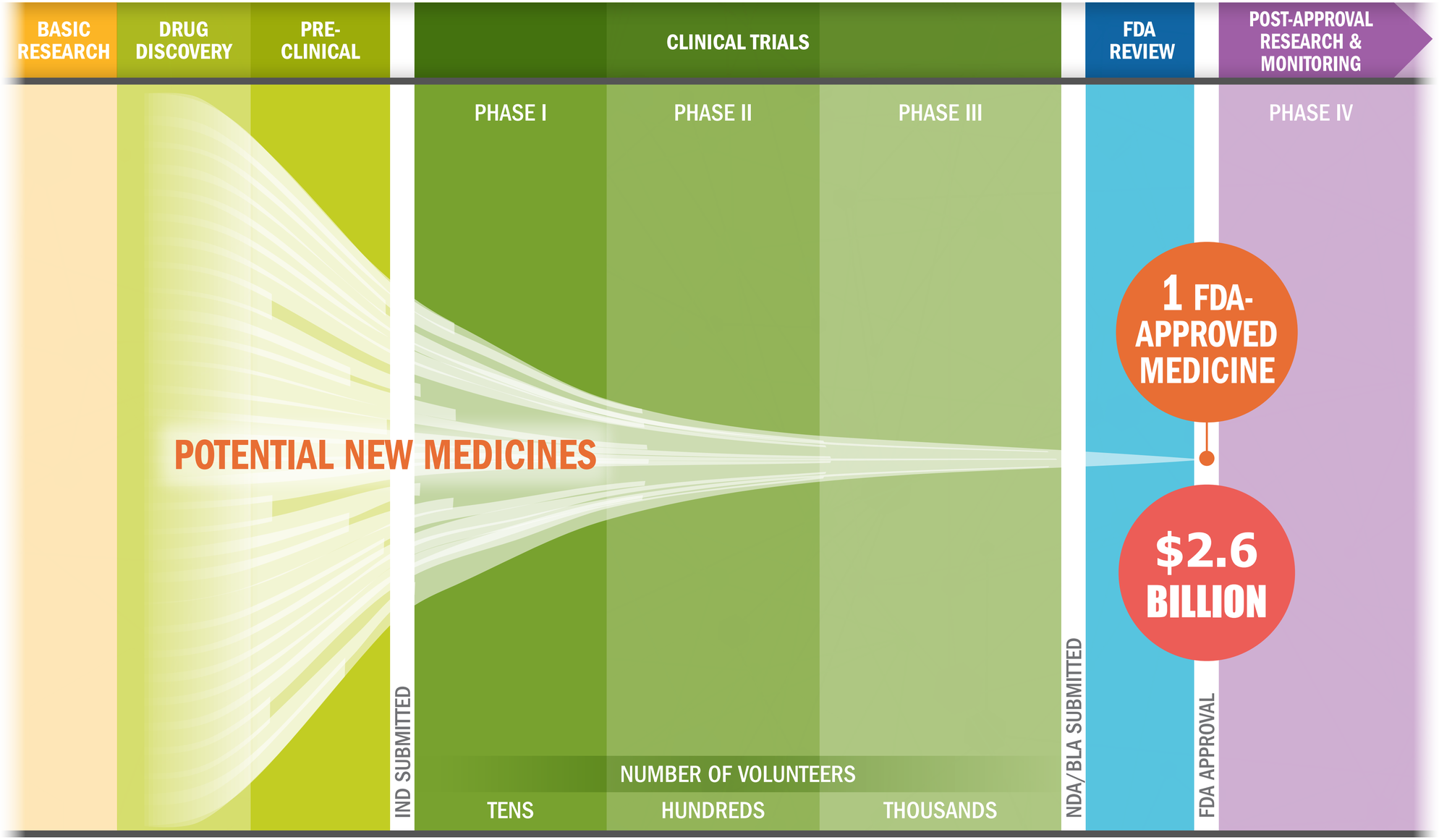
Deep Generative Models for Computational Drug Discovery

David Koes

@david_koes

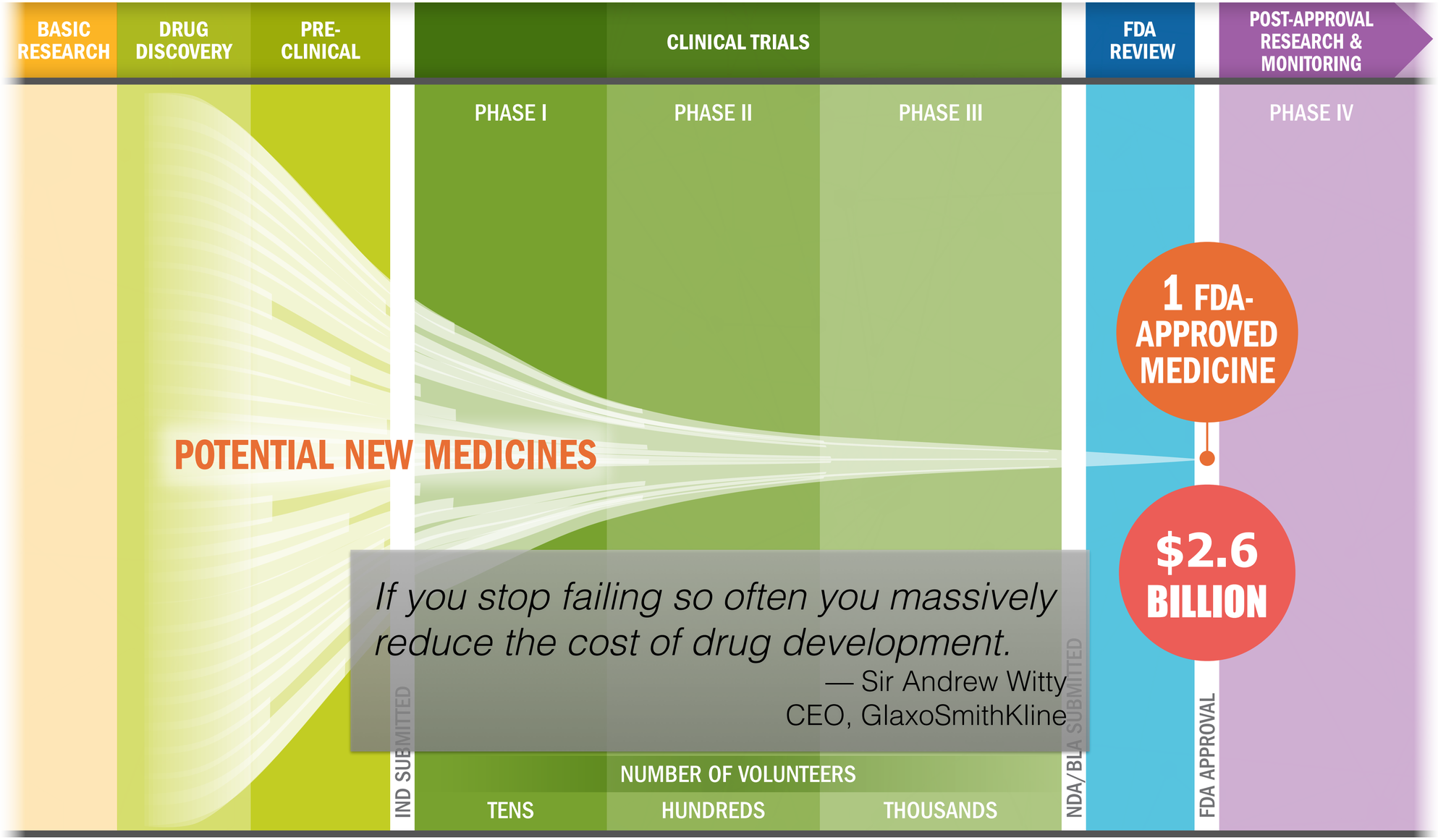
GTC 2019
San Jose, CA
March 20, 2019

THE BIOPHARMACEUTICAL RESEARCH AND DEVELOPMENT PROCESS



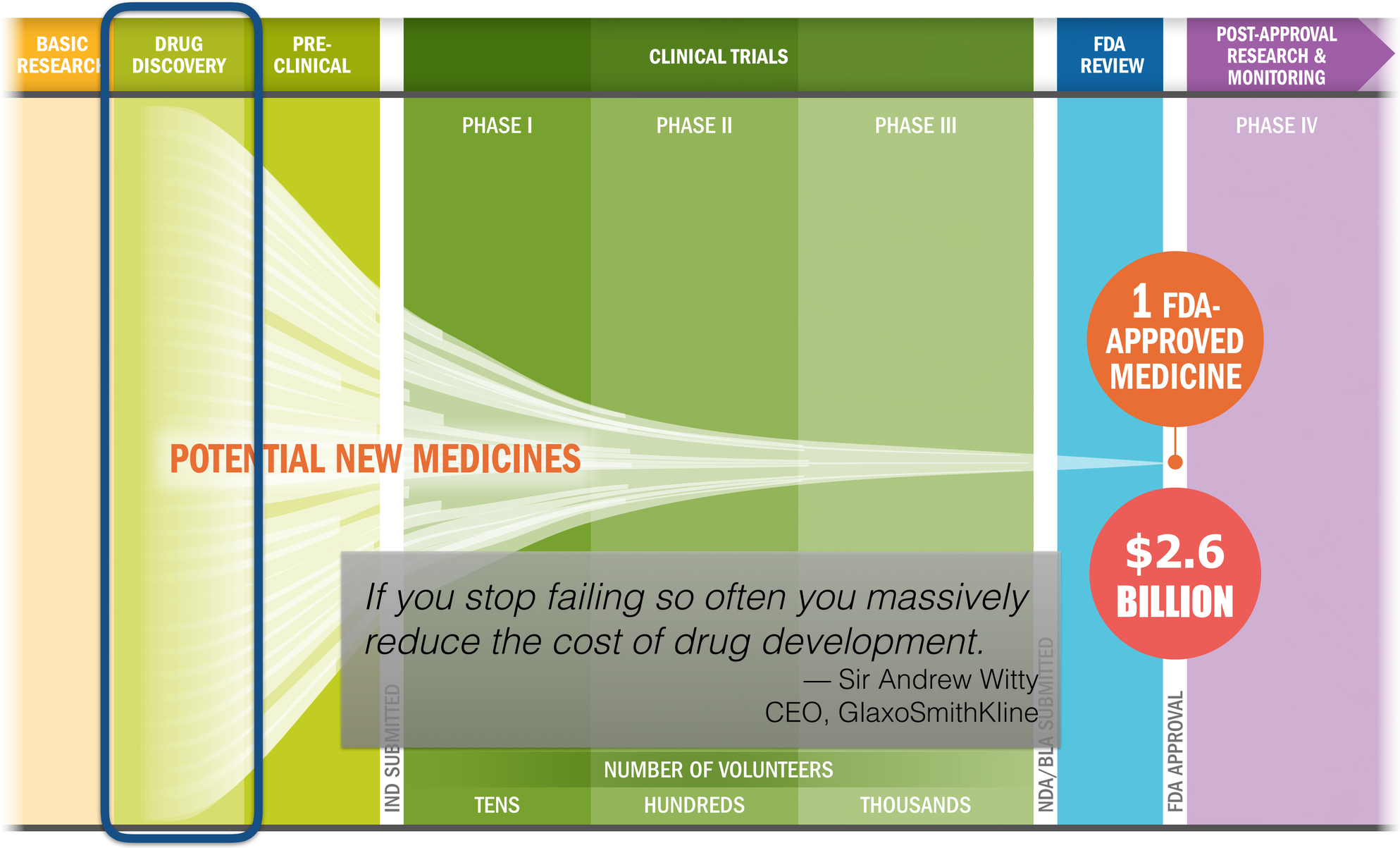
Source: Pharmaceutical Research and Manufacturers of America (http://phrma.org)

THE BIOPHARMACEUTICAL RESEARCH AND DEVELOPMENT PROCESS



Source: Pharmaceutical Research and Manufacturers of America (http://phrma.org)

THE BIOPHARMACEUTICAL RESEARCH AND DEVELOPMENT PROCESS



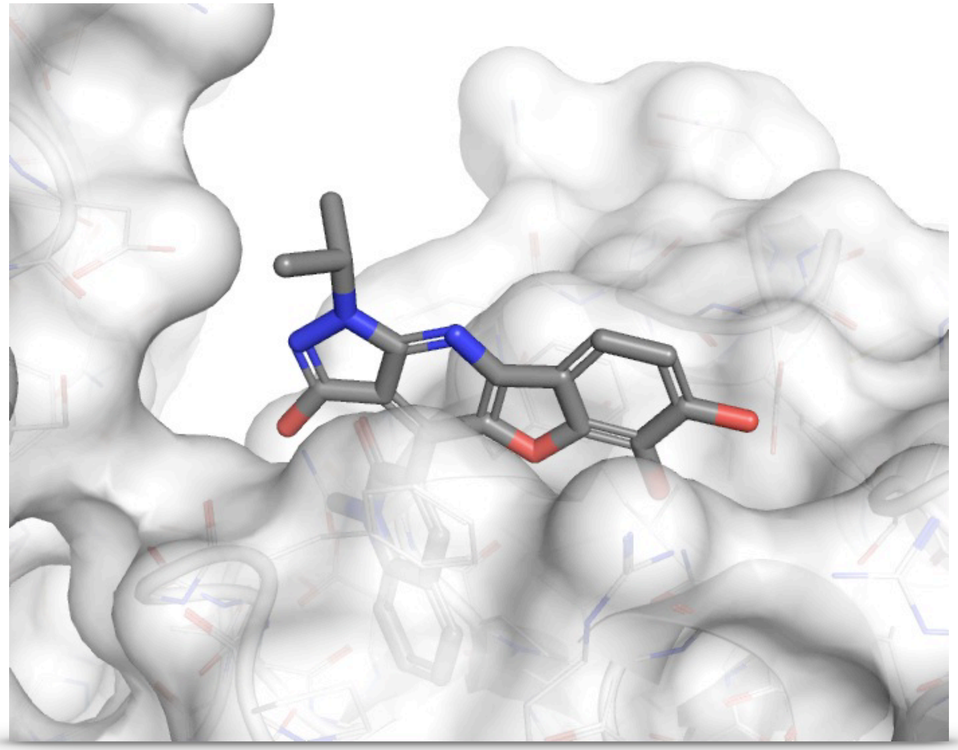
If you stop failing so often you massively reduce the cost of drug development.

— Sir Andrew Witty
CEO, GlaxoSmithKline

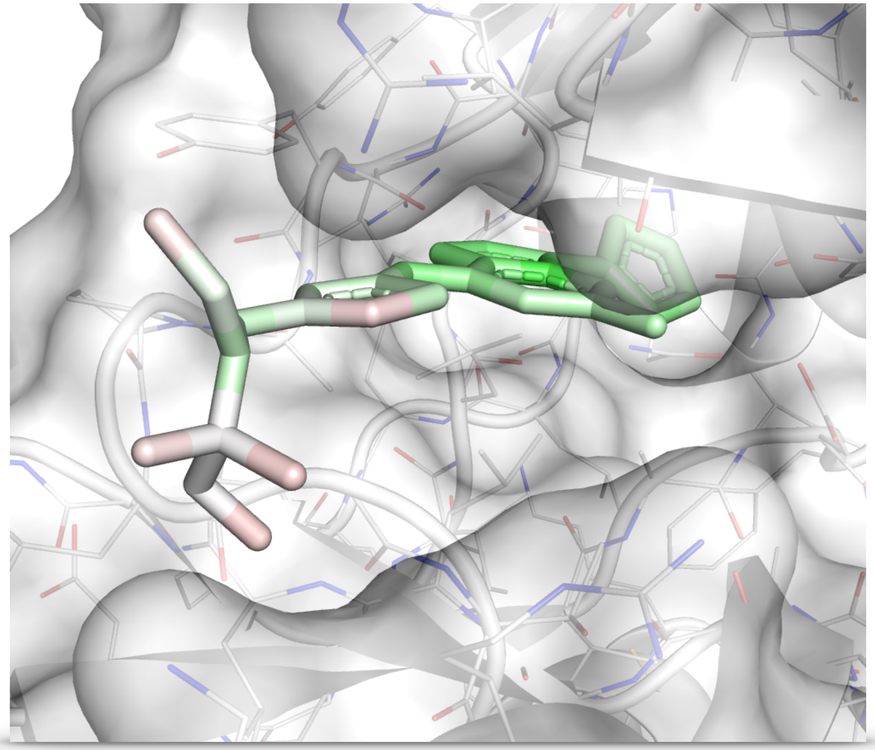
Source: Pharmaceutical Research and Manufacturers of America (http://phrma.org)

Structure Based Drug Design

Pose Prediction Binding Discrimination Affinity Prediction



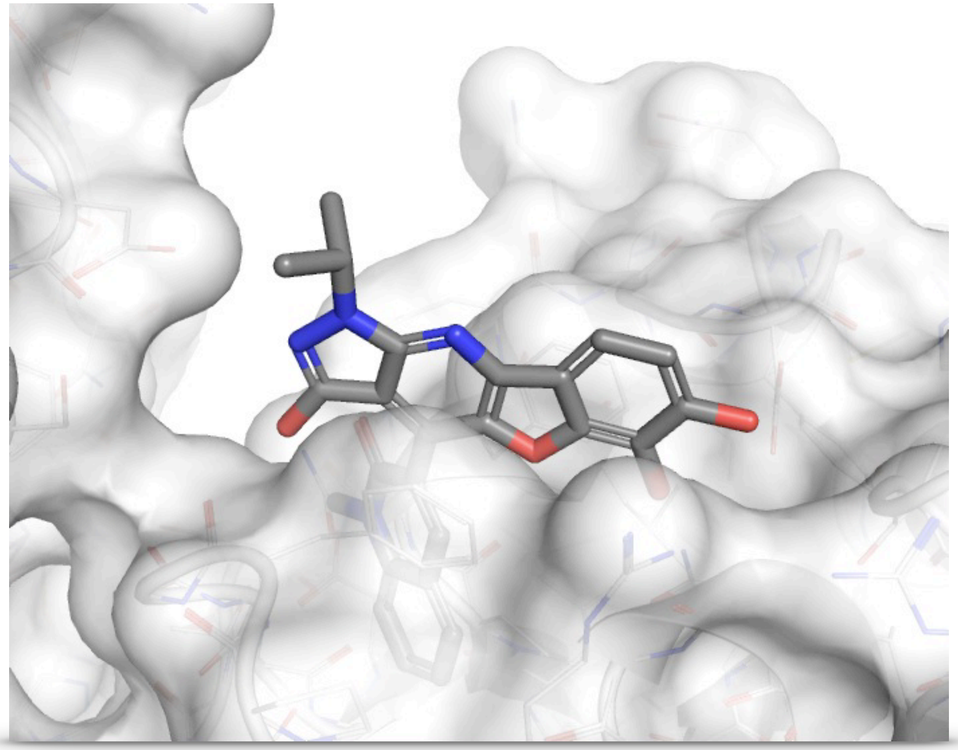
Virtual Screening



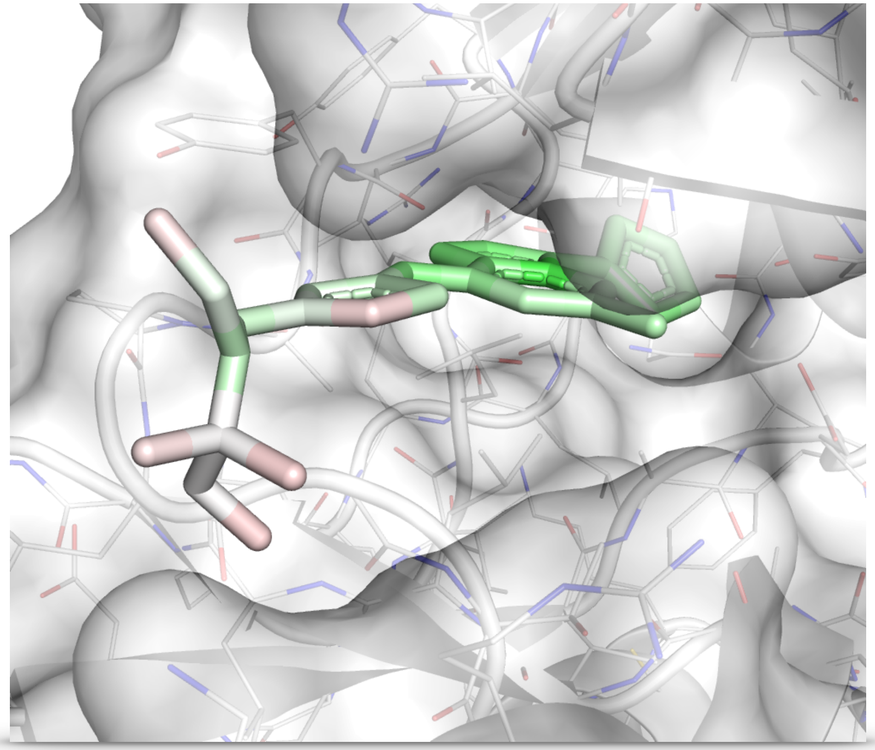
Lead Optimization

Structure Based Drug Design

Pose Prediction Binding Discrimination Affinity Prediction

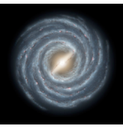


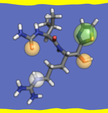
Virtual Screening



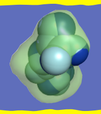
Lead Optimization

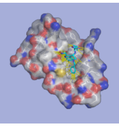
Drug Discovery Funnel

Purchasable  Accessible

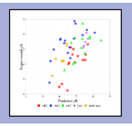


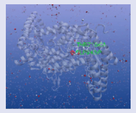
Matching



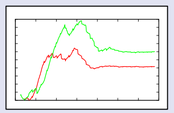


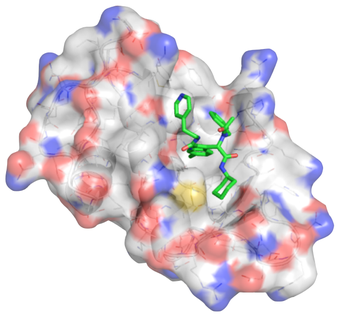
Scoring

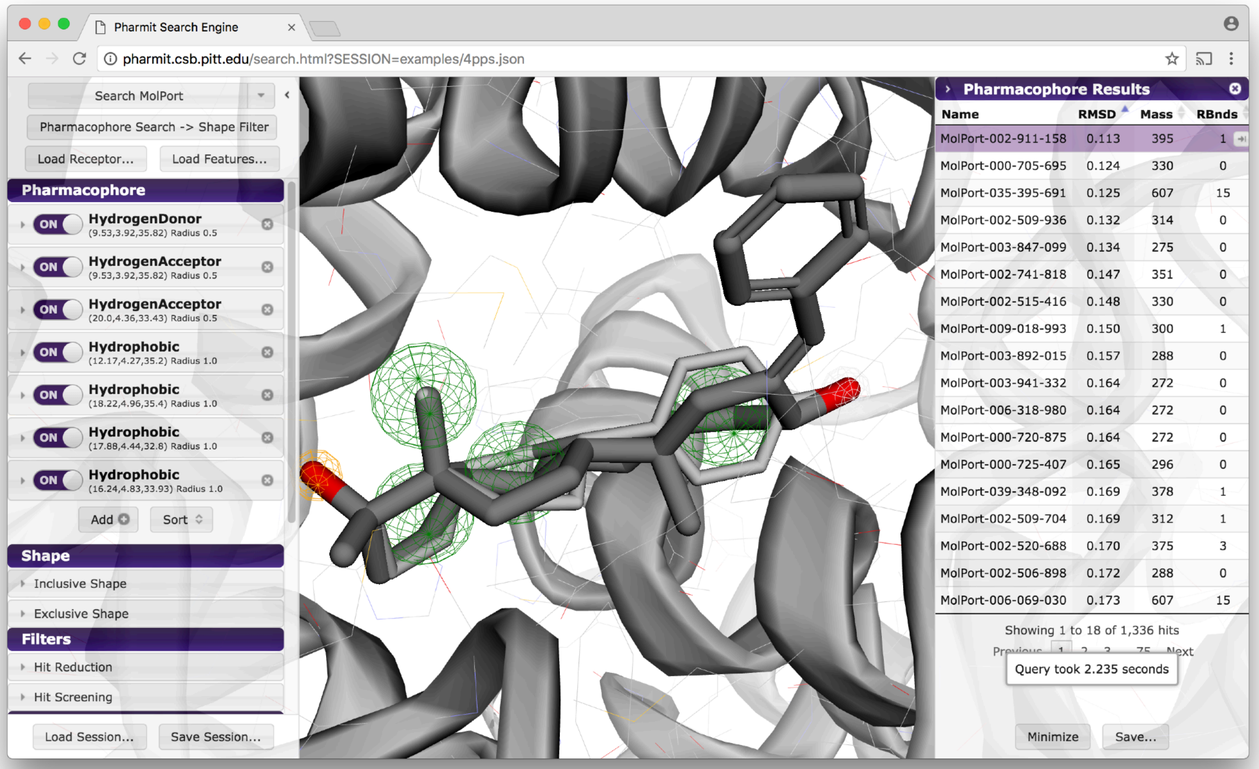




Dynamics

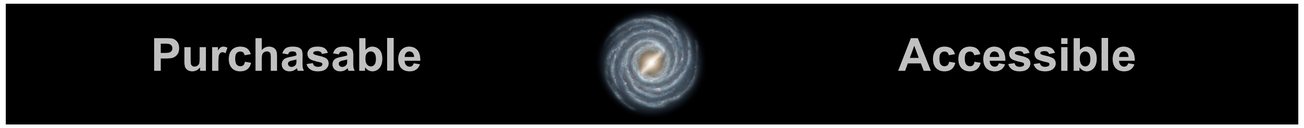


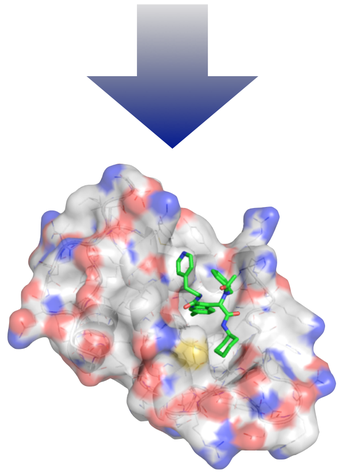


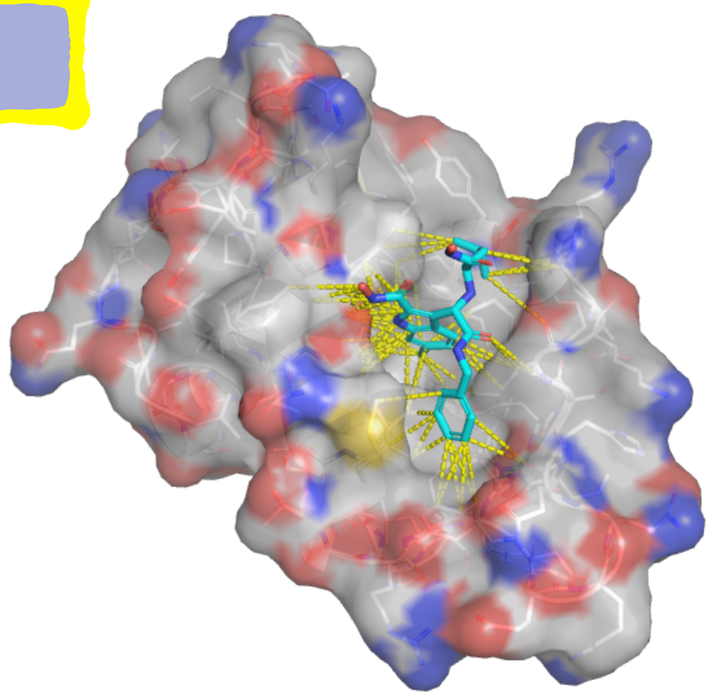


http://pharmit.csb.pitt.edu

Drug Discovery Funnel



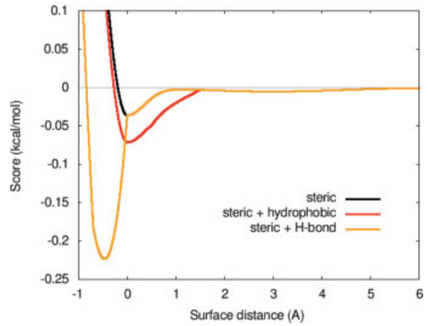




$$\text{gauss}_1(d) = w_{\text{guass}_1} e^{-(d/0.5)^2}$$
$$\text{gauss}_2(d) = w_{\text{guass}_2} e^{-((d-3)/2)^2}$$
$$\text{repulsion}(d) = \begin{cases} w_{\text{repulsion}} d^2 & d < 0 \\ 0 & d \geq 0 \end{cases}$$

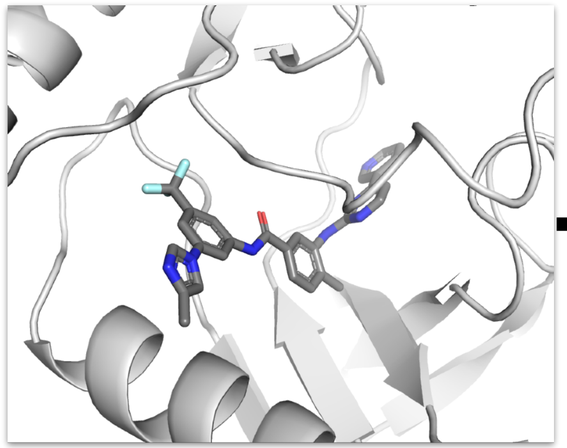
$$\text{hydrophobic}(d) = \begin{cases} w_{\text{hydrophobic}} & d < 0.5 \\ 0 & d > 1.5 \\ w_{\text{hydrophobic}}(1.5 - d) & otherwise \end{cases}$$

