JOHANNES GUTENBERG UNIVERSITY

DOCTORAL THESIS

Selectivity Determining Features in *N*-Myristoyltransferases – A Model System for Drug Targets with conserved Binding Sites

Author:
Frank Christian KERSTEN

Supervisors:
[redacted]

A thesis submitted in fulfillment of the requirements for the degree of Dr. rer. nat.

in

Pharmaceutical Chemistry
at the Institute of Pharmacy and Biochemistry

August 1, 2017

Selectivity Determining Features in *N*-Myristoyltransferases – A Model System for Drug Targets with conserved Binding Sites

DOCTORAL THESIS

submitted in fulfillment of the regulations
for attaining the degree of

"Doktor der Naturwissenschaften (Dr. rer. nat)"
in the subject of Pharmacy

at the Faculty of Chemistry, Pharmacy and Geoscience
at the Johannes Gutenberg University of Mainz

Frank Christian KERSTEN

born in Wiesbaden

Mainz, June, 25, 2017

Dean: [redacted]

Name of 1st reviewer: [redacted]
Name of 2nd reviewer: [redacted]

Date of doctoral examination: **27.07.2017**

Declaration of Authorship

I, Frank Christian KERSTEN, declare that this thesis titled, "Selectivity Determining Features in *N*-Myristoyltransferases – A Model System for Drug Targets with conserved Binding Sites" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the Johannes Gutenberg University, Mainz and the University in Bergen.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

"...if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms."

Richard Feynman

JOHANNES GUTENBERG UNIVERSITY

Abstract

Department of Medicinal and Pharmaceutical Chemistry
at the Institute of Pharmacy and Biochemistry

Dr. rer. nat.

Selectivity Determining Features in *N*-Myristoyltransferases – A Model System for Drug Targets with conserved Binding Sites

by Frank Christian KERSTEN

In rational drug design projects one major goal is to obtain high affinity ligands for a target, while maintaining selectivity over potential off-targets and thereby reducing unwanted side effects. However, this proves to be challenging when facing highly conserved binding sites. In this project, the focus was laid on a model system of two well investigated essential eukaryotic enzymes, which catalyze the transfer of myristate to a various number of substrates – the *N*-myristoyltransferases (NMT) of *Leishmania major* (*Lm*NMT) and the human homologue *Hs*NMT1. The enzymes share an overall sequence identity of over 40 % and an identical first protein-ligand interaction shell. Many non-selective NMT-inhibitors were found previously, but only a few selective ones are known. Further, the molecular basis for selective inhibition was unclear. A combination of molecular dynamic simulations (MDs), isothermal titration calorimetry (ITC), fluorescence-based enzyme inhibition assay and X-ray crystallography was used to analyze protein dynamics, water network formation and their changes upon ligand binding. Two different selectivity determining features were identified and validated by site-directed mutagenesis – the impairment of protein flexibility upon ligand binding close to the catalytically active C-terminus and a highly stable water molecule, only present in the binding site of *Hs*NMT1. Based on these findings, a virtual screening was conducted and three novel and selective *Lm*NMT inhibitors were revealed.

JOHANNES GUTENBERG UNIVERSITÄT

Zusammenfassung

Medizinische und Pharmazeutische Chemie
am Institut für Pharmazie und Biochemie

Dr. rer. nat.

Selektivitätsbestimmende Eigenschaften von *N*-Myristoyltransferasen - ein Modellsystem für Arzneistoff Zielstrukturen mit konservierten Bindetaschen

von Frank Christian KERSTEN

Ein Hauptziel in Projekten des rationalen Wirkstoffdesigns ist es, Liganden zu identifizieren, welche eine hohe Affinität zu einer Zielstruktur aufweisen und gleichzeitig Selektivität gegenüber potentiellen "off-targets" zeigen, um unerwünschte Nebenwirkungen zu reduzieren. Dies stellt im Falle hochkonservierter Bindetaschen jedoch eine Herausforderung dar. Der Fokus dieser Arbeit wurde auf ein Modellsystem zweier gut untersuchter, essentieller eukaryotischer Enzyme gelegt, welche den Transfer von Myristat auf verschiedene Substrate katalysieren – die *N*-Myristoyltransferasen (NMT) von *Leishmania major* (*Lm*NMT) und das humane Homologon *Hs*NMT1. Beide Enzyme weisen eine Sequenzidentität von über 40 % und identische direkte Protein-Ligand-Interaktionen auf. Bislang wurden viele unselektive, aber nur wenige selektive NMT-Inhibitoren identifiziert. Die molekulare Begründung im Falle einer selektiven NMT Inhibition blieb jedoch unklar. In dieser Arbeit wurde eine Kombination aus Moleküldynamik-Simulationen (MDs), isothermer Titrationskalorimetrie (ITC), einem fluoreszenzbasierten Assay und Kristallstrukturanalyse eingesetzt, um die Proteindynamik, Netzwerkbildung von Wassermolekülen und deren Veränderungen nach Ligandenbindung zu analysieren. Zwei unterschiedliche selektivtätsbestimmende Eigenschaften konnten identifiziert und mittels zielgerichteter Mutagenese validiert werden – die Einschränkung der Proteinflexibilität nahe des katalytisch aktiven C-Terminus nach Ligandenbindung einerseits, sowie ein hoch-stabil gebundenes Wassermolekül in der Bindetasche von *Hs*NMT1 andereseits. Auf Grundlage dieser Ergebnisse wurde ein virtuelles Screening (VS) durchgeführt, welches zur Entdeckung von drei neuartigen und selektiven *Lm*NMT Inhibitoren führte.

Acknowledgements

During the last few years as a doctoral candidate, I got the chance to work together with some great people and inspiring scientists, I would like to thank for their support and knowledge:

- [redacted].
- [redacted].
- [redacted].
- [redacted].
- [redacted].
- [redacted].
- [redacted].
- [redacted].

Contents

List of Abbreviations

AMBER	**A**ssisted **M**odel **B**uilding with **E**nergy **R**efinement
CADD	**C**omputer-**A**ided **D**rug **D**esign
CGenFF	**C**HARMM **Gen**eralized **F**orce **F**ield
CHARMM	**C**hemistry at **Har**vard **M**olecular **M**echanics
Cpd.	**C**om**p**oun**d**
COX	**C**yclo**ox**ygenase
CPM	7-Diethylamino-3-(4-**m**aleimido**p**henyl)-4-methyl**c**oumarin
DCCM	**D**ynamic **C**ross-**C**orrelation **M**atrix
DMSO	**Dim**ethyl **s**ulf**o**xide
ED	**E**lectron **D**ensity
FEP	**F**ree **E**nergy **P**ertubation
GAFF	**G**eneralized **A**MBER **F**orce **F**ield
GPU	**G**raphics **P**rocessing **U**nit
HA	**H**eavy **A**toms (all but hydrogen)
HAT	**H**uman **A**frican **T**rypanosomiasis (African Sleeping Sickness)
HEPES	(4-(2-**h**ydroxy**e**thyl)-1-**p**iperazine**e**thane**s**ulfonic acid
HPC	**H**igh-**P**erformance **C**omputing
Hs	***H**omo **s**apiens*
HTS	**H**igh **T**hroughput **S**creening
IC_{50}	**I**nhibitory **C**oncentration to achieve half maximum (**50**%) inhibition
IFST	**I**nhomogeneous **F**luid **S**olvation **T**heory
IPTG	**Iso**propyl β-D-1-**t**hio**g**alactopyranoside
ITC	**I**sothermal **T**itration **C**alorimetry
LE	**L**igand **E**fficiency
LJ	**L**ennard-**J**ones (potential)
Lm	*Leishmania **m**ajor*
MEP	**M**olecular **E**lectrostatic **P**otential
MetAP	**Met**hionin **A**mino **P**eptidase
MDs	**M**olecular **D**ynamic simulations
MMFF	**M**erck **M**olecular **F**orce **F**ield
MM	**M**olecular **M**echanics
MOE	**M**olecular **O**perating **E**nvironment
MPEG	**P**oly (**e**thylene **g**lycol) **m**ethyl ether
M_w	**M**olecular **W**eight
MyrCoA	**Myr**istoyl **Co**enzyme **A**
NAMD	**N**ot (just) **A**nother **M**olecular **D**ynamics program

NMT	***N*-M**yristoyl**T**ransferase
PAINS	**P**an **A**ssay **In**terference Compound**s**
PCA	**P**rincipal **C**omponent **A**nalysis
PDB	**P**rotein **D**ata **B**ank
PEG	**P**oly (**e**thylene **g**lycol)
PI3K	**P**hospho**i**nositide **3**-**k**inase
PTP	**P**rotein **T**yrosine **P**hosphatase
Pv	***P**lasmodium **v**ivax*
QM	**Q**uantum **M**echanics
RESP	**R**estrained **E**lectro**s**tatic **P**otential
S^2	Order Parameter
SAS	**S**olvent **A**ccessible **S**urface
SI	**S**electivity **I**ndex
SPR	**S**urface **P**lasmon **R**esonance
Tb	***T**rypanosoma **b**rucei*
TEV	***T**obacco **e**tch **v**irus*
vdW	**v**an **d**er **W**aals
VMD	**V**isual **M**olecular **D**ynamics
VS	**V**irtual **S**creening
wt	**W**ild **T**ype Protein

For my family.

Chapter 1

Introduction

1.1 Selectivity in Drug Design

One major goal of drug design projects is to obtain high affinity ligands for a certain target while maintaining selectivity over potential off-targets and thereby reducing side effects. This proves to be challenging when facing highly conserved binding sites. In rational approaches, the principles of electrostatic interactions or shape complementarity, specific to the target, or exploiting allosteric binding sites, not present in the off-target, are used. Further, the elucidation of binding-site hydration and protein flexibility opens possibilities for selectivity (Figure 1.1, Huggins, Sherman, and Tidor, 2012).

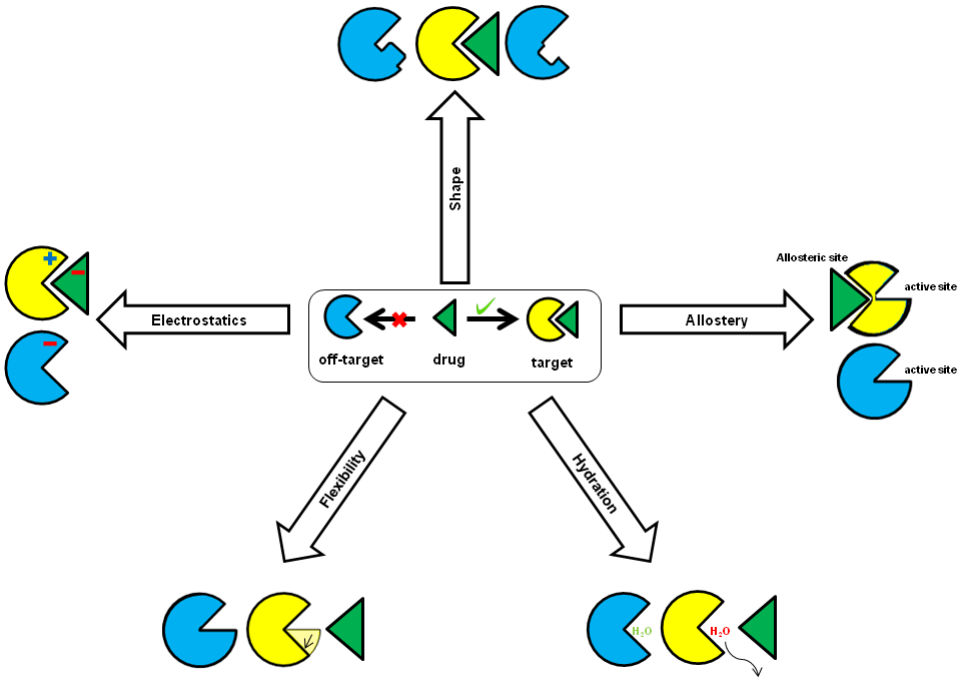


FIGURE 1.1: Strategies to increase selectivity in drug design: electrostatic interactions, shape complementarity, targeting an allosteric site, protein flexibility and hydration patterns (red: high energy "unhappy" water, green: low energy "happy" water). Modified from Huggins, Sherman, and Tidor, 2012.

Electrostatic interactions include charged groups and neutral polar groups to form hydrogen bonds (H-bonds), either directly with the target or indirectly mediated through water. A good electrostatic complementarity between target and ligand is needed as a desolvation penalty has to be compensated upon ligand binding (Lee and Tidor, 1997; Kangas and Tidor, 1998). One example for the exploitation of electrostatic interactions for selectivity is illustrated in Figure 1.2 a. While the target protein tyrosine phosphatase 1B (PTP1B) carries a negative charge on Asp48, the off-target PTPs contain an uncharged Asn at the corresponding position. This difference can be exploited by a positively charged moiety in the ligand to form a salt bridge only with the target Asp resulting in an selectivity of factor 20 (Iversen et al., 2000). Due to the orientation dependence of electrostatic interactions compared to non-polar interactions more polar or charged molecules tend to be less promiscuous than their hydrophobic counterparts (Radhakrishnan and Tidor, 2007).

Shape complementarity can cause selectivity as well. This accounts for less favourable interactions when one of the two binding sites is larger, resulting in a non-continuous contact surface between the ligand and off-target. This is even more severe in the opposite case, when a too small binding site would cause so-called *clashes* with the compound. Figure 1.2 b illustrates an example of unwanted selectivity due to shape complementarity. The I50V mutation in HIV-protease causes an affinity loss of the inhibitor darunavir from 0.2 nM to 2.0 nM due to less favourable van der Waals (vdW) interactions in the enlarged pocket (Kovalevsky et al., 2006). For selective cyclooxygenase-2 (COX-2) inhibition, one can exploit the V523I substitution in COX-1 (Figure 1.2 c). This difference results in an overlap of vdW-radii (clash) when COX-2 ligands bind to COX-1 and subsequently in an over 13 000-fold selectivity (Kurumbail et al., 1996; Biava et al., 2010; Sidhu et al., 2010).

The presence of an allosteric binding site whose crosstalk regulates the protein activity can be used for selective inhibitor design as well. The Bcr-Abl kinase inhibitor shown in Figure 1.2 d binds to the allosteric myristate pocket. This offers not only selectivity over kinases lacking this pocket, it also is a possibility to circumvent resistance arising from the active site T135I mutation (Zhang et al., 2010; Radi et al., 2016; Wylie et al., 2017).

In the absence of an allosteric binding site, the design of selective inhibitors for highly conserved active sites proves to be challenging. However, many examples of selective ligands facing a conserved binding site are known even though the underlying mechanism often remained unclear (Ikuta et al., 2001; Aliagas-Martin et al., 2009; Baba et al., 2003; Manas et al., 2004; Teng et al., 2013; Cleghorn et al., 2011). In some cases, protein flexibility or explicit solvent pocket hydration appeared to be contributing factors.

1.1.1 Binding-Site Water and the Hydrophobic Effect

Water can be denoted as the third binding partner of protein-ligand interactions with contributions to both enthalpy and entropy. This makes the simple molecule H_2O a long-standing challenge which deserves more consideration in computer-aided drug design (CADD; Ladbury, 1996; Klebe, 2006; Cheng et al., 2012; Snyder et al., 2014).

First to mention is the "classic" *hydrophobic effect*. Water molecules are forced into order around hydrophobic surfaces (like lipophilic pockets of binding sites or ligands). Upon protein-ligand binding (or aggregation of non-polar solutes), the water molecules are released resulting in a gain of entropy. However, "non-classic" effects with high enthalpic contributions have been observed as well (Snyder et al., 2011; Biela et al., 2013). These were caused by different reasons like a suboptimal solvation of the protein binding pocket, rearrangement of water networks or displacement of low-energy water molecules. Probing the binding site hydration is therefore an important point to address in structure based drug design to identify which regions to occupy for an increased binding free energy of ligand binding (Goodford, 1985; Bissantz, Kuhn, and Stahl, 2010).

The thermodynamic characterization of explicit water molecules within a binding pocket is of emerging importance. Upon ligand binding more or less tightly bound water molecules are displaced from the binding site with a corresponding higher or lower energetic effect. Besides easily displaceable water molecules, others might be highly coordinated and more likely to be considered as part of the binding pocket. Their entropy gain upon release might be more than compensated by an enthalpic penalty. On the other hand, examples are known where rather tightly bound water molecules retained high amounts of mobility within the binding site (Denisov et al., 1997; Matthews and Liu, 2009). Further, complete water network analysis can give hints to explain surprising thermodynamic signatures of non-additivity of functional groups within a series of ligands (Biela et al., 2012; Biela et al., 2013) or its contribution to enthalpy-entropy-compensation (Breiten et al., 2013). The influence of bound water on protein flexibility makes the question of which water to replace definitely a non-trivial one (Bissantz, Kuhn, and Stahl, 2010).

With this in mind, differences in location and thermodynamic profile of water molecules within similar binding sites can be a reason for selectivity. Due to the conserved ATP-binding site, kinases have been exhaustively studied to elucidate selectivity (Scheffzek et al., 1996; Knight and Shokat, 2005; Robinson, Sherman, and Farid, 2010). In case of the kinase inhibitor bosutinib (Figure 1.2 e) a H-bond from the nitrile moiety of the inhibitor mediated through the two water molecules W1 and W2 was identified to be crucial for selectivity (Levinson and Boxer, 2013). Ala403 is exchanged

with Thr in in many off-targets like EGFR tyrosine kinase, (Park et al., 2012). In this case, an additional low-energy water molecule (W3) is found that would have to be displaced by bosutinib. Further, the indirect interaction over W1 and W2 is no longer possible in EGFR tyrosine kinase.

As crystal structures do not always reveal the water-network within a binding site, much less the thermodynamic profile, computational methods to analyse binding site water energetics for kinase selectivity were occasionally used (Robinson, Sherman, and Farid, 2010; Barillari et al., 2011). These revealed which water molecules are favourable to be displaced. Also subtle changes like the destabilisation of a present water network without a real displacement of certain water molecules were shown to be an important contribution for selectivity between phosphoinositide 3-kinases (PI3K) β and δ (Robinson et al., 2016).

For a quantitative thermodynamic estimation, different theories can be used. Free energy pertubation methods (FEP) (Miyamoto and Kollman, 1993; Gouda et al., 2003) and inhomogeneous fluid solvation theory (IFST, Lazaridis, 1998) have been successfully applied to improve computational reproduction of experimental binding energies and to predict the influence of single water molecule replacement on ligand binding (Pearlman and Connelly, 1995; Tarek and Tobias, 2000; Ni, Sotriffer, and McCammon, 2001; Li and Lazaridis, 2003; Michel, Tirado-Rives, and Jorgensen, 2009a; Michel, Tirado-Rives, and Jorgensen, 2009b; Cappel et al., 2011). Maybe the most prominent method is WaterMap (Young et al., 2007; Abel et al., 2008) which is based on IFST. The entropic contribution for each hydration site is derived from explicit solvent MDs and described by the translational and rotational degrees of freedom of microscopic states for each water molecule. However, the complexity and computational demands prevented those methods to be applied routinely until recently.

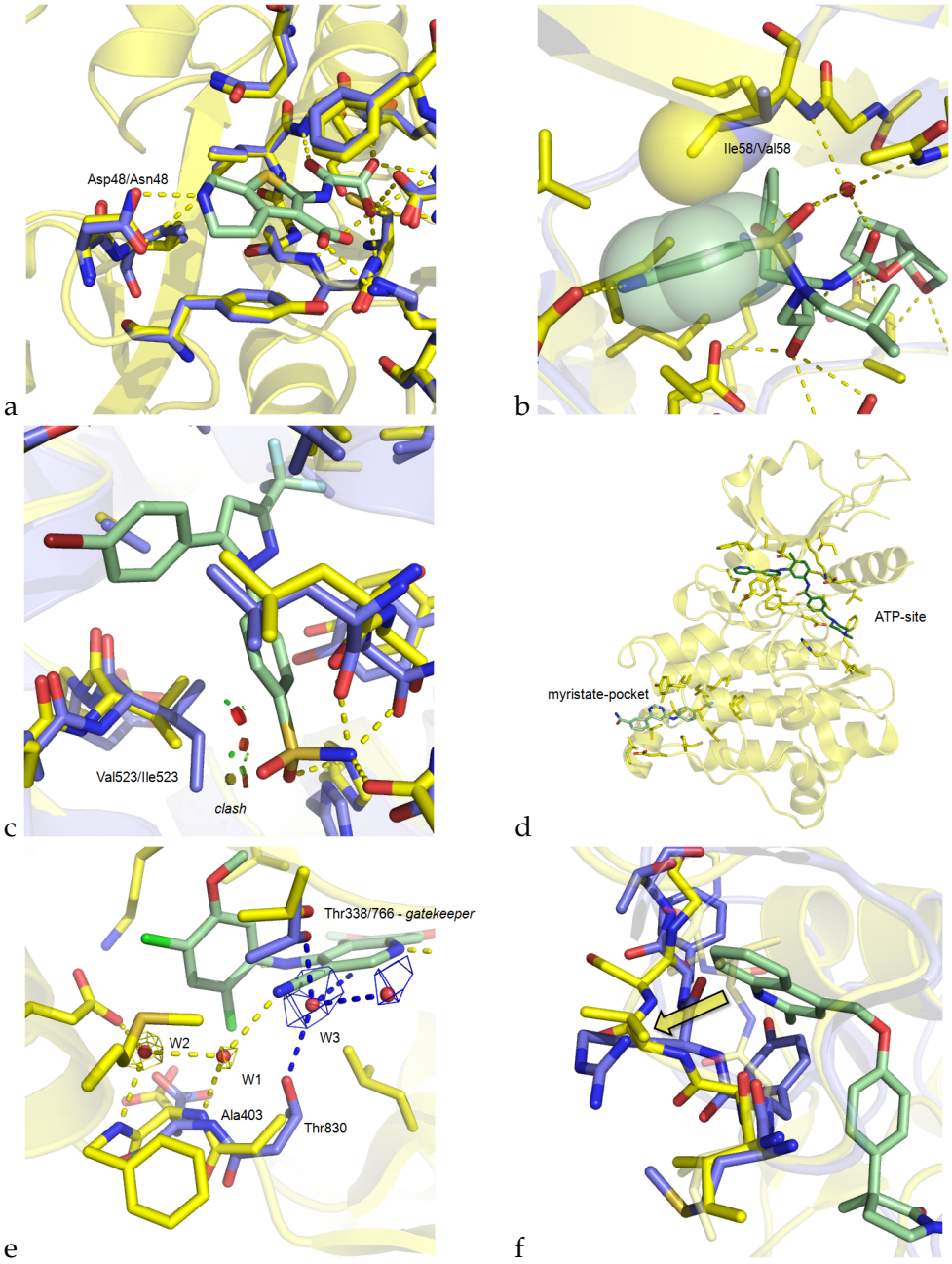


FIGURE 1.2: Selectivity strategies: a) Selectivity of PTP1B (carbon atoms colored in yellow) over other PTPs (carbon atoms colored in blue) by ionic interactions through Asp48 (PDB-IDs 1C88, 1C87). b) Selectivity caused by HIV protease I50V mutation. The mutant protein showed less favourable vdW interactions with the inhibitor darunavir, illustrated as transparent vdW spheres (PDB-IDs 2IEN, 2F8G). c) Selective COX-2 inhibition caused by a clash between Ile523 of COX-1 and the coxibe shown as vdW overlap discs (PDB-IDs 6COX, 3N8V). d) Allosteric binding site inhibitor GNF-2 (palegreen) bound to Bcr-Abl kinase (PDB-ID 3K5V) and active site inhibitor imatinib (dark green). e) Bosutinib (pale green) has a water-mediated interaction with target Src kinase (yellow carbon atoms, interaction indicated as dashed lines, and mesh, PDB-ID 4MXO) and off-target EGFR-kinase (blue carbon atoms, interactions and mesh, PDB-ID 4HJO). f) Induced fit mechanism of an inhibitor binding to TACE. Transparent pale yellow carbon atoms of unbound (PDB-ID 3KMC) and yellow carbon atoms for bound TACE (PDB-ID 2FV5) in overlay with off-target MMP-9 (blue carbon atoms, PDB-ID 2OW0).

1.1.2 Protein Flexibility

When it comes to protein-ligand binding the probably most used metaphor is *lock-and-key*, first postulated by Emil Fischer in 1894 (Fischer, 1894). However, today this picture is expanded by the knowledge of the flexibility of protein and ligand described as a *hand-in-glove* or *induced fit* model (Koshland, 1958). Another model is the *conformational selection* or *population-shift* model. Therein different possible conformations (e.g. active and inactive) are in equilibrium that is shifted upon ligand binding (Kumar et al., 2000, Figure 1.3).

The incorporation of protein flexibility and entropic effects in CADD is an emerging field (Carlson and McCammon, 2000; Teague, 2003; Alonso, Bliznyuk, and Gready, 2006). Flexibility is an intrinsic feature of proteins and can play an essential role in protein-ligand recognition. Thus, virtual screening (VS) benefits from modelling the protein flexibility. The methodological approaches therein account for different degrees of receptor flexibility ranging from so-called soft-docking over the use of different conformations (also referred to as conformational sampling or ensemble docking) up to real flexible target treatment (Antunes, Devaurs, and Kavraki, 2015). Molecular dynamic simulations (MDs) are a useful tool in this context as they help to understand structure and function of biological macromolecules, as well as underlying crosstalk in allostery (Leioatts et al., 2014) and also to cover the conformational space (Karplus and McCammon, 2002; Karplus and Kuriyan, 2005). Their implementation into drug design in initial or final stages led to successfully identified ligands, but the computational cost is still an issue. However, the constant improvements in computer power, algorithm design (parallelization) and force field parameters make MDs likely to play an increasingly important role in CADD (Durrant and McCammon, 2011; Zhao and Caflisch, 2015; De Vivo et al., 2016).

Again the field of kinases offers insights into selectivity derived from protein flexibility (Knight and Shokat, 2005; Thaimattam et al., 2007; "Protein Kinase Inhibitors : Insights into Drug Design from Structure"). The elucidation of structural rearrangement of the so called DFG-loop in a subset of kinases led to the development of DFG-out state (which is inactive) binding inhibitors like imatinib (Schindler et al., 2000). Also other examples are known like inhibitors for the TNF-α converting enzyme (TACE; Duan et al., 2002) in which selectivity over off-target metalloproteases (MMPs) was achieved by addressing a flexible loop of the target with an quinolone moiety in the ligand (Figure 1.2 f). The larger residues in off-target MMPs prevent this induced-fit mechanism as substantial rearrangement would be necessary (Niu et al., 2006; Alcaraz et al., 2007; Tochowicz et al., 2007; Rosner et al., 2010).

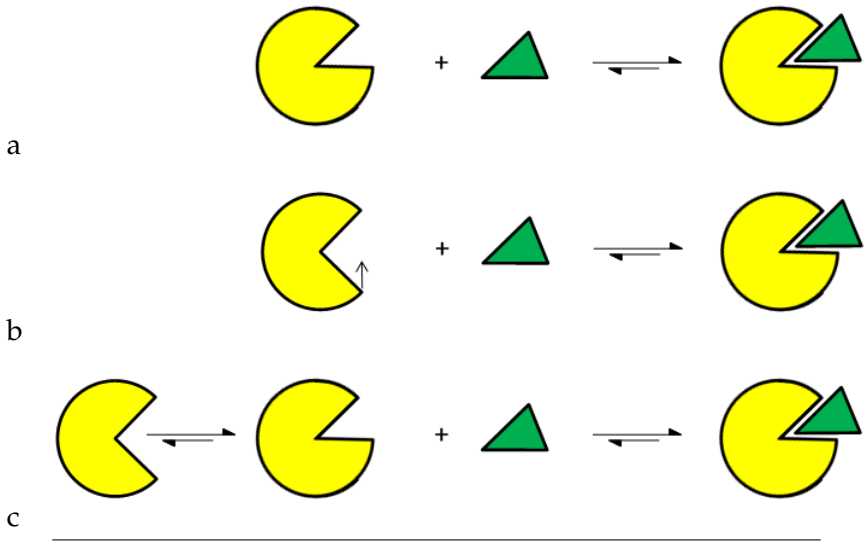


FIGURE 1.3: Illustration of protein-ligand binding models. a) static *lock-and-key*. b) *induced-fit*. c) *conformational selection*.

Besides those large domain movements, conformational changes can follow less obvious dynamics, like the flip of a single residue. This already is able to largely influence ligand binding as the H-bond donor/acceptor recognition pattern is altered (Fitzgerald et al., 2003; Brenk et al., 2003; Stengl et al., 2007). Even more subtle changes in flexibility of distal regions becoming more flexible or rigid have been reported (MacRaild et al., 2007; Frederick et al., 2007; Stockmann et al., 2008; Dhulesia, Gsponer, and Vendruscolo, 2008; Nilapwar et al., 2009; Mauldin, Carroll, and Lee, 2009; Hyeon et al., 2009). These changes have impact on the conformational entropy of the protein and thereby can have significant impact on affinity, too (Diehl et al., 2010; Marlow et al., 2010; Fenley, Muddana, and Gilson, 2012).

1.1.3 The Selectivity Challenge in CADD

In many of the examples used to illustrate strategies to obtain selectivity, the reason for selective binding became clear retrospectively, after detailed experimental results had been obtained. Testing for specific homologue off-target effects and common non-specific effects, e.g. via human ether-a-go-go channel (hERG, potassium ion channel, Cavalli et al., 2002; Farid et al., 2006) or cytochrome P450 interactions (CYP, oxidoreductases, De Graaf, Vermeulen, and Feenstra, 2005; Obach et al., 2006), is broadly performed and predictive models were developed. However, a generalized prospective computational approach remains challenging. This is due to the intrinsic design of computational methods which are limited in their ability to accurately predict the affinity of protein-ligand complexes, also known as scoring.

This accounts for complexes with the target, but is even more decisive for comparison of different complexes like target and off-target(s). CADD methods like docking programs are optimized towards a low rate of false positive hits at cost of higher false negatives (Kitchen et al., 2004; Cole et al., 2005; Leach, Shoichet, and Peishoff, 2006). For accurate selectivity prediction however, the false negative rate needs to be reduced without an increase of false positives. This challenge significantly increases when protein flexibility and explicit binding site water need to be addressed (Zentgraf et al., 2007). As current docking programs and scoring functions only incorporate first shell interactions, they are doomed to fail if the binding sites of target and off-target(s) are conserved.

For the elucidation of the selectivity determining features in proteins with conserved binding sites, a computationally and experimentally accessible model system is needed. Ideally, such a system should fulfill the following requirements: at least two proteins with a conserved binding site and known structure, as well as selective and non-selective ligands with known binding modes. A possible model system is the *N*-myristoyltransferase (NMT). High resolution structures of the target from *Leishmania major* (*Lm*NMT) and the human off-target homologue (*Hs*NMT1) in complex with and without ligands are freely available in the protein data bank (PDB, Berman et al., 2000) and a plethora of inhibitors are known (see below). The elucidation of protein flexibility and binding site water molecules and their influence on selectivity for this model system, aim to improve the understanding of these features and rationalise selective ligand design in general. Particularly in those cases where the high homology between target and off-target had prevented drug design approaches so far, this improved understanding may point out new opportunities.

1.2 The *N*-Myristoyltransferase

1.2.1 Enzyme Mechanism

The *N*-myristoyltransferase (NMT) catalyzes the transfer of C:14-saturated fatty acid myristate to the N-terminal glycine residue of its substrate proteins in a co- or post-translational reaction in eucaryotes (Wilcox, Hu, and Olson, 1987; Farazi, Waksman, and Gordon, 2001; Bowyer et al., 2008; Martin, Beauchamp, and Berthiaume, 2011). The reaction follows an ordered Bi-Bi mechanism (Rudnick et al., 1991; Bhatnagar et al., 1999). First, myristoyl-coenzyme A (MyrCoA, Figure 1.4) binds to NMT which induces a structural rearrangement to open a second binding pocket for the peptide substrate (Figure 1.5). The N-terminal glycine is deprotonated by the C-terminus of NMT (like Leu421 in *Lm*NMT) and rotates around its Ψ angle. The myristate

is then transferred to the N-terminal glycine residue of the substrate in an nucleophilic addition-elimination reaction (Figure 1.6). The carbonyl group of the MyrCoA thioester is activated by polarisation through an oxy-anion hole formed by the NMT amide backbone (Phe168, Leu169). The released thiolate is assumed to be stabilized by the aromatic amine of adenine prior protonation from glycine. Subsequently, the free CoA and the *N*-myristoylated substrate protein are released.

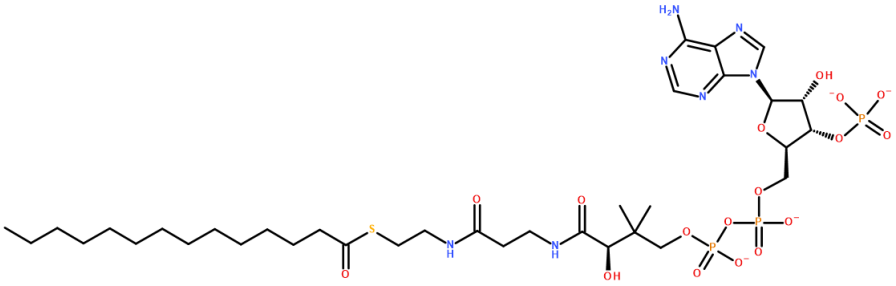


FIGURE 1.4: Myristoyl-Coenzyme A (MyrCoA)

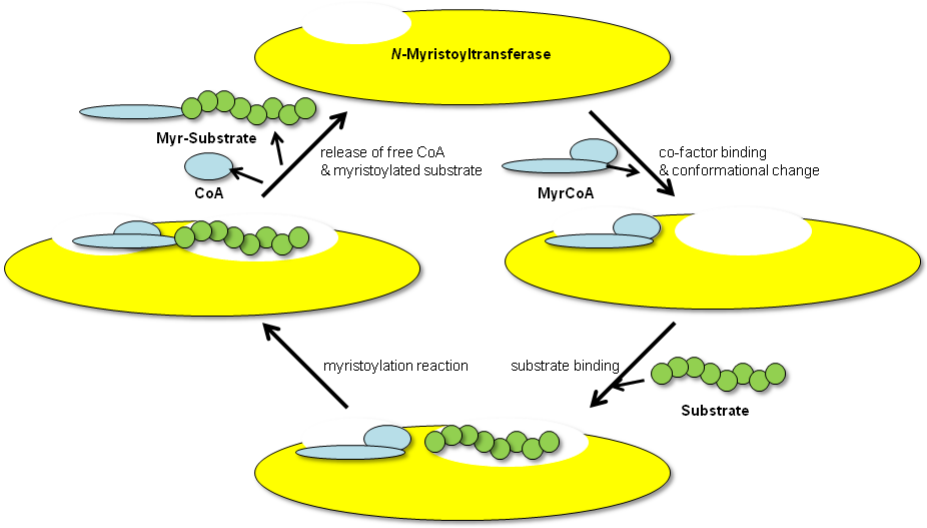


FIGURE 1.5: Catalytic myristoylation cycle.

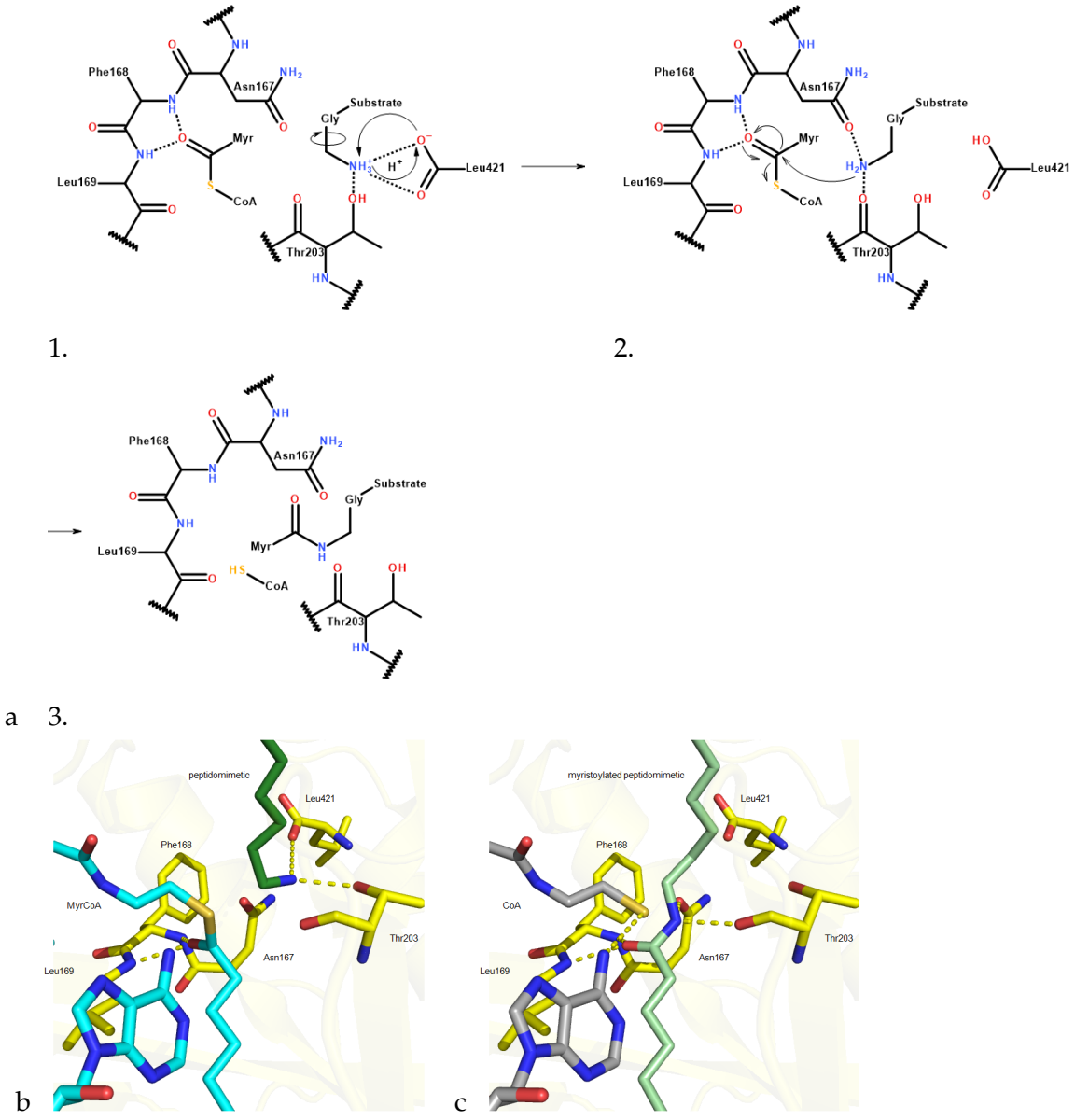


FIGURE 1.6: a) Myristoylation mechanism: 1. deprotonation of substrate N-terminal glycine. 2. nucleophilic addition-elimination reaction with the oxy-anion hole activated thioester of MyrCoA. 3. release of free CoA and myristoylated peptide. b) and c) myristoylation of peptidomimetic ligand, PDB-ID 4C7H (Olaleye et al., 2014). Yellow carbon atoms: *Lm*NMT, dark green: peptidomimetic ligand, cyan: MyrCoA; pale green: myristoylated peptidomimetc, grey: free CoA.

The upstream reactions prior myristoylation are the detachment of methionine by methionine aminopeptidase (MetAP) for co-translational myristoylation or a post-translational protease reaction to excavate "hidden" myristoylation recognition patterns. The sequence of NMT substrates is well described (Figure 1.7). The general consensus motif was identified to be Gly-X-X-X-Ser-X-X-X with some species specific preferences for "X" (Utsumi et al., 2004; Wright et al., 2010; Goldston et al., 2014). Global profiling of myristoylated proteome (myristome) shed light on the plethora of substrates

(>100 in human) and function of myristoylation (Thinon et al., 2014; Traverso, Giglione, and Meinnel, 2013; Xu et al., 2015; Wright et al., 2014). Further, computational tools for the prediction of potential substrates were developed (*MYR Predictor*, Maurer-Stroh, Eisenhaber, and Eisenhaber, 2002 and *Myristoylator*, Bologna et al., 2004).

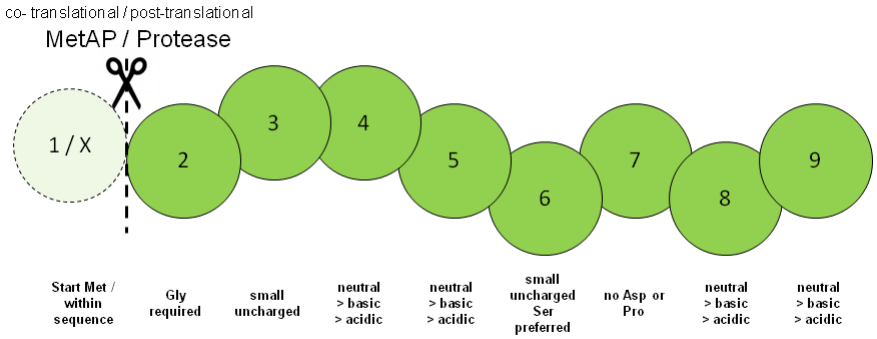


FIGURE 1.7: Myristoylation recognition sequence.

1.2.2 NMT as a Drug Target

The role of protein *N*-myristoylation includes cellular localization, signal transduction, tumorgenesis and apoptosis, virus assembly and morphological changes among others (Xu et al., 2015). Further, NMT was proven to be essential for parasitic survival and virulence (Price et al., 2003; Price et al., 2010). Thus, NMT is a drug target for cancer, viral diseases and neglected parasitic diseases (Ducker et al., 2005; Martin, Beauchamp, and Berthiaume, 2011; Thinon et al., 2016; Frearson et al., 2007; Bowyer et al., 2008; Brannigan et al., 2010; Goldston et al., 2014; Tate et al., 2013; Wright et al., 2014; Wright et al., 2015; Herrera et al., 2016; Ritzefeld, Wright, and Tate, 2017).

Drug Design projects targeting NMT include fungal infections of *Candida albicans* and *Aspergillus fumigatus* (also model system *Saccharomyces cerevisiae*; Devadas and Zupec, 1995; Bhatnagar et al., 1999; Fang et al., 2015), African sleeping sickness (or human African trypanosomiasis - HAT) and Chagas disease caused by *Trypanosoma brucei* and *T. cruzi*, respectively (Frearson et al., 2010; Brand et al., 2012; Brand et al., 2014; Roberts et al., 2014; Robinson and Wyatt, 2015; Spinks et al., 2015; Herrera et al., 2016), malaria (*Plasmodium falciparum* and *P. vivax*; Yu et al., 2012; Goncalves et al., 2012b; Rackham et al., 2013; Rackham et al., 2014; Yu et al., 2015) and leishmaniasis caused by *Leishmania donovani* and *L. major* (Bell et al., 2012; Hutton et al., 2014; Brannigan et al., 2014; Rackham et al., 2015). Further, drug-repurposing approaches against NMT from the filarial nematodes *Caenorhabditis elegans* and *Brugia malayi* were successful (Galvin et al., 2014).

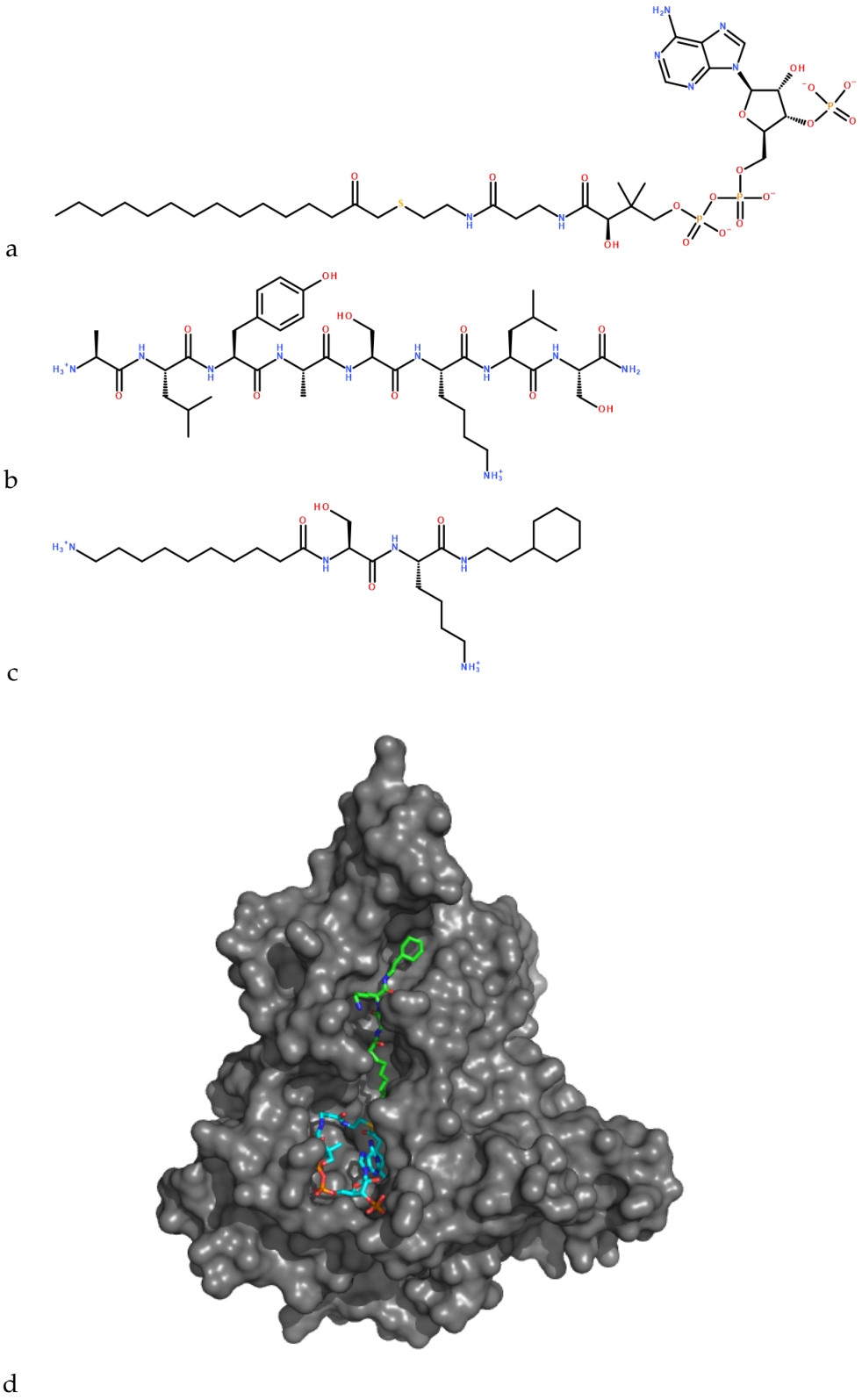


FIGURE 1.8: Initial NMT inhibitors: a) nonhydrolysable MyrCoA derivative. b) and c) peptidomimetics. d) peptidomimetic inhibitor c (green carbon atoms) and MyrCoA (cyan carbon atoms) bound to *Lm*NMT (grey surface, PDB-ID 4C7H)

With the exception of 2-oxopentadecyl-CoA (Figure 1.8 a; Pasha et al., 2004), a non-hydrolysable derivate of MyrCoA, the majority of designed inhibitors target the substrate binding site. First inhibitors developed against fungal NMTs were designed as peptidomimetics (Figure 1.8 b-d; Devadas and Zupec, 1995; Bhatnagar et al., 1999; Olaleye et al., 2014), but also many small drug-like molecules were found in high-throughput screenings and rational design projects (Zhao and Ma, 2014). Recent fragment-based approaches resulted in promising scaffolds as well (Robinson and Wyatt, 2015; Spinks et al., 2015).

Even though first inhibitors of *Trypanosoma brucei* NMT (*Tb*NMT) were able to cure HAT in mice (Frearson et al., 2010) the development of selective inhibitors is still a challenge (Goncalves et al., 2017). In cases where selective inhibition was achieved the reasons on a molecular level remained unclear. Thus, further studies are needed to elucidate the selectivity determining features.

1.2.3 Structure and Conservation

In this study, we focused on *Lm*NMT and *Hs*NMT1 to study the driving forces for selectivity in highly conserved binding sites for the following reasons: First of all, *Lm*NMT is not only a valid drug target, it is also broadly used as surrogate to obtain crystal structures, when crystallization conditions for the NMT under investigation are not or only hardly available (e.g. *Trypanosoma brucei* or *Leishmania donovani* - *Tb*NMT, *Ld*NMT; Frearson et al., 2010; Brand et al., 2012; Brannigan et al., 2014; Rackham et al., 2015). Several crystal structures of *Hs*NMT1 are also available, as it is regularly used as the off-target model.

The two enzymes share an overall sequence identity of 44 % and similarity of 65 % as well as a comparable shape (C_α RMSD = 0.613 Å, Figure 1.9 a). This conservation especially accounts for the binding site (defined by 5 Å around reference ligand compound **1**, Figure 1.9 b) with solely three different amino acids. Further, these three residues' side chains are oriented away from the binding site and form no interactions with the ligand. These residues are His398, Met420 and Leu421 for *Lm*NMT and Asn473, Leu495 and Gln496 at the corresponding positions in *Hs*NMT1 (Figure 1.9 c).

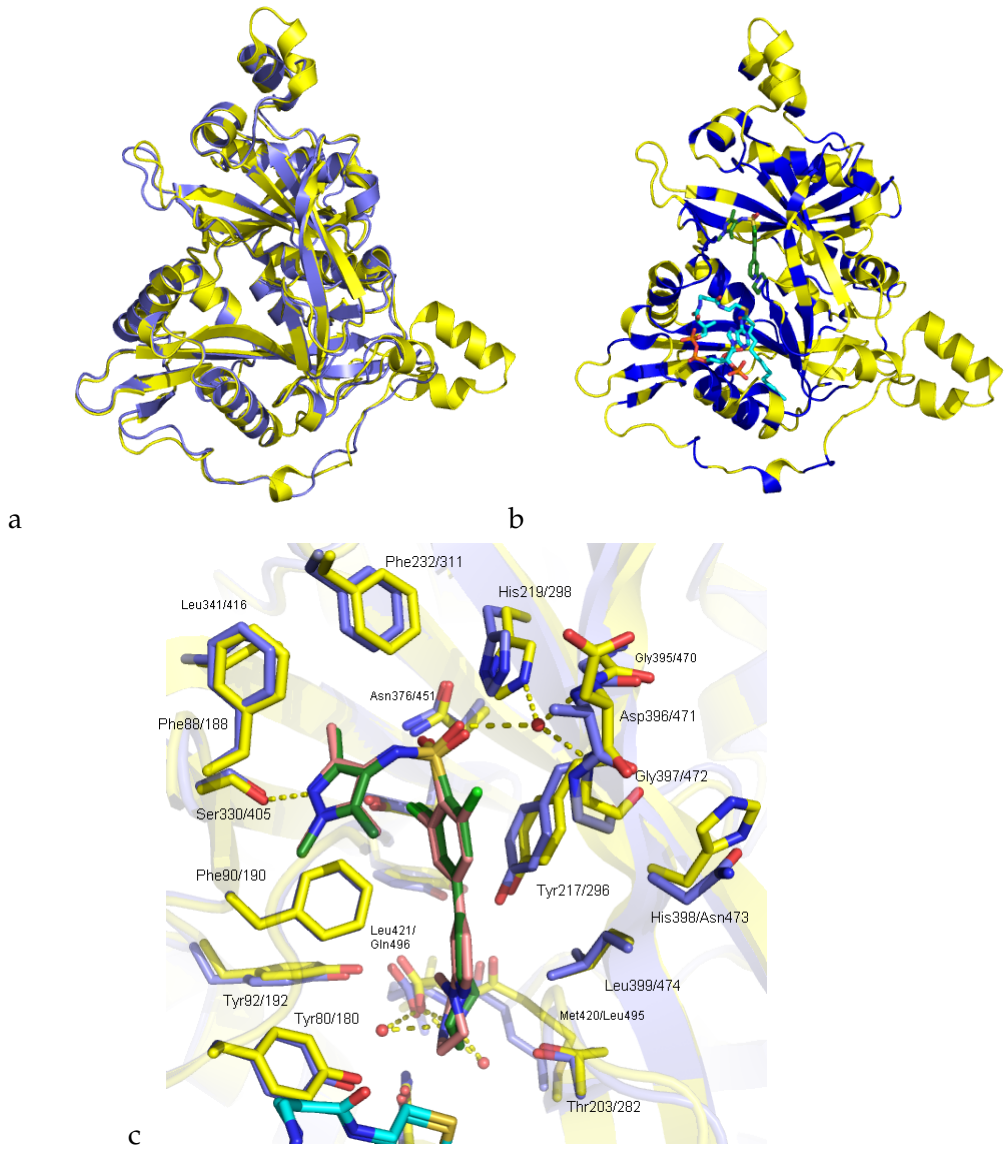
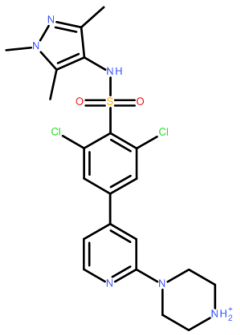


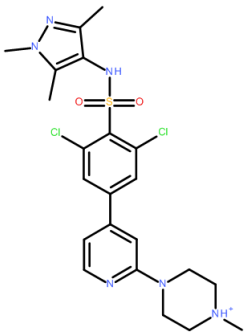
FIGURE 1.9: Conservation between *Lm*NMT and *Hs*NMT1: a) Structural alignment of *Lm*NMT in yellow and *Hs*NMT1 in blue. b) Conserved residues marked in dark blue within *Lm*NMT, compound **1** with dark green and MyrCoA with cyan carbon atoms to highlight the binding site. c) Enlarged binding site of *Lm*NMT (yellow) and *Hs*NMT1 (blue) around compound **1** - dark green in *Lm*NMT, light red in *Hs*NMT1, PDB-IDs 2WSA and 3IWE.

