Bioorganic & Medicinal Chemistry 17 (2009) 4153–4159

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

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

Beyond physical properties—Application of Abraham descriptors and LFER analysis in agrochemical research

Eric D Clarke *

Syngenta, International Research Centre, Jealott's Hill, Bracknell, Berkshire, RG42 6EY, UK

ARTICLE INFO

Article history:
Received 9 December 2008
Revised 18 February 2009
Accepted 28 February 2009
Available online 5 March 2009

Keywords:
Property-profiles
Abraham descriptors
LFER
Agrochemicals

ABSTRACT

This paper provides an overview of the Abraham approach to the determination of molecular descriptors for agrochemicals, and their potential in defining bioavailability related property-profiles and processes. The prediction of Abraham descriptors from structure has been demonstrated using agrochemical products and a simple guide to bioavailability defined in terms of the McGowan volume descriptor V ($\leqslant 3$), the hydrogen bond acidity descriptor A ($\leqslant 1$), and the hydrogen bond basicity descriptor, B ($\leqslant 3$). The use of measured physical property data to yield experimental descriptors is illustrated for 28 representative agrochemicals. Experimental descriptors were used to derive linear free energy relationships for soil binding and root uptake, to provide insights to these processes.

1. Introduction

Effective agrochemicals have diverse balances of intrinsic potency and bioavailability appropriate to their end use as herbicides, insecticides and fungicides. Consequently, a key challenge in agrochemical research is to identify the potential of a chemical class to achieve the potency-bioavailability balance required to meet a given target profile. Target profiles for agrochemicals can be complex, often involving selective control of multiple pest species in a wide variety of crops following application to soil and/or plant foliage.[1] Bioavailability itself can be considered a balance of mobility and stability related properties, the understanding of which is important at all stages of the lead generation, optimisation and product development process. Crop protection research chemists and biologists attempt to explore and define these balances through structure–activity, structure–property and structure–reactivity relationships; with expert support from biochemists, computational and physical chemists. Thus, a similar approach is taken in agrochemical and pharmaceutical research to deliver new crop protectants and drugs, albeit for different end uses.[2]

In 1997 Lipinski et al. published their highly cited paper defining physical chemical and structural properties profiles for optimal oral availability of drugs.[3] This gave rise to the well known Lipinski 'rule of 5' which placed upper limits for four properties; namely 500 for molecular weight, 5 for log$P$ octanol, 5 for the number of hydrogen bond donors and 10 for the number of hydrogen bond acceptors. Since then, the concepts of drug-likeness and lead-likeness have been well explored, notably in papers by Ghose,[4] Teague,[5] Hann,[6] Oprea,[7] Wenlock[8] and their co-workers, and recently reviewed by Muresan and Sadowski.[9] During the same time period physical and molecular property profiles were defined for agrochemicals, first in a SCI talk by Briggs[10] to give 'ground rules of 3', then in papers by Tice[11] and Clarke and Delaney.[12] With due regard to the complexities and subtleties of these studies, it would be reasonable to suggest that property profiles for agrochemicals and drugs, be they products or leads, have more similarities than differences. The shifts in profiles that were observed may well reflect the relative contribution of properties to mobility and/or metabolism in the particular bioavailability balance required. Taken overall, these studies suggest a pragmatic guide to bioavailability for 'bio-active' compounds could be as simple as molecular weight $\leqslant 450$, log$P$ octanol $\leqslant 4.5$ and number of hydrogen bond donors $\leqslant 3$.

The upper and lower limits of bioavailability profiles have included a wide range of whole molecule properties, such as molecular weight, octanol–water and alkane–water partition coefficients (log$P$, Δlog$P$), aqueous solubility (logSw), acid–base dissociation constant (pK_a); and simple structure-based counts of rotatable bonds, H-bond donors, H-bond acceptors, hetero-atom and aromatic proportions. In agrochemical research, aspects of these property guides have been applied to compound acquisition for screening and in the design of lead generation combinatorial libraries.[13] However, the differences in overall property profiles for herbicides, fungicides and insecticides have arguably proved too subtle for effective use by organic chemists in their lead optimisation programs. Ultimately, we will need to go beyond physical properties to explore chemically intuitive molecular descriptors

---

* Corresponding author. Tel.: +44 1344 414067; fax: +44 1344 455629.
E-mail address: eric.clarke@syngenta.com

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.02.061

to influence organic chemists in the design of new agrochemicals with optimal potency-bioavailability profiles.

Properties termed as mobility or transport, which relate to the passive movement of a compound from one phase to another, can be described via a small number of molecular descriptors, typically 5 to 6.[14,15] Whilst there are differences in the individual descriptors identified, for example, through independent research led by Klamt and Abraham, they do encode similar chemical information re dispersion interactions, dipolar/polarisability interactions, hydrogen bond acid and hydrogen bond base interactions and compound size.[16] In particular, Abraham and co-workers have made considerable progress over the past 20 years defining transport processes in terms of linear free energy relationships (LFER), the full details of which have recently been covered in two excellent reviews.[17,18]

The LFER method is based upon two solvation equations:

$$SP = c + eE + sS + aA + bB + vV \quad (1)$$

$$SP = c + eE + sS + aA + bB + lL \quad (2)$$

Eq. 1 is used for the movement of a compound between condensed phases such as octanol and water, whereas Eq. 2 is applied to transfer between the gas phase and a condensed phase such as air and water. Thus it is immediately apparent that these equations have the potential to expand our understanding of physical parameters such as octanol–water partition coefficient ($\log P$ octanol) and Henry's law constant ($H = RT/K_w$ where K_w is the water–air partition coefficient). In fact, these equations have the potential to predict and provide chemical insight to any transport related process dependant upon the equilibrium transfer or rate of transfer of a compound between phases. The terms *E*, *S*, *A*, *B*, *V* and *L* define the compound of interest (solute, *S*) and the coefficients *e*, *s*, *a*, *b*, *v* and *l* the solute property (*P*); *SP* therefore being the property of interest for a given solute, often expressed on a $\log_{10}$ scale.

