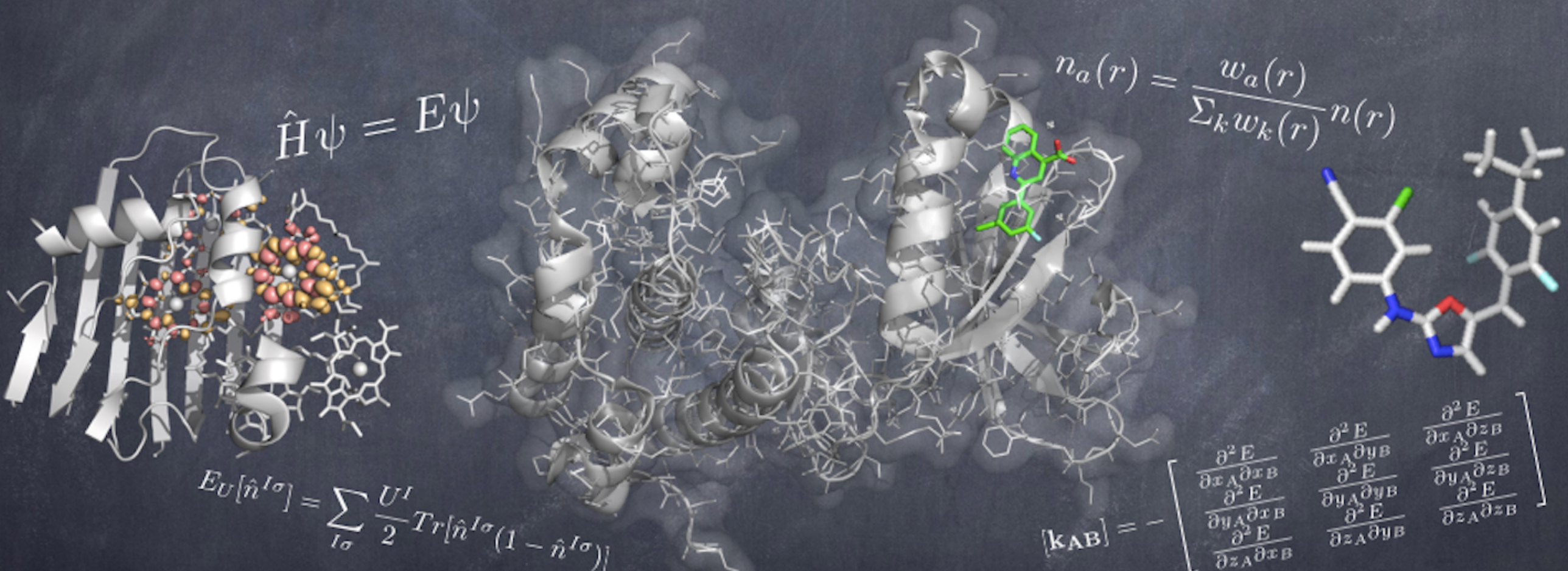


The Challenge of Biology

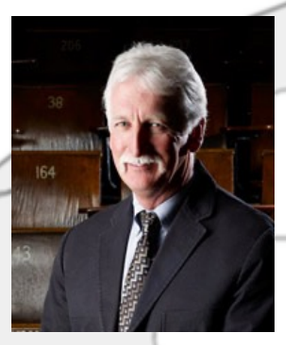
Daniel Cole
School of Natural & Environmental Sciences

Newcastle University

Biography

Prof. Mike Payne FRS

3



Prof. Bill Jorgensen

Yale

2

Newcastle University

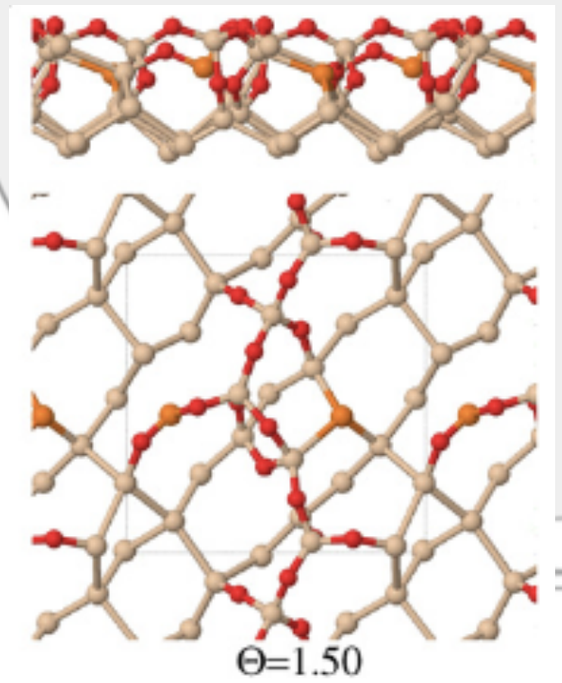
UNIVERSITY OF CAMBRIDGE

1

NUS National University of Singapore

SEVENTH FRAMEWORK PROGRAMME

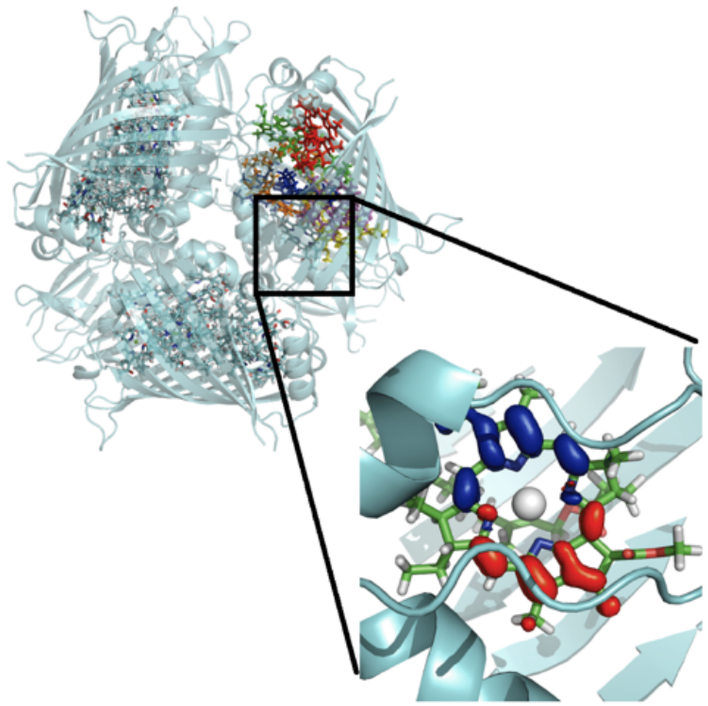
MARIE CURIE ACTIONS



DJ Cole, MC Payne, L Colombi Ciacchi, Surf Sci, (2007) 601, 488

Why did I keep going back?

