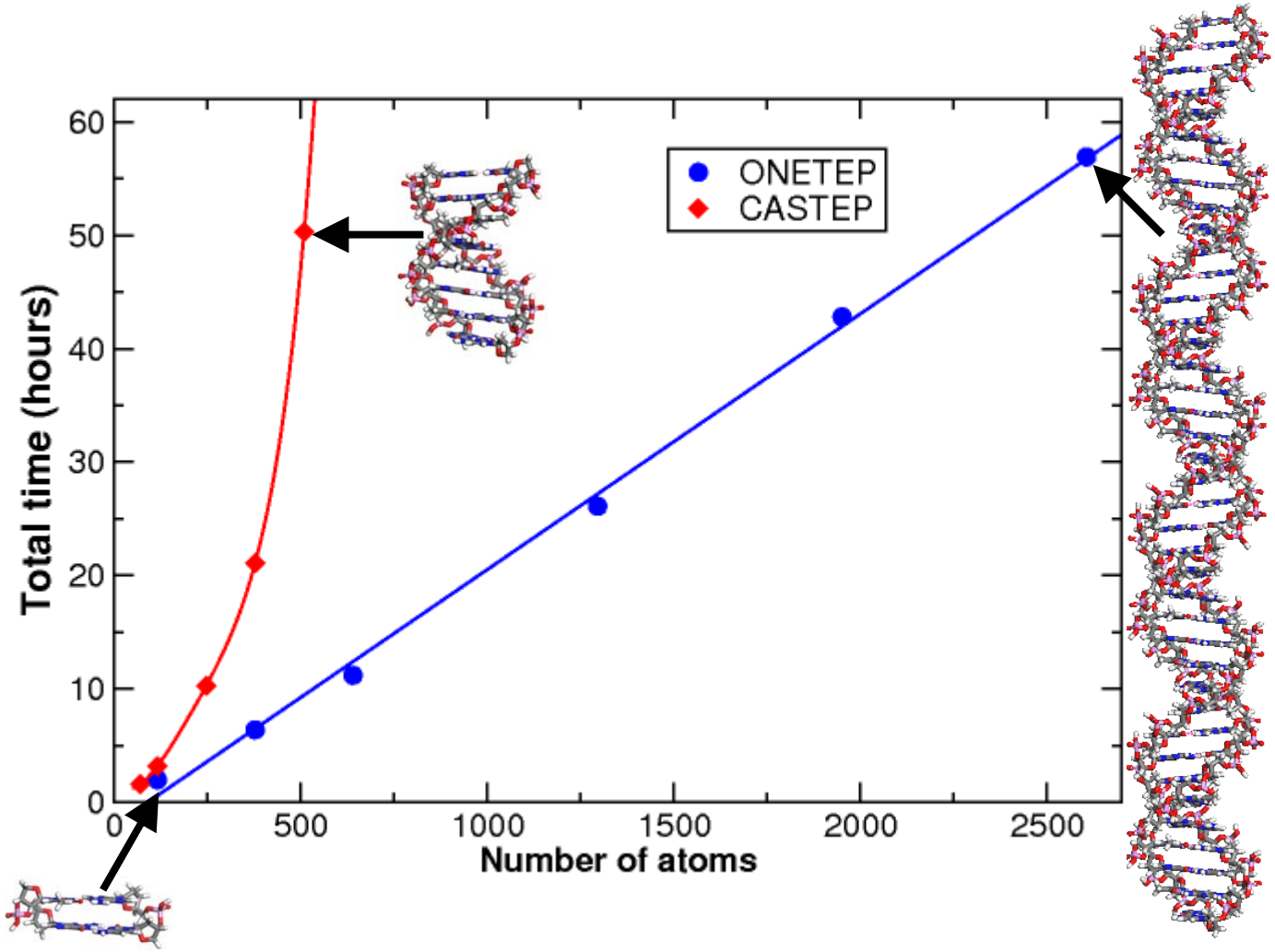
UNIVERSITY OF Southampton
School of Chemistry

Biomolecular simulations with ONETEP

Chris-Kriton Skylaris

2004: First version of ONETEP

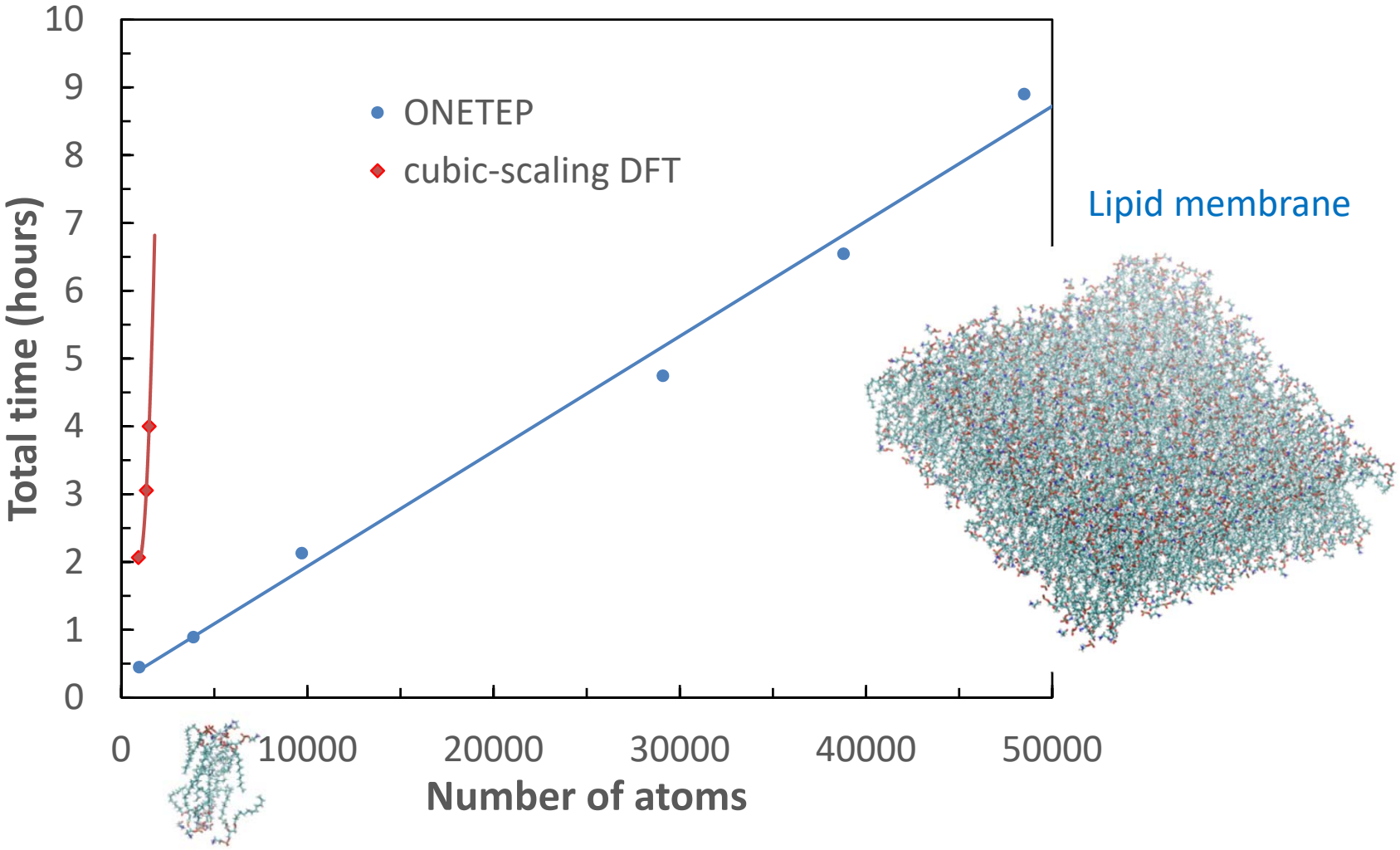
Calculations on 64 cores of “Franklin”, the 900-core SunFire 15K cluster of the CCHPCF



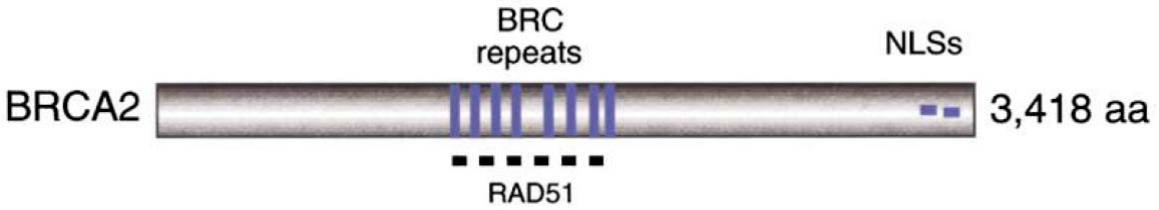
C.-K. Skylaris, P. D. Haynes, A. A. Mostofi and M. C. Payne, *J. Chem. Phys.* **122**, 084119 (2005).