The solute descriptors were carefully selected by Abraham, both with respect to the requirements of the cavity model for solvation and scope for chemical interpretation.[19] The interactions they cover are:

E: excess molar refraction relating to dispersion force interactions due to π and n-electrons relative to an alkane of equivalent molar volume, V—expressed as $(\text{cm}^3\ \text{mol}^{-1})/10$.

S: polarity/polarisability due to solute–solvent interactions between dipoles and induced dipoles.

A: overall hydrogen bond acidity defining the strength and number of H-bonds formed by solute donor groups with lone pairs of solvent acceptor groups.

B: overall hydrogen bond basicity defining the strength and number of H-bonds formed by lone pairs of solute acceptor groups with solvent donor groups.

V: McGowen characteristic volume—expressed as $(\text{cm}^3\ \text{mol}^{-1})/100$.

L: alternative size descriptor for gas-condensed phase systems defined as the gas–hexadecane partition coefficient at 298 K.

The non-polar descriptors *V* and *E* can be reliably calculated from structure, leaving the polar descriptors *A*, *B* and *S* to be determined from experimental data. This can be via analysis of three or more partition coefficients obtained from organic–water systems with significant differences in one or more of the coefficients of their respective solvation equations. Typically, the 'Solver' tool in Microsoft Excel is used to give the best fit for the 3 polar descriptors. Note that the *E* and *V* descriptors are scaled by a factor of 1/10 and 1/100, respectively to make all descriptors of similar numeric order.

For agrochemicals, the determination of Abraham descriptors from physical properties was first described in 2000 for 7 examples of 3-phenyl-1,1-dimethylurea herbicides.[20] Since then little has been published on this topic, the most recent paper in 2007 giving experimental descriptors for 19 commercial 1,3,5-triazine herbicides and effective predictions of Henry's law constant for this chemical class.[21] Though other material does exists on the application of Abraham descriptors and LFER to agrochemical research much of it is in the form of conference talks and posters. This paper will draw on this material and new work to provide a current overview of Abraham descriptors for agrochemicals and the processes for which LFER equations can be generated.

2. Results and discussion

2.1. Predicted Abraham descriptors for agrochemicals

Abraham descriptors can be predicted using the ABSOLV program,[22] which is a group contribution method based on a structural fragment analysis of measured descriptors for several thousand diverse chemicals compiled by Professor Abraham at University College London. In 2003, the version of ABSOLV then available from Sirius Analytical Instruments was used in the profiling of ~18,000 compounds representing agrochemical screen inputs, hits, leads and products. Predictions were made for the *V*, *E*, *A*, *B* and *S* descriptors, relating to transport properties between condensed phases, but only the *A* and *B* descriptors were selected for use in the comprehensive analysis of agrochemicals published.[12,23] For modern agrochemical products, defined as post 1967 compounds relative to their date of Chemical Abstract Service Registry Number (CAS RN), the range of *A* and *B* descriptors are given in Table 1. Modern agrochemicals as a whole, or sub-divided into herbicides, fungicides and insecticides had quite narrow and similar ranges for *A* and *B*. This was also the case for screen inputs, hits and leads, which all had predicted mean values of 0.3 for *A* and 1.4 for *B*. The

Table 1
Predicted 10**th** and 90**th** percentile limits and mean values for the *A* and *B* descriptor for modern agrochemical products

Descriptor type/value	*A* (products)	*A* (herbicides)	*A* (fungicides)	*A* (insecticides)	*B* (products)	*B* (herbicides)	*B* (fungicides)	*B* (insecticides)
10th percentile	0.00	0.00	0.00	0.00	0.61	0.72	0.72	0.65
90th percentile	0.86	0.90	0.57	0.65	1.95	2.06	1.63	1.84
Mean	0.29	0.32	0.25	0.20	1.26	1.32	1.19	1.26

Table 2
Predicted minimum, maximum and mean values for the non-polar descriptors *V*, *L* and *E* for agrochemical products

Descriptor type/value	*V* (all products)	*V* (modern products)	*L* (all products)	*L* (modern products)	*E* (all products)	*E* (modern products)
Minimum	0.25	0.65	1.01	1.08	−0.92	−0.53
Maximum	6.68	6.68	27.12	27.12	4.33	4.33
Mean	2.10	2.29	9.75	10.71	1.52	1.63

screen inputs result may seem surprising, but given the inputs were in fact a random selection from the Syngenta compound collection it is to be expected that they were biased in favour of 'agrochemical-like' properties.

