Acta Biochimica Polonica

Vol. 52 No. 3/2005, 737–740

on-line at: www.actabp.pl

Communication

Targeting drug-efflux pumps — a pharmacoinformatic approach*

Karin Pleban[1], Dominik Kaiser[1], Stefan Kopp[2], Michael Peer[2], Peter Chiba[2] and Gerhard F. Ecker[1]✉

[1]*Emerging Focus Pharmacoinformatics, Department of Medicinal Chemistry, University of Vienna, Wien, Austria;* [2]*Institute of Medical Chemistry, Medical University of Vienna, Wien, Austria;*
✉*e-mail: gerhard.f.ecker@univie.ac.at*

Received: 15 March, 2005; revised: 28 June, 2005; accepted: 18 July, 2005
available on-line: 04 August, 2005

In line with our studies on propafenone-type inhibitors of P-glycoprotein (P-gp), we applied several methods to approach virtual screening tools for identification of new P-gp inhibitors on one hand and the molecular basis of ligand–protein interaction on the other hand. For virtual screening, a combination of autocorrelation vectors and selforganising artificial neural networks proved extremely valuable in identifying P-gp inhibitors with structurally new scaffolds. For a closer view on the binding region for propafenone-type ligands we applied a combination of pharmacophore-driven photoaffinity labeling and protein homology modeling. On LmrA, a bacterial homologue of P-gp, we were able to identify distinct regions on transmembrane helices 3, 5 and 6 which show significant changes in the labeling pattern during different steps of the catalytic cycle.

Keywords: drug-efflux pumps, pharmacoinformatics, P-glycoprotein, self-organising maps, propafenone, photoaffinity labeling

Beginning at the early eighties, numerous active transport systems have been identified as being responsible both for the uptake and efflux of drugs. ATP-driven efflux pumps, such as P-glycoprotein (P-gp) are increasingly recognised as a major limiting factor for bioavailability and brain uptake (Terasaki & Hosoya, 1999). Additionally, these proteins are one of the basic mechanisms responsible for multiple drug resistance of tumour cells, bacteria and fungi. Up to now, hundreds of ABC-transporter have been identified, 49 of these being expressed in humans. Multidrug efflux pumps, such as P-gp, the breast cancer resistance protein (BCRP, MXR, ABCG2) and the multidrug resistance related proteins (MRPs, ABCCx family) very often show high promiscuity in ligand and substrate recognition. Thus, overexpression of P-gp in tumour cells leads to resistance to a broad variety of structurally and functionally diverse natural product toxins. An identical situation is seen for inhibitors, which resensitise multidrug resistant tumour cells. Currently, several compounds are in clinical phase three studies (Gottesman *et al.,* 2002). Although intensively studied, only little is known on the molecular basis of the drug–protein interaction. For P-gp, several 2D- and 3D-pharmacophore models have been described in the literature. Most of them relied on congeneric series of compounds (Wiese & Pajeva, 2001; Ecker & Chiba, 2001) and only few attempted to establish generalised models for ligand recognition (Klopman *et al.,* 1997; Pajeva & Wiese, 2002). In this paper, we describe several pharmacoinformatics approaches both for identification of new inhibitors of P-gp and for a deeper understanding of the molecular basis of the ligand–protein interaction.

VIRTUAL SCREENING PROTOCOLS BASED ON SELF-ORGANISING MAPS

Nowadays the drug development process starts with hits obtained in high throughput screening assays. Up to 1 000 000 compounds are biologically screened on a yes/no basis and the resulting

*Presented at the International Review Conference on Biotechnology, Vienna, Austria, November 2004.

