Systematic exploration of multiple drug binding sites

Mónika Enikő Bálint

Thesis of PhD Dissertation

Supervisor: Dr. Habil. Csaba Hetényi

Associate Professor

Department of Pharmacology and Pharmacotherpy,

Medical School, University of Pécs, Hungary

Structural biochemistry program

Doctoral school in Biology

Program Leader: Prof. Dr. Mihály Kovács

Head of school: Prof. Dr. Anna Erdei

Eötvös Loránd University, Faculty of Science

Institute of Biology, Department of Biochemistry

Budapest, 2018

Introduction

Computer-based drug design is a decade-old, now flourishing research area. Drug design starting from the molecular structure, requires precise calculations of target-ligand (drug) interactions, which can be achieved with the help of computational technologies. Although there are a number of methods available to us, these need improvement on several areas. Additionally, new strategies such as polypharmacology, allosteric or drug design for targets with multiple sites are a major challenge. The conformational change of the target and its role is gaining more attention during the design, therefore allosterically-induced conformational changes are particularly common in protein signalling pathways.

The blind docking (BD) approach was developed within our research group, and maps the potential binding sites on the surface of the entire protein, blindly, without knowing their position. Approximately 500 references were received for the initial blind docking procedure, which have been used in several studies for the search of allosteric binding sites or multiple drug binding sites. In addition to the benefits of the blind docking and its numerous applications, the limitations of our approach have also been established over the last decade. These limitations are caused, among other things, by simple water models, inadequate handling of the target flexibility, and random search algorithms, which altogether hinders the possibility of a systematic binding site search. Several studies have shown that the limitations of docking methods can be partially solved by molecular dynamics. MD methods are able to handle the flexibility of the target and account on the presence of water molecules as well. MD simulations are also able to track the entire drug binding pathway and the process of induced-fit as well. Although the use of molecular dynamics has been a major step forward, its application is still quite

cumbersome due to its long calculation time and cannot be considered a systematic approach.

Objectives

The main objectives of the presented thesis were the analysis of the limitations of the currently available and widely used docking and MD techniques. Moreover, learning from the observed limitations, my aim was the development of a fully systematic method which can find the binding sites and modes of ligands on protein surfaces, blindly, in a completely unbiased way.

1. Case study 1: Docking and Molecular dynamics on human serum albumin (HSA)

In order to achieve the final goal, challenging systems were initially studied, that are well known to present multiple binding sites. Accordingly, in our initial case study, human serum albumin (HSA) was investigated with blind docking and MD, to study binding site of zearalenone and compare our observation with the experimentally observed results.

2. Case study 2: Docking and Molecular dynamics on human estrogen receptor alpha (hERα)

In a second case study human estrogen receptor alpha (hERα) was analysed, with both docking and MD calculations.

a. The influence of the coactivator peptide (CA) in the stabilization and/or unbinding mechanism of E2 and EN, on the hERα was studied.

b. Structural description of the two known-, the classical and alternative binding site

(CBS and ABS) on hERα was one of our objective.

c. Beside the structural description of the hERα binding sites, following the docking calculations, MD simulations were performed to reveal the dynamic unbinding mechanism, and the kinetic stability of estradiol (E2) and estren (EN).

d. Unveiling and analysing the dynamic networking interaction, that lies behind the unbinding mechanism of ligands, from ABS binding site, was also one of our fundamental objective, in the study.

3. Method development: Wrap 'n' Shake

The third study describes the development of the fully systematic binding search method, exploiting the advantages of blind docking and MD simulations techniques. This main goal was divided into multiple conditions and objectives as follows in point **a-d**.

a. The initial stage was to cover (Wrapper stage) the target surface with docked ligand copies, with the purpose to solve the limitation of the current methods, regarding their inability to perform a truly systematic binding site search.

b. To develop a truly systematic method, my aim was to obtain a complete target surface coverage, with ligand monolayer. In order to obtain this thorough protein surface coverage with ligands, a few conditions and objectives were pre-set, such as:

- To minimize the ligand–ligand interactions through implementation of a weak repulsion between the docked ligand copies and therefore blocking the formation of ligand aggregates.