In 2006 Pharma Algorithms developed a new version of ABSOLV, which is available as a module in their ADME Boxes software.[24] Predictions have been made for the *V*, *L*, *E*, *A*, *B* and *S* descriptors for 1350 products listed in the Pesticide Manual,[25] of which 750 compounds were defined as modern agrochemicals. Table 2 shows that the difference between these two sets of products in terms of their maximum, minimum and mean values for the non-polar descriptors *V*, *L* and *E* are quite small. A similar result has been reported using molecular weight as a size parameter, the mean value for all products being 297 and modern products 322.[12]

Figure 1 gives the distributions of the *V*, *L* and *E* values for modern agrochemicals sub-divided into herbicides, fungicides and insecticides. The most significant differences observed are in McGowan volume, reflecting maximum values of 3.59 for fungicides, 4.31 for herbicides and 6.68 for insecticides. However, the three types of agrochemicals have similar mean values of 2.27 ± 0.16, which is, the same as shown in Table 2 for modern products. The trends in McGowan volume for agrochemical types are consistent with those previously reported using molecular weight.[12]

Again, as shown in Table 3 and Figure 2, the combination of minimum, maximum and mean values with distribution profiles are useful in the analysis of the polar descriptors, *A*, *B* and *S* for modern herbicides, fungicides and insecticides. The mean values of 0.21 ± 0.05 for *A* are very low, confirming the reported poor tolerance of agrochemicals for hydrogen bond donors.[11,12] Whilst the majority of these agrochemical products have *A* values of <1, there are exceptions, as shown by the maximum values of 1.7 for herbicides and 3.0 for fungicides.

There are clear differences in the distribution of the *B* descriptor, with insecticides having the widest range and fungicides the

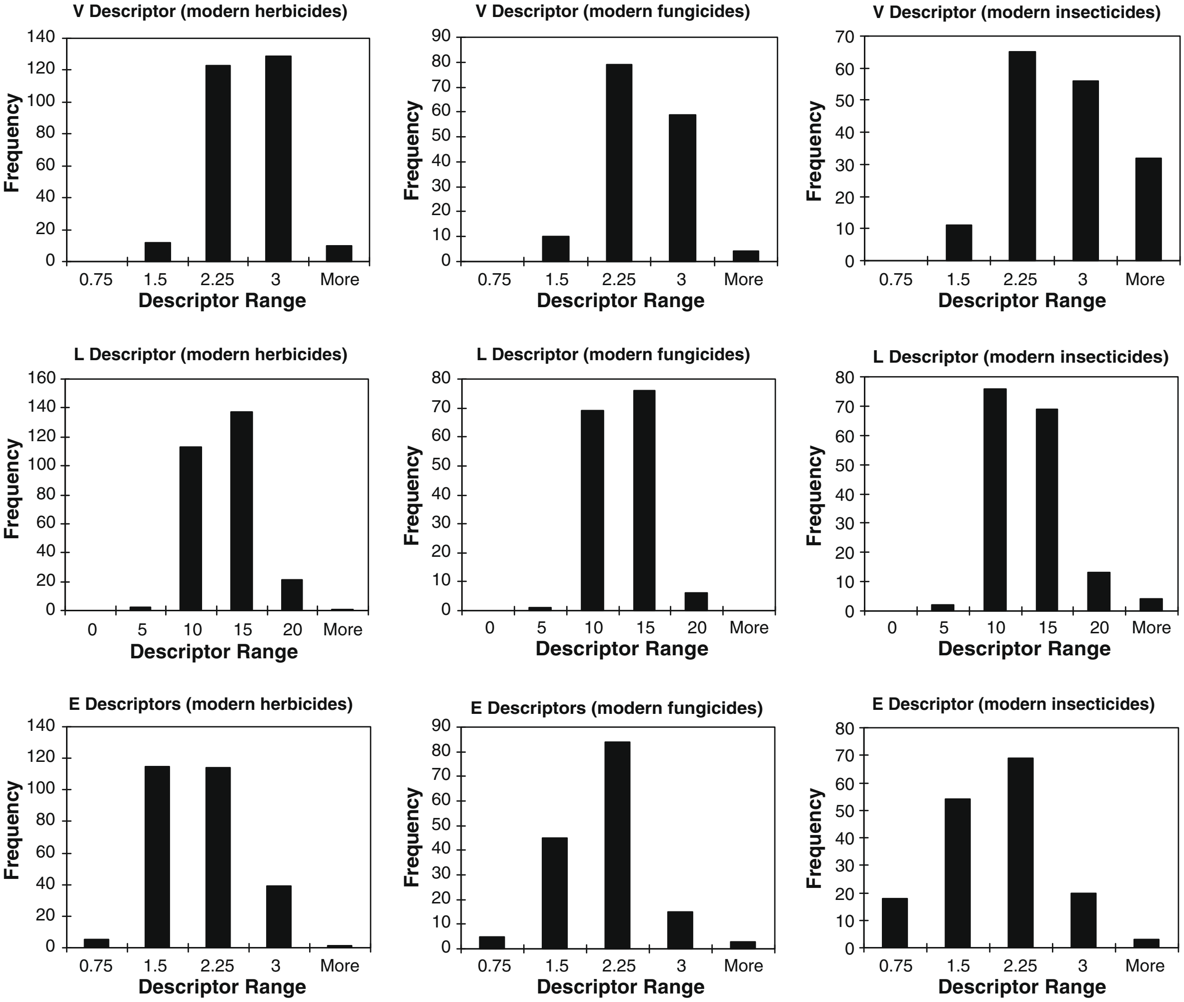

Figure 1. Distribution of the predicted non-polar descriptors *E*, *L* and *V* for modern pesticides.

