

## Calculation of molecular lipophilicity and drug likeness for few heterocycles

P. LALITHA\* and S. SIVAKAMASUNDARI

Department of Chemistry, Avinashilingam (Deemed University),  
Coimbatore - 641 043 (India).

(Received: December 11, 2009; Accepted: February 20, 2010)

### ABSTRACT

Log P value for 16 vinyl, nitrogen and sulphur heterocycles were calculated using eight computer programs: HyperChem 7.0 (based on atom contributions), XLOGP (based on atom contributions), KowWin (based on atom/fragment contributions), CLOGP (based on fragmental contributions), ALOGPS 2.1 (based on atom-type electro topological-state indices and neural network modeling), and IA logP (based on atom-type electrotopological-state indices and neural network modeling), miLogP using Molinspiration and MolSoft. The drug likeness has been calculated using Molinspiration and MolSoft programs. Almost all the heterocycles were found to obey Lipinski's rule of 5 for an orally active drug.

**Key words:** Molecular lipophilicity, drug likeness and heterocycles.

### INTRODUCTION

The partition coefficient for octanol–water ( $\log P_{ow}$ ) has become the preferred measure for lipophilicity in the development of biologically active molecules, in which transport across biological membranes is often critical<sup>1-3</sup>. The widespread application of lipophilicity to drug design explains the need for quick procedures to quantify molecular lipophilicity, particularly at the screening level<sup>4-5</sup>. Giving consideration to the significance of  $\log P$ , this paper is focused on the theoretical determination of  $\log P$  and other molecular properties like refractivity, polarisability, polar surface area,  $\log S$ , molecular volume, drug likeness and number of rotatable bonds, using computer programs.

### MATERIAL AND METHODS

Log P value for 16 vinyl nitrogen and

sulphur heterocycles (Fig 1) were calculated using eight computer programs: HyperChem 7.0 (based on atom contributions)<sup>6</sup>, XLOGP (based on atom contributions<sup>7</sup>), KowWin (based on atom/fragment contributions<sup>8</sup>), CLOGP (based on fragmental contributions<sup>9</sup>), ALOGPS 2.1 (based on atom-type electro topological-state indices and neural network modeling)<sup>10</sup>, IA logP (based on atom-type electrotopological-state indices and neural network modeling), miLogP using Molinspiration and MolSoft. The drug likeness has been calculated using Molinspiration and MolSoft programs.

### Calculation of Lipophilicity ( $\log P$ )

In our work, eight computer programs based on different calculation methods for computing  $\log P$  have been compared.

### Hyperchem 7.0.

The computer program HyperChem 7.0 calculates  $\log P$  values according to Ghose, Prichett

and Crippen<sup>11</sup>. Their method avoids correction factors and estimates lipophilicity on the individual atom basis. The program lists 51 atom contributions for each atom-type. The log P value is estimated by summing up all atom contributions.

### XLOGP

The XLOGP is another atom-additive method, based on summation of atomic contributions, but it includes ten additional correction factors for some intramolecular interactions. Atoms are classified by their hybridization states and neighboring atoms.

### (LogKow) Kowwin

The LogKow (KowWin) program calculates log P values of organic chemicals using the atom/fragment method. The log P value of the calculated compounds is by simply summing up all atom/

fragment contribution values and correction factors. CLOGP The CLOGP program is based on the fragmental method developed by Leo and Hansch.

### ALOGPS 2.1

The ALOGPS 2.1 package included a program for predicting the lipophilicity of chemical compounds. A method for predicting log P values is based on atom-type electro topological state (E-state) indices and neural network modeling developed by I. V. Tetko *et al.*,<sup>10</sup>.

### IA logP

The IA logP program predicts lipophilicity of compounds using neural network algorithms and E-state atom indices. The IAlogP predictor was developed by Interactive Analysis, using neural networks technology to derive a set of 10 fold cross-validated networks.

### Molinspiration

#### LogP (octanol/water partition coefficient)

LogP is calculated by the methodology developed by Molinspiration as a sum of fragment-based contributions and correction factors. Method is very robust and is able to process practically all organic and most organometallic molecules.

