Efficient and Accurate pK_a Prediction Enabled by Pre-Trained Machine-Learned Interatomic Potentials

Corin C. Wagen* and Arien M. Wagen

Rowan Scientific Corporation, Boston, Massachusetts 02134, United States

ABSTRACT: Quickly and accurately predicting the pK_a of small molecules is an important unsolved challenge in computational chemistry: while approaches based on electronic structure theory have shown great promise, the utility of these methods is limited by the considerable expense of the requisite computations. In this study, we investigate AIMNet2, a machine-learned interatomic potential, as a low-cost surrogate for electronic structure theory in pK_a prediction. The accuracy of the AIMNet2-based pK_a prediction workflow is evaluated over a wide range of compounds and functional groups, and potential sources of error are discussed.

INTRODUCTION

Knowing the pK_a of a molecule is key to understanding its structure and reactivity. In medicinal chemistry, pK_a values are used to predict a variety of important pharmacological properties, including solubility, hERG potassium channel blocking, phospholipidosis risk, CYP inhibition, and blood–brain-barrier penetration.[1–6] Efficient and accurate pK_a estimates play a pivotal role in the design of compounds with the desired physicochemical and pharmacological properties, and accordingly pK_a prediction has been recognized as a problem of "immense interest"[7] in computational chemistry for decades.[8,9]

The most common approach to pK_a prediction is to develop an empirical quantitative structure–property relationship (QSPR) based on experimentally measured data and use this model to predict pK_a values for new compounds.[10] QSPR approaches are typically fast and display excellent accuracy for compounds like those in the training data, making them an appealing choice for high-throughput pK_a prediction.[11] However, these methods struggle to extrapolate to unseen compounds and are generally unable to capture factors that arise from specific configurations of molecules in three-dimensional space, like conformational or stereoelectronic effects.[12]

A contrasting approach is to directly predict pK_a values using electronic structure theory, since microscopic pK_a values are directly proportional to the ΔG of the underlying acid dissociation reactions. This strategy possesses certain advantages over QSPR approaches: conformational and stereochemical effects are naturally incorporated, and new regions of chemical space can be modeled without new training data. In recent pK_a prediction challenges, approaches based on electronic structure theory have consistently been among the best-performing methods: in SAMPL6, submission *xvxzd* achieved a mean absolute error (MAE) of 0.58 pK_a units and a root-mean-squared error (RMSE) of 0.68,[13] while submission *EC_RISM* achieved an MAE of 0.53 and an RMSE of 0.73 in SAMPL7.[14]

Unfortunately, many electronic structure theory-based approaches are too computationally costly for routine usage: in their paper describing the *xvxzd* submission to SAMPL6, Pracht and Grimme write that "a macroscopic pK_a value for a molecule could be calculated within approximately one day on a 28 CPU node,"[15] which corresponds to roughly $25–30 with current cloud computing prices.[16] While semiempirical methods like PM6[17] and GFN2-xtb[18] have been used for rapid pK_a prediction, the resulting workflows display lower accuracy and often require judicious choice of reference compound or significant reparameterization. This is unsurprising: calculation of acid dissociation constants is "demanding and arduous,"[19] and both a sophisticated electronic structure method and a sizable basis set are required to consistently get accurate results even in the gas phase.[20]

Machine-learned interatomic potentials have recently emerged as an efficient alternative to conventional electronic structure theory.[21] We envisioned that AIMNet2, which displays accuracy comparable to routine density-functional theory methods for main-group thermochemistry,[22] might allow for rapid computation of pK_a values while preserving the advantages of electronic structure-based approaches. Here, we report the successful implementation of this strategy into a workflow ("Rowan pK_a"), which uses AIMNet2 to calculate microscopic pK_a values with good accuracy and minimal empiricism.
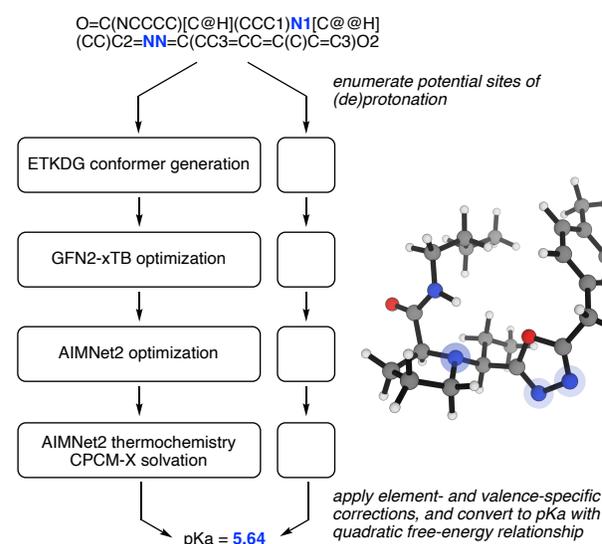
Figure 1. Visual overview of the Rowan pKa workflow.

