Protein-Ligand Interactions

Validation on a model Host-Guest System and Preliminary Results on T4 Lysozyme

Stephen Fox
School of Chemistry, University of Southampton
14th April 2010
Outline of talk

- **Validation on a model host-guest system**
  - The “Tennis Ball” dimer
  - Static structures and binding energies
  - MM-PBSA and QM-PBSA
  - Dispersion in ONETEP

- **Applications to the T4 lysozyme protein**
  - Hydrophobic cavity L99A mutant
  - Ligand classes
  - Further information from electronic density

- **Next Steps**
Validation on a model host-guest system

- The “Tennis Ball” dimer
- Static structures and binding energies
- MM-PBSA and QM-PBSA
- Dispersion in oneTEP

Applications to the T4 lysozyme protein

- Hydrophobic cavity L99A mutant
- Ligand classes
- Further information from electronic density

Next Steps
The "Tennis Ball" Dimer: Model of a Host-Guest System
Parameters for the "Tennis Ball" Dimer

ONETEP

- Kinetic energy cut-off 800eV
- XC functional: PBE
- Geometry optimisations: force threshold $0.003 \, E_h/a_0$
- Dispersion 1

AMBER

The generalised AMBER forcefield (gaff) in the AMBER10 program was used. Special parameters were taken from the paper by Fox et al. A DRMS value of 0.1 kcal/mol/Å was used for the geometry optimisation to correspond to the ONETEP force threshold.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>AMBER structures</th>
<th>ONETEP structures</th>
<th>Experimental $\Delta H$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AMBER</td>
<td>ONETEP</td>
<td>AMBER</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>-9.8</td>
<td>-10.3</td>
<td>-9.9</td>
</tr>
<tr>
<td>CHCl$_3$</td>
<td>-16.9</td>
<td>-2.8</td>
<td>-17.0</td>
</tr>
<tr>
<td>CF$_4$</td>
<td>-15.4</td>
<td>-8.4</td>
<td>-14.8</td>
</tr>
</tbody>
</table>

- **AMBER** binding energies vary little between structures
- **ONETEP** binding energies on ONETEP structures much more accurate (compared to experimental $\Delta H$)
MM-PBSA is a method for calculation of the free energy of binding.
Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA)

\[
\Delta G_4 = \Delta H + (\Delta G_3 - \Delta G_1 - \Delta G_2)
\]  

(1)

\[
\Delta G_4 = \Delta H + \Delta G_{sol}
\]  

(2)

\[
= \Delta E_{EL} + \Delta E_{VdW} + \Delta E_{Int} + \Delta G_{PB} + \Delta G_{SA}
\]  

(3)
Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA)

\[
\Delta G_{tot} = \langle \Delta H \rangle + \langle \Delta G_{sol} \rangle - T \langle \Delta S_{conf} \rangle
\]
In QM-PBSA the $\langle \Delta H_{MM} \rangle$ term is substituted by $\langle \Delta H_{QM} \rangle$ calculated with ONETEP. $\langle \Delta G^{QM}_{sol} \rangle$ is then calculated by scaling the electrostatic part of $\langle \Delta G^{MM}_{sol} \rangle$ thus,

$$\Delta G_{sol}^{QM} = \Delta G_{PB} \left( \frac{\Delta E_{DFT}}{\Delta E_{EL}} \right) + \Delta G_{SA},$$  \hspace{1cm} (5)

where

$$\Delta H_{QM} = \Delta E_{DFT} + \Delta E_{disp}.$$  \hspace{1cm} (6)

So

$$\Delta G_{QM} = \langle \Delta H_{QM} \rangle + \langle \Delta G^{QM}_{sol} \rangle.$$  \hspace{1cm} (7)
Differences between the MM and QM Optimised Structures

