ELSEVIER

In silico platform for xenobiotics ADME-T pharmacological properties modeling and prediction. Part II: the body in a Hilbertian space

Alexandre Jacob[1], Jaturong Pratuangdejkul[2,3,4], Sébastien Buffet[1], Jean-Marie Launay[2,3,4] and Philippe Manivet[2,3,4,5]

[1] Division of Structural Biology, BioQuanta, Paris 05, France
[2] Franco-Thai Molecular Modeling Unit (FTMMU), Department of Microbiology, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand
[3] APHP, Hôpital Lariboisière, Service de Biochimie et de Biologie Moléculaire, Unité de Biologie clinique structurale, Paris 10, France
[4] Université Paris Descartes, UFR des Sciences Pharmaceutiques et Biologiques, Paris 06, France
[5] INSERM U829, Université d'Evry-Val d'Essone, Evry, France

We have broken old surviving dogmas and concepts used in computational chemistry and created an efficient *in silico* ADME-T pharmacological properties modeling and prediction toolbox for any xenobiotic. With the help of an innovative and pragmatic approach combining various *in silico* techniques, like molecular modeling, quantum chemistry and in-house developed algorithms, the interactions between drugs and those enzymes, transporters and receptors involved in their biotransformation can be studied. ADME-T pharmacological parameters can then be predicted after *in vitro* and *in vivo* validations of *in silico* models.

Introduction

It is widely accepted that the accuracy of tools used in computational chemistry and biology for the prediction of ADME-T properties for any compound (xenobiotics) needs to be improved before they are used for modeling drugs' toxicity and biotransformation. Unfortunately, no efficient solution has yet been proposed. Some attempts have been made to identify weaknesses in computational methods. For example, poor chemical structural diversity or differences in the chemical space observed between external and training molecule sets used in QSAR studies have been incriminated in the lack of predictability of *in silico* ADME-T methods [1]. Also, people claim, in some publications, to have developed 'high tech' multidisciplinary platforms combining biological and informatics groups [2] with the intention of predicting ADME-T properties for any chemical compounds with the help of 'system biology' algorithms [3]. In fact, obtaining efficient and accurate *in silico* ADME-T modeling and predictive tools is not so simple. It cannot be solved by simply adjusting the chemical space of molecules of the training and external test sets, as carried out by computational chemists with only a basic knowledge in biology, using inappropriate 'magic algorithms' that unduly increase the complexity of choosing the ideal *in vitro* experimental model. Existing solutions succeeded neither in reducing drugs' clinical phase failures owing to toxicity discovered late in the process, nor in increasing optimization rate of pharmaceutical industry's focused libraries by monitoring drugs' ADME parameters. A real solution requires rethinking thoroughly the concept of ADME-T *in silico*, eliminating old persisting dogmas in computational chemistry, once useful but that now prevent real progress in the field. We have shown in this issue that a 'clear estimation of the accuracy of *in silico* methods' [1] is not sufficient for their application within the REACH regulations (*In silico* platform for xenobiotics ADME-T pharmacological properties modeling and prediction: Part I. Beyond the reduction of animal model use). Indeed, it is quite important to train qualified professionals with both expertise in computational chemistry and biology (structural clinical biology) who will create *in silico* tools adapted to ADME-T prediction. We also showed the obvious necessity of acceptance by the molecular modelers of professional and legal responsibility on the results they provide if they want to be considered as evaluators. In this review, we will detail the architecture of a multidisciplinary *in silico* platform for ADME-T modeling and prediction, using standardized methods, where total quality control management is an important part of the whole process ensuring the accuracy of *in*

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

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2009.01.013

silico methods being used. Control of information flows between biologists and molecular modelers is carried out with the help of a mathematic algorithm establishing *in fine* the road map of biotransformation processes of any drug in the whole body context.

'Diversity of compounds', 'chemical space' and 'drug-likeness': beautiful metaphors