- To maximize the target-ligand interactions in order to ensure that the

largest possible numbers of new ligand copies are placed on the surface in an actual BD cycle.

c. Using MD simulations (Shaker stage), filter out the unnecessary, weak binding ligands obtained from Wrapper, and retain only the ligand conformations; from the functional bindings sites.

d. Using the MD trajectories, validate the structural conformations of the calculated binding positions, to experimentally known binding sites, and binding conformations of specific ligands.

Results

Computational methods such as docking and molecular dynamics (MD) are becoming more affordable, than ever before, allowing a fast-paced solvation of drug discovery problems. With the computational improvements, the combination of GPUs and software codes that can exploit the newest innovations in hardware architectures, MD has become a fundamental tool for calculation of the structure of proteins and protein-ligand complexes.

The presented thesis is based on three papers. Two case studies are presented on challenging systems that have exploited both aforementioned computational methods (docking and MD). The third study, presents a new systematic binding site search method (Wrap'n'Shake), that works in synergy with popular open source program packages AutoDock 4.2.3 and GROMACS 5.0.2. Wrapper is freely available in a new open source package (WnS) as shell scripts and a C program Wrp available for download together with a User's Manual at www.wnsdock.xyz.

In our initial study, a challenging protein was investigated (HSA), with multiple

binding sites known on its surface. Beside fluorescence spectroscopy and ultrafiltration experiments, docking calculations, followed by MD were performed on the HSA. The binding sites resulted after docking calculations, were characterized and further improved with MD simulations. MD calculations allowed HSA-ZEN interaction energy re-evaluation, which was necessary to characterize the stability of the analysed binding sites. Docking calculations were insufficient to draw such conclusions regarding the stability of the functional binding sites.

MD was also required to investigate which protonation state of H464 is the most probable during ZEN binding, and the contact of which amino acids are mandatory for ZEN interaction in the A binding pocket. Combining the experimental methods such as fluorescence spectroscopy and ultrafiltration with theoretical calculation such as docking and molecular dynamics, we demonstrated that ZEN forms stable complex with albumin, occupying a non-conventional binding site on HSA (between Sudlow's site I and II).

In the second case study, human estrogen receptor alpha (hERα) was analysed, with both docking and MD calculations. As our initial objective, influence of the coactivator peptide (CA) in the stabilization and/or unbinding mechanism of E2 and EN were analysed. Our results supports the idea, that ABS is available for ligand binding, if AF2 is not occupied by CA, otherwise it is dynamically blocked by side-chains of amino acids composing the ABS. Our study also provided structural evidence, that non-classical effects of sex steroids require extranuclear, membrane-bound localization of the estrogen receptor, where the AF2 binding site is not occupied by CA.

Both classical (CBS) and alternative binding sites (ABS) were exhaustively mapped on the ligand-binding domain of human estrogen receptor alpha. Blind docking and molecular dynamics simulations allow an unbiased exploration of the whole surface

of hERα which is essential for finding multiple binding sites.

Using the ligand conformations obtained after docking calculations target – ligand complexes were submitted to MD. MD trajectories were analysed, and structural dynamics of non-classical effects of sex steroids was presented. Kinetic stability of the steroid –receptor complexes was investigated using molecular dynamics calculations. These investigations were possible by allowing fully flexible MD simulations on both the ligands, and hERα. Protein flexibility is particularly important in following the complete ligand binding pathway due to induced fit conformational changes. Real-time investigations of the complete interaction network at atomic resolution pointed to key residues (M357, H356 and L326) of steroid binding mechanism. Apart from a fully flexible hERα, the presence of explicit water molecules proved to be highly important especially in the estren binding/unbinding process.

