InSilicoTox Project

Deliverable 1 (D1)

Report on Reactivity Measurement and Estimation

November 2008

European Union Sixth Framework Programme Marie Curie Actions
Project No.: **42328**
Project Acronym: **InSilicoTox**
Project Full Name: ***In Silico*** **Tools to Reduce the Use of Animals in Toxicity Testing for Bioreactive Chemicals**

Deliverable Prepared by

Dr Johannes Schwöbel, Mark Cronin, Dr Yana Koleva, Dr Judith Madden,
School of Pharmacy and Chemistry
Liverpool John Moores University
Byrom Street
Liverpool
L3 3AF
England

Important Information

InSilicoTox – Deliverable 1

D1. Report on Reactivity Measurement and Estimation

Introduction

In the framework of the new European Union (EU) regulation Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), risk assessment of industrial chemicals is a very important issue in the upcoming decade.[1] REACH calls for persuasive data of produced and used chemicals to assure human health and to provide environmental risk assessment. Toxicological effects are, for example, the irritation potential, skin sensitisation, mutagenicity, acute and chronic toxicity, reproductive toxicity and carcinogenicity.[2] To reduce animal testing in the context of requested toxicological and ecotoxicological endpoints, there is an increasing demand to introduce alternative testing strategies. These alternative testing strategies include *in vitro*, *in chemico* and *in silico* methods.

It is a common perception, that there can be a link between the toxicity of a compound and the specific chemical reactivity of the compound.[3] An indication of chemical reactivity can be provided by different *in chemico* and *in silico* assays, which will be represented in this report.

Dealing with reactivity in connection with toxicity, one should consider different interaction types.[3] Weak reversible interactions, including hydrophobic van der Waals' interactions, hydrogen bonding and ionic bonds (e.g. by charged amino acids), are of great biological relevance, as they govern the interaction between receptors and the substrate. Reversible, non-covalent interaction with membrane components is called narcosis. Secondly, the hydrophobicity

of a compound is principally responsible for the bioavailability at a specific target site.

The formation of coordination bonds between metals or their salts and heteroatoms, such as nitrogen or oxygen, is a broad field in the disquisition of the allergenic potential of metals.

Covalent bonds between a substrate and a target molecule are formed by reactions between electron-rich nucleophiles and electron-poor electrophiles. Electron-rich groups usually contain nitrogen, oxygen, phosphorus or sulphur, especially in nucleic acids and proteins. Depending on which target site is attacked by the electrophilic molecules, this can lead to mutagenic effects, toxic effects or sensitisation effects.

The subject here is irreversible, reactive chemistry. Although one has to deal with a broad field of different mechanisms and biochemical pathways, the initiating steps often comprise a reactive, covalent binding between a small electrophile – the xenobiotic molecule – and an endogenous nucleophile.[4] This might be a target site at a peptide, a protein or an enzyme, the DNA molecule or related compounds. Besides covalent binding, oxidative or reductive modifications of the target sites are possible as well.[3]

Chemical Reactivity Assays

A very early relationship between toxicity and reactivity was published by Landsteiner *et al.* (1936).[5] This workgroup found a good correlation between the skin sensitisation potential of guinea pigs by a series of 20 benzene derivatives with halogen or nitro substituents and their reaction with aniline as a reference nucleophile. The chemicals, dissolved in a mixture of absolute ethanol and aniline, were heated in a steam bath for two hours and the liberation of halogen was measured. If the chemical did not possess labile halogens, the solution was stirred in the steam bath for 15 hours and the possible formation of the substitution compounds was inspected. If a chemical reaction was observed under these conditions, the chemical acted as a sensitiser in every case, otherwise not. The workgroup measured the velocity

of decomposition of the same compounds with methylate and ethylate as well, with a good qualitative correlation (95%) between high velocity constants and observed sensitisation potential in the animal tests.

