High Throughput Pharmacokinetic Modeling Using Computationally Predicted Parameter Values: Dissociation Constants

Cory L. Strope[1], Kamel Mansouri[2], Jayaram Kancherla[2], Caroline Stevens[3], John F. Wambaugh[4]

[1] The Hamner Institutes for Health Sciences, [2] ORISE Fellow at National Center for Computational Toxicology, EPA, [3] Ecosystems Research Division, National Exposure Research Laboratory, EPA, [4] National Center for Computational Toxicology, EPA

Abstract.

Estimates of the ionization association and dissociation constant (pKa) are vital to modeling the pharmacokinetic behavior of chemicals *in vivo*. Methodologies for the prediction of compound sequestration in specific tissues using partition coefficients requires appropriate characterization of the fraction of ionized molecules and the charge of any ionized state(s). Current methods for reporting the pKa report only the pH at which the pKa-associated atom will be ionized in 50% of the molecules. Important considerations for *in vivo* usage of the pKa such as (i) the chemical class (i.e., acid/base), and (ii) the interplay between ionization states at other atoms to determine the fraction of a chemical to exist in a particular ionization state, are reduced to "missing" information status. We propose a new method that more fully describes the process associated with reporting pKa values. Further, this new format is designed to support high-throughput applications. We compare the ionizable atom types between 8195 pharmaceutical and 24388 environmental compounds, and investigate the performance of several publically and commercially available pKa predictive models on these 32583 chemicals from the human exposome. Finally, the analysis methodology developed herein for efficient estimation of the parameters critical for predicting chemical pharmacokinetics are publicly-available as an R package.

INTRODUCTION

Current pKa prediction software is trained on pharmaceutical data sets. When using environmental data, pKa predictions can be made using the software, but there is little work on determining the validity of such predictions on environmental data. We compare three commercial pKa prediction software, ADMET Predictor, SPARC, and the pKa Plugin to ChemAxon, in order to quantify their predictions on pharmaceutical, near-field environmental, and far-field environmental chemicals.

We define the chemical classes based on their use category as defined by the ACToR database. We set the chemical class based on the use category defined exposure potential, as done in Wambaugh et al. (2013), further setting forth a precedence scheme respecting the intimacy of their exposure. The highest precedence is given to pharmaceuticals, i.e., if a chemical is defined as a pharmaceutical in any use category, we define it as a pharmaceutical, regardless of any further categories tied to the compound. Likewise, if the chemical has a near-field category (fragrance, food additive, consumer use, personal care; see Table 1), and does not have a pharmaceutical category, it is defined as a near-field environmental chemical. All others are defined as far-field.

We characterize the association and dissociation constants between the chemical classes to determine if there is a difference between the chemical classes.

OBJECTIVE

Determine and characterize if there is a marked difference between the ionization profiles of pharmaceutical and environmental chemicals.

If the association/dissociation constants between the pharmaceutical and environmental compounds are indistinguishable, predicting pKas with prediction software is a reasonable assumption for extrapolating the predictions to environmental compounds. If there is a marked difference between the pKa predictions, then caution must be exercised in using pKa prediction software with environmental compounds.

MATERIALS AND METHODS

Table 1: Chemicals in data set based on ACToR Use Categories.

Use Category	Assignment	Counts
Antimicrobial	Far-field	222
Colorant	Far-field	521
Drug	Pharmaceutical	8195
Flame-retardant	Far-field	32
Fragrance	Near-field	340
Industrial-manufacturing	Far-field	4424
Personal care	Near-field	1191
Petrochemical	Far-field	87
Chemical warfare	Far-field	47
Consumer use	Near-field	1624
Fertilizer	Far-field	3
Food additive	Near-field	2810
Herbicide	Far-field	51
Inert	Far-field	1513
Pesticide	Far-field	1966
No category	Far-field	15963
Near-field		4022
Far-field		20366
Pharmaceutical		8195

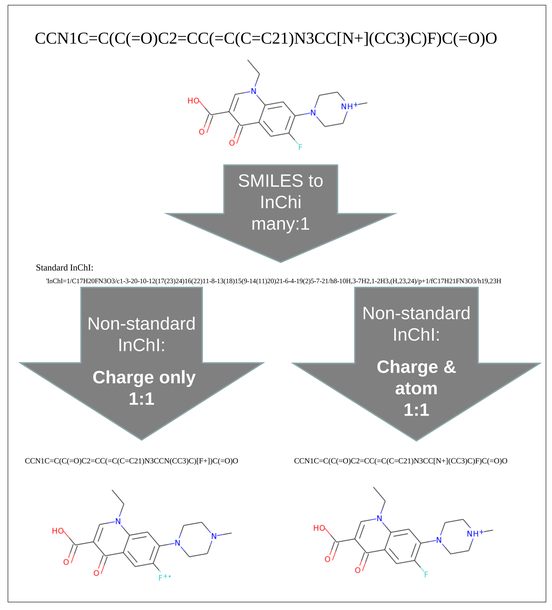

Figure 1: Canonicalization of a non-unique SMILES string of a charged compound through a SMILES→InChI→SMILES conversion routine. Beginning with a charged molecule and its associated, non-unique, SMILES string, we convert the SMILES string into a standard InChI string. (*Left*) The non-standard InChI string included with the InChI package provided by IUPAC. In this format, only the charge of the molecule is shown, and in the InChI→SMILE conversion the charged molecule is inferred by the routine. (*Right*) Our proposed InChI→SMILES conversion routine carries the molecule(s) that are to be charged, reproducing the input molecules SMILES string. This scheme will allow us to make direct comparisons between pKa prediction software packages.