'Chemical diversity' is a term widely used by drug designers and chemists. They all agree that 'the wider the chemical diversity of an initial compounds' library, the better the results obtained for the definition of a QSAR model for a typical protein target'. This is, in fact, not really true and certainly not so simple. It seems that the total number of chemically diverse or various small molecules within our own bodies may not exceed more than a few thousands [4,5]. It is thus useless to extend, indefinitely, the size of an initial chemical compound QSAR library. The key point lies more in the design of a clever library adapted to the protein of interest with a sufficient and adequate degree of diversity. Designing such a library requires a solid training, not only in chemistry, but also in biology and pharmacology, as well as knowledge of the functionality and regulation mechanisms of the protein of interest. An illustration of this concept is a QSAR study performed on the serotonin transporter (SERT) [6]. The compound training set library was built around the indole scaffold of serotonin. The degree of chemical diversity was small and compounds in the library looked similar. The selectivity degree of SERT is, however, so high that it is very sensitive to serotonin's conformation. The 5-HT *anti*-conformer is transported by SERT, whereas other local energy minima of serotonin are poorly transported (this result was confirmed later by others [7]). Simply changing the direction of one hydrogen bond from one compound to another radically modifies its uptake value. This study illustrates the gap between widely accepted concepts of chemical space and chemical diversity and a real and pragmatic QSAR approach, based on a strong knowledge of the biological system. It means that using a too highly diverse library of compounds for QSAR studies on SERT would potentially result in missing important information about those properties needed for a compound to be transported by SERT.

Another popular term used in the literature is 'drug-likeness'. In QSAR studies, authors usually recommend using only drug-like compounds for training and external test sets of compound libraries [8]. This term intrinsically has no meaning; it depends on a large number of parameters. General drug-likeness predictions are usually based on 1D or 2D structure–activity descriptors, such as molecular weight, number of donors or acceptors hydrogen bonds, number of aromatic rings, and connectivity indices. Such descriptors are too simplistic for a good appreciation of therapeutic potency. In addition, drug-likeness is only meaningful in the context of biological signal transduction pathways. A compound becomes a drug if it shows good affinity and selectivity when bound to a protein involved in a pathological signal transduction pathway. Not to mention the extreme difficulty in designing G-protein-coupled receptor (GPCR) agonists or inverse agonists, because they bind to a high-affinity-binding site revealed after conformational change of the GPCR. It is hard to see how simple 2D fragment-based drug-likeness prediction methods can provide ligand properties that would predict appropriate access to the binding site. By combining experimental assays and molecular modeling it can be demonstrated that all residues of the same binding site have different allosteric roles [9]. The crude approach of drug-likeness filtering is, thus, detrimental. It arbitrarily eliminates the possibility of selecting interesting chemical compounds with a potential therapeutic effect after optimization. Molecular modelers should not be scared by using random chemical libraries that mix known potential therapeutic compounds, extracted from pharmaceutical industry-focused libraries, with nontherapeutic compounds. The key is to build an accurate *in silico* model with a sufficiently diverse chemical library for adequate predictive power. Selection of drug-like compounds will be achieved later on.

Old concepts for poor ADME-T predictability

Some old medicinal chemistry dogmas still survive in spite of permanent scientific progress. In their time these concepts were established by pioneers in their field. They were very useful for the scientific community and allowed important advances in drug discovery. Of these the 'rule of five' [10] or the 'octanol/water partition coefficient', one of the 'rule of five' parameters [11], remains widely used in drug design. It is now clear that compounds with a molecular weight above 500 Da can be good drugs. Biological systems use such types of compounds. Good examples include antibiotics or neuronal peptides. One-dimensional descriptors in the 'rule of five', such as number of hydrogen bond centers, cannot be used to predict a drug's therapeutic potential. It actually depends on the target's molecular context. In addition, drugs can also exert their therapeutic activity exclusively through hydrophobic contacts.

The partition coefficient, called log *P*, is determined by the logarithm of concentration ratio of a nonionized solute in two solvents. The main problem with this coefficient is that it is restricted to nonionized compounds. Surprisingly, this coefficient is still used in QSAR studies for the prediction, for instance, of skin toxicity of compound libraries [12]. Toxicity cannot, however, be directly correlated to log *P*. Correct *in silico* prediction of the 'absorption' parameter of ADME-T using log *P* is, therefore, prone to error, because 'absorption' is a complex mechanism involving solubility and permeability (bioavailability), that includes (for solubility) a transition of phase of a drug from a solid state to a soluble state. In spite of all these weaknesses, publications of algorithms using log *P* or other one-dimensional empirical coefficients (log *D* for distribution or log *S* for solubility, for instance) are still serenely showing high predictability and excellent correlation coefficients between experimental and predicted values that approach, or even achieve, unity with sophisticated statistical algorithms for data treatment [13]. This is evidence for a lack of deep understanding of physical phenomena and an insufficient capacity to differentiate between actual causality and coincidence. This false perception of the relationship between log *P* and drug toxicity, blood–brain barrier transfer capability or any other biological property can, for instance, be compared to the nature of the link existing between diabetes and hyperglycemia. This analogy would unduly allow a clinician to classify diabetes type 1 or 2 based purely on glycemia. It is a fact that there is a direct correlation between glycemia and diabetes. In the clinical setting, glycemia never helps differentiate those patients suffering from type 1 or 2 diabetes. Only, for instance, the level of insulinemia, the study of

FIGURE 1

Allegoric representation of the false perception of the physical world by a molecular modeler unable to correctly use theoretical logical reasoning to adapt modeling to practical applications.

