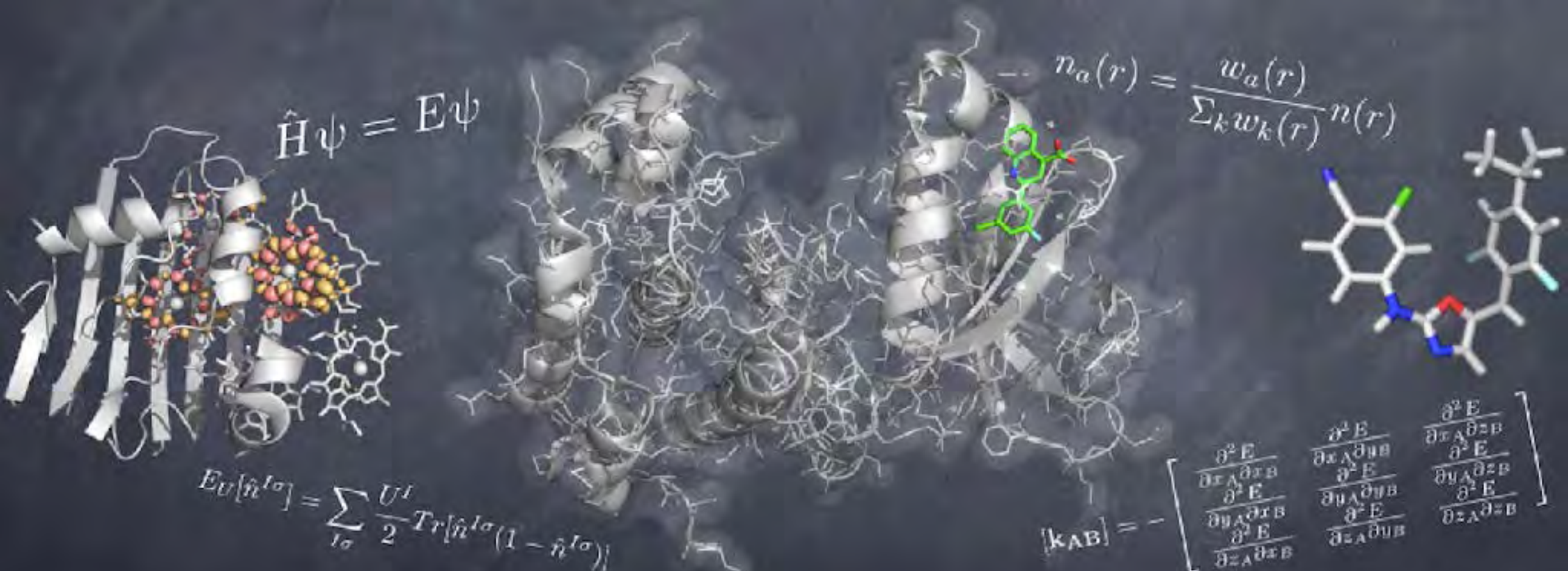


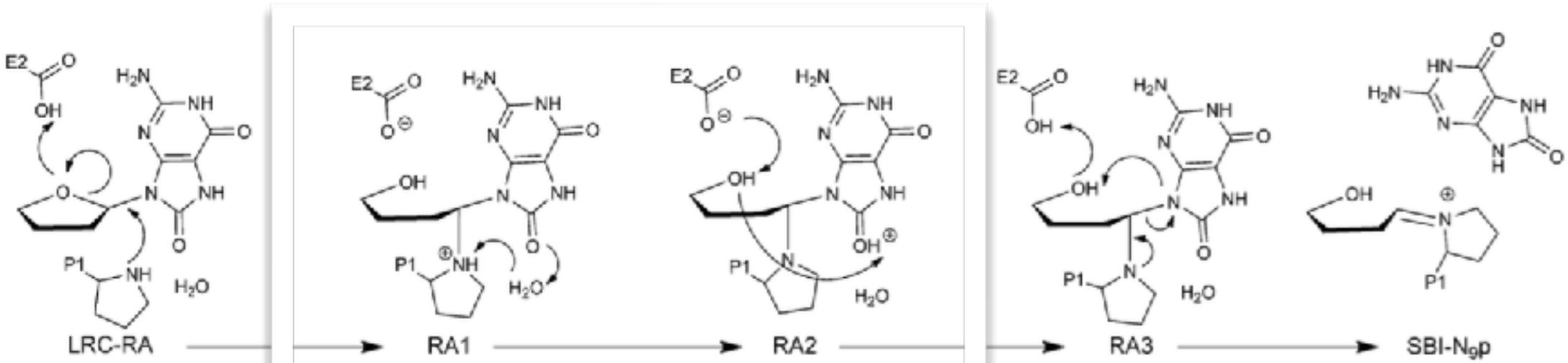
Simulating Biology with ONETEP

Daniel Cole
School of Natural & Environmental Sciences

Newcastle University

The Length Scale Problem

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



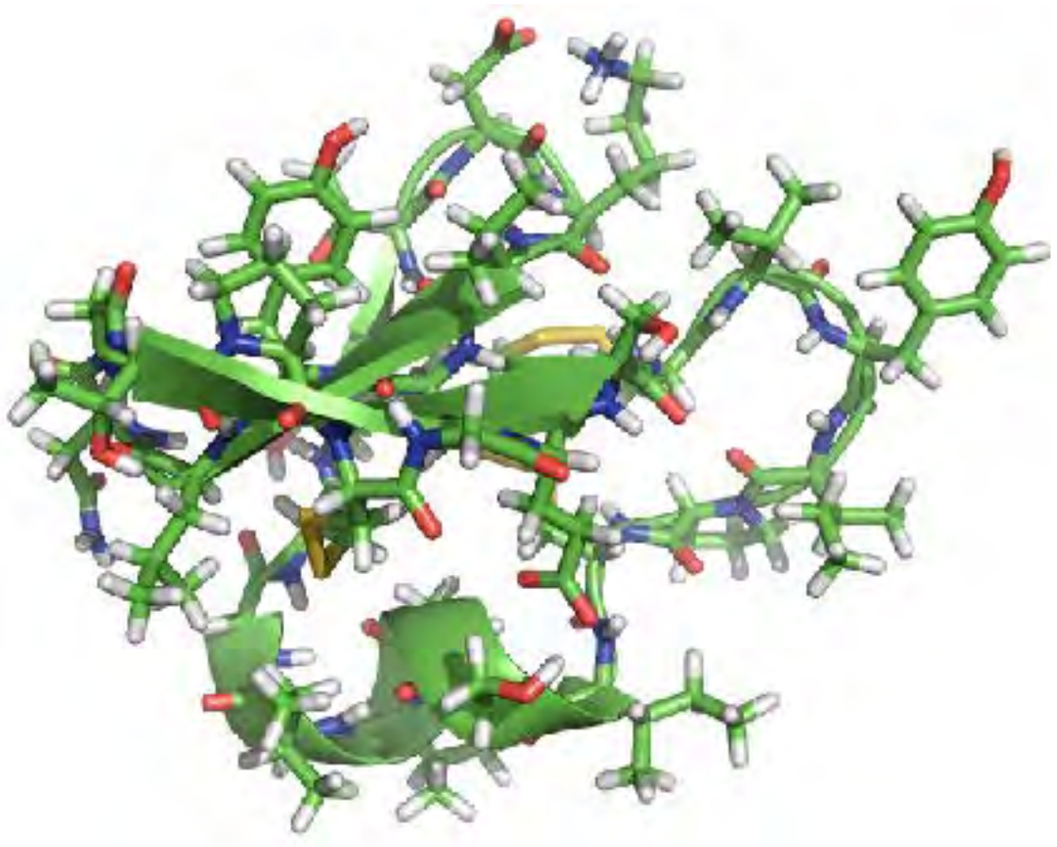
QM atoms	Barrier (kcal/mol)
143	28
278	6
493	14
606	14

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

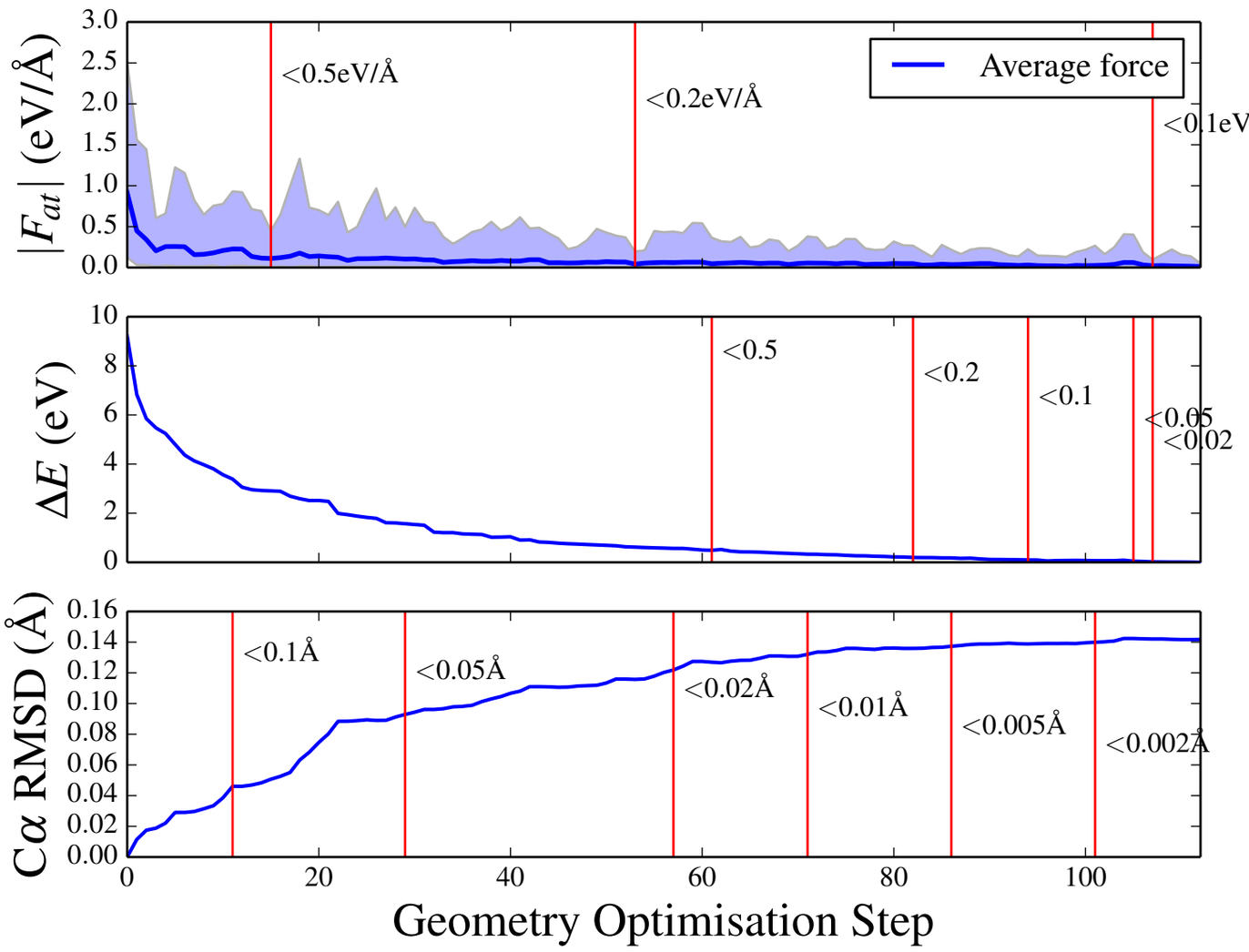
Protein Structure

Cyclotide kalata B5 (PDB: 2KUX)

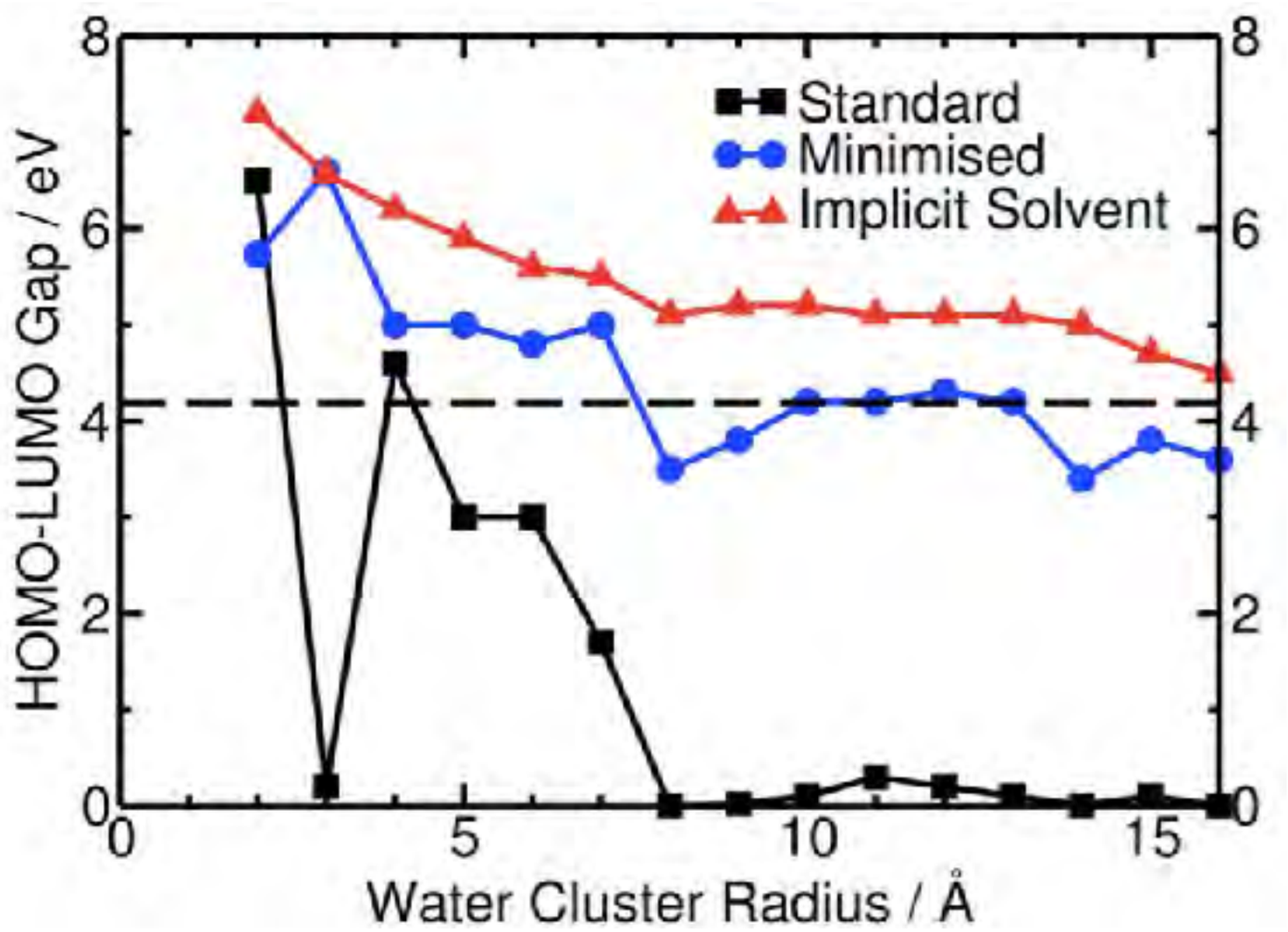
Structurally optimised using ONETEP (implicit solvent)

