International Journal of Drug Design and Discovery

***Volume* 5 • *Issue* 1 • *January – March* 2014. 1255-1263**

Refining the National List of Essential Medicines 2011: Accurate Prediction of Lipophilicity, Determination of Aqueous Solubility and Assessment of Molecular Complexity

Hiba Iqbal[1], Madambil Radhakrishnan Vineetha[1] and Ranganathan Balasubramanian[1*]

[1]Department of Pharmaceutical Chemistry, Amrita School of Pharmacy, Amrita Vishwa Vidyapeetham University, AIMS Health Sciences Campus, Kochi-682041, Kerala, India.

ABSTRACT: National List of Essential Medicines (NLEM) 2011 is a very important document in the Indian healthcare scenario. Currently in its third instalment, the list has undergone a systematic evolution keeping in mind the present medical needs of the average Indian. While it has grown to be clinically relevant, NLEM has not been able to gain popularity within the medicinal chemistry community. In the work presented herein, this list is viewed as an educational tool and therefore attempts made to augment its value as well as utility to the various stakeholders, particularly those within the pharmacy profession. Pharmaceutically handy physicochemical properties including partition coefficient and aqueous solubility have been computed for the small molecule subset within the NLEM. Besides this enrichment of the list per se, the work has resulted in attaining a workable solution for the dilemma faced by practising medicinal chemists on a regular basis with respect to picking and choosing an optimal software package for computation of lipophilicity.

KEYWORDS: NLEM, Partition coefficient, Aqueous solubility, Molecular complexity, Yalkowsky GSE.

Introduction

The term "essential medicine", first introduced by the World Health Organization (WHO) in 1977, refers to a medicine that satisfies the priority healthcare needs of the bulk of the population. In India, the importance of this term culminated with the formulation of a National List of Essential Medicines (NLEM) in the year 1996 under the aegis of the Ministry of Health & Family Welfare (MoHFW), Government of India (GOI). The philosophy behind the compilation of such a list that has since been revised twice, in 2003 and subsequently in 2011, lies in promoting rational use of medicines with respect to three significant parameters *viz.* cost, safety and efficacy.

One of the potential uses envisioned for this document is its ability to serve as a tool for public education and training of healthcare providers. As one of the many stakeholders of this document, we have attempted to add value to NLEM 2011 by incorporating relevant physico-chemical properties, especially partition coefficient and solubility[1,2,3]. Although these are handy for a medicinal chemist during the drug development process, most of the drug compendia and formularies do not have this information made available publicly. Moreover, not many compounds have reliable values for these properties in the literature because the experimental techniques to determine them are quite tedious and subject to significant errors[4]. As a result, many empirical and semi-empirical algorithms and software modules have been developed to reasonably predict partition coefficient and solubility of lead-like and drug-like molecules using molecular descriptors like molecular weight, number of hydrogen bond acceptors (HBAs)/donors (HBDs), topological polar surface area (TPSA), rotatable bonds etc[5]. The multitude of computational tools available for the assessment of such physicochemical properties however poses a big challenge with regard to choosing the most reliable software for one's analysis. In this context, we have also developed a simple protocol for identifying the most optimal software out of the various ones available by utilizing the compounds present in NLEM 2011 as data points.

While partition coefficient (log P) is a lipophilicity descriptor that is computed independently, solubility (log S) is accessed *via* the general solubility equation (GSE) established by Yalkowsky and co-workers[6]. Since one needs to know the melting temperature (Tm) and log P of a given substance for using this equation, the inclusion of melting point data for the medicines in NLEM 2011 becomes a prerequisite for estimation of solubility. A logical consequence of the above determinations is our computation of a dimensionless number called fraction sp^3 (Fsp^3) which has been introduced recently as a descriptor of molecular complexity and is intimately related to the aqueous solubility of substances.

Methods

NLEM 2011, a compilation of 348 unique medicines, is freely downloadable from the Central Drugs Standard

* *For correspondence:* Ranganathan Balasubramanian,
Tel.: +91-0484-2851234 (Ext. 7919); Fax: +91-0484-2802020
Email: ranga.balasubramanian@gmail.com

Control Organization (CDSCO) website (www.cdsco.nic.in). Consideration of the mono- as well as dinitro forms of isosorbide leads to a total of 185 small molecule entities in the list[7]. 168 of which have complete partition coefficient data and were therefore used as data points in **Fig. 1**. The experimental log P (written henceforth as log P exp.) values of these compounds were obtained from the unique bio- and cheminformatics resource database called DrugBank (www.drugbank.ca). Calculated partition coefficients mol log P and C log P were extracted from the free modules Molsoft (Molsoft LLC) and Bioloom (BioByte Corporation) respectively. DrugBank was again the primary source for getting the melting point data of the small molecules with the fourteenth edition of the Merck Index serving as a secondary source. Wherever both these resources did not contain the desired information, the melting point was obtained from Chemical Abstracts. Such a thorough search yielded reliable melting points of 181 out of the targeted 185 small molecule dataset. In cases where the melting points were not sharp and therefore recorded as a range, the average was taken and plugged into the GSE. Apart from the four compounds whose melting points could not be obtained from any of the three sources listed above, twelve other moieties which had a melting point of less than 25 °C were excluded from the aqueous solubility determination. Along with the 17 molecules that did not have at least one partition coefficient value (five members of this subset also had melting points lower than 25 °C while three did not have melting point data available), this group of 25 compounds was not considered in the log S analysis. Out of the 160 small molecules for which aqueous solubility was computed, fifteen were handpicked to improve the clarity in the depiction of the cluster plots in **Fig. 2** and **Fig. 3**. Care was taken to ensure that the selected drugs were structurally as well as pharmacologically different. More importantly, the choice of these fifteen small molecules was deemed rational as their partition coefficients ranged from approximately -6 to +6 thereby enabling the inclusion of the entire drug-like chemical space in this analysis. All 160 compounds were however used as data points in the ensuing error analysis (**Table 1**) wherein the absolute average error (AAE) and root-mean-square error (RMSE) were computed for the dataset using the following equations:

