Final Report

on the project entitled

Calculation of structure and thermodynamics of peptide complexes of epigenetic and therapeutic importance

Project No: NKFI K 123836
PI: Csaba Hetényi, PhD, Dr. (Pharm) Habil, Assoc. Prof., Unit Head
Department of Pharmacology and Pharmacotherapy, Pharmacoinformatics Unit
Medical School, University of Pécs

The project was originally planned for the 2017-2021 period and it was extended by one year due to administrative circumstances. In accordance with the suggestions of the NKFIH web page, the following report refers only to peer reviewed articles and omits citing conference participations (conferences are also listed among the publications attached to this report on-line). The results of the project are summarized in the following Sections with references to the articles attached in the publication list (the corresponding Action Points of the Work Plan are also indicated in the Section titles).

1 Development of the blind docking method (Action Point 1)

We aimed at the development and test of the blind docking (BD) methodology. Accordingly, a new method was published. A combination of fast BD with molecular dynamics (MD) was presented and tested (Bálint et al. 2017, Sci Rep) for the elucidation of changes in binding interaction networks of sex steroids responsible for their non-classical effects. MD also allowed the estimation of kinetic stability of binding poses found by fast BD. This work was a forerunner of our new method, Wrap 'n' Shake which combines the advantages of fast BD and explicit water MD methods (Bálint et al. 2017, J Cheminf). The method was published in J Cheminformatics and the corresponding source code was made publicly available under GNU Public License. The method provides a solution for the problematics of finding multiple binding sites with BD. Targets with multiple (prerequisite or allosteric) binding sites have an increasing importance in drug design. Experimental determination of atomic resolution structures of ligands weakly bound to multiple binding sites is often challenging. BD has been widely used for fast mapping of the entire target surface for multiple binding sites. Reliability of BD has been limited by approximations of hydration models, simplified handling of molecular flexibility, and imperfect search algorithms. To overcome such limitations, the Wrapper step of Wrap 'n' Shake systematically "wraps" the entire target into a monolayer of ligand molecules (**Fig. 1**) using several fast docking cycles. After including explicit water molecules (hydration), the functional binding sites are selected by a rapid molecular dynamics shaker. Wrap 'n' Shake can be considered as a solution of the above BD problems, and it was tested on biologically important systems such as mitogen-activated protein, tyrosine-protein kinases, key players of cellular signaling, and farnesyl pyrophosphate synthase,

a target of antitumor agents. A book chapter was also published with a practical guide to the usage of Wrapper (Hetényi et Bálint 2020) in the popular series Methods in Molecular Biology.

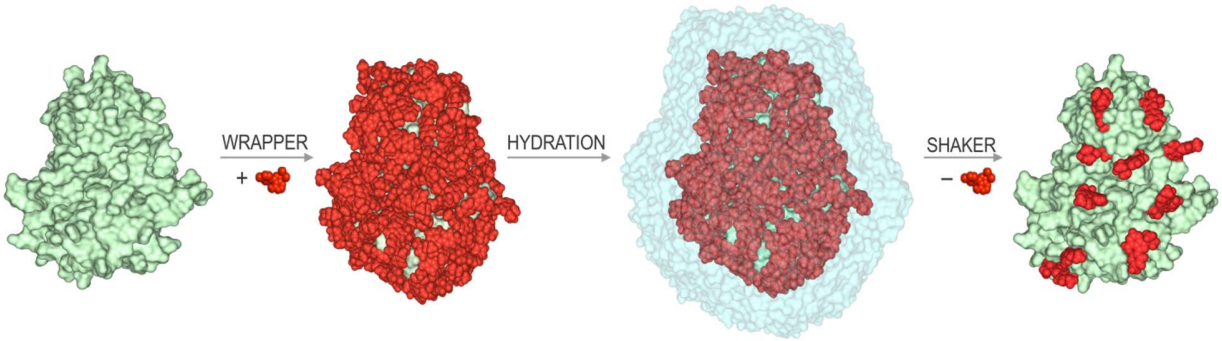

Fig. 1 The main steps of the Wrap 'n' Shake method. Target and ligand are shown in green and red, respectively. Documentation and source codes are available at http://www.wnsdock.xyz

