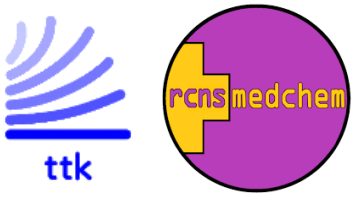
BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
FACULTY OF CHEMICAL TECHNOLOGY AND BIOTECHNOLOGY

Computational characterization of covalent enzyme inhibition

Thesis summary

Author: Levente Márk Mihalovits
Supervisor: György G. Ferenczy
Consultant: György Miklós Keserű



Medicinal Chemistry Research Group
Research Centre for Natural Sciences,
Loránd Eötvös Research Network

2021

1.INTRODUCTION AND THEORY

The daily usage of medication has become an essential part of our modern society and as a result of the continuous development in drug discovery more efficient pharmaceutics are produced, curing diseases, and improving life quality. With the exploding advancement of computer performance computer-aided drug design is now applied frequently in drug discovery campaigns, improving the effectiveness, and reducing the costs of the early phases. A good portion of these campaigns, resulted in officially approved drug compound, applied computational chemistry methods during the research process.[1,2]

The discovery of covalent drugs, a subset of inhibitor molecules that form a covalent bond with the targeted protein, also benefits from the computational methods. The purposeful search on covalent inhibitors since the early millennium raised the number of approved covalent drugs and they are now well represented in the treatment of multiple pathological conditions, most abundantly in cancer therapy; 10 out of the 14 FDA approved covalent drugs between 2011-2019 were cancer related medicines.[3]

Covalent inhibitors possess great advantages compared to the non-covalent counterparts, such as increased biochemical efficiency, longer residence time, lower dosage requirements, improved therapeutic index and the potential avoidance of some drug resistance mechanisms.[4]

The mechanism of action of covalent inhibition consists of two steps (**Eq. 1**), namely the molecular recognition and the chemical reaction. During the former the inhibitor molecule forms secondary interactions with the targeted protein, while the latter is referred to as the covalent step, in which the covalent bond between the reactive part of the molecule (warhead) and the targeted nucleophilic sidechain of the protein is formed.

$$\mathrm{E} + \mathrm{I} \overset{K_i}{\rightleftharpoons} \mathrm{E}\cdot\mathrm{I} \overset{k_{inact}}{\rightarrow} \mathrm{E}-\mathrm{I}\ (\mathrm{E}\text{: enzyme; I: inhibitor}) \qquad (1)$$

The two steps are characterized by the K_i equilibrium constant and for irreversible inhibition, the k_{inact} rate constant. Both binding steps may

[1] Talele T, Khedkar S, Rigby A.; *Curr Top Med Chem*, **2010,** *10*, 127–41.
[2] Sabe VT, Ntombela T, Jhamba LA, Maguire GEM, Govender T, Naicker T, et al.; *Eur J Med Chem*, **2021,** *224*, 113705
[3] Sutanto F, Konstantinidou M, Dömling A.; *RSC Med Chem*, **2020,** *11*, 876–84.
[4] Johnson DS, Weerapana E, Cravatt BF.; *Future Med Chem*, **2010**, *2*, 949–64

contribute significantly to the overall binding affinity, hence the complete description of the inhibition requires the evaluation of both constants. The determination of K_i and k_{inact} can be made experimentally, however, their computational evaluation is also possible by calculating the non-covalent binding free energy (ΔG) and the transition state free energy ($\Delta G^{\ddagger}$) and using **Eqs. 2** and **3**.

$$\Delta G = RT \ln(K_i) \tag{2}$$

$$\Delta G^{\ddagger} = -RT \ln\left(\frac{k_{inact}}{k_b T/h}\right) \tag{3}$$

where R is the universal gas constant, T is the absolute temperature, k_b is the Boltzmann constant and h is the Planck constant.

While K_i and k_{inact} are optimal for the comparison of experimental and calculated inhibition parameters, in specific cases only the IC_{50} — measuring the ligand concentration which halves the activity of the inhibited enzyme — is available experimentally. Nevertheless, the usage of IC_{50} in such comparisons is hampered by its time dependency in covalent inhibition.

