



$\text{Log}_{10}D^{\text{octanol/buffer}}$ Computation

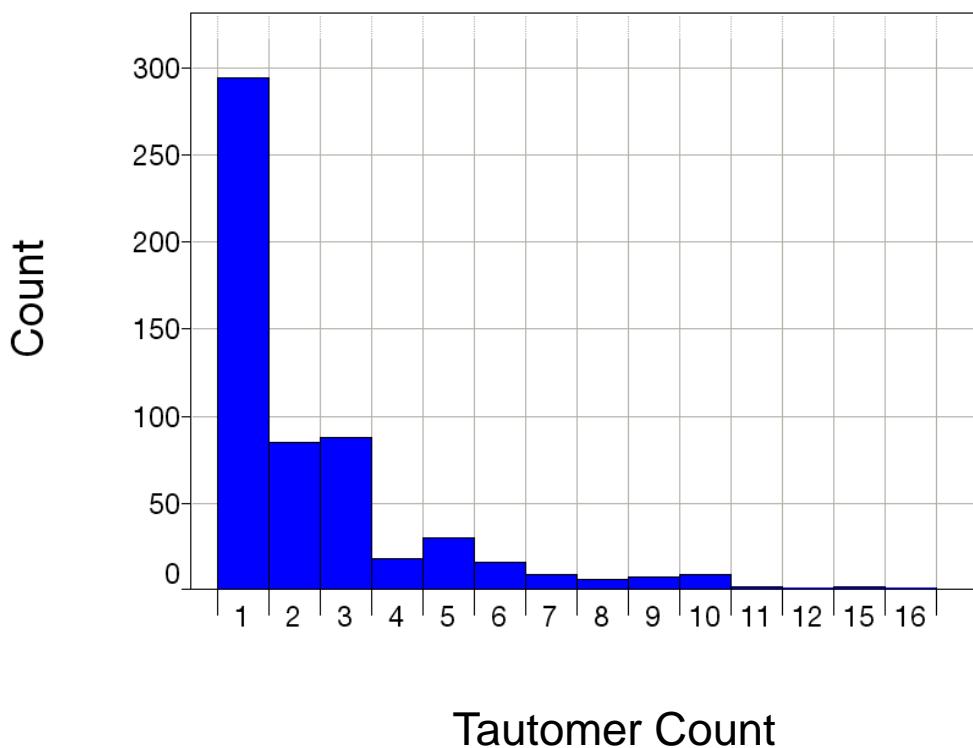
Which combination of computational predictors gives the best prediction of experimental results?

Yvonne Martin, William Porter, Jian Wu,
James Metz, Wentao Fu

Number of tautomers

Dataset:

Log *D*: 546 compounds, 286 have no tautomers



Dataset

286 Measured log D's, most at pH 7.4

Compounds have no tautomers using a combination of

- (1) The in-house Daylight toolkit program Tautomers that generates the common tautomers of heterocycles and also of beta-diketones and enols.
- (2) The resulting set of registered structures and tautomers in SMILES format was then processed with the tautomer generation component of Pipeline Pilot, version 7.5.2.300. For the latter, the following options were chosen: Carbon as donor, never; Make all sp² atoms acceptor, false; Amides tautomerization, only diamides; Perceive charge tautomers, false. Note that with these options the Pipeline Pilot component will not tautomerize beta-diketones to enols and vice versa.

Individual Regression Results: Experimental Value Predicted from Computed Value

Independent Variable	Model parameters	Estimate	Std. Error	<i>t</i>	<i>P</i> (2 Tail)	<i>R</i> ²	S.E.E.
Vendor 1, current	Intercept	1.327	0.133	9.998	< 0.001	0.603	1.260
	Slope	0.680	0.033	20.809	< 0.001		
Vendor 1, previous	Intercept	1.463	0.146	10.024	< 0.001	0.525	1.377
	Slope	0.671	0.038	17.778	< 0.001		
Vendor 2, previous	Intercept	0.216	0.171	1.259	0.209 *NS	0.633	1.221
	Slope	0.962	0.044	21.900	< 0.001		
Vendor 2, current	Intercept	0.241	0.168	1.436	0.152 *NS	0.632	1.212
	Slope	0.954	0.043	22.168	< 0.001		
Vendor 3	Intercept	0.697	0.185	3.774	< 0.001	0.520	1.384
	Slope	0.833	0.047	17.610	< 0.001		
LOGD_CHEMAXON	Intercept	1.050	0.137	7.645	< 0.001	0.793	1.220
	Slope	0.785	0.036	21.936	< 0.001		
Vendor 4	Intercept	1.627	0.186	8.723	< 0.001	0.349	1.612
	Slope	0.684	0.055	12.407	< 0.001		