METHODOLOGY

Our workflow begins by iteratively adding or removing hydrogens to the molecule of interest using RDKit and quickly estimating the proton affinity for the conjugate acid/conjugate base pair. If (de)protonation is predicted to yield a pK_a value close to the desired range, conformers are generated using the ETKDG algorithm[23,24] and optimized using the MMFF forcefield. After removing redundant geometries, low-energy conformers are then optimized using the GFN2-xTB semiempirical method,[25] single-point energies are calculated using the AIMNet2 model trained on ωB97M-D3/def2-TZVPP (henceforth abbreviated as AN2-ωB97MD3),[22] and the lowest N undergo full optimization and frequency calculations at the AN2-ωB97MD3 level of theory to generate a free energy for each conformer, including all standard entropy- and symmetry-based corrections.[26] (N is set to 10 by default but can be varied by the end user; *vide infra*.) Since AIMNet2 does not take solvation into account, ΔG_{solv} is computed from a single-point GFN2-xtb calculation with the CPCM-X implicit water model.[27]

Following Pracht and Grimme, we combine energies from each conformer to generate a single hybrid ΔG for the conjugate acid/conjugate base pair and apply a quadratic free-energy relationship to convert this into pK_a values.[18] Initial application of this strategy provided disappointing results, with considerable systematic error observed between different classes of acids. Functional group-dependent systematic error is common in *ab initio* pK_a prediction, and many different corrections have been employed: Schrödinger's Jaguar pK_a contains a different linear free-energy relationship for roughly a hundred different functional groups,[28] while Jensen and co-workers employed different reference molecules for each class of compound under study.[17] To minimize the number of empirical parameters employed in our workflow, we instead add an element- and valence-specific constant to ΔG for each conjugate acid/conjugate base pair. (The correction for divalent oxygen, i.e. deprotonation of ROH, is defined as zero to eliminate extra degrees of freedom.) These constants are determined by Levenberg–Marquardt least-squares minimization, not by comparison to any individual reference compound, and can be thought of as a correction for systematic errors in both thermochemistry and solvation. Similar corrections are often used to account for dispersion in mean-field electronic structure theory methods without substantially compromising the generality of the method.[29]

Workflow development and parameter optimization was conducted on a slightly modified version of the TR224 dataset, also used by Pracht and Grimme.[18] The final Rowan pK_a workflow contains 7 tunable parameters, shown in Table 1.

Table 1. Tunable Parameters for Rowan pK_a

Coefficients for Free-Energy Relationship	
C0:	-123.242550
C1:	0.50677935 mol/kcal
C2:	-1.7401e-04 mol^2/kcal^2
Element-/Valence-Specific Constants (kcal/mol)	
N3:	3.91398
N4:	5.13054
C4:	6.03872
O2:	0.00 *(defined)*
S2:	-5.14438

RESULTS

Quantifying the accuracy of property prediction under realistic conditions is challenging: evaluation over a large dataset, while superficially appealing, often leads to exaggerated correlation coefficients that are not reproduced in smaller, less diverse populations.[30] Instead, we follow Rich and Birnbaum in using "assay-stratified metrics" to evaluate the predictive ability of our pK_a workflow.[31] All datasets were scored using four metrics: mean absolute error (MAE), root mean square error (RMSE), the square of the Pearson correlation coefficient (r^2), and Kendall's tau (τ). Low values for MAE and RMSE reflect a high degree of absolute accuracy, while high values for r^2 and τ demonstrate relative accuracy within a given dataset (arguably more predictive of performance under real-world usage).

We began by evaluating Rowan pK_a on datasets from the SAMPL6 and SAMPL7 blind pK_a prediction challenges run several years ago. The SAMPL6 dataset consists of 24 molecules designed to resemble fragments of kinase inhibitors, many of which contain multiple potential sites of (de)protonation.[13] We computed microscopic pK_a values starting from each molecule's neutral state and compared these values to the relevant macroscopic pK_a values, excluding doubly ionized states (Figure 2). Rowan pK_a performed well, with an MAE of 0.94 and an RMSE of 1.21, comparable to the performance of commercial pK_a prediction tools like ChemAxon's Chemicalize and Schrödinger's Jaguar and Epik.[32] The high values obtained for r^2 and τ also illustrate the predictive utility of Rowan pK_a, although the large range of pK_a values in the SAMPL6 dataset makes obtaining a high value for r^2 relatively easy.
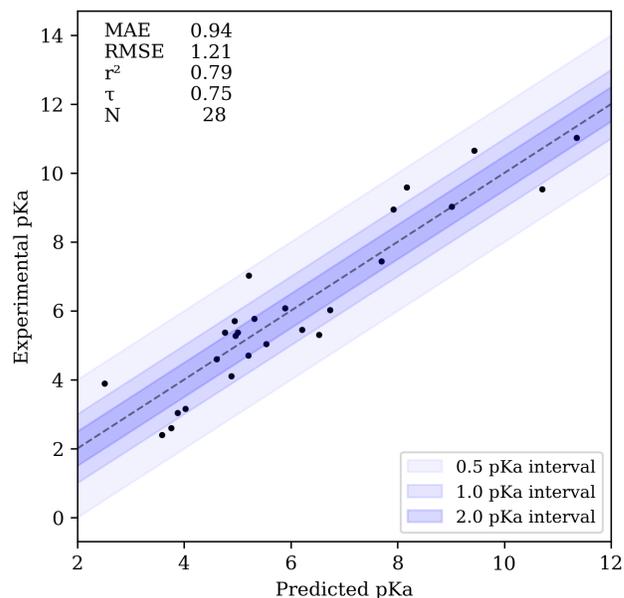
Figure 2. Correlation plot for the SAMPL6 dataset.

