Developing Predictive ML Models to Inform TPD Lead Optimization

Paul Novick, Mridula Bontha

TPD Europe

London

March 30, 2023

Important Notice and Disclaimers

This presentation contains statements that relate to future events and expectations and as such constitute forward-looking statements within the meaning of the Private Securities Litigation Reform Act of 1995. When or if used in this presentation, the words "anticipate," "believe," "could," "estimate," "expect," "intend," "may," "outlook," "plan," "predict," "should," "will," and similar expressions and their variants, as they relate to Nurix Therapeutics, Inc. ("Nurix", the "Company," "we," "us" or "our"), may identify forward-looking statements. All statements that reflect Nurix's expectations, assumptions or projections about the future, other than statements of historical fact, are forward-looking statements, including, without limitation, statements regarding our future financial or business plans; our future performance, prospects and strategies; future conditions, trends, and other financial and business matters; our current and prospective drug candidates; the planned timing and conduct of the clinical trial programs for our drug candidates; the planned timing for the provision of clinical updates and initial findings from our clinical studies; the potential advantages of our DELigase™ platform and drug candidates; the extent to which our scientific approach and DELigase™ platform may potentially address a broad range of diseases; the extent animal model data predicts human efficacy; and the timing and success of the development and commercialization of our current and anticipated drug candidates. Forward-looking statements reflect Nurix's current beliefs, expectations, and assumptions. Although Nurix believes the expectations and assumptions reflected in such forward-looking statements are reasonable, Nurix can give no assurance that they will prove to be correct. Forward-looking statements are not guarantees of future performance and are subject to risks, uncertainties and changes in circumstances that are difficult to predict, which could cause Nurix's actual activities and results to differ materially from those expressed in any forward-looking statement. Such risks and uncertainties include, but are not limited to: (i) risks and uncertainties related to Nurix's ability to advance its drug candidates, obtain regulatory approval of and ultimately commercialize its drug candidates; (ii) the timing and results of clinical trials; (iii) Nurix's ability to fund development activities and achieve development goals; (iv) the impact of macroeconomic conditions, including inflation, increasing interest rates and volatile market conditions, and global events, including the COVID-19 pandemic, on Nurix's clinical trials and operations; (v) Nurix's ability to protect intellectual property and (vi) other risks and uncertainties described under the heading "Risk Factors" in Nurix's Annual Report on Form 10-K for the fiscal year ended November 30, 2022, and other SEC filings. Accordingly, readers are cautioned not to place undue reliance on these forward-looking statements. The statements in this presentation speak only as of the date of this presentation, even if subsequently made available by Nurix on its website or otherwise. Nurix disclaims any intention or obligation to update publicly any forward-looking statements, whether in response to new information, future events, or otherwise, except as required by applicable law.

Certain information contained in this presentation relates to or is based on studies, publications, surveys and other data obtained from third-party sources and the Company's own internal estimates and research. While the Company believes these third-party sources to be reliable as of the date of this presentation, it has not independently verified, and makes no representation as to the adequacy, fairness, accuracy or completeness of, any information obtained from third-party sources. In addition, all of the market data included in this presentation involves a number of assumptions and limitations, and there can be no guarantee as to the accuracy or reliability of such assumptions. Finally, while we believe our own internal estimates and research are reliable, such estimates and research have not been verified by any independent source.

A Very Cool Place to Work!

nurix

About Platform Pipeline Scientific Resources Investors Careers & Culture

CAREERS & CULTURE

Blaze Your Trail With Us

We are looking for trailblazers who embody Nurix's values to join us to blaze a new trail in medicine. Nurix is a clinical-stage, science-focused company committed to bringing novel, first-in-class therapies to patients. We are leaders in discovering and developing a new class of drugs that modulate protein levels by harnessing or inhibiting a class of enzymes called E3 ubiquitin ligases that have a key role in the natural process of protein turnover.

Note to Employment Agencies: Please do not forward any agency resumes. The company will not be responsible for any fees related to resumes that are unsolicited.

nurix

The Nurix Machine Learning Team

Mridula Bontha
Primary ML Scientist on the project

Paul Novick
Director Cheminformatics and Machine Learning

Elena Caceres

Heta Gandhi

Jen Bone

Talk Outline

Introduction	Evaluation of existing tools in predicting molecular properties for Targeted Protein Degradation (TPD) molecules
Methods/Results	Performance and generalizability of Nurix Machine Learning (ML) models for Solubility and Permeability
Discussion	Interplay between model learnings and human interpretability

nurix

Introduction

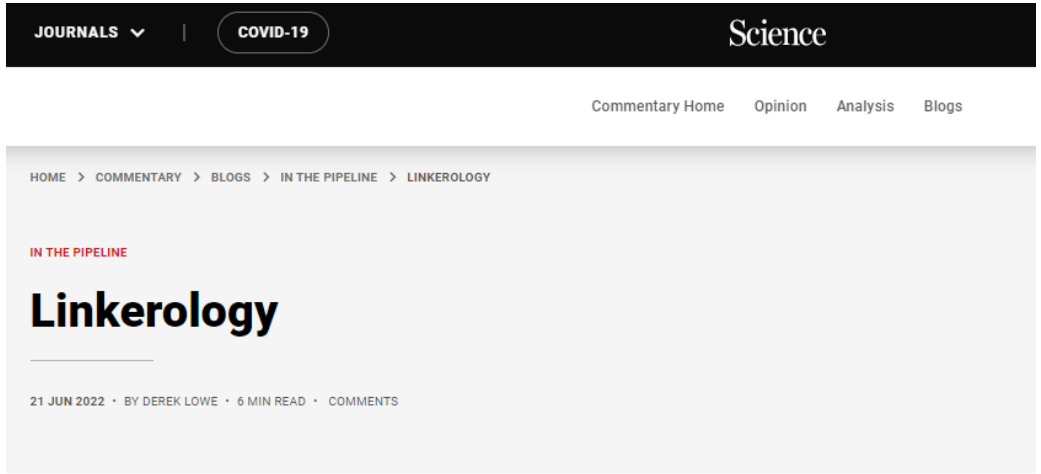
Evaluation of existing tools in predicting molecular properties for TPD molecules

nurix

Degrader Optimization Remains Largely Empirical

...**We're still in the "try 'em and see" stage of optimization**. ... It would be much, much nicer if we could stand at the whiteboard or look at a screen, pursing our lips thoughtfully and then pointing purposefully at The Compound To Make, but we ain't there yet.

Derek Lowe
In The Pipeline
6/21/22



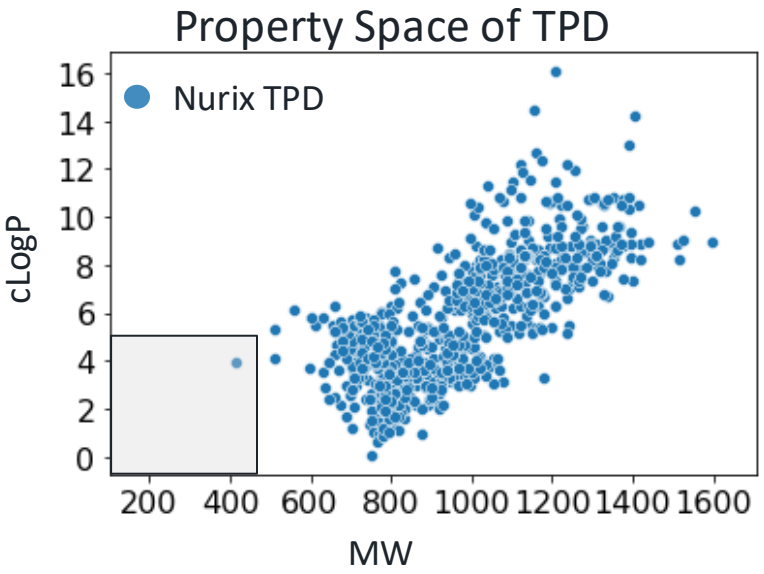
I've written a number of times here about bifunctional protein degraders, which have been a big topic in drug discovery for the past few years. There's a new paper that illustrates some of the challenges in this area, and it's worth using as an example.

For those outside the field, the idea behind these things is pretty straighforward, at least in principle. You find a protein that you think is involved in a disease process, one whose activity you would like to dial down. You find a small-molecule ligand that binds that protein - you may already have some inhibitors around, in house or from the literature, and for these purposes your small molecule doesn't even have to be an inhibitor, just a binder. (Of course, the way we run assays means that most of the time we're not set up to detect silent binders, so those are thinner on the ground). Now you break out your synthetic organic chemistry skills and build out a linker group from that known ligand, and at the other end of that linker you attach a known ligand for an "E3 ligase" enzyme. There are several possibilities, but so far the well-established ones for the enzymes cereblon and VHL are the ones that get used the most, by far.

Degrader Molecules Are Outside the Domain of Applicability for Commonly Applied Intuition and Rules

Targeted Protein Degrader molecules occupy a property space well beyond traditional Lipinski RO5 molecules.

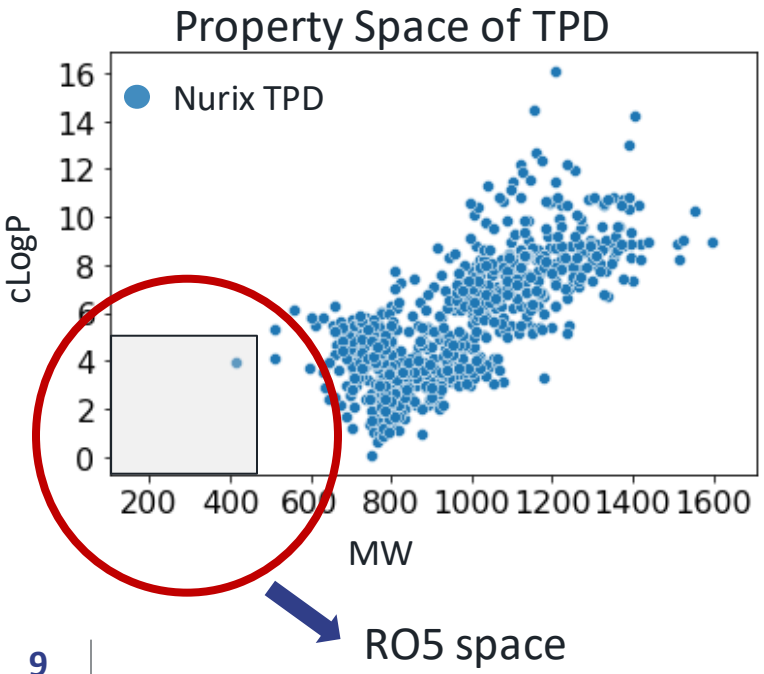
As a result, intuition and rules for predicting physicochemical properties developed based on RO5 small molecules do not generalize well when applied to TPD molecules.



Degrader Molecules Are Outside the Domain of Applicability for Commonly Applied Intuition and Rules

Targeted Protein Degrader molecules occupy a property space well beyond traditional Lipinski RO5 molecules.

As a result, intuition and rules for predicting physicochemical properties developed based on RO5 small molecules do not generalize well when applied to TPD molecules.

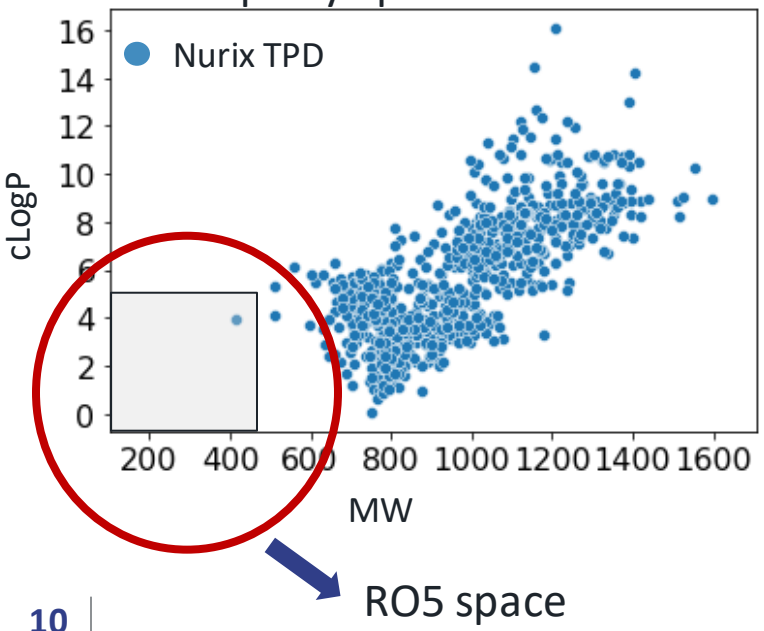


Degrader Molecules Are Outside the Domain of Applicability for Commonly Applied Intuition and Rules

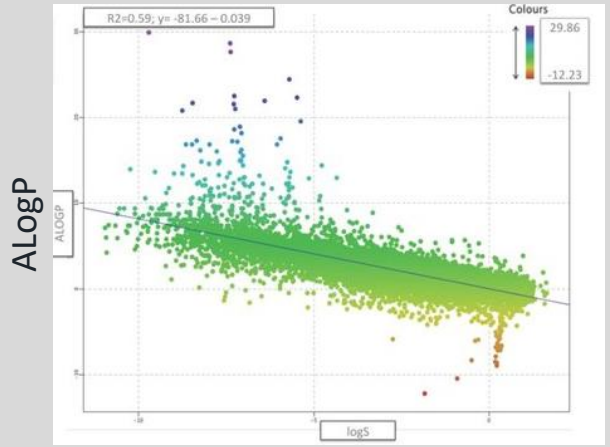
Targeted Protein Degrader molecules occupy a property space well beyond traditional Lipinski RO5 molecules.

As a result, intuition and rules for predicting physicochemical properties developed based on RO5 small molecules do not generalize well when applied to TPD molecules.

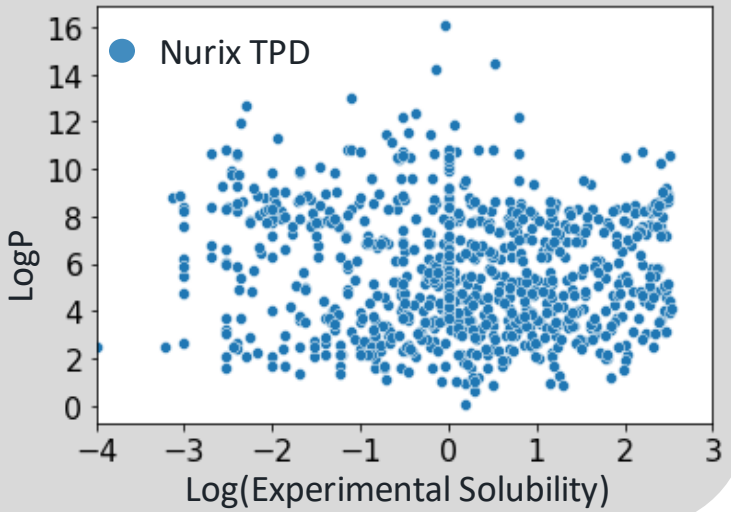
Property Space of TPD



Large Public Dataset*



Nurix TPD Ksol Data



Calculated LogP vs Experimental Solubility for small molecules (left) and Nurix TPD (right)

*Falcon-Cano, G., et al., ADMET & DMPK, 2020

Poor Correlation Between Ksol and 1D Molecular Properties

Extended view of weak correlation between molecular descriptors and solubility.