Error from combined results

Regression on average of five predictors	1.116
Regression on median of five predictors	1.162
Regression on average of three best predictors	1.090
Regression on average of Vendor 2 and ChemAxon	1.090
Regression on Vendor 2	1.214
Regression on ChemAxon	1.220

Conclusions

When tautomeric forms are ignored:

- ChemAxon and Vendor 2 are the best overall single predictor for Abbott compounds.

Conclusions

The simplest approach is to use Vendor 2 software or ChemAxon software to predict logD, based on results of analyzing data for $n = 286$ non-tautomeric Abbott compounds.

- for model:
logD (predicted) = $1.010 \times \text{Vendor 2 computed}$
 - $R^2 = 0.914$ (cannot be compared with R^2 for models with 2 or more parameters)
 - **S.E.E. = 1.214 log units**
- for model:
logD (predicted) = $1.050 + 0.785 \times \text{ChemAxon LogD computed}$
 - $R^2 = 0.793$
 - **S.E.E. = 1.220 log units**

Slightly greater precision in prediction can be achieved by averaging the predictions generated by Vendor 1, Vendor 2, and ChemAxon, based on results for these same Abbott compounds.

- for model:
logD (predicted) = $0.577 + 0.897 \times \text{Average}$
 - $R^2 = 0.702$
 - **S.E.E. = 1.090 log units**

Recommendations

- Implement tautomer screening for all Abbott structures entries and post warning if multiple tautomeric forms are identified.
- Post disclaimer that estimated error of computed logD values is ± 1.2 log units.
 - Provide link to R4P3 for experimental measurement of log D