2019: ONETEP v 5.2

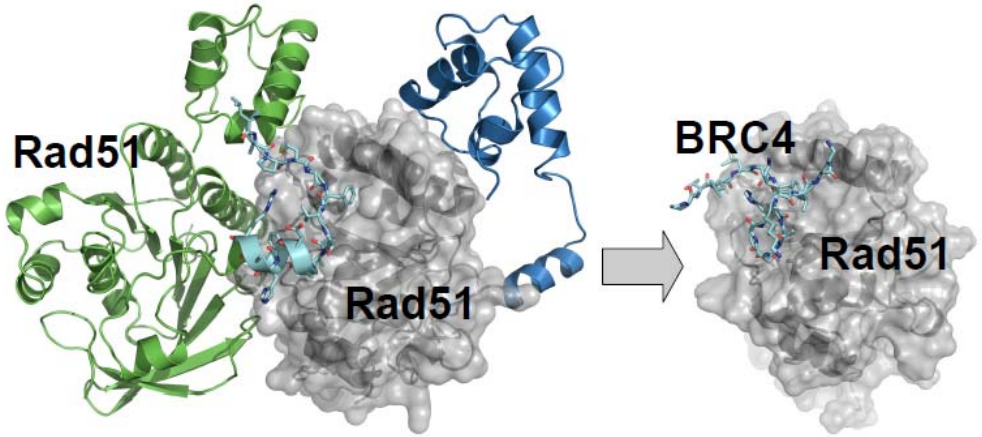
Calculations on1280 cores (MPI + OMP) Intel Skylake processors (Southampton)



Protein-protein interactions: Breast-cancer susceptibility protein (BRCA2)



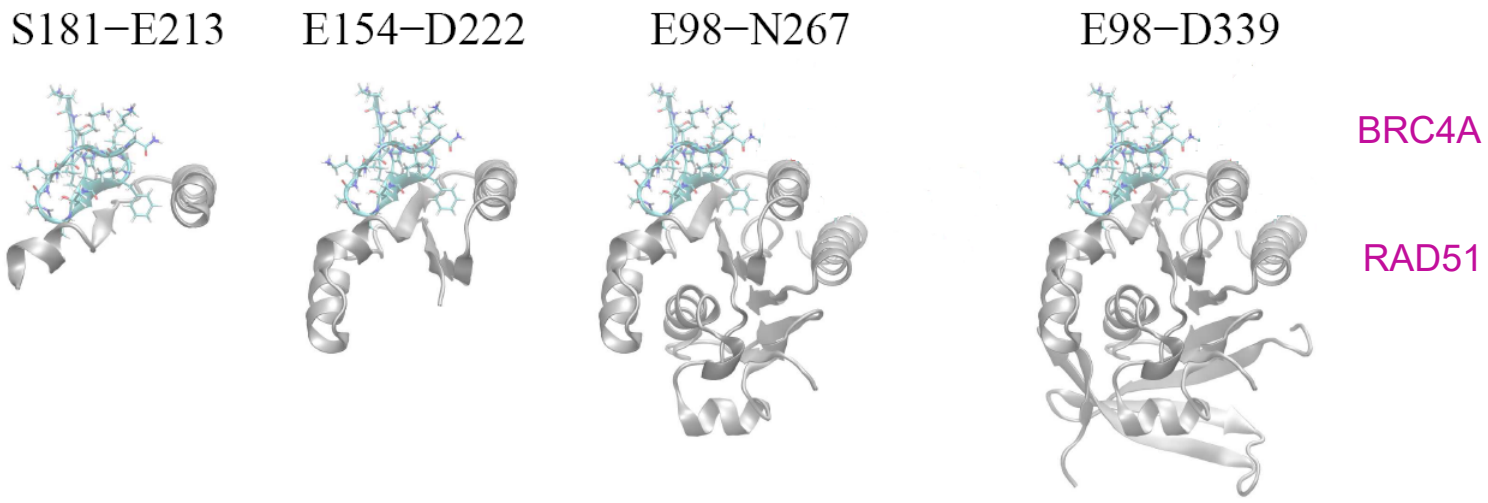
- Contains 8 BRC “repeats” of about 35-40 aa: BRC1, BRC2,...,BRC8
- Binds to the DNA recombination and repair protein Rad51 during DNA repair by Homologous Recombination (HR). The BRC repeats bind to RAD51 and compete for the Rad51-Rad51 self-oligomerisation interface:



- Despite significant sequence similarity, BRC repeats have varying affinities for Rad51 in experiments
- Want to compute the relative free energies of binding of each BRC repeat to Rad51 and compare to Rad51-Rad51 binding.

D. J. Cole, C.-K. Skylaris, E. Rajendra, A. R. Venkitaraman and M. C. Payne. *Europhysics Letters* **91** (2010) 3700.
D. J. Cole, E. Rajendra, M. Roberts-Thomson, B. Hardwick, G. J. McKenzie, M. C. Payne, A. R. Venkitaraman and C.-K. Skylaris. *PLoS Computational Biology* **7** (2011) e1002096

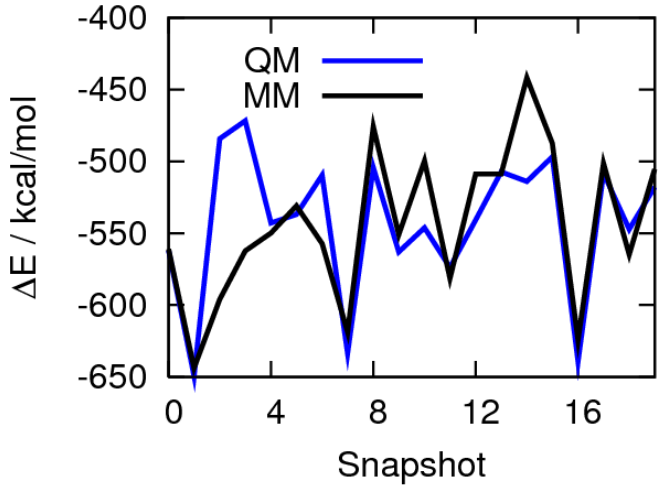
Convergence of binding free energies with RAD51 fragment size