HLA group markers or the detection of auto-antibodies in the patients' plasma, are able so to do. Figure 1 illustrates allegorically the false perception of the physical world for a molecular modeler lacking one sense (the means to use logic appropriately) leading to the loss of many dimensions (inability of theoretical logical reasoning to adapt to practical applications [14]). He has still two available hands that can help him to pull up his stringcourse, but he prefers to stay blind, somewhat satisfied with his comfortable scientific logic giving him a truncated perception of the world.

Another widely admitted idea is to use 2D descriptors for genotoxicity, ADME properties or any biological activity predictions [15]. This approach is very restrictive because, although two molecules can contain one or several common fragments in their chemical structure, they may not show the same activity and/or the same toxicity mechanism. A good example is promethazin (targeting histamine H1 receptors) and imipramine (targeting presynaptic serotonin and norepinephrine transporters). Both molecules have two similar fragments representing 60% of their structure. Their respective targets and activities are, however, completely different. The physical properties of a compound are not merely the addition of the properties of those fragments constituting the whole molecule. Similarity of activity or toxicity is an *N*-dimensional issue, where *N* is at least 3 and not a 2D one.

A Hilbertian view of the *in silico* ADME-T properties prediction

Conventional *in silico* techniques for the prediction of the ADME-T properties of drugs are too simplistic and inadequate to predict the behavior of a drug in a whole organism. Thus, it is not a solution compliant to regulatory requirements. Modeling the whole ADME-T process requires more audacity. It demands a clear understanding of the physiology of those organs (brain, liver, kidney, lung, intestines and skin) involved in drugs' pharmacology and which molecular entities (receptors, transporters and enzymes) participate in pharmacokinetics.

We have worked on a newly developed *in silico* ADME-T toolset. It is adapted to simulating biotransformation of any xenobiotic in the whole body. It contains: (i) a large panel of 3D models of 'enzymes, receptors and transporters' involved in drugs' pharmacokinetics [16–21]. These organs are major players in the three main biotransformation phases: phase I, phase II and phase III (Fig. 2). Hence, genetic and epigenetic sources of variation of those proteins activity are modeled. Xenobiotics can affect the activity of cytochromes and transporters through a direct interaction, or through modification of their expression level. In this latter category we find transcriptional factors and nuclear receptors. (ii) We also used a large library of pharmacophores specific for each enzyme, receptor and transporter involved in the biotransformation of xenobiotics. Modification of the pharmacophore definition [22] allowed the development of an algorithm that permitted high precision screening of any molecule library that allows an *N*-dimensional similarity measure. This algorithm quantitatively ranks molecules according to their biological activity. It replaces

