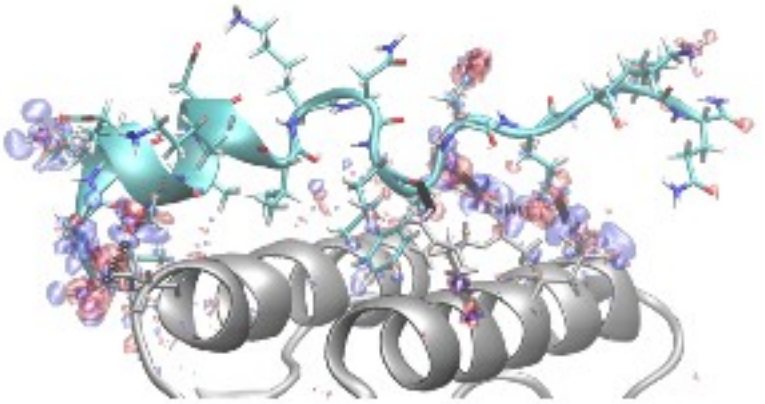
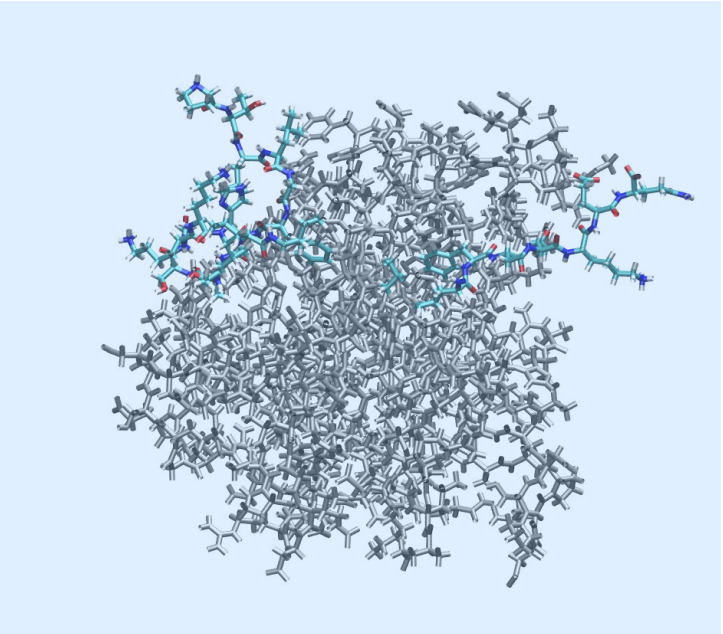
Protein-Protein Interactions from Linear-Scaling DFT Calculations



Danny Cole, Chris-Kriton Skylaris, Eeson Rajendra, Ashok Venkitaraman and Mike Payne

Department of Physics, University of Cambridge

Introduction



Protein-protein interactions are fundamental to all biological processes.

The interrogation of these interactions is important in the field of small molecule therapeutic intervention.

Classical approaches are capable of modelling real complexes of 1000s of atoms immersed in water over ns time scales.

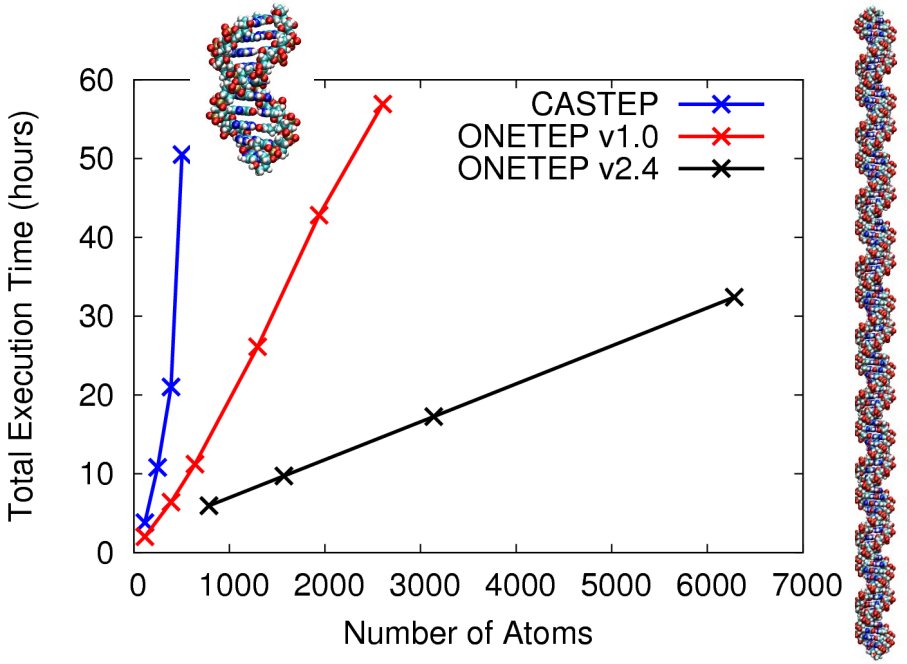
More accurate first principles approaches are traditionally limited by the number of atoms and limited time scales.

Combined approach allows the determination of relative free energies of binding of complexes of 1000s of atoms to high accuracy and with high computational efficiency.

ONETEP

ONETEP achieves plane wave accuracy with with a computational cost that scales linearly with the number of atoms.

Example: Timings for a DNA α-helix with counter-ions (N. Hine)



This allows us to perform DFT calculations on entire protein complexes consisting of 1000s of atoms.

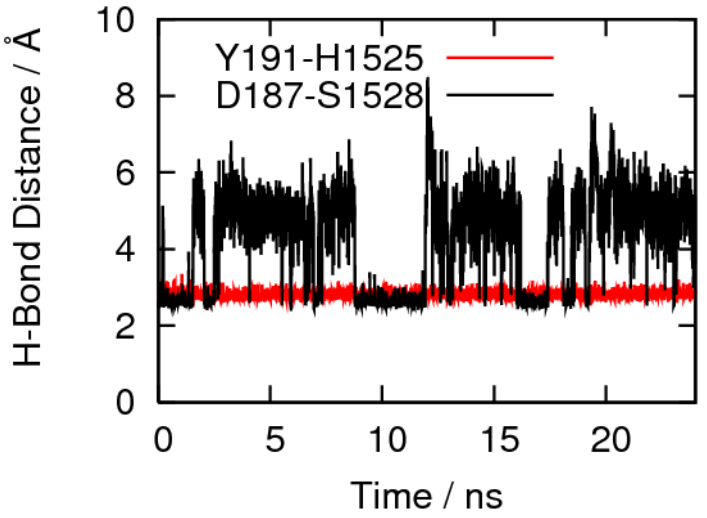
C. -K. Skylaris, P. Haynes, A. Mostofi, N. Hine, M. Payne

Problem Solved?

Time scale

Interactions in biological systems are dynamic over a much longer time scale than we can model using DFT. A single DFT calculation on the crystal structure would not account for this.

