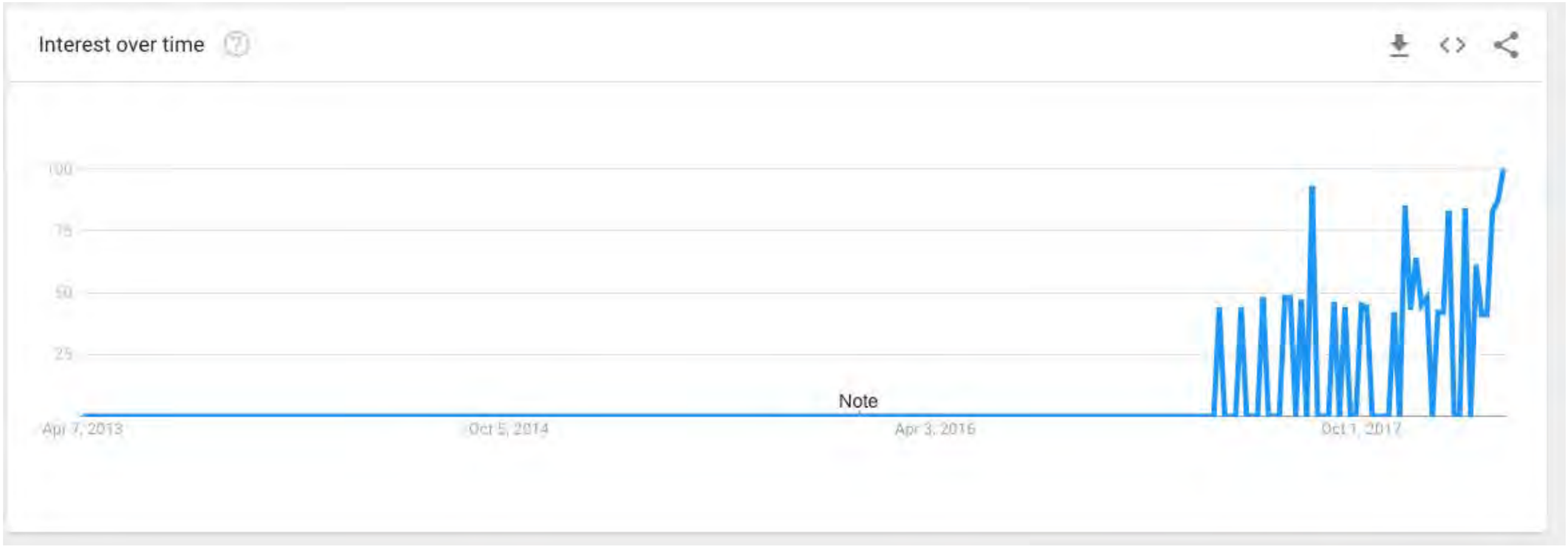
Artificial Intelligence Methods for Molecular Property Prediction

Evan N. Feinberg
Pande Lab

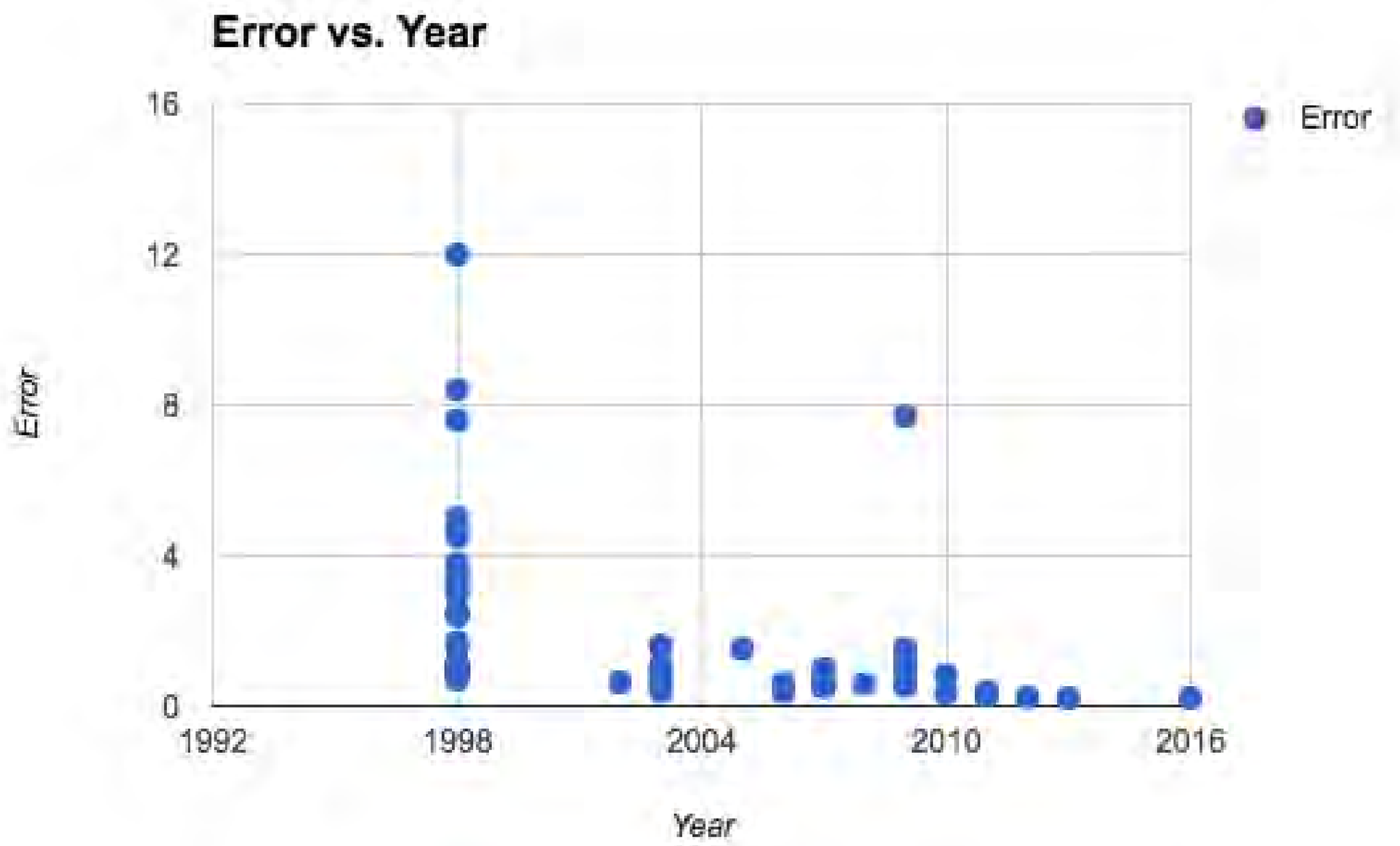
“AI Drug Discovery”