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



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



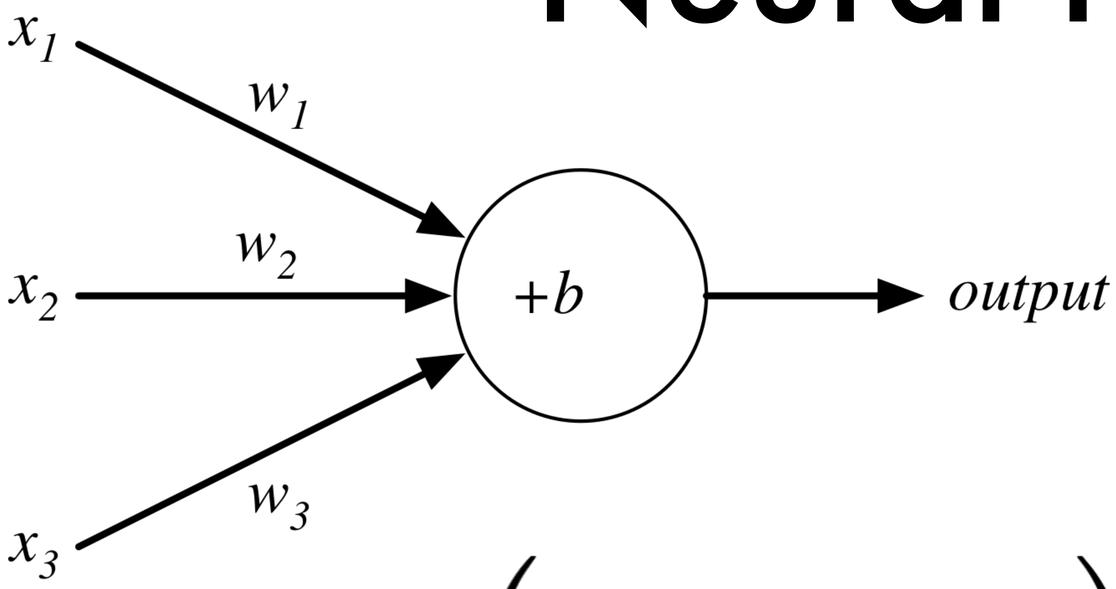
Model

Pose Prediction

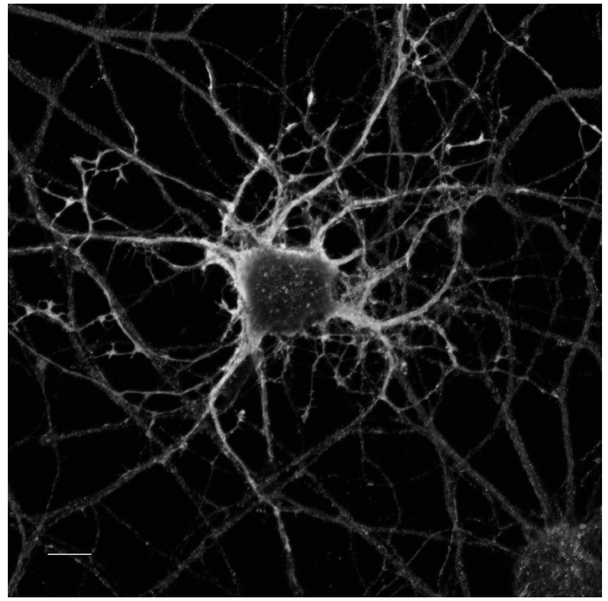
Binding Discrimination

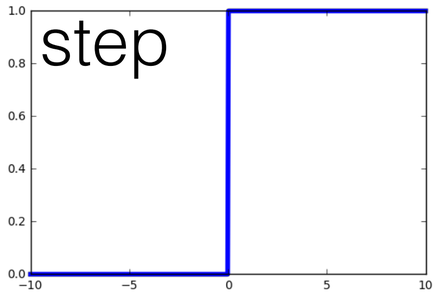
Affinity Prediction

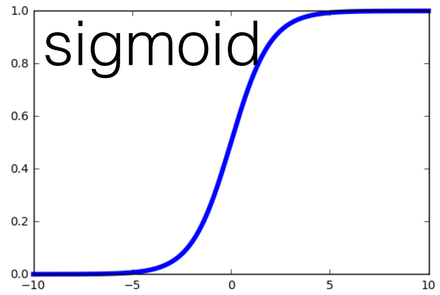
Neural Networks

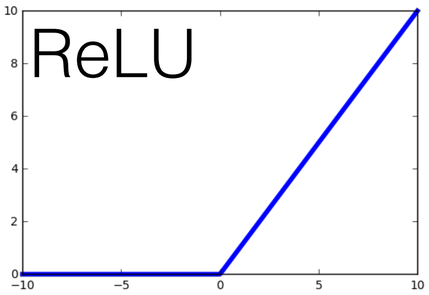


$$output = \sigma\left(\sum_i w_i x_i + b\right)$$

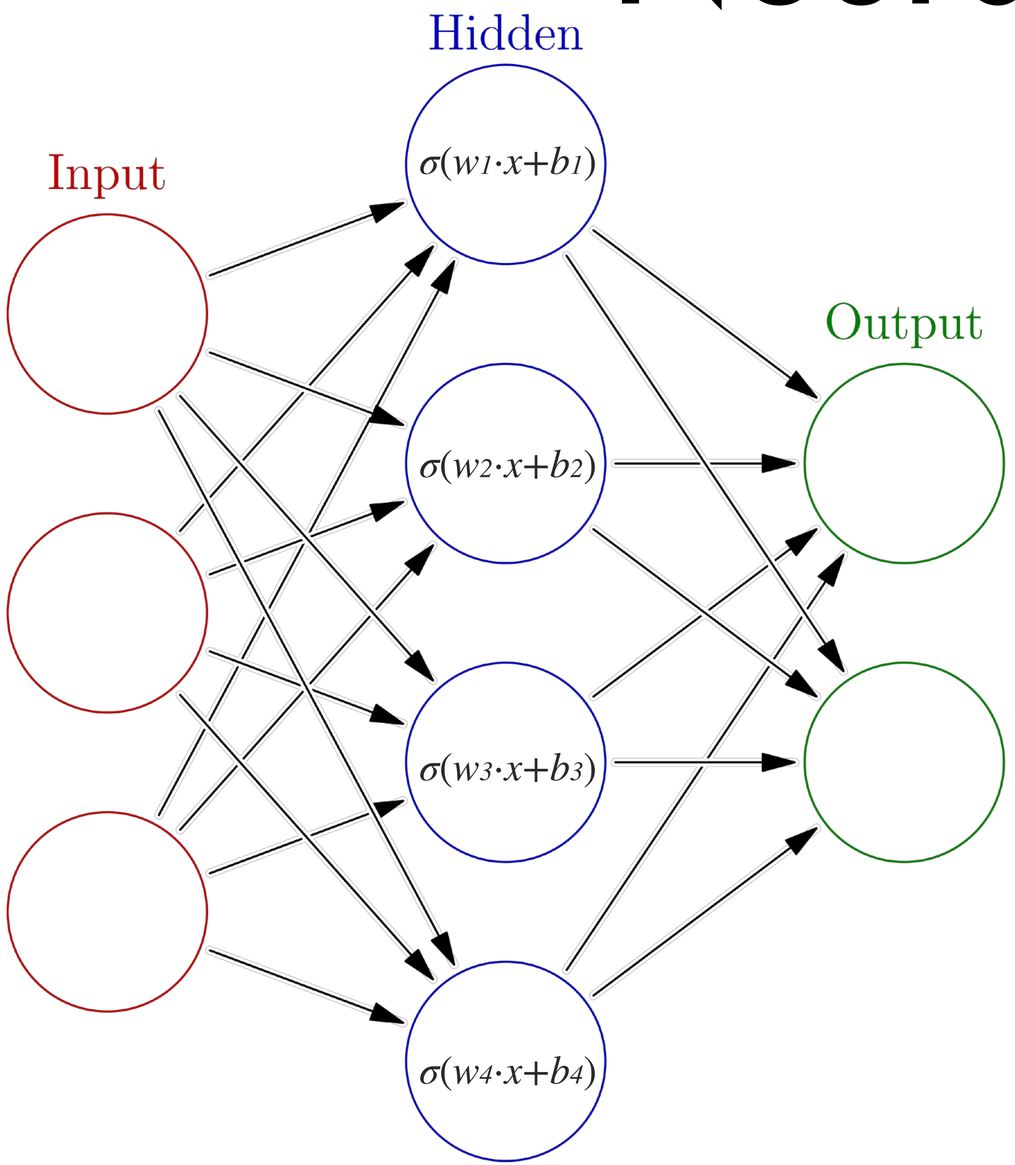


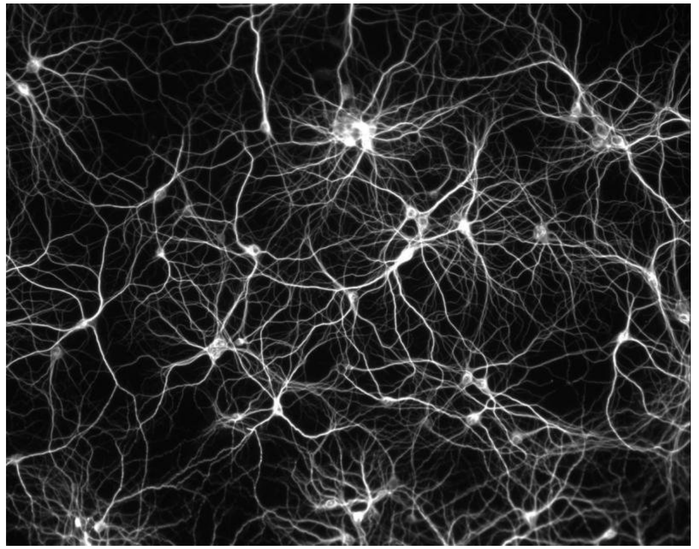






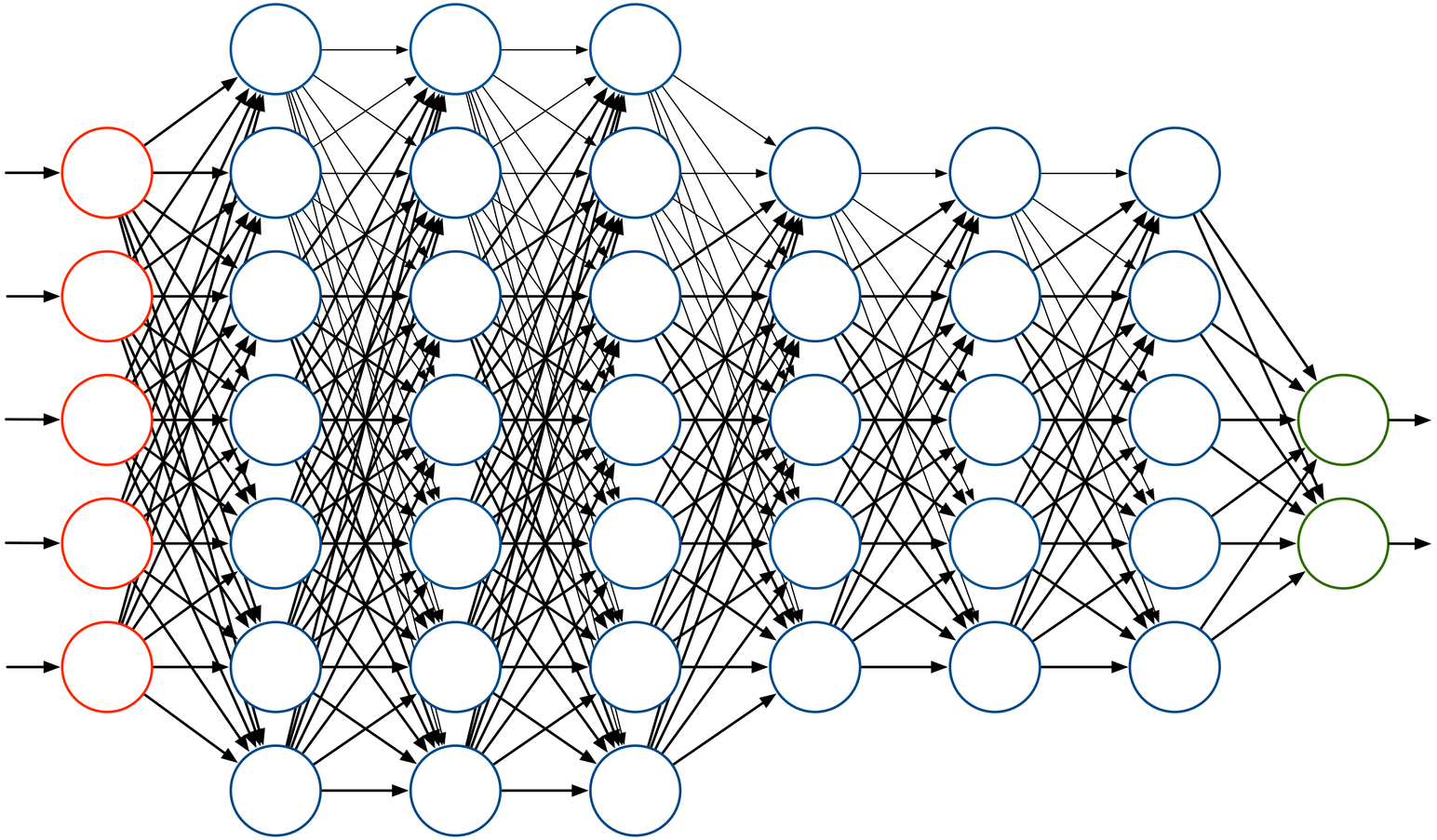
Neural Networks



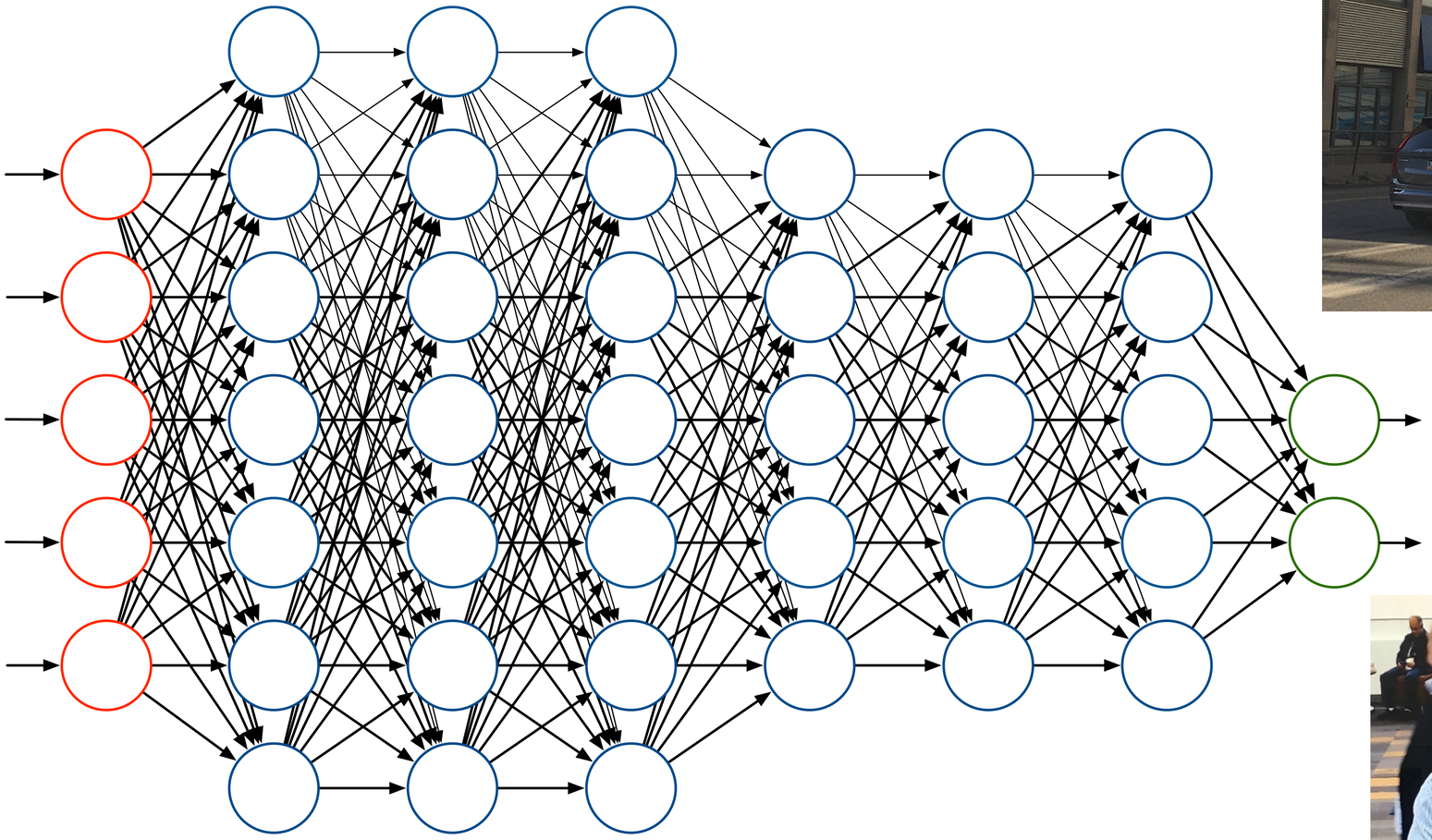


The **universal approximation theorem** states that, under reasonable assumptions, a feedforward **neural network** with a finite number of nodes **can approximate any continuous** function to within a given error over a bounded input domain.

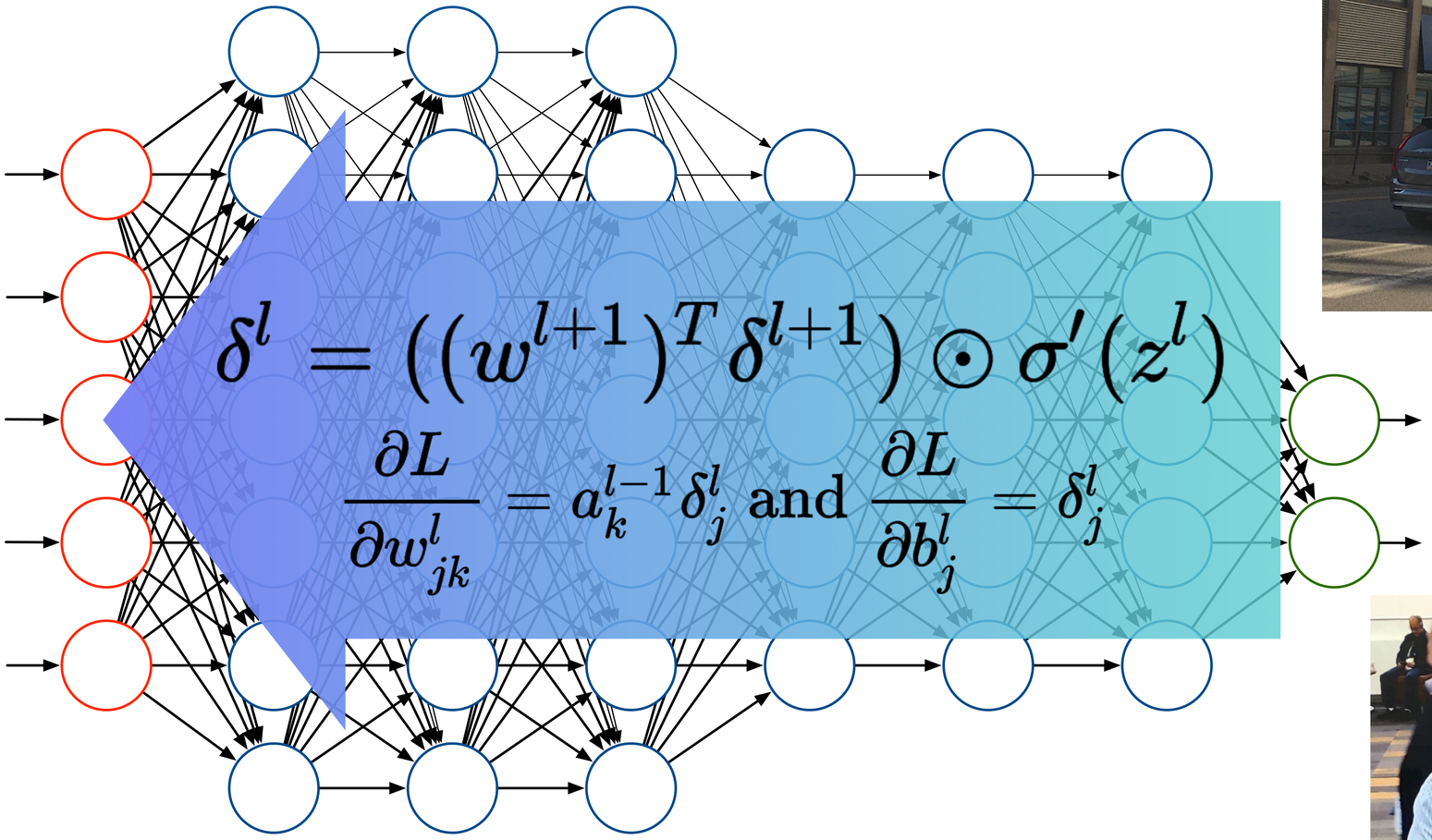
Deep Learning



Deep Learning



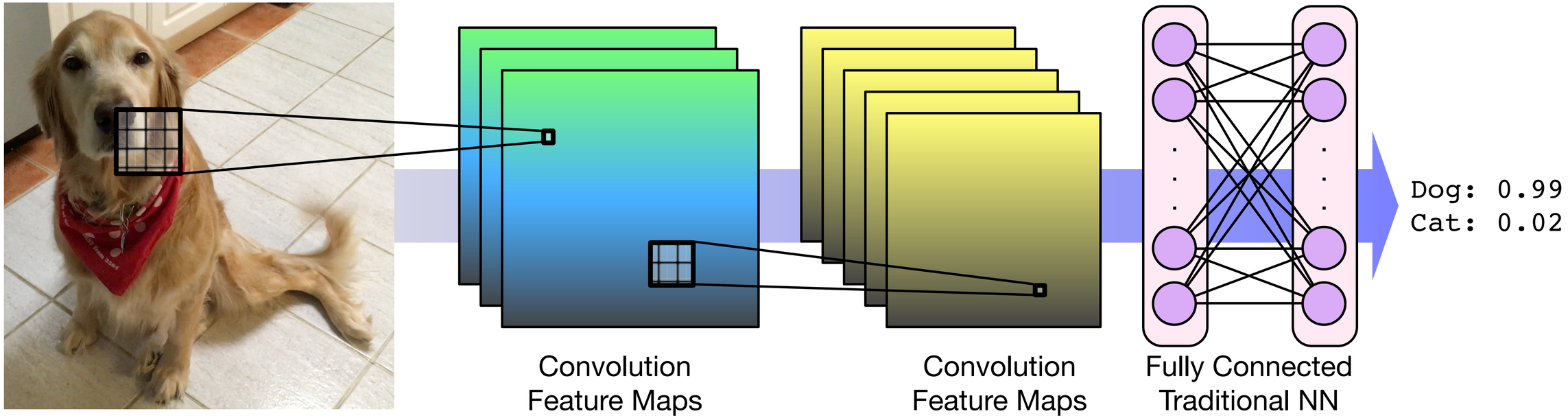
Deep Learning

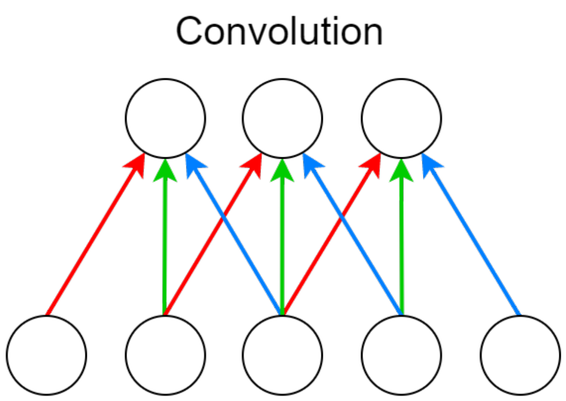


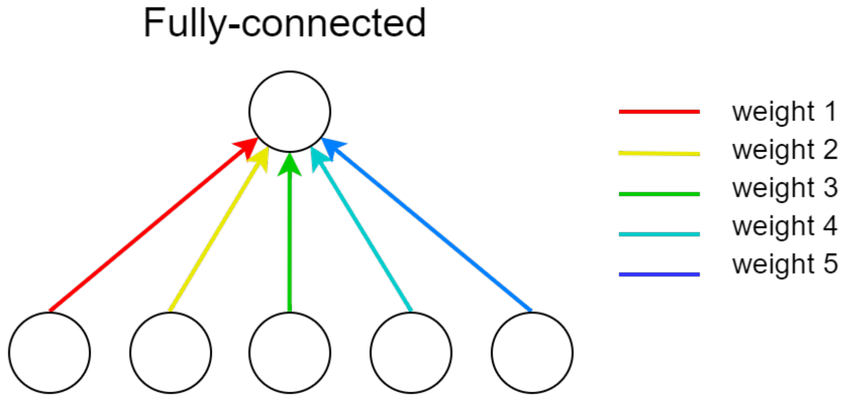
$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$$

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

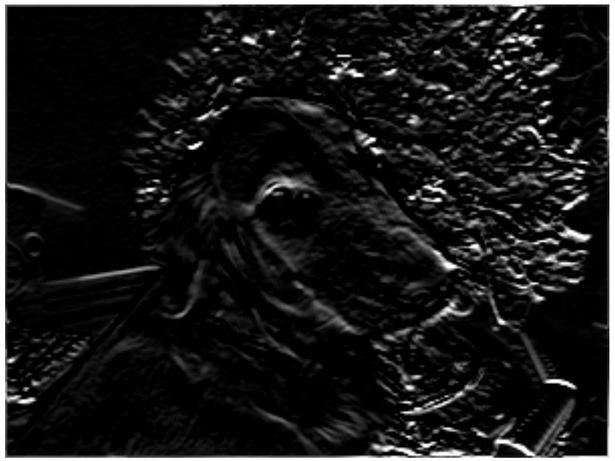
Convolutional Neural Networks







Convolutional Filters

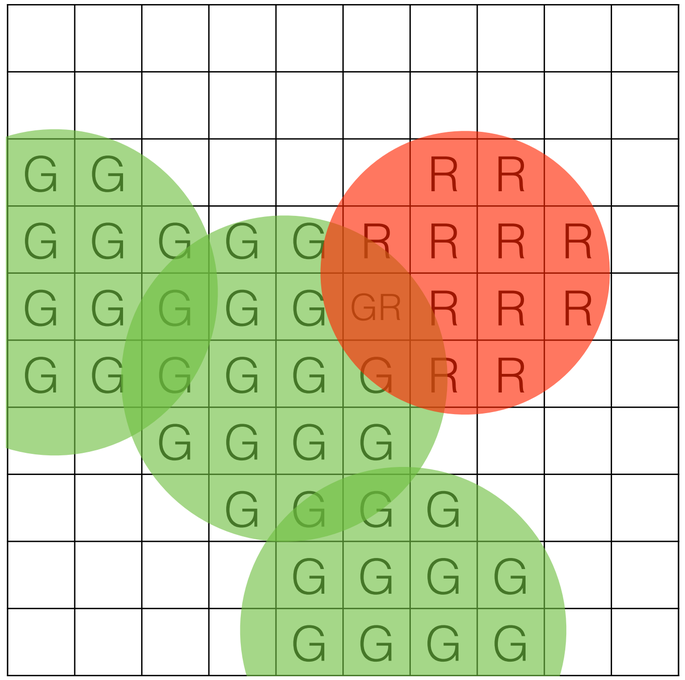


-1	-1	-1
0	0	0
1	1	1

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

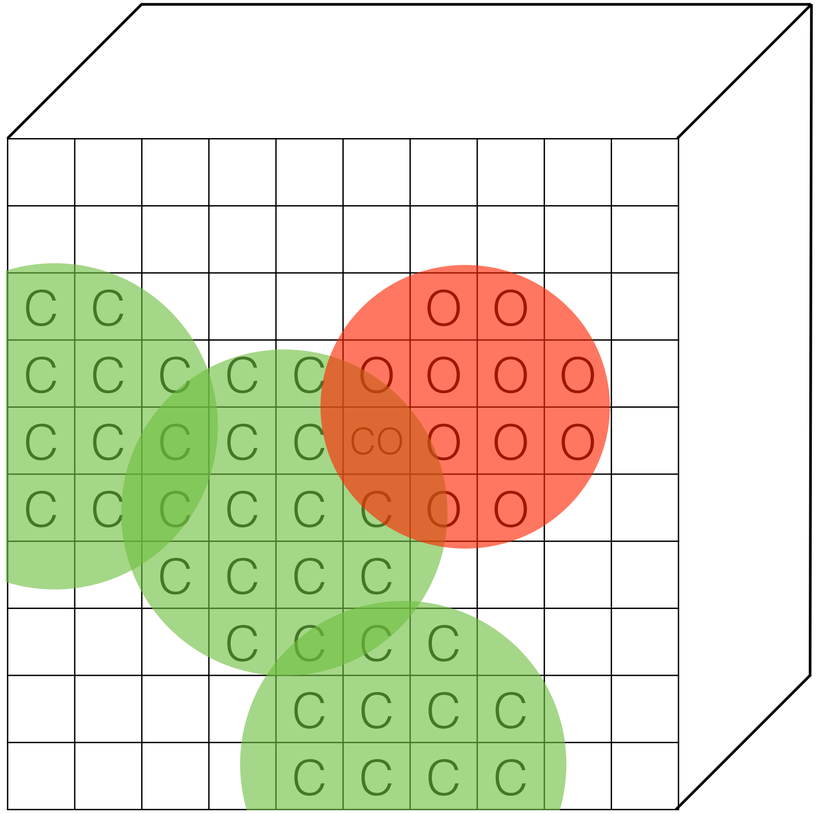
-1	-1	-1
-1	8	-1
-1	-1	-1

Protein-Ligand Representation



(R,G,B) pixel

Protein-Ligand Representation

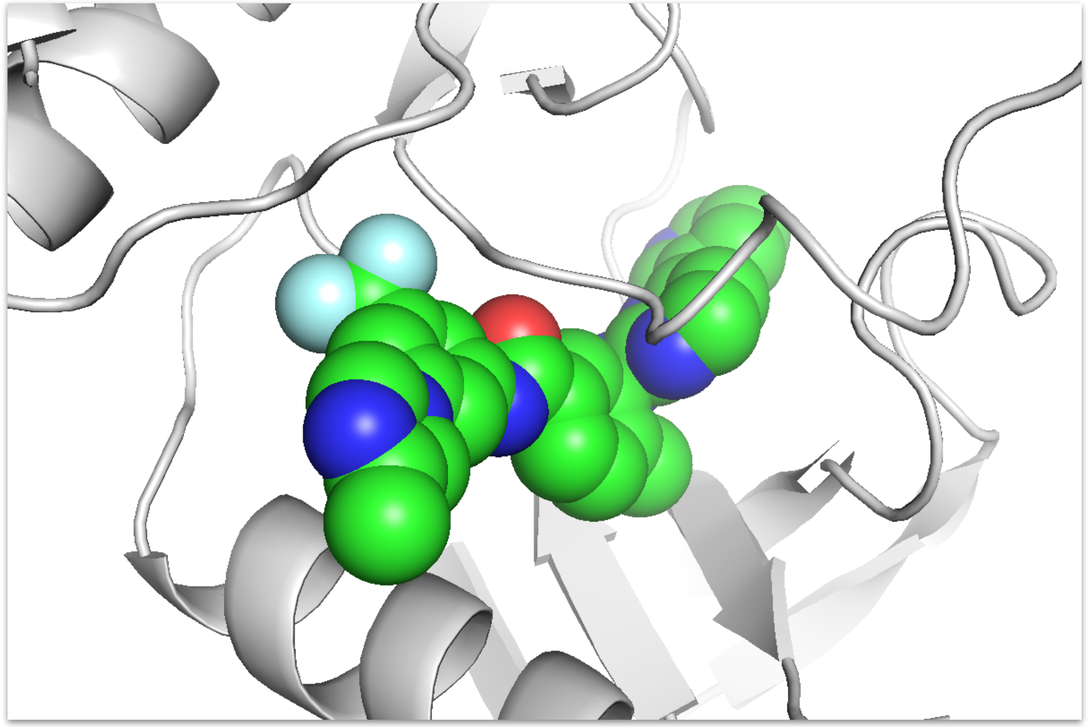


(R,G,B) pixel →

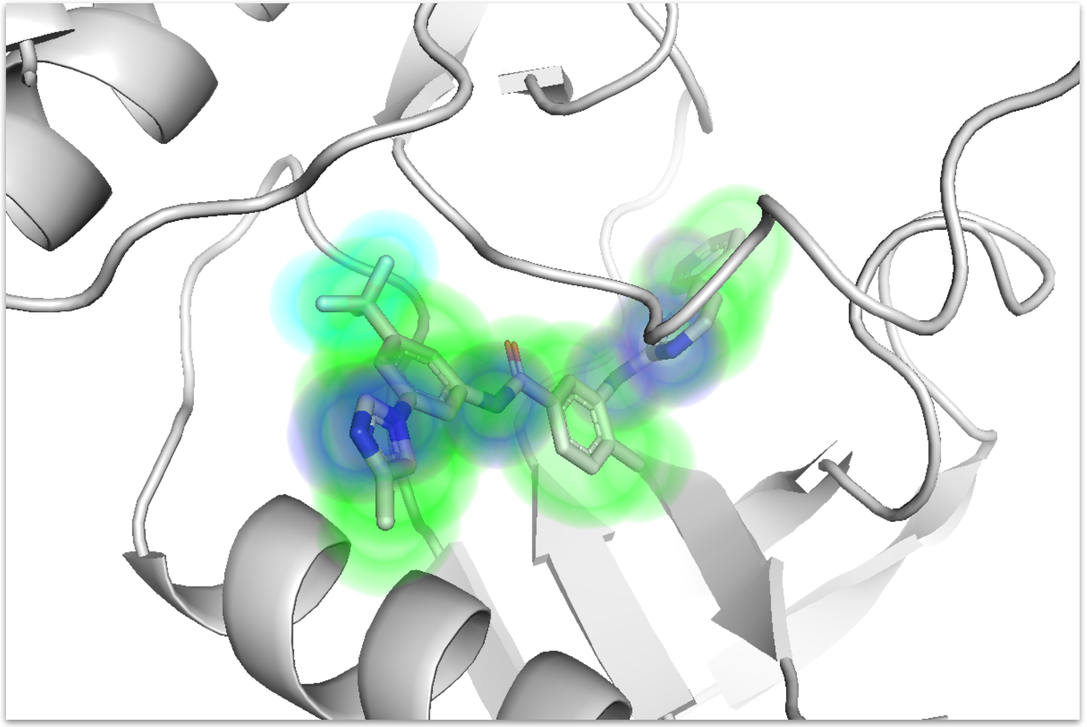
(Carbon, Nitrogen, Oxygen,…) **voxel**

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

Atom Density



Boolean



Gaussian

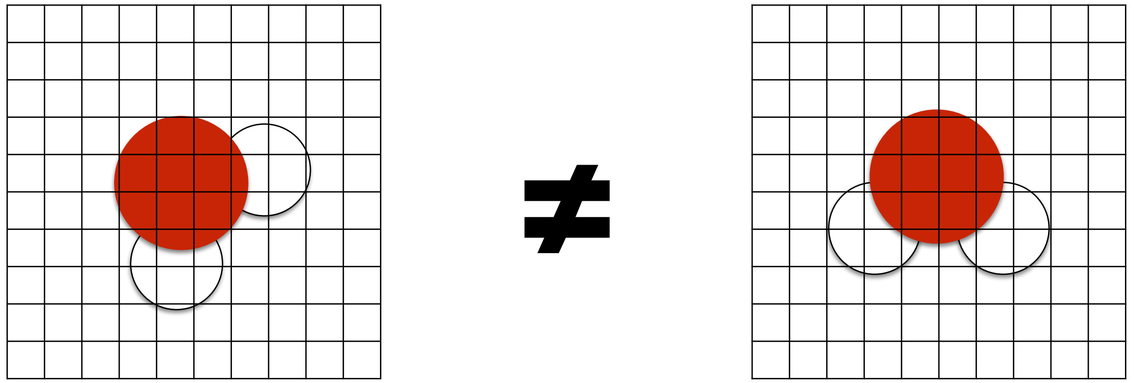
Why Grids?

