ELSEVIER

www.elsevier.nl/locate/farmac

Il Farmaco 56 (2001) 149–151

IL FARMACO

Prediction of drug toxicity

Mark T.D. Cronin

School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

Abstract

The use of computer-aided methods to predict the toxicity of drugs is described. These methods can assist in the identification of toxic compounds early in the drug development process. Thus, there is potential for these methods to be combined with combinatorial synthesis and library design. Quantitative structure–activity relationships allow for the prediction of individual endpoints, usually for restricted groups of compounds. Expert systems for toxicity prediction are based on a number of methodologies, each with its own strengths and weaknesses. The relative merit of each individual technique and methodology is described. However, more toxicity data are required, both to produce and to validate expert systems. Potential sources of new data include the use of high-throughput screening and microarrays for toxicology. © 2001 Elsevier Science S.A. All rights reserved.

Keywords: Toxicity; Quantitative structure–activity relationship; Expert system; Prediction

1. Introduction

A large number of time-consuming and costly in vivo animals tests are still required to assess the toxicity of a drug substance. Owing to the cost in both resources and time involved in in vivo toxicological assessment, there has recently been a considerable upsurge in interest in alternative methods to predict directly from physico-chemical structure the toxicity of drugs. This coincides with the reality of high throughput screening for both pharmacological and toxicological endpoints, and the availability of enormous compound libraries.

The interest in computer-aided toxicity prediction has been spawned as a result of the realisation that the drug development process is slowed, and often terminated, by an adverse toxicity profile. It would seem entirely reasonable, therefore, that requests should be made for predictions of toxicity, as soon as a lead compound is identified. The simplest method of doing this is to make a prediction directly from physico-chemical structure. The purpose of this paper, therefore, is to review briefly the use of computer-aided techniques for toxicity prediction, to assess their current level of usefulness, and to consider how they may be employed in the future.

2. Computer-aided toxicity prediction

The computer-aided prediction of toxicity is based upon the premise that the toxicity of a drug is related to its physico-chemical structure. Thus, if one knows the structure of potential new drugs (i.e. the members of a combinatorial library), then it should be possible to predict their toxicological profile. There are various methods to achieve this goal. Quantitative structure–activity relationships (QSARs) refer to the development of relationships between structure and toxicity of a series of chemicals. Normally the chemicals in the series are related in some manner, i.e. they are structurally similar or are considered to be acting by the same mechanism of toxic action. QSARs have the advantage that they can provide relatively accurate predictions of toxicity, but only normally for restricted series of compounds. Owing to its structural diversity, therefore, the use of QSARs may be inappropriate for the prediction of toxicity of a complete combinatorial library.

As an example of the prediction of basal cytotoxicity, the author and co-workers have developed highly significant QSARs. These indicate that toxicity can be related to the ability of a compound to reach the site of action, and its ability to react covalently there. Such phenomena can be modelled adequately using physico-chemical parameters for transport, i.e. hydrophobicity (described by the logarithm of the octanol–water parti-

E-mail address: m.t.cronin@livjm.ac.uk (M.T.D. Cronin).

PII: S0014-827X(01)01018-7

tion coefficient, log P) and electrophilicity (described by the energy of the lowest unoccupied molecular orbital, E_{LUMO}). Thus, Cronin et al. [1] found that the toxicity to the bacterium *Vibrio fischeri* (the so-called Microtox test, pT_{15}) of 63 aliphatic compounds, putatively with a range of different mechanisms of action, was well correlated to log P and E_{LUMO}:

$$pT_{15} = 0.76 \log P - 0.63E_{LUMO} - 0.47$$
$$n = 63,\ r^2 = 0.85,\ s = 0.46,\ F = 171 \quad (1)$$

where n is the number of observations, r^2 is the square of the correlation coefficient, s is the standard error of the estimate, and F is the Fischer statistic.

Cronin et al. [2] have also reported the development of similar models to Eq. (1) for a range of toxic endpoints. These were also based on the premise that parameters for hydrophobicity and electrophilicity are important for the description of toxicity. For instance, the hepatotoxic effects of branched and unbranched short-chain aliphatic alcohols are well predicted following the inclusion of a third parameter for the degree of molecular branching (third-order path cluster molecular connectivity, $^3\chi_{PC}$):

$$\log GPT = 0.58 \log P - 0.19E_{LUMO} + 0.49\ ^3\chi_{PC} - 1.19$$
$$n = 23,\ r^2 = 0.84,\ s = 0.18,\ F = 39 \quad (2)$$

where GPT is the extracellular release of glutamine-pyruvate transaminase from the perfused rat liver.