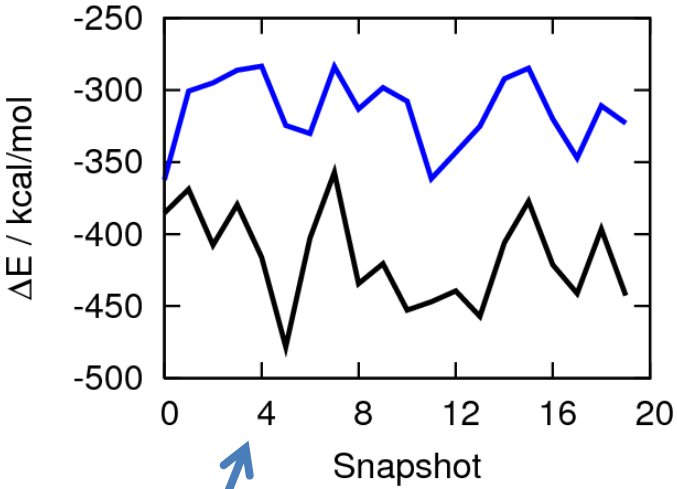


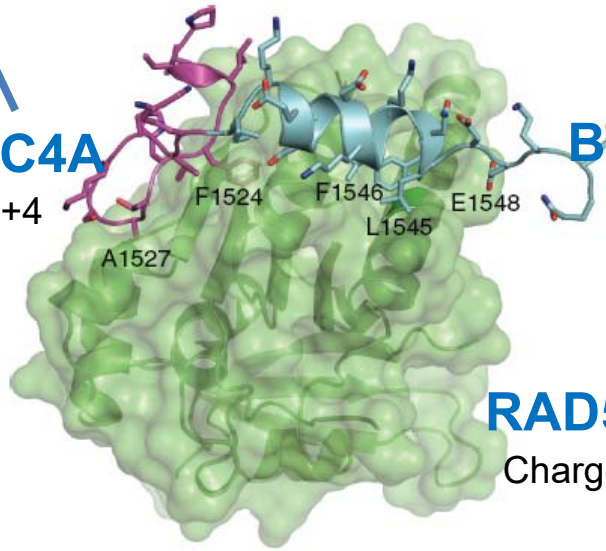
Residues	N_{Atoms}	ΔG_{MM}	ΔG_{QM}
S181-E213	737	-55.1	-63.1
E154-D222	1313	-57.7	-67.0
E98-N267	2780	-59.4	-65.6
E98-D339	3490	-59.5	-63.2

MM_PBSA and QM_PBSA calculations

Gas phase energies







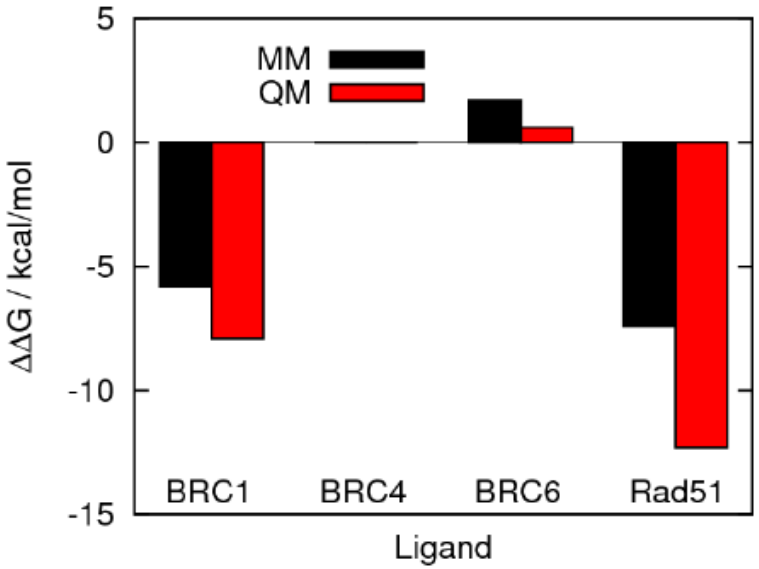
BRC4A
Charge +4

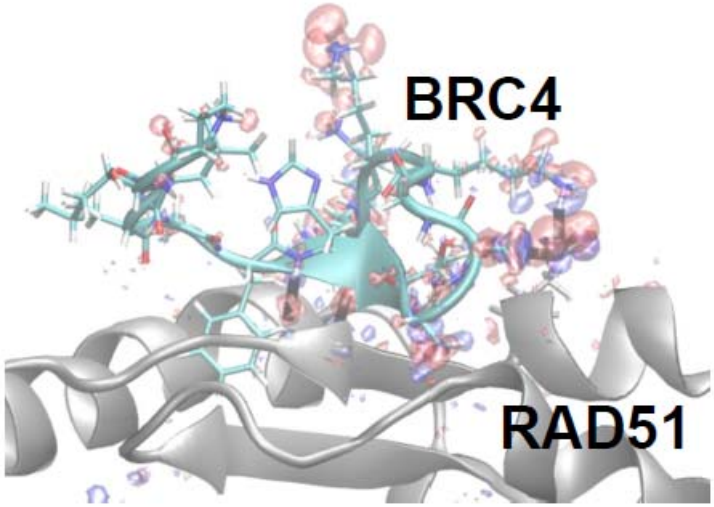
BRC4B
Charge 0

RAD51
Charge -4

UNIVERSITY OF Southampton
School of Chemistry

Relative free energies of binding





- Free energies of binding (relative to BRC4) in MM-PBSA and QM-PBSA are in good agreement and the order of binding affinity is unchanged.
- QM-PBSA still relies on the classical force field to sample the correct configurational space of the complex
- BUT, the solvent contribution has been taken into account in a very approximate way:

$$\Delta G_{PBSA}^{QM} = \Delta G_{PB} \times \left(\frac{\Delta E_{DFT}}{\Delta E_{EL}} \right)^{n_{PB}} + \Delta G_{SA}.$$

Implicit solvent model in ONETEP (2010 onwards)

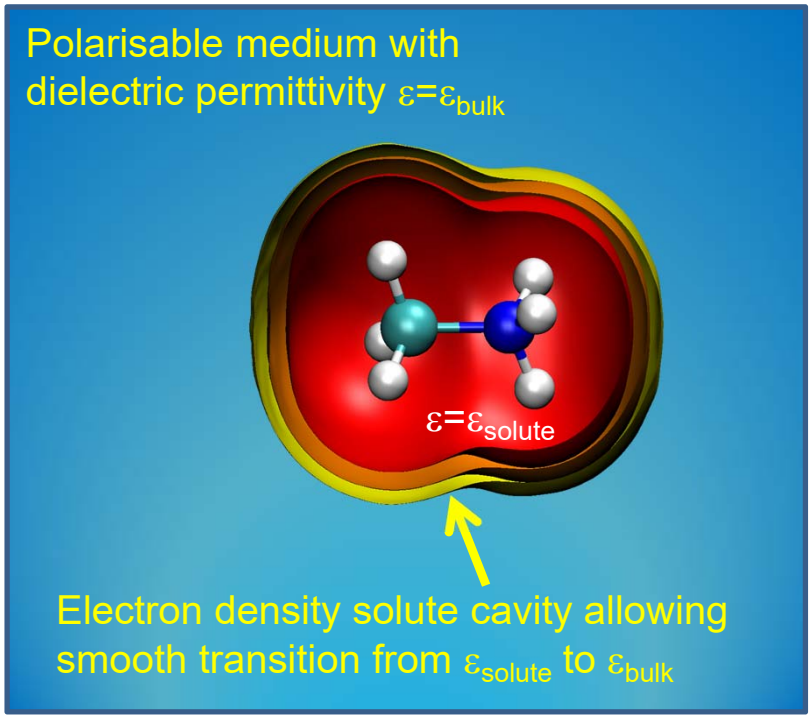
"Ab initio" or minimal parameter implicit solvent models

Fattebert and Gygi, *J. Comp. Chem.* **23**, 662 (2002)

- Solute cavity constructed from electron density of the solute
- Electrostatics by direct solution of the generalised Poisson equation
- Fully self-consistent: solute electrons polarised by solvent and vice versa

Our work [1]: accurate and highly parallel solvent model

- Reparametrized to provide accurate solvation free energies for neutral as well as for cationic and anionic species [1]
- Inclusion of dispersion-repulsion effects
- Parallel multigrid solver with high order defect correction – efficient use of supercomputing resources [2]
- Available within a robust linear-scaling DFT code (ONETEP) with near-complete basis set accuracy
- Can be part of multiscale simulation approaches, e.g. by combining with explicit solvent molecules