Another development of BD resulted a new methodology, the fragment blind docking (FBD) which establishes a combination of our previous Wrapper algorithm and a fragmenting approach (Bálint et al. 2019). The FBD method can be used for the calculation of atomic resolution structure of complexes of histone peptides and their reader/writer target proteins. Notably, the present project is aimed at the calculation of "… peptide complexes of epigenetic and therapeutic importance", and this proved to be a rather challenging task. The determination of such histone complexes is demanding due to the large number of their variations, considerable size, flexibility of peptides and the shallow binding surfaces of the readers. The paper (Bálint et al. 2019) presented a promising first attempt for solving these problems by FBD.

The FBD technique was applied in a modified form for the determination of the structure of the somatostatin peptide with its receptor SSTR4 (Börzsei et al. 2022, IJMS, paper 6878). Somatostatin is a regulatory peptide important for the proper functioning of the endocrine system, local inflammatory reactions, mood and motor coordination, and behavioral responses to stress. Somatostatin exerts its effects via binding to G-protein-coupled somatostatin receptors of which the fourth subtype (SSTR4) is a particularly important receptor mediating analgesic, anti-inflammatory, and anti-depressant effects without endocrine actions. Thus, SSTR4 agonists are promising drug candidates. Beyond the static complex structures, the binding mechanism of somatostatin was also elucidated in the explicit water molecular dynamics (MD) calculations, and key binding modes (external, intermediate, and internal) were distinguished in the paper. Similarly, a fragment-based approach led to the construction of complex structures of epigenetic importance the complexes of the unmodified human PRMT5-methylosome protein 50 (MEP50) structure and its T80-phosphorylated variant in complex with the full-length histone H4 peptide. The full-length histone H4 was in situ built into the human PRMT5-MEP50 enzyme using experimental H4 peptide fragments. Extensive MD simulations, structure and energy analyses

provided an atomic level explanation of important experimental findings (Börzsei et al. 2022, IJMS, in the press).

Besides the allosteric binding sites often captured by BD, prerequisite binding sites are also of great importance for drug design. In this project, we first tested a combination of focused fast docking methods with MD for investigation of agonist binding to TRPA1. Our study (Zsidó et al. 2021, Pharmaceuticals) identified the prerequisite binding modes of three agonists and showed how the binding of a ligand to the prerequisite site can forecast its successful docking to the final binding pocket. The prerequisite binding sites proved to be milestones on the association/dissociation pathway of the agonists, important in mechanism-based design. Thus, amino acids identified along the binding pathway will serve as new target sites for the design of reversible binding of future agonists, beyond the well-known target of the covalent binding pocket of TRPA1. The search for prerequisite binding modes was also a central question of our new NetBinder method (Bálint et al. 2022) which combined Wrap 'n' Shake (**Fig. 1**) with a network-based approach. NetBinder is based on atomistic simulations of the full inhibitor binding process and provides a networking framework on which to select the most important binding modes and uncover the entire binding mechanism, including previously undiscovered events. NetBinder (**Fig. 2**) was validated by a study of the binding mechanism of blebbistatin (a potent inhibitor) to myosin 2 (a promising target for cancer chemotherapy).

