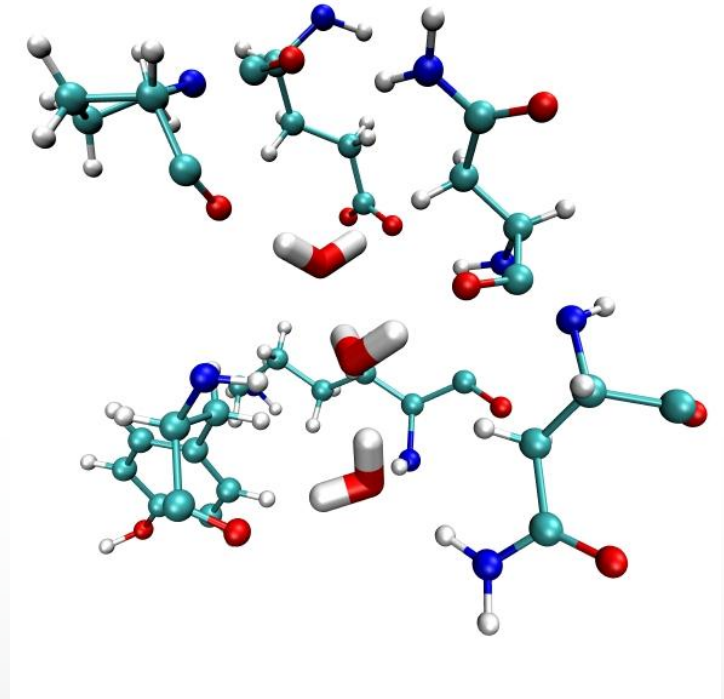


# Water and fragment binding by free energy simulations

Jonathan Essex

# Introduction

- Water molecules
  - Methods
  - Successes
  - Issues
- Fragment-Based Drug Discovery
  - Methodology
  - Factor Xa

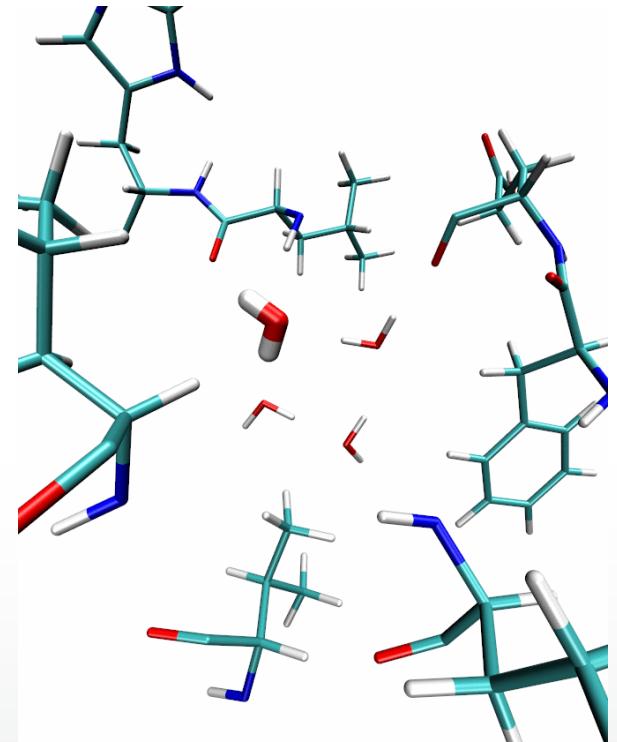


# Why worry about water?

- Important for rational drug design
  - Water displacement as a paradigm for drug development
  - Where are the waters?
  - How tightly do they bind?
  - Changes in hydration pattern with ligand substitution
  - Influence on docking and free energy calculations
- Location in x-ray structures often unclear – hydrophobic cavities

