The use and development of computational docking: from cytochrome inhibitors to viral ion channel blockers

PhD thesis

Dr. Balázs Zoltán Zsidó

Doctoral School of Pharmacology and Pharmaceutical Sciences

Neuropharmacology Program

Leader of doctoral school: **Prof. Dr. Erika Pintér**

Leader of doctoral Program: **Prof. Dr. Erika Pintér**

Supervisor: **Dr. Csaba Hetényi**

University of Pécs

Faculty of Medicine

Department of Pharmacology

Pharmacoinformatics Unit

Pécs, 2022.

1. The thesis is based on the following articles:

1. **Balázs Zoltán Zsidó**, Mária Balog, Nikolett Erős, Miklós Poór, Violetta Mohos, Eszter Fliszár-Nyúl, Csaba Hetényi, Masaki Nagane, Kálmán Hideg, Tamás Kálai, Balázs Bognár, Synthesis of spin-labelled bergamottin: a potent CYP3A4 inhibitor with antiproliferative activity. *Int. J. Mol. Sci.* 21 (**2020**) 508. [**IF: 5.923; Q1**]
2. **Balázs Zoltán Zsidó** , Rita Börzsei, Viktor Szél, Csaba Hetényi. Determination of Ligand Binding Modes in Hydrated Viral Ion Channels to Foster Drug Design and Repositioning. *J. Chem. Inf. Model. 61* (**2021**) 4011-4022 [**IF: 4.956; Q1**]

The sum impact factor of the articles as a basis of the thesis: 10.879

Cumulative impact factor: 43.749

Independent citations (mtmt.hu): 17

2. Other articles:

3. **Zsidó, B.Z.;** Hetényi, C. The role of water in ligand binding. *Curr. Opin. Struct. Biol.* **2021**, 67, 1–8. **IF: 6.809, Q1**
4. **Zsidó, B.Z.;** Hetényi, C. Molecular structure, binding affinity, and biological activity in the epigenome. *Int. J. Mol. Sci.* **2020**, 21, 1–40. **IF: 5.923, Q1**
5. **Zsidó B.Z.;** Börzsei R.; Pintér E.; Hetényi C. Prerequisite Binding Modes Determine the Dynamics of Action of Covalent Agonists of Ion Channel TRPA1. *Pharmaceuticals*, **2021**, *14*, 988-1000. **IF: 5.863, Q1**
6. Mohos, V.; Fliszár-Nyúl, E.; Lemli, B.; **Zsidó, B.Z.**; Hetényi, C.; Mladěnka, P.; Horký, P.; Pour, M.; Poór, M. Testing the pharmacokinetic interactions of 24 colonic flavonoid metabolites with human serum albumin and cytochrome P450 enzymes. *Biomolecules* **2020**, 10. **IF: 4.879, Q1**
7. Mohos, V.; Fliszár-Nyúl, E.; Ungvári, O.; Bakos, É.; Kuffa, K.; Bencsik, T.; **Zsidó, B.Z**.; Hetényi, C.; Telbisz, Á.; Özvegy-Laczka, C.; et al. Effects of chrysin and its major conjugated metabolites chrysin-7-sulfate and chrysin-7-glucuronide on cytochrome P450 enzymes and on OATP, P-gp, BCRP, and MRP2 transporters. *Drug Metab. Dispos*. **2020**, 48, 1064–1073. **IF: 3.231, D1**
8. Fliszár-Nyúl, E.; Faisal, Z.; Mohos, V.; Derdák, D.; Lemli, B.; Kalai, T.; Sár, C.; **Zsidó, B.Z**.; Hetényi, C. ; Horváth, Á.; Helyes, Zs.; Deme, R.; Bogdán, D.; Czompa,

A.; Matyus, P., Poór, M. Interaction of SZV 1287, a novel oxime analgesic drug candidate, and its metabolites with serum albumin. *Journal of Molecular Liquids* **2021**, *333*: 115945. **IF: 6.165, Q1**

9. **Zsidó B.Z.;** Hetényi C. A D-aminosavak élettani szerepéről: Előfordulásuk gyógyszerhatóanyagokban és étrendkiegészítőkben. *Gyógyszerészet*, **2021**, *5*: 293.