Further Analysis of the Agreement between Predictions

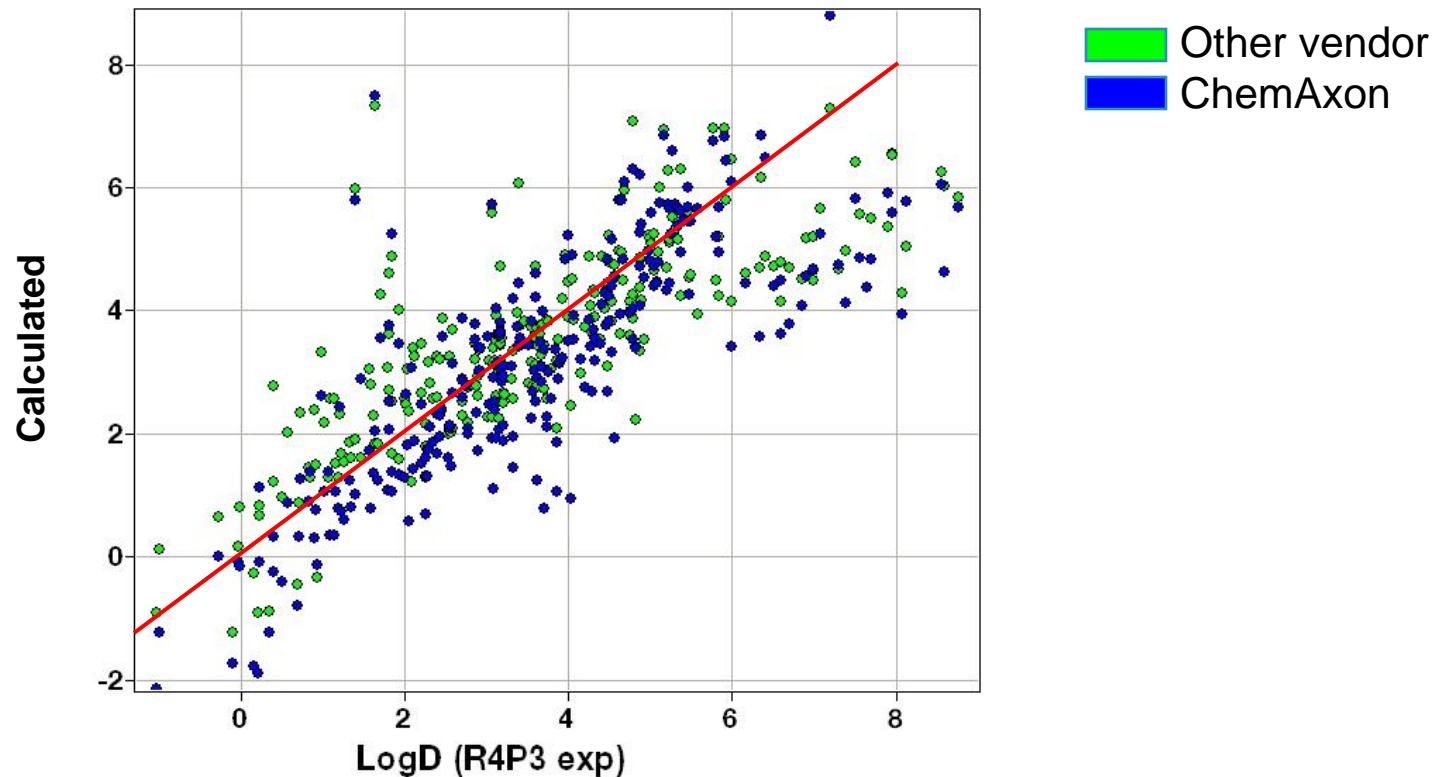
Although the two “best” programs have similar predictivity, how often do they both predict the same log D for a compound?

When they agree, is their prediction correct?

When they disagree, is there a pattern in which is closer to correct?

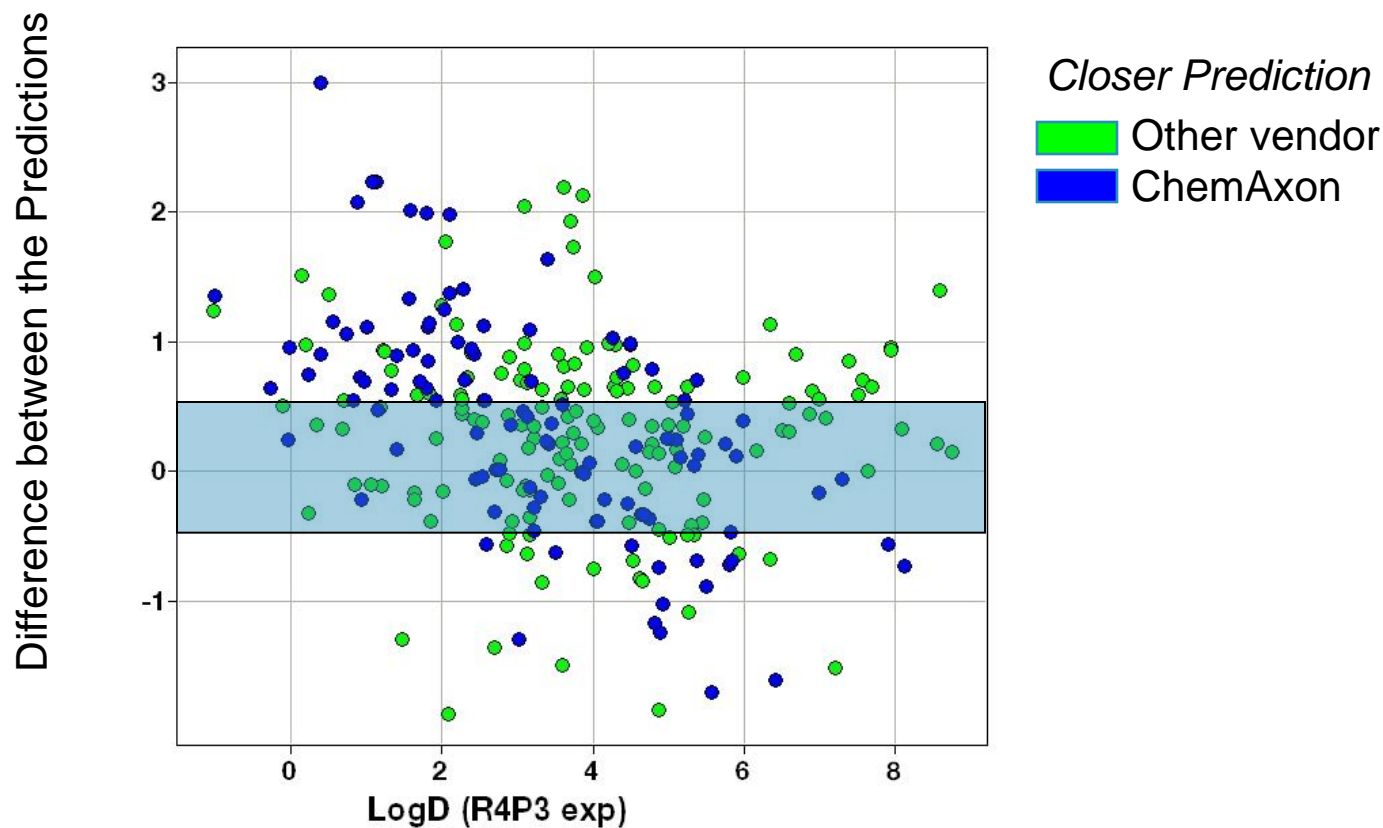
Observed log D of Compounds with no Tautomers vs. Predictions by the Two Programs

