Atomwise

Quantitatively predicting binding affinity globally across structures and ligands

Comparing voxel-based and point-based convolutional neural network models

Michael Mysinger, PhD
Principal Scientist
Atomwise Inc.

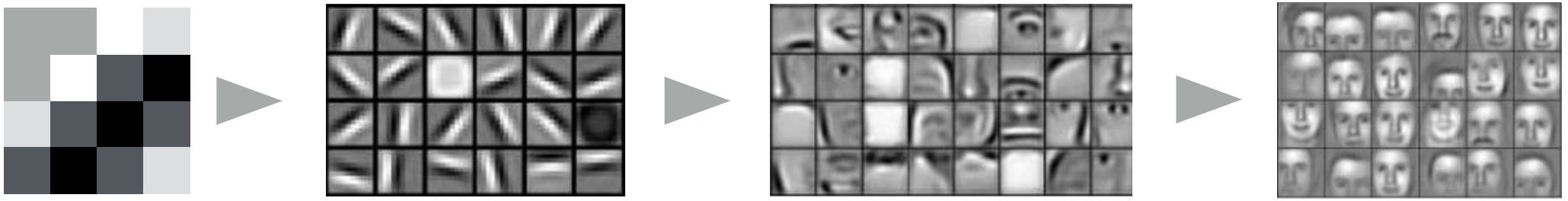
Agenda

- Model Architectures
- Global pKi Model
- Compound Series Benchmark
- Point-based versus Voxel-based
- Target versus Compound Splits
- Adding Inequality Data

Where does Deep Learning Fit In?

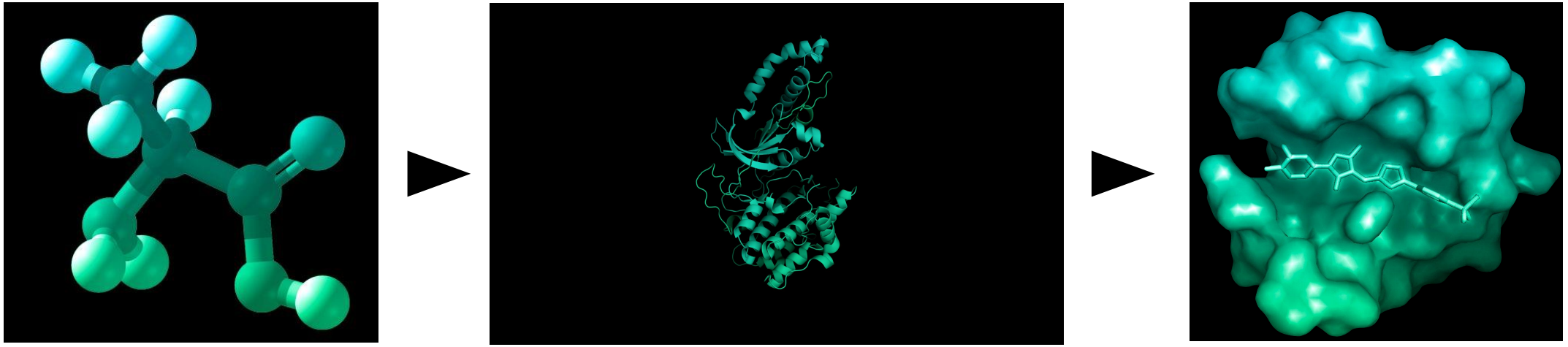
Adapting deep learning architectures to structure-based drug design

Image Recognition



Pixels → Edges → Eyes, Noses, Mouths → Faces

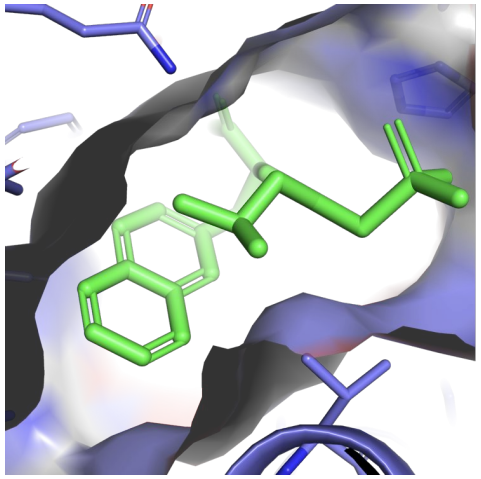
Molecular Recognition



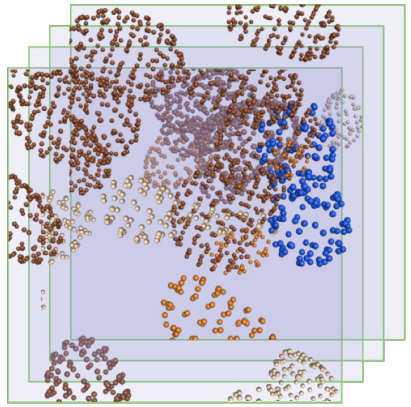
Atoms in 3D Space → Biochemical Features → Molecule-Protein Binding