Table 3

Predicted minimum, maximum and mean values for the polar descriptors *A*, *B* and *S* for modern agrochemical products

Descriptor value	*A* (herb.)	*A* (fung.)	*A* (insect.)	*B* (herb.)	*B* (fung.)	*B* (insect.)	*S* (herb.)	*S* (fung.)	*S* (insect.)
Minimum	0.00	0.00	0.00	0.29	0.37	0.10	−0.01	0.61	0.14
Maximum	1.71	3.02	1.22	3.46	5.28	4.66	4.46	4.59	4.19
Mean	0.24	0.22	0.16	1.44	1.28	1.53	2.13	1.95	1.98

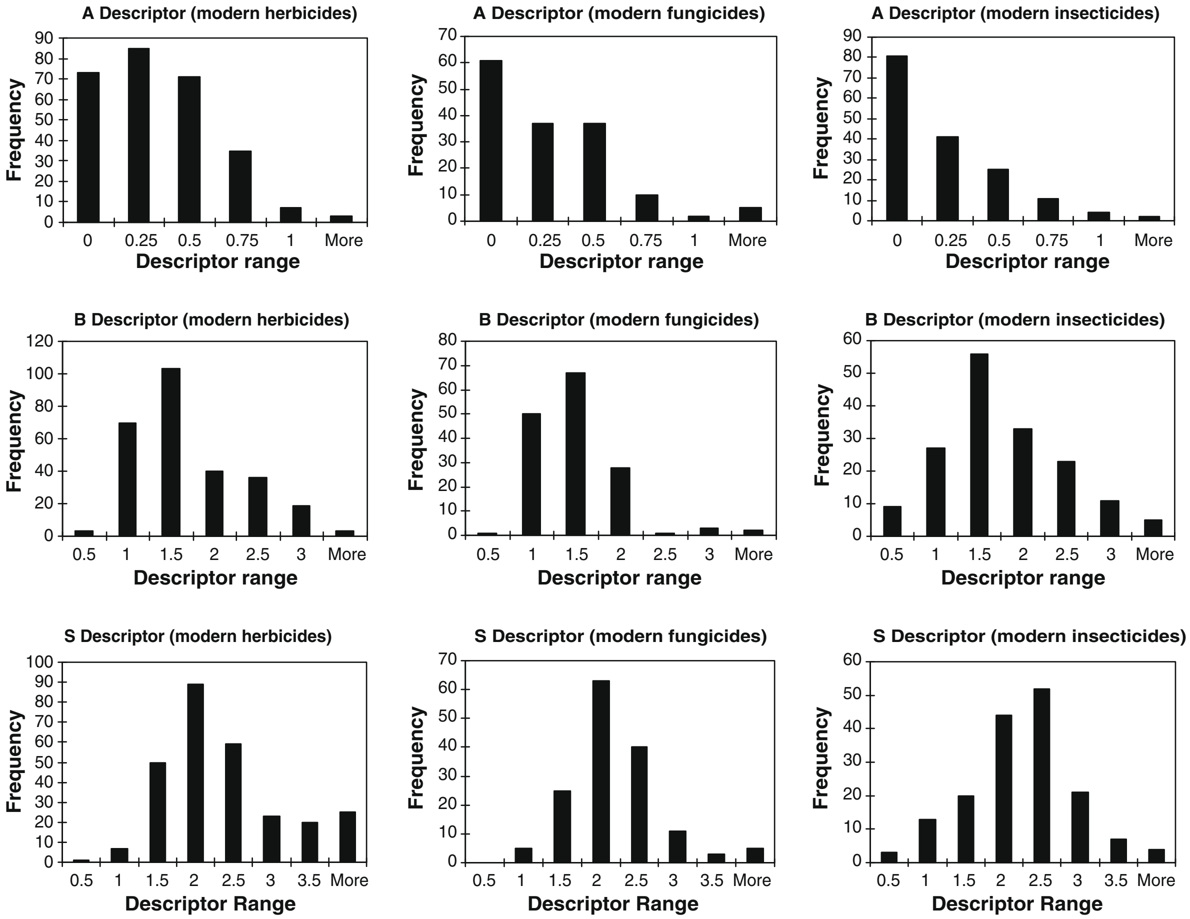

Figure 2. Distribution of the predicted polar descriptors *A*, *B* and *S* for modern pesticides.

lowest. These differences in hydrogen bond basicity are not apparent in the simple statistical data. Overall, little if any distinction can be made between the three types of agrochemicals with respect to the dipole dependent *S* descriptor. The sum of the *A* and *B* descriptors approximates to polar surface area.[19] The *A* + *B* sum of the mean values for herbicides, fungicides and insecticides are very similar with values of 1.6 ± 0.1, consistent with the conclusion of Tice[26] that surface area parameters offered no advantage as potential discriminating factors in agrochemical profiling.

The prediction of Abraham descriptors via ABSOLV provides a simple means of profiling agrochemicals, which can give information consistent with that previously reported via the combination of multiple approaches to the assessment of physical and molecular properties. A simple pragmatic guide to bioavailability of $V \leqslant 3$, $A \leqslant 1$ and $B \leqslant 3$ is suggested for agrochemicals.