Alex Chin, group leader,
Sorbonne Université



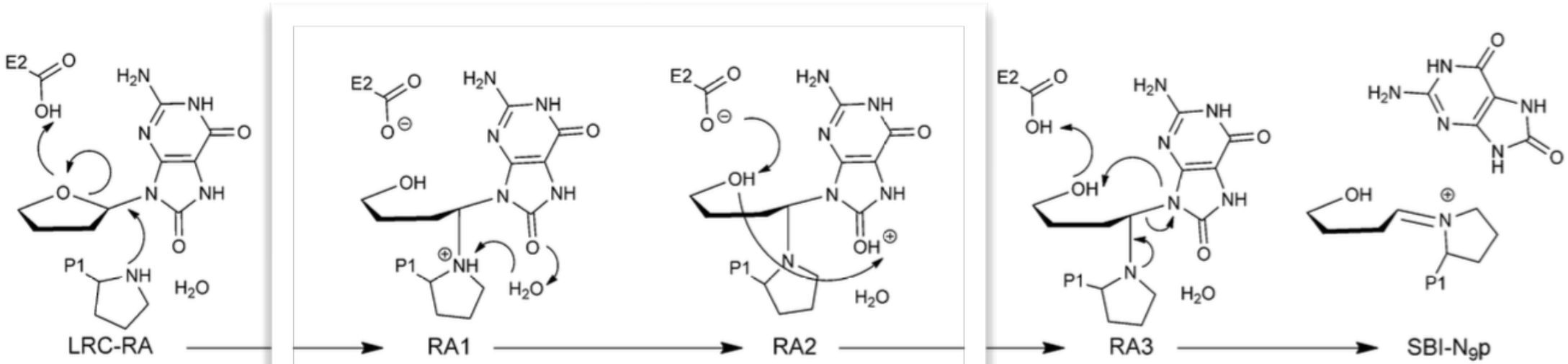
James Kermode, reader,
University of Warwick

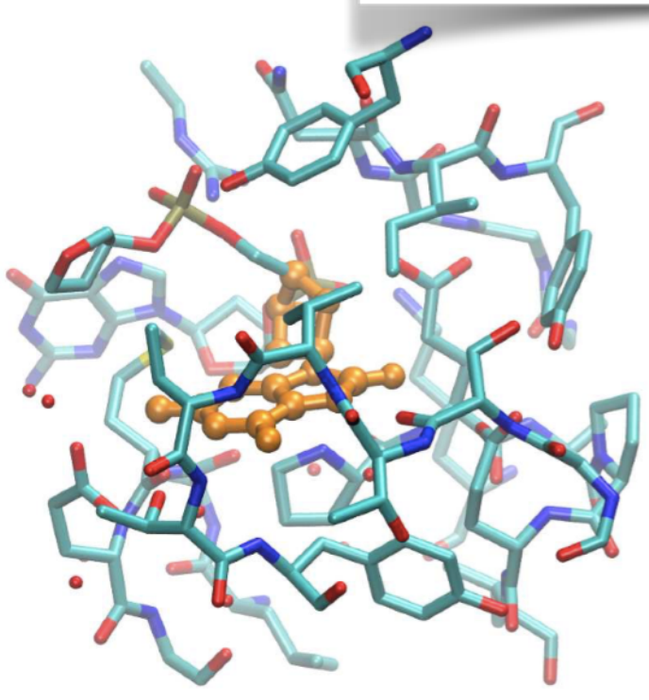
William Belfield: constrained geometric simulation of ion channel proteins
Louis Lee: electronic structure analysis in ONETEP, including protein-specific charges
Greg Lever: computational enzymology
Edward Linscott: strong correlation effects in metalloproteins
Alice Allen: biomolecular force field development
Lupeng Yang: excited state force fields

Can the same methods that we've been hearing about today be used to address problems in biology? (last stage of the grand agenda?)

The Length Scale Problem

Deprotonation of proline in the second stage of repair of oxidised guanine by the bacterial glycosylase, MutM

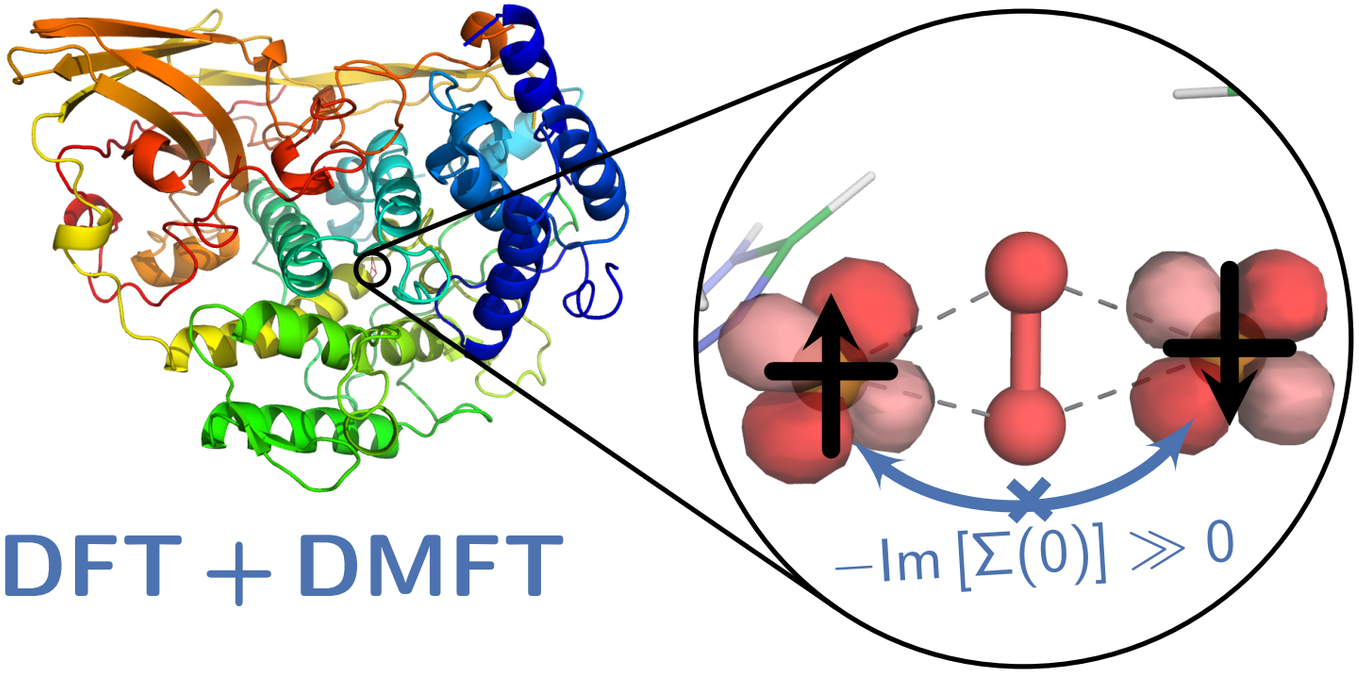




QM atoms	Barrier (kcal/mol)
143	28
278	6
493	14
606	14

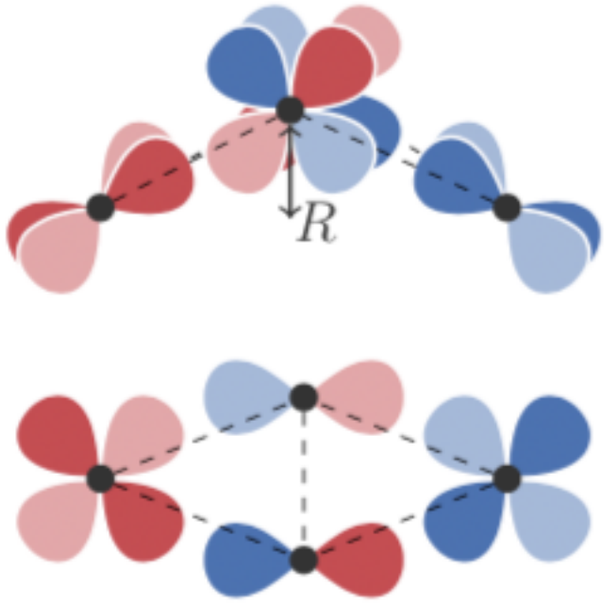
Sadeghian et al., Angew. Chem. Int. Ed. 53, 10044 (2014)

Strong correlations in metalloproteins



di-Cu oxo bridge in hemocyanin

DFT + DMFT

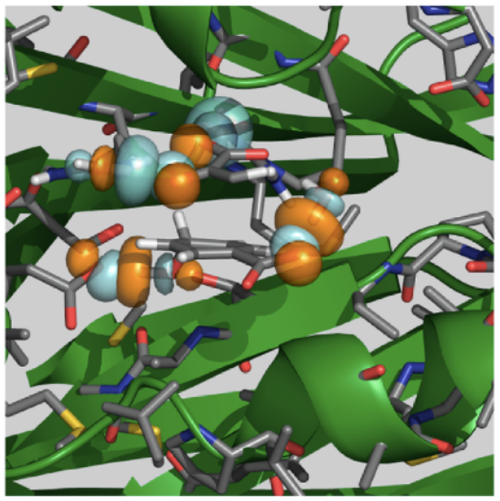


In hemocyanin, experimental observation of open shell singlet state is at odds with DFT calculations on Cu-O_2-Cu core.

