

# How to design selective ligands for highly conserved binding sites: A case study using *N*-myristoyltransferases as model system

Christian Kersten,<sup>ab</sup> Edmond Fleischer,<sup>a</sup> Josef Kehrein,<sup>bc</sup> Christoph Borek,<sup>a</sup> Elmar Jaenicke,<sup>d</sup> Christoph Sotriffer,<sup>c</sup> Ruth Brenk<sup>b,\*</sup>

<sup>a</sup> Institute of Pharmacy and Biochemistry, Johannes Gutenberg-Universität Mainz, Staudingerweg 5, 55128, Mainz, Germany

<sup>b</sup> Department of Biomedicine, University of Bergen, Jonas Lies vei 91, 5020, Bergen, Norway.

<sup>c</sup> Institute of Pharmacy and Food Chemistry, Julius-Maximilians-Universität Würzburg, Am Hubland, 97074, Würzburg, Germany

<sup>d</sup> Institute of Molecular Biophysics, Johannes Gutenberg University, Jakob-Welder-Weg 26, 55128, Mainz, Germany.

\* corresponding author: [ruth.brenk@uib.no](mailto:ruth.brenk@uib.no)

## Abstract

A model system of two related enzymes with conserved binding sites, namely *N*-myristoyltransferase from two different organisms, was studied to decipher the driving forces that lead to selective inhibition in such cases. Using a combination of computational and experimental tools, two different selectivity-determining features were identified. For some ligands, a change in side chain flexibility appears to be responsible for selective inhibition. Remarkably, this was observed for residues orienting their side chains away from the ligands. For other ligands, selectivity is caused by interfering with a water molecule that binds stronger to the off-target than to the target. Based on this finding, a virtual screen for selective compounds was conducted resulting in three hit compounds with the desired selectivity profile. This study delivers a guideline on how to assess selectivity-determining features in proteins with conserved binding sites and to translate this knowledge into the design of selective inhibitors.

## Introduction

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. The most common and routinely applied strategies<sup>1</sup> include the exploitation of differences in the overall shape of the binding site,<sup>2,3</sup> electrostatic interaction patterns,<sup>4-7</sup> displaceable water molecules<sup>8-11</sup> or, if available, addressing an allosteric functional binding pocket.<sup>12</sup>

However, the task of selective-ligand design becomes more challenging when facing a conserved binding-site between target and off-target(s). Even though for such challenging scenarios selective inhibitors were reported, the underlying molecular mechanism for selectivity often remained unclear.<sup>13-17</sup> In some of these cases, protein flexibility<sup>18,19</sup> or explicit water molecules within the binding site<sup>8-11</sup> of the enzymes appeared to be the main contributing factors.

Here, we embarked to reveal the selectivity-determining features in proteins with conserved binding sites using *N*-myristoyltransferase (NMT) as a model system. NMT catalyzes the transfer of the C:14 saturated fatty acid myristate to the *N*-terminal glycine residue of recognized proteins substrates. The enzyme has an ordered Bi-Bi reaction mechanism, binding first to myristoyl-CoA (MyrCoA) with the resulting conformational changes generating a peptide-binding site.<sup>20</sup> Subsequent formation of a ternary MyrCoA:NMT-peptide complex leads to catalysis and product release. The recognized peptides are species dependent, but all possess an *N*-terminal glycine, which is activated during the reaction through deprotonation by the carboxylate group of the C-terminus of enzyme.<sup>21</sup> NMT was shown to be essential for parasitic survival and virulence,<sup>22,23</sup> and is a drug target for cancer, and a range of parasitic and viral diseases.<sup>24-33</sup> In this study, we focused on NMT from the protozoan parasite *Leishmania major* (*Lm*NMT) and its human homologue *Hs*NMT1. Both enzymes share a sequence identity of 45.4 % but have highly conserved binding sites with only three residues differing (Figure 1). Further, chemically closely related

selective and unselective inhibitors were reported and the binding modes of some of these compounds were elucidated using X-ray crystallography.

Pyrazole sulfonamides constitute one large class of NMT inhibitors (e.g. **1** to **4** in Figure 2).<sup>34–36</sup> They bind to *Lm*NMT by forming hydrogen bonds with Ser330 from the pyrazole moiety which also forms  $\pi$ - $\pi$  stacking interactions with Phe90 (Figure 1). The trimethylation of the pyrazole improves packing within that pocket by addressing Phe88 and Leu341. The sulfonamide moiety interacts water-mediated with His219 and the backbone of Asp396 and Gly397. A hydrophobic, aromatic linker stacks with Tyr217 while a basic center, which mimics the N-terminus of the substrates, interacts directly or via a water molecule with the catalytically active C-terminal Leu421.

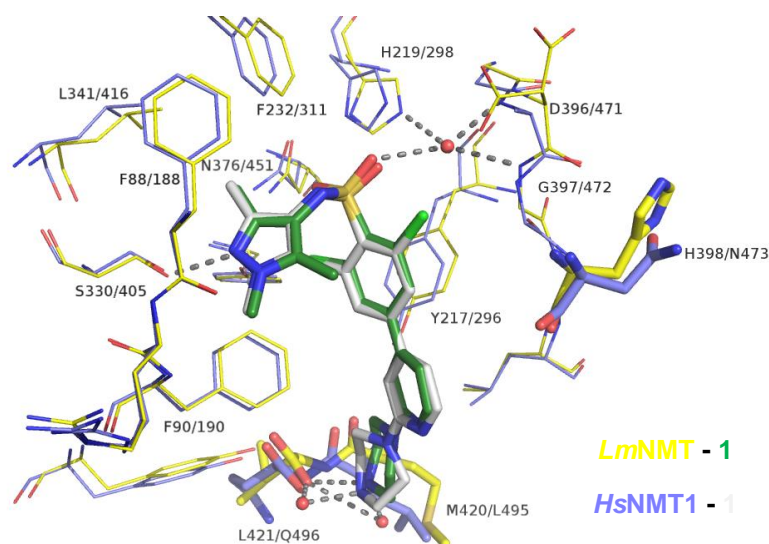


Figure 1: Superposition of binding sites of *Lm*NMT-1 (yellow carbon atoms for protein and green for ligand, PDB code 2WSA) and *Hs*NMT1-1 (blue carbon atoms for protein and white for ligand, PDB code 3IWE). For clarity, water molecules are only shown for *Lm*NMT-1 (red spheres). Hydrogen-bonds are indicated as dashed lines. Residues differing between both NMTs are highlighted as sticks. Residues are labeled as *Lm*NMT/*Hs*NMT1.

Further compound classes, containing piperidylindoles, aminoacylpyrrolidines and oxadiazole scaffolds, were also reported to inhibit NMT (e.g. **5** to **8** in Figure 2).<sup>37–39</sup> All these compounds contain a basic center to interact with the C-terminus but bind to the open binding site conformation. This is characterized by a rotation of Tyr217 which gives access to a mainly hydrophobic pocket (Figure 3 a). Additionally, compound **5** is lacking a functional group to form a hydrogen bond with Ser330.

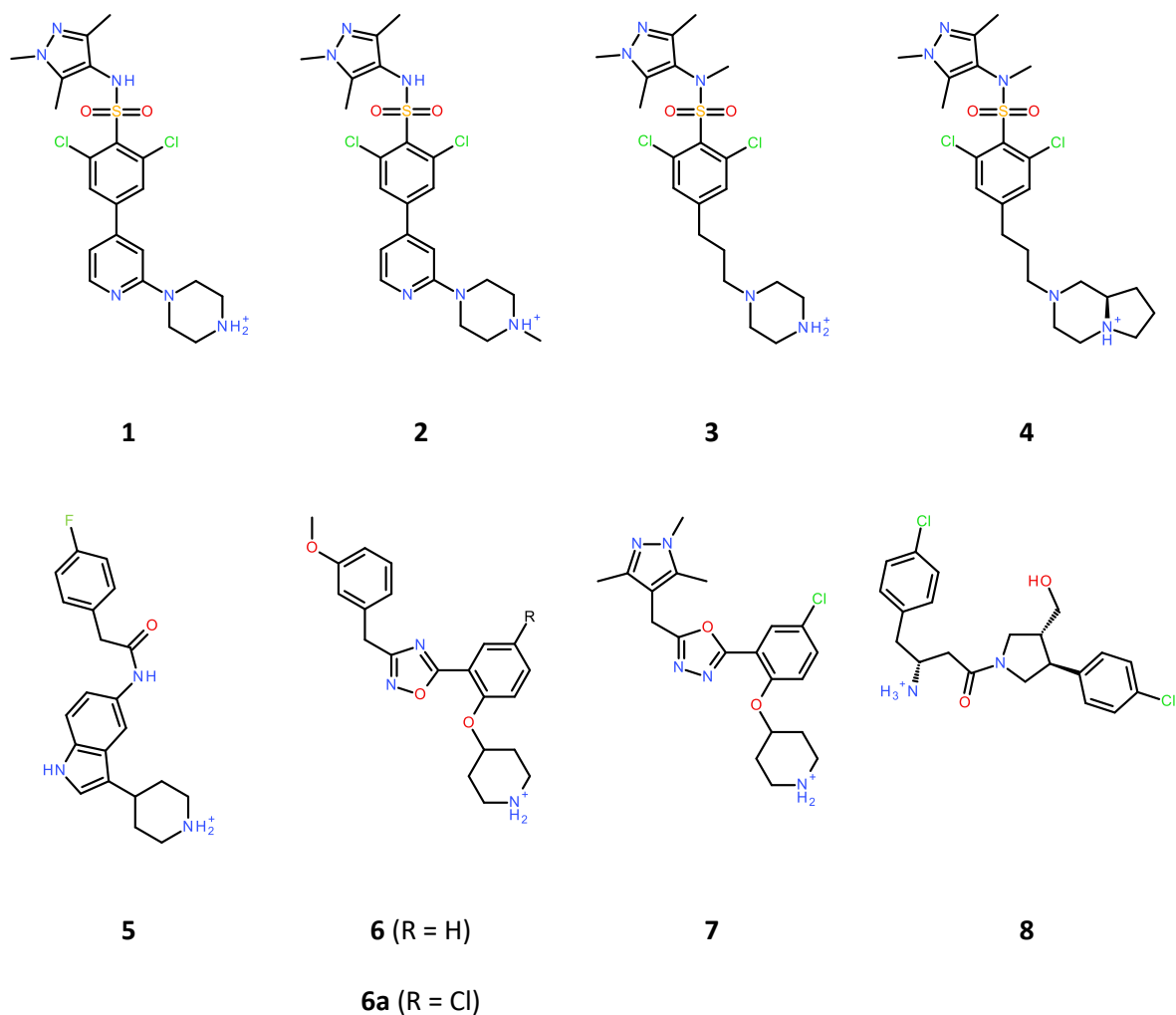


Figure 2: Chemical structures of investigated NMT inhibitors. Protonation states for neutral pH was used for modelling.

Several inhibitors selective for *Lm*NMT over *Hs*NMT1 were developed.<sup>34–40</sup> In the reported crystal structures, the NMT inhibitors only form main chain interactions with the three non-conserved binding

site residues (Figure 1). Hence, the reasons for selective inhibition remained unclear. For compound **5**, which is about 108-fold selective for *Lm*NMT over *Hs*NMT1, it was initially hypothesized that the selectivity arises from rotation of Tyr217 in *Lm*NMT which is required for binding (Figure 3 a).<sup>37,38</sup> It was suggested that the corresponding rotation of Tyr296 in *Hs*NMT1 could be unfavorable and thus cause selectivity.<sup>38</sup> However, recently the unselective *L. donovani* NMT (*Ld*NMT) inhibitors **6a** and **7** were reported (Figure 2; *Ld*NMT and *Lm*NMT share a sequence identity of 97.8% and inhibition data are comparable between the two enzymes.<sup>38</sup>). Compounds **6a** and **7** bind in a similar orientation to NMT as compound **5**, demonstrating that the orientation of Tyr217 alone cannot be the selectivity-determining feature (Figure 3).<sup>39</sup>

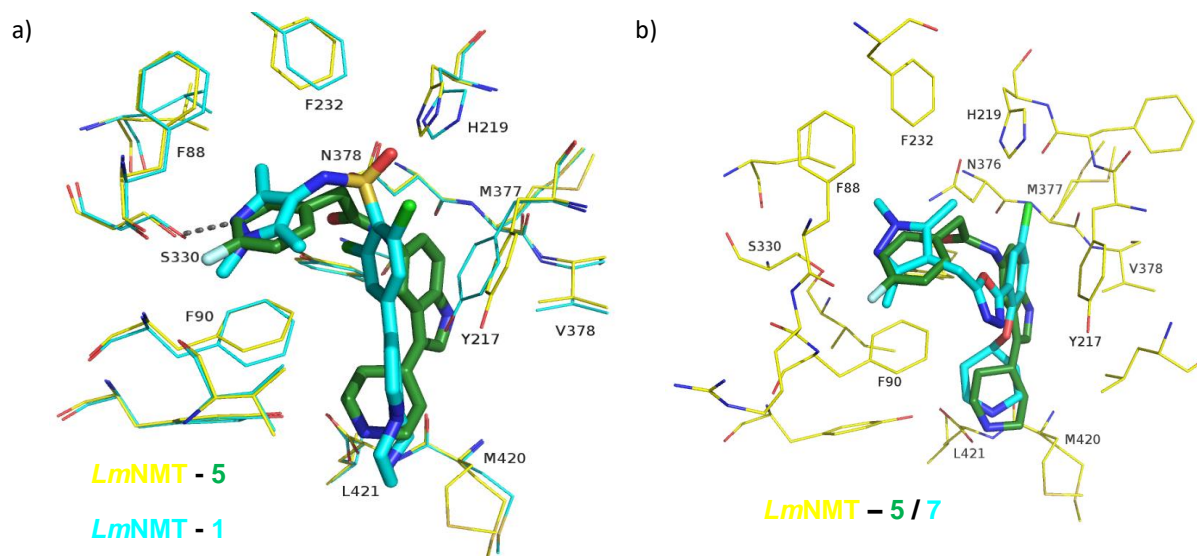


Figure 3: Binding mode of compound **5** in comparison to other ligands. a) Superposition of *Lm*NMT in complex with compound **1** (cyan carbon atoms of ligand and receptor, PDB code 2WSA) and compound **5** (green ligand carbon atoms and yellow receptor carbon atoms, PDB code 4CGN). The binding modes of the two compounds differ in compound **5** lacking a hydrogen-bond acceptor to interact with Ser330 and Tyr217 only adopting the open conformation when **5** is bound. b) Superposition of the binding modes of the selective compound **5** (green carbon atoms, PDB code 4CGN) and the unselective compound **7** (cyan carbon atoms, PDB code 5A28) binding to *Lm*NMT (yellow carbon atoms, for clarity, only shown from PDB code 4CGN). In both cases, Tyr217 adopts the open conformation.

To obtain a more detailed understanding of the molecular driving forces that lead to selectivity in proteins with conserved binding sites, we studied ligand binding to *Lm*NMT and *Hs*NMT1 in detail. Using a combination of molecular dynamics simulations, isothermal titration calorimetry, enzyme inhibition assay, site-directed mutagenesis and X-ray crystallography the thermodynamics of ligand binding, protein dynamics, water network formation and their changes upon ligand binding were analyzed. This approach led to the identification of two different selectivity-determining features for the compounds described above. Based on these findings, a virtual screening for selective compounds was conducted resulting in three hit compounds with the desired selectivity profile. Together with these results, the implications of this study for the rational design of selective inhibitors in general are also discussed.

## Results

### Selectivity for wild type proteins

First, we focused on the sulfonamides **1** – **4** and the indole derivative **5** (Figure 2). Different assays were previously used to determine inhibition constants.<sup>34,35,41</sup> Thus, to obtain consistent values, the inhibition constants of the compounds were re-determined using a fluorescence assay. As reported earlier, compounds **1-3** inhibited strongly both enzymes with only minor selectivity, whereas compounds **4** and **5** were selective for *Lm*NMT with a selectivity index (SI) of 215 and 16, respectively (Table 1).

Table 1: Inhibition constants and selectivity of NMT inhibitors **1-5**. All measurements were performed at least in duplicates.  $K_i$  values as mean with standard error were calculated from  $IC_{50}$  and  $K_M$  values using the Cheng-Prusoff equation.

