

Design, synthesis and evaluation of small molecule inhibitors of human *N*- myristoyltransferase

Mengjie Shen

Doctor of Philosophy Thesis

Institute of Chemical Biology

Department of Chemistry

Imperial College London

London SW7 2AZ

December 2018

Declaration

To the best of my knowledge the research presented in this thesis is original, and my own work, unless otherwise indicated in the text or references.

Mengjie Shen

Dec 2018

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non Commercial No Derivatives 4.0 International Licence (CC BY-NC-ND).

Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

N-myristoyltransferase (NMT) is an enzyme that catalyses the transfer of myristate to the *N*-terminal glycine of protein substrates. In human, two NMT proteins are expressed, playing crucial roles in protein membrane binding and targeting. They are involved in many diseases like cancer and common cold. This project aims to develop human NMT inhibitors as tools for NMT functional study and as potential treatments of *N*-myristoylation dependent diseases.

In subchapter 2.1, five previously developed inhibitors were re-purified and tested for enzyme activity by the CMP assay. All compounds show inhibitions against both human NMT1/2 with IC_{50} s between 1.3-25 μ M, but no inhibition against Lm NMT. The reason for the abnormal curves found in the CPM assay was systematically explored. Inhibitor-CPM adducts were identified and a likely side reaction mechanism was proposed.

In subchapter 2.2, ten inhibitors with a new linker were synthesized based on a developing series. The new linker is generally tolerant, brings better EC_{50}/IC_{50} ratios and the best compound shows an IC_{50} of 0.31 μ M. Crystallography structures of the best compound and its old-linker analogue were solved and used to explore SARs. Tagging assay showed dose-dependent labelling, providing direct evidence of target engagement.

In subchapter 2.3, a target repurposing approach was used to explore nematode NMT inhibitors from existing NMT inhibitors. Eight plasmodium or malaria NMT inhibitors were tested in the enzyme activity study and all showed inhibition. Their SARs were analysed using human NMT structures. Seven compounds were further tested for *in vivo* activities and one showed a higher potency than the positive control.

In Chapter 3, eleven ligand-protein structures of the imidazo[1,2-a]pyridin-6-phenoxy series and nine structures of azetidinopyrimidines series were solved by X-ray crystallography. Extensive binding modes analysis has revealed their key SARs. Multi-conformations are identified and investigated. Insights of crystal obtainment were summarized and discussed.

Acknowledgements

Firstly, many thanks to my supervisor, Ed, for giving me the opportunity to come to the UK and to Imperial College to work on such an interesting and varied project. The project has broadened my view of drug discovery in many dimensions and the years in the UK as my first abroad experience has broadened my life in many ways. You are always encouraging me to learn things that I am interested in and creating the opportunities. Your guidance and support are invaluable.

Secondly, I would like to thank the past and current members of the Tate group for your help and support. Thanks to Andy, Jennie and Jone for your advice on organic synthesis and the CPM assay. I am grateful to Inma and Chris for showing me all the techniques for crystallography. Thank you to Gee for teaching me the tagging assay. I am thankful to Markus for helping me with protein expression and purification.

Thirdly, my gratitude goes to my collaborators: Jim from the Wilkinson group for your advice on the crystallography and providing the NMT proteins and plasmids; Zhiru from the Carlow group for running the *In vivo* activity of *C. elegans* and providing the Ce and Bm NMTs.

Moreover, special thanks go to my proof-readers: Jone, Markus, Andy, Jim and Jennie.

I cannot thank more to my scholarship sponsor, Mr Lee for the generous financial support. It is not easy to get a full scholarship as a Chinese citizen and your donation made everything possible.

Finally, I want to thank my family for the love and support during my PhD, especially during the post-surgery period. Your selfless care keeps me motivated through all the ups and downs.

Abbreviations

Ala or A	Alanine
Arf2p	ADP ribosylation factor 2
Arg or R	Arginine
Asp or D	Aspartic acid
Asn or N	Asparagine
Az	Probe containing an azide functional group
AzTB	Azido-TAMRA-biotin capture reagent
Bm	<i>Brugia malayi</i>
C-	Carboxy-
Ca	<i>Candida albicans</i>
Ce	<i>Caenorhabditis elegans</i>
CHM	Cycloheximide
cLogD	Calculated octanol/water distribution coefficient
cLogP	Calculated octanol/water partition coefficient
CNS	Central nervous system
CoA	Coenzyme A
CPM	7-Diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin)
CuAAC	Copper-catalysed azide-alkyne [3+2] cycloaddition
Cys or C	Cysteine
DCM	Dichloromethane
DMF	<i>N,N</i> -Dimethylformamide
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
EC₅₀	Effective concentration: concentration of compound required to effect 50% of maximum biological activity in a cellular assay
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum

ESI	Electrospray ionisation
Et	Ethyl
GGPP	Geranylgeranyl pyrophosphate
GGTaseI	GGTaseI
Gln or Q	Glutamine
Glu or E	Glutamic acid
Gly or G	Glycine
GPI	Glycosylphosphatidylinositol
FBS	Fetal bovine serum
FFA	Free fatty acid
FPP	Farnesyl pyrophosphate
FTase	Farnesyltransferase
Hh	Hedgehog
Hhat	Hedgehog acyltransferase
His or H	Histidine
HNE	4-hydroxynonenal
HPG	Homopropargylglycine
HPLC	High performance liquid chromatography
Hs	<i>Homo Sapiens</i>
HRMS	High resolution mass spectrometry
HTS	High-throughput screening
Hz	Hertz
IC₅₀	Inhibitory concentration: concentration of inhibitor required to inhibit 50% of enzyme activity
Ile or I	Isoleucine
K_d	Dissociation constant: measured equilibrium constant describing the affinity of a ligand for a receptor.
K_i	Inhibition constant: calculated constant of inhibition which is independent of biological assay
K_m	Michaelis constant: concentration of substrate at which reaction rate is half the maximum rate
LB media	Lauria-Bertani broth media
LC-MS	Liquid chromatography – mass spectrometry

Ld	<i>Leishmania donovani</i>
LDE	lipid-derived electrophiles
Leu or L	Leucine
Lys or K	Lysine
Lm	<i>Leishmania major</i>
LPCAT1	Lysophosphatidylcholine acyltransferase
M	Molar
Me	Methyl
MeOH	Methanol
Met or M	Methionine
MRCT	Medical Research Council Technology
MTS	3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2 <i>H</i> -tetrazolium inner salt
MW	Molecular weight
Myr	Myristoyl
<i>m/z</i>	Mass to charge ratio
<i>N</i>-	Amino-
NHM	Non-hydrolysable myristoyl COA
NMR	Nuclear magnetic resonance
NMT	<i>N</i> -Myristoyltransferase
OD600	Optical density at 600 nm
PAT	Palmitoyl acyltransferase
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PEG	Polyethyleneglycol
Pf	<i>Plasmodium falciparum</i>
Ph	Phenyl
Phe or F	Phenylalanine
PMS	Phenazine methosulfate
ppm	Parts per million
Pro or P	Proline
PTM	Post-translational modification
Pv	<i>Plasmodium vivax</i>

R_f	Retention factor
R_t	Retention time
RNA	Ribonucleic acid
RNAi	RNA interference
rt	Room temperature
rpm	Revolutions per minute
SAR	Structure activity relationship
Ser or S	Serine
Sc	<i>Saccharomyces cerevisiae</i>
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SMO	Smoothened
spp.	Species
SPR	Surface plasmon resonance assay
STD	Saturation transfer difference spectroscopy
TAMRA	Carboxy-tetramethylrhodamine
Tb	<i>Trypanosoma brucei</i>
TBTA	Tris[(1-benzyl-1H-1,2,3- triazol-4-yl)methyl]amine
TC₅₀	Tagging Inhibitory concentration: concentration of inhibitor required to inhibit 50% of tagging
TCA cycle	Tricarboxylic acid cycle
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	<i>N,N,N',N'</i> -Tetramethylethylenediamine
Thr or T	Threonine
Tris	Tris(hydroxymethyl)aminomethane
Trp or W	Tryptophan
Tyr or Y	Tyrosine
Val or V	Valine
V_{max}	Maximum velocity
WaterLOGSY	Water-ligand observed by gradient spectroscopy
Yn	Probe containing an alkyne functional group
YnC12	Tetradec-13-ynoic acid

Contents

1	Introduction	1
1.1	Background	1
1.1.1	Post-translational modification	1
1.1.2	Protein lipidation.....	1
1.2	N-myristoylation	7
1.2.1	Mechanism and structure	7
1.2.2	Role of myristoylation	12
1.3	Human NMTs.....	14
1.3.1	Isoforms	14
1.3.2	Human NMTs as potential drug targets	15
1.4	NMT inhibitors.....	19
1.4.1	Fungal NMT inhibitors	19
1.4.2	Parasitic NMT inhibitors.....	25
1.4.3	Human NMT inhibitors.....	32

1.5	Objective	35
2	Medicinal Chemistry	36
2.1	Dihydrothiazole series.....	36
2.1.1	Previous work on dihydrothiazole series	36
2.1.2	CPM assay	38
2.1.3	Validation.....	39
2.1.4	Abnormal inhibition curves	41
2.1.5	Comparison assay adding inhibitors after quenching the reaction	42
2.1.6	Continuous CPM assay with or without enzyme.....	43
2.1.7	Comparison CPM assay with or without CPM.....	44
2.1.8	Proposed mechanism	45
2.1.9	Oxygen substitute.....	47
2.1.10	Conclusion and outlook.....	48
2.2	Azetidinopyrimidines series.....	49
2.2.1	Previous work	49
2.2.2	Crystal structure analysis	50

2.2.3	Change of the azetidine linker	51
2.2.4	Change of the trimethylpyrazole tail.....	53
2.2.5	Change of the diazepane head.....	55
2.2.6	Change of the halogen	56
2.2.7	The MTS assay	57
2.2.8	<i>In vitro</i> activity analysis.....	58
2.2.9	Crystal structure of IMP-0001096	61
2.2.10	Inhibitor validation.....	63
2.2.11	Conclusion and outlook.....	67
2.3	Nematode NMT.....	69
2.3.1	Introduction.....	69
2.3.2	Sequence analysis of NMT proteins	71
2.3.3	Kinetic study	73
2.3.4	Enzyme activity	74
2.3.5	<i>In vivo</i> activity.....	77
2.3.6	Conclusion and outlook	82

3	Crystallography	84
3.1	Workflow	84
3.1.1	Protein expression.....	84
3.1.2	Protein purification	85
3.1.3	Crystallization.....	87
3.1.4	Data collection	89
3.1.5	Data processing.....	90
3.1.6	Refinement and model building.....	90
3.1.7	Conclusion and outlook	91
3.2	Imidazo[1,2-a]pyridin-6-phenoxy series structure insights	93
3.2.1	IMP-0001029, IMP-0001114 and IMP-0001343	96
3.2.2	IMP-0001117, IMP-0001320 and IMP-0001336	99
3.2.3	IMP-0001222, IMP-0001223 and IMP-0001242	101
3.2.4	IMP-0001161 and IMP-0001180.....	102
3.2.5	C-terminus conformation	108
3.2.6	Conclusion and outlook	110

3.3	Azetidinopyrimidines series structure insights	111
3.3.1	Two conformations	114
3.3.2	Structure-activity relationship analysis	117
3.3.3	Conclusion and outlook	123
4	Conclusion and perspectives	125
5	Methods	127
5.1	Chemistry	127
5.1.1	General methods	127
5.1.2	Synthesis and characterization of compounds	128
5.2	Biology	140
5.2.1	General methods	140
5.2.2	Sequence alignment	140
5.2.3	Enzyme assay	144
5.2.4	Cell culture	146
5.2.5	MTS assay	147
5.2.6	Metabolic tagging assay	148

5.3	Crystallography	152
5.3.1	General methods	152
5.3.2	Transformation.....	152
5.3.3	Expression.....	153
5.3.4	Purification.....	153
5.3.5	Crystallization	155
5.3.6	Data processing.....	157
6	References	163
7	Appendix	176
7.1	Permission documents for figure reproduction	176
7.1.1	Figure 1	176
7.1.2	Figure 3	177
7.1.3	Figure 25	178
7.1.4	Figure 52 and Figure 53	178

Figures

Figure 1 Lipid metabolism regulates protein functions through protein lipidation	2
Figure 2 Structures of example lipids found in protein lipidation.	4
Figure 3 GPI anchors and its protein modification.....	5
Figure 4 Mechanism of <i>N</i> -myristoylation.....	7
Figure 5 Cartoon representation of the NMT catalytic cycle	7
Figure 6 Myristoyl-CoA binding in a "Question Mark" conformation in the NMT pocket.....	8
Figure 7 The Ab loop of Hs NMT1	9
Figure 8 The first mechanism hypothesis of Sc NMT catalysed <i>N</i> -myristoylation ⁸	10
Figure 9 New <i>N</i> -myristoylation mechanism based on Hs NMT1 crystal structures	11
Figure 10 Human NMT isoforms	14
Figure 11 Plot of dose-response activity against Hs NMT1 vs. Hs NMT2 for the primary screening hits and their analogues	15
Figure 12 Evidence of being a potential cancer target.....	16
Figure 13 Development of peptide mimic inhibitor of Ca NMT ⁸⁶	20
Figure 14 Development of Pfizer benzo[d]thiazole series inhibitor of Ca NMT ⁸⁰	22

Figure 15 Development of Roche benzo[d]furan series inhibitor of Ca NMT ⁷¹	23
Figure 16 A schematic drawing of the binding mode of C10 (an analogy of C8) in Ca NMT	24
Figure 17 Development of SSP benzo[d]thiazole series inhibitor of Ca NMT ⁷²	25
Figure 18 Development of Dundee pyrazole sulphonamide series inhibitor of Tb NMT	26
Figure 19 Development of Imperial College benzothiazole series	27
Figure 20 A schematic drawing of the binding mode of C23 (an analogy of C17) based on the crystal structure of the ternary complex of Pv NMT, Myr-CoA and C23	28
Figure 21 Development of Imperial College quinoline series	29
Figure 22 A schematic drawing of the binding mode of C26 based on the crystal structure of the ternary complex of Pv NMT, Myr-CoA and C26	30
Figure 23 The structure and IC ₅₀ s of the Pfizer screen hits	30
Figure 24 Development of Imperial College hybrid series.....	31
Figure 25 Development of the pyrrolo[1,2-a]pyrazine series	32
Figure 26 Development of Imperial College indazole series.....	33
Figure 27 Previous data of 20F4	36
Figure 28 Mechanism of CPM assay	38

Figure 29 Dose-response curves of four 20F4 analogues on Hs NMT1, Hs NMT2 and Lm NMT (inhibitor concentration in μM)	41
Figure 30 Comparison assay that adds inhibitors 20F4-12 and 20F4-13 after quenching the CPM reaction	42
Figure 31 Continues CPM assay with or without enzyme.....	43
Figure 32 Comparison assay with or without CPM on 20F4-13	44
Figure 33 High-performance liquid chromatography mass spectrometry results of CPM and inhibitor mixtures.....	45
Figure 34 Proposed mechanism of the side-reaction using 20F4-12 as an example.	46
Figure 35 Synthesis of (2-((2,6-dimethylphenyl)amino)-4-methyl-4,5-dihydrooxazol-4-yl)methanol (20F4-O).....	47
Figure 36 comparison assay between 20F4 and its oxygen substitute 20F4-O	47
Figure 37 Previous development of azetidinopyrimidines series.	49
Figure 38 Binding mode of IMP-0000898 in Hs NMT1	50
Figure 39 linker change of the azetidinopyrimidines series	51
Figure 40 Synthesis of the new azetidinopyrimidines series with the piperidinyl carbamate linker	52
Figure 41 Trimethylpyrazole tail change of the new azetidinopyrimidines series.	53

Figure 42 Diazepane head change of the new azetidinopyrimidines series.....	55
Figure 43 Halogen change of the new azetidinopyrimidines series	56
Figure 44 The workflow and mechanism of the MTS assay	57
Figure 45 Comparison of the binding modes of IMP-0001096 and IMP-0000180	61
Figure 46 The workflow and mechanism of the metabolic tagging assay.....	63
Figure 47 The SDS-PAGE result of the tagging assay in DMEM	64
Figure 48 The SDS-PAGE result of the tagging assay in Met (-) DMEM	65
Figure 49 Structure-based molecule development of the new azetidinopyrimidines series	67
Figure 50 The disabling lymphatic filariasis and onchocerciasis and the life circles of their causing parasites	69
Figure 51 Sequence analysis of relevant NMTs	71
Figure 52 The structures of NMT inhibitors tested in the enzyme activity assay and the <i>in vivo</i> assay.....	75
Figure 53 Inhibitor binding in Hs NMT1. Images created by PyMOL ²⁸	76
Figure 54 The life cycle of <i>C. elegans</i> at 22°C.....	78
Figure 55 Food consumption measured by OD ₆₀₀ of the L4 larvae treated with inhibitors	79

Figure 56 Food consumption measured by OD ₆₀₀ of the L1 larvae treated with IMP-0001031	82
Figure 57 Protein sequence alignment of Hs NMT1 (Hs 1_Full) and the modified version without <i>N</i> -terminus (Hs1_DeN).....	84
Figure 58 Purification workflow of Hs NMT1.	85
Figure 59 The SDS gel of Hs NMT1 purification	86
Figure 60 The SDS gel of His-tag presence analysis.....	86
Figure 61 Examples of crystals growing in sitting-drops	88
Figure 62 Examples of mounted crystals ready for data collection, with example diffraction images illustrating the raw data obtained.....	89
Figure 63 Electron density maps before building the ligand models.....	90
Figure 64 Structures and activities of IMP-0001029 , IMP-0001114 and IMP-0001343	96
Figure 65 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot ⁺ v1.4.5 ¹⁶¹	97
Figure 66 Structures and activities of IMP-0001117 , IMP-0001320 and IMP-0001336	99
Figure 67 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot ⁺ v1.4.5 ¹⁶¹	99

Figure 68 Binding modes of IMP-0001117 , IMP-0001320 and IMP-0001336 in Hs NMT1	100
Figure 69 Structures and activities of IMP-0001222 , IMP-0001223 and IMP-0001242	101
Figure 70 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot+ v1.4.5 ¹⁶¹	101
Figure 71 Binding modes of IMP-0001117 , IMP-0001222 , IMP-0001223 and IMP-0001242 in Hs NMT1	102
Figure 72 Structures and activities of IMP-0001161 and IMP-0001180	102
Figure 73 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot+ v1.4.5 ¹⁶¹	103
Figure 74 Binding modes comparison of IMP-0001161 with IMP-0001117	104
Figure 75 Binding mode comparison of IMP-0001161 , IMP-0001088 and IMP-0001117	105
Figure 76 Binding mode of IMP-0001161 at the <i>N</i> -methylethanamine end.....	106
Figure 77 Binding modes comparison of IMP-0001180 with IMP-0001114 (A) or IMP-0001222 (B)	107
Figure 78 Two conformations of the <i>C</i> -terminus with different ligand binding	109
Figure 79 Azetidinopyrimidines series structures with solved X-ray crystal structures. Structures are sorted by structural similarity and colored by different conformations.....	114

Figure 80 Two different binding conformations among azetidinopyrimidines series	114
Figure 81 The two conformations of IMP-0000180 in chain A of the human NMT1 protein	116
Figure 82 Structures and activities of IMP-0001171 , IMP-0001210 and IMP-0001232	117
Figure 83 Binding modes of IMP-0001171 , IMP-0001210 and IMP-0001232 highlighting the interactions with Ser405	118
Figure 84 Structures and activities of IMP-0000927 , IMP-0000899 and IMP-0000898	119
Figure 85 Binding modes of IMP-0000927 , IMP-0000899 and IMP-0000898 highlighting the interaction with Ser405	119
Figure 86 Structures and activities of IMP-0001227 , IMP-0000180 and IMP-0001096	120
Figure 87 Binding modes of IMP-0001227 , IMP-0000180 and IMP-0001096	121
Figure 88 Structures and activities of IMP-0000366 , IMP-0000974 and IMP-0001169	122
Figure 89 Binding modes of IMP-0000366 , IMP-0000974 and IMP-0001169	122
Figure 90 Converting 2D coordinates to scattering vectors.....	158
Figure 91 Positional refinement.....	159
Figure 92 Post-refinement.....	160

1 Introduction

1.1 Background

1.1.1 Post-translational modification

Protein post-translational modifications (PTMs) include a host of reversible or irreversible chemical transformations on proteins after or sometimes during ribosomal synthesis. PTMs currently found include the covalent attachment of small inorganic molecules (e.g. phosphorylation), carbohydrates (e.g. glycosylation), small proteins (e.g. ubiquitination) and lipids¹. These PTMs regulate protein function, activity and localization within the cell.

1.1.2 Protein lipidation

Protein lipidation is unique amongst the hundreds of classes of known protein PTMs, which enables direct interaction with cell membranes, ranging from constitutive and stable anchors that can withstand several rounds of endosomal recycling to transient membrane binders that permit rapid switching of subcellular localization². Lipidation modulates the function of targeted proteins by increasing their membrane binding affinities, thus rapidly switching their subcellular localization and association with other proteins, as well as affecting protein folding and stability³. Protein lipidation occurs in every form of life and has developed into its most sophisticated forms in eukaryotes, where lipidated protein families regulate vesicular trafficking pathways and membrane-bound signalling platforms². These PTMs also play an important role in diseases. Many enzymes involved in installing and processing protein lipidation have been targeted for drug discovery and resulted in a number of clinical trials².

Proteins can be covalently modified by at least six classes of lipids (Figure 1): fatty acids, lipid-derived electrophiles (LDEs), glycosylphosphatidylinositol (GPI) anchors, isoprenoids, sterols and phospholipids³.

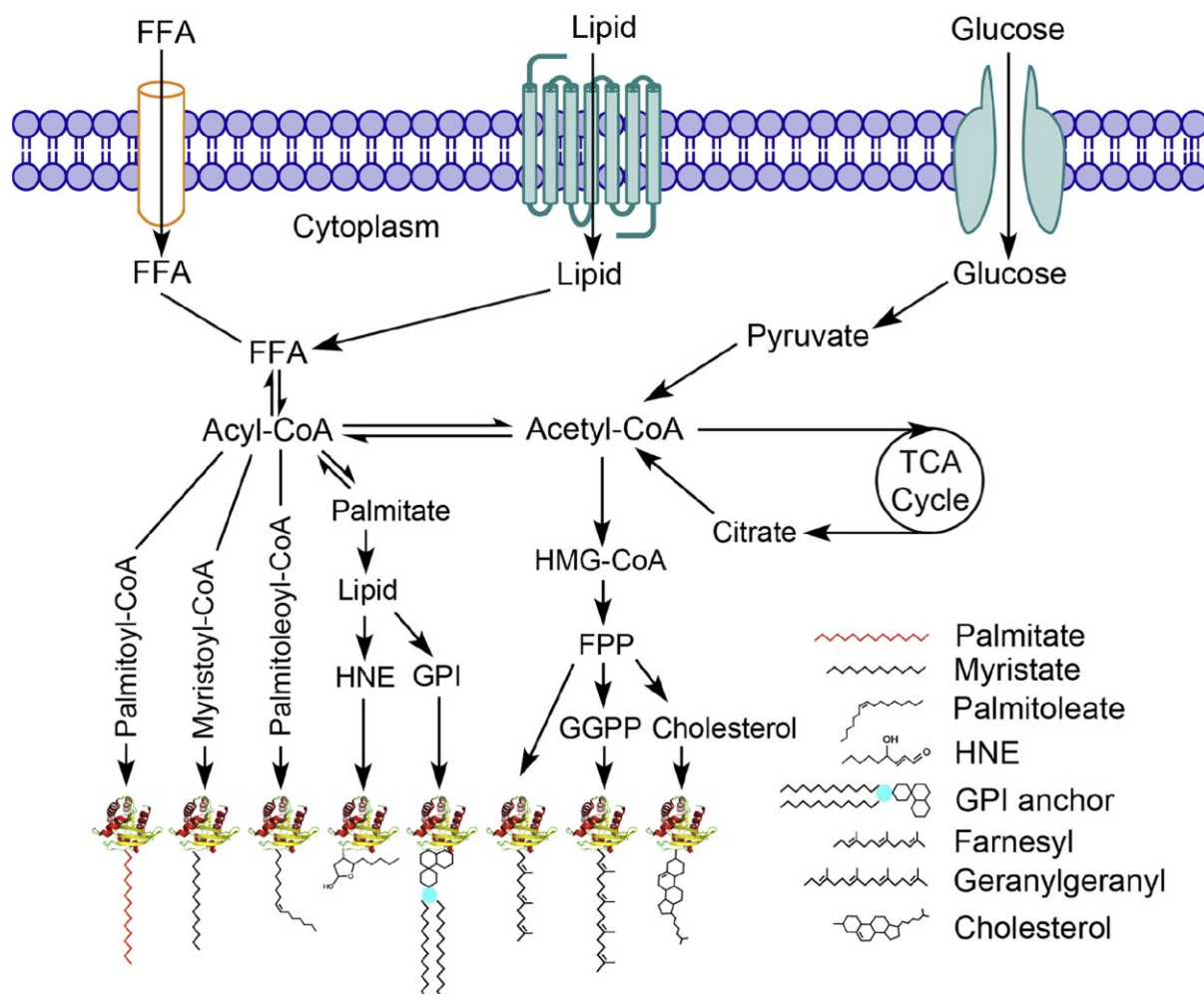


Figure 1 Lipid metabolism regulates protein functions through protein lipidation (Image taken from Chen *et al*³ with permission). Fatty acyl-CoA and other lipid derivatives are generated from cellular lipid metabolism and later used as substrates for protein lipidation. Lipid metabolism affects the availability of these lipids, including saturated and unsaturated fatty acids (palmitate, myristate, palmitoleate, etc.), lipid-derived electrophiles (HNE, etc.), GPI anchors, isoprenoids (farnesyl and geranylgeranyl), cholesterol and phospholipids (not shown here). Palmitate, a 16-carbon fatty acid, services as an intermediate for the biosynthesis of other lipids in the cells. FFA: free fatty acid; HNE: 4-hydroxynonenal; GPI: glycosylphosphatidylinositol; FPP: farnesyl pyrophosphate; GGPP: geranylgeranyl pyrophosphate; TCA cycle: tricarboxylic acid cycle or Krebs cycle.

Saturated and unsaturated fatty acids can be attached to the cysteine, serine or lysine residues of proteins, through a process called fatty acylation. *N*-myristoylation attaches the 14-carbon myristate (Figure 2) to the *N*-terminal glycine of its substrate proteins by *N*-myristoyltransferase (NMT), which is the main focus of this thesis and will be discussed in detail in 1.2. Another major form of fatty acylation, *S*-palmitoylation or *S*-acylation, attaches

the 16-carbon palmitate (Figure 2) or other fatty acids (e.g. the 18-carbon stearic acid) to cysteine residues of its substrate proteins. In humans, the *S*-palmitoylation on intracellular proteins is catalysed by a family of 23 palmitoyl acyltransferases (PATs) with a conserved sequence Asp-His-His-Cys (DHHC)⁴. *S*-palmitoylation is reversible by breaking the resulting thioester bond between palmitate and cysteine under the catalysation of thioesterases. Therefore, *S*-palmitoylated proteins can switch between acylated and diacylated forms in response to upstream signals⁵. More rarely, these fatty acyl groups can be attached to serine residues, namely *O*-palmitoylation, or to the *N*-terminus, namely *N*-palmitoylation. Only a few *O*- or *N*-palmitoylated proteins have been found. Mediated by membrane-bound *O*-acyltransferase called Porcupine, a conserved serine residue of Wnt proteins are modified by a monounsaturated fatty acid, *cis*-palmitoleic acid (Figure 1)⁶. Ser45 residue on histone H4 can be palmitoylated by lysophosphatidylcholine acyltransferase 1 (LPCAT1)⁷. By ghrelin *O*-acyltransferase, the serine residue of the signalling peptide ghrelin can be modified by the 8-carbon octanoyl group⁸. Catalysed by Hedgehog acyltransferase (Hhat), the Hedgehog (Hh) proteins are an example of *N*-palmitoylated proteins, on which the *N*-terminus of the signalling peptide is palmitoylated⁹.

LDEs, such as 4-hydroxy-2(E)-nonenal (HNE, Figure 2)¹⁰, are a group of endogenous reactive metabolites produced by non-enzymatic lipid oxidation/nitration or enzymatic oxidation mechanisms¹¹. Mis-regulation of lipid metabolism often leads to LDE accumulation. LDEs can attack nucleophilic residues of proteins (e.g. cysteine, lysine and histidine) and form covalent adducts with via Michael addition¹². LDE modification is considered to be involved in many chronic diseases, such as neurodegeneration¹³, atherosclerosis¹⁴, inflammation¹⁵ and cancer¹⁶.

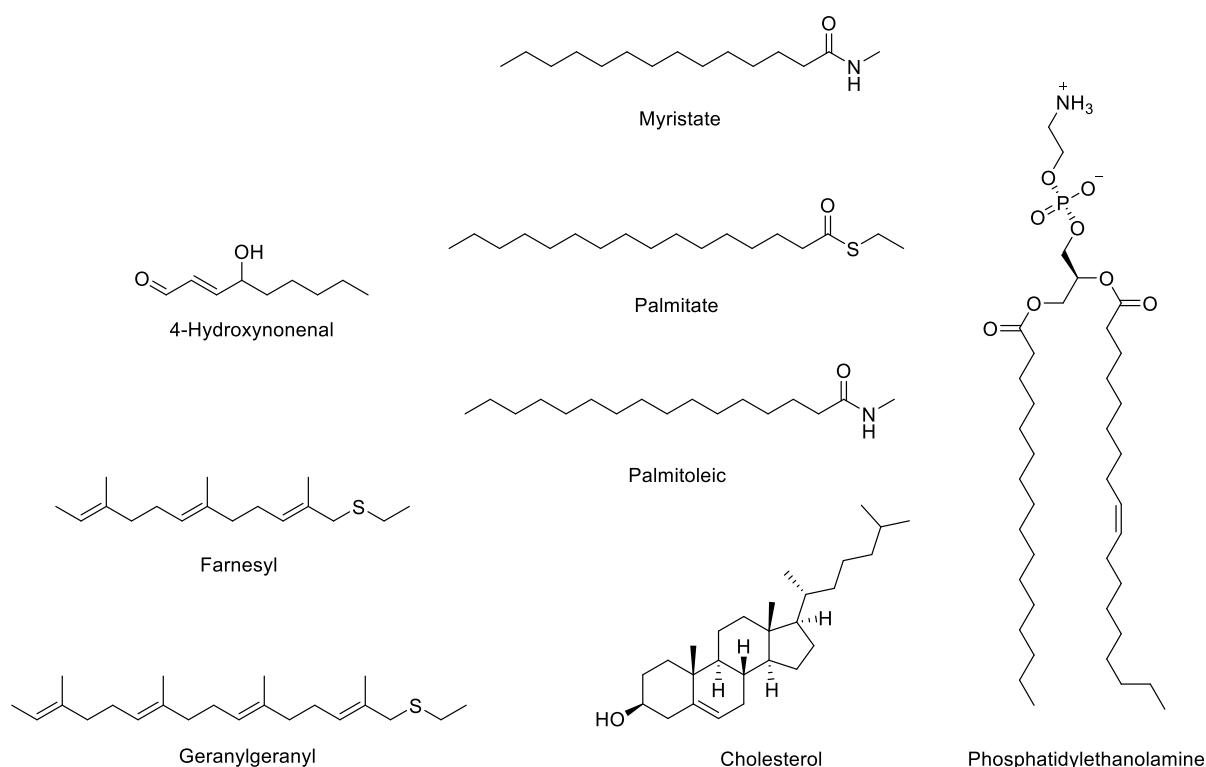


Figure 2 Structures of example lipids found in protein lipidation. There are at least six types of lipids including fatty acids (e.g. myristate, palmitate and palmitoleic), lipid-derived electrophiles (LDEs, e.g. 4-hydroxynonenal), glycosylphosphatidylinositol (GPI) anchors (example structure can be found in Figure 3), isoprenoids (e.g. farnesyl and geranylgeranyl), sterols (e.g. cholesterol) and phospholipids (e.g. phosphatidylethanolamine).

GPI anchor (Figure 3), a complex glycolipid composed of phosphatidylinositol and the carbohydrate residues, glucosamine and mannose, can be covalently attached to C-terminus of proteins via PTMs. About 1% of the eukaryotic proteins are GPI anchor modified¹⁷. GPIs are synthesized in the endoplasmic reticulum (ER) and attached to proteins by the GPI transamidase complex¹⁸. In addition to targeting proteins to the membrane, GPI anchors also have many other biological functions, including mediating immune response and inflammation¹⁹.

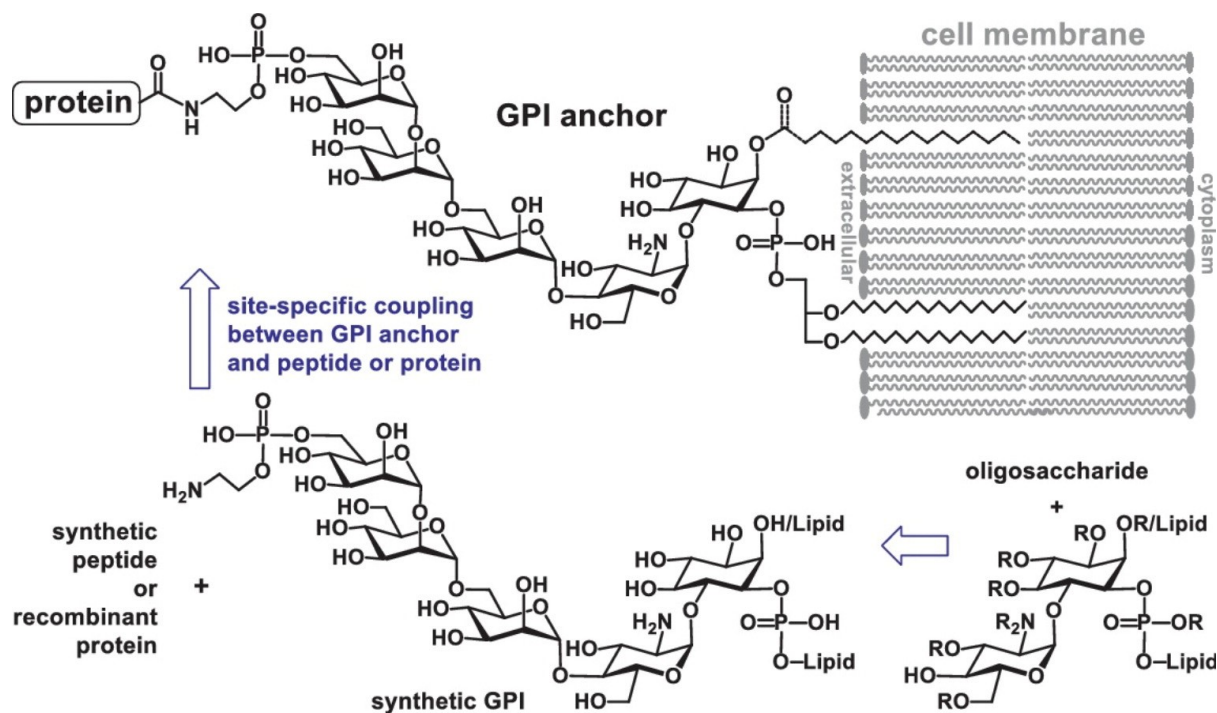


Figure 3 GPI anchors and its protein modification (image taken from Yu et al¹⁸ with permission). Glycosylphosphatidylinositols (GPIs) are synthesized on/in the endoplasmic reticulum (ER) and attaches to the protein C-terminus during PTM by their phosphoethanolamine group. In this way, GPIs anchor surface proteins to the cell membrane, as one of GPIs' most apparent functions.

Attachment of isoprenoids onto a Cys at or near the C-terminus of proteins is known as prenylation⁴. Most of the substrate proteins of prenylation contain a Cys-aliphatic-aliphatic-X (CAAX) motif on their C-terminus, where the consensus Cys residue can be modified by a 15-carbon farnesyl (Figure 2) by farnesyltransferase (FTase) or a 20-carbon geranylgeranyl group (Figure 1) by geranylgeranyl transferase I (GGTaseI). The "X" residue determines whether the protein will be farnesylated or geranylgeranylated. The "AAX" sequence is cleaved after prenylation and the prenylated Cys on the C-terminal is carboxymethylated. The Ras family, including *K*-Ras, *H*-Ras and *N*-Ras are well-studied prenylated proteins²⁰. Prenylation alters the membrane affinity of its substrate proteins and is generally considered as irreversible.

Covalent modification by cholesterol (Figure 2) is uncommon. To date, Hedgehog proteins are the only well-characterized cholesterol-modified proteins. After cleavage, cholesterol forms an ester bond at their C-terminal glycine. New chemical probes have been developed

recently for identification of other cholesterol-modified proteins²¹. Smoothened (SMO), a co-receptor of the Hh pathway, is found to be cholesterol modified on its Asp95 residue²².

Although phospholipids play important roles in membrane formation and signal transduction, phospholipid-modified proteins are rarely identified. Modified by phosphatidylethanolamine (Figure 2), autophagy-related protein Atg8/LC3 is the only known example, which is mediated by multi-step conjugation processes and essential for the formation of autophagosome's double-membrane²³.

1.2 N-myristoylation

1.2.1 Mechanism and structure

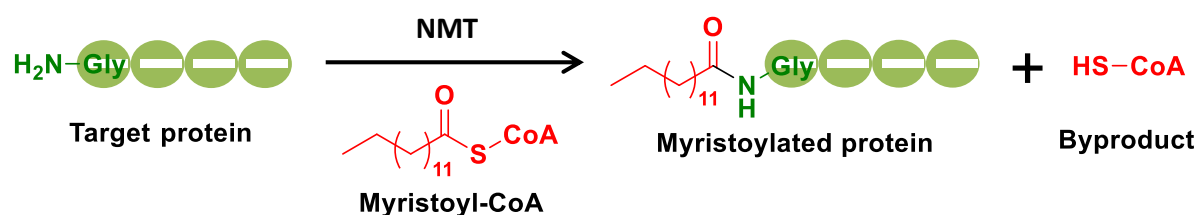


Figure 4 Mechanism of *N*-myristoylation

N-myristoylation is the transfer of myristate, a 14-carbon saturated fatty acid, from myristoyl coenzyme A to the *N*-terminal glycine of a range of protein substrates (Figure 4). It is found to be ubiquitous in eukaryotes²⁴. Approximately 0.5% of all eukaryotic proteins are *N*-myristoylated²⁵. The enzyme catalyses this reaction is *N*-myristoyltransferase.

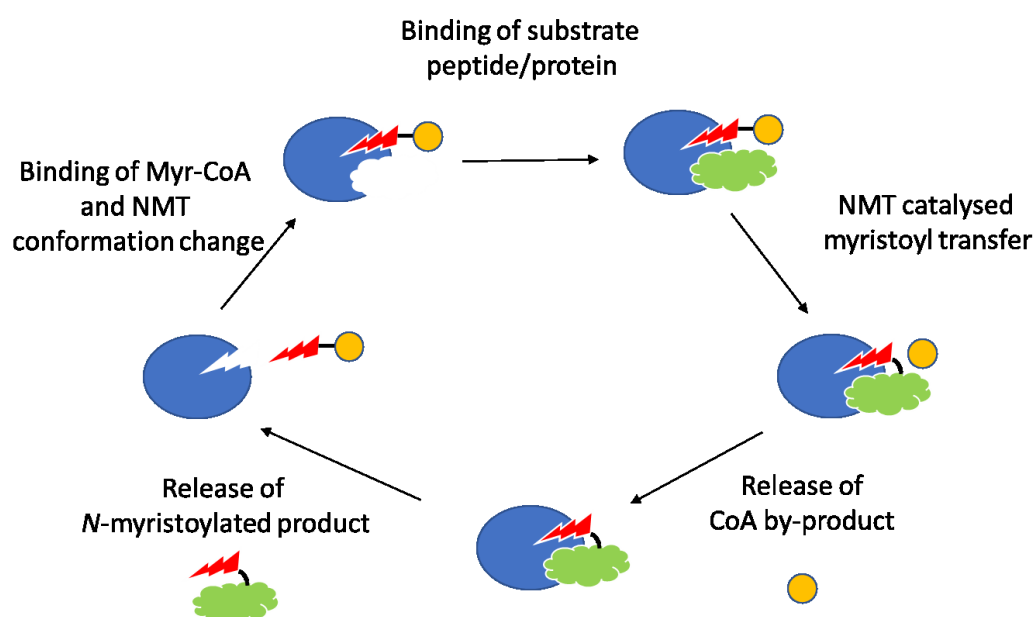


Figure 5 Cartoon representation of the NMT catalytic cycle. NMT is shown in blue, CoA in yellow, myristoyl in red and substrate peptide/protein in green.

Studies by Gordon *et al.* revealed the catalytic cycle of NMT and named it “Bi-Bi” mechanism²⁶⁻²⁷. The apo-enzyme first binds with the myristoyl-CoA, which induces conformational changes of the enzyme and opens a peptide binding pocket adjacent to the

myristoyl-CoA (Figure 5). The target protein or peptide substrate then binds to the enzyme which catalyses the transfer of myristate to the protein or peptide substrate, releasing CoA-SH as the by-product. Finally, the myristoylated protein or peptide is released from the enzyme and the apo-enzyme can start the next catalytic cycle.

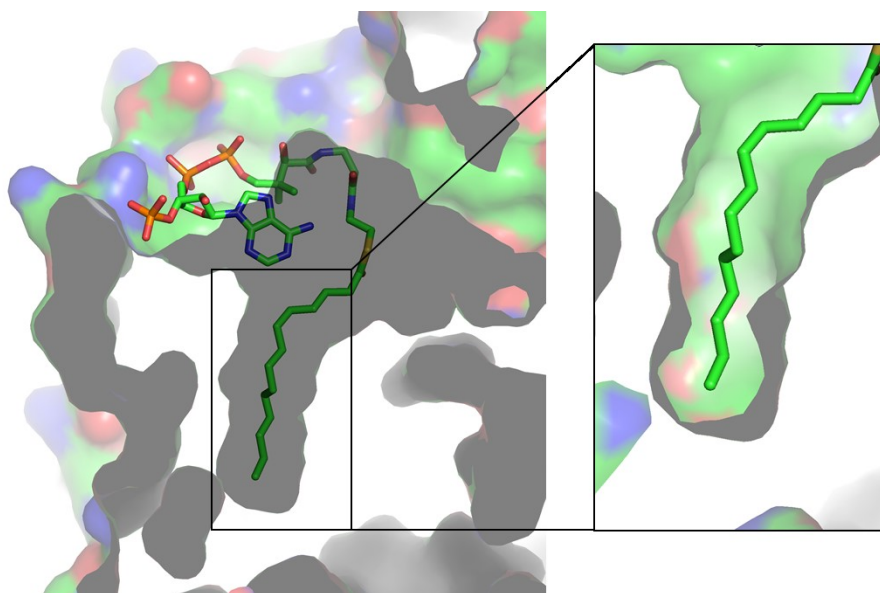


Figure 6 Myristoyl-CoA binding in a "Question Mark" conformation in the NMT pocket. Image generated using PyMOL²⁸ from the ternary structure of Hs NMT1 in complex with myristoyl-CoA and an inhibitor (PDB: 5O6H). Myr-CoA is shown as stick and Hs NMT1 is shown as surface at 50% transparency. Carbon is coloured in green, nitrogen in blue, oxygen in red, sulphur in yellow and phosphorus in orange.

Myristoyl-CoA binds to NMT with high affinity²⁹ and in a “Question Mark” conformation (Figure 6). The fatty acyl chain is binding in a narrow hydrophobic pocket in the enzyme. The size of the pocket matches the length of the myristoyl chain, providing a way for the enzyme to recognize and bind to myristoyl-CoA at high specificity, and position the thioester at the active site for reaction²⁴. The thioester carbonyl interacts with the enzyme via two main chain amides of a leucine and a phenylalanine, creating an “oxyanion hole”³⁰. This hole polarises the myristate carbonyl and activates it for nucleophilic attack from the *N*-terminus of the substrate, as well as stabilising the tetrahedral intermediate during the catalysed reaction³⁰. The thioester sulphur also forms an intramolecular hydrogen bond to the 6-*N* adenine amine of CoA,

stabilising the thiolate leaving group as well as facilitating the releasing of the CoA-SH by-product from the enzyme as a compact and globular structure³¹.

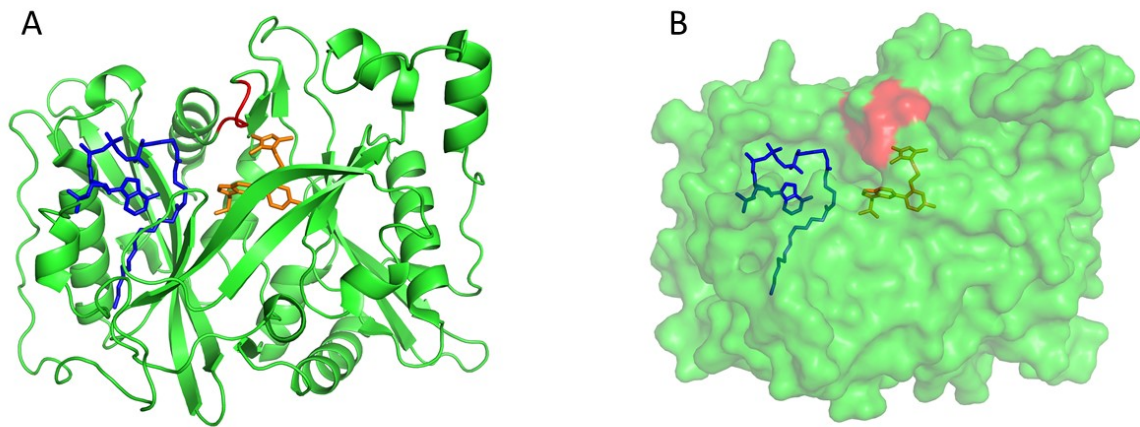


Figure 7 The Ab loop of Hs NMT1. Image generated using PyMOL²⁸ from the ternary structure of Hs NMT1 in complex with myristoyl-CoA and an inhibitor (PDB: 5O6H). Myr-CoA is shown as stick in blue, the inhibitor is shown as stick in orange and the Ab loop is shown in red. (A) Ribbon representation; (B) Surface representation.

Myristoyl-CoA forms part of the peptide binding pocket³⁰⁻³¹. The binding of myristoyl-CoA induces the ordering of the Ab loop (Figure 7)—a flexible, acidic loop of NMT (Glu-Asp-Asp-Asp) that forms part of the peptide binding pocket and contributes to peptide recognition^{30, 32}. Therefore, it is only after the binding of myristoyl-CoA, that the peptide or protein substrate can bind to the peptide binding pocket²⁷.

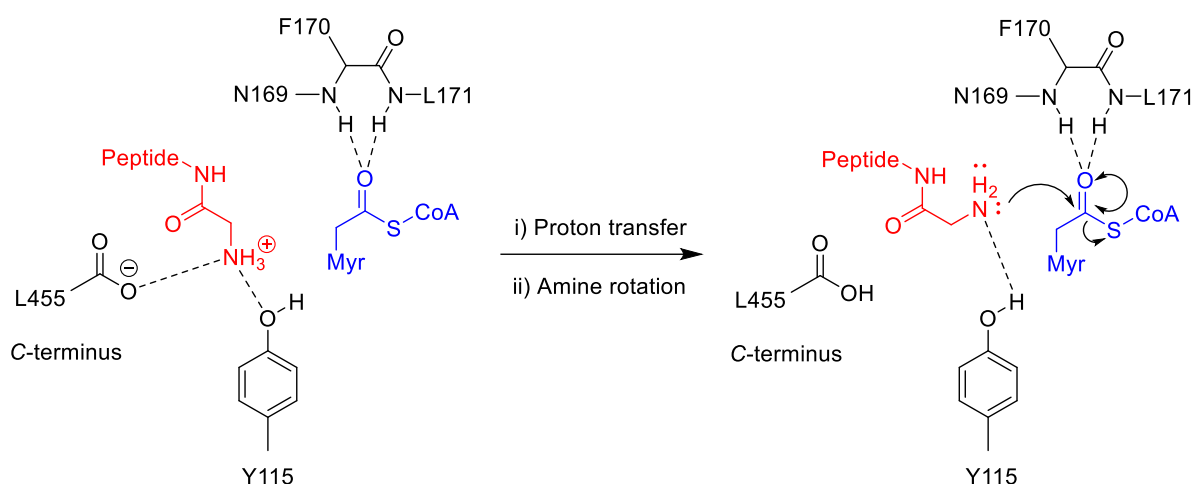


Figure 8 The first mechanism hypothesis of Sc NMT catalysed *N*-myristoylation⁸. NMT is shown in black, myristoyl CoA in blue and peptide in red.

Based on a crystal structure of *Saccharomyces cerevisiae* (Sc) NMT bound with a non-hydrolyzable Myr-CoA, S-(2-oxo) pentadecyl-CoA and a substrate peptide, GLYASKLA, the first reaction mechanism was hypothesised by Gordon *et al*^{24, 31}. This hypothesis suggests that the *N*-terminal glycine ammonium of the protein or peptide substrate interacts electrostatically with the carboxylate of the *C*-terminus of NMT (Figure 8)^{24, 31}. This carboxylate can deprotonate the substrate ammonium and generate a nucleophilic amine to attack the myristoyl-CoA thioester³¹. The amine then rotates around the backbone ψ angle after deprotonation and approaches the polarised myristoyl-CoA carbonyl at the correct geometry for the following nucleophilic attack³¹.

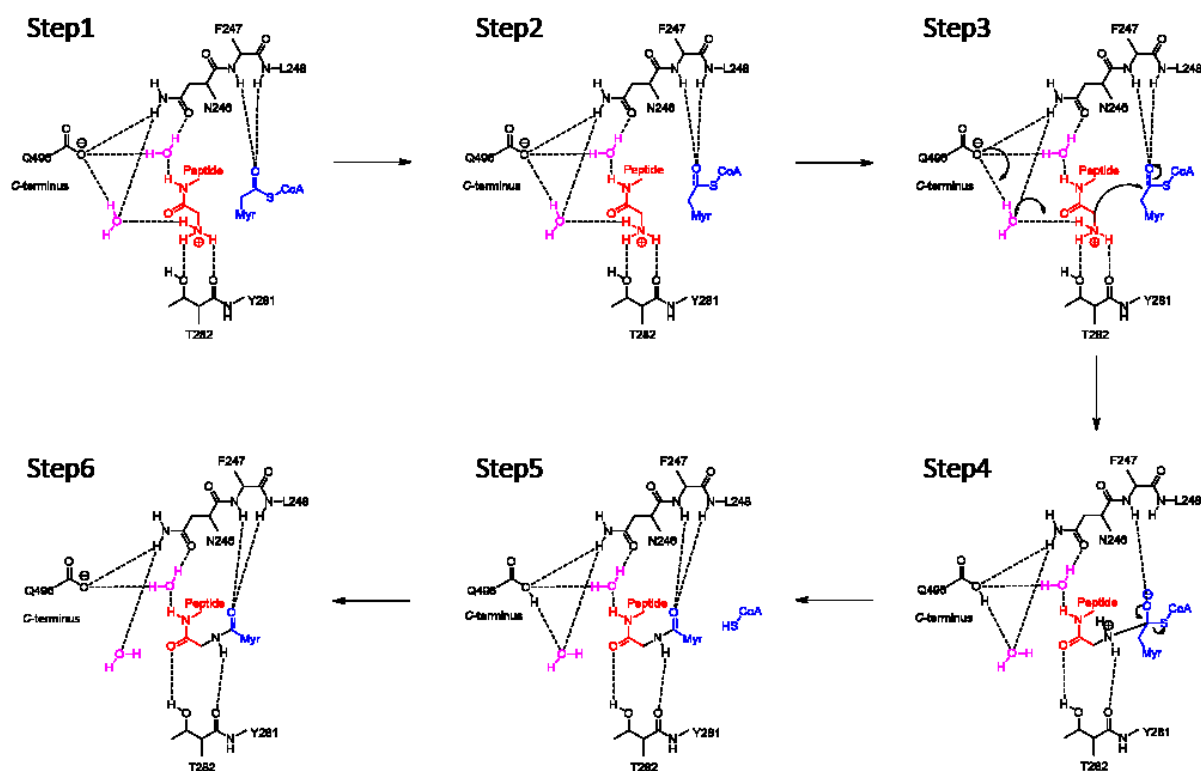


Figure 9 New *N*-myristoylation mechanism based on Hs NMT1 crystal structures (modified from Dian *et al*³³)

Recently, a new *N*-myristoylation mechanism (Figure 9) was suggested by Dian *et al* based on crystal structures of Hs NMT1 bound with Myr-CoA and substrate peptides at different reaction states³³.

The mechanism also agrees with the Bi-Bi mechanism that a peptide can only bind to NMT after the binding of Myr-CoA to the apo-NMT and the induction of the Ab loop conformation change from disordered to ordered-open.

However, in contrast to the previous mechanism, it suggests that the peptide is stabilized by water-mediated hydrogen bonding interactions with the C-terminus and N246 as well as two direct hydrogen bonds with T282 (Figure 9, Step 1). The binding of the peptide further induced the Ab loop to change its conformation from ordered-open to ordered-closed, which then distorts the thioester plane of the Myr-CoA. This tensed conformation of Myr-CoA facilitates the depolarization of the thioester by the “oxyanion hole” created by F247 and L248 and pre-

sets the acyl group up in an sp³ hybridization status ready for the nucleophile to attack (Figure 9 Step 2, a reaction status captured by crystallography). The ammonium *N*-terminus of the peptide is deprotonated by the carboxy group of Q496 through a water-mediated attack. In the meantime, the ammonium *N*-terminus of the peptide attacks the thioester carbonyl of Myr-CoA guided by the T282 hydrogen bonds (Figure 9 Step 3), forming a tetrahedral intermediate shown in Figure 9 Step 4 (a reaction status captured by crystallography). The following steps are the same in the two mechanisms. The breakdown of the tetrahedral intermediate is the rate-limiting step of the *N*-myristoylation reaction³⁴. As L248 forms a hydrogen bond with the thioester carbonyl again, the myristoylated peptide is synthesised as well as the by-product CoA in a tensed conformation (Figure 9 Step 5). Relaxation of the tensed CoA favours the release of the CoA moiety from the active site. Spontaneously, the carboxyl group of the *C*-terminus Q496 is deprotonated to get ready for a new reaction. Finally, the Ab-loop opens, the *N*-terminus of NMT unfolds and the myristoylated peptide is released (Figure 9 Step 6).

The difference between the two mechanisms proposed is likely a result of the use of non-hydrolysable myristoyl-CoA (NHM), S-(2-oxo) pentadecyl-CoA in the former study and Myr-CoA in the later study³³. The NHM analogue has a methylene group interposed between the reactive thioester carbonyl and sulphur to prevent *N*-myristoylation reaction during crystallization. However, this extra methylene group also causes steric hindrance to the *N*-terminus of the substrate peptide and prevents it from binding into its recognition pocket. This change causes a series of binding mode changes of the peptide substrate in the former study and leads to a different mechanism hypothesis.

1.2.2 Role of myristoylation

Myristoylation usually occurs co-translationally, following the removal of the starting methionine by methionine aminopeptidase to expose the *N*-terminal glycine³⁵. More rarely,

post-translational myristoylation can occur during apoptosis, where cleavage of pro-apoptotic proteins by caspases reveals a new *N*-terminal³⁶⁻³⁸.

Myristoylation plays crucial roles in the regulation of membrane targeting, trafficking and signalling, modulating the specificity and efficiency of signal transduction and influences the movement of a signalling protein within the cell and its final destination^{25, 39}. When happen post-translationally, it takes part in apoptotic signalling and induce programmed cell death^{2, 37}.

1.3 Human NMTs

1.3.1 Isoforms

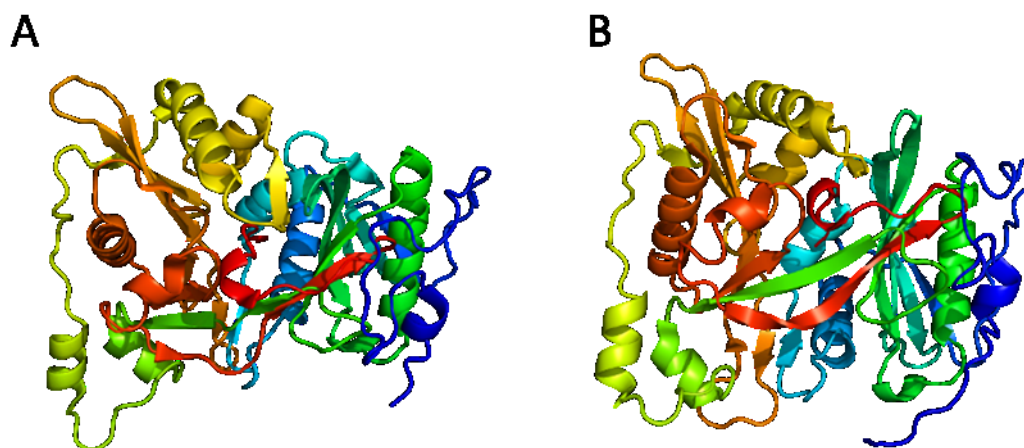


Figure 10 Human NMT isoforms. (A) Ribbon representation of Hs NMT1 (PDB code: 4C2Y); (B) Ribbon representation of Hs NMT2 (PDB code: 4C2X).

In higher organisms, such as humans, two NMT proteins are expressed in most tissues. The two distinct NMT genes, *Homo sapiens* (Hs) NMT1 and NMT2 share 91% sequence similarity and 83% sequence identity⁴⁰. The structures of both Hs NMT1 (PDB code: 4C2Y) and Hs NMT2 (PDB code: 4C2X) are solved by crystallography³⁶. Their protein substrate-binding pocket are highly conserved. The major differences are in their flexible *N*-terminal domains, which are not involved in substrate binding⁴¹⁻⁴². A screen on 2066 compounds in ten structural series by Pfizer found no significant differences between the IC₅₀s against Hs NMT1 and Hs NMT2 among all tested series, indicating that it is hardly possible to selectively inhibit just one isoform at the enzyme level (Figure 11)⁴⁰.

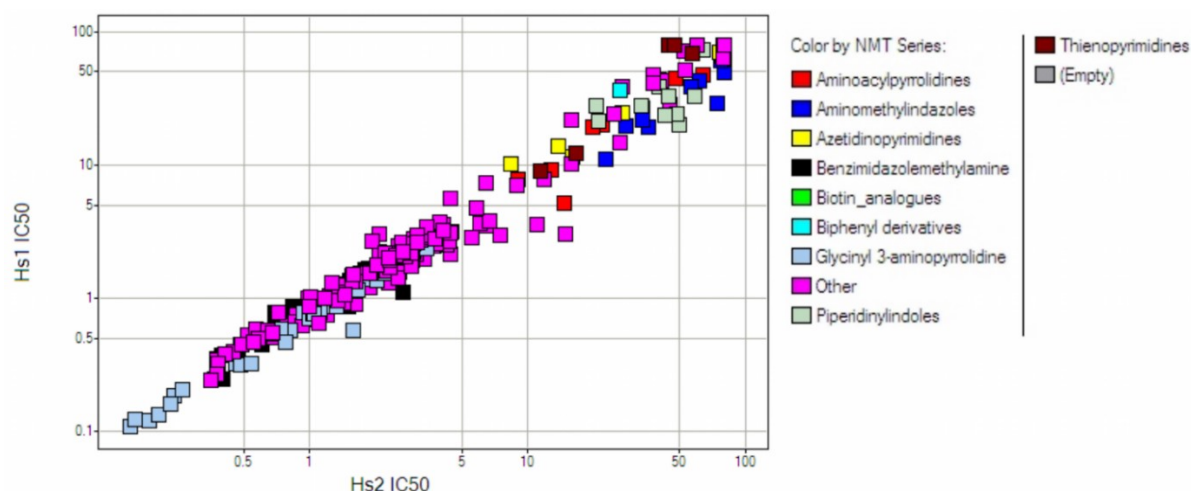


Figure 11 Plot of dose-response activity against Hs NMT1 vs. Hs NMT2 for the primary screening hits and their analogues (taken from Bell *et al*⁴⁰ under the Creative Commons Attribution License)

1.3.2 Human NMTs as potential drug targets

Studies have shown the involvement of human NMTs in the development and progression of diseases including epilepsy⁴³, cancer⁴⁴, Noonan-like syndrome⁴⁵, Alzheimer's disease^{43, 46}, viral⁴⁷⁻⁴⁸ and bacterial infections⁴⁸. Among them, cancer and common cold virus are the most investigated. Small molecule NMT inhibitors were shown to be cytotoxic to cancer cells⁴⁴ and inhibitory to the reproduction of common cold virus⁴⁷.