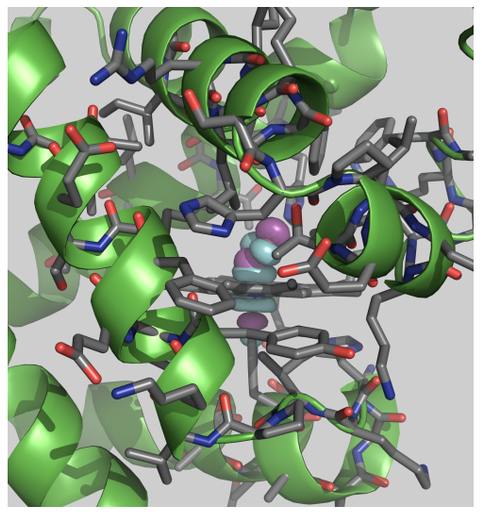
Using DFT+DMFT found that singlet is superposition of two localised magnetic moments in Heitler-London regime.

al-Badri MA, Linscott E, et al, *Communications Physics*, in press (2020)

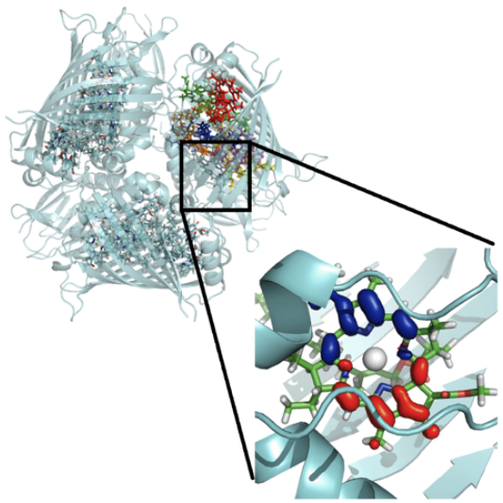
Biological Applications



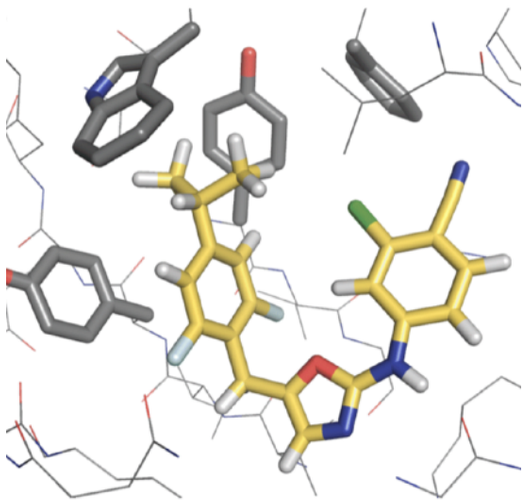
Transition state searching in enzymes



Protein-ligand binding in metalloproteins



Optical spectroscopy in a light-harvesting protein

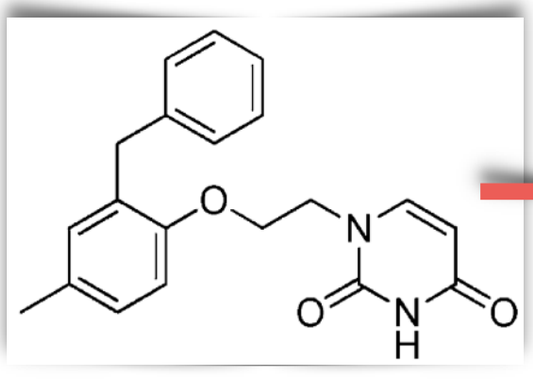


Classical force field parameterisation for drug discovery

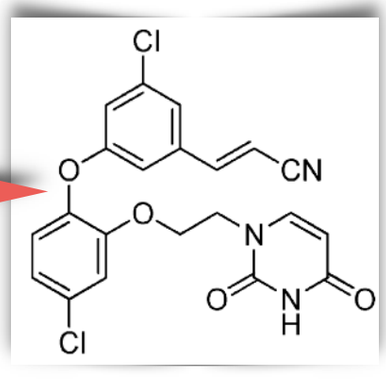
"Applications of Large-Scale DFT in Biology", D. J. Cole & N. D. M. Hine
Journal of Physics: Condensed Matter, 28, 393001 (2016)

Computer-Aided Drug Design

5μM docking hit



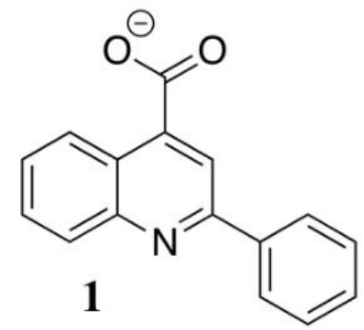
55pM inhibitor

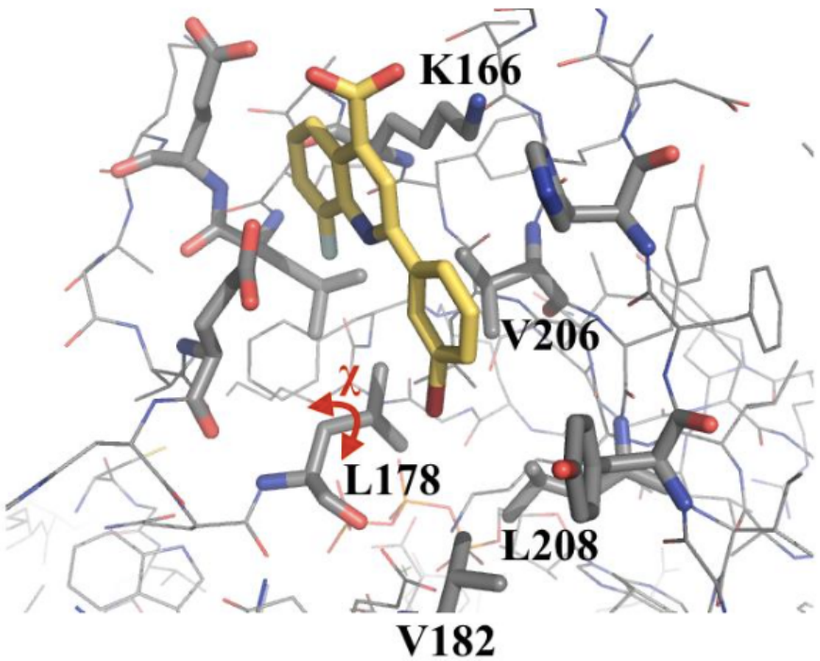


Notable successes include computationally-guided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase (NNRTIs).

