Peeking into the Black Box of Molecular Deep Learning

David Koes

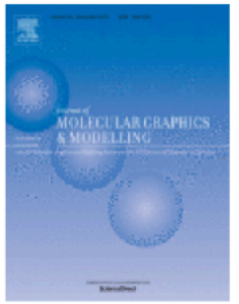
@david_koes

Gordon Research Conference: Computer Aided Drug Design
Mount Snow, VT
July 17, 2019

ELSEVIER

Journal of Molecular Graphics and Modelling

Volume 84, September 2018, Pages 96-108



Visualizing convolutional neural network protein-ligand scoring

Joshua Hochuli, Alec Helbling, Tamar Skaist, Matthew Ragoza, David Ryan Koes

But first…

Novel Insights from Data Visualization with Pedagogical Active Learning

Pedagogical Active Learning

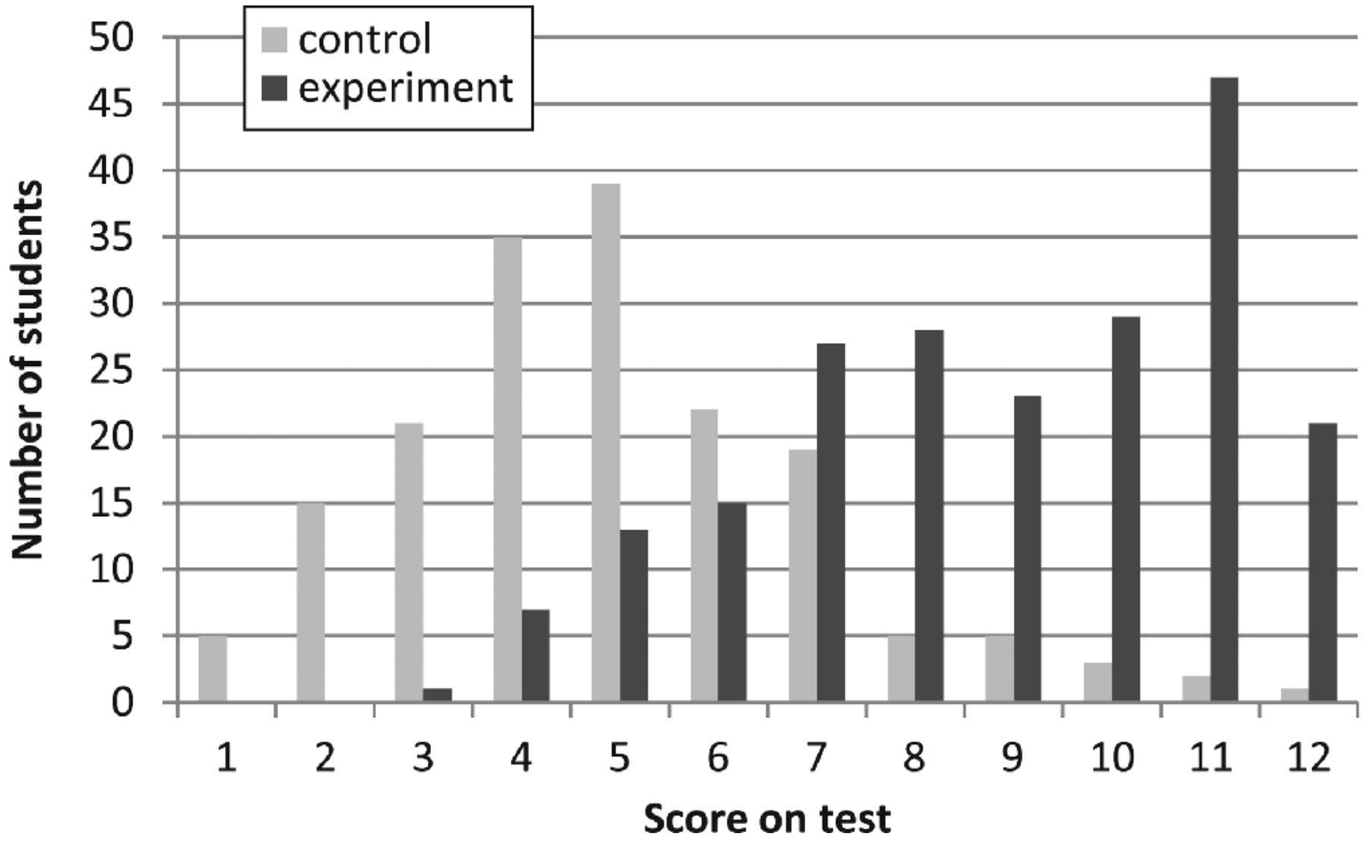
REPORT

Improved Learning in a Large-Enrollment Physics Class

Louis Deslauriers[1,2], Ellen Schelew[2], Carl Wieman[*,†‡]

+ See all authors and affiliations

Science 13 May 2011:
Vol. 332, Issue 6031, pp. 862-864
DOI: 10.1126/science.1201783



Pedagogical Active Learning

ELSEVIER

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect

Nathaniel J. Hunsu*, Olusola Adesope, Dan James Bayly

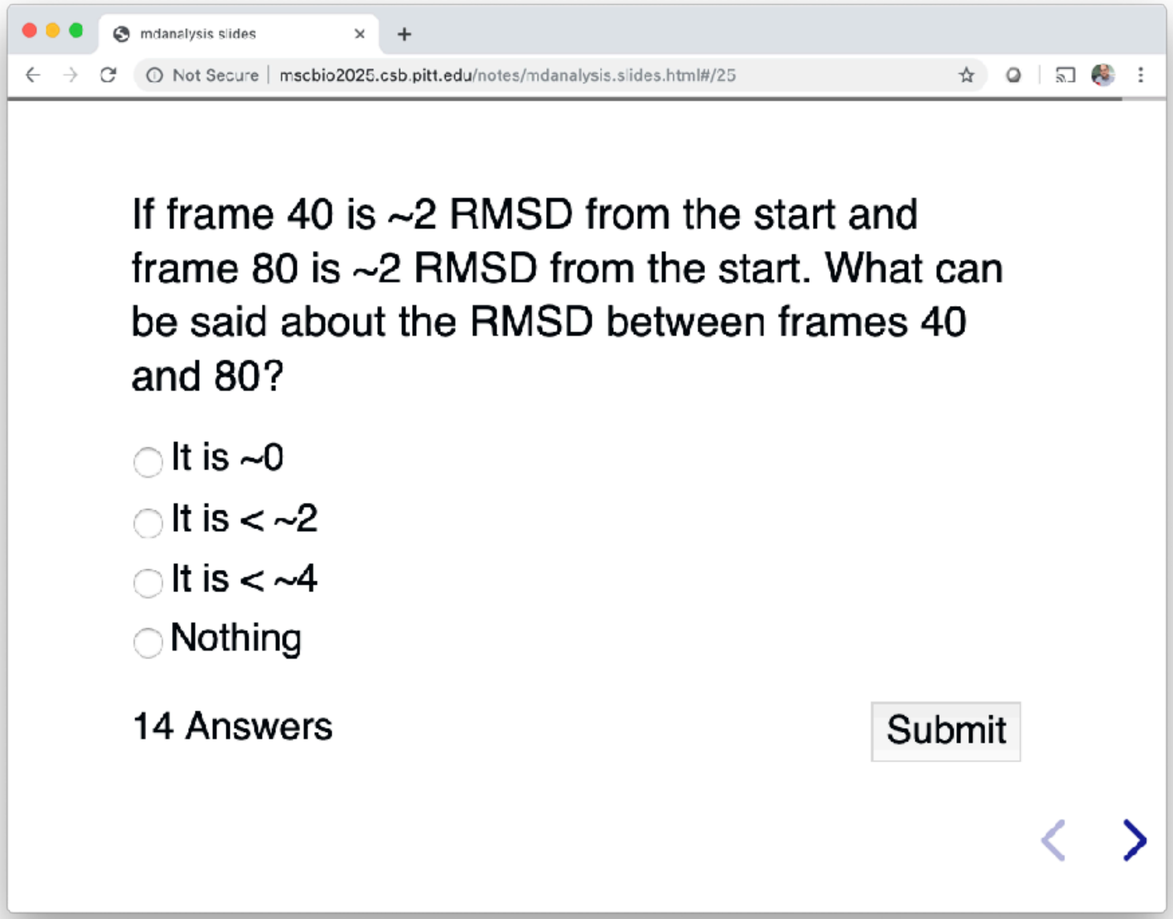
Educational Leadership, Sport Studies, Educational and Counseling Psychology, Washington State University, Pullman, WA 99164-4530, USA

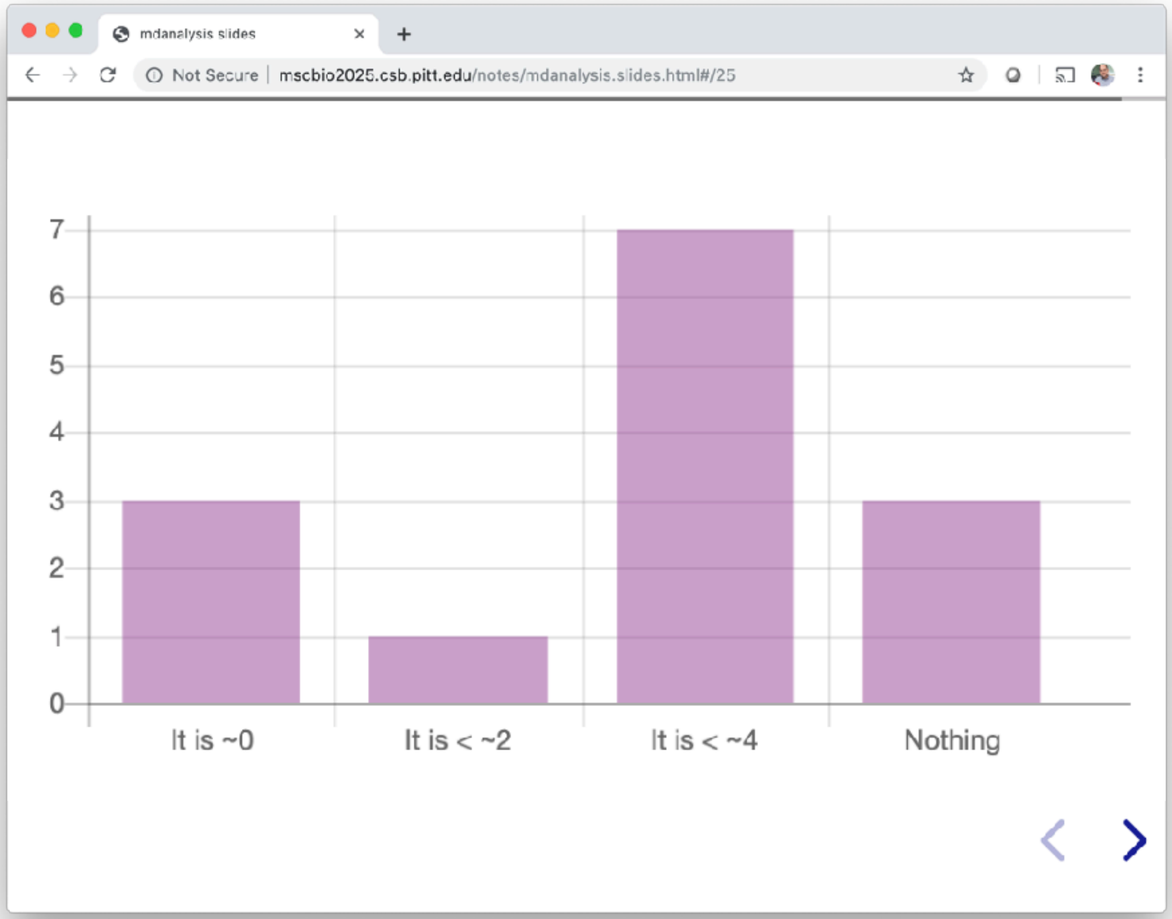
"Overall, we found small but significant effects of using ARS-based technologies on a number of desirable cognitive and non-cognitive learning outcomes."

Go to this URL: http://3dmol.csb.pitt.edu/viewer.html?session=GRC

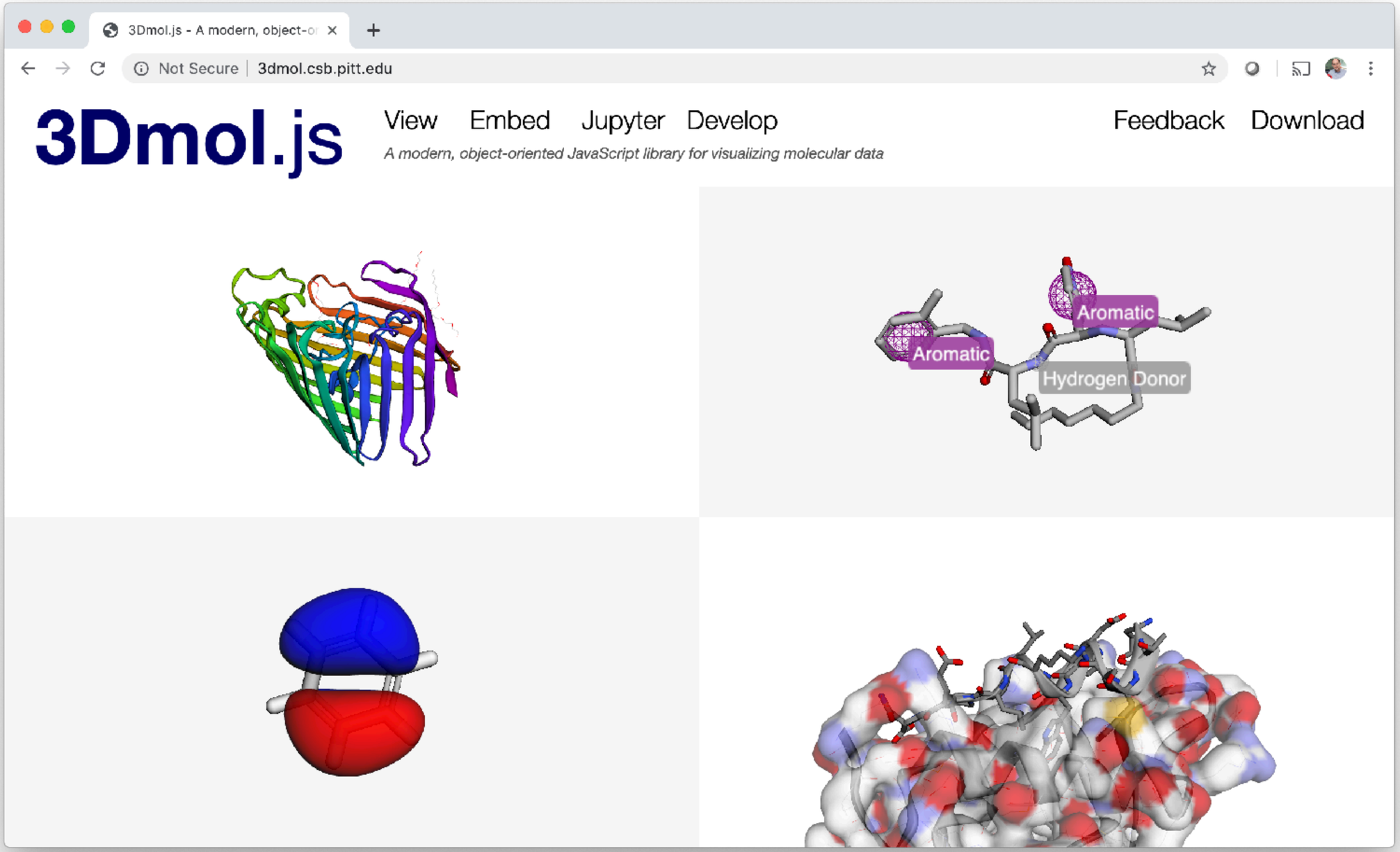
Pedagogical Active Learning

https://github.com/dkoes/asker.js





Go to this URL: http://3dmol.csb.pitt.edu/viewer.html?session=GRC



Go to this URL: http://3dmol.csb.pitt.edu/viewer.html?session=GRC

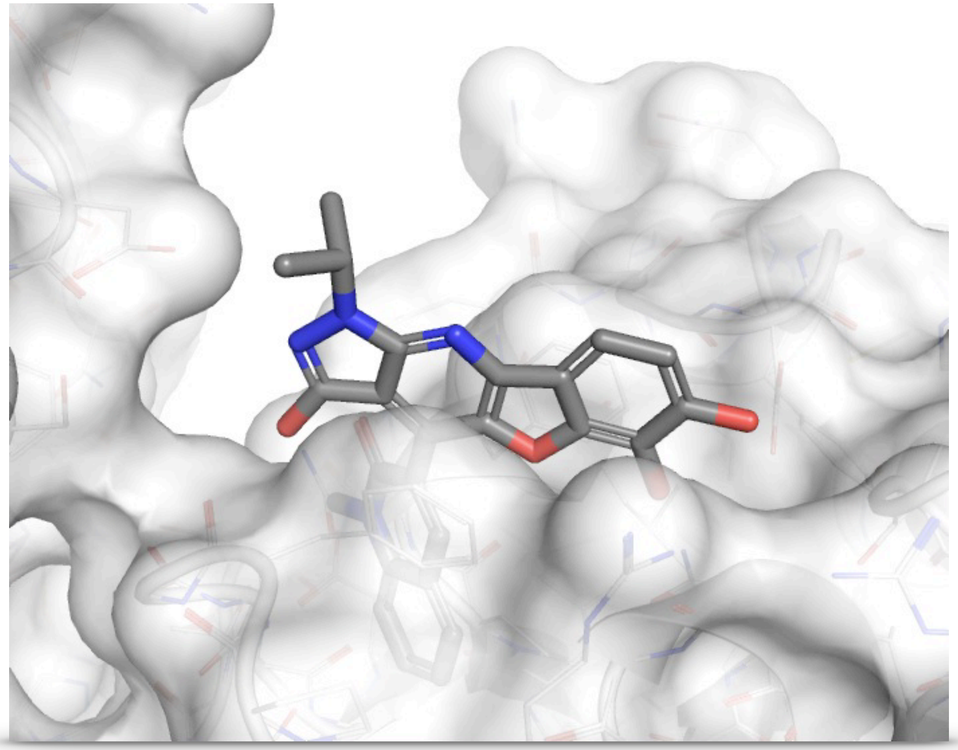
Molecular Active Learning

Demo

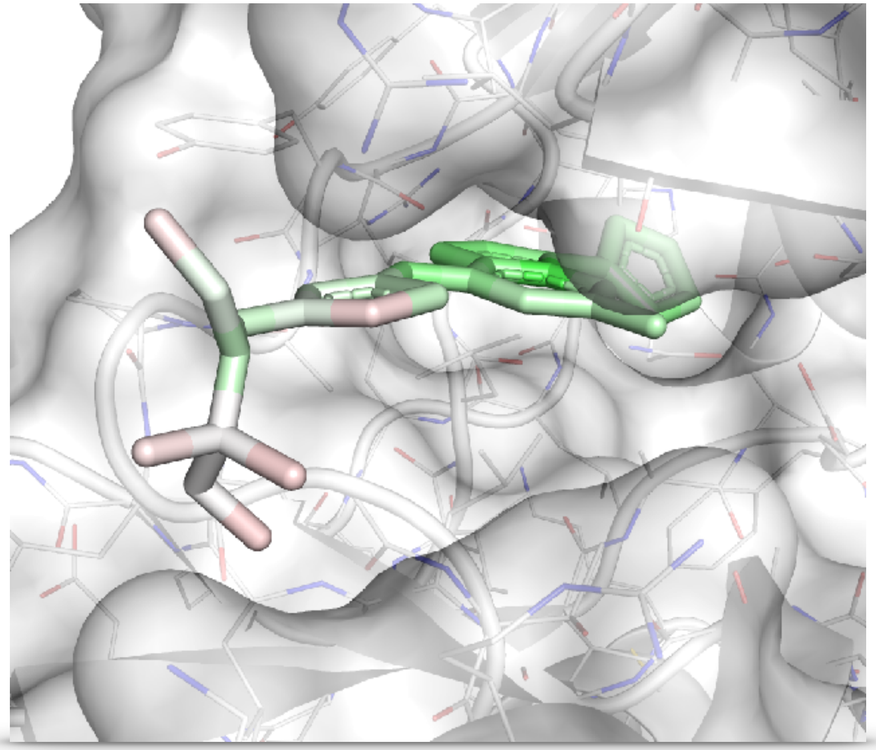
Go to this URL: http://3dmol.csb.pitt.edu/viewer.html?session=GRC