As with all structure-based prediction methods, the descriptors generated will be sensitive to the drawing convention used in vendor and corporate databases. Assignment of a physiologically relevant tautomeric form for compounds with ionisable functions is of obvious importance. Methods such as Permute and Leatherface have been described[27] which can be used to assess and reassign structures prior to analysis. In 2003 Leatherface was applied to all structures and used to assign formal charge @ pH 7 in the work of Clarke and Delaney on agrochemical profiles.[12] An indirect evaluation of predicted descriptors was presented by Japertas[28] at an ACS meeting in 2007 using a set of ~600 agrochemical products with measured log*P* octanol values taken from the Pesticide Manual.[25] Descriptors and log*P* octanol values were calculated using the corresponding LFER via the Pharma Algorithms ABSOLV module. A plot of measured vs. predicted log*P* values gave an r^2 of 0.73 and mean absolute error of 0.64, consistent with other log*P* prediction methods applied to this dataset.

2.2. Experimental Abraham descriptors for agrochemicals

Abraham and co-workers have shown that measured organic–water partition coefficients, expressed as log*P*, for octanol, cyclohexane, toluene and chloroform are sufficient for the determination of the *A*, *B* and *S* descriptors for drugs.[29] Enomoto in her Ph.D. studies supervised by Abraham and Clarke, used a similar set of log*P* values, that is, for octanol, hexane, toluene and dichloromethane, to determine descriptors for agrochemicals.[28,30,31] The solvation equations for each of these organic–water partition coefficients are given in Table 4.

Examples of the experimental *A*, *B* and *S* descriptors obtained for representative agrochemicals from best fits between calculated and measured log*P* value using the Excel Solver tool are given in Table 5, which also includes the independently calculated E and *V* descriptor values. Of the 46 sets of descriptor values originally determined by Enomoto, the 28 listed in Table 5 are accessible via the ABSOLV database option in ADME Boxes.[24] Whilst this minimal dataset approach can yield reliable descriptor values, the preferred method is to use all available partition coefficient data, including water–air (logK_w), to give the maximum number of solvation equations of types 1 and 2 for analysis, as recently described for triazine herbicides.[21] Such an analysis is in progress for the complete Enomoto compound set and other agrochemicals. The results of this comprehensive study will be reported by Abraham and Clarke.

At Syngenta, we are evaluating routinely measured properties such as octanol–water log*P*, hexane solubility, water solubility,

Table 4
Solvation equations for organic–water partition coefficients used to determine experimental descriptors

Partition coefficient (log*P*)	Constant	*A*	*B*	*S*	*E*	*V*
Octanol–water	0.088	0.034	−3.460	−1.054	0.562	3.814
Toluene–water	0.143	−3.010	−4.824	−0.720	0.527	4.545
Hexane–water	0.361	−3.599	−4.764	−1.723	0.579	4.344
Cyclohexane–water	0.159	−3.740	−4.929	−1.678	0.784	4.577
Dichloromethane–water	0.314	−3.238	−4.137	−0.022	0.001	4.259
Chloroform–water	0.327	−3.191	−3.437	−1.254	0.573	4.589

Table 5
Experimental descriptors for representative agrochemicals

Name	CAS-RN	*A*	*B*	*S*	*E*	*V*
Atrazine	1912-24-9	0.32	0.96	1.17	1.51	1.62
Simazine	122-34-9	0.33	0.95	1.20	1.55	1.48
Cyanazine	21725-46-2	0.45	0.97	2.24	1.73	1.77
Acetochlor	34256-82-1	0.00	1.36	1.11	1.16	2.14
Propachlor	1918-16-7	0.00	0.98	0.95	1.02	1.66
Flurochloridone	61213-25-0	0.00	0.86	1.41	1.06	1.87
Dimethirimol	5221-53-4	0.00	1.35	0.87	1.04	1.78
Ethirimol	23947-60-6	0.35	1.26	0.83	1.11	1.78
Pyrimethanil	53112-28-0	0.03	0.99	0.92	1.65	1.62
Cyprodinil	121552-61-2	0.07	0.93	0.92	2.06	1.80
Metalaxyl	57837-19-1	0.00	1.60	1.96	0.85	2.23
Furalaxyl	57646-30-7	0.00	1.60	1.47	1.49	2.32
Isoxaben	82558-50-7	0.18	1.40	2.07	1.54	2.60
Diphenamid	957-51-7	0.00	1.25	1.82	1.42	2.00
Flutriafol	76674-21-0	0.15	1.44	1.60	1.76	2.09
Tebuconazole	107534-96-3	0.26	1.42	1.52	1.54	2.41
Paclobutrazol	76738-62-0	0.26	1.41	1.49	1.53	2.27
Carbaryl	63-25-2	0.27	0.73	1.90	1.51	1.54
Pirimicarb	23103-98-2	0.00	1.34	1.34	1.18	1.89
Fenoxycarb	72490-01-8	0.12	0.94	1.95	1.31	2.32
Chlorotoluron	15545-48-9	0.47	0.81	1.50	1.11	1.62
Diuron	330-54-1	0.57	0.70	1.60	1.28	1.60
Fenuron	101-42-8	0.37	0.93	1.67	1.21	1.35
Azoxystrobin	131860-33-8	0.00	2.30	1.90	2.59	2.92
Kresoxim methyl	143390-89-0	0.00	1.45	0.89	1.15	2.42
Prosulfuron	94125-34-5	0.25	1.90	1.48	1.43	2.65
Diflubenzuron	35367-38-5	0.10	1.04	1.14	1.87	1.99
Hexaflumuron	86479-06-3	0.39	1.05	1.02	1.38	2.53

volatility[32] and chromatographic hydrophobicity index (CHI)[33] to set up equations of both types 1 and 2 to determine the *A*, *B* and *S* descriptors. In addition we are exploring the NMR method[34] reported by Abraham and co-workers to enable an independent determination of the *A* descriptor, thus reducing to two the unknowns in the Excel Solver method.