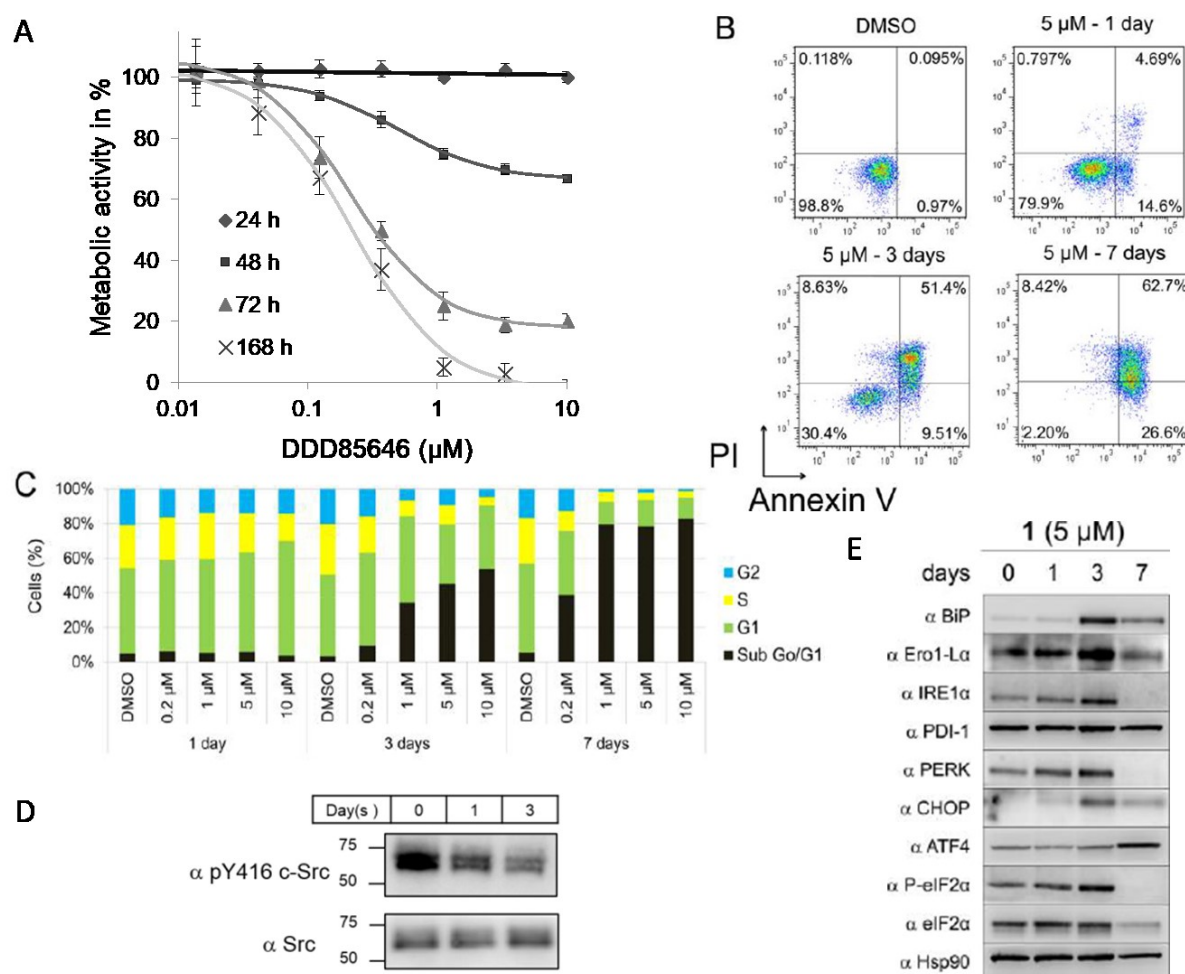


Figure 12 Evidence of being a potential cancer target (reproduced from Thinon *et al*^{36, 49} under a Creative Commons Attribution License). (A) Activity curves from cell cytotoxicity assay³⁶; (B) Graphs from Flow cytometry analysis⁴⁹; (C) Graph of cell cycle analysis⁴⁹; (D) Western blot profiling of overall c-Src expression and Y416 phosphorylation of c-Src after NMT inhibitor treatment⁴⁹; (E) Western blot analyses of markers of ER stress and unfolded protein response⁴⁹.

In 1995, Felsted and co-workers first proposed to consider myristoylation as a chemotherapeutic target for cancer⁵⁰. Since then, evidence has been found to suggest NMT's involvement in cancer. Overexpression of NMT is observed in colonic adenocarcinomas⁵¹⁻⁵², gallbladder carcinomas⁵³, oral squamous cell carcinomas⁵⁴ and brain tumours⁵⁵. The proliferative capacity of mammary epithelial cells has been found to correlate with NMT activity⁵⁶.

NMT inhibition by a small molecule inhibitor, **DDD85646 (C14)**, its drug discovery process can be found in 1.4.2), has shown dose-dependent cytotoxicity on HeLa cells in a time-dependent manner, which kills the cells in seven days at 5 μ M (Figure 12A), likely through apoptosis (Figure 12B) following or concurrent with G1 arrest (Figure 12C) ^{36, 49}.

NMT inhibition was shown to prevent the activation of oncogene Src⁴⁹. More than hundreds of NMT substrates were found in human cells, including the well-known oncogene Src³⁶. Tyrosine-protein kinase c-Src is a well-validated target for cancer therapy⁵⁷. c-Src promotes cell survival, proliferation and metastasis, and its overexpression is related to cancer progression⁵⁷. A Src inhibitor, Dasatinib, has been used in the clinic for cancer treatment, which also induces apoptosis and G1 arrest of cancer cells⁵⁸. *N*-myristoylation is required for the activation of c-Src. A study by Thinon *et al* has shown that **DDD85646** induced NMT inhibition decreases the level of phosphorylated c-Src (the activated version) after 1 and 3 days⁴⁹ (Figure 12D), suggesting that c-Src inhibition can be achieved alternatively through NMT inhibition. Therefore, NMT may be a promising target for cancer.

Thinon *et al* also found that NMT inhibition can induce apoptosis of cancer cells through ER-stress⁴⁹. The endoplasmic reticulum (ER) stress response is an adaptive cellular response to detrimental conditions. Cells in stress conditions switch on this mechanism to restore normal functions of the ER following an accumulation of misfolded proteins, thus supporting cell survival. When conditions become too severe and excessive, the ER stress pathway becomes over-burdened and cells turn on the pro-apoptotic module which triggers cell death⁵⁹. ER stress is chronically activated in many cancer cells, supporting continuous proliferation and survival of tumours in adverse microenvironments, such as hypoxia and chemotherapy⁶⁰. Drug-induced ER stress can add an additional burden on this pathway and force cancer cells into apoptosis. ER-stress inducing inhibitors of other protein targets, such as Bortezomib A or Brefeldin A⁶¹,

have already entered the clinic or clinic trial for cancer therapy. A quantitative proteomics study on four cancer cell lines has shown ER stress and unfolded protein response of these cells upon selective NMT inhibition by **DDD85646** (Figure 12E)⁴⁹. This finding suggests that NMT inhibition can promote the ER-stress of cancer cells and induce their apoptosis, therefore, can be an alternative way of treating cancer.

Apart from cancer therapy, pharmacological inhibition of human *N*-myristoylation was also shown to block viral replication.

Rhinovirus (RV) is the most frequent cause of the common cold. Its infection can also lead to exacerbation of other respiratory diseases, such as asthma, chronic obstructive pulmonary disease (COPD)⁶² and cystic fibrosis⁶³⁻⁶⁴. There is no approved antiviral therapy of RV due to its serotypic diversity, rapid replication and high mutation rate⁶⁵. The capsid assembly and infectivity of RV rely on the host cell NMTs for the *N*-myristoylation of a key protein VP0⁶⁶⁻⁶⁸. The host cell NMTs can be an attractive antiviral drug target minimally susceptible to serotypic variation and virus mutation. In a recent publication by Mousnier *et al*, the VP0 myristoylation is shown to be dose-dependently inhibited by a picomolar dual Hs NMT1/2 inhibitor (**C32**, its drug discovery process is described in 1.4.3) in chemical-tagging assays⁴⁷. In a two-day multicycle RV replication assay, this compound prevented virus-induced cytopathic effect (CPE) in a dose-dependent manner with an IC₅₀ value of 17 nM without affecting the cell viability of uninfected host cells at the meantime⁴⁷. In single-cycle RV infection assays, complete suppression of new infectious virus was observed at 125 nM, the antiviral potencies were demonstrated in five RV serotypes, and when added up to 3 hours post-infection⁴⁷. These data demonstrate that pharmacological inhibition of human *N*-myristoylation rapidly and completely prevents RV replication without inducing cytotoxicity, highlighting the potential of human myristoylation as a drug target in rhinoviral infections.

1.4 NMT inhibitors

Human NMT specific studies of small-molecule inhibitor development are limited^{47, 69}. However, several series of inhibitors have been discovered against NMTs of *Candida albicans* (Ca)⁷⁰⁻⁷², *Trypanosoma brucei*^{42, 73} (Tb), *Plasmodium falciparum*^{40, 74-76} (Pf), *Plasmodium vivax*^{75, 77} (Pv) and *Leishmania donovani*^{40, 78} (Ld), promising to tackle diseases, such as fungal infection, sleeping sickness, malaria and leishmaniasis. Human NMT activities were usually tested in these studies and sometimes inevitable due to high similarities of their binding pockets (analysed in 2.3.2). Therefore, their structure-activity relationships (SARs) and binding mode insights can be valuable knowledge to facilitate inhibitor development of human NMTs.

1.4.1 Fungal NMT inhibitors

In 1988, Gordon *et al* reported a group of peptide substrates of yeast NMT with kinetic measures to illustrate substrate specificity of NMT⁷⁹. Although later on this was considered the first publication of NMT inhibitors, the paper was largely overlooked by the medicinal chemistry community at the time, due to its heavy enzymology focus⁸⁰.

A limited number of myristic acid and myristic CoA analogues were published as NMT inhibitors between 1989 and 1995⁸¹⁻⁸⁵. However, due to concerns of their selectivity over Hs NMT1, these inhibitors were not further developed as drug candidates, though some have been used as tools for biological study⁸².

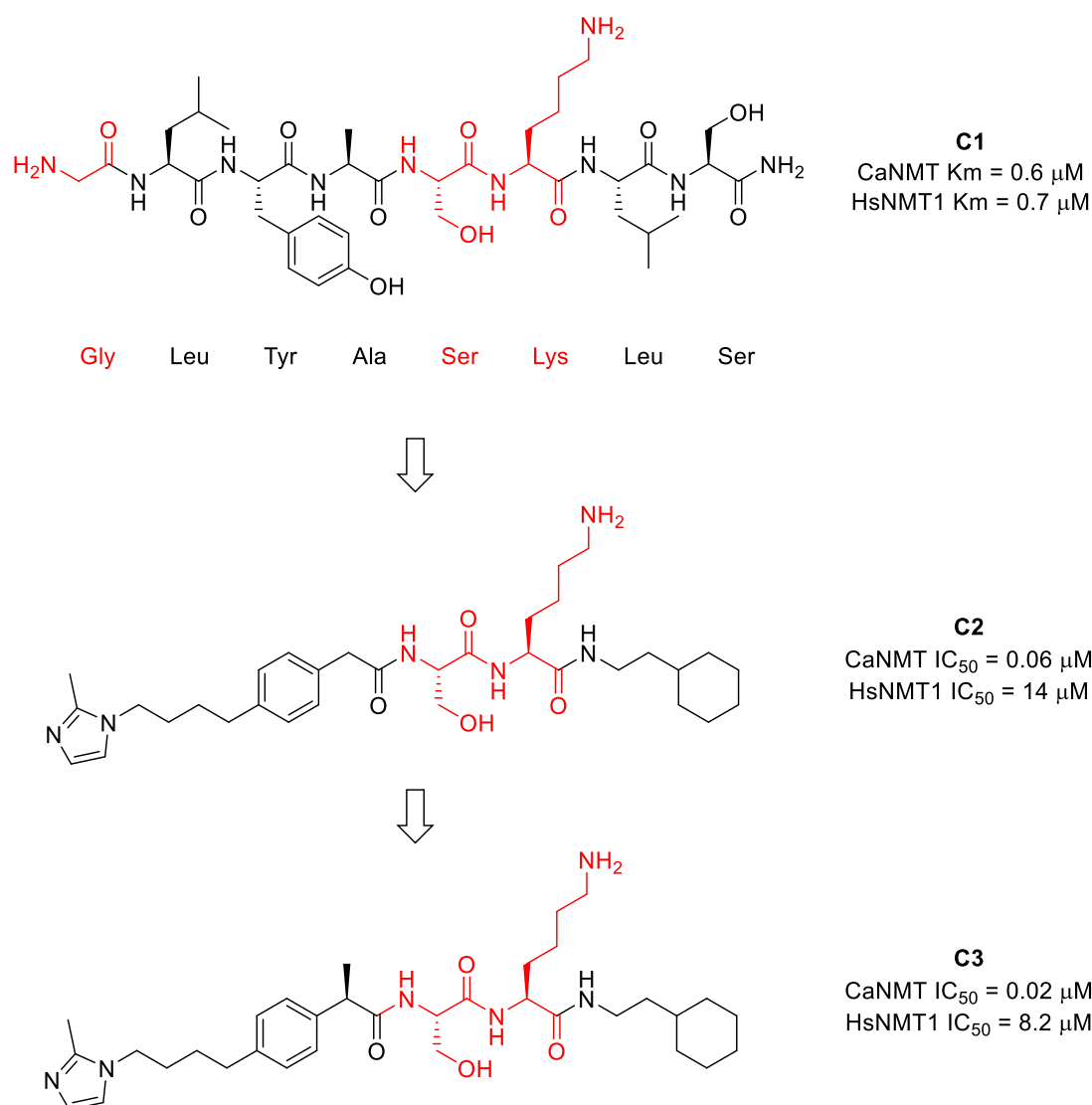


Figure 13 Development of peptide mimic inhibitor of Ca NMT⁸⁶. High-affinity amino acids and key interacting residues are coloured in red, and low-affinity amino acids and other parts of inhibitors in black.

In 1995, Gordon *et al.* reported a series of peptide mimic inhibitors of human pathogenic fungi, *C. albicans* (Ca) NMT⁷⁰, which was first expressed in yeast in 1992⁸⁷ and later proved to be essential for the vegetative growth of *C. albicans*⁸⁸.

The inhibitors were based on an octapeptide substrate GLYASKLS-NH₂ (**C1**, Figure 13) derived from the *N*-terminal fragment of NMT substrate protein Arf2p (ADP ribosylation factor 2)⁸⁹. By replacing each amino acid residue in **C1** with an alanine, the study revealed that

the primary glycine, the serine at position 5 and the lysine at position 6 are the key residues for tight binding⁷⁰.

The initial work kept the serine and lysine but depeptidized the rest of the residues and identified a dipeptide inhibitor **C2** with an IC₅₀ of 56 nM against Ca NMT and a 250-fold selectivity over Hs NMT1. The substitution of 2-methylimidazole for the primary glycine was well tolerated, which was later attributed to its similar basicity (pKa 7.1) to glycine (pKa 8.0), resulting in a good charge-charge interaction to Ca NMT⁹⁰.

Reintroduction of the α -methyl group in the original octapeptide substrate with an (R) configuration at the benzylic position (**C3**, Figure 13) improved the selectivity and potency against Ca NMT with an IC₅₀ of 20 nM and a 410-fold selectivity over Hs NMT1. This gave the most potent inhibitor at the enzyme level in this series⁹¹.

Subsequent investigations focused on further exploration of SARs and increased the half-life *in vivo* and bioavailability^{90, 92-95}, however, did not achieve good cell activities due to the poor permeability of the large and polar peptide analogues.

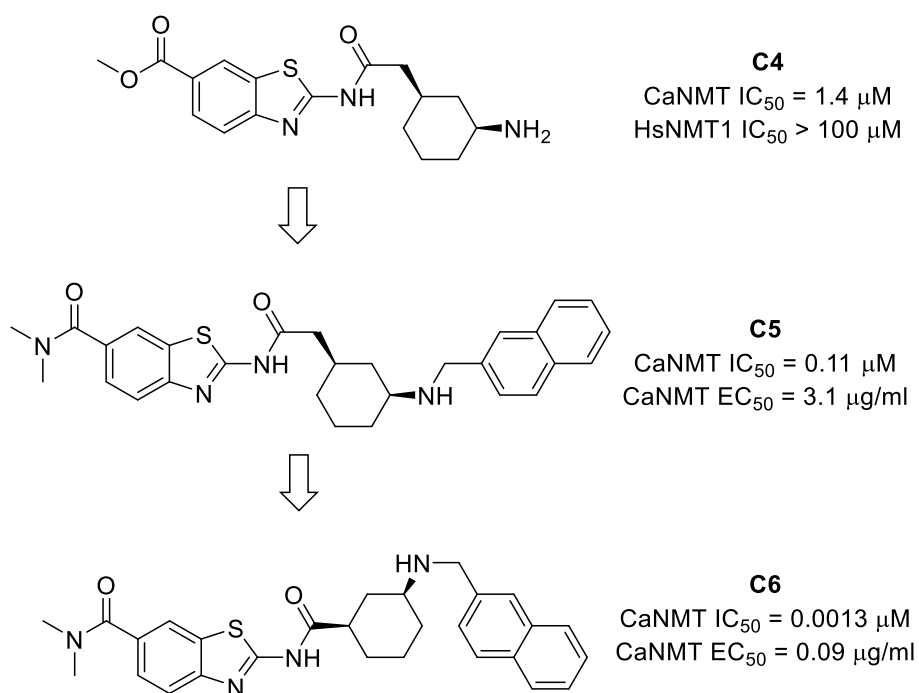


Figure 14 Development of Pfizer benzo[d]thiazole series inhibitor of Ca NMT⁸⁰.

In response to the discovery of the peptide mimic inhibitors of Ca NMT, Pfizer developed a benzo[d]thiazole series based on hits from a high throughput screen (HTS) against Ca NMT⁸⁰. The HTS identified the hit **C4** (Figure 14) with a moderate potency (1.4 μM) but a high selectivity over Hs NMT1 (>71-fold). By replacing the metabolically vulnerable ester with an amide and introduction of a bulky substituent onto the primary amine, **C5** (Figure 14) was identified as the first NMT inhibitor with potent antifungal activity (3.1 μg/ml). Shortening the linker between the cyclohexyl and amide yielded **C6** (Figure 14) with greatly improved enzyme activity (0.0013 μM) and cell activity (0.09 μg/ml). However, due to the limited potential clinical application resulted from its lack of broad-spectrum antifungal activity against other major fungal pathogens (*Aspergillus* and *Cryptococcus*), this series was halted for further development^{72, 96-97}.

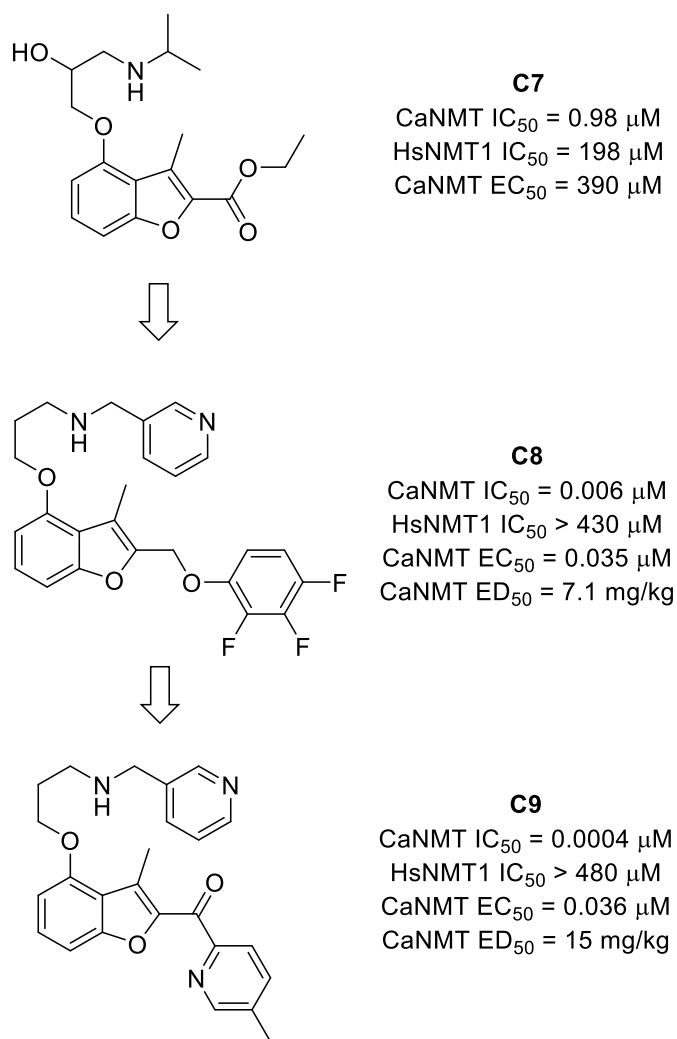


Figure 15 Development of Roche benzo[d]furan series inhibitor of Ca NMT⁷¹.

In parallel with the Pfizer study, Roche identified a benzo[d]furan series from their HTS findings⁷¹. The starting point was a screen hit **C7** (Figure 15), which is a moderately potent Ca NMT inhibitor (0.98 μM) with a good selectivity over Hs NMT1 (202-fold). However, **C7** displayed poor cellular activity (390 μM) and off-target pharmacology as a β-adrenergic receptor antagonist, therefore the development was discontinued.

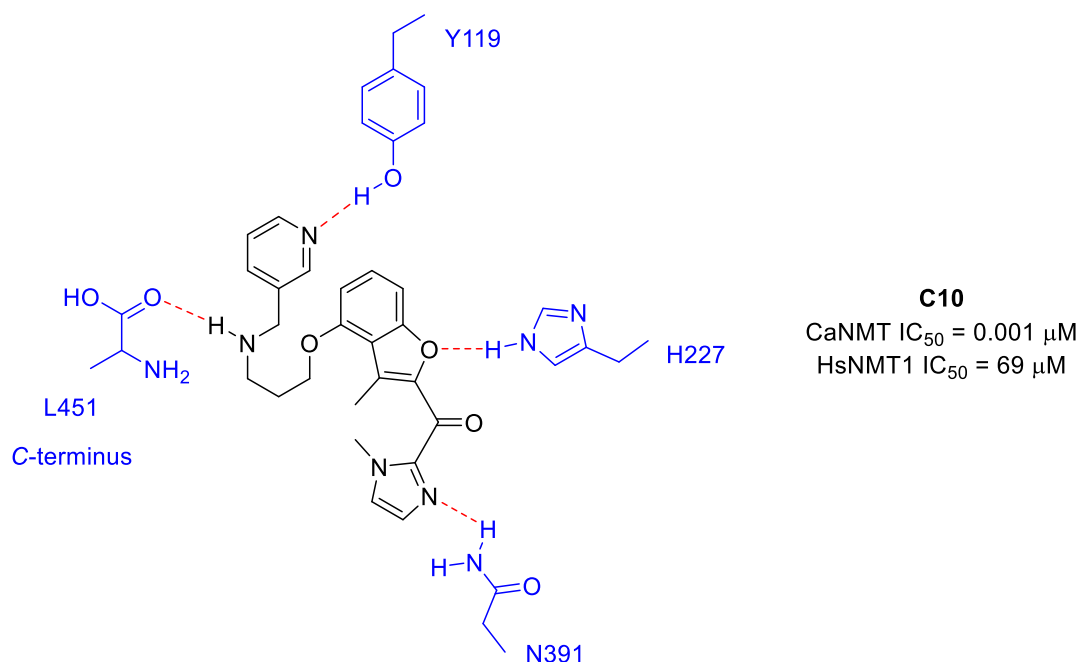


Figure 16 A schematic drawing of the binding mode of **C10** (an analogy of **C8**) in Ca NMT (modified from Sogabe *et al*⁹⁸ and based on the crystal structure of the ternary complex of Ca NMT with myristoyl-CoA and **C10**, PDB: 1IYL)

Around the same time, both the apo-Ca NMT and Ca NMT with a bound inhibitor were successfully crystallised⁹⁸⁻⁹⁹. With the aid of the binding modes revealed by crystallography (Figure 16), Roche explored the SARs of Ca NMT inhibitors, addressed the off-target problem by removing the hydroxyl group⁷¹, and identified the first NMT inhibitor (**C8**, Figure 15) that demonstrated efficacy in an animal infection model (ED₅₀ 7.1 mg/kg). To solve the instability problem of the ether in acidic conditions, the ether linker was replaced by a ketone to give **C9** (Figure 15) with subsequent optimization. Despite this series' potent activities against *Candida spp.* NMTs, its further development was halted due to lack of activity against other important systemic fungal pathogens¹⁰⁰.

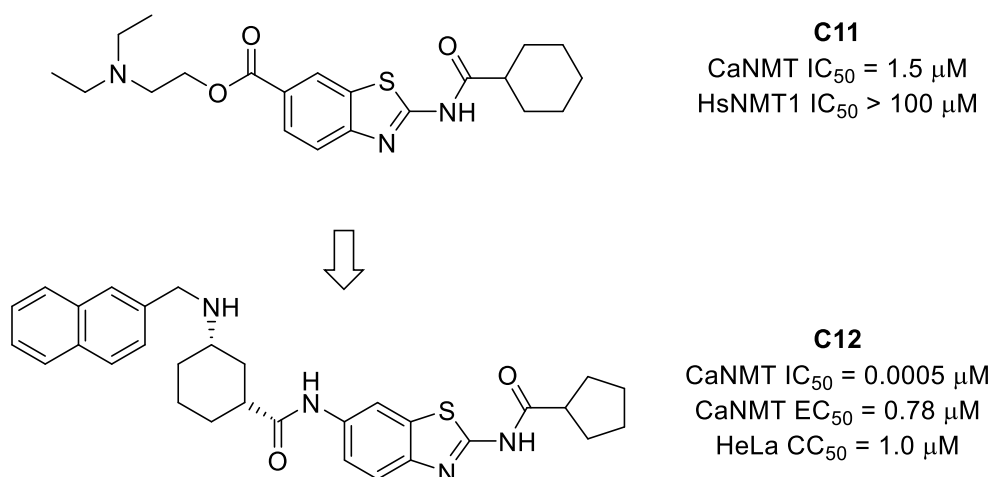


Figure 17 Development of SSP benzo[d]thiazole series inhibitor of Ca NMT⁷².

Following Pfizer and Roche's research, SSP developed another benzo[d]thiazole series of Ca NMT inhibitors⁷². **C11** (Figure 17) was identified as a hit from virtual screening as a moderately potent Ca NMT inhibitor (1.5 μM). Through extensive parallel synthesis, **C12** (Figure 17) was discovered with potent enzyme (0.5 nM) and cell activity (0.78 μM). However, this series has a narrow window between the 50% effective concentration (EC₅₀) and the 50% cytotoxic concentration (CC₅₀), indicating an off-target mechanism. The series also has a narrow antifungal spectrum, therefore was not developed further.

1.4.2 Parasitic NMT inhibitors

After the characterization of NMTs from *Plasmodium falciparum*¹⁰¹, *Leishmania*¹⁰² and *Trypanosoma brucei*¹⁰², NMT was validated by Price *et al* to be essential for the viability in *Leishmania major* (Lm) and *Trypanosoma brucei* (Tb) using gene targeting and RNA interference (RNAi)¹⁰². A panel of compounds originally developed as fungal NMT inhibitors were tested against Lm NMT and Tb NMT and showed *in vitro* activities in both NMTs and *in vivo* activities in Tb NMT¹⁰³.

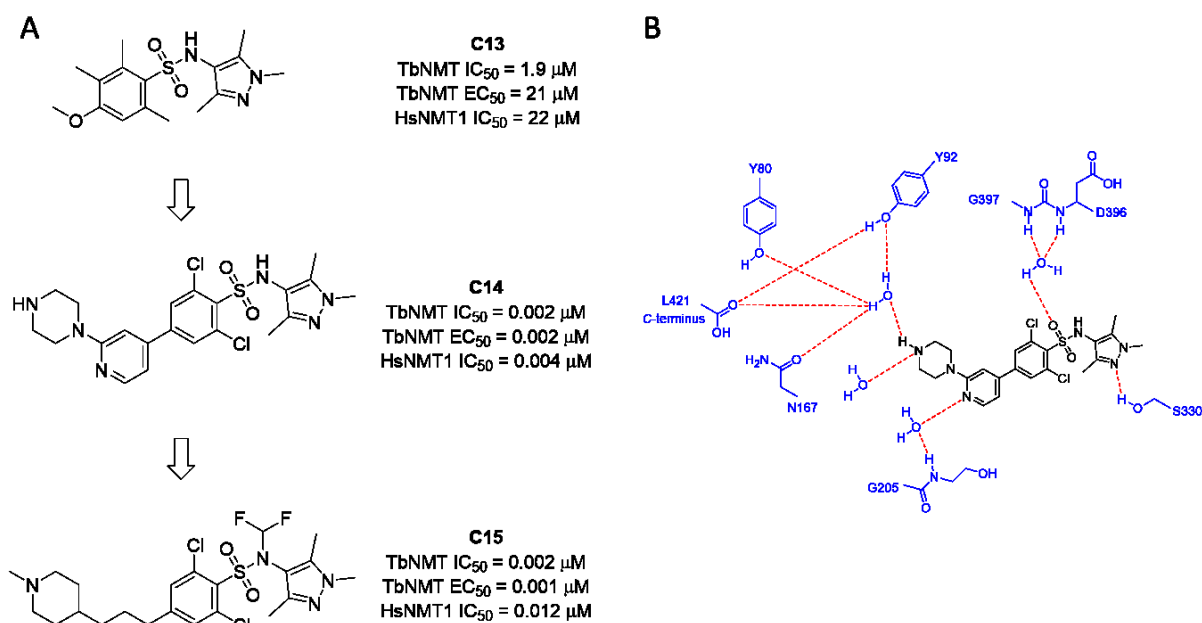


Figure 18 Development of Dundee pyrazole sulphonamide series inhibitor of Tb NMT. (A) Inhibitor structures and activities; (B) A schematic drawing of the binding mode of **C14** (**DDD85646**) in Lm NMT (modified from Brand *et al*⁷³ and based on the crystal structure of the ternary complex of Lm NMT with myristoyl-CoA and **C14**, PDB: 2WSA).

The successful “piggy-backing” (screen existing inhibitors originally designed for other targets to obtain inhibitors of new targets) attempt encouraged the HTS of the Dundee Drug Discovery compound collection against Tb NMT, which identified a moderate Tb NMT inhibitor, **DDD64558** (**C13**, Figure 18A)⁷³.

With the aid of crystallography using the highly homologous Lm NMT (72% similarity and 54% identity⁴⁰) as a Tb NMT substitute (Figure 18B), a potent inhibitor **DDD85646** (**C14**, Figure 18A) was discovered through extensive SAR studies⁷³. Crystallography revealed a water-mediated interaction between the *NH* of the piperazine heterocycle and the enzyme C-terminus, and the hydrogen bond between the pyrazole and the enzyme Ser330, which prove to be the two key interactions for tight binding (Figure 18B)⁷³. **DDD85646** was then tested and confirmed by Frearson *et al.* to be a trypanocidal Tb NMT inhibitor both *in vitro* and *in vivo*, and used as a tool to validate Tb NMT as an antiparasitic drug target⁴². Due to the high potency

of **DDD85646**, it was also used to validate other NMTs as potential drug targets, such as *Trypanosoma cruzi*¹⁰⁴, Filarial nematode¹⁰⁵, *Aspergillus fumigatus*¹⁰⁶ and cancer⁴⁴ (see 1.3.2).

However, **DDD85646** shows poor central nervous system (CNS) exposure, which restricts its clinical use. By reducing polar surface area and capping the sulfonamide to improve its blood-brain barrier permeability, and replacing the core aromatic with a flexible linker to improve selectivity over Hs NMT1, **DDD100097** (**C15**, Figure 18A) was developed, demonstrating partial efficacy in a central nervous system (CNS) mouse model¹⁰⁷.

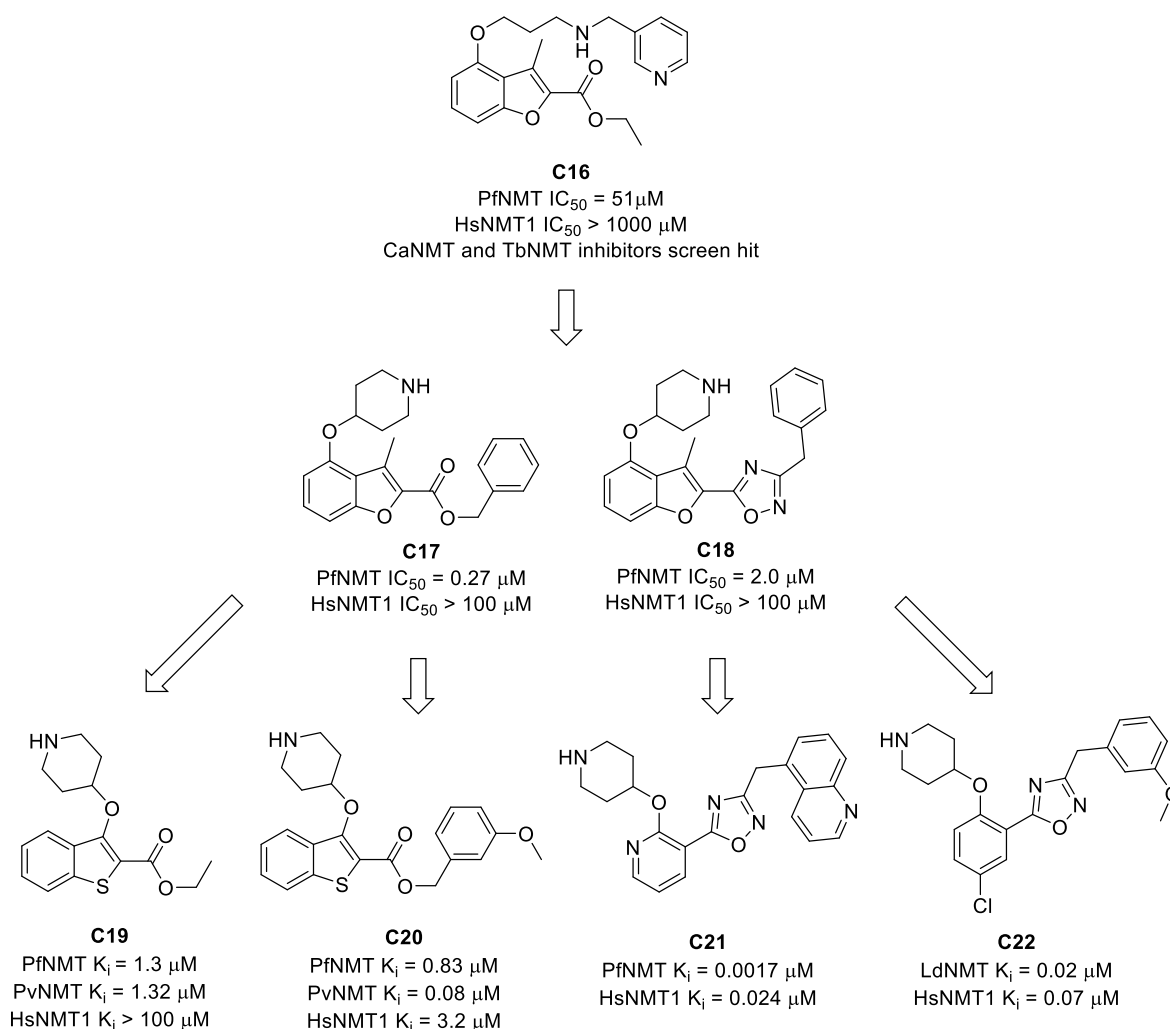


Figure 19 Development of Imperial College benzothiazole series

A “piggy-back” screening of a 25-inhibitor library was conducted on previous reported Ca NMT and Tb NMT series against Pf NMT and Hs NMT1^{76, 86, 108}. RO-09-4609 (**C16**, Figure

Moreover, to simplify the structure of **C18**, the Imperial College group “scaffold hopped” from benzo[*b*]furan to phenyl-based monocycle, and discovered **C21** with single-digit nM enzyme affinity¹⁰⁹ as well as a series of Ld NMT inhibitors represented by **C22** (Figure 19)⁷⁸.

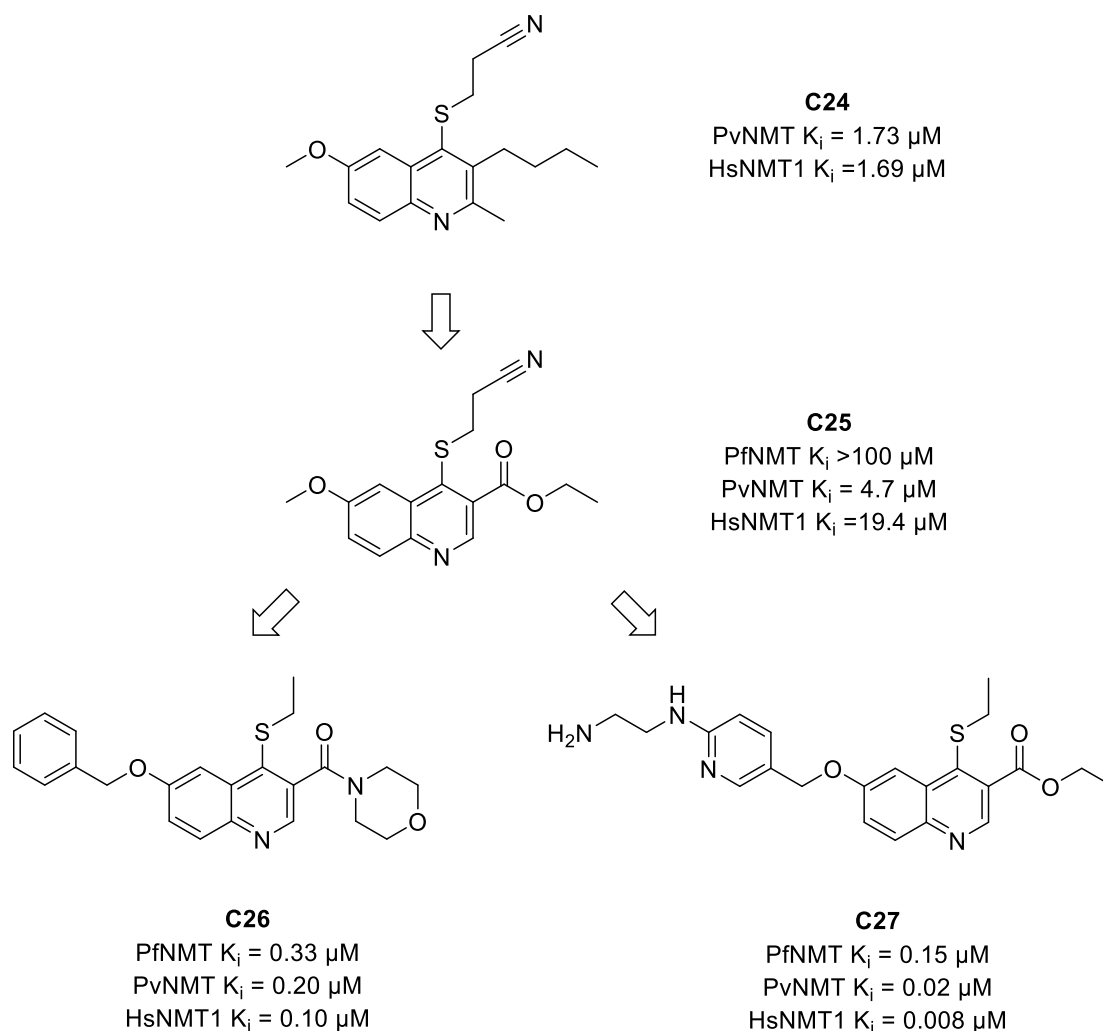


Figure 21 Development of Imperial College quinoline series

Goncalves *et al.* from the same Imperial College group conducted a high-throughput screen of 59.5k compounds from the Medical Research Council Technology (MRCT) collection and identified 107 inhibitors of Pv NMT⁷⁷. Among them, **C24** (Figure 21) was selected as a lead and the replacement of its aliphatic chain by an ethyl ester gave **C25** with better physicochemical properties⁷⁷. **C25** was further optimized into **C26** (Pf K_i = 0.33 μ M, Pv K_i =

0.20 μM) and **C27** (Pf K_i = 0.15 μM , Pv K_i = 0.02 μM) with improved potency against both Pv and Pf NMTs (Figure 21)⁷⁷.

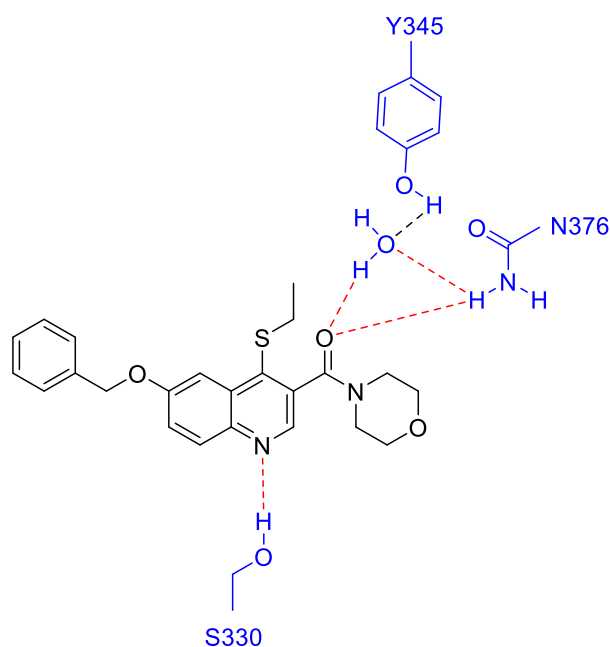


Figure 22 A schematic drawing of the binding mode of **C26** based on the crystal structure of the ternary complex of Pv NMT, Myr-CoA and **C26** (PDB: 5G20, modified from Goncalves *et al.*¹¹⁰).

Compound	Pf IC ₅₀	Ld IC ₅₀	Tb IC ₅₀	Hs1 IC ₅₀	Hs2 IC ₅₀
CP-854190	1.72	38.7	>80	0.57	1.59
PF-03665853	0.91	>80	35.4	0.51	0.56
PF-02378143	2.54	65.4	>80	38.9	56.3
PF-03531549	1.7	1.2	52.7	24.8	27.1
PF-03531814	0.48	72	>80	70.0	76.0
PF-03402619	>80	0.90	9.9	47.2	63.6
PF-03402623	71.9	0.0093	1.5	5.2	14.7
PF-03393842	9.0	0.10	35.9	73.2	64.1
PF-00349412	39.6	0.48	56.2	9.0	11.5
PF-00075634	>80	0.16	>80	36.1	26.7

Figure 23 The structure and IC₅₀s of the Pfizer screen hits (Modified from Bell *et al.*⁴⁰ under the Creative Commons Attribution License)

In 2012, Pfizer published a selection of ten screen hits (Figure 23) against Pf NMT, Ld NMT, Tb NMT and both Hs NMT1/2 in collaborations with universities and government to help

stimulate research into parasitic NMT inhibitors⁴⁰. Based on these hit compounds, several series of NMT inhibitors were published by Imperial College.

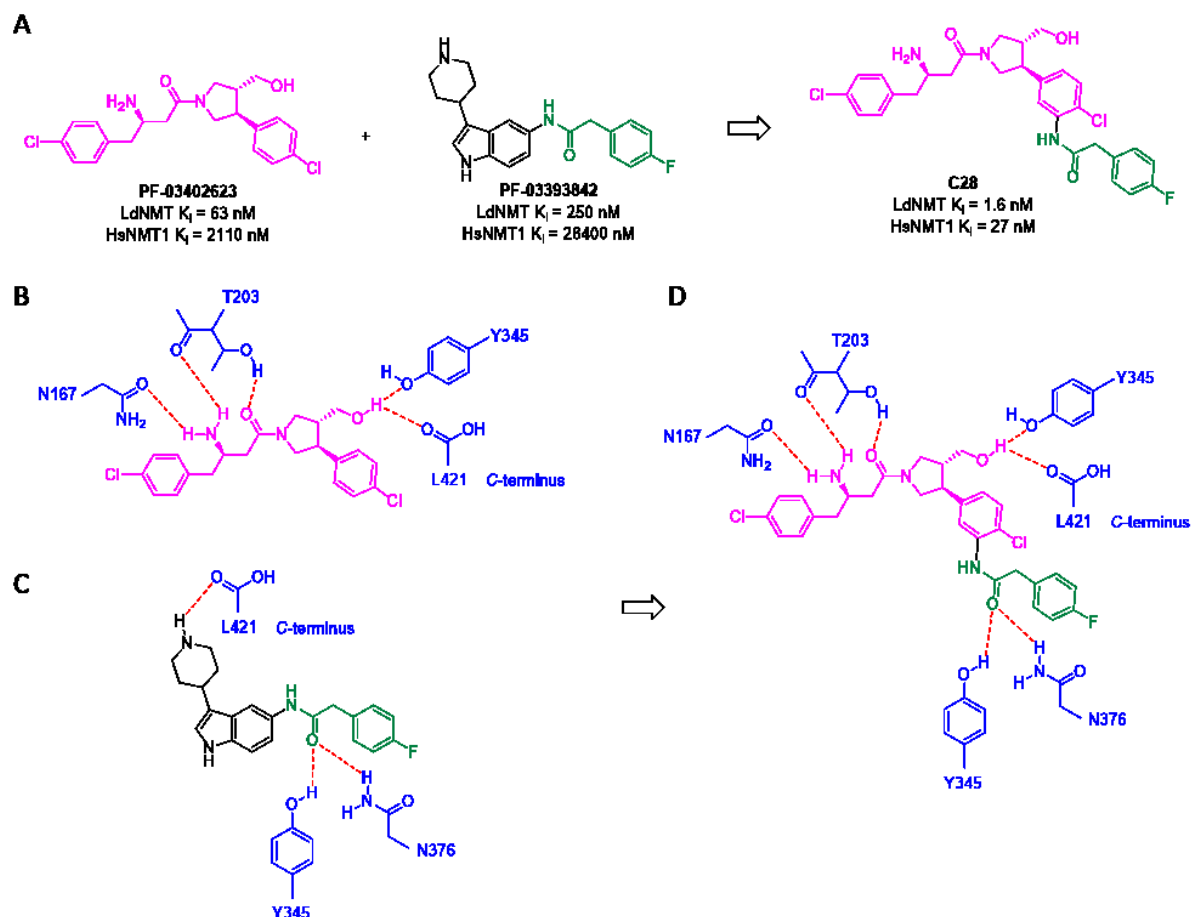


Figure 24 Development of Imperial College hybrid series (modified from Hutton *et al.*¹¹¹). (A) Hybridization of two complementary structures and the activity improvement; (B) A schematic drawing of the binding mode of **PF-03402623** in Lm NMT (PDB: 4CGL); (C) A schematic drawing of the binding mode of **PF-03402623** in Lm NMT (PDB: 4CGN); (D) A schematic drawing of the binding mode of **C28** in Lm NMT (PDB: 4CYO).

According to the overlapping and complementary binding modes of **PF-03402623** and **PF-03393842** revealed by crystallography (Figure 24), Hutton *et al.* hybridized these two inhibitors and discovered a novel Ld NMT inhibitor **C28** with good enzyme activity (K_i = 1.6 nM, Figure 24)¹¹¹. All interactions of the two inhibitors with the enzyme were conserved in the hybrid inhibitor **C28**, achieving a 39-fold and 156-fold activity improvement in Ld NMT inhibition, compared to the two original inhibitors respectively.

1.4.3 Human NMT inhibitors

In contrast to the discovery projects of fungal- or parasite-NMT-targeted inhibitors, there are very few studies searching for human NMT inhibitors.

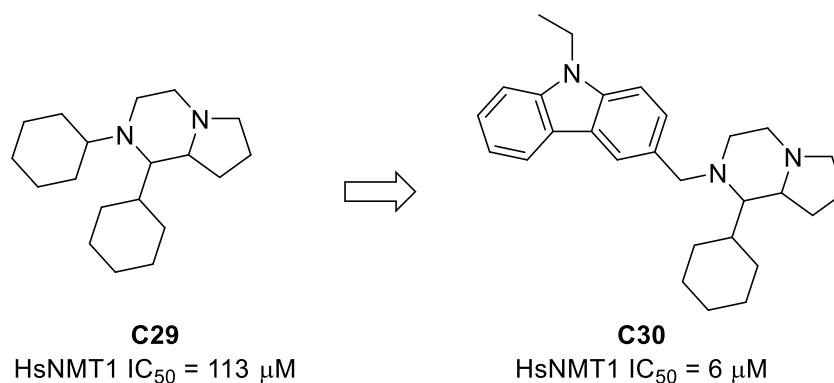


Figure 25 Development of the pyrrolo[1,2-a]pyrazine series

French *et al.* conducted an HTS of a library with approximately 16,000 diverse compounds against Hs NMT1 and identified the cyclohexyl-octahydropyrrolo[1,2-a]pyrazine (**C29**, Figure 25) with an IC₅₀ of 113 μM as one of its screen hits. Based on **C29**, French *et al.* modified it into 31 analogues and identified **C30** (Figure 25) with an IC₅₀ of 6 μM⁶⁹.

Interestingly, the Sharma group identified two protein Hs NMT1 inhibitors that inhibit human NMT1 in a concentration-dependent manner with 50% inhibition at 640 nM and 4.5 nM¹¹²⁻¹¹³. Both inhibitor proteins are susceptible to trypsin digestion and possess no protease activity. In addition, both inhibitor proteins are neither myristoyl-CoA binding proteins nor do they interfere with the binding of myristoyl-CoA and human NMT1. Furthermore, both inhibitor proteins do not have demyristoylase activities. Therefore, the Sharma group concluded that the inhibitor proteins may act directly on NMT as NMT inhibitors. The exact mechanism of action remains unclear, though the presence of an *N*-terminal alanine residue on the inhibitor proteins could make the proteins competitive to the peptide substrates of NMT.

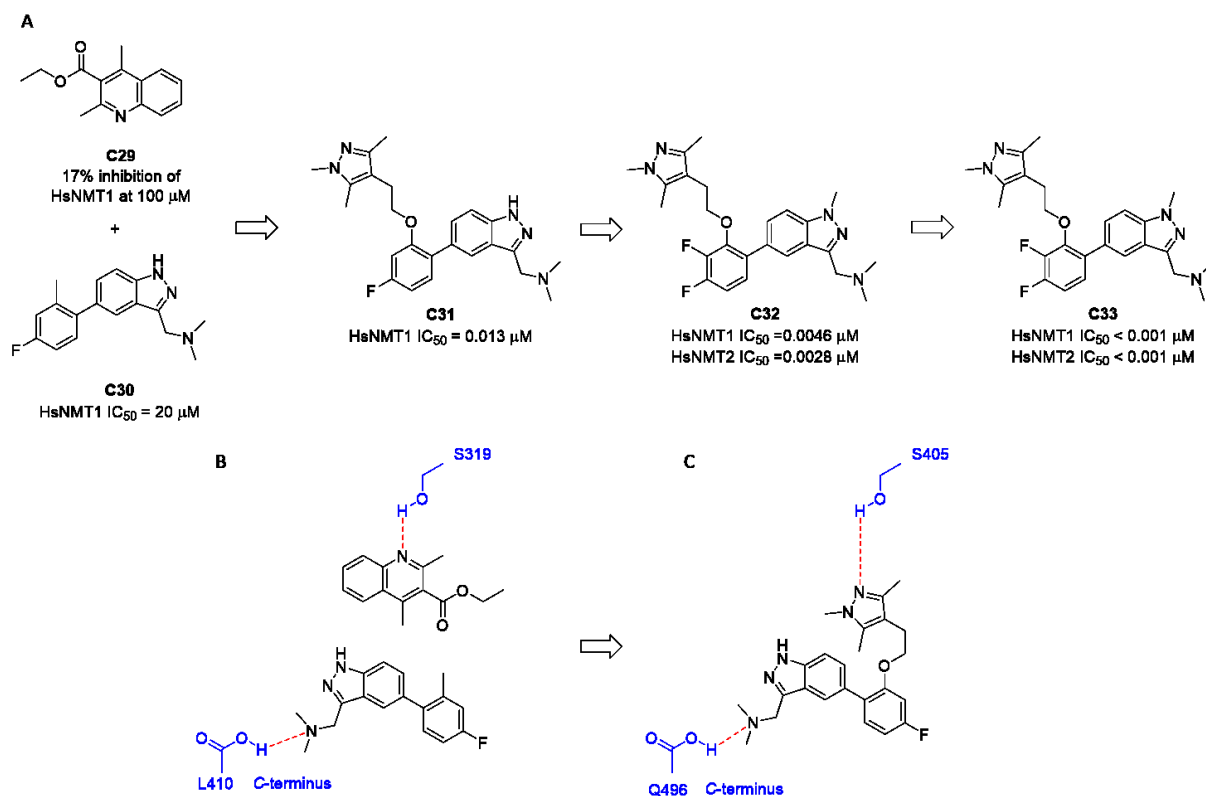


Figure 26 Development of Imperial College indazole series (Modified from Mousnier *et al*⁴⁷). (A) Hybridization of two complementary structures and the activity improvement; (B) A schematic drawing of the binding mode based on quaternary structure of Pv NMT bound with NHM, **C29** and **C30** (PDB: 5O4V); (C) A schematic drawing of the binding mode of **C31** in Hs NMT1 (PDB: 5O6H).

During the writing up period of this thesis, Mousnier *et al* published a series of potent human NMT inhibitors with picomolar IC_{50} s against Hs NMT1 and Hs NMT2 as well as low nanomolar antiviral activity. Based on the crystal structures of **C29** (Figure 26, a simplified fragment based on the Imperial College quinoline series) and **C30** (Figure 26, an analogue of the Pfizer screen hit **PF-02378143**), Mousnier *et al.* identified an opportunity to combine these two fragments and expanded **C30** into the pocket occupied by **C29**. The resulting indazole series represented by **C31** (**IMP-0000917**) picked up the key hydrogen bonds with Ser405 in Hs NMT1 (S319 in Pv NMT) and retained the key hydrogen bonds with Gln496 in Hs NMT1 (Leu410 in Pv NMT), showing a nano-molar inhibitory activities against Hs NMT1⁴⁷. *N*-methylation to reduce the potential for the elimination of dimethylamine and movement of the

fluorine substituent on the phenyl ring from meta-position to ortho-position relative to the ether linker gave **C32 (IMP-0001031)** with single-digit nanomolar IC₅₀s against both Hs NMT1/2. The ortho-fluorine substitution changed the conformation of its adjacent ether linker, enabling its pyrazole ring to form a stronger hydrogen bond with S405. Reintroducing the meta-fluorine discovered **C33 (IMP-0001088)** with subnano-molar IC₅₀, which also showed potent antiviral activity against Rhinoviruses by blocking the co-translational myristoylation of a specific virus-encoded protein (VP0)⁴⁷. This study suggested a new potential application of human NMT inhibitors as a treatment of myristoylation-dependent picornavirus infections (the target validation studies are described in 1.3.2).

1.5 Objective

To further explore the functions of Hs NMTs and examine their potential as drug targets, small molecule inhibitors of Hs NMTs are required as tool compounds. Although a few inhibitors targeting other NMTs show inhibition against Hs NMTs, only two specific studies of small-molecule inhibitor development targeting human NMTs have been reported^{47, 69}.

The binding mode insights revealed from existing NMT inhibitors, e.g. the importance of polar interactions with the enzyme C-terminus, and the established protein production and crystallography methods to access the binding modes of small molecule ligands in NMTs, enabled a structure-based drug design and development approach.

With the aid of crystallography, this project aims to develop potent Hs NMT inhibitors as useful tools for the future studies of human NMT functions and as potential treatments of *N*-myristoylation dependent diseases.

To make a full use of the knowledge of NMT inhibition accumulated and the testing methods developed in our group, this project also includes a minor aim to validate NMTs of less-explored species, e.g. Nematode NMTs, as drug targets and facilitate the development of their NMT inhibitors.

2 Medicinal Chemistry

2.1 Dihydrothiazole series

2.1.1 Previous work on dihydrothiazole series

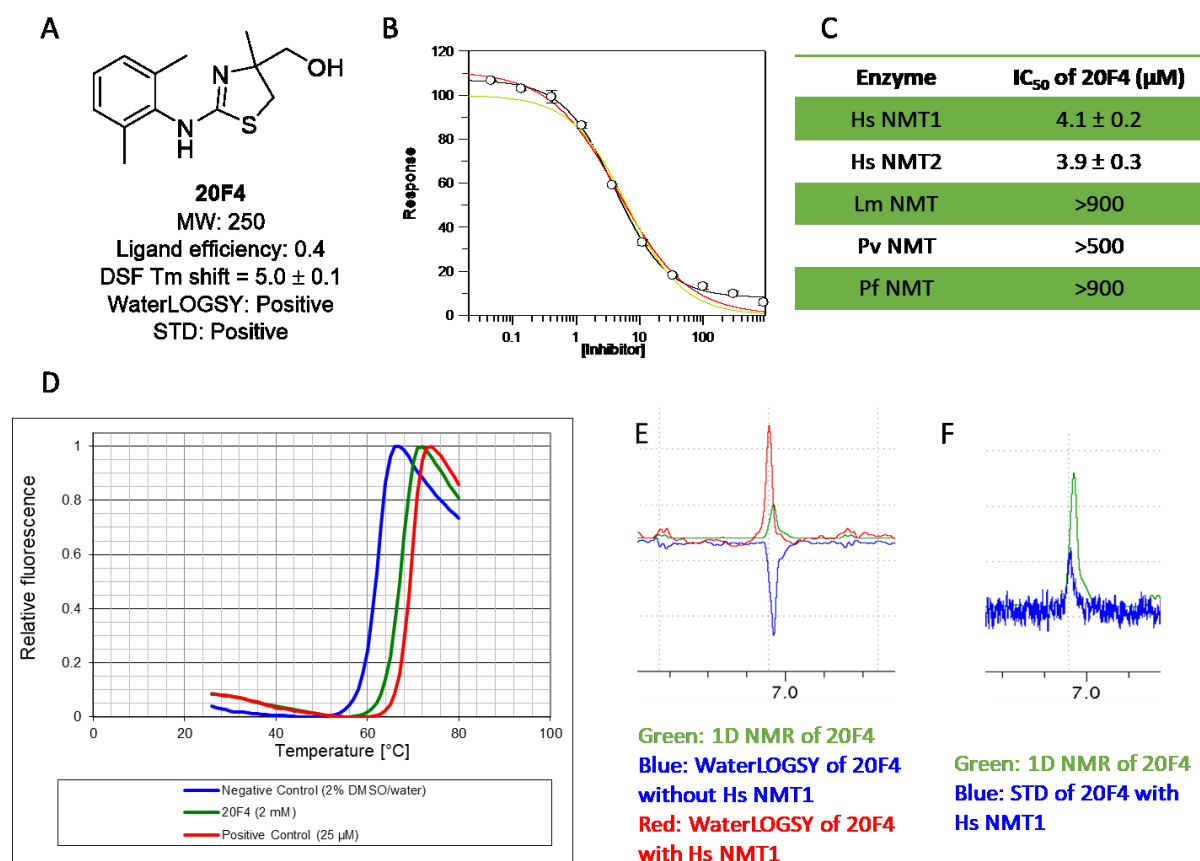


Figure 27 Previous data of **20F4** (reproduced from Eric Lee¹¹⁴ under Creative Commons use licence). (A) Structure of **20F4** and a summary of the validation assay results; (B) Dose-response curve of **20F4**; (C) IC₅₀s of **20F4** against five NMTs; (D) DSF curve of **20F4**; (E) WaterLOGSY assay result of **20F4**; (F) STD assay result of **20F4**.

The dihydrothiazole series was developed by a previous master student of our group, Eric Lee¹¹⁴, by screening 2320 fragments from the Maybridge RO3 Library using the differential scanning fluorimetry (DSF) assay¹¹⁵. With a good ligand efficiency (LE = $\text{pIC}_{50} \times 1.374 / \text{number of non-hydrogen atoms}$) of 0.4, the most potent fragment **20F4**, was further validated by Eric Lee using Water-Ligand Observed by Gradient Spectroscopy (WaterLOGSY) and Saturation Transfer Difference (STD) spectroscopy (Figure 27)¹¹⁴. **20F4** showed 5-degree

temperature shift in the DSF assay (Figure 27D). The WaterLOGSY (Figure 27E) signal flipped from negative to positive after adding Hs NMT1 and the STD (Figure 27F) also showed a positive signal after adding Hs NMT1, both results indicated the binding of **20F4** to Hs NMT1. The CPM assay results of **20F4** further indicate an IC₅₀ of 4.1 μ M against Hs NMT1 with a typical dose-dependent curve (Figure 27B). **20F4** was also tested on other NMTs (Figure 27C). Interestingly, **20F4** has a similar potency against Hs NMT2 but lacks activity against NMTs from *Leishmania major* (Lm), *Plasmodium vivax* (Pv) and *Plasmodium falciparum* (Pf). To explore the SARs, 18 analogues were synthesized based on **20F4** by Eric Lee (Imperial College London)¹¹⁴. 10 of them showed modest Hs NMT1 inhibition with IC₅₀s from 2.8-139 μ M (4 best ones are listed in Table 2). This study follows up Eric Lee (Imperial College London)'s findings to further develop this series.

