

· 研究论文 ·

杀菌剂、除草剂和杀虫剂的先导筛选规则

刘斌¹, 俞飞¹, 姚建华^{*1}, 廖泉¹, 范波涛²

(1. 中国科学院上海有机化学研究所, 上海 200032 2. IFODYS, 巴黎第七大学, 巴黎 75005, 法国)

摘要: 为了寻找适合于农药的先导筛选规则, 对 242 个杀菌剂、512 个除草剂和 399 个杀虫剂的 5 个描述符: 疏水常数 (LogP)、相对分子质量 (MW)、氢键给体数 (NHD)、氢键受体数 (NHA)、分子极性表面积 (PSA) 和毒性进行了计算和分析, 提出了杀菌剂、除草剂和杀虫剂先导化合物的筛选规则。测试结果显示: 杀菌剂、除草剂和杀虫剂的先导筛选规则总体正确率分别为 73.3%、53.3% 和 63.3%。

关键词: 杀菌剂; 除草剂; 杀虫剂; 筛选规则; 毒性; 先导化合物

中图分类号: O639 文献标志码: A 文章编号: 1008-7303(2007)03-0220-09

Screening Rules of Lead Compounds of Herbicide, Fungicide and Insecticide

LIU Bin¹, YU Fei¹, YAO Jian-hua^{*1}, LIAO Quan¹, Fan Bo-tao²

(1 Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032 China;
2 IFODYS, University of Paris 7, Paris 75005 France)

Abstract In order to find screening rules suitable for lead compounds of pesticide, five descriptors: logarithm of partition coefficient of octanol/water (LogP), relative molecular weight (MW), number of hydrogen-bond donors (NHD), number of hydrogen-bond acceptors (NHA), polar surface area (PSA) and carcinogenic and mutagenic toxicities of 242 fungicides, 512 herbicides and 399 insecticides were calculated and analyzed. Screening rules suitable for lead compounds of herbicide, fungicide and insecticide were proposed. The results showed that correct rate of the rules proposed for fungicides, herbicides and insecticides were 73.3%, 53.3% and 63.3%, respectively.

Key words herbicide; fungicide; insecticide; screening rules; toxicity; lead compounds

1996年, Lipinski通过对2000多种口服药物进行分析,提出了用于药物分子筛选的“5规则”,即化合物疏水常数(LogP)($CLogP \leq 5$ 或 $MLogP \leq 4.5$),相对分子质量(MW)不大于500,氢键给体数(NHD)不大于5和氢键受体数(NHA)不大于 10^{11} 。2003年,Andrew等报道了从高通量筛

选结果中获取先导化合物的规则: 化合物疏水常数(LogP)($CLogP \leq 3$) 相对分子质量小于 $450^{[2]}$ 。由于这些规则都是通过对药物分子进行分析得到的,它们并不一定适用于农药先导的筛选。对于农药先导筛选,Tice对136个除草剂以及243个杀虫剂和杀螨剂的疏水常数

收稿日期: 2007-02-06; 修回日期: 2007-07-25.

作者简介: 刘斌(1981-),男,江苏人,博士研究生;*通讯作者(Author for correspondence): 姚建华(1963-),女,上海人,副研究员,主要从事化学信息学方法和应用研究. 联系电话: 021-54925266; E-mail: yaojh@mail.sic.ac.cn

资助项目: 国家“973”项目(2003CB114400); 国家自然科学基金委项目(20473112, 20572120); 中科院海外基金项目(2003-1-3); 上海市科委项目(05ZR14054, 044307036); 国家“十一五”支撑项目(2006BAE01A05).

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

(LogP)、相对分子质量 (MW)、氢键受体数 (NHA)、氢键给体数 (NHD)、极性表面积 (PSA)、表面积 (SA) 和极性表面积与表面积的比例 (PSA/SA) 等描述符进行计算和分析, 提出了杀虫剂、除草剂的先导筛选规则^[3,4]。但是, 该除草剂筛选规则的正确率不大于 55%, 杀虫剂的正确率尚未见报道。

作者通过对 242 个杀菌剂、512 个除草剂和 399 个杀虫剂相应的描述符进行计算和分析, 在提出除草剂和杀虫剂的先导筛选规则之外, 还提出了杀菌剂的先导筛选规则。筛选规则中除了 Tice 提到的 5 个描述符: LogP, MW, NHA, NHD 和 PSA 外, 还包含了化合物的毒性指标。初步的测试结果表明, 本筛选规则的正确率高于 Tice 所提出的筛选规则。

