sabo, P. J.; Bansal, M. S.; Carles, A.; Dixon, J. R.; Farh, K.-H.; Feizi, S.; Karlic, R.; Kim, A.-R.; Kulkarni, A.; Li, D.; Lowdon, R.; Elliott, G.; Mercer, T. R.; Neph, S. J.; Onuchic, V.; Polak, P.; Rajagopal, N.; Ray, P.; Sallari, R. C.; Siebenthall, K. T.; Sinnott-Armstrong, N. A.; Stevens, M.; Thurman, R. E.; Wu, J.; Zhang, B.; Zhou, X.; Beaudet, A. E.; Boyer, L. A.; De Jager, P. L.; Farnham, P. J.; Fisher, S. J.; Haussler, D.; Jones, S. J. M.; Li, W.; Marra, M. A.; McManus, M. T.; Sunyaev, S.; Thomson, J. A.; Tlsty, T. D.; Tsai, L.-H.; Wang, W.; Waterland, R. A.; Zhang, M. Q.; Chadwick, L. H.; Bernstein, B. E.; Costello, J. F.; Ecker, J. R.; Hirst, M.;

Meissner, A.; Milosavljevic, A.; Ren, B.; Stamatoyannopoulos, J. A.; Wang, T.; Kellis, M.; Roadmap Epigenomics, C. *Nature* **2015**, *518*, 317.

(14) Chi, P.; Allis, C. D.; Wang, G. G. *Nature Reviews Cancer* **2010**, *10*, 457.

(15) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. *Nature* **2010**, *468*, 1067.

(16) Marks, P. A.; Breslow, R. *Nature biotechnology* **2007**, *25*, 84.

(17) Zwergel, C.; Valente, S.; Jacob, C.; Mai, A. *Expert opinion on drug discovery* **2015**, *10*, 599.

(18) Katz, J. E.; Dlakic, M.; Clarke, S. *Molecular Cell Proteomics* **2003**, *2*, 525.

(19) Paik, W. K.; Kim, S. *Biochemistry and Biophysics Research Communications* **1967**, *29*, 14.

(20) Murray, K. *Biochemistry* **1964**, *3*, 10.

(21) Friedman, M.; Shull, K. H.; Farber, E. *Biochemical and biophysical research communications* **1969**, *34*, 857.

(22) Kaye, A. M.; Sheratzky, D. *BBA Section Nucleic Acids And Protein Synthesis* **1969**, *190*, 527.

(23) Paik, W. K.; Kim, S. *Journal of Biological Chemistry* **1970**, *245*, 88.

(24) Kakimoto, Y.; Akazawa, S. *Journal of Biological Chemistry* **1970**, *245*, 5751.

(25) Wolf, S. S. *Cell Molecular Life Sciences* **2009**, *66*, 2109.

(26) Greer, E. L.; Shi, Y. *Nature Reviews Genetics* **2012**, *13*, 343.

(27) Boffa, L. C.; Karn, J.; Vidali, G.; Allfrey, V. G. *Biochemical and biophysical research communications* **1977**, *74*, 969.

(28) Yang, Y.; Bedford, M. T. *Nature Reviews Cancer* **2013**, *13*, 37.

(29) Cheng, D.; Côté, J.; Shaaban, S.; Bedford, M. T. *Molecular cell* **2007**, *25*, 71.

(30) Branscombe, T. L.; Frankel, A.; Lee, J. H.; Cook, J. R.; Yang, Z.; Pestka, S.; Clarke, S. *Journal of Biological Chemistry* **2001**, *276*, 32971.

(31) Zobel-Thropp, P.; Gary, J. D.; Clarke, S. *Journal of Biological Chemistry* **1998**, *273*, 29283.

(32) Musselman, C. A.; Lalonde, M.-E.; Cote, J.; Kutateladze, T. G. *Nature Structure Molecular Bioloy* **2012**, *19*, 1218.

(33) Klose, R. J.; Zhang, Y. *Nature Reviews Molecular Cell Biology* **2007**, *8*, 307.

(34) Kooistra, S. M.; Helin, K. *Nature Reviews Molecular Cell Biology* **2012**, *13*, 297.

(35) Mosammaparast, N.; Shi, Y. *Annual review of biochemistry* **2010**, *79*, 155.

(36) Byvoet, P.; Shepherd, G. R.; Hardin, J. M.; Noland, B. J. *Archives of biochemistry and biophysics* **1972**, *148*, 558.

(37) Duerre, J. A.; Lee, C. T. *Journal of neurochemistry* **1974**, *23*, 541.

(38) Shi, Y.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J. R.; Cole, P. A.; Casero, R. A.; Shi, Y. *Cell*, *119*, 941.

(39) Klose, R. J.; Kallin, E. M.; Zhang, Y. *Nature Reviews Genetics* **2006**, *7*, 715.

(40) Chang, B.; Chen, Y.; Zhao, Y.; Bruick, R. K. *Science* **2007**, *318*, 444.

(41) Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., 3rd; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A. *Science* **2009**, *325*, 90.

(42) Bottger, A.; Islam, M. S.; Chowdhury, R.; Schofield, C. J.; Wolf, A. *The Biochemical journal* **2015**, *468*, 191.

(43) Ge, W.; Wolf, A.; Feng, T.; Ho, C.-h.; Sekirnik, R.; Zayer, A.; Granatino, N.; Cockman, M. E.; Loenarz, C.; Loik, N. D.; Hardy, A. P.; Claridge, T. D. W.; Hamed, R. B.; Chowdhury, R.; Gong, L.; Robinson, C. V.; Trudgian, D. C.; Jiang, M.; Mackeen, M. M.; McCullagh, J. S.; Gordiyenko, Y.; Thalhammer, A.; Yamamoto, A.; Yang, M.; Liu-Yi, P.; Zhang, Z.; Schmidt-Zachmann, M.; Kessler, B. M.; Ratcliffe, P. J.; Preston, G. M.; Coleman, M. L.; Schofield, C. J. *Nature Chemical Biology* **2012**, *8*, 960.