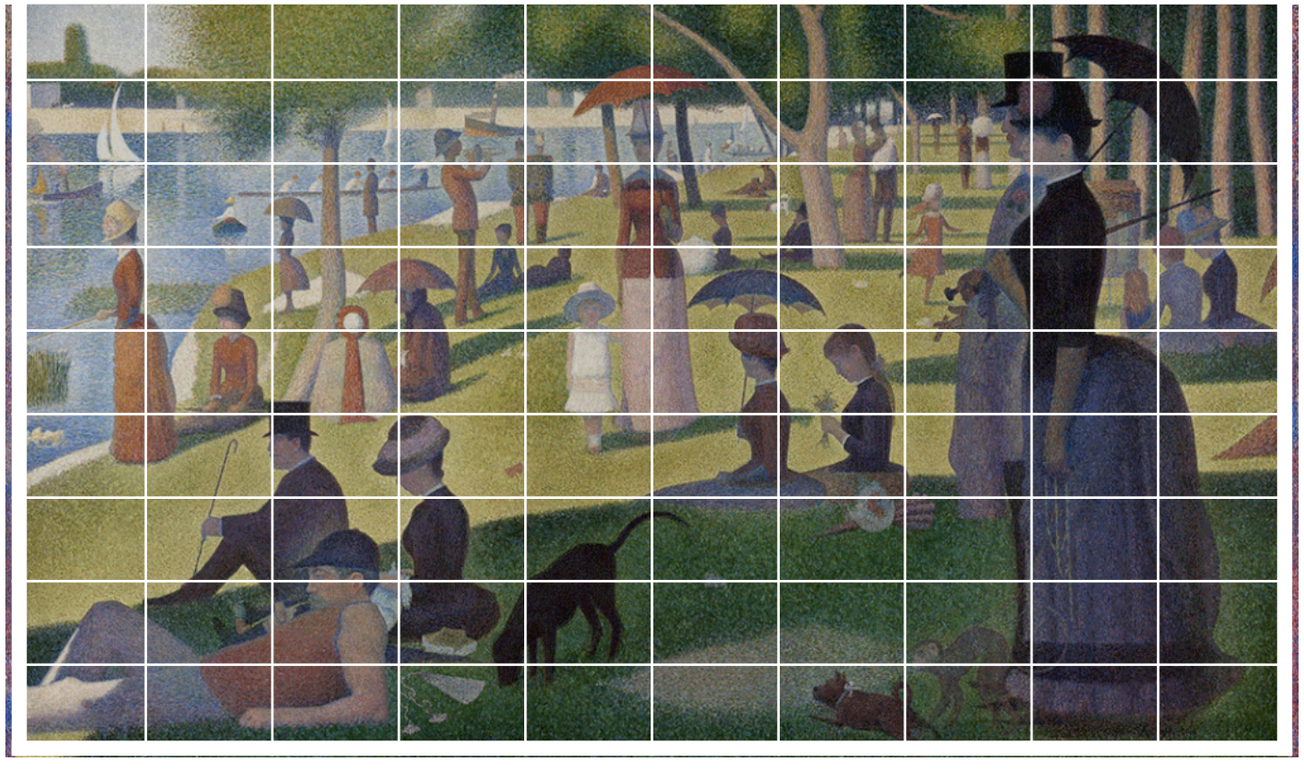
trends.google.com

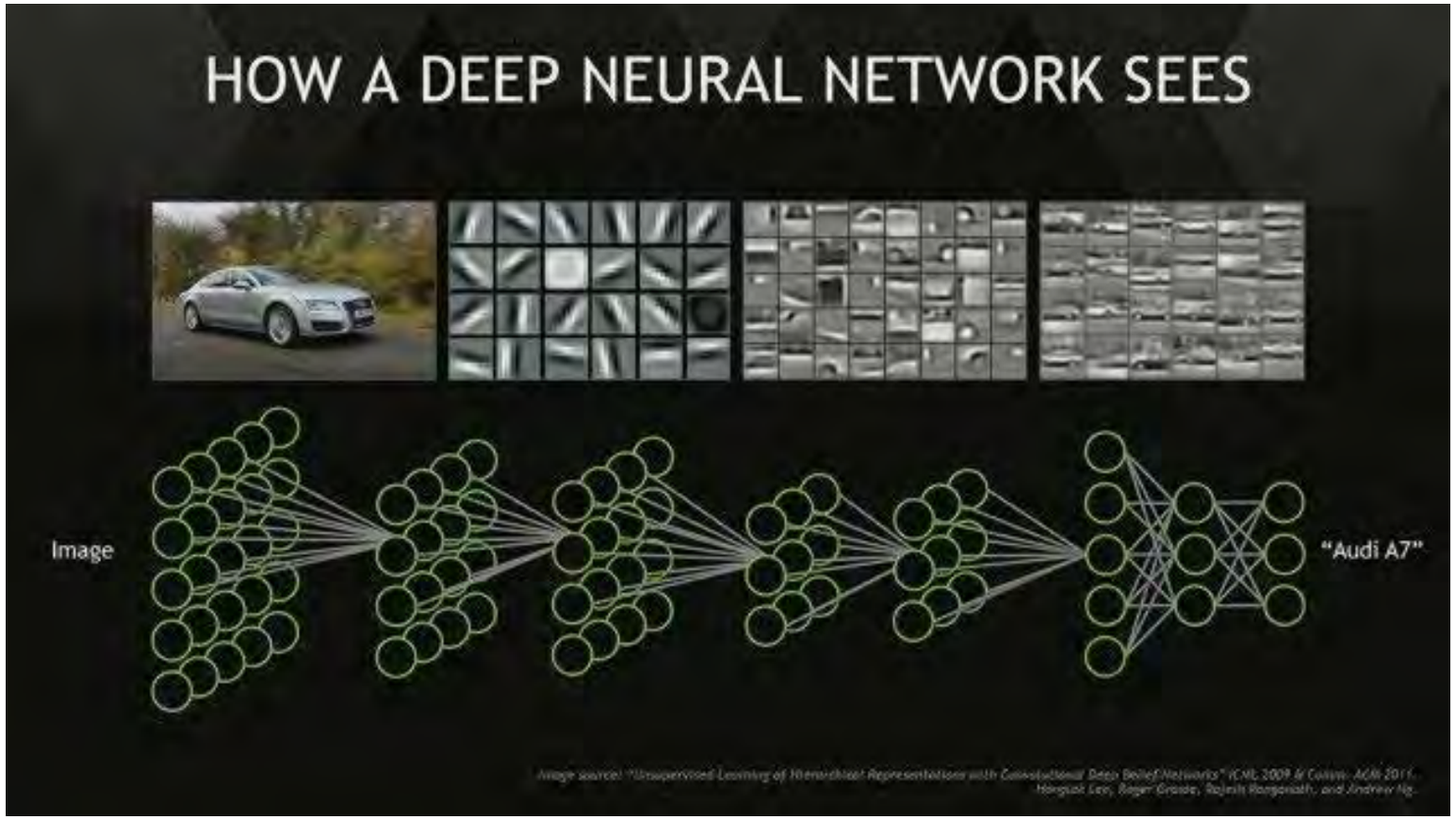
https://g.co/trends/3Y4Uf

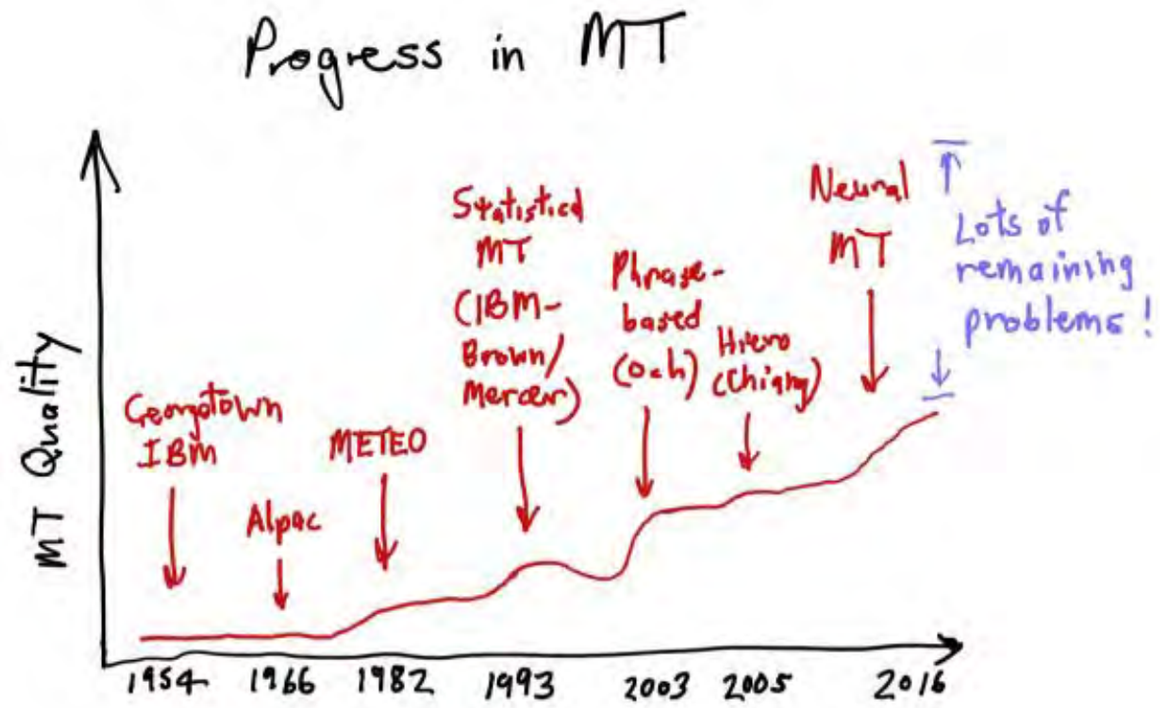
MNIST Performance over Time

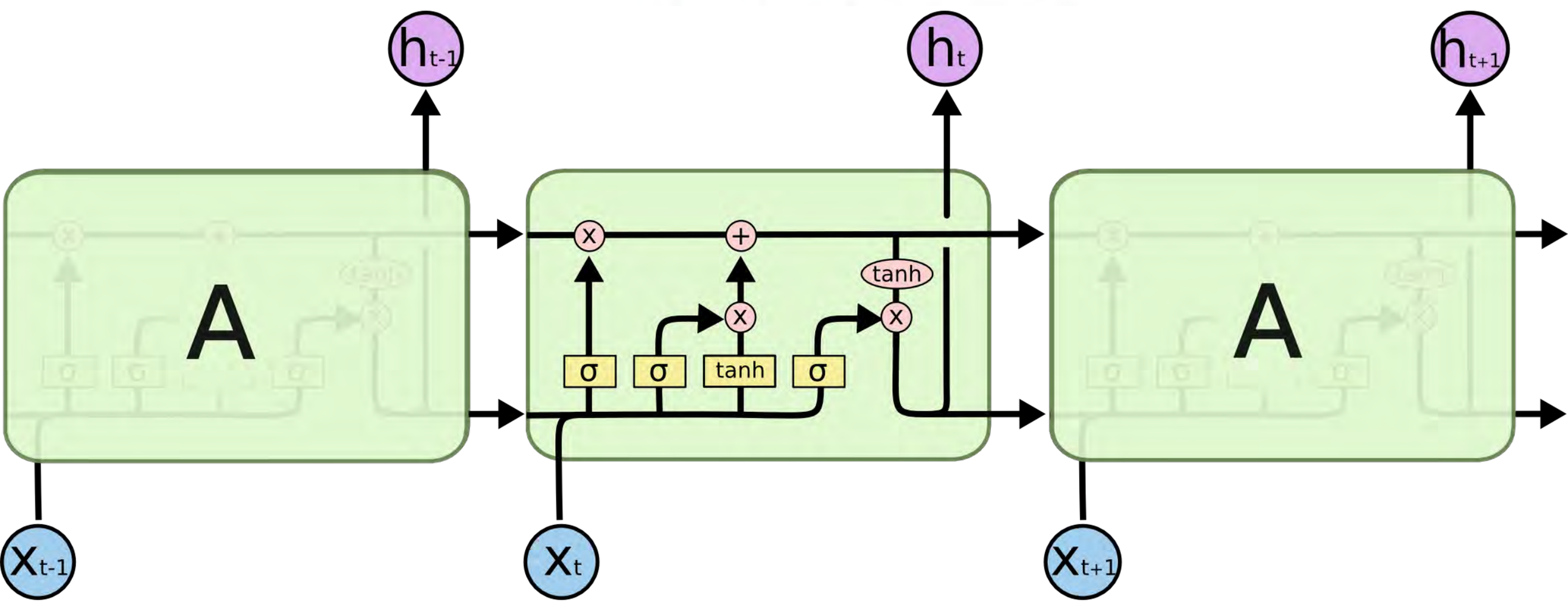


https://srconstantin.wordpress.com/2017/01/28/performance-trends-in-ai/



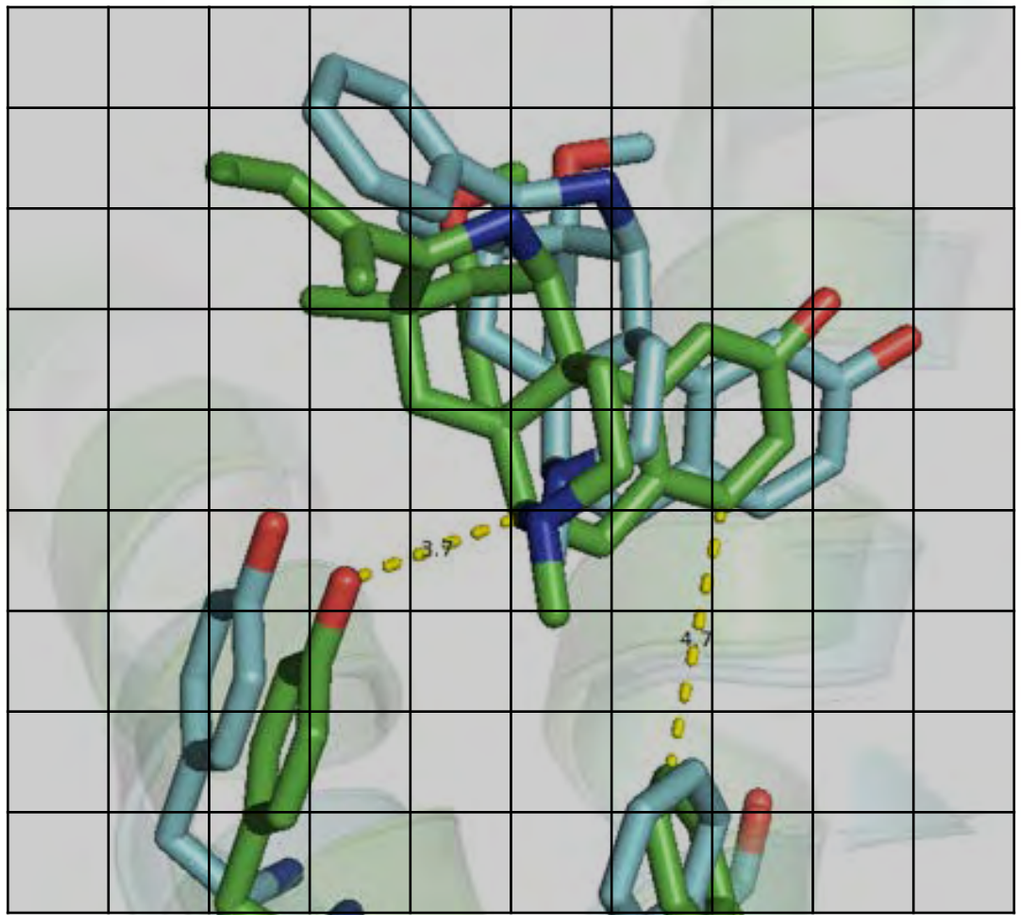

HOW A DEEP NEURAL NETWORK SEES
Image
"Audi A7"





http://colah.github.io/posts/2015-08-Understanding-LSTMs/

What’s the right approach for molecules?



