

SAMPL6 Part II Partition Coefficient Challenge Overview

Mehtap Isik

SAMPL6 Part II $\log P$ Challenge Virtual Workshop
May 16th , 2019

For more information
<https://github.com/MobleyLab/SAMPL6>

SAMPL blind challenges: Statistical Assessment of the Modeling of Proteins and Ligands

Physical properties

To test of force field accuracy and to isolate chemical effects.

Host-guest systems

Binding of small drug-like molecules without slow protein timescales.

Model protein-ligand systems

Isolate individual physical challenges (e.g. binding of charged ligands).

**SAMPL0
2007**

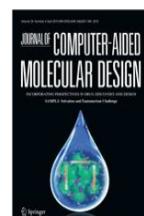
JNK3 kinase inhibitors
hydration free energies

**SAMPL1
2008**

CDK2 kinase inhibitors
hydration free energies

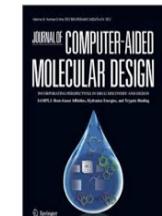
**SAMPL2
2009**

hydration free energies
tautomer ratios



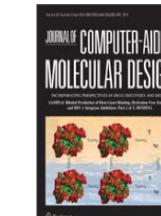
**SAMPL3
2011**

trypsin inhibitors
hydration free energies



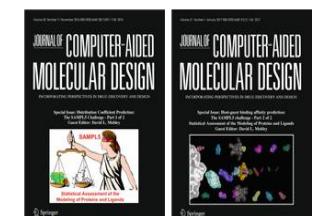
**SAMPL4
2013**

HIV-1 integrase inhibitors
hydration free energies
octoacid host-guest
CB7 host-guest



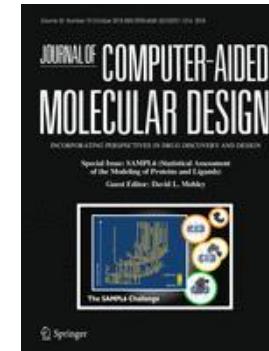
**SAMPL5
2016**

distribution coefficients
CBClip host-guest
CB7 host-guest



**SAMPL6
Part I - 2018**

pK_a
host-guest affinity
Sampling

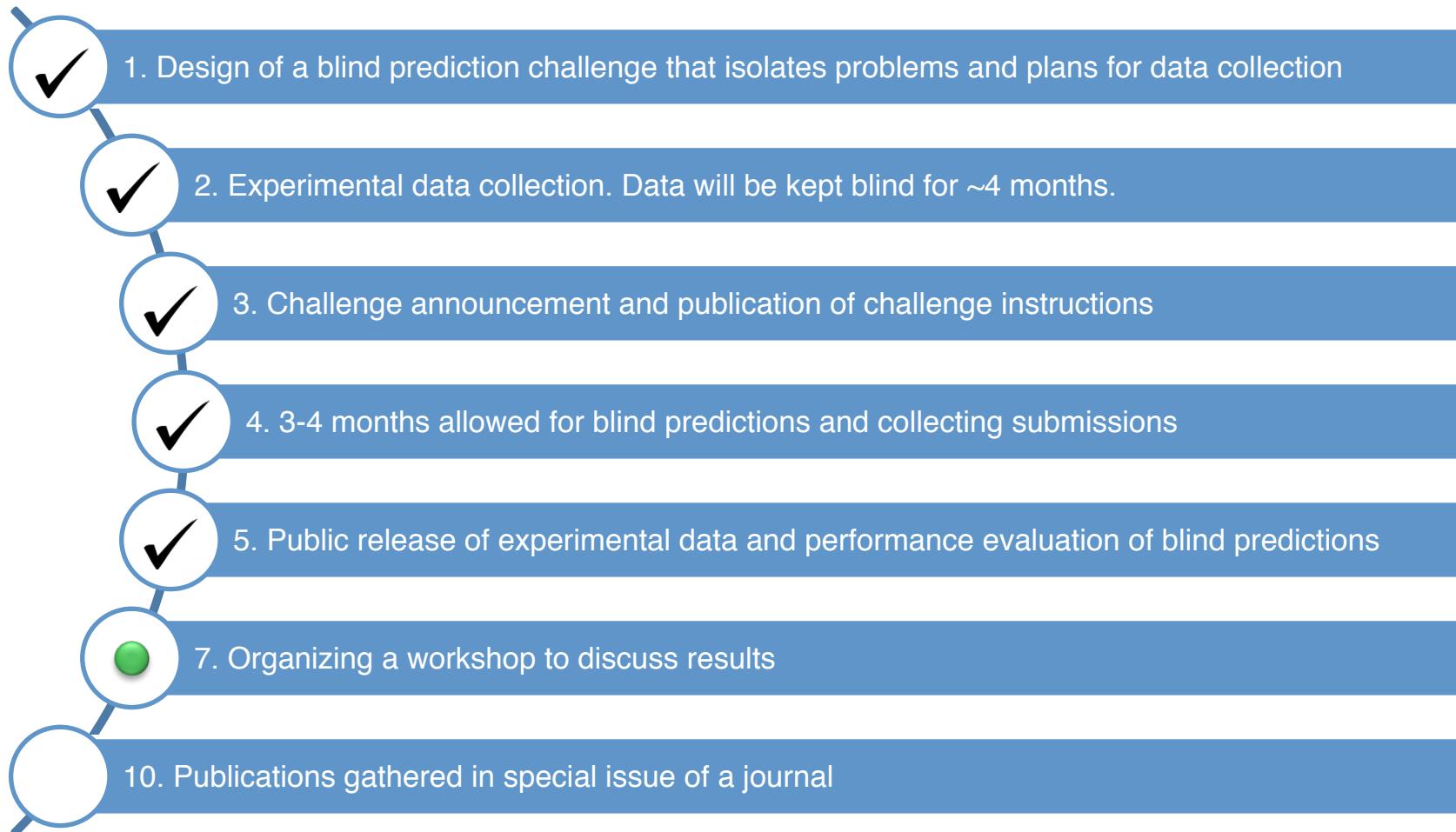


**SAMPL6
Part II - 2019**

partition coefficients



Roadmap for SAMPL6 logP Challenge



Experimental collaborators

Merck, MRL
Timothy Rhodes
Dorothy Levorse
Brad Sherborne

SAMPL6 organizers

David Mobley
John Chodera
Mehtap Isik
Daniel Bergazin

SAMPL6 logP challenge participants are encouraged to submit publications to JCAMD special issue.

In coordination with Terry Stouch, we are planning a special issue of *J. Comp. Aided Mol. Design* (JCAMD) focused on the SAMPL6 logP challenge.

All participants are welcome to submit manuscripts evaluating their methods.

The submission deadline for this is **September 15, 2019**.



- Overall performance analysis paper
- Paper that describes experimental data collection
- Individual method papers by participants

Why did we decide to organize a separate pK_a and $\log P$ prediction challenges in SAMPL6?

SAMPL5 logD challenge showed that the failure to account for ionization state (pK_a) effects impact prediction accuracy of $\log D$.

$$\log D = \log \frac{[X^0]_{\text{oct}} + [XH^+]_{\text{oct}}}{[X^0]_{\text{aq}} + [XH^+]_{\text{aq}}}$$

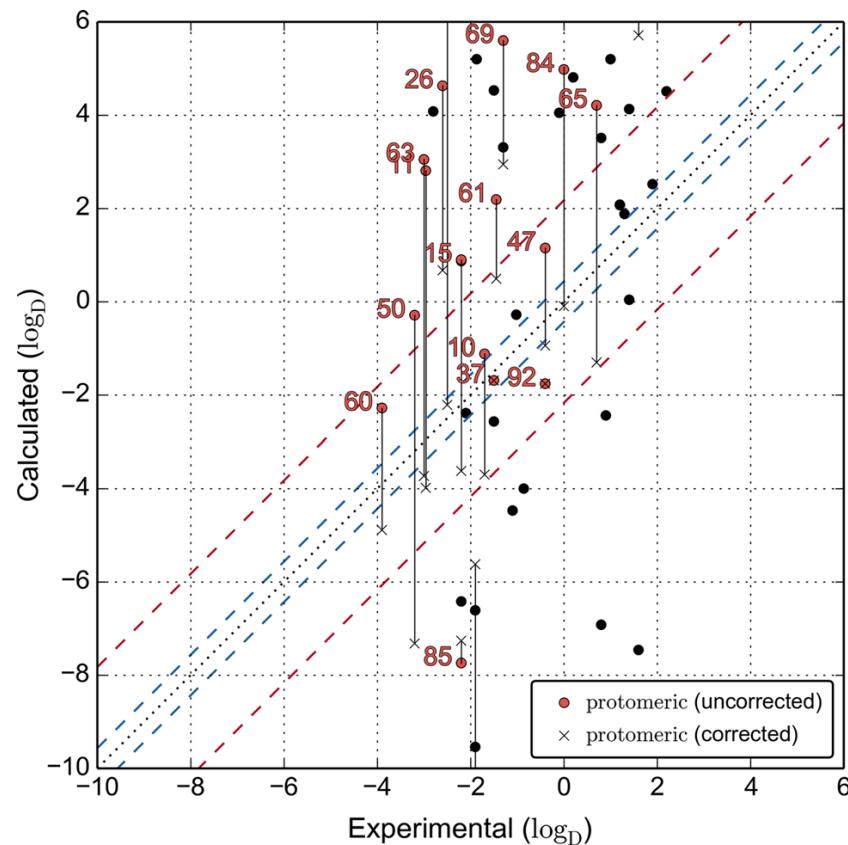
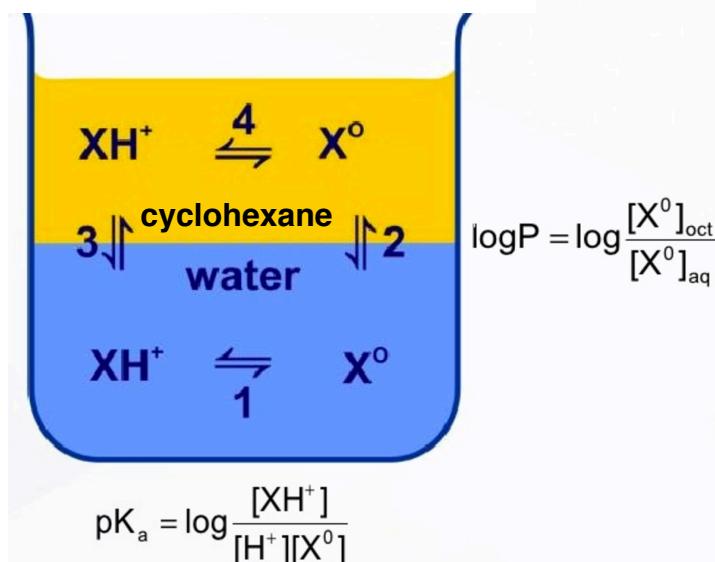
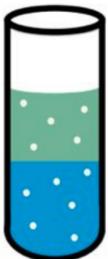