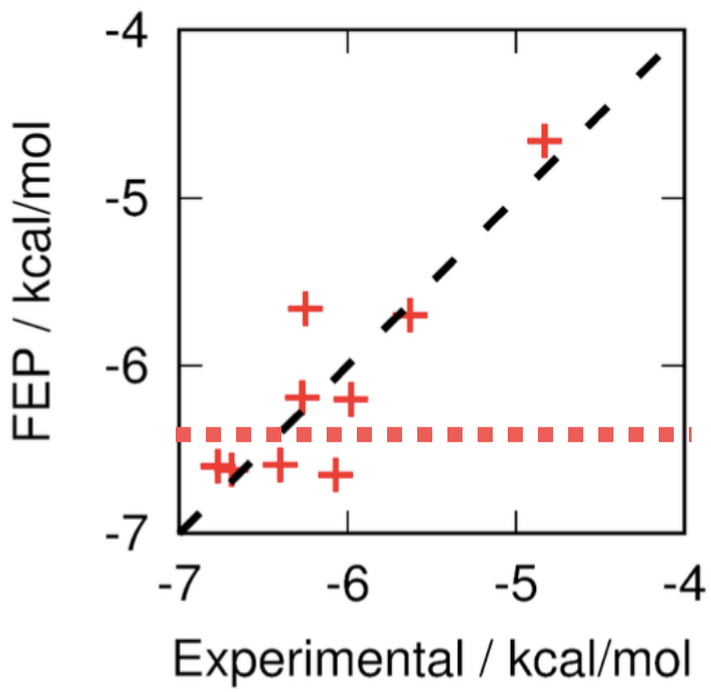
M. Bollini *et al.*, *J. Med. Chem.* **2011**, *54*, 8582

W. G. Lee *et al.*, *J. Am. Chem. Soc.* **2013**, *135*, 16705





aurora kinase

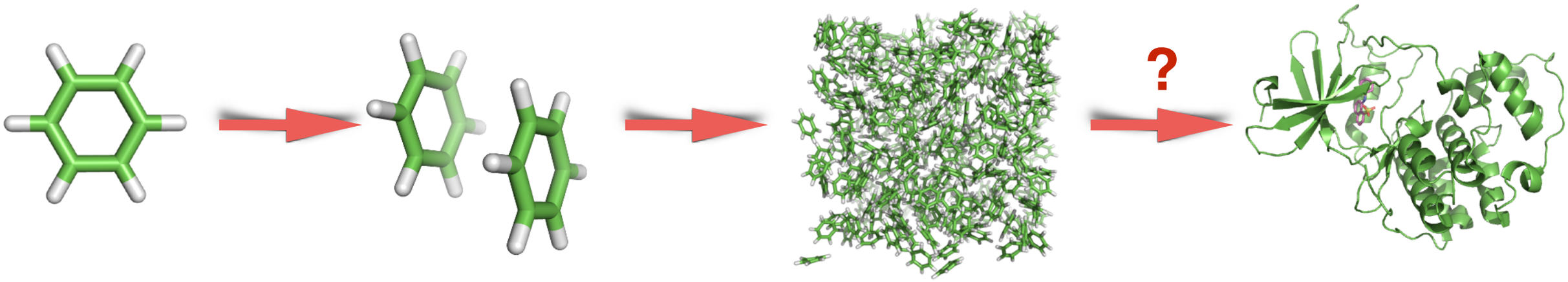


RMSE = 0.32 kcal/mol

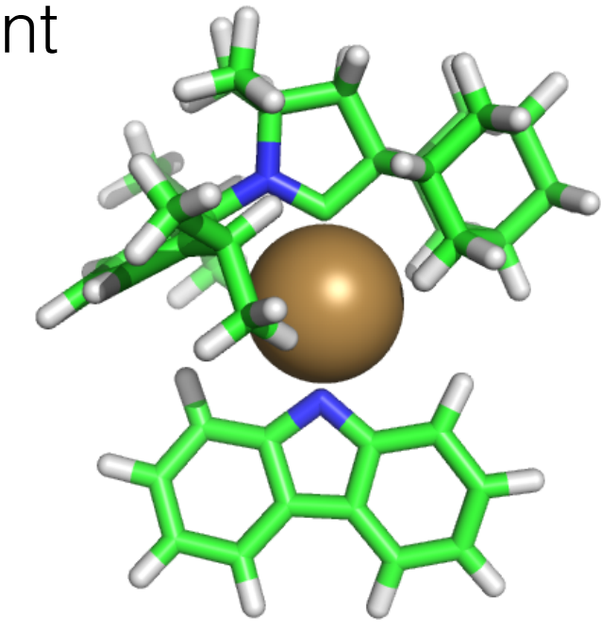
D. J. Cole et al., *Chem. Commun.* **2017**, *53*, 9372

Disadvantages of Force Field Design

Transferability: Force field parameters are obtained by fitting to properties of small molecules:



Labour-intensive: Typically requires large groups, significant time to produce a model useful for practical problems.

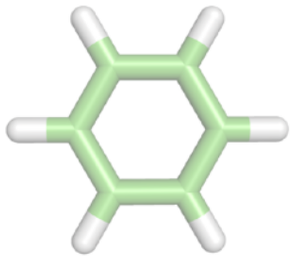


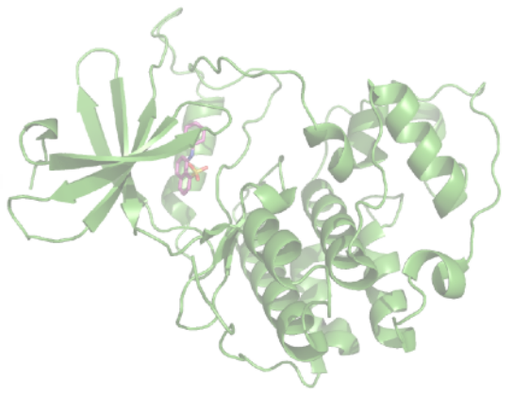
Functional form: Is all of the required physics contained within the simple functional form of MM force fields?

Ease of use: Missing parameters are problematic for inexperienced force field users (metals, excited states?).

Disadvantages of Force Field Design

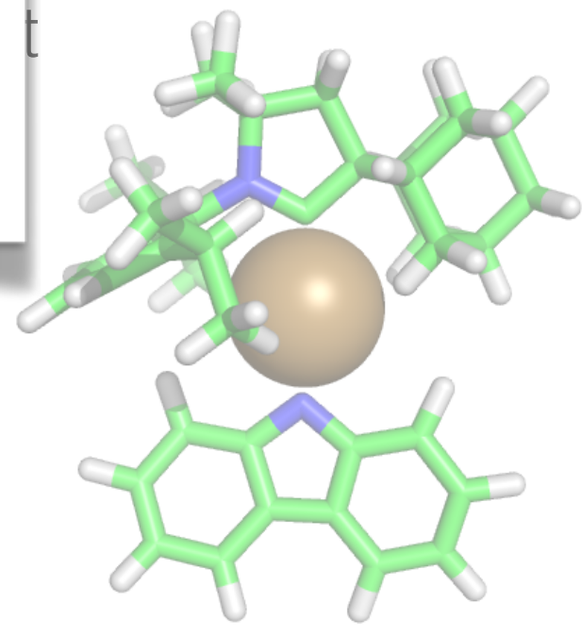
Transferability: Force field parameters are obtained by fitting to properties of small molecules:





> "Classical force field parameterization is a bottomless pit of despair"
>
> Twitter user

Labour-intensi t
time to prod



Functional form: Is all of the required physics contained within the simple functional form of MM force fields?

Ease of use: Missing parameters are problematic for inexperienced force field users (metals, excited states?).

