

In silico platform for xenobiotics ADME-T pharmacological properties modeling and prediction. Part I: beyond the reduction of animal model use

Lucie Huynh¹, Rosalinde Masereeuw², Thomas Friedberg³, Magnus Ingelman-Sundberg⁴ and Philippe Manivet^{5,6,7,8}

There is an urgent need for efficient *in silico* ADME-T prediction tools for the selection of potent therapeutic drugs as well as the elimination of toxic compounds. This is particularly important in view of the high costs and ethical issues inherent to the use of animal models for drugs filtering. To achieve this mission, not only does the accuracy of *in silico* tools need to be improved, but also new experts in the field with skills in theoretical chemistry, clinical and fundamental biology have to be trained. Similarly, clinical biologists committed to the obligation of means and legally responsible for the results they generate could establish a legal framework that defines legal responsibilities when performing *in silico* predictions.

Introduction

Since 1993 the European Union, through the Scientific Advisory Committee of European Centre for the Validation of Alternative Methods (ECVAM), has been promoting the scientific and regulatory acceptance of alternative methods to animal experimentation. They are of importance to biosciences as they can 'reduce, refine or replace the use of laboratory animals'. Recommendations were made about the best ways to use various types of *in vitro* tests for evaluating toxicity of xenobiotics. Besides *in vitro* tests, *in silico* methods have been recently proposed as an alternative. The recently published REACH regulation mentions QSAR (Quantitative Structure–Activity Relationship) techniques for 'predictive toxicology'. The scope of REACH¹ principles application concerns

the chemical industry. It requires the registration, over a period of 11 years, of some 30,000 chemical substances in use today, 'a process which will allow to fill information gaps on the hazards of substances and to identify appropriate risk management measures to ensure their safe use'. In 2007, the REACH principles induced the emergence of a comparable regulation in the USA from the U.S. Environmental Protection Agency named ToxCast.² It concerns 'the safeguard of public health and the environment from harmful effects that may be caused by exposure to pollutants in the air, water, soil, and food'. ToxCast also promotes 'the application of Computational Toxicology to assess the risk chemicals poses to human health and environment'. Besides the chemical industry, the application of REACH principles can be extended to the pharmaceutical industry, which also has an urgent need for new *in silico* 'solutions' to generate new, active and nontoxic

Corresponding author: Manivet, P. (philippe.manivet@lrb.aphp.fr)

¹ Division of Structural Biology, ADME-T Modeling Platform, BioQuanta, Paris 05, France

² Department of Pharmacology and Toxicology, Radboud University Nijmegen Medical Centre, Nijmegen Centre for Molecular Life Sciences, The Netherlands

³ Biomedical Research Centre, Cancer Research UK Molecular Pharmacology Unit Level 5, Ninewells Hospital and Medical School, University of Dundee, Scotland, United Kingdom

⁴ Section of Pharmacogenetics, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

⁵ APHP, Hôpital Lariboisière, Service de Biochimie et de Biologie Moléculaire, Unité de Biologie clinique structurale, Paris 10, France

⁶ Université Paris Descartes, UFR des Sciences Pharmaceutiques et Biologiques, Paris 06, France

⁷ Franco-Thai Molecular Modeling Unit (FTMMU), Department of Microbiology, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand

⁸ INSERM U829, Université d'Evry-Val d'Essone, Evry, France

¹ **REACH** (Registration, Evaluation and Authorisation of Chemicals) system, proposed by the Commission's White Paper on a Future Chemicals Policy (http://ec.europa.eu/enterprise/reach/reach/archives/white_paper/introduction/index_en.htm).

² **ToxCast** (Registration, Evaluation and Authorisation of Chemicals) predicting hazard, characterizing toxicity pathways, and prioritizing the toxicity testing of environmental chemicals (http://www.epa.gov/comptox/toxcast/).