1.2.4 Selected Inhibitors

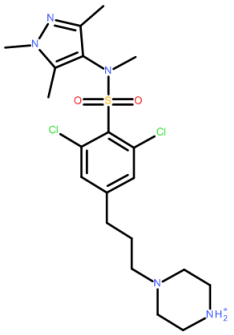
Several *Lm*NMT inhibitors are know. Those chosen for this investigation were selected based on their high affinity and their selectivity profile (Figure 1.10). Compounds **1-4** originate from a series of *Tb*NMT inhibitors from the *Drug Discovery Unit Dundee* (Frearson et al., 2010; Brand et al., 2012; Brand et al., 2014) derived from a small sulfonamide (Figure 1.11 a). Crystal structures of some of these and closely related inhibitors in *Lm*NMT and



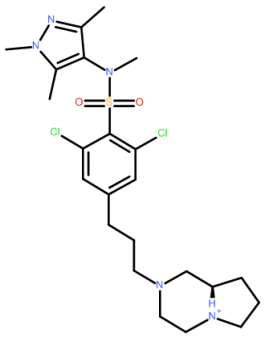
compound 1
646 (PDB-ID 2WSA)
*Lm*NMT IC_{50} 6 nM
*Tb*NMT IC_{50} 2 nM
*Hs*NMT1 IC_{50} 12 nM



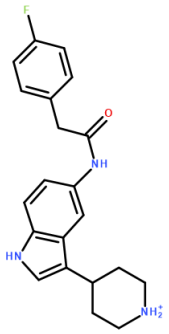
compound 2
*Tb*NMT IC_{50} 15 nM
*Hs*NMT1 IC_{50} 25 nM



compound 3
*Tb*NMT IC_{50} 3 nM
*Hs*NMT1 IC_{50} 51 nM



compound 4[a,b]



compound 5[a]
7AH (PDB-ID 4CGN)
*Lm*NMT IC_{50} 550 nM
*Hs*NMT1 IC_{50} 59.2 μM

FIGURE 1.10: Compounds included in this study. PDB three letter-code if available. [a] selective for *Lm*NMT over *Hs*NMT1. [b] Inhibition data obtained in this work.

*Hs*NMT1 were determined. They all share the same pyrazole sulfonamide scaffold and key interaction features (Figure 1.9 c). A hydrogen bond to Ser330 (Ser405 respectively in *Hs*NMT1) from the pyrazole moiety which also forms π-π stacking interactions with Phe90/190 (*Lm*NMT/*Hs*NMT1 residue numeration). The trimethylation of this moiety improves packing in that pocket especially in 3-position with Phe88/188 and Leu341/416. The sulfonamide moiety interacts water-mediated with His219/298 and the backbone Asp396/471 and Gly397/472 (Figure 1.12). This water mimics the Ser residue at the 6-position of the substrate. The Asp396/471 and Gly397/472 interactions are induced by a flip around the Asp Ψ angle upon ligand and substrate binding. Therefore, this region can be called a *"closing lid"*. A hydrophobic aromatic linker, stacking with Tyr217/296 connects the pyrazole sulfonamide moiety to a basic center which mimics the N-terminus of the substrate. This results in a large gain of affinity by

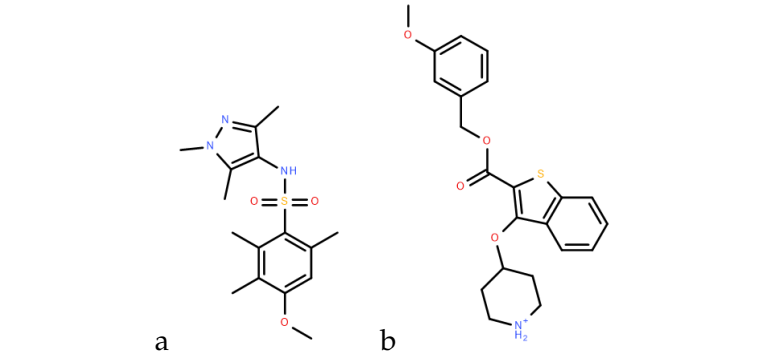


FIGURE 1.11: a) Initial hit for development of sulfonamide inhibitors. b) Selective *Pv*NMT inhibitor as starting point of oxadiazole *Ld*NMT inhibitor development

either direct or indirect interaction with the catalytically active C-terminal Leu421/Gln496 (Figure 1.12 b; Brand et al., 2012). From those sulfonamide inhibitors compound **1** is the most prominent as it was able to cure HAT in mice (Frearson et al., 2010) and is regularly used as reference compound and for analysis of binding modes. Compounds **2-4** show similar structure, but only compound **4** is selective for *Lm*NMT over *Hs*NMT1 (unpublished data).

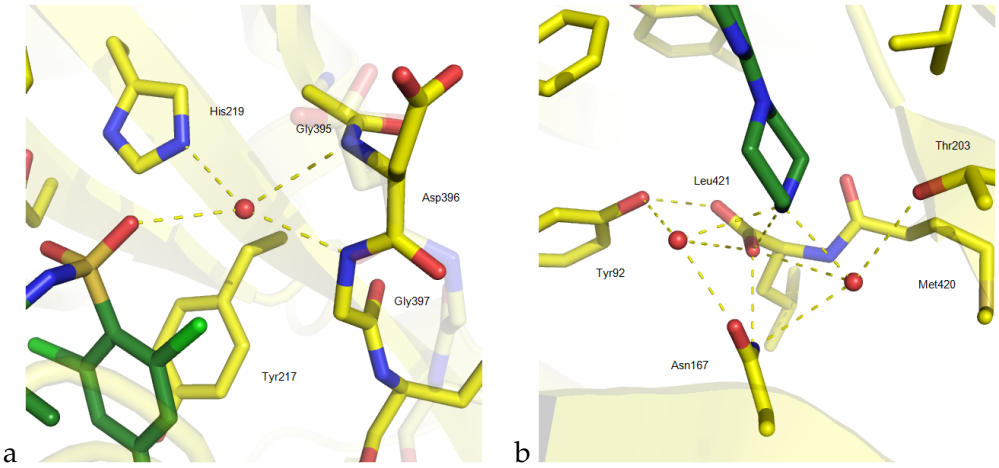


FIGURE 1.12: Water-mediated interactions between compound **1** (dark green carbon atoms) and *Lm*NMT. a) Sulfonamide moiety with closing lid. Transparend sticks show ligand-free conformation (PDB-ID 3H5Z). b) Interactions of secondary amine with C-terminal Leu421. Ligand bound PDB-ID 2WSA.

The selective inhibitor compound **5** was identified in a high-throughput screening (HTS) for *Ld*NMT inhibitors at *Pfizer* (Bell et al., 2012). *Ld*NMT and *Lm*NMT share a sequence identity of 97.8 % and it was shown that affinity of compound **5** for both enzymes was comparable (Brannigan et al., 2014). This compound adapts a different binding mode than the sulfonamide inhibitors. The key differences are the lack of a hydrogen bond to Ser330 and that it is binding to the "open" conformation of NMT, defined by the orientation of Tyr217 and the closing lid, which opens a hydrophobic pocket (Figure 1.13 a). Binding to the closed conformation would result in a clash

with Tyr217/296 (Figure 1.13 b). Therefore, the initial selectivity hypothesis for compound **5** was that the open conformation was preferred in *Lm*NMT and *Ld*NMT, whereas in *Hs*NMT1 the closed conformation was preferred. However, ligand-free NMT crystals showed flexibility for Tyr217/296 in both enzymes (Figure 1.13 c).

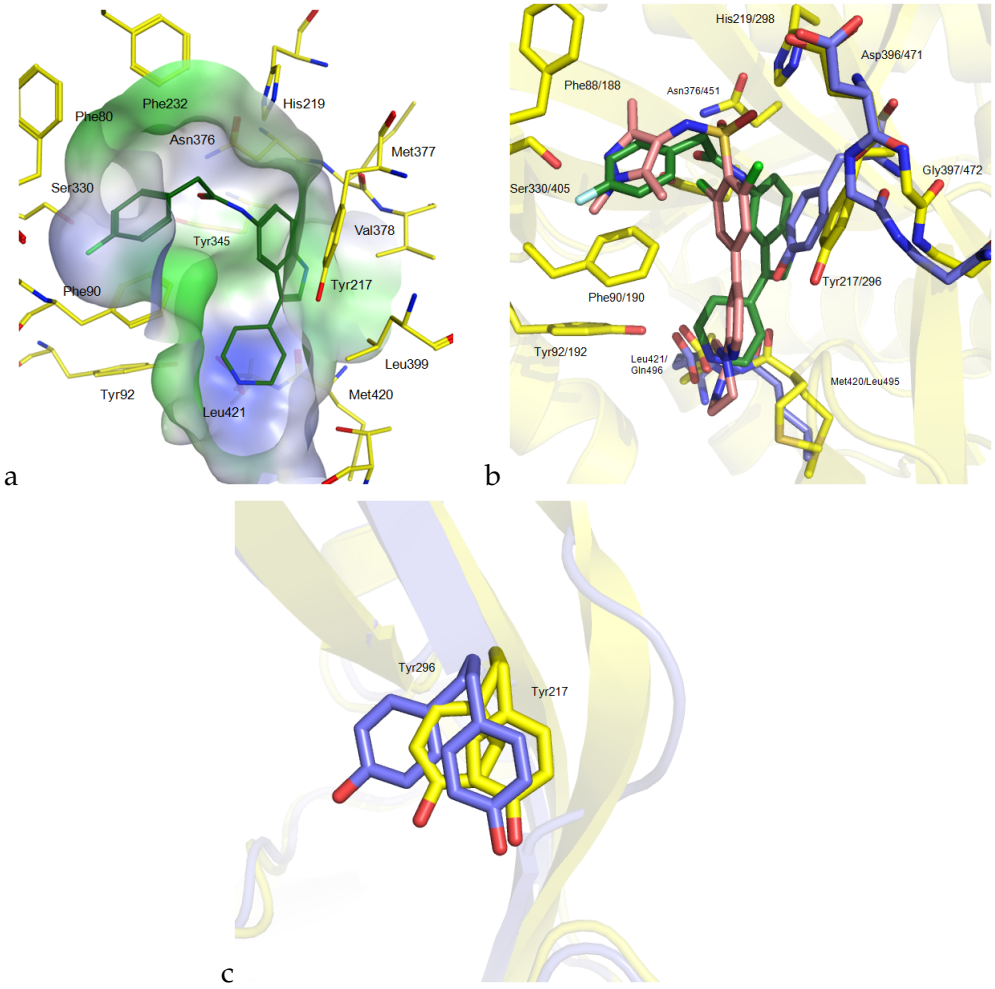


FIGURE 1.13: a) Binding mode of compound **5** (dark green carbon atoms) in complex with *Lm*NMT (PDB-ID 4CGN, yellow carbon atoms). Surface colored by lipophilicity to illustrate addressed hydrophobic pocket - green lipophilic, blue hydrophilic. Figure made with MOE 2015. b) Binding mode of compound **5** (green carbon atoms) in complex with *Lm*NMT (PDB-ID 4CGN, yellow carbon atoms) in overlay with compound **1** (pale red carbon atoms) bound to *Hs*NMT1 (PDB-ID 3IWE, blue carbon atoms).Compound **5** binds to the "open" conformation defined by the orientation of Tyr217/296 and the closing lid (Asp396/471 and Gly 397/472). c) Multiple conformations for Tyr217/296 in unbound *Lm*NMT (PDB-ID 3H5Z) and *Hs*NMT1 (PDB-ID 3IU1).

Additionally, recently found selective and non-selective inhibitors binding to the NMT-open conformation indicated that selectivity is not determined by the open conformation alone. These inhibitors included a hybrid molecule of compound **5** and another selective compound **9** from the same

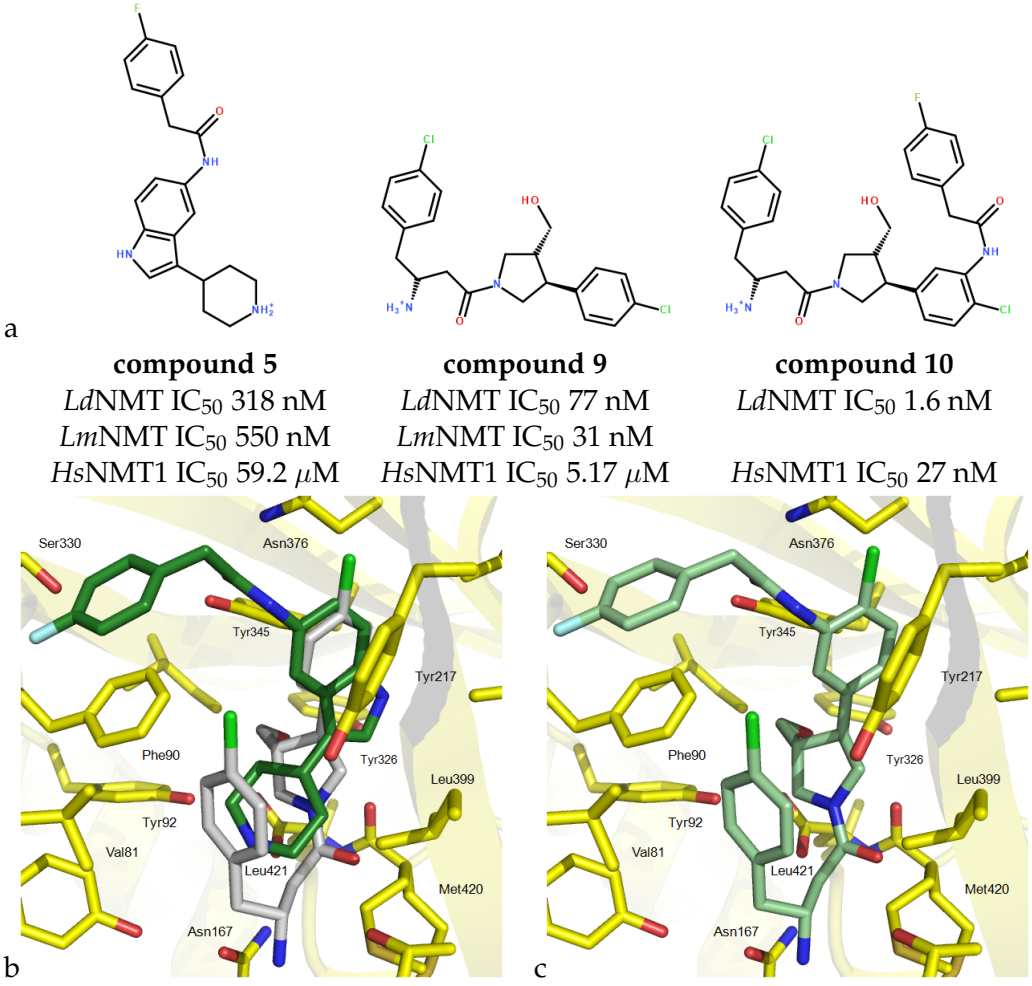
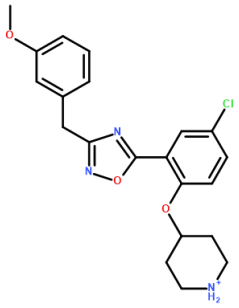


FIGURE 1.14: a) Rational hybridization design of compounds **5** and **9** to yield compound **10**. b) Overlay of binding modes from compound **5** (dark green carbon atoms, PDB-ID 4CGN) and compound **9** (light grey carbon atoms, PDB-ID 4CGL) bound to *Lm*NMT. c) Binding mode of hybrid compound **10** (pale green carbon atoms, PDB-ID 4CYO).

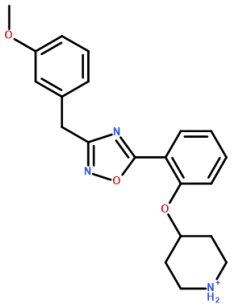
Pfizer HTS, which resulted in high affinity inhibitor compound **10**, but with lower selectivity (Figure 1.14; Hutton et al., 2014).

Furthermore, a series of oxadiazole inhibitors optimized from a benzothiophene inhibitor (Figure 1.11 b), which showed selectivity for *Pv*NMT over both *Ld*NMT and *Hs*NMT1, resulted in high affinity *Ld*NMT inhibitors with different selectivity over *Hs*NMT1 (Rackham et al., 2015). The selectivity of compound **7** (Figure 1.15 a) seemed to be caused by the removal of a single chlorine atom, but the reasons on the molecular level remained unclear.

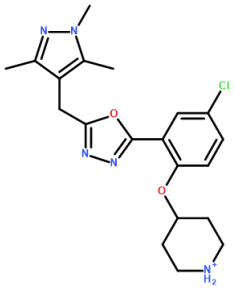
a



compound 6
*Ld*NMT IC_{50} 20 nM
*Hs*NMT1 IC_{50} 70 nM



compound 7[a]
*Ld*NMT IC_{50} 60 nM
*Hs*NMT1 IC_{50} 1.24 μM



compound 8
*Ld*NMT IC_{50} 10 nM
*Hs*NMT1 IC_{50} 20 nM

b

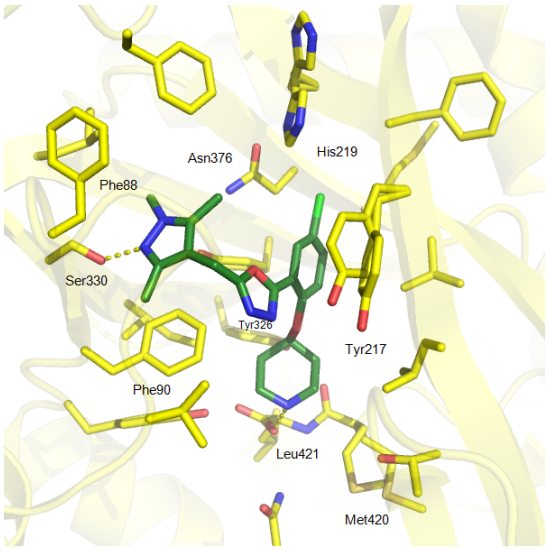


FIGURE 1.15: a) Non-selective and selective[a] open-conformation binding *Ld*NMT inhibitors. b) Binding mode of compound **8** in *Lm*NMT (PDB-ID 5A28).

1.3 Hypothesis

Due to the conserved binding sites between *Lm*NMT and *Hs*NMT1 direct protein-ligand interactions can not account for the selectivity of identified ligands. However the examples in Chapter 1.1 showed that selectivity can be described by other events than first shell interactions. This includes changes in binding site hydration (Chapter 1.1.1) and indirect interactions like those illustrated in Figure 1.12, as well as changes in protein dynamics (Chapter 1.1.2). The crosstalk between ligand and interacting protein residues to second and further shells of the protein can result in a change of dynamics making distal regions of the protein more rigid or flexible (Frederick et al., 2007; MacRaild et al., 2007; Stockmann et al., 2008; Dhulesia, Gsponer, and Vendruscolo, 2008; Nilapwar et al., 2009; Mauldin, Carroll, and Lee, 2009; Hyeon et al., 2009). These changes are associated with conformational entropy which can have a significant impact on protein-ligand affinity (Diehl

et al., 2010; Marlow et al., 2010; Fenley, Muddana, and Gilson, 2012).

Based on these considerations, we hypothesized that the reason for selective inhibition of NMT from different species is caused by changes in protein flexibility upon ligand binding. If these changes differ between *Lm*NMT and *Hs*NMT1, the correlated conformational entropy of the protein and subsequently the affinity to the ligand is different as well. Further, the indirect protein-ligand interactions (Figure 1.12) indicated the importance of binding-site water molecules in NMT. Previous work (Chapter 1.1.1) already showed the large impact of only minor differences in binding-site hydration on affinity and selectivity in kinases (Robinson, Sherman, and Farid, 2010; Barillari et al., 2011; Robinson et al., 2016). Therefore similar features might also determine selectivity in NMT.

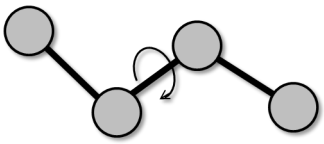
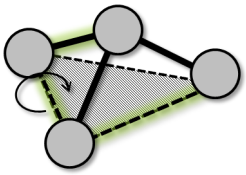
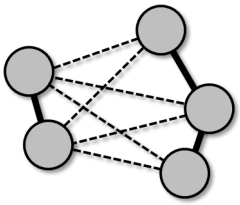
1.4 Molecular Dynamic Simulations and Force Fields

Molecular dynamic simulations (MDs) are a very useful tool to study protein flexibility (MacKerell et al., 1998; Leach, 2001; Karplus and McCammon, 2002; Karplus and Kuriyan, 2005; Klebe, 2009; Sotriffer, 2011). MDs are based on force fields and often referred to as molecular mechanics (MM). Using these empirical techniques, molecular properties like geometry and energy are calculated. The forces within a single molecule or between multiple molecules are described by functions with parameters for covalent and non-covalent interactions, including bond lengths, angles, dihedral angles and non-bonded interactions. These parameters are derived from experimental data and quantum mechanic (QM) calculations for high accuracy. The overall energy of a molecule or a multi-molecule complex is the sum of its single terms (Table 1.1).

The starting structure for force field calculations is of importance as well. As an energy minimization process is generally "down-hill" (e.g. towards lower energy states), different starting structures, e.g. of a protein-ligand complex, will result in different local minima of a multi-dimensional energy landscape. Therefore, different start-geometries for force field calculations should be used. MDs are a special case of force field calculations. In MDs not only potential energy, but also kinetic energy is considered. This is of relevance, because biological processes do not take place at 0 Kelvin (K), but at around 310 K resulting in dynamic structures. Especially for protein-ligand interactions, the flexibility, conformational changes and adaption processes play an important role. In MDs the movement of atoms and molecules is simulated under the given parameters of a force field. A standard work-flow begins with a starting structure, like a prepared protein or DNA structure from the PDB or a homology model, put into implicit or explicit solvent

TABLE 1.1: Description of energetic force field terms for bonded and non-bonded interactions. Bonds and angles are described by quadratic equations. For the bond potential function E_{bond} in $\frac{kcal}{mol}$, the bond force constant is K_b in $\frac{kcal}{mol*Å^2}$, the bond length b and minimal energy bond length b_0 are in Å. The angle potential function E_{angle} in $\frac{kcal}{mol}$ is described by the angle force constant K_θ in $\frac{kcal}{mol*deg^2}$ and the angle θ and minimal energy angle θ_0 in deg. For the dihedral potential function E_{tors} in $\frac{kcal}{mol}$, the dihedral force constant is K_ϕ in $\frac{kcal}{mol}$, the torsion angle ϕ and the phase shift δ are in degree. n is the multiplicity of the wave function. To describe torsion angles sometimes multiple interfering wave functions are used to discriminate different energetic minima and maxima. Improper torsion angles are a special case of torsion angles. They do not follow the order in which four atoms are connected, but they describe the orientation of a certain atom relative to the plain which is defined by the remaining three atoms that are connected to the first atom. This orientation of atom 1 to the "atom-2-3-4-plain" e.g. is used to describe the planarity of aromatic systems or to freeze stereo centers. The non-bonded term E_{nb} consists of electrostatic and - without the electrostatic term - van der Waals (vdW) energies. $A_{ij}r_{ij}^{-12}$ describes the repulsion of two atoms being very close to each other. A_{ij} is proportional to the sum of the radii of the atoms i and j. $-C_{ij}r_{ij}^{-6}$ describes the attraction between the two atoms. $q_iq_j/\epsilon r_{ij}$ finally describes the electrostatic interactions (coulombic potential) between i and j with the (partial-) charge of q and ϵ being the permittivity. r_{ij} is the distance between the two atoms i and j for all partial terms. Without the last electrostatic term, this non-bonded (vdW-) interaction potential is called the Lennard-Jones (LJ) potential. Table and illustrations modified from Klebe, 2009.

$$E = \sum_{i}^{bonds} E_{bond} + \sum_{i}^{angles} E_{angle} + \sum_{i}^{torsionangles} E_{tors} + \sum_{i}^{non-bonded} E_{nb} \quad (1.1)$$

Term	Illustration	Equation
Bond		$E_{bond} = K_b * (b - b_0)^2$
Angle		$E_{angle} = K_\theta * (\theta - \theta_0)^2$
Torsion		$E_{tors.} = K_\phi * (1 + (cos(n * \phi - \delta)))$
Improper torsion		same as torsion, but no consecutive atoms
Non-bonded		$E_{nb} = A_{ij}r_{ij}^{-12} - C_{ij}r_{ij}^{-6} + q_iq_j/\epsilon r_{ij}$

(e.g. a water box or sphere). The structure preparation includes the protonation of the structure, selection of residue conformations, if multiple are possible (see occupancy column of PDB files), system charge neutralisation by adding ions (Na^+ or Cl^-) and an energy minimization to avoid bias from geometrically unfavourable starting structures. The simulation is started by the application of random starting velocities to every atom of the system equivalent to a given temperature (Boltzmann distribution). Subsequently, the forces on every atom are calculated and the atoms are moved to their next position by solving Newton mechanic equations. This procedure is stepwise continued (one step is usually 1-2 fs) until a defined end point is reached (Figure 1.16). The resulting trajectory is then analyzed, e.g. to identify side-chain orientations, conformational changes or – on a larger time scale – the formation of secondary protein structures or a ligand binding to a protein. Further binding affinities in form of free enthalpy ΔG can be calculated by some advanced MD methods, as well .

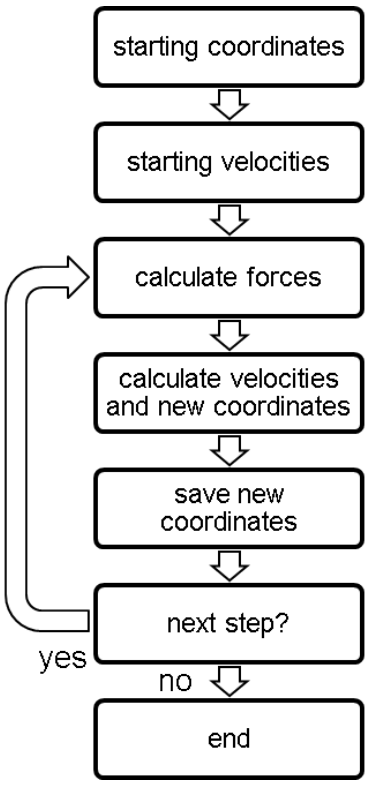


FIGURE 1.16: Exemplary standard work-flow of a MD.

1.4.1 Order Parameters - S^2

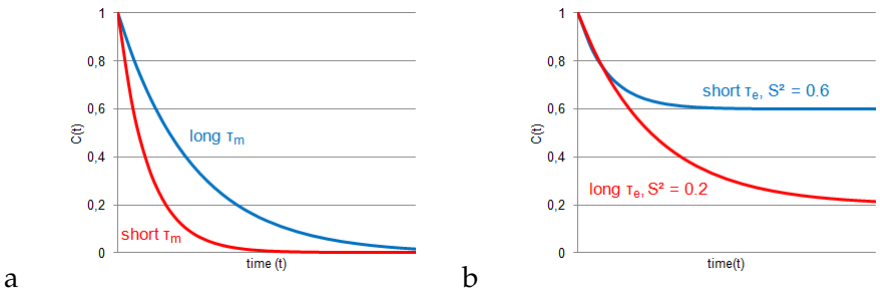
One metric to analyze the internal dynamics of a protein, are order parameters (S^2). Based on the so-called *model-free formalism* (Lipari and Szabo, 1982;

Clarkson, 2007) the tumbling of a A-X bond can be approximately described by a correlation function of the form

$$C_M(t) = e^{\frac{-t}{\tau m}} \tag{1.2}$$

with t as the time and τ_m as a time constant for how fast A-X tumbles called *rotational correlation time*. This equation describes the likeliness to find A-X at a certain time point i+t in the same orientation as at time point i. For an independent and free tumbling bond $C_M(t)$ will decay to 0 depending on how long τ_m is (Figure 1.17 a). For large systems (like proteins) the movement vector of A-X with X being part of the larger system (the protein) and A, an atom bound to this system (e. g. a side chain atom), and under the assumption that the protein is not tumbling in solution, C(t) does not decay to 0 (Figure 1.17 b). Instead, a static value between 0 (high flexibility) and 1 (very rigid) will be reached. This is called a *generalized order parameter* - S^2. This "extra" motion $C_E(t)$ is described by

$$C_E(t) = S^2 + (1 - S^2)e^{\frac{-t}{\tau e}} \tag{1.3}$$



a b

FIGURE 1.17: Form of correlation function a) for freely moving A-X bond and b) for larger system with S^2.

As the protein part of the bond (X) is not completely rigid the total correlation function $C_T(t)$ can be described for independent motions (hence it is called model-free) as the product of equation 1.2 and 1.3 for internal and global motion:

$$C_T(t) = C_M(t) * C_E(t) \tag{1.4}$$

The correlation function itself can not be measured, but NMR experiments give the *spectral density function* $J(\omega)$ from which order parameters can be obtained (Lipari and Szabo, 1982). Further, experimental side chain order parameters were shown to be proportional to the reciprocal value of conformational entropy (Yang and Kay, 1996; Li and Brüschweiler, 2009), from which an amino acid specific dictionary of side chain entropies was established. The general form is:

$$S = k_B M[A + Bf(1 - S^2_{NMR})] \tag{1.5}$$

With S as the conformational entropy, k_B the Boltzman constant (1. 380 648 527 9 * 10^{-23} J/K), M the number of side chain dihedral angles, S^2_{NMR} the NMR-derived side chain order parameters and A and B residue bond-vector specific fit parameters listed in the dictionary.

From a computational chemistry/biology point of view it was shown that MDs are fairly well able to reproduce experimental S^2 especially for backbone (N-H) bond vectors (Showalter et al., 2007). They can be calculated using

$$S^2 = \frac{3}{2}[(x^2)^2 + (y^2)^2 + (z^2)^2 + 2(xy)^2 + 2(xz)^2 + 2(yz)^2] - \frac{1}{2} \tag{1.6}$$

with x, y and z as the components of the unit vector along the investigated bond (e. g. N-H of the amide backbone or S_δ-C_ϵ for a methionine side chain). Squared parentheses imply the average over the whole trajectory (MacRaild et al., 2007).

However, due to the two assumptions: $S^2_{MD} \approx S^2_{NMR}$ and $S^2_{NMR} \propto f(1/S)$, conclusions from MD-derived S^2 to conformational entropy can be only drawn in a qualitative manner (Stockmann et al., 2008), even though improvements of the methods are under continuous investigation (Genheden, Akke, and Ryde, 2014; Allner, Foloppe, and Nilsson, 2015).

1.4.2 Quantitative Analysis of Binding Site Water

Water network analysis can also be performed using explicit solvent MDs (Betz et al., 2016). Density peaks of water molecules may highlight conserved water molecules present only in one of two related enzymes contributing to selectivity by their displacement or indirect interactions.

One method for the quantitative analysis of single binding site water molecules is called SPAM (maps spelled in reverse, Cui, Swails, and Manas, 2013). In SPAM the distribution of interaction energies (E_{water}) between a water molecule within a hydration site and its environment is analysed. The hydration site is defined as a water density peak, calculated by a grid based method from explicit solvent MDs, with a certain cut-off (usually 0.07). The free energy G_{SPAM} of the water molecule representing this density peak is subsequently described by

$$G_{SPAM} = -RT * lnQ_{SPAM} \tag{1.7}$$

Therefore the interaction energy (E_{water}) for a hydration site with a water molecule present, is calculated for every frame from a MD. For every hydration site, the average interaction energy ($\bar{E}_{water}$) with its standard deviation and subsequently the corresponding probability density function of a Gaussian distribution is derived. This function is used for the calculation of Q_{SPAM}:

$$Q_{SPAM} = \sum_{E_{water}} [P(E_{water}) * e^{\frac{-E_{water}}{RT}}] \tag{1.8}$$

with $P(E_{water})$ as the probability of the observed water molecule having the interaction energy E_{water} for a given frame.

The energy difference between bound and unbound water ($G_{SPAM,bound}$ - $G_{SPAM,bulk}$ derived from Equation 1.7) is calculated using

$$\Delta G_{SPAM} = G_{SPAM,bound} - G_{SPAM,bulk} \tag{1.9}$$

and

$$\Delta H_{SPAM} = \bar{E}_{water,bound} - \bar{E}_{water,bulk} \tag{1.10}$$

with $\bar{E}_{water,bound/bulk}$ being the average interaction energy during the MDs and

$$-T\Delta S_{SPAM} = \Delta G_{SPAM} - \Delta H_{SPAM} \tag{1.11}$$

to describe the temperature dependent entropy. By this method, the "affinity" as ΔG, as well as the thermodynamic properties (ΔH and $-T\Delta S$) of a single hydration site can be estimated.

1.5 Objectives

The lack of the molecular understanding of the underlying driving forces behind selective inhibition, when facing a conserved binding site, is a major hurdle for rational drug design. This hurdle may have prevented some potential targets to be further evaluated. Therefore, the overall objective of this project is the elucidation of these selectivity determining features in proteins with conserved binding sites using *Lm*NMT and *Hs*NMT1 as a model system.

To reach this goal, a combination of computational and experimental methods, namely the thermodynamic characterization by ITC experiments,

the structural analysis from crystallography and the dynamic evaluation from explicit solvent MDs was used. Such a complementary and synergistic approach can improve the understanding of the influence of protein flexibility and water thermodynamics on selectivity. The proof of concept was performed by site-directed mutagenesis to yield mutant NMT proteins with tailored selectivity profile. In this way, validated results with the aim to improve incorporation of protein flexibility and binding site water into predictive computer aided drug design (CADD) was obtained. The described methods can be transferred to new targets and their homologue off-targets for selectivity predictions in future projects.

Chapter 2

Materials and Methods

2.1 Experimental Methods

2.1.1 Reagents and Chemicals

Chemicals, media and reagents were purchased from Carl Roth, Sigma Aldrich, VWR and Alfa Aesar. Myristoyl-Coenzyme A (MyrCoA) was obtained from Santa Cruz Biotech, the substrate peptide $pp60^{src}_{(2\text{-}9)}$ with sequence GSNKSKPK-amidation was synthesized by SelleckChem. Inhibitors **1-5** were synthesized by Edmond Fleischer and Christoph Borek from MicroCombiChem according to reported methods (Brand et al., 2012, Brand et al., 2014 and Hutton et al., 2014). Inhibitors' identity and purity was confirmed by NMR and LC-MS.

2.1.2 Transformation and Site-Directed Mutagenesis

Plasmids coding for *Lm*NMT amino acids 11-421 (addgene #25200), *Hs*NMT1 amino acids 108-496 (#25277) and *Hs*NMT1 amino acids 115-496 (#25272) with an N-terminal hexahistidine tag and a *Tobacco etch virus* protease (TEV) cleavage site incorporated into a pET15-MHL vector were purchased from addgene. Plasmids were extracted from delivered stab cultures using the QIAprep Spin Miniprep Kit (QIAGEN) and their identities were checked using the sanger sequencing service from GATC Biotech with T7 and pET-RP primers.

The isolated plasmids were transformed using heat-shock into Rosetta2™ competent cells (Merck Millipore) or stored at -20°C until use for site-directed mutagenesis.

Site-directed mutagenesis to introduce point mutations into the coding region of NMT sequences was performed using the standard protocols of Quick Change® Lightnung Multi Site-Directed Mutagenesis Kit for multiple mutations within one PCR reaction or Quick Change® II Site-Directed Mutagenesis Kit for single mutations with a ^{3}Prime Thermal Cycler (Techne). For the amino acid exchange of *Hs*NMT1-A452M requiring a mismatch of three bases (GCA to ATG) the annealing temperature was raised by 5 °C. Primers were designed using Agilent's mutagenic primer design tool

(http://www.genomics.agilent.com/primerDesignProgram.jsp) with the help of the *E. coli* codon usage table (Sharp et al., 1988) and were purchased from eurofins Genomics at HPLC grade or higher (2.1). The PCR product was subsequently transformed into XL-10 gold ultracompetent cells and plated on LB agar containing ampicillin and chloramphenicol as resistance markers. Single colonies were picked the day after and grown in LB medium supplemented with the same antibiotics overnight. Subsequently, the plasmid was isolated using QIAprep Spin Miniprep Kit (QIAGEN) and sequenced at GATC Biotech using pET-RP primer. Sequencing results were analysed using web tools of Clustal Omega (Sievers et al., 2011, Li et al., 2015, McWilliam et al., 2013, http://www.ebi.ac.uk/Tools/msa/ clustalo/) and Reverse complement (www.bioinformatics.org) and compared to the addgene depositor sequence. Desired plasmids were transformed into Rosetta2^{TM} competent cells (Merck Millipore) for protein expression. Glycerol stocks of all NMT expressing cell lines were prepared and stored at -80 °C.

TABLE 2.1: List of mutagenic primers.

Mutation	**Primer**
*Lm*NMT H398N	TTGGCGCCGGCGACGGTAACCTTCGATATTACTTCTAC
*Lm*NMT M420L	CAAGCTTCGTCATCACAGCAGCACCAAGGCAAC, GTTGCCTTGGTGCTGCTGTGATGACGAAGCTTG
*Lm*NMT L421Q	GTTGCCTTGGTGATGCAGTGATGACGAAGCTTG, CAAGCTTCGTCATCACTGCATCACCAAGGCAAC
*Lm*NMT M420L:L421Q	TCAGGTTGCCTTGGTGCTGCAGTGATGACGAAGCTTG
*Hs*NMT1 R295Q	GACCGATGCCAATACTGGCAGGTGCCAACGGG, CCCGTTGGCACCTGCCAGTATTGGCATCGGTC
*Hs*NMT1 W297F	CCCGTTGGCACCTGCCAGTATTGGCATCGGTC, CCGTTGGCACCTGCAGGTATTTCCATCGGTCCCTA
*Hs*NMT1 A452M	TCGCCAAAATGAAAGGGTTTGATGTGTTCAATATGCTGGATCTCATGGAGAAC, AGCGGTTTTACTTTCCCAAACTACACAAGTTATACGACCTAGAGTACCTCTTG
*Hs*NMT1 L453V	AAGGGTTTGATGTGTTCAATGCAGTGGATCTCATGG, TTCCCAAACTACACAAGTTACGTCACCTAGAGTACC
*Hs*NMT1 A452M:L453V	TCTCCATGAGATCCACCATATTGAACACATCAAACCC, GGGTTTGATGTGTTCAATATGGTGGATCTCATGGAGA
*Hs*NMT1 L462V	GGGTTTGATGTGTTCAATATGGTGGATCTCATGGAGA, ATGGAGAACAAAACCTTCGTGGAGAAGCTCAAGTTTG
*Hs*NMT1 N473H	GGCATAGGGGACGGCCACCTGCAGTATTACC
*Hs*NMT1 L495M	ATGGAGAACAAAACCTTCGTGGAGAAGCTCAAGTTTG
*Hs*NMT1 Q496L	GGTTGGACTGGTGCTACTATGATGACGAAGCTTGC
*Hs*NMT1 L495M:Q495L	CAGAGAAGGTTGGACTGGTGATGCTGTGATGACGAAGCTTGCGGCC

2.1.3 Design of *Hs*NMT1 *8x* Mutant

In addition to successively increasing the number of exchanged residues to obtain *Hs*NMT1 R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L (*Hs*NMT1 *8x*), a plasmid with *Hs*NMT1 containing all 8 mutation sites in a pET-15b vector was designed and ordered from Genscript. The corresponding vector map is illustrated in Figure 2.1.

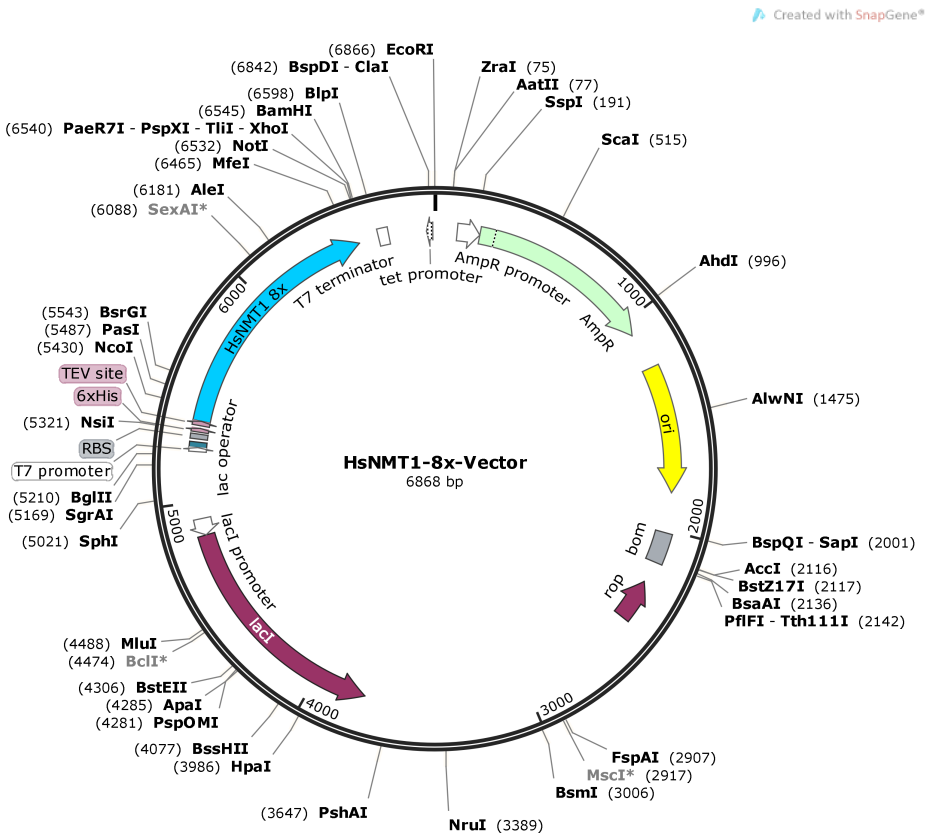


FIGURE 2.1: Vector map of *Hs*NMT1 *8x* (R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L). Illustrated with SnapGene® Viewer 2.8.2 (GSL Biotech LLC).

2.1.4 Recombinant Protein Expression and Purification

Protein purification was conducted according to previously reported methods (Frearson et al., 2010). NMT was expressed in *E.coli* Rosetta2^{TM} strains in the presence of 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. A single colony was inoculated in 8 ml LB medium for 8 h and transferred into 100 ml LB and grown at 37 °C overnight while shaking at 225 rpm. After a second transfer into 3-4 l TB medium, the optical density was checked in regular intervals. After reaching an OD_{600} of 0.6 - 0.8 the temperature was lowered to 20 °C and protein expression was induced using 500 μM

isopropyl β-D-1-thiogalactopyranoside (IPTG).

Lysogeny broth (LB medium, 1 l)

Tryptone	10 g
NaCl	5 g
Yeast extract	5 g

Terrific broth (TB medium, 1 l)

Tryptone	12 g
Yeast extract	24 g
Glycerol	4 ml
KH_2PO_4	2.3 g
K_2HPO_4	12.5 g

The next day, the culture was harvested by centrifugation (30 min at 4 000 g) and resuspended in up to 120 ml lysis/loading buffer. Lysozyme, bovine pancreas DNAse I and a *cOmplete*™ EDTA-free protease inhibitor cocktail tablet (Merck Millipore) were added for 30 min pretreatment prior sonication. After centrifugation for 45 min at 17 000 g and 4 °C the supernatant was filtered (0.22 μm) and the lysate was loaded onto a 5 ml Histrap HP column using an ÄKTA purifier system (GE Healthcare) and a 150 ml superloop at a flow rate of 4 ml/min. By a smooth gradient increasing the elution buffer's concentration the column was washed to remove histidine-rich *E. coli* proteins and NMT was eluted at around 170 mM imidazol. The fractions corresponding to the NMT peak were pooled and concentrated using Amicon® Ultra-15 (Merck Millipore) or Vivaspin® 20 (Sartorius) centrifugal filter units. Identity and purity were checked by SDS-PAGE gel with Coomassie blue staining and the concentration was determined by absorption at 280 nm using a Nanodrop 1000.

Lysis/Loading buffer

HEPES	50mM
NaCl	500 mM
Imidazole	5 mM
Glycerol	5 %
pH 7.5	

Elution buffer

HEPES	50 mM
NaCl	500 mM
Imidazole	250 mM
Glycerol	5 %
pH 7.5	

2.1.5 M9 Mineral Medium Expression and Purification

For upcoming NMR experiments (s. Chapter 7.2.3) *Hs*NMT1 wildtype protein was expressed in M9 mineral medium. For that purpose, an overnight culture from 150 ml LB medium was harvested by centrifugation and resuspended in 4 l of M9 mineral medium. Induction and purification was performed as described above. For labeling, the nitrogen source NH_4Cl or the carbon source glucose can be exchanged to $^{15}NH_4Cl$ and ^{13}C-glucose respectively. Alternatively, deionized water can be exchanged to D_2O. Purity and yield of 18.5 mg were comparable to standard purification procedure.

M9 mineral medium, 1 l

M9 salt solution (10X)	100 ml
20% glucose	20 ml
1M $MgSO_4$	1 ml
1M $CaCl_2$	0.3 ml
1 mg/ml biotin	1 ml
1 mg/ml thiamin	1 ml
trace element solution (100X)	10 ml
H_2O	ad 1000 ml

M9 salt solution (10X)

Na_2HPO_4*$2H_2O$	75.2 g/l
KH_2PO_4	30 g/l
NaCl	5 g/l
NH_4Cl	5 g/l

Trace element solution (100X)

EDTA	13.4 mM
$FeCl_3$*$6H_2O$	3.1 mM
$ZnCl_2$	0.62 mM
$CuCl_2$*$2H_2O$	76 μM
$CoCl_2$*$2H_2O$	42 μM
H_3BO_3	162 μM
$MnCl_2$*$4H_2O$	8.1 μM

All stock solutions were autoclaved or sterilized by filtration before use.

2.1.6 Activity and Inhibition Assays

A fluorescence based activity and inhibition assay was performed as described previously (Goncalves et al., 2012a). It is based on the detection of free CoA, that is released during myristoylation reaction, with 7-diethylamino-3-(4-maleimidyphenyl)-4-methylcoumarin (CPM, Figure 2.2).

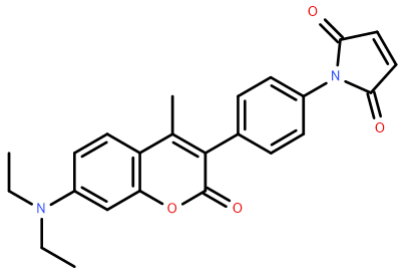


FIGURE 2.2: Structure of 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM)

The fluorescence intensity was recorded on a Tecan M200 infinite® pro or Spark® microplate reader with 380 nm excitation wavelength and 470 nm emission wavelength. First, 75 μl of assay buffer containing NMT and MyrCoA were placed in each well. For kinetic measurements and positive controls 10 μl of 10 % DMSO solution in assay buffer or 10 μl inhibitor and 10 % DMSO containing buffer were added. The reaction was started by injection of 25 μl containing CPM and the substrate $pp60^{src}{}_{(2\text{-}9)}$.

Assay buffer

Na_2HPO_4/NaH_2PO_4	50 mM
EDTA	0.5 mM
Triton-X 100	0.05 %
pH 7.8	

For IC_{50} determination final assay concentrations were: 8 nM NMT, 8 μM CPM, 4 μM MyrCoA, 4 μM $pp60^{src}{}_{(2\text{-}9)}$ and inhibitor at half-logarithmic step concentrations ranging from 0.1 nM to 100 μM (depending on the tested compound) resulting in a final DMSO content of 0.9 %. A negative control containing no NMT and a positive control without inhibitor was also performed by default. The assay was run over 30 min and stopped by adding 60 μl of 100 mM acetate buffer pH 4,5. IC_{50} values were calculated from the stable fluorescence signal against inhibitor concentration using the nonlinear

4 parameter logistic fit (Equation 2.1) of Grafit 5.0.13 (Leatherbarrow, 2009).

$$y = \frac{Range}{1 + (\frac{[I]}{IC_{50}})^s} + Background \tag{2.1}$$

where Range is corresponding to the difference in fluorescence signal between 0 % and 100 % activity, Background corresponds to the fluorescence signal of 0 % activity, [I] the inhibitor concentration and s the slope factor.

The inhibition constant K_i was determined from K_m and IC_{50} with the Cheng-Prusoff Equation 2.2 (Cheng and Prusoff, 1973).

$$K_\mathrm{i} = \frac{IC_{50}}{1 + \frac{[S]}{K_\mathrm{m}}} \tag{2.2}$$

For K_m determination, the final concentration of CPM was 20 μM. MyrCoA was added to a saturated concentration of 30 μM and the substrate $pp60^{src}_{(2\text{-}9)}$ was added in doubling steps from 1 to 32 μM or vice versa. The assay was run as duplicates over 30 min and K_m was calculated from the slope of the initial 10 min of reaction for the different substrate or Cofactor concentrations using the nonlinear fit of the Michaelis-Menten Equation 2.3 in Grafit with v as reaction speed, [S] the substrate concentration and v_{max} as the maximum reaction rate at saturation.

$$v = \frac{v_\mathrm{max} * [S]}{K_\mathrm{m} + [S]} \tag{2.3}$$

2.1.7 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were performed on a Microcal VP-ITC at 20 °C. The protein buffer was exchanged using Amicon® Ultra-15 (Merck Millipore) or Vivaspin® 20 (sartorius) centrifugal filter units. The inhibitors were diluted into the collected flow-through from a 10 mM stock solution containing 20 % or 50 % DMSO. Therefore an equivalent volume of pure DMSO was added to the protein solution to match the buffers. If not mentioned otherwise, the final concentration was 10 μM for NMT and 100 μM for inhibitor. To both solutions, MyrCoA was added to a final concentration of 40 μM. In displacement titration experiments, the low-affinity ligand (indole inhibitor, compound **5**) was added to result in 40 μM final concentration in both solutions. All ITC experiments were performed at least three times.

