Machine learning of interatomic potentials

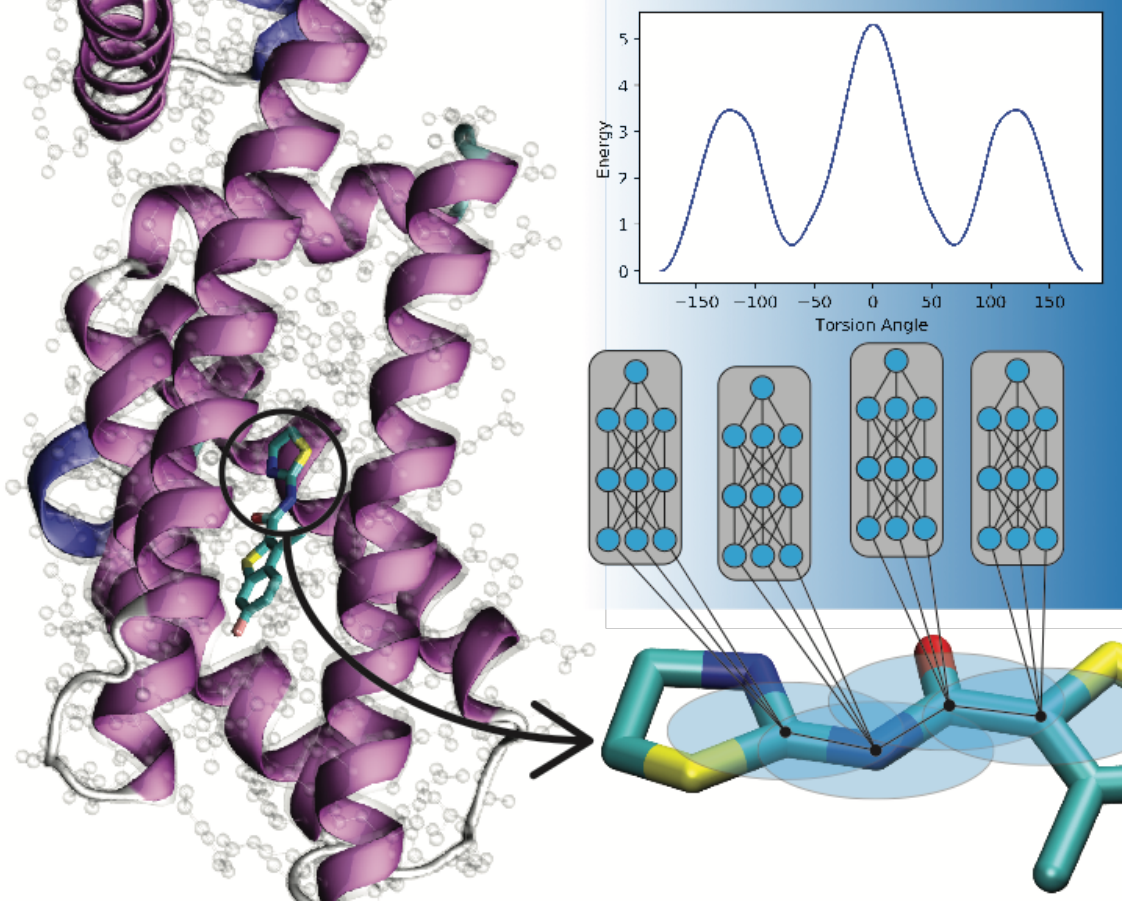
Justin S. Smith

Metropolis Postdoctoral Fellow
Los Alamos National Lab
(T-CNLS/T-1)

Los Alamos National Lab:
Kipton Barros
Sergei Tretiak
Benjamin Nebgen
Nicholas Lubbers
Saryu Fensin
Timothy Germann

UNC Chapel Hill:
Roman Zubatyuk
Olexandr Isayev

University of Florida:
Adrian Roitberg



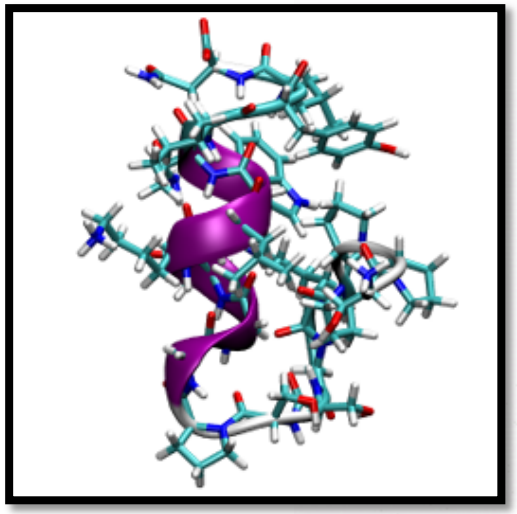
Los Alamos NATIONAL LABORATORY

LA-UR-19-30185

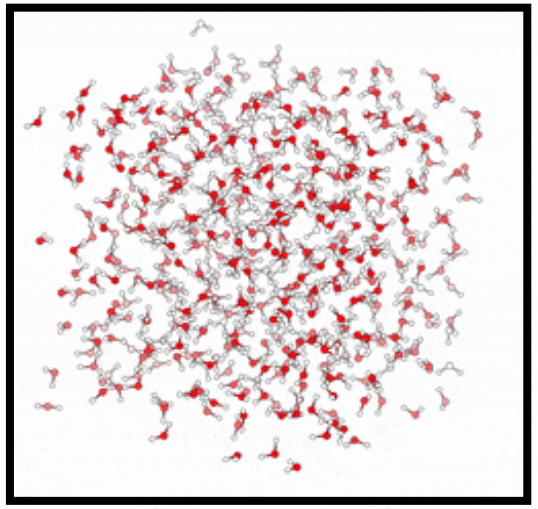
Center for Nonlinear Studies

Molecular (atomistic) dynamics

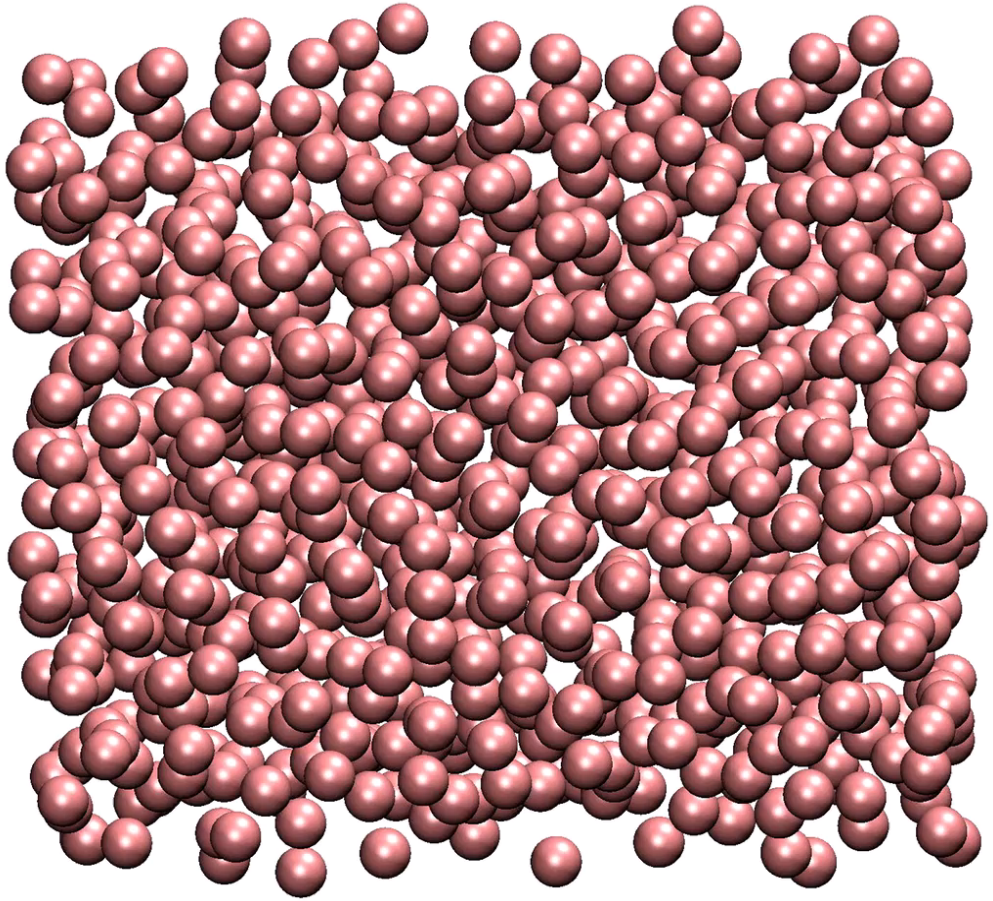
Proteins



Liquids



Materials



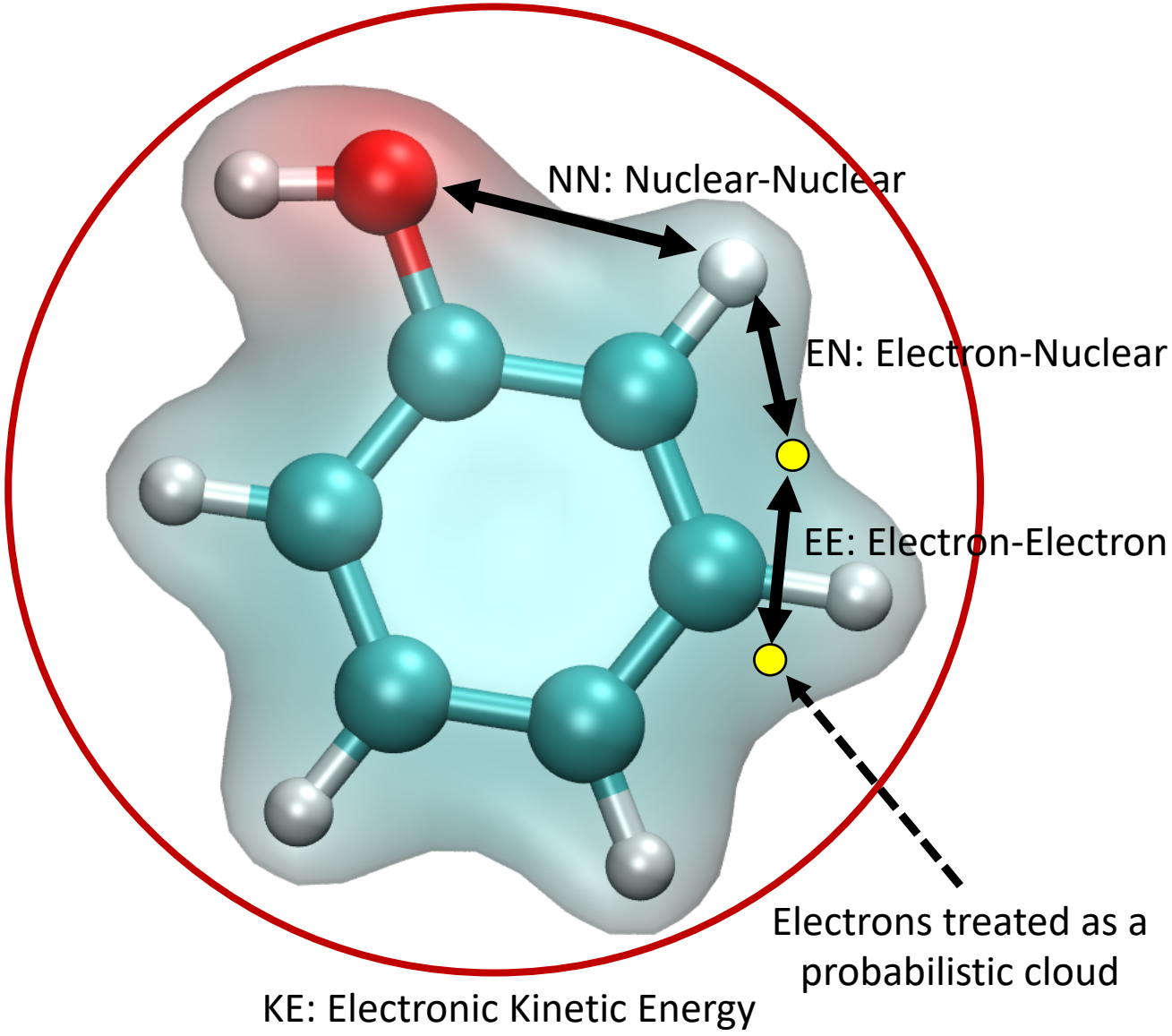
Energy $E[\mathbf{r}_1, \mathbf{r}_2, \ldots]$

Force $\mathbf{f}_i = -\nabla_i E$

In principle, requires a quantum mechanical calculation at *each* time step!

Dynamics $m\dfrac{d^2\mathbf{r}_i}{dt^2} = \mathbf{f}_i$

Ab initio Quantum Mechanics (QM)



Time-independent Schrödinger equation

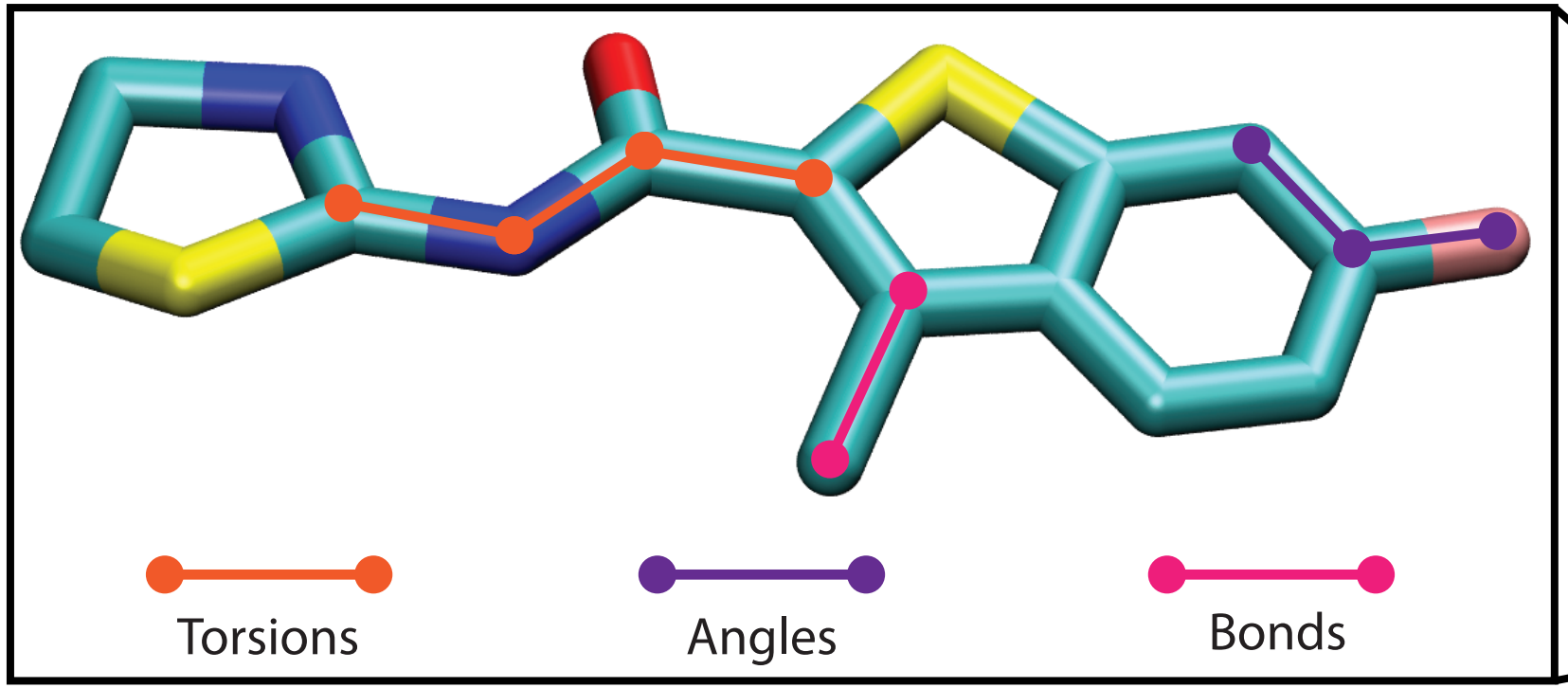
$$\hat{H}|\psi\rangle = E|\psi\rangle$$

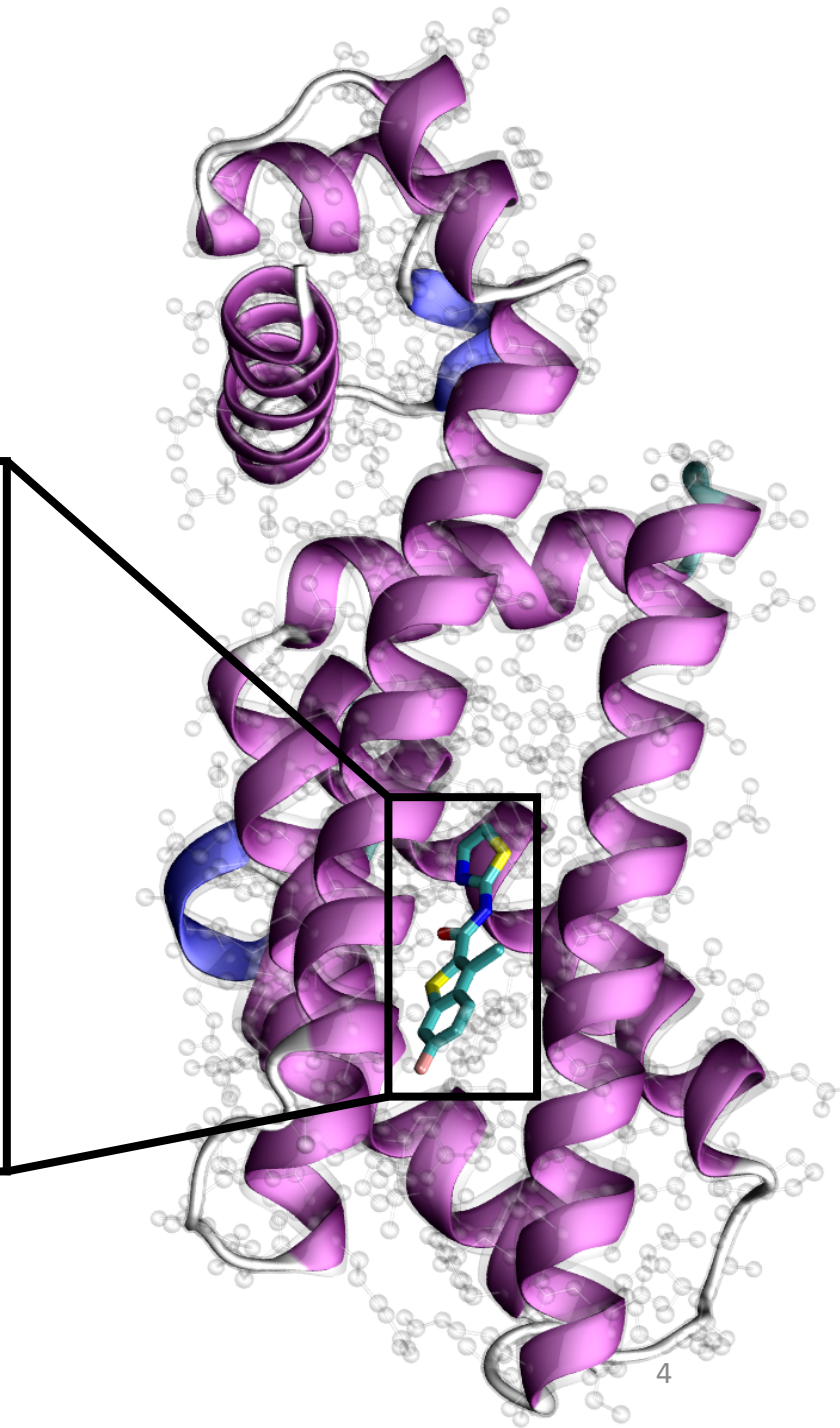
$$\hat{H} = KE + ER + \text{EN} + \text{NN}$$

Class	Method	Cost
Semi-empirical	AM1, PM6, DFTB	$O(N^2)$
Density functional theory	B3LYP, wB97x, PBE	$O(N^3)$
Post-Hartree Fock	MP2, Coupled Cluster	$O(N^4) >$

Classical force fields for Chemistry

$$E_{FF} = \sum E_{ij}^{\text{nonbonded}} + \sum E_{ijkl}^{\text{Torsions}} + \sum E_{ijk}^{\text{Angles}} + \sum E_{ij}^{\text{Bonds}}$$