1 研究方法

1.1 药物采集

从手册^[5]中选择杀菌剂 242 个, 除草剂 512 个, 杀虫剂 399 个。从 MDL^[6] CMC 2006 1 库中随机选择 999 个药物分子。

1.2 描述符的计算方法

在众多 LogP 的预测程序中^[7-16], 采用 CISOCLogP^[13,14] 预测化合物的 LogP 值; MW 由 MDL 公司的 ISIS_DRAW 软件计算得到; PSA 由 Chem-Office 计算而得, 所得的 PSA 为 TPSA (topological PSA)^[17]; NHD 和 NHA 均由 PoweM (V061)^[18,19] 计算得到。

由有机化合物致癌毒性预测系统 CISOCPSC^[20,21] 和致突变毒性预测系统 CISOCPST^[22] 分别预测化合物的致癌和致突变毒性。

2 计算与分析

2.1 疏水常数 (LogP)

疏水常数 (LogP) 是化合物在正辛醇-水体系中的分配系数, 表示化合物的疏水性, 常用于描述化合物对细胞的渗透能力^[1]。四类化合物的 LogP 值在不同数据范围内的分布比例见表 1。

将表 1 中单个数据范围内数据所占比例超过 10%、8% 和 5% 的区间分别相加, 得到的结果列于表 2。

表 1 四类化合物的 LogP 在各个数据范围内的比例

Table 1 Percentage of different LogP ranges of four types of compounds

数据范围 Range of LogP	杀菌剂 Fungicide (%)	除草剂 Herbicide (%)	杀虫剂 Insecticide (%)	药物 CMC (%)
≤ -1	1.24	1.56	0.75	2.60
(-1, 0]	2.48	1.56	4.51	4.82
(0, 1]	8.26	3.32	9.02	11.65
(1, 2]	16.53	14.65	12.78	15.46
(2, 3]	19.42	24.22	17.29	23.19
(3, 4]	29.34	29.30	15.04	19.38
(4, 5]	13.22	17.19	13.03	13.25
(5, 6]	6.61	6.25	14.29	7.03
(6, 7]	1.65	1.38	11.28	1.51
> 7	1.24	0.59	2.01	1.10

表 2 四类分子的 LogP 分布比例超过 5% 的数据范围

Table 2 Ranges of LogP of four types of chemicals with the percentage greater than 5%

百分比 Percent	杀菌剂 Fungicide		除草剂 Herbicide		杀虫剂 Insecticide		药物 CMC	
	数据范围	比例	数据范围	比例	数据范围	比例	数据范围	比例
	Range	Ratio (%)	Range	Ratio (%)	Range	Ratio (%)	Range	Ratio (%)
> 10	(1, 5]	78.51	(1, 5]	85.35	(1, 7]	83.70	(0, 5]	82.93
> 8	(0, 5]	86.78	(1, 5]	85.35	(0, 7]	92.73	(0, 5]	82.93
> 5	(0, 6]	93.39	(1, 6]	91.60	(0, 7]	92.73	(0, 6]	89.96

根据表 2 中的数据, 将选择范围定在大于 8%, 且各区间中化合物所占的比例之和应大于

80%。因此, 杀菌剂先导的 LogP 范围为 (0, 5], 除草剂的为 (1, 5], 杀虫剂的为 (0, 7], 而药物先导

的为 (0, 5]。

2.2 极性表面积 (PSA)

极性表面积是化合物中极性原子的表面积总和, 它是对化合物的溶解和渗透过程进行模拟时所使用的一个重要参数^[23, 24]。分子的极性表面积在化合物分子对生物细胞膜的渗透过程有很大影

响^[17, 25]。因此, 极性表面积常用于表征分子在生物体内的传输特性^[26, 27]。化合物的极性表面积数据分布见表 3。

与对表 1 的处理方法相同, 将表 3 中单个数据范围内数据所占比例超过 10%、8% 和 5% 的区间分别相加, 将得到的结果列于表 4。

表 3 四类化合物的 PSA 在各个数据范围内的比例

Table 3 Percentage of different PSA ranges of the four types of compounds

数据范围 Range of PSA	杀菌剂 Fungicide(%)	除草剂 Herbicide(%)	杀虫剂 Insecticide(%)	药物 CMC(%)
[0, 20)	8.16	3.44	8.21	11.81
[20, 40)	24.08	16.79	17.41	17.52
[40, 60)	29.39	20.61	24.13	23.32
[60, 80)	17.96	23.09	28.86	19.02
[80, 100)	9.80	16.03	13.68	14.81
[100, 120)	2.86	7.63	4.23	5.71
[120, 140)	4.08	2.10	0.75	4.60
[140, 160)	2.45	3.05	0.25	1.20
[160, 180)	0	3.63	0.50	0.90
≥ 180	1.22	3.63	1.99	1.10