We showed how steroid binding to the alternative binding site of non-classical action is facilitated by the presence of a ligand in the classical binding site and the absence of the co-activator peptide. We present an elucidation of the binding mechanism of estradiol and estren to hERα that can be further exploited in the structure-based design of new drugs with rapid, non-classical responses.

In the third study, a systematic strategy, the Wrap 'n' Shake was introduced for exploration of multiple binding sites and modes of drugs on their macromolecular targets. Ten test applications are presented with successful identification of multiple binding sites on biologically important systems such as mitogen activated- and tyrosine protein kinases, key players of cellular signalling as well as farnesyl pyrophosphate synthase, a target of antitumor agents.

Wrap 'n' Shake initially performs a systematic wrapping of the target into a

monolayer of ligand copies using a modified blind docking approach and finally selects the stable positions by shaking off loose binders. The method offers a computationally feasible solution for the present problems of the field.

In the first stage, Wrapper requires only fast blind docking cycles, to systematically and fully cover the whole protein surface. Wrapper is a fully automatized process, which combines the AutoDock 4.2.3 program package, with the Wrp program, in several consecutive cycles. A new atom type (X) was introduced in the force field used in the AutoDock program package, in order to minimize the ligand–ligand interactions. The purpose of the new atom type was to implement a weak repulsion between the docked ligand copies, and therefore block the formation of ligand aggregates. By inhibiting the formation of ligand aggregates, the target-ligand interactions are maximized to ensure that the largest possible numbers of new ligand copies are placed on the surface in an actual BD cycle.

The second stage (Shaker) is also fairly short and can be performed by available MD packages. Its purpose is to find the functional binding sites on the protein surface. This purpose is achieved with several Shaker cycles, which eliminates ("washes off") the weak binding ligands. The elimination of ligands is further accelerated by the presence of explicit water molecules and target flexibility. Shaker cycles were carried out until 75% of ligands were eliminated. After each shaker stage, target-ligand distance and energy calculation was carried out, in order to be able to follow the path of a ligand during the MD trajectory. The ligand elimination process was based on these calculated parameters. In the first cycles of the Shaker protocol, backbone restraint was implemented, to grossly filter out the ligand excess. In the following stages, simulated annealing was introduced in the MD protocols, in order to accelerate the dissociation process of the ligands. In the final Shaker stage, total protein flexibility was allowed, to refine the conformation of the cluster representatives. No position restraint was set on the target, in the final step to allow

any induced-fit movements to take effect. Since this step is considered a refinement step, the duration of the simulations at this stage can be entirely customizable, depending on the user's need, and computational availability. Shaker stages, has successfully filtered out the unnecessary, weak binding ligands obtained from Wrapper, and retain only the ligand conformations, from the functional binding sites in all ten test cases. The final calculated ligand conformations were validated to known, experimentally determined binding conformations, using RMSD as a metric number.

Conclusion

Several high quality articles proved that combining docking and MD methods, is a very efficient computational approach, in order to find the answer for diverse scientific questions. In each of my studies concerning the present thesis, blind docking was performed, followed by a subsequent molecular dynamics step.

The presented case study 1 used MD in order to refine and re-rank the zearalenone conformation obtained with BD. In the second case study, MD trajectories were evaluated in a more advanced way, by calculating the residence frequency of estren and estradiol, but also by exploring and shedding light on ligand binding pathways, and target-ligand interaction networks.

In the third study, a new method was developed (Wrap 'n' Shake) using multiple docking and MD protocols, and exploring the advantages for each of them. In its current stage Wrap 'n' Shake is suitable to study interactions of protein targets even with large peptide ligands, however further developments are on their way. We have started the extension of the method towards protein ligands using a fragment-based approach with post hoc reconstruction of the ligand. In future applications, Wrap 'n' Shake could be also used for general pocket search, besides docking of

individual ligands. We envision that Wrap 'n' Shake can become the tool of choice for systematic exploration of multiple binding sites and modes of ligands in drug design and structural biology.

Wrap'n'Shake was implemented in the software package WnS, released under the GNU General Public License, freely accessible with full documentation at www.wnsdock.xyz.

