# Measurement and Estimation of Electrophilic Reactivity for Predictive Toxicology

Johannes A. H. Schwöbel,[‡] Yana K. Koleva,[‡] Steven J. Enoch,[‡] Fania Bajot,[‡] Mark Hewitt,[‡] Judith C. Madden,[‡] David W. Roberts,[‡] Terry W. Schultz,[§] and Mark T. D. Cronin*,[‡]

[‡]School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, England

[§]College of Veterinary Medicine, Department of Comparative Medicine, The University of Tennessee, 2407 River Drive, Knoxville, Tennessee 37996-4543, United States

## CONTENTS

## 1. INTRODUCTION

There is an increasing interest in predicting the toxicological effects of chemicals from their structure and physicochemical properties. For new compounds, this will optimize the product development process by eliminating toxic compounds early. For existing compounds, these approaches enable the prioritization of potentially harmful compounds.[1] Key among these approaches is the efficient formation and use of toxicological databases, category formation and read-across, and (quantitative) structure−activity relationships (SARs and QSARs). Such computational, or in silico, approaches ultimately aim to decrease costs and reduce animal suffering for chemical risk assessment. These techniques have come to the fore recently due to mandates such as the categorization of the Canadian Domestic Substance List, the European Union's REACH and Cosmetics regulations, the Japanese Chemical Substance Control Law, as well as their continued use by the U.S. Environmental Protection Agency (EPA) and U.S. Food and Drug Administration (FDA).[2]

For in silico models to provide valid predictions of toxicity, they should, ideally, be based on a mechanistic framework. This requires knowledge of the mode or mechanism of action and the translation of the chemistry responsible for the toxicological effect into usable in silico models.[3] There are many modes and mechanisms of toxic action ranging from physical interference with cell membranes or skin, accumulation in cell membranes, through to receptor mediated effects. The focus of this Review, however, is to develop and report information for a number of toxic effects, which are brought about by so-called reactive mechanisms. In particular, these include the formation of covalent bonds between a reactive, electron-poor (electrophilic) substrate and a biological electron-rich (nucleophilic) target molecule, especially biological macromolecules such as nucleic acids and proteins.[4,5]

To build adequate models for end points and effects related to reactive mechanisms of toxic action, that is, those relating to electrophile−nucleophile covalent interactions, information is required on the likelihood and rate of reaction of a chemical with a particular biological nucleophile. Such information can be derived from the direct experimental assessment of reactivity, using what are termed in chemico assays. According to a European Centre for Validation of Alternative Methods (ECVAM) Expert Workshop, the term "in chemico" refers to the use of abiotic chemical reactivity methods as a replacement for animal (in vivo) assays; its use is analogous to that of in vitro (for cellular bioassays) and in silico (for computer predictions).[6] This definition has been broadened by Cronin et al.[4] to include not only experimental methods but also the possibility of capturing and applying in chemico information through a variety of computational approaches. Further, it was noted that in chemico information is unlikely to be a replacement (i.e., a one-to-one replacement such as may be expected with certain in vitro

assays), but is a source of information, which may provide a basis to understand the molecular initiating event (see section 2) of a toxicological action, along with other information. This may assist in making a rational conclusion regarding the toxicological hazard associated with a compound. Using this philosophy outlined by Cronin, careful and intelligent selection of compounds (sometimes referred to as intelligent testing strategies) for in chemico testing has enabled the exact definition of domains in terms of chemical structure for many reactive mechanisms.[7,8] Such in chemico testing can be performed in a series of assays with different peptides and other reference nucleophiles. This type of testing has historically been undertaken in basic chemistry research; however, no attempt has been made to compile and interpret these historical data. The objective of this research culminating in this Review, therefore, was to compile, interpret, and publish a database containing relevant reactivity data from the past 80 years. The database is provided as an openly available resource to assist in all areas of the prediction of reactive toxicity.

## 2. REACTIVITY IN TOXICOLOGY

The initiating steps for many mechanisms of toxicological action comprise the reactive, covalent binding between a (small) electrophile, the xenobiotic molecule, and an endogenous nucleophile.[9,10] The target sites for the electrophiles include peptides, proteins or enzymes, lipids, DNA, or related biological molecules. Of these, the formation of covalent adducts with proteins and DNA is well-established.[11] Therefore, to assist in the prediction of toxicity from structure for effects initiated by protein or DNA binding, toxicants may be categorized and classified on the basis of molecular reactivity.

The categorization of molecules by reactive mode or mode of toxic action provides a strong basis for in silico approaches.[3] Alternatives to animal testing based on mechanism or mode of action and an adverse outcome pathway provide the user with confidence in the prediction. In this context, the mode of action relates to a uniform biological or toxic response caused by exposure to a compound.[12,13] The term adverse outcome pathway denotes a set of chemical, biochemical, cellular, physiological, behavioral, etc., responses, which characterize the biological effects cascade resulting from a particular molecular initiating event. A mechanism of action denotes the molecular sequence of events leading from the absorption of an effective dose of a compound to the production of a specific biological response in the target organ or organism.[14−19] This term should not be confused with the (organic) chemical reaction mechanism, a subset of biological mechanisms where the molecular initiating event (MIE)[20] is the covalent binding between a toxicant and a biological target molecule. As noted in the next section, if this reaction does not occur, other mechanisms of action, pathways, or modes of toxic action may be present so it should not be taken to imply the absence of toxicity.

One important issue in the covalent binding of xenobiotics with biological macromolecules is the metabolism of nonreactive compounds to reactive species ("pro-reactive" compounds).[21] Metabolic activation is usually captured by binding studies utilizing liver microsomal fractions or other enzymatic and in vitro assays.[22,23] The classic Ames test is an efficient in vitro assay to screen for the carcinogenic potential of drugs and related compounds, which incorporates metabolic transformation.[24] In addition to metabolism, abiotic (nonbiological) transformation, for example, via air oxidation of a "prereactive" compound to a reactive species, is an important route of activation for chemicals. For instance, recent work has demonstrated this capability for poly-hydroxylated and -amino aromatic compounds: depending on the substitution pattern, such compounds can be air-oxidized into a reactive quinone or quinone imine.[25,26] It should also be remembered that, even if chemical reactivity is observed, organisms can respond (and be protected) via repair mechanisms, cell proliferation, and immune activity.[22] In addition to covalent binding, oxidative or reductive modifications of the target sites on proteins, which lead to a toxic effect, are possible.[6] Biotransformation, which is beyond the scope of this Review, should also be borne in mind; it includes a variety of detoxifying mechanisms, so the resulting metabolites are generally less toxic and more water-soluble than the parent molecules.[27,28]
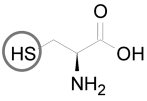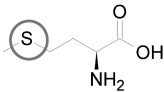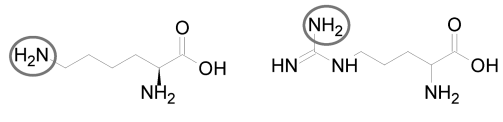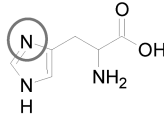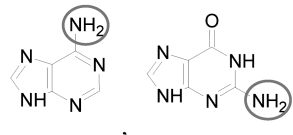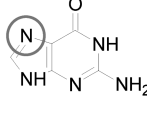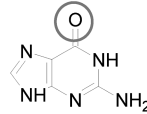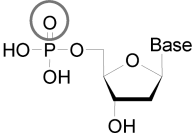
### 2.1. Defining the Spectrum of Electrophilic−Nucleophilic Reactions

It is well-known that not all electrophiles react with all potential biological nucleophiles; this is primarily because there are a number of different electrophilic reaction mechanisms. In addition, there are differing rates of reactivity within particular mechanisms of action.[6] It is, therefore, important to consider the selectivity of a compound toward a specific molecular site of action. This can be explained by the classification of electrophiles and nucleophiles according to their polarizability, in other words, the chemical "hardness" and "softness" of the electrophilic or nucleophilic center.[29,30] The principle of hard and soft acids and bases (HSAB) defined by Pearson[31] states that hard Lewis acids bind strongly to hard Lewis bases and soft Lewis acids bind strongly to soft Lewis bases. In this context, Lewis acids are electron acceptors that act as electrophiles, while Lewis bases are electron donors acting as nucleophiles.

Biological nucleophiles are predominantly hard, with their effects being brought about by their relatively hard oxygen or nitrogen containing moieties.[32] The softness of a nucleophile is associated with its ability to be easily oxidised; an example is the thiol residues of peptides and proteins with a sulfur atom. As the majority of reactive exogenous compounds (whose interactions result in toxicity) are soft electrophiles, the degree of softness assists in understanding their effects. Dissimilar hardness leads to a higher potential energy barrier for the reaction between an electrophile and a nucleophile, thus making them less reactive.

Generally speaking, hard electrophiles will prefer to react with DNA as well as with amino groups of, for instance, lysine; soft electrophiles will react preferentially with thio groups, for instance, cysteine on proteins. As a basis to understand reactivity, the spectrum of electro(nucleo)philic reactions considered in this Review corresponds well with "classic" Lewis acid−base theory and to the HSAB principle.[32,33] Further distinctions between electrophiles and nucleophiles are provided using the spectrum of reactivity ranging from soft to hard electro(nucleo)philic reactions.[29,34] Considering the spectrum of reactivity in more details, hard electrophiles are small molecules with low polarizability. Their electron deficiency is localized as a positive electrostatic charge. In contrast, soft electrophiles are often larger molecules and highly polarized, with their electron deficiency spread over a larger molecular region. The (nucleophilic) targets for electrophiles in biological molecules typically include electron-rich heteroatoms such as sulfur, nitrogen, and oxygen.[33] Biological molecules, in particular, amino and nucleic acids, have many such nucleophilic sites. These represent the continuum of Lewis bases from hard nucleophiles (low polarizability) to soft nucleophiles (high polarizability). Table 1 lists the specific nucleophilic sites in amino and nucleic acids in order of increasing hardness.

2563

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

**Table 1. Amino- and Nucleic-Acid Nucleophilic Sites Each Ordered Relative to Their Increasing Hardness[10]**

| Rank | Sites | Structure |
|------|-------|-----------|
| | **Amino acid sites** | |
| 1 | thiol-group of cysteine | |
| 2 | S-atoms of methionine | |
| 3 | primary amino-groups (e.g., lysine, arginine) | |
| 4 | secondary amino-group of histidine | |
| | **Nucleic acid sites** | |
| 1 | primary amine groups of purine bases (e.g., arginine, guanine) | |
| 2 | in-ring N-atoms of purine and pyrimidine bases (e.g., N7 of guanine) | |
| 3 | O-atoms of purine and pyrimidine bases (e.g., O6 of guanine) | |
| 4 | phosphate O-atom (P=O) | |

There has been no systematic attempt to establish experimentally based quantitative scales for hardness or softness, applicable across the whole range of electrophiles and the whole range of nucleophiles. As far as the electrophiles are concerned, some generalizations can be drawn up, based on their potential chemical mechanism: Michael acceptors and $S_NAr$ compounds are soft, acyl transfer agents are hard, and $S_N2$ can range from very soft to very hard (these mechanisms will be explained in further detail in section 2.2).

While dissimilarity in electro(nucleo)philic hardness results in a higher potential energy barrier of reactions between electrophiles and nucleophiles (implying lower chemical reactivity), this concept of "like-reacts-with-like" is not absolute: There are noted examples, such as acrolein, where reactivity is ubiquitous. Moreover, strong electrophiles such as nitroso-compounds and isothiocyanates react with a spectrum of biological nucleophiles. Because of the ability of some electrophiles to be reactive with a number of nucleophilic centers, the reactivity of an electrophile is best quantified by a relative reactivity profile with respect to a number of model nucleophiles. Although different nucleophiles can differ very substantially in their absolute reactivity toward a given electrophile, their relative reactivities are usually well correlated over a range of electrophiles within the same mechanism of action, as shown in Table 2.

While there are occasional exceptions, the general principle of HSAB illustrated in Table 1 is widely applicable. This applicability

**Table 2. Kinetic Rate Constants of α,β-Unsaturated Compounds with Thiolates (−SH) and Amines (−NH₂)[74]**

| CH₂=CH− | $k$ (−SH) $\times 10^{-3}$ M$^{-1}$ s$^{-1}$ | $k$ (−NH₂) $\times 10^{-5}$ M$^{-1}$ s$^{-1}$ | $k$(−SH)/ $k$(−NH₂) | $\log k$(−SH)/ $k$(−NNH₂) |
|---|---|---|---|---|
| −CO₂Me | 11.0 | 76.0 | 145 | 2.16 |
| −CN | 2.70 | 20.4 | 132 | 2.12 |
| −CONH₂ | 0.46 | 2.60 | 178 | 2.25 |

**Table 3. Swain−Scott Parameters[34] for Some Nucleophiles, the Nucleophilicity Parameter $n$, and the Corresponding Susceptibility Parameter $s$ of Some Electrophiles[a]**

| nucleophile | $n$ | electrophile | $s$ |
|---|---|---|---|
| aniline | 4.49 | ethyl tosylate | 0.66 |
| n-butylamine | 5.13 | benzyl chloride | 0.87 |
| HS⁻ | 5.08 | methyl bromide | 1.00 |
| cysteine | 5.08 | epichlorohydrin | 0.93 |
| cysteine ethyl ester | 5.24 | | |
| serum albumin | 5.37 | | |
| HO⁻ | 5.94 | | |
| acetate ion | 2.89 | | |

[a] The parameter $n$ quantifies the nucleophilic reactivity of the nucleophile relative to water. The higher is the $n$ value, the more reactive is the nucleophile. The parameter $s$ quantifies the selectivity of the electrophile; the higher is the $s$ value, the greater the degree of preference shown by the electrophile towards nucleophiles with higher $n$ values.

means that relative reactivities determined with one model nucleophile may be used to estimate the relative reactivities of electrophiles toward unknown biological nucleophiles. In terms of relative chemical reactivity, thiol- or amino-based compounds such as propanethiolate, aniline, or n-butylamine are just as good a model nucleophile as cysteine, lysine, glutathione, or other amino acids and peptides with reactive side chains. The relative reactivities are also a consequence of the Swain−Scott relationship, which is explained below.

To quantify relative reactivity, Mayr and Patz constructed scales of nucleophilicity $N$ and electrophilicity $E$.[35,36] These can describe kinetic rate constants $k$ (at 20 °C in dichloromethane) of a large variety of polar organic reactions, with $s$ being a system specific parameter, dependent on the reference nucleophile:

$$\log k = s(N + E) \quad (1)$$

The nucleophilicity parameter ranges from −4 (weak) to +24 (strong nucleophile), and the electrophilicity parameter ranges from −12 (weak) to +6 (strong electrophile). To this end, the electrophilicity scale has been applied for benzhydryl cations only, but the nucleophilicity parameter allows for a quick comparison of a variety of reference nucleophiles (e.g., water, $N = 5.1$; n-butylamine, $N = 11.7$; 2-hydroxyethanethiol, $N = 15.6$).[37]

In the context of electrophile−nucleophile reactions, it is convenient to consider the HSAB principle against the background of the Swain−Scott relationship, which predates the HSAB concept.[34] Swain and Scott argued that nucleophiles could be assigned a nucleophilicity parameter $n$ and that a parameter $s$ could be assigned to electrophiles to quantify their susceptibility to changes in $n$. On that basis, the Swain−Scott relationship was originally written as:

$$\log(k_{E/N} / k_{E/\text{water}}) = ns \quad (2)$$

In this expression, $k_{E/N}$ is the rate constant for reaction of the electrophile E with the nucleophile N, and $k_{E/\text{water}}$ is the rate constant for reaction of the electrophile E with water. Water may be regarded as the reference nucleophile, $k_{E/\text{water}}$ quantifying the intrinsic reactivity of the electrophile. Taking methyl bromide as the reference electrophile and assigning an $s$ value of 1, Swain and Scott measured rate constants at 25 °C and from these derived $n$ values for a series of nucleophiles and $s$ values for a series of electrophiles. Some $n$ and $s$ values for various nucleophiles and electrophiles are shown in Table 3.

A parallel set of $n$ and $s$ values, based on methanol as the reference nucleophile and methyl iodide as the reference electrophile, has also been established for reactions in methanol. It has been pointed out by Loechler that, toward the hard end of the hard−soft spectrum, deficiencies in the Swain−Scott equation become apparent.[38] Thus, the methyl diazonium ion, a hard methylating agent, is more selective for oxygen-centered nucleophiles of DNA than for nitrogen-centered nucleophiles, although with less hard electrophiles the nitrogen-centered nucleophiles are the more reactive. Loechler argued that $n$ values of nucelophiles are not completely constant, but may vary according to the electrophile. In particular, when the nucleophile and electrophile are well matched on the hard−soft spectrum, the nucleophile may exhibit an enhanced $n$ value.

Despite this and other deficiencies, the Swain−Scott equation provides a useful framework for considering the issues involved in trying to model an in vivo nucleophile, whose precise identity may not be known. Exchanging the in vivo nucleophile for a model nucleophile in chemico, the key questions are: if the relative reactivity of two electrophiles (E₁ and E₂, having $s$ values $s_1$ and $s_2$, respectively) is determined with a model nucleophile ($N_m$), is this relative reactivity the same for the in vivo nucleophile ($N_{iv}$)? If not, what is the difference, and how can the difference be minimized?

These questions can be addressed by rearranging the Swain−Scott equation to derive:
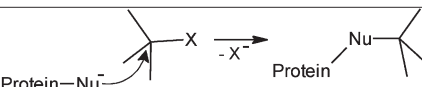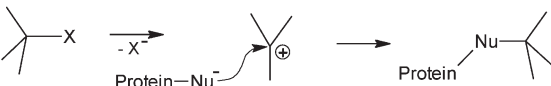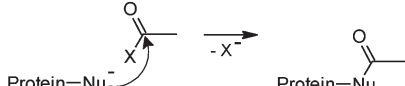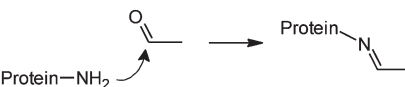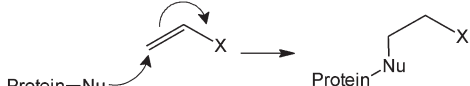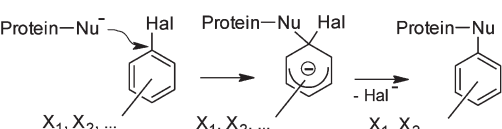
$$\text{relative reactivity in vivo}: \log(k_{N_{iv}/E1} / k_{N_{iv}/E2})$$
$$= n_{iv}(s_1 - s_2) + \log(k_{\text{water}/E1} / k_{\text{water}/E2}) \quad (3)$$

$$\text{relative model reactivity}: \log(k_{N_m/E1} / k_{N_m/E2})$$
$$= n_m(s_1 - s_2) + \log(k_{\text{water}/E1} / k_{\text{water}/E2}) \quad (4)$$

For the two expressions to be identical, it is necessary that either $s_1 = s_2$ or $n_{iv} = n_m$. If this condition is not met, there is a difference of $(n_{iv} - n_m)(s_1 - s_2)$. This difference, being dependent on the identities of the electrophiles, is not constant, so it will tend to weaken a predictive approach (i.e., a QSAR for an effect elicited by reactivity such as excess acute toxicity) in which reactivity is based on a model nucleophile. Depending how much is known or can be inferred about the in vivo nucleophile, selecting a model nucleophile with a similar $n$ value (if the $n$ value of the in vivo nucleophile is not known, this can be approached by aiming to match position both on the hard−soft scale and on the p$K_a$ value) will minimize $(n_{iv} - n_m)$ and hence reduce $(n_{iv} - n_m)$ $(s_1 - s_2)$. Working within the same reaction mechanistic domain,

**Table 4. Most Important Direct Acting Covalent Binding Mechanisms (Nu = Nucleophilic Site of a Protein)**

| Mechanism | Protein Binding Reaction |
| --- | --- |
| $S_N2$ |  X = leaving group (e.g. halogen) |
| $S_N1$ |  X = leaving group (e.g. halogen) |
| Acylation |  X = leaving group (e.g. halogen) |
| Schiff Base Formation |  |
| Michael Addition |  X = electron withdrawing group (e.g. -CHO, -COR) |
| $S_NAr$ |  Hal = (pseudo) halogen X = electron withdrawing groups (e.g. -NO2, -CN) |

or within subdomains (categories), will minimize differences in *s* values, and hence reduce $(n_{iv} - n_m)(s_1 - s_2)$.

Thus, there are a variety of methods available to assist in defining the spectrum of reactivity between electrophiles and nucleophiles and hence relate that to interactions at the biological level. This can be translated into practical tools and approaches to understand the molecular initiating steps that result in toxic effects.

## 2.2. Electrophilic Chemical Reaction Mechanisms

The concept of electrophilic reactivity is recognized as an important issue in predictive toxicology.[39−41] In its broadest sense, reactivity with biological molecules includes a spectrum of conjugation, substitution, and addition reactions,[10] which have their foundation in the principles of organic reactions where an electron-rich compound interacts with an electron-deficient one.[42] The importance of comprehending this process from a predictive standpoint is that the mechanistic capacity of electrophilic reactivity, leading to a particular toxicological end point, provides a means of grouping compounds into categories for that particular end point based on mechanistic similarity.[43]

In contrast to receptor-mediated chemical interactions, electrophiles are not specific in regards to their molecular targets. To illustrate this, Wong and Liebler, in their examination of mitochondrial proteins from cells treated with two different electrophiles, observed that adducts were formed with more than 800 proteins.[44] In addition, toxicants such as acrolein are able to react with several different nucleophilic targets. Therefore, mechanistic understanding is important for the appropriate regulation of industrial compounds.[3,4,45]

The toxicity of many electrophiles is related to their intrinsic capability to bind to proteins. The six most important direct acting covalent binding mechanisms are listed in Table 4, which may result, for example, in skin sensitization or elevated acute aquatic toxicity above baseline narcosis.[41] About 50 specific reactive protein binding mechanisms are known,[46] but the six organic chemistry mechanisms shown capture the general mechanisms of direct acting protein binding in the field of toxicology. These allow for a classification of electrophiles into appropriate mechanistic applicability domains, that is, chemical space associated with individual mechanisms. Further details on the six main mechanisms are provided below.

