Predicting protein-ligand binding using quantum mechanics

Iva Lukac
UKQSAR, Sep 2019
Quantum Mechanics in ligand optimisation

ΔG_{binding(aq)} = ΔE + ΔH_{corr(gas)} - TΔS(gas) + ΔG(solv)

ΔE = E_{complex} - E_{ligand} - E_{receptor}

Affinity = αΔE + β\log P + γ

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
Development of the QM model

Starting from the 3D structure of the protein and ligand, identify the key binding elements of the remainder of the protein.

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

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

\[
\Delta E = E_{\text{complex}} - E_{\text{ligand}} - E_{\text{receptor}}
\]

**Affinity** = \(a\Delta E + b\log P + \gamma\)

Lactate Dehydrogenase A (LDHA) theoceptor

- Interpretation of heteroatom positioning
- Assessing stereochemistry
- Addressing missing density
- Binding mode studies
- Binding affinity prediction

- 1 (Chair)
- 1 (Boat)
- 2 (Pseudo-equatorial)
- 2 (Pseudo-axial)

Lactate Dehydrogenase A (LDHA)
The approach has been successfully applied to several projects within the DDU

- Binding affinity prediction
  - Requirement: at least three compounds of known affinity are required - only relative affinity can be computed
  - A range of affinity assessment types (IC\textsubscript{50}, K\textsubscript{i}, K\textsubscript{D} etc) can be used

Theoceptors in practice

![Graph 1](image1)

- Predicted pXC\textsubscript{50} vs. measured pXC\textsubscript{50}
- $R^2 = 0.83$

![Graph 2](image2)

- Predicted pIC\textsubscript{50} vs. measured pIC\textsubscript{50}
- $R^2 = 0.93$

![Graph 3](image3)

- Predicted pIC\textsubscript{50} vs. experimental pIC\textsubscript{50}
- $R^2 = 0.91$

![Graph 4](image4)

- Predicted pK\textsubscript{D} vs. experimental pK\textsubscript{D}
- $R^2 = 0.70$
The approach has been successfully applied to several projects within the DDU

- Binding affinity prediction
  - Requirement: at least three compounds of known affinity are required - only relative affinity can be computed
  - A range of affinity assessment types (IC$_{50}$, K$_i$, K_D etc) can be used
- Binding mode studies
- Tautomeric and protonation state preferences as well as stereoselectivity
- Help resolve ambiguities in ligand refinement

No stereochemical preference
Fragment Molecular Orbital Method (FMO)

PIEDA (Pair interaction energy decomposition analysis)

Electrostatic

Charge transfer

Exchange repulsion

Dispersion

Coulomb “bath” of the full system

Converged Coulomb “bath” of the full system

FMO enables QM calculations to be performed and interpreted in a way that can drive SBDD


Assessing water molecules using FMO

Targeting crystallographically observed water molecules for displacement or specific interaction to improve ligand affinity has proven successful in drug discovery (increased potency, selectivity, better binding kinetics, etc.)

Many programs exit to analyse water locations and their energetics

Evaluating solvent energetic contribution is the most challenging aspect:

- Overall, the placement of water molecules was fairly accurate, with all programs predicting 60–90% of water oxygens within 1 Å of their observed locations.
- A major problem for all tools was a consistent prediction of energetic contributions of water molecules. The tools seldom agreed with each other and also the consistency within each tool was very low.

Poor correlation of the predicted water energies with the experimentally observed SAR

The quality of the prediction is indirectly read out as a relative affinity of a new ligand that displaces water molecule, and/or perturbs the existing water network

Can FMO help with the interpretation?

<table>
<thead>
<tr>
<th>Number</th>
<th>pIC&lt;sub&gt;50&lt;/sub&gt;</th>
<th>Total PIE</th>
<th>Compound Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.88</td>
<td>-70.7 kcal/mol</td>
<td><img src="image" alt="Structure 1" /></td>
</tr>
<tr>
<td>2</td>
<td>6.63</td>
<td>-95.6 kcal/mol</td>
<td><img src="image" alt="Structure 2" /></td>
</tr>
<tr>
<td>4</td>
<td>6.79</td>
<td>-95.6 kcal/mol</td>
<td><img src="image" alt="Structure 4" /></td>
</tr>
<tr>
<td>5</td>
<td>5.85</td>
<td>-95.7 kcal/mol</td>
<td><img src="image" alt="Structure 5" /></td>
</tr>
</tbody>
</table>
Expansion of the ligand leaves the water network unperturbed

Hydrophobic tails of 4 and 5 lead to a slight shift of Phe45, Tyr106 and Ile113

PIE correlates with the SAR going from 1 to 2 to 4

4 and 5 energetically equivalent (≠ #rot bonds, entropic penalty?)

The energies of water molecules should remain fairly similar for all BRD9 structures
  • The energies are within ~5 kcal/mol
Ligand extends into the water network, displacing water molecules and causing perturbations

PIE does not explain the observed SAR

Energetic contribution from perturbed water network
Overall energy change of the perturbed water network is unfavourable

Average ‘happiness’ of the water network gives better interpretation of the observed SAR
TAF1(2)

- Shift of water #0
- Displacement of waters #3 and #4

Total PIE -118.5 kcal/mol

pIC$_{50}$=7.22

Total PIE -127.7 kcal/mol

Shift of water #0
Displacement of waters #3 and #4

Total PIE -128.5 kcal/mol

5
pIC$_{50}$=7.33

2
pIC$_{50}$=7.22

4
pIC$_{50}$=6.38

Total PIE -118.5 kcal/mol
TAF1(2)

- PIE does not explain the observed SAR
- Energetic contribution from perturbed water network

**Shift of waters #0 and #3**

**Shift of water #0**

**Displacement of waters #3 and #4**

**Total PIE**

-118.5 kcal/mol

**pIC\textsubscript{50} = 7.22**

**Displacement of water #4**

**Total PIE**

-127.7 kcal/mol

**pIC\textsubscript{50} = 6.38**

**Total PIE**

-128.5 kcal/mol

**pIC\textsubscript{50} = 7.33**

**HOH0**

**HOH1**

**HOH2**

**HOH3**

**HOH4**

**Drug Discovery Unit**
2-4: Water network broken, decrease in potency

2-5: Re-establishing the water network
- HOH2 and HOH3 still ‘happy’
- HOH0 less ‘happy’

4-5: Alters the conformation of the ligand, disrupts the network, increase in potency
Correctly quantifying and describing interactions between protein and the ligand computationally requires a physically accurate description of the molecules.

Binding affinity prediction
- The combination of computed QM binding energies with measured or predicted logP values can provide usefully accurate predictions of protein-ligand binding energies.

Rationalising SAR
- Going beyond ‘visual inspection’ and identifying nonintuitive interactions.

Binding mode studies
- Predicting the correct binding pose
- Tautomeric and protonation state preferences as well as stereochemistry
- Help resolve ambiguities in ligand refinement.

Assessing water energetics

Current work:
- How to deal with the systems where there is a movement in the protein backbone?
- Throughput is getting better, but slowly
  - Scripts that enable quicker setup and analysis of the results
  - How different settings affect accuracy
  - Focus on group efficiency
- How to best communicate the results to the project?
Acknowledgments

Fabio Zuccotto
Juan Bueren-Calabuig
Michael Kiczun
Judith Hobrath
Carmine Morgillo
Riccardo Guareschi
David Robinson
Alain-Pierre Petit

DDU Management
Paul Wyatt
Kevin Read
Ian Gilbert
David Gray
Julie Brady

Andrew Leach
Steve St-Gallay
Richard Ward

Thank you for your attention...
Questions?