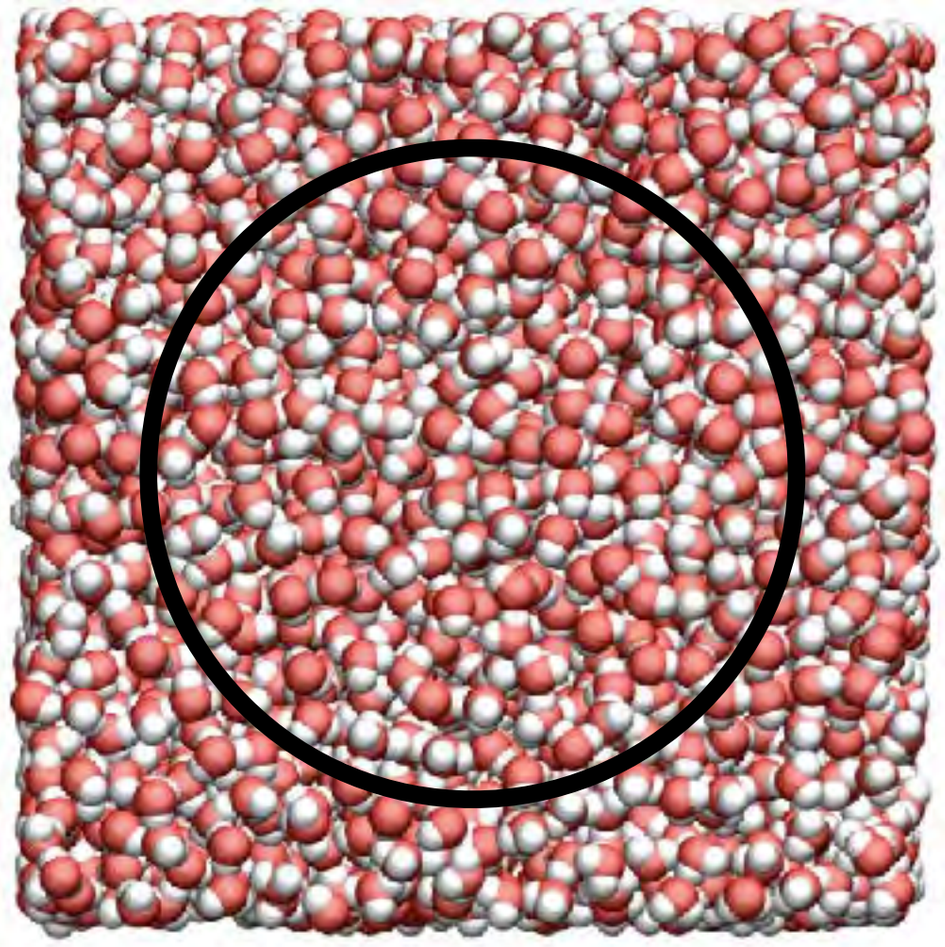


RMSD = 0.14 Å



HOMO-LUMO Gap

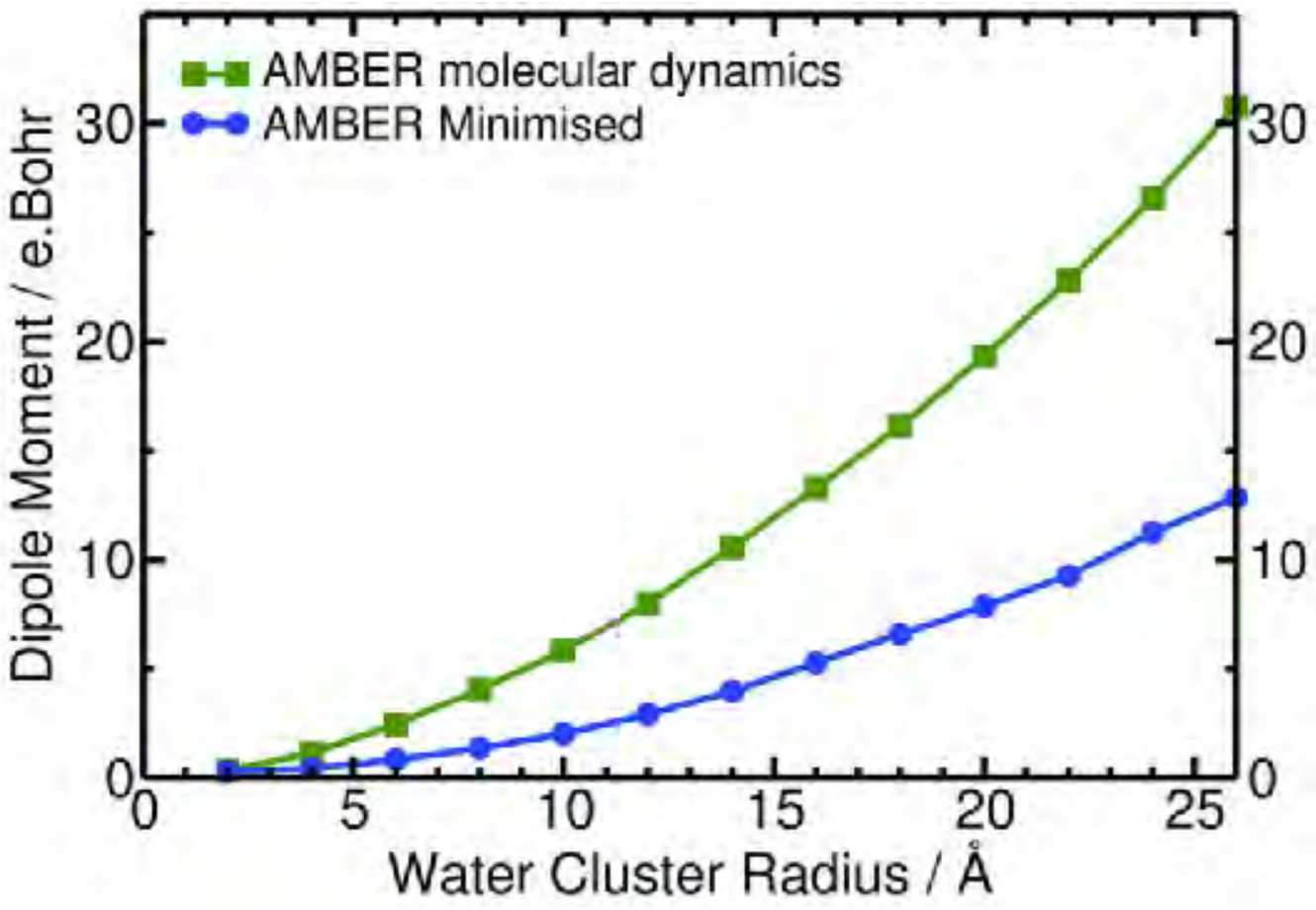


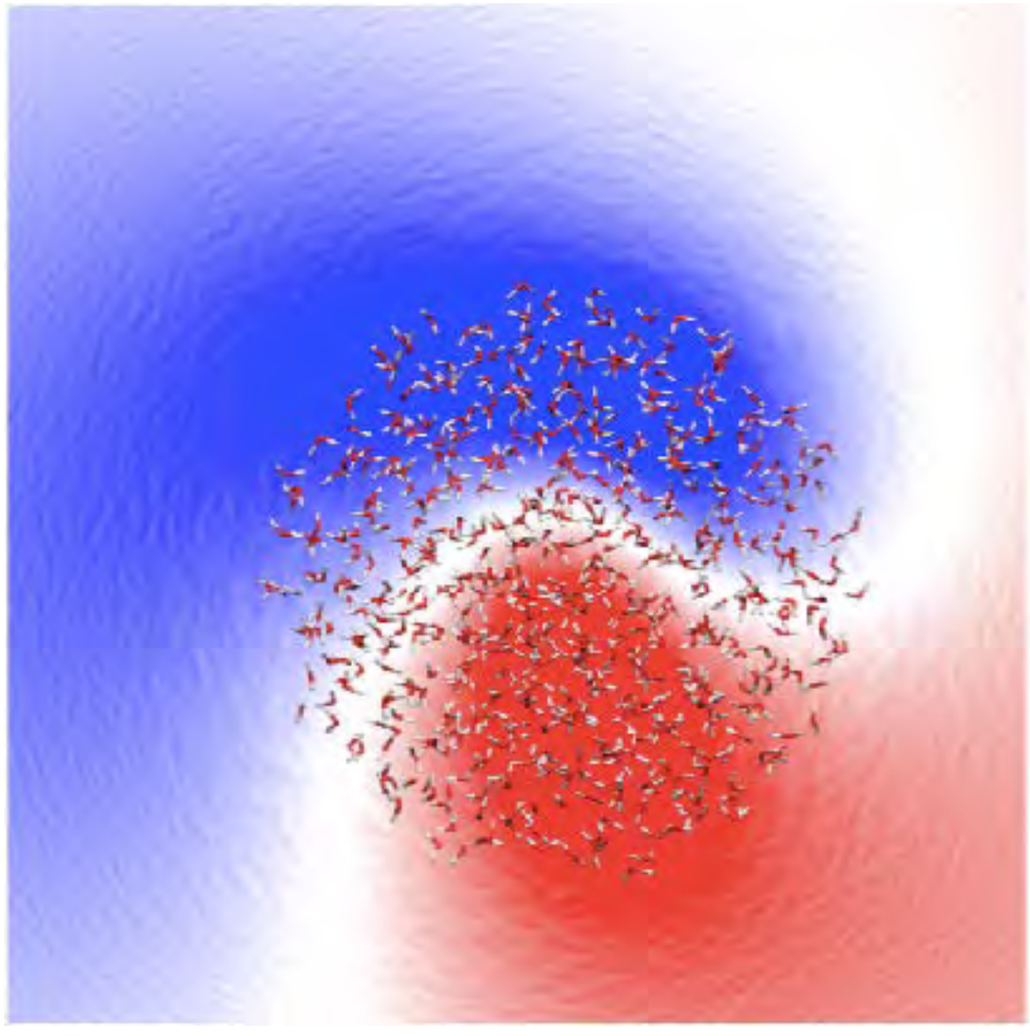


We know that DFT under-estimates band gap, but no reason why the effect should be worse for large systems?

G. Lever, D. Cole, N. Hine, P. Haynes, M. Payne, *J. Phys. Cond. Matt.*, **2013**, *25*, 152101

Dipole moment

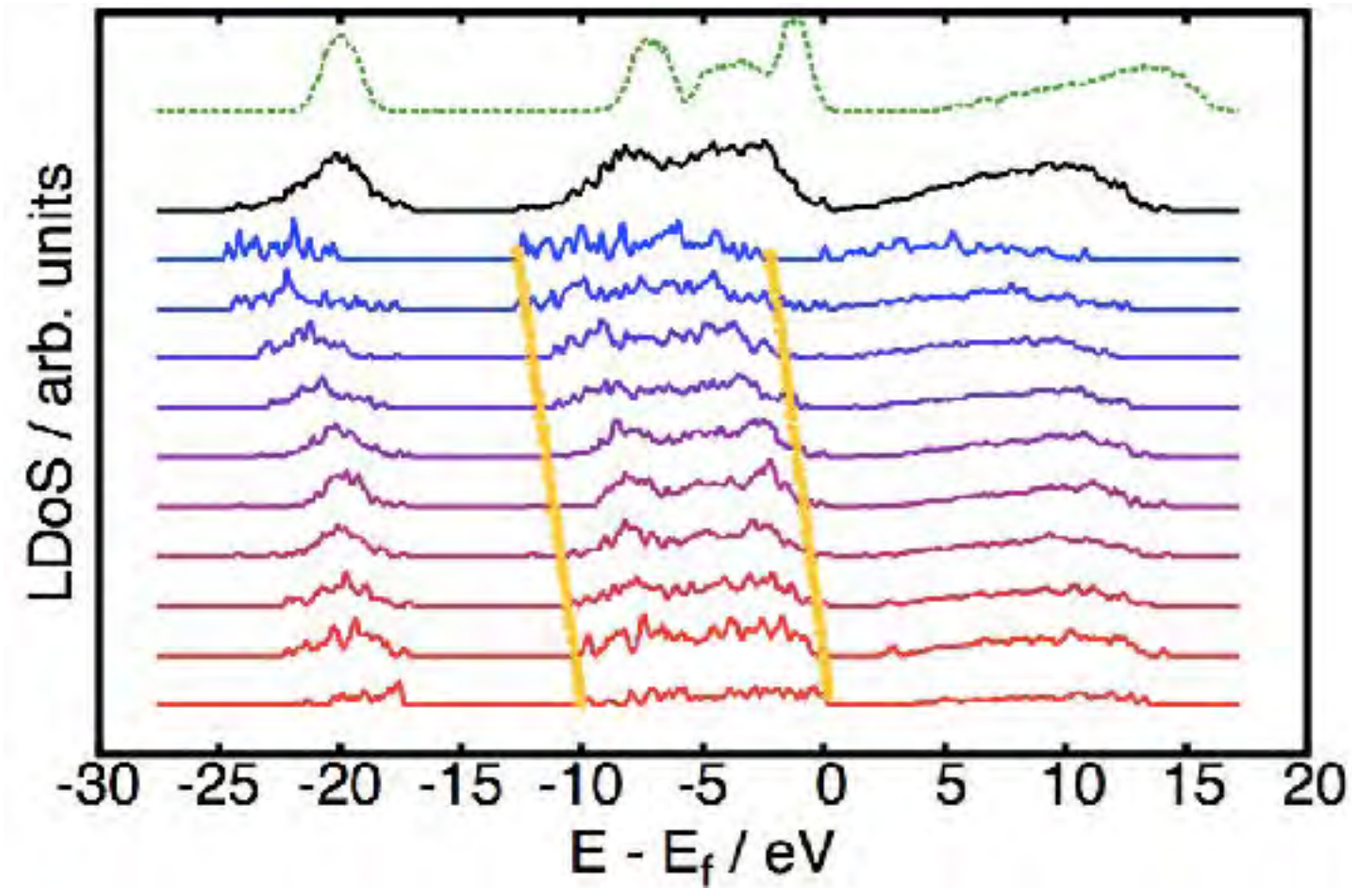


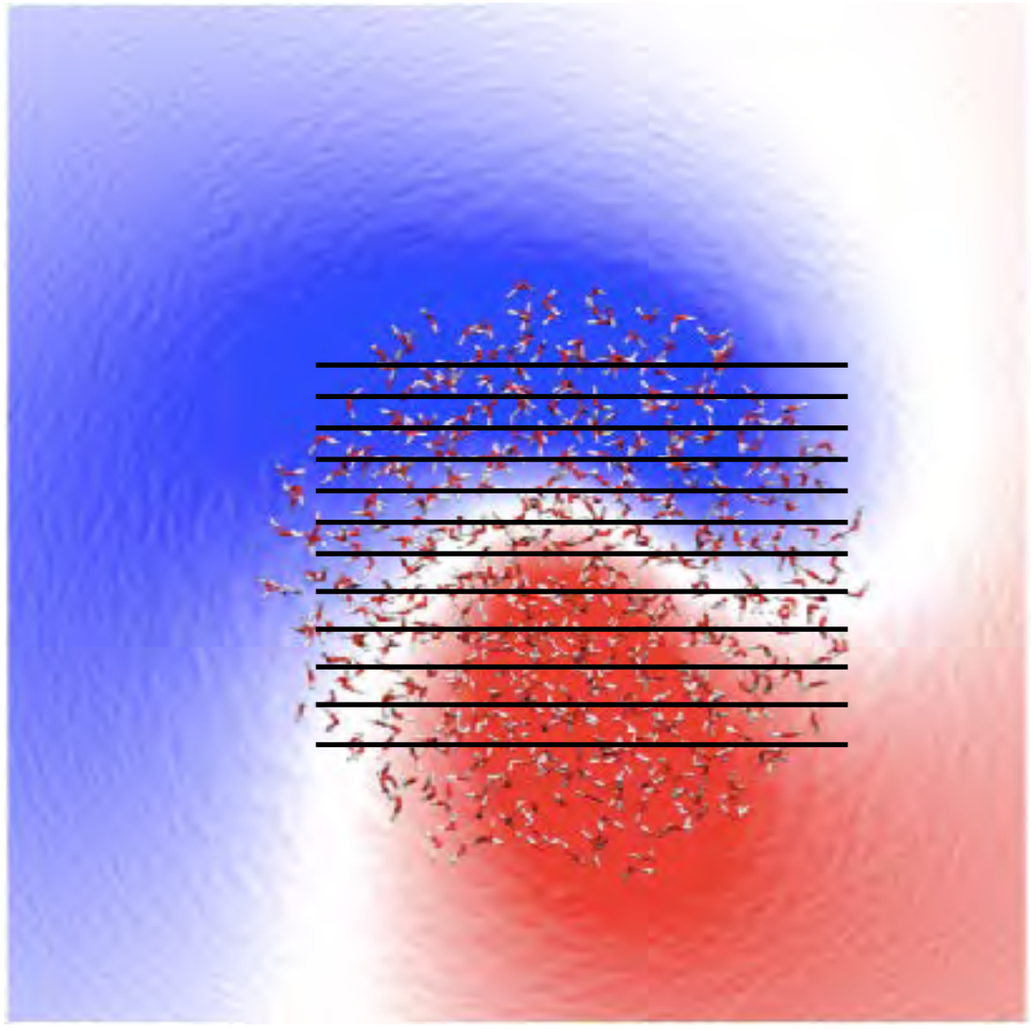


G. Lever, D. Cole, N. Hine, P. Haynes, M. Payne, *J. Phys. Cond. Matt.*, **2013**, *25*, 152101

HOMO-LUMO Gap

Dipole field leads to shifts in the electronic energy levels and closing of the HOMO-LUMO gap.

