Food and Chemical Toxicology 106 (2017) 595–599

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

# In silico prediction of genotoxicity

Jörg D. Wichard

*Bayer Pharma AG, Genetic Toxicology, Müllerstr. 178, D-13353, Berlin, Germany*

ARTICLE INFO

*Article history:*
Received 22 September 2016
Received in revised form
25 November 2016
Accepted 10 December 2016
Available online 12 December 2016

*Keywords:*
In silico methods
QSAR
Expert systems
Genotoxicity
Genotoxic impurities

ABSTRACT

The *in silico* prediction of genotoxicity has made considerable progress during the last years. The main driver for the pharmaceutical industry is the ICH M7 guideline about the assessment of DNA reactive impurities. An important component of this guideline is the use of *in silico* models as an alternative approach to experimental testing. The *in silico* prediction of genotoxicity provides an established and accepted method that defines the first step in the assessment of DNA reactive impurities. This was made possible by the growing amount of reliable Ames screening data, the attempts to understand the activity pathways and the subsequent development of computer-based prediction systems. This paper gives an overview of how the *in silico* prediction of genotoxicity is performed under the ICH M7 guideline.



## 1. Introduction

In pre-clinical development of pharmaceuticals, studies of the drug's toxicity need to be performed including testing for genotoxicity (ICH M3 (R2), ICH S2 (R1)). According to the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) Guideline S2 (R1), a bacterial reverse mutation assay, typically the Ames test, has to be performed for every new drug submission. Beside testing the new drug/active ingredient itself, impurities have to be identified, quantified, and evaluated as well. Impurities in pharmaceuticals can result from synthesis, e.g. degradation or by products, manufacturing, formulation or environmental factors such as temperature, humidity and light. In summary, the Ames test is central to the genotoxicity assessment of a drug substance and its impurities.

The ability to predict the outcome of the Ames test based on well-understood structure-activity relationships is the key to minimize the risk of drug attrition due to genotoxicity at later stages of the drug discovery process. In this context, the application of *in silico* tools for the evaluation of genotoxicity comes into play, in particular when very limited information on impurities is available.

There is a number of *in silico* tools/models available that are mostly based on the *in vitro* data that was gathered over the years. These *in silico* tools could be divided into statistical QSAR based approaches and expert systems, also known as knowledge based systems. The main driver in this field is the recent ICH M7 guideline "Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk" that provides the framework for assessing the DNA reactive potential of impurities in pharmaceutical products (ICHM7). The ICH M7 was intended to complement the to the ICH guidelines for industry Q3A (R2) that covers impurities in new drug substances and Q3B (R2) that deals with impurities in new drug products (ICHQuality).

The assessment of the DNA reactive potential of chemicals was made possible by the growing amount of reliable Ames screening data, the attempts to understand the activity pathways and the subsequent development of computer-based prediction systems demonstrating high negative predictivity (i.e., true Ames negatives among all negative predictions) on various sets of impurity data (>85%; Sutter et al., 2013).

A key aspect of the ICH M7 recommendations is the proposed use of (Q)SAR predictions as a substitute for an experimental Ames test. In detail, the ICH M7 guideline says: "Structure-based assessments are useful for predicting bacterial mutagenicity outcomes based upon the established knowledge. There are a variety of approaches to conduct this evaluation, including a review of the available literature and/or computational toxicology assessment."

And further: "A computational toxicology assessment should be performed using Quantitative Structure-Activity Relationship ((Q)SAR) methodologies that predict the outcome of a bacterial mutagenicity assay […]. Two (Q)SAR prediction methodologies that complement each other should be applied. One methodology

*E-mail address:* joerg.wichard@bayer.com.

http://dx.doi.org/10.1016/j.fct.2016.12.013
0278-6915/

should be expert rule-based, and the second methodology should be statistical-based."

With respect to expert knowledge the guideline says: "If warranted, the outcome of any computer system-based analysis can be reviewed with the use of expert knowledge in order to provide additional supportive evidence on relevance of any positive, negative, conflicting, or inconclusive prediction and to provide a rationale to support the final conclusion."

The aim of this paper is to give an overview of how the *in silico* prediction of genotoxicity is performed under the ICH M7 guideline.

## 2. Expert systems

The first expert systems were developed in computer science in the late 1970s and became popular in the early 1980s. In general, an expert systems is a computer based system that simulates the judgement and behavior of a human who has expert knowledge and experience in a particular field. Therefore it consists of two major components. A knowledge base that contains the accumulated experience necessary for understanding, formulating and solving the problems and an inference engine that includes the basic rules for the decision making process. A unique feature of an expert system is the capability to explain its own reasoning and to justify the decision making process. The advantage of this approach is that it can give a comprehensive output that can be understood, challenged and judged by the user.