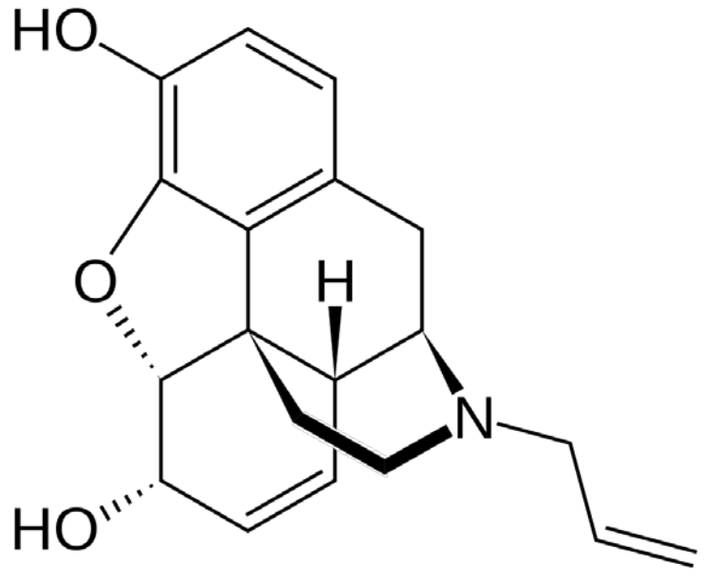
Talk Part 1 Roadmap

1. How to represent molecules?

2. How to represent systems of molecules?

3. New DNNs to leverage those representations

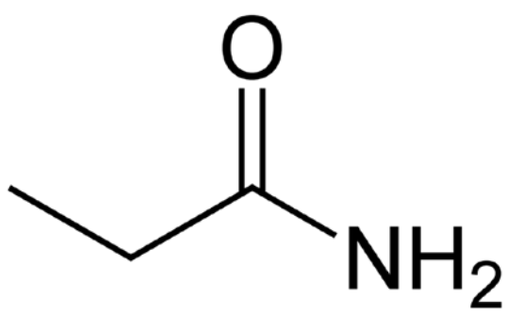
4. Applications to Drug Pharmacokinetics and Binding Prediction



$$x = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1f_{\text{in}}} \\ x_{21} & x_{22} & \cdots & x_{2f_{\text{in}}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nf_{\text{in}}} \end{bmatrix} \in \mathbb{R}^{N \times f_{\text{in}}}$$

$$A = \begin{bmatrix} 0 & A_{12} & \cdots & A_{1N} \\ A_{21} & 0 & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{N \times N},$$

$$\text{where: } A_{ij} = \begin{cases} 1, v_j \in N(v_i) \\ 0, \text{otherwise.} \end{cases}$$



$$x = \begin{array}{c} \begin{matrix} \text{Carbon} & \text{Oxygen} & \text{Nitrogen} & \text{sp3 hyb.} & \text{sp2 hyb.} \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Multi-Layer Perceptron / Fully Connected Neural Network

$$h^{(1)} = ReLU\left(W^{(1)} \cdot x\right)$$

$$h^{(2)} = ReLU\left(W^{(2)} \cdot h^{(1)}\right)$$

$$\vdots$$

$$h^{(K)} = W^{(K)} \cdot h^{(K-1)}$$

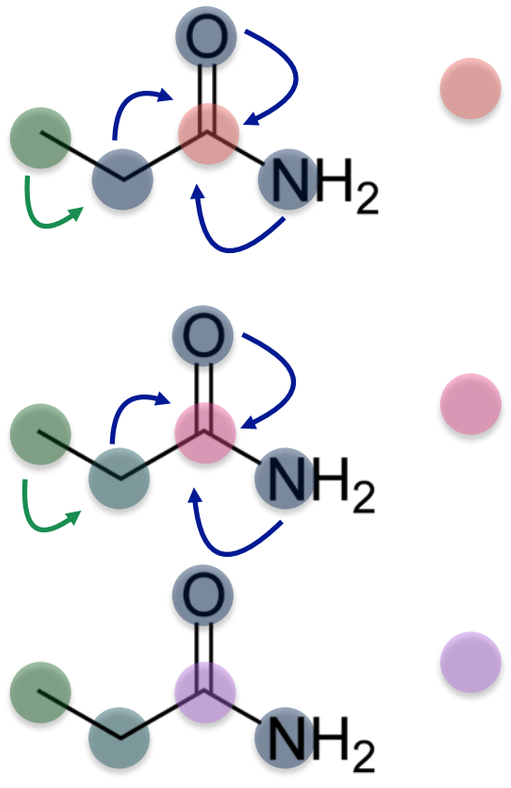
Graph Convolutional Neural Network (GCNN)

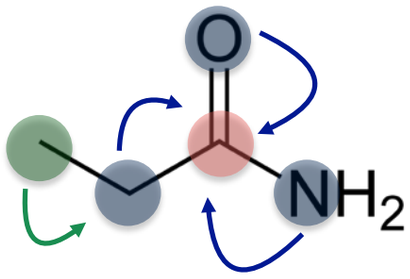
$$h^{(1)} = ReLU\left(W^{(1)} A \cdot x\right)$$

$$h^{(2)} = ReLU\left(W^{(2)} A \cdot h^{(1)}\right)$$

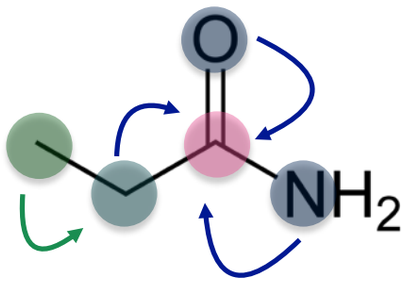
$$\vdots$$

$$h^{(K)} = ReLU\left(W^{(K)} A \cdot h^{(K-1)}\right)$$

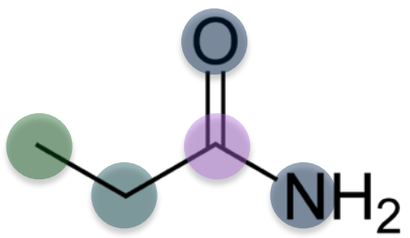




Input: SP2
Hybridized Carbon



After First Hidden Layer: RNN(SP2
Hybridized Carbon,
Weights * [Primary Amine Nitrogen +
Alkane Carbon with 2 Implicit Hydrogens
+ SP2 Hybridized Oxygen])



After Second Hidden Layer: Information
has Propagated from All Heavy Atoms
To Carbonyl Carbon

Gated Graph Neural Network (Fancy GCNN)

$$h^{(1)} = GRU\left(x, \sum_{e}^{N_{\text{et}}} W^{(e)} A^{(e)} x\right)$$

$$h^{(2)} = GRU\left(h^{(1)}, \sum_{e}^{N_{\text{et}}} W^{(e)} A^{(e)} h^{(1)}\right)$$

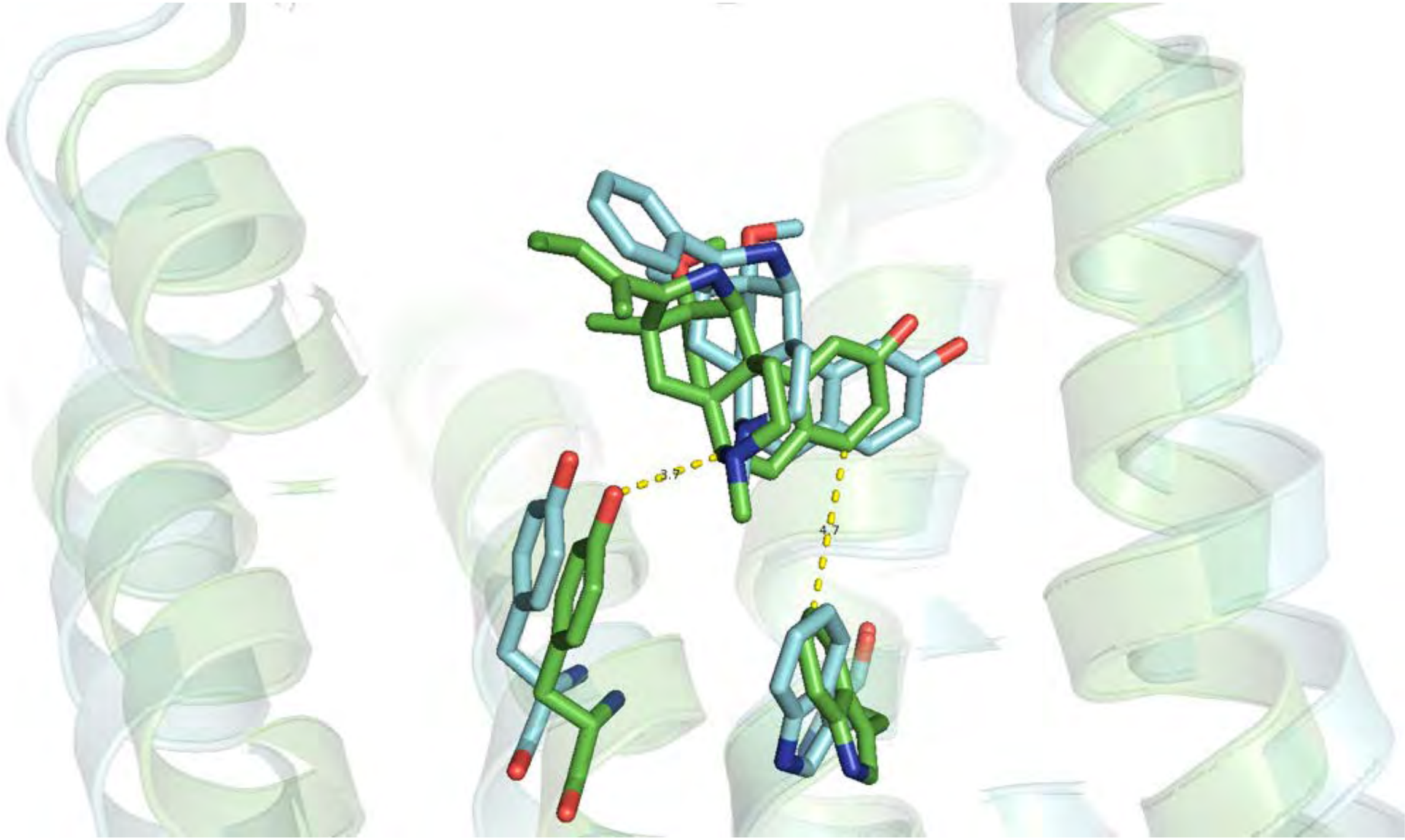
$$\vdots$$

$$h^{(K)} = GRU\left(h^{(K-1)}, \sum_{e}^{N_{et}} W^{(e)} A^{(e)} h^{(K-1)}\right),$$

Update

Message

What about two (or more!) interacting molecules?



$$A = \left(\begin{bmatrix} A_{111} & A_{121} & \cdots & A_{1N1} \\ A_{211} & A_{221} & \cdots & A_{2N1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N11} & A_{N21} & \cdots & A_{NN1} \end{bmatrix}, \cdots, \begin{bmatrix} A_{11N_{\text{et}}} & A_{12N_{\text{et}}} & \cdots & A_{1NN_{\text{et}}} \\ A_{21N_{\text{et}}} & A_{22N_{\text{et}}} & \cdots & A_{2NN_{\text{et}}} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1N_{\text{et}}} & A_{N2N_{\text{et}}} & \cdots & A_{NNN_{\text{et}}} \end{bmatrix} \right)$$