Voxel-based is Like Pixels and Images in 3D

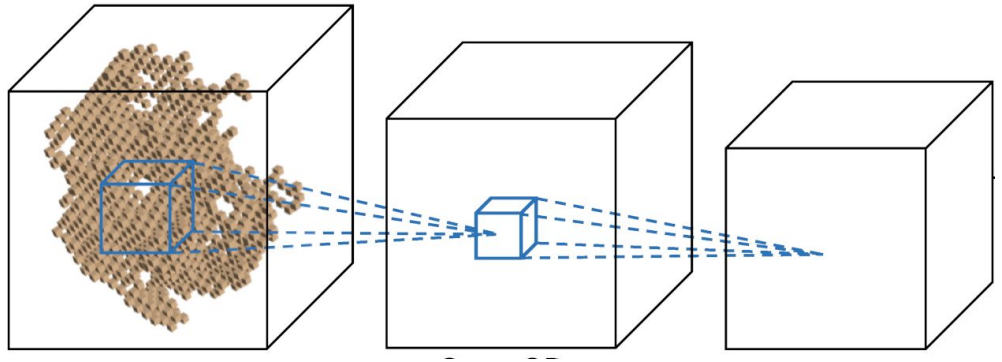
Classical convolutional neural networks



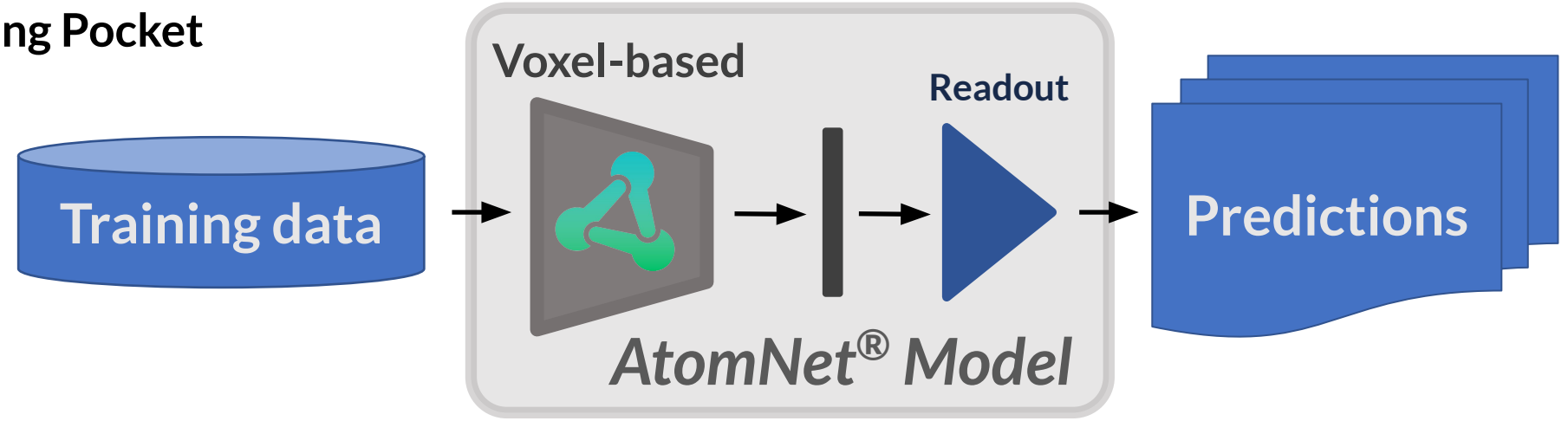
Binding Pocket



Voxelized Grid

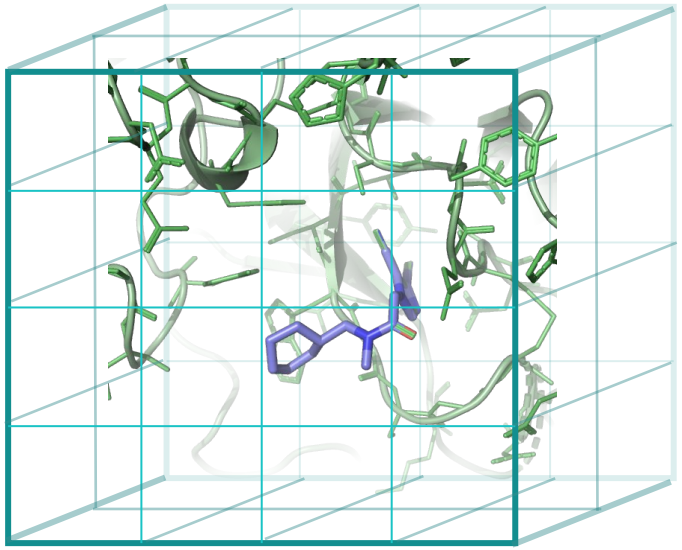


Convolutional Neural Network

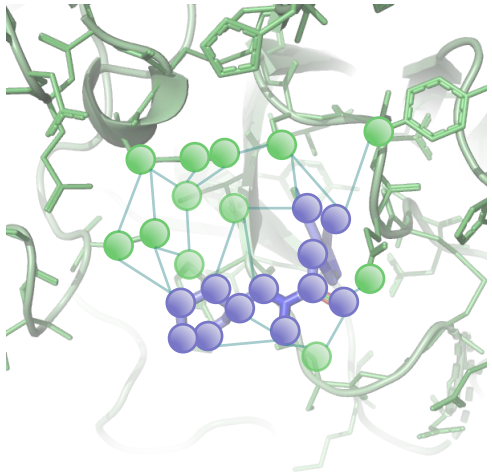


Point-based uses Atoms and Distances

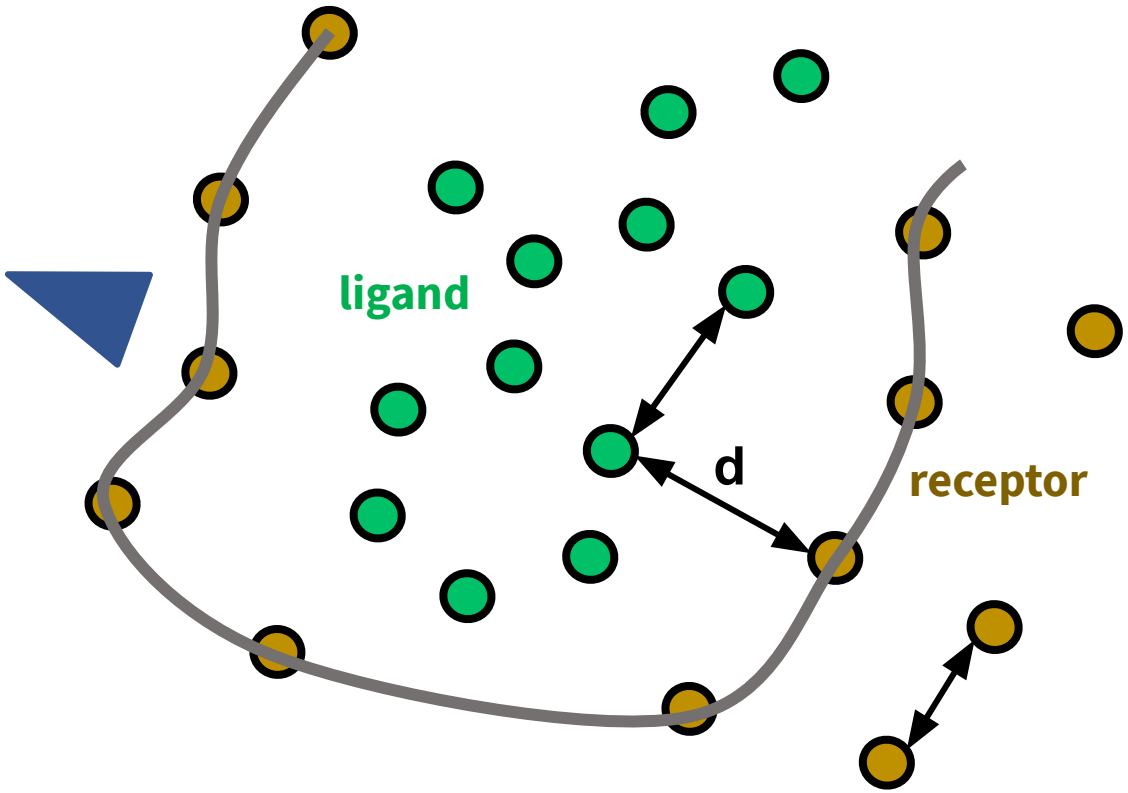
Message passing neural networks (MPNN) on atomic graphs



Voxel-based



Point-based



Directional Message Passing

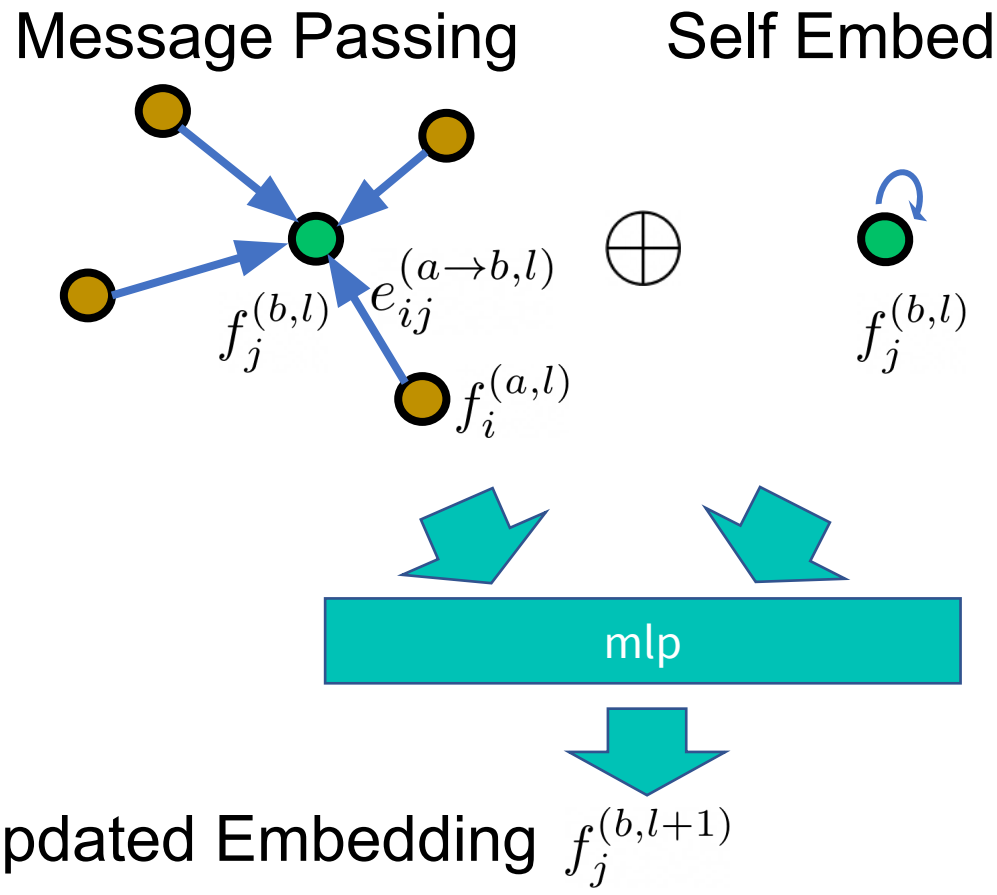
Point-based MPNN ideas

- Update messages depend on
  - Source atom embedding
  - Distance between atoms
- Messages are linear combinations of Radial Basis Functions:

$$e_{ij}^{(a\to b,l)} = \sum_c c_n^{(l)} R_n(r_{ij})$$

e.g., radial Bessel functions

$$R_n(r_{ij}) = j_0\left(\frac{z_{0n} r_{ij}}{R}\right)$$



Voxel-based versus Point-based Networks

Pros and cons for structure-based pKi prediction

- Voxel-based – 3D CNN
  - Maps easily to image-based literature
  - Needs rotational data augmentation
    - Harder to learn
    - Yet also more robust
- Point-based – MPNN or GCN
  - More natural representation for atoms
  - Rotationally invariant
  - Requires more overfitting controls (dropout, data augmentation, etc.)

Global pKi Model

Overview

- Classification model is trained to find initial hits (binds or not)
  - For more see Pawel's talk at https://info.atomwise.com/acs_spring2021
- Regression model is trained to find tightest binders

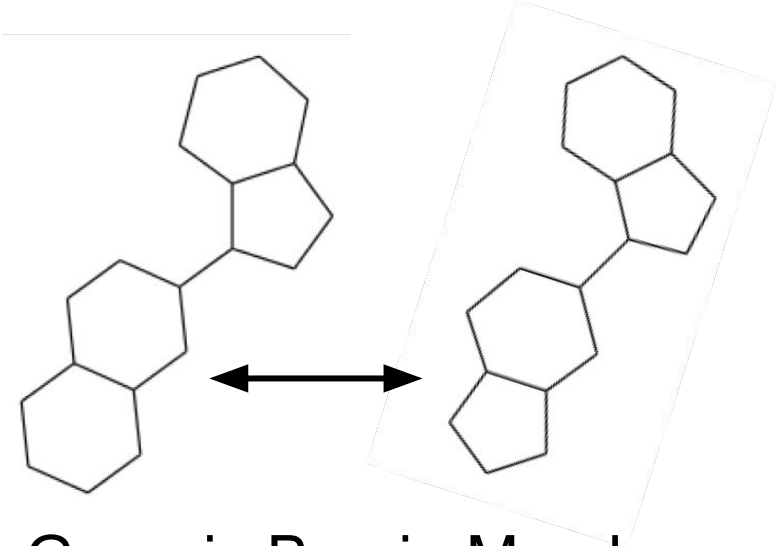
- Predict quantitative pKi ($-\log_{10} K_i$ [M])
- Global model – all targets versus all ligands
- IC_{50} and K_i quantitative binding data
  - Typically excluding inequalities

Compound Series Benchmark

Motivation

- Global pKi model primary use case is expanding initial screening hits into compound series
- Want to achieve ranking performance within a compound series
- Find molecules for each target that form chemical series from our quantitative data

How to define chemical series?

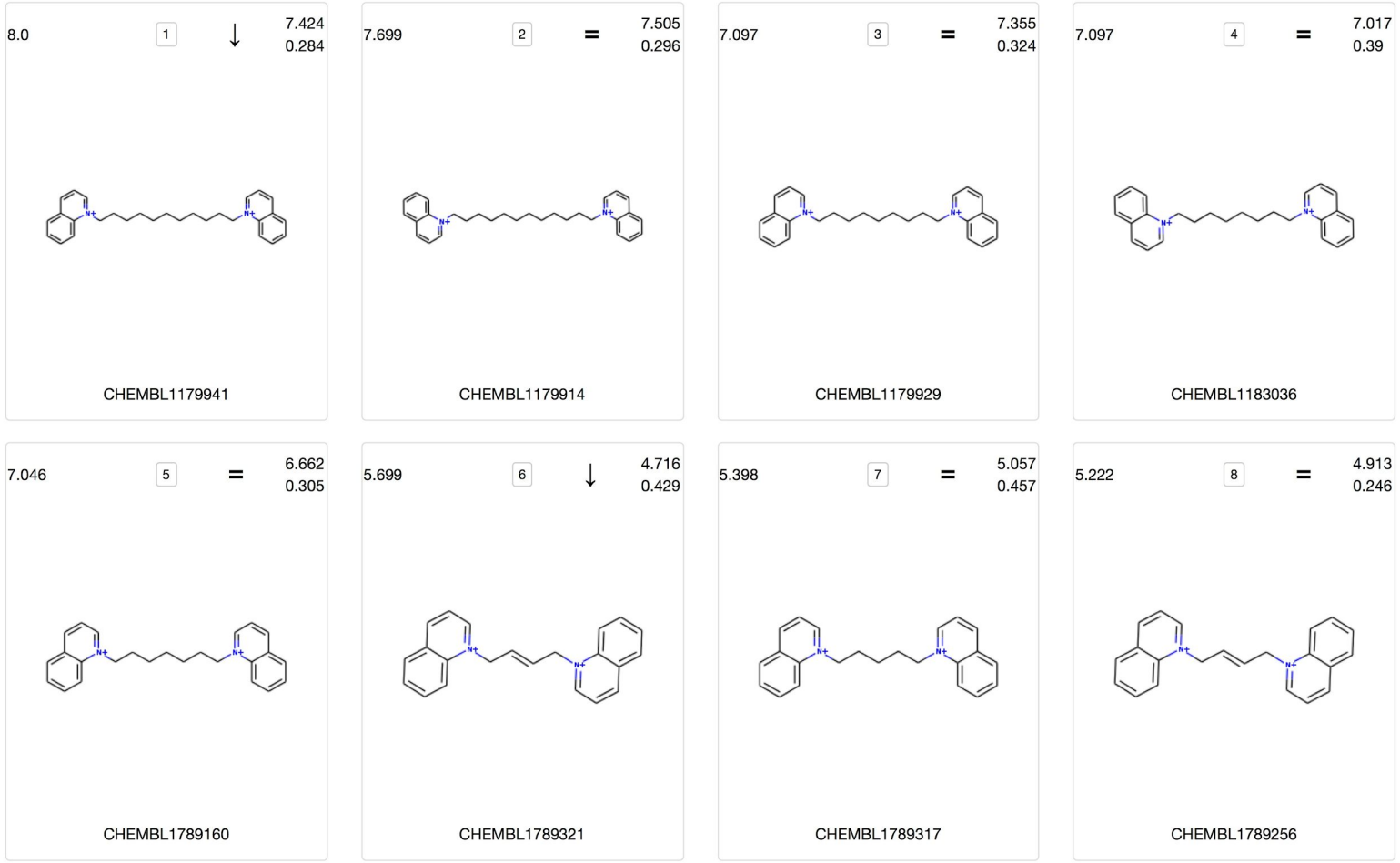


Generic Bemis-Murcko Scaffolds

Compound Series Benchmark

Matched molecular pair based Series

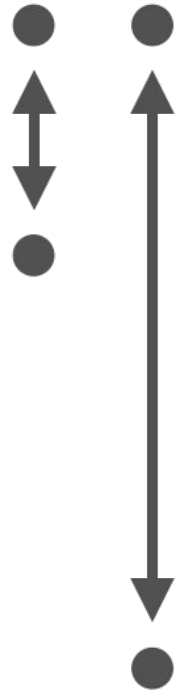
Matched Molecular Pair (MMP) Based Series



