Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 1037–1039

A neural network based classification scheme for cytotoxicity predictions: Validation on 30,000 compounds

László Molnár,[a] György M. Keserű,[a,*] Ákos Papp,[b] Zsolt Lőrincz,[c,d] Géza Ambrus[c] and Ferenc Darvas[b]

[a]*Department of Chemical Information Technology, Budapest University of Technology and Economics, Szent Gellért tér 4., H-1111 Budapest, Hungary*
[b]*ComGenex, Inc., Bem rkp. 33-34., H-1027 Budapest, Hungary*
[c]*RecomGenex, Ltd, Záhony u. 7., H-1031 Budapest, Hungary*
[d]*Institute of Enzymology, Hungarian Academy of Sciences, Karolina u. 29., H-1113 Budapest, Hungary*

Received 14 September 2005; revised 20 October 2005; accepted 24 October 2005
Available online 8 November 2005

Abstract—Elimination of cytotoxic compounds in the early phases of drug discovery can save substantial amounts of research and development costs. An artificial neural network based approach using atomic fragmental descriptors has been developed to categorize compounds according to their in vitro human cytotoxicity. Fragmental descriptors were obtained from the Atomic7 linear log*P* calculation method implemented in Pallas PrologP program. We used cytotoxicity values obtained from an in-house screening campaign of a diverse set of 30,000 drug-like molecules. The training set included only the most and least toxic 12,998 compounds, however, cytotoxicity data for all compounds were used for validation. The proposed approach can be safely used for filtering out potentially cytotoxic candidates from the development pipeline before synthesis or assays during lead development or lead optimisation. The trained neural network misclassified less than 5% percent of the non-toxic and 9% of the toxic compounds.

In silico prediction techniques are widely used to filter large compound libraries based on their ADME characteristics, but probably due to the questionable accuracy, these tools are not popular in prediction of more complex parameters, like toxicity. The non-linearity proven to be involved in the relationship between the molecular structure and ADME/Tox properties explains the successful application of neural networks on this field.[1–3] One of the keys in building a predictive neural network is to identify the appropriate molecular descriptor type, which is able to describe the SAR information encoded in the molecular structure. Although most of the neural network approaches use specifically tailored molecular descriptors for this purpose, our present method is based on simple and easily accessible atomic fragmental descriptors. The fragment definitions are identical to those used in the knowledge base of the Atomic7 linear log*P* calculation method implemented in Pallas PrologP[4] program. The descriptor set is based on Ghose-Crippen[5] fragmentation technique, but uses modified atomic contributions and additional correction terms. This approach is a new member of a novel class of pseudo-linear algorithms,[6] where the precision of the non-linear approaches is combined with the transparency of the early linear methods.

In this work, we report an artificial neural network based non-linear approach, which is able to predict cytotoxicity at high accuracy. We used a diverse set of 30,000 drug-like compounds composed of 19 distinct synthetic libraries.

The in vitro assay was based on measuring the toxicity and antiproliferative effects of small molecules on a human fibroblast cell line, MRC-5. The following criteria were considered for the choice of the cell line: (i) it should be a well-described human cell line, (ii) it should proliferate quickly, that is, the duplication time of the cells should not exceed 30 h and (iii) the cell line should be viable and devoid of any phenotypical changes after several passages.

Keywords: Cytotoxicity; Measurement; Human fibroblast cell line; MRC-5; Classification; Neural network; Atomic7 descriptor set.
* Corresponding author. Tel.: +36 1 4634141; fax: +36 1 4633953; e-mail: gykeseru@goliat.eik.bme.hu

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.10.079

There are many cell lines that fulfil these requirements; however, beyond these principles the selection of the cell line is not critical in HTS cytotoxicity measurements. According to previous studies, there are no significant differences between the basal cytotoxicity results on mammalian cell lines and on primer human cells.[7]

Measurements on human primer hepatocytes or HepG2 cell lines would more likely correlate better with in vivo toxicity data; nevertheless, the extensive metabolism in hepatocytes and HepG2 cells would render the modelling of structure based toxicity difficult, since the experiments would be biased by toxic metabolites of non-toxic compounds as well as by non-toxic metabolites of toxic compounds.

