Improving pKa modelling through active collaborations

Genentech
A Member of the Roche Group

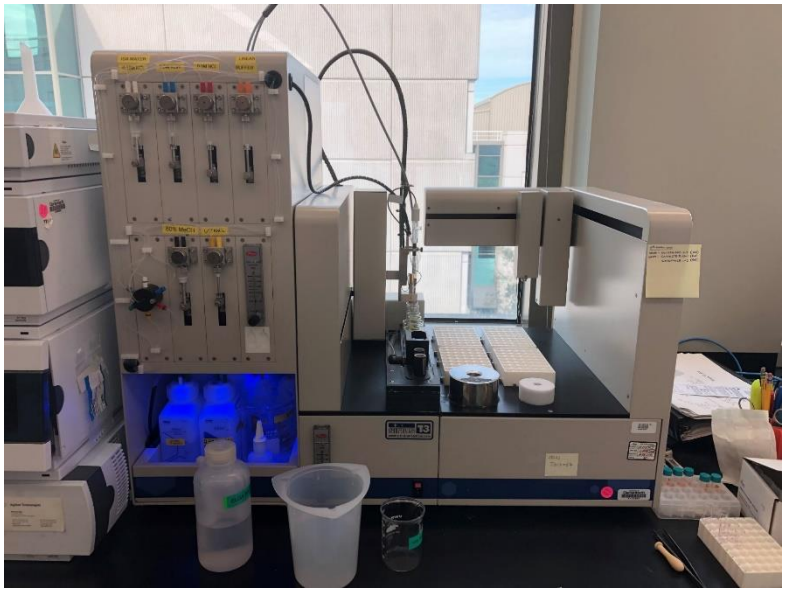
Huy Q. Nguyen
Ionization Summit
October 31st, 2023

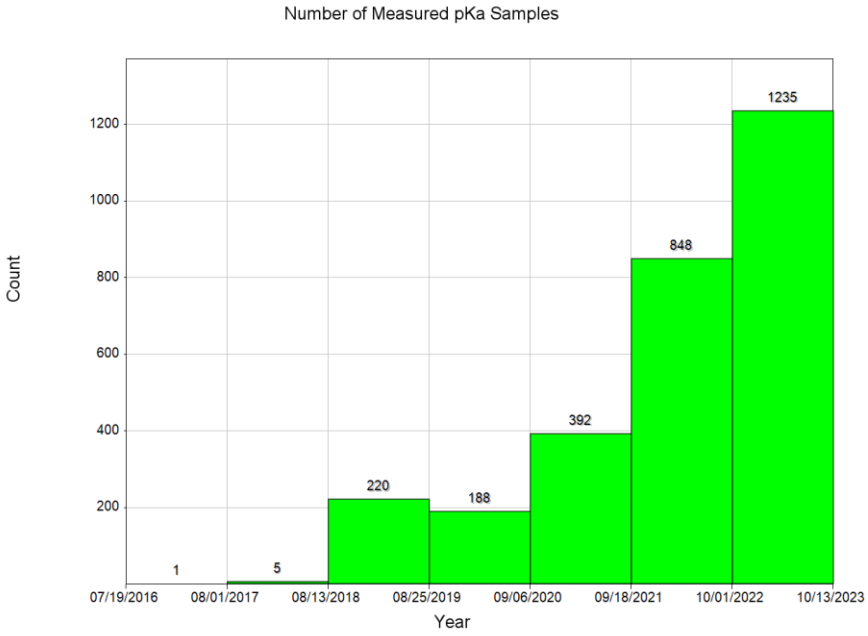
What needs to be completed for collaboration

- Consistency in pKa measurement methods
- “cleaning” experimental data to make sure it makes sense
- Working with legal team to identify what data can be shared

Measuring pKa

- All shared data measured with Pion Sirius T3 system
- UV and pH-metric
- Measuring pKa has increased in demand