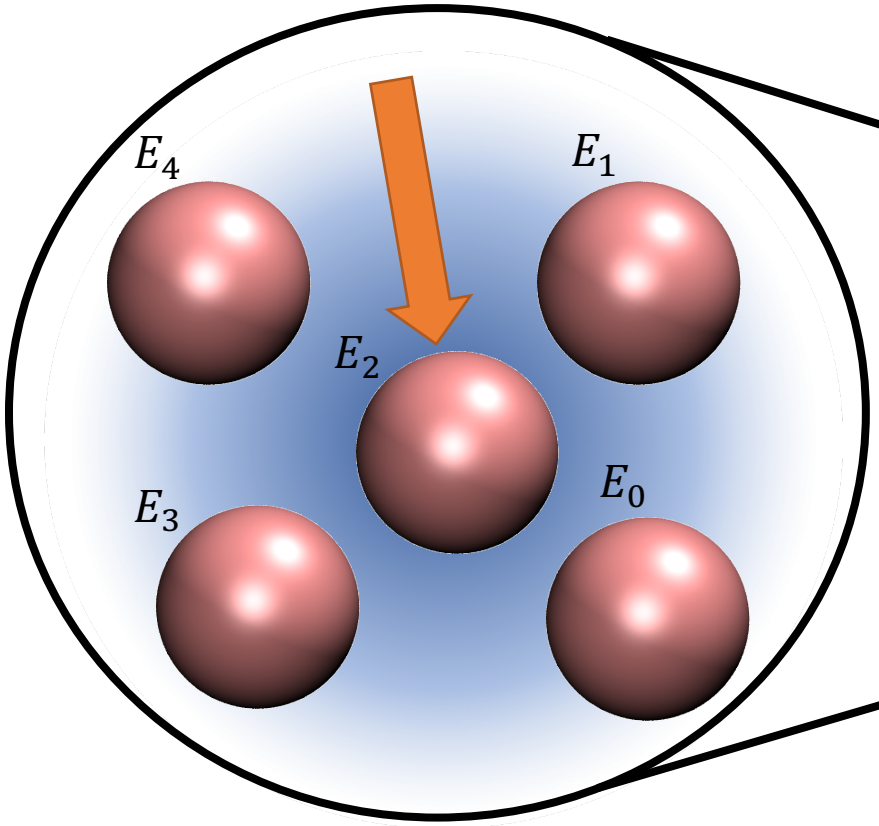


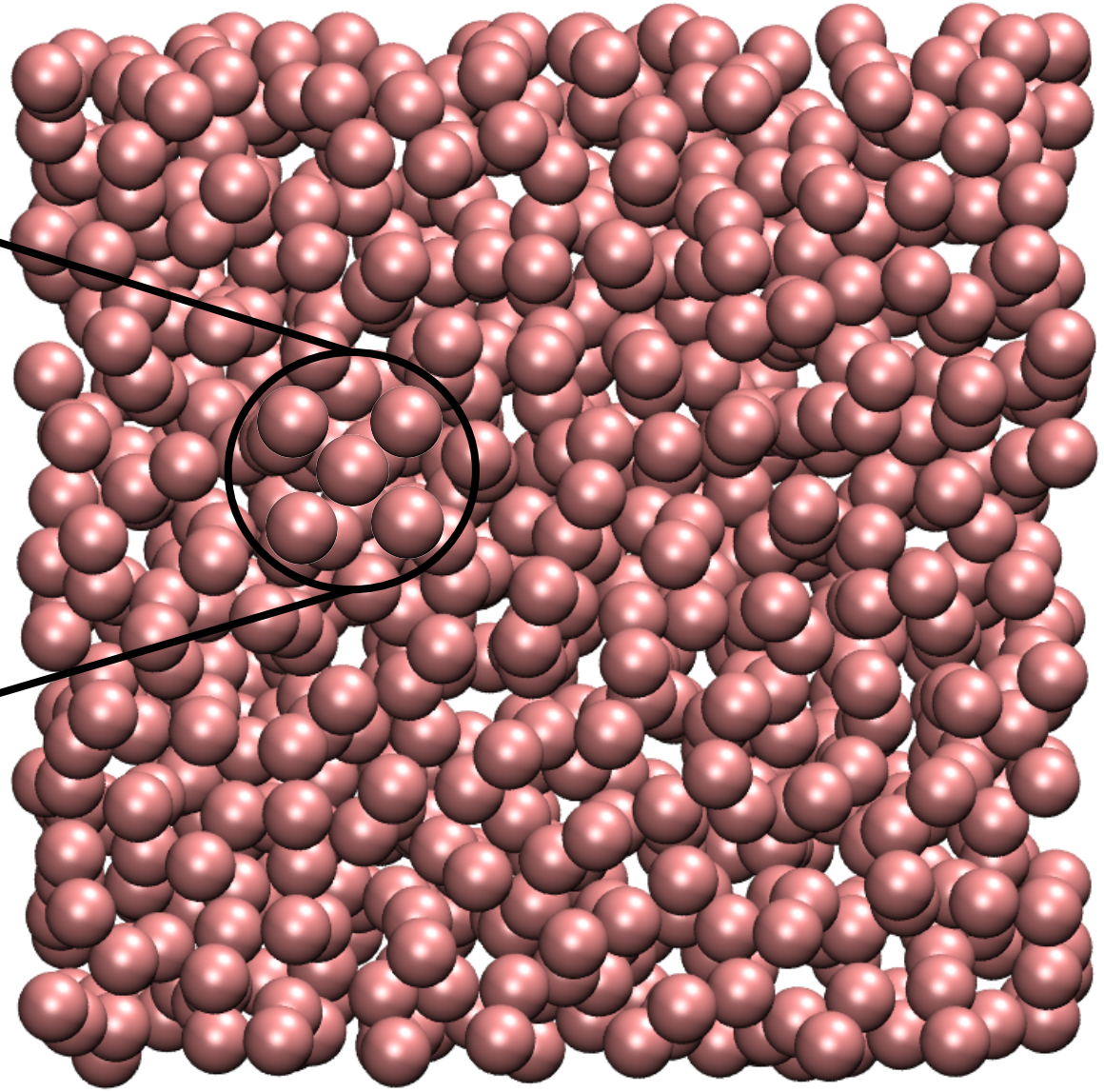
Classical Force Fields (FF):

CHARMM – OPLS – AMBER

Empirical potentials for materials

E_i represents the energy of embedding atom i within the electron field of its neighbors



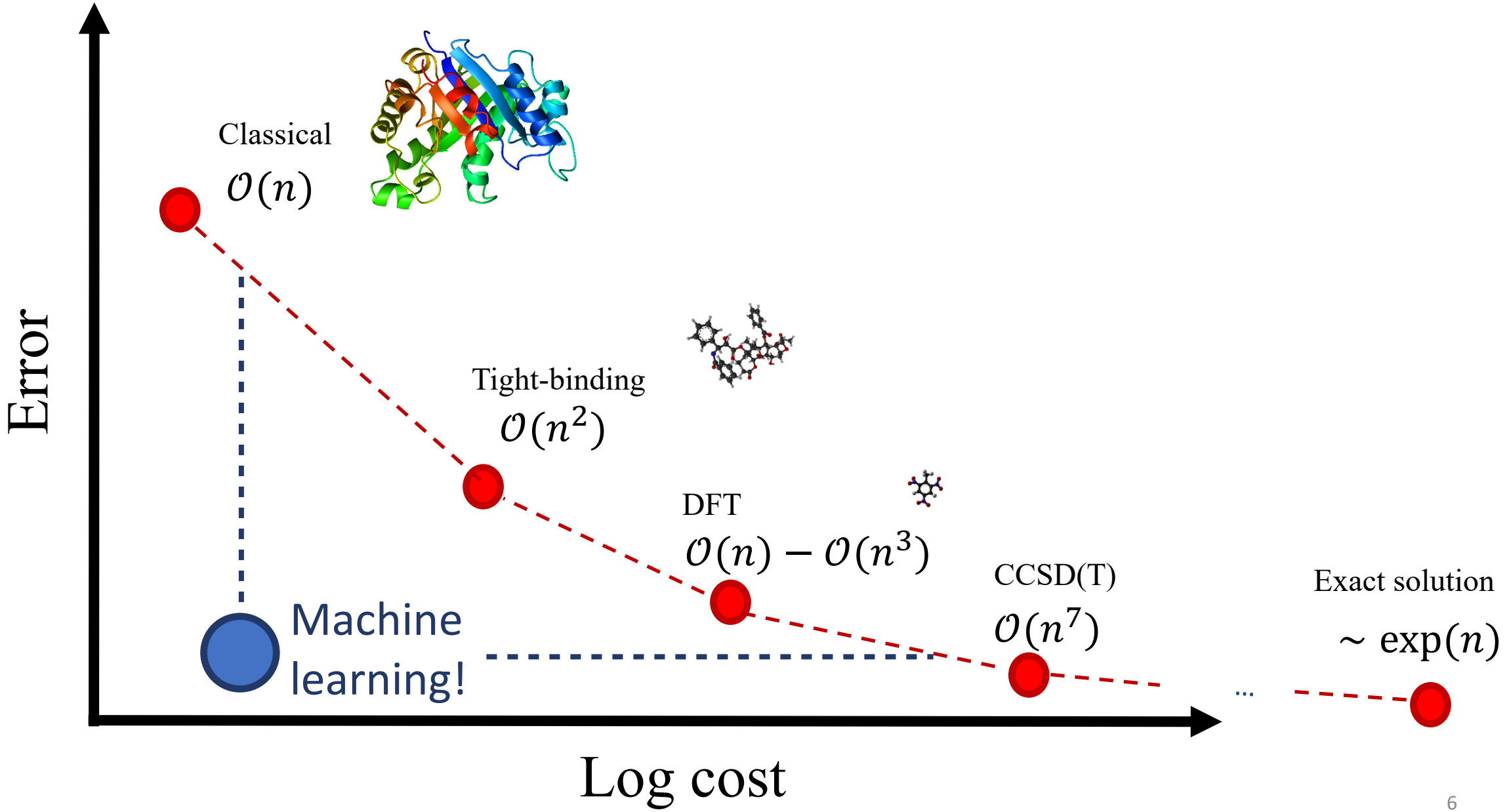


$$E = \sum_{i=0}^{N_{atoms}} F_\alpha \left(\sum_{i \neq j} \rho_\beta (r_{ij}) \right) + \frac{1}{2} \sum_{i \neq j} \phi_{\alpha\beta}(r_{ij})$$

Embedded Atom Model (EAM)

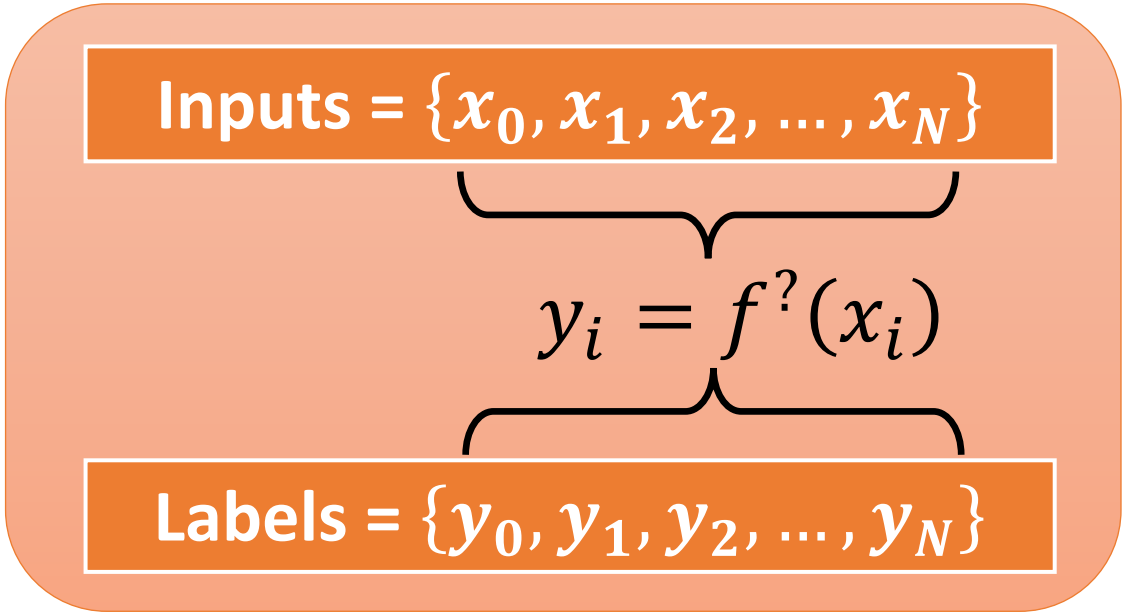
M.S. Daw and M.I. Baskes, Phys. Rev. Letters 50 (1983) 1285.
M.S. Daw and M.I. Baskes, Phys. Rev. B 29 (1984) 6443.

Levels of model chemistry

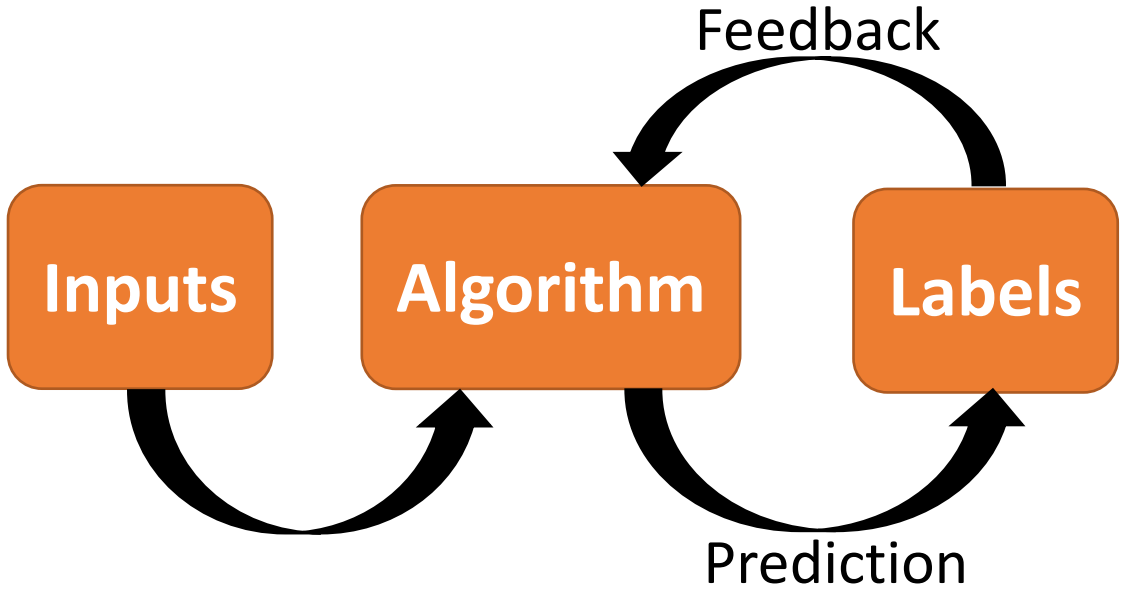


Supervised Machine Learning

Training dataset for supervised learning



Supervised Learning



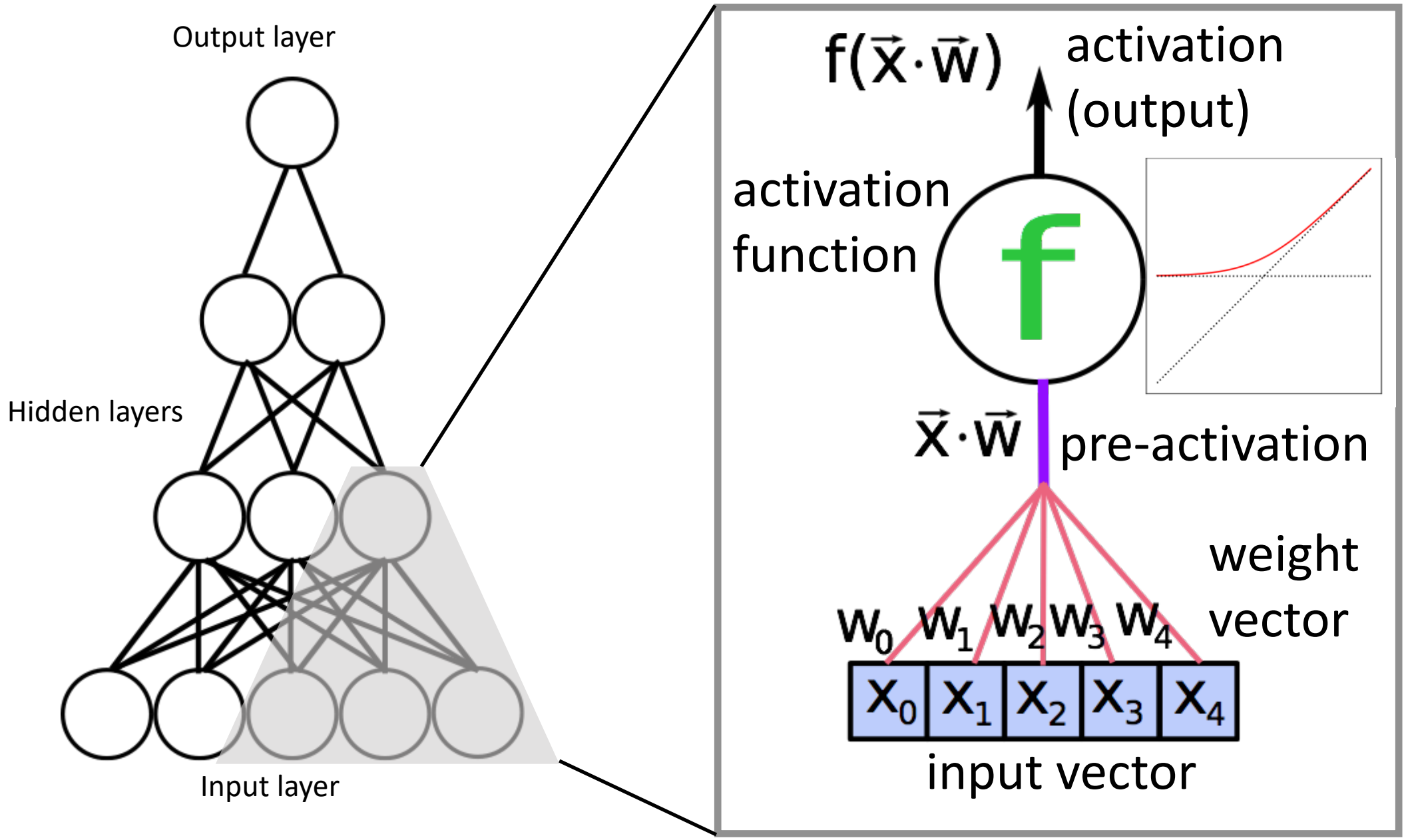
Types of tasks

- Regression
- Classification

A few applications

- Image recognition
- Social media moderation
- Stock market prediction

Deep learning (neural networks) in a nutshell



Deep Learning for Atomistic Potentials

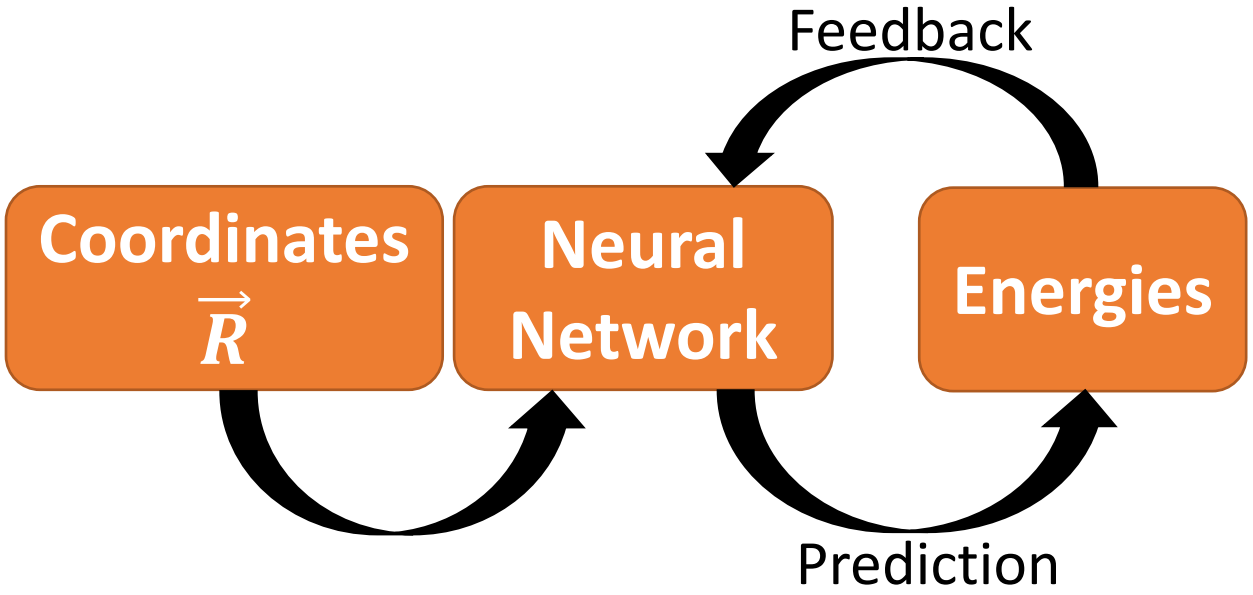
Training dataset for supervised learning