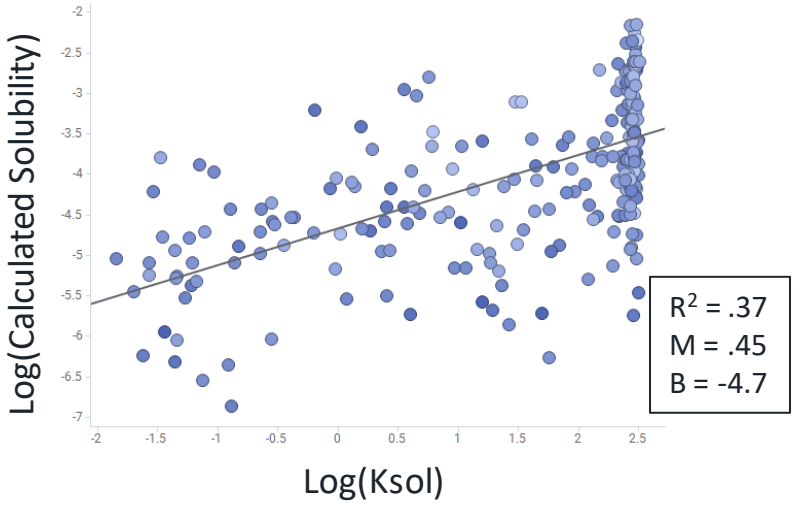
	R-squared
Kinetic Solubility X	1.000000
MolWt	0.117218
LogP	0.130907
NumHAcceptors	0.047313
NumHDonors	0.022847
NumHeteroatoms	0.079500
NumRotatableBonds	0.009230
NumHeavyAtoms	0.117752
NumAliphaticCarbocycles	0.016792
NumAliphaticHeterocycles	0.062368
NumAliphaticRings	0.085996
NumAromaticCarbocycles	0.032745
NumAromaticHeterocycles	0.044656
NumAromaticRings	0.132261
RingCount	0.185393
FractionCSP3	0.002377
TPSA	0.056036

A more thorough evaluation of 1D properties commonly thought to affect or relate to the solubility of small molecules shows very low correlation to experimental data.

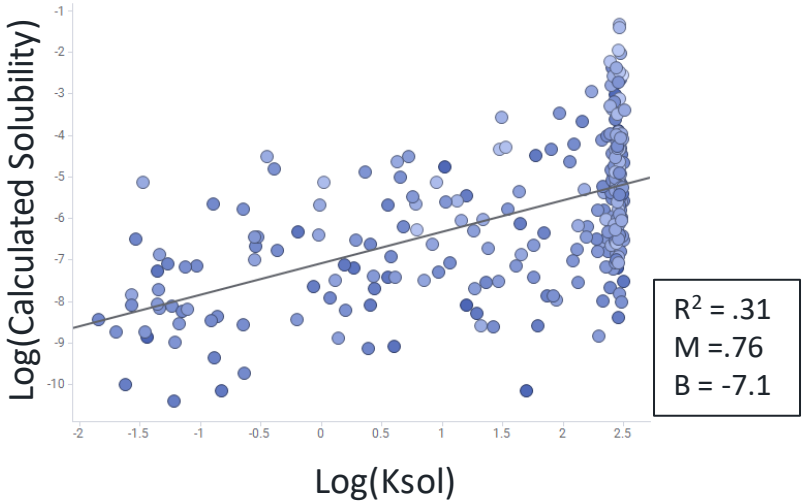
nurix

More Sophisticated Cheminformatics Calculators Similarly Struggle to Generalize to Degraders

Well Used Comp Chem Software #1

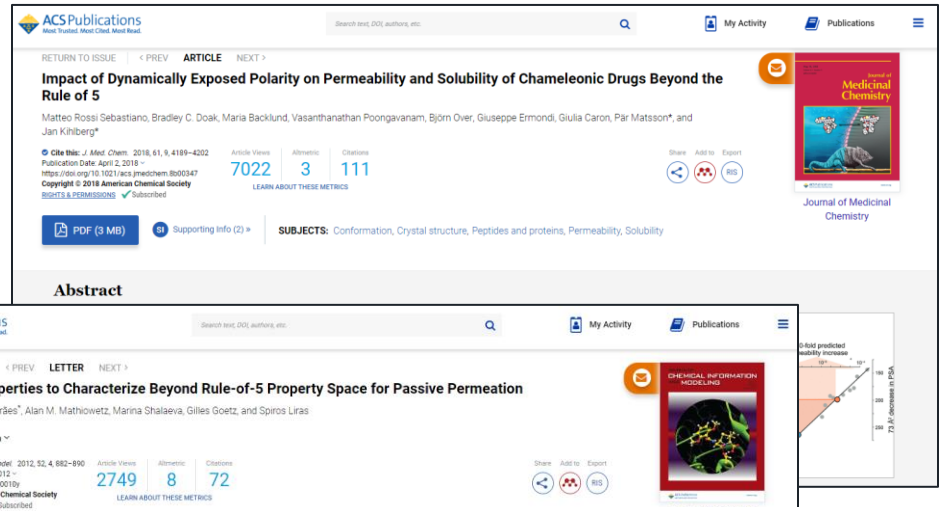


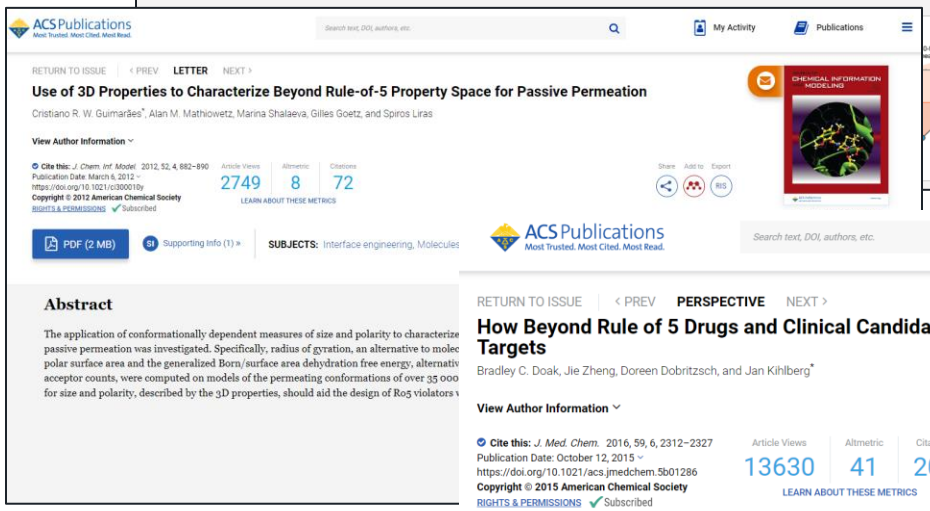
Well Used Comp Chem Software #2

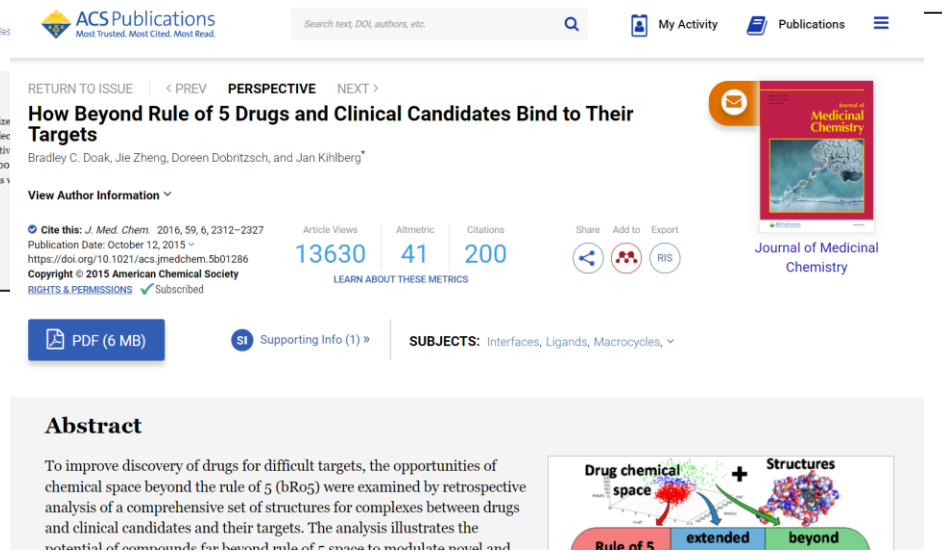


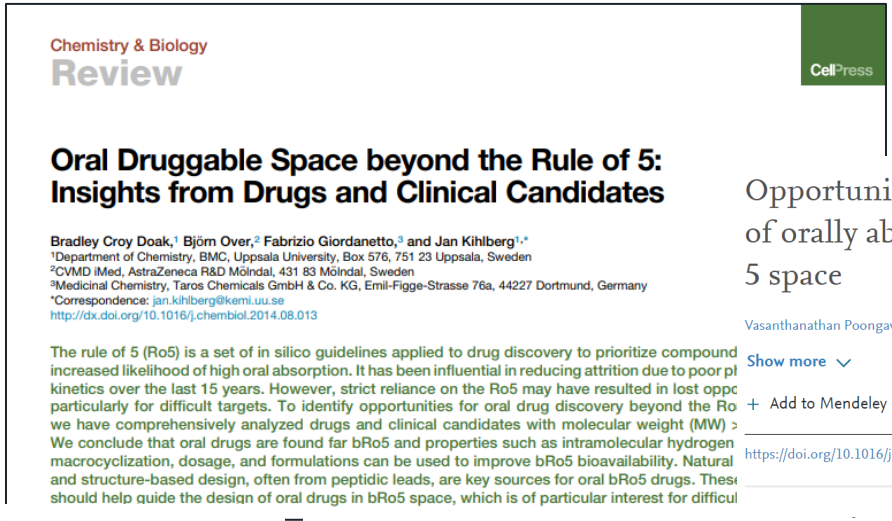
Popular CompChem and Cheminformatics software struggle to accurately predict relevant properties of Degrader compounds.

Active Literature Addressing Property Prediction in BRO5 Space

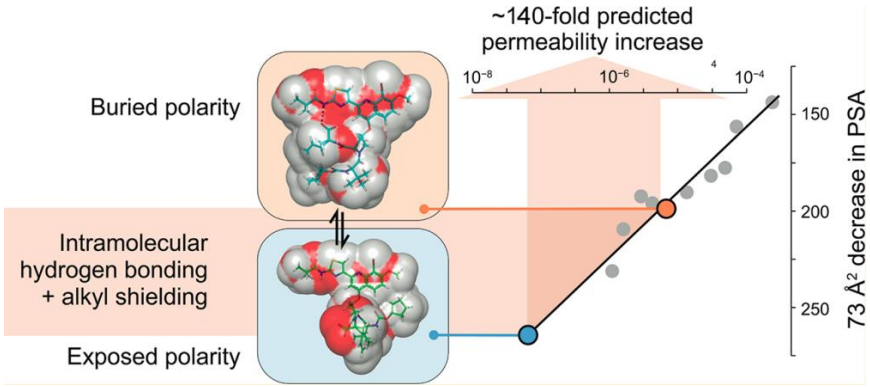






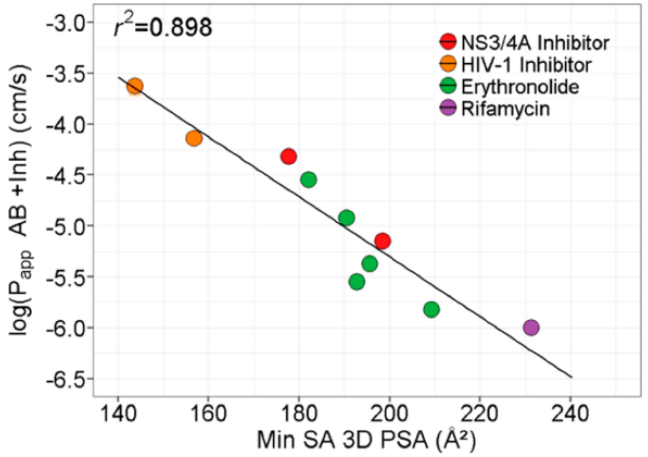


Literature BRO5 Models Struggle to Predict Nurix Internal Degrader Experimental Data



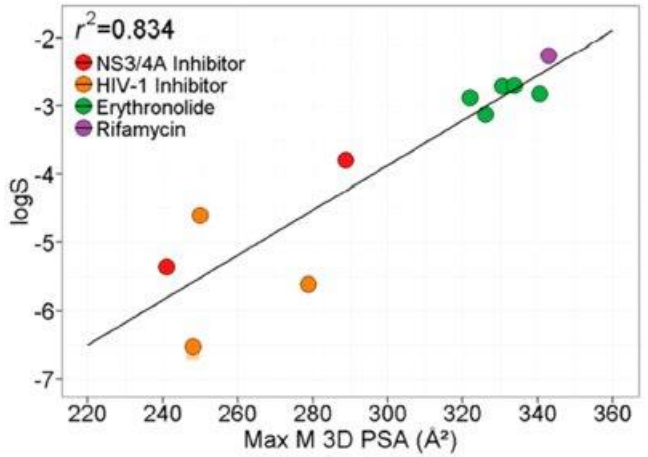
Impact of Dynamically Exposed Polarity on Permeability and Solubility of Chameleonic Drugs Beyond the Rule of 5

