The ICH M7 Guideline highlights that both safety and quality risk management are required to establish suitable levels of mutagenic impurities (MI) that will be expected to pose negligible carcinogenic risk in man. It recommends assessment and subsequent control strategies for MIs that may be carried over into the final active pharmaceutical ingredient (API) or drug product, taking into consideration the intended human use. The risk assessment involves the use of literature searches for carcinogenicity or mutagenicity in order to classify potential MIs (PMIs) or real MIs as class one, two or three or in silico predictions of mutagenicity in order to classify PMIs or MIs as class three, four or five (see Table 1, page 24).

Therefore, the major challenge facing Industry was how to perform genotoxic risk assessments (GRAs) and how to use quantitative structure analysis (QSAR) tools to predict PMIs or MIs. ICH M7 highlights that computational toxicology assessments should be undertaken using QSAR methodologies that will be highly predictive of the outcome.
of a bacterial mutagenicity assay. The guidance directs that two QSAR methodologies that are complementary in nature should be used. One approach uses an expert rule-based system (i.e., Derek Nexus, Leadscope Genetox Expert Alerts®) and the second (Q)SAR approach is statistically-based (i.e., Sarah Nexus®, Leadscope Genetox Statistical QSAR, Case ULTRA®, etc.). It is worth looking in depth at two representative computational tools from both of these classes: Derek Nexus and Leadscope Genetox Statistical QSAR system.

**Expert rule-based system**

Historically, Derek Nexus® has played a fundamental role in structure-based assessments of mutagenicity. Derek Nexus uses expert knowledge rules generated using both public and proprietary mutagenicity data, and applies those rules to facilitate in silico predictions about the mutagenicity of synthetic intermediates, impurities, degradants, as well as potential metabolites.

Proprietary information provided by Lhasa consortium members has been utilised in approximately one quarter of the bacterial in vitro mutagenicity alerts found in Derek Nexus®. Additionally, proprietary information is used to validate these alerts to demonstrate the predictive outcome of this tool within the chemical space of most interest to users (often proprietary in nature). These data donations can take several different forms:

1) Consortium members make data sets designed to either improve or validate existing alerts available to the rest of the consortium. This is then further analysed to identify lead compounds for alert development. Additionally, data sharing collaborations, such as Vitis Nexus® intermediates and Consortium for the Investigation of Genotoxicity of Aromatic Amines (CIGAA® have also been used for alert development.

2) Consortium members contact Lhasa to highlight that an existing alert is not sufficiently predictive within their proprietary chemical space, and they make available the data to support an alert modification. In some instances, this includes dialogues with Lhasa QSAR scientists that can lead to additional focussed testing within a specific class to either fill those data gaps or resolve contradictory outcomes within a data set.

In both these cases, the objective is to translate proprietary data held by a specific member into generic, publically-available structure-activity relationships. The level of disclosure is entirely negotiable. Some consortium members prefer to remain unidentified; for example, alert 760 for 2-halopyridines states that: ‘four compounds donated by a Lhasa member which have all been reported non-mutagenic’ whilst other consortium members disclose their identity; for example alert 475 for 3-aminomethyl-1,2,4-oxadiazoles states that ‘This alert is based on data from Hoffmann-La Roche AG and describes the Ames test activity observed for a series of 3-aminomethyl-1,2,4-oxadiazole compounds’.

This collaborative effort within the consortium improves the predictivity within the chemical space that is most important to member companies. As a result, Derek Nexus has been shown to exhibit higher sensitivity within the proprietary chemical space.

**Statistically-based (Q)SAR systems**

Leadscope uses a feature-based QSAR approach utilising molecular descriptors which include structural features and calculated properties (molecular weight, aLogP, polar surface area, hydrogen bond acceptors, hydrogen bond donors, number of rotational bonds and Lipinski score). The models use partial logistic regressions to encode the relationship between these descriptors and the bacterial mutagenesis endpoint. When making a prediction, the same structural features and properties held in the model are calculated for the query compound. These descriptors are then weighted and used to calculate the probability of a positive result. The applicability domain is also evaluated by measuring in vitro results and the level of disclosures made by the leadscope member companies.