<table>
<thead>
<tr>
<th>Bond Lengths (Å)</th>
<th>CH₄</th>
<th>CHCl₃</th>
<th>CF₄</th>
<th>Empty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter (internal)</td>
<td>-0.78</td>
<td>-0.32</td>
<td>0.00</td>
<td>0.25</td>
</tr>
<tr>
<td>Diameter (between monomers)</td>
<td>0.59</td>
<td>0.15</td>
<td>0.13</td>
<td>-0.54</td>
</tr>
<tr>
<td>H-bond</td>
<td>-0.17</td>
<td>-0.14</td>
<td>-0.23</td>
<td>0.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond</th>
<th>CH₄</th>
<th>CHCl₃</th>
<th>CF₄</th>
<th>Empty</th>
</tr>
</thead>
<tbody>
<tr>
<td>C=O</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>N-H</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>C-H</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C-C</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
<tr>
<td>C-N</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
</tbody>
</table>
### MM-PBSA and QM-PBSA Results on the “Tennis Ball” Dimer

<table>
<thead>
<tr>
<th>Ligands</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>AMBER</th>
<th>ONETEP</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_4$</td>
<td>$-8.7 \pm 0.2$</td>
<td>$-9.0 \pm 0.2$</td>
<td>-9</td>
<td></td>
</tr>
<tr>
<td>CHCl$_3$</td>
<td>$-16.8 \pm 1.0$</td>
<td>$-7.1 \pm 1.1$</td>
<td>-7</td>
<td></td>
</tr>
<tr>
<td>CF$_4$</td>
<td>$-12.6 \pm 0.3$</td>
<td>$-6.0 \pm 0.4$</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\Delta\Delta H$ (kcal/mol)</th>
<th>AMBER</th>
<th>ONETEP</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_4$ $\rightarrow$ CHCl$_3$</td>
<td>$-8.2 \pm 1.2$</td>
<td>$2.0 \pm 1.3$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CH$_4$ $\rightarrow$ CF$_4$</td>
<td>$-4.0 \pm 0.5$</td>
<td>$3.0 \pm 0.6$</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
Importance of Dispersion Forces in the "Tennis Ball"

Comparison of PBE binding energies with and without the dispersion correction.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta H_{vac}$ kcal/mol</th>
<th>Without Dispersion</th>
<th>With Dispersion</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_4$</td>
<td>1.4</td>
<td>-9.0</td>
<td>-9</td>
<td>-9</td>
</tr>
<tr>
<td>CHCl$_3$</td>
<td>14.1</td>
<td>-7.1</td>
<td>-7</td>
<td></td>
</tr>
</tbody>
</table>

Dispersion interactions are very important in biological systems.
Why is dispersion important?

- Weak interactions
  - Hydrogen bonding
  - Van der Waals interactions

- Strong interactions
  - Ionic bonds
  - Covalent bonds

Weak interactions are common in biological systems and very important to correctly describe.

Common DFT functionals can not fully account for dispersion interactions (no long range attractive forces).
Empirical Correction for Including Dispersion in onETEP

\[ E[n] = T_s[n] + \int n(r)\nu(r)d\mathbf{r} + J[n] + E_{xc}[n] - \frac{i}{j} \sum_{i,j,i \neq j} f_{damp}(r_{ij}) \frac{C_{6,ij}}{r_{ij}^6} \]

- \( C_{6,ij} = \frac{3}{2} \alpha'_i \alpha'_j \frac{I_i I_j}{I_i + I_j} \)
  - \( I_n \) is the ionisation potential
  - \( \alpha'_n \) is the polarizability and \( r_{ij} \) is the distance between the atoms
- \( f_{damp} \) is a damping function (ranges from 0 at small \( r_{ij} \) values to 1 at large \( r_{ij} \) values)

Theory,

Empirical Correction for Including Dispersion in ONETEP

\[ f_{damp}(r_{ij}) \]


Parameters optimised for: H, C, N, O, S (for all GGAs) against CCSD(T) and MP2 binding energies for 60 complexes.