Refining the Series Benchmark

Use Kendall's tau-b that ignores small differences

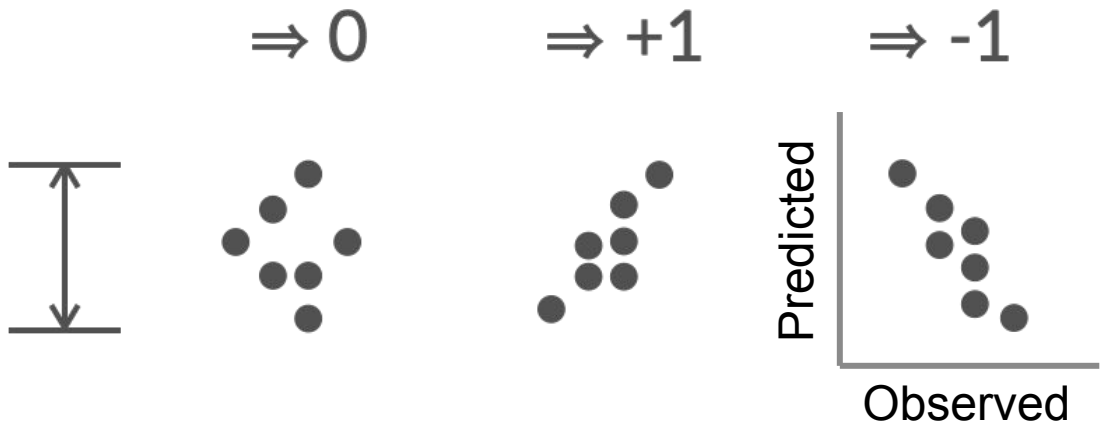
- Problem
  - Small pKi differences are too important for metrics
- Fix
  - Use a Kendall's Tau-b metric with ties
  - Measures concordance versus discordance over pairs of molecules
  - User tunable difference threshold (0.5 pKi here)



Refining the Series Benchmark

Range and dispersion criteria

- Problem
  - Smaller, tightly clustered series have highly variable performance without reason
- Fix 1
  - Effective size threshold
  - At least 10 molecules
  - 50 (~$_{10}C_2$) non-tied pairs above threshold
- Fix 2
  - Dynamic range threshold
  - Span at least 2 pKi units



Compound Series Benchmark

Putting it all together

- Result is a useful task-focused benchmark
- We use 6 cross-fold strategy to train and analyze benchmark
  - Predict each target-compound pair using the model that has not seen it
  - Sequence similarity 70% (seqsim70) splits unless otherwise noted

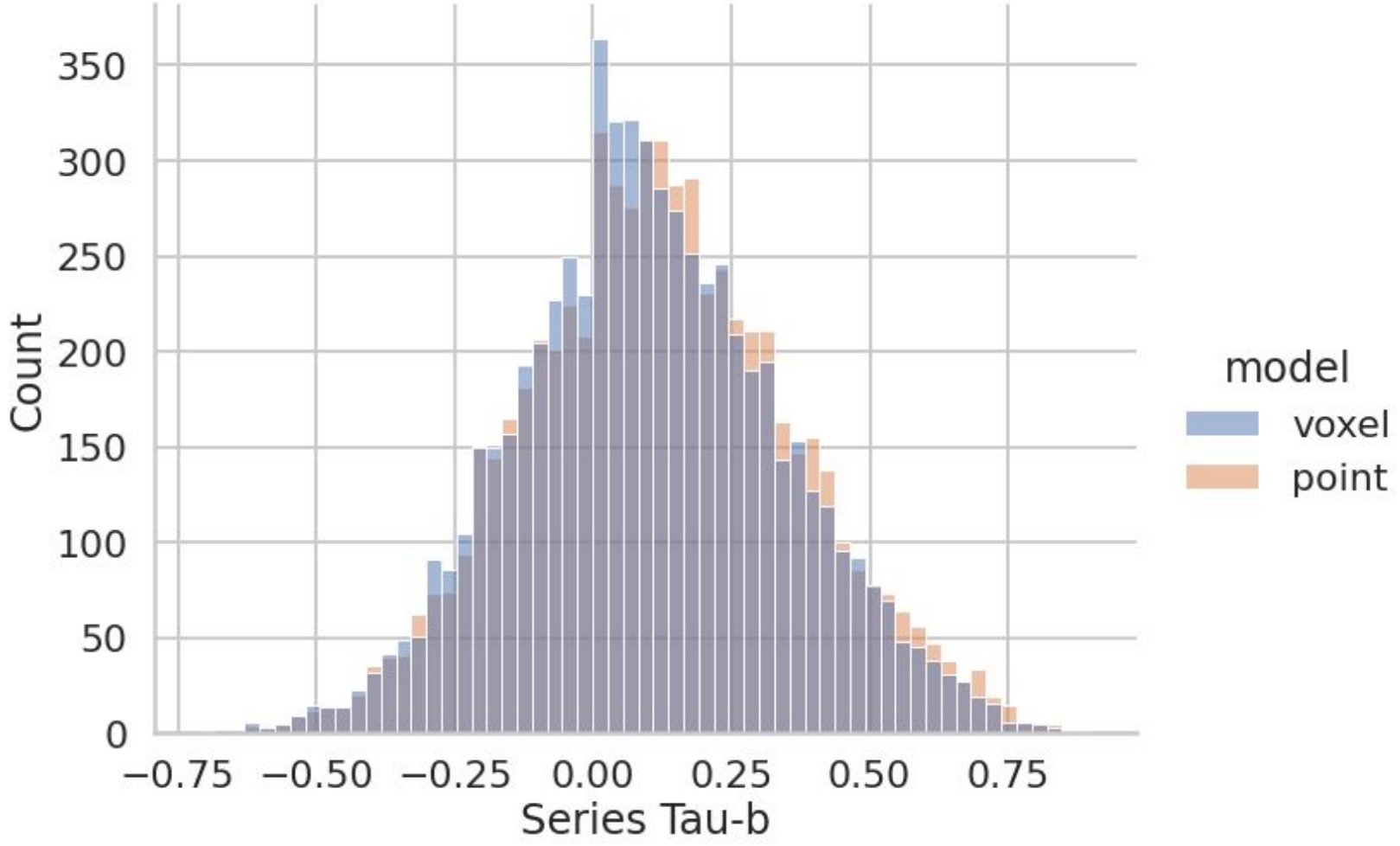
- Current benchmark composition
  - 6000 compound series
  - 700 targets
  - 200K molecules

Voxel-based versus Point-based

Which is better for pKi?