### Table 1: Impurity classification, based on mutagenicity or carcinogenicity and proposed control strategy (as per ICH M7)

<table>
<thead>
<tr>
<th>Classification</th>
<th>Classification Definition</th>
<th>Control Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Known mutagenic carcinogen</td>
<td>Control at or below acceptable limit (TTC or LTL** threshold)</td>
</tr>
<tr>
<td>2</td>
<td>Known mutagens with unknown carcinogenicity potential</td>
<td>Control at or below acceptable limit (TTC or LTL** threshold)</td>
</tr>
<tr>
<td>3</td>
<td>Alerting structure that is unrelated to the structure of API in the absence of mutagenicity data</td>
<td>Either control at or below acceptable limit (TTC or LTL** threshold) or conduct Ames test; if positive classify as Class 2; if negative classify as Class 5</td>
</tr>
<tr>
<td>4</td>
<td>Alerting structure that is related to the structure of API and where the API is non-mutagenic</td>
<td>Treat as an ICH Q3A or Q3B non-mutagenic impurity</td>
</tr>
<tr>
<td>5</td>
<td>No alerting structures or an alerting structure with inadequate data to demonstrate a lack of mutagenicity or carcinogenicity</td>
<td>Treat as an ICH Q3A or Q3B non-mutagenic impurity</td>
</tr>
</tbody>
</table>

*TTC: Threshold of Toxicological Concern, **LTL: Less than Lifetime
the distance between the query compound and similar compounds in the training set. Compounds that are found to be outside the predefined applicability domain will not return a prediction as they are considered out of domain (OOD).

Leadscope has introduced a knowledge-sharing collaboration to address any specific QSAR regulatory issues that have been identified via discussions with interested parties or regulatory agencies. Knowledge sharing can either be with the aim of improving the model’s predictivity for structurally similar compounds or more usually to improve the performance of an entire alert and/or the QSAR model; e.g., to improve the applicability domain and prediction accuracy of specific classes of compounds that are not well represented in the public domain, such as boronic acids. The extent of data disclosure is assessed on a case by case basis that is dependent on the sensitivity of the proprietary information provided. Full data disclosure, where both the structure and supporting data are incorporated into the QSAR statistical models, helps to maintain the model’s transparency. However, where required, compound/company confidentiality can be maintained.

For example, the specificity for the predictivity of primary aromatic amines was increased by 14% with no commensurate decrease in sensitivity, as a result of such data sharing with a single pharmaceutical company. This was achieved by the use of structural fingerprints for different compound classes. Nearly 600 chemical fingerprints based on the Leadscope fragment hierarchy were used. Subsequent data analysis of a reference set and existing external knowledge containing a variety of primary aromatic amines allowed an updated alert to be derived. This list of sub-structures included meta-, para-, ortho- and hetero-substituted, as well as more complex substitution patterns. Only those results for the pre-defined sub-structures in the fingerprint are summarised and it is therefore possible to apply these fingerprints to a proprietary database without revealing sensitive information for individual compounds or supporting data. This collaboration now includes 13 different pharmaceutical companies and regulatory agencies and has led to a continued improvement in the predictivity of this class of compounds. This fingerprint methodology is also being applied to various other chemical classes.

The use of two (Q)SAR systems

A transnational inter-company survey of eight companies showed an increased usage of (Q)SAR-based evaluations. The authors indicated that many publications have historically focused on the potential of these QSAR models either individually or in combination, to predict accurately the mutagenic effects of certain proprietary chemicals in the Ames assay. Typically, these assessments evaluate significant numbers of these compounds together with QSAR-based approaches to try and predict their likely mutagenic liability. However, these publications did not address the associated expert assessments that are also necessary to interpret any alert(s). The authors were concerned that structural assessments alone were being used to conclude that a given impurity was non-mutagenic. This perspective was assessed using an inter-company survey to understand companies’ individual success rates for accurately predicting mutagenicity. This survey established that the Negative Predictive Value (NPV) of these QSAR in silico approaches were 94%. However, when ‘expert knowledge’ was added to this process, the NPV was increased to 99%. The authors commented on the significance of expert interpretation of in silico predictions and they indicated that the use of multiple in silico models (as required by ICH M7) is not a significant factor in the outcome of these evaluations, with respect to NPV.