表 4 四类化合物 PSA 分布比例超过 5% 的数据范围

Table 4 Ranges of PSA of four types of chemicals with the percentage greater than 5%

百分比 Percent	杀菌剂 Fungicide		除草剂 Herbicide		杀虫剂 Insecticide		药物 CMC	
	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)
> 10	[20, 80)	71.43	[20, 100)	76.53	[20, 100)	84.08	[0, 100)	86.49
> 8	[0, 100)	89.39	[20, 100)	76.53	[0, 100)	92.29	[0, 100)	86.49
> 5	[0, 100)	89.39	[20, 120)	84.16	[0, 100)	92.29	[0, 120)	92.19

根据表 4 中的数据, 同样将选择范围定在超过 8%, 并且各区间中化合物所占比例之和超过 80%, 这样所得到杀菌剂和杀虫剂先导的 PSA 范围为 [0, 100), 与药物的相同。如果所有范围中化合物所占比例之和未达到 80%, 将导致其结果不足以表达大部分化合物的特征, 所以, 对于除草剂其选择范围定在大于 5%, 因此其 PSA 的范围是 [20, 120)。

2.3 相对分子质量 (MW)

相对分子质量是物质的分子或特定单元的平均质量与核素 ¹²C 原子质量的 1/12 之比。四类化合物的 MW 数据分布的详细信息列于表 5。对其数据的分析方法与 PSA 相同, 结果见表 6。选择范围与 LogP 相同, 因此, 杀菌剂、除草剂、杀虫剂和药物先导的 MW 范围分别为 (150, 400]、(200, 450]、(150, 400] 和 (150, 400]。

2.4 氢键受体数 (NHA)

氢键受体数是化合物中带有孤对电子的氮原

子和氧原子数之和。研究表明, 分子中氢键受体数与分子的溶解性和渗透性有密切关系^[2, 28, 29]。四类化合物计算得到的 NHA 的数据分布信息见表 7。数据的分析方法与 PSA 相同, 结果列于表 8。

根据表 7 和表 8 中列出的数据, 将选择范围定在大于 8%, 并且各区间中化合物所占比例之和应大于 80%。因此, 杀菌剂先导的 NHA 范围为 [1, 5]。与 PSA 的情况类似, 除草剂和杀虫剂所有大于 8% 的范围内化合物所占比例之和小于 80%, 因此将选择范围放宽到大于 5%, 这样, 除草剂先导的 NHA 范围为 [2, 6], 杀虫剂先导的 NHA 范围为 [1, 6], 而药物的为 [1, 5]。

2.5 氢键给体数 (NHD)

氢键给体数为分子中与氢原子相连的氮原子和氧原子数之和。NHD 也与化合物的溶解性和渗透性能有很密切的关系^[2, 28, 29]。四类化合物的 NHD 数据分布见表 9。数据分析方法与 PSA 相同, 结果见表 10。

表 5 四类化合物的 MW 在各个数据范围内的比例

Table 5 Percentage of different MW ranges of four types of compounds

数据范围 Range of MW	杀菌剂 Fungicide(%)	除草剂 Herbicide(%)	杀虫剂 Insecticide(%)	药物 CMC(%)
(0, 100]	1.22	1.53	1.74	0.40
(100, 150]	4.08	0.95	2.23	4.00
(150, 200]	8.57	7.44	8.19	14.01
(200, 250]	17.55	22.71	20.35	16.42
(250, 300]	25.31	20.61	19.60	23.62
(300, 350]	27.35	19.08	19.60	19.12
(350, 400]	10.61	14.84	14.14	11.21
(400, 450]	3.67	8.40	5.71	6.31
(450, 500]	0.82	3.44	2.73	2.10
> 500	0.82	0.95	5.71	2.80

表 6 四类化合物 MW 分布比例超过 5% 的数据范围

Table 6 Ranges of MW of four types of chemicals with the percentage greater than 5%

百分比 Percent	杀菌剂 Fungicide		除草剂 Herbicide		杀虫剂 Insecticide		药物 CMC	
	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)
> 10	(200, 400]	80.82	(200, 400]	77.29	(200, 400]	73.70	(150, 400]	84.38
> 8	(150, 400]	89.39	(200, 450]	85.69	(150, 400]	81.89	(150, 400]	84.38
> 5	(150, 400]	89.39	(150, 450]	93.13	(150, 450]	87.59	(150, 450]	90.69