3. Other presentations:

1. Zsidó, Balázs Zoltán ; Börzsei, Rita ; Hetényi, Csaba
   Calculation of complex structures of protein targets using a fragment blind docking approach (2019) – **oral presentation** From Protein Complexes to Cell-Cell Communication, Esztergom, Magyarország 2019.10.27. - 2019.10.29.
2. Zsidó, Balázs Zoltán ; Hetényi, Csaba
   A TRPA1 receptor ligandumokkal alkotott komplexeinek számítógépes vizsgálata – **oral presentation** Bioinformatika 2020 - Magyar Bioinformatikai Társaság Konferenciája
3. Zsidó, Balázs Zoltán ; Hetényi, Csaba
   Investigation of the receptor-ligand complexes of TRPA1 receptor by computational approaches (2020) – **oral presentation** Online Medical Conference for PhD Students and Experts of Clinical Sciences (MedPECS) 2020
4. Zsidó, Balázs Zoltán ; Zsidó, András Norbert ; Botz, Lajos
   „Real world" egészségügyi adatok felhasználása a gyógyszerterápia biztonságosságának vizsgálatára (2019) – **oral presentation** In: Kórházi Gyógyszerészek 2019. évi Szimpóziumán elhangzott előadások
5. Zsidó, Balázs Zoltán ; Zsidó, András Norbert ; Botz, Lajos
   A valós-életbeli egészségügyi adatok hozzájárulása a nem kívánt gyógyszerhatások értelmezéséhez In: Magyar Kísérleti és Klinikai Farmakológiai Társaság III. Gyógyszer Innovációs Kongresszusán bemutatott poszterek kivonata (2019) p. 30758670 Paper: 30758670
6. Zsidó, Balázs Zoltán; Vida, Róbert György ; Botz, Lajos
   Gyógyszerezéssel összefüggő problémák elemzése adattudományi metodikával In: Bódog, Ferenc; Csiszár, Beáta; Pónusz, Róbert (szerk.) Medical Conference for PhD Students and Experts of Clinical Sciences : Book of Abstracts Pécs, Magyarország : Pécsi Tudományegyetem Doktorandusz Önkormányzat (2018) Paper: 30333870
7. Zsidó, Balázs Zoltán ; Botz, Lajos

Új módszer a nemkívánatos gyógyszerterápiás események felderítésére egy klinikai rendszerben In: Kórházi Gyógyszerészek 2018. évi Szimpóziumán elhangzott előadások és bemutatott poszterek absztraktgyűjteménye (2018) p. 30373037 Paper: 30373037

8. Bayarsaikhan, Bayartsetseg ; Hetényi, Csaba ; Zsidó, Balázs Zoltán
   Evaluation of current peptide docking tools on systems of pharmaceutical relevance. In: Molnár, Dániel; Molnár, Dóra (szerk.) XXIV. Tavaszi Szél Konferencia 2021: Absztraktkötet Budapest, Magyarország : Doktoranduszok Országos Szövetsége (DOSZ) (2021) 667 p. pp. 516-516. **oral presentation**

1. Introduction, aims

Computational docking is a method, that involves the simultaneous conformational sampling of a ligand on the target molecule and calculation of the binding affinity to produce the receptor-ligand complex. Method development studies thrive to overcome the burdens of computational docking, which are the following: the errors of the original (experimental) structures, the flexibility and size of the receptor and the ligand, the correct handling of water molecules and appropriate use of experimental data, such as partial charge assignation. In the first article related to the thesis the ligand binding of metabolistic enzymes was investigated. The CYP enzyme family has a central iron ion in the binding pocket as a part of the heme, its correct partial charge assignation was of particular interest. In the experimental structure, the imidazole N of ketoconazole was 2.7 Å away from the Fe^{3+} ion of the heme. To accurately reproduce this binding, a parameter set produced by gas phase quantum chemical calculation by an other researcher group was used. This parameter set also allowed for the different partial charges observed in the catabolic cycle. The second study was focused on the involvement of interface water molecules to molecular docking. In this study, both the correction of experimental water positions and partial charge assignation was inevitable to reproduce the experimental ligand binding position. The energetical and structural role of water molecules were investigated using the viral ion channels of the influenza A virus (M2A) and the SARS-CoV-2 virus (EC2). Also, the role of water molecules was observed on drug reposition between the two targets. Interestingly, water molecules were proven to be at least as important in ligand binding, as the target proteins themselves. In both studies the main objective was to test and develop our methods on biologically important systems, that play a key role in the pathogenesis of diseases.