Publications

Research papers concerning the present thesis

1. **Bálint M**, Jeszenoi N, Horvath I, van der Spoel D, Hetenyi C (2017) **Systematic exploration of multiple drug binding sites.** Journal of Cheminformatics, 9(1): 65

2. Bálint M, Jeszenői N, Horváth I, Ábrahám IM, Hetényi C. (2017) Dynamic changes in binding interaction networks of sex steroids establish their non-classical effects. Scientific Reports, 7(1): 14847.

3. Poór M, Kunsági-Máté S, **Bálint M**, Hetényi C, Gerner Z, Lemli B. (2017) **Interaction of mycotoxin zearalenone with human serum albumin.** Journal of Photochemistry and Photobiology b-biology 170:16-24.

Conference lectures concerning the present thesis

1. **Bálint M**, Hetényi C. (2017). **Hogyan csomagoljunk be egy célpontot?** KeMoMo–QSAR Szimpózium: 2017.06.1-2, Szeged, Hungary (Presentation).

2. **Bálint M**, Hetényi C. (2017). **Hogyan csomagoljunk be egy célpontot?** Magyar

Tudományos Akadémia Bioinformatikai Osztályközi Állandó Bizottsága és a Magyar Bioinformatikai Társasága. 2017. Nov. 10. Budapest, Hungary. (Presentation).

3. Hetényi C, **Bálint M**, Jeszenői N, Horváth I. (2017). **Improvements of computational optimization of drug-target interactions.** Global Conference on Pharmaceutics and Drug Delivery Systems, 2017.06. Valencia, Spain. (Presentation).

4. **Bálint M.** (2016) Dinamikus "Blind docking" módszer, kötőhelyek keresésére szisztematikus módon. PhD conference. 2016-03-02 - 2016-03-04 (2016). Debrecen, Hungary, (Presentation).

5. **Bálint M** (2016). **Dinamikus "blind docking" módszer, kötőhelyek keresésére.** Tavaszi szél Nemzetközi multidiszciplináris konferencia, Doktoranduszok Országos Szövetsége. Keresztes Gábor (editor). Abstractbook. 485 p. ISBN: 978 615 5586 04 0. 2016.04.15-2016.04.17. Budapest, Hungary. (Presentation).

6. **Bálint M.** (2014) **Allosztérikus kötőhelyek keresése szerkezeti vizek bevonásával és blind docking módszerrel.** PhD konferencia, Balassi Intézet Márton Áron Szakkollégium. 2014-03-01 - 2014-03-03. Szeged, Hungary. (Presentation)

Other research papers

1. Jeszenői, N., Schilli, G., Bálint, M., Horváth, I. and Hetényi, C. (2018). **Analysis of the influence of simulation parameters on biomolecule-linked water networks**. Journal of Molecular Graphics and Modelling. 82: 117-128

2. Poór M, Boda G, Mohos V, Kuzma M, **Bálint M**, Hetényi C, Bencsike T. (2018). **Pharmacokinetic interaction of diosmetin and silibinin with other drugs: Inhibition of CYP2C9-mediated biotransformation and displacement from serum albumin.** Biomedicine & Pharmacotherapy, 102: 912–921

3. Poór M, Bálint M, Hetényi C, Gődér B, Kunsági-Máté S, Kőszegi T, Lemli B. (2017). Investigation of non-covalent Interactions of aflatoxins (B1, B2, G1, G2, and M1) with serum albumin. Toxins 9(11): 339

4. Herman BE, Szabó J, Bacsa I, Wölfling J, Schneider G, Bálint M, Hetényi C, Mernyák E, Szécsi M. (2016). Comparative investigation of the in vitro inhibitory potencies of 13-epimeric estrones and D-secoestrones towards 17β-hydroxysteroid dehydrogenase type 1. Journal of Enzyme Inhibition and Medicinal Chemistry 31 (Suppl. 3): 61-69.