2.3. LFER equations for physical and physiological parameters

A large number of LFER equations exist for a wide variety of physical, physiological and toxicological end points in which rate of diffusion (log*k*), equilibrium partitioning (log*P*, log*K*) or non-specific activity and interactions (−log toxicity, −log MIC, etc) play a significant role. The examples in Table 6 are selected to illustrate the scope of LFER's, in terms of the diversity of application and descriptor coefficients found.[14,19,35–40] Judgement as to the relevance and potential interpretation of a given LFER is not simply based on the statistical quality of the equations, but also knowledge and experience of the end-points. Consequently, the original papers cited frequently have cross-discipline authors, reflecting the breadth of thought required to produce credible analysis. It is advisable to refer back to such papers in order to understand fully the rationale and significance of any LFER, particularly those with multifaceted and/or species dependent end-points such as brain uptake[14] and non-specific toxicology.[38,39]

2.4. LFER equations for agrochemicals

2.4.1. Soil sorption

An LFER equation for the prediction of soil organic carbon sorption coefficient (K_{oc}) based on the dataset of Tao et al.[41] was presented in a talk by Clarke at QSAR 2004.[42] The dataset contained measured K_{oc} values for 592 compounds covering 17 chemical classes. Experimental descriptors were available for 209 of these compounds, which were representative of the complete dataset in terms of range of K_{oc} values, log*P* octanol values and structural diversity.

Multiple linear regression analysis gave the equation for $\log K_{oc}$ in Table 7, which had an r^2 of 0.92 and standard error of 0.38. The relative contributions of the *a*, *b*, *s* and *e* coefficients to $\log K_{oc}$ compared to log*P* octanol were assessed by scaling each equation to their v coefficient as shown in Table 7. There is a marked similarity

Table 6
Examples of LFER's for physical, physiological and toxicological processes

	Constant	*a*	*b*	*s*	*e*	*v*
Octanol–water (log*P*)[a]	0.088	0.034	−3.460	−1.054	0.562	3.814
Octanol–alkane (Δlog*P*)[a]	−0.072	3.655	1.396	0.528	−0.093	−0.521
Intestinal adsorption (log*k*)[b]	0.544	−0.409	−0.514	0.141	−0.025	0.204
Water–plant cuticle (log*K*)[c]	−0.415	−0.508	−4.096	−0.413	0.596	3.908
Skin permeation (log*K*)[d]	−5.426	−0.473	−3.000	−0.473	−0.106	2.296
Protozoa (−log IGC)[e]	2.706	0.938	−1.025	0.433	1.426	2.599
Tadpole (−log Tox)[f]	0.716	0.127	−2.914	−0.517	0.703	3.308
Fish (−log Tox)[g]	0.996	0.417	−3.574	−0.182	0.418	3.377
Oral bacteria (−log MIC)[h]	−3.320	0.727	−1.904	−0.605	1.111	2.423

[a] Ref. 14.
[b] Ref. 19.
[c] Ref. 35.
[d] Ref. 36.
[e] Ref. 37.
[f] Ref. 38.
[g] Ref. 39.
[h] Ref. 40.

Table 7
Comparison of LFER's for $\log K_{oc}$ and $\log P$ octanol

	Constant	*a*	*b*	*s*	*e*	*v*
Log K_{oc}	0.39	−0.36 −0.17 (a/v)	−1.98 −0.99 (b/v)	−0.34 −0.17 (s/v)	0.88 0.44 (e/v)	2.01 —
Log *P* octanol	0.09	0.03 0.01 (a/v)	−3.46 −0.91 (b/v)	−1.05 −0.28 (s/v)	0.56 0.15 (e/v)	3.81 —

in the scaled coefficients for $\log K_{oc}$ and $\log P$ octanol, the most significant difference being in the e coefficient.

The dataset of 209 compounds yields a $\log P$ dependent Eq. 3 similar to that reported by Briggs[43] 4. Interestingly, inclusion of the calculated *E* descriptor gave an improved Eq. 5 compared to $\log P$ octanol alone, indicating a contribution from dispersion force interactions in soil sorption:

$$\log K_{oc} = 0.5 \log P + 0.9 \quad r^2 = 0.73 \tag{3}$$

$$\log K_{oc} = 0.52 \log P + 0.62 \quad r^2 = 0.84 \tag{4}$$

$$\log K_{oc} = 0.52 \log P + 0.74E + 0.29 \quad r^2 = 0.91 \tag{5}$$

Poole and Poole[44] have published a LFER for $\log K_{oc}$ which is very similar to that obtained in this work, though it uses the alternative *Bo* descriptor[17], which is preferred to *B* for some chemical types in solvents which contain significant amounts of water. Analysis of the Tao dataset will be revisited when the number compounds with experimental descriptors in the Abraham database exceeds 300.