ITC buffer

Na_2HPO_4/NaH_2PO_4	50 mM
EDTA	0.5 mM
β-mercapto ethanol	1 mM
pH 7.5	

For elucidation of proton transfer effects upon ligand binding experiments in buffer with higher ionization enthalpy (HEPES) were performed as well.

ITC HEPES buffer

HEPES	50 mM
EDTA	0.5 mM
β-mercaptoethanol	1 mM
pH 7.5	

Results were analysed using the ITC routines of Origin® 7 (OriginLab, Northampton, MA). Curves were fitted with the one binding-site iterative nonlinear least square model provided by the software. Dissociation constants and free energies were calculated from the following equations:

$$\Delta G = \Delta H - T\Delta S \tag{2.4}$$

$$\Delta G = -RT * lnK_{\mathrm{D}} \tag{2.5}$$

with T as absolute temperature in Kelvin and the ideal gas constant being R = 8.3144598 J / mol*K.

In displacement titration experiments ΔH and K_d were calculated as described by Zhang and Zhang, 1998:

$$K_{\mathrm{d_2}} = (\frac{K_{\mathrm{d_1}}}{K_{\mathrm{d_{app}}}} - 1) * \frac{1}{[L_2]} \tag{2.6}$$

$$\Delta H_2 = (\Delta H_1 - \Delta H_{\mathrm{app}})(1 + \frac{1}{K_{\mathrm{d_2}}} * [L_2]) \tag{2.7}$$

with K_{d_1} and ΔH_1 being the dissociation constant and binding enthalpy of the stronger binding ligand (which displaces the low-affinity ligand L_2 during the experiment). K_{d_2} and ΔH_2 are the dissociation constant and binding enthalpy of the displaced ligand with the concentration $[L_2]$. $K_{d_{app}}$ and ΔH_{app} are the observed results during the displacement titration experiment.

2.1.8 Protein Crystallization

Starting conditions for protein crystallization were derived from the Protein Databank PDB (Berman et al., 2000, www.rcsb.org) entries 3H5Z, 3IU1, 4C2Y, 4C2Z and the corresponding publications (Frearson et al., 2010; Thinon et al., 2014). Screening around the published conditions by variation of pH, salt and PEG content led to the following crystallization buffers giving the best and reproducable crystallization results:

***Lm*NMT crystallization condition**

PEG 1500	32 %
NaCl	0.2 M
Sodium cacodylate	0.1 M
pH 5.5	

***Hs*NMT1 crystallization condition A**

PEG 3350	22 %
Diammonium hydrogen citrate	0.2 M
pH 5.7	

***Hs*NMT1 crystallization condition B**

PEG 4000	24 %
$NiCl_2$	5 mM
Sodium citrate	0.1M
Glycerol	2.5 %
pH 4.5	

***Hs*NMT1 crystallization condition C**

MPEG 2000	25 %
KBr	0.2 M
Sodium citrate	0.1M
Glycerol	5 %
pH 4.5	

The protein in 50 mM sodium phosphate buffer at pH 7.5 was concentrated to 10 mg/ml for *Lm*NMT and 5 - 8 mg/ml for *Hs*NMT1 and its mutants. NMT was always incubated with MyrCoA (1 mM) and with or without inhibitors (0.5-1 mM) for 1 h on ice. Experiments were performed with the hanging drop vapour diffusion method with crystallization drops consisting of 2-3 μl protein solution plus 2 μl reservoir at 8 °C or 20 °C. Crystals usually grew within 24 h (48 h at 8 °C for *Hs*NMT1 crystallization conditions C). Crystals were measured under cryo conditions. Thus, the

harvested crystals were transferred for around 20 s into cryoprotectant solutions, which contained *Hs*NMT1 crystallization conditions B or C with the glycerol concentration adjusted to 16 %, before flash freezing.

2.1.9 Structure Determination

*Hs*NMT1 X-ray diffraction data were collected from single crystals on synchrotron beamline ID29 at the European Synchrotron Radiation Facility (ESRF) in Grenoble with a Pilatus 6M (Decritis LTD) detector. Crystals were maintained at 100 K cryostream. The used wavelength was 0.976 Å.

*Hs*NMT1 in complex with compound **7** and *Hs*NMT1 *8x* crystals were collected with a Bruker AXS Microstar-H generator with a MAR Scanner 345 mm image-plate detector with a X-ray wavelength of 1.5417 Å.

Data were processed, integrated and scaled with ccp4 (Leslie, 2006; Winn et al., 2011; Battye et al., 2011) or XDS (Kabsch, 2010). The structures were solved by molecular replacement and the model built with PhaserMR (McCoy et al., 2007) within the phenix worksuite (Adams et al., 2010). PDB entries 3IWE and 4C2Y were used as template structures, which were cleaned by removing the cofactor MYA (MyrCoA), the ligand 646 (compound **1**) and all water molecules before use. Real space molecular replacement and ligand placement was performed using *Coot*-0.8.3 (Emsley et al., 2010). Final refinement and water molecule placement was performed with phenix.refine. The parameters for MyrCoA and ligands (**2**, **4**, **5** and **7**) were generated using the *electronic Ligand Builder and Optimization Workbench* (eLBOW; Moriarty, Grosse-Kunstleve, and Adams, 2009). Crystallographic data and refinement analysis was validated with phenix.

2.1.10 Surface Plasmon Resonance

Preliminary SPR data for the interactions and binding kinetics between NMT and compounds **1**, **2**, **5** were obtained using a Biacore^{TM} T200 (GE Healthcare). Histidine-tagged NMT was immobilized onto a NTA sensor chip using recommended standard protocols (Instruction 22-0519-97 AF, Sensor Chip NTA, Biacore^{TM}). Experiments were performed at 25 °C using HBS-P+ buffer enriched with 100 nM MyrCoA as described previously (Frearson et al., 2010). Interactions were measured as multi-cycle kinetics with a protein-free, but nickel-loaded reference cell. NMT was immobilized at at a concentration of 20 μM in HBS-P+ buffer on the sample cell. Binding was tested using varying inhibitor concentrations (compound 1: 0.5 nM - 1 μM, compound 2: 0.1 nM - 500 nM, compound 5: 100 nM - 500 μM) at a flow rate of 30 μl/min with final DMSO concentration of 0.1 %. Data analysis was performed with Biacore T200 evaluation software version 2.0.

HBS-P+ running buffer

HEPES	10 mM
NaCl	150 mM
Surfactant P20	0.005 %
pH 7.4	

2.2 Computational Methods

2.2.1 Pose Generation for MDs by docking

Molecular dynamic simulations (MDs) were started with modified crystal-structures. In those cases where no crystal structure of the protein-ligand complex was available in the PDB, the starting structures were generated by docking using LeadIT-2.1.6 (www.biosolveit.com) with the FlexX scoring function (Rarey et al., 1996). For the sulfonamide compounds **2-4** PDB entry 2WSA was chosen for *Lm*NMT and 3IWE, chain A, for *Hs*NMT1 as the receptor. Ligand-free 3IU1, chain B and 4C2Y, chain A were chosen for compound **5** to generate complexes with *Hs*NMT1 due to the different open-conformation binding mode.

The receptor was defined by choosing compound **1** as the reference ligand and including all residues within 6.5 Å of this compound, as well as the cofactor MyrCoA. Protonation was done by Protoss within the LeadIT worksuite and manually refined so that His219 δ-tautomer for *Lm*NMT formed a hydrogen bond to the neighbouring Asn376. The same was done for His298 and Asn451 in the human homologue. Water molecules with less than three interactions to the receptor were removed, those with three interactions were chosen to be displaceable. The receptor model was validated by redocking the reference ligand and inspection of the resulting pose and its root mean square deviation (RMSD, Equation 2.8)

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2} \tag{2.8}$$

where δ_{i} is the distance (in Å) between atom i and the reference atom and N the number of heavy atoms of the ligand.

Compounds **2-5** were energy minimized with the *Merck Molecular Force Field* (MMFF94; Halgren, 1996) and docked in their dominant protonation state determined by MarvinSketch (ChemAxon Ltd., http://www.chemaxon.com). The following default parameters were chosen for docking: hybrid approach of enthalpy and entropy for base placement, application of the same settings to all H-bond relevant contact types with full score contribution threshold 0.3 and no score contribution threshold 0.7. Chemical ligand parameters allowed no stereo mode generation. Clash handling had a hard allowed overlap volume of 2.9 $Å^3$ and an intra-ligand clash factor of 0.6 considering hydrogens in internal clash test. The maximum number of solutions per iteration and fragmentation was 200 for both. The top 10 poses for each ligand were rescored using the HYDE scoring function (Reulecke et al., 2008) for hydrogen bonding and dehydration. The binding modes with the highest HYDE score were chosen as starting structure for MDs.

2.2.2 Small Molecule Parameterization

Compounds **1-5** as well as MyrCoA were parameterized depending on the force field used for MDs. For use with Amber ff99SB (Cornell et al., 1995) atom types from antechamber (Wang et al., 2006) and parameters for bonds, angles, dihedrals and non-bonded parameters derived from the generalized amber force field (GAFF, Wang et al., 2004) using parmchk2 from the AmberTools (AMBER 2016, University of California, San Francisco.) were used. Atom partial charges were computed from the molecular electrostatic potential (MEP) calculated with gaussian09 (Gaussian 09, Revision C.01,Gaussian, Inc., Wallingford CT, 2009, Frisch et al., 2009) at a HF/6-31G* level of theory and the thereby obtained restrained electrostatic potential (RESP, Bayly et al., 1993) was incorporated with antechamber.

Small molecule parameters for use with CHARMM36 (MacKerell et al., 1998; Mackerell, Feig, and Brooks, 2004; Huang and Mackerell, 2013) were taken from CGenFF (Vanommeslaeghe et al., 2010; Mackerell, Feig, and Brooks, 2004; Vanommeslaeghe and MacKerell, 2012; Vanommeslaeghe, Raman, and MacKerell, 2012) using the web service (https://cgenff.paramchem.org) of the University of Maryland. Due to instabilities in binding mode during MDs for compounds **1** and **2** parameters for angles and dihedral angles were obtained and adapted from the MMFF using the web server from the molecular modeling group at the Swiss institute of bioinformatics (SIB, Zoete et al., 2011, http://www.swissparam.ch/).

Parameters for MyrCoA from CGenFF were of generally high quality because of its similarity to well parameterized molecules like ATP and fatty acids. Only the low quality parameters (high penalty score) for the thioester bond - which had been derived from standard ester parameters by closest analogy - were refined. Therefore, CHARMM-compatible parameters were calculated using the force field toolkit (FFtk, Mayne et al., 2013) within VMD 1.9.2 (Humphrey, Dalke, and Schulten, 1996) for fitting the gaussian09 quantum mechanic (QM) -derived (HF/6-31G*) parameters for bond lengths, angles, dihedrals and charges taking S-Ethyl thiopropionate as reference structure for QM calculations.

2.2.3 Molecular Dynamics Simulations

NMT starting structures were prepared from the PDB entries (IDs 3H5Z, 4CGP, 2WSA, 4CGN, 5A27, 5A28, 3IU1, 4C2Y, 3IWE) and corresponding docking poses as described in Chapter 3.3.1. As *Hs*NMT1's asymmetric unit is a dimer, generally chain A was chosen for structure preparation. Because of missing residues in chain A for 3IU1, chain B was prepared in this case. For 2WSA, Asp27 was changed to Thr to give the correct sequence for *Lm*NMT. Histidine residues were protonated in ϵ position besides those

mentioned in Chapter 2.2.1 and His12 in *Lm*NMT, which was chosen to be two times protonated to form ionic interactions with the phosphate moiety from cofactor MyrCoA. Alternative side chain conformations and atoms not belonging to the protein, cofactor, ligands and crystallographic water were removed.

For simulations of mutated proteins related amino acids were exchanged in PyMOL® (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.) and energy minimized with MOE2015.1001 (Molecular Operating Environment (MOE), 2015.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7, 2017.) using the AMBER12:EHT force field (Cornell et al., 1995) keeping the protein rigid except for the selected residues.

Simulation systems were subsequently built with tleap and minimized with Sander from AmberTools 13 over 200 minimization cycles using the generalized Born implicit solvent model (Onufriev, Bashford, and Case, 2004) with the ff99SB force field parameters. The structure was then neutralized with sodium ions and solvated in a TIP3P water box (Jorgensen et al., 1983) exceeding the protein by at least 10 Å. For use with the CHARMM 36 forcefield, the system was built similarly within VMD 1.9.2 and minimized after neutralization and solvation over 2000 time steps.

The system equilibration was performed with harmonic constraints applied to all non-water atoms and the system was heated from 100 to 300 K over 500 ps using NAMD 2.11 (Phillips et al., 2005). In the next step, the constraints were gradually released over the first 500 ps in a constant-volume box. Periodic boundary conditions were applied to the system. Van der Waals cut-off was set to 14 Å for use with CHARMM 36 and 12 Å for use with the AMBER force field ff99SB and the particle mesh Ewald methodology (Darden, York, and Pedersen, 1993) for electrostatic interactions. The different cut-offs are caused by the intrinsic difference in the origin of the force field parameters.

For the following production runs, constant pressure and constant temperature were achieved by the use of Nosé-Hoover Langevin piston pressure control (Martyna, Tobias, and Klein, 1994) and Langevin dynamics (Feller et al., 1995). 2 fs time-steps were used in combination with rigid bond lengths and trajectories were written every ps. Production runs were carried out for 50 ns on the graphics processing units (GPUs) on the high performance computing (HPC) cluster "Mogon" of the university Mainz.

2.2.4 Data Analysis

Frames of the trajectories were aligned to their starting structures protein backbone atoms C, C_α and N in VMD 1.9.2. (Two-dimensional) Root mean square deviation (2D-RMSD and RMSD), root mean square fluctuation

(RMSF, Equation 2.9), B-factors and water occupancy maps were calculated by VMD and included tcl-scripts.

$$RMSF = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\delta_t^2} \tag{2.9}$$

δ_t is the distance (in Å) between an observed atom at every frame and its average atom position, t the current frame and T the final frame.

Convergence of the simulations was checked using the 2D-RMSD (Koukos and Glykos, 2014).

Order parameters (S^2 or OP, MacRaild et al., 2007, Equation 1.6) for N-H bond vectors, side chain bond vectors and the corresponding conformational entropy were obtained by the *iRED* approach (Prompers and Brüschweiler, 2002) with cpptraj using the Brüschweiler dictionary for sidechain entropies (Li and Brüschweiler, 2009). Under the assumption that $S^2_{NMR} \approx S^2_{MD}$ (Showalter et al., 2007), MD-derived S^2 were used for an estimation of entropy changes upon ligand binding (see also Chapter 1.4.1).

Principal component analysis (PCA) for cross-correlation maps of concerted movements and community determination were performed with the Bio3D package (Grant et al., 2006; Skjaerven et al., 2014) for R (R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.).

For quantitative analysis of single water sites within the binding pocket the SPAM approach of Cui, Swails, and Manas, 2013 was used for water energy calculation within cpptraj (Chapter 1.4.2). For that purpose, bulk water energy terms were derived from a 10 ns MD of a pure TIP3P waterbox with the dimensions of 40 Å* 40 Å* 40 Å.

2.2.5 Virtual screening

The virtual screening (VS) was performed using the ZINC 15 database (Irwin et al., 2012; Sterling and Irwin, 2015) with its *Tranch Browser* to obtain prefiltered protomers. A further molecular properties filter step was performed using MOE within KNIME (Berthold et al., 2007). After conformer generation using omega2 (Hawkins et al., 2010) a pharmacophore query was performed within MOE. The passing molecules were docked against *Lm*NMT (PDB-ID 5A28) using LeadIT-2.1.6 with the parameters for posing and scoring as described above, keeping water molecules with three interactions as part of the receptor. The 300 molecules with the most negative score were rescored using the HYDE scoring function and visually inspected prior ordering. The exact filters and numbers of passing molecules are described in Chapter 6. The ordered compounds from ChemBridge / Hit2Lead

(**ChBr1-4**) and WuXi Apptec / LabNetwork (**WuXi1-2**) were checked for purity (at least 95 %) and identity by LC-MS.

2.2.6 Software

The following list summarizes all software packages that were used during this project (corresponding citations in the text):

AmberTools	Worksuite: antechamber, cpptraj, ParmED, pharmchk2, sander, tleap
antechamber	Assigning atomtypes and RESP fitting for use with AMBER force field
Bio3D	PCA, cross-correlation and community analysis of MDs
CGenFF	Assigning atomtypes and parameters for use with CHARMM force field
ClustalW	Sequence alignment
cpptraj	Order parameter and sidechain entropy calculation, SPAM energy calculation, part of AmberTools
gaussian09	QM calculations
Grafit	Analysis of enzymatic assay data
Knime	Workflow automation of database handling and docking
LeadIT/FlexX	Docking
MarvinSketch	pK_a prediction
MOE	Alignments, pharmacophore search and visualization, database handling
NAMD	Running MDs
omega2	Conformer generation
Origin	Analysis of ITC data
ParmED	CHARMM to AMBER parameter file conversion
pharmchk2	Small molecule parameterization for use with AMBER force field
PyMOL	Visualization
R	Statistics and evaluation
sander	Energy minimization, part of AmberTools
SnapGene Viewer	Vector map illustration
swissparam	Small molecule parameterization for use with CHARMM force field
tleap	Structure preparation, part of AmberTools
VMD	Preparation and analysis of MDs, visualization

Chapter 3

Characterization of *Lm*NMT and *Hs*NMT1 wildtypes

Even though the binding sites of *Lm*NMT and *Hs*NMT1 are highly conserved, selective inhibitors were reported (Chapter 1.2.4; Brand et al., 2012; Brand et al., 2014; Brannigan et al., 2014; Hutton et al., 2014; Rackham et al., 2015). For determination of the selectivity determining features in NMT, inhibition of *Lm*NMT and *Hs*NMT1 by compounds **1-5** were characterized in a fluorescence based enzyme inhibition assay (Chapter 3.1; Goncalves et al., 2012a). The thermodynamic profile of ligand binding was obtained from ITC experiments (Chapter 3.2). For first hints on selectivity, already published crystal structures were analysed (Chapters 1.2 and 3.3.4; Frearson et al., 2010; Brand et al., 2012; Brannigan et al., 2014) and the inhibitors with no crystal structure in complex with NMT available, were docked for binding mode identification and interaction energy estimation (Chapter 3.3.1).

3.1 Enzyme Inhibition

Kinetic and inhibitory data for *Lm*NMT, *Hs*NMT and inhibitors **1-5** were determined for validation of the established assay with substrate GSNKSKPK-amidation ($pp60^{src}_{(2\text{-}9)}$) and for comparison with previously reported results (Frearson et al., 2010; Brand et al., 2012; Brand et al., 2014; Bell et al., 2012; Brannigan et al., 2014; Hutton et al., 2014).

TABLE 3.1: Michaelis constants of *Lm*NMT and *Hs*NMT1 for MyrCoA and $pp60^{src}_{(2\text{-}9)}$. Values in parentheses are previously reported results. n.d. = not determined)

Enzyme	K_m (MyrCoA) [μM]	K_m ($pp60^{src}_{(2\text{-}9)}$) [$\mu$M]
*Lm*NMT	4.9 ± 0.8 (1.4)	5.8 ± 3.4 (n.d.)
*Hs*NMT1	2.9 ± 1.9 (8.2)	2.6 ± 0.4 (2.8)

The obtained Michaelis constants were in agreement with previously published data (Table 3.1; Goncalves et al., 2012b; Panethymitaki et al., 2006).

Inhibition data (IC_{50}) were obtained for inhibitors **1-5** and converted to K_i for better comparability between the two enzymes. The selectivity index (SI) was calculated by division of K_i from *Hs*NMT1 by K_i from *Lm*NMT (Table 3.1).

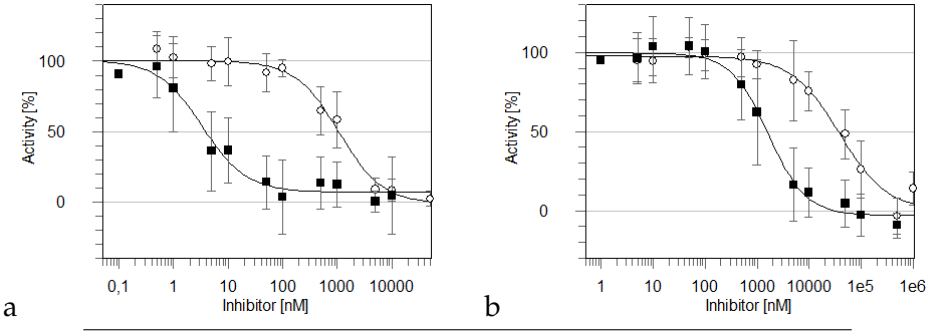
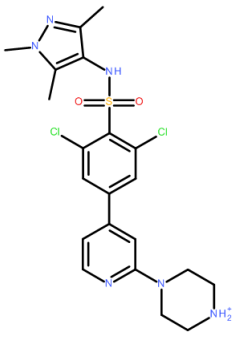
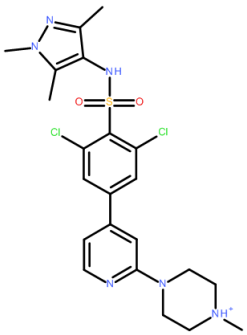
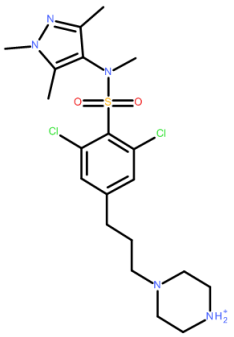
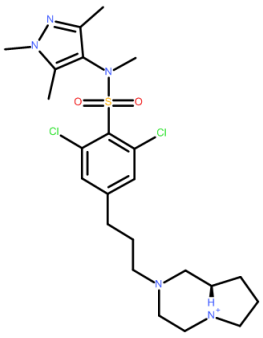
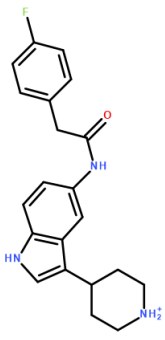


FIGURE 3.1: Inhibition of *Lm*NMT and *Hs*NMT1 by compound 4 (a) and compound 5 (b). White circles show datapoints for *Hs*NMT1, black filled squares for *Lm*NMT

The inhibitors of the sulfonamide series (**1-3**) show strong inhibition of 1.4 to 19.8 nM and 13.3 to 96.4 nM, respectively, but only low selectivity of factor 9 or less (Table 3.2). Compound **4** inhibited strongly *Lm*NMT (2 nM), comparable to the strongest ligand, compound **2**, but was highly selective over *Hs*NMT1 with a SI of 213. Compound **5** with an indole core adopting a different binding mode as described in Chapter 1.2.4 (Figure 1.13), exhibited moderate selectivity in the assay (SI 16), but lower affinity overall of 1.0 and 15.4 μM, respectively (Figure 3.1).

Additionally, a derivative of compound **5** with a pyridine replacing the p-fluorophenyl ring was tested (Figure 3.2). In the crystallographically determined binding mode of compound **5**, the fluorophenyl moiety is placed in the same pocket as the trimethyl-pyrazole moiety of the sulfonamides. However it is not able to form a hydrogen bond to Ser330/405. This interaction was previously found to be important for good binding affinity (Brand et al., 2012). In the designed pyridine derivate of compound **5**, the H-bond could be formed which might lead to an increase of the binding affinity. This compound (Figure 3.2) inhibited *Lm*NMT only slightly stronger than compound **5** with and IC_{50} of 953.0 $\pm$ 115.2 nM (K_i = 564.0 $\pm$ 68.2 nM) and *Hs*NMT1 weaker with 46.1 $\pm$ 22.7 μM (K_i = 18.3 $\pm$ 9.1 μM). However, it was not followed up because of instability and degradation issues.

TABLE 3.2: Inhibition of *Lm*NMT and *Hs*NMT1 by compounds (Cpd.) **1-5**. K_is in nM.

Cpd.	Structure	K_i (*Hs*NMT1)	K_i (*Lm*NMT)	SI
1		31.6 ± 4.5	8.4 ± 1.3	4
2		13.3 ± 2.8	1.4 ± 0.3	9
3		96.4 ± 11.7	19.8 ± 2.5	5
4		428.2 ± 90.8	2.0 ± 0.4	213
5		15 388.1 ± 4 3670.0	975.8 ± 102.1	16

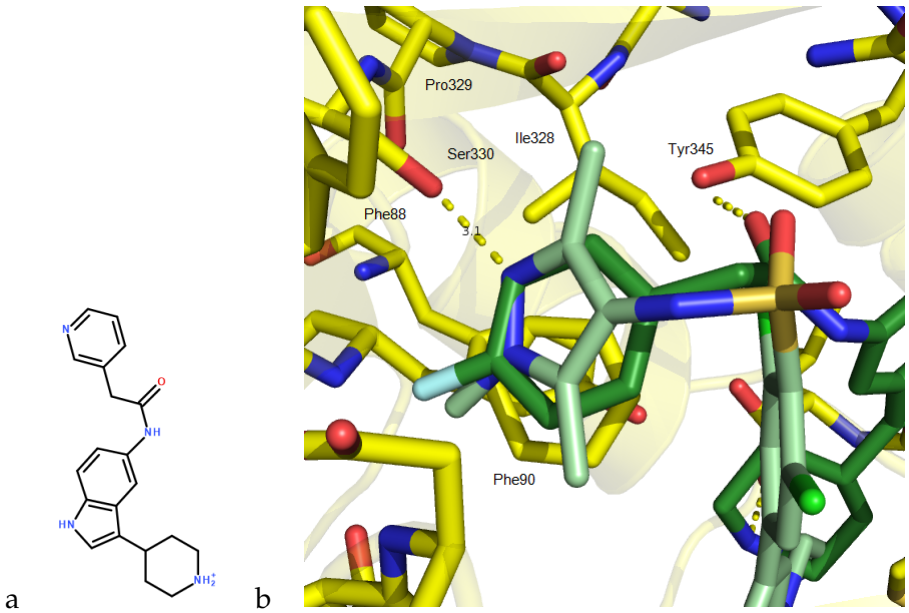


FIGURE 3.2: a) Structure of pyridine derivative of compound **5**. b) Illustration of reasoning to exchange fluorophenyl moiety by pyridine due to proximity to Ser330 in *Lm*NMT. Pale green ligand compound **1** (PDB-ID 2WSA), dark green ligand compound **5** (PDB-ID 4CGN). Ortho cabon atom from fluorophenyl ring overlaps with nitrogen of pyrazole moiety. Replacement of carbon by nitrogen in the pyridine derivate of **5** was expected to have favourable orientation and distance (3.1 Å) to form a H-bond with Ser330.

3.2 Thermodynamics of NMT-Ligand Binding

For thermodynamic characterization of ligand binding to NMT, ITC experiments were performed. The results (Figure 3.3, Table 3.3 and Table 3.4) confirmed the affinity and selectivity profile of inhibitors **1-5** derived from the inhibition assay. The slightly lower selectivity of compound **4** (K_d = 21.4 nM for *Lm*NMT and 183.8 nM for *Hs*NMT1) might be caused by limitations of the assay at very low IC_{50}. The higher SI of 114 for compound **5** can be explained by the ITC method. Affinity and thermodynamics were not directly detectable for this compound, but only by displacement titration. Thereby some inaccuracies may have arisen.

For *Lm*NMT a generally higher enthalpic contribution was observed for all ligands. The difference in ΔH ($\Delta\Delta H$) for non-selective ligands **1-3** binding to *Lm*NMT and *Hs*NMT1 were 15.0, 7.7 and 20.8 kJ/mol with a similar K_d for both enzymes. This indicated strong interaction, but slight conformational adaption in *Lm*NMT whereas the human homologue seemed to be highly preordered resulting in a lower entropic penalty. This led to a hypothesized loss of flexibility upon ligand binding to *Lm*NMT compared to *Hs*NMT1 (see also Chapter 3.3.4).

Compounds **3** and **4** both have a flexible alkyl linker to the basic ring system. For **3** the entropic contribution upon binding was reduced (-$T\Delta S$

of -8.3 kJ/mol for *Hs*NMT1 and +15.0 kJ/mol for *Lm*NMT) compared to the more rigid compounds **1** and **2** (-18.0 and -20.9 kJ/mol in *Hs*NMT1 and -4.5 and -13.7 kJ/mol in *Lm*NMT). This was expected due to the rigidification of the flexible linker. However, this did not account for compound **4** binding to *Hs*NMT1 with a higher ΔH of -16.9 kJ/mol compared to compounds **1** and **2** and a lower -TΔS of -21.0 kJ/mol. This indicated that selectivity was driven by the lower gain of enthalpy in comparison to the parasitic enzyme with a ΔH of -44.2 kJ/mol.

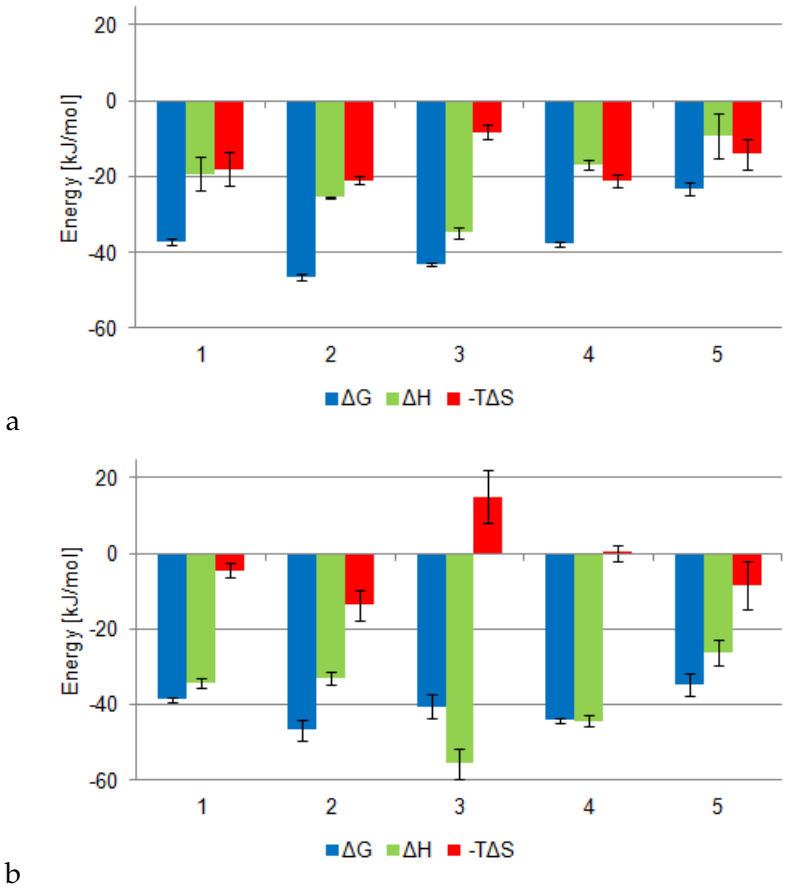


FIGURE 3.3: Thermodynamic profiles of ligand binding to *Hs*NMT1 and *Lm*NMT. a) shows results for *Hs*NMT1, b) for *Lm*NMT. Compound 2 was measured using HEPES buffer, otherwise sodium phosphate buffer was used (pH for both cases was 7.5). Parameters include the Gibbs free energy (ΔG, blue bar), enthalpy change (ΔH, green bar) and temperature dependent entropy change (-TΔS, red bar). Data are mean values of at least triplicate determination with standard error.

Additional experiments using altered buffer conditions with higher ionization ethalpy (HEPES instead of sodium phosphate, Goldberg, Kishore, and Lennen, 2002) indicated a protonation event for compound **3** binding to both NMTs, but not for compounds **1** and **2** (Appendix B). As the thermodynamic differences between the ligands binding were of minor interest

TABLE 3.3: Affinity and thermodynamic data of compounds **1-5** from ITC experiments with *Hs*NMT1. ΔG, ΔH and -TΔS in kJ/mol.

Compound	K_d [nM]	ΔG	ΔH	-TΔS
1	248.3 ± 108.4	-37.2 ± 0.9	-19.2 ± 4.6	-18.0 ± 4.6
2	5.6 ± 2.2	-46.5 ± 0.9	-25.5 ± 0.4	-20.9 ± 1.0
3	20.7 ± 3.2	-43.2 ± 0.4	-34.8 ± 1.4	-8.3 ± 1.8
4	183.8 ± 39.7	-37.8 ± 0.5	-16.9 ± 1.2	-21.0 ± 1.7
5	88 509.2 ± 70 664.4	-23.3 ± 1.9	-9.2 ± 5.8	-14.0 ± 3.9

TABLE 3.4: Affinity and thermodynamic data for compounds **1-5** from ITC experiments with *Lm*NMT. ΔG, ΔH and -TΔS in kJ/mol.

Compound	K_d [nM] (SI)	ΔG	ΔH	-TΔS
1	132.2 ± 40.0 (1.9)	-38.7 ± 0.7	-34.2 ± 1.2	-4.5 ± 1.9
2	6.8 ± 5.5 (0.8)	-46.8 ± 2.6	-33.1 ± 1.7	-13.7 ± 4.1
3	88.5 ± 67.6 (0.2)	-40.6 ± 3.2	-55.6 ± 4.0	15.0 ± 7.1
4	21.4 ± 4.0 (8.8)	-44.2 ± 0.7	-44.3 ± 1.5	0.1 ± 2.1
5	779.8 ± 719.6 (113.5)	-34.8 ± 3.0	-26.3 ± 3.4	-8.5 ± 6.4

compared to the differences of a ligand binding to different NMTs, this protonation enthalpy elucidation was not followed up. However, it should be considered when comparing thermodynamics of different ligands binding to one enzyme and be part of future elucidation.

3.3 Molecular Dynamic Simulations

To obtain insights into the dynamics of both NMTs molecular dynamic simulations (MDs) in the presence and absence of the inhibitors were performed. In cases where no crystallographic information was available, docking studies to obtain starting structures were performed. The obtained MD trajectories were analyzed to elucidate protein flexibility in a qualitative and quantitative manner and water network formation and changes upon different ligands binding to the enzyme (see also Chapters 4 and 5).

3.3.1 Docking derived Binding Modes and Scores

As starting structures for MDs binding poses of compounds **2-4** in *Lm*NMT and **2-5** in *Hs*NMT1 were generated with LeadIT-2.1.6. Redocking of compound **1** into 2WSA for *Lm*NMT and 3IWE, chain A for *Hs*NMT1 resulted in binding modes with a RMSD of 0.64 Å and 0.66 Å, respectively. Docking poses for compound **2-5** were rescored using HYDE and the pose for each

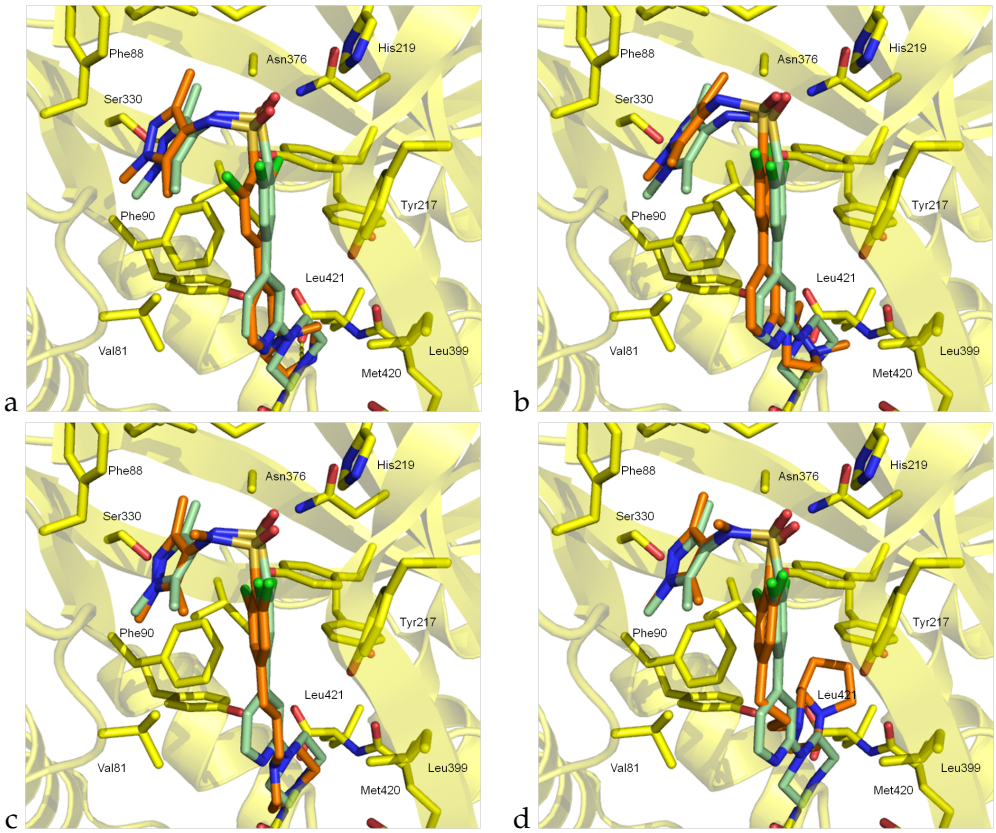


FIGURE 3.4: Compounds **1-4** binding to *Lm*NMT. Protein represented with yellow carbon atoms, reference ligand (compound **1**) with pale green carbon atoms, predicted binding mode with orange carbon atoms. PDB-ID = 2WSA. a) redocking of compound **1**, b) predicted pose for compound **2**, c) compound **3** and d) compound **4**.

ligand with the highest HYDE score and forming all key interactions (Chapter 1.2.4) was chosen as MD starting structure. All selected binding poses are shown in Figures 3.4 and 3.5. The corresponding scores are listed in Table 3.5. Divergent for compound **5**, *Hs*NMT1 structure 3IU1, chain B was chosen instead of 3IWE, as the open conformation of Tyr296 was mandatory for this ligand's binding mode. The RMSD compared to compound **5** binding to *Lm*NMT (PDB-ID 4CGN) was 1.52 Å.

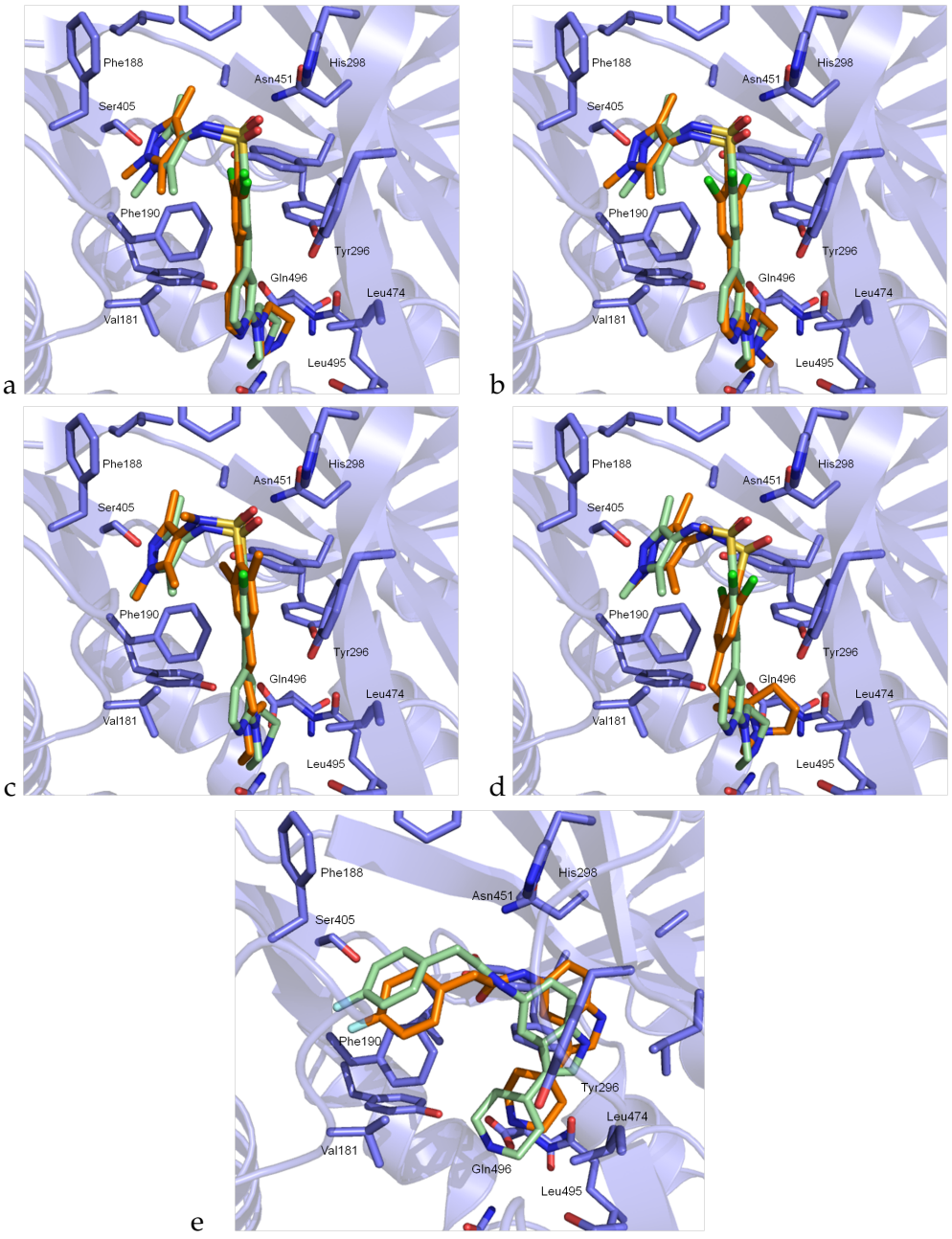


FIGURE 3.5: Compounds **1-5** binding to *Hs*NMT1. Protein represented with blue carbon atoms, reference ligand (compound **1**) with pale green carbon atoms, docking solutions with orange carbon atoms. PDB-ID = 3IWE, chain A. a) redocking of compound **1**, b) solution for compound **2**, c) compound **3** and d) compound **4**. e) binding of compound **5** using 3IU1, chain B as template. Reference ligand with pale green carbon atoms illustrates corresponding binding mode of compound **5** in *Lm*NMT, PDB-ID 4CGN.

When comparing the calculated scores with ITC-derived experimental results the HYDE scores overestimate binding of compound **5** to both enzymes (HYDE of -38 to -45 and -48 kJ/mol against -23.1 and -34.8 kJ/mol in ITC) and underestimate binding of compound **4**, not indicating any selectivity (HYDE score of -29 kJ/mol for both enzymes). The non-selective compounds **2** and **3**, however had different scores for human and parasitic

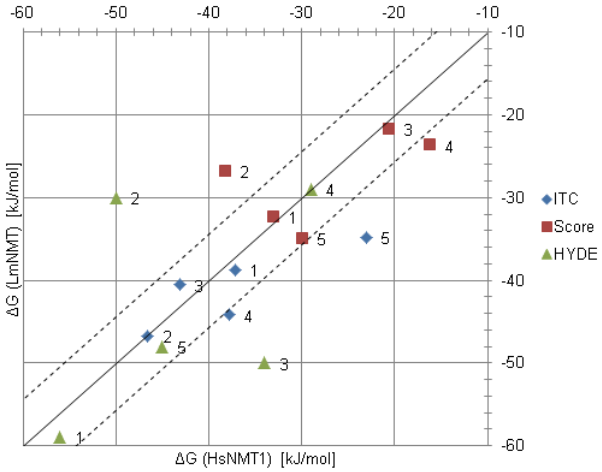


FIGURE 3.6: Comparison of selectivity for compounds **1-5** by ΔG from ITC with predicted selectivity from LeadIT-scores and HYDE scores. Diagonal straight line indicates no selectivity, dotted lines 10-fold selectivity cut-off. Blue data points from ITC, red from scores and green from HYDE.

NMT. Even though HYDE proved to be useful to discriminate realistic from less-favourable binding modes, it was only roughly able to predict binding affinity or selectivity between the two NMTs (Table 3.5, Figure 3.6). Therfore, a higher level of theory incorporating protein flexibility, water network formation and their changes upon ligand binding was needed.

TABLE 3.5: Docking derived binding scores for comparison with experimental data corresponding to binding poses shown in Figures 3.4 and 3.5. All values are in kJ/mol. [a] Crystallographic binding mode used for HYDE assessment. [b] binding mode transferred from *Lm*NMT used for HYDE assessment.

Protein	Ligand	ΔG_{ITC}	Score	HYDE
*Hs*NMT1	1	-37.2	-33.0	-56[a]
*Hs*NMT1	2	-46.6	-38.2	-50
*Hs*NMT1	3	-43.1	-20.6	-34
*Hs*NMT1	4	-37.8	-16.2	-29
*Hs*NMT1	5	-23.1	-29.9	-38 (-45[b])
*Lm*NMT	1	-38.7	-32.4	-59[a]
*Lm*NMT	2	-46.8	-26.9	-30
*Lm*NMT	3	-44.2	-21.7	-50
*Lm*NMT	4	-44.2	-23.7	-29
*Lm*NMT	5	-34.8	-35.0	-48[a]

3.3.2 Quantum Mechanic Thioester Parameterization

MDs were performed with AMBER and (mostly) CHARMM forcefields. For AMBER the small molecule parameters for MyrCoA and inhibitors were derived from GAFF using antechamber. CHARMM parameters were

derived from CGenFF or MMFF94. The CGenFF thioester parameters of the cofactor were derived from ester parameters by closest analogy and were refined on a QM level of theory using S-ethyl thiopropionate (Figure 3.7) as model system. Results are summarized in tables 3.6, 3.7 and 3.8. C(O) denotes sp^2 carbon atom of the thioester.

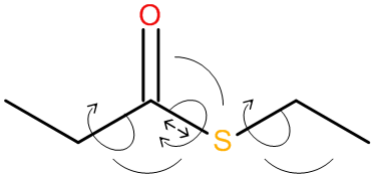


FIGURE 3.7: S-Ethyl thiopropionate as a model system for parameterization of the thioester in cofactor MyrCoA. Parameterized bond, angles and dihedrals are illustrated.

TABLE 3.6: Thioester bond parameter

Bond	K_b [kcal/mol/Å^2]	b_0 [Å]	bond (experimental) [Å]
C(O)-S	195	1.761	1.767

TABLE 3.7: Thioester angle parameters

Angle	K_θ [kcal/mol/deg]	θ_0 [deg]	angle (experimental) [deg]
C-C(O)-S	14.12	114.22	113.4
O=C(O)-S	103.68	125.07	124.1
C(O)-S-C	84.38	96.53	99.4

TABLE 3.8: Thioester dihedral angle parameters. Torsion angles are described by multiple wave functions.

Dihedral angle	K_ϕ [kcal/mol]	n	δ [deg]	dihedral (experimental) [deg]
S-C(O)-C-C	1.591	2	180	
S-C(O)-C-C	0.756	3	0	
S-C(O)-C-H	0.850	2	180	
S-C(O)-C-H	0.714	3	180	
C-C(O)-S-C	0.073	1	0	
C-C(O)-S-C	2.800	2	180	
O=C(O)-S-C	0.348	1	0	-3.5
O=C(O)-S-C	2.863	2	180	-3.5
C-C-S-C(O)	0.572	1	180	
C-C-S-C(O)	0.415	3	0	

The good agreement of energetic minima with experimental reported data (Zacharias et al., 1983) indicated well fitted parameters for implementation in CHARMM.

3.3.3 Protein Dynamics

Starting structures for MDs were equilibrated by increasing temperature from 100 to 300 K and releasing initial constraints on the protein over 1 ns as described in Chapter 2.2.3. Subsequent production runs were carried out for 50 ns. If not mentioned differently the simulations using CHARMM36 force field are described. All simulations showed a high stability with RMSD values lower than 2 Å compared to the starting structures. As illustrated in Figure 3.8 representing the backbone RMSD over time for MDs with CHARMM36 and crystallographic starting structures, the average RMSD for the parasitic enzyme was slightly lower than for the human homologue (highest average deviation for ligand free enzyme of 1.41 Å). This also accounted for the docking derived structures summarized in Table 3.9. All MDs showed good convergence (Koukos and Glykos, 2014) within the 50 ns of the simulation and were further analysed with respect to their conformational and dynamical changes in ligand-bound and -free form.

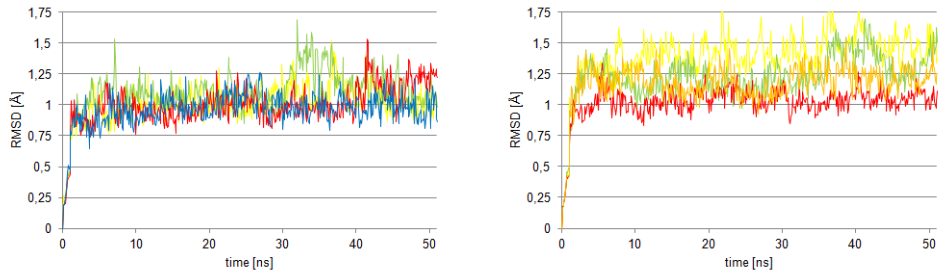


FIGURE 3.8: *Lm*NMT and *Hs*NMT1 backbone RMSD plot over 1 ns equilibration and 50 ns trajectories. a) for *Lm*NMT in complex with MyrCoA in yellow (PDB-ID 3H5Z) and green (4CGP). Bound to compound **1** (2WSA) in red and compound **5** in blue (4CGN). b) *Hs*NMT1 + MyrCoA in yellow (3IU1, chain B) and green (4C2Y chain A), + compound **1** in red (3IWE, chain A) and orange (3IWE, chain B).

TABLE 3.9: Average backbone RMSD (in Å) of 50 ns trajectories of *Lm*NMT and *Hs*NMT1 in complex with MyrCoA and with or without compounds **1-5** bound compared to starting structures. [a]docking derived starting structure.