Example: Time dependence of two hydrogen bonds in the RAD51-BRC4 interaction



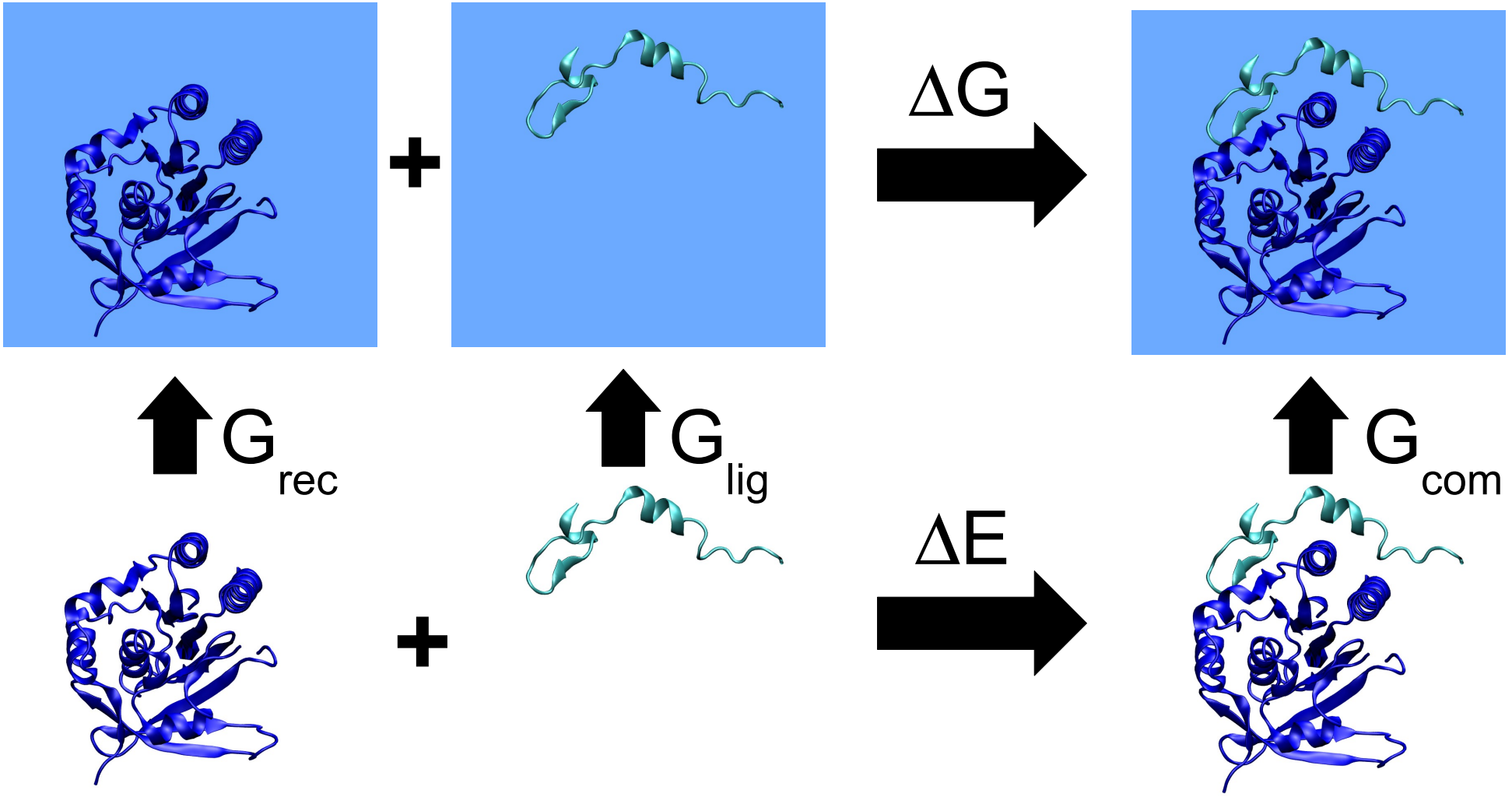
Solvent Effects

Gas phase binding energies massively over-estimate binding, since in biological systems the separated molecules are stabilised by interactions with the solvent.

We adapt the classical MM-PBSA technique, which accounts for these effects, to include gas phase binding energies calculated by QM

...QM-PBSA

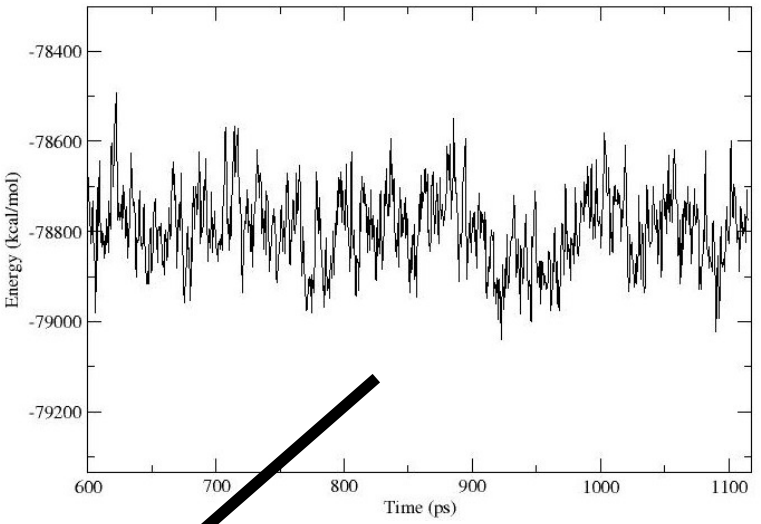
MM-PBSA

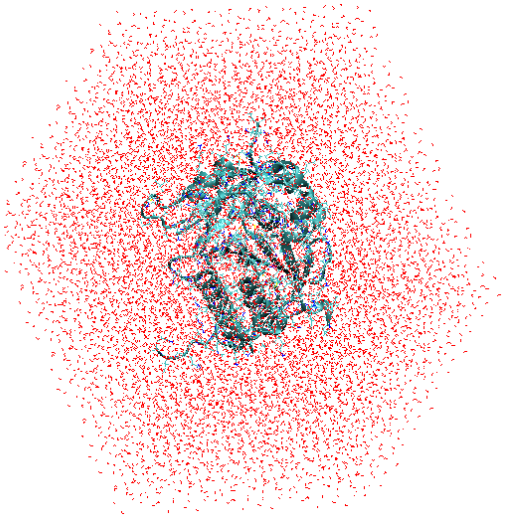


$$\Delta G = \Delta E + (G_{com} - G_{rec} - G_{lig}) = \Delta E + \Delta G_{PB}$$

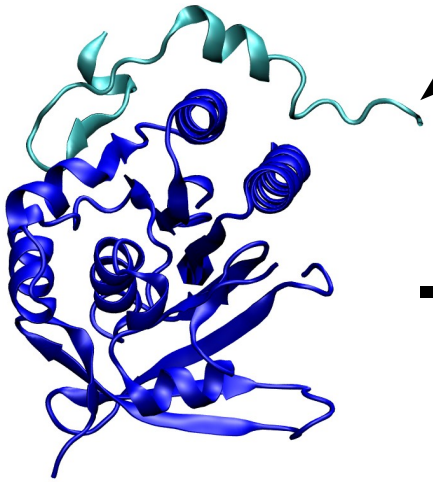
MM-PBSA (Single Trajectory)

1) Run MD simulation of complex in explicit water using classical force field.





2) Sample snapshots of trajectory, removing water molecules.



3) Calculate average binding free energy of complex in implicit solvent → ΔG

Massova and Kollman, JACS 121, 8133 (1999)

MM-PBSA approach

Free energy of binding calculated as averages over MD snapshots of gas phase energy, harmonic entropy and implicit solvation energy:

$$\Delta G_{MM} = \Delta E_{MM} - T\Delta S_{MM} + \Delta G_{PB}$$

Entropic terms are often assumed to cancel out in differences in free energy of binding when the ligands have structural similarity:

$$\Delta\Delta G_{MM} = \Delta\Delta E_{MM} + \Delta\Delta G_{PB}$$

QM-PBSA approach

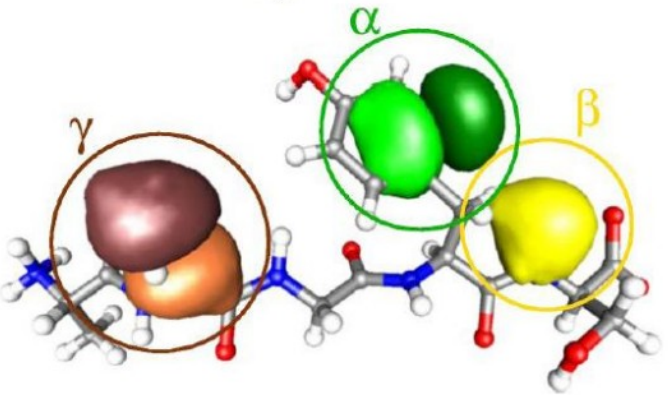
Based on a classical trajectory but obtaining the energies of the snapshots from quantum calculations (DFT) on the entire molecule:

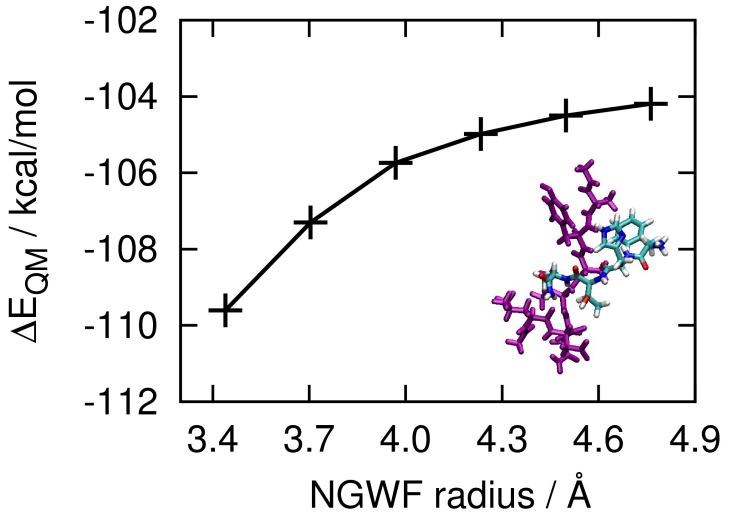
$$\Delta\Delta G_{QM} = \Delta\Delta E_{QM} + \Delta\Delta G_{PB,QM}$$

Some Notes on ΔE_{QM}

In ONETEP, the density matrix is expressed in terms of a set of non-orthogonal generalised Wannier functions (NGWFs) that are localised in real space:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}')$$



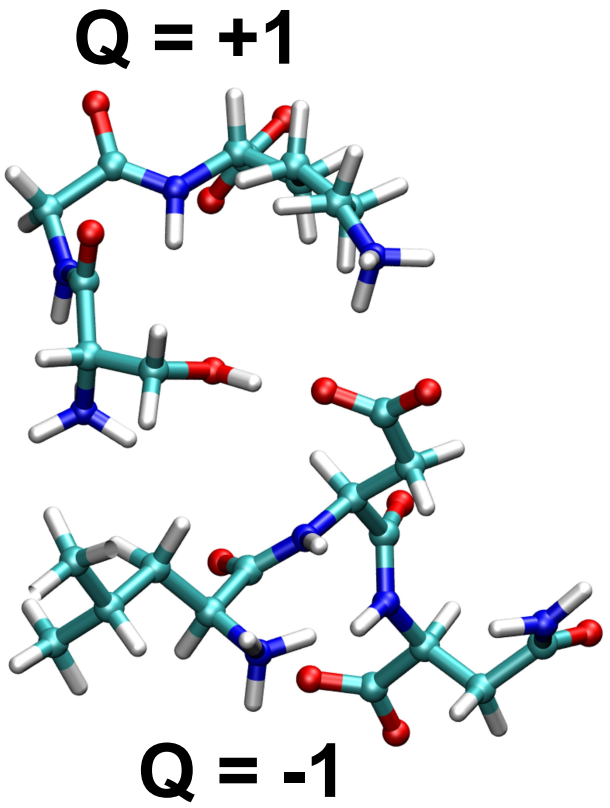


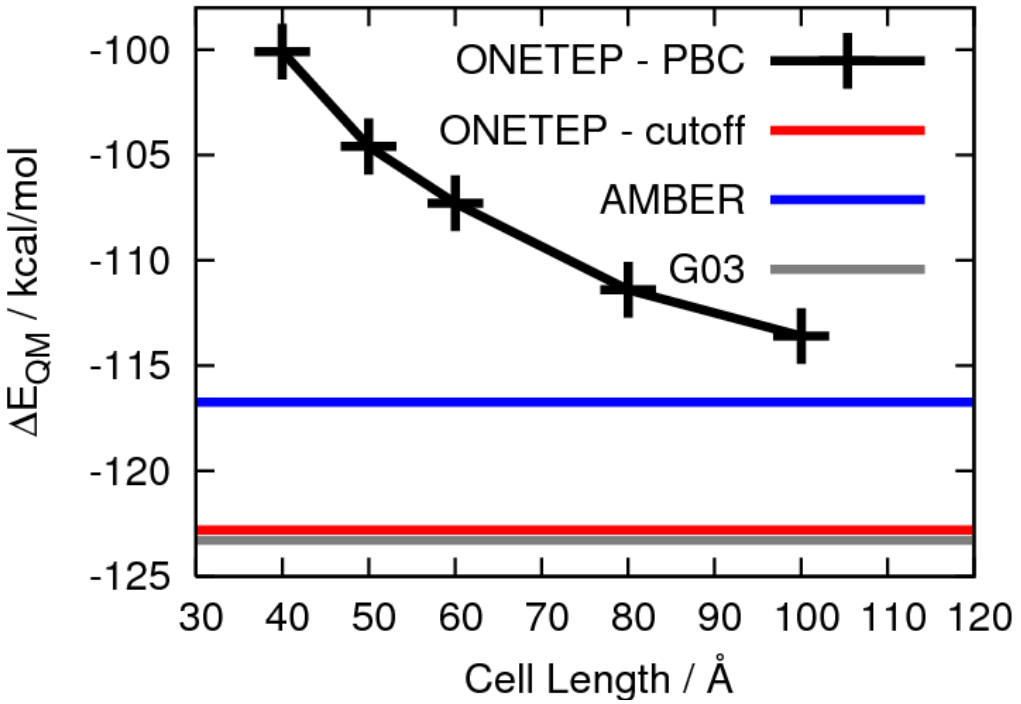
The convergence of ΔE_{QM} may be systematically improved by increasing the NGWF radii and the energy cut-off of the psinc basis functions.

ΔE_{DFT} is augmented by damped London potentials with parameters optimised for the PBE functional.

Q. Hill, C. -K. Skylaris

Cut-Off Coulomb Interaction

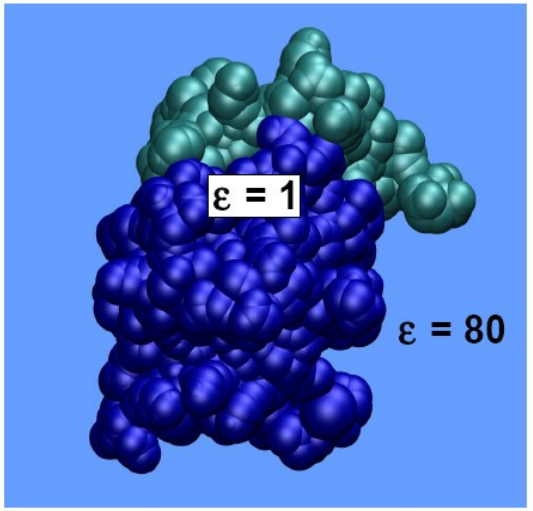




Cut-off Coulomb: use spherical cut-off and padded cell to remove spurious Hartree interactions between neighbouring cells.

G03: Gaussian03, PBE xc functional, 6-311G* basis set, counterpoise correction, augmented by ONETEP dispersion interaction.

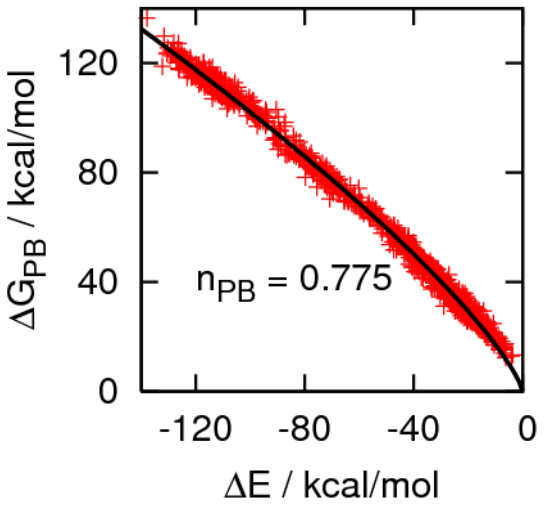
How to find $\Delta G_{PB,QM}$?



