

# **InSilicoTox Project**



## **Deliverable 2 (D2)**

### **Novel Descriptors for Reactivity**

**June 2009**

**European Union Sixth Framework Programme Marie Curie Actions**

**Project No.: 42328**

**Project Acronym: InSilicoTox**

**Project Full Name: *In Silico* Tools to Reduce the Use of Animals in Toxicity Testing for Bioreactive Chemicals**

#### **Deliverable Prepared by**

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

### **Important Information**

This document is a deliverable of the InSilicoTox Marie Curie Transfer of Knowledge Project. The copyright of this material belongs to Liverpool John Moores University. This document was prepared in good faith, Liverpool John Moores University or the authors of this document are not responsible for any inaccuracies or errors.

## **D2. Novel Descriptors for Reactivity**

### **1. Summary**

The prediction of the toxicity of reactive chemicals (those that bind covalently to biological macromolecules) requires the use of chemical descriptors that encapsulate how covalently reactive a given chemical is. Quantum chemical descriptors are related to the electronic structure of reactive chemicals and to the chemical mechanisms that are involved in covalent bond formation between biological nucleophiles and electrophilic chemicals. A number of computational methods are available for the calculation of theoretical descriptors; readers are directed to some excellent reviews for more information.<sup>1-3</sup> A list of commonly used quantum chemical descriptors used in QSAR applications is given in the literature (see Table 1),<sup>4</sup> and current conceptual DFT descriptors have been linked successfully with chemical reactivity.<sup>5</sup>

**Table 1.** Selection of Molecular Descriptors Used to Model Chemical Reactivity

Descriptor	Explanation	Ref.
<i>Global quantum chemical descriptors</i>		
$E_{\text{HOMO}}$	Energy of the highest occupied molecular orbital	1
$E_{\text{LUMO}}$	Energy of the lowest unoccupied molecular orbital	1
$IP (\approx -E_{\text{HOMO}}^{\text{a}})$	Ionisation potential for removing a electron from a molecular system X ( $X \rightarrow X^+ + e^-$ )	1
$EA (\approx -E_{\text{LUMO}}^{\text{a}})$	Electron affinity attaching an additional electron to a molecular system X ( $X + e^- \rightarrow X^-$ )	1
$\mu = \left( \frac{\partial E_{\text{el}}}{\partial N} \right)_v$	Chemical potential, defined as the change in electronic energy $E_{\text{el}}$ upon change in total number of electrons $N$	6
$\chi = -\mu \approx -\frac{1}{2}(E_{\text{HOMO}} + E_{\text{LUMO}})^{\text{a}}$	Absolute electronegativity	6
$\eta = -\left( \frac{\partial \mu}{\partial N} \right)_v \approx -(E_{\text{HOMO}} - E_{\text{LUMO}})^{\text{a}}$	Molecular hardness, defined as the change in chemical potential $\mu$ upon change in total number of electrons $N$	6
$S = \frac{1}{2\eta}$	Molecular softness	6
$\alpha$	Molecular polarisability; note that molecules arrange themselves toward a state of minimum polarisability and maximum hardness	6
$\omega_{\text{el}} = \frac{\mu^2}{2\eta} = \frac{\chi^2}{2\eta}$	Electrophilicity index	7
<i>Charge distribution</i>		
$Q_{\text{A}}(r)$	Net atomic charges (at atom $r$ )	1
$PSA$	Polar surface area, describing the spacial surface density distribution	1
$\mu$	Molecular dipole moment	1
$\sigma(\mathbf{r}) = \rho(\mathbf{r}) / N$	Electronic density ( $\rho$ ) per particle shape function, normalised to 1, expresses the	8

distribution of the total number of electrons  $N$  in a system among different parts  $\mathbf{r}$  of a system

### Site specific molecular descriptors

$$\bar{I}(\mathbf{r}) = \sum_i \frac{\rho_i(\mathbf{r})|\varepsilon_i|}{\rho(\mathbf{r})}$$