Cons

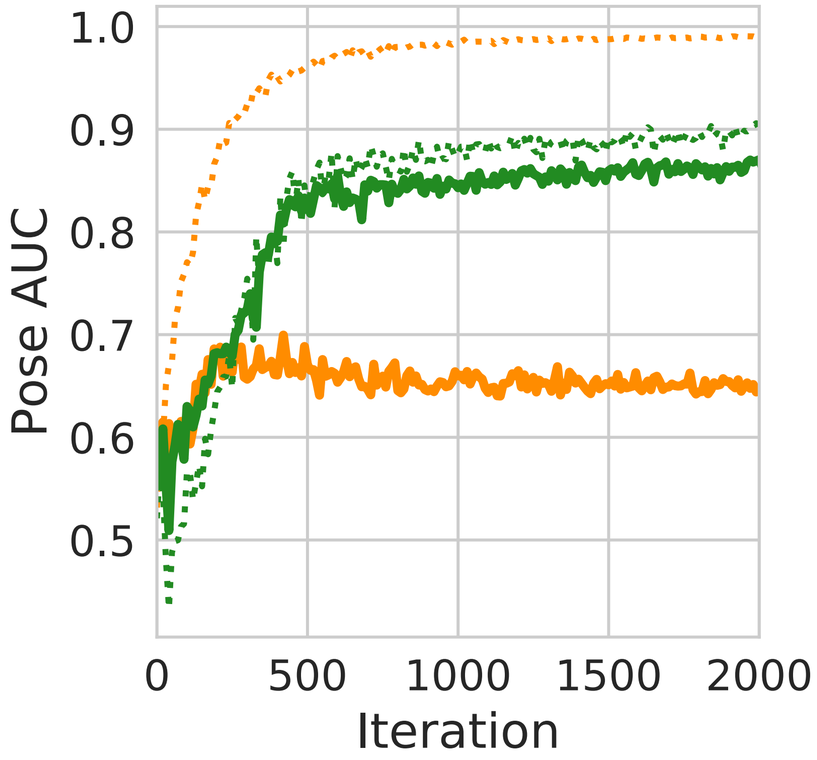
- *coordinate frame dependent*
- pairwise interactions not explicit

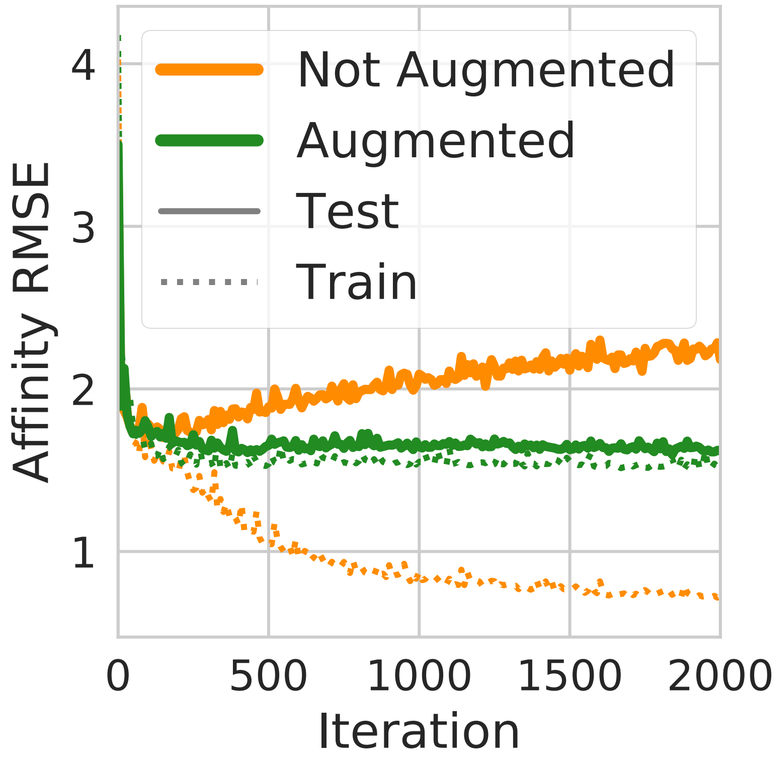
Pros

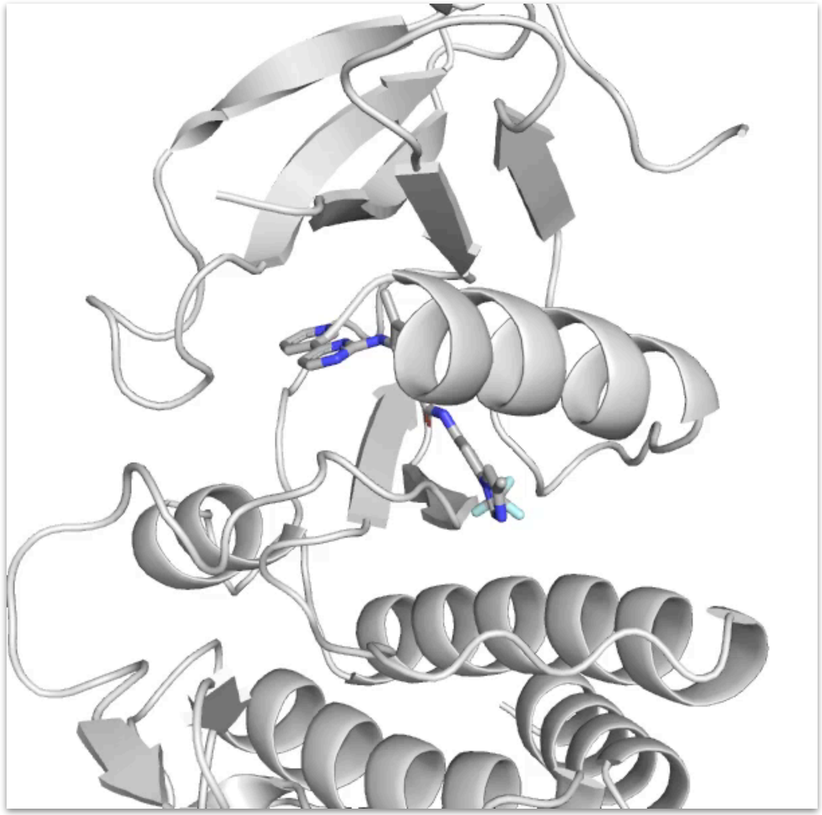
- clear spatial relationships
- amazingly parallel
- easy to interpret