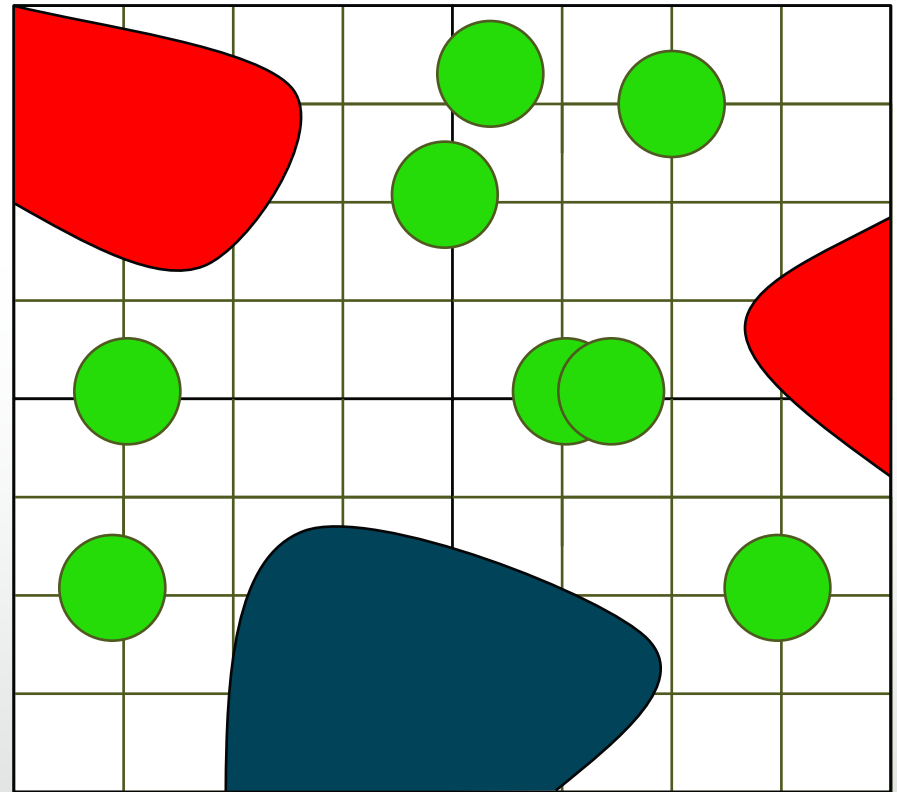
# What methods?

- Double decoupling
- JAWS
- GCMC
  
- Can we predict hydration sites?
- Can we calculate binding free energies of these waters?
- Could we use these locations and free energies to rationalise the SAR?
  
- What are the relative pros and cons of these approaches?



# Just Add Water Molecules (JAWS) - I

- 1) Define binding pocket
- 2) Place 3D grid over the pocket
- 3) Insert  $N$   $\theta$ -water molecules over the grid
- 4) Allow these molecules to move around grid and sample  $\theta$ :
  - $\theta = 1 = \text{On}$
  - $\theta = 0 = \text{Off}$
- 5) If  $\theta_i > 0.95$ , then increase a counter on nearest grid point
- 6) Collect statistics on the grid points



# Just Add Water Molecules (JAWS) - II

- 1) Restrain a  $\theta$ -water at each favourable hydration site
- 2) Apply a biasing potential to the potential energy function of each molecule and collect statistics in a MC run:

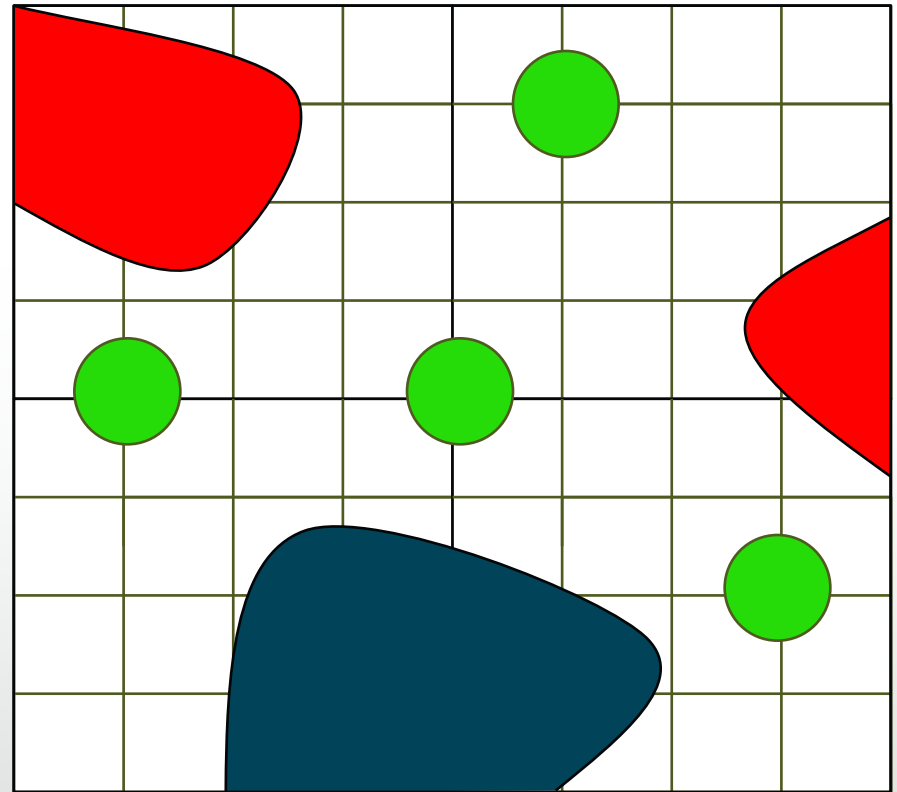
$$V(\text{bias}, i) = (-\Delta G_{\text{hyd}} - kT \ln(V^c/V^o))\theta_i$$

- Biasing term corrects for the desolvation penalty

- 3) Calculate the binding free energy as:

$$\Delta G = -kT \ln(P^{\text{on}}/P^{\text{off}})$$

- 4) Retain molecules with favourable binding free energies



# Grand Canonical Monte Carlo (GCMC) - I

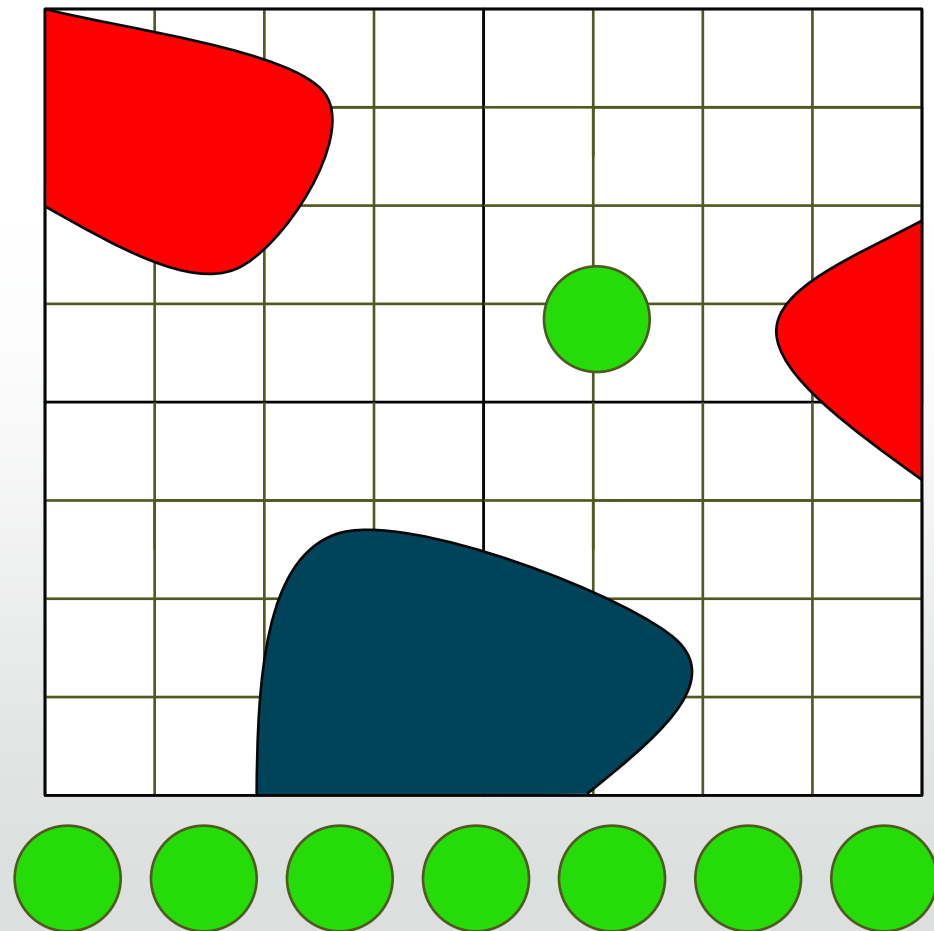
*J. Am. Chem. Soc.* **1996**, *118*, 8493