***Lm*NMT**	**PDB-ID**	**RMSD**	***Hs*NMT1**	**PDB-ID**	**RMSD**
+MyrCoA	3H5Z	1.03	+MyrCoA	3IU1, B	1.41
+MyrCoA	4CGP	1.24	+MyrCoA	4C2Y, A	1.24
+MyrCoA+**1**	2WSA	1.01	+MyrCoA+**1**	3IWE, A	1.04
			+MyrCoA+**1**	3IWE, B	1.19
+MyrCoA+**2**	2WSA[a]	1.00	+MyrCoA+**2**	3IWE, A[a]	1.08
+MyrCoA+**3**	2WSA[a]	1.03	+MyrCoA+**3**	3IWE, A[a]	1.07
+MyrCoA+**4**	2WSA[a]	1.18	+MyrCoA+**4**	3IWE, A[a]	1.29
+MyrCoA+**5**	4CGN	0.97	+MyrCoA+**5**	3IU1, B[a]	1.32

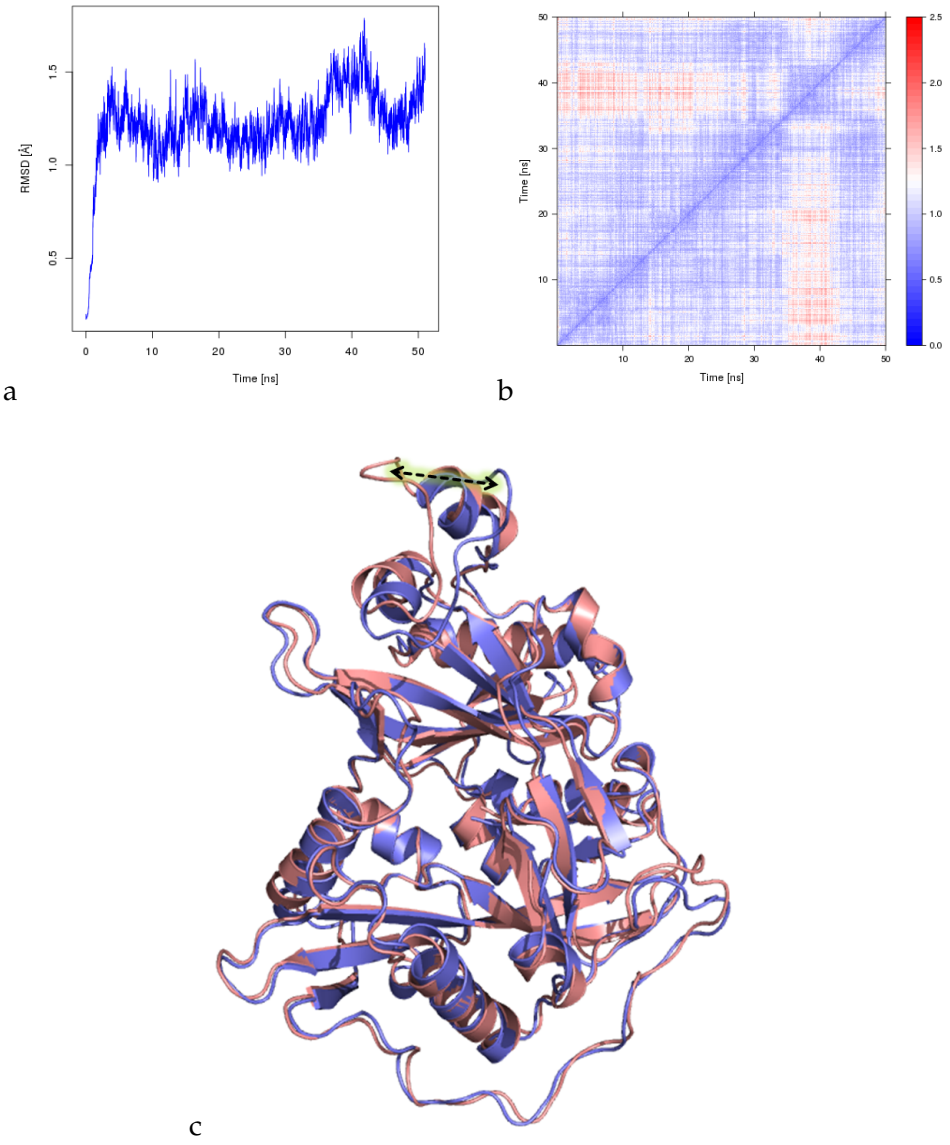
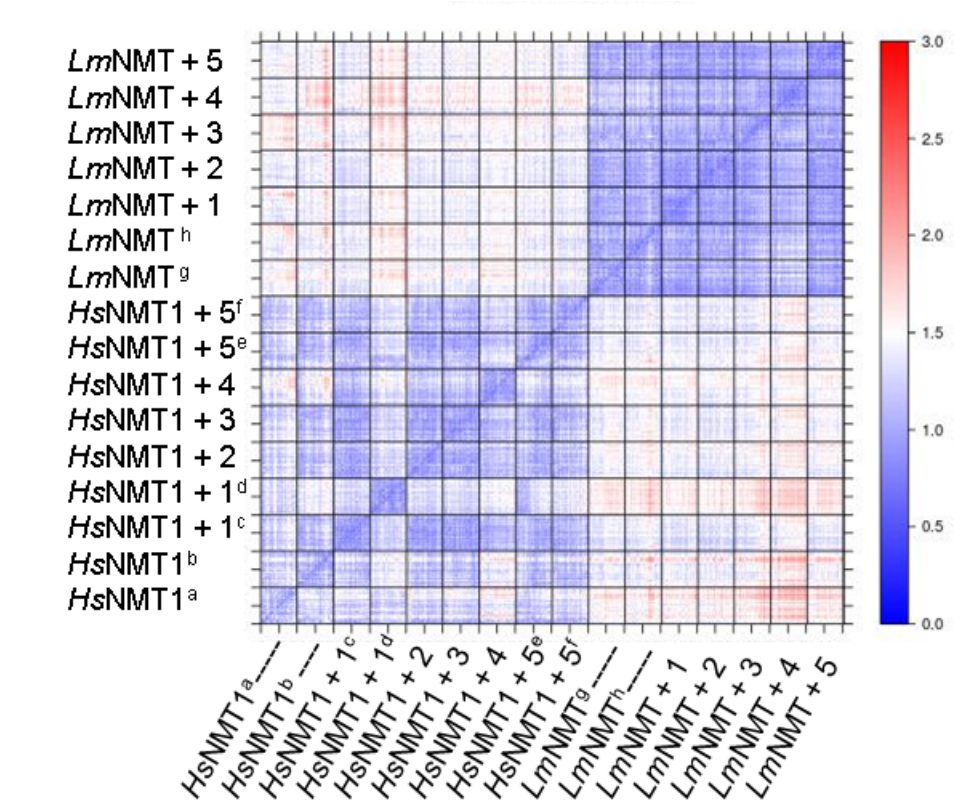
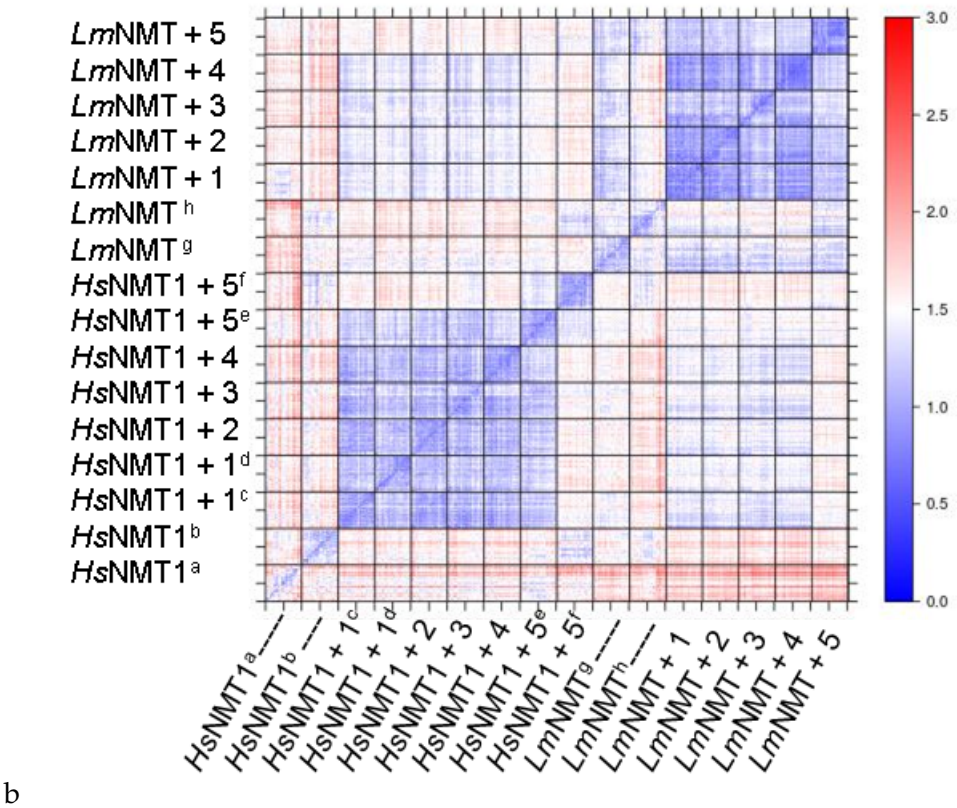


FIGURE 3.9: Exemplary illustration of a 2D RMSD plot of *Hs*NMT1 to reveal conformational changes during 50 ns MD. a) 1D RMSD plot including equilibration. b) corresponding 2D RMSD plot without equilibration. Comparison of all frames of the simulation with each other. Coloring according to the RMSD (in Å) between the compared frames (color bar on the right). c) Superposition of average structure from all frames (conformation 1, blue) with average structure between 37.3 and 41.3 ns (conformation 2, red) reveals upper loop movement.

For further investigation of conformational shifts two-dimensional RMSD plots (2D-RMSD) were calculated for all MDs to elucidate correlation between preferred conformation and the bound ligand(s). By this method one can not only compare the RMSD of a given frame to one defined structure, but also between all given frames to identify conformational changes. This is exemplary illustrated in Figure 3.9 where a slightly altered second conformation (between 37.3 and 41.3 ns) for *Hs*NMT1 was observed.



a



b

FIGURE 3.10: Comparative 2D-RMSD plots for each frame of 16 MDs. Each square represents a complete 50 ns MD simulation trajectory. Similar structure snapshots with low RMSD (in Å, right bar) are colored blue, high structure deviations are colored in red (over white). a) backbone RMSD of whole protein. b) all non-hydrogen atoms of binding site residues. Duplicate MDs starting from PDB-ID [a]3IU1, chain B, [b]4C2Y, chain A, [c]3IWE, chain A, [d]3IWE, chain B, [e]compound **5** docked into 3IU1, chain A, [f]compound **5** in 4C2Y, chain A, [g]3H5Z, [h]4CGP.

The 2D-RMSD comparison of all MDs (Figure 3.10 a) confirmed the overall high stability and preorganisation, especially for *Lm*NMT. It also indicated the influence of the starting structures. Even though the conformational differences for simulations starting from different chains of *Hs*NMT1 in complex with compound **1** (PDB-ID 3IWE, chain A vs chain B of the asymmetric unit, Figure 3.11, RMSD = 0.19 Å) were not high, larger differences throughout the trajectory were found between these chains than caused by different ligands binding to the enzyme. The increased RMSD between the two structures was caused by a flexible loop (close to Ser405, Figure 3.11). However, no correlation between binding of a selective or non-selective inhibitor to a preferred global conformation was identified. Therefore minor differences of binding-site and second-shell residues were elucidated.

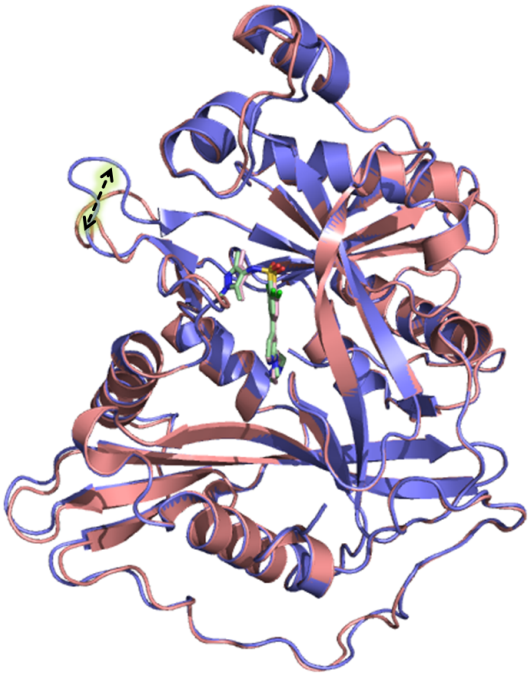
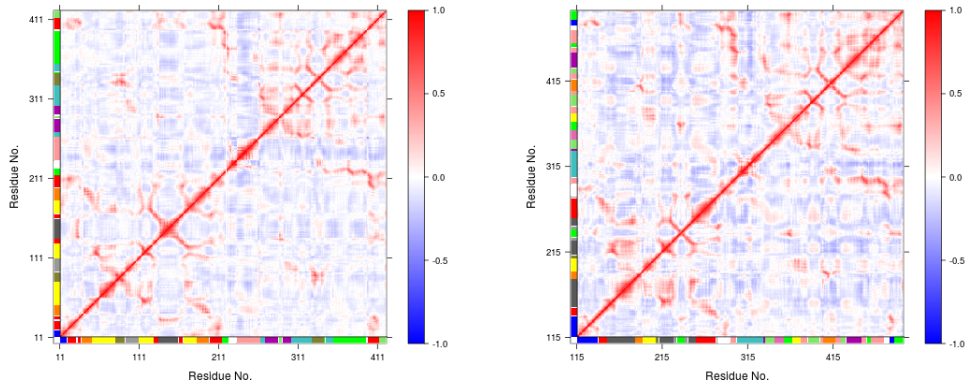


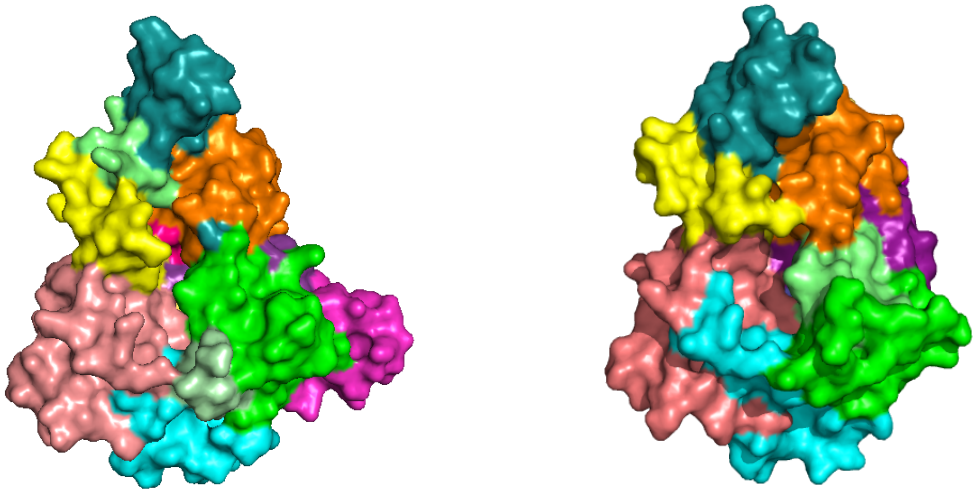
FIGURE 3.11: Superposition of chain A and B of the asymmetric unit of *Hs*NMT1 in complex with compound **1** (PDB-ID 3iwe). Chain A: blue cartoon and pale green ligand, chain B: red cartoon and light pink ligand. Arrow highlights different loop orientation.

For all binding site side chain residues the 2D RMSD plots were calculated (Figure 3.10 b). It is noteworthy that all sulfonamide ligands showed a consistent binding site conformation, even between *Lm*NMT and *Hs*NMT1 (blue areas in the upper left and lower right of the plot). The MDs of ligand-free enzymes generally showed more fluctuations with higher RMSD values, both within a single simulation as well as compared to simulations of ligand bound NMT (red areas). Again, this behaviour was more pronounced in the human enzyme. The MDs of NMT bound to indole ligand **5** showed higher

RMSDs compared to sulfonamide bound NMTs. This was mainly caused by the different orientation of Tyr217 and Tyr296 (Figure 1.13, Chapter 1.2.4), respectively and was stable within the 50 ns simulation, as well. The conformations of unbound NMTs did not correlate with conformations of NMT bound to one particular ligand. As no explicit conformation was dominantly preferred, more subtle changes in protein flexibility upon ligand binding might play a more important role for selectivity.



a b



c d

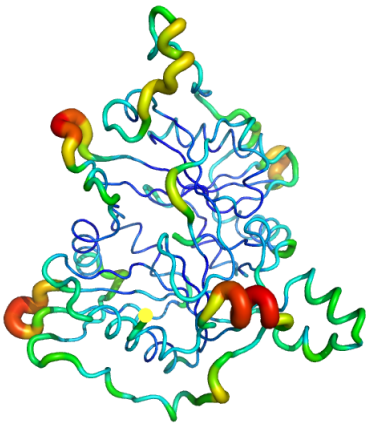
FIGURE 3.12: Dynamic cross-correlation matrices (DCCM) of *Lm*NMT (a) and *Hs*NMT1 (b) with red-colored areas for concerted movements, white areas for independent movement and blue regions for anti-correlated movements. Dedicated residue communities illustrated as three-dimensional network on structure with matched colors for corresponding communities between *Lm*NMT (c) and *Hs*NMT1 (d).

To elucidate the crosstalk between ligand binding and conformational changes in distal regions of the protein, cross-correlation maps were obtained to highlight concerted and independent movements of C_α carbon atoms and to identify networks of residues behaving like one domain. Figure 3.12 represents those networks exemplary for *Lm*NMT and *Hs*NMT1 without inhibitor. The networks generally included neighbouring residues of the protein core or isolated loop regions and often ended at changes in secondary structures. Again no large changes between the two NMTs were found and ligands binding led only to minor alterations in the cross-correlation network (Appendix D).

3.3.4 Regions of altered Flexibility upon Ligand Binding

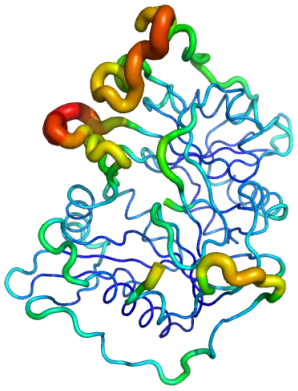
An initial hypothesis for selective NMT inhibition was that ligand binding causes alterations in protein flexibility - and thereby conformational entropy to different extends between *Lm*NMT and *Hs*NMT1 (Chapter 1.3). Analysis of normalized crystallographic B-factors (calculated in an median based approach; Smith et al., 2003) indicated that some regions indeed showed different changes in flexibility upon ligand binding between the two NMTs (Figure 3.13). While the *upper insert* and *α-helix insert* are structural differences between the two NMTs and can not be compared by numbers, major deviations in B-factors can be found in the loop regions. The so-called *acid loop* consisting of one glutamate and three aspartate residues (82-85 and 182-185 respectively), the *pocket loop* following serine 330/405 and a *distal loop* show rigidification to a different extend. However distal loop regions have an intrinsically high flexibility which makes these informations hard to interpret. Interestingly for the *closing lid* (Gly397/472) and the C-terminus with preceding residues, both close to the ligand, the changes in B-factors are in opposite directions, with an increase for *Hs*NMT1 and a decrease for *Lm*NMT upon ligand binding.

Further the root mean square fluctuation (RMSF) was calculated from MDs and compared with the crystallographic B-factors to confirm the correct identification of flexible and rigid regions in the simulations. As represented in Figure 3.14 a, strong qualitative agreement was obtained by comparison of experimental B-factors with MD-derived RMSF. As indicated by crystallographic B-factors the highest flexibility was found in the loop regions like the N-terminus up to residue 57/155 (residue numeration always *Lm*NMT/*Hs*NMT1), the acid loop from Glu82/182 to Asp85/185, the two α-helices of *Lm*NMT around residues 141-159, a quite large mobile region between Phe232/311 of the binding site and residue 246 and 320 respectively, the short loop connecting two β-sheets at residues 315-318/389-393 and the loop close to Ser330/405 of the binding site forming an important H-bond to the pyrazole moiety of the sulfonamide series of ligands (residues



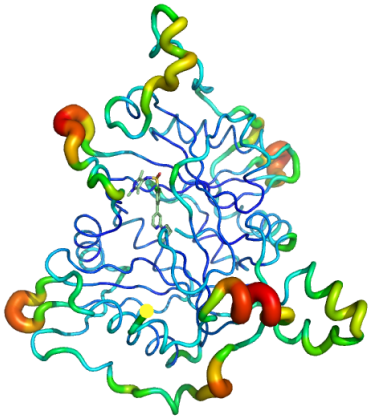
a

ligand-free *Lm*NMT
(PDB-ID 3H5Z)



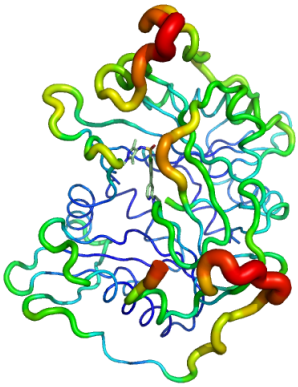
b

ligand-free *Hs*NMT1
(PDB-ID 3IU1)



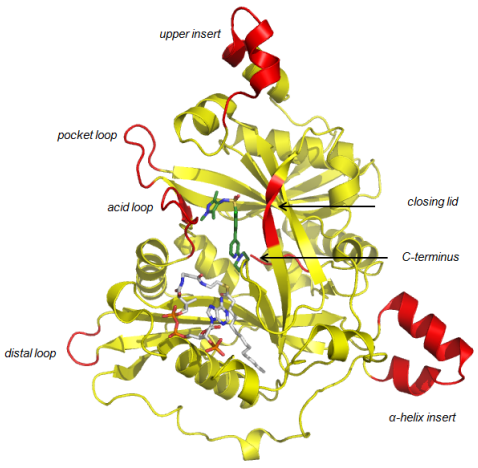
c

compound 1 bound to *Lm*NMT
(PDB-ID 2WSA)

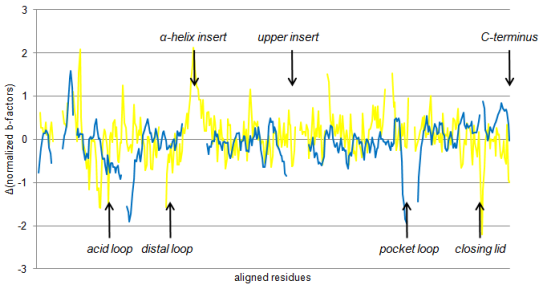


d

compound 1 bound to *Hs*NMT1
(PDB-ID 3IWE)



e



f

FIGURE 3.13: a-d) B-factor tube illustration of ligand-free and compound **1**-bound *Lm*NMT and *Hs*NMT1. Blue thin areas represent regions with low b-factors, thick yellow to red regions of high b-factors. e) Red color highlights regions described in the text with different changes of normalized b-factors upon ligand binding between NMTs, shown on *Lm*NMT structure. f) Difference in B-factors (Δ B-factor, normalized) upon ligand **1** binding. Yellow for *Lm*NMT, blue for *Hs*NMT1. Outliers excluded.

333-339/408-414). A last region of higher mobility in the unbound state was found around residues 395-398 and 470-473 respectively of the "closing lid" which will be discussed later (Chapter 5).

A direct quantitative correlation between experimental and MD-derived B-factors however was difficult because of the divergence in the origin of the numerical values. Experimental B-factors are related to the isotropic displacement of an atom in a "blurry" electron density map (Trueblood et al., 1996), but are highly influenced by the resolution and the quality of the model used for molecular replacement as well. For MDs, it can be directly calculated from the RMSF.

$$B = 8 * \pi^2 * RMSF^2 \tag{3.1}$$

It appeared that MD-derived B-factors around residues of high flexibility were overestimated. However, these residues might also show reduced B-factors in crystal structures due to the stabilisation by packing interactions (Figure 3.15). The higher temperature as well as better hydration during MD could also contribute to different values (Meinhold and Smith, 2005; Li and Brüschweiler, 2009; Janowski et al., 2013). Consequently, the RMSF and its changes upon different ligands binding was exclusively used for identification of regions of interest, but not for quantitative correlation with experimental B-factors or entropy changes in ITC experiments.

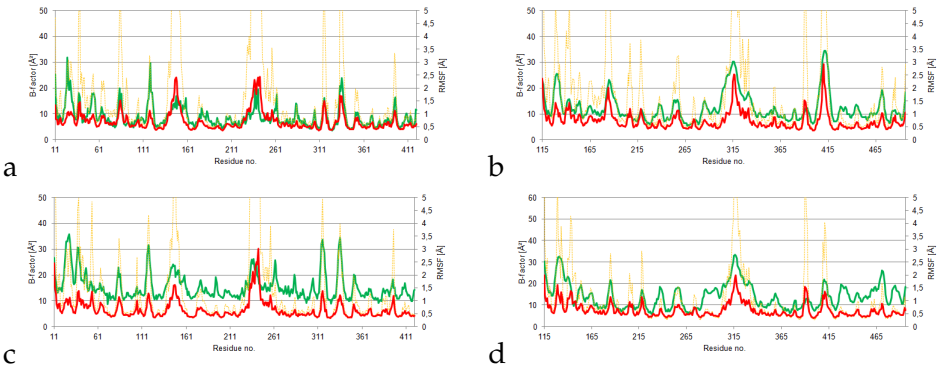


FIGURE 3.14: Comparison between experimental b-factors and MD-derived RMSF and b-factors for C_α cabon atoms. Green line represents experimental b-facors, red line MD-derived RMSF and dashed orange line MD-derived b-factors from RMSF according to equation 3.1. a) *Lm*NMT, PDB-ID 4CGP, b) *Hs*NMT1, 3IU1, c) *Lm*NMT in complex with compound **1**, 2WSA and d) *Hs*NMT1 in complex with compound **1**, 3IWE.

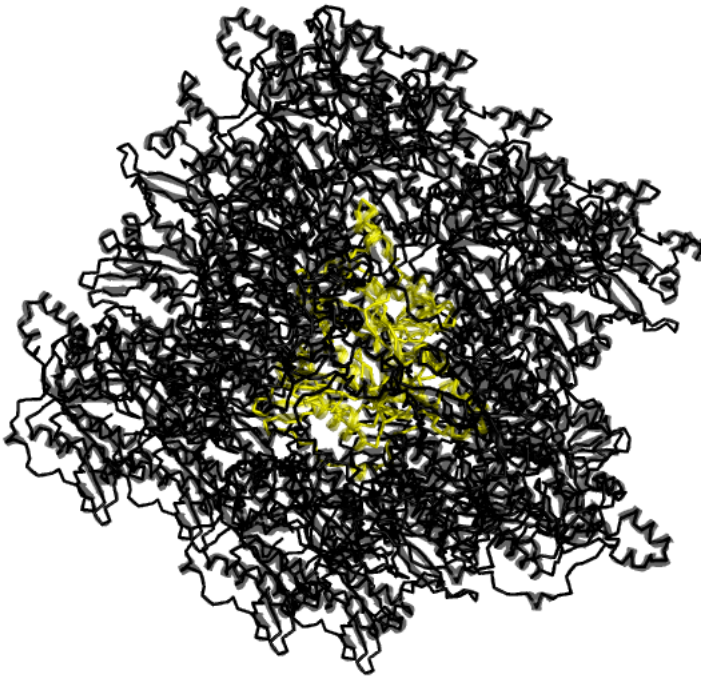


FIGURE 3.15: Crystal packing of *Lm*NMT. Due to packing effect of *Lm*NMT (yellow) with surrounding asymmetric units (black) within the crystal, surface loops might have lower B-factors than calculated from MDs.

Besides the RMSF, another metric of flexibility and MD quality assessment are order parameters (S^2, Chapter 1.4.1; Showalter et al., 2007; Trbovic et al., 2008; Gu, Li, and Brüschweiler, 2014; Genheden, Akke, and Ryde, 2014). An overview of all residues' flexibility throughout all 50 ns MDs is illustrated as heat maps for RMSF, backbone NH and side chain S^2 in Appendix D, Figure D.2.The intrinsically high flexibility of loop regions (like the *pocket loop*) makes minor changes in S^2 hard to interpret. Therefore, during analysis the focus was laid on on NH and side chain S^2 of first and second shell residues around the binding site (defined as 6 Å and 6-12 Å distance from reference ligand **1**). Together with the results from RMSF analysis, three different regions close to the binding site, which showed altered flexibility upon different ligands binding, were identified. Detailed results and their potential impact on selectivity are described in the corresponding Chapters 4.2.1 and 5.1.

3.4 Crystallization of *Hs*NMT1-Inhibitor Complexes

Under the crystallization conditions described in Chapter 2.1.8 protein crystals in presence and absence of inhibitors were grown. *Lm*NMT conditions

delivered only thin platelets (no figure available) and were not followed up as many relevant structures of protein-ligand complexes had been deposited in the PDB previously (PDB IDs 3H5Z, 4CGP, 2WSA, 4CGN, 4UWJ). For *Hs*NMT1 condition A, sea urchin-like needles appeared whereas conditions B and C resulted in three dimensional crystals which subsequently were optimized by variation of the PEG and salt content (Figure 3.16).

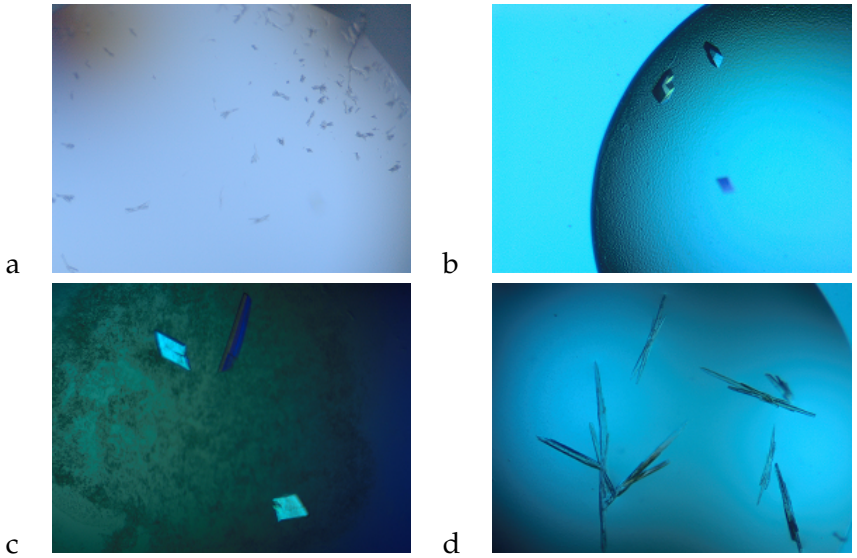


FIGURE 3.16: Representative crystals of *Hs*NMT1 with MyrCoA a) under condition A in presence of compound **1**, b) in presence of compound **5** under condition B, c) condition C in presence of compound **2** and d) condition B with compound **4**.

Crystal structures of *Hs*NMT1 in complex with compounds **2** and **4** were determined and confirmed binding similar to compound **1** with identical interactions of the pyrazole moiety to Ser405 and the tertiary amine to the C-terminal Gln496 (Figure 3.17). The crystal structures of *Hs*NMT1 in complex with compound **2** and **4** also proofed the accuracy of the docking method including the methyl orientation at the tertiary amine of compound **2** (RMSD = 0.81 Å) and (even though there are some deviations in hydrophobic linker orientation) the basic center of compound **4** (RMSD = 1.35 Å, Figure 3.18).

However, no crystals structures in complex with compound **5** (neither for co-crystallization nor soaking with ligand concentrations of up to 1 mM) could be determined. Even though a resolution of 1.8 Å was obtained, no electron density was present within the binding site for the ligand. The data collection and refinement statistics for the determined structures of *Hs*NMT1 in complex with compounds **2**, **4** and further structures (Chapter 5) are summarized in Appendix C.

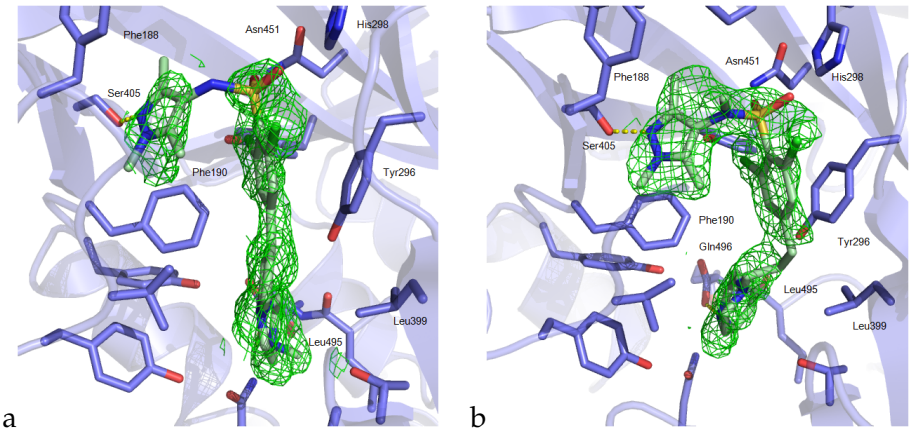


FIGURE 3.17: Crystal structures of *Hs*NMT1 in complex with compounds **2** (a) and **4** (b) with 2Fo-Fc electron density map for ligand contoured at 1 σ.

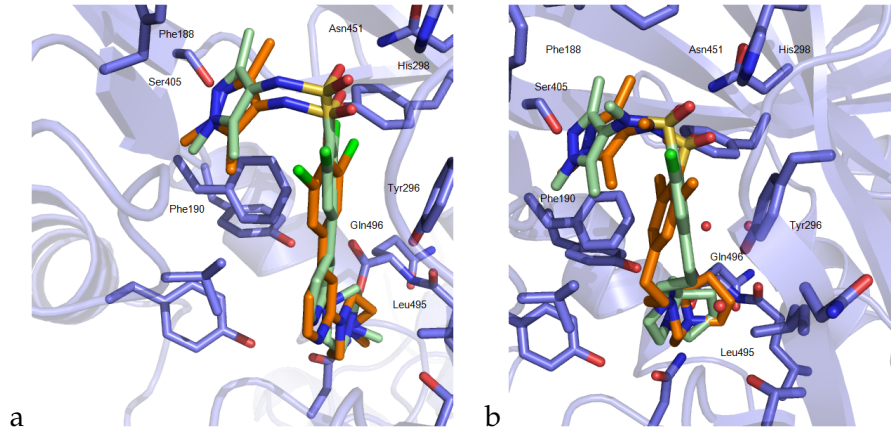


FIGURE 3.18: Comparison of crystallographic (pale green carbon atoms) and docking (orange carbon atoms) derived binding mode of compounds **2** (a) and **4** (b).

3.5 Discussion

The experimental characterization of compounds **1-5** identified sulfonamides **1-3** to be non- or less-selective. Compound **4** which was assumed to have a similar binding mode due to structural similarity however proved to be selective for *Lm*NMT. This binding mode was confirmed by crystallization. Compound **5** which adopts a different binding mode than **1-4**, also showed selectivity for *Lm*NMT over *Hs*NMT1 (Table 3.2).

ITC experiments confirmed these results and further revealed that ligand binding to *Lm*NMT is more enthalpy-driven than it is to *Hs*NMT1 (Table 3.3). This indicated stronger interactions for *Lm*NMT on the one hand and a lower entropic penalty for *Hs*NMT1 on the other hand which might be attributed to a higher preorganisation of the protein.

To elucidate this behaviour and the reasons for selectivity, binding interactions were investigated on the molecular level using computational methods. Docking with LeadIT (FlexX) and the HYDE scoring function

predicted realistic binding modes that were subsequently proved by crystallography for *Hs*NMT1 in complex with compounds **2** and **4** (Figure 3.17). The scores however gave no clear indications for selectivity between the two NMTs (Table 3.5 and Figure 3.6).

The more detailed characterization by MDs revealed that basically no large conformational changes could be observed for both NMTs within 50 ns. Therefore a detailed characterization of single residues' flexibility within and beyond the binding site, as well as the binding site hydration seemed to be necessary.

Chapter 4

Selectivity in Sulfonamide Inhibitor Series

4.1 Mutation Sites

All binding site residues that expose side chains towards the ligand are conserved in *Lm*NMT and *Hs*NMT1. From the other residues within 5 Å of reference ligand **1**, only three are different. These residues' side chains are oriented away from the binding site (Figure 4.1, see also Chapter 1.2.4) and are not directly interacting with the ligand. To answer the question if these three residues nevertheless cause the selective inhibition of some compounds, they were exchanged by site-directed mutagenesis to obtain *Lm*NMT with a "*Hs*NMT1 binding site" and *vice versa*. The three amino acids exchanged were His398, Met420 and Leu421 from *Lm*NMT and Asn473, Leu495 and Gln496 from *Hs*NMT1. Subsequently, kinetic parameters of the mutated NMTs and inhibition controls were determined.

4.2 Results

The results of the enzyme inhibition assay for *Lm*NMT H398N:M420L:L421Q and *Hs*NMT N473H:L495M:Q496L are summarized in Table 4.1 (complete Table with K_m and IC_{50} values in Appendix A). Interestingly, *Hs*NMT1 N473H:L495M:L496Q showed a similar affinity to substrate $pp60^{src}{}_{(2\text{-}9)}$ with K_m = 2.56 μM and an increased inhibition by all sulfonamide compounds compared to the *Hs*TNM1 wild type (wt: refers to unaltered protein binding site, but with hexahistidine tag and TEV protease cleavage site attached to the truncated N-terminus; see also Chapter 2.1.2). Especially for the slightly selective compound **3** and the selective compound **4** the K_i's were reduced from 96.4 to 25.8 nM and 428.2 to 18.4 nM, respectively (Figure 4.2). This indicated that the exchanged residues play a keyrole in selectivity determination for the sulfonamide series of inhibitors as the SI's for **1-4** were reduced from 4, 9, 5 and 213 to 2, 4, 1 and 9. The inhibitor with the indole core (**5**) however, was hardly affected by the mutations (14.8 μM against

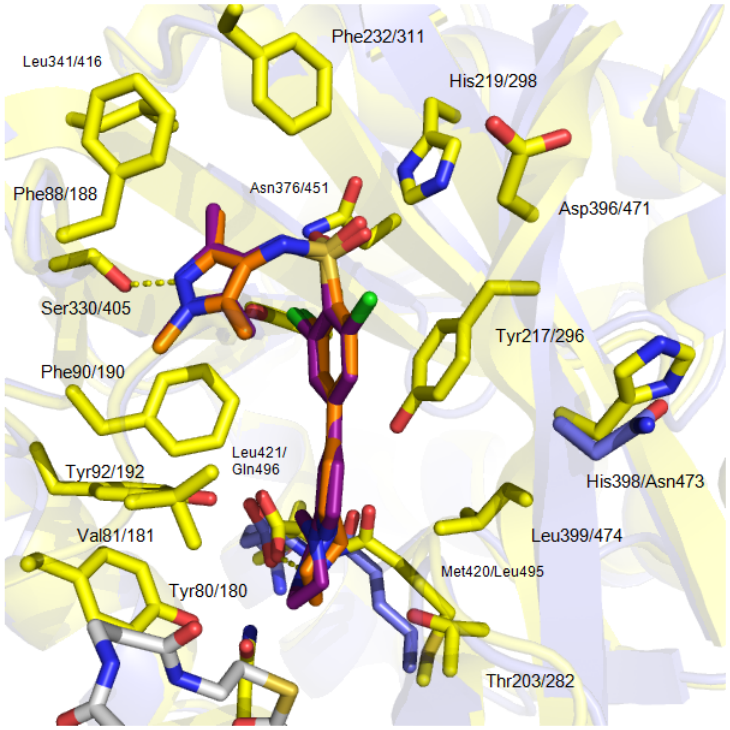


FIGURE 4.1: Illustration of binding site residues and first generation mutation sites. The three residues His398, Met420 and Leu421 from *Lm*NMT were exchanged with Asn472, Leu495 and Gln496 from *Hs*NMT1 and vice versa. Yellow carbon atoms of *Lm*NMT residues, 5 Å around compound **1** (PDB-ID 2WSA, orange ligand carbon atoms) superposed with *Hs*NMT1 (PDB-ID 3IWE, chain A, deep purple ligand carbon atoms). Different residues for *Hs*NMT1 are colored with blue carbon atoms, for identical residues only *Lm*NMT residues are shown. Cofactor Myr-CoA represented with white carbon atoms.

15.4 μM in *Hs*NMT1 altering the SI from 16 to 15). The *Lm*NMT with the *vice versa* exchanged residues H398N:M420L:L421Q however showed a nearly complete loss of activity and the K_m could not be determined.

To reveal which amino acid was causing the increased inhibition and loss of selectivity for compound **4** in *Hs*NMT1 N473H:L495M:Q496L and the loss of catalytic activity in *Lm*NMT H398N:M420L:L421Q - or if it is a combination of multiple residues - NMTs with only one exchanged residue were expressed and purified. For *Hs*NMT1, the focus was laid on the C-terminus, leaving out *Hs*NMT1 N473H, because only a small influence on the inhibition of the corresponding exchange in *Lm*NMT (H398N) was observed.

The results (lower panel of Table 4.1) revealed that L495M was the driving force of increased inhibition in *Hs*NMT1. The K_is for compounds **3** (19.1 nM) and **4** (18.0 nM) were comparable to those obtained for *Hs*NMT1 N473H:L495M:Q496L and close to *Lm*NMT. *Hs*NMT1 Q496L showed only a moderately increased inhibition by those compounds (48.9 nM and 282.4 nM) and even a slightly increased K_i for compound **1** and **2** of 42.9 nM and 40.2 nM compared to *Hs*NMT1. Further *Lm*NMT M420L was identified to

be the amino acid exchange causing the loss of activity. *Lm*NMT H398N and L421Q both were active and similarly inhibited by compounds **1-5** as the *Lm*NMT wt. One deviation was the slightly increased K_i of 36.6 nM for compound **4** in *Lm*NMT L421Q compared to 2.0 nM in *Lm*NMT wt. This slightly increased K_i for *Lm*NMT L421Q matched with the lowered K_i for *Hs*NMT1 Q496L. The main contribution to the relevant changes could be contributed to the C-terminus-1 mutations: M420L in *Lm*NMT for the loss of activity and L495M in *Hs*NMT1 for the stronger inhibition by compound **4**.

TABLE 4.1: Inhibition (K_i in nM) of *Lm*NMT and *Hs*NMT1 and their first and second generation mutants by compounds **1-5**. n.d. = not determined due to inactivity of the enzyme.

NMT	**1**	**2**	**3**	**4**	**5**
***Lm*NMT**	8.4 ± 1.3	1.4 ± 0.3	19.8 ± 2.5	2.0 ± 0.4	975.8 ± 102.1
***Hs*NMT1**	31.6 ± 4.5	13.3 ± 2.3	96.4 ± 11.7	428.2 ± 90.8	15 388.1 ± 4 270.0
***Lm*NMT H398N: M420L: L421Q**	n.d.	n.d.	n.d.	n.d.	n.d.
***Hs*NMT1 N473H: L495M: Q496L**	20.1 ± 4.3	6.2 ± 1.9	25.8 ± 6.5	18.4 ± 3.6	14 800.0 ± 4642.9
***Lm*NMT H398N**	13.9 ± 3.2	8.5 ± 3.6	1.6 ± 0.8	1.3 ± 0.3	517.5 ± 201.5
***Lm*NMT M420L**	n.d.	n.d.	n.d.	n.d.	n.d.
***Lm*NMT L421Q**	41.3 ± 8.6	6.0 ± 3.2	2.3 ± 0.8	36.6 ± 5.0	2 638.0 ± 1 032.9
***Hs*NMT1 L495M**	4.0 ± 2.8	0.9 ± 0.5	19.1 ± 3.0	18.0 ± 5.1	(> 15 μM)
***Hs*NMT1 Q496L**	42.9 ± 16.4	40.2 ± 9.2	48.9 ±20.1	282.4 ± 61.4	(> 15 μM)

ITC experiments were performed with *Hs*NMT1 N473H:L495M:Q496L, as well as with *Lm*NMT H398N:M420L:L421Q, to reveal if the residue exchange prevented proper folding or solely catalytic functionality in *Lm*NMT (Figure 4.3 and Table 4.2, for comparison with results for wt protein see Chapter 3.2). In addition, *Hs*NMT1 M495L was tested against compound

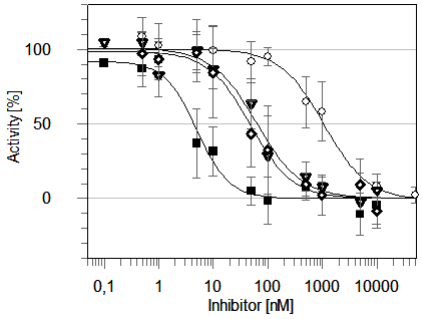


FIGURE 4.2: IC_{50} determination of compound **4**. White circles show data points for *Hs*NMT1, black filled squares for *Lm*NMT, white triangles with bold edges for *Hs*NMT1 N473H:L495M:Q496L and white rhombi whith bold edges for *Hs*NMT1 L495M.

4 due to the altered selectivity, and compound **1** as non-selective reference ligand.

First of all it became evident that *Lm*NMT H398N:M420L:L421Q was still able to bind ligands comparable to other NMTs and thus had to be properly folded. Additionally the thermodynamic split-off was altered as it showed a generally higher entropic contribution for all compounds binding, like it was observed in *Hs*NMT1 wt. Further a slight drop of affinity for all compounds could be observed, but was more pronounced for selective compounds **5** (K_d increased from 0.8 μM to 111.7 μM while it was 70.7 μM in *Hs*NMT1) and **4** where the K_d was increased from 21.4 nM to 46.5 nM while it was 183.8 nM in *Hs*NMT1.

The vice versa mutant *Hs*NMT1 N473H:L495M:Q496L revealed an increased affinity for compound **4** of 13.6 nM compared to *Hs*NMT1 wt, as observed in the inhibition assay. This affinity gain could be mainly attributed to the L495M mutation. However, it was not accompanied with an more negative enthalpy, that was dominant in compound **4** binding to *Lm*NMT (ΔH of -44.3 kJ/mol). Rather the affinity gain was due to an increased entropy between *Hs*NMT1 L473H:L495M:Q496L of -TΔS = -26.3 kJ/mol and *Hs*NMT1 of -21.0 kJ/mol with an basically equal ΔH (-16.9 to -18.1 kJ/mol).

Compound **3**, which showed some low selectivity in the assay (SI of 5) and ITC (SI of 6 in HEPES) was influenced by the selected mutation sites as well. The K_d was increased from 8.7 nM to 79.8 nM in *Lm*NMT by introduction of H398N:M420L:L421Q and decreased from 48.4 nM to 22.8 nM by introducing N473H:L495M:Q496L into *Hs*NMT1. Again, the moderate increase in affinity for the mutated *Hs*NMT1 was driven by an entropic gain (Δ-TΔS = 5.0 kJ/mol) rather than a gain in enthalpy.

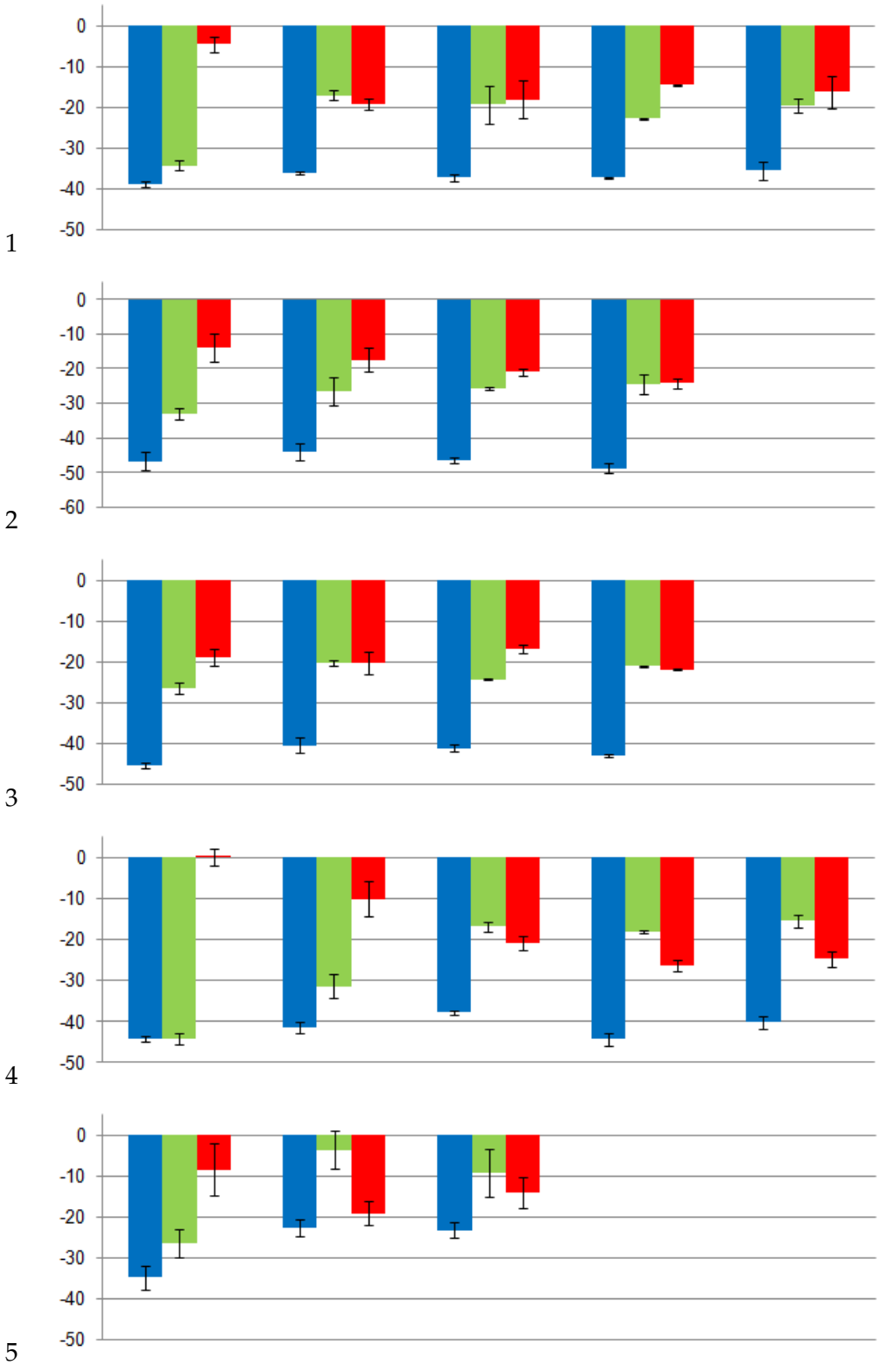


FIGURE 4.3: Thermodynamic characterization of ligand-binding to *Hs*NMT1 and *Lm*NMT and discussed mutants by ITC. Blue bar: Gibb's free energy (ΔG), green bar: enthalpy change (ΔH), red bar: temperature dependent entropy change (-TΔS) in kJ/mol. Cpd. = compound, y-axis unit is kJ/mol. Compounds **2** + **3** data from HEPES buffer, otherwise phosphate buffer was used.

TABLE 4.2: Thermodynamic characterization of *Lm*NMT H398N:M420L:L421Q, *Hs*NMT1 N473H:L495M:Q496L and *Hs*NMT1 L495M. Results in parentheses indicate results from HEPES buffer experiments, otherwise phosphate buffer was used. ΔG, ΔH and -TΔS are in kJ/mol, K_d is in nM. [a]displacement titration experiment.