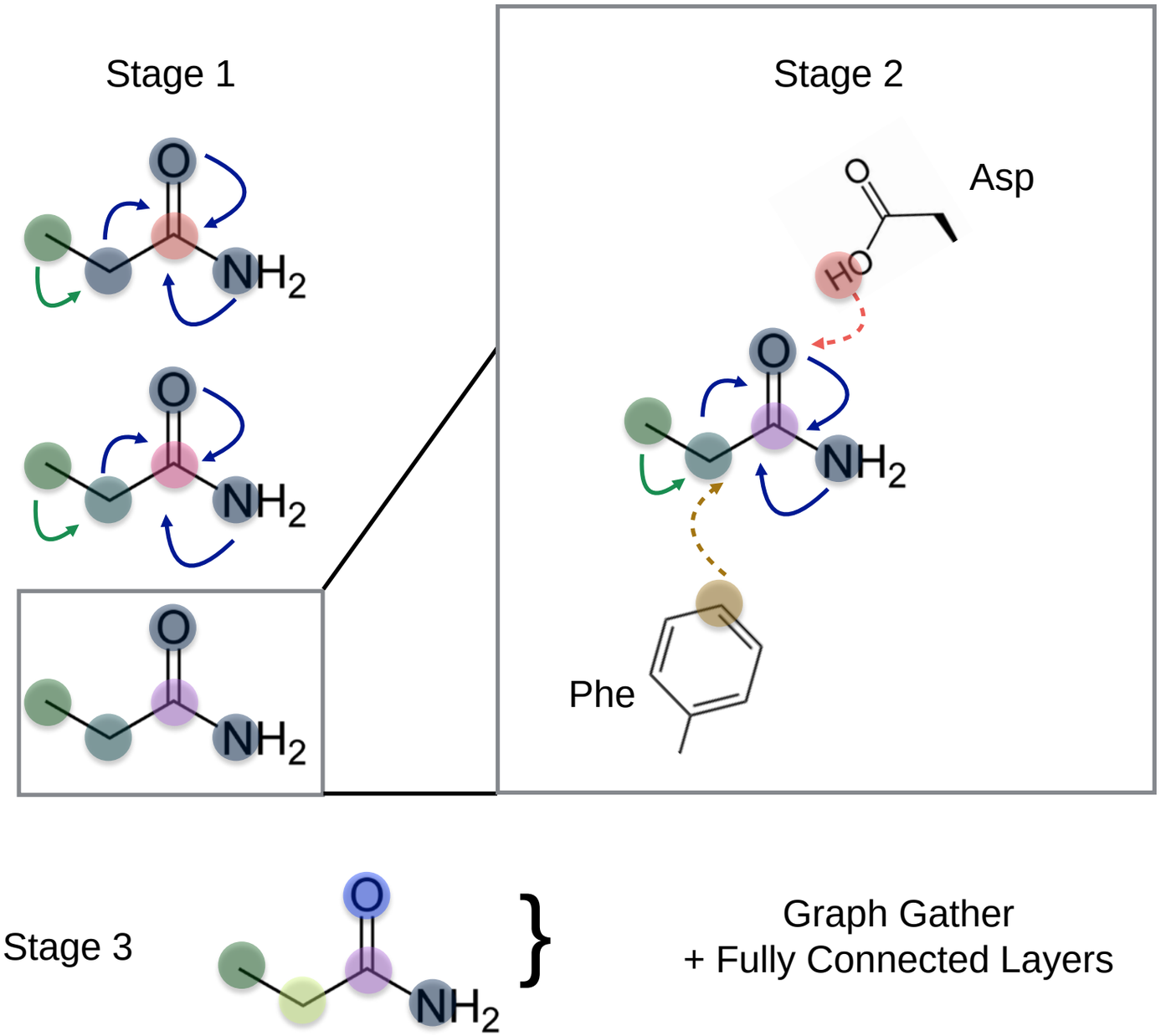
$$\in \mathbb{R}^{N \times N \times N_{\text{et}}} \text{ ,where: } A_{ijk} = \begin{cases} 1, v_j \in N(v_i) \text{ and } e_{i,j} = k \\ 0, \text{otherwise.} \end{cases}$$

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{bmatrix} \in \mathbb{R}^{N \times N}$$

$$= \begin{bmatrix} A_{L:L} & A_{L:P} \\ A_{P:L} & A_{P:P} \end{bmatrix},$$

$$\text{where: } A_{ij} = \begin{cases} 1, v_j \in N(v_i) \\ 0, \text{otherwise.} \end{cases}$$

Staged Spatial Graph Convolution (SSGC)



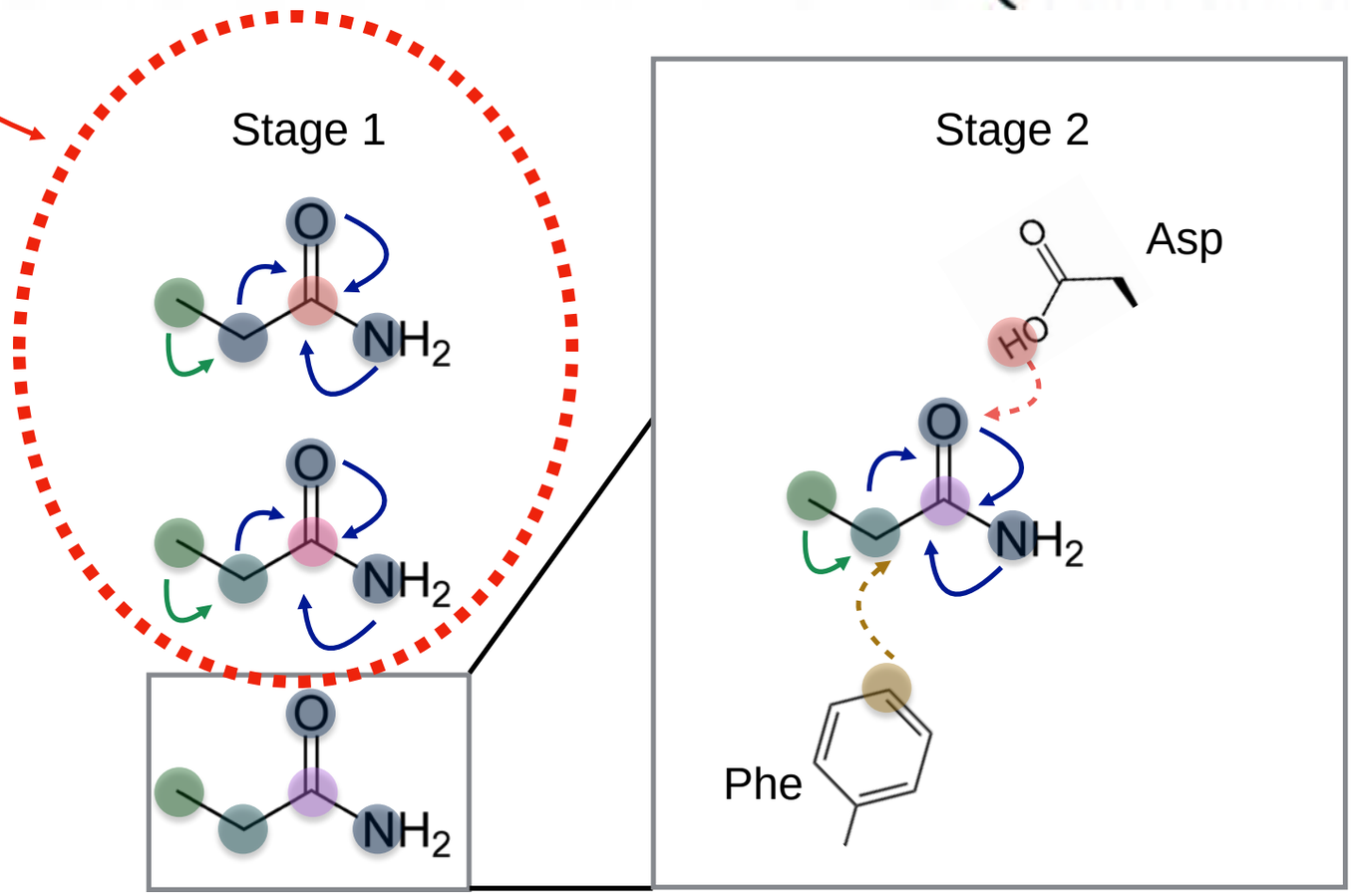
Feinberg, E.N., Sur, D., Husic, B.E., Mai, D., Li, Y., Yang, J., Ramsundar, B. and Pande, V.S., 2018. Spatial Graph Convolutions for Drug Discovery. *arXiv preprint arXiv:1803.04465*.

$$A = \left(\begin{bmatrix} A_{111} & A_{121} & \cdots & A_{1N1} \\ A_{211} & A_{221} & \cdots & A_{2N1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N11} & A_{N21} & \cdots & A_{NN1} \end{bmatrix}, \cdots, \begin{bmatrix} A_{11N_{\text{et}}} & A_{12N_{\text{et}}} & \cdots & A_{1NN_{\text{et}}} \\ A_{21N_{\text{et}}} & A_{22N_{\text{et}}} & \cdots & A_{2NN_{\text{et}}} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1N_{\text{et}}} & A_{N2N_{\text{et}}} & \cdots & A_{NNN_{\text{et}}} \end{bmatrix} \right)$$

$$\in \mathbb{R}^{N \times N \times N_{\text{et}}} \text{ ,where: } A_{ijk} = \begin{cases} 1, v_j \in N(v_i) \text{ and } e_{i,j} = k \\ 0, \text{otherwise.} \end{cases}$$

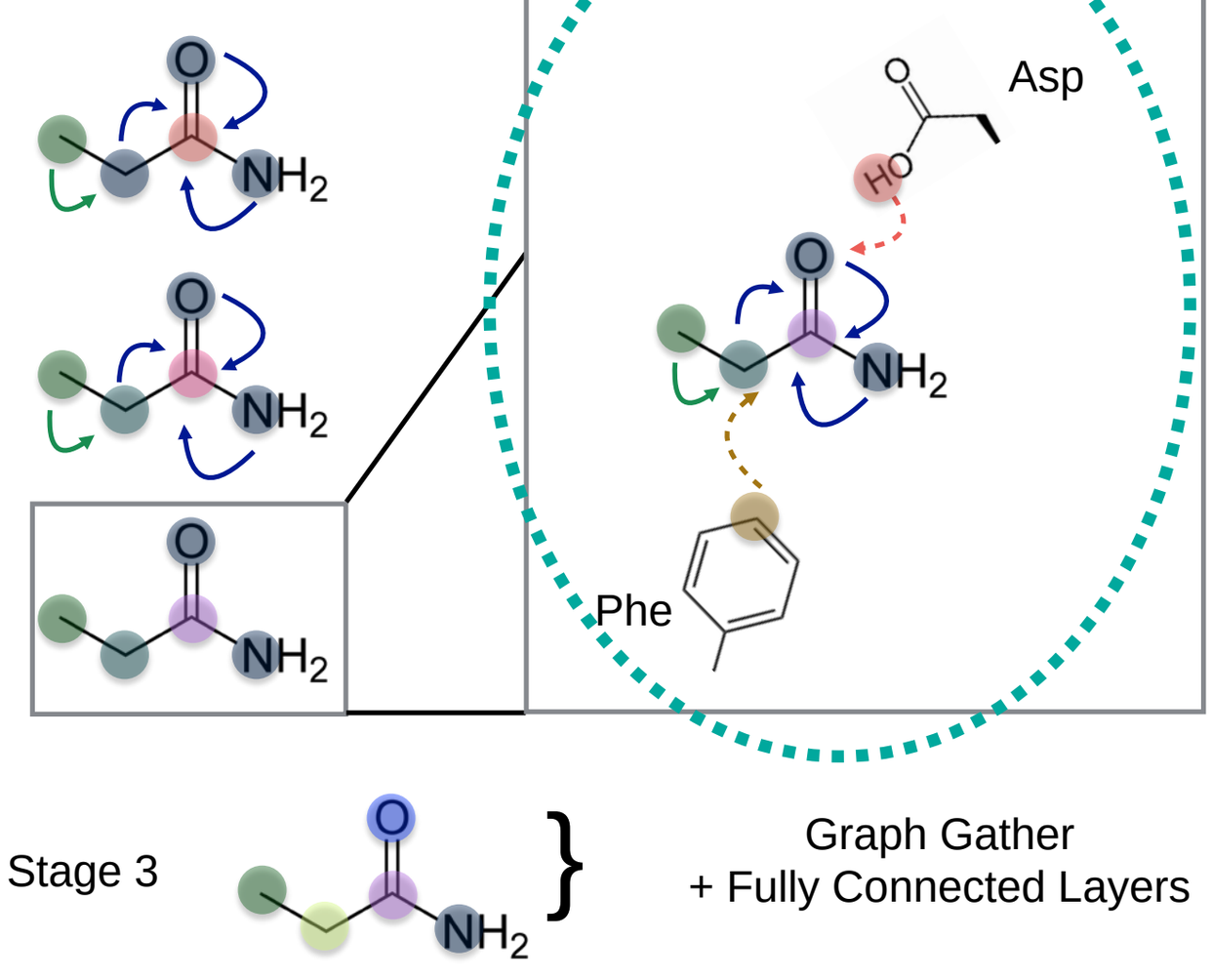
$$A = \left(\begin{bmatrix} A_{111} & A_{121} & \cdots & A_{1N1} \\ A_{211} & A_{221} & \cdots & A_{2N1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N11} & A_{N21} & \cdots & A_{NN1} \end{bmatrix}, \cdots, \begin{bmatrix} A_{11N_{\text{et}}} & A_{12N_{\text{et}}} & \cdots & A_{1NN_{\text{et}}} \\ A_{21N_{\text{et}}} & A_{22N_{\text{et}}} & \cdots & A_{2NN_{\text{et}}} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1N_{\text{et}}} & A_{N2N_{\text{et}}} & \cdots & A_{NNN_{\text{et}}} \end{bmatrix} \right)$$

