6 Bajorath, J. (2002) Integration of virtual and high-throughput screening. *Nature Rev. Drug Discov.* 1, 882–894

Jürgen Bajorath
Senior Director
Computer-Aided Drug Discovery
Albany Molecular Research
Bothell Research Center
and University of Washington
Seattle
Washington, USA
e-mail: jurgen.bajorath@albmolecular.com

Positioning ADMET *in silico* tools in drug discovery

Recently there has been a dramatic increase in the size of compound collections in pharmaceutical companies and also, because of ultra high-throughput screening, an increase in the rate at which biological activity data can be obtained. The generation of pharmacokinetic and safety data at the lead optimization stage has therefore struggled to keep up with the screening of compounds. Virtual screening of compound libraries has also become a component of many lead generation and optimization programmes. Large pharmaceutical and many smaller chemistry-oriented companies support substantial efforts in this area. The hope is that *in silico* screening and analysis will aid significantly in addressing the balance between drug potency and various ADMET properties.

Although ADMET *in silico* tools can be used for high-throughput screening, their main benefit is in predicting properties of compounds before they are synthesised and also in understanding the relationship between chemical structure and ADMET properties. In my personal experience I have seen a number of examples that show that blending measurements together with information generated by these assays can offer a better chance of success than just by testing more compounds alone. However, ADMET *in silico* models are frequently complex and it is perhaps not surprising that many experimentalists perceive this field as an 'algorithmic jungle'. Consequently, the benefits and limitations of virtual screening are sometimes misunderstood.

In silico predictions are probably no less predictive of what occurs *in vivo* than are *in vitro* tests. They have the decisive advantage of being cheap and they enable predictions to be made on virtual compounds. Although these *in silico* models do undoubtedly work, in many cases experimentalists often prefer to generate 'wet' data on all the compounds, irrespective of what the odds of success might be. If we are to fully capitalize on the opportunities presented by *in silico* tools, implementation and integration of these tools into drug discovery processes needs to be carried out in a rational and systematic manner. We need to better understand the relationship between the physicochemical properties and structure of a molecule and its likely fate in the body. For example, for intestinal absorption there are at least five different processes that can affect the absorption of a molecule. These are passive paracellular, passive transcellular, efflux and/or influx transporters, solubility and dissolution rate. As with any complex problem, the task of building an understanding becomes less daunting if it can be broken into different, simple processes.

We also have to resolve the issue of whether global or local models should be used. Global models that are based on data from several programmes are usually good in modeling general phenomena and trends but are less good at enabling an understanding of the effect of small structural differences. By contrast, local models, which are built on a particular chemical series, might work well within a closely related series but will quickly cease to be applicable as the synthetic direction of a programme changes. *In silico* modelers should therefore check the validity of their models by selecting and obtaining experimental data from new compound sets with structures different from those in the original model training sets. It is also important that *in silico* modelers choose the right data to model and use descriptors that convey a simple message to experimentalists. This is why Lipinski's 'Rule of Five' [1] is widely used and has gained acceptance with chemists and biologists. As these tools become more user-friendly, and as more examples of successful applications are shown, it seems highly probable that *in silico* approaches will evolve rapidly, as was the case with *in vitro* methods during the last two decades. It should also be acknowledged that *in silico* tools have been in existence for a relatively short time and it is therefore unrealistic to expect good predictions in every application. However, *in silico* modelers must adhere to best practice, performing adequate validation and testing the reliability of the predictions to ensure that these tools are used appropriately. It is also equally important for model developers to work with programme teams to explore and understand the reasons why these tools demonstrate limitations in certain cases.

Drug discovery has always been a competitive industry and now the stakes are even higher than ever. In terms of sales and profits, only one in five new products launched since 1997 has been 'significant' [2]. A combination of technologies such as HTS and *in silico* models can offer great advantages in improving the odds of success in a discovery programme. In the early stages this can be done by eliminating the guesswork and decreasing the experimental load. What needs to be avoided is using *in silico* models simply in addition to *in vitro* and *in vivo* experiments, rather than being used in

 1359-6446/04/$ – see front matter PII: S1359-6446(03)02956-8

decision making. Validation, awareness of the limitations and proper use of these tools is crucial if they are to impact drug discovery and development. By working alongside programme scientists to understand the processes being modeled, *in silico* modelers can help build confidence in these *in silico* tools, such that they become as integrated into the discovery process as *in vitro* methodologies currently are.

References

1 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev* 23, 3–25

2 Hecht, P. (2002) High-throughput screening: beating the odds with informatics-driven chemistry. *Curr. Drug Discov.* January issue, 21–24

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