Structure Based Drug Design

Pose Prediction Binding Discrimination Affinity Prediction

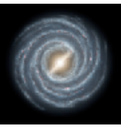


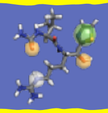
Virtual Screening



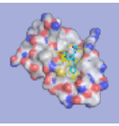
Lead Optimization

Drug Discovery Funnel

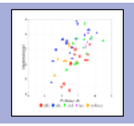
Purchasable  Accessible

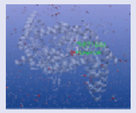


Matching

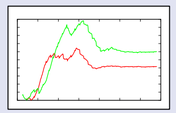


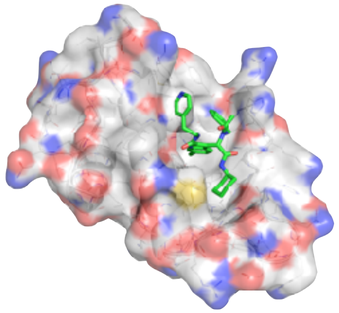
Scoring

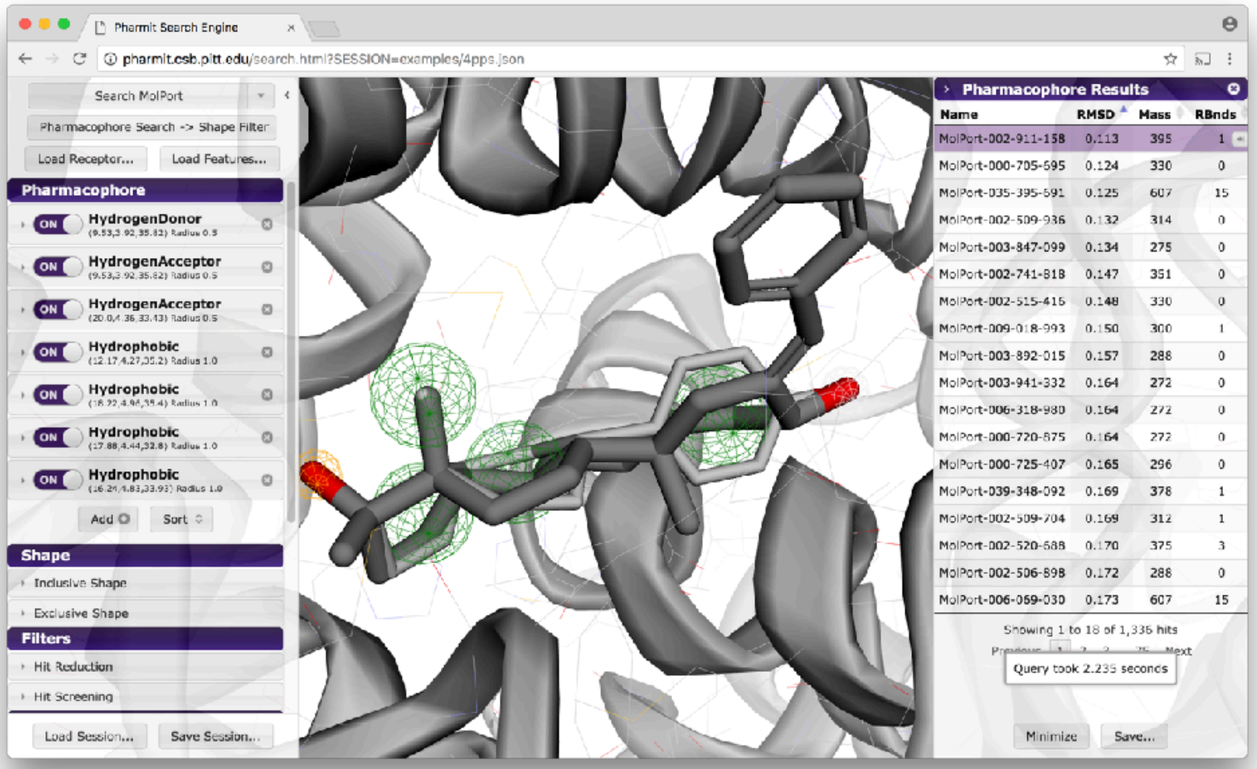




Dynamics

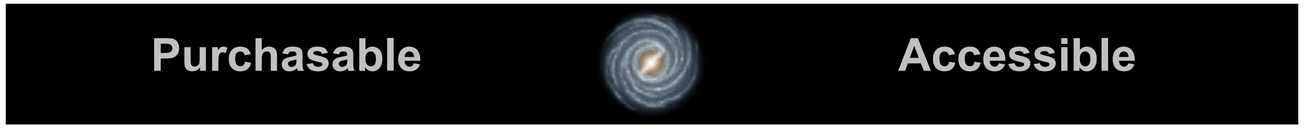


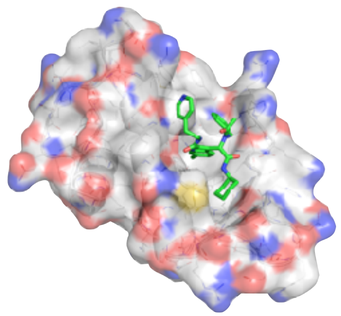


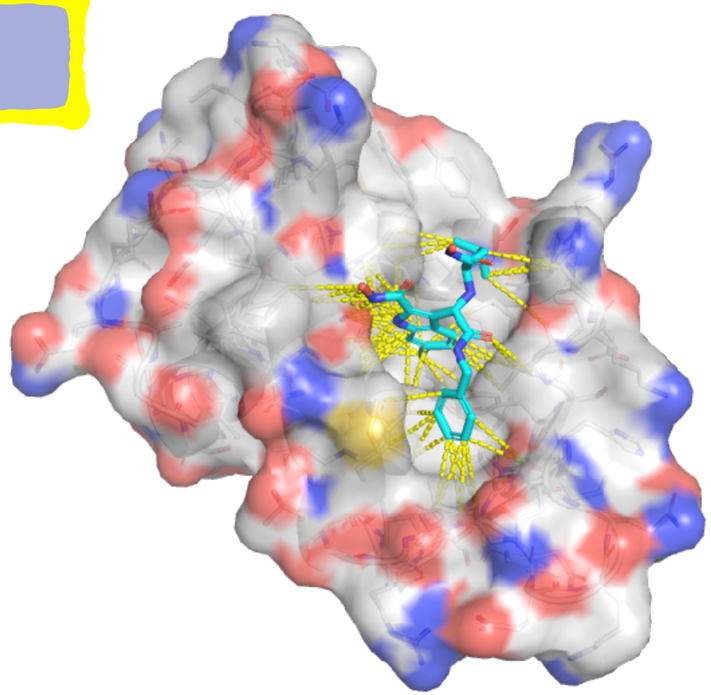


http://pharmit.csb.pitt.edu

Drug Discovery Funnel



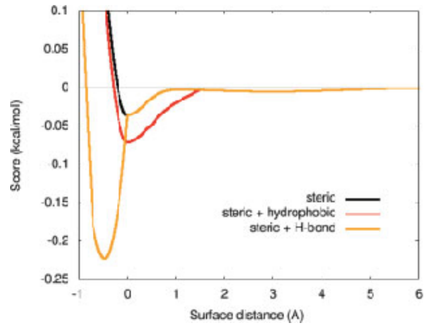




$$\text{gauss}_1(d) = w_{\text{guass}_1} e^{-(d/0.5)^2}$$
$$\text{gauss}_2(d) = w_{\text{guass}_2} e^{-((d-3)/2)^2}$$
$$\text{repulsion}(d) = \begin{cases} w_{\text{repulsion}} d^2 & d < 0 \\ 0 & d \geq 0 \end{cases}$$

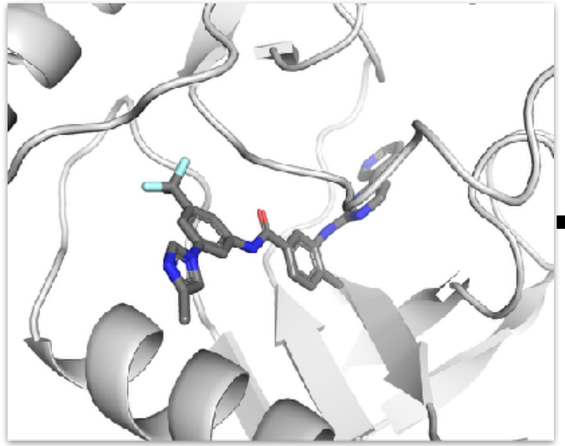
$$\text{hydrophobic}(d) = \begin{cases} w_{\text{hydrophobic}} & d < 0.5 \\ 0 & d > 1.5 \\ w_{\text{hydrophobic}}(1.5 - d) & otherwise \end{cases}$$

$$\text{hbond}(d) = \begin{cases} w_{\text{hbond}} & d < -0.7 \\ 0 & d > 0 \\ w_{\text{hbond}}(-\frac{10}{7} d) & otherwise \end{cases}$$



O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *Journal of Computational Chemistry* 31 (2010) 455-461

Protein-Ligand Scoring



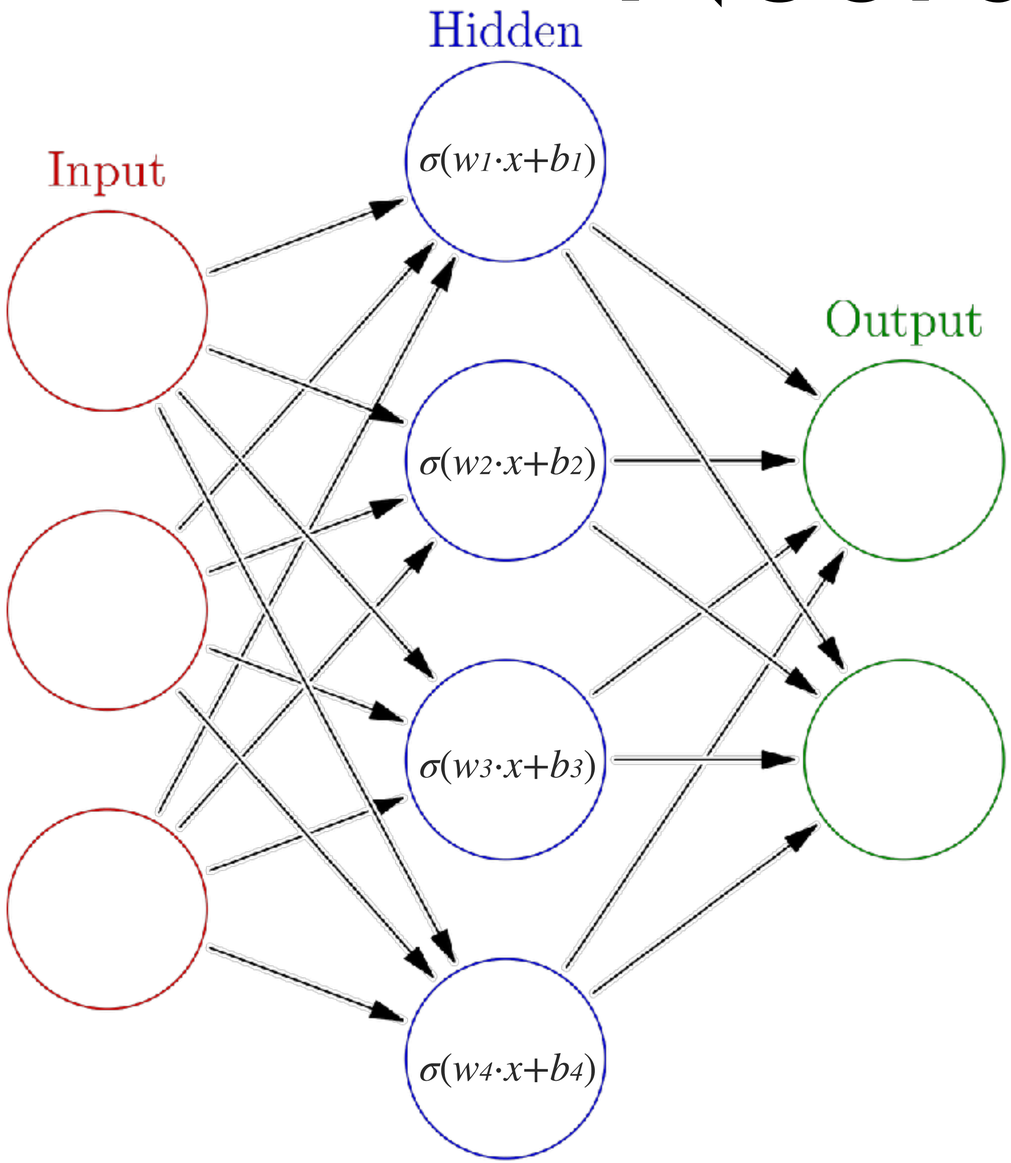
Model

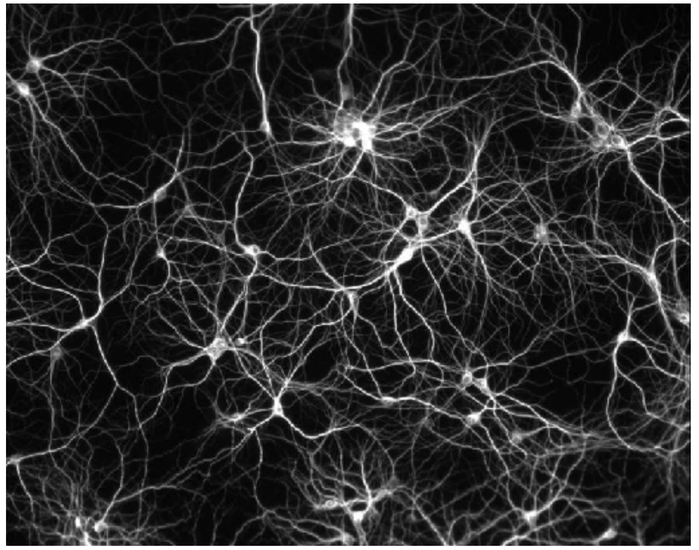
Pose Prediction

Binding Discrimination

Affinity Prediction

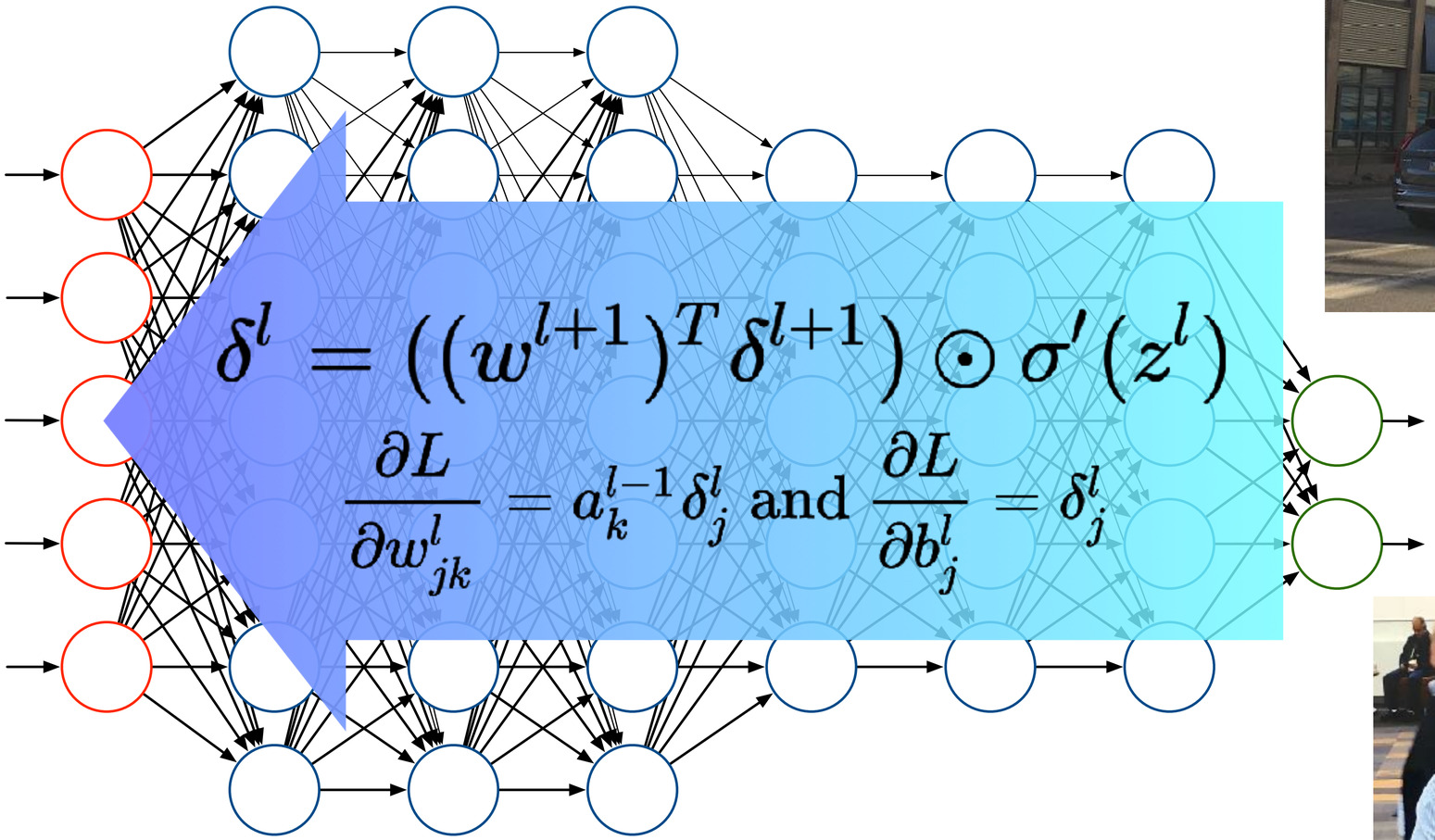
Neural Networks





The **universal approximation theorem** states that, under reasonable assumptions, a feedforward **neural network** with a finite number of nodes **can approximate any continuous** function to within a given error over a bounded input domain.