For the reaction between an electrophile and a particular nucleophile, reactivity can be quantified in terms of the rate constant k (or log k).[3] Instead of absolute rate constants, relative rate constants k_{rel} (or log k_{rel}) can compare a set of electrophilic chemicals among each other. Rate constants are defined for a particular temperature, however, the Arrhenius equation can be used to correct for different experimental temperatures.
The RC_{50} value (or pRC_{50}) is the concentration of electrophile, which gives a defined half-life t for the nucleophile. The electrophile concentration has to be constant (or in a large excess) throughout the reaction by definition. The RC_{50} value is inversely related to k, but independent of the initial nucleophile concentration:

$$k = \frac{\ln 2}{t \cdot RC_{50}}$$

One should keep in mind that extrapolated RC_{50} values should not exceed water solubility (or solubility in the particular reaction medium).
Another method to quantify reactivity is using percent depletion (%DP) after a fixed time t. If an excess initial electrophile concentration $[El]_0$ is used and the depletion of a nucleophile measured, the following expression can be used:

$$k = \frac{\ln\left(\frac{100}{100 - \%DP}\right)}{t \cdot [El]_0}$$

For example, Gerberick *et al.* (2004) and Natsch *et al.* (2007) use an approach based on this method.[6-7] One great disadvantage of this method is, when the percent depletion is close to 0 or close to 100, small errors in the depletion value lead to large errors in k. It is less accurate than a kinetics experiment, since only two data points are used.

Note, that reaction rates can only be compared in the context of the same mechanistic domain. This was not the case in the Gerberick assay, published in 2004 with 38 chemicals and updated in 2007, comprising 82 chemicals of different skin sensitisation potencies.[6,8]

Qualitative reactivity data can be obtained by the application of mass spectrometry (MS) and nuclear magnetic resonance (NMR) techniques:[3] With these sensitive and accurate methods it can be observed if any product formation takes place or not. Furthermore, they shed light on the possible mechanism. MS yields specific mass fragments of ionised products, which can be identified. NMR measures chemical shifts of nuclei (^{1}H, ^{13}C, ^{15}N, ^{19}F or ^{31}P), which depend on their electronic environment and shall be different for products and reactants.

Chromatographic methods, such as HPLC-UV, help to separate and to identify compounds in a quantitative manner:[3] The concentration in a solution can be related to the absorbance of conjugated organic compounds (by UV-VIS detection) or the fluorescence intensity (by fluorometric spectroscopy) at a specific wavelength.

One should keep in mind that reactivity measures do not reflect the specific chemical mechanism necessarily. Often it is not clear which aspect of the reaction might be relevant for potency: the selectivity (e.g. towards a special amino acid target), the reaction rate or the stability of the conjugates (e.g. of a hapten-protein conjugate). Nevertheless, reactivity assays have proven their reliability to predict appropriate toxicity over the past decades, even in a quantitative manner.

The tripeptide glutathione (GSH, *L*- -glutamyl-*L*-cysteinyl-glycine) is one of the most widely used nucleophilic reference molecules in reactivity assays. It is the most prevalent cellular thiol and the most abundant low molecular weight peptide in cells.[4] GSH protects cells by detoxifying electrophilic compounds and acts as an antioxidant. The concentration of GSH depletes during the attack by electrophilic compounds, commonly by alkylation. A high

GSH depletion rate makes other endogenous thiol groups susceptible to attack, especially soft cysteine –SH moieties. Therefore, if the GSH concentration falls below a critical level in a cell, this can cause accumulation of damage.[9] Straight GSH depletion is modelled by a pure chemical reactivity assay.

GSH is reactive towards soft electrophiles, for example , -unsaturated carbonyls, which act predominantly as Michael type acceptors. In general, soft molecules, which are readily polarisable and have a low electronegativity, are easily oxidised.