Sebastiano, et al., J Med Chem, 2018

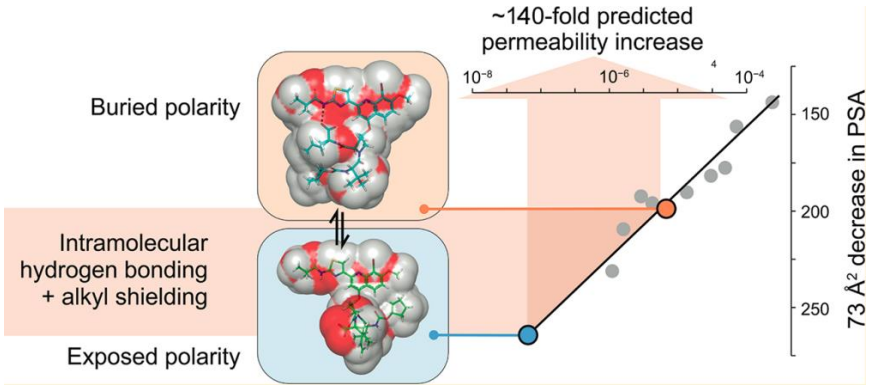


Publication Data

Literature model for predicting permeability and solubility of BRO5 compounds as a function of 3D polar surface area struggle to track with Nurix experimental data

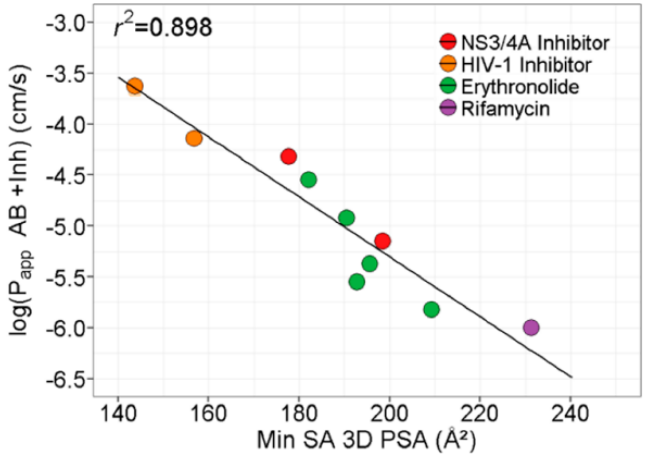


Literature BRO5 Models Struggle to Predict Nurix Internal Degrader Experimental Data

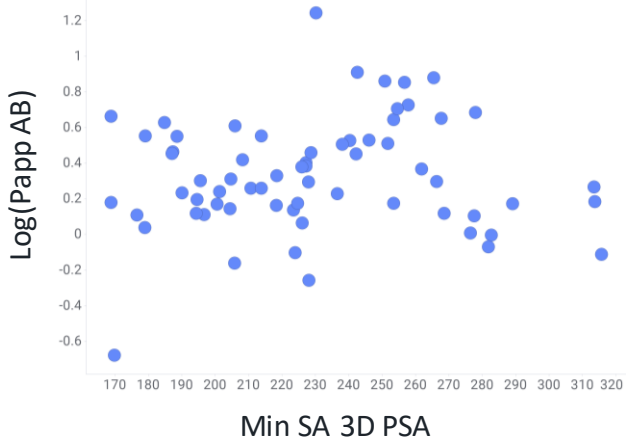


Impact of Dynamically Exposed Polarity on Permeability and Solubility of Chameleonic Drugs Beyond the Rule of 5

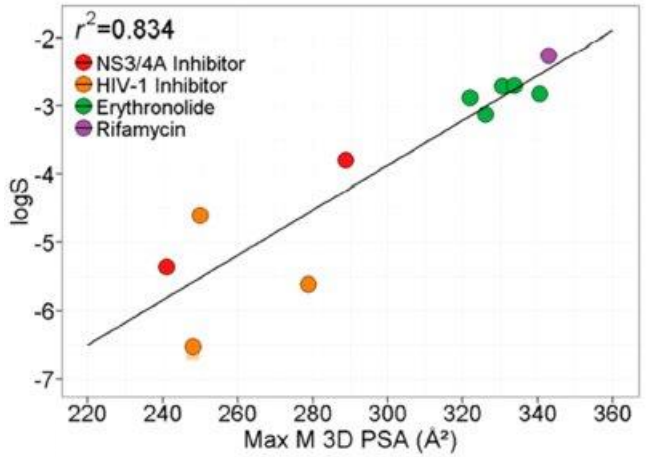
Sebastiano, et al., J Med Chem, 2018

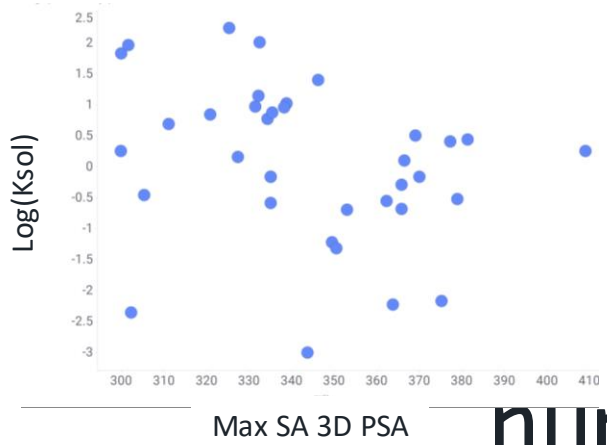


Publication Data



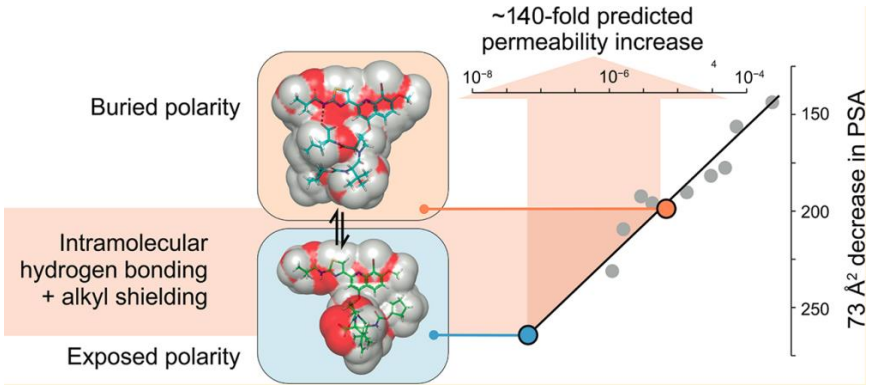
Nurix Data





Literature model for predicting permeability and solubility of BRO5 compounds as a function of 3D polar surface area struggle to track with Nurix experimental data

Literature BRO5 Models Struggle to Predict Nurix Internal Degrader Experimental Data

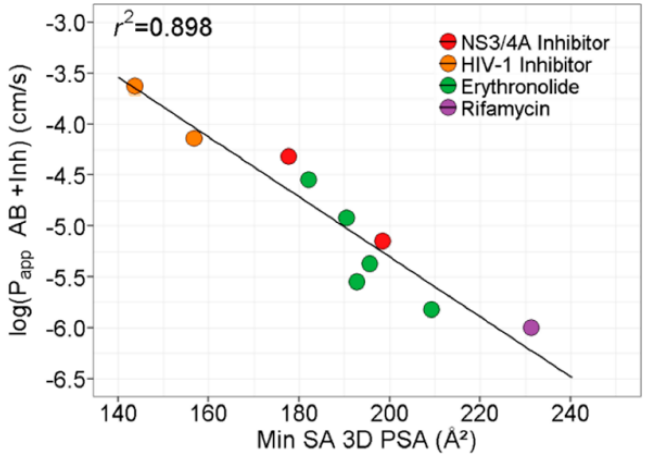


Impact of Dynamically E… on Permeability a… hic Drug…

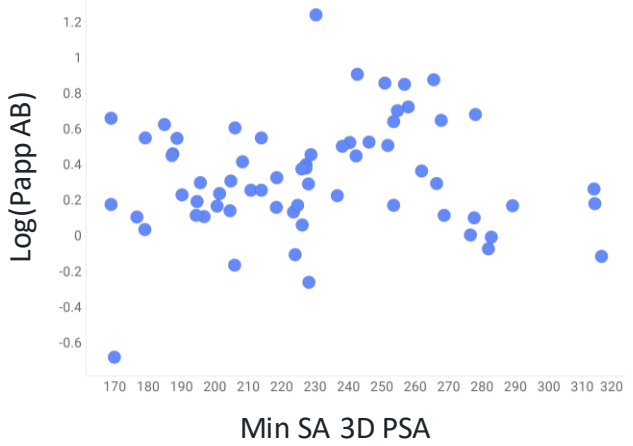
Se… …hem, 2018

Results representative of all published methods tried

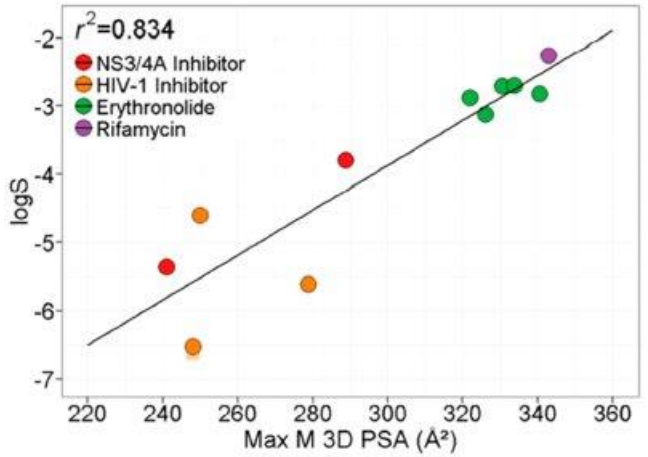
Literature model for predicting permeability and solubility of BRO5 compounds as a function of 3D polar surface area struggle to track with Nurix experimental data

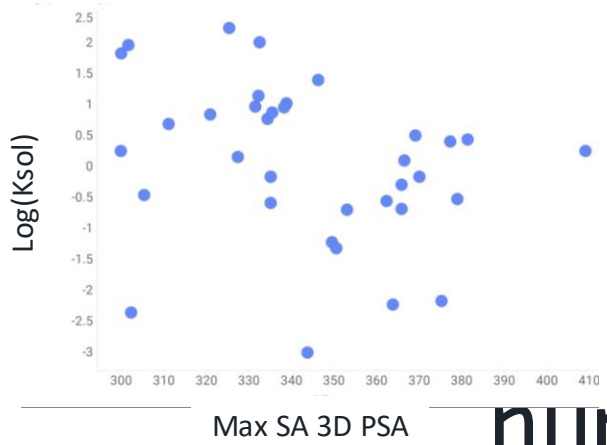


Publication Data



Nurix Data





Moving Away from the Streetlight

1. TPD compounds occupy a completely different region of chemical space than historical drugs

2. Tools and intuition developed for RO5-like compounds largely fail to be predictive for TPDs

Moving Away from the Streetlight

1. TPD compounds occupy a completely different region of chemical space than historical drugs

2. Tools and intuition developed for RO5-like compounds largely fail to be predictive for TPDs

Moving Away from the Streetlight

1. TPD compounds occupy a completely different region of chemical space than historical drugs

2. Tools and intuition developed for RO5-like compounds largely fail to be predictive for TPDs

Machine learning offers the opportunity to re-learn the rules and create models that are actually predictive.

Methods/Results

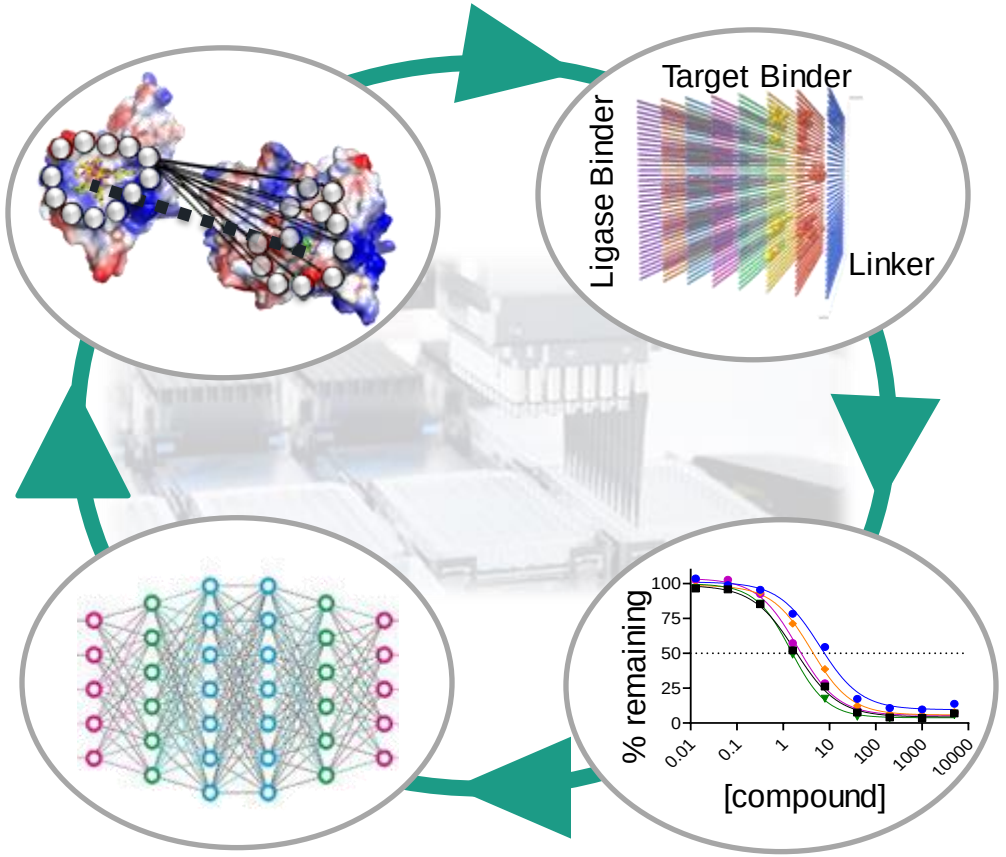
Performance and generalizability of Nurix Machine Learning models for Solubility and Permeability

Nurix Is Applying Automation to Better Define the Parameters of Degrader Design

- Challenge: Identifying and optimizing degraders remains largely an empirical process.
- Nurix platform is scaled to enable evaluation of broad and unique search spaces to identify and optimize degraders.