Abbreviations: ABC, ATP-binding cassette; ADMET, absorption, distribution, metabolism, elimination, toxicity; HTS, high throughput screening; NBD, nucleotide-binding domain; P-gp, P-glycoprotein; SOM, self-organising map; TM, transmembrane; TMD, transmembrane domain.

hits are prioritised on the basis of novelty, patentability, synthetic accessibility and data obtained in early ADMET profiling programs. In parallel, *in silico* screening approaches are gaining increasing importance. They are mainly used to select subsets of large virtual combinatorial libraries, which then should show a higher incidence for biological activity (or at least higher drug likeliness) and thus lead to increased hit rates.

However, both bioavailability and drug-likeliness are rather complex issues, being mainly influenced by interactions with so called non-target proteins. Some of the key non-target proteins responsible for poor ADMET properties and/or high toxicity are the multidrug efflux pump P-glycoprotein, the cytochrome P450 enzyme complex and the human ether-a-go-go related gene potassium channel. All these proteins share a sort of promiscuity (or polyspecificity) in their binding interaction with ligands, which makes the use of rational drug design approaches to avoid undesirable drug–protein interaction rather difficult. This promiscuity may arise from a simultaneous accommodation of more than one ligand, multiple separate binding sites, protein flexibility or a combination of several of these properties (Ekins, 2004).

In the case of ligand–protein interactions being as complex as those mentioned above, machine learning methods seem to be the method of choice. Thus, we chose self-organising maps, a type of non-supervised learning algorithms, to implement a virtual screening protocol for identification of new inhibitors of P-gp. The principal ability of SOMs is to obtain a 2D-rendering of a multi-dimensional space which brings similar compounds in close vicinity on the map (Kohonen, 1982). SOMs have been successfully applied to designing combinatorial libraries, filtering HTS libraries, and to distinguish drugs from non-drugs (Anzali *et al*., 1998). First of all, a training set of 131 propafenone-type inhibitors of P-gp was used to explore the general possibility of distinguishing between active and inactive compounds using self-organizing maps (Fig. 1).

Using 2D autocorrelation vectors based on atom properties (Gasteiger, 1988), two different models were retrieved which showed good discrimination between compounds with high and low activity. For subsequent identification of new lead compounds, the 131 propafenone analogs were merged with 134767 compounds from the SPECS database and the complete data set was presented to the SOM under conditions identical to the training conditions. Thus, the compounds of the SPECS library placed in close vinicity to highly active propafenone analogs should also show high activity.

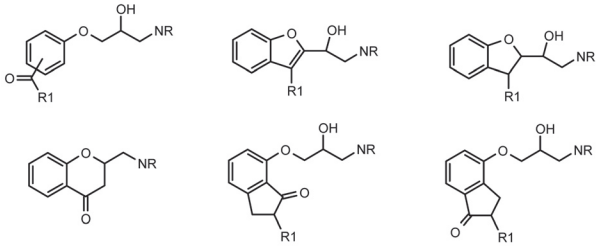

Figure 1. Structural scaffolds present in the propafenone-type training set.

Elimination of duplicates and restriction to compounds which co-localize with propafenones with an IC_{50}-value (half maximum inhibitory concentration) < 0.16 μmol/l gave a set of 43 compounds. Twelve out of these 43 compounds were located in the same neuron as the hitherto most active propafenone-type P-gp inhibitors. Some of these virtual screening hits show high structural analogy, which further reduced the number of hits to seven. These compounds were pharmacologically tested in the daunorubicin efflux assay. Briefly, P-gp-expressing cells are loaded with the fluorescent dye daunorubicin and the time-dependent decrease of cell-associated fluorescence in the presence of different concentrations of inhibitors is measured. Plot of the first order rate constants *vs*. inhibitor concentration leads to concentration/response curves which allow calculation of the respective IC_{50} values.

The results show that two of the compounds were highly active with IC_{50} values below 1 μmol/l (Fig. 2), four compounds had activities between 1 and 10 μmol/l, and only one compound was inactive. Thus, a combination of autocorrelation vectors and self-organising maps represents a useful tool for identification of new inhibitors of P-glycoprotein (Kaiser *et al.,* submitted).