The computational determination of the complete covalent inhibition, namely the calculation of ΔG and $\Delta G^{\ddagger}$ can be carried out by performing molecular dynamics simulations. The former can be estimated by thermodynamic integration or free energy perturbation calculations, while the determination of the latter requires the application of biased QM/MM based molecular dynamics simulations.[5]

In my thesis, I have combined the aforementioned methods and developed a complex, molecular dynamics based protocol in which both the non-covalent and covalent steps of the covalent inhibition are modeled and the related equilibrium and rate constants are calculated. The protocol was tested on a series of relevant protein targets, namely on MurA, KRAS, EGFR, ITK, BTK, BMX, immuno- and constitutive proteasome. The protocol is primarily designed for compound prioritization in covalent inhibitor design; however, I showed its applicability in mechanistic and selectivity studies, too.

2. METHODS

The evaluation of the covalent inhibition requires a wide range of computational chemistry tools. Quantum mechanics (QM) was applied in the

[5] De Vivo M, Masetti M, Bottegoni G, Cavalli A.; *J Med Chem,* **2016,** 59, 4035–61.

calculation of protein independent ligand reactivity prediction, while molecular mechanics (MM) based methods were used to construct initial protein-ligand complexes that serve as a starting point for the modeling of both the non-covalent recognition and the chemical bond formation. The binding poses of the examined compounds were estimated based on experimental protein-ligand X-ray structures and ligand docking.

Classical molecular dynamics simulations were carried out to refine and equilibrate the previously constructed complexes and to inspect the stability of the binding poses. Such simulations also give information about the stability of certain protein conformations, the dynamic properties of structural regions or the occurrence of certain events. The main drawback of unbiased simulations is the poor sampling of high energy regions. Therefore, the calculation of the binding events of covalent inhibition with higher energy barriers requires the application of enhanced sampling MD calculations, where an additional biasing potential is present during the simulations.[6]

We applied hybrid QM/MM based enhanced sampling MD methods, namely steered molecular dynamics (SMD) and umbrella sampling (US) to model the covalent reaction step. SMD introduces a time dependent biasing potential whose center is moved along a binding event related reaction coordinate, enabling the sampling of higher energy regions. US follows similar concept but instead of performing one simulation with moving biasing potential, it contains a series of MD simulations (windows) with constant bias potentials along the reaction coordinate. The energy profile as a function of the reaction coordinate, referred as the potential of mean force (PMF) curve can be generated based on both methods, however, we used SMD to generate starting structures to the subsequent US simulation windows evaluating the covalent binding step. The PMFs were constructed based on the biased distributions of the US simulation trajectories applying the weighted histogram analysis method (WHAM). The covalent reaction barriers ($\Delta G^{\ddagger}$) were determined according to the generated free energy profiles.

Thermodynamic integration (TI) is a method used to evaluate the free energy difference between two states. With specific thermodynamic cycles containing real chemical or alchemical processes the binding free energy differences between ligands ($\Delta\Delta G$) or in specific cases the binding free energy difference of the same ligand but in different proteins can be calculated. Performing TI simulations means the running of a series of MD simulations with increasing coupling parameter (λ) smoothly transforming the initial state

[6] Abrams C, Bussi G.; *Entropy*, **2013,** *16,* 163–99.

into the perturbed state. In our work the free energy difference (ΔΔG) for several ligand pairs were evaluated and an arbitrarily selected constant was added to obtain a set of absolute ligand binding free energies (ΔG).

3. RESULTS

3.1. REACTIVITY PREDICTION OF COVALENT INHIBITORS [T4]

At the initial stage of my work, we calculated the transition state energies of model systems, mimicking the covalent reaction between selected compounds and the most frequently targeted cysteine sidechain, substituted by a simple methyl-thiolate (**Figure 1**). We applied the B3LYP/6-311G**(d,p) level of theory and determined the transition barrier as the energy difference between the transition state and the reactant state structures. The calculated data were compared to the experimental GSH half-lifes, describing the ligand reactivity towards cysteine (**Figure 2**).