$$\in \mathbb{R}^{N \times N \times N_{\text{et}}} \text{ ,where: } A_{ijk} = \begin{cases} 1, v_j \in N(v_i) \text{ and } e_{i,j} = k \\ 0, \text{otherwise.} \end{cases}$$



$$A = \left(\begin{bmatrix} A_{111} & A_{121} & \cdots & A_{1N1} \\ A_{211} & A_{221} & \cdots & A_{2N1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N11} & A_{N21} & \cdots & A_{NN1} \end{bmatrix}, \cdots, \begin{bmatrix} A_{11N_{\text{et}}} & A_{12N_{\text{et}}} & \cdots & A_{1NN_{\text{et}}} \\ A_{21N_{\text{et}}} & A_{22N_{\text{et}}} & \cdots & A_{2NN_{\text{et}}} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1N_{\text{et}}} & A_{N2N_{\text{et}}} & \cdots & A_{NNN_{\text{et}}} \end{bmatrix} \right)$$

$$\in \mathbb{R}^{N \times N \times N_{\text{et}}} \text{ ,where: } A_{ijk} = \begin{cases} 1, v_j \in N(v_i) \text{ and } e_{i,j} = k \\ 0, \text{otherwise.} \end{cases}$$



Predicting Molecular Properties

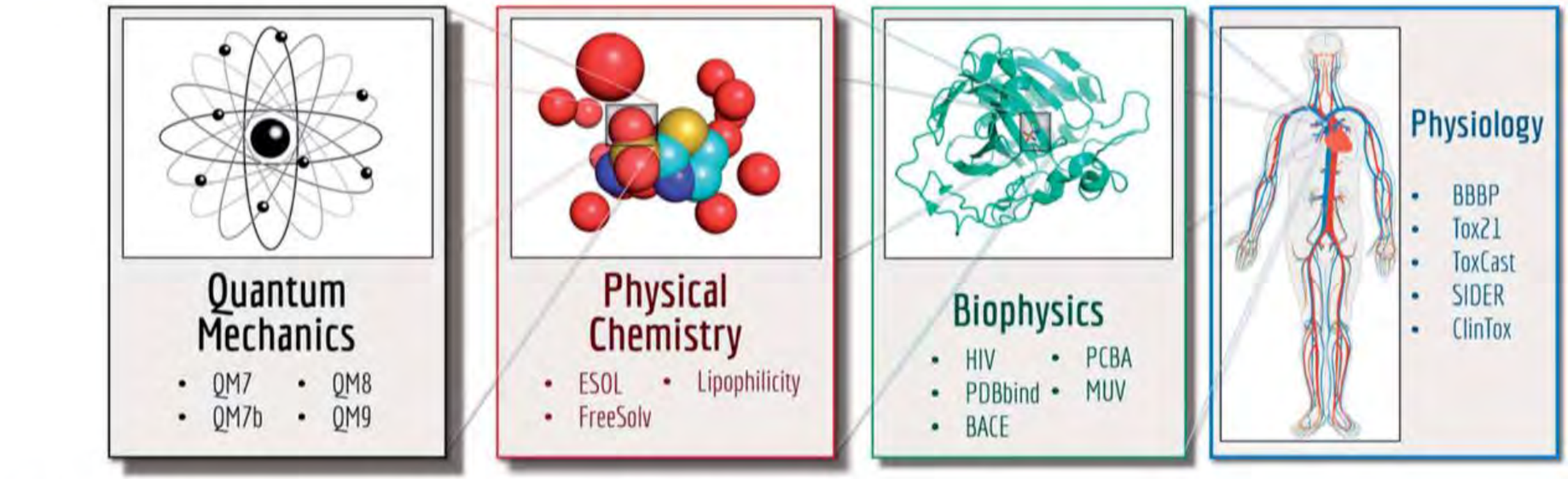
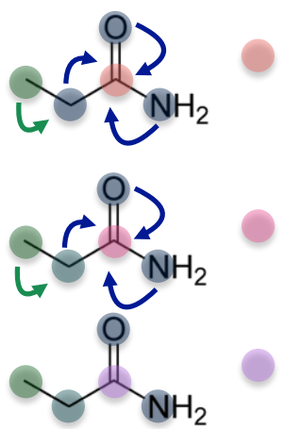


Fig. 2 Tasks in different datasets focus on different levels of properties of molecules.



Wu, Z., Ramsundar, B., **Feinberg, E.N**., Gomes, J., Geniesse, C., Pappu, A.S., Leswing, K. and Pande, V., 2018. MoleculeNet: a benchmark for molecular machine learning. *Chemical Science*, *9*(2), pp.513-530.

Challenge 1: Protein-Ligand Binding Affinity (PDBBind)

1,300 Compounds

Model Architecture	Test Pearson *R*
Staged Spatial Graph Convolution (Ours)	**0.822 (0.021)**
PotentialNet Spatial Graph Convolution	0.795 (0.040)
GGNN, Ligand-only control	0.650 (0.017)
TopologyNet *(No Validation Set)*	**0.826**
RF-Score	0.783
X-Score	0.643

Challenge 2: Ligand-based Toxicity (Tox21) Prediction

8,014 Compounds

Model Architecture	AUC, Valid	AUC, Test
Random Forests	0.763	0.769
Graph Convolution, Standard	0.825	0.829
Graph Convolution, Weave	0.828	0.820
Graph Convolution, PotentialNet (Ours)	0.878	**0.863**

Challenge 3: Ligand-based Solubility (ESOL) Prediction

1,128 Compounds

Model Architecture	RMSE, Valid	RMSE, Test
Random Forests	1.16	1.07
Graph Convolution, Standard	1.05	0.97
Graph Convolution, Weave	0.57	0.61
Graph Convolution, PotentialNet (Ours)	0.54	**0.50**

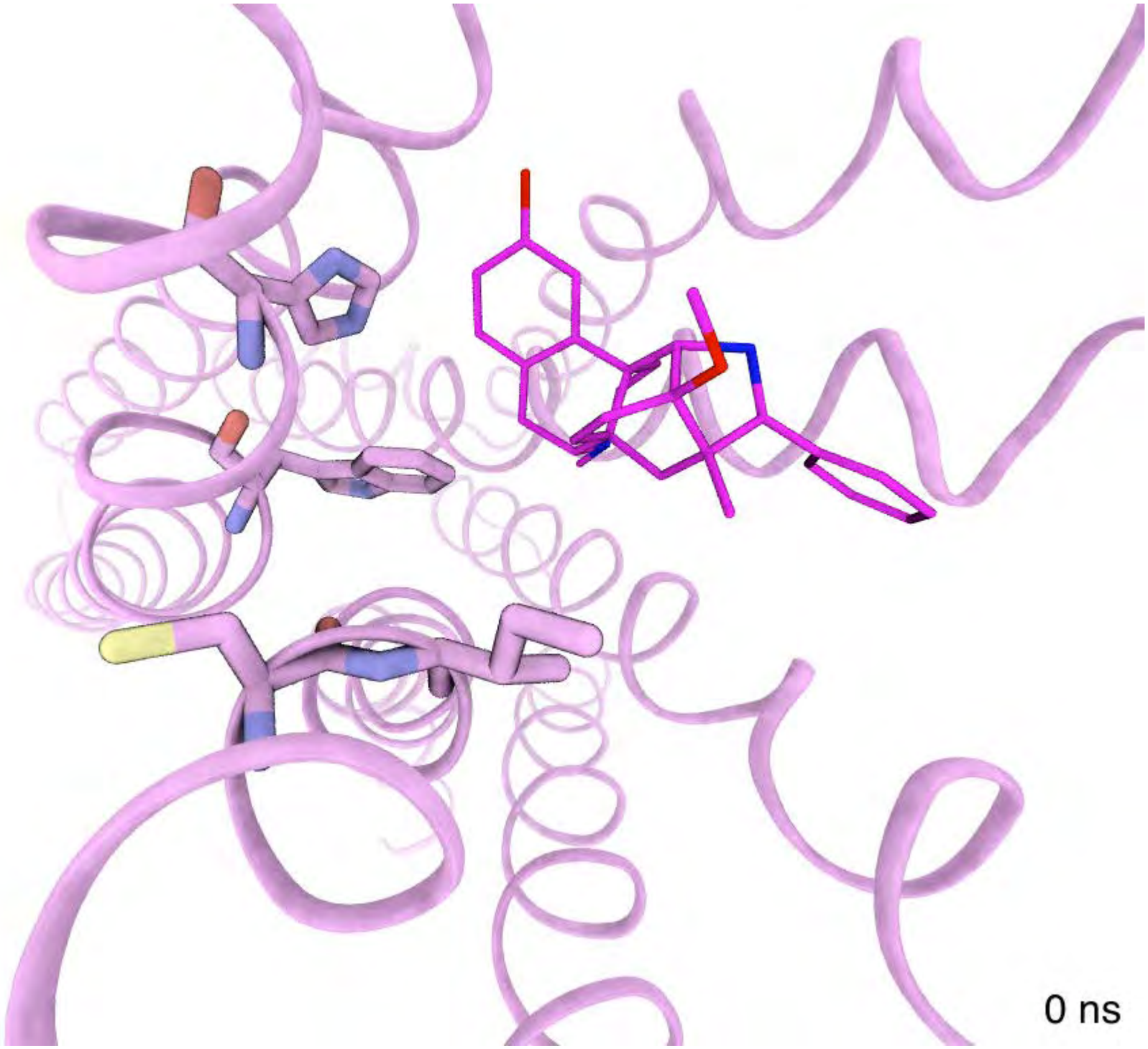
Challenge 4: QM8 Electronic Property Prediction

21,768 Compounds

Model Architecture	MAE, Valid	MAE, Test
Multitask	1.55E-02	1.50E-02
Graph Convolution, Standard	1.50E-02	1.48E-02
Graph Convolution, [SEP]MPNN	1.46E-02	1.43E-02
Graph Convolution, Staged Spatial (Ours)	1.14E-02	**1.12E-02**