In kinetic studies, an excess amount of the test chemical and GSH are dissolved in dimethyl sulfoxide (DMSO) and phosphate buffer. After a determined reaction time, the concentration of free thiol groups is measured, as described below. Depending on the particular assay, the conditions diverge from each other: The range of the pH value is 7-10, the temperatures range from 20°C to 37°C and maximum reaction times ranging from 10 minutes to 24 hours are given in the literature.[6-11] Typical assays have been introduced by Gerberick *et al.* (2004), Schultz *et al.* (2005) and Natsch *et al.* (2007). Gerberick's depletion assay is a fast screening 15 min method at 25°C and pH=7.4 with high excess of the electrophile (1:100).[6] Natsch works at higher temperatures (30°C) to avoid precipitation and less electrophile excess (1:20), but 24h are needed for a similar amount of depletion.[7] Schultz measures RC_{50} values after 180 min at 25°C; pH=7.4 and different electrophile concentrations.[11]

Free thiol groups are usually quantified by UV-VIS spectroscopy at 412 nm absorption after reaction with the chromophore DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid), also known as Ellman reagent. Since some GSH might exist in the oxidised form GSSG with a disulfide bridge, GSH and GSSG can be separated by reverse phase HPLC columns and detected. Alternatively, GSSG can be transformed to GSH selectively using a mixture of the enzyme

glutathione reductase and TPNH, triphosphopyridine nucleotide, as described by Tietze (1969).[12]

Different publications show a good qualitative relationship between the GSH reactivity of chemicals acting as Michael type acceptors and biological endpoints (e.g. acute fish toxicity or skin sensitisation).[4,9] For example, the different toxicity of methacrylate, acrylate and crotonate can be explained by their different reactivity, although the structures are very similar: methacrylate is only little reactive, while acrylate is very reactive and crotonate lies somewhere in between.[13]

The assays can be applied as long as the expected mechanism of action includes a Michael addition, and there is no activation required. This holds for other mechanisms (like nucleophilic substitution, aromatic nucleophilic substitution, Schiff base formation or acylation for skin sensitisation[14]), involving other reference nucleophiles, as well.

Abiotic activation for *pre*-electrophiles can occur within the reaction medium, e.g. by air oxidation. Metabolic activation for *pro*-electrophiles needs separate consideration and is best modelled by special enzymatic or disparate *in vitro* assays.[3]

The choice of the reference nucleophile depends on the expected mechanism and the site of action:

Soft interactions, involving the thiol group –SH, can be modelled by small molecules, like mercaptopropionate[15] or propanethiolate[16], and peptides, like GSH (see above), cysteine, acetylcysteine or peptides with a cysteine residue.[7]

Nucleophiles containing primary amino groups $–NH_2$, such as in the amino acids lysine and arginine, can undergo Schiff base formation or acylation by way of example. Reference nucleophiles are aniline (see above, Landsteiner

et al.) and butylamine.[17-18] The reactivity of butylamine gets close to the reactivity of lysine, acetyllysine or peptides with a lysine residue.[6,19]

Histidine is a further amino acid with a secondary aromatic amino group –NH, which should be attacked by electrophiles in principal. However, no significant correlation has been observed between biological endpoints and the reactivity of histidine or peptides containing histidine so far.[6] A possible model nucleophile for histidine is imidazole.[18]

The –OH group of the phenolic amino acid tyrosine could be modelled by phenolates or alkylates. Furthermore, methalkylates or ethalkylates are used as model nucleophiles for nucleophilic substitution of activated aromatics.[20] Tyrosine has not been used in reactivity assays so far.

Synthetic peptides, mimicking active protein sites, such as Cys-420 of the human
coronin 1 C, or containing all key nucleophiles have been investigated.[3] Their purpose is to reproduce the spectrum of reactivity simultaneously. On the other hand, synthetic peptides can be very cost-intensive, which impairs their use in standard screening assays. One alternative might be to use purchasable proteins, like the widely-used human serum albumin (HSA), but in this case the technique to detect and to identify adducts is very time-consuming and complex (MALDI-TOF-MS and nano-ES-MS/MS) and so this is not useful for a general screening at the present time.

For this reason, Gerberick *et al.* (2007) and Natsch *et al.* (2007) suggested not to use one peptide assay alone, but a battery of assays, in order to achieve a holistic assessment of reactivity, related to different toxicological endpoints.[7-8] Gerberick utilised a depletion assay of a cysteine containing peptide and a lysine containing peptide for skin sensitisation. Both soft and hard interactions are covered by this model, and it is a good compromise between accuracy and time and effort. The model is able to distinguish between strong or moderate contact allergens on the one hand and weak

contact allergens or non-sensitisers on the other hand. The concordance is 89% for 82 compounds.