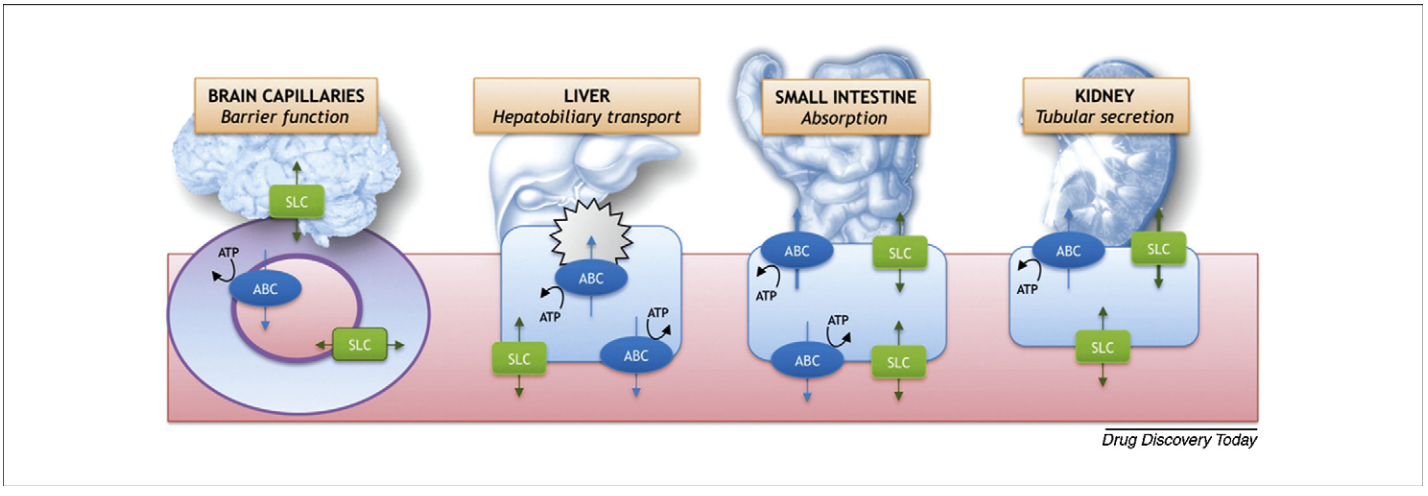

FIGURE 2

Four main organs involved in an intense xenobiotic biotransformation. SLC and ABC are abbreviations for those organic anion transporters belonging to the solute carrier (SLC) and ATP-binding cassette (ABC) efflux pumps families, respectively.

the classical, linear QSAR equation obtained by multiple linear regression, which truncates information by compressing 3D descriptors in 1 or 2D ones or Bayesian probabilistic approaches that cannot solve the problem of irrelevant and weak redundant features in the data sets; (iii) *ab initio* and quantum chemistry calculations are applied to decipher the mechanisms driving drug transformations by cytochromes and toxic reactional mechanisms involving electron transfer and allow the evaluation and prediction of the pK_a solubility [23] of any nonionizable or ionizable compound under different crystalline polymorph and salt states, by using phase transition Born–Haber thermodynamics.

Combining 3D models, *N*-dimensional pharmacophores, quantum chemistry and molecular dynamics on explicit physiological membrane models allows the prediction of ADME-T parameters with high degrees of accuracy.

In this strategy, the body is considered as an ensemble of organs composed of subensembles of molecular entities (Fig. 3). Hilbertian spaces formalism used in quantum mechanics was chosen as an elegant vectorial formulation of ADME-T properties for the mathematical integration of data flow. *In silico* 'measurable quantities' (drug interaction with a receptor and enzyme, drug uptake through a transporter and so on) defined by *N*-dimensional QSAR models represented by pharmacophores are linear operators. Integration of all measures for one drug was obtained by orthogonal projections. When a full analysis has been completed (with 3D models, ND-QSAR equations, explicit molecular dynamics simulations and so on) for a drug compound, then *in vivo* endpoints, for example pharmacokinetics parameters such as volume of distribution, and renal/hepatic clearance, can be obtained from mathematical integration.

Special attention has been focused on toxicity prediction. Toxicity cannot be separated from ADME because transformed drugs can induce side effects and tissue damage in the patient. Two important mechanisms contribute to toxicity: bioactivation through metabolism and accumulation through transport modulation. Thus, to predict the toxicity of drugs, one has to first predict accurately their ADME parameters. Biotransformation of xenobiotics may result in formation of highly reactive molecules that are known to play a role in damage to the constituents of cells, for example the induction of mutagenesis and carcinogenesis (through bioactivation). The mathematical algorithm integrates that a metabolite can interact with enzymes involved in the biosynthesis of endogenous compounds, with receptors, or with signal transduction mechanisms by connecting sub-Hilbertian spaces (Fig. 3). During phase I, reactive and potentially toxic intermediates can be generated. Phase II metabolism can produce highly reactive electrophilic intermediates that covalently bind to macromolecules. Phase III transporters play crucial roles in the cellular toxicity of these agents. Transporters in liver, kidney and the blood–brain barrier affect the exposure of organs to the drugs that they filter. These transporter-mediated adverse drug responses can generally be classified into three categories: increase of drug concentrations in plasma owing to a decrease in uptake and/or secretion in 'clearance organs', such as liver and kidney; increase of concentrations in targeted organs due either to enhanced uptake or reduced efflux of the drug; increased concentration of an endogenous compound in plasma owing to a drug uptake inhibition in an organ supposed to eliminate it.