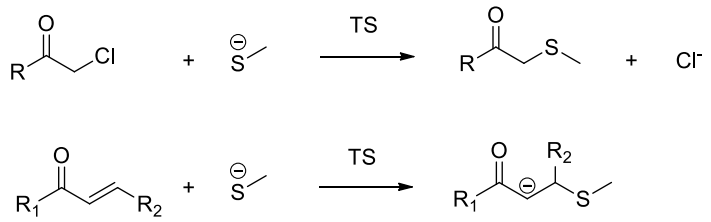


Figure 1. Modeled reactions

We found that better correlation is achieved if we consider molecules with different warheads separately. This indicates that the experimental reactivity of covalent inhibitors can be successfully predicted by quantum chemical reaction barrier calculations, however it is recommended to use it for molecules possessing the same warhead.

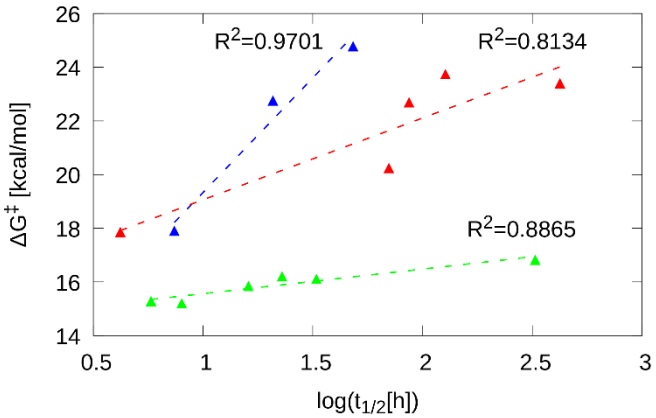


Figure 2. Calculated reaction barriers of model reactions as a function of experimental GSH half-lives. Different colors indicate compounds with different warheads: acrylamides (red), acrylesters (blue), chloro-acetamides (green)

3.2. Catalytic mechanism and covalent inhibition of MurA [T1]

The aim of the MurA related studies was to explore the mechanism of action of MurA and to develop a robust MD based method to evaluate covalent enzyme-ligand reactions. Though, the catalytic mechanism of MurA has been studied previously[7], no detailed investigations concerning the reaction energetics and the role of the active site residues during the catalysis have been performed. Former investigations proposed that Cys115 found in the highly flexible loop region take part in both catalytic and inhibition reaction mechanisms, forming covalent bond with phosphoenolpyruvate (PEP) or with cysteine targeting inhibitors, respectively.[8] As the nucleophilic attack of the cysteine sidechain is initiated by the deprotonation of the sulfur, we examined the possible mechanism of this process. We performed classical MD simulations to investigate the structure and dynamics of the enzyme of the Cys115 containing loop whose closed conformation forms the active site of the enzyme. We found that His394 approaches Cys115 in the loop closed conformation and adopts a position which is able to initiate the Cys115-His394 proton transfer. The unliganded enzyme structure has a loop open conformation, and the closure is proposed to happen upon the UDP-N-

[7] Gautam A, Rishi P, Tewari R.; *Appl Microbiol Biotechnol*, **2011,** *92*, 211–25
[8] Zhu JY, Yang Y, Han H, Betzi S, Olesen SH, Marsilio F, et al.; *J Biol Chem*, **2012,** *287*, 12657–67.

acetylglucosamine (UNAG) cofactor binding; therefore, both the catalytic mechanism and the covalent inhibition is initiated by UNAG binding. We evaluated the PMF curve of the proton transfer reaction (step I in **Figure 3**) and the 2D PMF surface of the Cys115-PEP reaction (steps II and III in **Figure 3**). The proton transfer turned out to be energetically feasible, while the two-step reaction of the PEP binding turned out to be consecutive, with the protonation of the double bond preceding the nucleophilic attack of the thiolate.