The SAMPL7 dataset consists of 20 N–H acids with experimentally measured pK_a values.[14] Rowan pK_a performed somewhat less well on the SAMPL7 dataset (Figure 3), with an MAE of 1.38 and an RMSE of 1.72—nevertheless, our workflow still outperformed all but the top two or three entries to the SAMPL7 challenge (depending on which metric is used for ranking), suggesting that SAMPL7 is a more challenging dataset. While most molecules were predicted with reasonable accuracy, Rowan frequently significantly overestimated the acidity of various

sulfonamides, which may point to a limitation of the underlying AN2-ωB97MD3 model.
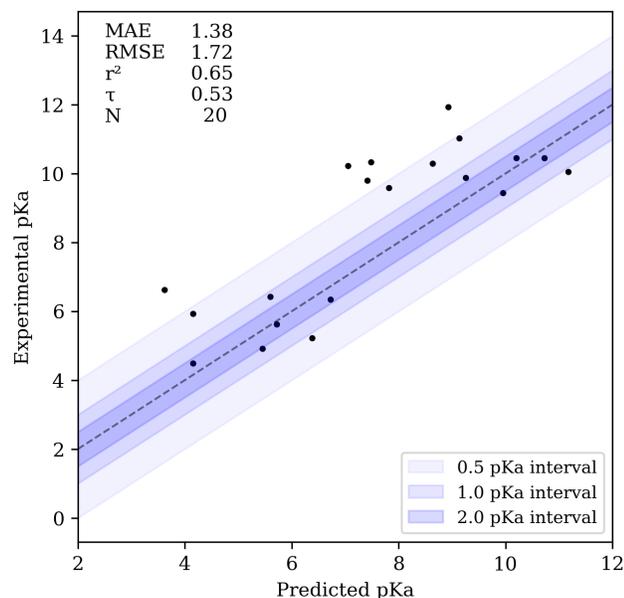
Figure 3. Correlation plot for the SAMPL7 dataset.

We next evaluated Rowan pK_a on three datasets selected to showcase changes in amine basicity. In 2010, Müller and co-workers reported that introducing spiro-oxetane substituents in saturated *N*-heterocycles could attenuate the basicity of amines.[33] We modeled their dataset using Rowan pK_a, with the change that the bulky *N*-piperonyl group (added as a UV-absorbing handle for experimental convenience) was replaced with *N*-methyl *in silico*. Overall, good agreement between experimental and predicted pK_a was observed, although the effect of β-oxetane substituents was systematically underestimated (Figure 4). More recently, scientists from Enamine reported a new route towards the synthesis of α-CF$_3$ bridged bicyclic amines and measured the experimental pK_a values for a variety of bridged bicyclic amines.[34] Rowan pK_a was able to reproduce the observed attenuation of basicity induced by the α-CF$_3$ group, although subtle changes in pK_a caused by differing ring size were not captured, leading to a poor value for τ (Figure 5).

In a 2017 paper, Leito and co-workers measured the basicity of a range of unsaturated *N*-heterocycles in several different solvents, including water.[35] We selected 28 of these values for a test set, excluding compounds which were in our initial training set (pyridine and pyridazine). Rowan pK_a was able to reproduce the measured pK_a values with good accuracy, including effects from stabilizing intramolecular hydrogen bonding which are challenging to identify without explicit consideration of 3D conformers (Figure 6). In fact, the largest errors came from overestimating the effect of intramolecular hydrogen bonding: for instance, the basicity of 1,10-phenanthroline was overestimated by about 2.5 pK_a units (Figure 6), which may point to inaccuracies in the directionality of N–H–N hydrogen bonding in AN2-ωB97MD3.

Exp. pKa: 10.0	Exp. pKa: 6.3	Exp. pKa: 8.1	Exp. pKa: 6.1
Pred. pKa: 9.40	Pred. pKa: 7.37	Pred. pKa: 9.56	Pred. pKa: 6.89
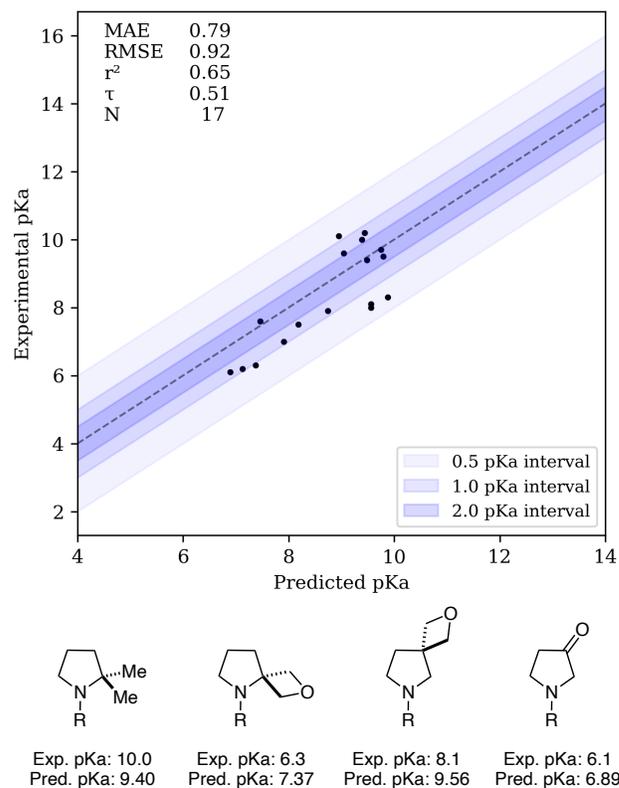
Figure 4. Correlation plot for the Müller spiro-oxetane dataset and illustration of Rowan's failure to model β-oxetane substituents (R = piperonyl for experiments, methyl for computations).
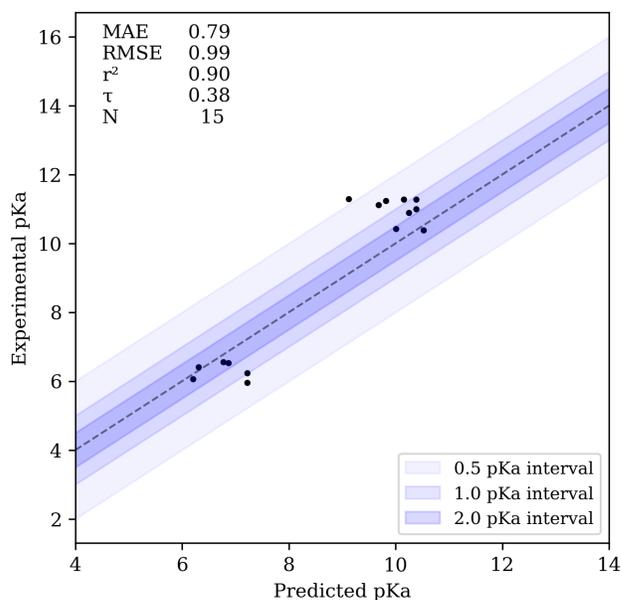
Figure 5. Correlation plot for the Enamine α-CF$_3$ amine dataset.
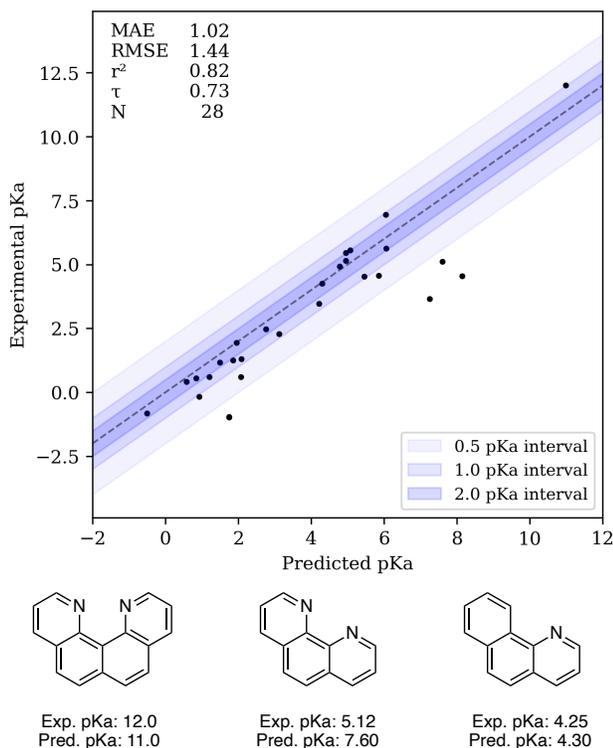
Figure 6. Correlation plot for the aromatic *N*-heterocycle dataset and illustration of the effect of intramolecular hydrogen bonding.