(44) Chowdhury, R.; Sekirnik, R.; Brissett, N. C.; Krojer, T.; Ho, C. H.; Ng, S. S.; Clifton, I. J.; Ge, W.; Kershaw, N. J.; Fox, G. C.; Muniz, J. R.; Vollmar, M.; Phillips, C.; Pilka, E. S.; Kavanagh, K. L.; von Delft, F.; Oppermann, U.; McDonough, M. A.; Doherty, A. J.; Schofield, C. J. *Nature* **2014**, *510*, 422.

(45) Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y. H.; Perlin, J. R.; Leonelli, L.; Sonbuchner, L. S.; McDonald, C. H.; Cook, R. G.; Dou, Y.; Roeder, R. G.; Clarke, S.; Stallcup, M. R.; Allis, C. D.; Coonrod, S. A. *Science* **2004**, *306*, 279.

(46) Knuckley, B.; Bhatia, M.; Thompson, P. R. *Biochemistry* **2007**, *46*, 6578.

(47) Fuhrmann, J.; Clancy, K. W.; Thompson, P. R. *Chemistry Review* **2015**, *115*, 5413.

(48) Ogawa, T.; Kimoto, M.; Sasaoka, K. *Journal of Biological Chemistry* **1989**, *264*, 10205.

(49) Marletta, M. A. *Journal of Biological Chemistry* **1993**, *268*, 12231.

(50) Bannister, A. J.; Schneider, R.; Kouzarides, T. *Cell* **2002**, *109*, 801.

(51) Schapira, M.; Ferreira de Freitas, R. *Medicinal Chemistry Communications* **2014**, *5*, 1779.

(52) Martin, G.; Ostareck-Lederer, A.; Chari, A.; Neuenkirchen, N.; Dettwiler, S.; Blank, D.; Ruegsegger, U.; Fischer, U.; Keller, W. *Rna* **2010**, *16*, 1646.

(53) Zhang, X.; Cheng, X. *Structure* **2003**, *11*, 509.

(54) Zhang, R.; Li, X.; Liang, Z.; Zhu, K.; Lu, J.; Kong, X.; Ouyang, S.; Li, L.; Zheng, Y. G.; Luo, C. *PLloS One* **2013**, *8*, e72424.

(55) Zhang, X.; Zhou, L.; Cheng, X. *The EMBO journal* **2000**, *19*, 3509.

(56) Obianyo, O.; Osborne, T. C.; Thompson, P. R. *Biochemistry* **2008**, *47*, 10420.

(57) Gui, S.; Wooderchak-Donahue, W. L.; Zang, T.; Chen, D.; Daly, M. P.; Zhou, Z. S.; Hevel, J. M. *Biochemistry* **2013**, *52*, 199.

(58) Sack, J. S.; Thieffine, S.; Bandiera, T.; Fasolini, M.; Duke, G. J.; Jayaraman, L.; Kish, K. F.; Klei, H. E.; Purandare, A. V.; Rosettani, P.; Troiani, S.; Xie, D.; Bertrand, J. A. *The Biochemical journal* **2011**, *436*, 331.

(59) Antonysamy, S.; Bonday, Z.; Campbell, R. M.; Doyle, B.; Druzina, Z.; Gheyi, T.; Han, B.; Jungheim, L. N.; Qian, Y.; Rauch, C.; Russell, M.; Sauder, J. M.; Wasserman, S. R.; Weichert, K.; Willard, F. S.; Zhang, A.; Emtage, S. *Proceedings of the National Academy of Sciences* **2012**, *109*, 17960.

(60) Sun, L.; Wang, M.; Lv, Z.; Yang, N.; Liu, Y.; Bao, S.; Gong, W.; Xu, R. M. *Proceedings of the National Academy of Sciences* **2011**, *108*, 20538.

(61) Gary, J. D.; Clarke, S. *Progress in Nucleic Acid Research Molecular Biology* **1998**, *61*, 65.

(62) Bedford, M. T.; Clarke, S. G. *Molecular cell* **2009**, *33*, 1.

(63) Di Lorenzo, A.; Bedford, M. T. *Febs Letters* **2011**, *585*, 2024.

(64) Cha, B.; Jho, E.-H. *Expert Opinion on Therapeutic Targets* **2012**, *16*, 651.

(65) Wang, R.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *Journal of the American Chemical Society* **2011**, *133*, 7648.

(66) McBride, A. E.; Silver, P. A. *Cell* **2001**, *106*, 5.

(67) Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57.

(68) Hanahan, D.; Weinberg, R. A. *Cell* **2011**, *144*, 646.

(69) Inche, A. G.; La Thangue, N. B. *Drug discovery today* **2006**, *11*, 97.

(70) Daser, A.; Rabbitts, T. H. *Seminars in cancer biology* **2005**, *15*, 175.

(71) Cheung, N.; Chan, L. C.; Thompson, A.; Cleary, M. L.; So, C. W. *Nature cell biology* **2007**, *9*, 1208.

(72) Yoshimatsu, M.; Toyokawa, G.; Hayami, S.; Unoki, M.; Tsunoda, T.; Field, H. I.; Kelly, J. D.; Neal, D. E.; Maehara, Y.; Ponder, B. A.; Nakamura, Y.; Hamamoto, R. *International Journal of Cancer* **2011**, *128*, 562.

(73) Avasarala, S.; Van Scoyk, M.; Karuppusamy Rathinam, M. K.; Zerayesus, S.; Zhao, X.; Zhang, W.; Pergande, M. R.; Borgia, J. A.; DeGregori, J.; Port, J. D.; Winn, R. A.; Bikkavilli, R. K. *Journal of Biological Chemistry* **2015**, *290*, 13479.

(74) Wilson, K. W., J *Principles and Techniques of Biochemistry and Molecular Biology*; Sixth ed.; Cambridge University Press, 2005.