There are several expert systems for *in silico* prediction of genotoxicity. One of the first and most famous is the commercial, off-the-shelf DEREK software (Marchant et al., 2008). In DEREK (an acronym for Deductive Estimation of Risk from Existing Knowledge) rules or facts are stored in a knowledge base and retrieved by an inference engine which considers the strength and direction of each assertion through a process known as reasoning (Judson and Vessey, 2003). This approach leads to different reasoning levels that reach from 'plausible' over 'equivocal' to 'negative' and could be judged by the user. The knowledge base is defined by Lhasa scientists based on publicly available data and data that was contributed by Lhasa Consortium members. It is updated on a regular basis and distributed among the consortium members.

Further commercial expert systems are Multicase (Multicase, 2016) or Leadscope Expert Alerts (Leadscope). Several rule based systems like ToxTree (ToxTree) are included in the OECD QSAR toolbox (QSAR Toolbox, 2016) which is freely available.

## 3. Statistical systems

In contrast to expert systems that are driven by expert rules, statistical QSAR systems are driven by the data that is provided to the system. Following the ICH M7 the user should utilize models that "predict the outcome of a bacterial mutagenicity assay …" and the models should follow the general validation principles set forth by the Organisation for Economic Co-operation and Development (OECD) (OECD, 2007). These principles state that the statistical QSAR model should have a defined endpoint, an unambiguous algorithm, a defined domain of applicability, appropriate measures of goodness-of-fit and (if possible) a mechanistic interpretation. Given the data set of a bacterial mutagenicity assay and an appropriate set of molecular descriptors that describe the features of the chemical structures, the statistical QSAR models identifies relationships between chemical structure and biological activity. A comprehensive overview is given by Varnek and Baskin (2012). Some examples for commercial systems are Leadscope Genetic Toxicity Suite (Leadscope), Multicase (Multicase, 2016), SARAH Nexus (SARAH, 2016) and TOPKAT (Topkat, 2016). A recent comparison of freely available QSAR models for predicting Ames genotoxicity is given in Cassano et al. (2014).

One of the key features of statistical systems for the use under the ICH M7 guideline is transparency, i.e. they should be amenable to expert judgment. It is not sufficient to provide a "black box system" that has no option for any reinterpretation or analysis of the prediction results, since statistical systems might find correlations that are coincident in nature rather than causally linked structure activity relations.

Despite the differences in these approaches, expert systems and statistical systems could be considered to be "two sides of the same coin". The rules that are included in the expert systems are based on statistical evidence. Conversely, the selection of molecular descriptors, models and data set are guided by an expert (see Table 1).

### 3.1. Data bases

The first step in assessing chemicals with respect to their genotoxic potential is a database and literature search for carcinogenicity and bacterial mutagenicity data. For impurities, these data are used for assignment to impurity classes 1, 2, or 5 as shown in Fig. 1a and Fig. 1b and to derive a compound-specific limit (see Addendum to ICH M7). In the pharmaceutical industry, a search covers proprietary in-house data as well as open access, commercial or shared data. A comprehensive list of databases and data services including a detailed description is given in Amberg et al., (2016) and Sutter et al., (2013). A short list of databases dealing with carcinogenicity and mutagenicity data includes the following ones:

- NTP: The US National Toxicology Program (NTP)
- TOXNET: US National Library of Medicine (NLM)
- IARC: The International Agency for Research on cancer (IARC)
- CPDB: The Carcinogenic Potency Database (CPDB)
- CCRIS: Chemical carcinogens, data covering the period 1985–2011
- VITIC: Commercial DB from published and unpublished sources.

Since it is impractical to search different sources individually, one might use a commercial database that contains up-to-date information from public sources and that is frequently updated by the vendor.

## 4. The use of two *in silico* systems under ICH M7

The ICH M7 guideline requires two complementary methodologies for genotoxicity assessments. The application of two complementary methodologies is predicated on the assumption that greater sensitivity in detecting potential mutagens will be achieved if a positive prediction from either methodology leads to an overall positive conclusion. This gain of sensitivity is achieved at the expense of specificity, which may be justified with a focus on patients' (or consumers') safety. However, development costs and the time needed to bring a new drug to the market may be unnecessary increased. Therefore one has to deal with the question, how to perform the assessment and how to deal with conflicting prediction results. The key to using two *in silico* systems under ICH M7 is to integrate supporting information, also referred to as expert knowledge, into the overall conclusion. This may comprise a database search, a detailed review of structural fragments determining the prediction. The critique is often raised that expert judgment is heavily biased by the human expert (Powley, 2015). Many researchers investigated so far the impact of expert knowledge in order to improve the expected outcome of *in silico* assessments. Dobo et al. ran a cross industry survey among eight