We next examined the ability of Rowan to recapitulate pK_a values from the medicinal chemistry literature. In 1972, Miller, Doukas, and Seydel reported that *N*-aryl sulfonamides served as potent inhibitors of folate synthesis and found that the inhibitory activity could be correlated with the pK_a of the sulfonamide N–H.[36] The Rowan pKa workflow recapitulates the experimental trend in pK_a values and thus also the observed trend in potency, illustrating the utility of pK_a calculations in this system (Figure 7).

Inhibition of BACE1, an enzyme involved in amyloid production, has been extensively investigated as a potential treatment for Alzheimer's disease. BACE1 inhibitors must possess very precise physicochemical properties to pass through the blood–brain barrier: while cyclic amidines have been demonstrated to be potent BACE1 inhibitors, substantial reduction of amidine basicity is necessary to achieve suitable brain penetration.[37] In 2013, Rombouts and co-workers at Janssen investigated the effect of electron-withdrawing groups on the basicity of tricyclic amidine BACE1 inhibitors.[38] We modeled this dataset with Rowan pK_a: while some compounds were modeled quite accurately, Rowan overestimated the effect that an α-nitrile substituent would have on lowering amidine basicity, resulting in large deviations from experiment (Figure 8). Nevertheless, the error induced by an α-nitrile is systematic, allowing for comparisons to be made between different nitrile-containing compounds.
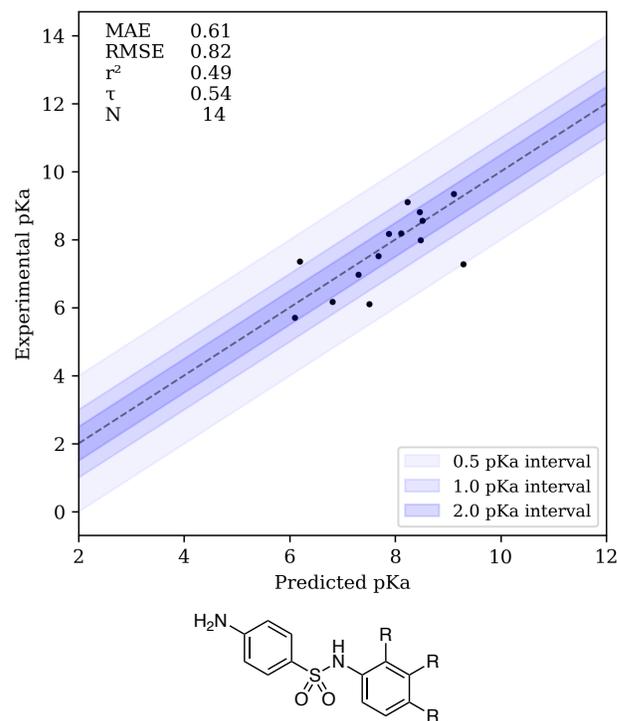
Figure 7. Correlation plot for the folate dataset and illustration of the sulfonamides investigated by Miller, Doukas, and Seydel.
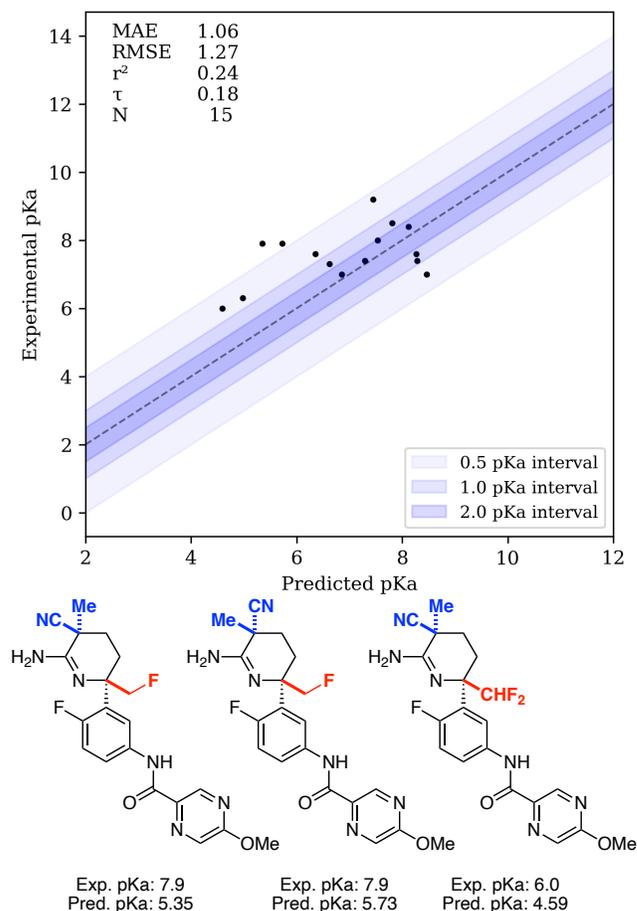
Figure 8. Correlation plot for the Rombouts dataset and illustration of Rowan's underestimation of the pK_a of α-cyano amidines.