(75) Frietze, S.; Lupien, M.; Silver, P. A.; Brown, M. *Cancer research* **2008**, *68*, 301.

(76) Kim, Y. R.; Lee, B. K.; Park, R. Y.; Nguyen, N. T.; Bae, J. A.; Kwon, D. D.; Jung, C. *BMC cancer* **2010**, *10*, 197.

(77) Al-Dhaheri, M.; Wu, J.; Skliris, G. P.; Li, J.; Higashimato, K.; Wang, Y.; White, K. P.; Lambert, P.; Zhu, Y.; Murphy, L.; Xu, W. *Cancer research* **2011**, *71*, 2118.

(78) Kim, K. Y.; Wang, D. H.; Campbell, M.; Huerta, S. B.; Shevchenko, B.; Izumiya, C.; Izumiya, Y. *Molecular Cell Biology* **2015**, *35*, 238.

(79) Habashy, H.; Rakha, E.; Ellis, I.; Powe, D. *Breast Cancer Res Treat* **2013**, *140*, 307.

(80) Wang, L.; Zhao, Z.; Meyer, M. B.; Saha, S.; Yu, M.; Guo, A.; Wisinski, K. B.; Huang, W.; Cai, W.; Pike, J. W.; Yuan, M.; Ahlquist, P.; Xu, W. *Cancer cell* **2014**, *25*, 21.

(81) Hong, H.; Kao, C.; Jeng, M. H.; Eble, J. N.; Koch, M. O.; Gardner, T. A.; Zhang, S.; Li, L.; Pan, C. X.; Hu, Z.; MacLennan, G. T.; Cheng, L. *Cancer* **2004**, *101*, 83.

(82) Limm, K.; Ott, C.; Wallner, S.; Mueller, D. W.; Oefner, P.; Hellerbrand, C.; Bosserhoff, A.-K. *European Journal of Cancer* **2013**, *49*, 1305.

(83) Osada, S.; Suzuki, S.; Yoshimi, C.; Matsumoto, M.; Shirai, T.; Takahashi, S.; Imagawa, M. *Oncology reports* **2013**, *30*, 1669.

(84) Jansson, M.; Durant, S. T.; Cho, E. C.; Sheahan, S.; Edelmann, M.; Kessler, B.; La Thangue, N. B. *Nature cell biology* **2008**, *10*, 1431.

(85) Yang, M.; Sun, J.; Sun, X.; Shen, Q.; Gao, Z.; Yang, C. *PLoS genetics* **2009**, *5*, e1000514.

(86) Pal, S.; Baiocchi, R. A.; Byrd, J. C.; Grever, M. R.; Jacob, S. T.; Sif, S. *The EMBO journal* **2007**, *26*, 3558.

(87) Chung, J.; Karkhanis, V.; Tae, S.; Yan, F.; Smith, P.; Ayers, L. W.; Agostinelli, C.; Pileri, S.; Denis, G. V.; Baiocchi, R. A.; Sif, S. *Journal of Biological Chemistry* **2013**, *288*, 35534.

(88) Li, Y.; Chitnis, N.; Nakagawa, H.; Kita, Y.; Natsugoe, S.; Yang, Y.; Li, Z.; Wasik, M.; Klein-Szanto, A. J.; Rustgi, A. K.; Diehl, J. A. *Cancer discovery* **2015**, *5*, 288.

(89) Nicholas, C.; Yang, J.; Peters, S. B.; Bill, M. A.; Baiocchi, R. A.; Yan, F.; Sif, S.; Tae, S.; Gaudio, E.; Wu, X.; Grever, M. R.; Young, G. S.; Lesinski, G. B. *PLloS One* **2013**, *8*, e74710.

(90) Shilo, K.; Wu, X.; Sharma, S.; Welliver, M.; Duan, W.; Villalona-Calero, M.; Fukuoka, J.; Sif, S.; Baiocchi, R.; Hitchcock, C. L.; Zhao, W.; Otterson, G. A. *Diagnostic pathology* **2013**, *8*, 201.

(91) Kim, J. M.; Sohn, H. Y.; Yoon, S. Y.; Oh, J. H.; Yang, J. O.; Kim, J. H.; Song, K. S.; Rho, S. M.; Yoo, H. S.; Kim, Y. S.; Kim, J. G.; Kim, N. S. *Clincal Cancer Research* **2005**, *11*, 473.

(92) Wei, T. Y. W.; Juan, C. C.; Hisa, J. Y.; Su, L. J.; Lee, Y. C. G.; Chou, H. Y.; Chen, J. M. M.; Wu, Y. C.; Chiu, S. C.; Hsu, C. P.; Liu, K. L.; Yu, C. T. R. *Cancer Sci.* **2012**, *103*, 1640.

(93) Yang, F.; Wang, J.; Ren, H. Y.; Jin, J.; Wang, A. L.; Sun, L. L.; Diao, K. X.; Wang, E. H.; Mi, X. Y. *Tumour Biology* **2015**.

(94) Powers, M. A.; Fay, M. M.; Factor, R. E.; Welm, A. L.; Ullman, K. S. *Cancer research* **2011**, *71*, 5579.

(95) Park, J. H.; Szemes, M.; Vieira, G. C.; Melegh, Z.; Malik, S.; Heesom, K. J.; Von Wallwitz-Freitas, L.; Greenhough, A.; Brown, K. W.; Zheng, Y. G.; Catchpoole, D.; Deery, M. J.; Malik, K. *Molecular oncology* **2015**, *9*, 617.

