In silico screening of inventories of high public concern. Mutagenicity predictions for HPVC OECD chemicals

Milen Todorov

Abstract: Risk assessment is a process that characterizes the magnitude of risk that chemicals pose to human and environmental health. Recent findings have suggested that certain chemicals have properties that make them harmful to human health or the environment which may require their restriction or even prohibition. One of the endpoints of major regulatory concern is mutagenicity defined as ability of chemicals to cause damages on DNA. Considering the large number of the registered industrial chemicals and taking in mind the cost and the time needed for testing it is evident that they could not be experimentally evaluated for possible mutagenic effect. However, the task could be accomplished by application of alternative in silico methods such quantitative structure–activity relationships (QSARs). In the present study the mutagenic potential of the high production volume HPV OECD chemicals (4843) have been computationally evaluated by DNA profiling schemes incorporated in the computer platform QSAR Toolbox. It was found that 10% of the organic compounds could possess mutagenic effect.

Key words: *mutagenicity, in silico toxicology, risk assessment, QSAR, industrial chemicals*

INTRODUCTION

The increasing production and availability of chemicals on the global market, is of high concern today. In general, the main problem is that there is a lack of toxicological information for huge number of the currently registered chemicals. For example, it was officially stated that there is no toxicological information for 43% of the high production volume chemicals managed by the US EPA authority [1].

A number of potential toxic effects need to be dealt with in the prioritization of a chemical during its safety assessment. Among them mutagenicity is considered as one of the most important endpoints due to the fact that It possesses significant similarity of action with genotoxic mechanism of carcinogenicity [2].

The challenge of evaluating a large number of chemicals for their mutagenic potential can be handled at least to a certain degree by the ability to construct models based on structure–activity relationships (SARs). Such models have great potential for use in the identification and classification of large number of potential chemical mutagens. At the very least, these models could be employed to establish a prioritization procedure for subsequent in vitro/in vivo testing.

Several commercial programs equipped with modules for mutagenicity predictions are available, probably the best known of which are OASIS TIMES [3], DEREK for Windows (DfW) [4], CASETOX [5], TOPKAT [6] and Leadscope model applier [7]. These are generally regarded as “expert systems” since they were developed using a non-congeneric set of chemicals encompassing a number of different biological mechanisms. The performance of each system has been exhaustively investigated and evaluated [8].

Beside mentioned commercial programs one should point out the growing use of free or open type in silico platforms for predictions of variety biological endpoints [9, 10, and 11]. Currently, a larger part of them are accepted and used in many companies, organizations and national authorities for in silico predictions for endpoints of interest including mutagenicity.

The present article examines the performance of built-in profilers for DNA damages in the most popular free in silico system – QSAR Toolbox applied for screening of OECD HPV inventory as integrated database in the system.

EXPERIMENTAL

OECD HPV database

The database (4843 chemicals) has been compiled based upon submissions from member countries including the European Union's high production volume (HPV) chemical list according to EC Regulation 793/93 [12]. This database include all chemicals reported to be produced or imported at levels greater than 1000 tons per year in at least one Member country or in the EU region. One of the strategic goals related to this collection is constant addition of toxicological data for each chemical which will allow ultimate evaluation of the whole toxicological profile of the chemicals in the list.

OECD QSAR Toolbox

The Toolbox [9] is a PC- or server-based expert system that incorporates the OECD guidance related to categorization, read-across, and QSAR models. It also incorporates a large number of data sets containing physical and chemical property data, molecular descriptors, mammalian and non mammalian toxicity test data, in vitro and high throughput data, and categorical and endpoint/mechanistic descriptors derived by variety organizations for thousands of chemicals. A GUI allows the user to enter or retrieve data on individual chemicals on a point-and-click basis; define category criteria; and conduct read-across, trend analyses or run QSAR models to fill data gaps for untested chemicals.

Another advantage of the system is the opportunity to investigate a chemical with account to its metabolic fate. It is well known that the chemical in its parent form may not exert toxic effect however after metabolism a reactive metabolite can be produced which may damage biological macromolecules. This became extremely important in assessment of mutagenic potential of various type of chemicals.

In the following two sections details will be given for current versions of the profilers associated with DNA damages and in vitro metabolic simulator incorporated in version 3.3 of the Toolbox.