Compound	$K_i$ [nM]		
	<i>Lm</i> NMT	<i>Hs</i> NMT1	Selectivity index (SI) <sup>a</sup>
<b>1</b>	8.4 ± 1.3	31.6 ± 4.5	4
<b>2</b>	1.4 ± 0.3	13.3 ± 2.8	9
<b>3</b>	19.8 ± 2.5	96.4 ± 11.7	5
<b>4</b>	2.0 ± 0.4	4.3 × 10 <sup>2</sup> ± 91	215
<b>5</b>	9.8 × 10 <sup>2</sup> ± 1.0 × 10 <sup>2</sup>	1.5 × 10 <sup>4</sup> ± 4.4 × 10 <sup>3</sup>	16

<sup>a</sup> SI was calculated as  $K_i$  (*Hs*NMT)/ $K_i$ (*Lm*NMT1)

ITC experiments were performed to elucidate the thermodynamic binding profiles. The measurements confirmed selectivity of compound **4** and **5**, albeit with an altered selectivity profile (SI of 8 for compound **4** and 113 for compound **5**, Figure 4 a and b, Table S 1 in supporting information). For all ligands, regardless of their SI, binding to *Lm*NMT was clearly dominated by enthalpy, whereas for *Hs*NMT1 enthalpic and entropic contributions were more balanced.

Crystal structures of ligand **1** in complex with *Lm*NMT and *Hs*NMT1, as well as ligand **5** in complex with *Lm*NMT were available in the protein databank (PDB<sup>42</sup>).<sup>35,38</sup> Structure determination of the remaining structures was attempted in order to elucidate the binding modes of the ligands. Co-crystallization and soaking experiments of *Hs*NMT1 with compound **5** were unsuccessful as no additional electron density for the ligand was observed within the binding site. For the selective ligand **4**, crystal structures of the complexes with *Lm*NMT and *Hs*NMT1 with a resolution of 1.50 and 1.89 Å, respectively, were determined.



In addition, for the unselective ligand **2** a structure in complex with *Hs*NMT1 was determined at 2.05 Å resolution (Table S 2). In all available structures, the binding modes were conserved between both protein species and no changes in binding site conformations were observed (Figure 1 and Figure 5a).

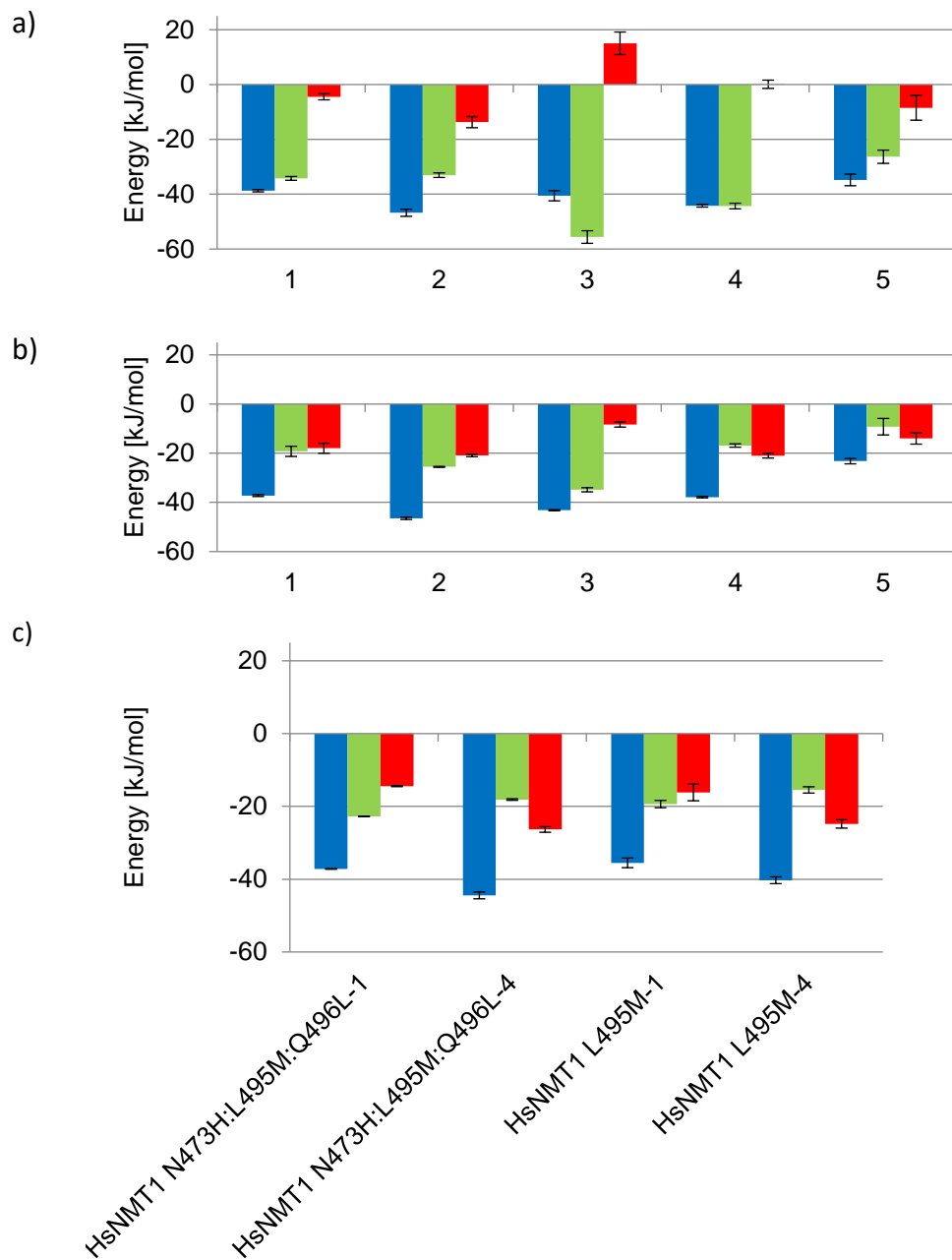


Figure 4: Thermodynamic profiles of compounds **1-5** binding to a) *Lm*NMT and b) *Hs*NMT based on ITC experiments. c) Thermodynamic profiles of compounds **1** and **4** binding to active-site mutants *Hs*NMT1 N374H:L495M:Q496L and *Hs*NMT1 L495M.  $\Delta G^\circ$ ,  $\Delta H^\circ$  and  $-\Delta S^\circ$  are depicted as blue, green and red colored bars, respectively.

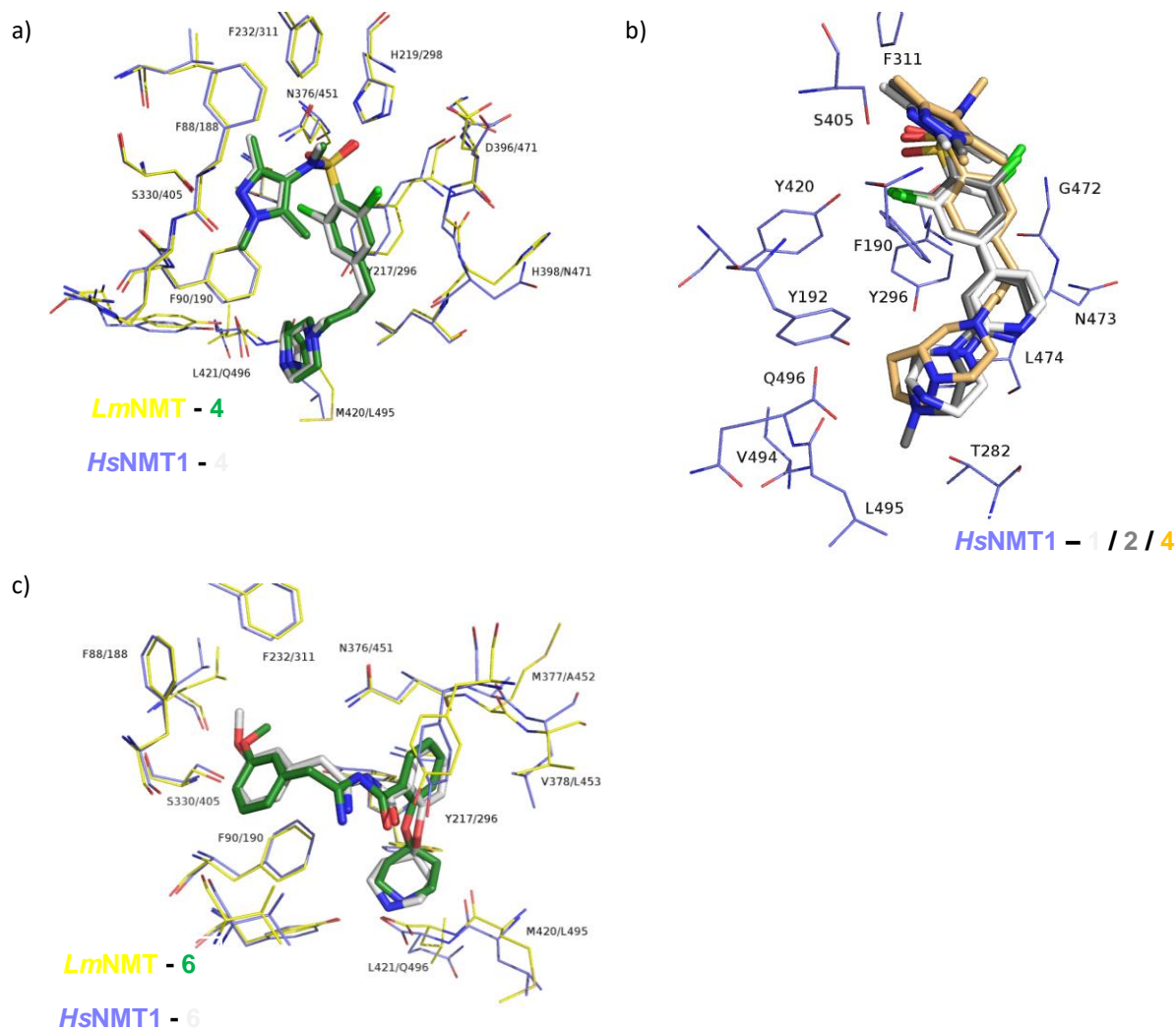


Figure 5: Superposition of a) compound **4** binding to *LmNMT* (green ligand carbon atoms and yellow protein carbon atoms, PDB code 6EU5) and *HsNMT1* (white ligand carbon atoms and blue protein carbon atoms, PDB code 6FZ5). b) Superposition of compounds **1** (white carbon atoms, PDB code 3IWE), **2** (grey carbon atoms, PDB code 6FZ3) and **4** (light orange carbon atoms, PDB code 6FZ5) in complex with *HsNMT1* (blue carbon atoms, residues only shown from *HsNMT1*-**4** complex, residues covering the ligands are hidden for clear view). c) Compound **6** binding to *LmNMT* (green ligand carbon atoms and yellow protein carbon atoms, PDB code 6EWF) and *HsNMT1* (white ligand carbon atoms and blue protein carbon atoms, PDB code 6FZ2). In a and c, residues are labeled as *LmNMT*/*HsNMT1*.

Purely based on the interactions formed in the binding sites, the selectivity profile of the ligands could not be explained. Therefore, site-directed mutagenesis experiments combined with MD simulations were carried out to obtain a more detailed picture of the molecular reasons for selective inhibition by these compounds.

#### Binding-site swap

The binding sites of *Lm*NMT and *Hs*NMT1 are highly conserved. Within 5 Å around compound **1**, the only difference are the change of His398, Met420 and Leu421 in *Lm*NMT to Asn473, Leu495 and Gln496 in *Hs*NMT1 (Figure 1). However, the ligands form no direct interactions with the side chain atoms of these residues. To elucidate the influence of these residues on selectivity, they were swapped with the corresponding residues in the orthologous protein to obtain *Lm*NMT with an *Hs*NMT1 binding site and *vice versa*. The inhibition profiles of the mutated proteins were subsequently determined. It turned out that *Lm*NMT H398N:M420L:L421Q was no longer catalytically active while *Hs*NMT1 N473H:L495M:Q496L was fully functional with a similar  $K_m$  as the wild type (w. t., Supporting information Table S 3). For the latter enzyme, the  $K_i$  values of the unselective inhibitors **1-3** and the selective compound **5** were only slightly altered compared to w. t. *Hs*NMT1 (Table 2). In contrast, the  $K_i$  of the previously selective compound **4** decreased from 428.2 nM to 18.4 nM, compromising its selectivity.

Table 2: Inhibition constants of compounds **1-5** against *Lm*NMT and *Hs*NMT1 active site mutants. All measurements were performed at least in duplicates.  $K_i$  values were calculated from  $IC_{50}$  and  $K_M$  values using the Cheng-Prusoff equation. (n.d. = not determined due to inactive enzyme)

Compound	$K_i$ [nM]						
	<i>Lm</i> NMT	<i>Hs</i> NMT1	<i>Lm</i> NMT	<i>Lm</i> NMT	<i>Lm</i> NMT	<i>Hs</i> NMT1	<i>Hs</i> NMT1
	H398N:M420L: L421Q	N473H:L495M: Q496L	H398N	M420L	L421Q	L495M	Q496L
<b>1</b>	n.d.	20.1 ± 4.3	13.9 ± 3.2	n.d.	41.3 ± 8.6	4.0 ± 2.8	42.9 ± 16.4
<b>2</b>	n.d.	6.2 ± 1.9	8.5 ± 3.6	n.d.	6.0 ± 3.2	0.9 ± 0.5	40.2 ± 9.2
<b>3</b>	n.d.	25.8 ± 6.5	1.6 ± 0.8	n.d.	2.3 ± 0.8	19.1 ± 3.0	48.9 ± 20.1
<b>4</b>	n.d.	18.4 ± 3.6	1.3 ± 0.3	n.d.	36.6 ± 5.0	18.0 ± 5.1	2.8x10 <sup>2</sup> ± 61.4
<b>5</b>	n.d.	1.5x10 <sup>4</sup> ± 4.6x10 <sup>3</sup>	5.2x10 <sup>2</sup> ± 2.0x10 <sup>2</sup>	n.d.	2.6x10 <sup>3</sup> ± 1.0 x10 <sup>3</sup>	>1.5x10 <sup>4</sup>	>1.5x10 <sup>4</sup>

To further narrow down the crucial amino acids for the changes in enzyme activity and inhibition, NMT variants with only one altered residue were generated and tested. The substitution M420L in *Lm*NMT resulted in an inactive enzyme. Altering the corresponding residue in *Hs*NMT1 (L495M) led to stronger inhibition by all investigated sulfonamides with the largest change observed for **4** while changing the C-terminus (Q496L) had only a minor influence on the inhibition constant of these compounds (Table 2).

MD simulations of the w. t. and mutated apo structures as well as the ligand-bound complexes were performed to elucidate the molecular reason for the altered selectivity profile. All simulations showed a

high stability of the proteins without larger conformational changes within 50 ns (Table S 5). Careful inspection of the trajectories did not reveal any changes in the ligand interactions or the water network around the C-terminal residues that could explain the selectivity data. Therefore, we turned our attention to flexibility changes. For quantification of flexibility impairment, side-chain order parameters ( $S^2$ ) were calculated with  $S^2 = 1$  indicating low and 0 high flexibility.<sup>43,44</sup> Among all simulations, the most pronounced differences in  $S^2$  for binding site residues between NMTs for both species were observed for the C-terminal residues (Leu421 in *Lm*NMT and Gln496 in *Hs*NMT1, Figure 6, Table S 6 and Table S 7). For *Hs*NMT1, binding of the unselective compound **1** only slightly reduced the flexibility of the side chain of Gln496 compared to the apo structure ( $S^2 = 0.16$  and  $0.34$  in ligand free system (values from two independent simulations) and  $S^2 = 0.45$  in the complex with compound **1**). In contrast, when binding the selective inhibitor **4**, the flexibility of this side chain was reduced more drastically ( $S^2 = 0.75$ ). However, for the parasitic enzyme, smaller changes in  $S^2$  were observed for the C-terminal residue when binding this ligand. Similar results as for the parasitic enzyme were obtained from simulating the complex of compound **4** with the less selective mutated enzyme *Hs*NMT1 L495M. In this case, in the unbound form,  $S^2$  was determined to be  $0.58$  while it only increased to  $0.63$   $S^2$  in the compound **4**-bound form and decreased to  $0.20$  when the unselective compound **1** was bound. Taken together, based on the MD simulations, in *Hs*NMT1 the side chain of the C-terminal residue Gln496 is relatively flexible. This flexibility is most drastically impaired in the complex with the selective compound **4** (Figure 6). In contrast, the side chains of the C-terminal residues in *Lm*NMT and the *Hs*NMT1 L495M variant are already more rigid in the apo form and their flexibility is less effected by the bound ligands. Based on these results, one would expect that **4** binds less potently to *Hs*NMT1 than to *Hs*NMT1 L495M due to entropic reasons (Figure 4).