In 2007, Müller and co-workers at Hoffmann–La Roche found unusual pK_a effects in a series of stereochemically complex tricyclic thrombin inhibitors.[39] This paper has subsequently been used as a test set for pK_a prediction and has been observed to be particularly challenging in several different cases.[9,17] We too found this to be a challenging dataset, with an MAE of 1.63 and an RMSE of 1.94. Visual analysis of the predictions shows that Rowan pK_a systematically overestimates the pK_a of the tricyclic amines and that substantial predictive power remains despite the high error, as demonstrated by the moderate r^2 and τ values. These errors may arise in part from the fact that the conjugate acid of these amines is doubly cationic, which may lead to inaccuracies in the computed free energy of solvation.
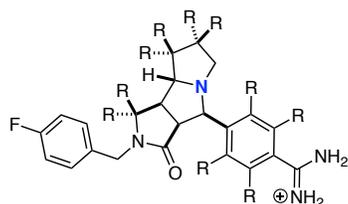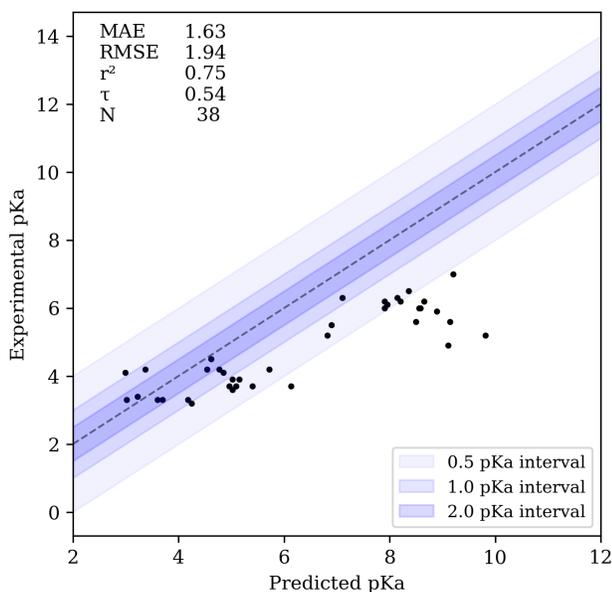
Figure 9. Correlation plot for the tricyclic thrombin inhibitor dataset and illustration of the general tricyclic scaffold investigated by Müller and co-workers, with the site of protonation illustrated in blue.

We next examined the ability of Rowan pK_a to predict the pK_a of macrocyclic peptides. Macrocyclic peptides inhabit different conformations from their linear congeners and frequently possess markedly different pK_a values, with significant consequences for solubility and bioavailability. In 2018, Yudin and co-workers reported that the pK_a of reduced amide bonds adjacent to oxadiazole rings was lowered in macrocycles, which they attributed to intramolecular hydrogen bonding that stabilizes the neutral amine.[40] Rowan pK_a was able to recapitulate the observed trend in pK_a values (Figure 10), which would be considerably more challenging for QSPR-type methods that do not explicitly consider conformational effects.
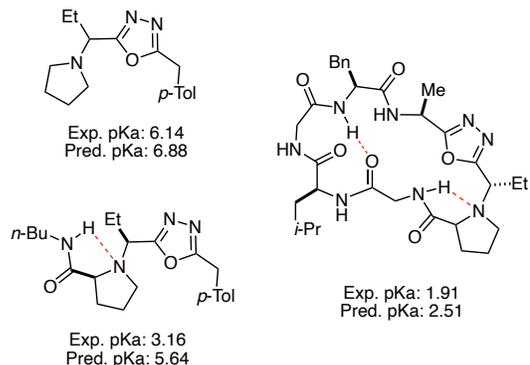
Figure 10. Comparison of predicted and experimentally measured pK_a values for various oxadiazole/pyrrolidine-containing structures, with the intramolecular hydrogen bonds proposed by Yudin and co-workers illustrated in red.