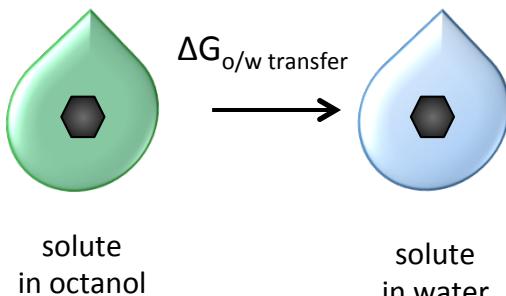
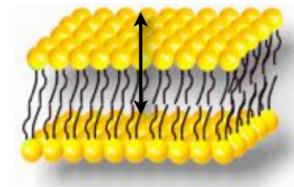
Fig. 4 Our partition estimations from MM BAR (submission 38) plotted against experiment. We have applied our QM based free energy corrections (adiabatic/absolute scheme, submission 10), shifting the predicted values towards more hydrophilic values. These corrections account for multiple protomeric states and for ligand ionization due to the presence of protonizable groups. These corrections substantially reduce the RMSD and increase the correlation of these predictions with respect to experimentally determined values

What is the motivation behind a blind octanol-water partition coefficient challenge?



$$\log P_{ow} = \log \frac{[A]_{oct}}{[A]_{wat}}$$

Log P values are a proxy for predicting behavior of small molecules in biologically interesting environments such as **water-lipid membrane interfaces**.



It provides a test system for evaluating force field errors in protein-ligand binding affinity prediction methods:

- Octanol phase is a **hydrophobic environment** similar to protein environment.
- Solvation free energies or transfer free energies** between two environments can be calculated to predict partitioning.
- $\log P$ prediction allows separating **force-field accuracy** from errors related to conformational sampling of proteins and protonation state predictions.

$$\Delta G_{transfer} = G_{oct} - G_{wat}$$

$$\Delta G_{transfer} = -RT \ln(10) \log P_{ow}$$

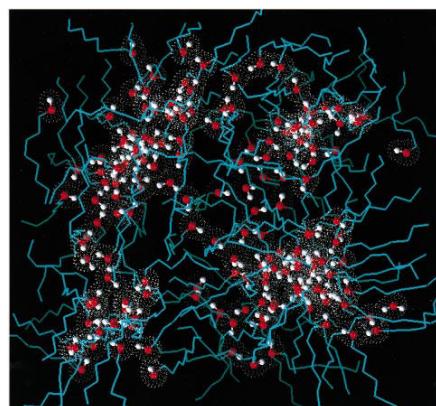
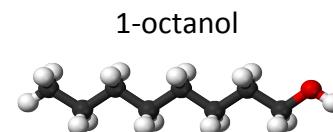
$$\log P_{ow} = -\frac{\Delta G_{transfer}}{RT \ln(10)} = -\frac{G_{oct} - G_{wat}}{RT \ln(10)}$$

Properties of octanol phase

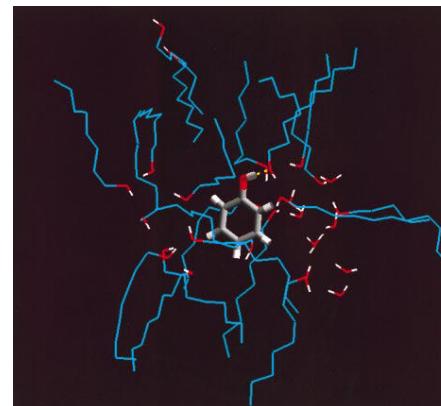
- Flexible molecule
- Octanol phase is wet

The mole fraction of water in octanol: 27.05% [2]

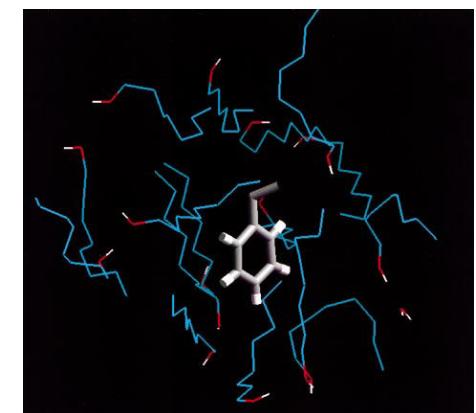
- Heterogeneous environment that has hydrophobic and hydrophilic regions



Polar and hydrophobic regions of water saturated octanol. [1]



Phenol in polar region



Ethylbenzene in nonpolar region

[1] Best, Scott A., Kenneth M. Merz, and Charles H. Reynolds. "Free Energy Perturbation Study of Octanol/Water Partition Coefficients: Comparison with Continuum GB/SA Calculations." *The Journal of Physical Chemistry B* 103, no. 4 (January 28, 1999): 714–26. <https://doi.org/10.1021/jp984215v>

[2] Lang, Brian E. "Solubility of Water in Octan-1-Ol from (275 to 369) K." *Journal of Chemical & Engineering Data* 57, no. 8 (August 9, 2012): 2221–26. <https://doi.org/10.1021/je3001427>.

Structure of SAMPL6 log *P* Challenge

- Predicted values to report
 - **log *P* value**
 - **log *P* SEM**: Captures statistical uncertainty of the predicted method.
 - **model uncertainty**: Predicted accuracy of a method, estimates of how well your predicted values are expected to agree with experimental values
- Participants were asked to categorize their methods
 - Physical
 - Empirical
 - Mixed
 - Other
- Detailed method description

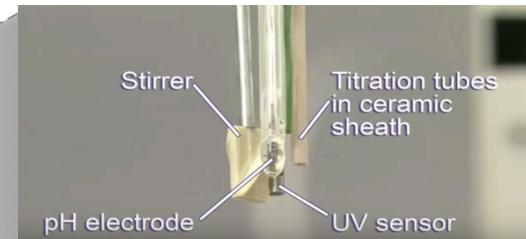
Challenge Timeline

- Nov 1, 2018 - SAMPL6 Part II Challenge start date
- Mar 22, 2019 - Challenge submissions due
- Mar 25, 2019 - Experimental data release date
- May 16, 2019 - SAMPL6 log *P* challenge virtual workshop
- Aug 22-23, 2019 - Joint D3R/SAMPL workshop, San Diego
- Sep 15, 2019 - JCAMD special issue submissions due

For more information on log*P* prediction challenge
<https://github.com/MobleyLab/SAMPL6>

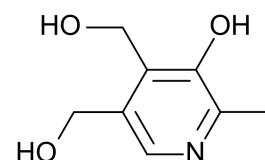
Potentiometric logP measurements were collected with Sirius T3

Collaborators from MRL
Dorothy Levorse
Timothy Rhodes

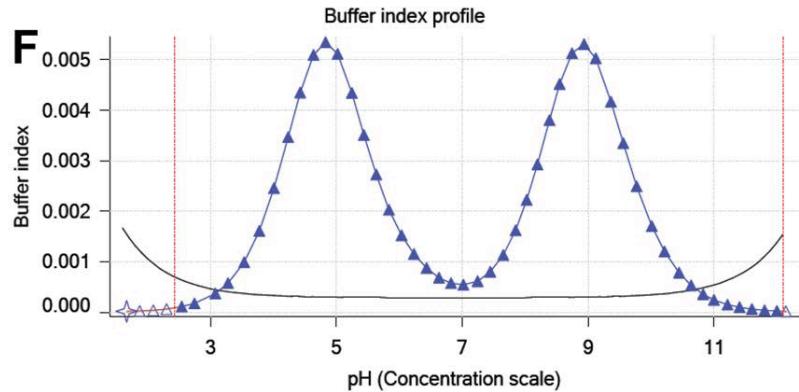
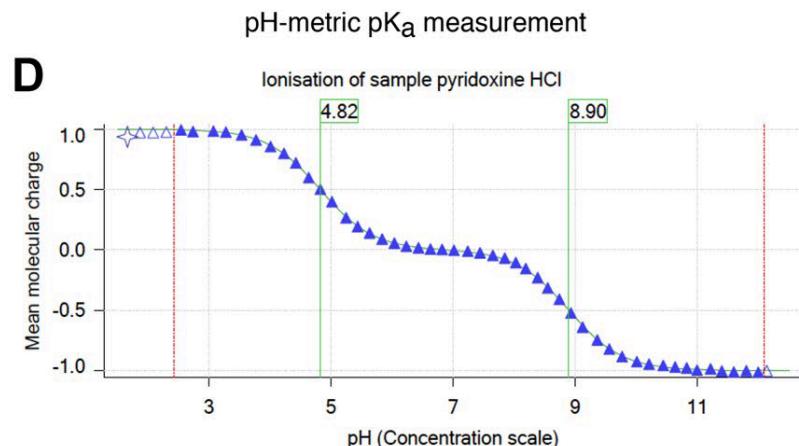


Potentiometric logP measurement relies on detecting apparent pK_a shift in octanol-water biphasic system.

