ALOGPS 2.1 (http://www.vcclab.org) is a Free on-line Program to Predict logP and logS of Chemical Compounds

Igor V. Tetko(1,2), Vsevolod Yu. Tanchuk(2)

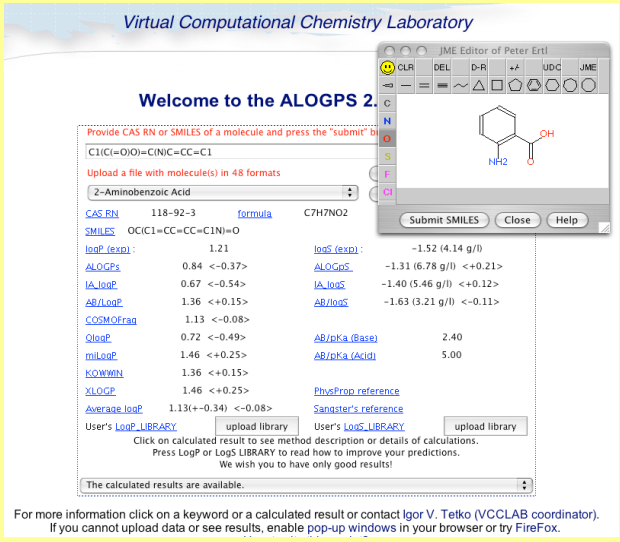
(1) - MIPS, GSF, Institute for Bioinformatics, Munich, Germany

(2) - IBPC, Ukrainian NAS, Kyiv, Ukraine, http://www.vcclab.org

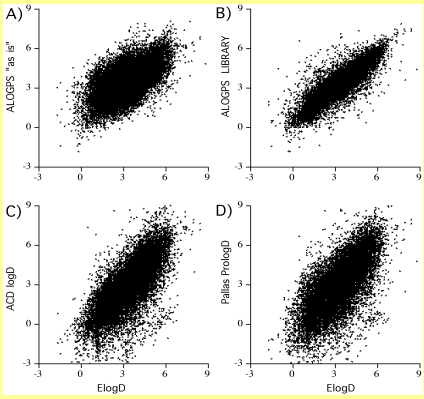
The ALOGPS 2.1 program was developed using Associative Neural Network method and combines models to predict lipophilicity and aqueous solubility of chemicals.

LogP (lipophilicity): training set of 12908 molecules from the PHYSPROP database. 75 E-state indicies. 64 neural networks were trained using 50% of molecules selected by chance from the whole set. The prediction accuracy for molecules not used in the training set is root mean squared error rms=0.35 and standard mean error s=0.26.

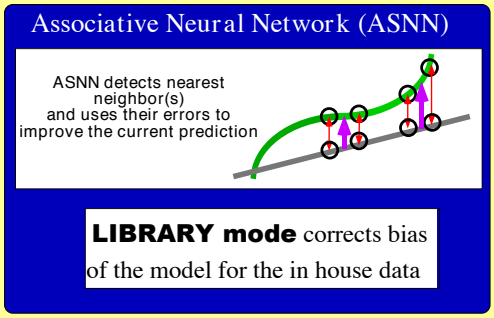
LogS (aqueous solubility): training set of 1291 molecules. Prediction results rms=0.49, s=0.38, i.e. significantly improved compared to ALOGPS 2.0



Poor prediction of new series?

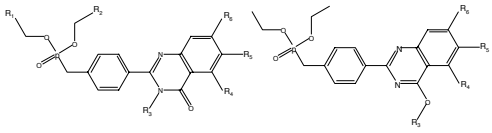


Solution (self learning)