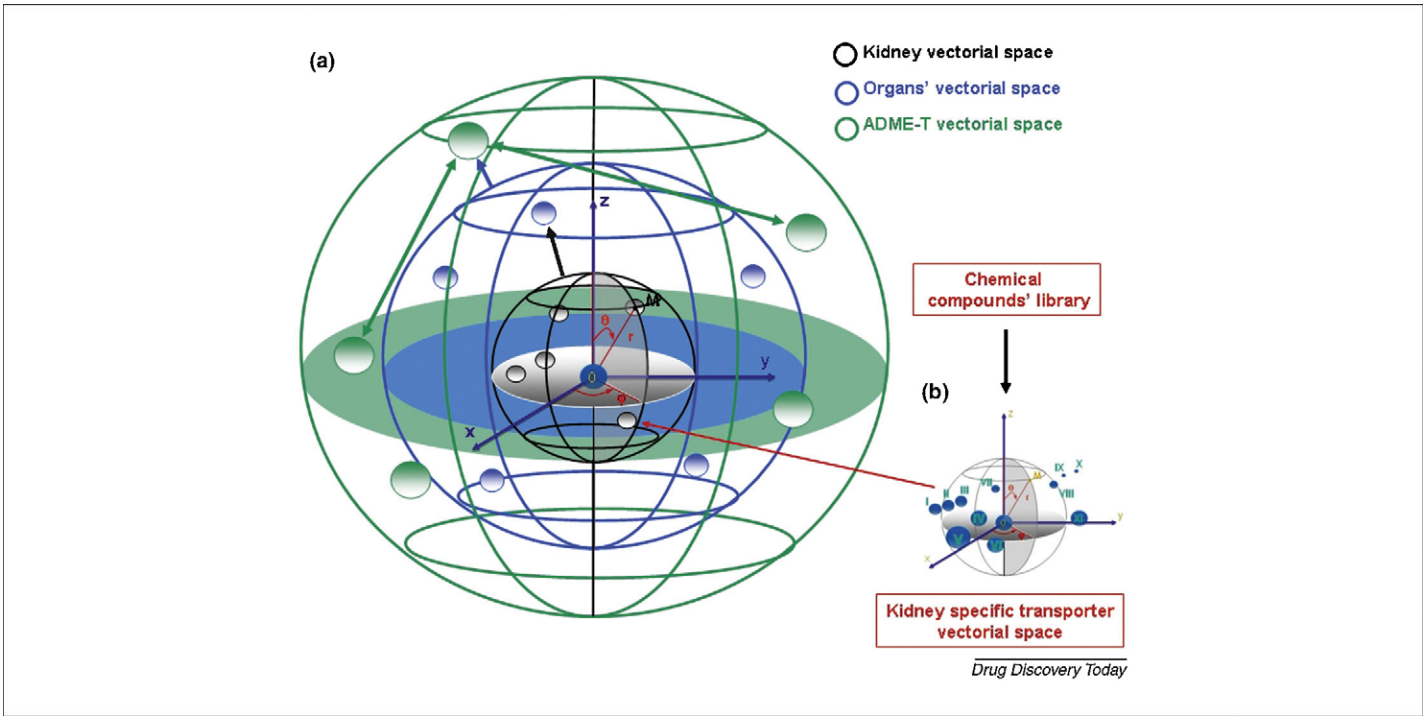

FIGURE 3

(a) Simplified schematic representation of the statistical integration of the whole *in silico* ADME-T parameters for xenobiotics or drugs in a multi-Hilbertian space. Subensembles are hierarchically ranked from **(b)** the elementary structural unit, that is a pharmacophore for a typical protein (here example of a transporter) influencing ADME-T parameters of kidney (black sphere), through the implication of a group of transporters in several different organs (blue sphere), until interconnections between ADME-T parameters. Spherical coordinates definition of pharmacophores is defined in Ref. [21].

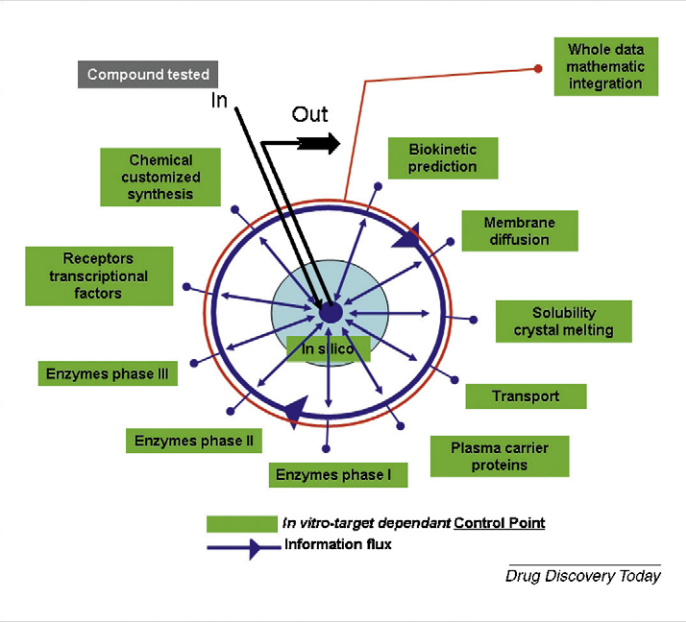

FIGURE 4

'Cell cycle' representation of the workflow for ADME-T parameters prediction of xenobiotic. Interaction of each xenobiotic is confronted with experimentally validated 3D models and pharmacophores of proteins involved in different phases of xenobiotics biotransformation.