Local ionisation energy

9

$\rho(\mathbf{r})$  = Total electronic density at point  $\mathbf{r}$ ;  $\rho_i(\mathbf{r})$  = Electronic density of occupied molecular orbital (MO)  $i$  with MO energy  $\varepsilon_i$

$$f^+(r) = \left( \frac{\partial \rho(r)}{\partial N} \right)_{v(r)}^+ \approx \rho_{N+1}(r) - \rho_N(r)$$

Electrophilic Fukui function, defined as the change in electron density  $\rho$  at atom  $r$  upon addition of electrons to the system ( $N$  = electron number)

6

$$\omega_{el}(r) = \omega_{el} \times f^+(r)$$

Local electrophilicity index

10

$$\Delta f(r) = f^+(r) - f^-(r)$$

$$\Delta f(r) \approx \rho_{\text{LUMO}}(r) - \rho_{\text{HOMO}}(r)^a$$

Reactivity-selectivity descriptor or dual descriptor, while  $f^+(r)$  measures reactivity towards nucleophilic and  $f^-(r)$  towards electrophilic attacks; therefore, electrophilic sites are identified by  $\Delta f(r) > 0$ .  $\rho_{\text{LUMO}}$  and  $\rho_{\text{HOMO}}$  are the electron densities of the LUMO and HOMO orbitals, respectively

11

$$FN(r) \left( = \sum_{\mu(r)} c_{\mu k}^2 \right)$$

Frontier orbital electron density (towards a nucleophile)

12

$c$  = Molecular orbital (MO) coefficients;  $\mu$  = Atomic orbitals (AO) of atom  $r$ ;  $k$  = LUMO

$$DN(r) \left( = 2 \sum_{k=\text{LUMO}}^{\max} \sum_{\mu(r)} \frac{c_{\mu k}^2}{\alpha - \varepsilon_k} \right)$$

Acceptor delocalisability (towards a nucleophile)

12

$c$  = MO coefficients;  $\mu$  = AOs of atom  $r$ ;  $k$  = unoccupied MO with MO energy  $\varepsilon_k$ ;  $\alpha$  = relevant orbital energy of attacking nucleophile ( $\approx E_{\text{HOMO}}$  of nucleophile<sup>a</sup>)

$$SN(r) \left( = 2 \sum_{k=\text{LUMO}}^{\max} \sum_{\mu(r)} \frac{c_{\mu k}^2}{(-\varepsilon_k)} \right)$$

Superdelocalisability (towards a nucleophile)

12

$$ALP(r) \left( = 4 \sum_{i=\text{HOMO}}^1 \sum_{k=\text{LUMO}}^N \sum_{\mu(r)} \frac{c_{\mu i}^2 c_{\mu k}^2}{(\varepsilon_i - \varepsilon_k)} \right)$$

Atomic self-polarisability

12

<sup>a</sup> According to Koopman's theorem.

A common way to describe the reactivity of an electrophilic compound in toxicity is the molecular electrophilicity index  $\omega$ . However, because chemical reactivity is induced by site-specific features (e.g. functional groups) within a molecule, an adequate site specific reactivity descriptor is desired for the prediction of reactivity, with regard to local electronic properties. In this deliverable, we describe a new site specific reactivity descriptor  $\omega_{EQ}$ , which is the local counterpart of the molecular electrophilicity index. The great advantage of site-specific descriptors over molecular descriptors is that they allow for a reliable reactivity prediction, even if the molecule contains more than one potential reaction site.