Inputs = $\{\vec{R}_0, \vec{R}_1, \vec{R}_2, \ldots, \vec{R}_N\}$

$$E_i = f^{QM}(\vec{r}_i)$$

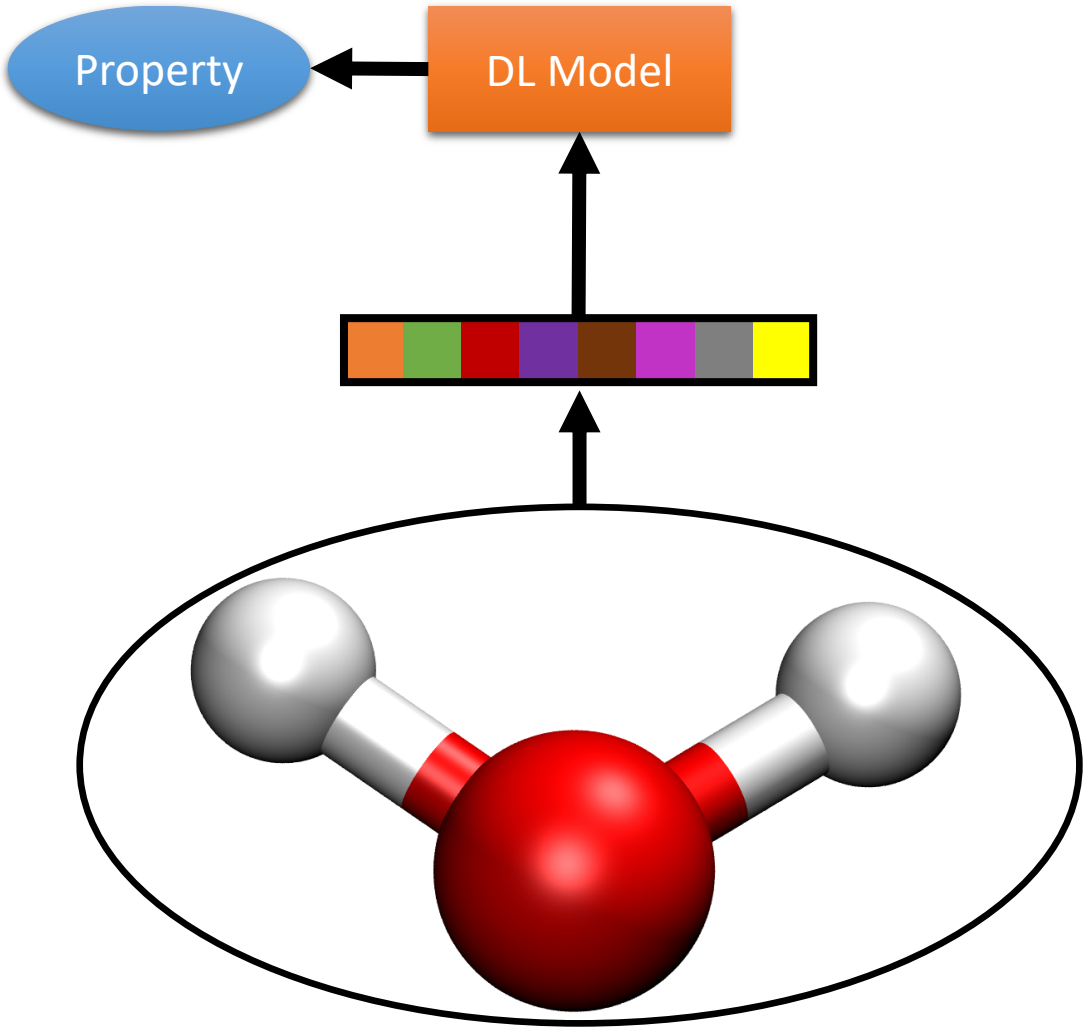
Labels = $\{E_0, E_1, E_2, \ldots, E_N\}$

Supervised Learning

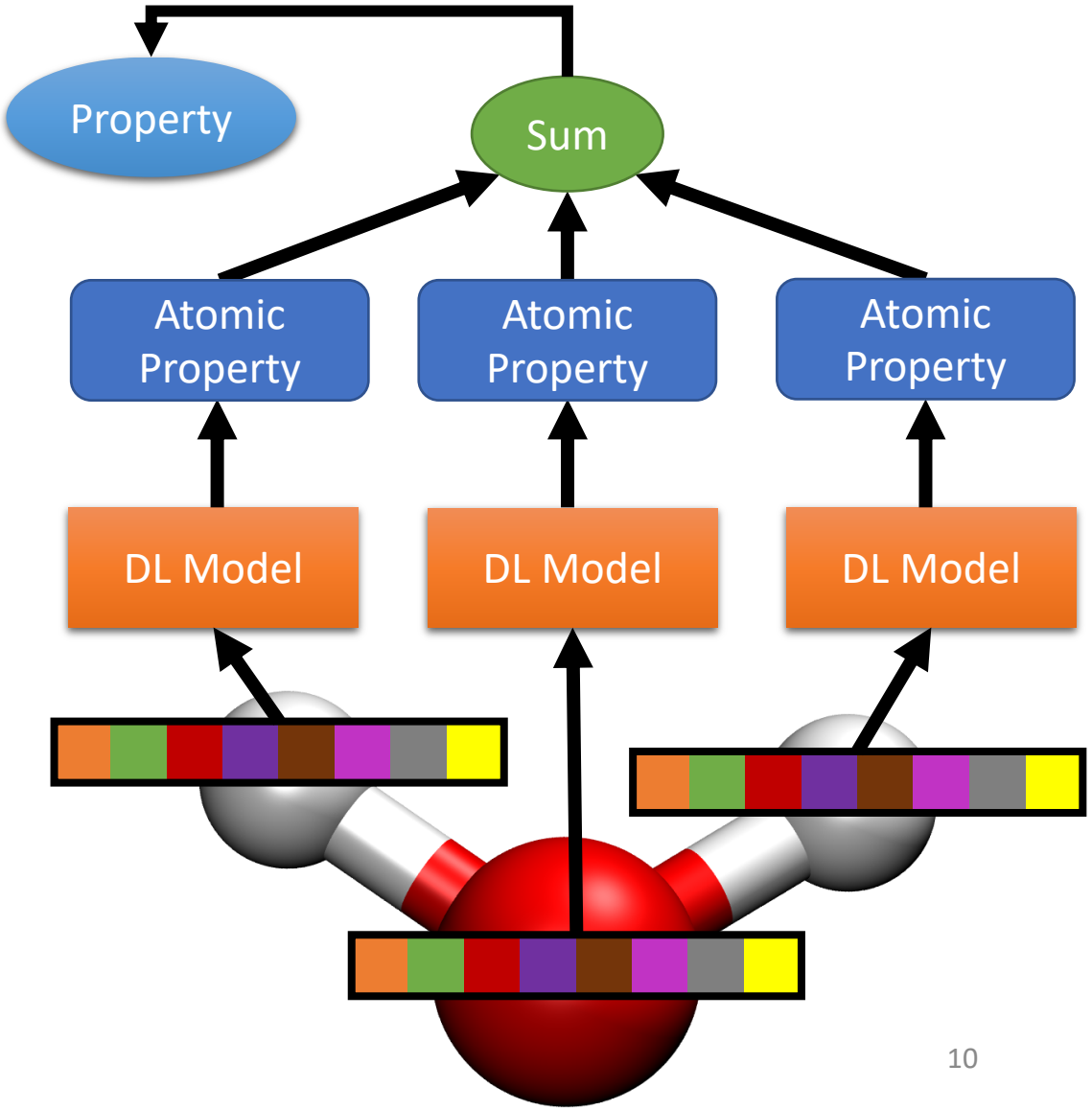


What are deep learning potentials?

Molecular Fingerprints

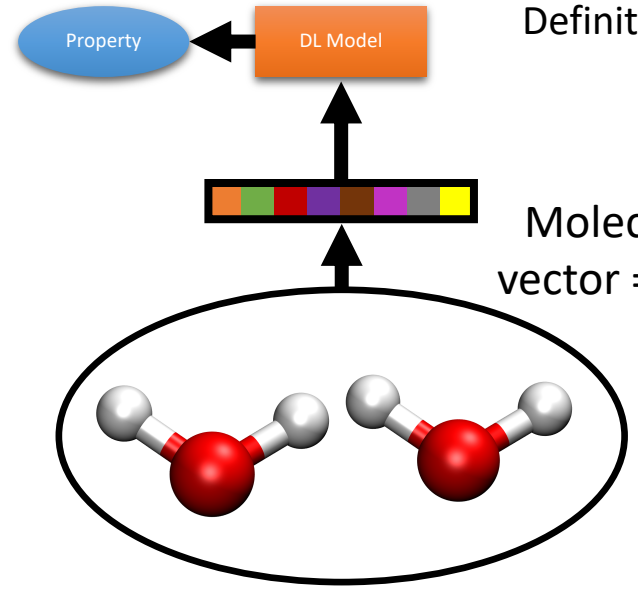


Atomic Environment Descriptors

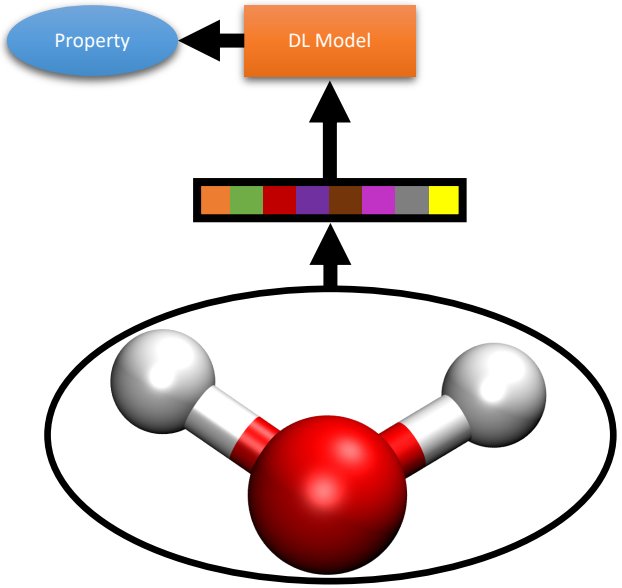


Transferability and extensibility for DL potentials

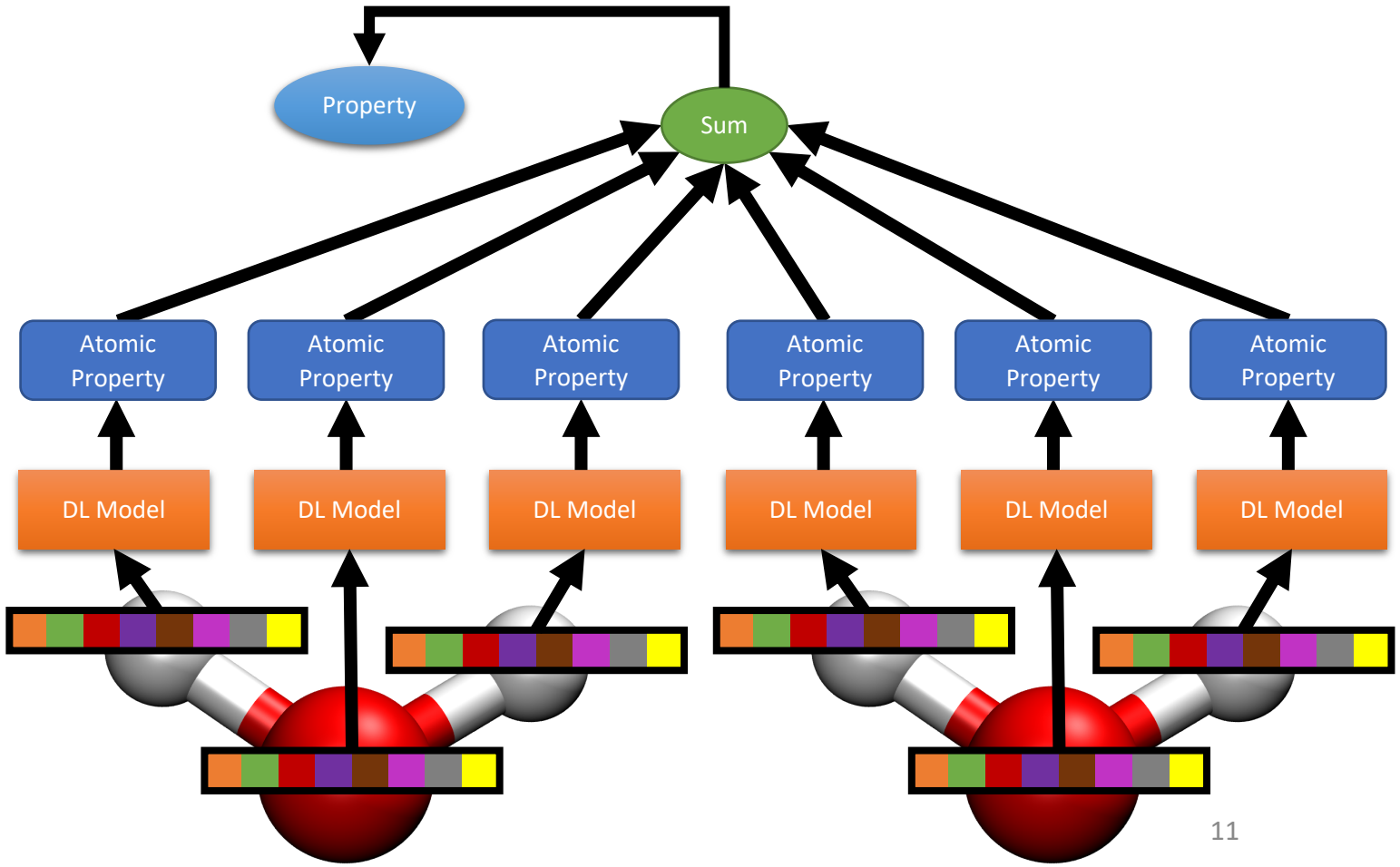
Definition: **Transferability** – The ability for the model to predict on systems *not* in the training dataset
Definition: **Extensibility** – The ability for a model to predict on systems *larger* than those in the training dataset



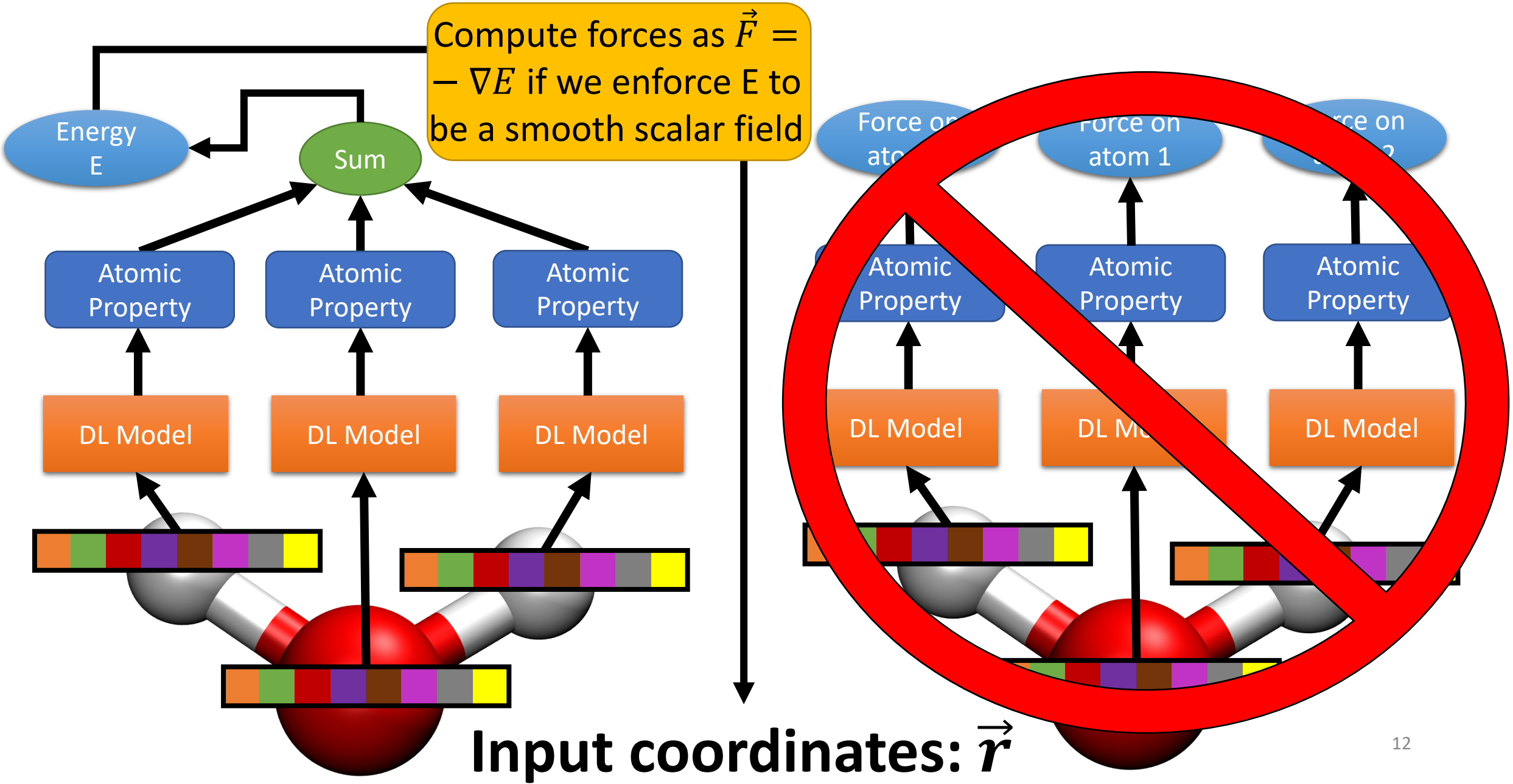
Molecular fixed sized vector = no extensibility



Atomic fixed sized vector = extensibility



Building an energy conservative DL potential: how to get atomic forces?



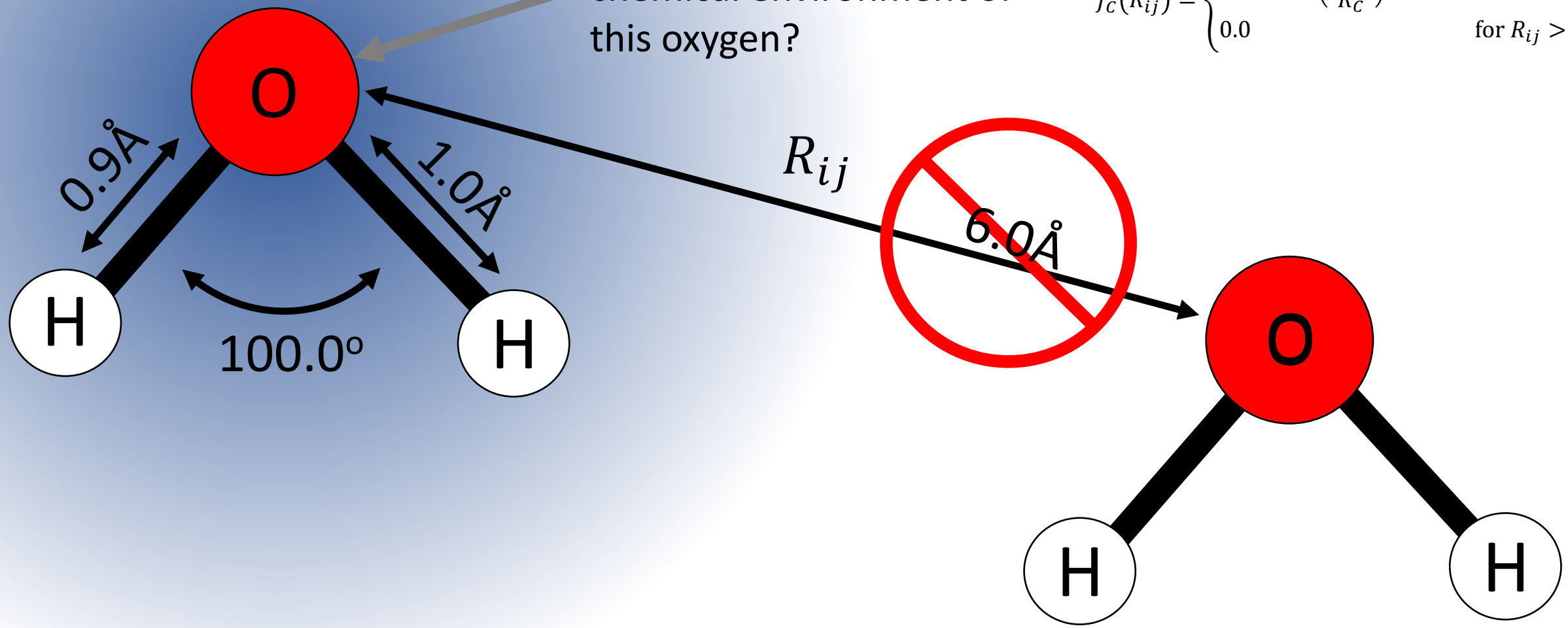
Input coordinates: $\vec{r}$

Atomic environment description

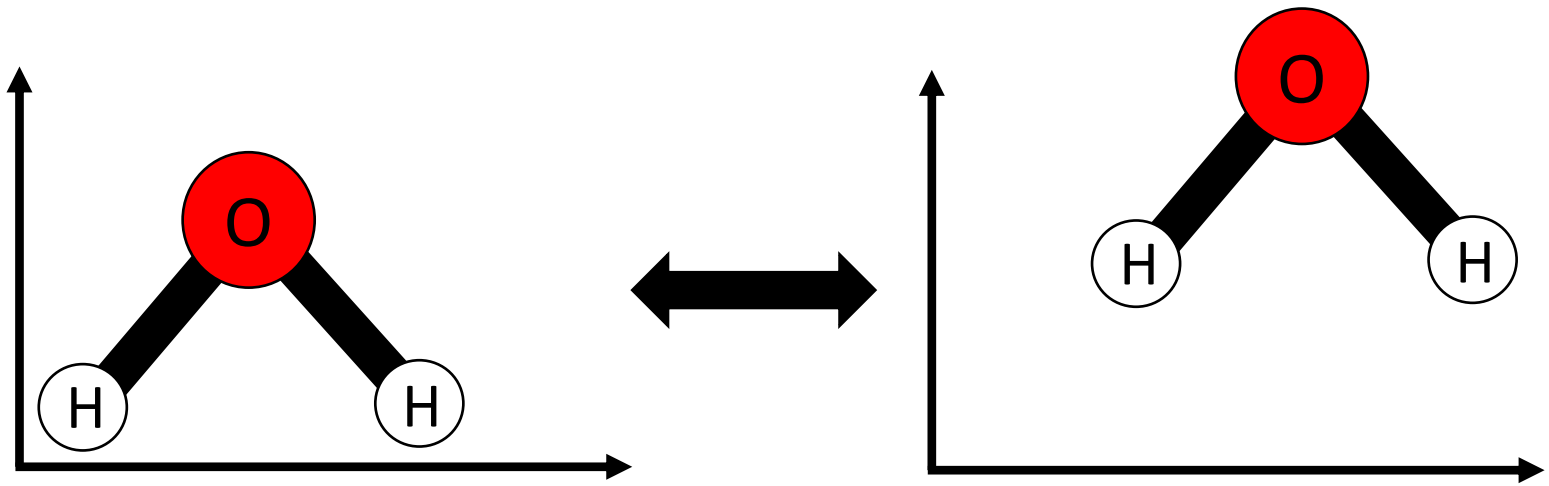
How do we represent the chemical environment of this oxygen?