To back-up the MD simulations, ITC measurements with the mutated proteins were performed (Figure 4 c, Table S 4). For compound **1** binding to *Hs*NMT1 N473H:L495M:Q496L the enthalpic contribution was increased whereas the entropic contribution was decreased compared to binding to the w. t. enzyme. In

contrast, the gain in affinity (and loss of selectivity) of compound **4** binding to *HsNMT1* N473H:L495M:Q496L, as well as *HsNMT1* L495M was rather driven by an increase of entropy upon ligand binding, whereas the difference in affinity was more pronounced for the former protein.

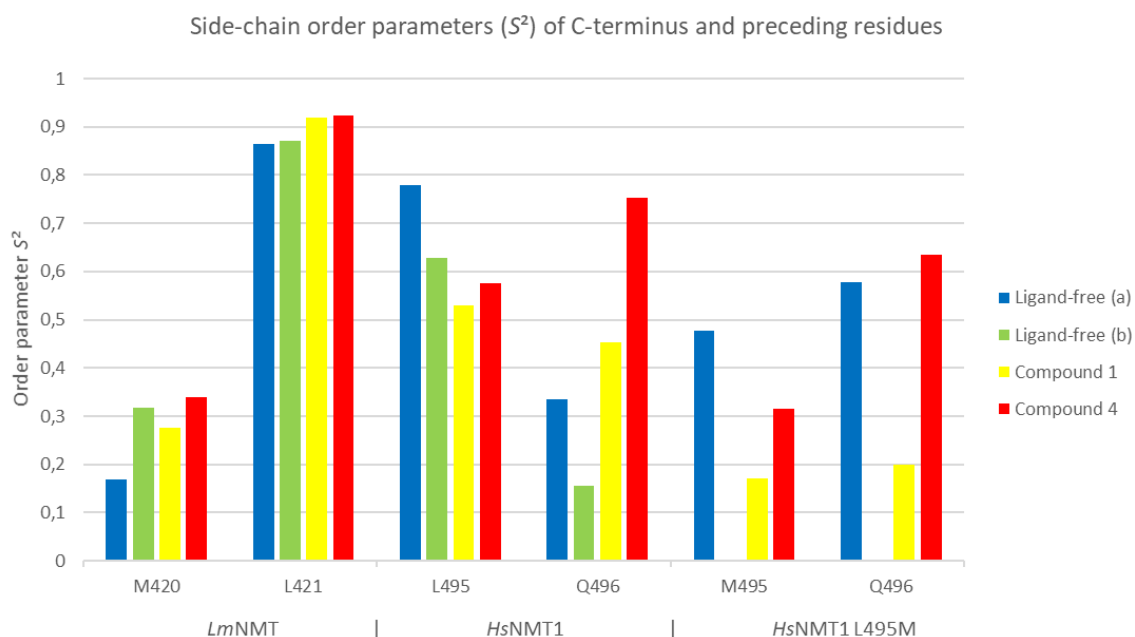


Figure 6: Side-chain order parameters of the C-terminal and preceding residues of *LmNMT*, *HsNMT1* and *HsNMT1* L495M. For unbound structures of *LmNMT* two simulations were run starting from PDB code 3H5Z (blue bar) and 4CGP (green bar), as well as for *HsNMT1* starting from PDB code 3IU1, chain B (blue bar) and 4C2Y, chain A (green bar). High values with a maximum of 1 indicate rigid side-chains, whereas low values represent high flexibility.

To investigate why the selective compound **4** influences the flexibility of the C-terminus stronger than the unselective ligand, we analyzed the available crystal structures. Superposition of the binding modes of **1** and **4** in NMT from both species revealed that the contact area between compound **4** and the C-terminus is larger than for compound **1** and the C-terminus (Figure 1, Figure 5a and b). Thus, the available space for movements of the C-terminus might become more restricted when **4** is bound compared to **1**.

#### Influence of Tyr217/296 on selectivity

In contrast to compound **4**, the affinity of compound **5** was not affected when swapping the binding sites of *Lm*NMT and *Hs*NMT1 (Table 1 and Table 2). Compared to the sulfonamides **1-4**, compound **5** lacks a hydrogen-bond acceptor for Ser330/405 and the orientation of the hydrophobic linker moiety within the binding site differs (Figure 3 a). Furthermore, in the complex structure with compound **5**, Tyr217 adopts an open conformation whereas Tyr217 is found in a closed conformation when binding the sulfonamides. Through the open conformation a sub-pocket is enlarged, which becomes partially occupied by the indole core of compound **5**. Therefore, it was hypothesized earlier that adopting the open conformation is energetically more favorable in *Lm*NMT compared to *Hs*NMT1 and, thus, causes selectivity of compound **5**.<sup>38</sup> However, recently non-selective *Ld*NMT inhibitors, which are believed to have similar affinity to *Lm*NMT and likewise bind to the open conformation, were found, defying this hypothesis.<sup>39</sup> Our MD simulations support the latter findings. In two independent MD simulations of unbound *Hs*NMT1, the open conformation was found in 93.2 and 72.3% of the time, respectively. In *Lm*NMT the open conformation was not favored compared to *Hs*NMT1 and occurred with similar frequency (81.2 and 75.8% of the time in two independent simulations of the unbound structure).

#### Influence of water molecules on selectivity

To identify the molecular reasons for selectivity of **5**, we turned our attention to the water network formed with the amino acids lining the binding site. Crystal structure analysis revealed differences between *Lm*NMT and *Hs*NMT1 in close proximity to Tyr217/296. As mentioned above, adopting the open conformation of Tyr 217/296 enlarges a sub-pocket. This sub-pocket is occupied by several water molecules (Figure 7 a). The depth of this pocket is limited by two residues that differ between the NMTs (Met377 and Val378 in *Lm*NMT; Ala452 and Val453 in *Hs*NMT1, Figure 7).<sup>39</sup> In this pocket, a specific water molecule, here referred to as W1 (corresponding for example to W629 in the PDB structure 3IU1, chain B), was found in crystal structures of unbound *Hs*NMT1, but not of unbound *Lm*NMT (the closest water



molecules are 2.6 and 2.8 Å away from W1 in the parasitic enzyme). However, in crystal structures with bound sulfonamides, the inhibitors trap Tyr217/296 in the closed conformation and the water molecule W1 is found in *LmNMT* as well as *HsNMT1* (Figure 7 b). When compound **5** is bound to *LmNMT*, this ligand occupies space close to the W1 hydration site. As a consequence, a water molecule is found at a slightly shifted position (Figure 7 c). Based on these observations, we hypothesized that W1 is causing selectivity of compound **5** for the following reason. W1 appears to be stronger bound to *HsNMT1* than to *LmNMT* and as binding of **5** requires displacing this water molecule (or alternatively adopting a different binding mode in *HsNMT1*), **5** should bind stronger to *LmNMT* than to *HsNMT1*. As we could not obtain a crystal structure confirming the binding mode of **5** in *HsNMT1*, we turned to the MD trajectories to investigate this hypothesis further.

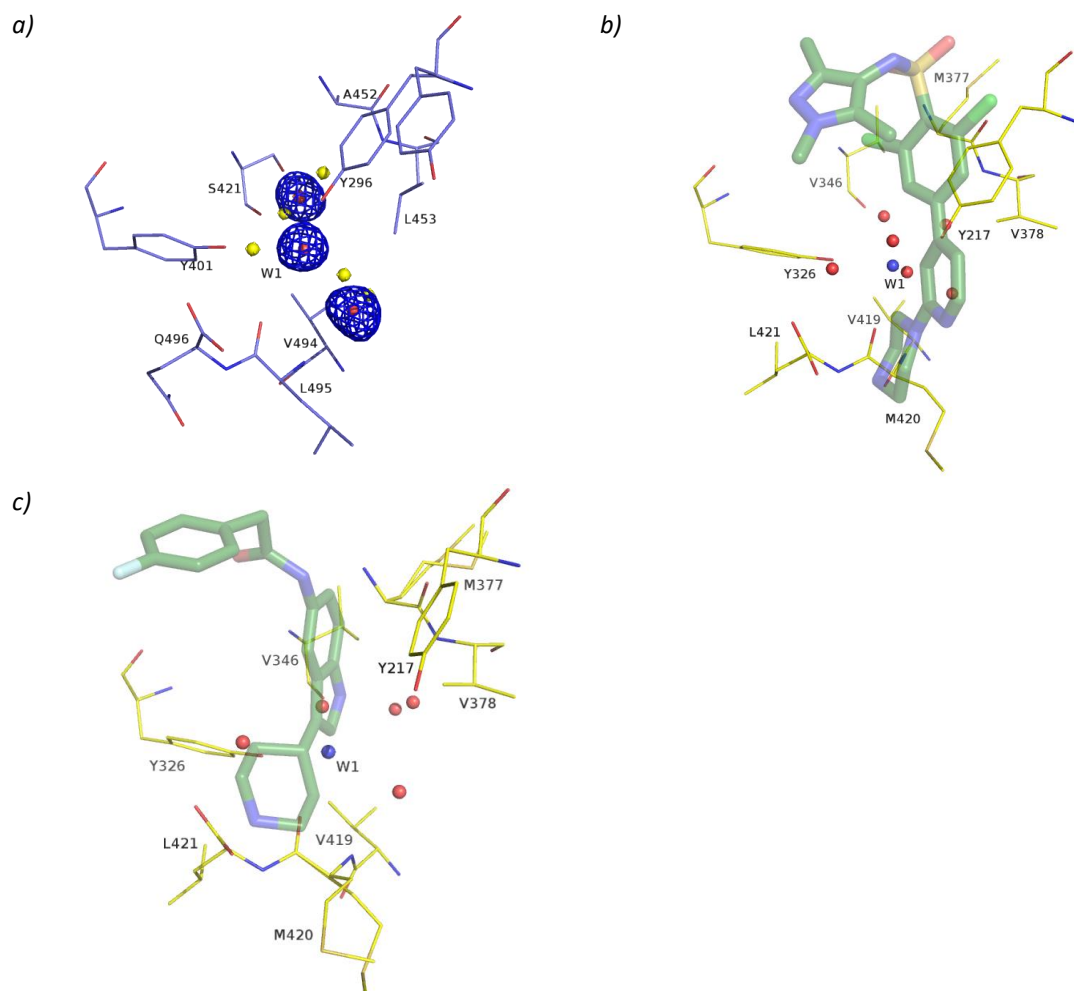


Figure 7: Illustration of the sub-pocket containing the water molecule W1 close to Y296 in *HsNMT1* a) Superposition of the sub-pocket containing W1 (PDB code 3IU1) with the positions of the water molecules found in this area in the *LmNMT* ligand-free crystal structure (yellow spheres, PDB code 3H5Z). The  $F_o - F_c$  electron density (countered at  $3\sigma$ ) for the water molecules in the *HsNMT1* structure (calculated with the water molecules in this area omitted from the structure), is also shown. b) Superposition of the sub-pocket close to Y217 (which corresponds to Y296 in *HsNMT1*) together with water molecules and the bound ligand 1 (PDB code 2WSA) with the position of the water molecule W1 from *HsNMT1* (blue sphere, PDB code 3IU1). c) Superposition of the sub-pocket close to Y217 (which corresponds to Y296 in *HsNMT1*) together with water molecules and the bound ligand 5 (PDB code 4CGN) with the position of the water molecule W1 from *HsNMT1* (blue sphere, PDB code 3IU1).

The analysis of the MD simulations supported the hypothesis that W1 is crucial for selectivity of compound **5**. At the W1 position, water density maps were calculated from the simulations for all investigated binding sites. A water molecule was considered to be present if the density was higher than 0.07 as described previously.<sup>45</sup> A water molecule was found in all simulations of *HsNMT1* with and without the sulfonamides **1** – **4** and in all simulations with these ligands bound to *LmNMT*, but not in simulations of apo *LmNMT*. Calculation of the free energy change of W1 relative to the bulk with the SPAM approach also indicated that W1 binding is more favored in *HsNMT1* than in *LmNMT*.<sup>45</sup> For complexes with the same ligands lower  $\Delta G$  values for that explicit hydration site were constantly found in *HsNMT1* compared to *LmNMT* (Table S 8). Additionally,  $\Delta G$  of W1 slightly decreased in complexes with ligands compared to apo structures in *HsNMT1*. Further, in simulations of *LmNMT-5*, the binding mode of the ligand was stable and W1 was not present. In contrast, in two independent simulations of *HsNMT1-5* (one starting from a docking pose generated in the absence of W1, the other one using a binding mode generated by a direct transfer of the binding mode of compound **5** in complex with *LmNMT* to *HsNMT1*), a different binding mode for the ligand was found. Throughout the trajectories, the W1 hydration site was predominantly occupied and the ligand binding mode was shifted by  $2.1 \pm 0.3 \text{ \AA}$  compared to *LmNMT* (calculated as heavy-atom RMSD of representative snapshots – defined as lowest protein backbone RMSD compared to the average coordinates throughout the trajectory - relative to binding mode from PDB code 4CGN, Figure 8). In this altered conformation, the secondary amine of the piperidine moiety is shifted by  $2.9 \pm 0.5 \text{ \AA}$ , abolishing hydrogen-bonding to the C-terminus. As it is known that a hydrogen-bond of a secondary amine with the C-terminus strongly contributes to binding affinity, the loss of this interaction might explain the reduced affinity of compound **5** for *HsNMT1*.<sup>34,38</sup>

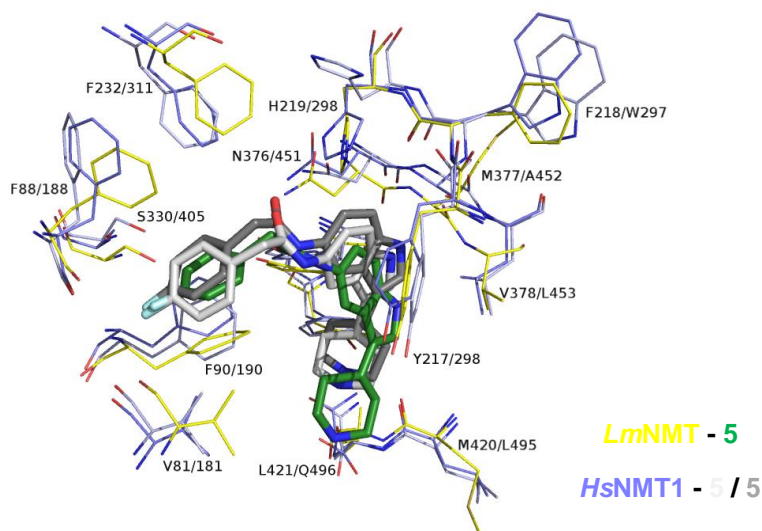


Figure 8: MD-derived binding modes of **5** binding to *LmNMT* and NMT1. For the *LmNMT*-**5** complex, the structure with PDB code 4CGN was used as starting structure for the simulations (ligand carbon atoms in green, protein carbon atoms in yellow). For the *HsNMT1*-**5** complexes, two different structures were used: 1) a structure obtained from docking (ligand carbon atoms in white, protein carbon atoms in blue) or 2) a structure generated by transferring the binding mode from *LmNMT* to *HsNMT1* (ligand carbon atoms in grey, protein carbon atoms in light blue). Residues are labeled as *LmNMT*/*HsNMT1*.

We further investigated the hypothesis that displacement of W1 contributes to selective binding of compound **5** by modulating the stability of W1 through the alteration of amino acids in its surrounding and characterizing the binding properties of the resulting mutants. The goal was to create an environment of W1 in *HsNMT1* that resembles that in *LmNMT*. To achieve this, first all residues in close proximity of W1 were swapped with those found in *LmNMT* at these positions. To provide space for these, additional residues had to be exchanged with smaller residues as well. Finally, further residues were exchanged to exclude the effect of other non-conserved residues on ligand selectivity. In total, up to eight residues in *HsNMT1* were altered (Figure 9). The binding site residues Asn473, Leu495 and Gln496 were included because of their proximity to the bound ligands and impact on the selectivity of compound **4**. Further, Ala452 was changed to the corresponding Met and Leu453 to the corresponding Val as these residues

define the depth of the sub-pocket to which W1 binds.<sup>39,46</sup> The second shell residue Leu462 was changed to Val to avoid a clash with the side chain of the introduced Met452. In addition, Trp297 was exchanged to Phe, as it is adjacent to Tyr296, which can adopt an open or closed conformation and thus determines the width of the entry of the sub-pocket. Finally, based on structure analysis it appeared possible that Arg295 influences the closing of the binding site upon ligand or substrate binding through hydrogen bonds with backbone oxygen atoms of Gly470, Asp471 and Gly472 (Figure S 1) and thus contributes to the selectivity of compound **5** by stabilization of the open conformation. Therefore, this residue was swapped as well with the corresponding Asn. In total, 12 different *HsNMT1* variants were generated, carrying up to 8 altered residues.