## 2. Development of a Local Electrophilicity Index

A common way to describe the reactivity of an electrophilic compound in toxicity is the molecular electrophilicity index  $\omega$ .<sup>7</sup> It can be calculated by the following equation:

$$\omega = \frac{\chi^2}{2\eta} = \frac{E_{HOMO}^2 + 2E_{HOMO}E_{LUMO} + E_{LUMO}^2}{4(E_{LUMO} - E_{HOMO})} \quad (1)$$

In this equation  $\chi$  is the molecular electrophilicity,  $\eta$  the molecular hardness, both of which can be replaced in terms of the energies of the highest occupied molecular orbital,  $E_{HOMO}$ , and the lowest unoccupied molecular orbital,  $E_{LUMO}$ , by applying Koopmans' theorem.<sup>10</sup> This descriptor has been widely used: Domingo *et al* calculated the electrophilicity index for a series of activated ethylenes.<sup>13, 14</sup> Enoch *et al* described a method to rank the reactivity of Michael acceptors by the electrophilicity index, using this ranking for a read-across approach to predict skin

sensitisation and respiratory sensitisation capabilities of compounds.<sup>15, 16</sup> In a related study, the electrophilic index was used to model the cytotoxicity of a series of sugars acting via Michael addition.<sup>17, 18</sup>

As here a site specific reactivity descriptor is desired for the prediction of reactivity, a local counterpart of the electrophilicity index is applied. In principle, this can be provided by multiplying the molecular electrophilicity index  $\omega$  with an atom centred Fukui index  $f^*$ , as described by Chattaraj *et al* (2001).<sup>10</sup> This is a method to obtain the most reactive site within a molecule or to describe trends in a qualitative manner. Because of the difficulties which arise from the application of several net atomic charge models needed for the calculation of the Fukui index  $f^*$ , we have followed a different route. Note that atomic charges are not quantum chemical observables.

We have defined the local electrophilicity index  $\omega_{\text{EQ}}(r)$  at an atomic reaction site  $r$  in the following way:

$$\omega_{\text{EQ}}(r) = \frac{EQ_{\text{occ}}(r)^2 + 2EQ_{\text{occ}}(r)EQ_{\text{vac}}(r) + EQ_{\text{vac}}(r)^2}{4(EQ_{\text{vac}}(r) - EQ_{\text{occ}}(r))} \quad (2)$$

In this equation,  $EQ_{\text{occ}}(r)$  is the effective (charge limited) electron donor energy, and  $EQ_{\text{vac}}(r)$  the analogous electron acceptor energy. These, and related, descriptors were originally invented by Klamt (1993).<sup>19</sup> They were recently reintroduced by Schüürmann and co-workers, in order to predict the rate constants of indirect photolysis by  $\bullet\text{OH}$  radicals,<sup>20, 21</sup> and the affinity to form hydrogen bonds.<sup>22-</sup>  
<sup>24</sup> Exact definitions for the electron donor and acceptor energies are given below. Furthermore, we define a maximum electrophilicity index  $\omega_{\text{max}}$  which accounts for the observation that after a certain point higher (calculated) electrophilicity indices do not

reflect higher (experimental) rate constants, because reaction rates become diffusion controlled, as previously observed by Aizman *et al.*<sup>25</sup>

### 3. Local Electron Donor and Acceptor Energies

Both the donor energy  $EQ_{\text{occ}}$  and the acceptor energy  $EQ_{\text{vac}}$  use a linear weight function  $w$ . This depends on whether the charge amount  $Q$ , induced by peptide or DNA adduct formation, is reached in the atomic orbitals considered or not. In general, Greek indices  $\mu, \rho, \dots$  denote the atomic orbitals, while Latin letters  $i, k, \dots$  run over the molecular orbitals (MOs) of an atomic centre  $r$ .

$$EQ_{\text{occ}}(Q, r) = \frac{1}{Q} \sum_{i=1}^{\text{HOMO}} E_i \cdot w_i(Q, r); \quad EQ_{\text{vac}}(Q, r) = \frac{1}{Q} \sum_{k=\text{LUMO}}^{\text{max}} E_k \cdot w_k(Q, r) \quad (3)$$

with:

$$w_i(Q_{\text{occ}}, r) = \begin{cases} p_i(r) & \text{if : } Q_{\text{occ}} > P_{\text{occ}}(i, r) + p_i(r) \\ Q_{\text{occ}} - P_{\text{occ}}(i, r) & \text{if : } P_{\text{occ}}(i, r) \leq Q_{\text{occ}} \leq P_{\text{occ}}(i, r) + p_i(r) \\ 0 & \text{if : } Q_{\text{occ}} < P_{\text{occ}}(i, r) \end{cases} \quad (4)$$