#### Milog P

Method for logP prediction developed at Molinspiration is based on group contributions. These have been obtained by fitting calculated logP with experimental logP for a training set more than twelve thousand, mostly drug-like molecules.

#### Molecular property predictions using MolSoft

In MolSoft all molecular property predictors are calculated using fragment-based contributions. Other properties calculated

#### HYDRATION ENERGY kcal/mol

Hydration energy is very important in the selectivity filter of ion channels where the drug is almost entirely strip from the hydration water.

#### Refractivity A<sup>3</sup>

There are several potential classes of parameters used in QSAR studies. The selection of parameters is an important first step in any QSAR study. If the association between the parameter(s)

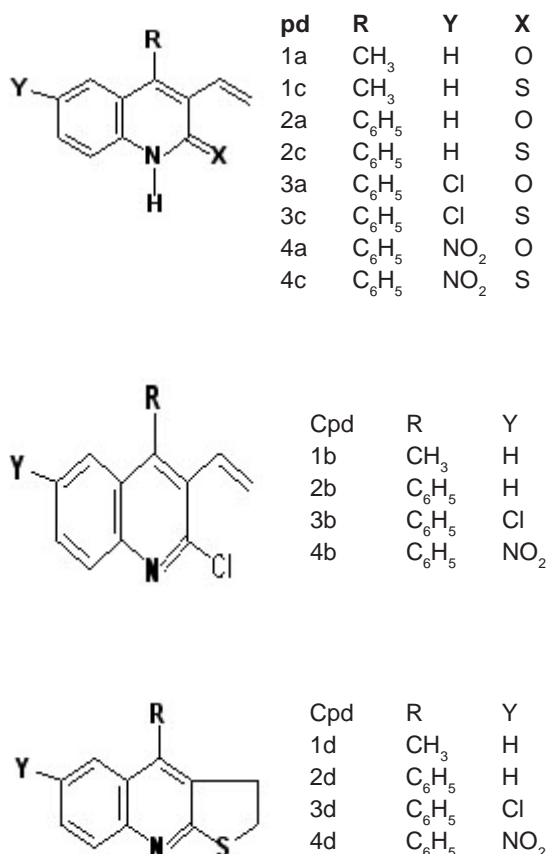


Fig. 1.

selected and activity is strong, then activity predictions will be possible. One such important parameter is refractivity.

### Polarisability A<sup>3</sup>

A challenge facing the practical application of computational alchemy to drug discovery is insuring the adequacy of the underlying potential energy function. Although existing potential functions that assume pairwise additivity of nonbonded interatomic interactions appear to be sufficient in many cases the description of certain interactions may require explicit treatment of electronic polarisability.

Molecular Polar Surface Area TPSA is calculated based as a sum of fragment contributions. O- and N- centered polar fragments are considered. Briefly, the procedure is based on the summation of tabulated surface contributions of polar fragments.

violations -"Rule of 5" Properties is set of simple molecular descriptors used by Lipinski in formulating his "Rule of 5". Molecules violating more

than one of these rules may have problems with bioavailability.

### nrotb-Number of Rotatable Bonds

This simple topological parameter is a measure of molecular flexibility. It has been shown to be a very good descriptor of oral bioavailability of drugs.

### Molecular Polar Surface Area PSA

Molecular polar surface area (PSA) is a very useful parameter for prediction of drug transport properties.

### Molecular volume

Molecular volume determines transport characteristics of molecules, such as intestinal absorption or blood-brain barrier penetration. Volume is therefore often used in QSAR studies to model molecular properties and biological activity.