QUBE Force Field

$$E^{QM} = \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle$$

ab initio

electrostatics
induction
dispersion
exchange-repulsion

accuracy

system size

QUantum mechanical BEspoke (QUBE) force field

fast to evaluate
large system sizes

Transferable force field

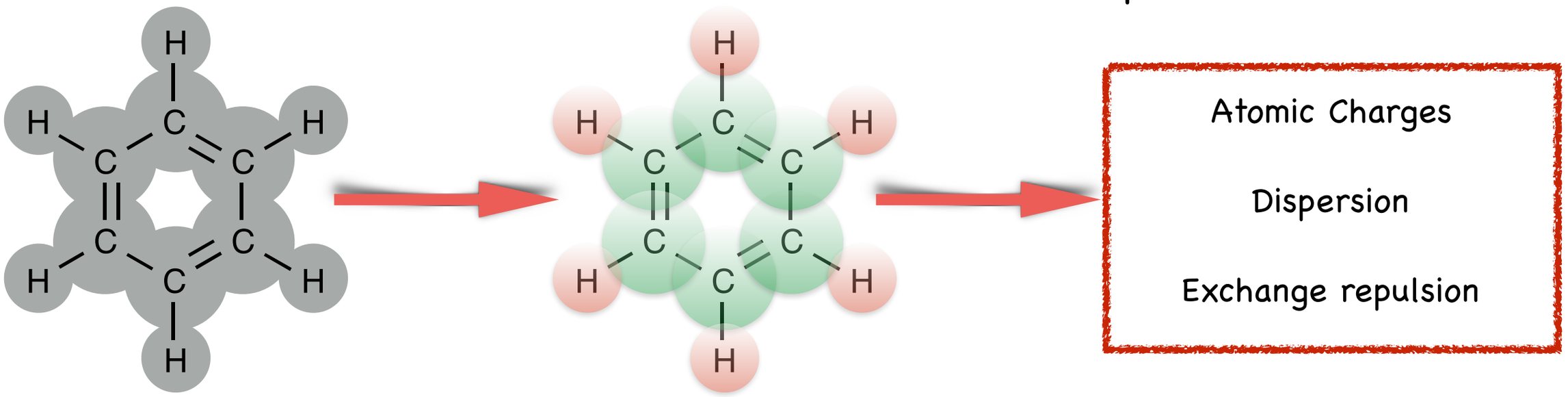
$$E^{MM} = E_{\text{bonded}} + E_{\text{non-bonded}} = E_{\text{bonded}} + \sum_{i<j} \frac{q_i q_j}{r_{ij}} + \left(\frac{A^{ij}}{r_{ij}^{12}} - \frac{C_6^{ij}}{r_{ij}^{6}} \right)$$

Atoms-in-Molecule

DFT Calculation computes total electron density

Electrons partitioned amongst the atoms in the system

Atomistic force field parameters computed directly from partitioned electron density



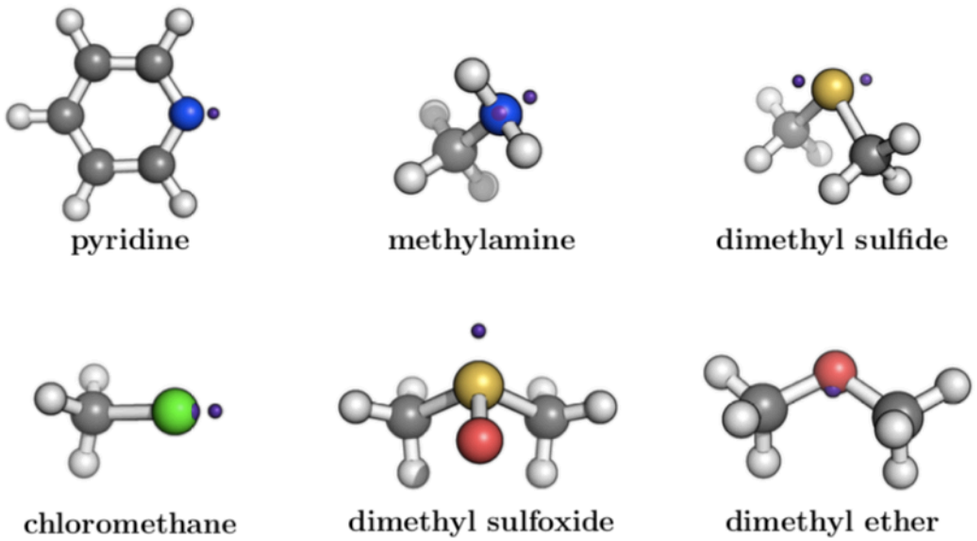
Density derived electrostatic and chemical (DDEC) electron density partitioning (good reproduction of the ESP and not too conformation dependent).

Charges are computed in implicit solvent to account for induction effects.

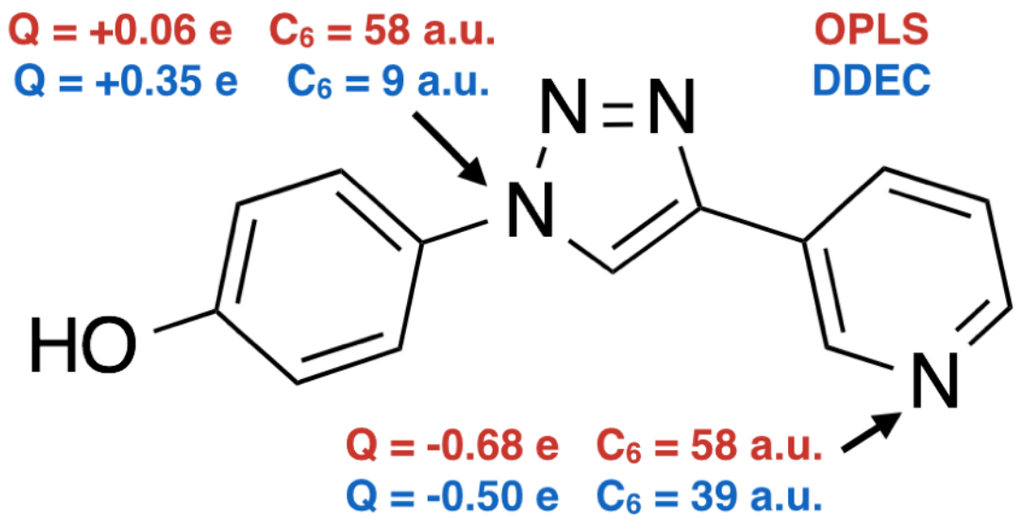
D. Cole, J. Vilseck, J. Tirado-Rives, M. Payne, W. Jorgensen, *JCTC,* **2016**, *12,* 2312

Force Field Parameters

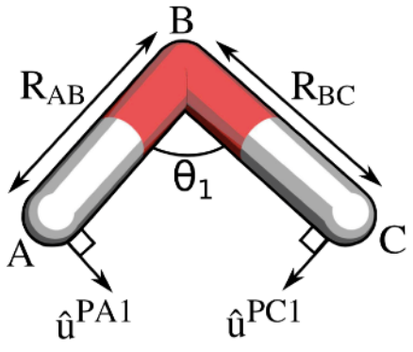
Off site charges (available in ONETEP)



Lennard-Jones (from Tkatchenko-Scheffler approach)

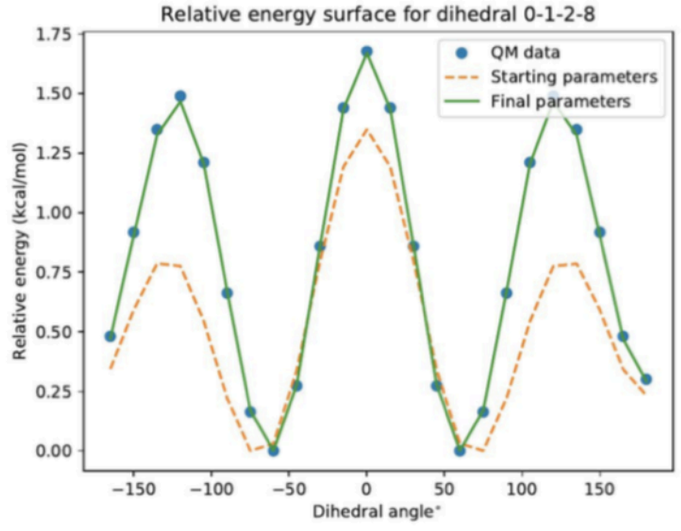


Bonded parameters (modified Seminario)