During my PhD thesis, I aimed to use and develop computational docking tools to produce target-ligand complexes.

The use of computational docking in its state-of-the-art form was tested on the CYP3A4 metabolic enzyme with a known strong inhibitor, ketoconazole, also to explain its structure activity relationships. Similarly, the *de novo* atomic resolution complex of a novel ligand was produced with the same method, to give a detailed description of its structure activity relationship.

The other great challenge of fast computational docking is the correct determination of the position of water molecules, that play an important role in ligand binding. There are important drug targets, that are sensitive to these water molecules, without which fast computational dry

docking approaches cannot produce accurate results, and might be even misleading. To solve the problem, we aimed to create the hydrated, target-ligand atomic resolution complex from scratch, only needing the experimental dry protein structure, and the Lewis structure of the ligand.

2. Methods

In the articles that give the basis of the thesis tools of computational chemistry were used. These involve computational docking, molecular dynamics simulations and quantum chemical calculations. In every case, the missing atoms of the experimental protein structures were corrected, and hydrogen atoms were added, and then the structures were energy minimized. To assign partial charges, the Gasteiger-Marsili partial charge system was used for the amino acids, and the partial charge system of the heme was used from an other study, where it was calculated by gas phase quantum chemical methods. The ligands were newly built, Gasteiger-Marsili partial charges were added and all active torsions were enabled on them. In the case of the CYP3A4 target, a dry docking approach was used, and the results were compared to experimental ligand binding modes. Then the validated method was performed to produce the binding modes of the novel ligands.

In the case of the viral ion channels, to incorporate explicit water molecules into computational docking a five step method was developed, which was named HydroDock. In the first step, a blind docking search was performed without water molecules, the docking box covered the whole surface of the target proteins. In the second step, the hydrated structure of the target was produced, based on the original experimental structure of the protein. In the third step, the results of the previous two steps were merged, the hydrated target and the ligand binding mode of the dry docking calculation were placed into the same coordinate systems, and overlapping water molecules were deleted, and then the hydrated complex was energy minimized. In the fourth step, the stability of the complexes, the movement of the ligands and the formation of the hydration structure were investigated by molecular dynamics simulations. In the fifth step, a representative binding mode was selected from the approximately 1000 snapshots from the molecular dynamics simulation. This was the binding mode that showed the least deviation from the average binding mode. The method was first validated on the ion channel of influenza A virus, because this target had an available experimental complex structure, to use as a reference. Then, the validated method was transferred to the E protein (ion channel) of the SARS-CoV-2 virus, that caused the global pandemic of COVID-19 disease.

3. Results

3.1. Small molecules binding to CYP enzyme proteins

In the ligand-bound CYP3A4 complex structure (2V0M), the imidazole N atom of ketoconazole is 2.7 Å away from the iron ion of the heme. This interaction was successfully reproduced at 2.3 Å RMSD with the approach described in the Methods section. This value is within the accepted threshold (≤2.5 Å) for excellent docking results as described in the literature. Also, the main interactions were reproduced: The phenol rings of F304 amino acid and ketoconazole were parallel, participating in π-π interactions, the A370 amino acid took part in hydrophobic interactions with the methyl group of ketoconazole. The positively charged side chain of R372 participated in an ionic interaction with the partially negatively charged oxo moiety of ketoconazole. After the successful reproduction of the binding mode of ketoconazole, the same protocol was transferred to bergamottin and SL-bergamottin. In both cases the first ranked binding mode of the docking showed similarity to the experimental binding mode of ketoconazole. The furo[3,2-g]cromene-7-on ring of both bergamottin and SL-bergamottin was placed similarly as the imidazole ring of ketoconazole within the binding pocket of CYP3A4 above the heme ring, coordinated to the iron ion. The oxygen of the furan ring of both new ligands interacted with the iron ion of the heme. In the case of bergamottin the distance of the two was 4.0 Å, and in the case of SL-bergamottin 4.2 Å. In this binding position the furo[3,2-g]cromene-7-on ring and the elongated side chain was similarly placed for both ligands. The furo[3,2-g]cromene-7-on rings were parallel with the heme, participating in π interactions with the heme, similarly as in the case of the imidazole ring of ketoconazole. 7 amino acids interacted with the described binding mode of bergamottin, and 8 with that of SL-bergamottin, of which 5 were the same. The hydrophilic T309 amino acid formed a hydrogen bond with the furan oxygen. The methyl groups of the ligands interacted with the side chains of M114, F241, I301 and F304. F304 played an important role in the binding of bergamottin, SL-bergamottin and ketoconazole also. In good agreement with the experimental results, the results of our calculations provided an atomic level explanation on why SL-bergamottin is a stronger inhibitor of the CYP3A4 enzyme, than the mother compound, bergamottin.