Finally, we explored the suitability of Rowan pK_a as a high-throughput method for pK_a prediction. We selected the first 100 rings from Peter Ertl's recently published library of medicinal chemistry-relevant ring systems as a test set (15.5 atoms/molecule on average) and investigated whether the accuracy of the conformation search/refinement process could be further reduced to increase throughput without significantly hurting the workflow's accuracy.[41] Two modes with reduced conformer searching were eventually developed: the key changes made for each new mode are summarized in Table 2. Modest gains in efficiency are possible, but at the cost of significantly increased error: if multiple low-energy conformers exist, these data suggest that they should be included.[42]

Table 2. Different Modes for Conformer Searching

Parameter	Careful	Rapid	Reckless
# initial confs.	250	100	50
initial energy cutoff (kcal/mol)	15.0	10.0	5.0
RMSD threshold (Å)	0.10	0.25	0.50
# confs. (xTB)	20	10	3
final energy cutoff (kcal/mol)	5.0	5.0	3.0
# confs. (AIMNet2)	10	3	1
MAE (relative to "Careful")	0.00	0.38	0.69
RMSE (relative to "Careful")	0.00	1.17	1.66
Average Time (s)[a]	23.05	16.01	14.13

[a] Timings reported for an EC2 c5.2xlarge instance.

DISCUSSION

In this work, we show that relatively aggressive approximations can be applied to "gold standard" approaches like Grimme's *xvxzd* workflow while still maintaining useful accuracy. The *xvxzd* workflow consists of extensive metadynamics-based conformation searching in *crest*, geometry optimization with PBEh-3c in implicit solvent, single-point energy calculations using the double-hybrid functional DSD-BLYP-D3(BJ)/def2-TZVPD, and subsequent implicit solvent calculations using COSMO-RS; here, we replace *crest* with RDKIT's ETKDG conformer generation algorithm, DSD-BLYP-D3(BJ) with AIMNet2, and COSMO-RS with CPCM-X.[15] While Rowan pK_a is not as fast as QSPR-based approaches or as accurate

 ORCID: https://orcid.org/0000-0003-3315-3524

as high-level electronic structure theory-based approaches like *xvxzd*, the combination of accuracy, speed, and minimal empiricism achieved by our workflow is nevertheless appealing. We anticipate that the ability to obtain fast and reasonable pK_a values for unseen regions of chemical space will be particularly useful in drug discovery. Owing to the lack of functional group-specific corrections, the pK_a values computed by Rowan are ideally suited to downstream manipulation or scaling, which can further reduce systematic error where desired.

What limits the accuracy of the Rowan pK_a workflow, and what can be done to improve it? At a high level, errors in *ab initio* microscopic pK_a prediction can come from several different sources: (1) insufficient consideration of conformational effects or analysis of the wrong conformations, (2) poor description of the underlying gas-phase thermochemistry, or (3) inaccurate description of solvation.[43] To achieve maximum efficiency, a Pareto-optimal pK_a prediction workflow ought to balance errors from all three sources without "overspending" for a high-accuracy answer in only one area. In specific cases, we have been able to track down anomalous predictions made by Rowan pK_a to all three underlying potential sources of error—conformational searching, gas phase thermochemistry, and implicit solvent model—indicating that no factor is solely responsible for the observed error. While improvements to conformational generation[44] and machine-learned interatomic potentials[45] are easy to envision based on current research, improvements to implicit solvation may prove to be more elusive: despite repeated calls for alternatives to "scandalous" implicit solvation,[46] no practical and general alternatives have yet been developed.[47] Recent work in coarse-grained machine-learned potentials may point towards a new paradigm for handling solvent effects, but more work is needed to investigate the strengths and limitations of such approaches.[48,49]

More generally, the approach described in this paper divides the task of predicting pK_a into two separate components: (1) a general pre-trained model which maps molecular structures to energies, thus "learning chemistry" at a high level, and (2) an application-specific workflow which translates geometries and energies generated by the pre-trained model into predictions relevant to the task at hand. This partitioning has certain advantages: the general pre-trained model (here AIMNet2) is systematically improvable and can be trained using purely *in silico* data, in principle allowing arbitrarily high levels of accuracy to be attained, while the application-specific workflow relies largely on conventional computational chemistry techniques and does not require extensive experimental training data like ML-based QSPR methods. As the accuracy of machine-learned interatomic potentials improves, we anticipate that the strategy demonstrated here—using machine-learned interatomic potentials as drop-in replacements for quantum chemistry—will become a general strategy for accelerating computational chemistry.

AUTHOR INFORMATION

Corresponding Author

* corin@rowansci.com

Notes

The authors declare the following competing financial interest: C.C.W and A.M.W are co-founders of the Rowan Scientific Corporation, which is commercializing the Rowan pK_a workflow disclosed in this paper.

REFERENCES

(1) Di, L.; Kerns, E. H. *Drug-Like Properties: Concepts, Structure Design, and Methods*, 2nd ed.; Academic Press, 2016.

(2) Charifson, P. S.; Walters, W. P. Acidic and Basic Drugs in Medicinal Chemistry: A Perspective. *J. Med. Chem.* **2014**, *57*, 9701–9717.

(3) Garrido, A. et al. hERG toxicity assessment: Useful guidelines for drug design. *Eur. J. Med. Chem.* **2020**, *195*, 112290.

(4) Ploeman, J.-P. H.T.M. et al. Use of physicochemical calculation of pK_a and CLogP to predict phospholipidosis-inducing potential: A case study with structurally related piperazines. *Exp. Toxicol. Pathol.* **2004**, *55*, 347–355.