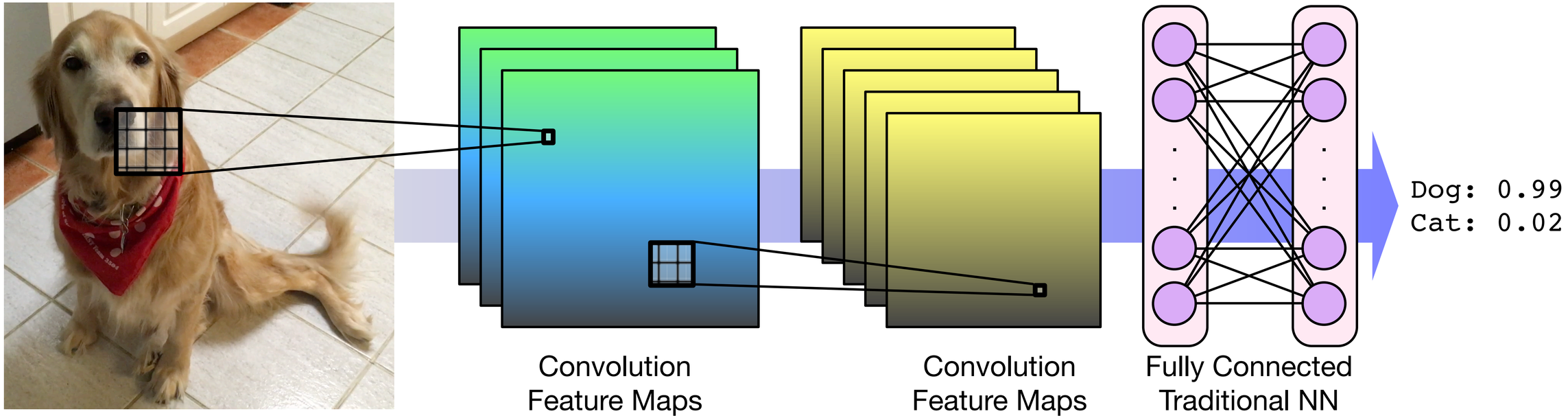
Deep Learning

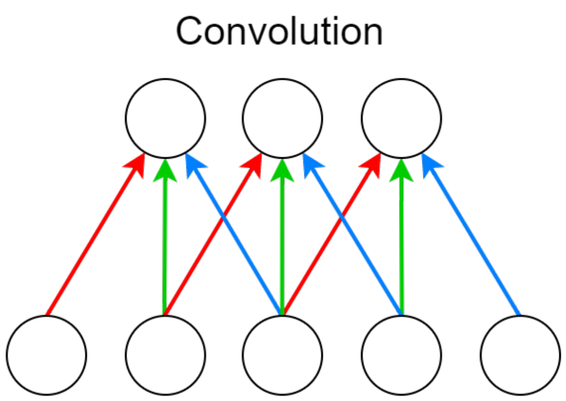


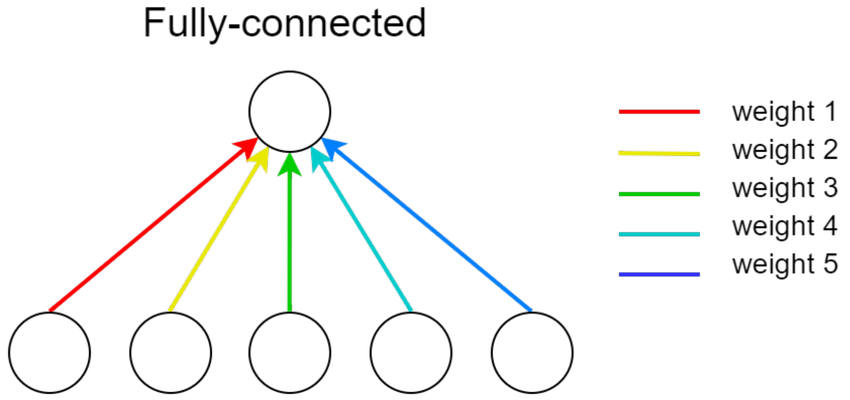
$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$$

$$\frac{\partial L}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j \text{ and } \frac{\partial L}{\partial b^l_j} = \delta^l_j$$

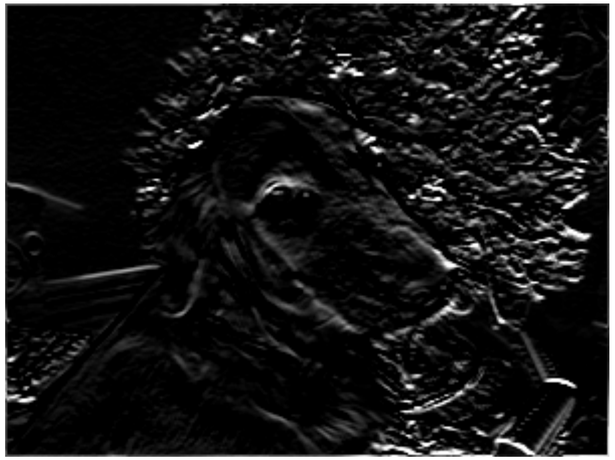
Convolutional Neural Networks



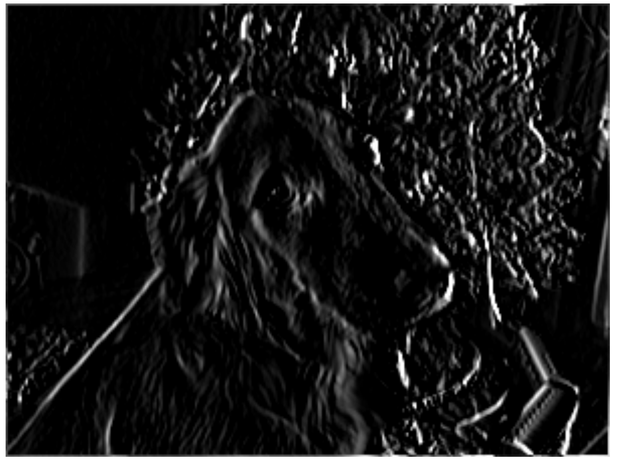




Convolutional Filters



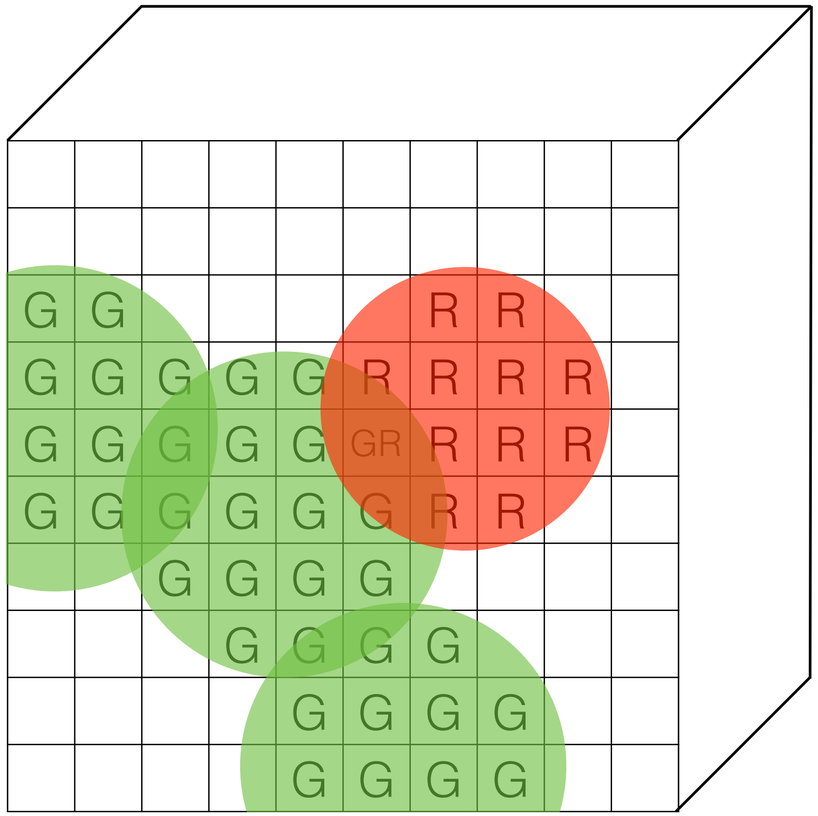
-1	-1	-1
0	0	0
1	1	1



-1	0	1
-1	0	1
-1	0	1

-1	-1	-1
-1	8	-1
-1	-1	-1

Protein-Ligand Representation



(R,G,B) pixel →

(Carbon, Nitrogen, Oxygen,…) **voxel**

The only parameters for this representation are the choice of **grid resolution**, **atom density**, and **atom types**.

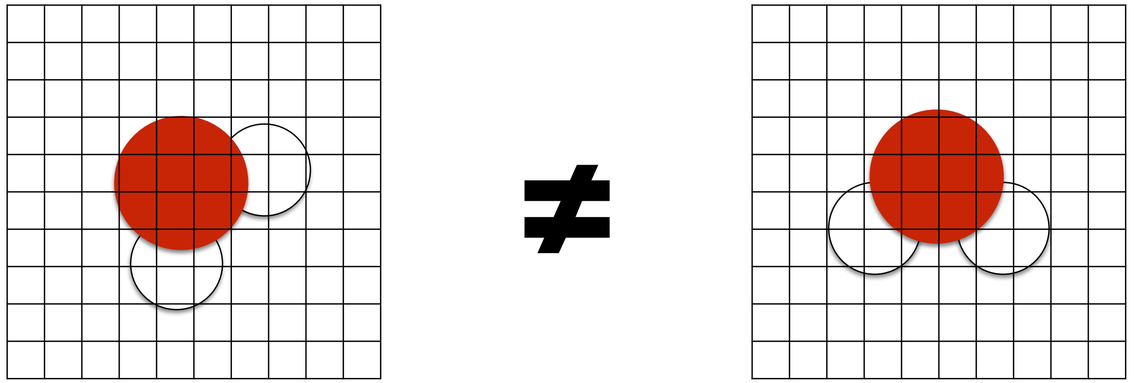
Why Grids?

Cons

- *coordinate frame dependent*
- pairwise interactions not explicit

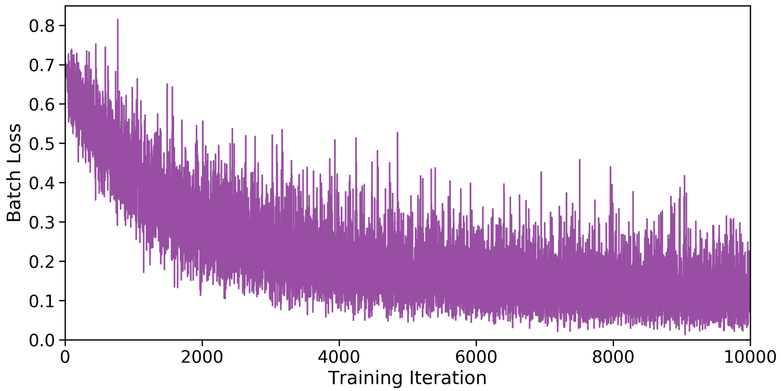
Pros

- clear spatial relationships
- amazingly parallel
- easy to interpret

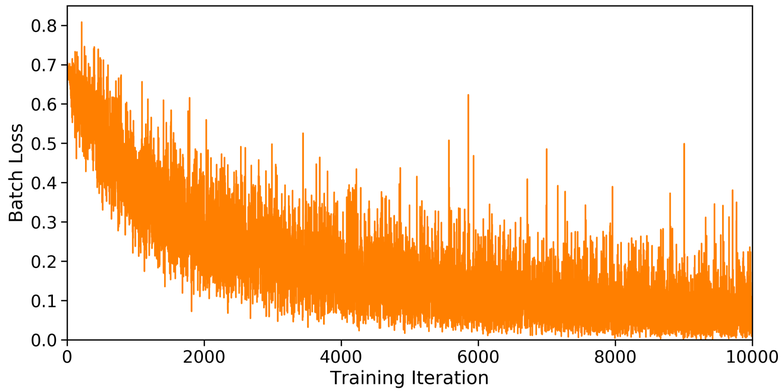


libmolgrid

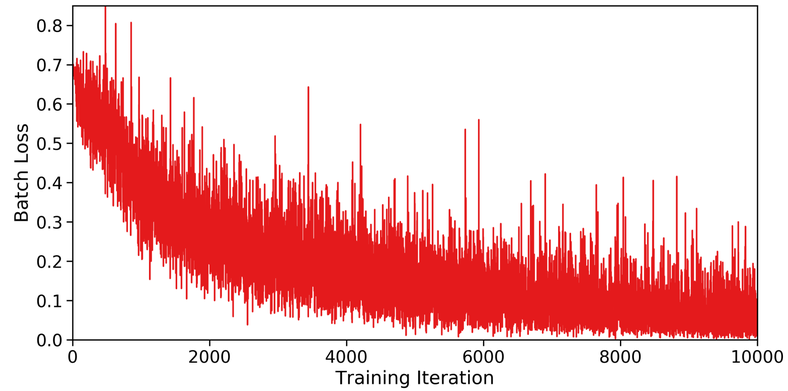
Caffe Training



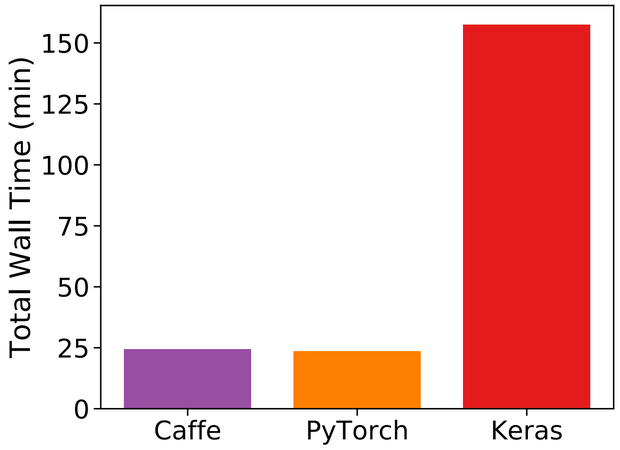
PyTorch Training



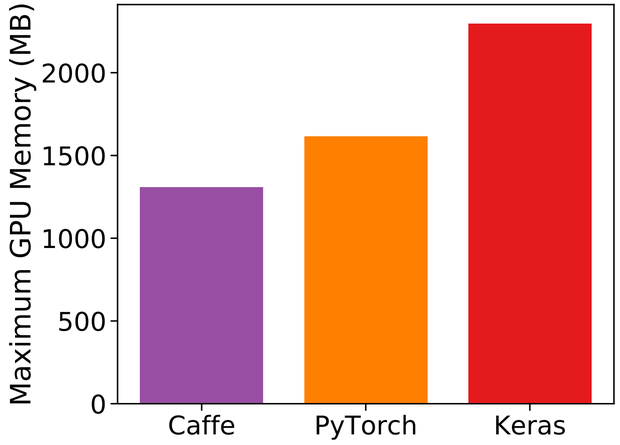
Keras Training



GPU Performance



GPU Memory Utilization



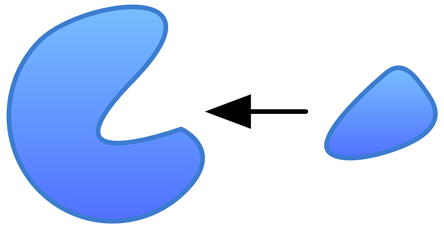

```
e = molgrid.ExampleProvider(balanced=True,shuffle=True)
e.populate('examples.txt')

gmaker = molgrid.GridMaker()

batch = e.next_batch(batch_size)
gmaker.forward(batch, input_tensor,
               random_translation=0, random_rotation=True)
```

github.com/gnina/libmolgrid

Training

PDBbind 2016 refined set

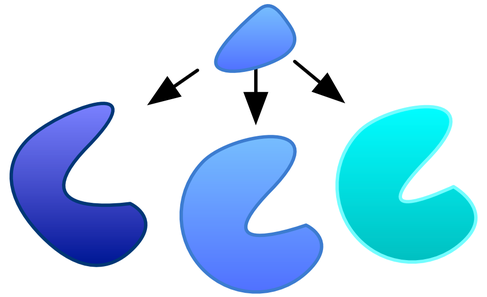
- 4056 protein-ligand complexes
- diverse targets
- wide range of affinities
- generate poses with AutoDock Vina
- include minimized crystal pose



Redocked Training Set

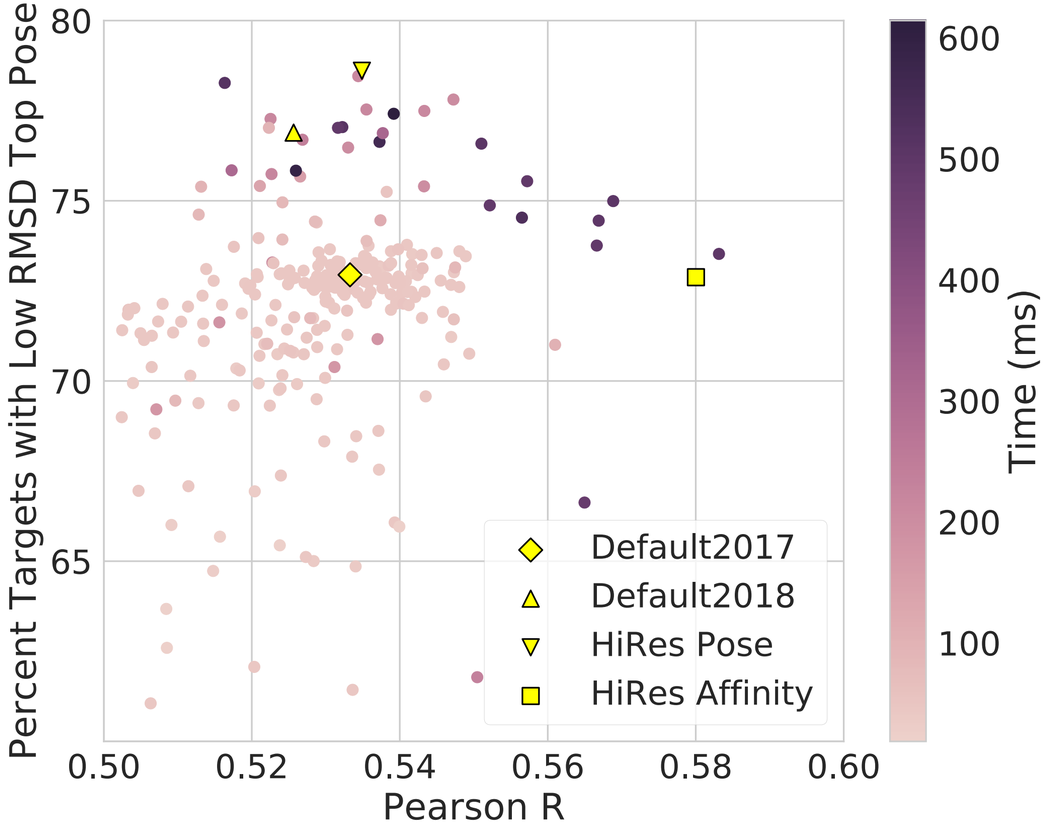
Pocketome

- 2923 distinct pockets
- 27,142 receptor structures
- 4,138,117 non-redundant poses
- generate poses with AutoDock Vina
- include minimized crystal pose

