Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supplementary Information

An interactive database to explore herbicide physicochemical properties

Michael N. Gandy,[a] Maxime G. Corral,[a,b] Joshua S. Mylne[a,b*] and Keith A. Stubbs[a*]

[a]*School of Chemistry and Biochemistry, The University of Western Australia, Crawley, WA, 6009, Australia.*

E-mail: keith.stubbs@uwa.edu.au, joshua.mylne@uwa.edu.au

[b]*ARC Centre of Excellence in Plant Energy Biology, 35 Stirling Highway, Crawley, Perth 6009, Australia*

Materials and Methods

Compound selection

To gather a comprehensive list of herbicides the literature was initially surveyed and all compounds listed in previous reviews[1] were incorporated. Compounds were also sourced from the World of Herbicides provided by Herbicide Resistance Action Committee and produced by Syngenta as well as from the EU pesticide database, Department of Horticulture database (University of Kentucky), Urban Integrated Pest Management database (University of Arizona) and Department of Agriculture, Forestry & Fisheries (Republic of South Africa) and the Pesticide Manual [2].

A textual list of the 334 compound names follows: 2,4,5-T; 2,4-D; 2,4-DB; acetochlor; acifluorfen; aclonifen; acrolein; alachlor; allidochlor; alloxydim; ametryne; amicarbazone; amidosulfuron; aminocyclopyrachlor; aminopyralid; amiprophos-methyl; amitrole; anilofos; asulam; atrazine; azafenidin; azimsulfuron; beflubutamid; benazolin; benazolin-ethyl; benfluralin; benfuresate; bensulfuron-methyl; bensulide; bentazon; benthiocarb; benzfendizone; benzobicyclon; benzofenap; bicyclopyrone; bifenox; bilanaphos; bispyribac; bromacil; bromobutide; bromofenoxim; bromoxynil; butachlor; butafenacil; butamifos; butenachlor ; butralin; butroxydim; butylate; cafenstrole; carbetamide; carfentrazone-ethyl; chlomethoxyfen; chloramben; chlorbromuron; chlorflurenol; chlorimuron-ethyl; chlorotoluron; chloroxuron; chlorphthalim; chlorpropham; chlorsulfuron; chlorthal-dimethyl; chlorthiamid; cinidon-ethyl; cinmethylin; cinosulfuron; clethodim; clodinafop; clodinafop-propargyl; clomazone; clomeprop; clopyralid; cloransulam-methyl; cumyluron; cyanazine; cycloate; cyclosulfamuron; cycloxydim; cyhalofop-butyl; dalapon; dazomet; desmedipham; desmetryne; diallate; dicamba; dichlobenil;

dichlorprop; diclofop-methyl; diclosulam; diethatyl-ethyl; difenzoquat; diflufenican; diflufenzopyr; dimefuron; dimepiperate; dimethachlor; dimethametryn; dimethenamid; dimethylarsinic acid; dinitramine; dinoseb; dinoterb; diphenamid; diquat; dithiopyr; diuron; DNOC; DSMA; dymron; endothall; EPTC; esprocarb; ethalfluralin; ethametsulfuron-methyl; ethidimuron; ethiolate; ethofumesate; ethoxyfen-ethyl; ethoxysulfuron; etobenzanid; fenoxaprop; fenoxaprop P-ethyl; fenoxasulfone; fentrazamide; fenuron; flamprop M-isopropyl; flamprop-methyl; flazasulfuron; florasulam; fluazifop; fluazifop-butyl; fluazolate; flucarbazone-sodium; flucetosulfuron; fluchloralin ; flufenacet; flufenpyr-ethyl; flumetsulam; flumiclorac-pentyl; flumioxazin; fluometuron; fluoroglycofen-ethyl; flupoxam; flupropacil; flupropanate; flupyrsulfuron-methyl; fluridone; flurochloridone; fluroxypyr; flurtamone; fluthiacet-methyl; fomesafen; foramsulfuron; fosamine; glufosinate; glyphosate; halosafen; halosulfuron-methyl; haloxyfop-methyl; hexazinone; imazamethabenz-methyl; imazamox; imazapic; imazapyr; imazaquin; imazethapyr; imazosulfuron; indanofan; indaziflam; iodosulfuron; iofensulfuron; ioxynil; ipfencarbazone; isopropalin; isoproturon; isouron; isoxaben; isoxachlortole; isoxaflutole; isoxapyrifop; karbutilate; lactofen; lenacil; linuron; MCPA; MCPA-thioethyl; MCPB; mecoprop; mefenacet; mefluidide; mesosulfuron; mesotrione; metam; metamitron; metazachlor; metazosulfuron; methabenzthiazuron; methazole; methiozolin; methoprotryne; methoxyphenone; methyldymron; metobenzuron; metobromuron; metolachlor; metosulam; metoxuron; metribuzin; metsulfuron-methyl; molinate; monalide; monolinuron; monuron; MSMA; naproanilide; napropamide; naptalam; NC-330; neburon; nicosulfuron; nitrofen; norflurazon; OK-8910; oleic acid; orbencarb; orthosulfamuron; oryzalin; oxadiargyl; oxadiazon; oxasulfuron; oxaziclomefone; oxyfluorfen; paraquat; pebulate; pelargonic acid; pendimethalin;