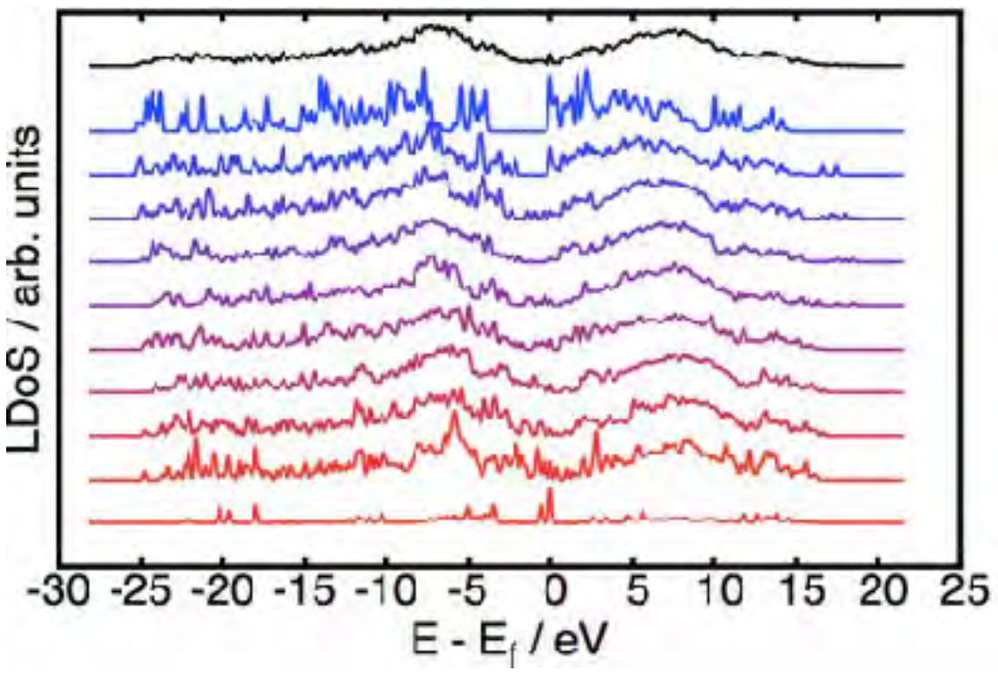




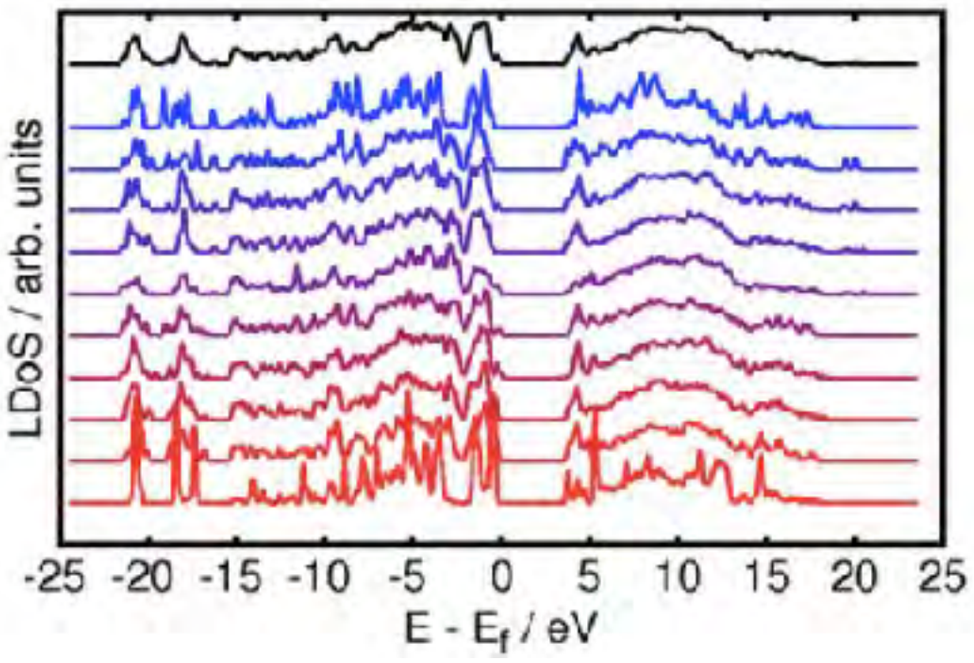
G. Lever, D. Cole, N. Hine, P. Haynes, M. Payne, *J. Phys. Cond. Matt.*, **2013**, *25*, 152101

HOMO-LUMO Gap

Before



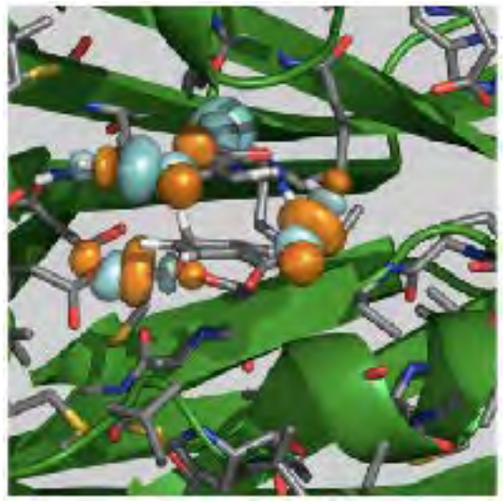
After



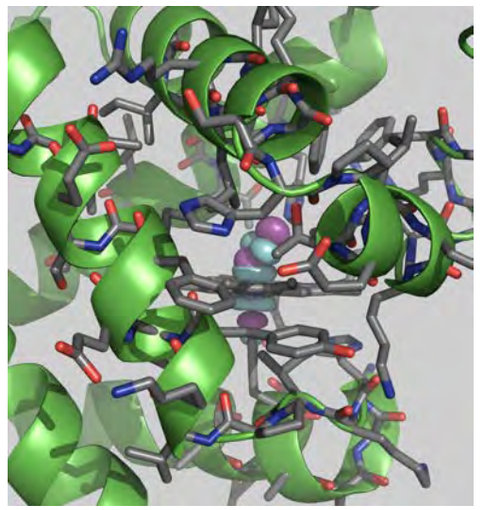
For large-scale cluster calculations, Kohn-Sham DFT works well, but:
1) check for presence of HOMO-LUMO gap;
2) use implicit solvent.

G. Lever, D. Cole, N. Hine, P. Haynes, M. Payne, *J. Phys. Cond. Matt.*, **2013**, *25*, 152101

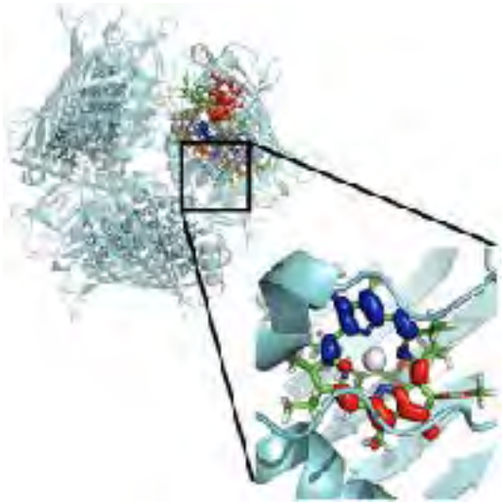
Biological Applications



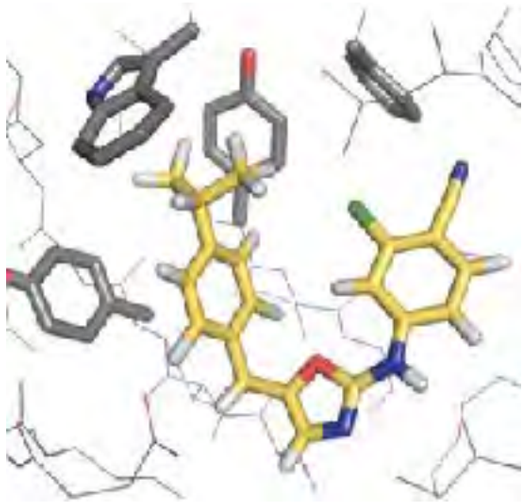
Transition state searching in enzymes



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)

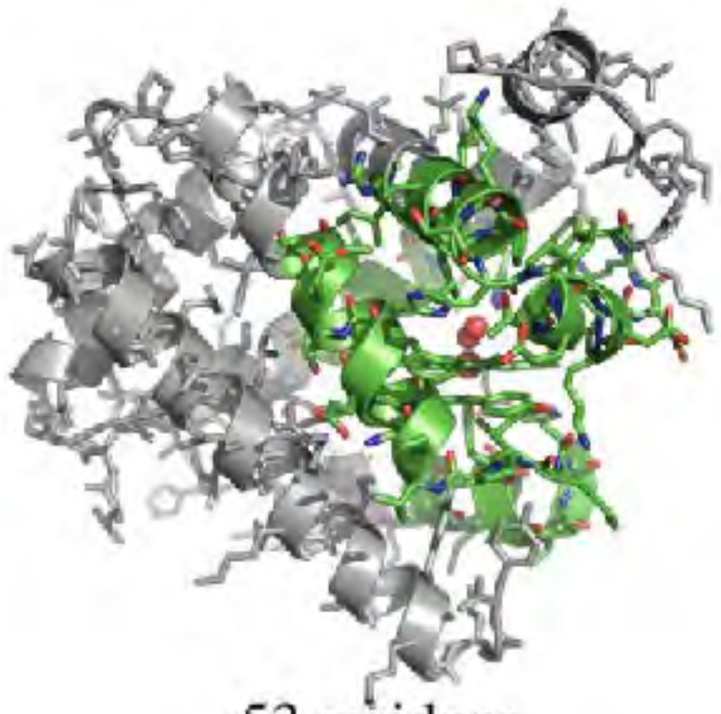
Metalloproteins

$K_{CO}:K_{O_2}$ ~ 20,000
ΔΔE = 5.9 kcal/mol

1 residue

3 residues

$K_{CO}:K_{O_2}$ ~ 20
ΔΔE = 1.9 kcal/mol



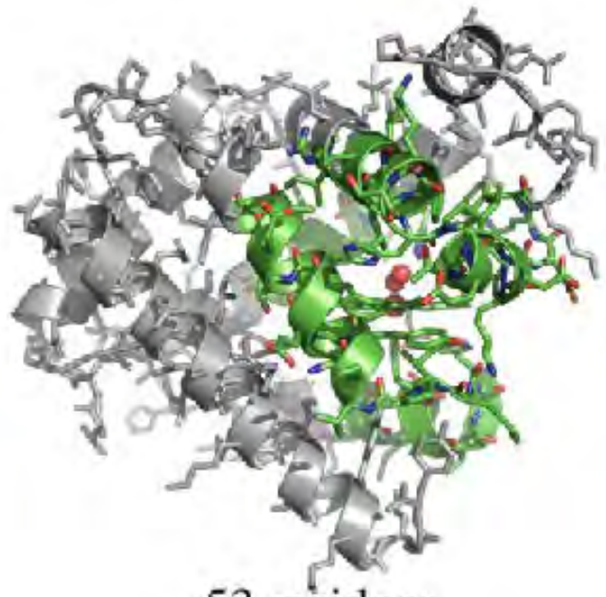
53 residues

Myoglobin is small protein responsible for storing O_2 in muscle tissue.
Function hindered by CO.

Olson, *J. Biol. Inorg. Chem.* **2**, 544 (1997)

Energetics

1 residue



53 residues

ΔΔE / kcal/mol

	Vacuum	Protein
U=0eV	18	15
U=3eV	12	9
Experiment	6	2

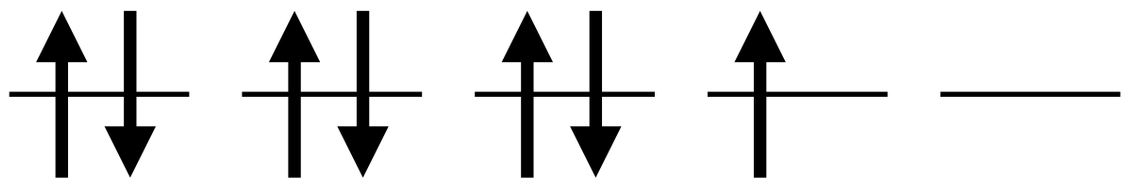
D. Cole, D. O'Regan, M. Payne, *J. Phys. Chem. Lett.* **3**, 1448 (2012)

Dynamical Mean Field Theory

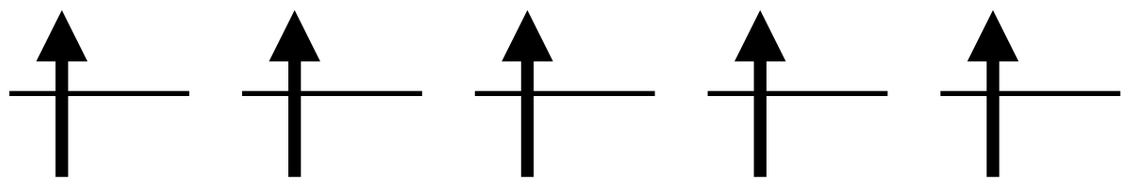
In systems with strongly-correlated electrons, DMFT treats the U correction at the many-body level, including multi-determinant and finite-temperature effects.

In addition, Hund's exchange coupling J is included, which tends to align the spins in the correlated subsystem.