- 1) Define binding pocket
- 2) Place 3D grid over the pocket
- 3) Start simulation at a set value of  $B$

$$B = \frac{\mu'}{kT} + \ln\langle\bar{n}\rangle$$

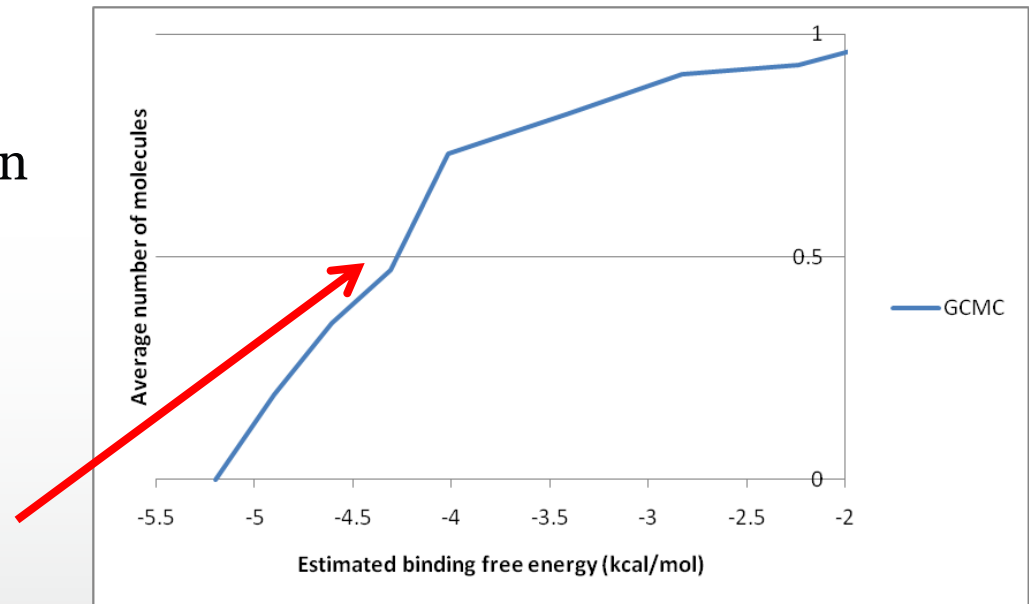
- 4) Three different types of move:

- a) Insertion
- b) Deletion
- c) Translation



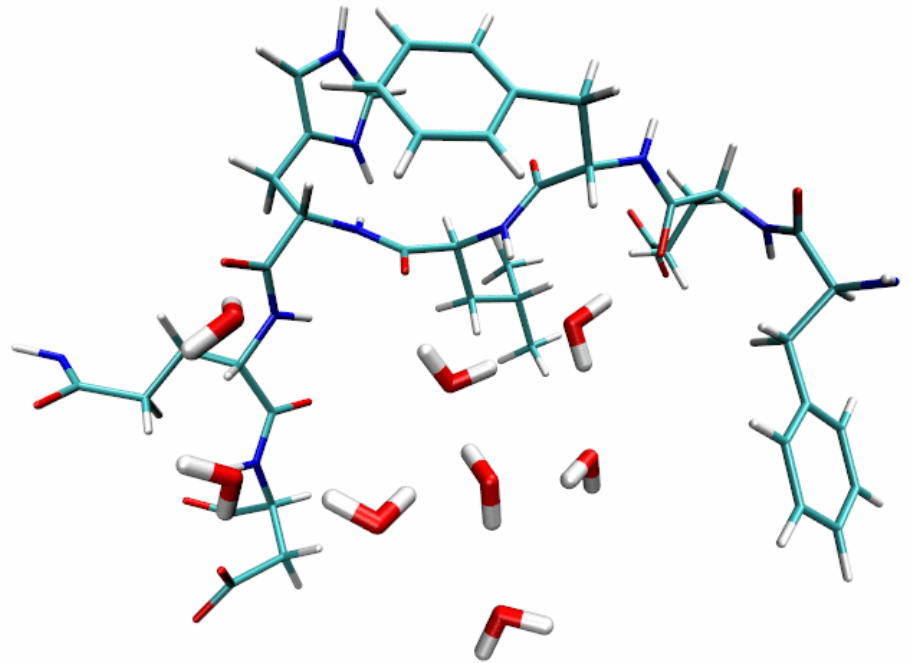
# Grand Canonical Monte Carlo (GCMC) - II