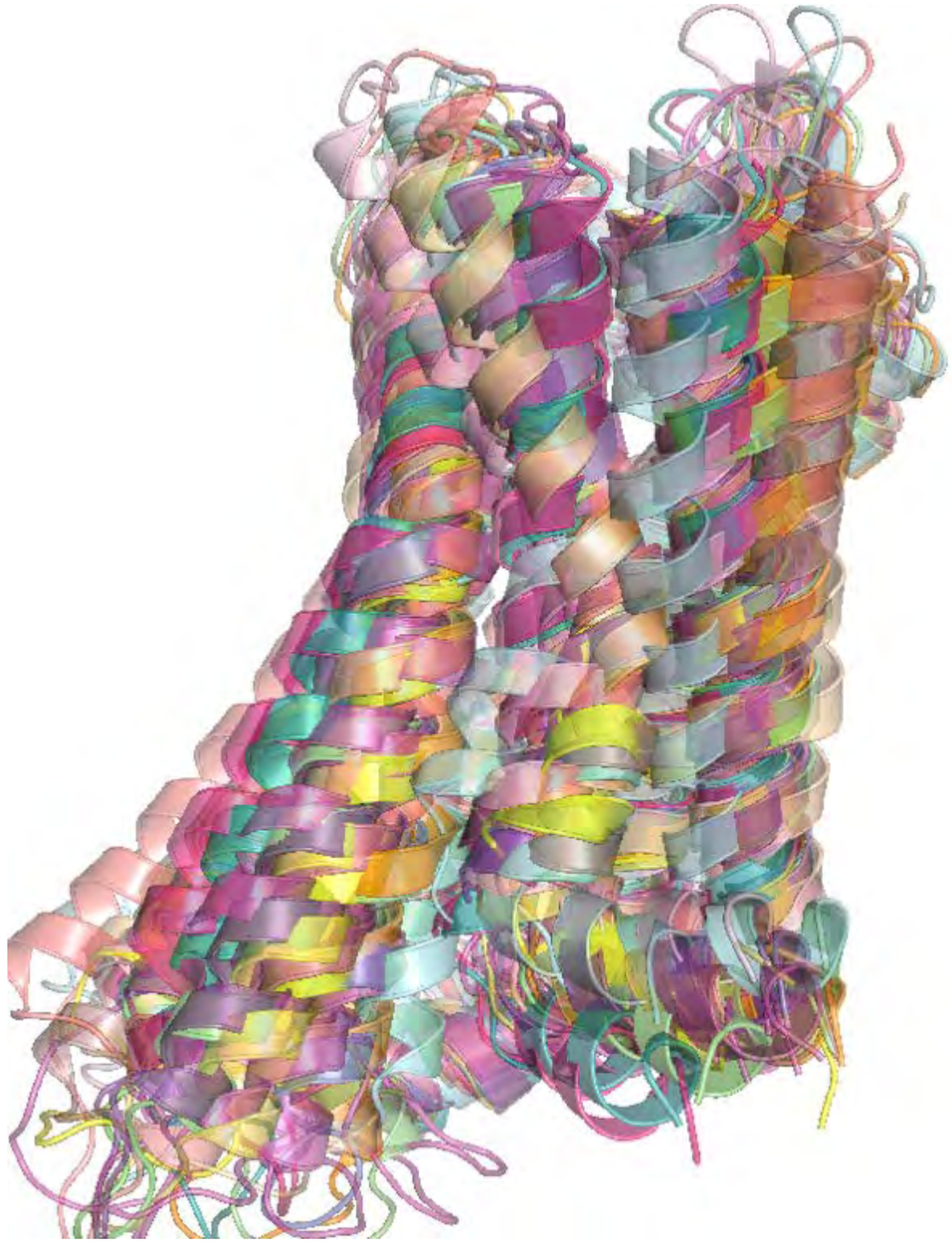
Part 2: Harnessing Molecular Dynamics with Machine Learning to Systematically Discover Lead Molecules

μ Opioid Receptor

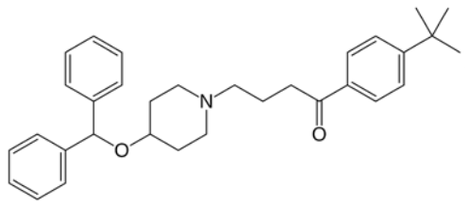


Feinberg, EN, Farimani AB, Hernandez CX, Pande VS. Kinetic Machine Learning Unravels Ligand-Directed Conformational Change of μ Opioid Receptor. biorXiv*. in Review.* http://www.biorxiv.org/content/early/2017/07/31/170886

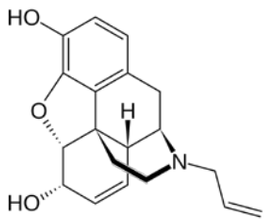
So, if Drug Targets, like GPCRs, sample many conformations….



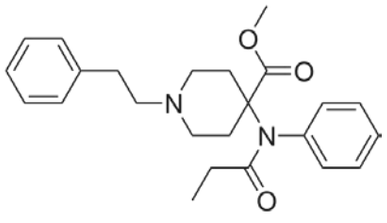
…then why does virtual screening dock ligands to only one receptor state?



Ebastine (non-binder)

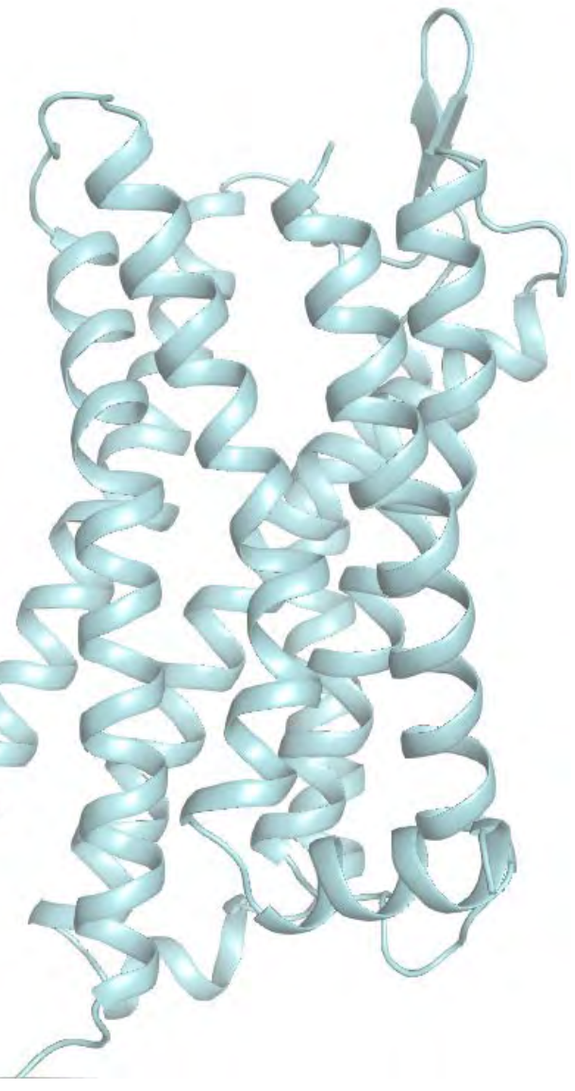


Nalorphine (antagonist)

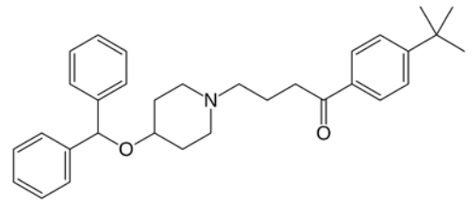


Carfentanil (Agonist)

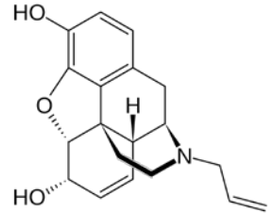
Molecular Docking



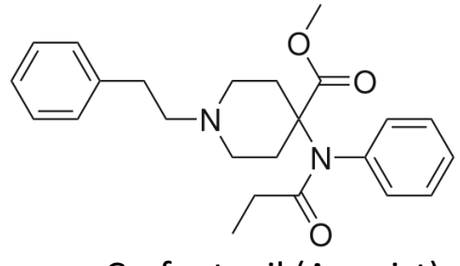
Model-Free Rank



Ebastine (non-binder)

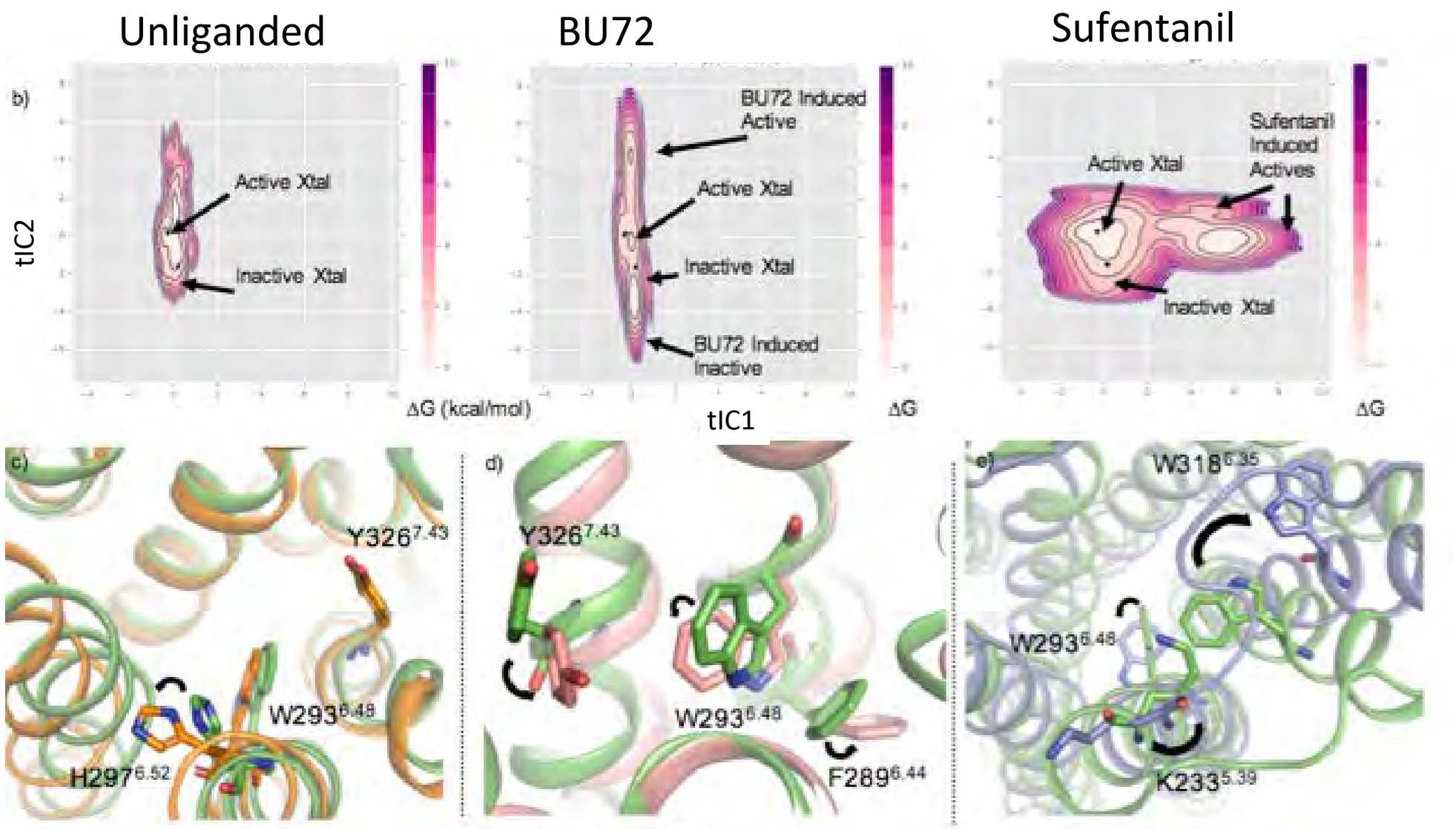


Nalorphine (antagonist)