(96) Arrowsmith, C. H.; Audia, J. E.; Austin, C.; Baell, J.; Bennett, J.; Blagg, J.; Bountra, C.; Brennan, P. E.; Brown, P. J.; Bunnage, M. E.; Buser-Doepner, C.; Campbell, R. M.; Carter, A. J.; Cohen, P.; Copeland, R. A.; Cravatt, B.; Dahlin, J. L.; Dhanak, D.; Edwards, A. M.; Frye, S. V.; Gray, N.; Grimshaw, C. E.; Hepworth, D.; Howe, T.; Huber, K. V. M.; Jin, J.; Knapp, S.; Kotz, J. D.; Kruger, R. G.; Lowe, D.; Mader, M. M.; Marsden, B.; Mueller-Fahrnow, A.; Muller, S.; O'Hagan, R. C.; Overington, J. P.; Owen, D. R.; Rosenberg, S. H.; Roth, B.; Ross, R.; Schapira, M.; Schreiber, S. L.; Shoichet, B.; Sundstrom, M.; Superti-Furga, G.; Taunton, J.; Toledo-Sherman, L.; Walpole, C.; Walters, M. A.; Willson, T. M.; Workman, P.; Young, R. N.; Zuercher, W. J. *Nature Chemical Biology* **2015**, *11*, 536.

(97) Arrowsmith, C. H.; Audia, J. E.; Austin, C.; Baell, J.; Bennett, J.; Blagg, J.; Bountra, C.; Brennan, P. E.; Brown, P. J.; Bunnage, M. E.; Buser-Doepner, C.; Campbell, R. M.; Carter, A. J.; Cohen, P.; Copeland, R. A.; Cravatt, B.; Dahlin, J. L.; Dhanak, D.; Edwards, A. M.; Frye, S. V.; Gray, N.; Grimshaw, C. E.; Hepworth, D.; Howe, T.; Huber, K. V. M.; Jin, J.; Knapp, S.; Kotz, J. D.; Kruger, R. G.; Lowe, D.; Mader, M. M.; Marsden, B.; Mueller-Fahrnow, A.; Muller, S.; O'Hagan, R. C.; Overington, J. P.; Owen,

D. R.; Rosenberg, S. H.; Roth, B.; Ross, R.; Schapira, M.; Schreiber, S. L.; Shoichet, B.; Sundstrom, M.; Superti-Furga, G.; Taunton, J.; Toledo-Sherman, L.; Walpole, C.; Walters, M. A.; Willson, T. M.; Workman, P.; Young, R. N.; Zuercher, W. J. http://www.chemicalprobes.org/, 2015, p The Chemical Probes Portal.

(98) Horiuchi, K. Y.; Eason, M. M.; Ferry, J. J.; Planck, J. L.; Walsh, C. P.; Smith, R. F.; Howitz, K. T.; Ma, H. *Assay and drug development technologies* **2013**, *11*, 227.

(99) Cheng, D.; Yadav, N.; King, R. W.; Swanson, M. S.; Weinstein, E. J.; Bedford, M. T. *Journal of Biological Chemistry* **2004**, *279*, 23892.

(100) Spannhoff, A.; Heinke, R.; Bauer, I.; Trojer, P.; Metzger, E.; Gust, R.; Schule, R.; Brosch, G.; Sippl, W.; Jung, M. *Journal of Medicinal Cheimstry* **2007**, *50*, 2319.

(101) Egelhofer, T. A.; Minoda, A.; Klugman, S.; Lee, K.; Kolasinska-Zwierz, P.; Alekseyenko, A. A.; Cheung, M.-S.; Day, D. S.; Gadel, S.; Gorchakov, A. A.; Gu, T.; Kharchenko, P. V.; Kuan, S.; Latorre, I.; Linder-Basso, D.; Luu, Y.; Ngo, Q.; Perry, M.; Rechtsteiner, A.; Riddle, N. C.; Schwartz, Y. B.; Shanower, G. A.; Vielle, A.; Ahringer, J.; Elgin, S. C. R.; Kuroda, M. I.; Pirrotta, V.; Ren, B.; Strome, S.; Park, P. J.; Karpen, G. H.; Hawkins, R. D.; Lieb, J. D. *Nature Structure Molecular Biology* **2011**, *18*, 91.

(102) Egelhofer, T. A.; Minoda, A.; Klugman, S.; Lee, K.; Kolasinska-Zwierz, P.; Alekseyenko, A. A.; Cheung, M.-S.; Day, D. S.; Gadel, S.; Gorchakov, A. A.; Gu, T.; Kharchenko, P. V.; Kuan, S.; Latorre, I.; Linder-Basso, D.; Luu, Y.; Ngo, Q.; Perry, M.; Rechtsteiner, A.; Riddle, N. C.; Schwartz, Y. B.; Shanower, G. A.; Vielle, A.; Ahringer, J.; Elgin, S. C. R.; Kuroda, M. I.; Pirrotta, V.; Ren, B.; Strome, S.; Park, P. J.; Karpen, G. H.; Hawkins, R. D.; Lieb, J. D. http://compbio.med.harvard.edu/antibodies/, 2011.

(103) Lakowski, T. M.; t Hart, P.; Ahern, C. A.; Martin, N. I.; Frankel, A. *ACS Chemical Biology* **2010**, *5*, 1053.

(104) Gehrig, P. M.; Hunziker, P. E.; Zahariev, S.; Pongor, S. *Journal of the American Society for Mass Spectrometry* **2004**, *15*, 142.

(105) LaMarr, L. E. F. W. A. *Agilent Technologies, https://*http://www.agilent.com/cs/library/applications/5991-5260EN.pdf **2014**.

(106) Collazo, E.; Couture, J. F.; Bulfer, S.; Trievel, R. C. *Analytical Biochemistry* **2005**, *342*, 86.

(107) Hendricks, C. L.; Ross, J. R.; Pichersky, E.; Noel, J. P.; Zhou, Z. S. *Analytical Biochemistry* **2004**, *326*, 100.

(108) Ragno, R.; Simeoni, S.; Castellano, S.; Vicidomini, C.; Mai, A.; Caroli, A.; Tramontano, A.; Bonaccini, C.; Trojer, P.; Bauer, I.; Brosch, G.; Sbardella, G. *Journal of Medicinal Cheimstry* **2007**, *50*, 1241.

(109) Nishioka, K.; Reinberg, D. *Methods* **2003**, *31*, 49.

