Ágnes Peragovics

Virtual affinity fingerprints in drug discovery: The Drug Profile Matching method

PhD Theses

Supervisor: András Málnási-Csizmadia DSc.
Associate Professor

Structural Biochemistry Doctoral Program
Doctoral School in Biology

Program Leader: Prof. László Nyitray DSc.
Head of the School: Prof. Anna Erdei DSc.

Eötvös Loránd University
Budapest, Hungary
2014

Introduction

Single target based approaches have dominated pharmacological research in the last few decades, but despite the great efforts, these strategies proved insufficient in many cases [1]. As a result, a considerable shift from the traditional one drug – one target paradigm towards the theory of polypharmacology has been taking place recently in the drug development field [2]. According to polypharmacology, many drugs exert their effects by multitarget interactions and these multiple actions seem to be essential to obtain efficacy in complex diseases [3].

This increasingly recognized theory also influences the methodological side of state-of-the-art drug design strategies. Accordingly, the simplified approach considering only a single interaction between a bioactive compound and the human proteome is not suitable for further development in many therapeutical fields. In order to achieve efficacy in complex diseases, interference of a compound on multiple sites is required and this issue must be included in current drug development methods.

In accordance with the theory of polypharmacology, molecules can be characterized by their affinities to a panel of proteins that can simulate the significant interactions occurring in a complex biological system. Even though the resulting pattern not necessarily contains the actual interactions of the molecule in a human organism, it encodes fundamental components relevant to binding and therefore carries the potential for biological activity prediction. It is important to note that the individual interaction values of the fingerprint are not examined in this case, but the whole profile serves as an approximation of the interaction ability of a compound with the human proteome. The literature does not have a uniform terminology for these bioactivity profiles but they are often referred to as *affinity fingerprints*. They can be classified as *in vitro* or *in silico* affinity fingerprints based on the applied fingerprint generation method.

Aims

Affinity fingerprints provide a unique way for characterizing molecules. This description of chemical compounds has not been fully exploited in drug development. A type of *in silico* affinity fingerprint was chosen to further study because calculated binding free energies can be generated relatively fast and only structural information is required from the ligands' site. Our starting hypothesis was that the *in silico* generated affinity fingerprint of a drug, *i.e.*, a series of calculated binding free energy values for a set of proteins, correlates with the bioactivity properties of the drug. The affinity fingerprint was termed as *Interaction Profile* (IP) in our work.

I. Affinity fingerprints were already successfully applied in target fishing experiments (predicting targets for chemical compounds), however, known studies only focused on a couple of targets. Therefore, our primary aim was to include a considerably larger target pool and **construct a systematic *in silico* prediction method that is able to uncover the complex Target Profiles of drug molecules**. Drug Profile Matching (DPM) was chosen as a name for our approach as it reflects the main feature of this methodology, namely that it compares complex Interaction Profiles of drug molecules to large bioactivity profiles. Taking into account the enormous amount of information to be processed, multidimensional approaches were adopted in the method.

II. The secondary goal of this study was to investigate whether **broad pharmacological effects of drug molecules can be predicted** by exploiting the information content of Interaction Profiles. This wider application area has never been presented in connection with affinity fingerprints.

III. Thirdly, we wanted to demonstrate the added value of DPM to traditional molecular similarity-based approaches by **comparing its pharmacological effect prediction performance to 2D and 3D similarity searches**. It is very important to investigate whether the more complex DPM methodology can outperform the conventional approaches and define its feasibility domain.

IV. Lastly, we carried out experimental work to **test our findings in *in vitro* measurements** for selected effect categories. Experimental data justifying computational predictions can demonstrate the practical value of the developed methodology. We were interested to find out whether DPM can identify valuable candidate molecules for drug discovery that could be further studied and optimized.

Methods

Creation of the Interaction Profile matrix

1,175 FDA-approved drugs were docked to 135 proteins by DOVIS 2.0 software (DOcking-based VIrtual Screening) [4], using AutoDock4 docking engine [5], Lamarckian genetic algorithm and X-SCORE [6] scoring function (average X-SCORE value). Special care was taken to ensure that none of the selected proteins is involved in the mechanism of the studied drugs according to our current knowledge of drug actions. The minima of the binding free energies for each drug-protein pair were imported to the Interaction Profile matrix (Figure 1.). In this matrix, each row represents the IP of a given drug molecule that is analogous to an *in silico* affinity fingerprint.