### Molecular properties predictions using MolSoft

Drug-likeness score predicts an overall drug-likeness score using a Support Vector Machine (SVM) classifier and MolSoft's fingerprints. Drug

**Table 1: Results of experimental determination of log P value and other properties using Hyperchem 7.0**

Cpd	LOG P	Hydration energy kcal/mol	Refractivity A <sup>3</sup>	Polarisability A <sup>3</sup>	Mass amu	Volume A <sup>3</sup>
1a	0.28	-4.21	60.83	21.73	185.23	570.02
1b	1.79	-2.32	64.93	23.15	203.67	582.98
1c	1.40	-5.07	69.12	24.88	201.29	587.44
1d	1.02	-1.92	65.9	23.64	201.29	604.23
2a	0.73	-5.58	85.07	29.55	247.3	689.08
2b	2.24	-3.71	89.16	30.97	265.74	715.61
2c	1.85	-6.61	93.35	32.71	263.36	722.95
2d	1.48	-3.61	90.14	31.46	263.36	735.73
3a	1.24	-4.94	88.92	31.57	282.75	741.69
3b	2.75	-3.09	93.01	33.00	301.19	767.94
3c	2.36	-5.95	97.21	34.73	298.81	774.26
3d	1.99	-3.04	93.99	33.49	298.81	788.95
4a	-2.01	-10.02	90.78	31.39	292.29	752.22
4b	-0.50	-8.07	94.88	32.81	310.74	763.3
4c	-0.89	-10.61	99.07	34.55	308.35	773.16
4d	-1.06	-8.65	101.31	36.04	316.42	815.23

Table 2: Log p (lipophilicity) and log S (aqueous solubility) calculations using ALOGPS 2.1 program

Compound	miLogP	Properties				Molinspiration drug-likeness score						
		TPSA	NON	NOHNH	Nviolations	Nrotb	Volume	GPCR ligand	Ion channel modulator	Kinase inhibitor or	Nuclear receptor ligand	
1a	2.440	32.862	2	1	0	1	176.297	-0.27	-0.33	-0.53	-1.51	
1b	4.242	12.892	1	0	0	1	181.707	-0.25	-0.21	-0.32	-0.64	
1c	3.230	15.791	1	1	0	1	186.176	-0.43	-0.33	-0.34	-1.12	
1d	3.530	12.892	1	0	0	0	181.33	-0.57	-0.68	-0.53	-1.43	
2a	3.787	32.862	2	1	0	2	231.144	0.03	-0.13	-0.07	-0.62	
2b	2.588	12.892	1	0	1	2	236.664	0.02	-0.07	0.06	0.00	
2c	4.576	15.791	1	1	0	2	240.022	-0.09	-0.13	0.07	-0.32	
2d	4.877	12.892	1	0	0	1	236.178	-0.17	-0.37	-0.05	-0.53	
3a	4.441	32.862	2	1	0	2	244.68	0.04	-0.12	-0.07	-0.63	
3b	6.242	12.892	1	0	1	2	250.09	0.03	-0.06	0.05	-0.03	
3c	5.230	15.791	1	1	1	2	263.668	-0.07	-0.12	0.06	-0.36	
3d	5.531	12.892	1	0	1	1	249.714	-0.15	-0.34	-0.05	-0.66	
4a	3.722	78.686	5	1	0	3	264.478	-0.13	-0.23	-0.19	-0.66	
4b	5.523	58.716	4	0	1	3	269.889	-0.15	-0.18	-0.08	-0.13	
4c	4.511	61.615	4	1	0	3	263.366	-0.24	-0.23	-0.07	-0.41	
4d	4.812	58.716	4	0	0	2	269.612	-0.31	-0.44	-0.17	-0.58	