- Simulated annealing for  $\sim 5M$  MC moves at each B level
- Continue until  $\langle N \rangle$  is less than 0.10
- Read free energy from this value
  - Free energy 'titration'
- Or convert into concentration and compare to bulk concentration

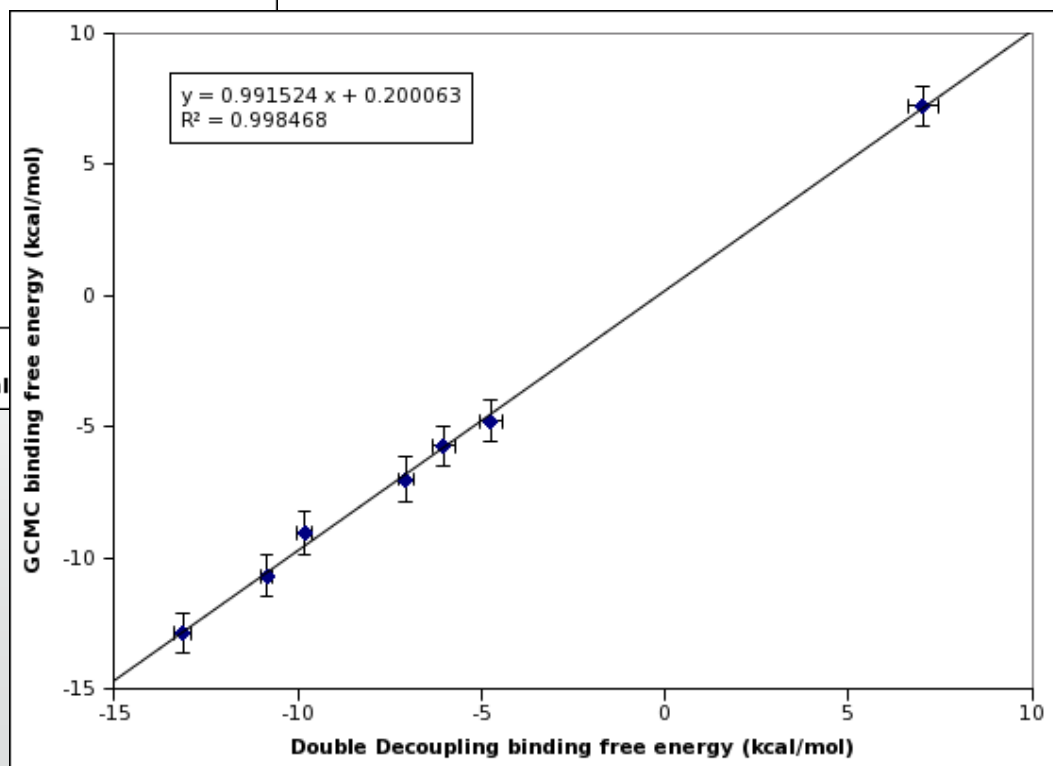
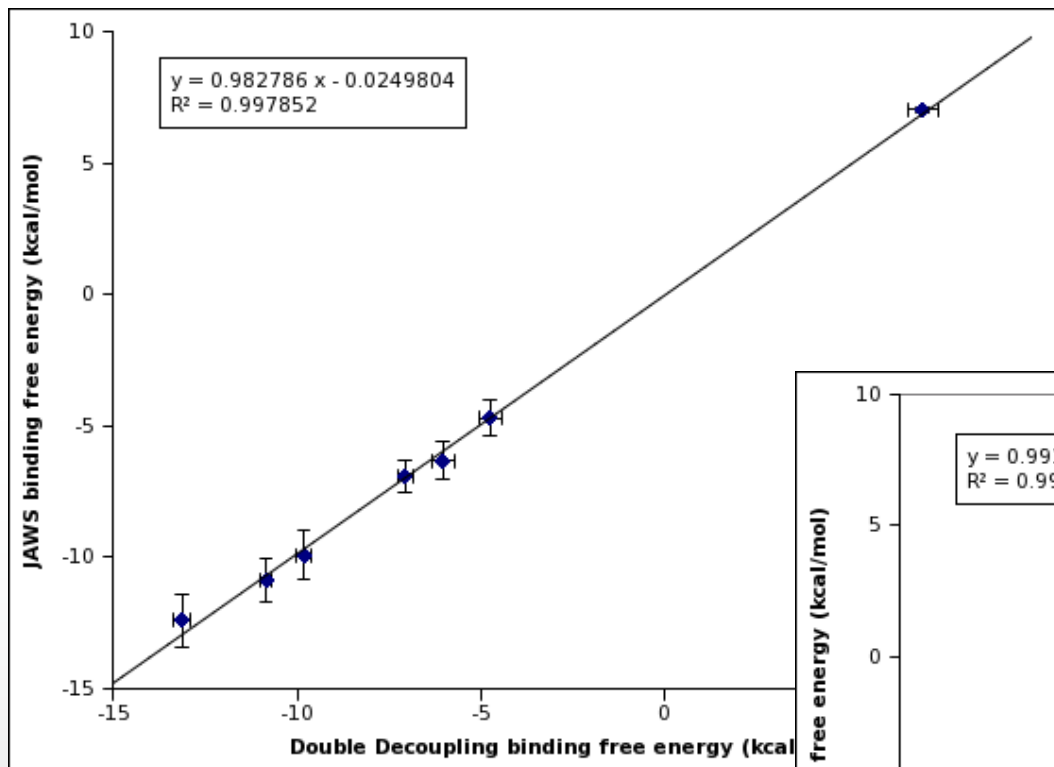