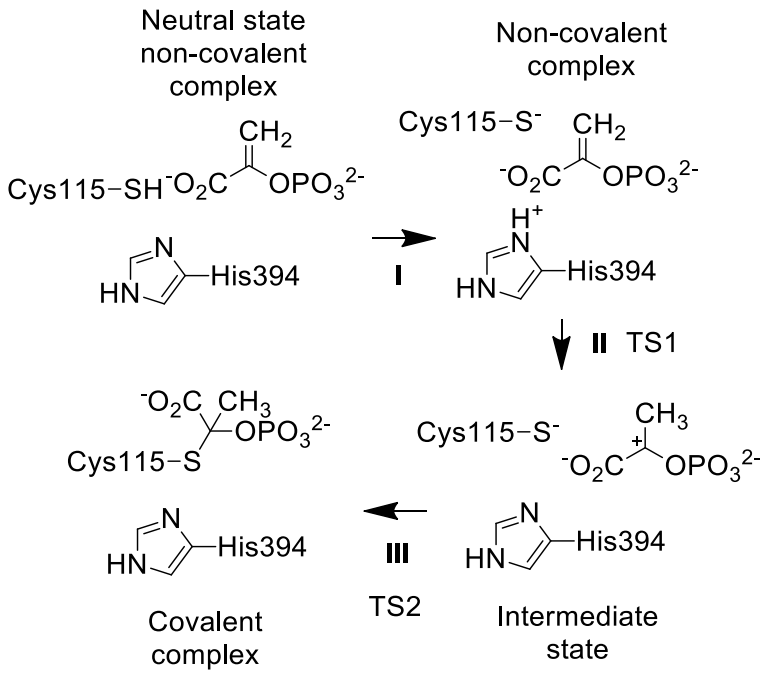


Figure 3. Proposed reaction mechanism for PEP binding to Cys115.

Assuming the same deprotonation mechanism, we modeled the reaction between Cys115 and three sets of potential cysteine targeting inhibitors containing two active and one inactive compounds against MurA in each set. We calculated the free energy barrier for every protein-ligand complex formation and compared them with the experimental activity data (**Table 1**). The active-inactive separation based on the calculated barriers were performed successfully. In all three cases the largest barrier belonged to the inactive compounds, highlighting the effectiveness of the applied method in covalent compound prioritization.

Table 1. Experimental and calculated properties of MurA targeting compounds.

Compound	Residual activity [%]	Type	GSH half life [h]	$\Delta G^{\ddagger}$ [kcal·mol^{-1}]
M1	NA[a]	Active	NA[a]	13.3
M2	NA[a]	Active	NA[a]	11.1
M3[b]	93	Inactive	77.5	19.6[c], 20.2[d]
M4	3	Active	0.0	22.3
M5	3	Active	0.0	19.9
M6	110	Inactive	32.8	27.8
M7	12	Active	0.2	10.8
M8	1	Active	0.0	11.3
M9	95	Inactive	127.0	16.5

[a]NA – not available, [b]Racemic, [c]R-enantiomer, [d]S-enantiomer

3.3. AFFINITY AND SELECTIVITY ASSESSMENT OF COVALENT INHIBITORS [T2]

The calculations performed in the MurA studies assumed that the energy contribution of the non-covalent binding is negligible compared to the covalent part. The aim of our next study was to take into account the complete covalent inhibition including both the non-covalent and covalent steps. We selected ten experimentally characterized $KRAS^{G12C}$ and five $EGFR^{T790M/L858R}$ inhibitors; and their binding free energies and reaction barriers were calculated by thermodynamic integration and QM/MM MD umbrella sampling, respectively. The calculated properties were converted into K_i and k_{inact} using **Eqs. 2** and **3**. Good correlation was found between the computed and experimental data, proving the effectiveness of the applied protocol (**Figure 4** and **5**).

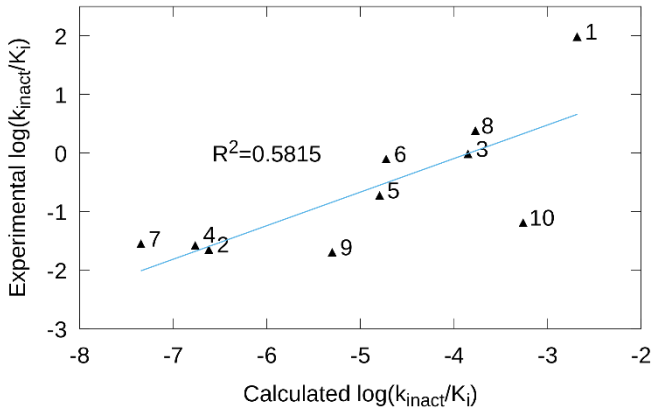


Figure 4. Comparison of the experimental and calculated $\log(k_{inact}/K_i)$ for $KRAS^{G12C}$ inhibitors