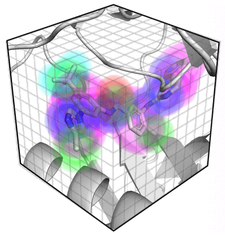


Crossdocked Training Set

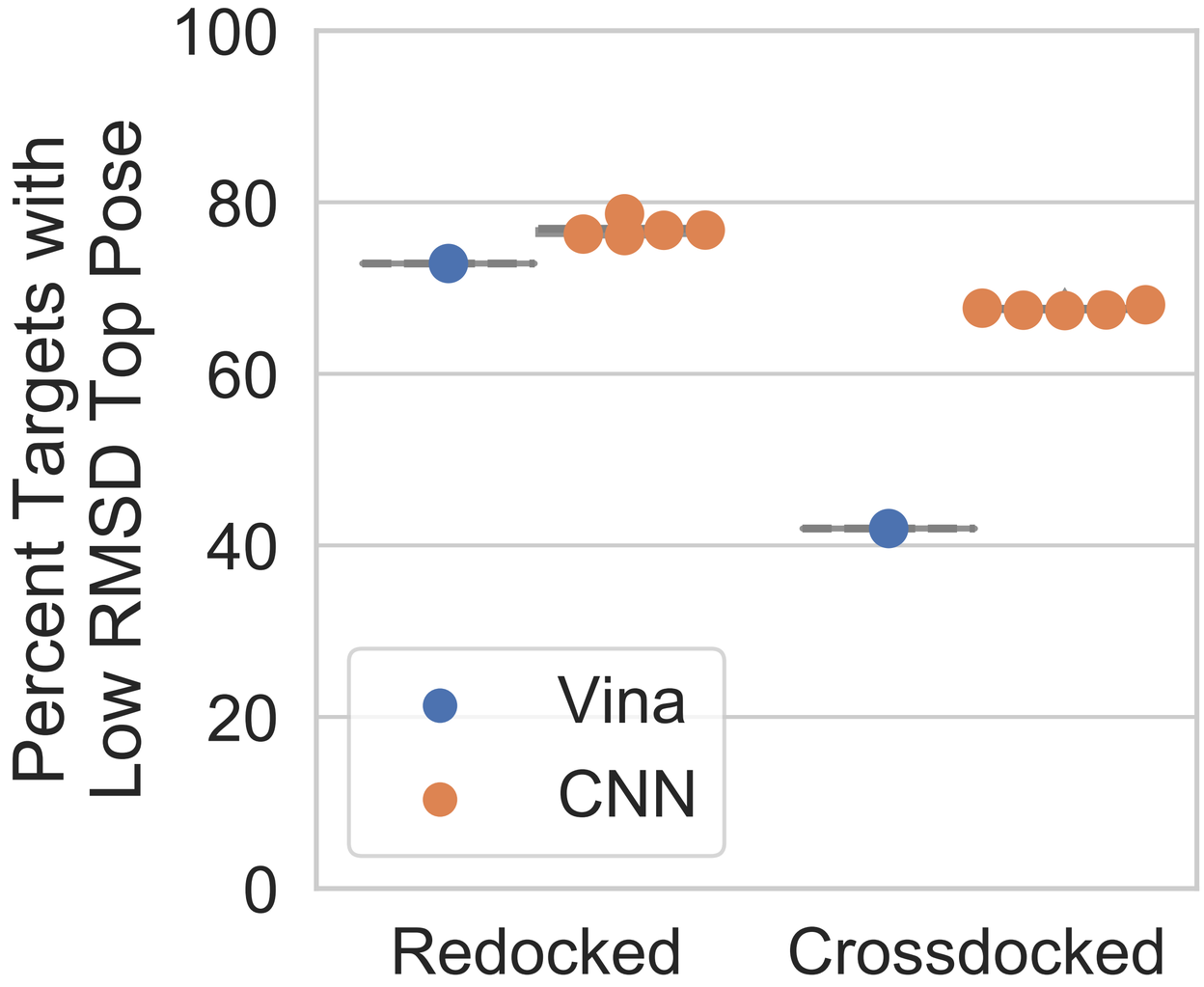
Optimized Models

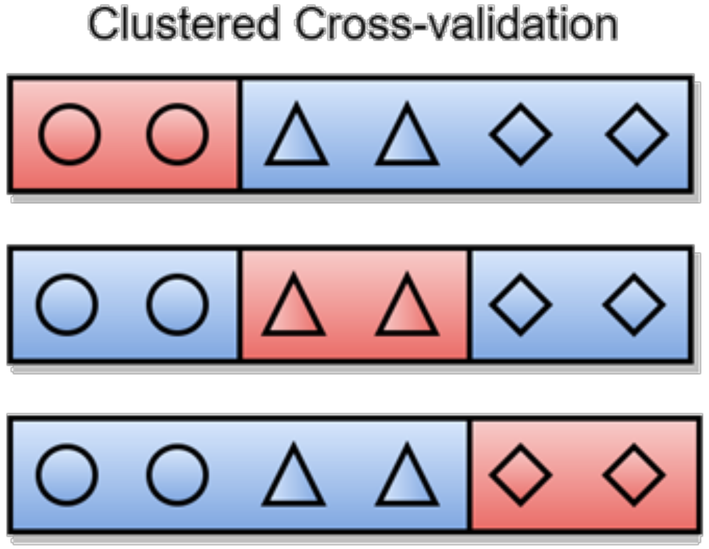


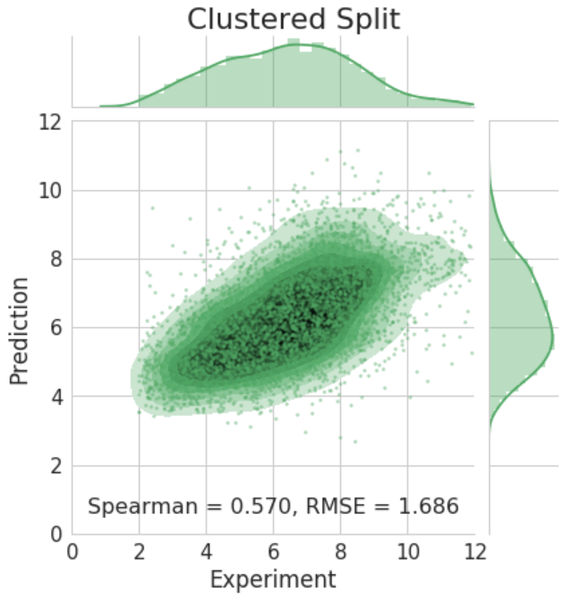
Default2018

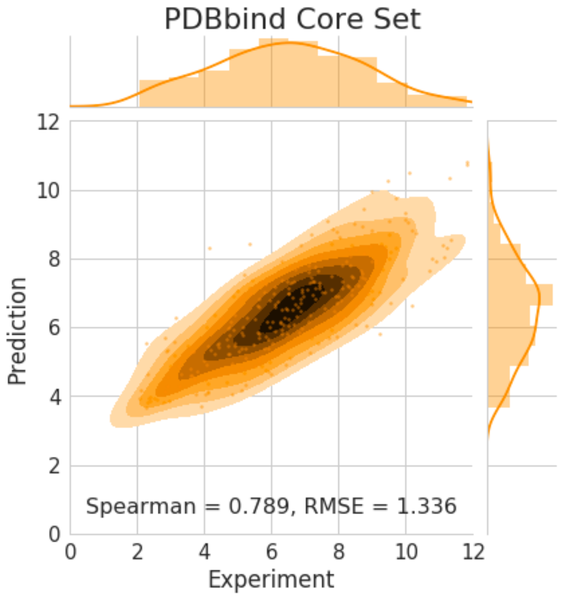


Results

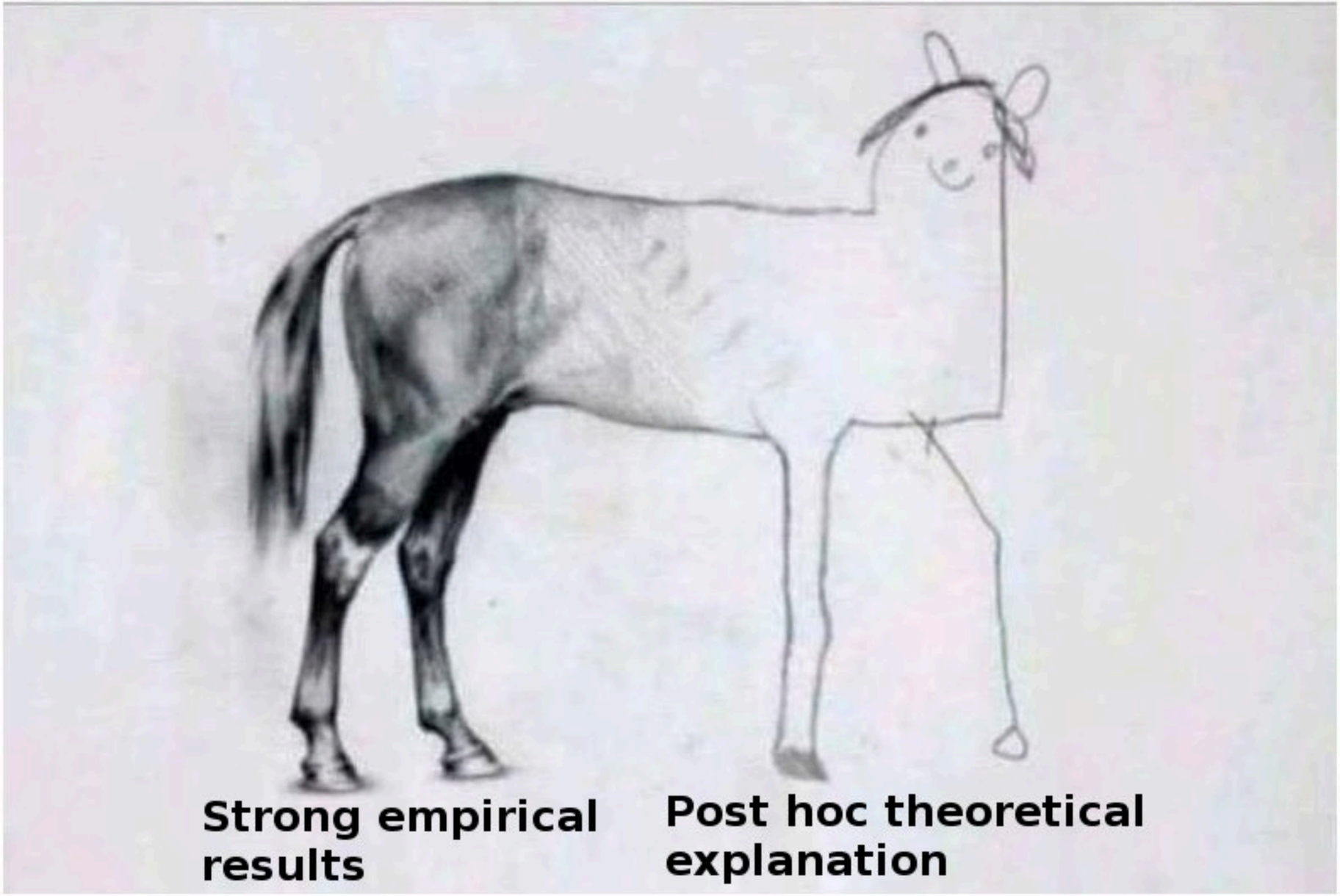

Percent Targets with Low RMSD Top Pose
100
80
60
40
20
0
Vina
CNN
Redocked
Crossdocked