$$AAE = \frac{\Sigma\left|\log S\,(\text{mol log p}) - \log S\,\text{exp.}\right|}{160} \qquad \text{…..(1)}$$

$$RMSE = \frac{\sqrt{\Sigma\left[\log S\,(\text{mol log p}) - \log S\,\text{exp.}\right]^2}}{160} \qquad \text{…..(2)}$$

Results and Discussion

Our research group had recently classified the medicines in NLEM 2011 into seven major chemical classes: small molecules, salts, inorganics, combination medicines, biologicals, enzymes and miscellaneous[7]. Small molecules which constitute more than half the medicines in the list were chosen for further analysis in this work because they represent the most structurally and pharmacologically diverse subset among the seven categories listed above. Two freely available software modules - Molsoft that gives mol log P and Bioloom that spits out C log P - were picked for addressing the issue of identifying the most reliable one from the many on offer in the market. The strategy developed herein and applied between these two modules can be extended to any number of commercial software. Along with enriching the NLEM 2011 and augmenting its role as an educational tool, this exercise also helps medicinal chemists to intelligently rely on software while having to make a choice during computational assessment of physicochemical properties of drug-like molecules.

One strategy to address the problem at hand was evolved by plotting log P exp. against computed mol log P and C log P values. These cluster plots depicted in **Fig. 1** clearly portray mol log P as a better descriptor of lipophilicity (the R^2 value in this case is higher and closer to 1) compared to C log P. This graphical analysis constitutes a direct approach to hit upon the better log P predicting software.

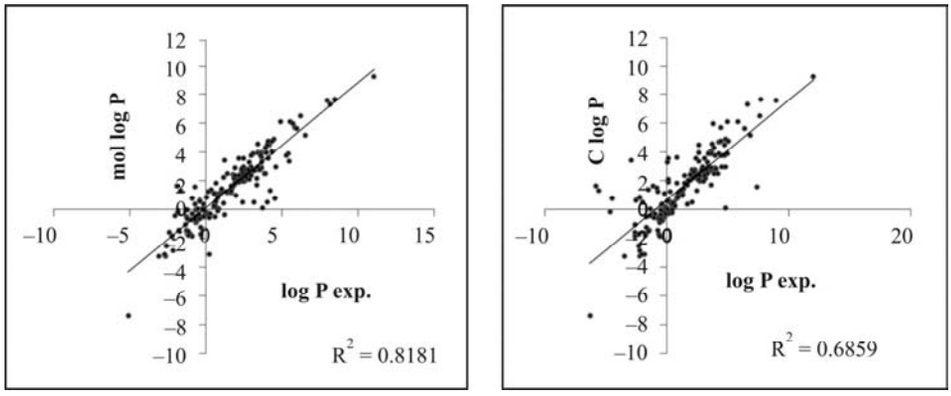

Fig. 1 Direct evaluation of computed partition coefficients. mol log P *vs.* log P exp. [left panel]; C log P *vs.* log P exp. [right panel].

A more circuitous approach to solving this problem involves the calculation of log S values for the entire small molecule dataset. As alluded to previously, the Yalkowsky GSE that relates solubility to melting point and partition coefficient has been employed for this purpose. Apart from being the most widely accepted consensus equation for the measurement of aqueous solubility of unionized compounds, it also satisfies our broader objective of value addition to the NLEM 2011 because its application entails the melting point data being made available.

$$\log S = 0.5 - 0.01 (Tm - 25) - \log P \quad \text{.....(3a)}$$

where Tm (melting point) is to be plugged in with units of degrees Celsius. For compounds that are liquids at room temperature (those that have a melting point of less than 25 °C), the term Tm – 25 is to be set as zero so that the GSE simplifies to

$$\log S = 0.5 - \log P \quad \text{.....(3b)}$$

In both forms of the equation, Yalkowsky recommends experimentally determined log P values to be used wherever available. We have therefore used log P exp. to calculate an experimental log S parameter (henceforth called as log S exp.) and used that value as a reference in further analysis. The GSE was also used to compute two other log S values, one from mol log P [referred to as log S (mol log P)] and the other from C log P [referred to as log S (C log P)]. Cluster plots (**Fig. 2**) analogous to those used in the graphical method to analyze log P were constructed. Inspection of the R^2 values demonstrated that the solubility estimates from mol log P values correlated better with log S exp. It is gratifying to note that such an indirect approach to the problem yields a result that is identical to the one obtained by the direct comparison of log P values described above. It is pertinent to reiterate here that in spite of being a little convoluted, this approach using GSE serves to add value to the NLEM 2011 *via* the customary inclusion of melting point and log S data for the medicines.