Leaving groups bonded to primary alkyl groups easily undergo nucleophilic attack via the second-order nucleophilic substitution ($S_N2$) mechanism. These are mainly halides, sulfonic acids,

**Table 5. Some Examples of Direct Acting DNA Alerts and Their Associated Mechanistic Domains**

| Functional group | Example structure | Mechanistic domain |
|---|---|---|
| $\alpha,\beta$-Unsaturated carbonyls | | Michael addition |
| Quinones | | Michael addition |
| Epoxides | | $S_N2$ |
| Aliphatic halogen | | $S_N2$ |
| Aliphatic aldehydes | | Schiff base |
| Acyl halide | | Acylation |
| Isocyanates | | Acylation |

sulfates, or epoxides, the latter activated by ring strain release in the transition state. Leaving groups located at secondary positions in a ring structure can also be effective; good examples include the alkane sultones. As a consequence of the $S_N2$ mechanism, inversion takes place at the reactive site. In general, $S_N2$ electrophiles span a wide range of the hard–soft spectrum.

In contrast to the $S_N2$ mechanism, the nucleophile does not participate in the initiating step of the first-order nucleophilic substitution ($S_N1$) mechanism. Here, an intermediate planar carbenium ion, which is independent from the original nature of the leaving group, is the highly reactive species attacked by the nucleophile.[47] Leaving groups at secondary or tertiary alkyl groups are likely to act via the $S_N1$ mechanism, where the carbenium ion is stabilized, as in allylic and benzylic halides.

Acylating agents form a reactive tetrahedral intermediate with the nucleophile and finally the ionic leaving group is released. This reaction only takes place if the corresponding acid of the leaving group is sufficiently acidic; therefore, simple alkyl esters are not acylating electrophiles. Phenyl esters, particularly when they have electronegative ring substituents, react as acylating electrophiles, as do carboxylic anhydrides and acyl halides.

Schiff-base formers are aliphatic aldehydes, a small number of aromatic aldehydes, and ketones in the neighborhood of electron-withdrawing groups. The reaction site can be seen as hard; this prefers to interact with harder nucleophiles (e.g., the lysine of a protein). The reaction mechanism is nucleophilic addition forming a hemiaminal, followed by imine generation. Cyanates, isothiocyanates, and all thio analogous compounds are able to react in a very similar way.

Polarized $\alpha,\beta$-unsaturated compounds can react via Michael-type nucleophilic addition with soft and basic electrophiles. This type of addition is usually considered a base-catalyzed two-step process.[47] After formation of the reactive ionic nucleophilic residue, the electrophilic $\beta$-carbon atom is attacked, leading to a nominally carbanionic species; it is only "enolic" when the activating group is –CHO or –COR. Proton abstraction from a protonated base is the final step. The lower is the electron density at the $\beta$-carbon, the greater is the likelihood of undergoing nucleophilic attack, especially by soft thiol groups. In addition, hard nucleophilic methoxide ions are reported to be highly reactive toward Michael acceptor compounds also due to their high basicity.[41]

$S_NAr$ electrophiles require two (or more) activating groups in the position ortho- or para- to the leaving group, due to convenient stabilization of the transition state, and thus can be easily identified. Reactivity is dependent on the ability of the activating groups, by a combination of inductive and mesomeric effects, to stabilize the intermediate anion after attack of the nucleophile, and on the ability of the leaving groups to stabilize this intermediate by an inductive effect. These effects can be modeled by the sum of the $\sigma^-$ of the activating groups and $\sigma^*$ of the leaving group, respectively;[48] these Hammett–Taft coefficients $\sigma^-$ and $\sigma^*$ are quantitative indices for inductive and mesomeric substituent effects and will be explained in section 4.3.

Many electrophiles contain more than one reactive group, which makes application of general rules regarding their mechanism of reactivity very difficult. For example, $\alpha,\beta$-unsaturated aldehydes might potentially react via the Michael acceptor or Schiff base reaction, and either mechanism can predominate. Sometimes the most reactive mechanism may be obvious from inspection of the structure (in case of $\alpha,\beta$-unsaturated aldehydes, the mechanism is dependent on the substituents at the $\beta$-carbon); in other cases, experiments with model nucleophiles and investigation of the products are required to make the decision.[41] This lack of high selectivity for protein adducts may give rise to multiple and seemingly conflicting results (i.e., a particular compound being both mutagen and sensitizing agent) observed with some electrophiles.

A number of direct acting covalent mechanisms are involved in genotoxicity including those also known to be involved in protein binding mediated toxicity. In contrast to the soft thiol electrophile that is predominantly involved in protein binding mechanisms, genotoxicity involves reactions with the harder nitrogen electrophiles contained within the DNA (or RNA) molecule. An additional complicating factor lies in the wide range of metabolic conversions that can result in non-electrophile compounds being converted into electrophiles. Such mechanisms are out of the scope of this Review (several literature sources cover these mechanisms[49–52]), which will focus on direct, non-metabolically activated mechanisms.

A recent compilation of genotoxic structural alerts, a set of structural features related to a molecular initiating event (see section 4.2), can be used to highlight the significant overlap in the mechanistic domains between the covalent reactions involving a soft thiol electrophile and those with a harder nitrogen electrophile.[52–55] An analysis of this compilation shows that these direct acting DNA alerts can be assigned to four of the six mechanistic domains suggested for protein binding.[41] The four domains are Michael addition, $S_N2$, Schiff base formation, and acylation (example alerts and associated domains are given in Table 5). In contrast, no alerts for direct acting DNA binding were suggested for the $S_NAr$ domain.

The lack of alerts within the $S_NAr$ domain can be explained by the fact that the electron-withdrawing activating groups (that are required to stabilize the intermediate in the $S_NAr$ reaction) are frequently metabolized in the liver, thus resulting in alternate genotoxic mechanisms. For example, 2,4-dinitrochlorobenzene (DNCB) has been previously assigned to the $S_NAr$ domain for protein binding in skin sensitization studies.[56] However, it is unlikely to exhibit its genotoxicity via the same mechanism as the nitro groups are readily metabolized into the significantly more reactive nitrenium ions, resulting in covalent bond formation via an $S_N1$ mechanism.[57] In keeping with the protein binding studies for skin sensitization, mutagenicity studies without the presence

2567

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

**Table 6. Some Examples for Electrophilic Reactions in Toxicity**

| toxicity | reaction site | main reaction mechanisms[a] | ref |
|---|---|---|---|
| skin sensitization | chemically modified skin proteins (e.g., Cys, Lys, or Ser residues) leading to T-cell-mediated allergic response | protein haptenation via $S_N2$, $S_NAr$, MA, SB, Ac | 41 |
| respiratory sensitization | chemically modified proteins in the lung (e.g., Lys residues) | SB, protein cross-linking, $S_N1$, $S_N2$, Ac | 179 |
| skin irritation | skin proteins and lipids in the stratum corneum | SB, $S_N2$, MA, Ac, $A_N$[b] | 69 |
| elevated acute toxicity and cytotoxicity (aquatic or terrestrial) | cellular GSH; interaction with nucleophiles ($-OH$, $-NH_2$, $-SH$ groups) in biological macromolecules (e.g., inhibition of acetylcholine esterase) | electrophilic reactivity via $S_N1$, $S_N2$, acylation, MA, SB (in contrast to polar and unpolar narcosis) | 67,186,187 |
| mutagenicity and carcinogenicity | DNA or RNA gene mutation via adduct formation, base pair substitutions, and frameshifts; interaction with regulatory molecules | $S_N1$, $S_N2$, acylation, MA, SB | 148 |
| chromosomal aberration | alteration of DNA and sequence of genetic material (number or structure of chromosomes), which often alters embryonic development; inhibition of topoisomerases and interaction with nuclear proteins associated with DNA (e.g., histone proteins) | DNA and protein binding mechanisms | 51,188,189 |
| hepatotoxicity | attack of hepatocytes, the bile duct, or sinusoidal endothelium, Kupffer, or Ito cells by: (1) direct cell stress, direct mitochondrial impairment, and specific immune reactions (2) direct and death receptor-mediated pathways leading to mitochondrial permeability transition (3) apoptosis and necrosis | protein binding and receptor-mediated mechanisms (e.g., interaction with P-450 enzyme family, leading to damaged mitochondrial functions and possible idiosyncratic effects) | 190,191 |

[a] Organic reaction mechanism: MA = Michael addition, SB = Schiff base formation, Ac = acylation, $S_N1/S_N2$ = first- or second-order nucleophilic substitution, $S_NAr$ = aromatic nucleophilic substitution, $A_N$ = nucleophilic addition. [b] Possible electrophilic mechanisms according to structural alerts of the German Federal Institute for Risk Assessment (BfR).[192]

of the metabolizing S9 fraction have shown DNCB to be mutagenic,[58,59] presumably by an $S_NAr$ mechanism. Given this, compounds with electron-withdrawing activating groups capable of undergoing $S_NAr$ reactions are likely to be genotoxic.

The above analysis has demonstrated that the six mechanistic domains that have been suggested to be responsible for protein binding-mediated toxicity are also important for direct acting genotoxicity. It is clear that if a compound falls into one of these domains then it should potentially be considered as genotoxic. However, it is important to realize that the mechanistic applicability domains are not always consistent between protein binding and DNA binding. For example, simple $\alpha,\beta$-unsaturated aldehydes are considered capable of both protein and DNA binding.[41,52] In principle, Michael acceptors that are genotoxic can act directly.[60,61] However, there is evidence that some polarized alkenes are metabolically activated by conversion to their epoxides, which are hard electrophiles and react with the DNA.[62−64] In those cases, the epoxide reaction products are probably more thermodynamically stable than the direct Michael addition products. (Note that epoxides are skin sensitizers and thiol reactive, as well.) Closer inspection of the mechanistic applicability domains reveals that cinnamic aldehydes are only capable of protein binding and do not bind to DNA.[65,66] This example of the difference in the mechanistic applicability domains highlights the need to consider all the available data for a given reactive toxicological end point and that one needs to be cautious when extrapolating mechanistic chemistry from protein to DNA binding.

## 2.3. Relating Electrophilic Chemistry to Toxicological End Points

The ultimate purpose of collecting and applying information regarding electrophilic reactivity is to assist in the prediction of toxicity from chemical structure or data from in chemico assays. As an illustration, recently there have been improvements in our understanding of the reactions involved between xenobiotics and proteins that result in immunogenicity and allergic contact dermatitis (skin sensitization)[26,56] as well as for the covalent reaction with fish gill membranes,[67] which will promote acute toxicity above narcosis. Many of these mechanisms of action are electrophilic in nature. There are many further examples of electrophilic toxicity, for example, respiratory sensitization, liver toxicity, skin irritation, etc.[68,69] Some examples for the relationship between toxicology and electrophilic reactivity are shown in Table 6.

The clustering of compounds into groups that are toxicologically relevant is based on the hypothesis that their properties, including electrophilic reactivity, are related to toxic effects and potency (see section 4.1). As noted by Borgert et al.,[14] this grouping of compounds typically presumes a common "toxic mechanism" or sequence of events leading from the absorption of an effective dose of a compound to the production of a specific biological response in the target organ. In the same way, the absence of a common mechanism of action also indicates that compounds are not likely to be members of the same toxicological group.

The recent success of the OECD QSAR Toolbox[46] in grouping compounds for assessment based on their intrinsic chemical reactivity, especially electrophilic reactivity, has spurred the quest to understand the molecular structure limitations for specific chemical reactions.[70] A compound may undergo an assortment of reactions with biological macromolecules. Schultz et al. described the decisive, and typically the first, reaction as the molecular initiating event (MIE).[5] As an example, for skin sensitization, the reaction of the xenobiotic with the protein rendering

immunogenicity is the MIE. This can lead to the formation of toxicologically meaningful groups based on reaction mechanisms.

The MIEs typically trigger changes in normal chemical and biochemical processes, which in turn lead to a progression of biological effects, some of which are critical. These effects can be defined in terms of their scientific significance to the pathway leading to an adverse outcome, such as mortality, that is of regulatory significance. When these critical effects are associated with the various levels of biological organization (cell, tissue, organ, etc.), an in vivo hazard end point can be traced back to one or more MIEs, hence the term pathway and, more specifically, adverse outcome pathway (see section 2). By focusing on a series of topics such as chemical reactivity, molecular sites of action, the MIE, the affected biochemical pathways, and the cellular- tissue- and organ-level responses, an outcome pathway leading to a specific in vivo outcome can be devised. Compounds that follow that pathway and elicit the same key responses along the pathway are likely to be in the same toxicological group.

Schultz has described a number of adverse outcome pathways for toxicological end points.[20] These include those for skin sensitization and respiratory irritation in acute fish mortality; in both of these examples of pathways, the seminal event is the electrophilic reactivity with cellular proteins. For that reason, to predict these and similar toxicities, measured or estimated electrophilic reactivity is vital. The structural domain of compounds associated with the MIE for these pathways can be defined in terms of their chemistry.

## 3. MEASURING CHEMICAL REACTIVITY

A systematic data compilation for intrinsic electrophilic reactivity would be of great use to allow for better development of categories and prediction of toxicity. In particular, there is a need to determine experimental reactivity for some or all of the following reasons:

(a) To determine if a compound is electrophilic in nature.
(b) To determine whether a compound is able to react covalently with a biologically important nucleophile and identify that nucleophile.
(c) To determine where a compound fits on the spectrum of hard to soft electrophiles.
(d) To assist in the definition of the structural boundaries associated with a mechanism of action.
(e) To determine relative potency of an electrophile.

For these five reasons, the measurement of the reactivity (e.g., in terms of reaction kinetics) of the compound with a reference nucleophile is required. The information derived from the reactivity measurements can also be used as a surrogate for toxicity testing itself; a compound that is reactive in the test tube may also have the capability of reactivity in vivo. A number of model nucleophiles have been utilized to investigate reactivity; some of the issues with using the data are described below.[6]

## 3.1. Scales of Reaction Kinetics

The extent and speed of reactivity of an electrophile and a particular nucleophile can be quantified in terms of the rate constant $k$ (or log $k$). Rather than using absolute rate constants, relative rate constants $k_{rel}$ (or log $k_{rel}$) can be used to compare the reactivity of a set of electrophilic compounds. Many assays use an excess amount of either the electrophile or the nucleophile, so pseudo-first-order kinetics are derived without significant excess compound depletion. Calculation of second-order rate constants can be derived according to Frost and Pearson.[71]

The pseudo-first-order rate constant $k_{1,pseudo}$ (stated in units of $min^{-1}$ in the database associated with this review) is quantified in

terms of the linear regression slope $(\ln c_{DP} - \ln c^0_{DP})$ versus the reaction time $t$. $c_{DP}(t)$ and $c^0_{DP}$ are the time-dependent concentration and initial concentration of the compound degraded.[72] The compound degraded is the nucleophile, if the electrophile is given in an excess amount and vice versa. A pseudo-first-order rate constant can be converted into a second-order rate constant $k_2$ ($M^{-1}$ $min^{-1}$) by dividing it by the initial concentration $c^0_{EX}$ of the compound in excess:

$$k_2 = \frac{k_{1,\,pseudo}}{c^0_{EX}} \quad (5)$$

If the depletion of the compound in excess cannot be neglected, the linear regression slope $k'$ of $(\ln c_{DP} - \ln c_{EX})$ versus $t$ can be used to calculate second-order rate constants:[73]

$$k_2 = \frac{k'}{(c^0_{DP} - c^0_{EX})} \quad (6)$$

Sometimes degradation half-lives $t_{1/2}$ of nucleophiles are given in the literature. Following pseudo-first-order kinetics, half-lives are related to rate constants as follows:

$$t_{1/2} = \frac{\ln 2}{k_{1,\,pseudo}} \quad (7)$$

According to second-order kinetics, this expression would be:

$$t_{1/2} = \frac{1}{k_2 \times (c^0_{EX} - c^0_{DP})} \times \ln\left(2 - \frac{c^0_{DP}}{c^0_{EX}}\right) \quad (8)$$

The $RC_{50}$ value is the concentration of electrophile, which gives a defined half-life $t_{1/2}$ for the nucleophile.[72] In other words, the time is fixed, and the concentration is changed so that one-half the reaction is complete at the fixed time. By definition, the electrophile concentration has to be constant (or in a large excess) throughout the reaction. Following the pseudo-first-order reaction equation, the $RC_{50}$ value is inversely related to $k_2$, but independent of the initial nucleophile concentration:

$$k_2 = \frac{\ln 2}{t_{1/2} \times RC_{50}} \quad (9)$$

In the case of second-order kinetics, the following equation is recommended instead:[73]

$$k_2 = \frac{1}{t_{1/2} \times (RC_{50} - c^0_{Nu})} \times \ln\left(2 - \frac{c^0_{Nu}}{RC_{50}}\right) \quad (10)$$

Here, $c^0_{Nu}$ is the initial concentration of the nucleophile. One should keep in mind that measured $RC_{50}$ values should not exceed water solubility (or solubility in the particular reaction medium); otherwise, they will need to be extrapolated from data measured below the solubility level.

Another method to quantify reactivity is to use percent depletion (%DP) of a degraded compound after a fixed time $t$. If an excess initial electrophile concentration $c^0_{EX}$ is used and the depletion of a nucleophile measured, the following expression can be used:[72]

$$k_2 = \frac{\ln\left(\frac{100}{100 - \%DP}\right)}{t \times c^0_{EX}} \quad (11)$$

One significant disadvantage of this method is when the depletion rate is close to 0% or 100%. At these times, small errors in the depletion value lead to large errors in $k$. It is less accurate than a kinetics experiment because only two data points are used.

## 3.2. Model Nucleophiles

The choice of the reference nucleophile depends on the expected mechanism and the site of action. In order of increasing hardness, the nucleophilic sites in biomolecules are as follows (as shown in Table 1):[10] first, thiol groups of cysteinyl residues and glutathione; second, sulfur atoms of methionyl residues; third, primary amino groups (arginine, lysine) and secondary amino groups (histidine); then there are amino groups of purine bases in RNA and DNA, followed by oxygen atoms of purines and pyrimidines; and finally, phosphate oxygen (P=O) of RNA and DNA are the hardest nucleophiles in biological systems. An overview of reference and model nucleophiles and their main properties is given in Table 7.

Soft electrophilic interactions, involving the thiol group −SH, can be modeled by small molecules, such as mercaptopropionate[74] or propanethiolate,[75] and peptides such as GSH, cysteine, acetylcysteine, or peptides with a cysteine residue.[76] The tripeptide glutathione (GSH, L-γ-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.[77] GSH protects cells by detoxifying electrophilic compounds and acts as an antioxidant. The concentration of GSH is depleted 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. In particular, thiol groups in $Ca^{2+}$ translocases might be affected, resulting in disruption in $Ca^{2+}$ homeostasis and of the cytoskeleton, and loss of plasma membrane integrity.[78] Therefore, if the GSH concentration falls below a critical level in a cell, this can cause accumulation of cellular damage.[79]

Biochemical functions and pathways of glutathione in the reduced form GSH, and as the disulfide GSSG, are reviewed in the literature.[80−86] The general reactivity of GSH to different compound classes was reviewed by Douglas.[87] GSH is odorless and nonhazardous to work with, unlike simple thiols. It is worth mentioning that GSH has some restrictions connected with its limited solubility in organic solvents. Straight GSH depletion is modeled by a pure chemical reactivity assay. GSH is reactive toward soft electrophiles, for example, polarized α,β-unsaturated compounds, which act predominantly as Michael-type acceptors. Schiff-base formers react poorly with GSH and only in the absence of amino groups.[88] In general, soft molecules, which are readily polarizable and have a low electronegativity, are easily oxidized.

