Molecular mechanisms of Phase II drug metabolizing enzymes SULT1 and their interactions with small ligands modeled through computational approaches

Ph.D. thesis

Tóth Dániel

Pharmaceutical Sciences and Health Technologies Division
Semmelweis University

Supervisors:

Erika Balog, PhD.

Maria A. Miteva, PhD habil.

Official reviewers:

Nathalie Lagarde, PhD habil.

Zoltán Gáspári, PhD habil.

Members of the Complex Examination Committee:

President: Romána Zelkó, DSc.

Members: András Czirók, PhD habil.

Christian Jelsch, PhD.

Budapest, 2024

1. Introduction

It is well known that drug discovery is a long and costly process, often taking over a decade and billions of dollars. Despite the many advancements in drug discovery, finding a good balance between efficacy and safety is a major challenge. Our research focuses on Sulfotransferases (SULTs), important members of drug metabolizing enzymes (DMEs). DMEs modulate the intracellular bioavailability and pharmacokinetics of drugs and other xenobiotics. Moreover, a high percentage of drug candidate failures are due to toxicity or undesirable drug-drug interactions, many of these are due to the inhibition of enzymes such as SULTs.

SULTs are Phase II metabolizing enzymes. They catalyze a sulfonation reaction, that transfer the sulfonate group from their universal cofactor 3′-Phosphoadenosine 5′-Phosphosulfate (PAPS) to an oxygen or nitrogen atom of the targeted small molecule (substrate), resulting in a sulfonated product. This process creates products with increased water solubility, making them easier to excrete in urine or bile, thereby preventing their buildup in the cells.

The first of the two isoenzymes focused in this research is SULT1A1, which is most abundant in the liver. It has a broad substrate specificity catalyzing the sulfonation of molecules

such as 17β-estradiol, 4'-hydroxy-nitrophenol and minoxidil with high affinity. The second isoenzyme, SULT1A3, on the other hand is highly specific for monoamine neurotransmitters, such as dopamine and adrenaline. Interest in the two isoenzymes arises from their structural similarity (93% sequence identity) even though they sulfonate some very different substrates.

2. Objectives

Objective1: Exploration of structural dynamics of SULT1A1 and SULT1A3 monomers

We aim to clarify the structural and dynamic differences of SULT1A1 and SULT1A3 monomers, with a particular focus on their conformational flexibility and substrate specificity using MD and MDeNM simulations.

Objective2: Structural bases of substrate specificity and selectivity

We aim to understand the impact of the structural variations of these isoenzymes, influencing the binding of different substrates and inhibitors, explaining their selectivity and specificity. A digital repository of substrates, inhibitors, selective substrates and selective inhibitors are gathered for docking into the conformation ensembles generated based on dynamics simulations.

Objective3: Impact of dimerization on SULT1A1 structure and function

We aim to elucidate the effect of dimerization on cofactor and substrate biding of SULT1A1. SULTs have been studied as monomers *in silico*, but their *in vivo* dimerization suggests functional importance. Therefore, we compare the behavior of SULT1A1 in monomeric and dimeric forms, while examining it at different occupancy states of PAPS and ligand in the dimer chains.

3. Methods

3.1. Structure preparation

Initial X-ray crystal structures for both SULT1A1 (PDBID 4GRA) and SULT1A3 (PDBID 2A3R) monomers were obtained from the Protein Data Bank. For both SULT1A1 and SULT1A3, an identical preparation protocol was followed: PAP, the inactive co-factor was replaced with PAPS by overlapping the shared atoms. PAPS as well as the substrates used during the simulations were parameterized using CGenFF. The online web tool CHARMM-GUI was used to generate a solvent box of TIP3 water molecules around the protein. Energy minimization was performed using the CHARMM C36m force field. In the case of the dimer systems, the same protocol was followed, however the SULT1A1 dimer crystal structure

(PDBID 2D06), which contains the canonical dimerization region, was used instead.

3.2. Parameters for MD and MDeNM simulations

All MD and MDeNM simulations were performed using NAMD with CHARMM C36m force field. Since SULT1A1 exhibited larger rigidity than SULT1A3 during classical MD simulations MDeNM simulation was carried out to comprehensively map its conformational surface.