- 3 independent replicates
- $25 \pm 0.5^\circ\text{C}$
- 1-3 mg analyte
- Ionic-strength adjusted water (0.15 M KCl)
- Octanol saturated with ionic-strength adjusted water
- Automated acid/base titrations



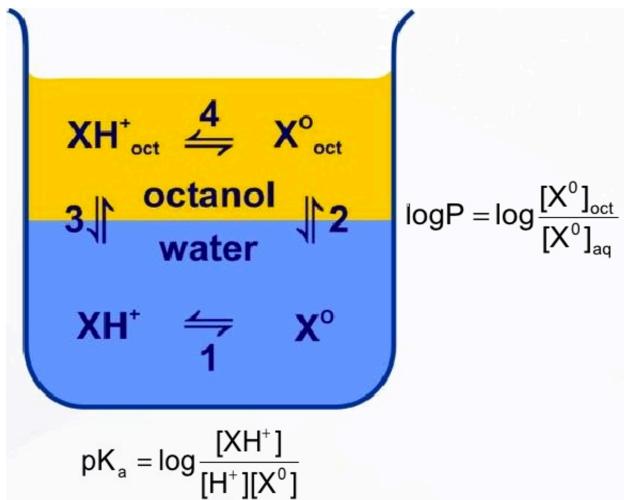
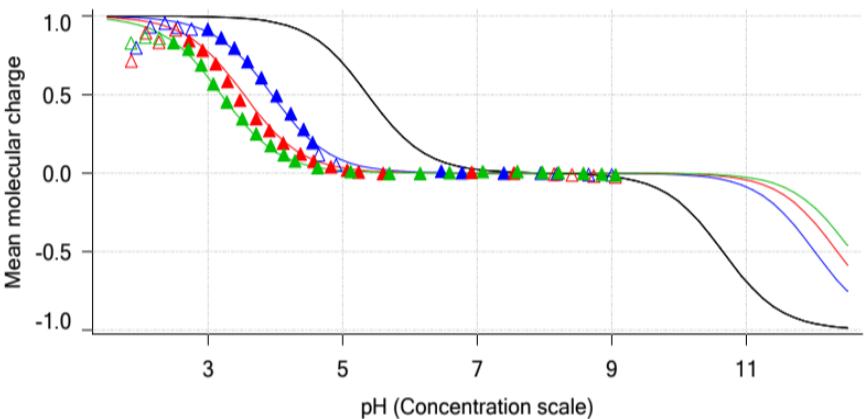
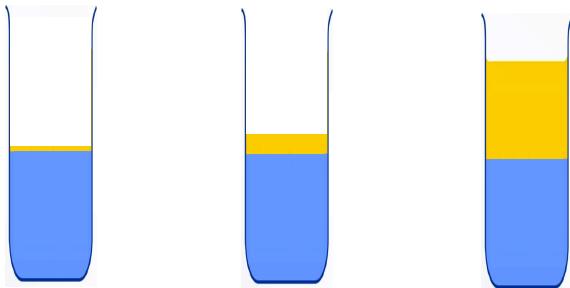
Pyridoxine HCl



Potentiometric logP measurement relies on detecting apparent pK_a shift with respect to relative octanol volume in a biphasic system.

Collaborators from MRL
Dorothy Levorse
Timothy Rhodes

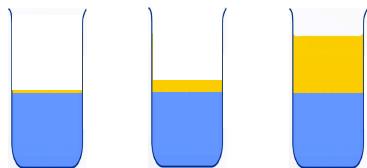
20 uL octanol 70 uL octanol 1070 uL octanol



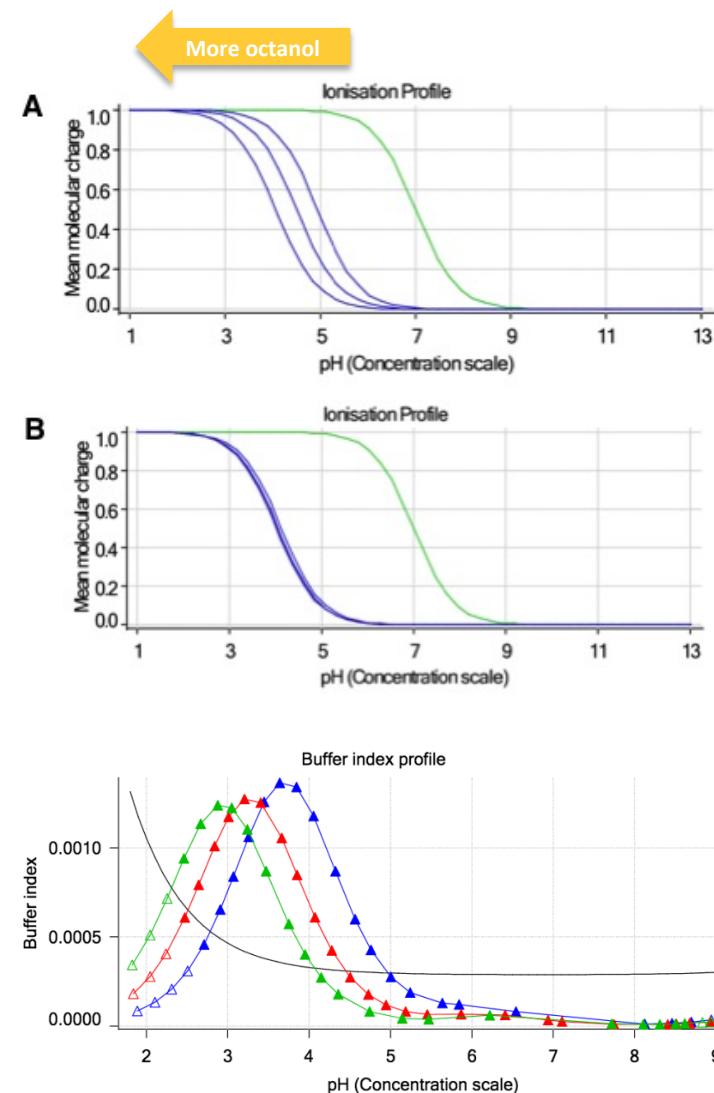
- Sample**
- M06 concentration factor
 - Sample pK_as
 - logP (XH²⁺)
 - logP (neutral XH)
 - logP (X⁻)
 - Carbonate
 - Acidity error

Limitations and challenges of potentiometric logP measurements

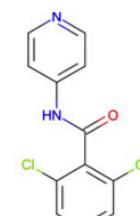
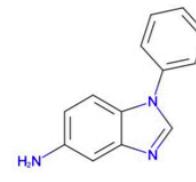
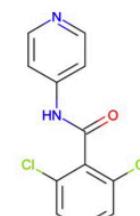
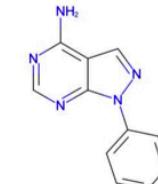
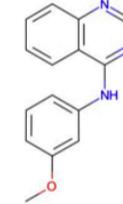
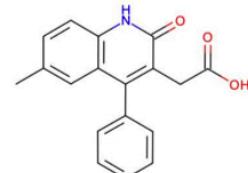
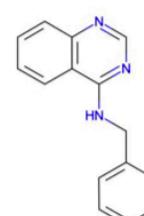
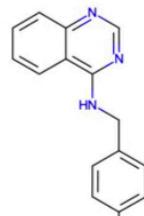
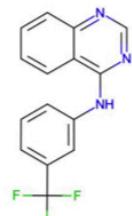
- How to optimize the correct octanol-water volume ratio?
 - Significant pK_a shift desired
 - Limited flask volume



- Measureable pK_a range and pK_a shift
 - At least one titratable group
 - Must adjust protocol for significant apparent pK_a shift
 - Can only measure pK_a s between 2-12
 - Molecules with low basic pKas and high acidic pKas are not suitable for this measurement
- How to optimize analyte concentration?
 - High enough buffering capacity required to measure pK_a
 - Limited by thermodynamic and kinetic solubility
 - pH titration requires visiting pH's where analyte has low solubility.
- Limited option of organic solvent
 - Cyclohexane is too volatile for this measurement method
 - Dodecane logP values of SAMPL6 compounds were beyond measurable range.



Octanol:water logP values for 11 compounds measured with potentiometric method of Sirius T3.



SAMPL6 logP challenge molecules are a subset of SAMPL6 pKa challenge molecules which were dictated by the experimental limitations of potentiometric logP measurement method.

- 4-amino quinazoline group: 6 molecules
- Benzimidazole group: 2 molecules

SAMPL6 logP challenge experimental data has been released after challenge deadline.

Molecule ID	logP mean	logP SEM	Assay Type
SM02	4.09	0.03	pH-metric octanol logP
SM04	3.98	0.03	pH-metric octanol logP
SM07	3.21	0.04	pH-metric octanol logP
SM08	3.10	0.03	pH-metric octanol logP (predose 80 uL octanol)
SM09	3.03	0.07	pH-metric octanol logP
SM11	2.10	0.04	pH-metric octanol logP
SM12	3.83	0.03	pH-metric octanol logP
SM13	2.92	0.04	pH-metric octanol logP (predose 100 uL octanol)
SM14	1.95	0.03	pH-metric octanol logP
SM15	3.07	0.03	pH-metric octanol logP
SM16	2.62	0.01	pH-metric octanol logP (predose 100 uL octanol)

- 3-4 independent replicates were performed for each molecule.
- Narrow dynamic range of logP values:

1.95 - 4.09

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$
$$SEM = \frac{\sigma}{\sqrt{N}} ; \sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$$

We received a significant number of blind predictions covering a variety of methods.

We collected **91 submissions** from **27 groups**.

Reporting Predictions

Molecule ID, $\log P$, $\log P$ SEM, $\log P$ model uncertainty

Method category

- Empirical: 17
- Physical: 48
- Mixed: 17
- Other: 9

How accurate were octanol-water $\log P$ predictions?

Statistics calculated for evaluating accuracy:

- Root mean square error(RMSE)
- Mean absolute error (MAE)
- Mean error (ME)
- Linear regression slope (m)
- R-squared (R^2)

95% confidence intervals of each statistic were calculated by bootstrapping.

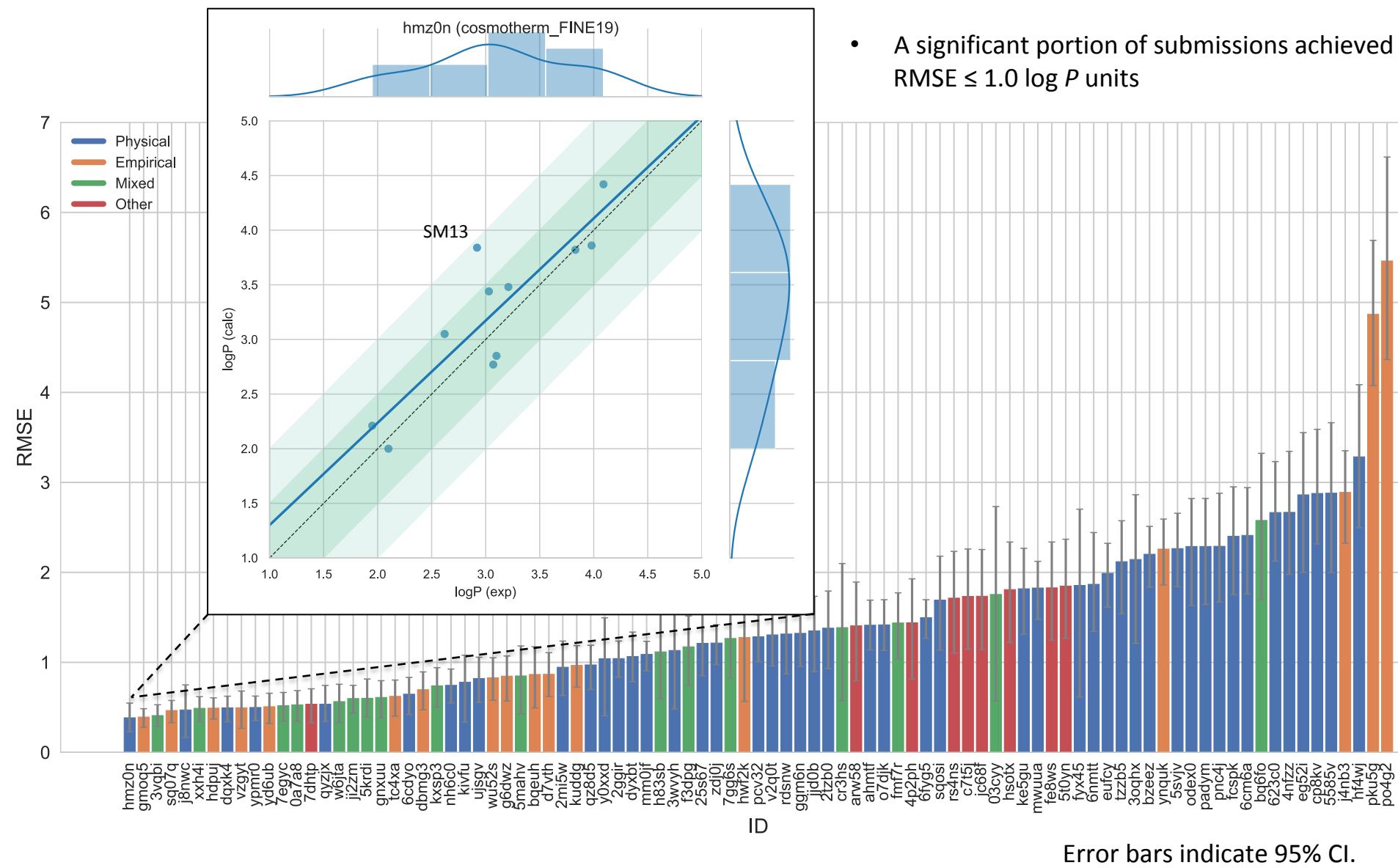
Github SAMPL6 repository:

[**SAMPL6 / physical_properties / logP / analysis / analysis_outputs / StatisticsTables / statistics.pdf**](#)

Because of limited dynamic range RMSE and MAE more important statistics for evaluating model accuracy than correlation-based metrics.

Accuracy evaluation of methods that participated SAMPL6 logP challenge

- 10 methods achieved RMSE $\leq 0.5 \log P$ units
- A significant portion of submissions achieved RMSE $\leq 1.0 \log P$ units

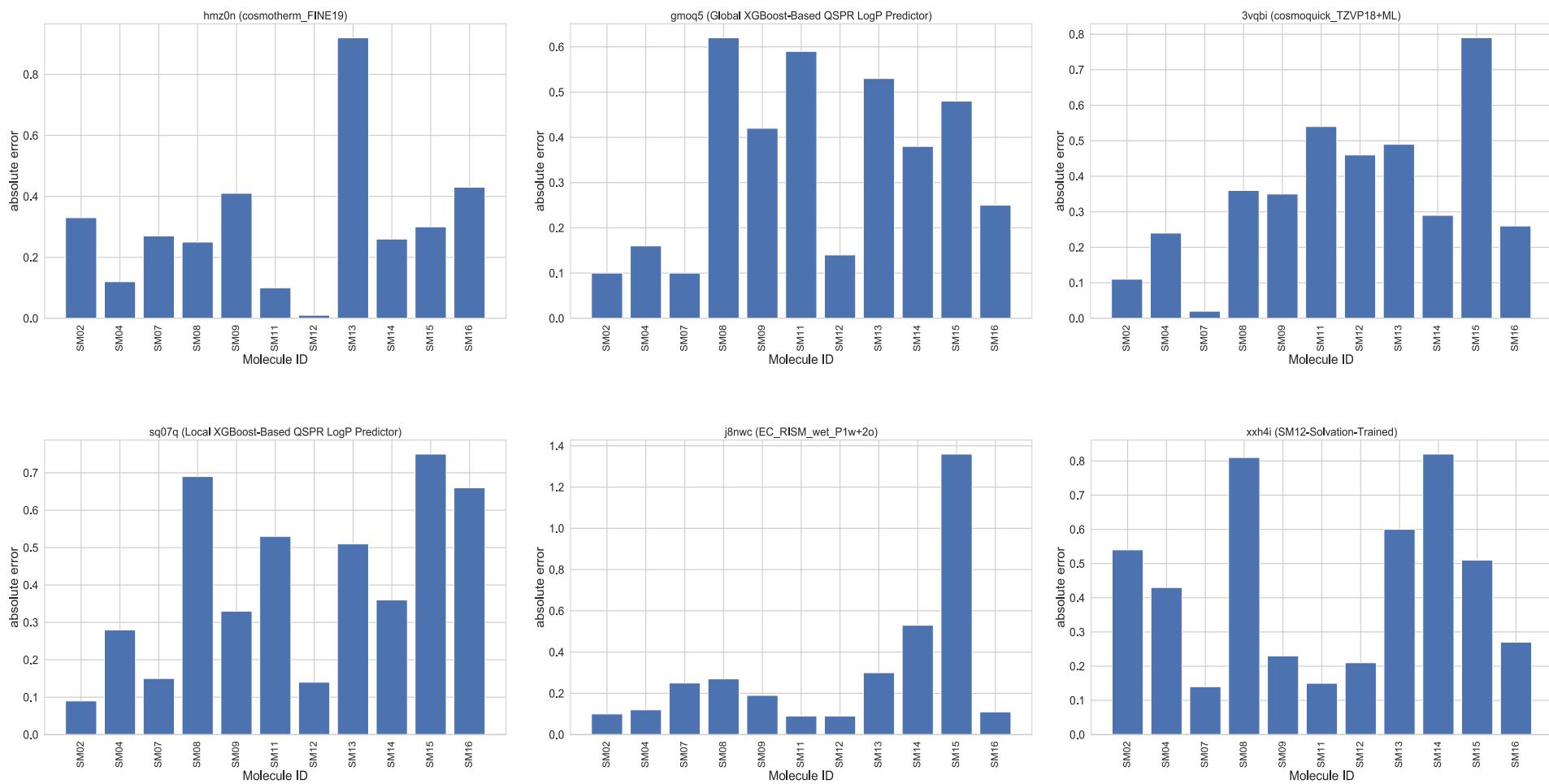


Is there a reason to suspect experimental error for SM13 logP value?

Absolute errors plots for calculated for each molecule for methods with lowest RMSE don't suggest a possibility of experimental error with SM13.

With additional NMR and MS measurements we confirmed that SM13 has the correct structure.