2.1.2 CPM assay

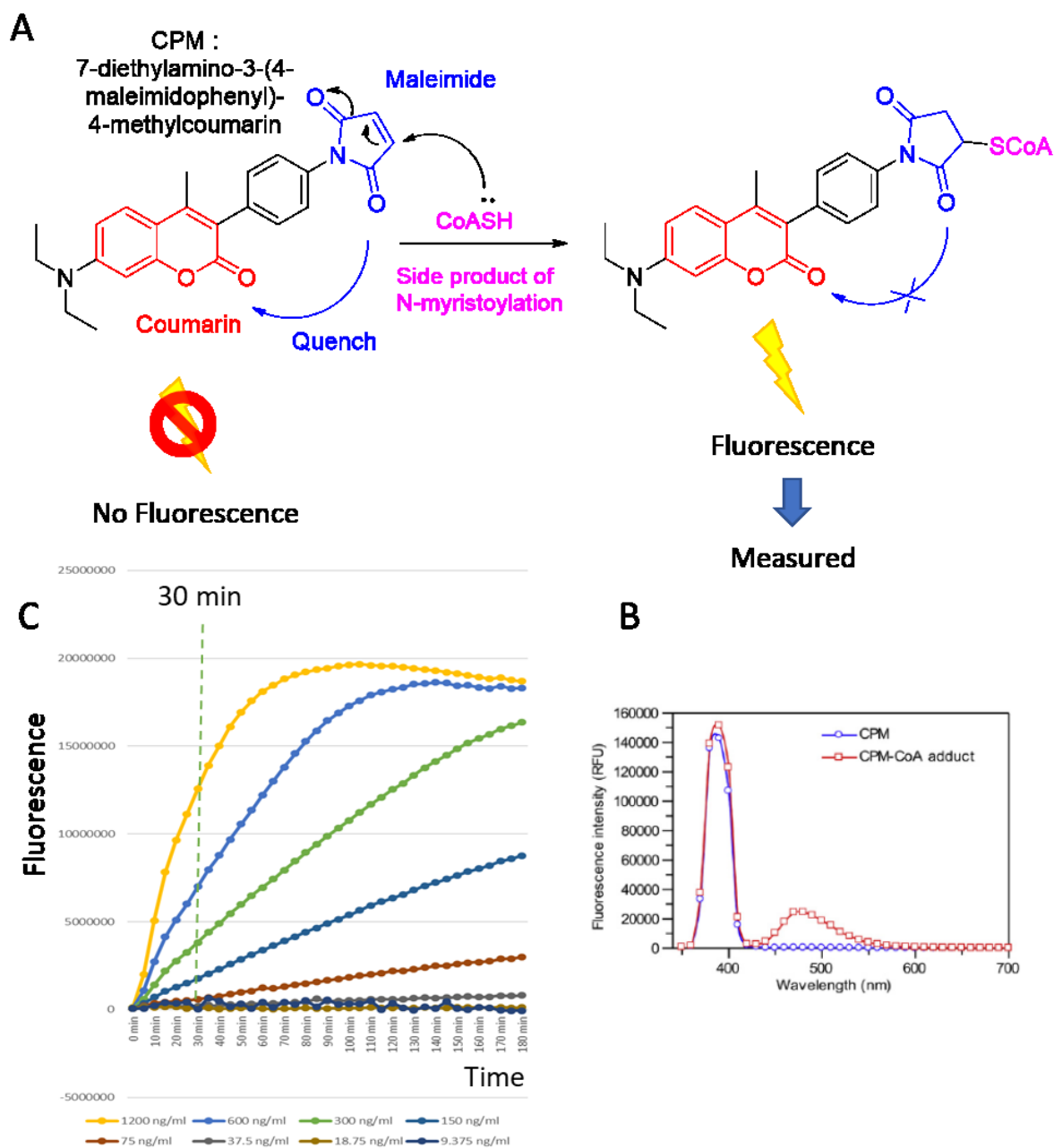


Figure 28 Mechanism of CPM assay. (A) The fluorescence of CPM can be released by reacting with the side product of *N*-myristoylation; (B) Changes in fluorescence emission spectra after CoA addition with excitation at 380 nm (reproduced with permission obtained from RightsLink¹¹⁶); (C) Changes in fluorescence over time during the CPM assay at eight different enzyme concentrations of Hs NMT1. 300 ng/ml (Green) was chosen as the assay concentration and 30 min was chosen as the assay endpoint. Graph generated by Excel 2017.

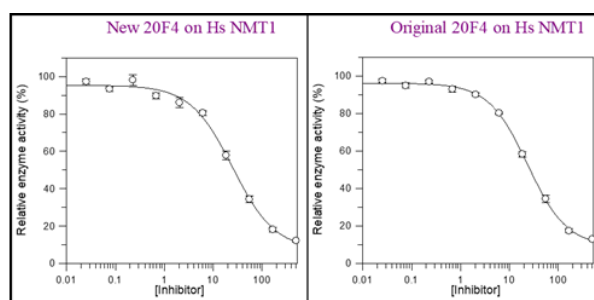
The CPM assay is a fluorogenic enzyme assay for NMT activity testing, which is developed in our group¹¹⁶, which was used to measure IC₅₀s in this project. It monitors CoA-SH production in real time using a commercially-available probe, 7-diethylamino-3-(4-maleimidophenyl)-4-

methylcoumarin (CPM). CPM contains a fluorogenic coumarin group which is quenched by the neighbouring maleimide group. Michael addition between the myristoylation by-product CoA thiol (Figure 28A) and the maleimide restores the fluorescence of CPM at 490 nm (Figure 28B). The substrates of the myristoylation used in the assay are two synthetic peptide substrates based on the *N*-terminal sequence of the catalytic subunit of *H. sapiens* proto-oncogene tyrosine kinase p60-Src¹¹⁷⁻¹¹⁸: Hs p60-Src₍₂₋₉₎ (8AA) for both human NMTs and Hs p60-Src₍₂₋₁₆₎ for Pv, Pf and Lm NMTs. The intensity of fluorescence is linear relative to the amount of CoA-SH generated by *N*-myristoylation. The speed of *N*-myristoylation is constant until one of its substrates, Myr-CoA or the peptide substrate, is used up. For example, the speed of *N*-myristoylation of Hs NMT1 at the testing condition (300ng/ml with 4 μ M of peptide substrate and Myr-CoA) keeps constant for the initial 2.5 hours (Figure 28C). Therefore, by measuring the fluorescence at 490 nm 30 minutes after the start of the reaction, dose-dependent inhibition curve can be generated to measure compounds' IC₅₀ values.

2.1.3 Validation

Table 1 IC₅₀s of the new **20F4** and original **20F4**.

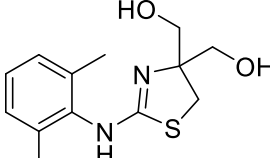
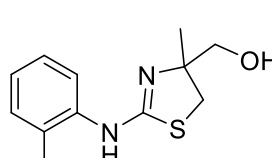
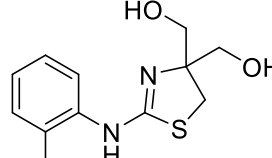
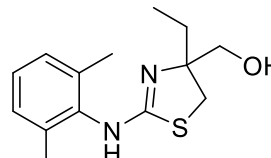
(μ M)	Hs NMT1 IC ₅₀	Standard Error
New 20F4	25	3.8
Original 20F4	24	1.4



All the previous results from Eric Lee (Imperial College London) suggested a potential Hs NMT1 inhibitor series with good ligand efficiencies¹¹⁴. The initial aim of this project is to validate this series and improve its potency. Firstly, the compound **20F4** was re-purified by LC-MS and its IC₅₀s was tested in parallel with the original **20F4** sample (Table 1).

Considering the standard error, there is no significant difference between the IC₅₀ values of the new and original **20F4** (25 μ M and 24 μ M), showing repeatable inhibition against Hs NMT1.

Table 2 IC₅₀s of four **20F4** analogues against Hs NMT1, Hs NMT2 and Lm NMT¹¹⁴.

 20F4-09 MW: 266  20F4-12 MW: 236  20F4-13 MW: 252  20F4-14 MW: 264				
Compound / IC ₅₀ (μ M)	Hs NMT1		Hs NMT2	Lm NMT
	Current study	Previous work		
20F4-09	3.6 $\pm$ 0.4	8.5 $\pm$ 0.7	3.9 $\pm$ 0.5	>200
20F4-12	11 $\pm$ 4.5	9.7 $\pm$ 1.0	10 $\pm$ 0.8	>200
20F4-13	3.2 $\pm$ 0.4	6.3 $\pm$ 0.5	5.8 $\pm$ 0.7	>200
20F4-14	1.7 $\pm$ 0.3	2.8 $\pm$ 0.3	1.3 $\pm$ 0.2	>200

To cover more compounds in the validation, the most potent four analogues of **20F4** from the previous work were chosen to be re-purified by LC-MS and tested against Hs NMT1 (Table 2) by the CPM assay. The results were comparable as well. These compounds were also tested against Hs NMT2, Lm, Pv and Pf NMTs to validate whether this series is human selective (Table 2). The IC₅₀s of each inhibitor on Hs NMT1 and NMT2 were comparable. The CPM assay results of Lm, Pv and Pf NMTs showed no dose-dependent decrease of enzyme activity, indicating no inhibition. However, the positive controls of Pv and Pf NMTs also did not show dose-dependent curves, which is likely a result of the loss of the activities of Pv and Pf NMTs. Therefore, their data are not reliable and not listed. Since the previous work used the same

batch of Pv and Pf NMTs, the previous Pv and Pf activity results are also invalid. To prove this series as human selective, new Pv and Pf NMTs should be produced and used for testing.

2.1.4 Abnormal inhibition curves

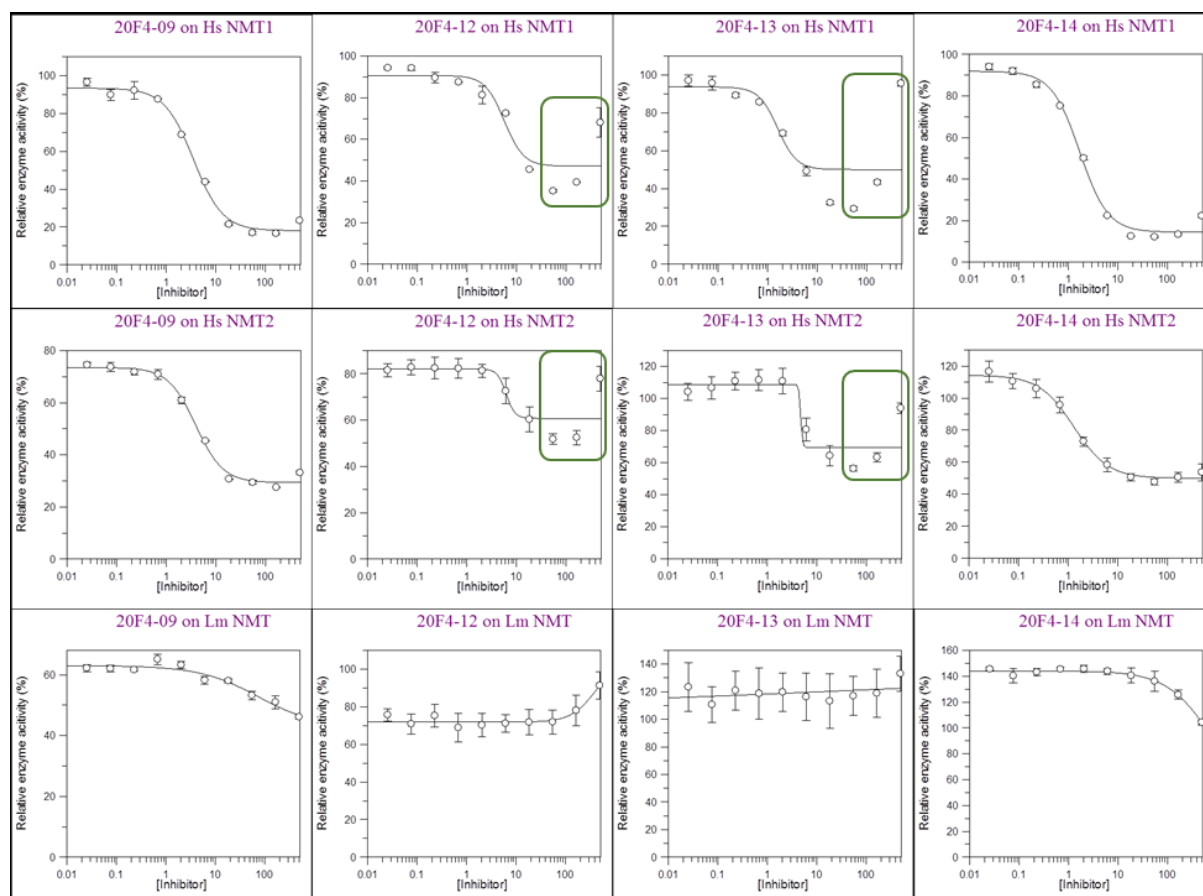


Figure 29 Dose-response curves of four **20F4** analogues on Hs NMT1, Hs NMT2 and Lm NMT (inhibitor concentration in μM). Inhibition Curves fitted by Grafit 7 (Erithacus Software Limited) to a 4 parameter equation $y = \text{Range}/[1+(x/\text{IC}_{50})^s] + \text{Background}$, where s is the Hill slope factor¹¹⁹. Green circles highlight the abnormal upwards bends. Where the abnormal upward bends occur, the data points at 167 μM and 500 μM were not used in the calculation of IC_{50} s.

Although the IC_{50} s of human NMTs and Lm NMT are consistent with the previous work, the inhibition curves were in an abnormal shape, especially for **20F4-12** and **20F4-13**. Their dose-response curves showed an increasing response at high inhibitor concentrations (highlighted in green circles in Figure 29). Since the fluorescence measured was generated by reacting with the *N*-myristoylation by-product, the intensity of the fluorescence should be positively related

to the enzyme activity. Therefore, a typical NMT inhibition curve should be a characteristic "S"-shaped curve (sigmoid curve) start from 100% activity at the lowest inhibitor concentration and gradually decrease to 0%¹¹⁶. At the end of the curve where inhibitor concentration reaches the highest, the fluorescence is expected to be the lowest. A series of assays were designed to investigate this phenomenon.

2.1.5 Comparison assay adding inhibitors after quenching the reaction

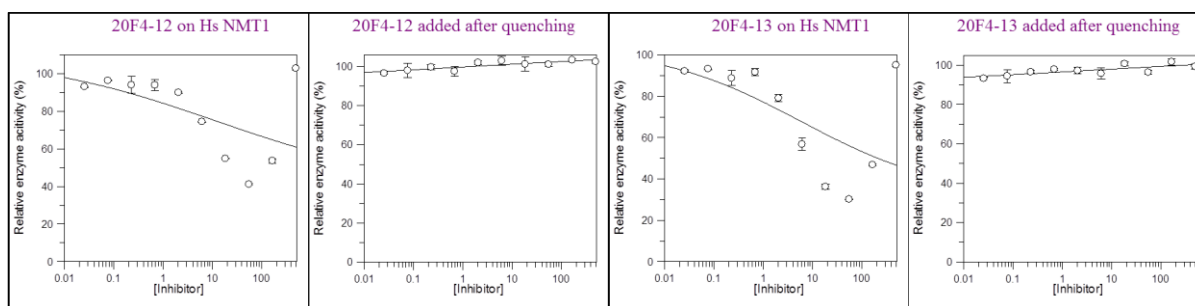


Figure 30 Comparison assay that adds inhibitors **20F4-12** and **20F4-13** after quenching the CPM reaction (inhibitor concentration in μM).

The two possibilities of the abnormal increase of fluorescence at inhibitor concentration 167 and 500 μM are fluorescence of the inhibitors and fluorescent products generated by certain side-reaction with the inhibitors. To investigate the first possibility, a comparison assay was done by adding the two inhibitors after quenching the CPM reaction (Figure 30). The inhibitors added were the same in concentration as the original CPM assay (from 500 μM to 0.025 μM) but were added after 30 minutes from the start of the assay and after the reaction was quenched by a acidifying buffer. On the inhibition curves, all inhibitor concentrations showed 100% fluorescence. Only a negligible increase in fluorescence was detectable with increasing inhibitor concentrations, suggesting that the inhibitors are not fluorescent and no fluorescence-generating side-reaction happens after quenching.

2.1.6 Continuous CPM assay with or without enzyme

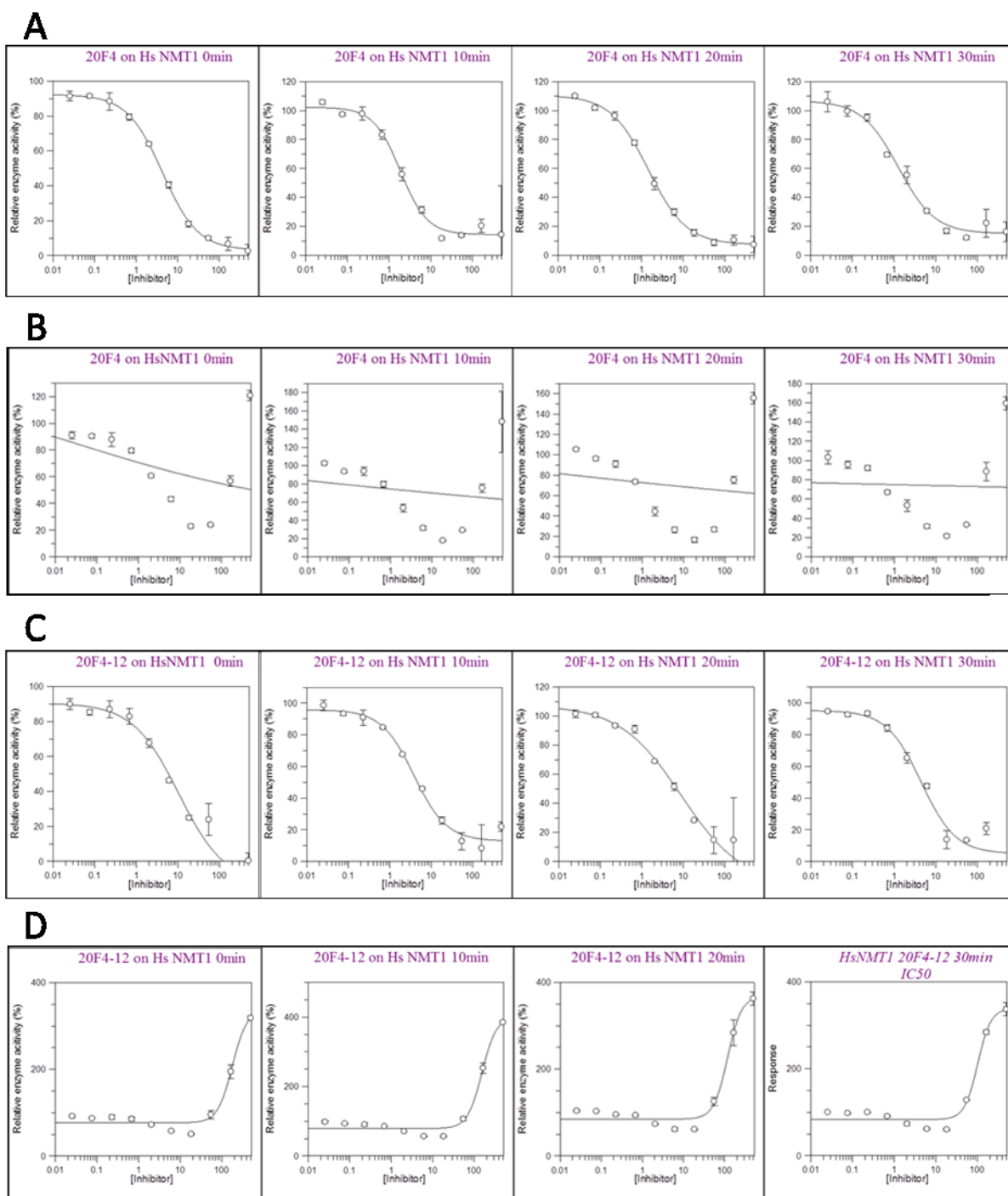


Figure 31 Continues CPM assay with or without enzyme (inhibitor concentration in μM). (A) CPM result of **20F4** corrected by data without adding enzyme; (B) Original CPM result of **20F4**; (C) CPM result of **20F4-12** corrected by data without adding enzyme; (D) Original CPM result of **20F4-12**.

In one CPM assay, **20F4** was also found to be having an upward bend at the end of the curve. The only difference from the original assay where **20F4** showed a normal curve is that there

were more compounds to test during the assay, thus the entire assay took longer and **20F4** was incubated with the assay solution longer. This phenomenon shows that the abnormal fluorescence is time dependent and could be generated by inhibitor side-reactions with enzyme, CPM, peptide or other assay reagents. Therefore, a continuous CPM assay with **20F4** and **20F4-12** was performed by measuring fluorescence at every ten minutes. For comparison, the assay was performed both in presence and absence of the enzyme. Surprisingly, subtracting the no-enzyme fluorescence measurements from the results in presence of NMT resulted in a typical dose-dependent curves (Figure 31). These findings proved that the increasing fluorescence signals at high inhibitor concentrations are enzyme independent.

Taking a closer look at the original data, the fluorescence at high inhibitor concentration kept increasing during the time and the high-concentration fluorescence of **20F4-12** was much stronger than **20F4**'s, which is consistent with previous results and indicated that **20F4-12** and **20F4-13** had more side-reactions than **20F4**.

2.1.7 Comparison CPM assay with or without CPM

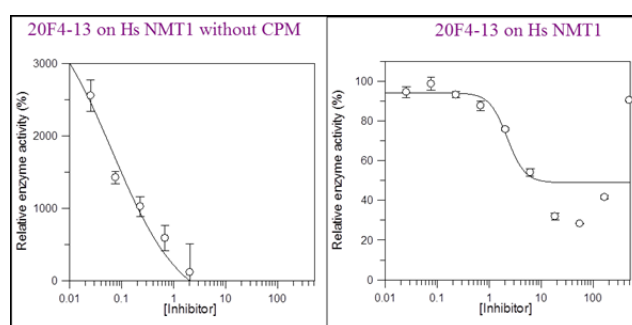


Figure 32 Comparison assay with or without CPM on **20F4-13** (inhibitor concentration in μM).

After ruling out the possibility of side-reactions between the inhibitor and the enzyme, the remaining potential substrates were CPM, peptide and assay buffer. Among them, CPM seems to be the most reactive one. Therefore, CPM assay with or without CPM was performed on **20F4-13** (Figure 32). The result showed that there is no end fluorescence generated when CPM

is not present, which proved that CPM is responsible for the abnormal fluorescence. Considering all the previous assay results, the initial conclusion was that the fluorescence is likely from side-reaction between CPM and the inhibitors.

2.1.8 Proposed mechanism

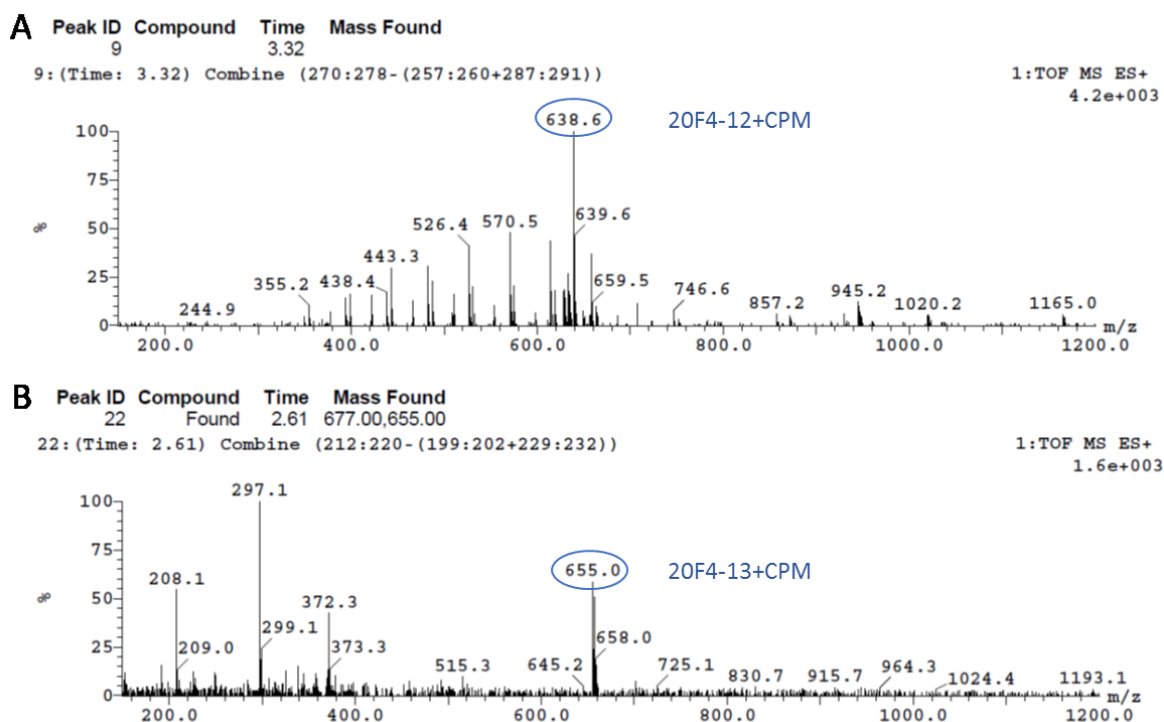


Figure 33 High-performance liquid chromatography mass spectrometry results of CPM and inhibitor mixtures.
(A) **20F4-12** with CPM; (B) **20F4-13** with CPM.

To analyse the fluorogenic side reaction, mixtures of CPM and inhibitors in the assay buffer are sent for high-resolution mass analysis by High-performance liquid chromatography (HPLC). In samples of the two high fluorescent inhibitors **20F4-12** and **20F4-13**, a peak of the mass that is exactly the sum of the molecular weight of inhibitor and CPM was found (Figure 33). This result indicated the formation of an adduct of the inhibitors and CPM. The sum peak was not found in **20F4**, possibly due to less adduct formation and the limitation of the LC-MS resolution. According to the CPM results, **20F4** has much weaker fluorescence increase at high inhibitor concentration, indicating less side-product generation.

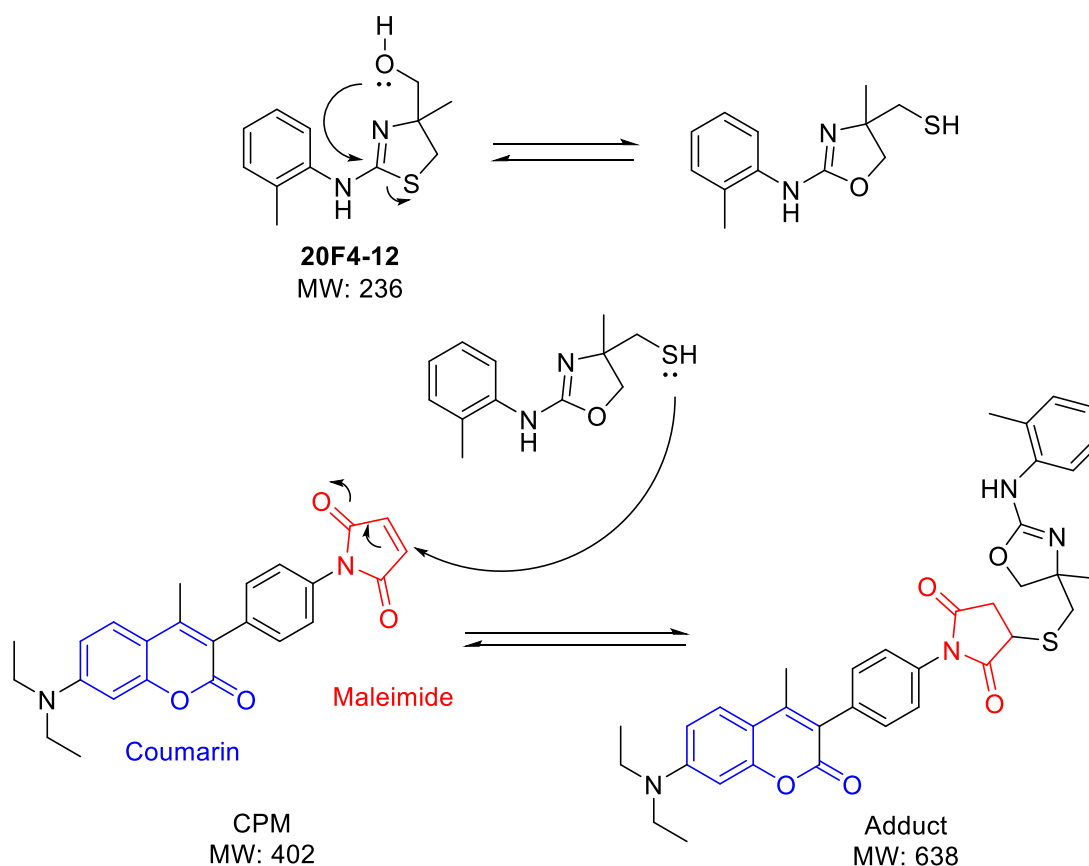


Figure 34 Proposed mechanism of the side-reaction using **20F4-12** as an example.

As the HPLC mass result shows, this reaction should be an addition between the inhibitor and CPM without adding or losing atoms. On the CPM compound, the possible reacting spot is the maleimide, where the Michael addition happens. On inhibitors, the two high fluorescence compounds both have one methyl group compared to two methyl groups of the other compounds tested. Their higher fluorescence intensity at high inhibitor concentrations indicates that steric hindrance is affecting the side-reaction rate and the reaction should be happening at a place where the methyl group's steric hindrance can affect. Based on all the observations, a mechanism of the side-reaction was proposed (Figure 34). The hydroxy group attacks the electron-poor C2 of the thiazol group and breaks the carbon-sulphur bond, generating a tautomer with a thiol group. This tautomer could then attack the maleimide in the CPM compound and generate the corresponding adduct. Due to the irreversibility of the Michael addition, the tautomeric equilibrium shifts towards the thiol containing derivative until

the inhibitor is used up. The slower reaction rates of inhibitors with two methyl groups can be explained by the slower tautomerization caused by steric hindrance.

2.1.9 Oxygen substitute

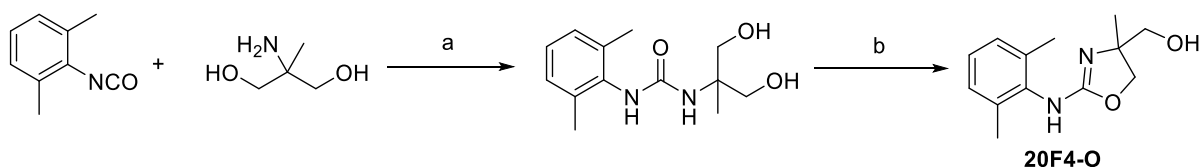


Figure 35 Synthesis of (2-((2,6-dimethylphenyl)amino)-4-methyl-4,5-dihydrooxazol-4-yl)methanol (**20F4-O**).

Reagents and Conditions: (a) Ethyl acetate, reflux, overnight; (b) 37% HCl, reflux, 2 h. Overall yield 6%.

To minimize the side reaction with CPM, an oxygen substitute of **20F4** was synthesized as shown in Figure 35. 2-isocyanato-1,3-dimethylbenzene was reacted with 2-amino-2-methylpropane-1,3-diol in ethyl acetate under reflux to obtain the phenylurea intermediate, which was then refluxed in hydrochloric acid to generate the target compound **20F4-O**¹²⁰. The yield over the two steps was 6%. This could be a result from the small-scale reaction (1mmol) and the steric hindrance of the two methyl groups.

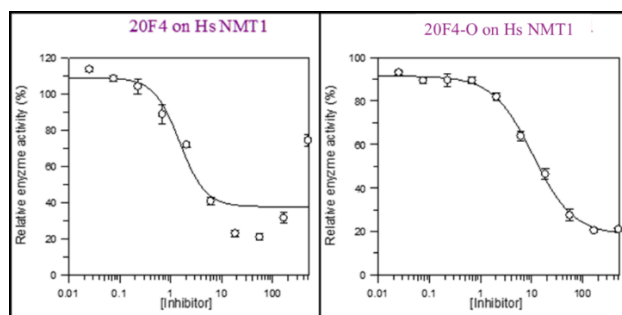


Figure 36 comparison assay between **20F4** and its oxygen substitute **20F4-O** (inhibitor concentration in μM).

The oxygen substitute was tested in parallel with **20F4** and showed a typical inhibition curve (Figure 36). The results indicate that the oxygen to sulphur substitution prevents the side reaction due to a decreased nucleophilicity of the hydroxy group, thereby supporting the proposed mechanism of the CPM adduct formation.

2.1.10 Conclusion and outlook

20F4 and its analogues are inhibitors of both Hs NMT1 and Hs NMT2 with IC_{50} s between 1.3-25 μ M. No selectivity is observed between the two human NMT isoforms, but all these inhibitors do not show inhibition against Lm NMT.

This series, including the oxygen substitute, was later found to be unstable in the standard stock condition (50mM DMSO stock in a Nitrogen atmosphere) and generate detectable impurity (approximately 5%) within one month after synthesis (extra peaks found by LCMS). Considering the side-reaction in the CPM assay interfering the enzyme activity determination and the short bench life of the inhibitors, this series was discontinued.

However, the interesting finding of the inhibitor-CPM side reaction reveals one limitation of the CPM assay. Inhibitors with strong nucleophilic groups or inhibitors that can generate such groups during the assay condition (pH 7.9) can react with the CPM compound and generate interfering fluorescence.

Although it is disappointing to find this seemingly promising series to be not developable, the problems and possible mechanisms were identified quickly with a few assays, saving the time from further synthesis and testing. The analysing methods described in this chapter can be used to analyse future problems of abnormal inhibition curves in the CPM assay.

2.2 Azetidinopyrimidines series

2.2.1 Previous work

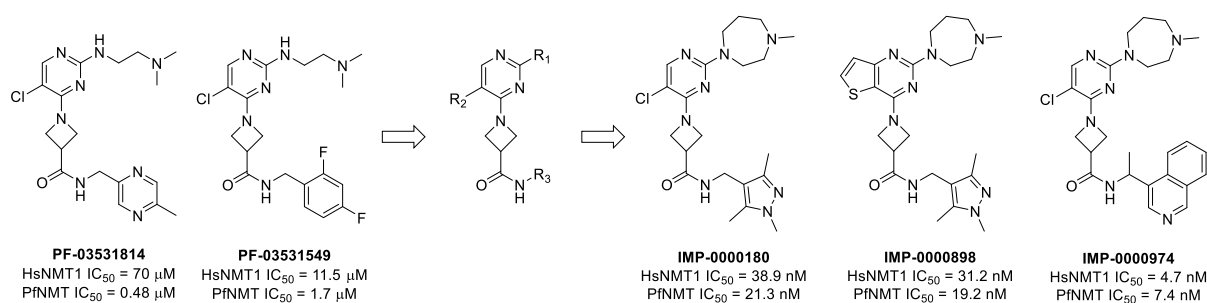


Figure 37 Previous development of azetidinopyrimidines series.

The azetidinopyrimidines series was derived from a screen on 150,000 compounds from the Pfizer Global Diverse Representative Set⁴⁰ against Pf and Ld NMTs (Described in 1.4.2). Two lead compounds **PF-03531814** and **PF-03531549** sharing the same azetidino-pyrimidine core showed IC_{50} s of 0.48 μ M and 1.7 μ M against PfNMT, and IC_{50} s of 70 μ M and 11.5 μ M against Hs NMT 1, respectively. They were chosen and developed by previous group members, *Dr* Zhiyong Yu and *Dr* Jennie Hutton, to target *Plasmodium* NMTs (Unpublished work). A series of modifications were performed on the R₁ and R₂ of the pyrimidine ring, and R₃ on the carboxamide (Figure 37). 81 compounds were made and the IC_{50} s of Pf NMT were lowered to single-digit nanomolar scale. However, despite the great efforts to avoid human NMT inhibition, the selectivity over Hs NMT1 was lost as the potency on *Plasmodium* NMT improved, showing a few nanomolar IC_{50} s against Hs NMT1. Considering the possible side-effects on the normal host NMTs, this is not ideal for *Plasmodium* inhibitors but a promising start for the development of drugs for human *N*-myristoylation relative diseases like cancer and biological tools for human NMT functional study. Another issue of this series is the high EC_{50}/IC_{50} ratios, which indicated that good enzyme activity does not efficiently transfer into good cell activity. Therefore, the aim of this study is to improve the EC_{50}/IC_{50} ratio of the azetidinopyrimidines series and retarget it to human NMT inhibitors.

2.2.2 Crystal structure analysis

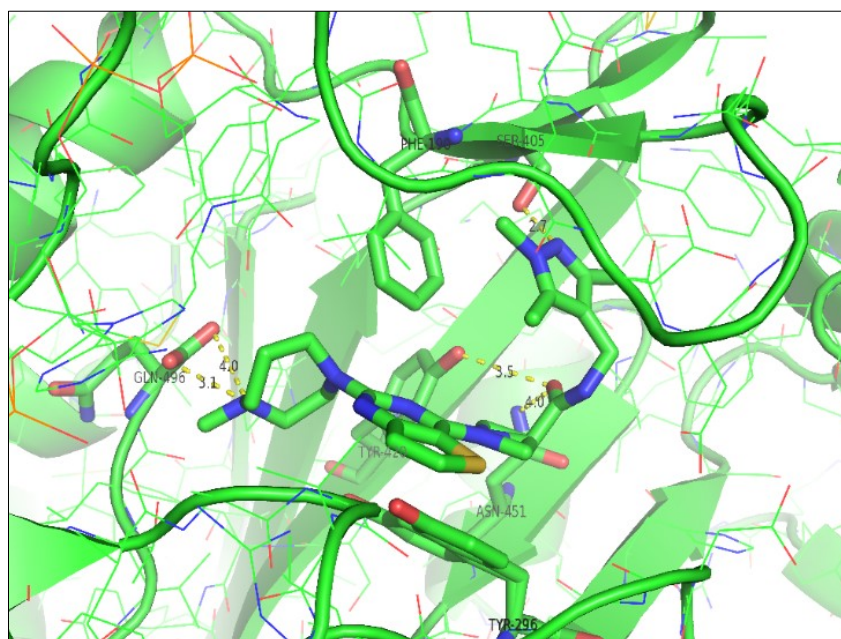


Figure 38 Binding mode of **IMP-0000898** in Hs NMT1 (Unpublished structure from *Dr James Brannigan*). The ligand binding mode was analysed using PyMOL²⁸. The interacting residues in the binding pocket were shown in stick and their names were labelled in black. The polar interactions were shown in yellow dashes and their lengths were shown in black labels.

At the beginning of this study, one co-crystal structure of Hs NMT1 with **IMP-0000898** was available (1.64 Å, solved by *Dr James A. Brannigan*, University of York). The binding mode was analysed by PyMOL²⁸ (Figure 38). The ligand is binding in the peptide binding pocket with its diazepane pointing towards the C-terminal and its trimethylpyrazole bending upwards to interact with the Ser405, forming an S-shape. The 1N on trimethylpyrazole forms a 2.7 Å hydrogen bond with the hydroxy group of Ser405. The amide oxygen is interacting with the phenolic hydroxyl group of Tyr420 by a 3.5 Å hydrogen bond and the amide of Asn451 by a 4.0 Å hydrogen bond. The 4N on the diazepane is weakly interacting with the carboxyl group of the C-terminal Glu496 at distances of 3.1 and 4.0 Å. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at the human blood pH 7.4, 90.46% of the ligand has a protonated 4N on the diazepane. The carboxyl group of the Gln496 is deprotonated under biological condition. Therefore, the interaction between the diazepane and the glutamate is ionic bonding,

which is stronger than hydrogen bonds and non-directional. There are also π - π stacking interactions of Phe190 with the trimethylpyrazole and Tyr296 with the thieno[3,2-d]pyrimidine.

2.2.3 Change of the azetidine linker

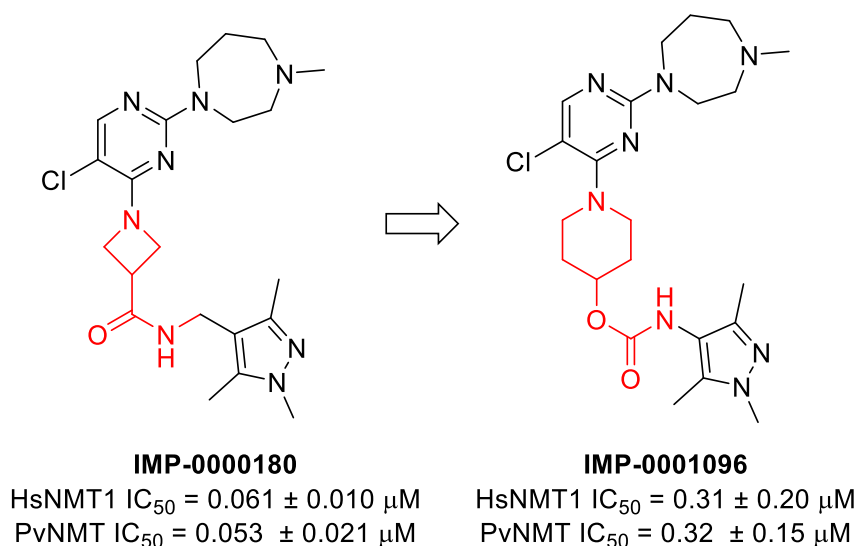


Figure 39 linker change of the azetidinopyrimidines series.

According to the crystal structure, there are three parts of **IMP-0000898** that form polar interactions with Hs NMT1: the diazepane head, the azetidine linker and the trimethylpyrazole tail. Both the diazepane and the trimethylpyrazole have been widely explored in previous work, but the azetidine linker had never been changed. The lengths of the hydrogen bonds of the azetidine linker, 3.5 Å to Tyr420 and 4.0 Å to Asn451, are longer than the ideal length and could be optimized by an alternative linker. Therefore, this study starts by changing the linker to explore wider chemical space and optimize its hydrogen bonds.

An ideal linker should keep the acyl group for hydrogen binding, link the pyrimidine and trimethylpyrazole with similar length, and be flexible enough to bend to form an S-shape. Considering all these requirements and the feasibility of synthesis, the piperidinyl carbamate was chosen as the new linker.

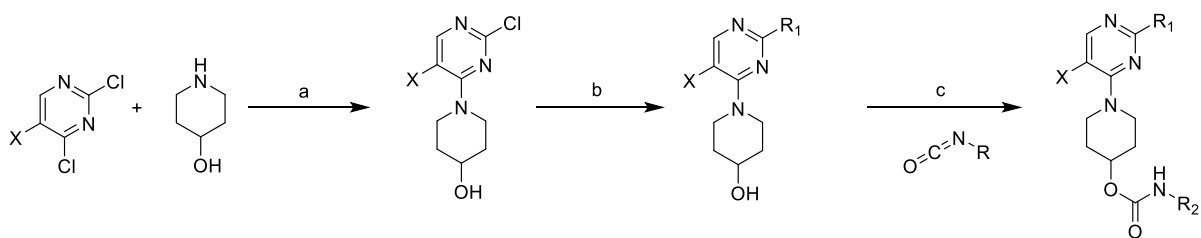


Figure 40 Synthesis of the new azetidinopyrimidines series with the piperidinyl carbamate linker. Reagents and Conditions: (a) DMF, Et₃N, rt, overnight, yield: 68.7-77.9%; (b) Amine, *n*-BuOH, microwave 200 °C, 1.5 h, yield: 8.8-77.6%; (c) Toluene, microwave 170 °C, 7 h, yield: 8.8%-84.1%.

Piperidin-4-ol selectively substituted the 4Cl on 2,4,5-trichloropyrimidine to form the 1-(2,5-dichloropyrimidin-4-yl)piperidin-4-ol, which was then reacted with 1-methyl-1,4-diazepane to be selectively substituted at the 2Cl position. The selectivity of 4Cl and 2Cl over 5Cl is due to the orientation effect of the two nitrogen atoms on the pyrimidine as the nucleophilic aromatic directing groups, activating their *ortho* and *para* positions. The selectivity of 4Cl over 2Cl is due to the lone pairs of the two nitrogen atoms on the pyrimidine. Locating in the sp² orbitals, these two pairs stick out of the aromatic ring, preventing the negatively charged nucleophile to approach the 2C atom to form the *ipso* substitution intermediate. The alcohol on the resulting intermediate, 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol, was then attacked by 4-isocyanato-1,3,5-trimethyl-1*H*-pyrazole to form the target compound **IMP-0001096** (Figure 40 where X=Cl) ¹²².

The enzyme activity of **IMP-0001096** was tested by the CPM assay described in 2.1.2 and showed an IC₅₀ of 0.31 ± 0.20 μM against Hs NMT1 and 0.32 ± 0.15 μM against Pv NMT, which is five-fold (Hs NMT1) and six-fold (Pv NMT) of **IMP-0000180** with the original azetidine linker. Therefore, the new linker was well tolerated by the binding pocket and retained comparable potencies.

2.2.4 Change of the trimethylpyrazole tail

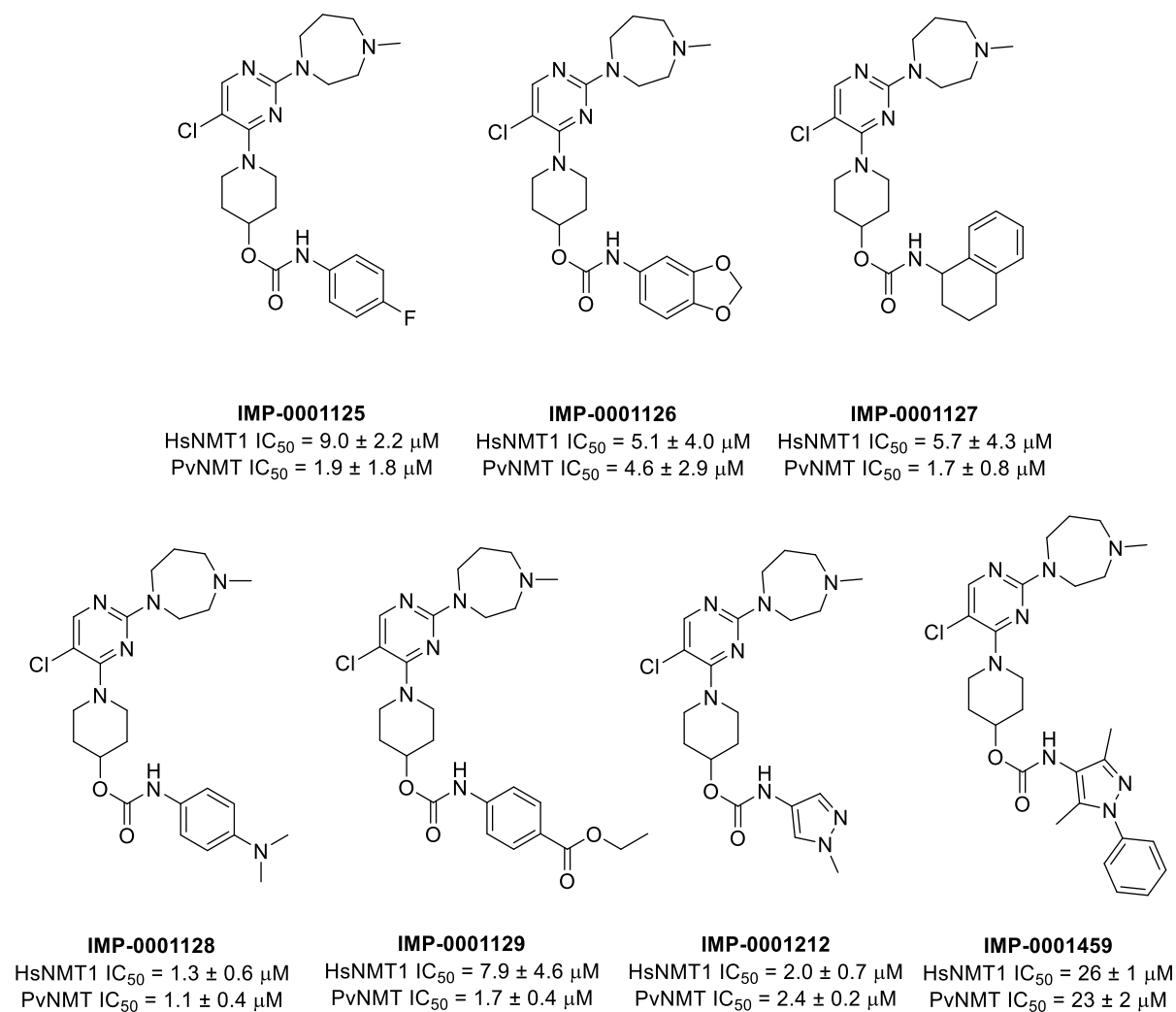


Figure 41 Trimethylpyrazole tail change of the new azetidinopyrimidines series.

According to the crystal structure of **IMP-0000898**, the loss of the potency with the new piperidinyl carbamate linker is likely due to the position change of the trimethylpyrazole tail, which forms two critical interactions – a hydrogen bond with Ser405 and π - π stacking interaction with Phe190. To optimize the potency and explore the SARs of the new linker, seven compounds with different tails were designed (Figure 41) and synthesised using the same synthesis route as **IMP-0001096** (Figure 40).

Their enzyme activities of Hs NMT1 and Pv NMT were tested using the CPM assay. **IMP-0001125** and **IMP-0001127** were designed without a hydrogen acceptor to test the importance

of this hydrogen bond. Their IC₅₀ values against Hs NMT1 were 43- and 27-fold of **IMP-0001096**'s, which showed the critical role of the original hydrogen bond. The π - π stacking interaction is more tolerant with the position of the tail, as **IMP-0001127** showed a similar potency to **IMP-0001125** with a phenyl group extending further away. **IMP-0001126** and **IMP-0001129** were designed with alternative hydrogen bonding acceptors and showed 25- and 38-fold IC₅₀s comparing to **IMP-0001096**'s. The similar IC₅₀s of **IMP-0001125** and **IMP-0001127** indicated that the alternative hydrogen bonding acceptors are either out of place or too weak that they do not form effective hydrogen bonds with Ser405. **IMP-0001212** and **IMP-0001459** were designed with the same pyrazole as **IMP-0001096** but was either only substituted by one methyl group or substituted with a phenyl group. Offset parallel π - π interaction is stabilized by the favouring London dispersion force and solvophobic effects, over the disfavouring repelling force between the negative quadrupole on the top and bottom of the aromatic rings¹²³. Removing the electron donating methyl groups on the pyrazole should therefore strengthen the π - π stacking interaction. The 10-fold potency loss of **IMP-0001212** further shows the low potency impact of the π - π stacking interaction and could be a result of the increased flexibility of the pyrazole. The 120-fold potency loss of **IMP-0001459** should be caused by the steric hindrance of the phenyl group. **IMP-0001128** with both a phenyl ring to form π - π stacking and a tertiary amine as a hydrogen bond acceptor is the best among these seven compounds, showing a 6-fold IC₅₀ of **IMP-0001096**'s. The weaker potency indicates that the position of the hydrogen donor is not optimal. The IC₅₀s of Pv NMT were also tested and they are comparable to the IC₅₀s of Hs NMT1.

The attempt to change the tail indicated the importance of the hydrogen bonding to the Ser405, the tolerance of the π -system positions, and the limit size of the binding pocket near the tail.

2.2.5 Change of the diazepane head

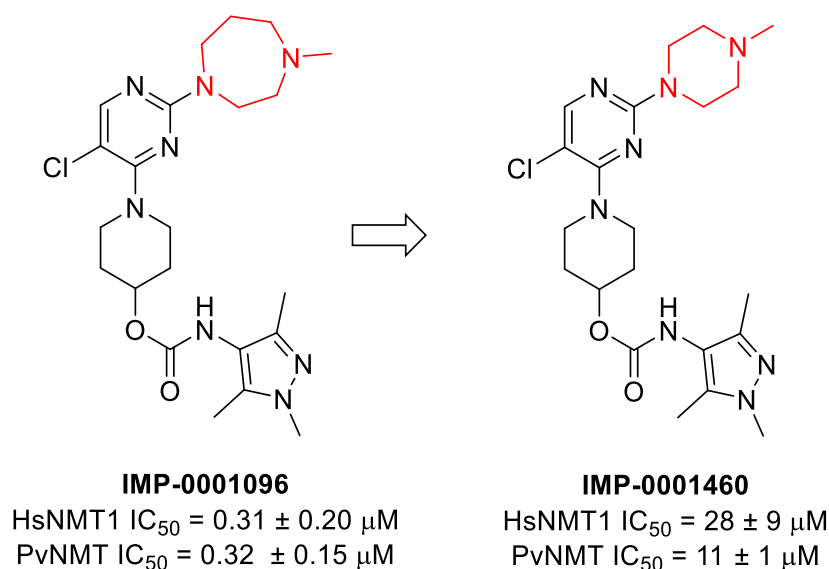


Figure 42 Diazepane head change of the new azetidinopyrimidines series.

According to the crystal structure and pKa analysis, the diazepane head is binding with the C-terminal Gln496 by ionic bonding. The force between two ions are inversely proportional to the square of their distance, therefore, shortening the distance can effectively strengthen the interaction. According to the crystal structure (discussed in 2.2.9), the original distances are 3.1 and 4.0 Å and the diazepane head is long enough to reach closer to the C-termini but bends away from it, which could be a result of steric hindrance. Therefore, **IMP-0001460** with a smaller piperazine, was made by the synthesis route in Figure 40.

The potencies of **IMP-0001460** against Hs NMT1 and Pv NMT were tested using the CPM assay. The 90-fold (Hs NMT1) and 34-fold (Pv NMT) of IC₅₀s were a big decrease in potency for a small change of one CH₂. More optimizations are needed to understand the SARs.

2.2.6 Change of the halogen

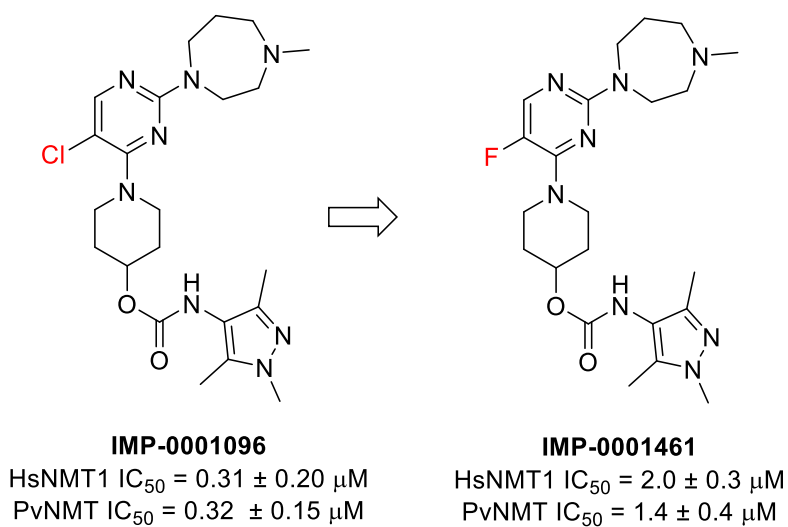


Figure 43 Halogen change of the new azetidinopyrimidines series

The chlorine on the pyrimidine shoulder was also substituted by fluorine to explore the SARs at the 5-position, using the synthesis route in Figure 40. The IC₅₀s of the resulting compound, **IMP-0001461**, were 2.0 ± 0.3 μM against Hs NMT1 and 1.4 ± 0.4 μM against Pv NMT, which were six-fold and four-fold of **IMP-0001096**'s. The potency was generally retained, and the difference could be a result from the electron withdrawing effect of the halogen on the π-π stacking between the pyrimidine and Tyr296.

2.2.7 The MTS assay

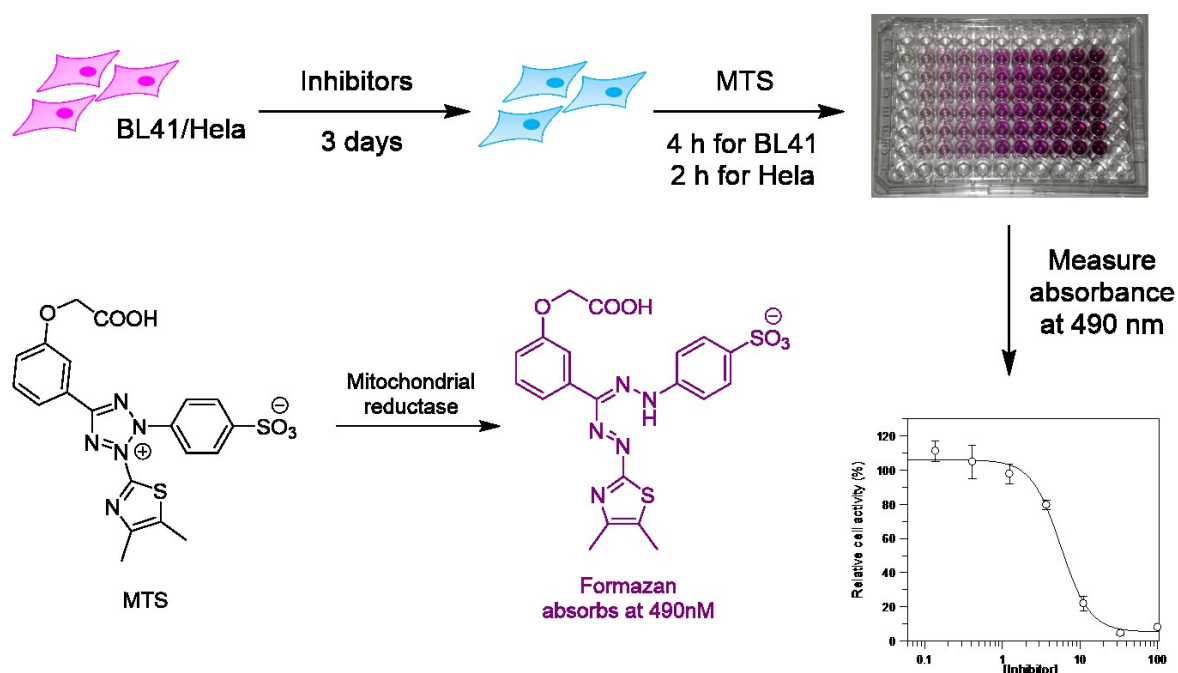


Figure 44 The workflow and mechanism of the MTS assay (inhibitor concentration in μM). The BL41/Hela cells were incubated with inhibitors at different concentrations for three days before being treated with MTS for another four/two hours. The mitochondrial reductase in living cells converts the MTS into aqueous, soluble formazan which absorbs at 490nm. The amount of absorbance at 490nm was measured by the EnVision Xcite Plate Reader and the dose-dependent activity curves is generated using Graft 7 (Erithacus Software Limited).

The EC_{50} s of the inhibitors were measured using the MTS assay, which was developed by previous group members from the CellTiter 96[®] Aqueous Non-Radioactive Cell Proliferation Assay (Figure 44)¹²⁴. MTS is a compound named 3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium inner salt. In a sensitivity screen in cooperation with the Genomics of Drug Sensitivity in Cancer project on two potent NMT inhibitors **IMP-0001031** (**C32**) and **IMP-0001036** against 237 cancer cell lines, BL41 (a cell line established from the tumour tissue of an 8-year-old Caucasian boy with Burkitt lymphoma) was identified as one of the most sensitive cancer cell lines towards NMT inhibition¹²⁵. Therefore, BL41 was chosen to be the main testing cell line in the MTS assay. It is a non-adherent suspension cell line established from the tumour tissue of an 8-year-old Caucasian boy with Burkitt lymphoma.

The BL41 cells were incubated with inhibitors at different concentrations for three days before being treated with MTS for another four hours. The mitochondrial reductase in living cells converts the MTS into aqueous, soluble formazan which absorbs at 490nm. The amount of absorbance at 490nm was measured by the EnVision Xcite Plate Reader and is proportional to the cell activity of BL41 remained. Therefore, dose-dependent activity curves can be generated from the readouts to calculate the EC₅₀s of the inhibitors in test.

HeLa cells are also sensitive to NMT inhibition and can be used in the MTS assay. To compare with the TC₅₀ (the concentration of the NMT inhibitor for 50% reduction in-gel fluorescence) discussed in 2.2.10, the IC₅₀ of **IMP-0001096** was measured on both BL41 and HeLa cells. The assay process is the same except that the HeLa cells were seeded one day before adding inhibitors and the incubation time with MTS was two hours instead of four.

2.2.8 *In vitro* activity analysis

Table 3 *In vitro* activity data of the new azetidinopyrimidines series. The cLogP values were calculated by ChemAxon¹²⁶.