The utility of applying experimental measurement of reactivity has been demonstrated by a number of different publications showing a good qualitative relationship between GSH reactivity of Michael type acceptors and biological end points, for example, acute fish toxicity[79] or skin sensitization.[77] For example, the different toxicity (effects and potency) of methacrylate, acrylate, and crotonate can be explained by their different reactivity, although the structures are very similar: methacrylate is only partially reactive, while crotonate has moderate reactivity and acrylate greater reactivity.[78]

Nucleophiles containing primary amino groups (−NH₂), such as in the amino acids lysine and arginine, can undergo Schiff base formation or acylation. Reference nucleophiles for such reactions are aniline[89] and butylamine.[90,91] The reactivity of n-butylamine approaches that of lysine, acetyllysine, or peptides

**Table 7. Reference Nucleophiles, Used in Chemical Reactivity Assays, and Their Main Properties**
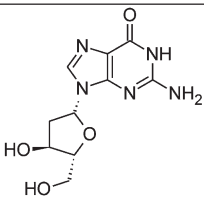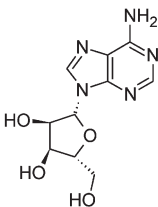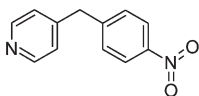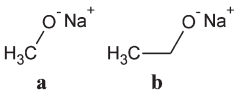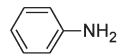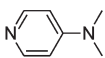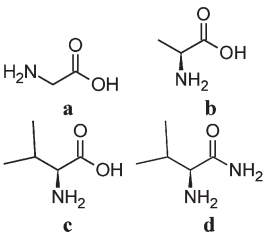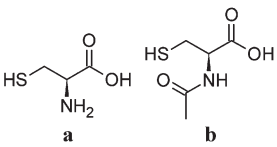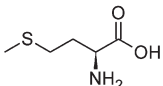
| # | Nucleophile (NU) | Structure | Modeling | Description |
|---|---|---|---|---|
| 1 | 2'-Deoxyguanosine (dG) | | DNA binding | Nucleoside, comprising guanine, attached to deoxyribose. Oxidative stress leads to its oxidized derivative as a major product of DNA oxidation. Most reactive nucleoside, which is capable for adduct formation with alkylating agents. |
| 2 | Adenosine | | DNA binding | Nucleoside, composed of adenine attached to ribofuranose. It is an inhibitory neurotransmitter and a model for ATP and ADP, which play a vital role in biochemical energy transfer. |
| 3 | 4-Nitrobenzylpyridine (NBP) | | $S_N2$ reactivity, DNA binding | Reagent for determination of organophosphorus pesticides and other electrophilic compounds spectrophotometrically. |
| 4 | Alkylates (**a**) Methylate (**b**) Ethylate | | $S_N$Ar reactivity | Nucleophilic alkylating agents which bind to electrophilic compounds. They are capable of displacing halide substituents in $S_N$Ar reactions. |
| 5 | Piperidine | | $S_N$Ar or nucleophilic addition ($A_N$) reactivity | Heterocyclic, secondary amine, which is commonly used for nucleophilic chemical degradation reactions, e.g. cleavage of nucleotides and deprotection of FMOC peptides. |
| 6 | Morpholine | | $A_N$ reactivity | Reactivity is typical towards secondary amines. The nitrogen is less nucleophilic than in piperidine due to the electron withdrawing oxygen. |
| 7 | Aniline | | $S_N$Ar or $A_N$ reactivity | Simplest aromatic amine and a weak base. Susceptible to electrophilic substitution reactions and formation of *N*-alkyl derivatives. |
| 8 | 4-Dimethylamino-pyridine | | Michael addition reactivity | Reference nucleophile for Michael addition reactions. Catalyst for acetylation of hydroxy compounds. |
| 9 | **a)** Glycine (Gly), **b)** Alanine (Ala) and derivatives **c)** Valine (Val) **d)** Valinamide | | Gly, Ala, Val, Ile, Leu reactivity | Non-polar amino acid and almost never involved in protein function or reactivity. In chemical reactivity assays, it is usually used as non-reactive reference systems. Valinamide (**d**) is a model for *N*-terminal valine residues of hemoglobin. |
| 10 | **a)** Cysteine (Cys) **b)** *N*-Acetylcysteine and derivatives | | Cys reactivity | Hydrophobic amino acid with a highly reactive soft thiol group. Together with cystine, the disulfide bonded dimer of cysteine is an important structural component of many proteins. *N*-Acetylcysteine is used for the comparison of reactivity with and without α-amine side reactions. |
| 11 | Methionine (Met) and derivatives | | Met reactivity | Second sulfur-containing amino acid, which is essential in the biosynthesis of peptides and phospholipids. Its sulfur group is significantly less reactive than in cysteine. |

**Table 7. Continued**
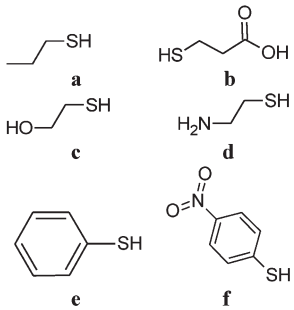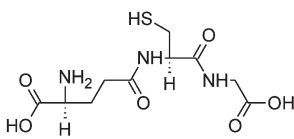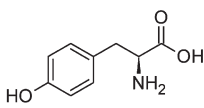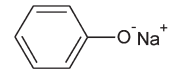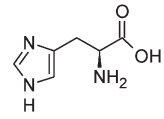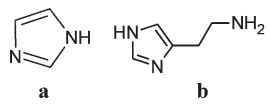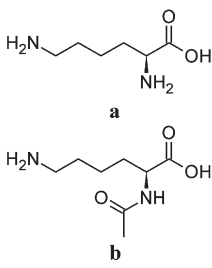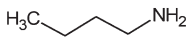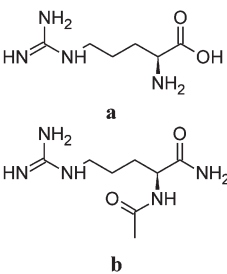
| # | Nucleophile (NU) | Structure | Modeling | Description |
|---|---|---|---|---|
| 12 | Thiolates, e.g.<br><br>**a)** Alkylthiolates,<br>**b)** Mercaptoalkylates,<br>**c)** 2-Mercaptoethanol,<br>**d)** Cysteamine and derivatives<br>**e)** Thiophenol<br>**f)** Nitrobenzenethiol (NBT) |  | Soft nucleophiles, Cys or GSH reactivity | Sulfur analogues of alcohols, the particular thiolate anions are quite nucleophilic. Thiolates are easily oxidized by electrophiles to form organic disulfides. Nitrobenzenethiol is considered as a surrogate for cutaneous proteins. |
| 13 | Glutathione (GSH) |  | Soft nucleophile | Tripeptide with vital cellular role. Conjugation of the markedly soft thiol residue to electrophiles usually leads to reductive detoxification by increased solubility. It is involved in disulfide bond rearrangements of proteins and is a cofactor for several enzymes. |
| 14 | Tyrosine (Tyr) |  | Tyr reactivity | Amino acid with phenol functionality. Tyrosine residues are part of cellular signal transduction, and their radicals part of the photosynthesis system. Can be modeled by phenolates. |
| 15 | Phenolates |  | Tyr and DNA binding | Phenolates are easily attacked by electrophiles in ortho- or para-position, because of high +M effect and low –I effect. |
| 16 | Histidine (His) |  | His reactivity | Only amino acid which acts as a proton donor and proton acceptor at the same time under physiological conditions due to the imidazole side chain. It is activated as a reactive nucleophile by proton reception, e.g. in catalytic sites of enzymes. |
| 17 | **a)** Imidazole<br>**b)** Histamine |  | His reactivity | Imidazole is an aromatic alkaloid with different tautomeric forms, present in the structure of histidine, the hemoglobin molecule, and part of many pharmaceuticals. Histamine is involved in local immune responses. Under physiological conditions, the imidazole ring is not protonated. |
| 18 | **a)** Lysine (Lys)<br>**b)** $N_\alpha$-Acetyllysine and derivatives |  | Lys reactivity | Basic amino acid with a hard, reactive ε-amino group, which is capable of alkylation or acetylation, hence proton functionalities can be modified. |
| 19 | *n*-Butylamine |  | Lys reactivity | Model nucleophile for the hard ε-amino group of lysine with very similar chemical behavior. |
| 20 | Dipeptide | H-Lys-Tyr-OH | Lys, Tyr reactivity | Model peptide, designed to measure reactivity of lysine or tyrosine residues. |
| 21 | Tripeptide | H-Gly-His-Lys-OH | Lys, His reactivity | Model peptide, designed to measure adduct formation at lysine or histidine residues. |
| 22 | Heptapeptides | **(a)** AcRFAA**C**AA,<br>**(b)** AcRFAA**K**AA,<br>**(c)** AcRFAA**H**AA | **(a)** Cys,<br>**(b)** Lys,<br>**(c)** His reactivity | Heptapeptide with cysteine, lysine or histidine as reactive site, respectively. |

**Table 7. Continued**

| # | Nucleophile (NU) | Structure | Modeling | Description |
|---|---|---|---|---|
| 23 | Cor1 heptapeptide | AcNKKCDLF | Cys and Lys reactivity | Reactive site of human Coronin 1 C (Cys420), a protein binding to cytoskeleton forming actin. |
| 24 | Peptide series I | AcVEL**X**VLL with **X** =A,C,D,F,H,K,M,N,P,Q,R,S,T,W,Y | Peptide reactivity | Series of model peptides to determine adduct formation products and reaction mechanisms for all relevant amino acids. |
| 25 | Peptide series II (with intramol. cross-linking capability) | AcL**X**EN**X'**LL**X''**F-NH$_2$ with **X,X',X''**=A,R,K,H in different mutations | Peptide reactivity | Series of model peptides to determine adduct formation products and reaction mechanisms, including (intramolecular) cross-linking. |
| 26 | Globin peptide | H-VLSPADKTNWGHEYRMFQIG-OH | Peptide reactivity | Peptide modeling parts of the hemeprotein globin containing all amino acids except Cys. |
| 27 | Nucleophilic peptide | H-PHCKRM-OH | Peptide reactivity | Peptide with essentially all nucleophilic amino acids. |
| 28 | Glutamic-C peptides (Glu-C) | (**a**) NYCN; ALYLVCGE; FVNQHLCGSHLVE (**b**) LYQLE; RGFFYTPKA | Peptide reactivity | Series of Glutamic-C peptides from digested insulin, with (**a**) and without (**b**) cysteine residues. |
| 29 | **a)** Arginine **b)** $N_\alpha$-Acetyl argininamide |  | Arg reactivity | Amino acid which enables the formation of multiple hydrogen bonds, which makes arginine containing peptides more soluble in water. The guanidinium group is positively charged in acidic, neutral and many basic environments. It plays a role in the detoxification of ammonia in the body. |
| 30 | Model peptides | (**a**) AcRRWWCR-NH$_2$; (**b**) AcFNleRF; AcPFRSVQ (**c**) AcWEHD-CHO; AcEHFRWG (**d**) AcRYYRIK-NH$_2$; For-NleLFNleYK (**e**) AcDYMGWM-NH$_2$; BOC-GWMDF | Peptide reactivity | Peptides with Cys (**a**), Arg (**b**), His (**c**), Lys (**d**) and Met (**e**) residues as potential reactive sites. |
| 31 | Acetate |  | Asp, Glu, Tyr reactivity | Model for carbon acids in aspartic acid, glutamic acid and acidic phenolate in tyrosine. |
| 32 | Hexapeptides | (**a**) AcFAA**C**AA, (**b**) AcFAA**K**AA, (**c**) AcFAA**H**AA, (**d**) AcFAA**R**AA, (**e**) AcFAA**Y**AA, (**f**) AcFAAAAA, (**g**) FAAAAA | (**a**) Cys, (**b**) Lys, (**c**) His (**d**) Arg (**e**) Tyr (**f**) Ala (**g**) N-term. reactivity | Acetylated hexapeptides with a cysteine, lysine, histidine, arginine or tyrosine as potential reactive sites, respectively, with alanine as a non-reactive reference peptide, and a hexapeptide determining N-terminal reactivity. |

with a lysine residue.[92] Histidine is a further amino acid with a secondary aromatic amino group (−NH), which should, in principle, be attacked by electrophiles. However, no significant correlation has so far been observed between biological end points and the reactivity of histidine or peptides containing histidine.[93] A possible model nucleophile for histidine is imidazole.[91]

The −OH group of the phenolic amino acid tyrosine may be modeled by phenolates or alkylates. For instance, certain 4-substituted phenols can behave as tyrosine analogues.[94] Furthermore, methylates or ethylates are used as model nucleophiles for nucleophilic substitution of activated aromatics.[95,96]

Synthetic peptides, mimicking reactive protein sites, such as Cys-420 of the human coronin 1 C, or containing all key nucleophiles have been investigated.[75,76] The purpose is to simultaneously produce a spectrum of reactivity. Synthetic peptides are capable of explaining interference effects between different amino

acid residues; for example, lysine residues have been shown to increase reactivity of cysteine residues.[97] One alternative might be to use readily purchasable proteins, such as the widely used human serum albumin (HSA). The use of an intact protein (either the target protein or closely related to it), rather than short peptide strings, has one great advantage: the influence of the three-dimensional protein environment is better represented as the surrounding hydrophobic and non-polar side chains could have an influence on the local pH, which might affect the ionization and reactivity of the particular nucleophilic site.[98] Despite the utility of HSA, the techniques to detect and to identify adducts are very time-consuming and resource demanding (e.g., MALDI-TOF-MS and nano-ES-MS/MS). Thus, this is not useful for general screening at the present time.[6]

The reactivity of compounds that are able of causing DNA damage has been measured with model nucleic acids, such as 2′-

deoxyguanosine and other deoxyribonucleosides. These nucleic acids have been used to model carcinogenicity[99] and aquatic fish toxicity.[100] Other reference nucleophiles are reported in the literature, which do not model peptide or DNA binding in particular, but model reactivity in general, for example, 4-nitrobenzylpyridine for acute fish toxicity.[101]

## 3.3. Chemical Reactivity Assays

There is a requirement to determine information regarding chemical reactivity from experimental measurement. Such information will support our capabilities to determine and predict the relative toxicities of compounds. A very early relationship between toxicity and reactivity was published by Landsteiner and Jacobs.[89] These workers found a good correlation between the skin sensitization potential in guinea pigs induced by a series of 20 benzene derivatives with halogen or nitro substituents and their reaction with aniline as a reference nucleophile. The compounds, dissolved in a mixture of absolute ethanol and aniline, were heated in a steam bath for 2 h, and the liberation of halogen was measured. If the compound did not possess labile halogens, the solution was stirred in the steam bath for 15 h, and the possible formation of the substitution compounds was inspected. If a chemical reaction was observed under these conditions, the compound acted as a sensitizer in every case, otherwise not. The authors 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 sensitization potential in the animal tests.

Since the 1930s, the ability to measure the extent and rate of reaction quantitatively has, of course, become much sophisticated. Qualitative reactivity data can be obtained by the application of mass spectrometry (MS) and nuclear magnetic resonance (NMR) techniques. Using these sensitive and accurate methods, it can be observed whether any product formation takes place or not. Furthermore, they shed light on the possible mechanism of covalent adduct formation. Typical MS yields specific mass fragments of ionized products that 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 will be different for products and reactants. Chromatographic methods, such as HPLC-UV, help to separate and to identify compounds in a quantitative manner: 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.

It should be kept in mind that reactivity measures do not necessarily reflect the specific chemical mechanism. Often it is not clear which aspect of the reaction might be relevant for potency: the selectivity (e.g., toward 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 in predicting appropriate toxicity over the past decades, even in a quantitative manner. As long as the biological target site is unknown, reaction rates should only be compared in the context of the same mechanistic domain of potential adduct formation, for example, Michael addition reactions, measured by glutathione depletion. This also holds for other mechanisms (such as nucleophilic substitution, aromatic nucleophilic substitution, Schiff base formation, or acylation for skin sensitization[102]), involving other reference nucleophiles. Increasing mechanistic understanding may help further refine in chemico reactivity assays as well as aid the interpretation of the reactivity data. When a specific biomolecular target site, leading to toxic effects, is identified, conclusions can be drawn directly from reactivity toward a reference nucleophile modeling the biological target, regardless of information about the chemical reaction mechanism.

Most of the known kinetic assays make several assumptions:[79] the endogenous electrophile concentration remains constant (steady state kinetics) and is equal to the concentration in the reaction medium. 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. Abiotic activation for pre-electrophiles can occur within the reaction medium, for example, by oxidation in air. Metabolic activation for pro-electrophiles needs separate consideration and is best modeled by special enzymatic or disparate in vitro assays.[6]

Cysteine and cysteine derivatives have often been reported in the literature for the prediction of peptide reactivity. Friedman et al. compared the reactivities of cysteinic thiol groups, measured by microamperometric titration with silver nitrate, and amino groups in a peptide, measured by manometric amino nitrogen determination.[74] The thiol moieties are usually more than 3 orders of magnitude more reactive as compared to the particular amino groups. The reactivities of both groups are approximately similar only if the thiol moieties are situated at tertiary carbon atoms (e.g., penicillamine). Hence, it is justifiable to put peptide reactivity on the same level with thiol reactivity. Presumably the transition state at the sulfur atom is energetically favored over the amino group, because it has empty 3d-orbitals. These orbitals may stabilize high-energy electrons of the polarized electrophilic site during transition state formation with the sulfur anion. The transition state with the amino group would lead to an awkward charge separation between the negative polarized electrophile and the positive polarized amino group. Among other end points, cysteine reactivity has been used to model cytotoxicity[94] and acute fish toxicity.[103]

The majority of kinetic peptide assays use glutathione as a reference nucleophile. Very early GSH and cysteine kinetic assays were reported by Dickens for halogenacetic acids, Saunders for nitrobenzenes and benzyl halogenides, and Goddard et al. for iodoethyl alcohol.[104−106] In these early studies, adduct formation under physiological conditions was observed by elemental analysis of reactants and products. Most GSH assays deal with Michael addition; however, other mechanisms (S$_N$2, S$_N$Ar, nucleophilic addition) have also been examined.[107−109]

For an in chemico assay, an excess amount of the test compound and GSH is usually dissolved in an aqueous phosphate buffer solution. After a determined reaction time, the concentration of free thiol groups is measured. Free thiol groups are usually quantified by UV/vis spectroscopy at 412 nm absorption after reaction with the chromophore 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB), also known as Ellman reagent.[110,111] Because some GSH might exist in the oxidized form GSSG with a disulfide bridge, GSH and GSSG can be separated by reverse phase HPLC columns and detected, as shown by Farris and Reed.[112] Alternatively, GSSG can be transformed selectively to

2574

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

GSH using a mixture of the enzyme glutathione reductase and TPNH, triphosphopyridine nucleotide, as described by Tietze.[113]

In 1975, Esterbauer et al. examined the equilibrium between adduct formation of conjugated carbonyls with GSH and the reverse reaction.[114] From this, a mechanistic understanding was provided, which was of particular use for the subsequent research in this field. The rate constants varied by 5 orders of magnitude, depending on the compound class (in the sequence aldehyde > ketone > ester > amide > carboxylate). An additional withdrawing group on the $\beta$-carbon led to an increased polarization and eventually to increased reactivity. In turn, reactivity falls by a factor of $10^4$ after occupation of the $\beta$-carbon by alkyl groups: for example, $k(\text{acrolein}) = 7.26 \times 10^3 \text{ M}^{-1} \text{ min}^{-1}$, but $k(\text{citral}) = 2.00 \text{ M}^{-1} \text{ min}^{-1}$.

Tanii and Hashimoto, as well as McCarthy et al., measured the glutathione reactivity of acrylates and methacrylates under physiological conditions (37 °C; aqueous phosphate buffer solution).[99,115] Acrylates and methacrylates are high volume compounds used in industry and medicine to manufacture polymers. One primary outcome was that methylacrylates are more than 1 order of magnitude less reactive (methyl methacrylate, $k = 0.325 \text{ M}^{-1} \text{ min}^{-1}$; ethyl acrylate, $k = 0.139 \text{ M}^{-1} \text{ min}^{-1}$) than their analogs without $\alpha$-methyl substitution (methyl acrylate, $k = 52.0 \text{ M}^{-1} \text{ min}^{-1}$; ethyl acrylate, $k = 26.6 \text{ M}^{-1} \text{ min}^{-1}$; butyl acrylate, $38.7 \text{ M}^{-1} \text{ min}^{-1}$) and no reaction with butyl methacrylate was observed. $\alpha$-Methyl substitution of $\alpha,\beta$-unsaturated compounds decreases reactivity toward nucleophiles significantly. The same outcome was observed in subsequent research (e.g., refs 73,88,116,117). As noted in section 3.2, the variation in reactivity (and related toxicity) is probably due to a combination of steric and electronic effects because the $\alpha$-substituent is an electron-donating group, which increases electron density on the (less electrophilic) $\beta$-atom. Note that ethylacrylate and methylmethacrylate possess similar frontier orbital energies and similar hydrophobicity.[88]

GSH has been used to determine the reactivity of agrochemicals. Clarke et al. examined second-order rate constants and half-life time constants for a broad range of compound classes, including pesticides.[118] Reactivities were studied both without an enzyme mix and with rat liver enzyme glutathione transferase[119] (GST). Reactivity was determined using an HPLC-based assay system with maximum reaction times of up to 16 h; different agrochemicals were measured at different pH values (7, 9, 10.5), and the temperatures ranged from 22 to 40 °C. This method provides a measure of the intrinsic reactivity of the agrochemicals as well as a rapid comparison of metabolically activated reactivity. GST facilitates dissociation of GSH at physiological pH values and plays a role in the detoxification of hard electrophiles.[10] The authors indicated that highly activating rat liver GST might overestimate the potential of a compound to react as an herbicide, fungicide, or insecticide. As the subject of this Review is intrinsic reactivity, this issue will not be considered in further detail at this point. There are several other GSH assays using enzymatic GST to model metabolic activation or detoxification (e.g., refs 27,120−122). Other enzymatic GSH assays use glutathione S-aryltransferase and S-epoxidetransferase.[123]

One attempt to measure the reactivity of a compound under standardized conditions within, and across different, compound classes has been introduced by Schultz et al.[88] They measured the 50% effect concentrations ($\text{RC}_{50}$) with a relatively rapid and inexpensive UV-photometric-based assay. Again, the $\text{RC}_{50}$ value is the concentration of electrophile, which gives a defined half-life

$t_{1/2}$ for the nucleophile. After 120 min incubation time at 25 °C and pH 7.4, DTNB is added to the vials, as described above, which then reacts with unconsumed GSH. Therefore, the free GSH concentration can be measured by absorption at 412 nm (as compared to a control with GSH and a blank without GSH). Different electrophile concentrations were applied to calculate static $\text{RC}_{50}$ effect levels. Effect levels, unable to be measured within the restrictions of this 2 h assay, were extrapolated.

There are a number of further applications of the Schultz assay. The structural domain for Michael-type acceptors and $S_N 2$ haloaliphatics has been defined using results obtained from it.[124,125] $\text{RC}_{50}$ values have been used to model skin sensitization[77] and aquatic toxicity[78] in terms of $\text{IGC}_{50}$ values for *Tetrahymena pyriformis*. Utilizing the Schultz assay, enhanced reactivity of ethynylene containing $\alpha,\beta$-unsaturated compounds was observed as compared to vinylene containing compounds. This can be explained as follows: the triple bond is intrinsically more electron-deficient and is, formally, formed of a strong $\sigma$-bond (ca. 250 kJ) and two weak $\pi$-bonds (ca. 135 kJ each). The double bond is formed of a strong $\sigma$-bond (ca. 250 kJ) and one weak $\pi$-bond (ca. 168 kJ).[88] Thus, the C≡C intermediate is more stable, explaining the higher reactivity toward GSH. Next, the reactivity of terminal C=C bonds is greater than of those compounds where the C=C bond is located in the interior of the molecule. This can be explained by reduced steric hindrance. It should be remembered that the critical cysteinyl moiety is in the middle of the tripeptide GSH, so its accessibility is hindered. In general, $\alpha,\beta$-unsaturated aldehydes are more reactive than $\alpha,\beta$-unsaturated ketones, due to better stabilization of the intermediate. Interestingly, $\alpha,\beta$-unsaturated esters with a vinyl group on the alcohol side of the ester possess an increased reactivity when compared to the esters without additional unsaturated moieties.

Böhme et al. recently discussed the advantages and disadvantages of the Schultz assay, in comparison with a kinetic GSH chemoassay, which also accounts for oxidative GSSG formation.[73] They were able to reproduce $\text{RC}_{50}$ values with an interlaboratory accuracy of approximately 0.3 log units. However, the static $\text{RC}_{50}$ assay appears to be limited by the water solubility and a lack of sensitivity for highly reactive compounds. The water solubility limits the maximum amount offered for reaction. Thus, compounds with similar kinetic rate constants, but different water solubilities, would have different $\text{RC}_{50}$ values. The minimum concentration required to degrade 50% of GSH is 0.07 mM, which is equal to 50% of the initial GSH concentration. This limit is reached by several highly reactive compounds. As a result, those compounds with similar $\text{RC}_{50}$ values might have differing kinetic rate constants, up to a factor of 10.

As indicated above, *n*-butylamine is used as a surrogate for hard nucleophiles such as lysine or peptides based on lysine. In this context, the relative reactivities of *n*-butylamine and lysine are likely to be similar. The acid−base equilibrium of the amine is usually established rapidly, and the nucleophilic addition to the electrophile is rate-limiting. Just as in the case of low pH values or low amine concentrations, the subsequent proton transfer reaction might become the rate-limiting step.[126] Despite the possible drawbacks of this method, a good correlation ($r^2 = 0.74$) was observed between hepatocyte cytotoxicity for 11 $\alpha,\beta$-unsaturated compounds and amine reactivity, as compared to that with GSH thiol reactivity ($r^2 = 0.21$).[127] Nevertheless, absolute rate constants for thiol reactivity are usually 10−100 times higher than rate constants for amine reactivity. Free primary amine

2575

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562−2596

groups were quantified by reaction with fluorescamine and detection of the highly fluorescent product (390 nm excitation, 475 nm emission).