In future, $\Delta G_{PB,QM}$ will be derived from a grid-based solution of the Poisson-Boltzmann equation assuming high dielectric solvent, low dielectric solute and with MM partial atomic charges replaced by the DFT electron density:

$$\nabla \cdot (\epsilon(r) \nabla \phi(r)) = -4\pi \rho(r)$$

see work by H. Helal, A. Mostofi and M.Payne and talk by J. Dziedzic



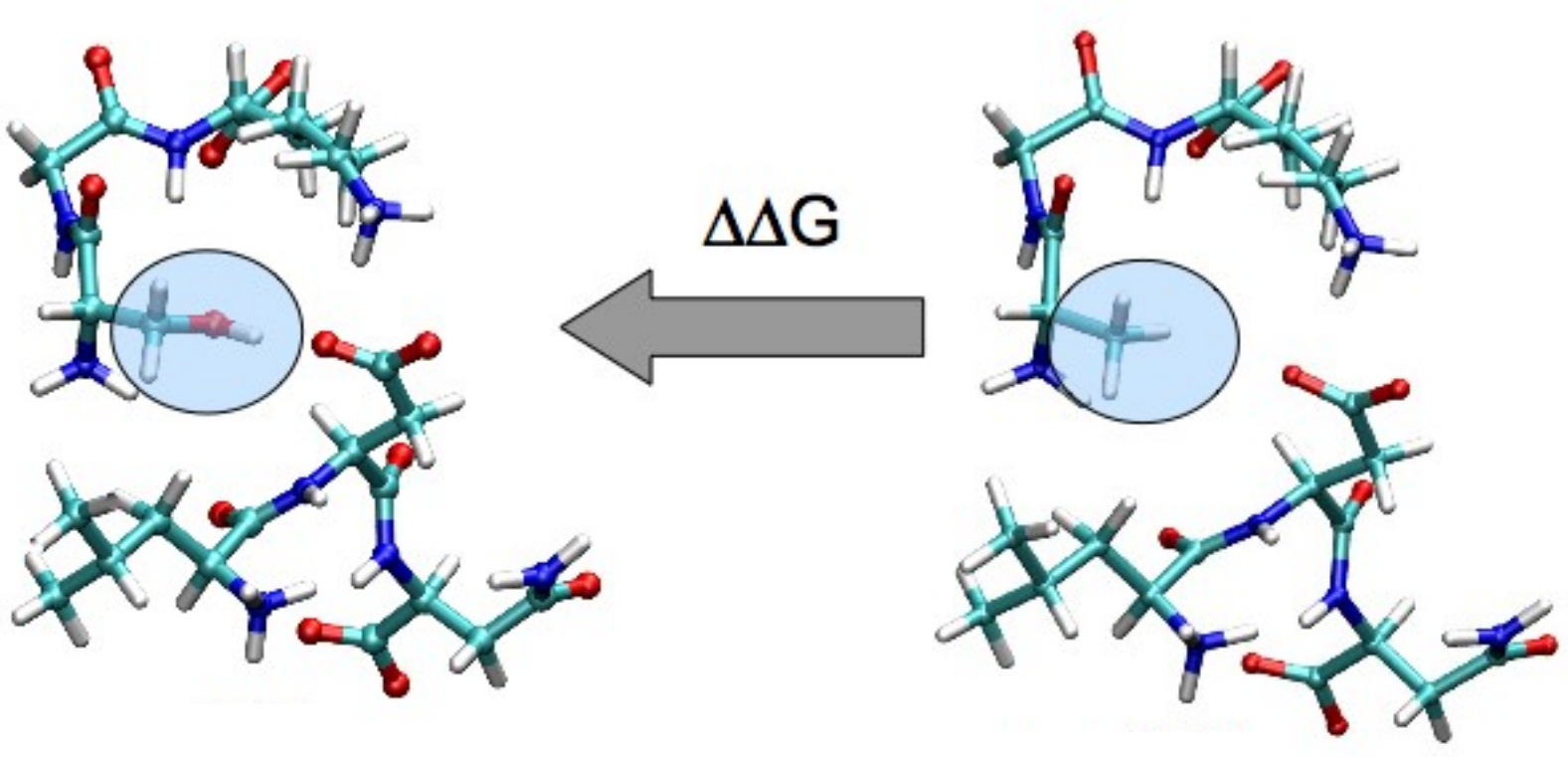
For now, we can see that there is a trend between ΔG_{PB} and ΔE and we get a good power law fit with:

$$\Delta G_{PB} \propto (\Delta E)^{npb} \qquad (n_{pb} < 1)$$

If we assume that $\Delta G_{PB,QM}$ follows the same power law, then:

$$\Delta G_{PB,QM} = \Delta G_{PB} \times \left(\frac{\Delta E_{DFT}}{\Delta E_{ELE}} \right)^{n_{PB}}$$

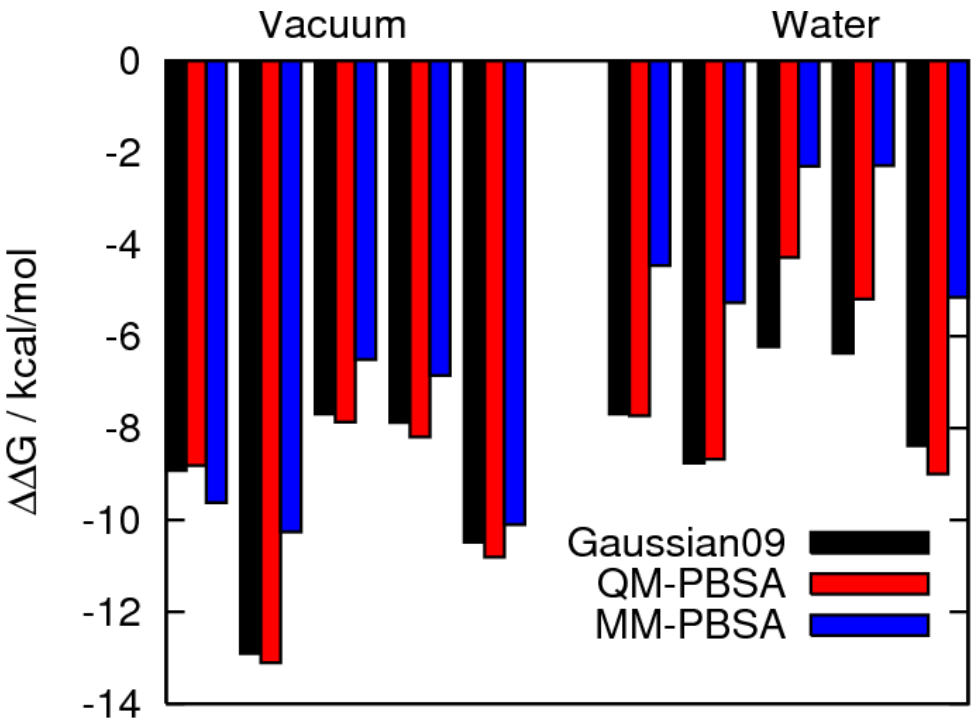
Computational Alanine Scan



Very common computational and experimental technique for determining contributions of different residues to binding. MM-PBSA can get reasonably accurate results.

Here we study a SER → ALA mutation and compare to MM-PBSA and benchmark Gaussian09 calculations.