Occupation of Fe 3d subshell in heme:



Low J (doublet)



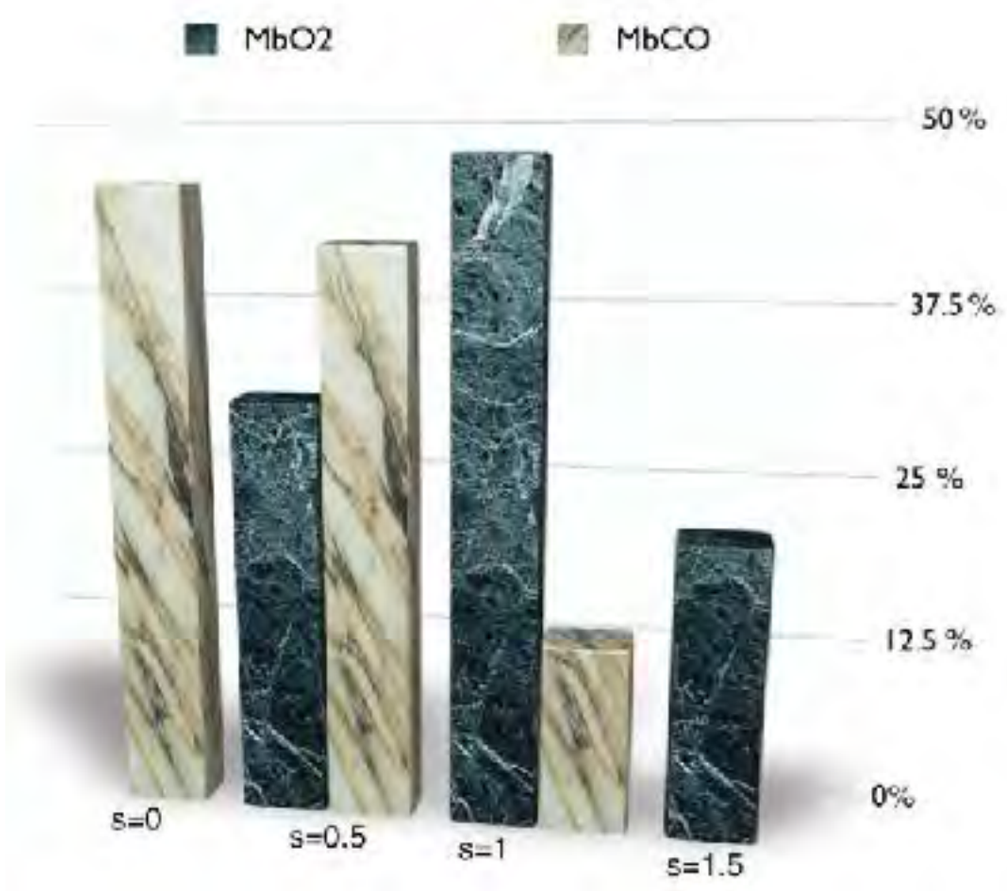
High J (sextet)

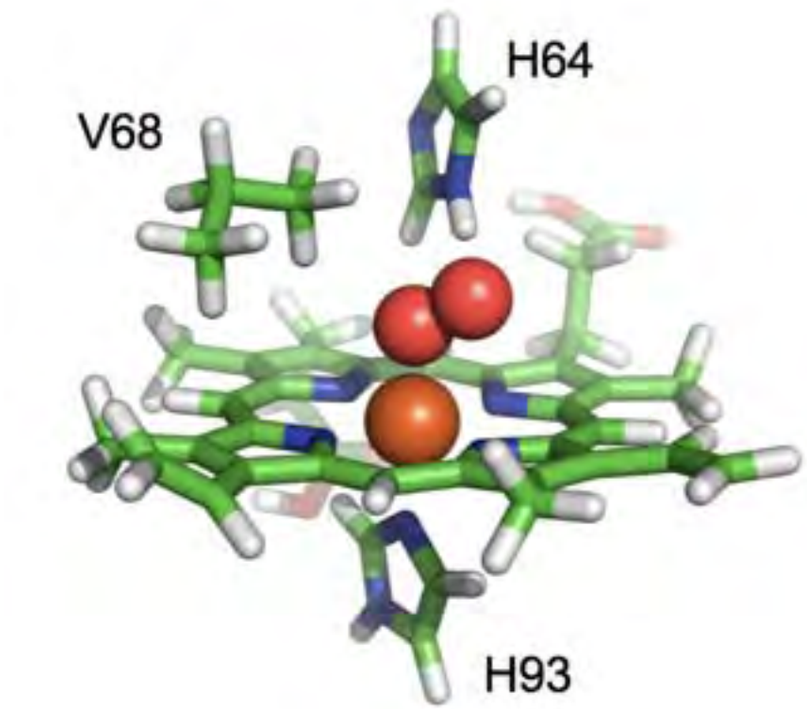
C. Weber et al., *Phys. Rev. Lett.* **110**, 106402 (2013)

Dynamical Mean Field Theory

Apply DFT+DMFT to realistic model of myoglobin (1,000 atoms).

The ground state wave function is not a pure state (entangled singlet, triplet … states). Larger valence fluctuations observed in MbO_2 than MbCO.





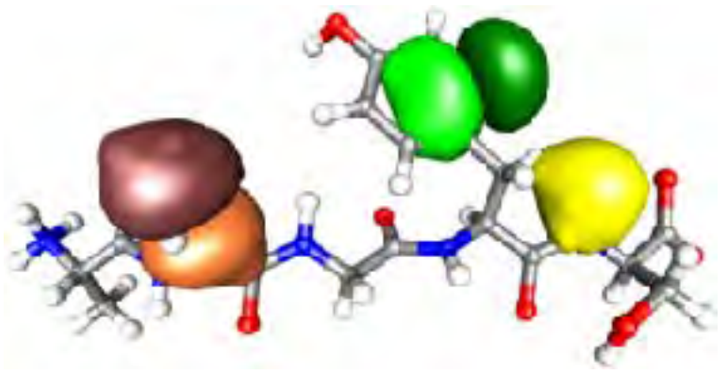
DFT+U: $Q(O_2) \sim -0.5\ e$

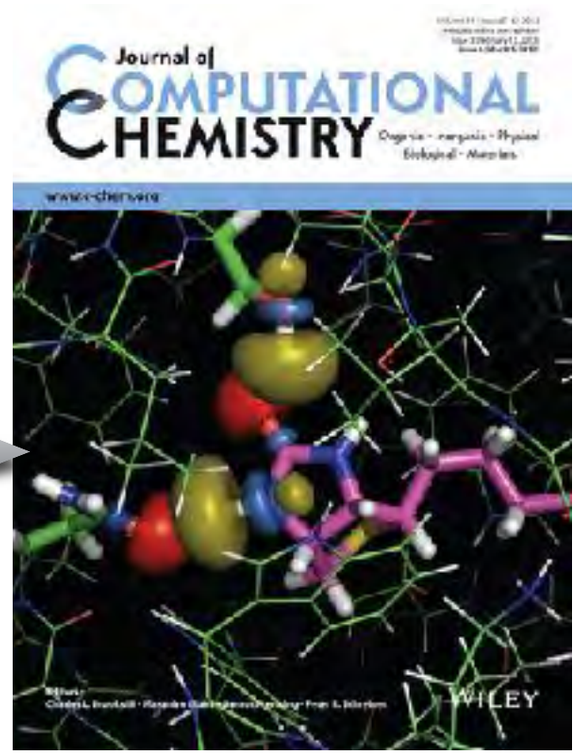
DFT+DMFT: $Q(O_2) \sim -1.0\ e$

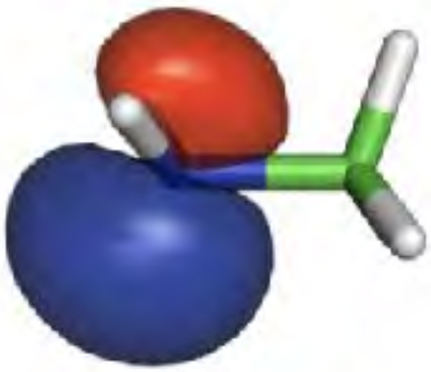
C. Weber, D. Cole, D. O'Regan, M. Payne, *Proc. Natl. Acad. Sci. USA* **111**, 5790 (2014)

Natural Bond Orbitals

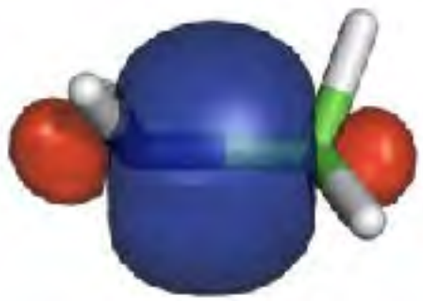
NBO translates accurate calculations into chemical insights
http://www.chem.wisc.edu/~nbo5



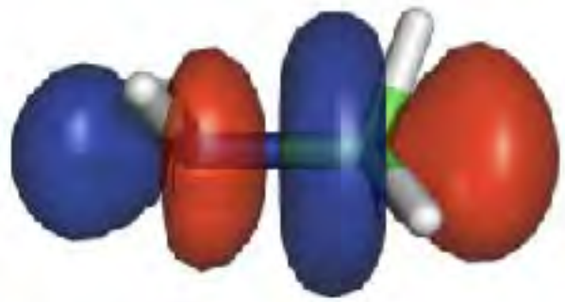




N lone pair



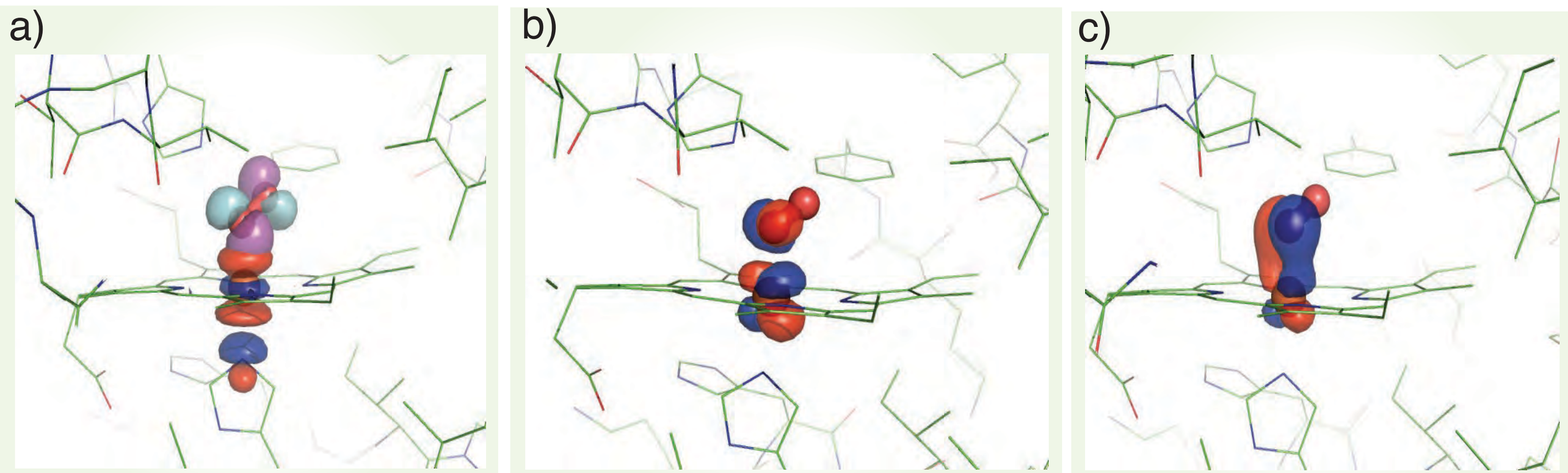
C-N σ bond



C-N σ* anti-bond

L. Lee, D. Cole, M. Payne, C.-K. Skylaris, *J. Comp. Chem.* **34**, 429 (2013)

Dynamical Mean Field Theory



ligand-to-metal back charge transfer

metal-to-ligand charge transfer

Enhanced π-bonding, charge transfer increases bonding to O_2.

dπ hole character is 19% using DFT+DMFT, compared with 15% from Fe L-edge X-ray absorption spectroscopy (50% for CASSCF/MM).

C. Weber, D. Cole, D. O'Regan, M. Payne, *Proc. Natl. Acad. Sci. USA* **111**, 5790 (2014)

Energetics

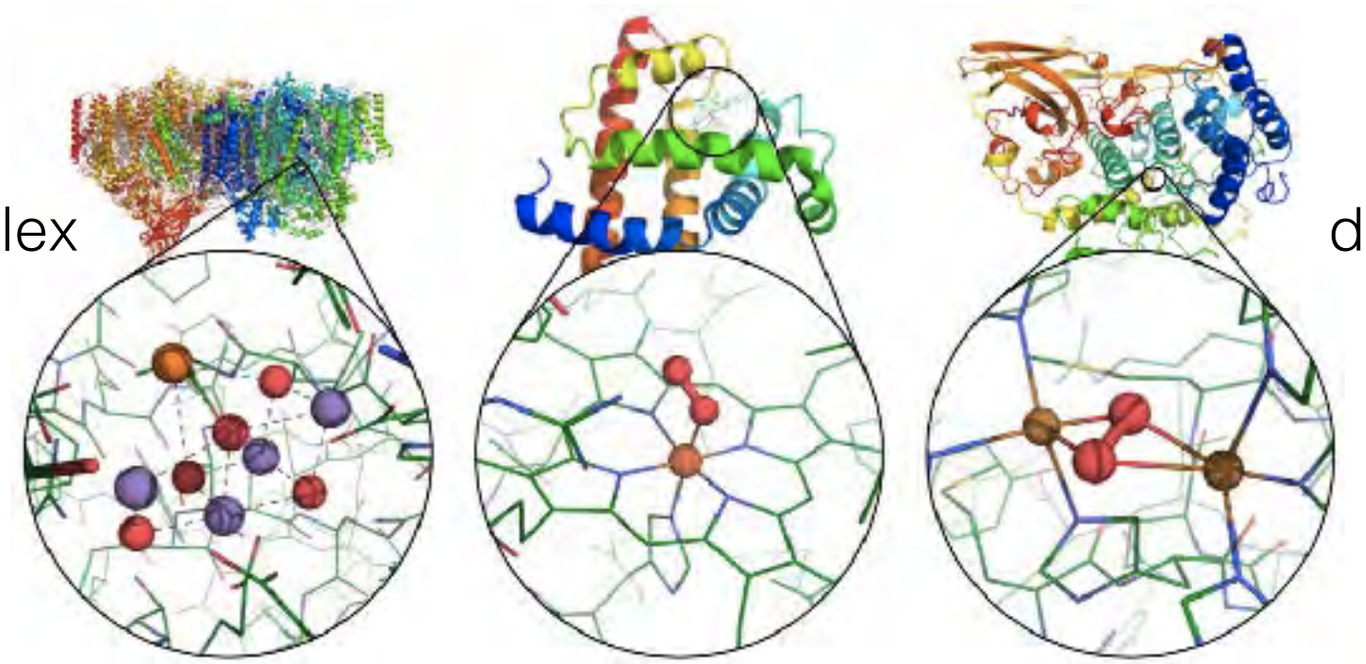
	ΔΔE / kcal/mol	
	Vacuum	Protein
DFT	18	15
DFT+U (U=3eV)	12	9
DFT+DMFT (J=0.7eV)	—	2
Experiment	6	2

Require both multi-determinant quantum effects and large system sizes to recover experimental relative binding energies

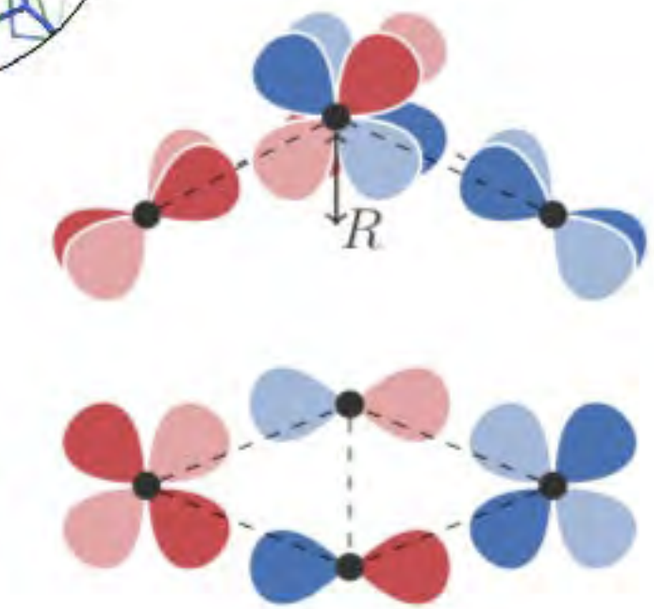
C. Weber, D. Cole, D. O'Regan, M. Payne, *Proc. Natl. Acad. Sci. USA* **111**, 5790 (2014)

Ongoing Work

Water oxidising complex of photosystem II



di-Cu oxo bridge in hemocyanin

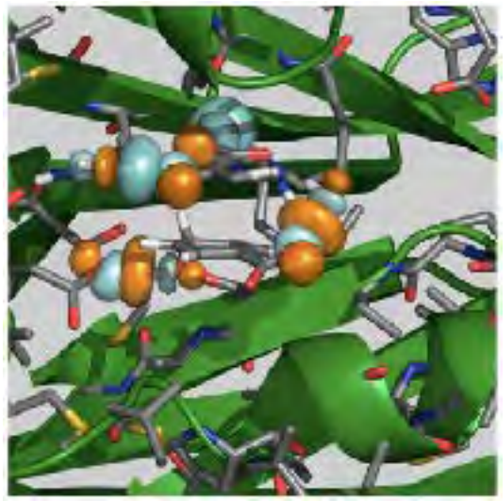


In hemocyanin, experimental observation of singlet state is at odds with DFT calculations on Cu-O_2-Cu core.