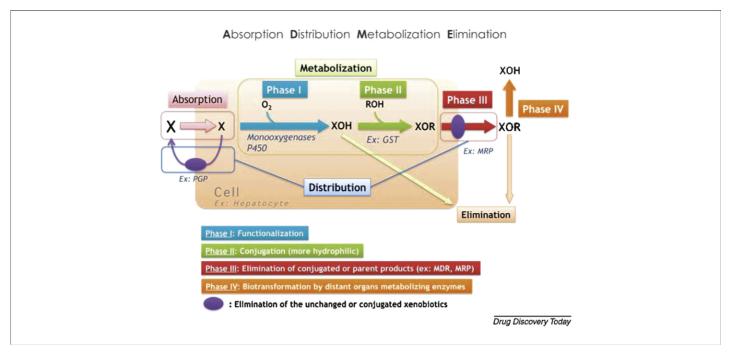


FIGURE 1

Schematic representation of Phases I–IV chemical biotransformation undergone by xenobiotics (X) in a hepatocyte. ADME parameters are shown to be involved in different phases of biotranformation. Immediately after absorption efflux pumps, like P-glycoprotein (ABCB1; PGP), may extrude X to limit the intracellular concentration of X. Intracellularly, X can be converted by Phase I enzymes (Cyp P450), directly eliminated or further conjugated by Phase II enzymes, for example by glutathione S-transferases (GST), and then eliminated by Phase III proteins. The latter step may include diverse transporters, like multidrug resistance proteins (ABCC; MRP) or PGP. Phase IV involves transformation of X in distant organs by other metabolizing enzymes.

compounds, while reducing the high cost of failures in clinical phases. A correct toxicity prediction requires to place xenobiotics in the "whole body context" and to keep in mind that bioactivation may take place. Modeling and prediction of ADME (absorption, distribution, metabolism and elimination) pharmacological parameters (see Fig. 1) for any molecule is an obvious necessity. In addition, the toxicity of a drug, or its metabolites, may modify one of the ADME parameters.

The use of biokinetics (or physiologically based pharmacokinetics - PBPK) modeling software is increasingly being adopted, particularly in the pharmaceutical industry as part of highthroughput screening strategies. These methods rely on a set of data collected in vitro and in vivo for a drug library that are modeled by simple mathematical equations (pharmacokinetic). The information obtained is really useful, but only partial, because it is based on biomarkers and drug partitioning in whole animals and, hence, is missing several dimensions relating to the chemical structure of the drugs and the macromolecules responsible for their biotransformation.

Questions that still remain: can we correlate an acute or chronic toxicity for a specific tissue to the three-dimensional structure of a compound if the statistical analysis was based on one or at most two-dimensional descriptors and general organ biological parameters? How should lead optimization be positioned with respect to ADME-T predictions?

The answer to the first question is clearly negative. A pertinent answer cannot be obtained by using methods currently available for ADME parameters and toxicity properties prediction. These methods are classified into two groups: those based on analyzing molecule similarities (pharmacophores and molecular fragments

based methods), and those based on the analysis of calculated physicochemical properties ($\log P$, $\log D$ and so on) (see [1] for review). Both methods suffer from a lack of accuracy and are sometimes based on old concepts and enduring false dogma. A serious overhaul is, therefore, needed before the application to ADME-T predictions.

Surprisingly, no answer has been proposed for the second question since, to date, it has not even been remotely addressed. A drug is predicted to be toxic and so what? Do we throw it away in the trash bin? If the origin of its toxicity is known exactly, could the drug structure not be amended to render it nontoxic while keeping its therapeutic activity? This is the essence of 'lead optimization'. The term ADME-T-LO (LO for lead optimization) would be more appropriate than ADME-T.

Under the principles of REACH, it is anticipated that QSAR will be used more extensively for time- and cost-effectiveness, as well as for animal welfare. Although the chemicals policy of the EU raises a clear regulatory need for scientifically valid QSAR models available to all stakeholders, it does not address the current issues about validity and applicability of QSAR. QSAR is a general term giving only a blurred vision of the realistic possibilities offered by in silico methods. Hence, it is necessary to develop a framework for the validation of QSAR. An adequate proposition to this issue would be to provide a platform combining various technologies and expertises in the field of computational chemistry (physics, quantum chemistry, force fields customization, molecular dynamics, new concept of pharmacophores-based screening algorithms, receptor activated state modeling, ND-QSAR and so on), cell biology, biochemistry and molecular biology, together with active participation of healthcare professionals. All such experts would work to guide molecular modelers on how accurately to model 'the body' through the summation of three-dimensional models and the pharmacophores of drugs and protein targets involved in ADME parameters. This multidisciplinary platform would help 'refresh' old concepts used in conventional molecular modeling, data mining and QSAR studies to model biological processes and ADMET properties, like octanol–water partition coefficient, similarity measures based on molecular 2D-fragments or crude distance matrix-based pharmacophore definition.