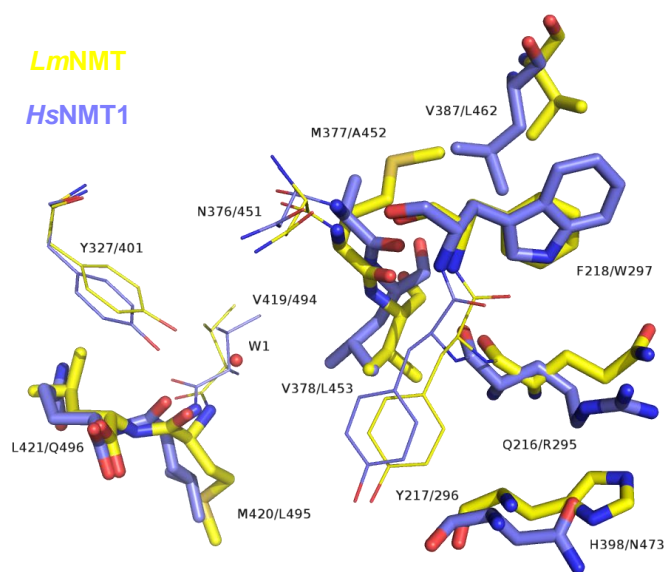


Figure 9: Residues in close proximity to W1 in *HsNMT1* (blue carbon atoms, PDB code 3IU1) superimposed with corresponding residues in *LmNMT* (yellow carbon atoms, PDB code 3H5Z). Amino acids depicted as sticks were exchanged in *HsNMT1* to the corresponding ones of *LmNMT* for the generation of *HsNMT1*-8x (R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L). Residues are labeled as *LmNMT*/*HsNMT1*.

The  $K_i$  value of the unselective compound **1** was only slightly altered for *HsNMT1* containing all eight swapped amino acids (*HsNMT1* 8x) compared to w. t. *HsNMT1* (26.4 vs. 31.6 nM, Table 3 and Table 1). Similar results were also observed for *HsNMT1* 8x with the unselective compounds **2** and **3** ( $K_i$  of 4.4 nM vs. 13.3 nM and 22.3 vs. 96.4 nM, resp.). As *HsNMT1*-8x includes the substitution L495M, increased affinity for compound **4** compared to w. t. *HsNMT1* was expected. This was indeed observed (79.4 nM vs. 430 nM), however not to the same extent as seen for *LmNMT* (2 nM) or *HsNMT1* L495M (18 nM). This indicates that the combination of multiple mutations can have opposing effects that partially cancel out each other. The most drastic change was observed for the selective compound **5**. Its  $K_i$  value was reduced from 15.4  $\mu$ M to 0.97  $\mu$ M which is very close to the  $K_i$  value for *LmNMT* inhibition ( $K_i = 0.98 \mu$ M). For none of the remaining *HsNMT1* variants bearing a subset of these eight mutated residues a  $K_i$  below 5  $\mu$ M was found for this compound.

In the subsequently determined crystal structure of *HsNMT1*-8x (Table S 2) electron density for W1 was only found in one of four chains of the asymmetric unit. The resolution of this structure (1.94 Å), was slightly lower than that of the w. t. apo-*HsNMT1* structure (1.42 Å). As water placement in electron density is resolution-dependent, we turned again to MD-simulations to investigate the water network further.<sup>47</sup> The simulations of *HsNMT1*-8x revealed a shift of the W1 coordinates by around 1 Å compared to the w. t. Collectively, the absence of the water molecule W1 in three out of four chains of *HsNMT1*-8x together with the shifted position in the MD simulations and the inhibition data obtained with the various *HsNMT1* variants strongly supports the hypothesis that W1 is crucial for selectivity of compound **5**.

Table 3: Inhibition constants of compounds **1** and **5** against *HsNMT1* mutants. *HsNMT1* 6x contains W297F:A452M:L453V:L462V:L495M:Q496L and *HsNMT1* 8x contains R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L. All measurements were performed at least in duplicates.  $K_i$  values were calculated from  $IC_{50}$  and  $K_M$  values using the Cheng-Prusoff equation.

NMT	K <sub>i</sub> [nM]	
	1	5
<i>HsNMT1</i> 8x	26.4 ± 2.4	9.7 x10 <sup>2</sup> ± 2.4 x10 <sup>2</sup>
<i>HsNMT1</i> A452M	79.2 ± 31.4	1.1x10 <sup>4</sup> ± 3.3 x10 <sup>3</sup>
<i>HsNMT1</i> L453V	2.3x10 <sup>2</sup> ± 2.0x10 <sup>2</sup>	> 1.5x10 <sup>4</sup>
<i>HsNMT1</i> A452M:L453V	3.0x10 <sup>2</sup> ± 96.5	> 1.5x10 <sup>4</sup>
<i>HsNMT1</i> A452M:L453V:L462V	1.2x10 <sup>2</sup> ± 40.5	1.0x10 <sup>4</sup> ± 5.5x10 <sup>3</sup>
<i>HsNMT1</i> A452M:L453V:L495M	2.8x10 <sup>2</sup> ± 1.5 x10 <sup>2</sup>	1.2x10 <sup>4</sup> ± 6.1x10 <sup>3</sup>
<i>HsNMT1</i> 6x	11.3 ± 10.5	8.7x10 <sup>3</sup> ± 2.9x10 <sup>3</sup>
<i>HsNMT1</i> R295Q	69.1 ± 32.2	> 1.5x10 <sup>4</sup>
<i>HsNMT1</i> R295Q:N473H:L495M:Q496L	38.1 ± 16.9	5.6x10 <sup>3</sup> ± 1.1 x10 <sup>3</sup>

## Profiling additional inhibitors

The generated *HsNMT1* variants allow profiling of further ligands to elucidate why they are selective. Interestingly, chloro-substituted derivatives of compound **6**, e. g. **6a** and **7**, are non-selective for *LdNMT* over *HsNMT1* while **6** is selective (Figure 2).<sup>39</sup> Due to the almost identical sequence of *LdNMT* and *LmNMT* (97.8 %) and comparable IC<sub>50</sub> values previously obtained for other inhibitors,<sup>38</sup> a similar selectivity pattern for *LmNMT* over *HsNMT1* was expected for these compounds. Further, **8** was also reported to be selective for *LmNMT*.<sup>37–39</sup> This data was confirmed by re-determining the inhibition profiles resulting in SI values of 46.8, 4.5 and 10.5 for **6**, **7** and **8**, respectively (Table 4).

To elucidate the reasons for selective inhibition, the compounds were subsequently tested against a selection of the generated *HsNMT1* variants (Table 4). *HsNMT1-8x* was inhibited more strongly by all compounds compared to w. t. *HsNMT1*, resulting in a loss of selectivity for compounds **6** and **8** (SI 0.5 and 0.1). When testing the compounds against *HsNMT1* L495M and *HsNMT1* N473H:L495M:Q496L, the affinity for **6** dropped again almost to the level of w. t. *HsNMT1* thus restoring selectivity. In contrast, affinity for **8** was still increased compared to w. t. *HsNMT1* maintaining the loss of selectivity (SI 2 and 1.5). Thus, **6** showed a similar inhibition profile in this panel as **5**, whereas **8** behaved more like compound **4** (Table 1, Table 2, Table 3).

The binding mode of compound **6** in complex with *LmNMT* and *HsNMT1* was confirmed using X-ray crystallography (Table S 2, Figure 5 c). The oxadiazole ring was opened in the structure as observed before for **6a** (PDB code 5A27). The ring opening is most likely caused by radiation damage during data collection (according to LC/MS data, **6a** was intact when used for crystallization trials).<sup>39</sup> Like **5**, compound **6** binds to the open Try217/296 conformation. Whereas in *HsNMT1-6* W1 is in the same position as in *HsNMT1*-apo, it is shifted by 0.5 Å in the *LmNMT-6* structure. As a consequence compared to *LmNMT-6*, the ligand in *HsNMT1* is slightly shifted (RMSD = 1.04 Å). MD simulations using intact structures of **6** also revealed a shifted binding mode (non-hydrogen RMSD = 1.3 Å for representative – closest to trajectory average



coordinates - snapshots, Figure S 3 a). For unselective compound **6a** similar binding modes were observed for both enzymes (RMSD = 0.91 Å, Figure S 3 b and c). Additionally, a W1 hydration site with a density of higher than 0.07 was found during simulations of both enzymes in complex with compound **6a**, but only for *HsNMT1* in complex with selective compound **6**. Taken together this suggests that selectivity of compound **6** is, like for compound **5**, caused by interference with water molecule W1.

Table 4: Inhibition constants and SI (in parentheses) relative to *LmNMT* of compounds **6-8**. All measurements were performed at least in duplicates.  $K_i$  values were calculated from  $IC_{50}$  and  $K_M$  values using the Cheng-Prusoff equation.

NMT	$K_i$ [nM] (SI)		
	<b>6</b>	<b>7</b>	<b>8</b>
<i>LmNMT</i>	70.8 ± 52.4	24.8 ± 2.3	2.8x10 <sup>2</sup> ± 0.5x10 <sup>2</sup>
<i>HsNMT1</i>	3.3x10 <sup>3</sup> ± 2.3x10 <sup>3</sup> (47.8)	1.1x10 <sup>2</sup> ± 0.3x10 <sup>2</sup> (4.5)	2.9x10 <sup>3</sup> ± 1.3x10 <sup>3</sup> (10.5)
<i>HsNMT1</i> 8x	33.0 ± 11.9 (0.5)	4.8 ± 1.4 (0.2)	29.2 ± 8.6 (0.1)
<i>HsNMT1</i> L495M	1.6x10 <sup>3</sup> ± 0.4x10 <sup>3</sup> (22.2)	36.3 ± 6.2 (1.5)	5.5x10 <sup>2</sup> ± 2.8x10 <sup>2</sup> (2.0)
<i>HsNMT1</i> N473H:L495M:Q496L	2.1x10 <sup>3</sup> ± 1.5x10 <sup>3</sup> (29.7)	42.5 ± 7.2 (1.7)	4.2x10 <sup>2</sup> ± 0.4x10 <sup>2</sup> (1.5)

The compounds **4** and **8** showed a similar inhibition profile in the panel. Like **4**, compound **8** places some atoms relatively close to the C-terminus (Figure 10). Thus, as observed for **4** (Figure 6), restricting the movements of the C-terminal residue in *HsNMT1*, but not in *LmNMT* or *HsNMT1* L495M and *HsNMT1* N473H:L495M:Q496L is likely the reason for selective inhibition of **8**.

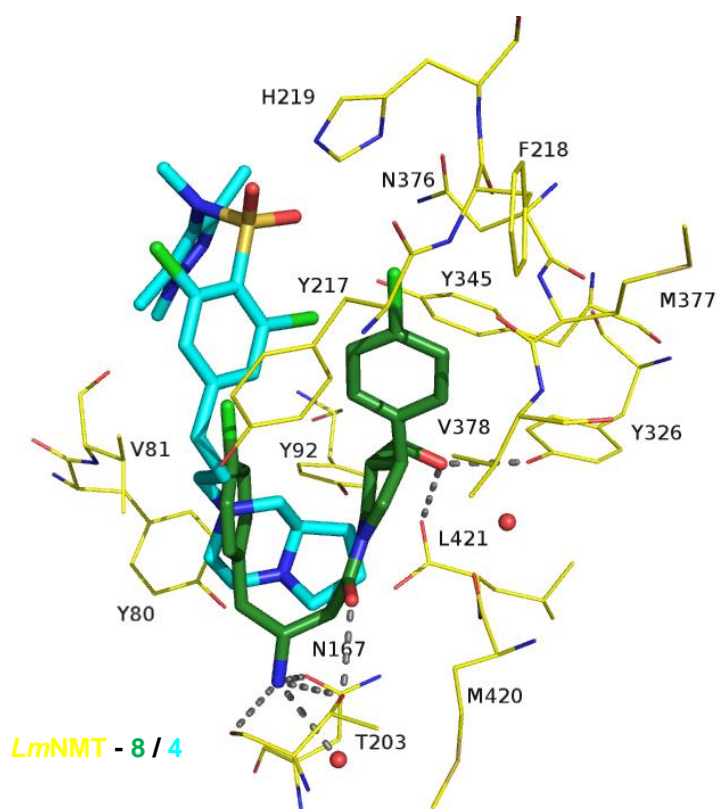


Figure 10: Binding mode of compound **8** (green carbon) in complex with *LmNMT* (yellow carbon, PDB code 4CGL) superimposed with binding mode of compound **4** (cyan carbon atoms, PDB code 6EU5).

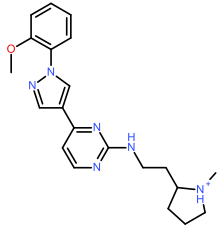
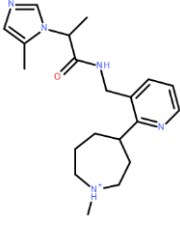
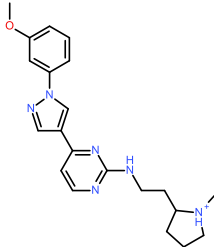
#### Virtual Screening for selective *LmNMT* inhibitors

A virtual screening for selective inhibitors was conducted based on the hypothesis that addressing the position of W1 in *HsNMT1* leads to selective inhibitors. The core feature of the postulated pharmacophore was the requirement of a group to displace or interfere with W1 in the off-target *HsNMT1* in a similar position as carbon C2 from compound **5** (grey sphere F1, Figure S 2). Further, interactions as observed for

the inhibitors **1-7** were included to obtain reasonable binding affinity. Molecules passing the pharmacophore query were subsequently docked into the binding site and 6 compounds were selected for testing (Table S 9).

An initial enzyme inhibition screen was performed with the purchased compounds against *LmNMT* and *HsNMT1* in concentrations of 10, 100 and 500  $\mu\text{M}$ . Compounds **9-11** revealed a concentration-dependent inhibition against *LmNMT*. Subsequently, the  $K_i$  values of these compounds against *LmNMT* were determined to be 19.6, 52.3 and 7.3  $\mu\text{M}$ , respectively (Table 5). To check for unspecific binding, the compounds were tested against the unrelated NS2B/NS3 protease of dengue II virus. No inhibition was found at 100  $\mu\text{M}$  suggesting no unspecific assay interference. Further, no inhibition at 100  $\mu\text{M}$  and only low inhibition at 500  $\mu\text{M}$  against *HsNMT1* was observed, suggesting that the compounds are specific and selective for *LmNMT*. To further support the hypothesis that selectivity is caused by W1-interference, compounds **9-11** were also tested against *HsNMT1-8x*. This *HsNMT1*-variant was inhibited with  $K_i$ -values of 16.7  $\mu\text{M}$  (cpd **9**), 14.5  $\mu\text{M}$  (cpd **10**) and 23.8  $\mu\text{M}$  (cpd **11**), indicating the same selectivity mechanism for these compounds as found for **5** and **6**.

Table 5: Inhibition constants and percentage of inhibition at 500  $\mu\text{M}$  of compounds **9-11** for *Lm*NMT and *Hs*NMT1. All measurements were performed at least in duplicates.  $K_i$  values were calculated from  $\text{IC}_{50}$  and  $K_M$  values using the Cheng-Prusoff equation.

NMT	$K_i$ [ $\mu\text{M}$ ] (% inhibition at 500 $\mu\text{M}$ )		
	<b>9</b>	<b>10*</b>	<b>11</b>
			
<i>Lm</i> NMT	19.6 $\pm$ 3.4 (91 %)	52.3 $\pm$ 7.3 (90 %)	7.3 $\pm$ 3.0 (97 %)
<i>Hs</i> NMT1	>> 40 (59 %)	> 200 (30 %)	>> 40 (59 %)
<i>Hs</i> NMT1-8x	16.7 $\pm$ 4.1 (98 %)	14.5 $\pm$ 4.4 (93 %)	23.8 $\pm$ 7.5 (92 %)

\* The compound that was docked contained a 2-methylimidazolyl group whereas in-house NMR analysis revealed that the purchased compound contained a 4-methylimidazolyl group.

## Discussion and Conclusions

A common goal in many drug discovery programs is to achieve high potency for the target and at the same time selectivity over off-targets. This is particularly challenging if the binding sites of the on- and off-targets are highly conserved. Here, we used *Lm*NMT and its human homologue *Hs*NMT1 as model systems for proteins with conserved binding sites to study selectivity-determining features in detail at the molecular level. Three key points stand out from this study: 1) For two different inhibitor series, two different selectivity-determining features were revealed. 2) Knowledge of these features could be exploited for the

design of selective inhibitors. 3) The chosen approach can serve as guideline on how to assess selectivity determining features in proteins with conserved binding site and to translate this knowledge into the design of selective inhibitors. In the following, we elaborate on each of these points.