Cleaning pKa data and data validation

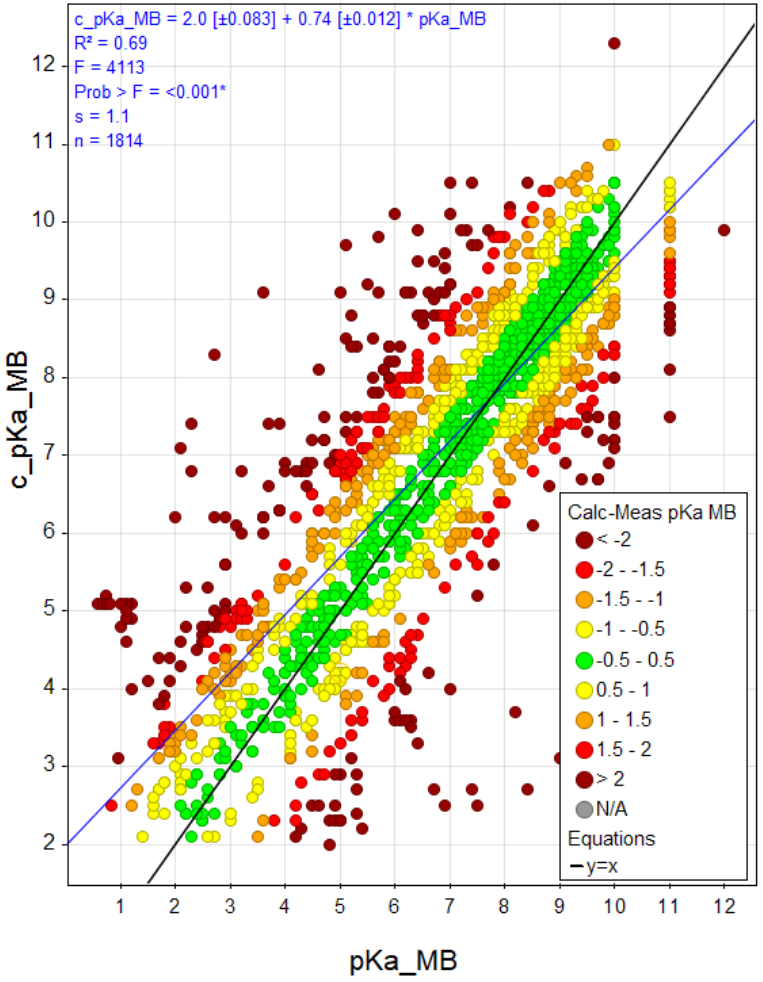
|Measured – calculated| pKa values: Issues Identified

Issue	MA > 2	MA 1.5-2	MA 1-1.5	MA < 1	MB >2	MB 1.5-2	MB 1-1.5	MB < 1	Total
Total	133	75	89	331	201	158	275	1180	2272
Calculation error	109	67	84		156	135	268		819
Need remeasuring	17	6	3		33	17	7		83
Data issues	5	2	2		18	8	0		35
OK as is	1								1