Profiling schemes for DNA damages. OASIS and ISS

The profiler OASIS DNA v. 1.3 is based on Ames mutagenicity model part of OASIS TIMES system [3]. The profiler contains exact definitions of 78 structural alerts responsible for interaction of chemicals with DNA. The scope of this profiler is to investigate the presence of alerts within the target molecules responsible for interaction with DNA, especially related to Ames mutagenicity.

The second ISS v.2.3profiler contains a list of 30 structural alerts (SAs). The SAs for mutagenicity are molecular functional groups or substructures known to be linked to the mutagenic activity of chemicals. As one or more SAs embedded in a molecular structure are recognized, the system flags the potential mutagenicity of the chemical.

In vitro metabolism simulator

The current in vitro rat liver metabolic simulator represents electronically designed set of 509 structurally generalized, hierarchically arranged biotransformation reactions, which are characteristic for the metabolism for in vitro experimental systems such as rodent (mostly rat) liver microsomes and S9 fraction. A training set of 647 xenobiotic chemicals of a wide structural diversity, with experimentally observed metabolic reactions and pathways has been built, using published data on their metabolism in rodent liver microsomes and S9 fraction. On the whole, the simulator contains 450 – 470 enzymatic phase I transformations, such as aliphatic C-oxidation, aromatic C-hydroxylation, oxidative N- and O-dealkylation, epoxidation, ester and amide hydrolysis, carbonyl group reduction, nitro and azo group reduction, N-hydroxylation, etc. Additionally, 15 – 20 enzymatic phase II transformations, such as glucuronidation, sulfation, glutathione conjugation, N-acetylation, etc. are included with significantly lower priority than phase I ones.

RESULTS AND DISCUSSION

The OECD high production volume (HPV) database was initially investigated in terms of available mutagenicity (Ames test) experimental data. This analysis was performed by retrieving mutagenicity data from a large collection of chemicals incorporated in the Toolbox. It was found that there is Ames data for 971 out of all 4843 (20%) chemicals. In order to assess the performance of both OASIS and ISS profilers – they were applied for profiling of all 971 compounds (120 mutagens and 851 non mutagens). The predictions results were analyzed to determine sensitivity and specificity. Sensitivity represents the Ames positive mutagens that were predicted to be positive whereas specificity reflects the Ames negative compounds predicted to be negative. The results are summarized in Table 1.

Table 1. Prediction results for Ames mutagenicity for 971 chemicals with experimental Ames mutagenicity based on: (a) parent structures and (b) with account for metabolism simulation.

	OASIS	ISS
Sensitivity	56%	66%
Specificity	93%	79%

(a)

	OASIS	ISS
Sensitivity	71%	84%
Specificity	76%	28%

(b)

As it can be seen the OASIS profiler performs well, especially in terms of specificity with and without metabolic activation. The obtained results showed that both profilers could not be applied in combination aiming to perform a kind of weight of evidence approach for ultimate mutagenicity prediction due to the high rate of false positives generated by the ISS profile. Therefore it was decided that the OASIS profiler should be used alone for prediction of the rest of 3872 OECD HPV chemicals without experimental Ames data.

As a result of preliminary structural screening it was found that 1008 out of 3872 chemical are inorganic or structures with unknown composition. Thus, the OASIS profiler has been applied for prediction of a total 2864 discrete organic compounds. The obtained result is presented in Table 2.

Table 2. Identified potential mutagenic compounds by OASIS DNA profiler.

OECD HPV chemicals without observed mutagenicity data	Identified DNA binding structural alerts by OASIS profiler	
	As parent structures	After metabolism simulation
2864	30	254

The application of OASIS profiler on parent structures identifies structural alert(s) related to mutagenic effect in 30 chemicals. In combination with metabolism simulator additional 254 chemicals were found to possess DNA binding alert in at least one of the generated metabolites. Taken together the predictions lead to identification of 284 (approx. 10%) potential mutagenic compounds among all discrete organic OECD HPV chemicals.

In order to be experimentally examined by standardized Ames test a total sum of 795,200$ will be required for all 284 chemicals (approx. 2800$ per chemical). However, as a result of additional prioritization the ultimate cost could be segregated in at least two

loops (tests with and without metabolic activation) of experimental evaluations. A good starting point could be additional analysis for identification of those structural alerts which bring positive mutagenic prediction for larger part of the investigated chemicals. In this respect it was found that 50% of the chemicals with DNA reactive structural alerts in parent structures are Epoxides or aziridines (8 chemicals) and Hydrazine derivatives (7 chemicals). Since there are other published QSAR models for these reactive groups they could be applied in a way to assess additionally the mutagenic potential by a kind of weight of evidence approach. Ultimately, only those chemicals predicted positive by the second model will be forwarded for experimental testing.