Our strategy of combining X-ray crystallography, ITC, inhibition assay and MD simulations revealed two different selectivity-determining features within the inhibitor series under study. First, we revealed that also amino acids which orient their side chains away from the ligand, can contribute to selectivity. Second, we demonstrated that displacing or interfering with a water molecule from the *HsNMT1* binding site contributes to selective inhibition of compounds from the indole and oxadiazole series.

We started our investigations by changing the three non-conserved residues Asn473, Leu495 and Gln496 in *HsNMT1* to the corresponding residues in *LmNMT* and *vice versa*. As the side chains of these residues are oriented away from the bound ligands we did not expect a change in binding affinity for the investigated ligands binding to the modified pockets (Figure 1). In contrast to our expectations, the inhibition constants for the selective compound **4** dropped when binding to *HsNMT1* N473H:L495M:Q496L, reaching the level of binding to w. t. *LmNMT* and, thus, abolishing the selectivity, while the inhibition constants for the unselective compounds **1-3** were unaltered (Table 2). In fact, the exchange of L495M in *HsNMT1* was enough to obtain an *HsNMT1* variant which compound **4** inhibited with a similar potency as w. t. *LmNMT*. Both, ITC data and side chain order parameters calculated by MD simulations point in the direction that in this case entropic effects are causing selective inhibition (Figure 4, Figure 6). Notably, it is not the exchanged residue for which the flexibility was altered but the neighboring one. The aminoacylpyrrolidine compound **8** shows a similar selectivity pattern with respect to w. t. NMTs and the variants generated in this study (Table 4). The inhibitors **4** and **8** place ligand atoms most closely to the C-terminus among the studied compounds and, thus, likely influence most drastically the flexibility of these residues in *HsNMT1*, leading to selectivity for *LmNMT* over *HsNMT1* (Figure 10). We suggest therefore as one strategy for the design of selective *LmNMT* inhibitors to incorporate a bulky group

into the ligands to restrict the flexibility of the C-terminus in *HsNMT1*. MD simulations to calculate the order parameters in the presence of potential ligands can be used to guide the selection of suitable groups. Further, the role of W1 as an alternative selectivity-determining feature was revealed. Interestingly, the affinity of the indole derivative **5** was not affected by swapping the three non-conserved binding site residues (Table 2). It was initially hypothesized that the selectivity of this compound is caused by binding to an open binding site conformation (as defined by the rotamers of Tyr217 and 296, respectively), which might be differently accessible in the investigated NMT variants.<sup>38</sup> However, the subsequent identification of non-selective open-conformation-binding compounds (e. g. **6a** and **7**) and analysis of our MD simulations of NMT apo structures invalidated this hypothesis.<sup>39</sup> Thus, we turned our attention to the water molecule W1 which is in close proximity to Tyr217/296 and likely influenced by binding of compound **5** (Figure 9). In crystal structures and throughout MD simulation trajectories, at this position a water molecule was always found for *HsNMT1* and for pyrazole sulfonamides bound to *LmNMT*. However, no water molecule was present for apo *LmNMT* and when the selective compound **5** is bound to *LmNMT*. Thus, we hypothesized that the water molecule W1 is more strongly bound to *HsNMT1* than to *LmNMT* and that as a consequence ligands for which binding either requires displacement of W1 or adaption of an alternative binding mode in *HsNMT1*, have a higher binding affinity for *LmNMT* than for *HsNMT1*. This hypothesis was supported by generating *HsNMT1-8x* to which W1 bound more weakly based on evidence from crystal structure analysis and MD simulations and which in fact was more strongly inhibited by compound **5**, eliminating its selectivity over *LmNMT* (Table 3). Compound **6** showed a similar inhibition pattern against NMT variants as **5**, suggesting that W1 is also crucial for the selectivity of this ligand (Table 4). MD simulations further suggested that for both ligands the binding mode in *HsNMT1* is shifted compared to *LmNMT*, resulting in weaker interactions with the C-terminus, which likely explains their observed selectivity pattern (Figure 4 c, Figure 8 and Figure S 3 a). Interestingly, swapping only a subset of the 8 residues was not enough to abolish selectivity of **5** (Table 3). While the exchange R295Q alone is not

able to alter the  $K_i$  value of **5** in the desired direction, this residue appears to become important when combined with other replacements as the exchange is contained in the two NMT variants with the most pronounced change in  $K_i$  values (*HsNMT1-8x* and R295Q:N473H:L495M:Q496L). Furthermore, the exchanges A452M and L453V, which define the depth of the sub-pocket, in combination with L495M as present in both *HsNMT1-8x* and *HsNMT-6x*, which also had a slightly reduced  $K_i$  for compound **5**, seemed also to be important. Therefore, we hypothesize that a complex interplay between multiple residues is necessary to interfere with the affinity of W1. To further elucidate this interplay, more studies with additional NMT variants are required. We can only speculate about the biological role of W1. Due to its location close to the peptide binding groove, it might be involved in substrate specificity. However, additional experiments are necessary to investigate this further.

Knowledge of one of the selectivity determining features was subsequently exploited for the design of *LmNMT*-selective compounds. A predictive virtual screening was performed using a pharmacophore model that included an expansion into the W1 binding groove as a selectivity marker (Figure S 2). This screening resulted in the identification of three novel compounds **9-11** with  $K_i$  values of 7.3 to 52.5  $\mu\text{M}$  for *LmNMT* and selectivity over the human off-target (Table 5). To the best of our knowledge, this is the first example where selectivity of *LmNMT* over the human homologue could be achieved in a rational way. The activities of these compounds were restored against *HsNMT1-8x*, further supporting that W1 is crucial for selectivity.

Finally, the procedure presented herein can be considered as a guideline for the determination of selectivity-determining features of proteins with conserved binding sites and to translate this knowledge into the design of selective ligands. Of particular importance in this study was the interplay between a range of experimental and computational methods that highlighted subtle differences in ligand binding which led to a better understanding of selectivity. While we chose to focus on enzymatic assays, ITC experiments, X-ray crystallography and MD simulations in other cases where selectivity might be caused

by other factors, additional methods such as surface plasmon resonance to determine binding kinetics or NMR to experimentally determine protein and ligand dynamics might also be relevant. The critical role of water molecules for selective binding has already been revealed for other enzymes.<sup>1</sup> However, unlike in many other studies, here computational analysis to predict the importance of certain water molecules for selectivity were underpinned with experimental data. More surprising in this study is probably the fact that selectivity can also arise from restricting side chain movements, even if the side chain is oriented away from the ligand. The protein variants generated in this study could also be used to decipher the selectivity-determining features for structurally unrelated inhibitors. Such information can be very valuable for the design of ligands with improved affinity without the loss of selectivity. This study also showed that a complex interplay between several residues in and outside the binding site determines selectivity and more studies are needed to better understand this interplay.

## Experimental Section

### General Methods for synthesis and analytics

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a *Bruker* Fourier 300 (300 MHz) spectrometer. Chemical shifts ( $\delta$ ) are expressed in ppm. Signal splitting patterns are described as singlet (s), broad singlet (bs), doublet (d), triplet (t), quartet (q), multiplet (m), or combinations thereof.

Electron spray ionization (ESI) mass spectra were recorded on a *Agilent* 1100 series LC/MSD trap or *Waters*-ZQ2000 or *Waters*-Quattro-Micro, run in positive ion mode, using either methanol, methanol/water (95:5), or water/acetonitrile (1:1) and 0.1% formic acid as the mobile phase. The columns used were either a *Waters* Atlantis T3 C18 column, 150 mm  $\times$  4.6 mm, 3  $\mu$ m particle size or *Agilent* Poroshell 120 EC-C18 column 150 mm  $\times$  2.1 mm, 4  $\mu$ m particle size for analytics or a *Waters* Atlantis T3 100 mm  $\times$  30 mm C18 column with 5  $\mu$ m particle size for preparative separation. The following method was used for analytics: water + 0.1% HCOOH (mobile phase A) and acetonitrile (mobile phase B) 95:5 to 2:98 gradient over 7 min, and then held at for 2 min; flow rate 0.8 mL/min.



The method for preparative separation was: A:B of 90:10 to 2:98 gradient over 8 min at a flow rate of 50 mL/min with UV detection at 210 and 254 nm using a *Waters* 2996 Photodiode Array Detektor and ELSD detector *Waters* Micromass ZQ.

All compounds had a measured purity of greater than 95% on this analytical HPLC-MS system (TIC and HPLC UV). HPLC retention times and  $M^+$  data are given below to substantiate the purity and integrity of the compounds.  $^1\text{H}$  and  $^{13}\text{C}$  NMR also confirmed compound identity and purity (with other organic components being absent).

Thin layer chromatography (TLC) was carried out on Merck silica gel 60 F254 plates using UV light for visualization. TLC data are given as the  $R_f$  value with the corresponding eluent system specified in brackets.

Column chromatography was performed using *Merck* silica gel (0.030-0.063 or 0.015-0.040 mm) prepacked columns using the mobile phases stated below.

All reactions were carried out under dry and inert conditions unless otherwise stated.

Synthesis and characterization of small molecules

Inhibitors **1**, **2**, **3** and **5** were synthesized according to published procedures.<sup>34–36,38</sup> Compound **4** was synthesized as described below. Compounds **6–8** were kindly provided by Prof. Ed Tate and co-workers from Imperial College, London. Compounds **9–14** (Supporting information) were purchased from Chembridge/hit2lead or WuXi Apptec/LabNetwork. Identity and purity of these compounds was confirmed by LC-MS analysis and NMR (only for compounds **9–11**).

2,6-dichloro-4-[2-(piperazin-1-yl)pyridin-4-yl]-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzene-1-sulfonamide (**1**)

$^1\text{H}$ -NMR 400 MHz, Methanol- $d_4$ )  $\delta$  1.807 (3H, s), 2.043 (3H, s), 3.565 (2H, tbr), 3.627 (3H, s), 3.637 (2H, s.c.), 3.652 (2H, s.c.), 3.712 (2H, tbr), 6.975 (1H, dd,  $J=5.3$  Hz,  $J=1.4$  Hz), 7.117 (1H, dbr,  $J=1.4$  Hz), 7.892 (2H, s), 8.12 (1H), 8.223 (1H, d,  $J=5.3$  Hz);

$^{13}\text{C}$ -NMR (400 MHz, Methanol)  $\delta$  163.35, 40.99, 46.56, 46.11, 47.30, 161.27, 149.89, 112.89, 147.22, 106.94, 145.05, 130.89, 137.35, 137.24, 113.82, 140.41, 9.17, 36.48, 146.77, 10.75;

C<sub>22</sub>H<sub>24</sub>Cl<sub>2</sub>N<sub>6</sub>O<sub>2</sub>S; Mass: 495.42 g/mol

2,6-dichloro-4-[2-(4-methylpiperazin-1-yl)pyridin-4-yl]-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzene-1-sulfonamide (2)

<sup>1</sup>H-NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ 1.814 (3H, s), 2.048 (3H, s), 2.827 (3H, s), 3.225 (4H, tbr), 3.627 (3H, s), 3.893 (4H, tbr), 7.031 (1H, dd, J=5.3 Hz, J=1.1 Hz), 7.160 (1H, dbr, J=1.1 Hz), 7.890 (2H, s), 8.250 (2H, d, J=5.3 Hz);

<sup>13</sup>C-NMR (400 MHz, Methanol) δ 44.29, 54.44, 44.03, 160.47, 149.95, 113.53, 147.27, 106.89, 144.76, 130.84, 137.34, 137.28, 113.75, 140.21, 9.15, 36.45, 146.67, 10.70;

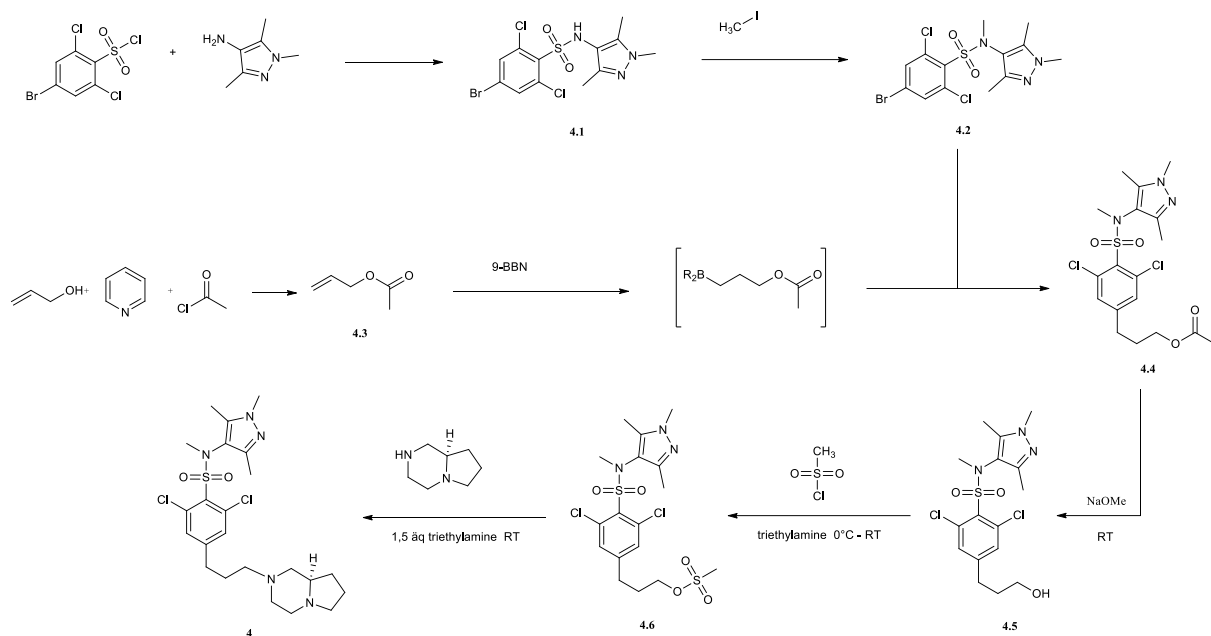
C<sub>22</sub>H<sub>26</sub>Cl<sub>2</sub>N<sub>6</sub>O<sub>2</sub>S; Mass: 509.45 g/mol

2,6-dichloro-N-methyl-4-[3-(piperazin-1-yl)propyl]-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzene-1-sulfonamide (3)

<sup>1</sup>H-NMR (400 MHz, Methanol-*d*<sub>4</sub>) δ 1.800 (3H, s), 2.002 (3H, s), 2.062 (2H, ddt, J=5.2 Hz, J=7.2 Hz, J=8.2 Hz), 2.722 (2H, dd, J=7.2 Hz, J=8.2 Hz), 3.213 (2H, dd, J=5.2 Hz, J=8.2 Hz), 3.381 (3H, s), 3.55 (2H, br), 3.56 (2H, br), 3.636 (s), 7.466 (2H, s);

<sup>13</sup>C-NMR (400 MHz, Methanol) δ 42.16, 49.83, 57.45, 25.79, 32.13, 148.36, 132.82, 137.48, 134.78, 40.32, 118.35, 140.86, 9.39, 36.46, 146.76, 11.04;

C<sub>22</sub>H<sub>29</sub>Cl<sub>2</sub>N<sub>5</sub>O<sub>2</sub>S; Mass: 474.45 g/mol



Scheme 1: Synthesis of compound **4**.

#### Synthesis of compound **4**

Compound **4** was synthesized according to Scheme 1.

#### 4-bromo-2,6-dichloro-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzenesulfonamide (**4.1**)

706 mg of 4-bromo-2,6-dichlorobenzoylchloride (2,18 mmol) were added in small portions to a solution of 420 mg 1,3,5-trimethyl-1H-pyrazole (2.18 mmol) in 7.5 mL of pyridine and the mixture was stirred at room temperature for 24 h. Then diethyl ether was added to precipitate side products which were filtered off. The filtrate was distributed between dichloromethane and aqueous sodium hydrogen carbonate solution, the water layer was reextracted with dichloromethane, the combined organic layers were washed with sodium hydrogen carbonate solution and dried over magnesium sulfate and concentrated under reduced pressure. The product was triturated with diethyl ether and ultrasonic irradiation to form 610 mg of the product as slightly yellow solid (67 %).

$^1\text{H NMR}$  (300 MHz,  $\text{DMSO-}d_6$ )  $\delta$  9.71 (s, 1H), 7.97 (s, 2H), 3.56 (s, 3H), 1.91 (s, 3H), 1.71 (s, 3H).

$^{13}\text{C}$  NMR (75 MHz, DMSO)  $\delta$  143.84, 137.16, 135.47, 135.39, 133.82, 125.87, 111.88, 39.52, 36.25, 10.49, 8.59.