Clustered Cross-validation

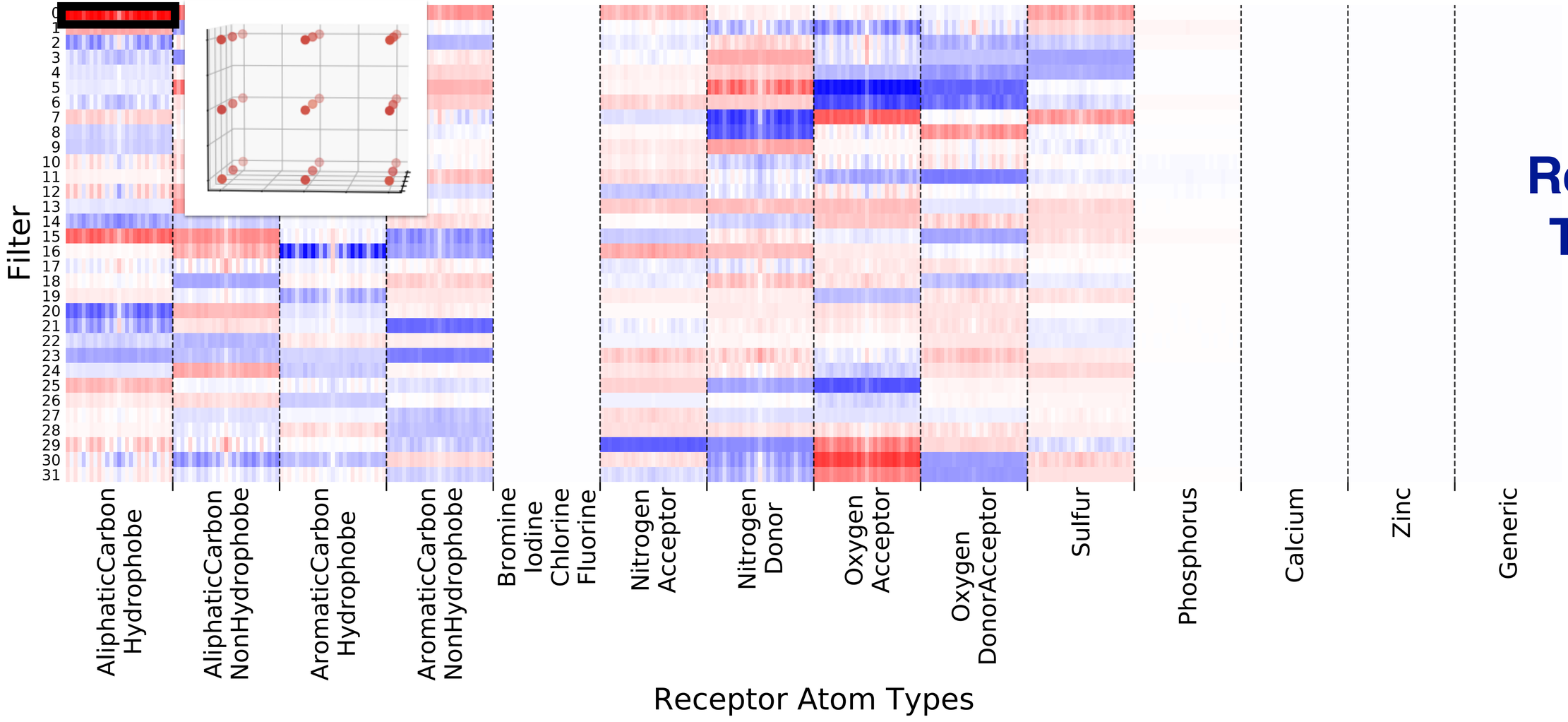

Clustered Split
Prediction
Experiment
Spearman = 0.570, RMSE = 1.686


PDBbind Core Set
Prediction
Experiment
Spearman = 0.789, RMSE = 1.336

Anatomy of a deep learning paper

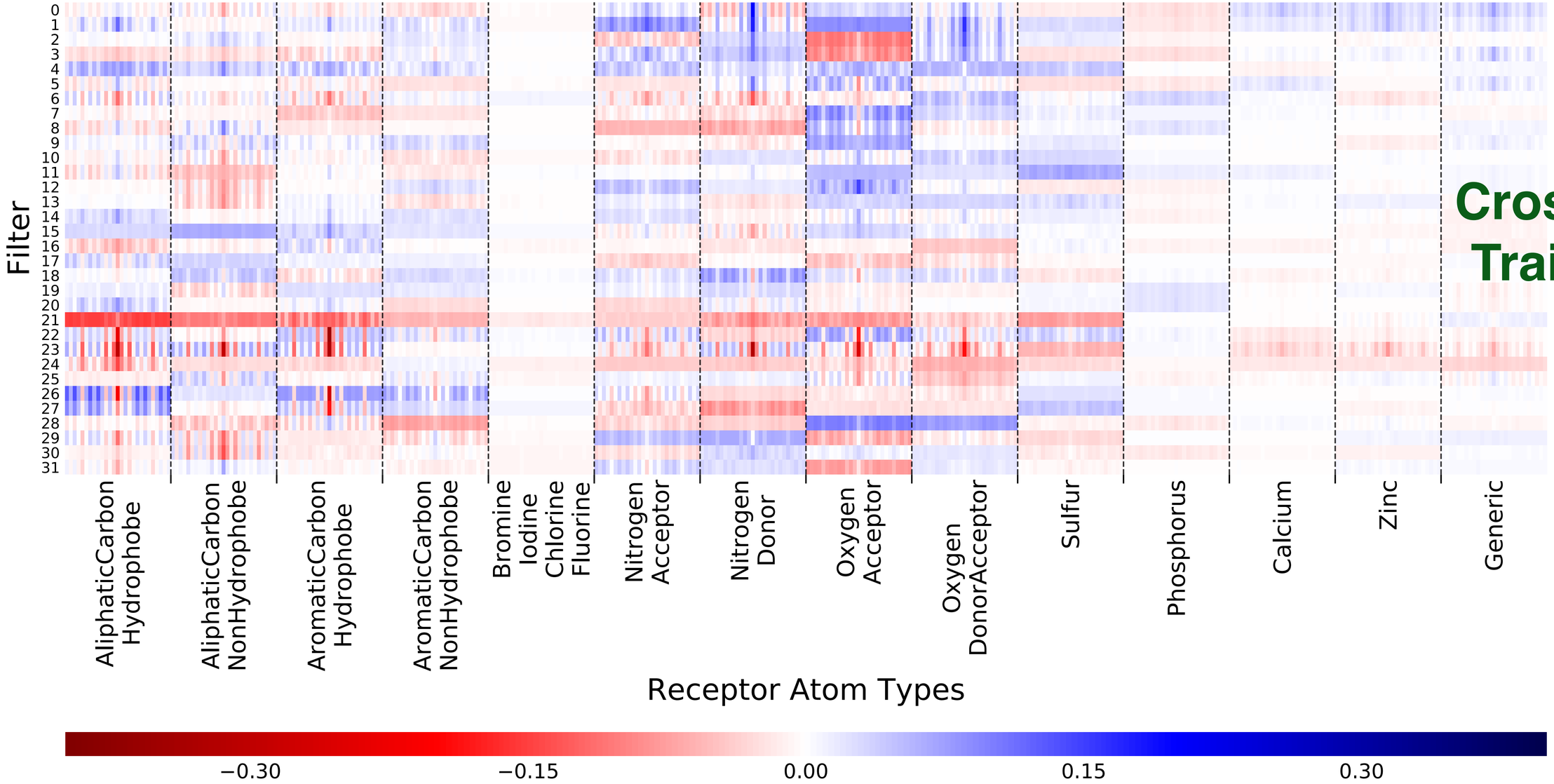


Visualizing the Network



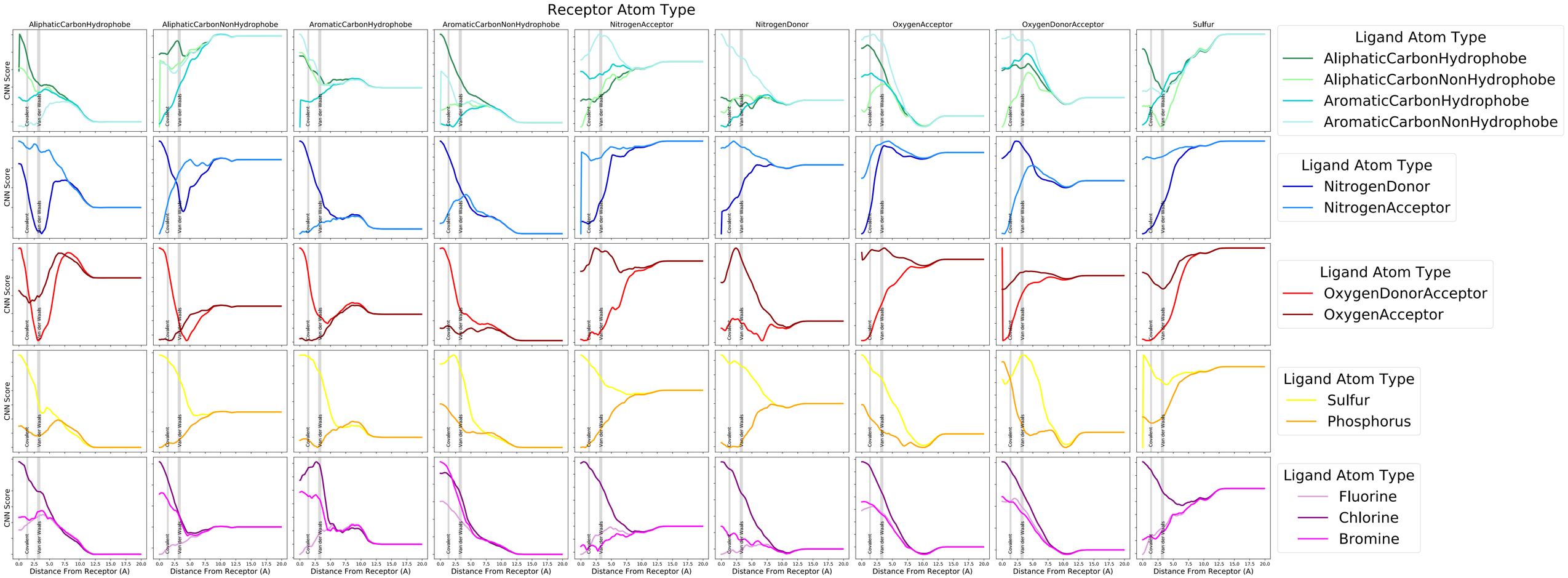
Redocked Training Set

Visualizing the Network

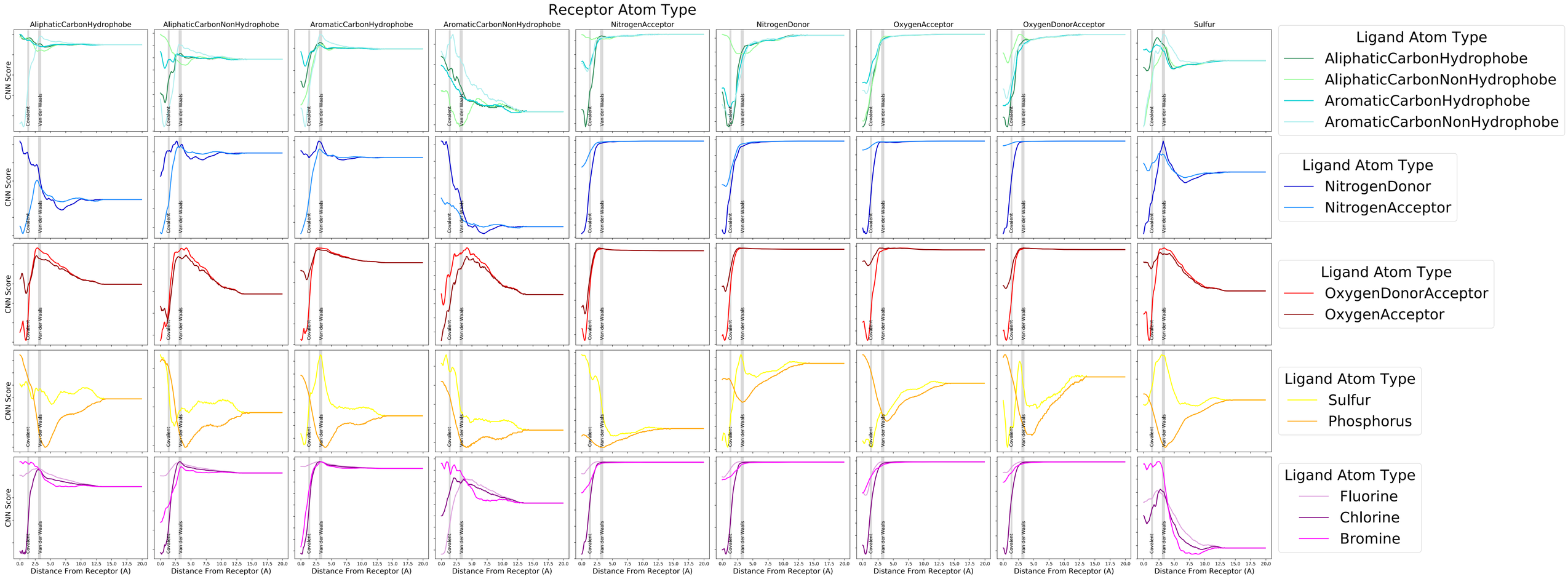


Crossdocked Training Set

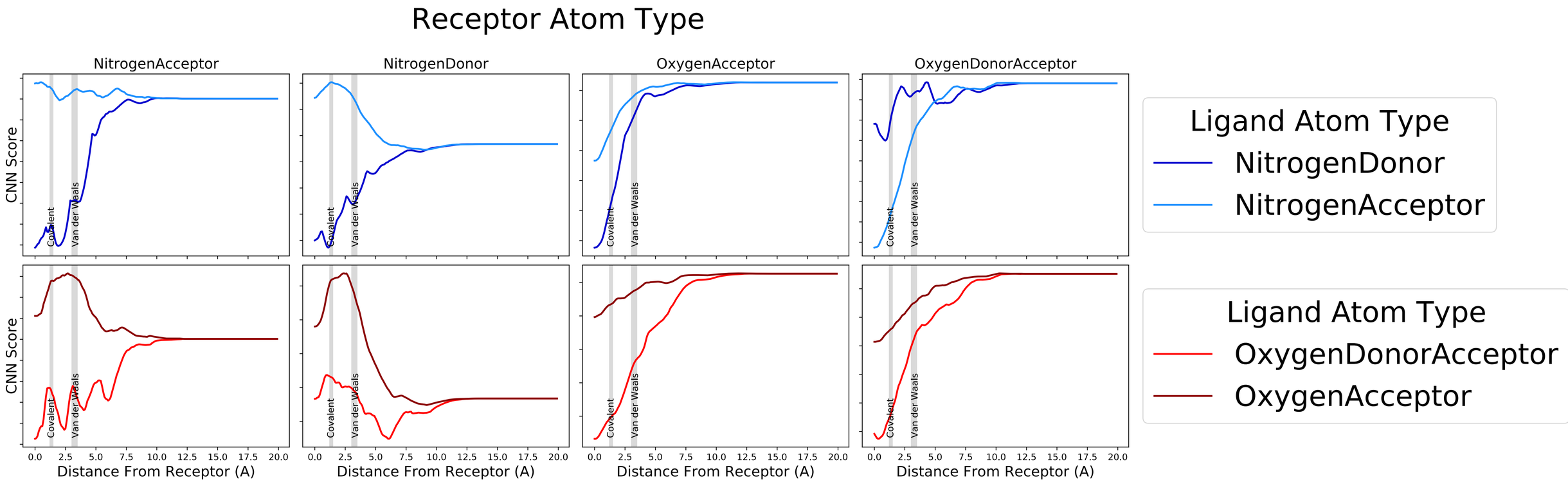
Visualizing with Atomistic Probes



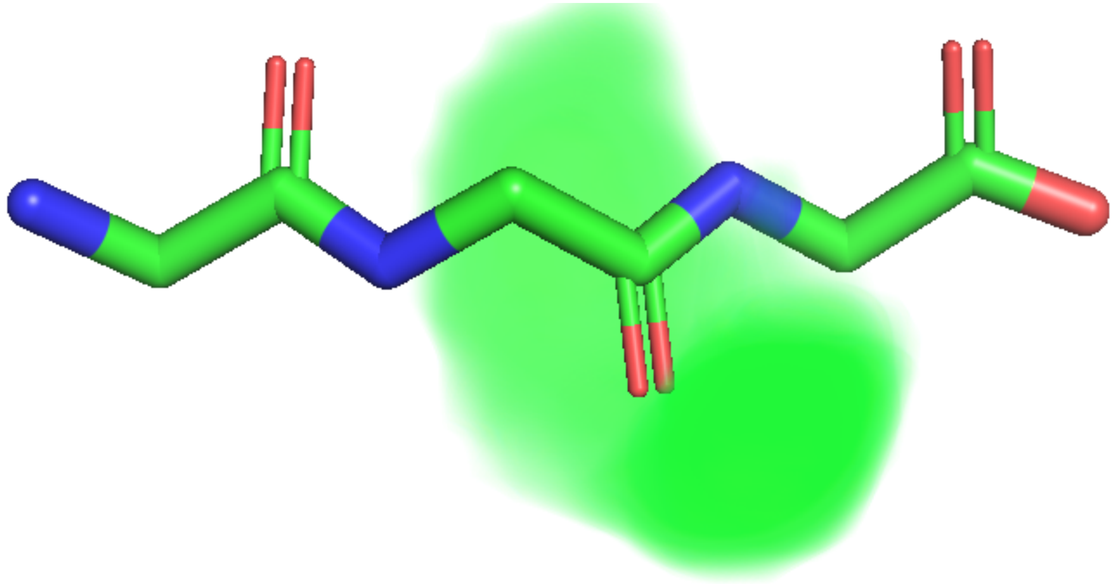
Visualizing with Atomistic Probes



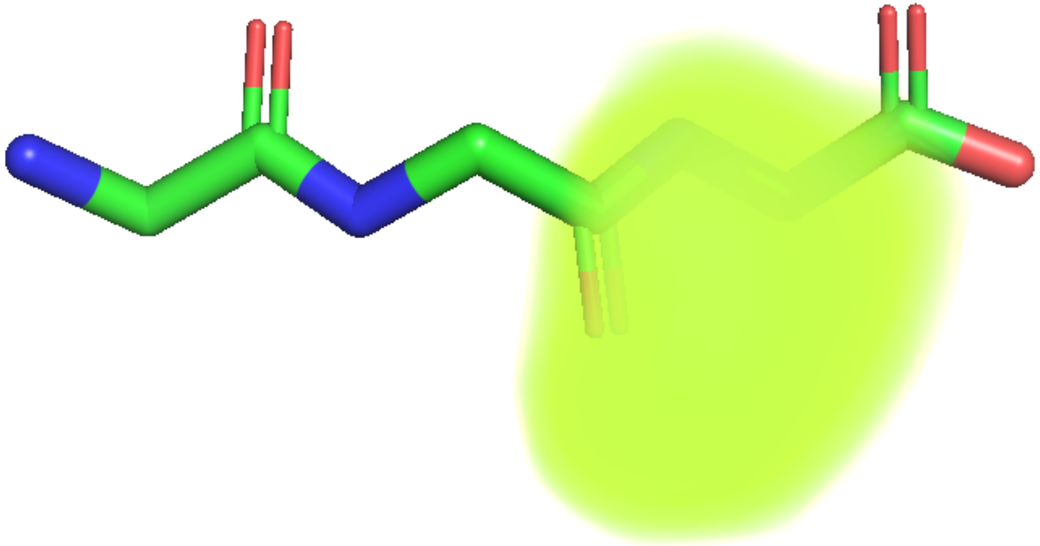
Hydrogen Bonds... or Not



Visualizing with Atomistic Probes

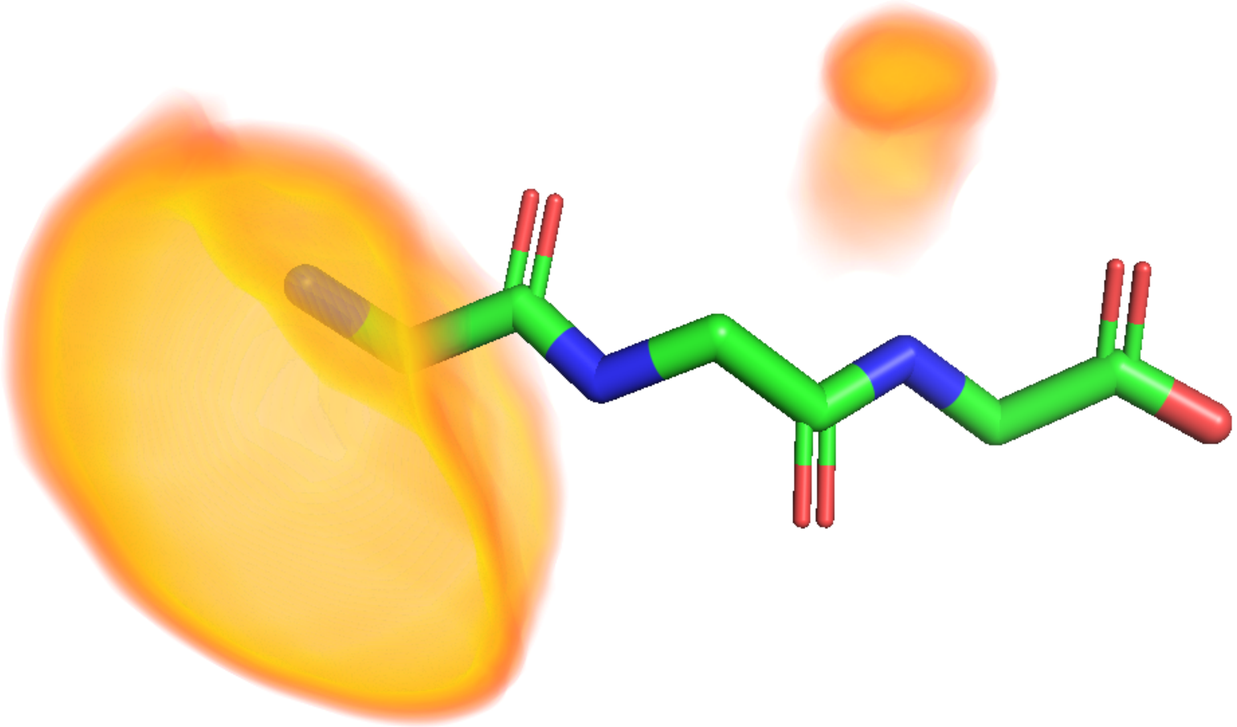


Aliphatic Carbon

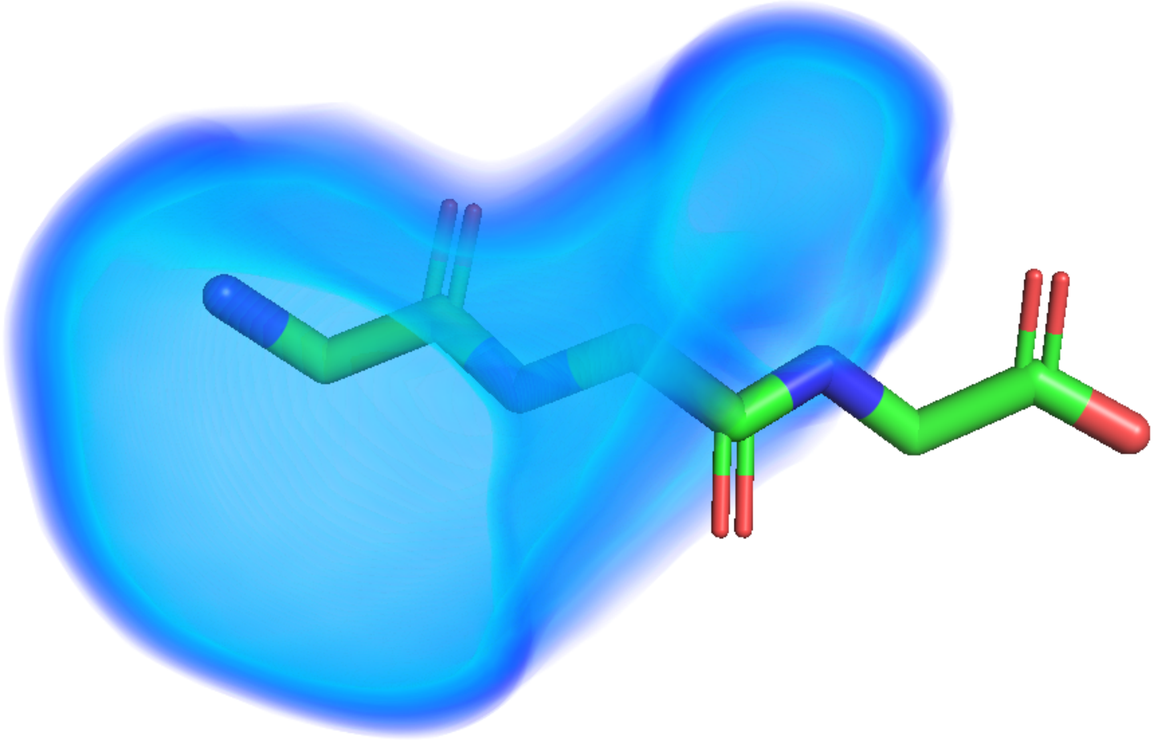


Aromatic Carbon

Visualizing with Atomistic Probes

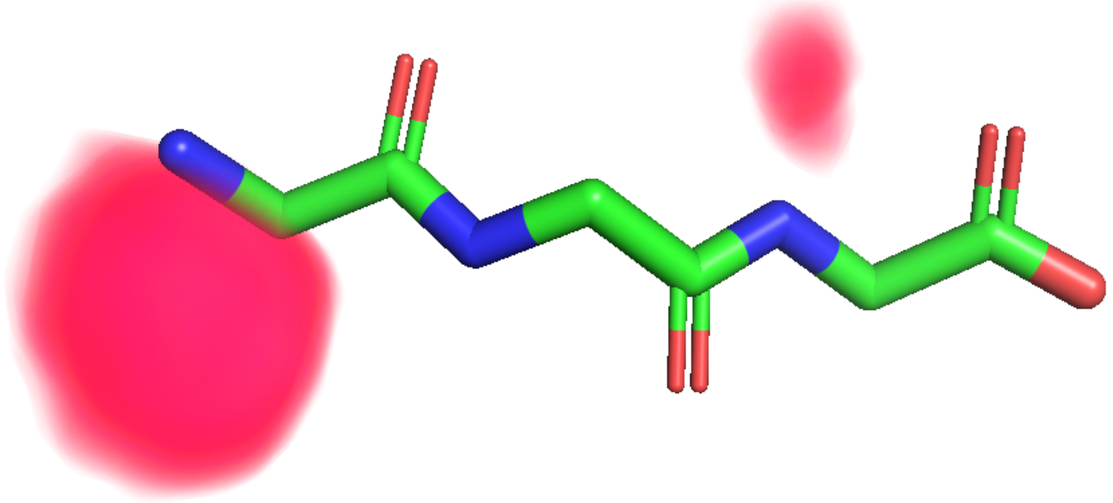


Oxygen Acceptor

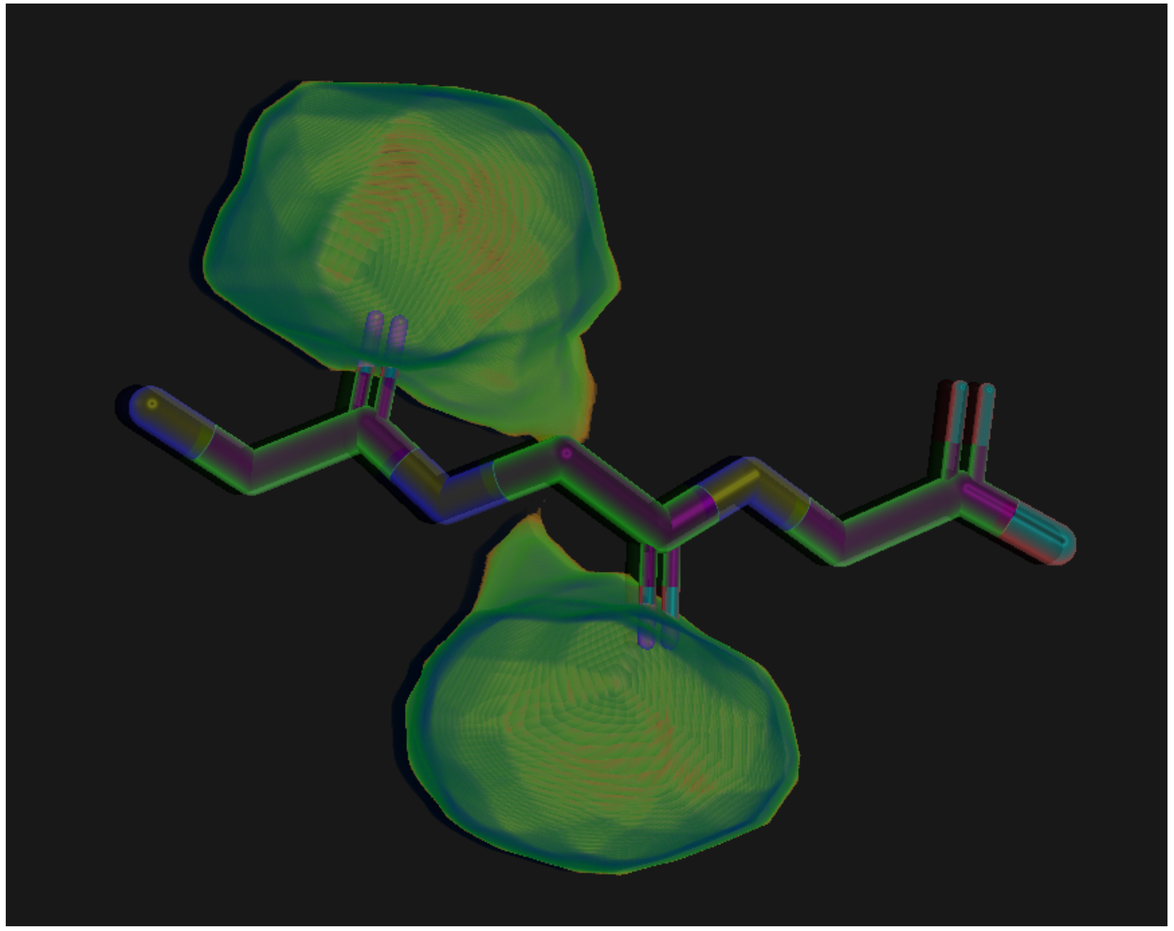


Nitrogen Acceptor

Visualizing with Atomistic Probes