(5) Pelletier, D. J. et al. Evaluation of a Published in Silico Model and Construction of a Novel Bayesian Model for Predicting Phospholipidosis Inducing Potential. *J. Chem. Inf. Model.* **2007**, *47*, 1196–1205.

(6) Manallack, D. T. The pK_a Distribution of Drugs: Application to Drug Discovery. *Perspect. Medicin. Chem.* **2007**, *1*, 25–38.

(7) Meloun, M.; Bordovská, S. Benchmarking and validating algorithms that estimate pK_a values of drugs based on their molecular structures. *Anal. Bioanal. Chem.* **2007**, *389*, 1267–1281.

(8) Liao, C.; Nicklaus, M. C. Comparison of Nine Programs Predicting pK_a Values of Pharmaceutical Substances. *J. Chem. Inf. Model.* **2009**, *49*, 2801–2812.

(9) Settimo, L.; Bellman, K.; Knegtel, R. M. A. Comparison of the Accuracy of Experimental and Predicted pK_a Values of Basic and Acidic Compounds. *Pharm. Res.* **2013**, *31*, 1082–1095.

(10) For state-of-the-art QSPR methods, see: (a) Johnston, R. C. et al. Epik: pK_a and Protonation State Prediction through Machine Learning. *ChemRxiv*, **2023**, 10.26434/chemrxiv-2023-c6z8t-v3. (b) Mayr, F.; Wieder, M.; Wieder, O.; Langer, T. Improving Small Molecule pK_a Prediction Using Transfer Learning With Graph Neural Networks. *Front. Chem.* **2022**, *10*, 866585. (c) Luo, W. et al. Uni-pK_a: An Accurate and Physically Consistent pK_a Prediction through Protonation Ensemble Modeling. *ChemRxiv*, **2023**, 10.26434/chemrxiv-2023-lw5k0.

(11) Rupp, M.; Körner, R.; Tetko, I. V. Predicting the pK_a of Small Molecules. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 307–327.

(12) Navo, C. D.; Jiménez-Osés, G. Computer Prediction of pK_a Values in Small Molecules and Proteins. *ACS Med. Chem. Lett.* **2021**, *12*, 1624–1628.

(13) Işık, M. et al. Overview of the SAMPL6 pK_a Challenge: Evaluating small molecule microscopic and macroscopic pK_a predictions. *J. Comput. Aided Mol. Des.* **2021**, *35*, 131–166.

(14) Bergazin, T. D. et al. Evaluation of log *P*, pK_a, and log *D* predictions from the SAMPL7 blind challenge. *J. Comput. Aided Mol. Des.* **2021**, *35*, 771–802.

(15) Pracht, P. et al. High accuracy quantum-chemistry-based calculation and blind prediction of macroscopic pK_a values in the context of the SAMPL6 challenge. *J. Comput. Aided Mol. Des.* **2018**, *32*, 1139–1149.

(16) Based on current AWS EC2 on-demand pricing (Feb. 26, 2024) a c5.4xlarge instance with 16 vCPUs costs $0.68/hour: $0.68/hour * 28/16 * 24 hours = $28.56.

(17) Kromann, J. C.; Larsen, F.; Moustafa, H.; Jensen, J. H. Prediction of pK_a values using the PM6 semiempirical method. *PeerJ*, **2016**, *4*, e2335.

(18) Pracht, P.; Grimme, S. Efficient Quantum-Chemical Calculations of Acid Dissociation Constants from Free-Energy Relationships. *J. Phys. Chem. A.* **2021**, *125*, 5681–5692.

(19) Alongi, K. S.; Shields, G. C. Theoretical Calculations of Acid Dissociation Constants: A Review Article. In *Annual Reports in Computational Chemistry*,; Wheeler, R. A., Ed.; Elsevier, 2010; Vol. 6, 113–138.

(20) Dékány, A. A.; Czakó, G. Benchmark ab initio proton affinity and gas-phase basicity of α-alanine based on coupled-cluster theory and statistical mechanics. *J. Comput. Chem.* **2021**, *43*, 19–28.

(21) See, *inter alia*, (a) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121*, 10037–10072. (b) Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.* **2021**, *20*, 750–761. (c) Anstine, D. M.; Isayev, O. Machine Learning Interatomic Potentials and Long-Range Physics. *J. Phys. Chem. A* **2023**, *127*, 2417–2431.

(22) Anstine, D. M.; Zubatyuk, R.; Isayev, O. AIMNet2: A Neural Network Potential to Meet your Neutral, Charged, Organic, and Elemental-Organic Needs. *ChemRxiv,* **2023**, 10.26434/chemrxiv-2023-296ch.

(23) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, *55*, 2562–2574.

(24) RDKit's ETKDG algorithm consistently performs well in benchmarks, particularly when speed is at a premium: see (a) Friedrich, N.-O. et al. Benchmarking Commercial Conformer Ensemble Generators. *J. Chem. Inf. Model.* **2017**, *57*, 2719–2728. (b) Ebejer, J.-P.; Morris, G. M.; Deane, C. M. Freely Available Conformer Generation Methods: How Good Are They? *J. Chem. Inf. Model.* **2012**, *52*, 1146–1158.

(25) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB-An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(26) Ochterski, J. W. Thermochemistry in Gaussian. *Gaussian*, 2022. https://gaussian.com/thermo/ (accessed Mar. 6, 2024).

(27) Stahn, M.; Ehlert, S.; Grimme, S. Extended Conductor-like Polarizable Continuum Solvation Model (CPCM-X) for Semiempirical Methods. *J. Phys. Chem. A*, **2023**, *127*, 7036–7043.

