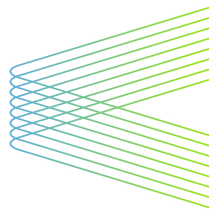


General description of covalent bonds with machine-learning potentials

Alessandro Lunghi and Stefano Sanvito

School of Physics, CRANN and AMBER, Trinity College Dublin 2, Ireland

email: lunghia@tcd.ie



CRANN



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin



Force Fields in Chemistry

Coordination Chemistry:

Applications

Molecular magnetism

Catalysis

Metallo-proteins

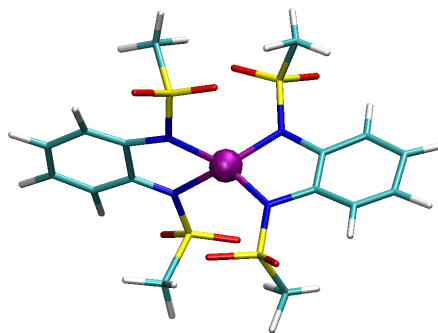
Metal-Organic-Frameworks

Force Fields in Chemistry

Coordination Chemistry:

Applications

Molecular magnetism
Catalysis
Metallo-proteins
Metal-Organic-Frameworks



Co S N O C

Co-N₄ Coordination bond

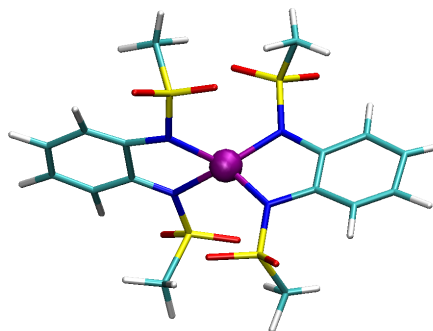
C-C sigma/pi-bonds

Force Fields in Chemistry

Coordination Chemistry:

Applications

Molecular magnetism
Catalysis
Metallo-proteins
Metal-Organic-Frameworks



Co S N O C

Co-N₄ Coordination bond
C-C sigma/pi-bonds

How to do MD? Solutions from “Chem. Rev., 2017, 117, 1564-1686”

Organic Chemistry:

Electrostatic and VdW Potentials:

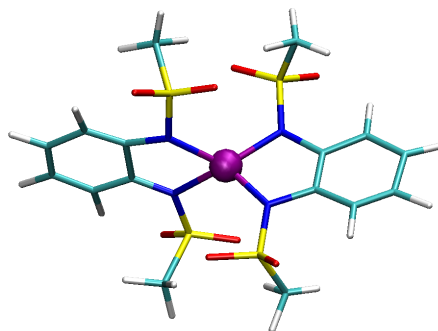
No general solution to simulation of organo-metallic compounds

Force Fields in Chemistry

Coordination Chemistry:

Applications

Molecular magnetism
Catalysis
Metallo-proteins
Metal-Organic-Frameworks



Co S N O C

Co-N₄ Coordination bond
C-C sigma/pi-bonds

How to do MD? Solutions from “Chem. Rev., 2017, 117, 1564-1686”

Organic Chemistry:

Stretching, bendings, torsions...

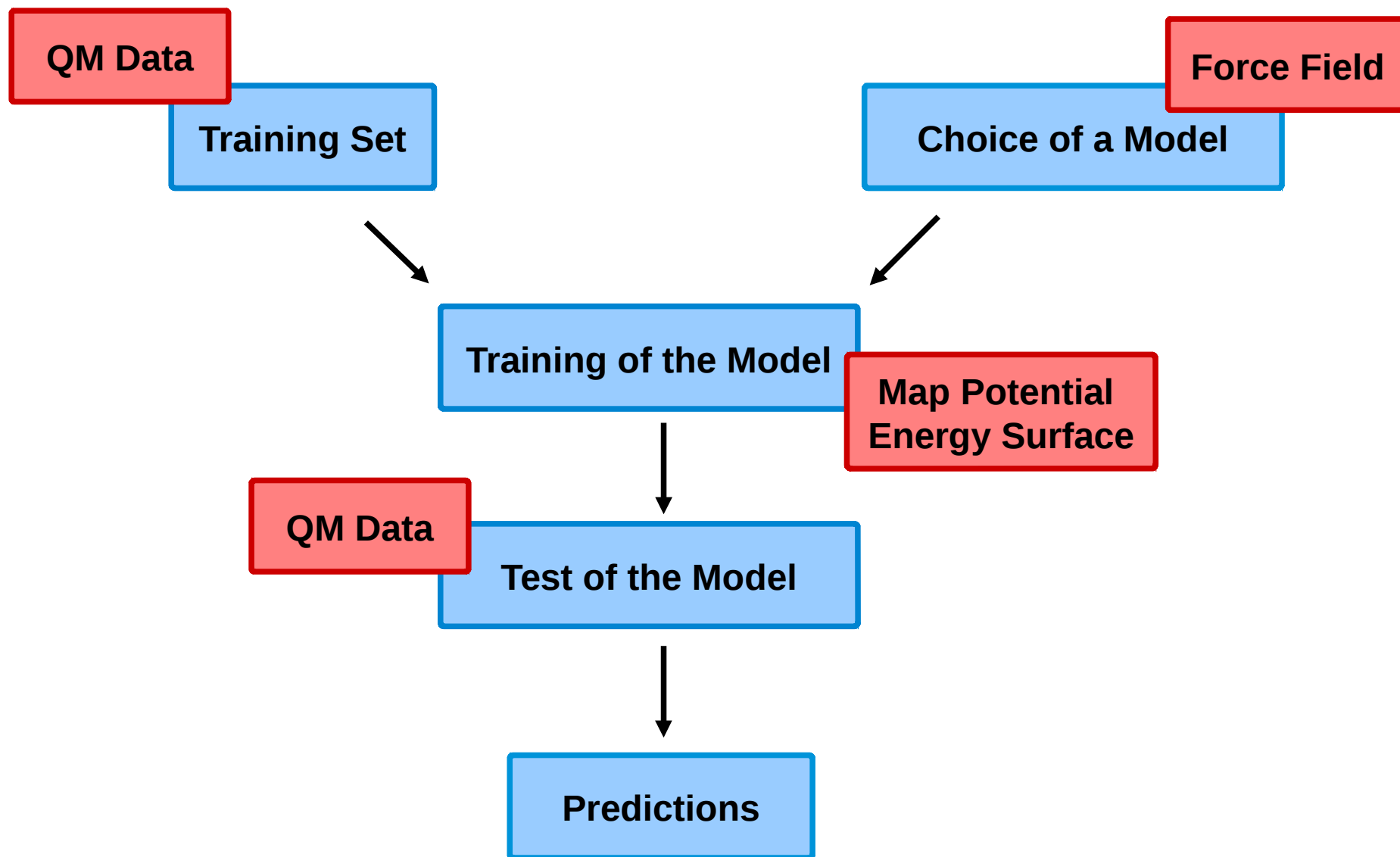
$$E_{AMBER} = \sum_i \alpha_i (r_i - r_{i0})^2 + \sum_j \beta_j (\theta_j - \theta_{j0})^2 + \dots$$