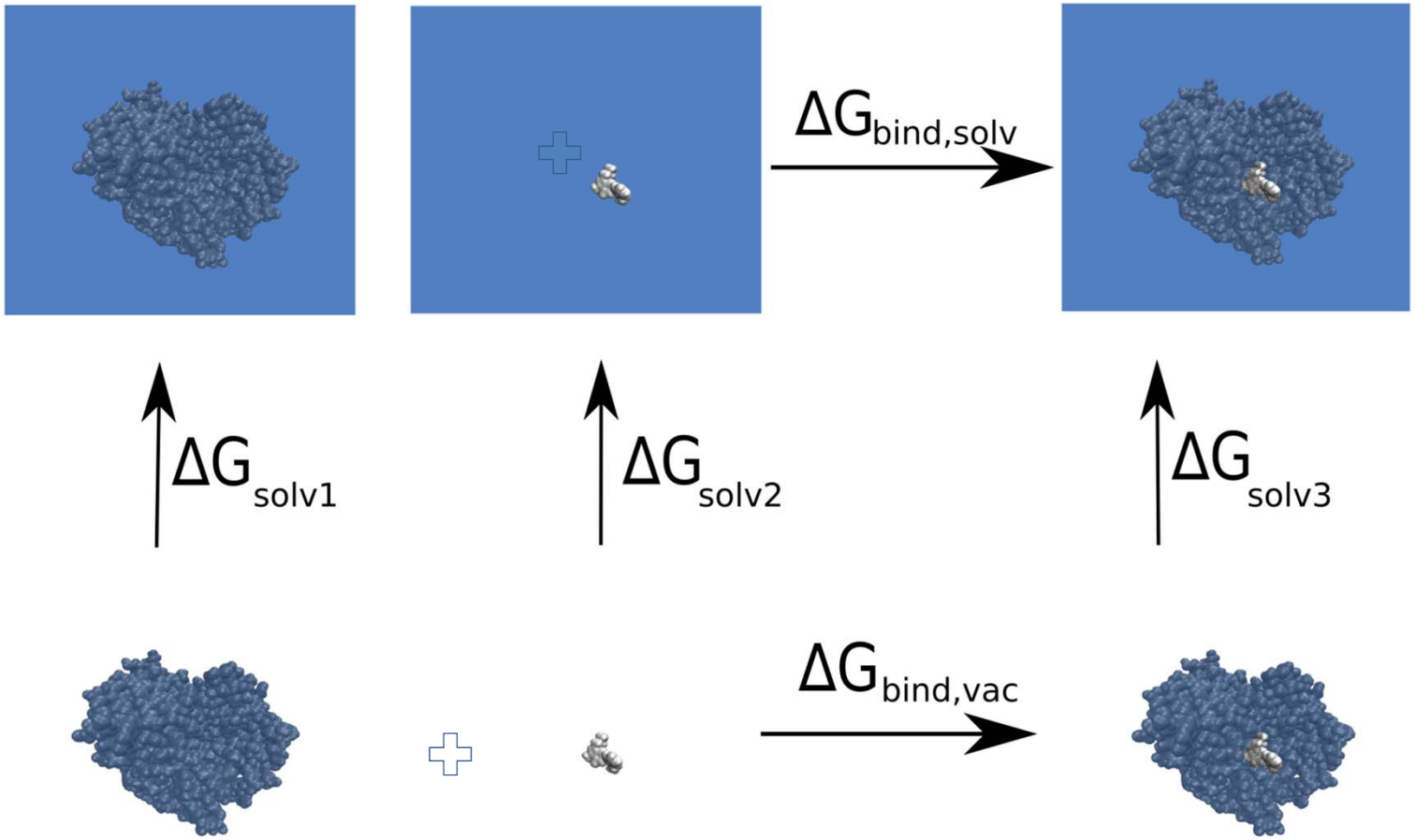


Approach	XC functional	rms error	max error	r
this work[a]	PBE	3.8	8.3	0.83
this work[b]	PBE	4.1	9.1	0.83
PCM	PBE	10.9	23.3	0.53
SMD	M05-2X	3.4	14.5	0.87
AMBER	(classical)	5.1	19.9	0.77

[1] Minimal parameter implicit solvent model for ab initio electronic structure calculations. J. Dziedzic, H. H. Helal, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne. *Europhysics Letters* **95** (2011) 43001

[2] DL_MG: A Parallel Multigrid Poisson and Poisson-Boltzmann Solver for Electronic Structure Calculations in Vacuum and Solution. J. C. Womack, L. Anton, J. Dziedzic, P. Hasnip, M. Probert, and C.-K. Skylaris. *J. Chem. Theor. Comput*. **14** (2018) 1412

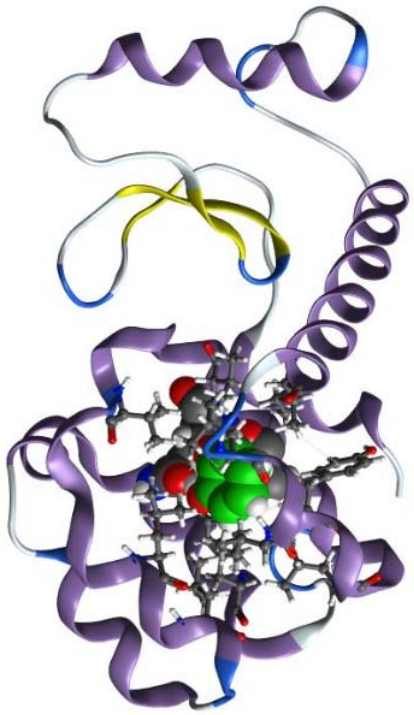
Protein – ligand free energies of binding



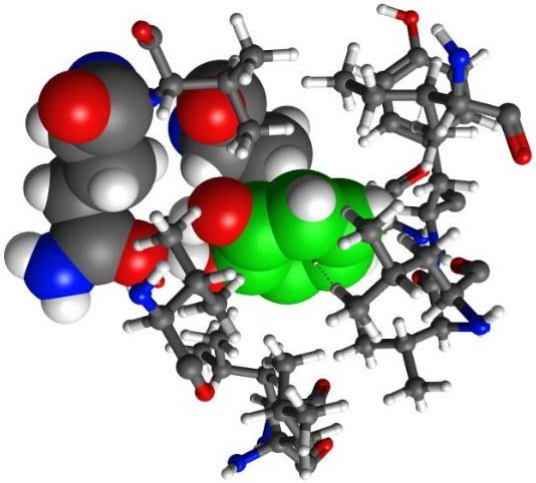
$$\Delta G_{bind,solv} = \Delta G_{bind,vac} + \Delta G_{solv3} - (\Delta G_{solv1} + \Delta G_{solv2})$$

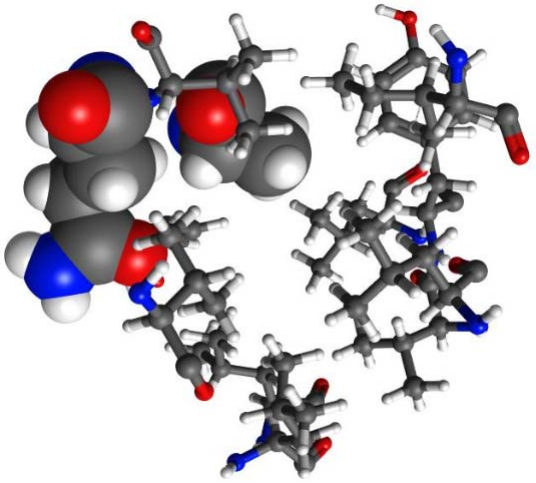
Applications to drug design: T4 Lysozyme L99A/M102Q protein