RESULTS

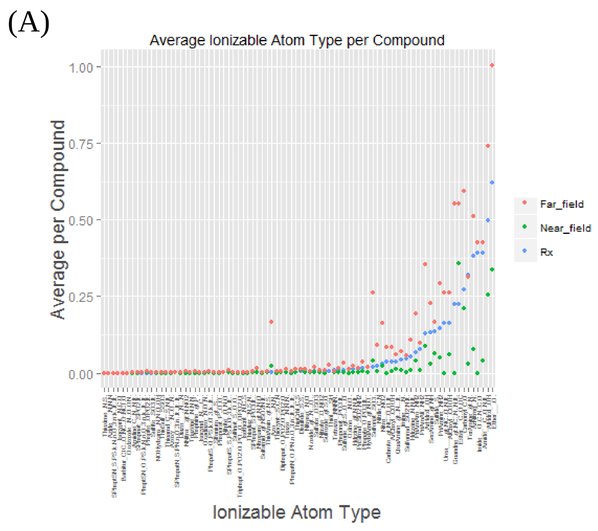

Figure 2: Average ionizable atom types per chemical for (blue) pharmaceutical, (green) near-field environmental, and (red) far field environmental chemicals. **(A)** Ionizable atom types are separated into ~80 different types of ionizable atoms, based on the context in which they appear in the compound (e.g., aliphatic hydroxyl vs. aromatic hydroxyl). **(B)** Ionizable atom types binned into "chemical classes" (e.g., ionizable hydroxyls can be separated into alcohols, ethers, etc)

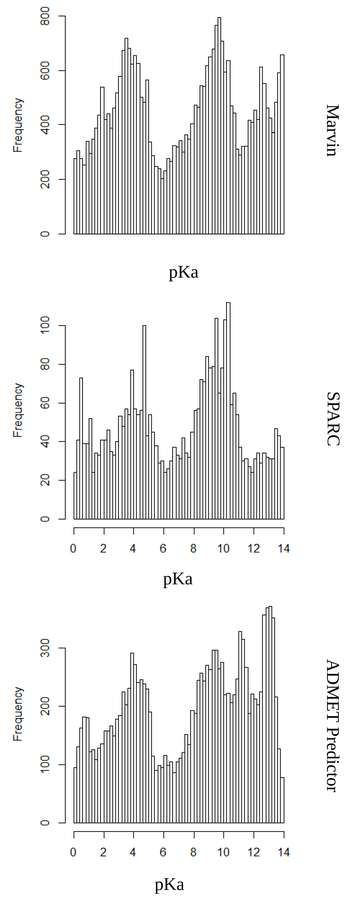

Figure 3: pKa predictions per predictor against the full dataset. (Top) ChemAxon pKa Plugin. (Middle) SPARC. (Bottom) ADMET Predictor version 7. The y-axis scale not normalized to show the promiscuity of predictions for each method. *Note: Due to circumstances, we were only able to run SPARC predictions against an incomplete dataset of only 4000 chemicals.*

REFERENCES

pKa Plugin: J Szegezdi and F Csizmadia. 2007. A method for calculating the pKa values of small and large molecules. In: Conference Abstracts, 233rd National Meeting of the American Chemical Society. Chicago, IL: American Chemical Society.

ADMET Predictor: R Fraczkiewicz. 2013. In silico prediction of ionization. Reference Module in Chemistry, Molecular Sciences and Chemical Engineering 5:603-626.

SPARC: P Lee, S Ayyampalayam, L Carreira, M Shalaeva, S Bhattachar, et. al. 2007. In silico prediction of ionization constants of drugs. Molecular Pharmaceutics 4:198—512.

ACToR Use Categories: K Dionisio, A Frame, M-R Goldsmith, JF Wambaugh, A Liddell, T Cathey, D Smith, J Vail, and RS Judson. In prep. Exploring exposure pathways with Chemical/Product Categorical (CPCat) Data.

CONCLUSIONS

Preliminary results suggest that there are some differences between the ionizations of pharmaceutical and environmental chemicals (Figure 2). We are currently running analyses to determine the impacts of these differences when applied to downstream analyses, e.g., pharmacokinetic partition coefficient inference. The pKa prediction methods used in this analysis appear to have a similar distribution of predictions on the data set. It is interesting to note that pKa Plugin makes a large number of predictions compared to ADMET Predictor, despite their similar distribution of predictions.

Future analyses are underway to compare the pKa predictions per ionization atom to further understand the difference between pKa prediction methods. We also are examining the distribution of pKa values predicted for each of the ionizable atom types to see if the predicted value for each atom type can be described by a distribution. Finally, future work is planned to incorporate the pKa prediction routine into partition coefficient inference methods for physiologically-based pharmacokinetic models.