Computational approaches to the understanding of ADMET properties and problems

Sandeep Modi, Computational and Structural Science Group (CASS), GlaxoSmithKline R&D, Stevenage, Herts, UK SG1 2NY; e-mail: sandeep.2.modi@gsk.com

Different approaches for the modelling of ADMET properties (absorption, distribution, metabolism, excretion and toxicity), together with their applications and limitations, were addressed at the Molecular Graphics Modelling Society (MGMS) meeting (http://www.mgms.org/) at Brasenose College, Oxford, which was held from 2nd to 4th April, 2003. The meeting consisted of a series of presentations given by acknowledged leaders in the field, followed by short oral presentations and poster sessions.

The need for *in silico* ADMET tools

It is now widely recognised that the definition of a successful medicine is an appropriate balance of potency, safety and favourable pharmacokinetics (PK). Several studies [1] in the late 1990s suggested that poor PK and toxicity were among the most important causes of late-stage failures of compounds in drug development. Drug metabolism, PK and toxicity testing is now carried out at a much earlier stage of drug discovery. However, there has recently been a dramatic increase in the size of compound collections in pharmaceutical companies and, owing to ultra high-throughput screening, an increase in the rate at which biological activity data can be obtained. Consequently, the drug discovery process has struggled to keep up with the generation of compounds and screening data owing to time consuming and costly PK and safety studies at the lead optimisation stage. This meeting covered a number of *in silico* ADMET tools that have been developed in order to try and address these issues. Many speakers discussed the advantages of these *in silico* tools, which include: (1) the ability to screen more compounds in less time, thereby accelerating the drug discovery process; (2) the ability to run these models on virtual compounds, so that an appropriate balance of various properties can be chosen from the large virtual libraries for chemical synthesis; and (3) an understanding of the relationship between these ADMET properties and the structure and physiochemical properties of a compound can be gained.

Han van de Waterbeemd (Pfizer R&D; http://www.pfizer.com) gave an overview of the need for *in silico* ADMET tools, the tools that are available and the state of the art in applying them. Chris Lipinski, also from Pfizer, talked about various biological targets and how, for some of these targets, it is easier to achieve a balance of favourable ADME properties and activity than it is for others. He showed that it is common for compounds from combinatorial libraries to display poor aqueous solubility and a low density of functionalisation. Problems such as these are more closely linked to problems in chemical synthesis than to problems in library design. As a result of this, no matter how smart a library design is, it can be fatally flawed by the fact that the synthesis might only work for a subset of the compounds required and that the compounds made will not have the designed profile. He also mentioned that for every 57 projects pursued only one succeeds.

Methods and tools for developing *in silico* ADMET models

These tools vary in their throughput, prediction accuracy and range of statistical methods and descriptors. Descriptors, for example, can be based on simple whole-molecule properties (e.g. hydrogen-bond donors and acceptors, size, logP/D, among others) or can be semi-empirical methods based on quantum theory. For data modelling, quantitative structure–property relationship studies have been applied since the 1960s. Several approaches were discussed in this meeting, including simple multiple linear regression, stepwise regression, rule-based methods and more complex techniques such as partial least square (PLS), neural networks (NN) and genetic programming (GP). Although the latter techniques are complex and result in models that are viewed as black boxes, William Bains's (Amedis Pharmaceuticals; http://www.amedis-pharma.com) discussion highlighted the fact that no complex biological process can be modelled from first principles; all biological models are approximations that treat their fundamental units as black boxes. The biological knowledge of the system and all the processes associated with it

1359-6446/03/$ – see front matter PII: S1359-6446(03)02739-9

describe how big the black box is. He also discussed the idea of a new descriptor called 'life-likeness', which relates any given compound to a known molecule already present in a cell via a similarity metric. Dave Winkler [Commonwealth Scientific and Industrial Research Organisation (CSIRO); http://www.csiro.au] described some models that were produced using Bayesian neural nets (BNNs) and discussed how these can be used to model important processes such as blood–brain barrier (BBB) penetration, intestinal absorption and acute toxicity. BNNs have the property of being difficult to over-train, which is one of the drawbacks of classical NNs. He also raised some topical questions concerning the role of complexity in bioactive modelling.

The descriptors used for most methods of correlation and prediction of *in vitro* or *in vivo* ADME data range from simple descriptors such as lipophilicity, hydrogen bonding, flexibility, molecular size and VolSurf approaches [2] to more complex ones, such as topological descriptors. Mike Abraham from UCL (http://www.ucl.ac.uk) described the use of his five experimentally derived descriptors to correlate a number of biological and physicochemical properties. He showed that no known water–solvent partition process (e.g. water–octanol) is a suitable model for human intestinal absorption. Tim Clark (Erlangen University, Germany; http://www.ccc.uni-erlangen.de) focused on new types of descriptors, semi-empirical molecular orbital (MO) techniques, and ways of using these descriptors in ADME modelling. He has developed new concepts of local (i.e. atom)-centred quantum mechanical (QM) properties such as polarisability. The use of these descriptors allowed a compact three-dimensional (3D) and electrostatic description of molecules and whole-molecule properties to be formed.

