Predicting the Toxic Potential of Drugs and Chemicals *In Silico*

Angelo Vedani, Markus A. Lill and Max Dobler
Biographics Laboratory 3R, Basel, Switzerland

Summary
Based on the 3D structure of the target protein (ERαβ, AR, PPARγ, TRαβ, GR; CYP3A4) or a surrogate thereof (AhR), the Biographics Laboratory 3R has generated a series of virtual test kits and validated them against 693 compounds. In a pilot project (ToxDataBase), both existing and new drugs or environmental chemicals can be screened for their endocrine-disrupting potential or the probability to trigger drug-drug interactions in silico. After peer testing (2007–8), it is planned to make the database available on the Internet.

Keywords: in silico, QSAR, toxicity, receptor, REACH, TocDataBase, Quasar, Raptor

Background Information

Quantifying the binding of a drug or chemical to a target protein *in silico*

In the past decade, computer-based (in silico) concepts matured into powerful tools for simulating and quantifying biochemical processes at the molecular level. This became possible due to over 40,000 protein structures known at atomic resolution, a more detailed understanding of the forces governing molecular interactions and the nowadays available computing power. The philosophy of structure-based design is based on the lock-and-key analogon — recognized as early as 1894 by Emil Fischer (Nobel Laureate, 1902) — the three-dimensional complementarity of drug and its target receptor or enzyme. In absence of structural information on the target protein, receptor-mapping technologies were developed allowing to construct 3D surrogates of the binding pocket. In a QSAR context, those can act as substitutes for the structure of the true biological receptor.

Poor pharmacokinetics, side effects and compound toxicity are not only frequent causes of late-stage failures in drug development but also a source for unnecessary animal tests. In silico methods are nowadays routinely used in the early stages of drug development. In the context of the REACH (Registration, Evaluation and Authorization of Chemicals) initiative of the European Union, computer-based experiments have received additional attention as they can predict the toxic potential of existing and hypothetical compounds. In silico techniques are fast, reproducible, and are typically based on human bioregulators, making the question of data transferability between species obsolete.

Reception of chemicals at biological relevant structures

Nuclear receptors are an important protein class in living organisms. They comprise a family of ligand-dependent transcription factors that transform extra- and intracellular signals into cellular responses by triggering the transcription of target genes. In particular, they mediate the effects of hormones (ligand) and hormonally active compounds (endocrine disruptors). Nuclear receptors are specific for the various steroid hormones, e.g. the estrogens (ER), androgens (AR), progesterones, and glucocorticoids. A number of receptor-mediated adverse effects by xenobiotics have been identified in the past. This includes toxicity mediated by the thyroid hormone receptor, the epidermal growth factor and aryl hydrocarbon receptor (AhR). The concern about chemicals which bind to these receptors and induce adverse, uncontrolled effects has created a need to both screen and monitor compounds before they are further developed as potential drugs or manufactured or released into our environment. At the Biographics Laboratory 3R, we have developed and validated a series of virtual test kits for the AhR, ERαβ, AR, PPARγ, TRαβ, and the GR. Models for the pregnan-X (PXR) and mineralocorticoid receptor are in preparation (Fig. 1).

Metabolic transformation

Competition of drugs for metabolization at Cytochrome P450 3A4 (CYP3A4) may result in undesired drug-drug interactions in patients. In addition CYP3A4 might transform chemicals into reactive metabolites. The development of a computational model to accurately predict the docking potential of a diverse set of ligand molecules was based on the X-ray crystal structure of the human CYP3A4 enzyme and a total of 48 structurally diverse