S. J. Fox, J. Dziedzic, T. Fox, C. S. Tautermann, and C.-K. Skylaris, *Proteins* **82** (2014) 3335-3346



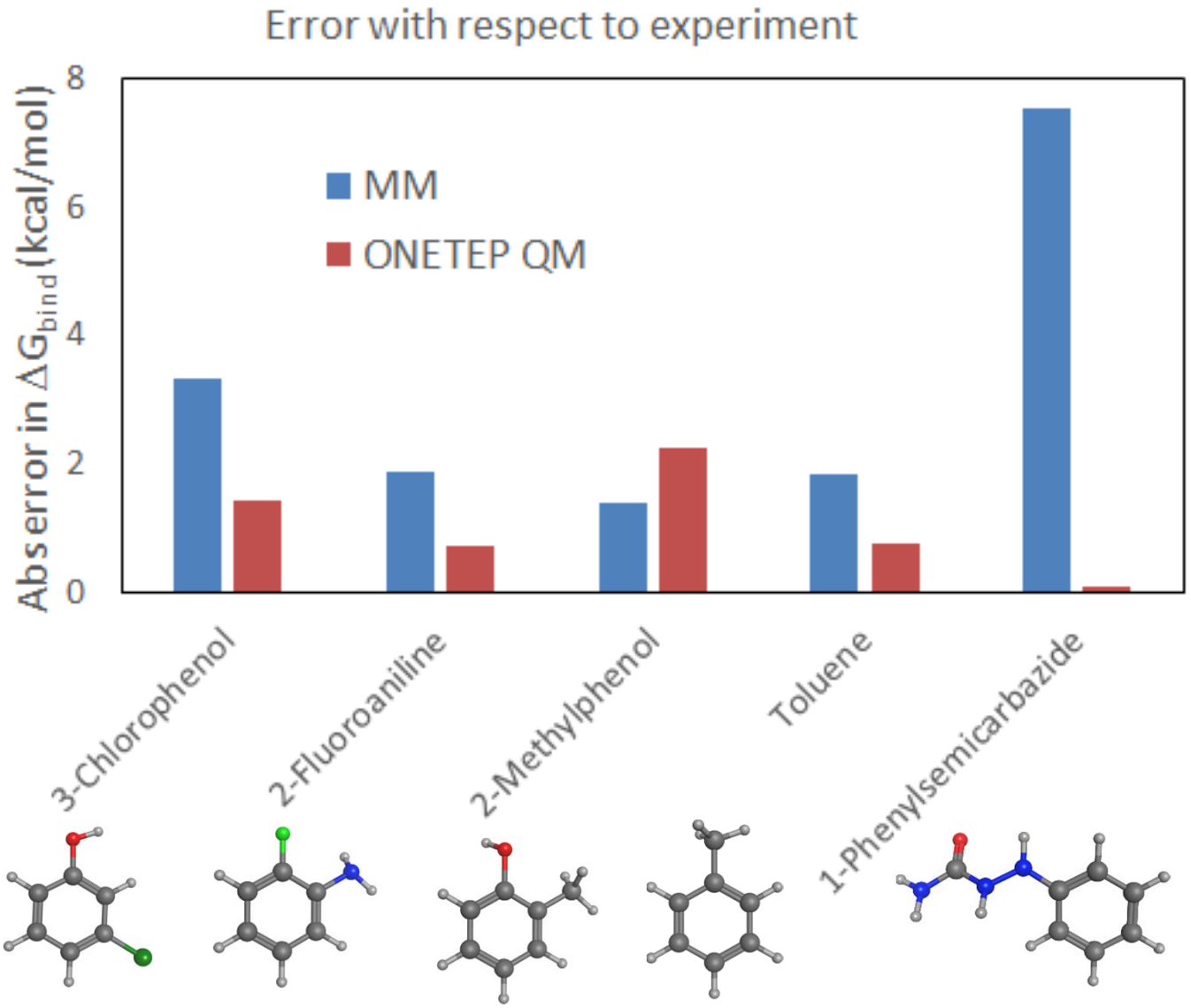
- 2616 atoms
- Polar binding site



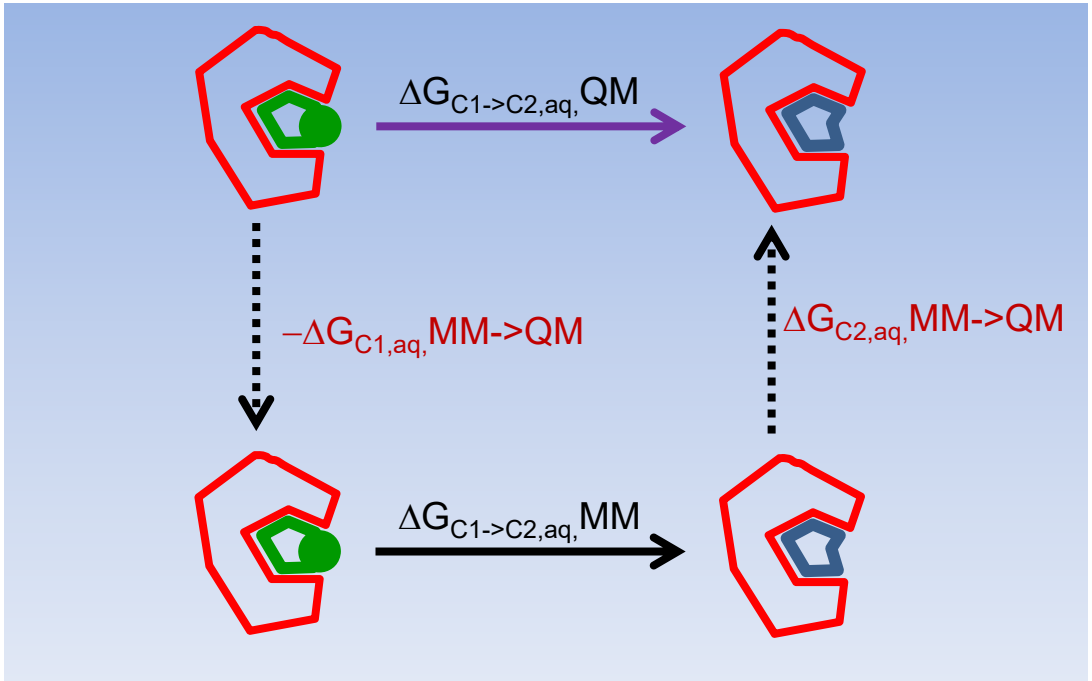


Complex of T4 Lysozyme L99A/M102Q and catechol

Free energies of binding: T4 Lysozyme L99A/M102Q



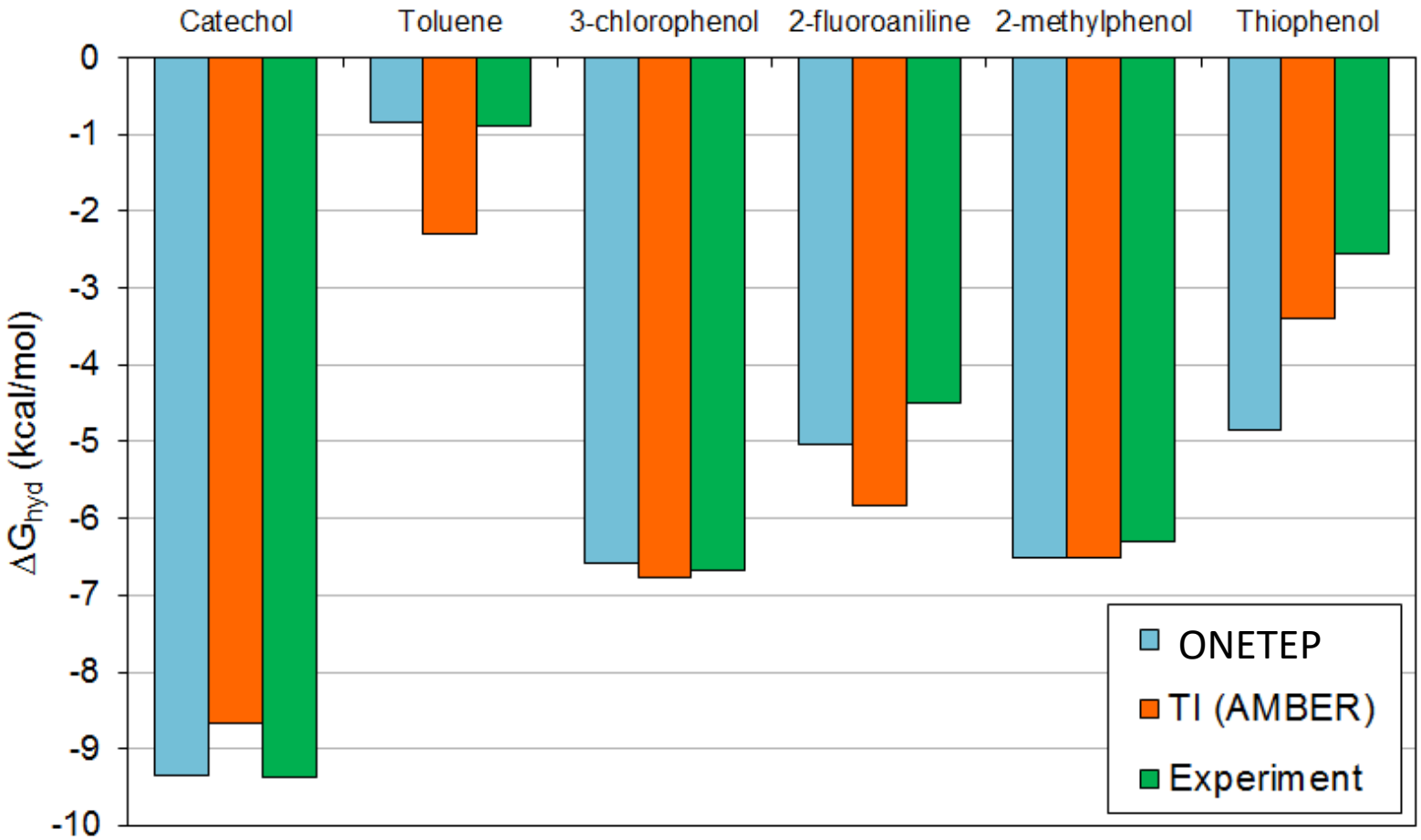
Rigorous methods for free energies in explicit water



