Investigating Bisubstrate Inhibitors of PRMT1 and CARM1

Emma Gunnell

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2018

Abstract

Protein arginine methyltransferases (PRMTs) are gaining traction as a novel drug target class for a wide range of diseases. The high sequence conservation within the active sites of the nine human PRMTs, particularly in the co-factor (S-adenosyl methionine [SAM]) binding site, makes the design of isozyme-selective PRMT inhibitors challenging. Linking guanidine or substrate peptides to SAM mimics has shown potential as an approach to designing isozyme-selective inhibitors, but the peptidic nature of such compounds limits their *in vivo* utility. Herein, the binding of a novel series of non-peptidic bisubstrate inhibitors to CARM1 and PRMT1 was evaluated.

The bisubstrate nature of the inhibitors was confirmed by crystal structures of the inhibitors bound to CARM1. Isothermal titration calorimetry (ITC) and activity assays showed that varying the alkyl linker length between the guanidine and adenosine moieties of the inhibitors tested does not confer selectivity for PRMT1 or CARM1. These findings were supported by co-crystal structures of the inhibitors bound to CARM1, which showed that the guanidinium group is thermally mobile within the CARM1 active site. Conversely, isosteric replacement of the substrate-mimic with an N-alkyl-2-aminopyridine group was found to confer selectivity for CARM1 over PRMT1 (K_d values of 1.1 μM and 43.7 μM, respectively, for the most selective inhibitor).

A CARM1 N265Y mutant displayed a significant reduction in affinity for the CARM1-selective aminopyridine inhibitor (judged by ITC), suggesting that this

residue, located in the substrate binding channel, may contribute to the CARM1-selectivity of this inhibitor.

The substitution of ribose for a morpholino group in PRMT bisubstrate inhibitors was found to abrogate inhibitor binding to PRMT1 and CARM1 in the majority of cases. Additionally, two putative FRET probes designed to bind to PRMT1 and CARM1 were designed and partially synthesised.

Acknowledgements

Firstly, I would like to thank Dr Ingrid Dreveny and Dr James Dowden for giving me the opportunity to complete this PhD and for their support over the last four years.

Additionally, I would like to thank Dr Clare Davies for her encouragement during my first year and help with the methylation assays. Thank you also to Prof. Steve Alexander for allowing me to work in his radioisotope lab, and to Dr Michael Garle for his help with all things radioactive. I would also like to thank Ben Pointer-Gleadhill for his assistance with mass spectrometry (which did not make it into this thesis) and Dr Huw Williams and Shazad Aslam for their NMR assistance.

I am very grateful for all the help and friendship from members of the Dreveny, Dowden and Emsley research groups. I couldn't have asked for a better group of lab mates. In particular, I would like to thank Dr Simon Caulton, for always taking the time to help others with their work and Danilo Belviso for his patience and assistance in solving my first crystal structure.

Thank you to Simon, Steve, Ingrid and James for help with proof-reading.

For the funding I have received to complete this PhD, I would like to thank the Medical Research Council (MRC).

List of Figures[1]

Figure 1: Arginine methylation patterns deposited by PRMTs 1

Figure 2: sDMA binding sites in a published crystal structure of an 'extended Tudor domain' from SND1 (PDB code 3OMC). 2

Figure 3: Structure of SAM-mimic sinefungin with SAM shown for comparison 4

Figure 4: Domain architecture of the PRMTs 5

Figure 5: Superposition of PRMT catalytic cores 6

Figure 6: Sequence alignment of human PRMTs 7

Figure 7: Structural elements of the SAM (top) and H3 (bottom) binding site of CARM1 (PDB code 5DX0) 8

Figure 8: Proposed mechanism of PRMT1-catalysed arginine methylation 9

Figure 9: Kinetic mechanism of PRMT1: Sequential ordered bi-bi. 12

Figure 10: CARM1 kinetic mechanism proposed by Jacques *et al.* based on kinetic studies.[43] 16

Figure 11: Structures of EPZ020411 and GSK3326595[57] 19

Figure 12: Recent examples of PRMT1 inhibitors 20

Figure 13: Selective, cell active CARM1 inhibitors TP-064, EZM2302 and SKI-73 22

Figure 14: Notable examples of bisubstrate inhibitors – muripirocin and isoniazid 23

Figure 15: Structure of inhibitors used in this thesis 27

Figure 16: Repurposed putative DNA methyltransferase inhibitors reported by Halby *et al.* 31

Figure 17: PCR protocol used for restriction enzyme cloning 34

Figure 18: PCR protocol used for site direction mutagenesis 37

Figure 19: Schematic representation of PRMT1 and CARM1 constructs used for this thesis 83

[1] All Figures in this thesis are originals generated using PyMOL Molecular Graphics system (version 2.1.0) Schrodinger LLC; ChemBioDraw, version 14; Microsoft Powerpoint and GraphPad Prism 7.03.

Figure 20: Schematic representation of cleavage products of cleavable CARM1 fusion proteins used in this thesis 85

Figure 21: Schematic representation of histone constructs used in this thesis 86

Figure 22: Purification of PRMT1-C-His 89

Figure 23: Purification of PRMT1-N-His 90

Figure 24: Initial activity test of PRMT1-N-His 91

Figure 25: Effect of various modifications to the purification procedure/construct on gel filtration elution profile 93

Figure 26: Anion exchange of pooled nickel affinity fractions. 95

Figure 27: Co-purification of PRMT1 and GB-H4 97

Figure 28: Appearance of apparent crystalline material in C3 Structure screen (entry 3, Table 15) 100

Figure 29: Purification of MBP-CARM1-His 102

Figure 30: TEV cleavage and subsequent isolation of CARM1-His 104

Figure 31: Activity screen of CARM1 constructs. 106

Figure 32: Nickel affinity purification of 2GKG-CARM1 109

Figure 33: Test expression of MBP-tagged CARM1 constructs in ArcticExpress (AX) and BL21 cells 110

Figure 34: Optimisation of 2GKG-CARM1-His expression conditions using 1 L expressions 110

Figure 35: SDS-PAGE analysis of samples from co-purification of MBP-CARM1-His and GB-H3 112

Figure 36: Purification of GST-mCARM1(FL) 114

Figure 37: Attempted purification of GST-mCARM1(FL) expressed in BL21 cells using the pCOLD vector 115

Figure 38: Optimisation of GST-mCARM1(FL) expression conditions using 1 L expressions 116

Figure 39: Analytical gel filtrations (Superdex 200 10/300) of eluates from GST-affinity column from Figure 38 (above) 118

Figure 40: Purification of TF-mCARM1(FL) (115 kDa) 120

Figure 41: Cleavage of TF-mCARM1(FL) and subsequent attempts at purification..... 122

Figure 42: CARM1-His crystallisation trials using PACT premier™ screen (Molecular Dimensions) both with (upper) and without (lower) the addition of 100 μM SAH..... 124

Figure 43: CARM1-His Crystals obtained using the PACT premier™ screen (Molecular Dimensions) 125

Figure 44: CARM1-SAH co-crystallisation optimisation plates. Light blue conditions gave crystals within 1 month, and dark blue squares gave crystals within 1 week..... 127

Figure 45: CARM1-His + UM249 co-crystallisation..... 128

Figure 46: Crystal structure of UM249 (green) bound to CARM1 (white). $2mF_o$-F_c map (contoured at 1.0 σ, carve radius = 1.5 around the ligand) is shown in blue and a published crystal structure of CARM1 (white) bound to histone 3 (magenta) and SAM-mimic sinefungin (green) is superposed (PDB 5Dx0) 130

Figure 47: ITC titration of SAH into PRMT1 and CARM1. Thermograms are shown in the upper panels and integrated heat plots with fitted binding curves in the lower panels..... 132

Figure 48: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM079, UM249 and UM251 with PRMT1 and CARM. 134

Figure 49: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for WH5C with PRMT1 and CARM..... 135

Figure 50: Signature plots for SAH, UM079, UM249, UM251 and WH5C binding to PRMT1 (upper) and CARM1 (lower). The inhibitor to which each dataset corresponds is shown above the graph. The data are given numerically in Table 20 and Table 21, below. 138

Figure 51: Purification of histone GB fusions 141

Figure 52: PRMT1 ^3H assay validation..... 143

Figure 53: Determining linear range of CARM1 assay signal..... 144

Figure 54: Estimate of Km^{app} of CARM1 histone substrate (GB-H3(3-31)His) under assay conditions. 144

Figure 55: IC_{50} assay data for linker length series UM079, UM249 and UM251 with PRMT1 and CARM1 146

Figure 56: IC_{50} assay data for SAH and WH5C with PRMT1 and CARM1. 147

Figure 57: Graphs showing IC_{50} (top) and K_d (bottom) values determined for PRMT1 (left) and CARM1 (right) 148

Figure 58: Co-crystal structures of amine guanidine inhibitors with CARM1.....151

Figure 59: Comparison of 'linker length' amine-guanidine inhibitor series with peptide-fused inhibitors reported by Van Haren *et al.*[74]..... 152

Figure 60: Full length CARM1 homology model superposed with CARM1·WH5C co-crystal structure 154

Figure 61: Full length CARM1 homology model superposed with CARM1 dimer from our CARM1·WH5C co-crystal structure..... 155

Figure 62: Structures of aromatic inhibitors UM305, AA175 and AA183..... 157

Figure 63: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM305, AA175 and AA183 with PRMT1 and CARM1.... 158

Figure 64: IC_{50} assay data for aromatic inhibitors UM305, AA175 and AA183 with PRMT1 and CARM1. The error bars show standard deviations of duplicate measurements at each inhibitor concentration. The IC_{50} values determined by curve-fitting are given in Table 23 (below). 159

Figure 65: Competition ITC studies of AA183 binding to PRMT1..... 161

Figure 66: Signature plots for SAH, UM305, AA175 and AA183 binding to PRMT1 (upper) and CARM1 (lower)..... 163

Figure 67: Predicted pK_a (red) and pK_{aH} (blue) values for protonation sites within inhibitors UM305, AA175 and AA183..... 165

Figure 68: Comparison of optimal ligand geometries output by GRADE (pink) and eLBOW (cyan)..... 167

Figure 69: Crystal structures of AA175 and AA183 bound to CARM1 169

Figure 70: (A) Binding pockets of adenosine (orange dashes), SAM amino acid (red dashes), and substrate binding channel (teal dashes) in the published CARM1·H3·sinefungin crystal structure (PDB code 5DX0);[15a] (B) Superposition of UM305 (pale pink), AA175 (purple) and AA183 (blue) co-crystals with a published structure of CARM1 in complex with histone 3 and sinefungin (magenta, PDB code 5DX0)..... 170

Figure 71: Superposition of the CARM1·AA183 crystal structure (grey) with PRMT1 structure 1or8 (cyan)..... 171

Figure 72: SDS-PAGE of single site mutants of PRMT1 (A) and CARM1 (B) 172

Figure 73: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM305, AA175 and AA183 with CARM(N161E) and CARM(N265Y) 173

Figure 74: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for SAH with CARM1(N161E) and CARM1(N265Y)..... 174

Figure 75: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for SAH and UM305 with PRMT1(E55N) and PRMT1(Y160N) 175

Figure 76: Signature plots for SAH, UM305, AA175 and AA183 binding to CARM1 and single site mutants thereof. 178

Figure 77: Signature plots for SAH, UM305, AA175 and AA183 binding to PRMT1 and single site mutants thereof. 179

Figure 78: Crystal structures of CARM1(N265Y) in complex with AA183 182

Figure 79: Crystal structure of CARM1(N265Y) in complex with AA175..... 184

Figure 80: Morpholino inhibitors synthesised by Dr U. Muhsen. 187

Figure 81: ITC titrations of morpholino inhibitors into PRMT1..... 189

Figure 82: Superposition of PRMT1 (PDB code 1OR8) and CARM1 (PDB code 5DX0) crystal structures showing location of tryptophans..... 191

Figure 83: Structure of tryptophan and examples of tryptophan FRET acceptors. 192

Figure 84: Proposed route to putative FRET probes. 193

Figure 85: Synthesis of protected propylamino scaffold 25 194

Figure 86: Synthesis of boc-protected amino-aldehyde 32 194

Figure 87: Synthesis of boc-protected naphthyl aldehydes 37 and 38 195

Figure 88: Second reductive amination and subsequent global deprotection..... 196

Figure 89: Proposed mechanisms for depurination and racemisation. 197

Figure 90: Inhibition screen of EG047 and EG048 vs PRMT1 and CARM1..... 198

Figure 91: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for EG047 and EG048 with PRMT1 and CARM1..... 198

Figure 92: Active site electron density for EG047 in active site of CARM1(WT) 200

Figure 93: Guanidine-containing bisubstrate CARM1 inhibitors for which there are co-crystal structures in the PDB..... 203

Figure 94: Comparison of guanidinium positioning in guanidine-containing bisubstrate inhibitors co-crystallised with CARM1 in the PDB..... 204

Figure 95: Guanidine-containing bisubstrate CARM1 inhibitors for which there are co-crystal structures in the PDB..... 206

Figure 96: Superposition of CARM1-AA175 co-crystal structure with aromatic bisubstrate inhibitors co-crystallised with CARM1 in the PDB..... 207

List of Tables

Table 1: IC_{50} data for bisubstrate SAM inhibitors reported by Dowden *et al.*[69] 24

Table 2: IC_{50} data for bisubstrate inhibitors reported by Van Haren *et al.*[70] 25

Table 3: Summary of PCR primers used for restriction enzyme cloning – PRMT1 and CARM1 39

Table 4: Summary of PCR primers used for side directed mutagenesis 40

Table 5: Summary of PCR primers used for restriction enzyme cloning - histone constructs 41

Table 6: Expression conditions for constructs used in this thesis. 42

Table 7: Buffers used for nickel affinity chromatography and gel filtration 46

Table 8: Western Blot antibodies used and working concentrations. 52

Table 9: Crystallisation conditions used to generate CARM1-inhibitor co-crystals 55

Table 10: Enzyme and substrate concentrations used for initial activity tests of purified recombinant constructs. 58

Table 11: Protein and inhibitor concentrations used for isothermal titration calorimetry. 64

Table 12: PRMT1 and CARM1 constructs used in this thesis 84

Table 13: CARM1 constructs generated by treating CARM1 fusion proteins with the appropriate protease. 85

Table 14: Histone 3 and histone 4 constructs used in this thesis 86

Table 15: Volume of bacterial culture grown, mass of cell pellet, and mass of protein at each stage of a representative purification of each construct. 87

Table 16: PRMT1 crystallisation trials set up to date. 100

Table 17: Purification and activity screening of MBP-CARM1 fusion constructs 106

Table 18: Composition of PACT premier™ (Molecular Dimensions) screen solutions which subsequent optimisation CARM1-His plates were based on. 126

Table 19: K_d and IC_{50} values for our series of inhibitors versus PRMT1 and CARM1 measured using ITC (present study) and enzyme activity assays (previous work), respectively 136

Table 20: ΔG, ΔH, -TΔS and stoichiometry (N) determined from titrations of SAH, UM079, UM249, UM251 and WH5C into PRMT1..... 139

Table 21: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titrations of SAH, UM079, UM249, UM251 and WH5C into CARM1 139

Table 22: IC_{50} values for our series of inhibitors versus PRMT1 and CARM1 measured using enzyme activity assays 148

Table 23: K_d and IC_{50} values for aromatic isostere inhibitors *versus* PRMT1 and CARM1 measured using ITC and enzyme activity assays, respectively. 160

Table 24: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into PRMT1 164

Table 25: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into CARM1..... 164

Table 26: K_d values for aminopyridine/aminopyrimidine inhibitors versus PRMT1 and CARM1 single site mutants measured using ITC..... 176

Table 27: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into CARM1 and mutants thereof..... 180

Table 28: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into PRMT1 181

Table 29: Inhibitor screening for morpholino analogues..... 187

List of Abbreviations

ΔH	Standard molar enthalpy change
2GKG	Receiver domain of myxococcus xanthus social motility protein FrzS (PDB code 2GKG)
Ac	Acetate
AD2	Activation domain 2
aDMA	Asymmetric dimethyl arginine
ADPR	Adenosine diphosphate ribose
AML	Acute myeloid leukemia
ATR	Attenuated total reflection
Boc	tert-Butyloxycarbonyl
CARM1	Co-activator associated arginine methyltransferase 1 (human)
COSY	Homonuclear correlation spectroscopy
Da	Daltons
DEPTQ	Distortionless enhancement by polarisation transfer with quaternaries
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DPM	Disintegrations per minute
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-mesenchymal transition
ESI^+	Electrospray ionization (positive ion mode)
Et	Ethyl
FDA	Food and Drug Administration
FL	Full length
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
GAR	Glycine arginine rich
GB	B1 domain of streptococcal protein G (PDB code 3GB1)
GRIP1	Glutamate receptor-interacting protein 1
GST	Glutathione S-transferase
H	Hour(s)
HepG2	Hepatoma G2 (liver cells)
HIV-1	Human immunodeficiency virus 1
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single-quantum correlation
IC_{50}	Half maximal inhibitor concentration
IPA	Isopropanol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KATG	Catalase-peroxidase
Kb	Kilobases
K_d	Dissociation constant
K_i	Inhibition constant
K_m	Michaelis-Menten constant
K_m^{app}	Apparent Michaelis-Menten constant
LB	Luria Bertani/Lysogeny broth
MBP	Maltose binding protein

mCARM1	Coactivator associated arginine methyltransferase (mouse)
MCF7	Michigan Cancer Foundation-7 (breast cancer cells)
Me	Methyl
MHRA	Medicines and Healthcare products Regulatory Agency
Min(s)	Minute(s)
MMA	Monomethyl arginine
MRC	Medical Research Council
MW	Molecular weight
NADH	Nicotinamide adenine dinucleotide
NEB	New England Biosciences
NMR	Nuclear magnetic resonance
NS3	Nonstructural protein 3
OD_{600}	Optical density at 600 nM
p/CIP	p300 CREB-binding protein/cointegrator-associated protein (aka NCoA-3)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG X	Polyethylene glycol with average molecular weight X Da
PH	Plextrin homology
PHD	Plant homeodomain
PRMT	Protein Arginine Methyltransferase (human)
REC	Restriction enzyme cloning
R_f	Retention factor
RNA	Ribonucleic acid
RPM	Rotations per minute
SAH	S-adenosyl homocysteine
SAM	S-adenosyl methionine
SDM	Site-directed mutagenesis
sDMA	Symmetric dimethylarginine
SDS	Sodium dodecyl sulphate
Sec(s)	Second(s)
SET	Su(var)3-9, enhancer-of-zeste and trithorax
SH3	SRC homology 3 domain
ShRNA	Short hairpin RNA
SiRNA	Small inhibitory ribonucleic acid
SWI/SNF	Switch/sucrose non-fermentable
TBST	Tris-buffered saline with TWEEN-20
TCA	Trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco etch virus (protease)
TF	Trigger factor
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
THP1	Tamm-Horsfall Protein-1 (leukemia cells)
TIM	Triosephosphateisomerase
TPR	Tetratricopeptide repeat
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
V/v	Volume-to-volume
W/v	Weight-to-volume
WBTB	Western blot transfer buffer
WD40	Tryptophan-aspartic acid 40
w.r.t.	With respect to

ZnF	Zinc finger

Contents

Abstract i

Acknowledgements iii

List of Figures iv

List of Tables x

List of Abbreviations xii

Contents xv

CHAPTER 1: Introduction 1

1.1. Arginine Methylation and PRMTs 1

1.2. PRMT1 11

1.2.1. Structure 11

1.2.2. Mechanism 12

1.2.3. Disease Links and Target Validation 13

1.3. CARM1 13

1.3.1. Structure 14

1.3.2. Mechanism 15

1.3.3. Disease Links and Target Validation 17

1.4. PRMT Inhibitors 17

1.4.1. PRMT1 Inhibitors 19

1.4.2. CARM1 Inhibitors 21

1.4.3. Bisubstrate Inhibitors 23

1.5. Aims and Objectives 26

1.6. Release of PRMT1 and CARM1 Structures in the PDB from 2014 - 2018 28

1.6.1. PDB Entries for PRMT1 to Date 28

1.6.2. PDB Entries for CARM1 to Date 29

CHAPTER 2: Materials and Methods 33

2.1. Cloning 33

2.1.1. General 33
2.1.2. Restriction Enzyme Cloning 33
2.1.3. Site-directed Mutagenesis 36
2.1.4. Glycerol Stock Preparation 37

2.2. Protein Expression 42

2.2.1. Preparative Protein Expression 42
2.2.2. Analytical Protein Expression 43
2.2.3. Test Expressions of 2GKG-CARM1-His 44
2.2.4. Test Expressions of GST-mCARM1(FL) 45

2.3. Protein Purification 45

2.3.1. General 45
2.3.2. Cell Lysis 46
2.3.3. Nickel Affinity Chromatography 47
2.3.4. GST Affinity Chromatography 47
2.3.5. MBP-Affinity Chromatography 47
2.3.6. Anion Exchange Chromatography 48
2.3.7. Gel Filtration 48
2.3.8. Co-purification with Histones 49

2.4. SDS-PAGE 49

2.5. Determining Protein and Inhibitor Concentrations 50

2.6. Western Blot 51

2.7. Crystallography 53

2.7.1. Crystallisation 53
2.7.2. Cryoprotection 55
2.7.3. Data Collection, Structure Determination and Refinement 56

2.8. Radiometric Activity Assays 57

2.8.1. Initial Test for Activity of Recombinant Constructs 57
2.8.2. Linear Range Determination 58
2.8.3. Determination of Histone K_m 59

2.8.4. IC_{50} Assays *60*
2.8.5. Inhibitor Screening Assays *61*

2.9. Isothermal Titration Calorimetry 62

2.10. Synthesis of Putative FRET Probes 65

CHAPTER 3: Result and Discussion 83

3.1. Summary of PRMT and Histone Constructs 83

3.3. PRMT1: Purification and attempted Crystallisation 88

3.3.1. PRMT1 Construct Design *88*
3.3.2. PRMT1 Purification *88*

3.4. CARM1: Purification and Crystallisation 101

3.4.1. CARM1 Construct Design *101*
3.4.2. CARM1 Purification *101*
3.4.3. Removal of His-tag and Addition of C-terminal Residues *105*
3.4.4. Attempts at Improving CARM1-His Yield *107*
3.4.5. Co-purification of CARM1 and GB-H3 ..... *111*
3.4.6. Attempted Expression and Purification of Full Length CARM1 Constructs *113*

3.5. CARM1 Crystallisation 123

3.5.1. CARM1-SAH Co-crystallisation *123*
3.5.2. CARM1-inhibitor Co-crystallisation *127*

3.6. Effect of Linker Length on Inhibitor Affinity 131

3.6.1. Isothermal Titration Calorimetry *131*
3.6.2. Radiometric (^3H) Activity Assays *140*
3.6.3. Homology Model of Full Length CARM1 ..... *153*

3.7. Replacement of Guanidine with Aromatic Isosteres 157

3.7.1. Binding of Aromatic Inhibitors *157*
3.7.2. Crystal Structures of CARM1 with Aromatic Inhibitors *165*
3.7.3. Mutagenesis Studies *172*
3.7.4. Aromatic Inhibitors: Conclusions *185*

3.8. Morpholino Inhibitors 187

3.9. Synthesis of putative FRET probes 191

3.9.1. Design of FRET Probes 191

3.9.2. Synthesis of Putative FRET probes 193

3.10. Comparison of Results in light of structures recently added to the PDB 201

3.10.1. Guanidine-containing bisubstrate inhibitors in the PDB compared with UM079, UM249, UM251 and WH5C 201

3.10.2. Aromatic-containing bisubstrate inhibitors in the PDB compared with UM305, AA175 and AA183 205

CHAPTER 4: Conclusions and Future Work 208

CHAPTER 5: References 210

CHAPTER 6: Appendix 221

6.1. Summary of CARM1 Structures available in the PDB at the time of writing (September 2018) 221

6.2. Alignment of PRMT1 and CARM1 Sequences Encoded by our Template Plasmids with Relevant Sequences 230

6.2.1. PRMT1 230

6.2.2. CARM1 231

6.3. Additional Analysis of Crystal Structures 232

6.3.1. Guanidine-containing Inhibitor B-factors 232

6.4. Protein and Ligand Concentration Determination 233

6.4.1. Bradford Assay to Verify ε_{260} Determined using Expasy Prot Param 233

6.4.2. Nanodrop Linearity Check 234

6.5. Crystallographic Data Processing Statistics 236

6.5.1. Final Structures 236

6.5.2. Additional Data Set Statistics (Refinement in Progress) 239

6.6. HPLC Traces of Putative FRET Probes Prior to Deprotection 240

6.7. Evidence for Synthesis of EG047 and EG048 241

6.7.1. EG047 241
6.7.2. EG048 243

CHAPTER 1: **Introduction**

1.1. **Arginine Methylation and PRMTs**

Arginine methylation is a posttranslational modification involved in numerous cellular processes including chromatin function,[1] RNA processing,[2] and DNA-damage response.[3] Methylated arginines are as prevalent in the human proteome as phosphorylated and ubiquitinated residues,[4] and have a similarly rich network of associated signalling pathways.[5] The proteins responsible for performing this modification are the protein arginine methyltransferases (PRMTs, E.C. 2.1.1.125). Nine confirmed human PRMTs have been identified in the human genome to date. These are typically classified according to the methylation pattern they deposit on the target arginine (Figure 1).[6]

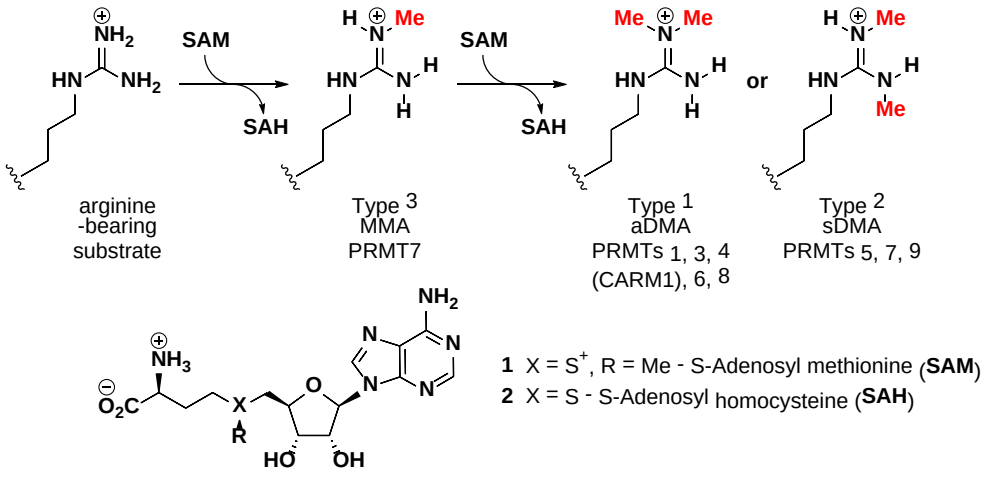


Figure 1: Arginine methylation patterns deposited by PRMTs. The PRMTs are classified as type 1, 2 or 3 according to whether they generate monomethylarginine (MMA), symmetric dimethyl arginine (sDMA) or asymmetric dimethylarginine (aDMA)

At a molecular level, installation of each arginine methyl group removes a hydrogen bond donor site and increases the hydrophobicity and Van der Waals radius of the guanidinium group, while the charge is unaffected. These changes influence the substrate protein's interactions with binding partners.[7] The primary readers of arginine methyl markers are proteins containing Tudor domains.[5a] Tudor domains harbour an 'aromatic cage', which confers methylation-state-dependent binding of arginine residues (Figure 2).[8] Individual PHD and WD40 domains have also been shown to bind methyl arginines, but this ability is limited to specific examples.[9]

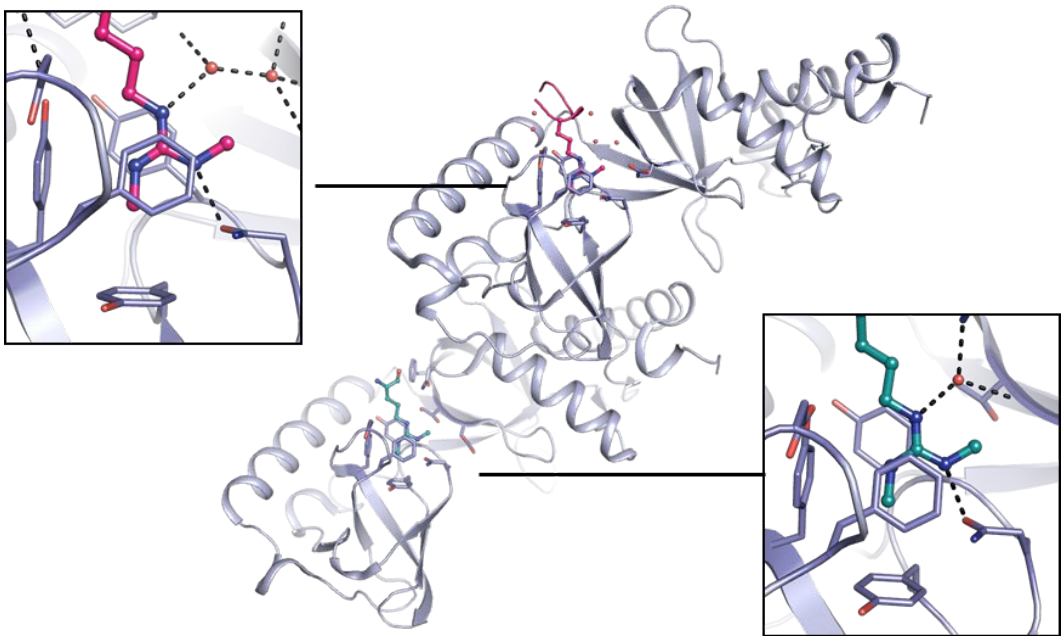


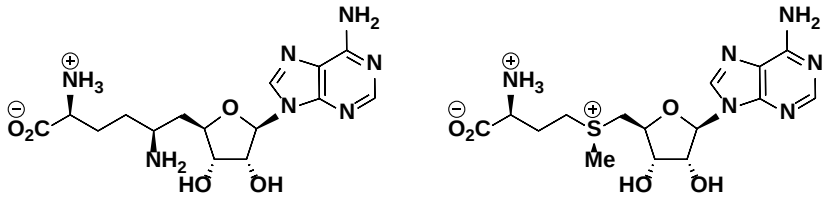
Figure 2: sDMA binding sites in a published crystal structure of an 'extended Tudor domain' from SND1 (PDB code 3OMC).

The nine PRMTs do not show obvious redundancy, with mice knockouts of each gene displaying distinct phenotypes; however PRMT1 knockout results in substrate scavenging by other PRMTs.[10] Chemogenetic studies suggest that there may be up to 35 additional PRMTs yet to be discovered and characterised.[11]

All PRMTs contain a 'catalytic core' of approximately 300 amino acids (Figure 4, page 5), which comprises (from N to C termini): N-terminal helices, a Rossman fold, and a β-barrel interrupted by a helical dimerisation arm. Several PRMTs have additional N- and C-terminal domains (Figure 4, page 5).[12] Most PRMTs form head-to-tail dimers, with dimer formation mediated by interaction between hydrophobic patches at the end of the dimerisation arm of each monomer and the Rossman fold of its partner. The size of the central dimer cavity varies depending on the length of the dimerisation arm.

S-Adenosyl methionine (SAM) binds in a pocket formed by the Rossman fold and N-terminal helices. This binding of SAM to a pocket composed of a Rossman fold plus an additional element is a common feature among methyltransferases, including non-SET-domain-containing lysine methyltransferases, DNA methyltransferases and small molecule methyltransferases.[13] Since SAM is unstable in aqueous solutions near physiological pH,[14] pan-PRMT-inhibitor sinefungin (**3**, Figure 3, page 4) and SAH (**2**, Figure 1, page 1) are commonly used as a surrogates for SAM in structural studies of methyltransferases (e.g. PDB entries 5DX0, 5YF1, 5YLT, 5WWR, 5EMK, 3SWR).[15] Structural data show that SAH/sinefungin binding to

PRMTs results in movement of the N-terminal helices to bury the bound cofactor and form the substrate binding groove.[12, 15a, 16] SAH and SAM have also been shown to affect the dissociation constants of PRMT1 and PRMT6 homodimers.[17]



Sinefungin (**3**) SAM (**1**)

Figure 3: Structure of SAM-mimic sinefungin with SAM shown for comparison

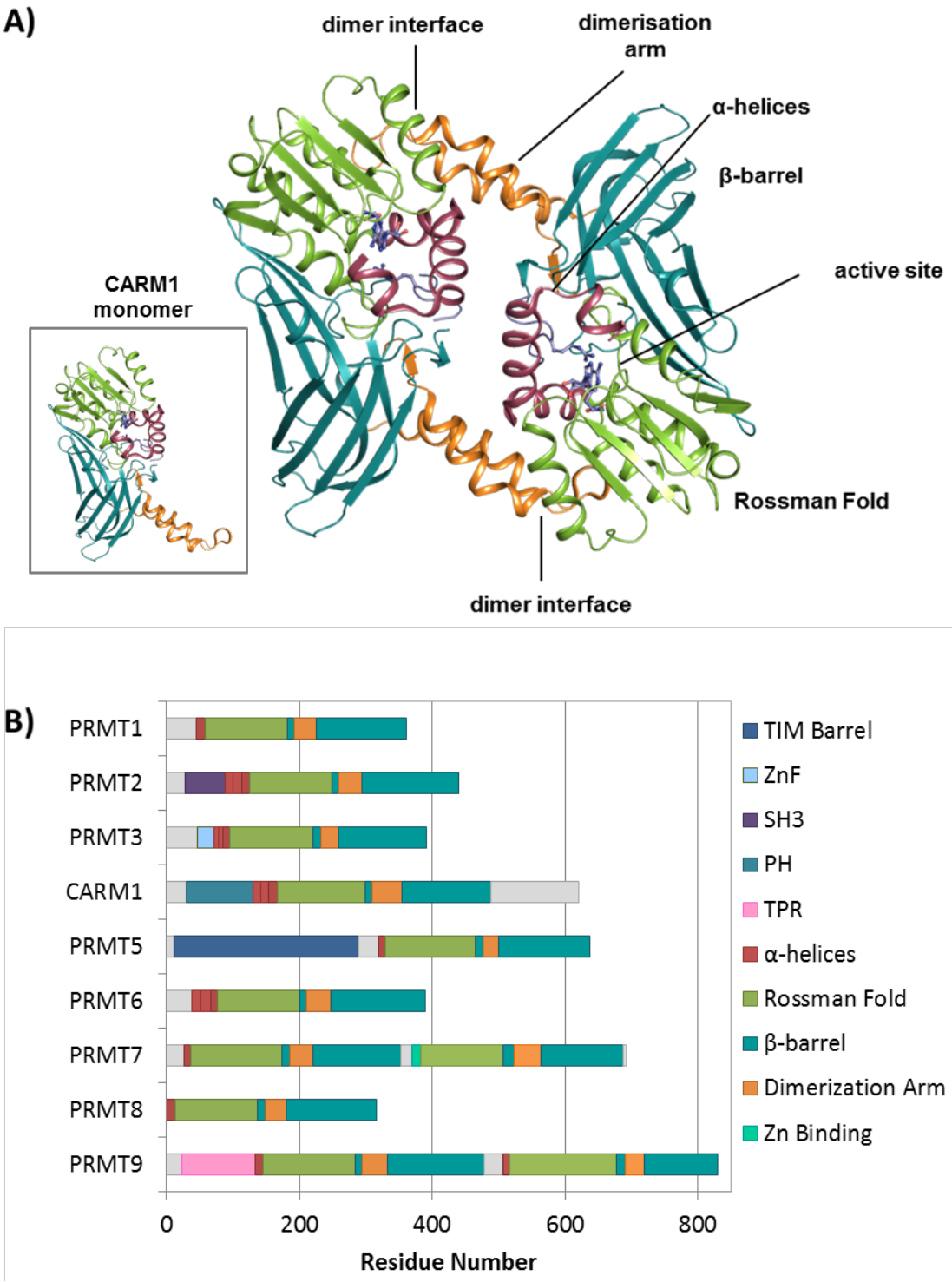


Figure 4: Domain architecture of the PRMTs. (A) Crystal structure of the CARM1 catalytic core dimer (PDB code 5DX0) with H3 substrate and SAM-mimic sinefungin bound. The domains are colour coded to correspond to (B). (B) Domains of the human PRMTs. Domain classifications are based on a combination of structural (X-ray) data and sequence homology. Regions that are unstructured, or for which the structure is unknown are shown in grey.

The catalytic cores of PRMTs 1-8 are structurally well conserved (Figure 5). In particular, the active site (between the Rossman fold and β-barrel domains) shows a high degree of structural and sequence similarity (Figure 6 and Figure 7). How each PRMT recognizes its specific set of substrates is still unclear. Structural evidence to date suggests that PRMT-substrate binding occurs primarily through interactions with the substrate's backbone.[15a] Comparison of *apo* and SAH-bound CARM1 structures suggests that SAM binding may be required for proper formation of the substrate binding site.[16a, 18] This observation is in line with more recent kinetic observations (see section 1.3.2, page 15).

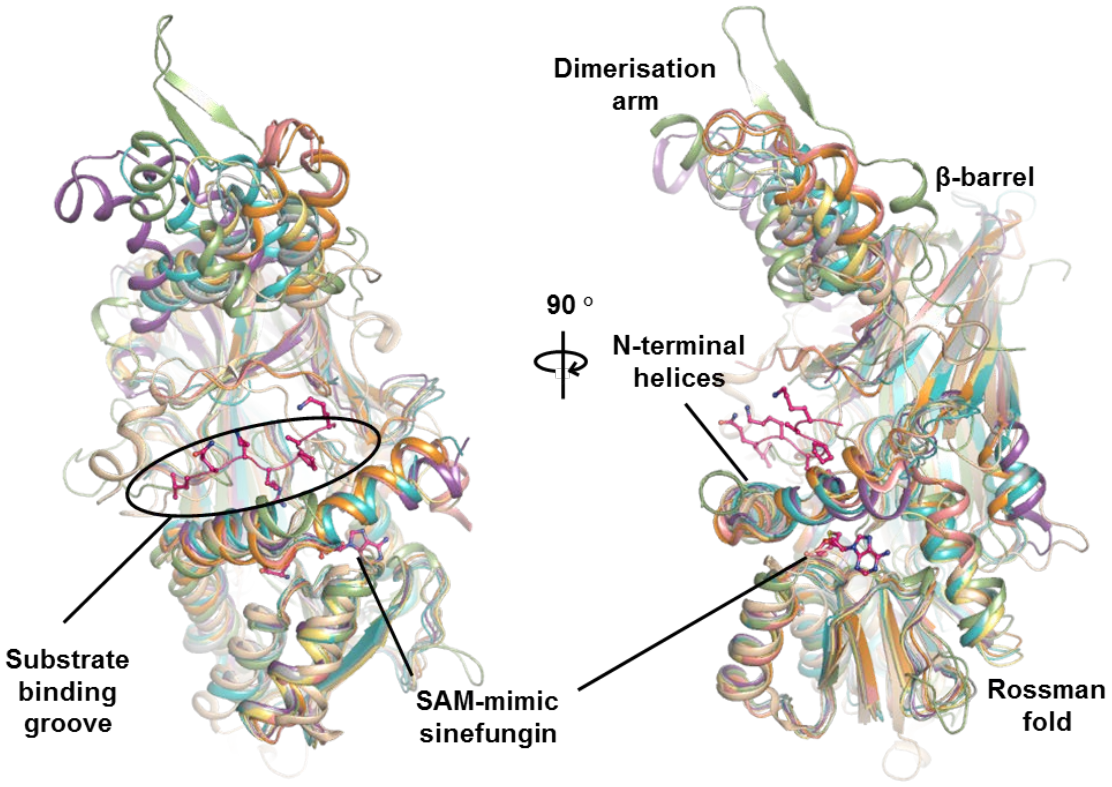


Figure 5: Superposition of PRMT catalytic cores. Generated using PDB entries 1or8 (PRMT1), 3SMQ (PRMT3), 5DXo (CARM1), 5C9Z (PRMT5), 5EGS (PRMT6), 5DST (PRMT8) and homology models of PRMT2 and PRMT7 generated using the PHYRE2 server.[19] No structure is available for PRMT9.

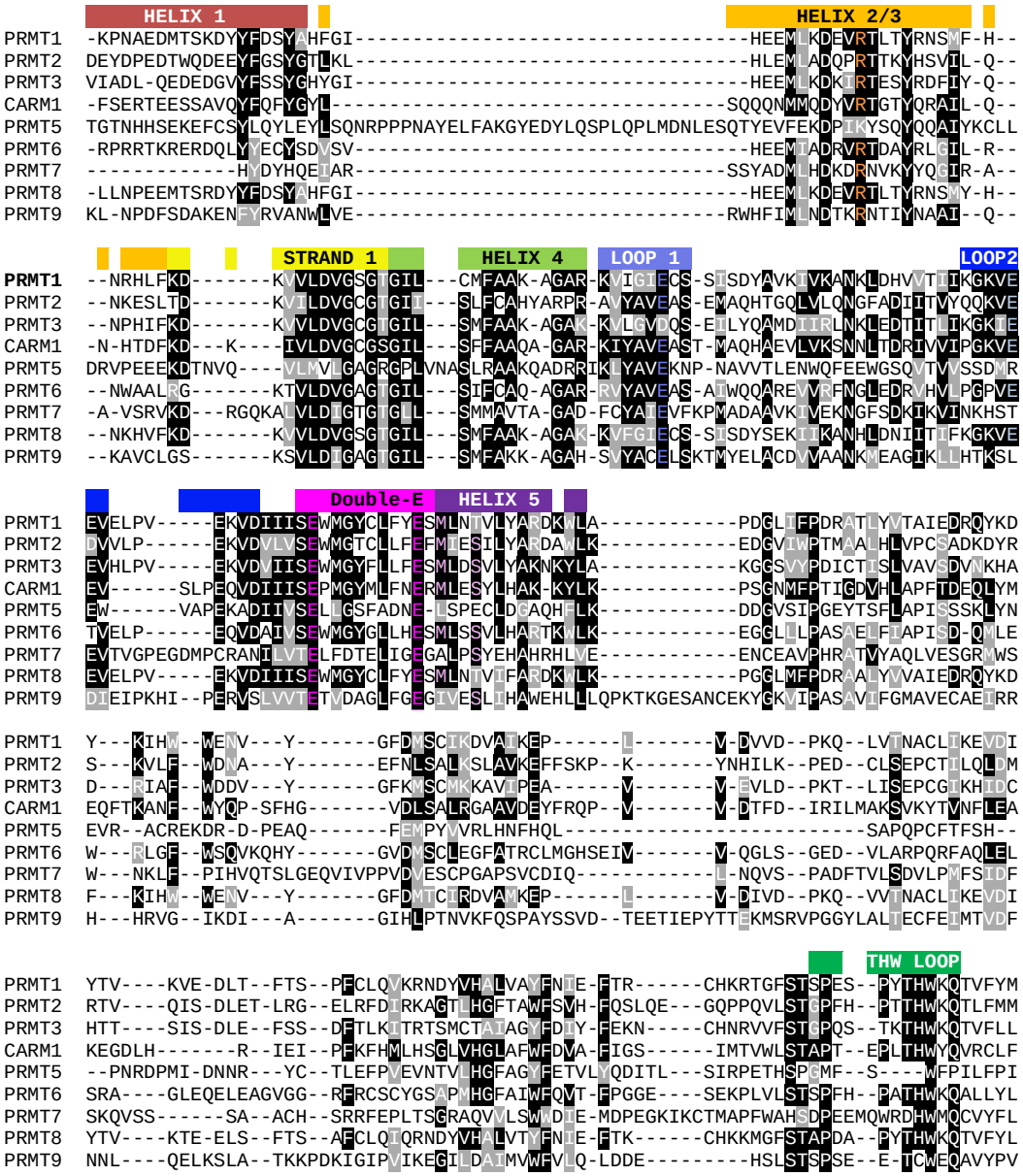


Figure 6: Sequence alignment of human PRMTs (canonical isoforms from Uniprot). Alignment was generated using ClustalW.[20] Consensus sequence determination and shading were performed using ExPASy BOXSHADE. Residues shaded in black match the consensus sequence; those shaded in grey are similar to the consensus residue. Structural elements that comprise the active site are highlighted above the alignment and are colour coded to match the structures shown in Figure 7 (below).

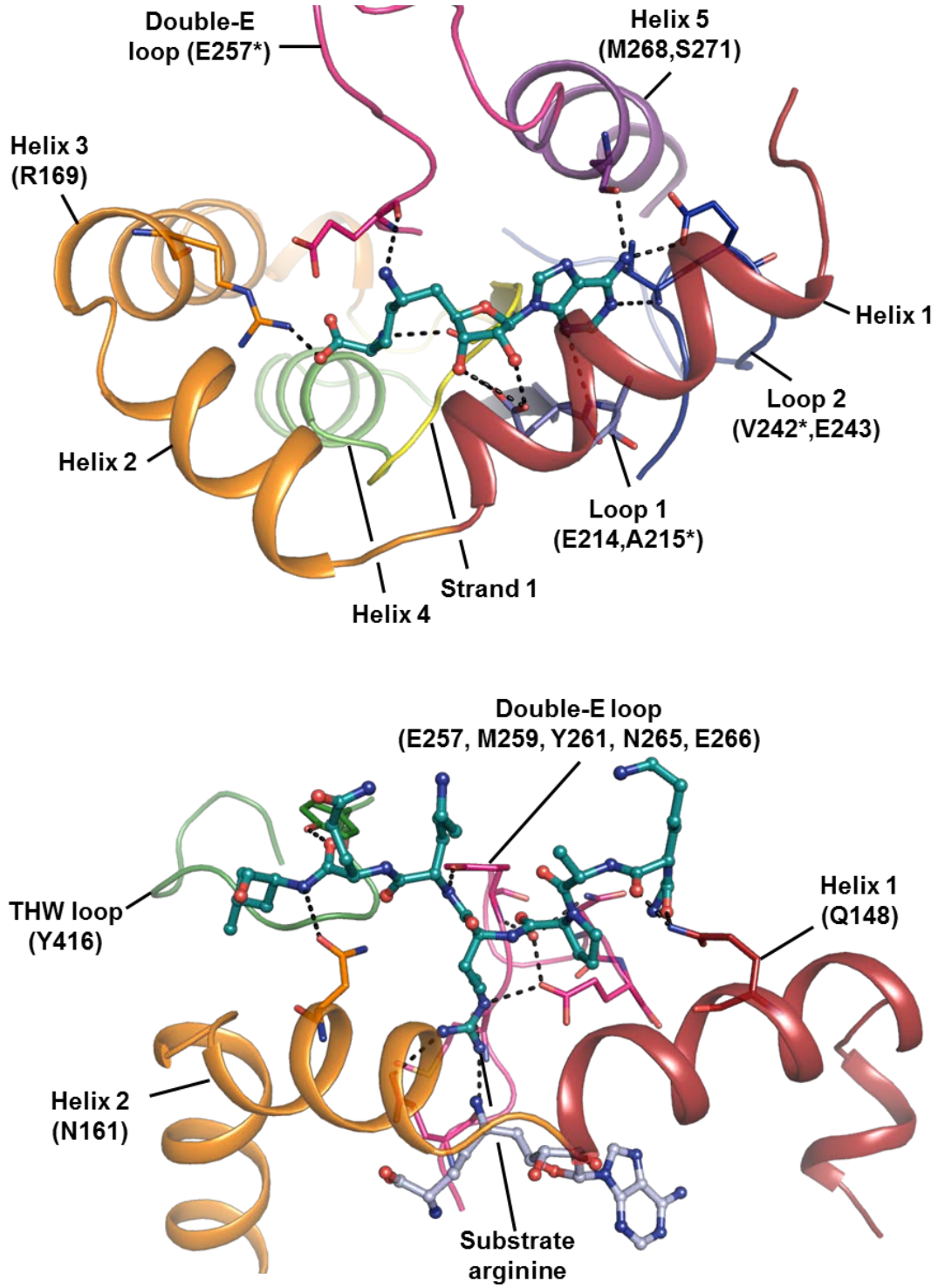


Figure 7: Structural elements of the SAM (top) and H3 (bottom) binding site of CARM1 (PDB code 5DX0).[15a] Structural elements are coloured to correspond to the sequence alignment in Figure 6. Polar contacts found using PyMOL are shown as dashes. Key active site residues are shown as lines. SAM-mimic sinefungin (3, Figure 3) and H3 are shown as ball-and-sticks. Asterisks indicate residues whose backbone atoms make contact with the substrate(s).

Detailed mechanistic studies have been carried out to decipher the catalytic mechanism of PRMT1, and identify the roles of key residues within the active site.[21] pH rate and solvent isotope effect studies suggest that methyl transfer is primarily driven by proximity, as opposed to general acid/base catalysis, with N-methylation occurring prior to deprotonation (Figure 8).[21]

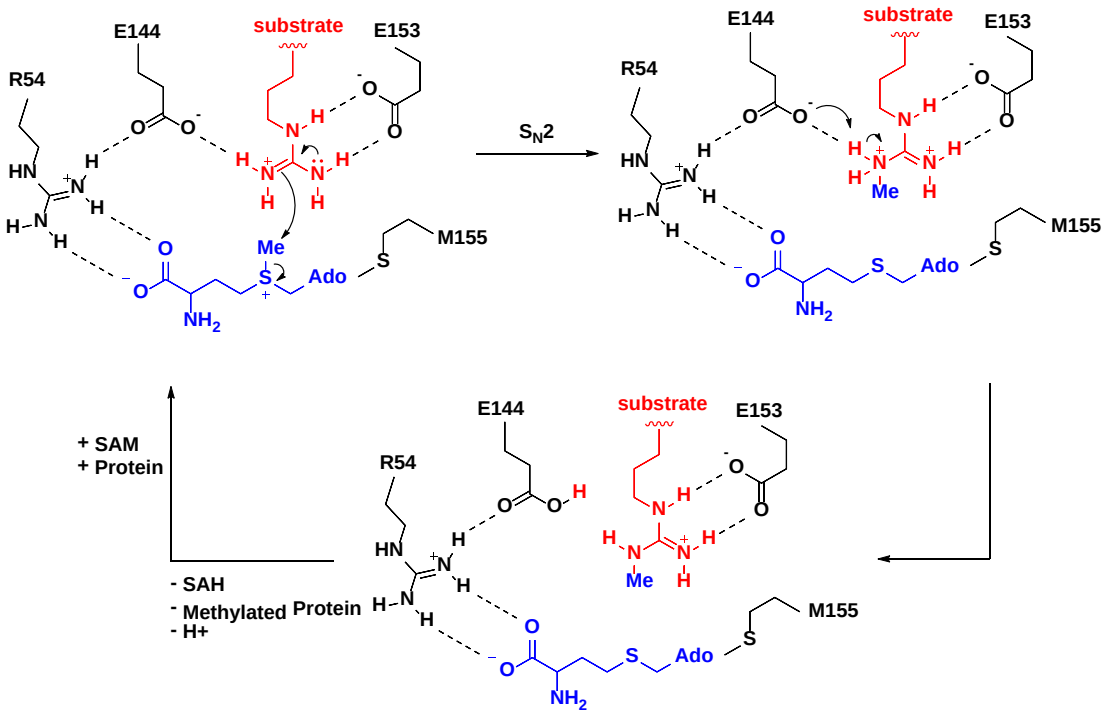


Figure 8: Proposed mechanism of PRMT1-catalysed arginine methylation (Rust *et al.*)[21] The numbering of the original paper has been used, which corresponds to the rat sequence (UniProtKB - Q63009 and see Appendix 6.1.1 for alignment with human sequence)

Mutation of R54 (R169 in CARM1, Figure 7, page 8) to lysine resulted in decreased enzymatic turnover, primarily through a reduction in k_{cat}, suggesting its primary role is correctly orientating E144. Surprisingly, mutation of E100,

which hydrogen bonds to the cis 1,2-diol in the ribose group of SAM/sinefungin in the crystal structure (E214 in CARM1, Figure 7, page 8), did not significantly affect PRMT1 activity. M155 (M268 in CARM1, Figure 7, page 8), was found to be important for productive SAM binding, with its mutation to alanine resulting in a 18-fold increase in SAM K_m and 36-fold decrease in k_{cat}/K_m. It has been postulated that M155 is responsible for regiospecific asymmetric dimethylation by type 1 PRMTs (Figure 1, page 1), however Rust *et al.*'s M155A mutant produced exclusively mono- and asymmetric dimethylarginine. The same result was obtained by Gui *et al.*, who additionally showed that mutation of a methionine residue at the opposite end of the active site (M48) is also insufficient to allow symmetric dimethylation.[22]

The two glutamate residues (E144 and E153 in PRMT1/E257 and E266 in CARM1) proposed to orientate the substrate arginine, are the only two active site residues that are strictly conserved throughout the nine human PRMTs (Figure 6, page 7), and are part of the conserved 'double E loop' within the Rossman fold. Mutation of either of these residues resulted in decreased activity. The most dramatic effects were seen for E153. Mutation to alanine resulted in greatly decreased k_{cat}, which is consistent with its proposed role in enhancing the nucleophilicity of the substrate guanidinium group. Interestingly, double mutation of E144 and E153 retained more activity than single mutants of either of these residues.[21] The authors propose this as evidence for the catalysis being primarily proximity driven.

1.2. **PRMT1**

PRMT1 - the founding member of the PRMT family - was discovered in 1968 as a component of calf thymus nuclear extract.[23] PRMT1 is responsible for >90% of asymmetric arginine methylation and its knockout is lethal to mice embryos.[10] Seven splice variants of PRMT1 have been identified to date.[24] These differ only at the N-terminus, but have distinct expression and localisation profiles. Biological processes mediated by PRMT1 include epigenetic gene regulation,[10a] protein trafficking, DNA repair,[25] signal transduction and pre-mRNA splicing.

1.2.1. Structure

Despite its status as the primary arginine methyltransferase, structural studies on PRMT1 are limited. The first crystal structure of PRMT1 was deposited in the Protein Data Bank (PDB) in 2003, and revealed that PRMT1 adopts a head-to-tail dimer conformation (as for CARM1, Figure 4, page 5).[26] Removal of the dimerisation arm results in abrogation of cofactor binding and methyl transferase activity *in vitro*, implying that the observed dimerisation is necessary for catalysis, and therefore physiologically relevant (as opposed to an artefact of crystallisation). The low pH (~4.7) under which the crystals were obtained and poorly defined electron density of the co-crystallised substrate limits the utility of these structural data. The only further structure of PRMT1 available is an M48L active site mutant, which showed altered substrate selectivity and was crystallised at pH 9.0 limiting its relevance for docking studies and structure-based inhibitor design.[22]

1.2.2. Mechanism

Until recently the kinetic mechanism of PRMT1 had not been reliably established, with conflicting data regarding the enzyme's processivity and order of substrate/co-factor binding (i.e. sequential ordered or rapid equilibrium random mechanisms).[21, 27] Earlier this year, Brown *et al.* performed an in depth study and concluded that PRMT1 kinetics vary depending on the substrate sequence; that the degree of processivity is dependent on several factors (primarily enzyme and SAM concentrations, but also substrate identity); and that the mechanism is sequential ordered bi-bi (Figure 9, below).[28] Further, data showed that while a histone substrate was able to bind PRMT1, this binary complex could not subsequently bind SAM (i.e. dead-end EB complex formation). Brown *et al.* propose that, based on the conservation of active site residues, *all* PRMTs are likely to follow the same kinetic mechanism.

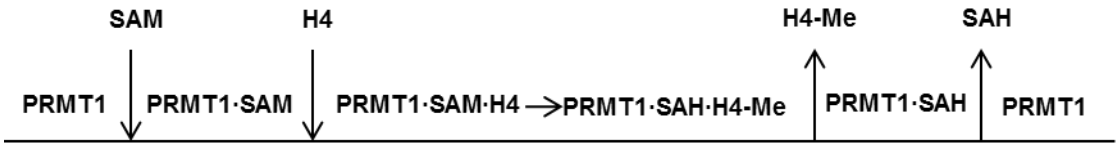


Figure 9: Kinetic mechanism of PRMT1: Sequential ordered bi-bi.

1.2.3. Disease Links and Target Validation

The dysregulation of PRMT1 has been associated with a host of diseases. PRMT1-v1 and v2 levels have been shown to correlate positively with breast cancer tumour grade and colon cancer aggressiveness.[29] PRMT1 is also implicated in epithelial-mesenchymal transition (EMT), cancer cell migration and metastasis.[30] SiRNA knockdown of PRMT1 has been shown to suppress the growth of bladder and lung cancer cells.[31] PRMT1 knock-down has also been shown to inhibit proliferation and induce apoptosis in four glioma cell lines (T98G, U87MG, U251, and A172).[32] Viral proteins including the HIV-1 proteins Rev and Tat, and the hepatitis C protein NS3 contain glycine arginine rich (GAR) motifs which are methylated by PRMT1 (or PRMT6), suggesting involvement of arginine methylation in viral pathogenesis. Inhibition of arginine methylation has been found to prevent replication of the virus hepatitis delta.[33] Arginine methylation has also been linked to pulmonary and cardiovascular diseases, multiple sclerosis and spinal muscular atrophy.[34]

1.3. **CARM1**

CARM1 (co-activator associated arginine methyltransferase 1, also known as PRMT4) was first identified using a yeast two hybrid experiment as a binding partner of GRIP1 (a transcriptional coactivator).[35] CARM1 knockout mice are born small and die shortly after birth.[35] Unlike other PRMTs, which typically methylate GAR (glycine arginine rich) motifs, CARM1 prefers PGM (proline, glycine, methionine)-rich regions.[36]

CARM1 is involved in transcriptional regulation. Its methylation of histone H3 at R17 and R26 is believed to promote transcription through interplay with histone acetylation.[37] CARM1's non-histone substrates also suggest its involvement transcriptional regulation. These substrates include transcriptional co-activators p/CIP and subunits of the SWI/SNF nucleosome remodelling complex.[38,39]

1.3.1. Structure

The longer dimerisation arm of CARM1 (Figure 4, page 5) results in a larger central cavity than PRMT1.[16a] Phosphorylation of S228 (located at the tip of this dimerisation arm) has been shown to inhibit dimerisation, with concomitant loss of CARM1 activity.[40] CARM1 also differs from the other PRMTs in that it possesses additional N- and C-terminal domains flanking the catalytic core (Figure 4, page 5).

N-terminal to the catalytic core is a plextrin homology (PH) domain. The structure of this domain in isolation has been solved (PDB code 2OQB), however when Troffer-Charlier *et al.* crystallised a construct comprising the catalytic core and N-terminal PH domain, no electron density was visible beyond the catalytic core.[18] A recent study by Shishkova *et al.* found this domain to be important for substrate recognition and binding.[36] Known CARM1 substrates were shown to interact with CARM1's PH domain using fluorescence polarisation. A truncated GST-fusion protein of CARM1 that lacks residues 1-138 was found to be less active, and bind substrates less effectively than its full

length counterpart.[36] In contrast to the conclusions of Shishkova *et al.*, Yue *et al.* have found that $CARM1_{146-489}$ (which lacks the PH domain) interacts with the AD2 domain of GRIP1 – a known CARM1 substrate - to a similar extent as full length CARM1. Further, the GRIP1 domain did not interact with $CARM1_{24-153}$ (the PH domain).[16a] The precise role of the N-terminal PH domain therefore remains unknown.

No structural information is available for the C-terminal domain of CARM1. Yue *et al.* found constructs with and without the C-terminal domain ($CARM1_{146-489}$ and $CARM1_{146-584}$) to interact with GRIP1 (AD2) to the same extent, indicating that the C-terminal domain is not essential for this interaction.[16a] Removal of residues 471-489 from the catalytic core (i.e. $CARM1_{146-471}$) was found to abrogate H3R17 methylation, though the effect on other methylation sites was not reported.[16a]

Teyssier *et al.* also showed that the catalytic domain of CARM1 was sufficient for its GRIP1 (AD2) interaction.[41] They also found that two constructs of the C-terminal domain ($CARM1_{461-608}$ and $CARM1_{501-608}$) fused to the Gal4 DNA binding domain induced expression of a reporter gene, indicating that the C-terminal domain is capable of autonomous autoactivation.[41]

1.3.2. Mechanism

As mentioned above, comparison of the CARM1 *apo* and SAH/sinefungin-bound structures shows that CARM1 undergoes structural changes upon SAH or

sinefungin binding. These changes are thought to enable substrate binding.[16a] Notably, attempts to co-crystallise peptide substrates with CARM1·sinefungin but *not* CARM1·SAH have been successful, suggesting structural differences between the two binary complexes.[15a, 16a] Further, several small molecule inhibitors have been reported to bind to CARM1 only in the presence of SAH (or display stronger binding with SAH present).[16b, 42] Together, these data suggest that CARM1, like PRMT1, may follow a sequential mechanism. In contrast, an extensive kinetic study by Jacques *et al.* (2016) has shown that CARM1 methylates histone H3 peptides distributively *via* a random sequential kinetic mechanism (Figure 10).

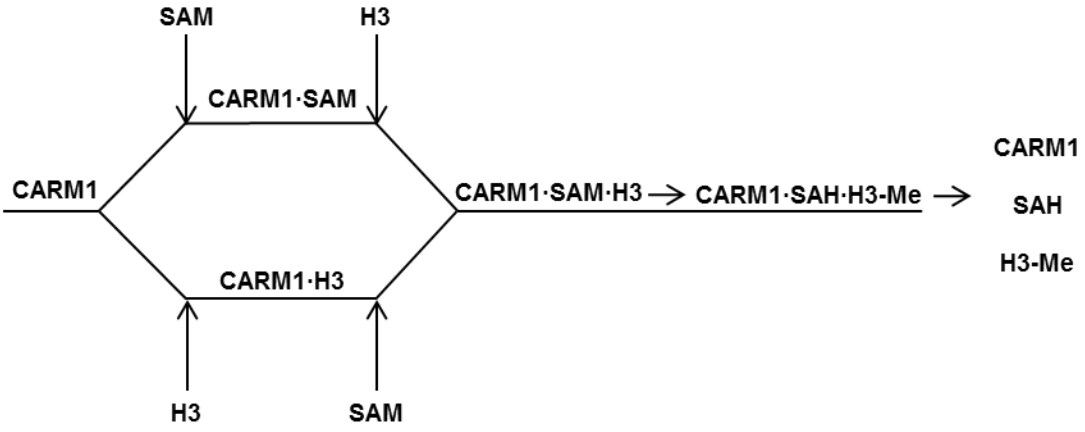


Figure 10: CARM1 kinetic mechanism proposed by Jacques *et al.* based on kinetic studies.[43]

The authors highlight the use of full length CARM1 for their kinetic analysis, in contrast to the truncated constructs used for all CARM1 crystal structures as a

possible cause for the apparent discrepancy between their results and structural data.[43]

1.3.3. Disease Links and Target Validation

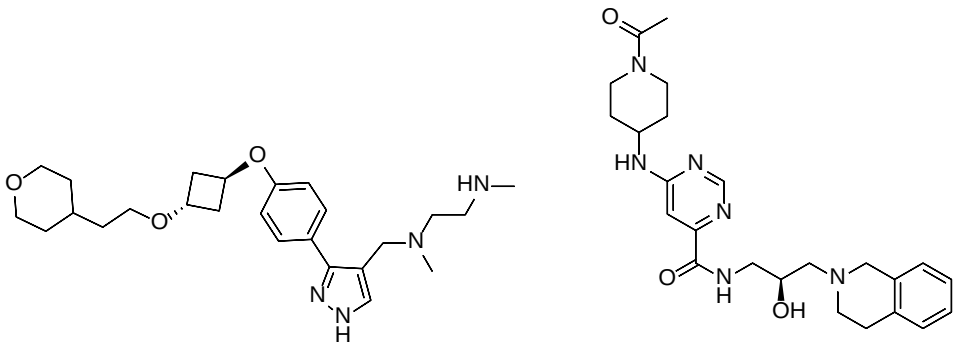
CARM1 is overexpressed in numerous cancer cell types, including prostate,[44] colorectal,[45] lung,[46] breast,[47] liver.[48] ShRNA CARM1 knockdown has been shown to inhibit liver cancer cell proliferation.[48] CARM1 has also been implicated in acute myeloid leukemia (AML).[49] Greenblatt *et al.* recently reported that CARM1 knockout, knockdown or small-molecule-inhibition in AML mouse models impaired AML initiation and proliferation.[50]

1.4. PRMT Inhibitors

Potent, selective inhibitors of individual PRMT members are valuable for use as chemical probes and therapeutics. In the context of chemical probes, selective inhibitors can be used to examine the effects of inhibiting the target enzyme in cells, tissues or animals. This in turn can be used to infer information about the biological role of the target enzyme, and whether inhibiting it may have therapeutic effect in diseased cells or organisms (i.e. target validation). In this application, probes are complementary to other techniques such as genetic knock-down or inhibitory RNAs (which bind to mRNA transcripts and prevent translation).[51] Further, selective compounds with fluorophores or other tags are useful to probe the tissue-distribution or subcellular localisation of target

enzymes. In all these instances, binding of the probe to unintended targets can confound the results. In target validation, for example, off-target binding would make it difficult to establish definitive links between the knockout of a particular enzyme and the resultant phenotypic readout. In drug discovery, off-target inhibitor binding may result in undesirable side-effects (though, notably, many MHRA/FDA-approved drugs do not have proven primary targets or mechanisms of action).[52] For a discussion of the importance of highly selective chemical probes, and the desirable characteristics of probes *versus* therapeutics, see Arrowsmith *et al.*[53]

The relatively high sequence conservation among the PRMTs makes the design of selective inhibitors challenging, nonetheless isozyme- and type- selective inhibitors have been emerging over the past 10 years (recently reviewed by Hu *et al.*[54]). A PRMT5 inhibitor GSK3326595 (also known as EPZ015938 - **5**, Figure 11) is currently in phase-1 clinical trials for patients with solid tumours and non-Hodgkin's lymphoma, highlighting the potential for therapeutic PRMT inhibition.[55] A tool compound for PRMT6 (EPZ020411, **4**, Figure 11) has also been reported.[56] This compound was shown to be 10 and 20-fold selective for PRMT6 over PRMTs 8 and 1, and over 100-fold selective over a panel of histone methyltransferases including PRMTs 3, 4, 5 and 7. It also showed good bioavailability in rats, making it suitable for *in vivo* studies. Both GSK3326595 and EPZ020411 target the substrate binding channels of their respective PRMT targets (as indicated by kinetic[57] and crystallographic[58] data, respectively).



EPZ020411 (**4**)
10 nM *vs* PRMT6

GSK3326595 (**5**)
6 nM *vs* PRMT5

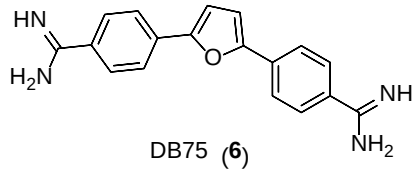
Figure 11: Structures of EPZ020411 and GSK3326595[57]

1.4.1. PRMT1 Inhibitors

There are currently no chemical probes for PRMT1 deposited in the Structural Genomics Consortium (SGC) database (a repository for reliable, validated probes).[59] Many PRMT1 inhibitors have been reported, but poor cellular activity or poorly established selectivity makes them inappropriate for use as PRMT1 probes. Several recent examples are given in Figure 12 and discussed below.

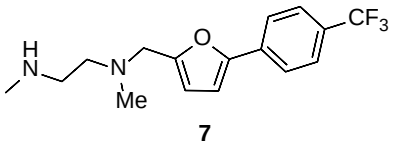
In 2017 Yang *et al.* synthesised a series of PRMT1 inhibitors designed to target Asp84 within the active site. Their most potent PRMT1 inhibitor (**7**, Figure 12) inhibited the proliferation of three cancer cell lines, with the strongest inhibition observed for HepG2 cells, however selectivity of **7** *versus* other PRMTs was not tested.[60] DCPR049_12 (**9**, Figure 12) – a type 1 PRMT inhibitor with selectivity over a panel of other methyltransferases – was also reported in 2017.[61] It inhibits proliferation of leukaemia cells and reduces cellular aDMA levels.[61] Type 1 PRMT inhibitor MS023 (**8**, Figure 12) was reported in 2016. It

inhibits PRMT1 with reasonable selectivity *vs* CARM1 and PRMT6; however concentrations of *ca.* 50 μM were needed to observed anti-proliferative effects in HepG2, MCF7 and THP1 cell lines.[62] DCLX069 (**10**, Figure 12) was reported in 2014. In 2014, Yan *et al.* screened an extensive series of diamidine compounds and found DB75 to be the most potent PRMT1 inhibitor. DB75 exhibited >10-fold selectivity for PRMT1 over CARM1, PRMT5 and PRMT6.



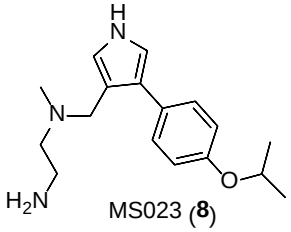
DB75 (**6**)

9.4 μM *vs* PRMT1
>100 μM *vs* CARM1, PRMT5, PRMT6
inhibits leukemia cell proliferation



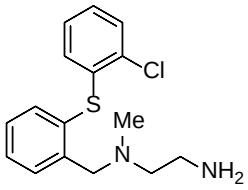
7

8.2 μM *vs* PRMT1
34.5 μM *vs* HepG2 cell line



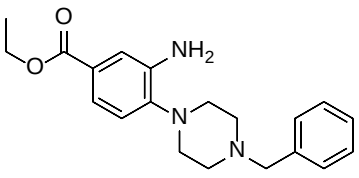
MS023 (**8**)

4 - 119 nM *vs* PRMTs 1, 3, 4, 6 and 8



DCPR049_12 (**9**)

1.1 - 63 nM *vs* PRMTs 1, 3, 4, 6 and 8
>100 μM *vs* PRMT5/panel of PKMTs/DNA-MTs
0.81 μM *vs* MV4-11 cell line (12 days)



DCLX069 (**10**)

17.9 μM *vs* PRMT1
<15% inhibition of CARM1/PRMT6
at 100 μM

Figure 12: Recent examples of PRMT1 inhibitors

Several apparent PRMT1 inhibitors (AMI-1, NS-1, A36 - structures not shown) have also been reported which have subsequently been determined to bind the histone substrate rather than PRMT1.[63] In the absence of crystal structures, binding-mode analysis of the above inhibitors is limited to docking studies, making structure-based optimisation unreliable.

1.4.2. CARM1 Inhibitors

As noted by Drew *et al.,* many CARM1 inhibitors have been reported,[64] but none of these have shown activity "in non-engineered cell lines or at concentrations below 5 μM".[42]

In December 2017, Drew *et al.* reported EZM2302 (**13**, Figure 13) - the first potent (6 nM *in vitro*), selective inhibitor of CARM1 to show *in vivo* anti-tumour activity.[42] The compound was reportedly designed through optimisation of a scaffold based on known arginine methyltransferase inhibitors. Earlier this year TP-064 (**12**, Figure 13) was reported. This compound, generated through optimisation of hits from a high throughput screening (HTS) library, was found to selectively inhibit CARM1 with an IC_{50} <10 nM and >100-fold selectivity over other PRMTs and a panel of lysine and DNA methyltransferases. It also slowed the growth of several multiple myeloma cell lines.[65] Additionally, CARM1-selective inhibitors SKI-72, and its pro-drug SKI-73 (**14** and **15**, Figure 13), have recently been added to the structural genomics consortium (SGC) chemical probes library.[66] Supporting data regarding SKI-72 and SKI-73 have not yet been published.

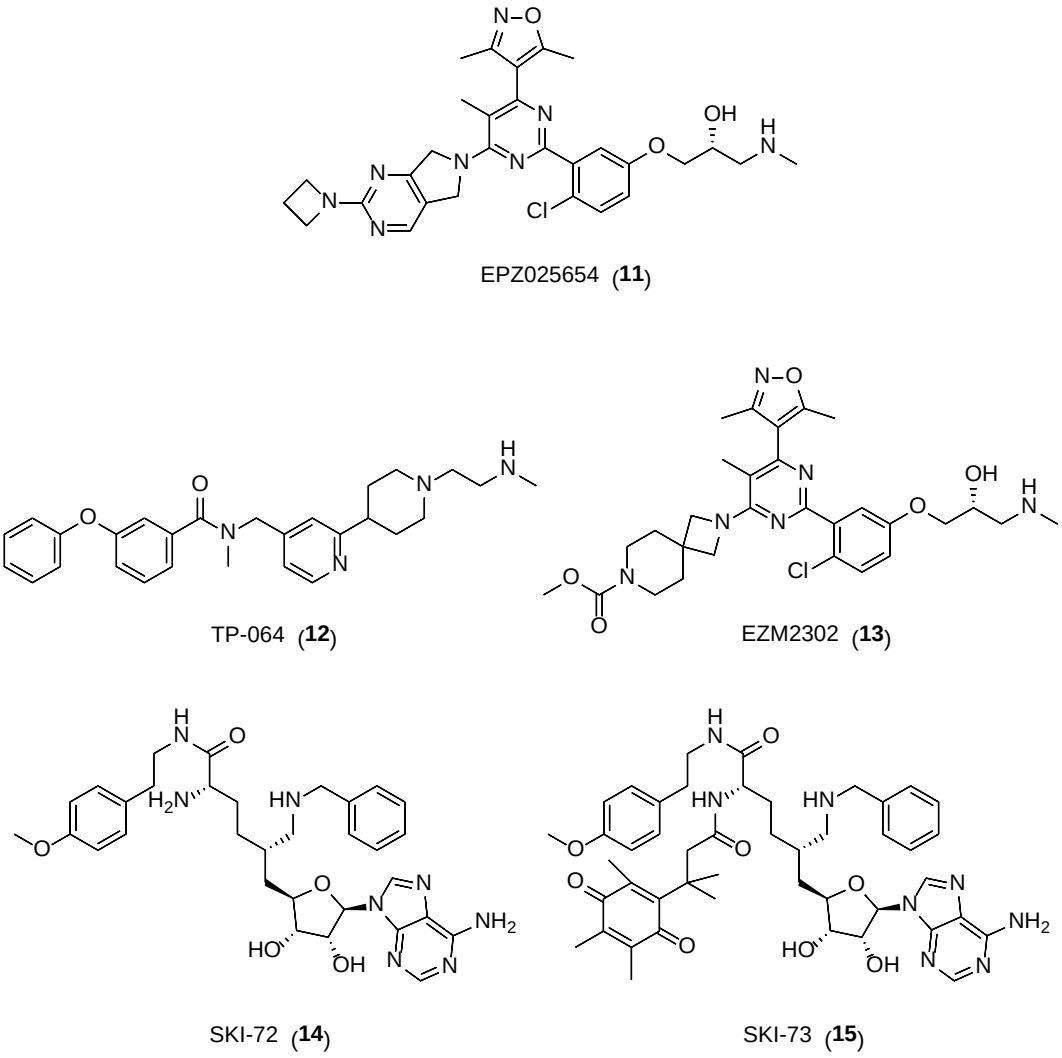


Figure 13: Selective, cell active CARM1 inhibitors TP-064, EZM2302 and SKI-73. SKI-73 is a prodrug of SKI-72.