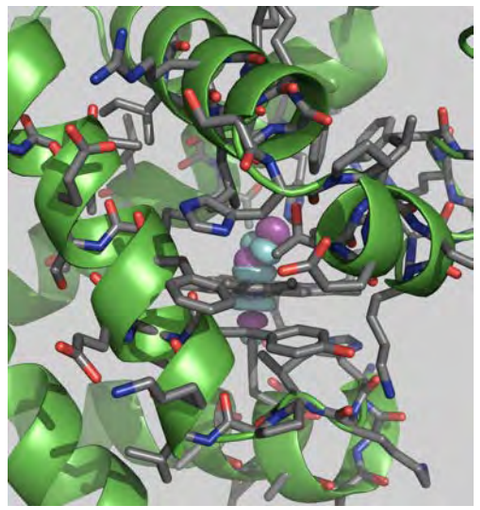
Aim to move away from ‘toy models’ towards predictive modelling of quantum effects in biology.

al-Badri MA, Linscott E, Georges A, Cole DJ, Weber C, https://arxiv.org/abs/1811.05739

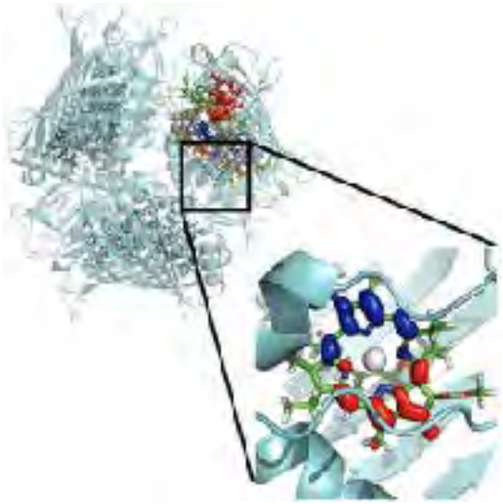
Biological Applications



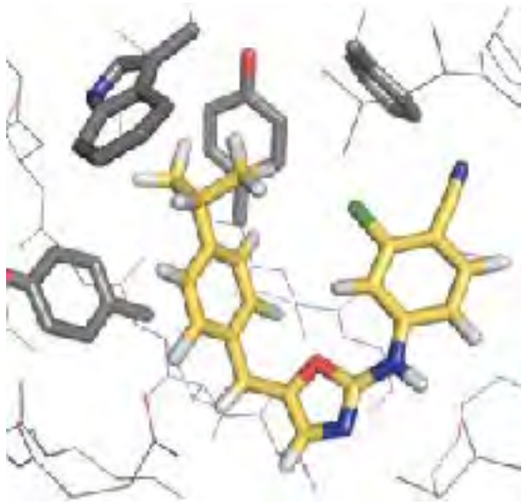
Transition state searching in enzymes



Protein-ligand binding in metalloproteins



Optical spectroscopy in a light-harvesting protein



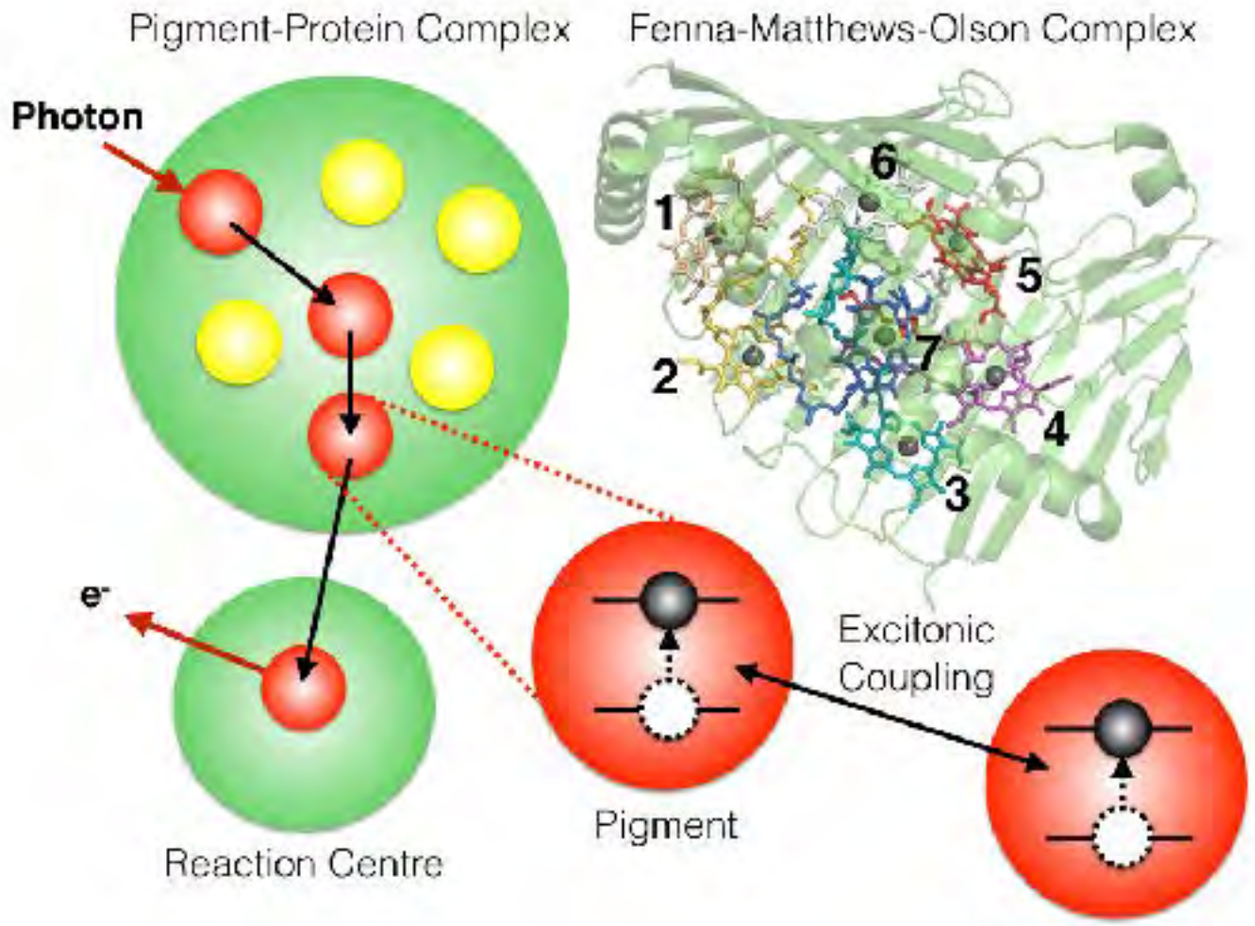
Classical force field parameterisation for drug discovery

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

FMO Complex

Fenna-Matthews-Olson (FMO) complex: funnels electronic excitations (excitons) to reaction centre where they are used to release electrons for photosynthesis.

X-ray diffraction reveals a trimeric structure. Each monomer contains 7 bacterio-chlorophyll pigments.



Exciton transfer through FMO modelled by pigment-protein complex (PPC) Hamiltonian:

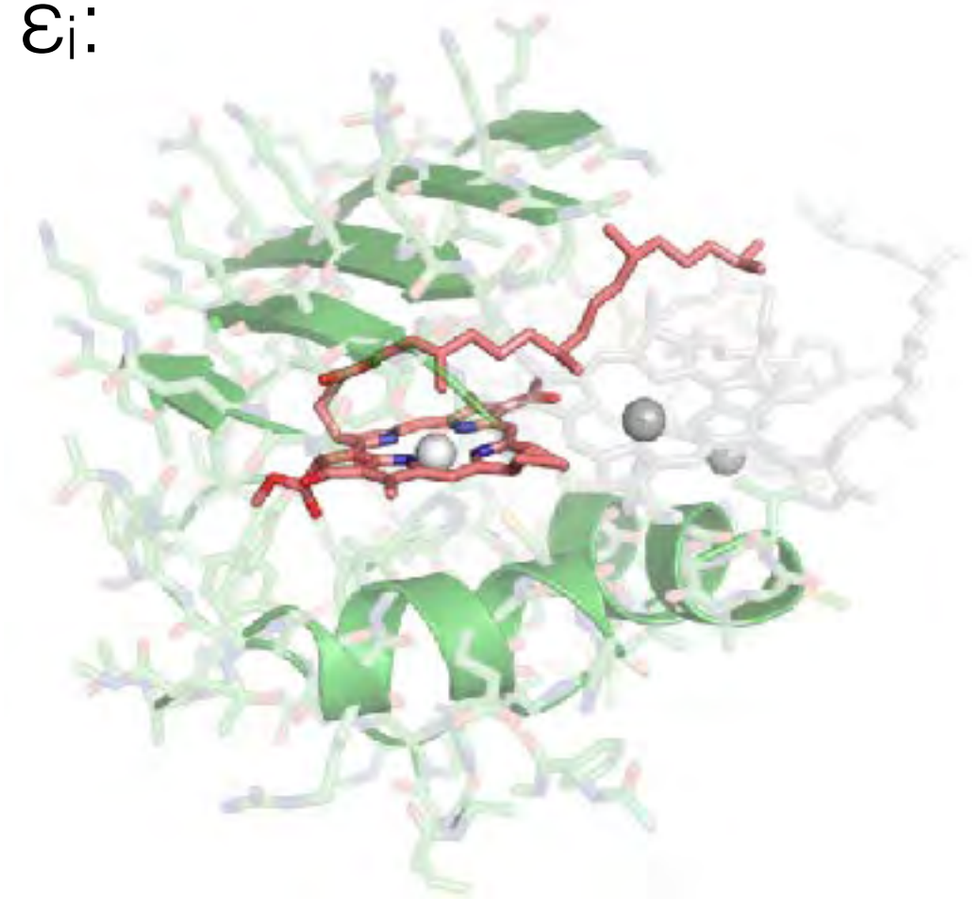
$$H = \sum_i \varepsilon_i |i\rangle\langle i| + \sum_{i \neq j} J_{ij} |i\rangle\langle j|$$

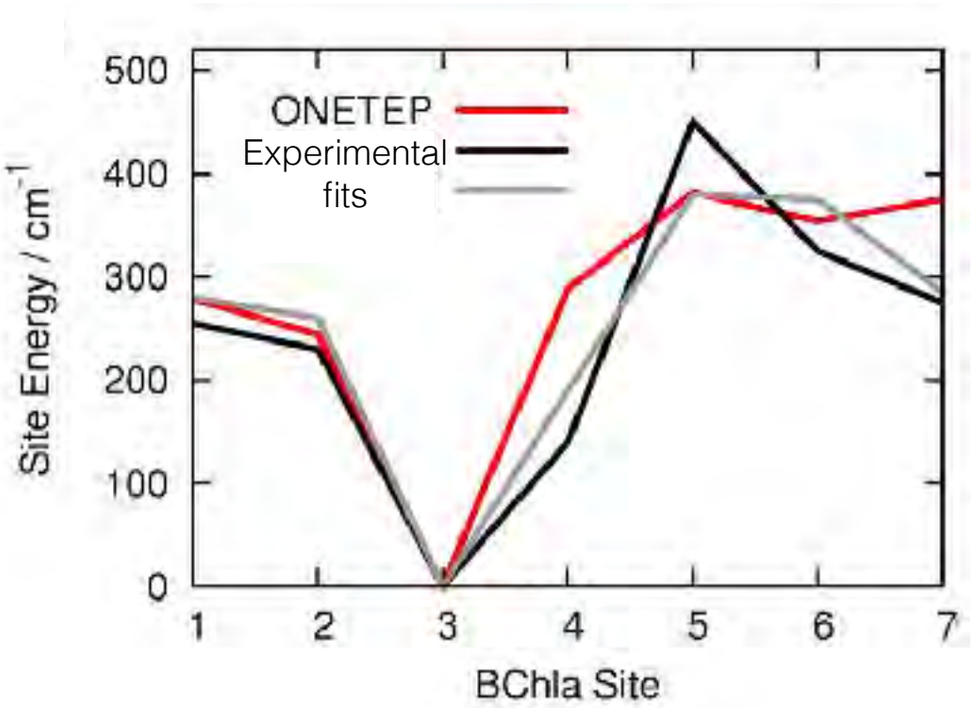
site energies — coupling between optical transitions

A. W. Chin et al., *Nature Phys.*, **2013**, *9*, 113-118

DFT results

ε_i:





Pigment site energies modulated by:

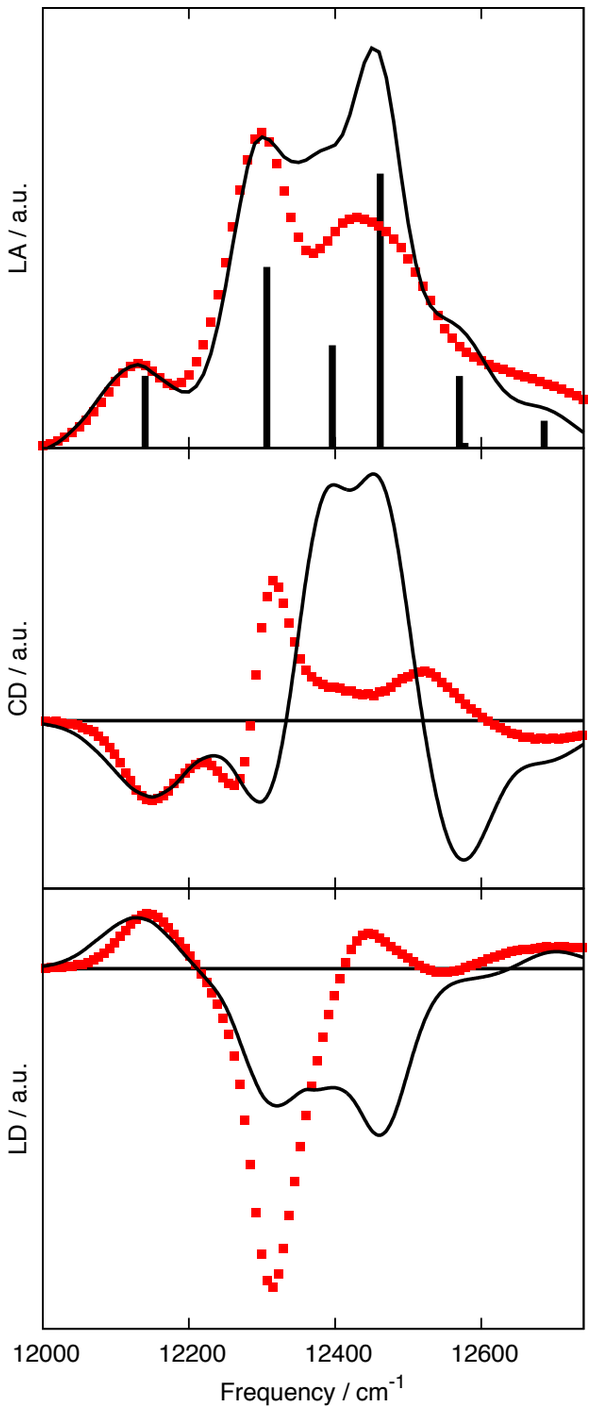
conformation of pigment

hydrogen bonds with environment

interactions with protein secondary structure

D. Cole, A. Chin, N. Hine, P. Haynes, M. Payne, J. Phys. Chem. Lett., ***4****, 4206 (2013)*

Optical Spectra



$$H = \sum_i \varepsilon_i |i\rangle\langle i| + \sum_{i \neq j} J_{ij} |i\rangle\langle j|$$

Linear optical absorption spectrum calculated from excitonic eigenstates of PPC Hamiltonian.