How can we obtain $\Delta G_{C1,aq,}$MM->QM and $\Delta G_{C2,aq,}$MM->QM ?

Relative free energies of binding $\Delta\Delta G_{bind} = \Delta G_{bind2,aq} - \Delta G_{bind1,aq} = \Delta G_{C1->C1,aq} - \Delta G_{L1->L2,aq}$ TI, FEP

Free energies of hydration (300K)



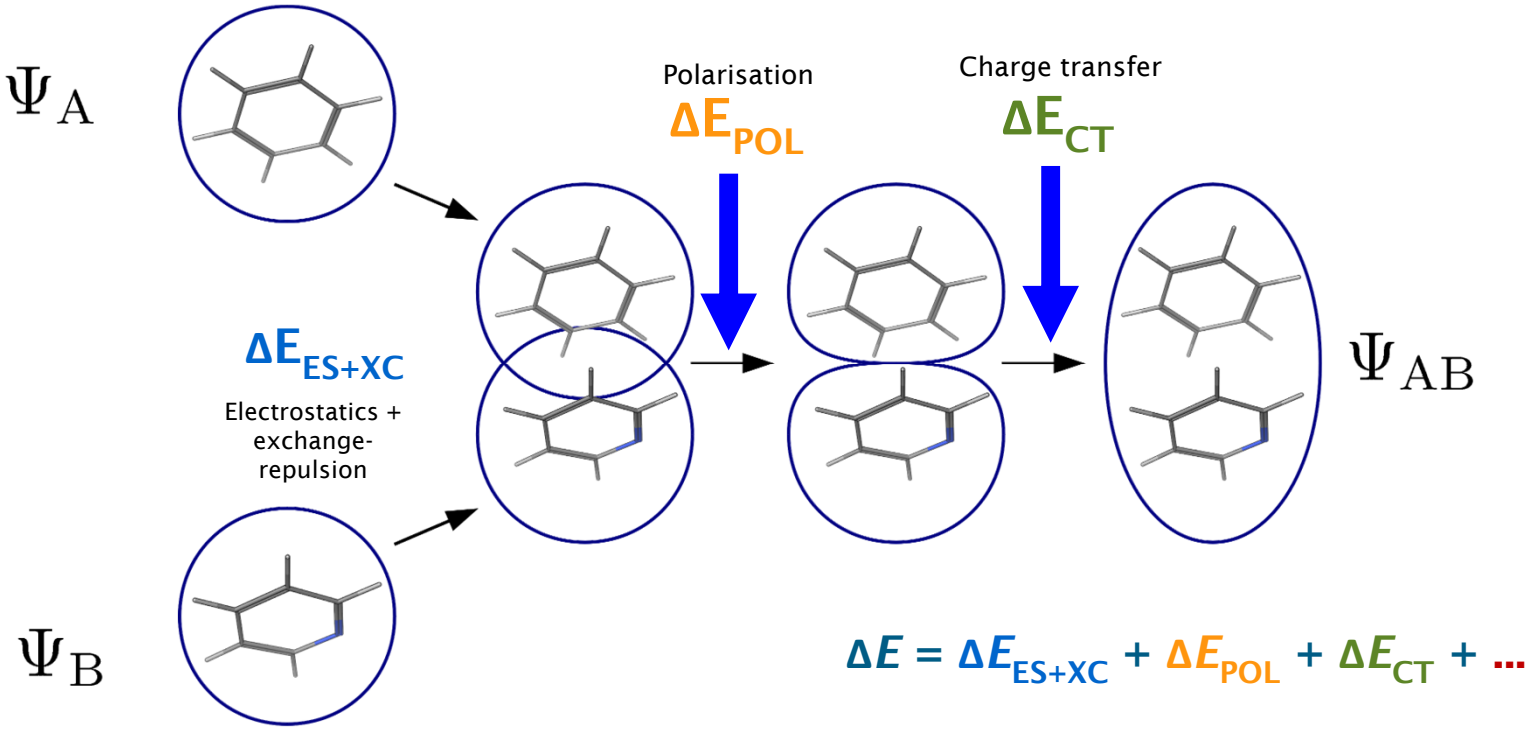
S. J. Fox, C. Pittock, C. S. Tautermann, T. Fox, C. Christ, N.O. J. Malcolm, J. W. Essex and C.-K Skylaris, *J. Phys. Chem. B* **117**, 9478 (2013)

Energy decomposition analysis (EDA)