We finally concluded to measure the in vitro human fibroblast cytotoxicity using human fibroblast cells that fulfils all of the criteria we set. MRC-5 foetal lung cell line has been grown in culture flasks to 90% confluency, then harvested in counted cell density and seeded into 384-well microtitre plates. The test compounds are added in various concentrations (in 1% DMSO final concentration) and incubated with the cells in CO_2 incubators at 37 °C for one cell cycle (24 h). During this period, we decreased the FBS content of the medium to 2% to avoid the masking effect of FBS on toxicity. For cell viability measurements we have used indicator alamarBlue, a fluorimetric/colourimetric indicator, which indicates metabolism of cells, is soluble in water, and is non-toxic. The alamarBlue is getting popular because it is easy to handle, it has good sensibility and it can be compared to earlier measurements.[8] Since it does not influence viability of cells,[9] it can be used for kinetic measurements. After 4 h of incubation with alamarBlue, we measured the cell viability by the fluorescent technique using exciting light at 544 nm and measuring the emitted light at 590 nm.[10] The negative control was 1% DMSO in culture medium, while the positive control was Triton X-100. The cytotoxicity assays were previously validated with compounds with known IC_{50} and LD_{50} values, and were optimised to provide *Z*'-factors generally above 0.6, which indicates the robustness of the assay.[11] The viability data at 60 μM have been determined for all compounds, and in case of viability less than 40% we determined IC_{50} values by measuring the compound either in 4 point duplicate or in 6 point triplicate, scenarios.

Table 1. 'Confusion' matrix for Training set (a), Test1 set (b), Test2 set (c) and Uncertain set (d)

	Predicted		
	Toxic	Non-toxic	Uncertain
(*a*) *Training set*			
Measured			
Toxic	2400	211	725
Non-toxic	162	4044	756
(*b*) *Test1 set*			
Measured			
Toxic	556	72	195
Non-toxic	52	912	213
(*c*)*Test2 set*			
Measured			
Toxic	553	78	192
Non-toxic	59	926	192
(*d*)*Uncertain set*			
Measured Uncertain	2886	10112	4704

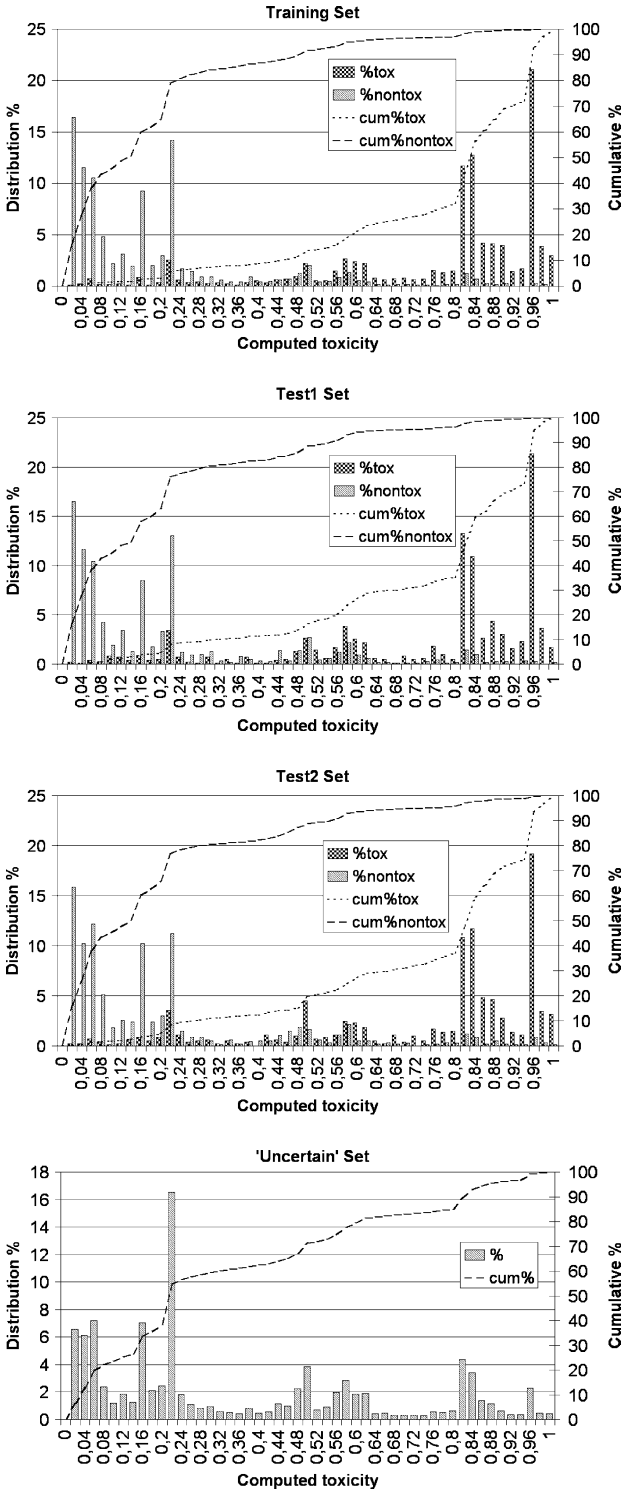

Figure 1. Distribution (left ordinate axis) and cumulative percentage (right ordinate axis) of computed toxicity (0 non-toxic, 1 toxic) for the toxic and non-toxic subsets of Training set, Test1 set, Test2 set and the Uncertain set.