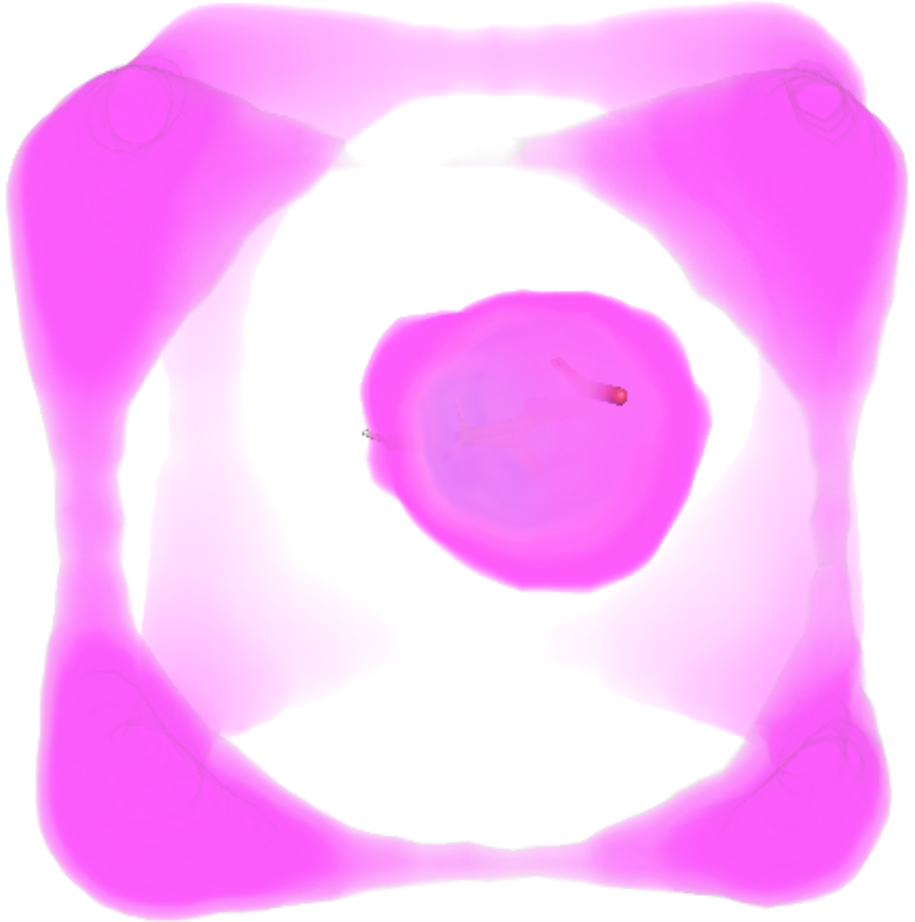


Oxygen Donor/Acceptor



Nitrogen Donor

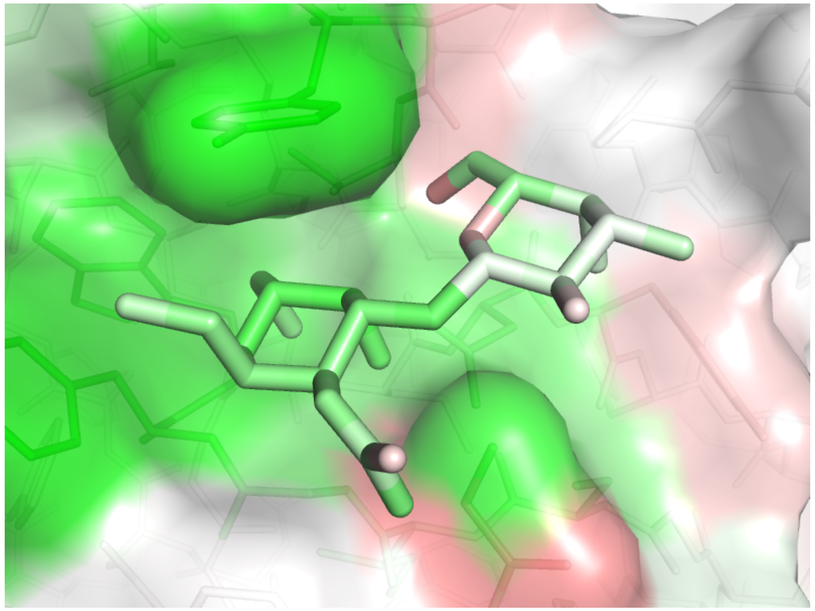
Visualizing with Atomistic Probes



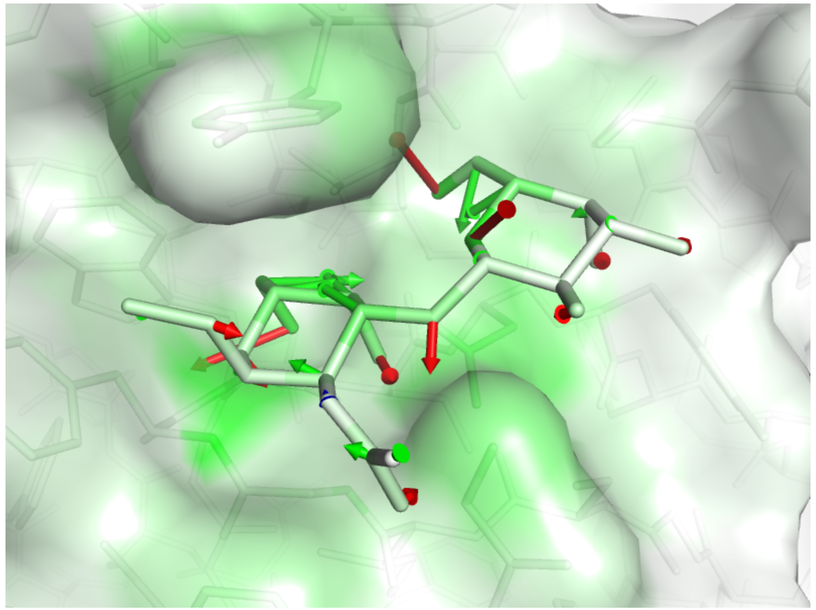
Fluorine

Carbon

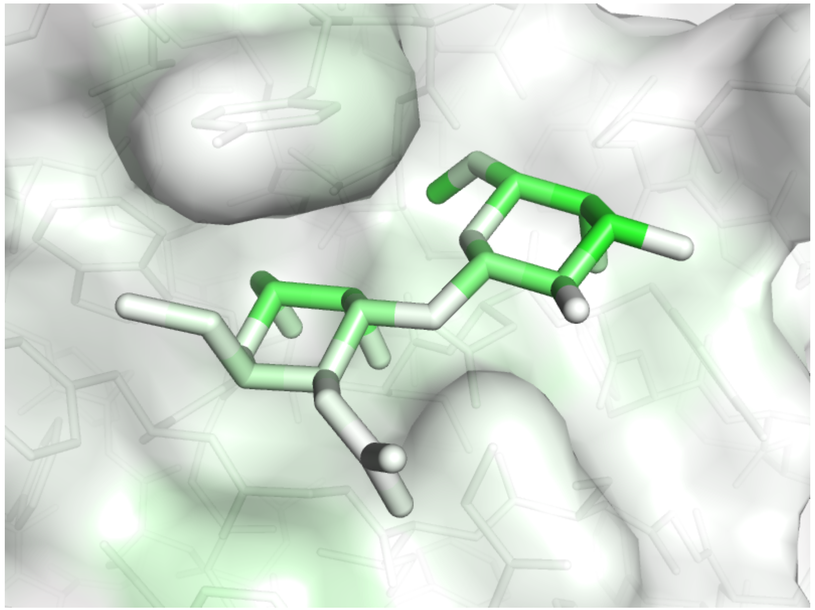
Visualizing Network Decisions



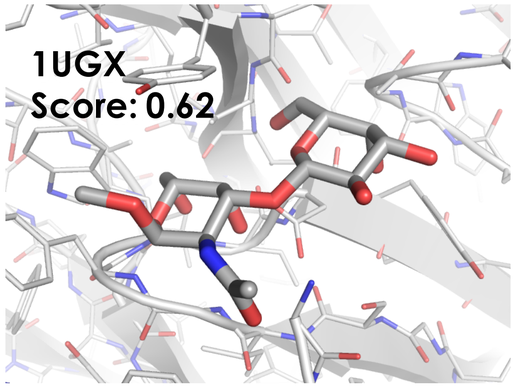
masking



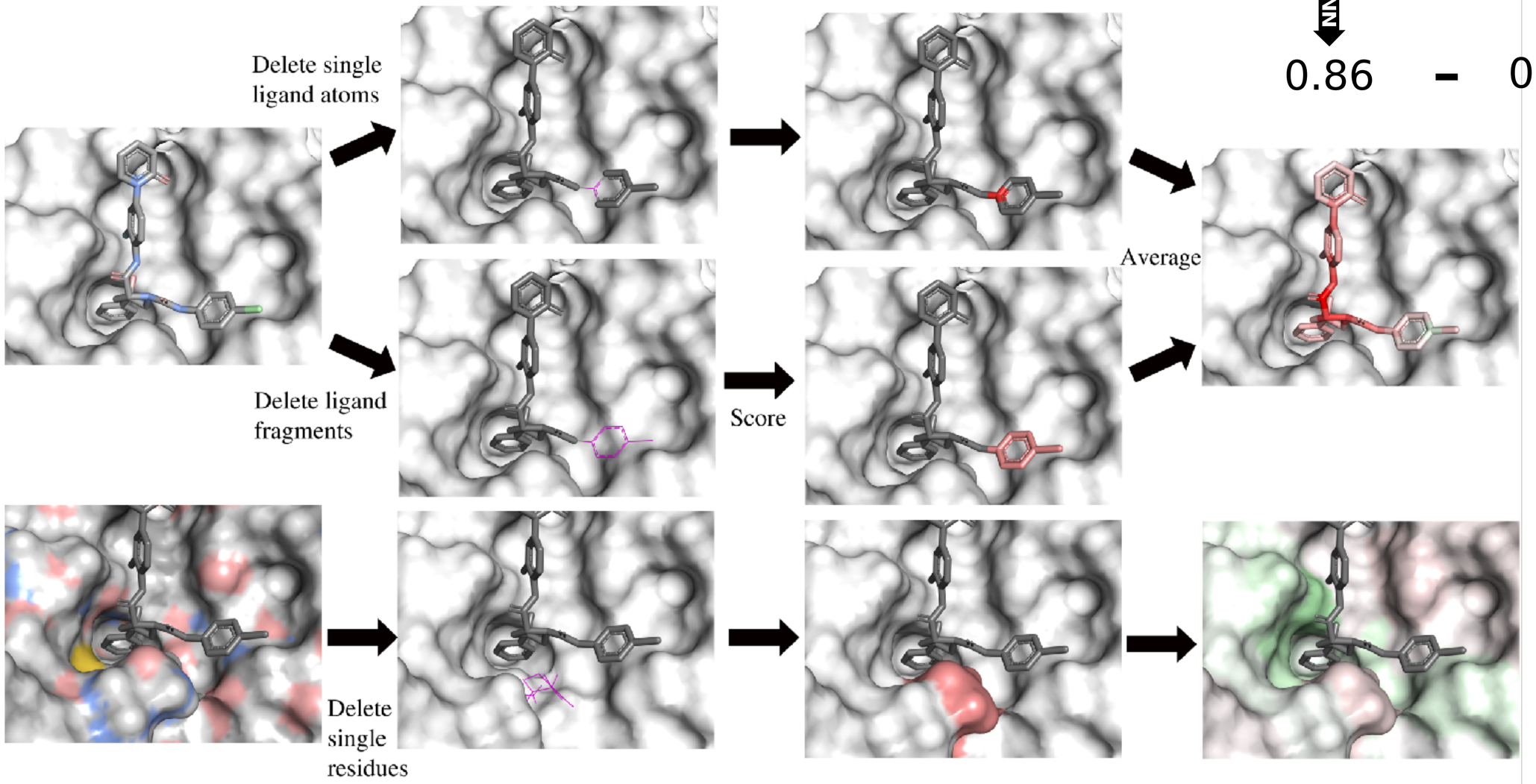
gradients

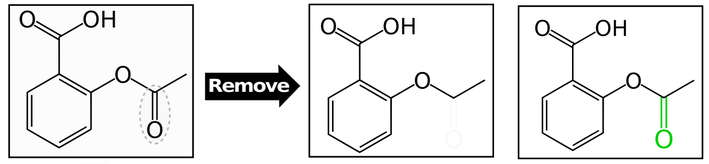


layer-wise relevance



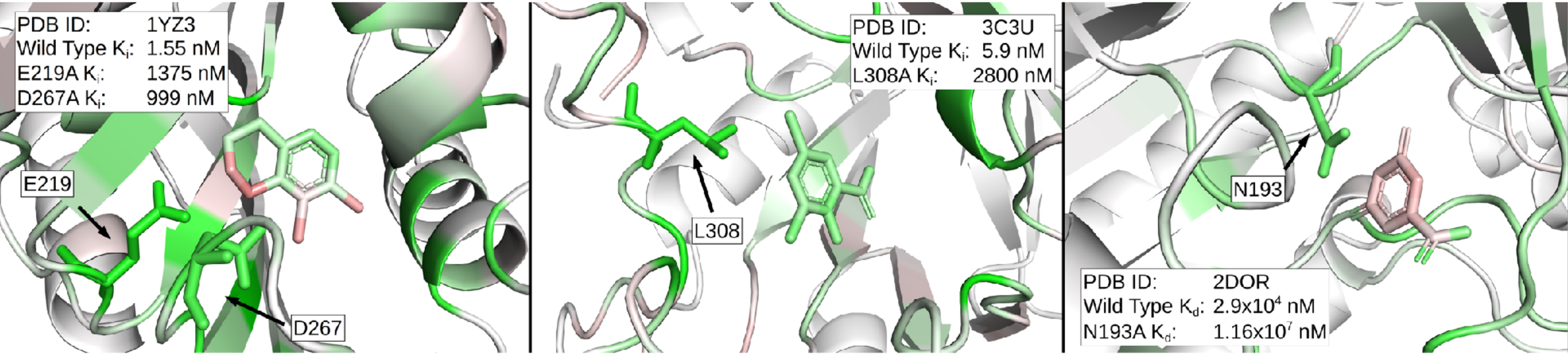
Masking



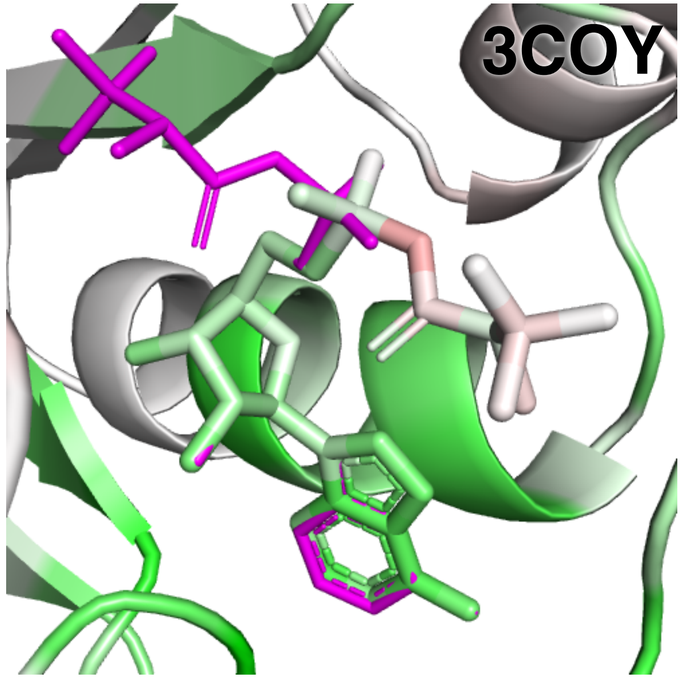


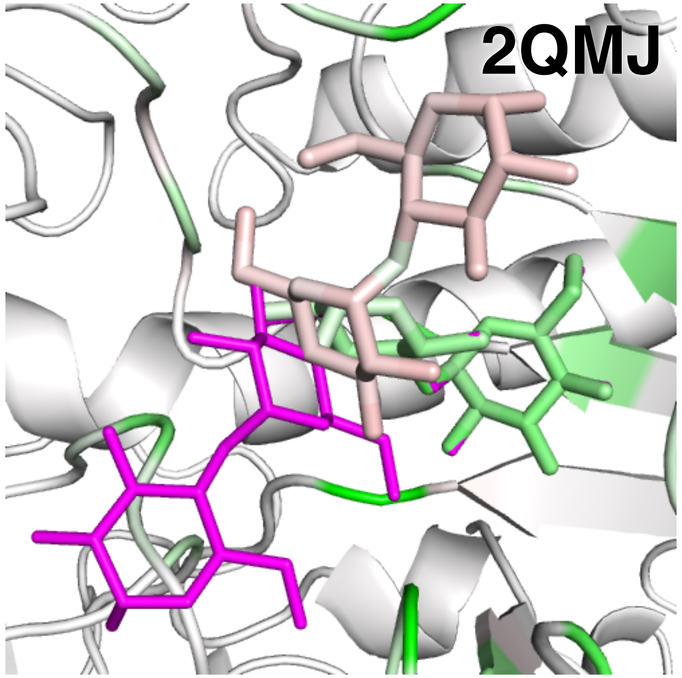
0.86 - 0.84 = 0.02

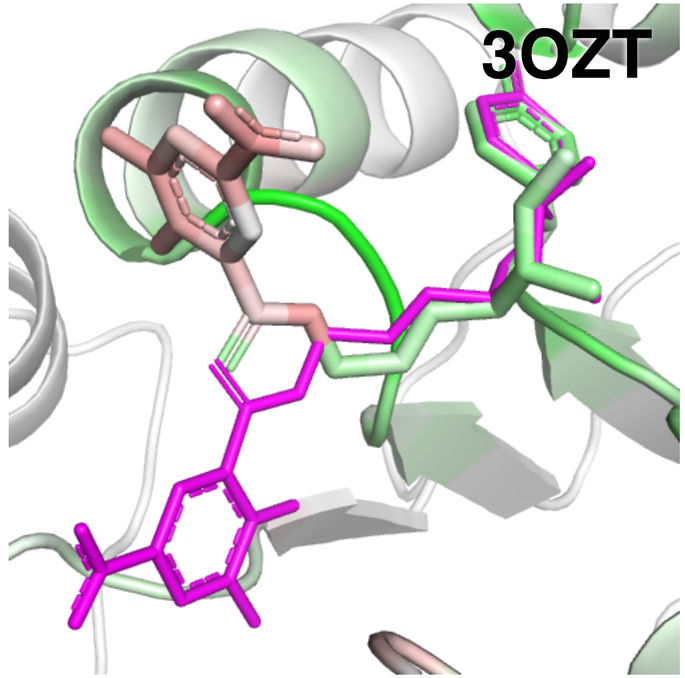
Masking: Enzyme Mutants



Pose Sensitivity

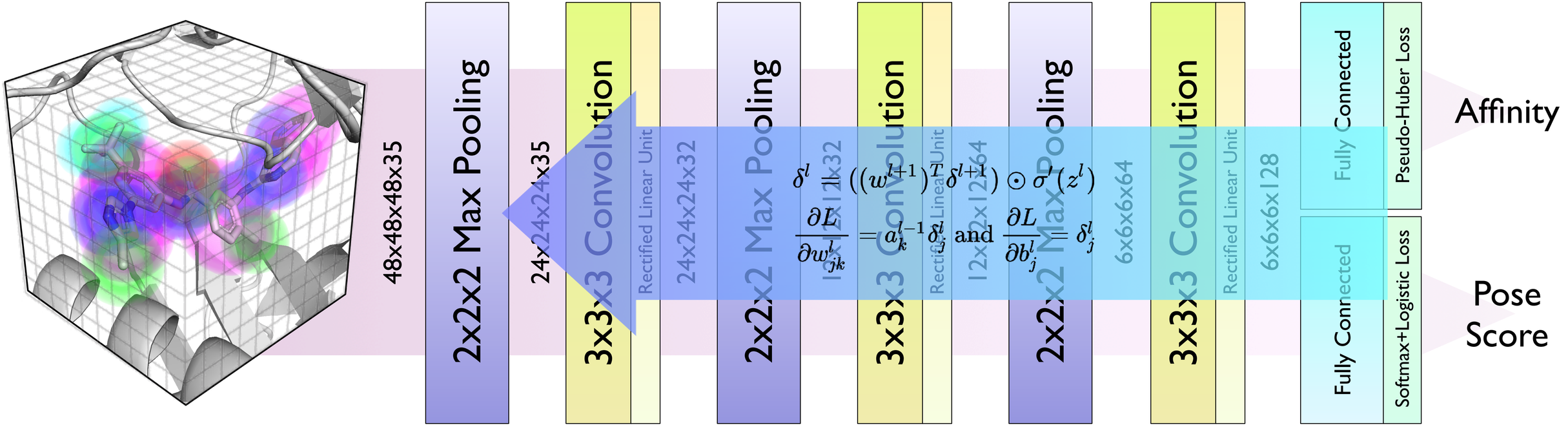




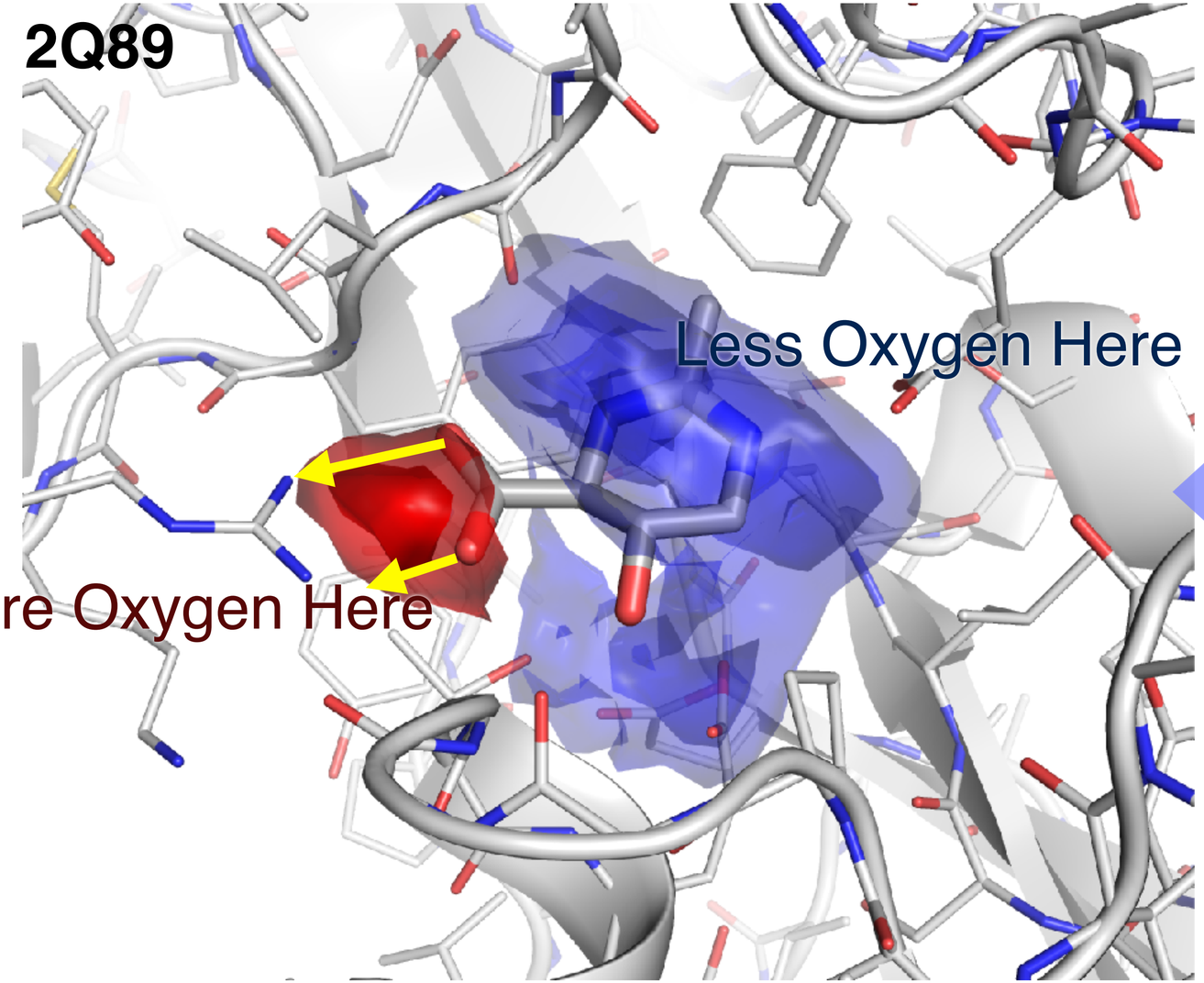


Partially Aligned Poses

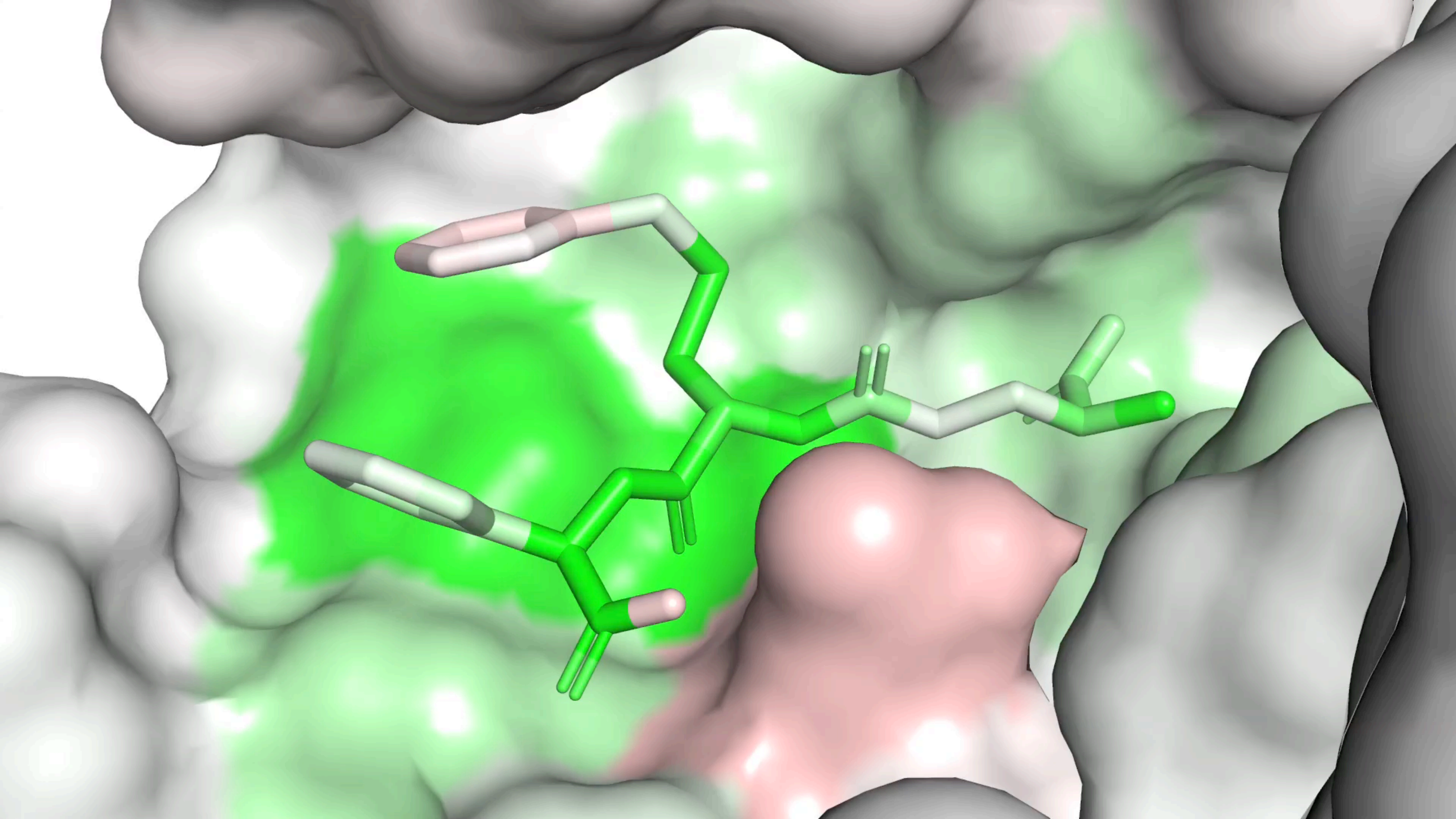
Gradients


48x48x48x35
2x2x2 Max Pooling
24x24x24x35
3x3x3 Convolution
Rectified Linear Unit
24x24x24x32
2x2x2 Max Pooling
12x12x12x32
3x3x3 Convolution
Rectified Linear Unit
12x12x12x64
2x2x2 Max Pooling
6x6x6x64