Enzyme+Co-factor+substrate	Type	Simulation time
SULT1A1	MD	3×1 μs
SULT1A1+PAPS	MD MDeNM	3×0.2/0.5/1 μs 240 x 0.2 ns
SULT1A1+PAPS+fulvestrant	MD	3×1 μs
SULT1A3+PAPS	MD	3×0.5 μs
SULT1A3+PAPS+dopamine	MD	3×0.5 μs
2SULT1A1	MD	3×1 μs
2SULT1A1+1PAPS	MD	3×1 μs
2SULT1A1+2PAPS	MD	3×1 μs
2SULT1A1+2PAPS+2fulvestrant	MD	3×1 μs
2SULT1A1+2PAPS+1fulvestrant	MD	3×1 μs

Table 1. Summary of performed simulations

3.3. Molecular Docking

A clustering approach was implemented based on the results of the MD and MDeNM simulations to create a conformational ensemble that represents the entirety of the mapped conformational space while significantly reducing the number of conformations for docking. Docking experiments were performed using AutoDock Vina software. During the docking process of SULT1A1, the binding site residues K106 and F247 which exhibited rapid side-chain conformational changes during the MD and MDeNM simulations were handled flexibly; the remainder of the protein and the co-factor were maintained in a rigid state. In the case of SULT1A3, only the ligands were treated flexibly. Filtering was performed to ensure that the distances between the substrate acceptor hydroxyl or primary amino functional group and the sulfate group of the co-factor PAPS and the catalytic residue H108 fell within 5 Å of all the substrates docked into SULT1A1 and SULT1A3.

4. Results

4.1 The role of conformational dynamics in substrate specificity of SULT1A1and SULT1A3 monomers

The structural dynamics of the isoenzymes SULT1A1 and SULT1A3 were investigated through a series of all-atom MD

and MDeNM simulations. These simulations aimed to shed light on the differences in flexibility and conformational behavior between the two isoenzymes, particularly in relation to their binding pockets and overall protein structures. The Root Mean squared Deviation (RMSD) values of MD simulations for SULT1A1 and SULT1A3 showed substantial differences between the two isoenzymes. SULT1A1 exhibited a more rigid structure, therefore, we performed MDeNM simulations to wider explore its conformational surface. Indeed, the MDeNM conformations distribution was more dispersed. In contrast, SULT1A3 demonstrated a broader range of conformational flexibility, particularly in the case of the dopamine-less structure. To identify and characterize the structural elements, the root mean squared fluctuation (RMSF) of the Cα atoms were calculated. For SULT1A1 MD and MDeNM results showed a higher flexibility of the functional loops L1 and L3.

For SULT1A3, L2 and L3 exhibited significantly higher fluctuations than SULT1A1. Simulating SULT1A3 with the selective substrate dopamine diminished the flexibility of L2 and L3 significantly, suggesting that dopamine binding stabilizes these functional loops by restricting the conformational space available for substrate binding.

To gain a deeper understanding of the collective motions and conformational transitions in SULT1A1 and SULT1A3, principal component analysis (PCA) was performed on all three MD simulation trajectories. The analysis revealed that SULT1A1 sampled a narrow conformational space along the first two principal components, supporting its more rigid behavior. In contrast, SULT1A3, both in its ligand-free and dopamine-bound forms sampled a much broader conformational space. To complement the PCA, the distribution of conformations described by PCA in the free energy landscape (FEL) representation was also calculated. These results suggest that SULT1A1 exists primarily in a single, stable conformational state, with low-energy barriers, while SULT1A3 displayed distinct energy minima, with higher energy barriers.

To gain insight into the mechanisms of SULT1A1 and SULT1A3 ligand interactions, ensemble docking was performed. For SULT1A1, 131 previously collected substrates and inhibitors were docked into the binding pocket of the centroid conformations generated by MD and MDeNM after clustering. These scores for SULT1A1 revealed a strong correlation between the Radius of Gyration (RGYR) of the binding pocket and the interaction energy (IE) of larger ligands. More extended MDeNM conformations exhibited better

docking results. For SULT1A3, 143 substrates and inhibitors were docked. Interestingly, SULT1A3 showed considerably more open states of the binding pocket than SULT1A1, and small ligands had worse IE scores than the comparable ligands of SULT1A1.