2.4.2. Root uptake

Polder et al.[45] have compiled measurements of root concentration factor (RCF), as defined in Eq. 6.

$$\text{RCF} = \text{conc. in roots (mg/kg fresh wt.)/conc. in external solution (mg/L)} \tag{6}$$

Experimental descriptors were available from, or kindly determined by Professor Abraham using appropriate measured physical property data, for 22 of the 26 compounds compiled. Multiple linear regression analysis gave an equation for log RCF with an r^2 of 0.85 and standard error of 0.44, which is compared with the equation for $\log P$ octanol in Table 8. In this case it is the *A* descriptor which differentiates log RCF from $\log P$ octanol.

Polder et al. also compiled data for transpiration stream concentration factor (TSCF) and stem concentration factor (SCF) but unfortunately less than 10 compounds had suitable datasets and experimental descriptors, ruling out credible LFER analysis.

3. Conclusions

The Pharma Algorithms ADME Boxes ABSOLV module has proved to be a useful tool for profiling agrochemicals with respect to Abraham descriptors. A simple guide to bioavailability has been defined in terms of McGowan volume *V* ($\leqslant 3$), hydrogen bond acidity *A* ($\leqslant 1$), and hydrogen bond basicity *B* ($\leqslant 3$).

Table 8
Comparison of LFER's for log RCF and $\log P$ octanol

	Constant	*a*	*b*	*s*	*e*	*v*
Log RCF	−0.76	1.01 0.41 (a/v)	−2.52 −1.03 (b/v)	−0.38 −0.16 (s/v)	0.06 0.02 (e/v)	2.44 —
Log *P* octanol	0.09	0.03 0.01 (a/v)	−3.46 −0.91 (b/v)	−1.05 −0.28 (s/v)	0.56 0.15 (e/v)	3.81 —

Experimental descriptors can be obtained via spreadsheet tools, such as Solver in Excel, using a wide range of measured physical properties with established LFER equations, the more the better. However, as few as four properties can be used, provided there is variation in one or more of their equation coefficients. Considerable care and experience is needed to ensure that the properties used are reliable, unfortunately there are significant errors in many databases, as illustrated in a recent study on triazine herbicides.[21]

The breadth of processes for which credible LFER's have been reported is exceptional, but this is not the case for end-points of relevance to agrochemical research. This situation reflects the paucity of suitable data in the literature. Apart from soil sorption related data, it is rare to find datasets for even 20–30 compounds, the minimum required, provided there is a reasonable distribution of descriptors. Whilst multiple linear regression analysis of end-point and descriptor values is easily done within Excel, the temptation should be resisted with small datasets to avoid over-fitting and misinterpretation of the significance of descriptors to a particular end-point. In this work an acceptable LFER for root concentration factor was obtained using 22 compounds but no attempt was made to generate LFER's for parameters such as transpiration stream concentration factor with datasets of less than 10 compounds.

The LFER equations for soil sorption and root concentration factor show marked similarities with the equation for $\log P$ octanol, as expected from the well established work of Briggs and co-workers.[43,46] It is however intriguing to see indications of the potential importance of dispersion force interactions in soil sorption and hydrogen bond acidity in root concentration factor.

It is in our interests in agrochemical research to make high quality physical property data available for use in the determination of experimental descriptor values. Not only could this expand the Abraham database and contribute to further improvement in prediction software such as ABSOLV, but also ensure the inclusion of more agrochemicals in the development of LFER's for physical, physiological and non-specific biological end-points.

Acknowledgements

I thank Professor Michael Abraham for his continued advice and support, Pranas Japertas and Pharma Algorithms for making pre-release versions of ABSOLV available to me for evaluation, and Lucy Entwistle for help with data analysis.

References and notes

1. Donovan, S.F. In *Synthesis and Chemistry of Agrochemicals VII*; Lyga, J. W., Theodoridis, G., Eds., ACS Symposium Series, 2007, 948, pp 7–22.
2. Delaney, J.; Clarke, E.; Hughes, D.; Rice, M. *Drug Discovery Today* **2006**, *11*, 839.
3. Lipinski, C. A.; Lombardo, F.; Dominy, P. J.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3.
4. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Comb. Chem.* **1999**, *1*, 55.
5. Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. I. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743.
6. Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856.
7. Oprea, T. I. *J. Comp. Aided Mol. Des.* **2002**, *16*, 325.
8. Wenlock, M. C.; Austin, R. P.; Barton, P.; Davies, A. M.; Leeson, P. D. *J. Med. Chem.* **2003**, *46*, 1250.
9. Muresan, S.; Sadowski, J. In *Molecular Drug Properties. Measurement and Prediction*; Mannhold, R., Ed.; Wiley-VCH: Weinheim, 2008; pp 441–459.
10. Briggs, G. G. SCI Meeting on Uptake of Agrochemicals and Pharmaceuticals, London, December, 1997.
11. Tice, C. M. *Pest. Manag. Sci.* **2001**, *57*, 3.
12. Clarke, E. D.; Delaney, J. S. *Chimia* **2003**, *57*, 731.
13. Lindell, S. D.; Scherkenbeck, J. *Combinatorial Chemistry and High Throughput Screening* **2005**, *8*, 555.
14. Abraham, M. H.; Chadha, H. S.; Martins, F.; Mitchell, R. C.; Bradbury, M. W.; Gratton, J. A. *Pestic. Sci.* **1999**, *55*, 78.
15. Klamt, A.; Eckert, F.; Hornig, M. *J. Comp. Aided Mol. Des.* **2001**, *15*, 355.
16. Zissimos, A. M.; Abraham, M. H.; Klamt, A.; Eckert, F.; Wood, J. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1320.
17. Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. *J. Chromatogr., A* **2004**, *1037*, 29.