Both the approaches described above portray mol log P to be a better descriptor of lipophilicity than C log P. To corroborate this consensus finding, two types of errors *viz.* AAE and RMSE were calculated for the entire dataset[8,9]. As seen from the lower error values (**Table 1**) for log S (mol log P) versus those for log S (C logP), it is evident that mol log P once again emerges as the superior candidate and thus predicts aqueous solubility better. This method represents yet another indirect mathematical approach to address the problem of choosing the better software module for expressing partition coefficient.

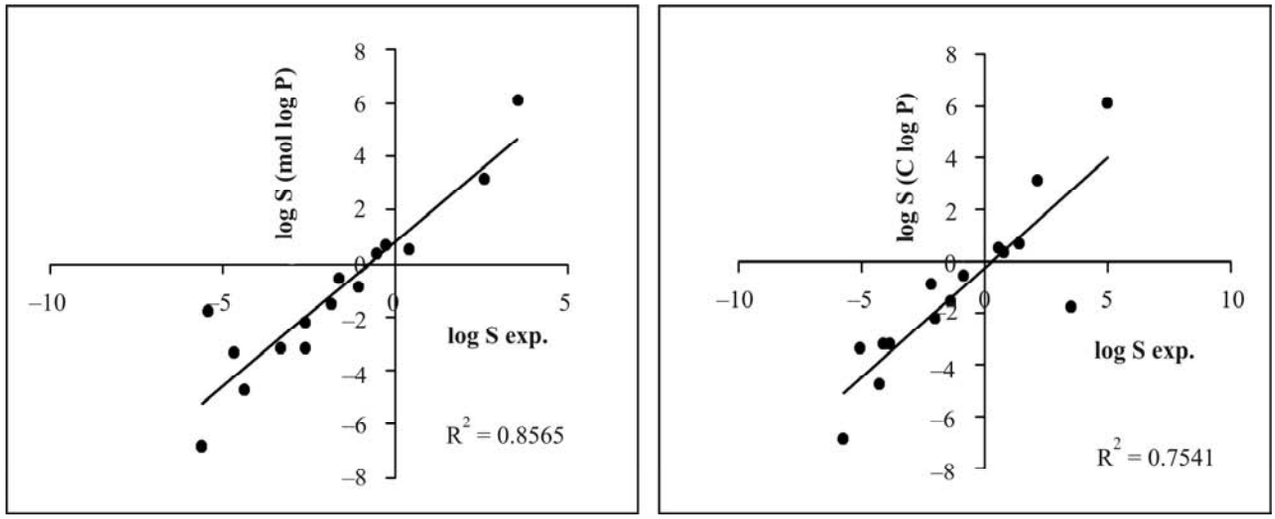

Fig. 2 Indirect evaluation of computed partition coefficients. log S (mol log P) vs. log S exp. [left panel]; log S (C log P) *vs.* log S exp. [right panel]. The Yalkowsky GSE-derived aqueous solubility determined from mol log P has a higher correlation to experimental values and therefore reinforces the superiority of mol log P over C log P as a lipophilicity descriptor for this dataset.

Table 1 Error estimates of aqueous solubility determination

Computed aqueous solubility	RMSE	AAE
log S (C log P)	0.1634	0.9601
log S (mol log P)	0.0408	0.9039

Medicinal chemistry efforts in the pharmaceutical industry have seen a deliberate shift towards relatively flat, achiral and aromatic compounds over the last couple of decades. This bias probably stemmed from the advances in coupling chemistry with respect to sp^2 hybridized carbons that made the preparation of molecules with greater unsaturation especially amenable to parallel synthesis. It is now realized that more highly complex molecules having a greater degree of saturation, and hence three-dimensionality, have better access to a larger chemical space. There is also a strong belief that a higher saturation level amounts to greater selectivity resulting in fewer off-target effects.

Recently, Lovering and co-workers[10] from Wyeth introduced a quantity Fsp^3, defined as the ratio of the number of sp^3 hybridized carbons to the total carbon count, to afford a readily interpretable measure of saturation. Since altering the saturation will also affect physical properties, we decided to explore how this parameter correlated with the solubility data calculated earlier. It is well known that aromatic molecules with extended ring systems and conjugated fragments are particularly water insoluble. Moreover, an increase in Fsp^3 concomitant with a decrease in flatness of the molecules tends to improve aqueous solubility[11]. The cluster plots (**Fig. 3**) precisely indicate such a relationship thus acting to not only corroborate the observations described above but also in reinforcing the legitimacy of our computed log S values.

The information contained in **Table 2** clearly presents the different subsets of data that have been used to carry out the various analyses described in this study. More significantly, such a comprehensive collection of properties represents a major step forward in our efforts to augment the utility of NELM to practitioners and students of medicinal chemistry.