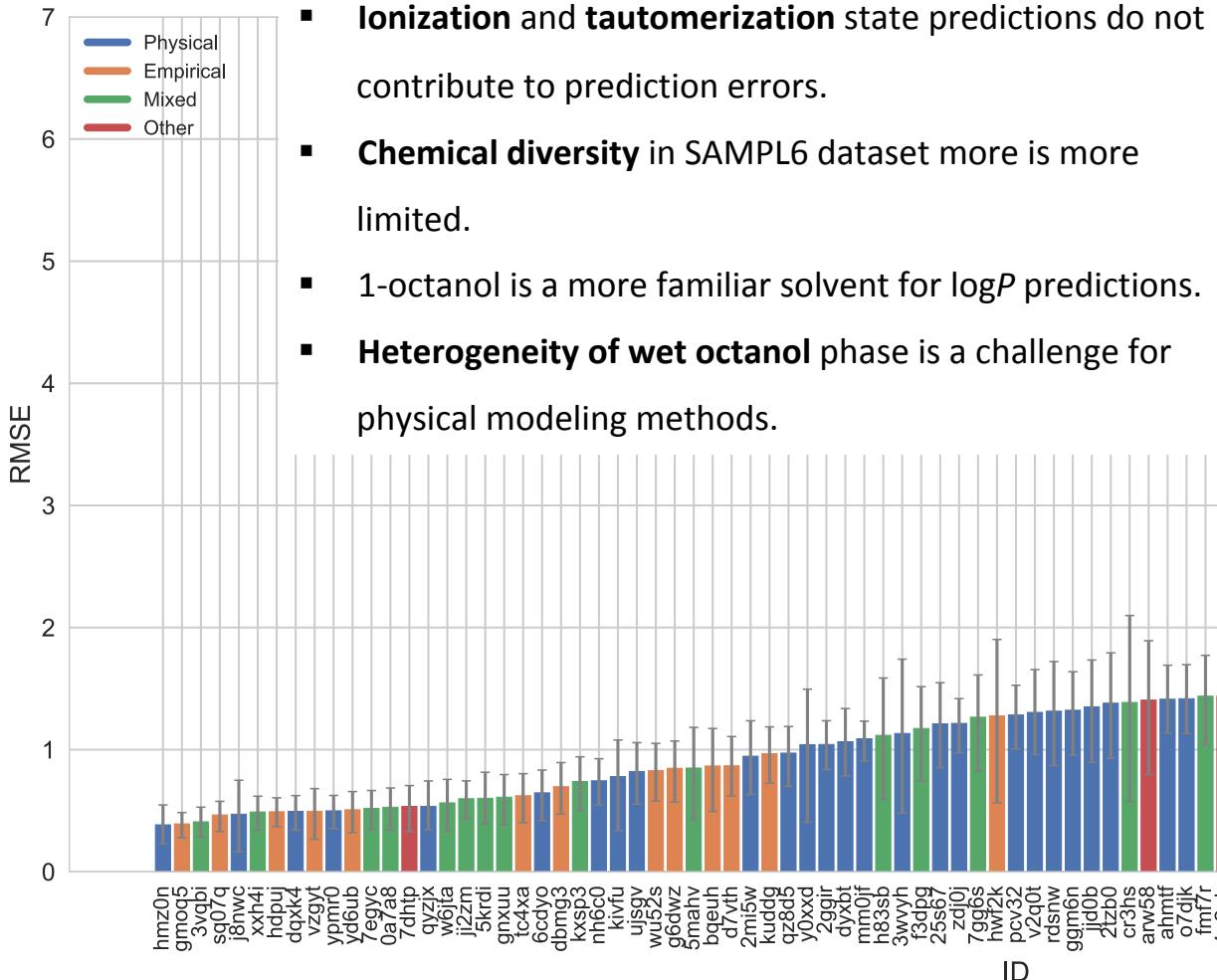
Dorothy Levorse



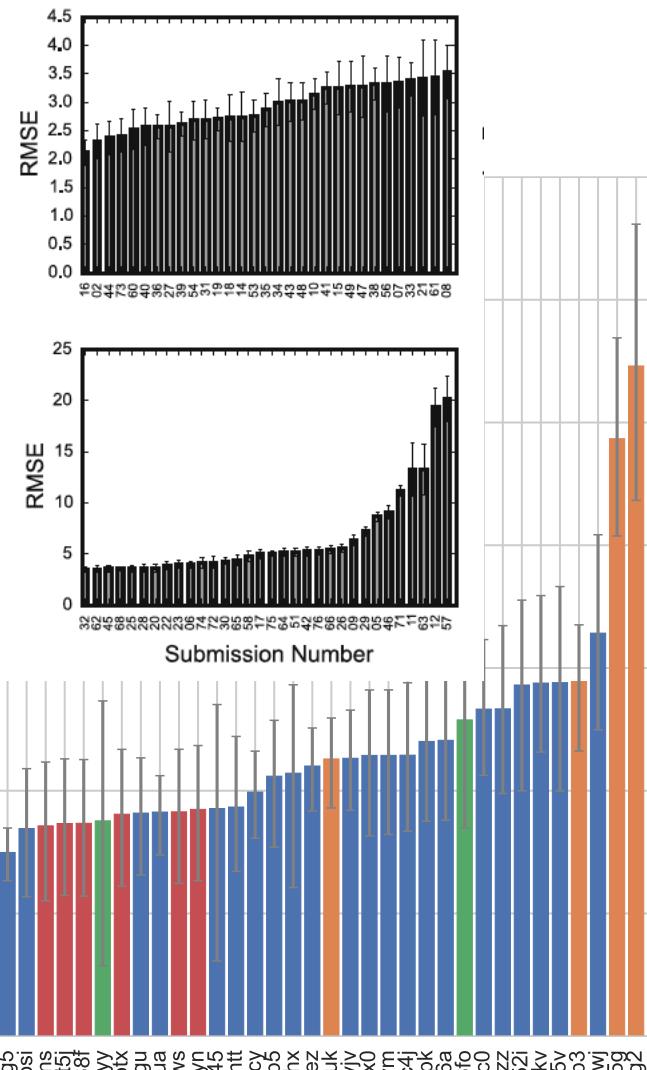
Accuracy evaluation of methods that participated SAMPL6 logP challenge

In SAMPL6 logP challenge:

- **Ionization and tautomerization state predictions** do not contribute to prediction errors.
- **Chemical diversity** in SAMPL6 dataset more is more limited.
- 1-octanol is a more familiar solvent for logP predictions.
- **Heterogeneity of wet octanol phase** is a challenge for physical modeling methods.

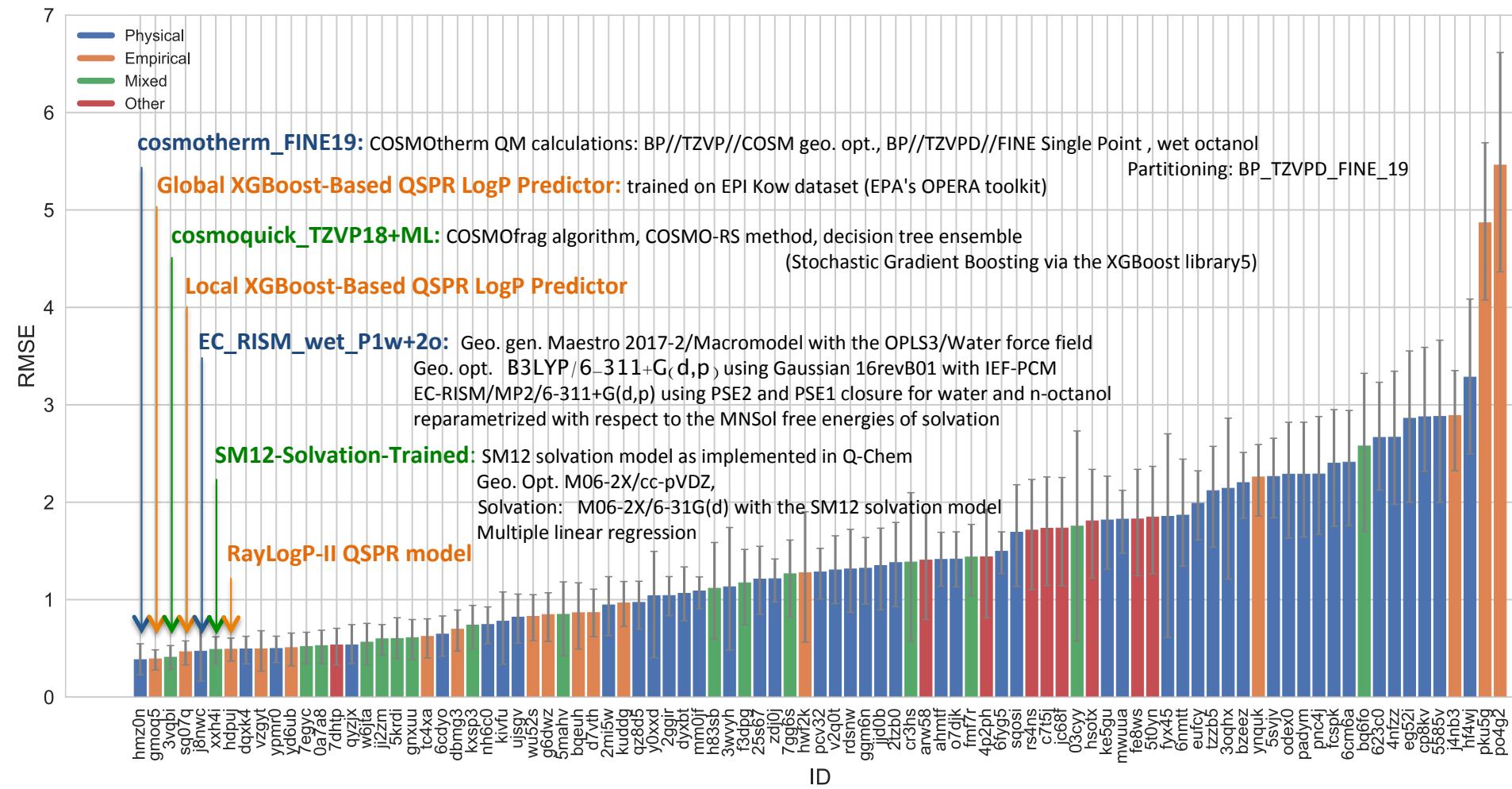


SAMPL5 prediction challenge for cyclohexane-water logD at pH 7.4

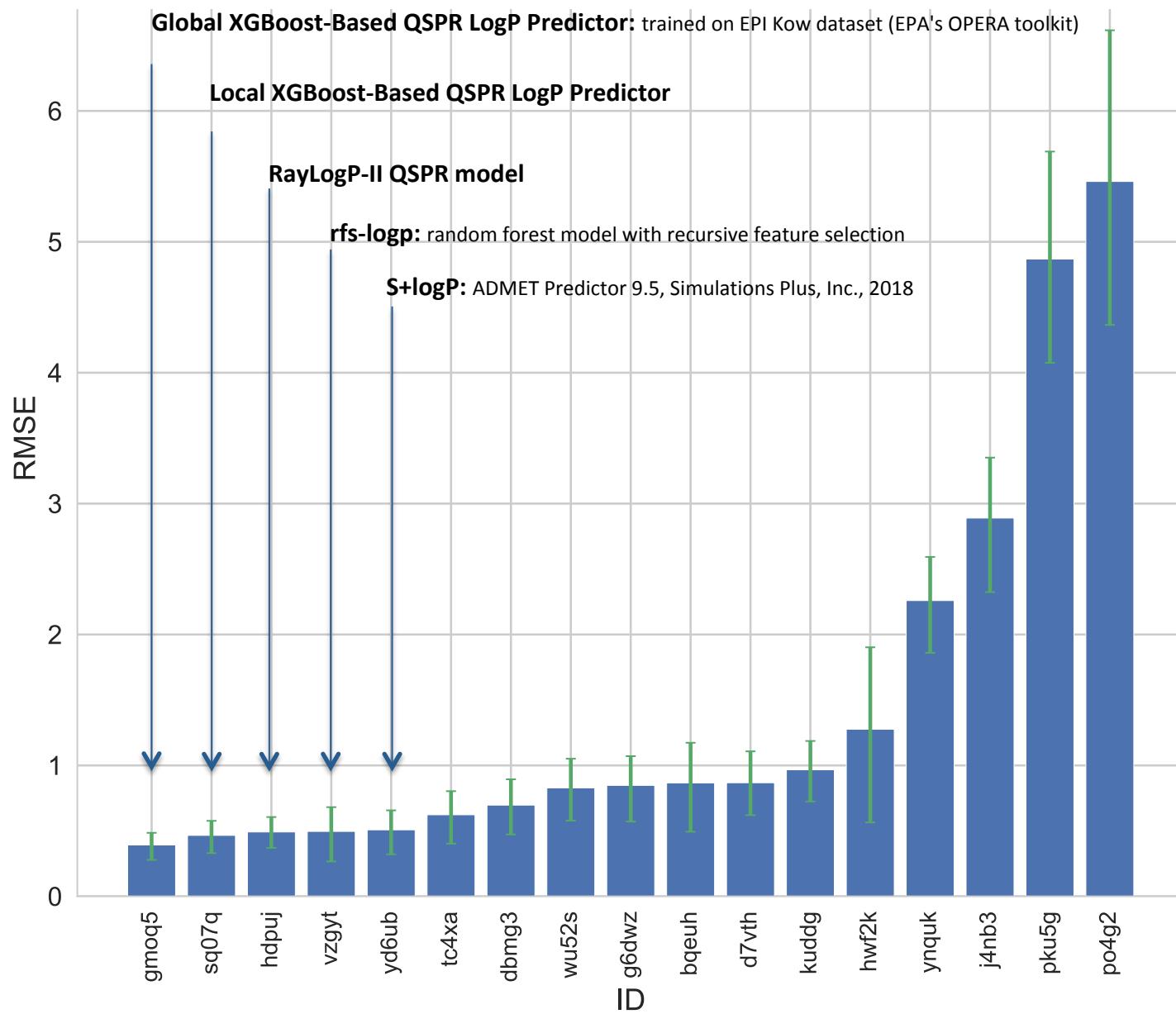


Error bars indicate 95% CI.