Interestingly, both EZM2302 and TP-064 require the presence of SAM or SAH to achieve full potency. This prerequisite cofactor binding is consistent with observed structural changes observed on cofactor binding, and the mechanistic studies discussed in Section 1.3.2, page 15. Given the enhanced affinity of CARM1 inhibitors for the CARM1-SAH binary complex (*vs* unliganded CARM1),

bisubstrate inhibitors that engage the cofactor and substrate binding have been explored (see below).

1.4.3. Bisubstrate Inhibitors

Bisubstrate inhibitors aim to simultaneously target both substrate-binding pockets of enzymes whose catalytic cycle involved ternary complex formation. The bisubstrate approach has been successfully applied to other enzymes that simultaneously bind a substrate and co-factor. Notable bisubstrate inhibitors (Figure 14) include topical antibiotic muripirocin (**16**) and anti-tuberculosis pro-drug isoniazid (**17**).

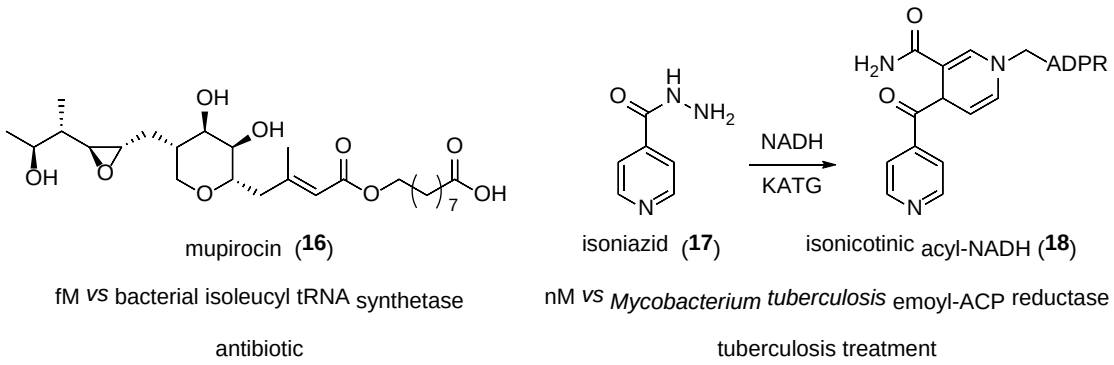
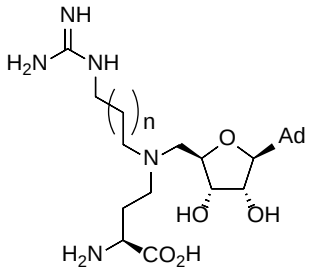


Figure 14: Notable examples of bisubstrate inhibitors – muripirocin and isoniazid. ADPR = adenosine diphosphate ribose.

In the context of PRMTs, the fusion of SAM and substrate mimics has been explored by several research groups. In 2011, Dowden *et al.* reported a series of bisubstrate inhibitors, consisting of a guanidinium moiety attached *via* an alkyl linker to adenosylhomocysteine (**WH3C**, **WH4C** and **WH5C**, Table 1, below) in

which the methionine sulfur has been replaced with a nitrogen atom. This replacement of sulfur with nitrogen avoids problems associated with stereochemical instability at the methylsulfonium centre of SAM (as discussed by Dr. U. Muhsen).[67]

Data indicated that varying the alkyl linker that connects the SAM and guanidine moieties could influence the inhibition of CARM1, while PRMT1 inhibition was unaffected. Notably, inhibitor **WH5C** was found to be a 5.6 μM inhibitor of PRMT1, but was inactive against CARM1 and lysine methyltransferase SET7 (Table 1).



WH3C: n = 1
WH4C: n = 2
WH5C: n = 3

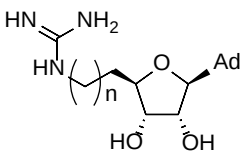
Inhibitor	IC_{50}		
	PRMT1	CARM1	SET7
WH3C	6.2 ± 3.8[68]	13.3 ± 8.7	>100[a]
WH4C	2.9 ± 0.8[68]	>100[a]	>100[a]
WH5C	5.6 ± 4.5[68]	-[b]	-[b]

[a]>50 % activity remaining at 100 μM inhibitor
[b]No inhibition observed at 100 μM inhibitor

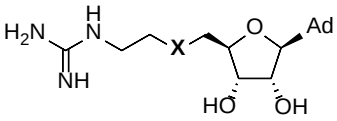
Table 1: IC_{50} data for bisubstrate SAM inhibitors reported by Dowden *et al.*[69]. SET7 is a lysine methyltransferase. Inhibitor codes used do not correspond to those used in the original paper.

In 2015, Van Haren *et al.* reported a closely related inhibitor series (referred to herein as **VH2C**, **VH3C**, **VHS**, **VHO**, **VHN** and **VHX**, Table 2).[70] They explored

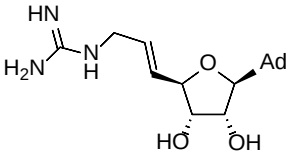
compounds with shorter linkers between guanidine and adenosine moieties, and found that the compound with the shortest linker, **VH2C**, was selective for CARM1 over PRMT1 (Table 2) and PRMT6 (data not shown).



VH2C n = 1
VH3C n = 2



VHS **X** = S
VHO **X** = O
VHN **X** = NH



VHX

Inhibitor	IC_{50}		
	PRMT1	CARM1	G9a
VH2C	11.9 ± 2.77	0.12 ± 0.02	>50[a]
VH3C	1.30 ± 0.38	0.56 ± 0.25	>50[a]
VHS	>50[a]	>50[a]	>50[a]
VHO	>50[a]	>50[a]	3.18 ± 2.67
VHN	>50[a]	>50[a]	>50[a]
VHX	16.96 ± 3.73	0.15 ± 0.05	>50[a]

[a]>50 % activity remaining at 100 μM inhibitor

Table 2: IC_{50} data for bisubstrate inhibitors reported by Van Haren *et al.*[70] G9a is a lysine methyltransferase. Inhibitor codes used do not correspond to those used in the original paper.

In both cases, the absence of crystal structures limited the structural analysis of inhibitor binding to docking studies. Taken together these studies suggest that the bisubstrate approach may be suitable for generating inhibitors selective for individual PRMT family members. Crystal structures of CARM1 and PRMT1 in complex with bisubstrate inhibitors would provide valuable structural information, enabling the fine-tuning of bisubstrate inhibitors to enhance their potency and selectivity.

Based on the success of this bisubstrate approach to PRMT inhibitors, Dr. U. Muhsen synthesised an expanded series of bisubstrate inhibitors during his PhD. The inhibitors used in this thesis (the majority of which were made by Dr. U. Muhsen) are summarised in Figure 15, page 27.

1.5. Aims and Objectives

This project aimed to determine the structural origins of the linker-length effects shown in Table 1 (page 23) and to assess the binding of a series of novel PRMT inhibitors (Figure 15, page 27).

To this end, the project objectives were to determine suitable conditions for the expression, purification and co-crystallisation of PRMT1 and CARM1 with the inhibitors in Figure 15, page 27. Crystallographic studies were envisaged to be supported by thermodynamic binding studies performed using isothermal titration calorimetry (ITC). ITC measures direct protein-ligand binding (i.e. K_d) and therefore avoids the complexities associated with determining inhibitor selectivity using enzyme activity assays.

In a related objective, we set out to examine the possibility of using the bisubstrate SAM inhibitor scaffold to generate PRMT imaging probes. This would entail determination of an appropriate reporter moiety to append to a PRMT inhibitor. Chemical synthesis and binding analysis of the resultant compound would be used to confirm that the appended tag did not interfere with PRMT binding.

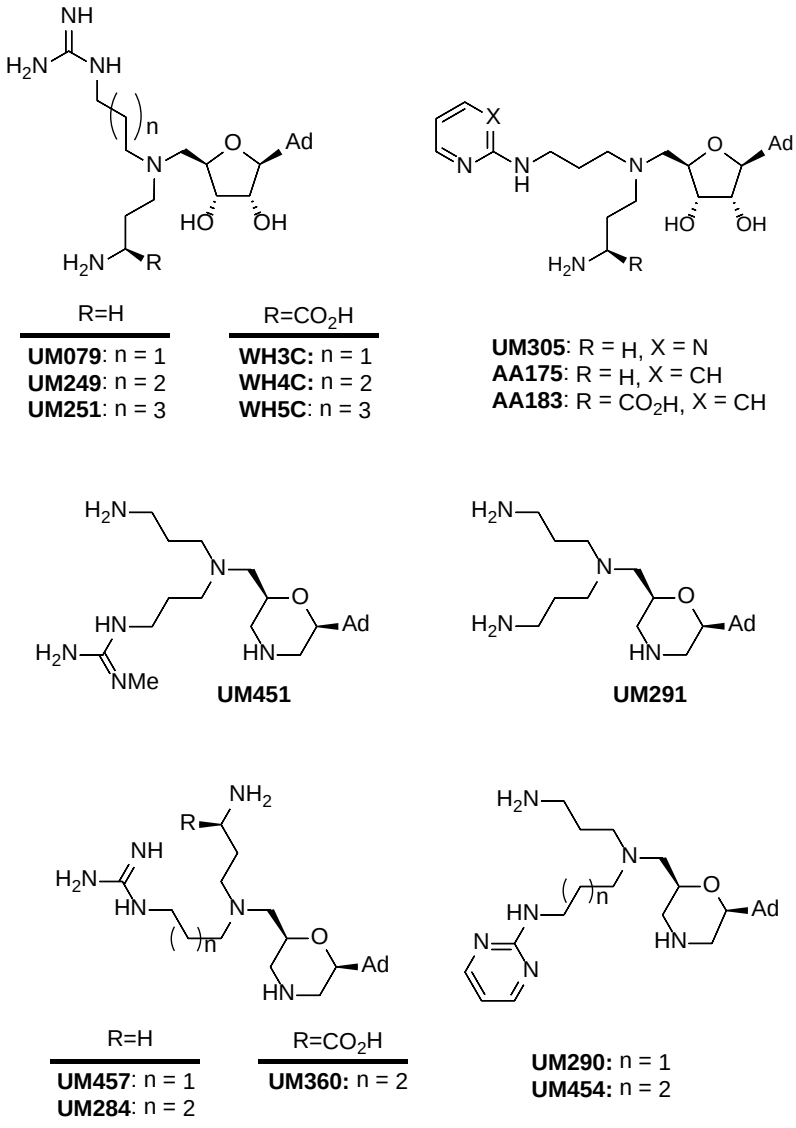


Figure 15: Structure of inhibitors used in this thesis. The name used to refer to each inhibitor throughout this thesis is given underneath.

1.6. Release of PRMT1 and CARM1 Structures in the PDB from 2014 - 2018

During this work the structural data available for PRMT1 remained unchanged, whereas a large number of new CARM1 structures were deposited in the PDB. All CARM1 structures currently available, the structures of bound inhibitors, and their publication dates are summarised in Appendix 6.1 and discussed in sections 1.6.1 and 1.6.2, below.

1.6.1. PDB Entries for PRMT1 to Date

At the beginning of this work (October 2014), three wild type PRMT1 structures were available in the PDB (1OR8, 1ORH, 1ORI).[26] These gave insight into the overall structure of PRMT1, but the low pH at which the crystals were grown likely altered the protonation states of many residues (relative to physiological conditions), limiting the usefulness of the resultant models (see section 1.2.1 for further discussion). A fourth structure (3Q7E),[71] containing an active site mutation was also available, but the active site mutation limits the use of this structure for drug design and inferring mechanistic information. No further structures became available from October 2014 to September 2018.

1.6.2. PDB Entries for CARM1 to Date

At the beginning of this work, 8 structures of the catalytic core of CARM1 were available in the PDB. These include PDB entries 2V74 and 2V7E – the first crystal structures of the catalytic core of CARM1 (with and without SAH bound, respectively);[72] 2OQB – a structure of CARM1's N-terminal PH domain; and 3B3F, 3B3G and 3B3J.[18] 3B3F and 3B3G comprise the CARM1 catalytic domain with and without SAH bound, respectively. 3B3J contains a construct comprising CARM1's catalytic core and N-terminal PH domain, though the PH domain was not modelled due to the absence of electron density (see Section 1.3.1 for further discussion). 4IKP (2013) was deposited by the SGC, and consists of CARM1 bound to methylenesinefungin (see Appendix 6.1 for inhibitor structure). Supporting data relating to 4IKP are yet to be published. The final CARM1 structures available were: 2Y1W, which contains an indole inhibitor and sinefungin in each active site, and 2Y1X, which has an indole-pyrazole inhibitor and SAH in each active site.[64g] Both indole inhibitors bind to the substrate binding channel.

Over the past 4 years (October 2014 – September 2018), 41 further CARM1 structures have been released in the PDB. In 2015, Boriack-Sjodin *et al.* published the first ternary complex structures of CARM1. These contained the CARM1 catalytic core bound to sinefungin and both methylated and unmethylated H3 and PABP1 peptide substrates (see Section 1.3.1, above, for further discussion). In 2016, 5K8V and 5U4X were added. 5K8V contains CARM1 with a bisubstrate inhibitor (**VH2C**, Table 2, page 25).[70, 73] This structure was the

first published example of a bisubstrate inhibitor bound to CARM1. 5U4X is the structure of CARM1 bound to TP-064 (see Section 1.4.2 and Figure 13, page 22).

In 2017, 26 further CARM1-inhibitor structures were released on the PDB. For 21 of these (5ISI-5IS6 and 5IH3-5LKJ, 5LV2, 5TBH-5TBJ, Appendix 6.1, deposited by Cura *et al.*), supporting data is yet to be published. PDB entries 5LGP, 5LGS, 5LGR, 5LGQ (Van Haren *et al.*) contain CARM1 bound to peptidic bisubstrate inhibitors comprising 5'-deoxyadenosine linked (*via* an alkyl linker extending from the 5' carbon) to peptides based on known CARM1 substrate PABP1 (see Appendix 6.1 for structures, including peptide fragment sequences).[74]

5LV2, 5TBI, 5TBH and 5TBJ (also deposited by Cura *et al.)* feature CARM1 bound to inhibitors comprising amino adenosine linked to a modified cytosine moiety, which occupies the arginine binding site (**DM1** – **DM7**, Figure 16, below).[75] Interestingly, this inhibitor series was designed to target DNA methyltransferases but showed no inhibitory activity against a panel of DNA and RNA methyltransferases. Of these inhibitors, only **DM2** and **DM4** inhibited CARM1 activity by more than 50% at inhibitor concentrations of 10 μM. The most potent of these inhibitors – **DM4** – inhibited CARM1 with an IC_{50} value of 1.5 μM. Compounds **DM5** and **DM6** were not co-crystallised with CARM1.

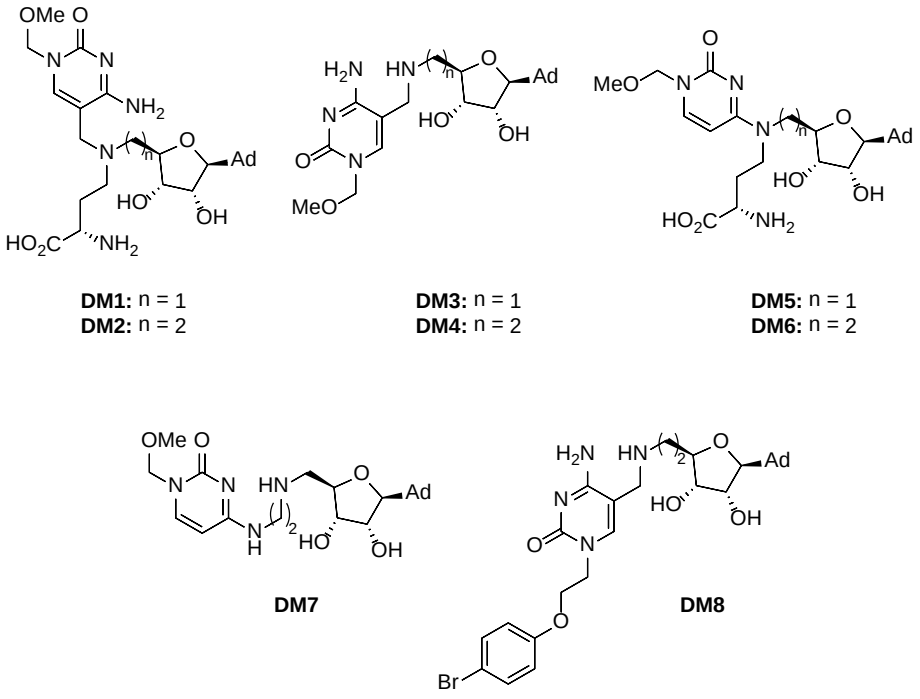


Figure 16: Repurposed putative DNA methyltransferase inhibitors reported by Halby *et al.*

5LV3 features CARM1 bound to a closely related inhibitor - **DM8**. **DM8** has the same structures as **DM2** but with an additional bromobenzene group, further extending the substrate-competitive portion of the inhibitor. The same primary citation as that for **DM1 – DM7** is given in the PDB entry for the CARM1-**DM8** complex structure; however this publication does not appear to contain inhibition data relating to **DM8**.

In 2018, 6ARJ was released, which contains CARM1 in complex with EPZ022302 (see Section 1.4.2 and Figure 13, page 22) and SAH. 6ARV, also released in 2018, contains one of the inhibitors made during the optimisation which eventually resulted in EPZ022302 (in addition to SAH).[42]

6D2L and 6DVR contain CARM1 in complex with CARM1 probes (S)-SKI-72 and (R)-SKI-72, respectively (See Appendix 6.1 for structures). Supporting inhibition data for these compounds are yet to be released. 5NTC (Cura *et al.*) – containing an inhibitor closely related to those deposited by the same authors in 2017 (see Appendix 6.1 for structure) was also added in 2018, but supporting data are yet to be published.

Notable trends throughout the current available CARM1 structures include near physiological pH and presence of PEG 2000 and PEG 3350 in the crystallization conditions; the near-universal P 21 21 2 space group; and consistent unit cell dimensions.

CHAPTER 2: **Materials and Methods**

2.1. **Cloning**

2.1.1. General

All water used for restriction enzyme cloning and site directed mutagenesis was de-ionized, and autoclaved prior to use.

2.1.2. Restriction Enzyme Cloning

2.1.2.1. PCR

For each desired PCR product, a 2-reaction master mix was made, which contained 0.5 µL template plasmid (50 – 250 ng of DNA), 22.5 µL 5 × Phusion HF buffer, 1.8 µL dNTPs and 66.25 µL H_2O. Two reaction mixtures were made, each of which contained 36.1 µL master mix, 2µL 20 µM forward primer, 2 µL 20 µM reverse primer, 0.5 µL Phusion polymerase (NEB), 2 µL 50% DMSO and 7.4 µL H_2O. 5 × HF buffer, Phusion polymerase and dNTPs were purchased from Thermo Fisher Scientific. The PCR protocol used is shown in Figure 17 (below), and was carried out using a PTC-200 Peltier Thermocycler.

Step	Temp (°C)	Duration	
Initial separation	95	5 mins	
Separation	95	30 secs	24 ×
Annealing	55 or 65[1]	30 secs	
Extension	72	1 min per kb[2]	
Final extension	72	10 mins	
End	4	infinite	

[1]one reaction was performed with an annealing temperature of 55 °C, and the other with an annealing temperature of 65 °C
[2]length of the plasmid template in kilobases

Figure 17: PCR protocol used for restriction enzyme cloning

Following PCR, agarose gel electrophoresis (1 - 2 % gel run for 60 min at 100 V) was used to confirm the molecular weight of the PCR product and assess the sample's purity. 5 µL Nancy-520 (Sigma-Aldrich) was added per 100 mL agarose solution (to visualise bands), and gels were imaged using a Dark Reader Transilluminator (Clare Chemical). Pure PCR products (indicated by the presence of a single band) were purified using a GenElute PCR Clean-Up kit (Sigma-Aldrich) eluting with 50 µL H_2O. Impure PCR products were purified using a GenElute Gel Extraction Kit (Sigma-Aldrich). The PCR primers used during the cloning of each construct are given in Table 3 (page 39) and Table 5 (page 41).

2.1.2.2. Restriction Enzyme Digest

Reactions contained 16 µL of DNA (purified PCR product or vector), 2 µL of an appropriate 10 × buffer and 1 µL of each restriction enzyme. The resulting mixture was incubated at 37 °C for 4 h. The digested vector was subsequently dephosphorylated by the addition of 2 µL of 10 × Antarctic Phosphatase Buffer

and 1 µL of Antarctic Phosphatase and incubation at 37 °C for 1 – 2 h. All enzymes and buffers were purchased from NEB. The digested PCR product and vector were purified using a GenElute PCR Clean-Up kit (Sigma-Aldrich).

2.1.2.3. Ligation

For each insert-vector pairing, two separate ligation reactions were performed, with the insert at either 3 × or 6 × the molar concentration of the vector. Each reaction contained 1 µL T4 DNA ligase (NEB), 1 µL 10 × T4 buffer (NEB), 1 µL vector and 3n or 6n µL insert, where n = [vector]/[insert] × insert size (kb)/vector size (kb). The mixture was made up to a volume of 10 µL with H_2O, then incubated at r.t. for 4 – 16 h. 2 µL of ligation mixture was used to transform competent NovaBlue (DE3) E. *coli* cells (Novagen, see below).

2.1.2.4. Transformation

Competent cells were prepared using the Inoue method.[76] 2 µL of plasmid DNA was added to 50 µL competent cells in a 1.5 mL polypropylene tube (Eppendorf), followed by incubation on ice for 10 mins. The tube was transferred to a heat block at 42 °C for 1 min then returned to ice for 5 mins. 800 µL LB was added to the tube, followed by incubation, with shaking, at 37 °C for 1 h. 20 – 200 µL of the resultant mixture was spread onto an LB agar plate containing the appropriate antibiotic(s). The plate was incubated at 37 °C for 16 h (to allow colony growth) then transferred to the fridge (4 °C).

2.1.2.5. Cell Lysis and Analytical Digest

To screen colonies of transformed Nova Blue cells for presence of the desired plasmid, cell lysis was performed. Up to 8 colonies were selected and each was used to inoculate 5 mL LB containing the appropriate antibiotic, followed by incubation at 37 °C for 16 h. 20 µL of each culture was removed, mixed with 6 µL cracking buffer (50 % sucrose, 0.5 % SDS, 0.5 M NaOH, bromophenol blue in H_2O), and run on an agarose gel. Positive clones were indicated by higher molecular weight band(s) in the agarose gel. Cells from 4 positive colonies were collected by centrifugation, and plasmid DNA extracted using a GenElute Plasmid Miniprep Kit (Sigma-Aldrich). Further confirmation of the presence of the desired plasmid was achieved using an analytical digest. 8 µL of each purified plasmid was digested using the appropriate restriction enzymes, as described above (Section 2.1.2.2), except that half the volume of restriction enzymes and buffer was used. 5 µL of each digestion reaction was run on an agarose gel, to determine whether the plasmid included an insert of the expected molecular weight. Plasmids were sequenced using Sanger sequencing (Source Bioscience), to confirm that they contained the correct DNA sequence.

2.1.3. Site-directed Mutagenesis

Site-directed mutagenesis was performed to alter single codons within plasmids. Each mutagenesis PCR reaction contained 25 – 50 ng template plasmid, 125 ng of each primer, 0.8 µL dNTPs (Thermo Fisher Scientific), 2 µL DMSO, 0.5 µL PfuUltra II (Agilent Technologies), 5 µL 10 × PfuUltra II buffer (Agilent Technologies), and H_2O to a total volume of 50 µL. The PCR protocol

used is shown in Figure 18, and was carried out using a PTC-200 Peltier Thermocycler.

Step	Temp (°C)	Duration	
Initial separation	95	30 secs	
Separation	95	30 secs	24 ×
Annealing	55	1 min	
Extension	68	1 min per kb[1]	
Final extension	68	10 mins	
End	4	infinite	

[1]length of the plasmid template in kilobases

Figure 18: PCR protocol used for site direction mutagenesis

Following PCR, 1 µL DpnI (NEB) was added to the sample, followed by incubation at 37 °C for 2 h. 2 µL of the DpnI digest reaction was then used to transform competent Nova Blue cells (Section 2.1.2.4). 4 colonies were selected and each was used to inoculate 5 mL LB containing the appropriate antibiotic, followed by incubation at 37 °C for 16 h. Cells were collected by centrifugation, and plasmid DNA extracted using a GenElute Plasmid Miniprep Kit (Sigma-Aldrich). Plasmids were sequenced using Sanger sequencing (Source Bioscience), and then used to transform BL21-CodonPlus competent cells (Agilent).

The mutagenesis PCR primers used during the cloning of each construct are given in Table 4, page 40.

2.1.4. Glycerol Stock Preparation

A glycerol stock of each clone was prepared by using a single colony to inoculate 100 mL LB containing appropriate antibiotics. The resulting culture

was grown at 37 °C for 16 h. 400 µL of the culture was then mixed with an equal volume of sterile 50 % glycerol in a polypropylene tube, and stored at -80 °C until needed.

Protein	Name	Forwards Primer	Reverse Primer	Template Plasmid	Insertion Vector	Restriction Sites
PRMT1	PRMT1-N-His	GCTACTGCCTCTTCAACGAGTCCATGCTC	GAGCATGGACTCGTTGAAGAGGCAGTAGC	GST-PRMT1	pET-26b(+)	NdeI + XhoI
	PRMT1-C-His	GGGAATTCCATATGAGTTGTGGCCAGGCGGAAAGCAG	CCGCTCGAGGCGCATCCGGTAGTCGGTG	GST-PRMT1	pET-26b(+)	NdeI + XhoI
	PRMT1(19-361)-N-His-	GGGAATTCCATATGGGTCACCACCACCACCACCACGAAGAAGTGAGTTGTGGCCAGGCGGAAAGC	CCGCTCGAGTCAGCGCATCCGGTAGTCGG	PRMT1-N-His	pET-26b(+)	NdeI + XhoI
CARM1	MBP-CARM1-His	ATAAGAATGCGGCCGCAGAAAACCTGTATTTTCAGTCTGTGTTCAGTGAGCGGACA	GGGAAGCTTTCAGTGGTGGTGGTGGTGGTGTGTACCTGTGTACCTGAAGAAGGG	GST-CARM1	pMALX(E)	NotI + HindIII
	MBP-CARM1	ATAAGAATGCGGCCGCAGAAAACCTGTATTTTCAGTCTGTGTTCAGTGAGCGGACA	GGGAAGCTTTCATGTACCTGTGTACCTGAAGAAGGG	GST-CARM1	pMALX(E)	NotI + HindIII
	MBP-CARM1(135-489)	ATAAGAATGCGGCCGCAGAAAACCTGTATTTTCAGTCTGTGTTCAGTGAGCGGACA	CCCAAGCTTTCAATAGTGGCTGCCCGGCGGCGGGCTCGGGGTTGTACCTGTGTACCTGAAGAAGGGG	GST-CARM1	pMALX(E)	NotI + HindIII
	MBP-CARM1(135-489)-N-His	ATAAGAATGCGGCCGCAGAAAACCTGTATTTTCAGTCTGTGTTCAGTGAGCGGACA	CCCAAGCTTTCAGTGGTGGTGGTGGTGGTGATAGTGGCTGCCCGGCGGCGGGCTCGGGGTTGTACCTGTGTACCTGAAGAAGGGG	GST-CARM1	pMALX(E)	NotI + HindIII
	TF-mCARM1 insert	GGGTTTCATATGGCCGCGGCGGCAGCGACGGCGGTGGGGC	CCCAAGCTTCTAACTCCCATAGTGCATGGTGTTGGTCG	GST-CARM1	TF-pCOLD[2]	NdeI + HindIII
	TF-mCARM1 vector	GAGGGATCCGAATTCAAGCTTGTCGACC	GGAATTCCATATGCGCGGAGGGCCCCTGGAACAG	TF-pCOLD	N/A	NdeI + HindIII
	GST-mCARM1(FL)	GGGTTTCATATGGCCGCGGCGGCAGCGACGGCGGTGGGGC	CCCAAGCTTCTAACTCCCATAGTGCATGGTGTTGGTCG	GST-CARM1	GST-pCOLD	NdeI + HindIII

Table 3: Summary of PCR primers used for restriction enzyme cloning – PRMT1 and CARM1

Protein	Name	Forwards Primer	Reverse Primer	Template Plasmid	Insertion Vector
PRMT1	PRMT1(E55N)	CGCACACTTTGGCATCCACGAGAACATGCTGAAGGACGAGGTGCG	CGCACCTCGTCCTTCAGCATGTTCTCGTGGATGCCAAAGTGTGCG	PRMT1-N-His	pET-26b(+)
	PRMT1(Y160N)	GCTACTGCCTCTTCAACGAGTCCATGCTC	GAGCATGGACTCGTTGAAGAGGCAGTAGC	PRMT1-N-His	pET-26b(+)
CARM1	MBP-CARM1(N161E)	CTATGGCTACCTATCCCAGCAGCAGGAAATGATGCAGGACTATGTGCGGAC	GTCCGCACATAGTCCTGCATCATTTCCTGCTGCTGGGATAGGTAGCCATAG	MBP-CARM1-His	pMALX(E)
	2GKG-CARM1(N161E)	CTACCTATCCCAGCAGCAGGAAATGATGCAGGACTATGTGC	GCACATAGTCCTGCATCATTTCCTGCTGCTGGGATAGGTAG	2GKG-CARM1-His	pRSF-13
	2GKG-CARM1(N265Y)	GGCTACATGCTCTTCTACGAACGAATGCTC	GAGCATTCGTTCGTAGAAGAGCATGTAGCC	2GKG-CARM1-His	pRSF-13

Table 4: Summary of PCR primers used for side directed mutagenesis

Protein	Name	Technique used[1]	Forwards Primer	Reverse Primer	Template Plasmid	Insertion Vector	Restriction Sites
Histone 3	GB-H3(21-44)His	REC	CGCGGATCCCTG GAAGTGCTGTTT CAGGGCCCGGCC ACCAAGGCAGCC AGG	CCGCTCGAGTTA GTGGTGGTGGTG GTGGTGGTGGCC GGGCCGGTAACG GTGAG	H3-FL-xenopus	GB-fusion pET-21a(+)	BamHI + XhoI
	GB-H3(21-44)	SDM	CCGGCCCGGCTA GCACCACCACC	GGTGGTGGTGCT AGCCGGGCCGG	GB-H3(21-44)His	GB-fusion pET-21a(+)	-
	GB-H3(3-31)	REC	GGGGGATCCGAA AACCTGTATTTTC AGGGCACCAAGC AGACCGCCCGTA AATCC	CCCCCTCGAGTC AGTGGTGGTGGT GGTGGTGAGCAG GAGCGGACTTCC TGGCTG	H3-FL-xenopus	GB-fusion pET-21a(+)	BamHI + XhoI
Histone 4	GB-H4(1-21)His	REC	CGCGGATCCGAA AACCTGTATTTTC AGTCTGGTCGTG GTAAAGGTGGTA AAG	CCGCTCGAGTTA GTGGTGGTGGTG GTGGTGGTGAAC TTTACGGTGACG TTTAGCAC	H4-FL-xenopus	GB-fusion pET-21a(+)	BamHI + XhoI

[1] REC = Restriction Enzyme Cloning, SDM = site-directed mutagenesis

Table 5: Summary of PCR primers used for restriction enzyme cloning - histone constructs

2.2. Protein Expression

2.2.1. Preparative Protein Expression

Expression of each construct followed the same general protocol. Due to anticipated difficulties in expressing un-tagged CARM1, both the catalytic core and full length CARM1 constructs were expressed as solubility-tag fusions. Variations are detailed in Table 6. A glycerol stock of the bacterial clone was used to inoculate 100 mL LB broth containing appropriate antibiotics. This was incubated for 16 – 18 h at 37 °C, with shaking at 180 RPM. 8 mL of this culture was used to inoculate 800 mL LB (containing the appropriate antibiotic) in a 2 L non-baffled Erlenmeyer flask. The resulting culture was incubated at 37 °C, with shaking at 180 RPM, until an OD_{600} of 0.5 – 0.6 was reached. Upon reaching this optical density, 1M IPTG was added to a final concentration of 0.4 mM, and the temperature dropped to 15 – 22 °C, depending on the construct (Table 6). After 16 – 24 h, cells were harvested by centrifugation, and stored at -20 ° C until needed. Single site mutants were expressed and purified using the same procedures as the corresponding wild type constructs.

Construct	Cell type	IPTG concentration/mM	T/°C	Expression Time/hours
PRMT1 (all constructs)	BL21-CodonPlus	0.4	22	16
2GKG-CARM1	BL21-CodonPlus	0.05 – 0.4[1]	20	16
MBP-CARM1 (all constructs)	BL21-CodonPlus	0.4	20	16
TF-mCARM1	BL21-CodonPlus	0.4	15	24
GST-mCARM1(FL) in pGEX vector	Arctic Express	0.4	10	16
GST-mCARM1(FL) in pCOLD vector	BL21-CodonPlus	0.4	15	24
Histones	BL21-CodonPlus	0.4	20	16

[1]0.4 mM IPTG was used for the CARM1 expressions presented herein. 0.05 mM IPTG concentration was subsequently found to increase the yield of 2GKG-CARM1.

Table 6: Expression conditions for constructs used in this thesis.

The expression conditions used for PRMT1 were based on those reported by Zhang and Cheng for similar PRMT1 constructs (though the OD_{600} at which the cells were induced was changed compared to their protocol).[26] Initial CARM1 expression conditions were based on those reported by Sack *et al.* for their MBP-CARM1 (catalytic core) fusion protein construct (though LB media was used, and the cells were induced at lower OD_{600}).[64g] The lower temperatures used for the full length CARM1 constructs are based on the supplier's recommended expression condition for the relevant expression vectors and cell lines (i.e. pCOLD and Arctic Express).[77]

2.2.2. Analytical Protein Expression

To determine whether the amount of soluble MBP-CARM1 constructs expressed could be increased using Arctic Express cells, small scale expressions were performed. For each clone, 5 mL taken from 100 mL culture grown overnight (as above) was used to inoculate 200 mL LB (containing the appropriate antibiotic) in a 500 mL Erlenmeyer flask. Cultures were then incubated, with shaking, at 37 °C until an OD_{600} of 0.5 – 0.6 was reached. Upon reaching this optical density, 1M IPTG was added to a final concentration of 0.2 mM and the temperature reduced to 10 °C (Arctic Express) or 15 °C (BL21-CodonPlus). After 16 h, 1 mL aliquots were removed from each culture, cells collected by centrifugation, and the supernatant discarded. Cell pellets were resuspended in 100 µL nickel buffer A (Table 7, page 46), sonicated (as described in Section 2.3.2) and the soluble and insoluble fractions separated by centrifugation at

12,000 × g for 10 min at 4 °C. The insoluble fraction was resuspended using a solution of 15 mM Tris, 2 M urea, 1 % SDS, 1.25 % β-mercaptoethanol, 2.5 % glycerol at a pH of 6.8 followed by boiling for 5 min. SDS-PAGE (see Section 2.4) was used to judge the MBP-CARM1 concentration in soluble and insoluble fractions. 5 µL of each sample was loaded, such that the band intensity was proportional to the total soluble and insoluble MBP-CARM1 content of the cells from each culture.

2.2.3. Test Expressions of 2GKG-CARM1-His

Test expressions were performed in an attempt to optimise the yield of 2GKG-CARM1-His. Expressions and cell lysis were performed using 800 mL media for each test condition (as described in sections 2.2.1 and 2.3.2). The conditions were varied from the standard expression protocol as detailed in Table 6, page 36. The clarified cell lysate was loaded onto a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with buffer A (300 mM NaCl, 50 mM Tris, 5 % glycerol, 20 mM imidazole, pH 7.5), using a peristaltic pump followed by washing with 25 mL buffer A. The column was then washed with 25 mL 20 % buffer B followed by 15 mL 100 % buffer B (300 mM NaCl, 50 mM Tris, 5 % glycerol, 500 mM imidazole, pH 7.5), collecting the eluent at each stage. Equal volumes of eluent from each elution stage were run on SDS-PAGE and the 2GKG-CARM1-His content analysed by visual inspection of the stained gel.

2.2.4. Test Expressions of GST-mCARM1(FL)

Test expressions were performed in an attempt to obtain an active sample of GST-mCARM1(FL). The conditions used are detailed in Figure 38, page 116. The clarified cell lysate was loaded onto a 1 mL GSTrap FF column (GE Healthcare) pre-equilibrated with buffer A (50 mM Tris, 150 mM NaCl, 0.05 % TWEEN-20, pH 7.5)[2] at a flow rate of 1.0 mL/min. The column was then washed with 10 mL buffer A, followed by 3 mL buffer B (100 mM Tris, 150 mM NaCl, 1 mM DTT, 24.4 mM reduced glutathione, pH 8.0), followed by eluting with 5 mL buffer B, collecting fractions 0 – 1 mL and 1 – 3 mL. 4 µL of flow through (collected during loading) and 8 µL of each fraction were run on SDS-PAGE and the mGST-CARM1(FL) content analysed by visual inspection of the stained gel.

2.3. **Protein Purification**

2.3.1. General

All chromatography was performed using ÄKTA FPLC systems (Prime, Purifier, Pure and Start). SDS-PAGE was used to analyse the protein content of fractions from chromatography. Fractions containing the desired protein were combined and concentrated using a Vivaspin 20 centrifugal concentrator with an appropriate molecular weight cut-off (GE Healthcare). All procedures were carried out on ice where possible, and samples stored at 4 °C between steps. FPLC systems were at room temperature. The buffers routinely used for the

[2] The buffers used were based on those reported by Shishkova *et al.*,[36] but modified by replacing NP-40 with TWEEN-20, and adding 150 mM mM NaCl and 1 mM DTT to buffer B.

purification of each construct are shown in Table 7, below. Any deviations from these buffers are indicated in the results section.

Construct	Nickel buffer A	Nickel buffer B	Gel filtration buffer
PRMT1 (all constructs)	300 mM NaCl, 50 mM tris, 5 % glycerol, 20 mM imidazole, pH 8.0	300 mM NaCl, 50 mM tris, 5 % glycerol, 500 mM imidazole, pH 8.0	300 mM NaCl, 20 mM tris, 5 % glycerol, pH 8.0
CARM1 (all constructs[1])	300 mM NaCl, 50 mM tris, 5 % glycerol, 20 mM imidazole, pH 7.5	300 mM NaCl, 50 mM tris, 5 % glycerol, 500 mM imidazole, pH 7.5	300 mM NaCl, 20 mM tris, 5 % glycerol, pH 7.5
Histones	150 mM NaCl, 50 mM tris, 5 % glycerol, 20 mM imidazole, pH 8.0	150 mM NaCl, 50 mM tris, 5 % glycerol, 500 mM imidazole, pH 8.0	150 mM NaCl, 20 mM tris, 5 % glycerol, pH 8.0

[1]Except GST-CARM1(FL) and MBP-CARM1, which were isolated from the lysate by GST and MBP affinity chromatography, respectively (see below) followed by gel filtration

Table 7: Buffers used for nickel affinity chromatography and gel filtration

2.3.2. Cell Lysis

Cells were thawed and resuspended in 10 mL buffer A per gram of cell pellet and distributed into 50 mL Falcon Tubes. The suspensions were sonicated at 12 microns for 20 seconds on and 20 seconds off for a total sonication time of 2 mins (4 mins in total). Both resuspension and sonication were performed on ice. The lysate was clarified by centrifugation at 15000 × g for 40 minutes. The supernatant was filtered using a 0.45 µm syringe filter (Sartorius).

For purification of histone proteins and full length CARM1 constructs, the buffer A into which the cell pellet was lysed was supplemented with cOmplete™ EDTA-free protease inhibitor cocktail (Roche).

2.3.3. Nickel Affinity Chromatography

Clarified lysate was loaded onto a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with buffer A, using a peristaltic pump. The column was washed with 50 – 100 mL buffer A, followed by a linear gradient elution of 0 – 100 % buffer B over 20 column volumes. 2.5 mL fractions were collected throughout the gradient elution, and analysed by SDS-PAGE to determine their protein content. Any modifications made to this protocol are discussed in the results section.

2.3.4. GST Affinity Chromatography

Clarified lysate was loaded onto a 1 mL GSTrap HP column (GE Healthcare) pre-equilibrated with buffer A (200 mM NaCl, 50 mM Tris, 0.4 mM EDTA, 5 % glycerol, pH 7.5) at a flow rate of 0.4 mL/min. The column was washed with 30 mL buffer A at a flow rate of 1 mL/min then eluted with 10 mL buffer B (200 mM NaCl, 50 mM Tris, 0.4 mM EDTA, 5 % glycerol, 20 mM reduced glutathione, pH 7.5) at a flow rate of 0.5 mL/min collecting 1 mL fractions.

2.3.5. MBP-Affinity Chromatography

Clarified lysate was loaded onto a 5 mL MBPTrap column (GE Healthcare) pre-equilibrated with buffer A (200 mM NaCl, 20 mM Tris, 2 % glycerol, 2 mM EDTA, pH 7.5) at a flow rate of 0.5 mL/min. The flow-through was loaded onto the same column (to maximise binding) at a flow rate of 0.5 mL/min. The

column was washed with 50 mL buffer A then eluted with 15 mL buffer B (200 mM NaCl, 20 mM Tris, 2 % glycerol, 2 mM EDTA, 10 mM maltose, pH 7.5) collecting 2.5 mL fractions.

2.3.6. Anion Exchange Chromatography

Combined fractions from nickel affinity chromatography were buffer exchanged by successive spin-concentration and re-dilution in buffer A (20 mM NaCl, 50 mM Tris, 5 % glycerol, pH 8.0). The sample was loaded onto a 6 mL Resource Q column (GE Healthcare), followed by a gradient elution of 0 – 100% buffer B (1 M NaCl, 50 mM Tris, 5 % glycerol, pH 8.0) collecting 2.5 mL fractions.

2.3.7. Gel Filtration

Preparative gel filtrations were carried out using HiLoad Superdex 16/60 columns (GE Healthcare). An S75 column was used to gel filtrate the histone constructs; an S200 was used for all other constructs. The column was pre-equilibrated with gel filtration buffer (Table 7, page 46). Pooled fractions from the previous chromatographic step were concentrated to 0.5 – 8 mL (depending on the solubility of the protein) then centrifuged for 10 min at 12000 × g. Samples were then loaded onto the column using an injection loop (1 – 10 mL depending on sample volume). An isocratic elution was performed and 2.5 mL fractions were collected. Fractions were analysed by SDS-PAGE, and those

containing pure protein were combined and spin-concentrated to the desired concentration.

Analytical gel filtrations were carried out using an S200 10/300 Superdex column (GE Healthcare).

2.3.8. Co-purification with Histones

Cell pellets containing the relevant PRMT and histone were mixed, followed by resuspension, lysis and nickel affinity chromatography. Typically, 800 mL bacterial culture for the relevant histone was used for every 4 – 5 L of culture expressing PRMT1 or CARM1.

2.4. **SDS-PAGE**

SDS-PAGE was used to visualise the protein content of samples including chromatography fractions and expression cultures. The SDS-PAGE gels presented in this thesis were 0.75 mm thick and comprised a stacking and resolving gel. The stacking gel contained 6 % acrylamide and had a pH of 6.8. All resolving gels had a pH of 8.8, but the acrylamide concentration varied from 10 – 20 % depending on the molecular weight of the protein of interest.

Samples were prepared by mixing 10 µL 3 × SDS-PAGE loading buffer (160 mM Tris pH 6.8, 5 % SDS, 30 % glycerol, 20 mM DTT, 0.03 % bromophenol blue) with 20 µL sample in a polypropylene tube. The samples were then heated at 95 °C for 5 min prior to loading into the gel. Gels were run at 200 V until the

dye front reached the bottom of the gel. Gels were stained using 0.25 % Coomassie Brilliant Blue R-250, 25 % IPA and 25 % acetic acid, and destained using 10 % IPA, 10 % acetic acid.

2.5. Determining Protein and Inhibitor Concentrations

Protein concentrations were determined by UV absorbance at 280 nm (with a NanoDrop 1000 v3.7) using extinction coefficients at 280 nm determined by the ExPASy ProtParam online tool. The protein concentrations calculated in this manner were shown to be in good agreement with the concentration determined using a Bradford assay with BSA as a standard (see Appendix 6.4.1).[78]

With the exception of the compounds discussed in section 3.8, the inhibitors used in this thesis were synthesized by Dr. U. Muhsen (**UM079**, **UM249**, **UM251**, **UM305**, **UM451**, **UM291**, **UM290**, **UM454**, **UM457**, **UM284**, **UM360**),[67, 69] Dr. W. Hong (**WH3C**, **WH4C**, **WH5C**)[68-69, 79] and Alaa Al-Noori (**AA175** and **AA183**).[80] A summary of these inhibitors is given in Figure 15, page 27.

The concentrations of inhibitor solutions were determined using a NanoDrop 1000 v3.7, using the molar extinction coefficient at 260 nm for S-adenosyl homocysteine (ε_{260} = 15400 $M^{-1}cm^{-1}$). **UM305** contains a 2-alkylaminopyrimidine group and **AA175** and **AA183** contain a 2-alkylaminopyridine. 2-alkylaminopyrimidines and 2-alkylaminopyridines absorb at 260 nm with extinction coefficients of <15400 $M^{-1}cm^{-1}$. The concentrations of **UM305**, **AA175**

and **AA183** may therefore have been overestimated up to two-fold, resulting in a corresponding over-estimation of the K_d and IC_{50} values of these inhibitors. The naphthylamine absorption spectrum of **EG047** and **EG048** also overlaps with the adenine absorption; therefore these concentrations may also have been overestimated. The Beer-Lambert law breaks down at high absorbance. A standard curve was generated to determine the concentration range within which the Beer-Lambert law applied (see Appendix 6.4.2). Dilutions were performed to determine the concentration of solutions which were outside the linear range.

2.6. Western Blot

Western Blot was used to confirm that the peak eluting first during gel filtration of the MBP-CARM1 TEV cleavage reaction corresponded to CARM1 (see Figure 30, page 104).

A 10 % SDS-PAGE gel using 3 µg of each protein was run (as described above). The protein was then transferred to a nitrocellulose membrane (pre-soaked with MeOH) using a Mini Trans-Blot cell (Biorad) at 100 mA for 2 h at 4 °C in Western Blot Transfer Buffer (WBTB: 192 mM glycine, 25 mM Tris pH 8.3, 20 % MeOH). All components of the transfer system were soaked in WBTB prior to running the transfer. After transfer, the membrane was incubated with 5 % skimmed milk in TBST (50mM Tris pH 7.5, 150 mM NaCl, 0.1% TWEEN-20) for 16 h at 4 °C, followed by incubation with a solution of the primary antibody in 5 % skimmed milk TBST for 1 h at r.t. The membrane was then washed for 45

min with 3 × 15 mL TBST, replacing the TBST every 15 mins. The membrane was then incubated with a solution of the secondary antibody in 5 % skimmed milk TBST for 1 h at r.t., followed by another 3 × 15 min washes with 3 × 15 mL TBST. The membrane was transferred to a clean container, then incubated for 1 min with enhanced chemiluminescence (ECL) solution (22 µL 90 mM *p*-coumaric acid, 50 µL 250 mM luminol and 3 µL 30 % H_2O_2 in 10 mL 0.1 M Tris pH 8.5). Chemiluminescence was detected using an ImageQuant LAS-4000 (GE Healthcare). After detection, the membrane was stained using Ponceau S stain to verify that adequate protein was transferred to the membrane.

A Western blot was also used to analyse GST-mCARM1(FL) fractions for presence of the desired construct (Figure 37, page 115). The Western blot was performed as above, except that 10 µL of each fraction and 10 µg of the negative control (MBP) was used.

The primary and secondary antibodies used for all Western blots are summarised in Table 8.

Epitope	Primary Antibody	Working concentration	Secondary Antibody	Working concentration
His Tag	mouse IgG anti-tetra-His (Qiagen)	100 µg/L	Anti-Mouse IgG (Fab specific)–Peroxidase antibody (Sigma)	0.1 % (v/v)
MBP tag	mouse IgG MBP-tag (GenScript)	100 µg/L	Anti-Mouse IgG (Fab specific)–Peroxidase antibody (Sigma)	0.1 % (v/v)
GST Tag	mouse IgG anti GST	100µg/L	Anti-Mouse IgG (Fab specific)–Peroxidase antibody (Sigma)	0.1 % (v/v)

Table 8: Western Blot antibodies used and working concentrations.

2.7. Crystallography

2.7.1. Crystallisation

All crystallisation trays were set up within 24 h of gel filtration of the relevant protein. Proteins were concentrated to the required concentration, then centrifuged for 10 min at 12,000 × g to remove any debris or aggregated protein.

Crystallisation was performed using the sitting drop, vapour diffusion method. Initial screens were performed in 96-well MRC Crystallization Plates (SWISSCI) using commercial crystallisation screens (Molecular Dimensions and Hampton Research). The reservoirs were filled with 80 µL of screen solution using a multi-channel pipette and subsequent solution transfers were performed using a Mosquito Crystal (TTP Labtech). Screen solution (120 – 200 nL) was transferred from each reservoir onto its corresponding shallow well(s). An equal volume of protein solution was then added to the drop in each shallow well, to form the sitting drops. The plate was sealed and incubated at 20 °C. Optimisations were carried out in 24-well Cryschem M Plates (Hampton Research). The reservoirs were filled with 400 - 500 µL of precipitant solution. 1 µL of solution was transferred from each reservoir onto its corresponding shallow well. An equal volume of protein solution was then added to the drop in each shallow well, to form the sitting drops. The plate was then sealed and incubated at 20 °C.

For co-crystallisations, proteins were mixed with an aqueous solution of inhibitor and incubated for 5 min at r.t. before adding to the drop of well solution. A summary of the crystallisation conditions used to generate the

CARM1-inhibitor crystals is given in Table 9, below. Generally, the CARM1 catalytic domain was concentrated to the concentration indicated, followed by the addition of a 500 μM aqueous solution of inhibitor such that the final inhibitor concentration was 167 μM. For the co-crystallisation of CARM1(N265Y) with **AA183**, excess **AA183** was added to the solution of CARM1(N265Y) prior to setting up the tray. This was to ensure that there was sufficient **AA183** present to saturate the protein (since this mutant had significantly lower affinity for **AA183** than wild type CARM1). For **AA175** with CARM1(N265Y), 1.4 mM **AA175** was used.

Inhibitor	[CARM1] /mg/mL[1]	[CARM1] /μM[1]	Well solution	Reservoir volume/ μL	Drop volume/ μL
UM079	1.7	44	0.1 M Bis-tris propane pH 7.0, 0.02 M sodium phosphate, 21 % (w/v) PEG 3350	500	2
UM249	1.7	44	0.1 M Bis-tris propane pH 7.0, 0.02 M dipotassium phosphate, 22 % (w/v) PEG 3350	400	2
UM251	1.7	44	0.1 M Bis-tris propane pH 7.0, 0.2 M sodium acetate, 24 % (w/v) PEG 3350	500	2
WH5C	1.2	31	0.1 M Bis-tris propane pH 7.0, 0.2 M sodium acetate, 20 % (w/v) PEG 3350	500	2
UM305	1.0	26	0.1 M Bis-tris propane pH 7.0, 0.02 M sodium phosphate, 22 % (w/v) PEG 3350	500	2
AA175	3.0	77	0.1 M Bis-tris propane pH 8.5, 0.2 M sodium formate, 20 % (w/v) PEG 3350[2]	80	0.4
AA183	3.0	77	0.1 M Bis-tris propane pH 8.5, 0.2 M sodium formate, 20 % (w/v) PEG 3350[2]	80	0.4
AA183 w/CARM 1(N265Y)	2.4	62	0.1 M Bis-tris propane pH 7.0, 0.02 M sodium phosphate, 26 % (w/v) PEG 3350	500	2
AA175 w/CARM 1(N265Y)	1.7	44	0.1 M Bis-tris propane pH 7.0, 0.02 M sodium phosphate, 22 % PEG 3350 (w/v)	500	2
EG047	1.0	26	0.1 M Bis-tris propane pH 7.0, 0.02 M sodium phosphate, 26 % PEG 3350 (w/v)	500	2

[1]The protein concentrations indicated are the concentrations measured prior to addition of a half-volume of 500 μM ligand solution, then mixing with the well solution, as detailed in the main text.
[2]CARM1 in complex with inhibitors **AA175** and **AA183** (blue box) was purified and crystallised by Alaa Al-Noori using the PACT premier™ HT-96 screen (Molecular Dimensions).

Table 9: Crystallisation conditions used to generate CARM1-inhibitor co-crystals

2.7.2. Cryoprotection

Crystals were cryoprotected in the well solution supplemented with 30 % (v/v) glycerol and 20 % (w/v) PEG 3350 prior to cryocooling in liquid nitrogen.

2.7.3. Data Collection, Structure Determination and Refinement

X-ray diffraction experiments were carried out at Diamond Light Source, Harwell, Oxford and ESRF, Grenoble. The beamlines used to collect the datasets used in this thesis are summarized in Appendix 6.5. All crystals belong to the space group P 2_12_12, with four monomers in the asymmetric unit.

Indexing, integration, space group determination scaling were performed using XDS and aimless.[81]

All CARM1 structures were solved by molecular replacement using Phaser MR.[82] The search model for CARM1·**UM249** was a published structure of CARM1 (PDB 5DX0, with all ligands removed) and the final model of CARM1·**UM249** (with ligands removed) was used as a search model for the remaining CARM1-inhibitor complexes. Model building was carried out using COOT[83] and refinement was carried out with Phenix.[84] Parameters for data collection, processing, structural refinement and validation can be found in Appendix 6.5.

At the time of writing, the structures of CARM1(WT) in complex with **EG047** and CARM1(N265Y) in complex with **AA175** have gone through several rounds of refinement and rebuilding, but final changes (e.g. precise ligand and water positioning) have not yet been made. All other structural data have been uploaded to the PDB deposition server and are pending submission.

2.8. Radiometric Activity Assays

2.8.1. Initial Test for Activity of Recombinant Constructs

Activity assays were performed to determine whether purified PRMT1 and CARM1 constructs were catalytically active.

Methylation reactions contained PRMT1 or CARM1 (2 μM), peptide substrate (30-60 μM) and Adenosyl-L-Methionine, S-[methyl-^{3}H] (Perkin Elmer, NET155H250UC at a final concentration of 0.23 μM) in PBS buffer at pH 7.3. The reactions were initiated by addition of Adenosyl-L-Methionine, S-[methyl-^{3}H] and incubated at 30 °C after mixing. After 60 minutes, the reaction mixture was pipetted onto a P81 phosphocellulose paper square (Perkin Elmer) and immediately immersed in a stirred solution of trichloroacetic acid (10 % w/v). The phosphocellulose squares were washed for 3 × 10 minutes, replacing the 10 % TCA after each wash. A final wash was carried out with 95 % ethanol, the squares removed and allowed to dry. Incorporation of tritium into the histone substrate was determined by liquid scintillation counting. This records the disintegrations per minute (DPM) of each sample, which can be converted to the concentration of radiolabelled methyl groups (see below) using the specific activity of the ^{3}H-SAM. Control reactions, in which the enzyme or substrate was replaced with H_2O, were run simultaneously.

The specific activity of the ^{3}H-SAM purchased varied from 78.0 to 82.7 Ci/mmol (NET155H, 250 μCi, lot numbers and 1912944 and 2149387). The isotope is packaged as a 0.55 mCi/mL solution in sulfuric acid. Owing to radioactive decay of tritium (half-life 12.3 years) and the differing date on which each experiment

was performed, the specific activity of ^{3}H-SAM at the time each experiment was performed ranged from 70.0 to 82.7 Ci/mmol. The lot number of the ^{3}H-SAM used for each individual experiment was not recorded, therefore the DPM signal was converted to the concentration of transferred methyl groups using an approximate specific activity of 80 Ci/mmol for all datasets. This approximation has no effect on the determined K_m or IC_{50} values. The substrates used to test the activity of each construct, their concentrations, and the concentrations of the enzyme being tested are given in Table 10, below. Histone peptides were custom synthesised by Alta Biosciences (University of Birmingham).[85] GB-fusion proteins were expressed and purified as described in section 3.6.2.1. The exact sequence of each GB-fused histone substrate is given in Table 14, page 86.

Construct	Concentration (μM)	Substrate	Concentration (μM)
PRMT1	2	H4(1-21) peptide	30
CARM1-His	2	H3(1-21) peptide	30
CARM1	2	H3(1-21) peptide	30
CARM1(135-489)	2	H3(1-21) peptide	30
GST-mCARM1(FL)	0.5	GB-H3(3-31)	120
TF-mCARM1(FL)	0.5	GB-H3(3-31)	60

Table 10: Enzyme and substrate concentrations used for initial activity tests of purified, recombinant constructs.

2.8.2. Linear Range Determination

Sample aliquots from frozen batches of pure PRMT1 or CARM1 (stored at -80°C) were thawed on ice. Methylation reactions and subsequent steps were performed as above (Section 2.8.1) except that the enzyme concentration was

varied from 0 – 1600 nM (PRMT1), 0 – 500 nM (CARM1, run 1), or 350 – 2600 nM (CARM1, run 2). Additionally, different substrates were used. For PRMT1, 40 µM GB-H4(1-21)His was used and for CARM1, 120 µM GB-H3(3-31)His was used (see Section 3.1 for sequence of and Section 3.5.2.1 for purification of histone substrates). Different enzyme concentrations were achieved by performing serial dilutions of the enzyme in PBS. The final concentration of Adenosyl-L-Methionine, S-[methyl-^{3}H] was 0.23 µM in all cases. The concentration of transferred methyl groups was plotted as a function of enzyme concentration. Assay signal varied linearly with enzyme concentration below and >5 × above the concentration subsequently used for IC_{50} assays (Section 3.6.2.2).

2.8.3. Determination of Histone K_m

Methylation reactions and subsequent steps were performed as above (Section 2.8.2) except that the substrate concentration was varied from 1.8 – 100 nM (PRMT1) or 0 – 200 nM (CARM1). Different substrate concentrations were achieved by performing serial dilutions of the substrate in PBS. The final concentration of Adenosyl-L-Methionine, S-[methyl-^{3}H] was 0.23 µM in all cases. PRMT1-N-His was used at a concentration of 120 nM and CARM1-His at a concentration of 320 nM. The concentration of transferred methyl groups was plotted as a function of enzyme concentration. The K_m and V_{max} were determined by curve fitting using non-linear regression in GraphPad Prism 7.03 using the Michaelis-Menten equation ($Y=V_{max}*X/(K_m+X)$). For the CARM1 data, the background DPM of the no-histone control samples was subtracted from

each histone-containing sample prior to curve-fitting. For the PRMT1 data, a no-histone control assay was not included, therefore the Michaelis-Menten equation was modified such that the y-intercept was fitted ($Y=V_{max}*X/(K_m+X)+c$).

2.8.4. IC_{50} Assays

Inhibition assays were performed using frozen aliquots of PRMT1 or CARM1 for which the linear range had been determined as described above (Section 2.8.2). A single batch of PRMT1 was used for all IC_{50} assays. One batch of CARM1 was used for inhibitors **UM079**, **UM249**, **UM251**, **UM305** and a second batch for SAH, **WH5C**, **AA175** and **AA183**. The uninhibited assay signal typically observed for both CARM1 batches was similar (25000-40000 DPM (i.e. 0.14 – 0.28 pmol methyl groups transferred)).

PRMT1 methylation reactions contained PRMT1-N-His (160 nM), GB-H4-His (40 µM) and Adenosyl-L-Methionine, S-[methyl-^{3}H] (Perkin Elmer, NET155H250UC) at a final concentration of 0.23 µM, in PBS buffer at pH 7.3 in the presence of the different inhibitors. For IC_{50} curves, a minimum of 10 inhibitor concentrations were used, and experiments were carried out in duplicate. Methylation reactions and subsequent steps were performed as described above (Section 2.8.1). For each inhibition series, two (i.e. duplicate) no-substrate control reactions were included which contained water in place of the histone substrate. This data point was included in the data used to fit the curve by setting the x-value (log[I]) to 4.0. Two (i.e. duplicate) no-inhibitor

controls were also included in the data, with the x-value (log[I]) set to -3.0. This approach was used (rather than using these values to convert the DPM of every other sample to percentage inhibition) because the DPM values of the control samples are no more accurate than the other data points: using them to calculate percentage inhibition would result in inaccurate values, whereas including them as data points increases the number of observations used for curve-fitting, thus increasing the accuracy of the determined IC_{50} values.[86]

CARM1 methylation reactions were carried out using 0.23 µM Adenosyl-L-Methionine, S-[methyl-^{3}H], 320 nM CARM1-His and 120 µM histone substrate (GB-H3(3-31)-His). Since CARM1 automethylates, a no-enzyme control was used as the maximum inhibition sample (log[I] set to 4.0).

Data were analysed in GraphPad Prism 7 and fit using non-linear regression using the 'log(inhibitor) vs. response -- variable slope (four parameters)' function (Y=Bottom + (Top-Bottom)/(1+10^((LogIC_{50}-X)*HillSlope))) with no constraints. The duplicate DPM values recorded for each inhibitor concentration (or control) were converted to pmol of methyl groups transferred (as described above), averaged, and the average values plotted and used for curve-fitting. IC_{50} curves are shown in the results section.

2.8.5. Inhibitor Screening Assays

Inhibitor screening assays were performed to determine whether the morpholino inhibitors showed inhibitory activity (prior to full IC_{50}

determination). Reactions were performed as for IC_{50} determination (see above) except that only 3 inhibitor concentrations were used. Background measurements were subtracted (no histone for PRMT1 or no enzyme for CARM1, as for IC_{50} assays). The background-subtracted DPM values were then divided by the average of two no-inhibitor control samples (which had also had the background subtracted). In several cases the measured DPM was higher for the lowest inhibitor concentration samples then the no-inhibitor control samples. In these instances, the duplicate samples with the highest DPM were used as the 'maximal activity' samples.

2.9. Isothermal Titration Calorimetry

Titrations were carried out using a MicroCal PEAQ-ITC Isothermal Titration Calorimeter (Malvern) or MicroCal iTC200 (Malvern) in the case of titrations of **UM249** and SAH into PRMT1. Data analysis was carried out using MicroCal PEAQ-ITC Analysis Software. Inhibitors were dissolved in gel filtration buffer to concentrations between 150 and 2220 μM as suitable for ITC titrations. 2.4 μL injections of inhibitor solutions into the protein solution at concentrations between 21 and 121 μM were performed, at 25°C. The differing protein concentrations reflect a balance between sample consumption and achieving sufficient curve sigmoidicity. Control titrations were carried out in which the inhibitor was titrated into gel filtration buffer. The control heats of dilution were subtracted from binding experiments prior to curve fitting using a one-site binding isotherm with fitted offset varied. The protein and ligand

concentrations used are summarized in Table 11 (below) and corresponding curves are shown in the results section. In some instances low c-value curves necessitated fixing the stoichiometry prior to curve-fitting and values of N were chosen based on observed N-values with the same batch of purified protein and stronger binding ligands. These cases are indicated in the results section.

Protein	Inhibitor	[protein] (cell)/μM	[inhibitor] (syringe)/μM
PRMT1	SAH	35	200
	UM079	42	269
	UM249	35	230
	UM251	74	480
	WH5C[1]	52	307
	UM305	52	343
	AA175	80	600
	AA183	80	600
	UM451	59	773
	UM454	74	2000
	UM457	83	2220
CARM1	SAH	30	150
	UM079	66	380
	UM249	66	380
	UM251	66	380
	WH5C	121	736
	UM305	61	421
	AA175	21	210
	AA183	24	240
	EG047	40	400
	EG048	40	400
CARM1 N161E	SAH	33	264
	UM305	46	320
	AA175	36	270
	AA183	30	256
CARM1 N265Y	SAH	35	220
	UM305	35	219
	AA175	37	270
	AA183	31	256
PRMT1 E55N	SAH	29	169
	WH5C	20	115
	UM305	52	473
PRMT1 Y160N	SAH	68	340
	UM305	53	285

[1]N value fixed during curve-fitting due to low c-value curve.

Table 11: Protein and inhibitor concentrations used for isothermal titration calorimetry.

A competition ITC titration was performed to determine the K_d of **AA183** and PRMT1. SAH (433 μM) was titrated into a solution containing PRMT1 (59 μM) and **AA183** (76 μM). A control titration of SAH (433 μM) into PRMT1 (69 μM) was performed to ensure that the preparation of PRMT1 was correctly folded. The K_d of **AA183** for PRMT1 was determined by curve fitting using the competitive binding model in the MicroCal PEAQ-ITC Analysis Software.

A titration of **UM457** into PRMT1 in the presence of SAH was performed to determine whether **UM457** bound to the PRMT1·SAH complex. SAH (220 μM) was titrated into PRMT1 (37 μM) until saturated. The syringe was re-loaded with **UM457** (516 μM), which was then titrated into PRMT1·SAH.

2.10. **Synthesis of Putative FRET Probes**

^{1}H and ^{13}C spectra were recorded on a Bruker AV-500 (500/125 MHz), Bruker AV-400 (400/100 MHz), Bruker AV(III)-400 (400/100 MHz) or Bruker DPX-300 (300 MHz) spectrometer. Proton (δ_H) and carbon (δ_C) chemical shifts are quoted in ppm and are internally referenced to the residual protonated solvent signal. Assignments were made on the basis of chemical shifts, coupling constants, COSY, DEPTQ, HSQC data and comparison with spectra of related compounds. Resonances are described as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), dd (double doublet) and so on. Coupling constants (*J*) are given in Hz and are rounded to the nearest 0.1 Hz. H and H' refer to

diastereotopic or otherwise non-equivalent protons attached to the same carbon and imply no particular stereochemistry.

High resolution mass spectra were recorded using the open access Mass Spectrometry Service at the University of Nottingham School of Chemistry using a Bruker MicroTOF mass spectrometer. Mass values (*m*/*z*) are reported in Daltons. High resolution values are calculated to four decimal places from the molecular formula.

Infrared spectra were recorded using a Bruker Fourier Transform spectrophotometer using thin films on a diamond ATR. Diagnostic absorption maxima (ν_{max}) are quoted in wavenumbers (cm^{-1}).

Melting points were determined using a Stuart SMP3 electrothermal melting point apparatus and are uncorrected.

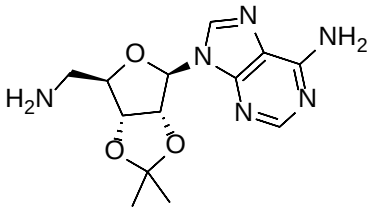
Analytical TLC was performed on pre-coated aluminium-backed silica gel 60 F_{254} plates (Merck) and visualised using UV or basic potassium permanganate dip. Retention factors (R_f) are reported with the solvent system used in parentheses.

Flash column chromatography was performed on silica gel (particle size 35–70 μm, pore diameter 60 Å) with the elution solvent system used in parentheses.

Reactions were carried out under a positive pressure of dry argon from a balloon during the course of the reaction. Reactions were stirred using Teflon-

coated magnetic stirrer bars. Elevated temperature reactions were maintained using a Thermowatch-controlled DrySyn heating block. Reagents were purchased from commercial suppliers and used without further purification unless otherwise stated. Dry solvents were purified using standard techniques. Water used experimentally was deionised. Organic solutions were concentrated using a rotary evaporator with a water bath temperature of 30 °C. Brine refers to a saturated solution of sodium chloride in water. 'Petrol' refers to the fraction of light petroleum ether with a boiling point in the range of 30–40 °C or 40–60 °C as stated.

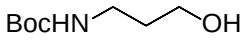
9-((3a*R*,4*R*,6*R*,6a*R*)-6-(Aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9*H*-purin-6-amine (24)[67]



Prepared in 3 steps from adenosine (4.50 g, 16.8 mmol) according to the procedure of Reeve and Townsend,[87] to give the title compound as an off-white solid (3.20 g, 10.5 mmol, 68 % over 3 steps). δ_H (400 MHz, $CDCl_3$) 8.32 (1H, s, Ar-H), 7.91 (1H, s, Ar-H), 6.03 (1H, d, *J* = 3.0 Hz 1'-C**H**), 6.02 (2H, br s, N**H**$_2$), 4.46 (1H, dd, *J* = 6.5, 3.0 Hz, 2'-C**H**), 5.00 (1H, dd, *J* = 6.5, 3.5 Hz, 3'-C**H**), 4.24 (1H, ddd, *J* = 5.9, 4.4, 3.5 Hz, 4'-C**H**), 3.02 (1H, dd, *J* = 13.4, 4.4 Hz, 5'-C**H**H'), 2.94 (1H, dd, *J* = 13.4, 5.9 Hz, 5'-CH**H'**), 1.73 (2H, br s, N**H**$_2$), 1.60 (1H, s, C**H**$_3$), 1.37 (1H, s, C**H**$_3$); δ_C (100 MHz, $CDCl_3$) 155.7 (**C**), 153.2 (**C**H), 149.4 (**C**), 139.9 (**C**H), 120.4 (**C**),

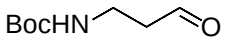
114.6 (**C**), 90.7 (1'-**C**H), 87.6 (4'-**C**H), 83.7 (2'-**C**H), 81.8 (3'-**C**H), 43.9 (5'-**C**H), 27.3 (**C**H_3), 25.4 (**C**H_3); HRMS (ESI$^+$) found 307.1518 ($C_{23}H_{19}N_6O_3^+$ requires 307.1513). Data are consistent with literature values.[67]

tert-Butyl (3-hydroxypropyl)carbamate (31)[88]



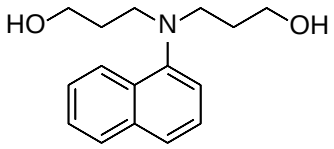
To a solution of 3-aminopropanol (3.00 g, 40.0 mmol) in dry CH_2Cl_2 (50 mL) at room temperature was added a solution of di-*tert*-butyl dicarbonate (9.60 g, 44.0 mmol) in dry CH_2Cl_2 (50 mL). The resulting mixture was stirred for 2 h after which the mixture was washed with sat. aq. $NaHCO_3$ (100 mL), water (100 mL) and brine (100 mL), dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash column chromatography (60 % EtOAc/petrol 40-60) to give the title compound as a colourless oil (3.5 g, 20.0 mmol, 50 %). R_f= 0.3 (60 % EtOAc/petrol 40-60); δ_H (400 MHz, $CDCl_3$) 4.75 (1H, br s, N**H**), 3.71 – 3.60 (2H, m, C**H**$_2$OH), 3.34 – 3.22 (2H, m, C**H**$_2$NHBoc), 1.71–1.58 (2H, m, CH_2C**H**$_2$$CH_2$), 1.44 (9H, s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 157.2 (**C**=O), 79.7 (**C**), 59.2 (**C**H_2), 36.8 (**C**H_2), 32.9 (**C**H_2), 28.4 (**C**H_3); HRMS (ESI$^+$) found 176.1283 ($C_8H_{18}NO_3^+$ requires 176.1281), found 198.1104 ($C_8H_{17}NNaO_3^+$ requires 198.1101); ν_{max}/cm^{-1} 3338 (br, O-H), 1682 (C=O), 1165 (C-O). Data are consistent with those reported by Muhsen.[67]

***tert*-Butyl (3-oxypropyl)carbamate (32)[67]**



General procedure for Swern Oxidation:[89] To a stirred solution of DMSO (0.60 mL, 8.45 mmol, 3.0 eq) in dry CH_2Cl_2 (12.5 mL) cooled to -78 °C was added oxalyl chloride (0.33 mL, 4.4 mmol, 1.5 eq.). The resulting mixture was stirred for 30 min after which a solution of alcohol (*tert*-butyl (3-hydroxypropyl)carbamate (**31**)) (500 mg, 2.85 mmol, 1.0 eq.) was added dropwise and stirred for 1 h. Triethylamine (1.85 mL, 13.3 mmol, 5.0 eq.) was added dropwise, stirred for 30 min and the mixture allowed to warm slowly to room temperature. 10 % aq. HCl (15 mL) and water (25 mL) were added to the mixture, the organic layer separated, and the aqueous layer further extracted with EtOAc (3 × 25 mL). The combined organic extracts were washed with sat. aq. $NaHCO_3$ (75 mL), brine (75 mL), dried ($MgSO_4$), filtered and concentrated under reduced pressure to give the title compound as a pale yellow oil (480 mg, 2.77 mmol, 97 %) which was used without further purification for subsequent reductive amination. δ_H (400 MHz, $CDCl_3$) 9.80 (1H, s, C**H**O), 4.90 (1H, br s, N**H**), 3.42 (2H, app q, J = 6.0 Hz, C$\mathbf{H_2}$N), 2.70 (2H, t, J = 5.8 Hz, $CH_2C\mathbf{H_2}CHO$), 1.42 (9H, s, $C(C\mathbf{H_3})_3$); δ_C (100 MHz, $CDCl_3$) 201.4 (**C**HO), 155.9 (**C**=O), 79.5 ($\mathbf{C}(CH_3)_3$), 44.3 ($\mathbf{C}H_2CHO$), 34.0 ($\mathbf{C}H_2N$), 28.4 ($C(\mathbf{C}H_3)_3$); HRMS (ESI^+) found 196.0946 ($C_8H_{15}NNaO_3^+$ requires 196.0944); ν_{max}/cm^{-1} 1687 (C=O), 1156 (C-O). Data are consistent with those reported by Muhsen.[67]

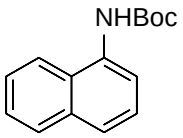
***3,3'-(Naphthalen-1-ylazanediyl)bis(propan-1-ol)* (39)**



Isolated when attempting to make 3-(naphthalen-1-ylamino)propan-1-ol (**35**) according to the procedure of Frade *et al.*[90] To naphthylamine (2 g, 14.0 mmol) was added a solution of 3-bromo-1-propanol (1 mL, 11.0 mmol) in ethanol (5 mL). The resulting solution was heated under reflux for 18 h. The reaction mixture was concentrated under reduced pressure and the resulting residue purified by flash column chromatography (0-5% MeOH/chloroform) to give the title compound as a purple oil (1.23 g, 4.7 mmol, 85 % [w.r.t. 3-bromo-1-propanol]). R_f = 0.2 (2 % MeOH/chloroform); δ_H (400 MHz, $CDCl_3$) 7.85-7.75 (2H, m, Ar-**H**), 7.59-7.39 (2H, m, Ar-**H**), 7.36 (1H, dd, *J* = 7.7, 7.6 Hz, Ar-**H**), 7.25 (1H, d, *J* = 8.1 Hz, Ar-**H**), 6.65 (1H, d, *J* = 7.6 Hz, Ar-**H**), 3.92 (2H, t, *J* = 5.7 Hz, C**H**$_2$OH^A)*, 3.81 (2H, t, *J* = 5.9 Hz, C**H**$_2$OH^B)*, 3.55 (2H, t, *J* = 6.4 Hz, C**H**$_2$N^B), 3.45 (2H, t, *J* = 6.2 Hz, C**H**$_2$N^A), 2.10 (2H, tt, *J* = 6.4, 5.9 Hz, CH$_2$C**H**$_2$CH$_2$^{B}), 2.05 (2H, tt, *J* = 6.2, 5.7 Hz, CH$_2$C**H**$_2$CH$_2$^{A}); δ_C (100 MHz, $CDCl_3$) 143.16 (Ar-**C**), 134.3 (Ar-**C**), 128.7 (Ar-**C**H), 126.6 (Ar-**C**H), 125.8 (Ar-**C**H), 124.9 (Ar-**C**H), 123.7 (Ar-**C**), 120.0 (Ar-**C**H), 118.0 (Ar-**C**H), 105.1 (Ar-**C**H), 62.2 (**C**H$_2$OH^A), 60.4 (**C**H$_2$OH^B), 43.0 (**C**H$_2$N^A), 35.0 (CH$_2$**C**H$_2$CH$_2$^{B}), 31.4 (CH$_2$**C**H$_2$CH$_2$^{A}), 30.4 (**C**H$_2$N^B); HRMS (ESI$^+$) found 202.1227 ($C_{13}H_{17}NO^+$ [3-(naphthalen-1-ylamino)propan-1-ol] requires 202.1226), expected ions not found; ν_{max}/cm^{-1} 3345 (O-H).