In the past decade, synthetic peptides have been introduced to investigate protein reactivity. Ahlfors et al. used a model peptide containing all nucleophilic amino acid residues.[128] After reaction with a series of potential allergens under physiological conditions, the different adducts were investigated with mass spectrometry and NMR. As expected, no lysine adducts were found for α,β-unsaturated compounds; only adducts at cysteine residues could be observed. Other examinations of adduct formation with synthetic peptides, revealed by MS or NMR techniques, have been reported in the literature (e.g., refs 75,76,129,130).

The Procter & Gamble Co. has introduced a combined glutathione and pentapeptide depletion assay, which will be called the Gerberick assay in this Review. Gerberick et al.[93] suggested not using one peptide assay alone, but a battery of assays, to achieve a holistic assessment of reactivity, which could then be related to different toxicological end points. Gerberick utilized four different depletion assays for the prediction of skin sensitization potency: with GSH, and with cysteine, lysine, or histidine containing synthetic heptapeptides. Both soft and hard interactions are captured by this approach, and it is a good compromise between accuracy, time, and effort. It should be noted that the Gerberick assay does not consider the underlying mechanistic domain of the compounds tested. The assay was first published in 2004 with results for 38 compounds and updated in 2007 with reactivity for 82 compounds of different skin sensitization potencies.[116] Gerberick's GSH depletion assay is a fast screening (15 min) method carried out at 25 °C and pH 7.4 with a high excess of the electrophile (1:100). The depletion time of the peptide assays is 24 h, the electrophile excess is much lower for the peptide assays (1:10 to 1:50), and for the lysine containing peptide a higher pH value of 10.2 was used. The assay is able to distinguish between strong or moderate contact allergens and weak contact allergens or nonsensitizers. The reported concordance is 89% for 82 compounds.[116]

Natsch et al. modified the Gerberick assay slightly and added some other peptides with a different molecular environment for the cysteine residue.[76] Natsch et al. worked at higher temperatures (30 °C) to avoid precipitation, the electrophile excess is 1:10 in general, and all samples were incubated for 24 h. Thereby, some of the depletion rates were adjusted from the extremes (0% or 100% depletion) to more meaningful rates in between. Gerberick's cysteine containing heptapeptide turned out to be far more reactive than acetylcysteine, but less reactive than the cystine residue of the model peptide for human Coronin 1 C. In addition, Natsch et al. examined adduct formation with LC-MS. In the case of some aldehydes (citral and α-methyl cinnamic aldehyde), peptide depletion did not correspond with adduct formation. On the other hand, the highly reactive isoeugenol formed a range of adducts, including dimer and trimer adducts. Oxidation of some adducts was indicated by mass shifts of +16 or +32. In general, mass spectrometry allows for further elucidation of adduct structures and underlying mechanisms, a feature missing in the original Gerberick assay.

Aleksic et al. extended the standardized approach and used a panel of six single nucleophile peptides and individually optimized the incubation conditions to favor chemical modification. Employing LC-MS/MS technique, simultaneously peptide depletion, adduct formation, and peptide dimerization for a cysteine containing peptide were obtained for 36 compounds of known skin sensitizing potency.[131]

Recently, Achilleos et al. passed various allergenic electrophiles (dissolved in a phosphate buffer solution at 25 °C) flowing continuously over surface-immobilized cysteine, poly-lysine and poly-histidine residues.[98] Electrophile-peptide binding was measured by surface plasmon resonance (SPR); more precisely, the local change of refractive index, indicating binding, is recognized in real-time by this technique. Binding of electrophiles, as well as their dissociation after the end of sample injection, was observed. Even if the standard deviations from this method are relatively high and the method lacks mechanistic considerations as in the Gerberick assay, the following outcomes arise: strong allergens have a high affinity for all three amino acids, moderate allergens mainly for cysteine, the weak ones mainly for lysine together with a high dissociation rate, and nonsensitizers showed no binding at all. In addition, none of the compounds reacted with glycine.

The potency of many carcinogens correlates with their binding to the nucleophilic moieties of DNA, as shown by Lutz.[132] Covalent DNA binding of xenobiotic molecules might be captured by in chemico assays, as long as the reaction is independent from metabolic activation. Other pathways for carcinogenicity might include intercalating agents, disturbance of the nucleotide precursor pool by base analogs, strand breaks and radical formation by ionizing radiation, activation of DNA binding pathways by enzyme inducers or inhibitors, and finally stimulation of cell division by growth factor hormones or peroxisome proliferators. Here, in vivo or in vitro assays will continue to help understanding of the mechanisms of carcinogenicity and mutagenicity. McCarthy et al. did not observe any binding between ethyl acrylate with deoxyribonucleosides (2′-deoxyguanosine, 2′-deoxycytidine, 2′-deoxyadenosine, and thymidine).[99] An aqueous phosphate buffer solution was used for these experiments with reaction times of up to 24 h, two different pH values (6.7 or 7.4) and temperatures (37 or 50 °C), but under no circumstances could any adduct formation be detected by RP-HPLC. So, even if carcinogenesis is reported for acrylates, the underlying mechanism is not direct alkylation of DNA. The authors suggest alkylation of protein thiols as electrophilic intermediates in tumor promotion. Conversely, trans,trans-muconaldehyde, which is a metabolite of benzene, reacted readily with deoxyguanosine and deoxyguanosine-5′-phosphate.[133,134] After 16 h incubation time at pH 7.4 and 37 °C, adducts of radiolabeled $^{14}$C-nucleotides were detected by applying HPLC and a downstream scintillation counter.

## 3.4. Influence of Experimental Factors on Reactivity

Reaction rate constants can be influenced strongly by the nature of the reaction medium, particularly when the charge distribution changes between the reactants and the transition state. However, provided comparisons are restricted to reactions that undergo similar changes in charge distribution between reactants and transition state, relative reactivities of different electrophiles toward a given nucleophile are not significantly affected by the reaction medium. The significance of relative reactivities ($k_{rel}$, or at least the qualitative similarity in rank ordering of reactivity) is that the reactivity of an electrophile can be modeled equally well by one nucleophile as by another, provided that both are reactive in the mechanistic domain to which the electrophile belongs and the nature of the solvent is not critical.

Rate constants, $k$, are defined at a particular temperature. However, the Arrhenius equation can be used to correct for

2576

dx.doi.org/10.1021/cr100098n |Chem. Rev. 2011, 111, 2562–2596

different experimental temperatures. The Arrhenius equation assumes that the activation energy $E_A$ is independent of temperature $T$.[135]

$$\ln k(T) = \ln A - \frac{E_A}{RT} \qquad (12)$$

Therefore, the rate constant for a particular temperature can be seen in a plot of observed $\ln k$ versus $1/T$. In the equation (12), $A$ is called the pre-exponential factor, which is also termed the molecular frequency factor of the reaction.

Clarke et al. examined the temperature dependence of the reaction of N,N-dipropyl-2-chloroacetamide and 1-chloro-3,4-dinitrobenzene with GSH.[118] According to these authors, a 10 °C rise in temperature may give a 3-fold increase in reactivity with GSH, in agreement with the Arrhenius equation.

It is well-known that variation in the reaction medium affects reactivity, particularly when the charge distribution changes in transition-state formation. The same compounds, N,N-dipropyl-2-chloroacetamide and 1-chloro-3,4-dinitrobenzene, were used to investigate the relationship between rate constants and pH, using the following equation:[118]

$$\log k(pH) = \log k_{max} - \log[1 + 10^{pK_a - pH}] \qquad (13)$$

Here, $k(pH)$ is the rate constant expected at a specific pH value, $k_{max}$ is the maximum rate constant, and $pK_a = 9.2$ is the (temperature-dependent) dissociation constant of GSH to the reactive thiolate anion $GS^-$.

Most reactions of thiols ($R-SH$) with electrophiles can be related by a Brønsted-type equation. There is a linear relationship between $\log k(RS^-$ attack) versus $pK_a(RSH)$; the slope is termed the Brønsted value $\beta_{nuc}$.[87] In aqueous media, $\beta_{nuc}$ values are usually between 0.2 and 0.7 for reactive compounds. More generally, rates increase with pH and reach an asymptotic value, after the highest $pK_a$ value is exceeded.

At a pH value of 7.4 in an aqueous phosphate buffer solution, more than 90% of GSH exists in the neutral, undissociated form. The rate-limiting step of adduct formation is assumed to be the proton transfer reaction by which the adduct-intermediate is stabilized.[114] The intermediate can react with all potent proton donors present in the medium, which might be water or buffers. The catalytic effects of different buffers were compared by Esterbauer et al. for a series of thiols: phosphate buffers turned out to be the most efficient, as compared to, for example, borate, ammonium, acetate, or formiate buffers, in agreement with the Brønsted catalysis law.[114]

In contrast, rate constants for cysteine are not influenced by buffers, because a fast intramolecular proton transfer from the $NH_3^+$ group of the cysteine residue might take place. Here, the rate-limiting step is probably the addition of $CyS^-$. Therefore, cysteine reacts approximately 100 times faster than other thiols with similar $pK_a$ values and nucleophilic strength of the mercaptide ion in non-ionic systems, where it acts as its own proton donor.[114] Ionic strength did not have a significant effect on rate constants.[74]

Böhme et al. have discussed the fact that change of the cosolvent (DMSO) concentration in GSH assays could, in principle, affect the rates of adduct formation. This is because transition state formation is stabilized by increasing solution polarity.[73] However, they did not observe any significant polar covalency under the experimental conditions. Apart from this, the rate of GSSG formation as side reaction increases with

increasing DMSO concentration. The authors suggested the following route of DMSO-mediated GSSG formation:

$$GSH + DMSO \rightarrow GS^\bullet + DMS^\bullet OH \xrightarrow{GSH} GSSG$$
$$+ DMS + H_2O \qquad (14)$$
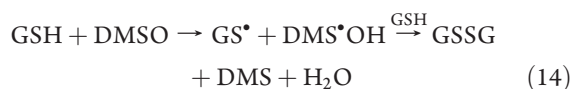
Consequently, GSH degradation through electrophilic attack tends to be overestimated without subsequent corrections. Using an oxygen-free aqueous phosphate buffer system with a fixed DMSO concentration ($c_{DMSO} = 0.279$ mol/L; $c^0_{GSH} = 1.40 \times 10^{-4}$ min$^{-1}$), a pseudo-first-order rate constant $k_{DMSO} = 7.41 \times 10^{-4}$ min$^{-1}$ was reported.

Ambient air oxygen, diluted in solution according to Henry's law constant, can also facilitate side-reaction GSSG formation:

$$GSH + O_2 \rightarrow GS^\bullet + HOO^\bullet \xrightarrow{GSH} GSSG + H_2O_2 \qquad (15)$$
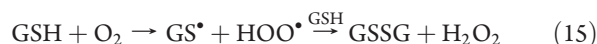
Here, the pseudofirst-order rate constant reported was $k_{O2} = 4.36 \times 10^{-4}$ min$^{-1}$ using an aqueous phosphate buffer system without DMSO ($c^0_{GSH} = 1.40 \times 10^{-4}$ min$^{-1}$).[73]

The determination of electrophilic reactivity is often limited by solubility of all reactants and products, because many peptides are soluble in water and insoluble in organic solvents, whereas many test compounds have low water solubility. In this context, Roberts and Natsch recently discussed interference from drowning out effects.[136] Therefore, five different scenarios for accounting for the levels of solubility are possible, of which the first two are favorable, and the latter three need mathematical corrections for the particular reaction competing the lack of solubility:

(1) All compounds are fully soluble in the reaction medium.
(2) The test compound is present at a nominal concentration higher than solubility, but remains supersaturated for the duration of the experiment.
(3) The test compound is dissolved only partly, and only the material in solution reacts. However, as it is consumed, more dissolves. In this case, the rate of dissolution is lower than the rate of reaction.
(4) The test compound may remain supersaturated for part of the experiment and then come out of solution.
(5) Compounds come out of solution throughout the whole reaction.

If the test compound coming out of solution (with the initial concentration $c_0$, the time $t$-dependent concentration $c(t)$ and its solubility $c_s$) is treated as a first-order process with the rate of separation $k_s$, the following equation can be written:

$$-\frac{dc}{dt} = k_s(c(t) - c_s) \qquad (16)$$

which can be integrated and solved for $c(t)$ by the following equation:

$$c(t) = c_s + (c_0 - c_s)\exp(-k_st) \qquad (17)$$

Calculating the rate of reference nucleophile depletion, the electrophile concentration needs to be corrected by eq 16 to account for precipitation.

### 3.5. Reactivity Database

Existing chemical reactivity assays from the literature are summarized in Table 8, including details of the reference nucleophiles and the number of electrophiles. For this review, a database of

## Table 8. Chemical Reactivity Assays

| assay[a] | NU[b] | electrophiles (EL)[c] | T[d]/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| k(dG) | 1 | 3 haloaliphatics 6 epoxy comp. | 30 | 7.65 | PBS | deoxyguanosine depletion measured by RP-HPLC-UV(252 nm) | 100 |
| adduct(dG) | 1 | ethyl acrylate and muconaldehyde | 37 | 7.4 | PBS | adduct formation of deoxyguanosine after 6−48 h observed by HPLC and radio chromatographic methods | 99,133 |
| adduct (adenosine) | 2 | chlorooxirane, vinyl chloride, 2-chloroacetaldehyde | 37 | 7.0 | ethanol, citrate buffer | gas chromatographic analysis of adenosine adducts after 15 min and 2 h reaction time | 193 |
| k(NBP) | 3 | 6 lactones | 15−35 | | dioxane, water | alkylation with NBP and adduct decomposition measured by UV($\lambda_{max}$) | 194−196 |
| k₁(NBP) | 3 | 19 haloaliphatics 4 halo- and nitroaromatics 16 ether and epoxy 1 aniline 4 carbonyls 11 phosphates | 80 | | 2-butanone | Hermens et al. assay: products with NBP measured by UV($\lambda_{max}$) ca. 530−550 nm) after 60 min reaction time | 101,172,193, 197−200 |
| ΔExt(NBP) | 3 | 17 haloaliphatics 1 epoxy comp. 2 sulfur comp. | 80 | | ethyl methyl ketone | alkylating activity for NBP measured by the change in extinction ($\lambda$ = 560 nm) after 10 min | 200−202 |
| k(MeO⁻) | 4a | 67 halo- and nitroaromatics | 0−175 | | methanol | determination of kinetic rate constants with methanolate in absolute methanol | 89,95 |
| k(EtO⁻) | 4b | 28 halo- and nitroaromatics | 0−150 | | ethanol | kinetic rate constants for ethanolate in absolute ethanol | 89,95 |
| k(aniline) | 7 | 17 haloaliphatics 2 haloaromatics 1 aromatic amine 2 carbonyls | 20−50 | | ethanol and DMSO:water | determination of kinetic rate constants with aniline or aniline derivatives | 95,126 |
| k_rel(aniline) | 7 | 3 haloaliphatics 1 ether and epoxy comp. 1 acyl halogenide 1 carboxyl 1 sulfonic halogenide | 0−50 | | ethanol:water acetone:water dioxane:water | relative rate constants of aniline and acetate as reference nucleophiles, as compared to water (being the hydroxide ion as reactive species) | 34 |
| k_rel(acetate) | 31 | | | | | | |
| EL-DP (aniline) | 7 | 20 halo- and nitroaromatics | 100 | | ethanol | Landsteiner et al. assay: depletion rate of electrophiles after 2 h in a steam bath, measured by halogen liberation | 89 |
| k(Me₂NPyr) | 8 | 11 Michael acceptor compounds | 25 | | aqueous solution | different classes of Michael acceptor compounds are measured by UV at $\lambda_{max}$ and the adducts elucidated by ¹H NMR | 203 |

**Table 8. Continued**

| assay[a] | NU[b] | electrophiles (EL)[c] | $T^{d}$/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| $k$(Gly) $k$(Ala) | 9 | 3 α,β-unsat. comp. (acrylonitrile, methyl acrylate, and acrylamide) | 30 | 8.1 | | determination of kinetic rate constants with glycine, alanine, and leucine derivatives | 74 |
| $k$(Cys) | 10 | 1 nitrile, 1 allyl alcohol, 23 α,β-unsat. comp., 16 carbonyls, 1 nitroso comp., 1 hydrazone | 20−37 | 6.8−8.1 | | Friedman et al. assay: determination of kinetic rate constants with cysteine and its derivatives | 74,94,103, 204−208 |
| $k$(Cys) | 10 | acrylonitrile | 30 | 8.1 | | kinetic rate constants for acrylonitrile with 10 different cysteine derivatives, including the −SH moiety and the −NH$_2$ moiety, measured by mikroamperometric titration | 74 |
| $k$(Met) | 11 | | | | | | |
| $k$(Cys) | 22a | 27 Michael acceptor compounds | 25 | 7.5 | PBS, ACN | Roberts/Natsch assay: LC-MS and kinetic analysis of Michael addition adducts, in addition free SH groups detected by fluorescence with monobromobimane ($\lambda_{ex.}$ 385/$\lambda_{em.}$ 480 nm); 24 h light protected reaction, results are corrected for drowning out effects | 136 |
| $k_1$(Cys) | 10a | 3 α,β-unsat. comp. | 25 | 7.4 | PBS | product concentrations measured by UV($\lambda$ = 230−280 nm) | 92 |
| $RC_{50}$(Cys) | 22a | 2 phenolic comp., 8 α,β-unsat. comp. | 30 | 7.4 | PBS | after 24 h reaction time, unreacted cysteine peptide concentration determined with monobromobimane fluorescence UV($\lambda_{ex.}$ 385 nm; $\lambda_{em.}$ 480 nm) | 76 |
| adduct (AcCys) | 10b | 4-oxo-2-nonenal | 37 | 7.4 | PBS | Michael addition products formed after 1 h, secondary reactions checked after 72 h by NMR and HMBC; reaction products compared to reactivity to bovine serum albumin and adduct formation confirmed with LC-MS | 209 |
| $k$(−SH) | 12a−e | 5 halo- and nitroaromatics, 4 α,β-unsat. comp., 1 nitroso comp., 3 hydrazones | 25−37 | 6−8 | | determination of kinetic rate constants with aliphatic and aromatic thiol containing compounds | 74,126,205,206, 97,210−214 |
| $k$(NBT) | 12f | 10 Michael acceptors, 8 S$_N$1/S$_N$2/S$_N$Ar, 4 acylating compounds | 25 | 7.4 (and 5.5) | PBS, ACN, acetone | Chipinda et al. assay: stopped-flow and UV measurements, loss of free thiol groups by UV-spectroscopy ($\lambda$ = 412 nm), using nitrobenzenethiol as a surrogate for cutaneous proteins | 215 |

**Table 8. Continued**

| assay[a] | NU[b] | electrophiles (EL)[c] | T[d]/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| $k$(PhOH) | 15 | 1-chloro-2,4-dinitrobenzene | 25 | | ethanol or methanol | rate constants of 1-chloro-2,4-dinitrobenzene with sodium phenolate and sodium cresolates | 95 |
| $k_1$(His) | 16 | 3 quinones (BDMP-QM, BHT-QM, and BHTOH-QM) | 25 | 7.4 | PBS | first-order rate constants with histidine, lysine, and $N_\alpha$-acetyllysine by HPLC-UV ($\lambda$ = 230 nm, 280 nm); histidine adduct formation confirmed with ¹H NMR of the 5.0 ppm imidazole-alkyl-singlet | 92 |
| $k_1$(Lys) | 18 | | | | | | |
| $k$(Gly) | 9a | 6 alkynes | | | self-buffered solution ($NH_3 \cdot HCl$ + NaOH) | determination of kinetic rate constants for various amines and glycine derivatives by maximum absorption of enaminone and ¹H NMR technique | 216,217 |
| $k$(piperidine, morpholine, amines) | 5, 6, 19 | | | | | | |
| $k$(BuNH₂) | 19 | 7 haloaliphatics, 4 halo- and nitroaromatics (for 2 of them, 6 additional amines as nucleophiles), 21 $\alpha,\beta$-unsat. comp. | 23−100 | | aqueous borate buffer (pH 8.5) or 1,4-dioxan or ethanol | Chan/O'Brien assay: determination of kinetic rate constants of $n$-butylamine; in general, free amine groups can be detected with addition of fluorescamine by fluorescence ($\lambda_{ex.}$ 390 nm; $\lambda_{em.}$ 475 nm) | 90,91,95,127, 153,214,218 |
| $k_{rel}$(BuNH₂) | 19 | 20 lactone derivatives | | | $CDCl_3$ | adduct formation rates of lactones with different side chains with $n$-butylamine, relative to the 3-methylchloride lactone derivative, by ¹H NMR technique | 91 |
| $k$ (Ppiperidine, morpholine, BuNH₂) | 5, 6, 19 | 3 nitroaliphatics, 2 nitroaromatics, 2 nitriles, 2 $\alpha,\beta$-unsat. comp. | 20 | | mixture of DMSO and water | reaction and reverse reaction kinetic rate constants for nucleophilic addition reactions with detailed analysis of underlying reaction mechanism and equilibrium constants | 126 |
| $k$ and $k_1$ (piperidine) | 5 | 28 halo- and nitroaromatics | 0−165 | | partly ethanol or benzene | second-order and pseudo first-order reaction rate constants for nucleophilic aromatic substitution reactions with piperidine | 95 |
| DP(Cys), DP(Lys), DP(His) | 22a, 22b, 22c | 4 hydrocarbons, 7 haloaliphatics, 3 haloaromatics, 3 heteroaromatics, 11 alcohols, 6 phenolic comp., 9 amines, anilines, 4 ether and epoxy, 26 carbonyls, 32 carboxyls, 30 $\alpha,\beta$-unsat. comp., 13 sulfur comp., 6 other | 25−30 | 7.5, 10.2, 7.5 | PBS acetate buffer PBS | Gerberick et al. assay: depletion assay with designed heptapeptides containing Cys, Lys, or His; after 24 h reaction time depletion rate in % measured by HPLC-DAD | 76,116,185 |