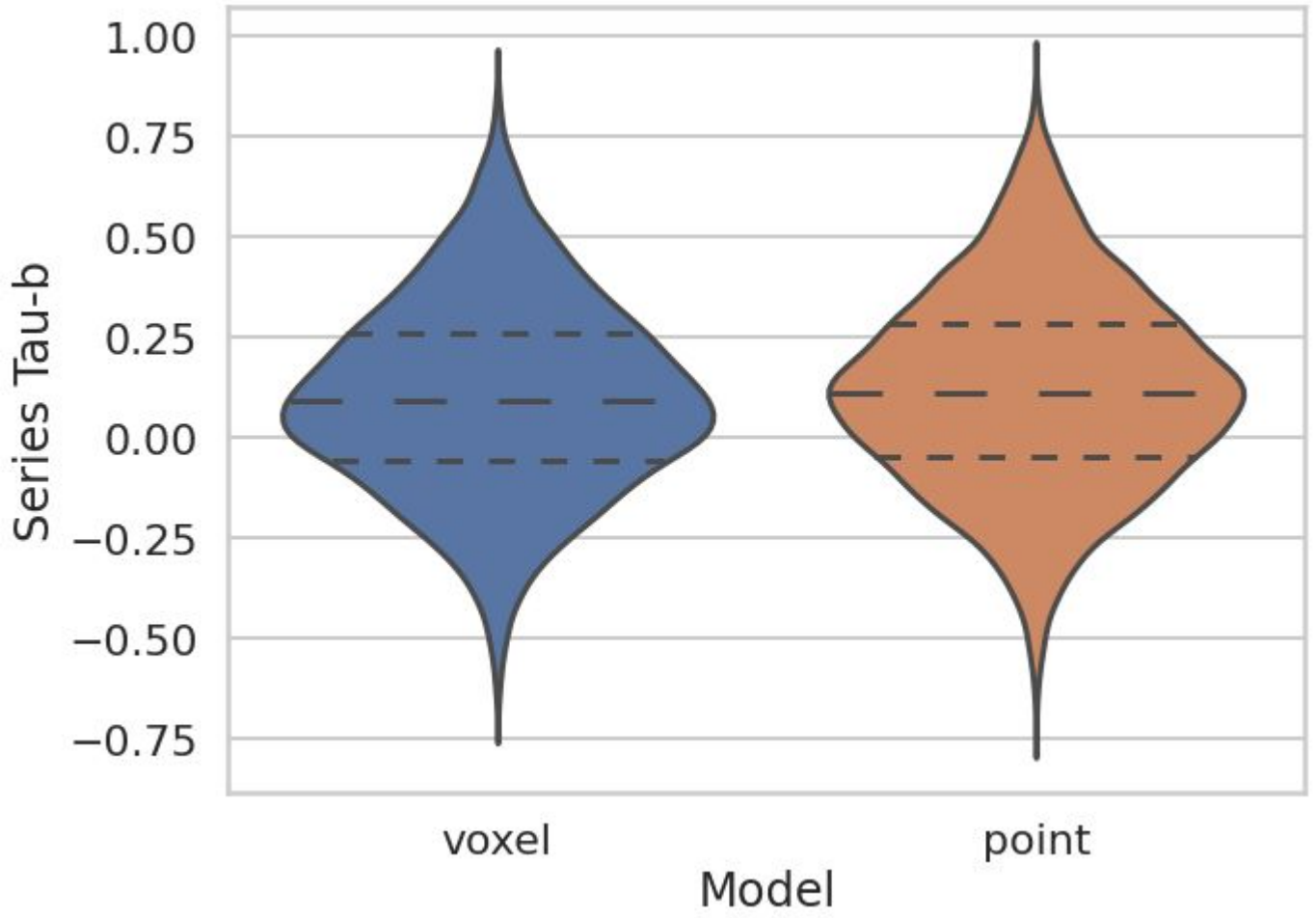
Voxel-based vs Point-based

Benchmark series tau-b histogram



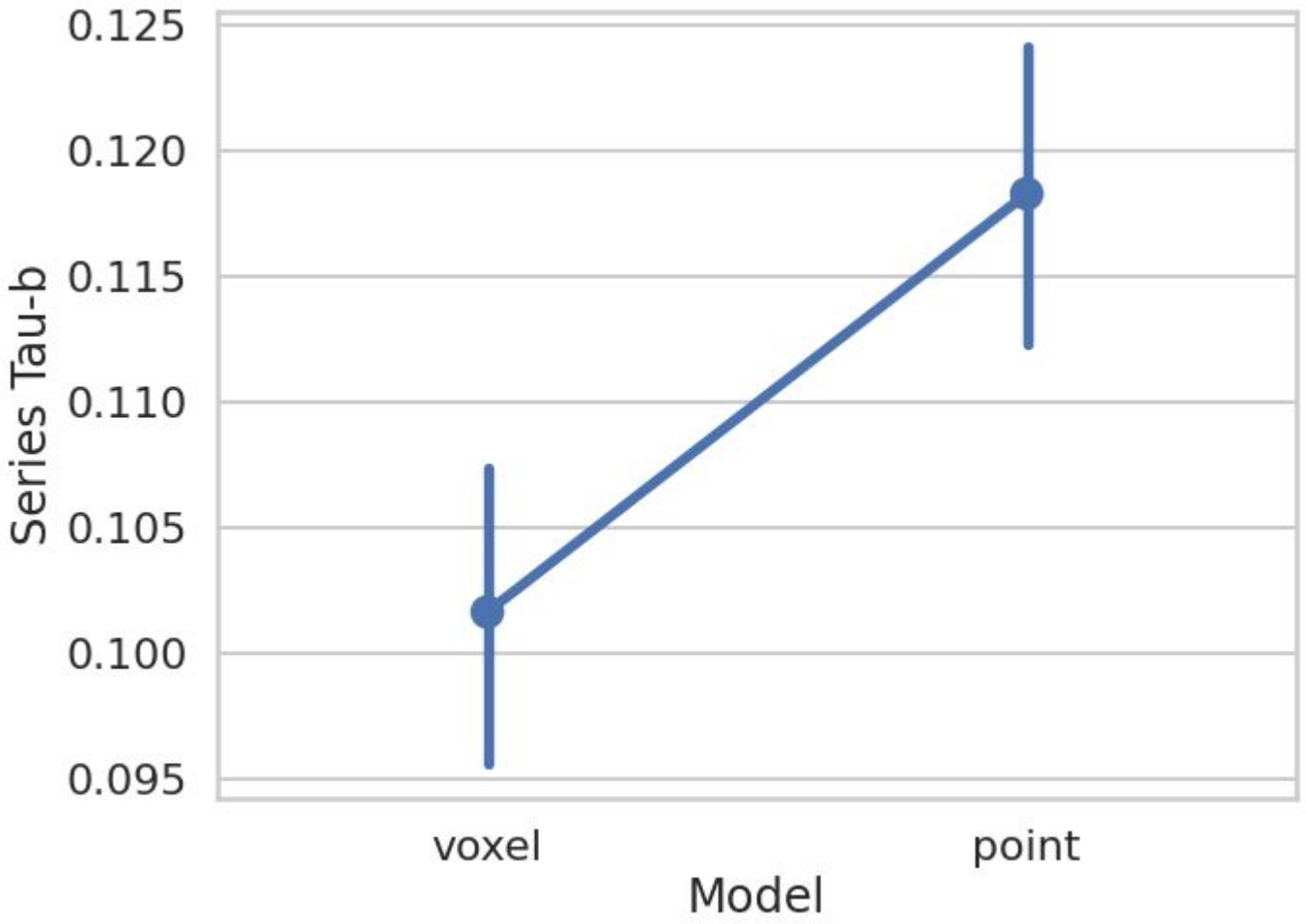
Voxel-based vs Point-based

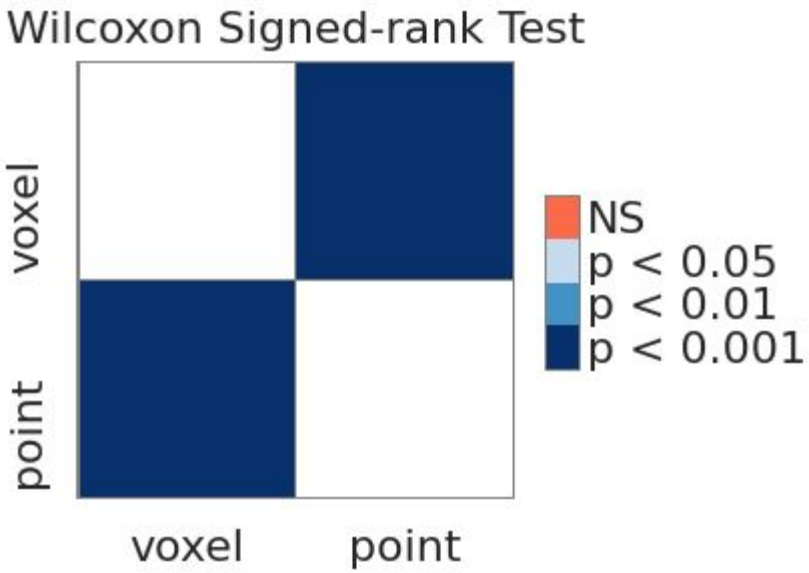
Benchmark series tau-b violin plot



Voxel-based vs Point-based

Benchmark series tau-b point plot



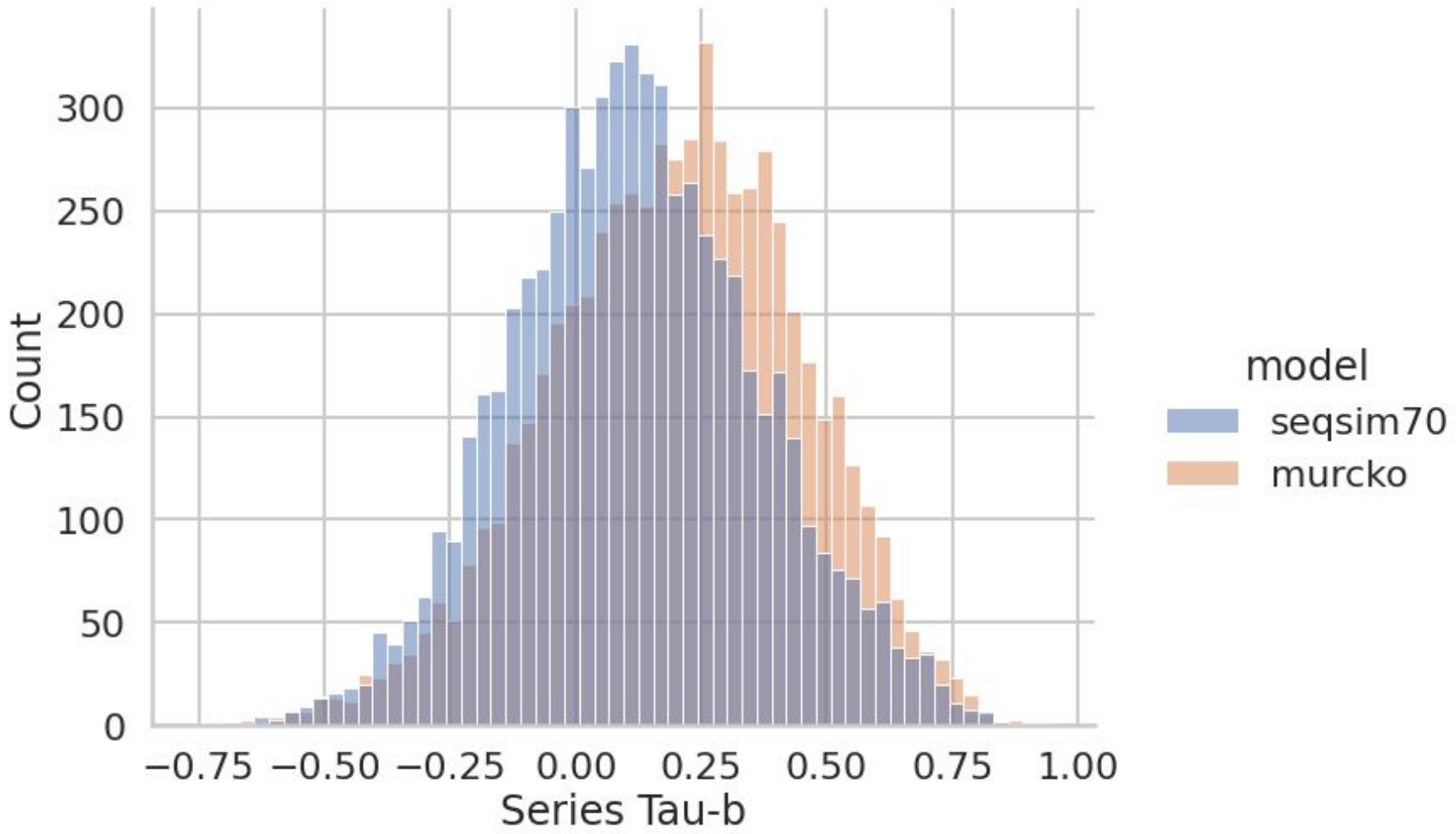


Compound versus Target Split