COMBINED PHOTOAFFINITY LABELING – PROTEIN HOMOLOGY MODELING APPROACHES

In the absence of high resolution structural data for P-glycoprotein, one alternative is to generate a homology model based on the structure of a known sequence homologue. Up to now three full-length bacterial ABC-transporters have been crystallised. These are the lipid A transporter MsbA from *Escherichia coli* (Chang & Roth, 2001) and from *Vibrio*

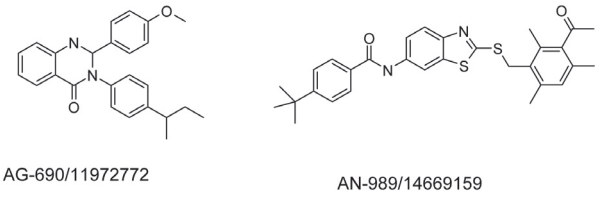

Figure 2. New hits for inhibitors of P-gp identified in an *in silico* screen of the SPECS compound library using a self-organising map.

cholerae (Chang, 2003) and the vitamin B12 transporter BtuCD from *E. coli* (Locher *et al*., 2002). Evaluation of the three full length transporters identified *Vc*-MsbA as the most suitable template for modeling the transmembrane (TM) regions, since BtuCD has 20 transmembrane segments while P-gp has a predicted number of 12 TMs. However, our first attempts in the field of protein homology modeling of drug efflux pumps focused on LmrA, a bacterial homologue of P-gp (Pleban *et al*., 2004). This is a half transporter showing six TMs, one NBD and acting as homodimer, responsible for resistance to 18 out of 21 clinically relevant antibiotics. As shown by van Veen *et al.* (1998), bacterial LmrA and human P-gp are functionally interchangeable with respect to both substrates and inhibitors. First, we built a homology model of LmrA using the template structure of *Vc*-MsbA. Although the resulting LmrA model correctly predicted polar amino-acid residues in TM segments to be oriented towards the central pore, the NBD:NBD interface failed to explain the catalytic requirements expected for an ABC transporter. Hence we replaced this NBD by an NBD generated from TAP1 as template. In order to correctly reestablish the TMD:NBD interface the intracellular loop 1 connecting TM helices 2 and 3 of each monomer was manually brought into position. The final model assembly was created by recalculating the linker region between TM helix 6 and NBD and subsequent splice repair with a cycle of 100 steps of Steepest Descent Algorithm minimization. The 3D model of the nucleotide-binding domain was checked in terms of geometry using ProStat and Verify 3D pull down of InsightII.

In ABC-transporters, binding and transport of substrates is mediated by the transmembrane domains. In our model, the TMDs are predicted to form a helical bundle which lines a central aqueous pore with access to the extracellular space. The TMD:TMD interface is formed by helices 3 and 5 of different monomers, which is in agreement with cross-linking data for P-glycoprotein (Stenham *et al.*, 2003). Previous substrate photoaffinity labeling experiments by our group indicated helices 3, 5 and 6 to be involved in substrate binding of propafenone-type ligands (Ecker *et al.*, 2004). These helices thus form the substrate binding domain and the model shows the close spatial proximity of these three helices. Additional support for an involvement of helices 5 and 6 in substrate binding comes also from photolabeling experiments with Rhodamine 123 (Alqwai *et al.*, 2003). Very recently, analogous result have been obtained for P-gp. Photoaffinity labeling experiments under different steps of the catalytic cycle clearly demonstrate that substrate binding occurs in the TMD:TMD interface (Pleban *et al.*, 2005). These protein homology models now may serve as tools for subsequent molecular dynamics simulations and docking studies, which will guide the target-based development of new efflux-pump inhibitors (Fig. 3).