Cutoff function

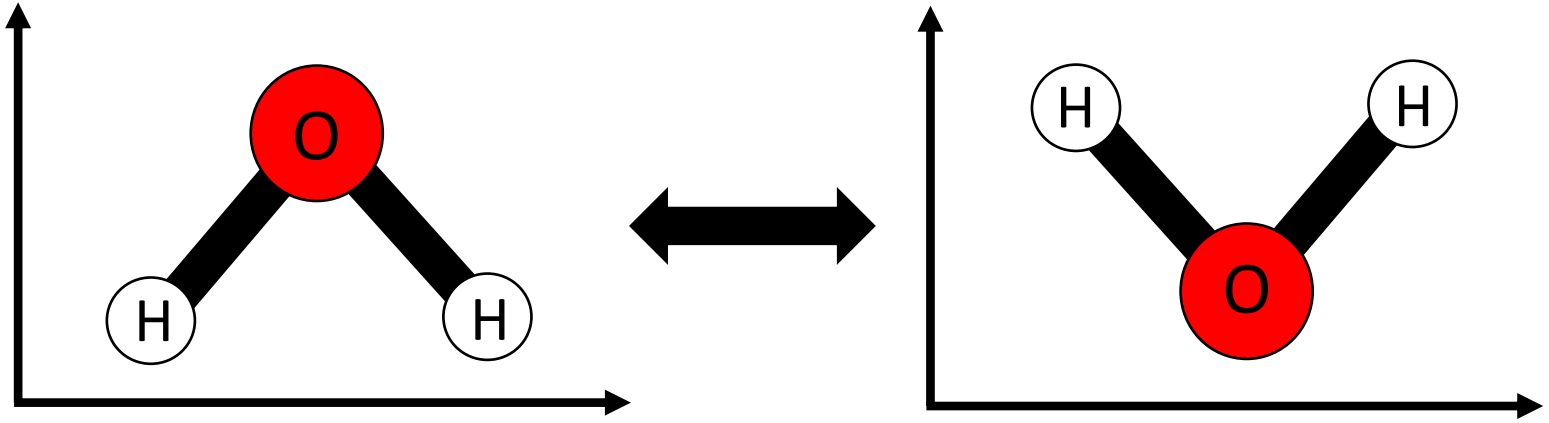
$$f_C(R_{ij}) = \begin{cases} 0.5 \times \cos\left(\frac{\pi R_{ij}}{R_C}\right) + 0.5 & \text{for } R_{ij} \leq R_C \\ 0.0 & \text{for } R_{ij} > R_C \end{cases}$$



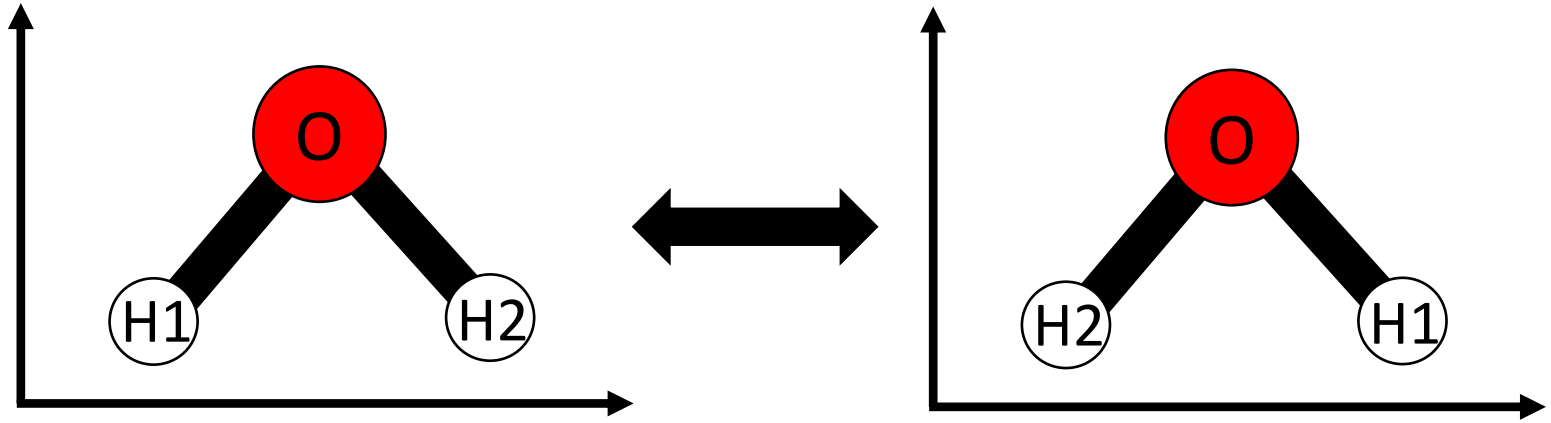
We choose physically inspired invariance:



Translation



Rotation



Permutation

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

ANI, the first ever general-purpose machine learning potential for molecules.

CrossMark click for updates

Cite this: *Chem. Sci.*, 2017, **8**, 3192

ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost†

J. S. Smith,[a] O. Isayev*[b] and A. E. Roitberg*[a]

HIP-NN, a first generation graph convolution based neural network for molecular potentials.

Hierarchical modeling of molecular energies using a deep neural network

Cite as: J. Chem. Phys. **148**, 241715 (2018); https://doi.org/10.1063/1.5011181
Submitted: 30 October 2017 . Accepted: 31 January 2018 . Published Online: 19 March 2018

Nicholas Lubbers, Justin S. Smith, and Kipton Barros

SCIENCE ADVANCES | RESEARCH ARTICLE

CHEMISTRY

Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network

Roman Zubatyuk[1,2,3], Justin S. Smith[2,4], Jerzy Leszczynski[3], Olexandr Isayev[1]*

AIM-Net, a multitask model for QM property prediction. In print at Science Advances.

Common invariant components of ML model potentials

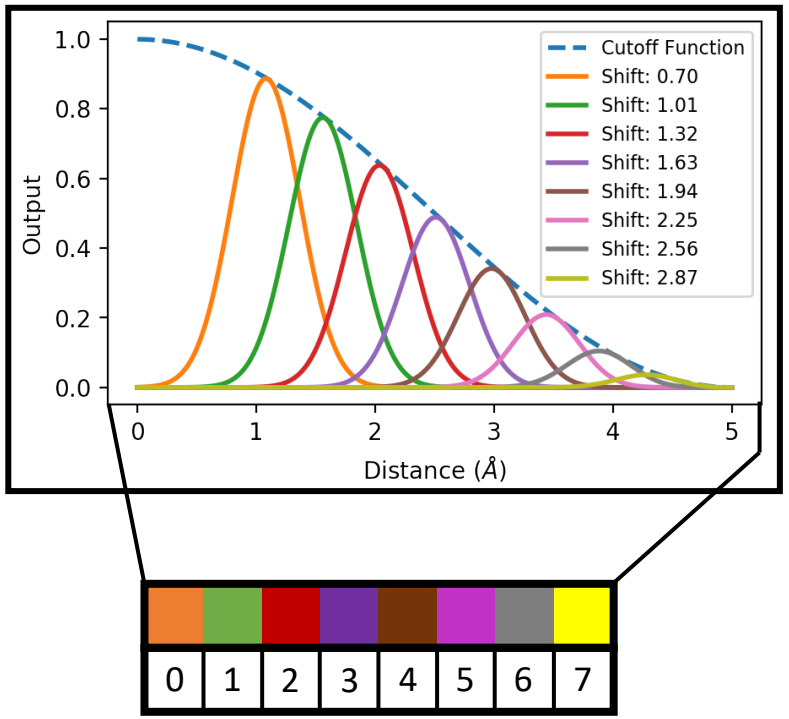
Radial symmetry functions

J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401

$$G_m^R = \sum_{j \neq i}^{\text{All Atoms}} e^{-\eta\left(R_{ij}-R_S\right)^2} f_C\left(R_{ij}\right)$$

N Lubbers, JS Smith, K Barros *J. Chem. Phys.*, 2018, **148**, 241715

$$G_m^R = \sum_{j \neq i}^{\text{All Atoms}} e^{-\left(R_{ij}^{-1}-\mu^{-1}\right)^2/2\sigma^{-2}} f_C\left(R_{ij}\right)$$



R_{ij}
Distance between atoms *i* and *j*

θ_{ijk}
Angle between atoms *j* and *k* centered on atom *i*

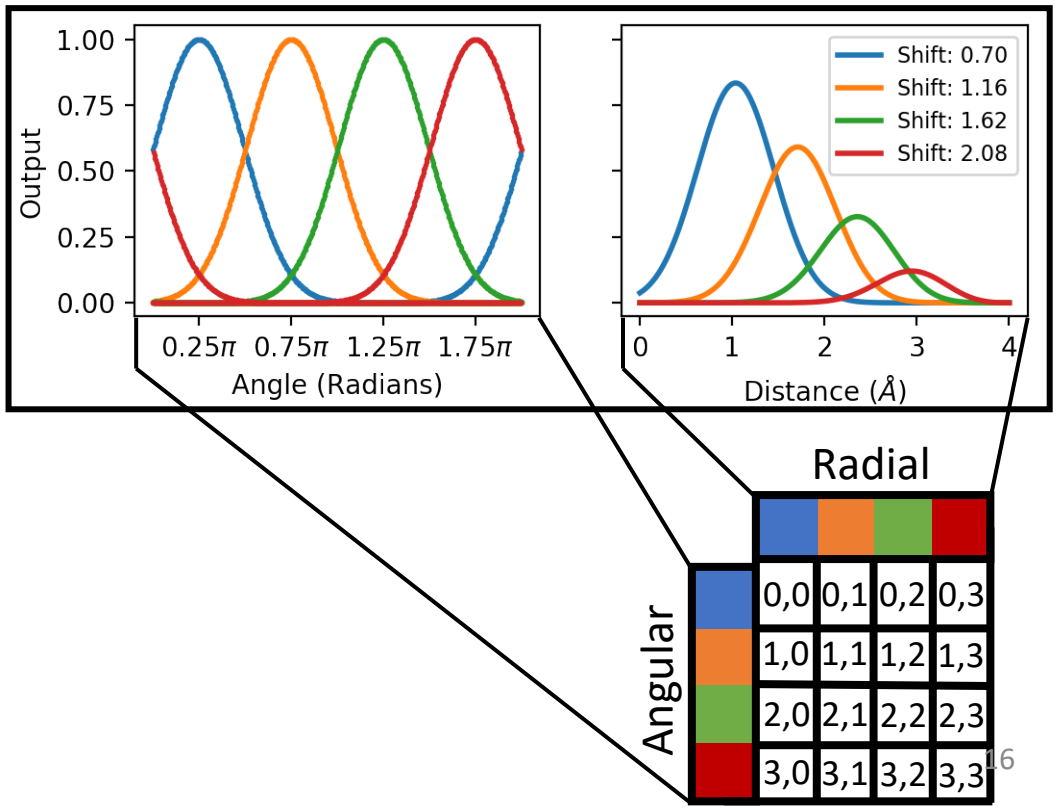
Angular symmetry functions

J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401

$$G_m^{A_{\text{mod}}} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{All Atoms}} \left(1+\lambda\cos\left(\theta_{ijk}\right)\right)^{\zeta} exp\left[-\eta\left(R_{ij}^2+R_{ik}^2+R_{jk}^2\right)^2\right] f_C\left(R_{ij}\right) f_C\left(R_{ik}\right) f_C\left(R_{jk}\right)$$

JS Smith, O Isayev, AE Roitberg, *Chem. Sci.*, 2017, **8**, 3192-3203

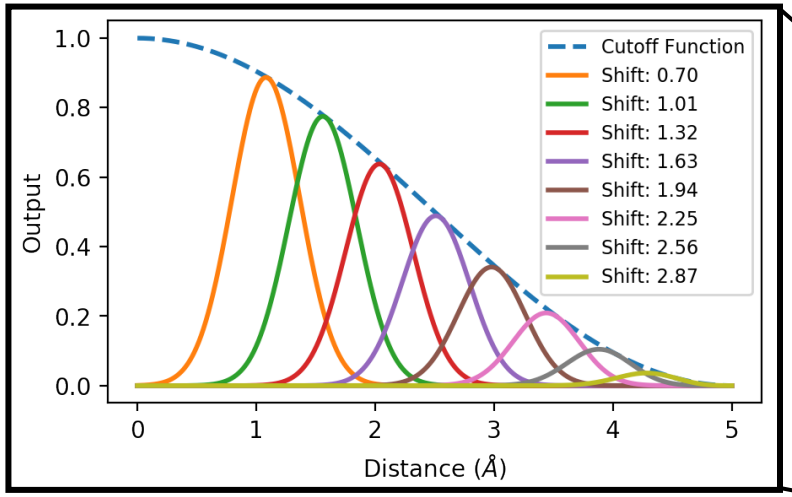
$$G_m^{A_{\text{mod}}} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{All Atoms}} \left(1+\cos\left(\theta_{ijk}-\theta_s\right)\right)^{\zeta} exp\left[-\eta\left(\frac{R_{ij}^2+R_{ik}^2}{2}-R_S\right)^2\right] f_C\left(R_{ij}\right) f_C\left(R_{ik}\right)$$



ANI style neural network model potential

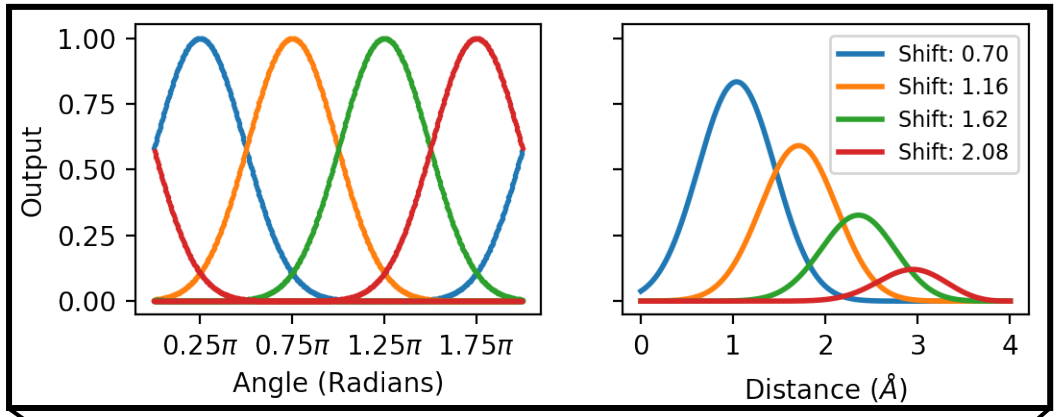
Radial environment

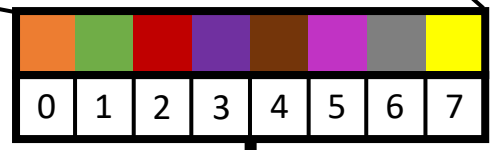
Distance over j neighbors centered on atom i

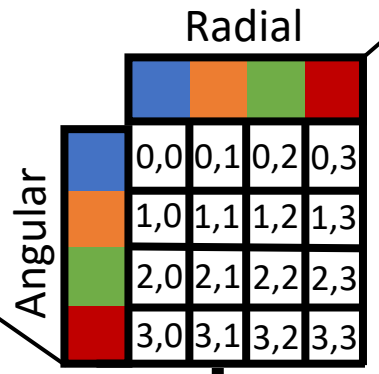


Angular environment

Angle over j,k neighbors centered on atom i

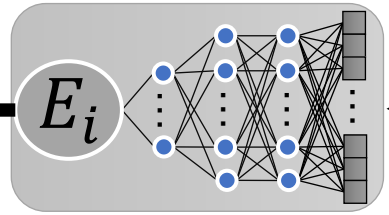






$$E = \sum_{i}^{N} E_i$$

Atomic Neural Network



Sum over j neighbors

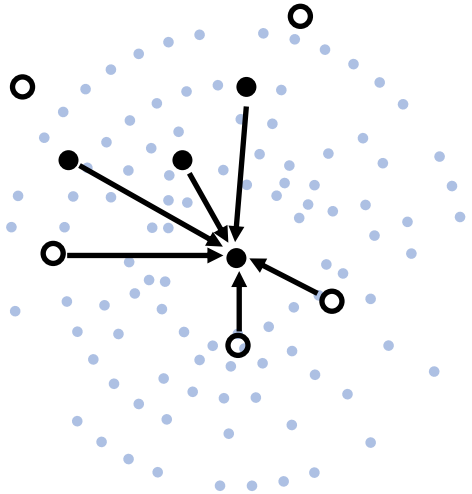
Z_j and Z_k type differentiation

Concatenate

JS Smith, O Isayev, AE Roitberg, *Chem. Sci.*, 2017, **8**, 3192-3203

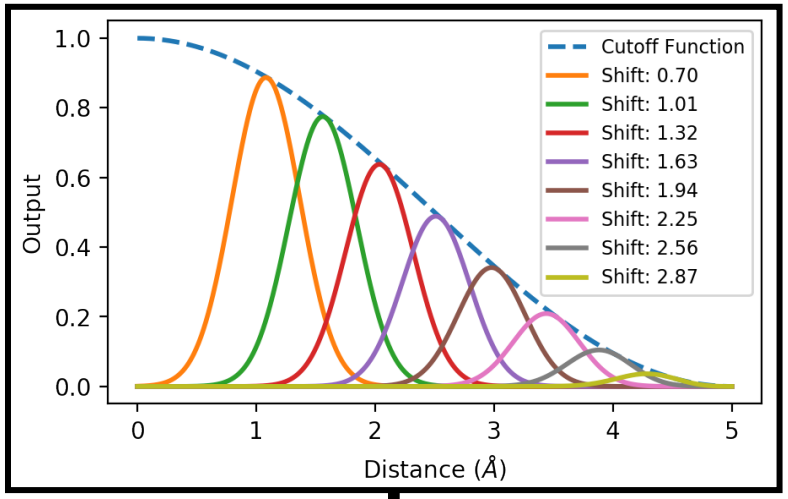
HIP-NN style neural network model potential