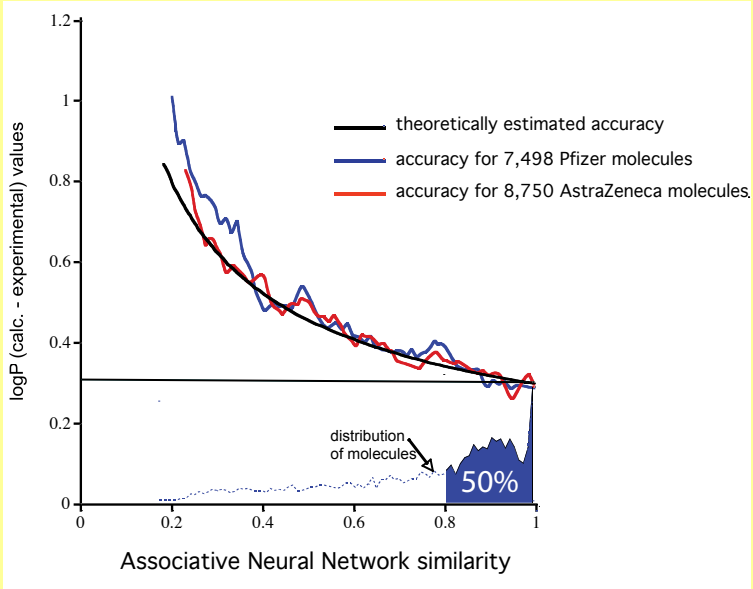
ALOGPS on-line provides comparison of 8 (soon 9) logP and 3 aqueous solubility models as well as pKa (courtesy of Pharma Algorithms) calculation. A consensus value is also calculated for logP models and everything is avaialble just **with one mouse click.**

Prediction of Pfizer compounds (Tetko & Poda, 2004) from A to B in less than <10 minutes!



(I)	ACD/logP		ALOGPS 2.1			(II)	ACD/logP		ALOGPS 2.1	
$logP_{exp}$	$logP_{calc}$	$logP_{calc}$ (after training)	$logP_{calc}$	$logP_{calc}$ (after training)		$logP_{exp}$	$logP_{calc}$	$logP_{calc}$ (after training)	$logP_{calc}$	$logP_{calc}$ (after training)
3.29	1.99	in training	2.91	in memory		5.26	3.01	in training	3.93	in memory
4.09	2.96	4.26	3.71	4		7.01	4.31	6.56	4.97	6.29
3.18	1.62	2.92	3.47	3.43		5.61	2.72	4.97	3.78	5.09
2.65	0.92	2.21	2.45	2.46		5.13	3.06	5.31	3.95	5.24
3.41	1.98	3.28	3.11	3.35		5.53	3.44	5.69	4.6	5.71
5.17	4.74	6.04	4.88	5.19		3.75	1.68	3.93	3.26	4.47
4.88	3.66	4.96	3.95	4.5		4.51	2.75	5.00	4.17	5.24
4.20	2.67	3.97	3.69	3.87		6.37	3.78	6.03	4.54	5.87
4.49	3.76	5.05	4.36	4.57		7.67	5.86	8.11	5.99	7.12
MAE	1.23	0.33	0.38	0.19		MAE	2.25	0.33	1.29	0.59

ALOGPS 2.1 program dramatically increased its prediction ability using just one molecule with known experimental values (Tetko & Tanchuk, 2002). The ALOGPS further increases its prediction ability (in 3-5 times) using larger number of molecules in LIBRARY. Notice, that NO NEW INDICES were created for neural network.



With just few measured molecules your can increase accuracy of predictions in 2-5 times!

>50% of Pfizer molecules predicted in the LIBRARY mode with < 0.4 MAE indeed had their error in this range (Tetko et al, 2006)

Literature

Tetko, I. V.; Bruneau, P.; Mewes, H. W.; Rohrer, D. C.; Poda, G. I. Can we estimate the accuracy of ADME-Tox predictions?, Drug Discov. Today, 2006, 11, 700-7.

Tetko, I. V.; Poda, G. I. Application of ALOGPS 2.1 to predict log D distribution coefficient for Pfizer proprietary compounds, J. Med. Chem., 2004, 47, 5601-4

Tetko, I. V.; Tanchuk, V. Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program, J. Chem. Inf. Comput. Sci., 2002, 42, 1136-45

pre-prints, poster and free analyses are available at http://www.vcclab.org