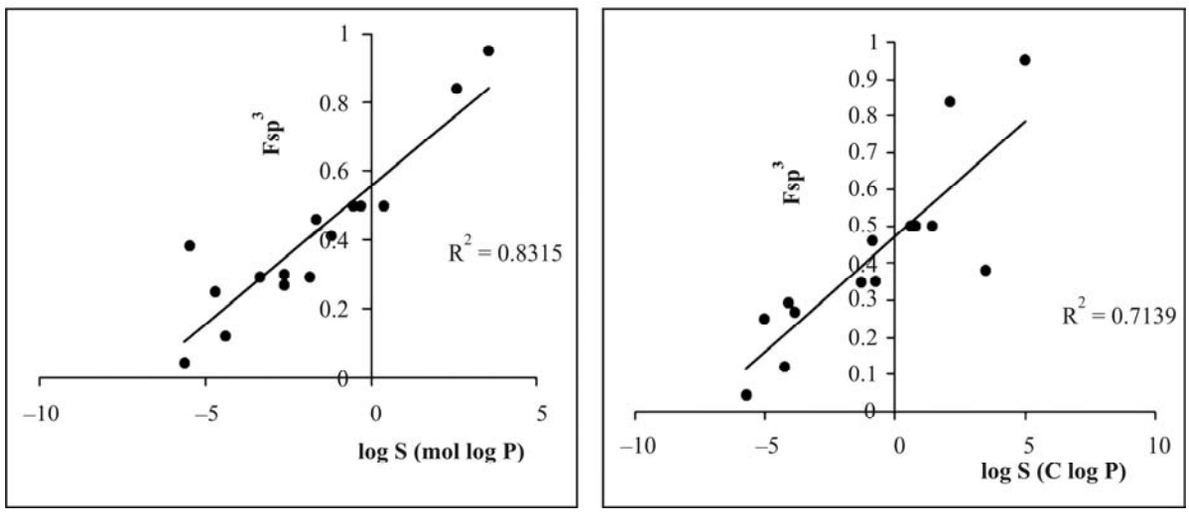

Fig. 3 The linear relationship between Fsp^3 and computed aqueous solubility values. Fsp^3 *vs.* log S (mol log P) [left panel]; Fsp^3 *vs.* log S (C log P) [right panel]. A better correlation between the flatness parameter and the aqueous solubility as estimated from mol log P supports the authenticity of the solubility determinations done earlier.

Table 2 Handy structural and physicochemical properties of the small molecules within NLEM 2011

Small molecule	No. of carbons	Fsp^3	Melting point (°C)	log P exp.	C log P	mol log P	log S exp.	log S (C log P)	log S (mol log P)
5-Amino salicylic acid	7	0	283	1.20	1.06	0.98	-3.28	-3.68	-3.06
5-Fluorouracil	4	0	283	-0.89	-0.58	-0.67	-1.18	-1.49	-1.41
Acetazolamide	4	0.25	260.5	-0.26	-0.98	-0.88	-1.59	-0.87	-0.98
Aspirin	9	0.11	135	1.19	1.02	1.41	-2.03	-1.86	-2.25
Actinomycin D	62	0.58	241.5 − 243	1.60	7.41	-0.71	-3.27	-9.08	-0.96
Acyclovir	9	0.34	255	-1.56	-2.61	-2.38	-0.24	0.81	0.58
Adenosine	10	0.50	235.5	-1.05	-2.61	-1.31	-0.54	1.01	-0.29
Albendazole	12	0.33	209	2.70	3.46	2.32	-4.09	-4.80	-3.66
Allopurinol	5	0	350	-0.55	-1.06	0.25	-2.20	-1.69	0.30
Alprazolam	17	0.11	228 − 228.5	2.12	2.55	2.33	-3.65	-4.08	-3.86