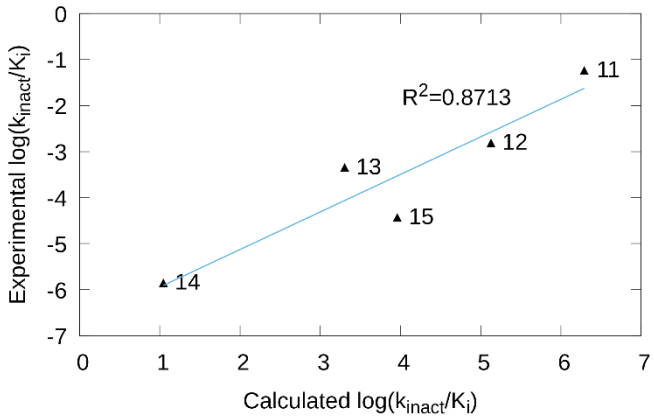


Figure 5. Comparison of the experimental and calculated $\log(k_{inact}/K_i)$ for $EGFR^{T790M/L858R}$ inhibitors

We evaluated the selectivity of a selected compound over three related kinases, ITK, BTK and BMX. We applied sidechain mutational TI simulations estimating the binding free energy differences. The covalent steps were modeled as previously described. The selectivity differences between two further inhibitor molecules towards EGFR single and double mutants were also characterized by sidechain mutational TI and QM/MM MD simulations.

Table 2. Experimental (exp) and calculated (calc) binding free energies, (ΔG) binding free energy differences (ΔΔG) and free energy barrier ($\Delta G^{\ddagger}$), for the binding of the selected compound to different kinases. (Energy values are in kcal·mol^{-1})

Enzyme	ΔG_{exp}	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{calc}$	$\Delta G^{\ddagger}_{exp}$	$\Delta G^{\ddagger}_{calc}$
ITK	-10.4	-	-	25.1	24.3
BTK	-5.8	-	-	21.1	19.6
BMX	-8.6	-2.8[a]	-2.3[a]	22.9	22.9

[a]BTK to BMX mutation

Table 3. Experimental ($\Delta\Delta G_{exp}$) and calculated ($\Delta\Delta G_{calc}$) binding free energies for the sidechain and ligand mutations of EGFR inhibition. (Energy values are in kcal·mol^{-1})

Ligand	Mutation	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{calc}$
17	T790->M790	-2.8	-0.7
18	T790->M790	-4.7	-3.3
Enzyme	**Transformation**	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{calc}$
L858R	**17->18**	1.5	0.5
L858R/T790M	**17->18**	-0.4	0.3

3.4. SELECTIVITY AND INHIBITION MECHANISM OF COVALENT IPS INHIBITORS [T3]

Finally, we tested the developed protocol for more complex covalent inhibition binding events. We selected the immunoproteasome (iPS) inhibition by oxathiazolone derivatives as the system to be examined. Firstly, we clarified the favored mechanism of the covalent reaction between the targeted Thr1 residue of iPS and the simplest oxathiazolone derivative. We also determined the rate-determining step of the covalent binding. Thereafter, we calculated the binding free energies and transition barriers of the rate-determining step of six oxathiazolones, converted them to K_i and k_{inact} values, and compared ratios with the experimental k_{inact}/K_i ratios. Sensible correlation was found between the calculated and experimental data, verifying the usefulness of the developed protocol, containing coupled TI MD and QM/MM US MD simulations (**Figure 6**).

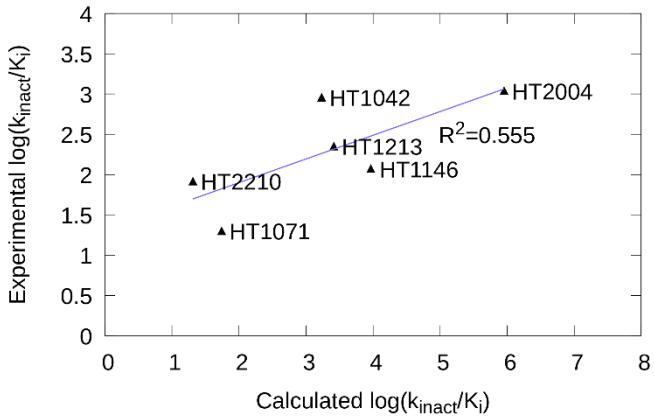


Figure 6. Comparison of the experimental and calculated $\log(k_{inact}/K_i)$ (k_{inact} in s^{-1} and K_i in M units).