Electrostatic and VdW Potentials:

$$E = \left[\left(\frac{A}{r^{12}} \right) - \left(\frac{B}{r^6} \right) \right] + \frac{C}{r}$$

No general solution to simulation of organo-metallic compounds

Machine Learning Force Fields



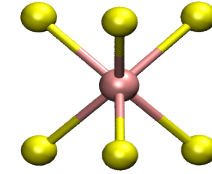
Spectral Neighbour Analysis Potential (SNAP)

Phys. Rev. Lett. 104 (2010) 136403

Descriptors: Bispectrum components

$$u_{m,m'}^j = U_{m,m'}^j(0,0,0) + \sum_{r_{ii'} < R_{ii'}} f_c(r_{ii'}) w_{i'} U_{m,m'}^j(\theta_0, \theta, \phi)$$

Atomic Environment



Wigner Matrix

$$B_{j_1, j_2, j} = \sum_{m_1, m_1' = -j_1}^{j_1} \sum_{m_2, m_2' = -j_2}^{j_2} \sum_{m, m' = -j}^j (u_{m, m'}^j)^* H_{\begin{matrix} j_1 m_1 m_1' \\ j_2 m_2 m_2' \\ j m m' \end{matrix}}^{j m m'} u_{m_1, m_1'}^{j_1} u_{m_2, m_2'}^{j_2}$$

Clebsh-Gordan coefficients

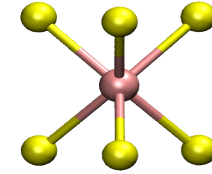
Spectral Neighbour Analysis Potential (SNAP)

Phys. Rev. Lett. 104 (2010) 136403

Descriptors: Bispectrum components

$$u_{m,m'}^j = U_{m,m'}^j(0,0,0) + \sum_{r_{ii'} < R_{ii'}} f_c(r_{ii'}) w_{i'} U_{m,m'}^j(\theta_0, \theta, \phi)$$

Atomic Environment



Wigner Matrix

$$B_{j_1, j_2, j} = \sum_{m_1, m_1' = -j_1}^{j_1} \sum_{m_2, m_2' = -j_2}^{j_2} \sum_{m, m' = -j}^j (u_{m, m'}^j)^* H_{\begin{matrix} j_1 m_1 m_1' \\ j_2 m_2 m_2' \\ j m m' \end{matrix}}^{j m m'} u_{m_1, m_1'}^{j_1} u_{m_2, m_2'}^{j_2}$$

Clebsh-Gordan coefficients

J. Comp. Phys. 285 (2015) 316-330

Energy Predictor: Linear Combination of B

$$E_{SNAP}^i(B_1^i, \dots, B_K^i) = \beta_0^{\alpha_i} + \sum_{k=1}^K \beta_k^{\alpha_i} B_k^i$$

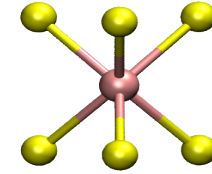
Spectral Neighbour Analysis Potential (SNAP)

Phys. Rev. Lett. 104 (2010) 136403

Descriptors: Bispectrum components

$$u_{m,m'}^j = U_{m,m'}^j(0,0,0) + \sum_{r_{ii'} < R_{ii'}} f_c(r_{ii'}) w_{i'} U_{m,m'}^j(\theta_0, \theta, \phi)$$

Atomic Environment



Wigner Matrix

$$B_{j_1, j_2, j} = \sum_{m_1, m_1' = -j_1}^{j_1} \sum_{m_2, m_2' = -j_2}^{j_2} \sum_{m, m' = -j}^j (u_{m, m'}^j)^* H_{\begin{matrix} j_1 m_1 m_1' \\ j_2 m_2 m_2' \\ j m m' \end{matrix}}^{j m m'} u_{m_1, m_1'}^{j_1} u_{m_2, m_2'}^{j_2}$$

Clebsh-Gordan coefficients

J. Comp. Phys. 285 (2015) 316-330

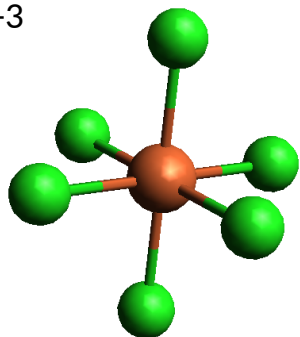
Energy Predictor: Linear Combination of B

$$E_{SNAP}^i(B_1^i, \dots, B_K^i) = \beta_0^{\alpha_i} + \sum_{k=1}^K \beta_k^{\alpha_i} B_k^i$$

Regression: Ridge

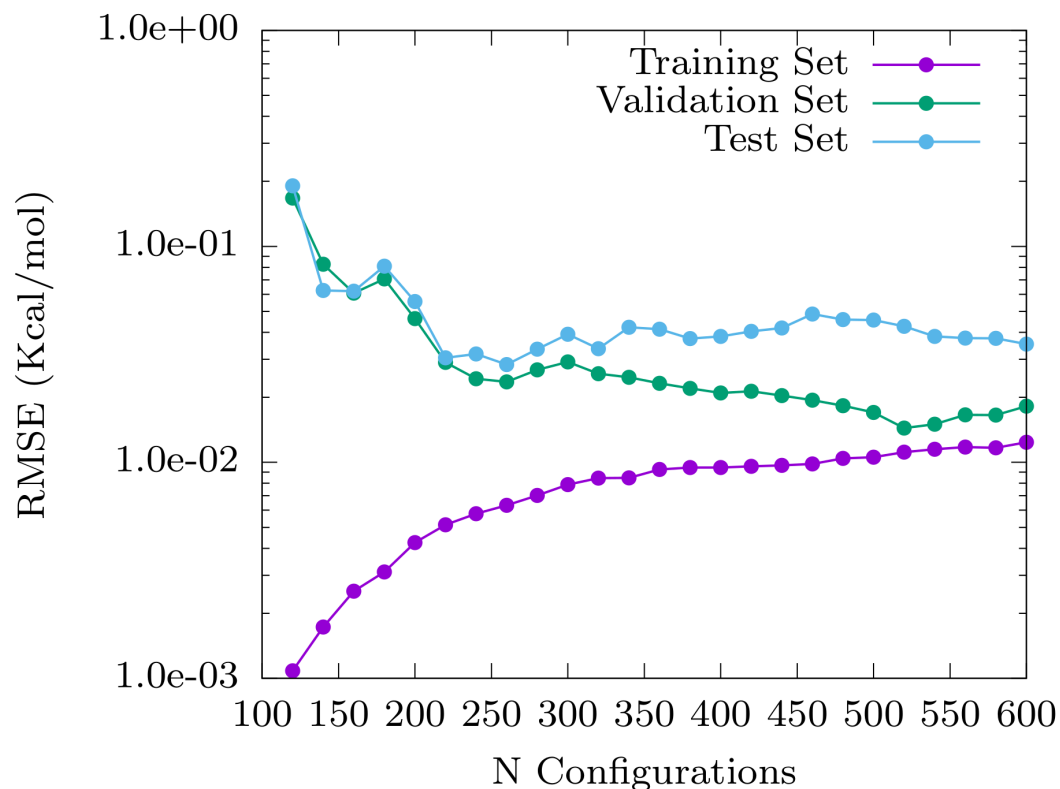
$$\text{Min}_{\beta} (E_{DFT} - E_{SNAP}(\beta_k))^2 + \lambda \sum_k \beta_k^2$$