In the same way the chemicals which are predicted to be mutagens after metabolic activation could be also prioritized. For completeness of the study it was found that the largest part of the metabolites responsible for positive prediction contains again structural functionality - Epoxides or aziridines. However, other structural alerts such as quinones, quinone methides, N-hydroxylamine, dicarbonyl compounds, C-nitroso compounds also contribute for positive mutagenic predictions.

In addition it should be pointed out that there is a documented metabolism database incorporated in the QSAR Toolbox which is suitable for comparison of observed and simulated metabolic maps for chemicals which contain DNA reactive group in their metabolites. Such comparison could be used as a measure of the reliability of the predictions related to generation of specific reactive metabolite(s).

CONCLUSIONS

Nowadays, industry and commerce have greatly increased the number and levels of chemicals in the environment. However, the lack of complete toxicological information rises a lot of safety concerns. One of them points out the necessity for evaluation the mutagenic potential of the chemicals used widely in industrial processes. Considering the large number of the registered industrial chemicals and taking in mind the cost and the time needed for testing it is evident that they could not be experimentally evaluated for possible mutagenic effect. Thus, commercially important compounds identified as potentially high risk mutagens by computational (in silico) methods would be further experimentally tested.

In the current study two profilers for prediction the mutagenic effect (DNA reactivity) were used for identification of mutagenic compounds among OECD HPV chemicals. It was found that 284 out of all 4843 chemicals could be mutagens. Given the ultimate expenses to cover all experimental tests a scheme for prioritization of potential mutagens is discussed. The results suggest that computational screening can be a rapid and inexpensive way to set priorities for further testing of OECD HPV chemicals that are associated with a high risk of being mutagens.

REFERENCES

[1] [http://www.epa.gov/HPV/pubs/general/hazchem.htm

[2] Hatch F.T., M.G Knize, D.H Moore, J.S Felton. Quantitative correlation of mutagenic and carcinogenic potencies for heterocyclic amines from cooked foods and additional aromatic amines Mutat. Res. Environ. Mutagen. Relat. Subj., 271 (1992), p. 269

[3] Mekenyan O., S. Dimitrov, T. Pavlov, Dimitrova G., Todorov M., Petkov P., Kotov S. Simulation of chemical metabolism for fate and hazard assessment. V. Mammalian hazard assessment. SAR and QSAR in Environmental Re-search (23) (5-6), (2012), p. 553-606.

[4]. Judson P., S. Stalford and J. Vessey. Assessing confidence in predictions made by knowledge-based systems Toxicology Research, (2), 1, (2013), p. 70-79.

[5]. Chakravarti, S.K., R.D. Saiakhov, and G. Klopman. Optimizing Predictive Performance of CASE Ultra Expert System Models Using the Applicability Domains of

Individual Toxicity Alerts. Journal of Chemical In-formation and Modeling, (52), (2012), p. 2609-2618.

[6]. Benigni R. Computational prediction of drug toxicity: the case of mutagenicity and carcinogenicity. Drug Discovery Today: Technologies, (1), 4, (2004), p. 457-463.

[7]. Valerio, L. G., K. P. Cross. Characterization and validation of an in silico toxicology model to predict the mutagenic potential of drug impurities. Toxicology and Applied Pharmacology, (2012), p. 209-221.

[8]. Cronin M., J. Jaworska, J. Walker, M. Comber, C. Watts and A. Worth. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. Environ. Health Persp. (111), (2003).

[9]http://www.oecd.org/chemicalsafety/risk-assessment/theoecdqsartoolbox.htm

[10] http://ambit.sourceforge.net

[11] http://toxtree.sourceforge.net

[12] http://webnet.oecd.org/hpv/ui/Default.aspx (last accessed 09.2015)

About the author:

Milen Todorov, PhD, Department of Inorganic Chemistry, Faculty of Natural Sciences, University "Prof. Assen Zlatarov", 1 Prof. Yakimov str., Burgas, Bulgaria, tel. 056/716 491, e-mail: mtodorov@btu.bg

This paper has been reviewed