www.forschung3r.ch/en/projects/pr_75_00.html

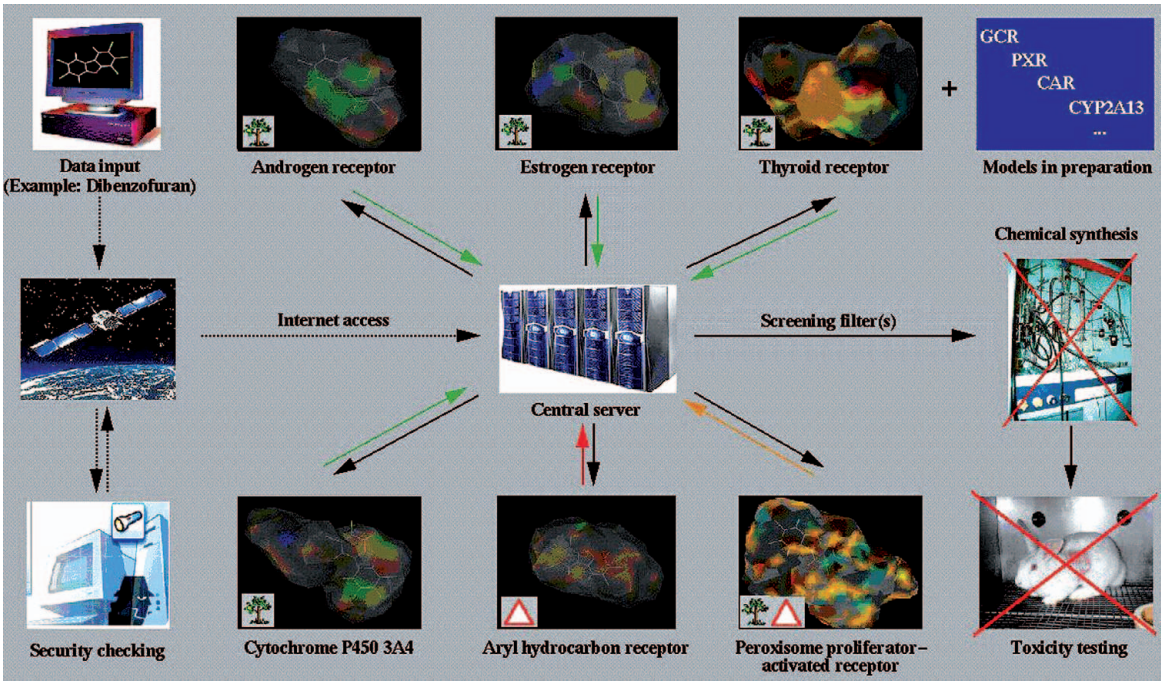

Fig. 1: Flowchart of the virtual laboratory with five receptor surrogates (estrogen, androgen, thyroid, and aryl hydrocarbon receptor and cytochrome P450 3A4) shown.
The hypothetical test compound (a dibenzofurane) is submitted *via* the Internet to a central server, which schedules the processors with the various receptor surrogates.
No significant binding affinities are calculated for the estrogen and androgen receptor and CYP3A4 (IC_{50} > 0.1 mM); against PPARγ, an IC_{50} of 58 µM is calculated. The interaction with the aryl hydrocarbon receptor (IC_{50} = 8 nM) indicates a high toxic potential of the compound; in this test example, the compound should consequently be removed from the evaluation pipeline at this point.

Tab. 1: Validation of virtual test kits.

System	# of compounds training+test=total	q^2	rms training	max. training	p^2	rms test	max. test
Aryl hydrocarbon	105 + 35 = 140	0.782	2.2	15.8	0.766	2.2	12.5
Estrogen α	80 + 26 = 103	0.895	2.0	8.6	0.892	2.9	9.5
Estrogen α*	80 + 23 = 106	0.787	0.9	2.5	0.505	1.7	7.3
Estrogen β*		0.703	1.4	6.7	0.561	1.9	5.7
Androgen	88 + 26 = 114	0.858	1.7	7.8	0.792	1.6	13.9
Thyroid α	64 + 18 = 82	0.919	1.8	4.3	0.814	2.5	10.0
Thyroid β		0.909	2.0	7.7	0.796	2.7	8.8
PPARg	75 + 20 = 95	0.832	1.4	6.2	0.723	1.4	3.9
Glucocorticoid**	80 + 30 = 110	0.743	1.2	6.1	0.623	2.3	6.1
CYP3A4	38 + 10 = 48	0.825	2.7	7.0	0.659	3.8	7.1

q^2 = cross-validated r^2, p^2 = predictive r^2; the rms and maximal deviation from the experimental binding affinity is given as factor in Ki or IC_{50}.
* model under development (ongoing PhD thesis); different compounds than for the 80/26 model above
** model under development (ongoing PhD thesis)

molecules (39 training and 9 test compounds). The results are given in Table 1 and were used to validate the predictability. A model for CYP2A13 is in preparation.

Construction of receptors

Model construction and optimization was achieved by combining protein and receptor modeling. First, the binding mode of each investigated compound was identified using the 3D structure of the target protein (for the Ah receptor, where no experimental structure is available, we used 4D pharmacophore generation instead) by automated, flexible docking combined with dynamic solvation of the binding pocket. Typically, the four lowest-energy orientations were composed in a 4D data set. In contrast to a 3D representation, each compound can be represented as an ensemble of conformations, orientations, protonation states, tautomers and stereoisomers (Tab. 2). This ligand superposition (the binding hypothesis) is then used for out multidimensional QSAR technologies named Quasar and Raptor. It is based on a quasi-atomistic model representation and explicitly allows for induced fit – the ligand-induced adaptation of the topology of the macromolecule (see, for example, www.biograf.ch and www.modeling.unibas.ch). Next, the models are validated using test compounds (ligands different from those in the training set), scramble testing and consensus scoring.

Validation of the virtual test kits

Vedani et al. (2006) gives details of model construction and validation. For generating a model a series of compounds (training set) is needed for which experimental binding affinities (Ki or IC_{50} values are available). The quality of the reproduction of these values is reflected by the q^2 value – the cross-validated r^2. Next, a series of ligands different from those of the training set (test set) is used to validate the model. The p^2 (the predictive r^2) value indicates the predictive power of the model. The predictivity can be given as $100 \times \sqrt{p^2}$ – e.g. 87.5% for the AhR.