The number of proteins acting as drivers of xenobiotic biotransformation that need to be modeled is quite large. The *in silico* toolset thus gradually and continuously is enriched by new pharmacophores and 3D models. Its accuracy increases exponentially with the number of evaluated chemical compounds. The dynamic process of this machinery is represented in Figs 4,5. *In silico* filtering is repeated for all those newly included compounds. The algorithm records all organs and proteins that are involved in biotransformation and establishes the cartography of the route followed by a drug inside the whole body. Mathematical integration is thus the key for successful *in silico* prediction. The same methodology can, for example, be applied to the generation of a 3D model of a new transporter, allowing it to be added to the toolset. In this case, *in vitro* experiments on cultures of cells, tissues or organs containing appropriate transporters, receptors and so on are placed in the 'cell cycle' represented in Fig. 4. A similar process can then be performed *in silico*. Such an approach highlights how *in silico* methods can assist in the development of validated 3D models by exchanging information with *in vitro* experiments or with different 3D models and pharmacophores. *In silico* methods are positioned at the cycle center and the information flow is radial. As a compound enters a test cycle it encounters several control points (CPs). Each CP contains a family of targets. For instance, for phase I, a set of cytochromes allows the evaluation of the potential for metabolism of the drug in question. This strategy rapidly increases the accuracy of the *in silico* model for CYP metabolism because information about a CYP's pharmacophore helps in the design and adjustment of the pharmacophores of other CYPs. On demand, pharmacophores can be modified (to be more or less selective) depending on the criteria established and imposed by the investigator. As in a real cell, a CP can block the workflow progression to adjust *in silico* or *in vitro* strategy. One compound is input in the cycle and subject to one rotation. The following compounds may then be introduced one after the other. It may be necessary to have more than one revolution of the cycle to obtain valid data.

Figure 5, a three-dimensional view of Fig. 4, illustrates a 'kaizen'-like representation of the quality management imposed upon the