DESIGN SCOPE

Theoretical range of degrader chemical space more fortuitous than rational



SYNTHESIZE AT SCALE

Automation enables Nurix to sample unprecedented chemical space

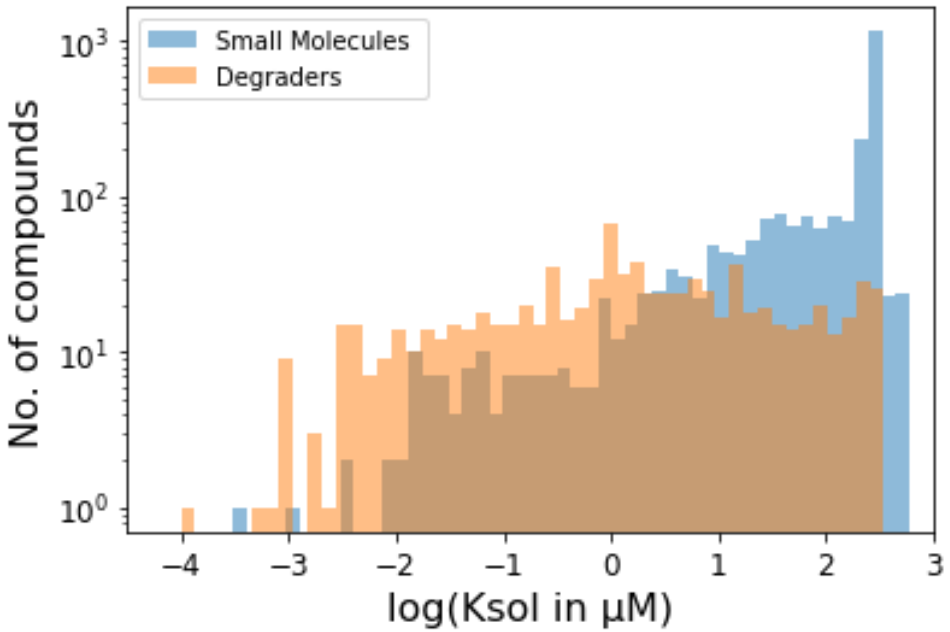
WRITE THE RULEBOOK

Machine Learning transforms large datasets into degrader rulebook for improved design

DISCOVER LEADS

Empirical data reveals degraders with optimal performance

Nurix Kinetic Solubility Dataset and Framing the ML Problem



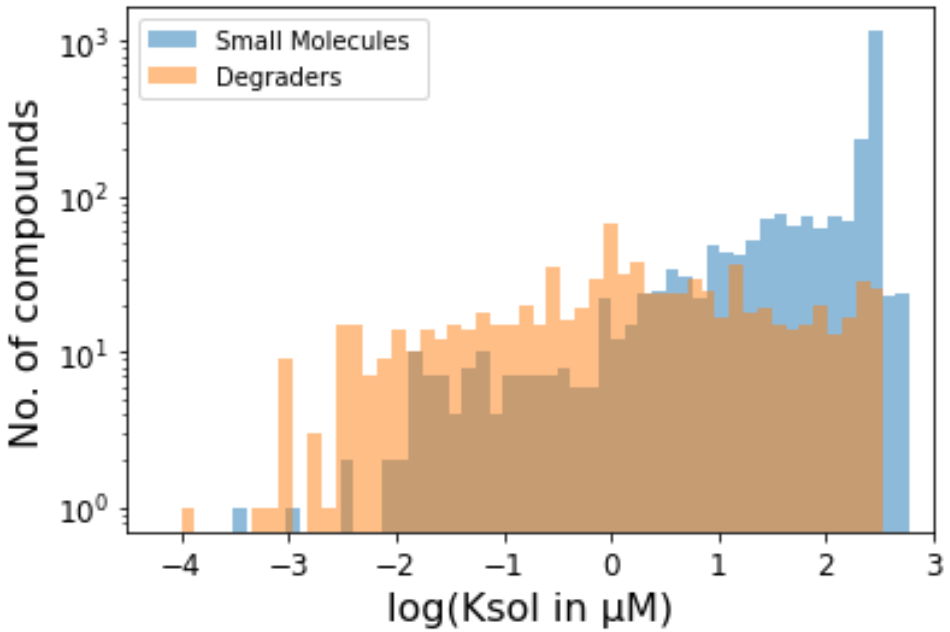
Binary Classification (Soluble/Insoluble)

Multiclass Classification (High/Med/Low)

Regression (Predict numeric Ksol value)

Leveraging the Nurix solubility dataset, we evaluated both Regression and Classification models (and a host of implementations and parameter sets) to identify the best performing models.

Nurix Kinetic Solubility Dataset and Framing the ML Problem



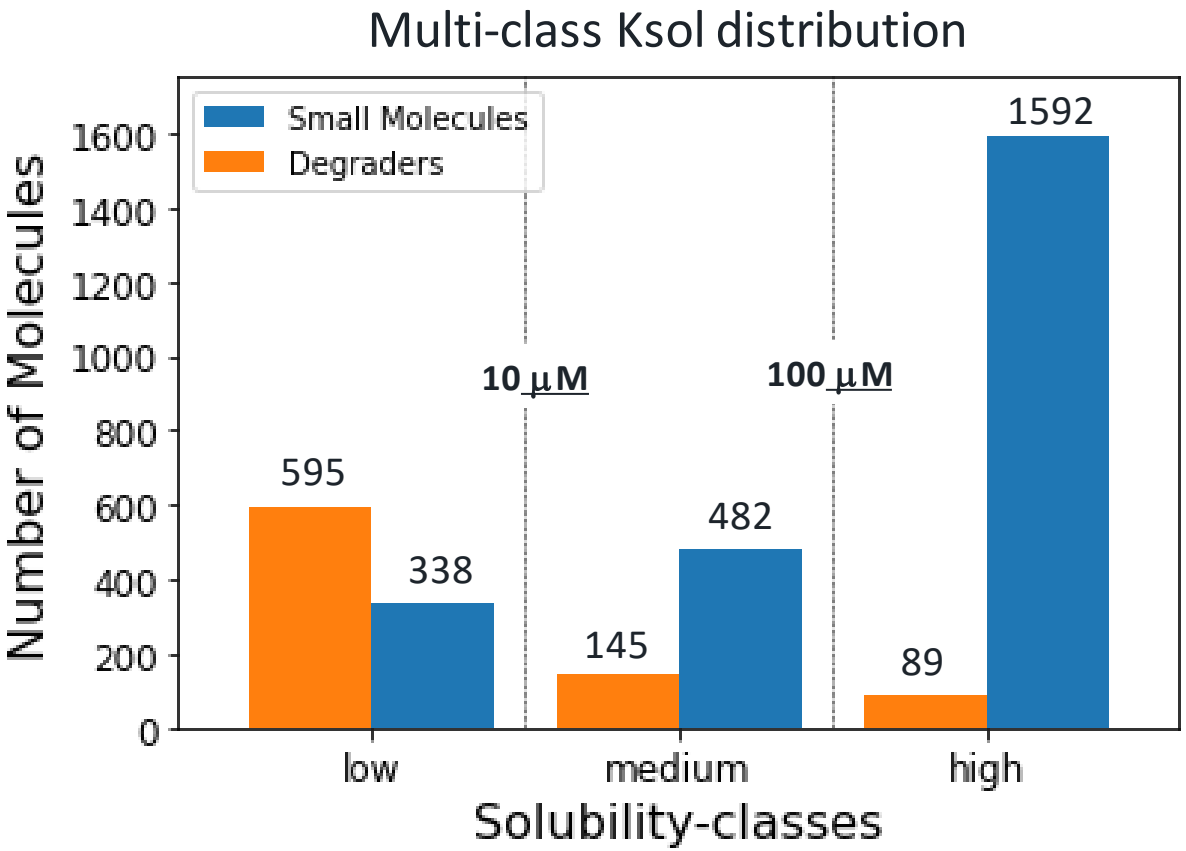
Binary Classification (Soluble/Insoluble)

Multiclass Classification (High/Med/Low)

Regression (Predict numeric Ksol value)

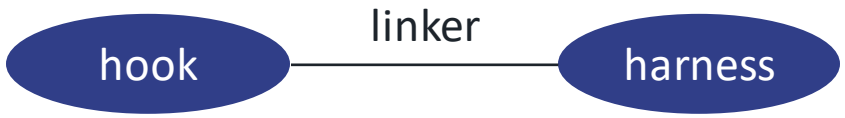
Leveraging the Nurix solubility dataset, we evaluated both Regression and Classification models (and a host of implementations and parameter sets) to identify the best performing models.

Partitioning the Nurix Ksol Dataset to Enable Lead Optimization

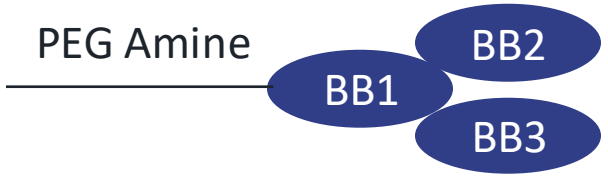


Combined small molecule and degrader dataset to increase generalizability of models

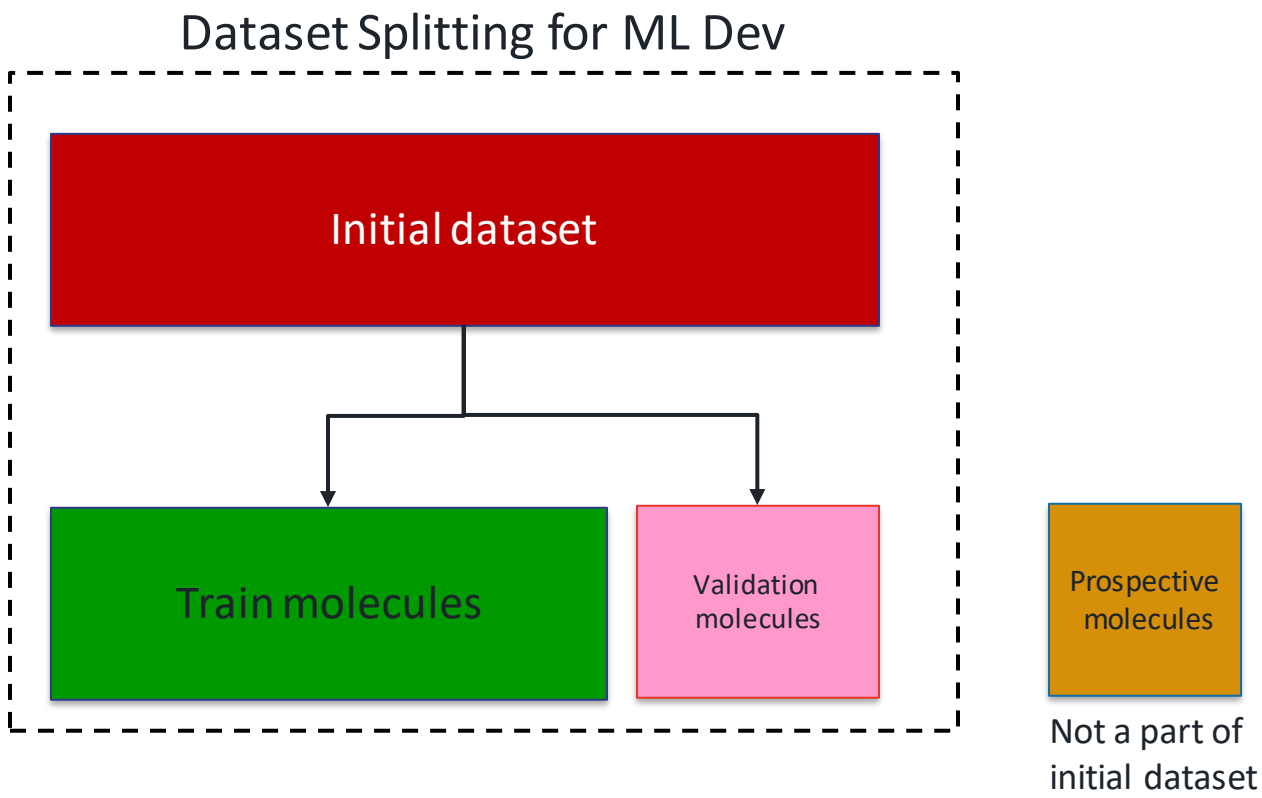
TPD compound anatomy



DEL 'Small Molecule'



Dataset Hygiene and Partitioning for the Multiclass ML Model

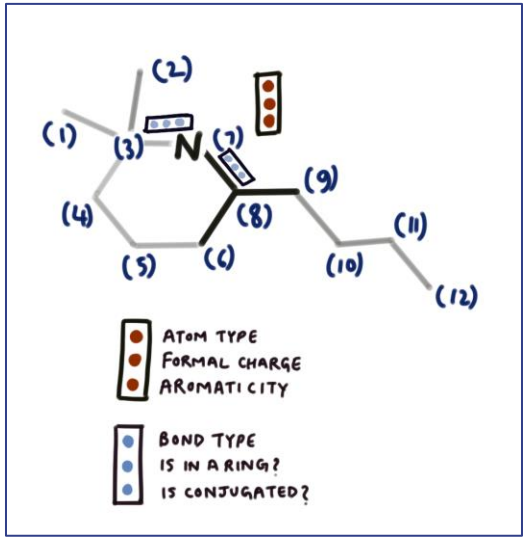


Proper dataset hygiene to ensure model is performant in prospective predictions

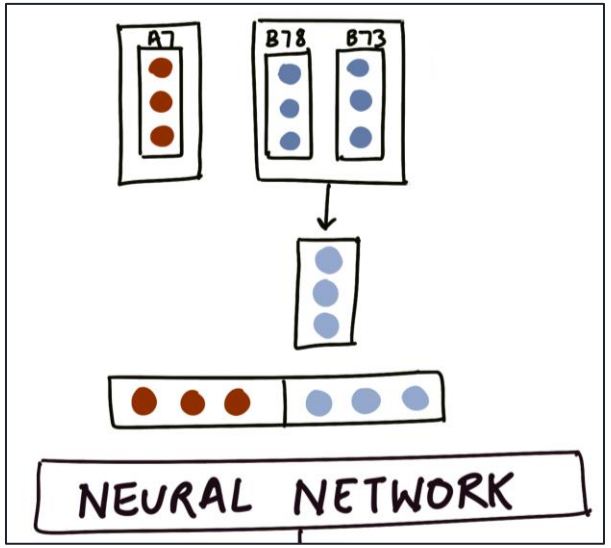
Model Overview of Best Performing Ksol ML Model

Input: SMILES

C(#N)C1=Cc2n(c(-c3ncc(C(=O)....