$$w_i(Q_{\text{vac}}, r) = \begin{cases} p_i(r) & \text{if : } Q_{\text{vac}} > P_{\text{vac}}(i, r) + p_i(r) \\ Q_{\text{vac}} - P_{\text{vac}}(i, r) & \text{if : } P_{\text{vac}}(i, r) \leq Q_{\text{vac}} \leq P_{\text{vac}}(i, r) + p_i(r) \\ 0 & \text{if : } Q_{\text{vac}} < P_{\text{vac}}(i, r) \end{cases} \quad (5)$$

The electron population  $p_i$  of an occupied molecular orbital  $i$ , or the electron space  $p_k$  available in an unoccupied molecular orbital  $k$ , is quantified by taking twice the sum



of the square of the coefficients  $c_{\mu i}$  ( $c_{\sigma k}$ ) of the LCAO-MO approach (linear combination of atomic orbitals to molecular orbitals) at an atomic centre  $r$ :

$$p_i(r) = 2 \cdot \sum_{\mu(r)} (c_{\mu i})^2 ; p_k(r) = 2 \cdot \sum_{\rho(r)} (c_{\rho k})^2 \quad (6)$$

The partial charge of selected occupied or vacant MOs  $i, k$  at an atomic centre  $r$  is:

$$P_{\text{occ}}(i, r) = \sum_{j=\text{HOMO}}^{i+1} p_j(r) ; P_{\text{vac}}(k, r) = \sum_{l=\text{LUMO}}^{k-1} p_l(r) \quad (7)$$

#### 4. Application in Reactivity and Toxicity Prediction

The application of the local electrophilicity index  $\omega_{\text{EQ}}(r)$  in reactivity and toxicity prediction will be discussed in further detail in the following deliverables:

- Deliverable D3: (Q)SARs for the Prediction of Reactive Toxicity.

Quantitative structure-activity relationship models for the prediction of reactivity and reactive toxicity, based on the calculation of the novel local electrophilicity index  $\omega_{\text{EQ}}$ , will be implemented into an InSilicoTox software tool, as described in the following deliverable:

- Deliverable D5: Software Tool for the Prediction of Reactive Toxicity.