Interaction is a message passed from neighboring atoms



Radial environment

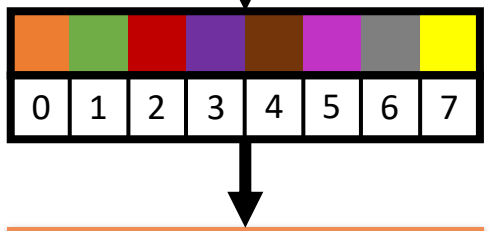
Distance over j neighbors centered on atom i



$$E = \sum_i^N E_i$$

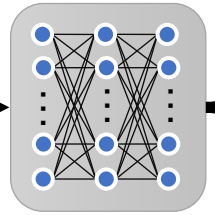
One-hot

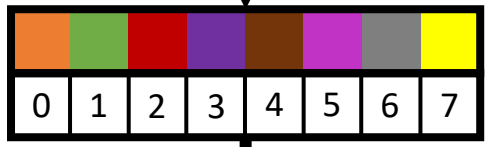
H	0
C	0
N	1
O	0



Interaction over j neighbors

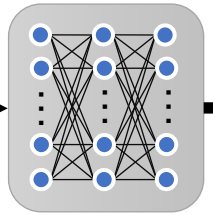
Atomic Neural Network





Interaction over j neighbors

Atomic Neural Network



E_0 + E_1 + $E_2 = E_i$

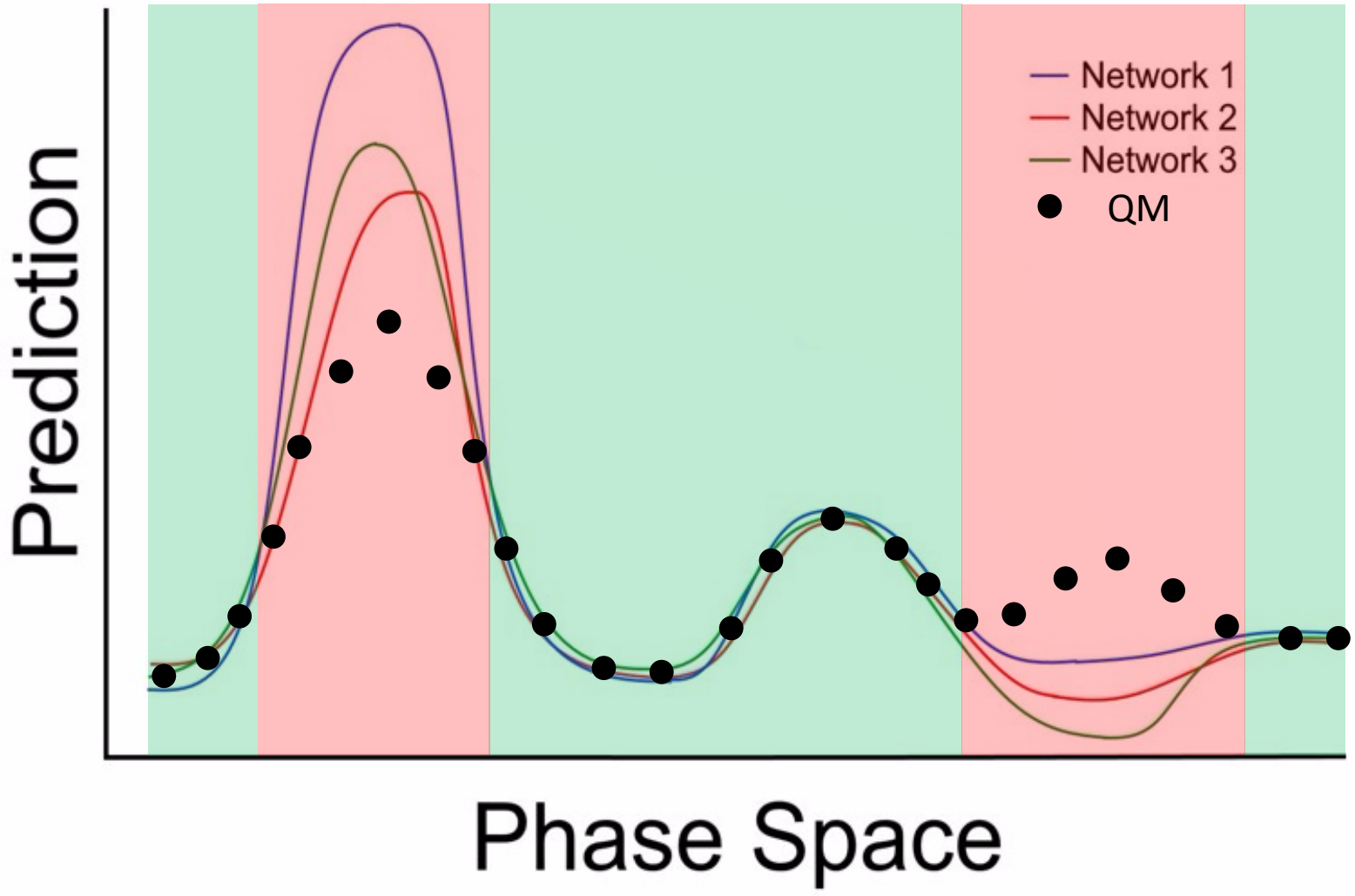
N Lubbers, JS Smith, K Barros; *J. Chem. Phys.*, 2018, **148**, 241715

Can we predict when the model is wrong?

Ensemble disagreement can drive data generation

Good data coverage

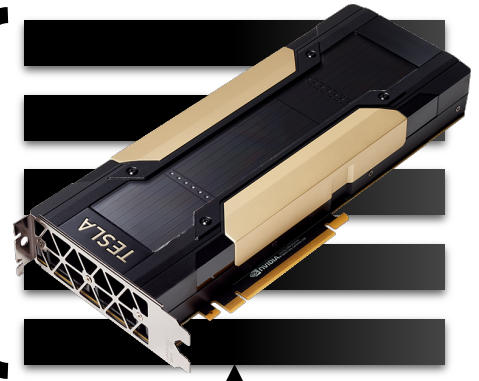
Bad data coverage



Active Learning - The Big Picture

An automated and self-consistent data generation framework

Ensemble of ANI networks



Non-equilibrium Conformational sampler

New test data

Check ensemble disagreement

Train network ensemble

Molecule Sampling

(e.g. GDB small molecule database, small peptides, drug like molecules)

Structure Pools

Compute Cluster

Computations with QM

ANI-1x Dataset (i.e. energies, forces, dipoles)

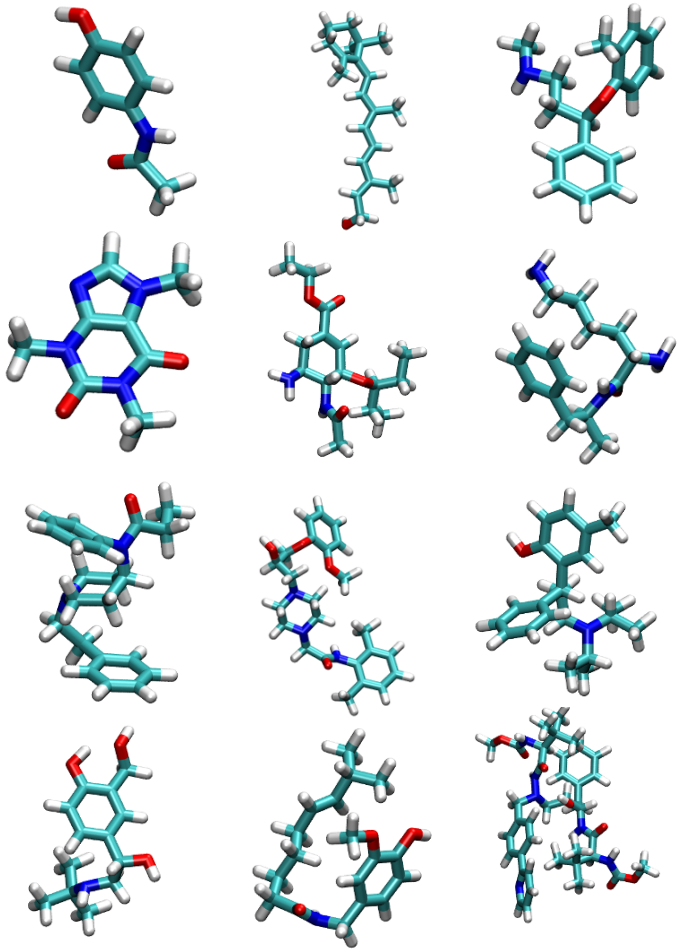
JS Smith, et al.; *The Journal of Chemical Physics*, (**2018**), 148 (24), 241733

Testing transferability and extensibility

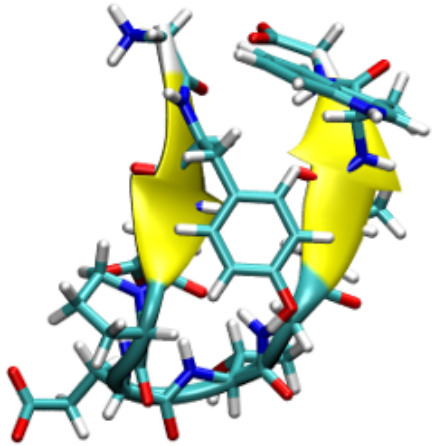
ANI-MD Benchmark

128 frames from 1ns trajectories @ 300K for each:

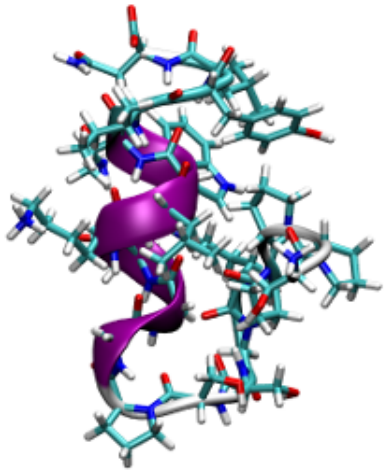
Various drug molecule



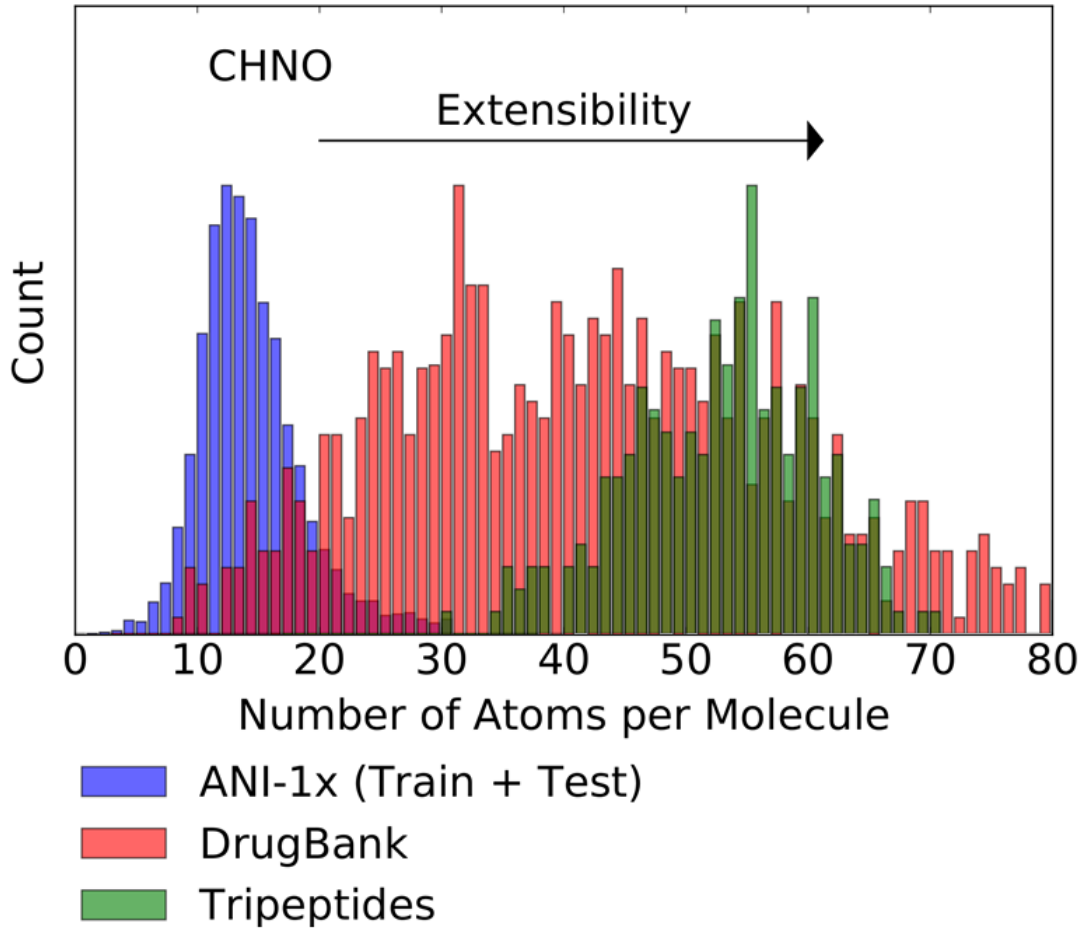
Chignolin



Trp-cage



DrugBank and Tripeptide Benchmarks



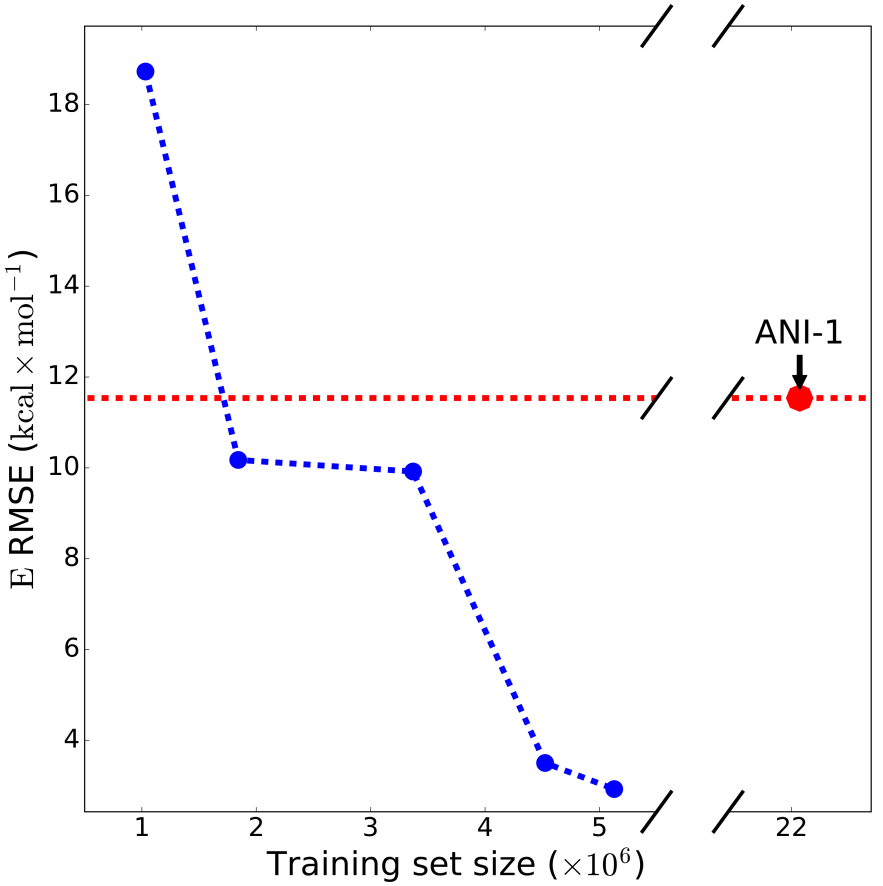
JS Smith, et al.; *The Journal of Chemical Physics*, (**2018**), 148 (24), 241733

Active-learning results vs. random sampling

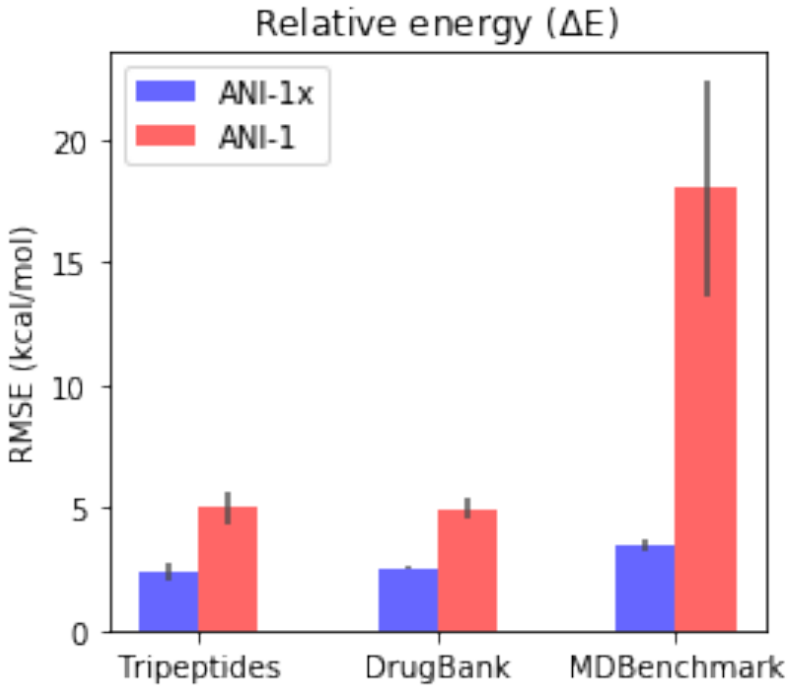
Dataset size comparison

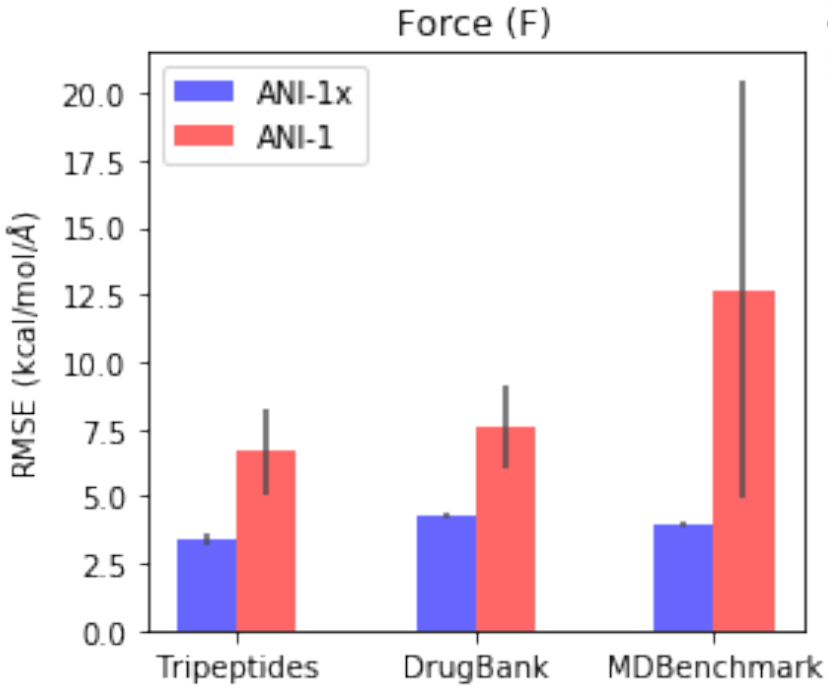
	ANI-1	ANI-1x
Datapoints	22M	5M

Active learning progression



Relative E and F RMSE comparison

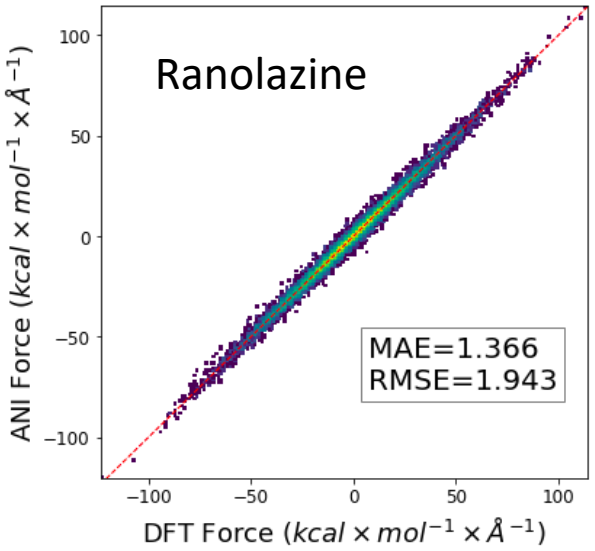


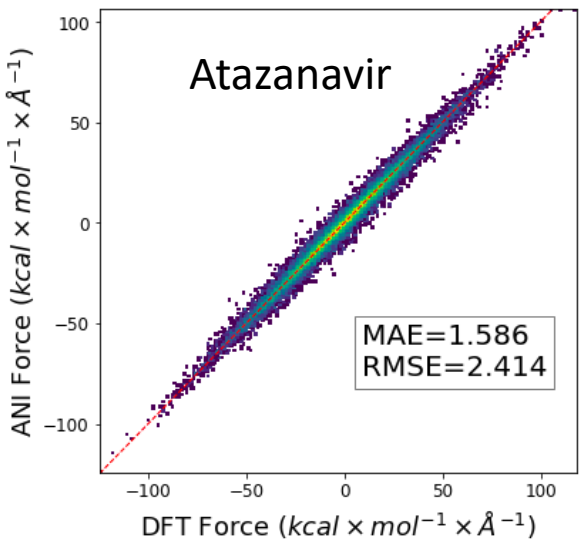


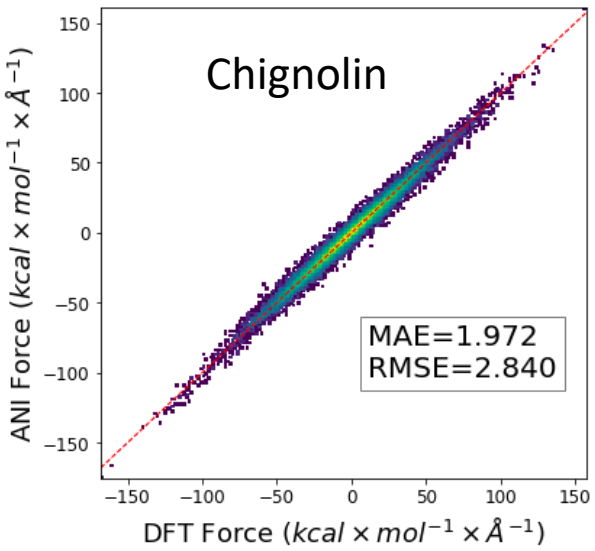
JS Smith, et al.; *The Journal of Chemical Physics*, (**2018**), 148 (24), 241733