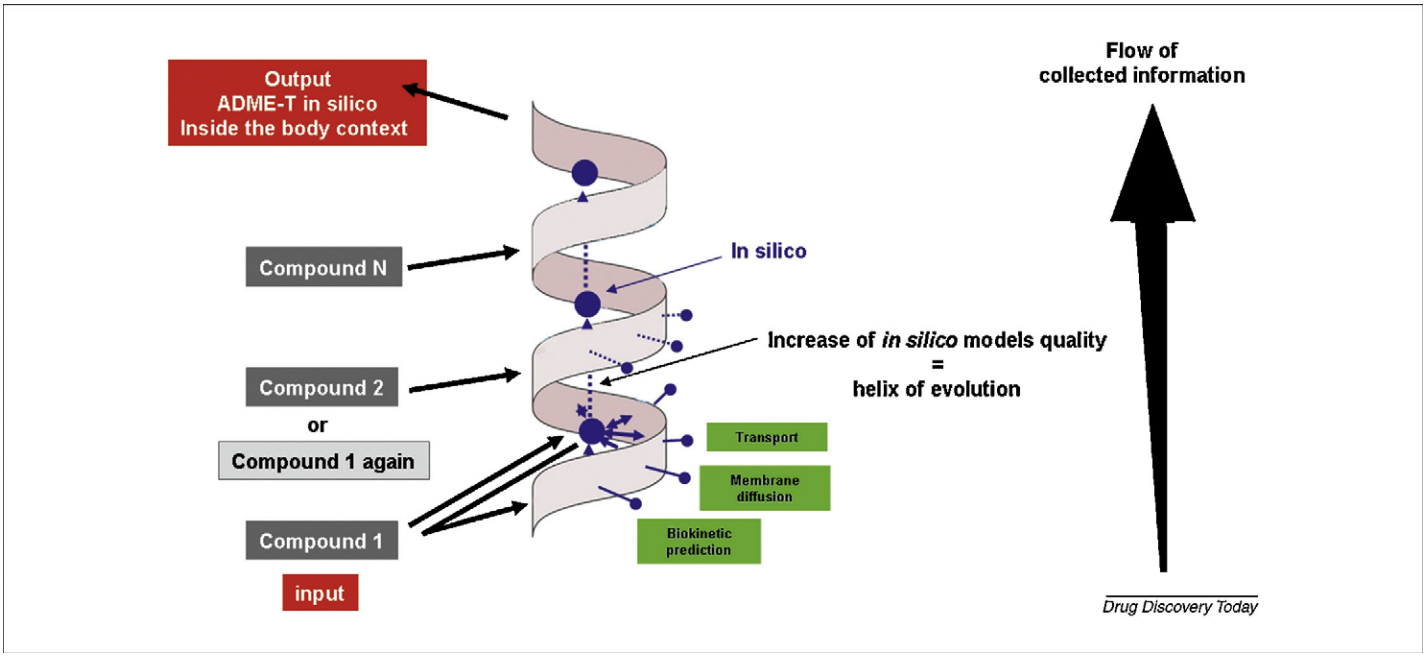

FIGURE 5

Three-dimensional representation of the 'gradual incremental improvement' of the predictive tools drawn in Fig. 4. The arrow indicates the increase of ADME-T predictive power of the *in silico* platform.

in silico toolset and methods [24]. Kaizen was invented by W Edwards Deming, an American statistician who has been associated with the rise of Japan as a manufacturing nation and with the invention of total quality management (TQM). Deming established the foundation of quality insurance in industry using 14 management points. He was inspired by Japanese methods. The Kaizen (gradual improvement) is originally a Japanese management concept for gradual continuous (incremental) change (improvement). It assumes that every aspect of our life deserves to be constantly improved. The kaizen philosophy lies behind many concepts such as: TQM, quality control cycles, small group activities and labor relations. Key elements of Kaizen are quality, effort, involvement of all members of each group, willingness to change and communicate. This is precisely the spirit of Fig. 5 that shows a constant improvement of the *in silico* models while spinning in an ascendant helix. Each helix turn passes over the previous one but at a higher quality level. The greater the number of compounds, the higher the quality of the *in silico* models and the accuracy of internal standards. The information flow is growing, compound after compound. Each expert (biologist and chemist) stores information about previous compounds and implements its knowledge while in the same helix turn all the experts communicate with each other.

Standardization of the *in silico* methods

Another 'innovative touch' to the *in silico* toolset was brought by technology transfer from hospitals. It deals with the standardization techniques existing in clinical biology where patients' biological fluids sample analyses require a high level of accuracy, reproducibility and repeatability. A recent report on the 'Economic Analysis of the Technology Infrastructure Needs of the U.S. Biopharmaceutical Industry' made by the US National Institute for Standards and Technology [25], underlines the urgent need and absolute necessity to create *in silico* methods standardization. A series of simulation protocols has been established to frame and reduce errors made in the predictability power. Standardization technologies will be further detailed elsewhere.

Conclusion

In this review we have shown that an *in silico* platform for modeling and prediction of ADME-T properties of any drug cannot be obtained by simply making computational chemists and biologists work together. Our technology relies on a small group of scientists that includes healthcare professionals with expertise in computational chemistry working close to the patient, who take on the careful quality control and management of their standardized tools and results obtained. A dynamic flow of information is maintained between *in silico* and *in vitro* experiments, with a limited but sufficient number of *in vivo* validations and with an interpretation in the 'patient whole body context'. It does not mean creating a perfect tool that makes no error in ADME-T prediction. It rather means, after technology transfer from clinical biology, establishing a clear protocol that allows the identification of an error, its origin, to quantify it and give the means, if possible, to eliminate it. Another prerequisite to standardization is the adaptation of tools to their specific task. That is the essence of the review. Along the gradually increasing number of modeled compounds and protein targets, our *in silico* ADME-T-Lead Op package of algorithms improves. The keystone of ADME-T-Lead Op modeling and prediction is a similarity measure performed with the help of an efficient pharmacophore definition in spherical coordinates that allows the determination of *N*-dimensional quantitative measures of the structure–activity relationships (not simply by linear regression of 2D descriptors), completed by quantum chemistry calculations, *à la carte* elaborated force fields and, finally, mathematical integration of all *in silico* simulations carried out along biotransformation phases I–III. Working on dozens of 3D models of proteins involved in biotransformation helps identifying origins of compounds' toxicities and may provide ways for overcoming them with lead optimization. Further, beyond toxicity prediction, *in silico* tools could help focusing chemical synthesis and repositioning of any drug.