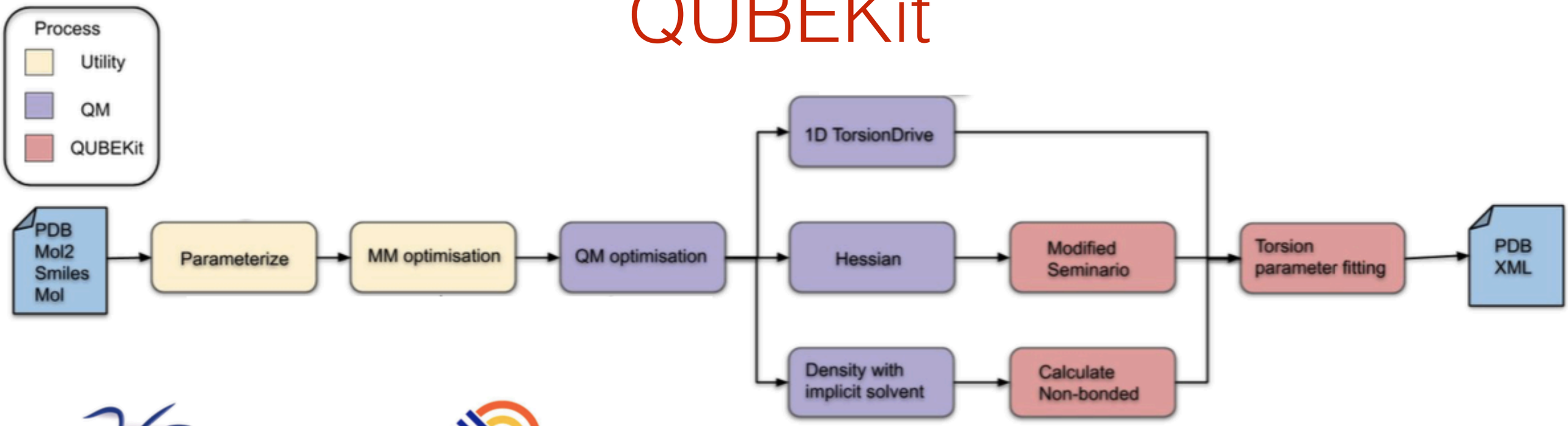
$$[\mathbf{k_{AB}}] = -\begin{vmatrix} \frac{\partial^2 E}{\partial x_A \partial x_B} & \frac{\partial^2 E}{\partial x_A \partial y_B} & \frac{\partial^2 E}{\partial x_A \partial z_B} \\ \frac{\partial^2 E}{\partial y_A \partial x_B} & \frac{\partial^2 E}{\partial y_A \partial y_B} & \frac{\partial^2 E}{\partial y_A \partial z_B} \\ \frac{\partial^2 E}{\partial z_A \partial x_B} & \frac{\partial^2 E}{\partial z_A \partial y_B} & \frac{\partial^2 E}{\partial z_A \partial z_B} \end{vmatrix}$$

Torsion parameters from QM scans



Horton JT, Allen AEA, Dodda L, Cole DJ. *J Chem Inf Model* 2019, **59**(4), 1366-1381.