Table 2 *Contd…*

Small molecule	No. of carbons	Fsp^3	Melting point (°C)	log P exp.	C log P	mol log P	log S exp.	log S (C log P)	log S (mol log P)
Amikacin[a]	22	0.95	203 − 204	-7.40	-6.30	-5.09	6.11	5.01	3.55
Amiodarone	25	0.40	154 − 158	7.57	8.95	7.96	-8.38	-9.76	-8.77
Amitriptyline	20	0.30	196.5	4.92	4.85	4.47	-6.13	-6.06	-5.68
Amlodipine	20	0.40	178 − 179	3.00	3.43	2.98	-4.03	-4.46	-4.01
Amoxicillin	16	0.43	194	0.87	-2.19	0.03	-2.06	1.00	-1.22
Amphotericin B[a]	47	0.38	170	0.80	-4.45	4.51	-1.75	3.50	-5.46
Ampicillin	16	0.43	208	1.35	-1.53	0.36	-2.68	0.20	-1.69
Artesunate[b(i)]	19	0.84	131 − 135	NA	2.93	3.06	NA	-3.51	-3.64
Ascorbic acid[a]	6	0.50	191	-1.85	-2.57	-0.87	0.69	1.41	-0.29
Atenolol	14	0.50	147	0.16	-0.11	0.56	-0.88	-0.61	-1.28
Atorvarstatin	33	0.27	159.2−160.7	5.70	4.46	5.82	-6.54	-5.30	-6.66
Azathioprine	9	0.11	243.5	0.10	0.51	0.72	-1.78	-2.19	-2.41
Azithromycin	38	0.97	114	4.02	2.64	3.94	-4.41	-3.02	-4.33
Benzyl benzoate[b(i),c]	14	0.07	21	NA	3.94	3.53	NA	-3.44	-3.03
Beclomethasone dipropionate	28	0.46	117 − 120	1.30	2.51	4.19	-1.73	-2.94	-4.62
Betamethasone	22	0.59	232	1.94	0.13	2.18	-3.51	-1.70	-3.75
Betamethasone dipropionate[b(ii)]	28	0.53	NA	NA	3.92	3.87	NA	NA	NA
Bisacodyl[b(i)]	22	0.13	138	NA	2.85	3.45	NA	-3.48	-4.08
Bleomycin[b(i)]	55	0.32	71	NA	-7.68	-7.76	NA	7.72	7.80
Busulphan	6	1	287	-0.52	-0.59	-0.39	-1.60	-1.53	-1.73
Carbamazepine	15	0	190.2	2.45	2.38	2.37	-3.60	-3.53	-3.52
Carbimazole[a]	7	0.42	123.5	0.40	1.68	0.58	-0.88	-2.16	-1.06
Carboplatin[b(i)]	6	0.50	228 − 230	1.06	0.03	NA	-2.60	-1.57	NA
Cefixime	16	0.25	218 − 225	-0.40	-0.42	-0.10	-1.06	-1.04	-1.36
Cefotaxime	16	0.31	162 − 163	-0.50	-0.53	-0.13	-0.38	-0.35	-0.75
Ceftriaxone	18	0.33	156	-1.70	-0.65	-1.03	0.90	0.15	0.23
Cephalexin	16	0.31	326.8	0.65	-2.51	-0.42	-3.16	0.01	-2.09
Cetrizine	21	0.38	112.5	2.80	2.08	3.30	-3.17	-2.45	-3.67
Chlorambucil	14	0.50	65	1.70	3.63	3.34	-1.60	-3.53	-3.24
Chloramphenicol	11	0.36	150.5	1.14	1.28	0.66	-1.18	-2.03	-1.41
Cisplatin[b(i)]	0	0	270	-2.19	-1.68	NA	0.24	-0.27	NA
Clindamycin	18	0.94	142	2.16	2.57	1.86	-2.83	-3.24	-2.53
Clofazimine	27	0.11	210 − 212	7.66	7.70	8.44	-9.02	-9.06	-9.80
Chlorhexidine	22	0.27	134	0.08	4.81	3.72	-0.67	-5.40	-4.31
Clopidogrel[a]	16	0.25	158	2.50	4.21	3.85	-3.33	-5.04	-4.68
Clotrimazole[a]	22	0.05	148	6.10	5.00	4.89	-6.83	-5.73	-5.62
Cloxacillin	19	0.36	170	2.48	2.20	2.62	-3.43	-3.15	-3.57
Colchicine	22	0.36	156	1.30	1.20	1.06	-2.12	-2.02	-1.88
Cylophosphamide	7	1	41 − 45	0.80	0.80	-0.90	-0.48	-0.48	1.22
Cyclosporine A[b(i)]	62	0.67	148 − 151	NA	14.36	2.17	NA	-15.11	-2.91
Cytosine arabinoside	9	1	212.5	-2.80	-2.20	-2.13	1.42	0.83	0.76
Dacarbazine	6	0.33	205	-0.24	0.48	0.06	-1.53	-2.25	-1.83
Danazol	22	0.63	225.6	0.51	0.60	4.02	-2.01	-2.10	-5.52
Dapsone	12	0	175.5	0.97	0.89	0.80	-2.02	-1.94	-1.85
Daunorubicin	27	0.41	208 − 209	1.83	0.84	0.69	-3.16	-2.17	-2.02
Dexamethasone	22	0.72	262	3.60	0.13	2.18	-3.70	-2.00	-4.05
Dextromethorphan	18	0.73	111	2.89	3.95	3.64	-3.25	-4.31	-4.00