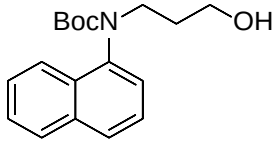
*A and B superscripts are used to distinguish between the two propanol groups.

***tert*-Butyl naphthalen-1-ylcarbamate (34)[91]**



Prepared according to the procedure of Kusakabe *et al.*[91] To a solution of naphthylamine (**33**) (6.00 g, 41.9 mmol) in dry THF (80 mL) was added a solution of di-*tert*-butyl dicarbonate (9.00 g, 41.9 mmol) in dry THF (20 mL) and the resulting mixture was stirred at 60 °C for 48 h then cooled to room temperature. The mixture was concentrated under reduced pressure and the resulting residue re-crystallised from petrol 40-60 to give the title compound as pale lilac needles (7.43 g, 30.6 mmol, 73 %). R_f = 0.7 (50 % Et_2O/petrol); δ_H (400 MHz, $CDCl_3$) 7.95–7.83 (3H, m, Ar-**H**), 7.63 (1H, d, *J* = 8.3 Hz Ar-**H**), 7.57–7.43 (3H, m, Ar-**H**), 6.84 (1H, br s, N**H**), 1.56 (9H, s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 153.6 (**C**=O), 134.1 (Ar-**C**), 133.0 (Ar-**C**), 128.7 (Ar-**C**H), 126.5 (Ar-**C**), 126.0 (Ar-**C**H), 125.9 (Ar-**C**H), 125.6 (Ar-**C**H), 124.5 (Ar-**C**H), 120.5 (Ar-**C**H), 118.7 (Ar-**C**H), 80.9 (C(**C**H$_3$)$_3$), 28.3 (**C**(CH$_3$)$_3$); HRMS (ESI$^+$) found 266.1153 ($C_{15}H_{17}NNaO_2^+$ requires 266.1151), found 312.1706 ($C_{18}H_{22}N_3O_2$ requires 312.1707); ν_{max}/cm^{-1} 1686 (C=O); m.p. = 79 – 80 °C (lit 96 – 99 °C)[92]. NMR data are consistent with literature values.[93]

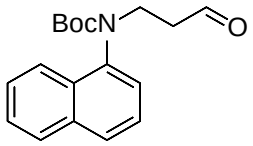
***tert-Butyl (3-hydroxypropyl)(naphthalen-1-yl)carbamate* (35)**



General procedure for alkylation: To a stirred suspension of sodium hydride (180 mg, 7.40 mmol) in dry DMF (20 mL), cooled to 0 °C was added dropwise a solution of *tert*-butyl naphthalen-1-ylcarbamate (**34**) (1.50 g, 6.20 mmol) in dry DMF (10 mL). The mixture stirred for 30 minutes at 0 °C and then warmed to room temperature. 3-Bromo-1-propanol (1.03 g, 7.40 mmol) was added dropwise and the mixture stirred for 3 h. Water (90 mL) was added to the mixture followed by extraction into Et_2O (3 × 120 mL). The combined organic extracts were washed with water (360 mL), brine (360 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by flash column chromatography (50 % Et_2O/petrol 40-60) to give the title compound as a colourless oil (1.00 g, 3.32 mmol, 54 %). R_f = 0.2 (50 % Et_2O/petrol); δ_H (400 MHz, $CDCl_3$) 7.92-7.85 (1H, m, Ar-**H**), 7.85-7.75 (2H, m, Ar-**H**), 7.57-7.48 (1H, m, Ar-**H**), 7.48-7.41 (2H, m, Ar-**H**), 7.22 (1H, d, *J* = 7.0 Hz, Ar-**H**), 4.21 (1H, dt, *J* = 14.0 Hz, 6.9 Hz, C**H**H'N), 3.85-3.73 (2H, m, C**H**$_2$OH), 3.70 (1H, br s, O**H**), 3.49 (1H, dt, *J* = 14.0, 5.0 Hz, CH**H**'N), 1.77-1.66 (2H, m, $CH_2C\mathbf{H}_2CH_2$), 1.24 (9H, s, $C(C\mathbf{H}_3)_3$); δ_C (100 MHz, $CDCl_3$) 138.5 (Ar-**C**), 134.6 (Ar-**C**), 130.4 (Ar-**C**), 128.4 (Ar-**C**H), 127.8 (Ar-**C**H), 126.6 (Ar-**C**H), 126.1 (Ar-**C**H), 125.5 (Ar-**C**H), 125.4 (Ar-**C**H), 122.7 (Ar-**C**H), 80.5 ($\mathbf{C}(CH_3)_3$), 58.9 ($\mathbf{C}H_2OH$), 46.1 ($\mathbf{C}H_2N$), 31.3 ($CH_2\mathbf{C}H_2\mathbf{C}H_2$), 28.0 ($C(\mathbf{C}H_3)_3$); HRMS ($ESI^+$) found 625.3260 ($C_{36}H_{46}N_2NaO_6^+$ requires

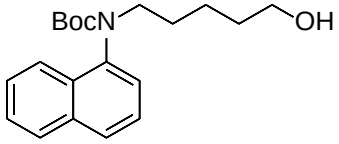
625.3248), found 324.1572 ($C_{18}H_{23}NNaO_3^+$ requires 324.1570); ν_{max}/cm^{-1} 3428 (br, O-H), 1672 (C=O).

tert-*Butyl naphthalen-1-yl(3-oxopropyl)carbamate* (37)



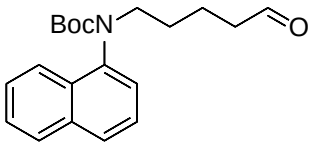
Prepared according to the general Swern Oxidation procedure (see above) using *tert*-butyl (3-hydroxypropyl)(naphthalen-1-yl)carbamate (**35**) (400 mg, 1.33 mmol) to give the title compound as a colourless oil (363 mg, 1.21 mmol, 91 %), which was used without further purification for subsequent reductive amination. δ_H (400 MHz, $CDCl_3$) 9.75 (1H, t, *J* = 1.8 Hz, C**H**O), 7.88 (1H, dd, *J* = 7.6, 1.7 Hz, Ar-**H**), 7.79 (2H, app. t, *J* = 8.3 Hz, Ar-**H**), 7.58-7.47 (2H, m, Ar-**H**), 7.46 (1H, apparent t, *J* = 7.6 Hz, Ar-**H**), 7.40-7.22 (1H, m, Ar-**H**), 4.28 (1H, dt, *J* = 13.9, 6.7 Hz, C**H**H'N), 3.86 (1H, dt, *J* = 13.9, 6.7 Hz, CH**H**'N), 2.73 (2H, td, *J* = 6.7, 1.8 Hz, C**H**$_2$CHO), 1.22 (9H, s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 200.8 (**C**HO), 155.3 (**C**=O), 138.2 (Ar-**C**), 134.5 (Ar-**C**), 130.6 (Ar-**C**), 128.5 (Ar-**C**H), 127.9 (Ar-**C**H), 126.6 (Ar-**C**H), 126.2 (Ar-**C**H), 125.5 (Ar-**C**H), 125.6 (Ar-**C**H), 122.7 (Ar-**C**H), 80.4 (**C**(CH$_3$)$_3$), 44.2 (**C**H$_2$N), 43.3 (**C**H$_2$CHO), 28.0 (C(**C**H$_3$)$_3$); HRMS did not find the expected ions; ν_{max}/cm^{-1} 1690 (C=O), 1720 (C=O).

tert-*Butyl (5-hydroxypentyl)(naphthalen-1-yl)carbamate* (36)



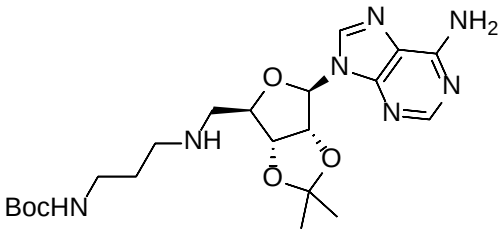
Prepared according to the general alkylation procedure (see above) using *tert*-butyl naphthalen-1-ylcarbamate (**34**) (1.50 g, 6.20 mmol) and 5-bromo-1-pentanol (1.24 g, 7.40 mmol) to give the title compound as a colourless oil (1.12 g, 1.21 mmol, 55 %). R_f = 0.2 (50 % Et_2O/petrol) δ_H (400 MHz, $CDCl_3$) 7.90-7.84 (1H, m, Ar-**H**), 7.84-7.74 (2H, m, Ar-**H**), 7.54-7.42 (3H, m, Ar-**H**), 7.33-7.27 (1H, m, Ar-**H**), 3.90 (1H, dt, *J*=14.5, 7.9 Hz, C**H**H'N), 3.60 (2H, d, *J*=6.5 Hz, C**H**$_2$OH), 3.51-3.38 (1H, m, CH**H**'N), 1.62-1.48 (4H, m, C**H**$_2$CH$_2$N and C**H**$_2$CH$_2$OH), 1.44-1.32 (2H, m, C**H**$_2$CH$_2$CH$_2$OH), 1.21 (9H, br s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 155.5 (**C**=O), 138.9 (Ar-**C**), 134.4 (Ar-**C**), 130.8 (Ar-**C**), 128.3 (Ar-**C**H), 127.4 (Ar-**C**H), 126.3 (Ar-**C**H), 126.0 (Ar-**C**H), 125.5 (Ar-**C**H), 125.4 (Ar-**C**H), 123.0 (Ar-**C**H), 79.8 (**C**(CH$_3$)$_3$), 62.7 (**C**H$_2$OH), 49.9 (**C**H$_2$N), 32.3 (C(**C**H$_3$)$_3$), 28.1 (**C**H$_2$CH$_2$OH), 24.4 (**C**H$_2$CH$_2$N), 22.9 (**C**H$_2$CH$_2$CH$_2$OH); HRMS (ESI$^+$) found 352.1878 ($C_{20}H_{27}NNaO_3$ requires 352.1883), found 681.3880 ($C_{40}H_{54}N_2NaO_6^+$ requires 681.3874); ν_{max}/cm^{-1} 3428 (br, O-H), 1672 (C=O).

tert-*Butyl naphthalen-1-yl(5-oxopentyl)carbamate* (38)



Prepared according to the general Swern Oxidation procedure (see above) using *tert*-butyl (5-hydroxypentyl)(naphthalen-1-yl)carbamate (**36**) (438 mg, 1.33 mmol) to give the title compound as a colourless oil (349 mg, 1.07 mmol, 80 %), which was used without further purification for subsequent reductive amination. δ_H (400 MHz, $CDCl_3$) 9.72 (1H, s, C**H**O), 7.90-7.83 (1H, m, Ar-**H**), 7.82-7.73 (2H, m, Ar-**H**), 7.55-7.40 (3H, m, Ar-**H**), 7.32-7.21 (1H, m, Ar-**H**), 3.99-3.81 (1H, m, C**H**H'N), 3.54-3.39 (1H, m, CH**H'**N), 2.49-2.37 (2H, m, C**H**$_2$CHO), 1.69-1.48 (4H, m, C**H**$_2$CH$_2$N and C**H**$_2$CH$_2$CHO), 1.21 (9H, br s, C(C**H**$_3$)$_3$); δ_C (100 MHz, $CDCl_3$) 202.2 (**C**HO), 155.4 (**C**=O), 138.8 (Ar-**C**), 134.4 (Ar-**C**), 130.7 (Ar-**C**), 128.4 (Ar-**C**H), 127.5 (Ar-**C**H), 126.4 (Ar-**C**H), 126.0 (Ar-**C**H), 125.5 (Ar-**C**H), 125.4 (Ar-**C**H), 122.9 (Ar-**C**H), 79.8 (**C**(CH$_3$)$_3$), 49.6 (**C**H$_2$N), 43.5 (**C**H$_2$CHO), 28.3 (CH$_2$**C**H$_2$CH$_2$), 28.1 (C(**C**H$_3$)$_3$), 19.3 (CH$_2$**C**H$_2$CH$_2$); HRMS (ESI$^+$) found 396.2285 ($C_{20}H_{25}NO_3$ + $C_3H_5N_2^+$ requires 396.2282); ν_{max}/cm^{-1} 1691 (C=O), 1722 (C=O).

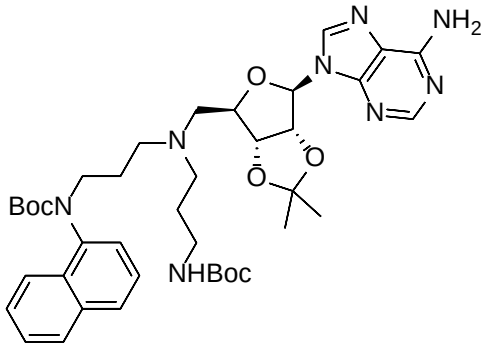
***tert*-Butyl (3-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)propyl)carbamate (25)[67]**



General procedure for first reductive amination:[69] A vigorously stirred suspension of 9-((3a*R*,4*R*,6*R*,6a*R*)-6-(aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9*H*-purin-6-amine (**24**) (858 mg, 2.80 mmol) and *tert*-butyl (3-oxypropyl)carbamate (400 mg, 2.30 mmol) in dry $(CH_2Cl)_2$ (14 mL) was warmed to 60 °C then allowed to cool to room temperature. Sodium triacetoxyborohydride (686 mg, 3.2 mmol) was added and the resulting mixture stirred at room temperature for 16 h. The reaction mixture was quenched with sat. aq. Na_2CO_3 (15 mL) and extracted with CH_2Cl_2 (4 × 20 mL). The combined organic extracts were dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (10 % sat. NH_3 in MeOH/EtOAc) to give the title compound as a white foam (534 mg, 1.15 mmol, 50 %). R_f = 0.2 (10 % sat. NH_3 in methanol/EtOAc); δ_H (400 MHz, $CDCl_3$) 8.34 (1H, s, Ar-H), 7.93 (1H, s, Ar-H), 5.99 (1H, d, *J* = 3.1 Hz, 1'-C**H**), 5.60 (2H, br s, N**H**$_2$), 5.50 (1H, dd, *J* = 6.3, 3.1 Hz, 2'-C**H**), 5.27 (1H, br s, N**H**$CO_2{}^tBu$), 5.04 (1H, dd, *J* = 6.3, 3.5 Hz, 3'-C**H**), 4.37 (1H, ddd, *J* = 6.8, 4.0, 3.5 Hz, 4'-C**H**), 3.34-3.08 (2H, m, C**H**$_2$N$HCO_2{}^tBu$), 2.92 (1H, dd, *J* = 12.5, 4.0, Hz, 5'C**H**H'), 2.82 (1H, dd, *J* = 12.5, 6.8 Hz, 5'CH**H**'), 2.76-2.65 (1H,

m, C**H**H'NH), 2.65-2.53 (CH**H**'NH), 1.74-1.66 (3H, m, $CH_2C\mathbf{H}_2CH_2N\mathbf{H}CH_2$), 1.62 (3H, br s, $C(C\mathbf{H}_3)_2$), 1.39 (9H, br s, $C(C\mathbf{H}_3)_3$), 1.41 (3H, br s, $C(C\mathbf{H}_3)_2$); δ_C (100 MHz, $CDCl_3$) 156.1 (**C**), 155.6 (**C**), 153.2 (Ar-**C**H), 149.5 (**C**), 140.1 (Ar-**C**H), 120.5 (**C**), 114.6 (**C**), 91.0 (1'-**C**H), 85.3 (4'-**C**H), 83.3 (2'-**C**H), 82.3 (3'-**C**H), 79.0 (**C**), 51.3 (5'-$\mathbf{C}H_2$), 47.9 ($CH_2\mathbf{C}H_2NHCH_2$), 39.3 ($\mathbf{C}H_2NHCO_2{}^t$Bu), 29.7 ($CH_2\mathbf{C}H_2CH_2$), 28.4 ($\mathbf{C}H_3$), 27.6 ($\mathbf{C}H_3$), 25.5 ($\mathbf{C}H_3$); HRMS (ESI^+) found 464.2617 ($C_{21}H_{34}N_7O_5^+$ requires 464.2616); ν_{max}/cm^{-1} 1694 (C=O). Data are consistent with those reported by Muhsen.[67]

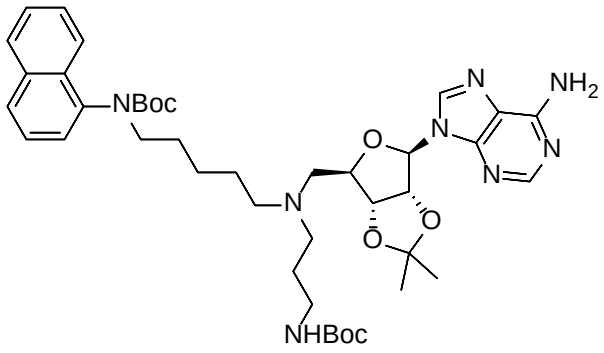
tert-*Butyl (3-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-((*tert-*butoxycarbonyl)amino)propyl)amino)propyl)(naphthalen-1-yl)carbamate* (26)



General procedure for second reductive amination:[69] A vigorously stirred suspension of *tert*-butyl (3-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)propyl)carbamate (**25**) (200 mg, 0.43 mmol), *tert*-butyl naphthalen-1-yl(3-oxopropyl)carbamate (**37**) (117 mg, 0.39 mmol) and $MgSO_4$ (470 mg, 4.30 mmol) in dry (CH_2Cl_2)

(4 mL) was stirred at 60 °C for 5 min then allowed to cool to room temperature and stirred for a further 1 h. Sodium triacetoxyborohydride (116 mg, 0.55 mmol) was added and the resulting mixture stirred for 4 h. The reaction mixture was quenched with sat. aq. Na_2CO_3 (25 mL) and then extracted with CH_2Cl_2 (4 × 25 mL). The combined organic extracts were dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash column chromatography (5% sat. NH_3 in MeOH/CH_2Cl_2) to give the title compound as a white foam (190 mg, 0.25 mmol, 65 %). R_f = 0.3 (5% sat. NH_3 in MeOH/CH_2Cl_2); HRMS (ESI^+) found 747.4175 ($C_{39}H_{55}N_8O_7$ requires 747.4188). Due to difficulties obtaining interpretable NMR data, the compound was subjected to global deprotection without further characterisation.

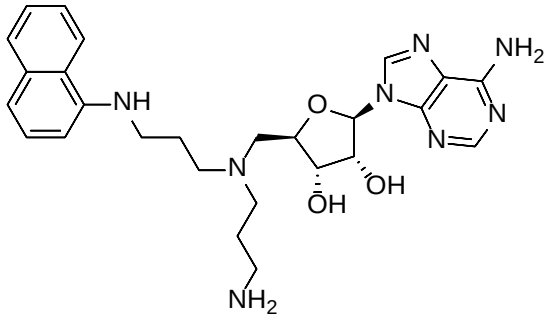
tert-*Butyl (5-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-((*tert-*butoxycarbonyl)amino)propyl)amino)pentyl)(naphthalen-1-yl)carbamate* (27)



Prepared according to general procedure for second reductive amination (see above) using *tert*-butyl (3-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-

dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)propyl)carbamate (**25**) (200 mg, 0.43 mmol), *tert*-butyl naphthalen-1-yl(5-oxopentyl)carbamate (**38**) (128 mg, 0.39 mmol) and $MgSO_4$ (470 mg, 4.30 mmol) to give the title compound as a white foam (199 mg, 0.26 mmol, 66 %). R_f = 0.3 (5% sat. NH_3 in MeOH/CH_2Cl_2); HRMS (ESI^+) found 775.4494 ($C_{41}H_{59}N_8O_7$ requires 775.4501). Due to difficulties obtaining interpretable NMR data, the compound was subjected to global deprotection without further characterisation.

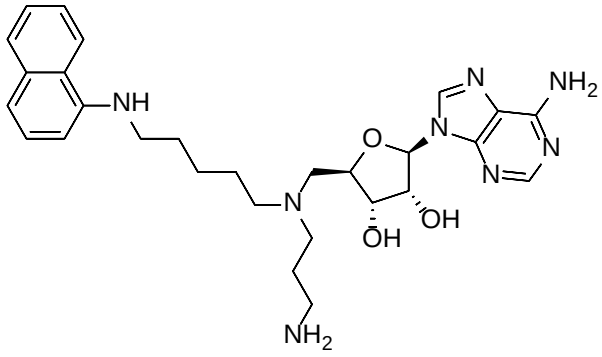
***(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-aminopropyl)(3-(naphthalen-1-ylamino)propyl)amino)methyl)tetrahydrofuran-3,4-diol* (EG047)**



General procedure for deprotection:[69] A solution of the intermediate tert-butyl (3-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-((*tert*-butoxycarbonyl)amino)propyl)amino)propyl)(naphthalen-1-yl)carbamate (**26**) in TFA (2 mL) and water (0.10 mL) was stirred for 6 h. The mixture was concentrated under reduced pressure, dissolved in water (5 mL) and washed with EtOAc (2 × 5 mL). To the aqueous layer was added Amberlite IRA-400(Cl^-) ion exchange resin. The mixture was stirred for 10 min then filtered. Water was

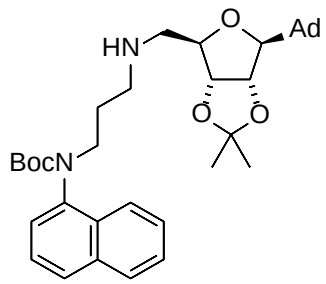
removed *in vacuo* to give an off-white powder (24 mg). HRMS (ESI^+) found 507.2834 ($C_{26}H_{35}N_8O_3$ requires 507.2827). Due to the presence of a contaminant, a full NMR assignment could not be made. NMR spectra and putative assignment is given in Appendix 6.6.

***(2R,3R,4S,5R)-2-(6-Amino-9H-purin-9-yl)-5-(((3-aminopropyl)(5-(naphthalen-1-ylamino)pentyl)amino)methyl)tetrahydrofuran-3,4-diol* (EG048)**



Prepared according to the general deprotection procedure (see above) using *tert*-butyl (5-((((3a*R*,4*R*,6*R*,6a*R*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-((*tert*-butoxycarbonyl)amino)propyl)amino)pentyl)(naphthalen-1-yl)carbamate (**27**) to give the title compound as an off-white solid (40 mg). HRMS (ESI^+) found 535.3146 ($C_{28}H_{39}N_8O_3$ requires 535.3140). Due to the presence of a contaminant, a full NMR assignment could not be made. NMR spectra and putative assignment is given in Appendix 6.6.

tert-*Butyl (3-((((3a*R*,4*R*,6*S*,6a*S*)-6-((1*s*,3*R*,5*R*,7*S*)-adamantan-1-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)propyl)(naphthalen-1-yl)carbamate*



Prepared according to general procedure for first reductive amination (see above) using 9-((3a*R*,4*R*,6*R*,6a*R*)-6-(aminomethyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine (**24**) (369 mg, 1.20 mmol) and *tert*-butyl naphthalen-1-yl(3-oxopropyl)carbamate (300 mg, 1.00 mmol) to give the title compound as a pale yellow foam (246 mg, 0.42 mmol, 42 %). R_f = 0.3 (5 % sat NH_3 in MeOH/Et_2O); δ_H (400 MHz, $CDCl_3$) 8.30 (1H, s, Ar-**H**), 7.92 (1H, br s, Ar-**H**), 7.89-7.82 (1H, m, Ar-**H**), 7.82-7.71 (2H, m, Ar-**H**), 7.54-7.37 (3H, m, Ar-**H**), 7.35-7.19 (1H, m, Ar-**H**), 5.99 (1H, d, *J* = 3.1 Hz, 1'-C**H**), 5.57 (2H, br s, N**H**$_2$), 5.40 (1H, ddd, *J* = 9.4, 6.5, 3.1 Hz, 2'-C**H**), 5.00-4.90 (1H, m, 3'-C**H**), 4.38-4.27 (1H, m, 4'-C**H**), 4.03-3.86 (1H, m, C**H**H'$NCO_2{}^tBu$), 3.57-3.40 (1H, m, CH**H**'$NCO_2{}^tBu$), 2.94-2.74 (2H, m, 5'-C**H**$_2$), 2.73-2.56 (2H, m, CH_2C**H**$_2$NH), 1.82-1.66 (3H, m, CH_2C**H**$_2CH_2$ and N**H**), 1.60 (3H, s, C(C**H**$_3$)$_2$), 1.52 (2H, br s, C(C**H**$_3$)$_3$), 1.40-1.33 (3H, m, C(C**H**$_3$)$_2$), 1.20 (7H, br s, C(C**H**$_3$)$_3$); δ_c (100 MHz, $CDCl_3$) 155.5 (**C**), 153.1 (Ar-**C**H), 150.1 (**C**), 149.5 (Ar-**C**), 147.3 (Ar-**C**), 139.9 (Ar-**C**H), 138.8 (Ar-**C**), 130.7 (Ar-**C**), 128.3 (Ar-**C**H), 127.5 (Ar-**C**H), 126.3 (Ar-**C**H), 126.0 (Ar-**C**H), 125.4 (Ar-**C**H), 123.0 (Ar-**C**H), 120.4 (Ar-**C**H),

114.6 (Ar-**C**H), 90.8 (3'-**C**H), 85.6 (4'-**C**H), 83.6 (2'-**C**H), 82.3 (3'-**C**H), 77.2 (**C**$(CH_3)_3$), 51.4 (5'-**C**H_2), 47.8 (**C**$H_2CO_2{}^tBu$), 47.1 (CH_2**C**H_2NHCH_2), 29.0 (CH_2**C**H_2CH_2), 28.1 (C(**C**$H_3)_2$), 27.3 (C(**C**$H_3)_2$); HRMS (ESI^+) found 590.3091 ($C_{31}H_{40}N_7O_5^+$ requires 590.3085); ν_{max}/cm^{-1} 1690 (C=O).

CHAPTER 3: **Result and Discussion**

3.1. **Summary of PRMT and Histone Constructs**

A summary of all PRMT1 and CARM1 constructs discussed in this thesis is given in Figure 19, and the exact sequences expressed are given in Table 12.

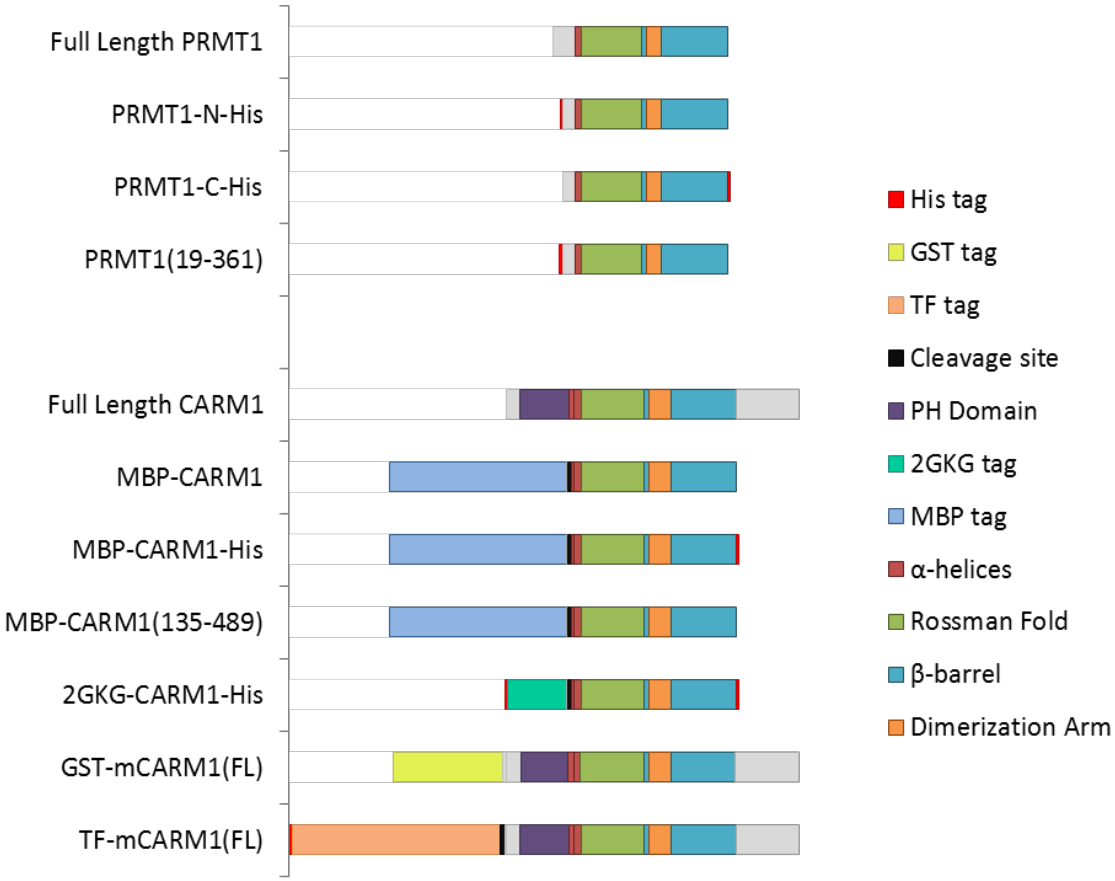


Figure 19: Schematic representation of PRMT1 and CARM1 constructs used for this thesis. Domains and tags are indicated using different colours. The length of each bar is proportional to the number of residues that comprise the domain it represents. Regions that are unstructured, or for which the structure is unknown are shown in grey.

Protein	Name	Residues	Vector	N-term modifications	C-term modifications	Mutations
human PRMT1	PRMT1-N-His	22-361	pET-26b	MGHHHHHH	-	-
	PRMT1-C-His	22-361	pET-26b	-	LGHHHHHH	-
	PRMT1(19-361)-N-His	19-361	pET-26b	MGHHHHHH	-	-
	PRMT1(E55N)	22-361	pET-26b	MGHHHHHH	-	E55N
	PRMT1(Y160N)	22-361	pET-26b	MGHHHHHH	-	Y160N
human CARM1	MBP-CARM1	135-479	pMALX(E)	(MBP) AALAAAQT NAAAENLYFQ*	-	-
	MBP-CARM1-His	135-479	pMALX(E)	(MBP) AALAAAQT NAAAENLYFQ*	HHHHHH	-
	MBP-CARM1(135-489)	135-489	pMALX(E)	(MBP) AALAAAQT NAAAENLYFQ*	-	-
	2GKG-CARM1-His	135-479	pSIMON2	(2GKG)AAAENLYFQ*	HHHHHH	-
	CARM1 (N161E)	135-479	pSIMON2	(2GKG)AAAENLYFQ*	HHHHHH	N161E
	CARM1 (N265Y)	135-479	pSIMON2	(2GKG)AAAENLYFQ*	HHHHHH	N265Y
mouse CARM1	GST-mCARM1-FL	3-608	pGEX-1λT	(GST)FAAAST	-	-
		1-608	GST-pCOLD	(GST)LG VFQ*GPGGH	-	-
	TF-mCARM1-FL	1-608	TF-pCOLD	(TF)SAGLG VLFQ*GPSAH[1]	-	-

[1]The TF vector has additional residues at the N-terminus of (TF), including a 6His tag

Table 12: PRMT1 and CARM1 constructs used in this thesis. Single letter codes are used to denote amino acids. Whole proteins are enclosed in brackets. Cleavage sites are underlined, with asterisks used to mark the point of cleavage.[3]

Cleavable fusion tags were used to enhance the solubility or ease of purification of several CARM1 constructs. Figure 20 and Table 13 summarise the products formed upon treatment with the indicated protease, and the names used to denote them herein. Though it contains a cleavage site, GST-mCARM1 was used as the fusion protein, therefore it is not included in Table 13.

[3] Numbering used throughout this thesis is relative to CARM1 isoform 3 (Uniprot code Q86X55-3). The sequence of PRMT1 isoform 1 (Uniprot code Q99873-1) has been updated during the preparation of this manuscript. The sequence to which the numbering corresponds is given in Appendix 6.2.

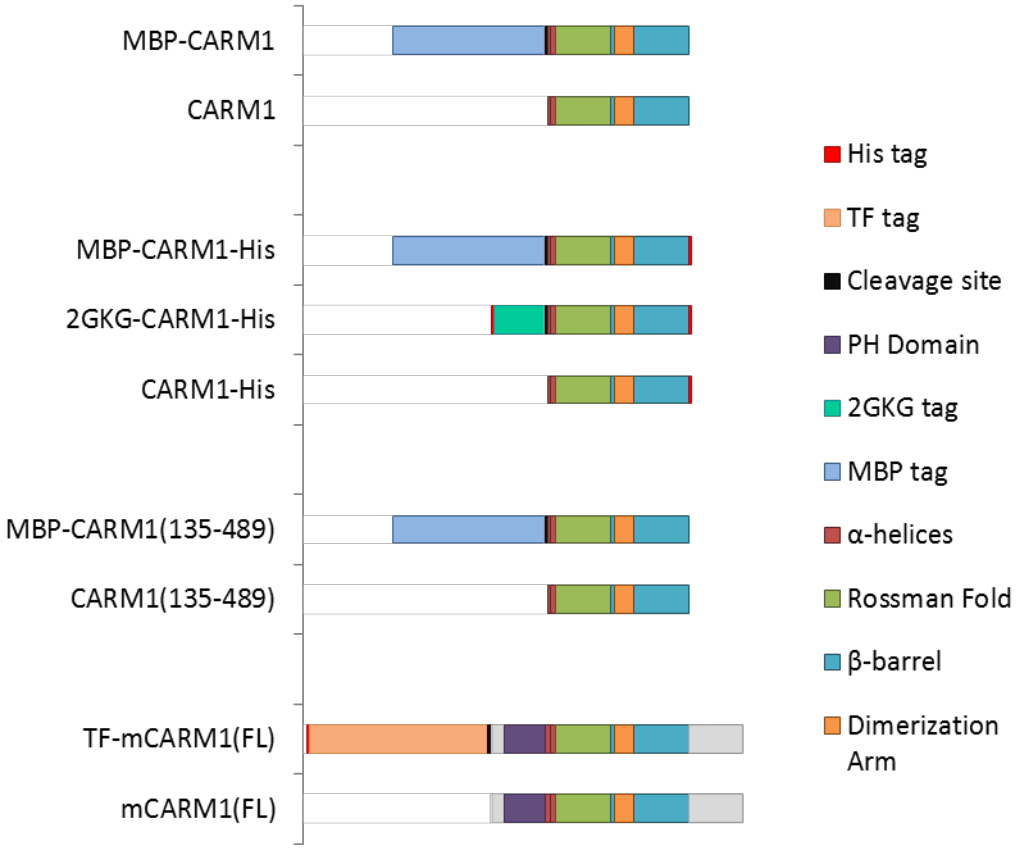


Figure 20: Schematic representation of cleavage products of cleavable CARM1 fusion proteins used in this thesis. Constructs that cleave to give the same CARM1 construct are grouped together.

Fusion Protein	Protease	Cleaved Protein	N-terminal modifications	C-terminal modifications
MBP-CARM1	TEV	CARM1	-	-
MBP-CARM1-His	TEV	CARM1-His	-	HHHHHH
2GKG-CARM1-His	TEV	CARM1-His	-	HHHHHH
MBP-CARM1(135-489)	TEV	CARM1(135-489)	-	-
TF-mCARM1-FL	HRV C3	CARM1(3-608)	GPSAH[1]	-

[1]Additional residues present as a consequence of cloning and the HRV C3 cleavage.

Table 13: CARM1 constructs generated by treating CARM1 fusion proteins with the appropriate protease.

In addition to PRMT1 and CARM1, several constructs of histones 3 and 4 were cloned, expressed and purified for use as substrates, or for co-purifications. These are shown in Figure 21 and Table 14.

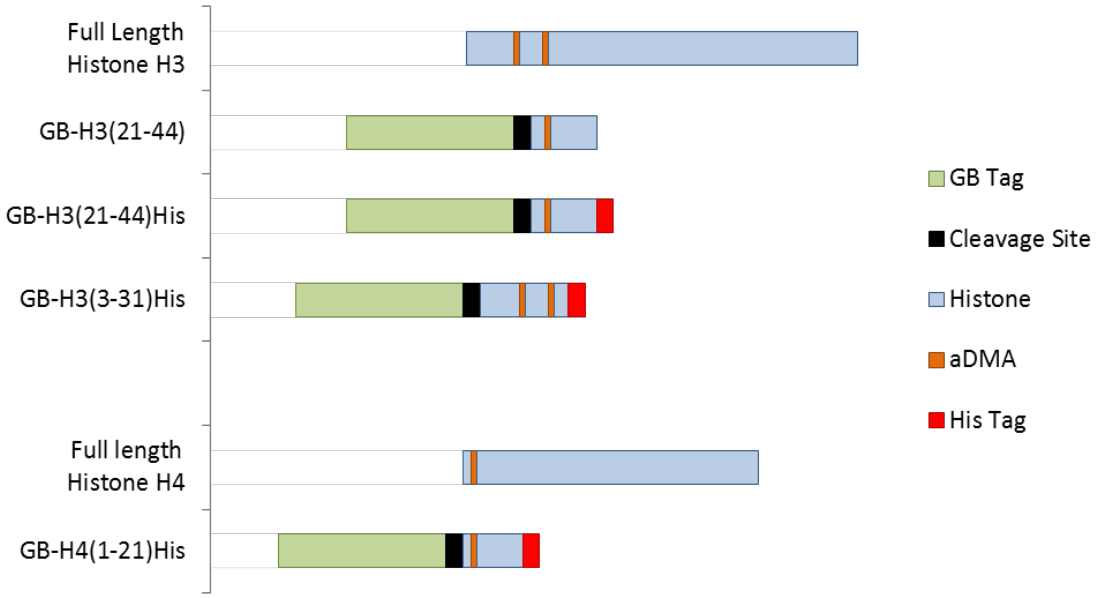


Figure 21: Schematic representation of histone constructs used in this thesis. Asymmetric dimethylation sites targeted by PRMT1 and CARM1 are highlighted in orange.

Protein	Name	Boundaries	Vector	N-terminal modifications	C-terminal modifications
Histone 3	GB-H3(21-44)	21-44	GBFusion pET-21a	(GB)LEVLFQ*GP	-
	GB-H3(21-44)His	21-44	GBFusion pET-21a	(GB)LEVLFQ*GP	HHHHHH
	GB-H3(3-31)His	3-33	GBFusion pET-21a	(GB)ENLYFQ*G	HHHHHH
Histone 4	GB-H4(1-21)His	135-479	GBFusion pET-21a	(GB)ENLYFQ*S	HHHHHH

Table 14: Histone 3 and histone 4 constructs used in this thesis. Single letter codes are used to denote amino acids. Whole proteins are enclosed in brackets. Cleavage sites are underlined, with asterisks used to mark the point of cleavage.

The initial conditions for the expression and purification of each construct were based on literature procedures (see Section 2.2.1). The volume of bacterial culture grown, mass of cell pellet, and mass of protein at each stage of a representative purification of each construct are summarized in Table 15, below and discussed further in subsequent sections.

Construct	Volume of bacterial culture (L)	Mass of cell pellet (g)	Approximate mass of protein after affinity column (mg)	Mass of protein after gel filtration (mg)	Mass of protein after TEV cleavage and subsequent gel filtration (mg)
PRMT1-C-His	-[1]	-	-	-	
PRMT1-N-His	4	13.1	13.8	10.4	10.4[2]
PRMT1(19-361)-N-His	4	10.8	4.0	Precipitated during concentration	N/A[3]
MBP-CARM1	0.8	2.9	4.0	N/A	Precipitated during TEV cleavage. 0.7 mg/mL left in supernatant. 0.2 mg after subsequent gel filtration.
MBP-CARM1-His	4	16.4	2.2	1.3	0.46
MBP-CARM1(135-489)	0.8	2.6	3.5	N/A	0.24
2GKG-CARM1-His[4]	8	-	-	N/A	2
GST-mCARM1-FL (pCOLD vector)	4	11.0	0.14	N/A	N/A
TF-mCARM1-FL	5	4.0	-	32	N/A
GB-H3(3-31)His	4	12.0	-	25	N/A
GB-H4	0.8	3.0	-	13	N/A
GB-H3 (21-44)	0.8	3.0	-	16	N/A

[1]Dashes indicate that the mass was not recorded
[2]However, on occasion, the protein would precipitate or stick to the concentrator resulting in significantly lower yields.
[3]N/A indicates that this step was not included in the purification
[4]0.4 mM IPTG used for induction of expression

Table 15: Volume of bacterial culture grown, mass of cell pellet, and mass of protein at each stage of a representative purification of each construct.

3.3. PRMT1: Purification and attempted Crystallisation

3.3.1. PRMT1 Construct Design

Rat $PRMT1_{14-353}$ has been successfully crystallised and shown to be catalytically active *in vitro*.[26] This sequence differs from human $PRMT1_{22-361}$ by a single residue (H161 in rat is Y in the human sequence), and both display comparable *in vitro* methyltransferase activity.[27f] We therefore chose the domain boundaries S22 to R361 for our PRMT1 constructs.

3.3.2. PRMT1 Purification

Initially, two PRMT1 constructs were cloned: one with a His tag at the N-terminus (PRMT1-N-His, Figure 19), and one with a His tag at the C-terminus (PRMT1-C-His, Figure 18). PRMT1-C-His expressed poorly. SDS-PAGE analysis (Figure 21C, below) did not show any clear increase in levels of PRMT1 in the lysate compared to the pre-induction sample (which would be evident as a band at 40.4 kDa). Subsequent nickel affinity chromatography and gel filtration achieved partial separation of a ~40 kDa protein; however the gel filtration UV trace comprised at least 5 overlapping peaks (Figure 22B, below). This elution profile is indicative of contaminant proteins or multiple oligomeric states of the target protein, both of which would likely be detrimental to crystallisation. Further optimisation was not carried out on this construct.

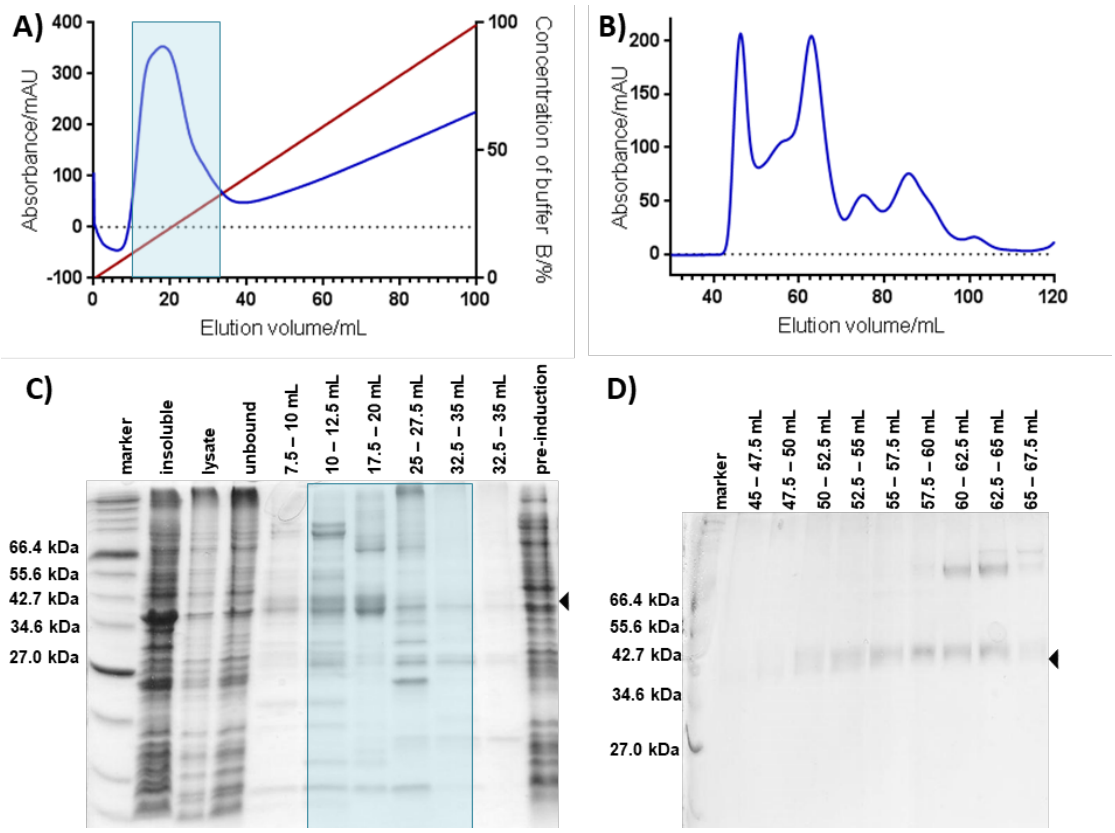


Figure 22: Purification of PRMT1-C-His. (A) Nickel column chromatogram showing UV trace (blue) and percentage buffer B (red). Fractions from 10 – 32.5 mL (turquoise box) were combined and concentrated for gel filtration; (B) Gel filtration (Superdex 200 16/60) chromatogram; (C) SDS-PAGE of expression, lysis samples and nickel chromatography fractions; (D) SDS-PAGE of gel filtration fractions. A band with (approximately) the correct MW for PRMT1-C-His (MW = 40.4 kDa) is present in fractions from 50 – 62.5 mL. SDS-PAGE of later fractions (not shown) did not show any bands of MW corresponding to PRMT1. An arrow is used on SDS-PAGE images to indicated the approximate anticipated height of PRMT1-C-His.

The N-terminally His-tagged construct (PRMT1-N-His) expressed well. PRMT1-N-His was obtained in high purity from nickel affinity chromatography, and eluted from gel filtration as an oligomer of molecular weight greater than 440 kDa (Figure 23). The isolated protein showed significant methyltransferase activity (compared to a substrate-only control) using a radiometric activity

assay with a histone 4 substrate (Figure 24, page 91). These results indicate that pure, active PRMT1-N-His can be obtained through nickel affinity chromatography followed by gel filtration.

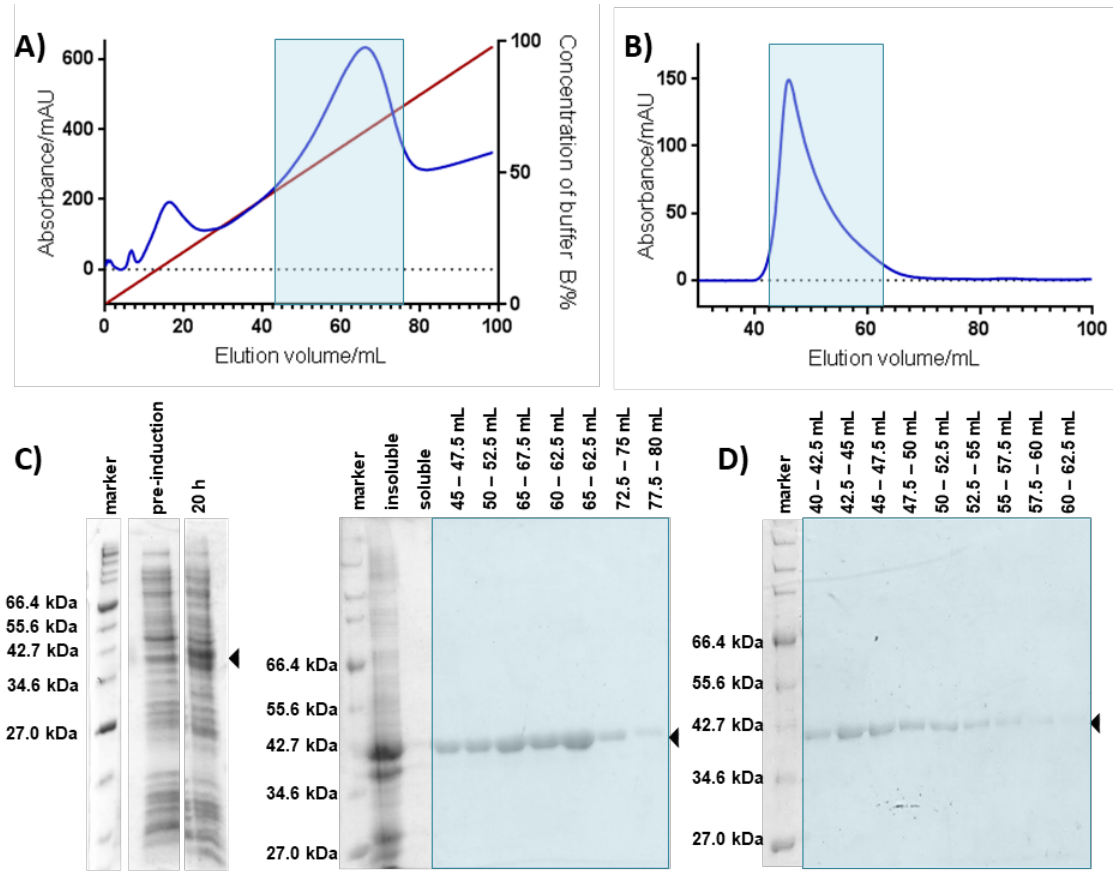


Figure 23: Purification of PRMT1-N-His. (A) Nickel column chromatogram with UV trace (blue) and percentage buffer B (red). Fractions from 45 – 77.5 mL (turquoise box) were combined and concentrated for gel filtration; (B) Gel filtration (Superdex 200 16/60) chromatogram. PRMT1 elutes at 47 mL, corresponding to an MW of >440 kDa, as an asymmetric peak with a long tail. All PRMT1-containing fractions were combined for subsequent steps; (C) SDS-PAGE of expression, lysis samples and nickel chromatography fractions; (D) SDS-PAGE of gel filtration fractions. An arrow is used on SDS-PAGE images to indicated the approximate anticipated height of PRMT1-N-His.

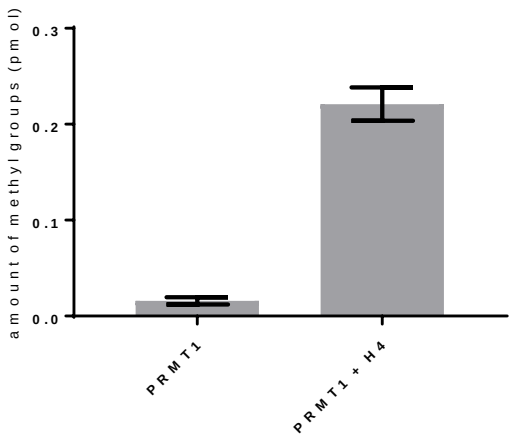


Figure 24: Initial activity test of PRMT1-N-His. Error bars show the standard deviation from 3 repeats.

The observed gel filtration elution volume of ~47 mL is indicative of species which are too large to permeate the gel filtration matrix. This is usually attributed to high molecular weight aggregates of misfolded protein. PRMT1 is known to form oligomers exceeding 600 kDa,[26] which is the upper size limit for proteins resolvable using an S200 gel filtration column. It was therefore difficult to distinguish between misfolded PRMT1 aggregates and high molecular weight oligomers based solely on the gel filtration elution profile. Additionally, the asymmetric peak shape is indicative of multiple oligomeric states, which would likely be detrimental to crystallisation. We therefore sought to modify the purification to obtain a monodisperse sample of PRMT1, which was expected to be indicated by a single, symmetrical peak in the gel filtration UV trace.

3.3.2.1. Modifications to PRMT1 Purification: Effect of SAH and DTT Incubation and Addition N-terminal Residues

The oligomerisation behaviour of PRMT1 has been studied by Zhang and Cheng,[26] who found that the size of PRMT1 oligomers (as judged by gel filtration and dynamic light scattering) is altered in the presence of SAH. In published crystal structures (PDB codes 1OR8, 1ORH, 1ORI), rat PRMT1 forms extended chains in which PRMT1 dimers are connected at either end by a disulphide bond between C254 in neighbouring proteins. Mutation of this cysteine to serine (i.e. C254S) has been shown to result in a reduction in oligomer size. With the hope of achieving reduced oligomer size without introducing non-native amino acids, we investigated the effect of incubating the pooled nickel fractions with either SAH or DTT (a reducing agent commonly used to prevent disulphide bond formation in proteins) prior to gel filtration. In both cases, a substantial change in the elution profile was observed, but the peak shapes still indicated polydisperse samples (Figure 25A and B, page 93). Crystallisation trials of PRMT1 with inhibitor UM249 (Figure 15, page 27) in the presence of TCEP (a reducing agent with a longer half-life in solution that DTT) were performed, but did not produce any protein crystals (Table 16, entry 4, page 101).

In a further bid to obtain a soluble, monodisperse sample of PRMT1, an alternative construct, PRMT1(19-361)-N-His (Figure 19, page 83) was cloned, expressed and purified. A similar elution profile to the shorter PRMT1 construct was observed during nickel affinity and gel filtration chromatography (Figure

25C, below); however solubility was greatly reduced, with the protein precipitating when concentrated to 1.20 mg/mL in gel filtration buffer. This low solubility is in line with the observations of Zhang and Cheng, who found the rat analogue of this construct to be poorly soluble,[26] and makes this construct unsuitable for use in crystallisation trials and assays.

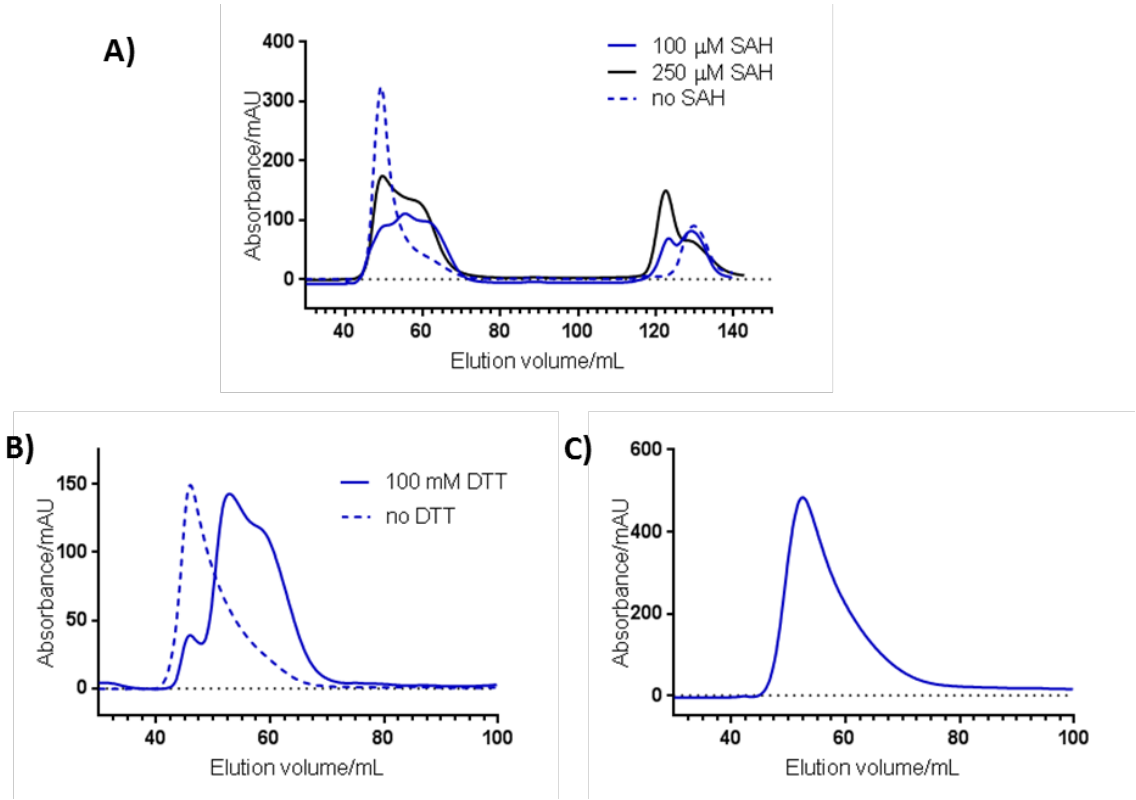


Figure 25: Effect of various modifications to the purification procedure/construct on gel filtration (Superdex 200 16/60) elution profile. (A) Pooled nickel fractions (45 µM PRMT1) incubated with 100 or 250 µM SAH for 30 min at r.t. prior to gel filtration; (B) Pooled nickel fractions incubated with 100 mM DTT for 40 min at r.t. prior to gel filtration; (C) Extension of PRMT1 construct (PRMT1(19-361)). For (A) and (B), the pooled, concentrated nickel fractions were split into two aliquots prior to the addition of SAH/DTT, and the sample without SAH/DTT was run as a control (dotted line).

Despite the failure to obtain a monodisperse sample of PRMT1, several crystallisation trials were set up, in the hope that the presence of multiple oligomeric states would not impede crystallisation (Table 16, page 101, entries 1, 2 and 4). Unfortunately, no protein crystals have formed.

3.3.2.2. Modifications to PRMT1 purification: Effect of Anion Exchange

In a further bid to obtain a monodisperse sample of PRMT1, the addition of an anion exchange step prior to gel filtration was tested. PRMT1-N-His eluted from anion exchange with an asymmetric elution profile with a long trailing tail (Figure 26B, page 95). Analytical gel filtration of selected anion exchange fractions (Figure 26C, page 95) suggested that later-eluting fractions from anion exchange contained larger oligomers (as indicated by earlier elution from gel filtration); however none of the fractions gave a symmetrical gel filtration peak. Unexpectedly, combining several anion exchange fractions and running on a preparative gel filtration column (Figure 26D, page 95) gave a broad but symmetrical peak. Crystallisation trays were set up with the eluting protein but no crystals grew. It is unclear why the gel filtration elution profile was altered, since the composition of the anion exchange buffer in which PRMT1 eluted is similar to that from nickel chromatography (Table 7, page 46 and Section 2.3.6) except for the absence of imidazole. It is possible that the high concentrations of imidazole required to elute PRMT1-N-His during nickel affinity chromatography influence PRMT1 oligomerisation. Another explanation is that

the lower concentration of the sample loaded following anion exchange resulted in lower order PRMT1 oligomers.

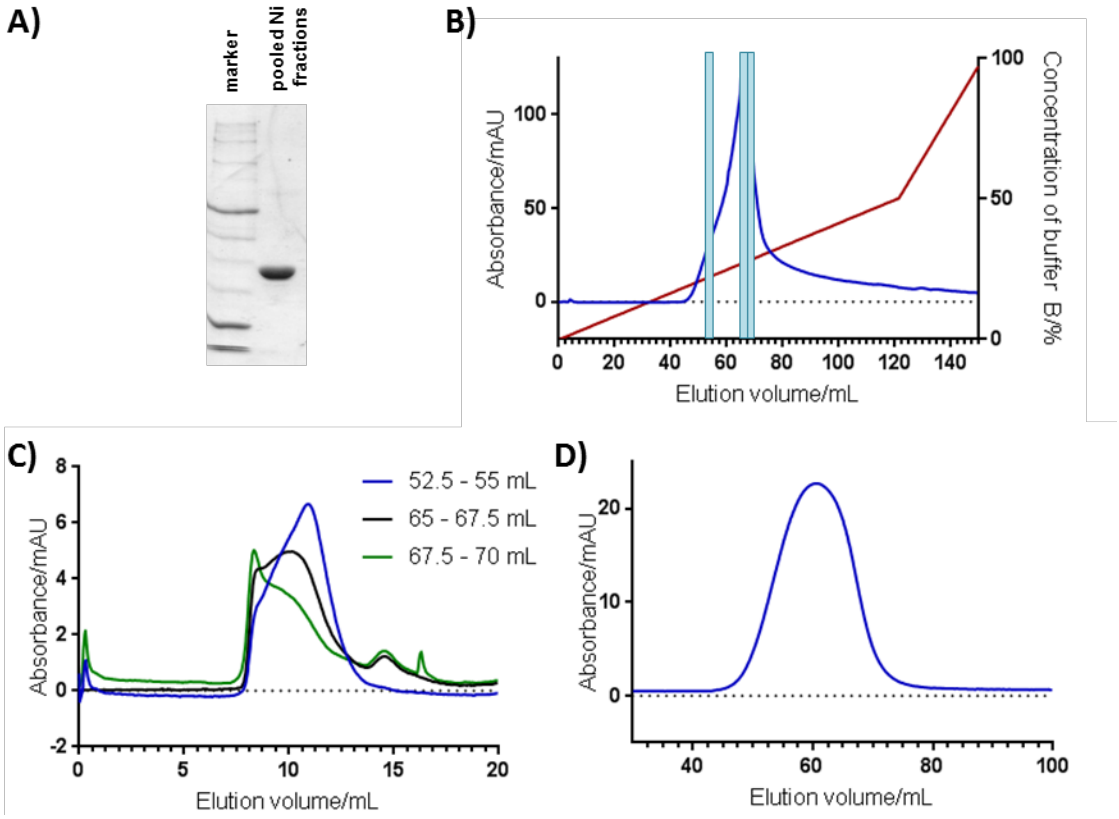


Figure 26: Anion exchange of pooled nickel affinity fractions. (A) SDS-PAGE of pooled nickel affinity fractions; (B) Anion exchange chromatogram. Fractions eluting between 65 – 70 mL were combined for preparative gel filtration. Samples from selected anion exchange fractions (turquoise boxes) were retained and for analytical gel filtration; (C) Analytical gel filtration (Superdex 200 10/300) chromatogram. The anion exchange fraction that corresponds to each gel filtration trace is indicated in the legend; (D) Preparative gel filtration (Superdex 200 16/60) chromatogram.

3.3.2.3. Modifications to PRMT1 Purification: Effect of Histone Co-purification

As a final attempt at obtaining a PRMT1 sample suitable for crystallisation, PRMT1 was co-purified with a GB-fused histone 4 peptide (GB-H4(1-21)-His, Figure 21, page 86). GB is the 56-residue B domain of *Streptococcal* protein G, and can be used as a fusion partner to enhance the expression of small recombinant proteins and peptides in *E. coli*.[94] PRMT1 is known to methylate H4 at arginine 3 *in vivo*,[95] and act on H4 peptides containing residues 1-21 *in vitro*.[27f] At the time of this experiment, there was still evidence to suggest a rapid equilibrium random mechanism (as opposed to the now confirmed sequential mechanism – see section 1.2.2) therefore it was a reasonable decision at the time. To this end, GB-H4(1-21)-His was cloned and expressed. A TEV-cleavable GB tag was included to aid expression and allow visualisation of the substrate using SDS-PAGE. Cell pellets from bacterial expression of PRMT1-N-His and GB-H4(1-21)-His were combined prior to sonication, centrifugation and chromatography (see section 2.3.8 for details). Both PRMT1-N-His and GB-H4(1-21)-His bound to and eluted from the nickel column, and all fractions containing either protein were combined and concentrated for gel filtration. The gel filtration elution profile showed a clear shift in the elution volume of PRMT1, and an improved peak shape when compared to other purification strategies (Figure 27). SDS-PAGE analysis indicated that the symmetrical peak at *ca*. 65 mL corresponds to the PRMT1·GB-H4 complex, while the shoulders at lower elution volume contain PRMT1 alone.

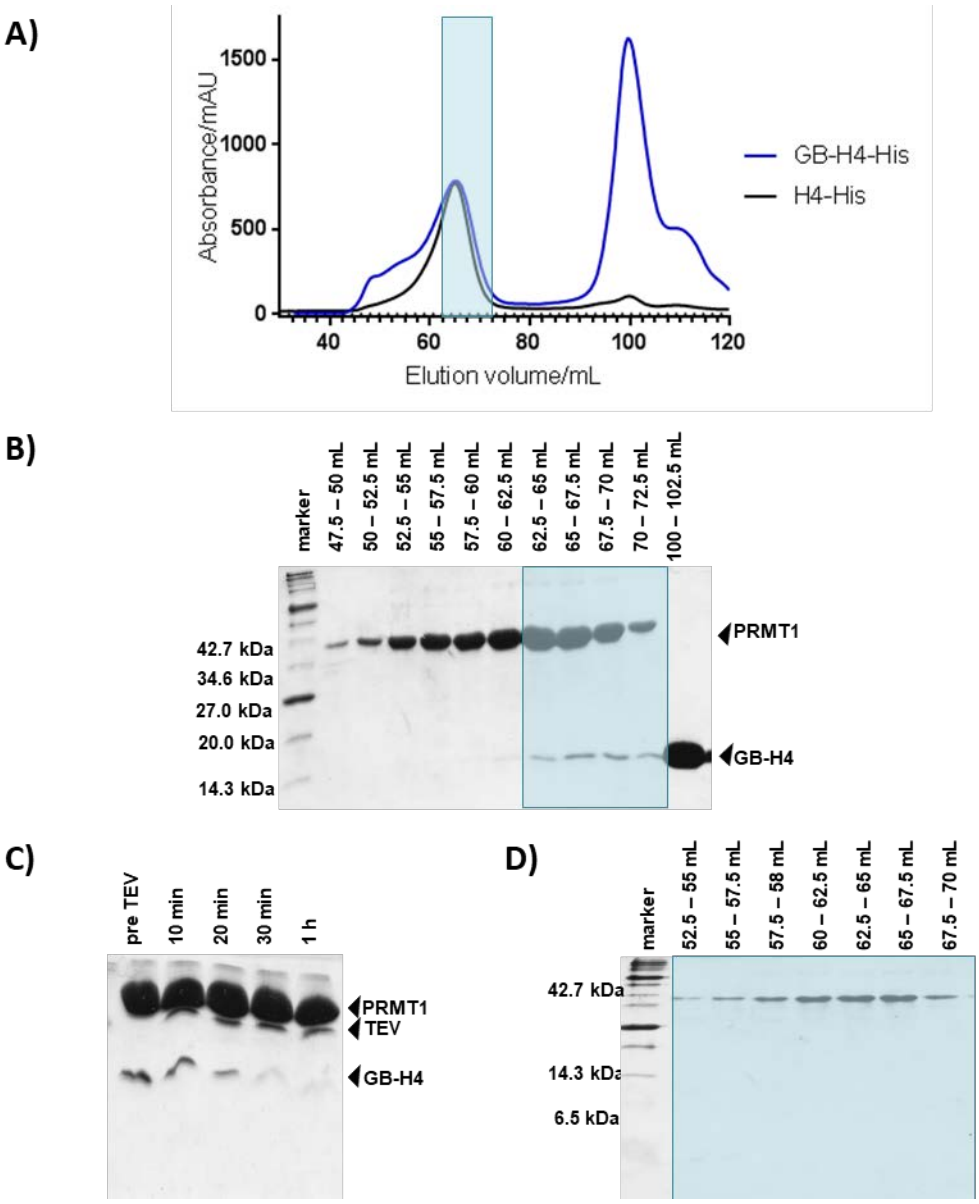


Figure 27: Co-purification of PRMT1 and GB-H4. (A) Gel filtration (Superdex 200 16/60) chromatogram both before (blue) and after (black) cleavage of the GB tag from GB-H4. Absorbance of the two traces has been normalised to allow comparison of peak shapes and elution volumes; (B) SDS-PAGE of gel filtration fractions. Fractions from 65 – 72.5 mL (turquoise box) were combined for TEV cleavage; (C) SDS-PAGE to show time course of TEV protease cleavage of GB-H4. The elution volume of PRMT1 is the same as that observed for the PRMT1·GB-H4 complex, suggesting that the PRMT1 remained bound to H4 once the GB tag had been cleaved. (D) SDS-PAGE of fractions eluting from gel-filtration after TEV cleavage of the GB tag. The H4 peptide is not visible, but this is likely due to its low MW (3.1 kDa) resulting in a diffuse band at the bottom of the gel. Fractions from 62.5 – 72.5 mL (turquoise box) were combined for crystallisation trials.

The NaCl concentration of all buffers was lowered to 150 mM for this co-purification.

The isolated PRMT1·GB-H4 complex (fractions 62.5 – 72.5 mL) was treated with TEV protease, which cleaved the GB tag from the histone (Figure 27C, above). Due to the small size of the non-tagged histone peptide, it was not possible to test for the presence of the cleaved peptide using SDS-PAGE. A second gel filtration of the cleavage mixture was performed. The elution profile closely matched that of the PRMT1-GB-H4 complex (Figure 27A, above). This suggests that the PRMT1-N-His eluting from gel filtration had the cleaved H4 peptide bound. Crystallisation trials were set up using the putative complex (Table 16, page 100, entry 3). Crystalline material was visible in well C3 of the PACT screen (Figure 28, page 100), but did not diffract and was not reproducible upon optimisation. Lower salt buffers (150 mM NaCl as opposed to the typical 300 mM NaCl - see Table 7, page 46) were used for this co-purification, since high salt concentrations can impede electrostatic enzyme-substrate interactions. It is therefore possible that the observed shifts in elution volume were due to this alteration.

The oligomerisation behaviour of PRMT1 is still poorly understood. Temperature, pH, concentration, ligands and reducing agents have all been shown to influence the apparent equilibria between PRMT1's oligomeric states.[26, 96] The reversibility of this oligomerisation means that a sample which appears monodisperse as judged by gel filtration may not remain so on addition of precipitant solution when setting up crystallisation trays. Since sample homogeneity is of vital importance to crystal growth, it may be prudent, to explore methods (e.g. dynamic light scattering, analytical gel filtrations) for

assessing PRMT1 speciation under conditions appropriate for crystal formation before further gel filtration studies and crystallisation trials. The addition of mild detergents such as Triton X-100, TWEEN-20 or more specialised detergents used for the purification of membrane proteins[97] may also be worth investigating. The only reported wild type crystal structures of PRMT1 (rat $PRMT1_{14-353}$, see Section 1.2.1) were obtained at pH 4.7, at which the enzyme is inactive.[26] Overall, further work is needed to develop a reliable method for generating diffraction quality PRMT1·inhibitor co-crystals under physiologically relevant conditions.

Entry	Gel Filtration Buffer	Components added prior to gel filtration	Components added prior to crystallisation trial	Protein concentration		Drop Size/nL	Screens	Crystals?
				mg/mL	µM			
1	300 mM NaCl, 20 mM Tris, 2% glycerol, pH 7.5	SAH (5 molar eq.)	SAH (25 µM)[1]	2.6 + 1.3	65 + 32	240	JCSG+	no
2	300 mM NaCl, 20 mM Tris, 5% glycerol, pH 8.0	DTT (100 µM)	DTT (100 µM)	10 + 5	248 + 124	400	JCSG+, PACT	no
			Sinefungin (500 µM)	5	124	400	JCSG+	no
3	150 mM NaCl, 20 mM Tris, 5% glycerol, pH 8.0	H4 (co-purification)	-	5	124	400	PACT, Structure	C3, structure[3]
			SAH (250 µM)[2]	2.5	62	400	PACT	no
			Sinefungin (500 µM)[2]	2.5	62	400	PACT	no
4	300 mM NaCl, 20 mM Tris, 5% glycerol, pH 8.0	DTT (50 mM)	UM249 (167 µM)	3.2	79	400	SG1, LMB, JCSG+, Morpheus, PACT	no
			UM249 (167 µM), TCEP (10 mM)	3.2	79	400	SG1, LMB, JCSG+, Morpheus, PACT	salt

[1]Right-hand shoulder of peak eluting from gel filtration was used for crystallisation, and was anticipated to have SAH bound prior to this supplementation.
[2]Protein purified in the gel filtration buffer stated was diluted 1:1 (v/v) with either 0.5 mM SAH or 1 mM Sinefungin.
[3]Did not diffract, and no crystals obtained upon attempted optimisation

Table 16: PRMT1 crystallisation trials set up to date.



C3 Structure Screen
1.5 M Li_2SO_4
0.1 M HEPES
pH 7.5

Figure 28: Appearance of apparent crystalline material in C3 Structure screen (entry 3, Table 16)

3.4. CARM1: Purification and Crystallisation

3.4.1. CARM1 Construct Design

At the time of beginning this work, 8 structures of the catalytic core of CARM1 were available in the PDB (PDB codes 2Y1X, 2Y1W, 2V7E, 2V74, 3B3G, 3B3F, 3B3J, 41KP). Residues beyond T479 are disordered in these CARM1 structures. Density for S135 (but not residues N-terminal to it) was visible when included in the crystallisation construct (e.g. PDB code 2Y1X). We therefore chose to use the domain boundaries S135 to T479. Sack *et al.* reported successful expression and purification of CARM1 with a cleavable MBP solubility tag.[16b] We therefore assembled an expression vector for MBP-CARM1 and MBP-CARM1-His (Table 12, page 84 and Figure 19, page 83).

3.4.2. CARM1 Purification

MBP-CARM-His was expressed in *E. coli* and purified by nickel affinity chromatography. MBP-CARM1-His typically eluted at ~30 – 50 % buffer B, with several contaminant proteins co-eluting in early fractions. To avoid the potential for these presumed contaminants to co-purify in subsequent purification steps, only fractions without contaminant proteins (25 – 37.5 mL in the representative purification shown in Figure 29, below) were combined and used for purification by gel filtration.