# Grand Canonical Monte Carlo (GCMC) - III

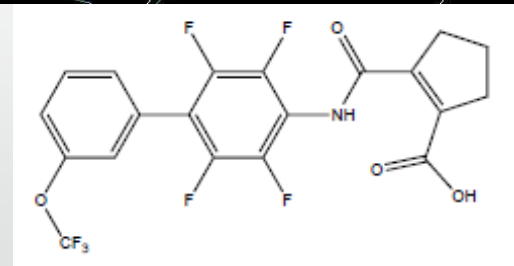
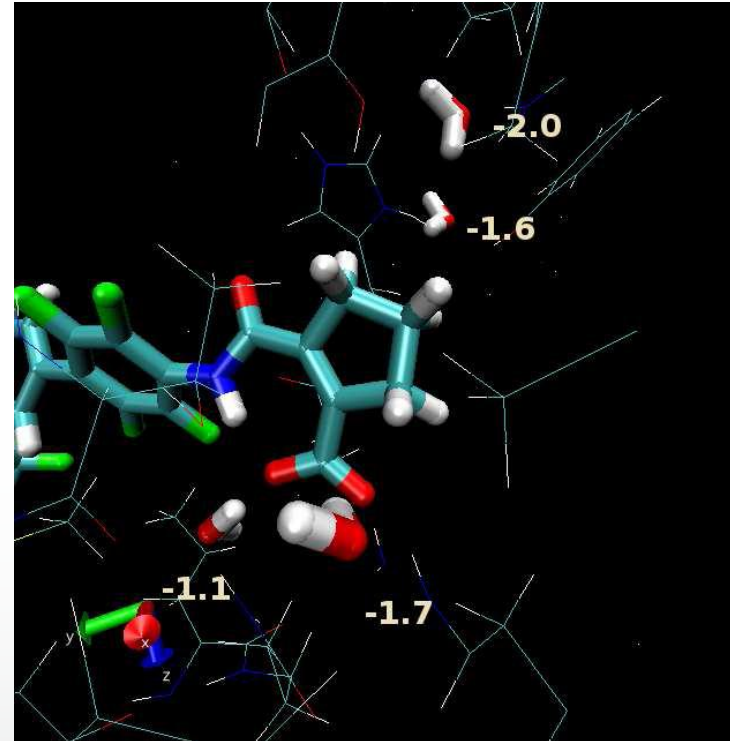
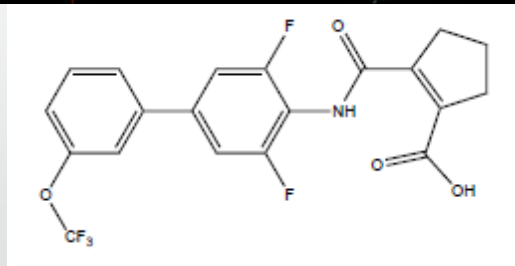
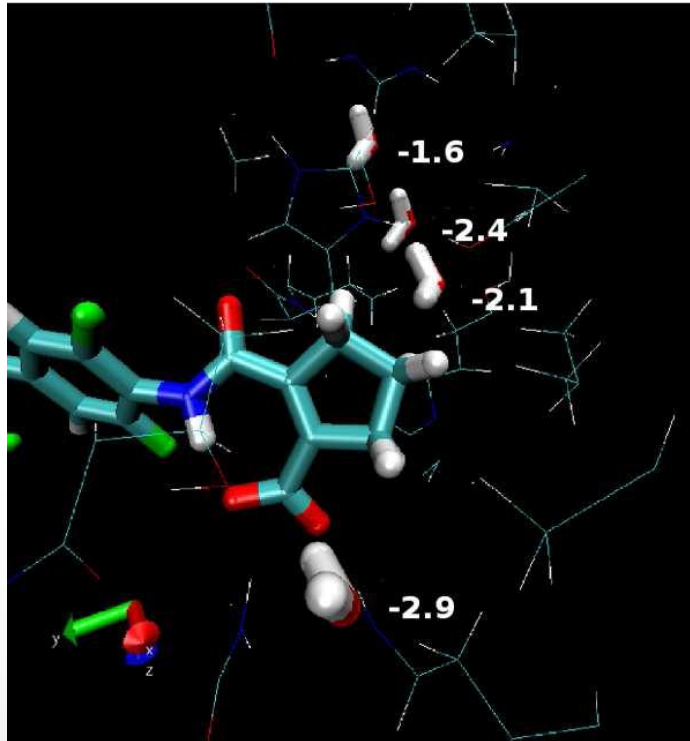
- Can visualise the number of waters throughout at the simulation at a set binding free energy
- Easy to see water networks forming throughout the simulation