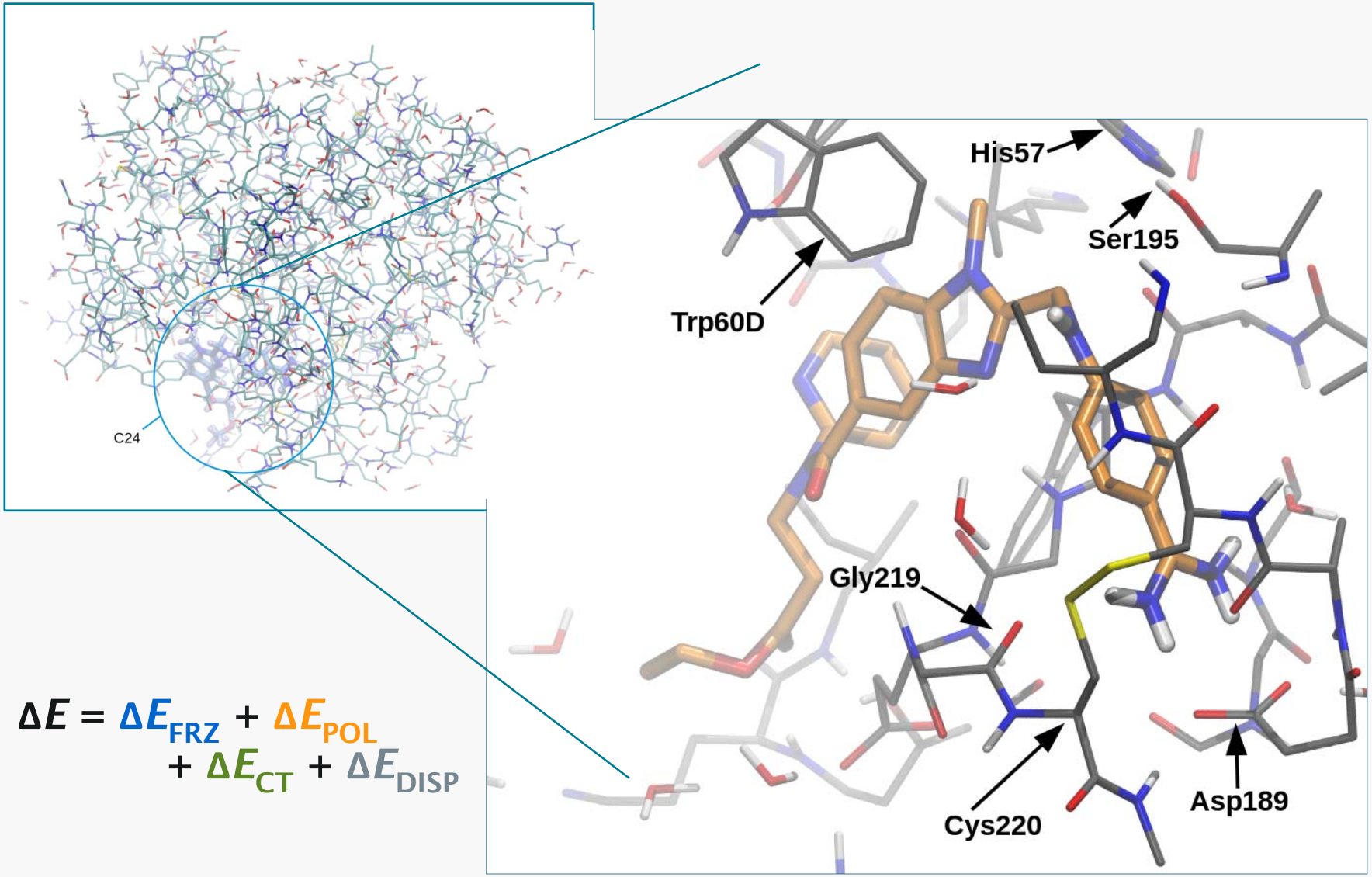
M. J. S. Phipps, T. Fox, C. S. Tautermann, C.-K. Skylaris, *Chem. Soc. Rev.* **44** (2015) 3177

M. J. S. Phipps, T. Fox, C. S. Tautermann, C.-K. Skylaris, *J. Chem.Theory Comput.*, **12** (2016) 3135

- Decompose any host-guest binding energy to chemically relevant components
- Inform the design of new host-guest ligands (e.g. in Pharma applications)
- Parameterise more approximate simulation methods (e.g. new MM approaches)



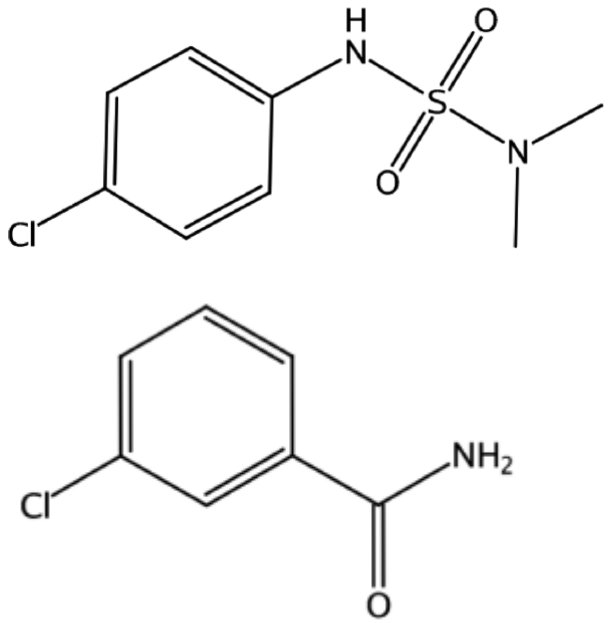
Thrombin protein-ligand complex (4975 atoms)



$$\Delta E = \Delta E_{FRZ} + \Delta E_{POL} + \Delta E_{CT} + \Delta E_{DISP}$$

Example: comparing L1 with L2

Energy Term (kcal/mol)	L1	L2
FRZ	2.1	2.8
ES	-35.2	-30.3
EX	-19.5	-16.5
REP	82.8	69.7
CORR	-26.0	-20.2
POL	-6.1	-5.3
CT	-11.4	-11.1
ΔE_{vac}(PBE/800eV)	-15.4	-13.6
SOLV	13.1	11.1
ΔE(PBE/800eV)	-2.3	-2.5



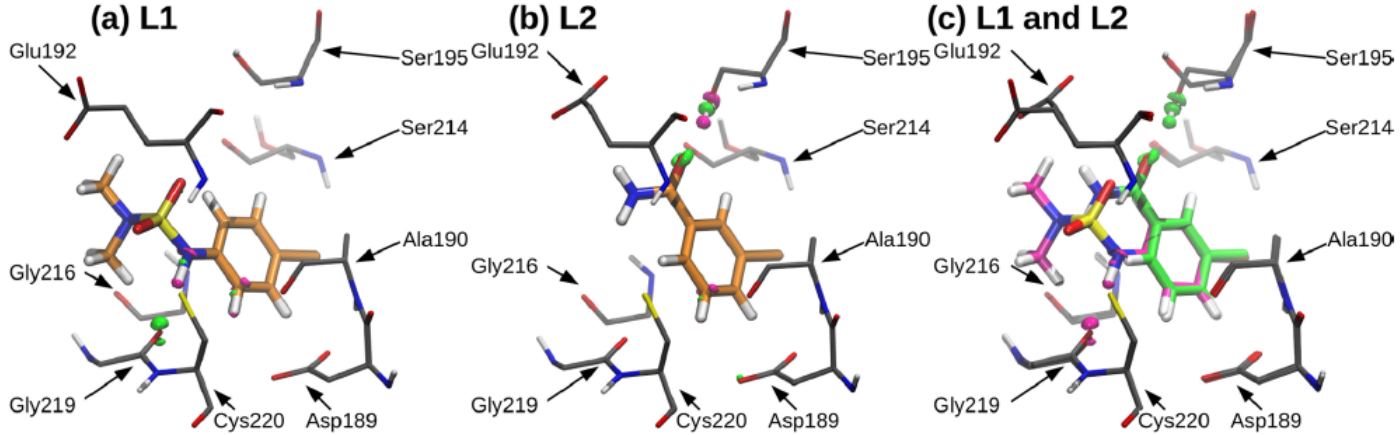
- Structurally comparable ligands with chlorobenzene group in common
- Essentially the same binding energies (-2.3 kcal/mol and -2.5 kcal/mol) but the actual EDA components are very different
- L1 has a Pauli (steric) repulsion which is 13.1 kcal/mol more unfavourable than L2. However this is counteracted by the other energy components which add up to eliminate this difference

Electron Density Difference (EDD)

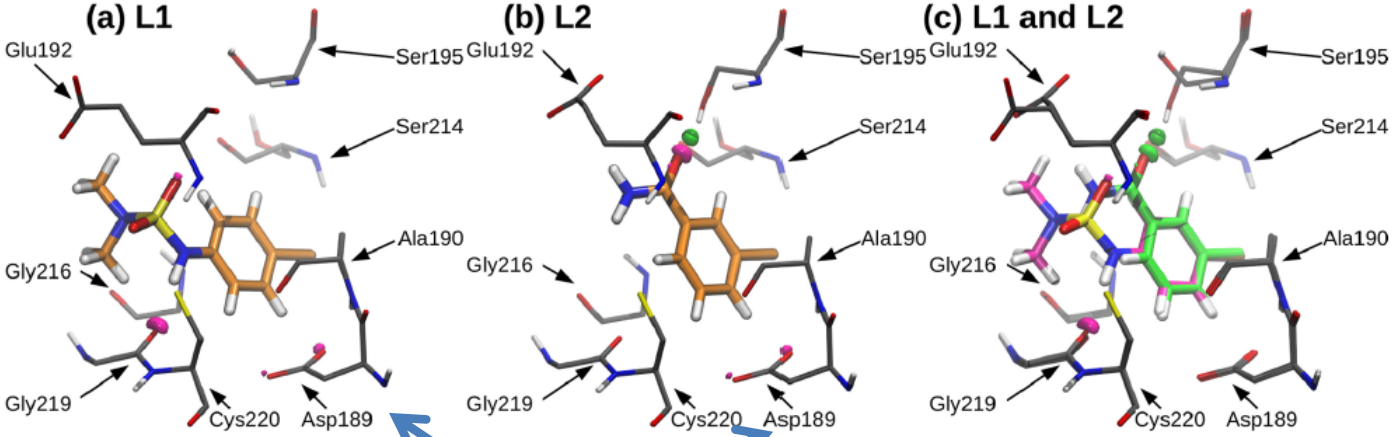
Visualising the interactions: L1 interacts with Gly219 L2 interacts with Ser195