How different are they?

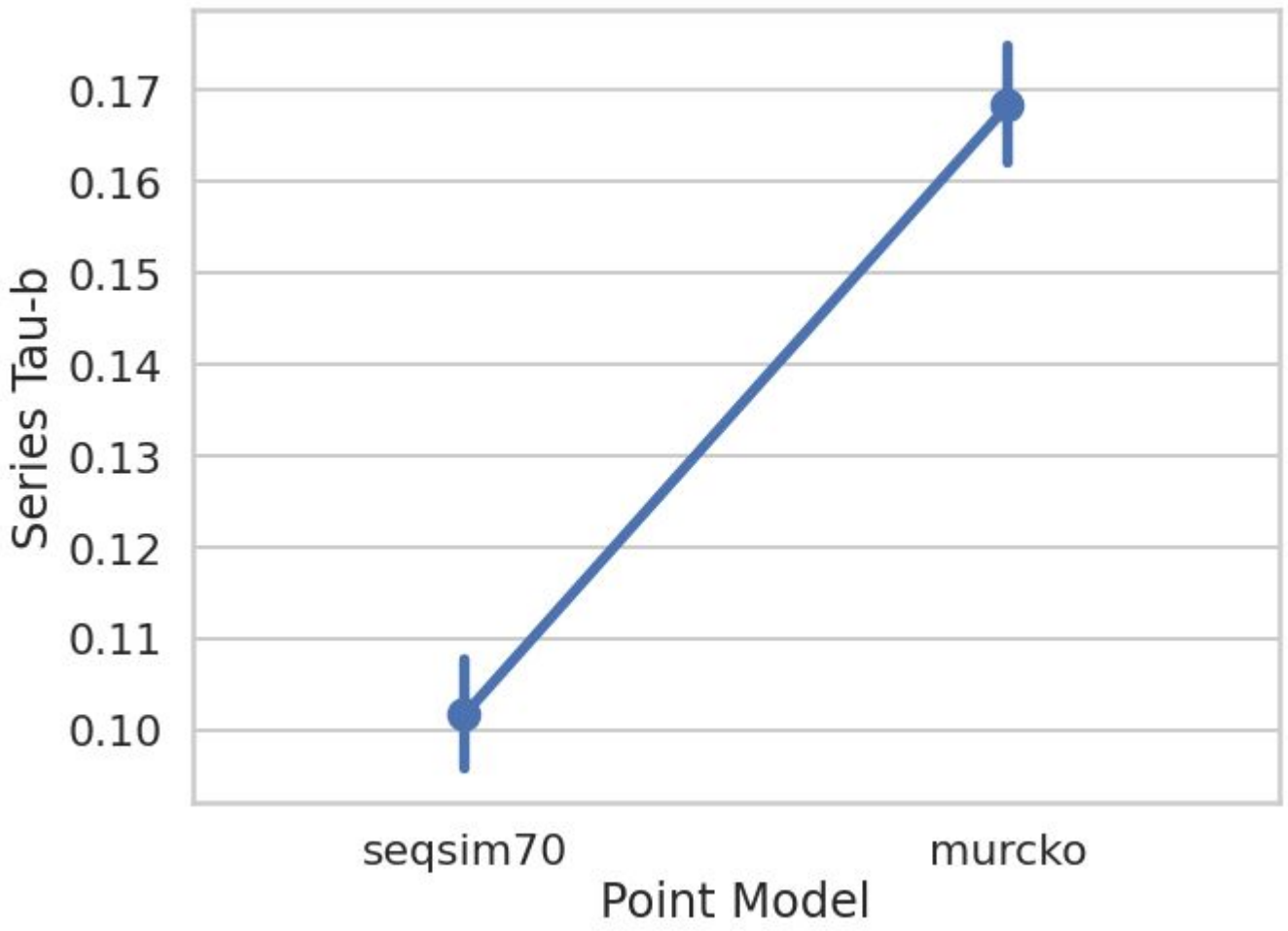
Murcko Splits Easier Than Target Splits

Splitting strategy for cross-folds histogram



Murcko Splits Easier Than Target Splits

Splitting strategy for cross-folds



Harder seqsim70 splits match our goal to tackle novel and undruggable targets

Inequality pKi Data

Does it help or hurt pKi performance?