Feed model with embeddings



Predict class

low

medium

high

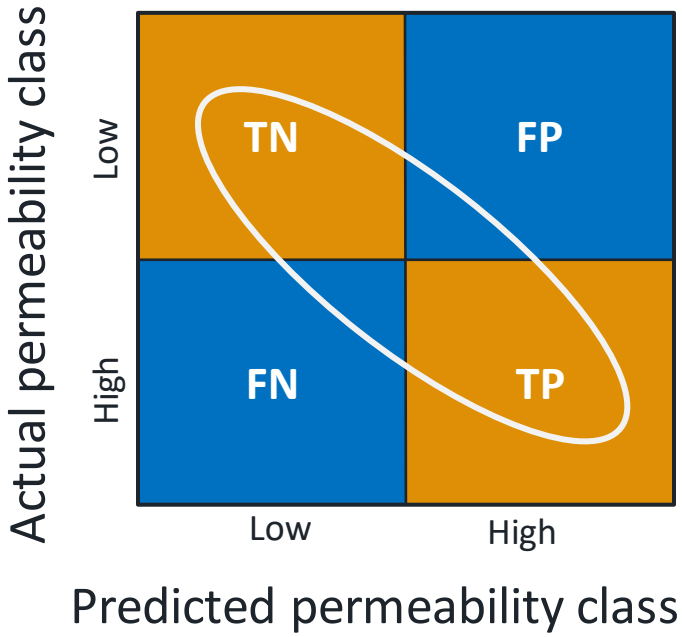
- This is a variant of commonly used molecular property prediction ML model called **Chemprop**
- Represents a molecule as a graph made up of atoms and bonds, with associated properties

nurix

Understanding Model Performance: Confusion Matrix

"All models are wrong, but some are useful" –George E. P. Box

Confusion Matrix*



Incorrect Predictions lie in off-diagonal boxes (we want small numbers here)

Correct Predictions lie along the axis (we want big numbers here)

*Adapted From: Das, C., Sahoo, A.K. and Pradhan, C., 2022. Multicriteria recommender system using different approaches. In *Cognitive Big Data Intelligence with a Metaheuristic Approach* (pp. 259-277). Academic Press.

nurix

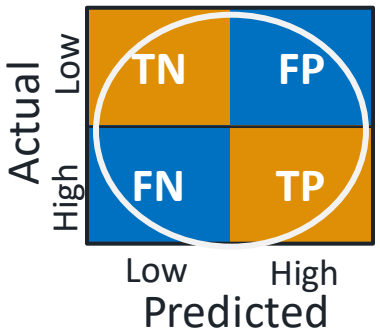
Understanding Model Performance: Metrics

"All models are wrong, but some are useful" –George E. P. Box

Accuracy

How many of the overall predictions were correct?

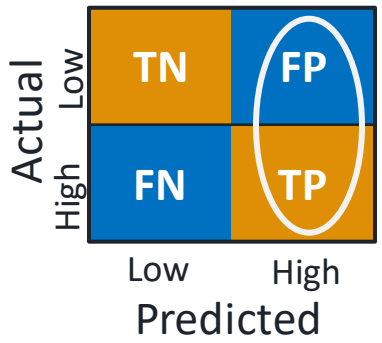
$$\frac{TP + TN}{TP + TN + FP + FN}$$



Precision

How likely is a predicted active to be active?

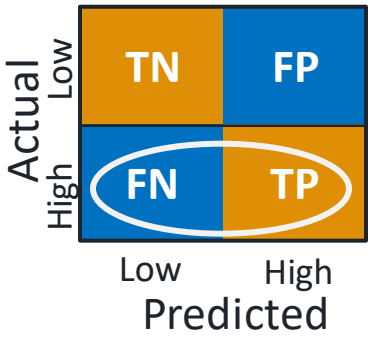
$$\frac{TP}{TP + FP}$$



Recall

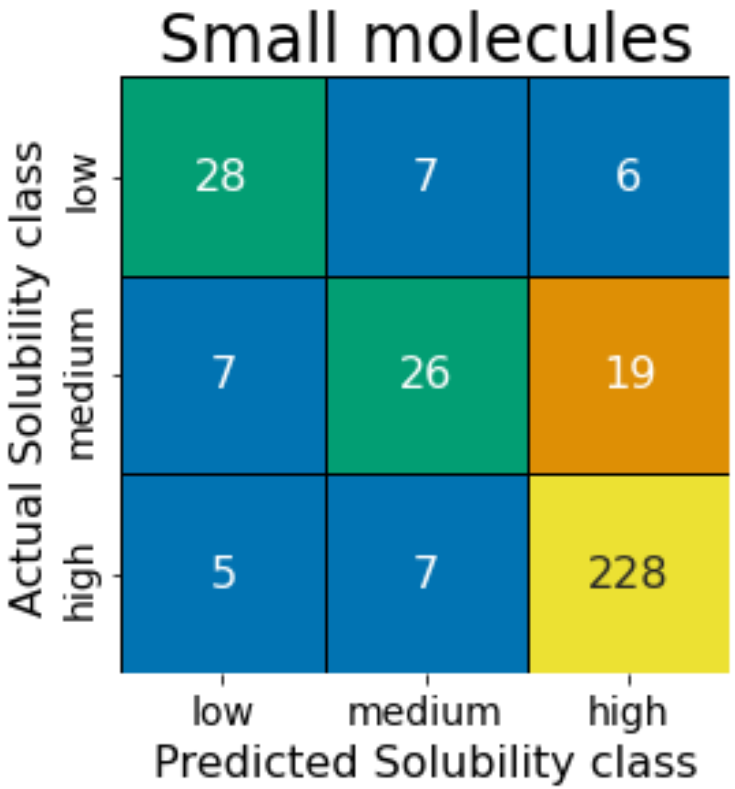
How many of the possible actives did our model find?

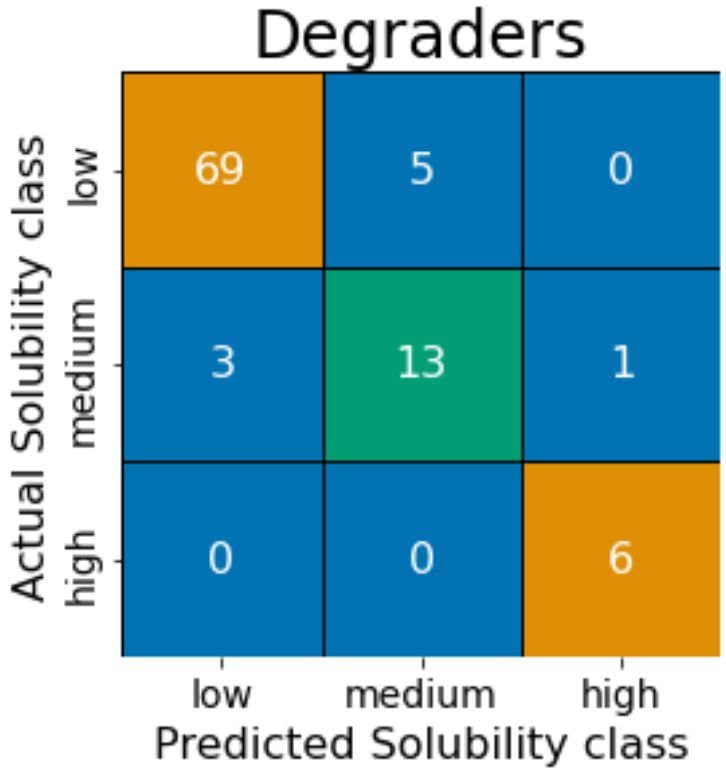
$$\frac{TP}{TP + FN}$$



nurix

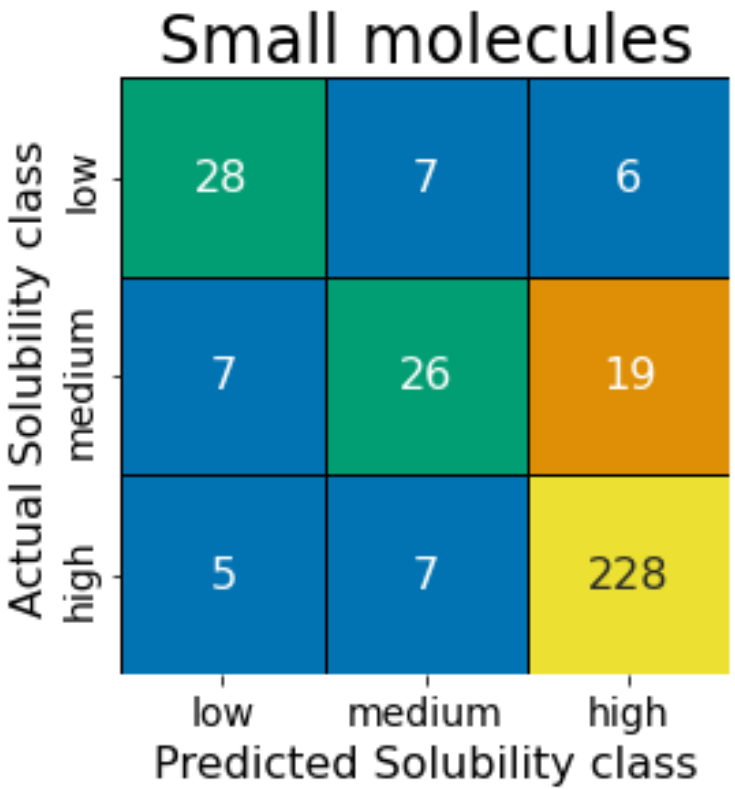
Prospective Performance Demonstrates the Best Performing Ksol Model Generalizes Well to Both SM and TPD

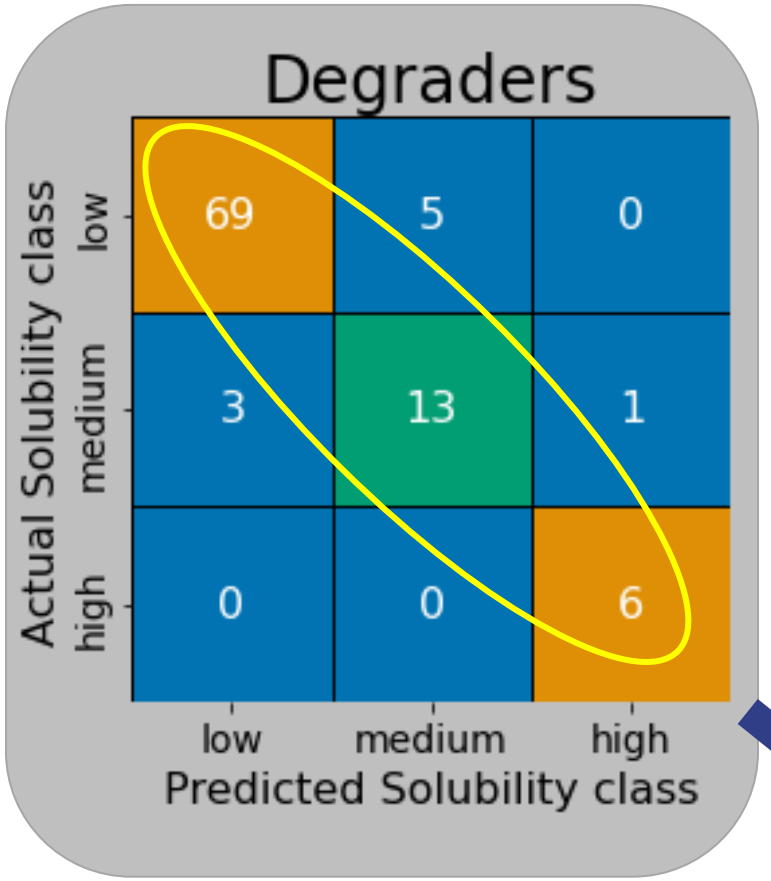




Overall Prospective Accuracy = 86%
Small Molecule Accuracy = 90%
Degrader Accuracy = 85%

Prospective Performance Demonstrates the Best Performing Ksol Model Generalizes Well to Both SM and TPD





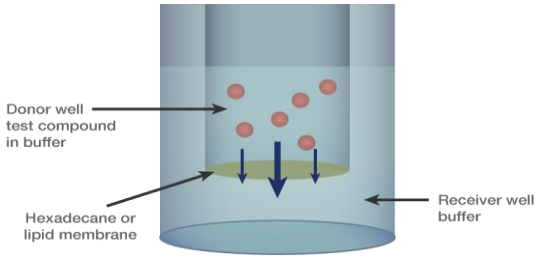
Overall Prospective Accuracy = 86%
Small Molecule Accuracy = 90%
Degrader Accuracy = 85%

High accuracy in prospective prediction of the true label of TPD compounds

nurix

Introduction to Permeability ML: The Assay

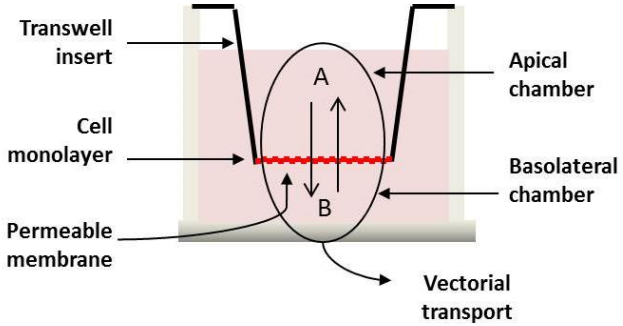
Artificial membrane assays (PAMPA)



Measures passive diffusion across an artificial membrane.