Another trans-national survey from 14 different companies focussed on the different QSAR methodologies used within the pharmaceutical industry, together with the predictive value of these different QSAR approaches. The authors indicated that most pharmaceutical companies used an expert rule-based system, e.g., Derek Nexus as their standard default methodology, and these approaches yielded negative predictivity values of ≥78% across all participating companies. A further augmentation in predictive power...
IN-DEPTH FOCUS: INFORMATICS

(>90%) was typically achieved by the use of a second usually statistically-based\(\text{TM}\) QSAR methodology, or an additional expert review. However, when two QSAR methods were used in tandem, an additional expert review was still required, since conflicting outcomes were often obtained.

There are five potential outcomes from any assessment using two in silico systems:\(\text{TM}\):

1. Positive/Positive. It would require very strong evidence during the expert review to dismiss both positive predictions (likely positive).
2. Positive/Equivocal or OOD. The lack of a prediction (equivocal or OOD) from the second system implies insufficient evidence to derive any other conclusion during the expert review (likely positive).
3. Positive/Negative. Very strong evidence from the expert review will be required to overturn the positive prediction (likely positive).
4. Negative/Equivocal or Out of Domain (OOD). The expert review may conclude this is a negative outcome where there is strong evidence showing that the structural motif responsible for the no prediction is also present in known negative examples (assuming there are no deactivating features). Some companies may assign this a positive outcome as negative is non-proven (likely uncertain).
5. Negative/Negative. The expert review needs to assess any concerning features, i.e. unclassified, misclassified, etc. (likely negative)

Based on the available data, it was concluded that an expert rule-based system augmented by either a second statistically based QSAR system or expert knowledge is an acceptable strategy. Transparency of the expert review process, i.e., methodologies, results, weight-of-evidence approaches, etc., were achieved by collaborative initiatives involving full data sharing. It was hoped that in the future, the use of standard reporting strategies would enable regulators to more fully understand the submitted data. The authors concluded that this approach was appropriate for regulatory submissions aligned with ICH M7\(\text{TM}\).

Conclusion

Computational, in silico tools that correlate chemical structure with bacterial mutagenicity play an important role in the risk assessment of reactive intermediates, impurities and breakdown products of drug substances and products. These methods are either expert rule-based systems or are statistically derived. However, irrespective of whether these QSAR methodologies are used singly or in tandem, an expert review is still required. Indeed, it has been reported\(\text{TM,}\) that if these ‘complementary’ (QSAR systems are used in tandem (as required by ICH M7) the expert review is more involved and takes significantly longer, because in addition to positive and negative outcomes, there are uncertain results where the two systems give conflicting conclusions. In addition, it has been reported that the use of multiple in silico models is not a significant factor in the outcome of these evaluations, with respect to negative predictive power\(\text{TM}\).

Acknowledgements

The author would like to acknowledge the support provided by Dr Angela White (GSK).

References


Dr Elder has BSc and MSc degrees in chemistry from Newcastle upon Tyne, before he moved to Edinburgh to study for a PhD in Crystallography. He is a visiting professor (King’s College, London). Dr Elder has 37 years’ experience at a variety of different pharmaceutical companies (Sterling, Syntex and GSK). He is currently a Director within the product development group in GSK R&D. Dr Elder is a member of the British Pharmacopoeia (Expert Advisory Group PCY: Pharmacy), a council member of the Analytical Division, Royal Society of Chemistry (RSC), UK and a council member of the Joint Pharmaceutical Analysis Group, UK. He is a fellow of the RSC (FRS) and a member of the Royal Pharmaceutical Society (SRPharmS). He has co-edited one book on the Analytical Characterisation and Separation of Oligonucleotides and their Impurities.

Dr Elder has BSc and MSc degrees in chemistry from Newcastle upon Tyne, before he moved to Edinburgh to study for a PhD in Crystallography. He is a visiting professor (King’s College, London). Dr Elder has 37 years’ experience at a variety of different pharmaceutical companies (Sterling, Syntex and GSK). He is currently a Director within the product development group in GSK R&D. Dr Elder is a member of the British Pharmacopoeia (Expert Advisory Group PCY: Pharmacy), a council member of the Analytical Division, Royal Society of Chemistry (RSC), UK and a council member of the Joint Pharmaceutical Analysis Group, UK. He is a fellow of the RSC (FRS) and a member of the Royal Pharmaceutical Society (SRPharmS). He has co-edited one book on the Analytical Characterisation and Separation of Oligonucleotides and their Impurities.