Testing via Internet

The Biographics Laboratory 3R is presently implementing an Internet database for the screening of adverse effects triggered by drugs and chemicals *in silico*. The bioregulators described so far in this account (AhR, ERαβ, AR, PPARγ, TRαβ, GR and CYP3A4) represent the backbone of this Internet Database; PXR, MCR and 2A13 are in preparation. Within this framework, hypothetical or existing compounds can now be tested for their activity towards the various virtual test kits (Fig. 1) and their toxic potential may be estimated therefrom. Adverse effects mediated by receptors other than those compiled in the database can, of course, not be identified. Accordingly, the present approach based on receptor modeling will result in the production of false-negative results for classes of toxic chemicals which do not interact via receptor or which interact via so far unknown receptor-based pathways. Therefore, QSAR technologies may be used to identify the harmful potential of a drug or chemical and no false positives are produced. However, they are not (yet) suited to prove its harmlessness.

Outlook

Up to date, our concept has not produced any false-positive results. At the current level, however, false-negative predictions are still obtained, as a compound of interest cannot be tested against all potential receptors it may bind to *in vivo*. Some macromolecular targets will remain unknown, for others no experimental structure exists or too few affinity data are available (prerequisites for a QSAR study). We are therefore extending the current concept by implementing a set of virtual filters, which can recognize benign compounds. These filters are based on criteria such as the molecular weight, drug-like properties, and the presence or absence of characteristic structural motifs. After successful completion of a peer testing, it is planned to make the database – along with all supporting software – freely available to universities, hospitals, governmental agencies and regulatory bodies worldwide.

3R relevance

The envisioned Internet laboratory and the already functional virtual test kits can contribute to a significant reduction in animal testing. In drug development, it allows for an early – even before compound synthesis – recognition of potentially harmful substances. By removing those candidate substances from the evaluation pipeline, they will not be forwarded to any *in vivo* toxicity tests. These expectations are supported by the fact that

Table 2. Dimensionality of QSAR approaches.

Dimension	Method	Protein
1D-QSAR	Affinity correlates with pKa, logP, electronic properties, etc.	no
2D-QSAR	Affinity correlates with structural patterns (connectivity, 2D pharmacophore)	no
3D-QSAR	Affinity correlates with the three-dimensional structure of the ligands	possible
4D-QSAR	Ligands are represented as an ensemble of conformers, orientations	typical
5D-QSAR	as 4D-QSAR + representation of different induced-fit models	yes
6D-QSAR	as 5D-QSAR + representation of different solvation scenarios	yes

our virtual experiments have so far not produced any false-positive results. In testing of industrial chemicals for toxicity – for example the 30,000 compounds that have to be retested within the REACH framework – and causing an estimated toll of 10 Million laboratory animals, our approach can be used to safely identify the most harmful compounds *in silico* and prevent their further testing *in vivo*.

Of course, with only a limited number of enzyme/receptor systems known to mediate adverse effects and even fewer accessible in a QSAR context (due to lacking experimental affinity data), false-negative results will always be present. It will selectively recognize potentially hazardous compounds associated with major mechanisms (e.g. metabolic degradation, endocrine disruption) and allow for discarding them early on. Second, a widely used database of this kind might reduce the number of otherwise doubly-conducted toxicity tests at research laboratories focusing on closely related biomedical targets. The main advantage of the proposed virtual laboratory is that it can be applied to hypothetical substances, produces reliable results and is fast and cheap.

References

Lill, M. A., Winiger, F., Vedani, A. and Ernst, B. (2005). Impact of induced fit on the ligand bin-ding to the androgen receptor: A multidimensional QSAR study to predict endocrine-disrupting effects of environmental chemicals. *J. Med. Chem. 48*, 5666-5674.

Lill, M. A., Dobler, M. and Vedani, A. (2006). Prediction of small-molecule binding to Cytochrome P450 3A4: Flexible docking combined with multidimensional QSAR. *ChemMedChem 1*, 73-81.

Lill, M. A., Vedani, A. Computational modeling of receptor-mediated toxicity. In S. Ekins (ed.), *Computational Toxicology: Risk assessment for pharmaceutical and environmental chemicals* (315-351). Hoboken NJ, USA: Wiley & Sons.

Vedani, A., Dobler, M. and Lill, M. A. (2005). Combining protein modeling and 6D-QSAR – Simulating the binding of structurally diverse ligands to the estrogen receptor. *J. Med. Chem. 48,* 3700-3703.

Vedani, A., Dobler, M. and Lill, M. A. (2006). The challenge of predicting drug toxicity in silico. *Pharmacol. Toxicol. 99,* 195–208.

Vedani, A., Zumstein, M., Lill, M. A. and Ernst, B. (2007). Simulating α/β specificity at the thyroid receptor: Consensus scoring in multidimensional QSAR. *ChemMedChem 2,* 78-87.

Correspondence to

Prof. Dr. Angelo Vedani
Biographics Laboratory 3R
Friedensgasse 35
4056 Basel
Switzerland
www.biograf.ch