Reactivity of compounds, which can cause DNA damage, can be measured by model nucleic acids, like 2'-deoxyguanosine (GUA). This nucleic acid has been used to model aquatic fish toxicity so far.[21]

These assays make several assumptions:[9] The endogenous electrophile concentration remains constant (steady state kinetics) and is equal to the concentration in the reaction medium (here: DMSO). For example, aqueous solutions can be used, if the considered reactions take place in the cytosol. The assumption of constant concentration might not hold for liver, kidney or other tissues with a high clearance rate. However, relative reactivity should not be significantly affected by the reaction medium, if the transition state is similar. The assumption is that the response is instantaneous with a first-order endogenous consumption. Finally, a crucial point is that the reaction should be dominated by chemical reactivity, not enzymatic conjugation and biotransformation, such as glutathione transferase.

Estimating chemical reactivity

Predictors of chemical reactivity, obtained by *in silico* methods, can be divided into two classes: qualitative and quantitative reactivity estimates.[22]

Firstly, qualitative structure-activity relationship models (SAR) evaluate the presence or absence of a specific substructure in a molecule.[23] These are called structural alerts. They imply a specific mechanism is associated with a substructure. For this purpose, expert knowledge has to be assembled: The substructures should be specific enough for reliable prediction of a particular mechanistic pathway. If , -unsaturated carbonyls are taken as an example, Michael addition does not occur in case of an alkylated -carbon atom, which opposes the withdrawing effect of the carbonyl group.[13] Here other mechanisms are more likely, e.g. Schiff base formation.

Computational expert systems are – to name just a few examples – ChemProp, Derek for Windows, M-CASE, TOPKAT and TIMES-SS.[24-28] These have in common that molecular structures are entered into the program, and the likelihood of potential chemical toxicity is estimated based on knowledge rules. In ChemProp and Derek, these knowledge rules are based on validated mechanistic considerations, while M-CASE and TOPCAT use an automatically learning algorithm. TIMES-SS is a hybrid, based on a combination of both approaches.

The use of SMARTS (SMiles ARbitrary Target Specification) patterns has been introduced by Enoch *et al.* (2008) to distinguish between different mechanisms of action.[29] Structural features are given by a two dimensional string, the SMARTS code, thereafter a dataset can be searched through specific strings, which are associated with a particular mechanism. However, it is difficult to assign a chemical a single mechanism, and the authors suggest adding some weight of evidence to the results.

Secondly, quantitative reactivity estimates can be achieved using electronic, geometry or topological descriptors.[22] Empirical statistical approaches and so-called global models, which can associate a number of (often non-interpretable) reactivity descriptors with toxicity, can lack mechanistic understanding, which leads to a disputable prediction quality. Models, which combine for example bioavailability (log *P*) and reasonable electronic properties, are more trustworthy in this context.

Local, mechanism-based models deal with electronic properties, which are connected to the underlying reaction mechanism. Taft's and Hammett's coefficients predict substituent effects of reactive chemicals. Mostly, the electronic properties are based on quantum chemical calculations.[30] Highest occupied and lowest unoccupied molecular orbital (HOMO and LUMO) energies are broadly used as indicator variables of molecular electron donor or acceptor affinity, which should represent the electrophilicity or nucleophilicity of a compound to a certain extent.

There exist more site-specific counterparts to describe electron donor or acceptor affinities of a reactive atom or functional group (see ref. [31]). Local softness and the electrophilic Fukui function are able to describe nucleophilic activation.[32] To this end, Parr's electrophilicity index and its local counterpart has been identified as the best descriptor for the electrophilicity of a compound.[33] This descriptor was used for an electrophilic ranking of reactive compounds, which seems to be related to toxic effects (here: skin sensitisation potential p*EC*3 values) and to reactivity of unsaturated compounds towards nucleophilic additions.[34]

Quantum chemical calculations might provide a closer insight in chemical reactivity, these are an alternative to labour-intensive *in chemico* methods: Thermodynamic properties can be obtained by the calculation of electrophilic and model nucleophilic reactants and their products as well as kinetic data. The latter is more difficult, because the calculation of appropriate transition states is needed. Nevertheless, these types of calculation are often very time-consuming, and there is no established routine method available to predict chemical reactivity to this end.