Coordination Chemistry



Code: ORCA
QM: DFT PBE

Configurations: 800 Random distortion
with 0.1/0.2 Angstrom of Max
displacement

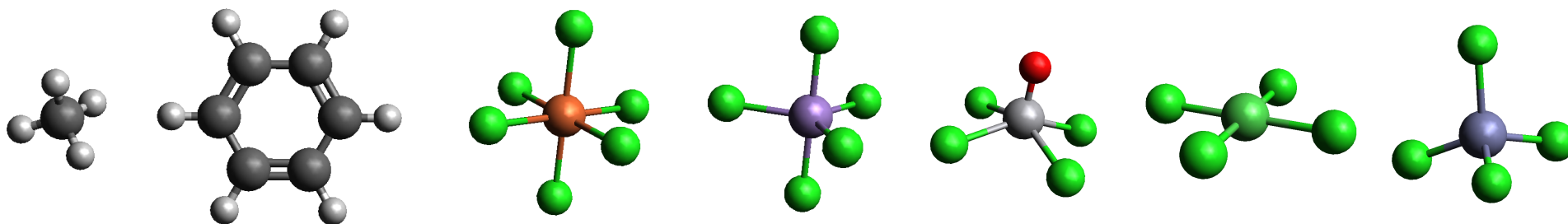
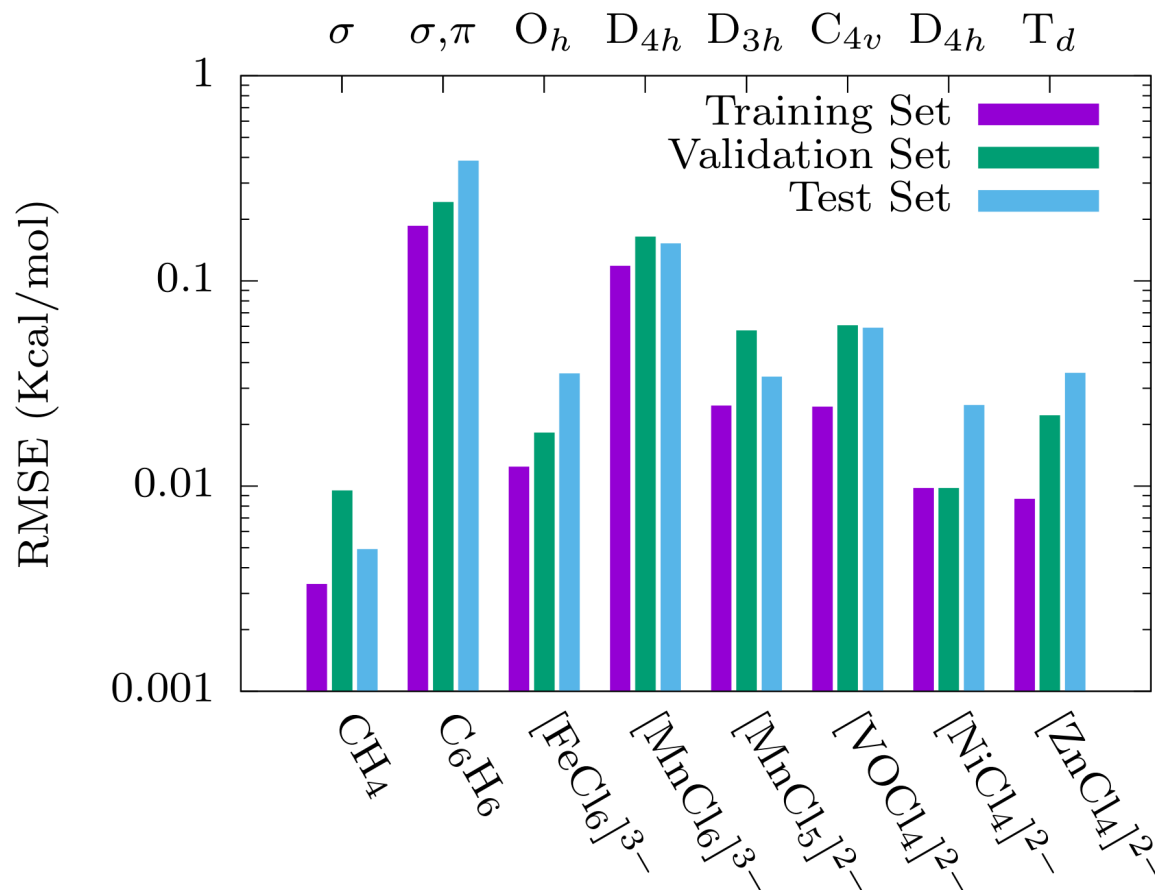