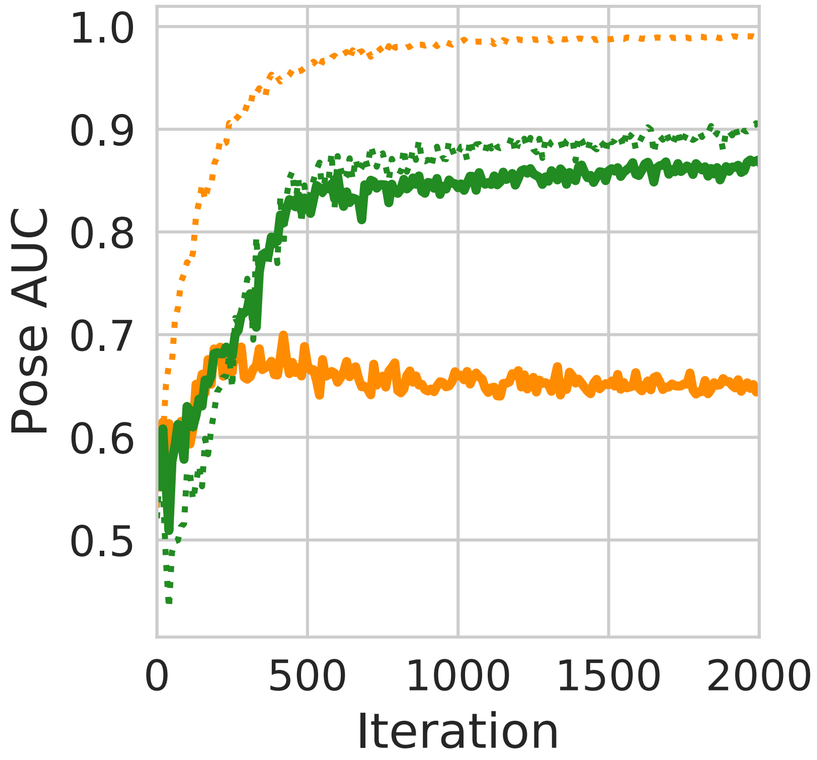
Data Augmentation

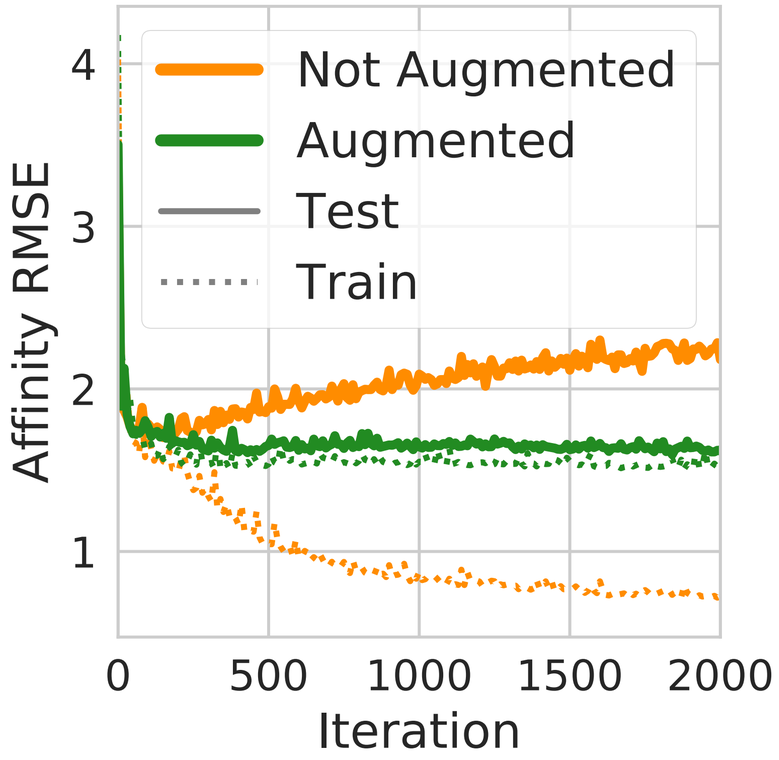


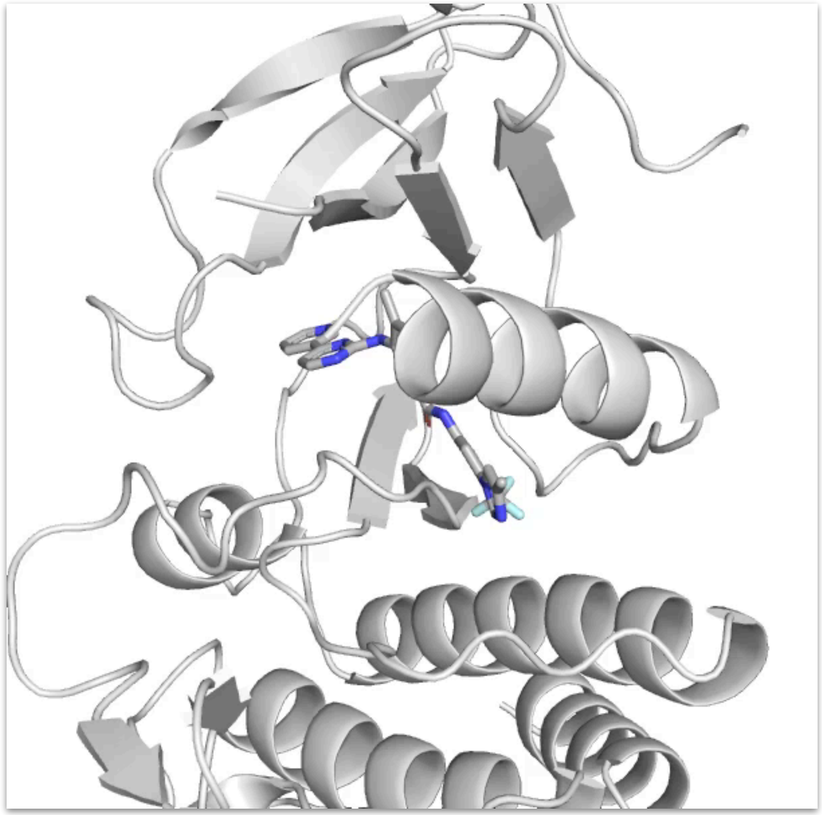




Data Augmentation



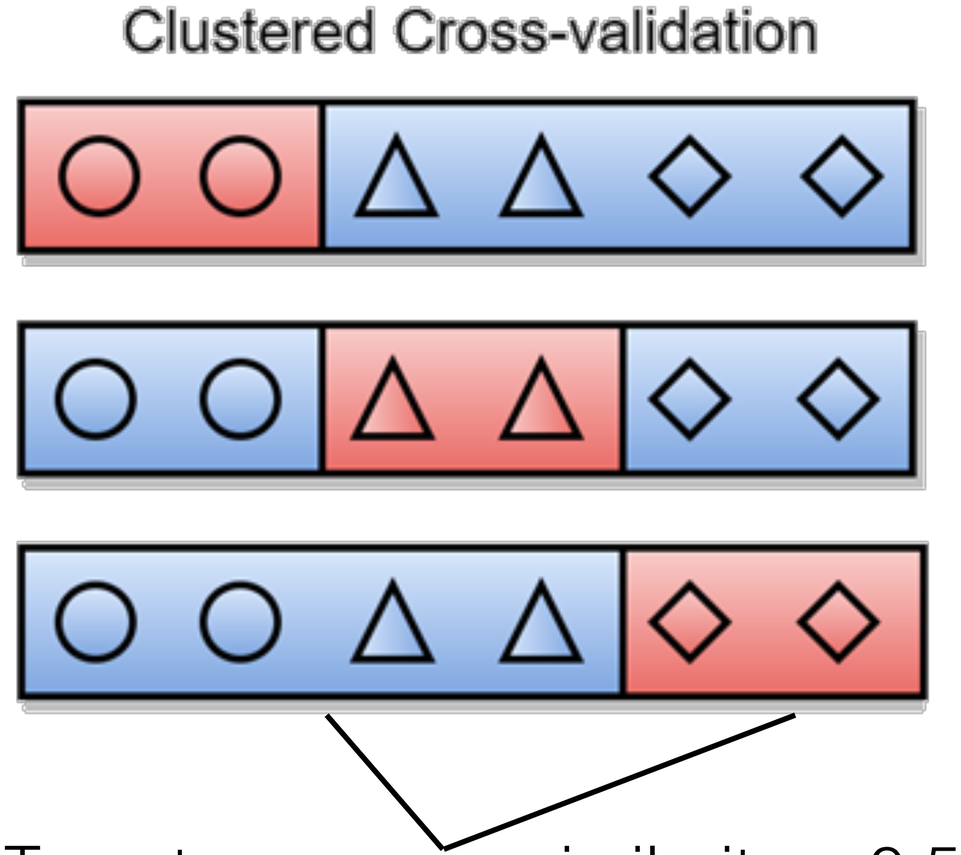




Training

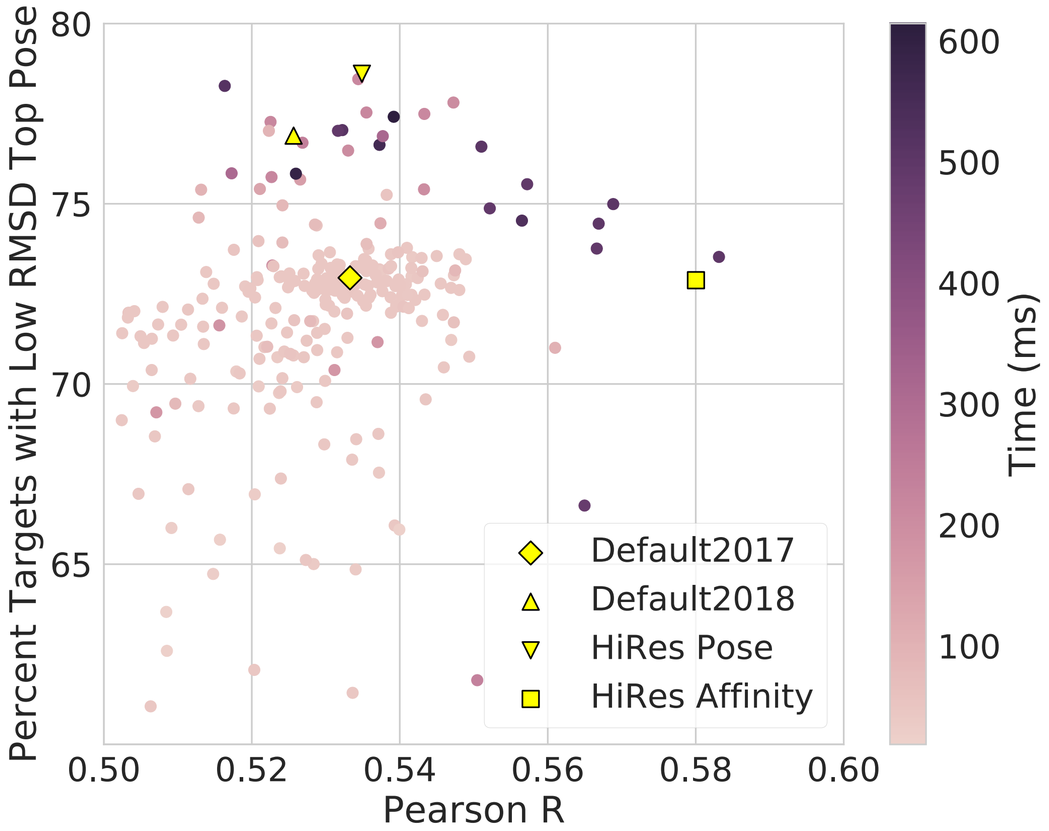
PDBbind 2016 refined set

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

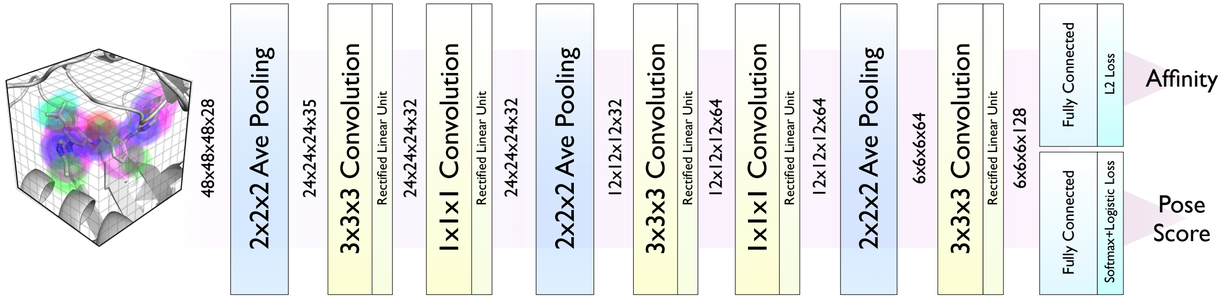


Target sequence similarity < 0.5
AND
Ligand similarity < 0.9

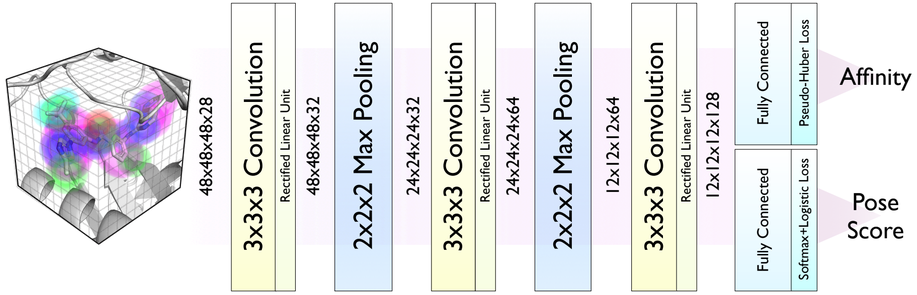
Optimized Models



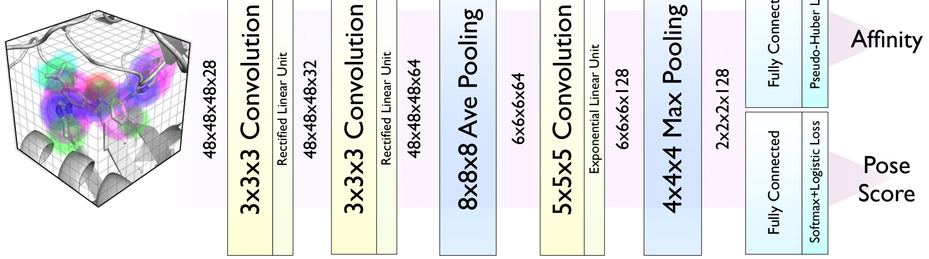
Default2018



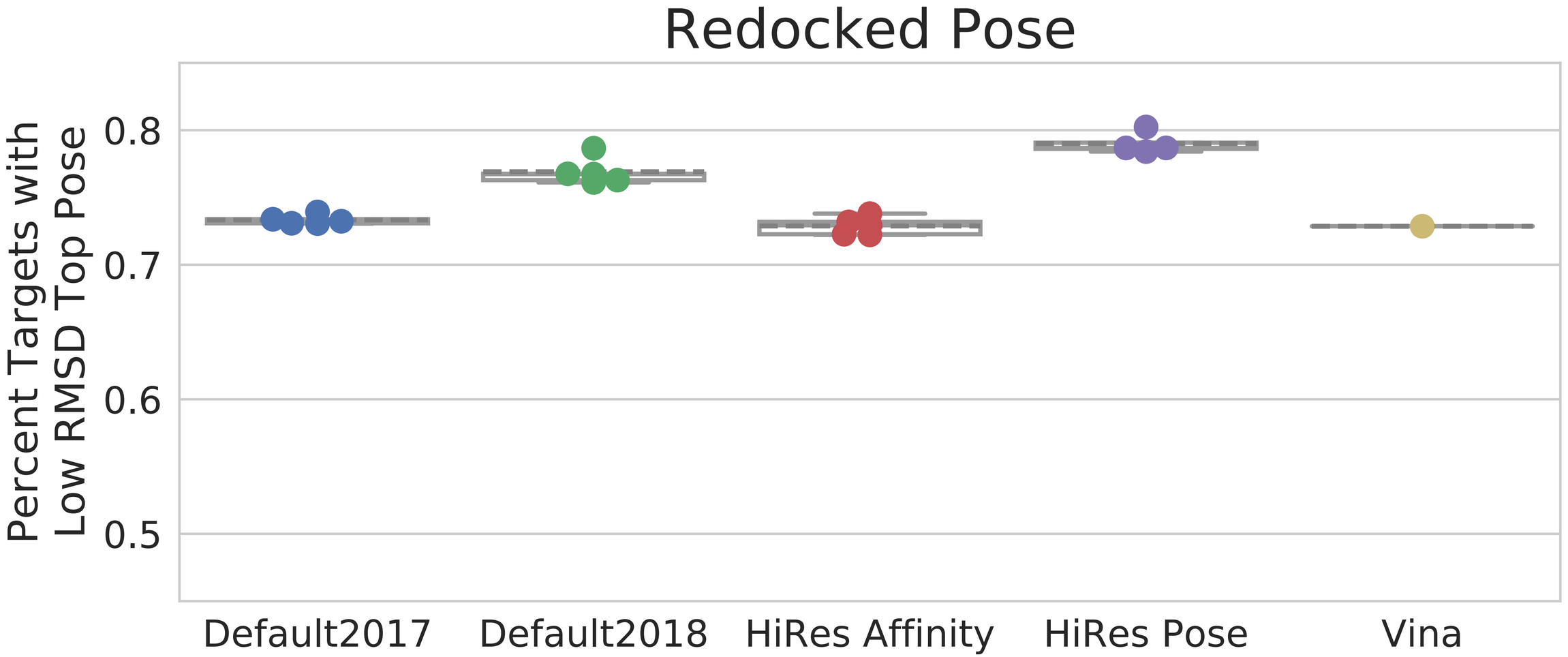
HiRes Pose



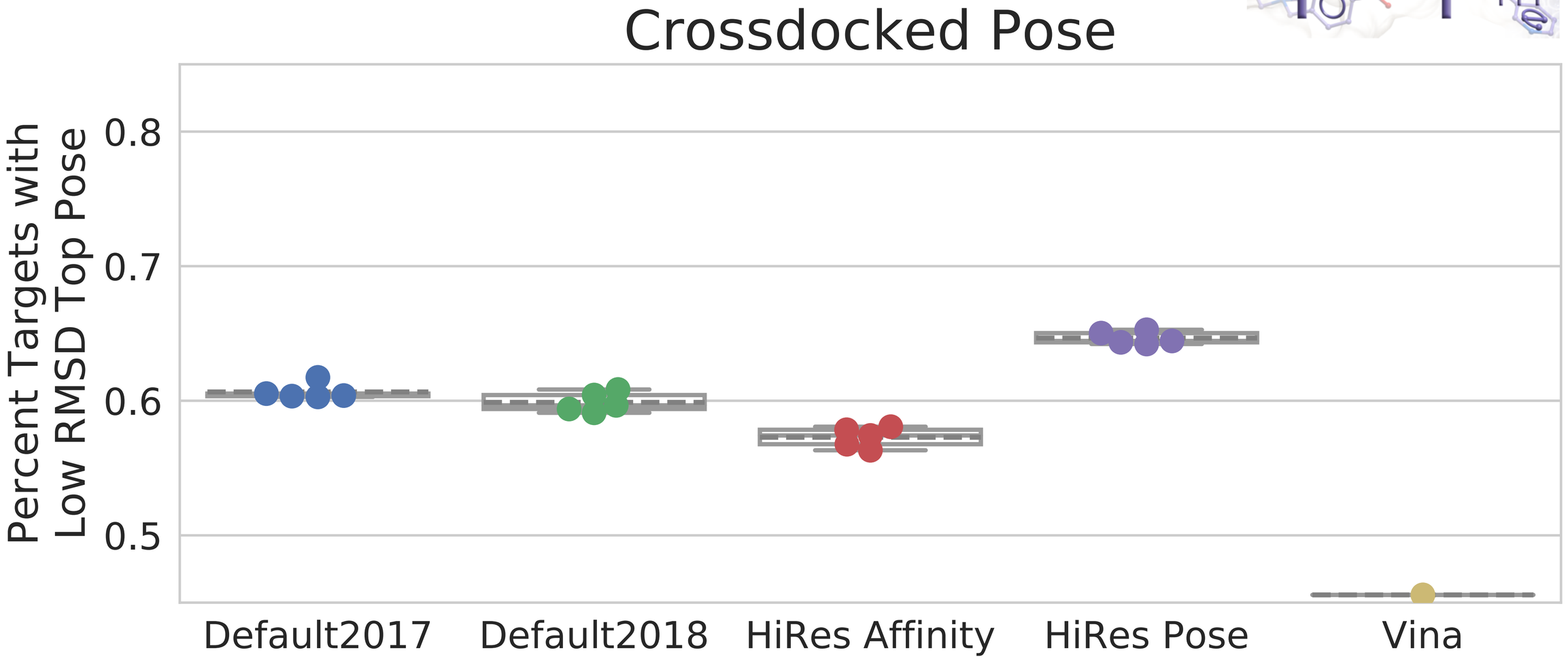
HiRes Affinity



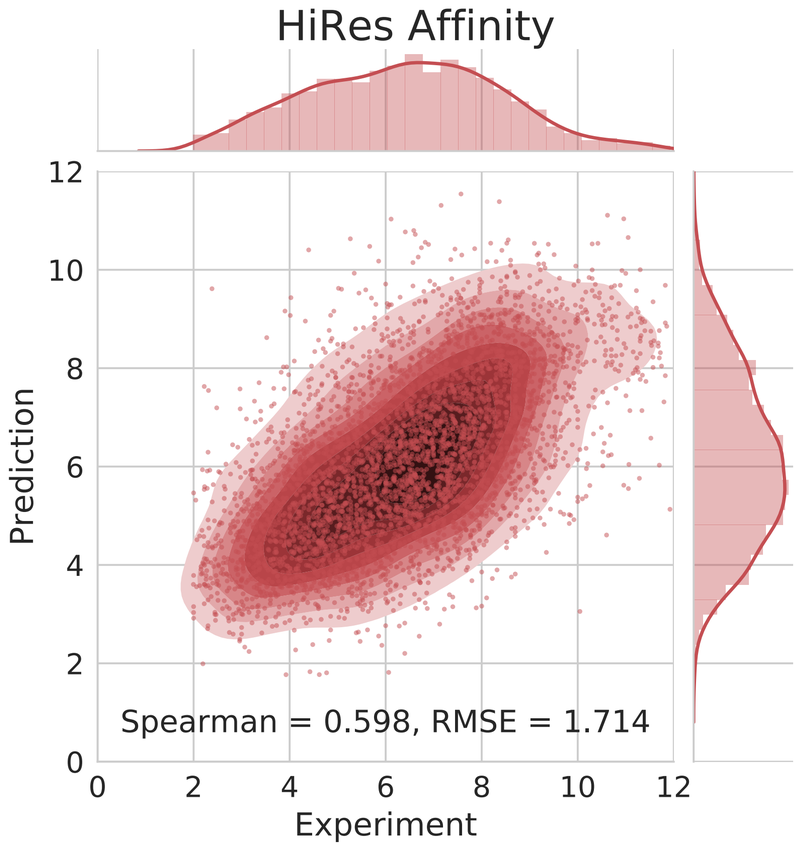
Pose Results

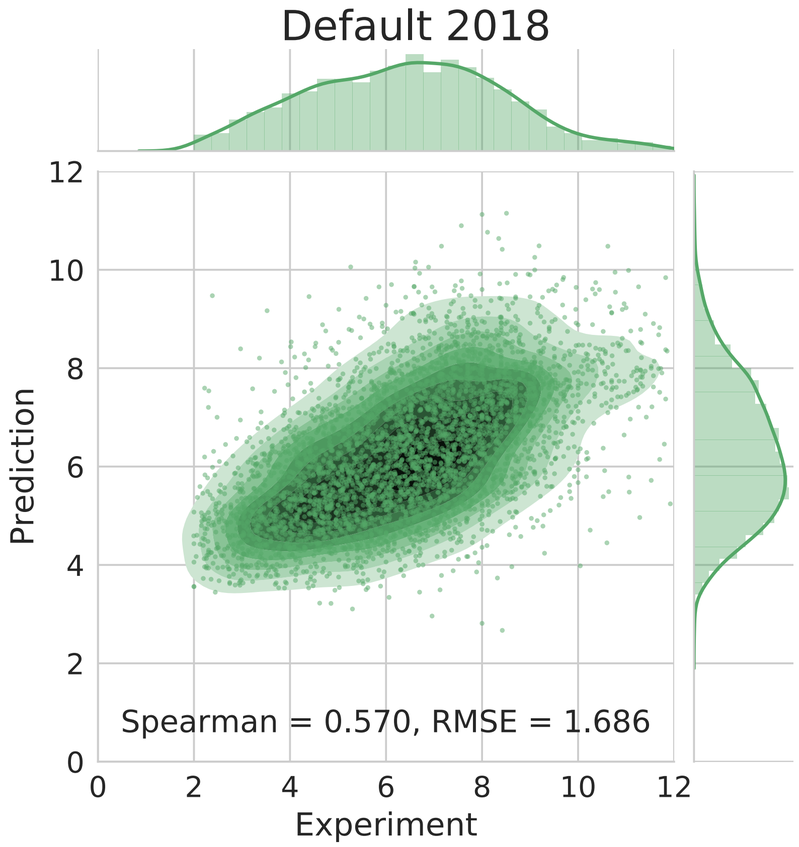


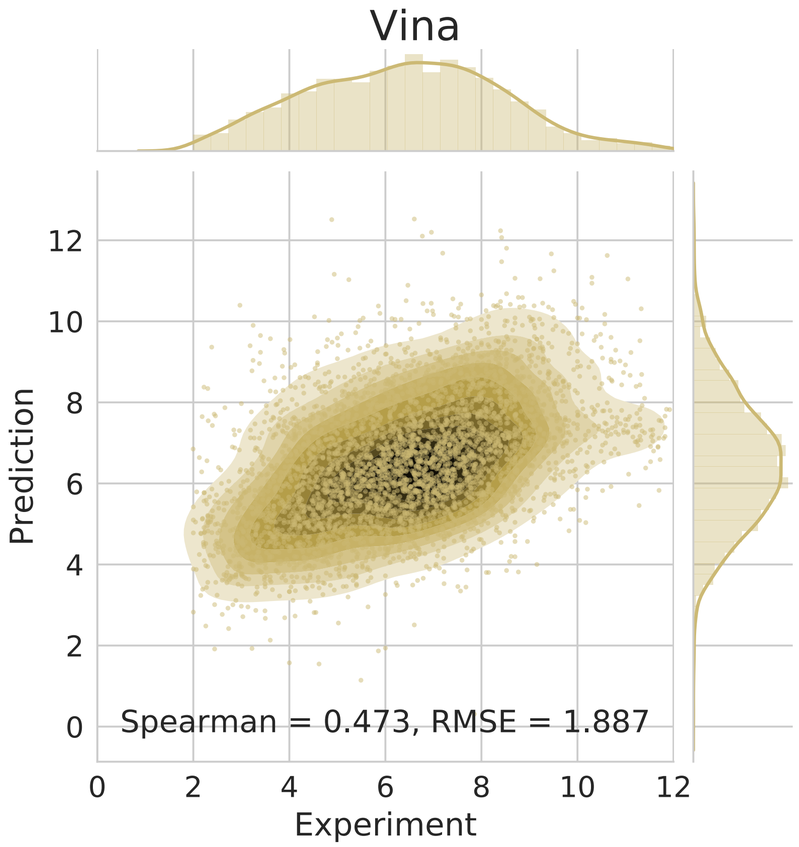
Pose Results



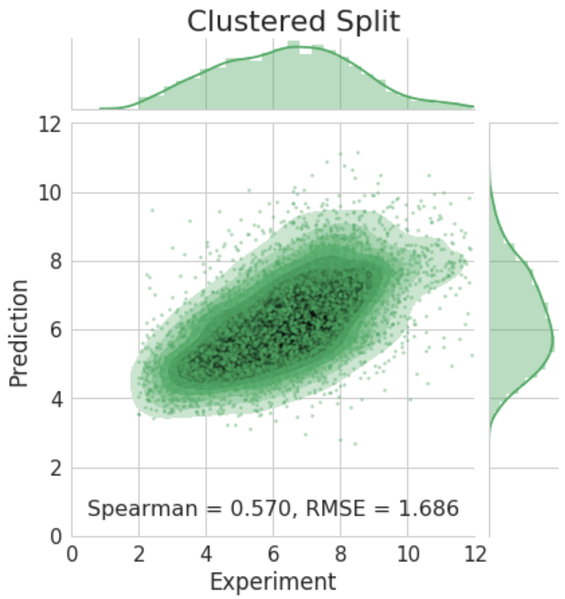
Affinity Results

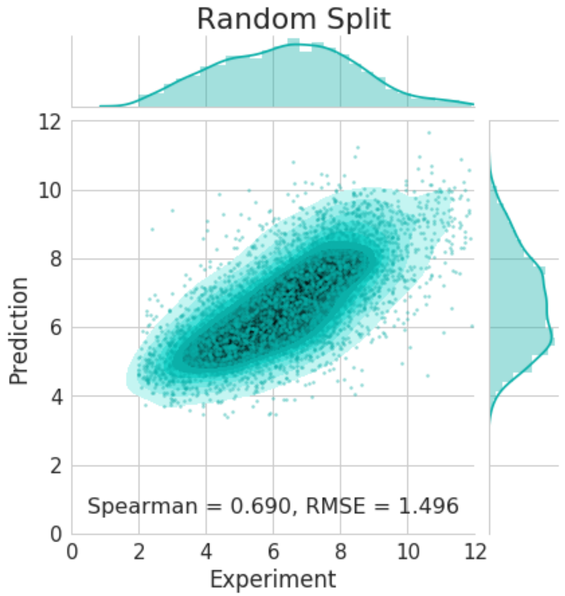


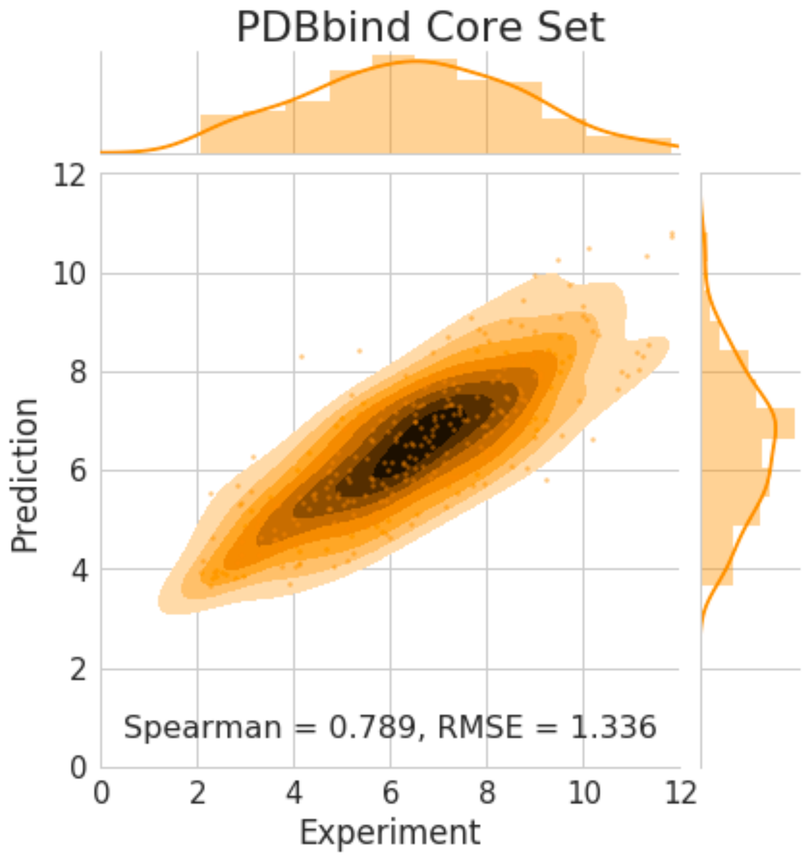


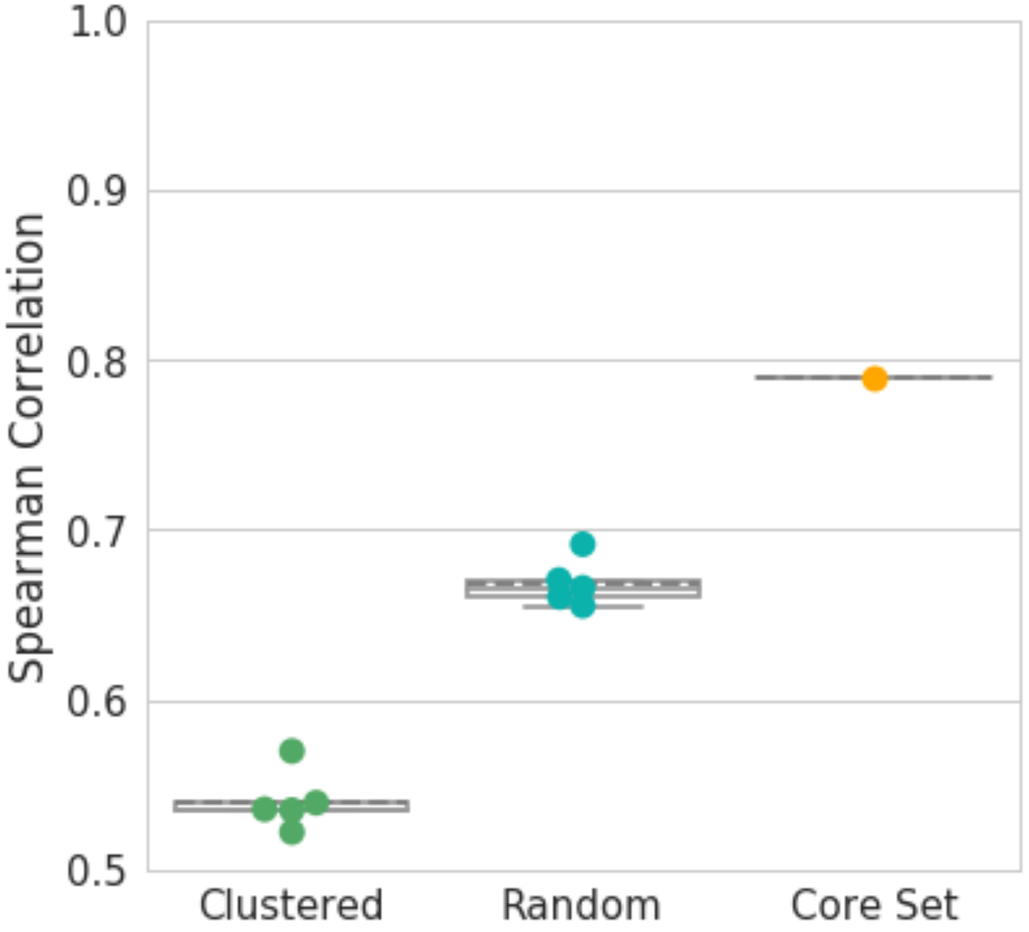


Affinity Results





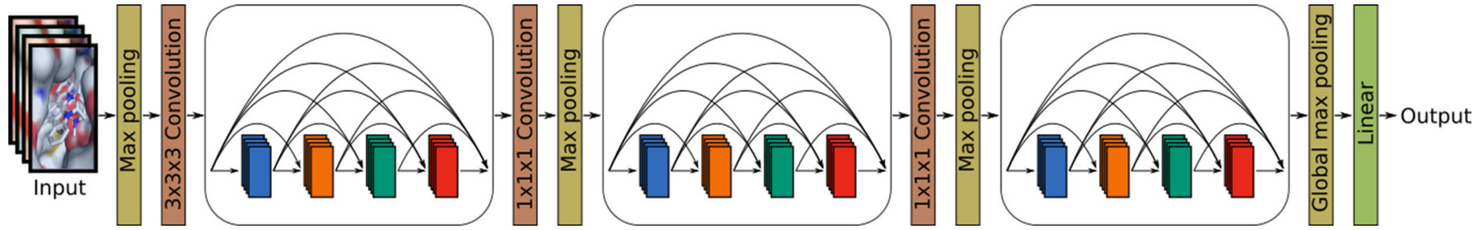


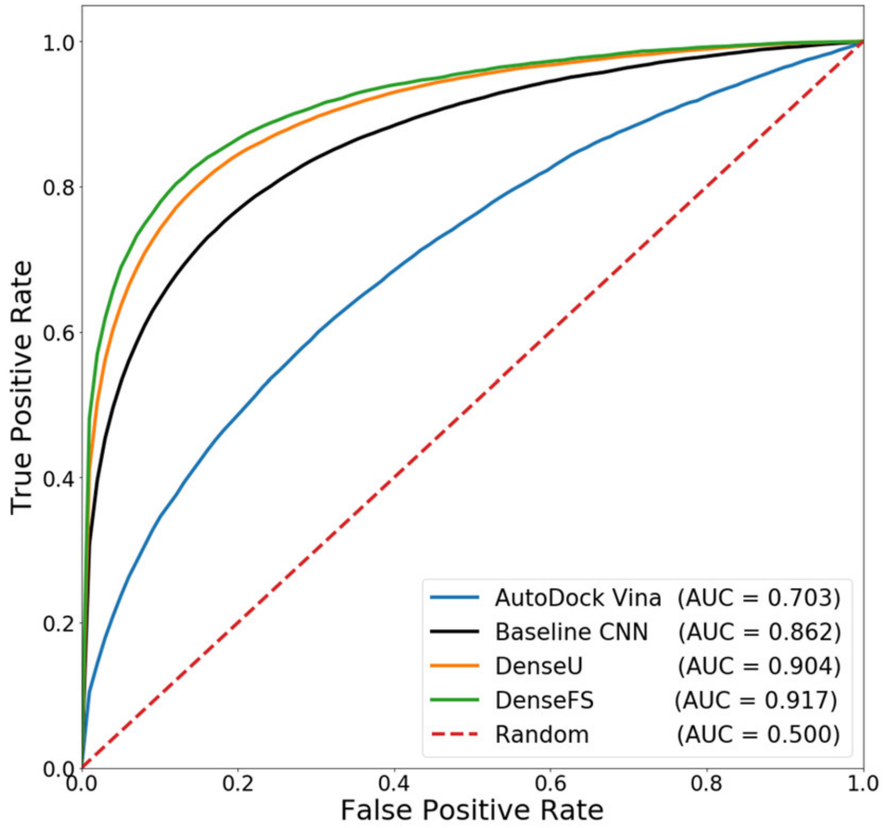


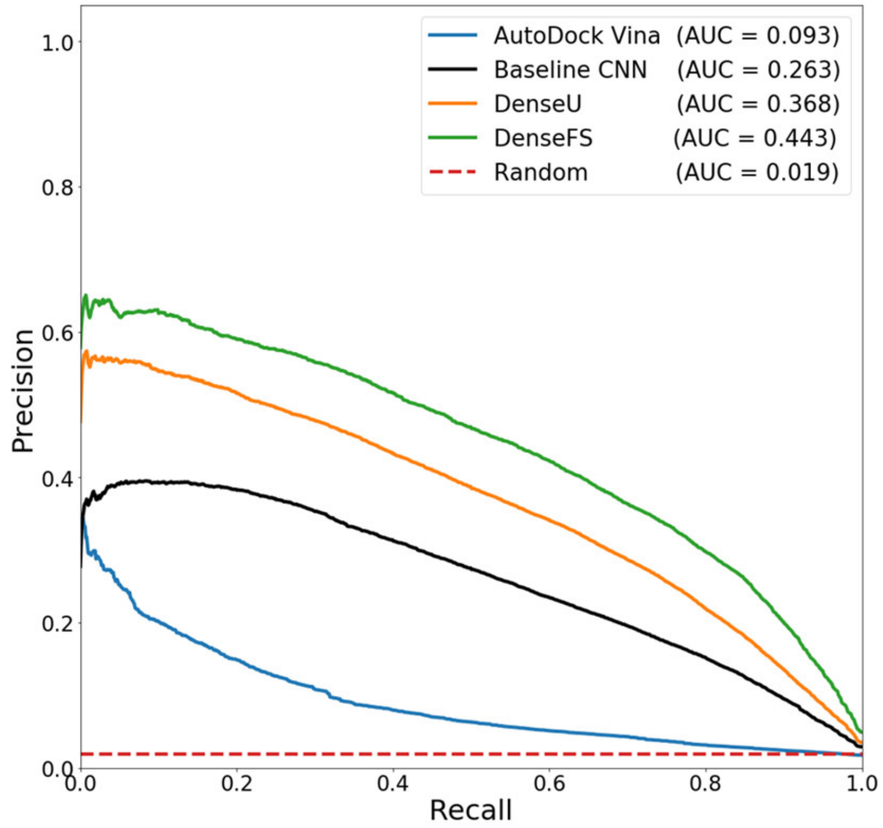
Virtual Screening

Protein Family-Specific Models Using Deep Neural Networks and Transfer Learning Improve Virtual Screening and Highlight the Need for More Data

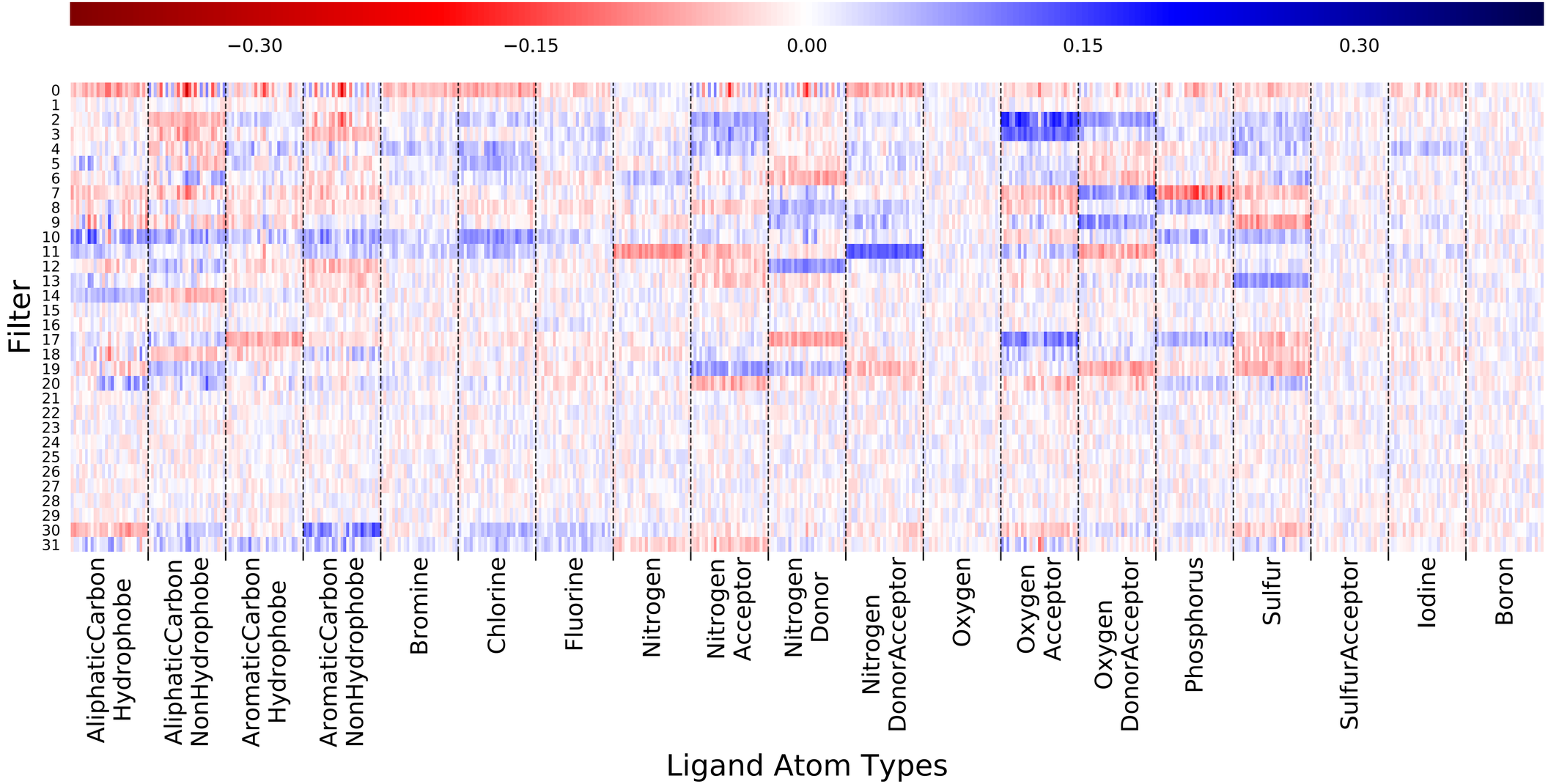
Fergus Imrie[†], Anthony R. Bradley[‡¶§], Mihaela van der Schaar[ǀ⊥], and Charlotte M. Deane[*†]
[†] Oxford Protein Informatics Group, Department of Statistics, University of Oxford, Oxford OX1 3LB, U.K.
[‡] Structural Genomics Consortium, University of Oxford, Oxford OX3 7DQ, U.K.
[¶] Department of Chemistry, University of Oxford, Oxford OX1 3TA, U.K.
[§] Diamond Light Source Ltd., Didcot OX11 0DE, U.K.
[ǀ] Department of Engineering, University of Oxford, Oxford OX1 3PJ, U.K.
[⊥] Alan Turing Institute, London NW1 2DB, U.K.



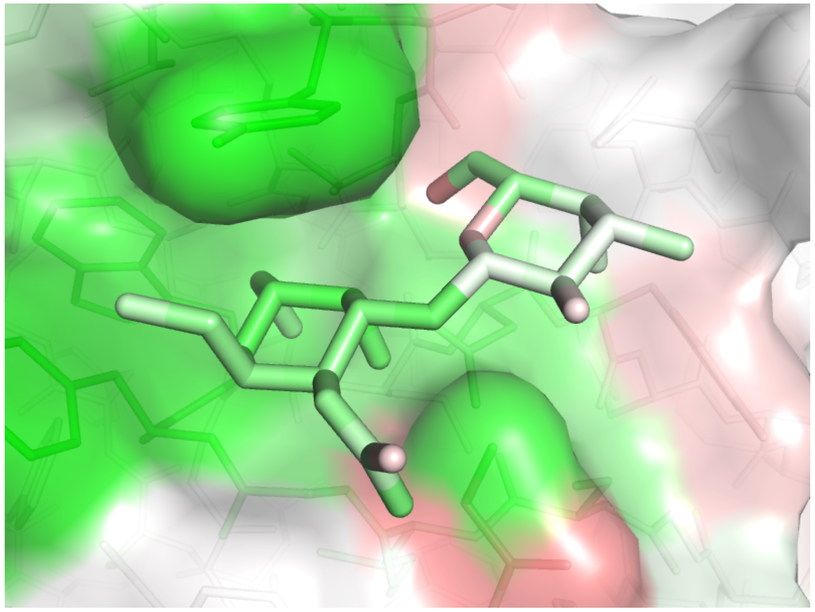




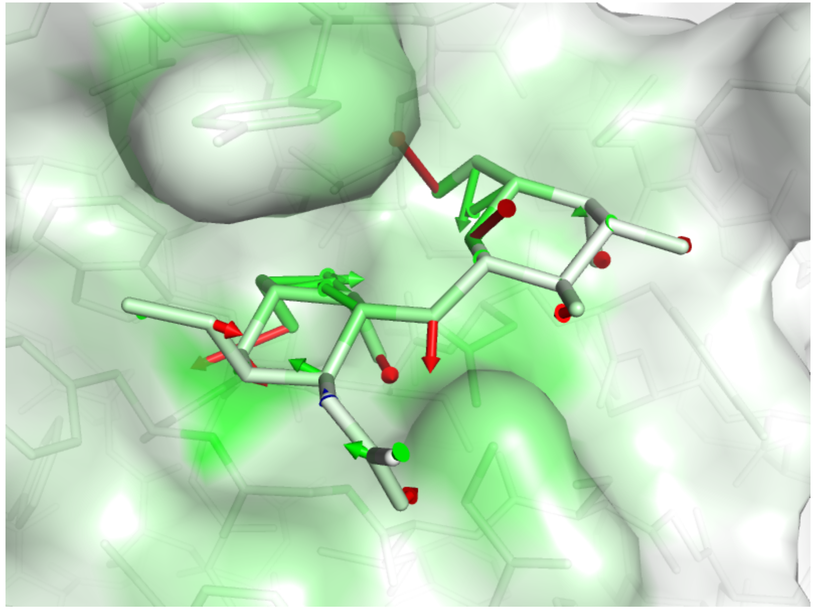
Filter Visualization



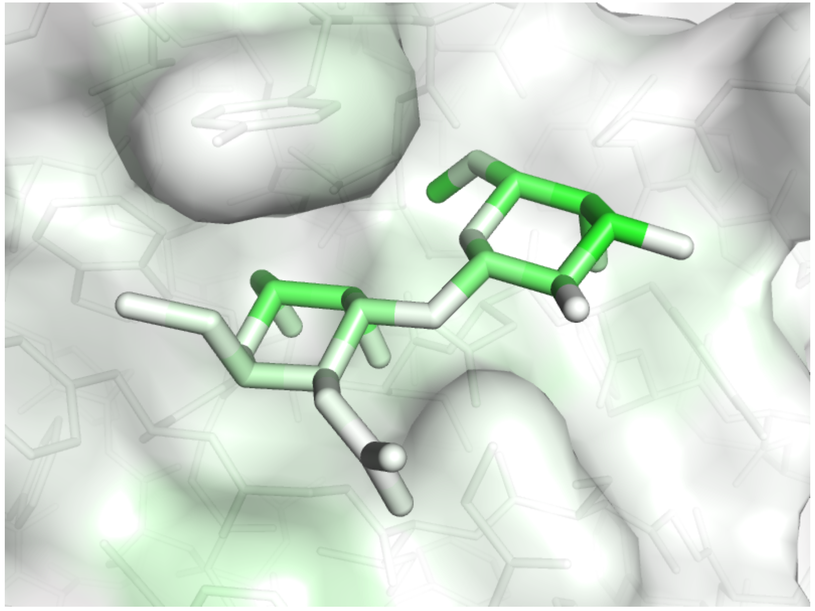
Visualization



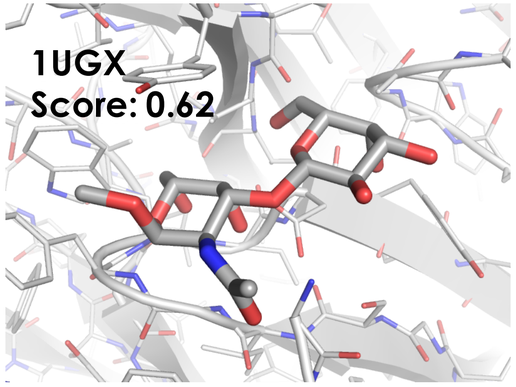
masking



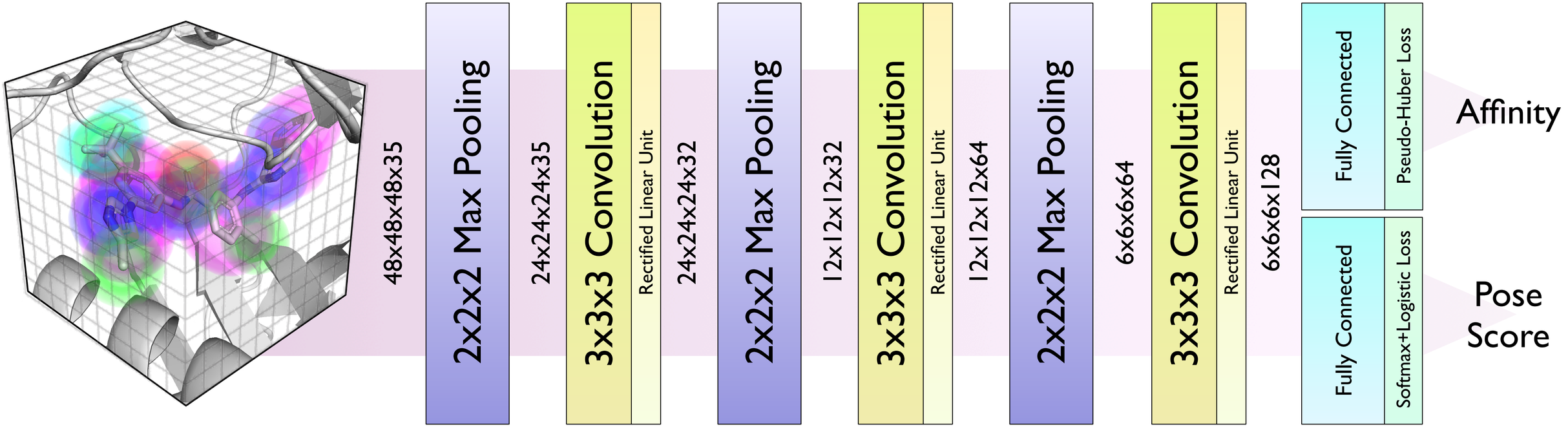
gradients



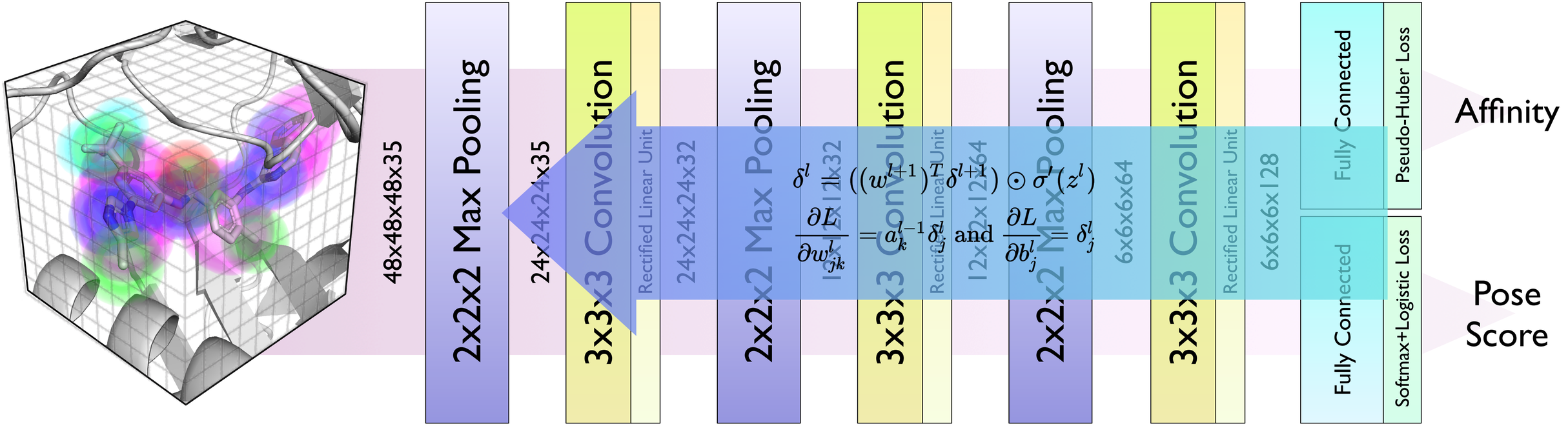
layer-wise relevance



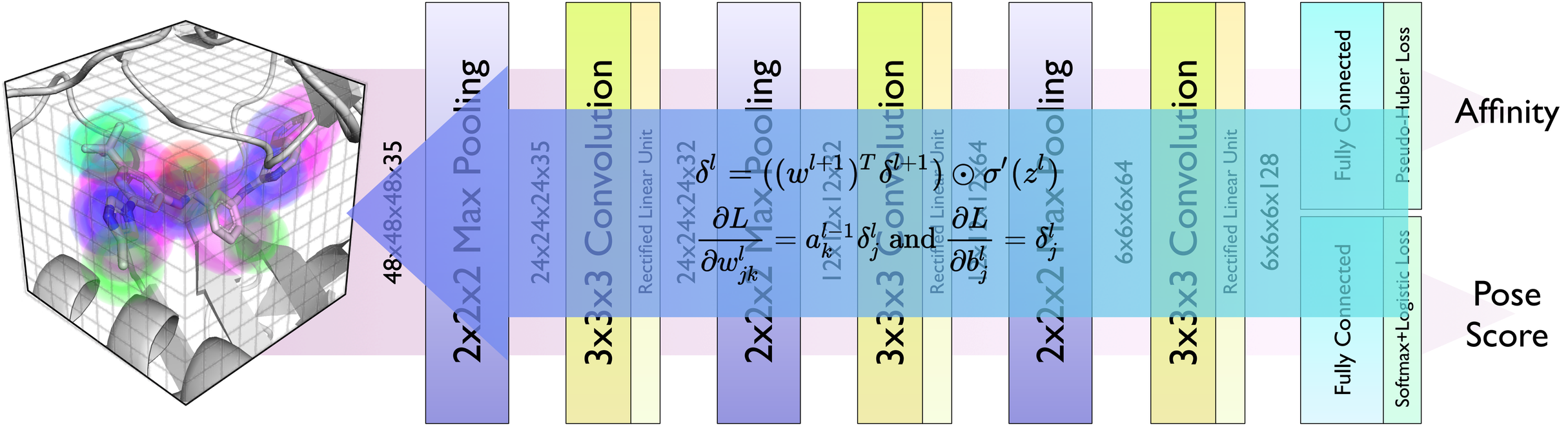
Beyond Scoring



Beyond Scoring



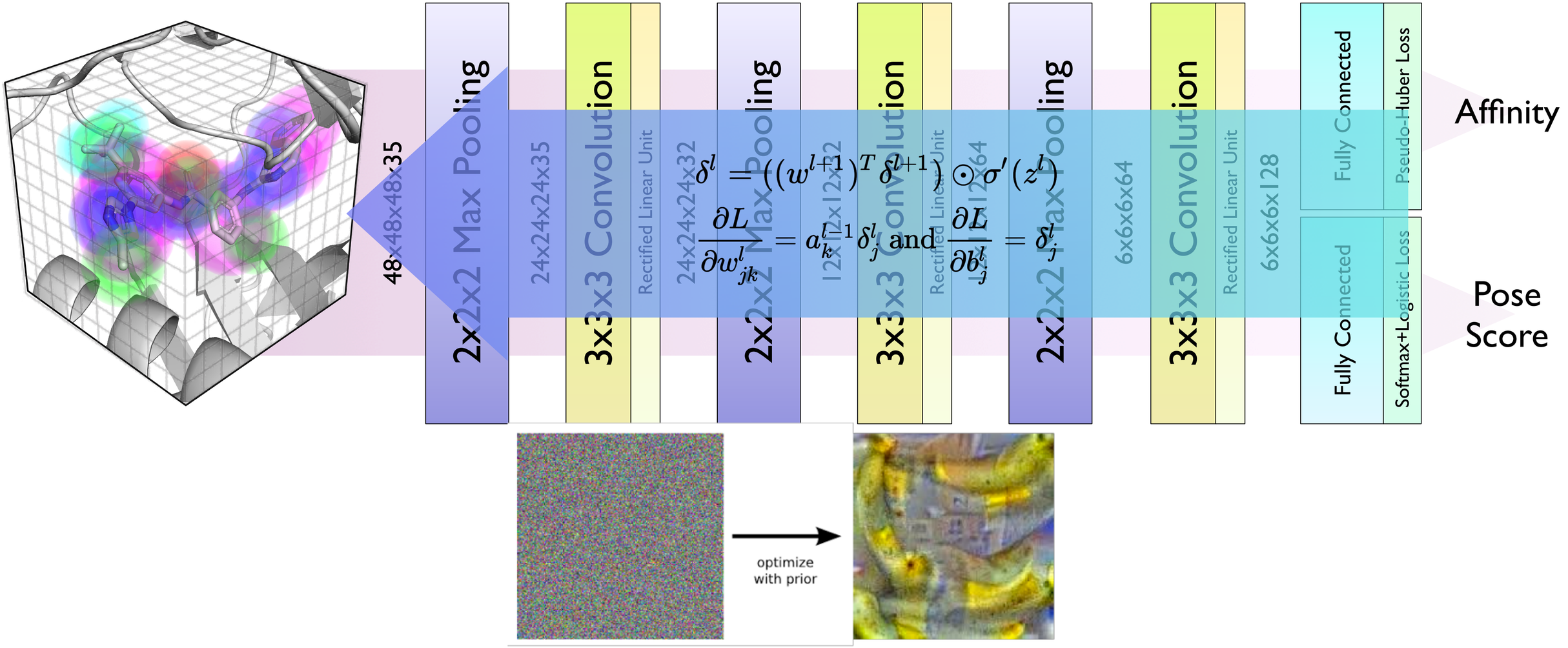
Beyond Scoring



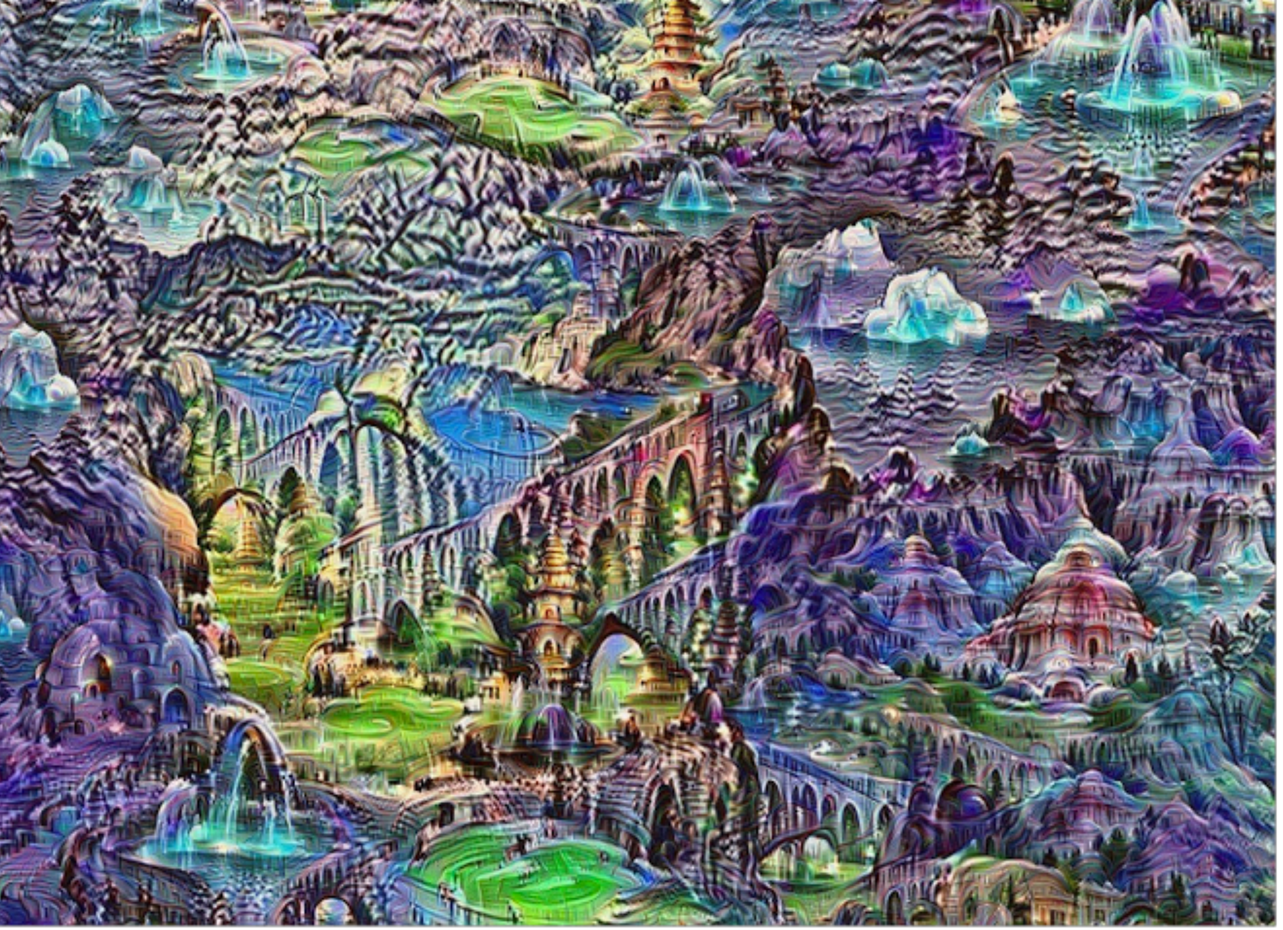
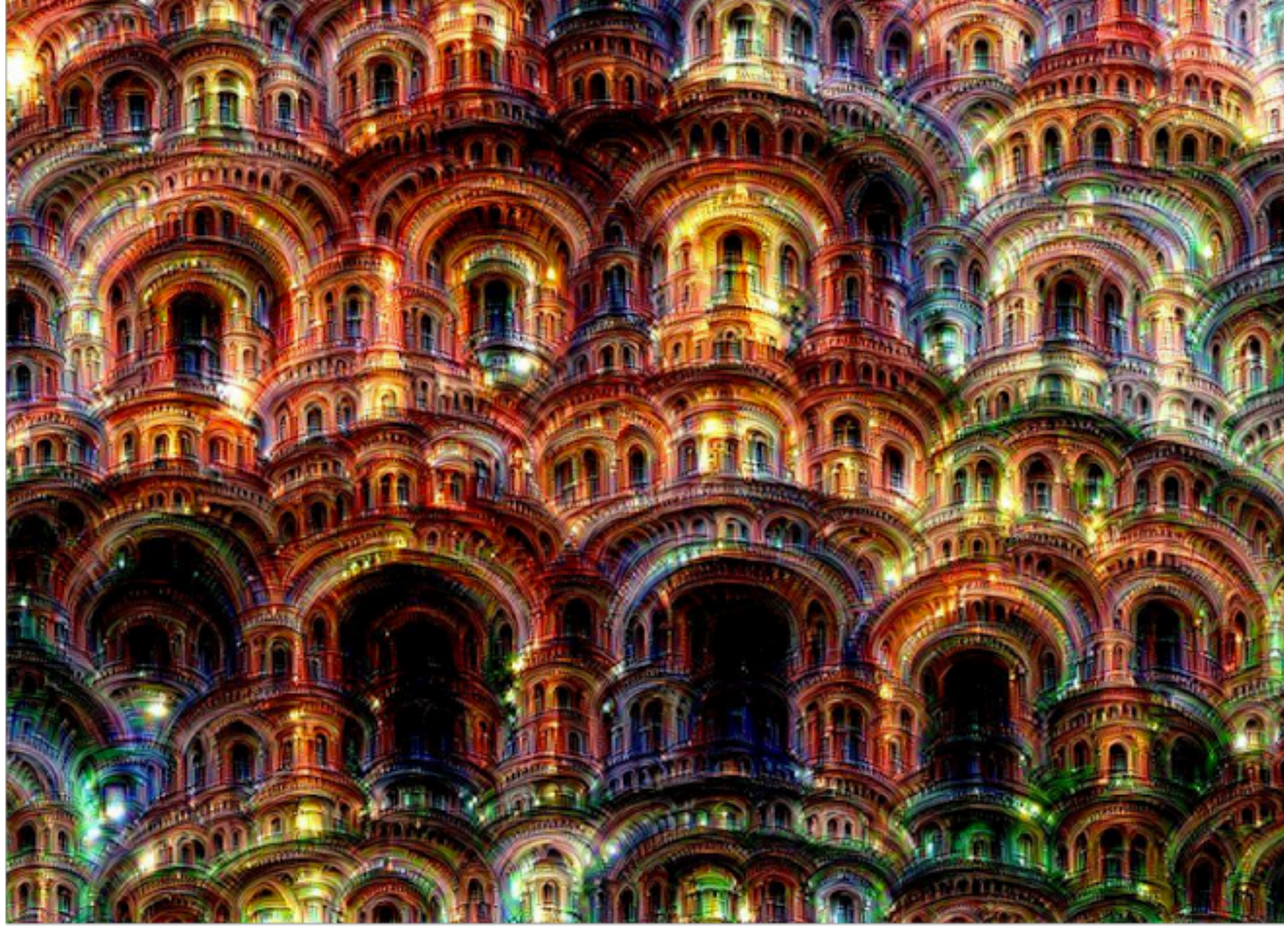
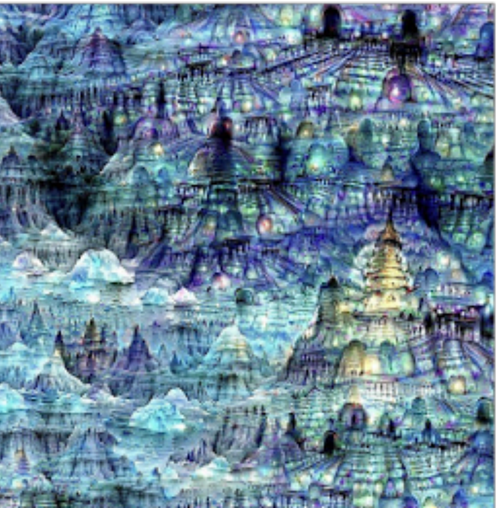
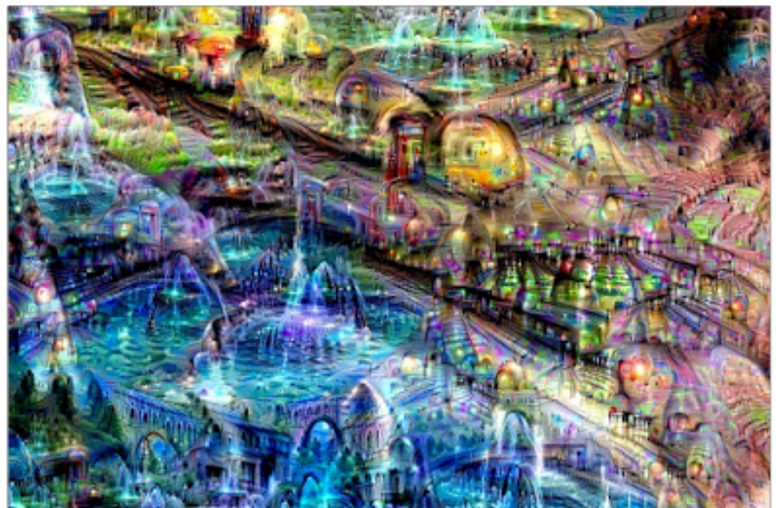
$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$$