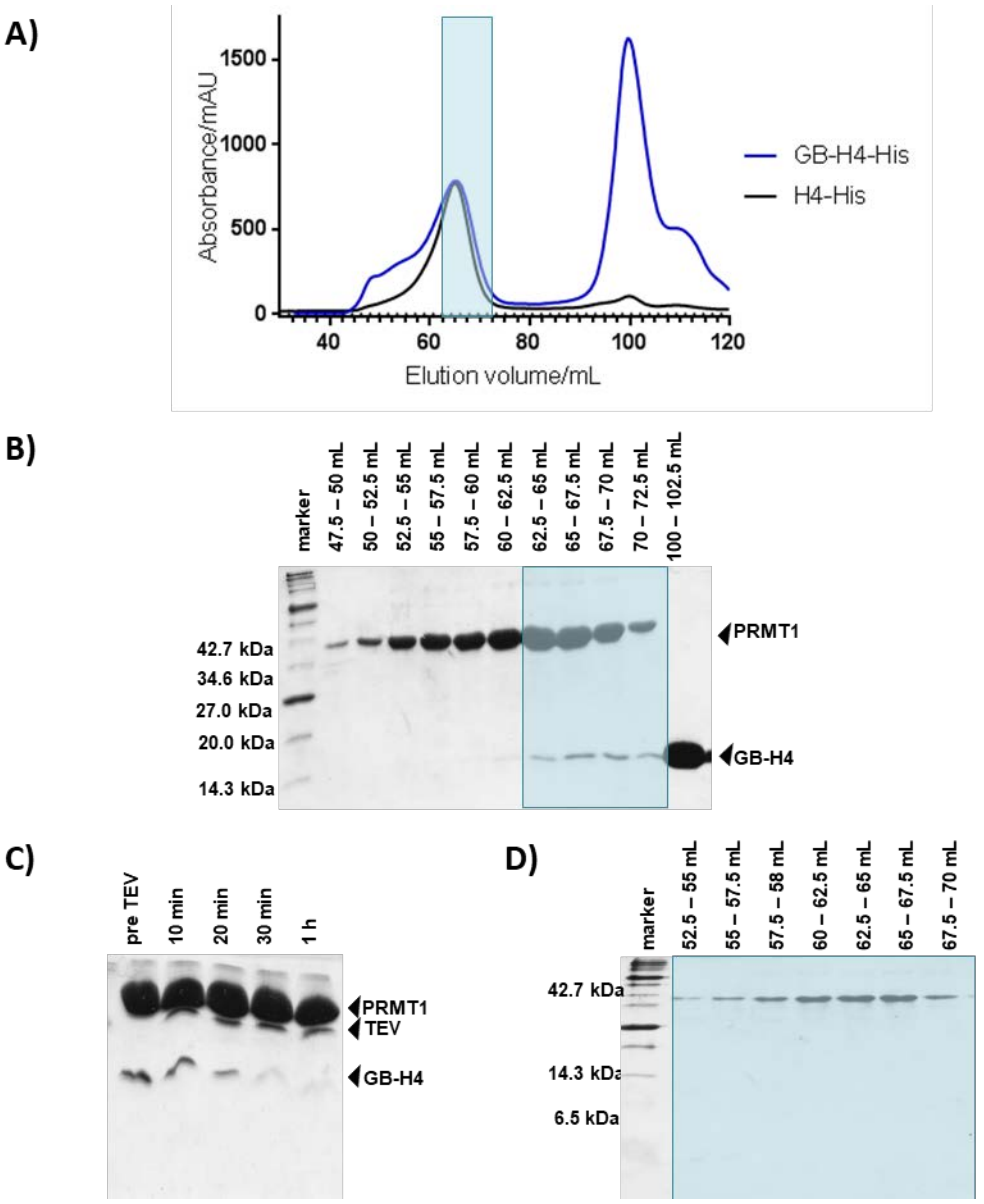


Figure 29: Purification of MBP-CARM1-His. (A) Nickel column chromatogram showing UV trace (blue) and percentage buffer B (red). Fractions from 20 – 37.5 mL (turquoise box) were combined for gel filtration; (B) Gel filtration (Superdex 200 16/60) chromatogram. MBP-CARM1 (81.0 kDa) elutes at a volume corresponding to a molecular weight of ~160 kDa, which is consistent with an MBP-CARM1 dimer. Fractions from 55-77.5 mL (turquoise box) were combined for subsequent TEV cleavage; (C) SDS-PAGE of expression samples; (D) SDS-PAGE of nickel chromatography fractions; (E) SDS-PAGE of gel filtration fractions.

TEV protease was used to remove the MBP tag. A small-scale optimisation was performed to determine how much TEV protease was needed to cleave the fusion protein (Figure 30A, below). While 0.05 molar equivalents were sufficient to achieve cleavage within 3 hours, 0.1 - 0.2 equivalents were typically used for purifications of this construct, since TEV protease (~28 kDa) is readily separable from CARM1-His by gel filtration, and it is desirable to minimise the total purification time.

Once cleaved, as judged by SDS-PAGE, a second gel filtration was used to separate cleaved CARM1 from TEV protease and the MBP tag. The similar molecular weights of CARM1 and MBP (40.0 and 41.0 kDa respectively) made it difficult to distinguish between them using SDS-PAGE. CARM1, like other PRMTs, is known to form oligomers, whereas MBP is a stable monomer.[98] This differing oligomerisation behaviour implied that the first peak in gel filtration corresponded to CARM1. This assignment was confirmed using two Western Blots (Figure 30C, below). Typically, 0.1 – 0.15 mg of CARM1-His was obtained per litre of bacterial culture, using this purification procedure.

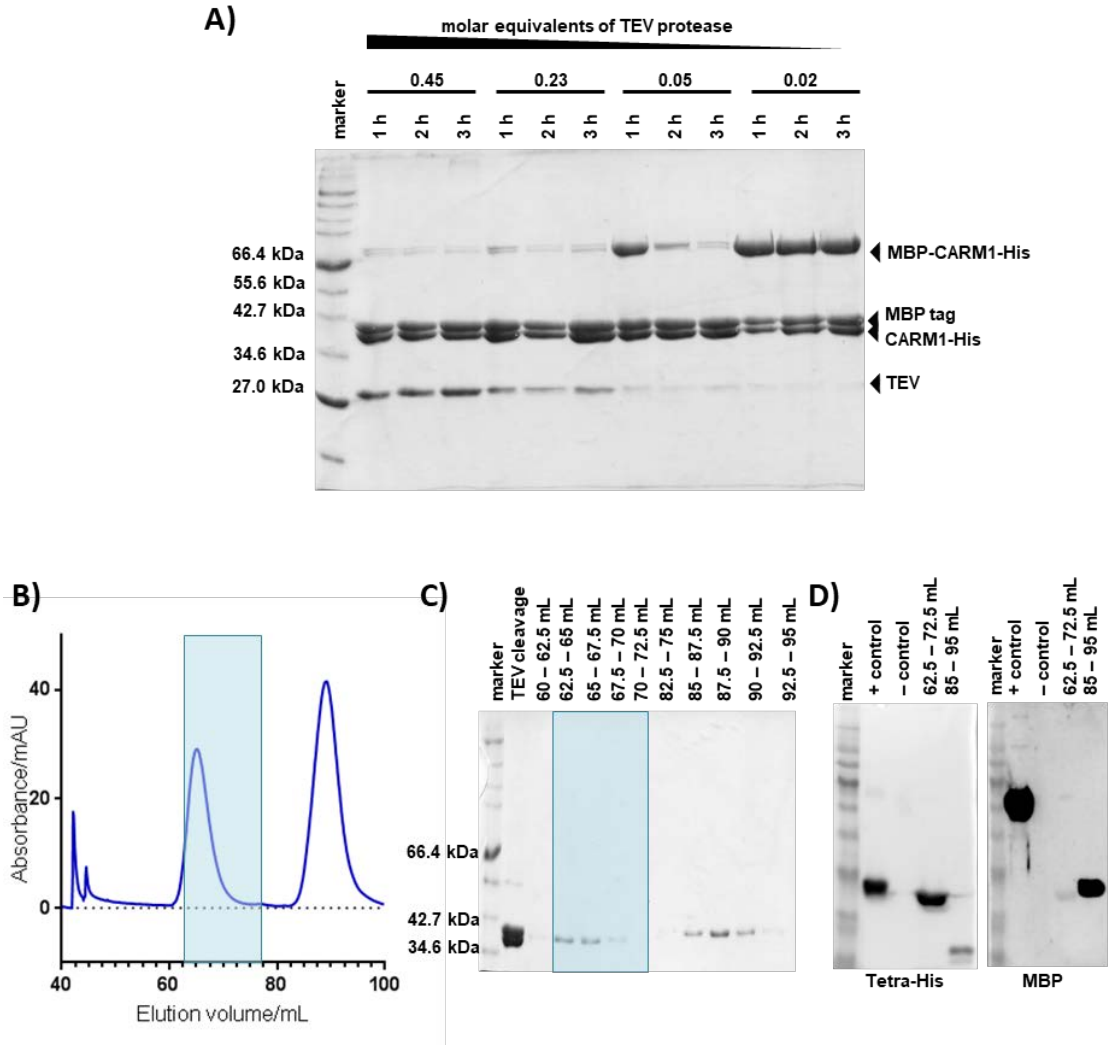


Figure 30: TEV cleavage and subsequent isolation of CARM1-His. (A) SDS-PAGE analysis of small scale reactions used to determine how much TEV is required to cleave the MBP tag from MBP-CARM1-His; (B) Gel filtration (Superdex 200 16/60) of cleavage mixture. CARM1 (MW = 40.0 kDa) elutes at 65 mL, corresponding to a molecular weight of ~160 kDa, which would correspond to a CARM1 tetramer. The MBP tag (MW = 41.0 kDa) elutes at ~90 mL, corresponding to an MBP monomer; (C) SDS-PAGE of fractions from gel filtration; (D) Anti tetra-his (left) and anti-MBP (right) Western blots confirm that the first gel filtration peak corresponds to CARM1. The small band visible at 85 – 95 mL on the anti tetra-his Western blot is presumably TEV protease, which is His-tagged. CARM1-His fraction combined for subsequent experiments are highlighted with a turquoise box on the gel filtration chromatogram and SDS-PAGE.

3.4.3. Removal of His-tag and Addition of C-terminal Residues

A construct containing the same residues as MBP-CARM1-His but lacking the C-terminal His Tag (MBP-CARM1, Figure 19, page 83) was cloned and expressed as for the His-tagged construct. This construct expressed well, but the solubility was reduced compared to the MBP-CARM1-His. This fusion protein precipitated when concentrated above ~1 mg/mL, and continued to precipitate during treatment with TEV protease. A construct with an additional 10 residues at the C-terminus (MBP-CARM1(135-489), Figure 19, page 83) was also cloned. This construct also expressed well but exhibited poor stability with respect to aggregation and precipitation.

An initial activity screen was performed, which indicated that CARM1-His had a higher specific activity than the untagged constructs (Figure 31, below). These results are summarized in Table 17, page 106. The activity data represent a one-off test of the activity of these constructs prepared as indicated, and are by no means proof that CARM1-His is inherently the most active. The high activity levels observed for the CARM1-only control samples are presumably a result of automethylation, which has been previously reported for CARM1.[99] Interestingly, the residue which CARM1 is known to automethylate (R550) is not present in any of the CARM1 constructs tested. This was not investigated further but identifying the residues targeted (e.g. by mass spectrometry of proteolysed samples) could shed light on this behaviour.

A summary of yields, specific activities and solubilities of the MBP-CARM1 constructs is given in Table 17, below. CARM1-His was chosen for use in subsequent assays and crystallisation trials, primarily due to its ease of purification and apparent improved stability when compared to the other constructs.

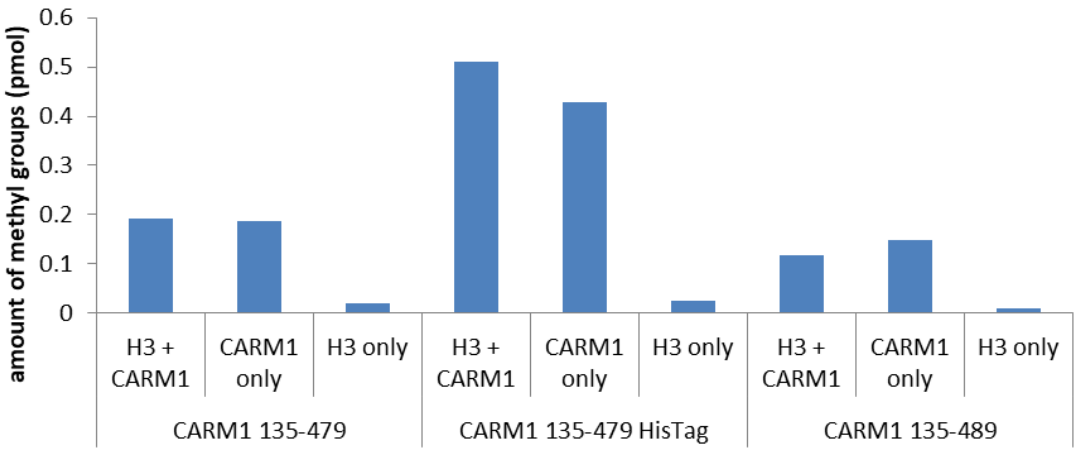


Figure 31: Activity screen of CARM1 constructs. The activity observed for the CARM1-only control is presumably due to automethylation.

Construct	Affinity Step used	Yield prior to TEV cleavage/L bacterial culture (mg)	Stability during purification	Yield/L bacterial culture (mg)	Approximate specific activity without tag
MBP-CARM1	MBP	5.2 (2)[1]	+	0.31 (2)	++
MBP-CARM1-His	nickel	0.37 (4)[2]	++++	0.12 (4)	++++
MBP-CARM1(135-489)	MBP	2.2 - 0.25 (2)	+	0.15 (2)	+

[1]Numbers in brackets indicate number of independent expressions and purifications this number is based on.
[2]Actual yield ~0.7 mg, but ~0.35 mg elutes without contaminants.

Table 17: Purification and activity screening of MBP-CARM1 fusion constructs. Average yields are used where yields were consistent, and ranges where yields showed large variation. The high pre-cleavage yields of the non-His-tagged constructs reflect the presence of misfolded protein (detected during gel filtration following TEV cleavage).

3.4.4. Attempts at Improving CARM1-His Yield

The yield of CARM1-His typically obtained using the described protocol was prohibitively low. A single ITC titration of a 10 μM (K_d) inhibitor into CARM1-His requires at least 1 mg of CARM1 to obtain a curve with sufficient sigmoidicity for accurate K_d determination.[100] We therefore sought to optimise the expression and purification protocol used to obtain CARM1-His. Performing a high salt wash during nickel chromatography (to try and remove co-eluting impurities, thereby increasing the number of nickel fractions that could be combined) was found to have little effect on the isolated yield (data not shown).

An alternative construct was designed, in which the MBP tag was replaced with a solubility tag designed by Dr Simon Caulton.[101] The tag is a receiver domain from *Myxococcus xanthus* social motility protein FrzS (PDB code 2GKG).[101-102] Cloning of this construct was performed by Alaa Al-Noori.[80] 2GKG-CARM1-His has an N-terminal His tag, but gives the same CARM1-His construct upon treatment with TEV protease. The 2GKG fusion protein eluted from the nickel column at a higher imidazole concentration (Figure 32, page 109) compared to the corresponding MBP fusion protein. This reduced contamination by co-eluting proteins, and meant that all 2GKG-CARM1-containing fractions could be used for subsequent purification steps (Figure 32, page 109). An additional improvement to the purification was made by diluting combined nickel fractions with gel filtration buffer prior to concentrating and treating with TEV protease (as opposed to buffer exchanging using gel filtration). This minimised loss due to sample handling required for gel filtration. This improved protocol

typically yielded ~0.25 mg CARM1-His per litre of bacterial culture. This improvement reflects the increased proportion of contaminant-free nickel fractions carried through to subsequent purification steps, rather than an increase in bacterial expression levels or solubility. This purification procedure is significantly quicker compared to that described for MBP-CARM1 since SDS-PAGE analysis of nickel fractions is not required, and only one gel filtration is performed.

Further attempts at improving the yield of CARM1 included the use of ArcticExpress cells (Agilent Technologies, Figure 33, page 110), using baffled flasks, and enriched media such as 2YT and Power broth™ (Molecular Dimensions). None of these changes significantly improved the isolated yield of CARM1. A slight improvement in the yield of 2GKG-CARM1 was made by reducing the IPTG concentration from 0.4 to 0.05 mM (Figure 34, page 110). This lower IPTG concentration was therefore used for subsequent 2GKG-CARM1 expressions.

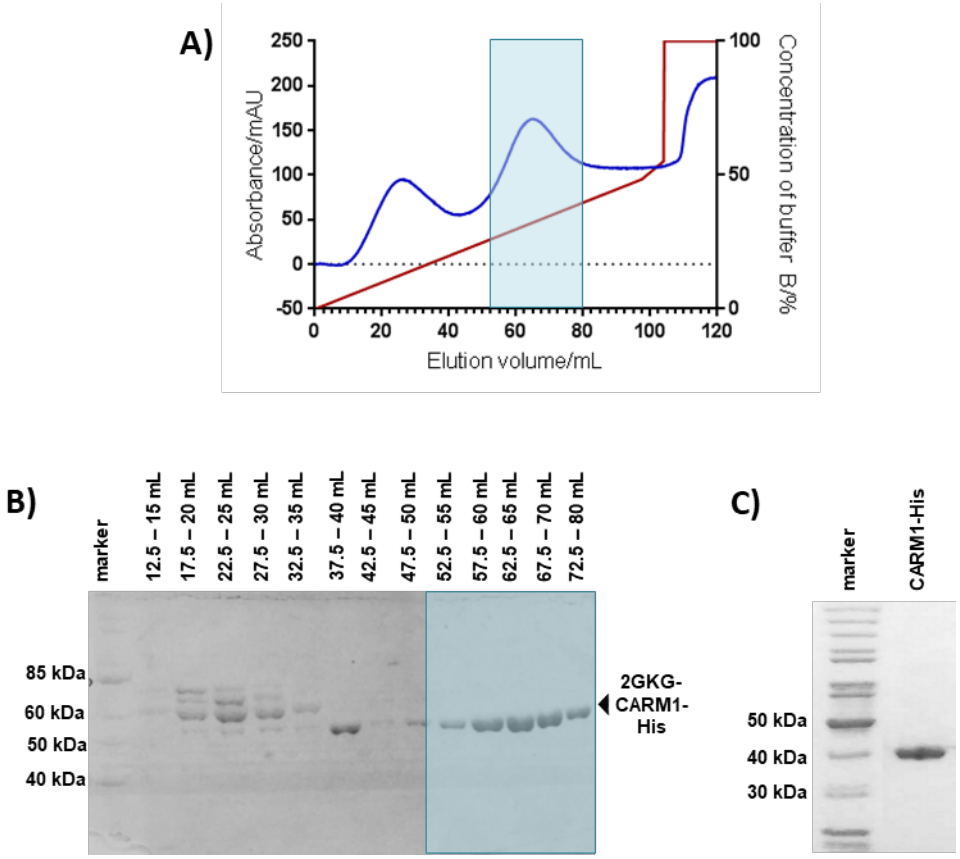


Figure 32: Nickel affinity purification of 2GKG-CARM1. (A) Nickel column chromatogram showing UV trace (blue) and percentage buffer B (red); (B) SDS-PAGE analysis of nickel fractions. All 2GKG-CARM1-containing fractions are pure as judged by SDS-PAGE. Fractions from 52.5 – 80 mL (turquoise box) were combined for treatment with TEV protease and subsequent gel filtration. (C) SDS-PAGE of CARM1-His after TEV cleavage and gel filtration.

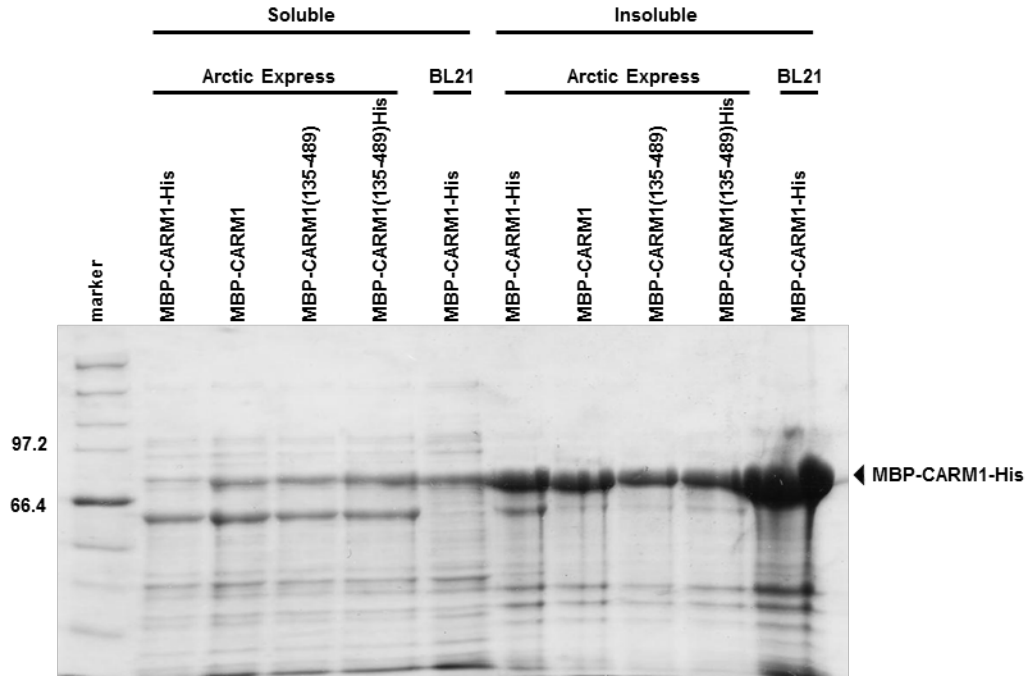


Figure 33: Test expression of MBP-tagged CARM1 constructs in ArcticExpress (AX) and BL21 cells. Equal volumes of lysate were separated into the soluble and insoluble fractions by centrifugation prior to loading (equal volumes) on SDS-PAGE. No improvement in the yield of soluble MBP-CARM1 is seen by using AX cells over BL21.

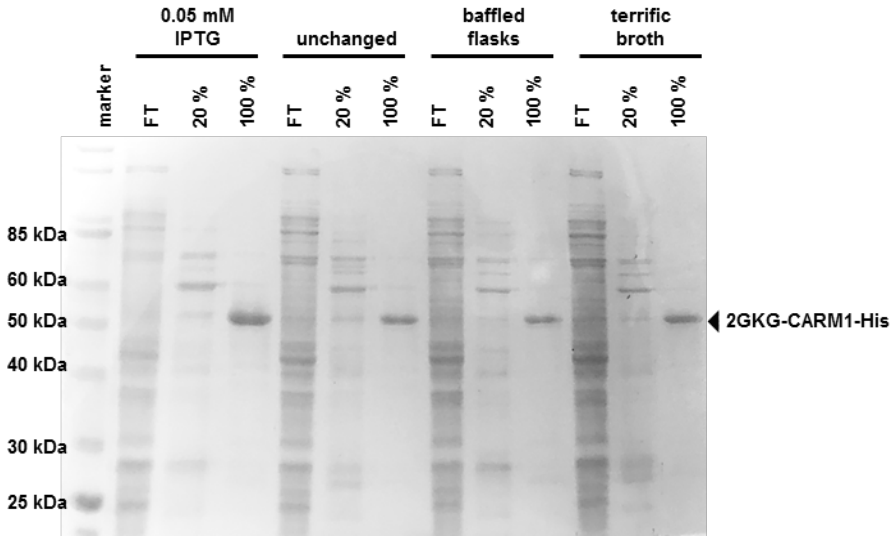


Figure 34: Optimisation of 2GKG-CARM1-His expression conditions using 1 L expressions, loading soluble protein onto a nickel column, and eluting with a step gradient of imidazole. Adjustments made to standard expression protocol (0.4 mM IPTG, 20 °C, non-baffled flasks, LB media, 16 h) are indicated. A slight increase in 2GKG-CARM1-His yield was seen on reducing the IPTG concentration.

3.4.5. Co-purification of CARM1 and GB-H3

Prior to cloning the 2GKG-CARM1 fusion construct, a co-purification of CARM1 with GB-fused histone 3 (GB-H3(21-44), Figure 21, page 86) was performed to try and increase the yield of CARM1-His (see section 2.3.8 for details). A band between the 14.3 and 6.5 kDa bands of the marker on SDS-PAGE implied that GB-H3 co-eluted with MBP-CARM1-His from the nickel column (Figure 35A, page 112). Since GB-H3 does not contain a His-tag this suggests that it had bound to CARM1 on the nickel column. Fractions eluting before 62.5 mL did not appear to contain GB-H3. Since the primary aim of this experiment was to determine whether the yield of CARM1-His could be increased by co-purification with GB-H3 (as opposed to isolating a CARM1-H3 complex), all MBP-CARM1-containing fractions were combined for subsequent gel filtration. The gel filtration elution profile was unchanged by the presence of GB-H3 (data not shown). Following gel filtration, the putative GB-H3 band was no longer visible on SDS-PAGE, suggesting that any bound GB-H3 had dissociated during the preceding spin-concentration and gel filtration (Figure 35B, page 112). It should be noted that small proteins and peptides typically stain poorly using Coomassie, therefore in future experiments it would be advisable to use silver stain to visualise relevant SDS-PAGE gels. The pure MBP-CARM1 was treated with TEV protease and separated by gel filtration. The isolated yield was 0.17 mg/L of culture, which represents no improvement over previously discussed purification methods. Despite this low yield, crystallisation trays were set up using CARM1-His (Section 3.5, page 123).

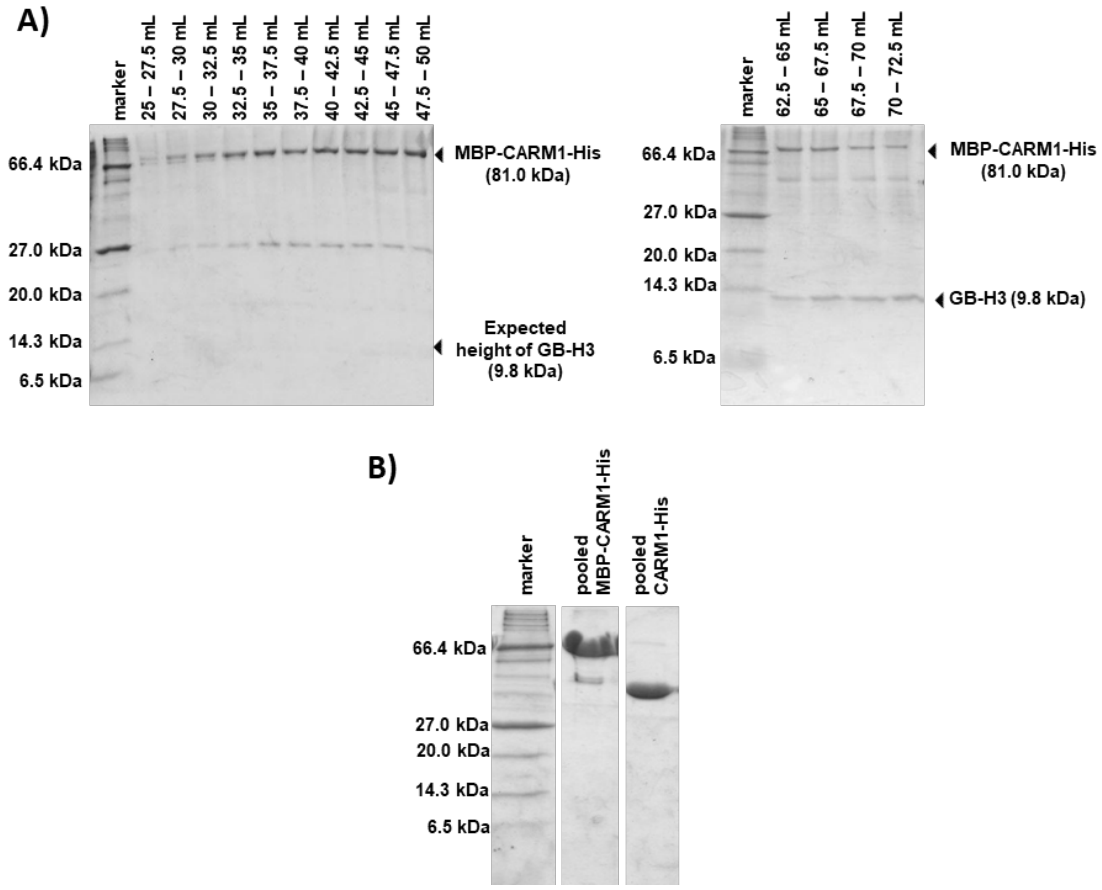


Figure 35: SDS-PAGE analysis of samples from co-purification of MBP-CARM1-His and GB-H3. (A) Fractions eluting from nickel affinity chromatography. GB-H3 was detected only in fractions eluting between 32.5 and 72.5 mL. SDS-PAGE of fractions 50 – 62.5 is not shown, but these fractions contained the same components as fractions from 25 to 50 mL). Fractions from 25 – 72.5 mL were combined for gel filtration; (B) Overloaded SDS-PAGE of pooled MBP-CARM1 fractions eluting from gel filtration. No band is visible between 6.5 and 14.3 kDa indicating that GB-H3 is not present in the sample.

3.4.6. Attempted Expression and Purification of Full Length CARM1 Constructs

As mentioned earlier (Section 1.3.1, page 14), the unique C-terminal domain of CARM1 (Figure 4, page 5) may influence inhibitor binding. To explore this possibility we sought to express and purify full length CARM1 constructs for use in inhibitor binding studies. Larger proteins are typically more challenging to express in *E. coli* due to greater potential for misfolding. Low temperature protein expression slows translation and can aid in the proper folding of recombinant proteins. Arctic Express cells, which express additional chaperonins and allow low temperature protein expression, were used to express GST-mCARM1(FL) (Figure 19, page 83). The pCOLD expression vector (TAKARA),[103] designed for low temperature protein expression, was also used for GST-mCARM1(FL) and TF-mCARM1(FL).

3.4.6.1. GST-CARM1(FL)

Initially, GST-mCARM1(FL) was chosen since it has been used for biochemical assays in several published studies.[16b, 69] A plasmid encoding the full length mouse sequence with an N-terminal GST tag in the pGEX-1λT vector was provided by Dr Clare Davies. The mouse sequence differs from the human sequence by only 2 residues (see Appendix 6.2 for sequence alignment). Expression was attempted using Arctic Express cells (transformed with the GST-mCARM1(FL) pGEX-1λT plasmid), but SDS-PAGE analysis showed no obvious differences in protein levels before and 16 hours after induction (Figure

36A, page 114). GST affinity chromatography was performed regardless, to see whether any GST-mCARM1(FL) could be detected. A Western blot was performed to analyse the content of fractions eluting from the GST column (Figure 36A, page 114). A low molecular weight band (presumably corresponding to native *E. coli* GST) was detected, but there was no clear band to indicate the presence of the CARM1 construct. This suggests that the construct was not present in the clarified lysate.

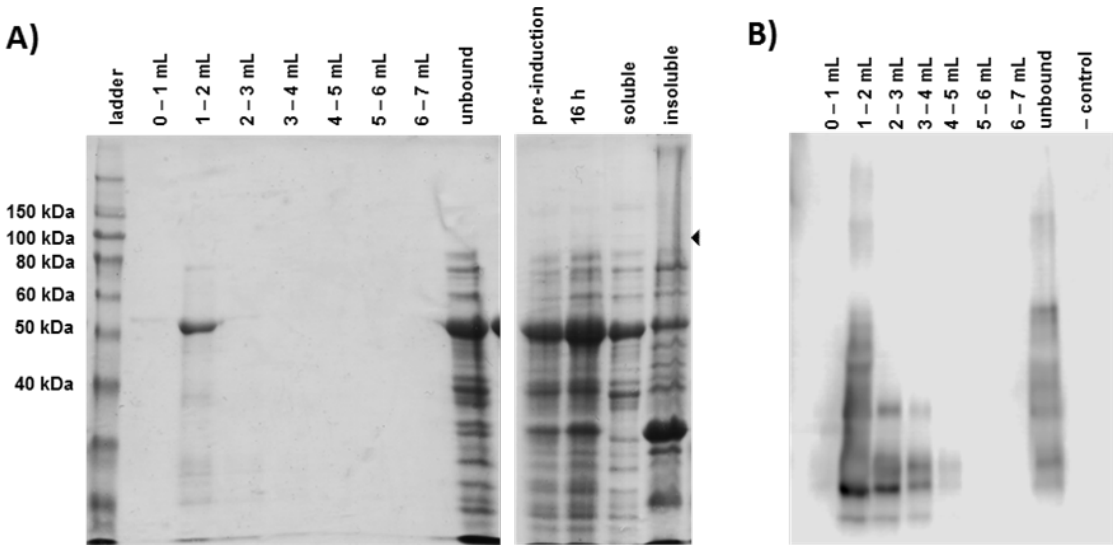


Figure 36: Purification of GST-mCARM1(FL) (MW = 92.7 kDa) (A) SDS-PAGE of expression samples and fractions eluting from GST column. An arrow is used to indicate the anticipated height of GST-mCARM1(FL); (B) Western blot of fractions eluting from GST column.

The construct was subcloned into the GST-pCOLD vector and expressed in BL21 cells (Figure 37, page 115). A protein with the approximate molecular weight of GST-mCARM1(FL) (92.7 kDa) was present in fractions eluting from the GST affinity column (Figure 37A, page 115). This protein eluted in the void volume

during subsequent gel filtration, and displayed no activity when tested using the radiometric activity assay (Section 2.8.1, page 50).

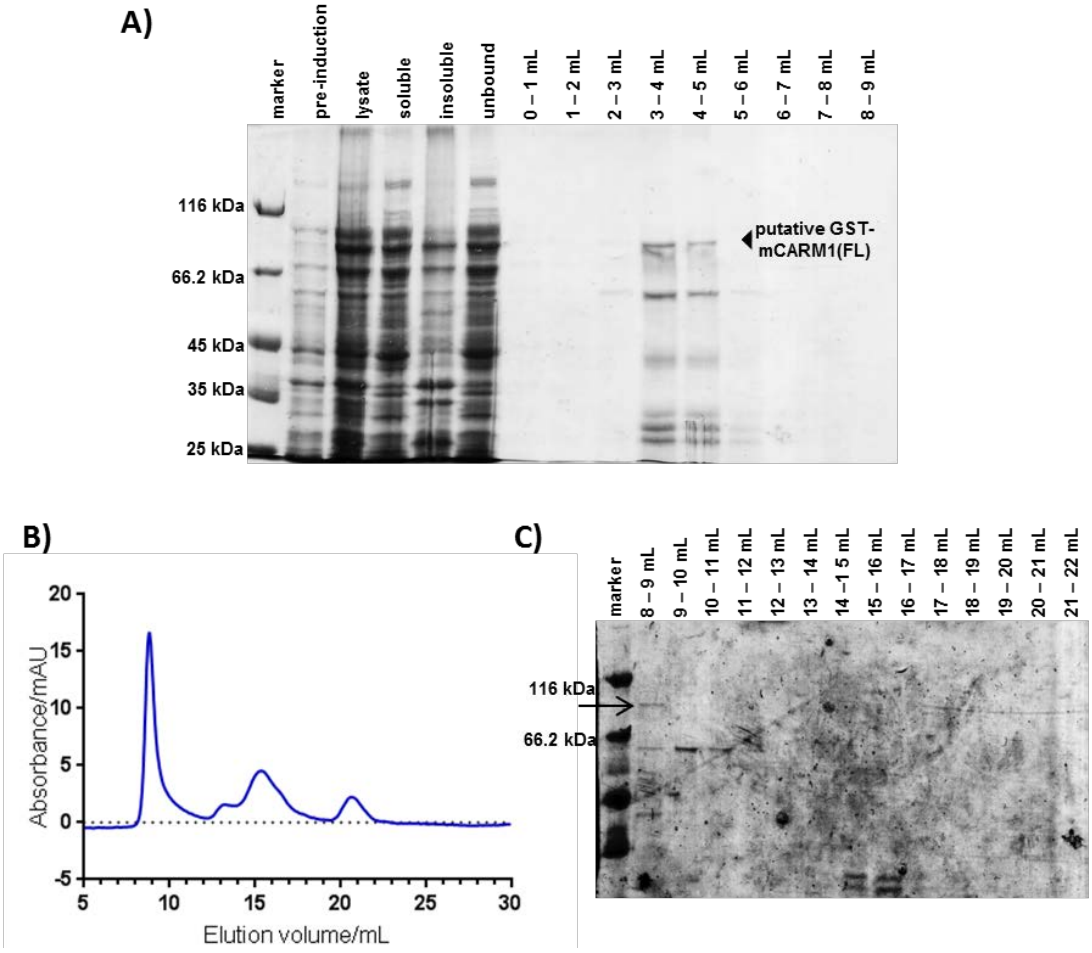


Figure 37: Attempted purification of GST-mCARM1(FL) expressed in BL21 cells using the pCOLD vector. (A) SDS-PAGE analysis of expression and GST-affinity fractions. Fractions eluting between 3 and 5 mL were combined and concentrated for gel filtration; (B) Gel filtration (Superdex 200 10/300) chromatogram; (C) SDS-PAGE analyis of gel filtration fractions. The putative GST-mCARM1(FL) band elutes in the void volume (~9 mL).

To explore whether adjusting the expression conditions might improve the expression levels of GST-mCARM1(FL), test expressions and purifications were performed (Figure 38, page 116). Additionally, the composition of the lysis and

elution buffers were changed, based on those reported by Shishkova *et al.,* who report to have expressed full length, active CARM1 in *E. coli* (see Section 2.2.4, page 38 for details).[36]

Condition #	vector	media	Induction temp (°C)	Expression time (h)	IPTG concentration (mM)
1	pGEX-1λT	LB	25	4	0.2
2	pGEX-1λT	LB	25	4	0.04
3	pGEX-1λT	LB	15	16	0.2
4	pGEX-1λT	LB	15	16	0.04
5	pGEX-1λT	Power broth™ (Molecular Dimensions)	25	4	0.2
6	pCOLD	LB	15	30	0.04

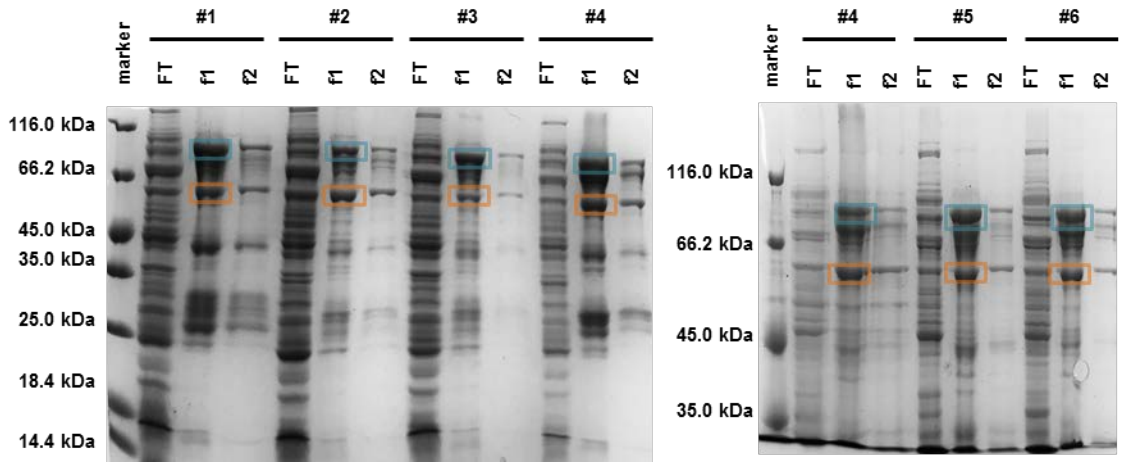


Figure 38: Optimisation of GST-mCARM1(FL) expression conditions using 1 L expressions, loading soluble protein onto a GST column, and eluting with a step gradient of elution buffer. The conditions were numbered and are given in the table above the SDS-PAGE gels. A band at the correct molecular weight for GST-mCARM1(FL) (92.7 kDa) is visible in all samples and is highlighted with blue boxes. Orange boxes are used to highlight a band at ~60 kDa, which may correspond to a break-down product.

Promisingly, all samples eluting from GST affinity chromatography contained a band at the correct molecular weight for the full length fusion protein. The

presence of a large number of lower molecular weight contaminants meant that these samples are not of sufficient purity for use in assays. To determine whether the putative GST-mCARM1(FL) was separable from these contaminants, analytical gel filtrations were performed with the eluate from conditions 3, 5 and 6 (Figure 39, page 118). These conditions were chosen since they appeared to contain the highest proportion of putative GST-mCARM1(FL).

Disappointingly, the major species eluting from gel filtration had a molecular weight of approximately 60 kDa, and very little protein with molecular weight corresponding to GST-mCARM1(FL) was observed by SDS-PAGE (Figure 39B, page 118). No precipitation was observed between GST affinity and gel filtration chromatography steps. It is therefore plausible that the 60 kDa protein eluting from gel filtration is a break-down product of GST-mCARM1(FL) (resulting from proteolytic cleavage by trace amounts of proteases).

Thus we have been unable to obtain a sample of active, GST-fused, full length CARM1.

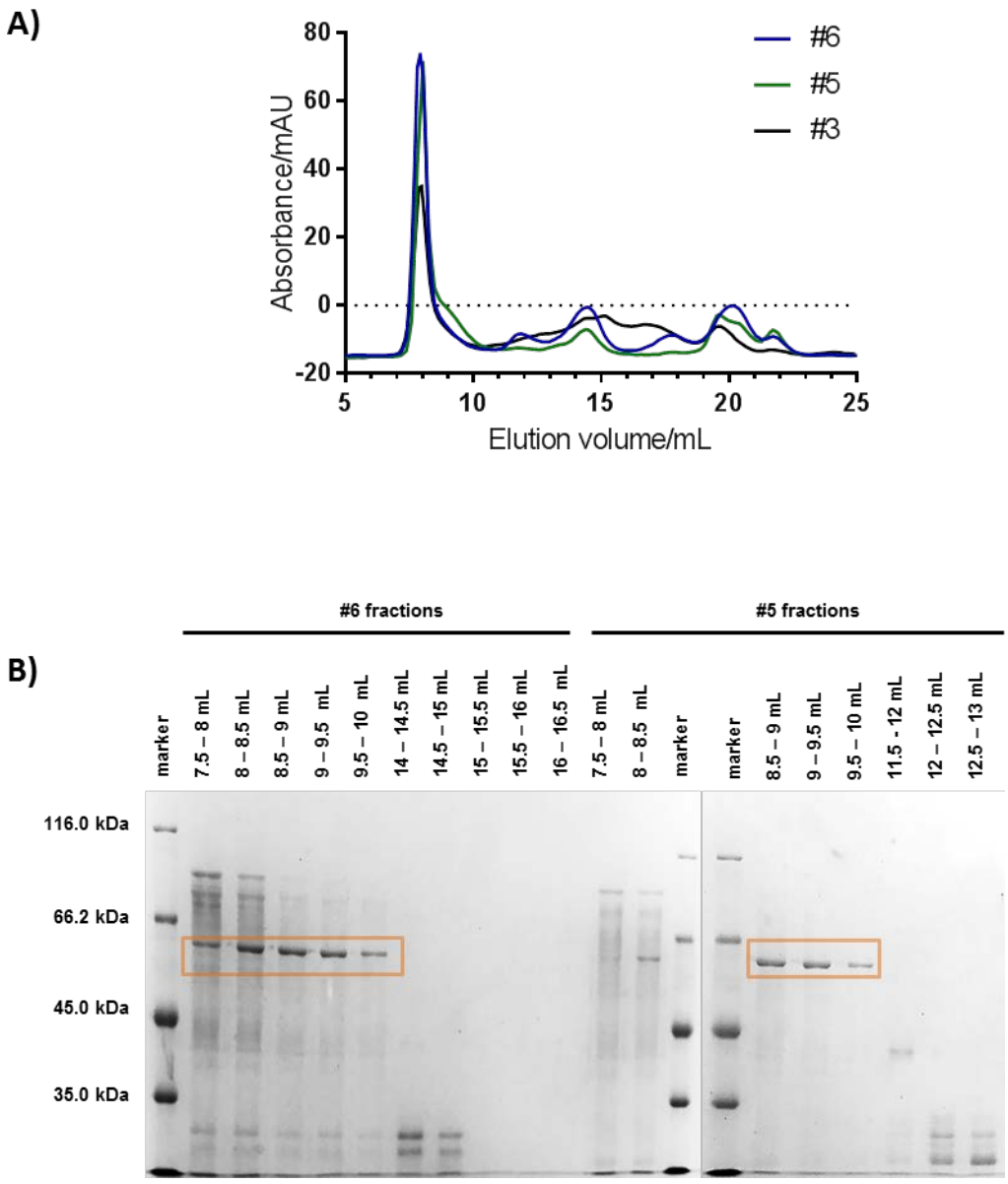


Figure 39: Analytical gel filtrations (Superdex 200 10/300) of eluates from GST-affinity column from Figure 37 (above). (A) UV chromatogram; (B) SDS-PAGE of eluting fractions. The major species eluting from gel filtration has a molecular weight of ~60 kDa. Orange boxes are used to highlight a band at ~60 kDa, which may correspond to a break-down product (see Figure 38).

3.4.6.2. TF-mCARM1(FL)

Having failed to obtain an active sample of full length CARM1 as a GST fusion, we next sought to express and purify CARM1 as a trigger factor fusion. Trigger factor (TF) is a 48 kDa prokaryotic chaperone protein, which promotes the correct folding of polypeptides as they are produced at the ribosome.[104] Trigger factor expresses well in *E. coli* and has been shown to increase total yields of correctly folded protein when used as a fusion tag. The full length CARM1 sequence was therefore sub-cloned into the TF-pCOLD vector. TF-mCARM1(FL) expressed well, with 32 mg obtained after nickel affinity chromatography and gel filtration from 5 L of culture (Figure 40, below).

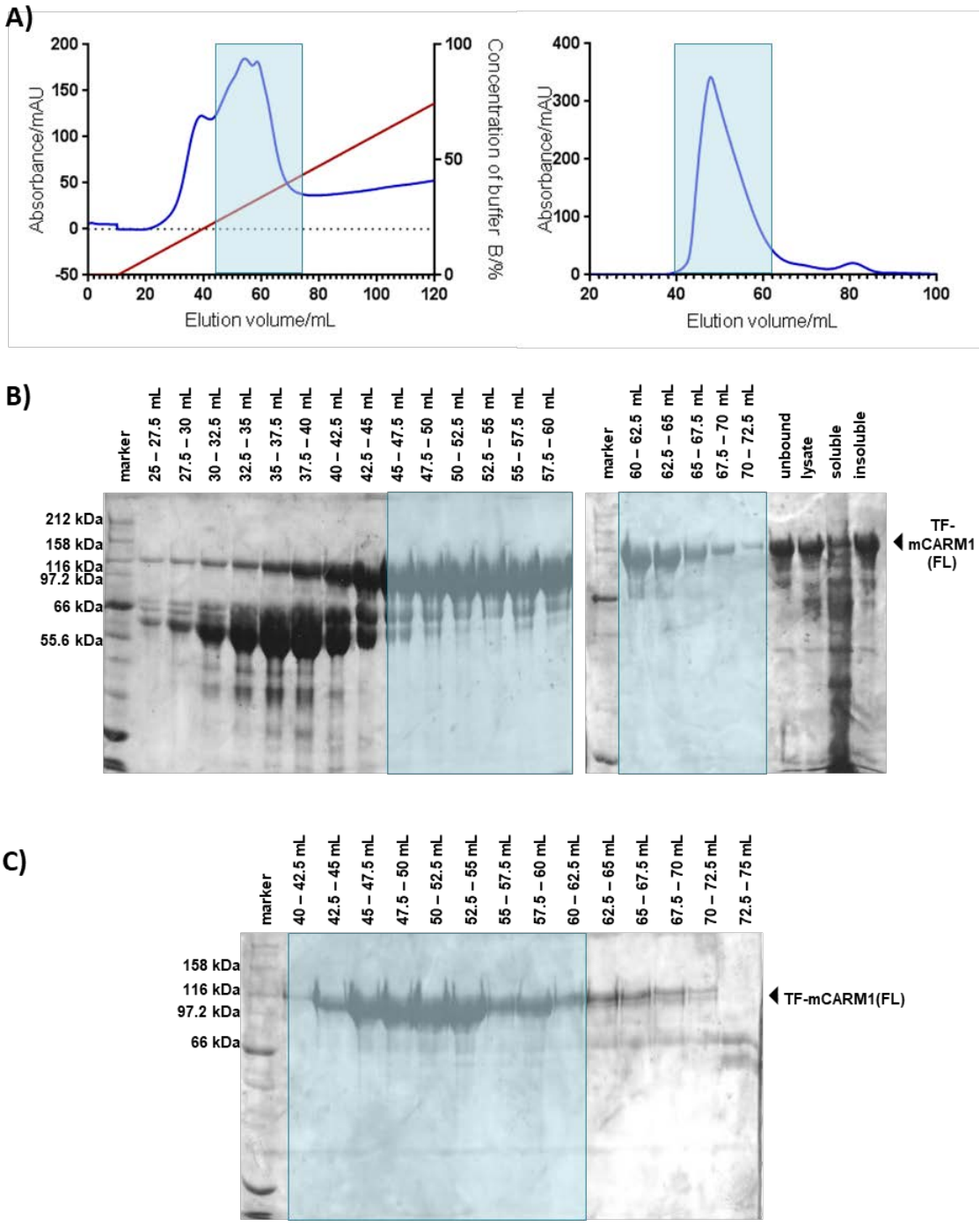


Figure 40: Purification of TF-mCARM1(FL) (115 kDa). (A) Nickel (left) and gel filtration (Superdex 200 16/60, right) chromatograms showing UV trace (blue) and percentage buffer B (red); (B) SDS-PAGE analysis of nickel fractions. Fractions from 45 – 72.5 mL (turquoise box) were combined for subsequent gel filtration; (C) SDS-PAGE of gel filtration fractions. Fractions from 40 – 62.5 mL (turquoise box) were combined for subsequent steps.

Attempted removal of the TF tag with HRV-C3 protease resulted in incomplete cleavage of the tag after 16 hours at 4 °C (Figure 41A, page 122). Increasing the temperature to r.t. for an additional hour resulted in partial precipitation of mCARM1(FL), indicated by the relative intensities of mCARM1(FL) and TF on SDS-PAGE (Figure 41A). Attempts were made to separate mCARM1(FL) from the uncleaved fusion TF-mCARM1(FL) and the TF tag (Figure 41B). Unfortunately, neither nickel affinity chromatography nor gel filtration were able to separate mCARM1(FL) from the other components of the digestion reaction, and all proteins eluted in the void volume during gel filtration. While preferable to perform assays using untagged full length proteins, tagged variants may still show activity and be useful for deriving biological information. An aliquot of TF-mCARM1(FL) was therefore analysed using the methylation assay (See Section 2.8.1 for details). Unfortunately, no activity was detected for the TF-mCARM1 fusion protein. This may be due to the TF tag obstructing the CARM1 binding site, or misfolding or aggregation of CARM1.

Further work is required to develop a reliable means of obtaining catalytically active, full length CARM1. For the remainder of this work, we continued with the catalytic core of CARM1 (CARM1-His, Figure 19, page 83).

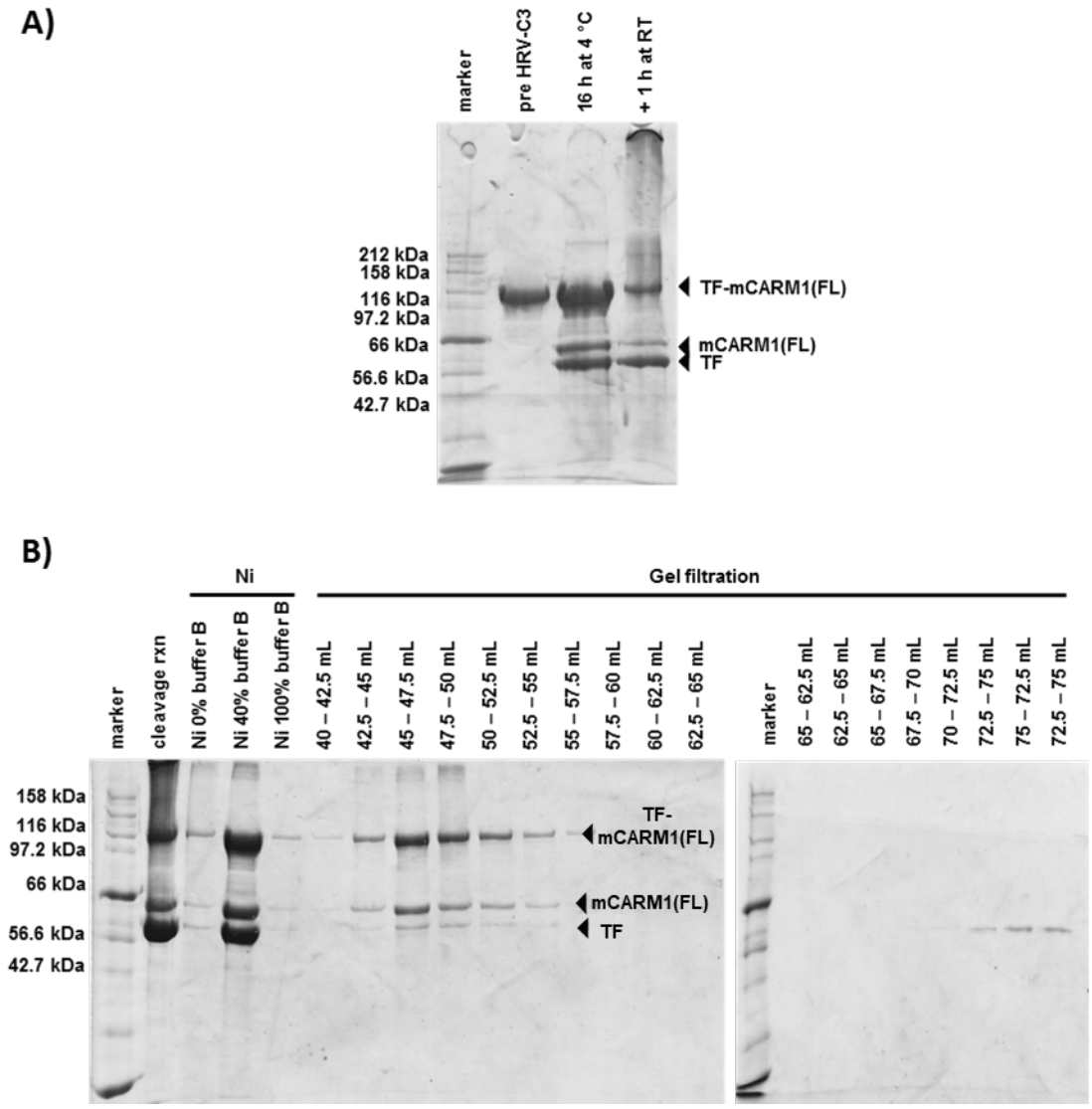


Figure 41: Cleavage of TF-mCARM1(FL) and subsequent attempts at purification. (A) SDS-PAGE of cleavage reaction at various time points. The reaction contained 0.32 mg HRV-C3 and 32 mg TF-mCARM1(FL) (i.e. 1:100 w/w) in a total volume of 10 mL; (B) SDS-PAGE of fractions collected during attempted separation of mCARM1(FL) (66 kDa) from TF-CARM1(FL) (115 kDa) and TF (49 kDa). The co-elution of the un-tagged TF with mCARM1(FL) during nickel and gel filtration chromatography indicate that it is sticking to TF-mCARM1(FL) or mCARM1(FL).

3.5. CARM1 Crystallisation

3.5.1. CARM1-SAH Co-crystallisation

Having explored CARM1 purification conditions and determined a route to a soluble, catalytically active CARM1 catalytic core construct (CARM1-His, Figure 19, page 83), attention turned to the crystallisation of CARM1.

Typically, protein concentrations of 2 – 50 mg/mL are used for protein crystallisation.[105] A single 96-well screening plate with 200 nL protein per condition therefore requires 0.04 – 1 mg of purified protein. Due to limited availability of CARM1-His (owing to its low expression levels), the first crystallisation screen was performed with CARM1-His at 0.8 mg/mL. This corresponds to a molar CARM1 concentration of 20 μM. The PACT screen was used, and drops were set up both with and without 100 μM SAH (this ligand concentration was based on the anticipated low micromolar affinity of SAH for CARM1 – see https://hamptonresearch.com/tip_detail.aspx?id=173). SAH was chosen as a model inhibitor to determine suitable conditions for generating co-crystals of CARM1-His and our putative bisubstrate inhibitors. SAH is commercially available, whereas the novel inhibitors we sought to co-crystallise require multi-step synthesis and are therefore more costly. We anticipated that the similar structure and anticipated binding mode of our inhibitors to SAH would make it an appropriate model ligand.

No crystal growth was observed in the absence of SAH (Figure 42, page 124) but, remarkably, crystals were observed in the SAH-containing drops within 5 days. Interestingly, precipitate was present in all drops set up without SAH. This may

indicate that SAH binding enhances the stability of CARM1 with respect to precipitation.

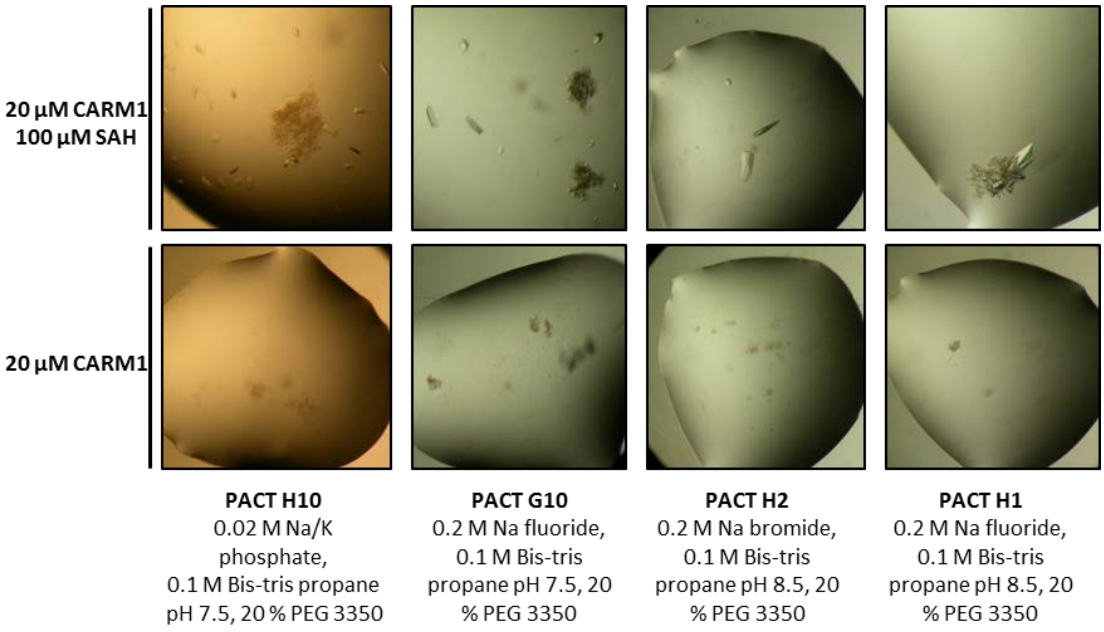


Figure 42: CARM1-His crystallisation trials using PACT premier™ screen (Molecular Dimensions) both with (upper) and without (lower) the addition of 100 µM SAH. Crystals with the most promising appearance are shown. The conditions are given underneath.

To determine how reproducible these CARM1-His crystals were, and identify additional hit conditions, a new batch of CARM1-His was purified and further crystallisation trials set up using the PACT and Morpheus screens (Figure 43, page 125). The same concentration of CARM1-His and SAH were used (20 µM and 100 µM, respectively). More than 20 PACT conditions gave crystals whereas a single condition from the Morpheus screen gave crystals. Since this was an isolated hit compared to the abundance of crystals formed in closely related

PACT conditions, subsequent optimisations were based solely on the most promising conditions from the PACT screen. All of these conditions contained 20 % (w/v) PEG 3350 and were buffered using bis-tris propane. There was a general trend of increasing crystal size at higher pH. The theoretical pI of CARM1-His (calculated using ExPASy ProtParam)[106] is 6.23. This trend may therefore reflect a reduction in solubility (and concomitant increase in seeding rate) of CARM1-His as the pH approaches its pI.

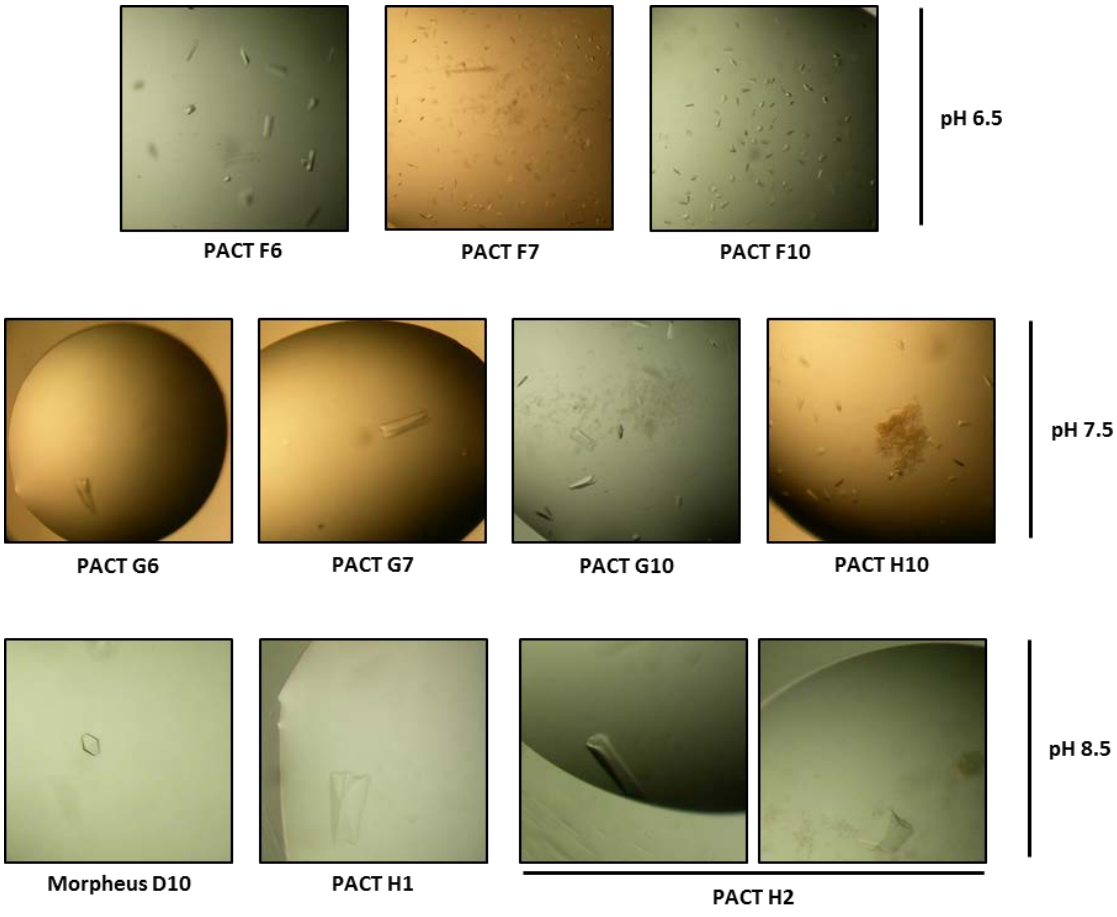


Figure 43: CARM1-His Crystals obtained using the PACT premier™ screen (Molecular Dimensions). The composition of each screen solution is given in Table 17, below. CARM1-His was at a concentration of 20 µM (0.8 mg/mL) and SAH was at 100 µM. There is an apparent trend towards fewer, smaller crystals at lower pH.

Screen	Additive	Buffer	pH	PEG
PACT F6	0.2 M Na formate	0.1 M Bis-tris propane	6.5	20 % PEG 3350
PACT F7	0.2 M Na acetate	0.1 M Bis-tris propane	6.5	20 % PEG 3350
PACT F10	0.02 M Na/K phosphate	0.1 M Bis-tris propane	6.5	20 % PEG 3350
PACT G6	0.2 M Na formate	0.1 M Bis-tris propane	7.5	20 % PEG 3350
PACT G7	0.2 M Na acetate	0.1 M Bis-tris propane	7.5	20 % PEG 3350
PACT G10	0.2 M Na fluoride	0.1 M Bis-tris propane	7.5	20 % PEG 3350
PACT H10	0.02 M Na/K phosphate	0.1 M Bis-tris propane	7.5	20 % PEG 3350
PACT H1	0.2 M Na fluoride	0.1 M Bis-tris propane	8.5	20 % PEG 3350
PACT H2	0.2 M Na bromide	0.1 M Bis-tris propane	8.5	20 % PEG 3350
Morpheus D10	0.12 M each 1,6-Hexanediol, 1-Butanol, 1,2-Propanediol, 2- Propanol, 1,4-Butanediol, 1,3-Propanediol	0.1 M each Tris (base), BICINE	8.5	20 % ethylene glycol, 10 % PEG 8000

Table 18: Composition of PACT premier™ (Molecular Dimensions) screen solutions which subsequent optimisation CARM1-His plates were based on.

Having confirmed the reproducibility of CARM1-SAH co-crystallisation under these conditions, we sought to optimise the crystallisation conditions. Optimisation is performed to increase the crystal size and ensure that crystals grow in each dimension. Needle or plate shaped crystals may result in lower resolution data along one or more of the unit cell axes, leading to difficulties in downstream analysis of the diffraction patterns. Smaller crystals are also more difficult to centre in the X-ray beam and may be more susceptible to radiation damage during data collection.

Optimisation plates were set up in which the pH and PEG concentration were varied (Figure 44, page 127). There was a general trend towards fewer crystals at higher pH and lower PEG concentration. There was no clear relationship between additive and crystal growth.

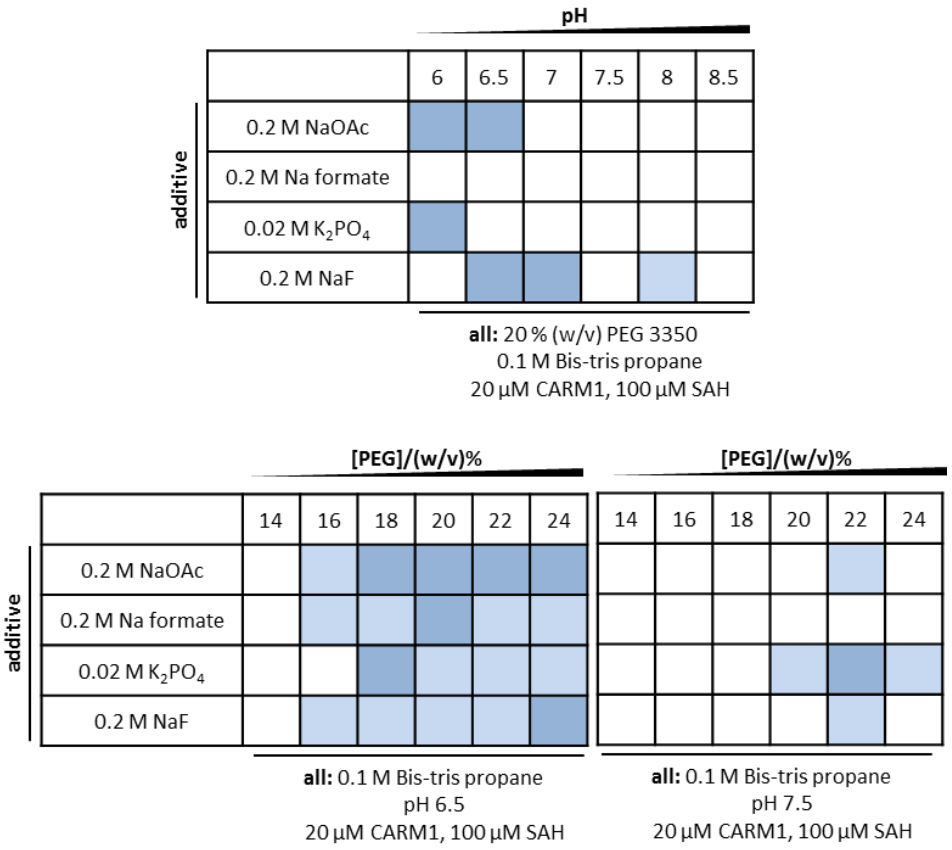


Figure 44: CARM1-SAH co-crystallisation optimisation plates. Light blue conditions gave crystals within 1 month, and dark blue squares gave crystals within 1 week.

3.5.2. CARM1-inhibitor Co-crystallisation

Having established conditions which reliably produce CARM1 crystals, we sought to generate CARM1 crystals in complex with one of our putative bisubstrate inhibitors. A 4-carbon linker amine-guanidine **UM249** (Figure 45A, page 128) was chosen as a starting point, since it was the compound we had the most of. A higher inhibitor concentration was used (167 µM as opposed to 100 µM used for SAH) due to the anticipated weaker binding of the inhibitor to CARM1 when compared to SAH. Trays were set up in which pH and PEG concentration were varied (Figure 45B, page 128) and crystals grew within 1 day

to 1 week. A selection of these crystals (Figure 45C, page 128) were fished, cryoprotected and their diffraction tested at Diamond Light Source.

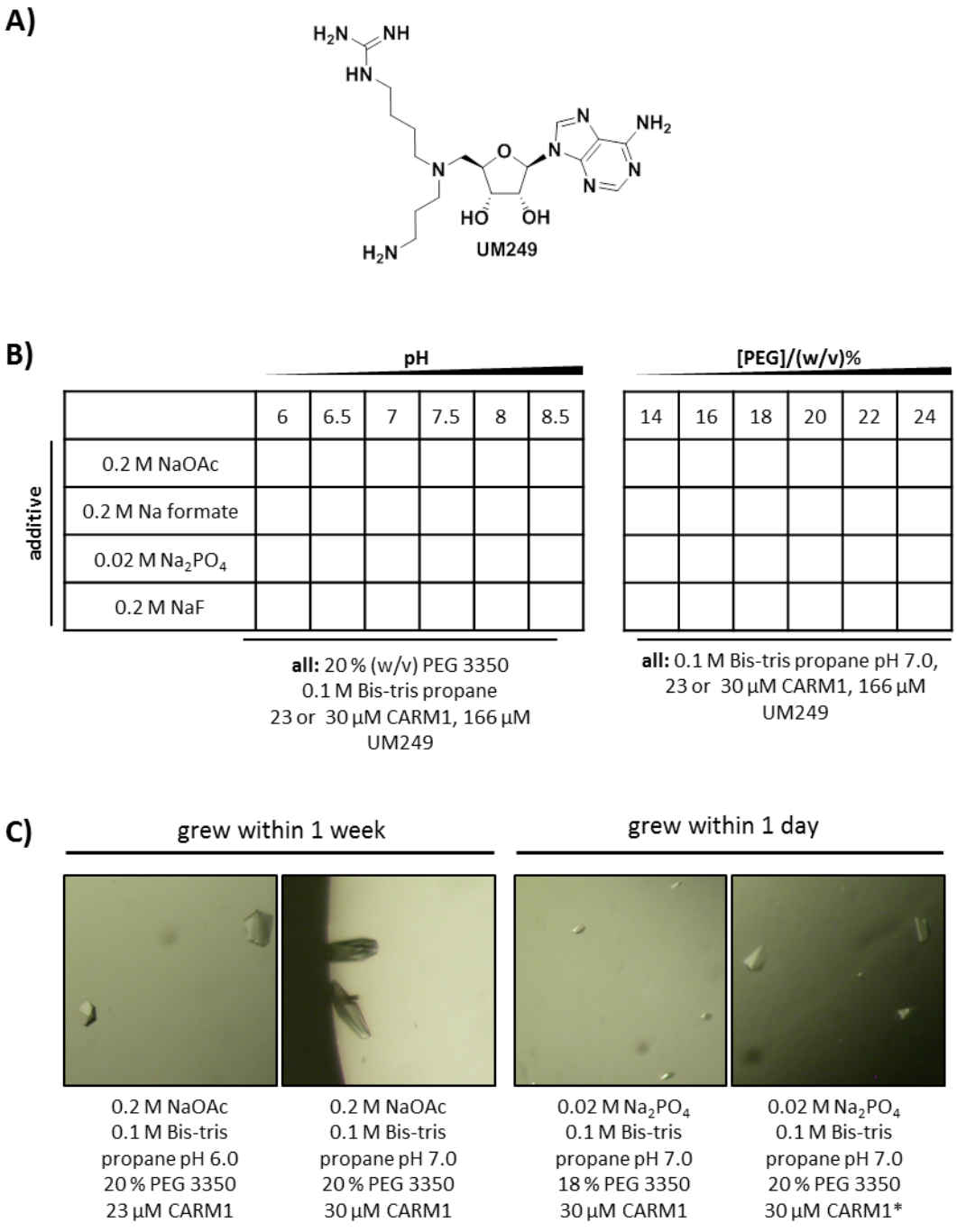


Figure 45: CARM1-His + UM249 co-crystallisation. (A) 24-well custom screens used to determine appropriate crystallisation conditions for the complex. Each plate shown was set up using both 23 and 30 µM CARM1. The concentration of UM249 was 167 µM; (B) Crystals which were taken to Diamond Light Source for X-Ray diffraction experiments.

***Crystals which gave the best diffraction data**

The crystal that gave the best diffraction data used CARM1 and **UM249** concentrations of 30 and 167 μM respectively, and a precipitant solution containing 0.02 M sodium phosphate, 20 % PEG 3350 and 0.1 M bis-tris propane at a pH of 7.0. This is in line with the crystallisation conditions used for other CARM1 structures available in the PDB (see Appendix 6.1). This dataset was solved by molecular replacement, using a published CARM1 structure (PDB code 5DX0, with ligands removed)[15a] as the search model. Crystals belonged to space group P 2_12_12 with two CARM1 dimers in the asymmetric unit, which is in line with CARM1 structures previously reported (see Appendix 6.1).[15a, 16b] Crystallographic data processing statistics are given in Appendix 6.4. Excitingly, examination of the active site revealed clear electron density for the inhibitor (Figure 46, page 130). This is the first evidence that these inhibitors occupy the CARM1 active site in a bisubstrate manner as intended.

Superposition of the determined structure with a published ternary structure of CARM1·H3·sinefungin revealed that the guanidinium group of the inhibitor extends into the substrate binding channel as intended. The high b-factors of the inhibitor guanidinium group (Appendix 6.3.1) relative to the adenosine moiety and surrounding amino acids may indicate a degree of disorder. This is likely due to inherent flexibility of this group within the active site. It is therefore inappropriate to interpret fine details of the modelled guanidinium position. Additionally, the modelled guanidinium position will be influenced by a weighted combination of the inhibitor restraints and the measured X-ray data. As hypothesized by Dowden *et al.*,[69] it appears that the four carbon linker

of **UM249** pushes the guanidinium group past key glutamates E257 and E266, which are proposed to hydrogen bond with the substrate arginine during catalysis (see Section 1.1). Therefore, while the butylguanidine group of **UM249** occupies the desired pocket within the substrate arginine binding site, it appears that a shorter linker may allow the inhibitor guanidinium to make the same interactions as the guanidinium substrate. This is consistent with previously reported IC_{50} values (Table 1, page 24), in which a 3-carbon linker amino acid-guanidine inhibitor inhibited CARM1 more strongly than its 4-carbon linker analogue.[69] Overall, the crystal structure shows that the conditions tested are suitable for obtaining CARM1-inhibitor co-crystals; **UM249** binds in the active site as intended and the propylguanidine and propylamine groups are flexible when bound.