Name	Hs1 IC ₅₀ (μM)	BL41 EC ₅₀ (μM)	EC ₅₀ /IC ₅₀	cLogP
IMP-0001096	0.31 ± 0.20	7.3 ± 2.4	24	2.5
IMP-0001125	9.0 ± 2.2	23 ± 6	2.6	3.9
IMP-0001126	5.1 ± 4.0	17 ± 7	3.3	3.3
IMP-0001127	5.7 ± 4.3	11 ± 7	1.9	4.4
IMP-0001128	1.3 ± 0.6	17 ± 12	13	3.8

IMP-0001129	7.9 ± 4.6	12 ± 8	1.5	4.1
IMP-0001212	2.0 ± 0.7	84 ± 40	42	2.1
IMP-0001459	26 ± 1	30 ± 12	1.2	4.1
IMP-0001460	28 ± 9	>150	-	2.4
IMP-0001461	2.0 ± 0.3	67 ± 24	34	2.0
IMP-0000180	0.061 ± 0.010	6.1 ± 0.7	100	1.5
IMP-0000898	0.031 ± 0.007	1.8 ± 0.6	58	2.1
IMP-0000974	0.0047 ± 0.0088	1.4 ± 0.6	299	2.9
DDD85646	0.013 ± 0.005	0.13 ± 0.06	10	2.0

The IC₅₀s of Hs NMT1 from the CPM assay and the EC₅₀s of BL41 from the MTS assay were listed in Table 3. The EC₅₀/IC₅₀ ratios of all inhibitors were calculated. The average EC₅₀/IC₅₀ ratio of the new inhibitors was 14, which is a much better conversion rate than the 100 of their closest analogue **IMP-0000180** with the old linker, and the 299 of the best analogue **IMP-0000974** with the old linker. This average EC₅₀/IC₅₀ ratio is also comparable to the EC₅₀/IC₅₀ ratio of **DDD85646 (C14)**, a well-validated NMT inhibitor³⁶. This is a promising finding that the new piperidinyl carbamate linker greatly improved the EC₅₀/IC₅₀ ratio. Although the IC₅₀s of the new inhibitors are worse than the original ones, their EC₅₀s are comparable. It is believed that with more optimization and a better understanding of the SARs, inhibitors with the new linker can achieve better EC₅₀s as well as IC₅₀s.

A possible reason for the improvement could be lipophilicity. **IMP-0000180** has a cLogP of 1.5. This high lipophilicity can result in high metabolic turnover and poor absorption. Its analogue, **IMP-0001096**, with the same structure but the new linker has a cLogP of 2.5 with a significant lipophilicity increase of 1.0 brought by the new linker. **IMP-0000898**, the analogue with the old linker but the thieno[3,2-d]pyrimidin shoulder, also showed a lower EC₅₀/IC₅₀ ratio of 58 with a higher cLogP of 2.1. However, when the lipophilicity increase is brought by the isoquinoline and the extra methyl group on the tail region, the EC₅₀/IC₅₀ ratio adversely increased to 299 for **IMP-0000974**. **DDD85646** with a cLogP of 2.0 showed an EC₅₀/IC₅₀ ratio of 10. For **DDD85646**, one can argue that it has a completely different structure, but the lipophilicity differences among the new inhibitors are also not correlated with the EC₅₀/IC₅₀ ratio. Therefore, the cause of the EC₅₀/IC₅₀ ratio improvement by the new linker is complicated and needs further study to understand.

A

B

C

IMP-0001096

D

E

F

IMP-000180

At later stage of the project, the crystal structure of **IMP-0001096** was solved at 2.03 Å after numerous trial and error (The crystallography details are discussed in 3.1). The electron density of **IMP-0001096** was nicely presented in the peptide binding pocket of Hs NMT1 (Figure 45A), which should give a convincing binding mode. The binding mode of **IMP-0001096** was analysed by PyMOL²⁸. The ligand is binding in a similar S-shape as **IMP-0000898** with its

diazepane interacting with the *C*-terminal and its trimethylpyrazole bending upwards to interact with the Ser405 (Figure 45B). The 1*N* on trimethylpyrazole formed a 2.7 Å hydrogen bond with the hydroxy group of Ser405. The 4*N* on the diazepane is interacting with the carboxyl group of the *C*-terminal Gln496 at distances of 2.8 and 3.9 Å. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at the human blood pH 7.4, 86.34% of the **IMP-0001096** has a protonated 4*N* on the diazepane. Therefore, the interaction between the diazepane and the glutamine is ionic binding. There are also π - π stackings interactions of the Phe188 and Phe190 interacting with the trimethylpyrazole and the Tyr296 interacting with the thieno[3,2-*d*]pyrimidin, respectively.

The crystal structure of the most resembled ligand **IMP-0000180** was also solved at 1.82 Å (The crystallography details are discussed in 3.1) and its binding mode was analysed by PyMOL²⁸. The binding modes of the two ligands are surprisingly alike: They present in very similar conformations; The positions of the 1*N*s on trimethylpyrazole and the 4*N*s on the diazepane are almost the same; The interacting residues also pose very similar orientations. The conformation difference lies in the middle of the ligands: the pyrimidine of **IMP-0001096** is pushed more towards the *C*-terminus due to its longer overall length, therefore its π - π stacking with Tyr296 is weakened; While the amide oxygen of **IMP-0000180** is weakly interacting with the phenolic hydroxyl group of Tyr420 by a 4.1 Å hydrogen bond and the amide of Asn451 by a 3.9 Å hydrogen bond as **IMP-0000898** does, the carbamate of **IMP-0001096** flipped to the opposite direction thus losing these interactions.

The five-fold IC₅₀ difference could come from the weakening of the π - π stacking with Tyr296 and two missing hydrogen bonds with Tyr420 and Asn451. Since the two hydrogen bonds of **IMP-0000180** were at 4.1 Å and 3.9 Å, which are longer than the typical hydrogen bonds and

around the maximum limit of possible length, the strength of them were very weak. Therefore, the potency loss was not significant.

2.2.10 Inhibitor validation

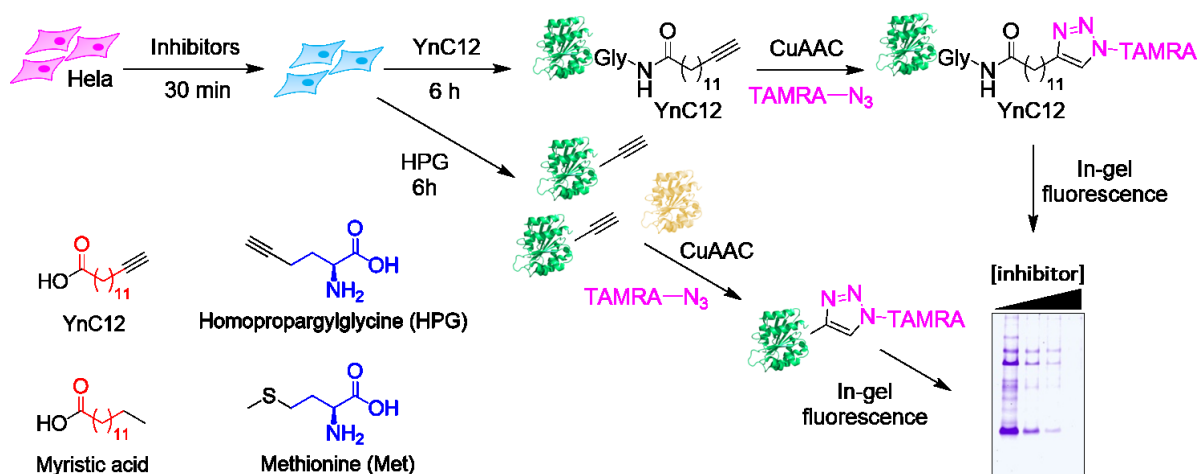


Figure 46 The workflow and mechanism of the metabolic tagging assay.

To validate whether the new inhibitor is working on target, the metabolic tagging assay^{36, 127} was used. YnC12 (tetradec-13-ynoic acid) is an alkyne version of the myristic acid, which can be absorbed and used as the myristic acid by the cells to form Myr-CoA, and then catalysed by NMT to label the newly-myristoylated proteins¹²⁸. HPG (homopropargylglycine) is an alkyne version of methionine, which can be absorbed by the cells and used as the methionine to label the newly-synthesized proteins¹²⁹. In this assay, HeLa cells in two groups were pre-treated with inhibitors in a series of concentrations (100-0.14 μ M) for 30 minutes before YnC12 or HPG were added. After six hours of incubation, the cells were washed and broken down. The protein concentrations of the cell lysates were tested and samples with an equal amount of total proteins were made. The samples were reacted with tetramethylrhodamine (TAMRA) azide under the copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) “click” reaction and analysed by the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The fluorescence on the resulting gels was detected by a Typhoon FLA 9500 Imager (GE Healthcare) and the protein loading was checked by Coomassie staining.

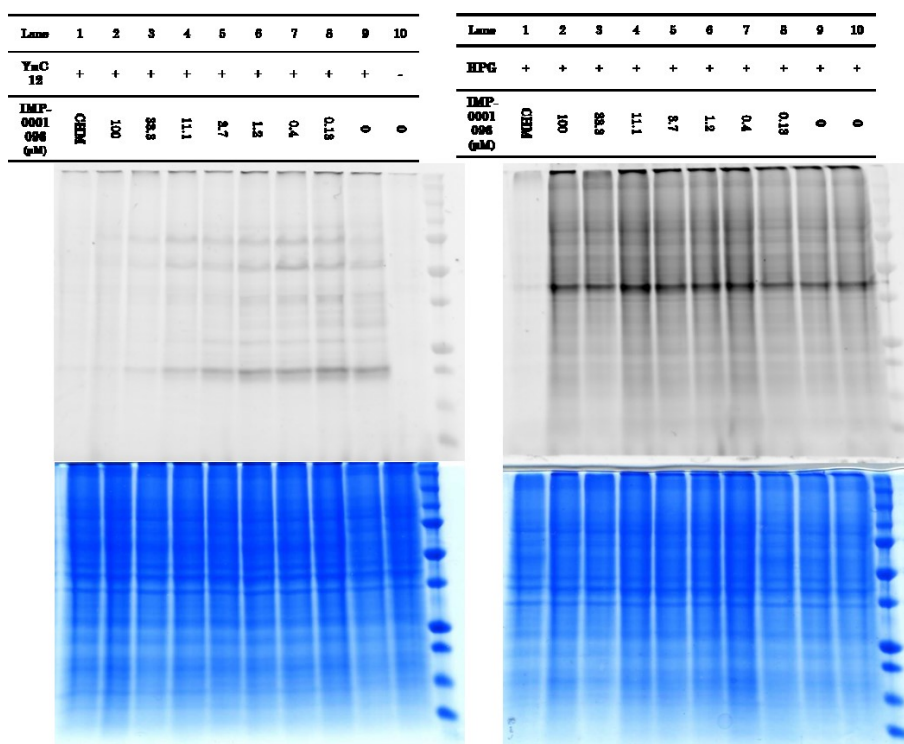


Figure 47 The SDS-PAGE result of the tagging assay in DMEM. The black/white images are generated using Typhoon FLA 9500 Imager (GE Healthcare) and the blue images are the same gels scanned after Coomassie blue staining. Probes (YnC12 (tetradec-13-ynoic acid) for the gel on the left and HPG (homopropargylglycine) for the gel on the right) are added to all lanes but lane 10 (lane 10 of the HPG gel was mistakenly added with HPG, please refer to Figure 48 for the correct result of the lane) Lane 1 is loaded with samples treated with cycloheximide (CHM), which inhibits global protein synthesis. Lane 2 to lane 8 are loaded with the inhibitor, **IMP-0001096**, at different concentrations. Lane 9 and lane 10 are controls, which do not have inhibitor.

The YnC12 tagging result showed an inhibitor-concentration-dependent decrease in labelling of the newly-myristoylated NMT substrates (Figure 47), indicating that **IMP-0001096** inhibits the myristoylation activity of HeLa cells. The HPG tagging assay done in parallel showed similar labelling across all conditions except the control group using cycloheximide (CHM) to inhibit global protein synthesis (Figure 47), proving that **IMP-0001096** did not prevent general protein synthesis in the cell.

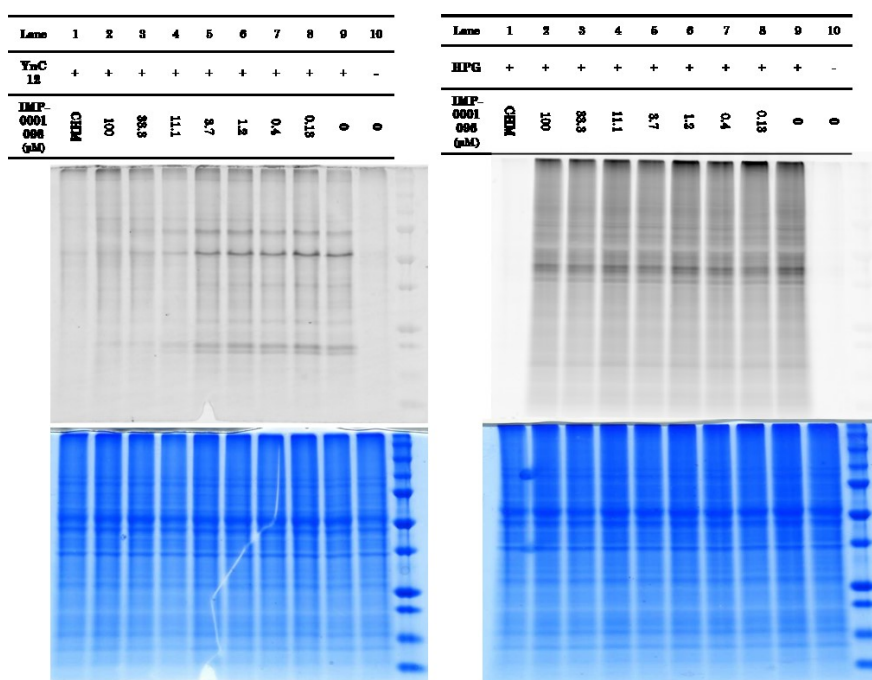


Figure 48 The SDS-PAGE result of the tagging assay in Met (-) DMEM. The black/white images are generated using Typhoon FLA 9500 Imager (GE Healthcare) and the blue images are the same gels scanned after Coomassie blue staining. Probes (YnC12 (tetradec-13-ynoic acid) for the gel on the left and HPG (homopropargylglycine) for the gel on the right) are added to all lanes but lane 10. Lane 1 is loaded with samples treated with cycloheximide (CHM), which inhibits global protein synthesis. Lane 2 to lane 8 are loaded with the inhibitor, **IMP-0001096**, at different concentrations. Lane 9 and lane 10 are controls, which do not have inhibitor.

In case the methionine contained in the Dulbecco's Modified Eagle Medium (DMEM) was interfering with the test results. This assay was repeated using a special formulated DMEM without methionine named Met (-) DMEM (Figure 48) and the result was similar to the ones done in normal DMEM. The labelling of the YnC12 group showed a dose-dependent decrease in intensity while the labelling of the HPG group showed similar intensity across different inhibitor concentrations.

Both results showed that **IMP-0001096** specifically inhibits myristoylation without inhibiting the global protein synthesis, therefore acting on target.

The quantification of fluorescent bands was carried out by ImageJ using the gel analyser function and normalized relative to no inhibition (Lane 9). The apparent in-cell tagging IC_{50} (TC_{50} , the concentration of the NMT inhibitor for 50% reduction in-gel fluorescence)³⁶ was determined using Graphpad Prism 5 by non-linear regression fitting. The mean TC_{50} of three repeats was $2.4 \pm 0.5 \mu M$, which suggested good uptake of **IMP-0001096** in mammalian cells. When comparing its EC_{50} s from the MTS assay ($13 \pm 3 \mu M$ for HeLa and $7.3 \pm 2.4 \mu M$ for BL41), it is roughly comparable, providing additional evidence for the on-target mode of action in cells.

Based on a developing series of *Plasmodium* NMT inhibitors and a crystal structure with Hs NMT1, ten compounds with a new linker were designed, synthesized and evaluated (Figure 49). Among them, the best compound, **IMP-0001096**, showed an IC₅₀ of 0.31 μM. Little potency loss comparing to the original series (IC₅₀ = 61 nM) indicates the flexibility of the linkage part. This compound also shows an excellent EC₅₀/IC₅₀ rate of 24 comparing to the EC₅₀/IC₅₀ rate of 100 of its old-linker analogue **IMP-0000180**. Crystallography study later

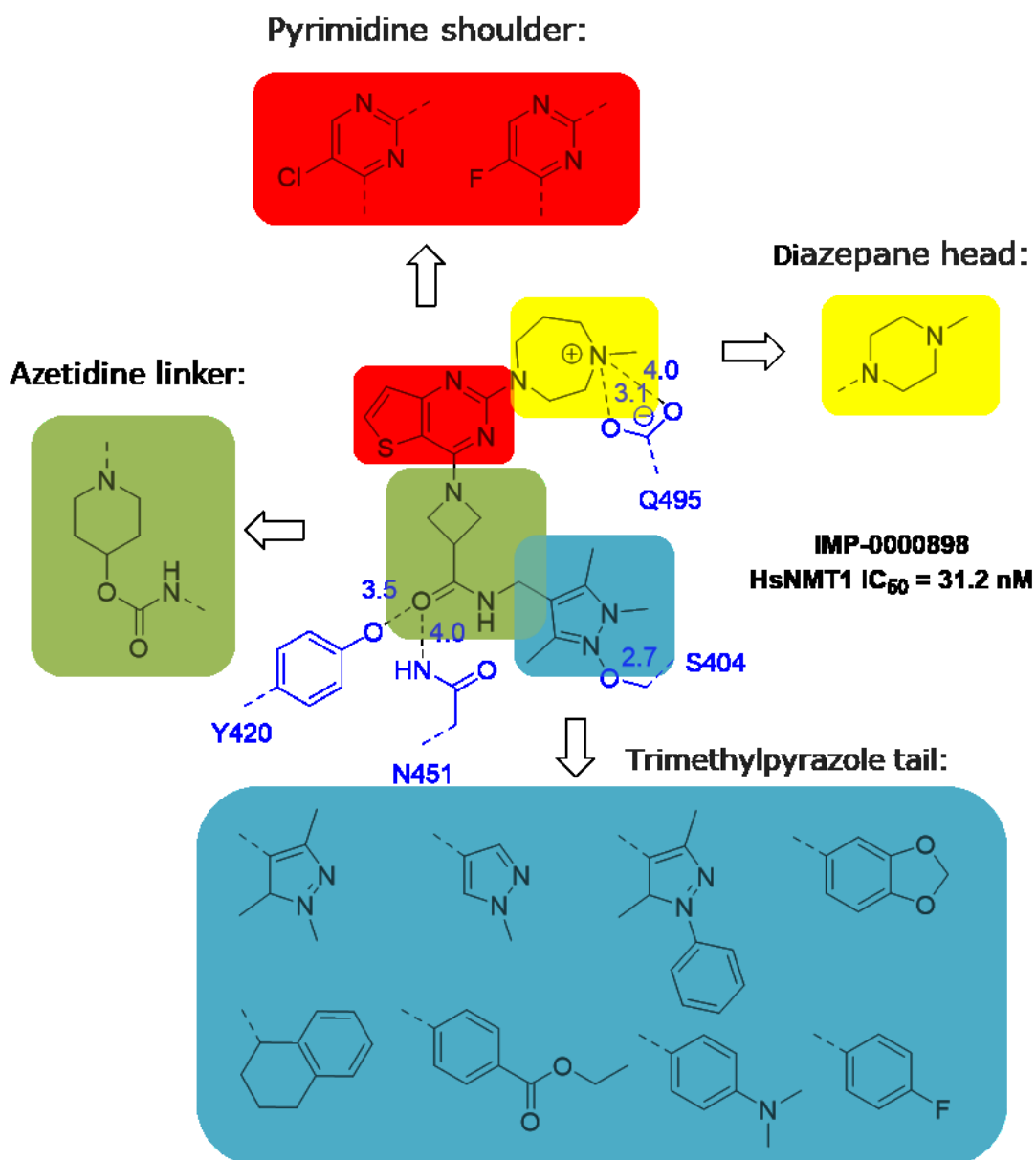


Figure 49 Structure-based molecule development of the new azetidinopyrimidines series.

Based on a developing series of *Plasmodium* NMT inhibitors and a crystal structure with Hs NMT1, ten compounds with a new linker were designed, synthesized and evaluated (Figure 49). Among them, the best compound, **IMP-0001096**, showed an IC₅₀ of 0.31 μM. Little potency loss comparing to the original series (IC₅₀ = 61 nM) indicates the flexibility of the linkage part. This compound also shows an excellent EC₅₀/IC₅₀ rate of 24 comparing to the EC₅₀/IC₅₀ rate of 100 of its old-linker analogue **IMP-0000180**. Crystallography study later

revealed similar binding modes between **IMP-0001096** and **IMP-0000180**, and the potency difference can be structurally explained. YnC12 tagging assay showed inhibitor-concentration-dependent labelling of the myristoylated NMT substrates, providing direct evidence that this compound acts on target. HPG tagging assay done in parallel ruled out the possibility of comprehensive protein-synthesis-inhibition from cell toxicity and confirmed the dose-dependent labelling shown in the YnC12 tagging assay is a result of selective NMT inhibition.

Due to time limitation, this study only explored limited examples of each of the four parts of the ligand. The established synthesis routes can be used to explore more possible substitutions and the crystal structures solved can be used to assist molecule design. Since the binding mode of **IMP-0001096** is very similar to **IMP-0000180**, the SARs established by previous work over 81 compounds and their binding modes (described in 0) can be a valuable guide for further development, especially on the diazepane head and trimethylpyrazole tail as their positions are very alike to the **IMP-0000180** with the old linker.

This study modified the linker of this series for the first time and showed the possibility to address the high EC_{50}/IC_{50} ratio problem while retaining potency by linker change. The new linker showed an average EC_{50}/IC_{50} ratio of 14, comparing to 100 of **IMP-0000180** and 299 of the most potent analogue **IMP-0000974**. According to the binding modes of **IMP-0001096** and **IMP-0000180**, although **IMP-0001096** is missing two hydrogen bonds on the linker and has a weakened π - π stacking on the pyrimidine, its potency is only slightly weaker than **IMP-0000180**'s (five-fold). This finding demonstrated that the two weak hydrogen bonds (3.9 and 4.1 Å) contributed little on the potency, therefore a proper linker change not only could retain IC_{50} and improve EC_{50}/IC_{50} ratio but also improve the potency by closer engagement with Tyr420 and Asn451 by shorter hydrogen bonds at an optimum angle. Further development on this series should emphasis on finding such linkers.

2.3 Nematode NMT

2.3.1 Introduction

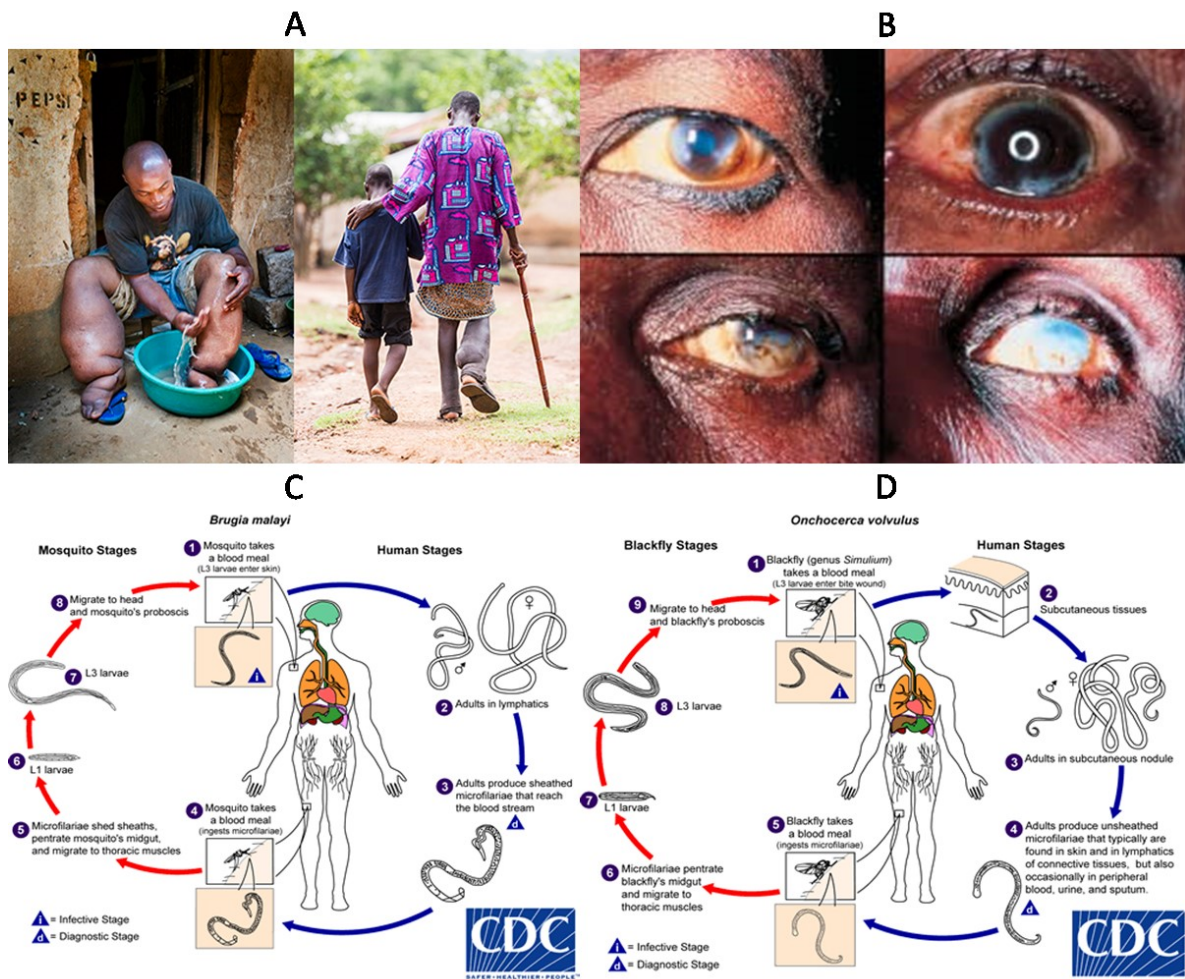


Figure 50 The disabling lymphatic filariasis and onchocerciasis and the life circles of their causing parasites. (A) Lymphatic filariasis causing lymphoedema (tissue swelling) or elephantiasis (skin/tissue thickening) of limbs and hydrocele (scrotal swelling)¹³⁰. Photos retrieved from the WHO website¹³¹; (B) Onchocerciasis causing visual impairment and permanent blindness¹³². Photos were taken by Ian Murdoch & Allen Foster¹³³; (C) Life circle of *Brugia malayi*¹³⁴; (D) Life circle of *Onchocerca volvulus*¹³⁵.

Lymphatic filariasis and onchocerciasis are chronic and disabling neglected tropical diseases caused by filarial nematodes (Figure 50). Lymphatic filariasis is caused by nematodes (*Wuchereria bancrofti*, *Brugia malayi* and *Brugia timori*) that inhabit the lymphatic vessels and lymph nodes of a human host. The infection spreads from person to person by mosquito bites (Figure 50C), threatening 947 million people in 54 countries¹³⁰. Onchocerciasis is an eye

and skin disease caused by *Onchocerca volvulus* and transmitted through bites of blackflies (Figure 50D), threatening 190 million people in 35 countries¹³². The current treatment is the mass drug administration (MDA), giving medicines annually to the entire at-risk population to reduce the larval parasite (microfilariae) and prevent its spread. However, these medicines have a limited effect on adult parasites and MDA needs to be conducted annually for 4-6 years (lymphatic filariasis) or 10-15 years (onchocerciasis) with effective coverage of the total population at risk to interrupt the transmission cycle. The limitation of the current therapy and the occurring drug resistance¹³⁶ call for novel drugs that are effective, safe and affordable.

In 2014, Carlow *et al.* validated NMT as a nematode drug target for the first time by performing genetic studies in *Caenorhabditis elegans* (*C. elegans*, Ce), and chemical inhibition studies in *C. elegans* and *Brugia malayi* (*B. malayi*, Bm)¹⁰⁵. NMT inhibitors **DDD85646 (C14)** and **DDD100870**, originally developed against *Trypanosoma brucei* (*T. brucei*, Tb)^{42, 137}, were tested by a radioisotopic labelling enzyme assay against nematode NMTs and showed potent enzyme activity with IC₅₀ values of 2.5-10 nM¹⁰⁵. In *in vivo* tests against *C. elegans*, **DDD100870** completely inhibited the nematode's food consumption at 100 µM, and both compounds interfered the reproduction of *C. elegans*, resulting in fewer and smaller progeny¹⁰⁵. In *in vivo* tests against *B. malayi*, the two inhibitors were active against microfilariae at 12.5 mM and adult worms at 50 mM, affecting their mobility and reproduction¹⁰⁵. RNA interference and gene deletion in *C. elegans* caused severe developmental defects and demonstrated that NMT is essential for nematode viability¹⁰⁵.

Encouraged by the above study, this study aims to screen inhibitors of plasmodium or malaria NMTs developed in our group, using both the CPM enzyme assay and the *in vivo* assay, to identify new nematode NMT inhibitors.

2.3.2 Sequence analysis of NMT proteins

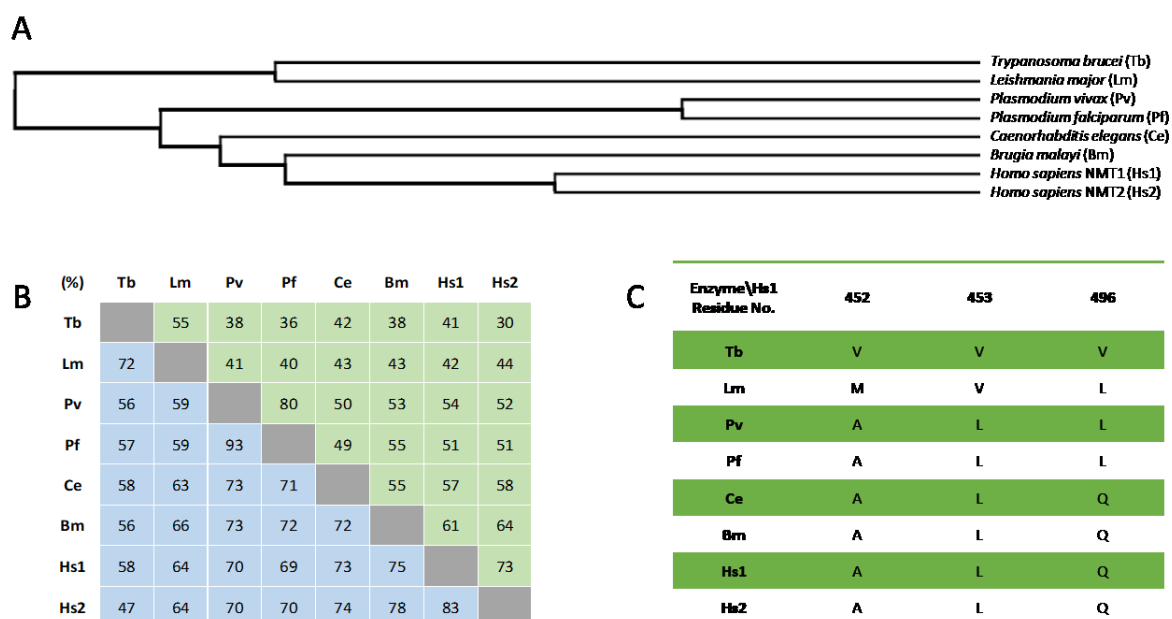


Figure 51 Sequence analysis of relevant NMTs. Protein sequences were obtained from UniprotKB (Sequence entry numbers are shown in brackets, the full sequences and alignment results can be found in 5.2.2)¹³⁸:

Trypanosoma brucei (D0A003), *Leishmania major* (Q4Q5S8), *Plasmodium vivax* (A5K1A2), *Plasmodium falciparum* (Q8ILW6), *Caenorhabditis elegans* (P46548), *Brugia malayi* (A0A0H5SBS7), *Homo sapiens NMT1* (P30419), *Homo sapiens NMT2* (O60551). (A) Rooted phylogenetic tree analysis generated by Uniplot¹³⁹; (B)

Comparison of protein identity and similarity. Percentages of sequence identity (on green background) and percentages of sequence similarity (on blue background) as defined by BLASTP; (C) Residue differences in the inhibitor binding pocket. Residues within 4.5 Å of the binding inhibitor in the co-crystal structure of DDD85646 and Hs NMT1 (PDB code: 3IWE) were analysed and the non-conserved ones listed.

An alignment of all the parasitic and human NMTs of interest to this study was analysed before the drug screen to examine their homology (Figure 51). The protein sequences of eight NMTs were aligned by Uniplot and a rooted phylogenetic tree was generated (Figure 51A)¹³⁸. The two human NMTs are more closely related to the two nematode NMTs, Ce and Bm NMTs, than to the other parasitic NMTs. The two plasmodium NMTs ranked in the middle, and the NMTs of Tb and Lm showed the lowest homology to the nematode NMTs. These findings indicated that a target repurposing method¹⁴⁰ to identify new nematode NMT inhibitors from inhibitors against parasite NMTs exhibits a potential for success, since the inhibitors of the

furthest related Tb NMT (**DDD85646** and **DDD100870**) were proved active against Ce and Bm NMTs¹⁰⁵.

Protein identities and similarities were calculated using BLASTP (Figure 51B)¹³⁹. Tb NMT shows 42%/38% sequence identity and 58%/56% sequence similarity with Ce and Bm NMTs, respectively. The rest of NMTs share comparable or more sequences with the nematode NMTs than Tb NMT, which also endorses the target repurposing method.

Co-crystal structure of **DDD85646** and Hs NMT1 (PDB code: 3IWE) was inspected in PyMOL to identify residues which interact with the inhibitor. 22 protein residues and the co-factor myristoyl-CoA were found within 4.5 Å distance of the inhibitor. The corresponding residues were compared with other NMT isoforms and three non-homologous residues were found among the eight analysed enzymes (Figure 51C). The alignment results agree with the phylogenetic tree and protein homology analysis. Ce, Bm, Hs1 and Hs2 NMTs share identical sequences in the drug binding pocket. Therefore, their binding sites should be identical. Due to the lack of the crystal structures of nematode NMTs, structures of Hs NMT1 were used in this study to explain the SARs of the inhibitors. The two plasmodium NMTs have the same sequences with one residue difference from the nematode NMTs, while Lm and Tb NMTs differ in all three residues. The side chain of residue 452 points away from the inhibitor, residue 453 affects the size of the binding pocket, and the C-terminal residue 410 interacts through a carboxyl group, which also affect the overall size of the pocket. These differences can result in the selectivity among NMT isoforms.

2.3.3 Kinetic study

Table 4 Kinetic parameters of nematode NMTs (the sequences and origins of the four peptides used can be found in Table 8 of 5.2.3)

Bm

Peptide	8AA		15AA		PfAA		CaAA	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Standard Error
V _{max} (Signal/min)	6.5 × 10 ⁵	1.0 × 10 ⁴	7.1 × 10 ⁵	4.0 × 10 ⁴	7.1 × 10 ⁵	9.6 × 10 ³	4.6 × 10 ⁵	5.9 × 10 ³
K _m (μM)	8.3	0.54	11	2.4	1.7	0.27	11	0.47

Ce

Peptide	8AA		15AA		PfAA		CaAA	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Standard Error
V _{max} (Signal/min)	7.9 × 10 ⁵	2.1 × 10 ⁴	7.6 × 10 ⁵	1.6 × 10 ⁴	4.8 × 10 ⁵	7.2 × 10 ⁴	2.5 × 10 ⁵	6.1 × 10 ³
K _m (μM)	14	1.1	8.9	1.1	0.67	1.4	3.2	0.34

The nematode NMT activities in the original paper¹⁰⁵ were measured by detection of a radio-isotopic label in the myristoylated product peptide. In this study, the CPM assay (described in 2.1.2), a robust fluorogenic assay avoiding radioactive reagents¹¹⁶, was used to determine enzyme potency. To evaluate the enzyme activities of the nematode NMTs, a kinetic analysis of Bm and Ce NMTs (enzymes kindly provided by the Carlow group at New England Biolabs) was performed by the CPM continuous assay using four synthesized peptide substrates, Hs p60-Src₍₂₋₉₎, Hs p60-Src₍₂₋₁₇₎, Pf Arf₍₂₋₁₇₎ and Ca Arf₍₂₋₁₇₎. 11 peptide concentrations from 32 to 0.03125 μM were tested together with the negative control without peptide. 30 μM Myr-CoA was added to 300 ng/ml of enzymes to saturate the co-factor pocket. The fluorescence was measured every minute for the initial 45 minutes. The results were fitted to the linear equation to obtain the initial rate constants and then the Michaelis-Menten equation for K_m (Michaelis Constant: concentration of substrate at which the reaction rate is half the maximum rate) and V_{max} (maximum velocity of the enzyme catalysed reaction) by OriginPro (Table 4). Although the values were not expected to be the same as the V_{max} and K_m values measured in

the radioactive assay using a peptide derived from ABL-1 (K_m of Bm = 4.4 μ M and K_m of Ce =12.2 μ M), the data obtained are of the expected magnitude.

2.3.4 Enzyme activity

Table 5 IC_{50} s of inhibitors against the nematode NMTs

	(nM)	IMP- 0001002	IMP- 0001031	IMP- 0000320	IMP- 0000964	IMP- 0000975	IMP- 0000623	IMP- 0000653	DDD85646
Ce	IC_{50}	1.3	<1	23.3	21.5	24.3	3424	~1	<1
	SE	0.3		6.2	4.3	3.4	1213		
Bm	IC_{50}	2.4	<1	15.0	12.6	33.7	4932	4.6	<1
	SE	1.9		7.9	4.7	6.9	738	3.7	

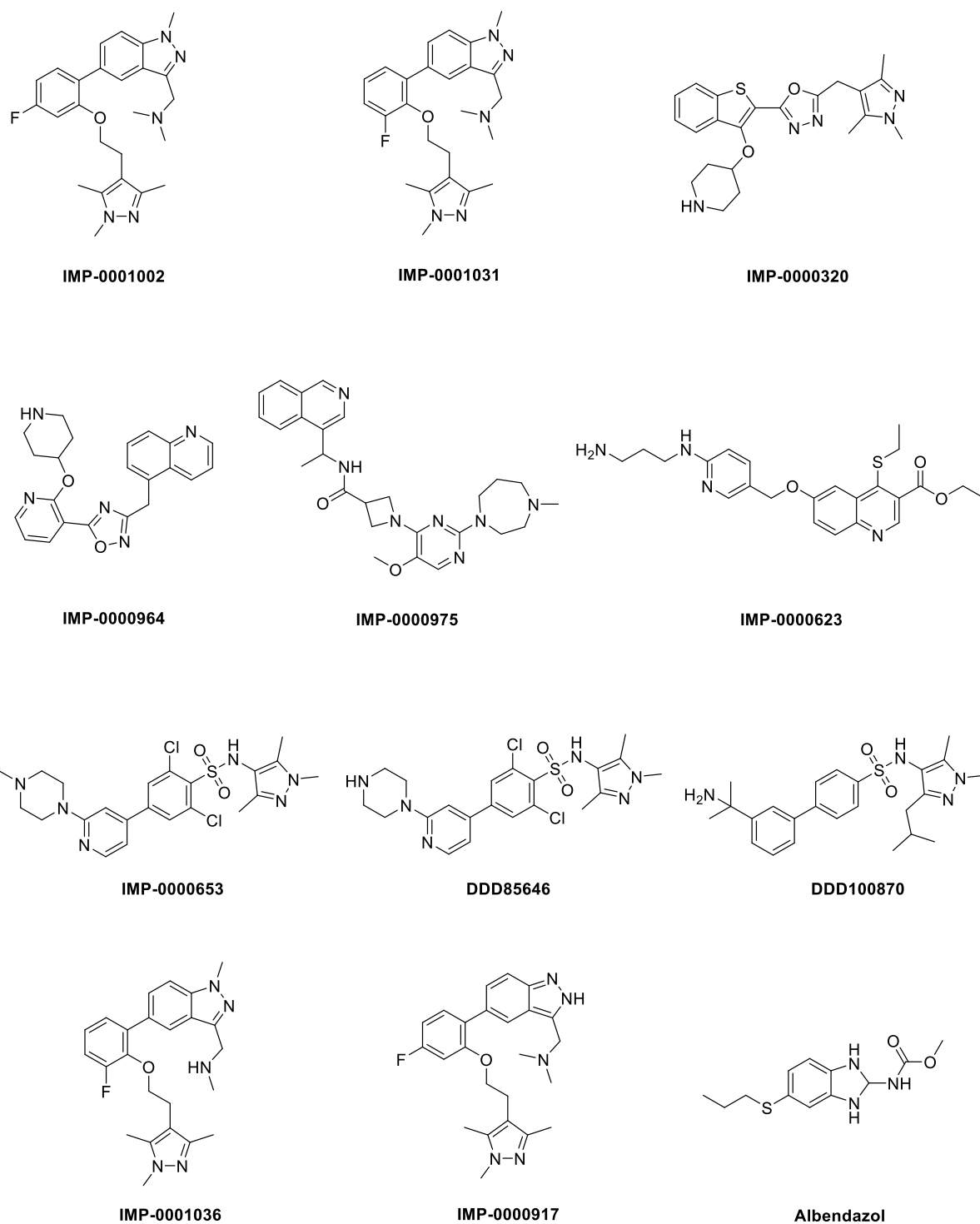


Figure 52 The structures of NMT inhibitors tested in the enzyme activity assay and the *in vivo* assay. **IMP-0001002, IMP-0001031, IMP-0000320, IMP-0000964, IMP-0000975, IMP-0000623, IMP-0000653, DDD85646** were tested against the nematode NMTs. **IMP-0001002, IMP-0001031, IMP-0001036, IMP-0000320, IMP-0000917, IMP-0000964, IMP-0000975** and Albendazol were tested against the *C. elegans* larvae.

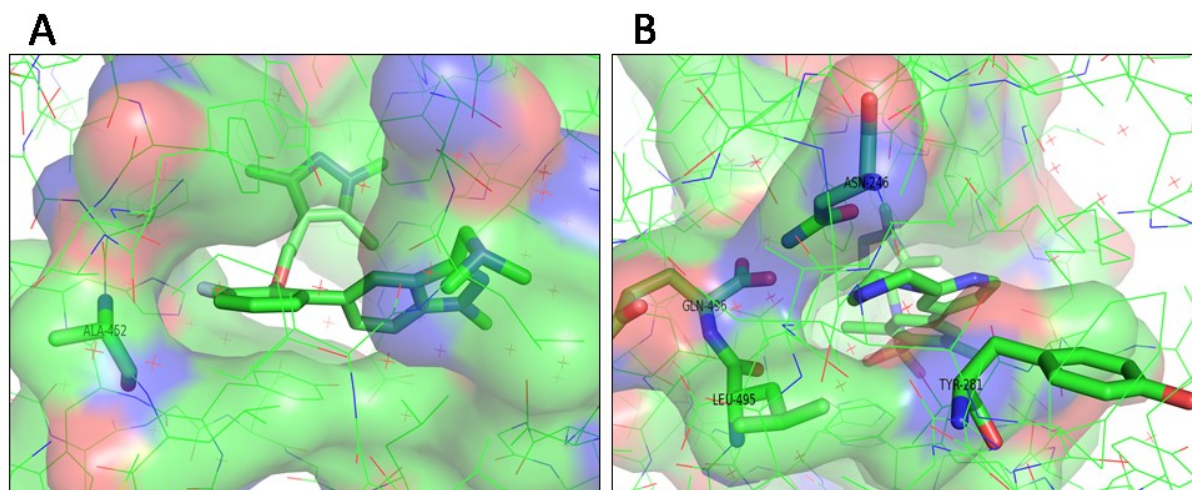


Figure 53 Inhibitor binding in Hs NMT1. Images created by PyMOL²⁸. The binding sites within 5 Å of the inhibitor were shown in surface with 50% transparency. (A) The binding mode of **IMP-0001031** (PDB code: 5O6J); (B) The binding mode of **DDD85646** (PDB code: 4C2Z).

Eight NMT inhibitors with six different scaffolds (Figure 52), synthesized by *Dr* Andy Bell, *Dr* Mark Rackham and *Dr* Zhiyong Yu, were selected and their IC₅₀s were tested against Bm and Ce NMTs by the CPM assay described in 2.1.2 (Table 5). All eight inhibitors are active against both nematode NMTs with similar IC₅₀ values, which accords with the same protein sequence of their binding pockets. **IMP-0001002**, **IMP-0001031** (C32), **IMP-0000653** and **DDD85646** (C14) are the most potent compounds with single-digit nanomolar or sub-nanomolar IC₅₀s. **IMP-0000320**, **IMP-0000964** and **IMP-0000975** followed with IC₅₀s of 12.6-33.7 nM. **IMP-0000623** is the least potent with an IC₅₀ of 3.4 μM for Ce NMT and an IC₅₀ of 4.9 μM for Bm NMT. **IMP-0001002** and **IMP-0001031** are very similar in structure and the change from ortho- to meta-fluoro decreased its potency against both Bm and Ce NMTs. According to **IMP-0001031**'s co-crystal structure with Hs NMT1 (Figure 53A), the meta-fluoro should point to the backbone of residue 452 and the tight space could cause steric issues or shift the whole inhibitor away from its optimal position, resulting in the potency loss. **IMP-0000653** and **DDD85646** are also similar in structure. The additional methyl group on the piperazine decreased the potency of **IMP-0000653** slightly against both Bm and Ce NMTs. According to **DDD85646**'s co-crystal structure with Hs NMT1 (Figure 52B), the methyl group

should be pointing to a small pocket surrounded by residues 246, 281, 495 and 496. Therefore, no steric clash should exist. However, the electrostatic interaction of the piperazine based amine with residues 246 and 496 could be altered when changing from a secondary to a tertiary amine.

2.3.5 *In vivo* activity

The *in vivo* activities of **IMP-0001002**, **IMP-0001031**, **IMP-0000320**, **IMP-0000964** and **IMP-0000975** were tested by Zhiru Li in the Carlow group. **IMP-0001036** and **IMP-0000917 (C31)** which have the same scaffold of **IMP-0001002** and **IMP-0001031** were also tested, as they might have similar high potencies (Figure 52). **IMP-0000623** was not included due to poor enzyme activity. Since **DDD85646** was already reported in the publication¹⁰⁵, **DDD85646** and its derivative **IMP-0000653** were not included.

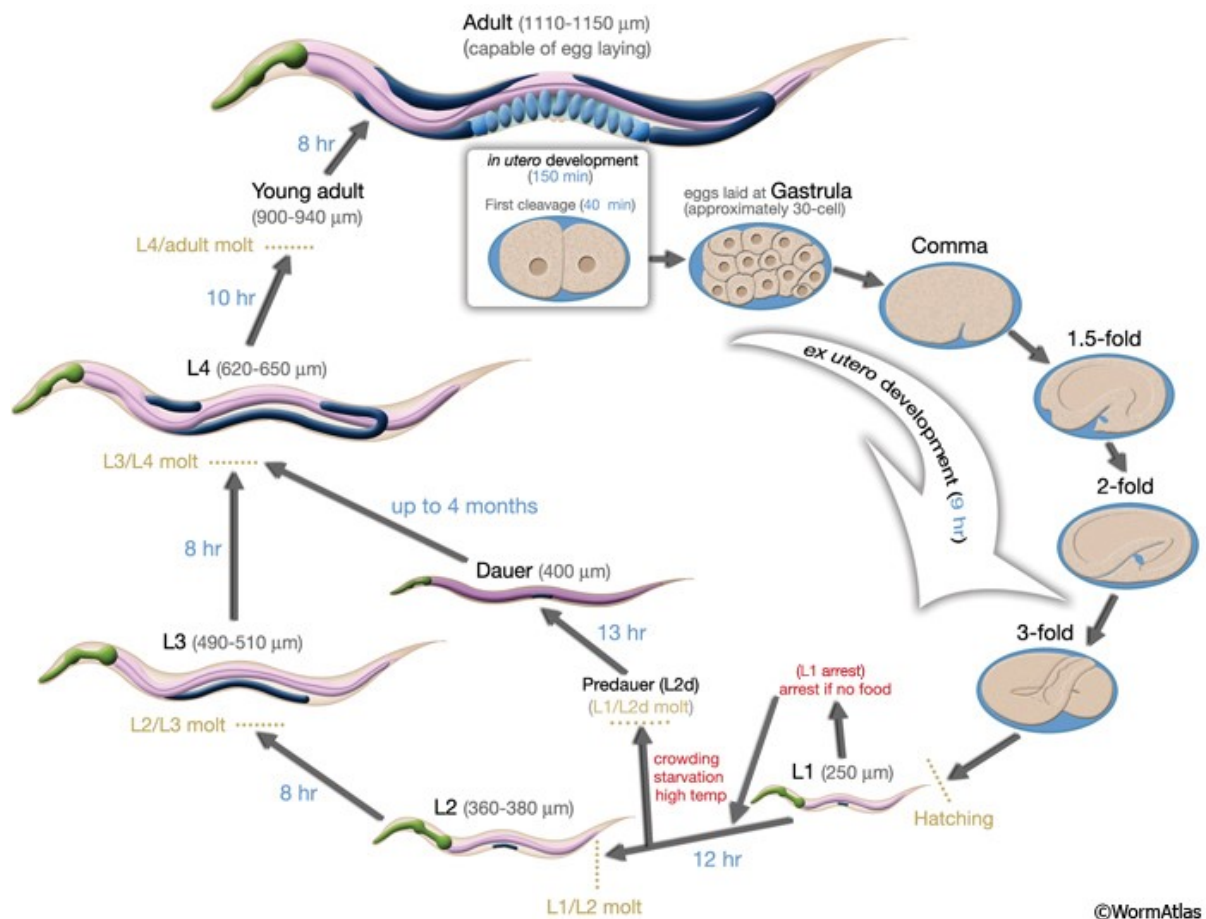


Figure 54 The life cycle of *C. elegans* at 22°C (Image reproduced from WormAtlas website¹⁴¹). Numbers in blue indicate the length of time the animal spends at a certain stage. 0 min is fertilization. The length of the animal at each stage is marked in grey next to the stage name in micrometres (μm). The adult worm can live for 2-3 weeks.

C. elegans was used as a nematode model in the *in vivo* assay. Like other nematodes, the life cycle of *C. elegans* is comprised of the embryonic stage, four larval stages (L1-L4) and adulthood¹⁴². The end of each larval stage is marked with a molt, during which the old cuticle is replaced by a newly-synthesized, stage-specific cuticle (Figure 54)¹⁴².

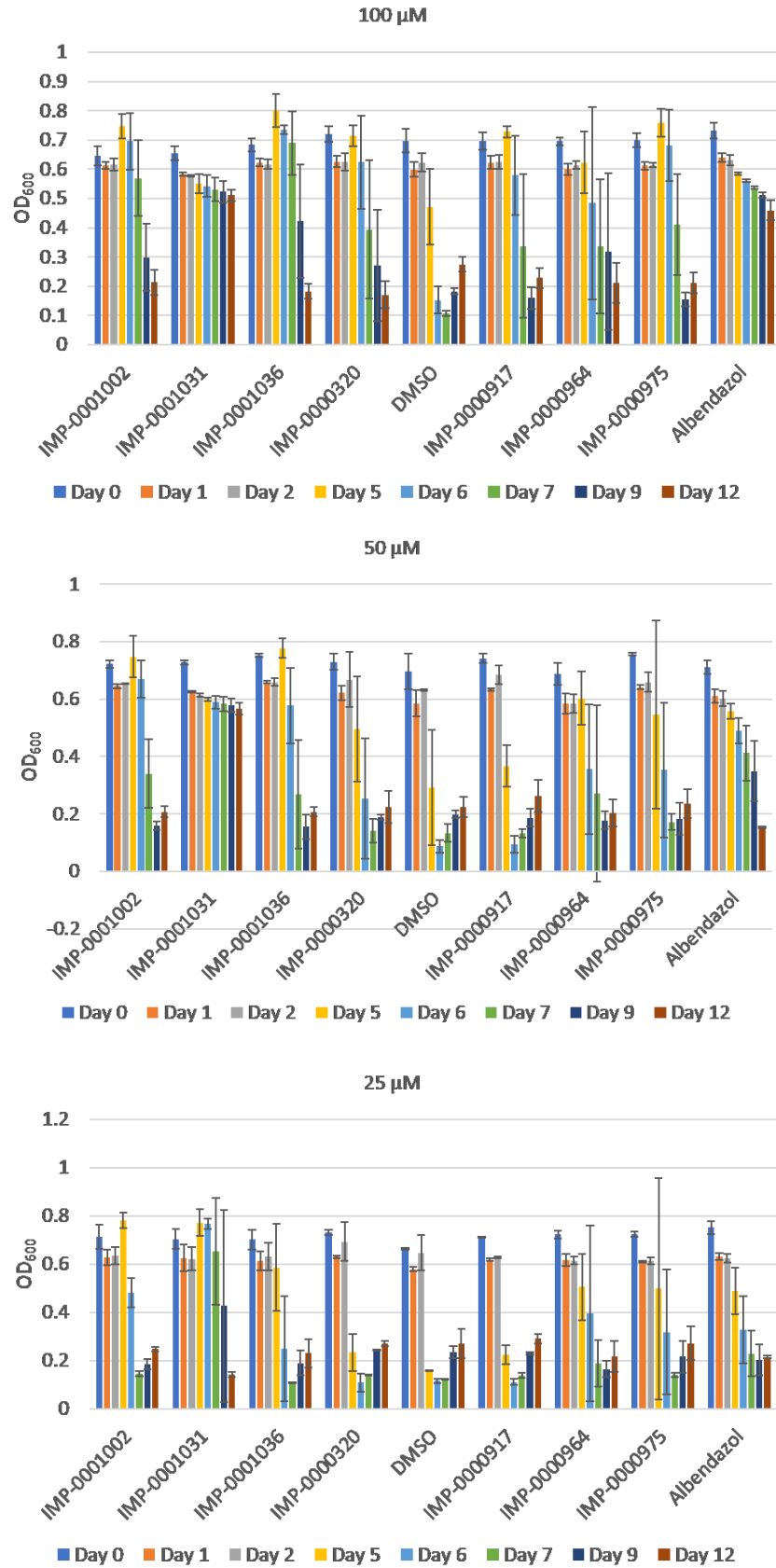


Figure 55 Food consumption measured by OD₆₀₀ of the L4 larvae treated with inhibitors (Images reproduced with kind permission of Zhiru Li in the Carlow group). DMSO and Albendazol were used as negative and positive controls. The OD₆₀₀ values were used as indicators of the food left and were measured at day 0, 1, 2, 5, 6, 7, 9 and 12.

One L4 larva was distributed into each well of a 96-well plate that contains HB101 *E. coli* bacteria as food source. L4 larvae were treated with 100 μ M, 50 μ M and 25 μ M of compounds in duplicate. Albendazol, the drug widely-used in MDA campaigns, was administrated as the positive control and DMSO was used as the negative vehicle control. The experiment lasted for 12 days and no fresh compounds were added during the experiment. Food consumption was measured by OD₆₀₀ as an indicator of worm health and viability. Each well was examined under the microscope to determine the development of F1 progeny and viability¹⁰⁵.

The OD₆₀₀ is not only affected by the remaining food source, HB101 *E. coli* bacteria, but also *C. elegans*. The bacteria were previously frozen and sterilized therefore they do not grow during the assay. In the DMSO control, the worms usually consume all the bacteria at day 6 and then live without food before they die at around day 16. The increase of OD₆₀₀ after day 6 is due to the hatching of progenies. Effective inhibitors can delay the food consumption and reproduction so that the lowest OD₆₀₀s occur later or are even not reached at the end of the assay at day 12. High concentrations of DMSO also have modest toxicity towards the worms therefore at 100 μ M the lowest OD₆₀₀ occurs at day 7. The potencies of inhibitors are ranked by comparing the OD₆₀₀s of the treated worms to the OD₆₀₀s of both controls.

At 100 μ M, all seven inhibitors showed inhibition by reducing the food consumption of the L4 larvae in comparison to the DMSO control. Interestingly, **IMP-0001031** showed higher activity than the positive control, Albendazol, at all three concentrations and its inhibition continued for the whole length of the assay at 100 μ M and 50 μ M. Its analogues **IMP-0001002** and **IMP-0001036**, and compounds **IMP-0000964** and **IMP-0000975** also showed inhibitory effect at all three concentrations, but their inhibition decreased as the assay proceeded. **IMP-0000320** and **IMP-0000917** only showed inhibition at 100 μ M. At 50 μ M and 25 μ M, their effects were not significant comparing to the DMSO control. Compared with the published

data of **DDD85646** and **DDD100870**¹⁰⁵, **IMP-0001031** is more potent. **IMP-0001002**, **IMP-0001036**, **IMP-0000964** and **IMP-0000975** showed comparable potency with **DDD100870**. The activities of **IMP-0000320** and **IMP-0000917** are similar to the potency of **DDD85646**.

Under the microscope, the worms in the DMSO control developed into adult worms and produced eggs that hatched to F1 progenies on day 2. On day 6, F2 progenies appeared. All inhibitor-treated worms showed delayed development. At 100 μ M and 50 μ M, **IMP-0001031** treated L4 larvae developed to adult worms but produced fewer eggs that didn't hatch. On the other conditions, the worms eventually produced F1 and F2 progenies.

The *in vivo* activity of **IMP-0001031** correlated with its enzyme activities. It is the most potent among the tested compounds. Although **IMP-0000964** and **IMP-0000975** had 17- and 19-fold higher IC₅₀ values than **IMP-0001002**, respectively, they showed similar *in vivo* activity against L4 *C. elegans*. **IMP-0000320** had similar IC₅₀ values as **IMP-0000964** and **IMP-0000975**, but the *in vivo* activity was lower. **DDD85646** had a similar IC₅₀ as **IMP-0001031**, but much lower *in vivo* activity. The deviations between enzyme activities and *in vivo* activities could be a result of the differences in compounds' uptake efficiencies.

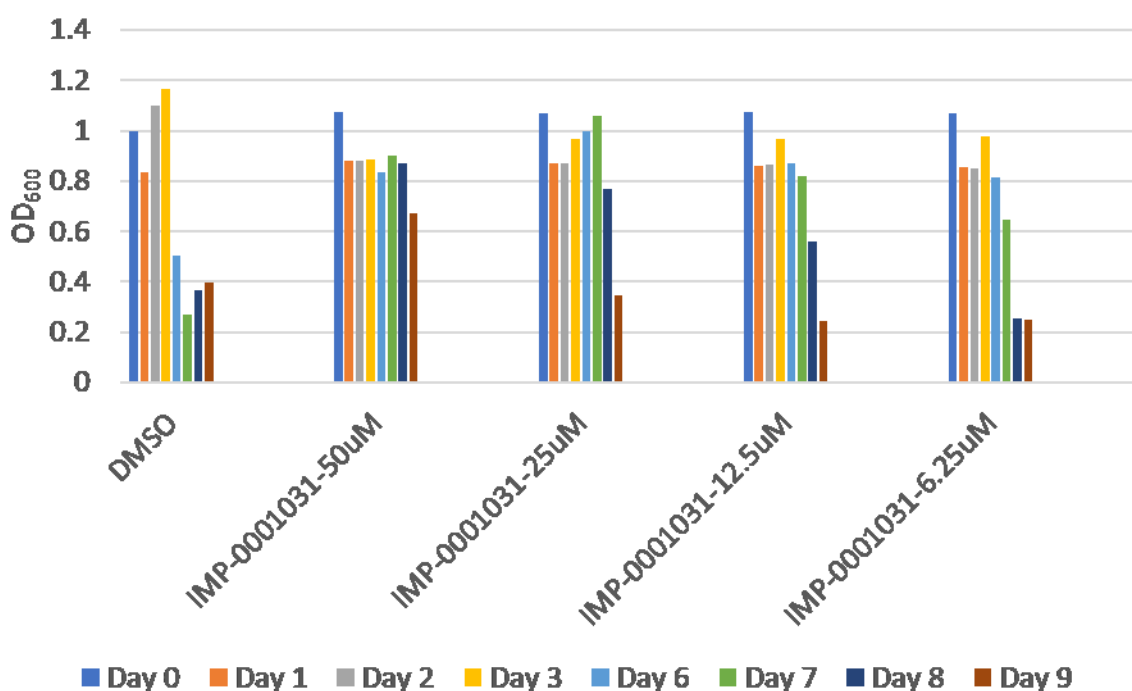


Figure 56 Food consumption measured by OD₆₀₀ of the L1 larvae treated with IMP-0001031 (Images reproduced with kind permission of Zhiru Li in the Carlow group). DMSO was used as the negative control. The OD₆₀₀ values were used as indicators of the food left and were measured at day 0, 1, 2, 3, 6, 7, 8 and 9.

The good potency of **IMP-0001031** against L4 larvae encouraged its test against L1 larvae. Lower concentrations, 50 μM, 25 μM, 12.5 μM and 6.25 μM, were used. Twenty L1 larvae were used in each well and the experiment lasted for 9 days¹⁰⁵.

The food consumption of L1 larvae showed a dose-dependent decrease as the concentration of IMP-0001031 increases (Figure 56). At 50 μM, **IMP-0001031** can inhibit the food intake of L1 larvae and the delay of its development was observed under the microscope.

2.3.6 Conclusion and outlook

Filarial nematodes are the parasites causing the neglected tropical diseases, lymphatic filariasis and onchocerciasis. The limitations of existing treatments highlight the need to develop new drugs.

This study analysed the homology across eight NMTs of parasite and human. The high protein identity and similarity of these organisms enabled the target repurposing approach to explore nematode NMT inhibitors from the existing inhibitors of other NMTs. Human and nematode NMTs are closely related with identical protein sequences of the inhibitor binding pockets. This presents a difficulty to achieve selectivity over human NMTs for nematode NMT inhibitors but enabled the utilisation of the available crystal structures of human NMTs to study the drug pocket of the nematode NMTs.