ΔEDD

Polarisation EDD



Charge transfer EDD

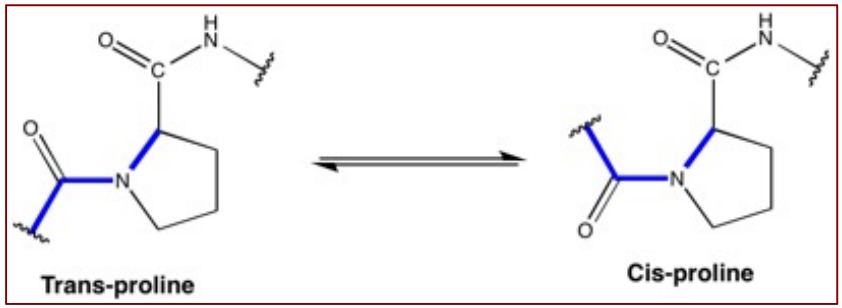


Common interactions (with Asp189) cancel out in ΔEDD

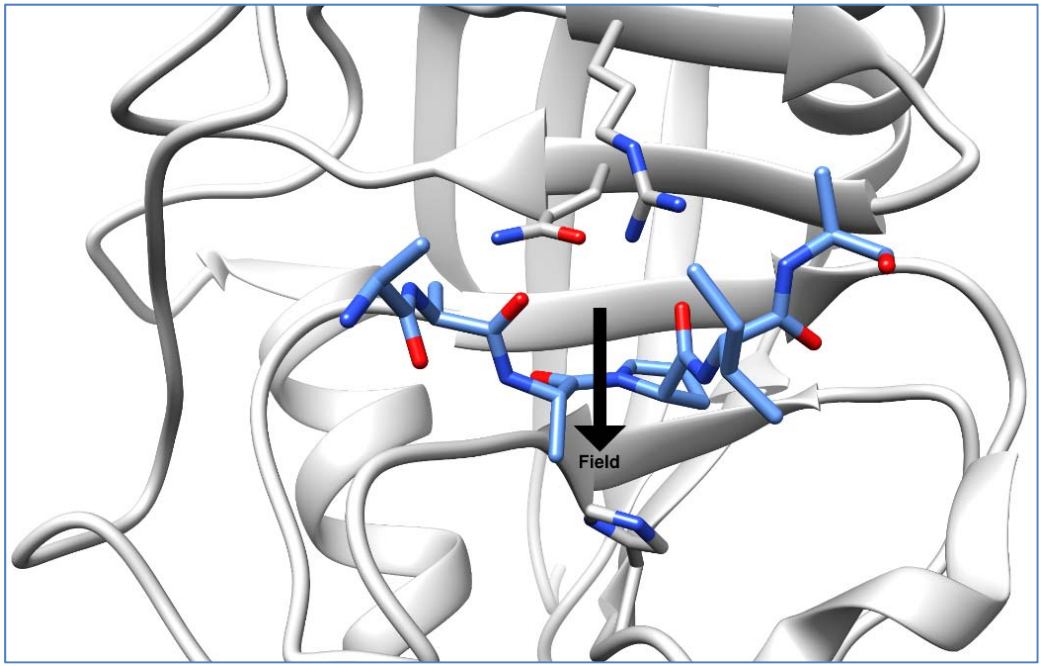
University of Southampton School of Chemistry

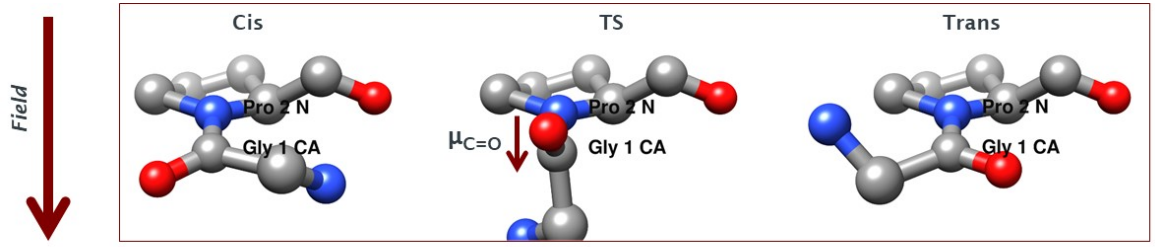
Using ONETEP to compute electric fields in enzymes: Cyclophilin A

Peptidylprolyl isomerase: Catalyses the cis/trans isomerisation of proline:



Proposed 'electrostatic handle'[1]





Environment field aligned with carbonyl dipole of preceding residue at transition state

1) Camilloni C. et al., PNAS, 2014, **28**, 10203-10208

UNIVERSITY OF Southampton
School of Chemistry

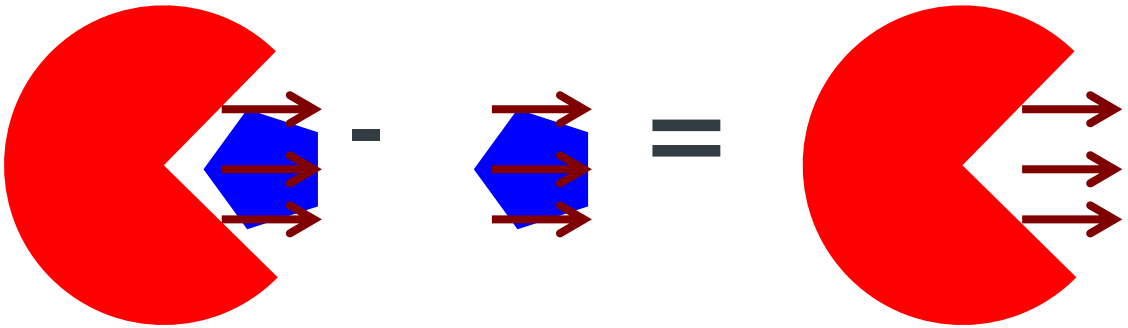
Calculation of electric fields in the enzyme active site with

- ONETEP
- A polarisable force field (AMOEBA)
- Fixed-charge force fields (AMBER, CHARMM)

$$U = U_{bond} + U_{angle} + U_{b\theta} + U_{oop} + U_{torsion} + U_{vdW} + U_{ele}^{perm} + U_{ele}^{ind}$$

Fundamentally, AMOEBA models polarisation *via* induced atomic dipoles:

'Environment field'



Calculate electric fields within the active site

UNIVERSITY OF Southampton
School of Chemistry

Comparison of calculated fields

- x,y,z field components at the carbonyl bond centre
- Combination of data from cis, trans, WT and R55A structures
- ONETEP calculations on 30 snapshots with ~20,000 atoms
- Fixed charge force fields slightly underestimate the magnitude of field components; weaker correlation; greater variability
- AMOEBA shows good correlation with DFT

Richard T. Bradshaw, Jacek Dziedzic, Chris-Kriton Skylaris, and Jonathan W. Essex, *In preparation*

UNIVERSITY OF Southampton
School of Chemistry

Thank you Mike!

Happy Birthday!