Protein	**Ligand**	**ΔG**	**ΔH**	**-TΔS**	**K_d**
*Lm*NMT 398N: M420L:L421Q	1	-36.1 ± 0.4	-17.0 ± 1.1	-19.1 ± 1.4	377.2 ± 68.6
*Lm*NMT 398N: M420L:L421Q	2	(-44.0 ± 2.4)	(-26.4 ± 4.1)	(-17.5 ± 3.4)	(19.8 ± 14.9)
*Lm*NMT 398N: M420L:L421Q	3	(-40.5 ± 2.0)	(-20.3 ± 0.7)	(-20.2 ± 2.7)	(76.8 ± 63.7)
*Lm*NMT 398N: M420L:L421Q	4	-41.5 ± 1.5	-31.4 ± 2.8	-10.1 ± 4.3	46.5 ± 30.0
*Lm*NMT 398N: M420L:L421Q	5[a]	-22.7 ± 1.9	-3.6 ± 4.7	-19.1 ± 2.9	111 749.9 ± 86 114.9
*Hs*NMT1 N473H: L495M:Q496L	1	-37.2 ± 0.2	-22.8 ± 0.1	-14.4 ± 0.3	234.2 ± 14.2
*Hs*NMT1 N473H: L495M:Q496L	2	(-48.9 ± 1.4)	(-24.6 ± 2.8)	(-24.2 ± 1.5)	(2.2 ± 1.4)
*Hs*NMT1 N473H: L495M:Q496L	3	-42.9 ± 0.2	-24.8 ± 0.5	-18.0 ± 0.7	23.6 ± 1.5
*Hs*NMT1 N473H: L495M:Q496L	3	(-42.9 ± 0.3)	(-21.0 ± 0.2)	(-21.9 ± 0.2)	(22.8 ± 2.6)
*Hs*NMT1 N473H: L495M:Q496L	4	-44.5 ± 1.6	-18.1 ± 0.4	-26.3 ± 1.3	13.6 ± 7.0
*Hs*NMT1 L495M	1	-35.5 ± 2.3	-19.4 ± 1.7	-16.2 ± 4.0	605.2 ± 458.8
*Hs*NMT1 L495M	4	-40.3 ± 1.6	-15.5 ± 1.5	-24.8 ± 2.0	78.1 ± 56.5

4.2.1 Flexibility of C-terminal Residues

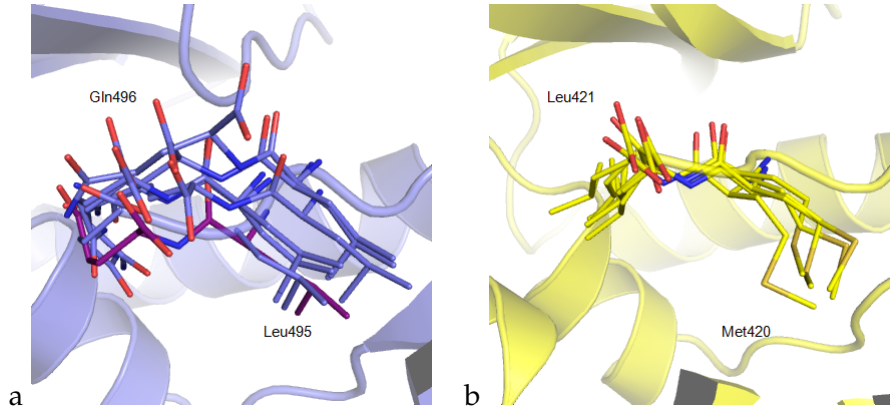


FIGURE 4.4: Illustration of the C-terminal residues' flexibility during 50 ns MDs a) for *Hs*NMT1 presenting Leu495 and Gln496 snapshots each 10 ns without ligand (blue carbon atoms) and average orientation while bound to compound **4** (deep purple carbon atoms) and b) for *Lm*NMT showing Met420 and Leu421 every 10 ns in unbound state.

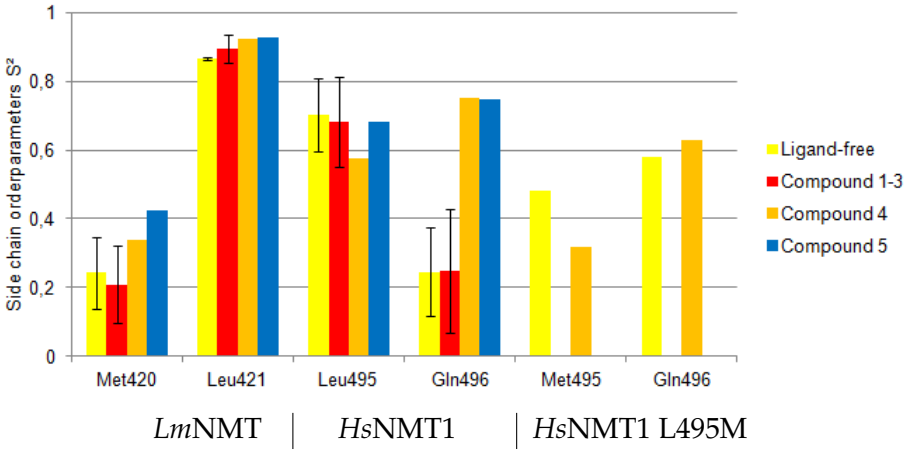


FIGURE 4.5: Side chain order parameters for C-terminal and preceding residues Met420 and Leu421 for *Lm*NMT, Leu495 and Gln496 for *Hs*NMT1. MDs of ligand-free (yellow) and non-selective ligand-bound (**1-3**, red) were pooled for each wt enzyme. MDs of NMT in complex with compound **5** (blue) and **4** (orange) as single run for all enzymes. Each bar represents average side-chain orderparameter for at least one 50 ns simulation with standard error.

To elucidate the molecular reason for the influence of the *Hs*NMT1 L495M mutation on the selectivity of compound **4** binding, the dynamic behaviour of the C-terminus was analysed. In MDs, the C-terminal residues, which contain the catalytically active center, appeared to be influenced in their flexibility upon different ligands binding (Appendix D, Figure D.2; Figure 4.4). This was also observed for crystallographic B-factors (Chapter 3.3.4, Figure 3.13). By ionic interactions with the secondary or tertiary amines of inhibitors **1-5**, one of the strongest and most crucial interactions needed for substrate

TABLE 4.3: Side chain order parameters for C-terminal and preceding residues for *Lm*NMT (Met420, Leu421), *Hs*NMT1 (Leu495, Gln496) and *Hs*NMT1 L495M (Met495, Gln496) for unbound, non-selective sulfonamide ligands **1-3** bound and selective ligands **4** and **5** bound.

NMT	**Ligand**	**C-terminus-1 S^2**	**C-terminus S^2**
		Met420	Leu421
*Lm*NMT	none	0.24 ± 0.10	0.86 ± 0.01
*Lm*NMT	non-selective (1-3)	0.21 ± 0.11	0.89 ± 0.04
*Lm*NMT	4	0.34	0.92
*Lm*NMT	5	0.43	0.93
		Leu495	Gln496
*Hs*NMT1	none	0.70 ± 0.11	0.25 ± 0.13
*Hs*NMT1	non-selective (1-3)	0.68 ± 0.13	0.25 ± 0.18
*Hs*NMT1	4	0.57	0.75
*Hs*NMT1	5	0.68	0.75
		Met495	Gln496
*Hs*NMT1 L495M	none	0.48	0.58
*Hs*NMT1 L495M	4	0.32	0.63

and inhibitor binding is formed there (Chapter 1.2.4). Figure 4.5 and Table 4.3 illustrate the S^2 of the side chain bond vectors of the C-terminus and the preceding residue (C-terminus-1), namely Met420, Leu421 in *Lm*NMT and Leu495, Gln496 in *Hs*NMT1. The most pronounced difference upon ligand binding was the rigidification of Gln496 in *Hs*NMT1 when bound to compound **4** or **5**, with S^2 of 0.75 and 0.69, respectively compared to 0.25 for unbound or non-selective ligand-bound. Leu421 side-chain of *Lm*NMT was already less flexible in unbound state (S^2 of 0.87) and only slightly rigidified upon binding. However, the S^2 of Met420 was slightly increased from 0.24 to 0.34 and 0.42 for compounds **4** and **5** compared to the unbound state. This different behaviour was further investigated, even though the side chains of the residues are not directly involved in the protein-ligand interactions.

To elucidate, if the affinity gain and thereby loss of selectivity of compound **4** by introduction of the L495M substitution in *Hs*NMT1 (SI from 214 to 9 in the assay) was correlated with the increased C-termial S^2, MDs of HsNMT1 L495M in presence and absence of compound **4** were performed. While for the wt *Hs*NMT1 the Gln496 side chain S^2 was increased from 0.25 to 0.75 upon compound **4** binding, it was only slightly increased, from 0.58 to 0.63 in *Hs*NMT1 L495M, but already more rigid in unbound state. The C-terminus-1 residue (Met495) S^2, was actually reduced from 0.48 to 0.32 upon binding (Table 4.3, lower panel).

4.3 Discussion

By the exchange of the three different binding site residues between *Lm*NMT and *Hs*NMT1 within 5 Å of reference ligand **1**, NMTs with the binding site residues of the other enzyme were obtained. Testing in inhibition assay and ITC revealed a loss of enzymatic activity in *Lm*NMT H398N:M420L:L421Q even though it was still properly folded and able to bind compounds **1-5**. However, affinity was decreased, especially for selective compounds **4** and **5** and the distribution between enthalpy and entropy approached towards the thermodynamic split-off found in *Hs*NMT1. Likewise *Hs*NMT1 N473H:L495M:Q496L was affected in the contrary direction, showing stronger inhibition by and higher affinity to compound **4**, but was not affecting compound **5**'s inhibitory effect. Further the increased affinity to compound **4** could be attributed to a gain of entropy (Tables 4.1, 4.2).

By subsequently reducing NMTs with three exchanged residues to only one mutated residue, the C-terminus-1 residue exchange (M420L and L495M, respectively) was identified to be crucial for both the loss of activity in *Lm*NMT H398N:M420L:L421Q and the loss of selectivity of compound **4** against *Hs*NMT1 N473H:L495M:Q496L.

The characterization by MDs with and without different ligands bound revealed, that basically no large conformational changes could be observed for both NMTs within 50 ns simulations (Chapter 3.6). Elucidation of backbone and side chain S^2 however, highlighted residues within and close to the binding site, whose flexibility was altered upon ligand binding. The flexibility of the C-terminus of *Hs*NMT1 was differently impaired depending on which ligand was bound. For non-selective ligands *Hs*NMT1 Gln496 S^2 was barely influenced, while it was increased upon compounds **4** and **5** binding. This indicated that Gln496 and surrounding residues might play a role in selectivity by rigidification (Table 4.3, Figure 4.5).

The entropic gain for *Hs*NMT1 L495M binding compound **4** compared to *Hs*NMT1 wt, does not negate the hypothesis of the rigidification of Gln496 upon ligand binding to be relevant for selectivity. Even though *Lm*NMT ligand binding was more enthalpy driven, the gained entropy may be partially contributed to a lower entropic penalty for Gln496 rigidification (indicated by S^2) in *Hs*NMT1 L495M compared to the *Hs*NMT1 wt.

Chapter 5

Selectivity of Indole Inhibitor Compound 5

While the *vice versa* exchange of binding site residues between *Lm*NMT and *Hs*NMT1 helped to explain the selectivity for compound **4**, the selectivity determining features for compound **5** remained unsolved. Even though the affinity of compound **5** was reduced in *Lm*NMT H398N:M420L:L421Q compared to the *Lm*NMT wt, the *vice versa* exchange in *Hs*NMT1 did not result in a stronger inhibition by this compound. Therefore, further MD analysis was performed.

5.1 Tyrosine 217/296 and the "closing lid"

Investigating NH backbone S^2, regions of non-uniform fluctuations in flexibility were identified. Figure 5.1 presents the evaluation of NH bond vector S^2 of ligand-free simulations and MDs of compound **5** bound to NMTs. A major difference was observed for Gly397 and Gly472 respectively for which an increased NH S^2 from 0.30 to 0.59 upon ligand binding was found in *Lm*NMT whereas it was reduced from 0.75 to 0.56 in the human enzyme. This residue was of high interest, as it is part of the "closing lid" of the binding site.

Accompanied with the change of NH S^2 of Gly397/472 was also the open and closed conformation of the binding site residue Tyr217/Tyr296. While it was quite stable in terms of its backbone NH S^2, the side chain S^2 showed a large increase upon ligand binding (Table 5.1). This residue was present in either the closed conformation (NMTs bound to compounds **1-4**) or the open conformation (compound **5**).

These findings for Gly397, Tyr217 in *Lm*NMT and Gly472, Tyr296 in *Hs*NMT1 were followed by a more detailed characterization of the affected residues and their local environment. The open conformation of the tyrosine offers space to a hydrophobic pocket which is addressed by the indole inhibitor **5** (Figure 1.13, Chapter 1.2.4). The occurrence of open and closed conformation was analysed to identify the preferred state for both NMTs.

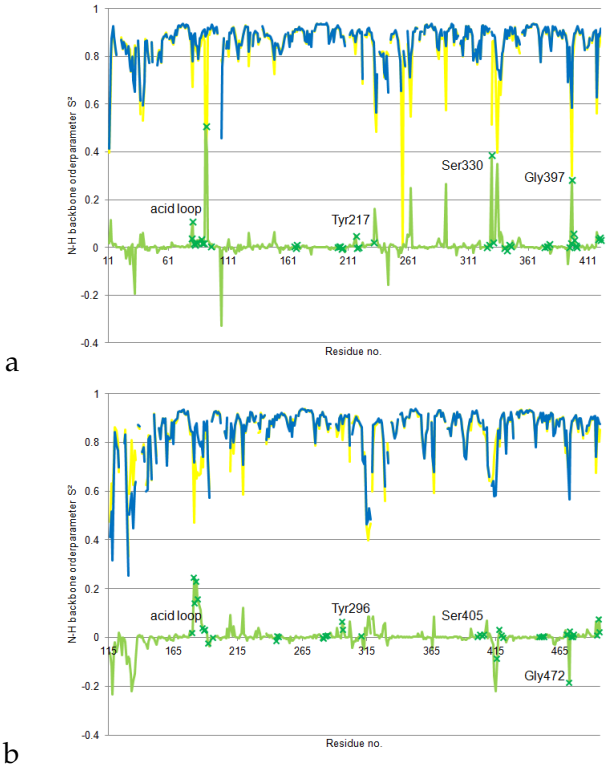


FIGURE 5.1: NH S^2 comparison of compound **5** binding to a) *Lm*NMT and b) *Hs*NMT1. Yellow line: S^2 derived from ligand-free MD, blue line: S^2 derived from MD with compound **5** bound to NMT, green line: difference in S^2, green crosses: highlighting binding pocket residues.

Different conformational preferences between *Lm*NMT and *Hs*NMT1 would have helped to explain the selectivity of different state binding inhibitors, like the open conformation binding compound **5**. For that purpose the distance between oxygen atoms of the phenolic oxygens of Tyr 217/296 and Tyr326/401 was measured (Figure 5.2). This distance described the conformation of Tyr217/296 very well, as Tyr326/401 was very stable and the O-O distance differed between open and closed conformation by around 5 Å. This made the open and closed conformation easy to discriminate. The torsion angle of Gly397 / Gly472 and the preceding Asp396/471 residue ((Asp)N-C-C_α-(Gly)N) of the closing lid located in the anti-parallel β-sheet next to the tyrosine was investigated as well, because it defines the closing lid by a flip around this angle.

In the entirely open conformation Gly397/472 has a dihedral angle of around 150 ° while this angle is approximately -33 ° in the closed conformation. Thus a cutoff at 58.5 ° was chosen to discriminate between the two conformations. For the distance between Tyr217/296 and Tyr326/401 the cutoff was 9.4 Å for *Lm*NMT and 8.3 Å for *Hs*NMT1 as it was around 7.5 Å in the closed and 11.3 Å in the open conformation in the parasitic enzyme and 6.7 to 10.0 Å in the human NMT. This already indicated a slightly reduced

TABLE 5.1: Tyrosine 217 / 296 side chain S^2 from MDs with ligand-bound and unbound NMT. NH S^2 in parenthesis.

***Lm*NMT + ligand**	**S^2**	***Hs*NMT1 + ligand**	**S^2**
no ligand	0.68 (0.83)	no ligand	0.32 (0.76)
1	0.82 (0.89)	**1**	0.88 (0.91)
2	0.83 (0.87)	**2**	0.83 (0.89)
3	0.89 (0.90)	**3**	0.74 (0.82)
4	0.90 (0.91)	**4**	0.87 (0.91)
5	0.89 (0.86)	**5**	0.79 (0.87)

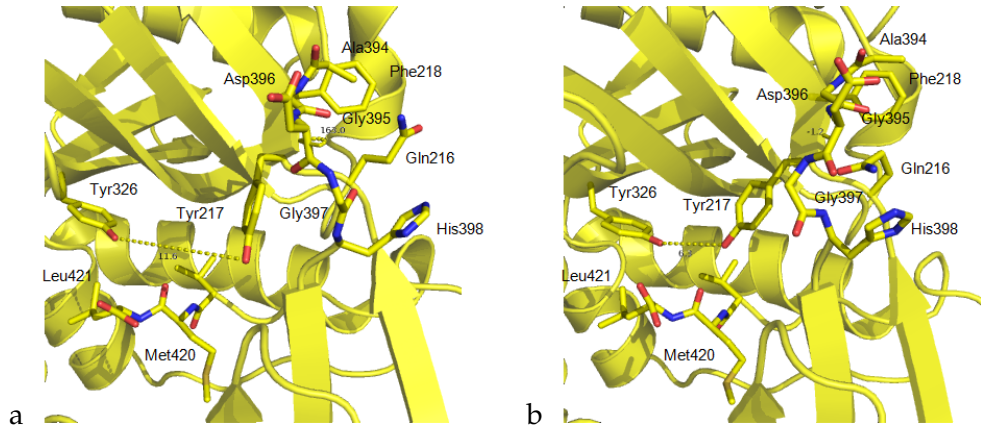


FIGURE 5.2: MD snapshots of a) open conformation of Tyr217 with a O-O distance to Tyr326 of 11.6 Å and a Gly397 dihedral angle of 163.0°. b) Closed conformation with distance of 6.3 Å and dihedral of -1.2 ° in *Lm*NMT.

space in the hydrophobic pocket for *Hs*NMT1. However, the given space was sufficient for compound **5** (as stated in Chapter 3.3.1) and other compounds (Chapter 1.2.4, Figure 1.15) to bind to *Hs*NMT1. Indeed, the given space in *Hs*NMT1 is slightly larger than in *Lm*NMT due to two different residues (Met377, Val378 in *Lm*NMT and Ala452, Leu453 in *Hs*NMT1, see Chapter 5.2). Furthermore, in MDs the dominant state of unbound *Hs*NMT1 was the open conformation for both Gly472 and Tyr296 (0.8 - 1.2 % closed for Gly472 and 6.8 - 27.7 % for Tyr296 derived from two independent MDs, Table 5.2, Figures 5.3 b and d). The trapping of Gly397 and Gly472, respectively in one or the other conformation by ligands, was also less distinct compared to observations from crystal structures (for ligand-free NMTs Gly397/472 was exclusively found in the open conformation, PDB-IDs 3H5Z, 4CGP for *Lm*NMT and 3IU1, 4C2Y for *Hs*NMT1). During MDs switching between Gly-open and -closed state, even in presence of a ligand, especially for compound **5** (closed Gly conformation in 15.2 % for *Lm*NMT and 29.3 % for *Hs*NMT1, Table 5.2, Figures 5.2 a and b), was observed. To elucidate if these two residues' flexibility was generally observable or specific for CHARMM, MDs of unbound NMTs using the AMBER force field ff99SB were performed for comparison. Those MDs showed a similar behaviour for Tyr217/296

and Gly397/472. Even though the percentage of closed conformation was increased with this setup (up to 69 % for Gly397 and 80 % for Tyr217 in *Lm*NMT and 36.7 % for Gly472 and 36.5 % for Tyr296 in *Hs*NMT1), the open conformation remained favoured for the human enzyme. In addition, a concerted movement of both residues was observed (Figures 5.3 e and f). Contrary, *Lm*NMT preferred the closed conformation (69.3 and 80.2 %).

Compound **5** binds to the Tyr217/296-open conformation of NMT (Chapter 1.2.4, Figure 1.13). Therefore, an initial hypothesis for its selectivity was, that this open conformation was preferred in *Lm*NMT, while the closed conformation was supposed to be privileged in *Hs*NMT1 (Brannigan et al., 2014). However, MDs showed that the open conformation was preferred in both enzymes. Together with recently found non-selective open conformation binding ligands (Rackham et al., 2015), these findings strongly indicate that preference of the open conformation itself can not be the selectivity determining feature for compound **5**.

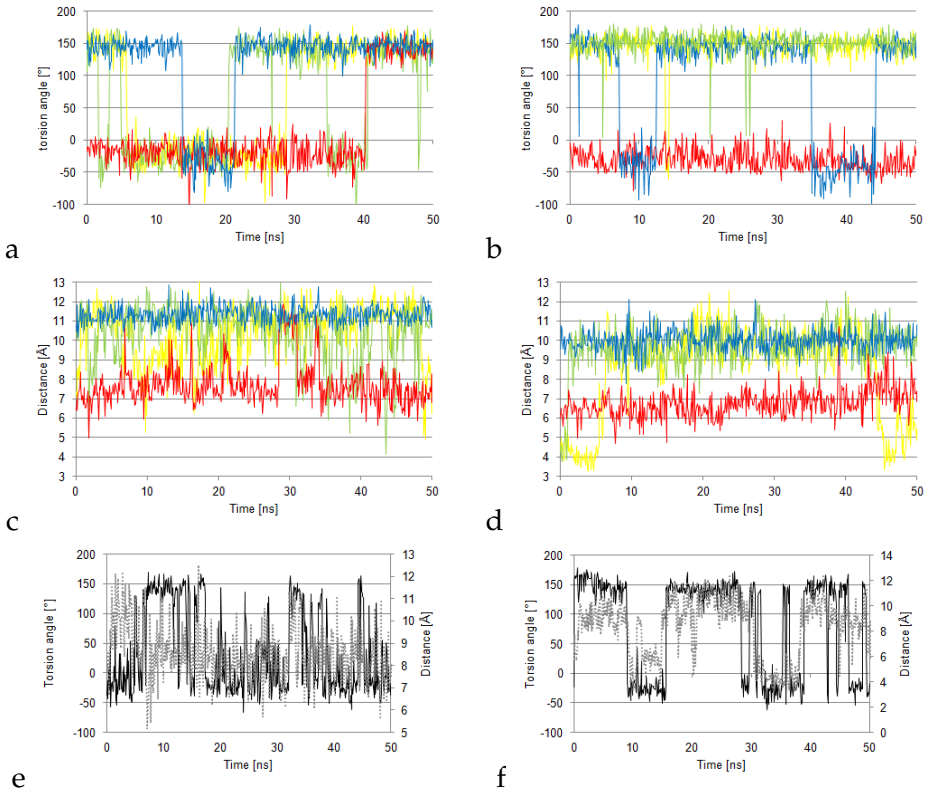


FIGURE 5.3: Gly397/472 and Tyr217/296 open and closed conformation distribution based on N-C-C_α-N torsion angle and phenolic O-O distance of tyrosines. Yellow and green represent inhibitor-free MDs, red compound **1** bound representative for all sulfonamide inhibitors, blue compound **5** bound. a) Gly397 torsion angle for *Lm*NMT, b) Gly472 torsion angle of *Hs*NMT1, c) Tyr271-Tyr326 distance *Lm*NMT and d) Tyr296-Tyr401 distance *Hs*NMT1. e) and f) represents comparative measurements of MDs using ff99SB for ligand-free *Lm*NMT (e) and *Hs*NMT1 (f). Black line for Gly torsion angle, dotted grey line for Tyr-Tyr distance.

TABLE 5.2: Tyr271/296 and Gly397/472 open and closed conformation distribution of unbound and inhibitor bound NMTs. [a]starting from open Gly472 conformation, PDB ID 3IWE, chain B

NMT	Gly397/472 [% closed]	Tyr217/296 [% closed]
*Lm*NMT	45.9	24.2
*Lm*NMT	47.1	18.8
*Lm*NMT (ff99SB)	69.3	80.2
*Lm*NMT + 1	80.4	92.4
*Lm*NMT + 5	15.2	0.0
*Hs*NMT1	1.2	27.7
*Hs*NMT1	0.8	6.8
*Hs*NMT1 (ff99SB)	36.7	36.5
*Hs*NMT1 + 1	100.0 (95.0[a])	97.2 (88.8[a])
*Hs*NMT1 + 5	29.3	0.2

5.2 Binding Site Water Network

To elucidate the influence of the displacement of explicit water molecules upon ligand binding, water density maps were calculated from MDs for all binding pockets of NMT. The obtained density peaks were compared to electron density maps of high resolution crystal structures to confirm correct water placement during MDs (Figure 5.4). The MD-derived density maps for explicit water molecules matched the crystallographic electron density very well and therefore proofed to be a useful tool to analyse water molecule positions and alterations for different ligands.

Using this approach, it was revealed that the hydration of the binding sites at the entrance to the pocket opened by Tyr217 and Tyr296, respectively, differed. While this pocket was occupied by several water molecules, two residues that differ between the NMTs (Met377, V378 in *Lm*NMT and Ala452, Val453 in *Hs*NMT1, Figure 5.5), define the depth of the pocket (Rackham et al., 2015) and are likely to be involved in the water network formation within that pocket. One explicit water molecule was present in unbound *Hs*NMT1 and not in unbound *Lm*NMT, which will be called "W1" in the following part and can be referred to W629 in PDB-ID 3IU1, chain B.

The W1 hydration site was also found in all simulations of ligated *Hs*NMT1 and in all but one simulation with ligand-bound *Lm*NMT. The sulfonamide inhibitors **1-4** trap the Tyr217/296 in the closed conformation and thereby W1, which is consistent with observations from crystal structures (Figure 5.6). At this coordinates, W1 has several potential binding partners to form H-bonds (Tyr217/296, Tyr326/401, backbone of Met420/Leu495 and further water molecules).

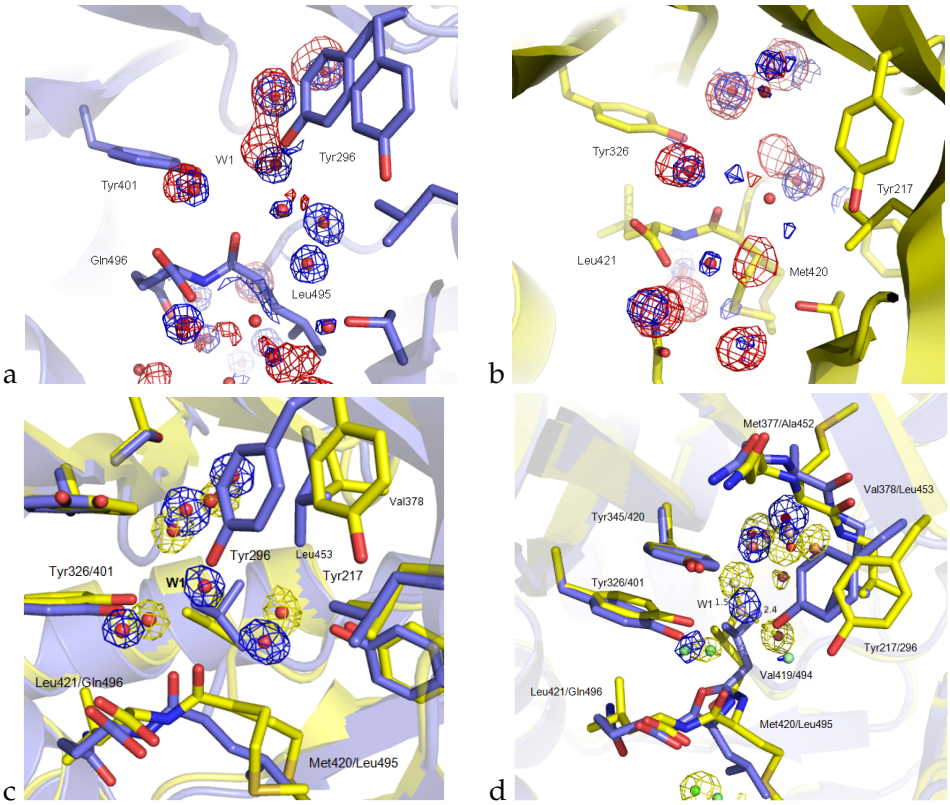


FIGURE 5.4: Comparison of water molecule electron density from crystal structures (blue mesh at 2 σ) with occupancy from MDs (red mesh, 70 % occupancy). a) *Hs*NMT1, PDB-ID 3IU1, b) *Lm*NMT, PDB-ID 4CGP. Water molecules (red spheres) from crystal structures. c) and d) Comparison of water network between *Hs*NMT1 and *Lm*NMT. c) Electron density 2Fo-Fc map of water molecules for *Lm*NMT, PDB-ID 4CGP with yellow carbon atoms and mesh and *Hs*NMT1, PDB-ID 3IU1, chain A, blue carbon atoms and mesh at 2 σ. d) MD-derived density maps for water oxygen at 0.07 cut-off, same protein and mesh coloring as in (c). Water molecules are colored according to their "SPAM" free energy ΔG_{SPAM}, see below. Closest water molecule distance in *Lm*NMT to position of W1 is 1.5 Å.

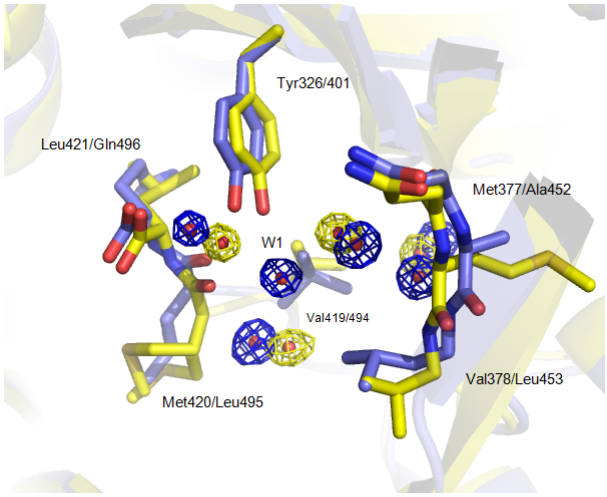


FIGURE 5.5: Residues of *Lm*NMT (yellow carbon atoms and 2Fo-FC electron density (ED) map at 2 σ) and *Hs*NMT1 (blue carbon atomas and ED map) close to binding site involved in altered water network.

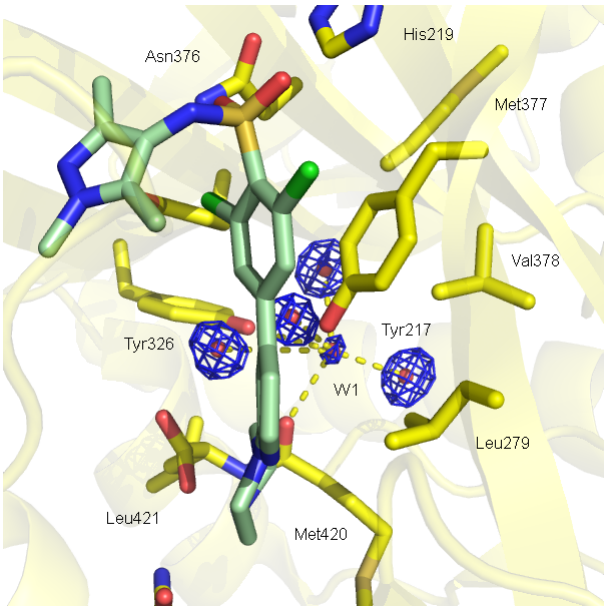


FIGURE 5.6: Trapped W1 in *Lm*NMT bound to compound 1, PDB-ID 2WSA. 2Fo-Fc electron density map as blue mesh for selected water molecules at 1.5 σ.

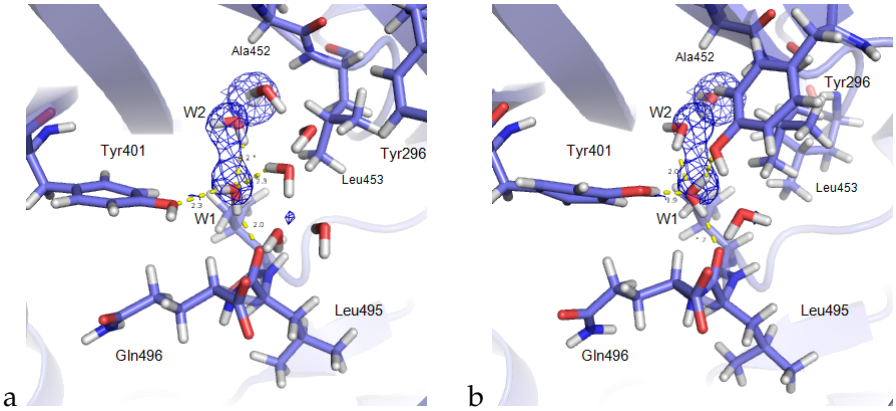


FIGURE 5.7: Influence of open (a) and closed (b) Tyr296 conformation on W1 orientation in MDs of *Hs*NMT1. Mesh at an occupancy level of 70 % for water oxygen.

Further, W1 was present independent of Tyr296's conformation and reoriented depending on the conformation of Tyr296, but it was never displaced. While W1 acted as an H-bond donor for Tyr296 and the backbone oxygen of Leu495 and as an acceptor for Tyr401 and another water molecule (W2) in the closed conformation, it was a H-bond donor to Tyr401 and acceptor for a an additional water molecule in the open conformation (Figure 5.7).

The selective inhibitor with the indole core (**5**) however binds to this pocket and thereby occupies space close to the position of W1 (Figure 5.8 a). A potential resulting displacement of the highly conserved W1 in *Hs*NMT1 might be less favourable due to its strong interactions with the protein

and surrounding water molecules. To elucidate the stability of this water molecule in *Hs*NMT1 in complex with compound **5**, it was further analysed.

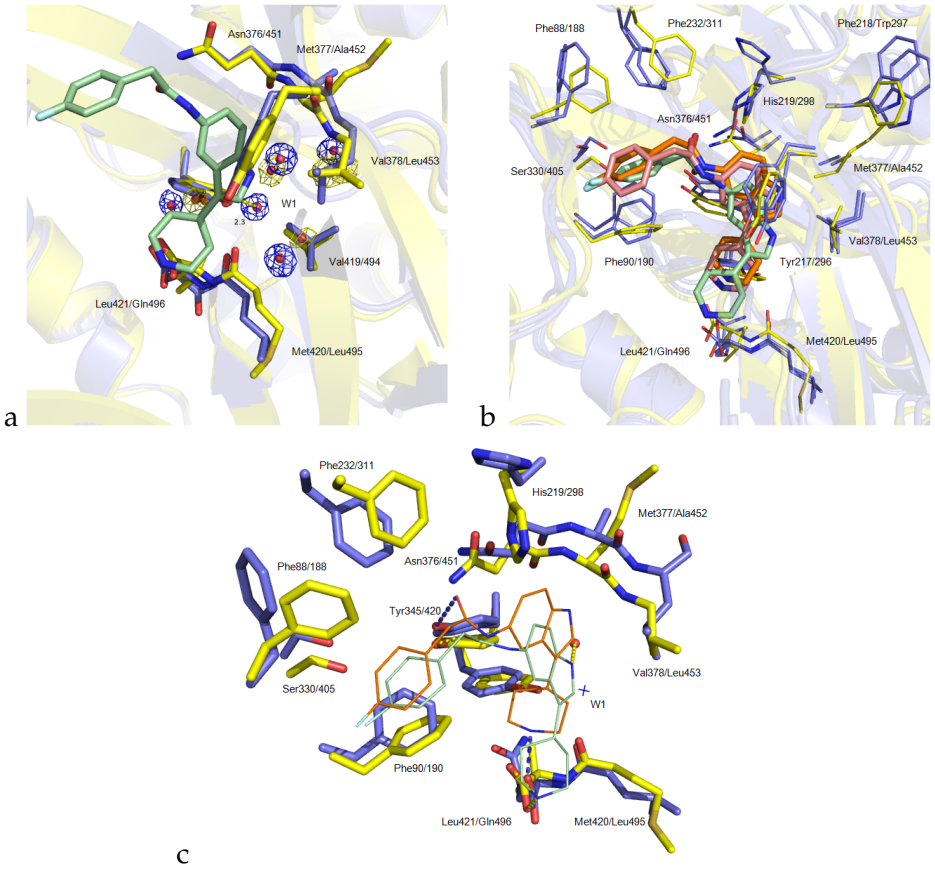


FIGURE 5.8: a) Superposition of *Lm*NMT with *Hs*NMT1. Compound **5** (pale green carbon atoms) in *Lm*NMT (yellow carbon atoms and mesh, PDB-ID 4CGN) is close to W1 in apo structure of *Hs*NMT1 (blue carbon atoms and mesh, PDB-ID 3IU1, chain B), electron density for selected water molecules at contour level of 2 σ. b) Overlay of compound **5** bound MD snapshots closest to average coordinates for *Lm*NMT with yellow carbon atoms and pale green ligand and two MDs of *Hs*NMT1 with blue carbon atoms and orange (starting structure derived from docking) and light pink ligand carbon atoms (starting structure derived by transferring compound **5**'s binding mode from *Lm*NMT). c) Altered interactions of shifted binding mode described in text for compound **5** in *Hs*NMT1 (blue carbon atoms for receptor and orange for ligand) compared to *Lm*NMT (yellow carbon atoms for receptor and pale green for ligand).

Two independent MDs of *Hs*NMT1 and compound **5** (one with a starting structure derived from a docking pose, the other transferring the binding mode from *Lm*NMT and minimizing the surrounding residues of *Hs*NMT1, Figure 5.8) were performed. W1 was not present in none of the starting structures due to its proximity to compound **5**, which would cause a steric clash. Contrary to the expectations, the MD results indicated that W1 was not displaced by the compound at all. Instead, compound **5** adapted its binding mode in *Hs*NMT1 with potentially less favourable interactions compared to *Lm*NMT binding (Figure 5.8 b). The shifted binding mode

altered the following interaction features (Figure 5.8): 1) the orientation of the piperidine moiety forming the salt bridge with the C-terminus, 2) the placement of the H-bond donor nitrogen atom of the indole core inside the hydrophobic pocket with an potential desolvation penalty, 3) a potential clash with His298 forcing this residue in another orientation, 4) an altered H-bond angle between oxygen from ligand's amide linker and Tyr420 and 5) different orientation of the fluorophenyl-moiety with altered π-π-stacking interactions with Phe188, Phe190 and Phe311. Thereby W1 could be assumed to be a selectivity determining feature by shifting compound **5**'s binding mode.

For quantification of W1 binding to NMTs, the WaterBase module of Relibase (Nayal and Di Cera, 1996; Gohlke, Hendlich, and Klebe, 2000; Sanschagrin and Kuhn, 1998; Bergner et al., 2001; Hendlich et al., 2003; Günther et al., 2003) was used as a knowledge based approach to score the water molecules found in the crystal structures. A first approximation for crystallographic W1 were the DrugScores for unbound and ligand-bound *Lm*NMT and *HsNMT1*. The scores (Table 5.3) indicated the strong binding of W1 based on the favourable Total DrugScore. For *Lm*NMT also crystal structures in complex with compounds **6-10** were included (see Chapter 1.2.4). It was observed that (besides PDB-ID 3IU1, chain B due to a not accounted interaction with the protein) W1 showed favourable total DrugScores in the unbound *Hs*NMT1 with a penalty for binding of compound 1 (PDB-IDs 3IWE and 4C2Z). Interestingly, the presence of W1 in *Lm*NMT in complex with open-conformation binders (**6**, **8-10**) resulted in a more negative score, too. Especially the selective inhibitors **9** and **10** (PDB-IDs 4CGL and 4CYO) showed lower values for W1 and a RMSD close to perfect tetrahedral coordination of 0.15 when compound **10** was bound. Such a good coordination was found in *Hs*NMT1, as well. These results gave further hints for the importance of W1.

Differences in MD-derived water occupancies between parasitic and human NMT were analysed subsequently to identify high and low energy water molecules that were displaced upon binding resulting in different gains or losses of Gibbs free enthalpy by the SPAM method (Cui, Swails, and Manas, 2013). To quantify the energy of W1 during MDs, ΔG_{SPAM}, ΔH_{SPAM} and $-T\Delta S_{SPAM}$ relative to bulk water were calculated (Table 5.4). The bulk water thermodynamics were derived from a 10 ns MD of a 40 Å * 40 Å * 40 Å pure waterbox using TIP3P water model and the CHARMM36 force field. The calculated G_{SPAM} for bulk water was -29.60 kcal/mol, H_{SPAM} -17.61 $\pm$ 5.43 kcal/mol and $-TS_{SPAM}$ -11.98 kcal/mol. Apparently W1 binding was accompanied with an entropic penalty. It was partially compensated by favourable interactions to the protein and other water molecules, resulting in a negative ΔH_{SPAM} (Table 5.4). Even though there were fluctuations in

TABLE 5.3: WaterBase descriptors for crystallographic water molecules corresponding to W1. SAS: solvent accessible area. Mobility: calculated from b-factor and occupancy of water molecule (W) and average b-factor and occupancy of surrounding residues (environment, env.) according to equation 5.1. $RMSD_{tet}$: RMSD of polyhedron with respect to ideal tetrahedron coordination. "-": no data available or W1 not present. [a]distance from W1 position 1.2 Å.[b]glycerol present in binding site

PDB-ID	**3H5Z**	**4CGP**	**2WSA**	**4CGN**	**5A27**	**5A28**	**4CGL**	**4CYO**
Protein + Ligand	*Lm*NMT	*Lm*NMT	*Lm*NMT + 1	*Lm*NMT + 5	*Lm*NMT + 6	*Lm*NMT + 8	*Lm*NMT + 9	*Lm*NMT + 10
Water ID	-	-	2541	-	2614	2356[a]	2340	2366
Polar Protein Contacts	-	-	2	-	2	1	1	2
Polar Water Contacts	-	-	3	-	2	2	3	2
Polar Ligand Contacts	-	-	0	-	0	0	0	0
Drug Score - Protein	-	-	-4941.01	-	-4156.44	-2863.94	-4516.61	-6353.54
Drug Score - Ligand	-	-	-232.51	-	-1927.61	+469.12	-2844.21	-2787.73
Drug Score - Water	-	-	-10904.31	-	-5441.14	-6746.64	-8234.85	-5592.94
Drug Score - Total	-	-	-16077.82	-	-11194.20	-9141.46	-16895.66	-14734.21
SAS [%]	-	-	71	-	53	50	51	55
Mobility	-	-	1.19	-	0.86	0.96	0.79	1.04
$RMSD_{tet}$ [Å]	-	-	0.22	-	0.49	0.53	0.52	0.15

PDB-ID	**3IU1, A**	**3IU1, B**	**4C2Y, A**	**4C2Y, B**	**3IWE, A**	**3IWE, B**	**4C2Z, A**	**4C2Z, B**
Protein + Ligand	*Hs*NMT1	*Hs*NMT1	*Hs*NMT1	*Hs*NMT1	*Hs*NMT1 + 1	*Hs*NMT1 + 1	*Hs*NMT1 + 1	*Hs*NMT1 +1
Water ID	628	629	2250	2650	540	618	2243	2149
Polar Protein Contacts	2	1	2	2	1	1	1	2
Polar Water Contacts	2	2	3	2	2	2	2	2
Polar Ligand Contacts	0	0	1[b]	0	0	0	0	0
Drug Score - Protein	-10223.55	-2444.57	-12700.21	-11118.76	-3888.76	-2583.89	-2532.61	-9608.17
Drug Score - Ligand	0	0	-2232.31[b]	-611.92[b]	369.43	0	+369.43	+478.79
Drug Score - Water	-2678.75	-2418.66	-3490.84	-2263.88	-2892.42	-2263.88	-5959.09	-395.38
Drug Score - Total	-12902.30	-4863.23	-18423.36	-13994.55	-6411.74	-4847.77	-8122.27	-9524.75
SAS [%]	55	68	46	47	67	68	68	59
Mobility	0.54	0.49	0.71	0.68	0.79	0.75	0.99	0.90
$RMSD_{tet}$ [Å]	0.15	0.11	0.14	-	-	-	0.11	-

$$Mobility_{\mathrm{w}} = \frac{b-factor_{\mathrm{w}}/b-factor_{\mathrm{env.}}}{occupancy_{\mathrm{w}}/occupancy_{\mathrm{env.}}} \tag{5.1}$$

TABLE 5.4: SPAM energies relative to bulk water calculated for W1 during 50 ns all atom MDs with explicit solvent. All values are kcal/mol. "-" no W1 density peak present at 0.07 cutoff. Values in parentheses represent density peak present only at lower cutoff of 0.05. Duplicate MDs starting from PDB-ID [a]3H5Z, [b]4CGP, [c]3IU1, [d]4C2Y, [e]3IWE, chain A, [f]3IWE, chain B, [g]compound **5** docked into 3IU1, chain B, [h] compound **5** in 4C2Y, chain A.

NMT + Ligand	ΔG_{SPAM}	ΔH_{SPAM}	$-T\Delta S_{SPAM}$
*Lm*NMT[a]	-	-	-
*Lm*NMT[b]	-	-	-
*Lm*NMT + 1	3.59	-3.32	6.91
*Lm*NMT + 2	5.64	0.45	5.19
*Lm*NMT + 3	(4.26)	(-1.76)	(6.02)
*Lm*NMT + 4	5.67	-2.30	7.97
*Lm*NMT + 5	-	-	-
*Hs*NMT1[c]	2.60	-2.69	5.29
*Hs*NMT1[d]	2.98	-3.16	6.14
*Hs*NMT1 + 1[e]	2.93	-3.93	6.76
*Hs*NMT1 + 1[f]	0.55	-4.21	4.76
*Hs*NMT1 + 2	0.05	-5.02	5.07
*Hs*NMT1 + 3	1.72	-2.24	3.96
*Hs*NMT1 + 4	0.66	-3.11	3.78
*Hs*NMT1 + 5[g]	2.42	-4.98	7.40
*Hs*NMT1 + 5[h]	2.39	-4.59	6.98

the numerical values, the resulting binding energy ΔG_{SPAM} was lower in *Hs*NMT1 than in *Lm*NMT for the corresponding protein-ligand complexes.

This also partially explained the differences in thermodynamic enthalpy-entropy fingerprint of ligands binding to *Lm*NMT, with an higher enthalpy gain and entropy loss compared to *Hs*NMT1 (Chapter 3.2). While in *Hs*NMT1 only small W1-related differences for ΔH and $-T\Delta S$ could be observed upon ligand binding, in *Lm*NMT the complete enthalpy/entropy of introduction of W1 into the pocket was observed upon compounds **1-4** of the sulfonamide series binding. The introduction of W1 was predicted to be accompanied with a negative ΔH and a positive $-T\Delta S$ (Table 5.4). This was also observed in ITC experiments of *Lm*NMT compared to *Hs*NMT1 (Figure 3.3 and Tables 3.3, 3.4).

Further, the absence of W1 only in unbound and compound **5**-bound *Lm*NMT might be the selectivity determining feature for this ligand. Alteration of its stability by mutation of *Hs*NMT1, shifting the W1 hydration site and thereby offering more space for compound **5** to result in a *Hs*NMT1 mutant with higher affinity for this ligand, was subsequently investigated.

5.3 Mutation Sites

To explain selectivity of compound **5**, two key differences between *Lm*NMT and *Hs*NMT1 were observed. First, the available space in the subpocket depending on the Tyr217/296 conformation, as described previously (Brannigan et al., 2014). Second, the water molecule W1, which is highly conserved in the human enzyme, as observed by crystal structure analysis and MDs. Therefore, the focus was laid on the residues that might alter the water network in *Hs*NMT1 and the orientation of Tyr296 to yield an *Hs*NMT1 variant with increased affinity to compound **5**.

Up to eight residues in *Hs*NMT1 were exchanged with the corresponding ones in *Lm*NMT (Figure 5.9). First of all the binding site residues Asn473, Leu495 and Gln496 as described in Chapter 4. However, the selectivity of compound **5** was in contrast to compound **4** not influenced by only these three mutations (Chapter 4.2). Additionally, the exchanges A452M and L453V were introduced as these residues define the depth of the hydrophobic pocket (Rackham et al., 2015; Yu et al., 2015, Figure 5.5). For MDs of *Hs*NMT1 A452M:L453V, a displacement for W1 by 1.3 Å with a lowered density (0.05 instead of 0.07, Figure 5.9 e) was observed. This position of the water molecule corresponds to the closest water position in *Lm*NMT.

Further, due to its proximity to A452M:L453V, Leu462 was changed to Val to avoid a clash with the introduction of Met at position 452 (side chain distance of 1 Å). In addition, Trp297 was exchanged to Phe, due to its proximity to the residues already mentioned and being the residue adjacent to Tyr296. Arg295 was mutated to Gln not only because it was the preceding residue prior to Tyr296, but it also formed several H-bonds with backbone oxygen of Gly470, Asp 471 and Gly 472 (Figure 5.9 c and d) which are likely to influence preference of open or closed conformation of the closing lid.

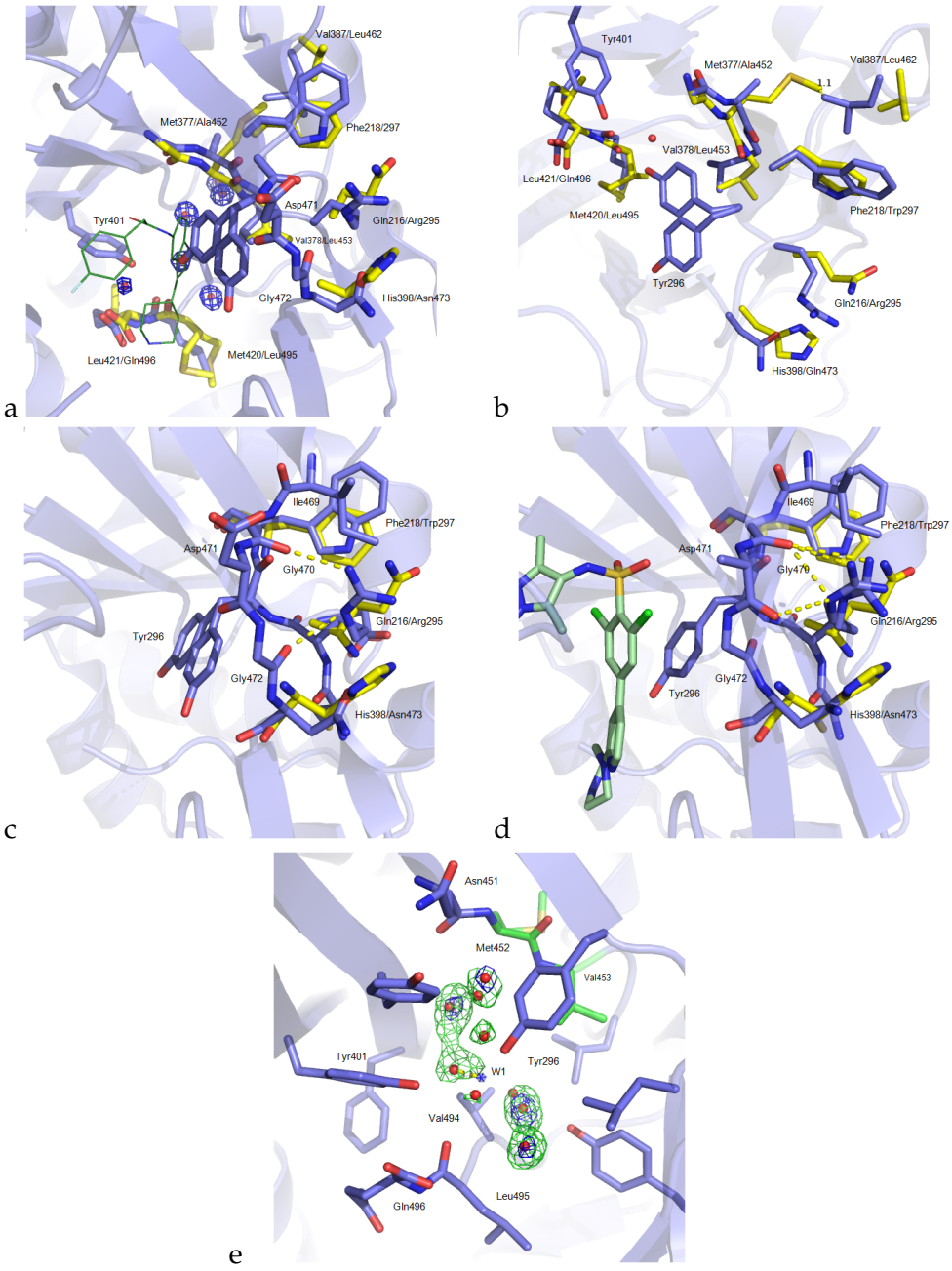


FIGURE 5.9: Illustration of mutation site selection to increase compound **5** affinity in *Hs*NMT1. a) *Hs*NMT1 in overlay with compoud **5** (bound to *Lm*NMT, PDB-ID 4CGN, ligand as lines with dark green carbon atoms) and mutation sites. *Hs*NMT1 blue carbon atoms, corresponding residues in *Lm*NMT with yellow carbon atoms. b) Rotated view to illustrate A452M, L453V and neighboring residues' exchange. c) Illustration of Arg295 interactions with closing lid backbone in open and (d) closed conformation when bound to compound **1** (palegreen carbon atoms). Equivalent residues Gln216, Phe218 and His398 of *Lm*NMT with yellow carbon atoms. e) MD-derived water densities of *Hs*NMT1 A452M:L453V. Green carbon atoms for mutated residues A452M and L453V, blue mesh density map at 0.07 contour level, green mesh at 0.05 contour level. Blue transparent sphere shows *Hs*NMT1 wt W1 position for comparison.