#### 4-bromo-2,6-dichloro-N-methyl-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzene-sulfonamide (4.2)

To a solution of 455 mg 4.1 (1.1 mmol) in 4 mL DMF, 33 mg sodium hydride (as 55 mg 60% dispersion in paraffin, 1.4 mmol) were added in small portions at 0 °C. After gas evolution subsided, 82  $\mu\text{L}$  methyl iodide (1,3 mmol) were added and the reaction mixture was slowly warmed to room temperature and stirred overnight. Then the solvent was removed and the residue was distributed between 10 mL of dichloromethane and 20 mL of water. The organic layer was separated, dried over magnesium sulfate and concentrated under reduced pressure. The raw product was recrystallized from pentane to form 352 mg of a slightly yellow solid (0.82 mmol, 75 %).

$^1\text{H}$  NMR (300 MHz, DMSO- $d_6$ )  $\delta$  7.99 (s, 1H), 3.58 (s, 2H), 3.30 (s, 2H), 1.91 (s, 2H), 1.74 (s, 1H).

$^{13}\text{C}$  NMR (75 MHz, DMSO)  $\delta$  143.86, 137.76, 135.99, 134.38, 134.08, 126.41, 115.89, 39.86, 36.27, 10.85, 8.83.

#### Allyl acetate (4.3)

2 g of allylic alcohol (34 mmol) and 4.5 g triethyl amine (44 mmol) in 20 mL of dichloromethane under argon were cooled in an ice bath. Then a solution of 3.5 g acetyl chloride (45 mmol) in 10 mL dichloromethane was added dropwise with intense cooling. The mixture was refluxed for 1 h and then stirred at room temperature for 12 h. Then the reaction mixture was washed with diluted sodium hydrogen carbonate solution, followed by concentrated sodium hydrogen carbonate solution and water. The organic layer was dried over magnesium sulfate and evaporated to yield 1.4 g of the product (14 mmol, 41 %) as yellow oil, which had sufficient purity for subsequent reactions.

$^1\text{H}$  NMR (300 MHz, Chloroform- $d$ )  $\delta$  5.91 (ddt,  $J = 17.1, 10.4, 5.8$  Hz, 1H), 5.45 – 5.13 (m, 3H), 4.56 (dt,  $J = 5.7, 1.3$  Hz, 3H), 2.07 (s, 3H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  170.82, 132.29, 118.34, 65.27, 21.03.

#### 3-(3,5-dichloro-4-(N-methyl-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)sulfamoyl)phenyl)-propyl acetate (4.4)

100 mg allyl acetate **4.3** (1 mmol) and 244 mg dimeric 9-borabicyclo[3.3.1]nonane (2 mmol 9BBN) were dissolved in 2 mL THF and heated under argon to 65 °C for 1 h. After cooling to room temperature 428 mg **4.2** (1 mmol), 50 mg of tetrakis(triphenylphosphin)palladium(0) (0,04 mmol) and 1 ml water were added and the mixture was heated to 60 °C with microwave irradiation. Afterwards, the reaction mixture was concentrated under reduced pressure and directly subjected to column chromatography (flash column, eluent: cyclohexane : ethyl acetate [1:1]) to yield the product, still containing the Pd-catalyst as a pale oil, containing 190 mg product (0,42 mmol, 42 %, determined via NMR).

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 7.20 (s, 2H), 4.02 (t, *J* = 6.3 Hz, 2H), 3.62 (s, 3H), 3.34 (s, 3H), 2.61 (t, *J* = 7.7 Hz, 2H), 2.02 (d, *J* = 4.4 Hz, 6H), 1.89 (p, *J* = 6.9 Hz, 2H), 1.75 (s, 3H).

<sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ 170.91, 147.03, 145.24, 138.56, 133.30, 133.14, 131.93, 131.77, 131.31, 117.00, 63.08, 39.95, 36.47, 31.21, 29.21, 20.90, 11.00, 9.48.

#### 2,6-dichloro-4-(3-hydroxypropyl)-N-methyl-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)-benzenesulfonamide (4.5)

240 mg **4.4**, containing 190 mg mg propyl acetate compound (0,42 mmol) were stirred in 5 mL methanol containing 400 mg sodium methanolate overnight. The next day, 50 mg of sodium hydroxide were added and the mixture was stirred for further 2 h. Then the mixture was neutralized with diluted hydrochloric acid, concentrated under reduced pressure and mixed with 20 ml dichloromethane and aqueous sodium hydrogen carbonate solution. The organic layer was separated and the aqueous layer was extracted with dichloromethane two times. All organic layers were concentrated under reduced pressure and purified by column chromatography (flash column, gradient: cyclohexane : ethyl acetate [1:2] → pure ethyl acetate) to yield 123 mg of the product as pale oil (0,30 mmol, 71 %).

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 7.19 (s, 2H), 3.58 (s, 3H), 3.54 (t, *J* = 6.2 Hz, 2H), 3.31 (s, 3H), 2.66 – 2.57 (m, 2H), 1.99 (s, 3H), 1.85 – 1.73 (m, 2H), 1.72 (s, 3H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  148.13, 145.30, 138.63, 135.88, 132.80, 131.42, 116.99, 60.79, 39.89, 36.35, 32.98, 31.00, 10.89, 9.39.

3-(3,5-dichloro-4-(N-methyl-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)sulfamoyl)phenyl)-propyl methanesulfonate (4.6)

123 mg **4.5** (0.3 mmol, 1 äq.) in 5 mL dichloromethane were mixed with 61 mg (84  $\mu\text{L}$ , 0.6 mmol, 2 äq.) at 0 °C. Under further cooling with an ice bath, 51 mg of methyl sulfonyl chloride (35  $\mu\text{L}$ , 0.45 mmol, 1.5 äq.) in 1 mL dichloromethane were added and the mixture was stirred at room temperature for two days. After this time, 20 mL dichloromethane were added and this mixture was washed with water three times, dried over magnesium sulfate and concentrated under reduced pressure. The raw product was assumed to be formed quantitatively and was used for the subsequent reaction without further analysis or purification.

(R)-2,6-dichloro-4-(3-(hexahydropyrrolo[1,2-a]pyrazin-2(1H)-yl)propyl)-N-methyl-N-(1,3,5-trimethyl-1H-pyrazol-4-yl)benzenesulfonamide (**4**)

Presumably 145 mg of the previous formed **4.6** (0.3 mmol, 1 äq.) was dissolved in 5 mL acetonitrile and mixed with 90 mg triethyl amine (124  $\mu\text{L}$ , 0.9 mmol) prior the addition of 70 mg (R)-octahydropyrrolo[1,2-a]pyrazine (0.56 mmol, 1.9 äq.). The mixture was refluxed for 4.5 h, concentrated under reduced pressure and preliminary purified by column chromatography (flash column, gradient: isopropanol : chloroform [1:1]) to yield 120 mg of the product as brownish resin, which was further purified by HPLC to yield 85 mg of highly pure title compound **4** (0.17 mmol, 57 %).

$R_f$  = 0.35 (methanol + 1%  $\text{NH}_3$ )

Mass: 513.17 (calculated), found: 514.6 (MS ES+)

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ )  $\delta$  7.46 (s, 2H), 3.62 (s, 3H), 3.38 (d,  $J$  = 1.8 Hz, 3H), 3.05 – 2.85 (m, 3H), 2.81 – 2.65 (m, 3H), 2.39 – 2.28 (m, 2H), 2.26 – 2.09 (m, 1H), 2.08 – 2.03 (m, 3H), 2.00 (s, 4H), 1.90 – 1.79 (m, 2H), 1.77 (s, 4H), 1.75 – 1.62 (m, 2H).

$^{13}\text{C}$  NMR (75 MHz, Acetone)  $\delta$  150.15, 145.42, 138.81, 136.48, 132.80, 63.51, 58.66, 57.40, 54.02, 53.36, 52.40, 40.42, 36.81, 32.86, 28.51, 28.45, 22.21, 11.59, 9.58.

**2-(4-fluorophenyl)-N-[3-(piperidin-4-yl)-1H-indol-5-yl]acetamide (5)**

$^1\text{H}$ -NMR (300 MHz,  $\text{DMSO-}d_6$ )  $\delta$  9.07<sub>eq</sub> (1H<sub>ax</sub>, d, J=10 Hz), 8.87<sub>ax</sub> (1H<sub>eq</sub>, 1H<sub>ax</sub>, dt, J=10 Hz, J=8.8 Hz), 3.30<sub>eq</sub> (1H<sub>ax</sub>, dbr, J=12.3 Hz), 3.04<sub>ax</sub> (1H<sub>eq</sub>, 2H<sub>ax</sub>, "q"br, J~10 Hz, J=12.3 Hz), 2.07<sub>eq</sub> (1H<sub>ax</sub>, dm, J=13.5 Hz), 1.84<sub>ax</sub> (1H, 2H<sub>eq</sub>, 1H<sub>ax</sub>, dq, J~13 Hz, J=3.6 Hz), 3.03<sub>ax</sub> (1H<sub>ax</sub>, 1H<sub>eq</sub>, tt, J=13, J<3), 7.09 (1H, d, J=1.8 Hz), 10.87 (1H, d, J=1.8 Hz), 7.27 (1H, d, J=8.7), 7.13 (1H, dd, J=8.7 Hz, J=1.7 Hz), 7.95 (1H, d, J=1.7 Hz), 10.14, 3.64, 7.39 (2H, dd s.c., J=8.8 Hz, J=5.5 Hz,  $^3J_{\text{CF}}$ =7.8 Hz), 7.15 (2H, dd s.c., J=8.8 Hz, J=9.2 Hz,  $^2J_{\text{CF}}$ =20.9 Hz);

$^{13}\text{C}$ -NMR (300 MHz,  $\text{DMSO-}d_6$ )  $\delta$  43.41, 29.14, 30.74, 117.99, 121.48, 133.20, 109.35, 115.15, 130.74, 111.39, 125.63, 168.39, 42.22, 132.14, 130.93, 114.91, 161.06;

$\text{C}_{21}\text{H}_{22}\text{FN}_3\text{O}$ ; Mass: 351.42 g/mol

*Analytic data of compounds 9-14*

Compound **9**, mass: 379.48 (calculated), 379.4 (found)

Compound **10**, mass: 357.48 (calculated), 357.4 (found)

Compound **11**, mass: 379.48 (calculated), 379.4 (found)

Compound **12**, mass: 396.46 (calculated), 396.5 (found)

Compound **13**, mass: 379.48 (calculated), 379.4 (found)

Compound **14**, mass: 366.44 (calculated), 366.4 (found)

Protein purification

Plasmids coding for *Lm*NMT (11-421) and *Hs*NMT1 (115-496) with an N-terminal hexahistidine tag and a TEV protease cleavage site were purchased from addgene. The gene coding for *Hs*NMT1 R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L was synthesized by GenScript and cloned into

a pET-15b vector. All other mutations were generated using the Quick Change (Multi-)Site directed Mutagenesis kits (QIAGEN).

NMT proteins were expressed and purified as described previously.<sup>34,35</sup> Briefly, the enzymes were expressed in *E.coli* Rosetta2 cells using TB medium and IPTG induction. Harvested cells were resuspended in 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol and 5 mM imidazole. Cells were lysed by sonication after addition of lysozyme, DNase I and a protease inhibitor cocktail tablet. The cleared lysate was loaded onto a 5mL HisTrap HP column (GE Healthcare) using an ÄKTA purifier or ÄKTA pure system and eluted by a 5-250 mM imidazole gradient. Fractions corresponding to NMT were pooled and concentrated. Identity and purity were checked by SDS-PAGE gel with Coomassie blue staining.

#### Crystallization

For crystallization proteins were rebuffered into 50 mM sodium phosphate buffer pH 7.5 and concentrated to 10 mg/mL for *Lm*NMT and 5-8 mg/mL for *Hs*NMT1, respectively. NMTs were incubated with 1mM MyrCoA and 0.5-1 mM inhibitor on ice for 1h. Crystallization was performed similarly as done recently<sup>48</sup> using the hanging drop vapor diffusion method with 22-24% PEG4000, 5mM NiCl<sub>2</sub>, 100 mM sodium citrate and 2.5% glycerol at pH 4.5 or 25% MPEG2000, 200 mM KBr, 100 mM sodium citrate, 5% glycerol at pH 4.5 for *Hs*NMT1 and 25% PEG1500, 200 mM NaCl, 100 mM sodium cacodylate at pH 5.6 for *Lm*NMT.<sup>34</sup>

#### X-ray data collection

For *Hs*NMT1-6, the X-ray diffraction data were collected in-house at the Johannes Gutenberg University on the generator Bruker AXS Microstar-H with Mar Scanner 345 mm image plate detector. Crystals of *Hs*NMT1-2 and *Hs*NMT1-4 were measured on the synchrotron beamline ID29 at the European Synchrotron Radiation Facility (ESRF) in Grenoble with a Pilatus 6M (Dectris LTD) detector, and data for *Lm*NMT-4, *Lm*NMT-6 and *Hs*NMT1-8x were collected on beamline ID30A-1 / MASSIF-1<sup>49,50</sup> with a Pilatus3 2M (Dectris LTD) detector.



## Structure solution and refinement

For *HsNMT1-2* data were processed using XDS<sup>51</sup> and SCALA<sup>52</sup> and for *HsNMT1-4* using Mosflm<sup>53</sup> and SCALA. For *HsNMT1-6*, xia2 with XDS and SCALA was used for processing, whereas autoPROC<sup>54</sup> with XDS/XSCALE was utilized for *LmNMT-4*, *LmNMT-6* and *HsNMT1-8x*. The scaled data were phased with the PHASER molecular replacement technique using *HsNMT1-4* (for *HsNMT1-2* and *-6*), 3IWE (for *HsNMT1-4*), 3H5Z (for *LmNMT-4* and *LmNMT-6*) or 4C2Y (for *HsNMT1-8x*) as search model.<sup>55,56</sup> Automated refinement for *HsNMT1-2*, *-4*, and *-6* was carried out with the PDBredo server.<sup>57</sup> In addition, manual refinement of the obtained structure was performed by REFMAC5<sup>58</sup> and COOT<sup>59</sup> based on the 2FO-FC and FO-FC electron density maps. Refinement of *LmNMT-4*, *LmNMT-6* and *HsNMT1-8x* was done using phenix.refine from the PHENIX software suite, as well as COOT. The model geometry was validated through the RSCB Protein Data Bank Validate Service.<sup>60</sup> The crystallographic data for all structures are listed in Table S 2.