Carfentanil (Agonist)

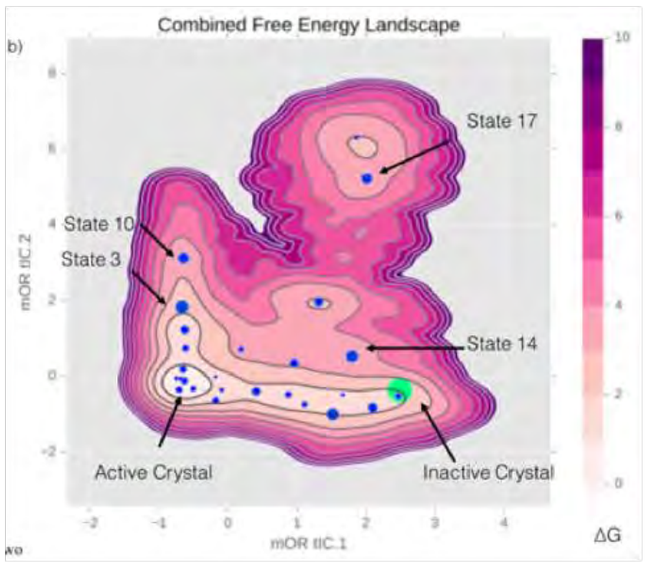
MD simulations & Unsupervised ML capture salient conformations of drug target



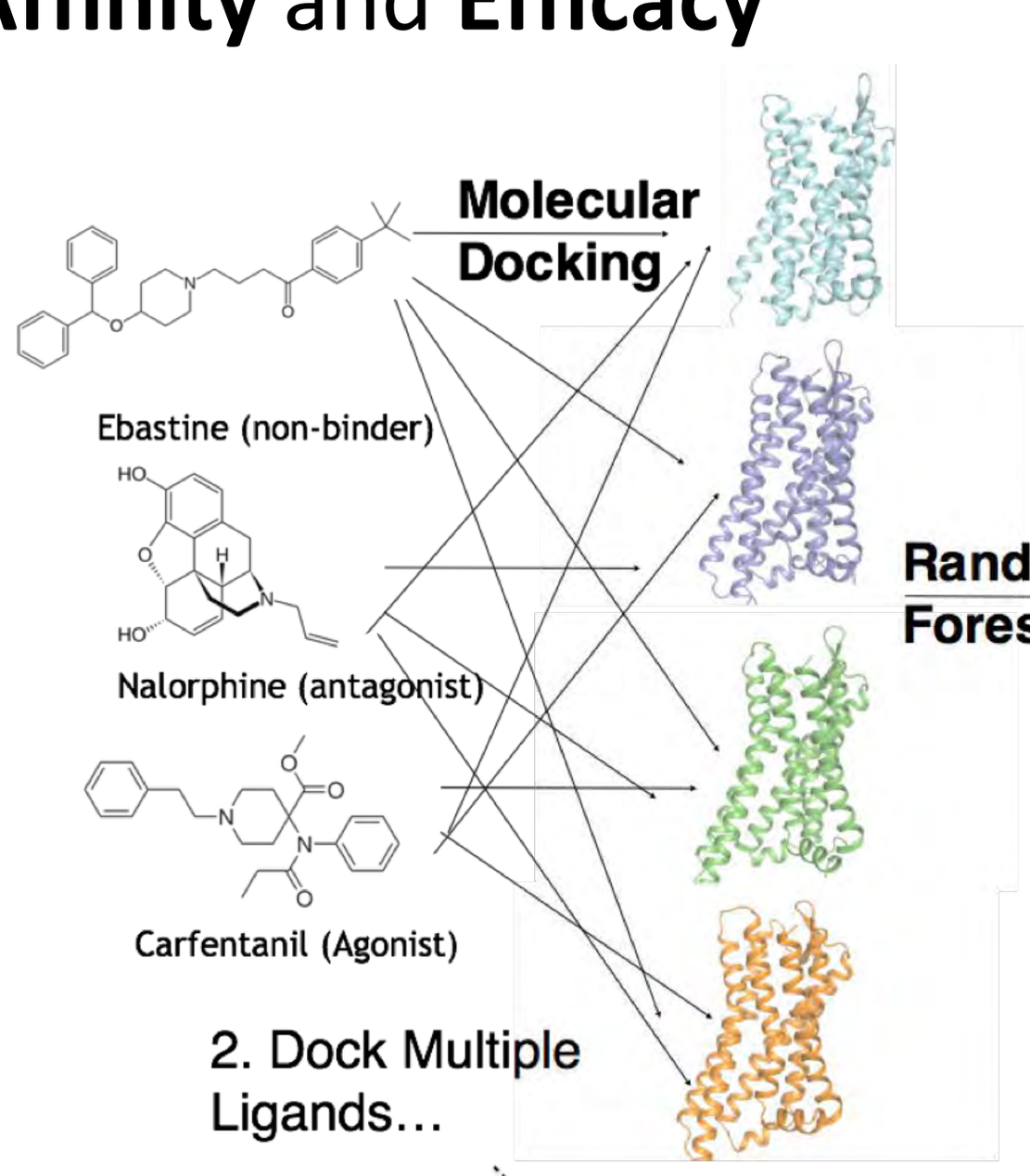
McGibbon, Robert T., Brooke E. Husic, and Vijay S. Pande. "Identification of simple reaction coordinates from complex dynamics." *The Journal of Chemical Physics* 146, no. 4 (2017): 044109.

Feinberg, EN, Farimani AB, Hernandez CX, Pande VS. Kinetic Machine Learning Unravels Ligand-Directed Conformational Change of μ Opioid Receptor. biorXiv*. in Review.* http://www.biorxiv.org/content/early/2017/07/31/170886

Supervised Machine Learning exploits Conformational Diversity from Molecular Simulation to Enhance the Prediction of **Affinity** and **Efficacy**



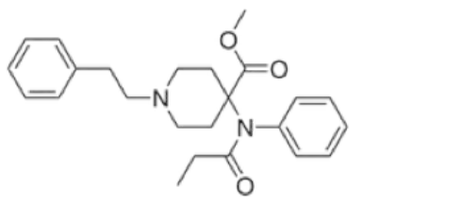
1.Simulate
+ tICA
+ Cluster

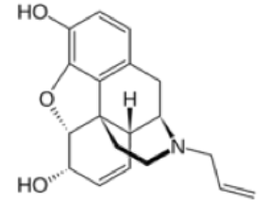


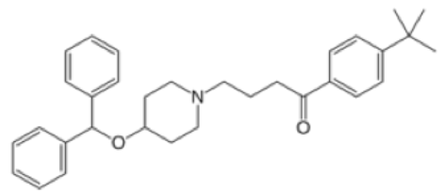
2. Dock Multiple Ligands…

…to Multiple Conformations from MD.

Random Forests


Carfentanil (Agonist)


Nalorphine (antagonist)


Ebastine (non-binder)

3. Train Model with Known Ligands
4. Virtually Screen Library of Unknowns

Task	**Split Type**	**AUC (Crystals alone)**	**AUC (Crystal + MD structures)**
Agonism	Random	0.73	**0.85**
Agonism	Scaffold (Fentanyl)	0.81	**0.91**
Agonism	Scaffold (Methadone)	0.89	**0.94**
Binding	Random	0.64	**0.79**
Binding	Scaffold	0.64	**0.78**

Stanford Compound Library (135,500 ligands)

↓

Dock to MD States

↓

Obtain 135,500 x 25 matrix of scores

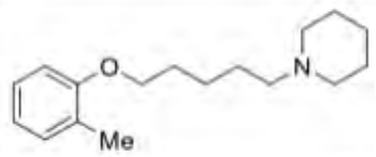
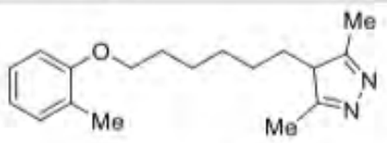
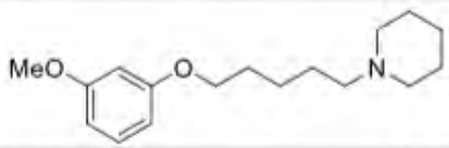
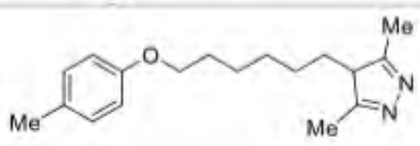
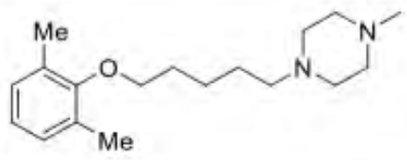
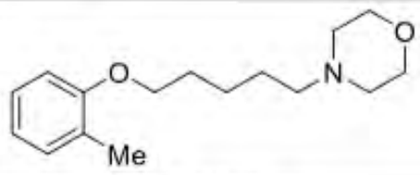
↓

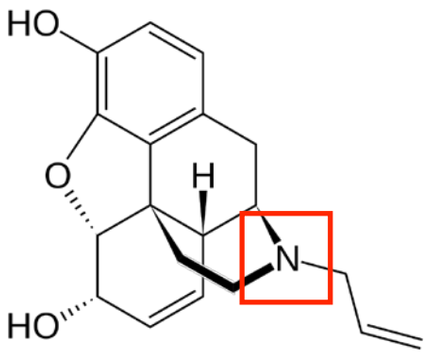
Apply Random Forest affinity and agonism models

↓

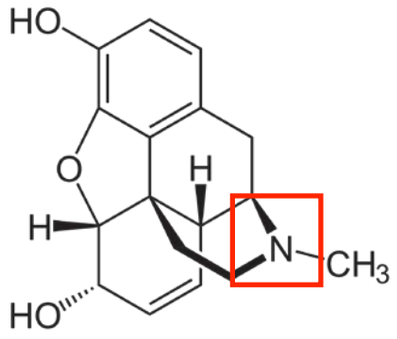
Collaborate with chemists at Sloan-Kettering to experimentally test

Extended Table Data 9. Receptor binding of FMP4 and structurally similar analogs

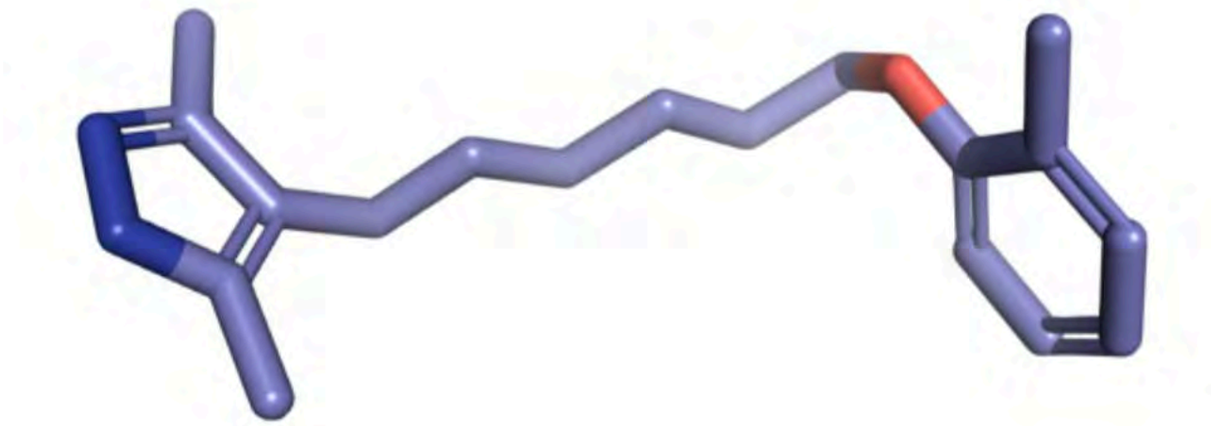
			Binding Data - K_i (nM)[a]		
FMP#	**Structure**	**MW**	**MOR-1**	**KOR-1**	**DOR-1**
1		261.409	841.4 ±241.9	7386 ±2376	17874 ±7615
4		286.419	3217 ±153	2503 ±523	8143 ±1398
6		277.408	>30 µM	6941 ±1970	>30 µM
16		286.419	1748 ±492	4918 ±2235	9058 ±566
21		290.451	>30 µM	5503 ±4194	>30 µM
30		263.381	11315 ±2262	6009 ±2163	12761 ±856
	DAMGO		3.3 ± 0.43 [b]		
	U50,488h			0.73 ± 0.32 [b]	
	DPDPE				1.39 ± 0.67 [b]