Table 2 *Contd…*

Small molecule	No. of carbons	Fsp^3	Melting point (°C)	log P exp.	C log P	mol log P	log S exp.	log S (C log P)	log S (mol log P)
Diazepam	16	0.13	132	2.82	2.96	2.82	-3.39	-3.53	-3.39
Diclofenac	14	0.07	283 − 285	4.51	4.73	4.15	-5.83	-5.55	-4.97
Didanosine[a]	10	0.50	160 − 163	-1.24	-1.65	-0.33	0.38	0.79	-0.53
Digoxin	41	0.31	249	1.26	-0.70	1.32	-3.00	-1.04	-3.06
Dihydroergotamine	33	0.34	239	2.00	2.71	2.53	-3.64	-4.35	-4.17
Diloxanide furoate[b(iii)]	14	0.14	NA	2.24	3.09	2.50	NA	NA	NA
Diltiazem	22	0.36	231	2.80	3.65	2.65	-4.36	-5.21	-4.21
Dimercaprol[c]	3	1	<25	0.58	0.18	-0.37	-0.08	0.32	0.87
Dithranol	16	0.06	176 − 181	2.81	3.29	2.79	-3.84	-4.32	-3.82
Dobutamine	18	0.34	184 − 186	3.60	2.43	2.50	-4.70	-3.53	-3.60
Domperidone	22	0.35	242.5	3.90	4.27	3.11	-5.57	-5.94	-4.78
Doxorubicin	27	0.45	229 − 231	1.27	0.32	0.15	-2.82	-1.87	-1.70
Doxycycline	22	0.36	201	-0.02	0.45	-1.21	-1.24	-1.71	-0.05
Efavirenz	14	0.28	139 − 141	4.60	3.73	3.42	-5.25	-4.38	-4.07
Esmolol	16	0.56	49	1.70	1.72	1.86	-1.44	-1.46	-1.60
Ethamutol	10	1	88	-0.30	0.12	0.17	0.17	-0.25	-0.30
Ethinylestradiol	20	0.55	183	3.67	0.83	3.67	-4.75	-1.91	-4.75
Ether[b(i),c]	4	1	-116.3	NA	0.87	7.72	NA	-0.37	-7.22
Etoposide	29	0.58	236 − 251	0.60	0.03	0.64	-2.28	-1.71	-2.32
Famotidine	18	0.17	163.5	-0.64	-0.58	-0.42	-0.24	-0.31	-0.46
Fentanyl	22	0.41	87.5	4.05	3.62	4.34	-4.17	-3.74	-4.46
Fluconazole	13	0.23	138 − 140	0.40	-0.44	-0.69	-1.04	-0.20	0.05
Flumazenil	15	0.27	201 − 203	1.00	1.29	1.97	-2.27	-2.56	-3.24
Fluorescein	20	0	315	3.40	2.66	5.50	-5.80	-5.06	-7.90
Flutamide	11	0.27	111 − 113	3.35	3.33	2.85	-3.71	-3.69	-3.21
Folic acid	19	0.21	250	-2.50	-2.31	-2.61	0.75	0.56	0.86
Folinic acid	20	0.20	245	-3.20	-3.49	-3.09	1.50	1.79	1.39
Formldehyde[c]	1	0	-92	0.35	0.35	-0.12	NA	NA	NA
Furosemide	12	0.08	295	2.03	1.90	2.12	-4.23	-4.10	-4.32
Gentamicin	21	0.95	105	-3.10	-1.80	0.24	2.80	0.70	-0.54
Glibenclamide	23	0.39	169	4.70	4.24	4.10	-5.64	-5.18	-5.04
Glucose[a]	6	0.84	83	-3.24	-2.21	-2.67	3.16	2.13	2.59
Glutaraldehyde[b(i),c]	5	0.60	-14	NA	-0.17	-0.37	NA	0.67	0.87
Glycerin[b(i),c]	3	1	17.8	NA	-1.54	-2.00	NA	2.04	2.50
Glyceryl trinitrate[c]	3	1	13.5	1.62	-5.76	-1.92	-1.12	6.26	2.42
Griseofulvin	17	0.41	220	2.18	2.05	3.33	-3.63	-3.50	-4.78
Haloperidol	21	0.38	151.5	4.30	3.85	3.93	-5.03	-4.58	-4.66
Homatropine	16	0.56	191 − 192	3.42	-2.90	1.21	-4.56	1.73	-2.37
Hydrochlorothiazide	7	0.14	274	-0.07	-0.36	-0.11	-1.92	-1.63	-1.88
Hydrogen peroxide[b(i),c]	0	NA	-0.43	NA	-1.50	-0.79	NA	2.00	NA
Ibuprofen	13	0.46	76	3.97	3.68	3.38	-3.98	-3.69	-3.39
Ifosfamide	7	1	39 − 41	0.86	0.92	0.77	-0.51	-0.57	-0.42
Imatinib	29	0.24	226	3.00	4.53	4.57	-4.51	-6.04	-6.08
Imipramine	19	0.36	174.5	4.80	5.04	4.07	-5.79	-6.03	-5.06
Indinavir	36	0.46	167.5 − 168	2.90	3.68	1.90	-3.82	-4.6	-2.82
Iopanoic acid[b(i)]	11	0.36	155.2 − 157	NA	4.70	2.45	NA	-5.51	-3.26
Isoflurane	3	1	48.5	2.06	1.76	2.41	-1.79	-1.49	-2.14
Isoniazid	6	0	171.4	-0.70	-0.67	-0.95	-0.26	-0.29	-0.01
Isosorbide mononitrate	6	1	88 − 91	-0.15	-4.59	-1.85	0.01	4.44	1.71