Eight plasmodium or malaria NMT inhibitors were tested in the enzyme activity study of Ce and Bm NMTs and all showed inhibition. **IMP-0001031**, **IMP-0001002**, **DDD85646** and **IMP-0000653** showed single-digit or sub-nanomolar inhibition. The SARs of two scaffolds were analysed by investigation of crystal structures of human NMTs.

Seven compounds were tested in the *in vivo* activity assays and all showed food consumption inhibition and development disruption effect on Ce L4 larvae at 100 μ M. **IMP-0001031** showed higher potency than the positive control Albendazol. At 50 μ M, it inhibited the food intake of both Ce L1 and L4 larva, delayed development and prevented reproduction.

The good enzyme and *in vivo* activity of **IMP-0001031** makes it a good tool for functional studies of the nematode NMTs. The compound could also be developed into a potential filarial parasite drug if the human selectivity can be addressed.

3 Crystallography

3.1 Workflow

3.1.1 Protein expression

Hs1_Full1	1	MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGLSPANDTGAKKKKKK
Hs1_DeN	1	-----
Hs1_Full1	61	QKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQKATETFSVGGQGPAAKTMEEASKRSYQFW
Hs1_DeN	1	-----MGSSH-----HHHHSSGIEV--LFC--CPHMEEASKRSYQFW
Hs1_Full1	121	DTQPVPKLGEVVNTHGVPDPKDNIRQEPYTLPPQGFTWDALDLGDRGVLKELYTLLNENY
Hs1_DeN	35	DTQPVPKLGEVVNTHGVPDPKDNIRQEPYTLPPQGFTWDALDLGDRGVLKELYTLLNENY
Hs1_Full1	181	VEDDDNMFRFDYSPEFLLWALRPPGWLPQWHCGVRVSSRKLVGFI SAIPANIHIYDTEK
Hs1_DeN	95	VEDDDNMFRFDYSPEFLLWALRPPGWLPQWHCGVRVSSRKLVGFI SAIPANIHIYDTEK
Hs1_Full1	241	KMVEINFLCVHKKLRSKRVAPVLIREITRRVHLEGIFQAVYTAGVVLKPKVGTCTRYWHR
Hs1_DeN	155	KMVEINFLCVHKKLRSKRVAPVLIREITRRVHLEGIFQAVYTAGVVLKPKVGTCTRYWHR
Hs1_Full1	301	LNPRKLIEVKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQLLTRYLKQF
Hs1_DeN	215	LNPRKLIEVKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQLLTRYLKQF
Hs1_Full1	361	HLTPVMSQEEVEHWFYPQENIIDTFVVENANGEVTDFLSFYTLPPSTIMNHPTHKSLKAA
Hs1_DeN	275	HLTPVMSQEEVEHWFYPQENIIDTFVVENANGEVTDFLSFYTLPPSTIMNHPTHKSLKAA
Hs1_Full1	421	SFYNVHTQTPLLDLMSDALVLAKMKGFDVFNALDLMENKTFLEKLFKFGIGDGNLQYYLYN
Hs1_DeN	335	SFYNVHTQTPLLDLMSDALVLAKMKGFDVFNALDLMENKTFLEKLFKFGIGDGNLQYYLYN
Hs1_Full1	481	WKCPSMGAEKVGLVLQ
Hs1_DeN	395	WKCPSMGAEKVGLVLQ

Figure 57 Protein sequence alignment of Hs NMT1 (Hs 1_Full) and the modified version without *N*-terminus (Hs1_DeN). The *N*-terminal deletions are replaced by a histidine-tag and protease HRV 3C cleavage site. The full length sequence was obtained from UniplotKB¹³⁸ (Sequence entry numbers: P30419). Sequences were aligned by Uniplot and displayed by BoxShade¹⁴³.

The recombinant plasmid was obtained from Dr James A. Brannigan, University of York, which was made by cloning the insert encoding residues 109-496 of *Homo sapiens* NMT 1 isoform (Hs NMT 1) into a modified pET28 vector, which incorporates the protein product

with an *N*-terminal hexahistidine tag to facilitate purification (Figure 57). The *N*-terminus is not involved in substrate binding and is usually disordered in NMT crystal structures due to its flexibility⁴¹⁻⁴². Therefore, the modified version of Hs NMT1 without the *N*-terminus was used in order to obtain crystals with better quality. The recombinant plasmid was transformed into *E. coli* BL21(DE3) pRareS, a host strain providing additional tRNA genes to enhance the expression of human genes in bacteria¹¹⁶, and expressed by autoinduction¹⁴⁴ with overnight growth at 18 °C.

3.1.2 Protein purification

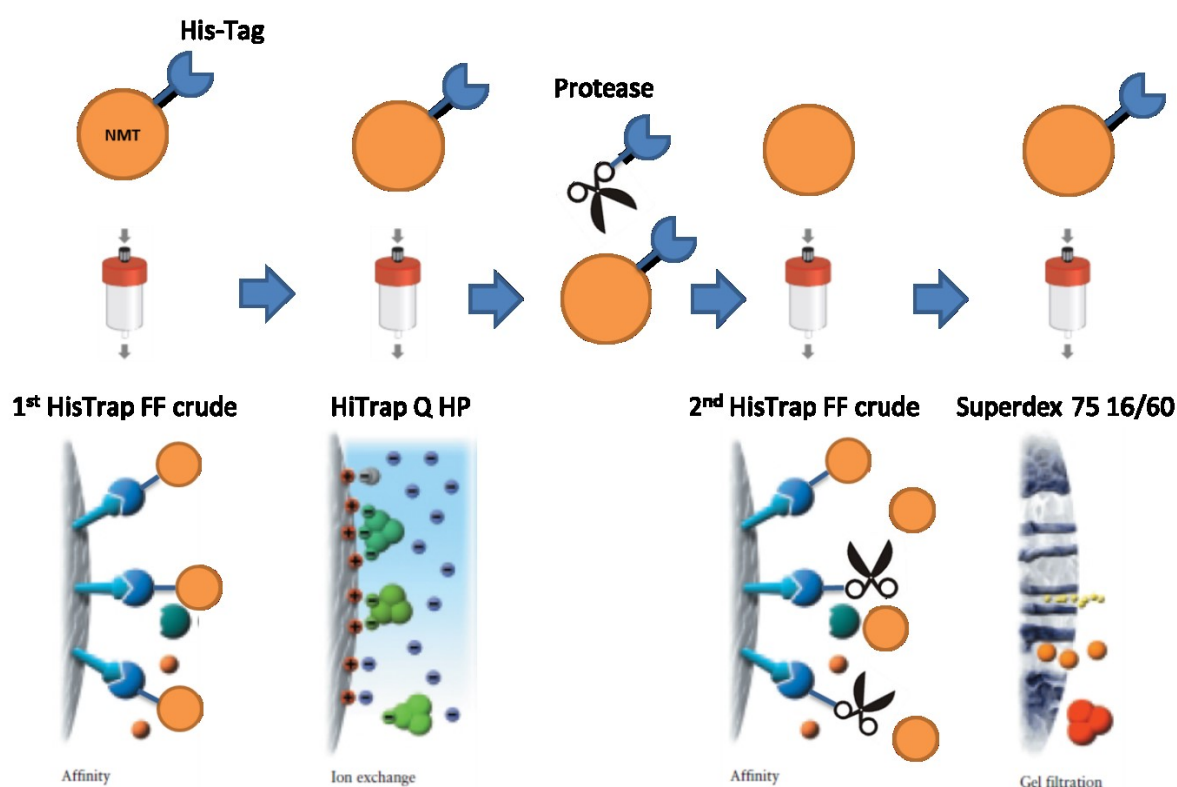


Figure 58 Purification workflow of Hs NMT1 (Images adapted from Handbooks from GE healthcare: Gel filtration principles and methods¹⁴⁵).

The purification follows the steps in the published method¹¹⁶, where the cells were sonicated in the lysis buffer and purified by Ni-NTA, ion exchange and gel filtration. This series of chromatography produces pure Hs NMT 1 protein with the His-tag ready for activity assays. To use the protein for crystallography, the His-tag needs to be removed³⁶. Therefore, a His tag

cleavage step by HRV 3C protease (kindly provided by *Dr James A. Brannigan*) and a second Ni-NTA column purification step by HisTrap FF crude was added before the final gel filtration step (Figure 58). The resulting Hs NMT1 sample was essentially pure (Figure 59) and His-tag free (Figure 60).

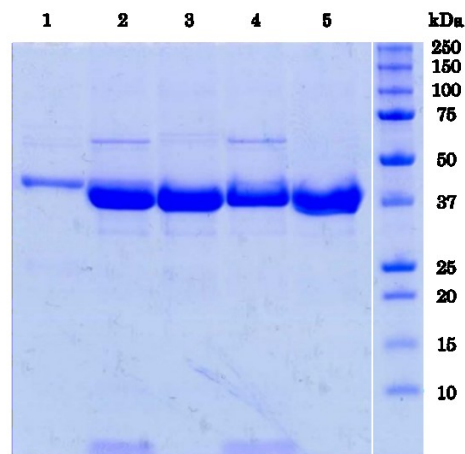


Figure 59 The SDS gel of Hs NMT1 purification. Lane 1 After HisTrap FF crude column. Lane 2 After ion exchange column. Lane 3 Main peak after His-tag cleavage. Lane 4 Shoulder peak after His-tag cleavage. Lane 5 After gel filtration column.

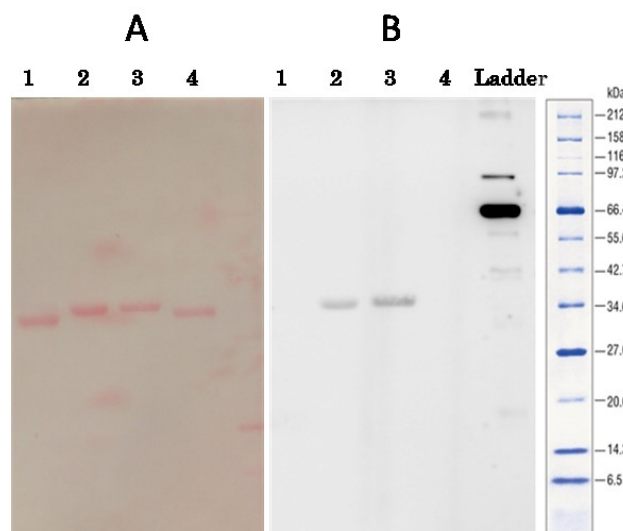


Figure 60 The SDS gel of His-tag presence analysis. (A) Ponceau S Stain shows the location of protein on gel; (B) His-tag in-gel stain shows the location of protein with His tag. Lane 1 and Lane 4 are loaded with samples purified with a 2nd Ni-NTA column after His-tag cleavage. Lane 2 and lane 3 are loaded with samples purified by the original method.

3.1.3 Crystallization

Crystals of apo-Hs NMT1 and 23 co-crystals of selected ligands from series A and B with Hs NMT1 were obtained by the sitting drop vapour diffusion technique. Conditions were modified from two published conditions³⁶ by varying the concentrations of the precipitant, polyethylene glycol (PEG): one with PEG 4K (from 12% to 28%) and the other with PEG monomethyl ether (mPEG) 2K (from 17% to 30%). The rest of the components in the two crystallization conditions were maintained the same.

For each new batch of protein, the binary complex of Hs NMT1 and Myr-CoA was crystallized to test its crystallization activity. For ligands with IC_{50} s lower than 100 nM, the ligand stock solutions were made into 5 mM solutions in 25% DMSO. For ligands with higher IC_{50} s, the original stock solution was used directly to maximize the ligand concentration in the drop. The pH of each ligand solution was tested and adjusted to between 5 and 8 before co-crystallization to prevent precipitation of protein at harsh pH conditions. Ligands were then tested under the two conditions in parallel in 96-well plates, stored at 20 °C and inspected every day. Conditions growing crystals were repeated in 48-well plates to achieve better crystals. For ligands that did not grow crystals in all conditions, a wider range of conditions were tested. All the successfully crystallized ligands grew crystals in the condition containing PEG 4K in 2-4 days (Figure 61A and C). Six of them also grew crystals in the condition containing mPEG 2K after 1-2 weeks from the initial set up (Figure 61B). Both series have ligands which grow crystals in mPEG 2K. The crystals in mPEG 2K were usually bigger and of better quality compared to crystals of the same ligand grown in PEG 4K. The slower growing speed is likely the reason for better crystals, which allows the protein to pack more regularly and uniformly.

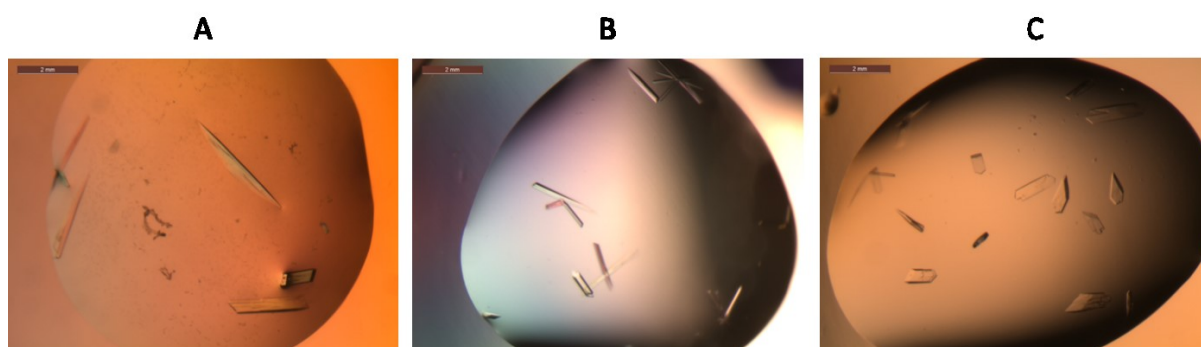


Figure 61 Examples of crystals growing in sitting-drops. (A) Binary complexes of Hs NMT1 and Myr-CoA crystals in PEG 4K; (B) Ternary complexes of Hs NMT1, Myr-CoA and IMP-0001222 crystals in mPEG 2K; (C) Ternary complexes of Hs NMT1, Myr-CoA and IMP-0001169 crystals in PEG 4K.

Some ligands of weaker potencies were crystallized, but their electron density in the binding pocket were too weak to build reliable models in. Their low occupancy in the binding pocket may due to their weaker binding affinity to the protein. Increasing ligand concentrations in the drop from 0.5 mM to 5 mM succeeded in solving structures of **IMP-0000899** and **IMP-0001171**. To prevent ligands diffusing out of crystals during mounting, ligands of the same concentration were added in the cryo-protectants.

However, increasing the ligand concentration of **IMP-0001096** by 10-fold still resulted in low ligand occupancy. Since the original ligand stock solution is 50 mM in 100% DMSO, increasing the ligand concentration will inevitably increase the concentration of DMSO as well. Crystallization of higher concentration of **IMP-0001096** was not successful. To solve this important structure, **IMP-0001096** was resynthesized. A 500 mM stock solution in DMSO was used for co-crystallization and succeeded in growing usable crystals. Soaking method was also tried with 500 mM of **IMP-0001096** in protein buffer without DMSO. In this method, ligand was soaked into pre-existing crystals of Hs NMT1 bound with Myr-CoA. Different soaking time (1 min, 5 min, 30 min, overnight and 3 days) was tested. Crystals soaked overnight or for three days did not survive. Crystals soaked for less time was sent for data collection together with the crystals from co-crystallization. The crystal soaked for thirty minutes was later found

to have the best resolution and the resulting electron density map has clear electron densities of the ligand presented in the protein pocket (Figure 63B).

3.1.4 Data collection

All data was collected at 100 K at Diamond Light Source synchrotron, Oxfordshire, U.K (Figure 62). A test dataset was collected first: three images were collected in exposure time of 0.15 s with an oscillation of 0.5° , a delta of 45° and 80% transmission. According to the test dataset, an estimated resolution was chosen for the complete dataset collection. If any of the three images did not show diffraction pattern, a manual re-centring or a grid scan was performed to re-locate the collection centre. For the complete datasets, 1000 images were collected in exposure time of 0.05 s at an oscillation of 0.2° , a delta of 0° and 40% transmission.

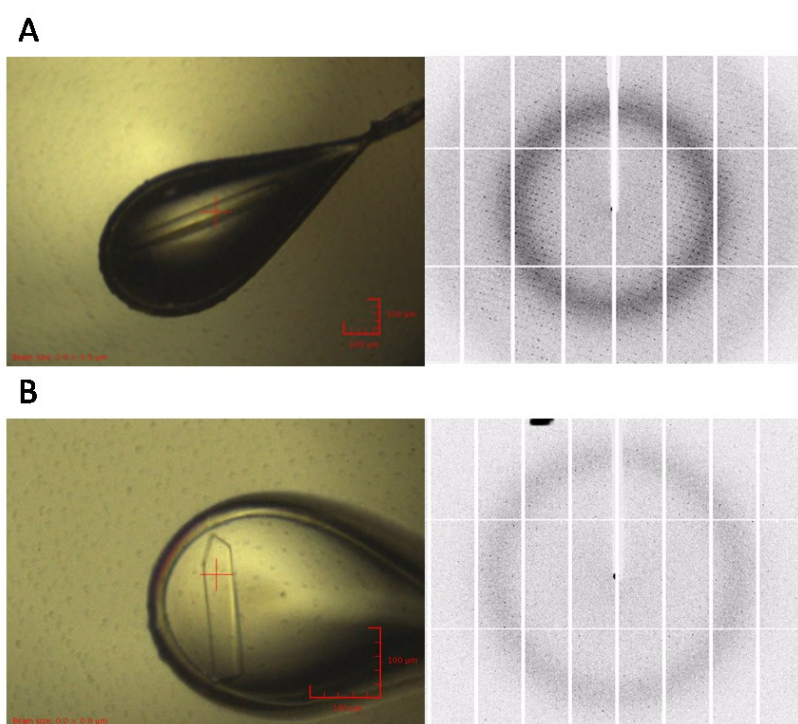


Figure 62 Examples of mounted crystals ready for data collection, with example diffraction images illustrating the raw data obtained. (A) Crystal of the ternary complex of Hs NMT1, Myr-CoA and IMP-0001161; (B) Crystal of the ternary complex of Hs NMT1, Myr-CoA and IMP-0000180.

3.1.5 Data processing

Data were processed by xia2¹⁴⁶ or autoPROC¹⁴⁷ automated data processing system, using XDS¹⁴⁸, DIALS¹⁴⁹, Aimless¹⁵⁰, XSCALE¹⁵¹, Pointless¹⁵², MOSFLM¹⁵³. Phasing of data by Molecular Replacement were achieved by PHASER¹⁵⁴ in CCP4i¹⁵⁵ using the structure of Hs NMT1 with Myr-CoA (PDB code 4C2Y) as the search mode. Two protein molecules are present in the asymmetric unit and the structures fall into the space group P2₁2₁2 or P22₁2₁.

3.1.6 Refinement and model building

After Molecular Replacement, models were visually inspected and manually adjusted according to their electron density using Coot¹⁵⁶. Myr-CoA was then built into the models after initial refinement in PHENIX¹⁵⁷ or REFMAC¹⁵⁸ in CCP4i¹⁵⁵.

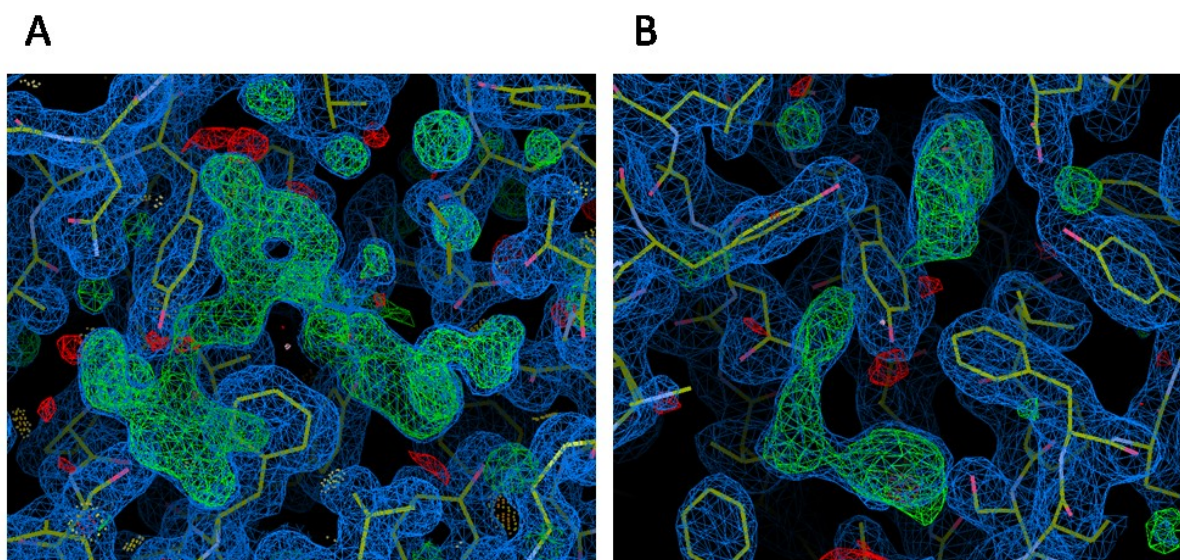


Figure 63 Electron density maps before building the ligand models. $2F_o-F_c$ map at 1σ is shown in blue mesh and F_o-F_c map at 3σ is shown in green (positive) and red (negative) mesh, where F_{obs} (F_o) is the observed amplitude and F_{calc} (F_c) is the model-based amplitude (A) Density map of **IMP-0001223** at 1.49 Å; (B) Density map of **IMP-0001096** at 2.02 Å.

After the second round of refinement, the density of ligands was clearly visible in the peptide binding pocket (Figure 63). Some high-resolution structures have well resolved maps where a

hole appeared in the centre of aromatic rings, while the maps of lower resolution show big blobs of density.

The models of ligands were generated by eLBOW¹⁵⁷ using simple optimization function and built into the density by Ligandfit¹⁵⁷. Where the ligand density was not clear enough, ligand was built after more runs of refinement. Where the ligand density was only clear in one of the protein molecules in the asymmetric unit (**IMP-0001242** only shows one ligand in chain B), one instead of two ligands were built. Where the electron density suggests two conformations of the ligand, the dominant conformation was built but the other conformation was discussed in section 3.3.1. After sufficient refinement to correct the models, water and other small molecules were built into the models and more runs of refinement were performed to lower the R_{work} and R_{free} to the possible minimum. The R-factor¹⁵⁹ is defined as $R = (\sum ||F_{\text{obs}}| - |F_{\text{calc}}||) / \sum |F_{\text{obs}}|$, which is a measure of the agreement between F_{calc} from the crystallographic model and F_{obs} from the experimental X-ray diffraction data. 5% of the experimental data was not included in any refinement, their R-factor (R_{free}) were calculated to see how well the model predicts the observation. By comparing to the R_{work} calculated from the 95% observation used in refinement, R_{free} is used as an indicator of over-interpretation of data.

3.1.7 Conclusion and outlook

Among all the ligand co-crystallizations tried, ligands with lower IC_{50} s in the enzyme assay were more likely to grow crystals and achieve higher structure resolutions. For imidazo[1,2-a]pyridin-6-phenoxy series, the highest IC_{50} of structurally-solved ligand was 1.06 μM and the one from azetidinopyrimidines series has an IC_{50} of 433 nM. This improvement effects of potent inhibitors on crystallization could be a result of their tighter binding. The binding of a ligand in the protein pocket usually stabilize the protein structure, preventing it to unfold, aggregate, or succumb to proteolysis thus facilitating crystallization¹⁶⁰.

High concentration of DMSO is found to disturb the crystallization process. In all attempts, conditions with DMSO concentration higher than 10% usually do not grow crystals. In the cases where crystals did grow, they grow slower and are usually short tiny needles that do not diffract enough for structure solving. For weaker ligands, more concentrated DMSO stock solution or stock solution without DMSO should be used instead of adding more volume of the original stock solution to achieve high concentration of ligands.

Since crystals of Hs NMT1 and Myr-CoA crystallized easily and of good quality, the successful attempt of solving the structure of **IMP-0001096** by soaking presents an alternative method to obtain crystals of ternary complexes of Hs NMT1, Myr-CoA and ligand. The resolution of structure is comparable to other co-crystallized ones, occupancy of the ligand in the protein pocket is high, and the ligand posts similar conformation as the co-crystallized one. The process of crystal preparation is simpler and faster. The data collection and interpretation can be longer since datasets of crystals with different soaking time need to be tested to obtain the best resolution and highest occupancy of ligands. For ligands that easily grow good crystals, co-crystallization can minimize the manipulation of crystals and achieve the best resolution. However, for weaker binding ligands of Hs NMT1 whose crystals are hard to obtain, soaking is a good alternative way.

The missing ligand in chain A of **IMP-0001242** could be the result of low occupancy. Chain A is the relatively more movable protein molecule among the two in the unit, where the protein is not in close contact with other proteins. Therefore, this protein is more dynamic, its ligand binding is looser or its ligand poses more than one conformation, resulting in weak or missing electron density.

3.2 Imidazo[1,2-a]pyridin-6-phenoxy series structure insights

The imidazo[1,2-a]pyridin-6-phenoxy series is a set of potent human NMT inhibitors developed by Mousnier *et al*⁴⁷. Its development started from a fragment-like screen hit of the Pfizer Global Diverse Representative Set1 on Pf and Ld NMTs⁴⁰. After fragment reconstruction, structure-guided fragment linking, and conformational control over linker geometry, several inhibitors reached potencies of single-digit nanomolar level against human NMT1 (More details of its drug discovery and activities are reviewed in 1.3.2 and 1.4.3). By solving eleven ligand-protein structures of this series (the data collection and refinement statistics of imidazo[1,2-a]pyridin-6-phenoxy series is shown in Table 6), this chapter aims to gain a better understanding of the ligand-NMT binding modes to guide future inhibitor development.

Table 6 Data collection and refinement statistics of the imidazo[1,2-a]pyridin-6-phenoxy series

Protein	Hs NMT1					
Ligand	IMP-0001029	IMP-0001114	IMP-0001117	IMP-0001161	IMP-0001180	IMP-0001222
Beamline at DLS ^a	IO4-1	IO3	IO4-1	IO4-1	IO4-1	IO3
Data collection						
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁
Cell dimensions						
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.72 178.85 58.58	79.18 178.65 58.27	58.98 79.06 180.12	58.92 78.84 180.18	59.12 78.91 180.51	58.40 78.87 179.10
Resolution (Å)	72.85 - 2.22	48.40-2.38	60.04-1.79	78.84-1.68	90.25-1.76	47.60-1.93
	(2.27-2.22) ^b	(2.42-2.38)	(1.84-1.79)	(1.72-1.68)	(1.81-1.76)	(1.98-1.93)
No. observations	613143 (45448)	244402 (18527)	580601 (36563)	702134 (52514)	601502 (39101)	462884 (462884)
Unique reflections	42220 (3069)	34027 (2481)	79849 (5836)	96418 (7043)	84200 (6065)	62230 (4406)
<i>R</i> _{merge} ^c	0.125 (0.781)	0.165 (0.490)	0.167 (1.690)	0.141 (1.917)	0.182 (1.512)	0.159 (0.954)
<i>R</i> _{meas} ^d	0.130 (0.810)	0.178 (0.526)	0.179 (1.845)	0.152 (2.059)	0.196 (1.646)	0.170 (1.032)
<i>R</i> _{pim} ^e	0.034 (0.210)	0.065 (0.190)	0.065 (0.723)	0.056 (0.745)	0.073 (0.637)	0.062 (0.387)
<i>R</i> _{merge} in top intensity bin ^f	0.041	0.061	0.030	0.035	0.056	0.040
Mean <i>I</i> /σ(<i>I</i>)	18.7 (4.2)	10.1 (4.3)	11.2 (1.1)	10.3 (1.2)	7.1 (6.4)	10.0 (2.0)
<i>CC</i> _{1/2}	0.999 (0.929)	0.994 (0.623)	0.996 (0.479)	0.998 (0.539)	0.995 (0.505)	0.997 (0.657)
Completeness (%)	99.7 (99.7)	100 (100)	99.6 (99.8)	99.9 (99.9)	99.6 (98.7)	98.8 (96.2)
Multiplicity	14.5 (14.8)	7.2 (7.5)	7.3 (6.3)	7.3 (7.5)	7.1 (6.4)	7.4 (7.0)
Refinement						
<i>R</i> _{work} / <i>R</i> _{free} ^g	0.18/0.21	0.19/0.24	0.18/0.21	0.20/0.23	0.26/0.30	0.19/0.22
No. atoms	6027	6685	7229	6887	6978	6871
Wilson <i>B</i> factor	29.48	24.65	18.59	19.07	13.74	21.32
R.m.s. deviations						
Bond lengths (Å)	0.008	0.008	0.008	0.009	0.009	0.008
Bond angles (°)	0.971	1.118	0.907	1.199	1.037	1.050
Ramachandran statistics						
Favored (%)	97.6	96.8	98.2	97.9	97.4	97.9
Allowed (%)	2.4	3.2	1.9	2.1	2.6	2.1
Outliers (%)	0	0	0	0	0	0

^a Data were collected at Diamond light source. ^b Values in parentheses are for highest-resolution shell. ^c $R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$, where *h, k, l* are the indices of the reflections and $\langle I_{hkl} \rangle$ is the average symmetry- (or Friedel-) related observations of a unique reflection. ^d $R_{\text{meas}} = \sum_{hkl} (n/n-1)^{0.5} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$, where *n* is the multiplicity of the reflection *hkl*. ^e $R_{\text{pim}} = \sum_{hkl} (1/n-1)^{0.5} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$. ^f *R*_{merge} in top intensity bin is provided for data sets scaled using xia2. ^g *R*_{free} test set size is 5%.

Protein Ligand	IMP-0001223	IMP-0001242	Hs NMT1 IMP-0001320	IMP-0001336	IMP-0001343
Beamline at DLS	IO3	IO4-1	IO3	IO3	IO3
Data collection					
Space group	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 22 ₁ 2 ₁
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	58.64 78.98 178.28	79.51 181.11 59.31	77.47 177.86 58.57	58.41 78.08 178.46	58.44 79.58 178.18
Resolution (Å)	89.14-1.18 (1.53-1.49)	90.55-1.84 (1.87-1.84)	71.02-1.45 (1.48-1.45)	47.32-1.95 (2.00-1.95)	48.87-2.22 (2.28-2.22)
No. observations	980038 (69147)	552668 (25719)	1046526 (47720)	445081 (30735)	306794 (20558)
Unique reflections	135732 (9842)	75440 (3709)	143532 (7035)	60424 (4389)	41914 (3045)
<i>R</i> _{merge}	0.141 (1.565)	0.098 (0.937)	0.102 (0.851)	0.163 (0.913)	0.139 (0.854)
<i>R</i> _{meas}	0.152 (1.692)	0.106 (1.013)	0.110 (0.922)	0.175 (0.986)	0.149 (0.926)
<i>R</i> _{pim}	0.056 (0.636)	0.039 (0.382)	0.040 (0.350)	0.064 (0.370)	0.054 (0.351)
<i>R</i> _{merge} in top intensity bin	0.052	-	-	0.046	0.040
Mean <i>I</i> /σ(<i>I</i>)	7.8 (1.4)	13.5 (2.1)	11.7 (2.1)	9.5 (2.1)	11.1 (2.1)
<i>CC</i> _{1/2}	0.997 (0.587)	0.999 (0.866)	0.999 (0.536)	0.996 (0.701)	0.997 (0.685)
Completeness (%)	99.9 (99.6)	100.0 (99.7)	100.0 (98.4)	99.9 (100.0)	99.8 (99.3)
Multiplicity	7.2 (7.0)	7.3 (6.9)	7.3 (6.8)	7.4 (7.0)	7.3 (6.8)
Refinement					
<i>R</i> _{work} / <i>R</i> _{free}	0.21/0.24	0.19/0.22	0.18/0.20	0.18/0.28	0.18/0.23
No. atoms	7318	6853	7213	6891	6721
Wilson <i>B</i> factor	13.31	23.50	13.68	21.03	31.19
R.m.s. deviations					
Bond lengths (Å)	0.006	0.010	0.006	0.007	0.008
Bond angles (°)	0.874	1.199	1.027	0.897	1.121
Ramachandran statistics					
Favored (%)	98.0	97.8	97.8	98.0	97.8
Allowed (%)	2.0	2.2	2.2	2.0	2.2
Outliers (%)	0	0	0	0	0

Inhibitors with similar structures are grouped in twos or threes, and their binding modes were examined to explore the SARs. To demonstrate the relationship between inhibitor potency and its interactions with NMT, especially the number and lengths of hydrogen bonds, 2D ligand-protein interaction diagrams were generated by LigPlot⁺ v1.4.5¹⁶¹. To demonstrate specific conformation differences, 3D binding modes were superposed and analysed by PyMOL v2.0²⁸.

3.2.1 IMP-0001029, IMP-0001114 and IMP-0001343

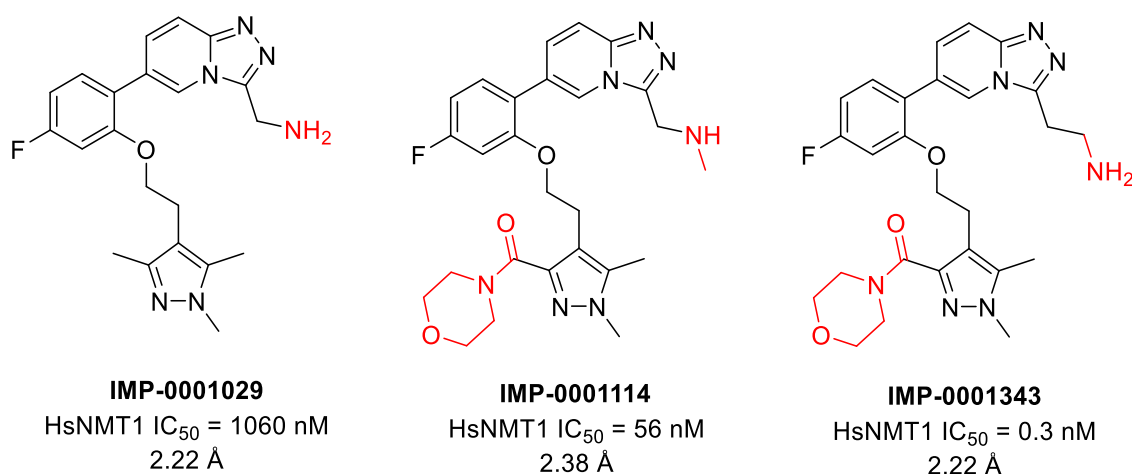


Figure 64 Structures and activities of **IMP-0001029**, **IMP-0001114** and **IMP-0001343**.

The structures of **IMP-0001029**, **IMP-0001114** and **IMP-0001343** differ at both the 3-position of the triazolo group and the 3-position of the pyrazole group (Figure 64). Their binding modes best illustrates the importance of the two key interactions for good potency: The C-terminal Gln496 and the Asn451.

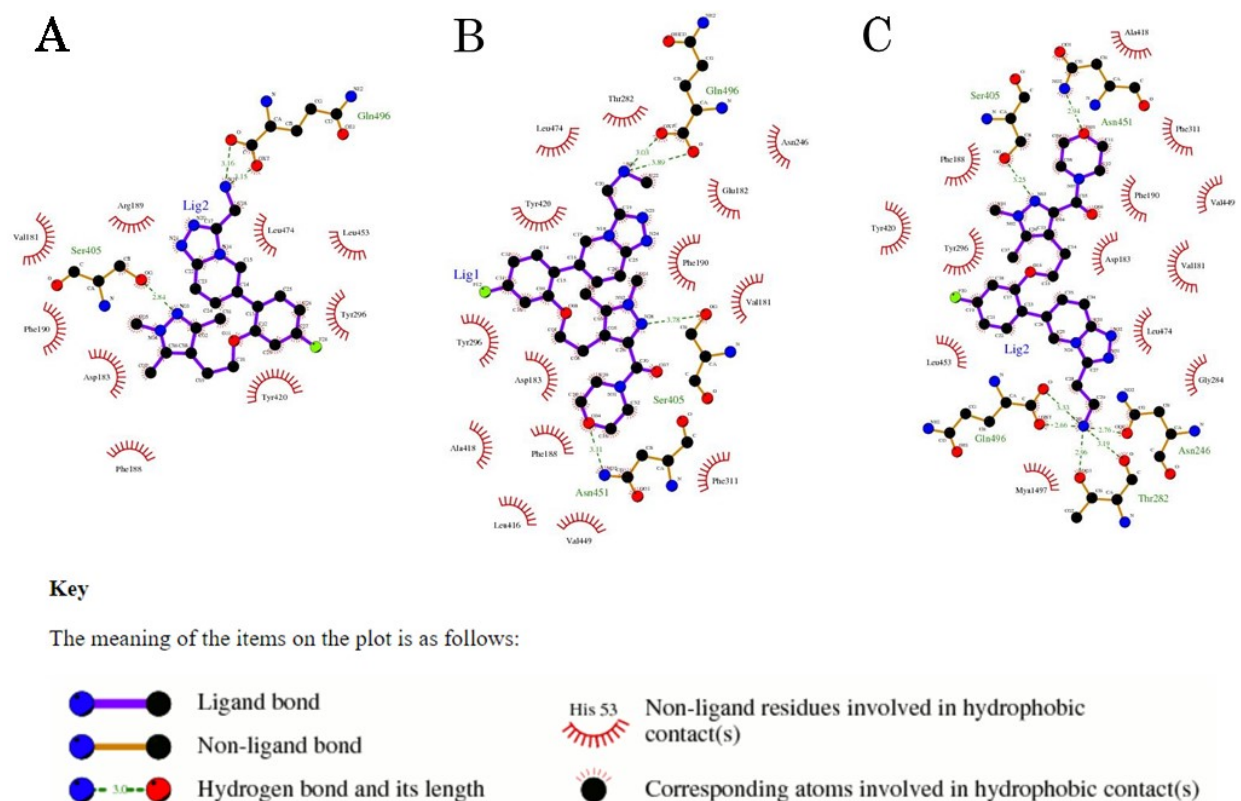


Figure 65 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot⁺ v1.4.5¹⁶¹. (A) IMP-0001029; (B) IMP-0001114; (C) IMP-0001343.

IMP-0001343 has three parts of the inhibitor making polar interactions with Hs NMT1. The oxygen on the morpholine formed a 2.94 Å hydrogen bond with Asn451. The nitrogen lone pair on the pyrazole formed a 3.25 Å hydrogen bond with Ser405. The ethanamine formed polar interactions with three residues: Gln496 (2.66 and 3.33 Å), Thr282 (2.96 and 3.19 Å) and Asn246 (2.76 Å). According to pKa analysis (default setting) of Marvin Sketch¹²¹, at the human blood pH 7.4, 99.33% of the ligand has a protonated primary amine on the ethanamine. Therefore, the polar interaction between ethanamine and the deprotonated carboxyl group of Gln496 is ionic bonding, which is stronger than hydrogen bonds and non-directional.

IMP-0001114 also has three parts making polar interactions. Like **IMP-0001343**, its morpholine formed a 3.11 Å hydrogen bond with Asn451 and its pyrazole formed a 3.78 Å hydrogen bond with Ser405. These two hydrogen bonds are both longer than **IMP-0001343**'s, therefore weaker. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at pH 7.4, 76.82% of the ligand has a protonated secondary amine on the *N*-methylethylmethanamine. Therefore, its interaction is still mostly ionic bonding but weaker than **IMP-0001343**'s. Moreover, because of the one-carbon-shorter distance between the secondary amine and the triazolo group, it can only reach the Gln496 and forms interactions at 3.03 and 3.89 Å lengths, which were longer than **IMP-0001343**'s. The longer and fewer interactions, and the less ionic-bonding character all weaken the interactions and result in the 187-fold decrease in potency comparing to **IMP-0001343**.

Without the morpholino methanone group, **IMP-0001029** has only two parts making polar interactions. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at pH 7.4, 67.76% of the ligand has a protonated primary amine on the methanamine. The lengths of interactions between the methanamine and the Gln496 are 3.15 and 3.16 Å. Therefore, the interaction on the C-terminal site is comparable to **IMP-0001114** but weaker than **IMP-0001343**. Without the morpholino methanone substituent, the pyrazole can optimize and form a shorter hydrogen bond to the Ser405 at 2.84 Å. However, the cost of the missing key interaction with Asn451 cannot be offset by a shorter hydrogen bond, resulting in a 19-fold decrease in potency than **IMP-0001114**.

3.2.2 IMP-0001117, IMP-0001320 and IMP-0001336

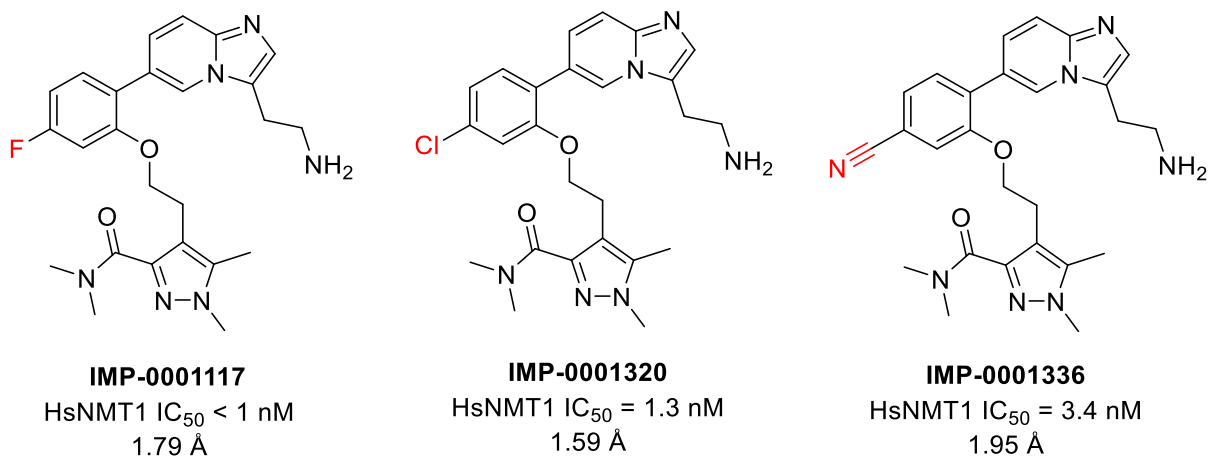


Figure 66 Structures and activities of **IMP-0001117**, **IMP-0001320** and **IMP-0001336**.

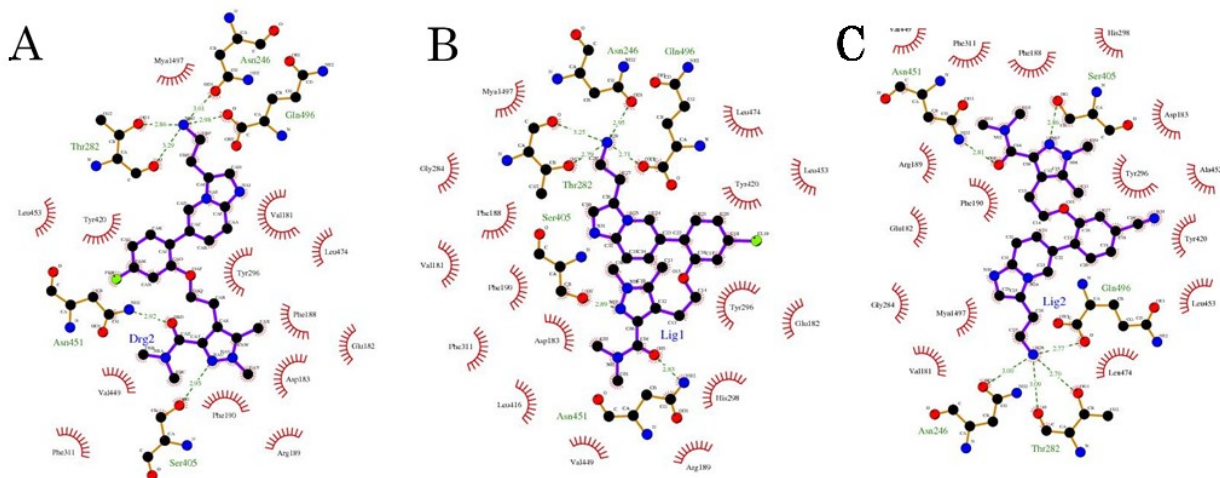


Figure 67 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot⁺ v1.4.5¹⁶¹. (A) **IMP-0001117**; (B) **IMP-0001320**; (C) **IMP-0001336**.

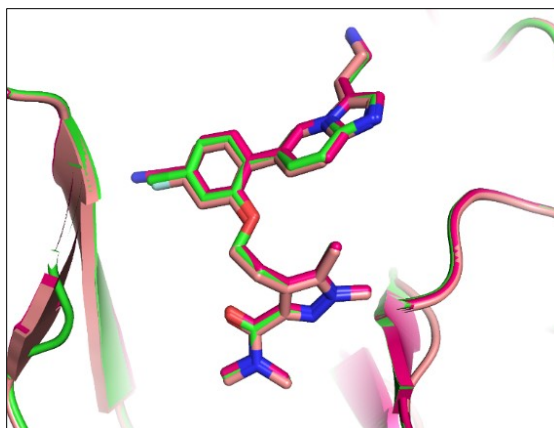


Figure 68 Binding modes of **IMP-0001117**, **IMP-0001320** and **IMP-0001336** in Hs NMT1. The ligand binding modes were superposed by PyMOL²⁸. The structure of **IMP-0001117** is coloured in light pink, **IMP-0001320** in green and **IMP-0001336** in hot pink.

Inhibitor **IMP-0001117**, **IMP-0001320** and **IMP-0001336** only differ in their substituents on the 5-position of the phenoxy group (Figure 66) and have similar interactions with the protein (Figure 67). According to their solved structures, they adopt near identical binding modes in the protein (Figure 68). The major difference is the lengths of their 5-substituents. Since the substituent is poking towards a backbone in the protein pocket, a longer substituent could cause steric hindrance, leading to a decrease in potency. This observation is in accordance with their structure-activity relationships: **IMP-0001117** with the shortest fluoro-substituent shows the highest potency, **IMP-0001336** with the longest cyano-substituent shows the lowest potency and **IMP-0001320** with the median chloro-substituent sit in the middle in potency. However, the potency difference is within test error range and the IC₅₀s reached the lower testing limit (5 nM). Therefore, the SAR found needs further verification on more examples with lower potencies or more accurate activity data is needed to confirm the SAR.

3.2.3 IMP-0001222, IMP-0001223 and IMP-0001242

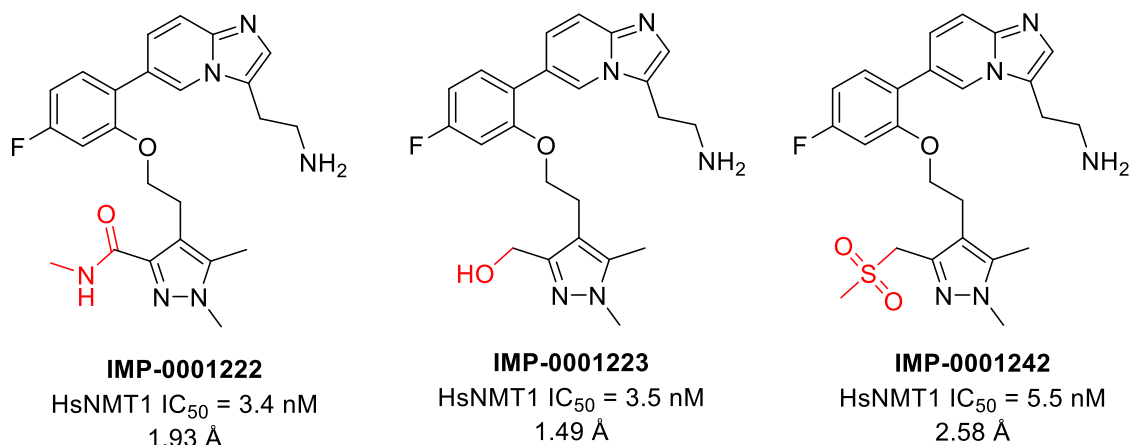


Figure 69 Structures and activities of **IMP-0001222**, **IMP-0001223** and **IMP-0001242**.

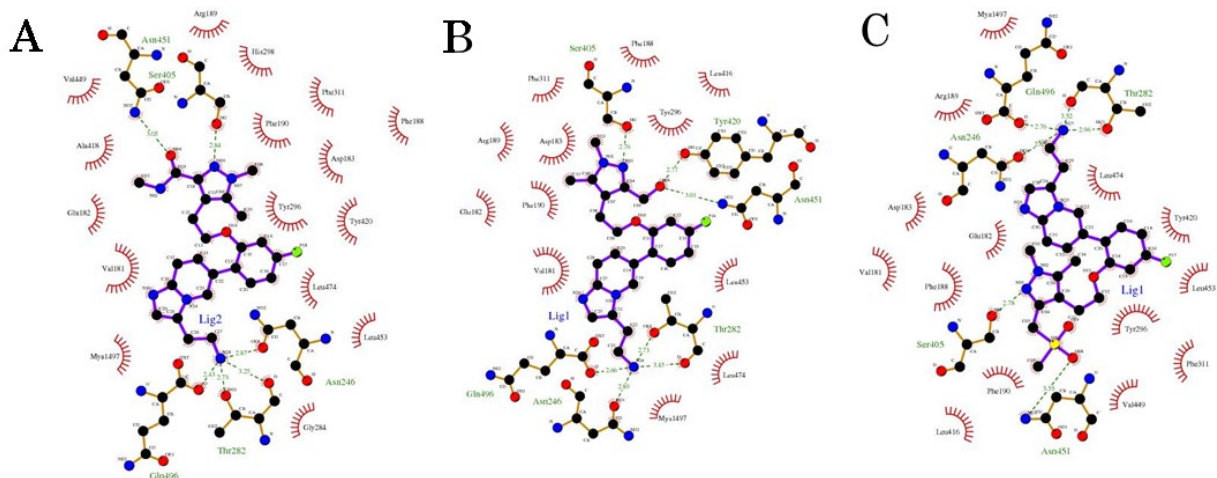


Figure 70 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot+ v1.4.5¹⁶¹.

(A) **IMP-0001222**; (B) **IMP-0001223**; (C) **IMP-0001242**.

IMP-0001222, **IMP-0001223** and **IMP-0001242** are good examples of alternative substituents on the pyrazole to maintain activity. They showed similar potency against Hs NMT1 (Figure 69) and forms similar interactions with Ser405 and Asn451 (Figure 70). **IMP-0001223** has an additional hydrogen bond between the methanol and the residue Tyr420. **IMP-0001242** has a longer hydrogen bond between the methylsulfonyl group and the residue Asn451. However, these differences do not affect the IC₅₀s reported. Since the enzyme assay has reached its limit (5 nM)

and cannot differentiate potency differences at single-digit nanomolar level, cell assay or SPR (surface plasmon resonance) assay need to be used for further activity investigation.

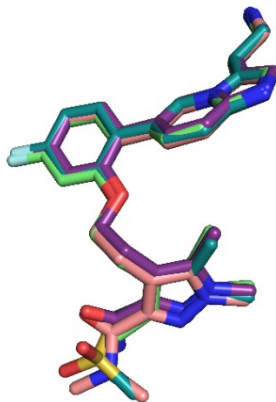


Figure 71 Binding modes of **IMP-0001117**, **IMP-0001222**, **IMP-0001223** and **IMP-0001242** in Hs NMT1. The ligand binding modes were superposed by PyMOL²⁸. The structure of **IMP-0001117** is coloured in light pink, **IMP-0001222** in purple, **IMP-0001223** in light green and **IMP-0001242** in dark green.

The structures of **IMP-0001117**, **IMP-0001222**, **IMP-0001223** and **IMP-0001242** are superposed by PyMOL²⁸ (Figure 71). As expected, they align well where the structure is identical and only differ at the different substituents on the pyrazole.

3.2.4 IMP-0001161 and IMP-0001180

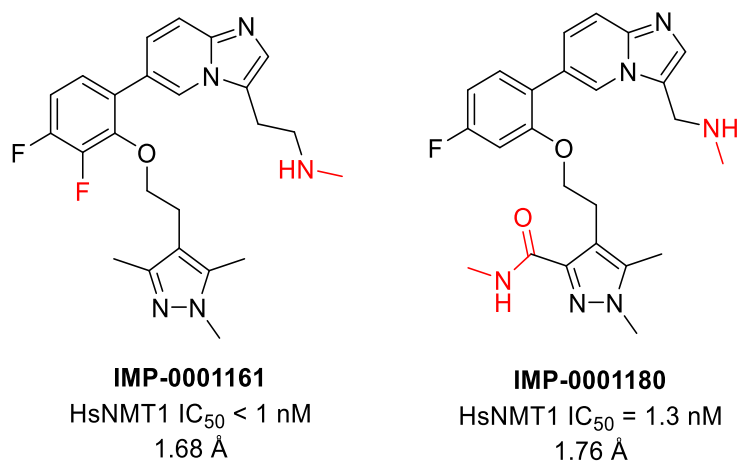


Figure 72 Structures and activities of **IMP-0001161** and **IMP-0001180**.

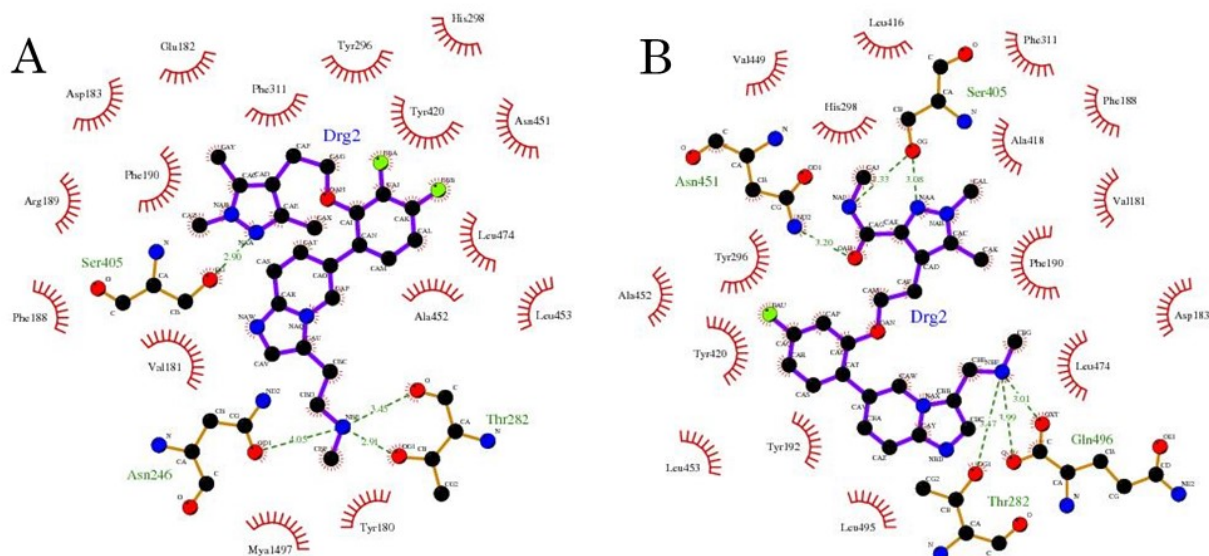


Figure 73 2D ligand-protein binding modes generated from solved X-ray crystal structures by LigPlot⁺ v1.4.5¹⁶¹.
(A) IMP-0001161; (B) IMP-0001180.

IMP-0001161 and **IMP-0001180** have unusually high potencies considering their seemingly suboptimal structures.

IMP-0001161 showed sub-nanomolar activity against Hs NMT1 without a polar substituent on the pyrazole, therefore, no hydrogen bond to Asn451. Figure 74 shows a superposition analysis between **IMP-0001161** and **IMP-0001117**.

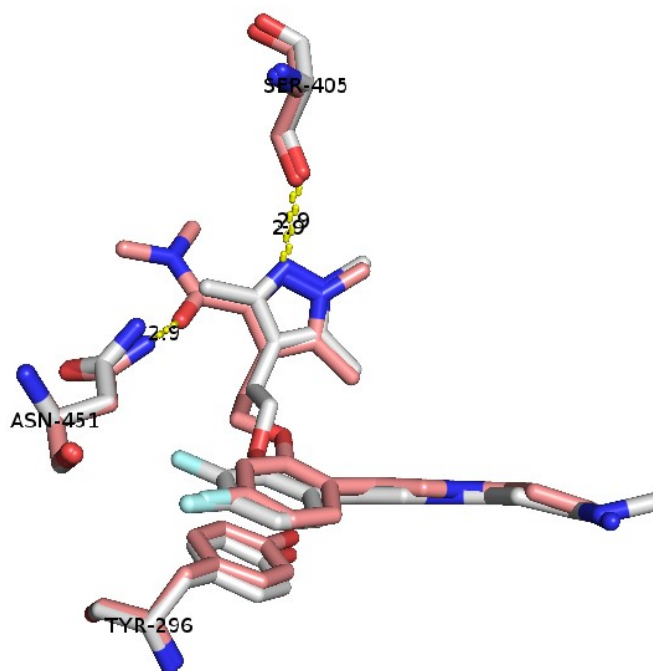


Figure 74 Binding modes comparison of **IMP-0001161** with **IMP-0001117**. The ligand binding modes were superposed and analysed by PyMOL²⁸. The structure of **IMP-0001161** is coloured in white and **IMP-0001117** is in pink.

The additional fluorine ortho to the ether group changed the conformation of the ether linker. This conformational change is also reported by Mousnier *et al.*⁴⁷, which enables the pyrazole to extend further to interact with Ser405 with a shorter distance between **C31** (2.9 Å according to 5O6H) and **C32** (2.6/2.7 Å according to 5O6J). However, the distances to Ser405 of **IMP-0001161** and **IMP-0001117** are both 2.9 Å (Figure 74), therefore the conformation change of ether linker does not give **IMP-0001161** a stronger hydrogen bond to Ser405.

The fluorine substituted ring is parallel to the phenol ring of Tyr296, forming an offset parallel π - π interaction. Offset parallel π - π interaction is stabilized by the favouring London dispersion force and solvophobic effects, over the disfavouring repelling force between the negative quadrupole on the top and bottom of the aromatic rings¹²³. Fluorine, as an electron-withdrawing group, reduces

the negative quadrupole of the aromatic ring and therefore stabilizes the offset parallel π - π interaction. With two fluorine atoms substituted on the benzene ring, **IMP-0001161** has a stronger π - π interaction.

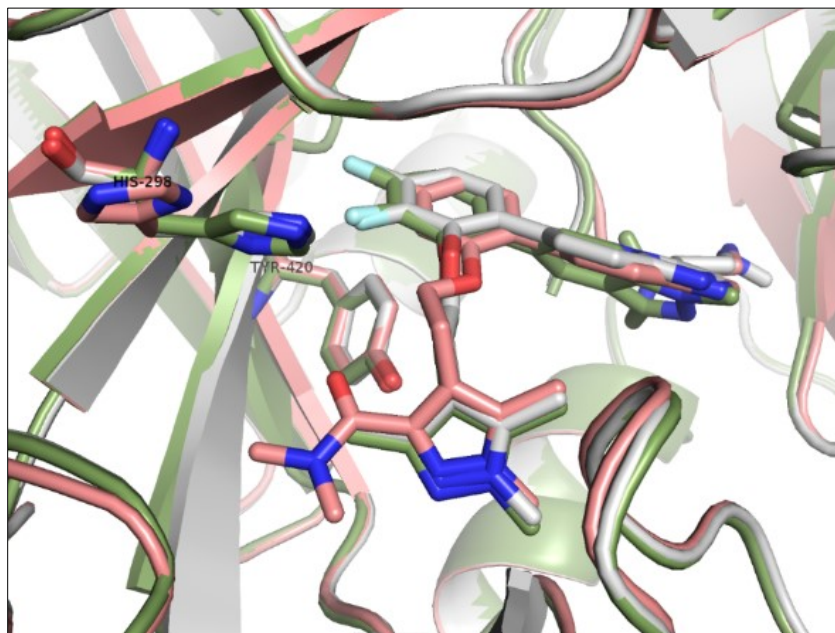


Figure 75 Binding mode comparison of **IMP-0001161**, **IMP-0001088** and **IMP-0001117**. The ligand binding modes were superposed and analysed by PyMOL²⁸. The structure of **IMP-0001161** is coloured in white, **IMP-0001088** in green and **IMP-0001117** is in pink.

The structure of **IMP-0001161** is also compared with the structure of **IMP-0001088** (C33) recently published by Mousnier *et al.*⁴⁷, which also has two fluorine substituents on the same position of the benzene ring. Both structures favour the protein conformation with the residue His298 folded back towards the inhibitor, compared to the His298 of **IMP-0001117** (Figure 75). The CH of the imidazole ring is delta positive, which could make a dipole-dipole with the 6-fluorine atom. The different position of the OCH₂ due to an adjacent fluorine brings the CH₂ closer to Tyr420, thus improving the Van der Waals contact.