To identify the key residues involved in ligand binding within the active sites of SULT1A1 and SULT1A3, clusters that accommodated the majority of substrates with good IE were selected. The analysis of competent docking positions revealed that phenolic groups in SULT1A1 such as F76, F84, and F247, helped to accommodate aromatic ligands. At the same time, SULT1A3 possess carboxylic groups D86 and E146 to stabilize the binding of catecholamine-like ligands such as dopamine, which would explain the SULT1A3 more selective nature.

4.2 Dimerization effects on SULT1A1 structural dynamics and ligand interactions

The dynamic behavior of SULT1A1 upon dimerization was investigated by comparing the apoenzyme, the PAPS-containing enzyme and the PAPS+fulvestrant-containing enzyme in both monomer and dimer forms. A comprehensive approach was employed, with the analysis of RMSD, RMSF and loop distances. These distances describe the openness of the loops of the active site. The distances L1-L2 and L1-L3 were used to

describe the extent of opening of the substrate binding gate. Moreover, the openness of the nucleotide binding gate was defined as the distance of L4-α15.

The results of the dimerization experiments indicated that, in the apo enzyme, there was an increase in the fluctuation and opening of the functional loops. The high flexibility of additional helix-loop regions indicated a less stable overall structure. Moreover, an asymmetric behavior was observed with regard to the two chains of the dimer. The monomer and dimer containing PAPS exhibited comparable behavior. The PAPS+fulvestrant-containing monomer and dimer exhibited an asymmetric opening of the L4.

Subsequently, the impact of ligand binding on the dimer was examined. Initially, the apo dimer was compared with the dimer containing one PAPS (in chain A) and with both chains containing PAPS. The results demonstrated that PAPS binding rigidifies the loop above PAPS, stabilizing the secondary structure of the whole enzyme and promoting the substrate gate opening towards the open conformations.

The L4 of the PAPS-less chain displayed large conformational shifts, spanning from a highly closed to a highly open states. The binding of the second PAPS resulted in the enzyme becoming more rigid, which in turn constrained both the population of L4

open structures and the conformational variety of the ligand binding gate.

To follow this allosteric effect in more detail we performed FDA calculations on the apo- and the 1PAPS dimer. First, to follow the effect of the perturbation caused by PAPS binding, we calculated the difference between the residue-based pairwise forces between the 1PAPS dimer and the apo dimer. Then, the residue based punctual stress was calculated by summing up the pairwise force differences sensed by a residue. The results of PAPS binding revealed that further regions involved in substrate binding in chain A exhibited considerable punctual stress, which could explain the larger opening of the substrate gate. Additionally, considerable punctual stress was observed on chain B at the PAPS binding site, which explains its larger conformational mapping.

The fulvestrant binding effects on the dimer demonstrated that fulvestrant-binding in chain A resulted in a shift of the ligand-binding gate and the L4 of the same chain towards the more open conformations. Interestingly, the fulvestrant-less B chain also exhibited more open conformations of the substrate binding gate. Moreover, additional helix loops demonstrated higher fluctuation upon fulvestrant binding.

The second fulvestrant binding resulted in a loosening of both chains of the enzyme, although the opening of L4 showed an asymmetric behavior. FDA calculation was also used to have a closer view of the allosteric effect of how 1fulvestrant binding to chain A affects gate opening of the ligand binding site of the fulvestrant-less chain B. Differences between the residue-based pairwise forces between the 2PAPS+1fulvestrant and the 2PAPS containing dimer were calculated, then summed up to calculate the residue based punctual stress. The analysis indicated both inter and intra-chain allostery, which appears to be a two-way street communication between the cofactor- and the ligand-binding sites. Specifically, if PAPS binds, the ligand-binding site will sense it, and vice versa: if a ligand binds, the PAPS-binding site is going to detect it.

5. Conclusions

In this comprehensive study, the dynamic behaviors of the sulfotransferase isoenzymes SULT1A1 and SULT1A3 were explored. MD and MDeNM simulations of SULT1A1 and SULT1A3 monomers were combined with ensemble docking of a collection of ligands. The advanced MDeNM simulations demonstrated better performance compared to classical MD simulations in generating a broader range of conformations, including "open-like" states of PAPS-bound SULT1A1. These

results demonstrated that SULT1A1 is capable of accommodating larger substrates, such as fulvestrant, regardless of co-factor occupancy.