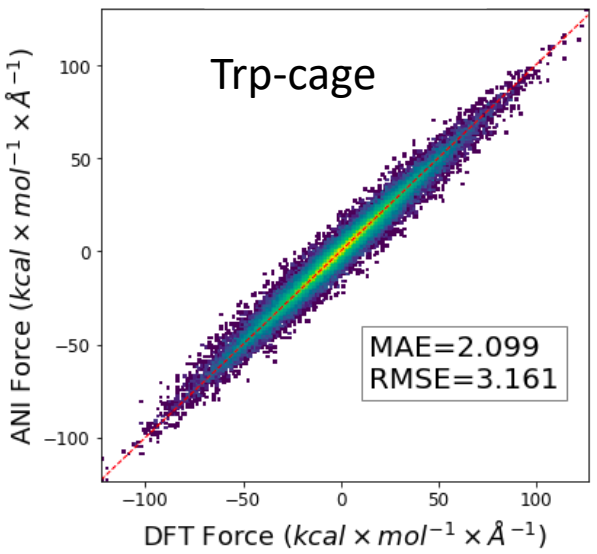
Molecular dynamics force errors vs. reference DFT

Correlation of force components for the ANI-MD Benchmark

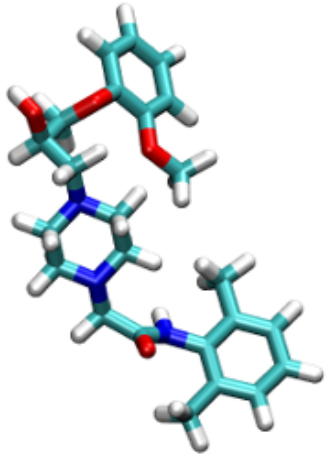




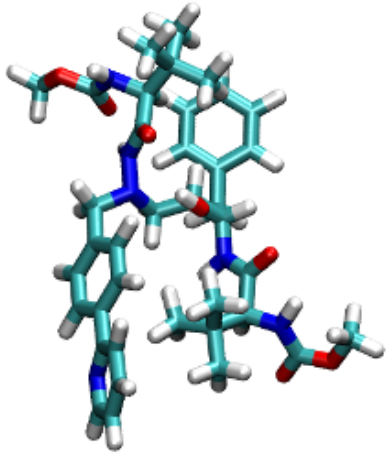




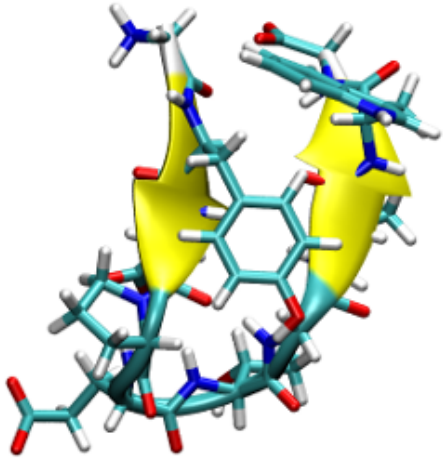
64 Atoms



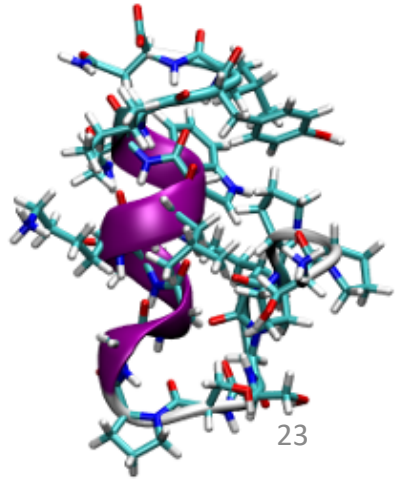
103 Atoms



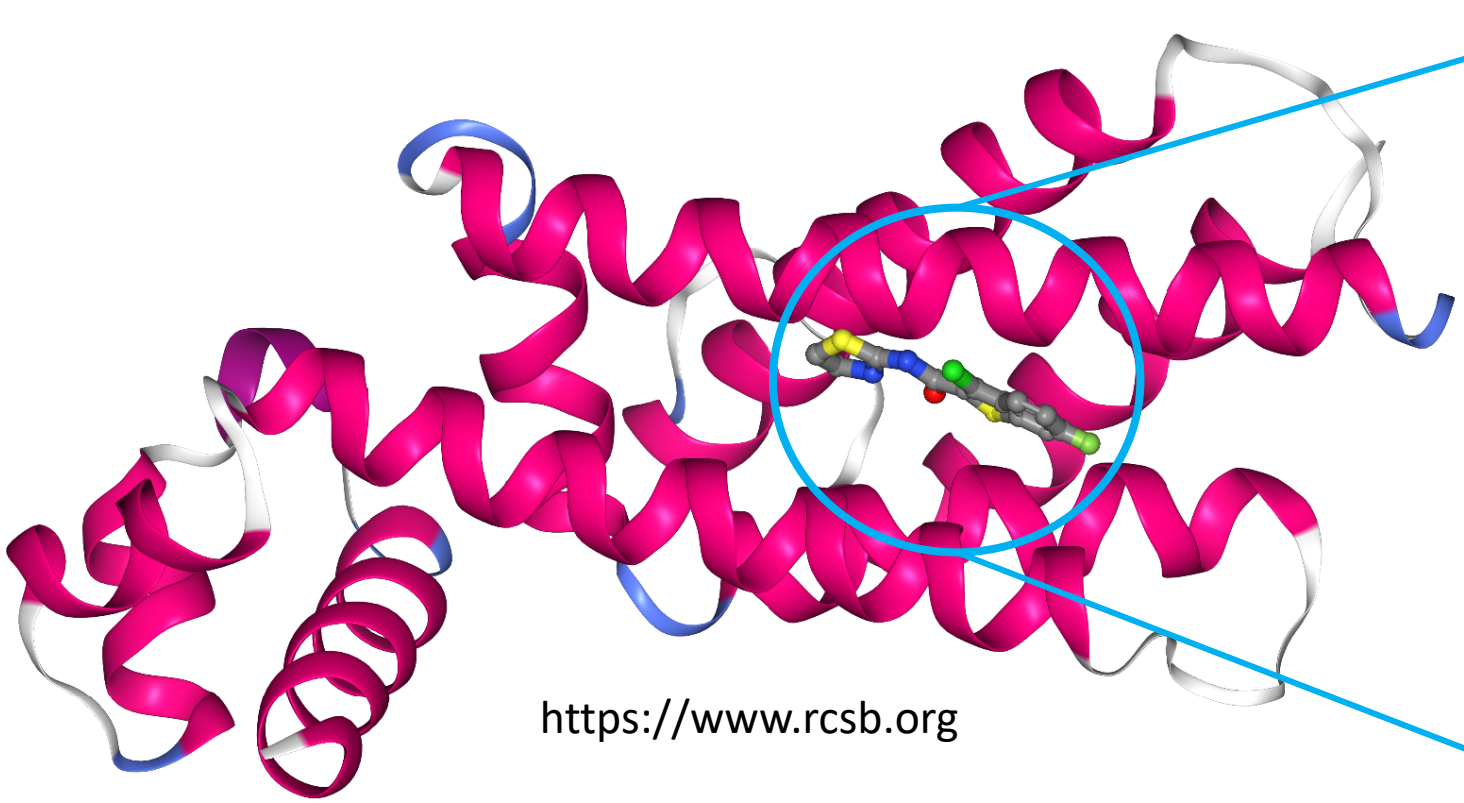
149 Atoms



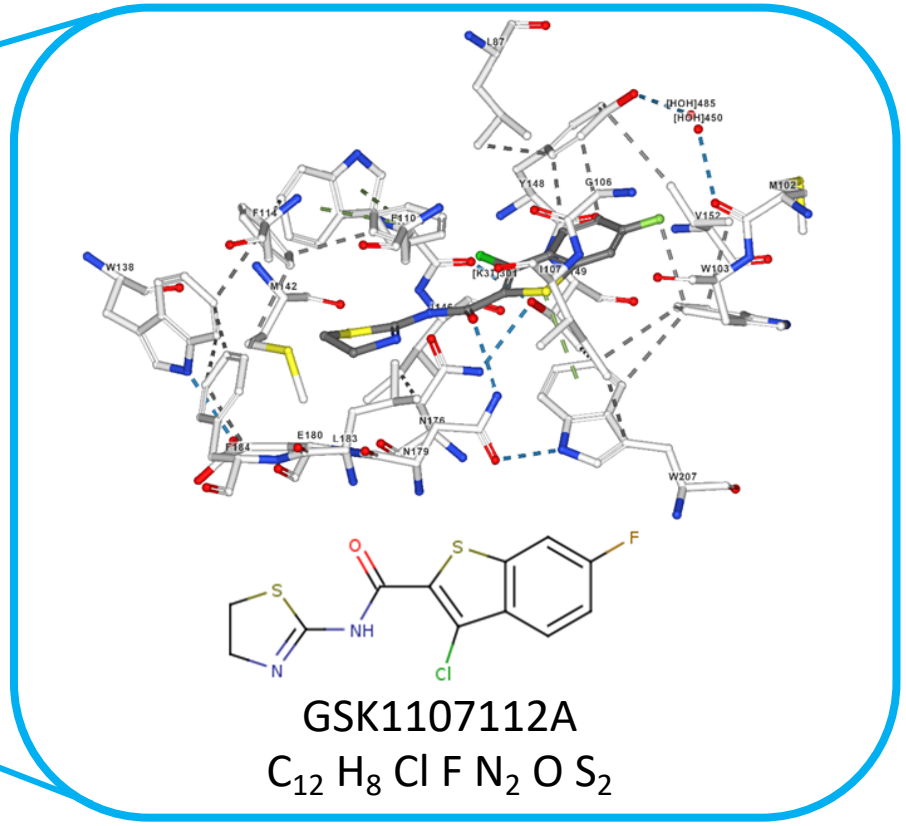
312 Atoms



Transcriptional Regulatory Repressor Protein (5MXV) in explicit water Simulated with ANI-2x (CHNOSFCl)



https://www.rcsb.org

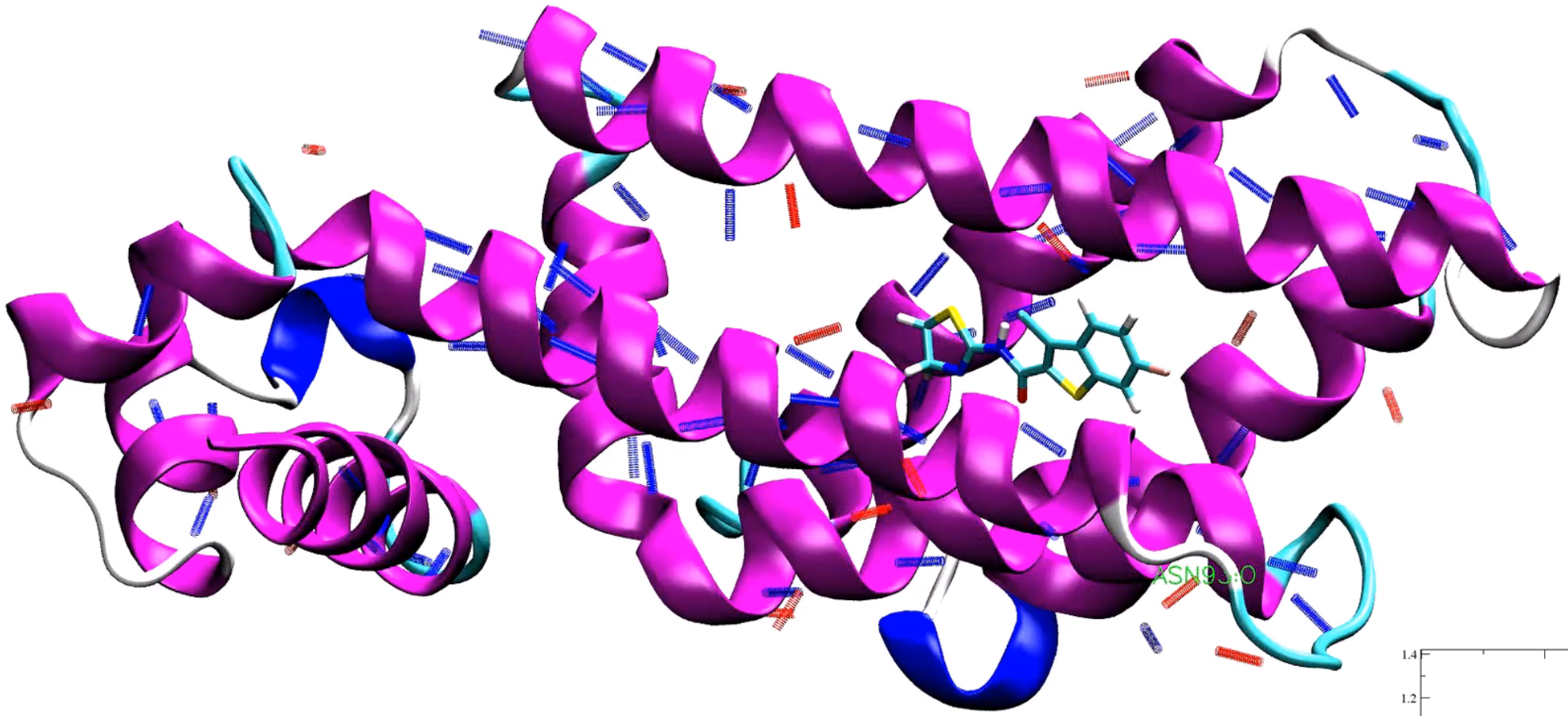


GSK1107112A

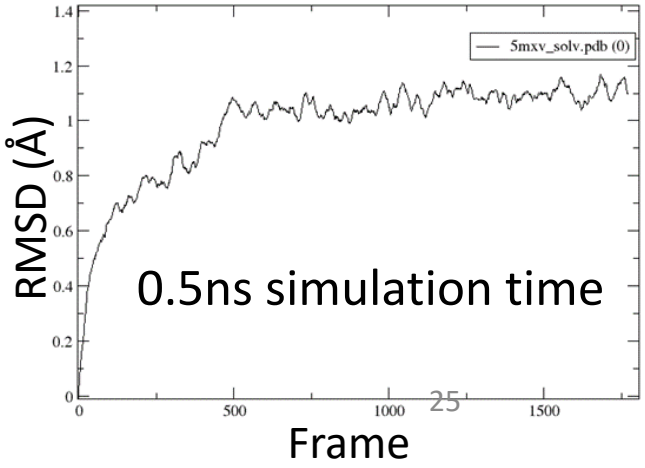
$C_{12}\ H_8\ Cl\ F\ N_2\ O\ S_2$

From Mycobacterium Tuberculosis

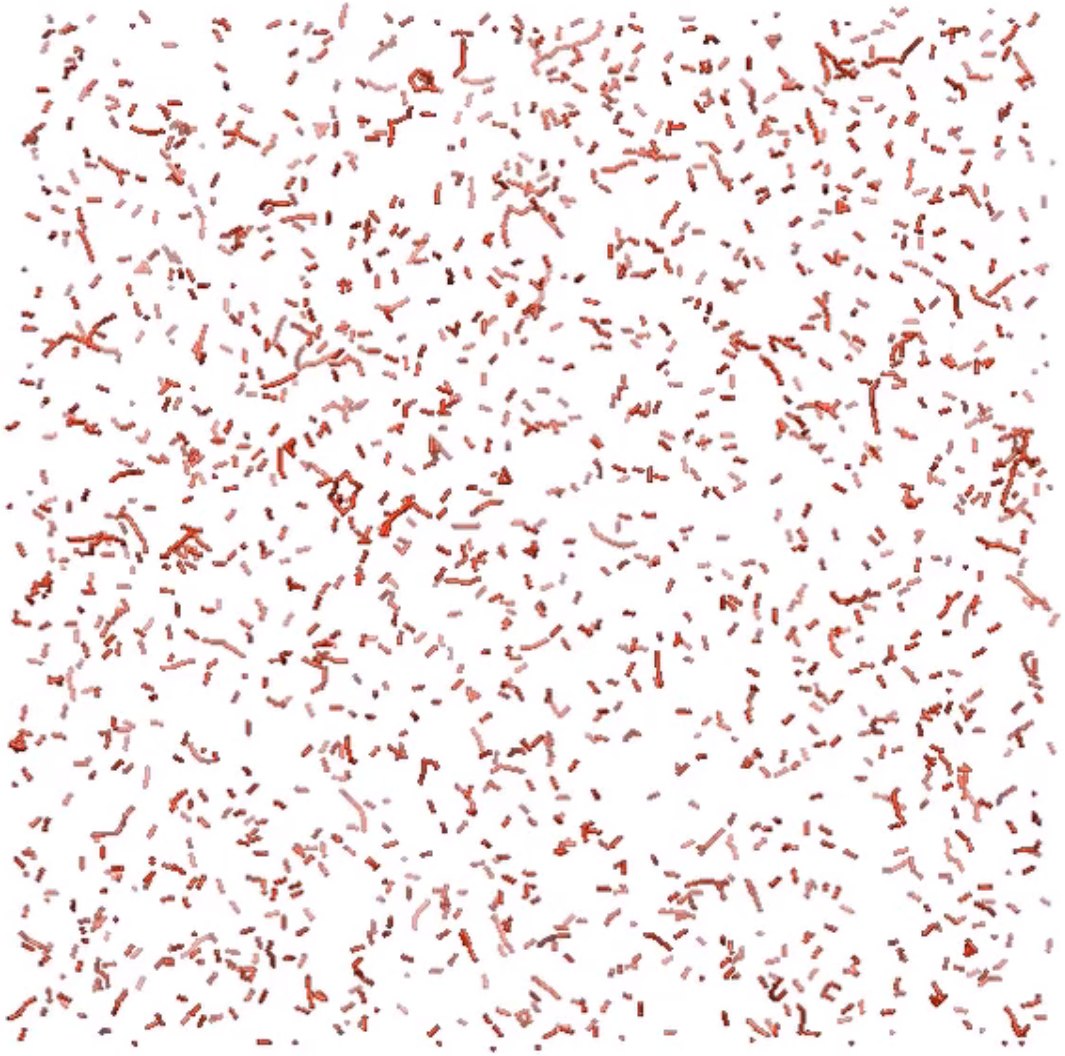
- ~35K atoms
- Explicit water
- No ions
- S, F and Cl in ligand



Timings for a 5x ensemble prediction for ANI-2x			
GPU	ANI-2x time per step	Total time per step	Steps per day
Tesla V100	297ms	317ms	**272k**



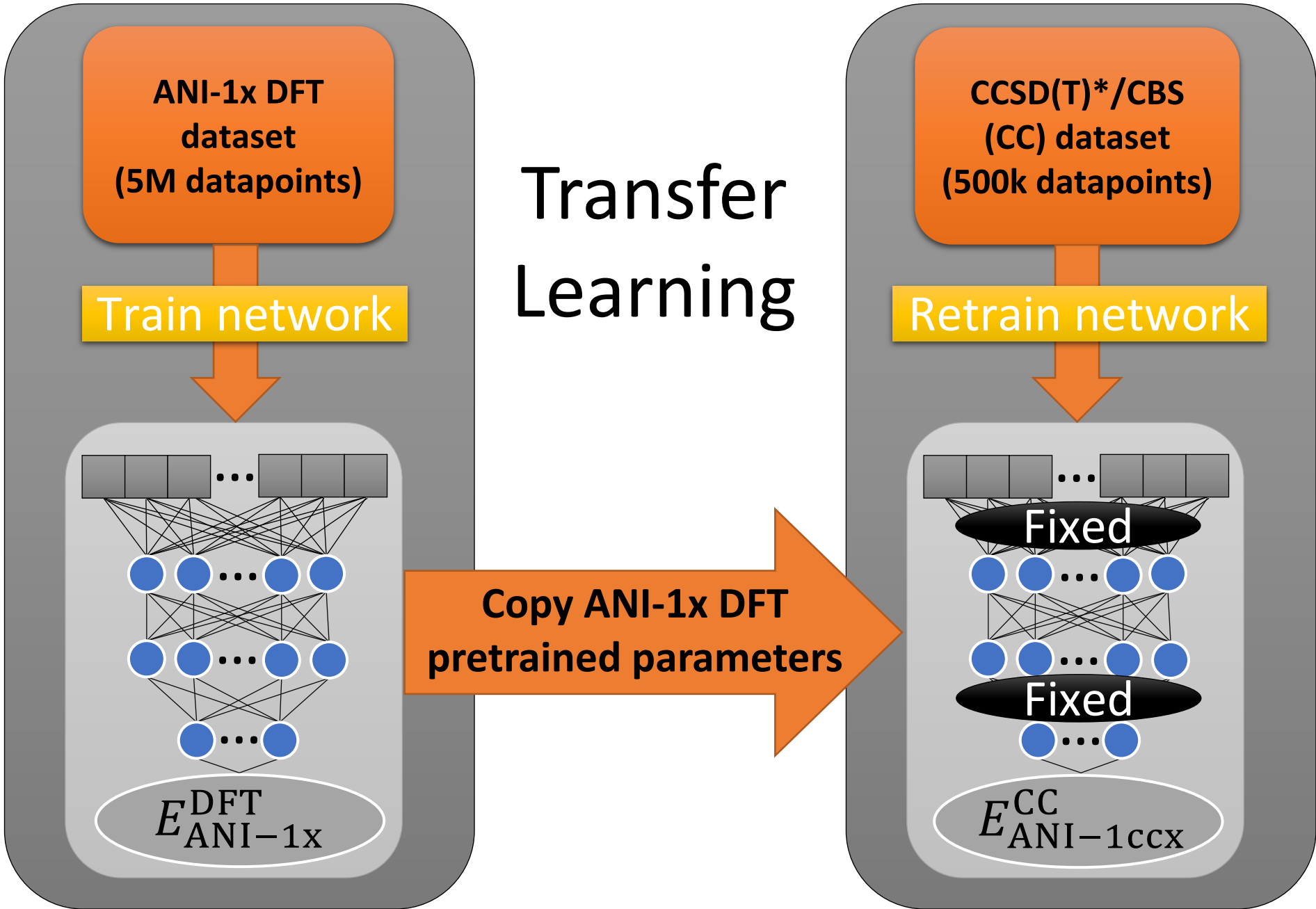
Simulation of Complex Chemical Reactions