表 7 四类化合物的 NHA 在各个数据范围内的比例

Table 7 Percentage of different NHA ranges of four types of compounds

数据范围 Range of NHA	杀菌剂 Fungicide(%)	除草剂 Herbicide(%)	杀虫剂 Insecticide(%)	药物 CMC(%)
0	6.12	1.15	4.96	1.40
1	18.78	4.58	6.95	9.51
2	22.45	22.33	13.90	19.72
3	17.55	22.71	25.31	25.33
4	13.06	17.75	21.09	20.02
5	8.57	9.35	14.14	12.21
6	5.31	7.06	7.20	6.11
7	2.04	4.39	3.23	2.50
8	2.45	5.15	0.25	0.60
> 8	3.67	5.53	2.98	0.10

根据表 9 和表 10 的数据, 将 NHD 的选择范围定在大于 8%, 并且各区间中的化合物所占的比例之和应大于 80%, 由此得到杀菌剂和除草剂先导的 NHD 范围为 [0, 2], 杀虫剂先导的 NHD 范围为 [0, 1], 而药物为 [0, 2]。

2.6 毒性

毒性系泛指化学物质对人体的毒害作用。这

种作用是潜在性的, 并与剂量和时间有关。根据已有的工作基础, 在本研究中, 主要关注化合物的致癌和致突变毒性。表 11 和表 12 列出了 7 个化合物的预测结果, 其结果均显示它们具有致癌和致突变毒性, 这与其实验结果相一致。目前, 这些农药均已被禁用^[31, 32]。

表 8 四类化合物 NHA 分布比例超过 5% 的数据范围

Table 8 Ranges of NHA of four types of chemicals with the percentage greater than 5%

百分比 Percent	杀菌剂 Fungicide		除草剂 Herbicide		杀虫剂 Insecticide		药物 CM C	
	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)
> 10	[1, 4]	71. 84	[2, 4]	62. 79	[2, 5]	74. 44	[2, 5]	77. 28
> 8	[1, 5]	80. 41	[2, 5]	72. 14	[2, 5]	74. 44	[1, 5]	86. 79
> 5	[0, 6]	85. 71	[2, 6]	79. 20	[1, 6]	88. 59	[1, 6]	92. 90

表 9 四类化合物的 NHD 在各个数据范围内的比例

Table 9 Percentage of NHD distributions of four types of compounds

数据范围 Range of NHD	杀菌剂 Fungicide(%)	除草剂 Herbicide(%)	杀虫剂 Insecticide(%)	药物 CM C(%)
0	55. 10	42. 18	69. 23	35. 54
1	32. 65	35. 50	22. 08	30. 43
2	8. 16	18. 70	5. 21	23. 62
3	0. 82	2. 86	2. 73	6. 51
> 3	3. 27	0. 76	0. 74	3. 90

表 10 四类化合物 NHD 分布比例超过 5% 的数据范围

Table 10 Ranges of NHD of four types of chemicals with the percentage greater than 5%

百分比 Percent	杀菌剂 Fungicide		除草剂 Herbicide		杀虫剂 Insecticide		药物 CM C	
	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)	数据范围 Range	比例 Ratio(%)
> 10	[0, 1]	87. 76	[0, 2]	96. 37	[0, 1]	91. 32	[0, 2]	89. 59
> 8	[0, 2]	95. 92	[0, 2]	96. 37	[0, 1]	91. 32	[0, 2]	89. 59
> 5	[0, 2]	95. 92	[0, 2]	96. 37	[0, 2]	96. 53	[0, 3]	96. 07

3 结果与讨论

基于以上的工作, 提出了杀菌剂、除草剂和杀虫剂先导的筛选规则, 见表 13。

表 13 中的数据表明: 1) 三种农药的 LogP、MW、PSA、NHA 和 NHD 的范围各不相同; 2) 杀虫剂与杀菌剂相关描述符的相似度比它们与除草剂的要大, 这也与经验说法: “虫菌不分家”相吻合。与 Lipinski 等提出的“5 规则”(见表 14)相比较可以知道, 用于药物的先导筛选规则并不适用于农药先导的筛选。另外, 表 2、4、6、8 和 10 中的信息显示, 用于药物分子的 5 个描述符的计算值与用于 3 类农药的并不相同。

本文提出的筛选规则与 Tice 的规则(见表 14)相比较, 两者在分类上有所不同。本筛选规则适用于杀菌剂、除草剂和杀虫剂先导筛选, 对除草剂未作进一步分类; 而 Tice 的筛选规则中未涉及杀菌剂, 但对除草剂作了芽前和芽后的分类。