- Validation on a model host-guest system
  - The “Tennis Ball” dimer
  - Static structures and binding energies
  - MM-PBSA and QM-PBSA
  - Dispersion in onetep

- Applications to the T4 lysozme protein
  - Hydrophobic cavity L99A mutant
  - Ligand classes
  - Further information from electronic density

- Next Steps
T4 Lysozyme

- Lysozymes: enzymes that act as a natural form of protection from pathogens. Attacks the peptidoglycans found in the cell walls of bacteria. When the Lysozyme binds the structure becomes strained making the glycosidic bond easy to break.
- L99A mutant
### T4 Lysozyme: Ligands for Tests

<table>
<thead>
<tr>
<th>Binder</th>
<th>Non-binder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrophobic</td>
<td></td>
</tr>
<tr>
<td>Polar</td>
<td></td>
</tr>
</tbody>
</table>

![Chemical structures of hydrophobic and polar ligands for binders and non-binders in T4 lysozyme tests](attachment:chemical_structures.png)
Parameters for T4 Lysozyme

**ONETEP**
- Kinetic energy cut-off 800eV
- XC functional PBE
- Dispersion 1
- Cut-off Coulomb

**AMBER**
AMBER10 package using ff99SB and gaff forcefields.

All calculations (**AMBER** and **ONETEP**) performed on the entire enzyme (>2600 atoms)
QM-PBSA calculations

On a polar ligand (2,3-benzofuran) and a non-polar ligand (benzene) using 80 snapshots through a 10ns MD simulation. (kcal/mol)

<table>
<thead>
<tr>
<th></th>
<th>MM-PBSA</th>
<th>QM-PBSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>$\Delta H$</td>
<td>$-16.8 \pm 0.5$</td>
</tr>
<tr>
<td>Benzofuran</td>
<td>$-23.7 \pm 0.9$</td>
<td>$-26.4 \pm 0.9$</td>
</tr>
<tr>
<td>Benzene</td>
<td>$\Delta G$</td>
<td>$-10.2 \pm 0.5$</td>
</tr>
<tr>
<td>Benzofuran</td>
<td>$-13.6 \pm 1.1$</td>
<td>$-16.6 \pm 3.5$</td>
</tr>
<tr>
<td>$\Delta \Delta G$</td>
<td>MM-PBSA</td>
<td>QM-PBSA</td>
</tr>
<tr>
<td></td>
<td>-3.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>
Density Differences on Static Structures

Benzene

Benzofuran
Density Differences through a 10ns MD Simulation

Purple: electron gain. Orange: electron depletion
Validation on a model host-guest system
- The “Tennis Ball” dimer
- Static structures and binding energies
- MM-PBSA and QM-PBSA
- Dispersion in ONETEP

Applications to the T4 lysozyme protein
- Hydrophobic cavity L99A mutant
- Ligand classes
- Further information from electronic density

Next Steps
Next Steps

- More snapshots for QM-PBSA
- Test binding free energy convergence with the number of snapshots
- More ligands
- QM solvation using the QM electronic density
- Further analysis of the density
  - Bader analysis
- Polar cavity as well as a hydrophobic cavity
- More flexible ligands
- Extend to methods for calculating the free energy of binding that are more rigorous than MM-PBSA
Conclusions

▶ “Tennis Ball” dimer
  ● Very promising results with the “Tennis Ball” dimer
  ● This is a small rigid model case, need to test on larger models and more flexible ligands

▶ T4 Lysozyme
  ● Preliminary results of the free energy of binding using QM-PBSA on the entire system (>2600 atoms)
  ● Improved results over MM-PBSA
  ● Needs to be extended to more ligands with different properties
Acknowledgements

- Southampton
  - Chris-Kriton Skylaris (Supervisor)

- Boehringer Ingelheim
  - Thomas Fox
  - Christofer Tauterman
  - Hannes Wallnoefer

- Cambridge
  - Daniel Cole

- BBSRC for CASE studentship supported by Boehringer Ingelheim