5.4 Results

TABLE 5.5: Inhibition (K_i in nM) of *Lm*NMT and *Hs*NMT1 mutants for compounds **1** and **5**. *Hs*NMT1 *8x*: R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L; *Hs*NMT1 *6x*: W297F:A452M:L453V: L462V:L495M:Q496L

NMT	1	5
***Lm*NMT**	8.4 ± 1.3	975.8 ± 102.1
***Hs*NMT1**	31.6 ± 4.5	15 388.1 ± 4 270.0
Hs*NMT1 *8x	26.4 ± 2.4	972.5 ± 243.7
***Hs*NMT1 A452M**	79.2 ± 31.4	10 899.3 ± 3 330.8
***Hs*NMT1 L453V**	234.7 ± 196.1	> 15μM
***Hs*NMT1 A452M:L453V**	307.0 ± 96.5	> 15 μM
***Hs*NMT1 A452M:L453V:L462V**	124.8 ± 40.5	10 358.0 ± 5 530.1
***Hs*NMT1 A452M:L453V:L495M**	284.6 ± 145.7	12 033.8 ± 6 104.1
***Hs*NMT1 N473H:L495M:Q496L**	20.1 ± 4.3	14 800 ± 4 642.9
Hs*NMT1 *6x	11.3 ± 10.5	8 727.9 ± 2 871.4
***Hs*NMT1 R295Q**	69.1 ± 32.2	> 15 μM
***Hs*NMT1 R295Q: N473H:L495M:Q496L**	38.1 ± 16.9	5635.1 ± 1096.8

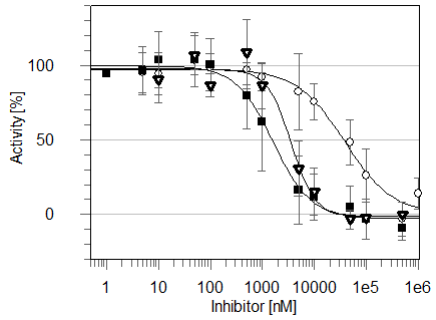


FIGURE 5.10: IC_{50} determination for compound **5**. White circles show data points for *Hs*NMT1, black filled squares for *Lm*NMT, white triangles with bold edges for *Hs*NMT1 *8x*.

The results of the enzyme inhibition assay of NMT with compound **1** as reference ligand and compound **5** as selective ligand are summarized in Table 5.5. The *Hs*NMT1 containing all the mutated residues listed above (R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L) was stronger inhibited by compound **5** with an K_i of 972.5 nM, close to *Lm*NMT wt (975.8 nM), than *Hs*NMT1 wt (15.4 μM, Figure 5.10).

When considering the *Hs*NMT1 mutants with only a subset of mutated residues, one can find only two with a lowered K_i for compound 5: *Hs*NMT1 R295Q:N473H:L495M:Q496L - 5.6 μM and *Hs*NMT1 W297F:A452M: L453V: L462V:L495M:Q496L (*Hs*NMT1 *6x*) - 8.7 μM. For all others the K_i was within the standard error of *Hs*NMT1 wt (15.4 $\pm$ 4.3 μM). Additionally the inhibitory effect of compound **1** was reduced in *Hs*NMT1 containing A452M (79.2 nM) and more pronounced L453V (234.7 nM) or both (307 nM). This was due to an increased IC_{50} and K_m for the substrate (Appendix A), but was partially attenuated by introduction of L462V (124.8 nM).

Together these results showed, that for altering the selectivity of compound **5**, a combination of several exchanged residues was needed, which was only found in *Hs*NMT1 *8x*. Further, *Hs*NMT1 N473H:L495M:Q496L and *Hs*NMT1 *6x* (W297F:A452M:L453V: L462V:L495M:Q496L) with an only slightly increased affinity, indicated that this combination at least needs C-terminal residue exchange (L495M $\pm$ Q496L) as well as R295Q (which alone had no or only minor impact) plus the residues defining the depth of the subpocket (A452M:L453V and eventually surrounding residues).

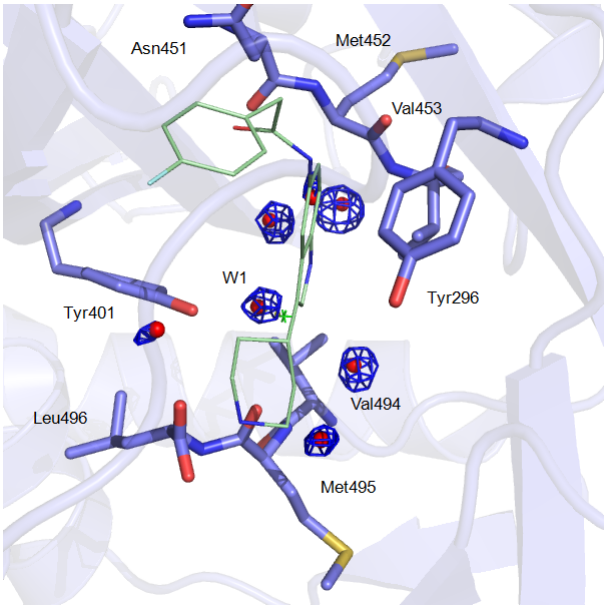


FIGURE 5.11: *Hs*NMT1 *8x* (blue carbon atoms) water density peak analysis from 34 ns explicit solvent MD, blue mesh at 0.07 density cut-off. W1 peak was shifted by 1.0 Å compared to the *Hs*NMT1 wt (green cross). Compound **5** (pale green carbon atoms as lines) binding mode from *Lm*NMT (PDB-ID 4CGN) in overlay for illustration of given space by W1 shift.

*Hs*NMT1 *8x* was also tested against compounds **1-4** with K_i of 26.5 $\pm$ 2.3, 4.4 $\pm$ 2.8, 22.3 $\pm$ 6.5 and 79.4 $\pm$ 33.4 nM. The affinities were increased compared to *Hs*NMT1 wt (Table 3.2, Chapter3.1). The crucial L495M mutation for compound **4**'s selectivity was included in the *8x* mutant as well. However, the inhibition by compound **4** of 18 nM as observed for *Hs*NMT1

N473H:L495M:Q496L and *Hs*NMT1 L495M (Chapter 4) was not completely achieved.

Water analysis of MDs with *Hs*NMT1 *8x* showed a similar variation of the W1 hydration site compared to *Hs*NMT1, as observed in *Hs*NMT1 A452M:L453V (Figure 5.9 e). The density peak of W1 was shifted by 1 Å in the direction of the closest water coordinates found in *Lm*NMT (Figure 5.4 d). This might be sufficient to not interfere with compound **5** binding (Figure 5.11). Additionally, differences in the energetic profile of W1 were identified. While ΔG_{SPAM} of 2.60 kcal/mol was comparable to the *Hs*NMT1 wt (Table 5.4), ΔH_{SPAM} was increased to +0.81 whereas $-T\Delta S_{SPAM}$ was reduced to 1.79 kcal/mol. This indicated that the 8 exchanged residues had impact on the W1 hydration site, and that W1 was less tightly bound in terms of enthalpy.

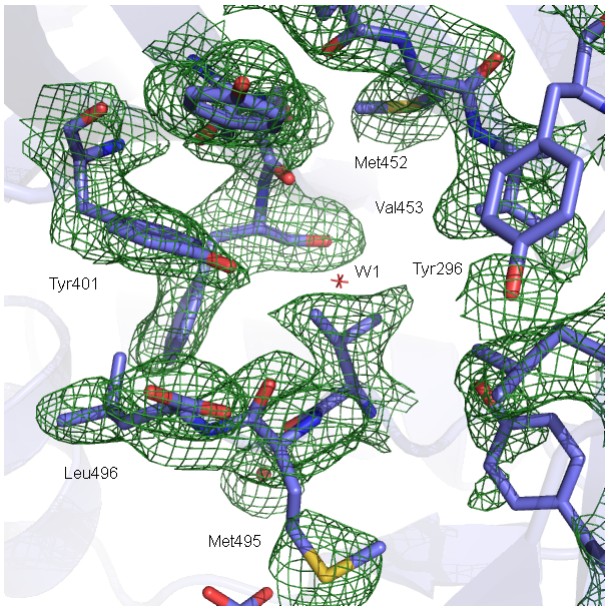


FIGURE 5.12: *Hs*NMT1 *8x* 2Fo-Fc electron density map at 1 σ. No electron density at W1 site (nonbonded cross from *Hs*NMT1 wt coordinates, PDB-ID 3IU1) was found.

For the molecular proof of concept, crystallization trials of *Hs*NMT1 *8x* in absence (for water network verification) and presence of compound **5** were performed. Crystallization conditions from *Hs*NMT1 wt gave three-dimensional crystals for *Hs*NMT1 *8x*, too. Similar to crystallization of *Hs*NMT with compound **5**, no electron density for the ligand within the binding site was found. The best resolution obtained was 2.4 Å. For this structure no electron density for W1 was found either (Figure 5.12). This resolution however, does not allow detailed insights in water positioning (for data collection and refinement statistics see Appendix C). Therefore further investigation on a molecular level has to be performed by obtaining higher resolution crystal structures.

5.4.1 Oxadiazole- and Aminoacylpyrrolidine Inhibitors

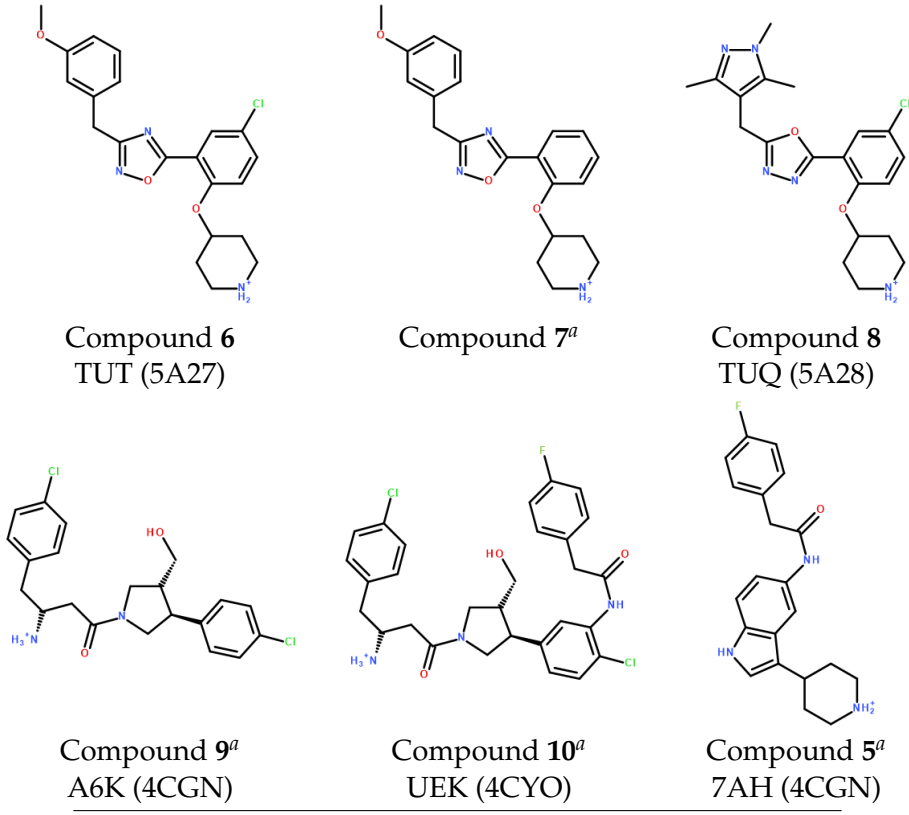


FIGURE 5.13: Structures of compounds **5-10**. PDB three-letter code, if available; corresponding PDB-ID of compound bound to *Lm*NMT in parentheses.[a] selective for *Ld*NMT over *Hs*NMT1.

Besides the indole inhibitor compound **5**, further *Ld*NMT over *Hs*NMT1 selective and non-selective open-conformation binding inhibitors are known (compounds **6-10**, Figure 5.13, see also Chapter 1.2.4, Figures 1.14 and 1.15). The binding mode of compounds **6** and **8** to *Lm*NMT was similar to compound **5**, but included the hydrogen bond of Ser330 formed to the methoxy- or trimethyl-piperazine-moiety, respectively (Chapter 1.2.4, Figure 1.15 b; Rackham et al., 2015). Further, compound **10** is a hybrid molecule of compounds **5** and **9**, which addresses the space close to the C-terminus and the co-factor MyrCoA in NMTs (Chapter 1.2, Figure 1.14). For all these ligands binding to *Lm*NMT, Tyr217 and the closing lid (Asp396, Gly397) were found to be in the open conformation (PDB-IDs 4CGL, 4CYO, 5A27, 5A28). Due to the almost identical sequence (97.8 %) and comparable IC_{50} values in previously reported data between *Ld*NMT and *Lm*NMT (compounds **5** and **9**, Brannigan et al., 2014), we assumed that the same selectivity pattern should be obtained for *Lm*NMT over *Hs*NMT1. To elucidate if the same mechanism for selectivity could be applied as for compound **5**, MDs of *Lm*NMT (starting from PDB-ID 5A27) and *Hs*NMT1 (docking solution in PDB-ID 4C2Y, chain A) in complex with the closely related compounds **6** and **7** were performed.

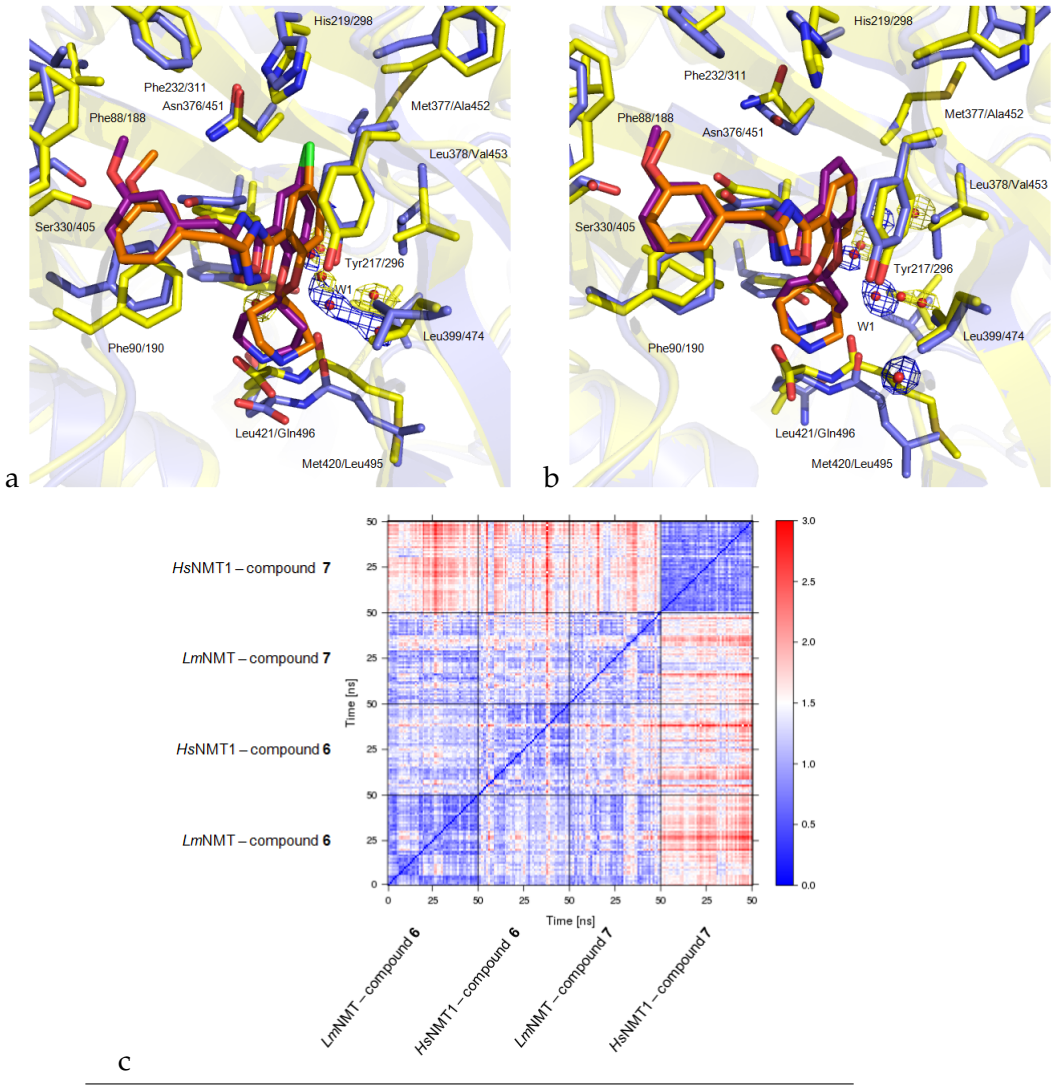


FIGURE 5.14: MD-derived average binding mode of compounds **6** (a) and **7** (b) to *Lm*NMT (yellow carbon atoms and mesh, orange ligand carbon atoms) and *Hs*NMT1 (blue carbon atoms and mesh, deep purple ligand carbon atoms). Density map contour level of 0.07. c) Heavy atom 2D-RMSD plot for compounds **6** and **7** bound to NMTs.

As already observed for compound **5**, a slightly displaced binding mode of the chlorine lacking phenyl and piperidine moiety for selective compound **7** was observed in *Hs*NMT1, which offered space for W1 in *Hs*NMT1, but not in *Lm*NMT (Figure 5.14 b). This binding mode was subsequently confirmed by X-ray analysis (Figure 5.15). Binding of non-selective compound **6** however, was fairly aligned for both enzymes and W1 was present in both NMTs (Figure 5.14 a). This does not only account for representative snapshots, but for the entire trajectories. As represented in the 2D-RMSD plot (Figure 5.14 c) for heavy atoms of compounds **6** and **7**, binding of compound **6** for both NMTs and **7** in *Lm*NMT is very similar (low RMSD, large blue areas) while compound **7** 's binding to *Hs*NMT1 is divergent (high RMSD, red areas), but highly stable within the trajectory. Further the W1 ΔG_{spam} of 2.3 to 3.1

kcal/mol (Table 5.6) are comparable to those described in Chapter 5.2, Table 5.4.

The work-group of Professor Ed Tate (Imperial College London) kindly provided compounds **7-9** for testing in enzyme inhibition assay and crystallography to validate the W1 hypothesis. By co-crystallization of *Hs*NMT1 in complex with compound **7**, this compound's binding mode was identified (Figure 5.15 a). As observed in MDs (Figure 5.14) binding was slightly shifted compared to compound **6** binding to *Lm*NMT (RMSD = 1.10 Å, Figure 5.15 b). Further, electron density was found for W1 which was hypothesized to be the reason for the shifted binding mode and thereby the selectivity of this inhibitor.

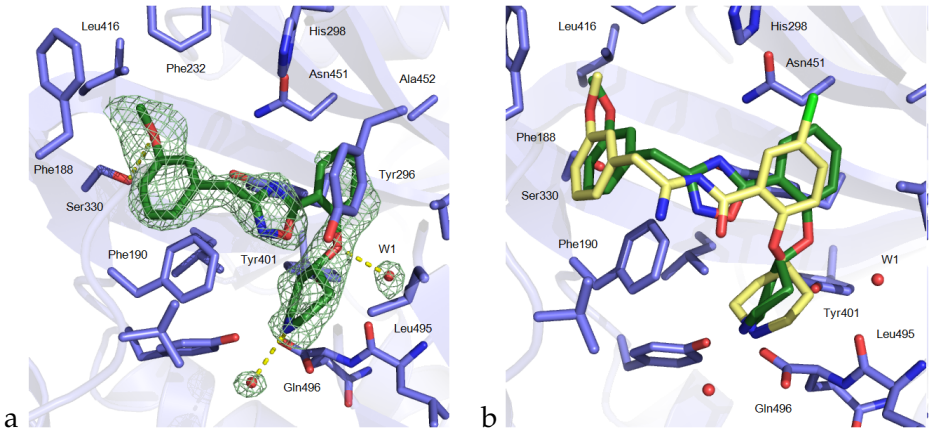


FIGURE 5.15: a) Binding mode of compound **7** in complex with *Hs*NMT1. Ligand with dark green carbon atoms, enzyme with blue carbon atoms. Electron density 2Fo-Fc map for ligand and water molecules at 1 σ. b) Compound **6**'s binding mode from *Lm*NMT (PDB-ID 5A27, pale yellow carbon atoms for ligand, oxadiazole ring opening caused by radiation damage, Rackham et al., 2015) in overlay with compound **7** bound to *Hs*NMT1 (coloring as in a).

TABLE 5.6: W1 spam scores derived from MDs of compounds **6** and **7** in complex with *Lm*NMT and *Hs*NMT1. Units are kcal/mol.[a] W1 density peak shifted by 1.4 Å compared to *Hs*NMT1, see also Figure 5.14 a.

NMT	Ligand	ΔG_{SPAM}	ΔH_{SPAM}	$-T\Delta S_{SPAM}$
*Lm*NMT[a]	6	3.06	-0.65	3.70
*Hs*NMT1	6	2.92	-3.71	6.63
*Lm*NMT	7	-	-	-
*Hs*NMT1	7	2.27	-4.39	6.66

TABLE 5.7: Inhibition data (K_i) for compounds **7-9**. All results in nM, values in parentheses describe SI compared to *Lm*NMT wt.

NMT	**8**	**7**	**9**
*Lm*NMT	24.8 ± 2.3	70.8 ± 52.4	276.5 ± 51.6
*Hs*NMT1	111.3 ± 31.0 (4.5)	3 384.5 ± 2 273.1 (47.8)	2 910.0 ± 1 292.9 (10.5)
*Hs*NMT1 *8x*	4.8 ± 1.4 (0.2)	33.0 ± 11.9 (0.5)	29.2 ± 8.6 (0.1)
*Hs*NMT1 L495M	36.3 ± 6.2 (1.5)	1 575.3 ± 363.9 (22.2)	554.8 ± 277.0 (2.0)
*Hs*NMT1 N473H: L495M: Q496L	42.5 ± 7.2 (1.7)	2 102.4 ± 1 520.7 (29.7)	417.3 ± 43.2 (1.5)

The results from the enzyme inhibition assay (Table 5.7) confirmed the selectivity to *Lm*NMT over *Hs*NMT1 for compounds **7** and **9** with SI of 46.8 and 10.5 respectively. Introduction of the eight point mutations in *Hs*NMT1 *8x* led to a stronger inhibition of all compounds and the loss of selectivity for **7** and **9** (SI 0.5 and 0.1). To elucidate if the underlying mechanism was the same as observed for indole compound **5** or sulfonamide compound **4**, inhibitors were tested against *Hs*NMT1 L495M and *Hs*NMT1 N473H:L495M:Q496L. Increased affinity (by lower K_i of 554.8 nM and 417.3 nM) and loss of selectivity (SI 2 and 1.5) were observed for compound **9**, but not for compound **7**.

This result indicated, that the W1 hypothesis for selectivity could be applied for compounds **5** and **7**, where only in *Hs*NMT1 *8x* the selectivity could be erased. For compound **9** however, the hypothesis of the impairment of the C-terminus' flexibility seemed more reasonable, because *Hs*NMT1 L495M was sufficient in this case, as it was for compound **4**. This finding was further related to the different binding mode of compound **9** compared to **6-8** addressing the area between the NMT C-terminus and MyrCoA instead of the W1 position (Figure 5.16). Even though compound **9** binds to the Tyr217-open conformation in *Lm*NMT, it does not enter deep into the hydrophobic pocket where W1 is located. This also explained the presence of W1 with its very negative DrugScore in the crystal structure of this compound in *Lm*NMT (Table 5.3).

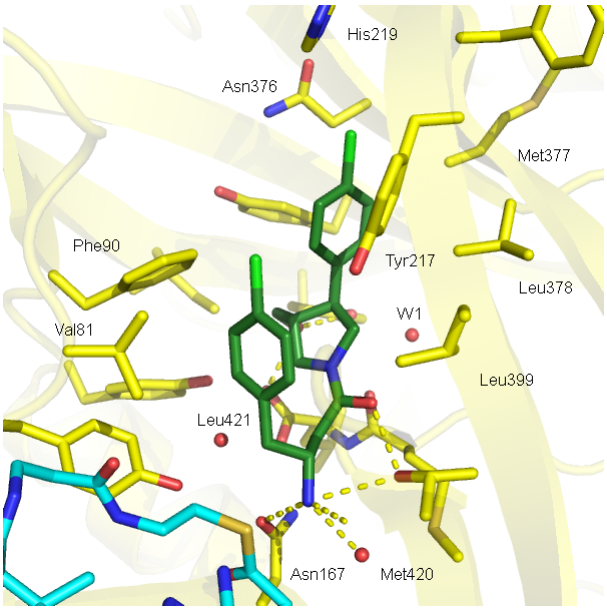


FIGURE 5.16: Crystallographic binding mode of compound **9** (dark green carbon atoms) in *Lm*NMT (yellow carbon atoms, PDB-ID 4CGL). Ligand occupies space between C-terminus and MyrCoA (cyan carbon atoms). W1 is present in crystal structure.

5.5 Discussion

The elucidation of backbone and side chain S^2 from MDs indicated residues within and close to the binding site, whose flexibility was altered upon ligand binding. Gly397/472 of the closing lid were identified to be flexible by it's N-H bond vector S^2 (Figures 5.1, 5.3). Closed- (compounds **1-4**) or open-conformation (compound **5**) binding ligands also forced Tyr217/296 in one or the other conformation and impaired its flexibility (side chain S^2, Table 5.1). While it was assumed that the orientation of Tyr217/296 and the preferrence of Tyr296-closed conformation in *Hs*NMT1 was the selectivity determining feature of compound **5** earlier (Brannigan et al., 2014), in a recent publication (Rackham et al., 2015) non-selective open-conformation binding inhibitors **6** and **8** were identified. This and the results from MDs, where the Tyr-open conformation was preferably found in unbound *Hs*NMT1 (Table 5.2), contradicted the hypothesis that Tyr217/296 orientation determines selectivity.

However, in close proximity to compound **5** and Tyr217/296, the water-density analysis revealed the highly conserved water molecule W1 in *Hs*NMT1 (Figures 5.4 to 5.6, Table 5.4). W1 density peaks corresponded to crystallographically observed electron density. Based on MDs, W1 was likely to alter compound **5**'s binding mode in *Hs*NMT1 to be less favourable compared to *Lm*NMT and thereby being the driving force behind selectivity (Figure 5.8).

Based on this hypothesis, that the highly conserved W1 in *Hs*NMT1 is the

selectivity determining feature in open-conformation binding ligands, eight point mutations were introduced in *Hs*NMT1 to destabilise W1. Testing of this *Hs*NMT1 *8x* against compound **5** revealed a loss of selectivity, which was not found for those mutants with a subset of exchanged residues (Table 5.5; not all potential 2^8 = 256 permutations tested). Further, a loss of selectivity was also found for compounds **7** and **9** (Table 5.7).

For the latter, already the M495L exchange was sufficient for altering selectivity, as it was found for (closed conformation binding) sulfonamide inhibitor **4**. This was most likely due to the different binding mode of compound **9** compared to the other open-conformation binders (Chapter 1.2.4, Figures 1.14, 5.16). Both compounds **4** and **9** address the area between C-terminal carboxylate and MyrCoA with more or less large parts of the molecule, instead of the hydrophobic pocket at W1 position. For those ligands the underlying selectivity mechanism was the impairment of the C-terminus flexibility (Chapter 4). This also explained the initially surprising presence of W1 with negative DrugScores in *Lm*NMT bound to selective compounds **9** and **10** (Chapter 5.2, Table 5.3, PDB-IDs 4CGL, 4CYO). Therefore the W1 hypothesis cannot be applied for selective open conformation binding NMT inhibitors in general, but only those which enter the hydrophobic pocket deep enough to interfere with W1 in *Hs*NMT1.

To proof the W1 hypothesis to be the selectivity determining feature for compounds **5** and **7** on a molecular level, crystallization trials were performed. Unfortunately crystals diffracting up to 2.4 Å had no electron density for compound **5** and allowed no final analysis of binding site water molecules in *Hs*NMT1 *8x*. Further, optimization and trials of co-crystallization of compound **7** with *Hs*NMT1 were successful. The shifted binding mode of this compound observed in MDs (Figures 5.8, 5.14), was proven by this crystal structure and confirmed the presence of W1 in the complex (Figure 5.15). This gave further evidence for W1 being the selectivity determining feature.

Chapter 6

Virtual Screening for novel and selective *Lm*NMT Inhibitors

6.1 Pharmacophore Hypothesis

Based on the results described in Chapter 5, a virtual screening (VS) was performed on *Lm*NMT (PDB-ID 5A28) with a strict pharmacophore query, including the features that were hypothesized to be crucial for high affinity and selectivity of open-conformation binding ligands. The pharmacophore model contained the following features (Figure 6.1): 1) A hydrogen-bond acceptor to interact with the hydroxyl group of Ser330 (blue sphere, F2). As illustrated in Figure 6.1 c), the ligand moieties can vary in this sub-pocket. Therefore a rather large radius of 1.4 Å (F2) was chosen for this feature. 2) An aromatic ring (F3, yellow sphere, like pyrazole or fluorophenyl from known ligands) was chosen to interact with Phe90 and further lipophilic residues in this pocket (Phe88, Phe232, Leu341). 3) Additionally, an aromatic moiety (F4, yellow sphere, (Figure 6.1 d) to interact with Tyr217. F4 was placed on carbon 3a of the indole core of compound **5**, which was in a similar position within the binding site as an aromatic atom of the linker moieties of compounds **6** and **8**. 4) The grey sphere of F1 represented any atom to displace or interfere with W1 in *Hs*NMT1 in a similar position as carbon c2 from compound **5** (see also Figure 5.8 a). This feature was chosen to increase selectivity over *Hs*NMT1 based on the W1 hypothesis (Chapter 5). 5) One basic center (F5, red sphere) to interact with the C-terminal residue Leu421. Known from different ligands binding (PDB-IDs 2WSA, 4CGN, 5A28), the position of this protonated moiety allowed some orientational variation, therefore F5 was chosen with a large radius of 1.7 Å. All receptor atoms were defined as excluded volume and all features were essential for molecules to pass the query.

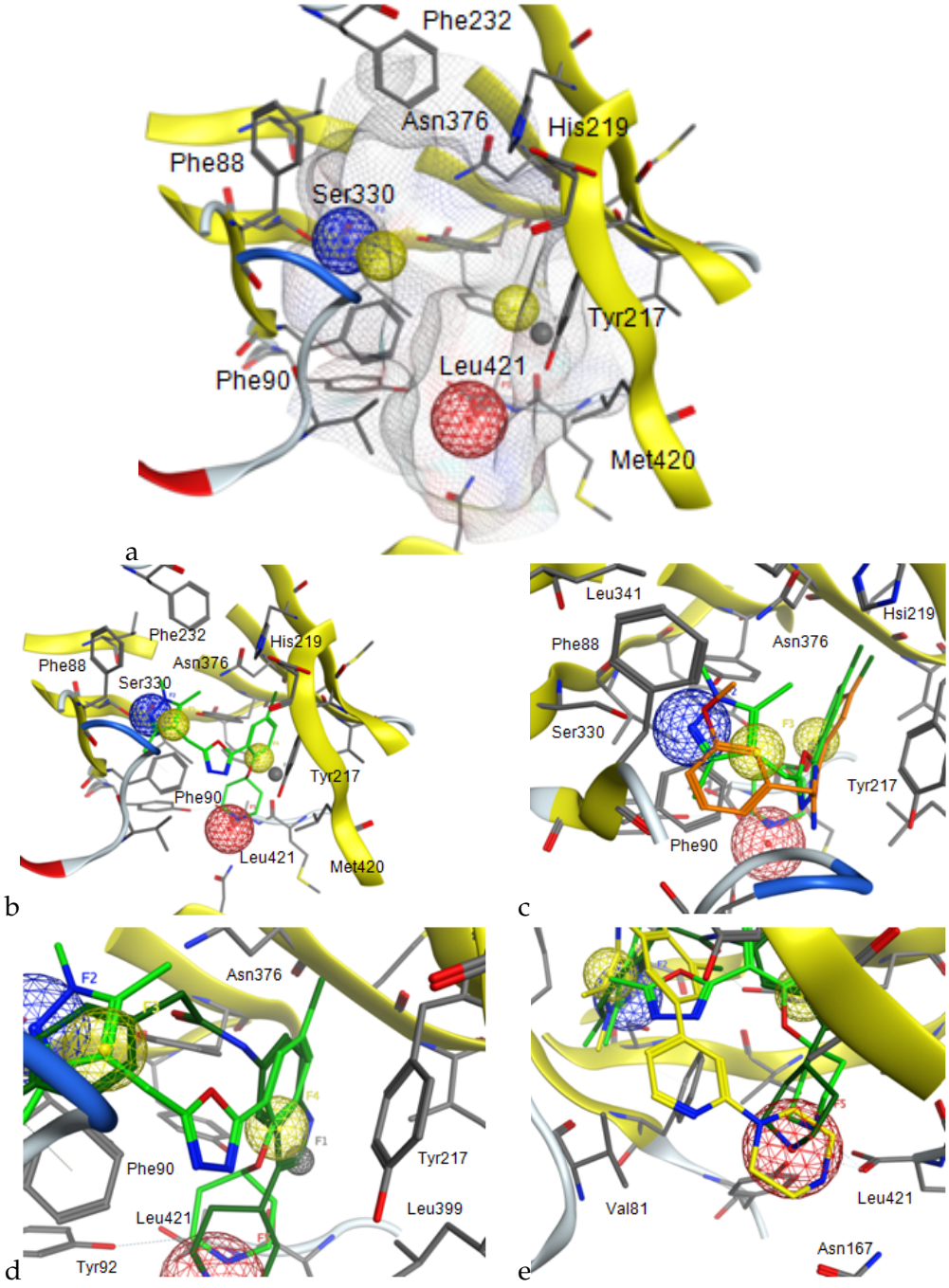


FIGURE 6.1: Pharmacophore features for virtual screening. a) Complete pharmacophore model containing all features together with protein surface. F1 (grey) any atom, F2 (blue) H-bond acceptor, F3 (yellow) aromatic, F4 (yellow) aromatic, F5 (red) basic center with positive charge. b) Same as (a), but with compound **8** as reference ligand (green carbon atoms). c) F2 and F3 in overlay with compound **8** (green carbon atoms) and **6** (orange carbon atoms). d) F1 and F4 with compound **8** (green carbon atoms) and **5** (darkgreen carbon atoms). e) F5 and compounds **8** (green), **5** (darkgreen) and **1** (yellow carbon atoms). Figures made with MOE 2015.1001

6.2 Virtual Screening of a large Compound Library derived from the ZINC Database

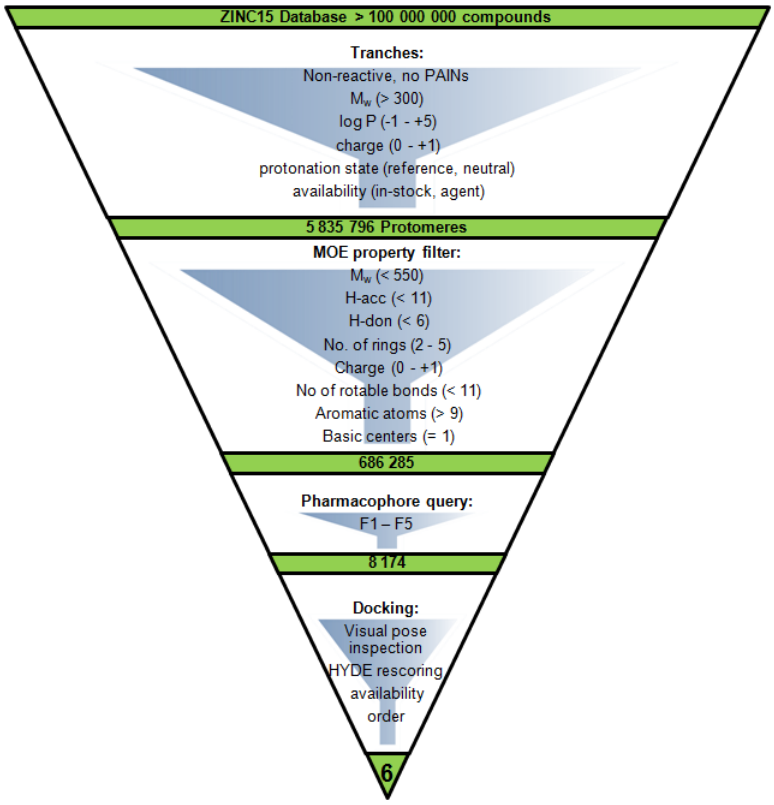


FIGURE 6.2: Schematic virtual screening worklow including property filters, pharmacophore search and docking.

The virtual small molecule compound library was derived from ZINC 15 database using the new 3D *tranches browser* (Irwin et al., 2012; Sterling and Irwin, 2015). As filters reactivity ("anodyne" to exclude reactive molecules and PAINs), molecular weight (M_w > 300 Da), log P (-1 – +5), charge (0 – +1), availability (in stock or via agent) and "pH" (reference and neutral to obtain only relevant protonation states) were applied. This resulted in a total of 5 835 796 protomeres which were additionally filtered using MOE 2015.1001. The following molecular properties were allowed to pass this additional filter step: M_w < 550 Da, H-bond acceptor count < 11, H-bond donor count < 6, no. of rings 2-5, charge 0 – +1, no. of rotable bonds < 11, aromatic atoms > 10 (corresponds to 2 or more aromatic rings) and exactly 1 basic atom. Generated conformations of the remaining 686 285 molecules were subjected to the pharmacophore query described above, which was passed by 8 174 compounds. These were docked without any restraints using LeadIT-2.1.6 after receptor and procedure validity check (Table 6.1). The resulting poses and scores were visually inspected and the 500 highest scoring molecules representing the desired binding interactions (H-bond with Ser330, direct or

indirect ionic interaction with C-terminal Leu421 and aromatic expansion into the hydrophobic pocket close to Tyr217) were rescored using the HYDE scoring function. After commercial availability check the 6 most promising compounds were purchased (Table 6.2). A schematic work-flow of the performed filtering steps is illustrated in Figure 6.2.

6.3 Docking Analysis

Validity of the prepared receptor (PDB-ID 5A28, Figure 6.3 a) and the docking procedure was ensured by redocking of reference ligand compound **8** with an RMSD of 1.05. Additionally, compounds **5** and **6** were docked and adopted their native binding mode, as well (Figure 6.3 b-d). Corresponding scores and HYDE are summarized in Table 6.1.

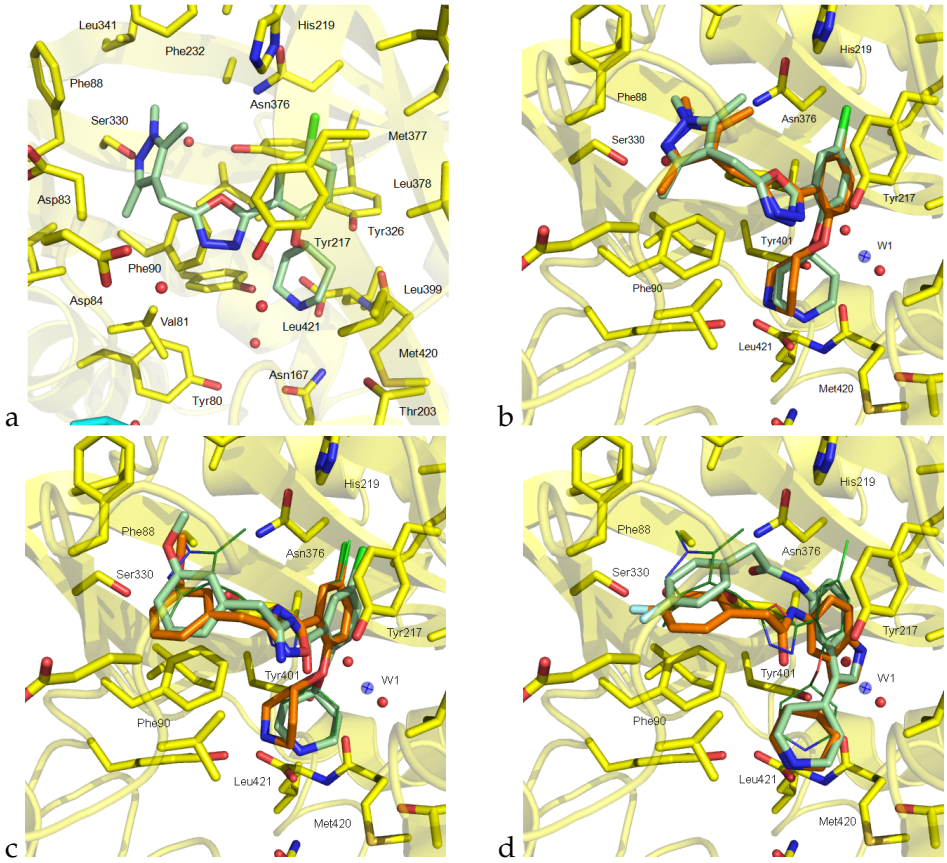


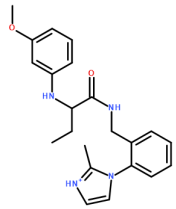
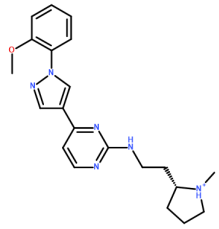
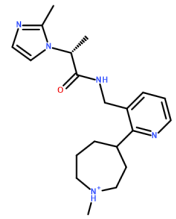
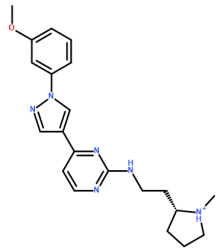
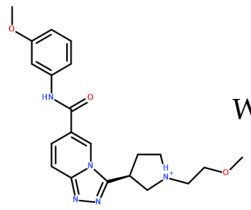
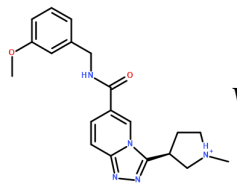
FIGURE 6.3: a) Illustration of prepared receptor (PDB-ID 5A28) including 4 water molecules, each with 3 interactions, but none in the hydrophobic pocket, as part of the receptor. b-c) Virtual screening receptor validation by redocking of open-conformation binding inhibitors compounds **8** (a), **6** (b, PDB-ID 5A27) and **5** (c, PDB-ID 4CGN). Reference ligands with palegreen carbon atoms, compound **8**'s pose with dark green carbon atoms as lines, docking poses with orange carbon atoms, receptor *Lm*NMT with yellow carbon atoms. Transparent blue sphere indicates W1 position from *Hs*NMT1. Red spheres for *Lm*NMT hydrations sites within that pocket (not included as part of receptor during docking, only for visualization purpose).

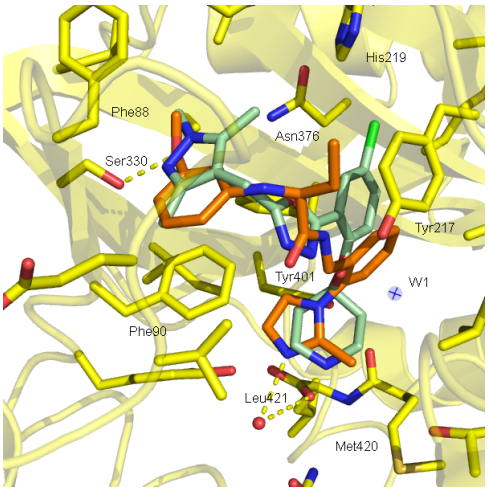
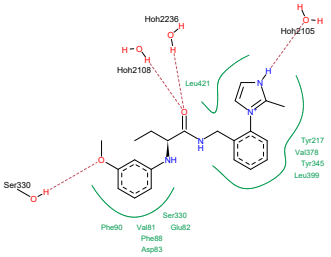
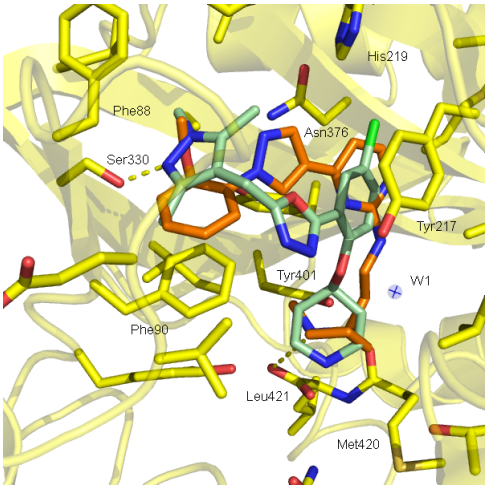
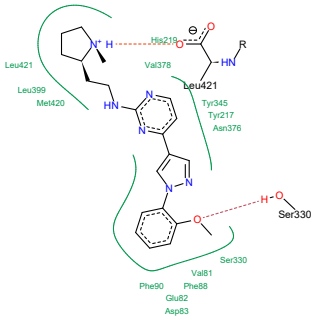
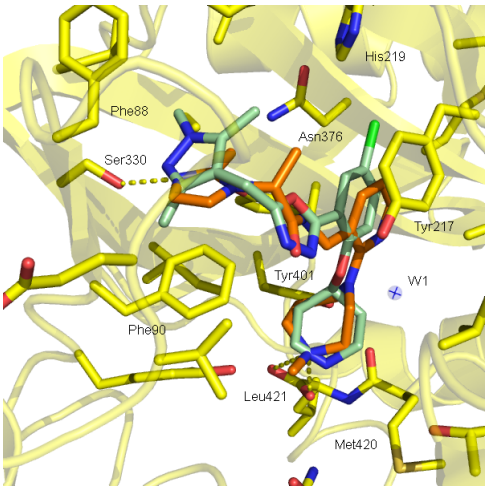
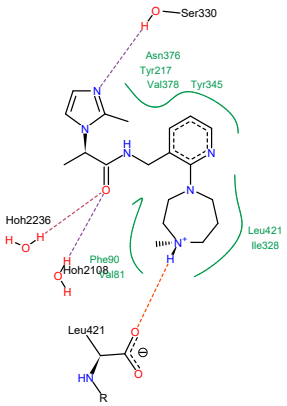
TABLE 6.1: Scores (in kJ/mol) for redocking of compounds **5**, **6** and **8**.

Compound	**Score**	**HYDE**	**RMSD [Å]**
8	-38.8	-58	1.05
6	-36.2	-52	1.36
5	-33.5	-46	1.70

The predicted binding modes of the purchased compounds (Table 6.2) are displayed in Figure 6.4. For **ChBr2**, **ChBr3** and **WuXi1** direct and for **ChBr1**, **ChBr4** and **WuXi2** indirect, water-mediated ionic interaction with Leu421 were predicted. The poses also included H-bonds with Ser330. Further, ligand atoms were placed close to the corresponding W1 position in *Hs*NMT1 indicating interferrence in the off-target.

TABLE 6.2: Virtual screening hits together with their scores (in kJ/mol).

Structure	Internal ID	ZINC ID	Supplier	Score	HYDE
	ChBr1	55369988	ChemBridge (racemic mixture)	-34.2	-53
	ChBr2	55060934	ChemBridge (racemic mixture)	-34.9	-50
	ChBr3	67446715	ChemBridge (racemic mixture)	-32.1	-49
	ChBr4	20869258	ChemBridge (racemic mixture)	-29.9	-45
	WuXi1	96267504	WuXi-AppTec (racemic mixture)	-30.2	-44
	WuXi2	96267435	WuXi-AppTec (racemic mixture)	-32.4	-43

Cpd.	Binding Pose	Poseview
ChBr1		
ChBr2		
ChBr3		

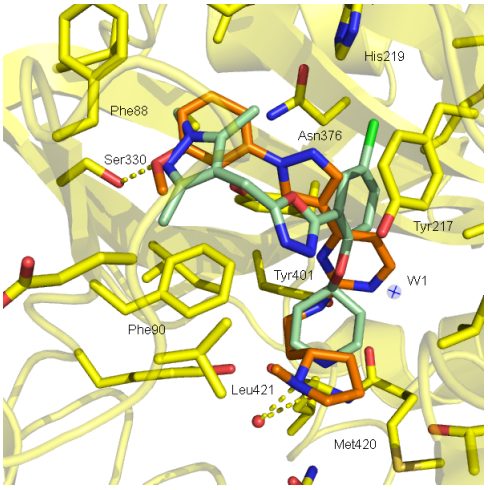
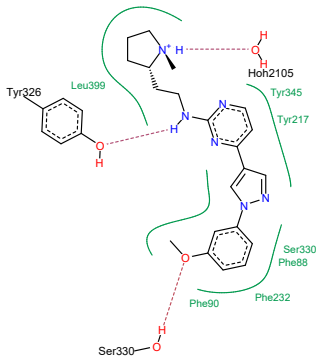
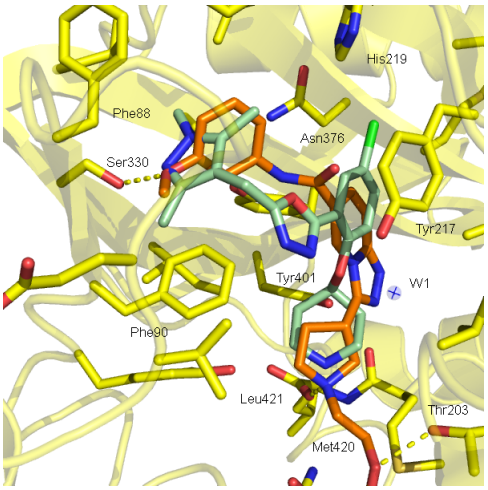
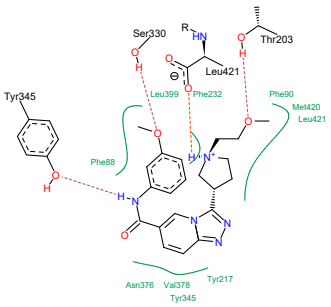
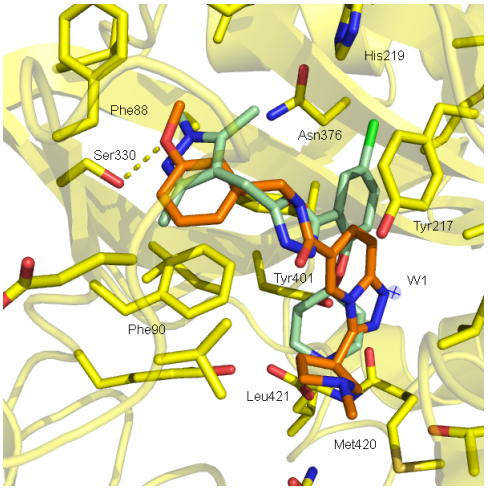
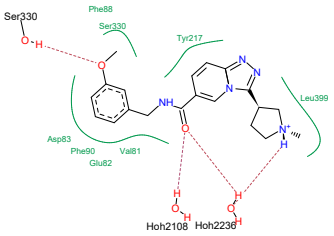
Cpd.	Binding Pose	Poseview
ChBr4		
WuXi1		
WuXi2		

FIGURE 6.4: Predicted binding poses and "pose view" interaction maps (Fricker, Gastreich, and Rarey, 2004; Stierand and Rarey, 2007; Stierand and Rarey, 2010) of VS hits **ChBr1-4** and **WuXi1-2**. Docking pose with orange carbon atoms, reference ligand compound **8** with pale green carbon atoms and receptor *Lm*NMT with yellow carbon atoms. Aligned W1 position in *Hs*NMT1 as blue transparent sphere.

