Calculation of target-ligand complexes using fragment-based methods of molecular design

Summary of the PhD thesis

István Horváth

Supervisor
Dr. Habil. Csaba Hetényi
associate professor
University of Pécs
Medical School
Department of Pharmacology and Pharmacotherapy

University of Szeged
Doctoral School of Chemistry
Szeged
2019.

1. Introduction

In the past half century, computer-aided chemical processes have enabled rapid modeling of the protein structure, their ligands and their interactions at atomic or more detailed resolution. Molecular mechanics (MM) handles the molecules at atomic level, the quantum mechanics (QM) studies the motion and interaction of the nuclei and electrons. The formation of complexes can be modelled and predicted by docking methods. The docking MM methods predict the preferred binding position and structure of a molecule when bound to each other to form a stable complex. Knowledge of the preferred binding position - either from experimental studies, or from previous computational calculation - may be used to predict the strength of association or binding affinity between two molecules. The molecular dynamics (MD) is usually applied for investigations on interactions between smaller ligand and large target molecules.

While the biomolecules act in water, but water is not only a solvent for target-ligand interactions but a medium as well, it is important to involve the water molecules into the investigations. The computational MD allows rapid calculations of the large biomolecular systems in surrounding water. New approaches to predicting hydration networks for biomolecular surfaces and complex interfaces are also based on molecular dynamics.

2. Aim of the work

1. Automated and standardized toolkits were elaborated to study the structural and thermodynamic properties of target-ligand complexes with fragmenting the structure of the target and the ligand.

2. We investigated for a direct relationship between structure of the target-ligand and thermodynamic properties. Involving the entire target structure in a QM calculation is not feasible within reasonable calculation time, thus, QM calculation necessitates an extraction of the interface region of the target-ligand complex. This extraction contains the ligand, the residues of the target near the ligand, and the water molecules in the interface region.

3. We investigated a hybrid water model that yields the best correlation between the heat of formation calculated at semi-empirical quantum chemical level and experimental binding enthalpy changes.

4. Our aim was to make the developed protocol accessible to anyone on a web server.

5. Our aim was to determine the structure of histone peptide complexes by fragmenting, using the blind docking, wrapping of the target and assembly of the resulting structures.

6. Our aim was to calculate the structure of water networks and their temperature dependence in complex interfaces and on the whole protein surface.

7. We have attempted to provide an interpretation of the success of the MD calculations of water structure.

3. Methods

During the calculations the first step was the hydration of the studied systems and complexes, these structures were calculated with the MobyWat based on MD. MobyWat calculates the surface and the interface explicit water molecules. We applied COSMO as implicit water model.

For docking, we used blind docking method, both docking and heat calculations were done on one of the supercomputers of Szeged, Debrecen or Budapest.

The heat of formation was calculated by using the MOPAC software package, the results of blind docking were processed on desktops.

The webserver is accessible on an Ubuntu Linux server running in the cloud, the web page is based on HTML and PHP, and the fragmented structure is rendered using JSmol based on JavaScript.

The structural conversions were performed by the Open Babel program.

The linking of the histone ligand was automated with a program written in Java. The calculation of the water networks was performed by MobyWat in NetDraw mode, the Gephi program was used to display the water network graph.

4. Results and Discussion

4.1. Calculating a simple scaling factor between the heat of formation and the binding enthalpy for target-ligand systems

1. We introduced a new an automated and standardized fragmentation protocol named Fragmenter, which cuts the interface region of the MM optimized protein-ligandum systems for QM calculations.

2. The protocol fostered a series of fast calculations of ligand binding enthalpies at the semi-empirical QM level and our results showed that the best correlation can be achieved using the COSMO implicit water model combining with explicit water molecules from the earlier published MobyWat method.

3. The proper distance of the explicit water molecules was determined to achieve the high correlation between the calculated heat of formation and the experimental binding enthalpy.

4. A single scaling factor was derived for the conversion of QM reaction heats to binding enthalpy values without the involvement of an intercept. In this way, a new, direct relationship was established between molecular structure and binding thermodynamics.

5. The fragmenting protocol, the hydration model and the scaling factor can be used as an automated and standardized toolkit for enthalpic optimization of drug candidates.

6. The Fragmenter is freely available as a web service.

4.2. Az FBD (Fragment Blind Docking) method

1. A new method, FBD, was introduced and tested on the examples of complexes of reader and writer proteins with histone peptide fragments, FBD benefits from the philosophy of its parent methods, fragment and blind docking.

2. After the fragmenting and wrapping cycle a new procedure was introduced for linking and welding of the fragments.

4.3. Investigation of water networks on surface and interface of biomolecule complexes in molecular dynamics (MD) calculations

1. It was demonstrated that the efficiency and reproducibility of the hydration structure determination of the in MD calculations are mainly temperature dependent.

2. The MobyWat program is effective in predicting hydration structures and creating water networks using MD trajectories.

3. We calculated the structure of water networks and revealed their temperature dependence for complex interfaces and whole protein surfaces. With network considerations we provided an interpretation of the success of molecular dynamics computation of the water structure.

5. Summary

The first topic was molecular design; the knowledge of atomic resolution structures is essential for predicting the binding affinity of drug candidates (ligands) to their targets.

The size of the target-ligand complexes and the lack of hydration structure make it difficult to simulate them at the quantum mechanical (QM) level, which often requires proper handling of electronic effects.

In order to help the QM design of new drugs, we have introduced a protocol to fragment complex interface structures with a defined number of explicit water molecules at specific positions.

The protocol promoted a series of rapid calculations of ligand binding enthalpies at the semi empirical QM level.