Training Set: Configurations used
for the fitting

Validation Set: To set regularization
parameters

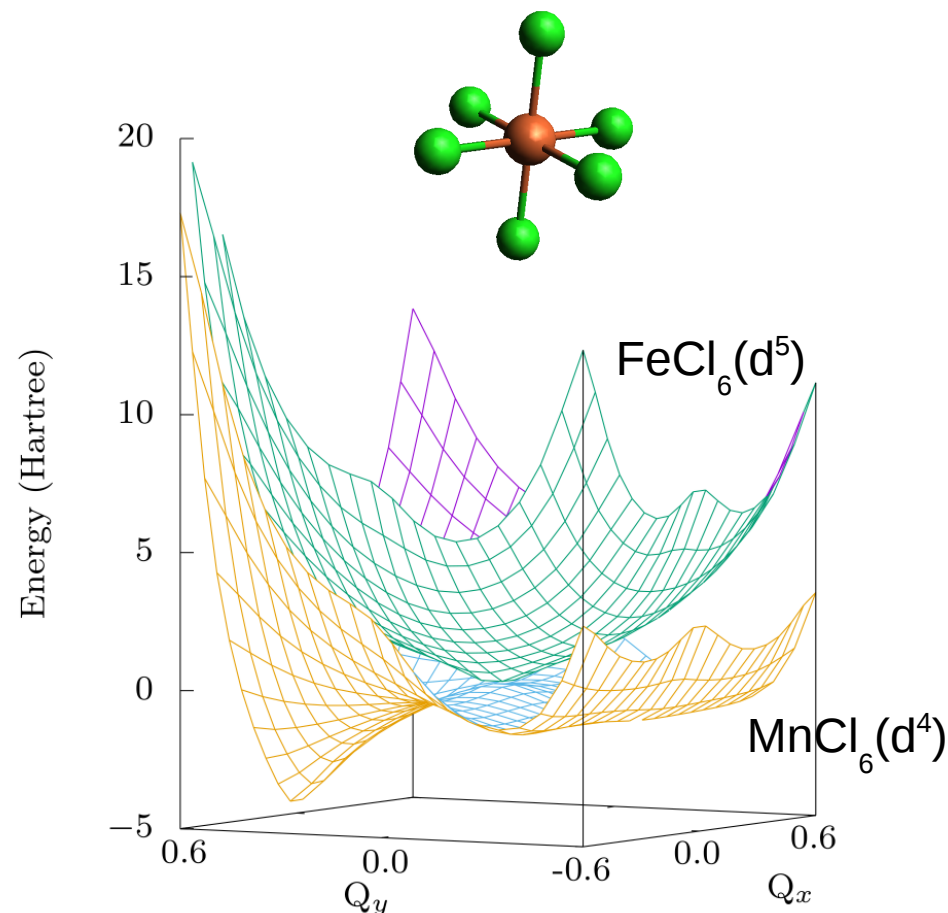
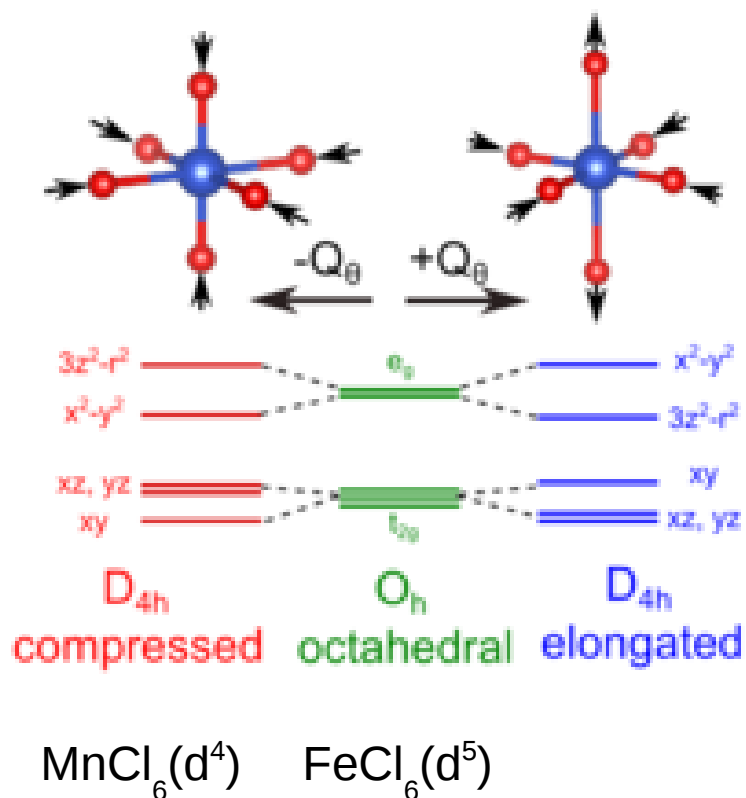
Test Set: Gives a unbiased estimation of the error

General Coordination Environments



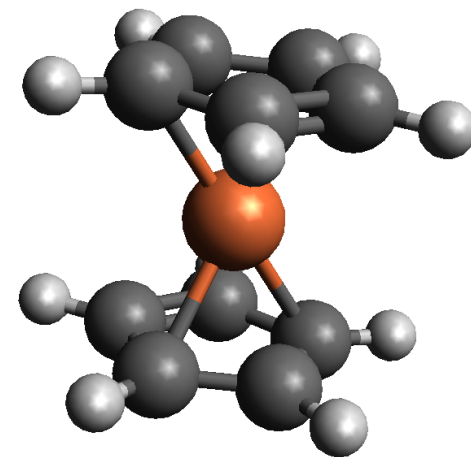
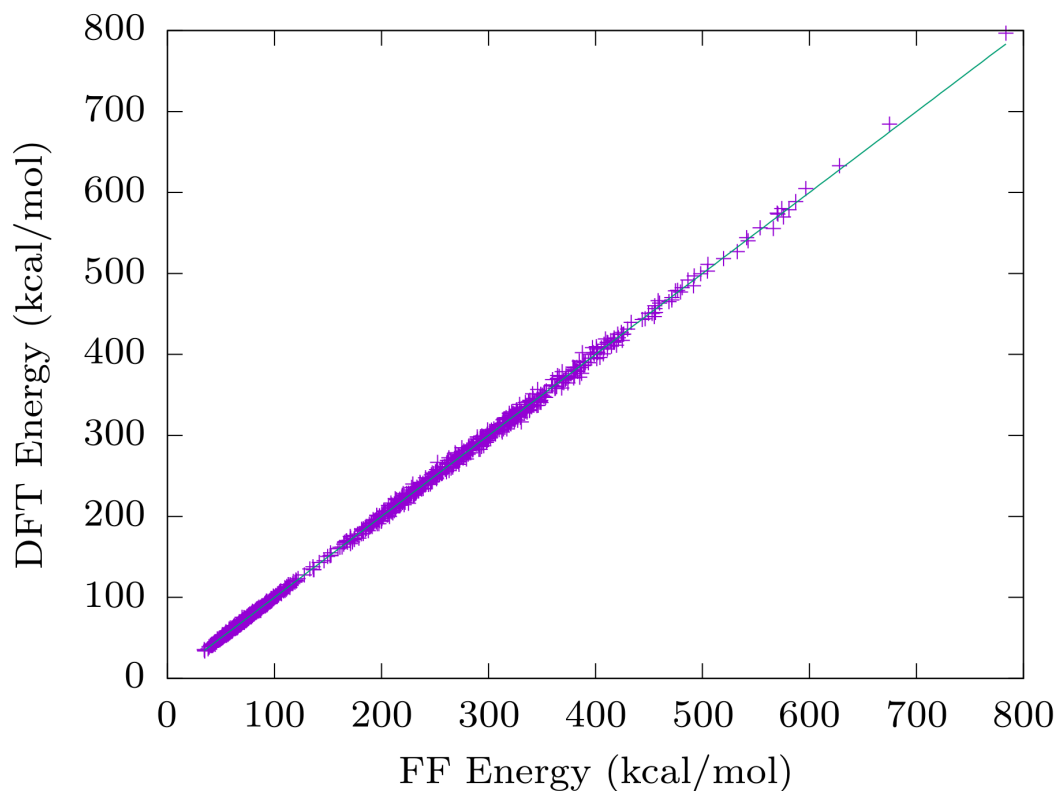
Jahn-Teller Effect

Jahn-Teller Theorem: A quantum mechanical electronic ground state is always non-degenerate.