A variety of hits with a protonated imidazole or thiazole moiety interacting with the C-terminal Leu421 were also observed. Even though there are NMT inhibitors known with this feature (*Candida albicans* NMT - *Ca*NMT, PDB-ID 1IYK, Figure 6.5; Sogabe et al., 2002), it was not ultimatively determined if the protonated state was dominant at physiological pH (predicted pK_a = 7.01). Thus, among these compounds, only **ChBr1** was representatively ordered, as it had the best score (HYDE = -53 kJ/mol).

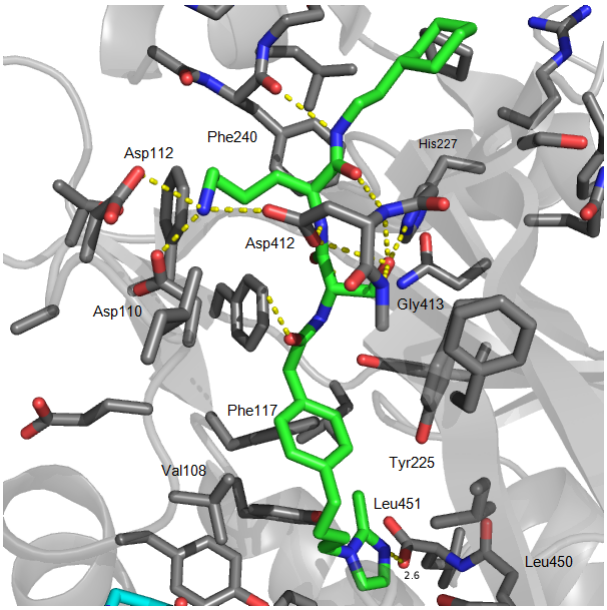


FIGURE 6.5: *Ca*NMT inhibitor with imidazole moiety, potentially interacting with C-terminal leucine, PDB-ID 1IYK.

6.4 Experimental Results of selected Compounds

Compounds **ChBr1-4** and **WuXi1-2** were tested in enzyme inhibition assay against *Lm*NMT and *Hs*NMT1 in an initial screening at concentrations of 10 and 100 μM. Only **ChBr2**, **ChBr3** and **ChBr4** revealed a concentration dependent inhibition against *Lm*NMT, whereas no inhibition at 100 μM (corresponds to an $K_i \gg 40$ μM) and lower inhibition at a further concentration of 500 μM against *Hs*NMT1 was observed. The subsequently determined K_is and ligand efficiencies (LE = 1.4*(-log(K_i))/heavy atoms in kJ/mol/HA; Hopkins, Groom, and Alex, 2004; Shultz, 2013) are summarized in Table 6.3.

TABLE 6.3: K_i values for of ChBr 2, ChBr3 and ChBr4 against *Lm*NMT and *Hs*NMT1. K_i in μM, values in parentheses are % inhibition at 500 μM. [a] Ligand efficiency (LE) in kJ/mol/HA.[b] *Hs*NMT1 $K_i \gg 40$ μM, but < 200 μM.

Compound	**K_i *Lm*NMT**	**hill slope**	**LE[a]**	**K_i *Hs*NMT1**
ChBr2	19.6 ± 3.4 (91)	1.05	0.24	$\gg$40[b] (59)
ChBr3	52.3 ± 7.3 (90)	0.81	0.23	>200 (30)
ChBr4	7.3 ± 3.0 (97)	1.31	0.26	$\gg$40[b] (59)

6.5 Discussion

The pharmacophore search and virtual screening revealed some compounds with known features as methoxy-phenyl or imidazol (**ChBr3**) moieties to interact with Ser330, and ring-bound secondary and tertiary amines to interact with C-terminal Leu421 (Brand et al., 2012; Brand et al., 2014; Rackham et al., 2015). Testing of the best scoring compounds delivered three hits against *Lm*NMT with low to mid μM-affinity and potential selectivity over *Hs*NMT1, as no inhibition up to 100 μM and lower inhibition at 500 μM was observed against the off-target (Table 6.3). The compounds' binding poses were predicted to occupy space close to W1 coordinates (Figure 6.4). Thus, a energetically unfavourable displacement of W1 in *Hs*NMT1 or a shifted binding mode due to the stability of W1 in the off-target were hypothesized. This supports the potential key-role of W1 being a selectivity determining feature in NMT (see also Chapter 5).

Further, the low molecular weight of **ChBr 2-4** (all below 380 g/mol) makes those compounds a potential starting point for further investigation and optimization (even though the LEs of 0.23 – 0.26 are not in an ideal range).

Chapter 7

Concluding Remarks and Future Perspectives

7.1 Summary and Conclusion

In this thesis, the driving forces for selective inhibition, if the binding sites of target and off-target are conserved, were elucidated. The model system pair *Lm*NMT and *Hs*NMT1 and a set of non-selective and selective inhibitors to identify these selectivity determining features were characterised. As pointed out in Chapter 1.1 protein flexibility, binding site water molecules and particularly their changes upon ligand binding, can play a crucial role for affinity and selectivity. By MDs and site-directed mutagenesis, the residues being relevant for selective inhibition were identified and validated by enzyme inhibition assays, ITC experiments and crystallography.

Depending on the type of inhibitors, two different reasons for selectivity were found: The three non-conserved binding site residues play a role in selectivity, even though their side chains are oriented away from the ligand and form no interactions with it (Figure 4.1). However, the C-terminus-1 residue *vice versa* exchange resulted not only in a loss of affinity and selectivity for compound **4** in *Lm*NMT, it also inactivated the enzyme (Table 4.1). In *Hs*NMT1 on the other hand it was identified to be the selectivity determining residue for inhibition by compounds **4** and **9** (Table 5.7). MDs strongly indicated that this residue largely influenced the flexibility of the catalytic C-terminus of NMT which also interacted with the inhibitors over a salt bridge. Upon selective ligand binding the C-terminus side chain S^2 was increased only in *Hs*NMT1 (Table 4.3, Figure 4.5). ITC experiments confirmed that the increased affinity of compound **4** to *Hs*NMT1 L495M was caused by a gain of entropy compared to the *Hs*NMT wt (Table 4.2, Figure 4.3).

Inhibition of *Hs*NMT1 by compounds **5** and **7** was not affected by the M495L mutation. For elucidation of those compounds' selectivity, binding site hydration analysis was performed. Qualitative characterisation of binding site water molecules identified W1 to be highly conserved in *Hs*NMT1,

but was only found in *Lm*NMT when bound to certain inhibitors and not in the ligand-free state (Figures 5.4 - 5.8). Quantification of W1 thermodynamics by the MD-based method SPAM revealed a favourable enthalpy and unfavourable entropy of W1 (Table 5.4). As this water molecule was found in *Lm*NMT bound to the sulfonamide inhibitor **1** as well, it can partially explain the thermodynamic fingerprint of ligand binding to *Lm*NMT found in ITC experiments. In ITC, ligand binding to *Lm*NMT was characterised by a higher enthalpic and lower entropic contribution compared to *Hs*NMT1 (Figure 3.3, Tables 3.3, 3.4). The trapping of W1 in the binding site of *Lm*NMT might cause that enthalpy gain with an entropic penalty.

Concerning selectivity of compound **5**, W1 might play a key-role as this inhibitor expands into the area close to W1 coordinates within the binding pocket. The exchange of eight residues, all chosen to destabilise W1 in unbound *Hs*NMT1, resulted in *Hs*NMT1 *8x*, a variant which showed increased affinity and loss of selectivity for all selective inhibitors powering the W1 hypothesis (Tables 5.5, 5.7). MDs gave hint that selective compounds **5** and **7** might not displace the conserved W1 in *Hs*NMT1, but adapt a slightly shifted - potentially less favourable - binding mode (Figures 5.8, 5.14). For compound **7** this shifted binding mode in *Hs*NMT1 and the presence of W1 was confirmed by crystallography (Figure 5.15). All together these results strongly support the hypothesis that W1 determines selectivity for those compounds.

Based on known ligand features and interactions (Brand et al., 2012; Brannigan et al., 2014; Rackham et al., 2015), as well as the "W1 hypothesis", a VS with a strict pharmacophore search (Figure 6.1) was performed and resulted in three low to mid μM inhibitors of *Lm*NMT with selectivity over *Hs*NMT1 (**ChBr 2-4**, Tables 6.2, 6.3). As these ligands were chosen based on their predicted binding mode addressing W1 coordinates comparable to compound **5**, this result was a further indication for the importance of W1 for selectivity and the accuracy of the W1 hypothesis.

Finally MDs proved their suitability not only in reproduction of experimental results (crystallographic B-factors, hydration sites, Figures 3.14, 5.4), they also gave the decisive hints by S^2, which regions and residues to focus on and to be mutated for identification of the selectivity determining features within the two NMTs (Figures 4.5, 5.1, Tables 4.3, 5.1). The strategy presented herein consisting of crystallographic analysis and MDs for hypothesis formulation and refinement, site-directed mutagenesis and experimental validation (inhibition assay, ITC, crystallography) for proof of concept, demonstrated success in the identification of the selectivity determining features when facing a conserved binding site. Even a predictive value of the methodology was demonstrated by the VS. Due to the highly refined pharmacophore hypothesis, not only a high hit-rate (3 out of 6) was

achieved, the selected compound were also selective.

The degree of how detailed protein flexibility and binding site hydration need to be incorporated in rational drug design may vary. If in some cases visual inspection of the crystal structure may reveal enough information for the design of high affinity ligands with acceptable selectivity, some examples from Chapter 1.1 highlight the importance of the incorporation of more complex methods like MDs, or Watermap (Young et al., 2007; Abel et al., 2008). In the end of the day, this thesis can be understood as a piece of a puzzle to be used for the improvement of the general understanding of selectivity determining features of proteins with conserved binding sites and to open perspectives in the prediction of selectivity based on protein flexibility and binding site water molecules. This approach may also give insights in cases, where no selectivity can be achieved and reveal the reasons for this. The transfer of the methods used, on new targets and their off-targets will reveal the general applicability of this strategy and hopefully reveal the selectivity determining features in those projects and potentially lead to novel selective ligands.

7.2 Preliminary Results and Perspectives

7.2.1 Proof of W1 Hypothesis

For further molecular proof of the W1-hypothesis, additional crystal structures are needed. Crystals of *Hs*NMT1 and *Hs*NMT1 *8x* in presence of compound **5** were grown. However, no electron density of a ligand was found within the binding site. To evaluate if W1 is displaced upon ligand binding in *Hs*NMT or if W1 is still present and the ligand adapts a shifted binding mode compared to *Lm*NMT, further optimization is necessary. As the same hypothesis accounts for compound **7**, also crystals were grown with this ligand and revealed the shifted binding mode compared to compound **6** bound to *Lm*NMT. Structures of compound **6** in *Hs*NMT1 and **7** in *Lm*NMT will give further insights and complete this elucidation.

An alternative computational method for binding site hydration may also be used for comparison with the results obtained by MDs and SPAM. The maybe most prominent approaches suitable for such a study might be Watermap (Young et al., 2007; Abel et al., 2008) or GRID (Goodford, 1985).

The selectivity of compounds **5** and **7** could only be erased in *Hs*NMT1 *8x*. The question if only a subset of exchanged residues and which ones determine the loss of selectivity can be answered by successively reducing the mutated amino acids from *Hs*NMT1 *8x*. Based on already designed mutants (especially *Hs*NMT1 *6x* and *Hs*NMT R295Q:N473H:L495M:Q496L, Table 5.5), a combination of different mutated residues seem necessary. This

combination might include the C-terminus (L495M ± Q496L), A452M ± L453V to define the depth of the hydrophobic pocket for W1 stability with potential further surrounding residues and R295Q, which showed some increased affinity, only in combination with other residues. A suitable first reduction step might therfore leave out the N473H and W297F mutations, but include R295Q, A452M, L453V, L462V, L495M and Q496L.

7.2.2 Binding Kinetics of NMT Inhibitors

Besides the binding thermodynamics also the emerging field of the elucidation of binding kinetics gives insights into protein-ligand interactions and driving forces behind affinity (Bernetti, Cavalli, and Mollica, 2017). Reported (Frearson et al., 2010) and our own preliminary results (Appendix B) proved the suitability of NMT for surface plasmon resonance (SPR) experiments. The detailed MD characterization can be even expanded by so-called steered MDs which gave insights into the molecular mechanisms of binding kinetics previously (De Vivo et al., 2016). In combination both methods are able to improve the understanding of selective binding and the underlying kinetics in NMT and in general.

7.2.3 Experimental Order Parameters

While S^2 from MDs are able to reproduce experimental S^2 and gave the decisive hints on which regions to focus within the two NMTs, they were still derived from a simulation. To quantitatively confirm MD-derived S^2, NMR experiments can be performed. First protein expression of *Hs*NMT1 using unlabeled M9 mineral medium was successful with comparable yield to standard expression an purification procedure (18.5 mg). The use of labeled nitrogen and carbon source is a reasonable next step. However, the size of the protein of around 50 kDa might be a challenge.

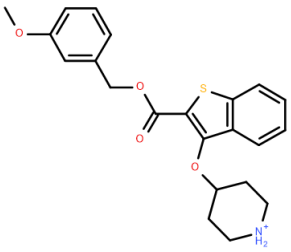
7.2.4 Ligand Optimization

Throughout this thesis two different selectivity determining features between *Lm*NMT and *Hs*NMT1 were identified. To improve the selectivity or the affinity of known inhibitors, a hybridization approach can be used. Such an approach was performed previously combining compounds **5** and **9** to yield compound **10** (Figure 1.14, Hutton et al., 2014). This compound had higher affinity, but lower selectivity than **5** and **9**. Based on the W1 hypothesis, this was most likely due to leaving out the part of compound **5** that addresses the W1 pocket. A combination of compound **5** or **7** addressing that pocket and compounds **4** or **9** addressing the area close to the C-terminus and MyrCoA might synergistically benefit from both features for selectivity.

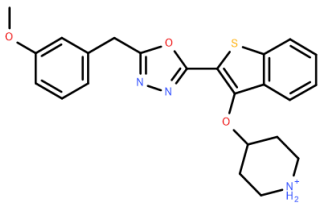
Further, the virtual screening (Chapter 6) revealed three low to mid μM-affinity inhibitors of *Lm*NMT with selectivity over *Hs*NMT1. Once again crystallography can give molecular evidence for the underlying mechanism and the correctness of the predicted binding mode. As **ChBr2-4** LEs are 0.23 – 0.26, these inhibitors hold some potential for further optimization to result in selective and high affinity inhibitors of *Lm*NMT. This can be done by some derivatisation without largely increasing the molecular weight to yield a more ligand efficient lead structure or by the hybridization approach, described above. **ChBr2-4** could be used instead of compounds **5** or **7** in that case, if the predicted binding mode is correct and W1 is addressed.

7.2.5 Long-Term Perspectives - New Targets

The successful application of the methodology on *Lm*NMT and *Hs*NMT1 makes this approach interesting to be transferred to new targets. A first expansion could incorporate NMT of *Plasmodium vivax* (*Pv*NMT). This protein shares high similarity with the NMTs already elucidated and also holds a plethora of crystal structures and non-selective and selective inhibitors, but with different selectivity profile compared to *Lm*NMT and *Hs*NMT1 (see Chapter 1.2 and references therein). Selectivity for *Pv*NMT is still a challenge (Goncalves et al., 2017) and the explanations for selective inhibition are not clear. For the benzothiophene series of *Pv*NMT-selective inhibitors, one argumentation for selectivity over *Lm*NMT includes the given space of the hydrophobic pocket addressed by the benzothiophene core (Figures 7.1 a and b, Rackham et al., 2013; Rackham et al., 2014; Rackham et al., 2015; Yu et al., 2015). While this might be possible for *Lm*NMT, *Hs*NMT1's residues at the corresponding positions are identical with *Pv*NMT, namely Ala366/452 and Leu367/453 (Figure 7.1 c). Hence, this argumentation is not sufficient to explain the selectivity over *Hs*NMT1. Certainly, the already obtained results and knowledge about the regions of interest (W1, C-terminus) can help to reveal the selectivity determining features of *Pv*NMT inhibition, too.

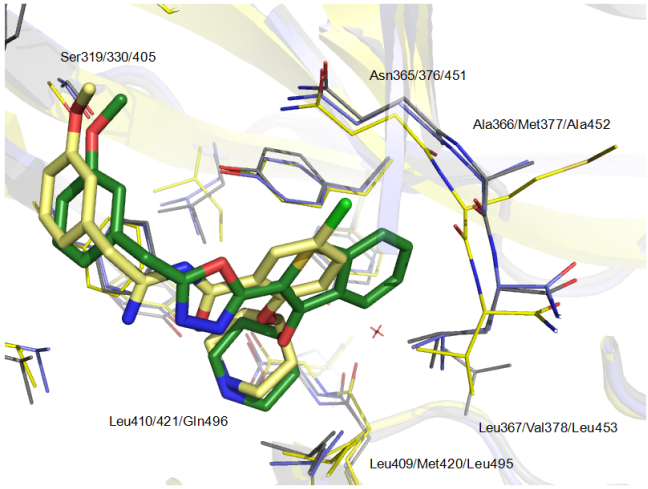


a



b

	K_i (μM)	K_i (nM)
*Pv*NMT	0.08	18
*Hs*NMT1	3.20	428
*Lm*NMT	5.82	



c

FIGURE 7.1: a) and b) selective benzothiophene *Pv*NMT inhibitors. c) Inhibitor (b, dark green carbon atoms) bound to *Pv*NMT (grey carbon atoms, PDB-ID 4CAE) in superposition with *Lm*NMT (yellow carbon atoms and pale yellow ligand **7**, PDB-ID 5A27) and *Hs*NMT1 (blue carbon atoms, PDB-ID 3IU1). W1 coordinates from *Hs*NMT1 as red cross. Residues labeled as *Pv*NMT/*Lm*NMT/*Hs*NMT1.

Another NMT of interest is from *Tyrpanosoma brucei*, the pathogen of human african trypanosomiasis (HAT or sleeping sickness, Frearson et al., 2010; Brand et al., 2012; Brand et al., 2014). The sulfonamide inhibitors of this thesis were derived from projects targeting *Tb*NMT, some of them even were able to cure HAT in mice (compound **1**). However, no crystal structure of *Tb*NMT is actually available. Further crystallization trials (eventually with a truncated enzyme) or the use of a homology model (Figure 7.2) may shed light on the reasons for selectivity trends in this enzyme's inhibition.

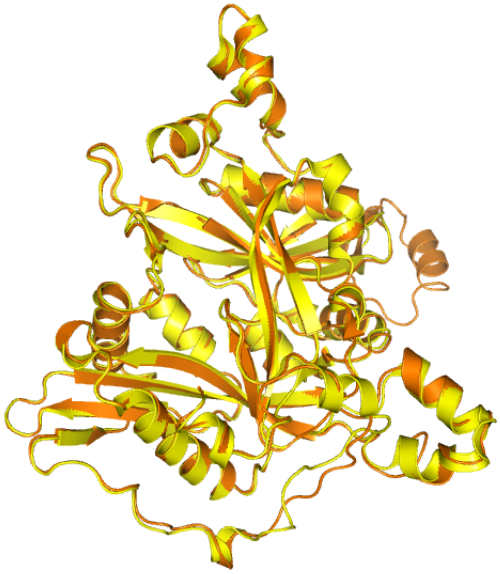


FIGURE 7.2: Homology model of *Tb*NMT (orange) in superposition with the template for model building *Lm*NMT (PDB-ID 3H5Z, yellow). Sequence identity of 54.3 % and similarity of 71.9%. Model built with MOE 2015

Moreover, completely new targets and off-targets can be investigated in a similar manner as described herein. Especially the field of parasitic diseases and cancer holds challenges due to a conserved binding sites between target and potential off-target(s). The subtle differences between those can often only be identified and exploited by the incorporation of protein flexibility and explicit binding site water analysis as it is shown within this thesis and some of the examples in Chapter 1.1 (Huggins, Sherman, and Tidor, 2012). However, a gerneralised approach is difficult because of the high individuality of each binding site. Thus special attention on details like single residues or water molecules has to be paid. Nevertheless the awareness of those details not only on selectivity, but also on affinity itself deserves more intensive implementation in standard CADD. Fortunately the continuous improvements in computation power (*Moore's law*) will allow the incorporation of those higher level of theory models in the future.

Appendix A

K_m and IC_{50} data

TABLE A.1: IC_{50} results for compounds **7-9** and **ChBr2-4**.

NMT	Ligand	IC_{50}
*Lm*NMT	**ChBr2**	33.2 ± 5.8 μM
*Lm*NMT	**ChBr3**	88.3 ± 12.4 μM
*Lm*NMT	**ChBr4**	12.4 ± 5.1 μM
*Lm*NMT	**7**	119.6 ± 88.5 nM
*Lm*NMT	**8**	41.9 ± 3.9 nM
*Lm*NMT	**9**	467.2 ± 87.2 nM
*Hs*NMT1	**7**	8 512.5 ± 5 717.3 nM
*Hs*NMT1	**8**	279.9 ± 78.1 nM
*Hs*NMT1	**9**	7 319.0 ± 3 251.8 nM
*Hs*NMT1 *8x*	**7**	121.0 ± 31.5 nM
*Hs*NMT1 *8x*	**8**	17.7 ± 5.3 nM
*Hs*NMT1 *8x*	**9**	106.9 ± 31.5 nM
*Hs*NMT1 L495M	**7**	5 588.8 ± 982.9 nM
*Hs*NMT1 L495M	**8**	128.8 ± 22.1 nM
*Hs*NMT1 L495M	**9**	1 968.5 ± 982.9 nM
*Hs*NMT1 N473H:L495M:Q496L	**7**	5 387.5 ± 3 896.9 nM
*Hs*NMT1 N473H:L495M:Q496L	**8**	109.0 ± 18.4 nM
*Hs*NMT1 N473H:L495M:Q496L	**9**	1 069.4 ± 110.8 nM

TABLE A.2: Enzyme kinetic and inhibition results of NMT and mutants. Compound **1-5** IC_{50} in nM, K_m in μM, nd = not determined

NMT	**K_m**	**1**	**2**	**3**	**4**	**5**
*Lm*NMT	5.8 ± 3.4	14.2 ± 3.3	2.4 ± 0.8	33.4 ± 6.4	3.4 ± 0.9	1 648.8 ± 259.1
*Lm*NMT H398N	3.06 ± 1.79	32.0 ± 7.4	19.7 ± 8.3	3.7 ± 1.9	3.0 ± 0.7	1 193.9 ± 465.0
*Lm*NMT M420L	nd	nd	nd	nd	nd	nd
*Lm*NMT L421Q	4.44 ± 2.14	78.6 ± 16.4	11.5 ± 6.1	4.3 ± 1.5	69.6 ± 9.5	5 014.6 ± 1 963.4
*Lm*NMT H398N:M420L:L421Q	nd	nd	nd	nd	nd	nd
*Hs*NMT1	2.64 ± 0.86	79.5 ± 11.4	33.5 ± 5.8	242.5 ± 29.7	1 077.0 ± 230.5	38 703.4 ± 10 839.2
*Hs*NMT1 L495M	1.57 ± 0.33	14.3 ± 9.8	3.3 ± 1.7	67.7 ± 10.6	63.9 ± 18.1	> 50 000
*Hs*NMT1 Q496L	2.92 ± 1.41	101.6 ± 38.9	95.2 ± 21.9	116.0 ± 47.6	669.3 ± 145.6	nd
*Hs*NMT1 N473H:L495M:Q496L	2.56 ± 0.79	51.5 ± 10.8	15.9 ± 4.9	66.2 ± 16.6	47.2 ± 9.2	37 925.0 ± 11 785.7
*Hs*NMT1 R295Q	2.92 ± 0.36	163.7 ± 76.4	nd	nd	nd	>50 000
*Hs*NMT1 R295Q:N473H:L495M:Q496L	3.85 ± 1.91	77.6 ± 34.4	nd	382.0 ± 170.3	nd	11 489.7 ± 2 236.3
*Hs*NMT1 A452M	2.22 ± 0.31	221.9 ± 88.1	27.9 ± 8.5	nd	nd	30 537.6 ± 9 332.3
*Hs*NMT1 L453V	2.09 ± 0.82	683.9 ± 571.3	nd	nd	nd	>50 000
*Hs*NMT1 A452M:L453V	7.00 ± 1.53	482.5 ± 151.7	14.2 ± 6.3	93.0 ± 73.1	736.1 ± 492.5	>50 000
*Hs*NMT1 A452M:L453V:L462V	3.59 ± 2.09	263.8 ± 85.7	21.1 ± 13.9	82.0 ± 40.7	274.6 ± 306.6	21 898.9 ± 11 691.8
*Hs*NMT1 A452M:L453V:L495M	4.38 ± 1.97	544.5 ± 278.7	0.9 ± 0.2	985.6 ± 850.9	nd	23 023.5 ± 11 678.7
*Hs*NMT1 *6x*	3.83 ± 0.85	23.1 ± 21.5	nd	468.4 ± 284.9	1 196.9 ± 852.9	17 843.3 ± 5 870.3
*Hs*NMT1 *8x*	1.50 ± 0.40	97.0 ± 8.6	16.2 ± 10.4	81.7 ± 23.8	329.4 ± 216.2	3 565.9 ± 893.5

Appendix B

ITC buffer ionization

Additional ITC experiments using altered buffer conditions with higher ionization ethalpy (HEPES instead of sodium phosphate, Table B.1, Goldberg, Kishore, and Lennen, 2002) were performed to identify protonation events during ligand binding. Affinity and thermodynamic profile show no significant difference for sulfonamide compounds **1** and **2**. For compound **3** however a decrease of the enthalpic contribution (-34.9 to -24.2 kJ/mol in *Hs*NMT1 and -55.6 to -26.5 kJ/mol in *Lm*NMT) and a similar sized increase in entropic contribution (-8.3 to 16.9 and +15.0 to -18.9 kJ/mol respectively) were observed for both the parasitic and the human NMT (Table B.2 and Figure B.1). This indicates a protonation event occuring upon protein-ligand complex formation that might be assigned to the piperazine moiety of the ligand.

TABLE B.1: Phosphate and HEPES buffer ionization reaction and standard molar enthalpy.

Buffer	**Reaction**	**pK_a**	**ΔH_{ion}**
Phosphate	$H_2PO_4^- = H^+ + HPO_4^{2-}$	7.198	3.6
HEPES	$HL = H^+ + L^-$ (HL: $C_8H_{18}N_2O_4S$)	7.564	20.4

TABLE B.2: ITC results of titrations with sodium phosphate buffer (P) and HEPES (H). ΔG, ΔH and $-T\Delta S$ in kJ/mol, K_d in nM.

Protein	**Ligand**	**Buffer**	**ΔG**	**ΔH**	**$-T\Delta S$**	**K_d**
*Hs*NMT1	1	P	-37.2 ± 0.9	-19.2 ± 4.6	-18.0 ± 4.6	248.3 ± 108.4
*Hs*NMT1	1	H	-38.5 ± 0.6	-18.2 ± 1.7	-20.3 ± 2.1	142.5 ± 41.2
*Hs*NMT1	2	P	-46.6 ± 4.3	-27.85 ± 7.9	-18.7 ± 6.8	15.4 ± 23.8
*Hs*NMT1	2	H	-46.5 ± 0.9	-25.5 ± 0.4	-20.9 ± 1.0	5.6 ± 2.2
*Hs*NMT1	3	P	-43.2 ± 0.4	-34.9 ± 1.4	-8.3 ± 1.8	20.7 ± 3.2
*Hs*NMT1	3	H	-41.1 ± 0.9	-24.2 ± 0.2	-16.9 ± 1.1	48.4 ± 15.8
*Lm*NMT	1	P	-38.7 ± 0.7	-34.2 ± 1.2	-4.5 ± 1.9	132.2 ± 40.0
*Lm*NMT	1	H	-39.7 ± 0.2	-39.8 ± 4.2	0.1 ± 4.2	87.3 ± 9.0
*Lm*NMT	3	P	-40.6 ± 3.2	-55.6 ± 4.0	15.0 ± 7.1	88.5 ± 67.6
*Lm*NMT	3	H	-45.4 ± 0.8	-26.5 ± 1.4	-18.9 ± 2.0	8.7 ± 2.7

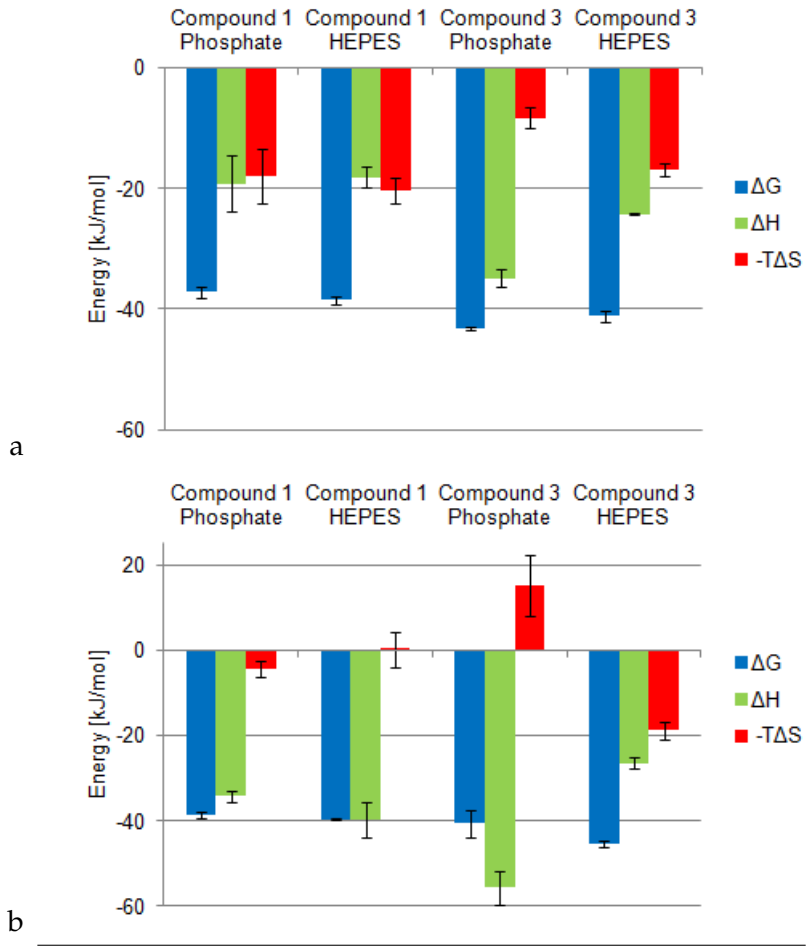


FIGURE B.1: Elucidation of the influence of buffer choice on binding thermodynamics. a) for *Hs*NMT1, b) for *Lm*NMT.

As the observed enthalpy (ΔH_{obs} consists of the binding enthalpy (ΔH_{bind}) of the ligand and n-times (n_{ion}: number of the accepted or released protons upon complex formation) the ionization enthalpy of the buffer (ΔH_{ion}, Equation B.1). The binding enthalpy was exemplarily determined from the results from two different buffer experiments (HEPES - H and sodium phosphate - P) for compound **3** binding to *Hs*NMT1 (Table B.3).

$$\Delta H_{obs} = \Delta H_{bind} + n_{ion} * \Delta H_{ion} \tag{B.1}$$

TABLE B.3: Compound 3's buffer-indipendent binding thermodynamics to *Hs*NMT1. ΔH in kJ/mol.

ΔH_{obs} (P)	ΔH_{obs} (H)	n_{ion}	ΔH_{bind}
-34.9	-24.2	0.64	-37.3

Appendix C

Crystallization - Data Collection and Refinement Statistics

TABLE C.1: Crystallographic data collection and refinement statistics.Values in parentheses indicate highest resolution shell.

Structure **Ligand(s)**	***Hs*NMT1** **MyrCoA + 2**	***Hs*NMT1** **MyrCoA + 4**
Data collection		
Space group	C121	C121
a, b, c (Å)	91.3, 58.0, 153.3	92.4, 58.17, 154.0
α, β, γ (°)	90, 92.2, 90	90, 92.3, 90
Resolution range (Å)	48.94 - 1.74 (1.80 - 1.74)	51.29 - 1.89 (1.96 - 1.89)
R_{merge}	0.889 (2.936)	0.7153 (2.507)
Mean I/I_{σ}	11.22 (1.20)	10.67 (1.46)
Completeness (%)	93 (94)	97 (99)
Refinement		
Total reflections	1288040 (130008)	1526306 (155101)
Unique reflections	76607 (7747)	64098 (6486)
R-work	0.282 (0.364)	0.231 (0.326)
R-free	0.346 (0.423)	0.277 (0.38)
No. of atoms	6552	6643
- Protein	6212	6210
- Ligands	192	192
RMS (bonds)	0.017	0.016
RMS (angles)	1.46	1.30
Ramachandran (%)		
- favoured	94	96
- allowed	5.4	4.2
- outliers	0.26	0.26
Average B-factor	23.35	26.15
- Protein	23.41	26.25
- Ligands	24.35	26.78
- Water	19.32	22.91

TABLE C.2: Crystallographic data collection and refinement statistics.Values in parentheses indicate highest resolution shell. [a] crystallized in presence of compound **5**, but with no electron density for ligand in binding site.

Structure **Ligand(s)**	***Hs*NMT1** **MyrCoA**[a]	***Hs*NMT1 *8x*** [a]	***Hs*NMT1** **MyrCoA + 7**
Data collection			
Space group	C121	P222	C121
a, b, c (Å)	87.1, 62.8,77.6	58.1, 78.1, 178.4	92.5, 58.4, 154.0
α, β, γ (°)	90, 117.8, 90	90, 90, 90	90, 92.6, 90
Resolution range (Å)	68.67 - 1.81 (1.87 - 1.81)	19.37 - 2.40 (2.49 - 2.40)	30.78 - 1.90 (1.97 - 1.90)
R_{merge}	0.652 (1.593)	0.220 (0.828)	0.765 (4.705)
Mean I/I_{σ}	25.85 (3.45)	4.83 (1.66)	24.30 (0.77)
Completeness (%)	93 (95)	96 (97)	98 (99)
Refinement			
Total reflections	1799165 (182094)	118308 (11879)	586679 (57170)
Unique reflections	31432 (3191)	31392 (3091)	64617 (6351)
R-work	0.236 (0.299)	0.415 (0.423)	0.279 (0.535)
R-free	0.280 (0.354)	0.500 (0.470)	0.358 (0.643)
No. of atoms	3369	3157	6637
- Protein	3130	3127	6272
- Ligands	63	-	180
RMS (bonds)	0.015	0.013	0.015
RMS (angles)	1.32	0.71	1.42
Ramachandran (%)			
- favoured	94	82	93
- allowed	5	13	6.8
- outliers	0.8	4.9	0.1
Average B-factor	30.47	30.93	37.89
- Protein	30.84	31.04	37.88
- Ligands	21.88	-	40.89
- Water	26.90	19.99	35.33

Appendix D

MDs: DCCMs, RMSF and S^2

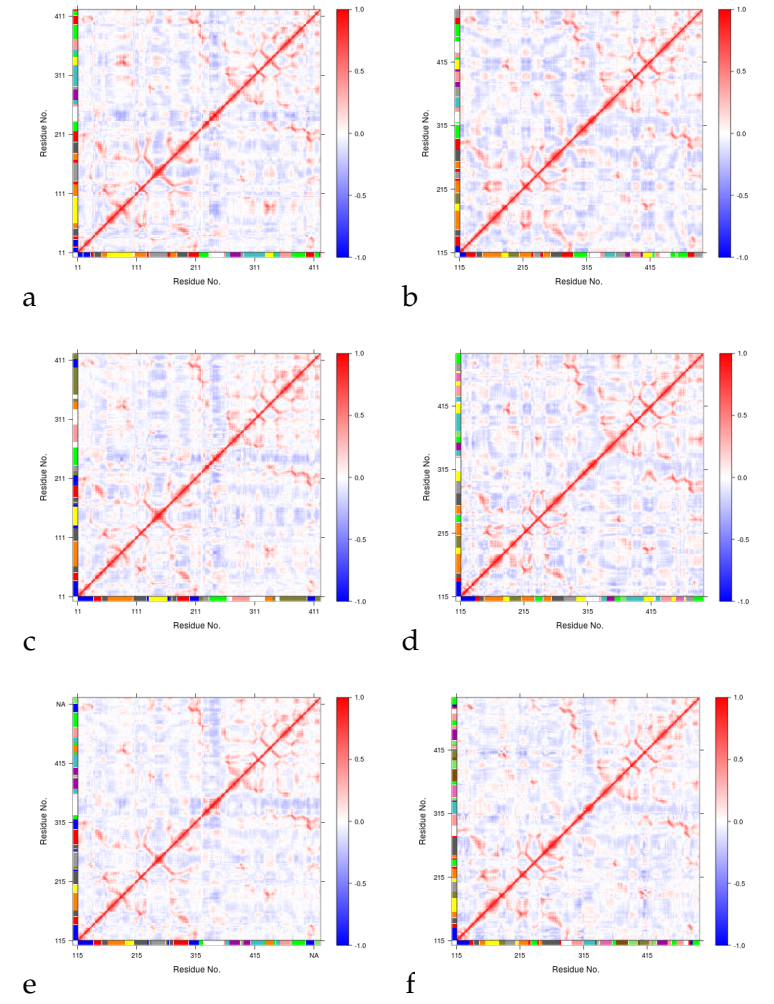


FIGURE D.1: Dynamic cross-correlation matrices (DCCM) of *Lm*NMT and *Hs*NMT1 with red-colored areas for concerted movements, white areas for independent movement and blue regions for anti-correlated movements. DCCMs revealed only minor changes for different ligands bound. a) *Lm*NMT bound to compound **1**, b) *Hs*NMT1 bound to compound **1**, c) *Lm*NMT bound to compound **4**, d) *Hs*NMT1 bound to compound **4**, e) *Lm*NMT bound to compound **5**, f) *Hs*NMT1 bound to compound **5**.

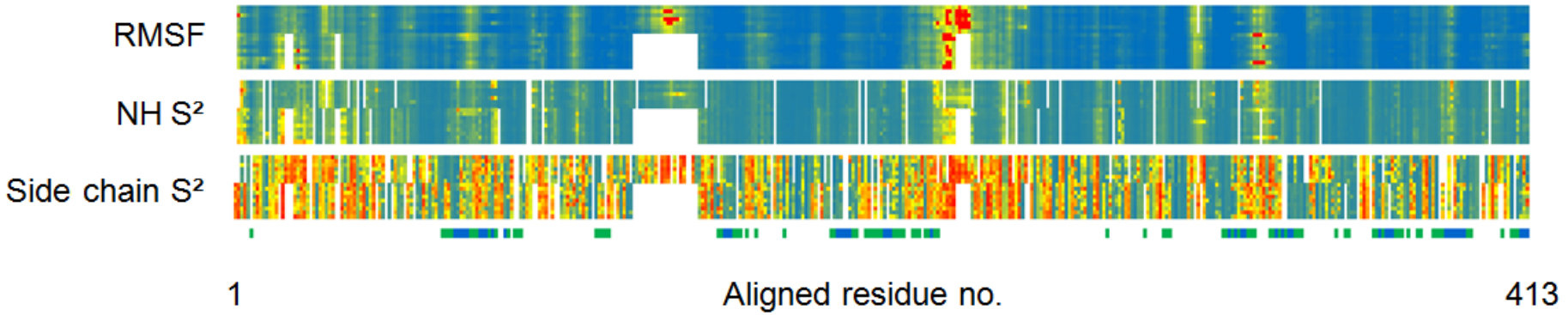


FIGURE D.2: All residue flexibility heat maps from all atom 50 ns MDs for *Lm*NMT and *Hs*NMT1 illustrating C_α RMSF, backbone S^2 and side chain S^2. Each colored datapoint represents one defined residue. Blue indicates rigid regions with low RMSF or high S^2, respectively. Red (over yellow) areas indicate more flexibility (high RMSF, low S^2). Descending order in each panel is: *Lm*NMT unbound (2 lines for 3H5Z, 4CGP), bound to compound 1, 2, 3, 4, 5 followed by *Hs*NMT1 unbound (2 lines for 3IU1, 4C2Y), bound to 1 (2 lines, 3IWE chain A and B), 2, 3, 4, 5 (2lines, docking in 3IU1 and transferred binding mode from 4CGN on 4C2Y). Blue and green color in bottom panel marks binding site (blue) and second shell (green) residues.

Appendix E

Preliminary SPR Results

TABLE E.1: Preliminary SPR results. nd = not determined

Protein	Ligand	K_d[nM]	k_{on}[$M^{-1}s^{-1}$]	k_{off} [s^{-1}]
*Hs*NMT1	1	17.6	$0.96*10^5$	$16.8*10^{-4}$
*Hs*NMT1	2	0.5	$17.6*10^5$	$7.8*10^{-4}$
*Hs*NMT1	5	38 460	nd	nd
*Lm*NMT	1	60.6	nd	nd
*Lm*NMT	2	2.1	nd	nd
*Lm*NMT	5	440	nd	nd

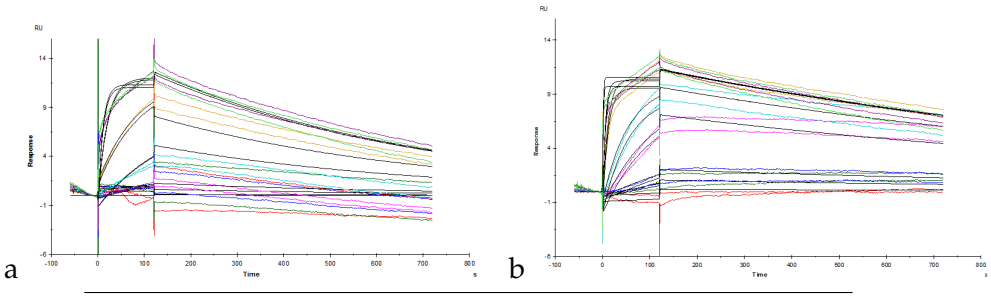


FIGURE E.1: Preliminary SPR chromatograms. a) *Hs*NMT1 - compound **1**, b) *Hs*NMT1 - compound **2**.

Bibliography

Abel, Robert, Tom Young, Ramy Farid, Bruce J Berne, and Richard A Friesner (2008). "Role of the active-site solvent in the thermodynamics of factor Xa ligand binding". In: *J. Am. Chem. Soc.* 130.9, pp. 2817–2831.

Adams, Paul D., Pavel V. Afonine, Gábor Bunkóczi, Vincent B. Chen, Ian W. Davis, Nathaniel Echols, Jeffrey J. Headd, Li Wei Hung, Gary J. Kapral, Ralf W. Grosse-Kunstleve, Airlie J. McCoy, Nigel W. Moriarty, Robert Oeffner, Randy J. Read, David C. Richardson, Jane S. Richardson, Thomas C. Terwilliger, and Peter H. Zwart (2010). "PHENIX: A comprehensive Python-based system for macromolecular structure solution". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66.2, pp. 213–221.

Alcaraz, Luis A, Lucia Banci, Ivano Bertini, Francesca Cantini, Antonio Donaire, and Leonardo Gonnelli (2007). "Matrix metalloproteinase-inhibitor interaction: The solution structure of the catalytic domain of human matrix metalloproteinase-3 with different inhibitors". In: *J. Biol. Inorg. Chem.* 12.8, pp. 1197–1206.

Aliagas-Martin, Ignacio, Dan Burdick, Laura Corson, Jennafer Dotson, Jason Drummond, Carter Fields, Oscar W Huang, Thomas Hunsaker, Tracy Kleinheinz, Elaine Krueger, Jun Liang, John Moffat, Gail Phillips, Rebecca Pulk, Thomas E Rawson, Mark Ultsch, Leslie Walker, Christian Wiesmann, Birong Zhang, Bing-Yan Zhu, and Andrea G Cochran (2009). "A class of 2,4-bisanilinopyrimidine Aurora A inhibitors with unusually high selectivity against Aurora B." In: *J. Med. Chem.* 52.10, pp. 3300–7.

Allner, Olof, Nicolas Foloppe, and Lennart Nilsson (2015). "Motions and entropies in proteins as seen in NMR relaxation experiments and molecular dynamics simulations". In: *J. Phys. Chem. B* 119.3, pp. 1114–1128.

Alonso, Hernán, Andrey A. Bliznyuk, and Jill E. Gready (2006). "Combining docking and molecular dynamic simulations in drug design". In: *Med. Res. Rev.* 26.5, pp. 531–568.

Antunes, Dinler A, Didier Devaurs, and Lydia E Kavraki (2015). "Understanding the challenges of protein flexibility in drug design." In: *Expert Opin. Drug Discov.* 10.12, pp. 1301–1313.

Baba, Yoshiyasu, Nozomu Hirukawa, Naoto Tanohira, and Mikiko Sodeoka (2003). "Structure-based design of a highly selective catalytic site-directed inhibitor of Ser/Thr protein phosphatase 2B (calcineurin)." In: *J. Am. Chem. Soc.* 125.32, pp. 9740–9.

Barillari, Caterina, Anna L Duncan, Isaac M Westwood, Julian Blagg, and Rob L M van Montfort (2011). "Analysis of water patterns in protein kinase binding sites." In: *Proteins* 79.7, pp. 2109–21.

Battye, T. Geoff G, Luke Kontogiannis, Owen Johnson, Harold R. Powell, and Andrew G W Leslie (2011). "iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67.4, pp. 271–281.

Bayly, CI, P Cieplak, W Cornell, and PA Kollman (1993). "A well-behaved electrostatic potential based method using charge restraints for deriving atomic . . . " In: *J. Phys. Chem.* Pp. 10269–10280.

Bell, Andrew S, James E Mills, Gareth P Williams, James a Brannigan, Anthony J Wilkinson, Tanya Parkinson, Robin J Leatherbarrow, Edward W Tate, Anthony a Holder, and Deborah F Smith (2012). "Selective inhibitors of protozoan protein N-myristoyltransferases as starting points for tropical disease medicinal chemistry programs." In: *PLoS Negl. Trop. Dis.* 6.4, e1625.

Bergner, Andreas, Judith Günther, Manfred Hendlich, Gerhard Klebe, and Marcel Verdonk (2001). "Use of relibase for retrieving complex three-dimensional interaction patterns including crystallographic packing effects". In: *Biopolymers* 61.2, pp. 99–110.

Berman, H M, J Westbrook, Z Feng, G Gilliland, T N Bhat, H Weissig, I N Shindyalov, and P E Bourne (2000). "The protein data bank." In: *Nucleic Acids Res.* 28.1, pp. 235–242.

Bernetti, M., A. Cavalli, and L. Mollica (2017). "Protein–ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling". In: *Med. Chem. Commun.* 8, pp. 534–550.

Berthold, Michael R., Nicolas Cebron, Fabian Dill, Thomas R. Gabriel, Tobias Kötter, Thorsten Meinl, Peter Ohl, Christoph Sieb, Kilian Thiel, and Bernd Wiswedel (2007). "KNIME: The Konstanz Information Miner". In: *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*. Springer.

Betz, Michael, Tobias Wulsdorf, Stefan G. Krimmer, and Gerhard Klebe (2016). "Impact of Surface Water Layers on Protein-Ligand Binding: How Well Are Experimental Data Reproduced by Molecular Dynamics Simulations in a Thermolysin Test Case?" In: *J. Chem. Inf. Model.* 56.1, pp. 223–233.

Bhatnagar, Rajiv S, Klaus Fütterer, Gabriel Waksman, and Jeffrey I Gordon (1999). *The structure of myristoyl-CoA:protein N-myristoyltransferase*.

Biava, Mariangela, Giulio C. Porretta, Giovanna Poce, Claudio Battilocchio, Fabrizio Manetti, Maurizio Botta, Stefano Forli, Lidia Sautebin, Antonietta Rossi, Carlo Pergola, Carla Ghelardini, Nicoletta Galeotti,

Francesco Makovec, Antonio Giordani, Paola Anzellotti, Paola Patrignani, and Maurizio Anzini (2010). "Novel Ester and acid derivatives of the 1,5-diarylpyrrole Scaffold as anti-inflammatory and analgesic agents. Synthesis and in vitro and in vivo biological evaluation". In: *J. Med. Chem.* 53.2, pp. 723–733.

Biela, Adam, Michael Betz, Andreas Heine, and Gerhard Klebe (2012). "Water makes the difference: rearrangement of water solvation layer triggers non-additivity of functional group contributions in protein-ligand binding." In: *ChemMedChem* 7.8, pp. 1423–34.

Biela, Adam, Nader N Nasief, Michael Betz, Andreas Heine, David Hangauer, and Gerhard Klebe (2013). "Dissecting the hydrophobic effect on the molecular level: the role of water, enthalpy, and entropy in ligand binding to thermolysin." In: *Angew. Chem. Int. Ed. Engl.* 52.6, pp. 1822–8.

Bissantz, Caterina, Bernd Kuhn, and Martin Stahl (2010). "A medicinal chemist's guide to molecular interactions". In: *J. Med. Chem.* 53.14, pp. 5061–5084.

Bologna, Guido, Cédric Yvon, Séverine Duvaud, and Anne Lise Veuthey (2004). "N-terminal myristoylation predictions by ensembles of neural networks". In: *Proteomics* 4.6, pp. 1626–1632.

Bowyer, Paul W, Edward W Tate, Robin J Leatherbarrow, Anthony a Holder, Deborah F Smith, and Katherine a Brown (2008). "N-myristoyltransferase: a prospective drug target for protozoan parasites." In: *ChemMedChem* 3.3, pp. 402–8.

Brand, Stephen, Laura a T Cleghorn, Stuart P McElroy, David a Robinson, Victoria C Smith, Irene Hallyburton, Justin R Harrison, Neil R Norcross, Daniel Spinks, Tracy Bayliss, Suzanne Norval, Laste Stojanovski, Leah S Torrie, Julie a Frearson, Ruth Brenk, Alan H Fairlamb, Michael a J Ferguson, Kevin D Read, Paul G Wyatt, and Ian H Gilbert (2012). "Discovery of a novel class of orally active trypanocidal N-myristoyltransferase inhibitors." In: *J. Med. Chem.* 55.1, pp. 140–52.

Brand, Stephen, Neil R. Norcross, Stephen Thompson, Justin R. Harrison, Victoria C. Smith, David A. Robinson, Leah S. Torrie, Stuart P. McElroy, Irene Hallyburton, Suzanne Norval, Paul Scullion, Laste Stojanovski, Frederick R C Simeons, Daan Van Aalten, Julie A. Frearson, Ruth Brenk, Alan H. Fairlamb, Michael A J Ferguson, Paul G. Wyatt, Ian H. Gilbert, and Kevin D. Read (2014). "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". In: *J. Med. Chem.* 57.23, pp. 9855–9869.

Brannigan, James A, Barbara A Smith, Zhiyong Yu, Andrzej M Brzozowski, Michael R Hodgkinson, Asher Maroof, Helen P Price, Franziska Meier,

Robin J Leatherbarrow, Edward W Tate, Deborah F Smith, and Anthony J Wilkinson (2010). "N-myristoyltransferase from Leishmania donovani: structural and functional characterisation of a potential drug target for visceral leishmaniasis." In: *J. Mol. Biol.* 396.4, pp. 985–99.

Brannigan, James A., Shirley M. Roberts, Andrew S. Bell, Jennie A. Hutton, Michael R. Hodgkinson, Edward W. Tate, Robin J. Leatherbarrow, Deborah F. Smith, and Anthony J. Wilkinson (2014). "Diverse modes of binding in structures of Leishmania major N-myristoyltransferase with selective inhibitors". In: *IUCrJ* 1, pp. 250–260.