To summarise, QSARs provide a means of predicting a range of toxicological endpoints. The use of parameters to describe hydrophobicity and electrophilicity seems appropriate to model many of these endpoints (for further illustrations of this approach, the reader is referred to citations listed in Cronin et al. [1,2]). A large number of other QSARs have also been reported. Many of these have been compiled into the C-QSAR database, which provides a comprehensive listing [3]. The downside of individual QSARs for toxicity prediction is that they provide predictions only for limited groups of defined compounds, and for single endpoints. These latter points make their usefulness for the prediction of toxicity of combinatorial libraries rather limited.

3. Expert systems for toxicity prediction

An expert system may be described as 'any formalised system, not necessarily computer-based, which enables a user to obtain rational predictions about the toxicity of chemicals' [4]. Typically they comprise a computational system that is able to make a prediction of toxicity automatically, following the input of a molecular structure. There are a number of well-established, commercially available expert systems. Such systems have been reviewed critically by Dearden et al. [4]

Table 1
Comparative properties of QSARs and expert systems for the prediction of toxicity

	Mechanistically based	Uses reliable and high quality data	Capable of operating in batch mode	Able to assign confidence to the prediction	Source of the knowledge	Capable of being integrated easily with computational library design	Corporate web-site (where applicable and available)
QSARs	Yes	Yes	No	No	Derived	No	
TOPKAT	No	Yes	Yes	Yes	Derived	Yes	www.oxfordmolecular.com/software/topkat/
DEREK	Yes	Yes	Yes	No	Human	Yes	www.chem.leeds.ac.uk/luk/derek/
Hazard Expert	Yes	Yes	Yes	No	Human	Yes	www.compudrug.com/hazard.html
CASE	No	Yes	Yes	No	Derived	Yes	www.multicase.com/

and Cronin [5], and their main attributes are summarised in Table 1.

There are a number of different philosophies behind the development of expert systems for toxicity prediction. A number of systems effectively automate QSAR predictions. Of these, TOPKAT is probably the best known and developed. TOPKAT provides predictions for a variety of endpoints, based on QSAR analyses of large heterogeneous databases. As such, however, the TOPKAT methodology does not allow for development of the models from a mechanistic standpoint. A unique feature of the software, however, is the optimum prediction space algorithm [6], which allows for an estimate of the confidence of the prediction to be assigned. Other expert systems, such as DEREK and HazardExpert, utilise a knowledge-base derived from human expertise. In these systems, structural alerts are associated with toxicological effects. The advantage of this approach is that it is highly mechanistically based, and a thorough analysis of the information means that the quality and reliability of the toxicological information is assured. A number of systems, including those based around the computer-automated structure-evaluation (CASE) methodology, derive rules and the association between molecular fragments and toxicity automatically. Such techniques are capable of analysing large databases of toxicological information rapidly, but do so at the expense of being able to add no mechanistic information.

4. Conclusion

To conclude, there is a wide variety of expert system technologies and approaches for the prediction of toxicity. However, these provide accurate predictions of toxicity only for compounds structurally similar to those on which the system is based. It should also be recognised that there is paucity of available data both for the production of models for toxicity prediction, and for the validation of extant models. No single toxicity prediction approach is likely to be successful, so a battery approach using a variety of systems is recommended. Further, we are now gaining the ability to create large amounts of toxicological information, both from high-throughput screening and microarray technologies [7]. Such data may well prove to be ideal for the development of the next generation of toxicity prediction systems.

References

[1] M.T.D. Cronin, G.S. Bowers, G.D. Sinks, T.W. Schultz, Structure–toxicity relationships for aliphatic compounds encompassing a variety of mechanisms of toxic action to *Vibrio fischeri*, SAR and QSAR in Environmental Research 11 (2000) 301–312.

[2] M.T.D. Cronin, J.C. Dearden, J.C. Duffy, R. Edwards, N. Manga, A.P. Worth, A.D.P. Worgan, The importance of hydrophobicity and electrophilicity descriptors in mechanistically-based QSARs for toxicological endpoints, SAR and QSAR in Environmental Research 11 (2000) in press.

[3] C. Hansch, A. Leo, Exploring QSAR. Fundamentals and Applications in Chemistry and Biology, American Chemical Society, Washington, 1995.

[4] J.C. Dearden, M.D. Barratt, R. Benigni, D.W. Bristol, R.D. Combes, M.T.D. Cronin, P.N. Judson, M.P. Payne, A.M. Richard, M. Tichy, A.P. Worth, J.J. Yourick, The development and validation of expert systems for predicting toxicity, Alternative to Laboratory Animals 25 (1997) 223–252.

[5] M.T.D. Cronin, Computational methods for the prediction of drug toxicity, Current Opinion in Drug Discovery and Development 3 (2000) 292–297.

[6] V.K. Gombar, Reliable assessment of log P of compounds of pharmaceutical relevance, SAR and QSAR in Environmental Research 10 (1999) 371–380.

[7] E.F. Nuwaysir, M. Bittner, J. Trent, J.C. Barrett, C.A. Afshari, Microarrays and toxicology: the advent of toxicogenomics, Molecular Carcinogenesis 24 (1999) 153–159.