Jahn Teller Distortion automatically arises from the PES

Ferrocene Fe(II)Cp₂



Training Set: 600 Conf

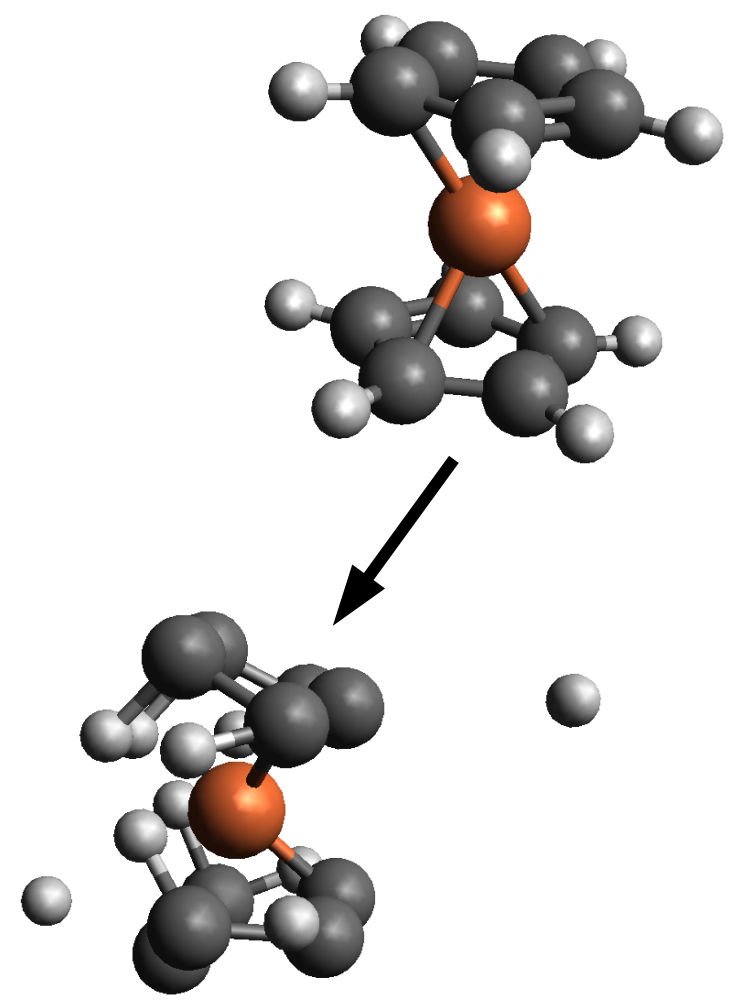
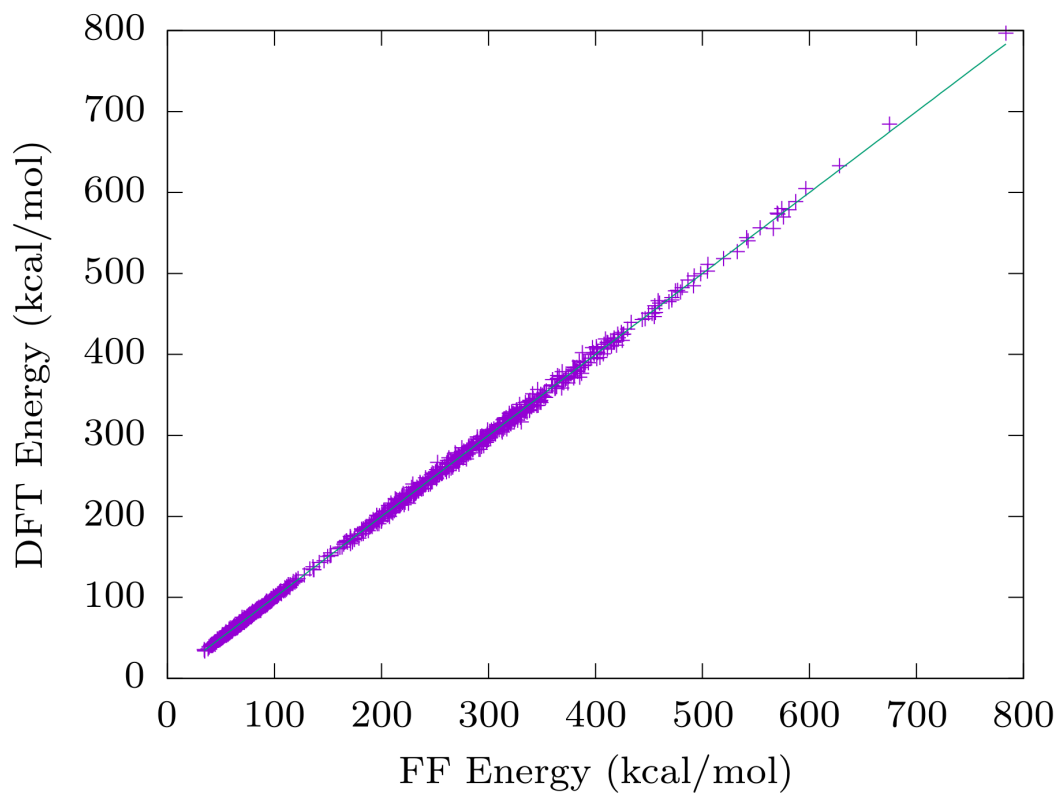
Val Set: 200 Conf