REACH or not REACH: that is not the question

We already mentioned the obvious necessity to design realistic and pragmatic in silico tools combined with in vitro experiments within a multi-technology platform, especially in view of a possible reduction of animals used for toxicity assays. Pressure on regulatory authorities from politicians, public opinion, media and the scientific community underpins REACH principles, which were designed for, and are directly applicable to, all industries using chemicals. Everyone agrees that the reduction of animal use is an important matter and the main argument addressed by REACH. Another one is equally, if not more, important: human health. Chemicals are often found in the 'customers' plate'. Drugs and chemicals can have side effects on human beings, which nowadays is difficult to ignore. Thus, it is obvious and logical for any chemical to be subjected to toxicological investigations and controlled in ways similar to drug compounds in the pharmaceutical industry. Xenobiotics can be released in the environment, taken up by animals and finally by humans, eventually metabolized, resulting in active and/or toxic metabolites. Hence media and authorities often report about dangerous chemical components in various consumer goods, such as toys, wallpaper or many housewares. Checks exerted by authorities on chemicals should force the chemical industry to move toward 'clean chemistry'.

Another interesting impact of REACH and in silico methods concerns drug development. Toxicity prediction made for, or by, the chemical industry can be taken up by the pharmaceutical or cosmetic industries. Indeed, some of the chemicals used are carcinogenic or cytotoxic and therefore generate human diseases. Yet, those poisons can be a source of inspiration for drug designers. For example, some chemicals can be used again after modifications for pathology treatments: some pesticides on the market were found to be potential anticancer drugs [2]. Though such a strategy may not appear very popular, it could help both the pesticide industry to become cleaner and the pharmaceutical industry to increase the diversity of its lead compounds' pipeline. Further, the pharmaceutical industry could extend the application of in silico methods to the numerous libraries that were rejected to the 'trash bin' for toxicity reasons and to save a good portion of the multibillion dollar expenses on drugs for which toxicity was revealed during clinical phases. Even if there is still a debate on the pertinence of REACH principles and their concrete applicability, their impact goes beyond the simple regulatory anecdote toward a public health issue. This revolution has already started. In silico is one of the alternative technologies promoted by ECVAM. In silico methods and technologies have here the opportunity to win their spurs; however, it is not so simple. There is a lack of complete in silico solutions adapted to the requirements for the safety evaluation of chemicals and drugs. Even the pharmaceutical industry cannot provide any real solution to the current need. Three main issues need to be addressed: the standardization of *in silico* approaches for modeling ADME-T properties of any compound; the choice of *in vitro* models and experiments designed to achieve the validation of the predictability power of *in silico* tools, and, finally, what position do *in silico* specialists want to adopt with respect to their responsibility on the results they provide from their *in silico* tools? This latter point is most crucial because it influences the two first parameters and opens the debate. As long as *in silico* sciences are viewed secondary to biological results as the main thrust supporting decisions and journal publications, wide acceptance as a really valuable tool will be difficult.

Contract Research Organization investigators provide controlled and standardized clinical results to their clients, the pharmaceutical industry. In a similar fashion, diagnostics equipment and kit manufacturers, committed to product liability, have set up relevant quality control procedures run by adequately trained personnel, and their supplies are also being checked by healthcare professionals. Clinical biologists (physicians) or healthcare professionals also pledge their civil and criminal responsibilities in the course of their daily professional activity. No error is acceptable when diagnosing an inherited genetic disease or when a prognosticator, vital to a patient is threatened. Similarly, in silico specialists should take on risks and responsibilities, should they wish to be considered as evaluators. For this purpose, in silico technology should be subject to the same evaluation as clinical biology: standardization, reproducibility, precision, accuracy, detection limit, and so on. In our proposed multidisciplinary platform, in silico investigations are performed by the people at the interface of biology and chemistry (see the accompanying article: 'In silico platform for xenobiotics ADME-T pharmacological properties modeling and prediction. Part II: the body in a Hilbertian space'). It is thus easy for the clinical biologist with skills in quantum chemistry and drug design to transfer standardization technology from hospital to in silico tools, raising their level of accuracy. Structural biology and clinical biology are merged together; Structural Clinical Biology is born.