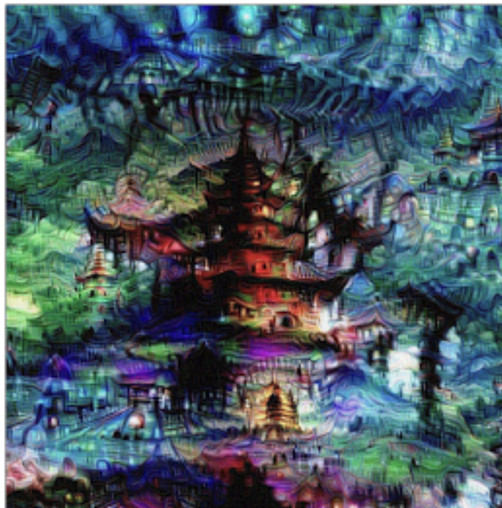
$$\frac{\partial L}{\partial w^l_{jk}} = a^{l-1}_k \delta^l_j \text{ and } \frac{\partial L}{\partial b^l_j} = \delta^l_j$$

Beyond Scoring

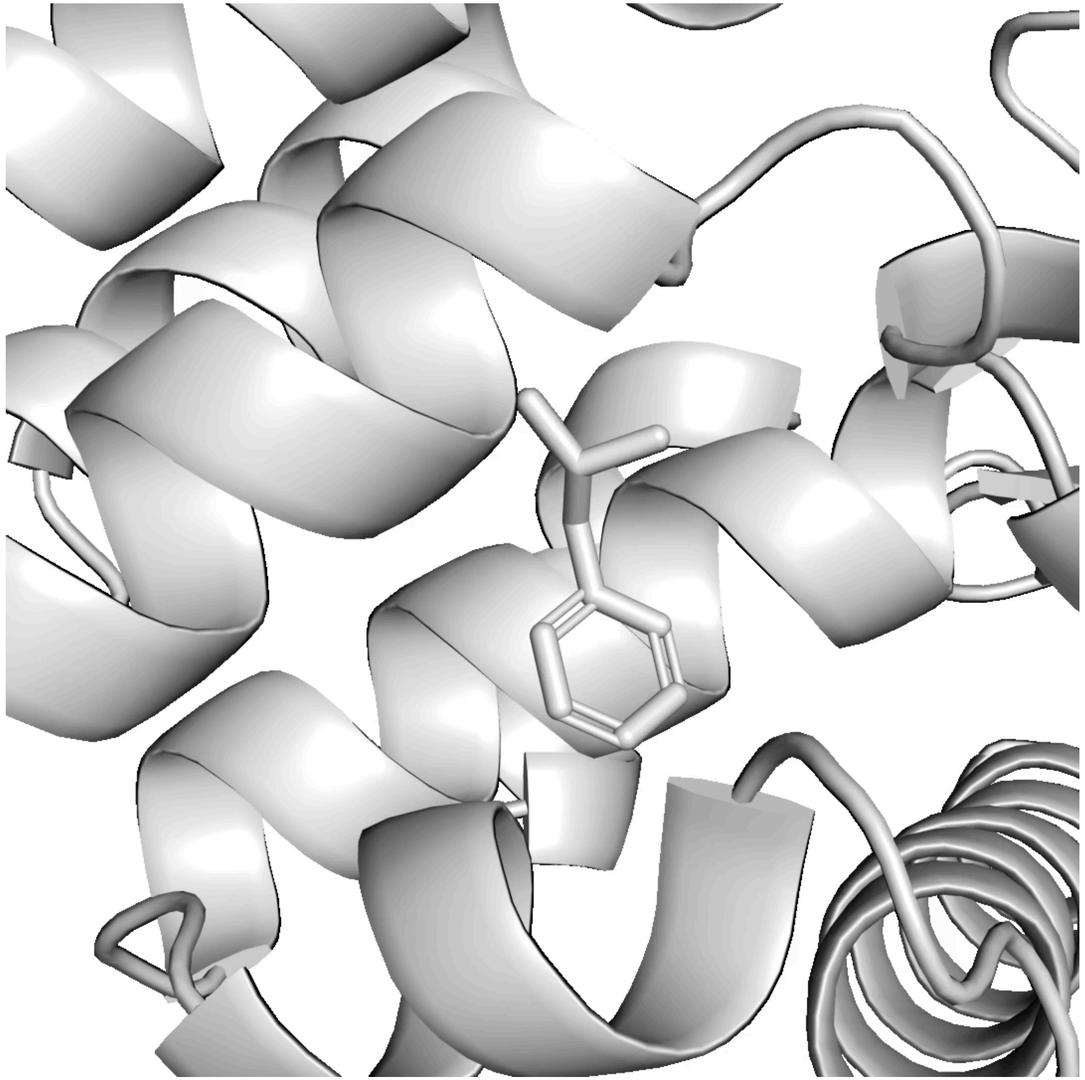


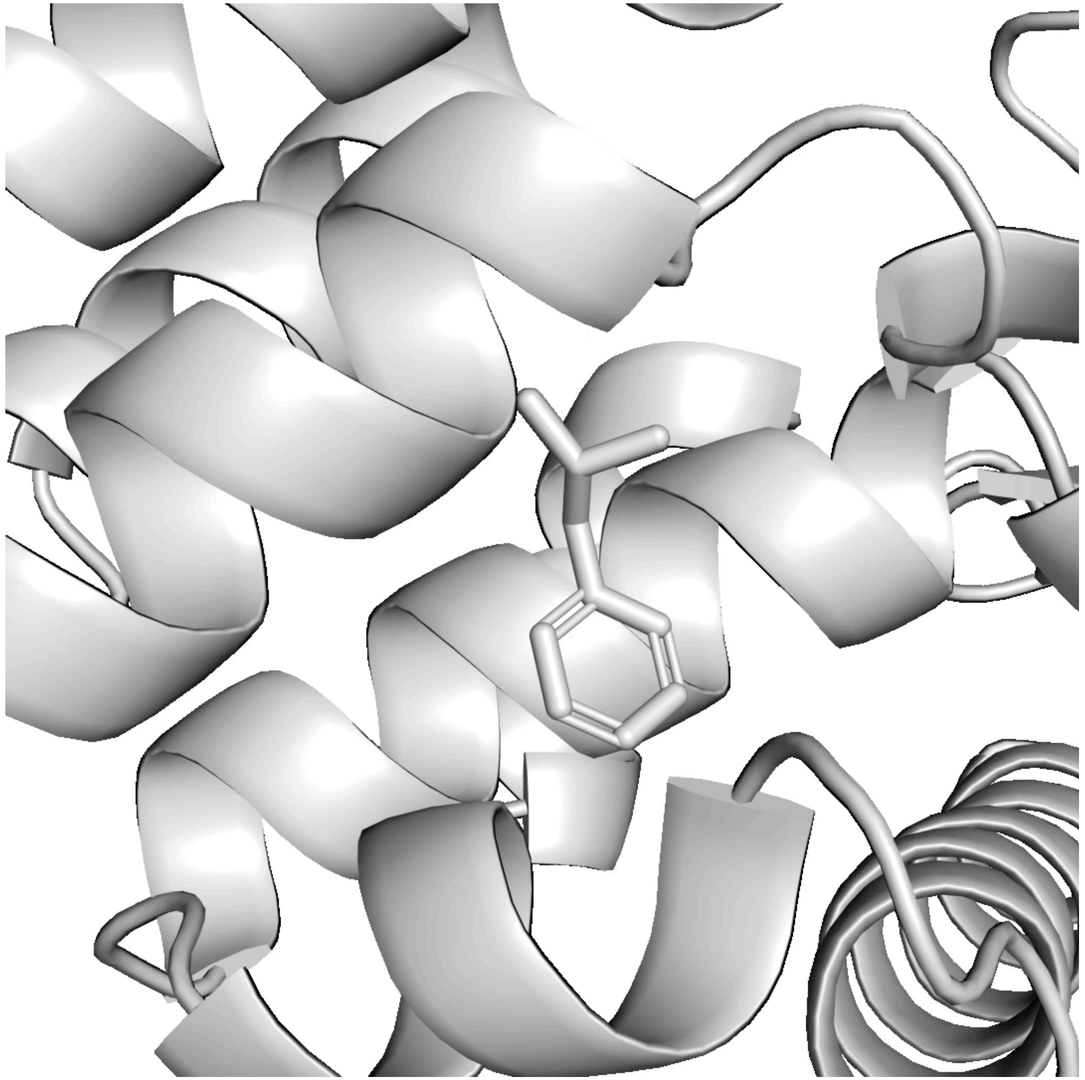
https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html



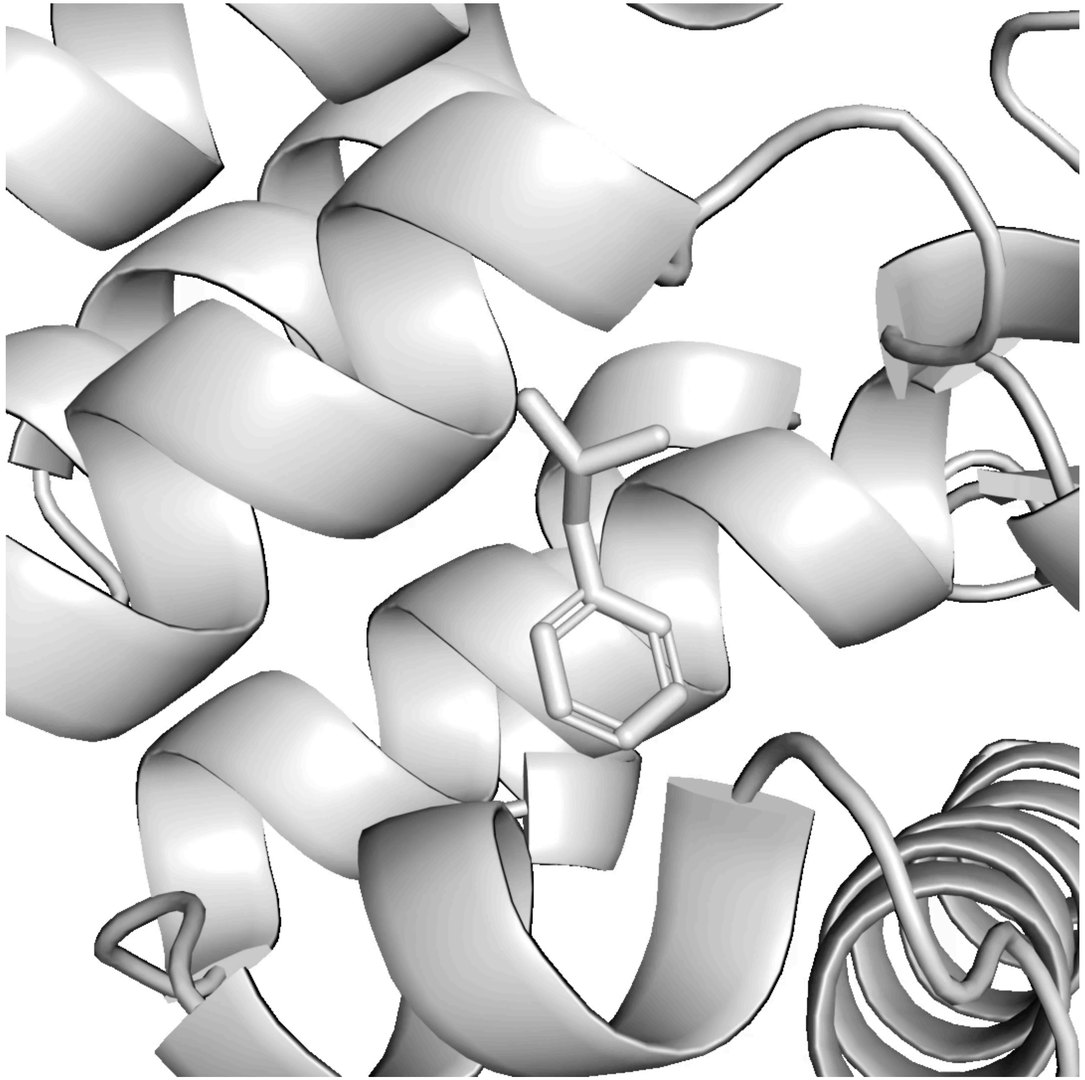
Deep Dreams of Molecules

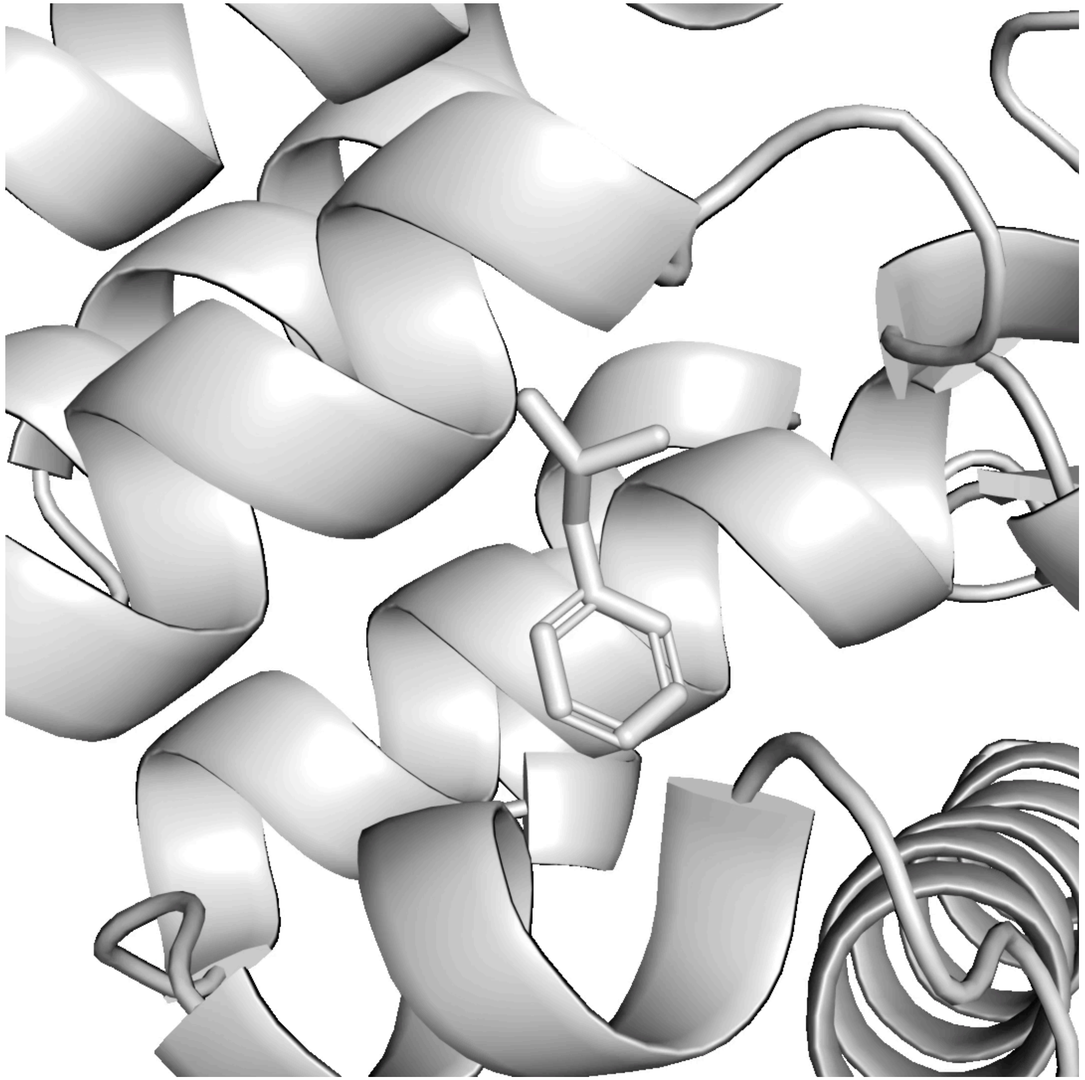




184l

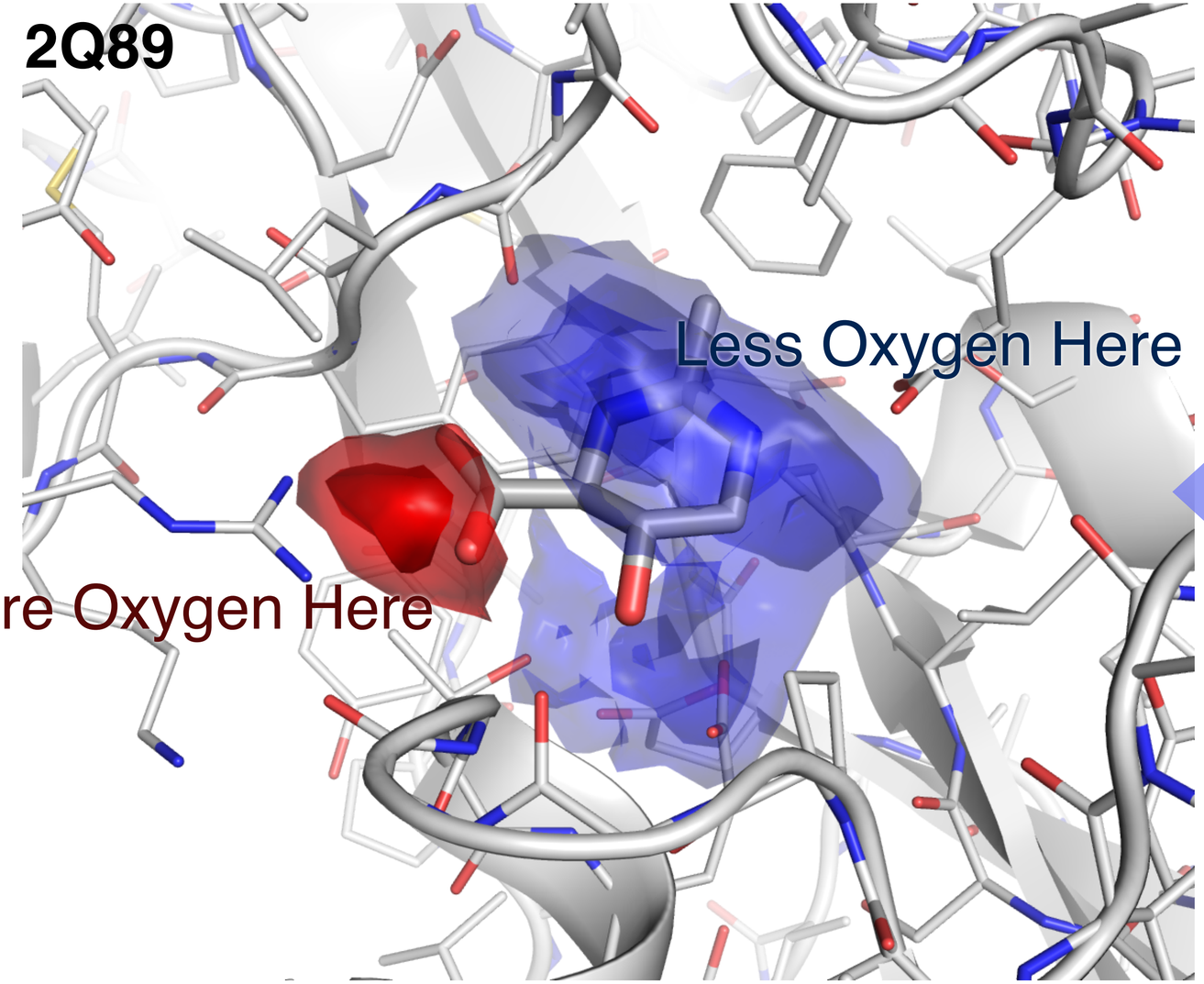
Deep Dreams of Molecules





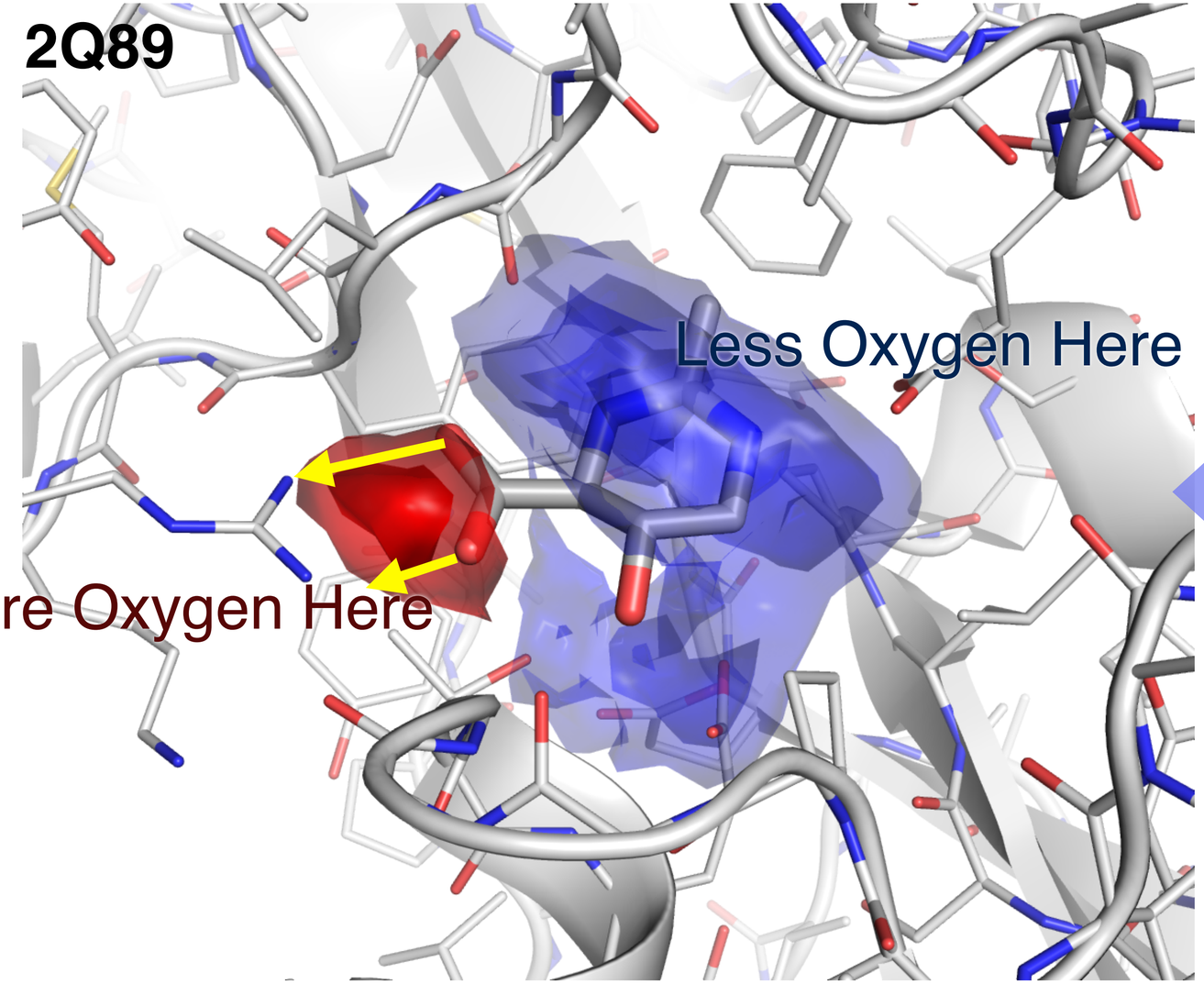
184l

Beyond Scoring

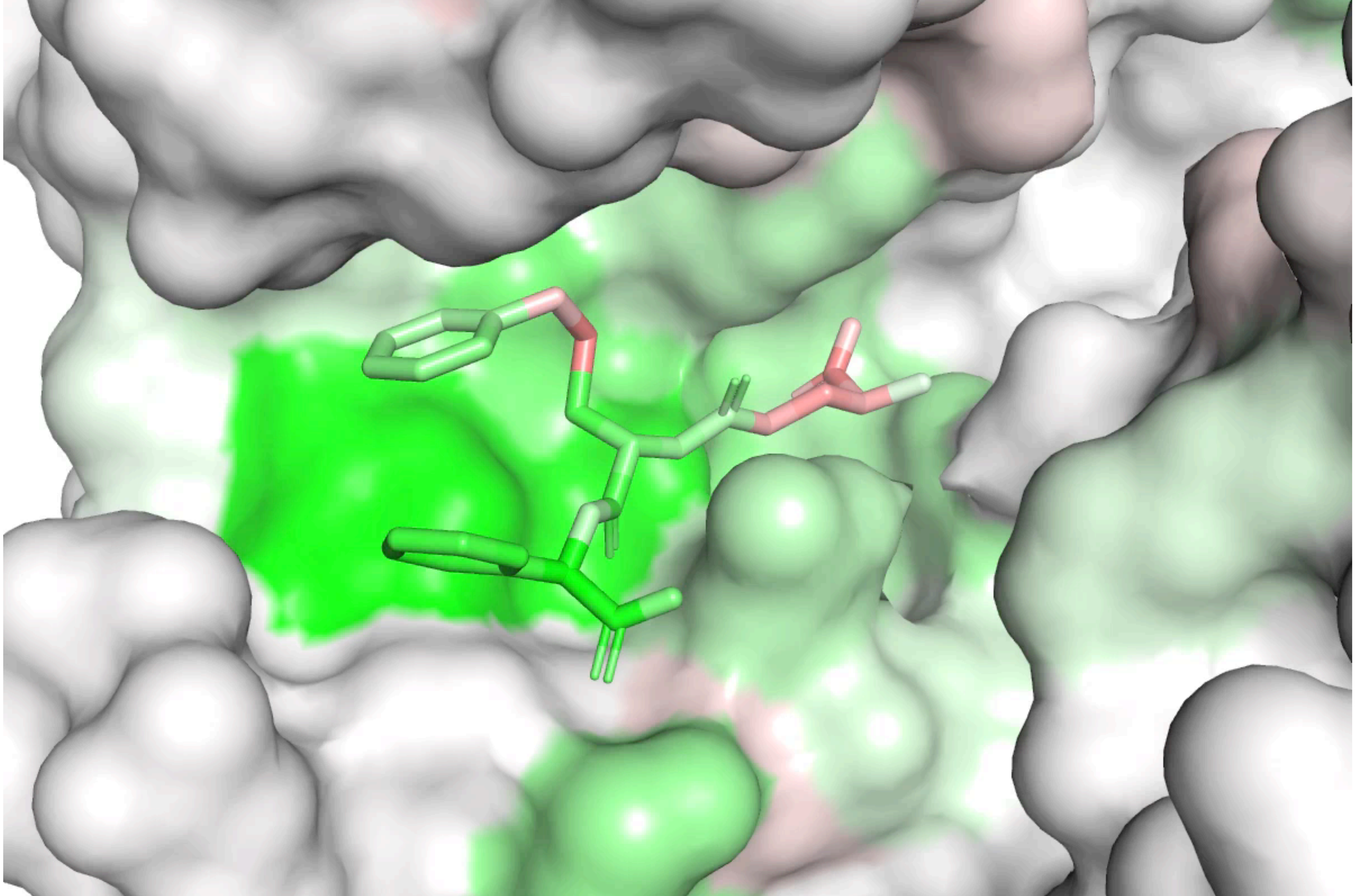


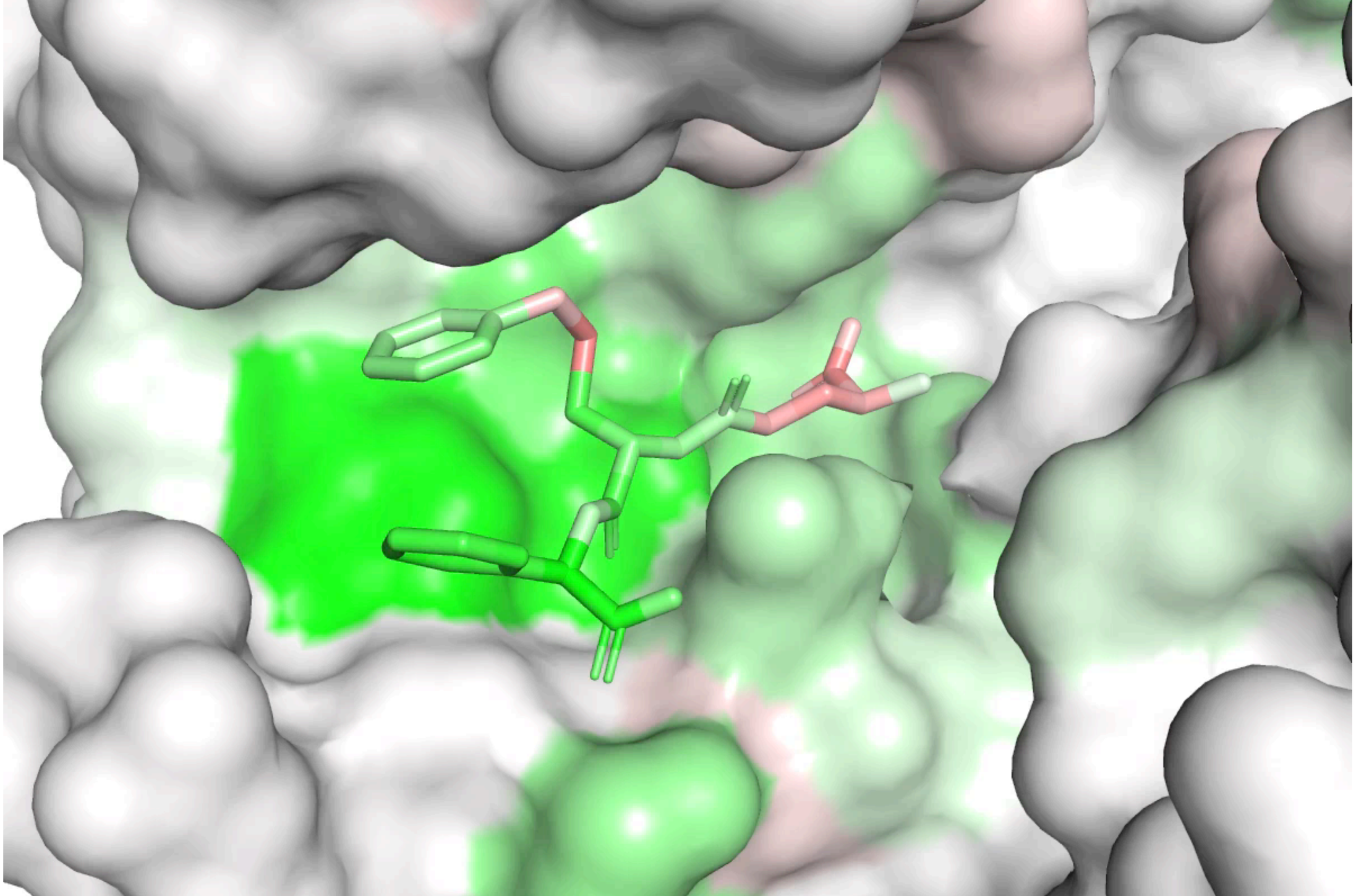
$$\frac{\partial L}{\partial A} = \sum_{i \in G_A} \frac{\partial L}{\partial G_i} \frac{\partial G_i}{\partial D} \frac{\partial D}{\partial A}$$

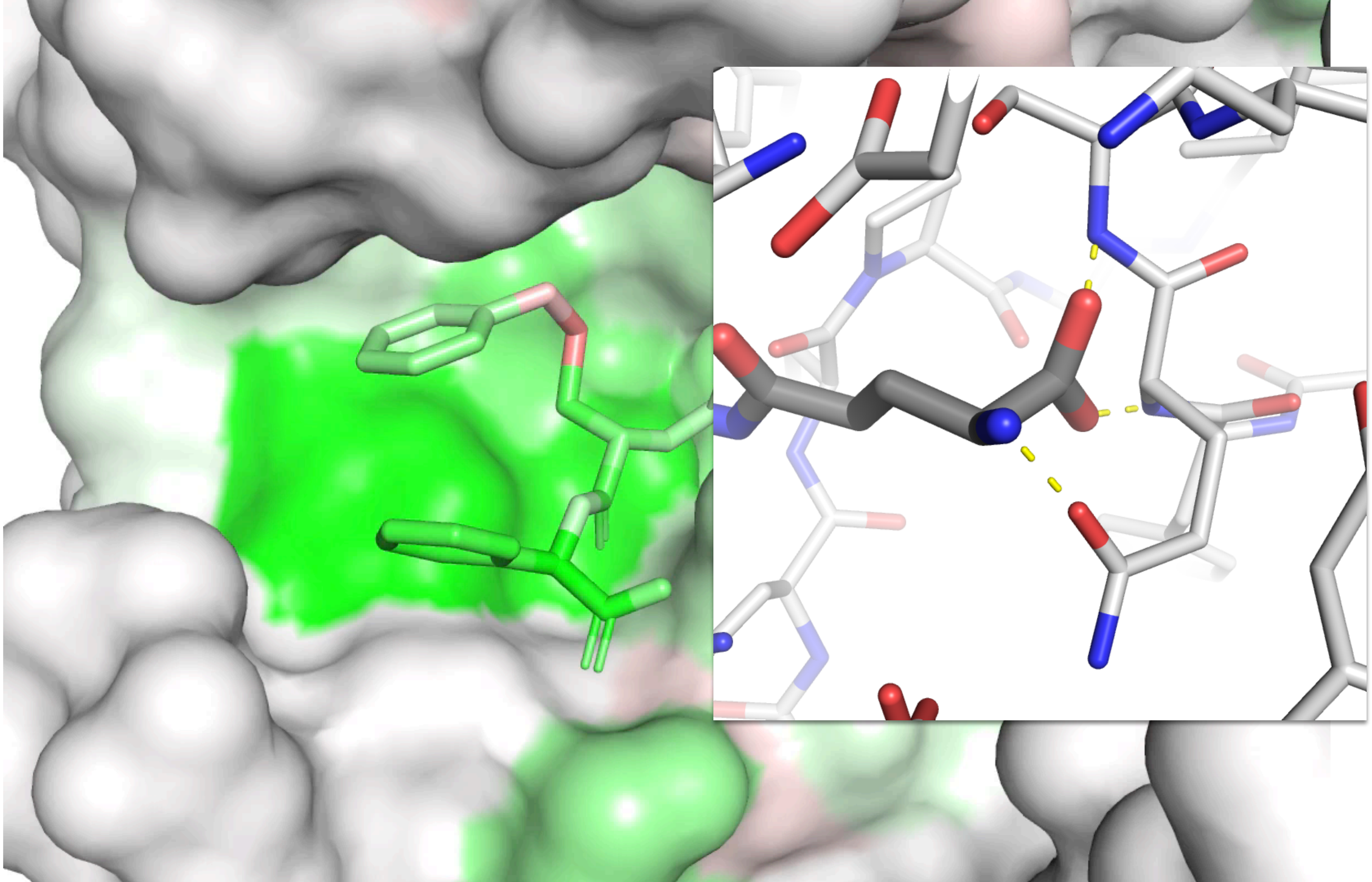
Beyond Scoring

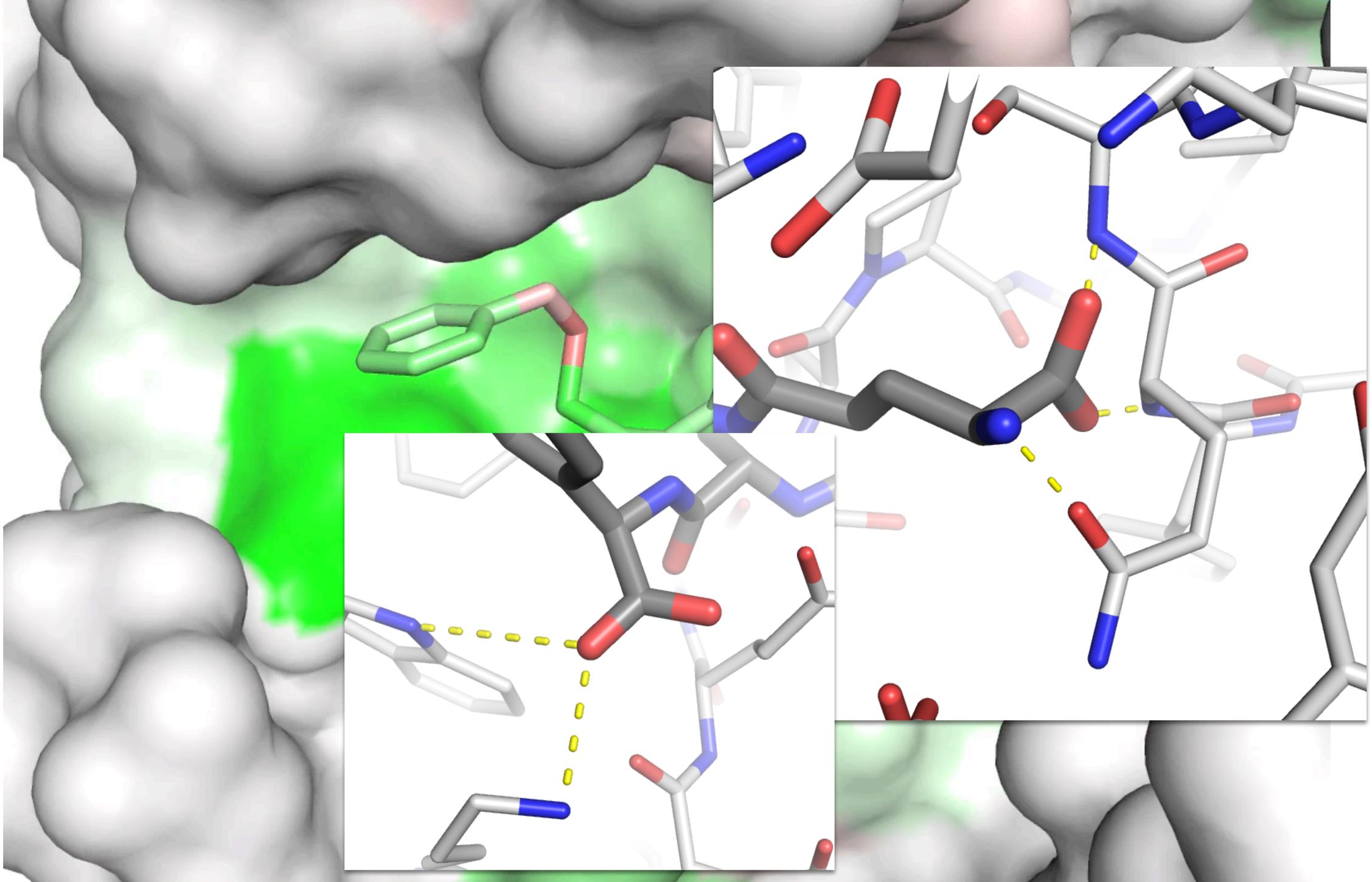


$$\frac{\partial L}{\partial A} = \sum_{i \in G_A} \frac{\partial L}{\partial G_i} \frac{\partial G_i}{\partial D} \frac{\partial D}{\partial A}$$