(28) Bochevarov, A. D.; Watson, M. A.; Greenwood, J. R.; Philipp, D. M. Multiconformation, Density Functional Theory-Based pK_a Prediction in Application to Large, Flexible Organic Molecules with Diverse Functional Groups. *J. Chem. Theory Comput.* **2016**, *12*, 6001–6019.

(29) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(30) Walters, P. We Need Better Benchmarks for Machine Learning in Drug Discovery. *Practical Cheminformatics*, 2023. https://practicalcheminformatics.blogspot.com/2023/08/we-need-better-benchmarks-for-machine.html (accessed Mar. 6, 2024).

(31) Rich, A.; Birnbaum, B. Get with the program. *Inductive Bio*, 2023. https://www.inductive.bio/blog/building-better-benchmarks-for-adme-optimization (accessed Mar. 6, 2024).

(32) Direct comparison to SAMPL6 is complicated by the sophisticated statistical procedures conducted by the contest organizers (and by the fact that we excludeδ doubly ionized microstates from our analysis, making our SAMPL6 dataset subtly different from that studied by other methods), but at a high level the error metrics for Rowan pK_a appear very comparable to those from the commercial programs mentioned in the text.

(33) Wuitschik, G. et al. Oxetanes in Drug Discovery: Structural and Synthetic Insights. *J. Med. Chem.* **2010**, *53*, 3227–3246.

(34) Smyrnov, O. et al. α-CF$_3$-Substituted Saturated Bicyclic Amines: Advanced Building Blocks for Medicinal Chemistry. *Eur. J. Org. Chem.* **2024**, *27*, e202300935.

(35) Lõkov, M. et al. On the Basicity of Conjugated Nitrogen Heterocycles in Different Media. *Eur. J. Org. Chem.* **2017**, *30*, 4475–4489.

(36) Miller, G. H.; Doukas, P. H.; Seydel, J. K. Sulfonamide structure-activity relationship in a cell-free system. Correlation of inhibition of folate synthesis with antibacterial activity and physicochemical parameters. *J. Med. Chem.* **1972**, *15*, 700–706.

(37) Oehlrich, D.; Prokopcova, H.; Gijsen, H. J. M. The evolution of amidine-based brain penetrant BACE1 inhibitors. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2033–2045.

(38) Rombouts, F. J. R. et al. Modulating physicochemical properties of tetrahydropyridine-2-amine BACE1 inhibitors with electron-withdrawing groups: A systematic study. *Eur. J. Med. Chem.* **2022**, *228*, 114028.

(39) Morgenthaler, M. et al. Predicting and Tuning Physicochemical Properties in Lead Optimization: Amine Basicities. *ChemMedChem*, **2007**, *2*, 1100–1115.

(40) Appavoo, S. D. et al. Development of Endocyclic Control Elements for Peptide Macrocycles. *J. Am. Chem. Soc.* **2018**, *140*, 8763–8770.

(41) Ertl, P. Database of 4 Million Medicinal Chemistry-Relevant Ring Systems. *J. Chem. Inf. Model.* **2024**, *64*, 1245–1250.

(42) These results contrast with Bochevarov and co-workers' finding (ref. 28) that decent results can be obtained from only a single conformer—the rapid workflows here may prune conformers too aggressively and often fail to find the global minimum. While electronic structure theory-based approaches can easily spend several minutes refining conformers without substantially impacting the overall runtime, the efficiency of AIMNet2 makes fast conformational searching crucial for speeding pK_a prediction.

(43) Implicit solvation has been demonstrated to be a leading source of error for conventional electronic structure theory-based pK_a prediction. See, *inter alia*, (a) Thapa, B.; Schlegel, H. B. Density Functional Theory Calculation of pK_a's of Thiols in Aqueous Solution Using Explicit Water Molecules and the Polarizable Continuum Model. *J. Phys. Chem. A* **2016**, *120*, 5726–5735. (b) Thapa, B.; Raghavachari, K. Accurate pK_a Evaluations for Complex Bio-Organic Molecules in Aqueous Media. *J. Chem. Theory Comput.* **2019**, *15*, 6025–6035.

(44) Jing, B. et al. Torsional Diffusion for Molecular Conformer Generation. *arXiv*, **2023**, 10.48550/arXiv:2206.01729.

(45) Liao, Y.-L.; Wood, B.; Das, A.; Smidt, T. EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations. *arXiv*, **2023**, 10.48550/arXiv.2306.12059.

(46) Grimme, S.; Schreiner, P. R. Computational Chemistry: The Fate of Current Methods and Future Challenges. *Angew. Chem. Int. Ed.* **2017**, *57*, 4170–4176.

(47) Cluster–continuum solvation shows promise but is difficult to run in a black-box manner and introduces considerable additional complexity. See (a) Pliego, J. R.; Riveros, J. M. The Cluster−Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species. *J. Phys. Chem. A* **2001**, *105*, 7241–7247. (b) Norjmaa, G.; Ujaque, G.; Lledós, A. Beyond Continuum Solvent Models in Computational Homogeneous Catalysis. *Top. Catal.* **2022**, *65*, 118–140. (c) Spicher, S. et al. Automated Molecular Cluster Growing for Explicit Solvation by Efficient Force Field and Tight Binding Methods. *J. Chem. Theory Comput.* **2022**, *18*, 3174–3189.

(48) Arts, M. et al. Two for One: Diffusion Models and Force Fields for Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.* **2023**, *19*, 6151–6159.

(49) Zhang, J.; Pagotto, J.; Gould, T.; Duignan, T. T. Accurate, fast and generalisable first principles simulation of aqueous lithium chloride. *arXiv*, **2023**, 10.48550/arXiv.2310.12535.