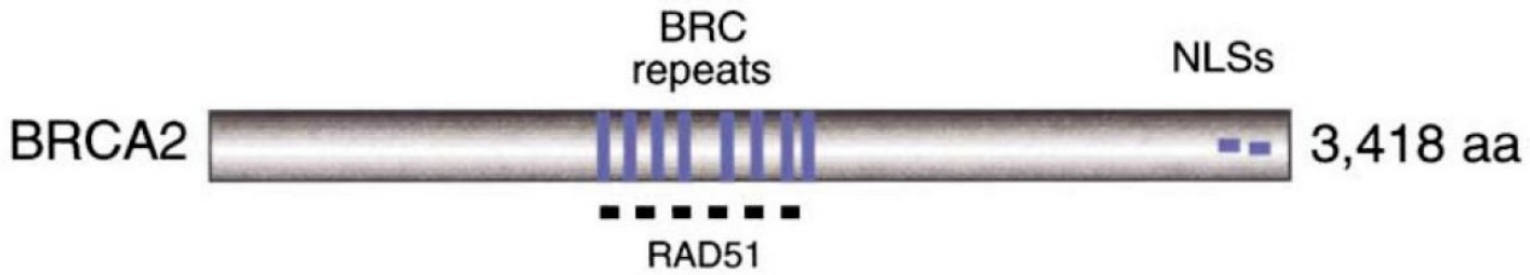
Computational Alanine Scan



Mean absolute error (kcal/mol) averaged over 5 snapshots:

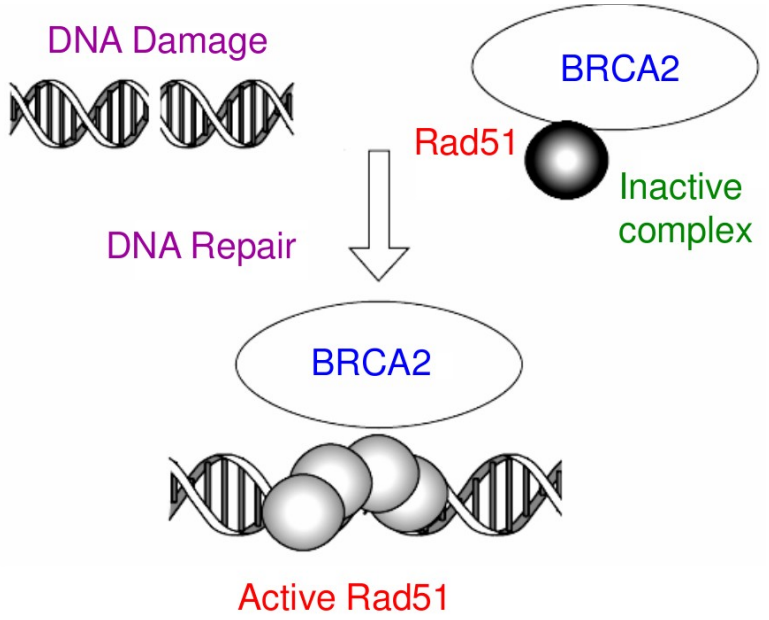
	Vacuum	G09 Solvent
MM-PBSA	1.0	3.9
QM-PBSA	0.2	0.8
QM-PBSA (npb=1)	--	1.3

BRCA2 Protein

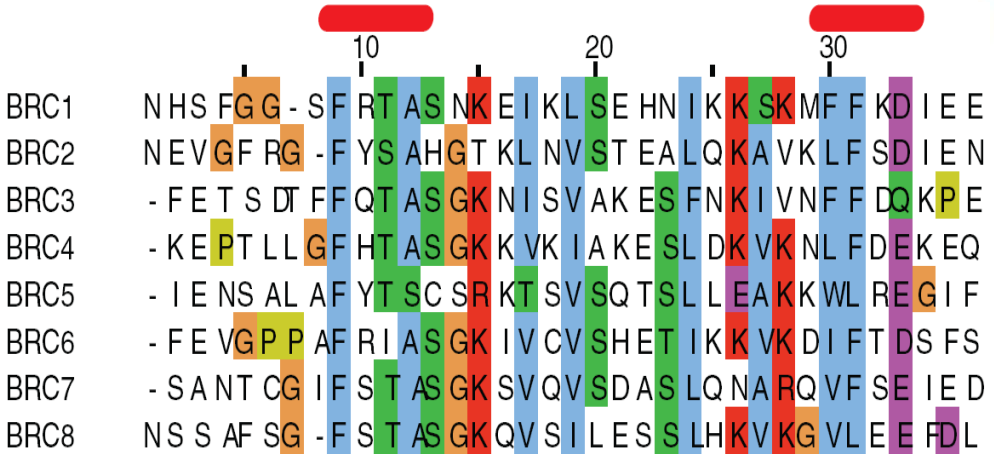


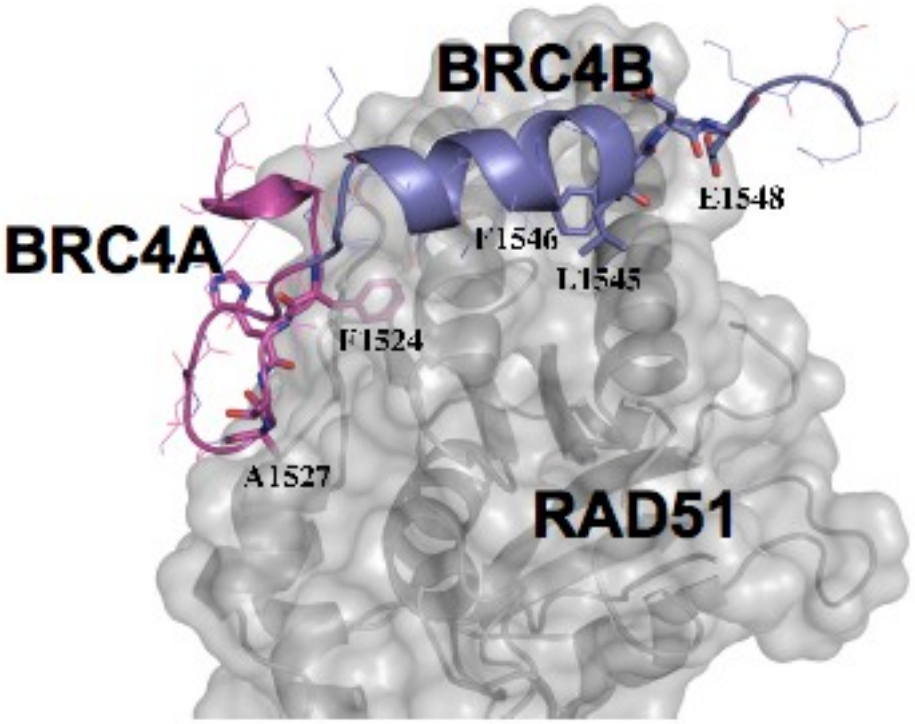
Contains 8 BRC repeats of about 35-40 amino acids: BRC1, BRC2,..., BRC8

Binds to the DNA recombination and repair protein RAD51 during DNA repair by Homologous Recombination:



Two Hotspot Model





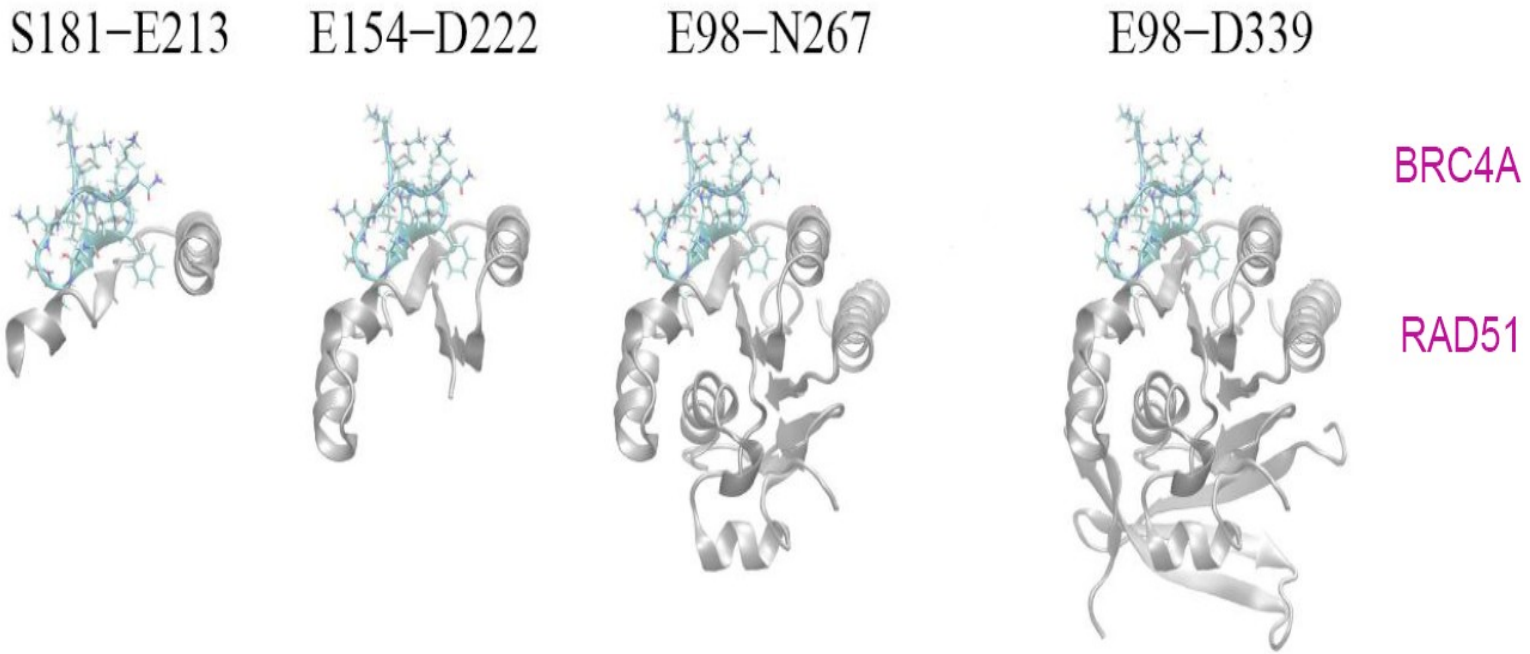
The crystal structure for the RAD51-BRC4 complex has been solved.

Sequence alignment and experimental ELISA assays suggest two hotspot interaction model.

Aim to measure computationally the difference in binding affinities between RAD51-BRC4A and RAD51-BRC4B.

L. Pellegrini, A Venkitaraman et al., Nature ***420****, 287 (2002)*
E. Rajendra and A. Venkitaraman, Nuc. Acids Res. ***38****, 82 (2010)*

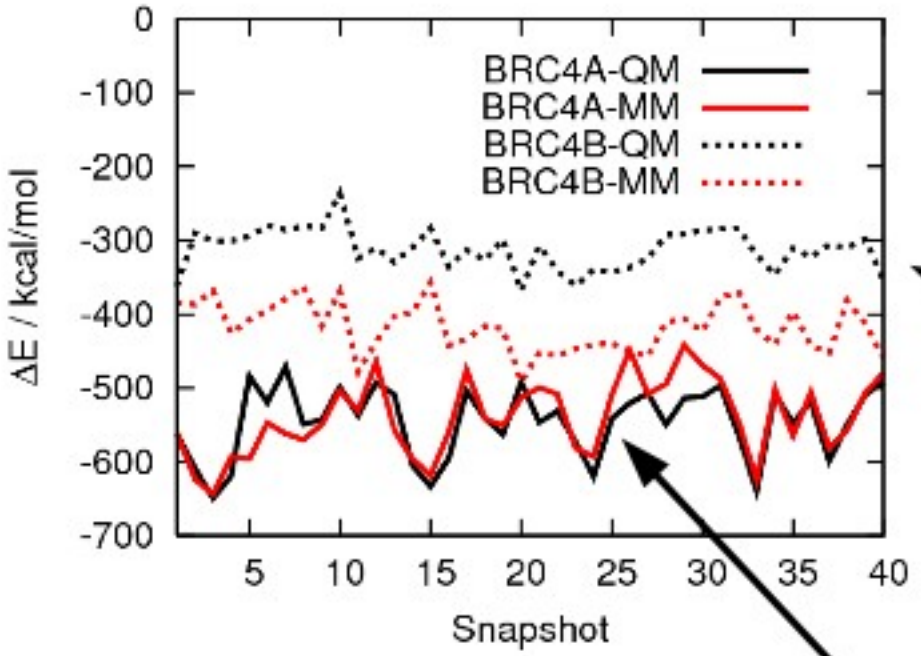
Convergence of ΔG with RAD51 size



Residues	Natoms	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

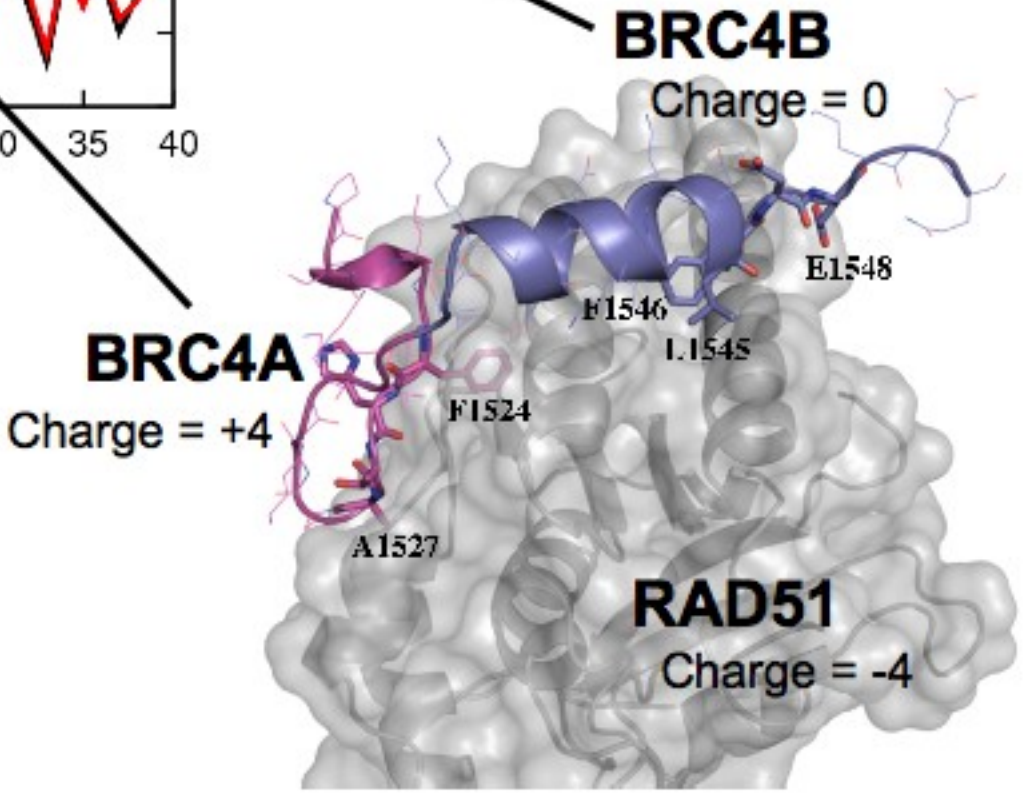
We use E98-N267 as a compromise between accuracy and expense.
$\Delta\Delta G_{QM}$ converges to within 1 kcal/mol of the full complex.

Gas Phase Results



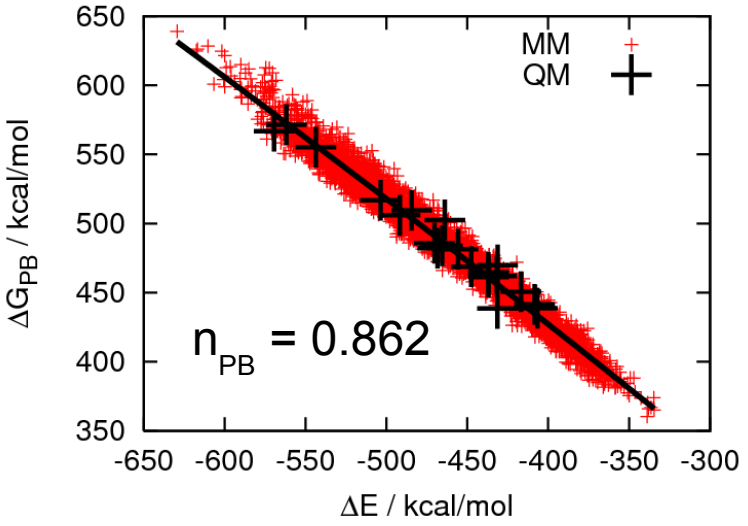
Good correlation between individual QM and MM snapshots.

