MAPPING PHYSICO-CHEMICAL PROPERTY SPACE FOR COMMERCIAL PESTICIDES AND DRUGS

Simona Funar-Timofei,[1] Ana Borota,[1] Sorin Avram,[1] Daniela Ionescu[2]

[1]Institute of Chemistry Timisoara of the Romanian Academy, Bul. Mihai Viteazul 24, 300223 Timisoara, Romania

[2]Department of Toxicology, Faculty of Pharmacy, University of Medicine and Pharmacy "Victor Babes" Timisoara, 2 Murgu E, 300041, Romania

Introduction

The concept of pesticide-likeness has been established in previous studies in terms of molecular properties ranges including molecular weight, calculated logP, the number of H-bond donors and acceptors, the number of aromatic bonds (considered to be correlated with pesticide photostability) and the number of rotatable bonds (related to bioavailability) [1]. Calculated molecular properties were used to predict oral drug-likeness in the field of drug discovery. So far the concept of pesticide-likeness did not yield same interest in comparison to the drug-likeness. Compared to drugs, pesticides are expected to have different absorption, distribution, metabolism, excretion and toxicity properties.

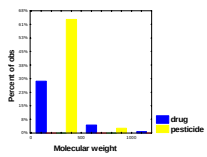
In this study drug-likeness and pesticide-likeness concepts were compared and discussed for a set of 1741 commercial pesticides and a series of previously published 902 marketed drugs [2]. Their physicochemical property distributions were comparatively analyzed and the chemical space covered by drugs and pesticides was described using principal component analysis. Trends in constitutive properties of pesticides are discussed.

Methods

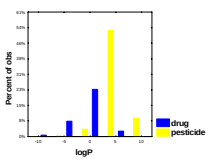
A set of 1741 pesticides were manually extracted from two sources [3, 4] and compared with a set of published drugs [2]. Molecular properties were generated using the InstantJChem software from Chemaxon (www.chemaxon.com) and solubility was calculated with the AlogPS 2.1 software (http://www.vcclab.org/lab/alogps/). Several constitutive properties used in drug-likeness criteria (Table 2), as provided by the InstantJChem program were analyzed with STATISTICA 7.1 (Statsoft, Inc., USA). Principal component analysis was perfomed with SIMCA P+ 12.0 (Umetrics-AB, Sweden).

Results and Discussion

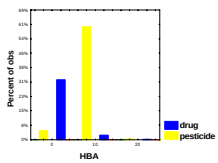
We analyzed twelve constitutive properties (MW, logP, TPSA, HBA, HBD, ABC, ARC, atom number, refractivity, logD, RC, logS) to determine which was the principal factor in differentiating drugs from pesticide. Few structural parameters presented normal distribution plots.



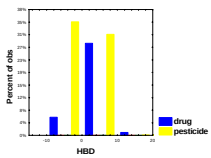
drug, Molecular weight: N = 898, Median = 339.3950, Max = 1620.67, Min = 46.07
Pesticide, Molecular weight: N = 1741, Median = 293.4100, Max = 1052.68, Min = 26.02



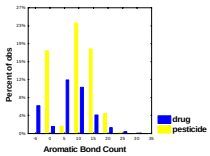
drug, logP: N = 898, Median = 1.9550, Max = 10.57, Min = -14.96
pesticide, logP: N = 1741, Median = 3.2100, Max = 10.33, Min = -8.99



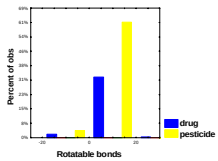
drug, HBA: N = 898, Median = 4.0, Max = 31, Min = 0
pesticide, HBA: N = 1741, Median = 2.0, Max = 22, Min = 0



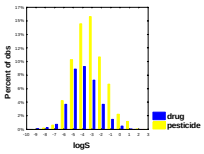
drug, HBD: N = 898, Median = 2, Max = 22, Min = 0
pesticide, HBD: N = 1741, Median = 0, Max = 20, Min = 0



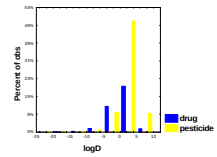
drug, Aromatic Bond Count: N = 898, Median = 10, Max = 32, Min = 0
pesticide, Aromatic Bond Count: N = 1741, Median = 6, Max = 30, Min = 0



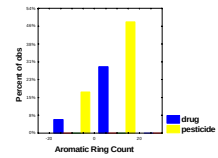
drug, Rotatable bonds: N = 898, Median = 5, Max = 38, Min = 0
pesticide, Rotatable bonds: N = 1709, Median = 5, Max = 21, Min = 0



drug, logS: N = 898, Median = -3.55, Max = 1.1, Min = -8.32
pesticide, logS: N = 1740, Median = -3.83, Max = 1.01, Min = -8.94



drug, logD: N = 898, Median = 0.89, Max = 10.57, Min = -24.25
pesticide, logD: N = 1741, Median = 2.98, Max = 10.33, Min = -16.31



drug, Aromatic Ring Count: N = 898, Median = 2, Max = 6, Min = 0
pesticide, Aromatic Ring Count: N = 1741, Median = 1, Max = 141, Min = 0