State of the art of ADME-T in silico

Modeling the ADME-T pharmacological parameters of a drug is a complex task. Difficulty emanates from the large number of enzymes, transporters and regulatory mechanisms involved in biotransformation, toxicity and inter-relationships existing between each ADME-T property. Phase I metabolizing enzymes are the most frequently studied by molecular modeling and to a lesser extent Phase II conjugating ones. However, Phase 0 and Phase III transporters involved in the modulation of the cellular entry and exit, respectively, of the parent or metabolized compounds are always neglected [3,4]. Besides the toxicity of a parent compound, in some cases, toxicity of a drug can be influenced by a transporter by concentrating, for instance, a drug or one of its metabolites in tissues.

Metabolism of a drug may take place virtually immediately in intestinal cells or during its first passage through the liver. Drug metabolites can be reabsorbed after their elimination in the bile and can become toxic by intra-organ accumulation. ADME-T is a schematic and convenient concept, but, in fact a too rapid shortcut for classifying a large number of integrated processes concerning many

tissues where drugs are partitioned and undergo chemical transformations; where each metabolite, in turn, may act in a distant organ after diffusion from the previous one. This concept finally increases the difficulty of interpreting and predicting drug biotransformation and toxicity rather than simplifying it. Two main strategies are currently used for modeling virtually each ADME parameter and predicting toxicity. The first one is based on PBPK data obtained from in vivo or in vitro experiments in animals. These make use of the simulations of the actual measurements of ADME-Tox properties in animals. Software currently used for biokinetics, include WinNonlin [5,6], GastroPlus [7], COSMOfrag [8] and OraSpotter [9]. Organs or groups of organs are considered as discrete interconnected compartments with physiological volumes and blood flows. These models account for physiological influences and are used to estimate internal tissue concentration, allowing extrapolation between species, doses and routes. In particular, physiological parameters include species dependence of tissue volumes and blood flows. Chemical-specific parameters include intrinsic clearance, enzyme affinity and rate of reaction, plasma membrane permeability, renal and biliary excretion clearances, plasma protein binding, tissue-toplasma partition coefficients and absorption parameters. Data on these metabolic parameters are obtained from in vitro studies (e.g. using hepatocytes and/or subcellular fractions like microsomes). Though widely used, these software applications are limited by the nature and physical dimensions of dependent variables (biological descriptors) upon which they are based. Finally, they give only general information on the biological behavior of an organ or a tissue. They are also limited by the mathematical formalism of PBPK equations that ignore structural and physical properties of drug compounds. In spite of a good predictability obtained with biokinetics algorithms, they fail correctly to classify a large number of drugs, cannot predict their ADME-T properties and cannot direct lead optimization.

The second in silico strategy for studying ADMET properties relies on molecular modeling and quantum chemistry techniques. This is a great progress toward prediction of the ADME-T properties of drugs, because those techniques have an edge on PBPK software: they allow the simulation of interactions between drugs and macromolecules acting along the ADME-T pathway at the atomic or molecular level. They are, however, limited by a lack of standardization and reproducibility and require further improvement before they can be considered acceptable for the prediction of ADME-T properties [10]. Hence, the number of scientific publications dedicated, for instance, to molecular dynamics (MD) study of cytochrome (CYP) inhibitor and substrate interactions or the study of interactions between a drug and the heme/iron ion complex with quantum chemistry, still remains modest although growing. In the case of MD simulation of drug/cytochrome interactions, force fields used for simulations, like those proposed by CHARMM [11] or AMBER [12], lack a correct mathematical term for modeling metal/ligand interactions (soft anharmonic potential for modeling iron coordination, induced polarization and anisotropic hydrogen bond terms, explicit and free water internal coordinates) needed to simulate the heme/iron complex, whether or not embedded in water solvent. Those simulations only permit the production of crude results, potentially leading to false interpretation of drug transformation mechanisms. No study using MD techniques is dedicated to the clear understanding of the allosteric activation