MM energies for RAD51-BRC4A are very good but are over-estimated for RAD51-BRC4B.

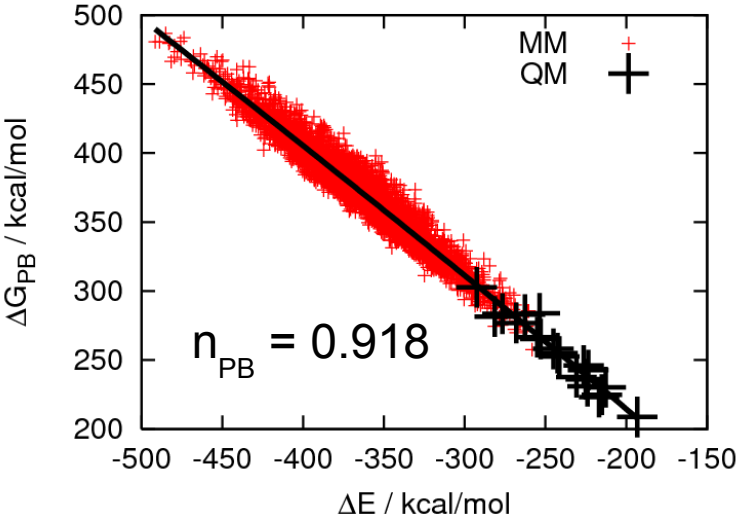


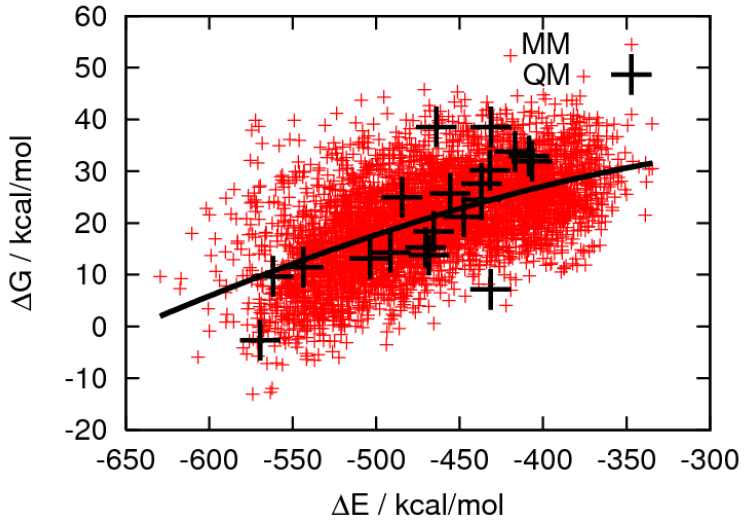
Calculation of $\Delta G_{PB,QM}$

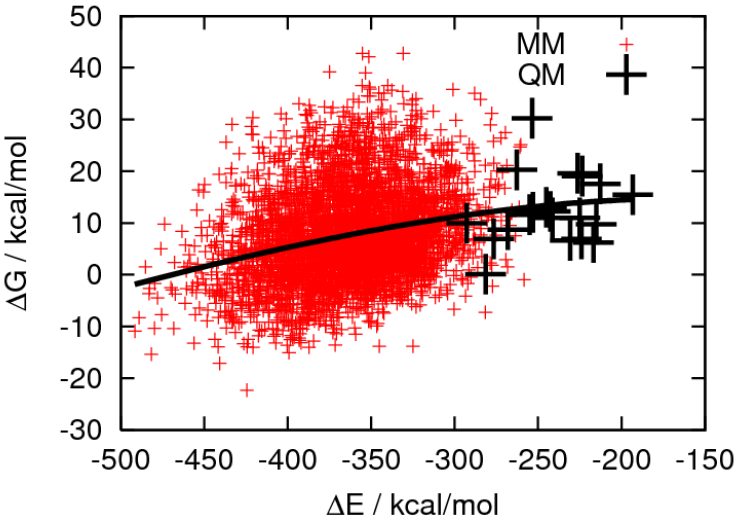
BRC4A



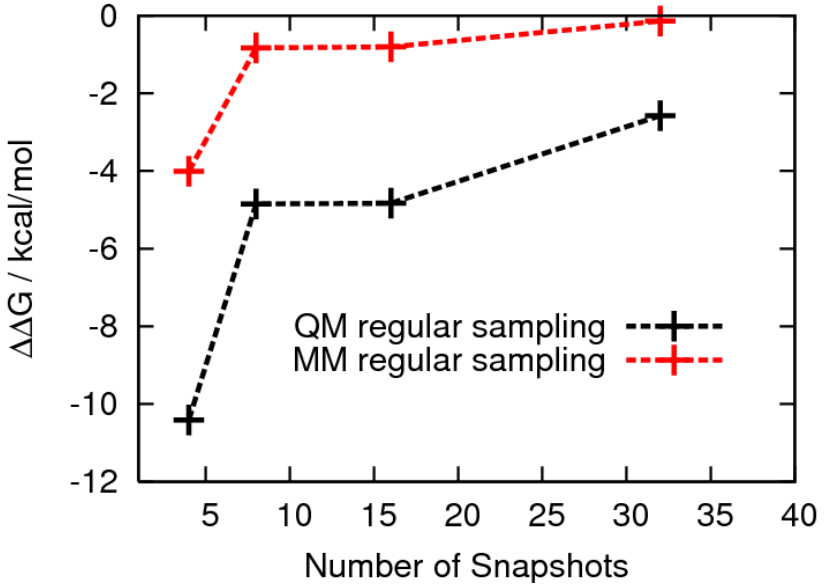
BRC4B

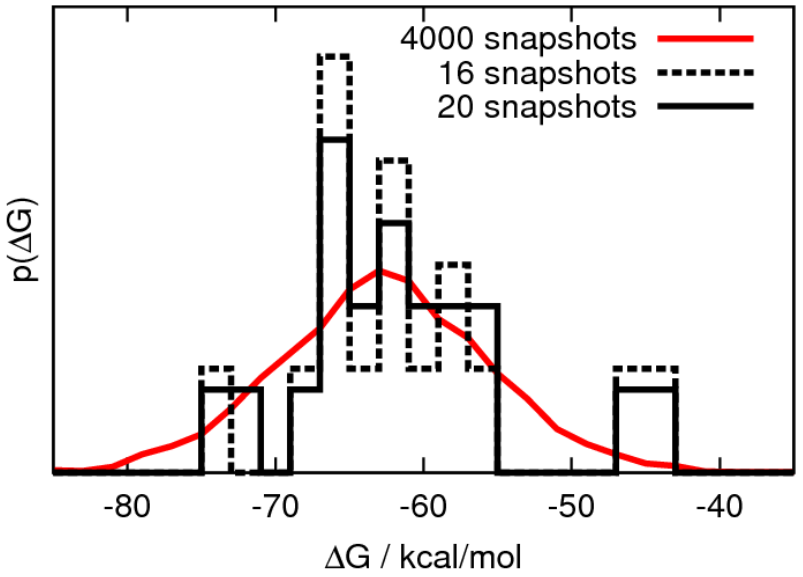






QM-PBSA Convergence with Snapshots



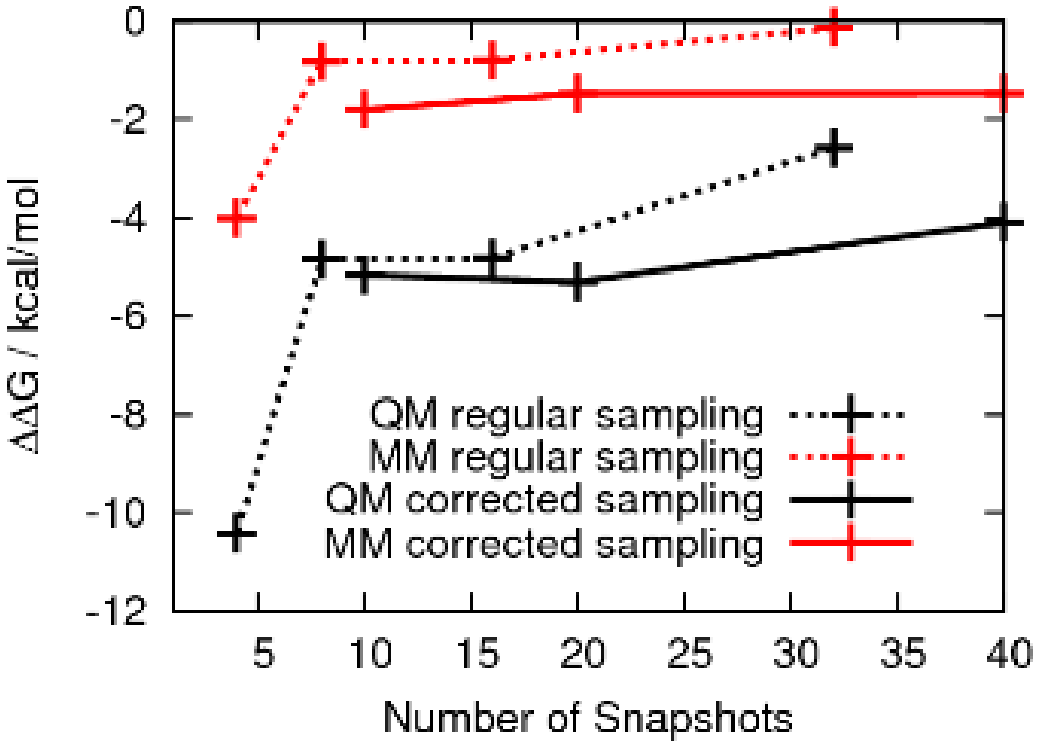


With evenly spaced sampling of the trajectory, random error in ΔG decreases only as √N.

However, there is strong correlation between MM and QM ΔGs, which we can use to select a sample that has the same properties as the full MM distribution of snapshots.

We choose mean and standard deviation of ΔG, and the occupancies of two intermittent hydrogen bonds.

QM-PBSA Convergence with Snapshots



Using corrected sampling, ΔGs are converged to within 1 kcal/mol with respect to number of snapshots:

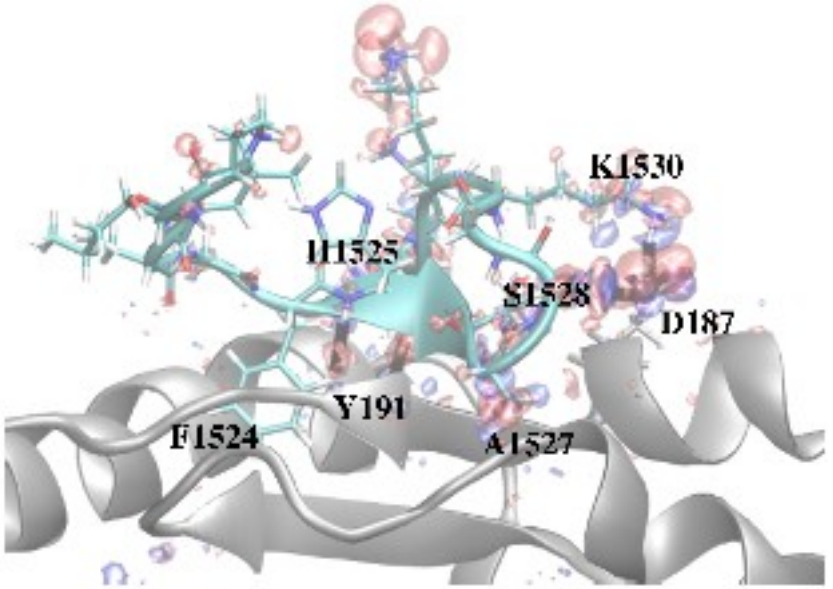
$\rightarrow \Delta\Delta G_{QM}$ = -4.1 kcal/mol

For these systems, QM-PBSA is in close agreement with MM:

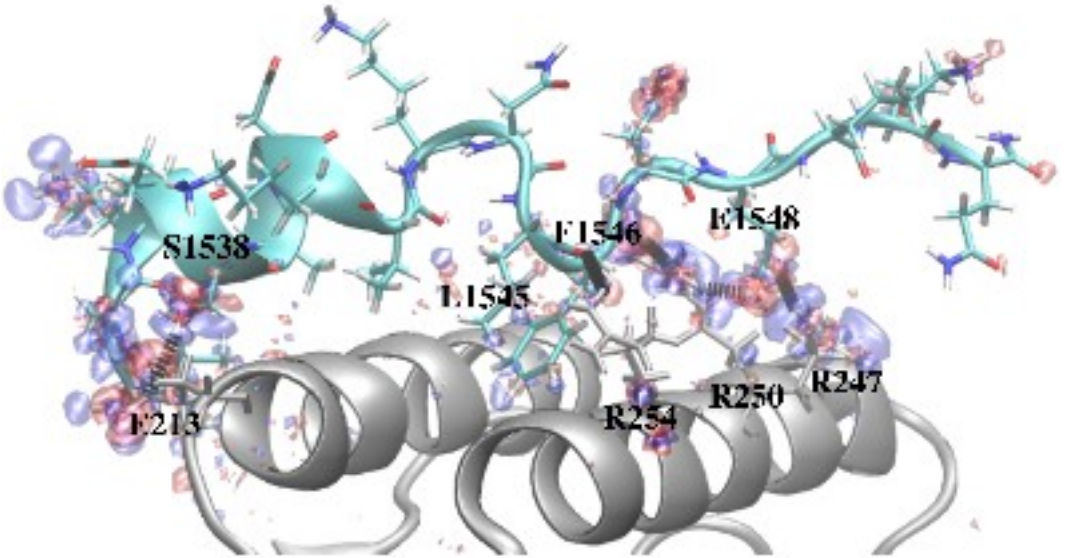
$\rightarrow \Delta\Delta G_{MM}$ = -1.7 kcal/mol

Charge Transfer and Forces

BRC4A



BRC4B



Discrepancy between QM and MM ΔE in BRC4B may be due to charge transfer and polarisation not accounted for in the MM force field.

The charge transfer is delocalised and particularly significant around the E1548 glutamate of BRC4B.