18. Virha, M.; Carr, P. W. *J. Chromatogr., A* **2006**, *1126*, 143.
19. Abraham, M. H.; Ibrahim, A.; Zissimos, A. M.; Zhao, Y. H.; Comer, J.; Reynolds, D. P. *Drug Discovery Today* **2002**, *7*, 1056.
20. Green, C. E.; Abraham, M. H.; Acree, W. E., Jr.; De Fina, K. M.; Sharp, T. L. *Pest. Manag. Sci.* **2000**, *56*, 1043.
21. Abraham, M. H.; Enomoto, K.; Clarke, E. D.; Roses, M.; Rafols, C.; Fuguet, E. *J. Environ. Monit.* **2007**, *9*, 234.
22. Platts, J. A.; Butina, D.; Abraham, M. H.; Hersey, A. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 835.
23. Clarke, E. D.; Delaney, J. S. In *Designing Drugs and Crop Protectants: Processes, Problems Solutions*; Ford, M., Livingstone, D., Dearden, J., Van de Waterbeemd, H., Eds.; Blackwell: Oxford, 2003; pp 51–54.
24. ADME Boxes v3.5. Pharma Algorithms Inc., www.ap-algorithms.com.
25. *The Pesticide Manual*; Tomlin, C. D. S., Ed., 12th ed.; Brit. Crop Protection Council: Farnham, UK, 2000.
26. Tice, C. M. *Pest. Manag. Sci.* **2002**, *58*, 219.
27. Kenny, P. W.; Sadowski, J. In *Chemoinformatics in Drug Discovery*; Oprea, T. I., Ed.; Wiley-VCH: Weinheim, 2004; pp 271–285.
28. Japertas, P.; Sazonovas, A.; Clarke, E. D.; Delaney, J. S. *Abstracts of Papers*, 234th National Meeting of the American Chemical Society, Boston, MA, August 19–23, 2007. (poster pdf available from E. D.Clarke).
29. Zissimos, A. M.; Abraham, M. H.; Barker, M. C.; Box, K. J.; Tam, K. Y. *J. Chem. Soc., Perkin Trans. 2* **2002**, 470.
30. Enomoto, K. PhD thesis, University College London, 2003.
31. Enomoto, K.; Clarke, E. D.; Abraham, M. H. In *Designing Drugs and Crop Protectants: processes, problems and solutions*; Ford, M., Livingstone, D., Dearden, J., Van de Waterbeemd, H., Eds.; Blackwell: Oxford, 2003; pp 20–22.
32. Smith, S. C.; Clarke, E. D.; Ridley, S. M.; Bartlett, D.; Greenhow, D. T.; Glithro, H.; Klong, A. Y.; Mitchell, G.; Mullier, G. W. *Pest. Manag. Sci.* **2005**, *61*, 16.
33. Valko, K.; My Du, C.; Bevan, C.; Reynolds, D. P.; Abraham, M. H. *Curr. Med. Chem.* **2001**, *8*, 1137.
34. Abraham, M. H.; Abraham, R. J.; Byrne, J.; Griffiths, L. *J. Org. Chem.* **2006**, *71*, 3389.
35. Platts, J. A.; Abraham, M. H. *Environ. Sci. Technol.* **2000**, *34*, 318.
36. Abraham, M. H.; Martins, F. *J. Pharm. Sci.* **2004**, *93*, 1508.
37. Bowen, K. R.; Flanagan, K. B.; Acree, W. E., Jr.; Abraham, M. H. *Sci. Total Environ.* **2006**, *369*, 109.
38. Bowen, K. R.; Flanagan, K. B.; Acree, W. E., Jr.; Abraham, M. H.; Rafols, C. *Sci. Total Environ.* **2006**, *371*, 99.
39. Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. *Chem. Res. Toxicol.* **2005**, *18*, 1497.
40. Mintz, C.; Acree, W. E., Jr.; Abraham, M. H. *QSAR Combinatorial Sci.* **2006**, *25*, 912.
41. Tao, S.; Piao, H.; Dawson, R.; Lu, X.; Hu, H. *Environ. Sci. Technol.* **1999**, *33*, 2719.
42. Abraham, M. H.; Clarke, E. D.; Enomoto, K. 11th Int. Workshop on QSAR, Liverpool, UK, 9–13 May, 2004 (talk pdf available from E. D. Clarke).
43. Brigg, G. G. *Philos. Trans. R. Soc. London. Ser. B* **1990**, 67.
44. Poole, S. K.; Poole, C. F. *J. Chromatogr., A* **1999**, *845*, 381.
45. Polder, M. D.; Hulzebos, E. M.; Jager, D. T. *Environ. Toxicol. Chem.* **1995**, *14*, 1615.
46. Briggs, G. G.; Bromilow, R. H.; Evans, A. A. *Pestic. Sci.* **1982**, *13*, 495.