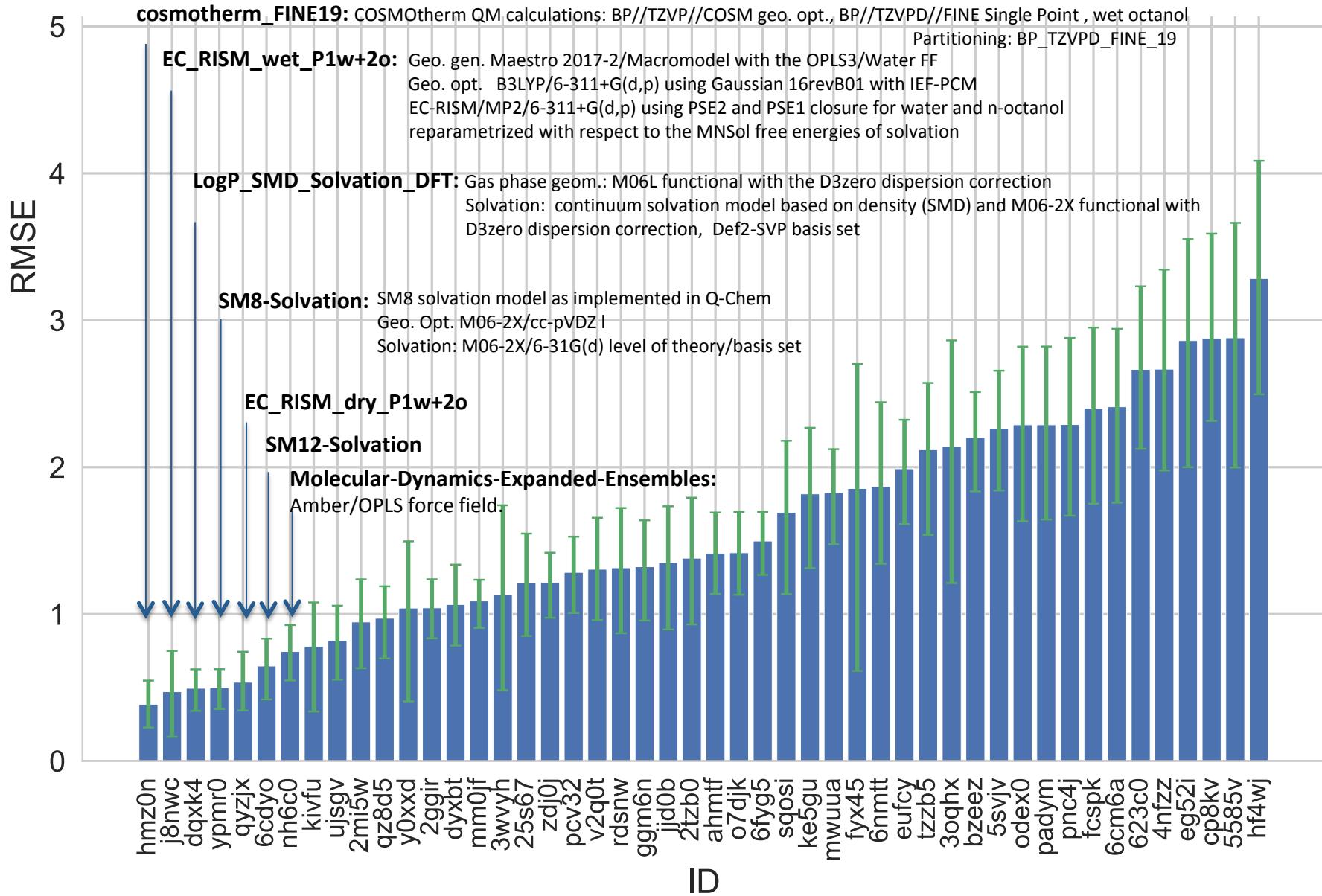
Performance evaluation of participated methods to logP challenge



Prediction accuracy of methods from “empirical” category



Prediction accuracy of methods from “physical” category



Prediction accuracy of methods from “mixed” category

cosmoquick_TZVP18+ML: COSMOfrag algorithm, COSMO-RS method, decision tree ensemble
(Stochastic Gradient Boosting via the XGBoost library⁵)

SM12-Solvation-Trained: SM12 solvation model as implemented in Q-Chem
Geo. Opt. M06-2X/cc-pVDZ, Solvation: M06-2X/6-31G(d) with the SM12 solvation model
Multiple linear regression

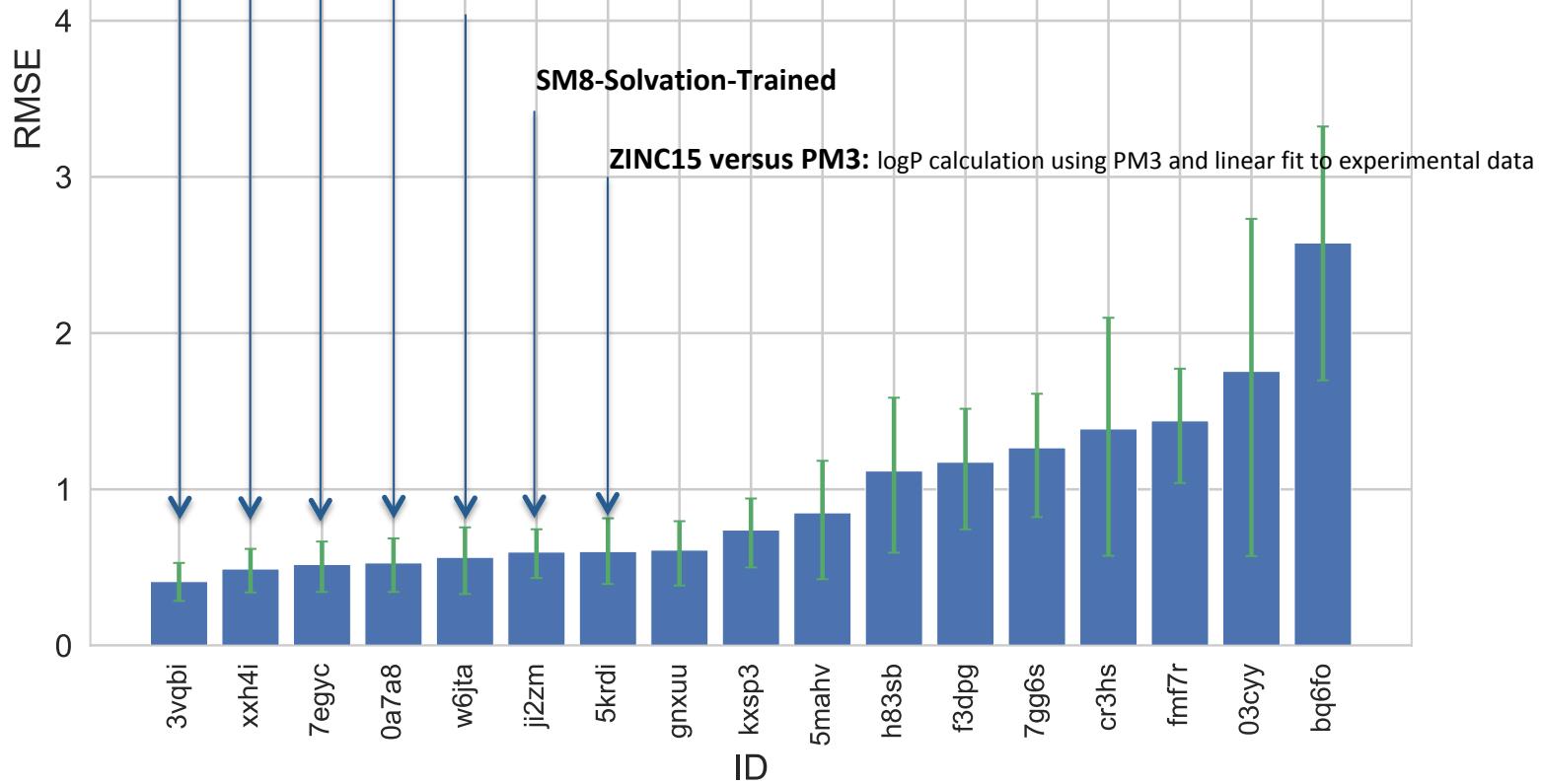
SMD-Solvation-Trained

**ML Prediction using MD Feature Vector Trained on logP_octanol_water,
with Additional Meta-learner**

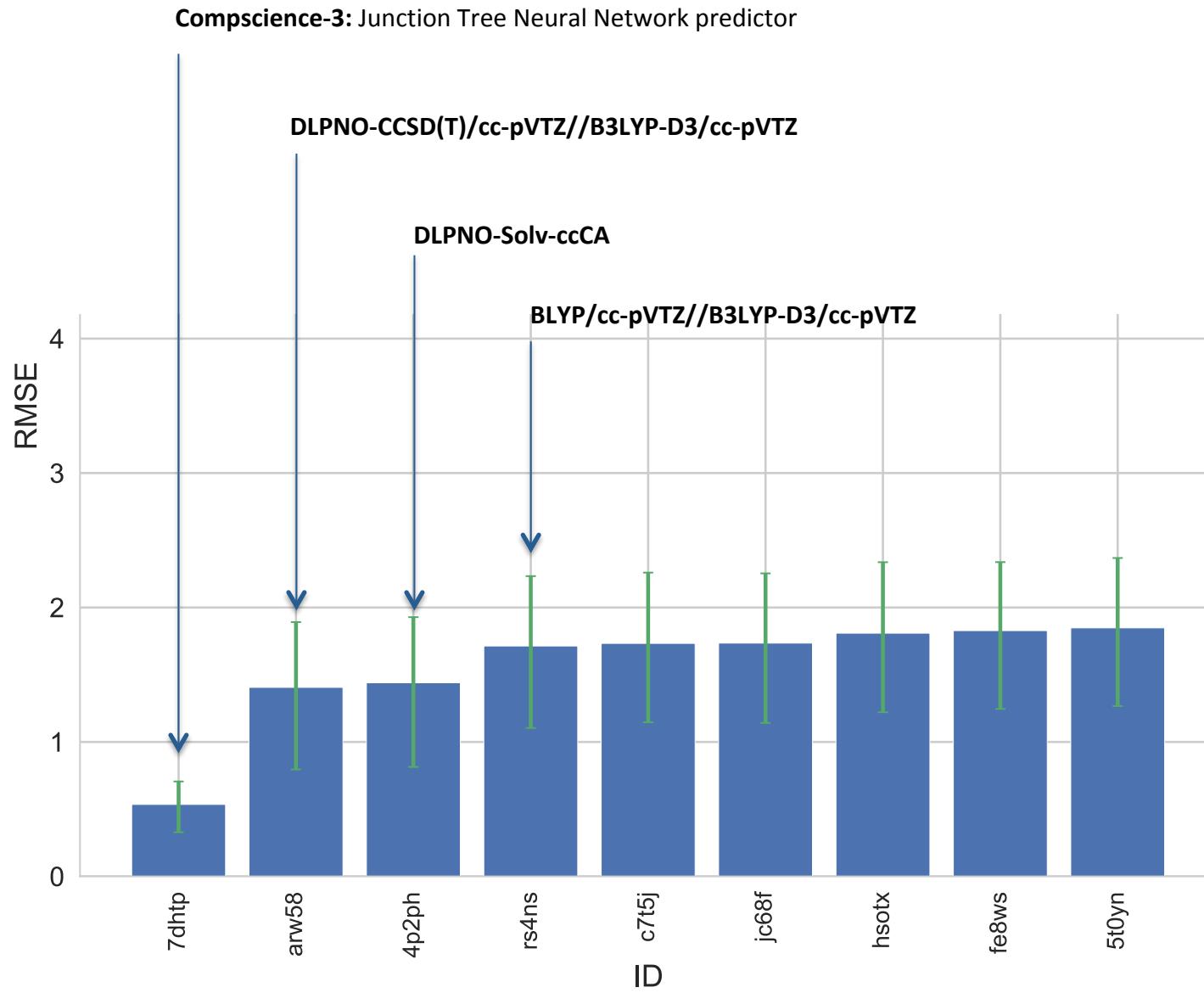
ML Prediction using MD Feature Vector Trained on logP_octanol_water