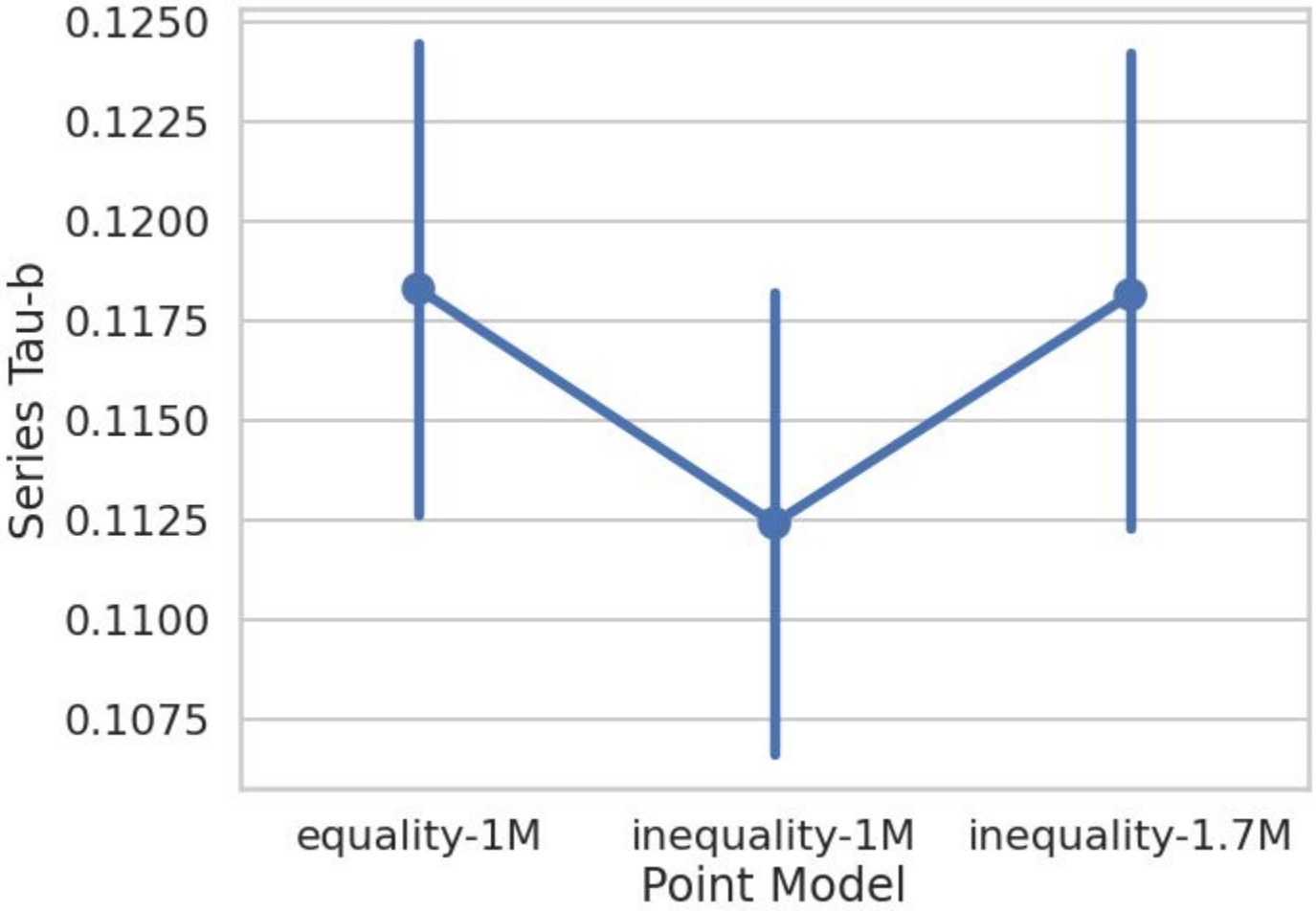
Non-binder Quantitative Data is Scarce

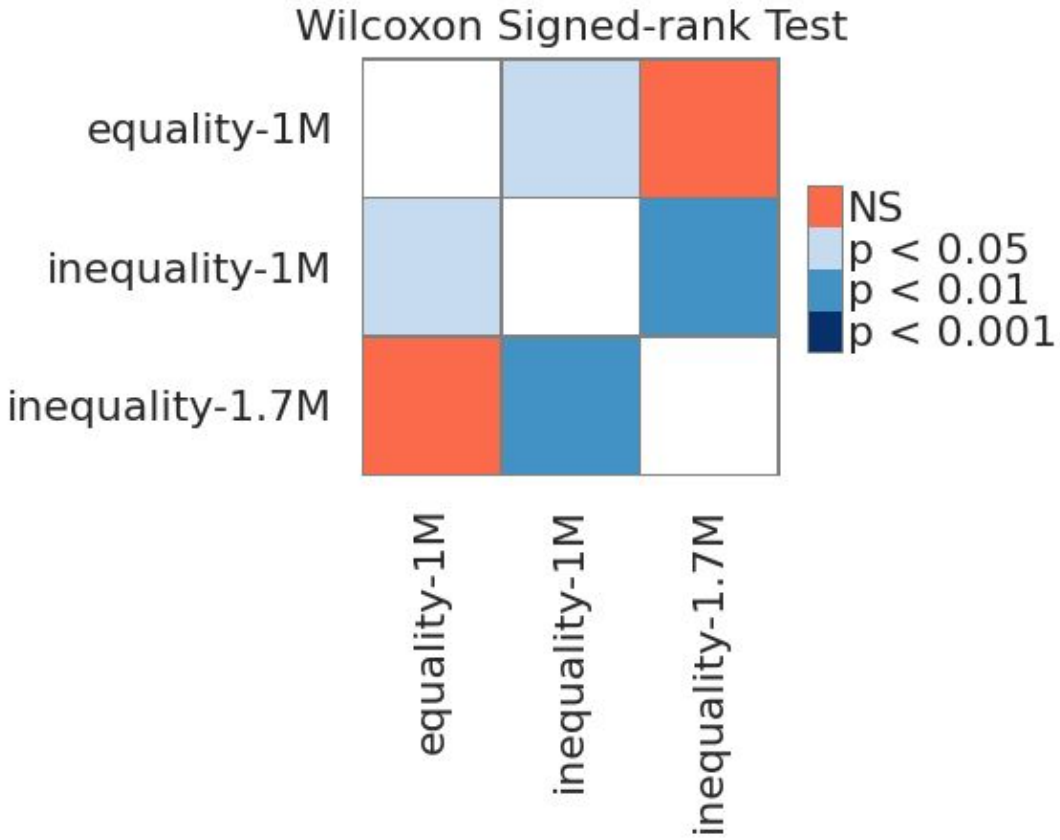
Global pKi model overestimates non-binders