Breiten, Benjamin, Matthew R. Lockett, Woody Sherman, Shuji Fujita, Mohammad Al-Sayah, Heiko Lange, Carleen M. Bowers, Annie Heroux, Goran Krilov, and George M. Whitesides (2013). "Water networks contribute to enthalpy/entropy compensation in protein-ligand binding". In: *J. Am. Chem. Soc.* 135.41, pp. 15579–15584.

Brenk, Ruth, Milton T. Stubbs, Andreas Heine, Klaus Reuter, and Gerhard Klebe (2003). "Flexible adaptations in the structure of the tRNA-modifying enzyme tRNA-guanine transglycosylase and their implications for substrate selectivity, reaction mechanism and structure-based drug design". In: *ChemBioChem* 4.10, pp. 1066–1077.

Cappel, Daniel, Rickard Wahlström, Ruth Brenk, and Christoph A. Sotriffer (2011). "Probing the dynamic nature of water molecules and their influences on ligand binding in a model binding site". In: *J. Chem. Inf. Model.* 51.10, pp. 2581–2594.

Carlson, H and J McCammon (2000). "Accommodating protein flexibility in computational drug design." In: *Mol. Pharmacol.* 57.2, pp. 213–218.

Cavalli, Andrea, Elisabetta Poluzzi, Fabrizio De Ponti, and Maurizio Recanatini (2002). "Toward a pharmacophore for drugs inducing the long QT syndrome: Insights from a CoMFA study of HERG K+ channel blockers". In: *J. Med. Chem.* 45.18, pp. 3844–3853.

Cheng, Tiejun, Qingliang Li, Zhigang Zhou, Yanli Wang, and Stephen H. Bryant (2012). "Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review". In: *AAPS J.* 14.1, pp. 133–141.

Cheng, Y and WH Prusoff (1973). "Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction." In: *Biochem. Pharmacol.* 22, pp. 3099–3108.

Clarkson, MW (2007). *Conformational Flux; http://conflux.mwclarkson.com/2007/11/the-model-free-dynamics-formalism-of-lipari-and-szabo/; accessed 2017-04-11*.

Cleghorn, Laura A T, Andrew Woodland, Iain T Collie, Leah S Torrie, Neil Norcross, Torsten Luksch, Chido Mpamhanga, Roderick G Walker, Jeremy C Mottram, Ruth Brenk, Julie A Frearson, Ian H Gilbert, and

Paul G Wyatt (2011). "Identification of inhibitors of the Leishmania cdc2-related protein kinase CRK3." In: *ChemMedChem* 6.12, pp. 2214–24.

Cole, Jason C., Christopher W. Murray, J. Willem M Nissink, Richard D. Taylor, and Robin Taylor (2005). "Comparing protein-ligand docking programs is difficult". In: *Proteins Struct. Funct. Genet.* 60.3, pp. 325–332.

Cornell, Wendy D, Piotr Cieplak, Christopher I Bayly, Ian R Gould, Kenneth M Merz, David M Ferguson, David C Spellmeyer, Thomas Fox, James W Caldwell, and Peter A Kollman (1995). "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules". In: *J. Am. Chem. Soc.* 117.19, pp. 5179–5197.

Cui, Guanglei, Jason M Swails, and Eric S Manas (2013). "SPAM: A simple approach for profiling bound water molecules". In: *J. Chem. Theory Comput.* 9.12, pp. 5539–5549.

Darden, Tom, Darrin York, and Lee Pedersen (1993). "Particle mesh Ewald: An N*log(N) method for Ewald sums in large systems". In: *J. Chem. Phys.* 98.12, p. 10089.

De Graaf, Chris, Nico P. E. Vermeulen, and K. Anton Feenstra (2005). *Cytochrome P450 in silico: An integrative modeling approach*.

De Vivo, Marco, Matteo Masetti, Giovanni Bottegoni, and Andrea Cavalli (2016). "Role of Molecular Dynamics and Related Methods in Drug Discovery". In: *J. Med. Chem.* 59.9, pp. 4035–4061.

Denisov, Vladimir P., Kandadai Venu, Jörg Peters, Hans Dietrich Hörlein, and Bertil Halle (1997). "Orientational disorder and entropy of water in protein cavities". In: *J. Phys. Chem. B* 101, pp. 9380–9389.

Devadas, Balekudru and ME Zupec (1995). "Design and syntheses of potent and selective dipeptide inhibitors of Candida albicans myristoyl-CoA: protein N-myristoyltransferase". In: *J. Med. ...* 38.11, pp. 1837–1840.

Dhulesia, Anne, Joerg Gsponer, and Michele Vendruscolo (2008). "Mapping of two networks of residues that exhibit structural and dynamical changes upon binding in a PDZ domain protein." In: *J. Am. Chem. Soc.* 130.28, pp. 8931–9.

Diehl, Carl, Olof Engström, Tamara Delaine, Maria Håkansson, Samuel Genheden, Kristofer Modig, Hakon Leffler, Ulf Ryde, Ulf J. Nilsson, and Mikael Akke (2010). "Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3". In: *J. Am. Chem. Soc.* 132.41, pp. 14577–14589.

Duan, James J-W, Lihua Chen, Zelda R Wasserman, Zhonghui Lu, Rui-Qin Liu, Maryanne B Covington, Mingxin Qian, Karl D Hardman, Ronald L Magolda, Robert C Newton, David D Christ, Ruth R Wexler, and Carl P Decicco (2002). "Discovery of gamma-lactam hydroxamic acids as selective inhibitors of tumor necrosis factor alpha converting enzyme:

design, synthesis, and structure-activity relationships." In: *J. Med. Chem.* 45.23, pp. 4954–7.

Ducker, Charles E., John J Upson, Kevin J French, and Charles D Smith (2005). "Two N-myristoyltransferase isozymes play unique roles in protein myristoylation, proliferation, and apoptosis." In: *Mol. Cancer Res.* 3.August, pp. 463–476.

Durrant, Jacob D and J Andrew McCammon (2011). "Molecular dynamics simulations and drug discovery." In: *BMC Biol.* 9, p. 71.

Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan (2010). "Features and development of Coot". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66.4, pp. 486–501.

Fang, Wenxia, David a. Robinson, Olawale G. Raimi, David E. Blair, Justin R. Harrison, Deborah E. a. Lockhart, Leah S. Torrie, Gian Filippo Ruda, Paul G. Wyatt, Ian H. Gilbert, and Daan M. F. van Aalten (2015). "N -Myristoyltransferase Is a Cell Wall Target in Aspergillus fumigatus". In: *ACS Chem. Biol.* 10.6, pp. 1425–1434.

Farazi, T. A., G. Waksman, and J. I. Gordon (2001). "The Biology and Enzymology of ProteinN-Myristoylation". In: *J. Biol. Chem.* 276.43, pp. 39501–39504.

Farid, Ramy, Tyler Day, Richard A. Friesner, and Robert A. Pearlstein (2006). "New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies". In: *Bioorganic Med. Chem.* 14.9, pp. 3160–3173.

Feller, Scott E., Yuhong Zhang, Richard W. Pastor, and Bernard R. Brooks (1995). "Constant pressure molecular dynamics simulation: The Langevin piston method". In: *J. Chem. Phys.* 103.11, pp. 4613–4621.

Fenley, Andrew T, Hari S Muddana, and Michael K Gilson (2012). "Entropy-enthalpy transduction caused by conformational shifts can obscure the forces driving protein-ligand binding." In: *Proc. Natl. Acad. Sci. U. S. A.* 109.49, pp. 20006–11.

Fischer, E. (1894). "Einfluss der Configuration auf die Wirkung der Enzyme". In: *Ber. Dtsch. Chem. Ges.* 27, pp. 2985–2993.

Fitzgerald, Catherine E, Sangita B Patel, Joseph W Becker, Patricia M Cameron, Dennis Zaller, Vasilis Bill Pikounis, Stephen J O'Keefe, and Giovanna Scapin (2003). "Structural basis for p38alpha MAP kinase quinazolinone and pyridol-pyrimidine inhibitor specificity." In: *Nat. Struct. Biol.* 10.9, pp. 764–9.

Frearson, Julie A., Paul G. Wyatt, Ian H. Gilbert, and Alan H. Fairlamb (2007). "Target assessment for antiparasitic drug discovery". In: *Trends Parasitol.* 23.12, pp. 589–595.

Frearson, Julie A, Stephen Brand, Stuart P Mcelroy, Laura A T Cleghorn, Ondrej Smid, Laste Stojanovski, Helen P Price, M Lucia S Guther, Leah S Torrie, David A Robinson, Irene Hallyburton, Chidochangu P Mpamhanga, James A Brannigan, Anthony J Wilkinson, Michael Hodgkinson, Raymond Hui, Wei Qiu, Olawale G Raimi, Daan M F van Aalten, Ruth Brenk, Ian H Gilbert, Kevin D Read, Alan H Fairlamb, Michael A J Ferguson, Deborah F Smith, and Paul G Wyatt (2010). "N-myristoyltransferase inhibitors as new leads to treat sleeping sickness." In: *Nature* 464.7289, pp. 728–32.

Frederick, Kendra King, Michael S Marlow, Kathleen G Valentine, and A Joshua Wand (2007). "Conformational entropy in molecular recognition by proteins." In: *Nature* 448.7151, pp. 325–9.

Fricker, Patrick C., Marcus Gastreich, and Matthias Rarey (2004). "Automated drawing of structural molecular formulas under constraints". In: *J. Chem. Inf. Comput. Sci.* 44.3, pp. 1065–1078.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, and 2016 Gaussian Inc. Wallingford CT (2009). *Gaussian 09 Revision C.01*. Gaussian Inc. Wallingford CT 2009.

Galvin, Brendan D., Zhiru Li, Estelle Villemaine, Catherine B. Poole, Melissa S. Chapman, Michael P. Pollastri, Paul G. Wyatt, and Clotilde K S Carlow (2014). "A Target Repurposing Approach Identifies N-myristoyltransferase as a New Candidate Drug Target in Filarial Nematodes". In: *PLoS Negl. Trop. Dis.* 8.9.

Genheden, Samuel, Mikael Akke, and Ulf Ryde (2014). "Conformational entropies and order parameters: Convergence, reproducibility, and transferability". In: *J. Chem. Theory Comput.* 10.1, pp. 432–438.

Gohlke, Holger, Manfred Hendlich, and Gerhard Klebe (2000). "Knowledge-based scoring function to predict protein-ligand interactions". In: *J. Mol. Biol.* 295.2, pp. 337–356.

Goldberg, Robert N., Nand Kishore, and Rebecca M. Lennen (2002). "Thermodynamic quantities for the ionization reaction of buffers." In: *J. Phys. Chem. Ref. Data* 31.2, pp. 231–370.

Goldston, Amanda M., Aabha I. Sharma, Kimberly S. Paul, and David M. Engman (2014). "Acylation in trypanosomatids: An essential process and potential drug target". In: *Trends Parasitol.* 30.7, pp. 350–360.

Goncalves, Victor, James a Brannigan, Emmanuelle Thinon, Tayo O Olaleye, Remigiusz Serwa, Salvatore Lanzarone, Anthony J Wilkinson, Edward W Tate, and Robin J Leatherbarrow (2012a). "A fluorescence-based assay for N-myristoyltransferase activity." In: *Anal. Biochem.* 421.1, pp. 342–4.

Goncalves, Victor, James a Brannigan, David Whalley, Keith H Ansell, Barbara Saxty, Anthony a Holder, Anthony J Wilkinson, Edward W Tate, and Robin J Leatherbarrow (2012b). "Discovery of Plasmodium vivax N-myristoyltransferase inhibitors: screening, synthesis, and structural characterization of their binding mode." In: *J. Med. Chem.* 55.7, pp. 3578–82.

Goncalves, Victor, James A. Brannigan, Alice Laporte, Andrew S. Bell, Shirley M. Roberts, Anthony J. Wilkinson, Robin J. Leatherbarrow, and Edward W. Tate (2017). "Structure-guided optimization of quinoline inhibitors of Plasmodium N-myristoyltransferase". In: *Med. Chem. Commun.* 8.1, pp. 191–197.

Goodford, P J (1985). "A computational procedure for determining energetically favorable binding sites on biologically important macromolecules". In: *J. Med. Chem.* 28.7, pp. 849–857.

Gouda, Hiroaki, Irwin D. Kuntz, David A. Case, and Peter A. Kollman (2003). "Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods". In: *Biopolymers* 68.1, pp. 16–34.

Grant, B. J., A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. D. Caves (2006). "Bio3d: an R package for the comparative analysis of protein structures". In: *Bioinformatics* 22.21, pp. 2695–2696.

Gu, Yina, Da Wei Li, and Rafael Brüschweiler (2014). "NMR order parameter determination from long molecular dynamics trajectories for objective comparison with experiment". In: *J. Chem. Theory Comput.* 10.6, pp. 2599–2607.

Günther, Judith, Andreas Bergner, Manfred Hendlich, and Gerhard Klebe (2003). "Utilising Structural Knowledge in Drug Design Strategies: Applications Using Relibase". In: *J. Mol. Biol.* 326.2, pp. 621–636.

Halgren, Thomas A (1996). "Merck Molecular Force Field". In: *J. Comput. Chem.* 17.5 & 6, pp. 490–519.

Hawkins, Paul C. D., A. Geoffrey Skillman, Gregory L. Warren, Benjamin A. Ellingson, and Matthew T. Stahl (2010). "Conformer generation with

OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database". In: *J. Chem. Inf. Model.* 50.4, pp. 572–584.

Hendlich, Manfred, Andreas Bergner, Judith Günther, and Gerhard Klebe (2003). "Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions." In: *J. Mol. Biol.* 326.2, pp. 607–20.

Herrera, Linda J., Stephen Brand, Andres Santos, Lilian L. Nohara, Justin Harrison, Neil R. Norcross, Stephen Thompson, Victoria Smith, Carolina Lema, Armando Varela-Ramirez, Ian H. Gilbert, Igor C. Almeida, and Rosa A. Maldonado (2016). "Validation of N-myristoyltransferase as Potential Chemotherapeutic Target in Mammal-Dwelling Stages of Trypanosoma cruzi". In: *PLoS Negl. Trop. Dis.* 10.4, pp. 1–20.

Hopkins, Andrew L., Colin R. Groom, and Alexander Alex (2004). "Ligand efficiency: a useful metric for lead selection". In: *Drug. Discov. Today* 9.10, pp. 430 –431.

Huang, Jing and Alexander D. Mackerell (2013). "CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data". In: *J. Comput. Chem.* 34.25, pp. 2135–2145.

Huggins, David J, Woody Sherman, and Bruce Tidor (2012). "Rational approaches to improving selectivity in drug design." In: *J. Med. Chem.* 55.4, pp. 1424–44.

Humphrey, William, Andrew Dalke, and Klaus Schulten (1996). "Visual molecular dynamics". In: *J. Mol. Graph.* 14.1, pp. 33–38.

Hutton, Jennie A., Victor Goncalves, James A. Brannigan, Daniel Paape, Megan H. Wright, Thomas M. Waugh, Shirley M. Roberts, Andrew S. Bell, Anthony J. Wilkinson, Deborah F. Smith, Robin J. Leatherbarrow, and Edward W. Tate (2014). "Structure-Based Design of Potent and Selective Leishmania N-Myristoyltransferase Inhibitors". In: *J. Med. Chem.* 57.20, pp. 8664–70.

Hyeon, Changbong, Patricia A Jennings, Joseph A Adams, and José N Onuchic (2009). "Ligand-induced global transitions in the catalytic domain of protein kinase A." In: *Proc. Natl. Acad. Sci. U. S. A.* 106.9, pp. 3023–8.

Ikuta, Mari, Kenji Kamata, Kazuhiro Fukasawa, Teruki Honma, Takumitsu Machida, Hiroshi Hirai, Ikuko Suzuki-Takahashi, Takashi Hayama, and Susumu Nishimura (2001). "Crystallographic Approach to Identification of Cyclin-dependent Kinase 4 (CDK4)-specific Inhibitors by Using CDK4 Mimic CDK2 Protein". In: *J. Biol. Chem.* 276.29, pp. 27548–27554.

Irwin, John J., Teague Sterling, Michael M. Mysinger, Erin S. Bolstad, and Ryan G. Coleman (2012). "ZINC: A free tool to discover chemistry for biology". In: *J. Chem. Inf. Model.* 52.7, pp. 1757–1768.

Iversen, L F, H S Andersen, S Branner, S B Mortensen, G H Peters, K Norris, O H Olsen, C B Jeppesen, B F Lundt, W Ripka, K B Møller, and N P Møller (2000). "Structure-based design of a low molecular weight, nonphosphorus, nonpeptide, and highly selective inhibitor of protein-tyrosine phosphatase 1B." In: *J Biol Chem* 275.14, pp. 10300–10307.

Janowski, Pawel A., David S. Cerutti, James Holton, and David A. Case (2013). "Peptide crystal simulations reveal hidden dynamics". In: *J. Am. Chem. Soc.* 135.21, pp. 7938–7948.

Jorgensen, William L., Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein (1983). "Comparison of simple potential functions for simulating liquid water". In: *J. Chem. Phys.* 79.2, p. 926.

Kabsch, Wolfgang (2010). "XDS". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66.2, pp. 125–132.

Kangas, Erik and Bruce Tidor (1998). "Optimizing electrostatic affinity in ligand-receptor binding: Theory, computation, and ligand properties". In: *J. Chem. Phys.* 109.17, pp. 7522–7545.

Karplus, M and J Kuriyan (2005). "Molecular dynamics and protein function". In: *Proc. Natl. Acad. Sci. U. S. A.* 102.19, pp. 6679–6685.

Karplus, Martin and Ja Andrew McCammon (2002). "Molecular dynamics simulations of biomolecules." In: *Nat. Struct. Biol.* 9.9, pp. 646–652.

Kitchen, DB, H Decornez, JR Furr, and J Bajorath (2004). "Docking and scoring in virtual screening for drug discovery: methods and applications." In: *Nat Rev Drug Discov* 3.11, pp. 935–949.

Klebe, Gerhard (2006). "Virtual ligand screening: strategies, perspectives and limitations". In: *Drug Discov. Today* 11.13-14, pp. 580–594.

— (2009). *Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen (2. Auflage)*. Spektrum Verlag, Heidelberg.

Knight, Zachary A. and Kevan M. Shokat (2005). *Features of selective kinase inhibitors*.

Koshland, D E (1958). "Application of a Theory of Enzyme Specificity to Protein Synthesis". In: *Proc. Natl. Acad. Sci.* 44.2, pp. 98–104.

Koukos, Panagiotis I. and Nicholas M. Glykos (2014). "On the application of good-turing statistics to quantify convergence of biomolecular simulations". In: *J. Chem. Inf. Model.* 54.1, pp. 209–217.

Kovalevsky, Andrey Yu, Yunfeng Tie, Fengling Liu, Peter I. Boross, Yuan Fang Wang, Sofiya Leshchenko, Arun K. Ghosh, Robert W. Harrison, and Irene T. Weber (2006). "Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M". In: *J. Med. Chem.* 49.4, pp. 1379–1387.

Kumar, S, B Ma, C J Tsai, N Sinha, and R Nussinov (2000). "Folding and binding cascades: dynamic landscapes and population shifts." In: *Protein Sci.* 9.1, pp. 10–9.

Kurumbail, Ravi G., Anna M. Stevens, James K. Gierse, Joseph J. McDonald, Roderick A. Stegeman, Jina Y. Pak, Daniel Gildehaus, Julie M. Iyashiro, Thomas D. Penning, Karen Seibert, Peter C. Isakson, and William C. Stallings (1996). "Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents". In: *Nature* 384.6610, pp. 644–648.

Ladbury, John E. (1996). "Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design". In: *Chem. Biol.* 3.12, pp. 973–980.

Lazaridis, Themis (1998). "Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory". In: *J. Phys. Chem. B* 102.18, pp. 3531–3541.

Leach, A. R. (2001). *Molecular Modelling. Principles and Applications (2nd Edition)*. Pearson Education Limited, Essex.

Leach, Andrew R, Brian K Shoichet, and Catherine E Peishoff (2006). "Prediction of protein-ligand interactions. Docking and scoring: Successes and gaps". In: *J. Med. Chem.* 49.20, pp. 5851–5855.

Leatherbarrow, R.J. (2009). *GraFit Version 7, Erithacus Software Ltd., Horley, U.K.*

Lee, L P and Bruce Tidor (1997). "Optimization of electrostatic binding free energy". In: *J. Chem. Phys.* 106.June 1997, pp. 8681–8690.

Leioatts, Nicholas, Pooja Suresh, Tod D. Romo, and Alan Grossfield (2014). "Structure-based simulations reveal concerted dynamics of GPCR activation". In: *Proteins Struct. Funct. Bioinforma.* 82.10, pp. 2538–2551.

Leslie, Andrew G W (2006). "The integration of macromolecular diffraction data". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62.1, pp. 48–57.

Levinson, Nicholas M and Steven G Boxer (2013). "A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity". In: *Nat. Chem. Biol.* 10.2, pp. 127–132.

Li, Da Wei and Rafael Brüschweiler (2009). "A dictionary for protein side-chain entropies from NMR order parameters". In: *J. Am. Chem. Soc.* 131.21, pp. 7226–7227.

Li, Weizhong, Andrew Cowley, Mahmut Uludag, Tamer Gur, Hamish McWilliam, Silvano Squizzato, Young Mi Park, Nicola Buso, and Rodrigo Lopez (2015). "The EMBL-EBI bioinformatics web and programmatic tools framework". In: *Nucleic Acids Res.* 43.W1, W580–W584.

Li, Zheng and Themis Lazaridis (2003). "Thermodynamic contributions of the ordered water molecule in HIV-1 protease". In: *J. Am. Chem. Soc.* 125.22, pp. 6636–6637.

Lipari, Giovanni and Attila Szabo (1982). "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity". In: *J. Am. Chem. Soc.* 104.17, pp. 4546–4559.

MacKerell, A D, D Bashford, M Bellott, R L Dunbrack, J D Evanseck, M J Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K

Kuczera, F T Lau, C Mattos, S Michnick, T Ngo, D T Nguyen, B Prodhom, W E Reiher, B Roux, M Schlenkrich, J C Smith, R Stote, J Straub, M Watanabe, J Wiórkiewicz-Kuczera, D Yin, and M Karplus (1998). "All-atom empirical potential for molecular modeling and dynamics studies of proteins." In: *J. Phys. Chem. B* 102.18, pp. 3586–616.

Mackerell, Alexander D., Michael Feig, and Charles L. Brooks (2004). "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation". In: *J. Comput. Chem.* 25.11, pp. 1400–1415.

MacRaild, Christopher A, Antonio Hernández Daranas, Agnieszka Bronowska, and Steve W Homans (2007). "Global changes in local protein dynamics reduce the entropic cost of carbohydrate binding in the arabinose-binding protein." In: *J. Mol. Biol.* 368.3, pp. 822–32.

Manas, Eric S, Rayomand J Unwalla, Zhang B Xu, Michael S Malamas, Chris P Miller, Heather A Harris, Chulai Hsiao, Tatos Akopian, Wah-Tung Hum, Karl Malakian, Scott Wolfrom, Ashok Bapat, Ramesh A Bhat, Mark L Stahl, William S Somers, and Juan C Alvarez (2004). "Structure-based design of estrogen receptor-beta selective ligands." In: *J. Am. Chem. Soc.* 126.46, pp. 15106–19.

Marlow, MS Michael S. MS Michael S., Jakob Dogan, Kendra K. Frederick, Kathleen G. Valentine, and A. Joshua Wand (2010). "The role of conformational entropy in molecular recognition by calmodulin". In: *Nat. Chem. ...* 6.5, pp. 352–358.

Martin, Dale D O, Erwan Beauchamp, and Luc G. Berthiaume (2011). "Post-translational myristoylation: Fat matters in cellular life and death". In: *Biochimie* 93.1, pp. 18–31.

Martyna, Glenn J, Douglas J Tobias, and Michael L Klein (1994). "Constant pressure molecular dynamics algorithms". In: 101.September, pp. 4177–4189.

Matthews, Brian W and Lijun Liu (2009). "A review about nothing: Are apolar cavities in proteins really empty?" In: *Protein Sci.* 18.3, pp. 494–502.

Mauldin, Randall V., Mary J. Carroll, and Andrew L. Lee (2009). "Dynamic Dysfunction in Dihydrofolate Reductase Results from Antifolate Drug Binding: Modulation of Dynamics within a Structural State". In: *Structure* 17.3, pp. 386–394.

Maurer-Stroh, Sebastian, Birgit Eisenhaber, and Frank Eisenhaber (2002). "N-terminal N-myristoylation of proteins: prediction of substrate proteins from amino acid sequence." In: *J. Mol. Biol.* 317.4, pp. 541–557.

Mayne, Christopher G., Jan Saam, Klaus Schulten, Emad Tajkhorshid, and James C. Gumbart (2013). "Rapid parameterization of small molecules using the force field toolkit". In: *J. Comput. Chem.* 34.32, pp. 2757–2770.

McCoy, Airlie J., Ralf W. Grosse-Kunstleve, Paul D. Adams, Martyn D. Winn, Laurent C. Storoni, and Randy J. Read (2007). "Phaser crystallographic software". In: *J. Appl. Crystallogr.* 40.4, pp. 658–674.

McWilliam, Hamish, Weizhong Li, Mahmut Uludag, Silvano Squizzato, Young Mi Park, Nicola Buso, Andrew Peter Cowley, and Rodrigo Lopez (2013). "Analysis Tool Web Services from the EMBL-EBI." In: *Nucleic Acids Res.* 41.Web Server issue, pp. 597–600.

Meinhold, Lars and Jeremy C. Smith (2005). "Fluctuations and Correlations in Crystalline Protein Dynamics: A Simulation Analysis of Staphylococcal Nuclease". In: *Biophys. J.* 88.4, pp. 2554–2563.

Michel, Julien, Julian Tirado-Rives, and William L. Jorgensen (2009a). "Energetics of displacing water molecules from protein binding sites: Consequences for ligand optimization". In: *J. Am. Chem. Soc.* 131.42, pp. 15403–15411.

— (2009b). "Prediction of the water content in protein binding sites". In: *J. Phys. Chem. B* 113.40, pp. 13337–13346.

Miyamoto, Shuichi and Peter A. Kollman (1993). "Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches". In: *Proteins Struct. Funct. Bioinforma.* 16.3, pp. 226–245.

Moriarty, Nigel W., Ralf W. Grosse-Kunstleve, and Paul D. Adams (2009). "Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 65.10, pp. 1074–1080.

Nayal, M and E Di Cera (1996). "Valence screening of water in protein crystals reveals potential Na+ binding sites." In: *J. Mol. Biol.* 256.2, pp. 228–234.

Ni, H., C. A. Sotriffer, and J. Andrew McCammon (2001). "Ordered water and ligand mobility in the HIV-1 integrase-5CITEP complex: A molecular dynamics study". In: *J. Med. Chem.* 44.19, pp. 3043–3047.

Nilapwar, Sanjay, Eleanor Williams, Christopher Fu, Christosmos Prodromou, Laurence H Pearl, Mark A Williams, and John E Ladbury (2009). "Structural-thermodynamic relationships of interactions in the N-terminal ATP-binding domain of Hsp90." In: *J. Mol. Biol.* 392.4, pp. 923–36.

Niu, Xiaoda, Shelby Umland, Richard Ingram, Brian M Beyer, Yan Hui Liu, Jing Sun, Daniel Lundell, and Peter Orth (2006). "IK682, a tight binding inhibitor of TACE". In: *Arch. Biochem. Biophys.* 451.1, pp. 43–50.

Noble, Martin E M, Jane A Endicott, and Louise N Johnson. "Protein Kinase Inhibitors : Insights into Drug Design from Structure". In: *Science* 303.5665, pp. 1800-5.

Obach, R. Scott, Robert L Walsky, Karthik Venkatakrishnan, Emily A Gaman, J Brian Houston, and Larry M Tremaine (2006). "The utility of in vitro cytochrome P450 inhibition data in the prediction of drug-drug interactions." In: *J. Pharmacol. Exp. Ther.* 316.1, pp. 336–348.

Olaleye, Tayo O, James a Brannigan, Shirley M Roberts, Robin J Leatherbarrow, Anthony J Wilkinson, and Edward W Tate (2014). "Peptidomimetic inhibitors of N-myristoyltransferase from human malaria and leishmaniasis parasites." In: *Org. Biomol. Chem.* 12.41, pp. 8132–7.

Onufriev, Alexey, Donald Bashford, and David A Case (2004). "Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model". In: *Proteins Struct. Funct. Genet.* 55.2, pp. 383–394.

Panethymitaki, Chrysoula, Paul W Bowyer, Helen P Price, Robin J Leatherbarrow, Katherine a Brown, and Deborah F Smith (2006). "Characterization and selective inhibition of myristoyl-CoA:protein N-myristoyltransferase from Trypanosoma brucei and Leishmania major." In: *Biochem. J.* 396.2, pp. 277–85.

Park, Jin, Yingting Liu, Mark Lemmon, and Ravi Radhakrishnan (2012). "Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain". In: *Biochem. J.* 423.3, pp. 417–423.

Pasha, Mohammed Khysar, Ponniah Selvakumar, Lakshmikuttyamma Ashakumary, Mabood Qureshi, Frank S Guziec, Jonathan R Dimmock, Ronald L Felsted, Constance J Glover, and Rajendra K Sharma (2004). "Lipidic inhibitors of human N-myristoyltransferase." In: *Int. J. Mol. Med.* 13.4, pp. 557–563.

Pearlman, David A. and Patrick R. Connelly (1995). "Determination of the differential effects of hydrogen bonding and water release on the binding of FK506 to native and Tyr82–>Phe82 FKBP-12 proteins using free energy simulations." In: *J. Mol. Biol.* 248.3, pp. 696–717.

Phillips, James C., Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten (2005). "Scalable molecular dynamics with NAMD". In: *J. Comput. Chem.* 26.16, pp. 1781–1802.

Price, Helen P., Malini R. Menon, Chrysoula Panethymitaki, David Goulding, Paul G. McKean, and Deborah F. Smith (2003). "Myristoyl-CoA:Protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites". In: *J. Biol. Chem.* 278.9, pp. 7206–7214.

Price, Helen P, M Lucia S Güther, Michael a J Ferguson, and Deborah F Smith (2010). "Myristoyl-CoA:protein N-myristoyltransferase depletion in trypanosomes causes avirulence and endocytic defects." In: *Mol. Biochem. Parasitol.* 169.1, pp. 55–8.

Prompers, Jeanine J. and Rafael Brüschweiler (2002). "General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation". In: *J. Am. Chem. Soc.* 124.16, pp. 4522–4534.

Rackham, Mark D., James A. Brannigan, David K. Moss, Zhiyong Yu, Anthony J. Wilkinson, Anthony A. Holder, Edward W. Tate, and Robin J. Leatherbarrow (2013). "Discovery of novel and ligand-efficient inhibitors of plasmodium falciparum and plasmodium vivax N-myristoyltransferase". In: *J. Med. Chem.* 56.1, pp. 371–375.

Rackham, Mark D., James A. Brannigan, Kaveri Rangachari, Stephan Meister, Anthony J. Wilkinson, Anthony A. Holder, Robin J. Leatherbarrow, and Edward W. Tate (2014). "Design and synthesis of high affinity inhibitors of Plasmodium falciparum and Plasmodium vivax N-myristoyltransferases directed by ligand efficiency dependent lipophilicity (LELP)". In: *J. Med. Chem.* 57.6, pp. 2773–2788.

Rackham, Mark D, Zhiyong Yu, James A Brannigan, William P Heal, Daniel Paape, K Victoria Barker, Anthony J Wilkinson, Deborah F Smith, Robin J Leatherbarrow, and Edward W Tate (2015). "Discovery of high affinity inhibitors of Leishmania donovani N-myristoyltransferase." In: *Medchemcomm* 6.10, pp. 1761–1766.

Radhakrishnan, Mala L. and Bruce Tidor (2007). "Specificity in molecular design: A physical framework for probing the determinants of binding specificity and promiscuity in a biological environment". In: *J. Phys. Chem. B* 111.47, pp. 13419–13435.

Radi, Marco, Ralf Schneider, Anna Lucia Fallacara, Lorenzo Botta, Emmanuele Crespan, Cristina Tintori, Giovanni Maga, Miroslava Kissova, Alessia Calgani, André Richters, Franesca Musumeci, Daniel Rauh, and Silvia Schenone (2016). "A cascade screening approach for the identification of Bcr-Abl myristate pocket binders active against wild type and T315I mutant". In: *Bioorganic Med. Chem. Lett.* 26.15, pp. 3436–3440.

Rarey, Matthias, Bernd Kramer, Thomas Lengauer, and Gerhard Klebe (1996). "A fast flexible docking method using an incremental construction algorithm." In: *J. Mol. Biol.* 261.3, pp. 470–89.

Reulecke, Ingo, Gudrun Lange, Jürgen Albrecht, Robert Klein, and Matthias Rarey (2008). "Towards an integrated description of hydrogen bonding and dehydration: Decreasing false positives in virtual screening with the HYDE scoring function". In: *ChemMedChem* 3.6, pp. 885–897.

Ritzefeld, Markus, Megan H Wright, and Edward W Tate (2017). "New developments in probing and targeting protein acylation in malaria, leishmaniasis and African sleeping sickness". In: *Parasitology*, pp. 1–18.

Roberts, Adam J., Leah S. Torrie, Susan Wyllie, and Alan H. Fairlamb (2014). "Biochemical and genetic characterization of Trypanosoma cruzi N-myristoyltransferase". In: *Biochem. J.* 459.2, pp. 323–332.

Robinson, Daniel, Thomas Bertrand, Jean Christophe Carry, Frank Halley, Andreas Karlsson, Magali Mathieu, Hervé Minoux, Marc Antoine Perrin, Benoit Robert, Laurent Schio, and Woody Sherman (2016). "Differential Water Thermodynamics Determine PI3K-Beta/Delta Selectivity for Solvent-Exposed Ligand Modifications". In: *J. Chem. Inf. Model.* 56.5, pp. 886–894.

Robinson, Daniel D, Woody Sherman, and Ramy Farid (2010). "Understanding kinase selectivity through energetic analysis of binding site waters." In: *ChemMedChem* 5.4, pp. 618–27.

Robinson, David A. and Paul G. Wyatt (2015). "Identification and structure solution of fragment hits against kinetoplastid N-myristoyltransferase". In: *Acta Crystallogr. Sect. FStructural Biol. Commun.* 71, pp. 586–593.

Rosner, Kristin E, Zhuyan Guo, Peter Orth, Gerald W Shipps, David B Belanger, Tin Yau Chan, Patrick J Curran, Chaoyang Dai, Yongqi Deng, Vinay M Girijavallabhan, Liwu Hong, Brian J Lavey, Joe F Lee, Dansu Li, Zhidan Liu, Janeta Popovici-Muller, Pauline C Ting, Henry Vaccaro, Li Wang, Tong Wang, Wensheng Yu, Guowei Zhou, Xiaoda Niu, Jing Sun, Joseph A Kozlowski, Daniel J Lundell, Vincent Madison, Brian McKittrick, John J Piwinski, Neng Yang Shih, M Arshad Siddiqui, and Corey O Strickland (2010). "The discovery of novel tartrate-based TNF-alpha converting enzyme (TACE) inhibitors". In: *Bioorganic Med. Chem. Lett.* 20.3, pp. 1189–1193.

Rudnick, D A, C A McWherter, W J Rocque, P J Lennon, D P Getman, and J I Gordon (1991). "Kinetic and structural evidence for a sequential ordered Bi Bi mechanism of catalysis by Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase." In: *J. Biol. Chem.* 266.15, pp. 9732–9.

Sanschagrin, P C and L A Kuhn (1998). "Cluster analysis of consensus water sites in thrombin and trypsin shows conservation between serine proteases and contributions to ligand specificity." In: *Protein Sci.* 7.10, pp. 2054–64.

Scheffzek, Klaus, Werner Kliche, Lisa Wiesmüller, and Jochen Reinstein (1996). "Crystal structure of the complex of UMP/CMP kinase from Dictyostelium discoideum and the bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-uridyl) pentaphosphate (UP5A) and Mg2+ at 2.2 A: Implications for water-mediated specificity". In: *Biochemistry* 35.30, pp. 9716–9727.

Schindler, Thomas, W Bornmann, Patricia Pellicena, W Todd Miller, Bayard Clarkson, and J Kuriyan (2000). "Structural mechanism for STI-571 inhibition of abelson tyrosine kinase." In: *Science* 289.5486, pp. 1938–42.

Sharp, Paul M., Elizabeth Cowe, Desmond G. Higgins, Denis C. Shields, Kenneth H. Wolfe, and Frank Wright (1988). "Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity". In: *Nucleic Acids Res.* 16.17, pp. 8207–8211.

Showalter, S A, E Johnson, M Rance, and R Brüschweiler (2007). "Toward Quantitative Interpretation of Methyl Side-Chain Dynamics from NMR by Molecular Dynamics Simulations". In: *J. Am. Chem. Soc.* 129.46, pp. 14146–14147.

Shultz, Michael D. (2013). "Setting expectations in molecular optimizations: Strengths and limitations of commonly used composite parameters". In: *Bioorganic Med. Chem. Lett.* 23.21, pp. 5980–5991.

Sidhu, Ranjinder S., Jullia Y. Lee, Chong Yuan, and William L. Smith (2010). "Comparison of cyclooxygenase-1 crystal structures: Cross-talk between monomers comprising cyclooxygenase-1 homodimers". In: *Biochemistry* 49.33, pp. 7069–7079.

Sievers, Fabian, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, Julie D Thompson, and Desmond G Higgins (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." In: *Mol. Syst. Biol.* 7.1, p. 539.

Skjaerven, Lars, Xin-Qiu Yao, Guido Scarabelli, and Barry J Grant (2014). "Integrating protein structural dynamics and evolutionary analysis with Bio3D." In: *BMC Bioinformatics* 15.1, p. 399.

Smith, David K, Predrag Radivojac, Zoran Obradovic, A Keith Dunker, and Guang Zhu (2003). "Improved amino acid flexibility parameters." In: *Protein Sci.* 12.5, pp. 1060–1072.

Snyder, Phillip W, Jasmin Mecinovic, Demetri T Moustakas, Samuel W Thomas, Michael Harder, Eric T Mack, Matthew R Lockett, Annie Héroux, Woody Sherman, and George M Whitesides (2011). "Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase." In: *Proc. Natl. Acad. Sci. U. S. A.* 108.44, pp. 17889–17894.

Snyder, Phillip W., Matthew R. Lockett, Demetri T. Moustakas, and George M. Whitesides (2014). "Is it the shape of the cavity, or the shape of the water in the cavity?" In: *Eur. Phys. J. Spec. Top.* 223.5, pp. 853–891.

Sogabe, Satoshi, Miyako Masubuchi, Kiyoaki Sakata, Takaaki A. Fukami, Kenji Morikami, Yasuhiko Shiratori, Hirosato Ebiike, Kenichi Kawasaki,

Yuko Aoki, Nobuo Shimma, Allan D'Arcy, Fritz K. Winkler, David W. Banner, and Tatsuo Ohtsuka (2002). "Crystal structures of Candida albicans N-myristoyltransferase with two distinct inhibitors". In: *Chem. Biol.* 9.10, pp. 1119–1128.

Sotriffer, Christoph a. (2011). "Molecular Dynamics Simulations". In: *Mol. Dyn. Simul. Drug Des.* 23.23, p. 235104.

Spinks, Daniel, Victoria Smith, Stephen Thompson, David A. Robinson, Torsten Luksch, Alasdair Smith, Leah S. Torrie, Stuart McElroy, Laste Stojanovski, Suzanne Norval, Iain T. Collie, Irene Hallyburton, Bhavya Rao, Stephen Brand, Ruth Brenk, Julie A. Frearson, Kevin D. Read, Paul G. Wyatt, and Ian H. Gilbert (2015). "Development of small-molecule Trypanosoma brucei N-myristoyltransferase inhibitors: Discovery and optimisation of a novel binding mode". In: *ChemMedChem* 10.11, pp. 1821–1836.

Stengl, Bernhard, Emmanuel A. Meyer, Andreas Heine, Ruth Brenk, François Diederich, and Gerhard Klebe (2007). "Crystal Structures of tRNA-guanine Transglycosylase (TGT) in Complex with Novel and Potent Inhibitors Unravel Pronounced Induced-fit Adaptations and Suggest Dimer Formation Upon Substrate Binding". In: *J. Mol. Biol.* 370.3, pp. 492–511.

Sterling, Teague and John J Irwin (2015). "ZINC 15 - Ligand Discovery for Everyone". In: *J. Chem. Inf. Model.* 55.11, pp. 2324–2337.

Stierand, Katrin and Matthias Rarey (2007). "From modeling to medicinal chemistry: automatic generation of two-dimensional complex diagrams." In: *ChemMedChem* 2.6, pp. 853–860.

— (2010). "Drawing the PDB: Protein-ligand complexes in two dimensions". In: *ACS Med. Chem. Lett.* 1.9, pp. 540–545.

Stockmann, Henning, Agnieszka Bronowska, Neil R. Syme, Gary S. Thompson, Arnout P. Kalverda, Stuart L. Warriner, and Steve W. Homans (2008). "Residual Ligand Entropy in the Binding of p-Substituted Benzenesulfonamide Ligands to Bovine Carbonic Anhydrase II". In: *J. Am. Chem. Soc.* 130.37, pp. 12420–12426.

Tarek, Mounir and Douglas J. Tobias (2000). "The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments." In: *Biophys. J.* 79.6, pp. 3244–57.

Tate, Edward W, Andrew S Bell, Mark D Rackham, and Megan H Wright (2013). "N-Myristoyltransferase as a potential drug target in malaria and leishmaniasis." In: *Parasitology*, pp. 1–13.

Teague, Simon J (2003). "Implications of protein flexibility for drug discovery." In: *Nat. Rev. Drug Discov.* 2.7, pp. 527–541.

Teng, Min, Mark T Hilgers, Mark L Cunningham, Allen Borchardt, Jeffrey B Locke, Sunny Abraham, Gregory Haley, Bryan P Kwan, Courtney Hall, Grayson W Hough, Karen J Shaw, and John Finn (2013). "Identification

of bacteria-selective threonyl-tRNA synthetase substrate inhibitors by structure-based design". In: *J. Med. Chem.* 56.4, pp. 1748–1760.

Thaimattam, Ram, Rahul Banerjee, Rajni Miglani, and Javed Iqbal (2007). "Protein kinase inhibitors: structural insights into selectivity". In: *Curr. Pharm. Des* 13.27, pp. 2751–2765.

Thinon, Emmanuelle, Remigiusz A Serwa, Malgorzata Broncel, James A Brannigan, Ute Brassat, Megan H Wright, William P Heal, Anthony J Wilkinson, David J Mann, and Edward W Tate (2014). "Global profiling of co- and post-translationally N-myristoylated proteomes in human cells." In: *Nat. Commun.* 5, p. 4919.

Thinon, Emmanuelle, Julia Morales-Sanfrutos, David J. Mann, and Edward W. Tate (2016). "N-Myristoyltransferase Inhibition Induces ER-Stress, Cell Cycle Arrest, and Apoptosis in Cancer Cells". In: *ACS Chem. Biol.* acschembio.6b00371.

Tochowicz, Anna, Klaus Maskos, Robert Huber, Ruth Oltenfreiter, Vincent Dive, Athanasios Yiotakis, Matteo Zanda, Wolfram Bode, and Peter Goettig (2007). "Crystal Structures of MMP-9 Complexes with Five Inhibitors: Contribution of the Flexible Arg424 Side-chain to Selectivity". In: *J. Mol. Biol.* 371.4, pp. 989–1006.

Traverso, José a, Carmela Giglione, and Thierry Meinnel (2013). "High-throughput profiling of N-myristoylation substrate specificity across species including pathogens." In: *Proteomics* 13.1, pp. 25–36.

Trbovic, Nikola, Byungchan Kim, Richard A Friesner, and Arthur G Palmer (2008). "Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation." In: *Proteins* 71.2, pp. 684–94.

Trueblood, K. N., H. B. Bürgi, H. Burzlaff, J. D. Dunitz, C. M. Gramaccioli, H. H. Schulz, U. Shmueli, S. C. Abrahams, and IUCr (1996). "Atomic Dispacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature". In: *Acta Crystallogr. Sect. A Found. Crystallogr.* 52.5, pp. 770–781.

Utsumi, Toshihiko, Kengo Nakano, Takeshi Funakoshi, Yoshiyuki Kayano, Sayaka Nakao, Nagisa Sakurai, Hiroyuki Iwata, and Rumi Ishisaka (2004). "Vertical-scanning mutagenesis of amino acids in a model N-myristoylation motif reveals the major amino-terminal sequence requirements for protein N-myristoylation". In: *Eur. J. Biochem.* 271.4, pp. 863–874.

Vanommeslaeghe, K. and A. D. MacKerell (2012). "Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing". In: *J. Chem. Inf. Model.* 52.12, pp. 3144–3154.

Vanommeslaeghe, K, E Prabhu Raman, and A D MacKerell (2012). "Automation of the CHARMM General Force Field (CGenFF) II: Assignment of

Bonded Parameters and Partial Atomic Charges". In: *J. Chem. Inf. Model.* 52.12, pp. 3155–3168.

Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell (2010). "CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields". In: *J Comput Chem* 31.4, pp. 671–690.

Wang, J M, R M Wolf, J W Caldwell, P a Kollman, and D a Case (2004). "Development and testing of a general amber force field". In: *J. Comput. Chem.* 25.9, pp. 1157–1174.

Wang, Junmei, Wei Wang, Peter A. Kollman, and David A. Case (2006). "Automatic atom type and bond type perception in molecular mechanical calculations". In: *J. Mol. Graph. Model.* 25.2, pp. 247–260.

Wilcox, Celeste, J S Hu, and Eric N Olson (1987). "Acylation of proteins with myristic acid occurs cotranslationally." In: *Science* 238.4831, pp. 1275–8.

Winn, Martyn D., Charles C. Ballard, Kevin D. Cowtan, Eleanor J. Dodson, Paul Emsley, Phil R. Evans, Ronan M. Keegan, Eugene B. Krissinel, Andrew G W Leslie, Airlie McCoy, Stuart J. McNicholas, Garib N. Murshudov, Navraj S. Pannu, Elizabeth A. Potterton, Harold R. Powell, Randy J. Read, Alexei Vagin, and Keith S. Wilson (2011). "Overview of the CCP4 suite and current developments". In: *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67.4, pp. 235–242.

Wright, Megan H, William P Heal, David J Mann, and Edward W Tate (2010). "Protein myristoylation in health and disease." In: *J. Chem. Biol.* 3.1, pp. 19–35.

Wright, Megan H, Barbara Clough, Mark D Rackham, Kaveri Rangachari, James A Brannigan, Munira Grainger, David K Moss, Andrew R Bottrill, William P Heal, Malgorzata Broncel, Remigiusz A Serwa, Declan Brady, David J Mann, Robin J Leatherbarrow, Rita Tewari, Anthony J Wilkinson, Anthony A Holder, and Edward W Tate (2014). "Validation of N-myristoyltransferase as an antimalarial drug target using an integrated chemical biology approach". In: *Nat. Chem.* 6.2, pp. 112–21.

Wright, Megan H., Daniel Paape, Elisabeth M. Storck, Remigiusz A. Serwa, Deborah F. Smith, and Edward W. Tate (2015). "Global analysis of protein N-myristoylation and exploration of N-myristoyltransferase as a drug target in the neglected human pathogen leishmania donovani". In: *Chem. Biol.* 22.3, pp. 342–354.

Wylie, Andrew A, Joseph Schoepfer, Wolfgang Jahnke, Sandra W Cowan-Jacob, Alice Loo, Pascal Furet, Andreas L Marzinzik, Xavier Pelle, Jerry Donovan, Wenjing Zhu, Silvia Buonamici, A Quamrul Hassan, Franco Lombardo, Varsha Iyer, Michael Palmer, Giuliano Berellini, Stephanie Dodd, Sanjeev Thohan, Hans Bitter, Susan Branford, David M Ross,

Timothy P Hughes, Lilli Petruzzelli, K Gary Vanasse, Markus Warmuth, Francesco Hofmann, Nicholas J Keen, and William R Sellers (2017). "The allosteric inhibitor ABL001 enables dual targeting of BCR–ABL1". In: *Nature*.

Xu, Mengmeng, Longxiang Xie, Zhaoxiao Yu, and Jianping Xie (2015). "Roles of Protein N -Myristoylation and Translational Medicine Applications". In: 25.3, pp. 259–268.

Yang, Daiwen and Lewis E Kay (1996). "Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding." In: *J. Mol. Biol.* 263.2, pp. 369–382.

Young, Tom, Robert Abel, Byungchan Kim, Bruce J Berne, and Richard A Friesner (2007). "Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding." In: *Proc. Natl. Acad. Sci. U. S. A.* 104.3, pp. 808–13.

Yu, Zhiyong, James A. Brannigan, Kaveri Rangachari, William P. Heal, Anthony J. Wilkinson, Anthony A. Holder, Robin J. Leatherbarrow, and Edward W. Tate (2015). "Discovery of pyridyl-based inhibitors of Plasmodium falciparum N-myristoyltransferase". In: *Medchemcomm* 6.10, pp. 1767–1772.

Yu, Zhiyong Y, James a Brannigan, David K Moss, a Marek Brzozowski, Anthony J Wilkinson, Anthony a Holder, Edward W Tate, and Robin J Leatherbarrow (2012). "Design and Synthesis of Inhibitors of Plasmodium falciparum N-Myristoyltransferase, A Promising Target for Antimalarial Drug Discovery". In: *J. Med. Chem.* 55.20, pp. 8879–8890.

Zacharias, David E., Peter Murray-Rust, Robert M. Preston, and Jenny P. Glusker (1983). "The geometry of the thioester group and its implications for the chemistry of acyl coenzyme a". In: *Arch. Biochem. Biophys.* 222.1, pp. 22–34.

Zentgraf, Matthias, Holger Steuber, Cornelia Koch, Concettina La Motta, Stefania Sartini, Christoph A. Sotriffer, and Gerhard Klebe (2007). "How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase". In: *Angew. Chemie - Int. Ed.* 46.19, pp. 3575–3578.

Zhang, Jianming, Francisco J Adrián, Wolfgang Jahnke, Sandra W Cowan-Jacob, Allen G Li, Roxana E Iacob, Taebo Sim, John Powers, Christine Dierks, Fangxian Sun, Gui-Rong Guo, Qiang Ding, Barun Okram, Yongmun Choi, Amy Wojciechowski, Xianming Deng, Guoxun Liu, Gabriele Fendrich, André Strauss, Navratna Vajpai, Stephan Grzesiek, Tove Tuntland, Yi Liu, Badry Bursulaya, Mohammad Azam, Paul W Manley, John R Engen, George Q Daley, Markus Warmuth, and Nathanael S Gray

(2010). “Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors.” In: *Nature* 463.7280, pp. 501–506.

Zhang, Y L and Z Y Zhang (1998). “Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B.” In: *Anal. Biochem.* 261.2, pp. 139–148.

Zhao, Can and Shutao Ma (2014). “Recent Advances in The Discovery of N-Myristoyltransferase Inhibitors”. In: *ChemMedChem* 9.11, pp. 2425–2437.

Zhao, Hongtao and Amedeo Caflisch (2015). “Molecular dynamics in drug design”. In: *Eur. J. Med. Chem.* 91, pp. 4–14.

Zoete, Vincent, Michel A Cuendet, Aurélien Grosdidier, and Olivier Michielin (2011). “SwissParam: A fast force field generation tool for small organic molecules”. In: *J. Comput. Chem.* 32.11, pp. 2359–2368.