of CYP, a prerequisite for the simulation of drug/CYP complexes and of putative reaction mechanisms involved in substrate transformation. All CYP/drug studies concern CYP/inhibitors (see [13] as an example) and very rarely CYP/substrate or inductor. It is well-known in structural pharmacology that modeling interactions between a targeted protein (receptor, transporter and enzyme) and an antagonist or an inhibitor is easier than modeling its interactions with an agonist or an allosteric change inducer. MD techniques are rarely well-documented with any information relating to the used force field and homology modeling methods of CYP are crudely performed. Furthermore, many improvements are needed in quantum mechanics (QM). For instance, most investigations performed on porphyrin-type structures as models for heme in cytochromes use very low levels of calculations with nonaccurate basis sets, which are recommended for describing very weak interactions, like charge transfers. Those calculations are usually performed on the whole tetrapyrolic molecule. Using an accurate level of calculations and extended quantum basis sets is hindered by the size of the system. Moreover, water solvent is always neglected in all calculations. Finally, QSAR, MD and QM are always focused on a single CYP or transporter involved in ADME-T.

Molecular fragments-based in silico methods [14] lead to unsatisfactory drug toxicity predictions. They show poor accuracy, because they are based on 2D molecular fragments that are not representative of the whole molecule's physical and electronic properties. Hence, toxicity data generally concerns one tissue. Moreover, no information is available on the enzyme, transporter or receptor responsible for the toxic mechanism in the tissue concerned. No structural data can thus be extracted from those studies, and multidimensional mathematical space is reduced to 2D or 1D with substantial losses of information.

Yet, the physical background of those techniques is excellent in that they have already proven their accuracy in other topics. The level of theory has, however, to be correctly adapted to the needs of pharmaceutical industry and the regulatory requirements of the European Union for ADME-T prediction. These methods need to be integrated inside a multidisciplinary platform including in silico and experimental assays. The results obtained have to be interpreted in the context of the whole process of drug biotransformation after mathematical integration. Hence, in silico methods must be standardized like in vitro biological experiments, their accuracy and reproducibility being validated by a clear protocol.

Understanding cell physiology transposed to simple in vitro models: the grail of ADME-T in silico prediction

In silico ADME-T predictive models exist thanks to biological or physicochemical data. In silico models must be validated by in vitro experimental assays. The availability of exploitable and reliable experimental data is a real bottleneck for building efficient in silico ADME-T predictive models. Many criticisms focus on transfected cells arguing that in vitro cell lines do not reflect the physiological situation. It is well-known that transfected cells overexpressing a protein of interest, like a receptor or a transporter, are useful for obtaining molecular information about the interaction of substrates with this protein of interest. Hence, it is an artificial system that requires information of endogenous expression levels for accurate in silico predictions. This is particularly true for transmembrane transporters using ionic gradients which are sensitive to the regulation

exerted by cell signaling cross-talks or trans-activations. Note that phosphorylation is only one way of regulation, proteins can also be glycosylated, inserted or subtracted from the membrane by various mechanisms. Therefore, caution must be taken when choosing cell type. For example, in a serotonin (5-HT) uptake study [15], we successfully used blood platelets naturally expressing the 5-HT transporter. Molecular modelers do not, however, intend to reexamine the cell signaling already described by others or to bypass in vivo assays. They use the cell as a tool for validating their in silico models. In this regard, they must have a complete understanding of the cell models. Experimental validation of in silico 3D or QSAR models is in fact a two-pronged strategy: (i) for instance, accuracy of the 3D-homology model or a crystal structure of a transmembrane receptor has to be checked for the rational design of ligand affinity prediction. Important residues for the ligand binding selected by molecular modeling are then mutated in the receptor and mutants are transfected in a cell line. Then radio-ligand binding assays (or uptake measures for any transporter) are performed and ligand affinity is measured in transfected cells. Finally, the 3D model is refined or rejected. Transfection is achieved in cells that do not naturally express the wild-type protein and do not contain additional coupling with the cytosolic proteins altering the natural function of the protein of interest. (ii) Once the 3D model is validated, predictability accuracy is tested in vivo with a new set of molecules. Predicted values of in vivo effects are finally confronted with in vivo experiments in animal models.

