Predicting protein-ligand binding using quantum mechanics

Iva Lukac, Andrew Leach, Steve St-Gallay, Fabio Zuccotto, Ian Gilbert, Paul Wyatt

wellcome centre anti-infectives research

University of Dundee

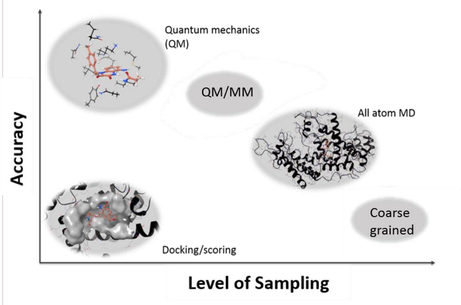
Introduction

Why Quantum Mechanics?

Classical mechanics, built upon Newton's second law ($F = ma$) accurately describes most systems that can be easily observed: movements of the objects of "normal" size, temperature and speed.

When the velocity of a particle is comparable to the speed of light, relativistic effects become important. Small light particles do not behave in a way that is consistent with the Newton's equations: we begin to think of the particle as being distributed through space like a wave, and we can calculate the probability of a particle being at a certain place at a certain time.

Atoms and molecules behave essentially as classic particles, while electrons can only be described by quantum mechanics.



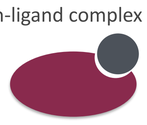
Compared to methods that use fixed atomic charges to model electrostatic interactions, QM has the advantage that it can represent charge fluctuation and dynamic polarisation. Despite significant progress, such calculations remain challenging because of their computational cost.

Xray structure checklist:

1) the protein structure is correct and known with high accuracy
2) the experimental conditions under which the crystal was obtained are relevant to the binding event
3) the ligand structure is accurate and that interactions between the binding partners are correct and well understood

Theoceptor model

protein-ligand complex



Starting from a 3D structure of the protein-ligand complex, identify key binding elements

Edit the structure; atoms are fixed only to mimic the scaffolding effect of the remainder of the protein

theoceptor

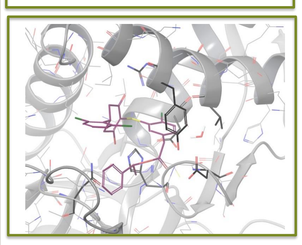
Perform QM optimisation. The result is a theoretical receptor constructed by computing the optimal geometry.

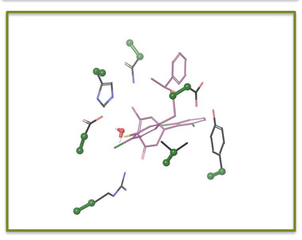
Binding affinity can be described as:

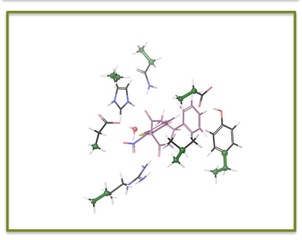
$Affinity = \alpha\Delta E + \beta logP + \gamma$

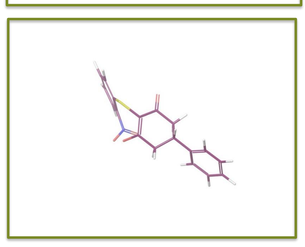
where α,β and γ are constants. ΔE is defined by:

$\Delta E = E_{complex} - E_{ligand} - E_{receptor}$









Why logP?