在数量上, 本筛选规则增加了毒性判断规则, 这对有效降低农药创制成本, 提高其创制效率非常重要。

为了验证本文提出的筛选规则的正确率, 在训练集之外分别收集了杀菌剂、杀虫剂和除草剂各 30 个, 对本文提出的规则进行测试, 即用前面提及的软件, 计算测试集中的化合物的 5 个描述符, 观察它们是否在筛选规则的范围, 若在规则的范围, 则表示测试结果正确, 否则为错误。测试结果见表 15。

与 Tice 规则的测试结果^[4]相比, 本文提出的除草剂先导的筛选规则的总体正确率(53.3%)高于 Tice 提出的芽后筛选规则(50.0%) (芽前的未见报道)。而对于芽后筛选规则中 4 个描述符(相对分子质量、疏水常数、氢键给体和受体数)组合的正确率而言, 本筛选规则的正确率为 56.7%, Tice 的为 70.8%; 两个描述符(疏水常数、极性表面积)组合

表 11 部分农药分子的致癌毒性预测与实验结果对照

Table 11 Information about carcinogenic toxicity by prediction and experiment

No	农药 Pesticides	预测结果 C ISOC-PSCT			实验结果 Toxic effect
		可预测性 Predictability (%)	致癌可能性 Carcinogenicity Possibility	不致癌可能性 Carcinogenicity impossibility	
1	五氯苯酚 Pentachlorophenol	100	0.996	0.001	致癌 ^a (肾上腺皮质肿瘤, 肝癌) Adrenal cortex tumors < Endocrine Carcinogenic by RTECS criteria < Tumorigenic Tumors < Liver ^a
2	六氯苯 Hexachlorobenzene (HCB)	100	0.990	0.001	致癌 (肝癌) ^a Carcinogenic by RTECS criteria < Tumorigenic Tumors < Liver ^a
3	甲胺磷 Methidathos	37	0.999	0.001	有害甚至致命 (吞入, 吸入或被皮肤吸收) ^b Harmful or fatal if swallowed, inhaled, or absorbed through the skin ^b
4	滴滴涕 DDT	100	0.997	0.001	致癌 (甲状腺、子宫肿瘤) ^a Equivocal tumorigenic agent by RTECS criteria < Tumorigenic Thyroid tumors < Endocrine Uterine tumors < Tumorigenic Effects < Reproductive Effects ^a
5	γ -1, 2, 3, 4, 5, 6-六 氯环己烷 Gamma-HCH	100	0.959	0.001	致癌 (肝癌, 肺癌) ^a Neoplastic by RTECS criteria < Tumorigenic Tumors < Liver Tumors < Lungs Thorax, or Respiration ^a
6	1, 2-二溴乙烷 1, 2-Dibromoethane	90	0.951	0.001	致癌 (对嗅觉及其他感觉器官有影响) ^a Equivocal tumorigenic agent by RTECS criteria < Tumorigenic Tumors < Olfaction < Sense organs and special Senses ^a
7	除草醚 Nitrofen	97	0.451	0.001	致癌 (肝癌, 肺癌) ^a Carcinogenic by RTECS criteria < Tumorigenic Tumors < Liver Tumors < Lungs Thorax or Respiration ^a

^a毒性数据收集自 MDL 毒性数据库 (MDL Toxicity Database), 其中 RTECS 是化学物质毒性作用登记; ^b美国环境保护署 (USEPA) ³²。

^a Data from MDL Toxicity Database RTECS Registry of toxic effects of chemical substance; ^b USEPA (U. S. Environment Protection Agency) ³².

表 12 部分农药分子的致突变毒性预测结果与实验结果对照

Table 12 Information about mutagenic toxicity of some pesticides by prediction and experiment

No	农药分子 Pesticides	预测结果 (沙门氏菌) C ISOC-PSCT (Salmonella typhimurium)			实验结果 Toxic Effect
		可预测性 Predictability (%)	致突变可能性 Mutagenicity- Possibility	不致突变可能性 Mutagenicity- Impossibility	
1	五氯苯酚 Pentachlorophenol	100	0.01	0.21	不致突变 (沙门氏菌) ^a Negative < Mutagenic (salmonella typhimurium) ^a
2	六氯苯 Hexachlorobenzene	100	0.01	0.39	不致突变 (沙门氏菌) ^a Negative < Mutagenic (salmonella typhimurium) ^a
3	甲胺磷 Methidathos	86	0.01	0.13	致突变 (鼠) ^b 不致突变 (沙门氏菌) ^b Positive < Mutagenic (rat no information about salmonella typhimurium) ^b Negative < Mutagenic (salmonella typhimurium) ^b
4	滴滴涕 DDT	100	0.01	0.62	不致突变 (沙门氏菌) ^a Negative < Mutagenic (salmonella typhimurium) ^a