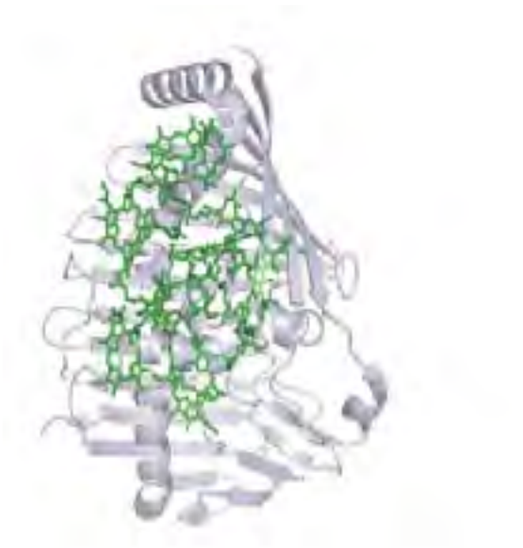
Also require spectral density function describing finite-temperature dissipative interactions with protein and electronic disorder.

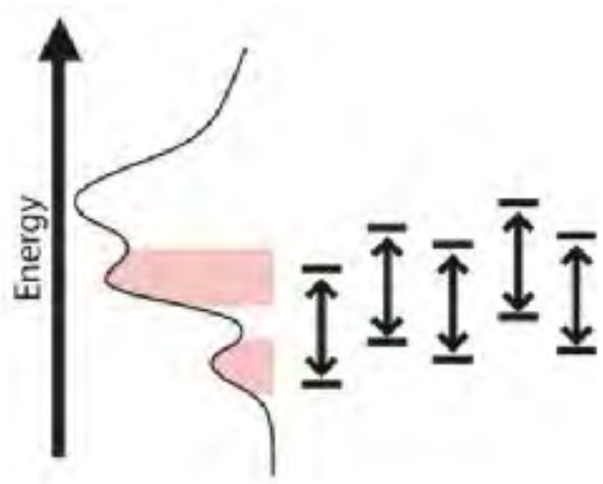
Red = experiment (courtesy of Dugan Hayes, Rienk van Grondelle and Markus Wendling)

D. Cole, A. Chin, N. Hine, P. Haynes, M. Payne, J. Phys. Chem. Lett., ***4****, 4206 (2013)*

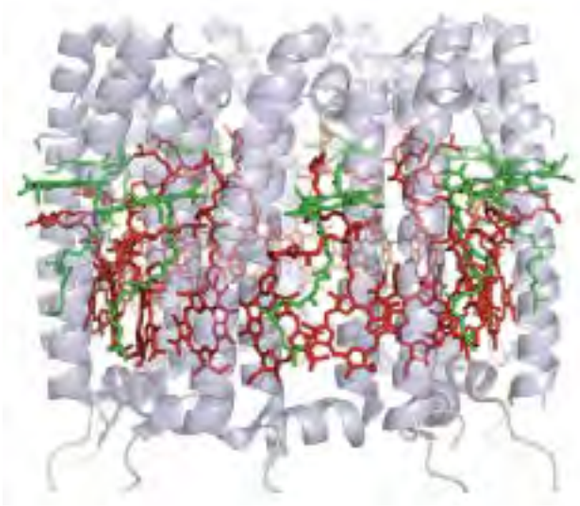
Static Disorder

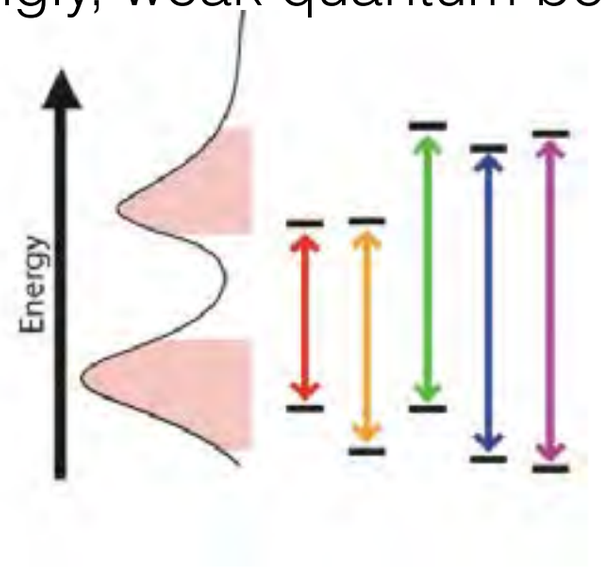
FMO: energy gaps remain constant, strong quantum beats





LH2: energy gaps vary strongly, weak quantum beats





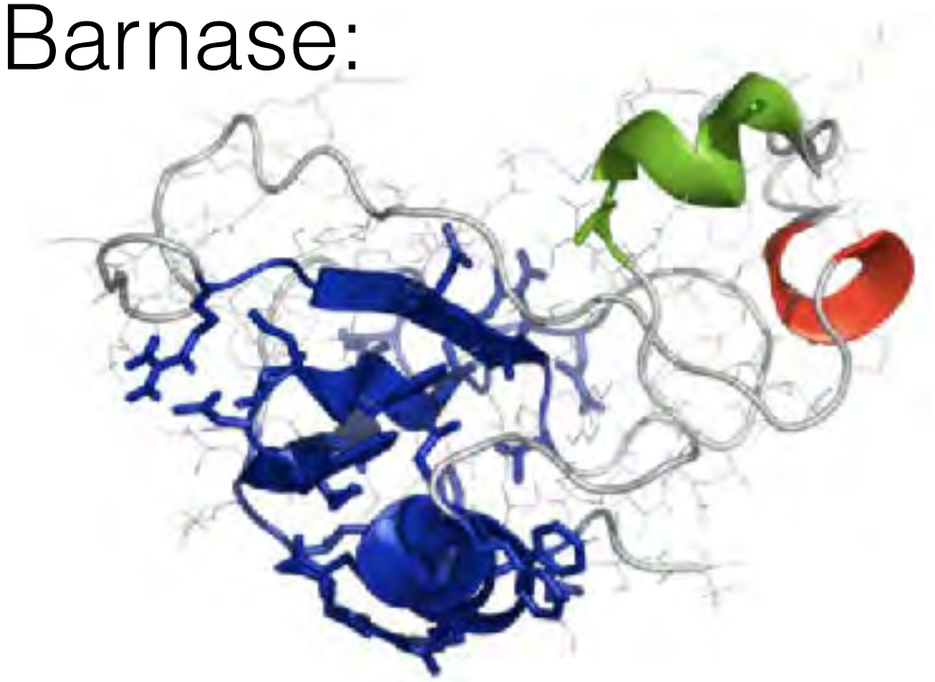
Exciton structure can be masked in experiments by static disorder across structural ensembles.

Structural motion associated with static disorder is on long time scale — difficult to access using traditional atomistic simulations.

Fidler et al., J. Phys. Chem. A, **116***, 282 (2012)*

Constrained Geometric Simulation

Barnase:



FIRST software identifies constraints due to bonds, angles, hydrogen bonds and hydrophobic interactions.

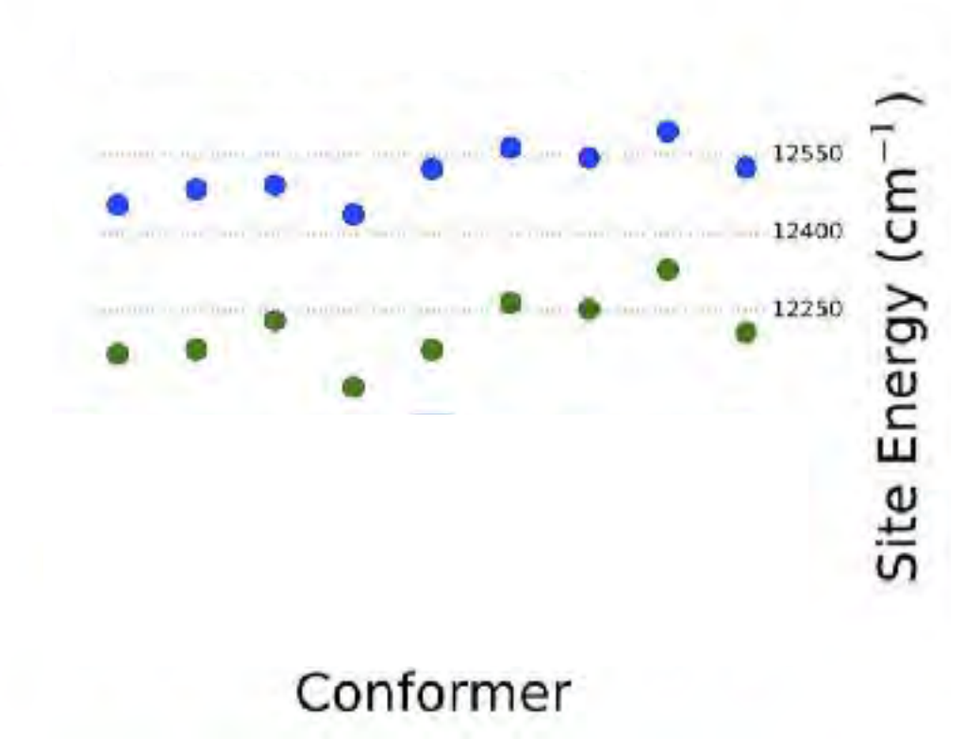
Determines flexible and rigid regions in the protein. Substantially reduces number of degrees of freedom to be explored.

FMO:

Ensemble of structures satisfying constraints is generated using FRODA — our best estimate of the large-amplitude fluctuations present in the experimental ensemble.

S. Wells, S. Menor, B. Hespenheide, M. Thorpe, Phys. Biol., ***2****, S127 (2005)*

Pigment-Protein Dynamics



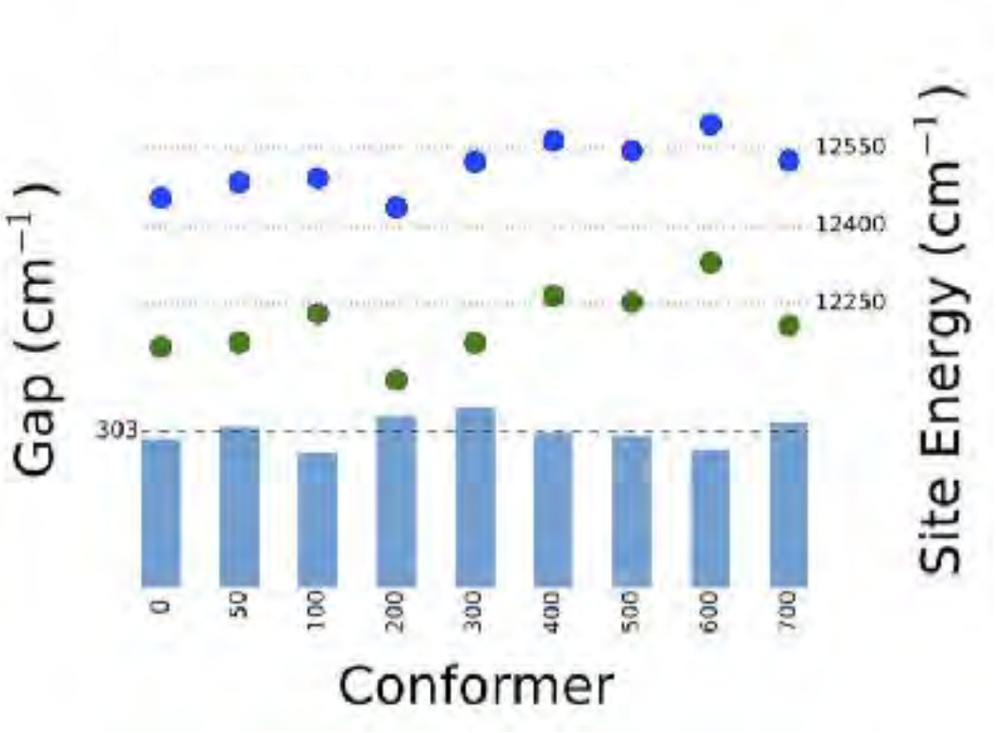
Snapshots post-processed using ONETEP to investigate effects of thermal disorder on excitonic energy landscape.

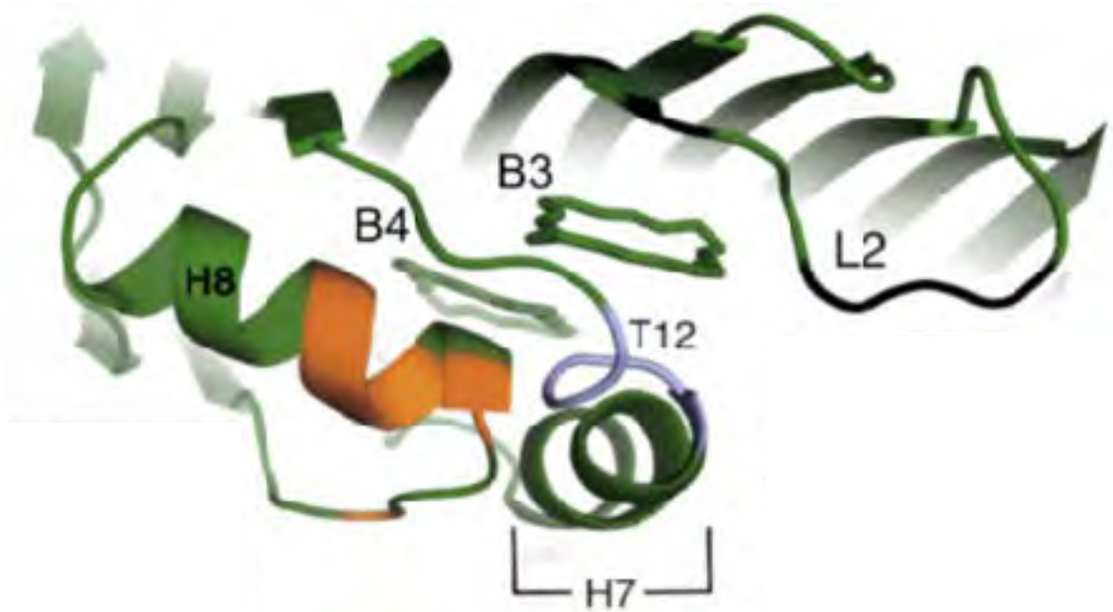
Substantial variation in site energies of pigments 3 (green) and 4 (blue).

If the site energies were uncorrelated we would expect gap variations of ~85/cm and rapid loss of exciton coherence.

A. Fokas, D. Cole, N. Hine, S. Wells, M. Payne, A. Chin, J. Phys. Chem. Lett., **8***, 2350 (2017)*

Pigment-Protein Dynamics





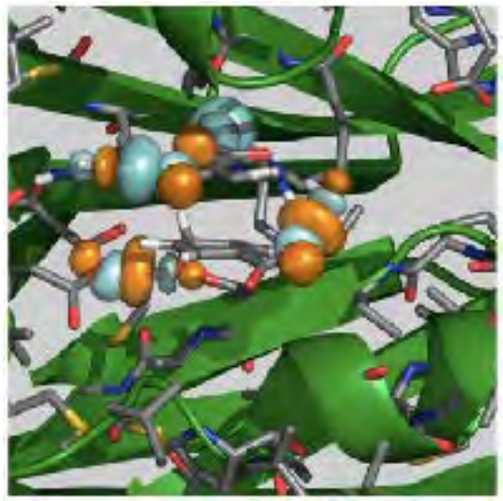
However, energy gap is extremely homogeneous (303 +/- 27/cm).

FMO appears to stabilise its excitonic structure by limiting thermal fluctuations to motions that do not affect energy gaps.

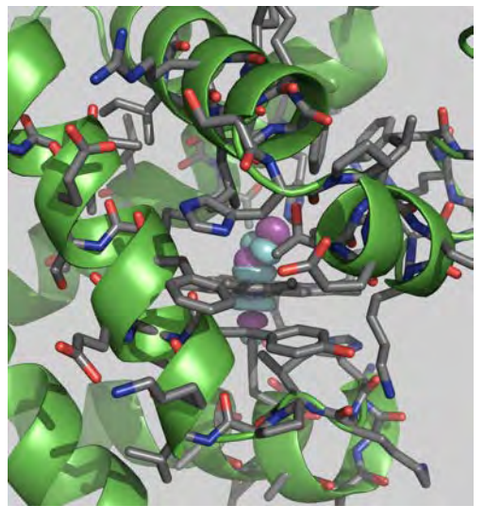
Exploitation in artificial light-harvesting devices?