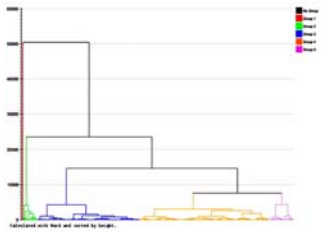
Fig. 1. Distributions of calculated molecular properties for pesticides

Similar distributions of molecular descriptors are seen for pesticides and drugs. Generally, pesticide have lower limits compared with drugs.

Table 1. Drug-likeness criteria estimated by the InstantJChem software

Compound	#molecules	Bioavailability (%)	Ghose filter (%)	Lipinski rule of 5 (4 of 4) (%)	Muegge filter (%)	Rule of three (%)	Veber filter (%)
Drug	902	87	61	77	63	92	80
Pesticide	1741	93	75	83	67	84	91

The low percentages of known drugs that passed the bioavailability filters are not surprising since we have, also, considered molecules that are administrated through other ways than oral. Surprisingly, a high number of pesticides were also predicted as orally absorbed. The chemical space described by the 64 descriptors from InstantJChem was reduced to 8 principal components by PCA analysis selected by 7 cross-validation groups. The principal components served as input variables in cluster analysis. We employed Ward's hierarchical agglomerative algorithm based on Euclidean distances. The resulting dendrogram indicates a good separation of five groups (Fig. 2). The first cluster comprises only a small number of drugs and was excluded in Fig. 3. Principal component analysis performed without these compounds gave following results: N=2639, 8PCs, R2X(cum)=0.909, Q2(cum)=0.829.



clusters	drug (%)	pesticide (%)
cl1	0.44	0.00
cl2	11.48	1.84
cl3	18.87	47.73
cl4	52.21	46.07
cl5	17.00	4.37

Fig 2. Dendogram calculated for the entire set of drugs and pesticides

Cl3 and Cl4 include more than 70% of drugs and 90% of pesticides. Compounds found herein occupy most of the common chemical space of both drugs and pesticides. Cl2 and Cl5 contain three, respectively two times more drugs than pesticides. Given the data used in this study, according to Fig. 3 and the corresponding mean descriptor values (Table 3), one can define the compounds comprised in Cl2 and Cl5 as belonging to low density chemical areas of pesticides and include marginal values of most of their molecular properties. Therefore, we focused on Cl3 and Cl4 which define the core space of pesticides. Compounds included in Cl3 and Cl4 might exhibit both pesticide and drug properties. In order to establish upper limits for pesticide-likeness, we computed the 95% confidence limits of several essential molecular properties calculated by Instant JChem software: MW ≤ 296, logP ≤ 3.23, TPSA ≤ 55, HBA (H bond acceptor) ≤ 3, HBD (H bond donor) ≤ 1, RB (rotatable bonds) ≤ 5, ABC (aromatic bond count) ≤ 7, ARC (aromatic ring count) ≤ 1, Atom count ≤ 34, Refractivity ≤ 74.8, logD ≤ 2.8, ring count ≤ 2, logS ≤ -3.67. We compared the pesticide-likeness criteria obtained by us with the filters previously proposed by Hao et al. [1] (according to them the following limits: MW ≤ 435, ClogP ≤ 6, HBA ≤ 6, HBD ≤ 2, RB ≤ 9, ABC ≤ 17 were established for 788 pesticides). From the 1741 pesticide, only three compounds passed Hao's filters, but not ours. We can therefore conclude that criteria proposed by us are closer to pesticide-likeness compounds.