Test Set: 200 Conf

Training RMSE: 3.1 kcal/mol

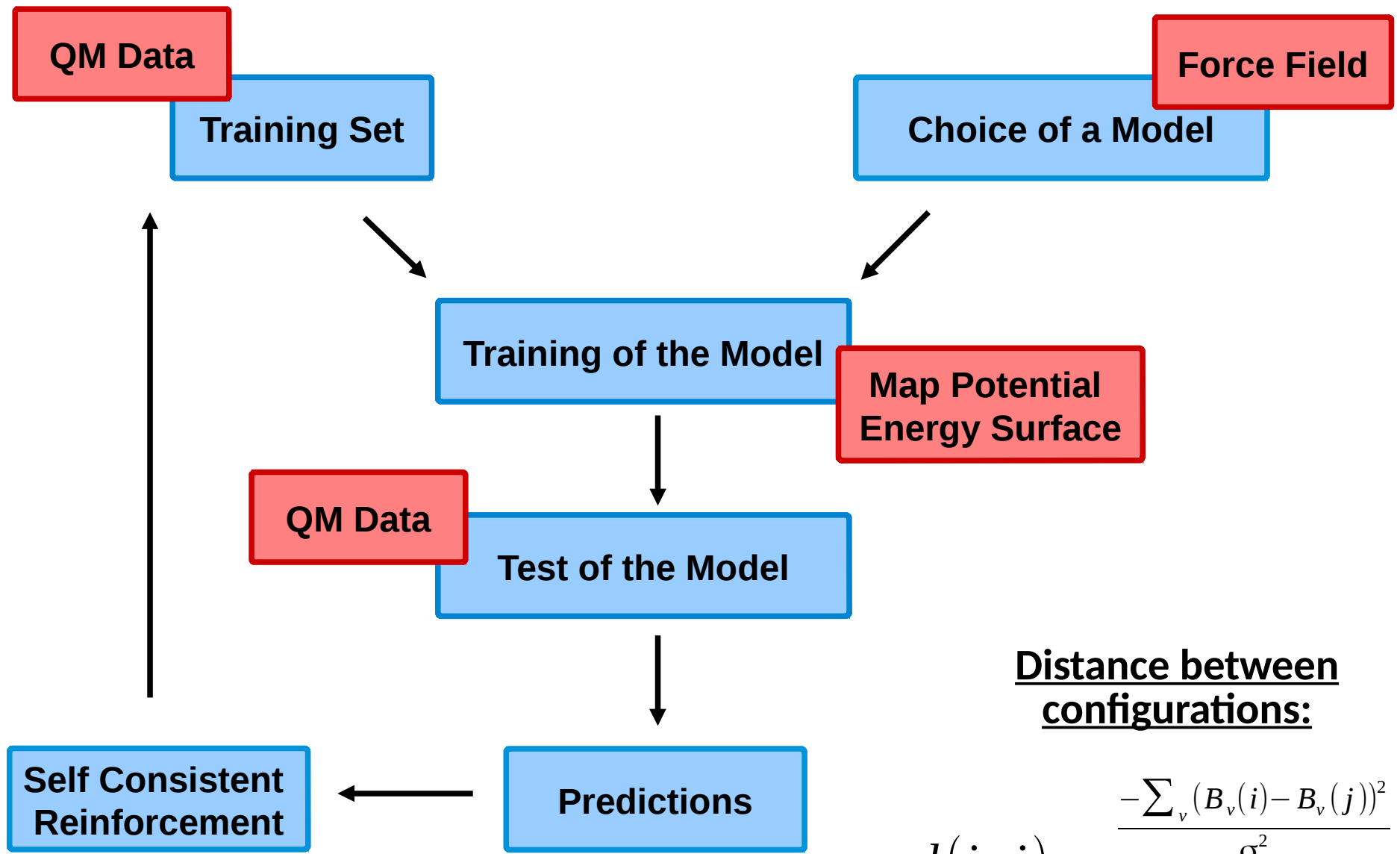
Test RMSE: 3.9 kcal/mol

Ferrocene Fe(II)Cp₂



The completeness of the TRAINING SET is essential

Machine Learning Force Fields

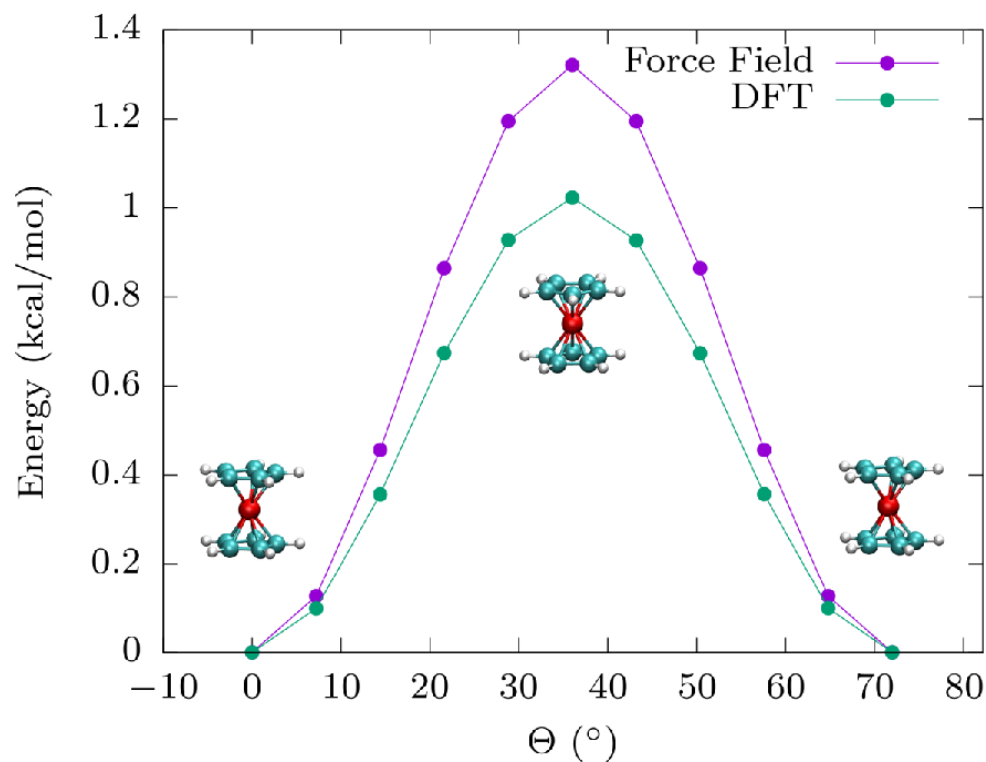
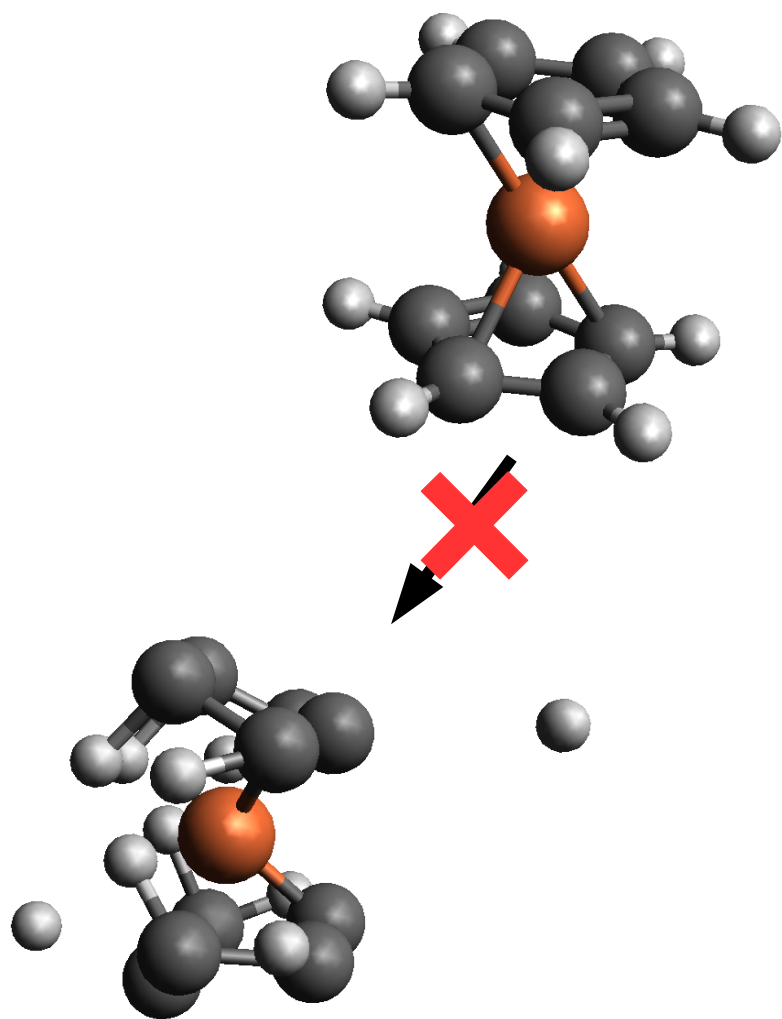