Endpoints: logPe, Recovery

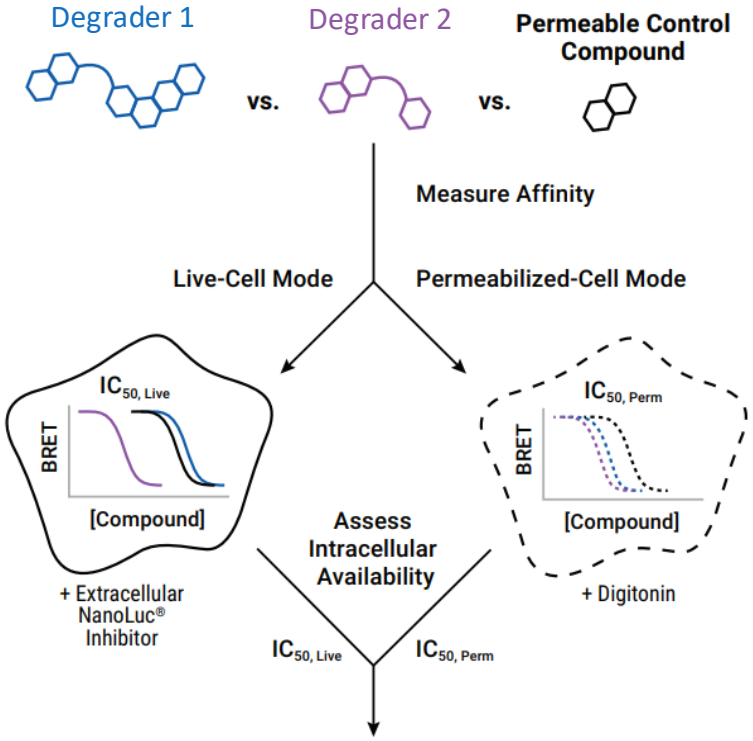
Cell monolayer assays (MDCK, Caco-2)



Directional transport owing to transporter proteins

Endpoints: Papp A->B, Papp B->A, Efflux Ratio, Recovery, Leakage, TEER

nanoBret Target Engagement assay



1. Relative Binding Affinity (RBA) = $\frac{IC_{50,\,Live}}{IC_{50,\,Perm}}$

2. Availability Index (AI) = $\frac{RBA_{PROTAC}}{RBA_{CONTROL}}$

3. Compare Values:

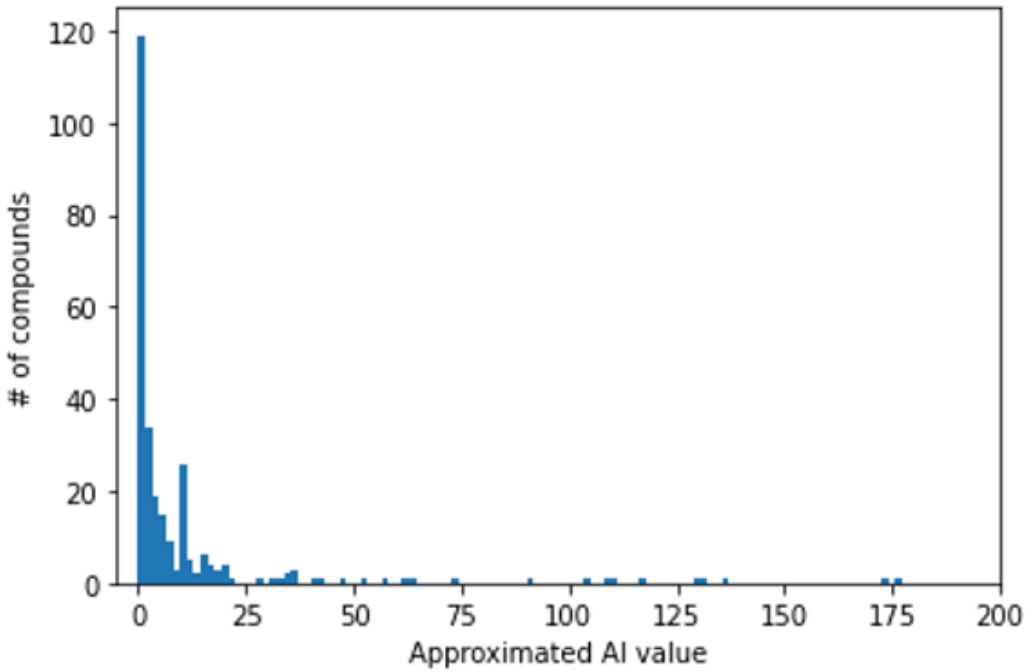
Compound	RBA	AI
Control	0.8	1.0
Degrader 1	16.0	20.0
Degrader 1	0.9	1.1

Evaluate shift in binding to intracellular target in live vs permeabilized cells

Endpoints: Relative Binding Affinity (RBA), Availability Index (AI)

nurix

Nurix Permeability Dataset and Framing the ML Problem



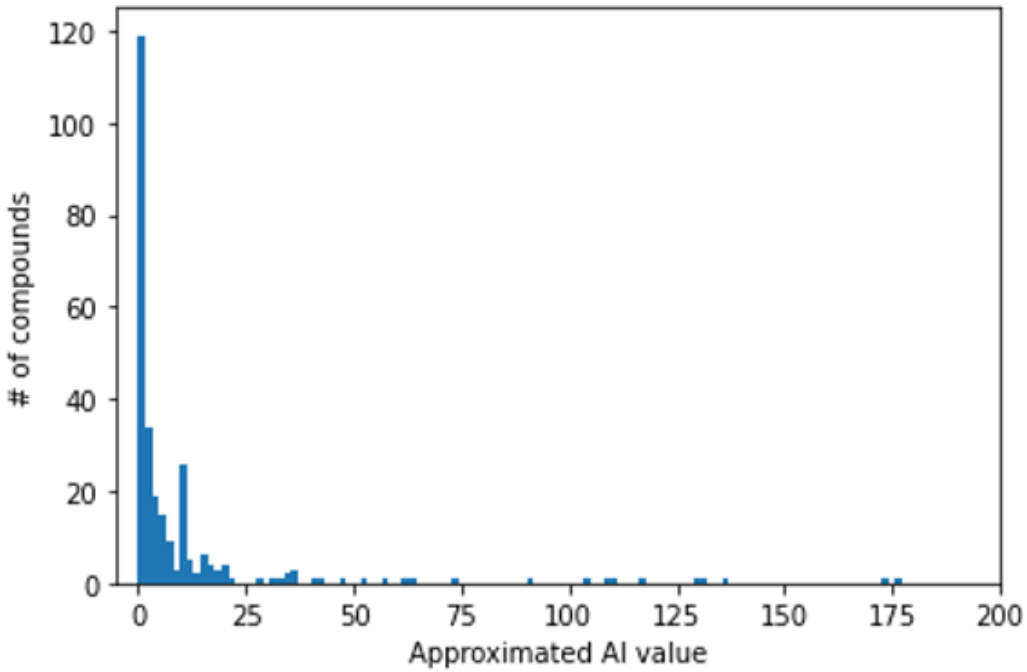
Binary Classification (Permeable/Impermeable)

Multiclass Classification (High/Med/Low)

Regression (Predict numeric TE AI value)

Due to the small size of the Nurix permeability dataset, we evaluated a binary-classification model

Nurix Permeability Dataset and Framing the ML Problem



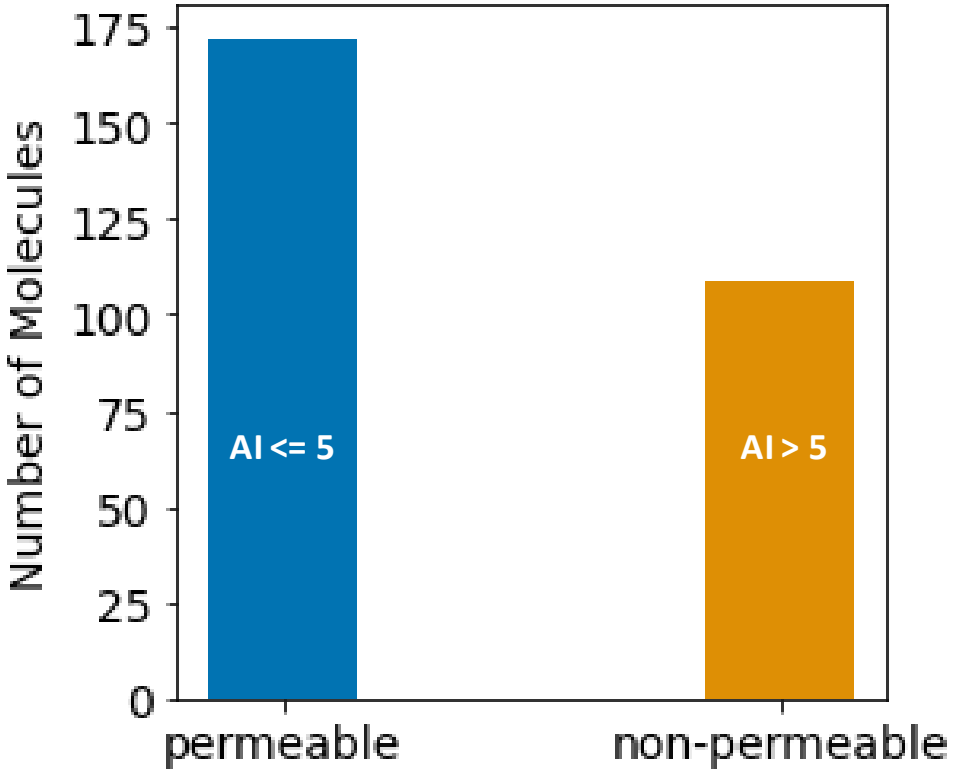
Binary Classification (Permeable/Impermeable)

Multiclass Classification (High/Med/Low)

Regression (Predict numeric TE AI value)

Due to the small size of the Nurix permeability dataset, we evaluated a binary-classification model

Permeability Data as a Binary-Class ML Dataset



- Dataset distribution : 281 degraders representing multiple ligases
- Data labeling: If **Availability Index (AI) > 5, then non-permeable** else **permeable.**

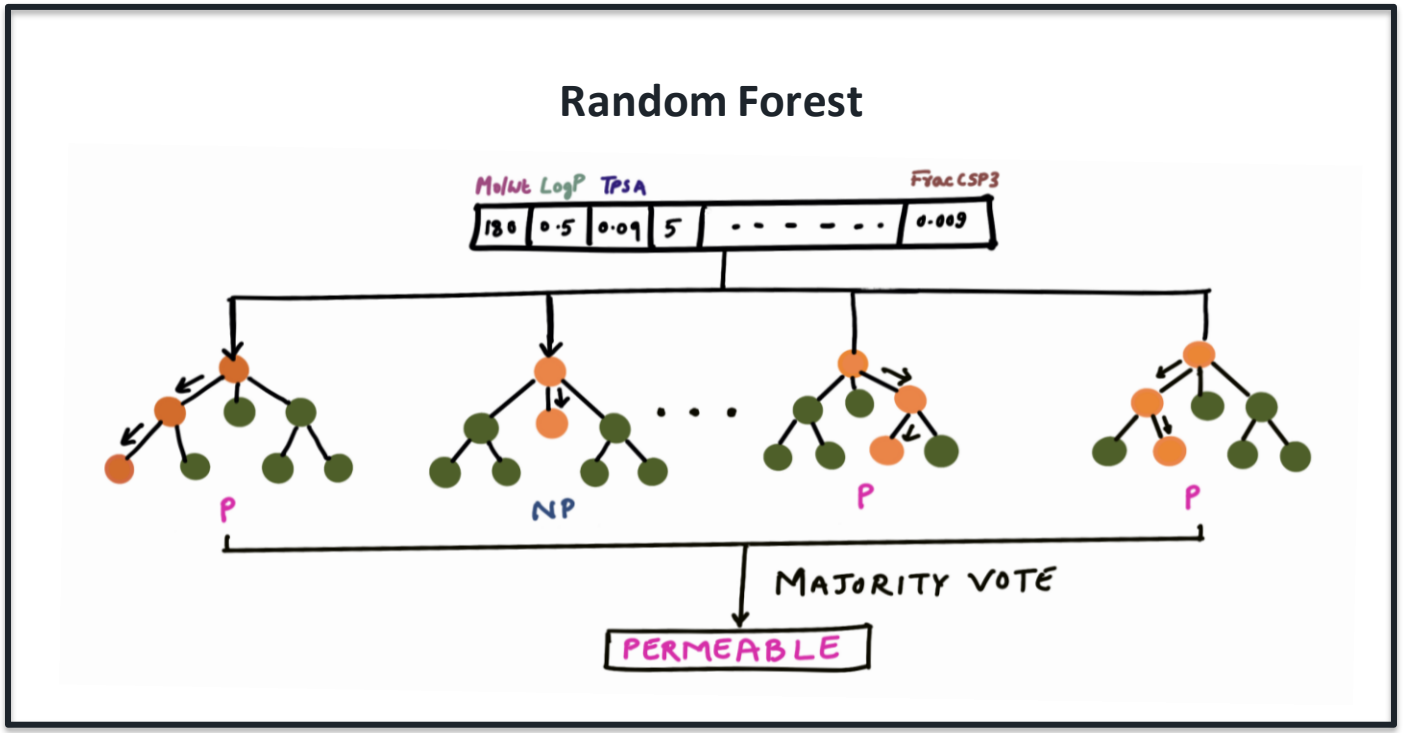
Feature Selection To Enable Human Interpretability

In determining how to featurize our molecules, we had a few considerations:

- Smaller dataset -> too many features can lead to overfitting
- Desire to enable human interpretability of ML model learned SAR

MW	#Rings
cLogP	FracSP3
#HBD	#HeteroAtoms
#HBA	TPSA

Model Overview of Best Performing Permeability ML Model

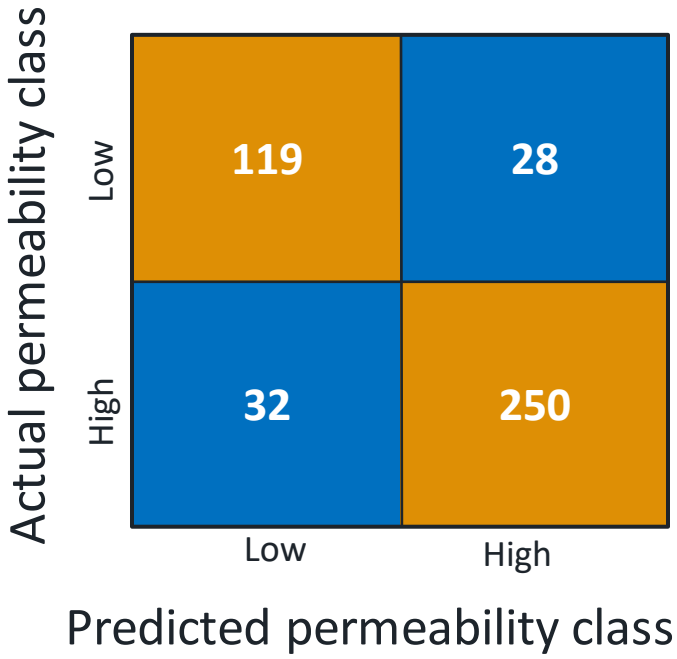


A collection of decision trees that predicts permeability based on molecular properties

nurix

Permeability Model Achieves High Levels of Accuracy in Prospective Predictions

RF Model achieves **86% accuracy and 90% precision in prospective** predictions generalizing across different ligases and protein targets



Permeability Model Achieves High Levels of Accuracy in Prospective Predictions

RF Model achieves **86% accuracy and 90% precision in prospective** predictions generalizing across different ligases and protein targets