Average absolute MM forces are 18 kcal/mol/Å in good agreement with ONETEP (26 kcal/mol/Å).

Summary of QM-PBSA

Sample a classical trajectory of a system of 1000s of atoms using linear-scaling DFT to improve the accuracy of relative binding affinities.

Still reliant on an accurate classical trajectory.

Currently use scaled solvation model and empirical dispersion corrections (this can be improved in the near future).

We see a large improvement in results over MM gas phase energies and have evidence that QM-PBSA is already better for ΔΔGs, such as in alanine scans.

First principles techniques are fully transferable (eg. small molecule ligands, systems containing transition metals).

All of the electronic structure information (eg. optical absorption, electron transport) is readily available.

see also talk by Stephen Fox

Acknowledgements

Cambridge Molecular Therapeutics Programme

Chris-Kriton Skylaris (University of Southampton, U.K.)
Eeson Rajendra (MRC Research Centre, Cambridge, U.K.)
Ashok Venkitaraman (MRC Research Centre, Cambridge, U.K.)
Mike Payne (University of Cambridge, U.K.)

Cambridge Molecular Therapeutics Programme

QM-PBSA Sampling
William Belfield
(University of Cambridge, U.K.)

Cut-Off Coulomb in ONETEP
Nick Hine
(Imperial College London, U.K.)

Computing
Cambridge HPC Service (Darwin)
Southampton HPC (Iridis 3)

EPSRC