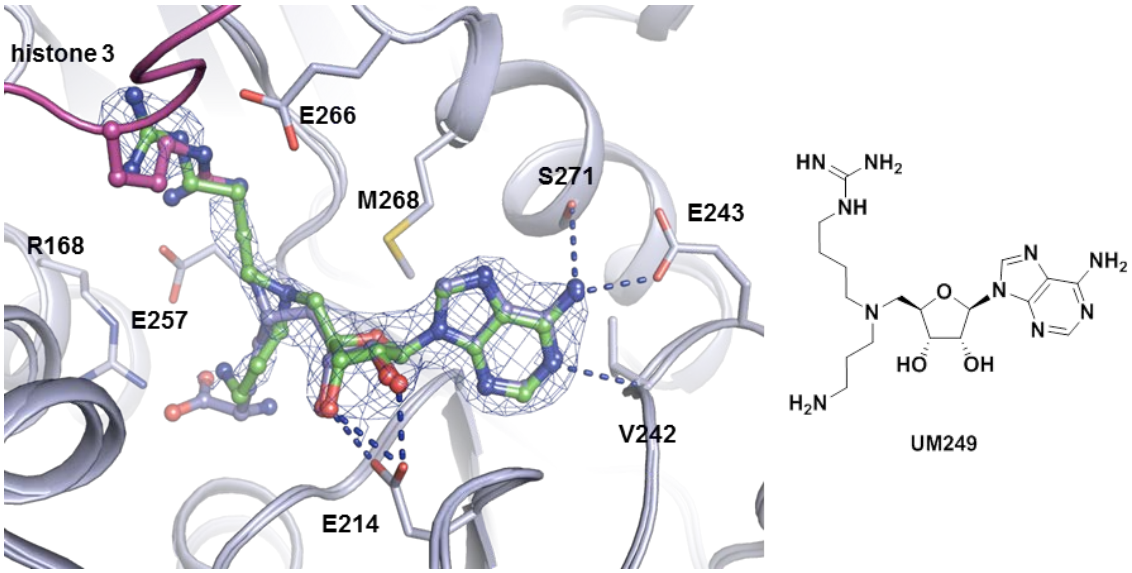


Figure 46: Crystal structure of UM249 (green) bound to CARM1 (white). $2mF_o$-F_c map (contoured at 1.0 σ, carve radius = 1.5 around the ligand) is shown in blue and a published crystal structure of CARM1 (white) bound to histone 3 (magenta) and SAM-mimic sinefungin (green) is superposed (PDB 5Dx0). Dashed lines indicate putative polar contacts between UM249 and CARM1. The N-terminal helix is hidden for clarity. The structure of UM249 is shown on the right.

3.6. Effect of Linker Length on Inhibitor Affinity

3.6.1. Isothermal Titration Calorimetry

Having established that the 4C amine-guanidine **UM249** binds in the CARM1 active site as intended, we next sought to examine how the length of the alkyl linker between the SAM and arginine mimicking moieties affects binding to PRMT1 and CARM1. To do this, isothermal titration calorimetry (ITC) was used.

To determine suitable conditions for inhibitor titrations, SAH was used as a model ligand (Figure 47, page 132). An initial titration was performed in which 200 μM SAH was titrated into 20 μM PRMT1. Binding was clearly detected, but the signal-to-noise levels of this titration were poor, and the determined stoichiometry (n = 0.661) was lower than the anticipated value (n = 1). An improved dataset was obtained by increasing the concentration of PRMT1 to 35 μM. A similar titration was performed into CARM1, and gave good signal-to-noise ratio and a sigmoidal curve. Determined K_d values are within the anticipated range, based on previously reported K_d and IC_{50} values.[74, 107] Interestingly, a stoichiometry of ~0.5 was determined for SAH with both PRMT1 and CARM1. This is unexpected given the observed 1:1 stoichiometry in published crystal structures and a previously reported ITC titration of SAM into CARM1.[16b] To explore the possibility that our calculated extinction coefficients for PRMT1 and CARM1 were inaccurate; a Bradford assay was performed (Appendix 6.4.1). The concentrations determined using the Bradford and UV methods were found to be in good agreement. A second possible cause of the observed low stoichiometry is that a fraction of the total protein is misfolded

and unable to bind to SAH. Since the single site binding model used for curve-fitting contains a variable stoichiometry term, the inactive protein fraction is accounted for and does not affect the determined K_d value. Determining whether the cause of the observed stoichiometry is due to the presence of misfolded protein or sub-stoichiometric ligand-binding to correctly folded PRMT1 and CARM1 is beyond the scope of this work, but would be interesting to explore further.

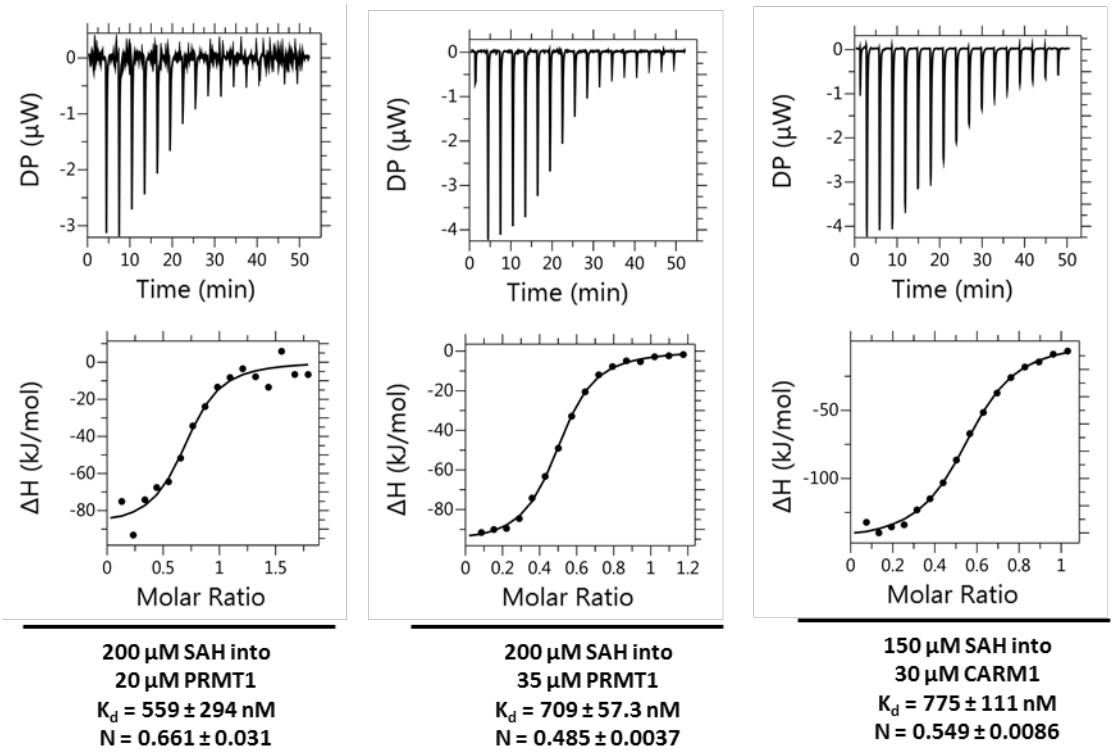


Figure 47: ITC titration of SAH into PRMT1 and CARM1. Thermograms are shown in the upper panels and integrated heat plots with fitted binding curves in the lower panels.

Having confirmed ITC as a suitable method for assessing inhibitor binding, we performed similar titrations of amine guanidine-inhibitors **UM079**, **UM249** and **UM251** into both PRMT1 and CARM1. The structures of these inhibitors are given in Figure 15 (page 27) and above Table 19 (page 136). Curves are given in Figure 48 and Figure 49 (below) and the values determined by curve fitting are given in Table 19 (page 136).

In previous work,[69] a closely related inhibitor series, in which the propylamine substituent has an additional carboxylic acid group, was tested. These inhibitors (synthesised by Dr. W. Hong[79] and referred to herein as **WH3C**, **WH4C** and **WH5C** to reflect the guanidine linker length) displayed decreasing CARM1-inhbition on increasing the length of the alkyl linker between the adenosine and guanidine moieties, with **WH5C** showing no inhibition of CARM1 (see section 1.4.3, and IC_{50} values given in Table 19, below, for ease of comparison). We anticipated that the present inhibitor series (**UM079**, **UM249** and **UM251**) would show a similar trend. Surprisingly, this was not the case.

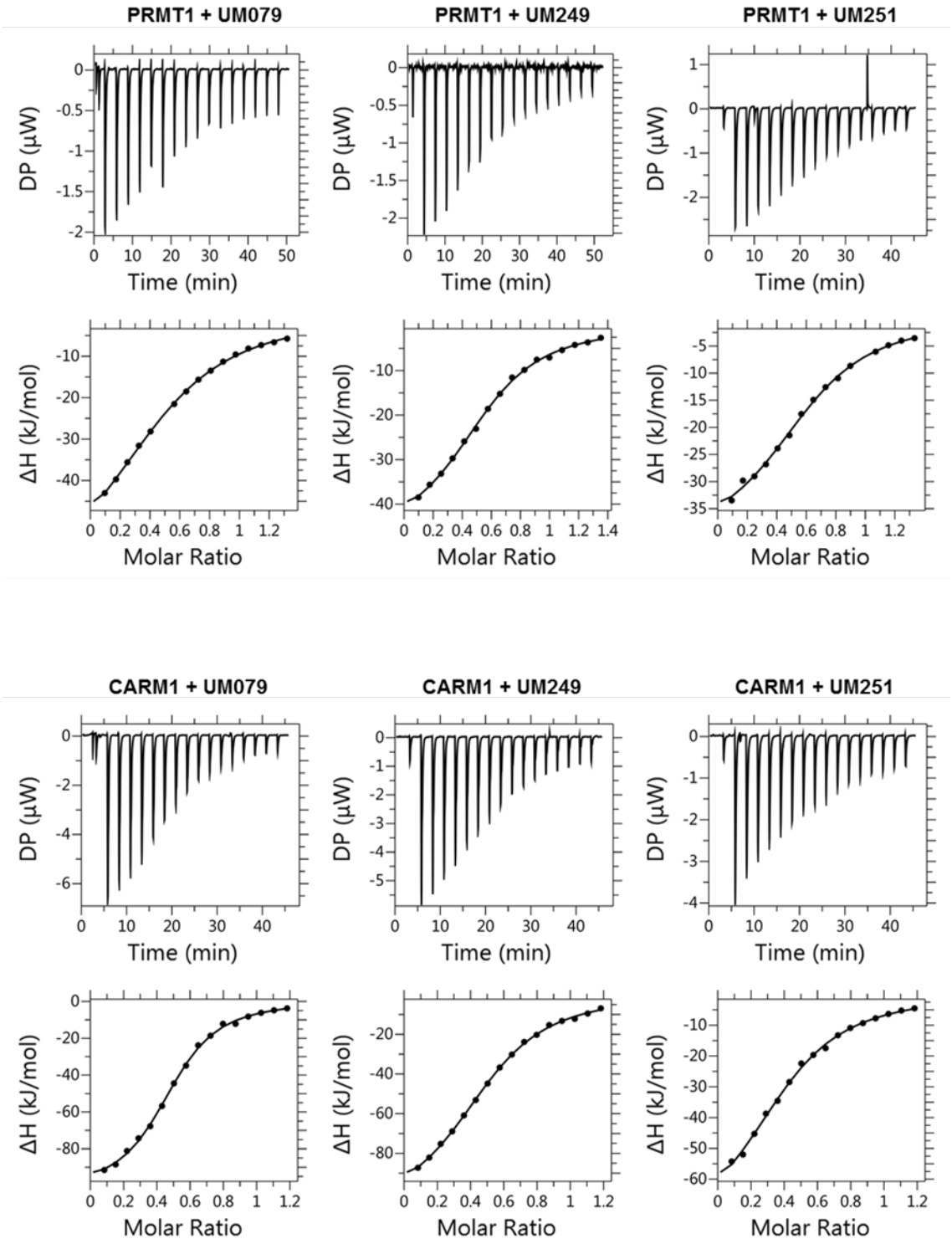


Figure 48: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM079, UM249 and UM251 with PRMT1 and CARM.

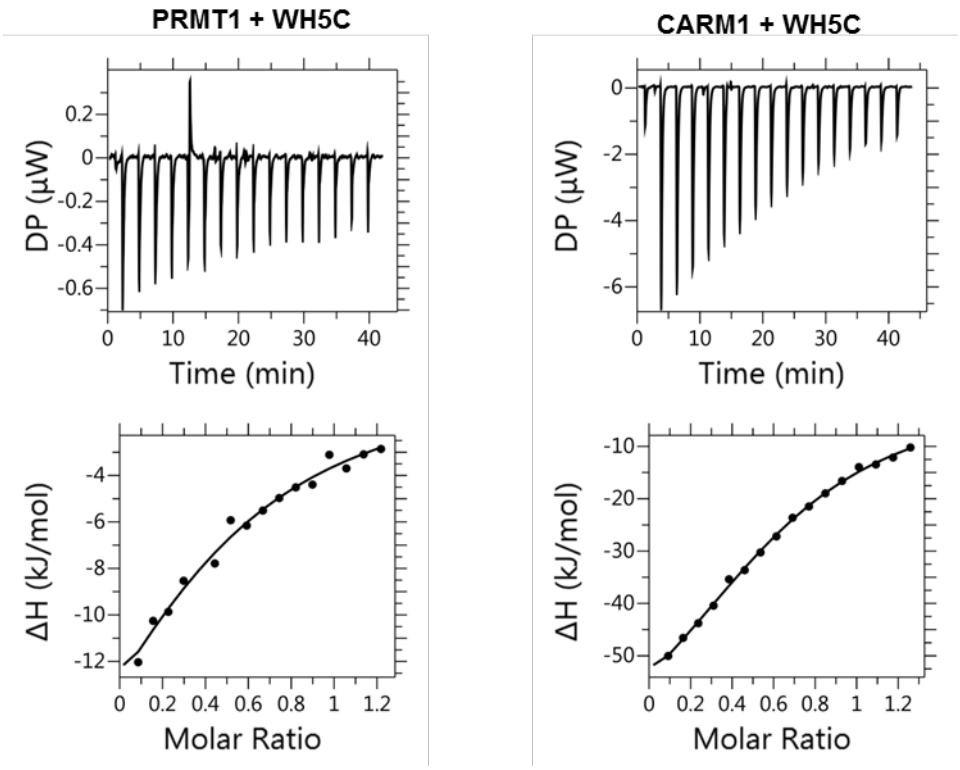
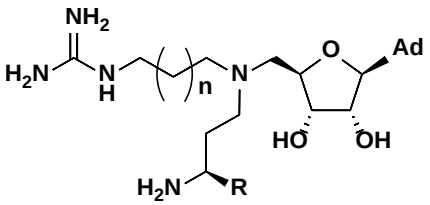


Figure 49: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for WH5C with PRMT1 and CARM.



R=H	R=CO_2H
UM079: n = 1	**WH3C:** n = 1
UM249: n = 2	**WH4C:** n = 2
UM251: n = 3	**WH5C**: n = 3

Inhibitor	K_d ± standard error (μM) from present study		IC_{50} ± standard deviation (μM) from previous study	
	PRMT1	CARM1	PRMT1	CARM1
UM079	10.2 ± 0.996	3.32 ± 0.335	2.2 ± 1.8[69]	3.3 ± 3.0[69]
UM249	4.02 ± 0.453	7.18 ± 0.557	-	-
UM251	9.99 ± 1.89	9.80 ± 1.83	-	-
WH3C	-	-	6.2 ± 3.8[68]	13.3 ± 8.7[69]
WH4C	-	-	2.9 ± 0.8[68]	>100[69]
WH5C	32.2 ± 3.53*	38.1 ± 9.61	5.6 ± 4.5[68]	no inhibition[69]

*Stoichiometry fixed to 0.5 due to low c-value curve

Table 19: K_d and IC_{50} values for our series of inhibitors versus PRMT1 and CARM1 measured using ITC (present study) and enzyme activity assays (previous work), respectively. Dashes indicate that values have not been determined.

Whilst the affinity of the inhibitor series for CARM1 decreases with increasing linker length (**UM079 → UM249 → UM251**), the three-fold increase in K_d is considerably less dramatic than the previously reported abrogation of inhibition on increasing the linker length from three to five carbons.[69] To determine whether the removal of the amino acid moiety was responsible for the 5C inhibitor unexpectedly binding to CARM1, an ITC titration of 5C amino acid-guanidine **WH5C** into PRMT1 and CARM1 was performed. **WH5C** was found to bind to both PRMT1 and CARM1 with similar K_ds (32.2 and 38.1 μM, respectively). A similar discrepancy between IC_{50} and K_d values was reported by Sack *et al*. for their indole and pyrazole inhibitors (structures not shown).[16b] As

noted by Sack *et al.*, IC_{50} and K_d are inherently different measurements: K_d is a direct measure of the protein-ligand affinity while IC_{50} reflects the effect of the inhibitor on the catalytic turnover of the enzyme under a specific set of assay conditions, so these parameters cannot be directly compared. Another possible cause for the difference in apparent potencies is the differing protein constructs used for the two studies (the catalytic cores of PRMT1 and CARM1 were used for K_d measurements reported herein, while full length PRMT1 and CARM1 GST-fusion proteins were used for the previously reported IC_{50} assays). To address this possibility, attempts were made at expressing two full length CARM1 constructs (GST-mCARM1(FL) and TF-CARM1(FL), Figure 19, page 83). These attempts were unsuccessful (see Section 3.4.6, page 113), therefore the possibility remains that the C-terminal extension (Figure 4, page 5) of CARM1 influences the binding of the inhibitors described herein.

In addition to K_d values, ITC data can also be used to determine the molar enthalpy and entropy changes (ΔH and ΔS, respectively) of ligand binding under the titration conditions. Signature plots (Figure 50, page 138) show that the binding of all ligands tested were enthalpically favourable (i.e. $\Delta H<0$) and entropically unfavourable ($-T\Delta S>0$). For PRMT1, both the unfavourable entropy of binding and favourable enthalpy of binding are reduced in magnitude as the alkyl linker length is increased. This trend was not observed for CARM1. Notably the increased affinity of SAH for both PRMT1 and CARM1 when compared to our inhibitors is entirely enthalpic in origin: the bisubstrate inhibitors have a more favourable binding entropy to both PRMT1 and CARM1

when compared to SAH. This may reflect the entropically unfavourable ordering of water molecules around the alkyl linker in solution, which is alleviated on inhibitor binding (i.e. the hydrophobic effect). This effect occurs irrespective of the enzyme target and is therefore not of use for informing selective inhibitor design.

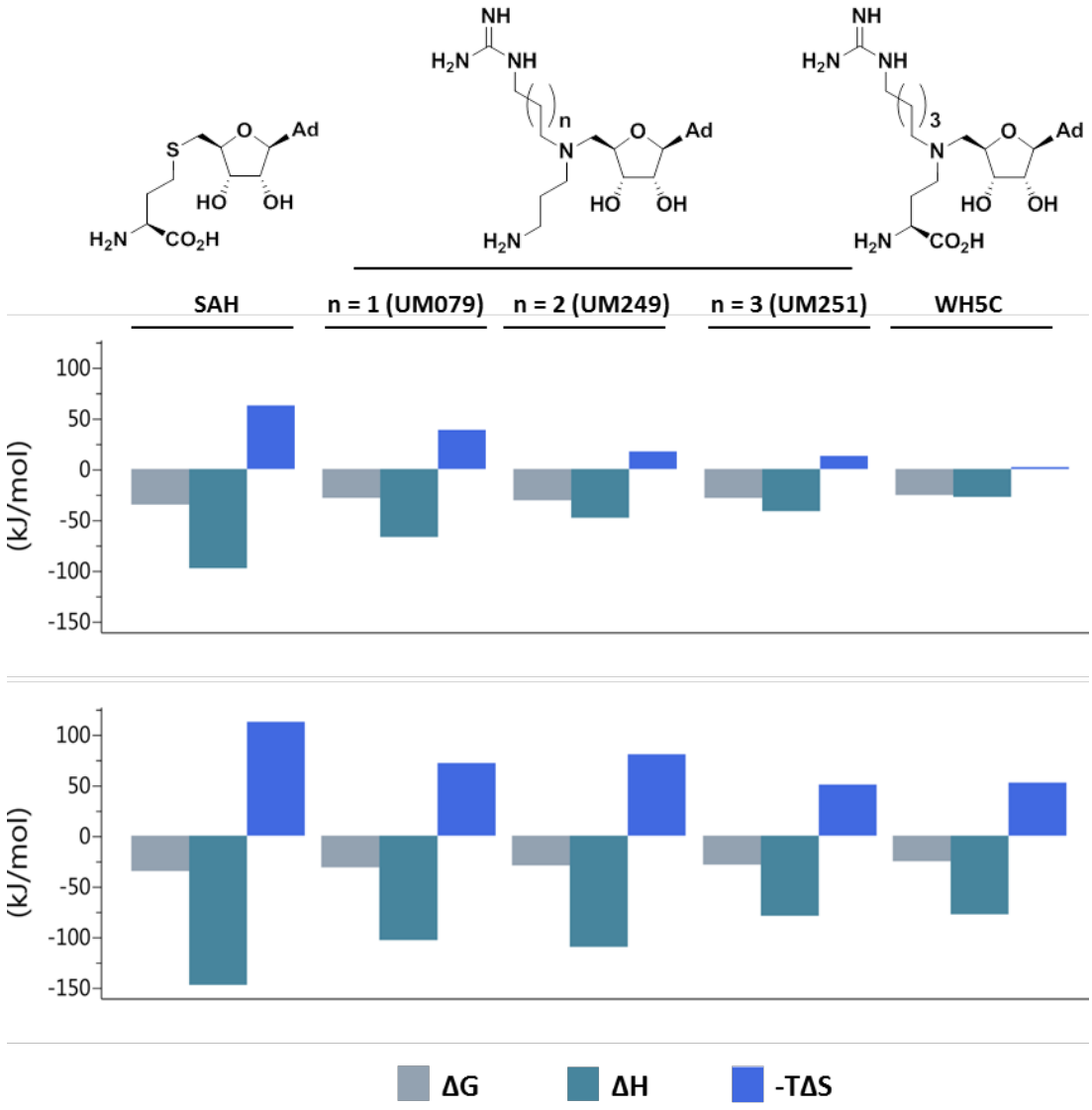


Figure 50: Signature plots for SAH, UM079, UM249, UM251 and WH5C binding to PRMT1 (upper) and CARM1 (lower). The inhibitor to which each dataset corresponds is shown above the graph. The data are given numerically in Table 20 and Table 21, below.

Inhibitor	PRMT1			
	ΔG (kJ/mol)	ΔH ± standard error from curve fitting (kJ/mol)	-TΔS (kJ/mol)	N ±standard error from curve fitting
SAH	-35.1	-97.7 ± 1.4	62.6	0.485 ± 0.0037
UM079	-28.5	-67 ± 3.22	38.5	0.519 ± 0.0082
UM249	-30.8	-48.1 ± 1.86	17.2	0.549 ± 0.0092
UM251	-28.6	-41.5 ± 2.79	12.9	0.607 ±0.019
WH5C	-25.7	-27.5 ± 1.4	1.84	N/A*

*The stoichiometry was fixed to 0.5 due to a low c-value curve.

Table 20: ΔG, ΔH, -TΔS and stoichiometry (N) determined from titrations of SAH, UM079, UM249, UM251 and WH5C into PRMT1. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

Inhibitor	CARM1			
	ΔG (kJ/mol)	ΔH ± standard error from curve fitting (kJ/mol)	-TΔS (kJ/mol)	N ±standard error from curve fitting
SAH	-34.9	-147 ± 4.62	113	0.549 ± 0.0086
UM079	-31.3	-103 ± 2.65	71.9	0.481 ± 0.0062
UM249	-29.4	-110 ± 2.98	80.5	0.501 ± 0.0060
UM251	-28.6	-79.2 ± 6.42	50.5	0.417 ± 0.012
WH5C	-25.2	-77.7 ± 9.65	52.5	0.648 ± 0.032

Table 21: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titrations of SAH, UM079, UM249, UM251 and WH5C into CARM1. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

3.6.2. Radiometric (^{3}H) Activity Assays

As a complimentary means for assessing inhibitor potency, we sought to use an activity assay to determine IC_{50} values for the amine-guanidine inhibitor series. An end-point radiometric activity assay was selected for its simplicity and diverse substrate compatibility.

3.6.2.1. Purification of GB-fused Histone Substrates

Histones 3 and 4 were expressed (as GB-fused peptides) and purified (Figure 51, below) for use as substrates. Brief optimisation of the assay was performed for both PRMT1 and CARM1, and is detailed in the following section.

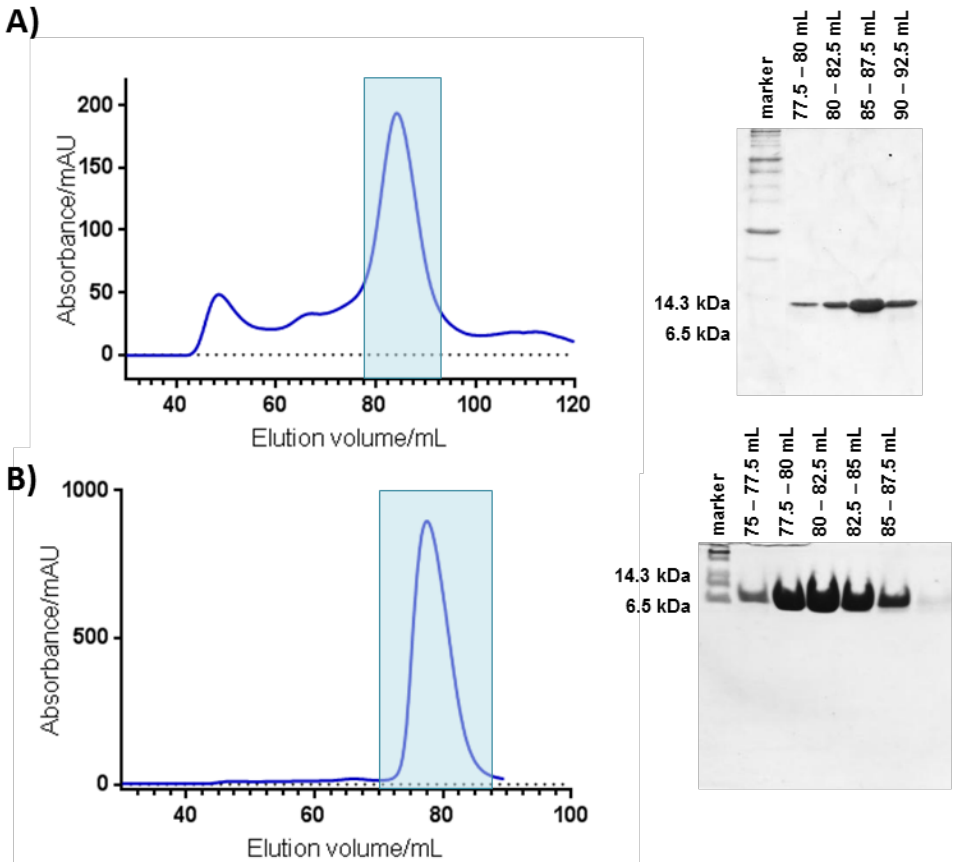
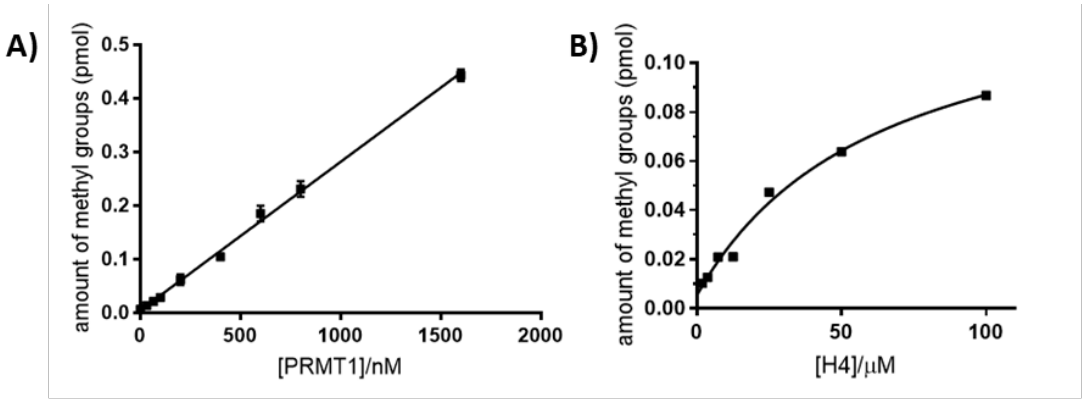


Figure 51: Purification of histone GB fusions. (A) GB-H4(1-21)-His gel filtration chromatogram and SDS-PAGE; (B) GB-H3(3-31)His gel filtration chromatogram and SDS-PAGE. Nickel chromatograms and SDS-PAGE analysis of nickel fractions are not shown. Turquoise boxes indicate fractions combined for assays. A Superdex S75 16/60 was used in both cases.

An alternative H3 construct (GB-H3(3-31)His) to the one used for co-purifications (GB-H3(21-44), Figure 21, page 86) was used for methylation assays. This was because we became aware of a published study, which showed that CARM1 constructs lacking the N-terminal domain preferentially methylate H3R17 over H3R26.[16a]

3.6.2.2. PRMT1 Assay Optimisation

Enzyme dilution assays were performed to determine the range over which the assay signal varied linearly with PRMT1 concentration (Figure 52A, page 143). The assay signal was found to vary linearly up to >1500 nM PRMT1. A PRMT1 concentration of 160 nM was chosen to use for IC_{50} assays. This represents a compromise between signal-to-noise ratio of the assay readout and the lowest measurable IC_{50} (which is equal to half the enzyme concentration). Substrate dilution assays were performed to determine an approximate value for the apparent Michaelis Menten (K_m^{app}) constant of the histone substrate under the assay conditions (Figure 52B, page 143). According to the Cheng-Prusoff equation ($K_i=IC_{50}+1/(1+[S]/K_m)$),[108] the observed IC_{50} values of substrate-competitive inhibitors tend towards the K_i values as the substrate concentration is lowered. Substrate concentrations significantly higher than their K_m values result in IC_{50} values that do not reflect true inhibitor potency. Since the apparent K_m of H4 and SAM are interdependent, a fixed, low, concentration of SAM (0.23 μM) was chosen and the K_m^{app} of the histone substrate determined to be approximately 60 μM. A histone concentration of 40 μM was therefore used for subsequent IC_{50} assays, which is a compromise between assay signal (achieved through high substrate concentration) and sensitivity to inhibition (achieved through lowering the substrate concentration below the K_m^{app}).



K_m = 63.4 ± 22.5 μM
R^2 = 0.988
V_{max} = 0.133 ± 0.0194 pmol/h/2.4 pmol enzyme
= 0.023 nmol/min/mg enzyme

Figure 52: PRMT1 ^{3}H assay validation. (A) Determining linear range of PRMT assay signal. Assays contained 0.23 μM ^{3}H-SAM, 40 μM GB-H4, (1-21)His, 0 – 1600 nM PRMT1 in a total volume of 15 μL PBS; (B) Estimate of K_m^{app} of histone substrate under assay conditions. The K_m^{app} and V_{max} determined by nonlinear regression fitting to the Michaelis-Menten equation and the standard errors of curve fitting are given underneath the graph. The R^2 value is also given. Assays contained 0.23 μM ^{3}H-SAM, 120 nM PRMT1 and 1.8 – 100 μM GB-H4(His) in a total volume of 15 μL PBS.

.

3.6.2.3. CARM1 Assay Optimisation

Similar enzyme dilutions were performed with CARM1 (Figure 53, below), and a CARM1 concentration of 320 nM was chosen. Determining a suitable histone concentration for CARM1 was more complex due to its apparent automethylation (Figure 31, page 106). A histone concentration of 120 μM was initially chosen arbitrarily, and was used for IC_{50} measurements. Subsequently, substrate dilution assays were performed and determined the K_m^{app} of GB-H3(3-31)His to be approximately 40 μM (Figure 54, below).

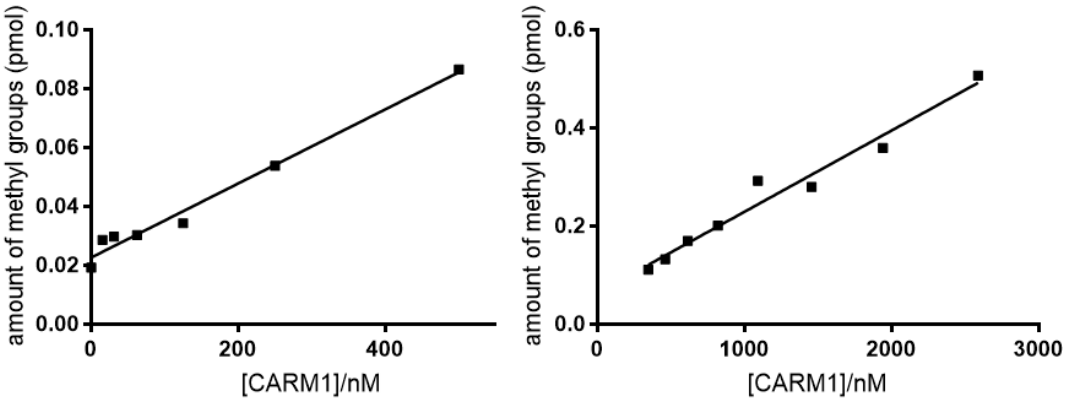


Figure 53: Determining linear range of CARM1 assay signal. Assays contained 0.23 µM ^{3}H-SAM, 120 µM GB-H3(1-31)His and the indicated concentration of CARM1 in a total volume of 15 µL PBS.

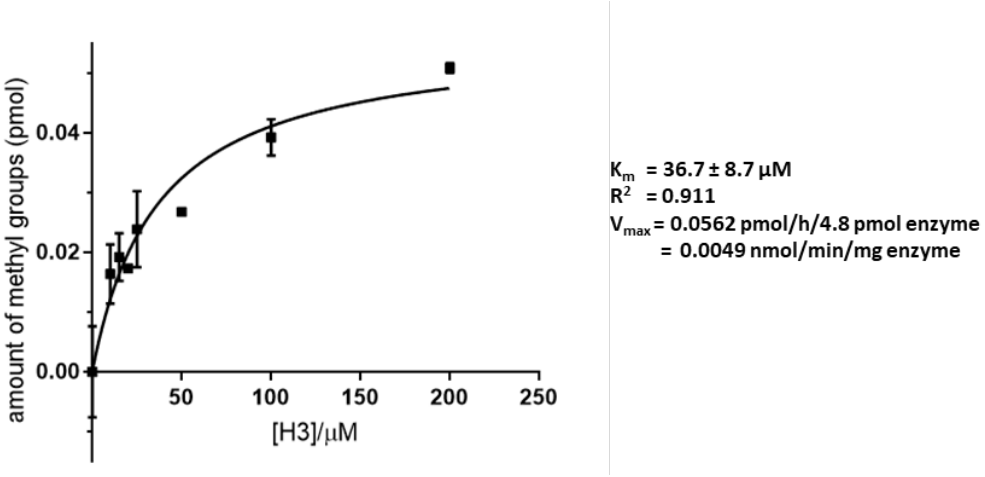


Figure 54: Estimate of Km^{app} of CARM1 histone substrate (GB-H3(3-31)His) under assay conditions. The K_m^{app} and V_{max} determined by nonlinear regression fitting to the Michaelis-Menten equation and the standard errors of curve fitting are given beside the graph. The R^2 value is also given.

Based on the Cheng-Prusoff equation, this may have resulted in CARM1 IC_{50} values that are significantly higher than the corresponding K_is (i.e. an

underestimation of inhibitor potency *versus* CARM1). The kinetics in our system are likely more complex than the ideal Michaelis Menten model, however detailed mechanistic studies are beyond the scope of this work. While further optimisation of the CARM1 assay would have been informative, the primary purpose of the assay was to determine whether the apparent discrepancies between the data from the present study and previous work are due to the inherent differences between IC_{50} and K_d measurements. The IC_{50} values determined using the activity assay are therefore useful, but require consideration of the caveats discussed above.

3.6.2.4. IC_{50} Assays

Having established acceptable assay conditions for both PRMT1 and CARM1, the 'linker length' amine-guanidine inhibitor series was tested. The IC50 curves for amine-guanidines UM079, UM249, UM251 and WH5C and results from curve fitting are given in Figure 55, Figure 56 and Table 22 (below).

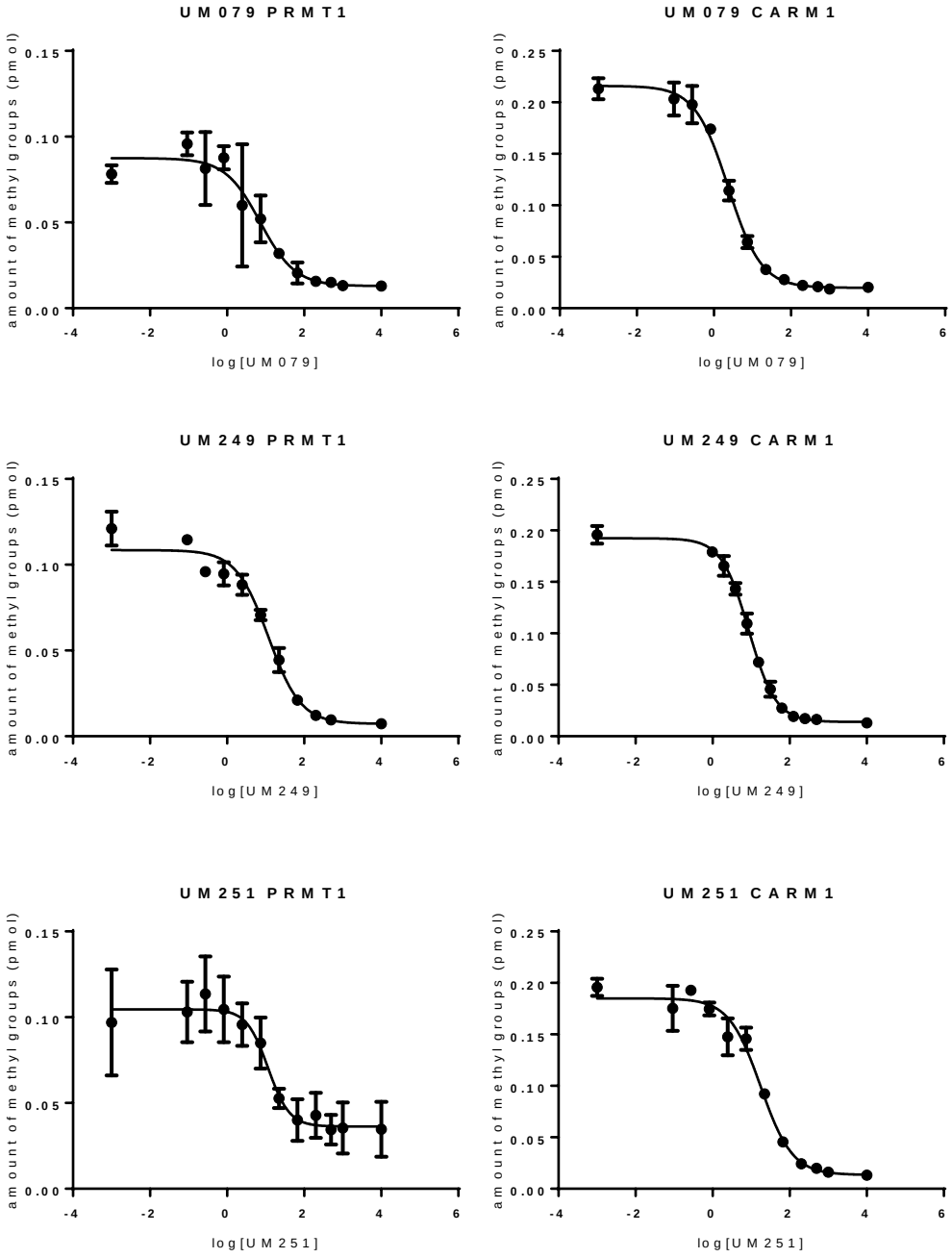


Figure 55: IC_{50} assay data for linker length series UM079, UM249 and UM251 with PRMT1 and CARM1. The error bars show standard deviations of duplicate measurements at each inhibitor concentration. The IC_{50} values determined by curve-fitting are given in Table 22 (below).

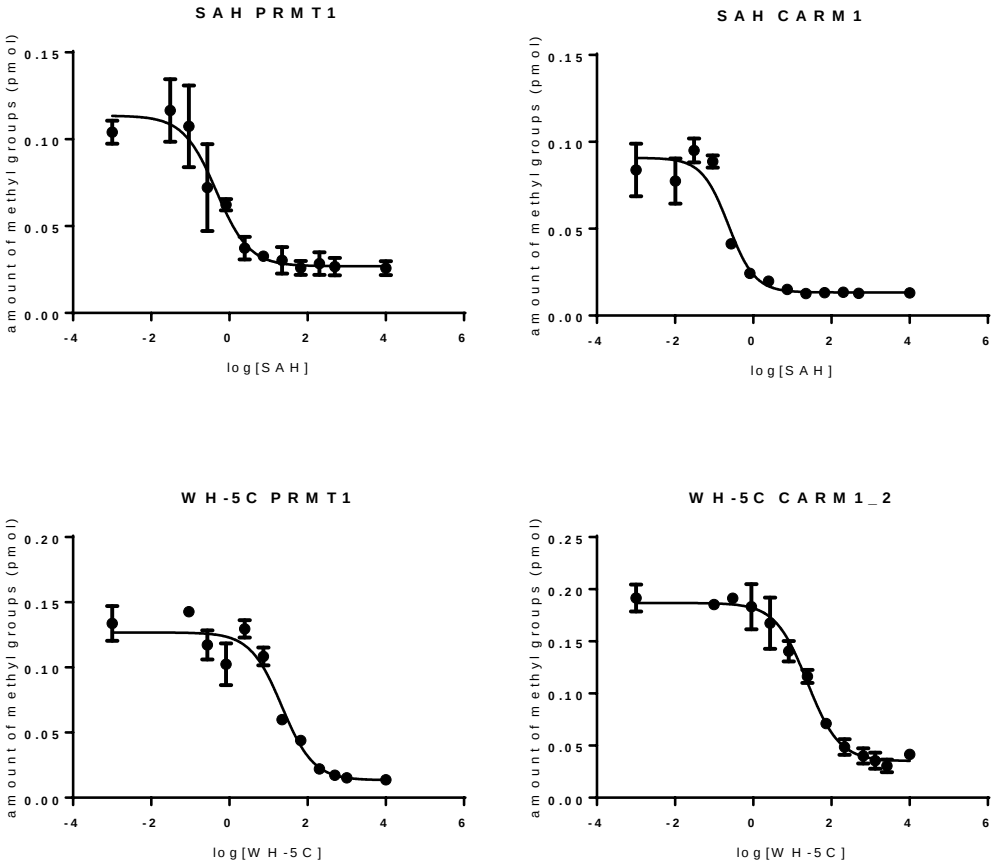


Figure 56: IC_{50} assay data for SAH and WH5C with PRMT1 and CARM1. The error bars show standard deviations of duplicate measurements at each inhibitor concentration. The IC_{50} values determined by curve-fitting are given in Table 22 (below).

Inhibitor	IC_{50} (95 % confidence interval) (μM) from present study		IC_{50} ± standard deviation (μM) from previous study	
	PRMT1	CARM1	PRMT1	CARM1
UM079 (3C)	7.2 (3.5 – 15.0)	2.3 (1.8 – 3.0)	2.2 ± 1.8[69]	3.3 ± 3.0[69]
UM249 (4C)	11.8 (9.0 – 15.5)	8.7 (6.9 – 10.8)	-	-
UM251 (5C)	11.6 (4.9 – 27.2)	17.7 (13.5 – 23.2)	-	-
WH5C	22.0 (14.4 – 35.5)	24.1 (15.4 – 37.6)	5.6 ± 4.5[68]	no inhibition[69]

Table 22: IC_{50} values for our series of inhibitors versus PRMT1 and CARM1 measured using enzyme activity assays. IC_{50} values reported in the study by Dowden *et al.* are given for comparison. Dashes indicate that the inhibitor was not tested.

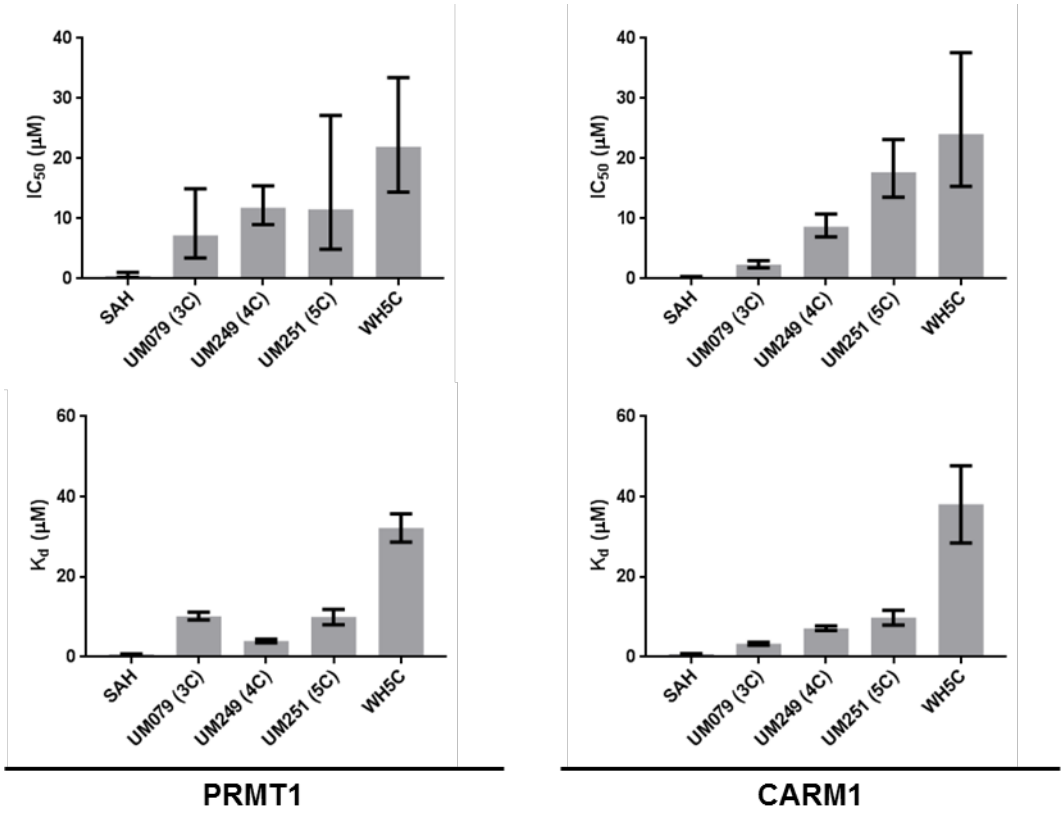


Figure 57: Graphs showing IC_{50} (top) and K_d (bottom) values determined for PRMT1 (left) and CARM1 (right). The error bars for IC_{50} values are 95 % confidence intervals. The error bars for K_d values are standard errors from curve-fitting.

A trend of decreasing potency with increasing alkyl linker length was observed for CARM1. The lower precision (i.e. wider 95 % confidence intervals) of the values determined for PRMT1 meant that a similar trend could not be inferred. Comparing the IC_{50} and K_d values (Figure 57, above), shows that they are in reasonable agreement. The same overall trend, of decreasing affinity with increasing linker length from 3 to 5 carbons, is seen for CARM1 IC_{50} and K_d values. The trends for PRMT1 are less clear, but the two datasets both show that the inhibitors are low micromolar inhibitors of PRMT1. Taken together, the IC_{50} and K_d values determined for this inhibitor series show that adjusting the alkyl linker length in this series of compounds is not a useful approach to achieving selectivity for PRMT1 or CARM1. As mentioned earlier, a significant caveat of this conclusion is that the CARM1 construct used for these data is not full length. Most reported studies of CARM1 inhibitors use the catalytic core for crystallography and ITC, but use full length CARM1 fusion proteins (obtained using mammalian or insect expression systems) for inhibition assays. We have found full length CARM1 to be difficult to obtain using bacterial expression (see Section 3.4.6, page 113). A crystal structure of full length CARM1 would provide valuable insight into how the CARM1 C-terminal domain influences the active site architecture, and whether the catalytic core is a suitable surrogate for the full length protein in studying inhibitor binding.

3.6.2.5. Crystal Structures of Full Linker Length Series

To gain structural insight into the effect of the alkylguanidine linker length, crystals were grown and diffraction data were collected for CARM1 in complex with amine-guanidines **UM079** (3C) and **UM251** (5C) and amino acid guanidine **WH5C** (5C). Since differences in inhibitor structure and protein batch may affect crystal growth, several PEG concentrations, CARM1-His concentrations and additives were used when setting up 24-well crystallisation trays with each new inhibitor. Details of the crystallisation conditions used for each structure are given in Section 2.7.1 (page 53) and crystallographic data processing statistics are given in Appendix 6.5. The structures were solved by molecular replacement using a monomer from the CARM1·**249** structure (Section 3.5.2, page 127) as the search model. The space group ($P2_12_12$) and number of molecules in the asymmetric unit (4) were the same for all structures.

As seen for the 4C inhibitor **UM249**, clear electron density for the relevant ligand was present in all active sites of each structure (Figure 58C, D, E, page 151). Superposing the structures with a published crystal structure of a CARM1·H3·sinefungin ternary complex showed that the 4 inhibitors in this series adopt very similar conformations within the CARM1 active site (Figure 58B, page 151). The adenosine portion of the inhibitor superposes well with that of sinefungin. The inhibitor guanidinium group does not superpose well with that of the histone arginine targeted for methylation (H3R17). In comparison to the position of H3 R17, the guanidinium groups of all inhibitors reach further out past E257 (which forms a bidentate interaction with the terminal nitrogens

of the substrate arginine) and E266 (which forms a hydrogen bond to the ε nitrogen of substrate arginine).

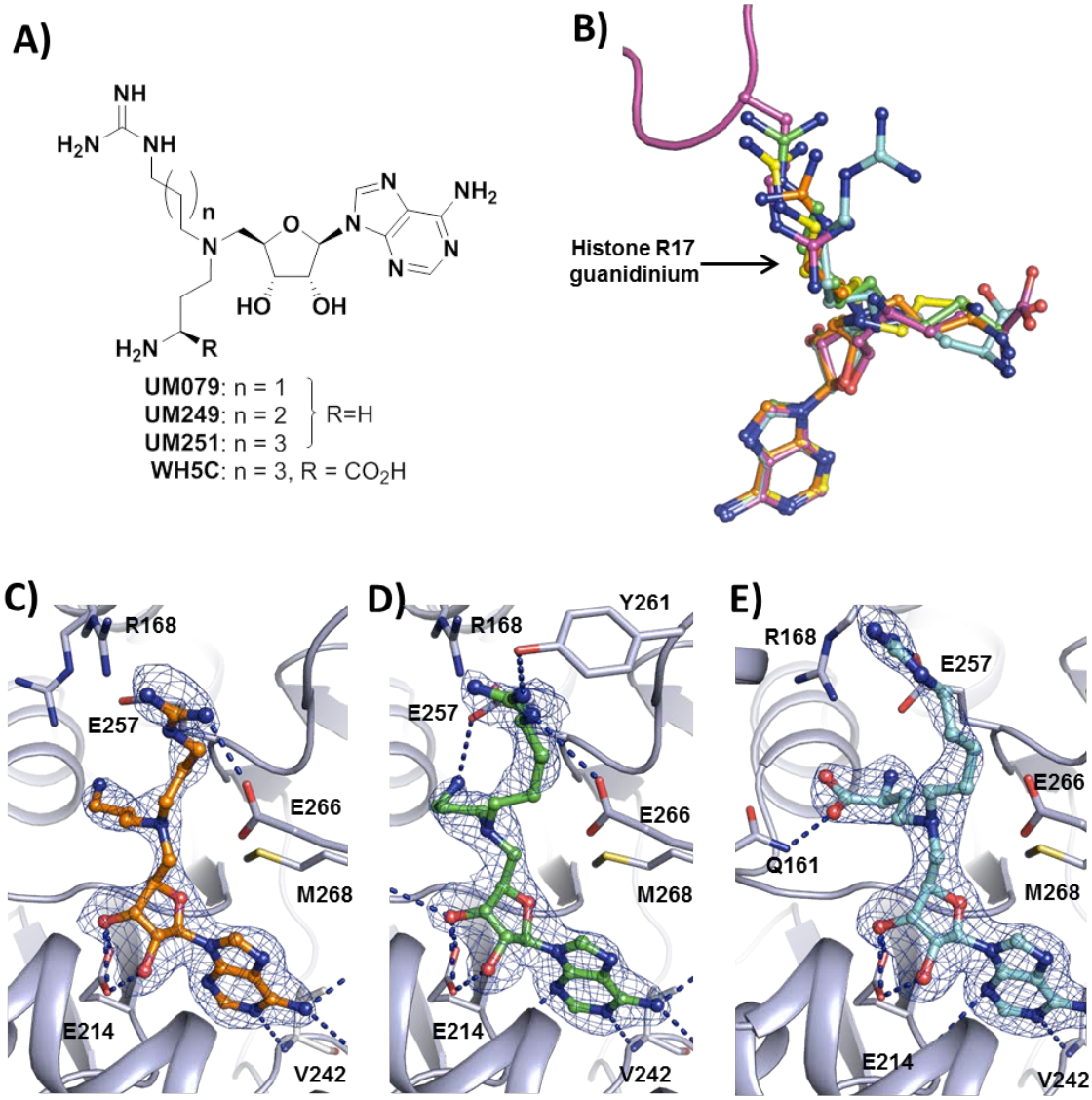


Figure 58: Co-crystal structures of amine guanidine inhibitors with CARM1. (A) Inhibitor structures; (B) Superposition of UM079 (orange), UM249 (yellow), UM251 (green) and WH5C (cyan) co-crystals with a published structure of CARM1 in complex with histone 3 and sinefungin (magenta, PDB code 5DX0). (C)-(E) 2mF_o-DF_c maps (contoured at 1 σ) for UM079 (C), UM251 (D) and WH5C (E) bound in the CARM1 active site. The monomers for which the ligand has the clearest electron density are shown. Polar contacts found using PyMOL are indicated with blue dashed lines and residues believed to be involved in SAM and substrate arginine binding (and those involved in the putative polar contacts) are shown.

As for the 4C inhibitor, **UM249**, the B-factors of the guanidinium groups are relatively high compared to the adenosine moiety and surrounding residues (Appendix 6.3.1). This suggests that this group is flexible when bound in the active site. This in turn may be due to the alkyl linker preventing the guanidinium from interacting optimally with the active site glutamates (E257 and E266). Van Haren *et al.* recently reported a closely related inhibitor series, in which the alkylguanidine group extends directly from adenosine (Figure 59, below).[74]

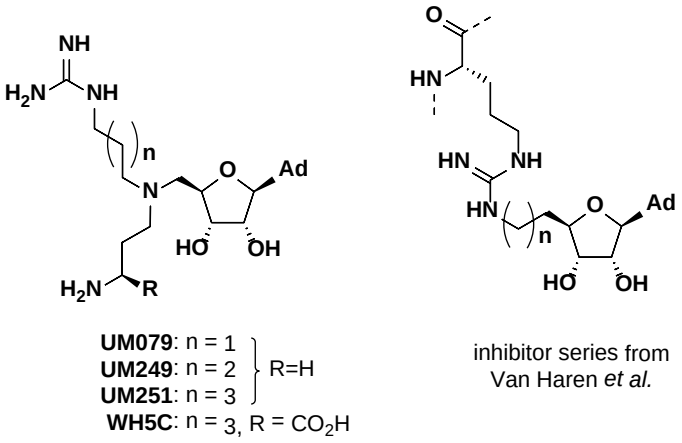


Figure 59: Comparison of 'linker length' amine-guanidine inhibitor series with peptide-fused inhibitors reported by Van Haren *et al.*[74]

They found that a 3-carbon methylene linker between the 4' carbon of adenosine and a guanidinium group (n=2, Figure 59, above; equivalent to a 1-carbon linker in our inhibitors) resulted in stronger inhibition of CARM1 than a 2-carbon linker, with the former appearing well-orientated to interact with both E257 and E266. This is consistent with our observation that the 3-carbon linker

of **UM079** is not sufficiently short to position the inhibitor guanidinium for optimal interaction with E257 and E266.

Taken together, our results show that alkyl linker adjustment may be useful for optimising affinity for individual PRMTs (as shown by Van Haren *et al.*), but inhibitors **UM079**, **UM249**, **UM251** and **WH5C** do not discriminate between the catalytic cores of PRMT1 and CARM1.

3.6.3. Homology Model of Full Length CARM1

A caveat of the experiments presented so far is that they were all performed using the catalytic core of CARM1, which lacks the N- and C- terminal domains of the full length protein (Figure 4, page 5). In the absence of access to full length CARM1 for crystallisation and assays (as discussed in Section 3.4.6, page 113), a homology model was generated using the PHYRE2 molecular fold recognition server.[19] To gain insight into whether the N- and C-terminal domains of full length CARM1 (which are absent in our CARM1-His construct) may influence inhibitor binding, we superposed our CARM1·**WH5C** crystal structure with the CARM1(FL) homology model (Figure 60, below).

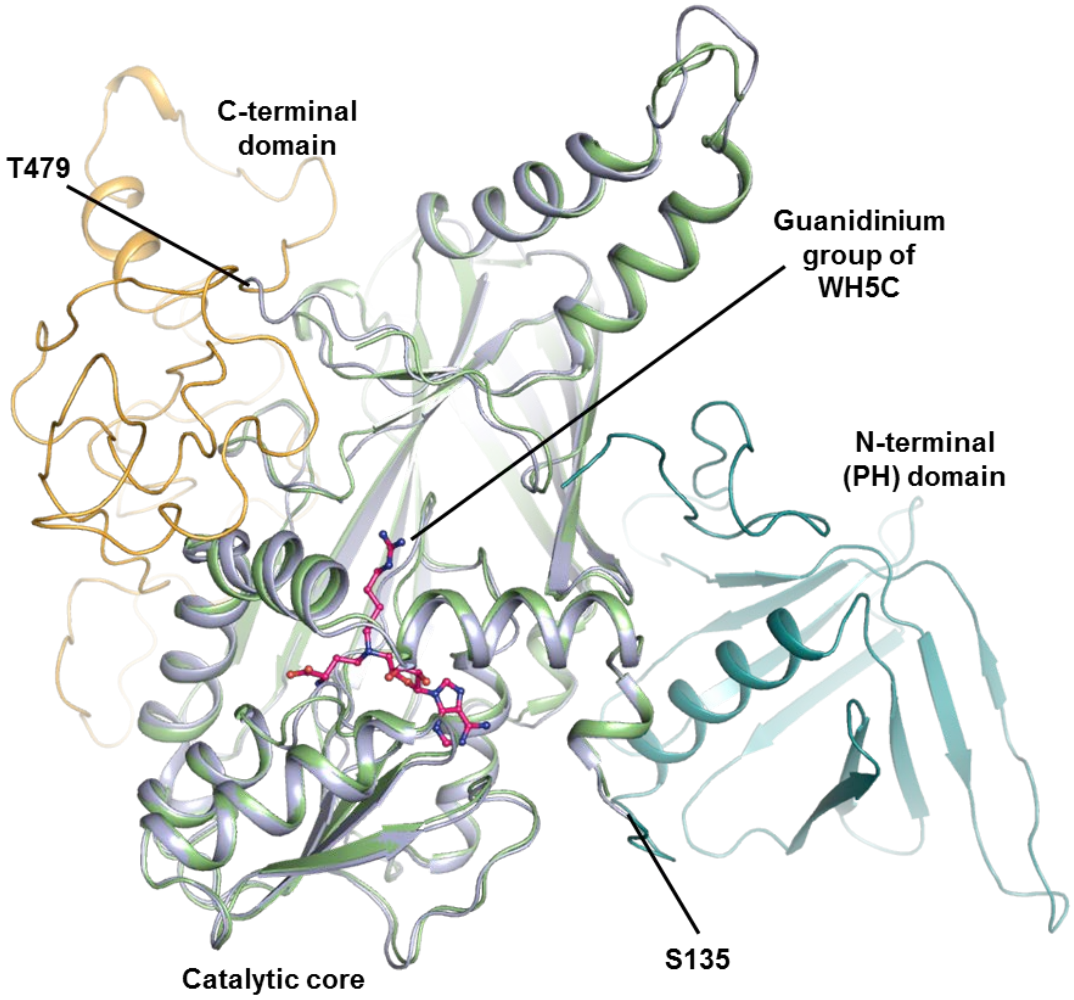


Figure 60: Full length CARM1 homology model superposed with CARM1·WH5C co-crystal structure. The N-terminal (PH) domain of the homology model is shown in teal; the C-terminal domain in orange, and the catalytic core in grey. The CARM1·WH5C co-crystal structure is shown in green (protein) and pink (WH5C).

In the predicted model, the N- and C-terminal domains are not sufficiently close to the active site to interact directly with the inhibitor.

To account for the dimerisation behaviour of CARM1, an additional model was built by superposing two CARM1(FL) models onto a CARM1 dimer from our CARM1·**WH5C** crystal structure (Figure 61, below). In this CARM1(FL) dimer

model, the N- and C-terminal domains of one CARM1 monomer are not sufficiently close to the active site of the opposing monomer to allow direct interaction with the bound inhibitor.

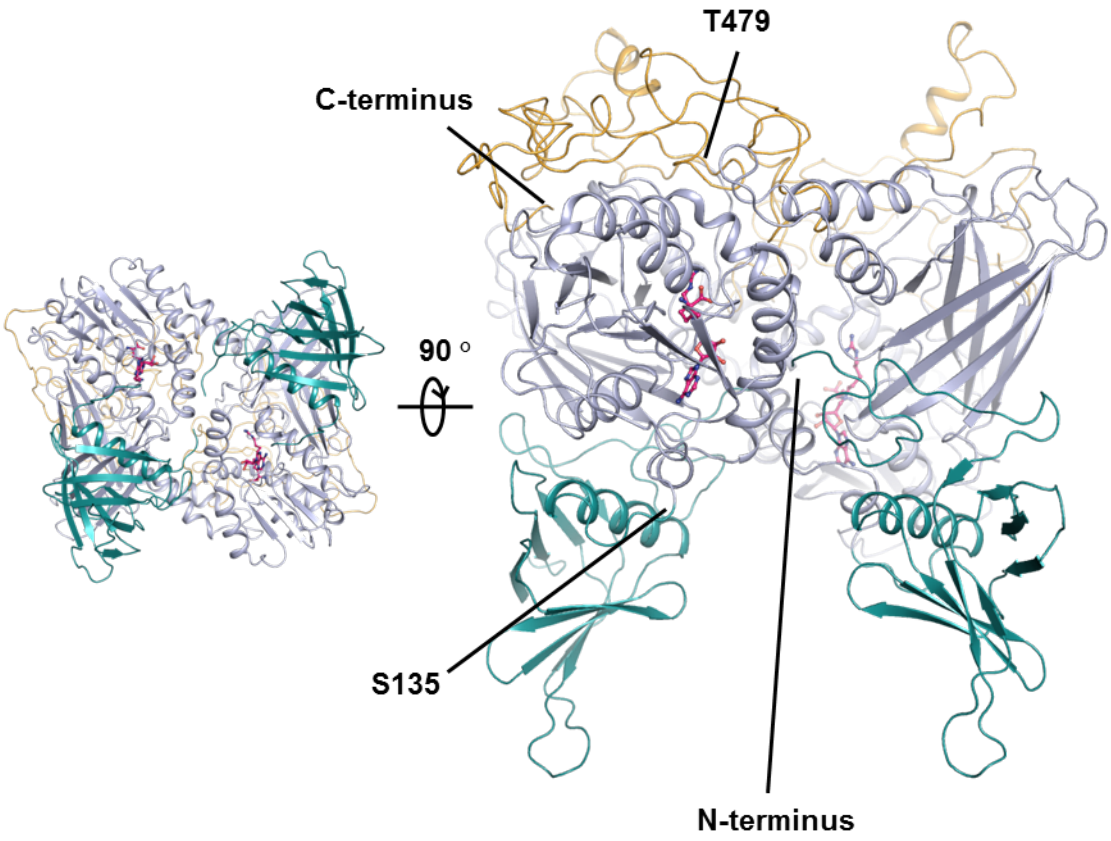


Figure 61: Full length CARM1 homology model superposed with CARM1 dimer from our CARM1·WH5C co-crystal structure. The N-terminal (PH) domain of the homology model is shown in teal; the C-terminal domain in orange, and the catalytic core in grey. The CARM1·WH5C protein is omitted for clarity and WH5C is shown as pink ball-and-sticks.

These models suggest that the N- and C-terminal domains are too far from the inhibitor binding site to directly influence inhibitor binding. It is, however,

important to note that in the absence of experimental, structural data regarding the C-terminal domain of CARM1, this area of the model is not reliable. While the structure of the N-terminal CARM1 domain is known (PDB code 2OQB),[109] the angle between the catalytic core and N-terminal domain is not known. Further, the presence of these peripheral domains could influence the structure of the catalytic core, thereby indirectly influencing inhibitor affinity (i.e. long range effects).

3.7. Replacement of Guanidine with Aromatic Isosteres

3.7.1. Binding of Aromatic Inhibitors

Having established that the amine-guanidine inhibitors were neither PRMT1 nor CARM1-selective, focus turned to a related series of compounds: **UM305**, **AA175** and **AA183** (Figure 62, below). Replacement of the guanidinium moiety was proposed as a means of reducing the number of charged groups, which may enhance the membrane-permeability of these inhibitors.[67] ITC and activity assays were used (respectively) to determine the K_d and IC_{50} values for this set of inhibitors, and the curves and determined values are given in Figure 63, Figure 64 and Table 23 (below).

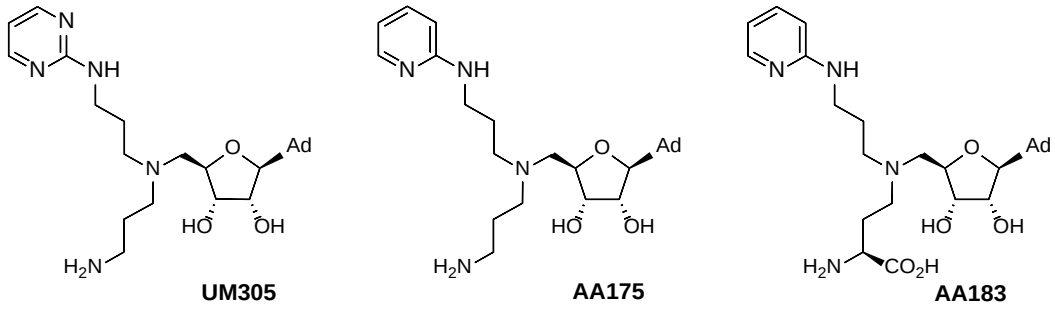


Figure 62: Structures of aromatic inhibitors UM305, AA175 and AA183

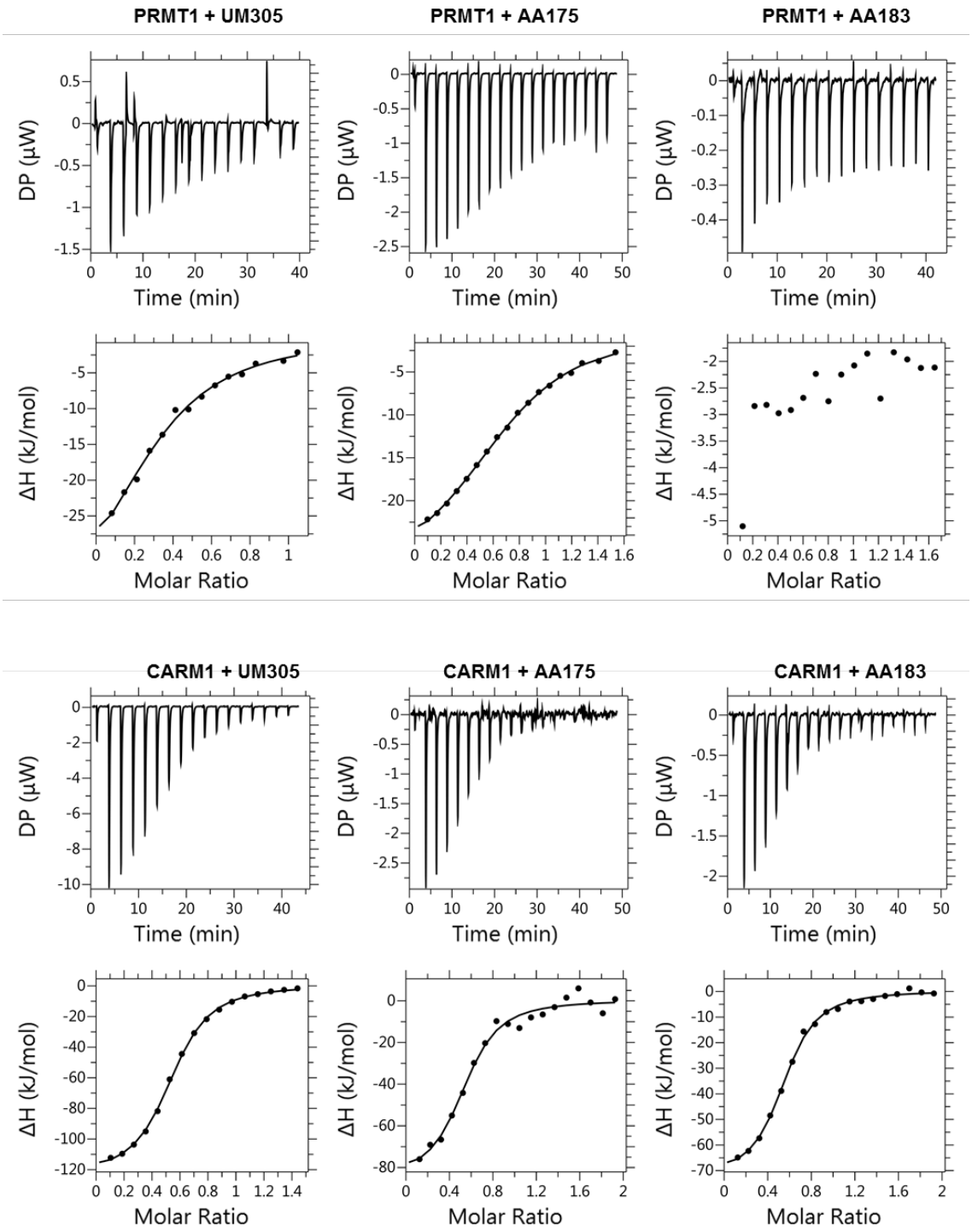


Figure 63: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM305, AA175 and AA183 with PRMT1 and CARM. The K_d values determined by curve fitting are given in Table 23 (below).

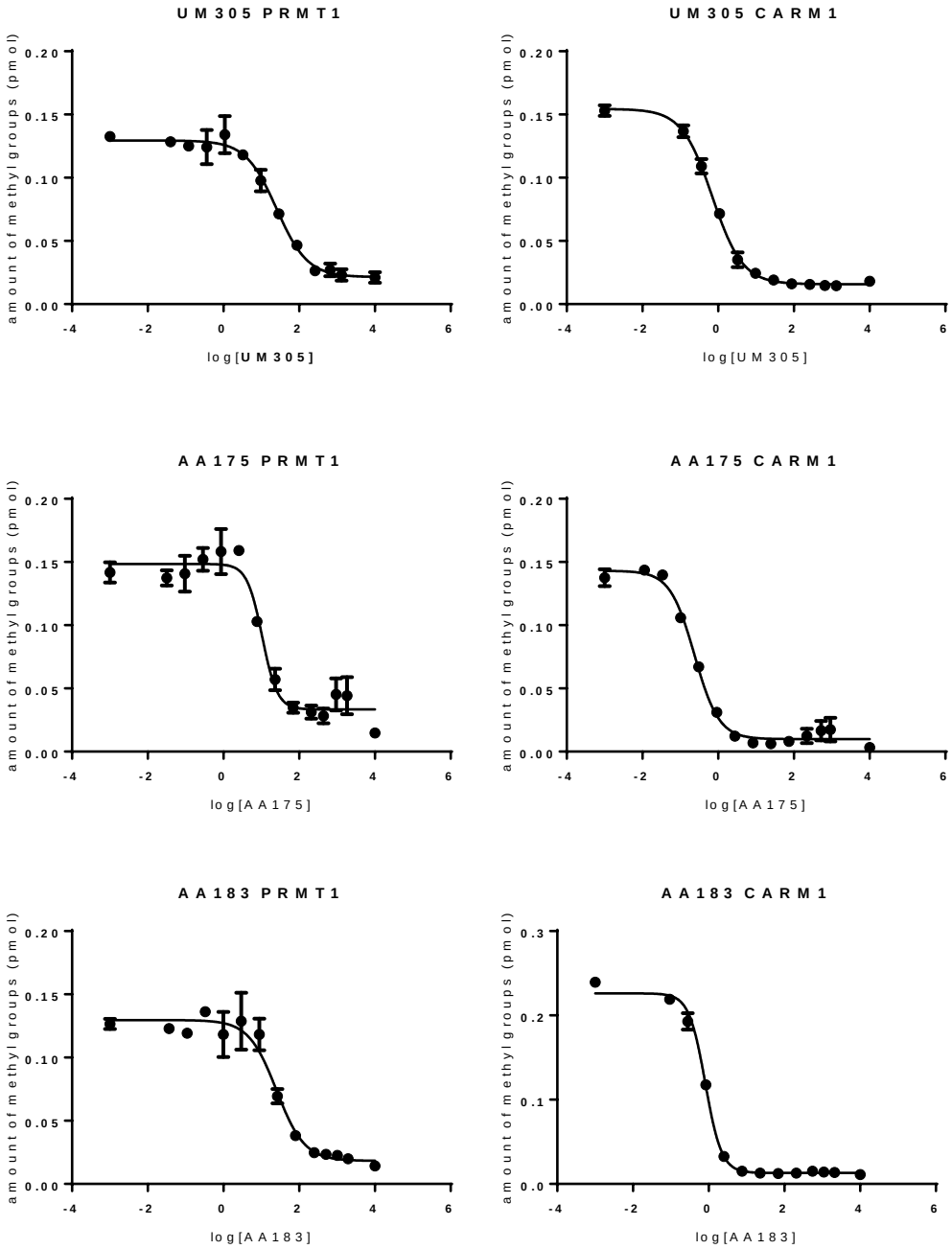
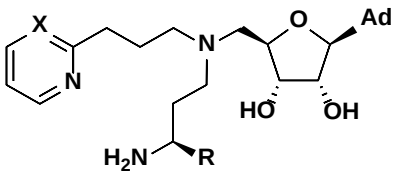


Figure 64: IC_{50} assay data for aromatic inhibitors UM305, AA175 and AA183 with PRMT1 and CARM1. The error bars show standard deviations of duplicate measurements at each inhibitor concentration. The IC_{50} values determined by curve-fitting are given in Table 23 (below).