3x3x3 Convolution
Rectified Linear Unit
6x6x6x128
Fully Connected
Pseudo-Huber Loss
Affinity
Fully Connected
Softmax+Logistic Loss
Pose Score
$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$
$\frac{\partial L}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j$ and $\frac{\partial L}{\partial b^l_j} = \delta^l_j$

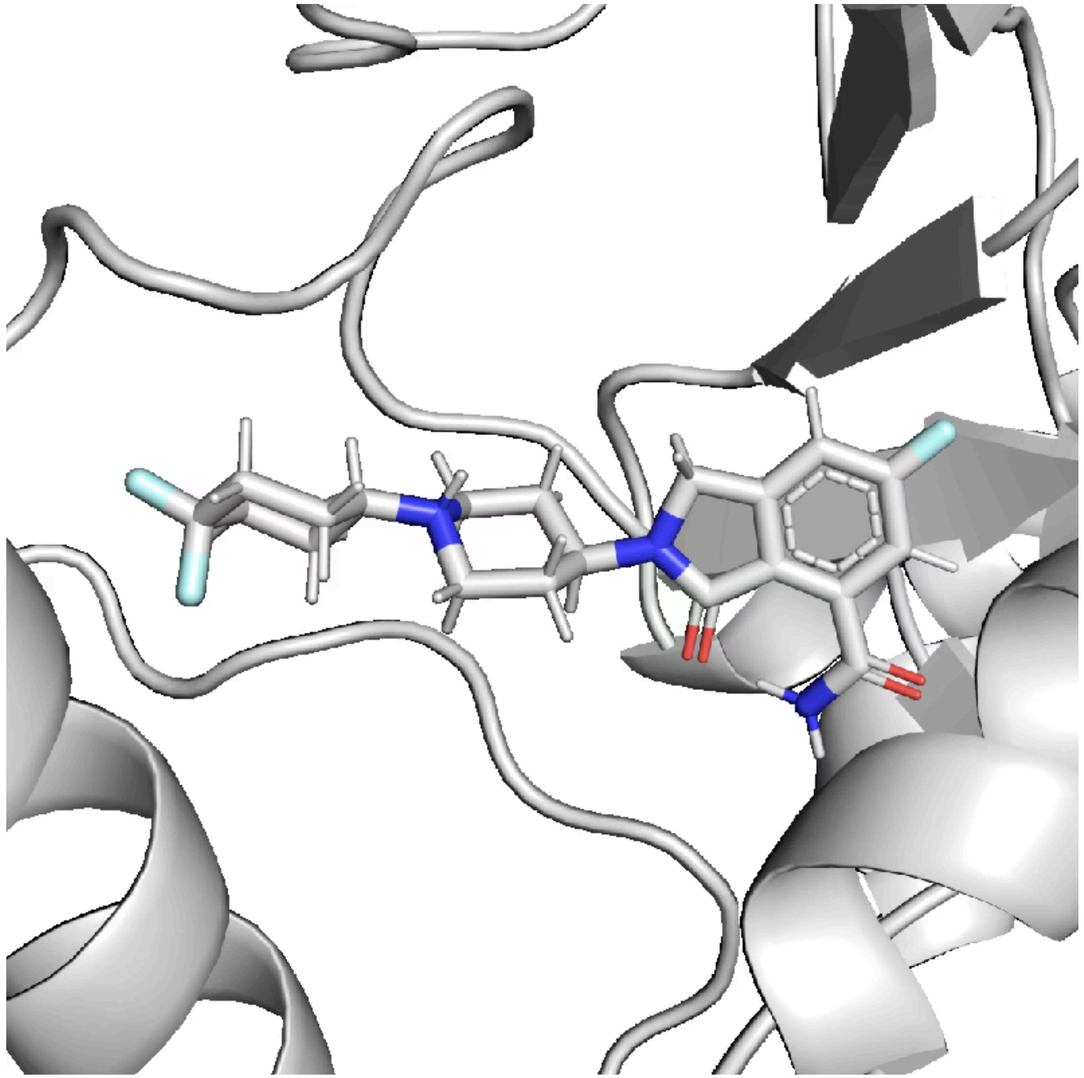
Gradients

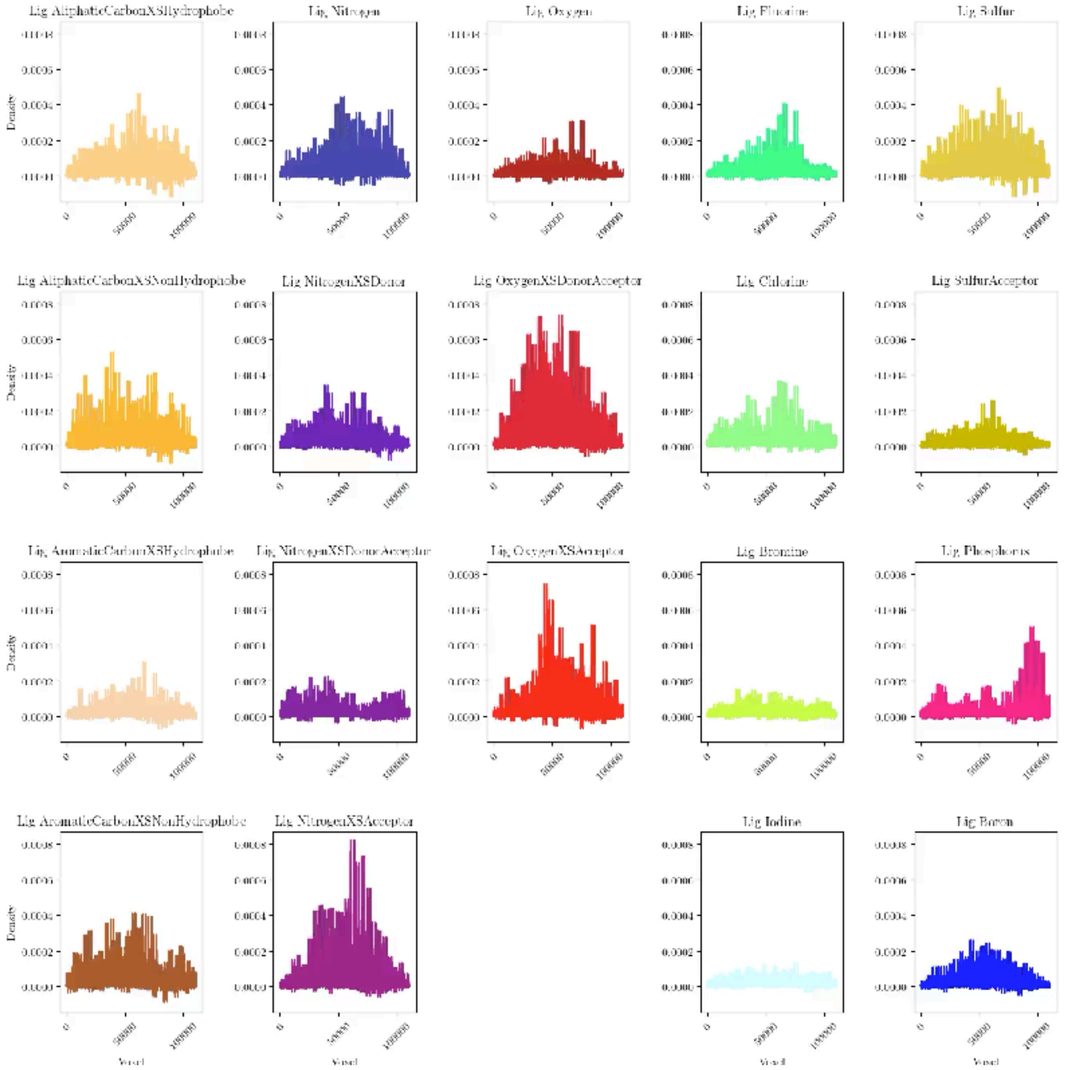


$$\frac{\partial L}{\partial A} = \sum_{i \in G_A} \frac{\partial L}{\partial G_i} \frac{\partial G_i}{\partial D} \frac{\partial D}{\partial A}$$



Pseudoligand From Gradients

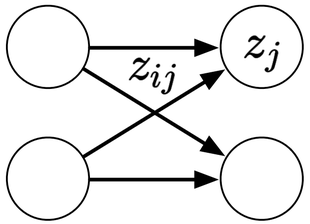




Layer-wise Relevance

$$f(x) = \ldots = \sum_{d \in l+1} R_d^{(l+1)} = \sum_{d \in l} R_d^{(l)} = \ldots = \sum_d R_d^{(1)}$$

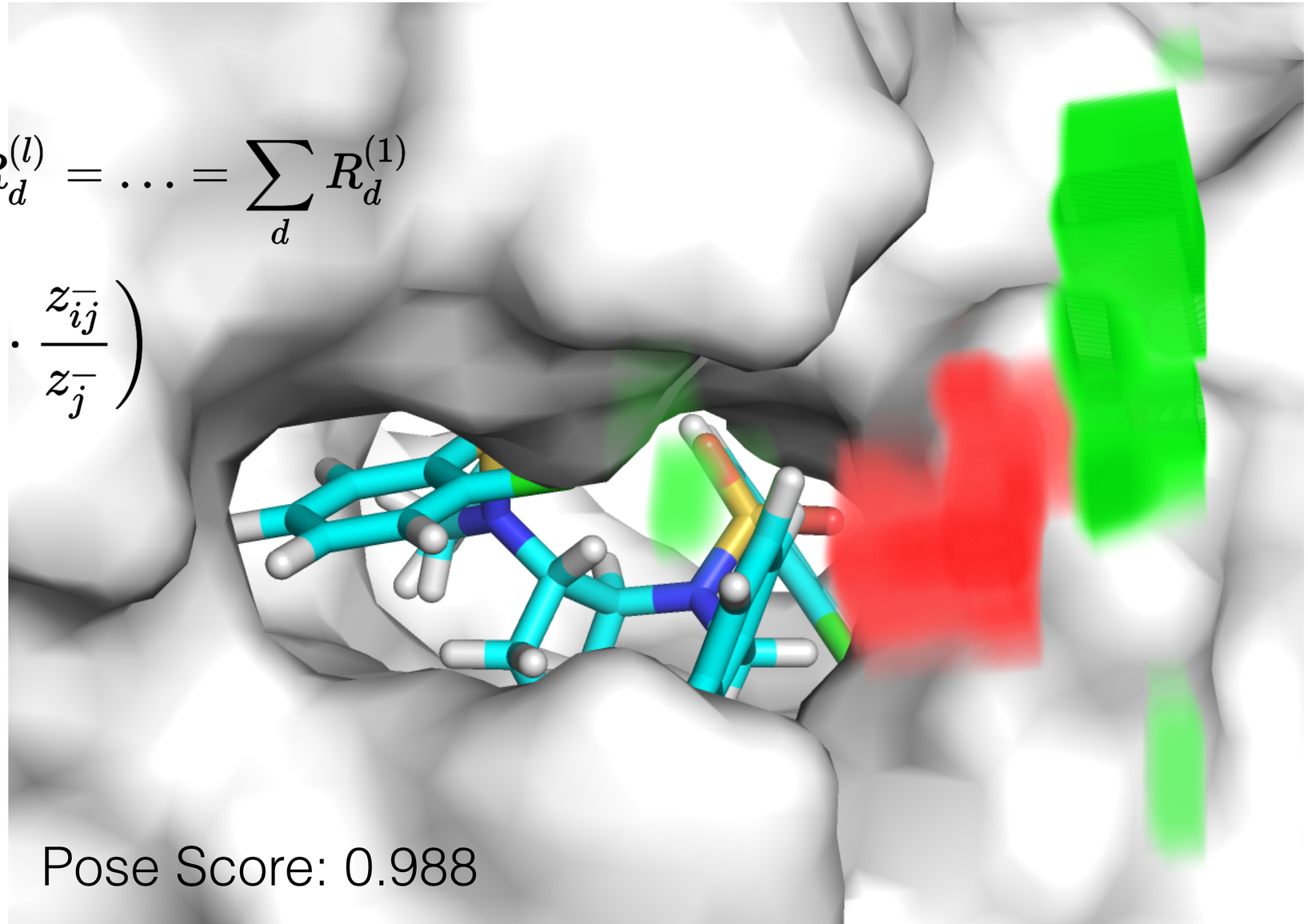
$$R_{i \leftarrow j}^{(l,l+1)} = R_j^{(l+1)} \cdot \left(\alpha \cdot \frac{z_{ij}^+}{z_j^+} + \beta \cdot \frac{z_{ij}^-}{z_j^-} \right)$$



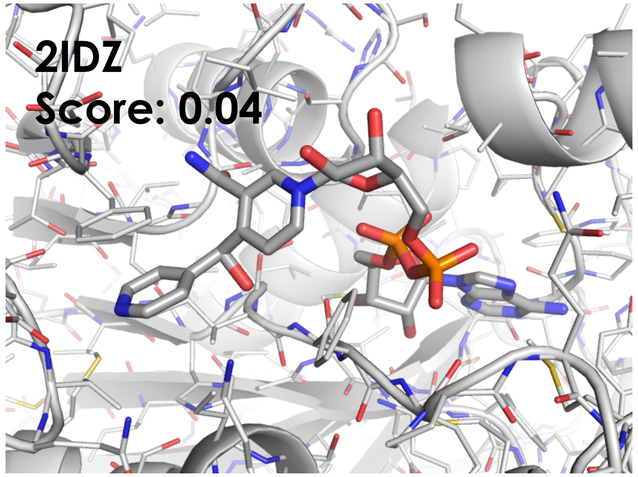
On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation

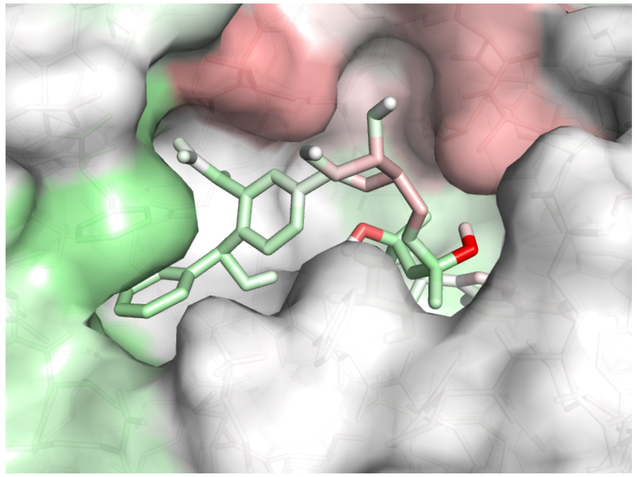
Sebastian Bach, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller, Wojciech Samek

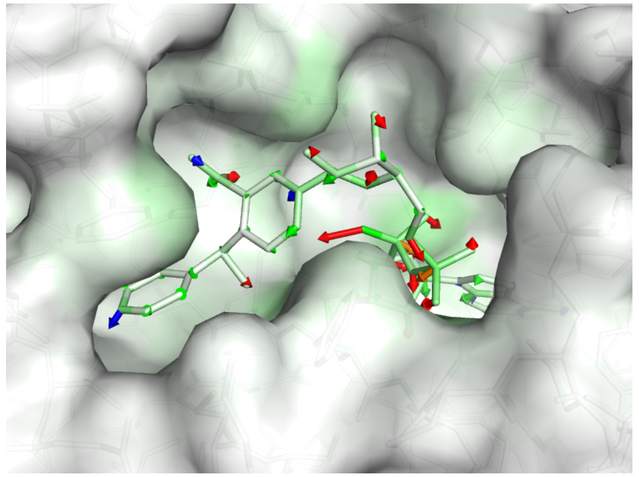
Published: July 10, 2015 • https://doi.org/10.1371/journal.pone.0130140

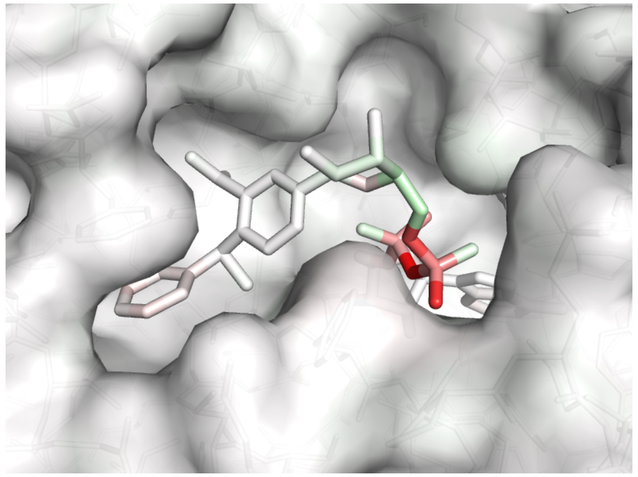


Visualizations





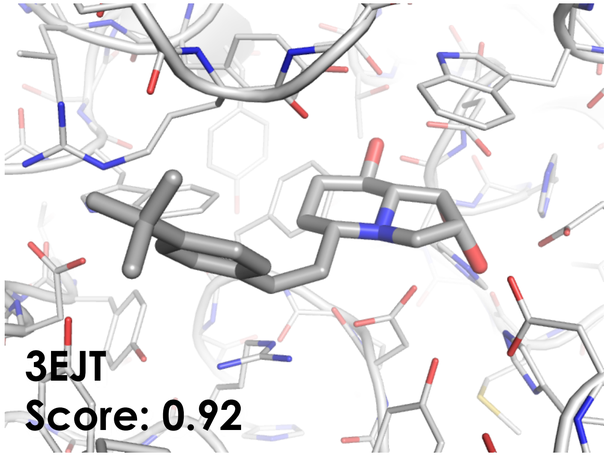


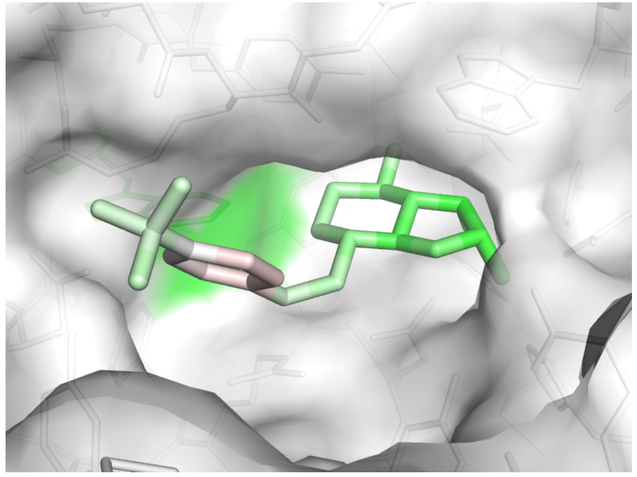


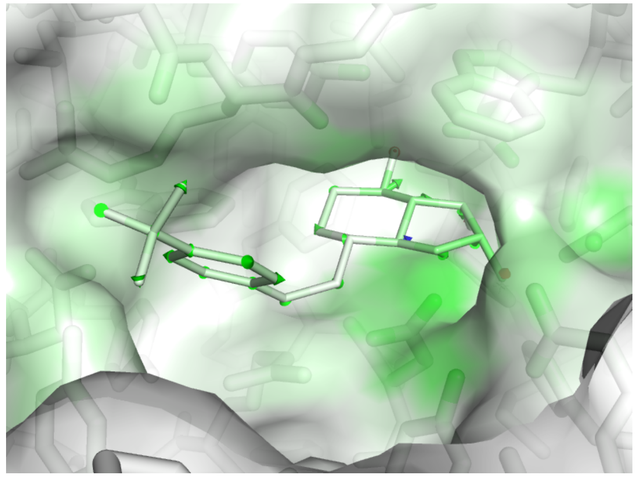
Masking

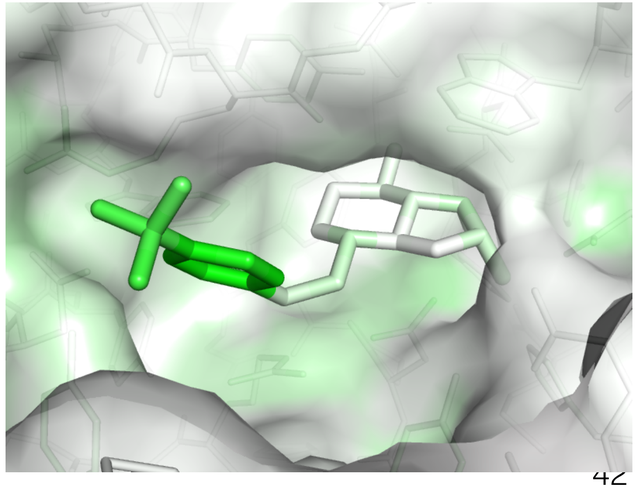
Gradients

LRP

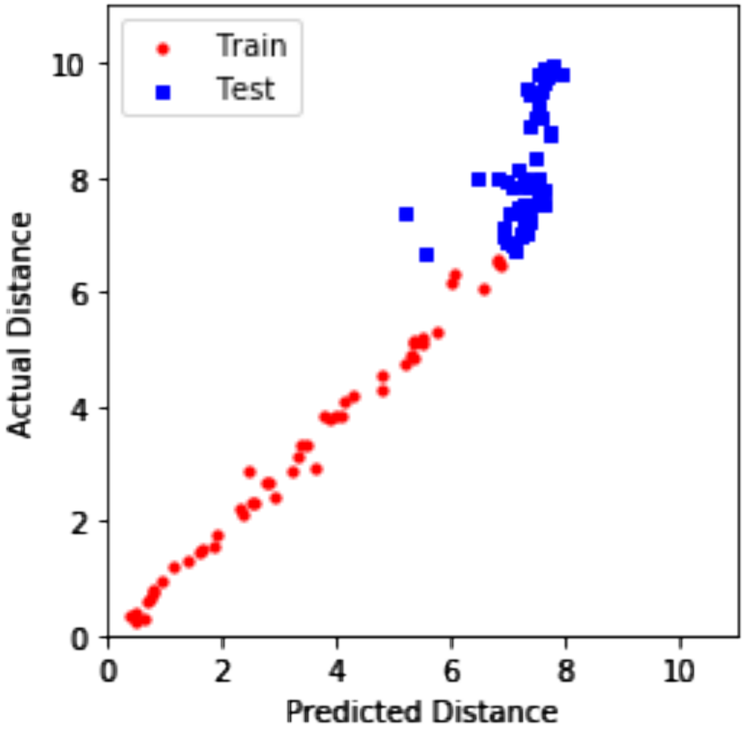


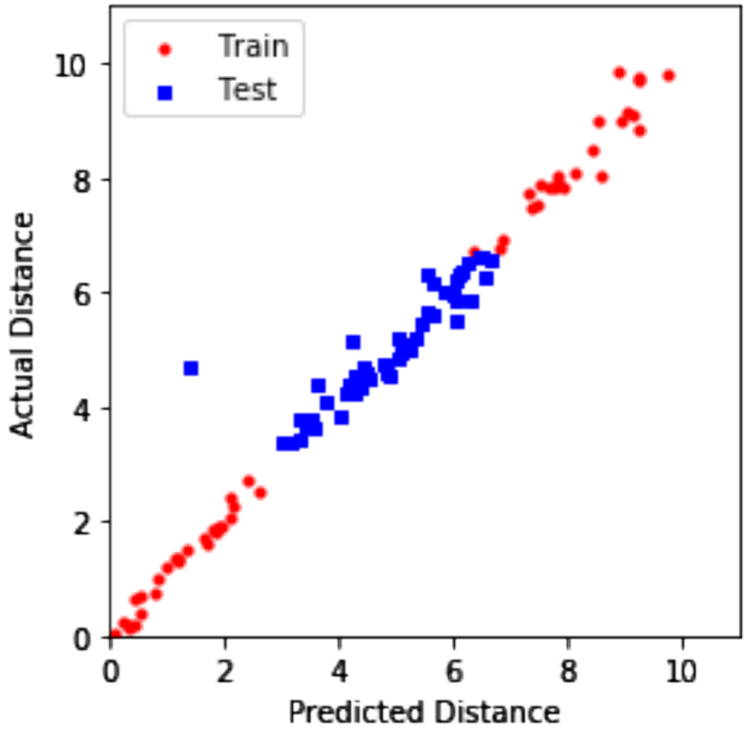


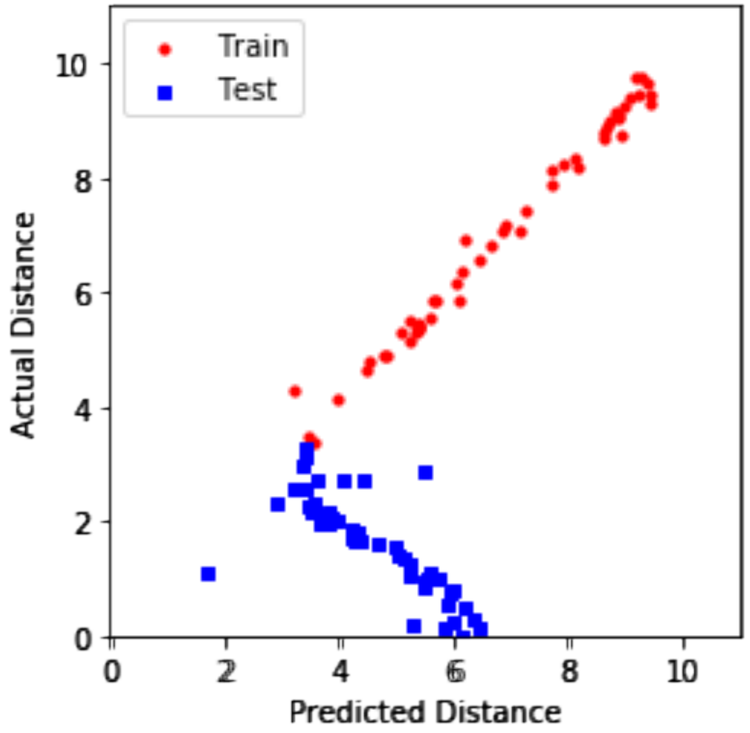


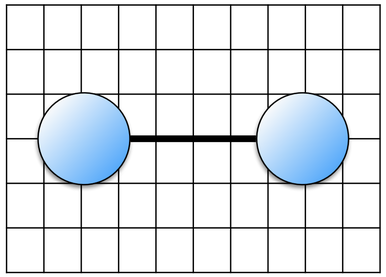


Another Aside...



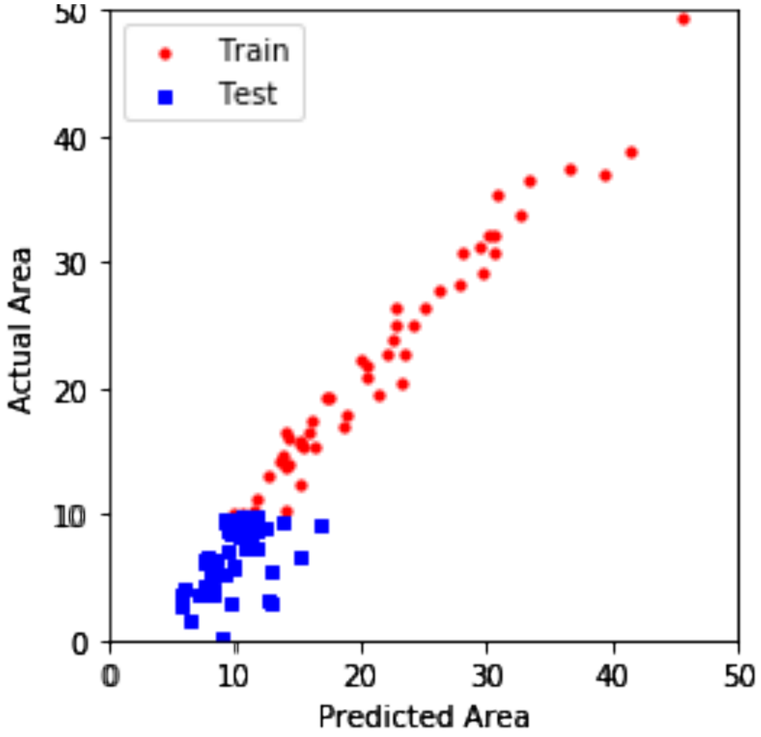


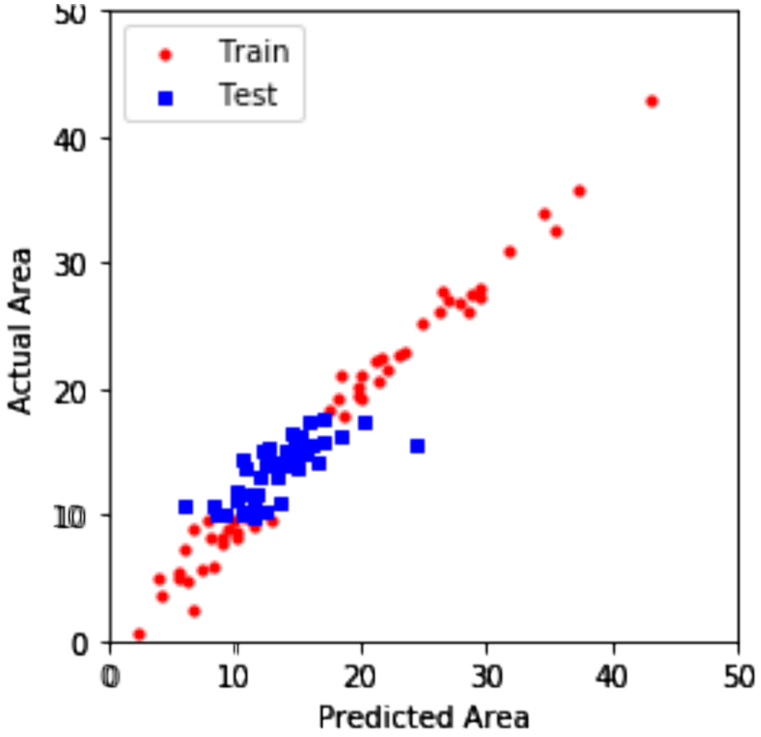


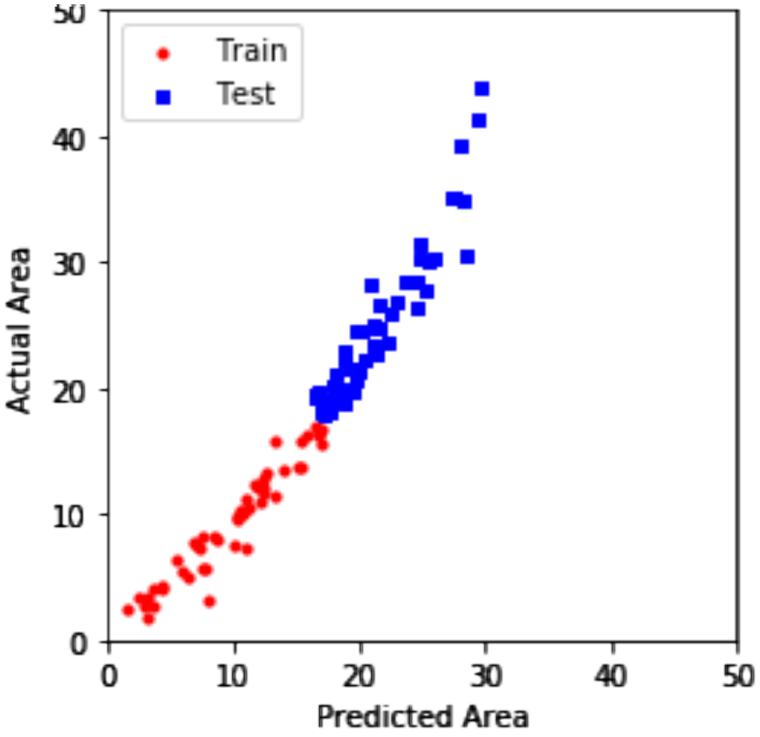


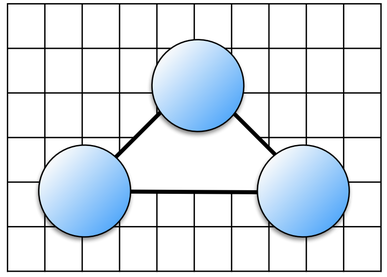
Distance

Another Aside...



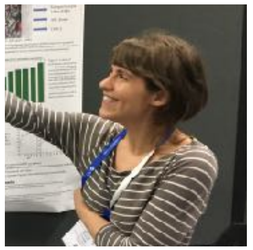






Area

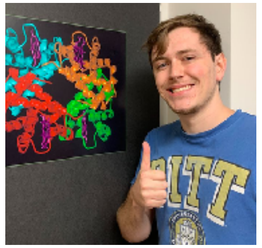
Acknowledgements



Jocelyn Sunseri

Paul Francoeur

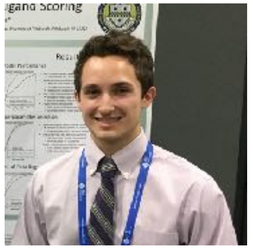
Josh Hochuli



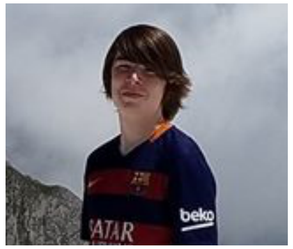
Hunter Haaf

Rachel Rosenzweig

Keshavan Seshadri



Matt Ragoza



Alec Helbling

Department of Computational and Systems Biology

National Institute of General Medical Sciences
R01GM108340

NVIDIA

Google Cloud

RELAY THERAPEUTICS

github.com/gnina

github.com/3dmol

http://bits.csb.pitt.edu

@david_koes