Table 3: Calculations of properties and drug likeness using Molinspiration

Cpd	M. F	Log P (exp)				Log S (exp)			
		ALOGPs	IA log P	CLOGP	MLogP	KOWWIN	XLOGP	ALOGPs	IA log S
1a	C <sub>12</sub> H <sub>11</sub> NO	2.01	3.01	2.04	2.44	2.77	2.24	-2.53 (0.55g/l)	-3.82 (28.04mg/l)
1b	C <sub>12</sub> H <sub>10</sub> NCI	4.16	4.00	4.05	4.24	4.24		-3.98 (21.23mg/l)	-3.94 (23.38mg/l)
1c	C <sub>12</sub> H <sub>11</sub> NS	2.49	3.25	2.47	3.23	3.19	2.49	-3.67(42.96mg/l)	-4.90 (2.53mg/l)
1d	C <sub>12</sub> H <sub>11</sub> NS	3.37	3.56	3.40	3.53	4.18	3.17	-4.23(11.88mg/l)	-4.42 (7.65mg/l)
2a	C <sub>17</sub> H <sub>13</sub> NO	3.45	4.04	3.42	3.79	3.18	3.74	-4.20(15.46mg/l)	-5.09 (2.01mg/l)
2b	C <sub>17</sub> H <sub>12</sub> NCI	5.61	4.86	5.44	5.59	5.45	5.50	-6.05(0.23mg/l)	-5.49 (0.86mg/l)
2c	C <sub>17</sub> H <sub>13</sub> NS	4.05	5.62	3.86	4.58	4.41	4.00	-4.85(3.71mg/l)	-3.08 (0.22g/l)
2d	C <sub>17</sub> H <sub>13</sub> NS	5.19	5.10	4.79	4.88	5.40	4.68	-5.64(0.61mg/l)	-6.36 (0.11mg/l)
3a	C <sub>17</sub> H <sub>12</sub> NOCI	4.19	4.62	4.28	4.44	3.83	4.37	-4.83(4.18mg/l)	-3.65 (63.07mg/l)
3b	C <sub>17</sub> H <sub>11</sub> NCI <sub>2</sub>	6.10	5.33	6.17	6.24	6.10	6.12	-6.47(0.10mg/l)	-6.10 (0.24mg/l)
3c	C <sub>17</sub> H <sub>12</sub> NSCI	4.65	5.31	4.58	5.23	5.05	4.62	-5.29(1.53mg/l)	-3.51 (92.03mg/l)
3d	C <sub>17</sub> H <sub>12</sub> NSCI	5.93	5.77	5.57	5.53	6.04	5.30	-6.21(0.18mg/l)	-7.21 (18.36mg/l)
4a	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>3</sub>	3.65	3.87	3.48	3.72	3.40	3.64	-4.40(11.56mg/l)	-3.45 (0.10g/l)
4b	C <sub>17</sub> H <sub>11</sub> N <sub>2</sub> O <sub>2</sub> Cl	5.29	4.53	5.26	5.52	5.27	5.39	-5.32(1.49mg/l)	-6.11 (0.24mg/l)
4c	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S	4.19	4.42	3.61	4.51	4.23	3.89	-5.05(2.74mg/l)	-6.65 (69.03mg/l)
4d	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S	4.78	4.75	4.68	4.81	5.21	4.57	-5.13(2.27mg/l)	-4.47 (10.45mg/l)

**Table 4: Molecular properties and drug likeness calculated using MolSoft**

CPD	MF	MW	No. of HBA	No. of HBD	LOG P	LOGS (In Log (Moles/L))	PSA (A <sup>2</sup> )	VOL (A <sup>3</sup> )	N.S	D.L
1a	C <sub>12</sub> H <sub>11</sub> NO	185.23	1	1	2.04	-2.53	25.19	229.16	0	-0.32
1b	C <sub>12</sub> H <sub>10</sub> NCI	203.67	1	0	3.32	-3.76	10.02	156.73	0	-1.85
1c	C <sub>12</sub> H <sub>11</sub> NS	201.29	1	1	2.31	-3.20	12.43	240.65	0	-0.22
1d	C <sub>12</sub> H <sub>11</sub> NS	201.29	2	0	3.70	-5.24	12.31	194.15	0	-0.60
2a	C <sub>17</sub> H <sub>13</sub> NO	247.30	1	1	3.36	-3.83	24.92	275.45	0	0.04
2b	C <sub>17</sub> H <sub>12</sub> NCI	265.74	1	0	4.48	-5.21	9.75	211.99	0	-2.01
2c	C <sub>17</sub> H <sub>13</sub> NS	263.36	1	1	3.63	-4.50	12.16	286.94	0	0.50
2d	C <sub>17</sub> H <sub>13</sub> NS	263.36	2	0	4.43	-6.44	12.04	245.88	0	-0.56
3a	C <sub>17</sub> H <sub>12</sub> NOCl	281.74	1	1	2.79	-3.41	25.19	246.43	0	-0.52
3b	C <sub>17</sub> H <sub>11</sub> NCI <sub>2</sub>	300.19	1	0	5.2	-6.10	9.75	229.26	0	-1.32
3c	C <sub>17</sub> H <sub>12</sub> NSCl	297.81	1	1	4.37	-5.38	12.16	304.21	0	0.28
3d	C <sub>17</sub> H <sub>12</sub> NSCl	297.81	2	0	5.15	-7.33	12.04	263.15	0	-0.55
4a	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>3</sub>	292.29	3	1	2.02	-3.30	58.59	254.22	0	-1.06
4b	C <sub>17</sub> H <sub>11</sub> N <sub>2</sub> O <sub>2</sub> Cl	310.74	3	0	4.42	-5.99	43.16	237.06	0	-1.53
4c	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S	308.36	3	1	3.60	-5.26	45.56	312.00	0	-0.56
4d	C <sub>17</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S	308.36	4	0	4.38	-7.22	45.44	270.94	0	-1.02