(110) Cheng, D.; Vemulapalli, V.; Bedford, M. T. *Methods Enzymology* **2012**, *512*, 71.

(111) Bedford, M. T. http://www.mdanderson.org/education-and-research/departments-programs-and-labs/labs/bedford-lab/protocols/index.html; Vol. 2015.

(112) Krause, C. D.; Yang, Z.-H.; Kim, Y.-S.; Lee, J.-H.; Cook, J. R.; Pestka, S. *Pharmacology and Therapeutics* **2007**, *113*, 50.

(113) Castellano, S.; Milite, C.; Ragno, R.; Simeoni, S.; Mai, A.; Limongelli, V.; Novellino, E.; Bauer, I.; Brosch, G.; Spannhoff, A.; Cheng, D.; Bedford, M. T.; Sbardella, G. *ChemMedChem* **2010**, *5*, 398.

(114) Heinke, R.; Spannhoff, A.; Meier, R.; Trojer, P.; Bauer, I.; Jung, M.; Sippl, W. *ChemMedChem* **2009**, *4*, 69.

(115) Bonham, K.; Hemmers, S.; Lim, Y. H.; Hill, D. M.; Finn, M. G.; Mowen, K. A. *The FEBS journal* **2010**, *277*, 2096.

(116) Fontan, N.; Garcia-Dominguez, P.; Alvarez, R.; de Lera, A. R. *Bioorganic and Medicinal Chemistry* **2013**, *21*, 2056.

(117) Allan, M.; Manku, S.; Therrien, E.; Nguyen, N.; Styhler, S.; Robert, M. F.; Goulet, A. C.; Petschner, A. J.; Rahil, G.; Robert Macleod, A.; Deziel, R.; Besterman, J. M.; Nguyen, H.; Wahhab, A. *Bioorganic and Medicinal Chemistry Letters* **2009**, *19*, 1218.

(118) Huynh, T.; Chen, Z.; Pang, S.; Geng, J.; Bandiera, T.; Bindi, S.; Vianello, P.; Roletto, F.; Thieffine, S.; Galvani, A.; Vaccaro, W.; Poss, M. A.; Trainor, G. L.; Lorenzi, M. V.; Gottardis, M.; Jayaraman, L.; Purandare, A. V. *Bioorganic and Medicinal Chemistry Letters* **2009**, *19*, 2924.

(119) Purandare, A. V.; Chen, Z.; Huynh, T.; Pang, S.; Geng, J.; Vaccaro, W.; Poss, M. A.; Oconnell, J.; Nowak, K.; Jayaraman, L. *Bioorganic and Medicinal Chemistry Letters* **2008**, *18*, 4438.

(120) Wan, H.; Huynh, T.; Pang, S.; Geng, J.; Vaccaro, W.; Poss, M. A.; Trainor, G. L.; Lorenzi, M. V.; Gottardis, M.; Jayaraman, L.; Purandare, A. V. *Bioorganic and Medicinal Chemistry Letters* **2009**, *19*, 5063.

(121) Cheng, D.; Valente, S.; Castellano, S.; Sbardella, G.; Di Santo, R.; Costi, R.; Bedford, M. T.; Mai, A. *Journal of Medicinal Cheimstry* **2011**, *54*, 4928.

(122) Sinha, S. H.; Owens, E. A.; Feng, Y.; Yang, Y.; Xie, Y.; Tu, Y.; Henary, M.; Zheng, Y. G. *European journal of medicinal chemistry* **2012**, *54*, 647.

(123) Hu, H.; Owens, E. A.; Su, H.; Yan, L.; Levitz, A.; Zhao, X.; Henary, M.; Zheng, Y. G. *Journal of Medicinal Cheimstry* **2015**, *58*, 1228.

(124) Osborne, T.; Roska, R. L.; Rajski, S. R.; Thompson, P. R. *Journal of the American Chemical Society* **2008**, *130*, 4574.

(125) Obianyo, O.; Causey, C. P.; Osborne, T. C.; Jones, J. E.; Lee, Y. H.; Stallcup, M. R.; Thompson, P. R. *Chembiochem : a European journal of chemical biology* **2010**, *11*, 1219.

(126) Bicker, K. L.; Obianyo, O.; Rust, H. L.; Thompson, P. R. *Molecular Biosystems* **2011**, *7*, 48.

(127) t Hart, P.; Lakowski, T. M.; Thomas, D.; Frankel, A.; Martin, N. I. *Chembiochem : a European journal of chemical biology* **2011**, *12*, 1427.

(128) Dowden, J.; Pike, R. A.; Parry, R. V.; Hong, W.; Muhsen, U. A.; Ward, S. G. *Organic and Biomolecular Chemistry* **2011**, *9*, 7814.

(129) van Haren, M.; van Ufford, L. Q.; Moret, E. E.; Martin, N. I. *Organic and Biomolecular Chemistry* **2015**, *13*, 549.

(130) Chan-Penebre, E.; Kuplast, K. G.; Majer, C. R.; Boriack-Sjodin, P. A.; Wigle, T. J.; Johnston, L. D.; Rioux, N.; Munchhof, M. J.; Jin, L.; Jacques, S. L.; West, K. A.; Lingaraj, T.; Stickland, K.; Ribich, S. A.; Raimondi, A.; Scott, M. P.; Waters, N. J.; Pollock, R. M.; Smith, J. J.; Barbash, O.; Pappalardi, M.; Ho, T. F.; Nurse, K.; Oza, K. P.; Gallagher, K. T.; Kruger, R.; Moyer, M. P.; Copeland, R. A.; Chesworth, R.; Duncan, K. W. *Nature Chemical Biology* **2015**, *11*, 432.

(131) Smil, D.; Eram, M. S.; Li, F.; Kennedy, S.; Szewczyk, M. M.; Brown, P. J.; Barsyte-Lovejoy, D.; Arrowsmith, C. H.; Vedadi, M.; Schapira, M. *ACS Medicinal Chemistry Letters* **2015**, *6*, 408.