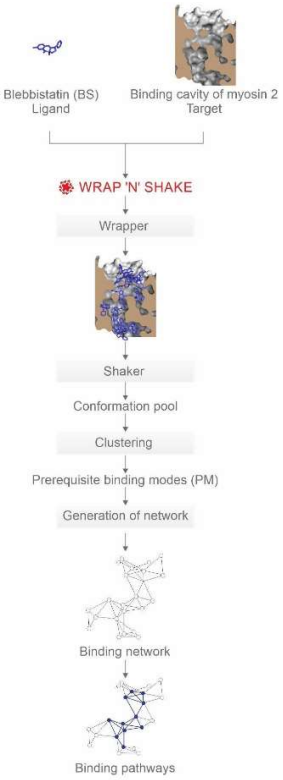

Fig. 2 The NetBinder method allows the mapping of alternative binding pathways via the determination of prerequisite binding modes with blind docking

2 Calculation of binding thermodynamics (Action Points 2, 4-8)

According to our plans, a new scheme was developed for quantum mechanical (QM) calculation of binding enthalpy. As it was indicated in the yearly report, with a slight modification of our plans, we focused on the enthalpic (complex free energy) calculators and skipped the plans for separate calculation of binding entropy. In our study (Horváth et al. 2019), a fragmentation and hydration protocol was elaborated (**Fig. 3**) and good correlations were achieved with experimental binding enthalpy data. The fragmentation protocol Fragmenter was implemented in a web server (http://www.fragmenter.xyz). The test systems were selected from our database of thermodynamic and structural data assembled in the previous steps of the research.

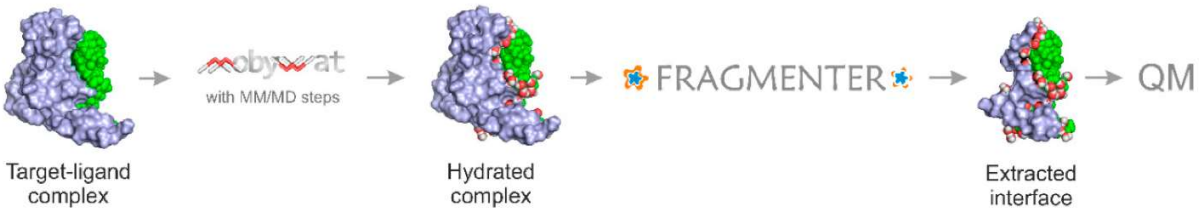

Fig. 3 The Fragmenter protocol prepares an extracted target-ligand interface applicable in QM calculations.

It was shown that a hybrid water model of our protocol including explicit water molecules provided the best results among the investigated solvent models. The study was selected for the cover illustration of IJMS (Vol. 20. Iss. 18). Besides small ligands, large peptides of epigenetic and therapeutic importance were also involved in the study. In our joint review (Zsidó et Hetényi 2020) the above Fragmenter study was surveyed in the context of other studies on epigenetic systems. The review is focused on structure, binding affinity and biological activity in the epigenome and nicely fits to the scope of this research project.

3 Development and tests of interface hydration techniques and their combination with the molecular docking engine (Action Point 3)

In a methodological study, we investigated the effect of key parameters of molecular dynamics simulations on the quality of prediction of hydration structure (Jeszenői et al. 2018) using our MobyWat program. Accordingly, systematic scans on temperature, pressure, force field, explicit water model and thermodynamic ensemble were performed. Explanations of optimal parameter values were provided using structural examples and analyses of the corresponding hydration networks.

The knowledge of atomic positions and explicit interactions of water molecules located in the target-ligand interface would be crucial for correct docking and scoring of the ligand molecules.

However, such atomic positions are often missing even from experimental structures, and therefore, calculation of the corresponding thermodynamics was an impossible mission from available experimental structures. To overcome this problem, a new protocol was developed and published (Zsidó et al. 2021, JCIM) for calculation of hydrated, atomic resolution target-ligand complex structures from scratch, i.e. using only the structures of the individual partners (ligand, target, waters) as input (**Fig. 4**). The protocol was named HydroDock and tested on binding of ligands to transmembrane ion channels of influenza and SARS-CoV-2 viruses.