Based on the viability data measured at 60 μM the most and least cytotoxic compounds (viability $\geqslant 1$ and $\leqslant 0.1$, respectively) have been selected. The resulted 12,298 compounds were divided into three sets: the largest set, containing 8298 molecules (3336 toxic and 4962 non-toxic), was used as training set, while the remaining 4000 molecules were divided evenly into two sets, containing 2000 members of each, namely Test1 and Test2 (each including 823 toxic and 1177 non-toxic compounds). The 17,002 compounds that were not included in the sets mentioned above have been categorized as 'uncertain' and have been used as an external validation set.

The three-layer feed-forward neural network was designed to contain 164 input neurons, 13 hidden neurons and one output neuron. The experimental viability values were changed to be 0.1 for non-toxic compounds and 0.9 for toxic compounds. The training of the neural network was carried out by the insertion of the 164 Atomic7 descriptors of the Training set onto the input layer and the scaled experimental viability values onto the output neuron of the neural network. All neural network operations were carried out using Stuttgart Neural Network Simulator (SNNS)[12] by standard backpropagation algorithm. Training set elements were mixed in each cycle to present them to the net in different orders. The training involved 5000 cycles with a learning rate of 0.2, which assured sufficient convergence in test runs. The Test1 set was used to monitor the quality of generalization ability of the neural network at each learning cycle. The best-trained neural network (which has the lowest sum of square errors /SSE/ on Test1 set) was saved. At the end of the training, the Test2 set was used as an external validation set.

Based on the results, the output range has been divided into 3 ranges: toxic (>0.75), non-toxic (<0.25), and uncertain ($\geqslant 0.25$ and $\leqslant 0.75$). Using these categories the accuracy is 77.6%, 73.4% and 73.4% for Training set, Test1 set and Test2 set, respectively (Table 1).

Finally, we used an additional external validation set containing the 'uncertain' compounds. The results show (Fig. 1) that the developed model can be safely applied for filtering out potentially cytotoxic candidates, because it tends to categorize the uncertain compounds rather to the non-toxic class, avoiding undesired exclusion of potential leads.

In conclusion, viability data obtained for a large and diverse set of compounds enabled us to develop a powerful classification method for the discrimination of cytotoxic and non-cytotoxic compounds. The developed algorithm is in daily use in library design projects of ComGenex and going to be commercialised as a module of the newest generation of the Pallas software family.

References and notes

1. Duprat, A. F.; Huynh, T.; Dreyfus, G. *J. Chem. Inf. Comput. Sci.* **1998**, *4*, 586.
2. Tetko, I. V.; Tanchuk, V. Y.; Villa, A. E. *J. Chem. Inf. Comput. Sci.* **2001**, *5*, 1407.
3. Wegner, J. K.; Zell, A. *J. Chem. Inf. Comput. Sci.* **2003**, *3*, 1077.
4. The PrologP program is a Trademark of CompuDrug, Inc., www.compudrug.com.
5. Ghose, A. K.; Crippen, G. M. *J. Comp. Chem.* **1986**, *7*, 565.
6. Molnár, L.; Keserű, G. M.; Papp, Á.; Gulyás, Zs.; Darvas, F. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 851.
7. Bugelski, P. J.; Atif, U.; Molton, S. *Pharm. Res.* **2000**, *17*, 1265.
8. Pagé, B.; Pagé, M.; Noël, C. *Int. J. Oncol.* **1993**, *3*, 473.
9. Ahmed, S. A.; Gogal, R. M.; Walsh, J. E. *J. Immun. Methods* **1994**, *170*, 211.
10. MRC-5 human foetal lung fibroblast cells ordered from European Collection of Cell Cultures (ECACC, Salisbury, UK). DMEM, FBS, PS, Trypsin–EDTA, L-glutamine and DMSO were Sigma products. Triton X-100 was purchased from Merck. alamarBlue was ordered from Biosource International. 384-well plates ordered from NUNC. We used an updated Beckman Biomek 2000 workstation for liquid handing with 200 μl tips from MBP. Fluorescence measurements were performed by a VICTOR2 1420 Multilabel Counter from PerkinElmer.
11. Zhang, J.; Chung, T. D. Y.; Oldenburg, K. R. *J. Biomol. Screen.* **1999**, *4*, 67.
12. SNNS: Stuttgart Neural Network Simulator, University of Stuttgart, 1995.