Continued

No	农药分子 Pesticides	预测结果(沙门氏菌) C ISOC-PSCT (<i>Salmonella typhimurium</i>)			实验结果 Toxic Effect
		可预测性 Predictability (%)	致突变可能性 Mutagenicity- Possibility	不致突变可能性 Mutagenicity- Impossibility	
		5	γ-1, 2, 3, 4, 5, 6- 六氯环己烷 Gamma-HCH	100	
6	1, 2-二溴乙烷 1, 2-Dibromoethane	100	0.44	0.01	致突变(沙门氏菌) ^a Positive< Mutagenic (<i>salmonella typhimurium</i>) ^a
7	除草醚 Nitrofen	100	0.19	0.01	致突变(沙门氏菌) ^a Positive< Mutagenic (<i>salmonella typhimurium</i>) ^a

^a毒性数据收集自 MDL 毒性数据库 (MDL Toxicity Database); ^b参考文献 [33]; ^c参考文献 [34].

^aData from MDL Toxicity Database; ^bRef. [33]; ^cRef. [34].

表 13 杀菌剂、除草剂和杀虫剂先导的筛选规则

Table 13 Screening rules of lead compounds for fungicide, herbicide and insecticide

描述符 Descriptor	杀菌剂 Fungicide	除草剂 Herbicide	杀虫剂 Insecticide
疏水常数 LogP	(0, 5]	(1, 5]	(0, 7]
极性表面积 PSA	[0, 100)	[20, 120)	[0, 100)
相对分子质量 MW	(150, 400]	(200, 450]	(150, 400]
氢键受体数 NHA	[1, 5]	[2, 6]	[1, 6]
氢键给体数 NHD	[0, 2]	[0, 2]	[0, 1]

表 14 Lipinski“5规则”和 Tice规则^[1, 3, 4]

Table 14 Lipinski's Rule of 5 and Tice's Rule

描述符 Descriptor	5规则 Rule of 5		Tice规则 Tice's Rules	
	口服药物 Oral drug	芽后除草剂 Post-emergence herbicide	芽前除草剂 Pre-emergence herbicide	杀虫剂 Insecticide
疏水常数 LogP	≤ 4.15	≤ 3.5	—	[0, 5]
极性表面积 PSA	—	[32, 7, 129, 1]	[16, 3, 74, 4]	[14, 8, 82, 7]
相对分子质量 MW	≤ 500	[169, 1, 465, 4]	[187, 3, 396, 4]	[185, 6, 502, 9]
氢键给体数 NHD	≤ 10	[0, 3]	[0, 2]	[0, 2]
氢键受体数 NHA	≤ 5	[2, 12]	[2, 7]	[1, 7]
致毒可能性 Tox. Possibility	—	—	—	—

的正确率为 67.7%, Tice的为 62.5%。可见, 总体而言, 本文提出的除草剂先导筛选规则的正确率高于 Tice筛选规则的正确率。本文的杀菌剂和杀虫剂先导筛选规则总体正确率分别为 73.3% 和 63.3%。

4 结论

通过对 242个杀菌剂、512个除草剂和 399个

杀虫剂的 5个描述符的分布进行分析, 得出适用于杀菌剂、除草剂和杀虫剂先导的筛选规则, 并进行了正确率测试。初步的测试结果表明, 其总体正确率分别为 73.3%、53.3% 和 63.3%。尽管其正确率比 Tice的有所提高, 但仍需要作进一步完善和优化, 如增加化合物的其他参数及对应的数据范围、增加其他毒性的预测、优化和发展相关参数计算和分析方法等。

表 15 筛选规则的正确率

Table 15 Correct rate of the screening rules

筛选规则 Screening rules	正确率 Correct rate (%)		
	杀菌剂 Fungicide	除草剂 Herbicide	杀虫剂 Insecticide
相对分子质量 MW	83.3	80.0	86.7
氢键受体数 NHA	90.0	83.3	96.7
氢键给体数 NHD	90.0	90.0	90.0
极性表面积 PSA	86.7	83.3	93.3
疏水常数 LogP	86.7	80.0	83.3
MW, LogP, NHD, NHA	73.3	56.7	63.3
PSA, LogP	80.0	67.7	76.7
MW, LogP, NHD, NHA, PSA	73.3	53.3	63.3