Comment: Neither program is especially accurate.



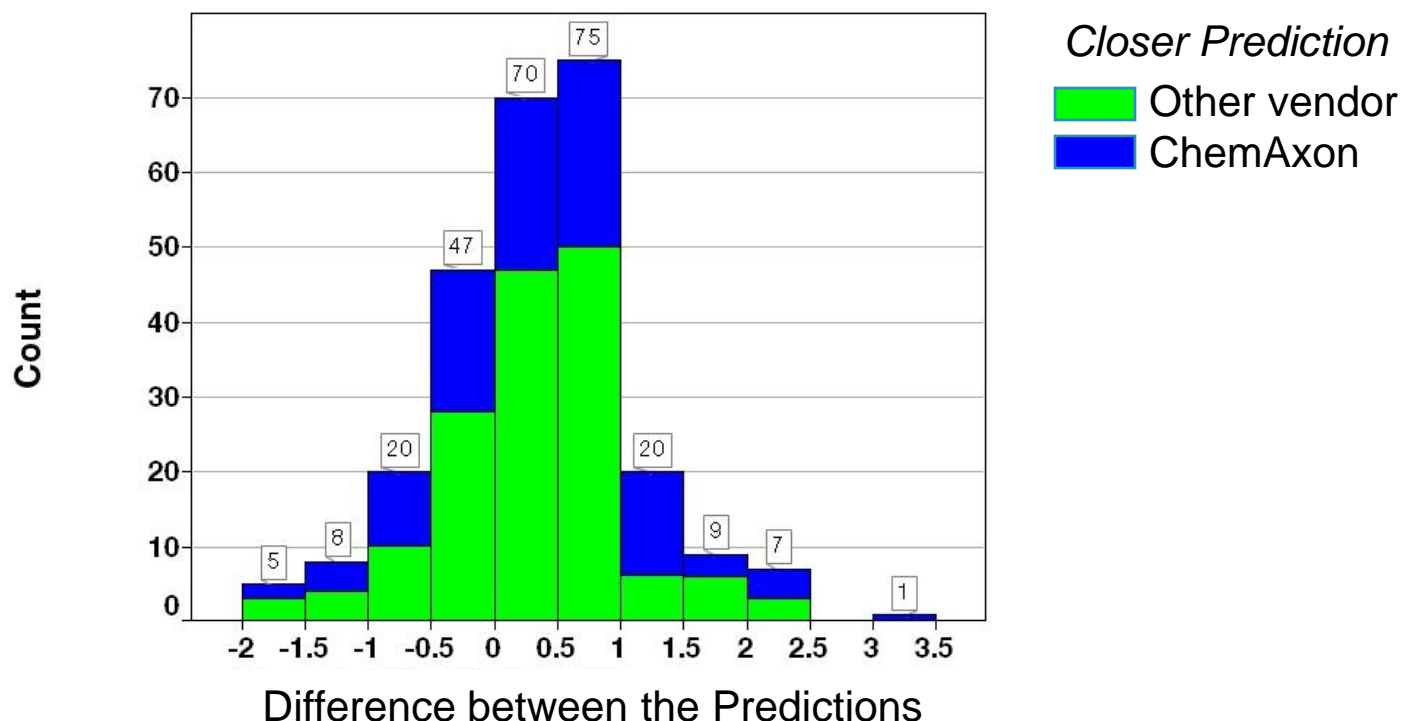
Observed log D of Compounds with no Tautomers vs. the Difference between the Predictions