UM305: R = H, X = N
AA175: R = H, X = CH
AA183: R = CO_2H, X = CH

3C aromatic isosteres

Inhibitor	K_d ± standard error (μM)		IC_{50} (95% confidence interval) (μM)	
	PRMT1	CARM1	PRMT1	CARM1
UM305	10.2 ± 4.21[1]	2.30 ± 0.162	22.0 (15.7 – 30.8)	0.7 (0.5 – 1.0)
AA175	15.3 ± 1.57[3]	1.14 ± 0.411[3]	11.1 (5.6 – 21.8)	0.3 (0.1 – 1.1)
AA183	No binding[2,3]	1.11 ± 0.132[3]	25.3 (17.3 – 37.0)	0.8 (0.6 – 1.0)

[1]Stoichiometry fixed to 0.5
[2]No binding detected using a standard titration; however a competition experiment with SAH determined the K_d to be 43.7 ± 14.9 μM
[3]Titrations performed by Alaa Al-Noori[80]

Table 23: K_d and IC_{50} values for aromatic isostere inhibitors *versus* PRMT1 and CARM1 measured using ITC and enzyme activity assays, respectively.

Remarkably, the aromatic inhibitors were all selective for CARM1 over PRMT1. Amino acid-aminopyridine AA183 was the most dramatic example, with a K_d of 1.11 μM for CARM1 and no binding detected for PRMT1 by ITC. Interestingly, on testing this compound using the methylation assay, an IC_{50} of 25.3 μM was determined. Throughout purification and crystallisation screening, we generally observed that PRMT1 was less stable compared to CARM1, and may therefore be more sensitive to denaturation. To exclude the possibility that AA183 inhibits PRMT1 through nonspecific destabilisation, SAH was titrated into a solution of

PRMT1 and AA183 to check that PRMT1 was still folded (Figure 65, below). If **AA183** caused precipitation of unfolding of PRMT1, SAH would no longer bind.

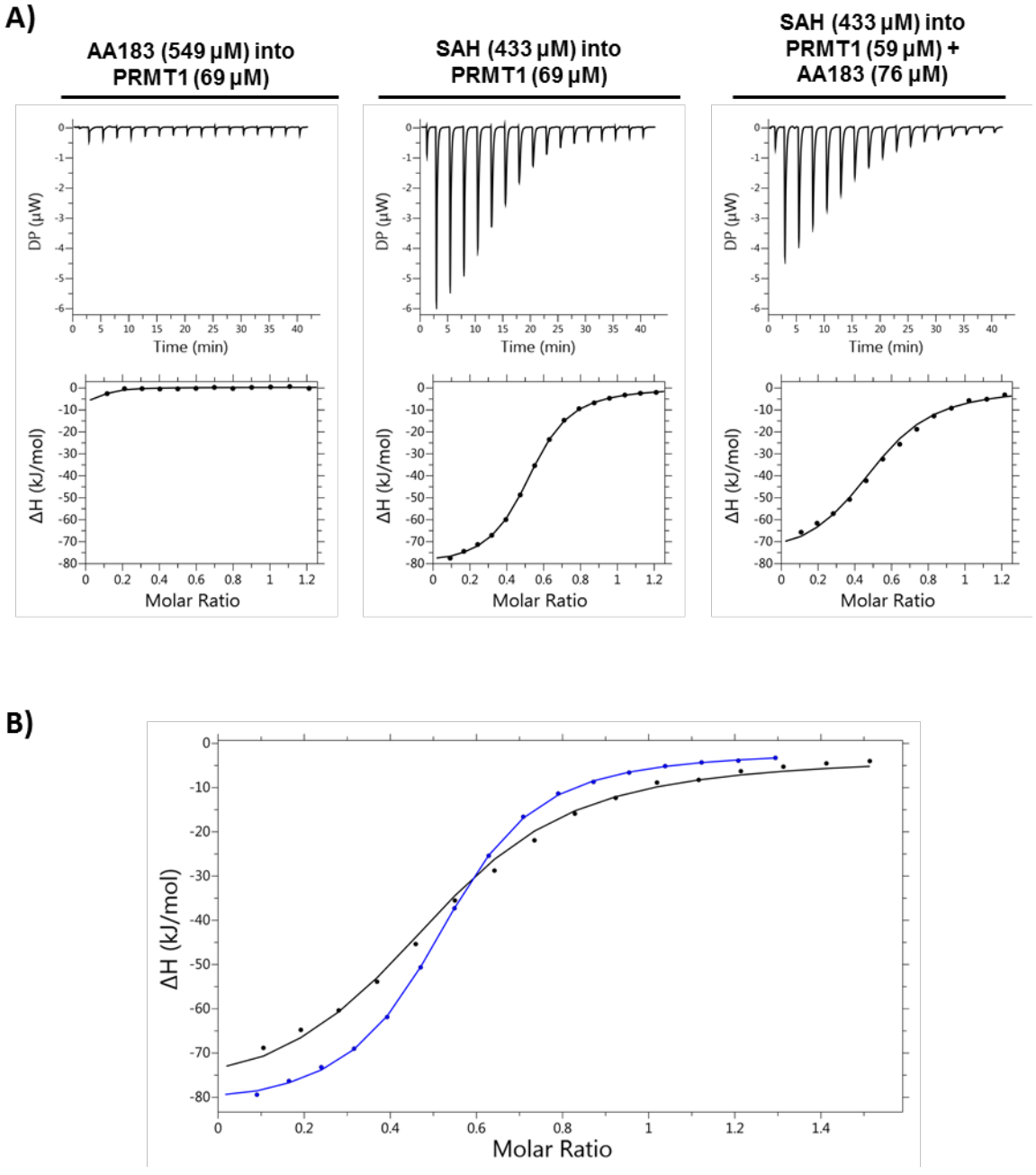


Figure 65: Competition ITC studies of AA183 binding to PRMT1. (A) From left to right: titration of AA183 into PRMT1; titration of SAH into PRMT1; titration of SAH into PRMT1 pre-incubated with AA183. (B) Superposition of integrated heats and fitted binding isotherms of SAH with PRMT1 in the presence (black) and absence (blue) of AA183. The change in gradient indicates that the affinity of SAH is lowered in the presence of AA183.

The presence of **AA183** in the cell resulted in an increase in the apparent K_d of SAH for PRMT1, while the stoichiometry remained unchanged. This indicates that **AA183** does not cause PRMT1 to unfold or precipitate. Fitting the resulting isotherm using a competitive binding model determined the K_d of **AA183** for PRMT1 to be 43.7 μM. This is broadly consistent with the IC_{50} value, and indicates that the lack of observed binding on titrating **AA183** into PRMT1 was due to the heat change associated with ligand binding being below the limit of detection. Hence, inhibitors **AA175** and **AA183** are, respectively, ~15- and ~40-fold selective for CARM1 (as judged by ITC) over PRMT1. While this degree of selectivity is suitable for compounds to be useful as chemical probes, further criteria (e.g. greater potency, confirmed selectivity over other methyltransferases and demonstrated cellular activity) would be required prior to using **AA175** and **AA183** to derive biological information related to CARM1 inhibition. The observed trend towards increased potency for CARM1 by using hydrophobic guanidine isosteres will be useful in the pursuit of CARM1 chemical probes.

As discussed earlier (Section 3.6.1, page 131), the enthalpic and entropic contributions to inhibitor binding can be extracted from ITC data. Signature plots (showing ΔG, ΔH and –TΔS) for the aromatic inhibitor series are given in Figure 66 (page 163). As for the guanidine inhibitors (Figure 50, page 138), the binding (to PRMT1 and CARM1) is less enthalpically favourable but more entropically favourable than the binding of SAH. For **UM305**, **AA175** and **AA183**, the binding enthalpies are more negative (i.e. more favourable) for CARM1 *vs*

PRMT1, while the –TΔS terms are more positive (i.e. less favourable). This suggests that the selectivity of this inhibitor series for CARM1 over PRMT1 is enthalpic in origin.

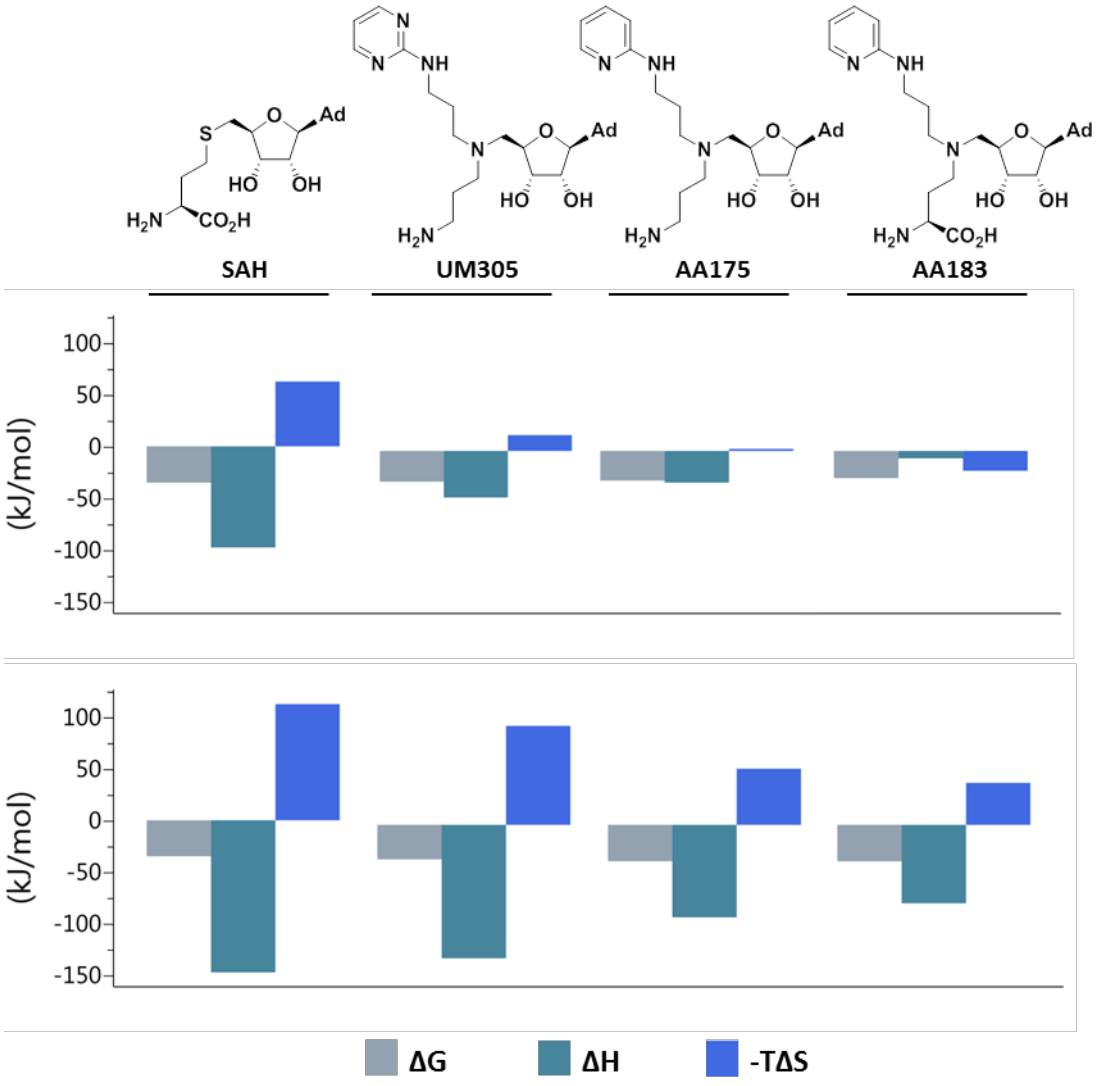


Figure 66: Signature plots for SAH, UM305, AA175 and AA183 binding to PRMT1 (upper) and CARM1 (lower). The inhibitor to which each dataset corresponds is shown above the graph. The data are given numerically in Table 24 and Table 25, below.

Inhibitor	PRMT1			
	ΔG (kJ/mol)	ΔH ± standard error (kJ/mol)	-TΔS (kJ/mol)	N ± standard error
SAH	-35.1	-97.7 ± 1.4	62.6	0.485 ± 0.0037
UM305	-28.5	-43.2 ± 10.5	14.7	0.322 ± 0.028
AA175	-27.5	-29.4 ± 1.16	39.1	0.716 ± 0.012
AA183	-24.9	-6.71*	-18.2	N/A*

*Data were determine using a competition experiment, therefore the ΔH error and binding stoichiometry are not calculated.

Table 24: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into PRMT1. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

Inhibitor	CARM1			
	ΔG (kJ/mol)	ΔH ± standard error (kJ/mol)	-TΔS (kJ/mol)	N± standard error
SAH	-34.9	-147 ± 4.62	113	0.549 ± 0.0086
UM305	-32.2	-124 ± 1.86	92.2	0.529 ± 0.0042
AA175	-34	-86.2 ± 7.42	52.3	0.529 ± 0.025
AA183	-34	-73.1 ± 1.92	39.1	0.536 ± 0.0079

Table 25: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into CARM1. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

3.7.2. Crystal Structures of CARM1 with Aromatic Inhibitors

To explore the structural origins of this selectivity, the aromatic inhibitors were co-crystallised with CARM1, diffraction data collected, and the structures solved by molecular replacement. As for the amine-guanidine inhibitor series, the space groups were $P2_12_12$, and there were 4 monomers in the asymmetric units. Clear electron density corresponding to the inhibitors was present in the active sites (Figure 69, page 169). An additional issue that required addressing with this inhibitor series was the generation of suitable ligand restraints.[110] First, the pK_a plugin within Marvin Sketch (Chem Axon) was used to predict pK_a (or pK_{aH}) values for the aminopyridine and aminopyrimidine nitrogen atoms (Figure 67, below). It is unclear whether the tertiary nitrogen is protonated at the pH at which crystallisations were performed, and it was left unprotonated. The primary amine was predicted to be protonated (pK_{aH}>pH). The aminopyridine and aminopyrimidine nitrogens were predicted to be neutral (pK_{aH}<pH).

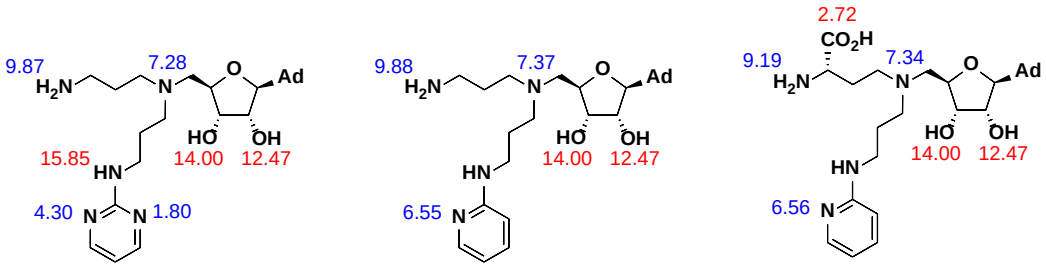


Figure 67: Predicted pK_a (red) and pK_{aH} (blue) values for protonation sites within inhibitors UM305, AA175 and AA183. Values were calculated using the pK_a plugin within Marvin Sketch (Chem Axon). Only the pK_as which fall between -2 and 16 are shown.

Next, eLBOW was used to generate ligand restraints, using SMILES strings generated from the ChemDraw sketches with the predicted protonation states (Figure 67). The restraints generated by eLBOW for the aminopyridine and aminopyrimidine groups were not stereochemically sensible (Figure 68, page 167). They were therefore edited manually to restrain the atoms comprising the pyridine/aminopyridine rings, the *exo* nitrogen, and its substituents to a single plane. This was found to be in reasonable agreement with related structures in the Cambridge Structural Database (e.g. CSD entries CITPUR, EMUFAV, GUXREX, CAGSUB, CAGTAI, EBUGAL, EBUGIT) and PDB (e.g. PDB entries 1ZGY, 5JIO). Restraints generated using the GRADE server (Global Phasing) also included a plane comprising these atoms, lending support to this decision (Figure 68, below).

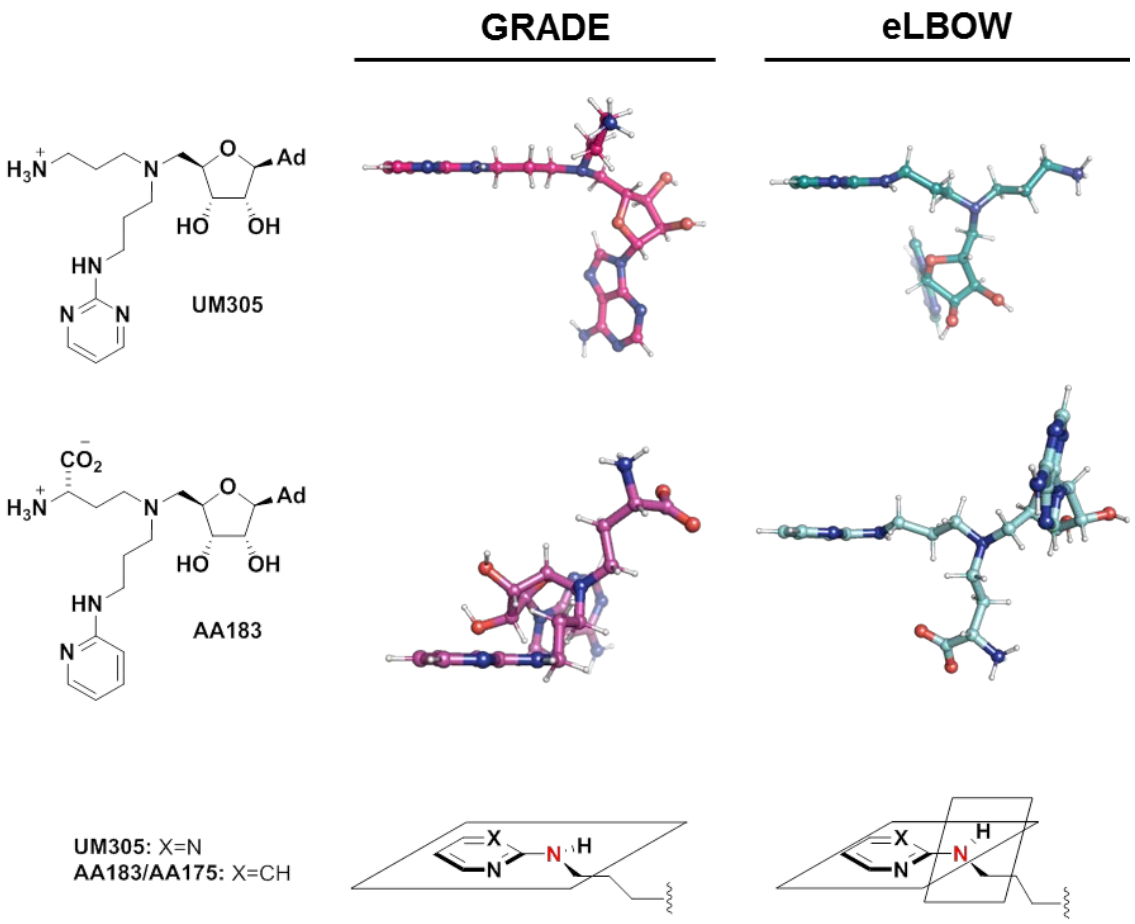


Figure 68: Comparison of optimal ligand geometries output by GRADE (pink) and eLBOW (cyan). Input structures are shown on the left (which were converted to SMILES prior to inputting into each restraints generator). GRADE confines the *exo* nitrogen substituents to the same plane as the ring atoms whereas eLBOW models them as separate planes that are rotated about the C-N bond.

Both eLBOW and GRADE use the CSD as their primary source of restraint information,[111] therefore it is unclear why the optimal output geometries generated by the two programs differ so strongly. In the absence of unambiguous electron density, it is important to keep in mind the influence of (potentially unreliable) restraints on the ligand stereochemistry when interpreting the final models for which they are used.

Having determined suitable ligand restraints, the ligand coordinates could be imported into the CARM1 model, and fitted to the active site electron density (Figure 69, below).

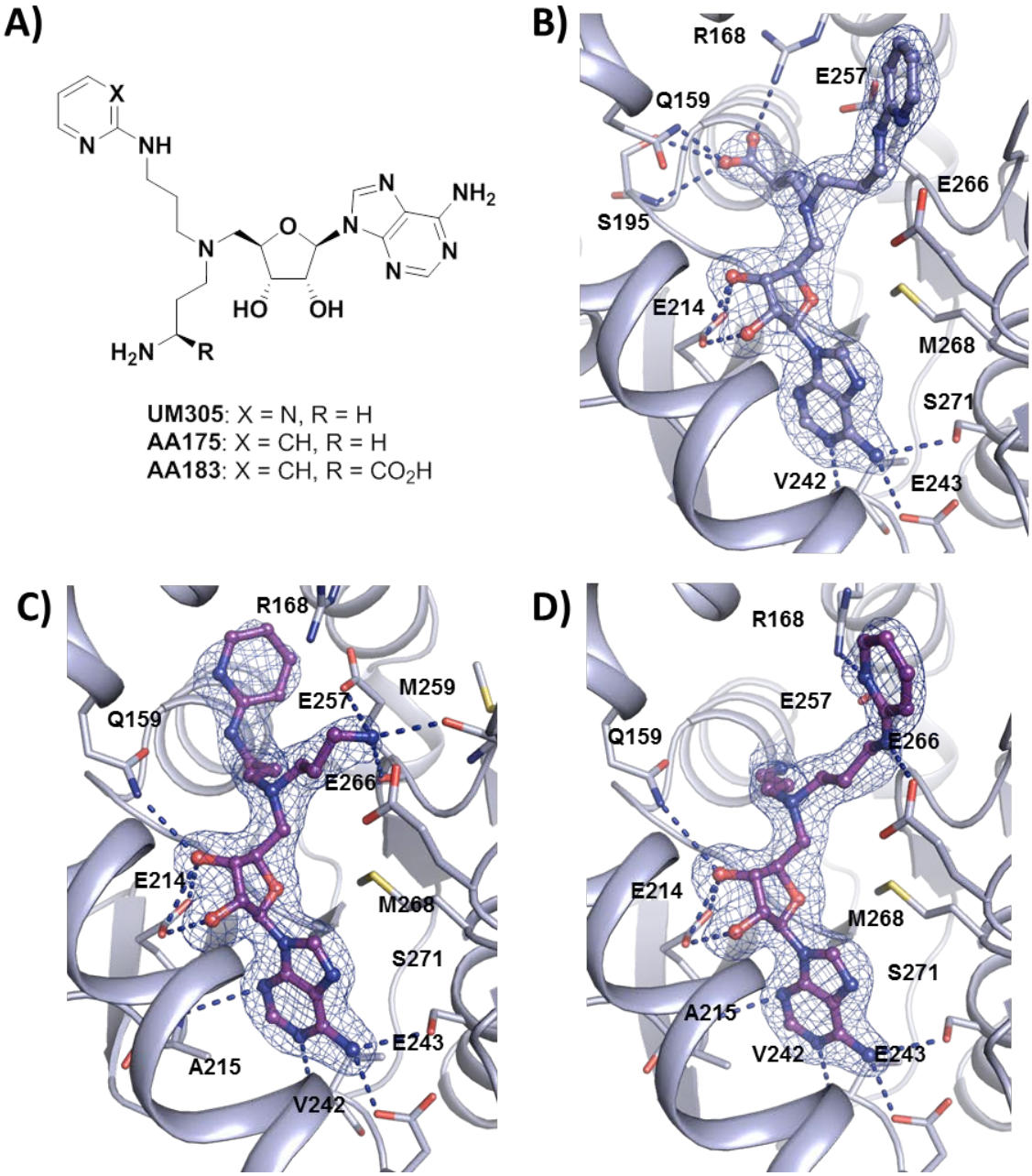


Figure 69: Crystal structures of AA175 and AA183 bound to CARM1. (A) Inhibitor structures; (C)/(D) 2mF_o-DF_c maps for AA175 in two monomers within the crystal structure. Polar contacts found using PyMol are shown as blue dashed lines and residues believed to be involved in SAM and substrate arginine binding (and those involved in the putative polar contacts) are shown as lines; (B) 2mF_o-DF_c map of AA183. Polar contacts and relevant residues are shown as for (C)/(D). All monomers within the asymmetric unit of the CARM1·AA183 structure have similar density, indicating the same ligand conformation as depicted in (E). All maps are contoured at 1 σ.

Intriguingly, the electron density for aminopyrimidine **UM305** (not shown) and aminopyridine **AA175** (Figure 69, above) suggests the presence of two alternate ligand conformations: one in which the aromatic group occupies the substrate binding channel and the amine occupies the SAM amino acid pocket, and one in which these positions are reversed (Figure 70, below and Figure 69C and D, above). Inhibitor **AA183**, which is identical to **AA175** except that it contains an additional amino acid moiety, does not appear to adopt such alternate conformations. This is presumably due to the interaction of this carboxylate group (which mimics that of SAM) with R168, though additional polar contacts with Q159 and S195 may also contribute. The position of the adenosine moiety was the same as seen for the amine-guanidine inhibitors in all cases (Figure 69B, above).

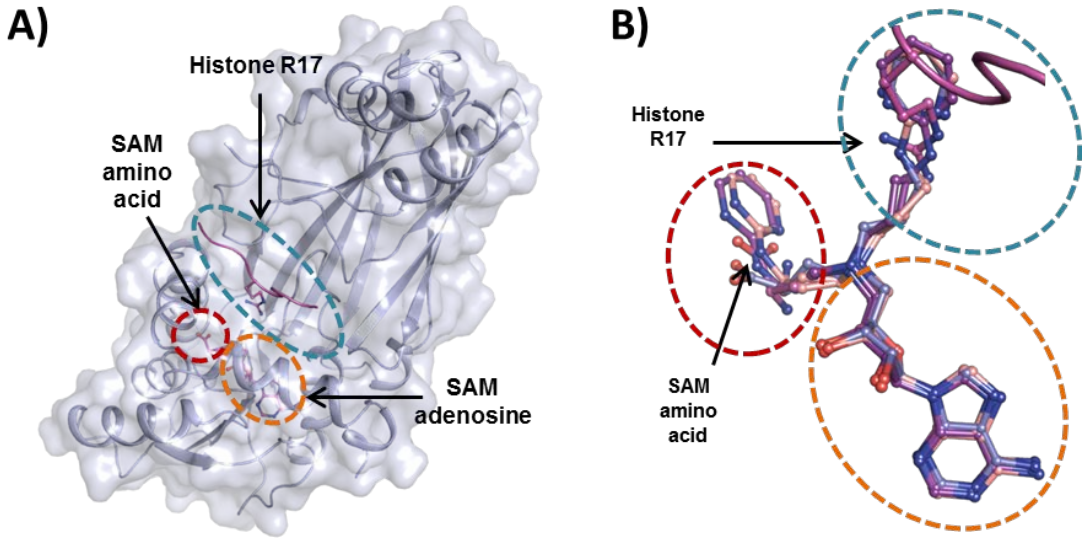


Figure 70: (A) Binding pockets of adenosine (orange dashes), SAM amino acid (red dashes), and substrate binding channel (teal dashes) in the published CARM1·H3·sinefungin crystal structure (PDB code 5DXo);[15a] (B) Superposition of UM305 (pale pink), AA175 (purple) and AA183 (blue) co-crystals with a published structure of CARM1 in complex with histone 3 and sinefungin (magenta, PDB code 5DXo)

Superposition of the CARM1·**AA183** co-crystal structure with a published crystal structure of PRMT1 (Figure 71, below) revealed two residues in the substrate binding channel which may contribute to the CARM1-selectivity of **AA183**. The first residue, N161, is a glutamate (E55) in PRMT1, which may hinder binding of **AA183** through unfavourable interaction of the charged side chain with the hydrophobic pyridine group of the inhibitor. The second, N265, is a tyrosine (Y160) in PRMT1. We hypothesized that Y160 may alter the orientation of the neighbouring conserved glutamate (E161), which may influence its interaction with the nearby **AA183** pyridine.

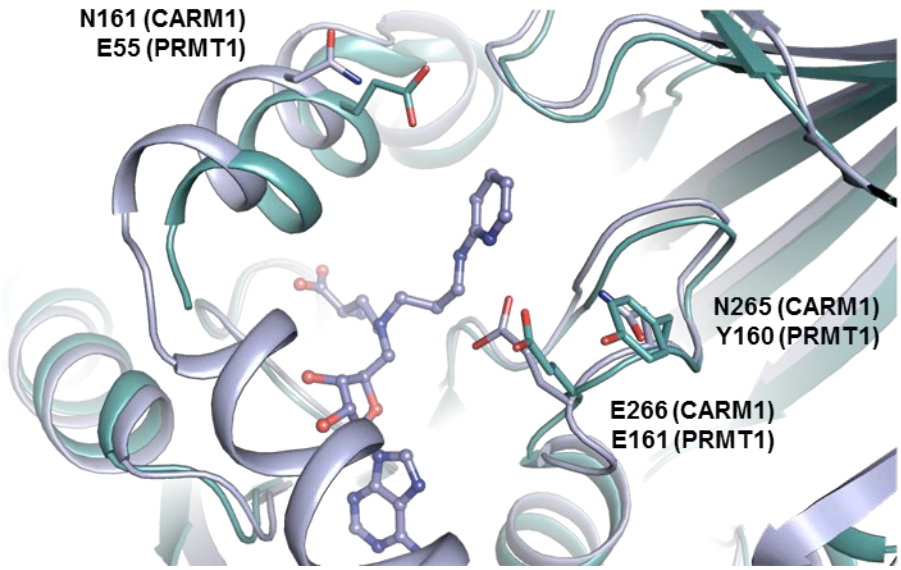


Figure 71: Superposition of the CARM1·AA183 crystal structure (grey) with PRMT1 structure 1or8 (cyan). Side chains of residues that may contribute to the CARM1-selectivity of AA183 are shown as lines.

3.7.3. Mutagenesis Studies

To examine the influence of these residues (Figure 71, above) on inhibitor binding, single site mutants - CARM1(N161E) and CARM1(N265Y) - were designed in which these CARM1 residues were changed to their corresponding residues in PRMT1. A complementary set of PRMT1 mutants - PRMT1(E55N) and PRMT1 (Y160N) - were also designed. These mutants were cloned (using site directed mutagenesis) then expressed and purified as for the wild type proteins. All mutants were obtained in high purity, and with similar overall yields to their corresponding wild type proteins (Figure 72, below), though the PRMT1 mutants were typically more prone to precipitation during purification and subsequent ITC (when compared to the wild type construct). The effect of the mutations on inhibitor binding was assessed using ITC and the results are given in Figure 73, Figure 74, Figure 75 and Table 26.

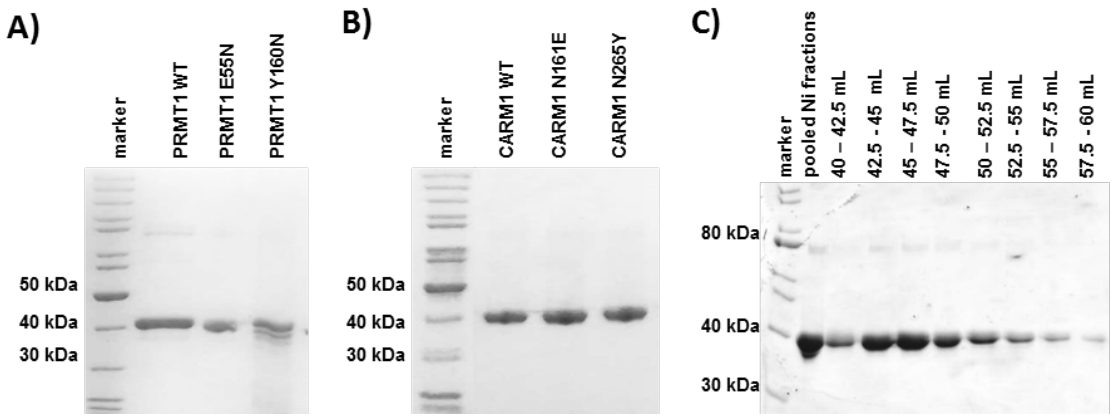


Figure 72: SDS-PAGE of single site mutants of PRMT1 (A) and CARM1 (B). The mutant proteins are the same molecular weight as the corresponding wild type proteins. Minor degradation of the PRMT1 Y160N sample during long-term storage was observed. A gel run immediately following gel filtration is shown in (C), indicating comparable purity to the other constructs immediately following purification.

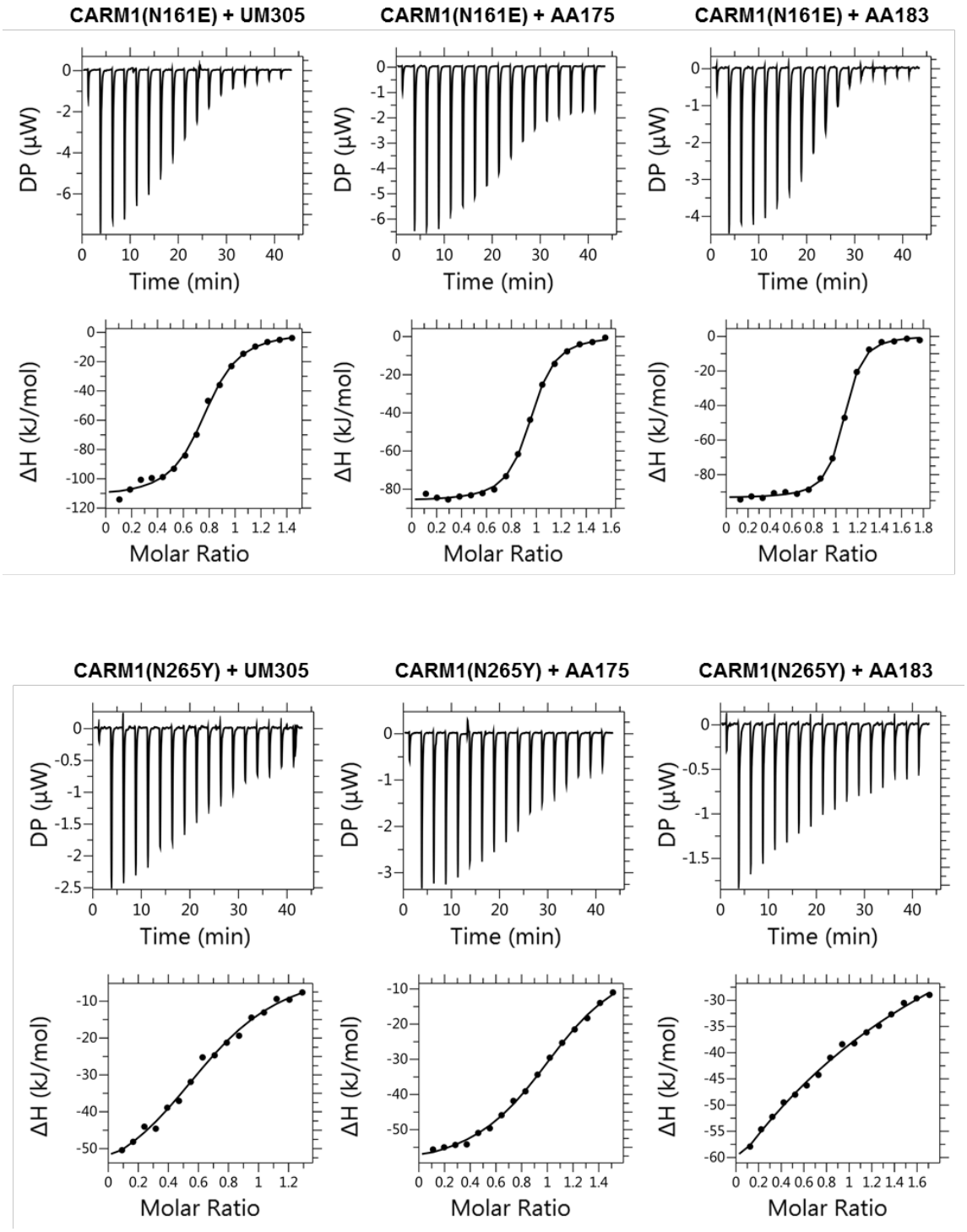


Figure 73: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for UM305, AA175 and AA183 with CARM(N161E) and CARM(N265Y). K_d values determined by curve fitting are given in Table 26 (below).

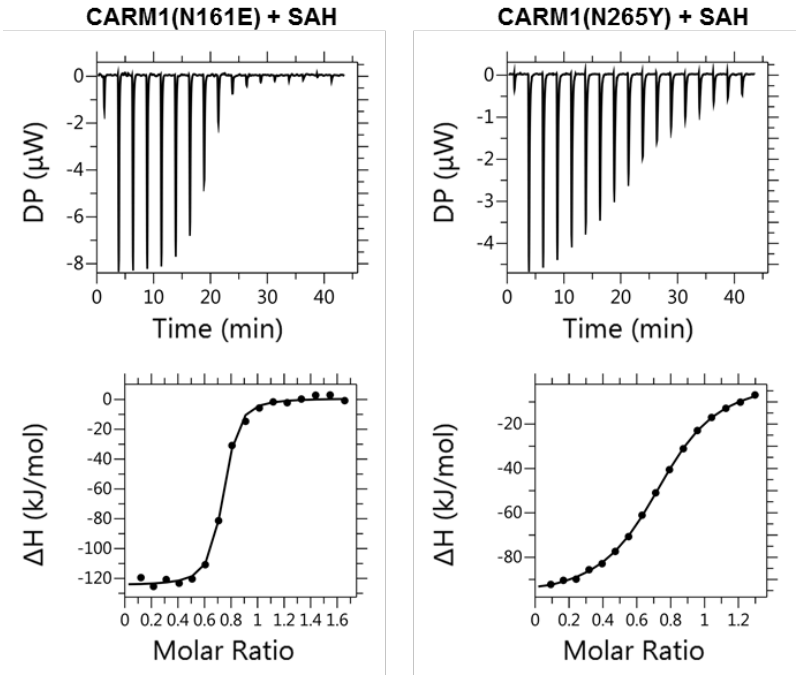


Figure 74: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for SAH with CARM1(N161E) and CARM1(N265Y). K_d values determined by curve fitting are given in Table 26 (below).

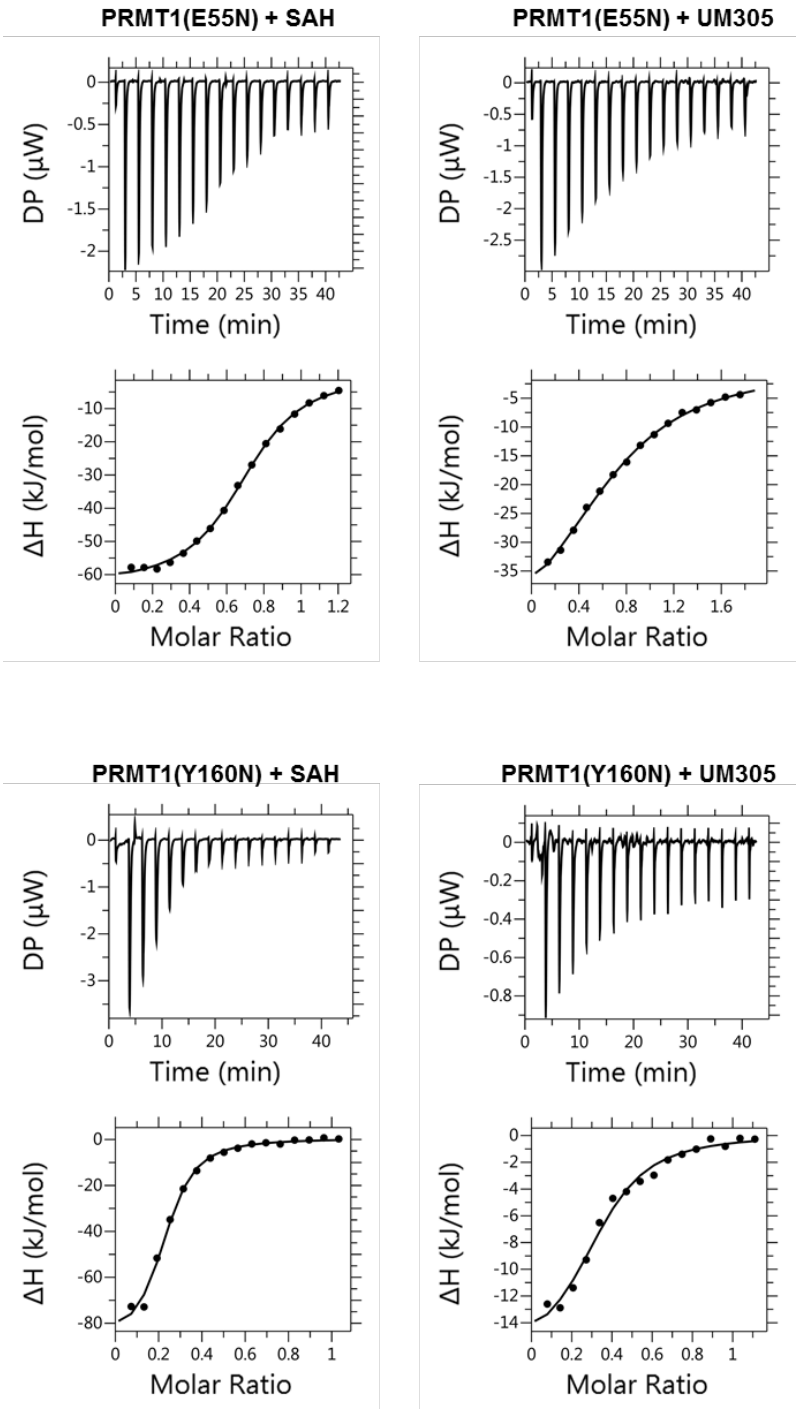


Figure 75: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for SAH and UM305 with PRMT1(E55N) and PRMT1(Y160N). K_d values determined by curve fitting are given in Table 26 (below).

Protein	Mutation	K_d ± standard error from curve fitting (μM)			
		UM305	**AA175**	**AA183**	**SAH**
CARM1	WT	2.30 ± 0.162	1.14 ± 0.411	1.11 ± 0.132	0.774 ± 0.111
	N161E	1.11 ± 0.201	0.354 ± 0.034	0.181 ± 0.0168	0.133 ± 0.022
	N265Y	5.18 ± 1.86	3.26 ± 0.458	**143 ± 75.9**[1]	1.49 ± 0.0628
PRMT1	WT	10.2 ± 4.21[1]	15.3 ± 1.57[3]	No binding[2,3]	0.709 ± 0.057
	E55N	15.8 ± 2.72	-	**-**	1.09 ± 0.106
	Y160N	(15.5 ± 6.98)[4]	13.7 ± 0.351	**-**	2.99 ± 0.245

[1]Stoichiometry fixed to 1
[2]No binding detected using a standard titration; however a competition experiment with SAH determined the K_d to be 43.7 ± 14.9 μM
[3]Titration performed by Alaa Al-Noori
[4]Noisy baseline therefore may be unreliable.

Table 26: K_d values for aminopyridine/aminopyrimidine inhibitors versus PRMT1 and CARM1 single site mutants measured using ITC.

The CARM1 N161E mutation resulted in a uniform increase in inhibitor affinity (including SAH) and is therefore unlikely to be responsible for the CARM1-selectivity of **AA183**. Notably, **AA183** showed significantly weaker binding to CARM1 N265Y *versus* the wild type protein. This suggests that the difference in this residue between CARM1 and PRMT1 contributes to the observed CARM1-selectivity of **AA183**. The binding of **AA175** and **UM305** to CARM1 is also weakened by the N265Y mutation, but to a lesser extent than **AA183**. This may be due to the ability of **UM305** and **AA175** to adopt alternate conformations in which the hydrophobic aminopyrimidine/aminopyridine groups are positioned in the SAM amino acid binding pocket (Figure 78, page 182), thereby avoiding any unfavourable interactions with the substrate binding channel.

The data obtained with the PRMT1 mutants show that the PRMT1 mutations do not significantly affect inhibitor binding. Both titrations (SAH and **UM305**) with PRMT1(Y160N) may have resulted in inaccurate K_ds due to a noisy baseline surrounding the first injection (Figure 75, page 175).

Signature plots for this set of ITC titrations (Figure 76 and Figure 77, below) show that the increased affinity of all inhibitors for CARM1 N161E is entropic in origin, with the exception of **AA183**. The low sigmoidicity of the CARM1 N265Y/**AA183** curve was not sufficient for reliable determination of enthalpy and entropy of binding, therefore it is not included in the figure. The weakened binding of **AA175** and **UM305** to CARM1 N265Y is enthalpic in origin.

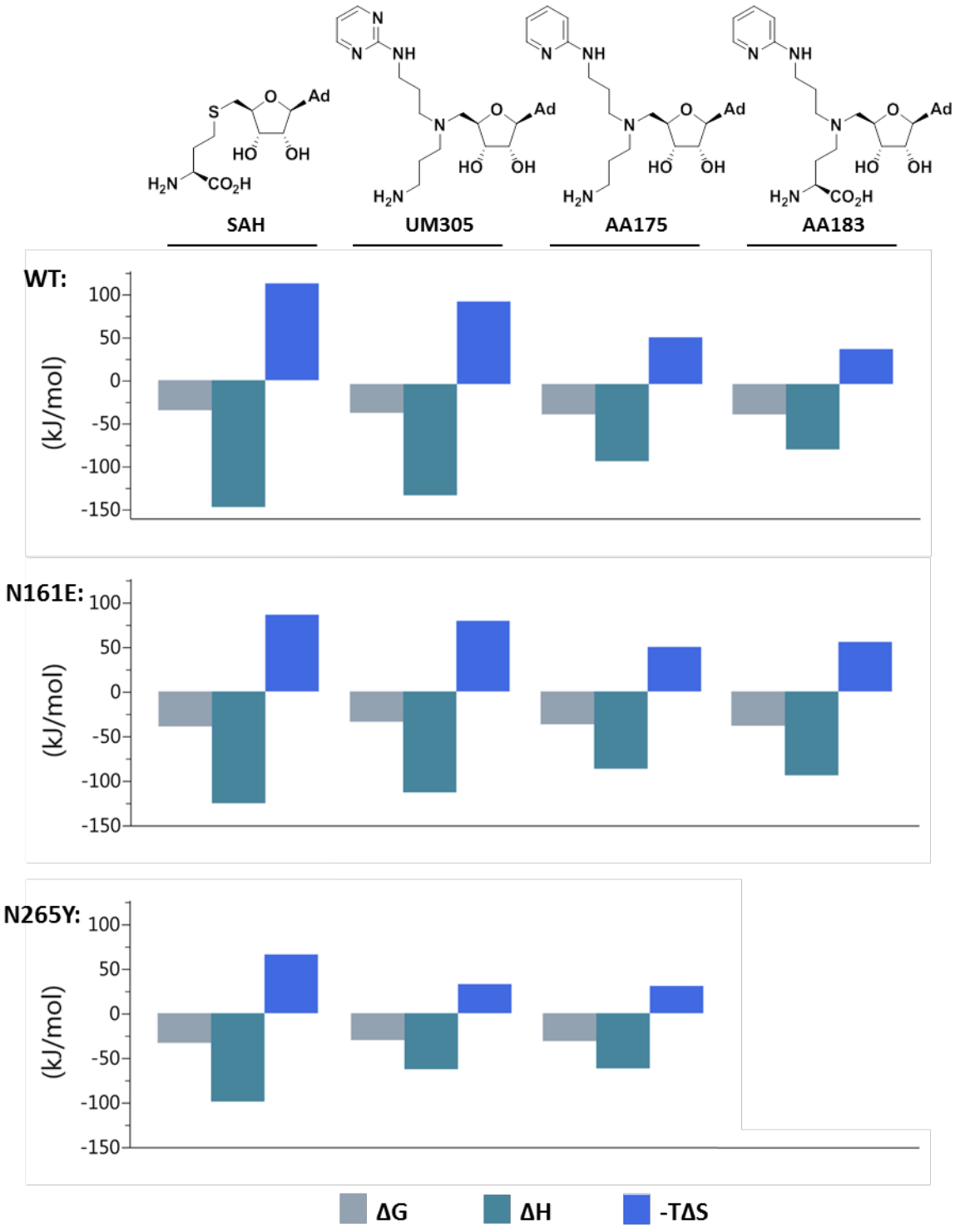


Figure 76: Signature plots for SAH, UM305, AA175 and AA183 binding to CARM1 and single site mutants thereof. The inhibitor to which each dataset corresponds is shown above the graph. The relevant mutation is shown on the left. The data are given numerically in Table 27, below.

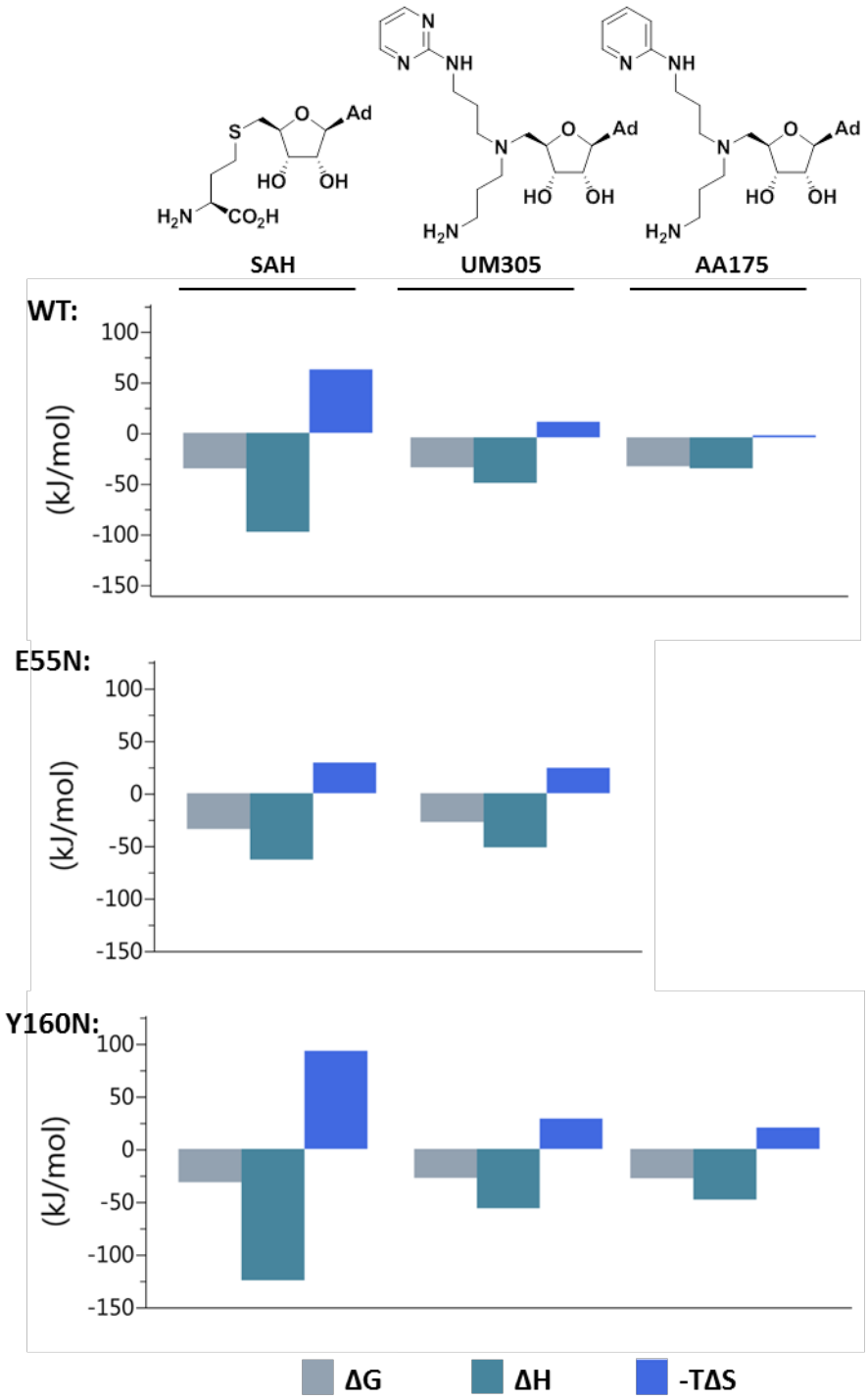


Figure 77: Signature plots for SAH, UM305, AA175 and AA183 binding to PRMT1 and single site mutants thereof. The inhibitor to which each dataset corresponds is shown above the graph. The relevant mutation is shown on the left. The data are given numerically in Table 28, below.

	Inhibitor	CARM1			
		ΔG (kJ/mol)	ΔH ± standard error (kJ/mol)	-TΔS (kJ/mol)	N± standard error
WT	**SAH**	-34.9	-147 ± 4.62	113	0.549 ± 0.0086
	UM305	-32.2	-124 ± 1.86	92.2	0.529 ± 0.0042
	AA175	-34	-86.2 ± 7.42	52.3	0.529 ± 0.025
	AA183	-34	-73.1 ± 1.92	39.1	0.536 ± 0.0079
N161E	**SAH**	-39.3	-125 ± 1.84	86.1	0.692 ± 0.0047
	UM305	-34	-113 ± 3.79	79.2	0.745 ± 0.012
	AA175	-36.8	-86.7 ± 1.17	49.9	0.914 ± 0.005
	AA183	-38.5	-94.1 ± 0.971	55.6	1.02 ± 0.004
N265Y	**SAH**	-33.3	-99.1 ± 1.05	65.9	0.737 ± 0.004
	UM305	-30.2	-62.8 ± 8.41	32.7	0.698 ± 0.051
	AA175	-31.3	-61.9 ± 3.36	30.6	1.07 ± 0.035
	AA183	-	-	-	-

Table 27: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into CARM1 and mutants thereof. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

	Inhibitor	PRMT1			
		ΔG (kJ/mol)	ΔH ± standard error (kJ/mol)	-TΔS (kJ/mol)	N ± standard error
WT	**SAH**	-35.1	-97.7 ± 1.4	62.6	0.485 ± 0.0037
	UM305	-28.5	-43.2 ± 10.5	14.7	0.322 ± 0.028
	AA175	-27.5	-29.4 ± 1.16	39.1	0.716 ± 0.012
	AA183	-24.9	-6.71*	-18.2	N/A*
E55N	**SAH**	-34.1	-63.2 ± 1.54	29.2	0.696 ± 0.0087
	UM305	-27.4	-51.6 ± 4.23	24.1	0.694 ± 0.018
Y160N	**SAH**	-31.6	-125 ± 5.73	93.1	0.18 ± 0.005
	UM305	-27.5	-56.3 ± 33.1	28.8	0.168 ± 0.066
	AA175	-27.8	-48.1 ± 0.38	20.3	0.448 ± 0.0019

*Data were determined using a competition experiment, therefore the ΔH error and binding stoichiometry are not calculated.

Table 28: ΔG, ΔH, -TΔS and stoichiometry (N) values determined from titration of SAH, UM305, AA175 and AA183 into PRMT1. Standard errors are not calculated for ΔG and -TΔS since they are calculated indirectly using the best-fit values of K_d and ΔH.

3.7.3.1. Crystal Structure of CARM1 Mutants in Complex with Aromatic inhibitors

To explore the proposed hypothesis (see above) that the carboxylate group of **AA183** forces the aminopyridine to make unfavourable interactions with the mutated substrate binding channel (and that **AA175** can avoid these interactions by adopting an alternate conformation) **AA183** and **AA175** were co-crystallised with CARM1(N265Y), and the structures solved. The CARM1(N266Y)·**AA183** crystal structure revealed that mutation N265Y induces a

conformational change in E266, in which the glutamate twists so that the plane of the carboxylate group is almost parallel to the phenol side chain of Y265 (Figure 78A, below). This change was consistent in all four CARM1(N265Y) monomers within the asymmetric unit (Figure 78B, below).

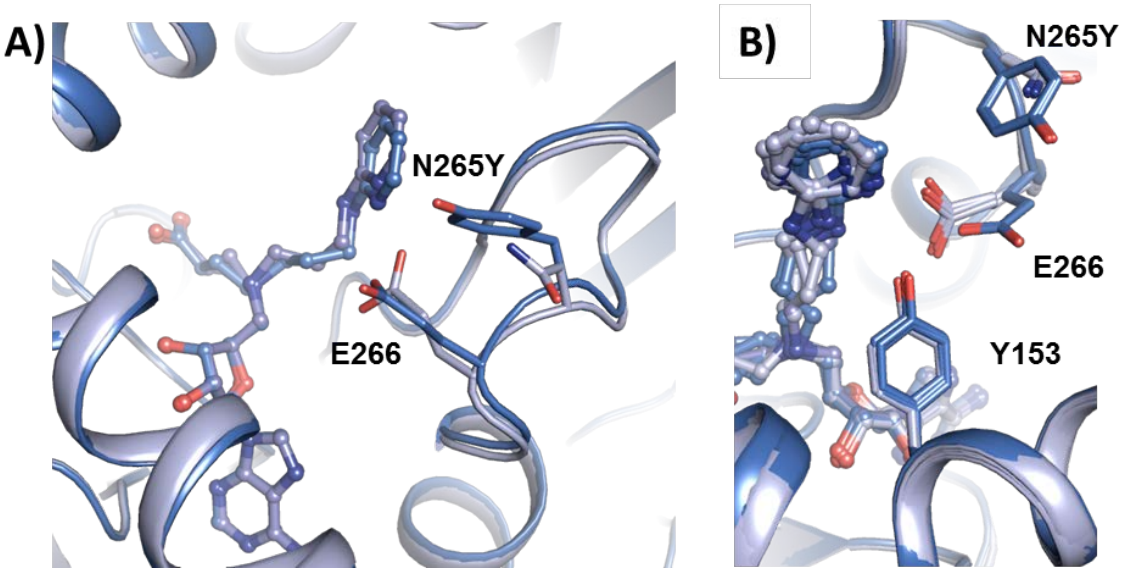


Figure 78: Crystal structures of CARM1(N265Y) in complex with AA183. (A) Superposition of CARM1·AA183 co-crystal structures with (blue) and without (white) the N265Y mutation; (B) Superposition of every monomer within the asymmetric unit of CARM1·AA183 (white) and CARM1 (N265Y)·AA183 (blue)

While this structural change is clear, it is not obvious whether the change in the glutamate position is responsible for the observed binding effects (Table 26, page 176). Since the aminopyridine and aminopyrimidine nitrogens are not expected to be protonated, it seems unlikely that a subtle change in the position of the E265 side-chain strongly affect the binding. It is possible that while the N265Y mutation alters the glutamate position, it is the change in side chain at position 265 that directly influences inhibitor binding (or the

combination of these two effects). Regardless of the exact mechanism which causes the CARM1-selectivity of the aromatic inhibitors, knowledge that this mutation affects binding could be useful for predicting how strongly this inhibitor series might bind other PRMTs.

The structure of CARM1(N265Y) in complex with **AA175** was less easy to interpret. The inhibitor electron density was less well defined, making modelling of the inhibitor difficult. In two of the four active sites within the asymmetric unit, absence of electron density for the propylamine and propylaminopyridine moieties meant that only adenosine (**19**, Figure 79B, below) could be modelled. In the other two active sites, the inhibitor appeared to be positioned with the aminopyridine occupying the substrate binding channel (Figure 79A, below).

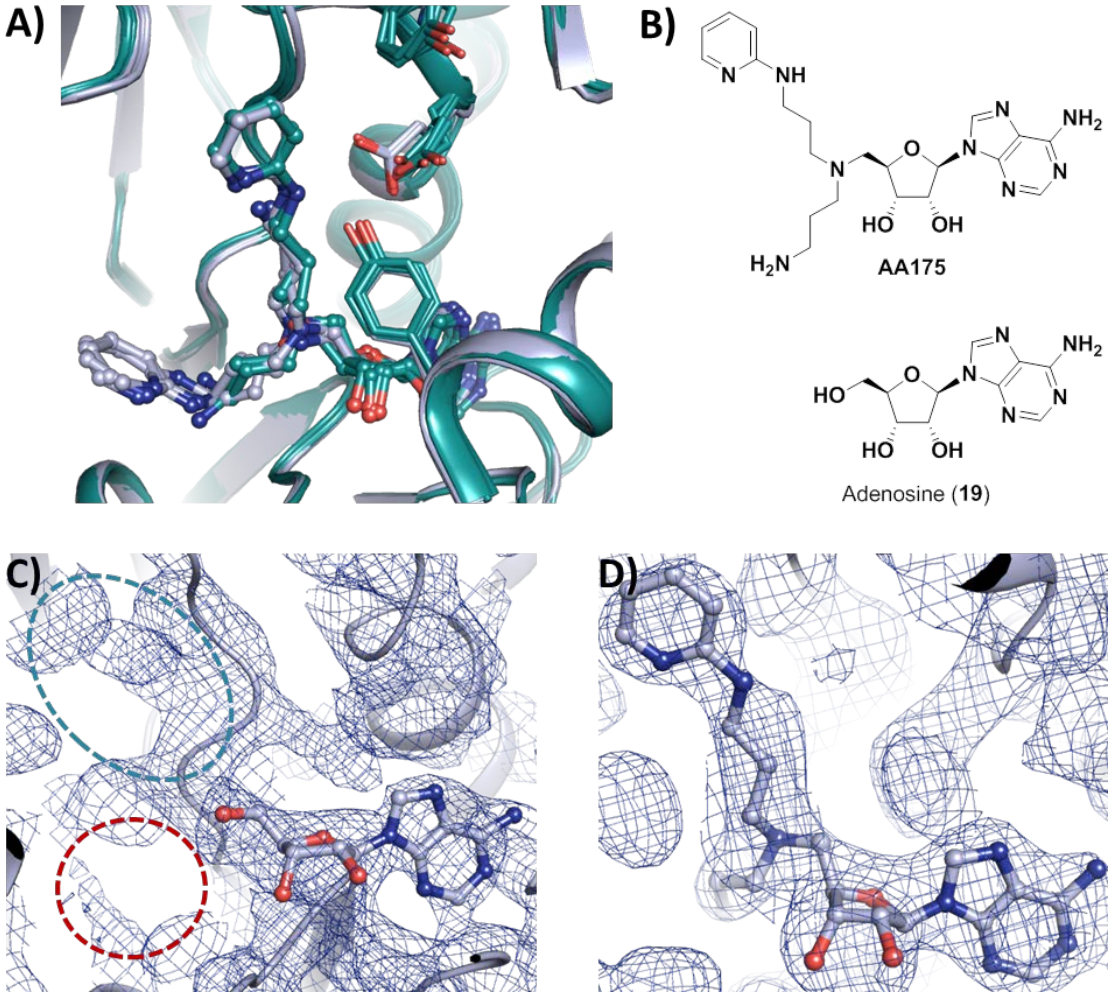


Figure 79: Crystal structure of CARM1(N265Y) in complex with AA175. (A) Superposition of every monomer within the asymmetric unit of CARM1·AA175 (white) and CARM1 (N265Y)·AA175 (teal); (B) Structure of AA175 and adenosine; (C) $2mF_o$-DF_c map of monomer in which there is no electron density for the propylamine and propyl aminopyridine groups of AA175. The substrate and SAM carboxylate binding pockets are shown with turquoise and red dashes, respectively; (D) $2mF_o$-DF_c map of monomer in which density is visible for all of AA175. Maps are contoured at 1σ and the carve radius around the ligand is 10 Å.

This is in direct disagreement with our previously proposed hypothesis (see above) that the aminopyridine group of AA175 avoids unfavourable interactions with the mutated substrate binding channel by adopting an alternate

conformation in which the pyridine group occupies the SAM carboxylate binding channel (Figure 70, page 170 and Figure 71, page 171).

3.7.4. Aromatic Inhibitors: Conclusions

Overall, the binding studies with wild type PRMT1 and CARM1 show that substituting the guanidine group for an aminopyridine or aminopyrimidine in our bisubstrate inhibitors confers 10 to 40-fold selectivity for CARM1 over PRMT1. Crystal structures show that **UM305** and **AA175** are able to adopt an alternate conformation in which the aminopyrimidine/aminopyridine are positioned in the SAM-amino acid binding pocket while **AA183** adopts a single conformation, with the aminopyridine in the arginine binding channel (Figure 69, page 169). This suggests that the inhibitor carboxylate group may be needed to ensure the arginine mimicking moiety occupies the desired pocket. Despite conferring this apparent conformational preference, the carboxylate group does not appear to enhance the affinity of our series of inhibitors for CARM1, with **AA175** and **AA183** displaying the similar K_d and IC_{50} values for CARM1 (Table 23, page 160). Furthermore, addition of a carboxylate group to 5C guanidine-amine **UM251** (i.e. inhibitor **WH5C**) resulted in a decrease in potency (Table 19, page 136). It is possible that the amine guanidine inhibitors **UM079**, **UM249** and **UM251** exhibit similar alternate conformations to **UM305** and **AA175**, but that they are not as well defined in the electron density due to the less pronounced size difference between the two terminal groups (amine-guanidine versus amine-aminopyridine/aminopyrimidine). In order to exploit features specific to

the CARM1 arginine binding channel, it may therefore be prudent to use an amino acid group (as in **WH5C** and **AA183**), to ensure that the arginine mimicking group exclusively occupies the substrate binding channel. The mutant binding and structural data show that residue N265Y (located in the substrate binding channel) is involved in conferring the CARM1-selectivity of aromatic inhibitors, especially **AA183**. This could aid in predicting the selectivity profiles of **AA183** (and related inhibitors) *versus* the other PRMTs family members.

3.8. **Morpholino Inhibitors**

In addition to the inhibitors discussed so far, Dr. U. Muhsen synthesised a series of related compounds in which the ribose moiety is replaced with a morpholino group (Figure 80, below). This substitution has been shown to be successful in antisense oligonucleotides, for use in a variety of contexts.[112] A selection of these inhibitors was screened for inhibitory activity against PRMT1 and CARM1 using the radiometric methylation assay (Section 2.8.5, page 61) and the results are given in Table 29 (below).

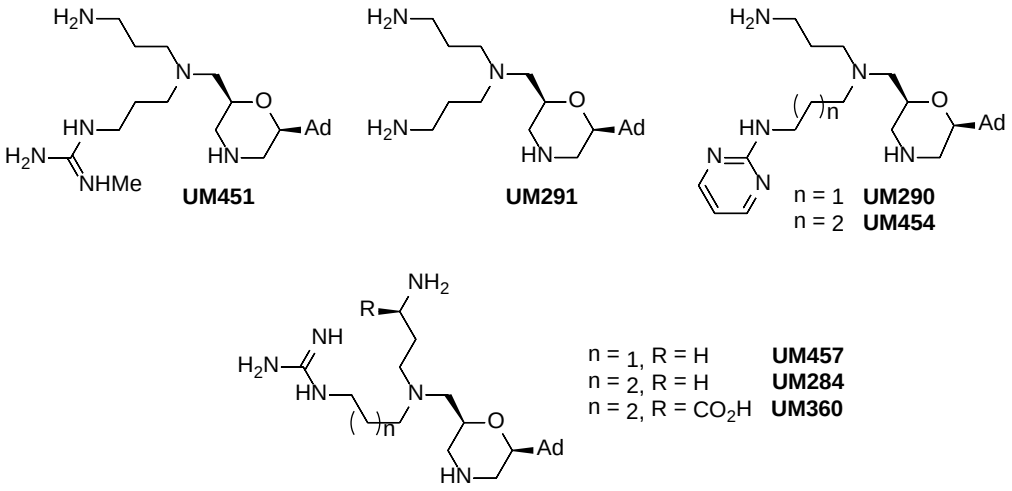


Figure 80: Morpholino inhibitors synthesised by Dr U. Muhsen.

Inhibitor	IC_{50} (μM)	
	PRMT1	CARM1
UM284	>100[1,3]	not tested
UM360	>100	>100
UM457	**<10[2,3]**	>100
UM451	**~10**	>100
UM290	>100	not tested
UM454	**<10**	>100
UM291	>100	>100

[1]>50 % maximal activity remaining in the presence of 100 μM inhibitor
[2]Full IC_{50} assay determined the IC_{50} to be 4.11 μM with 95 % confidence interval 3.80 – 1.97
[3]Two independent experiments (each performed in duplicate)

Table 29: Inhibitor screening for morpholino analogues.

Overall, it is apparent that the ribose-morpholino substitution drastically affected inhibitor binding. All morpholino analogues displayed IC_{50} values greater than 100 μM for CARM1. Only **UM457**, **UM451** and **UM454** had IC_{50} values lower than 100 μM for PRMT1. **UM451** and **UM457** have the same structure, except that **UM451** has a methyl substituent on the guanidinium group. Both contain a guanidine group tethered to the ribose moiety *via* a 3-carbon linker. **UM284** - the 4-carbon analogue of **UM457** – showed significantly reduced inhibition relative to its 3-carbon counterpart. This drastic effect is surprising given the linker length effects (i.e. lack thereof) discussed in Section 3.6 (page 131). The IC_{50} values of **UM290** (3C aminopyrimidine) and **UM454** (4C aminopyrimidine) with PRMT1 display the opposite trend, with **UM454** (4C) inhibiting PRMT1 up to 10-fold more strongly than its 3C counterpart.

To further investigate the binding of the morpholino inhibitors, ITC was used. Titrations of **UM451**, **UM454** and **UM457** into PRMT1 were performed, along with the corresponding ligand-to-buffer control titrations (Figure 81, page 189). Intriguingly, the inhibitors did not appear to bind well to PRMT1. In all cases the affinity was drastically lower than that anticipated based on their measured IC_{50} values (Table 29, page 187). Curve fitting for the inhibitor which showed the highest affinity – **UM457** – determined the K_d to be 642 ± 1020 μM. The wide error margins of curve-fitting are due to the low sigmoidicity of the curve: an inherent limitation of ITC. The heats of dilution for the other two inhibitors (**UM451** and **UM454**) are very large relative to the heats of binding (comparison of ligand-to-buffer and ligand-to-protein thermograms, Figure 81, page 189).

Overall, the data for **UM451** and **UM454** do not provide convincing evidence of specific binding to PRMT1.

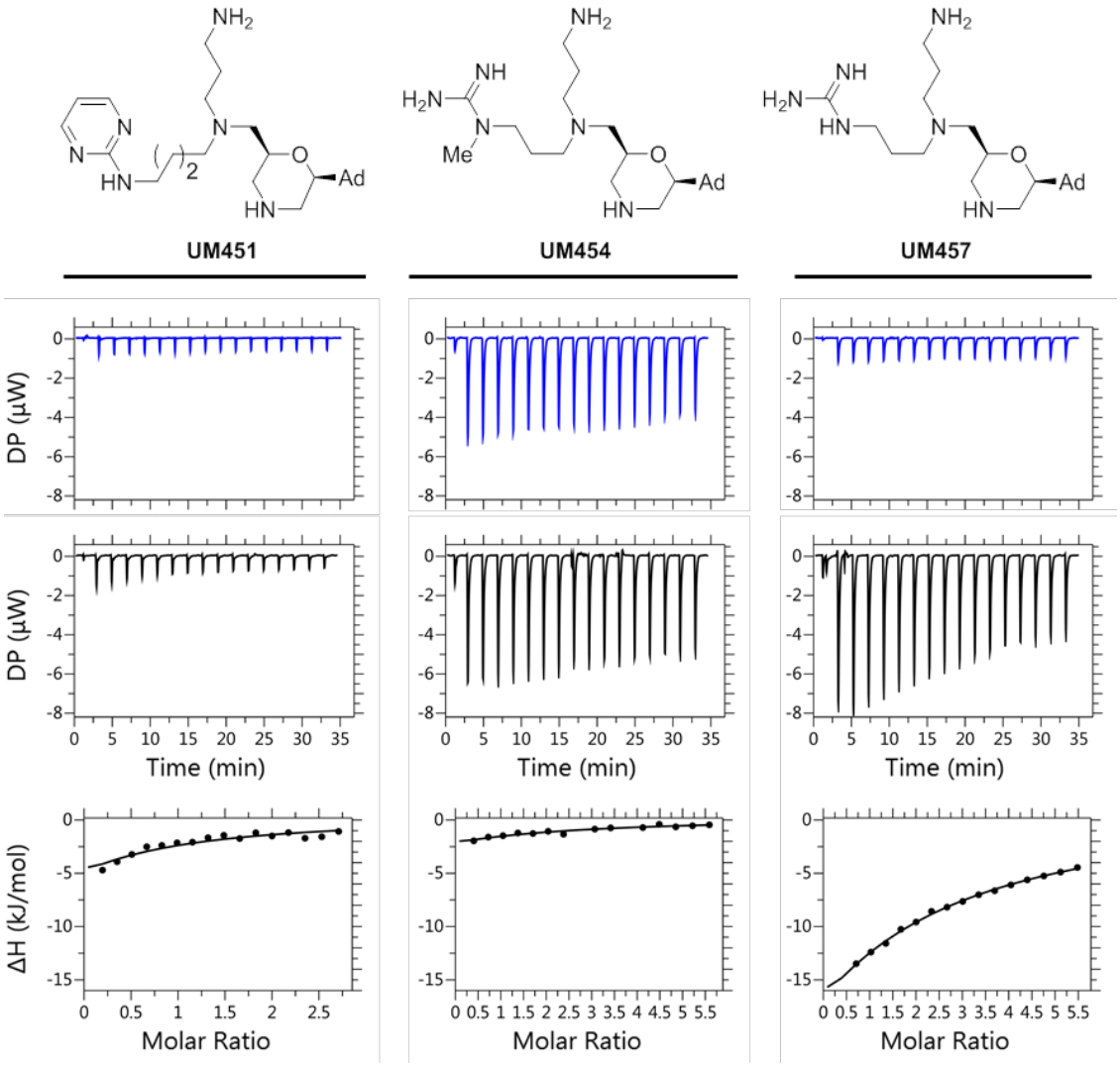


Figure 81: ITC titrations of morpholino inhibitors into PRMT1. Structures of the tested inhibitors are shown with corresponding ITC data underneath. From top to bottom: Ligand-to-buffer thermogram (blue); ligand-to-protein thermogram (black); ΔH values (black circles) with fitted curves (black line).

There are a number of plausible explanations for the apparent discrepancy between ITC and IC_{50} values determined for the morpholino analogues. One possibility is that these inhibitors target PRMT1 at one of the microstates along

its catalytic cycle. It is also possible that the inhibitors bind at a different site to their ribose counterparts. Several PRMT inhibitors have been reported which bind the PRMT more strongly in the presence of SAH.[16b, 42] To test this possibility, a titration of **UM457** into PRMT1 pre-saturated with SAH was performed (data not shown), but no binding was observed. This may indicate that **UM457** binds to the same site as SAH; however the loss of detectable binding could also be attributed to the lower PRMT1 concentration at which this titration was performed (37 μM compared to 87 μM). Nonetheless, this result rules out the possibility that SAH positively influences **UM457** binding. It remains possible that **UM457** binds to PRMT1 at some point along its catalytic cycle, or that it binds non-specifically to PRMT1 in a manner that prevents catalytic turnover.