The nitrogen on the *N*-methylethanamine of **IMP-0001161** is at a similar position to the nitrogen on the ethyl amine of **IMP-0001117**. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at pH 7.4, 99.66% of the ligand has a protonated secondary amine. It forms similar polar interactions with Thr282 and Asn246, but the *C*-terminal residue Gln496 is not directly bonded with the nitrogen. A closer look at the binding mode reveals a water mediated bonding network among the nitrogen, Asn246 and Gln496 (Figure 76), which should be comparable in strength to an ionic bond.

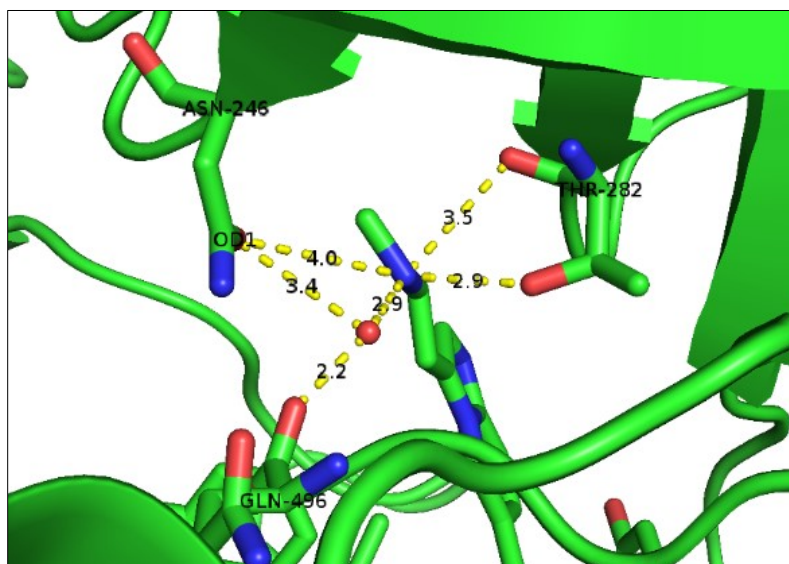


Figure 76 Binding mode of **IMP-0001161** at the *N*-methylethanamine end. The ligand binding modes were analysed using PyMOL²⁸. The ionic bonding interactions are shown as yellow dashes with their distances labelled in black.

IMP-0001180 is closest to **IMP-0001222** in structure. The only difference is the substituent on the 3-position of the imidazo[1,2-*a*]pyridine. This *N*-methylethanamine to ethanamine difference has also been seen in the **IMP-0001114** and **IMP-0001343**. Unlike **IMP-0001114** with the *N*-methylethanamine showing a 187-fold decrease in potency comparing to its analogue **IMP-0001343** with the ethanamine, **IMP-0001180** preserves the potency at the same level as **IMP-**

0001225. Binding modes analysis was done by PyMOL²⁸ to explore the reasons for the potency preservation.

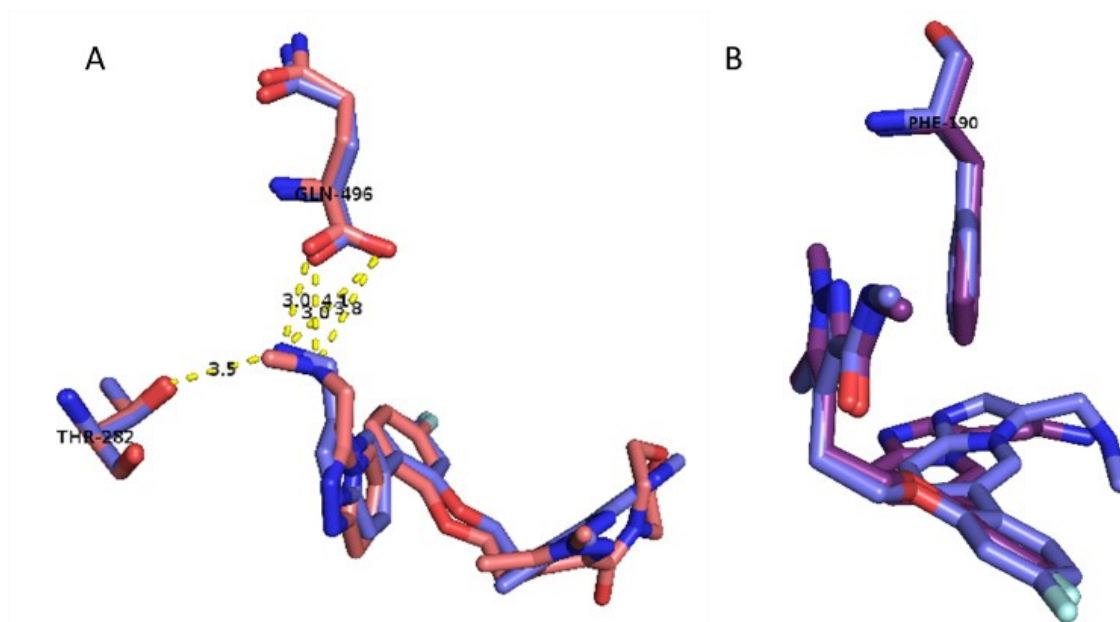


Figure 77 Binding modes comparison of **IMP-0001180** with **IMP-0001114** (A) or **IMP-0001222** (B). The ligand binding modes were superposed and analysed using PyMOL²⁸. The structure of **IMP-0001180** is coloured in blue, **IMP-0001114** in pink, **IMP-0001222** in light green and **IMP-0001242** in purple.

The biggest change from *N*-methylethanamine to ethanamine is the length that the amine can reach: the ethanamine of both **IMP-0001222** and **IMP-0001343** can reach an optimum point for the primary amine to form interactions with Asn246, Thr282 and Gln496, while the *N*-methylethanamine of **IMP-0001114** can only reach and interact with Gln496 alone. However, the *N*-methylethanamine of **IMP-0001180** is positioned differently from **IMP-0001114** due to the structure difference in the pyrazole ends, enabling it to interact with Thr282 as well (Figure 77A), which can contribute to the potency.

The second contributor to the maintained potency is the π - π stacking interaction with Phe190 (Figure 77B). Compared to **IMP-0001222**, the position of the imidazo[1,2-*a*]pyridine on **IMP-**

0001180 shifts considerably to maximize the reach of the *N*-methylethanolamine to form interactions. This shift also positions the imidazo[1,2-*a*]pyridine orthogonal to the phenyl group of the Phe190, forming an T-shaped π - π stacking interaction. This additional π - π stacking interaction compensates the loss of an ionic bonding with Asn246, resulting in a comparable potency to **IMP-0001222**.

3.2.5 *C*-terminus conformation

When fitting the electron density with protein modes, the *C*-terminus of structures **IMP-0001223**, **IMP-0001161** and **IMP-0001320** show areas of positive and negative density on the F_o-F_c difference map (contoured at 3 σ) after several rounds of refinement. These three structures happen to have the highest resolutions. It is likely that the *C*-terminus has multiple conformations and structures with higher resolutions show stronger signs of multi-conformations. This phenomenon is not seen in the azetidinopyrimidines series and the non-ligand bound protein structure, therefore the movement of the *C*-terminus is likely caused by the binding of imidazo[1,2-*a*]pyridin-6-phenoxy series ligands.

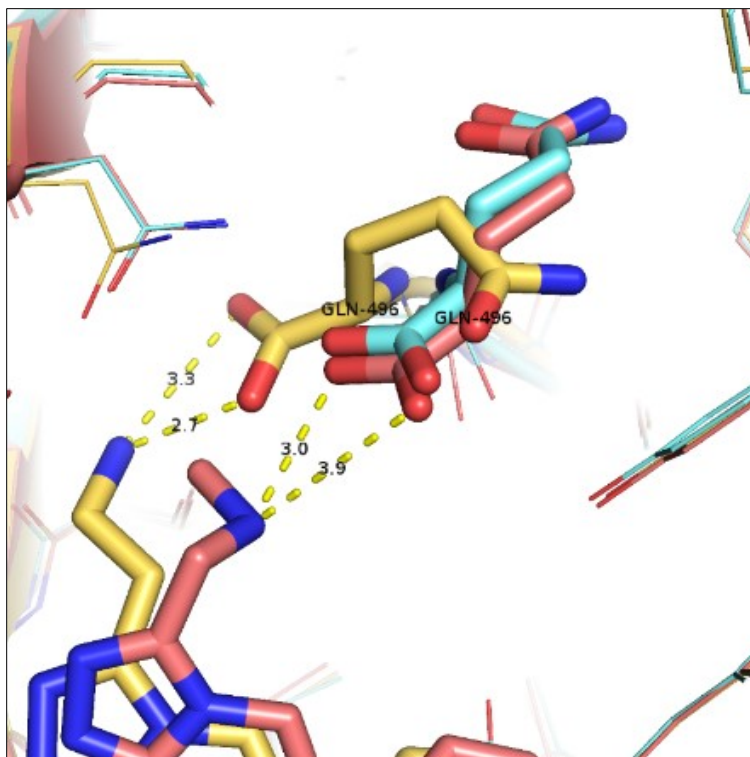


Figure 78 Two conformations of the *C*-terminus with different ligand binding. The ligand binding modes were superposed and analysed using PyMOL²⁸. The structure of **IMP-0001343** is coloured in yellow, **IMP-0001114** in pink and the no-ligand protein structure (PDB code: 3IUI) in light blue.

In structures with 2-aminoethyl on the imidazo[1,2-*a*]pyridine (**IMP-0001117**, **IMP-0001320**, **IMP-0001336**, **IMP-0001343**, **IMP-0001222**, **IMP-0001242** and **IMP-0001223**), the *C*-terminus has a conformation as **IMP-0001343** (Figure 78), where the Gln496 shifts to interact with the amine. In structures with other amino substituents on the imidazo[1,2-*a*]pyridine (**IMP-0001180**, **IMP-0001029**, **IMP-0001114** and **IMP-0001161**), the *C*-terminus has a conformation as **IMP-0001114** (Figure 78), which is the same conformation as the apo-protein structure (PDB code: 3IUI). Inhibitors with 2-aminoethyl substituents on the imidazo[1,2-*a*]pyridine all show very good IC_{50} s (<1 nM to 5.5 nM). As discussed in section 3.2.1, the good potencies are likely due to additional hydrogen bonds with residues Thr282 and Asn246, rather than the change of *C*-terminus conformation.

3.2.6 Conclusion and outlook

Eleven ligand-protein structures of the imidazo[1,2-a]pyridin-6-phenoxy series have been solved by X-ray crystallography. Binding modes analysis has revealed the key interactions to gain potency against Hs NMT1 at single-digit nanomolar level.

The C-terminal Gln496 requires a two-atom spacing between the interacting amine and the imidazo[1,2-a]pyridine core. At physiological conditions, the amine is protonated and forms ionic bonding interactions. Secondary and primary amines form interactions of similar strength.

On the other end, all the ligands form a hydrogen bond with Ser405 by the nitrogen lone pair on the pyrazole. A hydrogen donor on the 3-substituent of the pyrazole is required to form the key hydrogen bond with Asn451. Shorter bond lengths contribute to better potency.

Strengthened π - π stacking interactions also contribute to potency improvement. The main π - π stacking interactions are the offset parallel π - π interaction between Tyr296 and the benzene, and the T-shaped π - π stacking interaction between Phe190 and the imidazo[1,2-a]pyridine. Electron withdrawing substituent can strengthen offset parallel π - π interactions. The relative positions of the two aromatic systems are important for π - π stacking interactions.

3.3 Azetidinopyrimidines series structure insights

The azetidinopyrimidines series was developed from two screen hits as discussed in section 2.2. It shows good potencies against Hs NMT1 down to single-digit nanomolar values despite the initial intention to develop a series against *Plasmodium* NMTs. This subchapter aims to understand the binding modes of the azetidinopyrimidines series with Hs NMT1 by solving and analysing nine X-ray crystal structures. The data collection and refinement statistics of azetidinopyrimidines series is shown in Table 7.

Structures of azetidinopyrimidines series solved by others are also analysed and discussed in this subchapter to generate an overall view of the binding modes of this series. Structure **IMP-0000898** was solved by *Dr* James Brannigan (Unpublished). Structure **IMP-0000366** was solved by *Dr* Inmaculada Perez Dorado (Unpublished).

Table 7 Data collection and refinement statistics of the azetidinopyrimidines series

Protein	Hs NMT1				
Ligand	IMP-0000180	IMP-0000899	IMP-0000927	IMP-0000974	IMP-0001096
Beamline at DLS ^a	IO4-1	IO3	IO4-1	IO4-1	IO3
Data collection					
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	80.31 181.48 58.86	80.00 179.25 58.46	80.68 180.45 58.86	79.95 181.79 59.13	80.72 178.72 58.15
Resolution (Å)	60.49-1.82 (1.87-1.82) ^b	89.63-1.97 (2.02-1.97)	80.68-2.05 (2.10-2.05)	79.95-1.76 (1.81-1.76)	59.90-2.03 (2.08-2.03)
No. observations	559526 (45448)	444821 (33404)	404850 (30829)	559387 (28277)	405839 (30194)
Unique reflections	77687 (3069)	60426 (4430)	54860 (4032)	86121 (6303)	55292 (4052)
<i>R</i> _{merge} ^c	0.148 (1.510)	0.173 (1.769)	0.176 (2.162)	0.104 (1.292)	0.157 (1.930)
<i>R</i> _{meas} ^d	0.159 (1.656)	0.186 (1.899)	0.189 (2.319)	0.113 (1.462)	0.169 (2.075)
<i>R</i> _{pim} ^e	0.058 (0.664)	0.068 (0.684)	0.069 (0.832)	0.043 (0.670)	0.061 (0.754)
<i>R</i> _{merge} in top intensity bin ^f	0.034	0.043	0.028	0.027	0.035
Mean <i>I</i> /σ(<i>I</i>)	10.4 (1.1)	9.1 (1.1)	10.7 (1.1)	12.2 (1.0)	10.1 (1.2)
<i>CC</i> _{1/2}	0.997 (0.489)	0.996 (0.533)	0.997 (0.599)	0.999 (0.558)	0.998 (0.714)
Completeness (%)	99.6 (99.4)	100.0 (100.0)	99.6 (99.1)	99.7 (99.5)	99.9 (99.9)
Multiplicity	7.2 (6.0)	7.4 (7.5)	7.4 (7.6)	6.5 (4.5)	7.3 (7.5)
Refinement					
<i>R</i> _{work} / <i>R</i> _{free} ^g	0.21/0.25	0.19/0.23	0.21/0.25	0.19/0.23	0.20/0.25
No. atoms	6689	6961	6689	7178	6616
Wilson <i>B</i> factor	30.37	28.11	30.37	20.69	34.62
R.m.s. deviations					
Bond lengths (Å)	0.010	0.008	0.010	0.008	0.008
Bond angles (°)	1.204	1.140	1.204	0.982	0.915
Ramachandran statistics					
Favored (%)	98.0	97.4	98.0	98.2	96.5
Allowed (%)	2.0	2.6	2.0	1.8	3.5
Outliers (%)	0	0	0	0	0

^a Data were collected at Diamond light source. ^b Values in parentheses are for highest-resolution shell. ^c $R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$, where *h, k, l* are the indices of the reflections and $\langle I_{hkl} \rangle$ is the average symmetry- (or Friedel-) related observations of a unique reflection. ^d $R_{\text{meas}} = \sum_{hkl} (n/n-1)^{0.5} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$. ^e $R_{\text{pim}} = \sum_{hkl} (1/n-1)^{0.5} \sum_{j=1}^n |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$, where *n* is the multiplicity of the reflection *hkl*. ^f *R*_{merge} in top intensity bin is provided for data sets scaled using xia2. ^g *R*_{free} test set size is 5%.

Protein	Hs NMT1				
Ligand	IMP-0001169	IMP-0001171	IMP-0001210	IMP-0001227	IMP-0001232
Beamline at DLS	IO3	IO4-1	IO3	IO3	IO3
Data collection					
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 22 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.97 179.94 58.53	80.74 179.04 58.65	79.70 177.84 58.37	58.51 79.11 179.28	79.67 179.08 58.44
Resolution (Å)	41.82-2.06 (2.11-2.06)	73.60-2.16 (2.22-2.16)	79.70-2.04 (2.09-2.04)	89.64-1.81 (1.86-1.81)	47.12-1.80 (1.85-1.80)
No. observations	390665 (29471)	343156 (25383)	393127 (29676)	535643 (32903)	578585 (41552)
Unique reflections	53167 (3875)	343156 (25383)	53729 (3913)	76834 (5592)	78439 (5711)
<i>R</i> _{merge}	0.194 (1.911)	0.255 (2.155)	0.176 (1.839)	0.147 (1.255)	0.141 (2.201)
<i>R</i> _{meas}	0.209 (2.052)	0.274 (2.312)	0.190 (1.974)	0.158 (1.377)	0.152 (2.372)
<i>R</i> _{pim}	0.076 (0.741)	0.100 (0.828)	0.070 (0.712)	0.059 (0.560)	0.056 (0.876)
<i>R</i> _{merge} in top intensity bin	0.047	0.041	0.051	0.052	0.047
Mean <i>I</i> /σ(<i>I</i>)	8.3 (1.1)	8.2 (1.1)	8.5 (1.0)	7.5 (1.3)	8.9 (0.9)
<i>CC</i> _{1/2}	0.996 (0.550)	0.993 (0.525)	0.996 (0.661)	0.995 (0.675)	0.997 (0.509)
Completeness (%)	99.9 (99.8)	99.9 (99.9)	99.8 (99.6)	100.0 (100.0)	100.0 (100.0)
Multiplicity	7.3 (7.6)	7.4 (7.5)	7.3 (7.6)	7.0 (5.9)	7.4 (7.3)
Refinement					
<i>R</i> _{work} / <i>R</i> _{free}	0.20/0.24	0.21/0.25	0.20/0.25	0.22/0.25	0.19/0.22
No. atoms	6650	6271	6520	6887	6863
Wilson <i>B</i> factor	30.55	28.62	33.92	18.68	25.50
R.m.s. deviations					
Bond lengths (Å)	0.008	0.009	0.009	0.008	0.008
Bond angles (°)	0.895	1.161	1.010	0.939	1.167
Ramachandran statistics					
Favored (%)	97.0	97.3	97.3	97.2	97.6
Allowed (%)	3.0	2.7	2.7	2.8	2.4
Outliers (%)	0	0	0	0	0

3.3.1 Two conformations

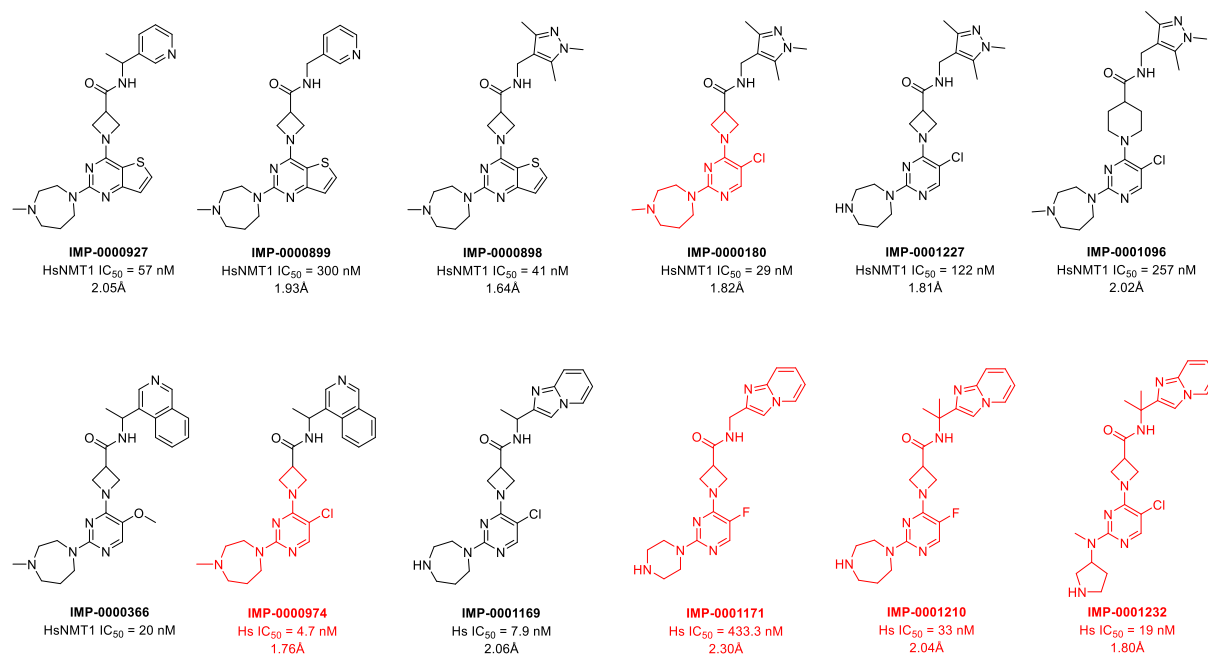


Figure 79 Azetidinopyrimidines series structures with solved X-ray crystal structures. Structures are sorted by structural similarity and colored by different conformations. Structures in black adopt an S-shape, structures in red adopt a C-shape and structures in half black and half red adopt both conformations.

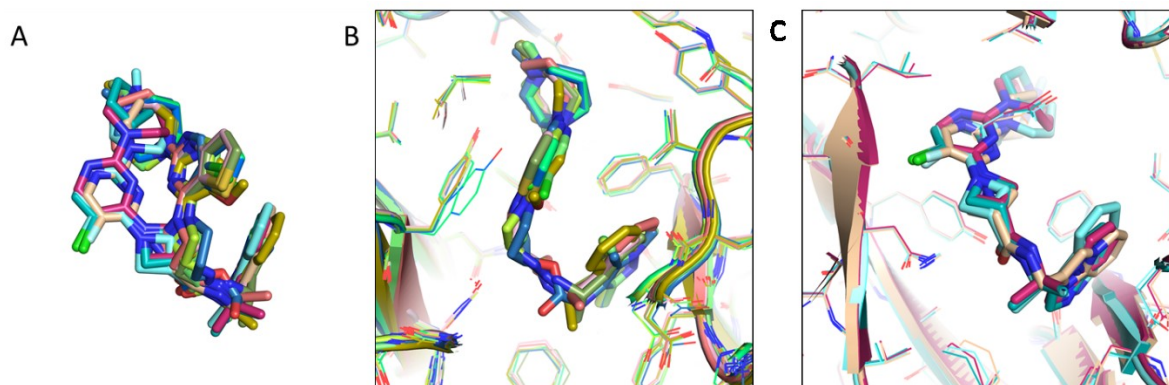


Figure 80 Two different binding conformations among azetidinopyrimidines series. (A) Over view of binding modes of all eleven structures; (B) Superimposed structures in an S-shape; (C) Superimposed structures in a C-shape.

Surprisingly, the eleven compounds with similar structures (Figure 79) adopt two different conformations in the binding pocket of the Hs NMT1 protein (Figure 80A). One conformation is in an S-shape where the pyrimidine ring bends in the same direction of the aromatic substitutes on the amide and the hydrophilic end of the ligand bends back to interact with the

C-terminal Gln496 (Figure 80B). The other conformation is in a C-shape where the pyrimidine ring bends in the opposite direction of the aromatic substitutes on the amide and the hydrophilic end points straight to interact with the C-terminal Gln496 (Figure 80C).

One observation is that all compounds adopting a C-shape (Structures in red in Figure 79) do not have a big substitution on the 5- and 6-positions of the pyrimidine. According to the X-ray crystal structures of the compounds in the C-shape, the 5-substitution on the pyrimidine is pointing towards the protein backbone (Figure 80C). A bigger 5-substitution can cause steric hindrance pushing the compound into an S-shape. In the eleven structures solved, all the compounds with a 5-substitution group bigger than one atom adopt the S-shape. Especially for **IMP-0000366**, it differs from **IMP-0000974** only in the 5-substitution but adopts the S-shape (Structures in black in Figure 79). However, this does not mean that the C-shape is at a lower energy state than the S-shape. Compounds in both conformations can reach single-digit nanomolar activity against human NMT1, for example, **IMP-0001169** has an IC_{50} of 7.9 nM and **IMP-0000974** has an IC_{50} of 4.7 nM.

Another observation is that compounds with a single aromatic ring on the amide mostly adopt the S-shape and compounds with a bicyclic aromatic ring on the amide mostly adopt the C-shape. However, **IMP-0000366** with a diazepane and **IMP-0001169** with a imidazo[1,2-a]pyridine appeared in an S-shape in the binding pocket. For **IMP-0000366**, this can be explained by the steric hindrance discussed in the previous paragraph. However, for **IMP-0001169** which is very similar in structure to **IMP-0001171** and **IMP-0001210**, it is strange that it adopts a different conformation.

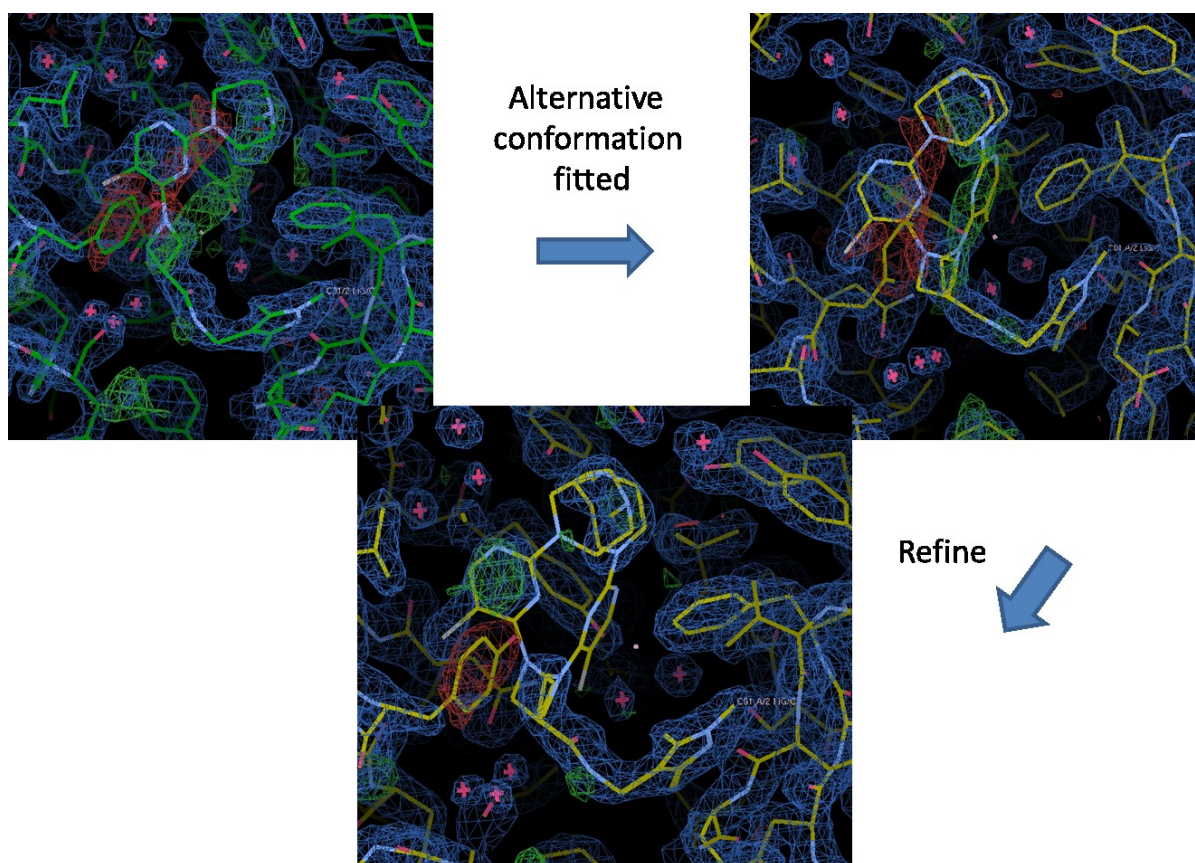


Figure 81 The two conformations of **IMP-0000180** in chain A of the human NMT1 protein. Images were generated by Coot with $2F_o-F_c$ map contoured at 1.0σ and F_o-F_c map contoured at 3.0σ .

The two conformations are also found in the crystal structure of the same compound. **IMP-0000180** and **IMP-0000974** both appear in two conformations in the binding pocket. **IMP-0000180**, for example, was fitted in a C-shape initially into the electron density. However, after refinement, red blocks of the F_o-F_c map appeared around the pyrimidine and green blocks of the F_o-F_c map appeared on the other side suggesting an S-shape conformation co-existing in the structure (Figure 81). When fitted with both conformations (50% occupation each), the F_o-F_c map appeared more agreeable after refinement (Figure 81). The constant red block on the Tyr296 showed the flexibility of the residue, which is likely due to its different positions corresponding to each ligand conformation.

Among all the compounds with a one-atom 5-substitution on the pyrimidine, **IMP-0000180** and **IMP-0000974** are the two structures with the highest resolutions. It is likely that all these

compounds adopt both conformations at various percentage occupations, but only the structures with the best resolutions show the two conformations clearly enough to be observed. Higher quality structures of other compounds in this series will be needed to confirm this hypothesis.

3.3.2 Structure-activity relationship analysis

Compounds of the azetidinopyrimidines series are arranged into groups of three according to their structural similarities and their SARs are explored.

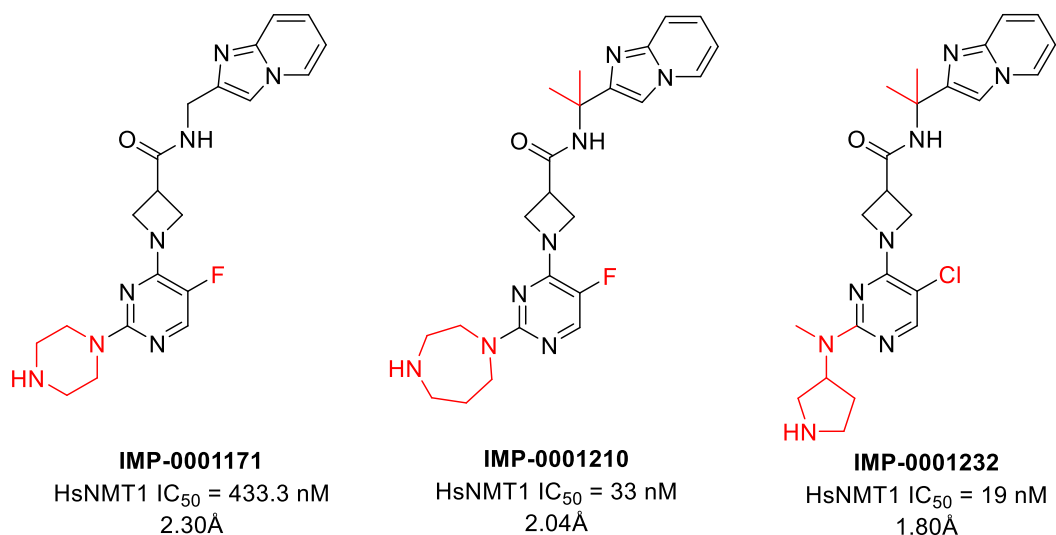


Figure 82 Structures and activities of **IMP-0001171**, **IMP-0001210** and **IMP-0001232**.



Figure 83 Binding modes of **IMP-0001171**, **IMP-0001210** and **IMP-0001232** highlighting the interactions with Ser405. The ligand binding modes were analysed using PyMOL²⁸. **IMP-0001171** is shown in beige, **IMP-0001210** in hot pink and **IMP-0001232** in light blue. The hydrogen bonds with Ser405 are shown in yellow dashes with their lengths labelled in black.

IMP-0001171, **IMP-0001210** and **IMP-0001232** all adopt the C-shape (Figure 83). All the secondary amines reach a similar position to interact with the C-terminus. Due to its longer length, **IMP-0001210** has a curvier conformation. **IMP-0001232** has a 5-chlorine on the pyrimidine, which is slightly longer and bigger than the 5-fluorine on **IMP-0001171** and **IMP-0001210**. However, the structural difference having the biggest impact on activities should be the two methyl groups on the carbon linker between the amide and the imidazo[1,2-a]pyridine. Without them, the imidazo[1,2-a]pyridine on **IMP-0001171** is more flexible and shift away from the Ser409, making a weak 4.0 Å hydrogen bond compared to 2.9 Å of **IMP-0001210** and 2.7 Å of **IMP-0001232** (Figure 83). The hydrogen binding with Ser409 is one of the two important interactions among all series against NMTs^{32, 75-76, 78, 111, 162}. A weaker hydrogen bond costs **IMP-0001171** thirteen-fold and twenty-three-fold potency loss compared to **IMP-0001210** and **IMP-0001232**, respectively.

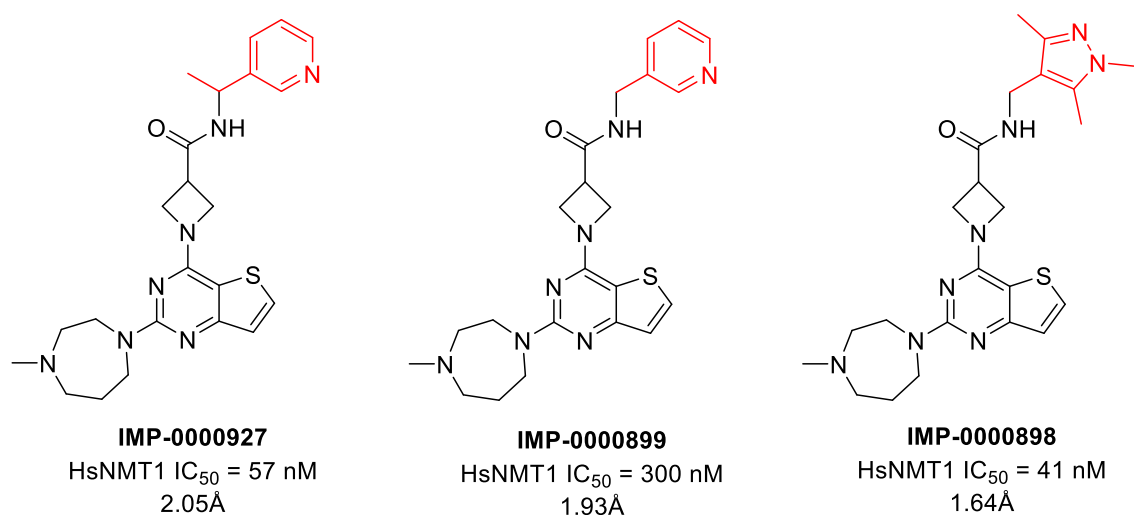


Figure 84 Structures and activities of **IMP-0000927**, **IMP-0000899** and **IMP-0000898**.

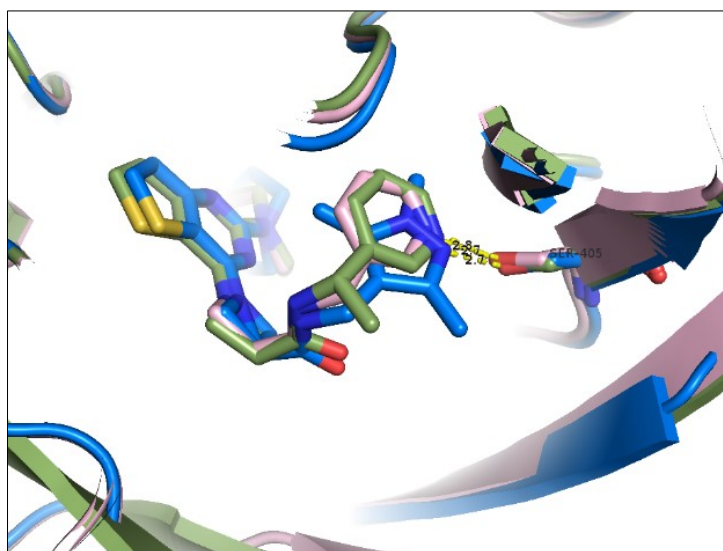


Figure 85 Binding modes of **IMP-0000927**, **IMP-0000899** and **IMP-0000898** highlighting the interaction with Ser405. The ligand binding modes were analysed using PyMOL²⁸. **IMP-0000927** is shown in green, **IMP-0000899** in pink and **IMP-0000898** in blue. Hydrogen bonds with Ser405 are shown in yellow dashes with their lengths labelled in black.

IMP-0000927, **IMP-0000899** and **IMP-0000898** all adopt the S-shape due to their bulky thieno[3,2-d]pyrimidine. Their only difference is the substitutions on the amide.

Although the X-ray crystal structures show similar lengths of the interactions between the secondary amines and Ser405, **IMP-0000927** with one methyl group on the one carbon linker between the amide and the pyridine showed five-fold better potency than its analogue, **IMP-0000899**, with no methyl substitution on the same position. Because of the identical structure

and binding mode in the rest of the structures, the potency difference could be a result from the different flexibility level of the pyridine. The more flexible the compound is, the more configurational entropy is lost when it binds to the protein, resulting in higher penalty in binding affinity¹⁶³. **IMP-0000927** with one methyl substitution is less flexible than **IMP-0000899** with no substitution, thus showing a better potency.

However, **IMP-0000898** with a 1,3,5-trimethyl-1*H*-pyrazole and no substitution on the one carbon linker does not seem to pay a higher penalty, showing similar potency as **IMP-0000927**. This may be due to the conformational restriction from the three methyl groups on the pyrazole, which reduces its flexibility. In addition, the potential steric hindrance of the *N*-methyl group with the protein backbone forces the compound to adopt a conformation where the *NH* group has a more approachable angle towards Ser409 to form the important hydrogen bond (Figure 85).

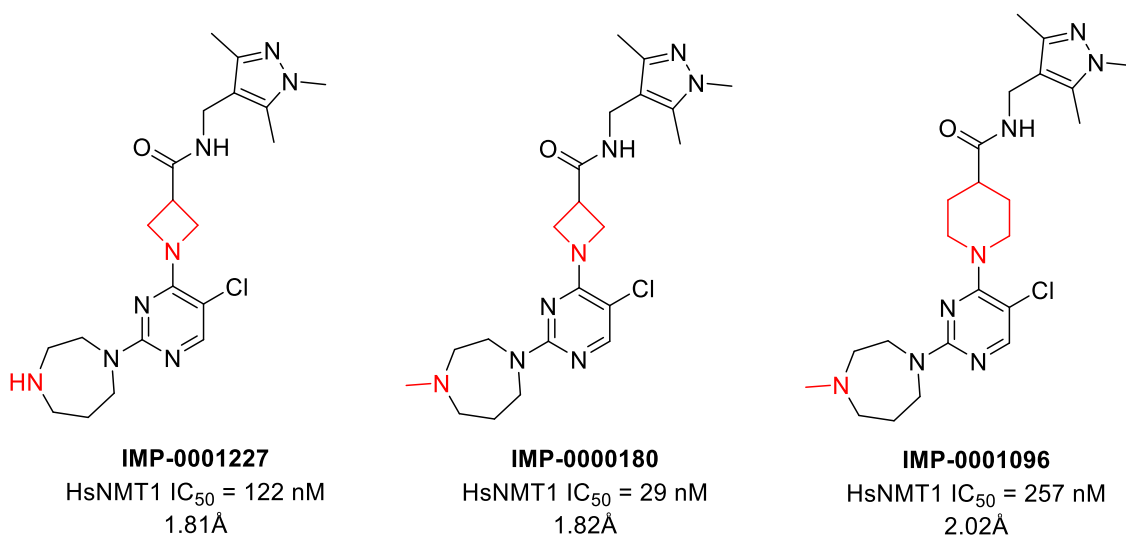


Figure 86 Structures and activities of **IMP-0001227**, **IMP-0000180** and **IMP-0001096**.

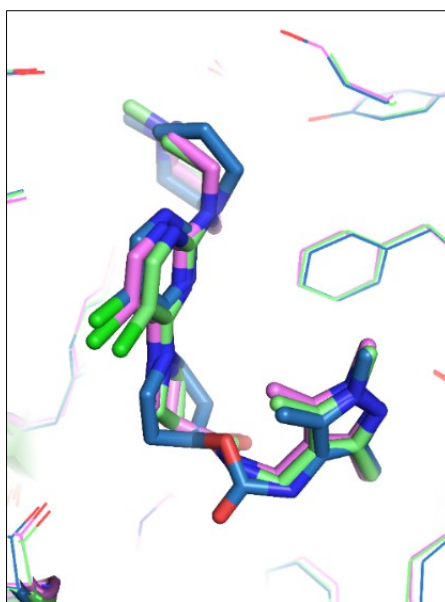


Figure 87 Binding modes of **IMP-0001227**, **IMP-0000180** and **IMP-0001096**. The ligand binding modes were analysed using PyMOL²⁸. **IMP-0001227** is shown in purple, **IMP-0000180** in green and **IMP-0001096** in blue.

IMP-0001227, **IMP-0000180** and **IMP-0001096** all have a 1,3,5-trimethyl-1*H*-pyrazole and a 5-chlorine pyrimidine substitution. According to pKa analysis (default setting) of Marvin Sketch¹²¹, at the human blood pH 7.4, 99.23% of **IMP-0001227**, 86.34% of **IMP-0000180** and 86.35% of **IMP-0001096** have a protonated 4*N* on the diazepane. Therefore, they all form ionic bindings with the *C*-terminus, which is stronger than hydrogen bonds and non-directional. **IMP-0001227** and **IMP-0000180** adopt identical conformations. According to the percentage protonation, **IMP-0001227** should have a stronger ionic bond than **IMP-0000180**. With a *N*-methyl diazepane having a pKa of 8.2, **IMP-0000180** has an cLogD of 0.58¹²¹. With an unsubstituted diazepane having a pKa of 9.51, **IMP-0001227** has an cLogD of -0.91¹²¹. The four-fold potency decrease of **IMP-0001227** could result from less non-specific binding due to a lower cLogD. The conformation comparison between **IMP-0000180** and **IMP-0001096** are discussed in detail in section 2.2.

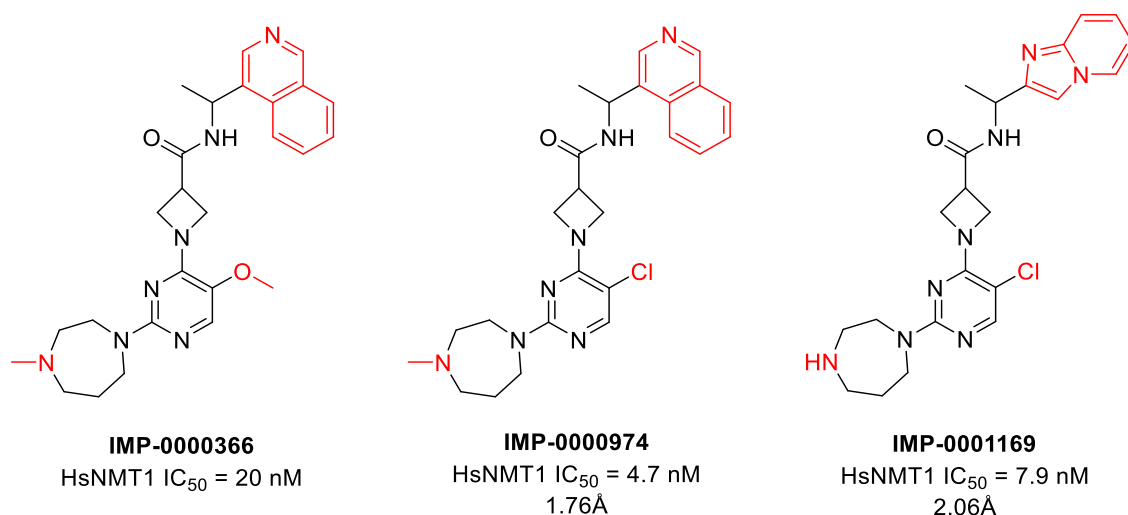


Figure 88 Structures and activities of **IMP-0000366**, **IMP-0000974** and **IMP-0001169**.

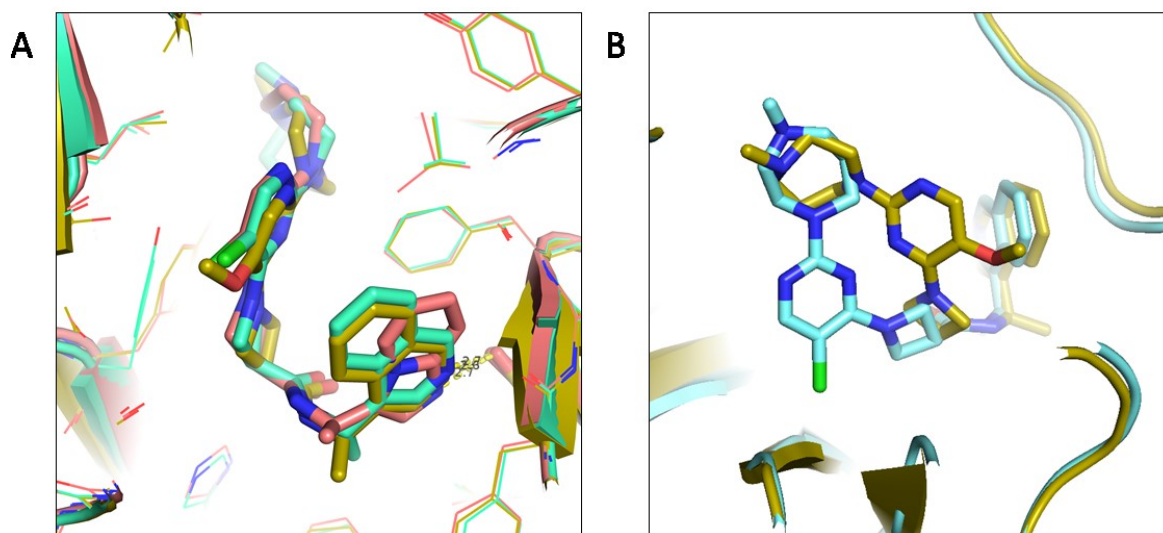


Figure 89 Binding modes of **IMP-0000366**, **IMP-0000974** and **IMP-0001169**. The ligand binding modes were analysed using PyMOL²⁸. **IMP-0000366** is shown in yellow, **IMP-0001169** in pink, **IMP-0000974** in S-shape in green and **IMP-0000974** in C-shape in blue. Hydrogen bonds with Ser405 are shown in yellow dashes with their lengths labelled in black. (A) Superimposed structures in S-shape; (B) Comparison of the **IMP-0000974** in C-shape and **IMP-0000366**.

IMP-0000366, **IMP-0000974** and **IMP-0001169** have comparable potencies. All of them have a bicyclic aromatic ring and one methyl substitution on the hydrophobic end of the ligands. When superimposing **IMP-0000974** in S-shape with **IMP-0000366** and **IMP-0001169**, they share similar conformations (Figure 89A). Although **IMP-0001169** has a different imidazo[1,2-a]pyridine on the hydrophobic end, it forms a hydrogen bond at similar length to Ser409. The hydrophilic ends all reach to the C-terminus. According to pKa analysis (default

setting) of Marvin Sketch¹²¹, at the human blood pH 7.4, 87.16% of **IMP-0000366**, 86.43% of **IMP-0000974** and 99.22% of **IMP-0001169** have a protonated 4*N* on the diazepane. Therefore, the hydrophilic ends form ionic bindings with the *C*-terminus.

IMP-0000366 and **IMP-0000974** only differ in the 5-substitution on the pyrimidine. As discussed in previous paragraphs, **IMP-0000974** appears in both conformations, but **IMP-0000366** only adopts the S-shape. Figure 89B shows the structure of **IMP-0000974** in C-shape superimposed with **IMP-0000366**. The 5-chlorine of **IMP-0000974** is pointing towards the protein backbone, where the bigger methoxy group of **IMP-0000366** could cause steric hindrance. The four-fold potency difference could be a result from **IMP-0000974**'s alternative conformation to bind, facilitating the compound to pose for binding.

3.3.3 Conclusion and outlook

Nine X-ray crystal structures of azetidinopyrimidines series were solved and analysed together with another two from the same series to explore the binding modes.

Azetidinopyrimidines series have two binding conformations: seven out of eleven structures were found in the S-shape, three out of eleven structures were found in the C-shape, and the other two were found in both shapes. All compounds with a big 5-substitution on the pyrimidine adopt the S-shape. The crystal structures also provide evidences of limited space for the 5-substitution group in the C-shape. Therefore, it is likely that compounds with big 5-substitutions on the pyrimidine can only adopt the S-shape. For compounds with a one-atom 5-substitution on the pyrimidine, the hypothesis is that they adopt both conformations at various percentage occupancies but only the structures with high resolutions clearly show both.

The hydrophobic end and the hydrophilic end of the ligand form two important interactions, which largely determine the potency. The hydrogen bonding between the hydrophobic end of

the ligand and the protein residue Ser405 is very sensitive to ligand conformation. One or two methyl substitutions on the one-carbon linker between the aromatic end and the amide can stabilize the ligand conformation and facilitate the formation of a strong interaction with Ser405. The three methyl substitutions on the pyrazole may have a similar conformational stabilization effect to keep good potencies without the methyl substitutions on the one carbon linker. The hydrophilic end of the ligand is highly protonated thus forming ionic bonds with the C-terminus. Because of the strong and non-directional nature of ionic bonding, the hydrophilic end is more tolerant with ligand conformations and structural modifications.

4 Conclusion and perspectives

N-myristoylation is the irreversible transfer of a myristate to the *N*-terminal glycine of protein substrates.

Based on established methods of NMT protein production and crystallography, twenty structures in two structure series of ternary complexes of Hs NMT1 bound with the ligand and Myr-CoA were solved with some adjustments. The binding mode insights revealed can facilitate the future drug discovery of NMT inhibitors, either to develop human NMT inhibitors or to avoid human NMT inhibition.

The high protein homology across NMTs from human, worm and parasite indicate the difficulty of achieving selective inhibition of specific NMT, but enabled the efficient target repurposing approach to explore NMT inhibitors of one organism from the existing inhibitors of other NMTs. A small-scale screen on eight parasite NMT inhibitors described here, identified several good nematode NMT inhibitors.

Guided by crystal structure insights, ten inhibitors with a new linker were synthesized based on a series of *Plasmodium* NMT inhibitors. The retention of the enzyme and cell activities of the best inhibitor as well as the maintenance of its binding mode comparing to its old linker analogue highlight the potential of potency improvement by linker changes. The tagging assay results prove direct evidence of its target engagement without affecting the overall protein synthesis in cells.

The CPM assay is a non-radioactive testing method used in many publications to obtain NMT enzyme activities and was conducted here for all IC₅₀ measurements. The abnormal inhibition curves found during the assay of the dihydrothiazole series were systematically investigated, providing an investigation protocol for similar problems. Inhibitor-CPM adducts were

identified and the formation mechanism was proposed, suggesting a limitation of the assay for testing strong nucleophiles.

With the aid of crystallography to reveal inhibitor binding mode insights, accumulated hit compounds from various drug discovery studies and high-throughput screens as starting points, established assay protocols to test inhibitors' *in vitro* and *in vivo* activities and to validate whether they are acting on target, this is a good time for medicinal chemists to develop human NMT inhibitors. The findings described here are built upon many previous studies and hopefully can also facilitate future NMT inhibitor development.

There is no doubt that human NMTs play import roles in many diseases. However, how to avoid side effects on normal human cells is the major problem faced to use them as drug targets. Due to the high similarity of the binding pockets between Hs NMT1 and Hs NMT2, it is unlikely to achieve selectivity at the enzymic level. The different urgencies of the need for *N*-myristoylation between normal cells and affected cells, such as cancer cells or virus-infected cells, creates a therapeutic window of several days without affecting the normal cells. It may suggest that NMT inhibitors could be used periodically in combination with other drugs, such as in a sequential chemotherapy program. Accurate drug delivery systems and suitable inhibitor pharmacokinetics can limit the inhibitor's in-body distribution and minimize its impact on normal cells.

5 Methods

5.1 Chemistry

5.1.1 General methods

All chemicals were purchased from Sigma-Aldrich Ltd (Gillingham, UK), Apollo Scientific (Stockport, UK), Acros Organics (Geel, Belgium) and Alfa Aesar (Heysham, UK) and used without further purification. Moisture sensitive reactions were performed under a nitrogen atmosphere using dried glassware, anhydrous solvents and standard syringe/septa techniques.

Compounds were purified by the Isolera (Biotage, UK) system using pre-packed SNAP silica cartridges (Biotage, UK). Elution method started with 1 column volume at x%, then a gradient from x% to y% for 15 column volumes and finishing with y% for 2 column volumes. Solvents, x and y will be defined in the characterization section.

Final compounds were purified by reverse phase LC-MS system Waters 2767 equipped with a photodiode array and an ESI mass spectrometer using an XBridge Prep C18 5 μm , 19 $\times$ 100 mm OBD column. Unless specified otherwise, all compounds were separated over a gradient of methanol in water (50–98% over 12 minutes then 98% methanol for 3 minutes), both containing 0.1% formic acid. The flow rate was 20 mL/min.

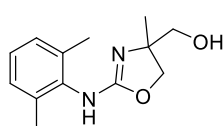
The purity of final compounds was verified by the Waters 2767 system using an XBridge Prep C18 5 μm , 4.6 $\times$ 100 mm OBD column over a gradient of methanol in water (5–98% over 12 minutes then 98% methanol for 3 minutes), both containing 0.1% formic acid. The flow rate was 1.2 mL/min.

^1H and ^{13}C NMR spectra were recorded at 400 MHz and 101 MHz respectively by Bruker AV instruments at room temperature unless specified otherwise and were referenced to residual

solvent signals. Data are presented as follows: chemical shift in ppm, integration, multiplicity (br = broad, app = apparent, s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet) and coupling constants in Hz.

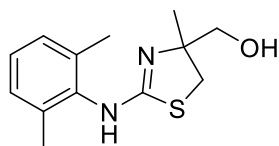
High-resolution mass spectra were obtained from the Mass Spectrometry Service of the Department of Chemistry, Imperial College London.

5.1.2 Synthesis and characterization of compounds



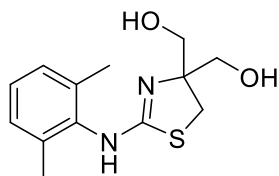
(+)/(-)-(2-((2,6-dimethylphenyl)amino)-4-methyl-4,5-dihydrooxazol-4-yl)methanol **20F4-O**

2-Isocyanato-1,3-dimethylbenzene (0.14 mL, 1.0 mmol) and 2-amino-2-methylpropane-1,3-diol (120 mg, 1.1 mmol) were added to ethyl acetate (5 mL) and the resulting solution was stirred at reflux overnight. Then the solvent was removed under reduced pressure. The residue was re-dissolved in 37% hydrochloric acid (5 mL) and heated under reflux at 95 °C for two hours. The mixture was cooled to room temperature and basified with 1 M aqueous sodium hydroxide (5 mL) in an ice bath. The resulting mixture was extracted with diethyl ether (5 mL) and washed with distilled water (5 mL x 2) followed by brine (5 mL). The organic layer was concentrated under reduced pressure and purified by column chromatography (2–10% MeOH in DCM) and LC-MS to give a white solid (14 mg, 6.0%). R_f = 0.5 (DCM/MeOH, 10:1); ^1H NMR (400 MHz, Chloroform- d) δ 8.57 (s, 1H), 7.79 (br, 1H), 7.13 (dd, J = 8.7, 6.0 Hz, 1H), 7.07 (d, J = 8.1 Hz, 2H), 4.64 (d, J = 8.5 Hz, 1H), 4.19 (d, J = 8.3 Hz, 1H), 3.75 (d, J = 11.9 Hz, 1H), 3.51 (d, J = 11.8 Hz, 1H), 2.27 (s, 6H), 1.44 (s, 3H); LC-MS R_t (retention time) = 8.31 min; HRMS found 235.1447 ($\text{C}_{13}\text{H}_{19}\text{N}_2\text{O}_2$, $[\text{M}+\text{H}]^+$, requires 235.1447).



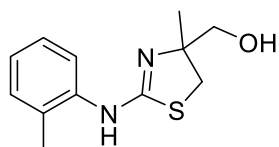
(+)/(-)-(2-((2,6-dimethylphenyl)amino)-4-methyl-4,5-dihydrothiazol-4-yl)methanol **20F4**

Compound was obtained from Eric Lee and purified by LC-MS to give a white solid. ^1H NMR (400 MHz, Chloroform-*d*) δ 10.21 – 9.82 (br, 1H), 8.52 (s, 1H), 7.14 – 7.01 (m, 3H), 3.70 (d, J = 11.3 Hz, 1H), 3.56 (d, J = 11.2 Hz, 1H), 3.46 (d, J = 11.2 Hz, 1H), 2.95 (d, J = 11.1 Hz, 1H), 2.24 (s, 6H), 1.45 (s, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 172.4, 170.4, 137.0, 135.3, 128.5, 128.3, 69.8, 66.3, 37.4, 22.8, 17.9, 17.7.



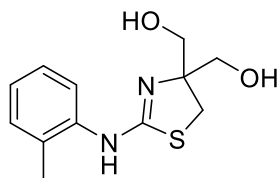
(2-((2,6-dimethylphenyl)amino)-4,5-dihydrothiazole-4,4-diyl)dimethanol **20F4-9**

Compound was obtained from Eric Lee and purified by LC-MS to give a white solid. ^1H NMR (400 MHz, Chloroform-*d*) δ 8.55 (br, 1H), 8.44 (s, 1H), 7.31 – 7.05 (m, 3H), 3.97 – 3.72 (m, 4H), 3.31 (s, 2H), 2.29 (s, 6H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 174.4, 169.4, 135.8, 135.6, 129.2, 128.8, 74.1, 62.8, 33.2, 17.9.



(+)/(-)-(4-methyl-2-(o-tolylamino)-4,5-dihydrothiazol-4-yl)methanol **20F4-12**

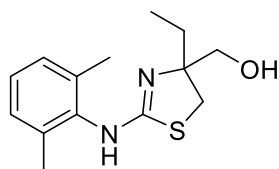
Compound was obtained from Eric Lee and purified by LC-MS to give a white solid. ^1H NMR (400 MHz, Chloroform-*d*) δ 8.59 (s, 1H), 8.24 (br, 1H), 7.30 – 7.12 (m, 4H), 3.72 (d, J = 11.4 Hz, 1H), 3.61 (d, J = 11.3 Hz, 1H), 3.44 (d, J = 10.8 Hz, 1H), 2.98 (d, J = 10.9 Hz, 1H), 2.30 (s, 3H), 1.46 (s, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 170.6, 169.4, 139.6, 133.5, 131.0, 127.3, 126.7, 124.8, 69.3, 66.7, 37.7, 22.8, 17.8.



(2-(o-tolylamino)-4,5-dihydrothiazole-4,4-diyl)dimethanol

20F4-13

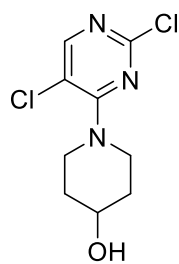
Compound was obtained from Eric Lee and purified by LC-MS to give a white solid. ¹H NMR (400 MHz, Chloroform-*d*) δ 8.50 (br, 1H), 8.47 (s, 1H), 7.38 – 7.15 (m, 4H), 3.88 – 3.79 (m, 4H), 3.32 (s, 2H), 2.32 (s, 3H). ¹³C NMR (101 MHz, Chloroform-*d*) δ 173.3, 169.8, 136.8, 134.7, 131.3, 128.8, 127.0, 126.0, 73.7, 62.8, 33.3, 17.7.



(+)(-)-(2-((2,6-dimethylphenyl)amino)-4-ethyl-4,5-dihydrothiazol-

4-yl)methanol 20F4-14

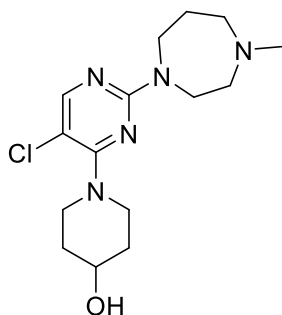
Compound was obtained from Eric Lee and purified by LC-MS to give a white solid. ¹H NMR (400 MHz, Chloroform-*d*) δ 8.58 (s, 1H), 8.21 (br, 1H), 7.15 (dd, *J* = 8.7, 6.3 Hz, 1H), 7.08 (d, *J* = 7.4 Hz, 2H), 3.80 (d, *J* = 11.5 Hz, 1H), 3.64 (d, *J* = 11.6 Hz, 1H), 3.43 (d, *J* = 11.1 Hz, 1H), 3.08 (d, *J* = 11.1 Hz, 1H), 2.29 (s, 6H), 1.94 – 1.73 (m, 2H), 1.08 (t, *J* = 7.5 Hz, 3H). ¹³C NMR (101 MHz, Chloroform-*d*) δ 170.5, 137.2, 135.3, 128.5, 128.3, 73.2, 65.8, 34.9, 29.0, 17.9, 8.0.