5. Jeszenői N, **Bálint M**, Horváth I, van der Spoel D, Hetényi C. (2016) **Exploration of Interfacial Hydration Networks of Target-Ligand Complexes.** Journal of Chemical Information and Modeling 56(1): 148-158.

6. Jeszenői N, Horváth I, **Bálint M**, van der Spoel D, Hetényi C. (2015) **Mobility-based prediction of hydration structures of protein surfaces.** Bioinformatics 31(12): 1959-1965.

7. Poór M, Lemli B, **Bálint M**, Hetényi C, Sali N, Kőszegi T, Kunsági-Máté S. (2015). **Interaction of citrinin with human serum albumin.** Toxins 7(12): 5155-5166.

8. Avram S, Crisan L, Pacureanu L, Bora A, Seclaman E, **Bálint M**, Kurunczi L. (2013). **Challenges in docking 2'-hydroxy and 2',4'-dihydroxychalcones into the binding site of ALR2.** Medicinal Chemistry Research 22(8): 3589-3605.

Other conference lectures and posters

1. **Bálint M**, Hetényi C. (2017) **Ligandumok kötődésének számítógépes vizsgálata a szomatosztatin receptor 4 (SST4) célponton.** A Magyar Élettani Társaság LXXXI. Vándorgyűlése. 2017.06.13-16. Debrecen, Hungary. (Presentation).

2. **Bálint M**, Hetényi C, Ábrahám I. (2017) **Exploration of ligand binding mechanisms to human estrogen receptor alpha.** Federation of European Neuroscience Society, Poster presentation, 2017.Sep.20-23, Pécs, Hungary. Poster presentation.

3. Hetényi C, Bálint M. (2017). Nem-természetes tirozin formák vélt szabályozó hatásainak szerkezeti magyarázata az inzulin receptor szubsztrát- 1 (IRS-1) esetében Magyar Diabetes Társaság XXV. Jubileumi Kongresszusa, 2017.04.20 – 23. Pécs, Hungary. Lecture presentation.

4. Mészáros N, Horváth I, **Bálint M**, Hetényi C. (2017) Unraveling the histone code by fragment blind docking ECBS 2017 5th European Chemical Biology Symposium 2017.07.02-04. Budapest, Hungary. Poster presentation.

5. Hetényi C, Jeszenői N, **Bálint M**, Horváth I (2016). **Answers to current challenges in target-based drug design.** BJMT Conference of Applied Mathematics. Laszlo Gelencser, Zoltan Horvath (editor.), Book of Abstracts. 28, 2016.06.01-2016.06.03. Győr, Hungary. Lecture presentation.

6. Jeszenői N, **Bálint M**, Horváth I, Hetényi C. (2016) **Komplex szerkezetek vízhálózatai,** KeMoMo–QSAR Szimpózium, 2017.05.12-13, Miskolc, Hungary. Lecture presentation.

7. Jeszenői N, Ábrahám I, Bálint M, Horváth I, Hetényi C. (2016) Az ösztrogén receptor alfa nem-klasszikus ösztrogén hatásokért felelős kötőhelyének azonosítása: in silico szerkezeti biokémiai vizsgálatok. III. Magyar Neuroendokrin Szimpózium, Magyar Endokrinológiai és Anyagcsere Társaság XXVI. Kongresszusának szatellit rendezvénye, 2016. 05.04, Szeged, Hungary. Lecture presentation.

8. Hetényi C, Jeszenői N, **Bálint M**, Horváth I. (2015). **Fehérjék hidrátszerkezete: modellezzem vagy mérjem?** KeMoMo–QSAR Szimpózium, Szeged, Hungary. Lecture

presentation.

9. **Bálint M**, Avram S, Kurunczi L. (2013). **Docking studies of 2' hydroxychalcones as inhibitors of aldose reductase.** Young People And Multidisciplinary Research, Proceedings of the XVth Symposium, Editura Politehnica 2013-11-14 - 2013-11-15, 7 p. Timisoara Romania. Lecture presentation.