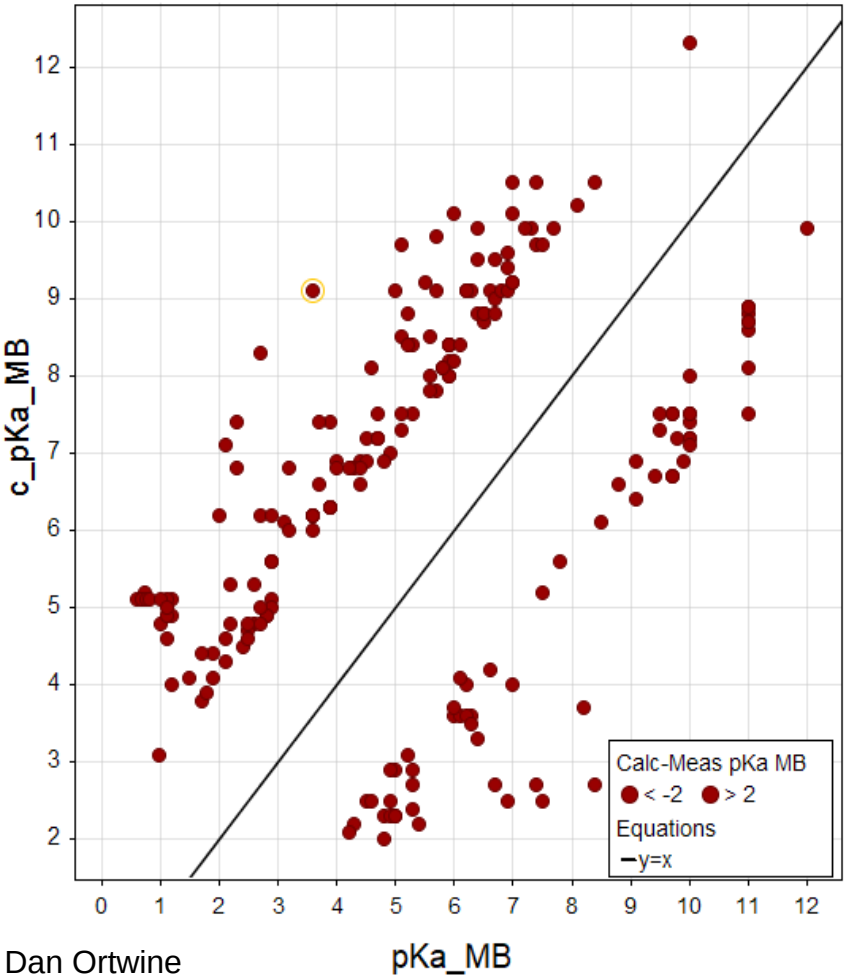
Basic pKa Values

201 with |measured – calculated| pKa_MB values ≥2 log units

All pKa_MB values



Miscalculated by > 2 log units

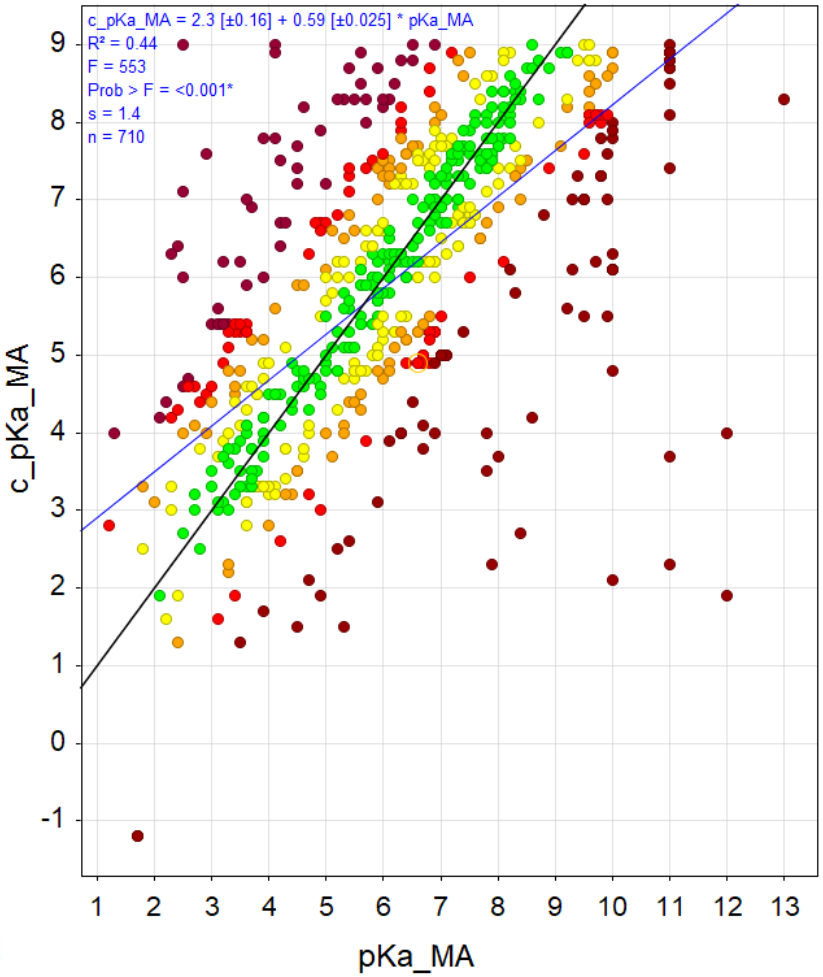


Analysis by Dan Ortwine

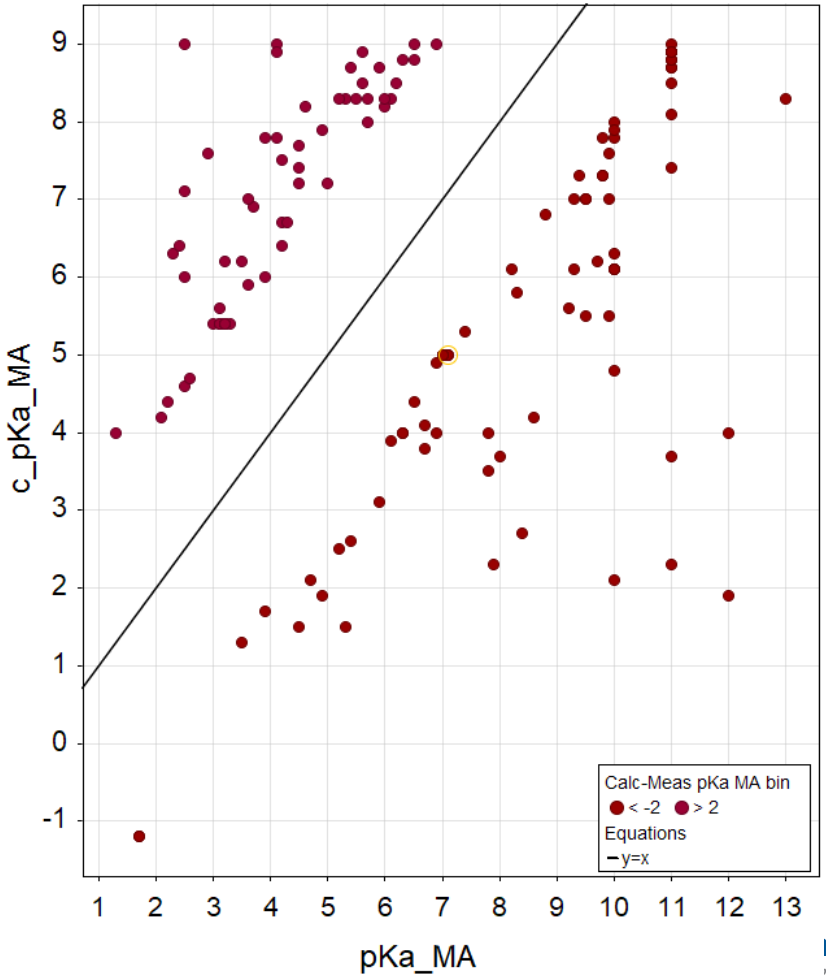
Acidic pKa values

133 with |measured – calculated| pKa_MA values ≥2 log units

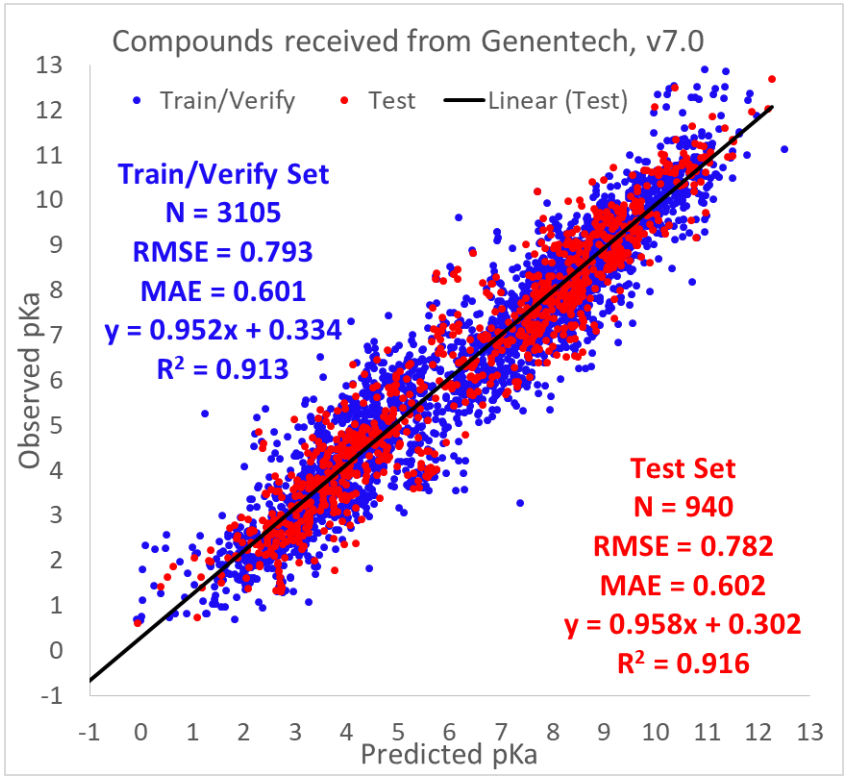
All pKa_MA values