1-(2,5-dichloropyrimidin-4-yl)piperidin-4-ol

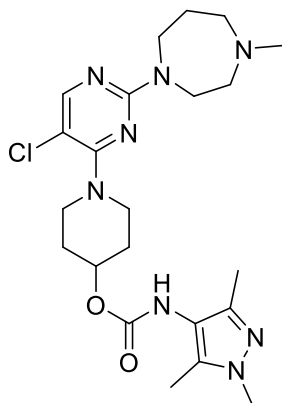
Piperidin-4-ol (1.5 g, 15 mmol) was stirred with triethylamine (18 mL, 15 mmol) in DMF (5 mL) for 10 min then 2,4,5-trichloropyrimidine (2.1 mL, 15 mmol) was added. The reaction was stirred overnight at room temperature. Water (10 mL) was then added to the suspension and extracted with ethyl acetate (10 mL). The organic layer was concentrated under reduced pressure and purified by column chromatography (20–50% ethyl acetate in

hexane) to give a light-yellow oil (2.9 g, 78%). $R_f = 0.4$ (ethyl acetate/hexane, 6 : 4); ^1H NMR (400 MHz, Chloroform- d) δ 8.02 (s, 1H), 4.14 (ddd, $J = 13.5, 6.7, 3.9$ Hz, 2H), 3.97 (tt, $J = 8.1, 4.0$ Hz, 1H), 3.45 (ddd, $J = 13.1, 9.0, 3.3$ Hz, 2H), 2.01 – 1.93 (m, 2H), 1.69 – 1.59 (m, 2H).



1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol

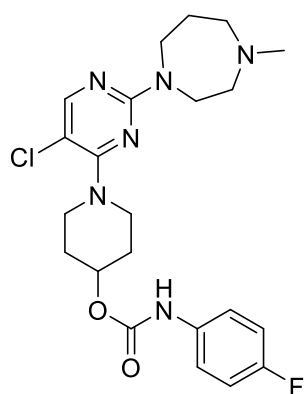
1-(2,5-dichloropyrimidin-4-yl)piperidin-4-ol (50 mg, 0.20 mmol) was added to *n*-BuOH (2 mL) followed by 1-methyl-1,4-diazepane (0.070 mL, 0.60 mmol). The suspension was heated by the microwave reactor (Biotage) for 1.5 hours at 200 °C. The suspension was then diluted with ethyl acetate (5 mL) and extracted with 10M NaOH (5mL). The organic layer was concentrated under reduced pressure and purified by column chromatography (0–5% MeOH in DCM) and LC-MS to give a yellow oil (54 mg, 78%). $R_f = 0.3$ (DCM/MeOH, 10:1); ^1H NMR (400 MHz, Chloroform- d) δ 7.92 (s, 1H), 4.07 (dt, $J = 13.5, 4.5$ Hz, 2H), 3.96 – 3.88 (m, 1H), 3.88 – 3.83 (m, 2H), 3.75 (t, $J = 6.4$ Hz, 2H), 3.25 – 3.14 (m, 2H), 2.72 – 2.64 (m, 2H), 2.62 – 2.53 (m, 2H), 2.38 (s, 4H), 2.03 – 1.93 (m, 4H), 1.63 (dtd, $J = 13.0, 9.2, 3.7$ Hz, 2H).



1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (1,3,5-trimethyl-1H-pyrazol-4-yl)carbamate IMP-0001096

1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (9.8 mg, 0.030 mmol) and 4-isocyanato-1,3,5-trimethyl-1H-pyrazole (4.5 mg, 0.030 mmol) in toluene (2 mL) were heated by the microwave reactor for 7

hours at 170 °C. The resulting suspension was washed with 10 M NaOH, water and brine then dried under reduced pressure. The resulting mixture was then purified by the LC-MS twice over a gradient of methanol in water, both containing 0.1% formic acid. The first time with 50–98% over 12 minutes then 98% methanol for 3 minutes, and the second time with 5–98% over 12 minutes then 98% methanol for 3 minutes to give a light-yellow oil (4.2 mg, 29%). ¹H NMR (400 MHz, Chloroform-*d*) δ 8.44 (s, 1H), 7.95 (s, 1H), 4.98 (tt, *J* = 8.3, 4.1 Hz, 1H), 4.09 (s, 3H), 4.00 (s, 2H), 3.90 – 3.77 (m, 4H), 3.73 (s, 3H), 3.46 – 3.31 (m, 2H), 3.22 (s, 2H), 3.14 (dd, *J* = 7.5, 3.1 Hz, 2H), 2.75 (s, 3H), 2.44 – 2.29 (m, 2H), 2.18 (s, 3H), 2.17 (s, 3H), 2.10 – 2.04 (m, 2H), 1.89 – 1.76 (m, 2H). ¹³C NMR (101 MHz, Chloroform-*d*) δ 159.3, 157.3, 71.1, 58.4, 57.3, 46.7, 46.2, 44.7, 36.4, 34.2, 30.9, 27.4, 11.1, 9.4. LC-MS *R*_t = 9.55 min; HRMS found 477.2516 (C₂₂H₃₄N₈O₂Cl, [M+H]⁺, requires 477.2493).

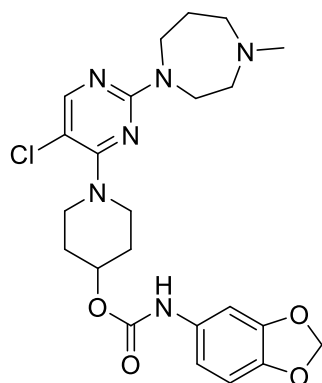


1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (4-fluorophenyl)carbamate IMP-0001125

Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (11 mg, 0.035 mmol) and 1-fluoro-4-

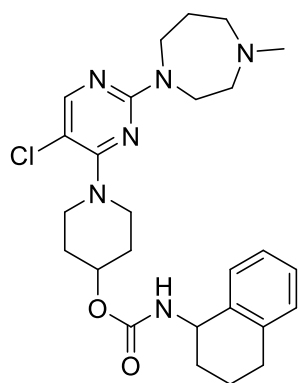
isocyanatobenzene (0.0039 mL, 0.035 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (8.3 mg, 65%). ¹H NMR (400 MHz, Chloroform-*d*) δ 8.48 (s, 1H), 7.96 (s, 1H), 7.48 – 7.31 (m, 2H), 7.10 – 6.97 (m, 2H), 5.03 (tt, *J* = 7.8, 3.7 Hz, 1H), 4.14 – 4.01 (m, 2H), 4.01 – 3.88 (m, 2H), 3.83 (t, *J* = 6.7 Hz, 2H), 3.50 – 3.37 (m, 2H), 3.23 – 3.13 (m, 2H), 3.13 – 3.03 (m, 2H), 2.72 (s, 3H), 2.38 – 2.26 (m, 2H), 2.15 – 2.01 (m, 2H), 1.83 (dtd, *J* = 12.6, 8.4, 3.5 Hz, 2H). ¹³C NMR (101 MHz, Chloroform-*d*) δ

167.6, 159.0, 157.5, 115.8, 115.6, 70.7, 57.1, 55.7, 45.0, 44.6, 44.4, 42.5, 30.9, 24.3. LC-MS R_t = 11.96 min; HRMS found 463.2030 ($C_{22}H_{29}N_6O_2ClF$, $[M+H]^+$, requires 463.2025).



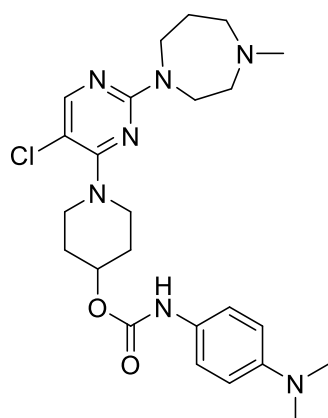
1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl benzo[d][1,3]dioxol-5-ylcarbamate **IMP-0001126**

Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (16 mg, 0.048 mmol) and 5-isocyanatobenzo[d][1,3]dioxole (7.9 mg, 0.048 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (9.5 mg, 41%). 1H NMR (400 MHz, Chloroform-*d*) δ 8.50 (s, 1H), 7.94 (s, 1H), 7.11 (s, 1H), 6.96 – 6.79 (m, 1H), 6.79 – 6.65 (m, 2H), 5.95 (s, 2H), 5.00 (tt, J = 7.9, 4.0 Hz, 1H), 4.14 – 4.00 (m, 2H), 4.00 – 3.87 (m, 2H), 3.82 (t, J = 6.6 Hz, 2H), 3.47 – 3.37 (m, 2H), 3.23 – 3.13 (m, 2H), 3.13 – 3.03 (m, 2H), 2.71 (s, 3H), 2.32 (tt, J = 9.1, 4.9 Hz, 2H), 2.14 – 1.98 (m, 2H), 1.81 (dtd, J = 12.4, 8.3, 3.5 Hz, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 167.9, 160.5, 159.0, 157.4, 148.0, 132.2, 108.1, 106.1, 101.3, 70.5, 57.1, 55.7, 44.9, 44.6, 44.4, 42.4, 30.9, 24.3. LC-MS R_t = 11.64 min; HRMS found 489.2005 ($C_{23}H_{30}N_6O_4Cl$, $[M+H]^+$, requires 489.2017).



1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (1,2,3,4-tetrahydronaphthalen-1-yl)carbamate **IMP-0001127**

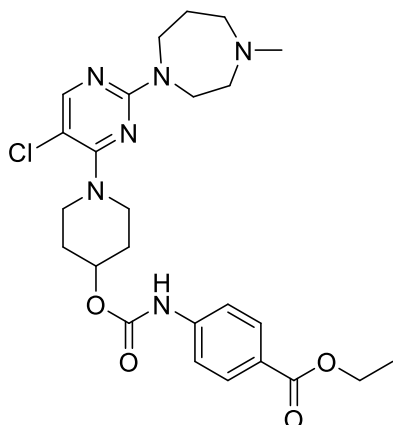
Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (25 mg, 0.077 mmol) and 1-isocyanato-1,2,3,4-tetrahydronaphthalene (0.012 mL, 0.077 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (3.4 mg, 8.8%). ¹H NMR (400 MHz, Chloroform-*d*) δ 8.47 (s, 1H), 7.95 (s, 1H), 7.25 (s, 1H), 6.73 (d, *J* = 8.8 Hz, 2H), 6.49 (s, 1H), 5.11 – 4.94 (m, 1H), 4.08 (s, 2H), 3.93 (s, 2H), 3.83 (t, *J* = 6.7 Hz, 2H), 3.50 – 3.36 (m, 2H), 3.26 – 3.16 (m, 2H), 3.16 – 3.06 (m, 2H), 2.93 (s, 6H), 2.73 (s, 3H), 2.41 – 2.27 (m, 2H), 2.13 – 1.99 (m, 2H), 1.82 (dtd, *J* = 12.3, 8.3, 3.5 Hz, 2H). ¹³C NMR (101 MHz, Chloroform-*d*) δ 167.9, 160.5, 159.1, 157.4, 155.4, 137.5, 129.2, 128.6, 127.4, 126.2, 105.8, 70.1, 57.3, 55.9, 49.2, 45.2, 44.8, 44.6, 43.2, 31.0, 30.4, 29.2, 25.0, 19.8. LC-MS *R*_t = 12.53 min; HRMS found 499.2570 (C₂₆H₃₆N₆O₂Cl, [M+H]⁺, requires 499.2588).



1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (4-(dimethylamino)phenyl)carbamate **IMP-0001128**

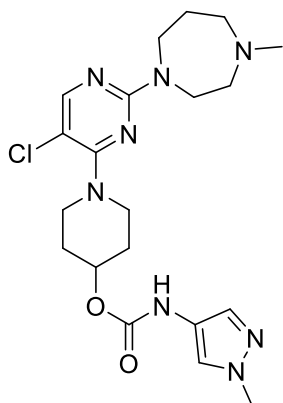
Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (13 mg, 0.040 mmol) and 4-isocyanato-*N,N*-dimethylaniline (6.3 mg, 0.040 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (11 mg, 62%). ¹H NMR (400 MHz, Chloroform-*d*) δ 8.53 (s, 1H), 8.08 – 8.00 (m, 2H), 7.96 (s, 1H), 7.53 – 7.46 (m, 2H), 7.00 (s, 1H), 5.05 (tt, *J* = 8.0, 3.9 Hz, 1H), 4.44 – 4.31 (m, 2H), 4.12 – 4.01 (m, 2H), 4.01 – 3.89 (m, 2H), 3.83 (t, *J* = 6.6 Hz, 2H), 3.50 – 3.38 (m, 2H), 3.12 (t, *J* = 4.8 Hz, 2H), 3.08 – 2.99 (m, 2H), 2.69 (s, 3H), 2.38 – 2.27 (m, 2H), 2.15 – 2.04 (m, 2H), 1.84 (dtd, *J* = 12.3, 8.3, 3.5 Hz, 2H), 1.41 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, Chloroform-*d*) δ 167.4, 159.0, 157.4, 113.4, 106.1, 57.1, 55.6, 44.9, 44.6, 44.3, 42.3,

41.0, 30.9, 24.2. LC-MS R_t = 8.93 min; HRMS found 488.2549 ($C_{24}H_{35}N_7O_2Cl$, $[M+H]^+$, requires 488.2541).



Ethyl 4-((((1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl)oxy)carbonyl)amino)benzoate **IMP-0001129**

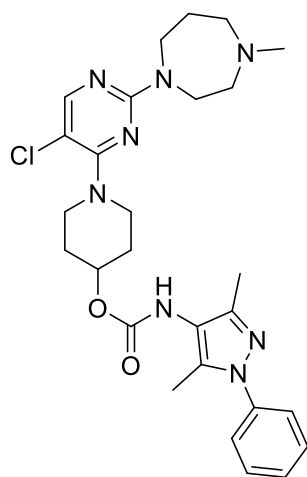
Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (7.7 mg, 0.020 mmol) and ethyl 4-isocyanatobenzoate (4.5 mg, 0.020 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (8.7 mg, 84%). 1H NMR (400 MHz, Chloroform- d) δ 8.45 (s, 1H), 7.95 (s, 1H), 7.64 (s, 1H), 7.37 (s, 1H), 4.99 (tt, J = 7.6, 3.7 Hz, 1H), 4.14 – 4.01 (m, 2H), 3.97 – 3.89 (m, 2H), 3.87 (s, 3H), 3.86 – 3.78 (m, 2H), 3.48 – 3.36 (m, 2H), 3.22 (app t, J = 4.6 Hz, 2H), 3.17 – 3.08 (m, 2H), 2.74 (s, 3H), 2.40 – 2.27 (m, 2H), 2.13 – 1.98 (m, 2H), 1.80 (dtd, J = 12.4, 8.1, 3.4 Hz, 2H). ^{13}C NMR (101 MHz, Chloroform- d) δ 168.1, 159.0, 157.5, 130.9, 117.6, 71.1, 60.9, 56.0, 45.1, 44.7, 44.6, 42.9, 30.8, 24.7, 14.4. LC-MS R_t = 12.56 min; HRMS found 517.2317 ($C_{25}H_{34}N_6O_4Cl$, $[M+H]^+$, requires 517.2330).



1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (1-methyl-1H-pyrazol-4-yl)carbamate **IMP-0001212**

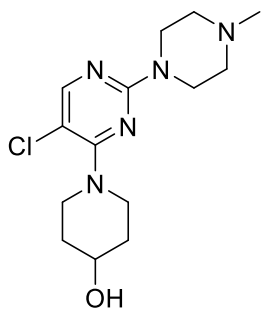
Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (27 mg, 0.082 mmol) and 4-isocyanato-1-methyl-1H-pyrazole (12 mg, 0.098 mmol), the compound was synthesized and purified by the same

method as **IMP-0001096** to give a yellow oil (10 mg, 28%). ^1H NMR (400 MHz, Chloroform-*d*) δ 8.45 (s, 1H), 7.95 (s, 1H), 7.64 (s, 1H), 7.37 (s, 1H), 4.99 (tt, $J = 7.6, 3.7$ Hz, 1H), 4.14 – 4.01 (m, 2H), 3.97 – 3.89 (m, 2H), 3.87 (s, 3H), 3.86 – 3.78 (m, 2H), 3.48 – 3.36 (m, 2H), 3.28 – 3.17 (m, 2H), 3.17 – 3.08 (m, 2H), 2.74 (s, 3H), 2.40 – 2.27 (m, 2H), 2.13 – 1.98 (m, 2H), 1.80 (dtd, $J = 12.4, 8.1, 3.4$ Hz, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 167.3, 160.6, 158.9, 157.4, 153.3, 130.3, 121.4, 121.2, 106.2, 70.7, 57.1, 55.6, 44.9, 44.6, 44.3, 42.3, 39.3, 30.9, 24.2. LC-MS $R_t = 9.75$ min; HRMS found 449.2176 ($\text{C}_{20}\text{H}_{30}\text{N}_8\text{O}_2\text{Cl}$, $[\text{M}+\text{H}]^+$, requires 449.2180).



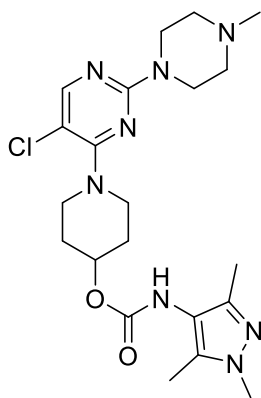
1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (3,5-dimethyl-1-phenyl-1H-pyrazol-4-yl)carbamate **IMP-0001459**

Using 1-(5-chloro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (39 mg, 0.12 mmol) and 4-isocyanato-3,5-dimethyl-1-phenyl-1H-pyrazole (76 mg, 0.36 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (30 mg, 46%). ^1H NMR (400 MHz, Chloroform-*d*) δ 8.38 (s, 1H), 7.97 (s, 1H), 7.53 – 7.42 (m, 5H), 7.42 – 7.32 (m, 1H), 5.09 – 4.93 (m, 1H), 4.19 – 4.06 (m, 2H), 4.06 – 3.97 (m, 2H), 3.89 – 3.78 (m, 3H), 3.49 – 3.33 (m, 2H), 3.31 – 3.21 (m, 2H), 3.21 – 3.12 (m, 2H), 2.78 (s, 3H), 2.38 (s, 2H), 2.32 – 2.21 (m, 6H), 2.15 – 2.04 (m, 2H), 1.91 – 1.75 (m, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 167.2, 159.0, 157.4, 139.8, 135.7, 129.1, 127.5, 124.6, 77.4, 77.0, 76.7, 57.3, 55.8, 44.9, 44.7, 44.6, 42.4, 30.9, 24.3, 11.3, 11.2, 10.7. LC-MS $R_t = 11.43$ min; HRMS found 539.2661 ($\text{C}_{27}\text{H}_{36}\text{N}_8\text{O}_2\text{Cl}$, $[\text{M}+\text{H}]^+$, requires 539.2650).



1-(5-chloro-2-(4-methylpiperazin-1-yl)pyrimidin-4-yl)piperidin-4-ol

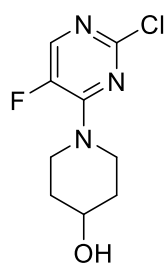
1-(2,5-dichloropyrimidin-4-yl)piperidin-4-ol (740 mg, 3.0 mmol) was added to *n*-BuOH (4 mL) followed by 1-methylpiperazine (1.0 mL, 9.0 mmol). The suspension was heated by the microwave reactor (Biotage) for 1.5 hours at 200 °C. The suspension was diluted with ethyl acetate and extracted with 10M NaOH. The organic layer was concentrated under reduced pressure and purified by column chromatography (1–10% MeOH in DCM) and LC-MS to give a yellow oil (420 mg, 45%). R_f = 0.4 (DCM/MeOH, 10:1); ^1H NMR (400 MHz, Chloroform-*d*) δ 8.02 (s, 1H), 4.19 – 4.10 (m, 2H), 3.98 (tt, J = 8.1, 3.9 Hz, 1H), 3.50 – 3.40 (m, 2H), 2.01 – 1.93 (m, 3H), 1.69 – 1.59 (m, 3H).



1-(5-chloro-2-(4-methylpiperazin-1-yl)pyrimidin-4-yl)piperidin-4-yl
(1,3,5-trimethyl-1H-pyrazol-4-yl)carbamate IMP-0001460

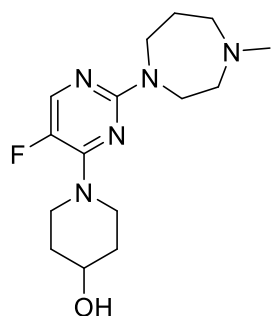
Using 1-(5-chloro-2-(4-methylpiperazin-1-yl)pyrimidin-4-yl)piperidin-4-ol (46 mg 0.15) and 4-isocyanato-1,3,5-trimethyl-1H-pyrazole (66 mg, 0.45 mmol), the compound was synthesized and purified by the same method as **IMP-0001096** to give a yellow oil (13 mg, 21%). ^1H NMR (400 MHz, Chloroform-*d*) δ 7.87 (s, 1H), 4.03 (dt, J = 13.4, 4.3 Hz, 2H), 3.85 (tt, J = 8.5, 4.0 Hz, 1H), 3.78 – 3.64 (m, 4H), 3.23 – 3.09 (m, 2H), 2.41 (t, J = 5.1 Hz, 4H), 2.29 (s, 3H), 1.97 – 1.86 (m, 2H), 1.58 (dtd, J = 12.9, 9.3, 3.7 Hz, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 159.2, 157.4, 53.9, 44.8, 44.7, 42.7, 36.3, 30.9, 11.0,

9.3. LC-MS $R_t = 10.27$ min; HRMS found 463.2332 ($C_{21}H_{32}N_8O_2Cl$, $[M+H]^+$, requires 463.2337).



1-(2-chloro-5-fluoropyrimidin-4-yl)piperidin-4-ol

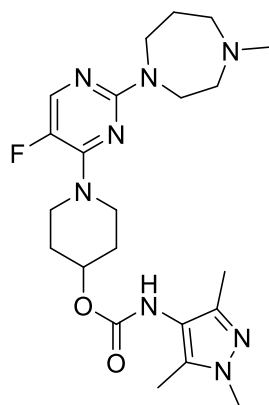
Piperidin-4-ol (1.5 g, 15 mmol) was stirred with triethylamine (18 mL, 15 mmol) in DMF (5 mL) for 10 min then 2,4-dichloro-5-fluoropyrimidine (1 g, 6 mmol) was added. The reaction was stirred overnight at room temperature. Water was then added dropwise to the suspension and extracted with ethyl acetate. The organic layer was concentrated under reduced pressure and purified by column chromatography (30–70% ethyl acetate in hexane) and LC-MS to give a light-yellow oil (0.95 g, 68.8%). $R_f = 0.1$ (Ethyl acetate/Hexane, 6 : 4); 1H NMR (400 MHz, Chloroform- d) δ 7.91 (d, $^2J_{C-F} = 6.3$ Hz, 1H), 4.25 – 4.11 (m, 2H), 4.04 (tt, $J = 7.9$, 3.9 Hz, 1H), 3.61 – 3.45 (m, 2H), 2.08 – 1.90 (m, 3H), 1.65 (dtd, $J = 12.4$, 8.3, 3.7 Hz, 2H).



1-(5-fluoro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol

1-(2-chloro-5-fluoropyrimidin-4-yl)piperidin-4-ol (707 mg, 3.00 mmol) was added to *n*-BuOH followed by 1-methyl-1,4-diazepane (1.1 mL, 9.0 mmol). The suspension was heated in a microwave reactor (Biotage) for 1.5 hours at 200 °C. The suspension was then diluted with ethyl acetate and extracted with 10M NaOH. The organic layer was concentrated under reduced pressure and purified by column chromatography (0–5% MeOH in DCM) and LC-MS to give a yellow oil (81.7 mg, 8.8%). $R_f = 0.1$ (DCM/MeOH, 10:1); 1H NMR (400 MHz, Chloroform- d) δ 8.43 (s, 1H), 7.71 (d, $^2J_{C-F} = 6.7$ Hz, 1H), 4.10 (dt,

$J = 13.6, 4.5$ Hz, 2H), 4.02 – 3.92 (m, 2H), 3.88 (tt, $J = 8.4, 4.0$ Hz, 1H), 3.74 (t, $J = 6.6$ Hz, 2H), 3.31 – 3.15 (m, 4H), 3.15 – 3.05 (m, 2H), 2.71 (s, 3H), 2.26 (dt, $J = 10.6, 6.5$ Hz, 2H), 1.96 – 1.83 (m, 2H), 1.53 (ddt, $J = 13.6, 9.3, 4.6$ Hz, 2H)..



1-(5-fluoro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-yl (1,3,5-trimethyl-1H-pyrazol-4-yl)carbamate **IMP-0001461**

Using 1-(5-fluoro-2-(4-methyl-1,4-diazepan-1-yl)pyrimidin-4-yl)piperidin-4-ol (29 mg, 0.10 mmol) and 4-isocyanato-1,3,5-trimethyl-1H-pyrazole (45 mg, 0.30 mmol), the compound was synthesized and purified by the same

method as **IMP-0001096** to give a yellow oil (18 mg, 38%). ^1H NMR (400 MHz, Chloroform-*d*) δ 8.46 (s, 1H), 7.82 (d, $^2J_{\text{C-F}} = 6.8$ Hz, 1H), 4.99 (dt, $J = 8.5, 4.3$ Hz, 1H), 4.09 – 3.99 (m, 3H), 3.87 – 3.76 (m, 2H), 3.72 (s, 3H), 3.53 – 3.38 (m, 2H), 3.27 – 3.17 (m, 2H), 3.17 – 3.05 (m, 2H), 2.74 (s, 3H), 2.40 – 2.27 (m, 2H), 2.17 (s, 3H), 2.16 (s, 3H), 2.11 – 2.00 (m, 2H), 1.85 – 1.69 (m, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 167.2, 143.4, 143.2, 57.2, 55.6, 45.1, 44.2, 42.4, 36.3, 31.0, 24.3, 11.0. LC-MS $R_t = 7.81$ min; HRMS found 461.2782 ($\text{C}_{22}\text{H}_{34}\text{N}_8\text{O}_2\text{Cl}$, $[\text{M}+\text{H}]^+$, requires 461.2789).

5.2 Biology

5.2.1 General methods

All reagents and solvents were purchased from Sigma-Aldrich, NovaBiochem or VWR and used without further purification unless stated otherwise. Culture media for cell culture was obtained from Sigma Aldrich, Gibco (Life technologies) or A&E Scientific (PAA). Ultrapure water was obtained through filtration with Milli-Q[®] Millipore water purification system. In-gel fluorescence was recorded using a Typhoon FLA 9500 Imager (GE Healthcare). Absorbance in 96-well plates was measured using a SpectraMax M2/M2e Microplate Reader (Molecular Devices) or an EnVision Xcite Plate Reader (PerkinElmer).

BL-41 was obtained from Gregor Lueg (Francis Crick Institute/Imperial College London). HeLa was obtained from *Dr* David Mann's group (Imperial College London). YnC12 was synthesized by *Dr* Yunyun Sun. The capture reagents TAMRA azide was kindly synthesized and provided by *Dr* Julia Morales-Sanfrutos. **DDD85646** was synthesized by *Dr* William Heal.

5.2.2 Sequence alignment

Protein sequences were obtained from UniProtKB:

```
>tr|D0A003|D0A003_TRYB9 Glycyl peptide N-tetradecanoyl transferase OS=Trypanosoma brucei gambiense (strain MHOM/CI/86/DAL972) GN=TbgDal_X16640 PE=3 SV=1
MTDKAFTEHQFWSTQPVRQPGAPDADKVGFIEMSSLDVPAEPYSLPSTFEWWSPDVANPE
DLRGVHELLRDNYVEDSESMFRFNYSEEFRLWALMPPGYHQSWHVGVRKSNKSVLGFVA
GVPITMRLGTPKMVLEKREHGEGGEEVINDYLEPQTICEINFLCVHKKLRQRRLLGPILIKEVT
RRVNLNMNIWHAVYTS GTLLPTPFAKGHYFHRSLNSQKLVDVKFSGIPPHYKRFQNPVAVME
RLYRLPDKTKTRGLRLMEPADVPQVTQLLKRLASFDVAPVFNEEEVAHYFLPREGVVFSYV
VESPVGPGKDEENAGKTSKGTPTGTKCVTGGCEKVITDFFSFYSLPSTIIGNSNHSLLKVAYV
YYTAATSVSITQLVNDLLIIVKLNGFDVCNVVDIYDNGTYLKEKLFSPGDGNLYYYFYNWSY
PSIPANEVGLVMV
```

```
>tr|Q4Q5S8|Q4Q5S8_LEIMA Glycyl peptide N-tetradecanoyl transferase OS=Leishmania major
GN=NMT PE=1 SV=1
MSRNPSNSDAAHAFWSTQPVPQTEDETEKIVFAGPMDEPKTVADIPEEPYPIASTFEWWTPN
MEAADDIHAIYELLRDNYVEDDDSMFRFNYSEEFRLQWALCPPNYIPDWHVAVRRKADKLL
AFIAGVPVTLRMGTPKYMVKVKAQEKGEGEAAKYDEPRHICEINFLCVHKQLREKRLAPILIK
EATRRVNRTNVWQAVYTAGVLLPTPYASGQYFHRSLNPEKLVEIRFSGIPAQYQKFQNPMA
```

MLKRNYQLPSAPKNSGLREMKPSDVPQVRRILMNYLDSFDVGPVFSDAEISHYLLPRDGVVF
TYVVENDKKVTDFFSFYRIPSTVIGNSNYNLLNAAAYVHYAATSIPLHQLLDLLIVAHSRGFD
VCNMVEILDNRSFVEQLKFGAGDGHLYFYFNWAYPKIKPSQVALVML

>tr|A5K1A2|A5K1A2_PLAVS Glycyl peptide N-tetradecanoyl transferase OS=*Plasmodium vivax*
(strain Salvador I) GN=PVX_085815 PE=1 SV=1

MNDDNKEFSGRDIYQLIKNAKDKIKIDYKFWYTQPVPKINDEFNESPNEPFISDNKVEDVRKD
EYKLPPGYSWYVCDVKDEKDRSEIYTLLTDNYVEDDDNIFRFNYSAEFLLWALTSPNYLKT
WHIGVKYDASNKLIGFISAIPTDICIHKRTIKMAEVNFLCVHKTLRSKRLAPVLIKEITRRINLE
NIWQAIYTAGVYLPKPVSDARYYHRSINVKKLIEIGFSSLNSRLTMSRAIKLYRVEDTLNIKN
MRLMKKKDVEGVHKLGSYLEQFNLYAVFTKEEIAHWFLPIENVITYYVNEENGKIKDMISF
YSLPSQILGNDKYSTLNAAYSFYNTTTATFKQLMQDAILLAKRNNFDVFNALVMDNYSVFDLQ
EDLKFGEGDGS�KYLYNWKASFAHVGIVLL

>sp|Q8ILW6|NMT_PLAF7 Glycyl peptide N-tetradecanoyl transferase OS=*Plasmodium falciparum*
(isolate 3D7) GN=PF14_0127 PE=3 SV=1

MNDDKKDFVGRDLYQLIRNAKDKIKIDYKFWYTQPVPKINDEFDENVNEPFISDNKVEDVRK
EYKLPSGYAWCVCDITKENDRSDIYNLLTDNYVEDDDNVFRFNYSSEFLLWALSSPNYVKN
WHIGVKYESTNKLVGFIAPIDMCVNKNIKMAEVNFLCVHKSLSKRLAPVLIKEITRRINLE
SIWQAIYTAGVYLPKPISTARYFHRSINVKKLIEIGFSLNTRLTMSRAIKLYRIDDTLNIKNLR
LMKKKDIDGLQKLLNEHLKQYNLHAIFSKEDVAHWFTPIDQVIYTYVNEENGEIKDLISFYSL
PSKVLGNKNYINLNAAFSYNITTTTTFKNLIQDAICLAKRNNFDVFNALVMDNYSVFDLQ
FGEGDGS�KYLYNWKASCHPSKIGIVLL

>sp|P46548|NMT_CAEL Probable glycyl peptide N-tetradecanoyl transferase OS=*Caenorhabditis elegans*
GN=nmt-1 PE=3 SV=1

MSHGSHDGAAPCGGHHGDDGAGGSRPSVNDVQALVDQLRLAGVDVSNMPNIPTAPRDMDE
ARSKSFQFWSQPVPQMDETVPADVNCIAEENIALDKVRAEPFSLPAGFRWSNVLDLSDDEQL
NELYNLLTRNYVEDDDSMFRFDYSADFLKWALQVPGFRPEWHCGVRADSNRLLAFIGAVP
QTVRVYDKTVNMVEINFLCVHKNLRSRRVAPVLIREITRRVNTGIFQAAFTAGIVIPKPVSV
CRYYHRSNLNPRKLIDVRFSLSAKMTMARTIKLYKLPEETATRNLRMKSTDVPQVFKLLTT
SLKQYSLAPVYNSEELAHALVPKKGVVYSYVAENQNGKITDFVSFYSLPSTVMGHTTHKTI
YAAALYYYVAGSVTPKQLINDSLILANREKFDVFNALDLMHNEKIFSDLKFGKGDNLQYY
LYNWKCADMKPSQIGLVLQ

>tr|A0A0H5SBS7|A0A0H5SBS7_BRUMA Glycyl peptide N-tetradecanoyl transferase OS=*Brugia malayi*
GN=bma-nmt-1 PE=3 SV=1

MKEPPAGNIKLDDEDEKMMAGNSKMDQNNDKENDAGDESIQNVTGRNSPTAVIDLKELT
ESTLMKKFEMLTVGGTSAAKHITEAQRHKYLFWDTPVPKINEMVTENRAIEPPLDISEVREE
PFSLPDPFCWCIEINSVKELTELYTLLTENYVEDDDNMFRFDYSPEFLLWALKAPGWMKNW
HCGVRAKSNGKLIAFISAIPSVIRVYDKQIKMVEINFLCVHKKLSKRVAPVLIREITRRVNRE
GIFQAVFTAGVVLPKPIATCRYWHRSLNPKKLVKFSHLSRKMTMQRTLKLYKLDPHPRTA
NLVPMKKCHIDGAYGLLQCYLKKFDLSPQFTRADFEHFFMPREDVIYSYVALNEEDSKVSDL
ISFYSLPSSVMHHPQYKSIRAAYSFYNVATSVTLKQLINDALILARNCGFDVFNALDLMNDKE
ILEDLKFGIGDGNLQYYLYNWKCPDIIPEKIGLVLQ

>sp|P30419|NMT1_HUMAN Glycyl peptide N-tetradecanoyl transferase 1 OS=*Homo sapiens*
GN=NMT1 PE=1 SV=2

MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGSPANDTGAKKKKKK
QKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQAIELFSVGQPAKTMEASKRSYQFWD
TPVPKLGEVVNTHGPVEPDKNIRQEPYTLQGFQWTDALDLGDRGVKELYTLLNENYVED
DDNMFRFDYSPEFLLWALRPPGWLPQWHCGVRVSSRKLVGFIAPANIHIYDTEKKMVEI
NFLCVHKKLSKRVAPVLIREITRRVHLEGIFQAVYTAGVVLPKPVGTCRYWHRSLNPRKLIE
VKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVHQLLTRYLKQFHLTPVMSQE

EVEHWFYPQENIIDTFVVENANGEVTDFLSFYTLPLSTIMNHPTHKSLKAAYSFYNVHTQTPLL
DLMSDALVLAKMKGFDVFNALDLMENKTFLEKLKFGIGDGNLQYYLYNWKCPSMGAEKV
GLVLQ

>sp|O60551|NMT2_HUMAN Glycyl peptide N-tetradecanoyl transferase 2 OS=*Homo sapiens*
GN=NMT2 PE=1 SV=1

MAEDSESAASQQSLELDDQDTCGIDGDNEEETEHA KGSPGGYLGAKKKKKKQKRKKEKPNS
GGTKSDSASDSQEIKIQPSKNPSVPMQKLQDIQRAMELLSACQGPARNIDEAAKHRYQFWD
TQVPVKLDEVITSHGAIEPDKNVRQEPYSLPQGF MWDTLDLSDAEVLKELYTLLNENYVED
DDNMFRFDYSPEFLWLALRPPGWLLQWHCGVRVSSNKKLVGFISAIPANIRIYDSVKKMVEI
NFLCVHKKLRSKRVAPVLIREITRRVNLEGIFQAVYTAGVVL PKPIATCRYWHRSLNPRKLVE
VKFSHL SRNMTLQRTMKLYRLPDVTKTSGLRPMEPKDIKSVRELINTY LKQFHLAPVMDEEE
VAHWFLPREHIIDTFVVESPNGKLTDFLSFYTLPLSTMHHPAHKSLKAAYSFYNIHTETPLLDL
MSDALILAKSKGFDVFNALDLMENKTFLEKLKFGIGDGNLQYYLYNWRCPGTDSEKVGLVL
Q

>The modified *Homo sapiens* NMT1 without *N*-terminus

MGSSHHHHHHSSGLEVL FQGPHMEEASKRSYQFWD TQVPVKLGEVVNTHGPVEPDKNIRQ
EPYTLPQGF TWDALDLGDRGV LKELYTLLNENYVEDDDNMFRFDYSPEFLWLALRPPGWLP
QWHCGVRVSSRKL VGFISAIPANIHIYDTEKKMVEINFLCVHKKLRSKRVAPVLIREITRRV
HLEGIFQAVYTAGVVL PKPVGTCRYWHRSLNPRKLIEVKFSHL SRNMTMQRTMKLYRLPETP
KTAGLRPMETKDIPVVHQLLTRYLKQFHLTPVMSQEEVEHWFYPQENIIDTFVVENANGEVT
DFLSFYTLPLSTIMNHPTHKSLKAAYSFYNVHTQTPLL DLMSDALVLAKMKGFDVFNALDLM
ENKTFLEKLKFGIGDGNLQYYLYNWKCPSMGAEKVGLVLQ

The protein sequences were aligned by Uniplot¹³⁸ and displayed by BoxShade¹⁴³. The alignment result of the eight NMTs was as follows:

Tb	1	-----
Lm	1	-----
Pv	1	-----
Pf	1	-----
Ce	1	-----MSHGHSHDGAPC-----
Bm	1	-----MKEPPAG----NIKLDDEDEKMM-AAG-----NSKM
Hs1	1	MADESETAVKPPAPPLPQMMEGNGNGHEHCSD-CENEEDNSYNRGGLSPANDTGAKKKKK
Hs2	1	MAEDSESAASQQSLEL-----DDQDTCGIDGDNEEET--EHAKGSPGGYLGAKKKKK

Tb	1	-----
Lm	1	-----M
Pv	1	-----MNDNKEFSGRDI-----YQLI
Pf	1	-----MNDKKDFVGRDL-----YQLI
Ce	13	-----GGHHGDDGA-----GGSRPS-VNDVQALVDQLRLAGVDVSNMPNIPTAP
Bm	27	DQN--NDKENDAGDESIQNV----TGRNSPTAVIDLKELTSTLMKKFEMLT-VGGTSAA
Hs1	60	KQKKKKEKG---SETDS-----AQDQPVKMNSLPAERIQEIQKAIELFS--VGQGPA
Hs2	51	KQKRKKEKPNSGGTKSDSASDSQEIKIQPSKNPSVPMQKLQDIQRAMELLS--ACQGPA

Tb	1	--MTLKAFTEHQFWSTQVPVRQPGAPDAKVGF----IMESSLDAPPAEPYSLPSTFEWWS
Lm	2	SRNPSNSDAAHAFWSTQVPVQTEDETEK-IVFAGPMDEPKTVADIPEEPYPIASTFEWWT
Pv	18	RNAKDKIKIDYKFWYTQVPVKINDEFNES--VNEPFISDNKVEDVRKDEYKLPPGYSWYV
Pf	18	RNAKDKIKIDYKFWYTQVPVKINDEFDEN--VNEPFISDNKVEDVRKEEYKLPSGYAWCV
Ce	56	RDMDEARSKSHQFWSTQVPVQMDETVPAD--VNCAIEENIALDKVRAEPFSLPAGFRWSN
Bm	80	KHTEAQREHKYLFWD TQVPVKINEMVTE-----NRAIEPPLDISVREEPFSLPDPFCWCD
Hs1	107	KTMEEASKRSYQFWD TQVPVKLGEVNT----HGPVEPDK--DNIRQEPYTLPQGF TWDA
Hs2	109	RNIDEAAKHRYQFWD TQVPVKLDEVITS----HGAIPEPK--DNVRQEPYSLPQGF MWDT

Tb	55	PDMANPEDLFGVHELLRDNYVEDSSMFRFNYSSEFLRWALMPPGYHQSWHVGVRCLKSNK
Lm	61	PNMEAADLHAIFYELLRDNYVEDDDSMFRFNYSSEFLQWALCPPNYTPDWHVAVRRKADK
Pv	76	CDVKDEKDRSEITYTLTDNYVEDDDNIFRFNYSSEFLRWALTSPNYLKTWHIGVKYDASN
Pf	76	CDITKENDRSDIYNLLTDNYVEDDDNIFRFNYSSEFLRWALSSPNYVKNWHIGVKYESTN
Ce	114	VDSLDEEQINELYNNLRTNYVEDDDSMFRFDYSADFLKWALQVPGFRPEWHCGVRADSNN
Bm	136	IEINSVKELTELYTLLTENYVEDDDNMFRFDYSPEFLLWALKAPGWMKNWHCGVRAKSNG
Hs1	161	LDLGDRGVLELYTLLTENYVEDDDNMFRFDYSPEFLLWALRPPGWLQWHCGVRVVSRR
Hs2	163	LDLSDAEVLKELYTLLTENYVEDDDNMFRFDYSPEFLLWALRPPGWLQWHCGVRVSSNK

Tb	115	SVIGFVAGVPIITMRIGTPKMVLEKREHGEGGEEVINDYLEPQTIICEINFLCVHKKLRQR
Lm	121	KLLAFIAGVPVITMRIGTPKYMKVKAQE--KGEGEAAKYDEPRHICEINFLCVHKQLREK
Pv	136	KLVGFISAIPTDICIT-----HKRTIKMAEVNFLCVHKTLRSK
Pf	136	KLVGFISAIPIDMCV-----NKNIIKMAEVNFLCVHKSLRSK
Ce	174	RLAFIAGVPQTVRV-----YDKTVNMVEINFLCVHKNLRSR
Bm	196	KLIAFISAIPSVIRV-----YDKQIKMVEINFLCVHKKLRSK
Hs1	221	KLVGFISAIPANIHI-----YDTEKKMVEINFLCVHKKLRSK
Hs2	223	KLVGFISAIPANIRI-----YDSVKKMVEINFLCVHKKLRSK

Tb	175	RLGPILLIKENTRRVNLNMIHNAVYTSGLTLPPTFAKGHYFHRSLSNSOKLVDVKFSGIPPH
Lm	179	RLAPILLIKEATRRVNRNTNVOAVYTAGVILPTPYASGOYFHRSLNPEKLVEIRFSGIPAQ
Pv	173	RLAPVLIKEITRRINLENINQAIYTAGVYLPKPVSDARYYHRSINVKKLIEIGFSSLSNR
Pf	173	RLAPVLIKEITRRINLESINQAIYTAGVYLPKPISTARYFHRSINVKKLIEIGFSCLNTR
Ce	211	RVAPVLIKEITRRVNLNMGIFQAATAGVILPKPVSVCRYYHRSLNPRKLIDVRFSHLSAK
Bm	233	RVAPVLIKEITRRVNLREGIFQAVTAGVVLPKPIATCRYWHRSLNPKKLIEVKFSHLSRK
Hs1	258	RVAPVLIKEITRRVNLLEGIFQAVYTAGVVLPKPVGTCRYWHRSLNPRKLIEVKFSHLSRN
Hs2	260	RVAPVLIKEITRRVNLLEGIFQAVYTAGVVLPKPIATCRYWHRSLNPRKLIEVKFSHLSRN

Tb	235	YKRFQNPVAVMERLYRLPDKTKTGLRLMEPADVPQVTQLLLKRLASFVAPVFN-EEEV
Lm	239	YQKFQNPMAMLKENYQLPSAPKNSGLREMKPSDVQVRRILMNYLDSFDVGPVFS-DAEI
Pv	233	L-----TMSRAIKLYRMEDTLNINLRLMKKKDVEGVHKLKLSYLEQFNLYAVFT-KEEI
Pf	233	L-----TMSRAIKLYRIDDTLNINLRLMKKKDIDGLQKLLNEHLKQYNLHAIFS-KEDV
Ce	271	M-----TMARTIKLYKLPEETATNLREMKSTDVPQVFKLLTTSLKQYSLAPVYNSEEL
Bm	293	M-----TMORTIKLYKLDPHTANTLVPMKKCHIDGAYGLLQCYLKKFIDLSPQFT-RADF
Hs1	318	M-----TMORTIKLYRLPEPTKTAGLRPMETKDI PVVHQLLTRYLKQFHLLTPVMS-QEEV
Hs2	320	M-----TMORTIKLYRLPDVTKTSGLRPMPEPKDKSVRELINTYLYKQFHLLAPVMD-EEEV

Tb	294	AHWFLPREGVVFSYVVESPVGPCKDEENAGKTSKGTPTGTKCVTGGCEKVIITDFFSFYSL
Lm	298	SHMLLPRDGVVETVYVEN-----DKKVTDFFSFYRI
Pv	287	AHWFLPTENVITYVNE--E-----NGKIKDMISFYSL
Pf	287	AHWETPIDQVIYTYVNE--E-----NGEIKDLISFYSL
Ce	326	AHALVPRKGVYYSYVAE-NQ-----NGKITDFVSFYSL
Bm	347	EHWFMPREDVIYSYVALNEE-----DSKVSDLISFYSL
Hs1	372	EHWFYFQENIIDTEVVE-NA-----NGEVTDFLSFYTL
Hs2	374	AHWFLPREHIIIDTEVVE-SP-----NGKLTDFLSFYTL

Tb	354	PSTLIGNSNHSLLKVAAYVYTAATSVSTIQLVNDLLIIVKLNFGFDVCNVVDIYDNGTYLK
Lm	329	PSTVIGNSNYNLLNAAVYHYAATSIPLHQLIIDLLIVAHSRGFDVCNMVEIIDNRSFVE
Pv	318	PSQLIGNDKYSTLNAAYSFYNVTTATFKQLMODATILAKRNNFDVFNALEVMDNKSQVFE
Pf	318	PSKVLGNKNYNILNAAVSFYNTTTTTFKNLIQDATICLAKRNNFDVFNALEVMDNYSVFQ
Ce	358	PSTVMCHTTHKITYAAYLYYVAGSVTPKQLINDSLILANREKFDVFNALEDLMDHNEKIFS
Bm	380	PSSVMHHPQYKSIRAAYSFYNVATSVTLKQLINDALILARNCGFDVFNALEDLMDNKEILE
Hs1	404	PSTIMNHPTHKSLAAYSFYNVHTQTPLLDLMSDALVLAKMKGFDVFNALEDLMDNKTIFLE
Hs2	406	PSTVMHHPAHKSLAAYSFYNVHTTETPLLDLMSDALILAKSKGFDVFNALEDLMDNKTIFLE

Tb	414	ELKFSPGDGNLYYYFYNNWSYPSIPANEVGLVMV
Lm	389	QLKFGAGDGHRLRYFYNNWAYPKIKPSQVALVML
Pv	378	DLKFGEGDGSLKYYLYNWKCAAFAPAHVGIVLL
Pf	378	DLKFGEGDGSLKYYLYNWKCAFSCHPSKIGIVLL
Ce	418	DLKFGEKGDGNLQYYLYNWKCADMKPSQIGLVLQ
Bm	440	DLKFGIGDGNLQYYLYNWKCPDIIEPKIGLVLQ
Hs1	464	KLKFGIGDGNLQYYLYNWKCPMSGAEKVGLVLQ
Hs2	466	KLKFGIGDGNLQYYLYNWKCPGTDSEKVGVLVLQ

5.2.3 Enzyme assay

Inhibitors in 50 mM DMSO stock solutions were dissolved in 10% DMSO/water (v/v) solution to the target well concentrations by serial dilution. The inhibitor solution (10 μ L), Myr-CoA (25 μ L, final concentration 4 μ M) and NMT (50 μ L) were added to each well of the Greiner black microplate 96-well plate (Sigma-Aldrich). Inhibitors were not added to the positive control wells and both the inhibitor and enzyme were absent from the negative control wells. The reaction was started by adding the peptide/CPM solution (25 μ L, final concentration 4 μ M for peptide and 8 μ M for CPM). In the endpoint assay, the plates were covered and left to react for 30 min before adding the quench solution (60 μ L, 50 mM acetic acid, 50 mM sodium acetate, pH 4.8). The fluorescence was read using an EnVision Xcite Plate Reader (PerkinElmer, excitation at 380 nm and emission at 470 nm). In the continuous assay, the fluorescence was read at set time intervals without quenching. The readouts were corrected using positive controls (as 100%) and negative controls (as 0%) in Excel 2017 and fitted by non-linear regression fitting using Grafit 7 (Erithacus Software Limited).

In the kinetic study, Myr-CoA (35 μ L, final concentration 30 μ M), NMT (50 μ L) and CPM (10 μ L, final concentration 8 μ M) were added to each well of the Greiner black microplate 96-well plate (Sigma-Aldrich). Four peptides in 10 mM DMSO stock solutions were dissolved in 5% DMSO/buffer (v/v) solution to the target well concentrations by serial dilution. The reaction was started by adding peptide solution (15 μ L, final concentration 32 to 0.03125 μ M).

5% DMSO buffer without peptides (15 μ L) were added to the negative control wells. The fluorescence was measured every minute for the initial 45 minutes using an EnVision Xcite Plate Reader (PerkinElmer, excitation at 380 nm and emission at 470 nm) and the results were fitted to the linear equation to obtain $V_{\max}$ and then to the Michaelis Menten equation for K_m by OriginPro.

All IC_{50} s were determined using a CPM assay and are the mean values of at least two determinations. The assay mechanism is described in section 2.1.2.

The assay buffer mentioned above is a Milli-Q water solution with 20 mM phosphate, 0.5 mM EDTA, 0.1% (v/v) Triton X-100 and pH 7.9.

The final concentration of NMTs are 300 ng/mL for Hs NMT1, 200 ng/mL for Hs NMT2, 400 ng/mL for Lm NMT, 200 ng/mL for Pv NMT, 300 ng/mL for Ce NMT and 300 ng/mL for Bm NMT.

Table 8 List of peptide substrates

Peptide	Origin	Sequence	Synthesized by
8AA	Hs p60-Src ₍₂₋₉₎	GSNKSKPK-NH ₂	Emmanuelle Thinon
15AA	Hs p60-Src ₍₂₋₁₆₎	GSNKSKPKDASQRRR-NH ₂	Jone Yu
PfAA	Pf Arf ₍₂₋₁₆₎	GLYVSRLFNRLFQKK-NH ₂	Victor Goncalves
CaAA	Ca Arf ₍₂₋₁₆₎	GLTISKLFASLLGRR-NH ₂	Tayo Olaleye

8AA was used to determine the IC₅₀s of Hs NMT1 and Hs NMT2, 15 AA was used to determine the IC₅₀s of Ld NMT, Pv NMT, Pf NMT, Bm NMT and Ce NMT, and the four peptides were all used in the kinetic study of nematode NMTs.

5.2.4 Cell culture

Cell culture procedures were carried out aseptically in a class II laminar flow cabinet.

HeLa was grown in DMEM supplemented with 10 % FBS and incubated in a humidified atmosphere containing 10 % of CO₂ at 37 °C. It was cultured on 6 cm or 10 cm cell culture plates (Falcon or Corning). When the cells reached confluency, the culture media was aspirated off. Cells were washed using PBS and detached using trypsin (0.25 %). When the cells were seen to be fully detached under a microscope, DMEM was added and the cell suspension was used for subculture.

BL41 was grown in RPMI-1640 supplemented with 10 % FBS and incubated in a humidified atmosphere containing 5 % of CO₂ at 37 °C. It was cultured in T25 or T75 flasks (Falcon or Corning). Cells were maintained at a density between 3×10^5 to 2×10^6 cells/mL. To subculture

and change medium, the cells were transferred into 50 mL Falcon tubes and centrifuged at $1,000 \times g$ for 5 min at 25 °C. After completion, the old medium was aspirated, and the new medium was added. Cell density was counted, and the cells were seeded at the desired density.

When not in use, cells were frozen with cell freezing medium (10 % DMSO in FBS) and stored at 150 °C.

5.2.5 MTS assay

10 mL of HeLa cells in fresh medium (32000 cells/mL) were made after cell counting. 50 μ L of HeLa cells were added into the inner wells of 96-well plates (wells B-G in columns 2-11). 150 μ L medium were added to the outer wells. The plates were incubated for 24 hours before dosing.

The inhibitor stocks were made in deep-well 96-well plates. 400 μ L of medium with 0.2% DMSO is added to wells 2-7. 598 μ L of medium was added to well 1 and mixed with 1.8 μ L of inhibitor stock solution (50 mM in DMSO). Serial dilution was done by transferring 200 μ L of medium from well 1 to well 2 with thorough mixing and repeating this transfer until well 7. The puromycin control was made by mixing 3 μ L puromycin stock (1.5g/mL in DMSO) with 1497 μ L medium.

Two inhibitors were dosed in one 96-well plate pre-seeded with HeLa cells. For the first inhibitor, 100 μ L of the 7 dilutions were added to well B4-B10, C4-C10 and D4-D10. For the second inhibitor, 100 μ L of the 7 dilutions were added to well E4-E10, F4-F10 and G4-G10. For positive control, 100 μ L of 0.2% DMSO/medium was added to well B2-G2 and B11-G11. For negative control, 100 μ L of the puromycin solution was added to well B3-G3 to get a final concentration of 2 μ g/mL. The dosed plates were returned to the incubator.

After 72 hours, the plates were treated with PMS (phenazine methosulfate) and MTS (3-(4,5-Dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium inner salt) for two hours. The PMS/MTS solution was made by adding 150 μ L PMS solution (0.92 mg/mL in PBS) to 3 mL MTS solution (2 mg/mL in PBS). 20 μ L of the PMS/MTS solution was added to each assay well.

The absorbance of each plate was measured at 490 nm. The average readout of the negative control was subtracted from each readout and the relative metabolic activity was calculated as a percentage of the average readout of the positive control. The EC₅₀s were calculated using Grafit 7.0 (Ethacus Software Ltd, UK).

For BL41 cells, the process was the same except that cells were seeded just before dosing at 6000 cells/mL. The incubation time with the PMS/MTS solution was four hours.

5.2.6 Metabolic tagging assay

Four 10 cm plates of HeLa cells at 80% confluency were split into sixteen 6 cm plates and incubated for 24 hours.

The inhibitor solutions in medium were made in six 15 mL Falcon tubes at 10, 3.33, 1.11, 0.37, 0.12 and 0 μ M with 0.2% DMSO. 4 mL of each inhibitor solution was added into two of the pre-seeded 6 cm plates (one for YnC12 and one for HPG). 4 mL of 0.2% DMSO in medium were added to the four control plates. The fourteen plates were incubated for 30 min before treated with YnC12 or HPG. 1.6 μ L of the YnC12 DMSO stock at 50 mM was added into each plate in the YnC12 group to achieve a final concentration of 20 μ M, except for the negative control plate, in which 1.6 μ L DMSO was added. 120 μ L of the HPG stock at 200 mM in Milli-Q water was added into each plate in the HPG group to get the final concentration of 2.4 mM,

except for the negative control plate, in which 120 μ L of Milli-Q water was added. The treated plates were incubated for another 6 hours.

The culture media was aspirated, and cells were washed with 3 mL of PBS. The plates were placed on ice and 150 μ L of the lysis buffer (PBS 1x, 0.1% SDS, 1% Triton X-100, 1 EDTA-free Complete protease inhibitor (Roche Diagnostics)) was added to each plate. The cells were manually lifted by a cell lifter. The lysate was collected in 2 mL Eppendorfs and spun down at $17,000 \times g$ for 10 min at 4 $^{\circ}$ C. 200 μ L of the supernatant of each Eppendorf was transferred into a new 0.5 mL Eppendorf and stored at -80 $^{\circ}$ C.

The protein concentration of each sample was determined using the commercially available Bio-Rad DC Protein Assay Kit. A calibration line was prepared using BSA (bovine serum albumin) solutions in lysis buffer at 2, 1, 0.5, 0.25 and 0 mg/mL. The absorbance at 750 nm was measured using the EnVision Xcite Plate Reader (PerkinElmer).

Samples with 50 μ g total protein in 47 μ L lysis buffer were made for the click labelling. 3 μ L of click reagents containing TAMRA-capture reagent (0.5 μ L of 10 mM stock), CuSO_4 (1 μ L of 50 mM stock), TCEP (1 μ L of 50 mM stock) and TBTA (0.5 μ L of 10 mM stock), was added to each sample. The samples were vortexed for 1 hour at room temperature before 2 μ L of EDTA (100 mM) was added to quench the reaction.

100 μ L of MeOH, 25 μ L of CHCl_3 and 50 μ L of water were added to each sample. The protein in the samples was precipitated by vortexing and centrifuging at $17,000 \times g$ for 10 min at room temperature. The supernatant was carefully removed without touching the protein pellets. The pellets were washed with 500 μ L of MeOH twice by vortexing and centrifuging at $17,000 \times g$ for 3 min.

2% SDS solution was made by mixing 2 mL of sodium dodecyl sulfate (10% in water), 1 mL phosphate buffered saline and 7 mL of Milli-Q water. 3.8 μ L of the 2% SDS was added to each sample. The samples were vortexed for 10 min at 4 °C before 33.7 μ L of PBS was added to each sample. The sample loading buffer was made by mixing 4 $\times$ sample loading buffer (NuPAGE LDS sample buffer) and β -mercaptoethanol at a 5:1 ratio. The samples were added with 12.5 μ L of the sample loading buffer, heated at 95 °C for 5 min and centrifuged at 1,000 $\times$ g for 1 min. The samples were loaded onto Bis-Tris gels made by the recipe in Table 9.

Table 9 Recipe of Bis-Tris gels (for two gels)

Component	Resolving gel	Stacking gel
Milli-Q water	2.3 mL	2.55 mL
1M Bis-Tris (PH 6.7)	3.6 mL	1.8 mL
30% acrylamide/bis-acrylamide	4 mL	0.67 mL
10% APS	100 μ L	25 μ L
TEMED	4 μ L	4 μ L

The gels were run for 15 min at 100 V for stacking then 50 min at 160 V. The fluorescence on the gels were imaged by Typhoon FLA 9500 Imager (GE Healthcare) and the protein loading was checked by Coomassie.

The intensity of fluorescent bands was analyzed by ImageJ by integrating the area under each band of interest using the ‘gel analyzer’ function and normalized relative to the positive

controls with YnC12/HPG but without inhibitors. The in-cell tagging IC_{50} s (TC_{50} s) were determined using GraphPad Prism 5 by non-linear regression fitting.

5.3 Crystallography

5.3.1 General methods

Protein purification was performed on AKTA pure (GE healthcare life sciences). Sitting drops were set up by Mosquito LCP nanolitre crystallisation robot (TTP Labtech). Structural data was collected at Diamond Light Source.

Polyethylene glycols and plates used in crystallography were purchased from Molecular Dimensions. Except **IMP-0001096**, inhibitors used in crystallography study were made by *Dr* Andy Bell, *Dr* Zhiyong Yu or *Dr* Jennie Hutton (Imperial College London). The plasmid of Hs NMT1 was a kind gift from *Dr* James Brannigan (University of York).

The structure of **DDD85646**, **IMP-0000917**, **IMP-0001031** and **IMP-0000899** was solved by *Dr* James Brannigan. The structure of **IMP-0001088** was solved by *Dr* Inma Perez Dorado⁴⁷.

All 3D images of protein structures and ligand binding modes were analysed and generated by PyMOL v2.0²⁸.

5.3.2 Transformation

The plasmid of Hs NMT1 is a kind gift from *Dr* James A. Brannigan, University of York and its preparation is detailed in previous publication¹¹⁶.

10 μ L of RosettaTM(DE3) pLysS Competent Cells (Novagen) was mixed with 1 μ L of plasmid of Hs NMT1 in a clean and autoclaved 1.5 mL Eppendorf. The Eppendorf was kept on ice for 5 min, transferred to a heating block at 42 °C for 45 seconds, then moved back onto ice for another 5 min. 90 μ L of SOC media was added and the Eppendorf was stuck onto a shaker to mix at 37 °C for one hour. The suspension was then spread evenly on a LB-agar plate supplemented with kanamycin at 50 μ g/mL and incubated at 37 °C overnight. A negative

control plate was treated the same way using cell suspensions without transformation of the plasmid.

5.3.3 Expression

In 50 mL sterile Falcon tubes, colonies picked from the plate were added into 5 mL of Lysogeny broth (LB) media, prepared by adding 10 g tryptone, 5 g yeast extract, 5 g NaCl in 1 litre of double-distilled H₂O, with kanamycin at 50 µg/mL. The culture was kept growing overnight at 37 °C with shaking. On the next day, 20 mL of fresh LB media with 50 µg/mL kanamycin was added to the overnight culture. The culture was left to grow for another one hour at 37 °C while pre-warming 500 mL of the autoinduction media (FORMEDIUM). After pre-warming, 50 µg/mL kanamycin and the starter culture were added into the autoinduction media. The larger culture was incubated at 37 °C for 6 hours and then 20 °C for 20 hours. The final culture was divided into two centrifuge bottles and centrifuged at 4000 rpm for 15 min. Pellets were re-suspended in the lysis buffer and used for purification. If purification was not directly followed, the pellets was frozen without lysis buffer at -20 °C for later use.