3.2. HydroDock

To explore the importance of the position and partial charges of the structural water molecules, a systematic investigation was performed, first on the transmembrane ion channel of the influenza A virus (M2A) as a target. The dry docking of mainly the binding of small amphipathic drugs were performed. In this step no water molecules were included, and the search space of the docking program was extended to the whole surface of the target protein (blind docking). The results did not meet the literature consensus RMSD value mentioned in the previous section. Then, the experimental water oxygen positions were included into docking calculations, with added hydrogen atoms, but the orientation of the hydrogens did not favour the binding of the ligands, which resulted in still inacceptable results. A special energy minimization protocol was performed to establish the correct hydrogen bonding network, that is suitable for ligand binding, yet with Gasteiger-Marsili partial charges, the results of the docking still lagged behind the expectations. The final solution was the assignation of TIP3P partial charges, after which excellent matches were observed with the experimental ligand binding modes. Since for the new target there were no experimentally available water oxygen positions, the hydration structure had to be predicted first. To validate this new approach first we started with the "old" M2A target. Of the experimentally captured 10 water molecules, 9 were found by MobyWat. These were energy minimized, and the same docking protocol was performed as described before. Similarly, excellent matches were observed with experimental ligand binding modes. Finally, to reach a realistic situation in drug design projects, the 5 step-protocol detailed in the Methods section was applied to produce the fully hydrated target-drug complexes only from scratch, which is the Lewis structure of the drugs, and the experimental protein structure without waters. Excellent agreement was reached with experimental results by using the HydroDock protocol. After validation on the "old" M2A target, the protocol was transferred to the new EC2 target, which is the viral ion channel of the SARS-CoV-2 virus that caused the COVID-19 pandemic. The experimental solid state NMR structure of the EC2 viral ion channel was available, from which we produced the fully hydrated target-ligand complexes with the amphipathic drugs. In the study, where the experimental structure of EC2 was determined, similarly, an amantadine derivate was used. We also achieved good agreement with identifying the experimentally determined interacting amino acids using the HydroDock protocol.

4. Summary

During my PhD work, for computational docking calculations targets of current interest and of great pathophysiological importance were selected, such as the selective toxicity of SL-bergamottin on human cervix cancer cell lines and the involvement of SL-bergamottin in the metabolism performed by CYP3A4, the enzyme that participated in the metabolism of 80% of the commercially available drugs, making it an important target to check for drug-drug interactions. First we successfully re-created the experimental binding mode of ketoconazole, a known strong inhibitor of CYP3A4, and then after this validation step, the target-ligand complex of bergamottin and SL-bergamottin were *de novo* described with the same method. A structural explanation, that is in good agreement with the *in vitro* results was given on why is SL-bergamottin a stronger inhibitor of CYP3A4, than the mother compound, bergamottin. Furthermore, the novel information on the binding mechanism of the investigated compounds offer valuable help in the detection of plausible off-target and metabolic drug-drug interactions. Binding to CYP3A4 is not sensible to water molecules, in this case computational docking can be used as it is. The same approach would not have been successful in the case of viral ion channels. The binding and potential repositioning of clinically used drugs, amantadine, rimantadine and a third derivate, spiro-adamantyl-amine was investigated to the viral ion channel of the influenza A virus (M2A) and to that of the SARS-CoV-2 virus (EC2). For these two targets we had to face two of the greatest challenges of computational docking, modelling induced fit binding and the inclusion of water molecules into computational docking (position and partial charges). To create the fully hydrated target-ligand complex structure of the viral ion channels, a new protocol was made, which is called HydroDock. HydroDock is based on well-known, popular and freely available software and program packages, which were used and combined in a novel way. During the development of the HydroDock protocol, we showed, that the often used fast, dry docking methods lead to misleading results in the case of viral ion channels, that are not sufficient for drug design projects. The correct positioning of water molecules, the orientation of their hydrogen atoms and selecting the appropriate partial charge system proved to be inevitable to reproduce experimental binding modes of the drugs. The developed protocol was validated on the experimentally available target-drug complexes of the ion channel of the influenza A virus, and after successful reproduction of the reference systems, the protocol was transferred to the new target, SARS-CoV-2 E protein. With a target-based drug design approach, the amantadine derivates were repositioned to the unfortunately current viral target, that caused