Conclusion

In the present review, we showed that the recent emergence of REACH principles gives the opportunity to *in silico* methods to win their spurs. It is clear that applicability of *in silico* methods goes beyond the chemical industry and can be extended to pharmaceutical, cosmetics and food industries. Although pharmaceutical industry claims to have already used efficient ADME-T prediction tools through 'scientists who possess both biological intuition and computational skills' [16], current tools need to be improved

because failures in clinical phases are frequently met because of unpredicted late toxicity of leads, leading to billions of dollars wasted in abandoned drug development programs. Existing solutions are thus inadequate. A large gap exists between current in silico techniques and real adapted methods directly applicable for ADME-T prediction. Weaknesses relate to scientific background and in silico results interpretation. A revolution is needed in researchers mentalities to wash out old obsolete surviving dogma in physicochemistry and drug design that prevent rapid evolution of modeling techniques. Hence, scientists who use and develop in silico ADME-T modeling tools must have more than a biological intuition of human body organ functions. They need to collaborate and share knowledge with experts in fundamental biology and specialists of protein families involved in xenobiotics biotransformation in organs where metabolism and transport are intense. In addition, they also have to practice clinical biology 'close to the patient' to grasp the essence of clinical biology and interpret multidimensional information of xenobiotics biotransformation in the patient's body context. Experience grasped on cells, tissues or small animals are good, but ADME-T concerns human beings. In this way, it seems logical to get closer to the patient for ADME-T interpretation and to guide development of accurate and adapted in silico tools. People with skills in both theoretical chemistry and clinical biology are rare. A strong partnership between pharmaceutical and chemical industries, biotechnology companies, universities, hospitals and regulatory bodies will be necessary for training the new type of professionals in our schools, and allowing them to work in the framework of 'clinical structural biology'. All partners will benefit from this new born 'translational science'. Another important question we address in our review is the legal responsibility that molecular modelers will have to accept to assume if they want their methods to be considered as equivalent to in vitro ones. Official recognition of the in silico tools value is a great chance. However, this means a higher level of constraints in evaluating quality of results and a complete acceptance from molecular modelers in assuming the legal consequences.

References

- 1 Tetko, I.V. *et al.* (2006) Can we estimate the accuracy of ADME-Tox predictions? *Drug Discov. Today* 11, 700–707
- 2 Clément, M-.J. et al. (2008) Benomyl and colchicine synergistically inhibit cell proliferation and mitosis: evidence of distinct binding sites for these agents in tubulin. Biochemistry 47, 13016–13025
- 3 Giacomini, K.M. and Sugiyama, Y. (2005) Membrane transporters and drug response. In Brunton L. (11th edn) (McGraw-Hill Professional New York, eds), pp 41–70
- 4 Szakács, G. et al. (2008) The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). Drug Discov. Today 13, 379–393
- 5 Metzler, C.M. (1989) PCNONLIN Version 3.0. Software for the statistical analysis of nonlinear models on micros. SCI Software, Lexington, KY, USA
- 6 Campos Moreno, E. et al. (2007) Population modelling to describe pharmacokinetics of amiodarone in rats: relevance of plasma protein and tissue depot binding. Eur. J. Pharm. Sci. 30, 190–197
- 7 De Buck, S.S. et al. (2007) The prediction of drug metabolism, tissue distribution, and bioavailability of 50 structurally diverse compounds in rat using mechanism-based absorption, distribution, and metabolism prediction tools. *Drug Metab. Dispos.* 35, 649–659
- 8 Honiq, M. and Klamt, A. (2005) COSMOfrag: a novel tool for high-throughput ADME property prediction and similarity screening based on quantum chemistry. *Chem. Inf. Model.* 45, 1169–1177

- 9 Svava Ósk, J. et al. (2005) Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. Bioinformatics 21, 2145–2160
- 10 Boobis, A. et al. (2002) In silico prediction of ADME and pharmacokinetics: report of an expert meeting organised by COST B 15. J. Pharm. Sci. 17, 183–193
- 11 Brooks, B.R. et al. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4, 187–212
- 12 Weiner, S.J. *et al.* (1986) An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* 7, 230–252
- 13 Afzelius, L. et al. (2001) Competitive CYP2C9 inhibitors: enzyme inhibition studies. Protein homology modelling and three-dimensional quantitative structure–activity relationship analysis. Mol. Pharmacol. 59, 909–919
- 14 Hulzebos, E. et al. (2005) Validity and validation of expert (Q)SAR systems. SAR QSAR Environ. Res. 16, 385–401
- 15 Pratuangdejkul, J. et al. (2005) Definition of an uptake pharmacophore of the serotonin transporter through 3D-QSAR analysis. Curr. Med. Chem. 12, 2393–2410
- 16 Kumar, N. et al. (2006) Applying computational modelling to drug discovery and development. Drug Discov. Today 11, 806–811