5.3.4 Purification

Table 10 Buffers for protein purification

Lysis buffer	20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 10 mM MgCl ₂ , 1 mM MnCl ₂ , supplemented with protease inhibitors, bovine pancreatic DNase I and 0.1% (v/v) Triton X-100.
IMAC washing buffer	20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole.
IMAC elution buffer	20 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole.
IEX buffer A	20 mM Tris-HCl pH 8.9, 25 mM NaCl, 1 mM DTT.

IEX buffer B	20 mM Tris-HCl pH 8.9, 500 mM NaCl, 1 mM DTT.
Gel filtration buffer	20 mM Tris-HCl pH 8.5, 250 mM NaCl, 2 mM DTT, 2 mM EGTA, 2 mM EDTA.
Crystallization buffer	20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT.

Cells pellets were thawed (if frozen) and resuspended in the lysis buffer (100 mL per 1 L of cell culture). The suspension was sonicated at full power and 50% time on/off for 20 min in an ice/water bath. The resulting mixture was centrifuged at 17000 rpm for 15 min. The supernatant was loaded directly onto an IMAC (immobilized metal-affinity chromatography) column (1 mL HisTrap FF Crude, GE Healthcare) and eluted in a linear gradient of 10-500 mM imidazole. Fractions were collected and checked by SDS-PAGE gel. Fractions containing NMT were pooled together and dialyzed into IEX buffer A overnight in the cold room. On the next day, the sample was centrifuged at 10K rpm for 10 min. The supernatant was loaded onto an ion-exchange (IEX) column (1mL HiTrap Q HP, GE Healthcare) previously washed and equilibrated in IEX buffer A using AKTA pure (GE Healthcare). The column was then eluted in a linear gradient of 0-100% buffer B (25-500 mM NaCl). The protein sample from the ion-exchange column was dialyzed into the washing IMAC buffer and digested by HRV 3C protease (0.1 µg/mL) overnight in the cold room. Then the mixture was passed through the IMAC column again to remove protease, His-tag and undigested Hs NMT1. The protein sample was concentrated to 2 mL and passed through a size-exclusion column (Superdex 75, 16/60, GE Healthcare) with the gel filtration buffer. Fractions containing Hs NMT1 was pooled together and stored at -20 °C with 50% glycerol.

5.3.5 Crystallization

Protein was dialyzed into the crystallization buffer and concentrated to 7.2 mg/mL. 4-8% of 10% HCl or 1 M NaOH was used to adjust the pH of ligands to between 5 and 8.

Table 11 The two PEG conditions

NO.	Reservoir condition
1	16-26% w/v mPEG 2K, 0.1M NaCit PH 4.5, 0.2M KBr, 5% v/v glycerol
2	17-27% w/v PEG 4K, 0.1M NaCit PH 4.5, 5mM NiCl ₂ , 2.5% v/v glycerol

The two PEG conditions (Table 11) were tested on ligands in parallel in 96-well 2-drop MRC Crystallization Plates by Mosquito LCP. On each plate, 12 different reservoir conditions (6 of each PEG) were tested on 8 ligands. In each well, 0.15 μ L of the complex solution containing 7.2 mg/mL Hs NMT1, 0.3 mM Myr-CoA and 0.5-5 mM ligand was mixed with 0.15 μ L of reservoir solution. The reservoir was filled with 80 μ L of reservoir solution. The plates were stored at 20 °C and inspected every day. Conditions in which crystals grew were repeated in 48-well MRC Maxi Optimization plates using 0.5 μ L of both complex solution and reservoir solution in the drop and 80 μ L in reservoir to achieve better crystals. Crystals for data collection were fished out from the drops where they grew, transferred into the cryo-protectant for 2 seconds and frozen in liquid nitrogen. Cryo-protection solutions were made by adding the same concentration of ligands and 20-35 % of glycerol to the corresponding reservoir solutions. The reservoir conditions of all solved ligands are show in Table 12.

Table 12 Reservoir conditions of all ligands

Ligand	Reservoir condition
IMP-0001029	30% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol
IMP-0001114	24% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001117	26% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001161	26% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001180	16% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001222	18% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001223	29% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol
IMP-0001242	14% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001320	27% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol
IMP-0001336	22% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001343	22% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0000180	12% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0000899	27% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol
IMP-0000927	26% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol

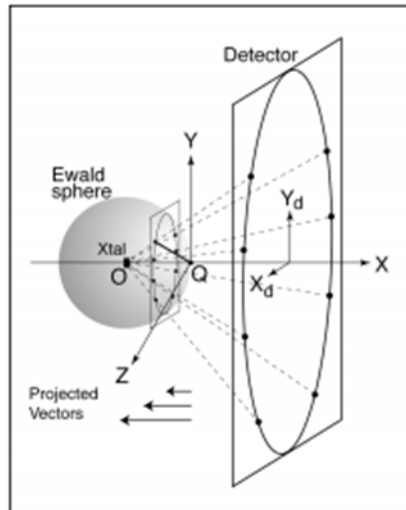
IMP-0000974	12% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001096	Socking crystal in 22% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001169	20% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001171	16% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001210	22% w/v PEG 4K, 0.1 M NaCit pH 4.5, 5 mM NiCl ₂ , 2.5% v/v glycerol
IMP-0001227	30% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol
IMP-0001232	27% w/v mPEG 2K, 0.1 M NaCit pH 4.5, 0.2 M KBr, 5% v/v glycerol

5.3.6 Data processing

The X-ray diffraction images collected as described in 3.1.5 was processed in four major steps by the automated data processing pipelines xia2¹⁴⁶ or autoPROC¹⁴⁷: spot finding and autoindexing; parameter refinement; integration; scaling and merging¹⁶⁴.

First of all, the likely diffraction spots on the image or images need to be located. XDS¹⁴⁸, DIALS¹⁴⁹ or MOSFLM¹⁵³ were used for this “spot finding” step. Depending on the default setting of each software, a selected group of images or the entire dataset were used.

After spot finding, indexing was carried out using an automated procedure (hence “autoindexing”), starting from mapping the spot positions obtained using the Ewald sphere construction. The 2D image coordinates of the spots were converted to scattering vectors that correspond to the 3D reciprocal lattice coordinates (Figure 90).



$$s = \begin{pmatrix} D/r - 1 \\ X_d/r \\ Y_d/r \end{pmatrix}$$

$$r = \sqrt{D^2 + X_d^2 + Y_d^2}$$

Figure 90 Converting 2D coordinates to scattering vectors (images taken from Harry Powell's slides named "Integration and Scaling"¹⁶⁵). s is the scattering vector, D is the crystal to detector distance, X_d and Y_d are the spot coordinates relative to the beam centre on the image.

Each scattering vector obtained was then oriented to reflect the orientation of the crystal by applying a simple rotation about the origin. There are two methods for indexing. In the real space method, the reciprocal space scattering vectors are transformed into real space before searching for a unit cell described by the diffraction pattern. In the reciprocal space methods, the search for a unit cell is performed directly with the scattering vector. Only after successful indexing, the results were converted into real space.

Taking MOSFLM¹⁵³ for example, the scattering vectors calculated from the diffraction spots are projected on to a set of several thousand vectors covering all directions of the whole hemisphere of the Ewald sphere. If the scattering vectors are projected along a real space axis direction, all the projections (the scalar products of the scattering vector and the vectors that describe the direction of the projection) for spots in the same reciprocal space plane will have the same length, as will for all those spots in the next plane, etc. These projections will cluster at lengths which are multiples of lattice spacing. If the scattering vectors are projected along other directions, the lengths of projections will be random without clustering. MOSFLM¹⁵³

applies a 1D Fourier transform on the projections calculated, and examines a sample of those giving the largest non-origin peaks to determine which can be combined to give a real space unit cell with reasonable axial lengths and interaxial angles¹⁶⁶⁻¹⁶⁷.

Whatever program was used for autoindexing, the initial result is a primitive triclinic cell. Via a set of standard matrix transformations, all programs compare the fit of the initial cell with the 44 characteristic lattices¹⁶⁸, each corresponds to one of the 14 Bravais lattices. Each program uses a different penalty function to describe the goodness-of-fit, giving multiple solutions of the unit cell geometry with different penalty scores. The highest symmetry from the low-penalty bunch was selected as an initial estimate. However, the true Bravais lattice can only be determined by analysing the intensities of symmetry-equivalent reflections.

The approximate parameters describing the geometry of the unit cell and the diffractometer derived from autoindexing were improved by refinement. The most common method, positional refinement, uses the spot positions on each image and optimizes the fit of the observed to calculated spot coordinates on the image (Figure 91).

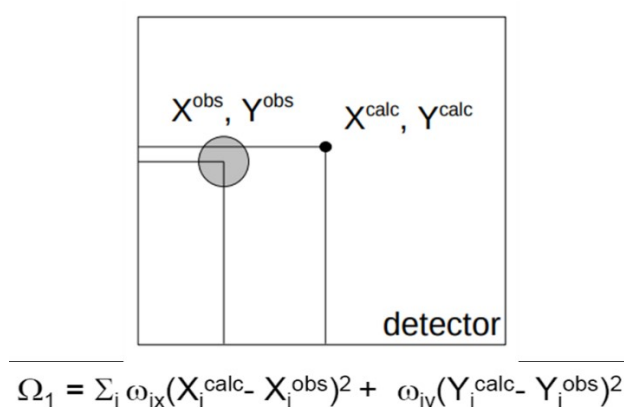
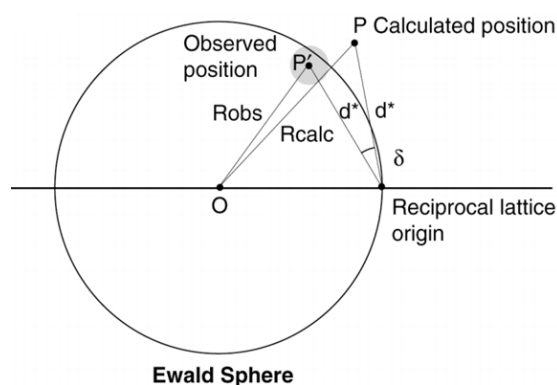


Figure 91 Positional refinement (images taken from Judit É Debreczeni's slides named "Data processing"¹⁶⁹). Ω_1 is the positional residual for least-squares minimisation. ω_{ix} and ω_{iy} are weights for X_i and Y_i , respectively. X_{calc} and Y_{calc} are the calculated spot coordinates and X_{obs} and Y_{obs} are the observed spot coordinates on the detector.

The second approach, post-refinement, analyses the relative intensity of partially recorded reflections across two or more contiguous images to optimize the distance of the reciprocal lattice points from the Ewald sphere (Figure 92).



$$\Omega_2 = \sum_{i=1}^n w_i \left[\frac{(R_i^{calc} - R_i^{obs})}{d_i} \right]^2$$

Figure 92 Post-refinement (images taken from Harry Powell's slides named "Integration and Scaling"¹⁶⁵). Ω_2 is the angular residual for least-squares minimisation. w_i is the weight for R_i . R_i^{calc} and R_i^{obs} are the calculated and observed distances of the reciprocal lattice point $\mathbf{d}^*$ from the centre of the Ewald sphere. d_i is the length of this $\mathbf{d}^*$.

Since this approach requires the intensities of the spots, it was carried out during or after the integration process (hence called "post-refinement") or substituted by other refinements in different programs.

After the crystal and instrument parameters have been optimised by refinement, integration was performed to measure the diffracted intensities. There are two common methods of integration being used in combination: summation integration and profile fitting. In summation integration, a reflection mask, the region used for integration that distinguishes the spots from the background, was determined by different methods in different programs. The pixel values of all pixels in the peak area of a spot are added together. The background region around the spot is used to estimate the background level under the spot and subtracted from the spot's

apparent intensity. This simple and fast method can be used for most spots except weak, overloaded or partially overlapping reflections. In profile fitting, the profiles of strong, well-measured reflections are scaled and used to improve the estimate of the intensity of neighbouring weak spots on a detector. This method gives a significant benefit in the measurement of weak reflections on a high background, where the shoulders of the weak spots are hidden by the background, as well as identifying outliers and overloaded spots.

Although sometimes the true space group is only known when the structure has been successfully solved and refined, a small number of possible space group can be identified from the diffraction experiment¹⁵¹. Space group assignment was carried out once integrated intensities of all reflections are available. For the reflection file produced by MOSFLM¹⁵³, Pointless¹⁵² was used to determine the Laue group from unmerged intensities. Since there is only one point group corresponding to each Laue group for chiral space groups, it also gives the possible space groups by analysing axial systematic absences¹⁶⁹.

The data obtained from integration are not on a common scale due to various physical reasons, including changes in incident radiation intensity, the illuminated volume of the crystal, anisotropic absorption of X-rays by the crystal, radiation damage to the crystal and variations in the response of the detector¹⁶⁴. Therefore, Aimless¹⁵⁰ or XSCALE¹⁵¹ were used to scale the intensities of equivalent reflections before merging.

Once the observations are on a common scale, the symmetry-related reflections were merged into unique observations and the individual parts of partially recorded reflections, if not merged in the integration step, were also merged into complete reflections. These multiple observations were reduced to a weighted mean intensity and standard deviation, which can be modified by the level of agreement between multiple observations¹⁷⁰. Statistics on the agreement between

multiple observations and data completeness were also calculated at this stage. Finally, the intensities were converted into structure factor amplitudes ($|F|$) and the distribution of $|F|$ were analysed to give information on B-factor, twinning, etc.

6 References

1. Storck, E. M.; Serwa, R. A.; Tate, E. W., Chemical proteomics: a powerful tool for exploring protein lipidation. *Biochemical Society transactions* **2013**, *41* (1), 56-61.
2. Tate, E. W.; Kalesh, K. A.; Lanyon-Hogg, T.; Storck, E. M.; Thinon, E., Global profiling of protein lipidation using chemical proteomic technologies. *Current opinion in chemical biology* **2015**, *24*, 48-57.
3. Chen, B.; Sun, Y.; Niu, J.; Jarugumilli, G. K.; Wu, X., Protein Lipidation in Cell Signaling and Diseases: Function, Regulation, and Therapeutic Opportunities. *Cell chemical biology* **2018**, *25* (7), 817-831.
4. Resh, M. D., Targeting protein lipidation in disease. *Trends in Molecular Medicine* **2012**, *18* (4), 206-214.
5. Rocks, O.; Gerauer, M.; Vartak, N.; Koch, S.; Huang, Z. P.; Pechlivanis, M.; Kuhlmann, J.; Brunsvel, L.; Chandra, A.; Ellinger, B.; Waldmann, H.; Bastiaens, P. I., The palmitoylation machinery is a spatially organizing system for peripheral membrane proteins. *Cell* **2010**, *141* (3), 458-71.
6. Clevers, H.; Nusse, R., Wnt/beta-catenin signaling and disease. *Cell* **2012**, *149* (6), 1192-205.
7. Zou, C.; Ellis, B. M.; Smith, R. M.; Chen, B. B.; Zhao, Y.; Mallampalli, R. K., Acyl-CoA:lysophosphatidylcholine acyltransferase I (Lpcat1) catalyzes histone protein O-palmitoylation to regulate mRNA synthesis. *J Biol Chem* **2011**, *286* (32), 28019-25.
8. Gutierrez, J. A.; Solenberg, P. J.; Perkins, D. R.; Willency, J. A.; Knierman, M. D.; Jin, Z.; Witcher, D. R.; Luo, S.; Onyia, J. E.; Hale, J. E., Ghrelin octanoylation mediated by an orphan lipid transferase. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (17), 6320-5.
9. Buglino, J. A.; Resh, M. D., Hhat is a palmitoylacyltransferase with specificity for N-palmitoylation of Sonic Hedgehog. *J Biol Chem* **2008**, *283* (32), 22076-88.
10. Esterbauer, H.; Cheeseman, K. H., Determination of aldehydic lipid peroxidation products: malonaldehyde and 4-hydroxynonenal. *Methods in enzymology* **1990**, *186*, 407-21.
11. Schopfer, F. J.; Cipollina, C.; Freeman, B. A., Formation and signaling actions of electrophilic lipids. *Chemical reviews* **2011**, *111* (10), 5997-6021.
12. Chen, Y.; Qin, W.; Wang, C., Chemoproteomic profiling of protein modifications by lipid-derived electrophiles. *Current opinion in chemical biology* **2016**, *30*, 37-45.
13. Reed, T. T., Lipid peroxidation and neurodegenerative disease. *Free radical biology & medicine* **2011**, *51* (7), 1302-19.
14. Butterfield, D. A.; Gu, L.; Di Domenico, F.; Robinson, R. A., Mass spectrometry and redox proteomics: applications in disease. *Mass spectrometry reviews* **2014**, *33* (4), 277-301.

15. Delmastro-Greenwood, M.; Freeman, B. A.; Wendell, S. G., Redox-dependent anti-inflammatory signaling actions of unsaturated fatty acids. *Annual review of physiology* **2014**, *76*, 79-105.
16. Liebler, D. C., Protein damage by reactive electrophiles: targets and consequences. *Chemical research in toxicology* **2008**, *21* (1), 117-28.
17. Orlean, P.; Menon, A. K., Thematic review series: lipid posttranslational modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycosphospholipids. *J Lipid Res* **2007**, *48* (5), 993-1011.
18. Yu, S.; Guo, Z.; Johnson, C.; Gu, G.; Wu, Q., Recent progress in synthetic and biological studies of GPI anchors and GPI-anchored proteins. *Current opinion in chemical biology* **2013**, *17* (6), 1006-13.
19. Tsai, Y. H.; Liu, X.; Seeberger, P. H., Chemical biology of glycosylphosphatidylinositol anchors. *Angewandte Chemie (International ed. in English)* **2012**, *51* (46), 11438-56.
20. Wang, M.; Casey, P. J., Protein prenylation: unique fats make their mark on biology. *Nature reviews. Molecular cell biology* **2016**, *17* (2), 110-22.
21. Heal, W. P.; Jovanovic, B.; Bessin, S.; Wright, M. H.; Magee, A. I.; Tate, E. W., Bioorthogonal chemical tagging of protein cholesterylation in living cells. *Chemical communications (Cambridge, England)* **2011**, *47* (14), 4081-3.
22. Xiao, X.; Tang, J. J.; Peng, C.; Wang, Y.; Fu, L.; Qiu, Z. P.; Xiong, Y.; Yang, L. F.; Cui, H. W.; He, X. L.; Yin, L.; Qi, W.; Wong, C. C.; Zhao, Y.; Li, B. L.; Qiu, W. W.; Song, B. L., Cholesterol Modification of Smoothed Is Required for Hedgehog Signaling. *Molecular cell* **2017**, *66* (1), 154-162.e10.
23. Nakatogawa, H.; Ichimura, Y.; Ohsumi, Y., Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. *Cell* **2007**, *130* (1), 165-78.
24. Wright, M. H.; Heal, W. P.; Mann, D. J.; Tate, E. W., Protein myristoylation in health and disease. *Journal of chemical biology* **2010**, *3* (1).
25. Resh, M. D., Trafficking and signaling by fatty-acylated and prenylated proteins. *Nature chemical biology* **2006**, *2* (11), 584-90.
26. Towler, D. A.; Adams, S. P.; Eubanks, S. R.; Towery, D. S.; Jackson-Machelski, E.; Glaser, L.; Gordon, J. I., Purification and characterization of yeast myristoyl CoA:protein N-myristoyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **1987**, *84* (9), 2708-12.
27. Rudnick, D. A.; McWherter, C. A.; Rocque, W. J.; Lennon, P. J.; Getman, D. P.; Gordon, J. I., Kinetic and structural evidence for a sequential ordered Bi Bi mechanism of catalysis by *Saccharomyces cerevisiae* myristoyl-CoA:protein N-myristoyltransferase. *J Biol Chem* **1991**, *266* (15), 9732-9.
28. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

29. Rioux, V.; Legrand, P., Saturated fatty acids: simple molecular structures with complex cellular functions. *Current opinion in clinical nutrition and metabolic care* **2007**, *10* (6), 752-8.
30. Bhatnagar, R. S.; Futterer, K.; Farazi, T. A.; Korolev, S.; Murray, C. L.; Jackson-Machelski, E.; Gokel, G. W.; Gordon, J. I.; Waksman, G., Structure of N-myristoyltransferase with bound myristoylCoA and peptide substrate analogs. *Nat Struct Biol* **1998**, *5* (12), 1091-7.
31. Farazi, T. A.; Waksman, G.; Gordon, J. I., The Biology and Enzymology of Protein N-Myristoylation. *J Biol Chem* **2001**, *276* (43), 39501-39504.
32. Brannigan, J. A.; Roberts, S. M.; Bell, A. S.; Hutton, J. A.; Hodgkinson, M. R.; Tate, E. W.; Leatherbarrow, R. J.; Smith, D. F.; Wilkinson, A. J., Diverse modes of binding in structures of Leishmania major N-myristoyltransferase with selective inhibitors. *IUCr* **2014**, *1* (Pt 4), 250-260.
33. Cyril Dian, I. P. D., Frédéric Rivière, Markus Ritzefeld, Mengjie Shen, Ernesto Cota, Thierry Meinel, Edward Tate, Carmela Giglione, Atomic resolution snapshots of human N-myristoyltransferase in action reveal a unique enzymatic mechanism. *To be published*.
34. Farazi, T. A.; Manchester, J. K.; Waksman, G.; Gordon, J. I., Pre-steady-state kinetic studies of *Saccharomyces cerevisiae* myristoylCoA:protein N-myristoyltransferase mutants identify residues involved in catalysis. *Biochemistry* **2001**, *40* (31), 9177-86.
35. Wilcox, C.; Hu, J. S.; Olson, E. N., Acylation of proteins with myristic acid occurs cotranslationally. *Science* **1987**, *238* (4831), 1275-8.
36. Thinon, E.; Serwa, R. A.; Broncel, M.; Brannigan, J. A.; Brassat, U.; Wright, M. H.; Heal, W. P.; Wilkinson, A. J.; Mann, D. J.; Tate, E. W., Global profiling of co- and post-translationally N-myristoylated proteomes in human cells. *Nature Communications* **2014**, *5*, 4919.
37. Zha, J.; Weiler, S.; Oh, K. J.; Wei, M. C.; Korsmeyer, S. J., Posttranslational N-myristoylation of BID as a molecular switch for targeting mitochondria and apoptosis. *Science* **2000**, *290* (5497), 1761-5.
38. Perinpanayagam, M. A.; Beauchamp, E.; Martin, D. D. O.; Sim, J. Y. W.; Yap, M. C.; Berthiaume, L. G., Regulation of co- and post-translational myristoylation of proteins during apoptosis: interplay of N-myristoyltransferases and caspases. *The FASEB Journal* **2013**, *27* (2), 811-821.
39. Resh, M. D., Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1999**, *1451* (1), 1-16.
40. Bell, A. S.; Mills, J. E.; Williams, G. P.; Brannigan, J. A.; Wilkinson, A. J.; Parkinson, T.; Leatherbarrow, R. J.; Tate, E. W.; Holder, A. A.; Smith, D. F., Selective inhibitors of protozoan protein N-myristoyltransferases as starting points for tropical disease medicinal chemistry programs. *PLoS neglected tropical diseases* **2012**, *6* (4), e1625.
41. Brannigan, J. A.; Smith, B. A.; Yu, Z.; Brzozowski, A. M.; Hodgkinson, M. R.; Maroof, A.; Price, H. P.; Meier, F.; Leatherbarrow, R. J.; Tate, E. W.; Smith, D. F.; Wilkinson, A. J., N-myristoyltransferase

from *Leishmania donovani*: structural and functional characterisation of a potential drug target for visceral leishmaniasis. *Journal of molecular biology* **2010**, *396* (4), 985-99.

42. Frearson, J. A.; Brand, S.; McElroy, S. P.; Cleghorn, L. A.; Smid, O.; Stojanovski, L.; Price, H. P.; Guther, M. L.; Torrie, L. S.; Robinson, D. A.; Hallyburton, I.; Mpamhanga, C. P.; Brannigan, J. A.; Wilkinson, A. J.; Hodgkinson, M.; Hui, R.; Qiu, W.; Raimi, O. G.; van Aalten, D. M.; Brenk, R.; Gilbert, I. H.; Read, K. D.; Fairlamb, A. H.; Ferguson, M. A.; Smith, D. F.; Wyatt, P. G., N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature* **2010**, *464* (7289), 728-32.

43. Selvakumar, P.; Lakshmikuttyamma, A.; Charavaryamath, C.; Singh, B.; Tucheck, J.; Sharma, R. K., Expression of myristoyltransferase and its interacting proteins in epilepsy. *Biochem Biophys Res Commun* **2005**, *335* (4), 1132-9.

44. Selvakumar, P.; Lakshmikuttyamma, A.; Shrivastav, A.; Das, S. B.; Dimmock, J. R.; Sharma, R. K., Potential role of N-myristoyltransferase in cancer. *Progress in lipid research* **2007**, *46* (1), 1-36.

45. Cordeddu, V.; Di Schiavi, E.; Pennacchio, L. A.; Ma'ayan, A.; Sarkozy, A.; Fodale, V.; Cecchetti, S.; Cardinale, A.; Martin, J.; Schackwitz, W.; Lipzen, A.; Zampino, G.; Mazzanti, L.; Digilio, M. C.; Martinelli, S.; Flex, E.; Lepri, F.; Bartholdi, D.; Kutsche, K.; Ferrero, G. B.; Anichini, C.; Selicorni, A.; Rossi, C.; Tenconi, R.; Zenker, M.; Merlo, D.; Dallapiccola, B.; Iyengar, R.; Bazzicalupo, P.; Gelb, B. D.; Tartaglia, M., Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair. *Nature genetics* **2009**, *41* (9), 1022-6.

46. Su, R.; Han, Z. Y.; Fan, J. P.; Zhang, Y. L., A possible role of myristoylated alanine-rich C kinase substrate in endocytic pathway of Alzheimer's disease. *Neuroscience bulletin* **2010**, *26* (4), 338-44.

47. Mousnier, A.; Bell, A. S.; Swieboda, D. P.; Morales-Sanfrutos, J.; Pérez-Dorado, I.; Brannigan, J. A.; Newman, J.; Ritzefeld, M.; Hutton, J. A.; Guedán, A.; Asfor, A. S.; Robinson, S. W.; Hopkins-Navratilova, I.; Wilkinson, A. J.; Johnston, S. L.; Leatherbarrow, R. J.; Tuthill, T. J.; Solari, R.; Tate, E. W., Fragment-derived inhibitors of human N-myristoyltransferase block capsid assembly and replication of the common cold virus. *Nature Chemistry* **2018**.

48. Maurer-Stroh, S.; Eisenhaber, F., Myristoylation of viral and bacterial proteins. *Trends in microbiology* **2004**, *12* (4), 178-85.

49. Thinon, E.; Morales-Sanfrutos, J.; Mann, D. J.; Tate, E. W., N-Myristoyltransferase Inhibition Induces ER-Stress, Cell Cycle Arrest, and Apoptosis in Cancer Cells. *ACS Chemical Biology* **2016**.

50. Felsted, R. L.; Glover, C. J.; Hartman, K., Protein N-myristoylation as a chemotherapeutic target for cancer. *Journal of the National Cancer Institute* **1995**, *87* (21), 1571-3.

51. Magnuson, B. A.; Raju, R. V.; Moyana, T. N.; Sharma, R. K., Increased N-myristoyltransferase activity observed in rat and human colonic tumors. *Journal of the National Cancer Institute* **1995**, *87* (21), 1630-5.

52. Raju, R. V.; Moyana, T. N.; Sharma, R. K., N-Myristoyltransferase overexpression in human colorectal adenocarcinomas. *Experimental cell research* **1997**, *235* (1), 145-54.

53. Rajala, R. V.; Radhi, J. M.; Kakkar, R.; Datla, R. S.; Sharma, R. K., Increased expression of N-myristoyltransferase in gallbladder carcinomas. *Cancer* **2000**, *88* (9), 1992-9.
54. Shrivastav, A.; Sharma, A. R.; Bajaj, G.; Charavaryamath, C.; Ezzat, W.; Spafford, P.; Gore-Hickman, R.; Singh, B.; Copete, M. A.; Sharma, R. K., Elevated N-myristoyltransferase activity and expression in oral squamous cell carcinoma. *Oncology reports* **2007**, *18* (1), 93-7.
55. Lu, Y.; Selvakumar, P.; Ali, K.; Shrivastav, A.; Bajaj, G.; Resch, L.; Griebel, R.; Fournay, D.; Meguro, K.; Sharma, R. K., Expression of N-myristoyltransferase in human brain tumors. *Neurochemical research* **2005**, *30* (1), 9-13.
56. Clegg, R. A.; Gordge, P. C.; Miller, W. R., Expression of enzymes of covalent protein modification during regulated and dysregulated proliferation of mammary epithelial cells: PKA, PKC and NMT. *Advances in enzyme regulation* **1999**, *39*, 175-203.
57. Aleshin, A.; Finn, R. S., SRC: a century of science brought to the clinic. *Neoplasia (New York, N.Y.)* **2010**, *12* (8), 599-607.
58. Wheeler, D. L.; Iida, M.; Dunn, E. F., The role of Src in solid tumors. *The oncologist* **2009**, *14* (7), 667-78.
59. Hetz, C., The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature reviews. Molecular cell biology* **2012**, *13* (2), 89-102.
60. Schonthal, A. H., Targeting endoplasmic reticulum stress for cancer therapy. *Frontiers in bioscience (Scholar edition)* **2012**, *4*, 412-31.
61. Yadav, R. K.; Chae, S.-W.; Kim, H.-R.; Chae, H. J., Endoplasmic reticulum stress and cancer. *J Cancer Prev* **2014**, *19* (2), 75-88.
62. Ritchie, A. I.; Farne, H. A.; Singanayagam, A.; Jackson, D. J.; Mallia, P.; Johnston, S. L., Pathogenesis of Viral Infection in Exacerbations of Airway Disease. *Annals of the American Thoracic Society* **2015**, *12* Suppl 2, S115-32.
63. Kieninger, E.; Singer, F.; Tapparel, C.; Alves, M. P.; Latzin, P.; Tan, H. L.; Bossley, C.; Casaulta, C.; Bush, A.; Davies, J. C.; Kaiser, L.; Regamey, N., High rhinovirus burden in lower airways of children with cystic fibrosis. *Chest* **2013**, *143* (3), 782-790.
64. Flight, W. G.; Bright-Thomas, R. J.; Tilston, P.; Mutton, K. J.; Guiver, M.; Morris, J.; Webb, A. K.; Jones, A. M., Incidence and clinical impact of respiratory viruses in adults with cystic fibrosis. *Thorax* **2014**, *69* (3), 247-53.
65. Thibaut, H. J.; De Palma, A. M.; Neyts, J., Combating enterovirus replication: state-of-the-art on antiviral research. *Biochemical pharmacology* **2012**, *83* (2), 185-92.
66. Marc, D.; Masson, G.; Girard, M.; van der Werf, S., Lack of myristoylation of poliovirus capsid polypeptide VP0 prevents the formation of virions or results in the assembly of noninfectious virus particles. *Journal of virology* **1990**, *64* (9), 4099-107.

67. Marc, D.; Drugeon, G.; Haenni, A. L.; Girard, M.; van der Werf, S., Role of myristoylation of poliovirus capsid protein VP4 as determined by site-directed mutagenesis of its N-terminal sequence. *The EMBO journal* **1989**, *8* (9), 2661-8.
68. Marc, D.; Girard, M.; van der Werf, S., A Gly1 to Ala substitution in poliovirus capsid protein VP0 blocks its myristoylation and prevents viral assembly. *The Journal of general virology* **1991**, *72* (Pt 5), 1151-7.
69. French, K. J.; Zhuang, Y.; Schrecengost, R. S.; Copper, J. E.; Xia, Z.; Smith, C. D., Cyclohexyl-octahydro-pyrrolo[1,2-a]pyrazine-Based Inhibitors of Human N-Myristoyltransferase-1. *Journal of Pharmacology and Experimental Therapeutics* **2004**, *309* (1), 340-347.
70. Devadas, B.; Zupiec, M. E.; Freeman, S. K.; Brown, D. L.; Nagarajan, S.; Sikorski, J. A.; McWherter, C. A.; Getman, D. P.; Gordon, J. I., Design and Syntheses of Potent and Selective Dipeptide Inhibitors of *Candida albicans* Myristoyl-CoA:Protein N-Myristoyltransferase. *J Med Chem* **1995**, *38* (11), 1837-1840.
71. Masubuchi, M.; Kawasaki, K.; Ebiiike, H.; Ikeda, Y.; Tsujii, S.; Sogabe, S.; Fujii, T.; Sakata, K.; Shiratori, Y.; Aoki, Y.; Ohtsuka, T.; Shimma, N., Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 1. *Bioorganic & medicinal chemistry letters* **2001**, *11* (14), 1833-7.
72. Yamazaki, K.; Kaneko, Y.; Suwa, K.; Ebara, S.; Nakazawa, K.; Yasuno, K., Synthesis of potent and selective inhibitors of *Candida albicans* N-myristoyltransferase based on the benzothiazole structure. *Bioorg Med Chem* **2005**, *13* (7), 2509-22.
73. Brand, S.; Cleghorn, L. A. T.; McElroy, S. P.; Robinson, D. A.; Smith, V. C.; Hallyburton, I.; Harrison, J. R.; Norcross, N. R.; Spinks, D.; Bayliss, T.; Norval, S.; Stojanovski, L.; Torrie, L. S.; Frearson, J. A.; Brenk, R.; Fairlamb, A. H.; Ferguson, M. A. J.; Read, K. D.; Wyatt, P. G.; Gilbert, I. H., Discovery of a Novel Class of Orally Active Trypanocidal N-Myristoyltransferase Inhibitors. *J Med Chem* **2011**, *55* (1), 140-152.
74. Roman, G.; Crandall, I. E.; Szarek, W. A., Synthesis and Anti-Plasmodium Activity of Benzimidazole Analogues Structurally Related to Astemizole. *ChemMedChem* **2013**, *8* (11), 1795-1804.
75. Rackham, M. D.; Brannigan, J. A.; Moss, D. K.; Yu, Z.; Wilkinson, A. J.; Holder, A. A.; Tate, E. W.; Leatherbarrow, R. J., Discovery of Novel and Ligand-Efficient Inhibitors of *Plasmodium falciparum* and *Plasmodium vivax* N-Myristoyltransferase. *J Med Chem* **2013**, *56* (1), 371-375.
76. Yu, Z. Y.; Brannigan, J. A.; Moss, D. K.; Brzozowski, A. M.; Wilkinson, A. J.; Holder, A. A.; Tate, E. W.; Leatherbarrow, R. J., Design and Synthesis of Inhibitors of *Plasmodium falciparum* N-Myristoyltransferase, A Promising Target for Antimalarial Drug Discovery. *J Med Chem* **2012**, *55* (20), 8879-8890.
77. Goncalves, V.; Brannigan, J. A.; Whalley, D.; Ansell, K. H.; Saxty, B.; Holder, A. A.; Wilkinson, A. J.; Tate, E. W.; Leatherbarrow, R. J., Discovery of *Plasmodium vivax* N-Myristoyltransferase Inhibitors: Screening, Synthesis, and Structural Characterization of their Binding Mode. *J Med Chem* **2012**, *55* (7), 3578-3582.

78. Rackham, M. D.; Yu, Z.; Brannigan, J. A.; Heal, W. P.; Paape, D.; Barker, K. V.; Wilkinson, A. J.; Smith, D. F.; Leatherbarrow, R. J.; Tate, E. W., Discovery of high affinity inhibitors of Leishmania donovani N-myristoyltransferase. *Medchemcomm* **2015**, 6 (10), 1761-1766.
79. Towler, D. A.; Gordon, J. I.; Adams, S. P.; Glaser, L., The Biology and Enzymology of Eukaryotic Protein Acylation. *Annu. Rev. Biochem.* **1988**, 57 (1), 69-97.
80. Bell, A., N-Myristoyltransferase as a Drug Target: A (Chemical) Space Odyssey. *Thesis* **2015**.
81. Devadas, B.; S. Kishore, N.; P. Adams, S.; I. Gordon, J., *Synthesis of myristic acid analogs with anti-HIV activity that may be resistant to metabolic processing by β -oxidation*. 1993; Vol. 3, p 779-784.
82. Paige, L. A.; Zheng, G. Q.; DeFrees, S. A.; Cassady, J. M.; Geahlen, R. L., S-(2-oxopentadecyl)-CoA, a nonhydrolyzable analog of myristoyl-CoA, is a potent inhibitor of myristoyl-CoA: protein N-myristoyltransferase. *J Med Chem* **1989**, 32 (8), 1665-1667.
83. Paige, L. A.; Zheng, G. Q.; DeFrees, S. A.; Cassady, J. M.; Geahlen, R. L., Metabolic activation of 2-substituted derivatives of myristic acid to form potent inhibitors of myristoyl CoA:protein N-myristoyltransferase. *Biochemistry* **1990**, 29 (46), 10566-10573.
84. M. Neder, K.; A. French, S.; P.F. Miller, S., *ChemInform Abstract: Synthesis and Inhibitory Activity of Difluoroketone Substrate Analogues of N-Myristoyltransferase*. 1995; Vol. 50, p 9847-9864.
85. Guo-Qiang, Z.; Xiufeng, H.; M., C. J.; A., P. L.; L., G. R., Synthesis of myristoyl CoA analogues and myristoyl peptides as inhibitors of myristoyl CoA: Protein N-myristoyltransferase. *Journal of Pharmaceutical Sciences* **1994**, 83 (2), 233-238.
86. Rackham, M., Design and Synthesis of N-Myristoyltransferase Inhibitors: A Promising Treatment for Malaria. *Thesis* **2012**.
87. Wiegand, R. C.; Carr, C.; Minnerly, J. C.; Pauley, A. M.; Carron, C. P.; Langner, C. A.; Duronio, R. J.; Gordon, J. I., The Candida albicans myristoyl-CoA:protein N-myristoyltransferase gene. Isolation and expression in Saccharomyces cerevisiae and Escherichia coli. *J Biol Chem* **1992**, 267 (12), 8591-8.
88. Robin, W.; Charles, M.; Sandra, F.; David, W.; Jeffrey, G.; Stephen, L., Genetic studies reveal that myristoylCoA:protein N - myristoyltransferase is an essential enzyme in Candida albicans. *Molecular Microbiology* **1995**, 16 (2), 241-250.
89. Rudnick, D. A.; Rocque, W. J.; McWherter, C. A.; Toth, M. V.; Jackson-Machelski, E.; Gordon, J. I., Use of photoactivatable peptide substrates of Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase (Nmt1p) to characterize a myristoyl-CoA-Nmt1p-peptide ternary complex and to provide evidence for an ordered reaction mechanism. *Proceedings of the National Academy of Sciences* **1993**, 90 (3), 1087-1091.
90. Brown, D. L.; Devadas, B.; Lu, H.-F.; Nagarajan, S.; Zupec, M. E.; Freeman, S. K.; McWherter, C. A.; Getman, D. P.; Sikorski, J. A., Replacements for lysine in l-seryl-l-lysyl dipeptide amide inhibitors of candida albicans myristoyl-CoA:protein N-myristoyltransferase. *Bioorganic & medicinal chemistry letters* **1997**, 7 (3), 379-382.

91. Devadas, B.; Freeman, S. K.; McWherter, C. A.; Kuneman, D. W.; Vinjamoori, D. V.; Sikorski, J. A., A chiral recognition element with an unusual size constraint affects the potency and selectivity for peptidomimetic inhibitors of *Candida albicans* myristoyl-CoA:protein N-myristoyltransferase. *Bioorganic & medicinal chemistry letters* **1996**, 6 (16), 1977-1982.
92. Devadas, B.; Freeman, S. K.; Zupec, M. E.; Lu, H.-F.; Nagarajan, S. R.; Kishore, N. S.; Lodge, J. K.; Kuneman, D. W.; McWherter, C. A.; Vinjamoori, D. V.; Getman, D. P.; Gordon, J. I.; Sikorski, J. A., Design and Synthesis of Novel Imidazole-Substituted Dipeptide Amides as Potent and Selective Inhibitors of *Candida albicans* MyristoylCoA:Protein N-Myristoyltransferase and Identification of Related Tripeptide Inhibitors with Mechanism-Based Antifungal Activity. *J Med Chem* **1997**, 40 (16), 2609-2625.
93. McWherter, C. A.; Rocque, W. J.; Zupec, M. E.; Freeman, S. K.; Brown, D. L.; Devadas, B.; Getman, D. P.; Sikorski, J. A.; Gordon, J. I., Scanning alanine mutagenesis and de-peptidization of a *Candida albicans* myristoyl-CoA:protein N-myristoyltransferase octapeptide substrate reveals three elements critical for molecular recognition. *J Biol Chem* **1997**, 272 (18), 11874-80.
94. Nagarajan, S. R.; Devadas, B.; Zupec, M. E.; Freeman, S. K.; Brown, D. L.; Lu, H. F.; Mehta, P. P.; Kishore, N. S.; McWherter, C. A.; Getman, D. P.; Gordon, J. I.; Sikorski, J. A., Conformationally constrained [p-(omega-aminoalkyl)phenacetyl]-L-seryl-L-lysyl dipeptide amides as potent peptidomimetic inhibitors of *Candida albicans* and human myristoyl-CoA:protein N-myristoyl transferase. *J Med Chem* **1997**, 40 (10), 1422-38.
95. Sikorski, J. A.; Devadas, B.; Zupec, M. E.; Freeman, S. K.; Brown, D. L.; Lu, H. F.; Nagarajan, S.; Mehta, P. P.; Wade, A. C.; Kishore, N. S.; Bryant, M. L.; Getman, D. P.; McWherter, C. A.; Gordon, J. I., Selective peptidic and peptidomimetic inhibitors of *Candida albicans* myristoylCoA: protein N-myristoyltransferase: a new approach to antifungal therapy. *Biopolymers* **1997**, 43 (1), 43-71.
96. Georgopapadakou, N. H., Antifungals targeted to protein modification: focus on protein N-myristoyltransferase. *Expert opinion on investigational drugs* **2002**, 11 (8), 1117-25.
97. Ebara, S.; Naito, H.; Nakazawa, K.; Ishii, F.; Nakamura, M., FTR1335 is a novel synthetic inhibitor of *Candida albicans* N-myristoyltransferase with fungicidal activity. *Biol Pharm Bull* **2005**, 28 (4), 591-5.
98. Sogabe, S.; Masubuchi, M.; Sakata, K.; Fukami, T. A.; Morikami, K.; Shiratori, Y.; Ebiike, H.; Kawasaki, K.; Aoki, Y.; Shimma, N.; D'Arcy, A.; Winkler, F. K.; Banner, D. W.; Ohtsuka, T., Crystal Structures of *Candida albicans* N-Myristoyltransferase with Two Distinct Inhibitors. *Chemistry & Biology* **2002**, 9 (10), 1119-1128.
99. Weston, S. A.; Camble, R.; Colls, J.; Rosenbrock, G.; Taylor, I.; Egerton, M.; Tucker, A. D.; Tunnicliffe, A.; Mistry, A.; Mancina, F.; de la Fortelle, E.; Irwin, J.; Bricogne, G.; Pauptit, R. A., Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat Struct Biol* **1998**, 5 (3), 213-21.
100. Masubuchi, M.; Ebiike, H.; Kawasaki, K.; Sogabe, S.; Morikami, K.; Shiratori, Y.; Tsujii, S.; Fujii, T.; Sakata, K.; Hayase, M.; Shindoh, H.; Aoki, Y.; Ohtsuka, T.; Shimma, N., Synthesis and biological activities of benzofuran antifungal agents targeting fungal N-myristoyltransferase. *Bioorg Med Chem* **2003**, 11 (20), 4463-78.

101. Gunaratne, R. S.; Sajid, M.; Ling, I. T.; Tripathi, R.; Pachebat, J. A.; Holder, A. A., Characterization of N-myristoyltransferase from *Plasmodium falciparum*. *Biochemical Journal* **2000**, *348* (Pt 2), 459-463.
102. Price, H. P.; Menon, M. R.; Panethymitaki, C.; Goulding, D.; McKean, P. G.; Smith, D. F., Myristoyl-CoA:protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J Biol Chem* **2003**, *278* (9), 7206-14.
103. Panethymitaki, C.; Bowyer, P. W.; Price, H. P.; Leatherbarrow, R. J.; Brown, K. A.; Smith, D. F., Characterization and selective inhibition of myristoyl-CoA:protein N-myristoyltransferase from *Trypanosoma brucei* and *Leishmania major*. *The Biochemical journal* **2006**, *396* (2), 277-85.
104. Roberts, A. J.; Torrie, L. S.; Wyllie, S.; Fairlamb, A. H., Biochemical and genetic characterization of *Trypanosoma cruzi* N-myristoyltransferase. *The Biochemical journal* **2014**, *459* (2), 323-32.
105. Galvin, B. D.; Li, Z.; Villemaine, E.; Poole, C. B.; Chapman, M. S.; Pollastri, M. P.; Wyatt, P. G.; Carlow, C. K. S., A Target Repurposing Approach Identifies N-myristoyltransferase as a New Candidate Drug Target in Filarial Nematodes. *PLoS neglected tropical diseases* **2014**, *8* (9), e3145.
106. Fang, W.; Robinson, D. A.; Raimi, O. G.; Blair, D. E.; Harrison, J. R.; Lockhart, D. E.; Torrie, L. S.; Ruda, G. F.; Wyatt, P. G.; Gilbert, I. H.; van Aalten, D. M., N-myristoyltransferase is a cell wall target in *Aspergillus fumigatus*. *ACS Chem Biol* **2015**, *10* (6), 1425-34.
107. Brand, S.; Norcross, N. R.; Thompson, S.; Harrison, J. R.; Smith, V. C.; Robinson, D. A.; Torrie, L. S.; McElroy, S. P.; Hallyburton, I.; Norval, S.; Scullion, P.; Stojanovski, L.; Simeons, F. R.; van Aalten, D.; Frearson, J. A.; Brenk, R.; Fairlamb, A. H.; Ferguson, M. A.; Wyatt, P. G.; Gilbert, I. H.; Read, K. D., Lead optimization of a pyrazole sulfonamide series of *Trypanosoma brucei* N-myristoyltransferase inhibitors: identification and evaluation of CNS penetrant compounds as potential treatments for stage 2 human African trypanosomiasis. *J Med Chem* **2014**, *57* (23), 9855-69.
108. Gelb, M. H.; Van Voorhis, W. C.; Buckner, F. S.; Yokoyama, K.; Eastman, R.; Carpenter, E. P.; Panethymitaki, C.; Brown, K. A.; Smith, D. F., Protein farnesyl and N-myristoyl transferases: piggy-back medicinal chemistry targets for the development of antitrypanosomatid and antimalarial therapeutics. *Mol Biochem Parasitol* **2003**, *126* (2), 155-63.
109. Yu, Z.; Brannigan, J. A.; Rangachari, K.; Heal, W. P.; Wilkinson, A. J.; Holder, A. A.; Leatherbarrow, R. J.; Tate, E. W., Discovery of pyridyl-based inhibitors of *Plasmodium falciparum* N-myristoyltransferase. *Medchemcomm* **2015**, *6* (10), 1767-1772.
110. Goncalves, V.; Brannigan, J. A.; Laporte, A.; Bell, A. S.; Roberts, S. M.; Wilkinson, A. J.; Leatherbarrow, R. J.; Tate, E. W., Structure-guided optimization of quinoline inhibitors of *Plasmodium* N-myristoyltransferase. *Medchemcomm* **2017**, *8* (1), 191-197.
111. Hutton, J. A.; Goncalves, V.; Brannigan, J. A.; Paape, D.; Wright, M. H.; Waugh, T. M.; Roberts, S. M.; Bell, A. S.; Wilkinson, A. J.; Smith, D. F.; Leatherbarrow, R. J.; Tate, E. W., Structure-based design of potent and selective leishmania N-myristoyltransferase inhibitors. *J Med Chem* **2014**, *57* (20), 8664-70.

112. Gowda, S.; Shrivastav, A.; Selvakumar, P.; Dimmock, J. R.; Sharma, R. K., A novel inhibitor protein of N-myristoyltransferase from Escherichia coli. *Biochem Biophys Res Commun* **2004**, 314 (4), 984-7.
113. Shrivastav, A.; Pasha, M. K.; Selvakumar, P.; Gowda, S.; Olson, D. J.; Ross, A. R.; Dimmock, J. R.; Sharma, R. K., Potent inhibitor of N-myristoylation: a novel molecular target for cancer. *Cancer research* **2003**, 63 (22), 7975-8.
114. Eric Lee, E. C., Edward Tate, Targeting human N-myristoyltransferase 1 using fragment based drug discovery approach. *Thesis* **2014**.
115. Lo, M. C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G., Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **2004**, 332 (1), 153-9.
116. Goncalves, V.; Brannigan, J. A.; Thinon, E.; Olaleye, T. O.; Serwa, R.; Lanzarone, S.; Wilkinson, A. J.; Tate, E. W.; Leatherbarrow, R. J., A fluorescence-based assay for N-myristoyltransferase activity. *Analytical Biochemistry* **2012**, 421 (1), 342-344.
117. Lacal, P. M.; Pennington, C. Y.; Lacal, J. C., Transforming activity of ras proteins translocated to the plasma membrane by a myristoylation sequence from the src gene product. *Oncogene* **1988**, 2 (6), 533-7.
118. McIlhinney, R. A. J.; Patel, P. B.; McGlone, K., Characterization of a polyhistidine-tagged form of human myristoyl-CoA: protein N-myristoyltransferase produced in Escherichia coli. *Eur J Biochem* **1994**, 222 (1), 137-146.
119. Hill, A. V., The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *The Journal of Physiology* **1910**, 40, i--vii.
120. Xu, X.; Qian, X.; Li, Z.; Song, G.; Chen, W., Synthesis and fungicidal activity of fluorine-containing phenylimino-thiazolidines derivatives. *Journal of Fluorine Chemistry* **2004**, 125 (7), 1159-1162.
121. Marvin was used for drawing, displaying and characterizing chemical structures, substructures and reactions. *Marvin 17.23.0, ChemAxon* (<http://www.chemaxon.com>) **2017**.
122. Chubb, N. A. L.; Critcher, D. J.; Eshelby, J. J.; Lunn, G.; Rudge, A. J.; Walshe, N. D.; Wiedenau, P. H.; Williams, D. H., Antiparasitic terpene alkaloids. Google Patents: 2004.
123. Hunter, C. A.; Sanders, J. K. M., The nature of .pi.-.pi. interactions. *Journal of the American Chemical Society* **1990**, 112 (14), 5525-5534.
124. Promega, *Technical Bulletin* **2012**, No. TB169.
125. Lim, S. G., Elucidating the roles of N-myristoyltransferase in cancers. *Thesis* **2016**.
126. <https://docs.chemaxon.com/display/docs/LogP+and+logD+calculations>.

127. Dieterich, D. C.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M., Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proceedings of the National Academy of Sciences* **2006**, *103* (25), 9482-9487.
128. Heal, W. P.; Wickramasinghe, S. R.; Leatherbarrow, R. J.; Tate, E. W., N-Myristoyl transferase-mediated protein labelling in vivo. *Organic & biomolecular chemistry* **2008**, *6* (13), 2308-15.
129. Dieterich, D. C.; Hodas, J. J. L.; Gouzer, G.; Shadrin, I. Y.; Ngo, J. T.; Triller, A.; Tirrell, D. A.; Schuman, E. M., In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nature neuroscience* **2010**, *13* (7), 897-905.
130. Lymphatic filariasis, Fact sheets. *WHO media centre* **2017**, 102.
131. WHO, What is lymphatic filariasis. http://www.who.int/lymphatic_filariasis/disease/en/ **2017**.
132. Onchocerciasis, Fact sheets. *WHO media centre* **2017**, 374.
133. Etyaale, D., Vision 2020: update on onchocerciasis. *Comm Eye Health* **2001**, *14* (38), 19-21.
134. CDC, Biology - Life cycle of *Brugia malayi*. https://www.cdc.gov/parasites/lymphaticfilariasis/biology_b_malayi.html **2017**.
135. CDC, Onchocerciasis. <https://www.cdc.gov/dpdx/onchocerciasis/index.html> **2017**.
136. Nana-Djeunga, H.; Bourguinat, C.; Pion, S. D.; Kamgno, J.; Gardon, J.; Njiokou, F.; Boussinesq, M.; Prichard, R. K., Single nucleotide polymorphisms in beta-tubulin selected in *Onchocerca volvulus* following repeated ivermectin treatment: possible indication of resistance selection. *Mol Biochem Parasitol* **2012**, *185* (1), 10-8.
137. Brand, S.; Cleghorn, L. A.; McElroy, S. P.; Robinson, D. A.; Smith, V. C.; Hallyburton, I.; Harrison, J. R.; Norcross, N. R.; Spinks, D.; Bayliss, T.; Norval, S.; Stojanovski, L.; Torrie, L. S.; Frearson, J. A.; Brenk, R.; Fairlamb, A. H.; Ferguson, M. A.; Read, K. D.; Wyatt, P. G.; Gilbert, I. H., Discovery of a novel class of orally active trypanocidal N-myristoyltransferase inhibitors. *J Med Chem* **2012**, *55* (1), 140-52.
138. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **2017**, *45* (D1), D158-D169.
139. Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezuk, Y.; Raytselis, Y.; Sayers, E. W.; Tao, T.; Ye, J.; Zaretskaya, I., BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* **2013**, *41* (Web Server issue), W29-33.
140. Target repurposing for neglected diseases. *Future Medicinal Chemistry* **2011**, *3* (10), 1307-1315.
141. Wormatlas, Introduction to *C. elegans* anatomy. <http://www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.htm> **2017**.

142. Wood, W. B., The nematode *C. elegans*. *Cold Spring Harbor Laboratory Press, New York* **1988**, 337-391.
143. Baron, K. H. a. M., BoxShade 3.21. https://embnet.vital-it.ch/software/BOX_form.html **2017**.
144. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **2005**, 41 (1), 207-34.
145. Adapted from Handbooks from GE healthcare: Gel filtration principles and methods. https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma-Aldrich/General_Information/1/ge-gel-filtration.pdf.
146. Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **2010**, 43 (1), 186-190.
147. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G., Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D: Biological Crystallography* **2011**, 67 (Pt 4), 293-302.
148. Kabsch, W., XDS. *Acta Crystallographica Section D: Biological Crystallography* **2010**, 66 (Pt 2), 125-132.
149. Waterman, D. G.; Winter, G.; Gildea, R. J.; Parkhurst, J. M.; Brewster, A. S.; Sauter, N. K.; Evans, G., Diffraction-geometry refinement in the DIALS framework. *Acta Crystallographica Section D* **2016**, 72 (4), 558-575.
150. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography* **2013**, 69 (Pt 7), 1204-1214.
151. Kabsch, W., Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D* **2010**, 66 (2), 133-144.
152. Evans, P., Scaling and assessment of data quality. *Acta Crystallographica Section D* **2006**, 62 (1), 72-82.
153. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography* **2011**, 67 (Pt 4), 271-281.
154. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007**, 40 (4), 658-674.
155. 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 Crystallographica Section D: Biological Crystallography* **2011**, 67 (Pt 4), 235-242.
156. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **2004**, 60 (12 Part 1), 2126-2132.

157. Adams, P. D.; Afonine, P. V.; Bunkoczi, 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 Crystallographica Section D* **2010**, *66* (2), 213-221.
158. Nicholls, R. A.; Long, F.; Murshudov, G. N., Low-resolution refinement tools in REFMAC5. *Acta Crystallographica Section D: Biological Crystallography* **2012**, *68* (Pt 4), 404-417.
159. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M., Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics* **1992**, *12* (4), 345-364.
160. Vedadi, M.; Niesen, F. H.; Allali-Hassani, A.; Fedorov, O. Y.; Finerty, P. J.; Wasney, G. A.; Yeung, R.; Arrowsmith, C.; Ball, L. J.; Berglund, H.; Hui, R.; Marsden, B. D.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A. M., Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences* **2006**, *103* (43), 15835-15840.
161. Laskowski, R. A.; Swindells, M. B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **2011**, *51* (10), 2778-86.
162. Rackham, M. D.; Brannigan, J. A.; Rangachari, K.; Meister, S.; Wilkinson, A. J.; Holder, A. A.; Leatherbarrow, R. J.; Tate, E. W., Design and Synthesis of High Affinity Inhibitors of Plasmodium falciparum and Plasmodium vivax N-Myristoyltransferases Directed by Ligand Efficiency Dependent Lipophilicity (LELP). *J Med Chem* **2014**, *57* (6), 2773-2788.
163. Chang, C.-e. A.; Chen, W.; Gilson, M. K., Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences* **2007**, *104* (5), 1534-1539.
164. Powell, H. R., X-ray data processing. *Biosci Rep* **2017**, *37* (5), 1-14.
165. Powell, H., Integration and Scaling. https://www2.mrc-lmb.cam.ac.uk/download/lectures/lmb_crystallography_course_2013/6_Integration_and_Scaling_-_Harry-Powell.pdf.
166. Steller, I.; Bolotovskiy, R.; Rossmann, M. G., An Algorithm for Automatic Indexing of Oscillation Images using Fourier Analysis. *Journal of Applied Crystallography* **1997**, *30* (6), 1036-1040.
167. Powell, H. R., The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **1999**, *55* (Pt 10), 1690-5.
168. de Wolff, P. M., Reduced bases. In *International Tables for Crystallography Volume A: Space-group symmetry*, Hahn, T., Ed. Springer Netherlands: Dordrecht, 2002; pp 750-755.
169. Debreczeni, J. É., Data Processing. http://www.ccp4.ac.uk/schools/China-2011/talks/JED-Data_processing.pdf.
170. Leslie, A. G. W., Data collection. https://www-structmed.cimr.cam.ac.uk/Course/Basic_collection/Datacol.html **2009**.

7 Appendix

7.1 Permission documents for figure reproduction

7.1.1 Figure 1



Thank you for your order!

Dear Ms. Mengjie Shen,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: Mengjie Shen
Order Date: Jun 3, 2019
Order Number: 4601540911505
Publication: Cell Chemical Biology
Title: Protein Lipidation in Cell Signaling and Diseases: Function, Regulation, and Therapeutic Opportunities
Type of Use: reuse in a thesis/dissertation
Order Total: 0.00 GBP

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
<https://myaccount.copyright.com>



RightsLink®

7.1.2 Figure 3



Thank you for your order!

Dear Ms. Mengjie Shen,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: Mengjie Shen
Order Date: Jun 11, 2019
Order Number: 4606091237627
Publication: Current Opinion in Chemical Biology
Title: Recent progress in synthetic and biological studies of GPI anchors and GPI-anchored proteins
Type of Use: reuse in a thesis/dissertation
Order Total: 0.00 GBP

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customer@copyright.com
<https://myaccount.copyright.com>



Copyright
Clearance
Center

RightsLink®

7.1.3 Figure 28

Thank you for your order!

Dear Ms. Mengjie Shen,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

Licensee: Mengjie Shen
Order Date: Nov 26, 2018
Order Number: 4476701163351
Publication: Analytical Biochemistry
Title: A fluorescence-based assay for N-myristoyltransferase activity
Type of Use: reuse in a thesis/dissertation
Order Total: 0.00 GBP

View or print complete [details](#) of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Tel: +1-855-239-3415 / +1-978-646-2777
customercare@copyright.com
<https://myaccount.copyright.com>



RightsLink®

7.1.4 Figure 55 and Figure 56

From: Carlow, Tilde <carlow@neb.com>
Sent: 05 September 2017 15:47
To: Tate, Ed W
Cc: Bell, Andy; Li, Zhiru
Subject: Re: NMT inhibitors, B. malayi

Hi Ed

Nice to hear from you. Are you referring to the data that you sent to us in April 2016 or have there been any further advances? If you have new compounds that are more potent are you able to send along the enzyme data? Liz will be willing to test these compounds in *C. elegans*. It would be great to see this work result in a publication in the future!

I see no problem including the enzyme work and *C. elegans* viability data in the thesis.

Good luck with the viva!

Take care,

Tilde

Clotilde Carlow

New England Biolabs
Tel: 978.380.7263
<mailto:carlow@neb.com>

From: "Tate, Ed W" <e.tate@imperial.ac.uk>
Date: Saturday, September 2, 2017 at 6:30 PM
To: "Carlow, Tilde" <carlow@neb.com>
Cc: "Bell, Andy" <a.bell11@imperial.ac.uk>
Subject: NMT inhibitors, B. malayi

Dear Tilde,

I hope all is well with you? It is about 18 months since we were last in touch about testing novel NMT inhibitors on *C. elegans* and *B. malayi*. Jessica (the PhD student in our lab who characterised the enzymes you sent, and measured IC50s for compounds) is about to complete her PhD thesis, and she would like to include a brief chapter describing her work on this collaboration. This would include at minimum presenting work using the enzymes you provided (which I think were already disclosed in your PLoS paper?), and to put the work in context she would also like to present the parasite viability data.

The data will of course be credited appropriately in the thesis, but the thesis will also be placed in the public domain (open access) at some point in the future, so we must ask your permission to include these data. The thesis will have an embargo of at least 1 year, potentially 2 years. In general, presenting the data in a thesis does not preclude future publication in a journal (although we have no plans for that at present), and I don't believe there is any IP-sensitive information involved.

Please could you let me know as soon as you are able whether this would be OK by you? Also, if you have any interest in testing more inhibitors, we have new molecules with likely improved bioavailability which may have the potential for enhanced whole-worm activity.

Best wishes,

Ed.