Table 2 *Contd…*

Small molecule	No. of carbons	Fsp^3	Melting point (°C)	log P exp.	C log P	mol log P	log S exp.	log S (C log P)	log S (mol log P)
Isosorbide dinitrate	6	1	70	1.31	-5.54	-1.67	-1.26	5.59	1.72
Lamivudine[a]	8	0.50	160 − 162	-1.40	-1.46	-1.24	0.54	0.60	0.38
Leflunomide	12	0.16	165 − 166	2.80	2.32	2.55	-3.70	-3.22	-3.45
Levothyroxine	15	0.13	235.5	4.00	3.51	3.64	-5.60	-5.11	-5.24
Lignocaine	14	0.50	68.5	2.44	1.95	1.78	-2.37	-1.90	-1.71
Lorazepam	15	0.07	167	2.39	2.37	1.88	-3.31	-3.29	-2.80
Mannitol	6	1	168	-3.10	-2.05	-2.68	2.10	1.12	1.75
Medroxy progesterone acetate[b(ii)]	24	0.75	NA	NA	4.20	4.09	NA	NA	NA
Mefloquine	17	0.35	175	3.90	3.67	3.58	-4.90	-4.67	-4.58
Melphalan[a]	13	0.46	182.5	-0.52	-0.21	0.58	-0.55	-0.86	-1.65
Mercaptopurine	5	0	313	0.01	-1.33	-0.22	-0.23	-1.05	-2.16
Metformin	4	0.50	223 − 226	-0.50	-1.63	-1.30	-0.99	0.13	-0.20
Methotrexate	20	0.25	195	-1.85	-0.53	-2.20	0.65	-0.67	1.00
Methyl ergometrine[a]	20	0.30	172	1.20	1.02	1.66	-2.17	-1.99	-2.63
Methyldopa	10	0.30	300	-1.70	-2.26	-0.34	-0.55	0.01	-1.91
Methylprednisolone	22	0.73	232.5	1.50	-0.82	1.75	-3.07	-0.75	-3.32
Metoclopramide	14	0.50	147.25	2.62	2.23	2.38	-3.34	-2.95	-3.10
Metaprolol	15	0.60	120	1.88	1.49	1.94	-2.33	-1.94	-2.30
Metronidazole	6	0.50	160.5	-0.02	-0.46	-0.65	-0.83	-0.39	-0.20
Miconazole	18	0.17	159 − 163	6.10	5.81	5.56	-6.96	-6.75	-6.42
Midazolam[a]	18	0.12	159	3.89	3.42	3.53	-4.73	-4.26	-4.37
Mifepristone	29	0.45	191 − 196	4.50	2.60	4.02	-5.25	-3.35	-4.77
Misoprostol	22	0.82	261 − 263	3.60	3.07	4.07	-5.47	-4.94	-5.94
Mitomycin C[d]	15	0.54	>360	-0.40	-1.35	-2.03	-2.45	-1.50	-0.82
N-acetylcysteine[b(i)]	5	0.60	109.5	NA	-0.62	-1.60	NA	0.27	1.25
Naloxone	19	0.53	200 − 205	2.09	0.16	2.84	-3.36	-1.43	-4.11
Nelfinavir	32	0.31	349.84	6.00	3.77	5.71	-8.74	-6.51	-8.45
Nevirapine[a]	15	0.27	196.06	2.50	2.65	1.39	-3.17	-3.86	-2.60
Nicotinamide	6	0	130	-0.37	-0.21	-0.43	-0.18	-0.34	-0.12
Nifedipine[a]	17	0.29	172 − 174	2.20	3.21	2.37	-3.18	-4.10	-3.35
Nitrofurantoin	8	0	263	-0.47	-0.47	-0.63	-1.49	-1.49	-1.33
Norethisterone	20	0.80	203.5	2.97	2.97	2.47	-4.25	-4.25	-3.75
Nystatin	47	0.38	160	0.50	-1.95	3.17	-1.35	1.10	-4.02
Ofloxacin	18	0.45	254	-0.39	-0.51	1.21	-1.40	-1.28	-3.00
Olanzapine	17	0.35	195	2.00	3.01	2.43	-3.20	-4.21	-3.63
Omeprazole	17	0.24	156	2.23	2.57	2.63	-3.04	-3.38	-3.44
Ondansetron	18	0.34	231.5	2.40	2.72	2.14	-3.96	-4.28	-3.71
Oxaliplatin[b(i)]	8	0.75	198.5−199.7	0.04	0.35	NA	-1.28	-1.59	NA
Paclitaxel	47	0.34	213 − 216	3.00	4.83	3.38	-4.39	-6.22	-4.77
Pantoprazole	16	0.25	139.5	0.50	2.11	2.44	-1.14	-2.75	-3.08
Paracetamol	8	0.13	170	0.46	0.49	1.14	-1.41	-1.44	-2.09
Penicillamine	5	0.80	198.5	-1.78	-1.73	-1.14	0.49	0.44	-0.15
Permethrin	21	0.29	34	6.50	7.61	6.21	-6.09	-7.20	-5.80
Phenobarbitone	12	0.25	174	1.47	1.37	1.04	-2.46	-2.36	-2.03
Phenylephrine	9	0.34	140 − 145	-0.31	-0.09	0.27	-0.70	0.92	-1.28
Phytomenadione[c]	31	0.61	-20	9.30	12.01	11.04	-8.35	-11.06	-10.09
Pilocarpine	11	0.55	204 − 205	1.10	-0.20	0.54	-2.30	-1.09	-1.83
Piperazine	4	1	106	-1.50	-1.48	-1.14	1.19	1.17	0.83

Table 2 *Contd…*

Small molecule	No. of carbons	Fsp^3	Melting point (°C)	log P exp.	C log P	mol log P	log S exp.	log S (C log P)	log S (mol log P)
Praziquantel	19	0.57	136	2.50	3.36	2.83	-3.11	-3.97	-3.44
Prednisolone	21	0.57	235	1.62	-1.14	1.69	-3.27	-0.52	-3.34
Prednisolone acetate[b(ii)]	23	0.61	NA	NA	1.96	2.27	NA	NA	NA
Primaquine	15	0.40	199 − 205	2.10	2.60	2.31	-3.37	-3.87	3.58
Procarbazine	12	0.42	223	0.06	-0.08	0.87	-1.54	-1.40	-2.35
Promethazine	17	0.24	60	4.81	4.40	4.34	-4.66	-4.25	-4.19
Propofol[c]	12	0.25	18	3.79	3.93	3.59	NA	NA	NA
Propyliodone	10	0.40	186.5	0.41	1.85	1.68	-3.29	-2.96	-2.79
Pyrazinamide	5	0	192	-0.60	-0.68	-1.19	-0.57	-0.49	0.02
Pyridoxine	8	0.38	159 − 162	-0.77	-0.35	-0.58	-0.09	-0.50	-0.28
Pyrimethamine	12	0.17	233.5	2.69	3.00	3.08	-4.27	-4.58	-4.66
Raloxifene	28	0.25	143 − 147	5.20	6.86	6.56	-5.90	-7.56	-7.26
Ranitidine	13	0.54	69 − 70	0.27	0.67	0.56	-0.21	-0.61	-0.51
Riboflavin	17	0.41	280	-1.46	-0.73	-1.75	-0.59	-1.32	-0.30
Rifampicin	43	0.30	183	2.70	3.02	3.19	-3.80	-4.12	-4.29
Ritonavir	37	0.29	120 − 122	3.90	4.94	5.41	-4.36	-5.40	-5.87
Salicylic acid	7	0	158	2.26	2.19	1.18	-3.09	-3.02	-2.01
Saquinavir	38	0.39	349.84	3.80	4.73	2.96	-6.54	-7.47	-5.71
Sevoflurane[c]	4	0.75	<25	2.40	1.45	1.93	-1.90	-0.95	-1.43
Spironolactone	24	0.63	134.5	2.78	2.84	3.54	-3.37	-3.43	-4.13
Stavudine	10	0.30	159 − 160	-0.72	-0.49	-0.73	-0.12	-0.35	-0.12
Sulfasalazine	18	0	220	2.50	3.88	3.89	-3.92	-5.33	-5.34
Sulphadiazine	10	0	255.5	-0.09	0.10	0.80	-1.71	-1.91	-2.61
Testosterone	19	0.84	155	3.32	0.08	3.69	-4.12	-0.88	-4.49
Tramadol	16	0.56	180 − 181	2.40	3.10	2.99	-3.45	-4.15	-4.04
Tropicamide[a]	17	0.29	96.5	1.30	1.18	1.64	-1.51	-1.39	-1.85
Verapamil[c]	27	0.89	<25	3.79	4.47	5.28	-3.29	-3.97	-4.78
Vincristine	46	0.43	220	2.82	4.04	2.33	-4.27	-5.49	-3.78
Vitamin A (Retinol)	20	0.50	63.5	5.68	6.40	5.98	-5.56	-6.28	-5.86
Vitamin D (Ergocalciferol)	28	0.71	116.5	7.30	6.59	8.17	-7.71	-7.01	-8.58
Zidovudine	10	0.60	113 − 115	0.05	0.04	-0.52	-0.44	-0.43	0.13