2580

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

## Table 8. Continued

| assay[a] | NU[b] | electrophiles (EL)[c] | $T^d$/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| DP(Cys), DP(Lys), DP(His), DP(Arg), DP(Tyr), DP(Ala), DP(N-term) | **32a, 32b, 32c, 32d, 32e, 32f, 32g** | 4 haloaliphatics; 2 halo- and nitroaromatics; 3 alcohols; 2 phenolic comp.; 1 aniline; 8 carbonyls; 7 carboxyls; 9 $\alpha,\beta$-unsat. comp. | 25−30 | 7.5, 10, 7.5, 10, 10, 7.5, 7.5 | PBS, ACN carbonate buffer PBS, ACN carbonate buffer carbonate buffer PBS, ACN PBS, ACN | Aleksic et al. assay: depletion assay with designed hexapeptides containing Cys, Lys, His, Arg, Tyr; after 24 h reaction time depletion rate in % measured by LC-MS/MS with AcFAGAGA as internal standard, delivering detailed mechanistic information from adducts formed | 131 |
| DP(Cor1) | 23 | 4 haloaliphatics; 1 halo- and nitroaromatics; 1 heteroaromatic; 2 ethers and anisols; 4 alcohols; 3 phenolic comp.; 4 amines and anilines; 15 carbonyls; 12 carboxyls; 2 ureas; 23 $\alpha,\beta$-unsat. comp.; 9 sulfur comp. | 37 | 7.5 | PBS, ACN | Natsch et al. assay: Depletion of the Cor1-C420 heptapeptide measured by LC-MS spectrometry after 24 h reaction time, and qualitative elucidation about type of adducts formed; free SH groups are detected by fluorescence with monobromobimane (385/480 nm) | 219 |
| k(GSH) | 13 | 13 haloaliphatics; 14 halo- and nitroaromatics; 1 nitrile; 7 ether and epoxy comp.; 2 carbonyls; 9 carboxyls; 96 $\alpha,\beta$-unsat. comp.; 4 sulfur comp.; 2 hydrazones; 8 azo comp. | 25 [10−65] | 7.4 [7.0− 11.5] | PBS, DMSO [partly tetraborate, or EDTA buffer; partly ethanol instead of DMSO] | Böhme/Schüürmann [and related] assays: determination of kinetic rate constants with GSH; unreacted GSH was measured in general by UV-spectroscopy ($\lambda$ = 412 nm) after DTNB-addition as chromophore [partly UV measurement of GSH conjugates or by fluorescence after GSH-derivatization with $o$-phthalaldehyde ($\lambda_{ex.}$ 350 nm; $\lambda_{em.}$ 420 nm)] | 73,74,79,94,99, 100,104, 107,117, 118, 121,122, 127,205,206, 218,220− 223 |
| k(GSH) | 13 | 29 $\alpha,\beta$-unsat. comp. | 20 | 7.4 | PBS | Esterbauer et al. assay: determination of second-order kinetic rate constants with glutathione as well as first-order constants for the reverse reaction and equilibrium constants; unreacted glutathione was measured in general by UV-spectroscopy ($\lambda$ = 412 nm) after DTNB-addition as chromophore, in addition the carbonyl absorption was measured ($\lambda$ = 222 nm) | 114 |

2581

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562−2596

**Table 8. Continued**

| assay[a] | NU[b] | electrophiles (EL)[c] | $T^d$/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| $\beta$(GSH) | 13 | 1 haloaliphatic<br>1 nitrile<br>5 epoxy comp.<br>2 carboxyls<br>2 $\alpha,\beta$-unsat. comp.<br>6 sulfur comp.<br>1 hydrazone<br>2 other | RT | | aqueous medium | Brønsted data for attacks of GS⁻ on various electrophiles | 87 |
| $k_{rel}$(GSH): DCNB | 13 | 5 halo- and nitroaromatics | 25 | 6.5–7.5 | PBS, partly DMSO | rate constants relative to 1-chloro-2,4-dinitrobenzene (DCNB), UV-spectroscopy of the GSH conjugate after 15 min reaction time | 107 |
| $k_{rel}$(GSH): DEM | 13 | 16 $\alpha,\beta$-unsat. comp. | 50 | 7.4 | ethanol | rate constants relative to diethylmaleate (DEM); detection of electrophile depletion by RP-HPLC ($\lambda = 254$ nm); reactive compounds measured at pH 3–5, low reactive compounds at pH up to 9, then extrapolated to pH 7.4 | 224 |
| $t_{1/2}$(GSH) | 13 | 9 haloaliphatics<br>6 haloaromatics<br>2 ether and epoxy<br>3 carboxyls<br>59 $\alpha,\beta$-unsat. comp.<br>3 sulfur comp. | 10–65 | 7–11.5 | PBS or EDTA buffer, partly with ethanol or DMSO | degradation half-life for glutathione; either unreacted glutathione was measured by UV-spectroscopy ($\lambda = 412$ nm) after DTNB-addition as chromophore, or depletion of the carbonyl group at $\lambda_{max}$ (210–300 nm) | 73,114,118,224 |
| $RC_{50}$(GSH) | 13 | 7 hydrocarbons<br>55 halo- and nitroaliphatics<br>4 aromatic hydrocarbons<br>4 halo- and nitroaromatics<br>2 nitriles<br>4 heteroaromatics<br>3 alcohols<br>7 phenolic comp.<br>5 amines, anilines, imines<br>1 anisole<br>15 carbonyls<br>15 carboxyls<br>104 $\alpha,\beta$-unsat. comp.<br>6 sulfur comp.<br>28 isothiocyanates | RT | 7.4 | PBS, max. 20% DMSO (some unsoluble compounds were measured in PBS plus 10% ethanol or an 1:1 mixture of methanol and PBS) | Yarbrough/Schultz assay: reactive concentration of electrophile to deplete 50% of 0.07 mM glutathione within 120 min; unreacted glutathione was measured by UV-spectroscopy ($\lambda = 412$ nm) after DTNB-addition as chromophore; highly reactive chemicals ($RC_{50} \leq 0.07$ mM) needed to be extrapolated | 5,73,77,78,88,108,109,114,124,125,164,1-65,218,8 |
| adduct (GSH) | 13 | 3 haloaliphatics<br>1 haloaromatic<br>1 sulfonic halogenide | | | aqueous medium | adduct formation observed by elemental analysis, and titration of released iodides with alcoholic iodine | 105,106,210 |

2582

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

**Table 8. Continued**

| assay[a] | NU[b] | electrophiles (EL)[c] | $T^d$/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| DP(GSH): Gerberick | 13 | 1 hydrocarbon<br>2 haloaliphatics<br>3 haloaromatics<br>1 heteroaromatic<br>4 alcohols<br>2 phenolic comp.<br>5 amines, anilines<br>1 ether and epoxy<br>15 carbonyls<br>24 carboxyls<br>21 α,β-unsat. comp.<br>8 sulfur comp. | 25 | 7.4 | PBS | Gerberick et al. glutathione depletion assay: after 15 min reaction time, not reacted glutathione was detected with the method of Farris and Reed (1987) by RP-HPLC after derivatization of GSH and GSSH with DNFB | 93,116 |
| DP(GSH): Natsch | 13 | 3 hydrocarbons<br>1 haloaromatic<br>2 heteroaromatics<br>10 alcohols<br>2 phenolic comp.<br>2 ether and epoxy<br>15 carbonyls<br>15 carboxyls<br>14 α,β-unsat. comp.<br>5 other | 30 | | PBS | Natsch et al. glutathione depletion assay; not reacted glutathione was detected with Tietze's (1969) enzymic test after 24 h reaction time[113] | 76,185 |
| $t_{1/2}$ (amino acids) | 10b, 18b, 9d | hydroxymethylvinyl ketone | 37−60 | 7.4−9.0 | PBS | after 4−6 h, formation of mono- and bis-Michael addition adducts determined by HPLC-UV | 225 |
| adduct (amino acids) | 13, 16, 18a, 29a | p-benzoquinone | RT | | | visual intensity of color after reaction of p-benzoquinone with GSH, histidine, lysine, proline, arginine, and ornithine (purple solution, if reaction takes place), as compared to control (yellow control) | 226 |
| DP by SPR | 9a, 10a, 16, 18a | 1 haloaliphatic<br>2 phenolic comp.<br>2 amines<br>6 carbonyls<br>3 carboxyls<br>4 α,β-unsat. comp.<br>3 sulfur comp. | 25 | | max 10% DMSO, partly PBS; peptides immobilized | Achilleos et al. assay: surface plasmon resonance (SPR) detects binding affinity of analyte in solution by its association rate to immobilized Gly, Cys, Lys, and His (in %, after 50 s), and stability by its dissociation rate (in %, 115 s after end of analyte injection) | 98 |
| $EC_{50}$ (dipeptide) | 20 | 1 haloaromatic<br>1 nitrile<br>4 carbonyls<br>2 carboxyls<br>2 sulfur comp. | 25 | 7.5 | ACN | factor of reactivity $EC_{50}$ (mM/mM), measured by HPLC | 227 |
| adduct (tripeptide) | 21 | 3 quinones (BDMP-QM, BHT-QM, and BHTOH-QM) | 25 | 7.4 | aqueous medium | adduct formation identified by HPLC-thermospray MS (m/z 120−600) | 92 |

**Table 8. Continued**

| assay[a] | NU[b] | electrophiles (EL)[c] | $T^d$/°C | pH | solvent[e] | description[f] | ref |
|---|---|---|---|---|---|---|---|
| adduct (peptide series) | 24, 25 | formaldehyde | 35 | 7.2 | PBS | adduct formation and cross-linking capability tested with a series of peptides including all relevant amino acids, detected by HPLC-MS | 130 |
| adduct (Cys, GSH, phenolate imidazole, BuNH$_2$ peptide) | 10b, 13, 15, 17a, 19, 26 | 2 isothiazols (MCI, MI) | RT | 7.7 | PBS, ACN (13, 15, 17: aqueous) | adduct formation (after up to 50 d) with a globin model peptide, glutathione, N-acetylcysteine, phenolate, imidazole, and n-butylamine; elucidation of reaction mechanism observed by $^{13}$C NMR and $^1$H($^{13}$C) HMBC and HSQC; in addition, measurement of detoxification for glutathione (adduct cleavage) | 75,97 |
| adduct (nucleoph. peptide) | 27 | 2 quinones (p-benzoquinone and 4-tert-butyl-1,2-benzoquinone) | 26 | 3.3, 7.4, 8.4 | acetate buff., PBS, borate buffer | observation of redox behavior of quinone adducts by different reaction conditions; measurement by HPLC, $^1$H and $^{13}$C NMR, FAB-MS | 128 |
| adduct (Glu−C) | 28 | acrolein | 37 | 7.4 | NH$_4$HCO$_3$ buffer | acrolein adducts and the stability of Michael addition intermediates versus Schiff base products detected by ESI-MS/MS; reactivity at terminal −NH$_2$ checked by blocking with SFB | 228 |
| k (N-Ac-peptide-amide) | 12d, 17b, 18b, 29b, 30a | 4-hydroxy-2-nonenal 4-oxo-2-nonenal | RT | 7.4 | PBS | depletion of electrophiles by N-acylated peptideamides, measured by UV spectroscopy (at $\lambda_{max}$ and $\lambda$ = 412 nm after DTNB-addition as −SH-chromophore) | 222 |
| adduct (model peptides) | 30 | 4-hydroxy-2-nonenal 4-oxo-2-nonenal | 37 | 7.4 | tricine buffer | comparison of 4-hydroxy- and 4-oxo-2-nonenal in the reaction with a series of model peptides; detection by MALDI-TOF-MS and postsource decay analysis | 222 |

[a] Assay measures: $k$ = second-order rate constant (in M$^{-1}$ min$^{-1}$), $k_1$ = (pseudo) first-order rate constant (in min$^{-1}$), $k_{rel}$ = relative rate constant (in M$^{-1}$ min$^{-1}$), DP = depletion rate, $\Delta Ext$ = change in extinction, DP = depletion rate of a reference nucleophile (in %), EL-DP = depletion rate of the particular electrophile (in %), RC$_{50}$ = reactive concentration needed for 50% depletion of reference nucleophile after fixed reaction time (in mM), $\beta$ = Bronsted index (log $k$(GS$^-$)/pK$_a$), $t_{1/2}$ = nucleophile degradation half-life (in min), EC$_{50}$ = factor of reactivity. [b] Reference nucleophiles, shown in Table 7. [c] Electrophile compound classes. If an electrophile contains more than one compound class, the functional group is listed that reacts most likely with peptides or DNA. [d] RT = room temperature. [e] PBS = aqueous phosphate buffer solution, ACN = acetonitrile. [f] Further abbreviations explained in Abbreviations section.

various reactivity data for numerous electrophiles forming peptide and DNA adducts, or reference nucleophiles for the evaluation of protein binding mechanisms, has been compiled from a full range of in chemico assays. The database contains both qualitative and quantitative reactivity data, experimental conditions, kinetic rate constants, and qualitative information about adducts formed. The associated values (e.g., kinetic rate constants) are listed in the Supporting Information. The compound class is listed and refers to what may be thought of as the most reactive biomolecular site of adduct formation. These classes are relevant for profiling compounds according to mechanistic categories of protein binding and DNA binding. This is similar to what may be achieved by the OECD QSAR Toolbox.[46]

More specifically, the database contains chemical reactivity data for 868 electrophiles (comprising 52 compound classes and 57 different mechanisms of action for protein and DNA binding) and their reactions with 108 reference nucleophiles. In total, 3089 quantitative and qualitative data arising from 100 different assays are presented here for the first time as a compilation.

# 4. SAR AND QSAR MODELS

The use of computational "in silico" techniques to predict toxicity varies in sophistication from the relatively simplistic approach of forming chemical groupings (category formation) to the more complex development of SARs (qualitative identification of chemical (sub)structures with the potential of being reactive or toxic) and QSARs (quantitative prediction of relative reactivity or toxicity).[137] There is a rich diversity of in silico techniques; however, it is generally acknowledged that a mechanistic basis to developing models allows for easier interpretation and provides greater confidence to the user.[138] The prediction of the toxicity of reactive compounds (those that bind covalently to biological macromolecules) requires the use of chemical descriptors that encapsulate how covalently reactive a given compound is. However, currently there is a lack of both toxicological and chemical reactivity data to allow for prediction of toxicity from reactivity data across manifold mechanisms of action in a broader sense.[26] Because of this issue, two main classes of reactivity descriptors have been used in the predictive toxicology literature: those derived from experimental data and those derived from topological or quantum chemical calculations. The aim of this section is to provide details on some of the (quantitative) structure—activity relationship ((Q)SAR) models that have been published in the literature to predict reactivity in the context of toxicology. This section is not intended to be a comprehensive review (as a review of available (Q)SAR models is not the main focus of this Review), rather that the intention was to provide illustrative examples of how reactivity may be predicted for several important toxicological end points.

## 4.1. Structure—Activity Relationships (SAR): Direct Toxicity Prediction and Category Formation

The simplest structure—activity relationship method for the prediction of chemical reactivity is via the use of expert knowledge. This knowledge arises from a detailed understanding of the mechanisms of action for a given end point and allows computational rule bases consisting of so-called structural alerts to be developed.[41,124,139] The structural alerts within these rule bases represent fragments within molecules that are associated with covalent bond formation and thus potentially toxicity. An example for structural alerts is shown in Table 5, regarding a selection of direct binding DNA alerts. Such knowledge is captured in

so-called expert systems, such as DEREK Nexus,[140] the OECD QSAR Toolbox,[46] ChemProp,[141] Oncologic,[142] and TIMES-SS,[143] among others.

Recently, there has been a growth of interest in forming groups of compounds (called categories) with common structural features presumed to be associated with a common mechanism of action.[144] Such groupings can be achieved by consideration of close structural analogs or can be formed using knowledge of the chemistry underpinning the mechanistic basis. The question of when is a compound in the same mechanistic domain is closely connected with the question of chemical similarity, which is reviewed elsewhere.[145] However, structural alerts can facilitate the identification of similar analogs for use in a read-across evaluation (for the prediction of toxicity). If a robust grouping or category can be formed, interpolation of effects can take place, a process called "read-across".[146] This can be qualitative or quantitative in nature, either in chemico or in silico. Once one leaves the "safety" of a closely related series of structural analogs to form a chemical grouping, a much greater emphasis is placed on developing the mechanistic basis to the category. Developing mechanistically based categories can, in some cases, allow for the formation of larger and more distinct categories; that is, the mechanistic basis can account for a greater structural diversity and hence increase applicability. The approach to forming these categories is based around determining and defining the reaction chemistry associated with the end point of interest. The grouping of compounds or category formed may vary according to end point, for example, respiratory versus skin sensitization. Reactivity with different electrophiles may assist in the assignment of a compound to the correct category.[4] Thus, an appreciation of the spectrum of reactivity discussed in section 2.2 is vital to be able to define categories for different toxicities. Examples of software that assist in the process of grouping compounds together include the OECD QSAR Toolbox, LeadScope, ToxMatch, and AMBIT Discovery, among others.[147]

The chemistry associated with forming a chemical grouping is usually associated with chemical fragments, often compiled into "profilers".[146,148] Related to this, SARs tend to link a particular (sub)structural fragment to a biological effect (usually a toxicity). These fragments are more appropriate when they are related directly to a mechanism of action and not assigned through correlation. In addition, QSARs tend to be more significant if they are developed on a single mechanism of action.[149] Other approaches to capture the chemistry associated with mechanisms of action have included the use of SMARTS strings.[139] SMARTS strings encode the relevant structural features of a molecule in a two-dimensional representation, which have been related to mechanisms of action. Thereafter, a data set can be searched for specific strings, which are associated with a particular mechanism. However, it can be difficult to assign a compound to a single mechanism if more than one string is present. Therefore, the authors suggest adding some weight of evidence to the results.[139]

There are a number of other computational expert systems for toxicity prediction that to a greater or lesser extent include reactivity information explicitly or implicitly. Such systems include, to name a few examples, M-CASE, TOPKAT, and ToxBoxes.[150−152] These have in common that molecular structures are entered into the program, and the likelihood of potential chemical toxicity is estimated on the basis of knowledge rules. In the programs associated with grouping, these knowledge rules are based on mechanistic considerations, while M-CASE and TOPKAT use

2585

dx.doi.org/10.1021/cr100098n |Chem. Rev. 2011, 111, 2562−2596

automatically learning algorithms. TIMES-SS is a hybrid system, based on a combination of both approaches and incorporating metabolic considerations.

## 4.2. Toxicity Prediction by Experimentally Determined Descriptors of Reactivity

Because of the historical problems of capturing information relating to reactivity from theoretical descriptors alone, there have been a number of attempts to measure reactivity experimentally and use the data to form models. One of the first reactivity descriptors applied in a QSAR approach for toxicity prediction was developed by Roberts and Williams.[153] It connects the dose of an alkylating agent ($D$, in mol/L), reactivity toward a reference nucleophile ($\log k$), and hydrophobicity ($\log P$, in terms of the octanol/water partition coefficient) and is called the relative alkylation index (RAI):

$$\text{RAI} = \log D + a \log k + b \log P \quad (18)$$

The RAI is a measure of the potency of protein−electrophile interactions. It has been used for a wide range of diverse data sets for structurally related compounds, for example, alkyl bromides, halobenzenes, sultones, sulfonate esters, acrylates, aldehydes, and diketones.[154]

The data from peptide reactivity assays have also been used to identify and model reactive toxicants in a qualitative manner with a particular emphasis on skin senstiziation.[6,93,164] These studies have demonstrated that high reactivity in a peptide assay is a good indicator of skin sensitizing potential and has the ability to identify (and separate) strongly and moderately reactive skin sensitizing compounds. However, one should be cautious; a lack of reactivity in a peptide assay does not necessarily indicate a lack of skin sensitizing potential. This is because a number of compounds require metabolic activation to cause skin sensitization. The use of peptide reactivity, specifically using glutathione data, has also been shown to be of use in the quantitative prediction of toxicity to the aquatic organism *Tetrahymena pyriformis*[165] for a series of compounds acting via the $S_N2$ mechanism, as with the skin sensitization studies the compounds were all direct acting toxicants:

$$\log(\text{IGC}_{50}{}^{-1}) = 0.85 \log(\text{RC}_{50}{}^{-1}) + 1.40$$
$$n = 19, r^2 = 0.91, s = 0.25, F = 199 \quad (19)$$

where $\log(\text{IGC}_{50}{}^{-1})$ is the 40 h toxicity to *T. pyriformis* and $\log(\text{RC}_{50}{}^{-1})$ is glutathione reactivity (in mmol L$^{-1}$).

## 4.3. Toxicity Prediction Using Theoretical and Calculated Descriptors

To reduce reliance on experimental measurements of reactivity to predict toxicity, a number of authors have derived QSAR models using molecular descriptors related to the electronic structure of reactive compounds.

The correlation of the relative reactivities is well recognized in organic chemistry and is the basis of the well-established linear free energy relationships (LFERs), in particular the use of Hammett and Taft constants in physical organic chemistry.[155] Thus, molecular interactions between electrophiles and nucleophiles are governed by their properties, which can be understood in terms of general patterns of hardness and softness.

A number of authors have utilized the well-developed Hammett−Taft substituent constants[156,157] to model skin sensitization potential within well-defined mechanistic categories.[41,56,65,158]

The Hammett−Taft parameters were originally developed from experimental data to model the electronic effects (both the inductive and the mesomeric) of a variety of organic functional groups.[155,159] This equation is a linear free-energy relationship relating reaction rates of reactions with varying substituents to each other (where the change in free energy of activation is proportional to the change in Gibbs free energy):

$$\log k_{\text{rel}} = \rho \cdot \sigma + E_s \quad (20)$$

Here, $k_{\text{rel}}$ is the relative rate constant of a series of differently substituted compounds, all following the same mechanism, as compared to the rate constant with a hydrogen atom instead of a particular substituent. $\rho$ is the sensitivity of the rate constants, $\sigma$ is the Hammett constant that represents inductive and mesomeric substituent effects, and $E_s$ is the steric substituent constant, that depends on the steric requirements in the particular reaction. Friedman et al. were the first workers to apply the Hammett−Taft relationship to amino acids and peptides,[160] and it has been used successfully since then.[94,161,162]

Toxicity can be modeled (and hence predicted) by chemical reactivity within well-defined mechanistically driven chemical categories. For example, Aptula et al.[163] modeled the skin sensitization potential ($pEC_3$) of 11 aliphatic aldehydes, one α-ketoester, and four α,β-diketones with $\sum \sigma^*$, the latter being the sum of Taft substituent constants (of the two groups R and R′ in RCOR′) for aliphatic inductive effects. The $pEC_3$ value is the logarithm of the molar concentration causing a 3-fold increase in T cell proliferation in murine lymph nodes after repeated application of compounds to the skin (in the local lymph node assay). The authors derived the QSAR:

$$pEC_3 = 1.12(\pm 0.07) \sum \sigma^* + 0.42(\pm 0.04) \log P - 0.62(\pm 0.13)$$
$$n = 16, r^2 = 0.952, r_{\text{adj}}{}^2 = 0.945, s = 0.12, F = 129.6$$
$$(21)$$

In addition to the development of QSAR models, the Hammett−Taft constants have been used as measures of chemical reactivity in toxicity prediction by applying the read-across paradigm.[56] Roberts et al. illustrated the use of either the Hammett or the Taft parameters for compounds within the Michael domain as a measure of reactivity to estimate the skin sensitizing potential of a compound. The same study also illustrated the ability of a combination of Hammett and Taft reactivity parameters to allow read-across to be performed within the $S_N$Ar mechanistic domain.