Table 2. Mean values of molecular properties of drugs and pesticides*

	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Both series	
	D	P	D	P	D	P	D	P	D	P	D	P
N	4		101	32	171	831	472	802	154	76	902*	1741*
MW	4095.27 ± 318.7	-	762.4 ± 273.3	689.7 ± 163.21	224.49 ± 74.67	247.6 ± 77.07	346.05 ± 89.16	337.32 ± 83.7	400.76 ± 111.05	354.80 ± 131.0	395.59 ± 14.05	301.74 ± 110.21
logP	-23.60 ± 6.65	-	1.74 ± 3.94	4.72 ± 3.1	0.36 ± 2.31	2.58 ± 2.03	2.51 ± 1.87	3.72 ± 1.57	0.55 ± 3.56	0.82 ± 4.43	1.56 ± 3.25	3.07 ± 2.17
TPSA	1716.76 ± 158.81	-	203.78 ± 135.01	134.24 ± 56.26	74.79 ± 35.33	45.95 ± 28.32	70.14 ± 39.49	61.74 ± 34.14	110.28 ± 81.13	75.26 ± 94.79	100.14 ± 133.01	56.12 ± 39.44
HBA	63 ± 4.62	-	10.35 ± 5.06	8.66 ± 3.97	3.84 ± 1.98	2.21 ± 1.67	4.13 ± 2.02	3.35 ± 2.08	5.90 ± 4.27	4.26 ± 4.84	5.33 ± 5.26	2.94 ± 2.37
HBD	57.25 ± 4.5	-	5.13 ± 5.1	3.22 ± 3.81	2.04 ± 1.37	0.65 ± 0.86	1.54 ± 1.34	0.70 ± 0.86	2.88 ± 2.81	1.84 ± 3.51	2.51 ± 4.49	0.78 ± 1.29
RB	134.0 ± 12.57	-	11.07 ± 8.39	6.59 ± 6.2	3.75 ± 2.88	4.97 ± 3.55	5.32 ± 3.25	4.87 ± 2.74	4.69 ± 3.39	3.40 ± 3.49	6.13 ± 9.72	4.88 ± 3.29
ABC	32.5 ± 10.5	-	13.56 ± 9.71	7.28 ± 11.61	4.04 ± 3.19	3.13 ± 3.01	11.66 ± 4.17	11.17 ± 3.71	3.60 ± 3.55	1.54 ± 3.21	9.15 ± 6.49	6.84 ± 5.47
RC	7.75 ± 1.5	-	5.31 ± 1.77	6.78 ± 3.84	1.1 ± 0.71	0.72 ± 0.57	2.89 ± 0.96	2.31 ± 0.63	3.75 ± 0.99	3.29 ± 1.14	2.99 ± 1.61	1.68 ± 1.37
LogD	35.8 ± 7.11	-	0.87 ± 4.40	4.15 ± 3.3	-0.77 ± 2.75	2.09 ± 2.50	1.33 ± 2.04	3.23 ± 2.08	-1.33 ± 5.17	0.37 ± 5.04	0.26 ± 4.18	2.58 ± 2.62
ARC	6 ± 2.16	-	2.41 ± 1.74	5.63 ± 24.78	0.69 ± 0.55	0.53 ± 0.51	2.06 ± 0.77	1.94 ± 0.66	0.64 ± 0.64	0.26 ± 0.55	1.61 ± 1.17	1.26 ± 3.48
logS	-	-	-4.64 ± 1.33	-4.96 ± 1.62	-2.0 ± 1.47	-3.19 ± 1.71	-3.76 ± 1.18	-4.34 ± 1.22	-3.5 ± 1.26	-3.39 ±1.99	-3.48 ± 1.49	-3.76 ± 1.62

* D – drug; P – pesticide; N – number of compounds included in the cluster; logS values are not available for all compounds

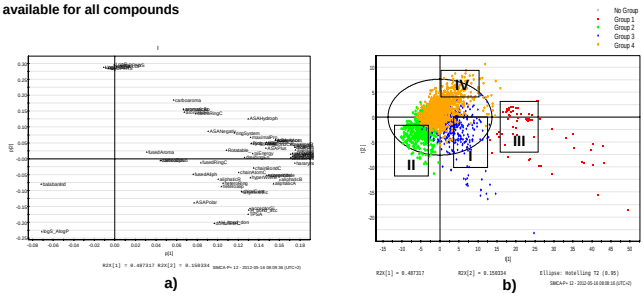


Fig 3. Principal component loadings (a) and scores (b) plots for the reduced set of drugs and pesticides

Four regions of low density including drugs and pesticides, corresponding to edge compounds, were defined in the score plot of PC1 versus PC2 (Fig. 3b). Molecular weight and number of rings increase in the positive direction of PC1. Higher compound hydrophobicity and flexibility are included in the positive direction of PC2. Polar compounds and the number of H bond acceptors and donors increase in positive direction of PC2.

Conclusions

Pesticides and drugs chemical space defined by calculated molecular properties was explored with principal component analysis. Most pesticides occupy smaller regions in comparison to drugs and might exhibit similar physicochemical and biological properties. The drug-likeness limits are larger in comparison to the pesticide ones. New upper limits for pesticide-likeness criteria were proposed.

Acknowledgements

We thank Dr. Sorel Muresan, AstraZeneca R&D Mölndal, for valuable discussions.

References

[1] Hao, G.; Dong, Q.; Yang, G., A Comparative Study on the Constitutive Properties of Marketed Pesticides. *Molecular Informatics* 2011, 30 (6-7), 614-622.
[2] Leslie Z. Benet, Fabio Broccatelli, Tudor I. Oprea, BDDCS Applied to Over 900 Drugs, *The AAPS Journal* 2011, 13(4), 519-547.
[3] The Pesticide Manual, Twelfth Edition, Ed. C.D.S. Tomlin, British Crop Protection Council, 2000
[4] Compendium of Pesticide Common Names, http://www.alanwood.net/pesticides/index.html