(132) Siarheyeva, A.; Senisterra, G.; Allali-Hassani, A.; Dong, A.; Dobrovetsky, E.; Wasney, G. A.; Chau, I.; Marcellus, R.; Hajian, T.; Liu, F.; Korboukh, I.; Smil, D.; Bolshan, Y.; Min, J.; Wu, H.; Zeng, H.; Loppnau, P.; Poda, G.; Griffin, C.; Aman, A.; Brown, P. J.; Jin, J.; Al-Awar, R.; Arrowsmith, C. H.; Schapira, M.; Vedadi, M. *Structure* **2012**, *20*, 1425.

(133) Liu, F.; Li, F.; Ma, A.; Dobrovetsky, E.; Dong, A.; Gao, C.; Korboukh, I.; Liu, J.; Smil, D.; Brown, P. J.; Frye, S. V.; Arrowsmith, C. H.; Schapira, M.; Vedadi, M.; Jin, J. *Journal of Medicinal Cheimstry* **2013**, *56*, 2110.

(134) Bruggeman, F. J.; Westerhoff, H. V. *Trends in Microbiology*, *15*, 45.

(135) Chern, M. K.; Chang, K. N.; Liu, L. F.; Tam, T. C.; Liu, Y. C.; Liang, Y. L.; Tam, M. F. *Journal of Biological Chemistry* **2002**, *277*, 15345.

(136) Niewmierzycka, A.; Clarke, S. *Journal of Biological Chemistry* **1999**, *274*, 814.

(137) Rust, H. L.; Zurita-Lopez, C. I.; Clarke, S.; Thompson, P. R. *Biochemistry* **2011**, *50*, 3332.

(138) Merrifield, R. B. *Journal of the American Chemical Society* **1963**, *85*, 2149.

(139) Voet, D. V., J *Biochemistry*; Fourth ed.; Wiley, 2011.

(140) Rink, H.; Sieber, P.; Raschdorf, F. *Tetrahedron Lett.* **1984**, *25*, 621.

(141) Isidro-Llobet, A.; Alvarez, M.; Albericio, F. *Chemistry Review* **2009**, *109*, 2455.

(142) Ramage, R.; Green, J.; Blake, A. J. *Tetrahedron* **1991**, *47*, 6353.

(143) Isidro, A.; Latassa, D.; Giraud, M.; Alvarez, M.; Albericio, F. *Organic and Biomolecular Chemistry* **2009**, *7*, 2565.

(144) Hirs, C. H.; Moore, S.; Stein, W. H. *Journal of Biological Chemistry* **1952**, *195*, 669.

(145) Wiejak, S.; Masiukiewicz, E.; Rzeszotarska, B. *Chemical and Pharmaceutical Bulletin* **1999**, *47*, 1489.

(146) Wiejak, S.; Masiukiewicz, E.; Rzeszotarska, B. *Chemical and Pharmaceutical Bulletin* **2001**, *49*, 1189.

(147) Neumann, C. S.; Jiang, W.; Heemstra, J. R., Jr.; Gontang, E. A.; Kolter, R.; Walsh, C. T. *Chembiochem : a European journal of chemical biology* **2012**, *13*, 972.

(148) Kurtz, A. C. *Journal of Biological Chemistry* **1949**, *180*, 1253.

(149) Katritzky, A. R. R., Boris V. *ARKAT USA, Inc* **2005**, *4*, 49.

(150) Yong, Y. F.; Kowalski, J. A.; Lipton, M. A. *Journal of Organic Chemistry* **1997**, *62*, 1540.

(151) Ramadas, K.; Srinivasan, N. *Tetrahedron Lett.* **1995**, *36*, 2841.

(152) Kent, D. R.; Cody, W. L.; Doherty, A. M. *Tetrahedron Lett.* **1996**, *37*, 8711.

(153) Hamilton, D. J.; Sutherland, A. *Tetrahedron Lett.* **2004**, *45*, 5739.

(154) Ma, D.; Xia, C.; Jiang, J.; Zhang, J.; Tang, W. *Journal of Organic Chemistry* **2003**, *68*, 442.

(155) Ferrario, F.; Levi, S.; Sala, A.; Trupiano, F. *Synthetic Communications* **1991**, *21*, 99.

(156) Bada, J. L. *Earth and Planetary Science Letters* **1972**, *15*, 223.

(157) Bernatowicz, M. S.; Wu, Y.; Matsueda, G. R. *Journal of Organic Chemistry* **1992**, *57*, 2497.

(158) Bernatowicz, M. S.; Wu, Y.; Matsueda, G. R. *Tetrahedron Lett.* **1993**, *34*, 3389.

(159) Asami, T. *Patent* **2009**, *Application number: WO2009JP58102 20090423*

(160) Zahariev, S.; Guarnaccia, C.; Lamba, D.; Čemažar, M.; Pongor, S. *Tetrahedron Lett.* **2004**, *45*, 9423.

(161) Martin, N. I.; Liskamp, R. M. J. *Journal of Organic Chemistry* **2008**, *73*, 7849.

(162) Huang, Z. P.; Du, J. T.; Su, X. Y.; Chen, Y. X.; Zhao, Y. F.; Li, Y. M. *Amino Acids* **2007**, *33*, 85.

(163) Schade, D.; Töpker-Lehmann, K.; Kotthaus, J.; Clement, B. *Journal of Organic Chemistry* **2008**, *73*, 1025.

(164) Phuan, P.-W.; Kozlowski, M. C. *Journal of Organic Chemistry* **2002**, *67*, 6339.

(165) Locke, J. M.; Griffith, R.; Bailey, T. D.; Crumbie, R. L. *Tetrahedron* **2009**, *65*, 10685.

(166) Greever, J. C. *Journal of Chemical Education* **1995**, *72*, A151.

(167) Green, T. W. W., P. G. M; *Protective Groups in Organic Synthesis*; Wiley-Interscience, 1999.