Distance between configurations:

$$d(i, j) = e^{-\frac{\sum_v (B_v(i) - B_v(j))^2}{\sigma^2}}$$

Ferrocene Fe(II)Cp_2

After 100 Refinement Configurations



Alanine: Self Consistent FF Generation

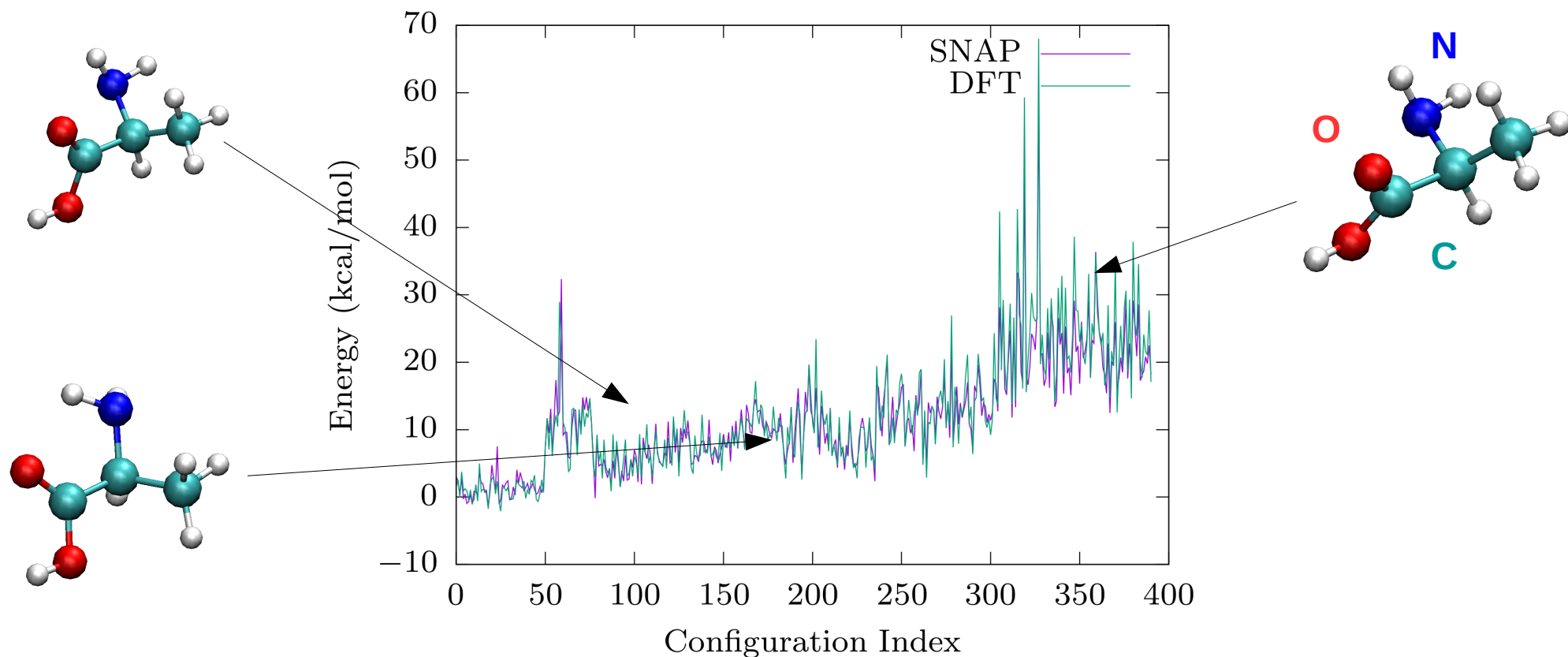
Training Set: 600 Conf

Val Set: 200 Conf

Test Set: 200 Conf

Training RMSE: 2.3 kcal/mol

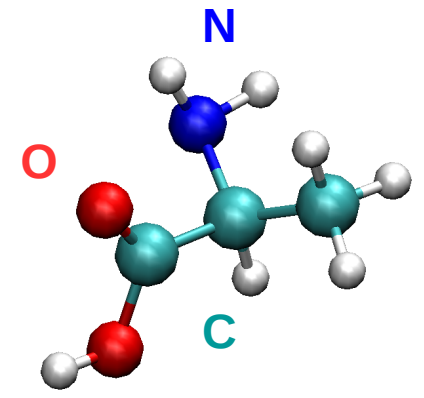
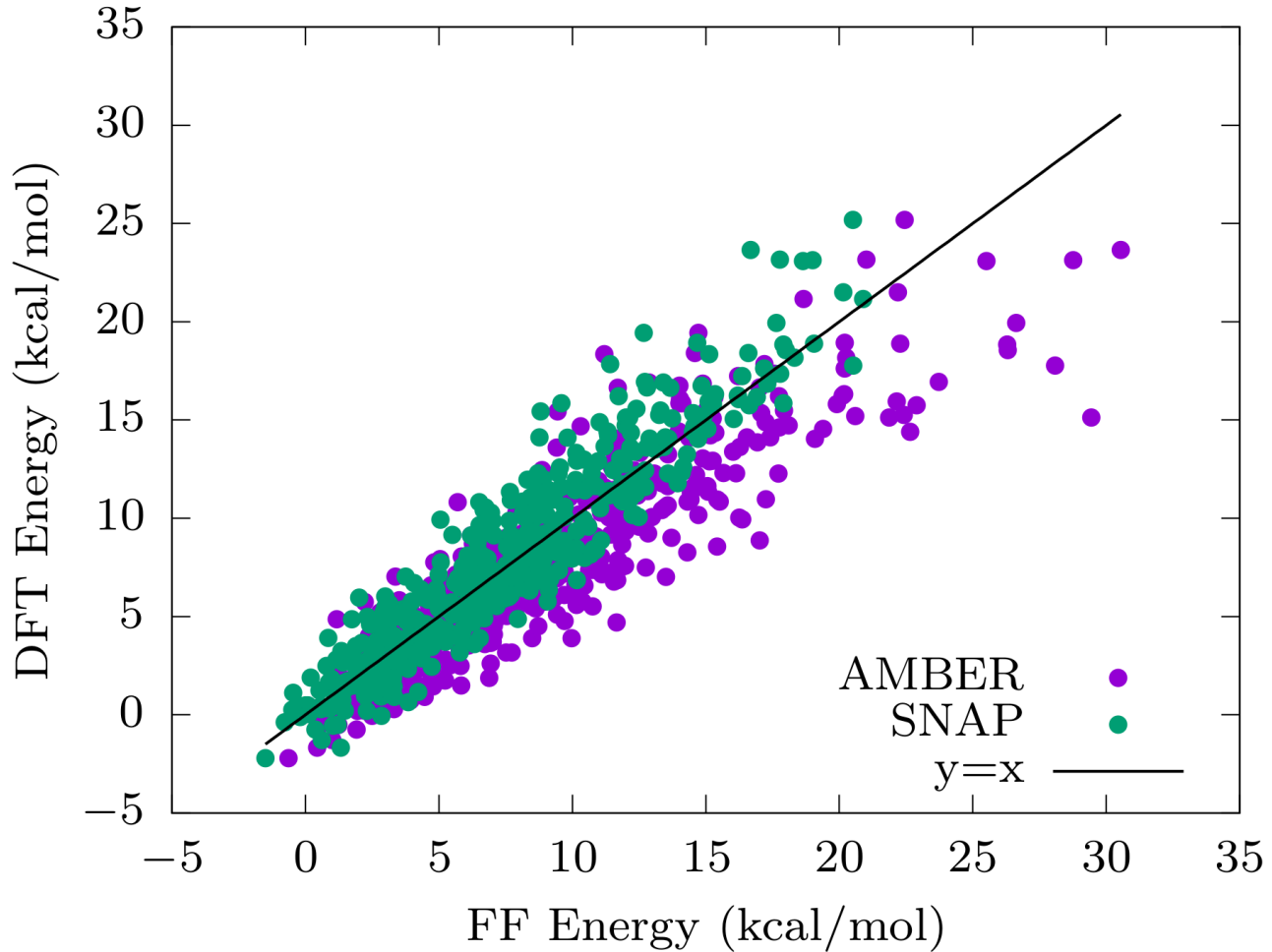
Test RMSE: 2.3 kcal/mol