Conclusions

The chemical reactivity of xenobiotic electrophiles towards nucleophilic reference compounds is a model for their biochemical behaviour towards bioactive sites, such as proteins or DNA. This leads to a promising way to make statements about their potency of toxicity.

This report summarised *in chemico* reactivity assays as well as *in silico* estimation methods.

In the past, various reactivity assays have proven their reliability to predict biochemical activity of a compound. Mostly a compound is able to undergo several chemical mechanisms. In order to get a holistic view about a particular compound, the results of different reactivity assays should be compared to each other. This includes the use of different reference nucleophiles (e.g. soft and hard ones) or different experimental conditions (e.g. change of pH range).

Computational estimation methods are a promising alternative to labour expensive experimental measurements. Based on a reliable database with experimental reactivity data, it is intended to model and estimate these data based on computer-aided methods.

References

1. Regulation (EC) No **1907/2006** concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC
2. Summary of different endpoints available at: http://osiris.ufz.de/.
3. Gerberick, F.G. et al. Chemical Reactivity Measurement and the Predictive Identification of Skin Sensitisers. *ALTA* **36** (2008) 215-242.
4. Aptula, A.O. et al. Non-enzymatic Glutathione Reactivity and in vitro Toxicity: A Non-Animal Approach to Skin Sensitization. Toxicology in Vitro **20** (2006) 239-247.
5. Landsteiner, K. et al. Studies on the Sensitization of Animals with Simple Chemical Compounds. II. *J. Exp.Med*. **64** (1936) 625-639.
6. Gerberick, F.G. et al. Development of a Peptide Reactivity Assay for Screening Contact Allergens. *Toxicological Sciences* **81** (2004) 332-343.
7. Natsch, A. et al. Utility and Limitations of a Peptide Reactivity Assay to Predict Fragrance Allergens in vitro. *Toxicology in Vitro* **21** (2007) 1220-1226.
8. Gerberick, F.G. et al. Quantification of Chemical Peptide Reactivity for Screening Contact Allergens: A Classification Tree Model Approach. *Toxicological Sciences* **97** (2007) 417-427.
9. Freidig, A.P. et al. Comparing the Potency of Chemicals with Multiple Modes of Action in Aquatic Toxicology: Acute Toxicity Due to Narcosis versus Reactive Toxicity of Acrylic Compounds*. Environ. Sci. Technol.* **33** (1999) 3038-3043.
10. Clarke, E.D. et al. Metabolism-Related Assays and Their Application to Agrochemical Research: Reactivity of Pesticides with Glutathione and Glutathione Transferases. *Pestic. Sci*. **54** (1998) 385-393.
11. Schultz, T.W. et al. Structure-Activity Relationships for Reactivity of Carbonyl Containing Compounds with Glutathione. *SAR QSAR Environ. Res.* **16** (2005) 313-322.
12. Tietze, F. Enzymic Method for Quantitative Determination of Nanogram Amounts of Total and Oxidized Glutathione. *Anal. Biochem*. **27** (1969) 502-522.
13. Yarbrough, J.W. et al. Abiotic Sulfhydryl Reactivity: A Predictor of Aquatic Toxicity for Carbonyl-Containing α,,β-Unsaturated Compounds. *Chem. Res. Toxicol*. **20** (2007) 558-562.
14. Aptula, A.O. et al. Skin Sensitization: Reaction Mechanistic Applicability Domains for Structure-Activity Relationships. *Chem. Res. Toxicol*. **18** (2005) 1420-1426.
15. Friedman, M. et al. Relative Nucleophilic Reactivities of Amino Groups and Mercaptide Ions in Addition Reactions with α,β-Unsaturated Compounds. *J. Am. Chem. Soc*. **87** (1965) 3672-3682.