(168) Olofson, R. A.; Martz, J. T.; Senet, J. P.; Piteau, M.; Malfroot, T. *Journal of Organic Chemistry* **1984**, *49*, 2081.

(169) Applegate, H. E.; Cimarusti, C. M.; Dolfini, J. E.; Funke, P. T.; Koster, W. H.; Puar, M. S.; Slusarchyk, W. A.; Young, M. G. *Journal of Organic Chemistry* **1979**, *44*, 811.

(170) Luzzi, S. D.; Marletta, M. A. *Bioorganic and Medicinal Chemistry Letters* **2005**, *15*, 3934.

(171) Fukuyama, T.; Cheung, M.; Jow, C.-K.; Hidai, Y.; Kan, T. *Tetrahedron Lett.* **1997**, *38*, 5831.

(172) Kaehler, H.; Rehse, K. *Tetrahedron Lett.* **1968**, *9*, 5019.

(173) Wenzel, T. J. *Discrimination of Chiral Compounds Using NMR Spectroscopy*; Wiley, 2007.

(174) Walport, L. J., University of Oxford, 2013.

(175) Wang, H.; Huang, Z. Q.; Xia, L.; Feng, Q.; Erdjument-Bromage, H.; Strahl, B. D.; Briggs, S. D.; Allis, C. D.; Wong, J.; Tempst, P.; Zhang, Y. *Science* **2001**, *293*, 853.

(176) Sigma-Aldrich
http://www.sigmaaldrich.com/catalog/product/sigma/a7007?lang=en&region=GB.

(177) Fassett, J. D.; Paulsen, P. J. *Analytical chemistry* **1989**, *61*, 643A.

(178) Nelson, R. W.; McLean, M. A.; Hutchens, T. W. *Analytical chemistry* **1994**, *66*, 1408.

(179) Eisenthal, R.; Danson, M. J.; Hough, D. W. *Trends in biotechnology* **2007**, *25*, 247.

(180) Lakowski, T. M.; Frankel, A. *Biochem J* **2009**, *421*, 253.

(181) Mann, H. B. a. W., D. R. *The Annals of Mathematical Statistics* **1947**, *18*, 50.

(182) Challis, B. C.; Iley, J. N. *Journal of the Chemical Society, Perkin Transactions 2* **1985**, 699.

(183) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Advanced Drug Delivery Reviews* **1997**, *23*, 3.

(184) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. *Journal of chemical information and computer sciences* **2001**, *41*, 1308.

(185) Millipore, M. http://www.merckmillipore.com/GB/en/product/Histone-Methyltransferase-Assay-Reagent-Kit,MM_NF-17-330?bd=1 - anchor_Biological Information.

(186) Ullmann, F. a. B., J. *Ber. Dtsch. Chem. Ges.* **1901**, *34*, 2174.

(187) Kim, S. S. N., Kasi; Kim, Sang Soo; Kim, Dong Won; Jung, Hyun Chul *Synthesis* **2002**, 2484.

(188) Prileschajew, N. *Ber. Dtsch. Chem. Ges* **1909**, *42*, 4811.

(189) Kulig, K.; Boba, A.; Bielejewska, A.; Gorska, M.; Malawska, B. *Tetrahedron: Asymmetry* **2009**, *20*, 322.

(190) Snyder, D. C. *Journal of Organic Chemistry* **1995**, *60*, 2638.

(191) Appel, R. *Angewandte Chemie International Edition* **1975**, *14*, 801.

(192) Dahl, T.; Tornøe, C. W.; Bang-Andersen, B.; Nielsen, P.; Jørgensen, M. *Angewandte Chemie International Edition* **2008**, *47*, 1726.

(193) Gschwend, H. W.; Rodriguez, H. R. In *Organic Reactions*; John Wiley & Sons, Inc.: 2004.

(194) Zhu, X.-Q.; Dai, Z.; Yu, A.; Wu, S.; Cheng, J.-P. *The Journal of Physical Chemistry B* **2008**, *112*, 11694.

(195) Hauck, M.; Schönhaber, J.; Zucchero, A. J.; Hardcastle, K. I.; Müller, T. J. J.; Bunz, U. H. F. *Journal of Organic Chemistry* **2007**, *72*, 6714.

(196) Madrid, P. B.; Polgar, W. E.; Toll, L.; Tanga, M. J. *Bioorganic and Medicinal Chemistry Letters* **2007**, *17*, 3014.

(197) Roe, A.; Montgomery, J. A.; Yarnall, W. A.; Hoyle, V. A. *Journal of Organic Chemistry* **1956**, *21*, 28.

(198) He, C.; Chen, C.; Cheng, J.; Liu, C.; Liu, W.; Li, Q.; Lei, A. *Angewandte Chemie International Edition* **2008**, *47*, 6414.

(199) Tao, C. A. L., Nan Zhaoa, Shuai Yanga, Xiaolang Liua, Jian Zhoua, Weiwei Liua, Jing Zhao *Synthetic Letters*, *1*, 134.

(200) He, C. 2009.

(201) Winter, D. K.; Drouin, A.; Lessard, J.; Spino, C. *Journal of Organic Chemistry* **2010**, *75*, 2610.

(202) Ma, D.; Geng, Q.; Zhang, H.; Jiang, Y. *Angewandte Chemie International Edition* **2010**, *49*, 1291.

(203) Ramanathan, R.; Su, A. D.; Alvarez, N.; Blumenkrantz, N.; Chowdhury, S. K.; Alton, K.; Patrick, J. *Analytical chemistry* **2000**, *72*, 1352.

(204) Parham, F.; Austin, C.; Southall, N.; Huang, R.; Tice, R.; Portier, C. *Journal of biomolecular screening* **2009**, *14*, 1216.

(205) Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P. *Journal of Medicinal Cheimstry* **2007**, *50*, 2385.

(206) Foster, M. P.; McElroy, C. A.; Amero, C. D. *Biochemistry* **2007**, *46*, 331.