Comment: Neither program is especially accurate.



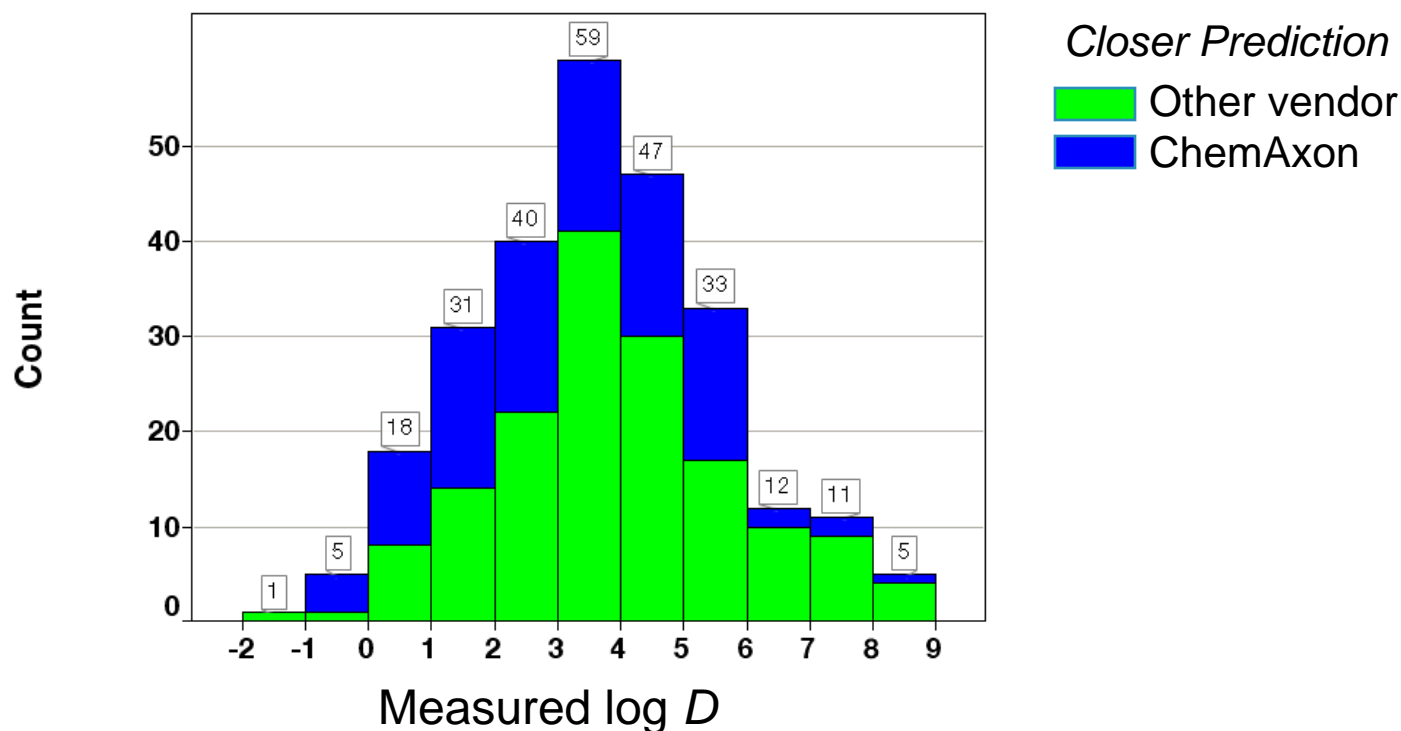
Frequency Distribution of Deviation of Predictions of $\log D$ of Compounds with no Tautomers

Comment: The two programs don't agree on their predictions.