Quantum chemical descriptors are related to the chemical mechanisms that are involved in covalent bond formation between biological nucleophiles and electrophilic compounds. A number of computational methods are available for the calculation of theoretical descriptors, ranging from semiempirical to ab initio. A detailed discussion of these methods is out of the scope of this Review; readers are directed to some excellent reviews for more information.[166−168] A list of commonly used quantum chemical descriptors used in QSAR applications is given in the literature,[169] and current conceptual DFT descriptors have been linked successfully with chemical reactivity.[170]

Quantitative estimates of reactivity can be obtained using electronic, geometry, or topological descriptors.[72] Empirical statistical approaches, and so-called global models, which can associate a number of (often non-interpretable) reactivity

2586

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

descriptors with toxicity, can lack mechanistic understanding, which leads to a disputable prediction quality. Models that combine, for example, bioavailability (modeled by log $P$) and reasonable electronic properties are more trustworthy in this context.

The continuum from hard to soft reactivity can be quantified by a number of calculated quantum chemical parameters.[169] For predictive toxicology, the energies of the frontier molecular orbitals of the reactants are among the most significant.[171] In particular, the lowest unoccupied molecular orbital and highest occupied molecular orbital (LUMO and HOMO, respectively) are the most critical. This is because it is the interaction of the HOMO of the nucleophile with the LUMO of the electrophile that is central to covalent bond formation.[42] Hard electrophiles have relatively high energies $E$ of the LUMO (a high $E_{LUMO}$), while soft electrophiles have relatively low, or negative, $E_{LUMO}$. In contrast, a hard nucleophile has a relatively low, and soft nucleophiles a relatively high, $E_{HOMO}$. Because the greater is the difference in $E_{LUMO}$ and $E_{HOMO}$ the greater is the likelihood the reaction will occur, there is a preference for like electrophiles to react with like nucleophiles (i.e., hard with hard and soft with soft). This preference or selectivity is best quantified by a relative reactivity probability profile of a particular electrophile with respect to a variety of model nucleophiles such as those noted in Table 1. As $E_{LUMO}$ and $E_{HOMO}$ are easily calculated by a number of methods, this provides a potential route into understanding and estimating reactivity and toxicity.[168]

Given the mechanistic importance of the frontier molecular orbitals, and especially the LUMO, it is unsurprising that a number of authors have utilized these descriptors based on these properties to model reactive toxicity. An early first example is the modeling of acute fish toxicity for a group of pesticides inhibiting the enzyme acetylcholinesterase using the parameter $E_{LUMO}$ together with the effective molecular diameter (employing the semiempirical MNDO methodology).[172] For a group of nitroaromatic compounds, $E_{LUMO}$ together with partition coefficients modeled algae toxicity reasonably well at the semiempirical AM1 level of calculation ($n = 19$, $r^2 = 0.91$, $r^2_{adj} = 0.90$, $s = 0.39$, $F = 82$).[173] Another example is the toxicity of several classes of compounds to *Tetrahymena pyriformis*, known to react via the Michael addition mechanism, which were modeled by several parameters including $E_{LUMO}$, again using AM1.[174] In the case of α,β-unsaturated aldehydes, this modeling resulted in an equation with excellent statistical fit (the authors of the study acknowledged that the ratio of three descriptors to 14 compounds is at the limit of acceptability):

$$\log(IGC_{50}^{-1}) = 0.29 \log P - 21.7(Q_{C4} + Q_{C3}) - 1.19E_{LUMO}$$
$$n = 14, r^2_{adj} = 0.97, r^2_{CV} = 0.93, s = 0.12, F = 123$$
$$(22)$$

where $(Q_{C4} + Q_{C3})$ is the sum of the atomic charges on the alkene carbon atoms.

Models for mutagenicity and carcinogenicity have also been investigated using theoretical descriptors.[175] As with all modeling of reactive toxicants, the most successful and interpretable QSAR models for these end points are through the investigation of chemical classes with a common mechanism. For example, the mutagenicity of 14 benz[a]anthracenes was found to correlate with a single descriptor, $E_{LUMO}$ (calculated using the CNDO/2 method):

$$\log(MP) = -13.38E_{LUMO} + 12.26$$
$$n = 14, r = 0.82, s = 0.27$$
$$(23)$$

where MP is the experimental potency for mutagenicity in a hepatocyte test system.

The recently proposed electrophilic index ($\omega_{el}$) has also been used by a number of authors to model reactive toxicants.[176] This index is derived from two further quantum mechanical properties, chemical potential ($\mu$) and chemical hardness ($\eta$), which themselves are related to the energies of the lowest unoccupied and highest occupied molecular orbitals ($E_{LUMO}$ and $E_{HOMO}$):[176−178]

$$\omega_{el} = \frac{\mu^2}{2\eta}$$
$$(24)$$

The concept behind eq 24 can be understood by analogy to classical electrostatics, with the relationship: $P = V^2/R$. Thus, "electrophilic power" ($P$) is related to the potential ($V$) and the resistance ($R$). Using Koopman's theorem, they can be calculated by:

$$\mu = \frac{E_{HOMO} + E_{LUMO}}{2}$$
$$(25)$$

$$\eta = E_{LUMO} + E_{HOMO}$$
$$(26)$$

A recent study illustrated the use of the electrophilic index in the modeling of the skin sensitization potential of a series of compounds within the Michael acceptor mechanistic domain.[43,179] The studies showed that within the Michael domain the electrophilic index $\omega_{el}$ (calculated at the B3LYP/6-31Gd level) could be used to rank a series of skin and respiratory sensitizing compounds correctly. In addition, it was demonstrated that the electrophilic index could be used to predict the potency of compounds using a simple read-across methodology. In a related study, the electrophilic index $\omega_{el}$ was used to model the cytotoxicity of a series of sugars acting via Michael addition.[180,181] In keeping with the findings calculated for skin sensitization, the authors of this study also demonstrated the ability of the electrophilic index (calculated using at the HF/6-31G(d) level) to rank the cytotoxicity of the compounds studied successfully. Good quality QSAR models for two classes of sugars were also constructed. The electrophilicity index can also be used to estimate the electronic contribution to Hammett substituent constants.[182]

Solvent effects on the electrophilicity index were described by Pérez et al.[183] They showed that there is a linear relationship between the change in electrophilicity index and the solvation energy. The solvation energy was calculated by the Self-Consistent Isodensity Polarized Continuum Model (SCI-PCM) with a dielectric constant of $\varepsilon = 78.5$ to mimic water as solvent and $\varepsilon = 1.0$ to mimic a highly non-polar, hydrophobic solvent.

Quantum chemical calculations provide a closer insight into chemical reactivity; they are able to provide supporting information, in addition to in chemico methods. Thermodynamic properties can be obtained by the calculation of the properties of electrophilic and model nucleophilic reactants and their products, as well as kinetic data. For the latter, the calculation of appropriate transition states is required. Further quantum and molecular mechanical (QM/MM) simulations might provide a further insight into the molecular initiating event of interest.[184] Nevertheless, these types of calculation are often very time-consuming, and there is no

2587

dx.doi.org/10.1021/cr100098n |Chem. Rev. 2011, 111, 2562–2596

established routine method available to predict chemical reactivity to this end.

## 4.4. Summary of (Q)SAR Models

The above sections have highlighted the ability of both experimentally and theoretically derived descriptors to model the formation of covalent bonds that play an important role in the toxicity of reactive compounds. Table 9 gives an overview of the mechanistically interpretable descriptors commonly used to model chemical reactivity in toxicity.

It is clear from the above brief discussion that a number of authors have utilized both experimental and theoretical descriptors to model the toxicity of reactive compounds successfully for a number of end points. The examples used illustrate the importance of mechanistically driven category formation within which structure−activity relationships can be successfully developed. It is important to realize that experimentally and theoretically derived descriptors for reactivity and electrophilicity both have their limitations and that successful implementation of either requires detailed mechanistic understanding of the toxicological event being modeled.

## 5. USE OF EXPERIMENTAL AND IN SILICO REACTIVITY DATA TO PREDICT TOXICOLOGICAL END POINTS

This Review has focused on the use of the data from experimental reactivity measurements to assist in the prediction of a number of toxicological end points. The main roles for the determination of reactivity through experimental in chemico measurement and in silico calculations to predict toxicity can be summarized as follows:[4,116]

(a) Confirming whether a compound is "reactive" to a particular endogenous nucleophile. This may provide an indication of hazard, for example, a structural alert for a particular effect, or whether a compound is associated with a mode/mechanism of action.

(b) Assisting in the definition of the *domains of reactivity*, especially the structural features associated with a particular domain or mechanism of chemical reactivity. This requires a process of intelligent testing to determine the effect of altering functional groups and substituents on compounds with known mechanisms of action. Ultimately this may form part of a profiler (based on, e.g., reactive functional groups), allowing for the formation of a category or group of compounds from which read-across may be performed.

(c) Determining the *relative potency* of a reactive compound. This may enable semi-quantitative (for example, distinguishing between low, moderate, and strong skin sensitizers), or even some quantitative estimate of potency, through either quantitative read-across or the formation of QSARs.

Information relating to reactivity must be used in a context and toxicity end point dependent manner. It will require a suitable framework to use the information, for instance, as part of an Integrated Testing Strategy (ITS).[7] The corollary to identifying a compound as being reactive is the absence of such reactivity. Should a compound not be associated with a reactive mechanism, then it could be categorized as "non-reactive". It may still require confirmatory in chemico, in vitro, or in vivo testing, or a compound may be classified as possibly "narcotic" (aquatic toxicity), "non-sensitizer", "not genotoxic", etc.

While there is a use for in chemico information, currently there is little or no guidance about how it can be used directly, and more case studies and understanding of weight of evidence are required. There are three important points to note when considering the use of in chemico data for toxicological assessment, whether part of an ITS or not. In terms of logic and practical use, the key steps are the following:

- A key decision is whether a compound is reactive or not. If this question can be answered with confidence by any means, it will provide an insight into toxicity.
- If a compound is found to be reactive, which organic chemical reaction mechanism is it reactive by? Thus, if a compound can be placed, with confidence, into a reactive mechanism and this is associated with a toxic effect, then it becomes a powerful predictive tool.
- For a prediction to be made from in chemico information, there must be a direct relationship to an adverse outcome pathway; otherwise, there is no physiological relevance.

There are a number of steps that can be applied to answer the questions stated above. The initial part of the process will involve in silico profiling. Subsequently, any category formed could be supported by existing reactivity data for the compounds within it. This latter point was one of the driving forces for the collation of data within this publication. It should be noted that (currently) relatively few data are available and even if data are available, they may exist in isolation; therefore, their application may be problematic. However, even historic data from a non-standard assay could provide some evidence regarding whether a class of compounds is reactive or not. Should reactivity data be available, then predictive models are required, an example of such a model being given by Gerberick et al.,[116] Natsch et al.,[185] and Aleksic et al.[131]

The exact role of the in chemico data will be determined by the toxicological end point being considered. For instance, within the spectrum of soft−hard electrophilic reactions, the reactions associated with soft electrophilicity may be related to skin sensitization and those with hard electrophilicity to mutagenicity. There is a requirement for more information on the relative influence of this spectrum of electrophilic interaction. However, if and when the association between a toxicological effect and reactivity can be established, then it should be possible to develop the strategies for use. For instance, if a toxicological event relies on a compound being "reactive" or falling into a specific domain of reactivity, then this is key knowledge. This can be supported by:

- well-defined chemical domains associated with relevant electrophilic domains,
- well-developed computational profilers for reactivity,
- existing reactivity data, that is, ad hoc data from historical assays, which may assist in determining whether a compound is reactivity, and
- reactivity data with exactly defined experimental conditions; currently there is no "standard" reactivity assay; however, some are becoming more established, for example, Schultz's $RC_{50}$ values or the depletion assays from Natsch, Aleksic, and Gerberick.

Different approaches may be required depending on whether it is felt there is sufficient evidence to determine if a compound is "reactive" or not, and even to which domain it may belong.

## Table 9. Selection of Molecular Descriptors Used To Model Chemical Reactivity in Toxicity

| descriptor | explanation | ref |
|---|---|---|
| | *Hammett/Taft* | |
| $\sigma$ | Hammett constant for inductive and mesomeric substituent effects in aromatic systems, where the reaction center is not conjugated with the aromatic ring | 155 |
| $\sigma^-$ | Hammett constant for inductive and mesomeric substituent effects in aromatic systems, where the reaction center is conjugated with the aromatic ring | 155 |
| $\sigma^*$ | Taft constant for inductive effects of substituents in non-conjugated aliphatic systems | 155 |
| $\rho$ | sensitivity for specific reaction and molecular system | 229 |
| $E_s$ | steric substituent constant | 229 |
| | *Global Quantum Chemical Descriptors* | |
| $E_{HOMO}$ | energy of the highest occupied molecular orbital | 166 |
| $E_{LUMO}$ | energy of the lowest unoccupied molecular orbital | 166 |
| IP $(\approx -E_{LUMO})$ [a] | ionization potential for removing a electron from a molecular system X $(X \rightarrow X^+ + e^-)$ | 166 |
| EA $(\approx -E_{HOMO})$ [a] | electron affinity attaching an additional electron to a molecular system X $(X + e^- \rightarrow X^-)$ | 166 |
| | | 230 |
| $\mu = \left(\dfrac{\partial E_{el}}{\partial N}\right)_v$ | chemical potential, defined as the change in electronic energy $E_{el}$ upon change in total number of electrons $N$ | |
| $\chi = -\mu \approx -\dfrac{1}{2}(E_{HOMO} + E_{LUMO})$ [a] | absolute electronegativity | 230 |
| | | 230 |
| $\eta = -\left(\dfrac{\partial \mu}{\partial N}\right)_v \approx -(E_{HOMO} - E_{LUMO})$ [a] | molecular hardness, defined as the change in chemical potential $\mu$ upon change in total number of electrons $N$ | |
| | | 230 |
| $S = \dfrac{1}{2\eta}$ | molecular softness | |
| $\alpha$ | molecular polarizability; note that molecules arrange themselves toward a state of minimum polarizability and maximum hardness | 230 |
| | | 176 |
| $\omega_{el} = \dfrac{\mu^2}{2\eta} = \dfrac{\chi^2}{2\eta}$ | electrophilicity index | |
| | *Charge Distribution* | |
| $Q_A(r)$ | net atomic charges (at atom $r$) | 166 |
| PSA | polar surface area, describing the spacial surface density distribution | 166 |
| $\mu$ | molecular dipole moment | 166 |
| $\sigma(\mathbf{r}) = \rho(\mathbf{r})/N$ | electronic density ($\rho$) per particle shape function, normalized to 1, expresses the distributionof the total number of electrons $N$ in a system among different parts $\mathbf{r}$ of a system | 231 |
| | *Site-Specific Molecular Descriptors* | |
| | | 232 |
| $\mathrm{I}^-(\mathbf{r}) = \sum_i \dfrac{\rho_i(\mathbf{r})|\varepsilon_i|}{\rho(\mathbf{r})}$ | local ionization energy $\rho(\mathbf{r})$ = total electronic density at point $\mathbf{r}$; $\rho_i(\mathbf{r})$ = electronic density of occupied molecular orbital (MO) $i$ with MO energy $\varepsilon_i$ | |
| | | 230 |
| $f^+(r) = \left(\dfrac{\partial \rho(r)}{\partial N}\right)^+_{v(r)} \approx \rho_{N+1}(r) - \rho_N(r)$ | electrophilic Fukui function, defined as the change in electron density $\rho$ at atom $r$ upon addition of electrons to the system ($N$ = electron number) | |

**Table 9. Continued**

| descriptor | explanation | ref |
|---|---|---|
| $\omega_{el}(r) = \omega_{el} \times f^+(r)$ | local electrophilicity index | 233 |
| $\Delta f(r) = f^+(r) - f^-(r)$ <br> $\Delta f(r) \approx \rho_{LUMO}(r) - \rho_{HOMO}(r)\,^a$ | reactivity-selectivity descriptor or dual descriptor, while $f^+(r)$ measures reactivity toward nucleophilic and $f^-(r)$ toward electrophilic attacks; therefore, electrophilic sites are identified by $\Delta f(r) > 0$. $\rho_{LUMO}$ and $\rho_{HOMO}$ are the electron densities of the LUMO and HOMO orbitals, respectively | 234 |
| $EE_{occ}(r)$, $EQ_{occ}(r)$, $QE_{occ}(r)$; <br> $EE_{vac}(r)$, $EQ_{vac}(r)$, $QE_{vac}(r)$ | energy weighted ($EE$) or charge-limited ($EQ$) electron donor energy and acceptor energy (occ or vac, respectively), and energy-limited ($QE$) electron donor and acceptor charge | 223,235– 237 |
| $FN(r)\left(= \sum\limits_{\mu(r)} c_{\mu k}^2\right)$ | frontier orbital electron density (toward a nucleophile) $c$ = molecular orbital (MO) coefficients; $\mu$ = atomic orbitals (AO) of atom $r$; $k$ = LUMO | 238 |
| $DN(r)\left(= 2\sum\limits_{k=LUMO}^{max} \sum\limits_{\mu(r)} \dfrac{c_{\mu k}^2}{\alpha - \varepsilon_k}\right)$ | acceptor delocalizability (toward a nucleophile) $c$ = MO coefficients; $\mu$ = AOs of atom $r$; $k$ = unoccupied MO with MO energy $\varepsilon_k$; $\alpha$ = relevant orbital energy of attacking nucleophile ($\approx E_{HOMO}$ of nucleophile$^a$) | 238 |
| $SN(r)\left(= 2\sum\limits_{k=LUMO}^{max} \sum\limits_{\mu(r)} \dfrac{c_{\mu k}^2}{(-\varepsilon_k)}\right)$ | superdelocalizability (toward a nucleophile) | 238 |
| $ALP(r)\left(= 4\sum\limits_{i=HOMO}^{1} \sum\limits_{k=LUMO}^{N} \sum\limits_{\mu(r)} \dfrac{c_{\mu i}^2 c_{\mu k}^2}{(\varepsilon_i - \varepsilon_k)}\right)$ | atomic self-polarizability | 238 |

$^a$ According to Koopman's theorem.

The type of evidence that may be required to demonstrate that a compound is reactive could include positive association in the computational profiler or known reactivity in an in chemico assay. To determine whether a compound is nonreactive, it could be outside of known chemistry of reactivity processes and not reactive in the in chemico assays. Both conclusions will be stronger when made using a weight of evidence approach.

As a final comment, it should not be assumed that there is a definitive relationship between reactive and non-reactive compounds and hazard. A reactive compound may not be toxic due to any number of considerations including ADME properties. Conversely, it should not be categorically concluded that a non-reactive compound will not have hazard associated with it. There are numerous other mechanisms of toxicological action for chronic end points, and in addition to (reactive) acute toxicity there is already a baseline toxic effect.

## 6. CONCLUSIONS AND OUTLOOK

The chemical reactivity of xenobiotic electrophiles toward nucleophilic reference compounds is a model for their biochem-

ical behavior toward bioactive sites, such as proteins or DNA. This provides a promising method to make assumptions regarding their potential toxic effects. Both in chemico reactivity assays and in silico estimation methods are capable of assisting in this process, as summarized in this Review.

Historically, various reactivity assays have proven their reliability to predict the biochemical activity of a compound. To obtain a holistic view about a particular compound, the results of different reactivity assays should be compared to each other: more than one chemical reaction mechanism might be involved in reactivity, and reactivity toward different nucleophilic sites is diverse. The comparison includes the use of different reference nucleophiles (e.g., soft and hard) or different experimental conditions (e.g., change of pH range accounting for varying microenvironments in a particular cellular compartment).

A database of experimentally determined values for reactivity has been compiled, which contains a list of electrophilic compounds and their chemical structure, reactivity data, and kinetic rate constants of various in chemico assays, related to peptide binding or DNA binding, and qualitative information about

**2590**

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

adducts formed. More specifically, the database contains more than 3000 quantitative and qualitative reactivity data for almost 900 different compounds and their reaction toward around 100 reference nucleophiles and assays, respectively. In combination with toxicological data, their mechanistic applicability domain, and physicochemical properties, this data collection will assist in toxicity profiling through the grouping of compounds into reactive categories as required for prioritization for testing, classification and labeling, and risk assessment. It will also assist in the development of sound screening tools, for example, to refine structural alerts in expert software systems, for the application of read-across and ITS. Computational estimation methods are able to provide supporting information, in addition to experimental measurements. On the basis of a reliable database with experimental reactivity data, it is intended to model and estimate in chemico data based on computer-aided methods.

The compiled database provides a tool to rank electrophilic compounds by examining the variation in reactivity to different nucleophilic sites and by comparison of related compounds. Analysis of the data distribution within the database reveals that a series of further in chemico tests is needed. While there are many data with regard to their reactivity to soft nucleophiles (especially, thiol groups), only sparse, quantitative, data are published about their reactivity to hard nucleophiles (e.g., amines and DNA/RNA nucleotides). Further development and systematical application of reactivity assays, especially with regard to hard nucleophiles, will improve understanding of electrophilic mechanisms for reactive toxicity and therefore will allow for classification of compounds into reactive categories with more confidence.

## ASSOCIATED CONTENT

### Ⓢ Supporting Information

Table of the reactivity data together with all experimental values, assay conditions, and underlying reaction mechanisms. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*Tel.: +44 151 231 2402. Fax: +44 151 231 2170. E-mail: m.t.cronin@ljmu.ac.uk.