NA: not available

[a] the fifteen compounds selected for log S correlation analysis depicted in **Fig. 2** as well as for studying the relationship between Fsp^3 and log S that is illustrated in **Fig. 3**

[b(i)] one or more of the log P values not available; even if one such entry is found in any of the three log P columns, this particular compound was excluded from the log P correlation analysis shown in **Fig. 1**

[b(ii)] at least one log P value as well as melting point data not available leading to the exclusion of these compounds from both the log P correlation analysis and the subsequent aqueous solubility determination

[b(iii)] only melting point data for such compounds not available enabling them to be included in the log P correlation analysis but excluding them from being used in the GSE; diloxanide furoate is the only member in this category

[c] these compounds are liquids at room temperature and have been left out from the error-based indirect approach whose results are summarized in **Table 1**

[d] a melting point of 360 °C was plugged into the GSE for this case

Conclusion

Our overall objective of value addition to NLEM 2011 from a medicinal chemist's point of view has been accomplished successfully in this work with the compilation of crucial physicochemical properties including partition coefficient and aqueous solubility of the small molecule subset within the list. In the absence of reliable experimental values of these properties in the literature, the chemist community has to depend upon a robust computational tool out of the multitude of them available presently. To address this issue, we have generated a simple and workable protocol for assessing the

most optimal software package for such predictions. By consistently projecting mol log P as the better choice for predicting partition coefficient, this proof-of-concept study has clearly demonstrated its utility in enabling the medicinal chemist to make an informed choice with regard to picking a desirable computational tool for such purposes.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgement

The authors would like to thank Dr. Sreeraj Gopi, Principal Scientist, Plant Lipids Ltd., Kochi for providing uninterrupted access to Chemical Abstracts and for useful technical discussions that aided the drafting of this manuscript.

References

[1] Huuskonen J, Salo M, Taskinen J (1998). Aqueous solubility prediction of drugs based on molecular topology and neural network modelling. *J. Chem. Inf. Comput. Sci.* **38(3):** 450-456.

[2] Abraham M.H, Le J (1999). The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship. *J. Pharm. Sci.* **88(9):** 868-880.

[3] Mitchell B.E, Jurs P.C (1998). Prediction of aqueous solubility of organic compounds from molecular structure. *J. Chem. Inf. Comput. Sci.* **38(3):** 489-496.

[4] Chin Y.P, Weber W.J, Voice T.C (1986). Determination of partition coefficients and aqueous solubilities by reverse phase chromatography--ii. Evaluation of partitioning and solubility models. *Wat. Res.* **20(11):** 1443-1450.

[5] Lipinski C.A, Lombardo F, Dominy B.W, Feeney P.J (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46(13):** 3-27.

[6] Jain N, Yalkowsky S.H (2001). Estimation of the aqueous solubility I. Application to organic nonelectrolytes. *J. Pharm. Sci.* **90(2):** 234-252.

[7] Iqbal H, Gopi S, Balasubramanian R (2014). National List of Essential Medicines 2011: A medicinal chemistry perspective. *Int. J. Pharm. Sci. Rev. Res.* **24(1):** 110-14.

[8] Kenny P.W, Montanari C.A, Prokopczyk I.M (2013). $ClogP_{alk}$: a method for predicting alkane/water partition coefficient. *J. Comput. Aided Mol. Des.* **7(5):** 389-394.

[9] Ran Y, Yalkowsky S.H (2001). Prediction of drug solubility by the general solubility equation (GSE). *J. Chem. Inf. Comput. Sci.* **41(2):** 354-357.

[10] Lovering F, Bikker J, Humblet C (2009). Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **52(21):** 6752-6756.

[11] Wassvik C.M, Draheim A.G.H, Artursson P, Bergström C.A.S (2008). Molecular characteristics for solid-state limited solubility. *J. Med. Chem.* **51(10):** 3035-3039.