Even better correlation was found between the calculated $\log(1/K_i)$ and experimental $\log(k_{inact}/K_i)$ values (**Figure 7**), suggesting that the differences in the binding affinities are mainly governed by the non-covalent step.

The examination of the selectivity differences of two oxathiazolones (**HT1146** and **HT2004**) towards the iPS and cPS was carried out, too. We performed TI simulation coupled with the "hybrid" active site approach to estimate the binding free energy differences of the selected compounds in iPS and cPS. Simulation of the covalent reactions were also carried out characterizing the rate-determining step in cPS.

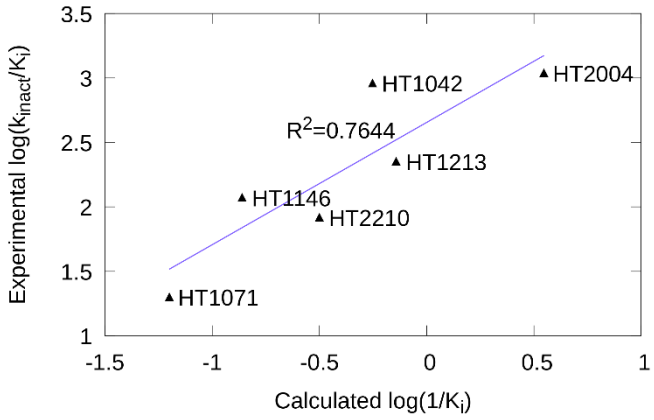


Figure 7. Comparison of the experimental $\log(k_{inact}/K_i)$ and calculated log(1/Ki) (K_i in M units).

Eq. 4 relates k_{inact}/K_i based and free energy difference based description of selectivity.

$$log\left(\frac{k_{inact}[A]}{K_i[A]}\right) - log\left(\frac{k_{inact}[B]}{K_i[B]}\right) = \frac{log(e)}{RT}\left(-\Delta G^{\ddagger}[B] + \Delta G^{\ddagger}[A] - \Delta G[B] + \Delta G[A]\right) = \frac{log(e)}{RT}\left(-\Delta\Delta G^{\ddagger} - \Delta\Delta G\right) \quad (4)$$

We evaluated selectivities both from computed and experimental data. The results collected in **Table 4** show good agreement between the computed and experimental values. The MD simulations highlighted that the main reason of the selectivity differences is a Gln53Ser mutation causing different Met48 conformation in cPS compared to iPS and this has a ligand size dependent effect on the binding and reactions of inhibitors.

Table 4. Comparison of experimental and calculated free energy differences between **HT1146** and **HT2004** binding (top two rows) and between binding to iPS and cPS (bottom two rows). (Energy differences are in kcal·mol^{-1})

Transformation	Experimental $-\Delta\Delta G^{\ddagger}-\Delta\Delta G$	Calculated $-\Delta\Delta G^{\ddagger}-\Delta\Delta G$
HT1146->**HT2004** iPS	-1.4	-2.0
HT1146->**HT2004** cPS	3.0	4.5
HT1146 iPS->cPS	0.8	0.4
HT2004 iPS-> cPS	5.1	6.7

4. THESIS POINTS

1. I have proposed a mechanism for the reaction catalyzed by MurA. This includes the loop closure upon UNAG binding, followed by a proton transfer between Cys115 and His394 that activates the cysteine residue for the nucleophilic attack towards the natural substrate PEP or the cysteine targeting inhibitors. I found that PEP is first protonated by His 394 and Cys115 forms a covalent adduct with the protonated PEP [T1].

2. I have shown that the active-inactive separation of ligands is possible based on the calculated barriers of the inhibition reaction between MurA and nine selected compounds containing oxirane, α-haloketone, α-haloacetamide, acrylamide and vinyl-sulfone warheads [T1].