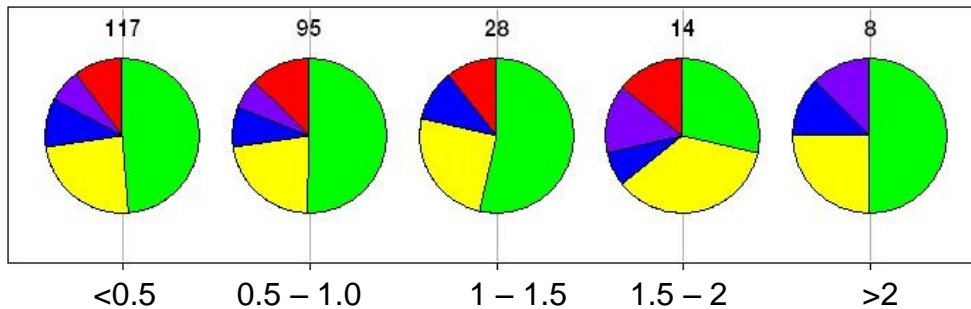


Frequency Distribution of Observed log D of Compounds with no Tautomers Showing Which Prediction is Closer

Comment: ChemAxon seems to do better at low observed log D and the other predictor at high log D .



There is no Relationship between the Agreement of Predictions and the Accuracy of the Predictions from the Average

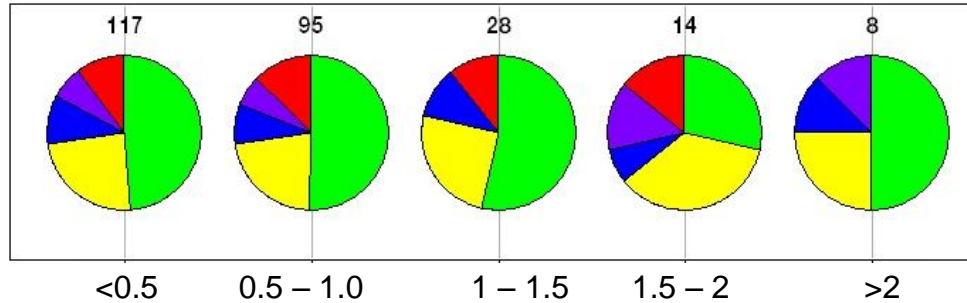


Absolute value of the difference between the predictions: *colored by the deviation of the observed log D from the average prediction.*

Coloring:

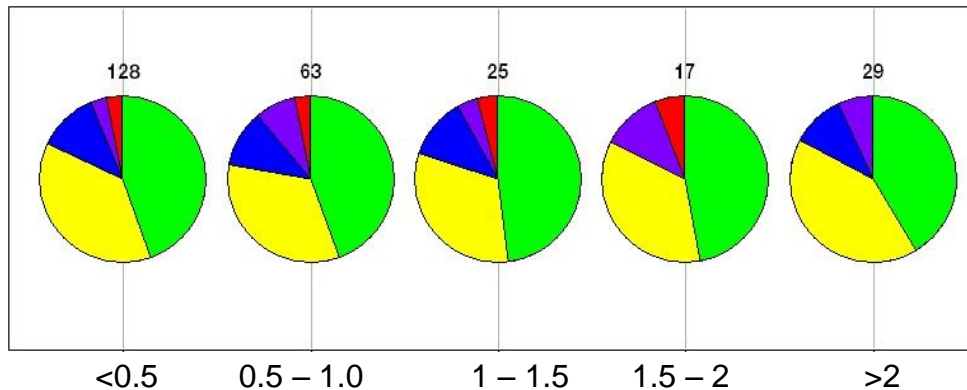
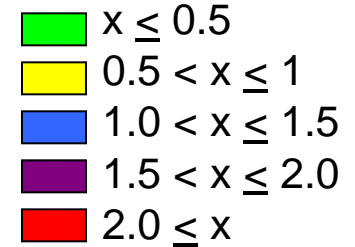