参考文献:

- 577.
- [1] LIPINSKI C A, LOMBARDO F, DOMINY B W, et al. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings [J]. *Adv Drug Delivery Rev*, 2001, 46: 3-26.
- [2] BAXTER A, BENNEN C, BENT J et al. Hit-to-lead Studies: The Discovery of Potent Orally Bioavailable Triazolethiol CXCR2 Receptor Antagonists [J]. *Bioorg Med Chem Lett*, 2003, 13: 2625-2628.
- [3] TICE C M. Selecting the Right Compounds for Screening: Does Lipinski's Rule of 5 for Pharmaceuticals Apply to Agrochemicals [J]. *Pest Manage Sci*, 2001, 57: 3-16.
- [4] TICE C M. Selecting the Right Compounds for Screening: Use of Surface-area Parameters [J]. *Pest Manage Sci*, 2002, 58: 219-233.
- [5] TOMLIN C D S. The e-Pesticide Manual [M]. Version 2.2, Ed., British Crop Protection Council, 2002-2003.
- [6] [EB/OL]. [2007-01-20]. <http://www.mdli.com>.
- [7] xlogp [CP/OL]. (2006-12-24), [2007-01-02]. <ftp://ftp2.ipc.pku.edu.cn/pub/software/xlogp/>.
- [8] WANG R X, FU Y, LAI L H. A New Atom-additive Method for Calculating Partition Coefficients [J]. *J Chem Inf Comput Sci*, 1997, 37: 615-621.
- [9] CambridgeSoft. C S Chem 3D 9.0 for Windows User's Guide [CP/OL]. [2007-01-20]. http://www.cambridgesoft.com/services/DesktopSupport/Documentation/Manuals/files/chem3d_9_english.pdf.
- [10] LEO A J. Calculating $\log P_{\text{oct}}$ from Structures [J]. *Chem Rev*, 1993, 93: 1281-1306.
- [11] HyperCube Inc. HyperChem 7.5 Feature Summary [EB/OL]. (2007-05-18), [2007-6-20]. <http://www.hyper.com/Products/HyperChem/Professional/ProFeatures/tabid/362/Default.aspx>.
- [12] GHOSE A K, CRIPPEN G M. Atomic Physicochemical Parameters for Three-dimensional Structure-directed Quantitative Structure-activity Relationships: I. Partition Coefficients as a Measure of Hydrophobicity [J]. *J Comput Chem*, 1986, 7: 565-577.
- [13] LIAO Q, YAO J H, YUAN S G. SVM Approach for Predicting $\log P$ [J]. *Mol Divers*, 2006, 10: 301-309.
- [14] CISOC-LogP: 有机化合物疏水常数预测系统 [Z]. 软著登字第 040394号 (No. 040394), 登记号为 2005SR08893 (Register No. 2005SR08893).
- [15] REKKER R. The Hydrophobic Fragmental Constant [M]. Amsterdam: Elsevier Scientific Publishing, 1977.
- [16] REKKER R, MANNHOLD R. Calculation of Drug Lipophilicity [M]. Berlin: VCH Weinheim, 1992.
- [17] ERTL P, ROHDE B, SELZER P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-based Contributions and its Application to the Prediction of Drug Transport Properties [J]. *J Med Chem*, 2000, 43: 3714-3717.
- [18] LIU J, FENG J, BROOKS A, et al. PoweMV: A Software Environment for Statistical Analysis, Molecular Viewing, Descriptor Generation, and Similarity Search [EB/OL]. (2006-06-26), [2006-10-23]. <http://www.niss.org/PoweMV/>.
- [19] LIU K J, FENG J, YOUNG S. PoweMV: A Software Environment for Molecular Viewing, Descriptor Generation, Data Analysis and Hit Evaluation [J]. *J Chem Inf Model*, 2005, 45: 515-522.
- [20] LIAO Q, YAO J H, LI F, et al. CISOC-PSCT: A Predictive System for Carcinogenic Toxicity [J]. *SAR and QSAR in Environmental Research*, 2004, 15(3): 217-235.
- [21] CISOC-PSCT, 化合物致癌毒性预测系统 [Z]. 软著登字第 022239号 (No. 022239), 登记号为 2004SR03838 (Register No. 2004SR03838).
- [22] CISOC-PSMT, 化合物致突变毒性预测系统 [Z]. 软著登字第 042160号 (No. 042160), 登记号为 2005SR10659 (Register No. 2005SR10659).
- [23] HERMANN R B. Theory of Hydrophobic Bonding. II. Correlation of Hydrocarbon Solubility in Water with Solvent Cavity Surface Area [J]. *J Phys Chem*, 1972, 76: 2754-2759.
- [24] PEARLMAN R S. Molecular Surface Areas and Volumes and their Use in Structure-Activity Relationships: Physical Chemical Properties of Drugs. *Medicinal Research Series Vol 10* [M].