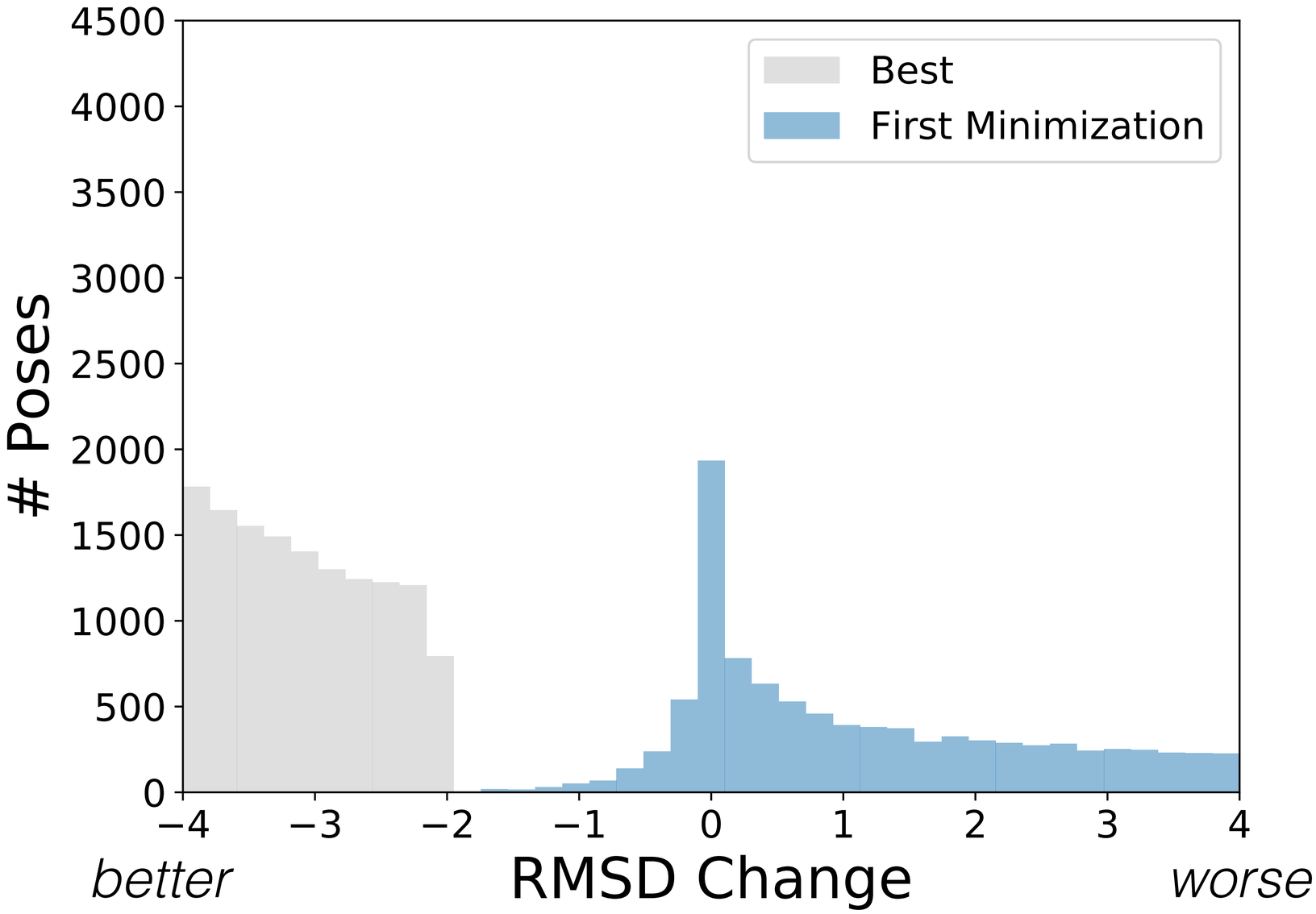




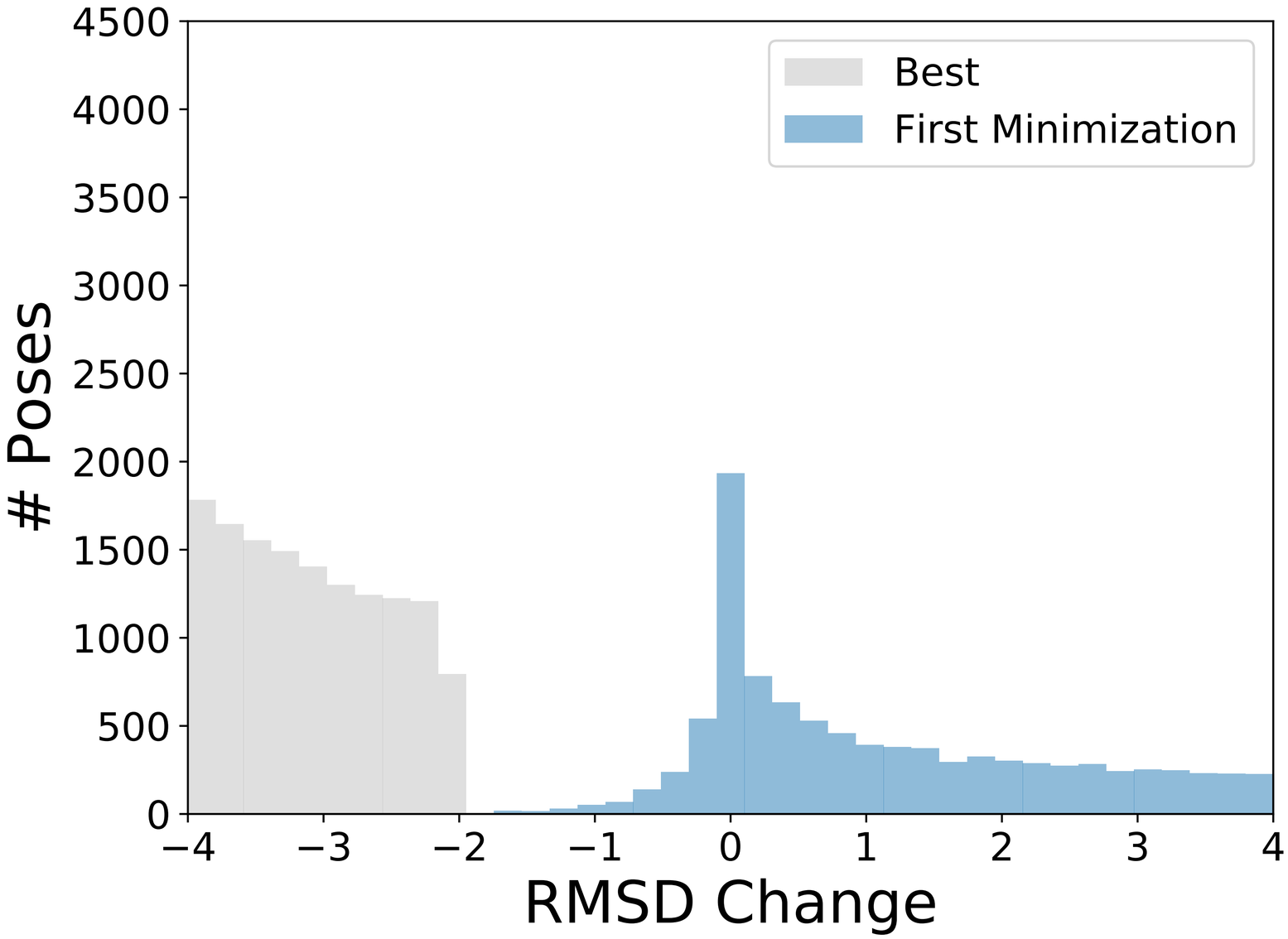




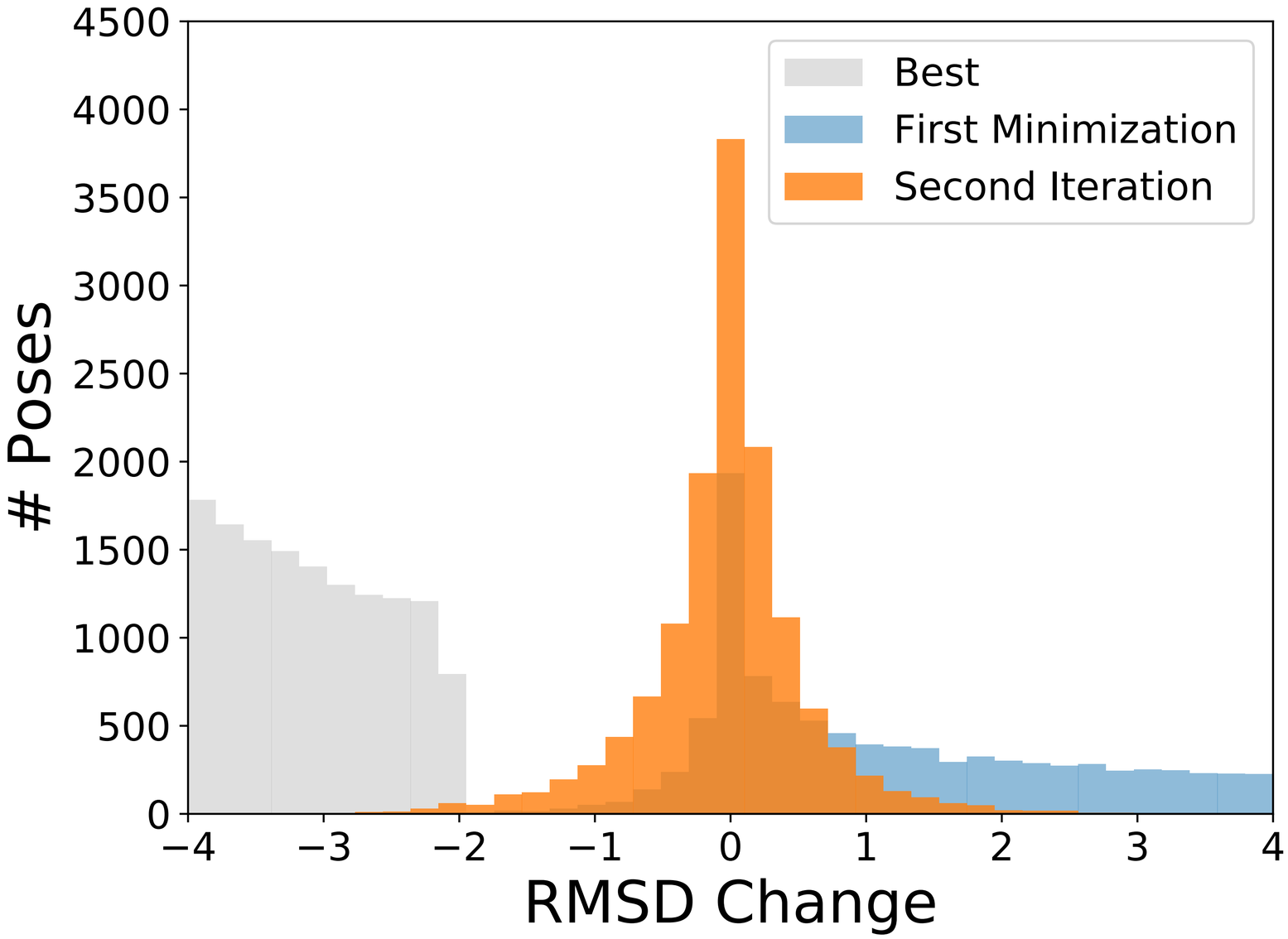
Minimizing Low RMSD Poses



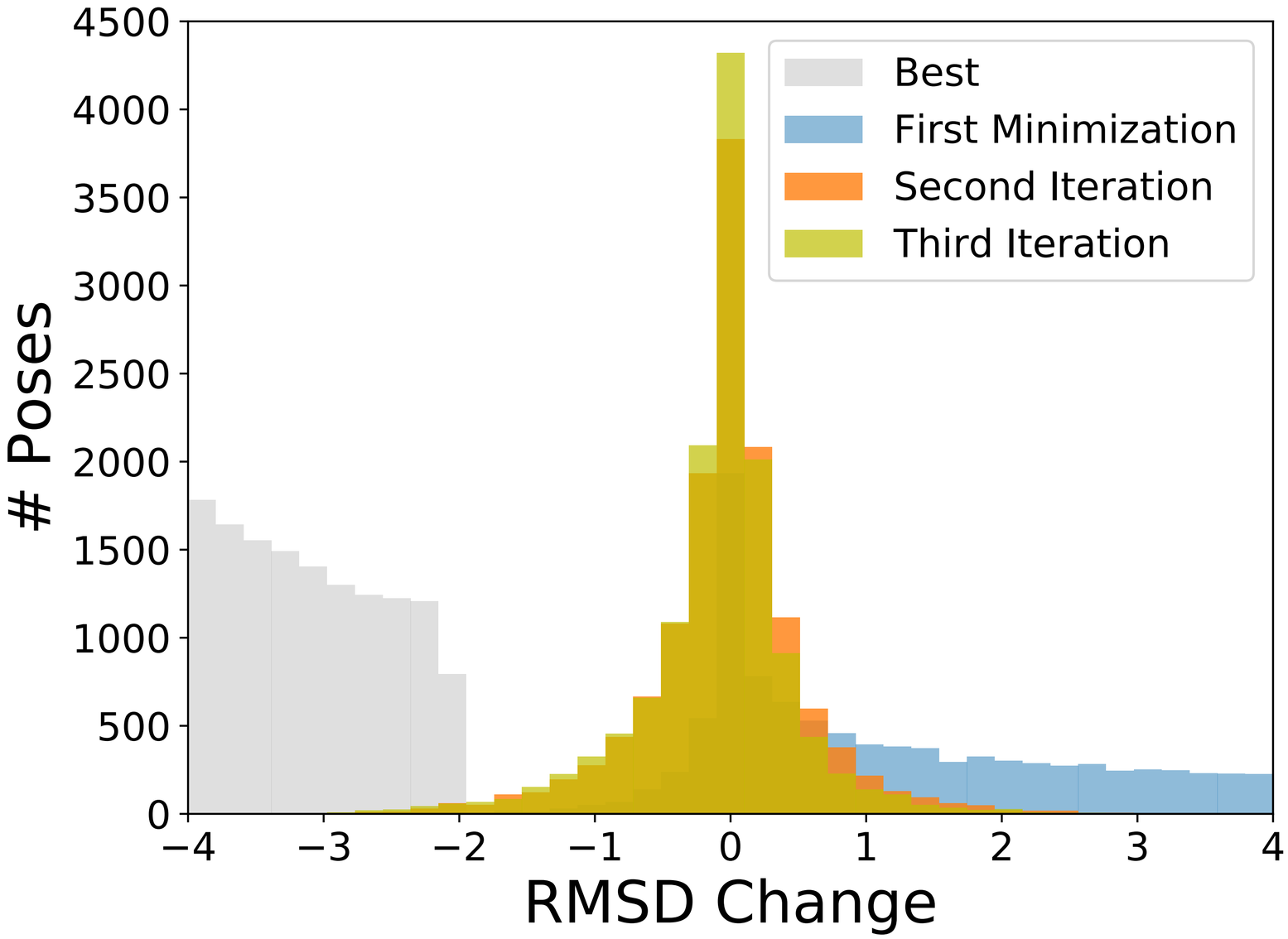
Iterative Refinement

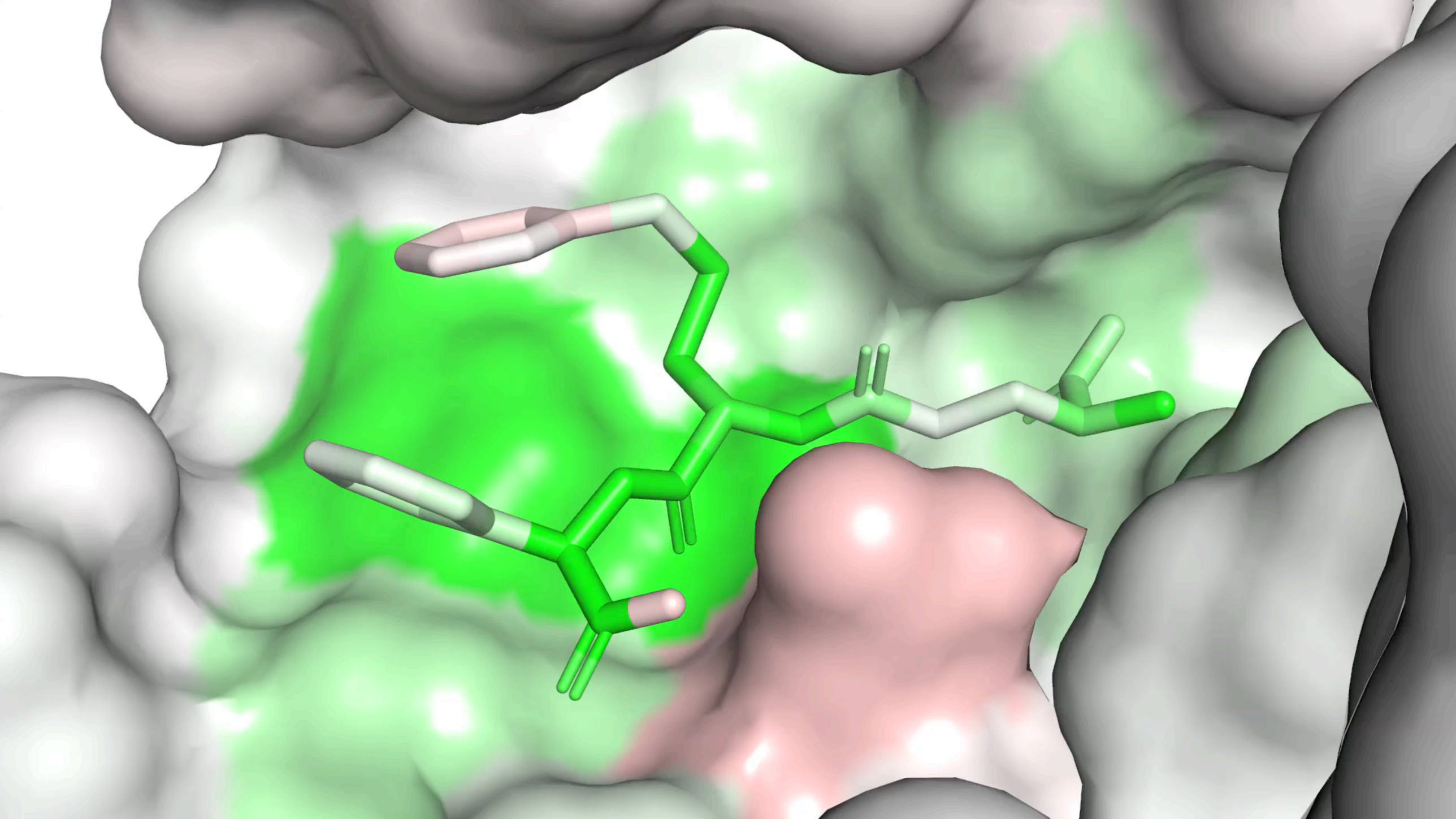


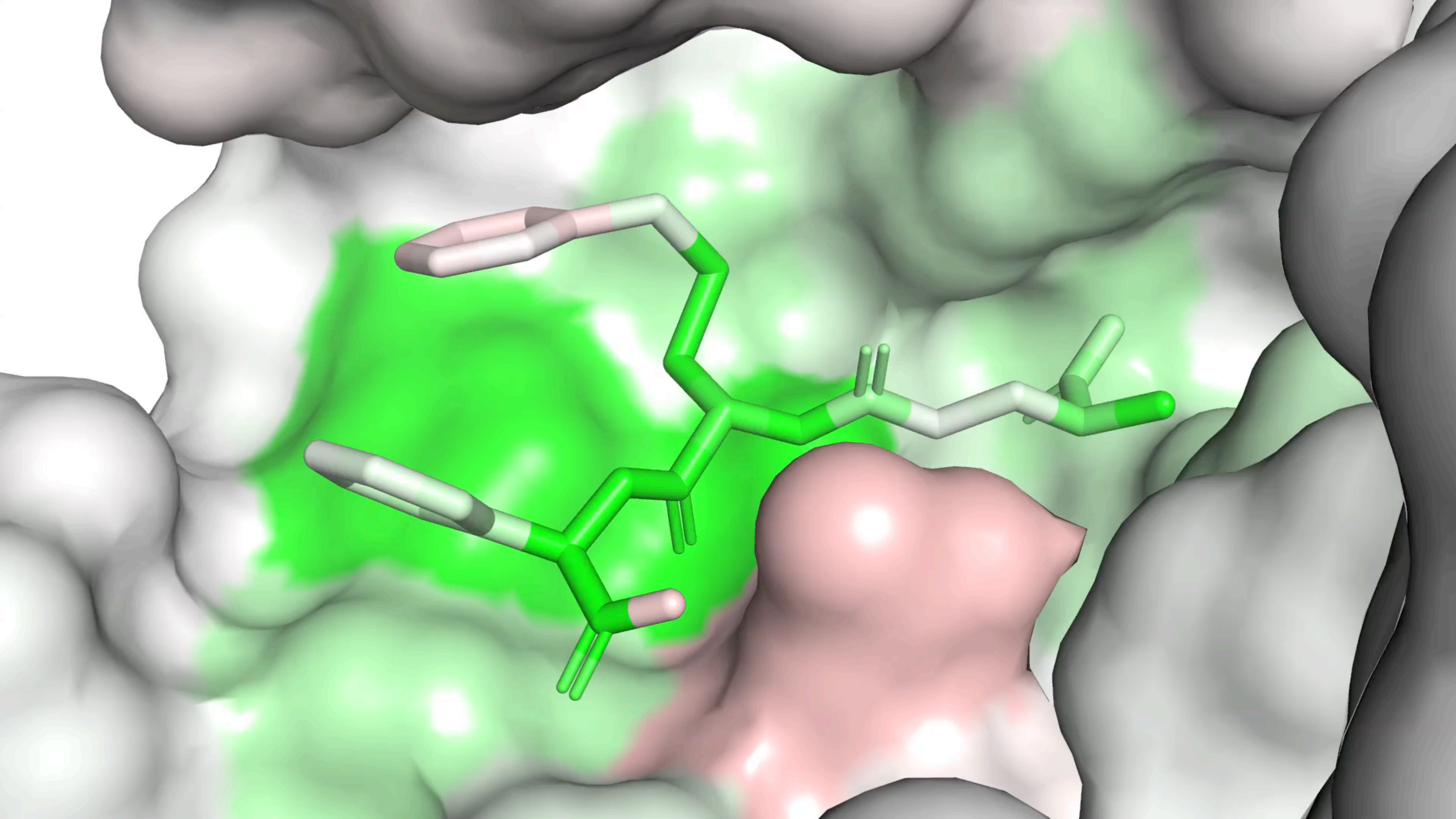
Iterative Refinement



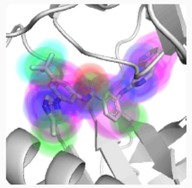
Iterative Refinement







Search or jump to... / Pull requests Issues Marketplace Explore



gnina

Repositories 4 People 9 Teams 1 Projects 0 Settings

Find a repository... Type: All Language: All

gnina

gnina is not smina/vina

cheminformatics computational-chemistry drug-discovery

C++ ★ 35 26 Updated 11 days ago

scripts

Jupyter Notebook ★ 1 13 BSD-3-Clause Updated on Sep 17, 2018

models

Trained caffe models

★ 6 9 Updated on Aug 23, 2018

Protein–Ligand Scoring with Convolutional Neural Networks

Matthew Ragoza†‡, Joshua Hochuli‡¶, Elisa Idrobo§, Jocelyn Sunseriǁ, and David Ryan Koes*ǁ
†Department of Neuroscience, ‡Department of Computer Science, ¶Department of Biological Sciences, and ǁDepartment of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States
§ Department of Computer Science, The College of New Jersey, Ewing, New Jersey 08628, United States

J. Chem. Inf. Model., **2017**, *57* (4), pp 942–957
DOI: 10.1021/acs.jcim.6b00740
Publication Date (Web): April 3, 2017

Cite this: *J. Chem. Inf. Model.* 2017, 57, 4, 942-957
RIS Citation GO

arXiv.org > stat > arXiv:1710.07400

Statistics > Machine Learning

Ligand Pose Optimization with Atomic Grid-Based Convolutional Neural Networks

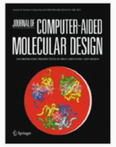
Matthew Ragoza, Lillian Turner, David Ryan Koes

(Submitted on 20 Oct 2017)

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



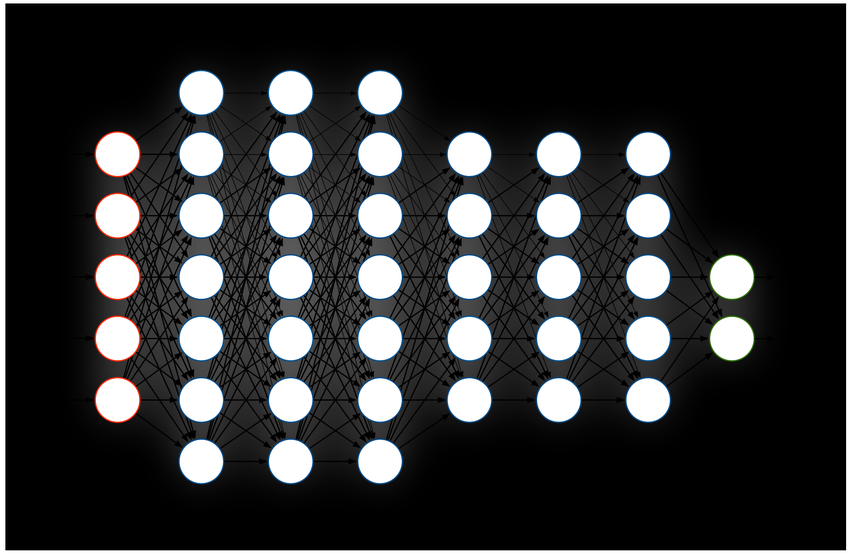
Journal of Computer-Aided Molecular Design
pp 1–16 | Cite as

Convolutional neural network scoring and minimization in the D3R 2017 community challenge

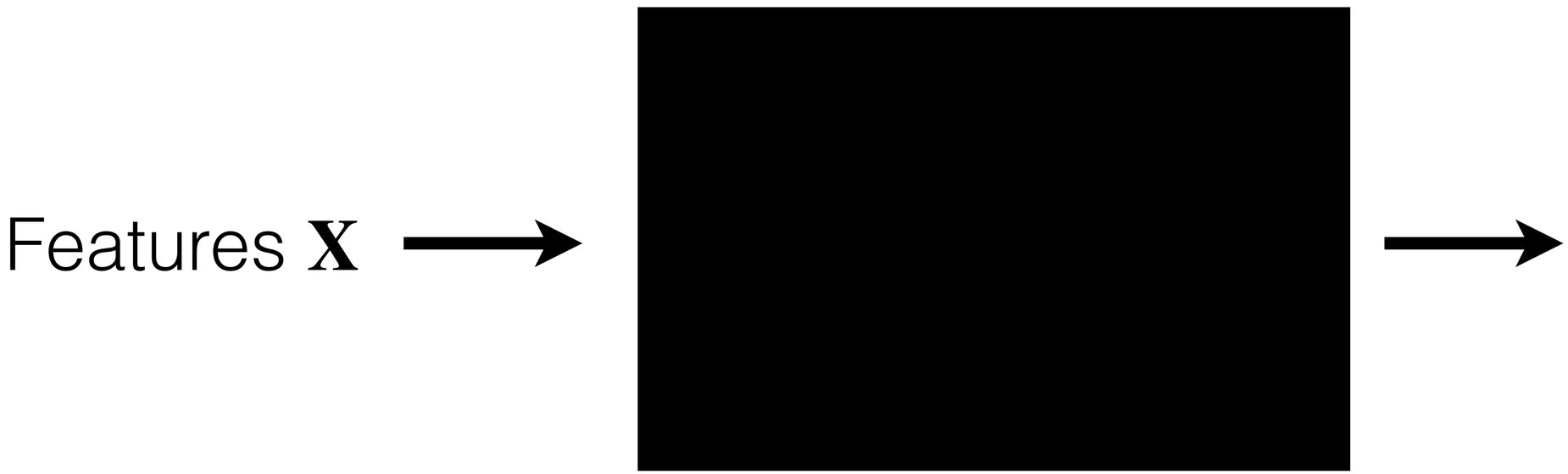
Authors Authors and affiliations
Jocelyn Sunseri, Jonathan E. King, Paul G. Francoeur, David Ryan Koes