(207) Attwood, D.; Mosquera, V.; Lopez-Fontan, J. L.; Garcia, M.; Sarmiento, F. *Journal of Colloid and Interface Science* **1996**, *184*, 658.

(208) Baell, J. B.; Holloway, G. A. *Journal of medicinal chemistry* **2010**, *53*, 2719.

(209) Saubern, S.; Guha, R.; Baell, J. B. *Molecular Informatics* **2011**, *30*, 847.

(210) Pangborn, A.; Giardello, M.; Grubbs, R.; Rosen, R.; Timmers, F. *Organometallics* **1996**, *15*, 1518.

(211) Tiainen, M.; Maaheimo, H.; Niemitz, M.; Soininen, P.; Laatikainen, R. *Magnetic Resonance in Chemistry* **2008**, *46*, 125.

(212) Sigma-Aldrich 2015.

(213) De Marco, R.; Di Gioia, M. L.; Leggio, A.; Liguori, A.; Perri, F.; Siciliano, C.; Viscomi, M. C. *Amino Acids* **2010**, *38*, 691.

(214) Millipore, M. **2015**.

(215) Halim, D.; Caron, K.; Keillor, J. W. *Bioorganic and Medicinal Chemistry Letters* **2007**, *17*, 305.

(216) Masiukiewicz, E.; Wiejak, S.; Rzeszotarskat, B. *Organic Preparations and Procedures International* **2002**, *34*, 531.

(217) Hickey, S. M. A., T.D; Khosa, S.K; Pfeffer, F.M; *Synthetic Letters* **2012**, *23*, 1779.

(218) Moore, W. M.; Webber, R. K.; Fok, K. F.; Jerome, G. M.; Kornmeier, C. M.; Tjoeng, F. S.; Currie, M. G. *Bioorganic and Medicinal Chemistry* **1996**, *4*, 1559.

(219) Lespagnol, A.; Cuingnet, E.; Debaert, M. *Bulletin de la Société Chimique de France* **1960**, *2*, 383.

(220) Lecher, H.; Heuck, C. *Justus Liebigs Annalen der Chemie* **1924**, *438*, 169.

(221) de Gracia Lux, C.; Olejniczak, J.; Fomina, N.; Viger, M. L.; Almutairi, A. *Journal of Polymer Science Part A: Polymer Chemistry* **2013**, *51*, 3783.

(222) Nowshuddin, S.; Rao, M. N. A.; Reddy, A. R. *Synthetic Communications* **2009**, *39*, 2022.

(223) Rossiter, S.; Smith, C. L.; Malaki, M.; Nandi, M.; Gill, H.; Leiper, J. M.; Vallance, P.; Selwood, D. L. *Journal of medicinal chemistry* **2005**, *48*, 4670.

(224) Jakobsche, C. E.; McEnaney, P. J.; Zhang, A. X.; Spiegel, D. A. *ACS Chemical Biology* **2012**, *7*, 316.

(225) AKScientific *aksci.com/item_detail.php?cat=N940* **2015**.

(226) Cai, M.; Xiao, R.; Yan, T.; Zhao, H. *Journal of Organometallic Chemistry* **2014**, *749*, 55.

(227) Laali, K. K.; Nagvekar, D. S. *Journal of Organic Chemistry* **1991**, *56*, 1867.

(228) Campbell, R.; Hatton, R. *Journal of Organic Chemistry* **1961**, *26*, 2480.

(229) McNeil, S. K.; Kelley, S. P.; Beg, C.; Cook, H.; Rogers, R. D.; Nikles, D. E. *ACS applied materials & interfaces* **2013**, *5*, 7647.

(230) Danilevicius, A.; Ostrauskaite, J.; Grazulevicius, J. V.; Gaidelis, V.; Jankauskas, V.; Tokarski, Z.; Jubran, N.; Sidaravicius, J.; Grevys, S.; Dzena, A. *Journal of Photochemistry and Photobiology A: Chemistry* **2004**, *163*, 523.

(231) Gilman, H.; Ingham, R.; Champaigne, J.; Diehl, J.; Ranck, R. *Journal of Organic Chemistry* **1954**, *19*, 560.

(232) Tu, X.; Fu, X.; Jiang, Q. *Displays* **2010**, *31*, 150.

(233) Gaina, L.; Csampai, A.; Turos, G.; Lovasz, T.; Zsoldos-Mady, V.; Silberg, I. A.; Sohar, P. *Organic and Biomolecular Chemistry* **2006**, *4*, 4375.

(234) Gilman, H.; Ingham, R.; Wu, T. C. *Journal of the American Chemical Society* **1952**, *74*, 4452.

(235) Albagli, O.; Flourens, A.; Crepieux, P.; Begue, A.; Stehelin, D.; Leprince, D. *Oncogene* **1992**, *7*, 1435.

(236) Burkamp, F.; Fletcher, S. R. *Journal of Heterocyclic Chemistry* **2002**, *39*, 1177.

(237) Bohme, T. 2008.

(238) Ramalingan, C.; Park, Y.-T. *Journal of Organic Chemistry* **2007**, *72*, 4536.

(239) Barsyte-Lovejoy, D.; Li, F.; Oudhoff, M. J.; Tatlock, J. H.; Dong, A.; Zeng, H.; Wu, H.; Freeman, S. A.; Schapira, M.; Senisterra, G. A.; Kuznetsova, E.; Marcellus, R.; Allali-Hassani, A.; Kennedy, S.; Lambert, J. P.; Couzens, A. L.; Aman, A.; Gingras, A. C.; Al-Awar, R.; Fish, P. V.; Gerstenberger, B. S.; Roberts, L.; Benn, C. L.; Grimley, R. L.; Braam, M. J.; Rossi, F. M.; Sudol, M.; Brown, P. J.; Bunnage, M. E.; Owen, D. R.; Zaph, C.; Vedadi, M.; Arrowsmith, C. H. *Proceedings of the National Academy of Sciences* **2014**, *111*, 12853.