N.S- no.of stereo centers; d.l- drug-likeness model score

likeness may be defined as a complex balance of various molecular properties and structure features, which determine whether particular molecule is drug or non-drug.

## DISCUSSION

In the relation between the structure and activity, lipophilic character log P appears to be among the most important physicochemical parameter. This paper is intended to determine the lipophilicity of 16 vinyl nitrogen and sulphur heterocycles.

It is evident from the results of HyperChem calculation of log P (Table I) that the nitro substituted compounds have a negative log P as is expected from the nature of the nitro group. The compounds 3b, 3c and 2b are found to have high log P values compared to the other compounds. It is also seen from the calculations that the refractivity of the phenyl substituents is greater than the 1series of compounds. So is the polarisability of the compounds. The log P (exp) calculated by all other programs (Table II) also show the same trend – the value being

highest for 3b compound. Compounds 2a and 2b show a positive value for kinase inhibitor score (Table III). Greater the score, greater is the inhibiting activity. Similarly compounds 2a, 2b, 3a and 3b show a positive value for GPCR ligand score.

2a,3c,2c compounds showed a positive value and hence more drug likeness as seen from the calculations of MolSoft program(**Table IV**). In order to increase the drug likeness score various modifications were tried with the compound 2c and the score was found to increase from 0.5 to 0.84 when the structure 2c had chloro substituents at 5 and 7 positions in the quinoline ring. All the computer programs were found to be relatively simple and applicable to QSAR studies. The programs are highly useful for predicting the solvation properties.

## ACKNOWLEDGEMENTS

The authors sincerely thank the authorities of Avinashilingam University for Women, Coimbatore for providing facilities to carry out this work.

## REFERENCES

1. Carrupt P. Testa B. and Gaillard P., in *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, Eds., Wiley-VCH, New York, 1899, **11**, 241(1997).
2. Hansch C. and Leo A., *Exploring QSAR Fundamentals and Applications in Chemistry and Biology*, American Chemical Society, Washington, D.D. (1995).
3. Martin Y. C., *Quantitative Drug Design: A Critical Introduction*, Marcel Dekker, New York (1978.)
4. Mannhold R. and Dross K., *Quant. Struct.-Act. Relat.*, **15**: 403 (1996).
5. van de Waterbeemd H. and Mannhold R., *Quant. Struct.-Act. Relat.*, **15**: 410 (1996).
6. Ghose K., Pritchett A., and Crippen G. M., *J. Comput.Chem.* **9**: 80 (1988).
7. Wang R., Fu Y., and Lai L., *J. Chem. Inf. Comput. Sci.* **37**: 615 (1997).
8. Meylan W. M. and Howard P. H., *J. Pharm. Sci.* **84** (1995).
9. Chou J. T. and Jurs P. C., *J. Chem. Inf. Comput. Sci.* **19**: 172 (1979).
10. Tetko V., *J. Chem. Inf. Comput. Sci.*, **42**, 717 (2002).