A. Fokas, D. Cole, N. Hine, S. Wells, M. Payne, A. Chin, J. Phys. Chem. Lett., **8**, *2350 (2017)*

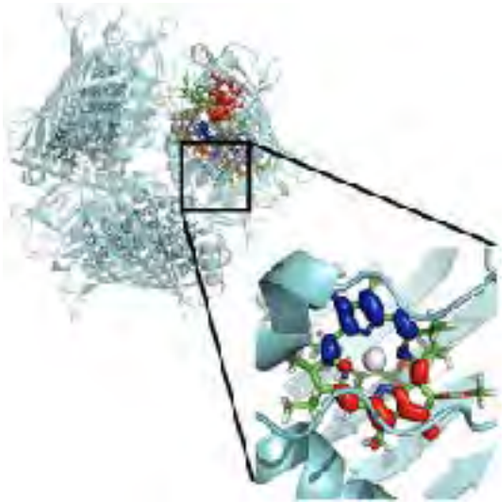
Biological Applications



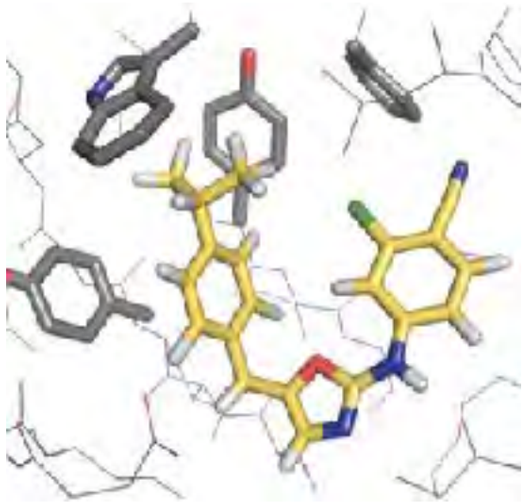
Transition state searching in enzymes



Protein-ligand binding in metalloproteins



Optical spectroscopy in a light-harvesting protein



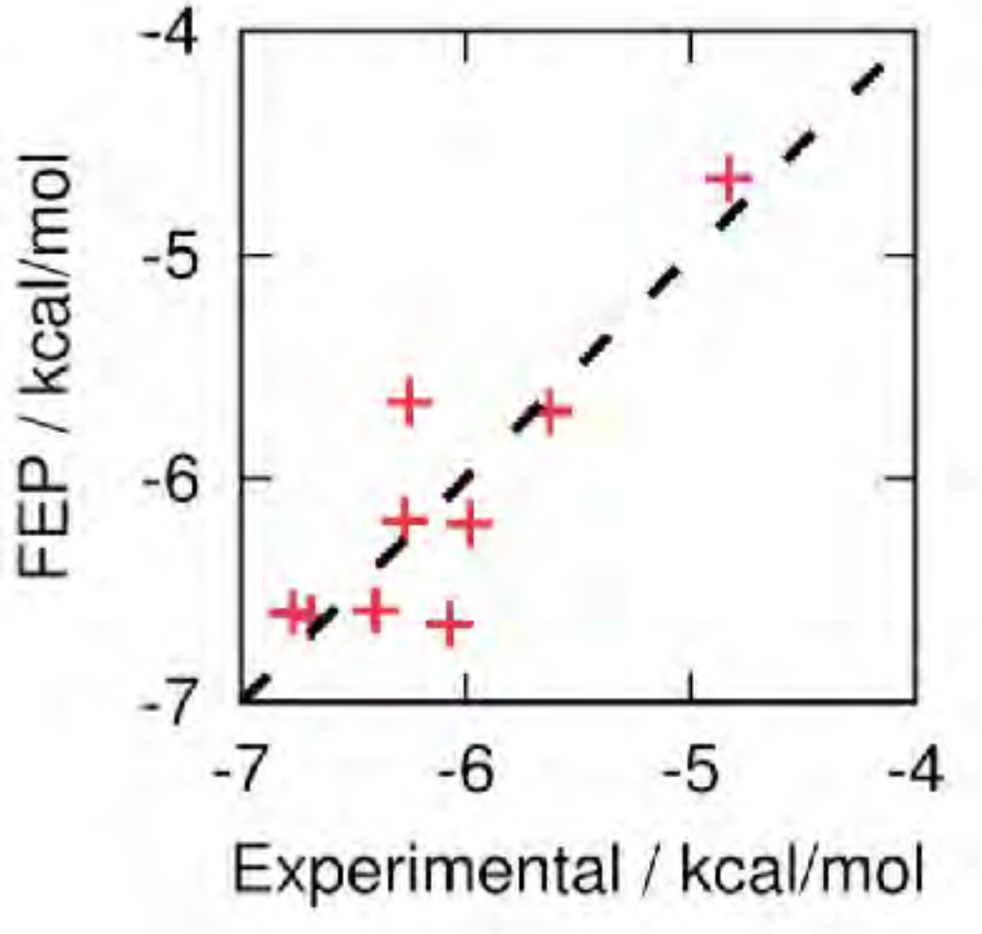
Classical force field parameterisation for drug discovery

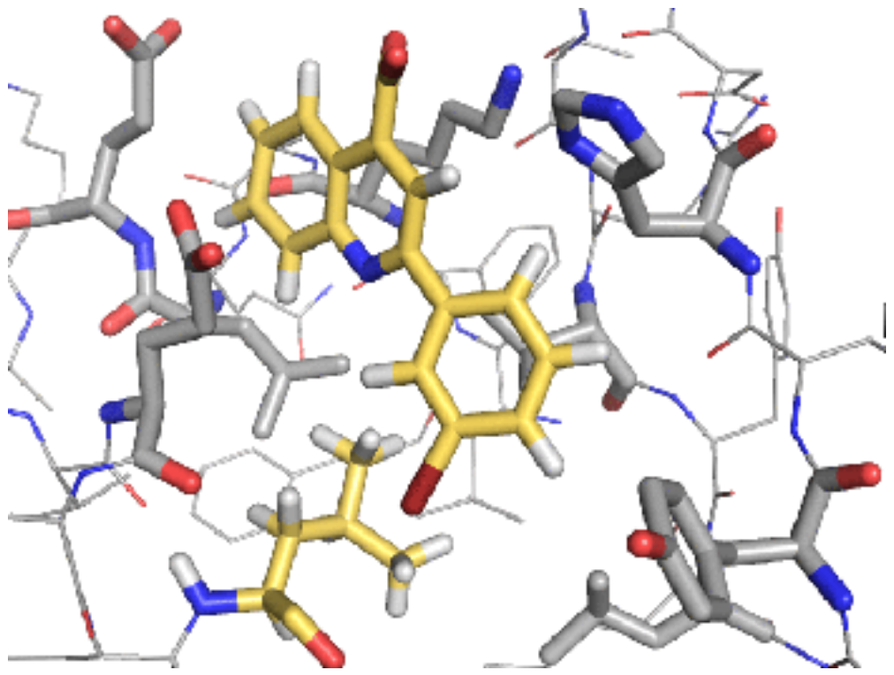
"Applications of Large-Scale DFT in Biology", D. J. Cole & N. D. M. Hine
Journal of Physics: Condensed Matter, 28, 393001 (2016)

Introduction

Molecular mechanics (MM) force fields are widely used in computer-aided drug design, protein folding, protonation states, protein-surface interactions, allosteric mechanisms, QM/MM….

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





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

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

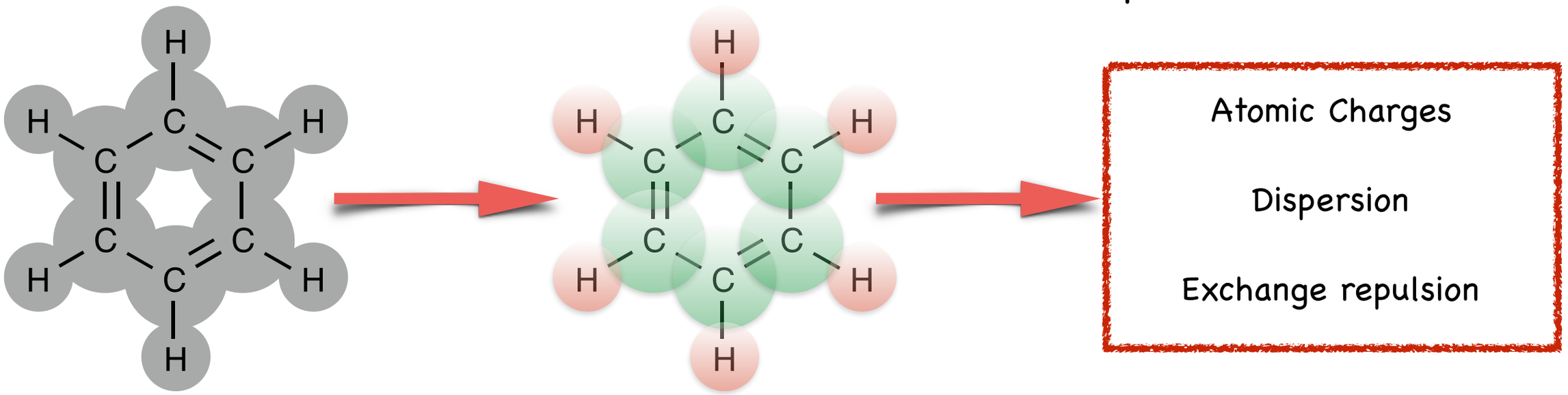
$$E^{MM} = E_{\text{bonded}} + E_{\text{non-bonded}} = E_{\text{bonded}} + \sum_{i<j} \frac{q_i q_j}{r_{ij}} + \left(\frac{A^{ij}}{r_{ij}^{12}} - \frac{C_6^{ij}}{r_{ij}^{6}} \right)$$

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



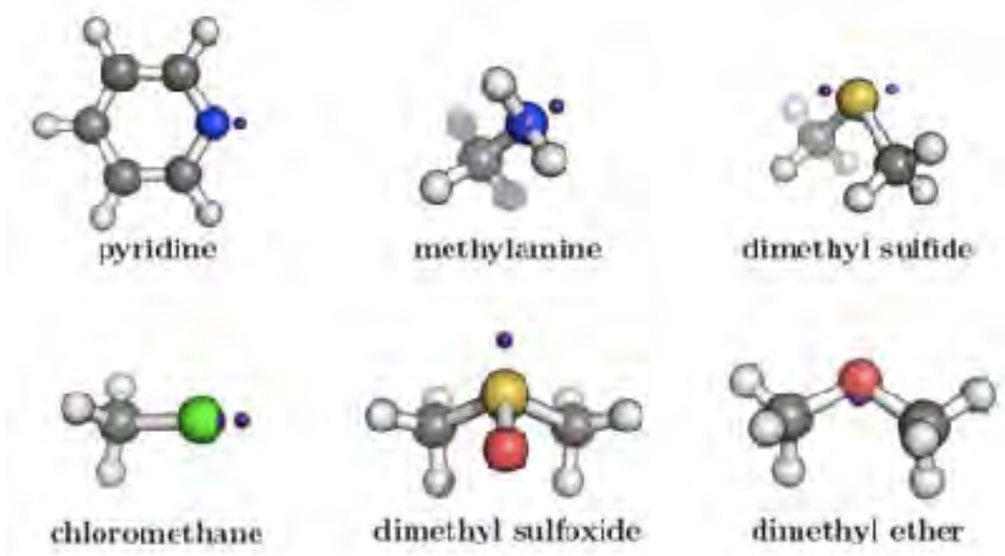
Density derived electrostatic and chemical (DDEC) electron density partitioning (good reproduction of the ESP and not too conformation dependent).

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

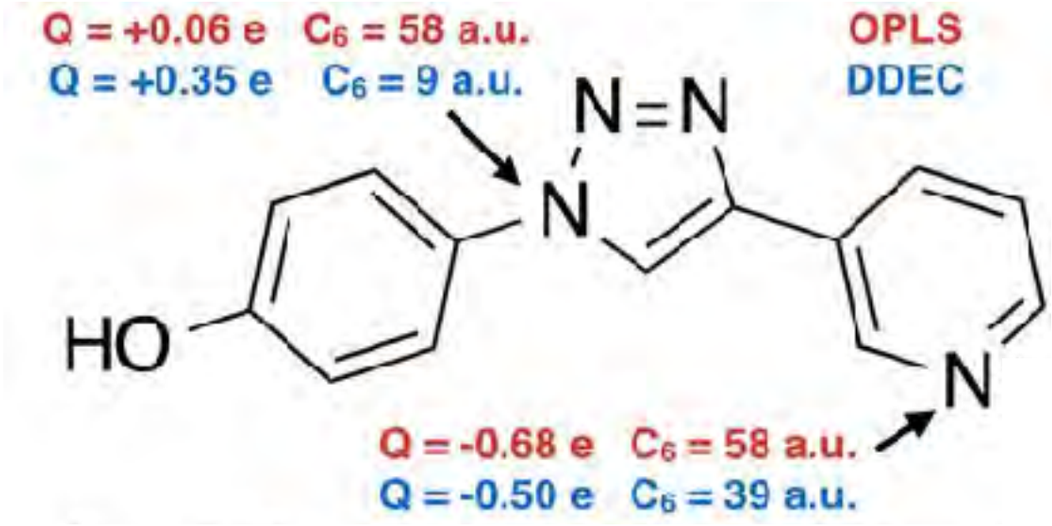
D. Cole, J. Vilseck, J. Tirado-Rives, M. Payne, W. Jorgensen, *JCTC,* **2016**, *12,* 2312

Force Field Parameters

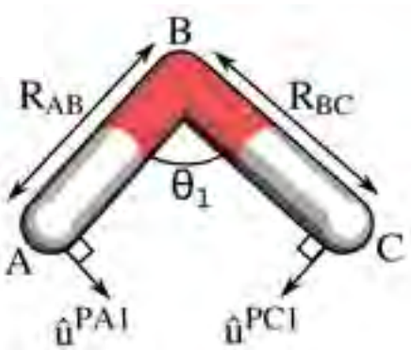
Off site charges (available in ONETEP)



Lennard-Jones (from Tkatchenko-Scheffler approach)

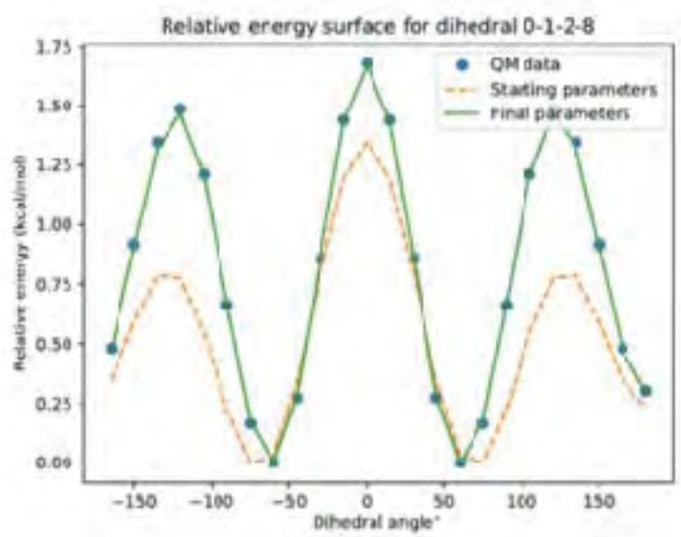


Bonded parameters (modified Seminario)