penoxsulam; pentachlorophenol; pentanochlor; pentoxazone; pethoxamid; phenmedipham; picloram; picolinafen; pinoxaden; piperophos; pretilachlor; primisulfuron-methyl; prodiamine; profluazol; profluralin; profoxydim; prometon; prometryne; propachlor; propanil; propaquizafop; propazine; propham; propisochlor; propoxycarbazone-sodium; propyrisulfuron; propyzamide; prosulfocarb; prosulfuron; pyraclonil; pyraflufen-ethyl; pyrasulfotole; pyrazolynate; pyrazon; pyrazosulfuron-ethyl; pyrazoxyfen; pyribenzoxim; pyributicarb; pyridafol; pyridate; pyriftalid; pyriminobac-methyl; pyrimisulfan; pyrithiobac; pyroxasulfone; pyroxsulam; quinclorac; quinmerac; quinoclamine; quizalofop; quizalofop-P-ethyl; quizalofop-P-tefuryl; rimsulfuron; saflufenacil; sethoxydim; siduron; simazine; simetryne; S-metolachlor; sulcotrione; sulfentrazone; sulfometuron-methyl; sulfosate; sulfosulfuron; TCA; TCBA; tebutam; tebuthiuron; tembotrione; tepraloxydim; terbacil; terbucarb; terbumeton; terbuthylazine; terbutryne; thenylchlor; thiazafluron; thiazopyr; thidiazimin; thiencarbazone; thifensulfuron-methyl; tiocarbazil; topramezone; tralkoxydim; triafamone; triallate; triasulfuron; triaziflam; triazofenamide; tribenuron-methyl; triclopyr; trietazine; trifloxysulfuron; trifluralin; triflusulfuron-methyl; tritosulfuron; vernolate.

Conversion of compounds to SMILES format

SMILES structures were generated from the corresponding CAS registry numbers for each herbicide using the molconvert utility provided by Marvin Beans software (ver 6.1.0, ChemAxon Ltd. and available from www.chemaxon.com), which was installed locally and used in batch processing mode. In rare cases where this tool failed, SMILES structures were instead generated individually by entering the skeletal formula into the web-based converter (available from Marvin, Calculator Plugin and Chemical Terms Demo (Chemaxon (http://www.chemaxon.com/marvin/sketch)). For

compounds composed of multiple fragments (e.g. compounds composed of cation(s) and anion(s)), only the active ingredient was considered when calculating physicochemical properties.

Physicochemical property computation tools

Various physicochemical properties were calculated for each compound using a combination of two tools. The first is a web-based applet named Virtual Computational Chemistry Laboratory (VCCLAB (www.vcclab.org))[3], which is capable of returning both experimental and calculated Log *P* and Log *S* values derived from a variety of sources. All other properties were calculated from SMILES strings using the cxcalc function provided by Marvin Beans software operated in batch processing mode. For properties where pH is specified, cxcalc internally estimates pK_a and pK_b values in order to determine the major protonation state. The specific parameters used for each command are listed in the supporting information.

Molar mass

The molar mass of a particular compound was determined from its corresponding SMILES string using cxcalc.

Proportion of aromatic atoms

The number of non-hydrogen atoms within aromatic systems as a ratio ranging from 0 to 1 of the total number of non-hydrogen atoms was calculated using cxcalc.