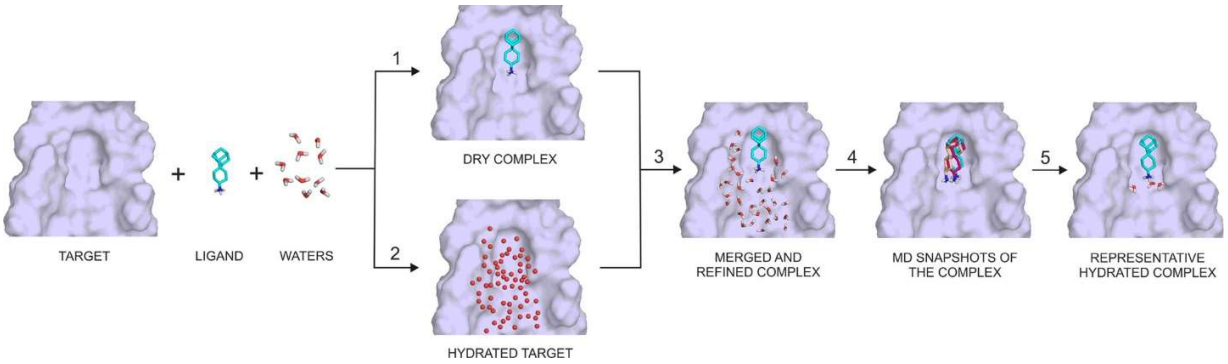

Fig. 4 The HydroDock protocol is a combination of dry docking (1) methodology with computational hydration (2) and MD calculations (4).

Excellent agreements were found with available experimental data. In the cases of viral ion channels, interfacial water structure is a key to drug action. Thus, the new HydroDock protocol will help antiviral drug design against not only COVID-19, but also other viral diseases. Using HydroDock, we designed new molecules with similar or higher binding affinity than the existing ones used in the above tests (experimental researchers were contacted to test the new molecules in vivo). The above problems regarding water structure were also discussed in details in our joint review (Zsidó et Hetényi 2021) accepted in the prestigious journal Current Opinion in Structural Biology.

4 Applications

In various collaborations, docking methodologies were extensively applied for elucidation of the binding of ligands of therapeutic importance to various target proteins including serum albumin (Poór et al. 2017; Faisal et al. 2018, both papers, 2019; Mohos et al. 2018, 2020, Biomolecules; Fliszár-Nyúl et al. 2021, 2022), cytochromes (Mohos et al. 2020, DMD; Zsidó et al. 2020), SSTR4 (Szőke et al. 2020; Kántás et al. 2021), and Ndr/Lats kinases (Parker et al. 2020). Such collaborations proved to be very useful providing new test systems and the experiences obtained helped to outline the limitations of the docking methodologies. They were especially important for the developments of our new approaches of BD and prerequisite binding pocket search described in **Section 1**.

5 Scientometry and dissemination of the results

The outcomes and scientometry of the present project are summarized in **Table 1**. Besides peer reviewed articles, the results were disseminated on web sites, several conferences, and in three PhD dissertations.

Table 1 The planned and actual outcomes of the project

Planned Outcomes	Actual Outcomes
Preparation of **a web site**, scripts and programs for the docking method	**Two web sites** were launched. www.wnsdock.xyz for the Wrap 'n' Shake blind docking method, with download option for scripts and an open source program. www.fragmenter.xyz for the on-line implementation of the Fragmenter protocol.
Submission of a total of **8 manuscripts** to international peer-reviewed journals	A total of **26 articles** were accepted in international peer-reviewed journals with a total impact factor of **127.985** (only articles with NKFIH acknowledgments were counted). The results were also disseminated in **1** book chapter, **21** conference lectures, and **3** posters.
The results form the basis of **2 PhD Dissertations**	**3 PhD Dissertations** of the following colleagues were defended successfully in three different doctoral schools, based on the results of this project. • Mónika Bálint, ELTE, 2018 • István Horváth, SZTE, 2019 • Balázs Zoltán Zsidó, PTE, 2022