Generative Modeling

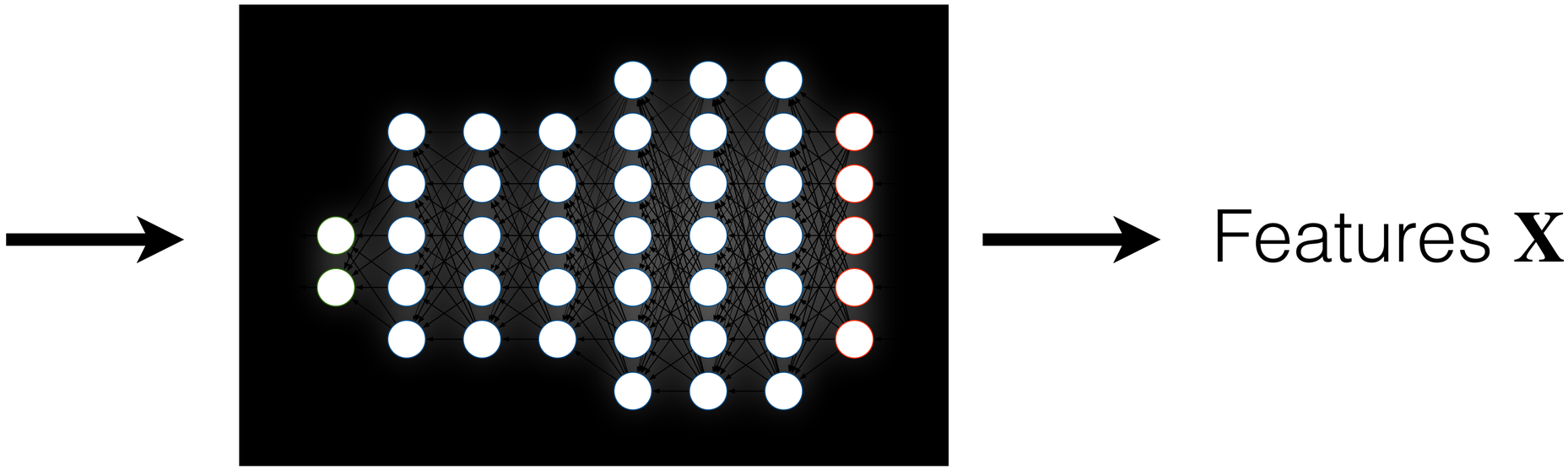
Discriminative Model

Features $\mathbf{X}$  Prediction $\mathbf{y}$

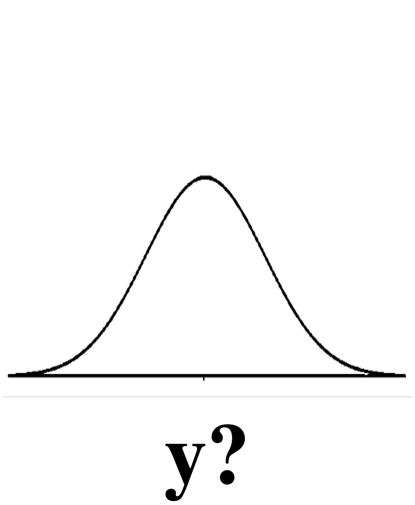
Generative Model

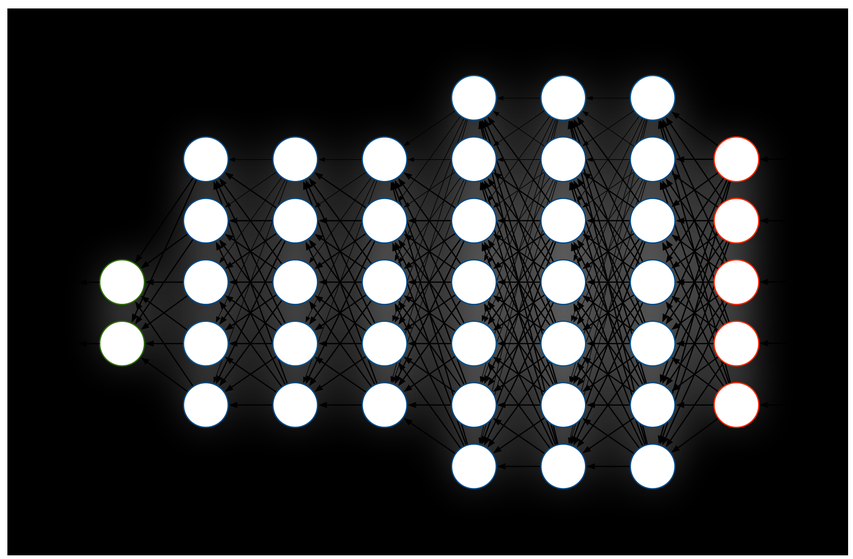


Generative Model



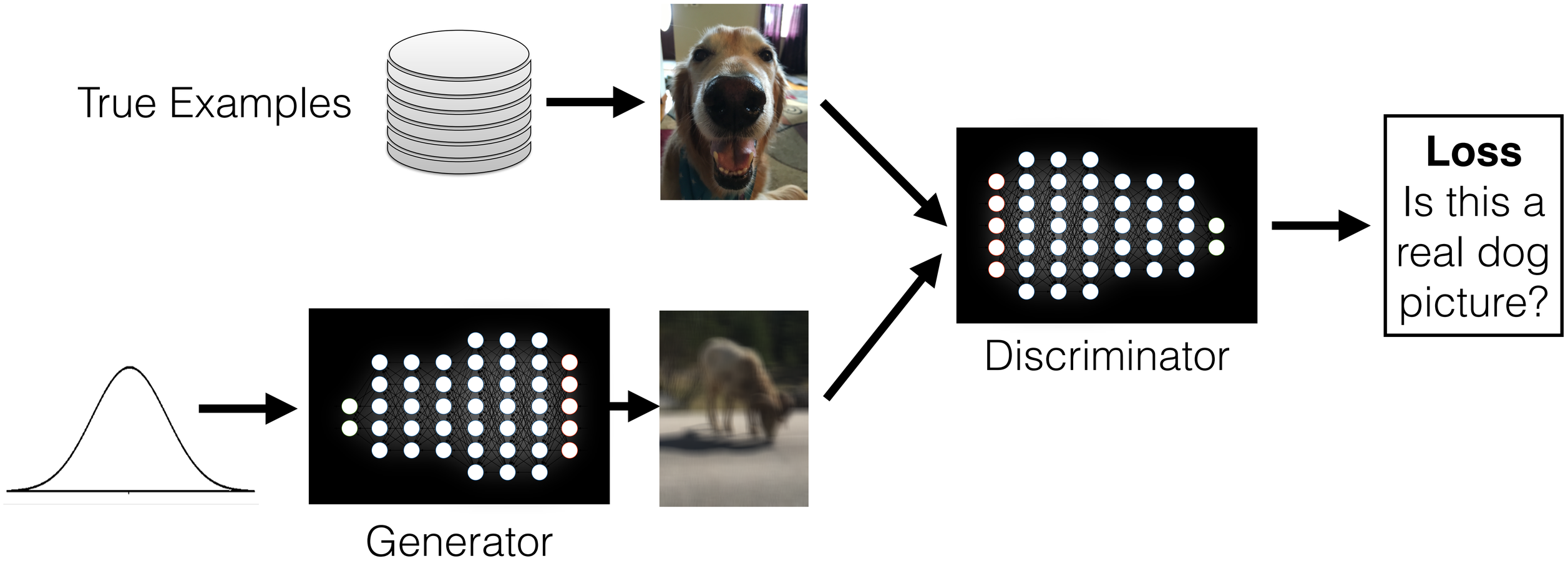
Generative Model





Features $\mathbf{X}$

Generative Adversarial Networks



Generative Adversarial Networks

2014

2015

2016

2017

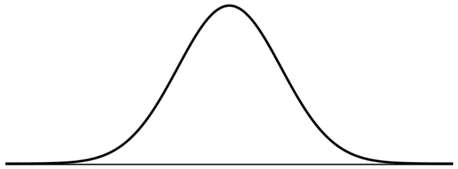
2018

Ian Goodfellow @goodfellow_ian · 2h
4.5 years of GAN progress on face generation. arxiv.org/abs/1406.2661 arxiv.org/abs/1511.06434 arxiv.org/abs/1606.07536 arxiv.org/abs/1710.10196 arxiv.org/abs/1812.04948

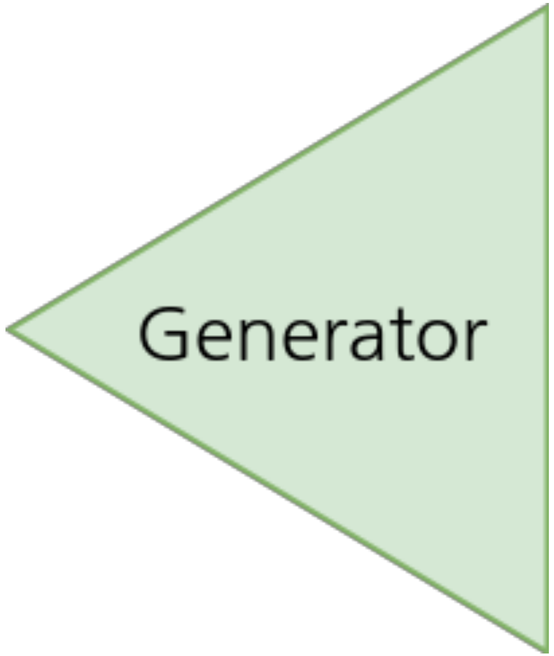
https://thispersondoesnotexist.com

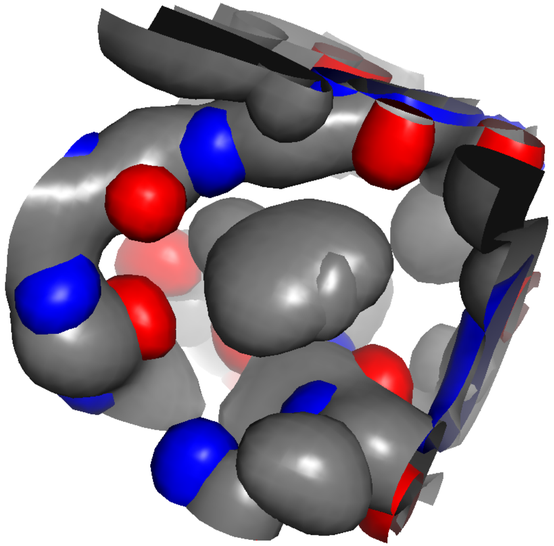
Generative Models

Generative models approximate a data distribution directly. They can map samples from one distribution (noise or input data) to realistic samples from an output distribution of interest.



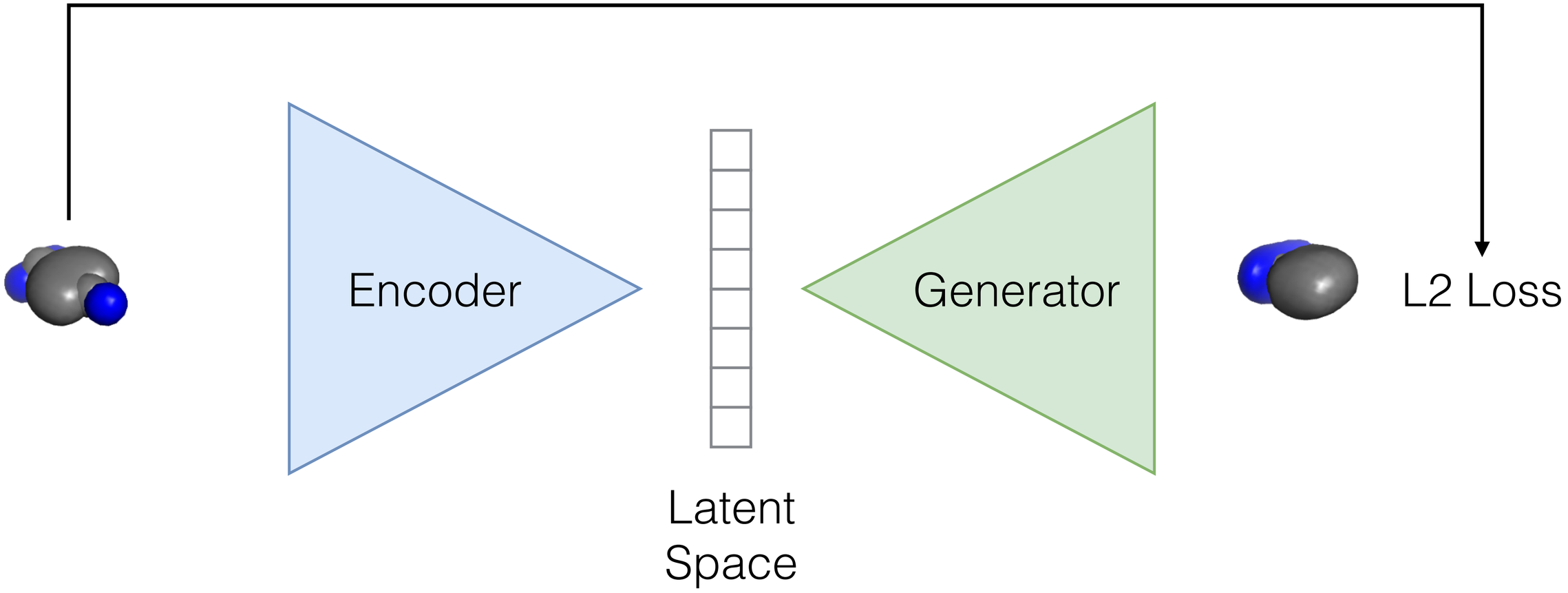
noise sample





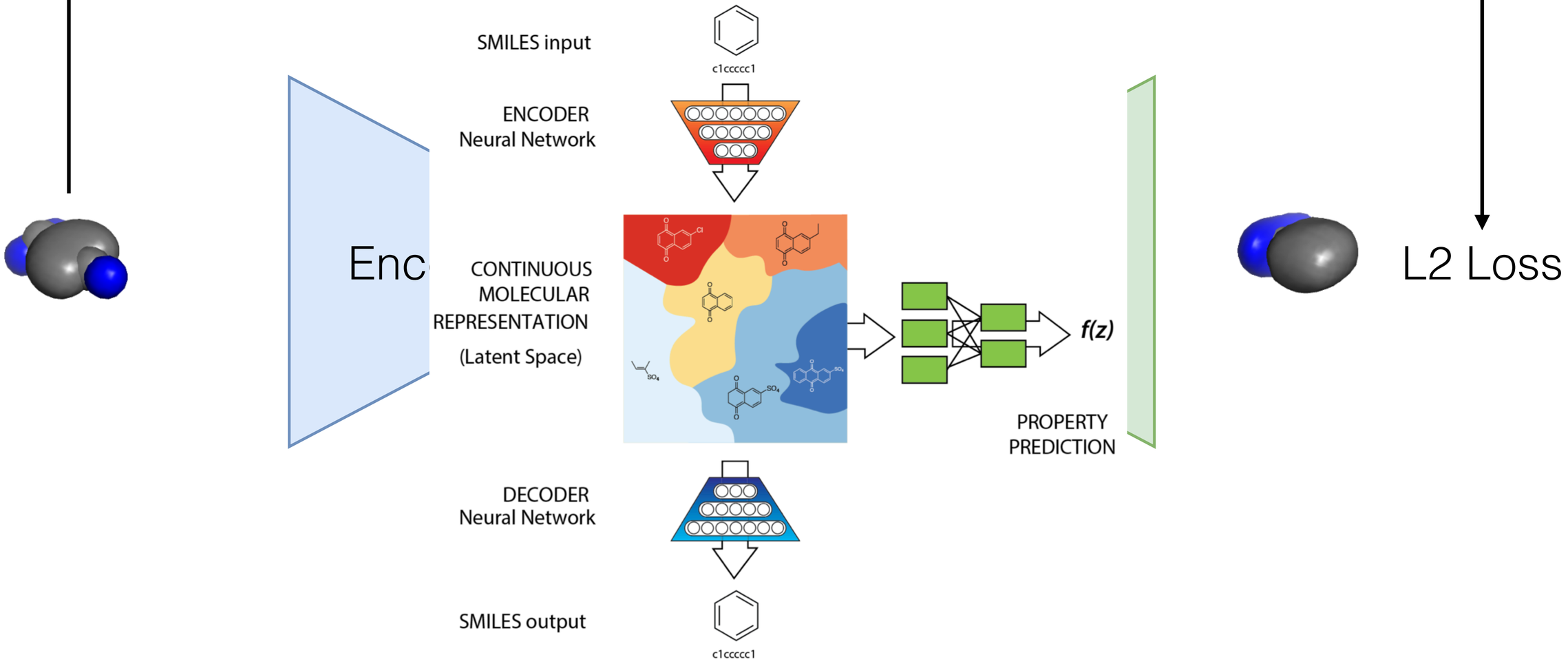
generated receptor & ligand grid

Autoencoding

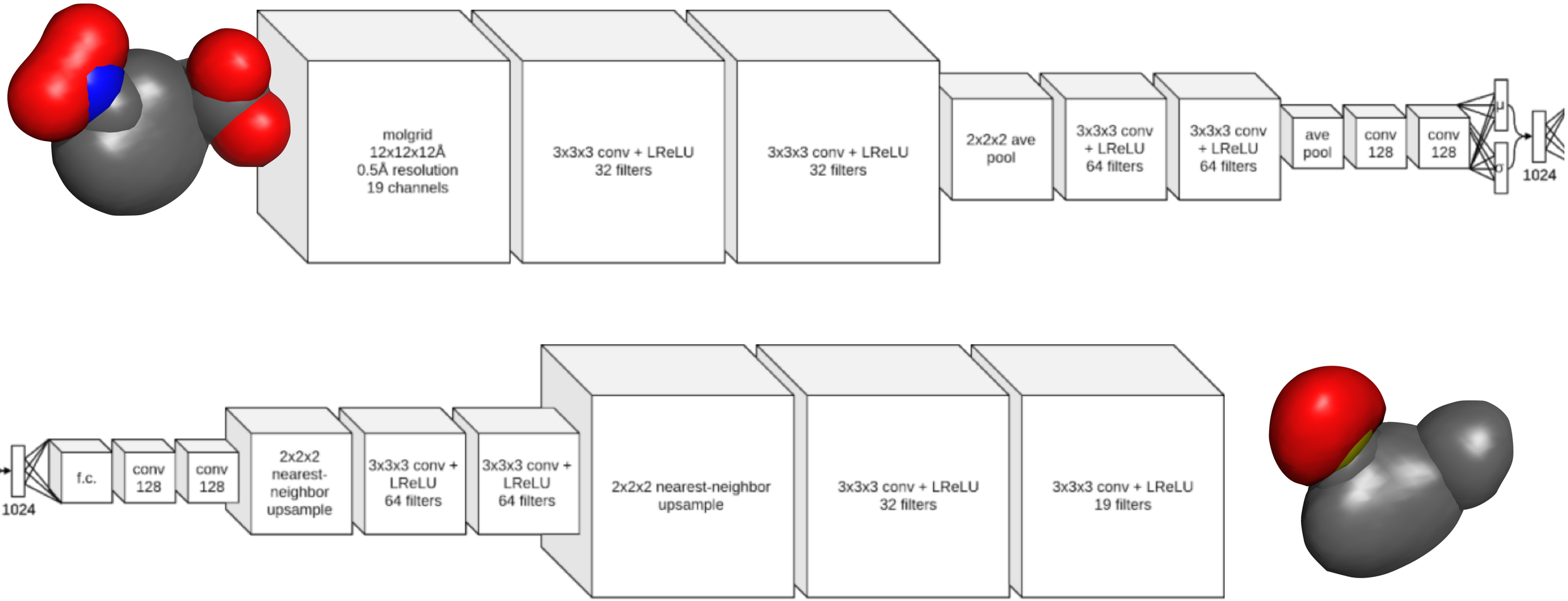


Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules

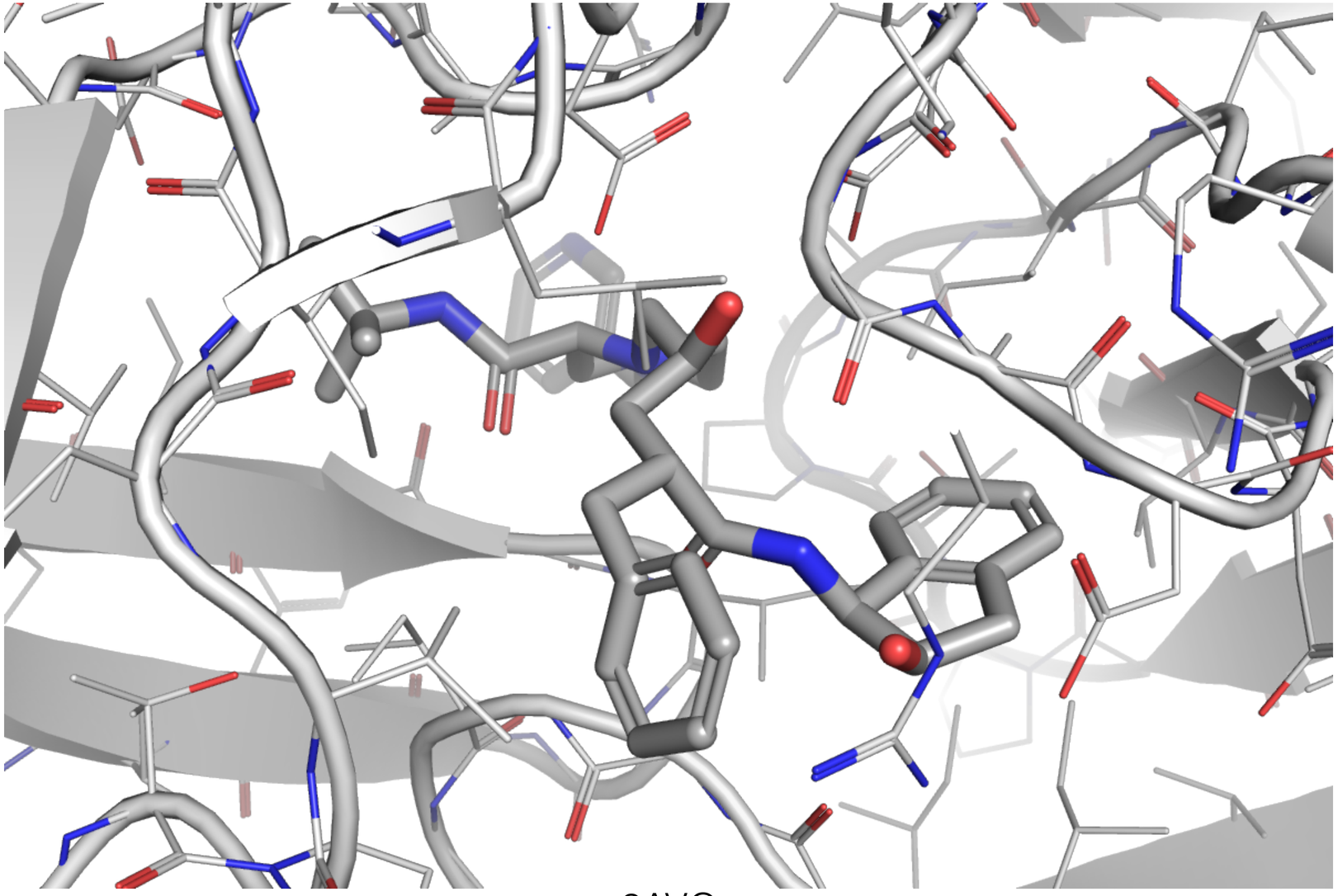
Rafael Gómez-Bombarelli[†#], Jennifer N. Wei[‡#], David Duvenaud[¶#], José Miguel Hernández-Lobato[§#], Benjamín Sánchez-Lengeling[‡], Dennis Sheberla[‡], Jorge Aguilera-Iparraguirre[†], Timothy D. Hirzel[†], Ryan P. Adams[∇⊥], and Alán Aspuru-Guzik[*‡⊥]



Model

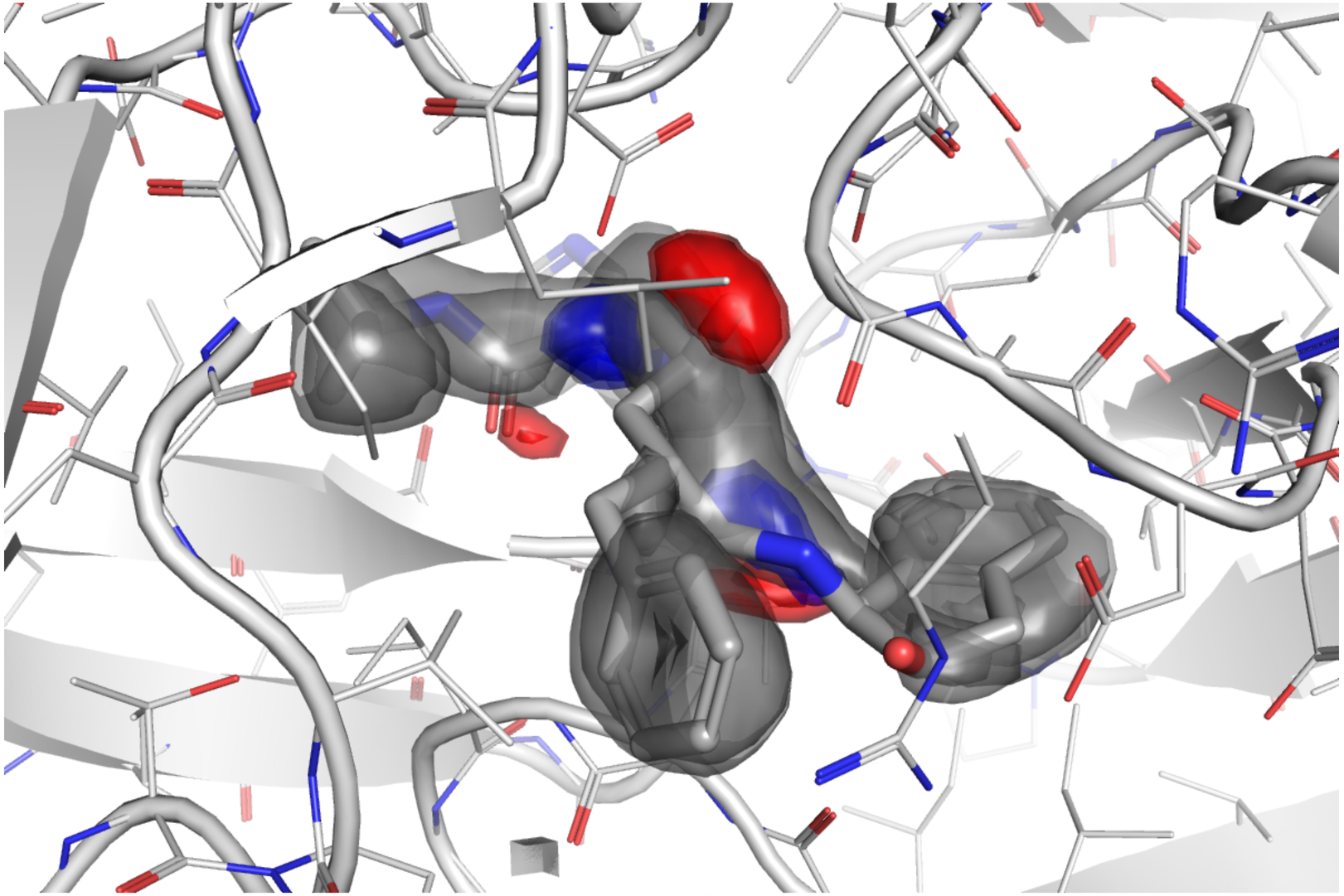

molgrid
12x12x12Å
0.5Å resolution
19 channels
3x3x3 conv + LReLU
32 filters
3x3x3 conv + LReLU
32 filters
2x2x2 ave
pool
3x3x3 conv
+ LReLU
64 filters
3x3x3 conv
+ LReLU
64 filters
ave
pool
conv
128
conv
128
μ
σ
1024
1024
f.c.
conv
128
conv
128
2x2x2
nearest-
neighbor
upsample
3x3x3 conv +
LReLU
64 filters
3x3x3 conv +
LReLU
64 filters
2x2x2 nearest-neighbor
upsample
3x3x3 conv + LReLU
32 filters
3x3x3 conv + LReLU
19 filters