16. Alvarez-Sanchez, R. et al. Covalent Binding of the ^{13}C-Labeled Skin Sensitizers 5-Chloro-2-methylisothiazol-3-one (MCI) and 2-Methylisothiazol-3-one (MI) to a Model Peptide and Glutathione. *Bioorg. Med. Chem. Lett.* **14** (2004) 365–368.
17. Roberts, D.W. Correlations between Skin Sensitization Potential and Chemical Reactivity for p-Nitrobenzyl Compounds. *Fd. Chem. Toxic*. **21** (1983) 811-813.
18. Franot, C. et al. Structure-Activity Relationships for Contact Allergenic Potential of γ,γ-Dimethyl- -butyrolactone Derivatives. *Chem. Res. Toxicol.* **7** (1994) 297-306.
19. Bolton, J.L. et al. Influence of Quinone Methide Reactivity on the Alkylation of Thiol and Amino Groups in Proteins: Studies Utilizing Amino Acid and Peptide Models. *Chem. Biol. Interact*. **107** (1997) 185-200.
20. Bunnett, J.F. et al. Aromatic Nucleophilic Substitution Reactions. *Chem. Rev*. **49** (1951) 273-412.
21. Harder, A. et al. Applicability and Limitation of QSARs for the Toxicity of Electrophilic Chemicals. Environ. Sci. Technol. **37** (2003) 4955-4961.
22. Roberts, D.W. et al. Chemical Reactivity Indices and Mechanism-based Read-across for Non-animal Based Assessment of Skin Sensitisation Potential. *J. Appl. Toxicol*. **28** (2008) 443-454
23. Schultz, T.W. et al. Verification of the Structural Alerts for Michael Acceptors*. Chem. Res. Toxicol*. **20** (2007) 1359–1363.
24. Schüürmann, G. Kühne, R. Kleint, F. Ebert, R.-U. Rothenbacher, C. Herth, P. A software system for automatic chemical property estimation from molecular structure. In: Chen, F. Schüürmann, G. (eds) *Quantitative Structure-Activity Relationships in Environmental Sciences. VII* SETAC Press, Pensacola FL, USA (1997) 93-114.
25. Derek for Windows, Lhasa Ltd., Leeds, UK (2005).
26. Klopman, G. et al. In-Silico Screening of High Production Volume Chemicals for Mutagenicity Using the MCASE QSAR Expert System. *SAR QSAR Environ. Res*. **14** (2003) 165-180.
27. Enslein, K. et al. Use of SAR in Computer-Assisted Prediction of Carcinogenicity and Mutagenicity of Chemicals by the Topkat Program. *Mutation Res*. **305** (1994) 47.
28. Patlewicz, G. et al. TIMES-SS – A Promising Tool for the Assessment of Skin Sensitization Hazard. A Characterization with Respect to the OECD Validation Principles for (Q)SARs and an External Evaluation for Predictivity. *Regul. Toxicol. Pharmacol*. **48** (2007) 225-239.
29. Enoch, S. et al. Identification of Mechanisms of Toxic Action for Skin Sensitisation Using a SMARTS Pattern Based Approach. *SAR QSAR Environ. Res.* **19** (2008) 555-578.
30. Schüürmann, G. Quantum Chemical Descriptors in Structure-Activity Relationships – Calculation, Interpretation, and Comparison of Methods. In: Cronin, Mark T.D. Livingstone, David (eds.) *Predicting Chemical Toxicity and Fate.* CRC Press LLC, Boca Raton FL, USA (2004) 85-149.
31. Schwöbel, J. et al. Modeling the H Bond Donor Strength of -OH, -NH, and -CH Sites by Local Molecular Parameters. *J. Comput. Chem.* (2008), in press.

32. Funtealba, P. et al. On the Condensed Fukui Function. *J. Chem. Phys.* **113** (2000) 2544-2551
33. Parr, R.P. et al. Electrophilicity Index. *J. Am. Chem. Soc*. **121** (1999) 1922-1924
34. Enoch, S. et al. Quantitative and Mechanistic Read Across for Predicting the Skin Sensitization Potential of Alkenes Acting via Michael Addition. *Chem. Res. Toxicol*. **21** (2008) 513-520.