Carbon nanoparticles/sheets nucleation [4000 atoms in 60A box at 2500K, 5ns MD simulation]

Transferring knowledge from DFT to CCSD(T)

- Subsample 10% of ANI-1x training data (0.5M of 5M)
- Recompute CCSD(T)/CBS level
- 340k parameters fixed, re-train 60k
- 10^7 faster than DFT

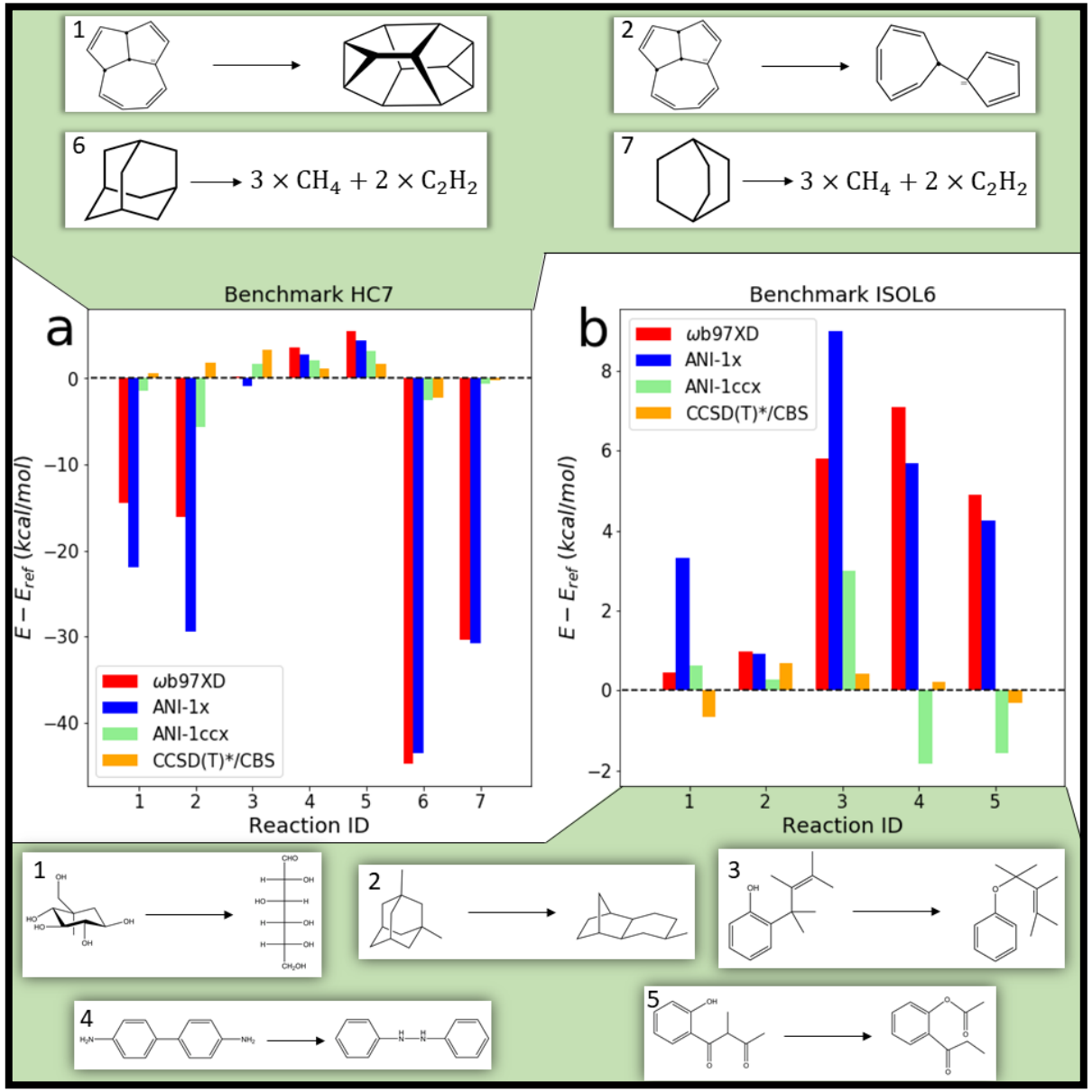


Smith et al.; Nat. Comm. **2019**
https://doi.org/10.1038/s41467-019-10827-4

Outsmarting Quantum Chemistry Through Transfer Learning

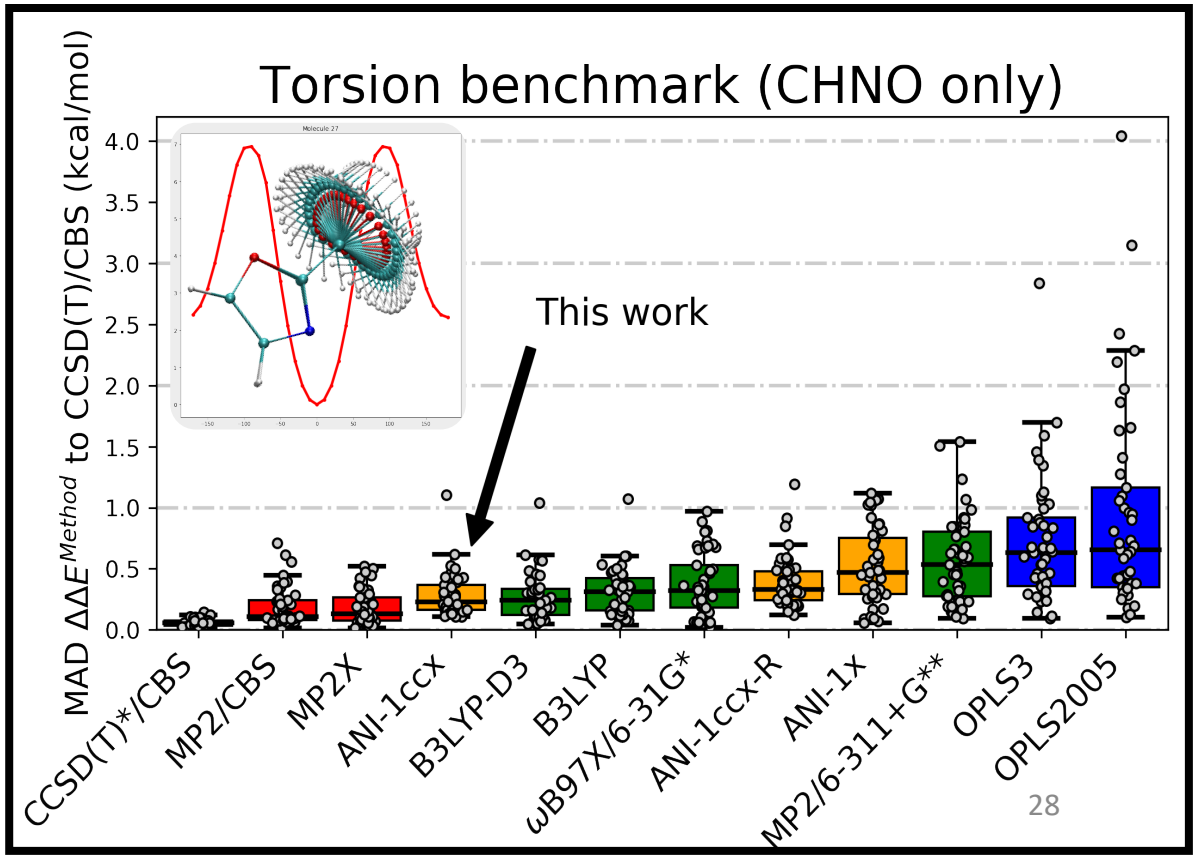
JS Smith, B Nebgen, R Zubatyuk, N Lubbers, C Devereux, K Barros, S Tretiak, O Isayev, A Roitberg

https://doi.org/10.1038/s41467-019-10827-4 Nat. Comm. **2019**



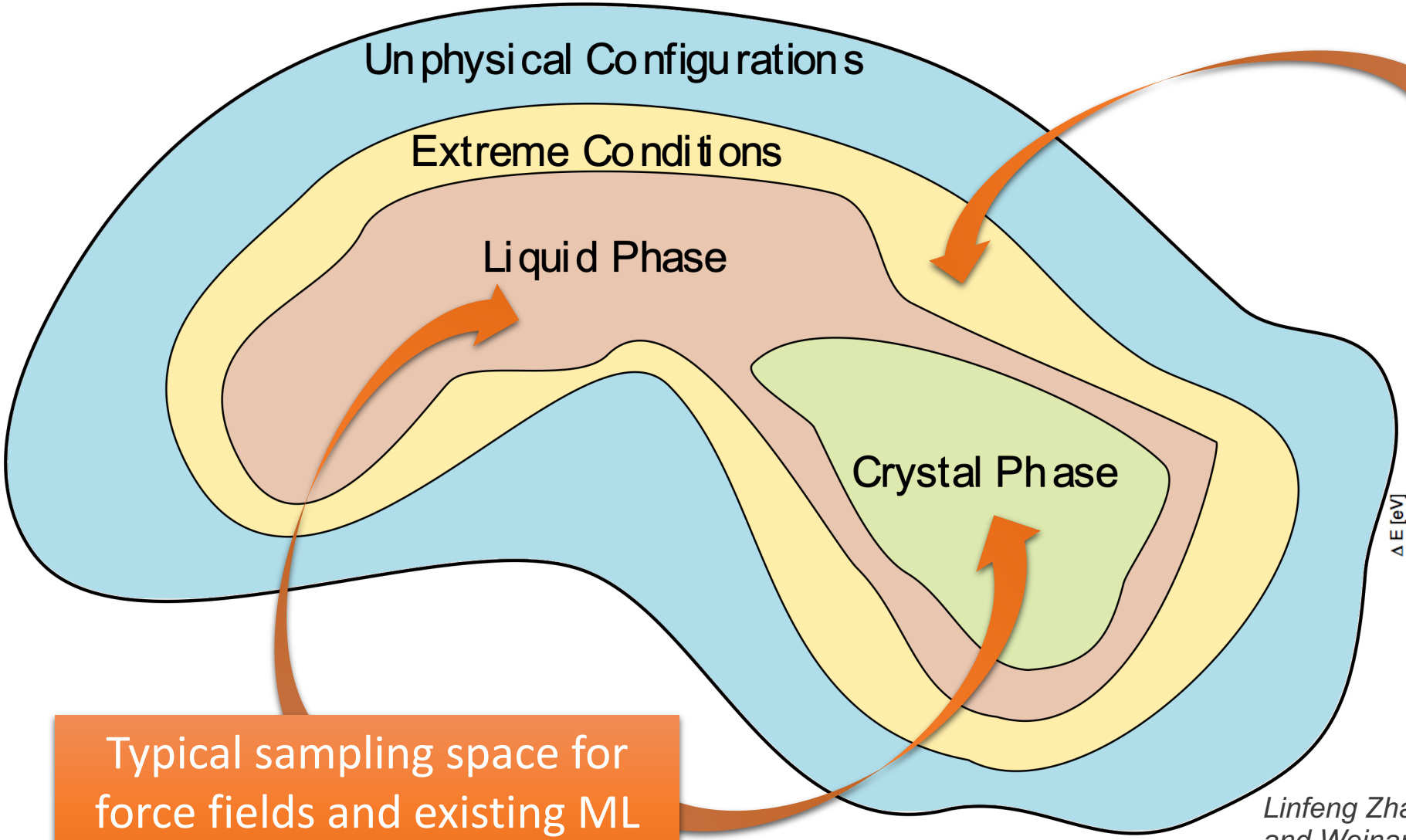
- New ANI-1ccx model outperforms DFT on reaction energies and torsional profiles

- A 24 core hours calculation for CCSD(T)/CBS takes 2 GPU microseconds for ANI-1ccx



How to build a general potential for metal?

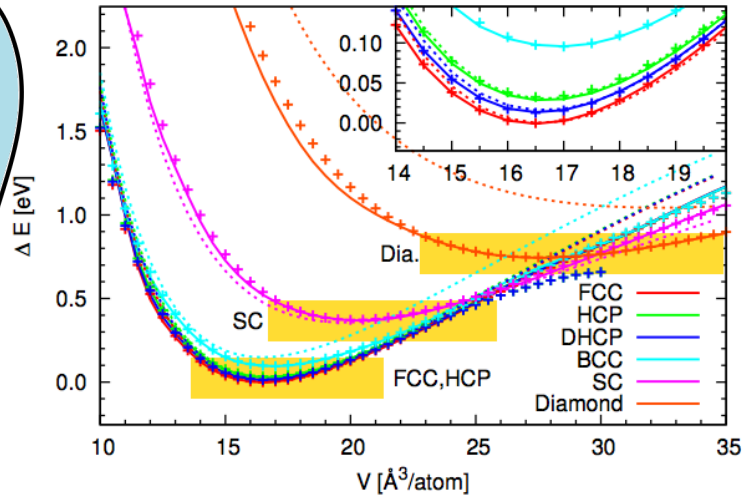
All possible configurations for a metal



Where extreme conditions and rare events exist (e.g. shock simulations)

Typical sampling space for force fields and existing ML potentials

Recent published work

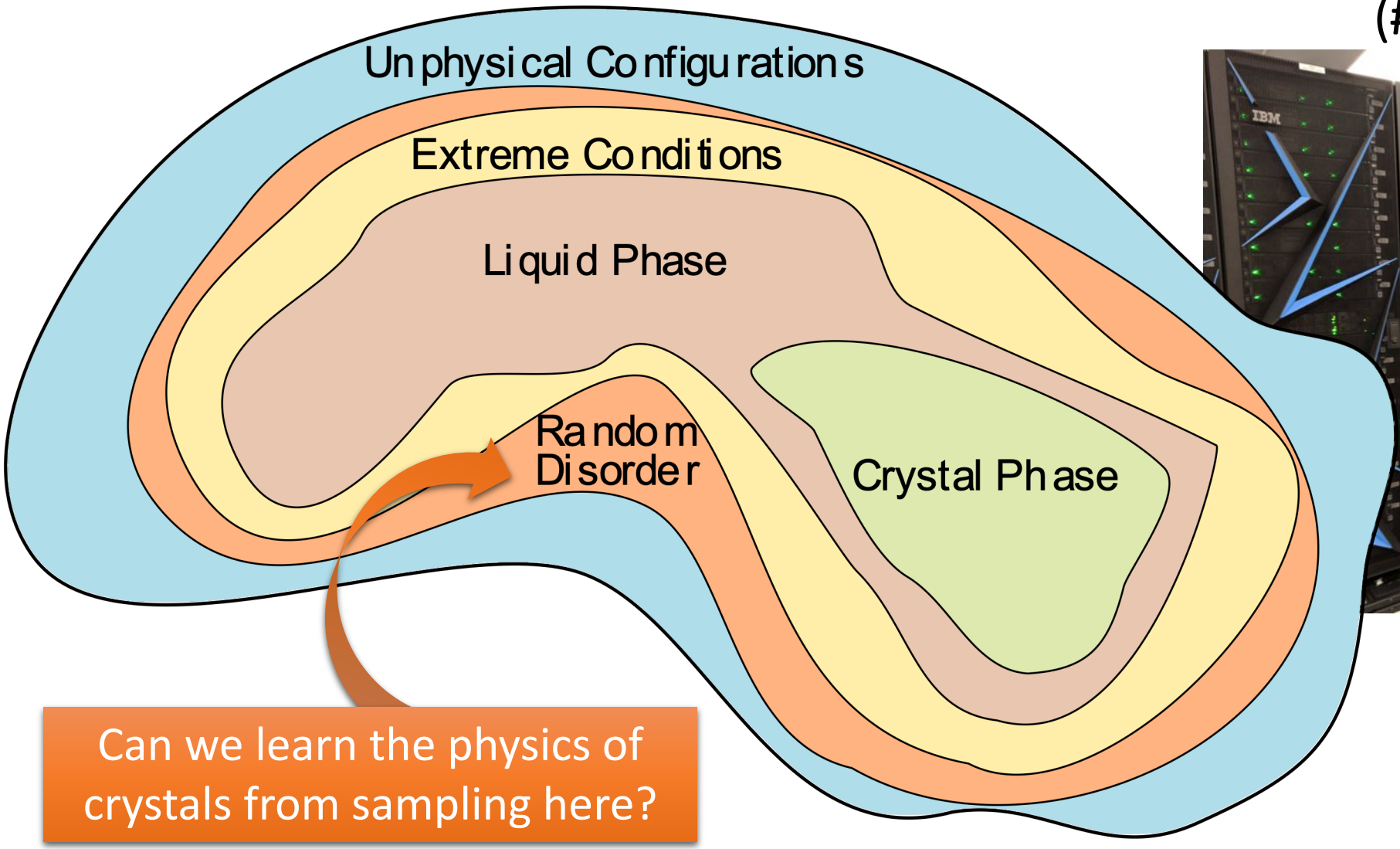


= training data

Linfeng Zhang, De-Ye Lin, Han Wang, Roberto Car, and Weinan E, Active Learning of Uniformly Accurate Inter-atomic Potentials for Materials Simulation, [arXiv:1810.11890]

How should we sample to build a general model?

All possible configurations for a metal



Can we learn the physics of crystals from sampling here?

LLNL Sierra
(#2 on TOP500 list)

Thanks for the open science early access allocation!

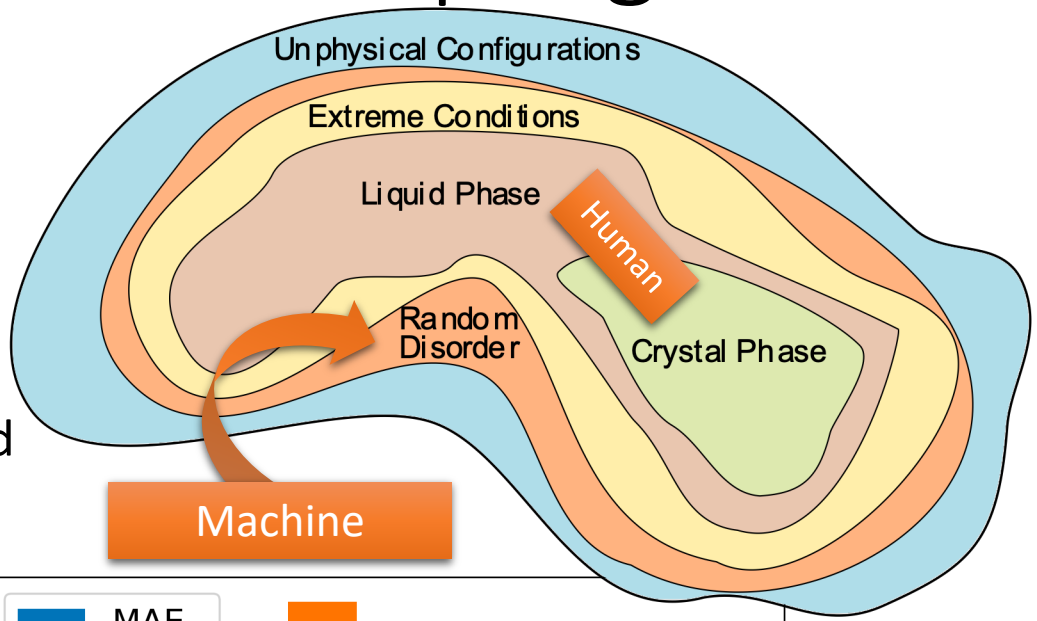
Human vs machine driven sampling

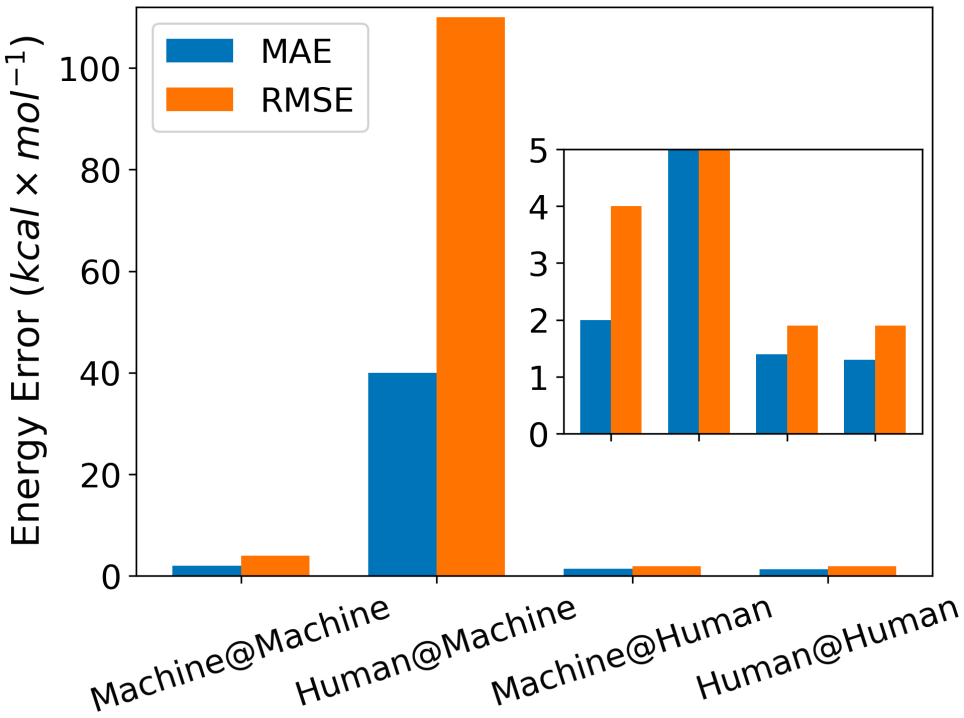
Datasets:

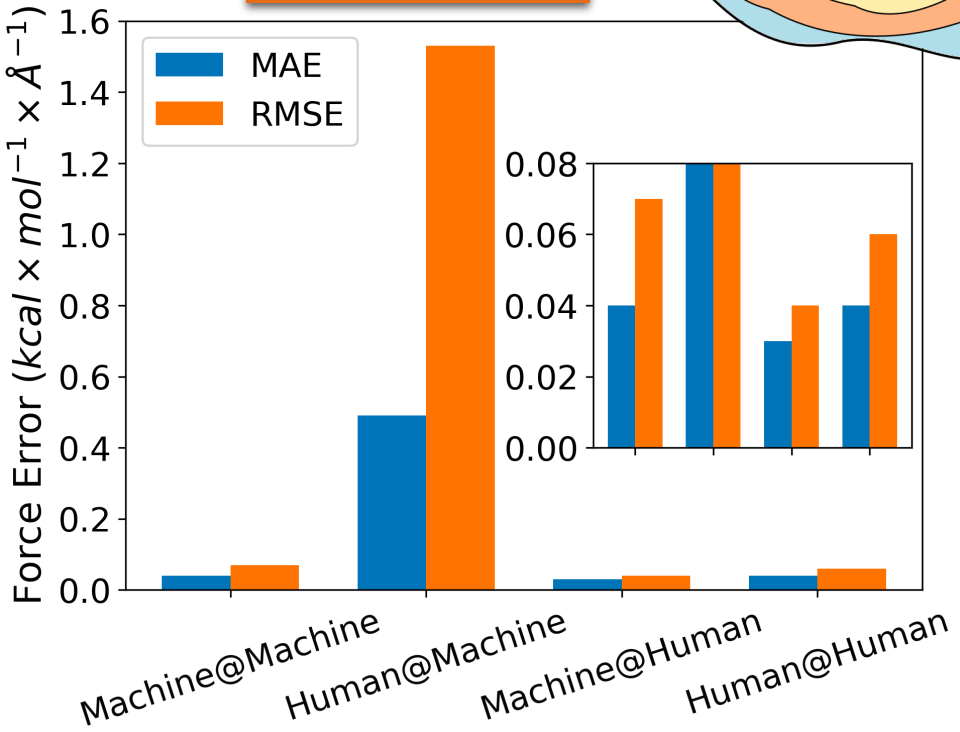
Machine = MD simulation of **Random configurations** and active learning configuations

Human = MD melt simulation of **human selected crystal** with systematic configurations

Machine@Human: trained to **machine selected data** and tested on **human selected data**

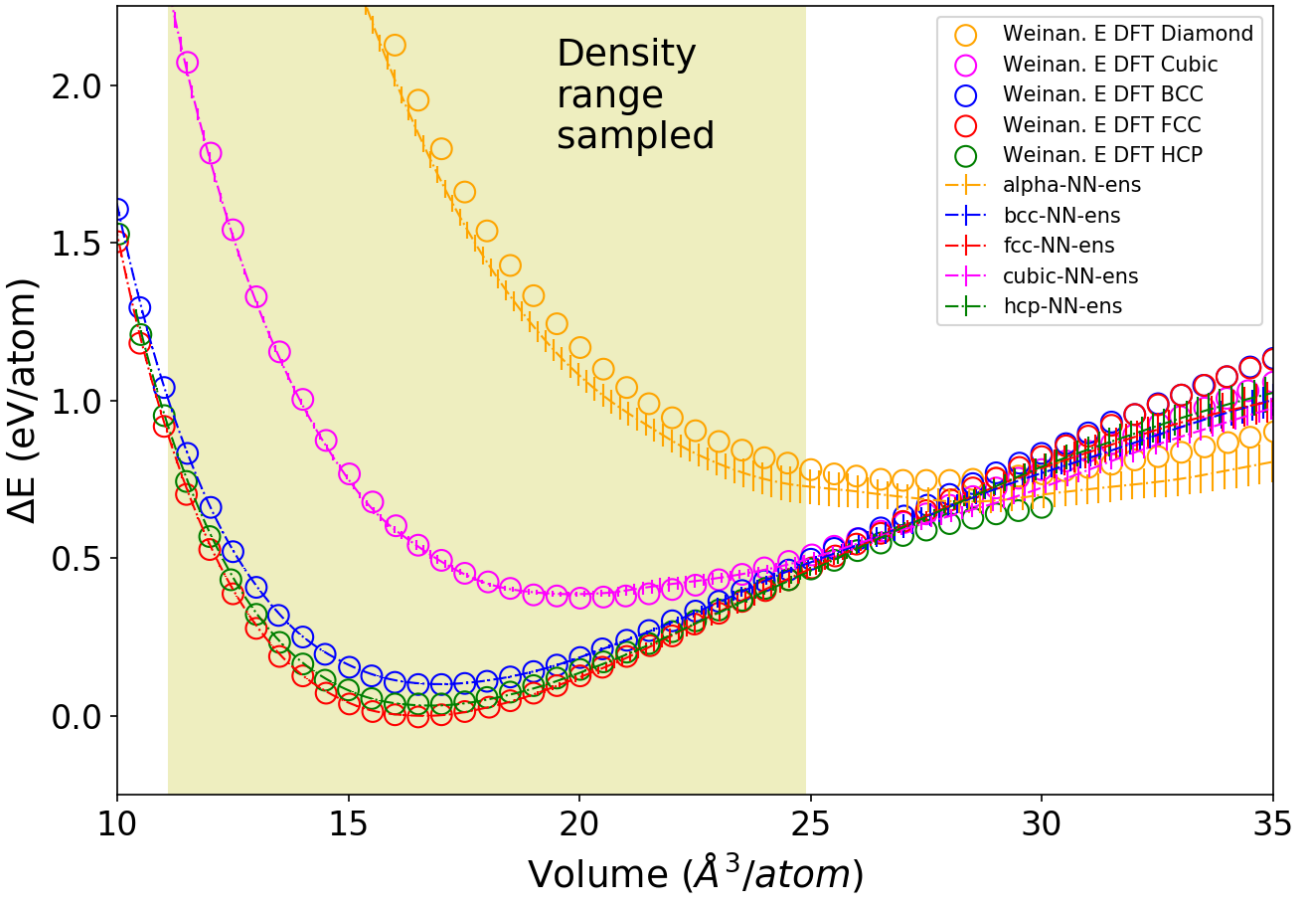




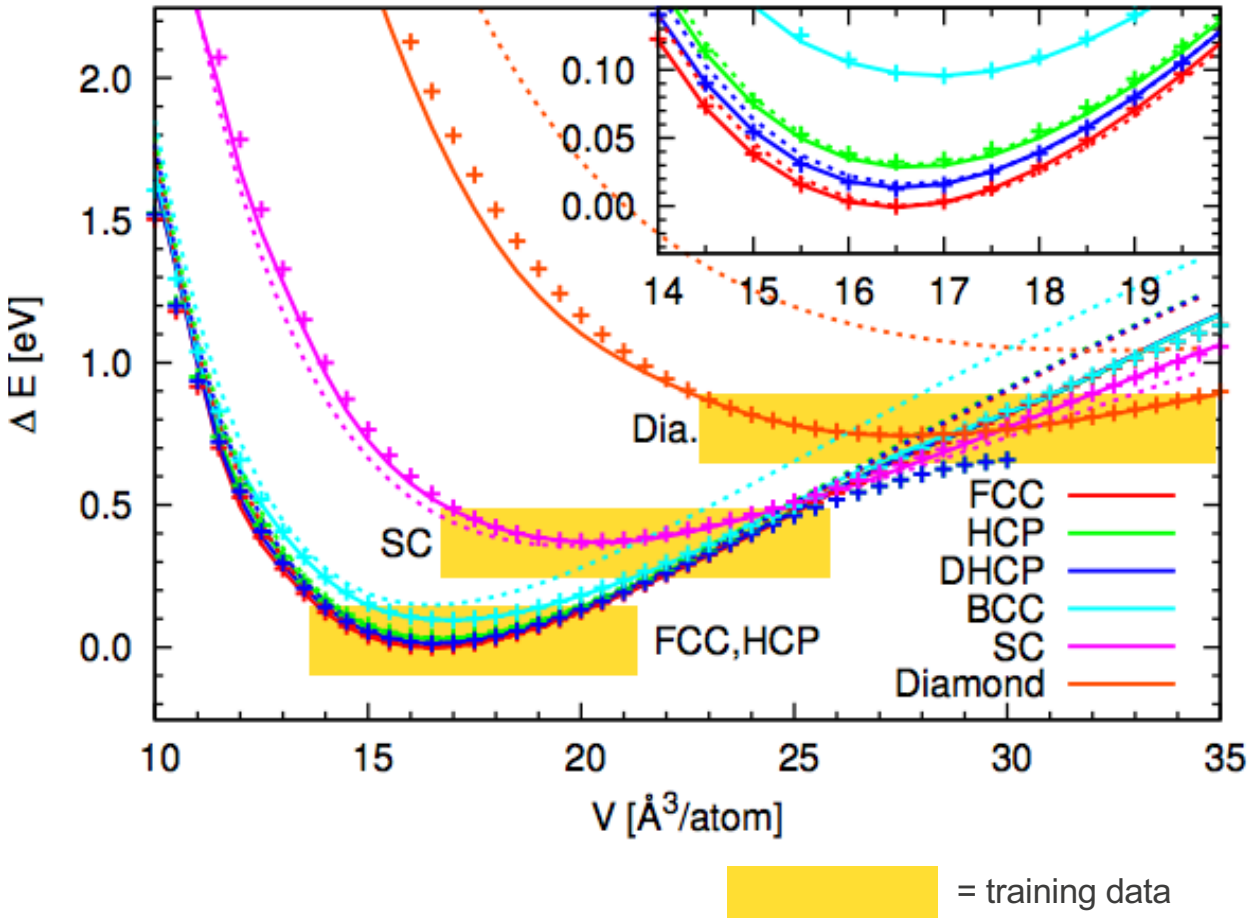


Select crystal vs. random disorder MD sampling for Al

No human knowledge used in sampling
Sampling technique: Disorder only
(our current work)

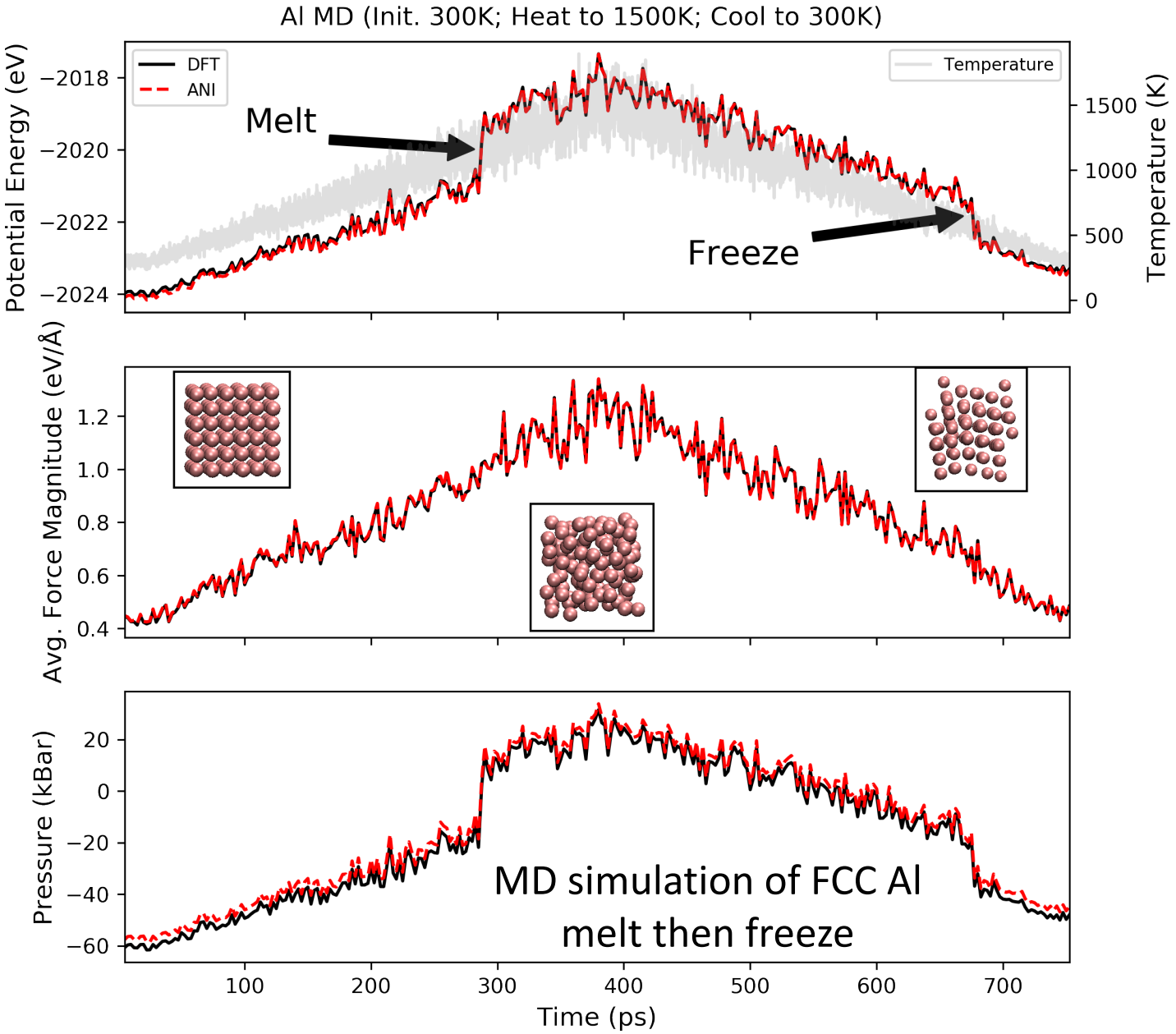


Crystals chosen based on human knowledge (previous literature)

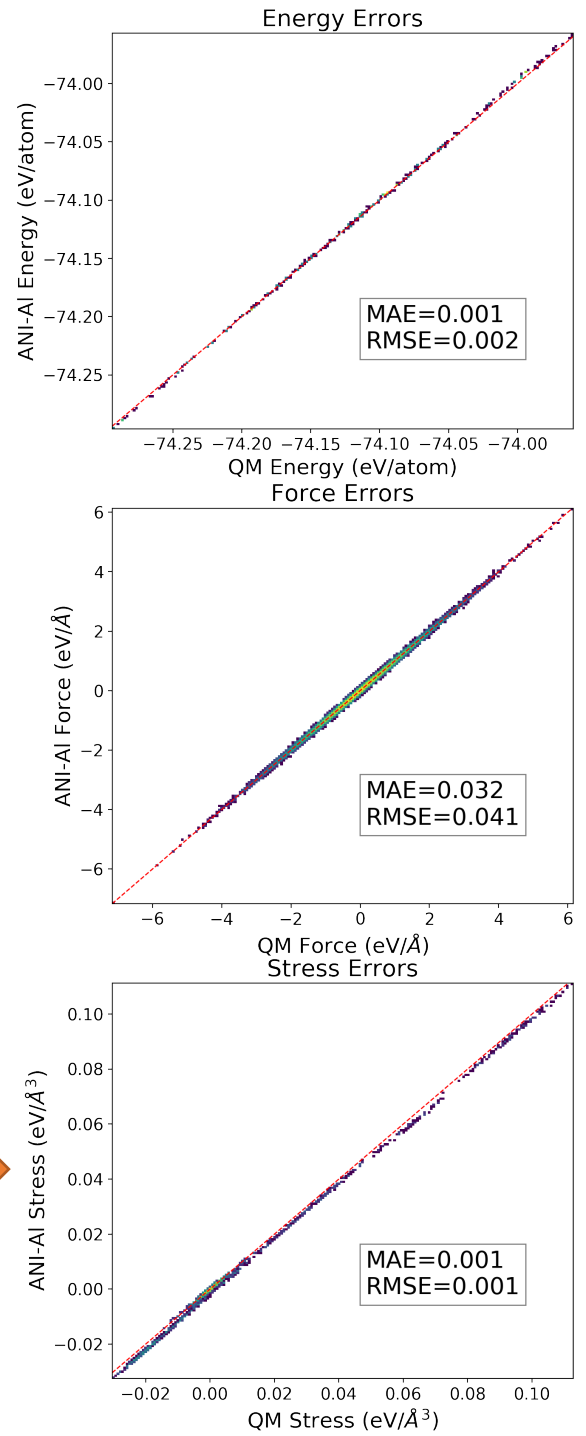


Linfeng Zhang, De-Ye Lin, Han Wang, Roberto Car, and Weinan E, Active Learning of Uniformly Accurate Inter-atomic Potentials for Materials Simulation, [arXiv:1810.11890]

ML model for Al is accurate and general



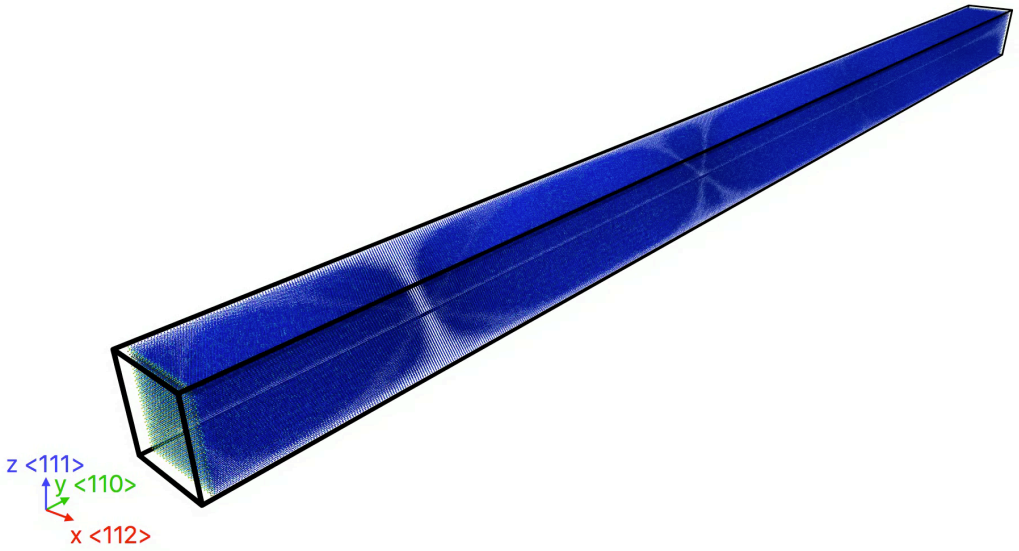
Energy, force, and stress errors for 25x 100ps MD simulations of liquid Al with temperatures in the range of 1000K to 2000K and volumes between 0.8 and 1.2 equilibrium.



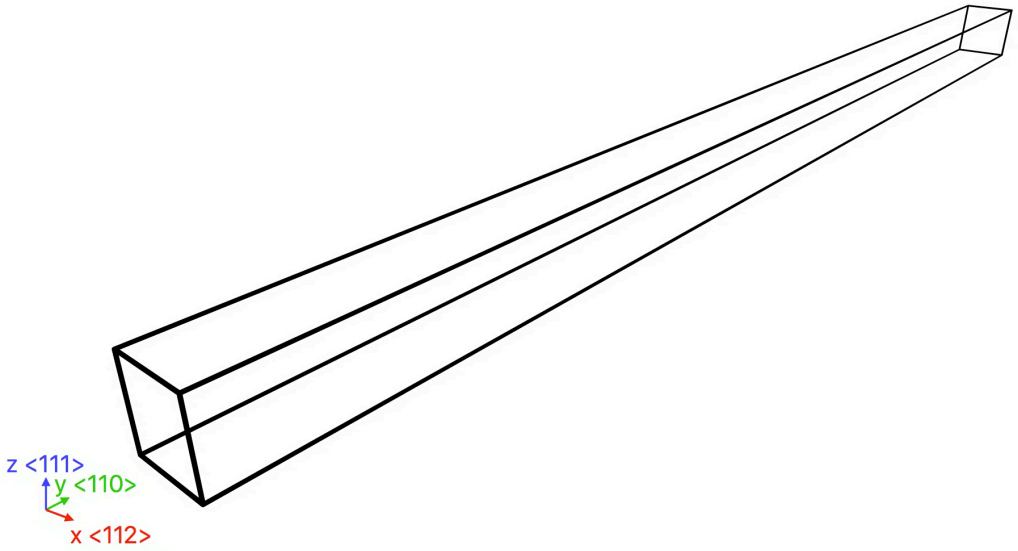
<110> shock in Al using LAMMPS with the ANI potential

1.3 million Al atom simulation

Potential Energy



Dislocation Formation



- Potential composed of an ensemble of 2 Neural Networks potentials
- Simulations carried out on 80x Titan V GPUs in about 10 hours
- 2 kms^{-1} shock

Los Alamos NATIONAL LABORATORY

Thank you!

ASC™

Center for Nonlinear Studies