3. I have developed a complex molecular dynamics based protocol characterizing the complete covalent protein inhibition including both the non-covalent and covalent steps. The protocol was validated successfully on a set of relevant protein targets, namely on KRAS, EGFR, ITK, BTK and BMX. Selectivity assessments were also carried out explaining binding affinity differences of specific compounds towards the related kinases and EGFR mutants [T2,T4].

4. I have clarified the mechanism of the covalent reaction between the Thr1 residue of immunoproteasome and an oxathiazolone derivative using QM/MM molecular dynamics umbrella sampling calculations. I have determined that the carbonate pathway is favored over the carbonthioate route. I showed that the nucleophilic attack of the activated Thr1 Oγ and the proton transfer between the terminal amino group of Thr1 and the negatively charged sulfur of the ligand are asynchronous events composing the rate determining step [T3].

5. I have shown that the differences in the iPS inhibitory activity of six examined oxathiazolone compounds are primarily affected by the molecular recognition, and less by the chemical reaction. I interpreted the immunoproteasome versus constitutive proteasome selectivity of two compounds. I found that the Gln53Ser mutation causes different Met48 conformation in cPS compared to iPS and this has a ligand size dependent effect on the binding and reactions of inhibitors [T3].

5. APPLICATIONS

Though, the primary aim of the developed protocol is to aid covalent drug design through an enhanced compound ranking process, the applicability is not limited to activity based compound prioritization. We successfully used the developed scheme in selectivity studies evaluating the binding free energy difference between the same ligand binding to related proteins, performing sidechain mutational TI simulations. The QM/MM free energy simulations are well suited to explore reaction mechanisms in complex systems, and the identified reaction mechanisms of covalent inhibitors are used in the characterization of the covalent inhibition process, in compound ranking and in structure optimization.

6. PUBLICATIONS

6.1. JOURNAL ARTICLES IN THE TOPIC OF THE PRESENT THESIS

T1. Catalytic Mechanism and Covalent Inhibition of UDP-N-Acetylglucosamine Enolpyruvyl Transferase (MurA): Implications to the Design of Novel Antibacterials
L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű,
Journal of chemical Information and Modeling, **2019**, *59*, 5161–5173.
DOI: 10.1021/acs.jcim.9b00691. IF(2019): 4.549 FI: 4

T2. Affinity and Selectivity Assessment of Covalent Inhibitors by Free Energy Calculations
L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű
Journal of chemical Information and Modeling, **2020**, *60*, 6579–6594.
DOI: 10.1021/acs.jcim.0c00834. IF(2020): 4.956 FI: 3

T3. Mechanistic and thermodynamic characterization of oxathiazolones as potent and selective covalent immunoproteasome inhibitors
L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű
Computational and Structural Biotechnology Journal, **2021**, *19*, 4486-4496.
DOI: 10.1016/j.csbj.2021.08.008. IF: 7.271 FI: 0

T4. The role of quantum chemistry in covalent inhibitor design
L. M. Mihalovits, G. G. Ferenczy, G. M. Keserű
International Journal of Quantum Chemistry, **2021** (early view).
DOI: 10.1002/qua.26768. IF(2020): 2.444 FI: 0

6.2. LECTURES IN THE TOPIC OF THE PRESENT THESIS

1. Kovalens UDP-N-acetilglükozamin enolpiruvil transzferáz (MurA) inhibitorok vizsgálata vegyes QM/MM molekuladinamikai szimulációk segítségével
   Levente Márk Mihalovits, György G. Ferenczy, György Miklós Keserű, *KeMoMo-QSAR 2019 szimpózium,* **Szeged,** 2019. június 6.
2. Az UDP-N-Acetilglükózamin Enolpiruvil Transzferáz (MurA) enzim katalitikus mechanizmusa és kovalens gátlása
   Levente Márk Mihalovits, György G. Ferenczy, György Miklós Keserű, *TTK SzKI szakmai előadói napok,* **Telki**, 2020. január 22.
3. Kovalens inhibitorok vizsgálata szabadentalpia számításokkal
   Levente Márk Mihalovits, György G. Ferenczy, György Miklós Keserű, *KeMoMo-QSAR 2021 szimpózium,* **Szeged**, 2021. szeptember 30.