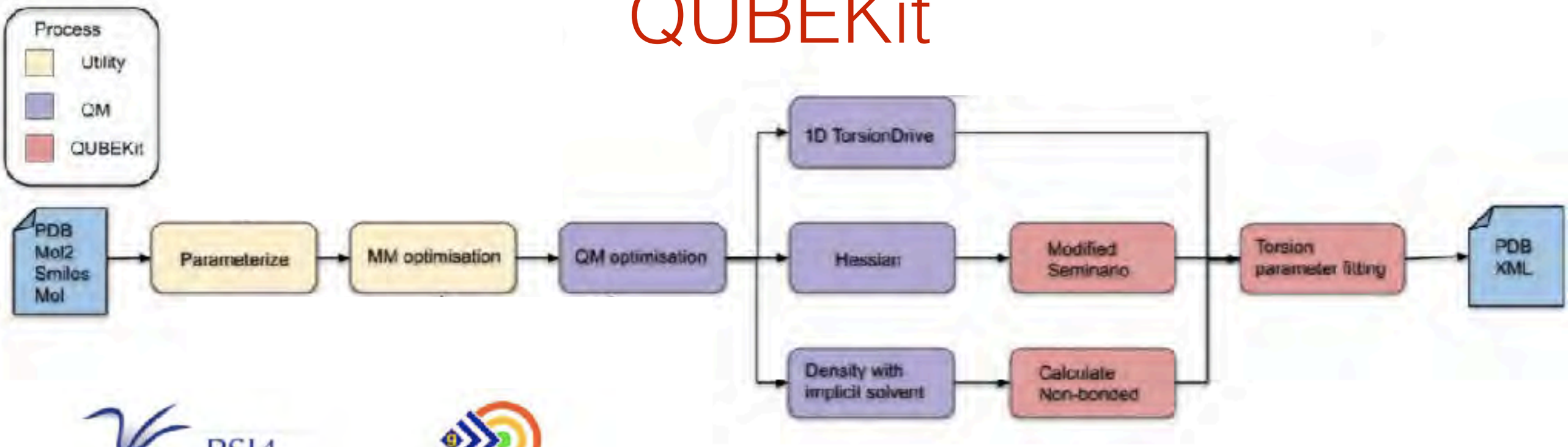
$$[\mathbf{k_{AB}}] = -\begin{vmatrix} \frac{\partial^2 E}{\partial x_A \partial x_B} & \frac{\partial^2 E}{\partial x_A \partial y_B} & \frac{\partial^2 E}{\partial x_A \partial z_B} \\ \frac{\partial^2 E}{\partial y_A \partial x_B} & \frac{\partial^2 E}{\partial y_A \partial y_B} & \frac{\partial^2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

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

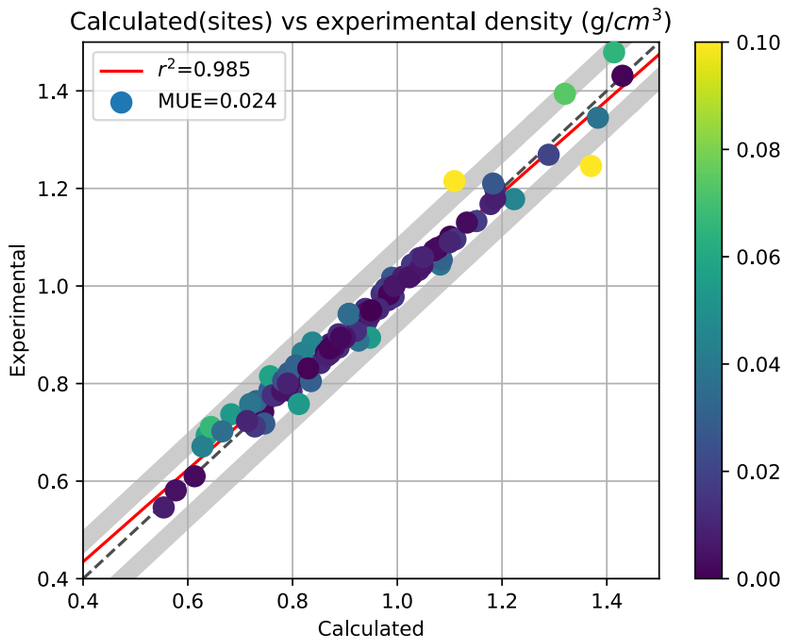
geomeTRIC

QUBE Kit

Chargemol

QC Archive

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

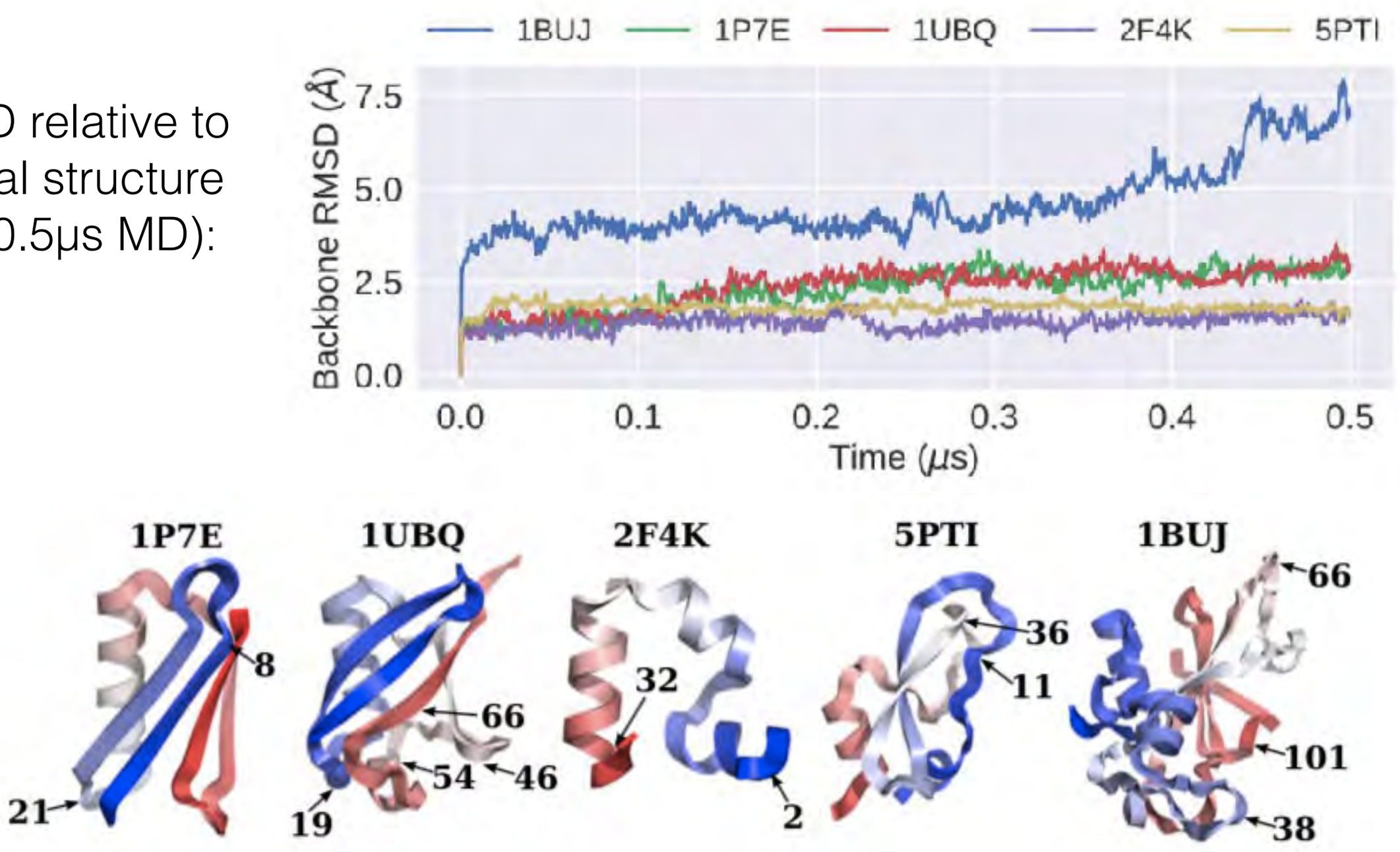


Josh Horton, Chris Ringrose

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

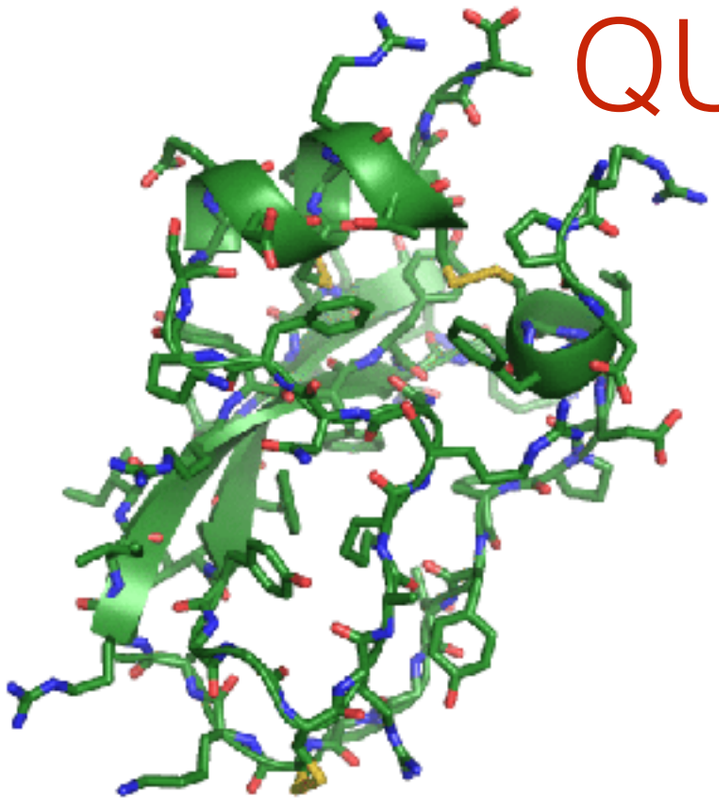
QUBE Protein Force Field

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



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

A. Allen, M. Robertson, M. Payne, D. Cole, *ACS Omega (2019), in press*



QUBE Protein Force Field

MD simulations can be benchmarked against experimental NMR J-couplings which describe conformational fluctuations.

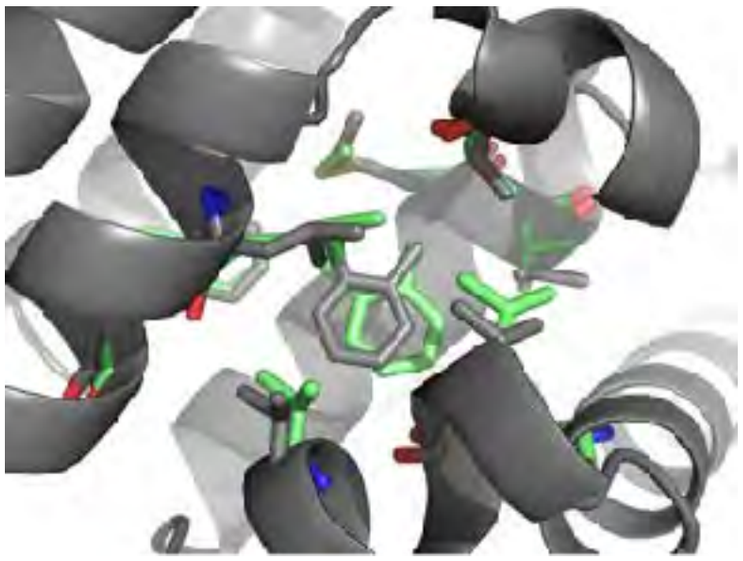
	OPLS-AA	OPLS-AA/M	QUBE
Dipeptides	1.0	0.4	0.4
Ala_5	2.3	1.2	0.9
1UBQ	1.8	1.1	1.5
1P7E	1.5	0.9	1.2

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

A. Allen, M. Robertson, M. Payne, D. Cole, *ACS Omega (2019), in press*

Free Energy Perturbation

Absolute free energies of binding of small molecules to the L99A mutant of T4 lysozyme (kcal/mol):



	OPLS	**QUBE**	**Experiment**
benzene	-7.7	-6.0	-5.2
p-xylene	-5.0	-4.4	-4.7
o-xylene	-2.9	-5.0	-4.6
benzofuran	-7.2	-7.0	-5.5
indole	-4.4	-3.8	-4.9
indene	-5.9	-4.0	-5.1
MUE	**1.3**	**0.9**	**—**

Cole DJ, Cabeza de Vaca I, Jorgensen WL. *MedChemComm* 2019, **10**, 1116-1120.

Acknowledgements

Newcastle University
Joshua Horton
Lauren Nelson
Chris Ringrose
Vadiraj Kurdekar

Yale University
Jonah Vilseck
Leela Dodda
Michael Robertson
Israel Cabeza de Vaca
Julian Tirado-Rives
William Jorgensen

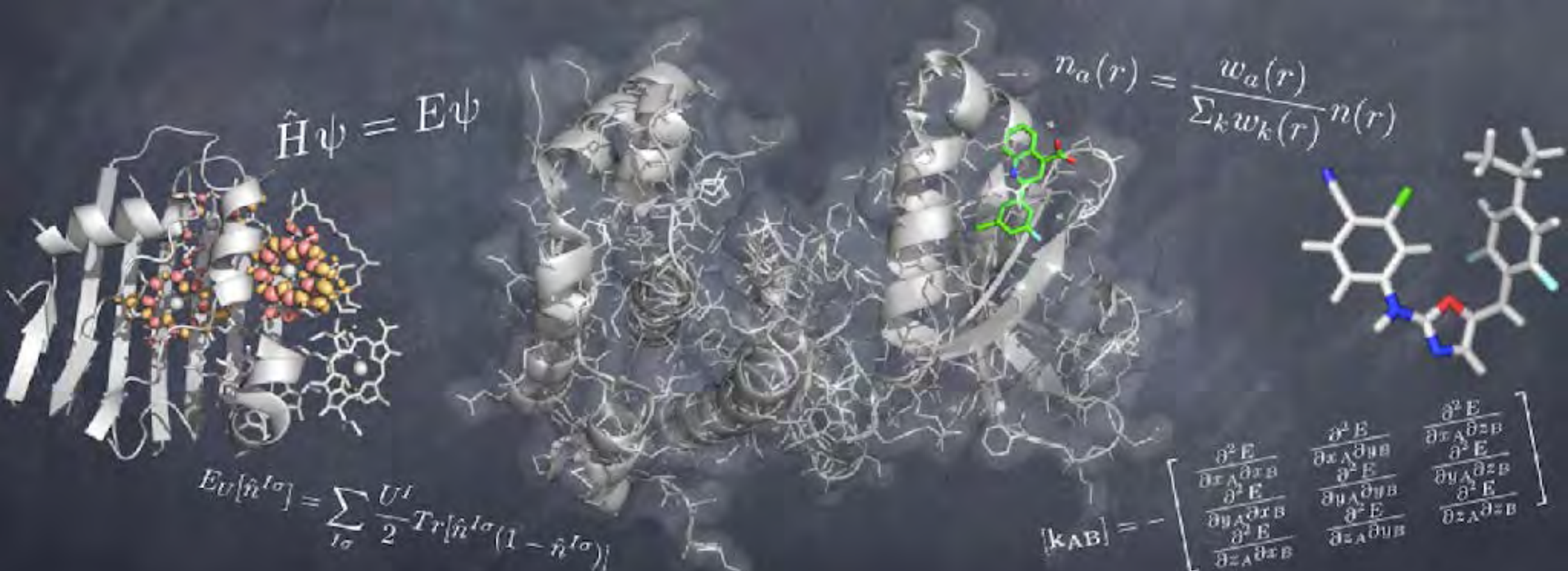
University of Cambridge
Alice Allen
Edward Linscott
Lupeng Yang
Louis Lee
Greg Lever
Alexander Fokas
Mike Payne

Collaborators
Alex Chin
Nick Hine (Warwick)
Cedric Weber (KCL)
Mohamed Ali al-Badri (KCL)
David O'Regan (TCD)
Stephen Wells (Bath)
David Huggins (Cornell)
Ashok Venkitaraman (Cambridge)
Gábor Csányi (Cambridge)
Chris-Kriton Skylaris (Southampton)
Peter Haynes (Imperial)
Thomas Manz (New Mexico State)

EPSRC
Engineering and Physical Sciences Research Council

SEVENTH FRAMEWORK PROGRAMME

MARIE CURIE ACTIONS



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

Newcastle University

Thank you for your attention