Various ligand sizes were involved: from small aromatic ligands to large, peptide helix ligands.

Comparison of different solvent models showed that the presence of the explicit water molecules with precisely predicted positions in the complex interface significantly improved the correlation with the experimental results.

A simple scaling factor was obtained for the relationship between heat of formation in QM and experimental binding enthalpy values.

This factor establishes a relationship between molecular structure and bonding enthalpy using QM calculations.

The results of the fragmenting protocol, hydration model, and scaling factor can be used as an automated and standardized toolkit for enthalpy optimization of drug candidate complexes.

The second topic was the determination of the structure of histone reading complexes using a new method, since this information is the key for the discovery of histone code and the design of new drugs.

The large number of variations of histones yields a large number of possible complexes; large scale structure determination of such complexes is challenging even for high throughput techniques.

The computation of such complexes is difficult due to the large size and flexibility of the peptides and the shallow binding surface of the readers.

Moreover, the location of the binding sites is often unknown, which requires blind docking to search the entire surface on the target protein.

In order to accelerate the work in this area, we have introduced a new method for determining the structure of the histone H3 peptide tail on their targets.

Both of these topics are closely related to the water already mentioned in the introduction, its role as a solvent, as a mediator and as a networking medium in the formation of complex structures.

Calculations using explicit solvent models allow us to track water molecules individually on a real-time scale, providing information on their mobility and predicting their networking with biomolecular solutes and other water partners.

Here we investigated the effect of the key parameters of MD simulations on the quality of such predictions, and systematically examined the effect of temperature.

6. Publications related to the scientific topic of the dissertation (MTMT ID: 10052611)

Horváth, I.; Jeszenői, N.; Bálint, M.; Paragi, G.; Hetényi, C., A fragmenting protocol with explicit hydration for calculation of binding enthalpies of target-ligand complexes at a quantum mechanical level. *Int. J. Mol. Sci.* **2019**, 20, 4384. IF_{2018}: 4,183

Bálint, M.; **Horváth, I.**; Mészáros, N.; Hetényi, C., Towards unraveling the histone code by fragment blind docking. *Int. J. Mol. Sci.* **2019**, 20, 422. IF_{2018}: 4,183

Jeszenői, N.; Schilli, G.; Bálint, M.; **Horváth, I.**; Hetényi, C., Analysis of the influence of simulation parameters on biomolecule-linked water networks. *J. Mol. Graph. Model.* **2018**, 82, 117-128. IF: 1,863

The cumulative impact factor of the referenced articles: 10,229

7. Publications related to the present thesis

Bálint, M.; Jeszenői, N.; **Horváth, I.**; van der Spoel, D.; Hetényi, C., Systematic exploration of multiple drug binding sites. *J. Cheminform.* **2017**, 9, Article number: 65.

Jeszenői, N.; Bálint, M.; **Horváth, I.**; van der Spoel, D.; Hetényi, C., Exploration of interfacial hydration networks of target−ligand complexes. *J. Chem. Inf. Model.* **2016**, 56 (1), 148-158.

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

8. Presentations and posters related to the present thesis

Hetényi, C.; Bálint, M.; Schilli, G. ; **Horváth, I.**, Computational blind docking of ligands to drug targets: Methodology and applications (**2018**), 2nd Edition of Global Conference on Pharmaceutics and DrugDelivery Systems 2018, Roma, Italy.

Bálint, M.; Hetényi, C.; Schilli, G. ; **Horváth, I.**, Discovering multiple drug binding sites and modes with Wrap 'n' Shake (**2018**), 2nd Edition of Global Conference on Pharmaceutics and DrugDelivery Systems, Roma, Italy.

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

Hetényi, C.; Bálint, M.; Jeszenői, N.; **Horváth, I.**, Improvements of computational optimization of drug-target interactions, 93 p. (**2017**), Global Conference on Pharmaceutics and Drug Delivery Systems (PDDS 2017), Valencia, Spain.

Jeszenői, N.; Bálint, M.; **Horváth, I.**; Hetényi, C., Komplex szerkezetek vízhálózatai (**2016**) Előadás, KeMoMo–QSAR Szimpózium (**2016**), Az MKE Szerves- és Gyógyszerkémiai Szakosztályának QSAR és Modellezési Szakcsoportja és az MTA Szegedi Akadémiai Bizottságának Kemometria és Molekulamodellezés Munkabizottsága által szervezett kétnapos szimpózium, Miskolc, Hungary.

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

Hetényi, C.; Jeszenői, N.; Bálint, M.; **Horváth, I.**; Lopata, A. (szerk.) Fehérjék hidrátszerkezete: modellezzem vagy mérjem? (**2015**), Előadás, KeMoMo–QSAR Szimpózium.

9. Other publications

Bálint, M.; Jeszenői, N.; **Horváth, I.**; Ábrahám, M. I.; Hetényi, C., Dynamic changes in binding interaction networks of sex steroids establish their non-classical effects. *Sci. Rep.* **2017**, 7, Article number: 14847.

Hetenyi, C.; Jeszenoi, N.; Balint, M.; **Horvath, I.**, Answers to current challenges in target-based drug design, In: Laszlo, Gelencser; Zoltan, Horvath (szerk.) (**2016**), BJMT Conference of Applied Mathematics 2016: Book of Abstracts.

Solymosi, F.; Bugyi, L.; Oszkó, A.; Horváth, I., Generation and Reactions of CH2 and C2H5 Species on Mo2C/Mo(111). *J. Catal.* **1999**, 1, 160-169.