## Enzyme activity and inhibition assay

Enzyme activity and inhibition was determined with a fluorescence based assay<sup>41</sup> on a Tecan M200 infinite pro or Tecan Spark with 380 nm excitation and 470 nm emission wavelength. The assay was performed with 8 nM NMT in 110  $\mu$ L sodium phosphate buffer pH 7.8 containing 0.5 mM EDTA and 0.05% triton-X 100 using GSNKSKPK-amidation (pp60<sup>src</sup><sub>(2-9)</sub>) as substrate and the cofactor MyrCoA. 7-diethylamino-3-(4-maleimidyphenyl)-4-methylcoumarin (CPM) was used as fluorescent dye for free CoA detection after enzymatic reaction. For substrate  $K_M$  determination a 30 min continuous assay was performed at a saturation concentration of MyrCoA of 30  $\mu$ M and substrate concentration of 1 to 32  $\mu$ M and a CPM concentration of 20  $\mu$ M. The  $K_M$  was calculated from the slope of the initial 10 min of the reaction. For  $IC_{50}$  determination substrate and MyrCoA concentration were 4  $\mu$ M and CPM concentration was 8  $\mu$ M. Inhibitors were tested at ten different concentrations varying from 0.1 nM to 500  $\mu$ M depending on the inhibitor. In all assays, DMSO content was at 0.9%. The inhibition assay was run as an endpoint assay and stopped after 30 min with 60  $\mu$ L 0.1 M sodium acetate buffer pH 4.5. All assays were performed at least in

duplicates. Inhibition constants ( $K_i$ ) for comparability between different NMTs and mutants were calculated from  $IC_{50}$  and  $K_M$  values using the Cheng-Prusoff equation.<sup>61</sup>

## ITC

All ITC experiments were performed with a MicroCal VP-ITC instrument at 20 °C using a buffer with 50 mM sodium phosphate at pH 7.5 containing 0.5 mM EDTA and 1 mM beta-mercaptoethanol. Enzyme concentration of 10  $\mu$ M and inhibitor concentrations of 100  $\mu$ M were used. MyrCoA was present at 40  $\mu$ M for all titrations and the final DMSO content was at 0.5 % for compounds **1** and **5** and at 0.2 % for all other compounds tested. Affinity and binding thermodynamics of compound **5** binding to *Hs*NMT1 were determined using a displacement experiment. For that purpose, compound **1** was titrated against *Hs*NMT1 in the presence of 40  $\mu$ M compound **5**.  $K_d$  and thermodynamic profile of the displaced compound **5** were calculated from the observed  $K_d$  and  $\Delta H^\circ$  from the displacement experiment and the standard experiment for the higher affinity ligand **1**.<sup>62</sup>

## Molecular Dynamics Simulations

MD simulations were set up using the crystal structures available at the time (Table S 5). For protein-ligand complexes where no crystal structure was available, docking poses generated using LeadIT-2.1.6<sup>63</sup> were used as starting points. For compounds **2**, **4** and **6** the docking poses were later confirmed by crystallography for complexes with *Hs*NMT1. For *Hs*NMT1 which crystallized with two monomers in the asymmetric unit, chain A was used for MD simulations as biological assembly, except for 3IU1 where chain B was used because of missing residues in the A chain. Histidines were protonated at the epsilon nitrogen except for His12 in *Lm*NMT which was modelled as positively charged and His219/298 (*Lm*NMT/*Hs*NMT1) which was modelled as delta protonated due to the observed hydrogen-bond interactions with neighboring residues. For simulations of mutant proteins the amino acids were exchanged using PyMOL.<sup>64</sup> The simulation systems were subsequently built within VMD 1.9.2<sup>65</sup> keeping all crystallographic water molecules. Parameters for MyrCoA and inhibitors were generated with the CGenFF webservice.<sup>66,67</sup>

Parameters for the thioester moiety of the co-factor MyrCoA (which is lacking in CGenFF) were generated from QM-derived (HF/6-31G\*) parameters for bond lengths, angles, dihedrals and charges calculated for S-ethyl thiopropionate with Gaussian09<sup>68</sup> and fitted for compatibility with CHARMM using the force field toolkit (FFtk)<sup>69</sup> within VMD. Due to an instable binding mode throughout the trajectory of complexes with compounds **1** and **2** (data not shown), these compounds' angles and dihedrals were taken from MMFF94 and adapted to compatibility with CHARMM using the SwissParam-webserver.<sup>70</sup> The complexes were solvated in a TIP3P<sup>71</sup> water box exceeding the complex structure by 10 Å. The solvated complexes were charge neutralized with sodium ions and minimized over 2,000 time steps with NAMD 2.11<sup>72</sup> using the CHARMM36 force field.<sup>73,74</sup> The minimized complexes were subsequently equilibrated with harmonic constraints applied to all non-water atoms and the systems were heated from 100 to 300 K over 500 ps. Constraints were gradually released over the following 500 ps in a constant-volume box. Periodic boundary conditions were applied to the systems. The production runs were carried out for 50 ns using constant pressure and temperature.<sup>75,76</sup> Throughout equilibration and production the van der Waals cut-off was set to 14 Å and the particle mesh Ewald methodology for electrostatic interactions was applied. Time steps of 2 fs were used in combination with rigid bond lengths and trajectories were written every ps. All simulations were performed on the graphic processing units of the high performance cluster "Mogon" at the Johannes Gutenberg University of Mainz.

MD simulations were analyzed using VMD 1.9.2 and compatible scripts. Convergence of the simulations was checked using an RMSD-based approach.<sup>77</sup> Order parameters  $S^2$  were calculated as described previously using the isotropic reorientational eigenmode dynamics (iRED) approach<sup>78</sup> in cpptraj.<sup>79</sup> Water density maps and peaks were calculated with ccptraj<sup>79</sup> from AmberTools17 as well.<sup>80</sup> For quantitative analysis of hydration sites the SPAM approach<sup>45</sup> was used within cpptraj. The bulk water thermodynamics were derived from a 10 ns MD simulation of a 40 Å \* 40 Å \* 40 Å pure water box using the TIP3P water

model and the CHARMM36 force field. The calculated  $G_{\text{SPAM}}$  for bulk water was -29.60 kcal/mol,  $H_{\text{SPAM}}$  -17.61 +/- 5.43 kcal/mol and  $-TS_{\text{SPAM}}$  -11.98 kcal/mol.

#### Virtual screening and docking

A virtual small-molecule compound library was derived from the ZINC 15 database<sup>81,82</sup> using the following filters: reactivity = "anodyne" (to exclude reactive molecules and PAINS),<sup>83</sup> molecular weight (Mw) > 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 5,835,796 protomers, which were additionally filtered using MOE 2015.1001.<sup>84</sup> Molecules with the following properties were allowed to pass this additional filter step: Mw < 550 Da, H-bond acceptor count < 11, H-bond donor count < 6, number of rings 2-5, number of rotatable bonds < 11, aromatic atoms > 10 (corresponds to 2 or more aromatic rings) and exactly 1 basic atom. This resulted in 686,285 molecules.

Conformers for the subsequent pharmacophore query were generated using Omega with default settings (OMEGA 2.5.1.4: OpenEye Scientific Software, Santa Fe, NM).<sup>85</sup> The pharmacophore model contained the following features (Figure S 2): A hydrogen-bond acceptor to interact with the hydroxyl group of Ser330, adjacent to or part of an aromatic ring to interact with lipophilic residues (Phe88, Phe90, Phe232), an additional aromatic moiety to interact with Tyr217, and a basic center to interact with the C-terminal residue Leu421. Any atom at a similar position as C2 of the indole moiety of compound **5** was included to achieve selectivity for *Lm*NMT over *Hs*NMT1. 8,174 compounds passed this pharmacophore query.

The remaining molecules were docked without any restraints using LeadIT-2.1.6.<sup>63</sup> The docking strategy was validated by redocking of compounds **5**, **6a** and **7**, which resulted in RMSD values between docked and crystallographically determined binding modes between 1.05 and 1.70 Å. The obtained poses were visually inspected and the 200 highest scoring molecules predicted to form crucial interactions (H-bond with Ser330, direct or indirect ionic interaction with the C-terminal Leu421, and placement of an aromatic

moiety into the sub-pocket close to Tyr217) were rescored using the HYDE<sup>86</sup> scoring function. After commercial availability check, compounds **9-14** (Table S 7), which had all crucial interactions and the best HYDE scores, were purchased from Chembridge/hit2lead or WuXi Apptec/LabNetwork). These compounds have not been flagged as pan assay interference compounds (PAINS).

## Supporting Information

The following is supplied as Supporting information:

- tables providing ITC data, data collection and refinement statistics,  $K_M$  values for *Lm*NMT, *Hs*NMT1 and their variants, backbone 1D-RMSD values and side-chain order parameters from MD simulations, calculated binding energies of W1 binding to various NMT variants, and structures and scores of tested compounds from the virtual screening
- figures describing the interactions of Arg295 in *Hs*NMT1, the pharmacophore used for virtual screening and the binding modes of **6** and **6a**
- coordinates of the predicted binding modes of compounds **9- 11**
- Molecular Formula Strings

## Accession Codes

The described structures have been deposited in the PDB with the following codes: *Hs*NMT1-**2**: 6FZ3, *Hs*NMT1-**4**, 6FZ5, *Hs*NMT1-**6**, 6FZ2, *Hs*NMT1-8x: 6F56, *Lm*NMT-**4**: 6EU5, *Lm*NMT-**6**: 6EWF. The authors will release the atomic coordinates and experimental data upon article publication.

## Acknowledgments

This work was funded by the German Research Foundation (DFG, project 252061573 to R.B.). We are grateful to Prof. Ed Tate (Imperial College London) for samples of compound **6**, **7** and **8**, to Benjamin Merget for support with the MDs, to Sabine Maehrlein and Khanh Kim Dao for help with protein purification, to Jarl Underhaug for NMR analysis, and to Openeye for free software licenses. The diffraction

experiments were performed on beamline ID29 and ID30-A1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Matthew Bowler as local contact at the ESRF for providing assistance in using beamline ID30-A1 / MASSIF-1. For the computational work, we used resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway (project 246687) and the Mogon cluster of the Johannes Gutenberg University Mainz. For crystallization, we made use of the Facility for Biophysics, Structural Biology and Screening at the University of Bergen (BiSS) which has received funding from the Research Council of Norway (RCN) through the NORCRYST and NOR-OPENSREEN consortia (projects 245828 and 245922).

#### Abbreviations used

CPM 7-diethylamino-3-(4-maleimidylphenyl)-4-methylcoumarin

*Hs* *Homo sapiens*

*Hs*NMT1-8x *Hs*NMT1 carrying the point mutations

R295Q:W297F:A452M:L453V:L462V:N473H:L495M:Q496L

*Ld* *Leishmania donovani*

*Lm* *Leishmania major*

MyrCoA myristoyl coenzyme A

n.d. not determined due to inactive enzyme

NMT *N*-myristoyltransferase

SI selectivity index

$S^2$  order parameter

w. t. wild type

## References

- (1) Huggins, D. J.; Sherman, W.; Tidor, B. Rational Approaches to Improving Selectivity in Drug Design. *J. Med. Chem.* **2012**, *55* (4), 1424–1444.
- (2) Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; McDonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Iyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C. Structural Basis for Selective Inhibition of Cyclooxygenase-2 by Anti-Inflammatory Agents. *Nature* **1996**, *384* (6610), 644–648.
- (3) Rutaganira, F. U.; Barks, J.; Dhasan, M. S.; Wang, Q.; Lopez, M. S.; Long, S.; Radke, J. B.; Jones, N. G.; Maddirala, A. R.; Janetka, J. W.; El Bakkouri, M.; Hui, R.; Shokat, K. M.; Sibley, L. D. Inhibition of Calcium Dependent Protein Kinase 1 (CDPK1) by Pyrazolopyrimidine Analogs Decreases Establishment and Reoccurrence of Central Nervous System Disease by *Toxoplasma Gondii*. *J. Med. Chem.* **2017**, *60* (24), 9976–9989.
- (4) Kangas, E.; Tidor, B. Optimizing Electrostatic Affinity in Ligand-Receptor Binding: Theory, Computation, and Ligand Properties. *J. Chem. Phys.* **1998**, *109* (17), 7522–7545.
- (5) Radhakrishnan, M. L.; Tidor, B. Specificity in Molecular Design: A Physical Framework for Probing the Determinants of Binding Specificity and Promiscuity in a Biological Environment. *J. Phys. Chem. B* **2007**, *111* (47), 13419–13435.
- (6) Grimster, N. P.; Anderson, E.; Alimzhanov, M.; Bebernitz, G.; Bell, K.; Chuaqui, C.; Deegan, T.; Ferguson, A. D.; Gero, T.; Harsch, A.; Huszar, D.; Kawatkar, A.; Kettle, J. G.; Lyne, P.; Read, J. A.; Rivard Costa, C.; Ruston, L.; Schroeder, P.; Shi, J.; Su, Q.; Throner, S.; Toader, D.; Vasbinder, M.; Woessner, R.; Wang, H.; Wu, A.; Ye, M.; Zheng, W.; Zinda, M. Discovery and Optimization of a Novel Series of Highly Selective JAK1 Kinase Inhibitors. *J. Med. Chem.* **2018**, *61* (12), 5235–5244.
- (7) Humphrey, J. M.; Yang, E.; Ende, C. W. A.; Arnold, E. P.; Head, J. L.; Jenkinson, S.; Lebel, L. A.; Liras, S.; Pandit, J.; Samas, B.; Vajdos, F.; Simons, S. P.; Evdokimov, A.; Mansour, M.; Menniti, F. S. Small-Molecule Phosphodiesterase Probes: Discovery of Potent and Selective CNS-Penetrable

- Quinazoline Inhibitors of PDE1. *Medchemcomm* **2014**, *5* (9), 1290–1296.
- (8) Robinson, D. D.; Sherman, W.; Farid, R. Understanding Kinase Selectivity through Energetic Analysis of Binding Site Waters. *ChemMedChem* **2010**, *5* (4), 618–627.
- (9) Levinson, N. M.; Boxer, S. G. A Conserved Water-Mediated Hydrogen Bond Network Defines Bosutinib's Kinase Selectivity. *Nat. Chem. Biol.* **2013**, *10* (2), 127–132.
- (10) Robinson, D.; Bertrand, T.; Carry, J. C.; Halley, F.; Karlsson, A.; Mathieu, M.; Minoux, H.; Perrin, M. A.; Robert, B.; Schio, L.; Sherman, W. Differential Water Thermodynamics Determine PI3K-Beta/Delta Selectivity for Solvent-Exposed Ligand Modifications. *J. Chem. Inf. Model.* **2016**, *56* (5), 886–894.
- (11) Spyrakis, F.; Ahmed, M. H.; Bayden, A. S.; Cozzini, P.; Mozzarelli, A.; Kellogg, G. E. The Roles of Water in the Protein Matrix: A Largely Untapped Resource for Drug Discovery. *J. Med. Chem.* **2017**, *60* (16), 6781–6827.
- (12) Wylie, A. A.; Schoepfer, J.; Jahnke, W.; Cowan-Jacob, S. W.; Loo, A.; Furet, P.; Marzinzik, A. L.; Pelle, X.; Donovan, J.; Zhu, W.; Buonamici, S.; Hassan, A. Q.; Lombardo, F.; Iyer, V.; Palmer, M.; Berellini, G.; Dodd, S.; Thohan, S.; Bitter, H.; Branford, S.; Ross, D. M.; Hughes, T. P.; Petruzzelli, L.; Vanasse, K. G.; Warmuth, M.; Hofmann, F.; Keen, N. J.; Sellers, W. R. The Allosteric Inhibitor ABL001 Enables Dual Targeting of BCR-ABL1. *Nature* **2017**, *543* (7647), 733–737.
- (13) Aliagas-Martin, I.; Burdick, D.; Corson, L.; Dotson, J.; Drummond, J.; Fields, C.; Huang, O. W.; Hunsaker, T.; Kleinheinz, T.; Krueger, E.; Liang, J.; Moffat, J.; Phillips, G.; Pulk, R.; Rawson, T. E.; Ultsch, M.; Walker, L.; Wiesmann, C.; Zhang, B.; Zhu, B.-Y.; Cochran, A. G. A Class of 2,4-Bisanilinopyrimidine Aurora A Inhibitors with Unusually High Selectivity Against Aurora B. *J. Med. Chem.* **2009**, *52* (10), 3300–3307.
- (14) Baba, Y.; Hirukawa, N.; Tanohira, N.; Sodeoka, M. Structure-Based Design of a Highly Selective Catalytic Site-Directed Inhibitor of Ser/Thr Protein Phosphatase 2B (Calcineurin). *J. Am. Chem. Soc.* **2003**, *125* (32), 9740–9749.



- (15) Manas, E. S.; Unwalla, R. J.; Xu, Z. B.; Malamas, M. S.; Miller, C. P.; Harris, H. A.; Hsiao, C.; Akopian, T.; Hum, W.-T.; Malakian, K.; Wolfrom, S.; Bapat, A.; Bhat, R. A.; Stahl, M. L.; Somers, W. S.; Alvarez, J. C. Structure-Based Design of Estrogen Receptor-Beta Selective Ligands. *J. Am. Chem. Soc.* **2004**, *126* (46), 15106–15119.
- (16) Cleghorn, L. A. T.; Woodland, A.; Collie, I. T.; Torrie, L. S.; Norcross, N.; Luksch, T.; Mpamhanga, C.; Walker, R. G.; Mottram, J. C.; Brenk, R.; Frearson, J. A.; Gilbert, I. H.; Wyatt, P. G. Identification of Inhibitors of the Leishmania Cdc2-Related Protein Kinase CRK3. *ChemMedChem* **2011**, *6* (12), 2214–2224.
- (17) Yang, H.; Chennamaneni, L. R.; Ho, M. W. T.; Ang, S. H.; Tan, E. S. W.; Jeyaraj, D. A.; Yeap, Y. S.; Liu, B.; Ong, E. H.; Joy, J. K.; Wee, J. L. K.; Kwek, P.; Retna, P.; Dinie, N.; Nguyen, T. T. H.; Tai, S. J.; Manoharan, V.; Pendharkar, V.; Low, C. B.; Chew, Y. S.; Vuddagiri, S.; Sangthongpitag, K.; Choong, M. L.; Lee, M. A.; Kannan, S.; Verma, C. S.; Poulsen, A.; Lim, S.; Chuah, C.; Ong, T. S.; Hill, J.; Matter, A.; Nacro, K. Optimization of Selective Mitogen-Activated Protein Kinase Interacting Kinases 1 and 2 Inhibitors for the Treatment of Blast Crisis Leukemia. *J. Med. Chem.* **2018**, *61* (10), 4348–4369.
- (18) Noble, M. E. M.; Endicott, J. A.; Johnson, L. N. Protein Kinase Inhibitors : Insights into Drug Design from Structure. *Science* **2013**, *303* (5665), 1800-1805.
- (19) Tziridis, A.; Rauh, D.; Neumann, P.; Kolenko, P.; Menzel, A.; Bräuer, U.; Ursel, C.; Steinmetzer, P.; Stürzebecher, J.; Schweinitz, A.; Steinmetzer, T.; Stubbs, M. T. Correlating Structure and Ligand Affinity in Drug Discovery: A Cautionary Tale Involving Second Shell Residues. *Biol. Chem.* **2014**, *395* (7–8), 891–903.
- (20) Rudnick, D. A.; McWherter, C. A.; Rocque, W. J.; Lennon, P. J.; Getman, D. P.; Gordon, J. I. Kinetic and Structural Evidence for a Sequential Ordered Bi Bi Mechanism of Catalysis by *Saccharomyces Cerevisiae* Myristoyl-CoA:Protein N-Myristoyltransferase. *J. Biol. Chem.* **1991**, *266* (15), 9732–9739.
- (21) Wright, M. H.; Heal, W. P.; Mann, D. J.; Tate, E. W. Protein Myristoylation in Health and Disease. *J. Chem. Biol.* **2010**, *3* (1), 19–35.