Since none of the inhibitors appear to bind to CARM1 and attempts at crystallising PRMT1 were unsuccessful, crystallography could not be used to examine the binding mode. Further work is needed to determine whether the observed PRMT1 inhibition of the morpholino compounds is due to a specific binding interaction. The main conclusion from the presented results is that the ribose-to-morpholino substitution drastically alters the binding of our bisubstrate inhibitors.

3.9. Synthesis of putative FRET probes

3.9.1. Design of FRET Probes

Throughout the work reported above, one key drawback was the low throughput nature of the techniques used to assess inhibitor binding and potency. A number of high throughput assays for PRMT inhibitors exist,[113] however many of these are enzyme coupled (with high potential for false positives), require special equipment (e.g. scintillation top-counters) or are prohibitively expensive for academic research. Having shown that our bisubstrate PRMT inhibitors can tolerate a variety of tertiary amine substituents without detriment to PRMT1- and CARM1- binding, such inhibitors have potential as scaffolds for general PRMT probes. Both PRMT1 and CARM1 have several tryptophan residues in close proximity to the substrate binding channel (Figure 82, below).

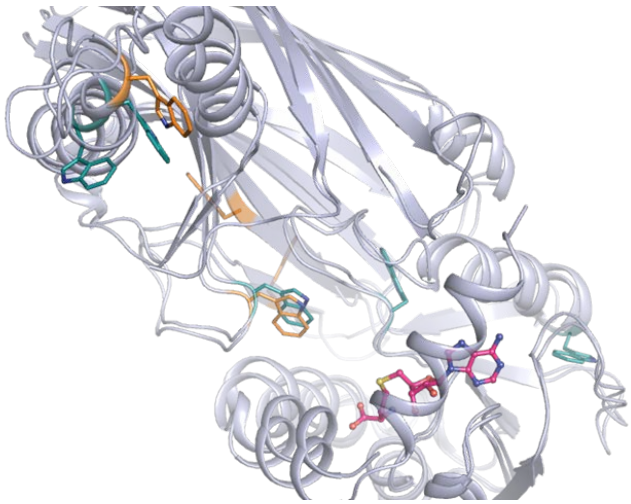
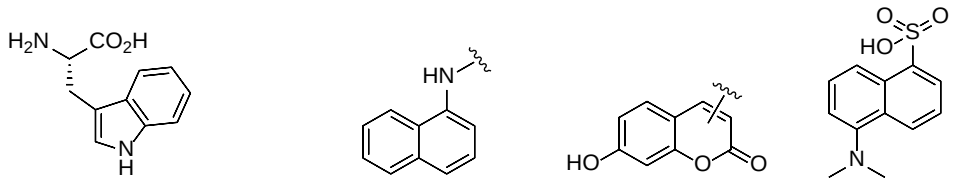


Figure 82: Superposition of PRMT1 (PDB code 1OR8) and CARM1 (PDB code 5DX0) crystal structures showing location of tryptophans. PRMT1 tryptophans are shown in teal and CARM1 in orange. SAH bound in the active site of PRMT1 is shown as magenta ball-and-sticks. Cropped parts of the structures do not contain any tryptophans.

Native tryptophans can be used as FRET donors to determine whether ligands containing suitable FRET acceptors are binding to the target protein.[114] Such 'analytical FRET probes'[115] may then be used in displacement assays to assess the binding of competing, unlabelled inhibitors. Applying this approach to PRMTs would provide a high-throughput means of screening compounds for binding to the active site. Inhibitor titrations could also be used to quantify binding (i.e. K_d values).

Several probes have been shown to be suitable FRET acceptors for tryptophan in the context of enzyme binding studies (Figure 83, below).[115-116]



tryptophan (**20**) naphthylamine (**21**) 7-hydroxy coumarin (**22**) dansyl (**23**)

Figure 83: Structure of tryptophan and examples of tryptophan FRET acceptors.

Putative FRET probes **EG047** and **EG048** were designed. These naphthylamine derivatives were chosen for their anticipated ease of synthesis, following the same synthetic strategy used to generate the adenosine-based bisubstrate inhibitors discussed above (i.e. **UM079**, **UM249** etc., Figure 15, page 27).[68-69] The route to protected propylamino scaffold **25** (Figure 84, page 193) is well established.[67-69, 87] Subsequent reductive aminations using boc-protected

naphthyl aldehydes followed by global deprotection would furnish compounds such as **EG047** and **EG048**, with various linker lengths (Figure 84, below).

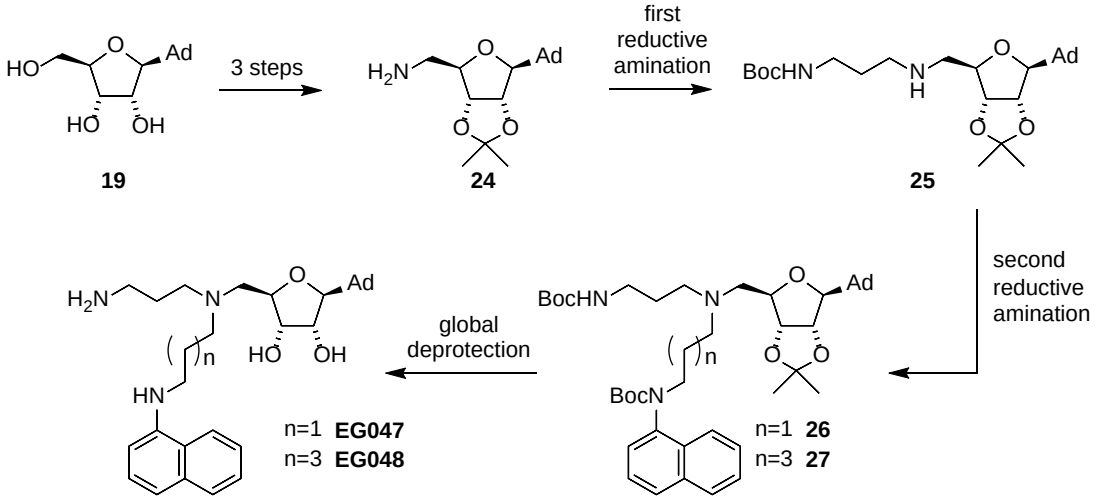


Figure 84: Proposed route to putative FRET probes.

3.9.2. Synthesis of Putative FRET probes

Propylamino scaffold **25** was synthesised from adenosine using previously reported procedures (Figure 85 and Figure 86, page 194) in an overall yield of 34 % over 4 steps.[67] The Mitsunobu reaction and reductive amination steps were low-yielding, but this was not of concern as sub-milligram quantities were sufficient for initial biochemical testing of the putative FRET probes. The boc-protected propylamino aldehyde (**32**) used for reductive amination was made in 2 steps from 3-aminopropan-1-ol (Figure 86, page 194), as described by Dr. U. Muhsen.[67]

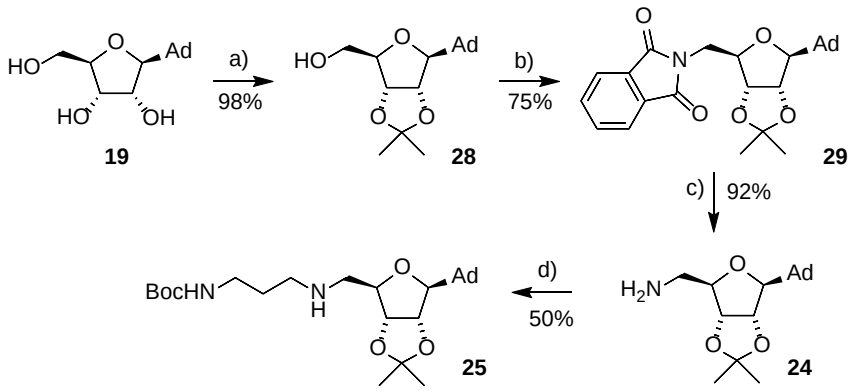


Figure 85: Synthesis of protected propylamino scaffold 25. a) 2,2-dimethoxypropane, acetone, *p*-TsOH, reflux, 16 h; b) phthalimide, PPh3, DIAD, THF, 0 °C to r.t., 16 h; c) $NH_2NH_2{\cdot}xH_2O$, EtOH, reflux, 1 h; d) 32, $(CH_2Cl)_2$, 60 °C to r.t., then $Na(OAc)_3BH$, 16 h.

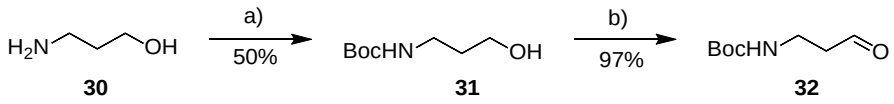


Figure 86: Synthesis of boc-protected amino-aldehyde 32. a) Boc_2O, CH_2Cl_2, r.t., 2 h; b) DMSO, $(COCl)_2$, Et_3N, CH_2Cl_2, -78 °C.

Attention next turned to generating the naphthyl aldehydes required for the second reductive amination (Figure 84, page 193). Initially, two aldehydes were prepared (Figure 87, page 195). One had a 3 carbon linker to the naphthylamine group (**37**) and the other, a 5-carbon linker (**38**). Attempts at direct alkylation of naphthylamine using 3-bromopropanol (according to the procedure of Frade *et al.*)[90] gave the product of dialkylation (**39**), in 85 % yield with respect to 3-bromo-1-propanol. Boc protection of naphthylamine prior to alkylation gave the desired boc-protected naphthyl alcohols **35** and **36**, in 54 and 55 % yields,

respectively. Swern oxidation of these alcohols gave the desired aldehydes **37** and **38**, in 91 and 80 % yields, respectively.

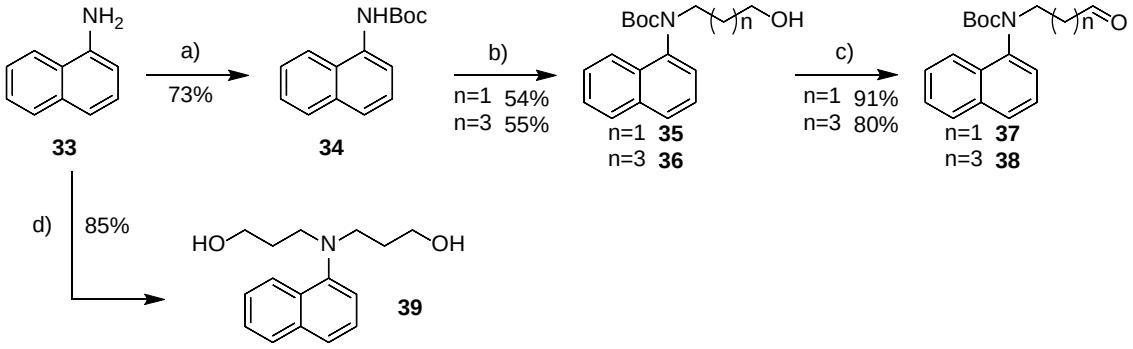


Figure 87: Synthesis of boc-protected naphthyl aldehydes 37 and 38. a) Boc_2O, THF, 60 °C, 48 h; b) NaH, DMF, 0 °C then 3-bromo-1-propanol (n = 1) or 5-bromo-1-pentanol (n = 3), THF, r.t., 3 h; c) DMSO, (COCl)2, Et3N, CH_2Cl_2, -78 °C; d) 3-bromo-1-propanol (0.8 eq), EtOH, reflux, 18 h.

With aldehyde building blocks in hand, installation of the alkylnaphthylamine group onto the propylamino scaffold (**25**) was attempted (Figure 88, page 196). Following reductive amination, broadening of all signals was observed in the ^{1}H-NMR spectrum of the resulting adducts in deuterated chloroform and methanol, hindering NMR assignment. LCMS traces indicated the presence of one major species (>95 % by HPLC-UV), of the correct molecular weight (Appendix 6.6). TFA deprotection was therefore performed in the hope that the final compounds would give interpretable ^{1}H-NMR spectra, enabling unambiguous characterisation.

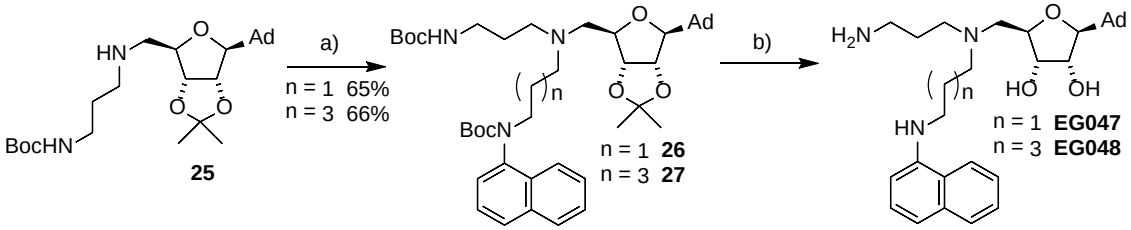


Figure 88: Second reductive amination and subsequent global deprotection. a) 37 or 38, $(CH_2Cl)_2$, 60 °C to r.t., 1 h, then $Na(OAc)_3BH$, 4 h; b) TFA, H_2O, r.t., 6 h.

While it was possible to obtain interpretable NMR spectra for the final compounds (Appendix 6.7), the presence of two doublets at ~6.0 ppm (characteristic of the 1'-CH of the ribose moiety) indicated the presence of structurally related impurity. The ratio of the compound:impurity peak integrals was 1:0.28. A corresponding pair of adenine CH signals in the ^{1}H NMR and splitting of ^{13}C NMR signals of the 1' and 2' ribose carbons suggests that the related species may be epimers of the desired compounds.

A second possibility is that the species are the depurinated analogues of the final compounds. A plausible mechanism for acid-catalysed de-purination and epimerisation is given in Figure 89 (below).

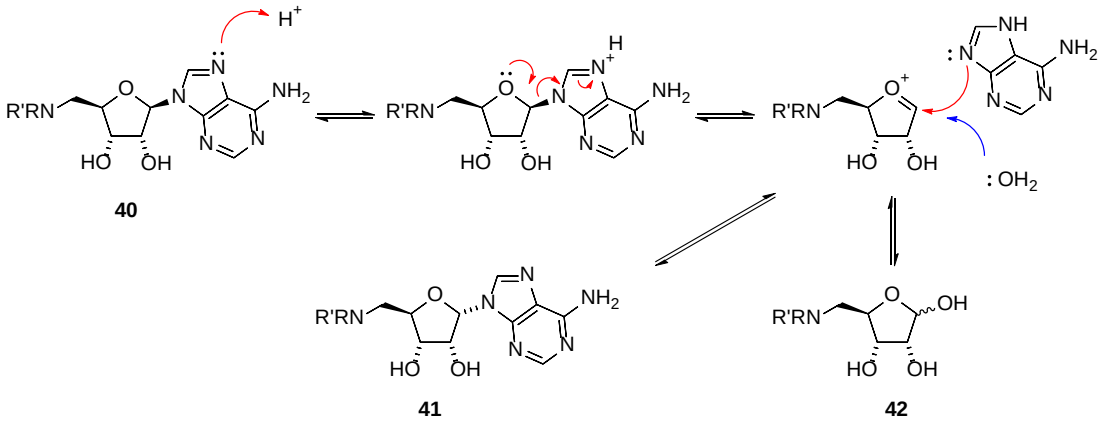


Figure 89: Proposed mechanisms for depurination and racemisation.

A third possibility is that a small amount of **25** was present in the starting material, resulting in the presence of the corresponding deprotected compound in the final product.

Initial binding tests on the crude **EG047** and **EG048** were performed, to determine whether optimising the synthesis was worthwhile. **EG047** and **EG048** were found to inhibit both PRMT1 and CARM1 (Figure 90, below). ITC titrations of **EG047** and **EG048** into CARM1 were performed (Figure 91, below).

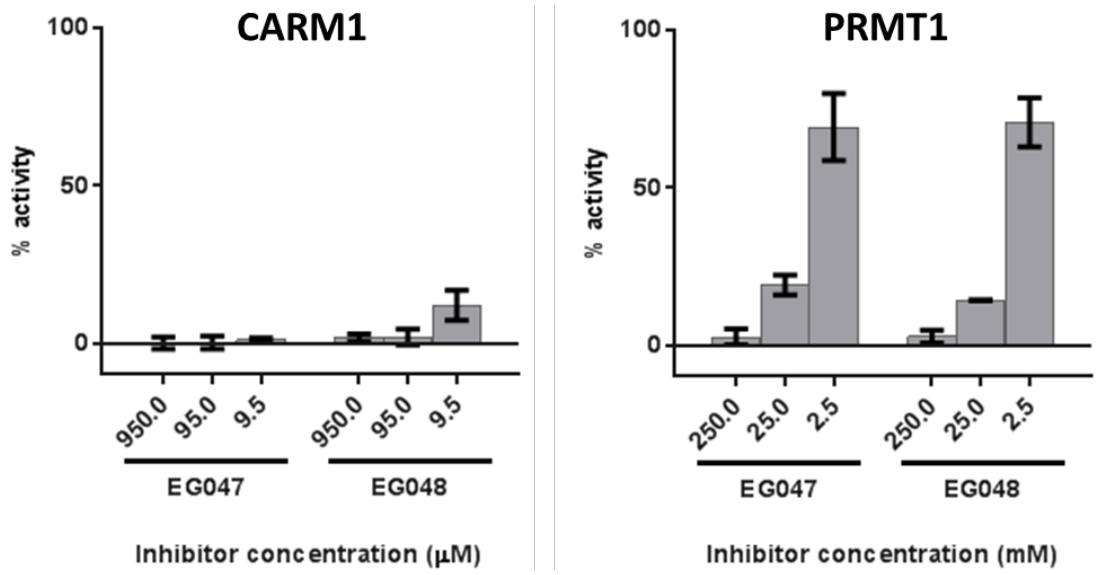


Figure 90: Inhibition screen of EG047 and EG048 vs PRMT1 and CARM1. Both inhibitors have IC_{50} values <10 µM for CARM1 and from 2.5 - 25 µM for PRMT1.

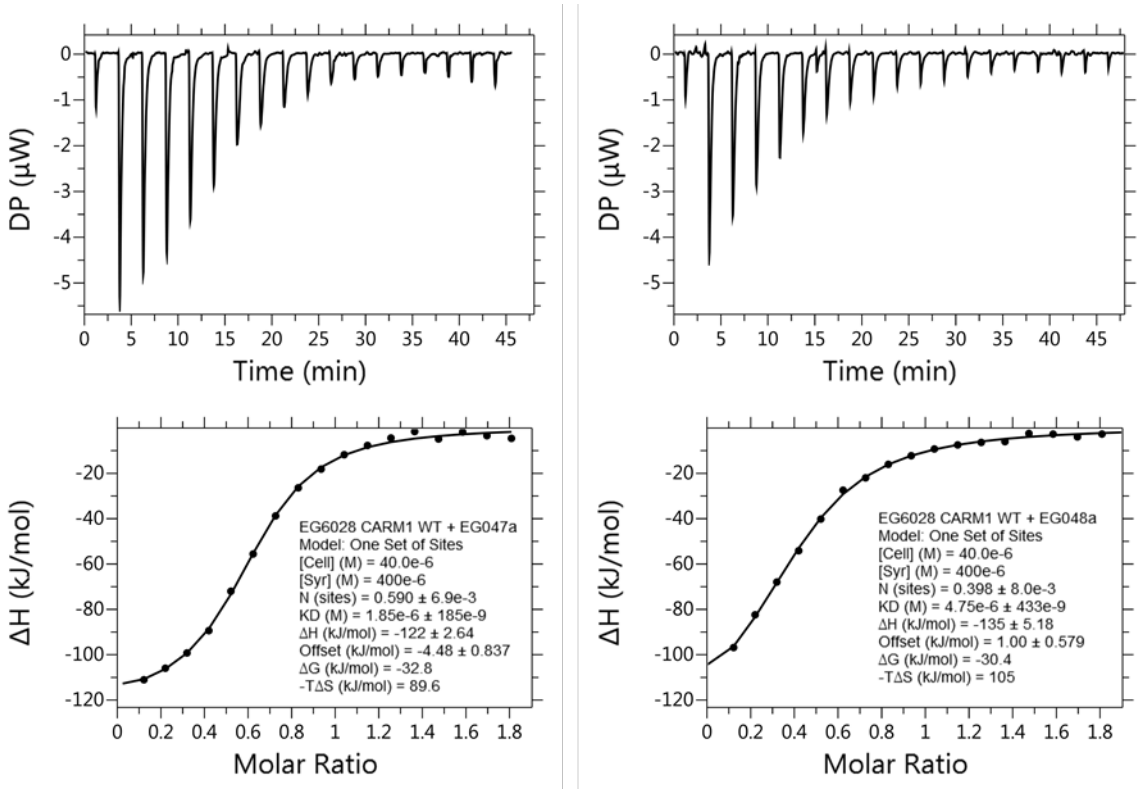


Figure 91: ITC thermograms (upper graphs) and integrated heat plots (lower graphs) with binding curves for EG047 and EG048 with PRMT1 and CARM1. Values from curve fitting and associated errors are given on the graphs.

The determined stoichiometry values are between 0.4 and 0.6, in line with those observed for other inhibitors. This is encouraging, and suggests that the major components of crude **EG047** and **EG048** are responsible for the observed inhibition. It is, however, important to note that the HPLC and LCMS data do not conclusively show that **EG047** and **EG048** are the major components of the tested samples.

Additionally, **EG047** was co-crystallised with CARM1 and the structure solved as for the other inhibitors (Figure 92, page 200). There is clear electron density for the adenosine moiety of the inhibitor, but none for the propylamine or propylnaphthylamine groups. This may be due to inherent flexibility of these moieties, or the presence of a related compound lacking these groups (presumably in the sample of **EG047**) preferentially binding to CARM1. Notably the refined b-factors of the adenosine atoms are comparable to those comprising the surrounding side chains, suggesting that full occupancy of CARM1 has been achieved.

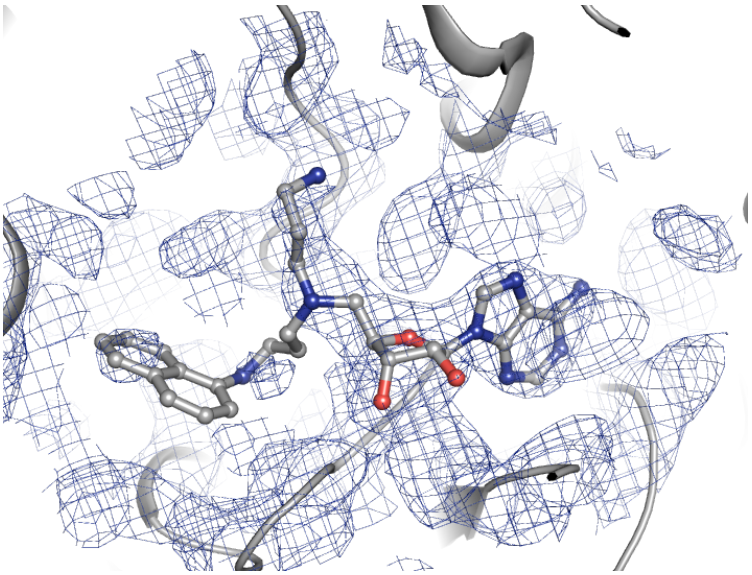


Figure 92: Active site electron density for EG047 in active site of CARM1(WT). The $2mF_o$-DF_c map is contoured to 1σ with a carve radius of 6 Å around the modelled ligand. There is no density into which the propylnaphthylamine or propylamine groups can be modelled (though they are shown for clarity).

Overall, further work is needed to obtain high purity samples of **EG047** and **EG048** to confirm the binding, after which initial FRET studies can be performed.

3.10. Comparison of Results in light of structures recently added to the PDB

As discussed in Section 1.6 (page 28), a large number of CARM1-inhibitor co-crystal structures became available during the course of this work. Of the structures currently available in the PDB (November 2018), 9 feature guanidine-containing bisubstrate inhibitors (Figure 93) and 9 feature bisubstrate inhibitors with aromatic groups (Figure 95).

3.10.1. Guanidine-containing bisubstrate inhibitors in the PDB compared with UM079, UM249, UM251 and WH5C

The IC_{50} values shown in Figure 93 (below) suggest that the optimal number of atoms between the 4' ribose carbon and the guanidine group is two. This suggests (as noted in Section 3.6.2.5) that the linker length in **UM079**, **UM249**, **UM251** and **WH5C** are too long for optimal affinity. In the above discussion this was attributed to the guanidinium group being pushed too far up the arginine binding channel to interact with key residues in the double-E loop of CARM1. Interestingly, the inhibitor guanidinium of **WH3C** in the now-available crystal structure 5ISE is not quite as far up the channel as seen in our **UM079**-CARM1 structure (Figure 94A, below). This is unexpected since both inhibitors have the same structure, except that **UM079** lacks the carboxylic acid group of **WH3C**. Since the b-factors of the guanidinium atoms are high (relative to the ribose moiety) in both cases, it is unlikely that the guanidinium group forms strong

interaction with the CARM1 active site in either case. The 2-atom linker compound in 5K8V (**VH2C**) has similarly high b-factors for the guanidinium group, though – notably – has the second lowest IC_{50} value of all the inhibitors Figure 93 (120 nM). **VH3C**, which has the same structure as **VH2C** except for one additional methylene group in the linker, is a weaker CARM1 inhibitor (560 nM), despite the guanidinium group appearing perfectly orientated to mimic the substrate's interactions with the double-E loop. Thus, the number of apparent interactions made between the guanidinium group and the CARM1 active site does not necessarily correlate with inhibitor potency. Regardless, based on the inhibition and structural data presented in Figure 93 and Figure 94 below, if potency and accurate 'substrate-micking' are desired – the linker lengths in **UM079**, **UM249** and **UM251** (and their amino acid counterparts **WH3C**, **WH4C**, and **WH5C**) are too long.

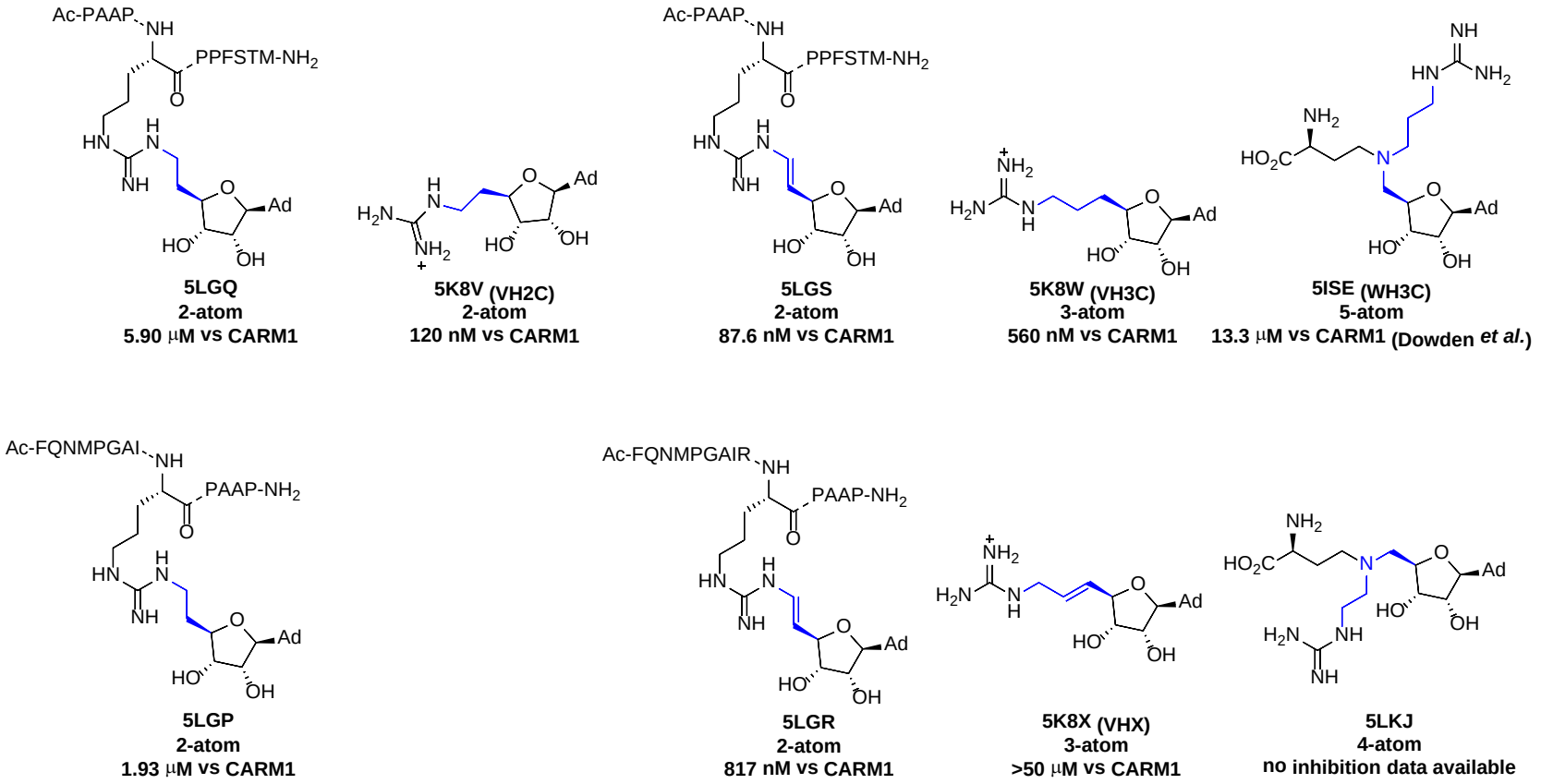


Figure 93: Guanidine-containing bisubstrate CARM1 inhibitors for which there are co-crystal structures in the PDB. The inhibitors are grouped according to the number of atoms between the 4' ribose carbon and the guanidinium group and the relevant bonds and atoms coloured blue. The relevant PDB codes and reported IC_{50} values are given underneath. In several cases, reported IC_{50} values do not originate from a paper with the same primary reference as the related PDB entry. The references for IC_{50} values are: 5LGQ,[74] 5K8V,[70] 5LGS,[74] 5K8W,[70] 5ISE,[69] 5LGP, [74] 5LGR,[74] 5K8X.[70] Where the inhibitor has been mentioned earlier in this thesis, its name is given in brackets after the PDB code. Primary references or author lists for each structure are given in Appendix 6.1.

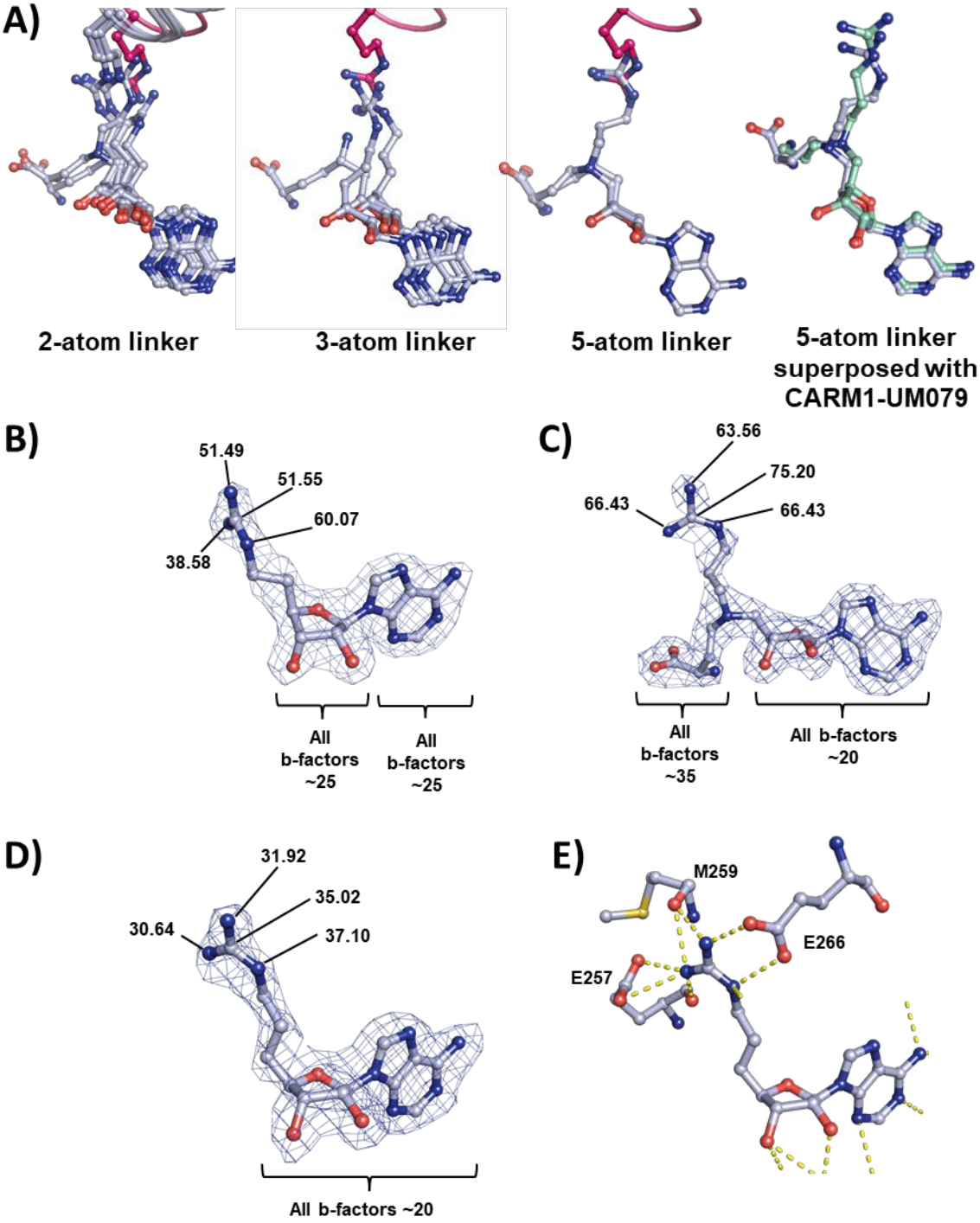


Figure 94: Comparison of guanidinium positioning in guanidine-containing bisubstrate inhibitors co-crystallised with CARM1 in the PDB. (A) Superposition of bisubstrate inhibitors with 5DX0 (CARM1-H3-sinefungin). Inhibitors are shown in pale blue, and the arginine side-chain of H3 is shown in pink. For PDB codes, see Figure 93, above; (B)/(C)/(D) $2mF_o$-DF_c maps of inhibitors in 5K8V (2-atoms), 5ISE (5-atoms + propyl amino acid group) and 5K8W (3-atoms), respectively. The sigma level is 1.0, and the carve radius (around the inhibitor atoms) is 1.5 Å; (E) Polar contacts between guanidine group of VH3C and the double-E loop of CARM1.

3.10.2. Aromatic-containing bisubstrate inhibitors in the PDB compared with UM305, AA175 and AA183

9 CARM1 co-crystal structures are currently available in the PDB with aromatic-containing bisubstrate inhibitors. For most of these inhibitors, supporting data have not yet been published; therefore the strength with which they inhibit PRMT1 and CARM1 cannot be compared to that of **UM305**, **AA175** and **AA183**. The structures of these inhibitors, and available IC_{50} values are given in Figure 95 (below). To look at how the aromatic substituents in these inhibitors are positioned compared to those of **AA175** (our most potent CARM1 inhibitor), each structure was superposed with our CARM1-**AA175** co-crystal structure (Figure 96, below). There is reasonably good overlap of the pyridine of **AA175** and the aromatic groups of related inhibitors, particularly in the SAM-carboxylate binding pocket. Whilst no biochemical data has yet been published for SKI-72, the SGC chemical probes entry states that it is highly potent, CARM1-selective and cell-active.[66] This supports the notion that bisubstrate mimics containing aromatic groups can be used as a basis for CARM1-selective inhibitors.

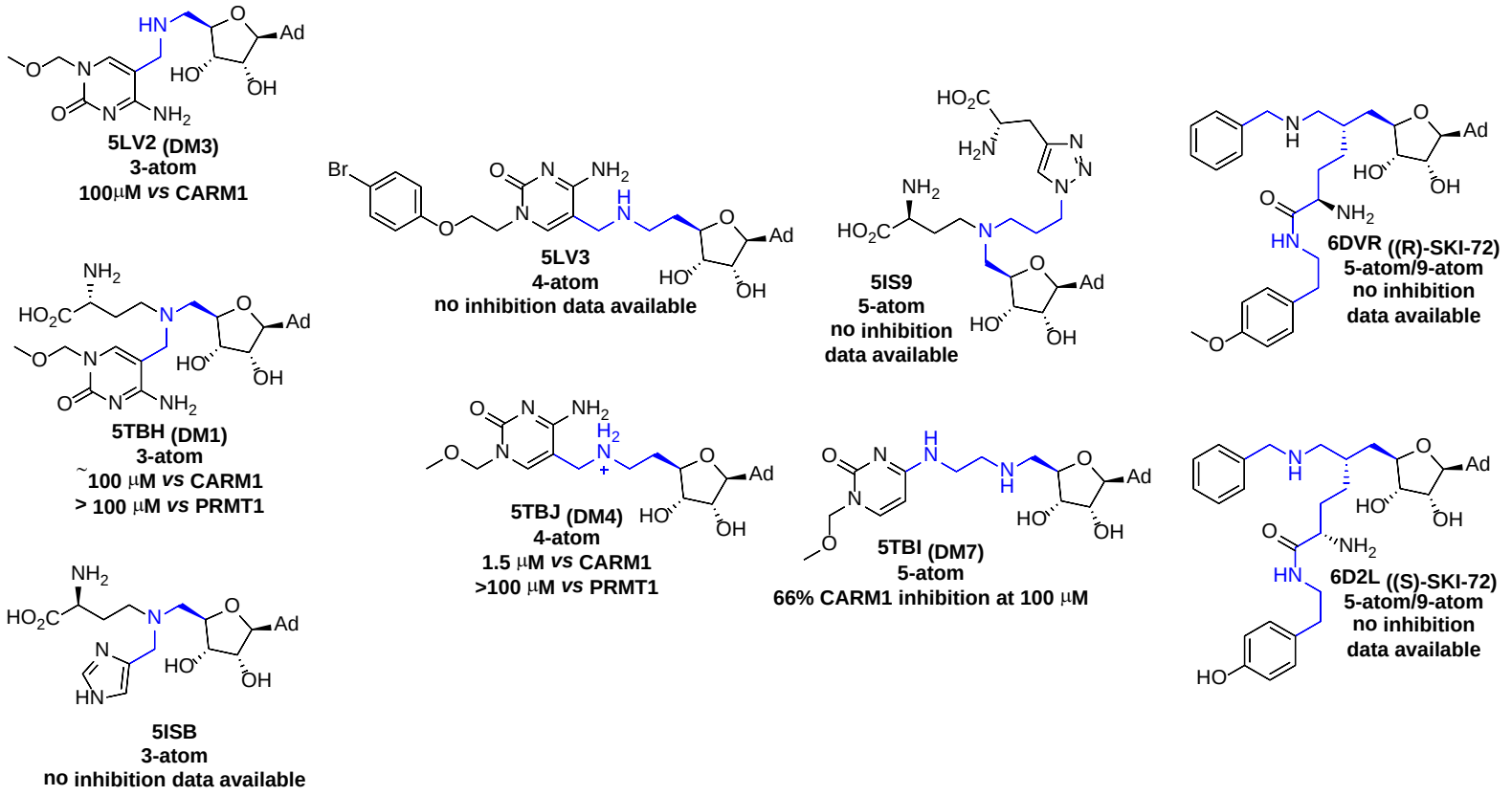


Figure 95: Guanidine-containing bisubstrate CARM1 inhibitors for which there are co-crystal structures in the PDB. The relevant PDB codes and reported IC_{50} values are given underneath (where available). The references for IC_{50} values are: 5LV2,[75] 5TBH, [75] 5TBJ, [75] 5TBI, [75], Where the inhibitor has been mentioned earlier in this thesis, its name is given in brackets after the PDB code. Primary references or author lists for each structure are given in Appendix 6.1.

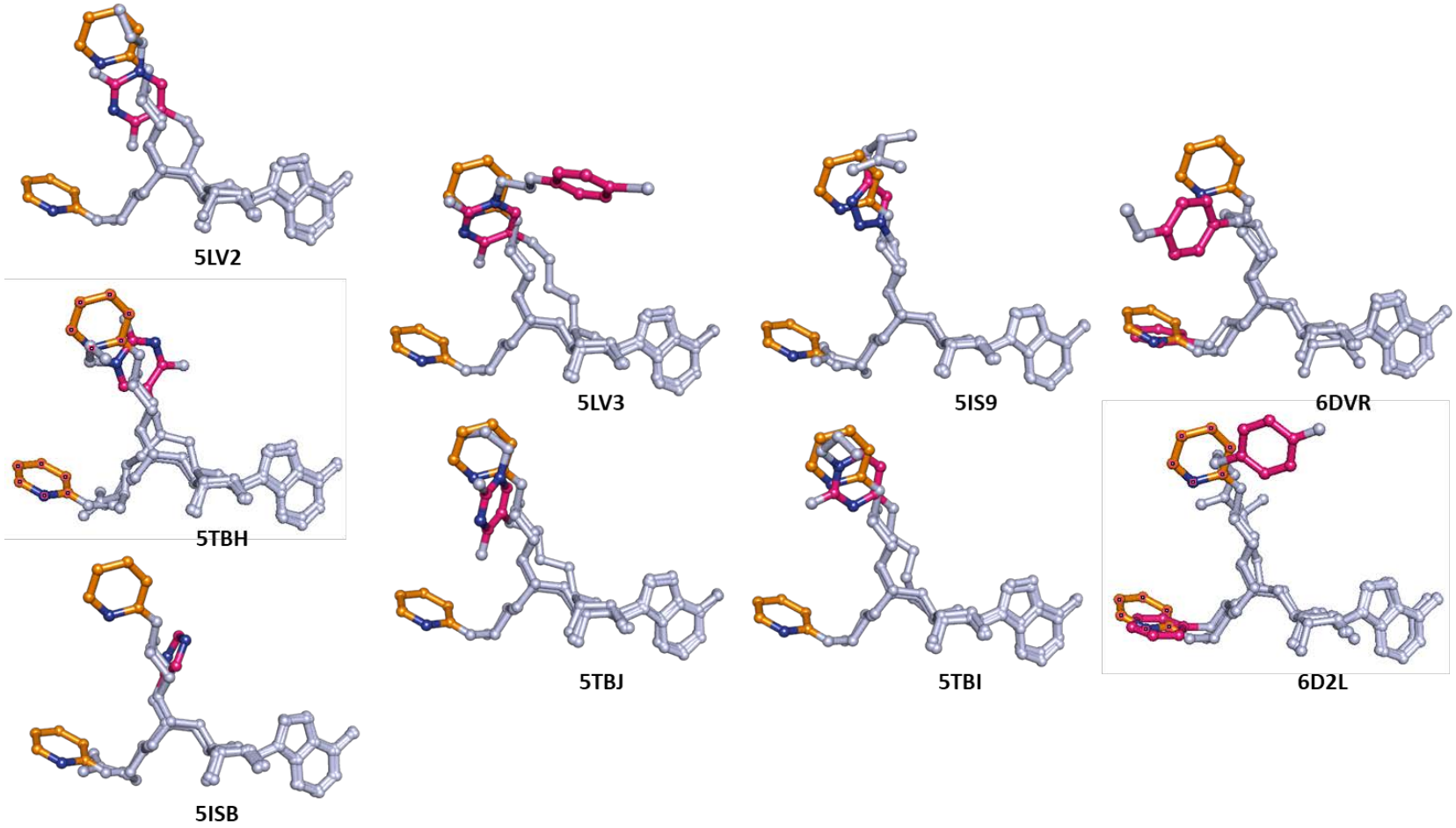


Figure 96: Superposition of CARM1-AA175 co-crystal structure with aromatic bisubstrate inhibitors co-crystallised with CARM1 in the PDB. Ligands are shown in grey, with aromatic groups highlighted in orange (AA175) and hot pink (superposed inhibitors). Relevant PDB codes are given underneath the relevant superposition and literature references for each structure are given in Appendix 6.1.

CHAPTER 4: **Conclusions and Future Work**

The binding of a novel series of bisubstrate PRMT inhibitors (**UM079**, **UM249**, **UM251**) to PRMT1 and the catalytic core of CARM1 have been tested. Co-crystal structures of these inhibitors with CARM1 confirm that the inhibitors occupy the SAM and substrate binding sites as intended. Isothermal titration calorimetry showed that all of the inhibitors tested have low micromolar (3.32 – 9.99 μM) affinity for both PRMT1 and CARM1, but do not display significant selectivity for either enzyme. Activity assays confirmed this result, with IC_{50} values ranging from 2.3 – 17.7 μM and no clear selectivity for PRMT1 or CARM1.

Replacement of the guanidine moiety with aminopyrimidine or aminopyrimidine resulted in inhibitors with modest preference for CARM1, as judged by ITC and activity assays. The most promising compound (**AA183**) was ~30-fold selective for CARM1 over PRMT1 (K_d = 1.11 vs 43.7 μM and IC_{50} = 25.3 vs 0.8 μM). Co-crystallisation of this compound with CARM1 was used to identify two residues (N161 and N265) in the CARM1 binding site that may be responsible for the observed CARM1-selectivity. ITC studies with single site mutants at these positions indicated that N265 is important for **AA183** binding, but N161 is not.

A related series of inhibitors, in which the ribose group is substituted for a morpholino group were tested. This substitution was found to drastically alter the binding. Modest PRMT1 inhibition by several of these compounds was detected using activity assays, but only very weak binding was detected by ITC.

A major limitation of this study is the use of the catalytic core of CARM1 throughout binding studies (as opposed to full length CARM1). Developing a high-yielding expression and purification protocol for an active, full length CARM1 construct would therefore provide reassurance that the data presented herein are relevant to the full length protein. Further, a crystal structure of full length CARM1 would be invaluable in elucidating the roles of the terminal domains in substrate binding, catalysis and protein function in a broader cellular context.

Similarly, developing a reliable method for obtaining PRMT1-inhibitor co-crystals under physiologically relevant conditions would be immensely helpful in the design of PRMT1-selective inhibitors. Further understanding of the PRMT1 oligomerisation would aid in its crystallisation, and provide interesting biological information in itself.

Structure-based modification of the aromatic inhibitors to enhance their CARM1-selectivity is an obvious next step. Several CARM1-selective inhibitors which contain aromatic moieties fused to adenosine mimics have been reported recently, reinforcing the potential for this type of inhibitor (e.g. SKI-72 and SKI-73 [Figure 13, page 22] and inhibitors recently reported by Halby *et al.*).[75]

Finally, large-scale re-synthesis of putative FRET probes **EG047** and **EG048** could be performed, and the TFA-catalysed deprotection optimised to avoid the apparent epimerisation or depurination. While the initial binding studies are promising, a fully characterised, pure sample of each compound is crucial to continue biochemical studies using these compounds.

CHAPTER 5: **References**

1. 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., Methylation of Histone H4 at Arginine 3 Facilitating Transcriptional Activation by Nuclear Hormone Receptor. *Science* **2001,** *293* (5531), 853.

2. Cheng, D.; Côté, J.; Shaaban, S.; Bedford, M. T., The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing. *Molecular Cell* **2007,** *25* (1), 71-83.

3. Vadnais, C.; Chen, R.; Fraszczak, J.; Yu, Z.; Boulais, J.; Pinder, J.; Frank, D.; Khandanpour, C.; Hébert, J.; Dellaire, G.; Côté, J.-F.; Richard, S.; Orthwein, A.; Drobetsky, E.; Möröy, T., GFI1 facilitates efficient DNA repair by regulating PRMT1 dependent methylation of MRE11 and 53BP1. *Nature Communications* **2018,** *9* (1), 1418.

4. Larsen, S. C.; Sylvestersen, K. B.; Mund, A.; Lyon, D.; Mullari, M.; Madsen, M. V.; Daniel, J. A.; Jensen, L. J.; Nielsen, M. L., Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Science Signaling* **2016,** *9* (443), rs9-rs9.

5. (a) Gayatri, S.; Bedford, M. T., Readers of histone methylarginine marks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2014,** *1839* (8), 702-710; (b) Cui, W.; Yoneda, R.; Ueda, N.; Kurokawa, R., Arginine methylation of translocated in liposarcoma (TLS) inhibits its binding to long noncoding RNA, abrogating TLS-mediated repression of CBP/p300 activity. *J. Biol. Chem.* **2018**; (c) Li Kim, K. C.; Chau Bess, L.; Lee Kevin, A. W., Differential interaction of PRMT1 with RGG-boxes of the FET family proteins EWS and TAF15. *Protein Sci.* **2017,** *27* (3), 633-642; (d) Yang, Y.; McBride, Kevin M.; Hensley, S.; Lu, Y.; Chedin, F.; Bedford, Mark T., Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation. *Molecular Cell* **2014,** *53* (3), 484-497.

6. Fuhrmann, J.; Clancy, K. W.; Thompson, P. R., Chemical Biology of Protein Arginine Modifications in Epigenetic Regulation. *Chem. Rev.* **2015,** *115* (11), 5413-5461.

7. (a) Bedford, M. T.; Frankel, A.; Yaffe, M. B.; Clarke, S.; Leder, P.; Richard, S., Arginine Methylation Inhibits the Binding of Proline-rich Ligands to Src Homology 3, but Not WW, Domains. *J. Biol. Chem.* **2000,** *275* (21), 16030-16036; (b) Hughes, R. M.; Waters, M. L., Arginine Methylation in a β-Hairpin Peptide: Implications for Arg–π Interactions, ΔCp°, and the Cold Denatured State. *J. Am. Chem. Soc.* **2006,** *128* (39), 12735-12742.

8. Gayatri, S.; Bedford, M. T., Readers of histone methylarginine marks. *Biochimica et biophysica acta* **2014,** *1839* (8), 702-10.

9. (a) Migliori, V.; Mapelli, M.; Guccione, E., On WD40 proteins: propelling our knowledge of transcriptional control? *Epigenetics* **2012,** *7* (8), 815-22; (b) Shindo, H.; Suzuki, R.; Tsuchiya, W.; Taichi, M.; Nishiuchi, Y.; Yamazaki, T., PHD finger of the SUMO ligase Siz/PIAS family in rice reveals specific binding for methylated histone H3 at lysine 4 and arginine 2. *FEBS Lett.* **2012,** *586* (13), 1783-1789.

10. (a) Bedford, M. T.; Clarke, S. G., Protein Arginine Methylation in Mammals: Who, What, and Why. *Mol. Cell* **2009,** *33* (1), 1-13; (b) Dhar, S.; Vemulapalli, V.; Patananan, A. N.; Huang, G. L.; Di Lorenzo, A.; Richard, S.; Comb, M. J.; Guo, A.; Clarke, S. G.; Bedford, M. T., Loss of the major Type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. *Sci. Rep.* **2013,** *3*, 1311.

11. Richon, V. M.; Johnston, D.; Sneeringer, C. J.; Jin, L.; Majer, C. R.; Elliston, K.; Jerva, L. F.; Scott, M. P.; Copeland, R. A., Chemogenetic analysis of human protein methyltransferases. *Chemical biology & drug design* **2011,** *78* (2), 199-210.

12. Schapira, M.; Ferreira de Freitas, R., Structural biology and chemistry of protein arginine methyltransferases. *Medchemcomm* **2014,** *5* (12), 1779-1788.

13. (a) Vidgren, J.; Svensson, L. A.; Liljas, A., Crystal structure of catechol O-methyltransferase. *Nature* **1994,** *368* (6469), 354-8; (b) Spåhr, H.; Habermann, B.; Gustafsson, C. M.; Larsson, N.-G.; Hallberg, B. M., Structure of the human MTERF4–NSUN4 protein complex that regulates mitochondrial ribosome biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109* (38), 15253-15258; (c) Song, J.; Rechkoblit, O.; Bestor, T. H.; Patel, D. J., Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. *Science (New York, N.Y.)* **2011,** *331* (6020), 1036-1040.

14. Hoffman, J. L., Chromatographic analysis of the chiral and covalent instability of S-adenosyl-L-methionine. *Biochemistry* **1986,** *25* (15), 4444-4449.

15. (a) Boriack-Sjodin, P. A.; Jin, L.; Jacques, S. L.; Drew, A.; Sneeringer, C.; Scott, M. P.; Moyer, M. P.; Ribich, S.; Moradei, O.; Copeland, R. A., Structural Insights into Ternary Complex Formation of Human CARM1 with Various Substrates. *ACS Chemical Biology* **2016,** *11* (3), 763-771; (b) Cao, R.; Zhang, X.; Liu, X.; Li, Y.; Li, H., Molecular basis for histidine N1 position-specific methylation by CARNMT1. *Cell Res.* **2018,** *28* (4), 494-496; (c) Hirano, T.; Fujiwara, T.; Niwa, H.; Hirano, M.; Ohira, K.; Okazaki, Y.; Sato, S.; Umehara, T.; Maemoto, Y.; Ito, A.; Yoshida, M.; Kagechika, H., Development of Novel Inhibitors for Histone Methyltransferase SET7/9 based on Cyproheptadine. *ChemMedChem* **2018,** *13* (15), 1530-1540; (d) Liu, R. J.; Long, T.; Li, J.; Li, H.; Wang, E. D., Structural basis for substrate binding and catalytic mechanism of a human RNA:m5C methyltransferase NSun6. *Nucleic Acids Res.* **2017,** *45* (11), 6684-6697; (e) Duncan, K. W.; Rioux, N.; Boriack-Sjodin, P. A.; Munchhof, M. J.; Reiter, L. A.; Majer, C. R.; Jin, L.; Johnston, L. D.; Chan-Penebre, E.; Kuplast, K. G.; Porter Scott, M.; Pollock, R. M.; Waters, N. J.; Smith, J. J.; Moyer, M. P.; Copeland, R. A.; Chesworth, R., Structure and Property Guided Design in the Identification of PRMT5 Tool Compound EPZ015666. *ACS Med. Chem. Lett.* **2016,** *7* (2), 162-6.

16. (a) Yue, W. W.; Hassler, M.; Roe, S. M.; Thompson-Vale, V.; Pearl, L. H., Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J.* **2007,** *26* (20), 4402-12; (b) Sack, John S.; Thieffine, S.; Bandiera, T.; Fasolini, M.; Duke, Gerald J.; Jayaraman, L.; Kish, Kevin F.; Klei, Herbert E.; Purandare, Ashok V.; Rosettani, P.; Troiani, S.; Xie, D.; Bertrand, Jay A., Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem. J.* **2011,** *436* (2), 331.

17. Thomas, D.; Lakowski, T. M.; Pak, M. L.; Kim, J. J.; Frankel, A., Forster resonance energy transfer measurements of cofactor-dependent effects on protein arginine N-methyltransferase homodimerization. *Protein Sci.* **2010,** *19* (11), 2141-51.

18. Troffer-Charlier, N.; Cura, V.; Hassenboehler, P.; Moras, D.; Cavarelli, J., Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *The EMBO Journal* **2007,** *26* (20), 4391-4401.

19. Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E., The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015,** *10*, 845.

20. (a) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and Clustal X version 2.0. *Bioinformatics* **2007,** *23* (21), 2947-2948; (b) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R., A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Res.* **2010,** *38* (suppl_2), W695-W699.

21. Rust, H. L.; Zurita-Lopez, C. I.; Clarke, S.; Thompson, P. R., Mechanistic Studies on Transcriptional Coactivator Protein Arginine Methyltransferase 1. *Biochemistry* **2011,** *50* (16), 3332-3345.

22. Gui, S.; Wooderchak, W. L.; Daly, M. P.; Porter, P. J.; Johnson, S. J.; Hevel, J. M., Investigation of the molecular origins of protein-arginine methyltransferase I (PRMT1) product specificity reveals a role for two conserved methionine residues. *J. Biol. Chem.* **2011,** *286* (33), 29118-26.

23. Paik, W. K.; Kim, S., Protein Methylase I: PURIFICATION AND PROPERTIES OF THE ENZYME. *J. Biol. Chem.* **1968,** *243* (9), 2108-2114.

24. (a) Goulet, I.; Gauvin, G.; Boisvenue, S.; Côté, J., Alternative Splicing Yields Protein Arginine Methyltransferase 1 Isoforms with Distinct Activity, Substrate Specificity, and Subcellular Localization. *Journal of Biological Chemistry* **2007,** *282* (45), 33009-33021; (b) Patounas, O.; Papacharalampous, I.; Eckerich, C.; Markopoulos, G. S.; Kolettas, E.; Fackelmayer, F. O., A novel splicing isoform of protein arginine methyltransferase 1 (PRMT1) that lacks the dimerization arm and correlates with cellular malignancy. *Journal of Cellular Biochemistry* **2018,** *119* (2), 2110-2123.

25. Auclair, Y.; Richard, S., The role of arginine methylation in the DNA damage response. *DNA repair* **2013,** *12* (7), 459-65.

26. Zhang, X.; Cheng, X., Structure of the Predominant Protein Arginine Methyltransferase PRMT1 and Analysis of Its Binding to Substrate Peptides. *Structure* **2003,** *11* (5), 509-520.

27. (a) Obianyo, O.; Causey, C. P.; Jones, J. E.; Thompson, P. R., Activity-based protein profiling of protein arginine methyltransferase 1. *ACS Chem. Biol.* **2011,** *6* (10), 1127-35; (b) Obianyo, O.; Osborne, T. C.; Thompson, P. R., Kinetic mechanism of Protein Arginine Methyltransferase 1. *Biochemistry* **2008,** *47* (39), 10420-10427; (c) Hu, H.; Luo, C.; Zheng, Y. G., Transient Kinetics Define a Complete Kinetic Model for Protein Arginine Methyltransferase 1. *J. Biol. Chem.* **2016,** *291* (52), 26722-26738; (d) Lakowski, Ted M.; Frankel, A., Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem. J.* **2009,** *421* (2), 253; (e) Feng, Y.; Xie, N.; Jin, M.; Stahley, M. R.; Stivers, J. T.; Zheng, Y. G., A Transient Kinetic Analysis of PRMT1 Catalysis. *Biochemistry* **2011,** *50* (32), 7033-7044; (f) Osborne, T. C.; Obianyo, O.; Zhang, X.; Cheng, X.; Thompson, P. R., Protein Arginine Methyltransferase 1: Positively Charged Residues in Substrate Peptides Distal to the Site of Methylation Are Important for Substrate Binding and Catalysis. *Biochemistry* **2007,** *46* (46), 13370-13381; (g) Dorgan, K. M.; Wooderchak, W. L.; Wynn, D. P.; Karschner, E. L.; Alfaro, J. F.; Cui, Y.; Zhou, Z. S.; Hevel, J. M., An enzyme-coupled continuous spectrophotometric assay for S-adenosylmethionine-dependent methyltransferases. *Anal. Biochem.* **2006,** *350* (2), 249-55; (h) Gui, S.; Wooderchak-Donahue, W. L.; Zang, T.; Chen, D.; Daly, M. P.; Zhou, Z. S.; Hevel, J. M., Substrate-induced control of product formation by protein arginine methyltransferase 1. *Biochemistry* **2013,** *52* (1), 199-209.

28. Brown Jennifer, I.; Koopmans, T.; van Strien, J.; Martin Nathaniel, I.; Frankel, A., Kinetic Analysis of PRMT1 Reveals Multifactorial Processivity and a Sequential Ordered Mechanism. *Chembiochem* **2017,** *19* (1), 85-99.

29. (a) Mathioudaki, K.; Scorilas, A.; Ardavanis, A.; Lymberi, P.; Tsiambas, E.; Devetzi, M.; Apostolaki, A.; Talieri, M., Clinical evaluation of PRMT1 gene expression in breast cancer. *Tumour Biol.* **2011,** *32* (3), 575-82; (b) Baldwin, R. M.; Bejide, M.; Trinkle-Mulcahy, L.; Cote, J., Identification of the PRMT1v1 and PRMT1v2 specific interactomes by quantitative mass spectrometry in breast cancer cells. *Proteomics* **2015,** *15* (13), 2187-97; (c) Papadokostopoulou, A.; Mathioudaki, K.; Scorilas, A.; Xynopoulos, D.; Ardavanis, A.; Kouroumalis, E.; Talieri, M., Colon cancer and protein arginine methyltransferase 1 gene expression. *Anticancer Res.* **2009,** *29* (4), 1361-6; (d) Mathioudaki, K.; Papadokostopoulou, A.; Scorilas, A.; Xynopoulos, D.; Agnanti, N.;

Talieri, M., The PRMT1 gene expression pattern in colon cancer. *Br. J. Cancer* **2008,** *99* (12), 2094-9.

30. 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., PRMT1 Is a Novel Regulator of Epithelial-Mesenchymal-Transition in Non-small Cell Lung Cancer. *J. Biol. Chem.* **2015,** *290* (21), 13479-89.

31. 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., Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. *Int. J. Cancer* **2011,** *128* (3), 562-73.

32. Wang, S.; Tan, X.; Yang, B.; Yin, B.; Yuan, J.; Qiang, B.; Peng, X., The role of protein arginine-methyltransferase 1 in gliomagenesis. *BMB Rep* **2012,** *45* (8), 470-5.

33. Li, Y.-J.; Stallcup, M. R.; Lai, M. M. C., Hepatitis Delta Virus Antigen Is Methylated at Arginine Residues, and Methylation Regulates Subcellular Localization and RNA Replication. *Journal of Virology* **2004,** *78* (23), 13325-13334.

34. (a) Zakrzewicz, D.; Zakrzewicz, A.; Preissner, K. T.; Markart, P.; Wygrecka, M., Protein Arginine Methyltransferases (PRMTs): promising targets for the treatment of pulmonary disorders. *Int J Mol Sci* **2012,** *13* (10), 12383-400; (b) Sen, S.; He, Z.; Ghosh, S.; Dery, K. J.; Yang, L.; Zhang, J.; Sun, Z., PRMT1 Plays a Critical Role in Th17 Differentiation by Regulating Reciprocal Recruitment of STAT3 and STAT5. *The Journal of Immunology* **2018**.

35. Chen, D.; Ma, H.; Hong, H.; Koh, S. S.; Huang, S. M.; Schurter, B. T.; Aswad, D. W.; Stallcup, M. R., Regulation of transcription by a protein methyltransferase. *Science* **1999,** *284* (5423), 2174-7.

36. Shishkova, E.; Zeng, H.; Liu, F.; Kwiecien, N. W.; Hebert, A. S.; Coon, J. J.; Xu, W., Global mapping of CARM1 substrates defines enzyme specificity and substrate recognition. *Nat Commun* **2017,** *8*, 15571.

37. Wu, J.; Cui, N.; Wang, R.; Li, J.; Wong, J., A Role for CARM1-Mediated Histone H3 Arginine Methylation in Protecting Histone Acetylation by Releasing Corepressors from Chromatin. *PLOS ONE* **2012,** *7* (6), e34692.

38. Naeem, H.; Cheng, D.; Zhao, Q.; Underhill, C.; Tini, M.; Bedford, M. T.; Torchia, J., The activity and stability of the transcriptional coactivator p/CIP/SRC-3 are regulated by CARM1-dependent methylation. *Mol Cell Biol* **2007,** *27* (1), 120-34.

39. 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., CARM1 methylates chromatin remodeling factor BAF155 to enhance tumor progression and metastasis. *Cancer Cell* **2014,** *25* (1), 21-36.

40. Higashimoto, K.; Kuhn, P.; Desai, D.; Cheng, X.; Xu, W., Phosphorylation-mediated inactivation of coactivator-associated arginine methyltransferase 1. *Proceedings of the National Academy of Sciences* **2007,** *104* (30), 12318.

41. Teyssier, C.; Chen, D.; Stallcup, M. R., Requirement for multiple domains of the protein arginine methyltransferase CARM1 in its transcriptional coactivator function. *J. Biol. Chem.* **2002,** *277* (48), 46066-72.

42. Drew, A. E.; Moradei, O.; Jacques, S. L.; Rioux, N.; Boriack-Sjodin, A. P.; Allain, C.; Scott, M. P.; Jin, L.; Raimondi, A.; Handler, J. L.; Ott, H. M.; Kruger, R. G.; McCabe, M. T.; Sneeringer, C.; Riera, T.; Shapiro, G.; Waters, N. J.; Mitchell, L. H.; Duncan, K. W.; Moyer, M. P.;

Copeland, R. A.; Smith, J.; Chesworth, R.; Ribich, S. A., Identification of a CARM1 Inhibitor with Potent In Vitro and In Vivo Activity in Preclinical Models of Multiple Myeloma. *Sci. Rep.* **2017,** *7* (1), 17993.

43. Jacques, S. L.; Aquino, K. P.; Gureasko, J.; Boriack-Sjodin, P. A.; Porter Scott, M.; Copeland, R. A.; Riera, T. V., CARM1 Preferentially Methylates H3R17 over H3R26 through a Random Kinetic Mechanism. *Biochemistry* **2016,** *55* (11), 1635-1644.

44. 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., Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. *Cancer* **2004,** *101* (1), 83-89.

45. (a) Ou, C. Y.; LaBonte, M. J.; Manegold, P. C.; So, A. Y.; Ianculescu, I.; Gerke, D. S.; Yamamoto, K. R.; Ladner, R. D.; Kahn, M.; Kim, J. H.; Stallcup, M. R., A coactivator role of CARM1 in the dysregulation of beta-catenin activity in colorectal cancer cell growth and gene expression. *Mol. Cancer Res.* **2011,** *9* (5), 660-70; (b) Kim, Y.-R.; Lee, B. K.; Park, R.-Y.; Nguyen, N. T. X.; Bae, J. A.; Kwon, D. D.; Jung, C., Differential CARM1 expression in prostate and colorectal cancers. *BMC Cancer* **2010,** *10* (1), 197.

46. Elakoum, R.; Gauchotte, G.; Oussalah, A.; Wissler, M. P.; Clement-Duchene, C.; Vignaud, J. M.; Gueant, J. L.; Namour, F., CARM1 and PRMT1 are dysregulated in lung cancer without hierarchical features. *Biochimie* **2014,** *97*, 210-8.

47. (a) Morettin, A.; Baldwin, R. M.; Cote, J., Arginine methyltransferases as novel therapeutic targets for breast cancer. *Mutagenesis* **2015,** *30* (2), 177-89; (b) Cheng, H.; Qin, Y.; Fan, H.; Su, P.; Zhang, X.; Zhang, H.; Zhou, G., Overexpression of CARM1 in breast cancer is correlated with poorly characterized clinicopathologic parameters and molecular subtypes. *Diagn. Pathol.* **2013,** *8*, 129.

48. Osada, S.; Suzuki, S.; Yoshimi, C.; Matsumoto, M.; Shirai, T.; Takahashi, S.; Imagawa, M., Elevated expression of coactivator-associated arginine methyltransferase 1 is associated with early hepatocarcinogenesis. *Oncol. Rep.* **2013,** *30* (4), 1669-74.

49. Vu, L. P.; Perna, F.; Wang, L.; Voza, F.; Figueroa, M. E.; Tempst, P.; Erdjument-Bromage, H.; Gao, R.; Chen, S.; Paietta, E.; Deblasio, T.; Melnick, A.; Liu, Y.; Zhao, X.; Nimer, S. D., PRMT4 blocks myeloid differentiation by assembling a Methyl-RUNX1-dependent repressor complex. *Cell Reports* **2013,** *5* (6), 1625-1638.

50. Greenblatt, S. M.; Man, N.; Hamard, P.-J.; Asai, T.; Karl, D.; Martinez, C.; Bilbao, D.; Stathais, V.; McGrew-Jermacowicz, A.; Duffort, S.; Tadi, M.; Blumenthal, E.; Newman, S.; Vu, L.; Xu, Y.; Liu, F.; Schurer, S. C.; McCabe, M. T.; Kruger, R. G.; Xu, M.; Yang, F.-C.; Tenen, D.; Watts, J.; Vega, F.; Nimer, S. D., CARM1 Is Essential for Myeloid Leukemogenesis but Dispensable for Normal Hematopoiesis. *Cancer Cell* **2018,** *33* (6), 1111-1127.e5.

51. Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L., Principles of early drug discovery. *British journal of pharmacology* **2011,** *162* (6), 1239-1249.

52. Drews, J., Drug discovery: a historical perspective. *Science (New York, N.Y.)* **2000,** *287* (5460), 1960-4.

53. 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.; Frederiksen, 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.; Müller, S.; O'Hagan, R. C.; Overington, J. P.; Owen, D. R.; Rosenberg, S. H.; Ross, R.; Roth, B.;

Schapira, M.; Schreiber, S. L.; Shoichet, B.; Sundström, 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., The promise and peril of chemical probes. *Nature Chemical Biology* **2015,** *11*, 536.

54. Hu, H.; Qian, K.; Ho, M. C.; Zheng, Y. G., Small Molecule Inhibitors of Protein Arginine Methyltransferases. *Expert Opin Investig Drugs* **2016,** *25* (3), 335-58.

55. Rasco, D.; Tolcher, A.; Siu, L. L.; Heinhuis, K.; Postel-Vinay, S.; Barbash, O.; Egger, J. L.; Gorman, S.; Horner, T.; Dhar, A.; Kremer, B. E., Abstract CT038: A phase I, open-label, dose-escalation study to investigate the safety, pharmacokinetics, pharmacodynamics, and clinical activity of GSK3326595 in subjects with solid tumors and non-Hodgkin's lymphoma. *Cancer Res.* **2017,** *77* (13 Supplement), CT038.

56. Mitchell, L. H.; Drew, A. E.; Ribich, S. A.; Rioux, N.; Swinger, K. K.; Jacques, S. L.; Lingaraj, T.; Boriack-Sjodin, P. A.; Waters, N. J.; Wigle, T. J.; Moradei, O.; Jin, L.; Riera, T.; Porter-Scott, M.; Moyer, M. P.; Smith, J. J.; Chesworth, R.; Copeland, R. A., Aryl Pyrazoles as Potent Inhibitors of Arginine Methyltransferases: Identification of the First PRMT6 Tool Compound. *ACS Medicinal Chemistry Letters* **2015,** *6* (6), 655-659.

57. Gerhart, S. V.; Kellner, W. A.; Thompson, C.; Pappalardi, M. B.; Zhang, X.-P.; Montes de Oca, R.; Penebre, E.; Duncan, K.; Boriack-Sjodin, A.; Le, B.; Majer, C.; McCabe, M. T.; Carpenter, C.; Johnson, N.; Kruger, R. G.; Barbash, O., Activation of the p53-MDM4 regulatory axis defines the anti-tumour response to PRMT5 inhibition through its role in regulating cellular splicing. *Scientific Reports* **2018,** *8* (1), 9711.

58. Mitchell, L. H.; Drew, A. E.; Ribich, S. A.; Rioux, N.; Swinger, K. K.; Jacques, S. L.; Lingaraj, T.; Boriack-Sjodin, P. A.; Waters, N. J.; Wigle, T. J.; Moradei, O.; Jin, L.; Riera, T.; Porter-Scott, M.; Moyer, M. P.; Smith, J. J.; Chesworth, R.; Copeland, R. A., Aryl Pyrazoles as Potent Inhibitors of Arginine Methyltransferases: Identification of the First PRMT6 Tool Compound. *ACS Medicinal Chemistry Letters* **2015**.

59. 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.; Frederiksen, M.; Frye, S. V.; Gray, N.; Grimshaw, C. E.; Hepworth, D.; Howe, T.; Huber, K. V.; 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., The promise and peril of chemical probes. *Nat. Chem. Biol.* **2015,** *11* (8), 536-41.

60. Yang, H.; Ouyang, Y.; Ma, H.; Cong, H.; Zhuang, C.; Lok, W. T.; Wang, Z.; Zhu, X.; Sun, Y.; Hong, W.; Wang, H., Design and synthesis of novel PRMT1 inhibitors and investigation of their binding preferences using molecular modelling. *Bioorg Med Chem Lett* **2017,** *27* (20), 4635-4642.

61. Wang, C.; Jiang, H.; Jin, J.; Xie, Y.; Chen, Z.; Zhang, H.; Lian, F.; Liu, Y.-C.; Zhang, C.; Ding, H.; Chen, S.; Zhang, N.; Zhang, Y.; Jiang, H.; Chen, K.; Ye, F.; Yao, Z.; Luo, C., Development of Potent Type I Protein Arginine Methyltransferase (PRMT) Inhibitors of Leukemia Cell Proliferation. *Journal of Medicinal Chemistry* **2017,** *60* (21), 8888-8905.

62. Xie, Y.; Zhou, R.; Lian, F.; Liu, Y.; Chen, L.; Shi, Z.; Zhang, N.; Zheng, M.; Shen, B.; Jiang, H.; Liang, Z.; Luo, C., Virtual screening and biological evaluation of novel small molecular inhibitors against protein arginine methyltransferase 1 (PRMT1). *Org Biomol Chem* **2014,** *12* (47), 9665-73.

63. (a) Wang, J.; Chen, L.; Sinha, S. H.; Liang, Z.; Chai, H.; Muniyan, S.; Chou, Y.-W.; Yang, C.; Yan, L.; Feng, Y.; Kathy Li, K.; Lin, M.-F.; Jiang, H.; George Zheng, Y.; Luo, C., Pharmacophore-Based Virtual Screening and Biological Evaluation of Small Molecule Inhibitors for Protein Arginine Methylation. *Journal of Medicinal Chemistry* **2012,** *55* (18), 7978-7987; (b) Feng, Y.; Li, M.; Wang, B.; Zheng, Y. G., Discovery and Mechanistic Study of a Class of Protein Arginine Methylation Inhibitors. *Journal of Medicinal Chemistry* **2010,** *53* (16), 6028-6039.

64. (a) 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., N-Benzyl-1-heteroaryl-3-(trifluoromethyl)-1H-pyrazole-5-carboxamides as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg Med Chem Lett* **2009,** *19* (4), 1218-23; (b) Cheng, D.; Valente, S.; Castellano, S.; Sbardella, G.; Di Santo, R.; Costi, R.; Bedford, M. T.; Mai, A., Novel 3,5-bis(bromohydroxybenzylidene)piperidin-4-ones as coactivator-associated arginine methyltransferase 1 inhibitors: enzyme selectivity and cellular activity. *J Med Chem* **2011,** *54* (13), 4928-32; (c) Ferreira de Freitas, R.; Eram, M. S.; Smil, D.; Szewczyk, M. M.; Kennedy, S.; Brown, P. J.; Santhakumar, V.; Barsyte-Lovejoy, D.; Arrowsmith, C. H.; Vedadi, M.; Schapira, M., Discovery of a Potent and Selective Coactivator Associated Arginine Methyltransferase 1 (CARM1) Inhibitor by Virtual Screening. *J Med Chem* **2016,** *59* (14), 6838-47; (d) 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., Optimization of pyrazole inhibitors of Coactivator Associated Arginine Methyltransferase 1 (CARM1). *Bioorganic & Medicinal Chemistry Letters* **2009,** *19* (11), 2924-2927; (e) John S. Sack, S. T., Tiziano Bandiera, Marina Fasolini, Gerald J. Duke, Lata Jayaraman, Kevin F. Kish, Herbert E. Klei, Ashok V. Purandare, Pamela Rosettani, Sonia Troiani, Dianlin Xie and Jay A. Bertrand, Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochemical Journal* **2011,** *436*, 331-339; (f) Purandare, A. V.; Chen, Z.; Huynh, T.; Pang, S.; Geng, J.; Vaccaro, W.; Poss, M. A.; Oconnell, J.; Nowak, K.; Jayaraman, L., Pyrazole inhibitors of coactivator associated arginine methyltransferase 1 (CARM1). *Bioorganic & Medicinal Chemistry Letters* **2008,** *18* (15), 4438-4441; (g) 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., Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *The Biochemical journal* **2011,** *436* (2), 331-9; (h) Therrien, E.; Larouche, G.; Manku, S.; Allan, M.; Nguyen, N.; Styhler, S.; Robert, M. F.; Goulet, A. C.; Besterman, J. M.; Nguyen, H.; Wahhab, A., 1,2-Diamines as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg Med Chem Lett* **2009,** *19* (23), 6725-32; (i) 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., Benzo[d]imidazole inhibitors of Coactivator Associated Arginine Methyltransferase 1 (CARM1)—Hit to Lead studies. *Bioorganic & Medicinal Chemistry Letters* **2009,** *19* (17), 5063-5066.