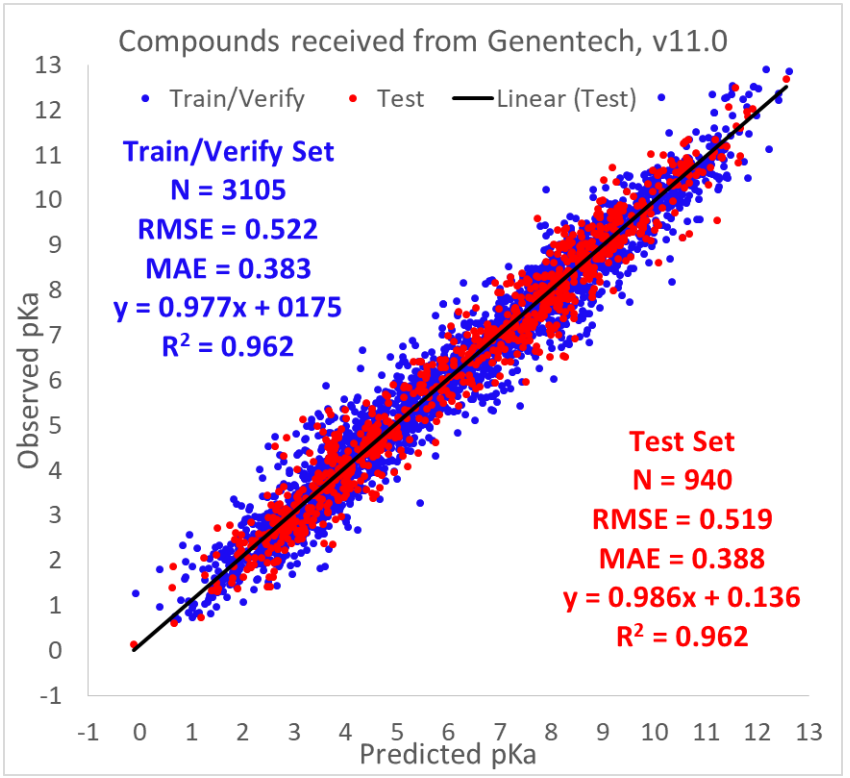


Miscalculated by > 2 log units



Results from the collaboration



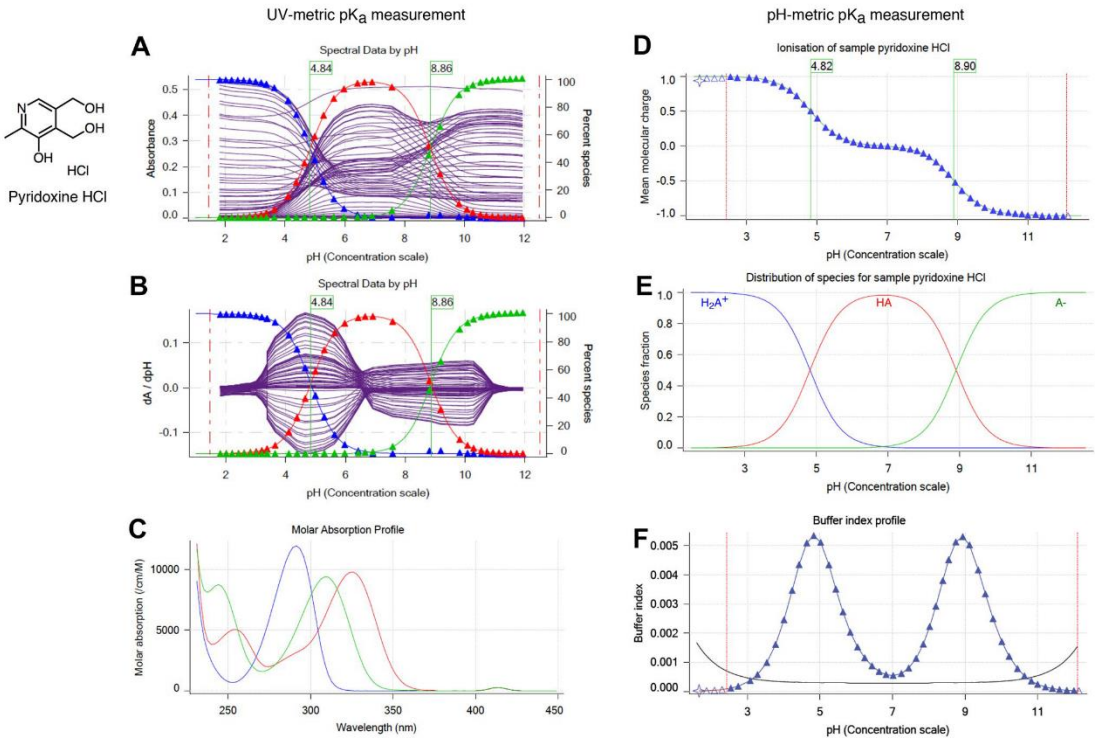


Genentech Discovery Chemistry

Genentech
A Member of the Roche Group

How is the cpKa data utilized?

- Predicted pKa data utilized for experimental setup
- Data validation – does the measured pKa make sense?
- Prioritize instrument and analysis time for challenging pKa samples



Mehtap Isik, *et al., Journal of Computer-Aided Molecular Design*, **2018**, 32, 1117-1138.

Genentech Discovery Chemistry

Genentech
A Member of the Roche Group

Concluding Remarks

- Most challenging aspect was cleaning/validating internal data – requires a team
- Sharing data with SimulationPlus was a valuable exercise
- ADMET Predictor v11 has been performing really well predicting cpKa

Thank you! Questions?

Newton Wu