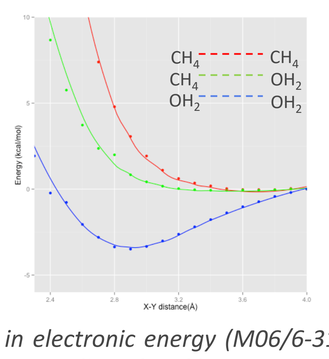
Quantum mechanical calculations generally deal with the gas phase electronic energy of the system whereas experimental observations in biological systems are related to free energies in solution:

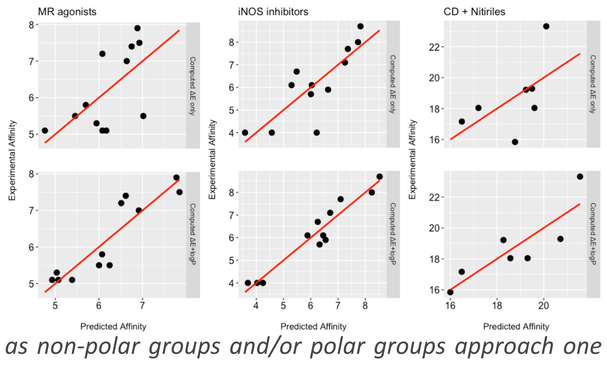
$$\Delta G_{binding}(aq) = \Delta E + \Delta H_{corr}(gas) - T\Delta S(gas) + \Delta\Delta G(solv)$$

$clogP$

$$\Delta E = E_{complex} - E_{ligand} - E_{receptor} \qquad Affinity = \alpha\Delta E + \beta logP + \gamma$$

The electronic energies (ΔE) can account for the energy of the complex but do not indicate the shape of the potential energy surface (PES), which describes how rigid the complex is.

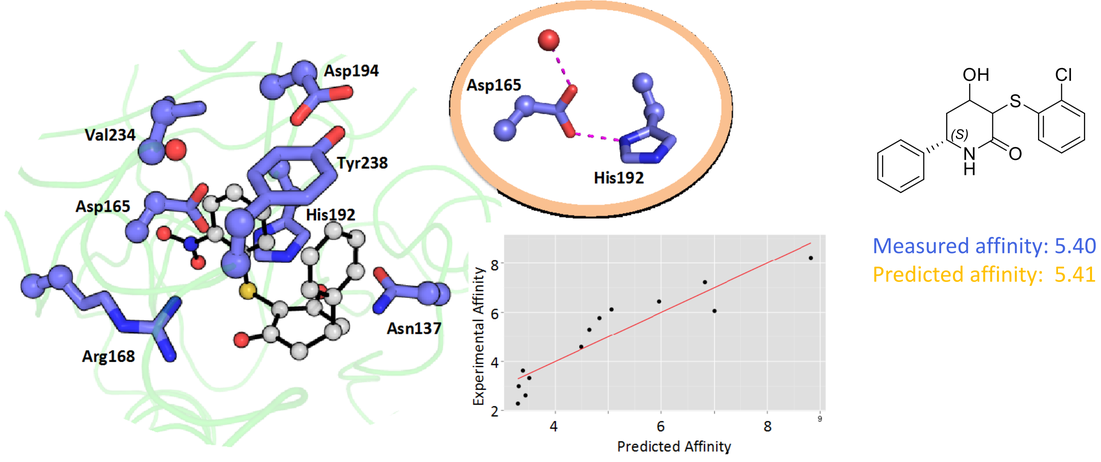




*Variation in electronic energy (M06/6-31+G**) as non-polar groups and/or polar groups approach one another. Example protein-ligand and host-guest systems studied with QM calculations.*
Top: experimental affinity is plotted against computed energies, ΔE; bottom: logP term is added