Variational Autoencoding Examples



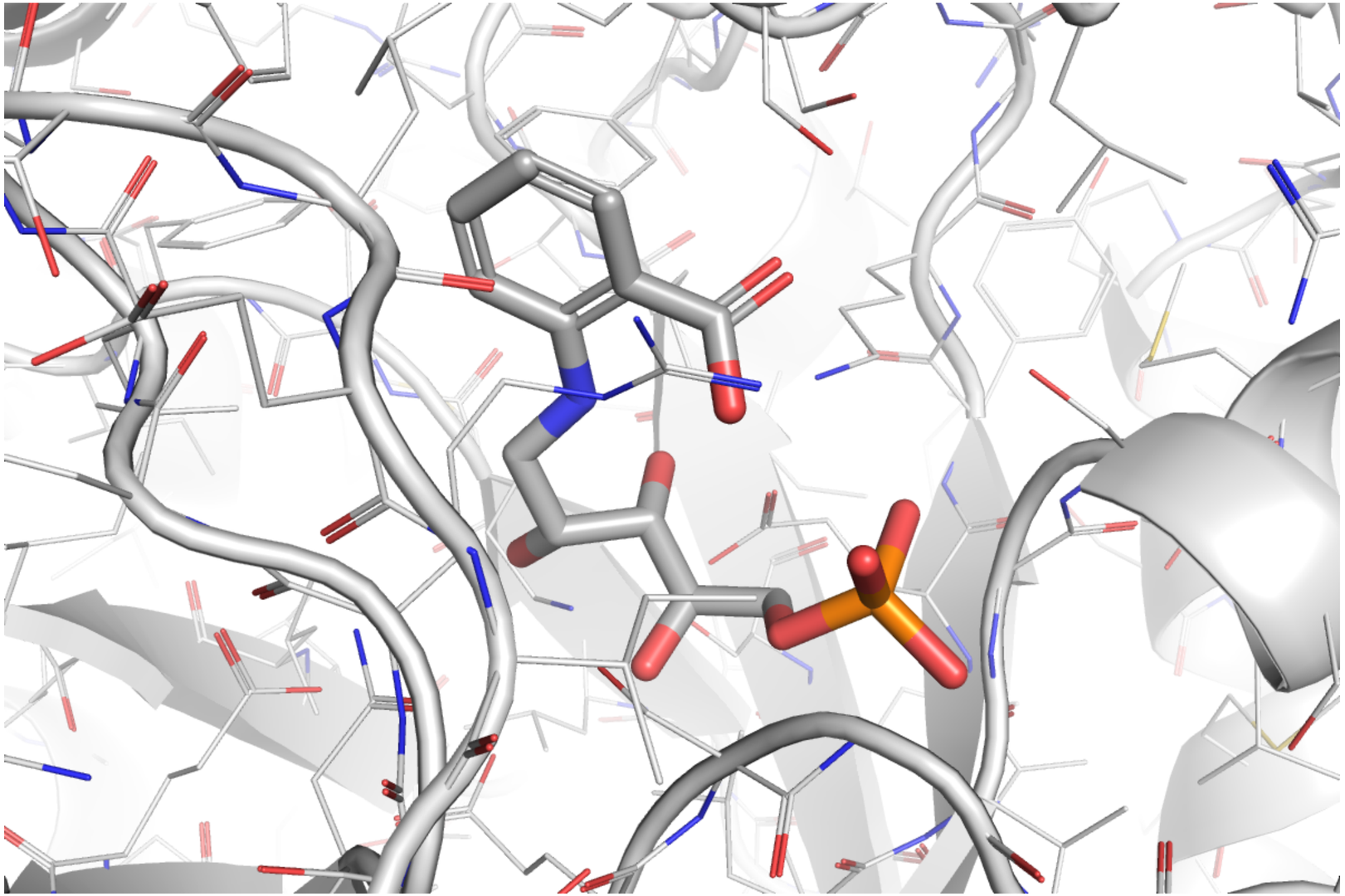
2AVO

Variational Autoencoding Examples



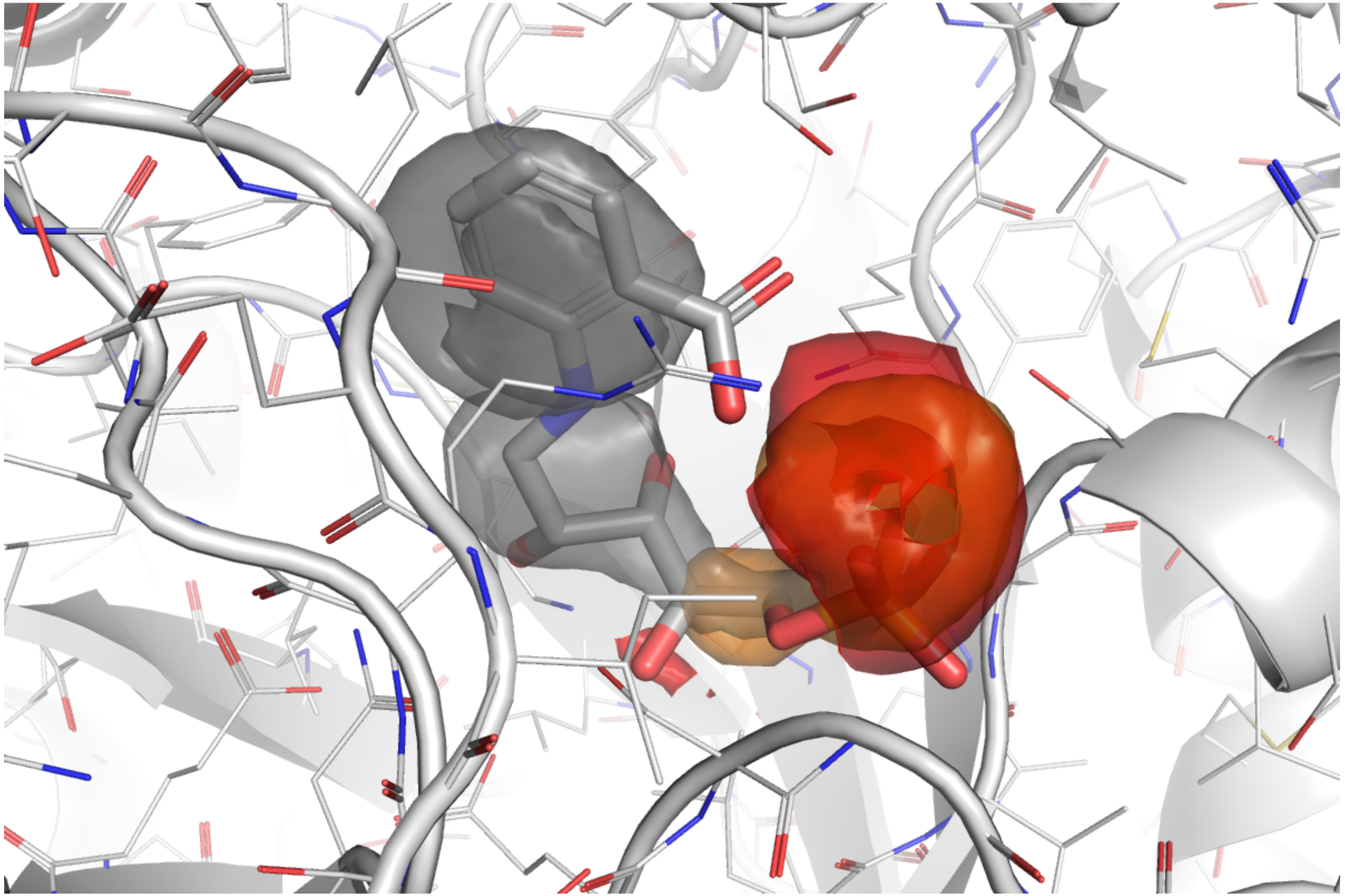
2AVO

Variational Autoencoding Examples



1LBF

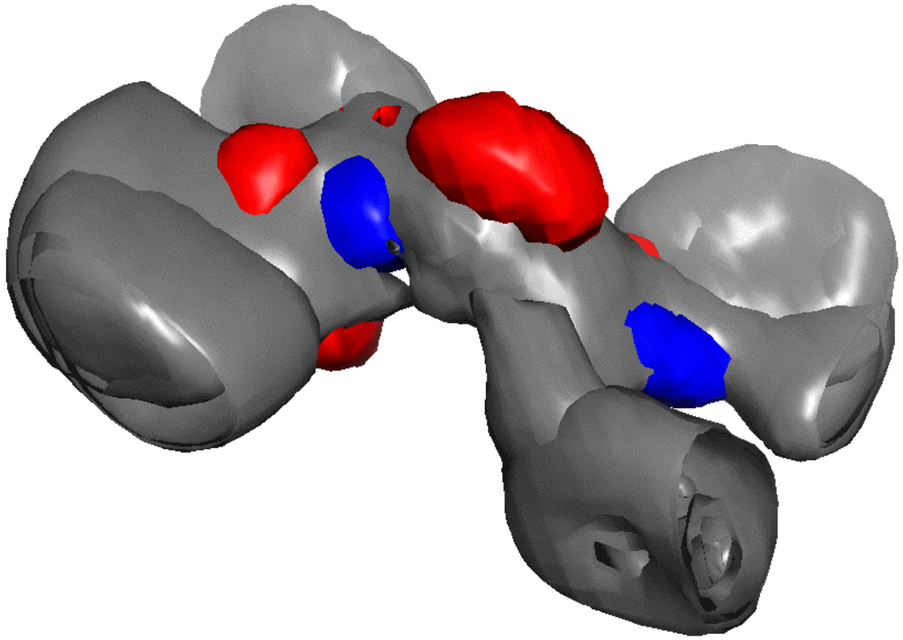
Variational Autoencoding Examples



1LBF

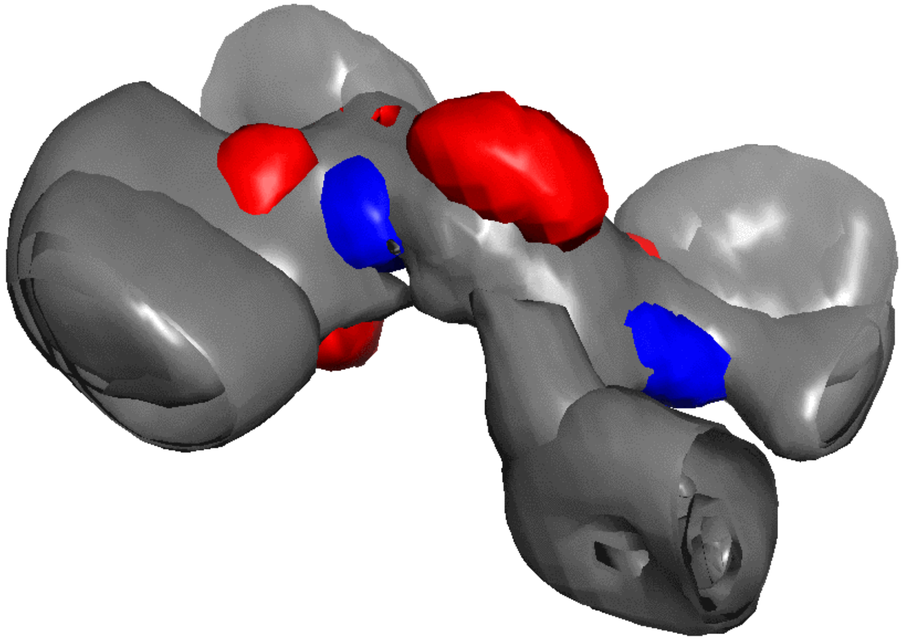
Atom Fitting

$$a^* = \underset{a}{\operatorname{argmin}} \|d - D(a)\|_2^2 + \lambda E(a)$$

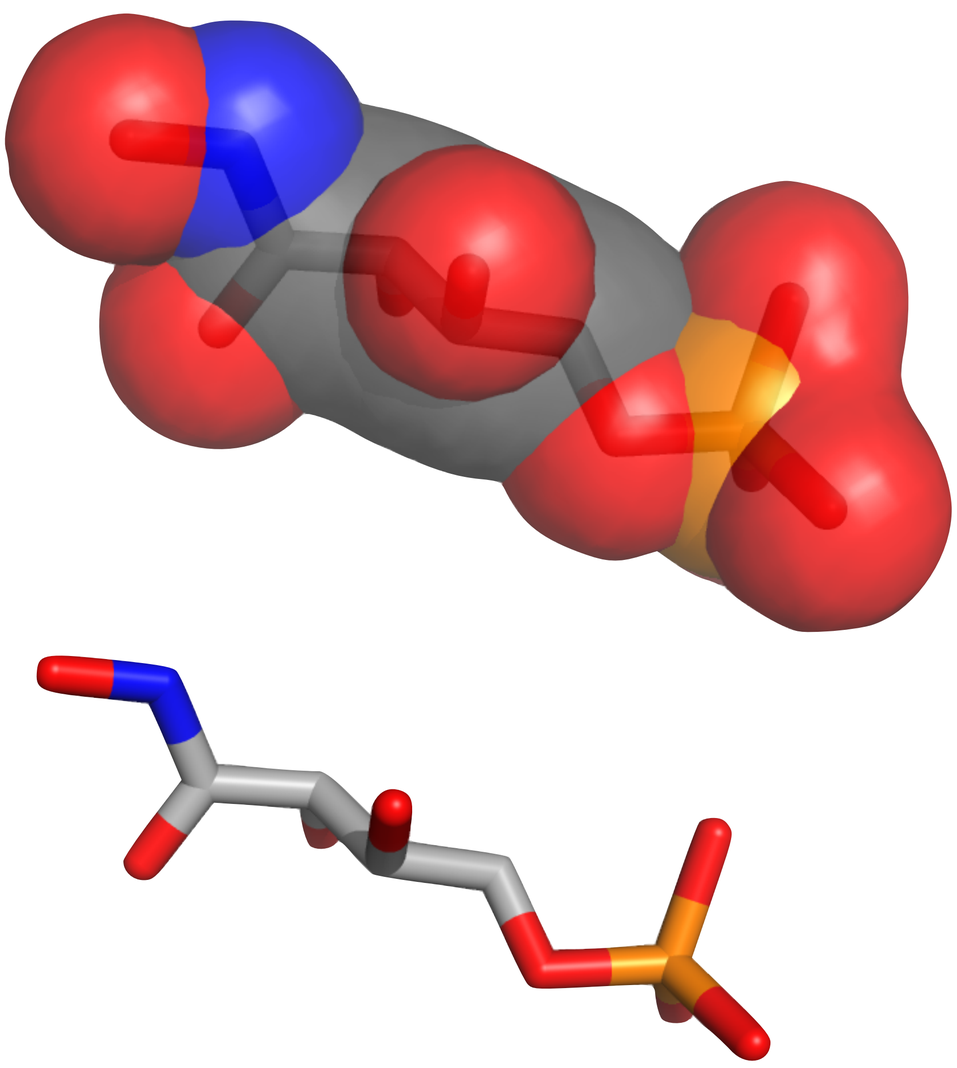


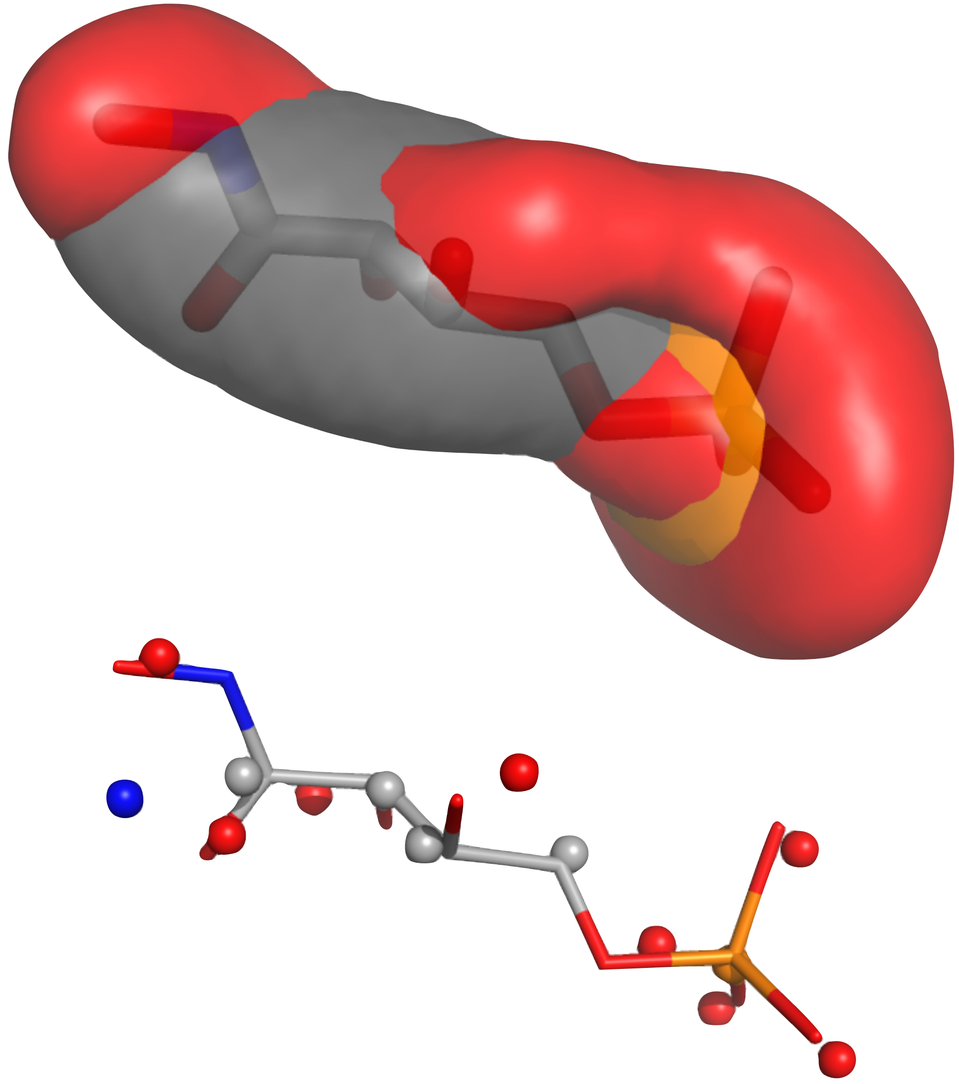
Atom Fitting

$$a^* = \underset{a}{\operatorname{argmin}} \|d - D(a)\|_2^2 + \lambda E(a)$$



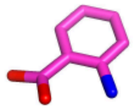
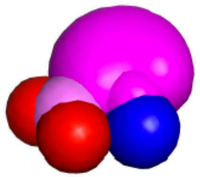
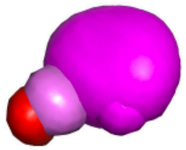
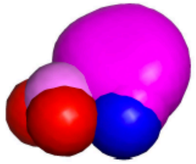
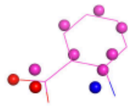
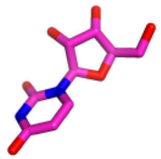
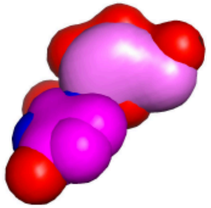
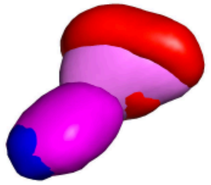
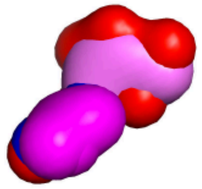
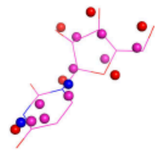
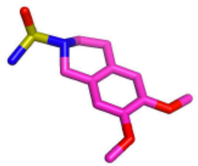
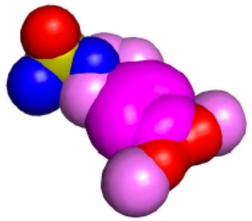
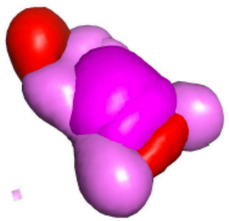
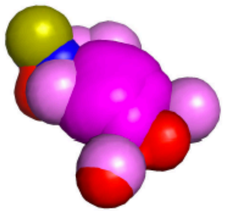
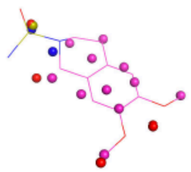
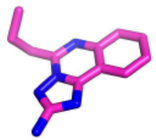
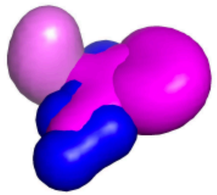
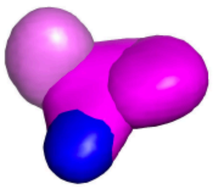
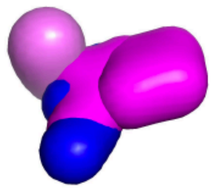
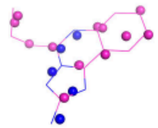
Variational Autoencoding Examples



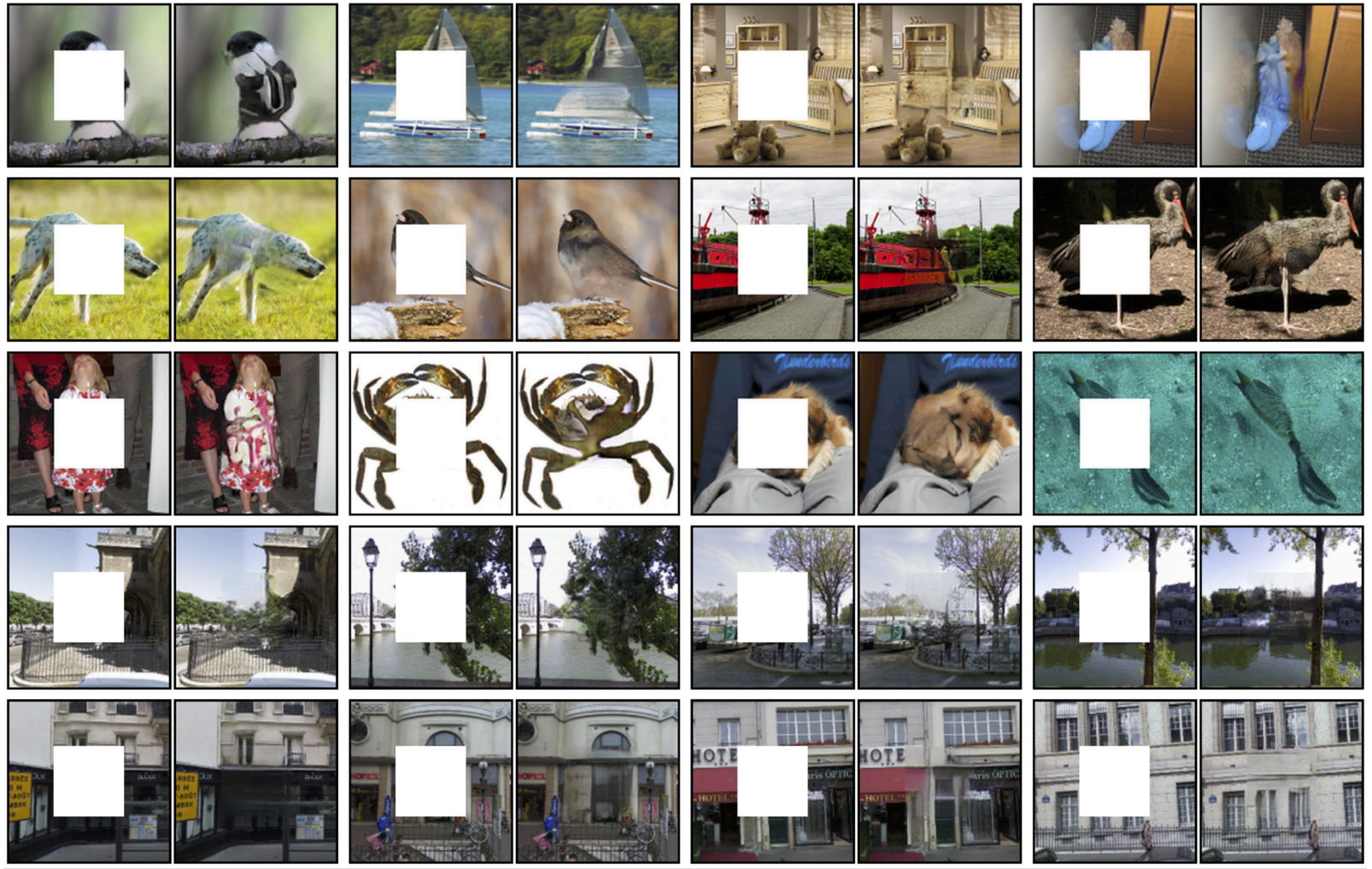


2BES

Variational Autoencoding Examples

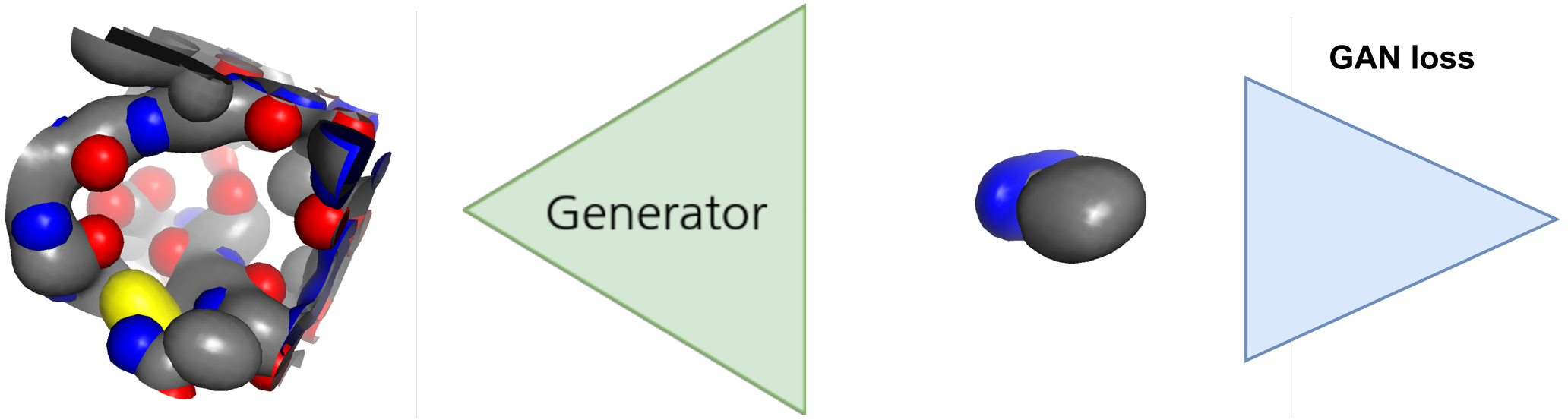
PDB	True structure	True density	Gen. density	Fit density	Fit structure	Gen. L2 distance	Fit L2 distance	Fit RMSD
3h78						9.4053	8.3141	0.6160
4jx9						13.8545	9.7198	0.8820
3igp						14.8525	12.5245	1.2066
4cwf						11.4730	9.0564	0.6725

Context Encoding



http://people.eecs.berkeley.edu/~pathak/context_encoder/

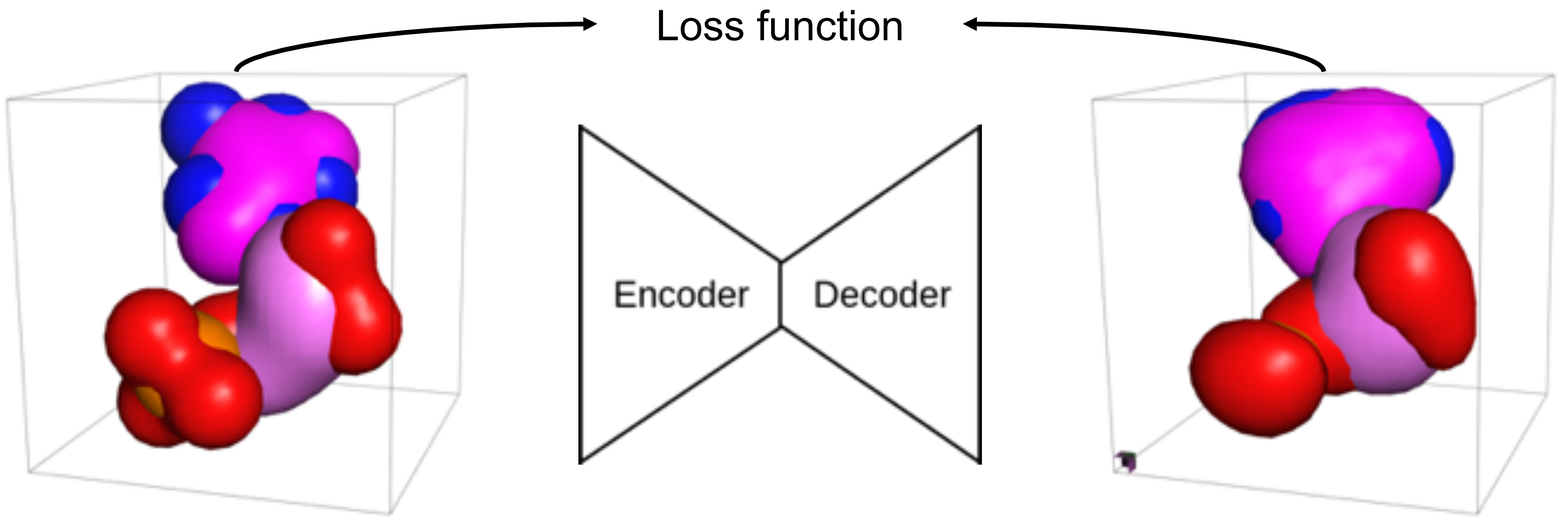
Context Encoding

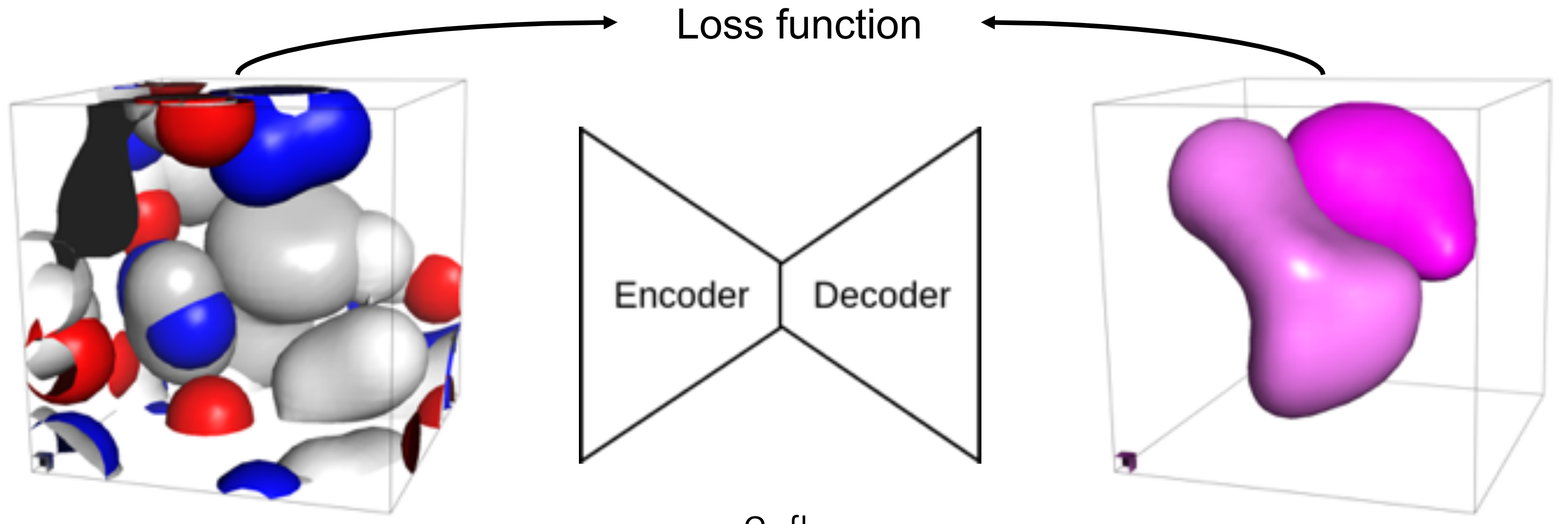


receptor grid

generated ligand grid

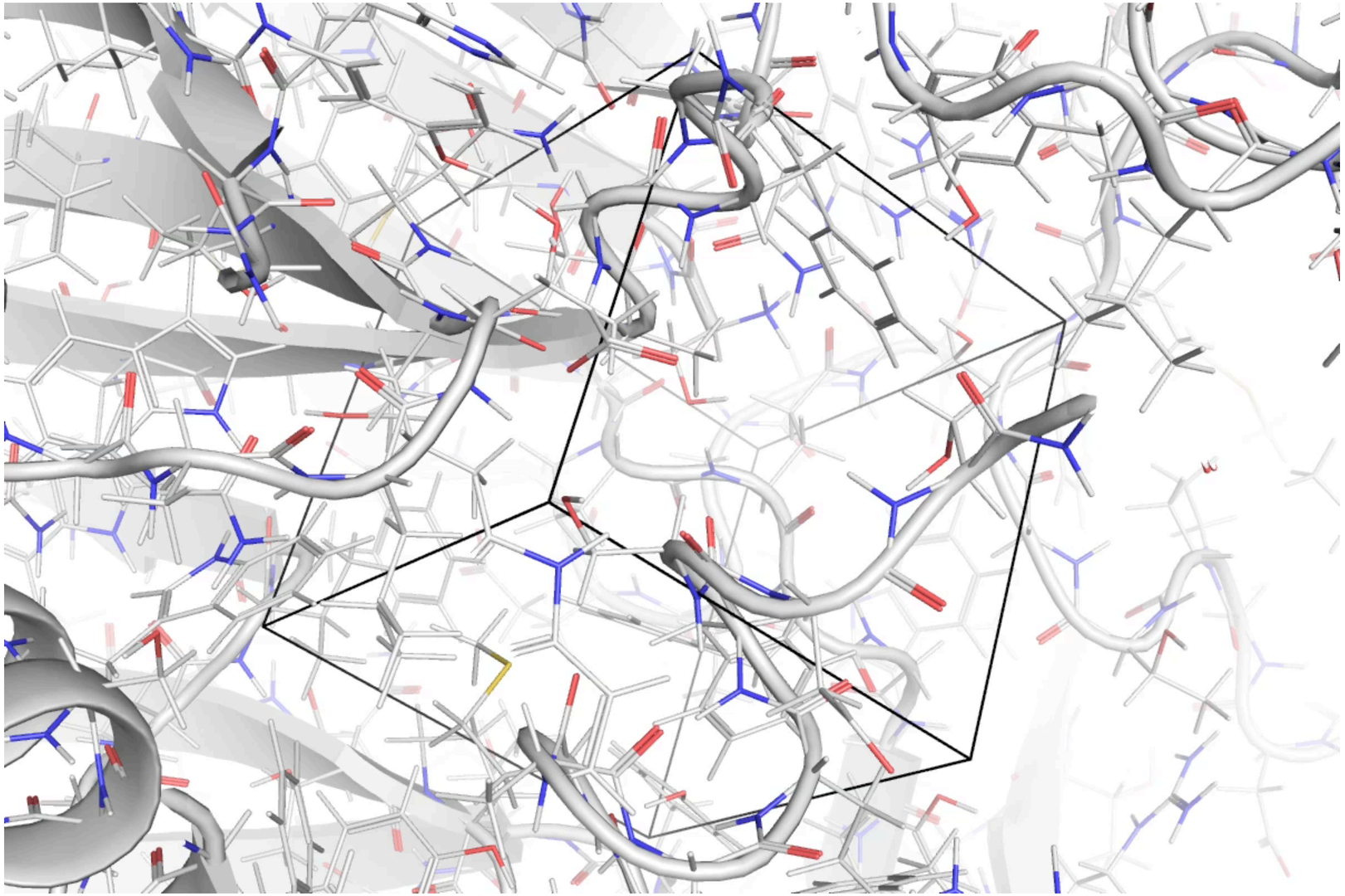
VAE vs. CE



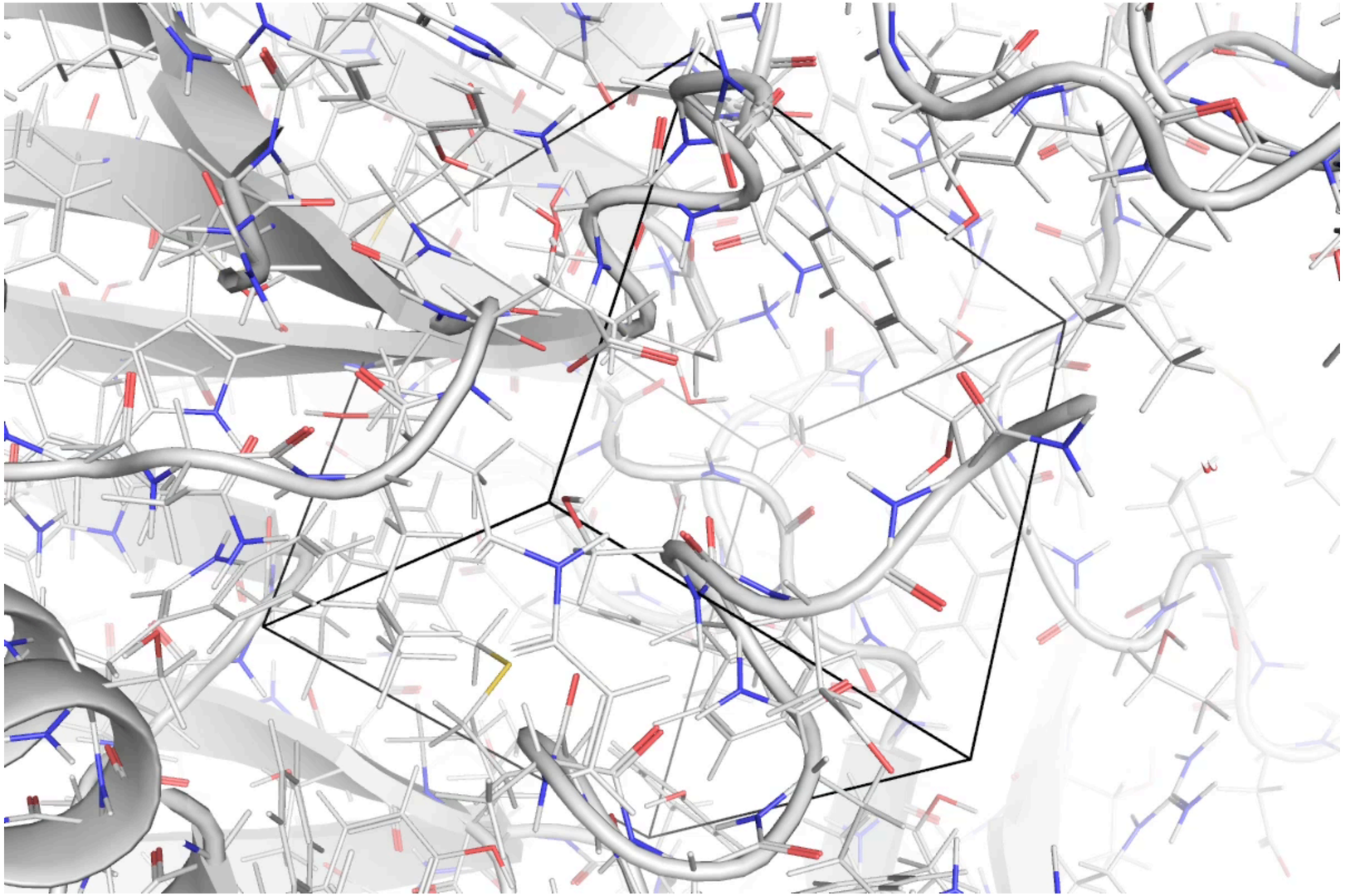


2vfk

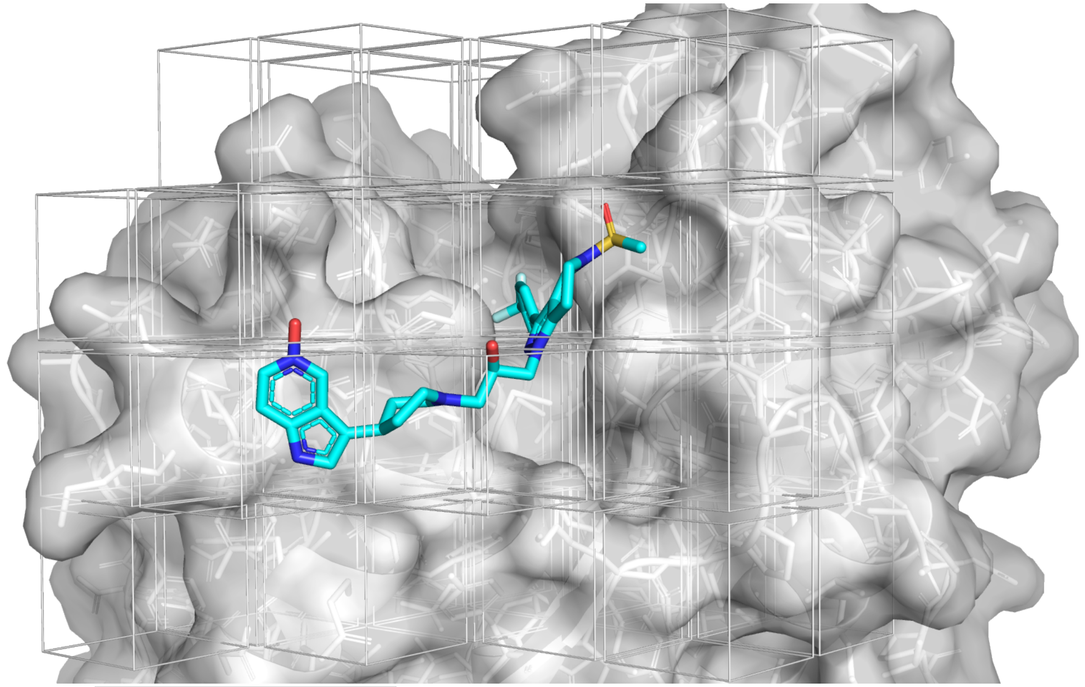
Conditioning on the Receptor



Conditioning on the Receptor



Molecular ML - Now With Promiscuity!



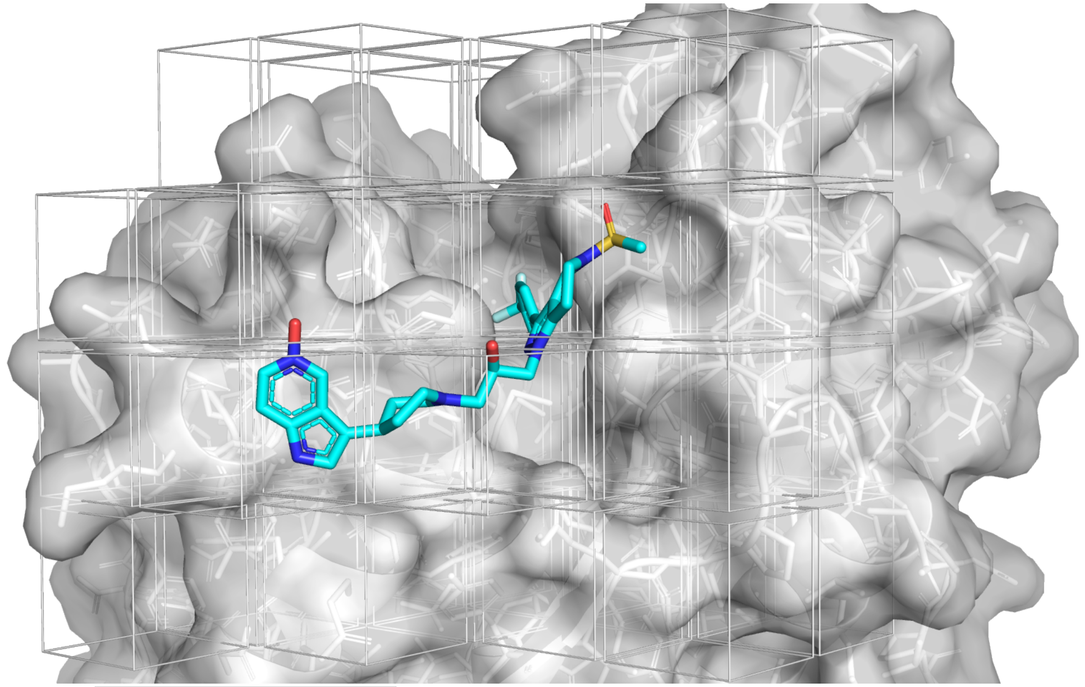
Caffe

libmolgrid

providing support for:

- balanced, randomized, stratified batches
- temporal and spatial recurrences
- generation of tensors from molecular input data, and not just grids either!

Molecular ML - Now With Promiscuity!



PYTORCH

Deep Learning with PyTorch

Caffe

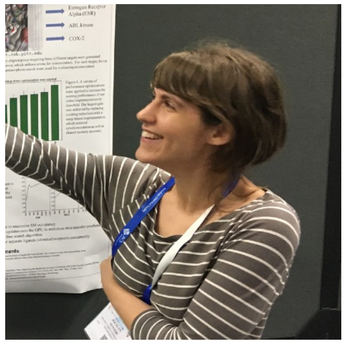
TensorFlow

libmolgrid *Coming soon!*

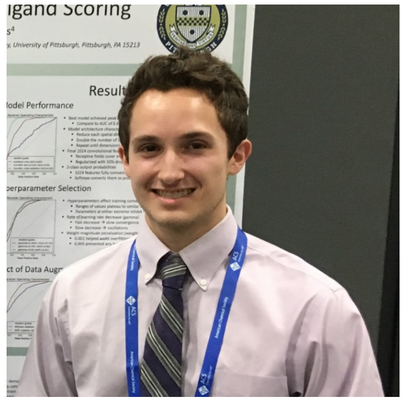
providing support for:

- balanced, randomized, stratified batches
- temporal and spatial recurrences
- generation of tensors from molecular input data, and not just grids either!

Acknowledgements



Jocelyn Sunseri



Matt Ragoza

Tomohide Masuda

Paul Francoeur

Department of Computational and Systems Biology

RELAY THERAPEUTICS

AI GRANT

National Institute of General Medical Sciences
R01GM108340

NVIDIA

Google Cloud

github.com/gnina

http://bits.csb.pitt.edu

@david_koes