## BIOGRAPHIES

Johannes Schwöbel started his studies of chemistry in 2000 at the Technical University of Darmstadt, Germany. He graduated under the guidance of Prof. H. Morgner at the University of Leipzig, investigating phospholipid layers using molecular dynamics computer simulations and X-ray photoelectron spectroscopy. The focus of his Ph.D. thesis, under the supervision of Prof. G. Schüürmann at the Helmholtz Centre for Environmental Research − UFZ in Leipzig, was on quantum chemical modeling of hydrogen-bond interactions for the prediction of the partitioning of compounds in vivo and the environment. Since 2008, he has been a Marie Curie postdoctoral research fellow working on the InSilicoTox Project in the QSAR and Modelling Research Area of the School of Pharmacy and Chemistry at Liverpool John Moores University, England. In this position, he investigated electrophilic reactivity and its relationship to a range of toxicological effects using in silico techniques.

Yana Koleva studied Chemical Technologies at the University "Prof. Assen Zlatarov", Bourgas, Bulgaria, and received her M.Sc. with professional classification as engineer-chemist in 1998. In the same year, she joined the Laboratory of Mathematical Chemistry at the University "Prof. Assen Zlatarov". She obtained her Ph.D. degree at the Institute of Organic Chemistry of the Bulgarian Academy of Science, Sofia, Bulgaria, under the guidance of Prof. O. Mekenyan and Assoc. Prof. S. Dimitrov. The main focus of her doctoral work was modeling the toxicity of organic compounds. She was a Marie Curie postdoctoral research fellow working on the InSilicoTox Project in the QSAR and Modelling Research Area of the School of Pharmacy and Chemistry at Liverpool John Moores University, England, from 2007−2009.

Steven Enoch is a chemist with a Ph.D. from Heriot-Watt University in Edinburgh. His research interests are focused on the use of organic chemistry to aid in the prediction of a number of toxicological end points, with a specific interest in human health. As such, he has worked on a number of projects aimed at developing the mechanistic chemistry associated with human health end points so that it can be used in a regulatory environment for risk assessment. He has published 15 journal articles, three book chapters, and has presented at both national and international scientific meetings.

Fania Bajot graduated in biochemistry from the Université Pierre Et Marie Curie (Paris VI) and in chemistry from the Université Aix-Marseille II. After specializing in drug design at the Université Lille II, she joined the University of Geneva where she obtained her Ph.D. under the supervision of Prof. Pierre-Alain Carrupt. The main achievement of her doctoral work was the development of in silico models to derive the pharmaco-kinetic profiles of new chemical entities. Since 2008, she has been a postdoctoral research fellow at Liverpool John Moores University with the main focus of her work being in silico and in chemico methodologies for the risk assessment of chemicals.

Mark Hewitt is currently undertaking a postdoctoral research position at Liverpool John Moores University focusing on the application of nontesting prediction methods in toxicology. He was trained as a pharmaceutical scientist, obtaining a Ph.D. in 2008 on the implementation of Quantitative Structure−Activity Relationships (QSARs) in reproductive toxicity. Being a partner of two European Union-funded projects (ReProTect and OSIRIS) has also led to a wide range of research interests including human placental transfer, applicability domain definition, category formation and read-across, and the development and utilization of Weight of Evidence (WoE) systems.

Judith Madden is Reader in Molecular Design at Liverpool John Moores University. Her Ph.D. studies were in the area of computer-aided drug design, and her postdoctoral research centered on pharmacokinetic profiling of drugs. Currently her research is directed toward the use of in silico tools to predict the biological activity of xenobiotics, both desirable therapeutic effects and toxicity. She also carries out research in predicting the uptake and time-course of compounds within the body, investigating pharmacokinetic and toxicokinetic profiles. She has approximately 40 publications in the area of QSAR and modeling.

David Roberts is an organic chemist with a Ph.D. from Manchester University. He spent more than 30 years at Unilever Research Port Sunlight working at the interfaces of chemistry

with toxicology and chemical engineering and has published more than 150 papers in these fields. Since 2003, he has worked as a consultant in manufacturing and toxicological chemistry and is an honorary researcher at Liverpool John Moores University. His major research interest is applying mechanistic organic chemistry to modeling of toxic end points.

Terry Schultz is Emeritus Professor at The University of Tennessee, Knoxville, TN. He has more than 30 years experience in predictive toxicology and serves as director of the Biological-Activity Testing and Modeling Laboratory in the College of Veterinary Medicine. He has authored more than 300 papers on a variety of topics including acute toxicity, teratogenesis, and receptor-mediated endocrine disruption. His current research interests include generating in chemico reactivity data from simple, rapid abiotic methods and using these data to form categories and conducting read-across analysis in toxicity assessments.

Mark Cronin has held an academic position at Liverpool John Moores University, England, since 1994. His training is at the interface of biology and chemistry with a Ph.D. from Liverpool Polytechnic in the application of quantitative structure–activity relationships (QSARs) and interspecies extrapolations for the toxicity of chemicals in the environment. He has published over 170 articles and coedited three books in the application of QSAR and in chemico methods to predict the toxicity of chemicals to both environmental species and human health effects.

## ACKNOWLEDGMENT

## ABBREVIATIONS

| | |
|---|---|
| ACN | acrylnitrile |
| ADME | absorption, distribution, metabolism, and excretion |
| ADP | adenosine diphosphate |
| AM1 | Austin Method 1 |
| ATP | adenosine-5′-triphosphate |
| B3LYP | Becke, three-parameter, Lee–Yang–Parr hybrid functional |
| BfR | Federal Institute for Risk Assessment (Germany) |
| BOC | N-blocking butyloxycarbonyl group in peptides |
| ECVAM | European Centre for Validation of Alternative Methods |
| EDTA | ethylenediaminetetraacetic acid |
| EPA | Environmental Protection Agency (United States) |
| FAB | fast atom bombardment |
| FATS | fish acute toxicity syndrome |
| FDA | Food and Drug Administration (United States) |
| FMOC | fluorenylmethyloxycarbonyl |
| CNDO | complete neglect of differential overlap |
| DAD | diode array detector |
| DFT | density functional theory |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acid |
| DNCB | dinitrochlorobenzene |
| DNFB | 1-fluoro-2,4-dinitrobenzene; Sanger's reagent |
| DTNB | 5,5′-dithiobis-(2-nitrobenzoic acid) |
| ES-MS | electrospray mass spectrometry |

| | |
|---|---|
| For | N-blocking formyl group in peptides |
| GSH | glutathione; L-$\gamma$-glutamyl-L-cysteinyl-glycine |
| GSSG | glutathione disulfide |
| GST | glutathione transferase |
| HF | Hartree–Fock |
| HMBC | heteronuclear multiple bond coherence |
| HOMO | highest occupied molecular orbital |
| HPLC | high performance liquid chromatography |
| HSA | human serum albumin |
| HSAB | hard and soft acid and bases principle |
| HSQC | heteronuclear single quantum coherence |
| $IGC_{50}$ | inhibition growth concentration for 50% impairment |
| ITS | integrated testing strategy |
| LC | liquid chomatography |
| LFER | linear free energy relationship |
| LLNA | mouse local lymph node assay |
| LUMO | lowest unoccupied molecular orbital |
| MALDI-MS | matrix-assisted laser desorption/ionization mass spectrometry |
| MIE | molecular initiating event |
| MNDO | modified neglect of differential overlap |
| MOA | mode of action |
| MS | mass spectrometry |
| NBP | 4-nitrobenzylpyridine |
| NMR | nuclear magnetic resonance spectroscopy |
| OECD | Organization for Economic Co-operation and Development |
| PBS | aqueous phosphate buffer solution |
| QSAR | quantitative structure–activity relationship |
| $RC_{50}$ | reactive electrophile concentration for a defined half-life of a nucleophile |
| RAI | relative alkylation index |
| REACH | Registration, Evaluation, Authorization, and restriction of Chemicals (European Union) |
| RNA | ribonucleic acid |
| RP-HPLC | reversed phase high performance liquid chromatography |
| S9 | supernatant postmitochondrial liver fraction |
| SAR | structure–activity relationship |
| SCI-PCM | self-consistent isodensity polarized continuum model |
| SFB | succinimido-4-fluorobenzoate |
| SMARTS | Smiles Arbitrary Target Specification |
| $S_N1$ | first-order aliphatic nucleophilic substitution |
| $S_N2$ | second-order aliphatic nucleophilic substitution |
| $S_NAr$ | nucleophilic aromatic substitution |
| SPR | surface plasmon resonance |
| TAE | thioglycolic adid ethylester |
| TOF | time of flight |
| UV/vis | ultraviolet–visible spectroscopy |
| WoE | weight of evidence |

## REFERENCES

(1) Cronin, M. T. D.; Livingstone, D. J. *Predicting Chemical Toxicity and Fate*; CRC Press: Boca Raton, FL, 2004.

(2) Bassan, A.; Worth, A. P. *QSAR Comb. Sci.* **2008**, *27*, 6.

(3) Vonk, J. A.; Benigni, R; Hewitt, M.; Nendza, M.; Segner, H.; van de Meent, D.; Cronin, M. T. D. *ATLA, Altern. Lab. Anim.* **2009**, *37*, 557.

(4) Cronin, M. T. D.; Bajot, F.; Enoch, S. J.; Madden, J. C.; Roberts, D. W.; Schwöbel, J. *ATLA, Altern. Lab. Anim.* **2009**, *37*, 513.

(5) Schultz, T. W.; Carlson, R. E.; Cronin, M. T. D.; Hermens, J. L. M.; Johnson, R.; O'Brien, P. J.; Roberts, D. W.; Siraki, A.; Wallace, K. D.; Veith, G. D. *SAR QSAR Environ. Res.* **2006**, *17*, 413.

(6) Gerberick, F.; Aleksic, M.; Basketter, D. A.; Casati, S.; Karlberg, A.-T.; Kern, P. S.; Kimber, I.; Lepoittevin, J.-P.; Natsch, A.; Ovigne, J. M.;

2592

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

Rovida, C.; Sakaguchi, H.; Schultz, T. W. *ATLA, Altern. Lab. Anim.* **2008**, *36*, 215.

(7) Cronin, M. T. D. *Altern. Anim. Exp.* **2010**, *27*, 269.

(8) Schultz, T. W.; Rogers, K.; Aptula, A. O. *Contact Dermatitis* **2009**, *60*, 21.

(9) Veith, G. D. *SAR QSAR Environ. Res.* **2004**, *15*, 323.

(10) Coles, B. *Drug Metab. Rev.* **1984**, *15*, 1307.

(11) Miller, E. C.; Miller, J. A. *Cancer Res.* **1947**, *7*, 468.

(12) Rand, G. M.; Wells, P. G.; McCarty, L. S. *Fundamentals of Aquatic Toxicology: Effects, Environmental Fate and Risk Assessment*; Taylor & Francis: Washington, DC, 1995; p 3.

(13) Spycher, S.; Nendza, M.; Gasteiger, J. *QSAR Comb. Sci.* **2004**, *23*, 779.

(14) Borgert, C. J.; Quill, T. F.; McCarty, L. S.; Mason, A. M. *Toxicol. Appl. Pharmacol.* **2004**, *201*, 85.

(15) Butterworth, B. E.; Conolly, R. B.; Morgan, K. T. *Cancer Lett.* **1995**, *93*, 129.

(16) Dellarco, V. L.; Wiltse, J. A. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **1998**, *405*, 273.

(17) Schlosser, P. M.; Bogdanffy, M. S. *Regul. Toxicol. Pharmacol.* **1999**, *30*, 75.

(18) US EPA. *Draft Dioxin Reassessment, Part III: Integrated Summary and Risk Characterization for 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and Related Compounds*; U.S. Environmental Protection Agency: Washington, DC, 2000; p 196.

(19) US EPA. *Office of Research and Development: Draft Final Guidelines for Carcinogen Risk Assessment*; U.S. Environmental Protection Agency: Washington, DC, 2001; p 123.

(20) Schultz, T. W. In *In Silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; Royal Society of Chemistry: Cambridge, 2010.

(21) Gibbs, S.; van de Sandt, J. J. M.; Merk, H. F.; Lockley, D. J.; Pendlington, R. U.; Pease, C. K. *Curr. Drug Metab.* **2007**, *8*, 758.

(22) Guengerich, F. P. *Arch. Biochem. Biophys.* **2005**, *433*, 369.

(23) Evans, D. C.; Watt, A. P.; Nicoll-Griffith, D. A.; Baillie, T. A. *Chem. Res. Toxicol.* **2004**, *17*, 3.

(24) Ames, B. N.; Durston, W. E.; Yamasaki, E.; Lee, F. D. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 2281.

(25) Bajot, F.; Cronin, M. T. D.; Roberts, D. W.; Schultz, T. W. *SAR QSAR Environ. Res.* **2010**, in press.

(26) Roberts, D. W.; Aptula, A. O.; Patlewicz, G. *Chem. Res. Toxicol.* **2007**, *20*, 40.

(27) Labrou, N. E.; Karavangeli, M.; Tsaftaris, A.; Clonis, Y. D. *Planta* **2005**, *222*, 91.

(28) Bradbury, S. P.; Christensen, G. M. *Environ. Toxicol. Chem.* **1991**, *10*, 1155.

(29) Pearson, R. G.; Songstad, J. *J. Am. Chem. Soc.* **1967**, *89*, 1827.

(30) Ho, T.-L.; Ho, H. C.; Hamilton, L. D. *Chem.-Biol. Interact.* **1978**, *23*, 65.

(31) Pearson, P. G. *J. Am. Chem. Soc.* **1963**, *85*, 3533.

(32) Carlson, R. M. *Environ. Health Perspect.* **1990**, *87*, 227.

(33) Hinson, J. A.; Roberts, D. W. *Annu. Rev. Pharmacol. Toxicol.* **1992**, *32*, 471.

(34) Swain, C. G.; Scott, C. B. *J. Am. Chem. Soc.* **1953**, *75*, 141.

(35) Mayr, H.; Patz, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 938.

(36) Mayr, H.; Bug, T.; Gotta, M. F.; Hering, N.; Irrgang, B.; Janker, B.; Kempf, B.; Loos, R.; Ofial, A. R.; Remennikov, G.; Schimmel, H. *J. Am. Chem. Soc.* **2001**, *123*, 9500.

(37) Minegishi, S.; Mayr, H. *J. Am. Chem. Soc.* **2003**, *125*, 286.

(38) Loechler, E. L. *Chem. Res. Toxicol.* **1994**, *7*, 277.

(39) Roberts, D. W. In *Aquatic Toxicology and Environmental Fate*; G. W., S. I., Lewis, M. A., Eds.; American Society for Testing and Materials: Philadelphia, PA, 1989; Vol. 11 (ASTMSTP1007), p 490.

(40) Lipnick, R. L. In *QSAR in Environmental Toxicology-IV*; Hermans, J. L. M., Opperhuizen, A., Eds.; Elsevier: Amsterdam, The Netherlands, 1991; p 131.

(41) Aptula, A. O.; Roberts, D. W. *Chem. Res. Toxicol.* **2006**, *19*, 1097.

(42) Jacobs, A. *Understanding Organic Reaction Mechanisms*; Cambridge University Press: Cambridge, United Kingdom, 1997; p 304.

(43) Enoch, S. J.; Cronin, M. T. D.; Schultz, T. W.; Madden, J. C. *Chem. Res. Toxicol.* **2008**, *21*, 513.

(44) Wong, H. L.; Liebler, D. C. *Chem. Res. Toxicol.* **2008**, *21*, 796.

(45) Boobis, A. R.; Doe, J. E.; Heinrich-Hirsche, B.; Meek, M. E.; Munn, S.; Ruchirawat, M.; Schlatter, J.; Seed, J.; Vickers, C. *Chem. Rev. Toxicol.* **2008**, *38*, 87.

(46) OECD Organization for Economic Co-operation and Development. *QSAR Toolbox, 2.0*; OASIS, Laboratory of Mathematical Chemistry: Bourgas, Bulgaria, 2010.

(47) Sykes, P. *Guidebook to Mechanism in Organic Chemistry*; Longman Group: London, United Kingdom, 1986.

(48) Roberts, D. W. *Chem. Res. Toxicol.* **1995**, *8*, 545.

(49) Serafimova, R.; Todorov, M.; Pavolo, T.; Kotov, S.; Jacob, E.; Aptula, A.; Mekenyan, O. *Chem. Res. Toxicol.* **2007**, *20*, 662.

(50) Mekenyan, O.; Dimitrov, S.; Serafimova, R.; Thompson, E.; Kotov, S.; Dimitrova, N.; Walker, J. D. *Chem. Res. Toxicol.* **2004**, *17*, 753.

(51) Mekenyan, O.; Todorov, M.; Serafimova, R.; Stoeva, S.; Aptula, A.; Finking, R.; Jacob, E. *Chem. Res. Toxicol.* **2007**, *20*, 1927.

(52) Benigni, R.; Bossa, C. *Mutat. Res.* **2008**, *659*, 248.

(53) Ashby, J.; Tennant, R. W. *Mutat. Res.* **1988**, *204*, 17.

(54) Miller, E. C.; Miller, J. A. *Cancer* **1981**, *47*, 2327.

(55) Miller, J. A.; Miller, E. C. In *Origins of Human Cancer*; Hiatt, H. H., Watson, J. D., Winsten, J. A., Eds.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1977; p 605.

(56) Roberts, D. W.; Patlewicz, G.; Kern, P. S.; Gerberick, F.; Kimber, I.; Dearman, R. J.; Ryan, C. A.; Basketter, D. A.; Aptula, A. O. *Chem. Res. Toxicol.* **2007**, *20*, 1019.

(57) Benigni, R.; Giuliani, A.; Franke, R.; Gruska, A. *Chem. Rev.* **2000**, *100*, 3697.

(58) Strobel, R.; Rohrborn, G. *Mutat. Res.* **1980**, *74*, 156.

(59) Goggelmann, W.; Summer, K.-H. *Mutat. Res.* **1980**, *74*, 155.

(60) Wang, H.-T.; Zhang, S.; Hu, Y.; Tang, M.-s. *Chem. Res. Toxicol.* **2009**, *22*, 511.

(61) Stein, S.; Lao, Y.; Yang, I.-Y.; Hecht, S. S.; Moriya, M. *Mutat. Res.* **2006**, *608*, 1.

(62) Rannug, U.; Göthe, R.; Wachtmeister, C. A. *Chem.-Biol. Interact.* **1976**, *12*, 251.

(63) Mosier, P. D.; Jurs, P. C.; Custer, L. L.; Durham, S. K.; Pearl, G. M. *Chem. Res. Toxicol.* **2003**, *16*, 721.

(64) Lake, B. G. *Food Chem. Toxicol.* **1999**, *37*, 423.

(65) Patlewicz, G.; Basketter, D. A.; Smith, C. K.; Hotchkiss, S. A. M.; Roberts, D. W. *Contact Dermatitis* **2001**, *44*, 331.

(66) Benigni, R.; Conti, L.; Crebelli, R.; Rodomonte, A.; Vari, M. R. *Environ. Mol. Mutagen.* **2005**, *46*, 268.

(67) von der Ohe, P. C.; Kühne, R.; Ebert, R.-U.; Altenburger, R.; Liess, M.; Schüürmann, G. *Chem. Res. Toxicol.* **2005**, *18*, 536.

(68) Cronin, M. T. D.; Dearden, J. C.; Duffy, J. C.; Edwards, R.; Manga, N.; Worth, A. P.; Worgan, A. D. P. *SAR QSAR Environ. Res.* **2002**, *13*, 167.

(69) Saliner, A. G.; Patlewicz, G.; Worth, A. P. *Review of Literature-Based Models for Skin and Eye Irritation and Corrosion*; Institute for Health and Consumer Protection, Joint Research Centre, European Commission: Ispra, Italy, 2006.

(70) OECD Organization for Economic Co-operation and Development. *OECD Environmental Health and Safety Publications, Series on Testing and Assessment No. 80: Guidance Document on Grouping of Chemicals*; ENV/JM/MONO(2007)28, 2007.

(71) Frost, A. A.; Pearson, R. G. *Kinetics and Mechanism. A Study of Homogenous Chemical Reactions*, 2nd ed.; Wiley: New York, 1961; p 9.

(72) Roberts, D. W.; Aptula, A. O.; Patlewicz, G.; Pease, C. *J. Appl. Toxicol.* **2008**, *28*, 443.

(73) Böhme, A.; Thaens, D.; Paschke, A.; Schüürmann, G. *Chem. Res. Toxicol.* **2009**, *22*, 742.

(74) Friedman, M.; Cavins, J. F.; Wall, J. S. *J. Am. Chem. Soc.* **1965**, *87*, 3672.

2593

dx.doi.org/10.1021/cr100098n |*Chem. Rev.* 2011, 111, 2562–2596

(75) Alvarez-Sánchez, R.; Basketter, D. A.; Pease, C.; Lepoittevin, J.-P. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 365.

(76) Natsch, A.; Gfeller, H.; Rothaupt, M.; Ellis, G. *Toxicol. Vitro* **2007**, *21*, 1220.

(77) Aptula, A. O.; Patlewicz, G.; Roberts, D. W.; Schultz, T. W. *Toxicol. Vitro* **2006**, *20*, 239.

(78) Yarbrough, J. W.; Schultz, T. W. *Chem. Res. Toxicol.* **2007**, *20*, 558.

(79) Freidig, A. P.; Verhaar, H. J. M.; Hermens, J. L. M. *Environ. Sci. Technol.* **1999**, *33*, 3038.

(80) Penninckx, M. J.; Elskens, M. T. *Adv. Microb. Physiol.* **1993**, *34*, 239.

(81) Reed, D. J. *Annu. Rev. Pharmacol. Toxicol.* **1990**, *30*, 603.

(82) Forman, H. J.; Zhang, H.; Rinna, A. *Mol. Aspect Med.* **2009**, *30*, 1.

(83) Mistry, P.; Harrap, K. R. *Pharmacol. Ther.* **1991**, *49*, 125.

(84) Meister, A. In *Coenzymes and Cofactors: Glutathione - Chemical, Biochemical and Medical Aspects*; Dolphin, D. A., Avramovic, O., Eds.; John Wiley & Sons: New York, 1989; Vol. 3A, p 1.

(85) Blair, I. A. *Curr. Drug Metab.* **2006**, *7*, 853.

(86) Franco, R.; Schoneveld, O. J.; Pappa, A.; Panayiotidis, M. I. *Arch. Physiol. Biochem.* **2007**, *113*, 234.

(87) Douglas, K. T. In *Glutathione Conjugation: Methods and Biological Significance*; Sies, H., Ketterer, B., Eds.; Academic Press: London, United Kingdom, 1988; p 1.

(88) Schultz, T. W.; Yarbrough, J. W.; Johnson, E. L. *SAR QSAR Environ. Res.* **2005**, *16*, 313.