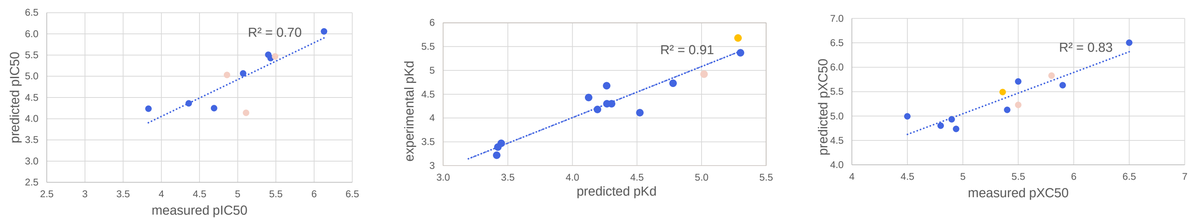
Theoceptor application

Binding affinity prediction



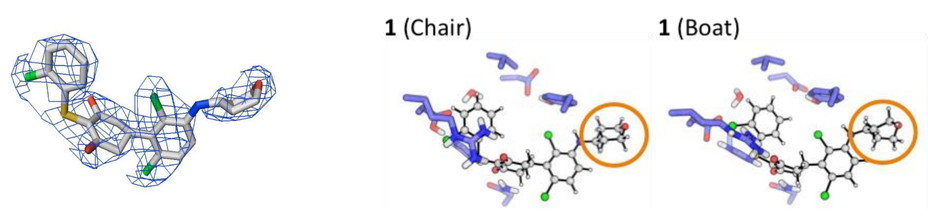
Theoceptor for the LDHA nicotinamide site. Cα, Cβ (spheres) and water oxygen (red sphere) were fixed during the QM optimization. The insert shows the hydrogen bonding network involving a water molecule, Asp165 and His192. Measured affinity is plotted against computed affinity R^2= 0.85 and RMSE = 0.76.

The approach has been successfully applied to several projects within the DDU:

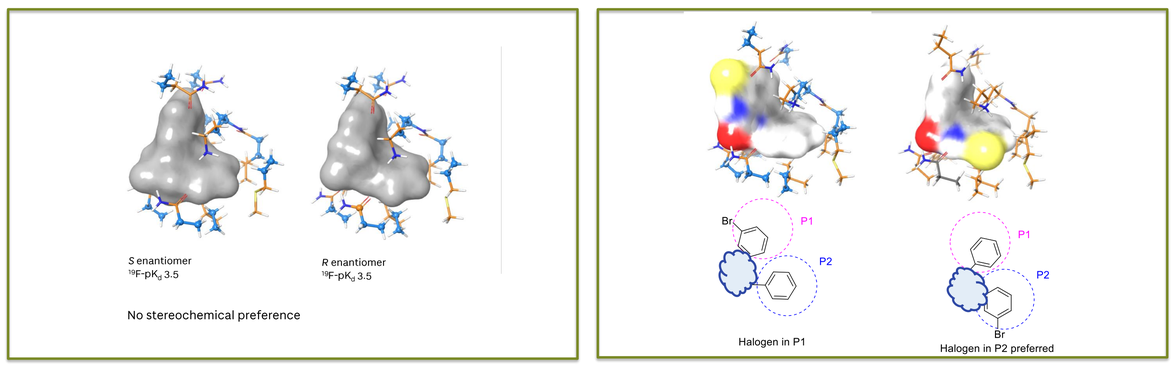


Binding mode studies

QM calculations can help discriminate amongst plausible interpretations of the observed electron density:



The chemical rigour demanded by the theoceptor calculations can improve: ligand conformations, interpretation of the positioning of heteroatoms, stereoselectivity, positioning of ligands where electron density is missing, assessment of tautomeric and ionisation states.



Both enantiomers of the fragment were optimised in the context of the receptor. The energy difference between the two flipped complexes was found to be ~0.4 kcal/mol and suggests that both binding modes can exist. This is supported with experimental binding affinity measurements.

In the example of the fragment on the right, two different binding modes were examined. There was a clear preference for a binding mode in which the halogen is placed in the P2 pocket, rather than P1, which was later confirmed by crystallography.

Conclusions

The combination of computed QM binding energies with measured or predicted log *P* values can provide usefully accurate predictions of protein-ligand binding energies.
By assigning relative energies to protein-ligand structures in which different conformations, tautomers, protonation states and stereoisomers are sampled, useful insights about the protein and ligand are generated. This can include regions that are not observed in the experimental density.

Acknowledgements: Richard A. Ward and Hend Abdelhakim

Our vision: To help tackle the urgent unmet medical need and lack of drug discovery for Neglected Tropical Diseases