The comparative analysis demonstrated that SULT1A3 displays enhanced structural flexibility in comparison to SULT1A1, particularly within the functional loops L2 and L3. The distinctive characteristics within their binding site were investigated. The results of the ensemble docking demonstrated key residues for SULT1A1 that can accommodate a larger variety of ligands. Additionally, the substrate specificity of SULT1A3 was also explained by its key residues.

The application of MD simulations with PAPS and fulvestrant to the dimer SULT1A1 revealed an increased flexibility upon dimerization. The presence of the large ligand fulvestrant further enhanced this flexibility, while the cofactor stabilized the enzyme. An interesting asymmetric behavior was observed between the two subunit chains in the apo form. The results obtained from the FDA indicated the presence of an allosteric two-way-street communication between both the active site substrate accommodation, and the cofactor binding site. This communication occurred both intra- and inter-chain, which may explain the experimentally observed but poorly understood behavior and role of dimerization in the SULT enzyme family.

6. Bibliography of the candidate's publications

Publications related to the thesis:

B. Dudas#, D. Toth #, D. Perahia, A. B. Nicot, E Balog, M. A. Miteva. "Insights into the Substrate Binding Mechanism of SULT1A1 through Molecular Dynamics with Excited Normal Modes Simulations." Scientific Reports 2021 11, no. 1 13129.
# contributed equally

D. Toth, B. Dudas, M. A. Miteva, E. Balog. "Role of Conformational Dynamics of Sulfotransferases SULT1A1 and SULT1A3 in Substrate Specificity." Int J Mol Sci 2023 24, no. 23.

In preparation:

D. Toth, B. Dudas, A. B. Nicot, M. A. Miteva, E. Balog "Dimerization effects on SULT1A1 structural dynamics and ligand interactions."

Conference publications related to the thesis:

D. Toth, B. Dudas, D. Perahia, E. Balog, M. A. Miteva "From Structural Dynamics to Machine Learning: SULT1 Isoenzymes substrate binding and selectivity" PhD Scientific Days 2022, Semmelweis University, 6-7 July 2022, Budapest

D. Toth, B. Dudas, D. Perahia, E. Balog, M. A. Miteva "*In silico* methods from structural analysis to virtual screening: SULT1

Isoenzymes substrate binding and selectivity" Reginal Biophysics Conference, 22-26 August, 2022, Pécs

D. Toth, B. Dudas, D. Perahia, E. Balog, M. A. Miteva "Mechanistic understanding of drug-drug interactions of Sulfotransferases" 4éme Forum Franco-Hongrois de la recherche scientifique, 13-14, Octobre 2022, Budapest

D Toth, B Dudas, D. Perahia, E. Balog, M. A. Miteva "Substrate Selectivity Of Sulfotransferase Isoenzymes, Results Based On Molecular Dynamics And Virtual Screening" GGMM 2023 - Young Modellers Conference 15-17 May 2023 Toulouse

D Toth, B Dudas, D. Perahia, E. Balog, M. A. Miteva "Substrate selectivity of Sulfotransferase Isoenzymes, results based on Molecular Dynamics and Virtual Screening" PhD Scientific Days 2023, Semmelweis University, 22-23 June 2023, Budapest

D. Toth, B. Dudas, D. Perahia, M. A. Miteva, Erika Balog "Role of conformational dynamics of sulfotransferases SULT1A1 and SULT1A3 in substrate specificity" European Biophysical Societies' Association (EBSA) Congress, 31 July -04 August 2023, Stockholm

D Toth, B Dudas, D. Perahia, E. Balog, M. A. Miteva "Substrate selectivity of Sulfotransferase Isoenzymes, results based on Molecular Dynamics and Ensemble Docking" Hungarian

Biophysics Society XXIX. Congress, 28-31 August, 2023, Budapest

D Toth, B Dudas, D. Perahia, E. Balog, M. A. Miteva "Conformational dynamics of sulfotransferase isoenzymes SULT1A1 and SULT1A3, a study in substrate specificity" Biophysical Society (BPS) 68th Annual Meeting, 10 -14 Februaury, 2024, Philadelphia

Publications not related to the thesis:

D. Toth, "Passages from the History of Molecular Dynamics." Kaleidoscope Művelődés-, Tudomány- és Orvostörténeti Folyóirat 12, no. 24 (2022): 295–304.

∑**IF: 9,897** (related to the thesis)