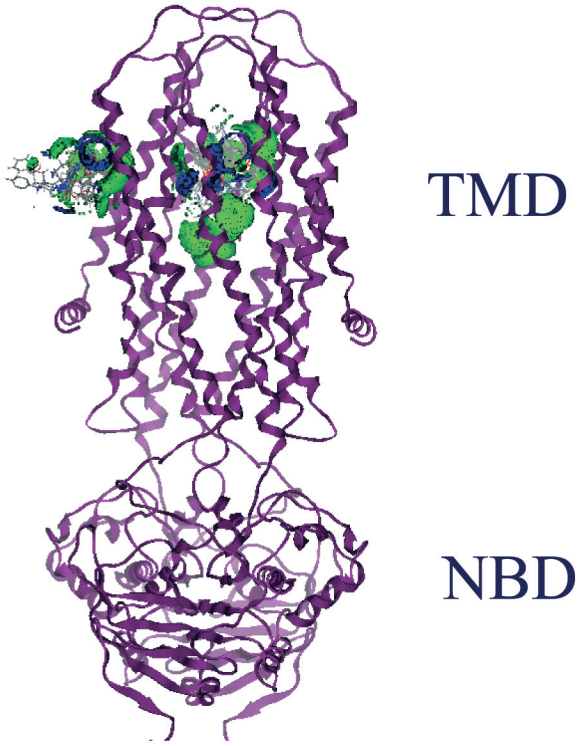

Figure 3. Preliminary results of docking of propafenone-type benzophenones to the LmrA homology model.

Acknowledgements

We gratefully acknowledge financial support by the Austrian Science Fund (grant P17014-B11) and the Austrian National Bank (grant 10654).

REFERENCES

Alqwai O, Poelarends G, Konings WN, Georges E (2003) *Biochem Biophys Res Commun* **311**: 696–701.
Anzali S, Gasteiger J, Holzgrabe U, Polanski J, Sadowski J, Teckentrup A, Wagener M (1998) *Persp Drug Discov Design* **9/10/11**: 273–299.
Chang G (2003) *J Mol Biol* **330**: 419–430.
Chang G, Roth CB (2001) *Science* **293**: 1793–1800.
Ecker GF, Chiba P (2001) *Recent Res Dev Med Chem* **1**: 121–137.
Ecker GF, Pleban K, Kopp S, Csaszar E, Poelarends GJ, Putman M, Kaiser D, Konings WN, Chiba P (2004) *Mol Pharmacol* **66**: 1169–1179.
Ekins S (2004) *Drug Discov Today* **9**: 276–285.
Gasteiger J (1988) In *Physical Property Prediction in Organic Chemistry*. Jochum C, Hicks MG, Sunkel J, eds, pp 119–138. Springer Verlag, Heidelberg.
Gottesman MM, Fojo T, Bates SE (2002) *Nat Rev Cancer* **2**: 48–58.
Klopman G, Shi LM, Ramu A (1997) *Mol Pharmacol* **52**: 323–334.
Kohonen T (1982) *Bioorg Med Chem Lett* **8**: 11–16.
Locher KP, Lee AT, Rees DC (2002) *Science* **296**: 1091–1098.
Pajeva I, Wiese M (2002) *J Med Chem* **45**: 5671–5686.

Pleban K, Macchiarulo A, Costantino G, Pellicciari R, Chiba P, Ecker GF (2004) *Bioorg Med Chem Lett* **14**: 5823–5826.
Pleban K, Kopp S, Csaszar E, Peer M, Hrebicek T, Rizzi A, Ecker GF, Chiba P (2005) *Mol Pharmacol* **67:** 365–374.
Stenham DR, Campbell JD, Sansom MS, Higgins CF, Kerr ID, Linton K (2003) *FASEB J* **17**: 2287–2289.
Terasaki T, Hosoya K (1999) *Adv Drug Deliv Rev* **36**: 195–209.
van Veen HW, Callaghan R, Soceneantu L, Sardini A, Konings WN, Higgins CF (1998) *Nature* **391**: 291–295.
Wiese M, Pajeva I (2001) *Curr Med Chem* **8**: 685–713.