Creation of the Target/Effect Profile matrix

Target and effect information on the 1,175 FDA-approved small-molecule drugs was exhaustively collected from DrugBank [7] and was further refined. In order to provide sufficient amount of information about the actives, only target/effect categories having at least 10 registered molecules were kept. A binary matrix called *Target Profile* (TP) matrix was then created based on the remaining groups that displays whether a drug interacts with a given target according to DrugBank (white cell marks the presence of the interaction while blue cell indicates that a given drug-target interaction is not documented, Figure 1.). A binary *Effect Profile* (EP) matrix was created analogously.

Multidimensional Analyses

Multidimensional analyses were applied to relate these two matrices (IP-TP and IP-EP, in separate analyses). In the first step, *canonical correlation analysis* (CCA) was performed between the whole IP matrix and a selected target/effect. As an output, maximally correlated factor pairs were created for each target/effect. These factor pairs were subsequently subjected to *linear discriminant analysis* (LDA) that identified the best discriminating surfaces and produced a classification function for each case. Using the mathematical formula of these functions, probabilities were calculated for every drug-target/effect pair and were collected in a new matrix termed as the Target/Effect Probability matrix. Each probability value in this matrix indicates the likelihood of exerting a given effect or interacting with a studied target for the drug in question.

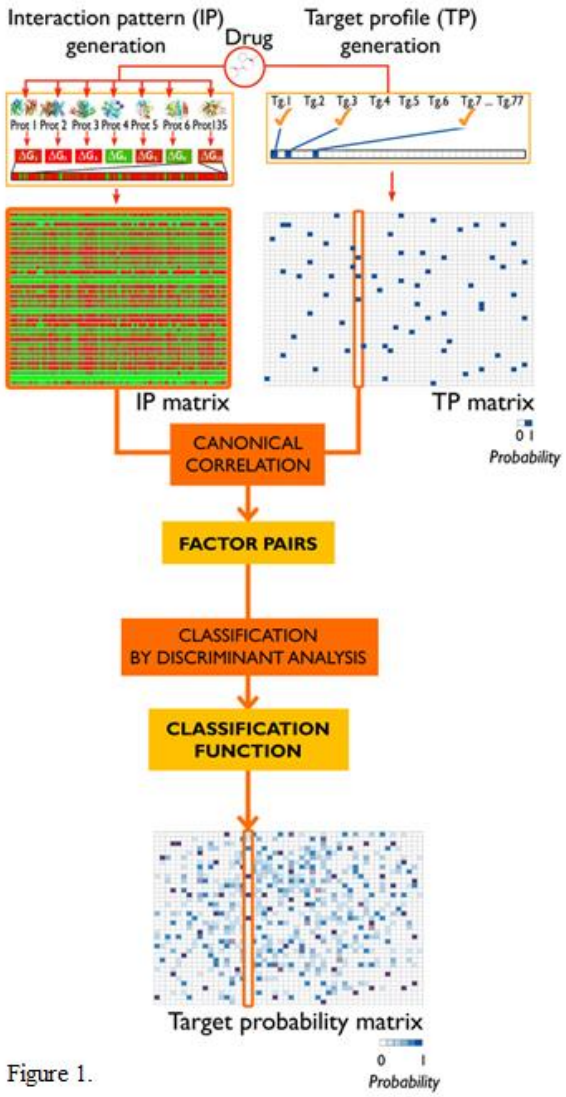

Figure 1.

The accuracy of the classification functions was determined by *Receiver Operating Characteristics* (ROC) analysis while the robustness of the results was investigated by the *10-fold cross-validation* procedure. Classification efficacy was compared to that of traditional 2D and 3D similarity searches by *random splitting* experiments. In two effect categories (*Angiotensin-converting enzyme inhibitor* and *Cyclooxygenase inhibitor*), false positive

molecules (previously not connected to the effect/target but predicted with high probability by DPM) were tested in *in vitro experiments* to prove the correctness of the predictions.

Results

The Drug Profile Matching method was developed that relates the complex drug-protein Interaction Profiles with known target/effect profiles of drugs by multidimensional techniques. The aims of my PhD work defined in the beginning were achieved as follows:

I. **A systematic *in silico* prediction method was constructed, which is able to uncover the complex Target Profiles of drug molecules**.

   I could establish a statistical relationship between the two different types of profiles (IP-TP) that enabled to construct target predictions for the drugs in question.