SM8-Solvation-Trained

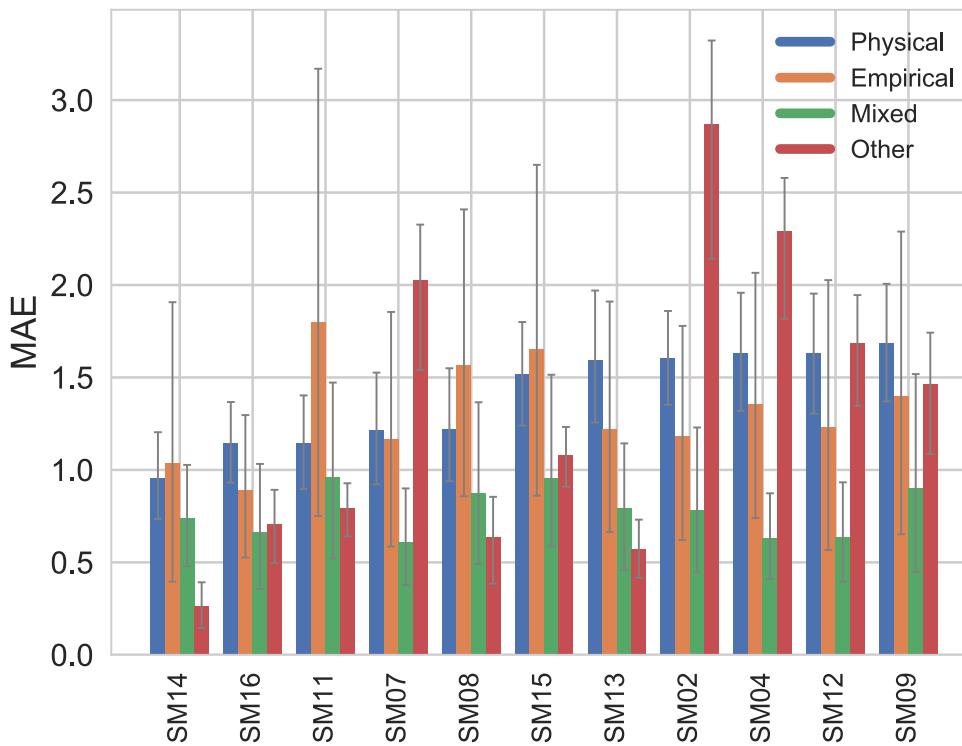
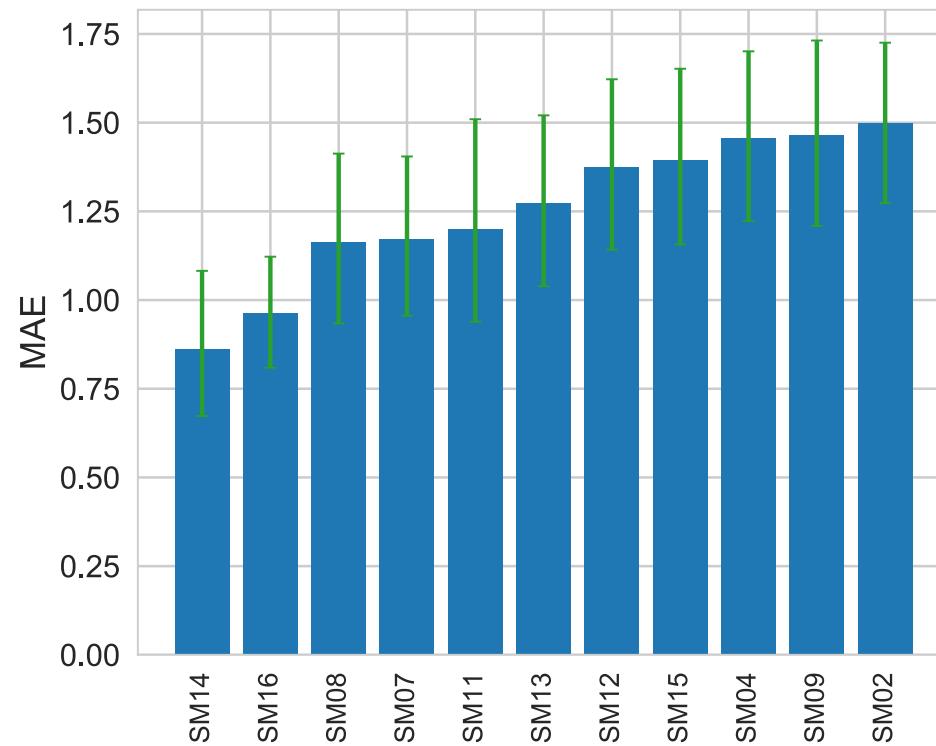
ZINC15 versus PM3: logP calculation using PM3 and linear fit to experimental data



Prediction accuracy of methods from “other” category



Analysis of average prediction accuracy of each molecule



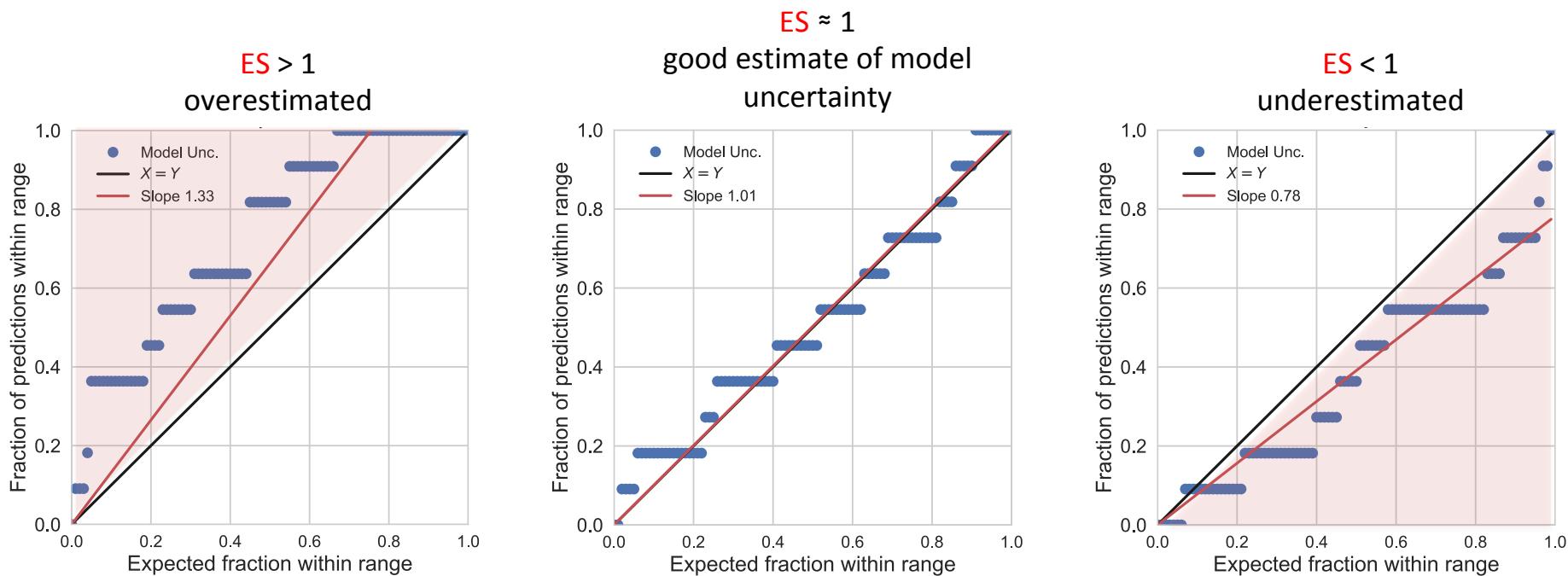
There wasn't any molecule that was significantly harder to predict than others.

To see the distribution of accuracy across models for each molecule more detail please see the violin plots found in SAMPL6 repository:

[SAMPL6 / physical_properties / logP / analysis / analysis_outputs / error_for_each_logP.pdf](#)

How accurate were model uncertainty predictions?

To evaluate if model uncertainty is **overestimated** or **underestimated** we calculated **error slope (ES)**:
The slope calculated from linear least squares fit to Q-Q Plot.