(89) Landsteiner, K.; Jacobs, J. *J. Exp. Med.* **1936**, *64*, 625.

(90) Roberts, D. W.; Goodwin, B. F. J.; Williams, D. L.; Jones, K.; Johnson, A. W.; Alderson, J. C. E. *Food Chem. Toxicol.* **1983**, *21*, 811.

(91) Franot, C.; Roberts, D. W.; Smith, R. G. *Chem. Res. Toxicol.* **1994**, *7*, 297.

(92) Bolton, J. L.; Turnipseed, S. B.; Thompson, J. A. *Chem.-Biol. Interact.* **1997**, *107*, 185.

(93) Gerberick, F.; Vassallo, J. D.; Bailey, R. E.; Chaney, J. G.; Morall, S. W.; Lepoittevin, J.-P. *Toxicol. Sci.* **2004**, *81*, 332.

(94) Cooksey, C. J.; Land, E. J.; Rushton, F. A. P.; Ramsden, C. A.; Riley, P. A. *Quant. Struct.-Act. Relat.* **1996**, *15*, 498.

(95) Bunnett, J. F.; Zahler, R. E. *Chem. Rev.* **1951**, *49*, 273.

(96) Rodionov, P. P.; Furin, G. G. *J. Fluorine Chem.* **1990**, *47*, 361.

(97) Alvarez-Sánchez, R.; Basketter, D.; Pease, C.; Lepoittevin, J.-P. *Chem. Res. Toxicol.* **2003**, *16*, 627.

(98) Achilleos, C.; Tailhardat, M.; Courtellemont, P.; Le Varlet, B.; Dupont, D. *Toxicol. Vitro* **2009**, *23*, 308.

(99) McCarthy, T. J.; Hayes, E. P.; Schwartz, C. S.; Witz, G. *Fundam. Appl. Toxicol.* **1994**, *22*, 543.

(100) Harder, A.; Escher, B. I.; Schwarzenbach, R. P. *Environ. Sci. Technol.* **2003**, *37*, 4955.

(101) Hermens, J.; Busser, F.; Leeuwanch, P.; Musch, A. *Toxicol. Environ. Chem.* **1985**, *9*, 219.

(102) Aptula, A. O.; Patlewicz, G.; Roberts, D. W. *Chem. Res. Toxicol.* **2005**, *18*, 1420.

(103) Deneer, J. W.; Seinen, W.; Hermens, J. L. M. *Aquat. Toxicol.* **1988**, *12*, 185.

(104) Dickens, F. *Biochem. J.* **1933**, *17*, 1141.

(105) Saunders, B. C. *Biochem. J.* **1934**, *18*, 1977.

(106) Goddard, D. R.; Schubert, M. P. *Biochem. J.* **1935**, *29*, 1009.

(107) Hulbert, P. B.; Hamoodi, N. M. *J. Pharm. Biomed. Anal.* **1990**, *8*, 1009.

(108) Gagan, E. M.; Hull, M. W.; Schultz, T. W.; Pöch, G.; Dawson, D. A. *Arch. Environ. Contam. Toxicol.* **2007**, *52*, 283.

(109) Schultz, T. W.; Yarbrough, J. W.; Pilkington, T. B. *Environ. Toxicol. Pharmacol.* **2007**, *23*, 10.

(110) Ellman, G. L. *Arch. Biochem. Biophys.* **1958**, *74*, 443.

(111) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70.

(112) Fariss, M. W.; Reed, D. J. *Methods Enzymol.* **1987**, *143*, 101.

(113) Tietze, F. *Anal. Biochem.* **1969**, *27*, 502.

(114) Esterbauer, H.; Zollner, H.; Scholz, N. *Z. Naturforsch.* **1975**, *30c*, 466.

(115) Tanii, H.; Hashimoto, K. *Toxicol. Lett.* **1982**, *11*, 125.

(116) Gerberick, F.; Vassallo, J. D.; Foertsch, L. M.; Price, B. B.; Chaney, J. G.; Lepoittevin, J.-P. *Toxicol. Sci.* **2007**, *97*, 417.

(117) Freidig, A. P.; Verhaar, H. J. M.; Hermens, J. L. M. *Environ. Toxicol. Chem.* **1999**, *18*, 1133.

(118) Clarke, E. D.; Greenhow, D. T.; Adams, D. *Pestic. Sci.* **1998**, *54*, 385.

(119) Hayes, J. D.; Pulford, D. J. *Crit. Rev. Biochem.* **1995**, *30*, 445.

(120) Guengerich, F. P. *Methods Enzymol.* **2005**, *401*, 342.

(121) Potter, D. W.; Tran, T.-B. *Toxicol. Lett.* **1992**, *62*, 275.

(122) van der Aar, E. M.; Bouwman, T.; Commandeur, J. N. M.; Vermeulen, N. P. E. *Biochem. J.* **1996**, *320*, 531.

(123) Boyland, E.; Chasseaud, L. F. *Biochem. J.* **1967**, *104*, 95.

(124) Schultz, T. W.; Yarbrough, J. W.; Hunter, R. S.; Aptula, A. O. *Chem. Res. Toxicol.* **2007**, *20*, 1359.

(125) Roberts, D. W.; Schultz, T. W.; Wolf, E. M.; Aptula, A. O. *Chem. Res. Toxicol.* **2010**, *23*, 228.

(126) Bernasconi, C. F. *Tetrahedron* **1989**, *45*, 4017.

(127) Chan, K.; Poon, R.; Brien, P. J. *J. Appl. Toxicol.* **2008**, *2008*, 1027.

(128) Ahlfors, S. R.; Sterner, O.; Hansson, C. *Skin Pharmacol. Appl. Skin Physiol.* **2003**, *16*, 59.

(129) Baker, A.; Źídek, L.; Wiesler, D.; Chmelík, J.; Pagel, M.; Novotny, M. V. *Chem. Res. Toxicol.* **1998**, *11*, 730.

(130) Metz, B.; Kersten, G. F. A.; Hoogerhout, P.; Brugghe, H. F.; Timmermans, H. A. M.; de Jong, A.; Meiring, H.; ten Hove, J.; Hennink, W. E.; Crommelin, D. J. A.; Jiskoot, W. *J. Biol. Chem.* **2004**, *279*, 6235.

(131) Aleksic, M.; Thain, E.; Roger, D.; Saib, O.; Davies, M.; Li, J.; Aptula, A.; Zazzeroni, R. *Toxicol. Sci.* **2009**, *108*, 401.

(132) Lutz, W. K. *Cancer Res. Clin. Oncol.* **1986**, *112*, 85.

(133) Schatz-Kornbrust, E.; Goldstein, B. D.; Witz, G. *Toxicologist* **1991**, *11*, 253.

(134) Latriano, L.; Witz, G.; Goldstein, B. D.; Jeffrey, A. M. *Environ. Health Perspect.* **1989**, *82*, 249.

(135) Blandamer, M. J.; Burgess, J.; Robertson, R. E.; Scott, J. M. W. *Chem. Rev.* **1982**, *82*, 259.

(136) Roberts, D. W.; Natsch, A. *Chem. Res. Toxicol.* **2009**, *22*, 592.

(137) Cronin, M. T. D.; Madden, J. C. *In Silico Toxicology - Principles and Applications*; Royal Society of Chemistry: Cambridge, United Kingdom, 2010.

(138) Freidig, A. P.; Hermens, J. L. M. *Quant. Struct.-Act. Relat.* **2001**, *19*, 547.

(139) Enoch, S. J.; Madden, J. C.; Cronin, M. T. D. *SAR QSAR Environ. Res.* **2008**, *19*, 555.

(140) Ridings, J. E.; Barratt, M. D.; Cary, R.; Earnshaw, C. G.; Eggington, C. E.; Ellis, M. K.; Judson, P. N.; Langowski, J. J.; Marchant, C. A.; Payne, M. P.; Watson, W. P.; Yih, T. D. *Toxicology* **1996**, *106*, 267.

(141) Schüürmann, G.; Kühne, R.; Kleint, F.; Ebert, R.-U.; Rothenbacher, C.; Herth, P. In *Quantitative Structure-Activity Relationships in Environmental Sciences*; Chen, F., Schüürmann, G., Eds.; VII SETAC Press: Pensacola, FL, 1997; p 93.

(142) NCCT. *OncoLogic - An Expert System for Prediction of The Carcinogenic Potential of Chemicals, 7.0*; U.S. Environmental Protection Agency: Washington, DC, 2010.

(143) Patlewicz, G.; Dimitrov, S. D.; Low, L. K.; Kern, P. S.; Dimitrova, G. D.; Comber, M. I. H.; Aptula, A. O.; Phillips, R. D.; Niemelä, J.; Madsen, C.; Wedebye, E. B.; Roberts, D. W.; Bailey, P. T.; Mekenyan, O. G. *Regul. Toxicol. Pharmacol.* **2007**, *48*, 225.

(144) Koleva, Y. K.; Madden, J. C.; Cronin, M. T. D. *Chem. Res. Toxicol.* **2008**, *21*, 2300.

(145) Nikolova, N.; Jaworska, J. *QSAR Comb. Sci.* **2003**, *22*, 1006.

(146) Diderich, R. In *In Silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; Royal Society of Chemistry: Cambridge, 2010, Chapter 16, p 385.

(147) Cronin, M. T. D.; Madden, J. C. In *In Silico Toxicology - Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; Royal Society of Chemistry: Cambridge, United Kingdom; 2010.

(148) Enoch, S. J.; Cronin, M. T. D. *Crit. Rev. Toxicol.* **2010**, *40*, 728.

(149) Cronin, M. T.; Bowers, G. S.; Sinks, G. D.; Schultz, T. W. *SAR QSAR Environ. Res.* **2000**, *11*, 301.

(150) Klopman, G.; Chakravarti, S. K.; Harris, N.; Ivanov, J.; Saiakhov, R. D. *SAR QSAR Environ. Res.* **2003**, *14*, 165.

(151) Enslein, K.; Gombar, V. K.; Blake, B. W. *Mutat. Res.* **1994**, *305*, 47.

(152) Dearden, J. C. Expert systems for toxicity prediction. In *In Silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; Royal Society of Chemistry: Cambridge, 2010; p. 478.

(153) Roberts, D. W.; Williams, D. L. *J. Theor. Biol.* **1982**, *99*, 807.

(154) Patlewicz, G.; Aptula, A. O.; Roberts, D. W.; Uriarte, E. *QSAR Comb. Sci.* **2008**, *27*, 60.

(155) Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165.

(156) Hammett, L. P. *Physical Organic Chemistry*; McGraw-Hill Book Co., Inc.: New York, 1940; Chapter VII.

(157) Taft, R. In *Steric Effects in Organic Chemistry*; Newman, M. S., Ed.; John Wiley and Sons: New York, 1956; Chapter 13.

(158) Patlewicz, G.; Roberts, D. W.; Walker, J. D. *QSAR Comb. Sci.* **2003**, 23.

(159) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; John Wiley & Sons: New York, 1979; p 350.

(160) Friedman, M.; Wall, J. S. *J. Am. Chem. Soc.* **1964**, *86*, 3735.

(161) Schüürmann, G. *Sci. Total Environ.* **1991**, *109/110*, 221.

(162) Roberts, D. W.; Aptula, A. O.; Patlewicz, G. *Chem. Res. Toxicol.* **2006**, *19*, 1228.

(163) Aptula, A. O.; Enoch, S. J.; Roberts, D. W. *Chem. Res. Toxicol.* **2009**, *22*, 1541.

(164) Aptula, N.; Roberts, D. W.; Schultz, T. W.; Pease, C. *Toxicology* **2007**, *231*, 104.

(165) Schultz, T. W.; Ralston, K. E.; Roberts, D. W.; Veith, G. D.; Aptula, A. O. *SAR QSAR Environ. Res.* **2007**, *18*, 21.

(166) Schüürmann, G. In *Predicting Chemical Toxicity and Fate*; Cronin, M. T. D., Livingstone, D. J., Eds.; Taylor and Francis: London, United Kingdom, 2004; p 85.

(167) Enoch, S. J. In *Recent Advances in QSAR Studies*; Puzyn, T. L., Leszczynski, J., Cronin, M. T. D., Eds.; Springer: New York, 2010; p 13.

(168) Enoch, S. J. The use of frontier molecular orbitals calculations in predictive reactive toxicology. In *In Silico Toxicology: Principles and Applications*; Cronin, M. T. D., Madden, J. C., Eds.; Royal Society of Chemistry: Cambridge; 2010; p 193.

(169) Karelson, M.; Lobanov, V. S.; Katritzky, A. R. *Chem. Rev.* **1996**, *96*, 1027.

(170) Geerlings, P.; De Proft, F. *Int. J. Mol. Sci.* **2002**, *3*, 276.

(171) Cronin, M. T. D.; Manga, N.; Seward, J. R.; Sinks, G. D.; Schultz, T. W. *Chem. Res. Toxicol.* **2001**, *14*, 1498.

(172) Schüürmann, G. *Environ. Toxicol. Chem.* **1990**, *9*, 417.

(173) Schmitt, H.; Altenburger, R.; Jastorff, B.; Schüürmann, G. *Chem. Res. Toxicol.* **2000**, *13*, 441.

(174) Schultz, T. W.; Netzeva, T. I.; Roberts, D. W.; Cronin, M. T. D. *Chem. Res. Toxicol.* **2005**, *18*, 330.

(175) Passerini, L. In *Quantitative Structure-Activity Relationships (QSAR) Models of Mutagens and Carcinogens*; Benigni, R., Ed.; CRC Press: Boca Raton, FL, 2003; Chapter 3, p 81.

(176) Parr, R. G.; v. Szentpály, L.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922.

(177) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801.

(178) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.

(179) Enoch, S. J.; Roberts, D. W.; Cronin, M. T. D. *Chem. Res. Toxicol.* **2009**, *22*, 1447.

(180) Campodónico, P. R.; Contreras, R. *Bioorg. Med. Chem.* **2008**, *16*, 3184.

(181) Campodónico, P. R.; Fuentealba, P.; Castro, E. A.; Santos, J. G.; Contreras, R. *J. Org. Chem.* **2005**, *70*, 1754.

(182) Domingo, L. R.; Pérez, P.; Contreras, R. *J. Org. Chem.* **2003**, *68*, 6060.

(183) Pérez, P.; Toro-Labbe, A.; Contreras, R. *J. Am. Chem. Soc.* **2001**, *123*, 5527.

(184) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142.

(185) Natsch, A.; Emter, R.; Ellis, G. *Toxicol. Sci.* **2009**, *107*, 106.

(186) Hermens, J. L. M. *Environ. Health Perspect.* **1990**, *87*, 219.

(187) Russom, C. L.; Bradbury, S. P.; Broderius, S. J.; Hammermeister, D. E.; Drummond, R. A. *Environ. Toxicol. Chem.* **1997**, *16*, 948.

(188) Benigni, R.; Bossa, C.; Tcheremenskaia, O.; Worth, A. *Development of Structural Alerts for the In Vivo Micronucleus Assay in Rodents*; Institute for Health and Consumer Protection, Joint Research Centre, European Commission: Ispra, Italy, 2009.

(189) Estrada, E.; Molina, E. *J. Mol. Graphics Modell.* **2006**, *25*, 275.

(190) Treinen-Moslen, M. In *Casarett & Doull's Toxicology: The Basic Science of Poisons*, 6th International ed.; Klassen, C. D., Ed.; McGraw-Hill: New York, 2001; p 471.

(191) Russmann, S.; Kullak-Ublick, G. A.; Grattagliano, I. *Curr. Med. Chem.* **2009**, *16*, 3041.

(192) Hulzebos, E.; Walker, J. D.; Gerner, I.; Schlegel, K. *QSAR Comb. Sci.* **2005**, *24*, 332.

(193) Barbin, A.; Brésil, H.; Croisy, A.; Jacquignon, P.; Malaveille, C.; Montesano, R.; Bartsch, H. *Biochem. Biophys. Res. Commun.* **1975**, *67*, 596.

(194) Gómez-Bombarelli, R.; González-Pérez, M.; Pérez-Prior, M. T.; Manso, J. A.; Calle, E.; Casado, J. *Chem. Res. Toxicol.* **2008**, *21*, 1964.

(195) Manso, J. A.; Pérez-Prior, M. T.; García-Santos, M. D. P.; Calle, E.; Casado, J. *Chem. Res. Toxicol.* **2005**, *18*, 1161.

(196) Fernández-Rodríguez, E.; Manso, J. A.; Pérez-Prior, M. T.; García-Santos, M. D. P.; Calle, E.; Casado, J. *Int. J. Chem. Kinet.* **2007**, *39*, 591.

(197) Verhaar, H. J. M.; Rorije, E.; Borkent, H.; Seinen, W.; Hermens, J. L. M. *Environ. Toxicol. Chem.* **1996**, *15*, 1011.

(198) Bardos, T. J.; Datta-Gupta, N.; Hebborn, P.; Triggle, D. J. *J. Med. Chem.* **1965**, *8*, 167.

(199) Deneer, J. W.; Sinnige, T. L.; Seinen, W.; Hermens, J. L. M. *Aquat. Toxicol.* **1988**, *13*, 195.

(200) Eder, E.; Neudecker, T.; Lutz, D.; Henschler, D. *Chem.-Biol. Interact.* **1982**, *38*, 303.

(201) Neudecker, T.; Lutz, D.; Eder, E.; Henschler, D. *Biochem. Pharmacol.* **1980**, *29*, 2611.

(202) Eder, E.; Neudecker, T.; Lutz, D.; Henschler, D. *Biochem. Pharmacol.* **1980**, *29*, 993.

(203) Heo, C. K. M.; Bunting, J. W. *J. Org. Chem.* **1992**, *57*, 3570.

(204) Esterbauer, H.; Ertl, A.; Scholz, N. *Tetrahedron* **1976**, *32*, 285.

(205) Eyer, P. In *Biological Oxidation of Nitrogen in Organic Molecules*; Gorrod, J. W., Ed.; VCH Publishers: Chichester, UK, 1985; Chapter 43, p 386.

(206) Drobnica, L.; Šturdík, E. *Biochim. Biophys. Acta* **1979**, *585*, 462.

(207) LoPachin, R. M.; Gavin, T.; Geohagen, B. C.; Das, S. *Toxicol. Sci.* **2007**, *98*, 561.

(208) LoPachin, R. M.; Geohagen, B. C.; Gavin, T. *Toxicol. Sci.* **2009**, *107*, 171.

(209) Shimozu, Y.; Shibata, T.; Ojika, M.; Uchida, K. *Chem. Res. Toxicol.* **2009**, *22*, 957.

(210) Saunders, B. C. *Biochem. J.* **1933**, *27*, 397.

(211) Klehr, H.; Eyer, P.; Schäfer, W. *Biol. Chem. Hoppe-Seyler* **1985**, *366*, 755.

(212) Bernasconi, C. F.; Killion, R. B., Jr. *J. Am. Chem. Soc.* **1988**, *110*, 7506.

(213) Bernasconi, C. F.; Ketner, R. J.; Chen, X.; Rappoport, Z. *J. Am. Chem. Soc.* **1998**, *120*, 7461.

(214) Bernasconi, C. F.; Leyes, A. E.; Rappoport, Z. *J. Org. Chem.* **1999**, *64*, 2897.

(215) Chipinda, I.; Ajibola, R. O.; Morakinyo, M. K.; Ruwona, T. B.; Simoyi, R. H.; Siegel, P. D. *Chem. Res. Toxicol.* **2010**, *23*, 918.

(216) Um, I.-H.; Lee, E.-J.; Seok, J.-A.; Kim, K.-H. *J. Org. Chem.* **2005**, *70*, 7530.

(217) Um, I.-H.; Hwang, S.-J.; Lee, E.-J. *Bull. Korean Chem. Soc.* **2008**, *29*, 767.

(218) Chan, K.; O'Brien, P. J. *J. Appl. Toxicol.* **2008**, *28*, 1004.

(219) Natsch, A.; Gfeller, H. *Toxicol. Sci.* **2008**, *106*, 464.

(220) Freidig, A. P. Ph.D. thesis, Utrecht University, Utrecht, The Netherlands, 2000.

(221) Kosover, E. M. *Chem. Biochem. React.* **2009**, *6*, 549.

(222) Doorn, J. A.; Petersen, D. R. *Chem.-Biol. Interact.* **2003**, *143−144*, 93.

(223) Wondrousch, D.; Böhme, A.; Thaens, D.; Ost, N.; Schüürmann, G. *J. Phys. Chem. Lett.* **2010**, *1*, 1605.

(224) Portoghese, P. S.; Kedziora, G. S.; Larson, D. L.; Bernard, B. K.; Hall, R. L. *Food Chem. Toxicol.* **1989**, *27*, 773.

(225) Barshteyn, N.; Elfarra, A. A. *Chem. Res. Toxicol.* **2009**, *22*, 918.

(226) Woker, G.; Antener, I. *Helv. Chim. Acta* **1937**, *20*, 1260.

(227) Reichardt, P.; Schreiber, A.; Wichmann, G.; Metzner, G.; Efer, J.; Raabe, F. *Environ. Toxicol.* **2003**, *18*, 29.

(228) Cai, J.; Bhatnagar, A.; Pierce, W. M., Jr. *Chem. Res. Toxicol.* **2009**, *22*, 708.

(229) Sykes, P. *Guidebook to Mechanism in Organic Chemistry*, 6th ed.; Pearson Education Ltd.: Essex, United Kingdom, 1986; p 358.

(230) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793.

(231) Geerlings, P.; de Proft, F.; Ayers, P. W. In *Theoretical Aspects of Chemical Reactivity*; Toro-Labbé, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2007; p 1.

(232) Politzer, P.; Murray, J. S. In *Theoretical Aspects of Chemical Reactivity*; Toro-Labbé, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2007; p 119.

(233) Chattaraj, P. K.; Roy, D. R. *Chem. Rev.* **2007**, *107*, PR46.

(234) Morell, C.; Grand, A.; Gutiérrez-Oliva, S.; Toro-Labbé, A. In *Theoretical Aspects of Chemical Reactivity*; Toro-Labbé, A., Ed.; Elsevier: Amsterdam, The Netherlands, 2007; p 101.

(235) Klamt, A. *Chemosphere* **1993**, *26*, 1273.

(236) Klamt, A. *Chemosphere* **1995**, *32*, 717.

(237) Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G. *J. Phys. Chem. A* **2009**, *113*, 10104.

(238) Fukui, K.; Yonezawa, T.; Nagata, C. *Bull. Chem. Soc. Jpn.* **1954**, *27*, 423.