Acknowledgement

We warmly thank Miss Véronique Legrand, the artist who kindly drew Fig. 1.

References

1 Tekto, I.V. *et al.* (2006) Can we estimate the accuracy of ADME–Tox predictions? *Drug Discov. Today* 11, 700–707
2 Kumar, N. *et al.* (2006) Applying computational modelling to drug discovery and development. *Drug Discov. Today* 11, 806–811
3 Mack, G.S. (2004) Can complexity be commercialized? *Nat. Biotechnol.* 22, 1223–1229
4 Dobson, C.M. (2004) Chemical space and biology. *Nature* 432, 824–828
5 Goto, S. *et al.* (2002) LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.* 30, 402–404
6 Pratuangdejkul, J. *et al.* (2005) Definition of an uptake pharmacophore of the serotonin transporter through 3D-QSAR analysis. *Curr. Med. Chem.* 12, 2393–2410
7 Jørgensen, A.M. *et al.* (2007) Molecular dynamics simulations of Na^+/Cl^--dependent neurotransmitter transporters in a membrane–aqueous system. *Chem. Med. Chem.* 2, 827–840
8 Clark, D.E. and Pickett, S.D. (2000) Computational methods for the prediction of drug-likeness. *Drug Discov. Today* 5, 49–58
9 Manivet, P. *et al.* (2002) The serotonin binding site of human and murine 5-HT2B receptors: molecular modeling and site-directed mutagenesis. *J. Biol. Chem.* 277, 17170–17178
10 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliver. Rev.* 23, 3–25
11 Leo, A. *et al.* (1971) Partition coefficients and their uses. *Chem. Rev.* 71, 525–616
12 Roberts, D.W. *et al.* (2006) Mechanistic applicability domains for non-animal based prediction of toxicological endpoints QSAR analysis of the Schiff base applicability domain for skin sensitisation. *Chem. Res. Toxicol.* 19, 1228–1233
13 Obrezanova, O. *et al.* (2008) Automatic QSAR modeling of ADME properties: blood–brain barrier penetration and aqueous solubility. *J. Comput. Aided Mol. Des.* 22, 431–440
14 Darré, J.P., ed. (1999) *La production de la connaissance*, MSH – Editions de la maison des Sciences de l'Homme
15 Snyder, R.D. *et al.* (2004) Assessment of the sensitivity of the computational programs DEREK, TOPKAT, and MCASE in the prediction of the genotoxicity of pharmaceutical molecules. *Environ. Mol. Mutagen* 43, 143–158
16 Duarte, M.P. *et al.* (2007) The stimulatory role of human cytochrome b5 in the bioactivation activities of human CYP1A2 2A6 and 2E1: a new cell expression system to study cytochrome P450-mediated biotransformation. *Mutagenesis* 22, 75–81
17 Kool, J. *et al.* (2007) Cytochrome P450 bio-affinity detection coupled to gradient HPLC: on-line screening of affinities to cytochrome P4501A2 and 2D6. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 858, 49–58
18 Russel, F.G. *et al.* (2008) Multidrug resistance protein 4 (MRP4/ABCC4): a versatile efflux transporter for drugs and signalling molecules. *Trends Pharmacol. Sci.* 29, 200–207

19 Spears, K.J. *et al.* (2005) Directional *trans*-epithelial transport of organic anions in porcine LLC-PK1 cells that co-express human OATP1B1 (OATP-C) and MRP2. *Biochem. Pharmacol.* 69, 415–423
20 Ingelman-Sundberg, M. (2008) Pharmacogenomic biomarkers for prediction of severe adverse drug reactions. *N. Engl. J. Med.* 358, 637–639
21 Teubner, W. *et al.* (2007) Identification and localization of soluble sulfotransferases in the human gastrointestinal tract. *Biochem. J.* 404, 207–215
22 Guérin, G.A. *et al.* (2006) Rational and efficient geometric definition of pharmacophores is essential for the patent process. *Drug Discov. Today* 11, 991–998
23 Pratuangdejkul, J. *et al.* (2006) Conformational dependence of serotonin theoretical pK_a prediction. *Chem. Phys. Lett.* 420, 538–544
24 Deming, W.E., ed. (1986) *Out of the Crisis*, MIT Press
25 Gallaher, M., eds *et al.* (2007) Economic analysis of the technology infrastructure needs of the U.S. biopharmaceutical industry. *Final Planning Report 07-1*. US National Institute for Standards and Technology