By Target

Target	Accuracy
1	88%
2	100%
3	92%
4	64%
5	78%

← No data in training set for this target

By Ligase

Ligase	Accuracy
CRBN	93%
VHL	81%
Others	60%

← Low representation in training set

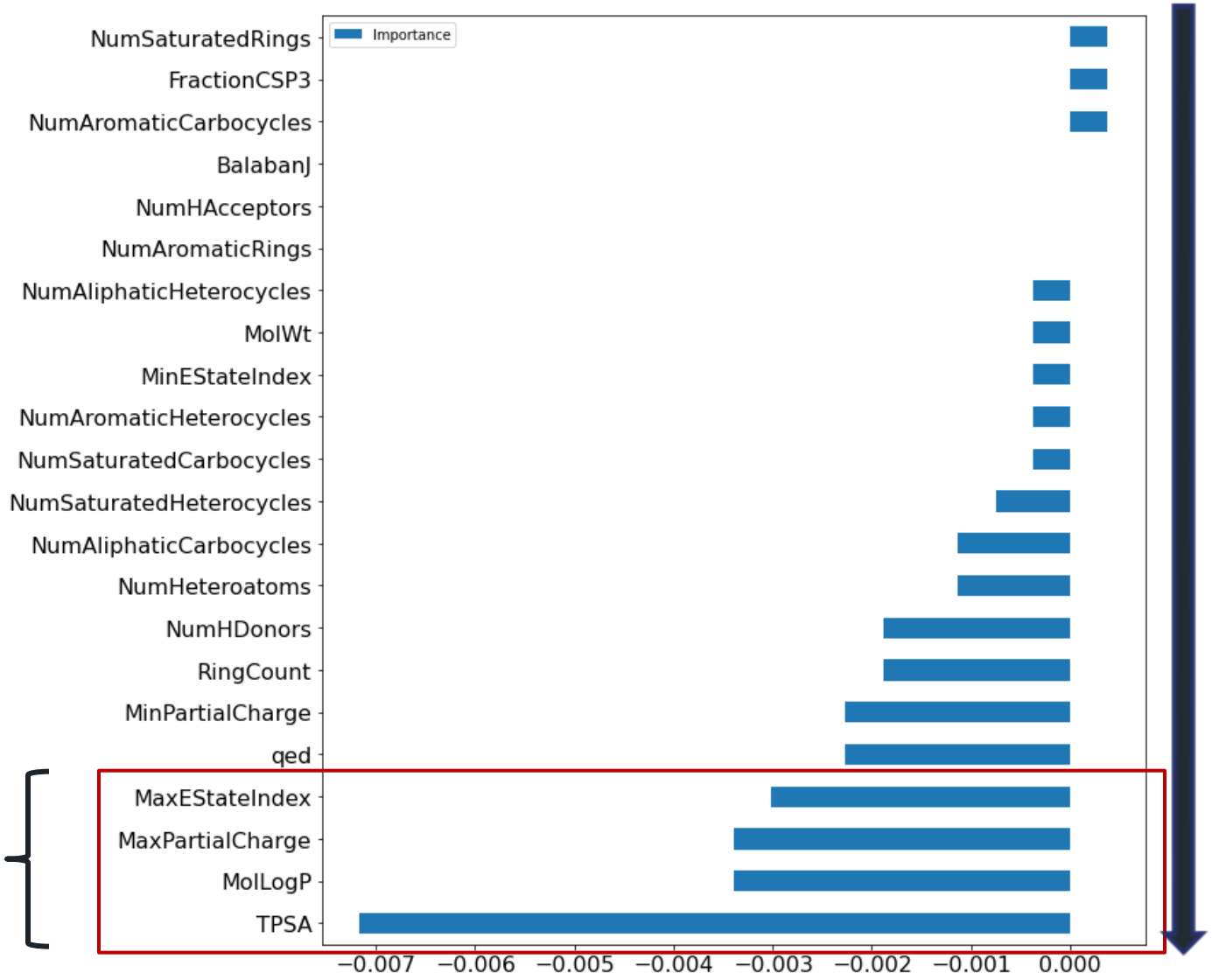
Discussion

Interplay between model learnings and human interpretability

nurix

Investigating What the Model Has Learned

- Feature importance tells which features contributed to model's learning process.
- Bars to the left are negatively correlated and to the right are positively correlated to permeability.
- **Interpretability analysis can allow chemists to build intuition and design better compounds.**



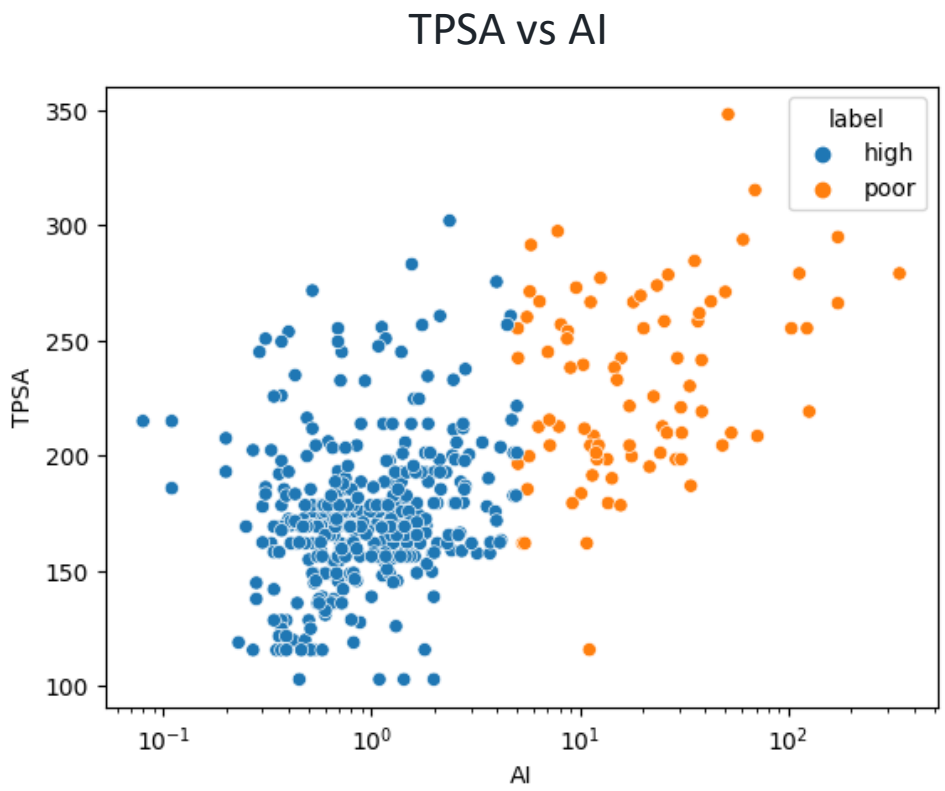
Positive correlation

Negative correlation

Important features

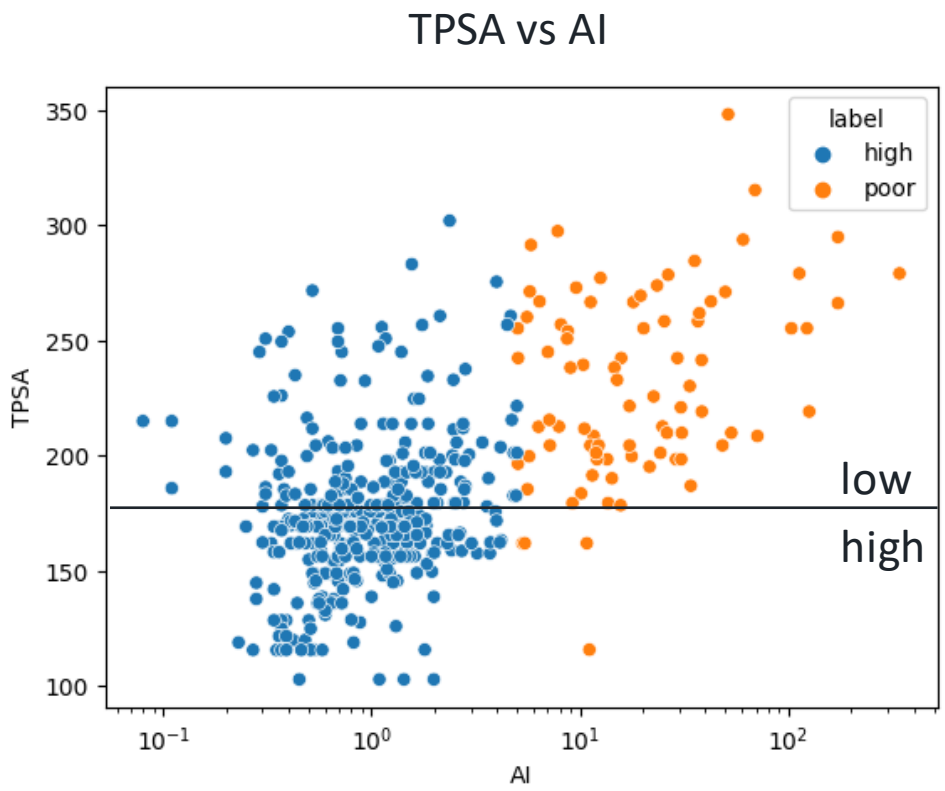
Building Human Intuition Leveraging Feature Importance Analysis

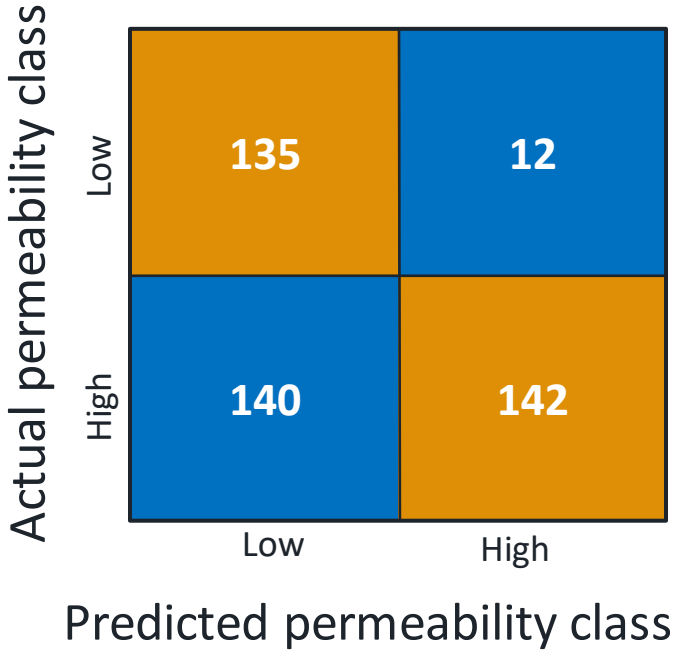
Evaluating the features suggested as important offers learnings to inform human guided molecular design



Building Human Intuition Leveraging Feature Importance Analysis

Evaluating the features suggested as important offers learnings to inform human guided molecular design

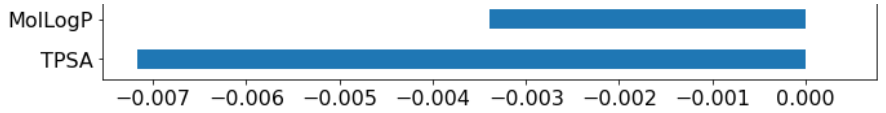




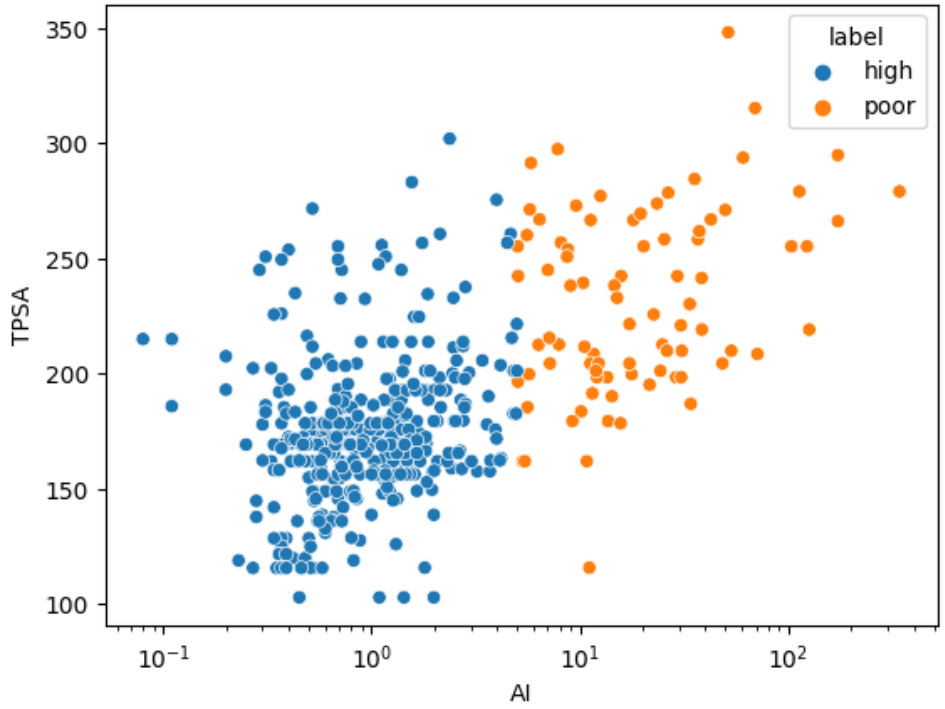
65% accuracy and 92% precision with a simple model based solely on TPSA

Building Human Intuition Leveraging Feature Importance Analysis

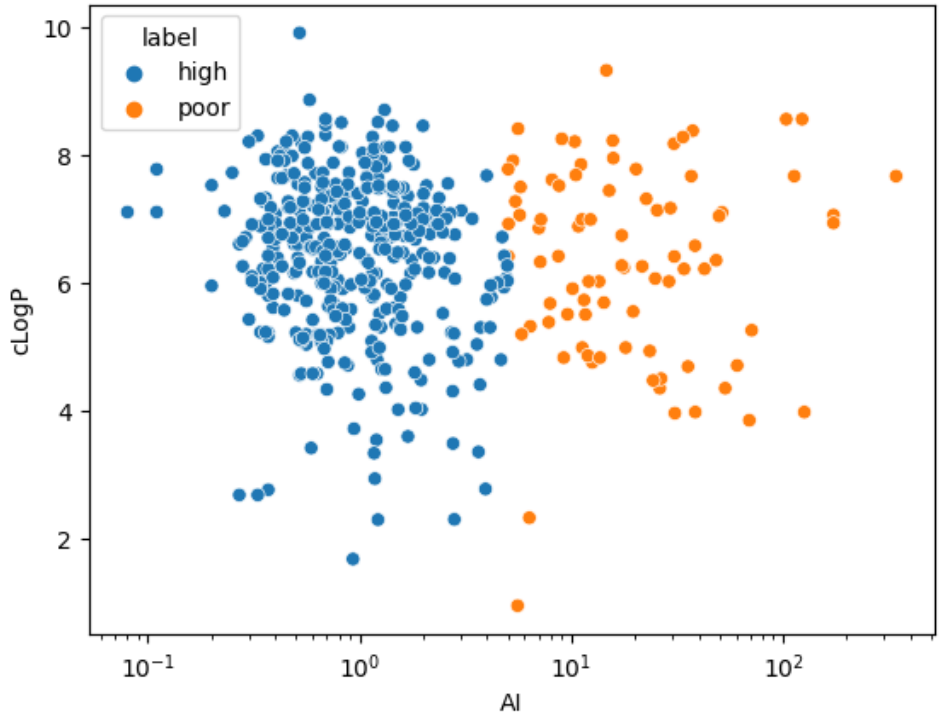
Evaluating the features suggested as important offers learnings to inform human guided molecular design