- New York: Marcel Dekker Inc, 1980: 321-347.
- [25] HOU T J, ZHANG W, XIA K, et al. ADMET Evaluation in Drug Discovery. 5. Correlation of Caco-2 Permeation with Simple Molecular Properties [J]. J Chem Inf Comput Sci, 2004, 44: 1585-1600.
- [26] PAIM K, STENBERG P, LUTHMAN K, et al. Polar Molecular Surface Properties Predict the Intestinal Absorption of Drugs in Humans [J]. Pharmaceutical Research, 1997, 14: 568-571.
- [27] DAVID E C. Rapid Calculation of Polar Molecular Surface Area and its Application to the Prediction of Transport Phenomena. 1. Prediction of Intestinal Absorption [J]. J Pharm Sci, 1999, 88: 807-814.
- [28] ABRAHAM M H, CHADHA S H, WHITING G S, et al. Hydrogen Bonding Octanol and Water-alkane Partitioning and the Delta logP Parameter of Seiler [J]. J Pharm Sci, 1994, 83: 1085-1100.
- [29] PATERSON D A, CONRAD I R A, HILGERS A R, et al. A Non-aqueous Partitioning System for Predicting the Oral Absorption Potential of Peptides [J]. Quant Struct-Act Relatsh, 1994, 13: 4-10.
- [30] BANNED Pesticides [EB/OL]. (2007-10-21), [2007-10-31]. http://thailand.ipm-info.org/pesticides/pesticides_banned.htm.
- [31] History of POPs Discovery, Use and Ban [EB/OL]. (2004-04-05), [2006-10-15]. <http://www.cneac.com/article/list.asp?id=25>.
- [32] Chemical Emergency Preparedness and Prevention [EB/OL]. (2006-05-05), [2006-11-11]. <http://yosemite.epa.gov/oswer/ceppoehnsf/profiles/10265-92-09/openDocument>.
- [33] Ipsc International Programme on Chemical Safety [EB/OL]. [2006-11-14]. <http://www.inchem.org/documents/hsg/hsg/hsg079.htm>.
- [34] NANTEL A. J. LINDANE [EB/OL]. (2001-09-17), [2006-11-14]. <http://www.inchem.org/documents/pim/s/chemical/pim859.htm#L1%20Substance>.

(Ed. JIN SH)

欢迎订阅《现代农药》(双月刊)

《现代农药》是中国农药行业科技类重点刊物, 网站 <http://nyxd.chinajournal.net.cn>, 刊号 CN 32-1639/TQ (国内)、ISSN 1671-5284 (国际)。《现代农药》发行量大, 影响面广, 已成为《美国化学文摘》(CA) 收录期刊、中国科技论文统计源期刊(中国科技核心期刊)、中国核心期刊(遴选)数据库收录期刊、中国学术期刊综合评价数据库(CAJCED)统计源期刊、中国期刊全文数据库(CJFD)全文收录期刊、中文科技期刊数据库收录期刊等。

《现代农药》及时报道中国农药研究技术最新进展, 密切关注国内外农药机构、农药品种及市场的最新动态, 着力展示中国农药行业发展水平、促进农药界的技术交流与进步。主要栏目有: 专论与综述、研究与开发、创制与生测、环境与残留、农药分析、世界农药、品种介绍、生物农药与生物技术、农药应用、试验简报等。本刊是展示现代农药最新科技进展的前沿平台, 适合于广大从事农药生产、科研、开发、教学、管理、销售、推广及应用等人员订阅。

《现代农药》定价 60.00 元/年, 10.00 元/期。邮发代号 28-304, 您可在全国各邮局订阅, 也可直接通过编辑部订阅。

请订阅者按下列地址汇款:

开户行: 工商银行南京汉中门支行

帐号: 4301030109100015508

收款单位: 江苏省农药研究所股份有限公司

邮汇地址: 南京市螺丝桥 80 号《现代农药》编辑部

邮编: 210019

E-mail: agrochem@263.net

联系电话: 025-86581148

传真: 025-86581147

联系人: 顾群

网址: <http://www.agroinfo.com.cn>