Lipophilicity and aqueous solubility

Experimentally determined Log *P* and Log *S* values were obtained by entering individual SMILES stings into VCCLAB, which returned values sourced from the Physical Properties Database (PHYSPROP, Syracuse Research Corp). Where experimental values were not available, certain estimated values returned by

VCCLAB were used. In our case, the partition coefficient was calculated as the consensus value (arithmetic mean) of Log *P* estimates generated using two methods, namely XLOGP3 (ref. [4]) and ALogPs 2.1 (ref. [5]), whereas Log *S* was calculated using AlogpS 2.1 (ref. [6]). Distribution coefficient (Log *D*) values were estimated for the major protonation state of the major tautomer at pH 7.4 using cxcalc, which returned Log *D* as the arithmetic mean of three estimations, made using the ALOGP (ref. [7]) method trained with one of three different training sets (Chemaxon (www.chemaxon.com/marvin-archive/6.1.0/marvin/help/calculations/partitioning.html)).

Formal charge

Formal charge, defined as the sum of charges carried by ionisable atoms for the major protonation state at pH 7.4, was estimated using cxcalc.

Hydrogen bonding

Counts of hydrogen bond acceptor atoms, defined as the total number of oxygen and nitrogen atoms able to accept hydrogen bond(s) for the major protonation state at pH 7.4, were made using cxcalc. Also, counts of hydrogen bond donors, defined as the total number of N–H and O–H bonds present in the major protonation state at pH 7.4, were made using cxcalc.

Polar surface area

Polar surface area (PSA) refers to the total exposed (non-bonding) surface area of all polar groups i.e. nitrogen and oxygen, and any hydrogen atoms directly attached to these. The method used here (via cxcalc), known as topological polar surface area (TPSA), was used to calculate PSA for the major protonation form at pH 7.4. TPSA calculates PSA using only two-dimensional descriptors (SMILES strings were used

here), yet provides a good approximation to PSA values obtained from other methods that take the three-dimensional structure of the molecule in account[8].

Number of rotatable bonds

Counts of rotatable bonds were made using cxcalc. Non-rotatable bonds, as defined by cxcalc, are bonds that were unsaturated, single bonds connecting hydrogens or terminal atoms, single bonds of amides and sulphonamides, or single bonds connecting two hindered aromatic rings (having at least three ortho substituents). (Chemaxon (http://www.chemaxon.com/marvin-archive/6.1.0/marvin/help/calculations/geometrygroup.html))

Mode of action and application stage

A recent analysis of pesticides by Klein and Lindell[9] observed clustering of physicochemical properties for specific modes of action. To include and extend this type of analysis the mode of action and recommended application stage (pre/post emergence) for the 334 herbicides were sought in the most recent edition of The Pesticide Manual[2]. If the mode of action was unknown, vague or ambiguous the term "unknown mode of action" was entered into the database.

References

1 (a) E. Clarke and J. Delaney, *Chimia*, 2003, **57**, 731; (b) C. M. Tice, *Pest Manag. Sci.*, 2001, **57**, 3; (c) *Pest Manag. Sci.*, 2002, **58**, 219.

2 C. MacBean, *The Pesticide Manual: A World Compendium*, British Crop Production Council, Hampshire, 2013.

3 I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk and V. V. Prokopenko, *J Comput Aided Mol Des*, 2005, **19**, 453.

4 T. Cheng, Y. Zhao, X. Li, F. Lin, Y. Xu, X. Zhang, Y. Li, R. Wang and L. Lai, *Journal of Chemical Information and Modeling*, 2007, **47**, 2140.

5 I. V. Tetko, V. Y. Tanchuk and A. E. Villa, *J Chem Inf Comput Sci*, 2001, **41**, 1407.

6 I. V. Tetko, V. Y. Tanchuk, T. N. Kasheva and A. E. Villa, *J Chem Inf Comput Sci*, 2001, **41**, 1488.

7 V. N. Viswanadhan, A. K. Ghose, G. R. Revankar and R. K. Robins, *Journal of Chemical Information and Computer Sciences*, 1989, **29**, 163.

8 P. Ertl, B. Rohde and P. Selzer, *J. Med. Chem.*, 2000, **43**, 3714.
9 R. Klein and S. D. Lindell, in *Plant Chemical Biology*, eds. D. Audenaert and P. Overvoorde, John Wiley & Sons, Inc., New Jersey, 2014, pp. 40-63.