TPSA vs AI



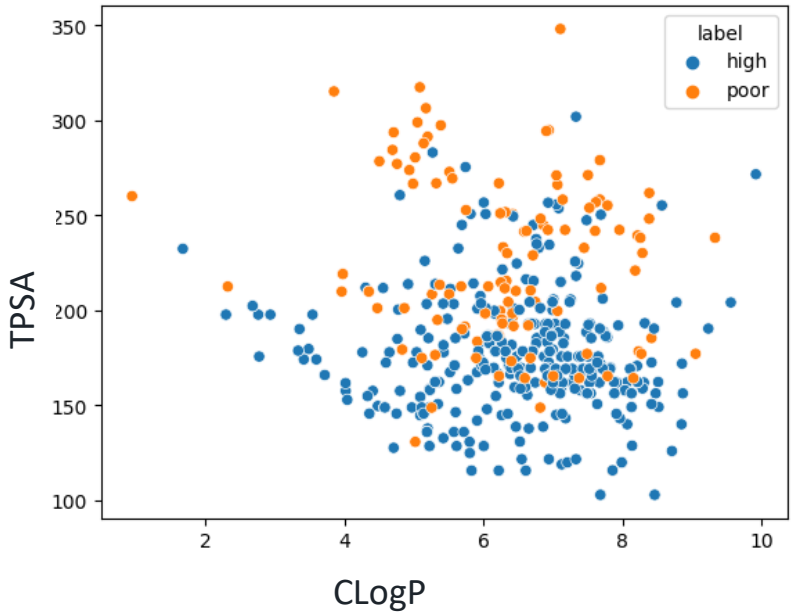
cLogP vs AI



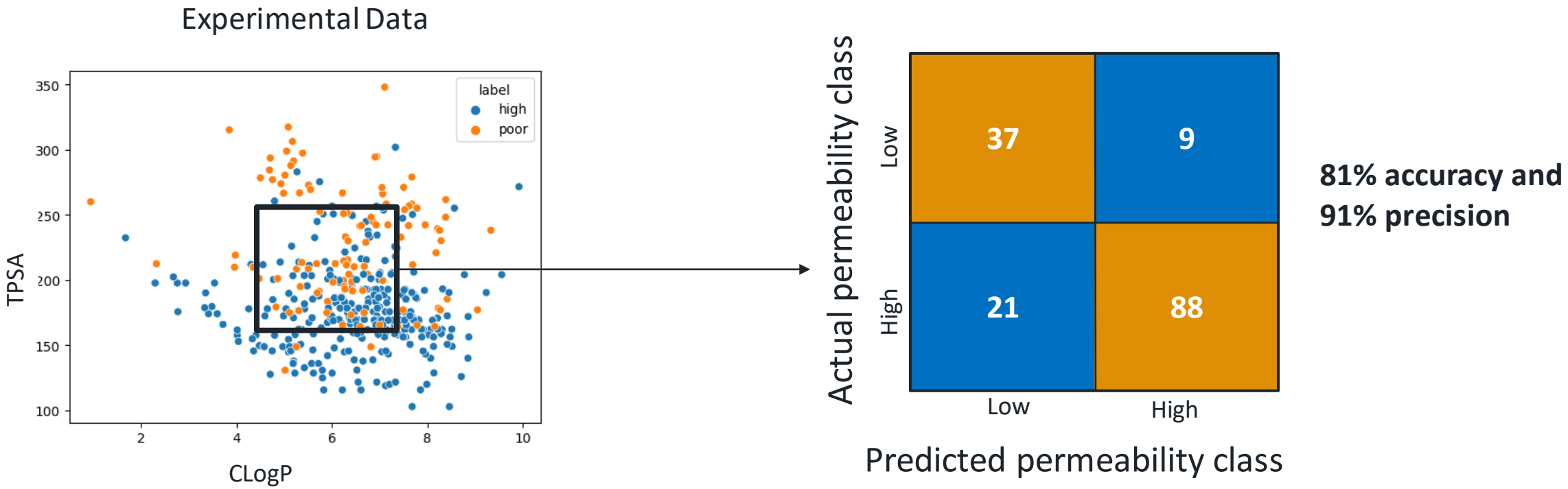
nurix

ML Model Complements Intuition In Non-Obviously Separable Property Space

Experimental Data

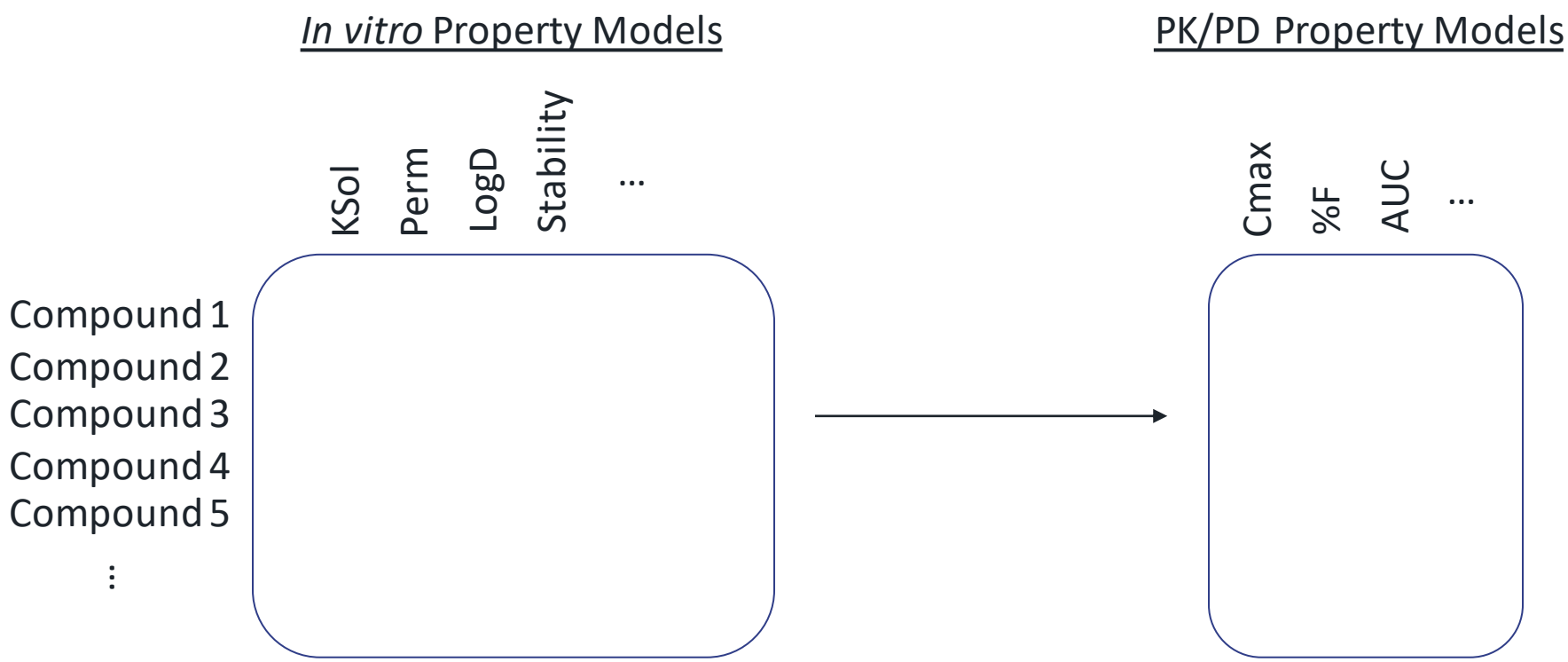


ML Model Complements Intuition In Non-Obviously Separable Property Space

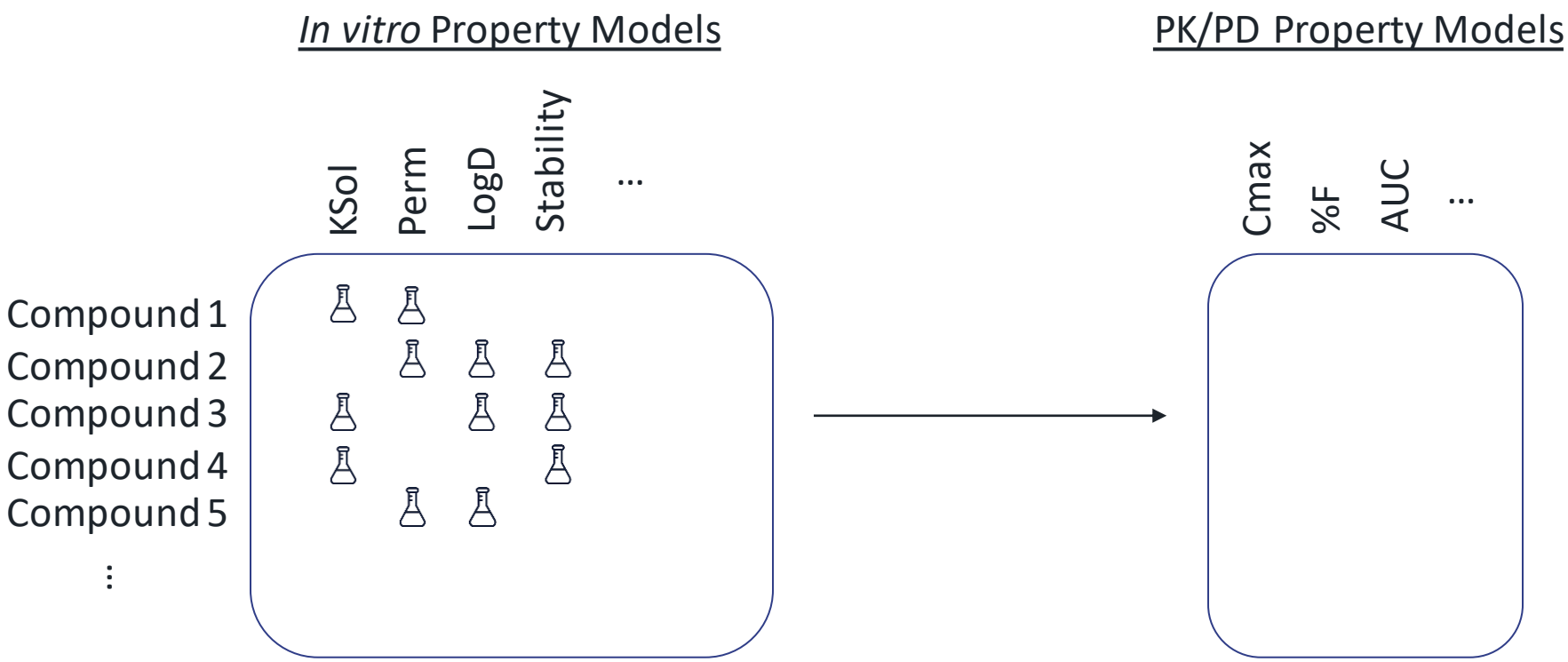


Permeability ML model predicts correct class membership despite no obvious separation of High/Low permeable compounds in property space defined by highest importance features.

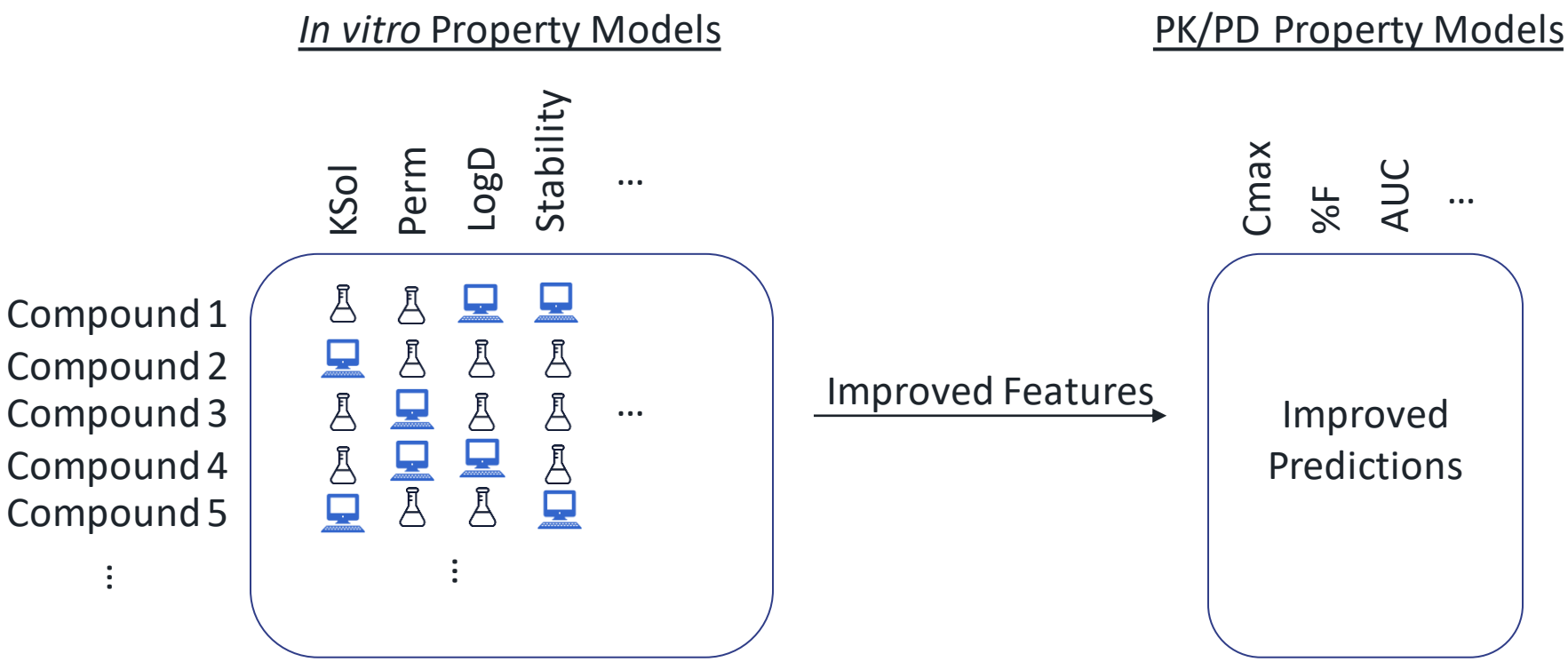
Rebuilding Understanding to Enable Improved Predictions



Rebuilding Understanding to Enable Improved Predictions



Rebuilding Understanding to Enable Improved Predictions



Conclusions

- Many off-the-shelf property calculators do not generalize well to TPD molecules given the novel chemical space occupied by these compounds
- Machine Learning can re-learn the rules of structure-property relationships and offer a useful tool to inform molecular design during Lead Optimization
- ML models for structure-property prediction can enable and complement human understanding of learned SPR

nurix

Thank you