the COVID-19 pandemic, reaching good agreements with experimental results on the new target also. The need to elucidate the role of water molecules in ligand binding often occurs but is rarely met appropriately. In the light of these findings, hopefully HydroDock will take part in future drug design projects to handle this great burden.

5. Thesis points

In the light of my previously listed aims, and the results of the dissertation, I summarize the most important findings of my doctoral dissertation work in the following thesis points.

1. The atomic resolution structures of bergamottin and SL-bergamottin in complex with CYP3A4 enzyme were de novo produced, as pillars of the binding mechanisms.

2. Knowing the results of the previous points, a comparative structural elucidation was given to the in vitro effects of bergamottin and SL-bergamottin. In good agreement with the experimental results, SL-bergamottin has a more favourable calculated energy of binding and is a stronger inhibitor of CYP3A4, than bergamottin.

3. It was shown that the correct partial charge system and H-atom orientation of water molecules are inevitable in accurately performing computational docking to produce the atomic resolution target-ligand complexes.

4. With systematic investigation, the protocol of molecular mechanics calculations and parameters were determined to improve the positional precision of dry docked ligand binding modes.

5. A new protocol, HydroDock was developed, which can handle one of the greatest burdens of computational docking, the incorporation of water molecules when docking on influenza A and SARS-CoV-2 viral ion channel targets. The protocol merges the methods of dry docking and the hydration of the target protein with explicit water molecules (using the conclusion of the previous two points) and produces hydrated target-ligand complexes.

1.

6. Acknowledgements

This work was funded by the Hungarian National Research, Development and Innovation Office (K123836, K134214), 2017-1.2.1 NKP-2017-00002 (NAP-2; Chronic Pain Research Group), EFOP-3.6.1.-16-2016-0004 and GINOP 2.3.2-15-2016-00050 “PEPSYS”. This work was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. We acknowledge the support from the Governmental Information Technology Development Agency, Hungary. The work was supported by the ÚNKP-21-5. New National Excellence Program of the Ministry for Innovation and Technology. Supported by the PTE ÁOK-KA No:2019/KA-2019-31. The project has been supported by the European Union, co-financed by the European Social Fund. Project name and code: Comprehensive Development for Implementing Smart Specialization Strategies at the University of Pécs, EFOP-3.6.1-16-2016-00004.

I am thankful to my PhD supervisor, Dr. Csaba Hetényi, from whom I have learned a lot, and who was always ready to help with my questions and to bring my work further. I am especially thankful for he always listened to my ideas and helped to work them out. I am grateful to all my colleagues of the Unit of Pharmacoinformatics Viktor Szél, Dr. Rita Börzsei, Bayartsetseg Bayarshaikan. Furthermore, I am thankful to Dr. Balázs Bognár, Prof. Dr. Tamás Kálai, Dr. Miklós Poór, Dr. Violetta Mohos, and Dr. Eszter Fliszár-Nyúl, who raised the standard of my work with their experimental collaboration. Finally, I am grateful to my family, my partner and my friends, who always supported me during my time as a PhD student. I owe a distinguished gratitude to Via, who was always by my side. I would like to offer this thesis to the memory of my belated Grandmother.