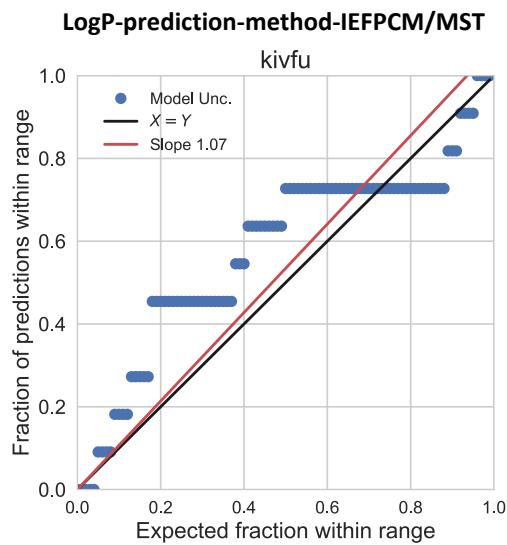
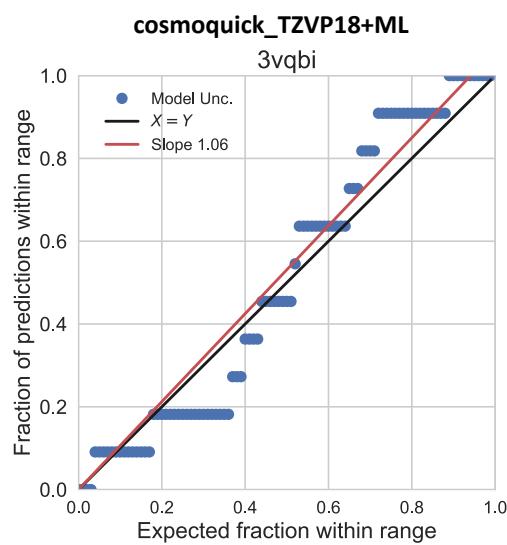
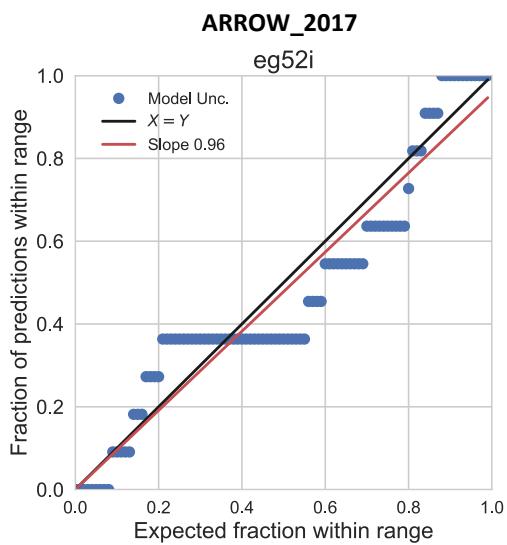
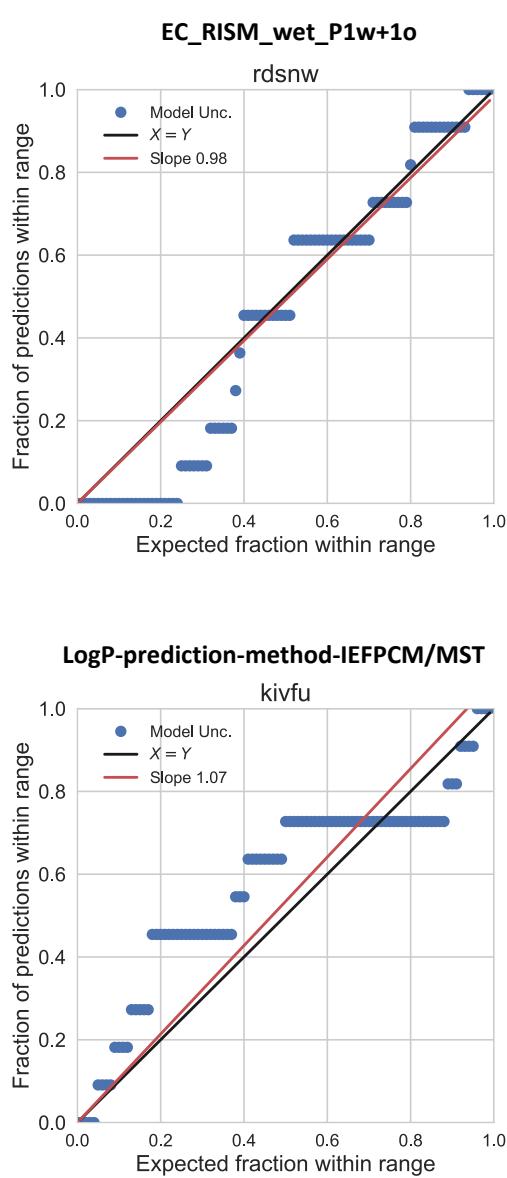
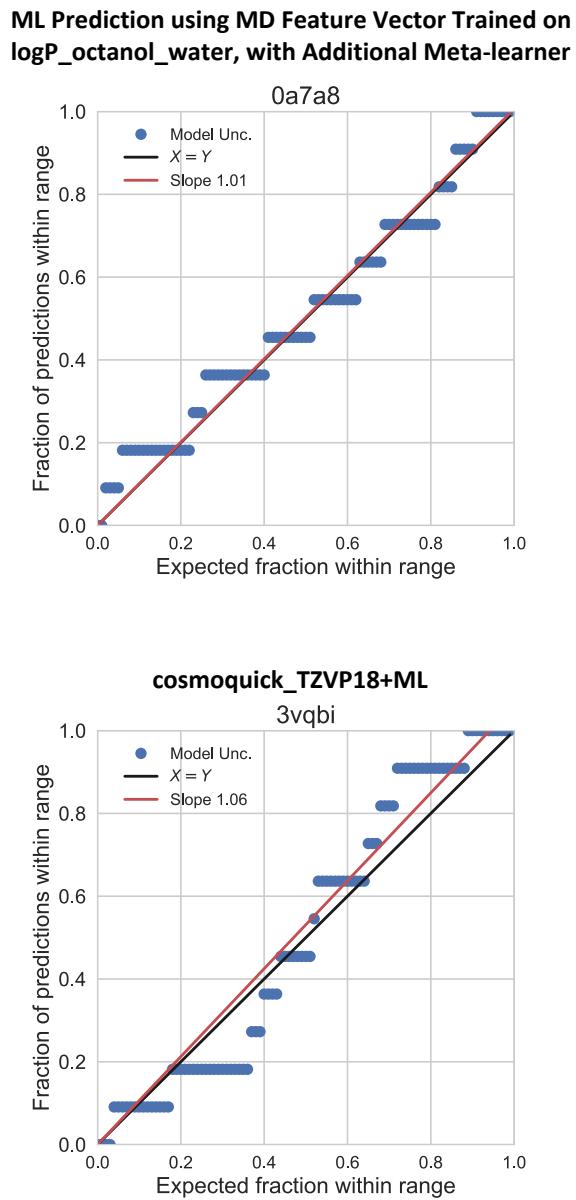
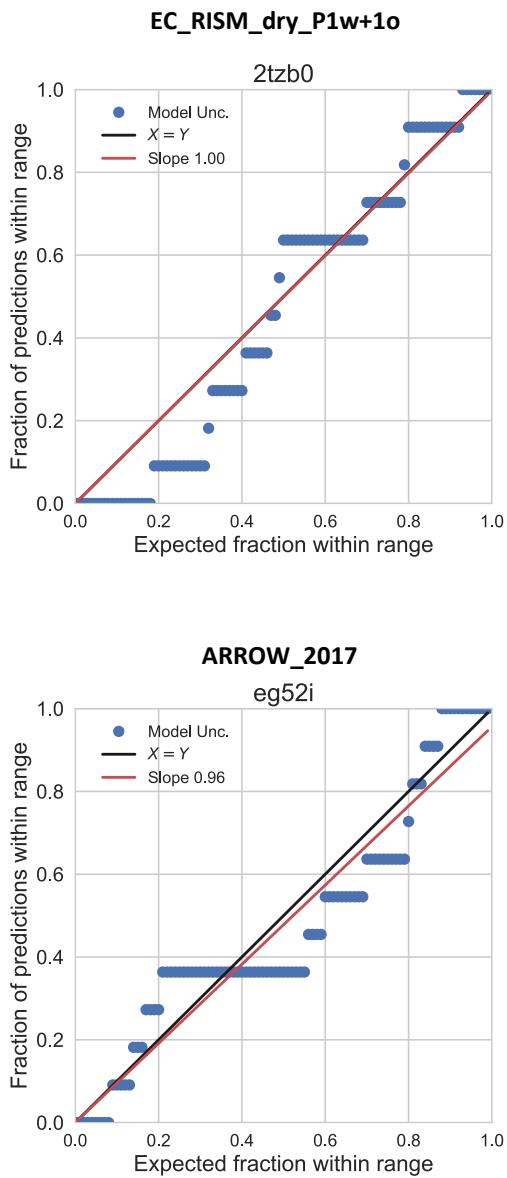
Q-Q Plots for each method can be found in SAMPL6 GitHub repository:

[SAMPL6 / physical_properties / logP / analysis / analysis_outputs / QQPlots /](#)

For comparison of error slope values please refer to:

[SAMPL6 / physical_properties / logP / analysis / analysis_outputs / StatisticsTables / statistics.pdf](#)

Which methods were very good at estimating their model uncertainty?



Suggested directions for analysis of method performance

1. Due to limited dynamic range of experimental data, prefer **RMSE** and **MAE** over correlation statistics to judge model performance.
2. Is your method able to capture $\log P$ rank order of **4-amino quinazoline** series? Is it capable of capturing substituent effects to $\log P$?
3. Are there any **molecules with lower prediction accuracy** than others?
4. Is **bias or variance** the limiting factor for the accuracy of your method?
5. How does your method perform compared to similar methods in SAMPL6 challenge?

Refer to submissions table to find similar methods: [SAMPL6](#) / [physical_properties](#) / [logP](#) / [analysis](#) /
[Submission IDs for log P prediction methods](#)

SAMPL6 $\log P$ challenge submissions were listed in the ascending order of RMSE.

Submission ID	Method Name	Category
hmz0n	cosmotherm_FINE19	Physical
gmoq5	Global XGBoost-Based QSPR LogP Predictor	Empirical
3vqbi	cosmoquick_TZVP18+ML	Mixed
sq07q	Local XGBoost-Based QSPR LogP Predictor	Empirical
j8nwc	EC_RISM_wet_P1w+2o	Physical

6. When relevant, we suggest pairing up with other participants and writing papers that compare methods together.

Refer to user map to learn who submitted a method you are interested in:

[SAMPL6](#) / [physical_properties](#) / [logP](#) / [predictions](#) / [SAMPL6-user-map-logP.csv](#)

Acknowledgments

SAMPL6 organizers and advisors

David Mobley
John Chodera
Daniel Bergazin
Caitlin Bannan
Andrea Rizzi
Bas Rustenburg

D3R

Michael Gilson
Rommie Amaro
Michael Chiu

Merck Preformulation Group

Timothy Rhodes
Dorothy Levorse
Heather Wang
Brad Sherborne

NIH R01 GM124270

Tri-Institutional PhD Program in Chemical Biology

Doris J. Hutchison Fellowship

Participants of logP challenge

Alexei Nikitin
Andreas Kraemer
Andrew Paluch
Bogdan Iorga
Boris Sattarov
Chris Loschen
Chung Kee-Choo
Davy Guan
Diogo Santos Martins
Esteban Vohringer-Martinez
Evrim Arslan
Floriane Montanari
Ganesh Kamath
Hwangseo Park
Jae Hong Shin
Jing Huang
Jose Jimenez Luna
Michael Jones
Mohammed Ali
Piero Procacci
Prajay Patel
Raymond Lui
Robert Fraczkiewicz
Samarjeet Prasad
Shuzhe Wang
Stefan Kast
William Zamora

Participants who will be presenting today at the virtual workshop and their submissions