## 5. Literature

1. Schüürmann, G., Quantum chemical descriptors in structure-activity relationships - calculation, interpretation and comparison of methods. In *Predicting Chemical Toxicity and Fate*, Cronin, M. T. D.; Livingstone, D. J., Eds. Taylor and Francis: London, United Kingdom, 2004; pp 85-149.
2. Enoch, S. J., The use of quantum mechanical derived descriptors in computational toxicology. In *Recent Advances in QSAR Studies*, Puzyn, T. L.; Leszczynski, J.; Cronin, M. T. D., Eds. Springer: New York, NY, USA, 2010; pp 13-26.
3. 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. RSC: Cambridge, 2010; Vol. in press.
4. Karelson, M.; Lobanov, V. S.; Katritzky, A. R., Quantum-chemical descriptors in QSAR/QSPR studies. *Chem. Rev.* **1996**, *96*, 1027-1043.
5. Geerlings, P.; De Proft, F., Chemical reactivity as described by quantum chemical methods. *Int. J. Mol. Sci.* **2002**, *3*, 276-309.
6. Geerlings, P.; De Proft, F.; Langenaeker, W., Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793-1873.
7. Parr, R. G.; v. Szentpály, L.; Liu, S., Electrophilicity index. *J. Am. Chem. Soc.* **1999**, *121*, 1922-1924.
8. Geerlings, P.; de Proft, F.; Ayers, P. W., Chemical reactivity and the shape function. In *Theoretical Aspects of Chemical Reactivity*, Toro-Labbé, A., Ed. Elsevier: Amsterdam, The Netherlands, 2007; pp 1-18.
9. Politzer, P.; Murray, J. S., The average local ionization energy: concepts and applications. In *Theoretical Aspects of Chemical Reactivity*, Toro-Labbé, A., Ed. Elsevier: Amsterdam, The Netherlands, 2007; pp 119-138.
10. Chattaraj, P. K.; Roy, D. R., Update 1 of: Electrophilicity Index. *Chem. Rev.* **2007**, *107*, PR46-PR74.
11. Morell, C.; Grand, A.; Gutiérrez-Oliva, S.; Toro-Labbé, A., Using the reactivity-selectivity descriptor  $Df(r)$  in organic chemistry. In *Theoretical Aspects of Chemical Reactivity*, Toro-Labbé, A., Ed. Elsevier: Amsterdam, The Netherlands, 2007; pp 101-118.
12. Fukui, K.; Yonezawa, T.; Nagata, C., Theory of substitution in conjugated molecules. *Bull. Chem. Soc. Japan* **1954**, *27*, 423-427.
13. Domingo, L. R.; Aurell, M. J.; Pérez, P.; Contreras, R., Quantitative characterization of the global electrophilicity power of common diene/dienophile pairs in Diels–Alder reactions *Tetrahedron* **2002**, *58*, 4417-4423.
14. Domingo, L. R.; Pérez, P.; Contreras, R., Electronic contributions to the  $\sigma_p$  parameter of the Hammett equation. *J. Org. Chem.* **2003**, *68*, 6060-6062.
15. Enoch, S. J.; Cronin, M. T. D.; Schultz, T. W.; Madden, J. C., Quantitative and mechanistic read across for predicting the skin sensitization potential of alkenes acting via Michael addition. *Chem. Res. Toxicol.* **2008**, *21*, 513-520.
16. Enoch, S. J.; Roberts, D. W.; Cronin, M. T. D., Electrophilic reaction chemistry of low molecular weight respiratory sensitizers. *Chem. Res. Toxicol.* **2009**, *22*, 1447-1453.

17. Campodónico, P. R.; Contreras, R., Structure–reactivity relationships for electrophilic sugars in interaction with nucleophilic biological targets. *Bioorg. Med. Chem.* **2008**, *16*, 3184-3190.
18. Campodónico, P. R.; Fuentealba, P.; Castro, E. A.; Santos, J. G.; Contreras, R., Relationships between the electrophilicity index and experimental rate coefficients for the aminolysis of thiolcarbonates and dithiocarbonates. *J. Org. Chem.* **2005**, *70*, 1754-1760.
19. Klamt, A., Estimation of gas-phase hydroxyl radical rate constants of organic compounds from molecular orbital calculations. *Chemosphere* **1993**, *26*, 1273-1289.
20. Böhnhardt, A.; Kühne, R.; Ebert, R.-U.; Schüürmann, G., Indirect photolysis of organic compounds: prediction of OH reaction rate constants through molecular orbital calculations. *J. Phys. Chem. A* **2008**, *112*, 11391-11399.
21. Böhnhardt, A.; Kühne, R.; Ebert, R.-U.; Schüürmann, G., Predicting rate constants of OH radical reactions with organic substances: advances for oxygenated organics through a molecular orbital HF/6-31G\*\* approach *Theor. Chem. Acc.* **2010**, *in press*.
22. Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G., Modeling the H bond donor strength of -OH, -NH, and -CH sites by local molecular parameters. *J. Comput. Chem.* **2009**, *30*, 1454-1464.
23. Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G., Prediction of the intrinsic hydrogen bond acceptor strength of chemical substances from molecular structure. *J. Phys. Chem. A* **2009**, *113*, 10104-10112.
24. Schwöbel, J.; Ebert, R.-U.; Kühne, R.; Schüürmann, G., Prediction of the intrinsic hydrogen bond acceptor strength of organic compounds by local molecular parameters. *J. Chem. Inf. Model.* **2009**, *49*, 956-962.
25. Aizman, A.; Contreras, R.; Pérez, P., Relationship between local electrophilicity and rate coefficients for the hydrolysis of carbenium ions. *Tetrahedron* **2005**, *61*, 889-895.