- Initially aimed to use only higher quality data for pKi model
- As you climb the experiment chain – % inhibition to IC_{50} to K_i
  - pKi data becomes more devoid of non-binders
- So pKi models trained with only equality data tends to overestimate non-binders
- How can we get more negative experimental pKi data?
  - Incorporate inequality data
  - Change cost function from MSE to censored (one-sided) MSE

Can Add Inequality Data for Free

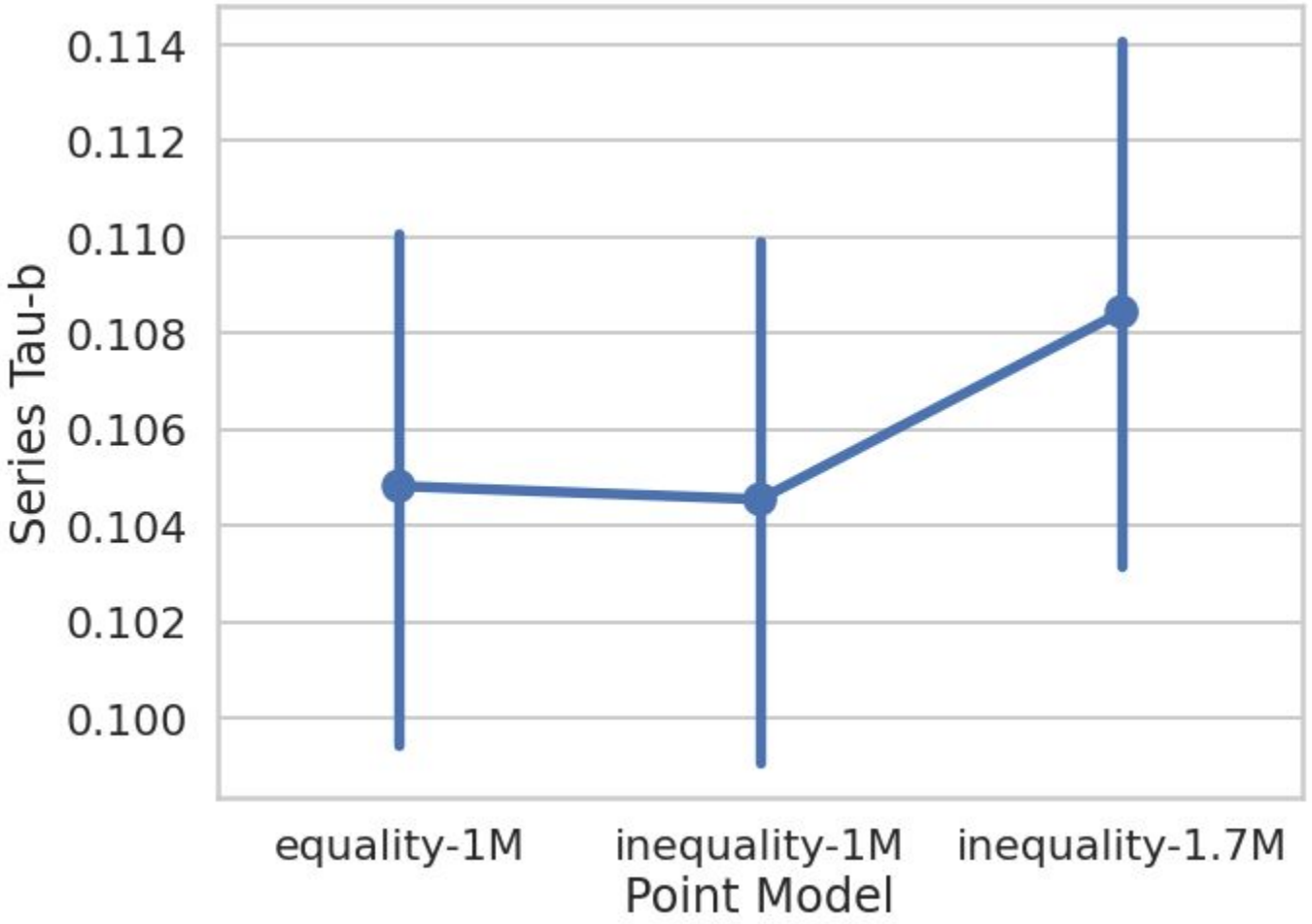
70% more data mostly from non-binders

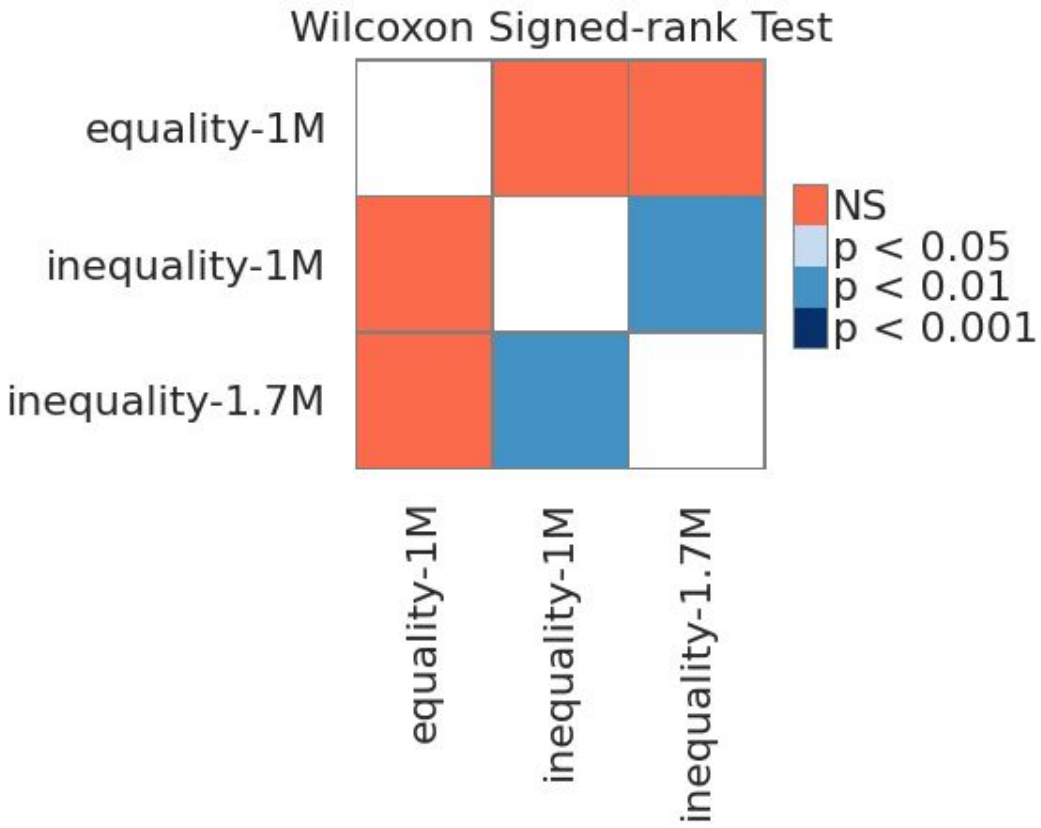




Better on Series Including Inequalities

Using a version of the compound series benchmark that includes inequalities





Takeaways

Conclusions

- We measure performance using a compound series based benchmark evaluated with a Kendall's Tau-b metric
- Our best point-based networks outperform our best voxel-based networks for global pKi prediction at this time
- Target-based (seqsim70) splits are harder to predict than compound-based (Murcko) splits
  - Seqsim70 splits more representative for novel and undruggable targets
- Can add inequality data for free in terms of performance
  - 70% more data
  - Mostly non-binders which are scarce

Acknowledgments

Kate Stafford & Brandon Anderson, plus the excellent team of Atomwise scientists

Our Other Talks & Posters

https://info.atomwise.com/acs_spring2021

Join Us!

https://www.atomwise.com/careers/