Intestinal and CNS permeation modelling

Sandeep Modi (GlaxoSmithKline; http://www.gsk.com) described several intestinal absorption models [e.g. Rule of 5, logD/calculated molar refractivity (CMR) model and polar surface area (PSA)] that can predict the passive permeation of compounds across membranes. However, most of these methods ignore the effects of other biological processes, such as membrane transport and metabolism. More advanced programs, such as Gastroplus™ (Simulations Plus; http://www.simulations-plus.com) [3], Idea™ (Lion Bioscience, http://www.lionbioscience.com) [4] and Cloe™ (Cyprotex; http://www.cyprotex.com) can predict the rate and the extent of intestinal absorption using simulation models of mammalian physiology and these programs utilise a number of *in vitro* data inputs. David Leahy from Cyprotex described the methods for predicting *in vivo* clearance, bioavailability, volume of distribution and plasma–time profile in rat and humans.

Several published *in silico* methods of BBB penetration prediction were reviewed by David Clark (Argenta; http://www.argentadiscovery.com). It was also shown that most of the drugs that cross the BBB are lipophilic, have few hydrogen-bond donors and acceptors and are also less flexible. Clark described that from a test set of approximately 30 compounds, >80% of the 'good' BBB penetrating compounds were identified using a model based on VolSurf descriptors.

Drug metabolism modelling

Several aspects of metabolism modelling (e.g. rate, extent, the enzymes involved in metabolism, the products and metabolites formed, cytochrome P450 inhibition, among others) were covered in this meeting. Patrizia Crivori (Pharmacia; http://www.pfizer.com) discussed the modelling of CYP3A4 metabolic stability data using the VolSurf and Almond packages. Her model was 80% correct in predicting the CYP3A4 metabolic stability of the external test set. Dave Lewis (University of Surrey; http://www.surrey.ac.uk) described the molecular modelling of important human cytochrome P450s based on the X-ray structure of mammalian CYP2C5. He also compared the results of his homology models with the *in vitro* metabolic fate of different substrates and recent site-directed mutagenesis studies. Chris Murray (Astex; http://www.astex-technology.com) described the first human X-ray structure of CYP2C9. He compared the active site of the 2C9 X-ray structure with existing homology models and also pharmacophores derived from known 2C9 substrates and/or inhibitors. The major differences that were observed were in the conformation of the BC loop and the role of Arg 97 in binding to the propionates of the heme rather than as a binding site for acidic ligands. Early predictions of the positions in a drug molecule that are metabolised might help to improve lead selection.

Gabriele Cruciani (University of Perugia; http://www.unipg.it) presented the Metasite program, which is based on a pharmacophore representation obtained from interaction fields for the 2C9 and 2D6 enzymes and a 'fingerprint' of their substrates.

There is now a large amount of information available on the enzymatic pathways involved in foreign compound biotransformations. All of this information can be codified into a series of rules, based on either empirical or on mechanistic considerations, for predicting the probable routes of metabolism. Alan Boobis from Imperial College, London (http://www.ic.ac.uk) discussed the advantages and disadvantages of such methods, using the commercially available program

METEOR (Lhasa; http://www.chem.leeds.ac.uk/luk/) as an example. Dr Amin Rostami (University of Sheffield; http://www.shef.ac.uk) discussed the problems of paediatric ADME studies and the need for better experimental data. He noted that in the adult population, the issue of varying levels of CYP2D6 in any individual owing to factors such as health, age and ethnic origin is as much a problem in clinical practice as are problems owing to polymorphism of this P450.

There was also talk by Bernard Pirard (Aventis Pharma; http://www.aventis.com) who described the pharmacophore model for Kv1.5 channel activity. Blockers of the human Kv1.5 channel could be used to treat atrial fibrillation, which is one of the most common sustained cardiac arrythmias.

Concluding remarks

This conference highlighted the importance of computational methods, informatics, modelling and predictions in extracting useful information from the large volume of experimental data. The levels of attrition of new drugs is still as high as ever and it is also becoming much more difficult to bring new compounds to the market due to regulatory concerns over safety. Applying appropriate filters (e.g. *in silico* ADME models) to large libraries should improve the hit rate and also reduce the duration of the R&D process. However, the models are clearly only as good as the data they are based on, and, unfortunately, in most cases, the data sets are rather limited. This meeting also raised the importance of good experimental data sharing and molecular descriptors in order to develop better global models. It is also important that models are continuously validated and refined. In many areas, a combination of models should be tried. For example, in the case of oral absorption modelling a combination of permeation, solubility and membrane transport models should be employed. These developments will have a fundamental impact on the future of drug discovery. It is accepted that these tools represent a fast method to enrich a biological screen. Many pharmaceutical companies have already adopted virtual screening methodologies to complement *in vitro* HTS methods.

References

1 Kennedy, T. (1997) Managing the drug discovery/development interface. *Drug Discov. Today* 2, 436–444
2 Cruciani, G. *et al.* (2000) VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. *Eur. J. Pharm. Sci.* 11 (Suppl.2), S29–S39
3 Agoram, B. *et al.* (2001) Predicting the impact of physiological and biochemical processes on oral drug bioavailability, *Adv. Drug Del. Rev.* 90, S42–S67
4 Noris, D.A. *et al.* (2000) Development of predictive pharmacokinetics simulation models for drug discovery. *J. Control. Release* 65, 55–62