65. Nakayama, K.; Szewczyk, M. M.; Dela Sena, C.; Wu, H.; Dong, A.; Zeng, H.; Li, F.; de Freitas, R. F.; Eram, M. S.; Schapira, M.; Baba, Y.; Kunitomo, M.; Cary, D. R.; Tawada, M.; Ohashi, A.; Imaeda, Y.; Saikatendu, K. S.; Grimshaw, C. E.; Vedadi, M.; Arrowsmith, C. H.; Barsyte-Lovejoy, D.; Kiba, A.; Tomita, D.; Brown, P. J., TP-064, a potent and selective small molecule inhibitor of PRMT4 for multiple myeloma. *Oncotarget* **2018,** *9* (26), 18480-18493.

66. https://www.thesgc.org/chemical-probes.

67. Muhsen, U. A. Toward the Development of Novel and Selective Bisubstrate Inhibitors of Protein Arginine Methyltransferases. University of Nottingham, 2014.

68. Dowden, J.; Hong, W.; Parry, R. V.; Pike, R. A.; Ward, S. G., Toward the development of potent and selective bisubstrate inhibitors of protein arginine methyltransferases. *Bioorg. Med. Chem. Lett.* **2010,** *20* (7), 2103-2105.

69. Dowden, J.; Pike, R. A.; Parry, R. V.; Hong, W.; Muhsen, U. A.; Ward, S. G., Small molecule inhibitors that discriminate between protein arginine N-methyltransferases PRMT1 and CARM1. *Organic & Biomolecular Chemistry* **2011,** *9* (22), 7814-7821.

70. van Haren, M.; van Ufford, L. Q.; Moret, E. E.; Martin, N. I., Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Organic & Biomolecular Chemistry* **2015,** *13* (2), 549-560.

71. Shanying Gui, W. L. W., Michael P. Daly, Paula J. Porter, Sean J. Johnson and Joan M. Hevel Investigation of the Molecular Origins of Protein-arginine Methyltransferase I (PRMT1) Product Specificity Reveals a Role for Two Conserved Methionine Residues. *Journal of Biological Chemistry* **2011,** *286*, 29118-29126.

72. Yue, W. W.; Hassler, M.; Roe, S. M.; Thompson-Vale, V.; Pearl, L. H., Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *The EMBO Journal* **2007,** *26* (20), 4402-4412.

73. Cura, V.; Marechal, N.; Troffer-Charlier, N.; Strub, J. M.; van Haren, M. J.; Martin, N. I.; Cianferani, S.; Bonnefond, L.; Cavarelli, J., Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors. *The FEBS journal* **2017,** *284* (1), 77-96.

74. van Haren, M. J.; Marechal, N.; Troffer-Charlier, N.; Cianciulli, A.; Sbardella, G.; Cavarelli, J.; Martin, N. I., Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. *Proceedings of the National Academy of Sciences* **2017,** *114* (14), 3625-3630.

75. Halby, L.; Marechal, N.; Pechalrieu, D.; Cura, V.; Franchini, D. M.; Faux, C.; Alby, F.; Troffer-Charlier, N.; Kudithipudi, S.; Jeltsch, A.; Aouadi, W.; Decroly, E.; Guillemot, J. C.; Page, P.; Ferroud, C.; Bonnefond, L.; Guianvarc'h, D.; Cavarelli, J.; Arimondo, P. B., Hijacking DNA methyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors. *Philos Trans R Soc Lond B Biol Sci* **2018,** *373* (1748).

76. Inoue, H.; Nojima, H.; Okayama, H., High efficiency transformation of Escherichia coli with plasmids. *Gene* **1990,** *96* (1), 23-28.

77. (a) https://www.takarabio.com/products/protein-research/expression-vectors-and-systems/protein-folding-kits/pcold-dna-cold-shock; (b) https://www.agilent.com/cs/library/usermanuals/Public/230191.pdf.

78. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **1976,** *72* (1), 248-254.

79. Hong, W. Design and synthesis of protein arginine methyltransferase inhibitors. 2010.

80. Al-Noori, A. Structural investigation of protein arginine methyltransferase bisubstrate inhibitors. University of Nottingham, 2018.

81. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2011,** *67* (Pt 4), 235-242.

82. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** *40* (4), 658-674.

83. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010,** *66* (Pt 4), 486-501.

84. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010,** *66* (Pt 2), 213-221.

85. https://www.birmingham.ac.uk/maps/buildings/edgbaston-campus/yellow-zone/alta.aspx.

86. https://www.graphpad.com/guides/prism/7/curve-fitting/index.htm?reg_50_of_what__relative_vs_absolu.htm.

87. Reeve, A. M.; Townsend, C. A., Inhibitors of an AdoMet-dependent 3-amino-3-carboxypropyl transferase and their use as ligands for protein affinity chromatography. *Tetrahedron* **1998,** *54* (52), 15959-15974.

88. Ensch, C.; Hesse, M., Total Syntheses of the Spermine Alkaloids (–)-(R,R)-Hopromine and (±)-Homaline. *Helv. Chim. Acta* **2002,** *85* (6), 1659-1677.

89. Mancuso, A. J.; Huang, S.-L.; Swern, D., Oxidation of long-chain and related alcohols to carbonyls by dimethyl sulfoxide "activated" by oxalyl chloride. *J. Org. Chem.* **1978,** *43* (12), 2480-2482.

90. Frade, V. H. J.; Coutinho, P. J. G.; Moura, J. C. V. P.; Gonçalves, M. S. T., Functionalised benzo[a]phenoxazine dyes as long-wavelength fluorescent probes for amino acids. *Tetrahedron* **2007,** *63* (7), 1654-1663.

91. Kusakabe, T.; Sekiyama, E.; Ishino, Y.; Motodate, S.; Kato, S.; Mochida, T.; Kato, K., Cyclization–Carbonylation–Cyclization Coupling Reactions of N-Propargylanilines and o-Alkynylphenols with Palladium(II)–Bisoxazoline Catalysts. *Synthesis* **2012,** *44* (12), 1825-1832.

92. Ingale, A. P.; More, V. K.; Gangarde, U. S.; Shinde, S. V., Chemoselective N-tert-butyloxycarbonylation of amines in glycerol. *New J. Chem.* **2018,** *42* (12), 10142-10147.

93. (a) Huang, L.; Sun, X.; Li, Q.; Qi, C., Palladium-Catalyzed C8 Alkylation of 1-Naphthylamides with Alkyl Halides via Bidentate-Chelation Assistance. *J. Org. Chem.* **2014,** *79* (14), 6720-6725; (b) Zhang, X.; Si, W.; Bao, M.; Asao, N.; Yamamoto, Y.; Jin, T., Rh(III)-Catalyzed Regioselective Functionalization of C–H Bonds of Naphthylcarbamates for Oxidative Annulation with Alkynes. *Org. Lett.* **2014,** *16* (18), 4830-4833.

94. Cheng, Y.; Patel, D. J., An efficient system for small protein expression and refolding. *Biochemical and Biophysical Research Communications* **2004,** *317* (2), 401-405.

95. Strahl, B. D.; Briggs, S. D.; Brame, C. J.; Caldwell, J. A.; Koh, S. S.; Ma, H.; Cook, R. G.; Shabanowitz, J.; Hunt, D. F.; Stallcup, M. R.; Allis, C. D., Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* **2001,** *11* (12), 996-1000.

96. Price, O. M. N., Damon; Gui, Laurel; Hevel, Joan M., Investigation into the kinetics and oligomerization of PRMT1. *Abstracts of Papers, 247th ACS National Meeting & Exposition, Dallas, TX, United States* **2014**.

97. Seddon, A. M.; Curnow, P.; Booth, P. J., Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2004,** *1666* (1), 105-117.

98. Reuten, R.; Nikodemus, D.; Oliveira, M. B.; Patel, T. R.; Brachvogel, B.; Breloy, I.; Stetefeld, J.; Koch, M., Maltose-Binding Protein (MBP), a Secretion-Enhancing Tag for Mammalian Protein Expression Systems. *PLoS One* **2016,** *11* (3), e0152386.

99. Kuhn, P.; Chumanov, R.; Wang, Y.; Ge, Y.; Burgess, R. R.; Xu, W., Automethylation of CARM1 allows coupling of transcription and mRNA splicing. *Nucleic Acids Res.* **2011,** *39* (7), 2717-26.

100. Damian, L., Isothermal titration calorimetry for studying protein-ligand interactions. *Methods Mol. Biol.* **2013,** *1008*, 103-18.

101. Caulton, S. Characterisation of the deubiquitinating enzyme USP20. University of Nottingham, 2017.

102. Fraser, J. S.; Merlie, J. P., Jr.; Echols, N.; Weisfield, S. R.; Mignot, T.; Wemmer, D. E.; Zusman, D. R.; Alber, T., An atypical receiver domain controls the dynamic polar localization of the Myxococcus xanthus social motility protein FrzS. *Mol. Microbiol.* **2007,** *65* (2), 319-32.

103. Qing, G.; Ma, L.-C.; Khorchid, A.; Swapna, G. V. T.; Mal, T. K.; Takayama, M. M.; Xia, B.; Phadtare, S.; Ke, H.; Acton, T.; Montelione, G. T.; Ikura, M.; Inouye, M., Cold-shock induced high-yield protein production in Escherichia coli. *Nat. Biotechnol.* **2004,** *22*, 877.

104. Hoffmann, A.; Bukau, B.; Kramer, G., Structure and function of the molecular chaperone Trigger Factor. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2010,** *1803* (6), 650-661.

105. Dessau, M. A.; Modis, Y., Protein Crystallization for X-ray Crystallography. *Journal of Visualized Experiments : JoVE* **2011,** (47), 2285.

106. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S. e.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, Walker, J. M., Ed. Humana Press: Totowa, NJ, 2005; pp 571-607.

107. (a) Copeland, R. A.; Solomon, M. E.; Richon, V. M., Protein methyltransferases as a target class for drug discovery. *Nature Reviews Drug Discovery* **2009,** *8*, 724; (b) Qian, K.; Hu, H.; Xu, H.; Zheng, Y. G., Detection of PRMT1 inhibitors with stopped flow fluorescence. *Signal Transduction and Targeted Therapy* **2018,** *3* (1), 6.

108. Yung-Chi, C.; Prusoff, W. H., Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical Pharmacology* **1973,** *22* (23), 3099-3108.

109. Troffer-Charlier, N.; Cura, V.; Hassenboehler, P.; Moras, D.; Cavarelli, J., Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *The EMBO Journal* **2007,** *26* (20), 4391-4401.

110. Kleywegt, G. J., Crystallographic refinement of ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **2007,** *63* (Pt 1), 94-100.

111. Steiner, R. A.; Tucker, J. A., Keep it together: restraints in crystallographic refinement of macromolecule-ligand complexes. *Acta Crystallographica Section D* **2017,** *73* (2), 93-102.

112. (a) Summerton, J., Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1999,** *1489* (1), 141-158; (b) Corey, D. R.; Abrams, J. M., Morpholino antisense oligonucleotides: tools for investigating vertebrate development. *Genome Biol.* **2001,** *2* (5),

reviews1015.1-reviews1015.3; (c) Heasman, J., Morpholino Oligos: Making Sense of Antisense? *Dev. Biol.* **2002,** *243* (2), 209-214.

113. Hu, H.; Qian, K.; Ho, M.-C.; Zheng, Y. G., Small Molecule Inhibitors of Protein Arginine Methyltransferases. *Expert Opinion on Investigational Drugs* **2016,** *25* (3), 335-358.

114. Ghisaidoobe, A. B. T.; Chung, S. J., Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *International Journal of Molecular Sciences* **2014,** *15* (12), 22518-22538.

115. Liao, F.; Xie, Y.; Yang, X.; Deng, P.; Chen, Y.; Xie, G.; Zhu, S.; Liu, B.; Yuan, H.; Liao, J.; Zhao, Y.; Yu, M. a., Homogeneous noncompetitive assay of protein via Förster-resonance-energy-transfer with tryptophan residue(s) as intrinsic donor(s) and fluorescent ligand as acceptor. *Biosensors and Bioelectronics* **2009,** *25* (1), 112-117.

116. (a) Doody, M. C.; Gotto, A. M.; Smith, L. C., 5-(Dimethylamino)naphthalene-1-sulfonic acid, a fluorescent probe of the medium chain fatty acid binding site of serum albumin. *Biochemistry* **1982,** *21* (1), 28-33; (b) Kim, J. H.; Sumranjit, J.; Kang, H. J.; Chung, S. J., Discovery of coumarin derivatives as fluorescence acceptors for intrinsic fluorescence resonance energy transfer of proteins. *Molecular bioSystems* **2014,** *10* (1), 30-3; (c) Lee, M. M.; Peterson, B. R., Quantification of Small Molecule–Protein Interactions using FRET between Tryptophan and the Pacific Blue Fluorophore. *ACS Omega* **2016,** *1* (6), 1266-1276.

117. (a) A. Klee, W.; Mudd, S., *The Conformation of Ribonucleosides in Solution. The Effect of Structure on the Orientation of the Base* *. 1967; Vol. 6, p 988-98; (b) K. Shapiro, S.; J. Ehninger, D., *Methods for the analysis and preparation of S-adenosylmethionine and S-adenosylhomocysteine*. 1966; Vol. 15, p 323-33; (c) Iwig, D. F.; Booker, S. J., Insight into the Polar Reactivity of the Onium Chalcogen Analogues of S-Adenosyl-l-methionine. *Biochemistry* **2004,** *43* (42), 13496-13509; (d) Dalhoff, C.; Lukinavicius, G.; Klimasauskas, S.; Weinhold, E., *Synthesis of S-adenosyl-L-methionine analogs and their use for sequence-specific transalkylation of DNA by methyltransferases*. 2006; Vol. 1, p 1879-86.

CHAPTER 6: **Appendix**

6.1. **Summary of CARM1 Structures available in the PDB at the time of writing (September 2018)**

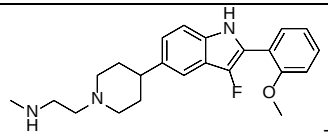
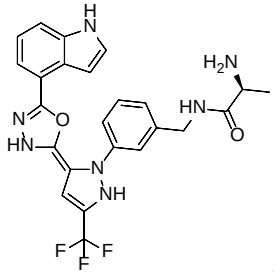
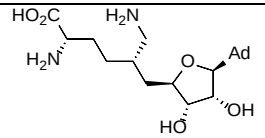
See following page for summary table.

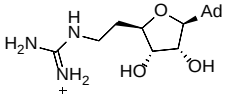
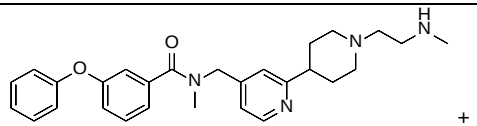
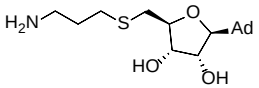
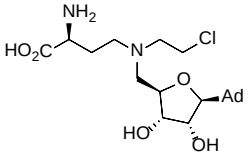
Entries are given in chronological order according to structure release date.

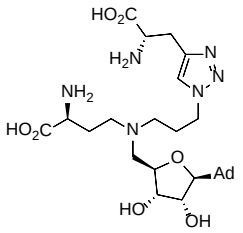
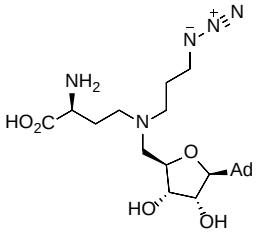
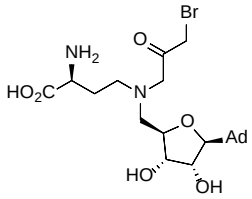
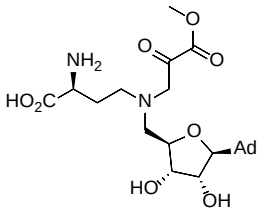
Inhibitor structures, crystal growth conditions, space groups and unit cell dimension were taken directly from the corresponding PDB entries.

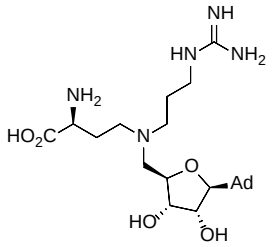
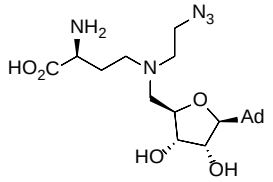
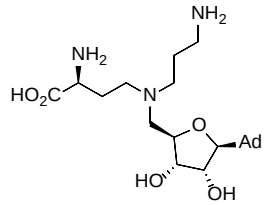
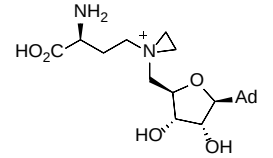
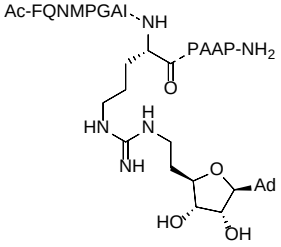
Only ligands that are relevant to this work are included (i.e. components of the crystallisation cocktails/heavy atoms/cryoprotectants are excluded).

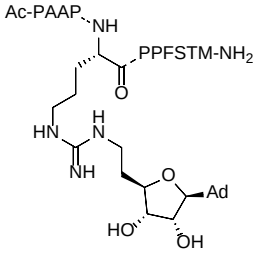
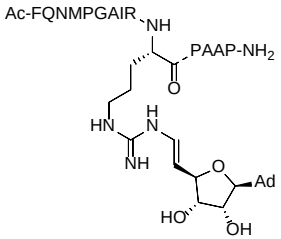
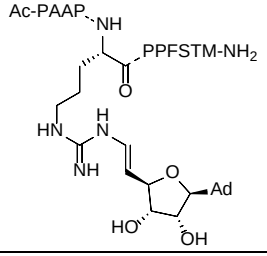
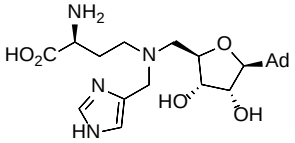
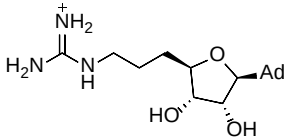
Primary citations relating to each PDB entry are indicated with the PDB code. Absence of a citation indicates that no primary citation is yet available for the PDB entry.

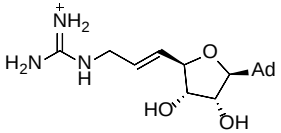
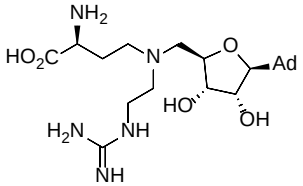
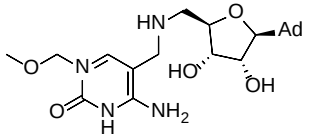
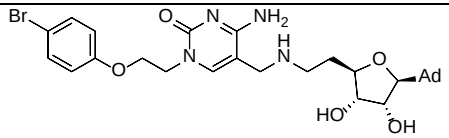
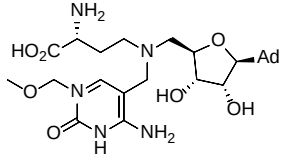
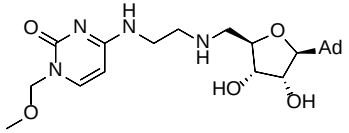
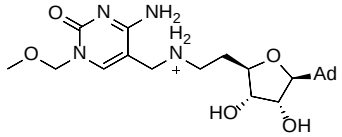
Date	PDB ID	Ligand(s)	Crystal Growth Conditions	Space Group	Unit Cell Angles (°)			Unit Cell Dimensions (Å)		
					α	β	γ	a	b	c
2007	2V74[18, 72]	SAH	0.1 M HEPES, 1.6M $(NH_4)_2HPO_4$, pH 7	P 21212	90	90	90	74.88	98.69	207.22
2007	2V7E[72]	-	0.1 M HEPES , 1.6M $(NH_4)_2HPO_4$, pH 7.5	I222	90	90	90	74.15	98.02	206.91
2007	2OQB[18]	-	13%-16% PEG 6000, pH 8	P41212	90	90	90	46.13	46.13	201.32
2007	3B3F[18]	SAH	0.1 5M Na malate, 19% PEG 3350, pH 7	P21212	90	90	90	74.82	98.69	206.93
2007	3B3G[18]	-	0.15 M Na malate, 19% PEG 3350, pH 7	I222	90	90	90	73.82	98.12	207.13
2007	3B3J[18]	-	0.1 M Na citrate, 0.1 M NaCl 0.1 M benzamidine chloride, 5% MPD, pH 5.6	P6222	90	90	120	136.02	136.02	125.3
2011	2Y1W[64g]	 + SFG	0.1 M Tris, 0.2 M trimethylamine N-oxide dihydrate 20-30% PEG 2000K, pH 8.5	P21212	90	90	90	74.9	98.47	207.18
2011	2Y1X[64g]	 + SAH	0.1 M Tris, 0.2 M trimethylamine N-oxide dihydrate 20-30% PEG 2000K, pH 8.5	P21212	90	90	90	75.56	98.76	207.46
2013	4IKP		0.2 M di-ammonium tartrate, 20% PEG 3350	P21212	90	90	90	75.06	98.84	206.63

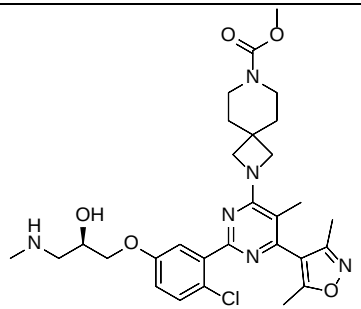
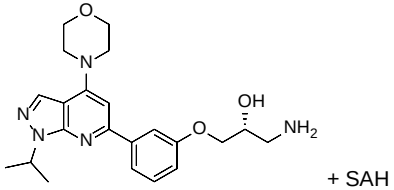
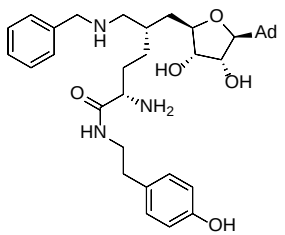
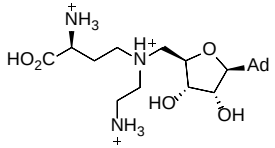
2015	5DWQ[15a]	SFG + methylated H3 peptide	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	75.42	98.4	208.37
2015	5DX0[15a]	SFG + H3 peptide	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	75.72	98.85	208.67
2015	5DX1[15a]	SFG + PABP1 peptide	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	75	98.32	208.15
2015	5DX8[15a]	SFG + methylated PABP1 peptide	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	74.86	98.27	207.54
2015	5DXA[15a]	SFG + methylated PABP1 peptide	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	75.18	98.69	207.48
2015	5DXJ[15a]	SFG	0.1 M Tris, 0.2 M NH_4SO_4, 18% w/v PEG 3350, pH 8.5	P21212	90	90	90	74.93	97.88	207.38
2016	5K8V[73]		0.1 M Tris, 0.1 M NaCl, 15 % PEG 2000 MME, pH 8	P21212	90	90	90	75.08	98.09	206.26
2016	5U4X[65]	 SAH	0.1 M Tris, 0.2 M AmSO4, 20% PEG 3350, pH 6.5	P21212	90	90	90	75.09	98.92	207.55
2017	5IS6*	SFG	0.1 M Tris, 0.1 M NaCl, 28 % PEG 2000 MME, pH 8	P21212	90	90	90	74.68	98.09	206.37
2017	5IS7*		0.1 M Tris, 0.167 M NaCl, 25 % PEG 2000 MME, pH 8	P21212	90	90	90	74.53	98.93	206.91
2017	5IS8*		0.1 M Bis-tris propane, 15 % PEG 3350, pH 7.5	P21212	90	90	90	74.54	98.62	207.47

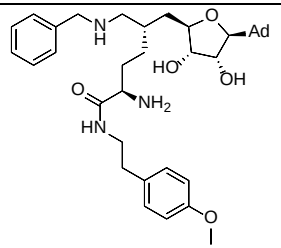
2017	5IS9*		0.1 M Tris, 0.1 M NaCl, 13 % PEG 2000 MME	P21212	90	90	90	74.89	98.37	206.94
2017	5ISA*		0.1 M MIB, 0.1 M NaCl, 17.5 % PEG 1500, pH 7	P21212	90	90	90	74.54	98.41	205.8
2017	5ISC*		0.1 M Tris, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	79.12	101.04	211.35
2017	5ISD*		0.1 M Tris, 0.133M NaCl, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	74.56	98.45	205.79

2017	5ISE*		0.1 M Tris, 0.133 M NaCl, 15 % PEG 2000 MME, pH 8	P21212	90	90	90	74.74	98.3	206.23
2017	5ISF*		0.1 M Tris, 0.167 M NaCl, 10 % PEG 2000 MME, pH 8	P21212	90	90	90	74.74	98.14	206.67
2017	5ISG*		0.1 M Tris, 0.2 M NaCl, 22 % PEG 2000 MME, pH 8	P21212	90	90	90	74.69	98.16	205.11
2017	5ISH*		0.1 M Tris, 67 mM NaCl, 17 % PEG 2000 MME, pH 8	P21212	90	90	90	74.68	98.41	206.85
2017	5ISI*	5'-amino-5'-deoxyadenosine	0.1 M Tris, 67 mM NaCl, 16 % PEG 2000 MME, pH 8	P21212	90	90	90	74.74	98.52	207.05
2017	5LGP[74]		0.1 M Tris, 0.2 M NH_4SO_4, 16% PEG 3350, pH 8.5	P21212	90	90	90	74.79	98.71	207.25

2017	5LGQ[74]		0.1 M Tris, 0.2 M NH_4SO_4, 14% PEG 3350, pH 8.5	P21212	90	90	90	74.71	99.01	207.31
2017	5LGR[74]		0.1 M Tris, 0.2 M NH_4SO_4, 14% PEG 3350, pH 8.5	P21212	90	90	90	74.83	98.54	208.34
2017	5LGS[74]		0.1 M Tris, 0.2 M NH_4SO_4, 20% PEG 3350, pH 8	P21212	90	90	90	74.99	98.6	208.17
2017	5IH3*	SAH	0.1 M Tris, 0.167 mM NaCl, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	74.63	97.92	204.96
2017	5ISB*		0.1 M Tris, 0.133 mM NaCl, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	74.43	98.08	205.68
2017	5K8W*		0.1 M Tris, 15 %PEG 2000 MME, pH 8	P21212	90	90	90	74.92	98.82	207.47

2017	5K8X*		0.1 M Tris, 0.1 M NaCl, 15 % PEG 2000 MME, pH 8	P21212	90	90	90	74.83	98.49	205.53
2017	5LKJ*		0.1 M Bis-tris propane, 30% PEG 1500, pH 7.5	P21212	90	90	90	74.64	98.16	206.5
2017	5LV2*		0.1 M Tris, 0.1 M NaCl, 17 % PEG 2000 MME, pH 8	P1211	90	90.53	90	98.68	74.83	206.64
2017	5LV3[75]		0.1 M Tris, 0.1 M NaCl, 15 % PEG 2000 MME, pH 8	P21212	90	90	90	75.09	98.52	208.1
2017	5TBH*		0.1 M Tris, 0.1 M NaCl, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	74.65	98.53	206.4
2017	5TBI*		0.1 M Tris, 0.1 M NaCl, 21 % PEG 2000 MME, pH 8	P21212	90	90	90	74.49	98.03	205.73
2017	5TBJ*		0.1 M Tris, 0.1 M NaCl, 19 % PEG 2000 MME, pH 8	P21212	90	90	90	74.84	99.2	207.26

2018	6ARJ[42]	 (EPZ022302) + SAH	0.1 M Tris , 0.2 M NH_4SO_4, 18% w/v PEG3350, pH 8.5	P21212	90	90	90	75.16	98.8	208.29
2018	6ARV[42]	 + SAH 	0.1 M Tris , 0.2 M NH_4SO_4, 18% w/v PEG3350, pH 8.5	P21212	90	90	90	75.23	99.12	208.85
2018	6D2L		0.1 M HEPES , 0.2 M NH_4SO_4, 25% PEG 3350, pH 7.5	P1211	90	101.04	90	75.6	155.64	95.34
2018	5NTC*		0.1 M Bis-tris propane, 50 mM NaCl, 25 % PEG 2000 MME, pH 7.5	P21212	90	90	90	74.91	99.15	206.87

2018	6DVR		0.1 M HEPES , 0.2 M NH_4SO_4, 25% PEG 3350, pH 7.5	I222	90	90	90	74.37	98.58	207.38

*Supporting publications not yet available, but all PDB entries have the same authors: Cura, V., Marechal, N., Mailliot, J., Troffer-Charlier, N., Hassenboehler, P., Wurtz, J.M., Bonnefond, L., Cavarelli, J.

6.2. **Alignment of PRMT1 and CARM1 Sequences Encoded by our Template Plasmids with Relevant Sequences**

6.2.1. PRMT1

```
PRMT1_22-361        ------------------------------SCGQAESSEKPNAEDMTSKDYYFDSYAHF 29
PRMT1_Rat           ----------MAAAEAANCI--------MEVSCGQAESSEKPNAEDMTSKDYYFDSYAHF 42
PRMT1_Human         ----------MENFVATLANGMSLQPPLEEVSCGQAESSEKPNAEDMTSKDYYFDSYAHF 50
PRMT1_Template      ------------DLVPRGSPGISGGG--GGIPCGQAESSEKPNAEDMTSKDYYFDSYAHF 46

PRMT1_22-361        GIHEEMLKDEVRTLTYRNSMFHNRHLFKDKVVLDVGSGTGILCMFAAKAGARKVIGIECS 89
PRMT1_Rat           GIHEEMLKDEVRTLTYRNSMFHNRHLFKDKVVLDVGSGTGILCMFAAKAGARKVIGIECS 102
PRMT1_Human         GIHEEMLKDEVRTLTYRNSMFHNRHLFKDKVVLDVGSGTGILCMFAAKAGARKVIGIECS 110
PRMT1_Template      GIHEEMLKDEVRTLTYRNSMFHNRHLFKDKVVLDVGSGTGILCMFAAKAGARKVIGIECS 106

PRMT1_22-361        SISDYAVKIVKANKLDHVVTIIKGKVEEVELPVEKVDIIISEWMGYCLFYESMLNTVLYA 149
PRMT1_Rat           SISDYAVKIVKANKLDHVVTIIKGKVEEVELPVEKVDIIISEWMGYCLFYESMLNTVLHA 162
PRMT1_Human         SISDYAVKIVKANKLDHVVTIIKGKVEEVELPVEKVDIIISEWMGYCLFYESMLNTVLYA 170
PRMT1_Template      SISDYAVKIVKANKLDHVVTIIKGKVEEVELPVEKVDIIISEWMGYCLFYESMLNTVLYA 166

PRMT1_22-361        RDKWLAPDGLIFPDRATLYVTAIEDRQYKDYKIHWWENVYGFDMSCIKDVAIKEPLVDVV 209
PRMT1_Rat           RDKWLAPDGLIFPDRATLYVTAIEDRQYKDYKIHWWENVYGFDMSCIKDVAIKEPLVDVV 222
PRMT1_Human         RDKWLAPDGLIFPDRATLYVTAIEDRQYKDYKIHWWENVYGFDMSCIKDVAIKEPLVDVV 230
PRMT1_Template      RDKWLAPDGLIFPDRATLYVTAIEDRQYKDYKIHWWENVYGFDMSCIKDVAIKEPLVDVV 226

PRMT1_22-361        DPKQLVTNACLIKEVDIYTVKVEDLTFTSPFCLQVKRNDYVHALVAYFNIEFTRCHKRTG 269
PRMT1_Rat           DPKQLVTNACLIKEVDIYTVKVEDLTFTSPFCLQVKRNDYVHALVAYFNIEFTRCHKRTG 282
PRMT1_Human         DPKQLVTNACLIKEVDIYTVKVEDLTFTSPFCLQVKRNDYVHALVAYFNIEFTRCHKRTG 290
PRMT1_Template      DPKQLVTNACLIKEVDIYTVKVEDLTFTSPFCLQVKRNDYVHALVAYFNIEFTRCHKRTG 286

PRMT1_22-361        FSTSPESPYTHWKQTVFYMEDYLTVKTGEEIFGTIGMRPNAKNNRDLDFTIDLDFKG--- 326
PRMT1_Rat           FSTSPESPYTHWKQTVFYMEDYLTVKTGEEIFGTIGMRPNAKNNRDLDFTIDLDFKG--- 339
PRMT1_Human         FSTSPESPYTHWKQTVFYMEDYLTVKTGEEIFGTIGMRPNAKNNRDLDFTIDLDFKG--- 347
PRMT1_Template      FSTSPESPYTHWKQTVFYMEDYLTVKTGEEIFXXIGMRPNAXEQPGXXXXHRXXXXGPAX 346

PRMT1_22-361        ---QLCELSCSTDYRMR--- 340
PRMT1_Rat           ---QLCELSCSTDYRMR--- 353
PRMT1_Human         ---QLCELSCSTDYRMR--- 361
PRMT1_Template      RXVXLHRLXXAXNXXXXGXT 366
```

6.2.2. CARM1

```
CARM1_135-479       ------------------------------------------------------------
CARM1_Human         ------------------MAAAAAAVGPGAGGAGSAVPGGAGPCATVSVFPGARLLTIG 41
CARM1_Mouse         ------------------MAAAAATAVGPGAGSAGVAGPGGAGPCATVSVFPGARLLTIG 42
CARM1_Template      XXXXSDLVPRGSPEFAAASTAAAATAVGPGAGSAGVAGPGGAGPCATVSVFPGARLLTIG 60

CARM1_135-479       ------------------------------------------------------------
CARM1_Human         DANGEIQRHAEQQALRLEVRAGPDSAGIALYSHEDVCVFKCSVSRETECSRVGKQSFIIT 101
CARM1_Mouse         DANGEIQRHAEQQALRLEVRAGPDAAGIALYSHEDVCVFKCSVSRETECSRVGRQSFIIT 102
CARM1_Template      DANGEIQRHAEQQALRLEVRAGPDAAGIALYSHEDVCVFKCSVSRETECSRVGRQSFIIT 120

CARM1_135-479       --------------------------------SVFSERTEESSAVQYFQFYGYLSQQQN 27
CARM1_Human         LGCNSVLIQFATPNDFCSFYNILKTCRGHTLERSVFSERTEESSAVQYFQFYGYLSQQQN 161
CARM1_Mouse         LGCNSVLIQFATPHDFCSFYNILKTCRGHTLERSVFSERTEESSAVQYFQFYGYLSQQQN 162
CARM1_Template      LGCNSVLIQFATPHDFCSFYNILKTCRGHTLERSVFSERTEESSAVQYFQFYGYLSQQQN 180

CARM1_135-479       MMQDYVRTGTYQRAILQNHTDFKDKIVLDVGCGSGILSFFAAQAGARKIYAVEASTMAQH 87
CARM1_Human         MMQDYVRTGTYQRAILQNHTDFKDKIVLDVGCGSGILSFFAAQAGARKIYAVEASTMAQH 221
CARM1_Mouse         MMQDYVRTGTYQRAILQNHTDFKDKIVLDVGCGSGILSFFAAQAGARKIYAVEASTMAQH 222
CARM1_Template      MMQDYVRTGTYQRAILQNHTDFKDKIVLDVGCGSGILSFFAAQAGARKIYAVEASTMAQH 240

CARM1_135-479       AEVLVKSNNLTDRIVVIPGKVEEVSLPEQVDIIISEPMGYMLFNERMLESYLHAKKYLKP 147
CARM1_Human         AEVLVKSNNLTDRIVVIPGKVEEVSLPEQVDIIISEPMGYMLFNERMLESYLHAKKYLKP 281
CARM1_Mouse         AEVLVKSNNLTDRIVVIPGKVEEVSLPEQVDIIISEPMGYMLFNERMLESYLHAKKYLKP 282
CARM1_Template      AEVLVKSNNLTDRIVVIPGKVEEVSLPEQVDIIISEPMGYMLFNERMLESYLHAKKYLKP 300

CARM1_135-479       SGNMFPTIGDVHLAPFTDEQLYMEQFTKANFWYQPSFHGVDLSALRGAAVDEYFRQPVVD 207
CARM1_Human         SGNMFPTIGDVHLAPFTDEQLYMEQFTKANFWYQPSFHGVDLSALRGAAVDEYFRQPVVD 341
CARM1_Mouse         SGNMFPTIGDVHLAPFTDEQLYMEQFTKANFWYQPSFHGVDLSALRGAAVDEYFRQPVVD 342
CARM1_Template      SXNMFPTIGDVHLAPFTDEQLYMXQFTKANFWYQPSFHGVDLSALRGAAVDEYFRQPVVD 360

CARM1_135-479       TFDIRILMAKSVKYTVNFLEAKEGDLHRIEIPFKFHMLHSGLVHGLAFWFDVAFIGSIMT 267
CARM1_Human         TFDIRILMAKSVKYTVNFLEAKEGDLHRIEIPFKFHMLHSGLVHGLAFWFDVAFIGSIMT 401
CARM1_Mouse         TFDIRILMAKSVKYTVNFLEAKEGDLHRIEIPFKFHMLHSGLVHGLAFWFDVAFIGSIMT 402
CARM1_Template      TFDIRILMAKSVKYTVNFLEAKEGDLHRIEIPFKFHMLHSGLVHGLAFWFDVAFIGSIMT 420

CARM1_135-479       VWLSTAPTEPLTHWYQVRCLFQSPLFAKAGDTLSGTCLLIANKRQSYDISIVAQVDQTGS 327
CARM1_Human         VWLSTAPTEPLTHWYQVRCLFQSPLFAKAGDTLSGTCLLIANKRQSYDISIVAQVDQTGS 461
CARM1_Mouse         VWLSTAPTEPLTHWYQVRCLFQSPLFAKAGDTLSGTCLLIANKRQSYDISIVAQVDQTGS 462
CARM1_Template      VWLSTAPTEPLTHWYQVRCLFQSPLFAKAGDTLSGTCLLIANKRQSYDISIVAQVDQTGS 480

CARM1_135-479       KSSNLLDLKNPFFRYTGT---------------------------------------- 345
CARM1_Human         KSSNLLDLKNPFFRYTGTTPSPPPGSHYTSPSENMWNTGSTYNLSSGMAVAGMPTAYDLS 521
CARM1_Mouse         KSSNLLDLKNPFFRYTGTTPSPPPGSHYTSPSENMWNTGSTYNLSSGVAVAGMPTAYDLS 522
CARM1_Template      KSSNLLDLKNPFFRYTGTXXXXXXXXXXXXXX---------------------------- 512

CARM1_135-479       ------------------------------------------------------------
CARM1_Human         SVIASGSSVGHNNLIPLANTGIVNHTHSRMGSIMSTGIVQGSSGAQGSGGGSISAHYAVN 581
CARM1_Mouse         SVIAGGSSVGHNNLIPLANTGIVNHTHSRMGSIMSTGIVQGSSGAQGGGGSS-SAHYAVN 581
CARM1_Template      ------------------------------------------------------------

CARM1_135-479       ---------------------------
CARM1_Human         SQFTMGGPAISMASPMSIPTNTMHYGS 608
CARM1_Mouse         NQFTMGGPAISMASPMSIPTNTMHYGS 608
CARM1_Template      ---------------------------
```

6.3. Additional Analysis of Crystal Structures

6.3.1. Guanidine-containing Inhibitor B-factors

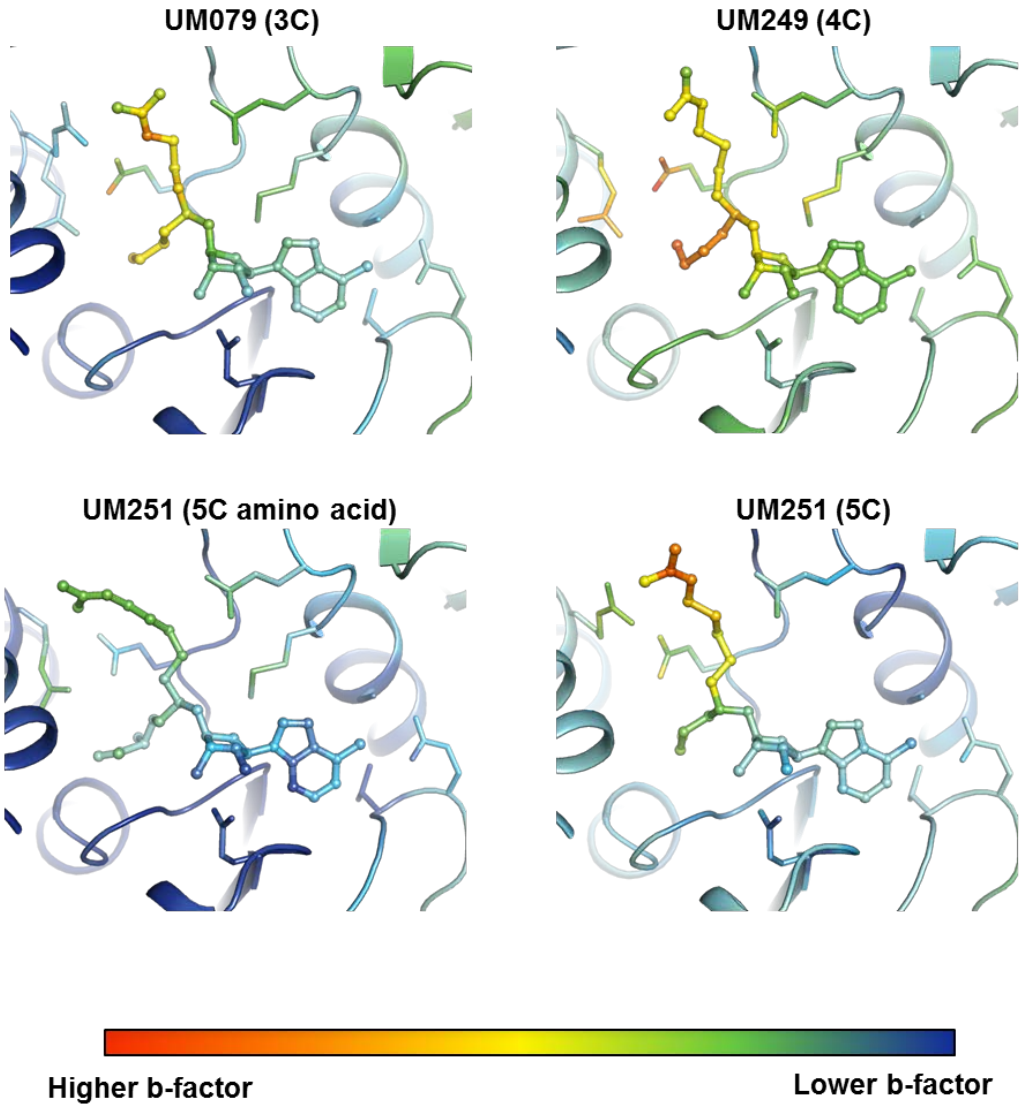


Figure A1: CARM1 co-crystal structures with guanidine inibitors series coloured by b-factor. Colours are coloured using the rainbow spectrum in PyMOL, with colours closer to blue indicating lower b-factors and colours closer to red indicating higher b-factors. Absolute b-factors depend on additional factors such as crystal mosaicity and should therefore not be compared between structures.

6.4. Protein and Ligand Concentration Determination

6.4.1. Bradford Assay to Verify ε_{260} Determined using Expasy Prot Param

Bovine serum albumin (BSA) was serially diluted from 1 to 0.3125 mg/mL in an appropriate buffer and used to generate a standard curve by mixing 10 µL of protein solution with 200 µl of Bradford reagent (Thermo Fisher Scientific). 10 µL of the test protein (in the same buffer) was mixed with the same reagent volume and all samples were incubated at r.t. for 10 minutes. The absorbance at 595 nm (A_{595}) was measured, and the concentrations of test samples determined by comparison with the BSA standard curve.

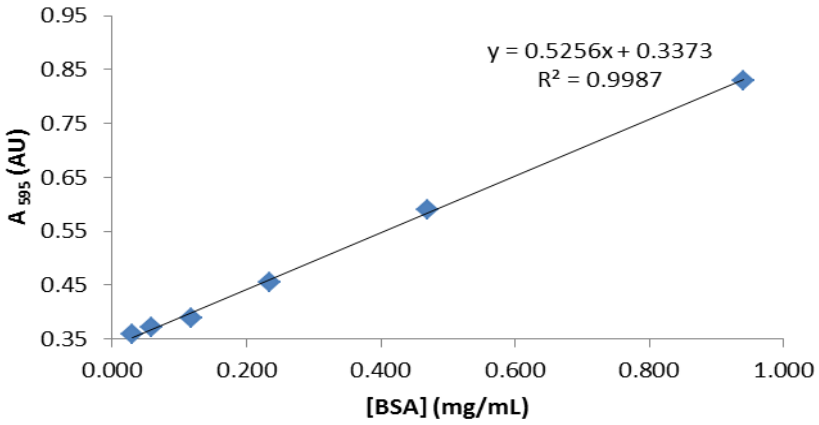


Figure A2: BSA standard curve

Protein	Calculated concentration (mg/mL)	
	Bradford Assay	Nanodrop
PRMT1	0.53	0.42
CARM1	0.58	0.56

Figure A3: Concentrations of PRMT1 and CARM1 determined using the Bradford assay and Nanodrop

6.4.2. Nanodrop Linearity Check

Throughout this thesis, the ligand concentrations were calculated using the Beer-Lambert equation (A: absorbance in AU; ε: extinction coefficient in $M^{-1}cm^{-1}$; c: concentration in M; l: the path-length in cm).

$$A = \varepsilon c l$$

The published molar extinction coefficient for SAH and SAM ($15400 M^{-1}cm^{-1}$)[117] was used. At high absorbance, the Beer-Lambert law breaks down. A dilution series was performed to determine the concentration range within which Beer's Law is valid. A small amount of **WH5C** was dissolved in 200 µL H_2O and serially diluted (100 µL + 100 µL) to give 8 solutions of unknown concentration. The A_{260} of these solutions was measured in triplicate using a NanoDrop 1000 v3.7. The A_{260} of the most dilute solution was found to be 1.3, corresponding to a concentration of 0.084 mM (calculated using $\varepsilon_{260} = 15400 M^{-1}cm^{-1}$). In general, all ligand solutions prepared for ITC, assays and crystallisations were less than or equal to 2 mM. In cases where more concentrated solutions were required, an aliquot was removed and serially diluted to a concentration below 2 mM. The concentration of this sub-sample was then scaled-up to determine the concentration of the parent solution.

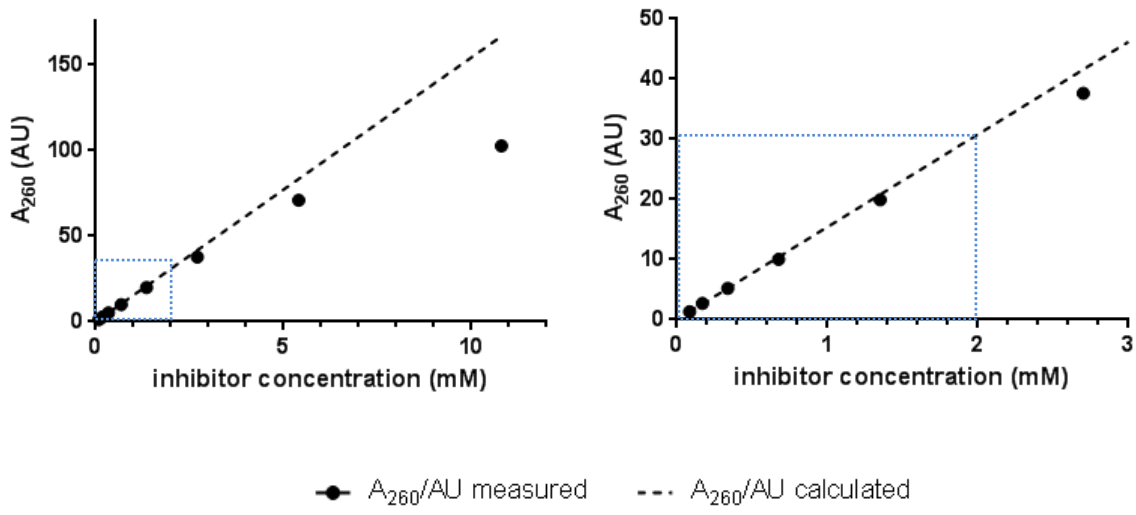


Figure A4: Serial dilution of inhibitor WH5C to determine the concentration at which the Beer-Lambert law can no longer be used to calculate the concentration of ligand in solution. The average A_{260} of each sample is shown as a black dot. The A_{260} calculated for each sample using the Beer-Lambert law is shown as a dashed line. Error bars were too small to be visible. The data on both graphs are the same, but are shown on different scales. The blue dashed line indicated the highest concentration for which the Beer-Lambert law was used to determine concentration.

6.5. **Crystallographic Data Processing Statistics**

6.5.1. Final Structures

	UM079	UM249	UM251	WH5C	UM305	AA175	AA183	AA183 (N265Y mutant)
Beamline	ID30B (ESRF)[1]	I02 (DLS)[2]	ID30B (ESRF)	ID30B (ESRF)	ID30B (ESRF)	I03 (DLS)	I03 (DLS)	ID30-A1 (ESRF)
Wavelength (Å)	0.976	0.976	0.976	0.976	0.976	0.976	0.976	0.976
Resolution range	51.86 - 2.3 (2.382 - 2.3)	61.12 - 2.659 (2.754 - 2.659)	57.23 - 2.3 (2.382 - 2.3)	88.99 - 2.062 (2.136 - 2.062)	48.17 - 2.1 (2.175 - 2.1)	71.42 - 1.86 (1.926 - 1.86)	69.48 - 2.1 (2.175 - 2.1)	48.01 - 2.12 (2.196 - 2.12)
Space group	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2
Unit cell a (Å)	74.697	75.455	74.684	74.72	75.393	74.784	75.479	74.988
b (Å)	98.629	99.367	98.677	98.63	99.029	98.624	99.196	98.718
c (Å)	207.438	208.463	207.1	206.36	207.952	207.137	208.439	207.071
Total reflections	279649 (18740)	294369 (2801)	276271 (17387)	342936 (15243)	302969 (16836)	852511 (33682)	612585 (30578)	350341 (18242)
Unique reflections	66343 (4477)	45903 (437)	66028 (4361)	91634 (4501)	80206 (4644)	129298 (6329)	92162 (4483)	87715 (4472)
Multiplicity	4.2 (4.2)	6.4 (6.4)	4.2 (4.0)	3.7 (3.4)	3.8 (3.6)	6.6 (5.3)	6.6 (6.8)	4.0 (4.1)
Completeness (%)	96.7 (98.2)	99.9 (98.0)	96.2 (95.8)	96.7 (96.5)	87.8 (90.1)	100.0 (100.0)	100.0 (99.9)	99.7 (99.8)
Mean I/sigma(I)	11.8 (2.2)	10.1 (2.8)	13.8 (2.4)	11.8 (2.1)	11.6 (2.2)	9.1 (0.8)	7.4 (0.9)	8.8 (1.2)

R-merge	0.075 (0.741)	0.210 (1.052)	0.065 (0.546)	0.058 (0.538)	0.064 (0.496)	0.095 (1.829)	0.186 (0.2.516)	0.096 (1.176)
R-meas	0.086 (0.855)	0.229 (1.146)	0.074 (0.625)	0.067 (0.634)	0.073 (0.568)	0.103 (2.032)	0.202 (2.725)	0.111 (1.353)
R-pim	0.041 (0.412)	0.090 (0.450)	0.044 (0.295)	0.033 (0.325)	0.034 (0.269)	0.040 (0.876)	0.078 (1.040)	0.054 (0.655)
$CC_{1/2}$	0.998 (0.684)	0.994 (0.781)	0.999 (0.765)	0.999 (0.780)	0.999 (0.818)	0.998 (0.393)	0.996 (0.287)	0.998 (0.431)
Reflections used in refinement	66203 (6645)	45808 (4499)	65966 (6428)	91546 (9039)	80113 (8094)	129154 (12752)	92043 (9068)	87609 (8680)
Reflections used for R-free	3320 (297)	2236 (237)	3257 (312)	1999 (198)	4026 (389)	6444 (616)	4544 (422)	2754 (331)
R-work	0.2068 (0.3141)	0.1945 (0.3060)	0.1851 (0.2571)	0.2336 (0.3088)	0.1921 (0.2845)	0.1993 (0.3201)	0.1989 (0.3268)	0.2039 (0.3111)
R-free	0.2355 (0.3306)	0.2278 (0.3406)	0.2129 (0.2893)	0.2542 (0.3531)	0.2151 (0.3084)	0.2194 (0.3504)	0.2314 (0.3406)	0.2464 (0.3230)
Number of non-hydrogen atoms	11384	11235	11552	11755	11804	11870	11521	11605
macromolecules	11083	10999	11031	11039	11035	11014	11026	11006
ligands	126	124	146	182	216	168	228	144
solvent	175	112	375	534	553	688	267	455
Protein residues	1368	1368	1368	1368	1368	1368	1368	1368
RMS (bonds)	0.016	0.016	0.020	0.012	0.018	0.015	0.032	0.013
RMS (angles)	1.76	1.72	1.73	1.52	1.51	1.66	2.07	1.62
Ramachandran favoured (%)	96.84	96.76	97.21	97.57	96.76	97.35	97.28	96.69

Ramachandran allowed (%)	3.16	3.24	2.79	2.43	3.24	2.65	2.72	3.16
Ramachandran outliers (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15
Rotamer outliers (%)	2.33	2.59	1.50	1.92	1.33	1.09	1.84	2.18
Clashscore	3.98	2.56	2.09	4.02	2.35	2.22	2.90	2.69
Average B-factor	46.66	42.78	41.97	44.58	36.68	44.85	45.42	44.59
macromolecules	46.61	42.64	41.84	44.18	36.58	44.58	45.35	44.56
ligands	58.92	62.01	52.48	59.28	35.15	50.37	52.01	45.21
solvent	40.99	35.68	41.72	47.77	39.36	47.73	42.83	45.26
Number of TLS groups	16	21	15	15	15	15	17	12

Statistics for the highest-resolution shell are shown in parentheses.

6.5.2. Additional Data Set Statistics (Refinement in Progress)

	AA175 (N265Y mutant)	EG047
Beamline	ID30B (ESRF)	ID30B (ESRF)
Wavelength (Å)	0.976	0.976
Resolution range	47.78 – 2.35 (2.43 – 2.35)	60.46 - 2.67 (2.77 – 2.67)
Space group	P 21 21 2	P 21 21 2
Unit cell a (Å)	74.654	74.678
b (Å)	98.265	98.385
c (Å)	206.094	206.019
Total reflections	385412 (32737)	268821 (19979)
Unique reflections	63534 (5999)	43221 (3695)
Multiplicity	6.1 (5.5)	6.2 (5.4)
Completeness (%)	99.3 (97.0)	97.9 (81.5)
Mean I/sigma(I)	4.4 (0.6)	5.8 (0.8)
R-merge	0.199 (1.171)	0.358 (4.860)
R-meas	0.218 (1.291)	0.391 (5.370)
R-pim	0.087 (0.533)	0.156 (2.223)
$CC_{1/2}$	0.991 (0.536)	0.990 (0.275)
Reflections used in refinement	63849 (5811)	43134 (3451)
Reflections used for R-free	3203 (281)	2104 (184)
R-work	0.2139 (0.3265)	0.2188 (0.4373)
R-free	0.2625 (0.3520)	0.2882 (0.4695)
Number of non-hydrogen atoms	11156	11168
macromolecules	11016	11008
ligands	104	138
solvent	36	22
Protein residues	1368	1368
RMS (bonds)	0.012	0.011
RMS (angles)	1.90	1.72
Ramachandran favoured (%)	96.47	96.69
Ramachandran allowed (%)	3.53	3.31
Ramachandran outliers (%)	0.00	0.00
Rotamer outliers (%)	0.08	0.17
Clashscore	18.04	9.75
Average B-factor	88.57	52.59
macromolecules	88.71	52.27
ligands	83.20	78.83
solvent	63.39	51.13
Number of TLS groups	20	16

Statistics for the highest-resolution shell are shown in parentheses.

6.6. HPLC Traces of Putative FRET Probes Prior to Deprotection

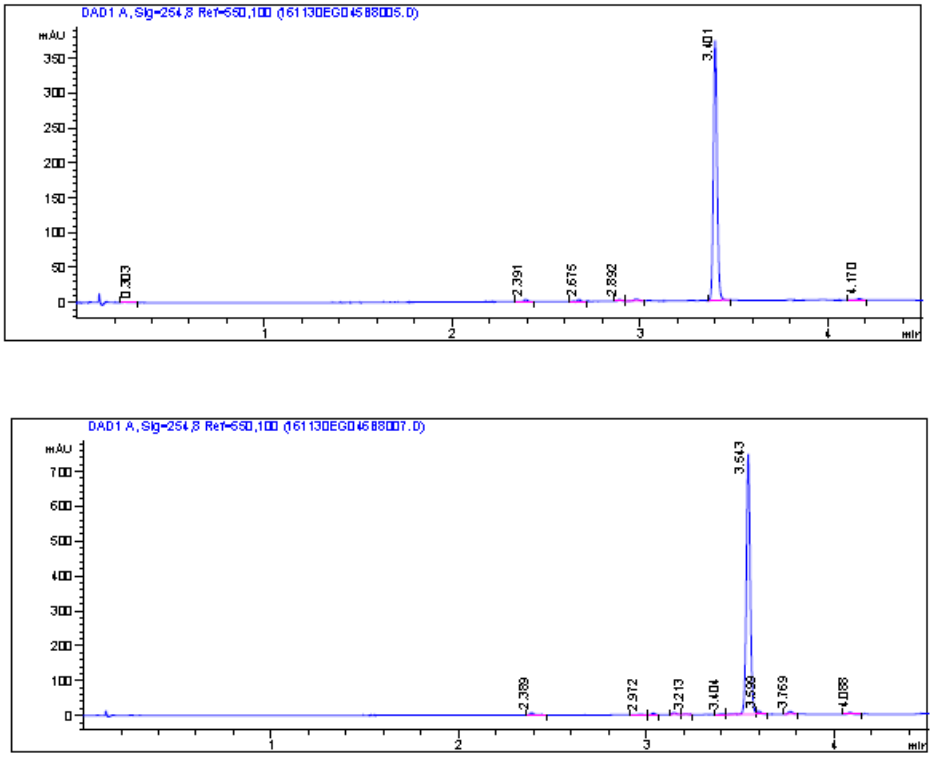


Figure A5: HPLC trace of EG047 (top) and EG048 (bottom) prior to TFA deprotection.

6.7. Evidence for Synthesis of EG047 and EG048

6.7.1. EG047

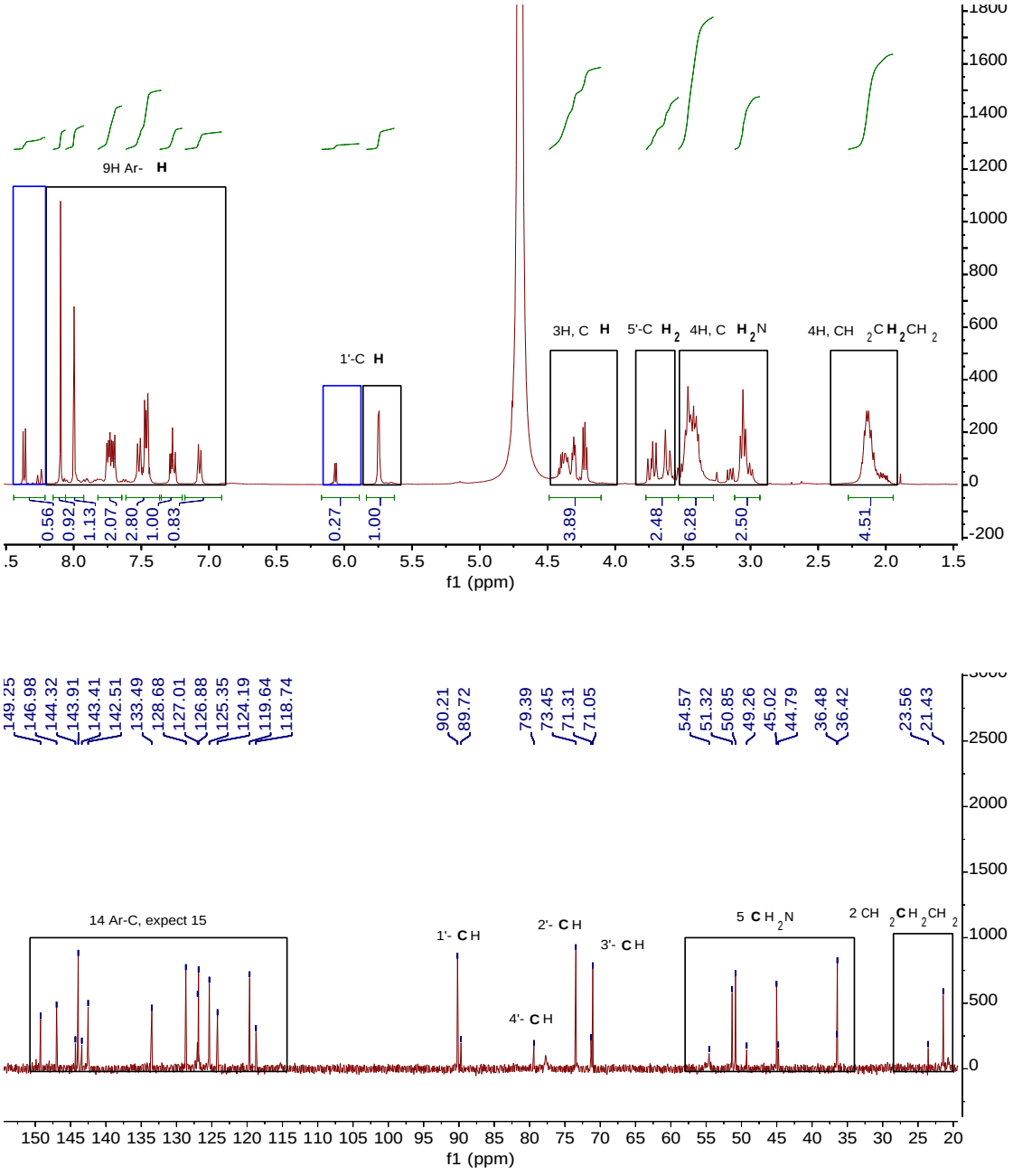


Figure A6: ^{1}H (top) and ^{13}C (bottom) NMR spectra for EG047. Blue boxes are used to indicate peaks of the unknown contaminant. Additional contaminant peaks are present, and overlap with those of EG047.

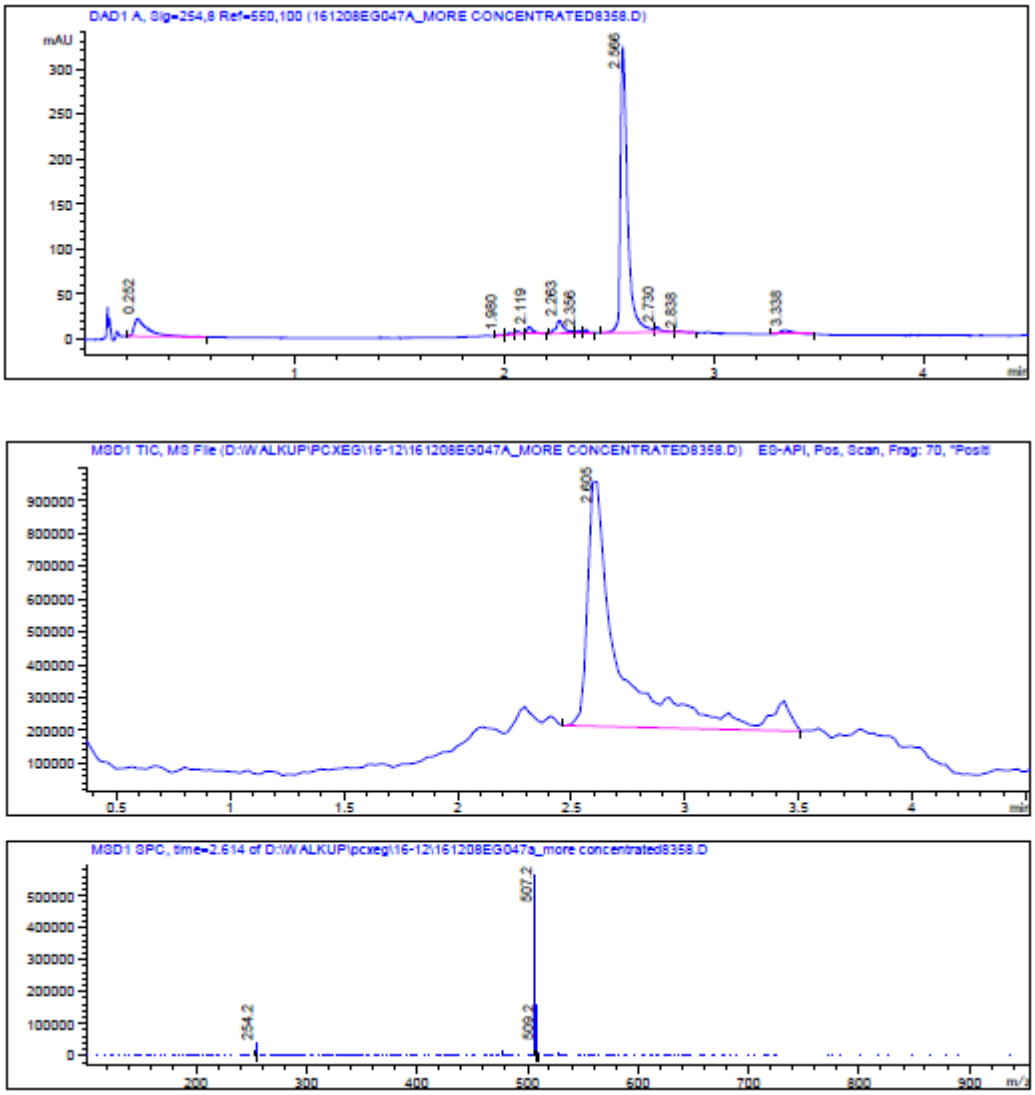


Figure A7: LCMS of EG047. The major peak has mass corresponding to EG047 [M+H$^+$]

6.7.2. EG048

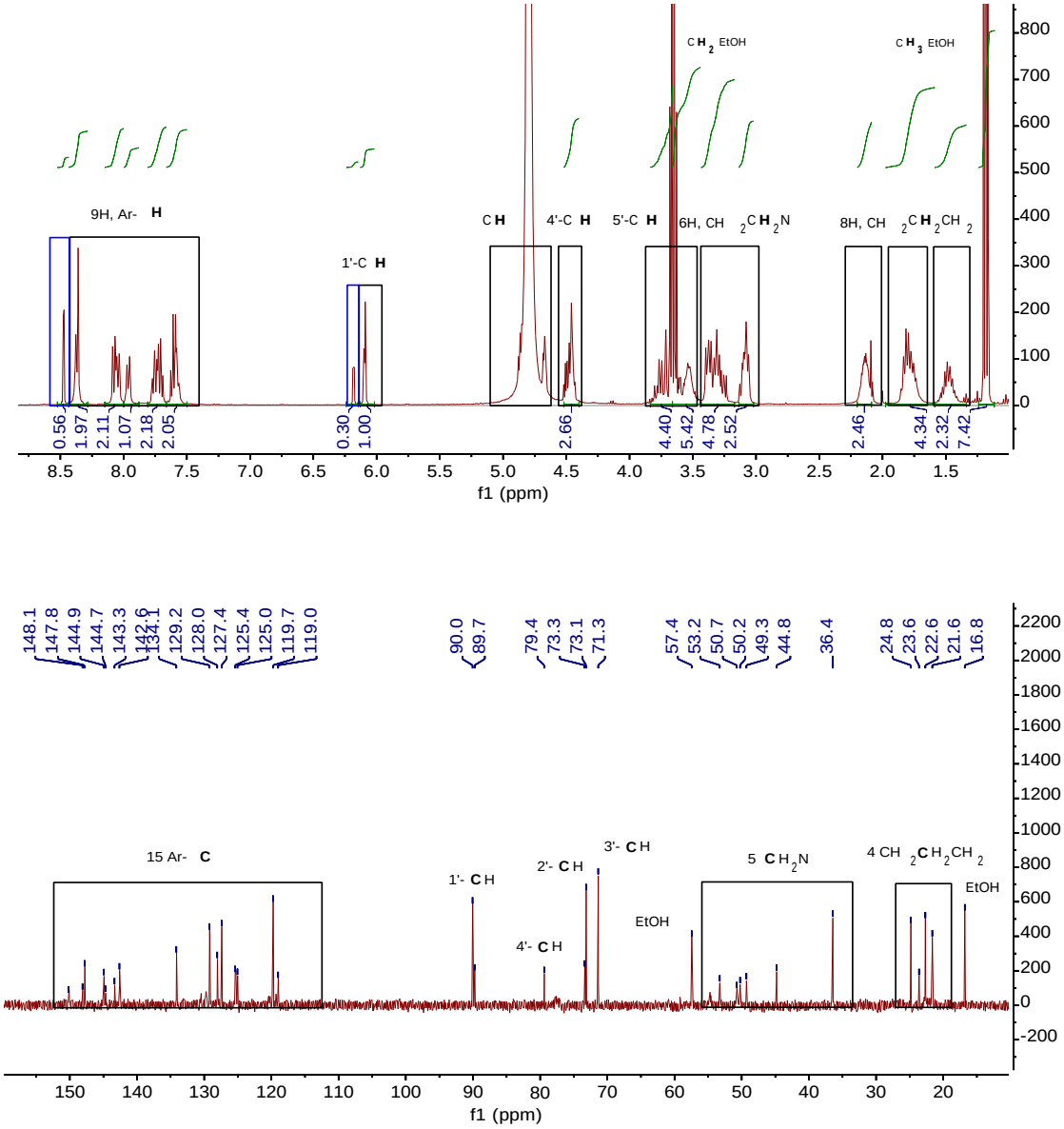


Figure A8: ^{1}H (top) and ^{13}C (bottom) NMR spectra for EG048. Blue boxes are used to indicate peaks of the unknown contaminant.

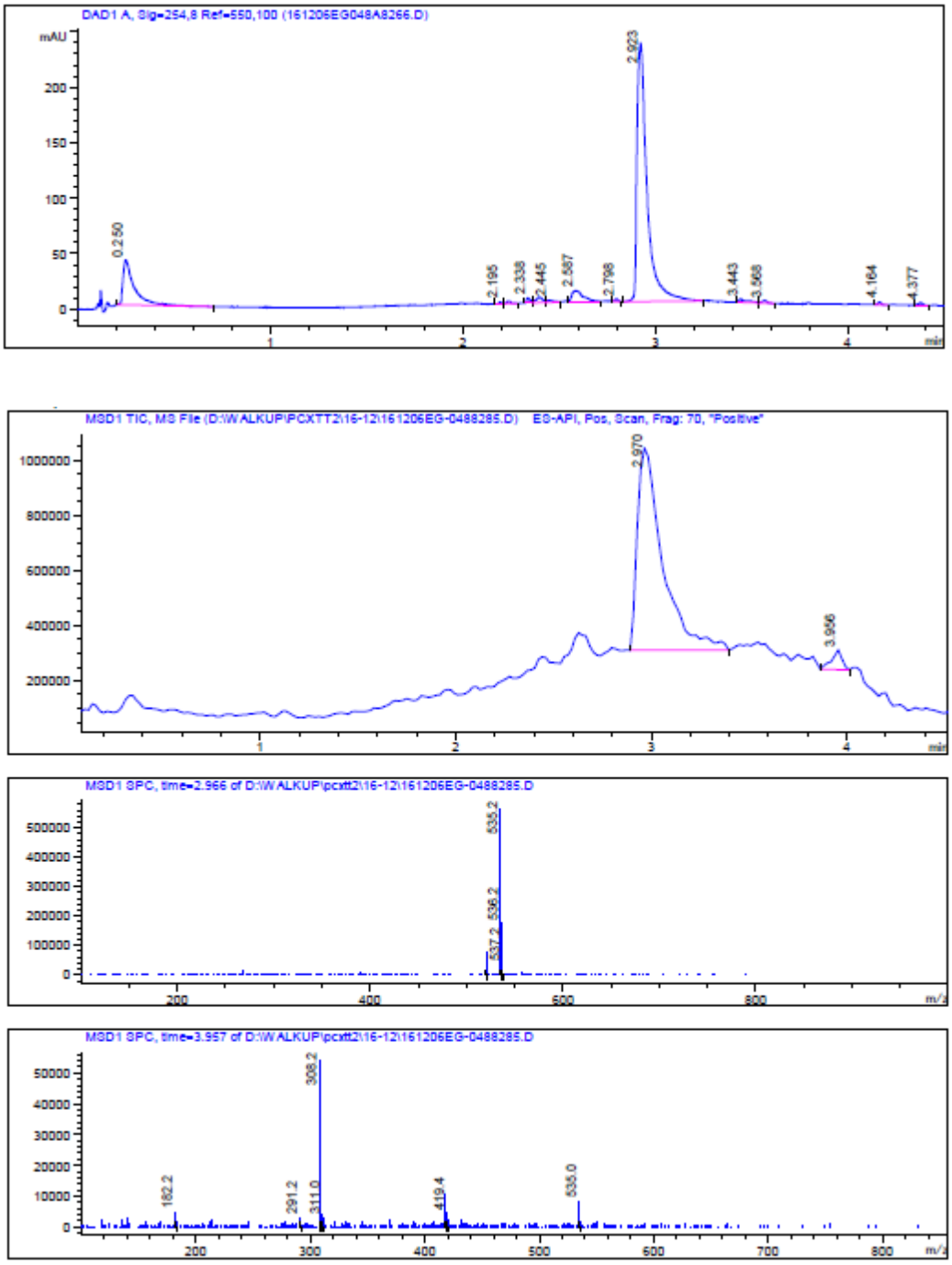


Figure A9: LCMS of EG048. The major peak has mass corresponding to EG048 [M+H$^+$]