QUBEKit



PSI4
AB INITIO QUANTUM CHEMISTRY

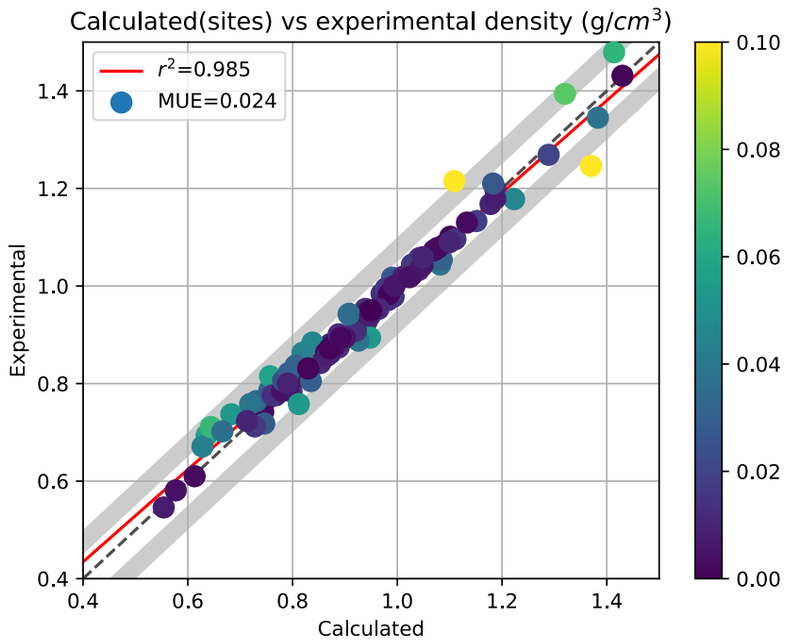
geomeTRIC

QUBE Kit

Chargemol

QC Archive

https://github.com/cole-group/QUBEKit

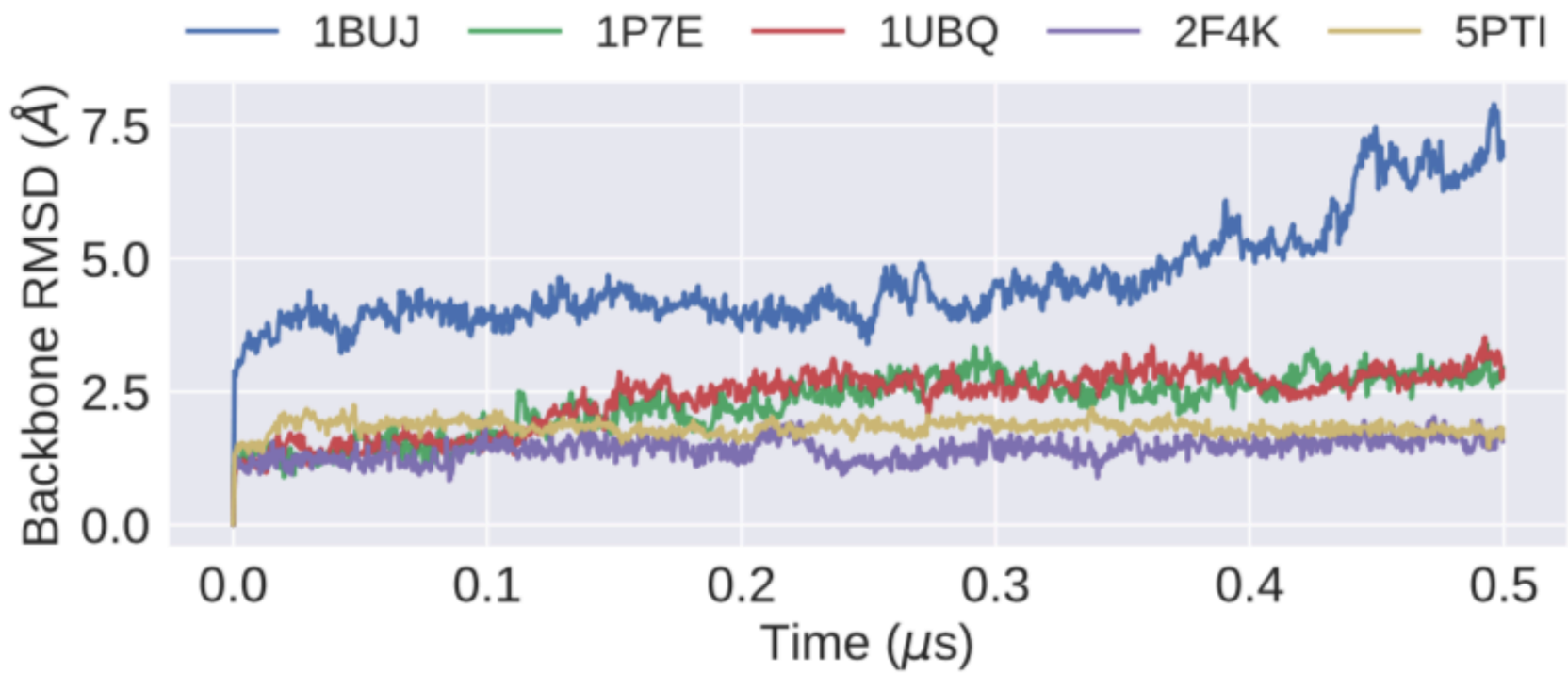


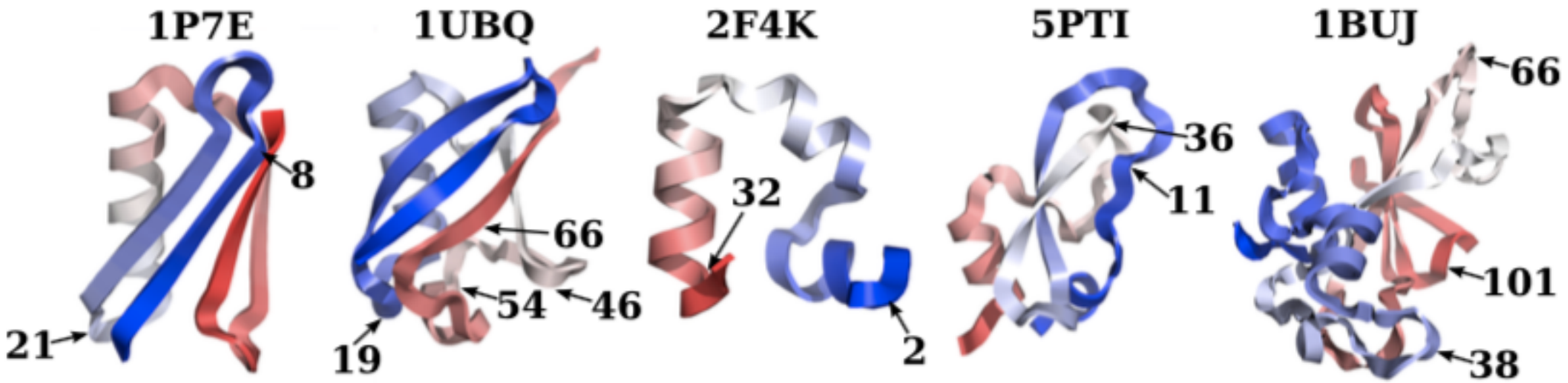
Josh Horton, Chris Ringrose

Horton JT, Allen AEA, Dodda L, Cole DJ. *J Chem Inf Model* 2019, **59**(4), 1366-1381.

QUBE Protein Force Field

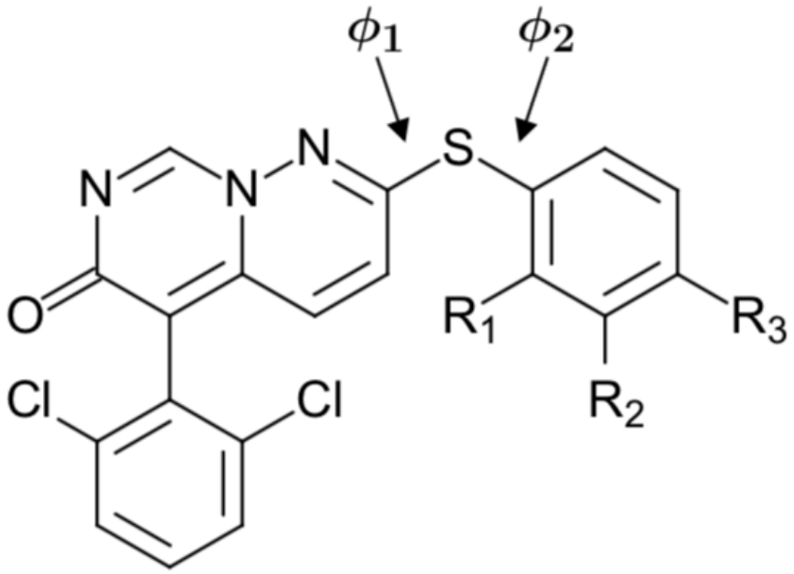
RMSD relative to crystal structure (3 x 0.5μs MD):

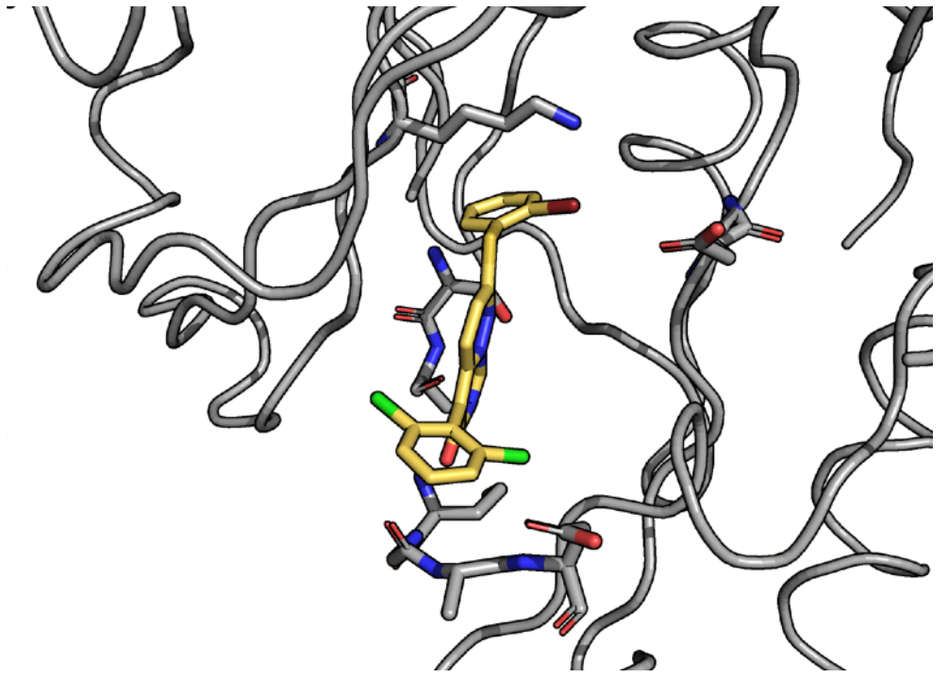




Allen AEA, Robertson MJ, Payne MC, Cole DJ, *ACS Omega* 2019, **4**, 14537-14550.