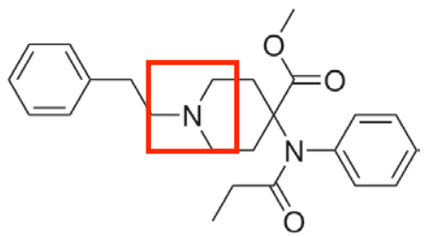
Nalorphine



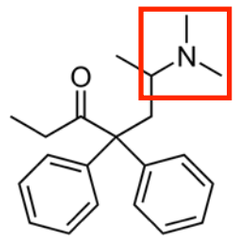
Morphine



FMP4 achieves binding and agonism at μ Opioid Receptor *without* a tertiary amine group



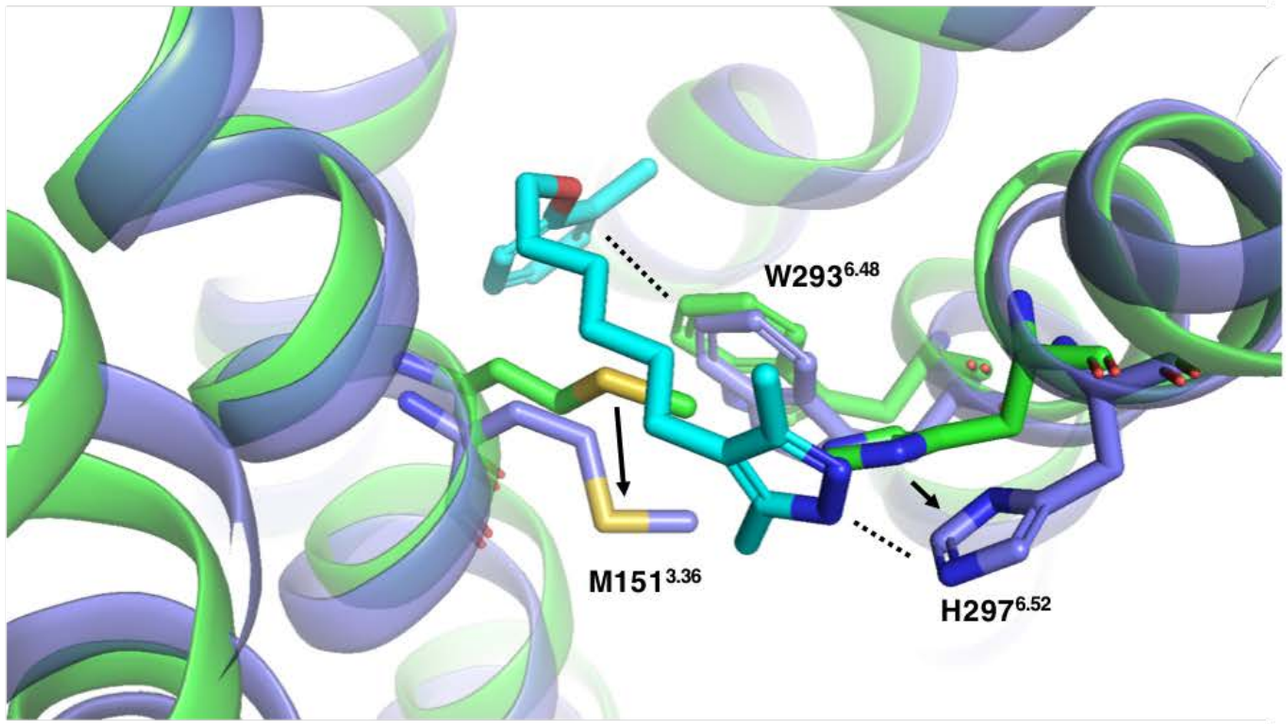
Carfentanil

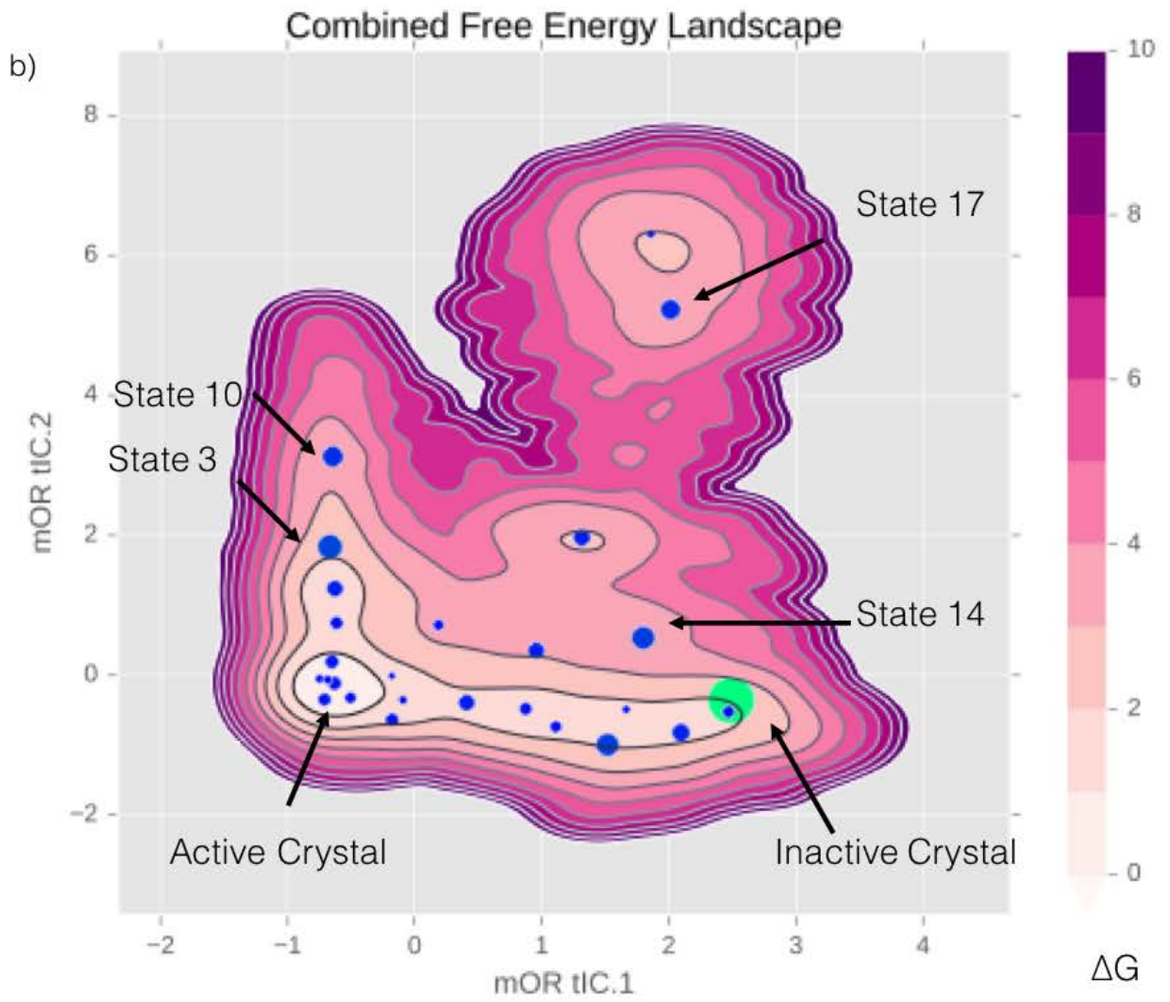


Methadone

Crystal Structure (5C1M)

Simulated FMP4-Specific Structure





Feinberg, E.N., Farimani, A.B., Uprety, R., Hunkele, A., Pasternak, G.W., Majumdar, S. and Pande, V.S., 2018. Machine Learning Harnesses Molecular Dynamics to Discover New mu Opioid Chemotypes. *arXiv preprint arXiv:1803.04479*.

Conclusions

PAPERS

1. A.B. Farimani*, E.N. Feinberg*, Vijay S. Pande. "Binding Pathway of Opiates to μ-Opioid Receptors Revealed by Unsupervised Machine Learning." arXiv. April, 2018.
2. E.N. Feinberg, et al, V.S. Pande. "Spatial Graph Convolutions for Drug Discovery." arXiv. March, 2018. In Review.
3. E.N. Feinberg, et al, V.S. Pande. "Machine Learning Harnesses Molecular Dynamics to Discover New μ Opioid Chemotypes." arXiv. March, 2018. In Review.
4. E.N. Feinberg, et al, V.S. Pande. "Kinetic Machine Learning Unravels Ligand-Directed Conformational Change of μ Opioid Receptor." bioRxiv. July, 2017. In Review.
5. B. Yi, et al, E.N. Feinberg, V.S. Pande, M. Shamloo. "Discovery of novel brain permeable and G protein-biased beta-1 adrenergic receptor partial agonists for the treatment of neurocognitive disorders." PLoS One. July, 2017.
6. J. Gomes, B. Ramsundar, E.N. Feinberg, V.S. Pande. "Atomic Convolutional Networks for Predicting Protein-Ligand Binding Affinity." arXiv. March, 2017.
7. Z. Wu, B. Ramsundar, E.N. Feinberg, et al. "MoleculeNet: A Benchmark for Molecular Machine Learning." Chemical Science. March, 2017.
8. A. Manglik, W. Huang, A.J. Venkatakrishnan, T. Laeremans, E.N. Feinberg, et al, R.O. Dror, B.K. Kobilka. "Structural Insights into Mu-Opioid Receptor Activation." Nature. August 20, 2015.
9. J.S. Burg, J.R. Ingram, A.J. Venkatakrishnan, K.M. Jude, A. Dukkipati, E.N. Feinberg, et al, R.O. Dror, H.L. Ploegh, K.C. Garcia. "Structural basis for chemokine recognition and activation of a viral G-protein coupled receptor." Science. March 6, 2015.

PATENTS

1. E.N. Feinberg and V.S. Pande. "Machine Learning and Molecular Simulation Based Methods for Enhancing Binding and Activity Prediction" U.S. Provisional Patent Application No. 62/638,805.
2. E.N. Feinberg and V.S. Pande. "Spatial Graph Convolutions with Applications to Drug Discovery". U.S. Provisional Patent Application No. 62/638,803.
3. E.N. Feinberg and V.S. Pande. "Ligands of the Mu, Kappa, and Delta Opioid Receptors." U.S. Provisional patent application No. 62/585,796.

Acknowledgments

XStream GPU Cluster

The Stanford Research Computing Facility
Foundation Major Research Instrumentati
Abel (Physics/KIPAC), Margot Gerritsen (IC
grant were from the Schools of Earth, Ene
XStream is allocated to researchers acros

XStream Homepage »

Sherlock

Sherlock is a share
the SLURM resour
system for tempor

Faculty can supple

We acknowledge the generous support of Dr. Anders G. Frøseth and Mr. Christian Sundt for our work on machine learning. The Pande Group is broadly sup

Acknowledgments

Vijay S. Pande

Ron O. Dror

Brian Kobilka

Mehrdad Shamloo

Kevin Grimes

Acknowledgments