II. **Broad pharmacological effects of drug molecules could be predicted by extending the applicability domain of DPM.**

   In addition to successfully predicting targets, the DPM method was also applied for investigating pharmacological effects of drugs. Despite getting lower performance for effect prediction, encouraging results were obtained for a well-defined set of effects.

III. **The pharmacological effect prediction performance of DPM was compared to 2D and 3D similarity searches**.

   DPM was also compared to traditional molecular similarity-based methods to demonstrate their common features and the added value of the more complex methodology. The main strength of DPM was found to be its ability to handle groups of structurally diverse active molecules.

IV. **Part of the findings was tested in *in vitro* measurements for selected effect categories.**

50 drug repurposing candidate molecules were experimentally tested for two selected effect categories (ACE and COX inhibition) and a considerable part of the DPM predictions was justified. 33% of the predictions were confirmed experimentally for ACE and 23% for COX inhibition.

Publications concerning this thesis

Végner, L.; **Peragovics, Á.**; Tombor, L.; Jelinek, B.; Czobor, P.; Bender, A.; Simon, Z.; Málnási-Csizmadia, A.: Experimental confirmation of new drug-target interactions predicted by Drug Profile Matching, *J Med Chem* 2013, 56 (21), 8377.

Peragovics, Á.; Simon, Z.; Tombor, L.; Jelinek, B.; Hári, P; Czobor, P; Málnási-Csizmadia A.: Virtual affinity fingerprints for target fishing: a new application of Drug Profile Matching, *J Chem Inf Model* 2013 53 (1), 103.

Peragovics, Á.; Simon, Z.; Brandhuber I.; Jelinek B.; Hári, P.; Hetényi, C.; Czobor, P.; Málnási-Csizmadia A.: Contribution of 2D and 3D Structural Features of Drug Molecules in the Prediction of Drug Profile Matching, *J Chem Inf Model* 2012, 52 (7), 1733.

Simon, Z.; **Peragovics, Á.**; Vigh-Smeller, M.; Csukly, G.; Tombor, L.; Yang, Z.; Zahoranszky-Kohalmi, G.; Végner, L.; Jelinek, B.; Hari, P.; Hetényi, Cs,; Bitter, I.; Czobor, P.; Malnasi-Csizmadia, A.: Drug effect prediction by polypharmacology-based interaction profiling, *J Chem Inf Model* 2012, 52 (1), 134.

Other publications:

Simon, Z.; Vigh-Smeller, M.; **Peragovics, Á.**; Csukly, G.; Zahoranszky-Kohalmi, G.; Rauscher, A. A.; Jelinek, B.; Hari, P.; Bitter, I.; Malnasi-Csizmadia, A.; Czobor, P.: Relating the shape of protein binding sites to binding affinity profiles: Is there an association?, *BMC Struct Biol* 2010, 10, (1), 32.

Móczár, I.; **Peragovics Á.**, Baranyai P., Tóth K.; Huszthy P.: Synthesis and fluorescence studies of novel bis(azacrown ether) type chemosensors containing an acridinone unit, *Tetrahedron* 2010, 66, 16.

References

1. Knight, Z.A., H. Lin, and K.M. Shokat, *Targeting the cancer kinome through polypharmacology*. Nat Rev Cancer, 2010. **10**(2): p. 130-7.
2. Medina-Franco, J.L., et al., *Shifting from the single to the multitarget paradigm in drug discovery.* Drug Discov Today, 2013. **18**(9-10): p. 495-501.
3. Hopkins, A.L., *Network pharmacology.* Nat Biotechnol, 2007. **25**(10): p. 1110-1.
4. Jiang, X., et al., *DOVIS 2.0: an efficient and easy to use parallel virtual screening tool based on AutoDock 4.0.* Chem Cent J, 2008. **2**: p. 18.
5. Huey, R., et al., *A semiempirical free energy force field with charge-based desolvation.* J Comput Chem, 2007. **28**(6): p. 1145-52.
6. Wang, R., L. Lai, and S. Wang, *Further development and validation of empirical scoring functions for structure-based binding affinity prediction.* J Comput-Aided Mol Des, 2002. **16**: p. 11-26.
7. Wishart, D.S., et al., *DrugBank: a knowledgebase for drugs, drug actions and drug targets.* Nucleic Acids Res, 2008. **36**(Database issue): p. D901-6.