More flexible molecules need longer refinements

Alanine: SNAP vs GAFF

Configurations from MD@200-400K



Anharmonic effects are automatically included

O₂-Fe(II)(Porphyrin)(Hys): Chemical Reactions

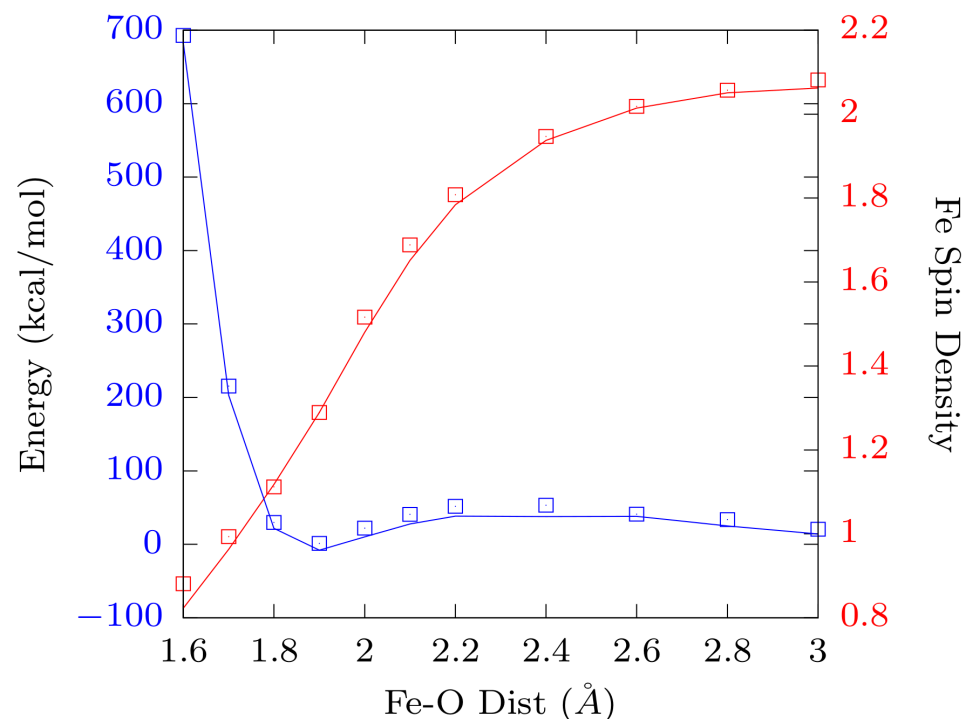
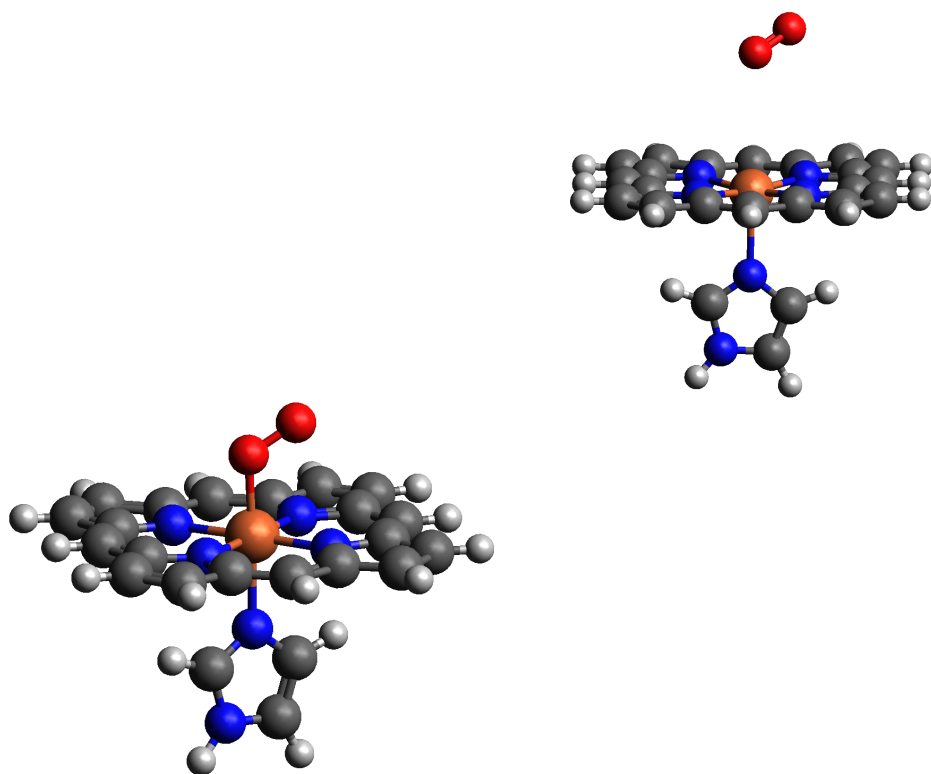
Training Set: 1300 Conf

Val Set: 140 Conf

Test Set: 150 Conf

Training RMSE: 1.2 kcal/mol

Test RMSE: 2.4 kcal/mol



Bonds breaking and prediction of other properties is possible

Conclusions

Take-Home Messages:

SNAP can be trained very easily and quickly converge wrt the training set size

SNAP can accurately describe all the covalent bond coordination environments

Self Consistent refinement of the training set efficiently enforces thermal stability

Further Development:

More robust starting training set

Description of multiple local environments

Extension to other properties other than energy

ACKNOWLEDGMENTS



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Trinity College Dublin
www.spincomp.com

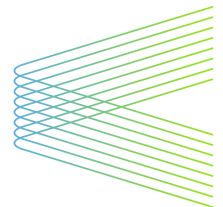


Thanks for Your kind attention

Alessandro Lunghi and Stefano Sanvito

School of Physics, CRANN and AMBER, Trinity College Dublin 2, Ireland

email: lunghia@tcd.ie



CRANN