There is no Relationship between the Agreement of Predictions and the Accuracy of the Predictions from the Average



Absolute value of the difference between the predictions, colored by the deviation of the observed log D from the average prediction

Coloring:



Absolute value of the deviation of the observed log D and the average of the predictions, colored by the deviation of the predictions

Conclusion

As computational chemists, we still cannot predict $\log D$ with greater than 1 log accuracy.

- Failure of the modeling procedures?
- Inadequate chemical descriptors?
- Limitation in the variety of structures in the training sets?
- Problems with the data used to develop the models?

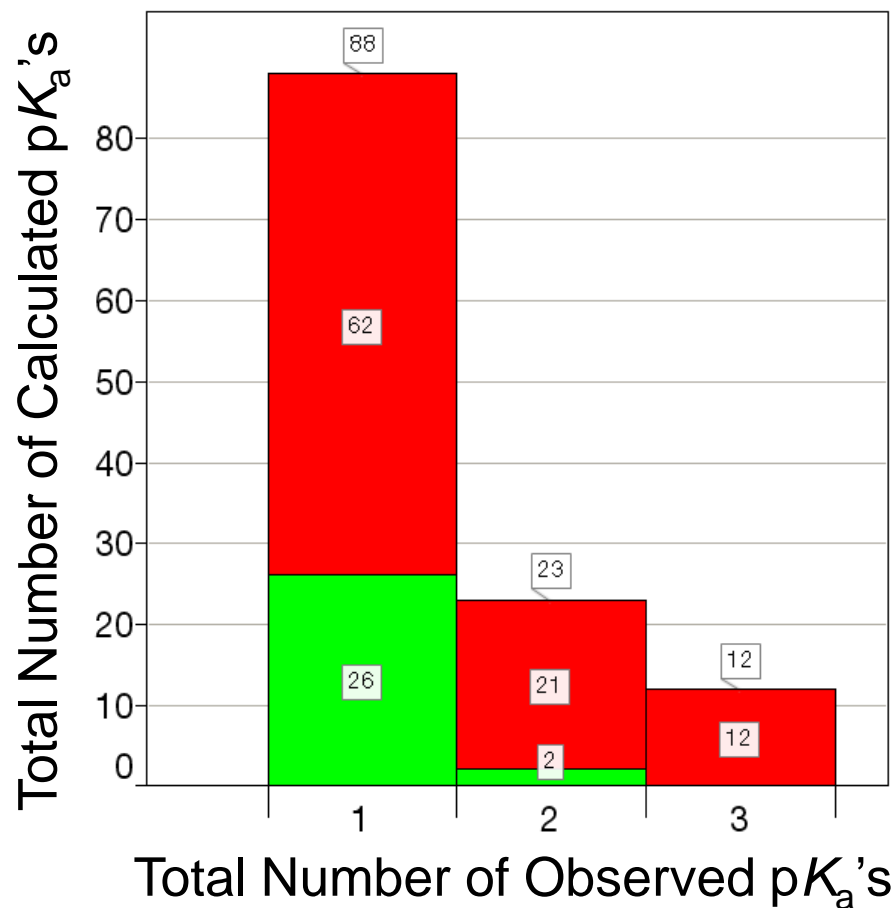
Conclusion

As computational chemists, we still cannot predict $\log D$ with greater than 1 log accuracy.

- Failure of the modeling procedures?
- Inadequate chemical descriptors?
- Limitation in the variety of structures in the training sets?
- Problems with the data used to develop the models?

If only we really understood solvation!

Why We Didn't Evaluate pKa Predictions



Only 28 of the 124 Compounds Have One Observed Tautomer and an Equal Number of Observed and Calculated pKa's