- (22) Price, H. P.; Menon, M. R.; Panethymitaki, C.; Goulding, D.; McKean, P. G.; Smith, D. F. Myristoyl-CoA:Protein N-Myristoyltransferase, an Essential Enzyme and Potential Drug Target in Kinetoplastid Parasites. *J. Biol. Chem.* **2003**, *278* (9), 7206–7214.
- (23) Price, H. P.; Güther, M. L. S.; Ferguson, M. a J.; Smith, D. F. Myristoyl-CoA:Protein N-Myristoyltransferase Depletion in Trypanosomes Causes Avirulence and Endocytic Defects. *Mol. Biochem. Parasitol.* **2010**, *169* (1), 55–58.
- (24) Xu, M.; Xie, L.; Yu, Z.; Xie, J. Roles of Protein N -Myristoylation and Translational Medicine Applications. [Crit. Rev. Eukaryot. Gene Expr.](#) **2015**, *25* (3), 259–268.
- (25) Thinon, E.; Morales-Sanfrutos, J.; Mann, D. J.; Tate, E. W. N-Myristoyltransferase Inhibition Induces ER-Stress, Cell Cycle Arrest, and Apoptosis in Cancer Cells. *ACS Chem. Biol.* **2016**, *11* (8), 2165–2176.
- (26) Goldston, A. M.; Sharma, A. I.; Paul, K. S.; Engman, D. M. Acylation in Trypanosomatids: An Essential Process and Potential Drug Target. *Trends Parasitol.* **2014**, *30* (7), 350–360.
- (27) Tate, E. W.; Bell, A. S.; Rackham, M. D.; Wright, M. H. N-Myristoyltransferase as a Potential Drug Target in Malaria and Leishmaniasis. *Parasitology* **2013**, *144* (1), 37–49.
- (28) Wright, M. H.; Clough, B.; Rackham, M. D.; Rangachari, K.; Brannigan, J. A.; Grainger, M.; Moss, D. K.; Bottrill, A. R.; Heal, W. P.; Broncel, M.; Serwa, R. A.; Brady, D.; Mann, D. J.; Leatherbarrow, R. J.; Tewari, R.; Wilkinson, A. J.; Holder, A. A.; Tate, E. W. Validation of N-Myristoyltransferase as an Antimalarial Drug Target Using an Integrated Chemical Biology Approach. *Nat. Chem.* **2014**, *6* (2), 112–121.
- (29) Herrera, L. J.; Brand, S.; Santos, A.; Nohara, L. L.; Harrison, J.; Norcross, N. R.; Thompson, S.; Smith, V.; Lema, C.; Varela-Ramirez, A.; Gilbert, I. H.; Almeida, I. C.; Maldonado, R. A. Validation of N-Myristoyltransferase as Potential Chemotherapeutic Target in Mammal-Dwelling Stages of *Trypanosoma Cruzi*. *PLoS Negl. Trop. Dis.* **2016**, *10* (4), 1–20.
- (30) Ritzefeld, M.; Wright, M. H.; Tate, E. W. New Developments in Probing and Targeting Protein Acylation in Malaria, Leishmaniasis and African Sleeping Sickness. *Parasitology* **2018**, *145* (2), 157–

174.

- (31) Mousnier, A.; Bell, A. S.; Swieboda, D. P.; Morales-Sanfrutos, J.; Pérez-Dorado, I.; Brannigan, J. A.; Newman, J.; Ritzeveld, M.; Hutton, J. A.; Guedán, A.; Asfor, A. S.; Robinson, S. W.; Hopkins-Navratilova, I.; Wilkinson, A. J.; Johnston, S. L.; Leatherbarrow, R. J.; Tuthill, T. J.; Solari, R.; Tate, E. W. Fragment-Derived Inhibitors of Human N-Myristoyltransferase Block Capsid Assembly and Replication of the Common Cold Virus. *Nat. chem.* **2018**, *10* (6), 599–606.
- (32) Corpas-Lopez, V.; Moniz, S.; Thomas, M.; Wall, R. J.; Torrie, L. S.; Zander-Dinse, D.; Tinti, M.; Brand, S.; Stojanovski, L.; Manthri, S.; Hallyburton, I.; Zuccotto, F.; Wyatt, P. G.; Rycker, M. de; Horn, D.; Ferguson, M. A. J.; Clos, J.; Read, K. D.; Fairlamb, A. H.; Gilbert, I. H.; Wyllie, S. Pharmacological Validation of N-Myristoyltransferase as a Drug Target in *Leishmania donovani*. *ACS Infect. Dis.* **2019**, *5* (1), 111–122.
- (33) Schlott, A. C.; Holder, A. A.; Tate, E. W. N -Myristoylation as a Drug Target in Malaria: Exploring the Role of N -Myristoyltransferase Substrates in the Inhibitor Mode of Action. *ACS Infect. Dis.* **2018**, *4* (4), 449–457.
- (34) Frearson, J. A.; Brand, S.; McElroy, S. P.; Cleghorn, L. A. T.; Smid, O.; Stojanovski, L.; Price, H. P.; Guther, M. L. S.; Torrie, L. S.; Robinson, D. A.; Hallyburton, I.; Mpamhanga, C. P.; Brannigan, J. A.; Wilkinson, A. J.; Hodgkinson, M.; Hui, R.; Qiu, W.; Raimi, O. G.; van Aalten, D. M. F.; Brenk, R.; Gilbert, I. H.; Read, K. D.; Fairlamb, A. H.; Ferguson, M. A. J.; Smith, D. F.; Wyatt, P. G. N-Myristoyltransferase Inhibitors as New Leads to Treat Sleeping Sickness. *Nature* **2010**, *464* (7289), 728–732.
- (35) Brand, S.; Cleghorn, L. A. T.; McElroy, S. P.; Robinson, D. A.; Smith, V. C.; Hallyburton, I.; Harrison, J. R.; Norcross, N. R.; Spinks, D.; Bayliss, T.; Norval, S.; Stojanovski, L.; Torrie, L. S.; Frearson, J. A.; Brenk, R.; Fairlamb, A. H.; Ferguson, M. A. J.; Read, K. D.; Wyatt, P. G.; Gilbert, I. H. Discovery of a Novel Class of Orally Active Trypanocidal N-Myristoyltransferase inhibitors. *J. Med. Chem.* **2012**, *55* (1), 140–152.

- (36) Brand, S.; Norcross, N. R.; Thompson, S.; Harrison, J. R.; Smith, V. C.; Robinson, D. A.; Torrie, L. S.; McElroy, S. P.; Hallyburton, I.; Norval, S.; Scullion, P.; Stojanovski, L.; Simeons, F. R. C.; van Aalten, D.; Frearson, J. A.; Brenk, R.; Fairlamb, A. H.; Ferguson, M. A. J.; Wyatt, P. G.; Gilbert, I. H.; Read, K. D. Lead Optimization of a Pyrazole Sulfonamide Series of *Trypanosoma brucei* N-Myristoyltransferase Inhibitors. Identification and Evaluation of CNS Penetrant Compounds as Potential Treatments for Stage 2 Human African Trypanosomiasis. *J. Med. Chem.* **2014**, *57* (23), 9855–9869.
- (37) Bell, A. S.; Mills, J. E.; Williams, G. P.; Brannigan, J. a; Wilkinson, A. J.; Parkinson, T.; Leatherbarrow, R. J.; Tate, E. W.; Holder, A. a; Smith, D. F. Selective Inhibitors of Protozoan Protein N-Myristoyltransferases as Starting Points for Tropical Disease Medicinal Chemistry Programs. *PLoS Negl. Trop. Dis.* **2012**, *6* (4), e1625.
- (38) Brannigan, J. A.; Roberts, S. M.; Bell, A. S.; Hutton, J. A.; Hodgkinson, M. R.; Tate, E. W.; Leatherbarrow, R. J.; Smith, D. F.; Wilkinson, A. J. Diverse Modes of Binding in Structures of *Leishmania major* N-Myristoyltransferase with Selective Inhibitors. *IUCrJ* **2014**, *1* (Pt 4), 250–260.
- (39) Rackham, M. D.; Yu, Z.; Brannigan, J. A.; Heal, W. P.; Paape, D.; Barker, K. V.; Wilkinson, A. J.; Smith, D. F.; Leatherbarrow, R. J.; Tate, E. W. Discovery of High Affinity Inhibitors of *Leishmania Donovanii* N-Myristoyltransferase. *Medchemcomm* **2015**, *6* (10), 1761–1766.
- (40) Brand, S.; Wyatt, P.; Thompson, S.; Smith, V.; Bayliss, T.; Harrison, J.; Norcross, N.; Cleghorn, L.; Gilbert, I.; Brenk, R. N-Myristoyl Transferase Inhibitors. US patent 9156811B2 **2011**.
- (41) Goncalves, V.; Brannigan, J. a; Thinon, E.; Olaleye, T. O.; Serwa, R.; Lanzarone, S.; Wilkinson, A. J.; Tate, E. W.; Leatherbarrow, R. J. A Fluorescence-Based Assay for N-Myristoyltransferase Activity. *Anal. Biochem.* **2012**, *421* (1), 342–344.
- (42) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242.
- (43) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance

- Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **1982**, *104* (17), 4546–4559.
- (44) Stockmann, H.; Bronowska, A.; Syme, N. R.; Thompson, G. S.; Kalverda, A. P.; Warriner, S. L.; Homans, S. W. Residual Ligand Entropy in the Binding of P-Substituted Benzenesulfonamide Ligands to Bovine Carbonic Anhydrase II. *J. Am. Chem. Soc.* **2008**, *130* (37), 12420–12426.
- (45) Cui, G.; Swails, J. M.; Manas, E. S. SPAM: A Simple Approach for Profiling Bound Water Molecules. *J. Chem. Theory Comput.* **2013**, *9* (12), 5539–5549.
- (46) Yu, Z.; Brannigan, J. A.; Rangachari, K.; Heal, W. P.; Wilkinson, A. J.; Holder, A. A.; Leatherbarrow, R. J.; Tate, E. W. Discovery of Pyridyl-Based Inhibitors of *Plasmodium falciparum* N-Myristoyltransferase. *Medchemcomm* **2015**, *6* (10), 1767–1772.
- (47) Carugo, O.; Bordo, D. How Many Water Molecules Can Be Detected by Protein Crystallography? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55* (2), 479–483.
- (48) Thinon, E.; Serwa, R. A.; Broncel, M.; Brannigan, J. A.; Brassat, U.; Wright, M. H.; Heal, W. P.; Wilkinson, A. J.; Mann, D. J.; Tate, E. W. Global Profiling of Co- and Post-Translationally N-Myristoylated Proteomes in Human Cells. *Nat. Commun.* **2014**, *5*, 4919.
- (49) Svensson, O.; Malbet-Monaco, S.; Popov, A.; Nurizzo, D.; Bowler, M. W. Fully Automatic Characterization and Data Collection from Crystals of Biological Macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2015**, *71* (8), 1757–1767.
- (50) Bowler, M. W.; Nurizzo, D.; Barrett, R.; Beteva, A.; Bodin, M.; Caserotto, H.; Delagenière, S.; Dobias, F.; Flot, D.; Giraud, T.; Guichard, N.; Guijarro, M.; Lentini, M.; Leonard, G. A.; McSweeney, S.; Oskarsson, M.; Schmidt, W.; Snigirev, A.; Stetten, D. von; Surr, J.; Svensson, O.; Theveneau, P.; Mueller-Dieckmann, C. MASSIF-1. A Beamline Dedicated to the Fully Automatic Characterization and Data Collection from Crystals of Biological Macromolecules. *J. Synchrotron Radiat.* **2015**, *22* (6), 1540–1547.
- (51) Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 125–132.

- (52) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 235–242.
- (53) Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. IMOSFLM: A New Graphical Interface for Diffraction-Image Processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 271–281.
- (54) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data Processing and Analysis with the AutoPROC Toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 293–302.
- (55) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.
- (56) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (4), 352–367.
- (57) Joosten, R. P.; Long, F.; Murshudov, G. N.; Perrakis, A. The PDB\_REDO Server for Macromolecular Structure Model Optimization. *IUCrJ* **2014**, *1* (Pt 4), 213–220.
- (58) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 355–367.
- (59) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D*

*Biol. Crystallogr.* **2004**, *60* (12 I), 2126–2132.

- (60) Gore, S.; Sanz García, E.; Hendrickx, P. M. S.; Gutmanas, A.; Westbrook, J. D.; Yang, H.; Feng, Z.; Baskaran, K.; Berrisford, J. M.; Hudson, B. P.; Ikegawa, Y.; Kobayashi, N.; Lawson, C. L.; Mading, S.; Mak, L.; Mukhopadhyay, A.; Oldfield, T. J.; Patwardhan, A.; Peisach, E.; Sahni, G.; Sekharan, M. R.; Sen, S.; Shao, C.; Smart, O. S.; Ulrich, E. L.; Yamashita, R.; Quesada, M.; Young, J. Y.; Nakamura, H.; Markley, J. L.; Berman, H. M.; Burley, S. K.; Velankar, S.; Kleywegt, G. J. Validation of Structures in the Protein Data Bank. *Structure* **2017**, *25* (12), 1916–1927.
- (61) Cheng Y, P. W. Relationship between the Inhibition Constant ( $K_i$ ) and the Concentration of Inhibitor Which Causes 50 per Cent Inhibition ( $I_{50}$ ) of an Enzymatic Reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- (62) Zhang, Y. L.; Zhang, Z. Y. Low-Affinity Binding Determined by Titration Calorimetry Using a High-Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B. *Anal. Biochem.* **1998**, *261* (2), 139–148.
- (63) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261* (3), 470–489.
- (64) Schrödinger LLC. The PyMOL Molecular Graphics System, Version 1.7.2.1.
- (65) Humphrey, W.; Dalke, A.; Schulten, K. Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.
- (66) Vanommeslaeghe, K.; MacKerell, A. D.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. [\*J. Chem. Inf. Model.\*](#) **2012**, *52* (12), 3144–3154.
- (67) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52* (12), 3155–3168.
- (68) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani,

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

- (69) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid Parameterization of Small Molecules Using the Force Field Toolkit. *J. Comput. Chem.* **2013**, *34* (32), 2757–2770.
- (70) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* **2011**, *32* (11), 2359–2368.
- (71) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926.
- (72) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.
- (73) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. A Program for Macromolecular Energy Minimisation and Dynamics Calculations. *J. Comp. Chem.* **1983**, *4* (2), 87–217.
- (74) Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145.
- (75) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem.*



*Phys.* **1994**, *101* (5), 4177–4189.

- (76) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621.
- (77) Koukos, P. I.; Glykos, N. M. On the Application of Good-Turing Statistics to Quantify Convergence of Biomolecular Simulations. *J. Chem. Inf. Model.* **2014**, *54* (1), 209–217.
- (78) Prompers, J. J.; Brüschweiler, R. General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *J. Am. Chem. Soc.* **2002**, *124* (16), 4522–4534.
- (79) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095.
- (80) Case, D. A.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco. **2017**.
- (81) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768.
- (82) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55* (11), 2324–2337.
- (83) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53* (7), 2719–2740.
- (84) Molecular Operating Environment (MOE), 2015.01; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2018**.

- (85) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50* (4), 572–584.
- (86) Reulecke, I.; Lange, G.; Albrecht, J.; Klein, R.; Rarey, M. Towards an Integrated Description of Hydrogen Bonding and Dehydration: Decreasing False Positives in Virtual Screening with the HYDE Scoring Function. *ChemMedChem* **2008**, *3* (6), 885–897.

Table of Content Graphic