p38 Kinase FEP

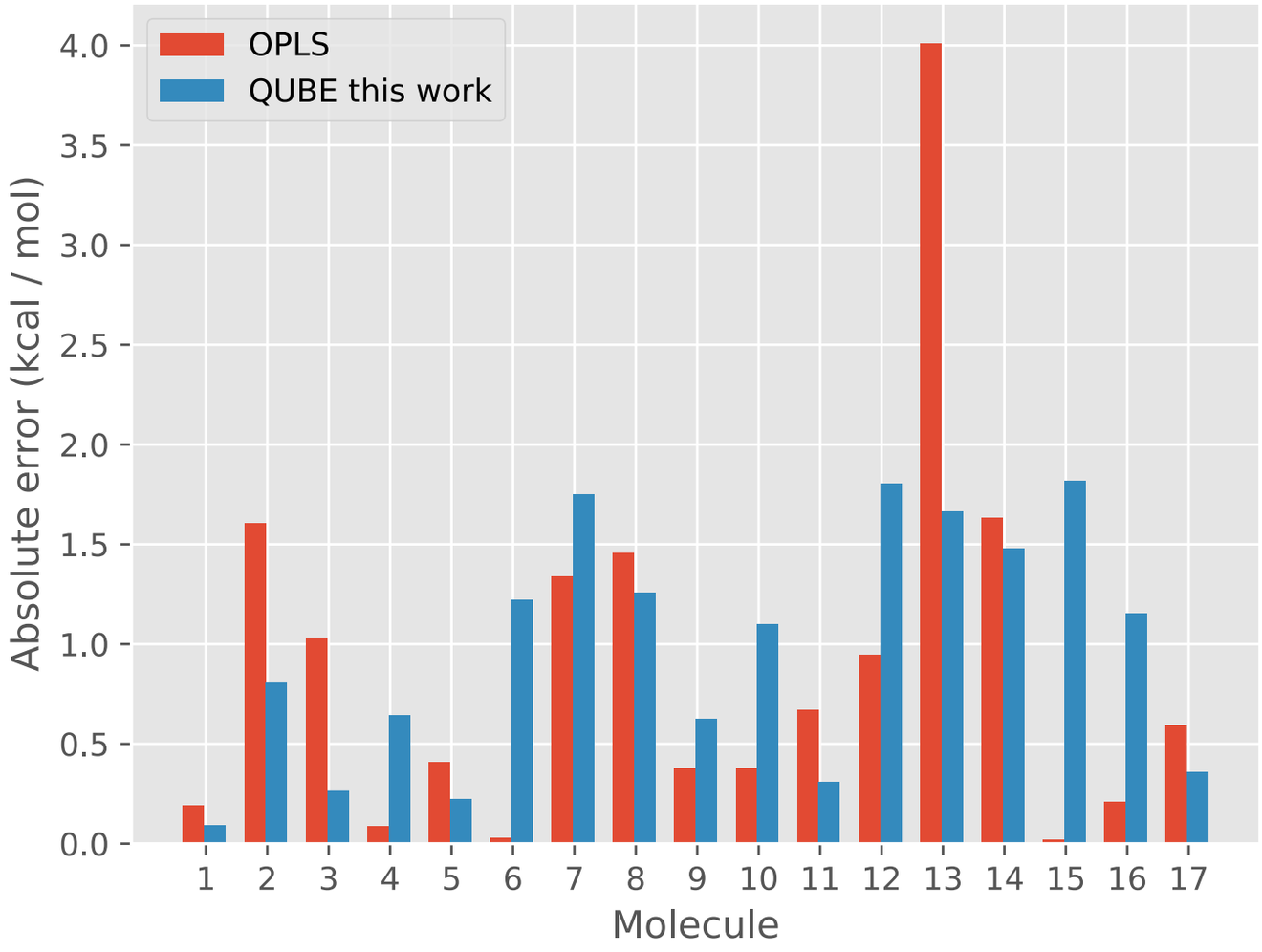


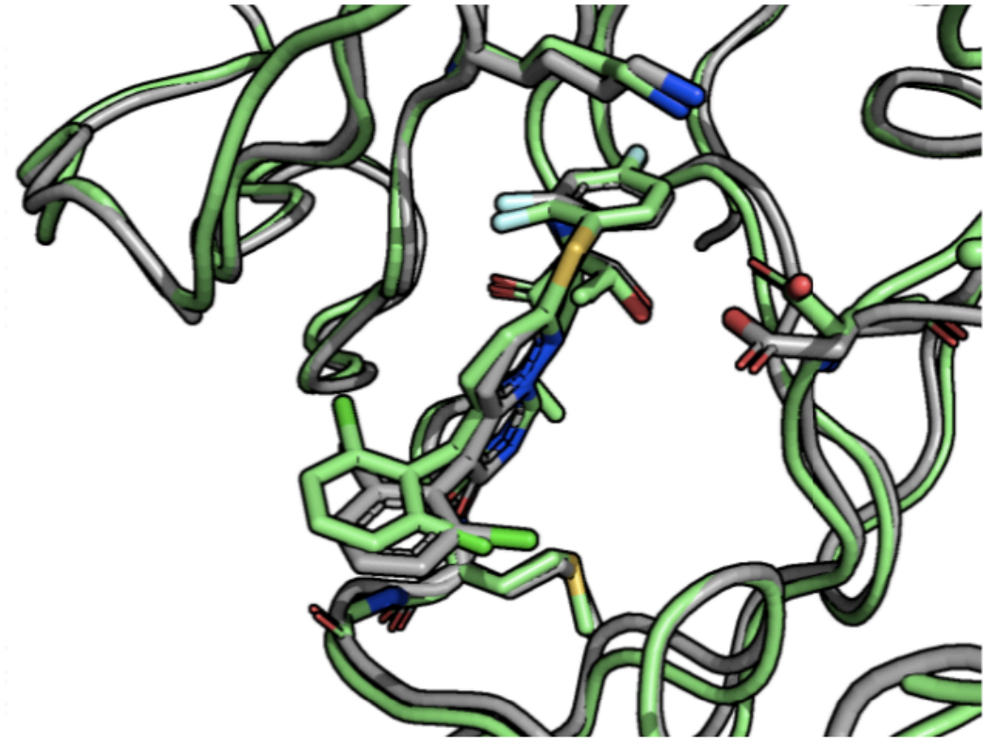


Compound	R1	R2	R3	pIC_{50}
1	H	H	H	6.602
2	H	H	F	7.000
3	H	H	CH_3	5.854
4	H	Cl	Cl	6.097
5	H	CH_3	H	5.854
6	H	CH_3	CH_3	5.721
7	H	F	H	6.347
8	CH_3	H	H	6.699
9	H	Cl	F	6.301
10	H	Cl	H	6.553
11	CH_3	H	Cl	6.745
12	Br	H	H	6.602
13	CH_3	H	CH_3	6.577
14	OH	H	H	6.444
15	NH_2	H	F	6.658
16	Cl	H	F	7.444
17	F	F	F	8.046
18	F	H	H	N/A

Horton JT, Allen AEA, Cole DJ. *ChemComm (2019) doi:10.1039/C9CC08574B*

p38 Kinase FEP





RMSE(OPLS) = **1.3** kcal/mol
RMSE(QUBE) = **1.1** kcal/mol

Good structural/energetic agreement for 1st generation MM force field.

Horton JT, Allen AEA, Cole DJ. *ChemComm (2019) doi:10.1039/C9CC08574B*

Future Work

Continuing development of **QUBEKit** to improve ease-of-use and widen applicability.

Interfaces with **Open Force Field Initiative** infrastructure will give access to large quantum chemistry databases for force field derivation and workflows for force field validation.

Investigating **machine learning** approaches for learning both force field parameters or entire potential energy surfaces from large quantum chemistry datasets.

There are very few empirical parameters, so the force field can easily be **re-designed** with a different functional form. E.g. we have been investigating the derivation of screened C_6, C_8 and polarisability coefficients from QM.

Potential to move into areas of **unmet need** for force fields,
e.g. metals, excited state potential energy surfaces, transition states.

Manz T, Chen T, Cole DJ, Gabaldon Limas N, Fiszbein B. *RSC Advances* 2019, **9**, 19297-19324.

https://blogs.ncl.ac.uk/danielcole/
https://github.com/cole-group/
@ColeGroupNCL

Newcastle University

Thank you for your attention