# Comparison of methods



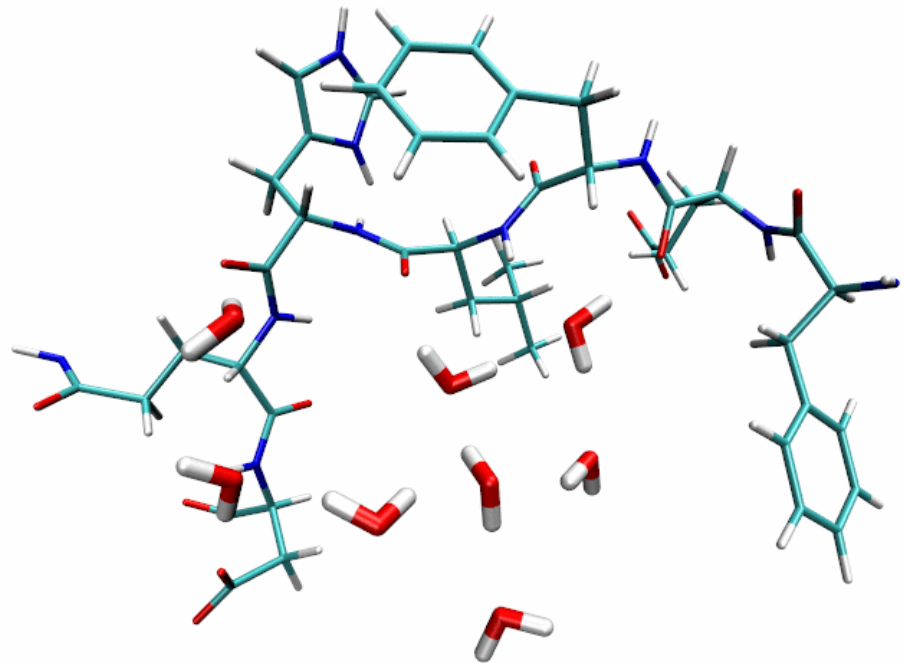
# DHODH – effect of ligand



- GCMC followed by JAWS stage II

# CDK2 kinase

- GCMC method used to look at the binding of water molecules to CDK2 kinase
- Pseudo-apo structure
- Can look at hydration patterns as a function of the binding free energy...

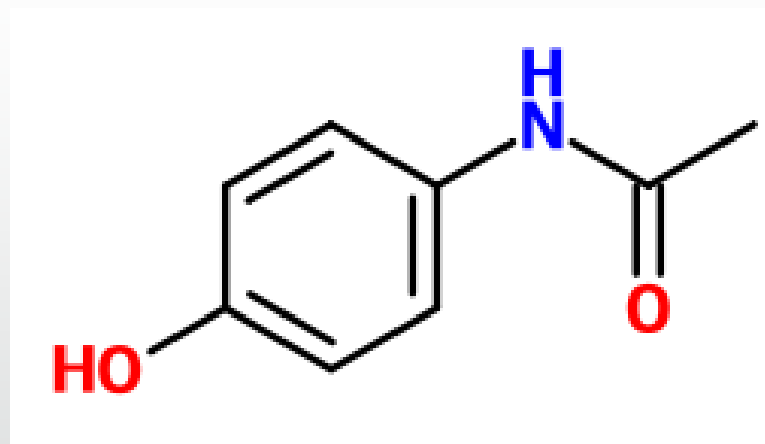
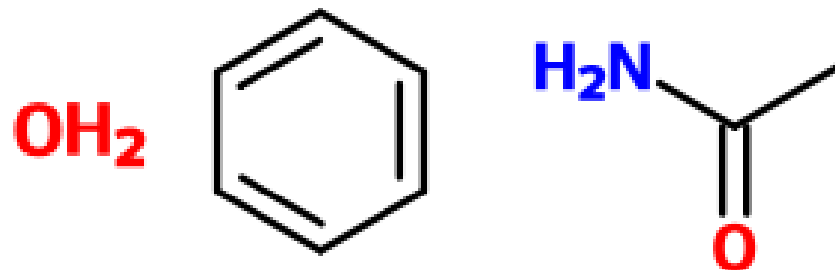


# Technique comparison

Technique	CPU Hours	Advantages	Disadvantages
Double decoupling	320 CPU hours	<ul style="list-style-type: none"> <li>• Theoretically rigorous and well-behaved</li> </ul>	<ul style="list-style-type: none"> <li>• Slow and computationally intensive</li> <li>• Know water locations</li> </ul>
GCMC	10 hours per B level x 10 levels = 100 CPU hours	<ul style="list-style-type: none"> <li>• Can locate water molecules</li> </ul>	<ul style="list-style-type: none"> <li>• Poor acceptance rates, sampling concerns</li> </ul>
JAWS	Stage one (5 hours) + stage two (15 hours) = 20 CPU hours	<ul style="list-style-type: none"> <li>• Can locate water molecules</li> <li>• Fast</li> <li>• Quick estimate of whether a water should be present</li> </ul>	<ul style="list-style-type: none"> <li>• Fails for tight binding molecules without biasing</li> <li>• Locating waters in diffuse maps an unresolved challenge</li> </ul>

# Fragment-Based Drug Discovery

- Relatively new technique
- Relies on efficient sampling of chemical space
  - $10^7$  common fragments against  $10^{60}$  ligands
- Hit rates higher than HTS



# Computational methods

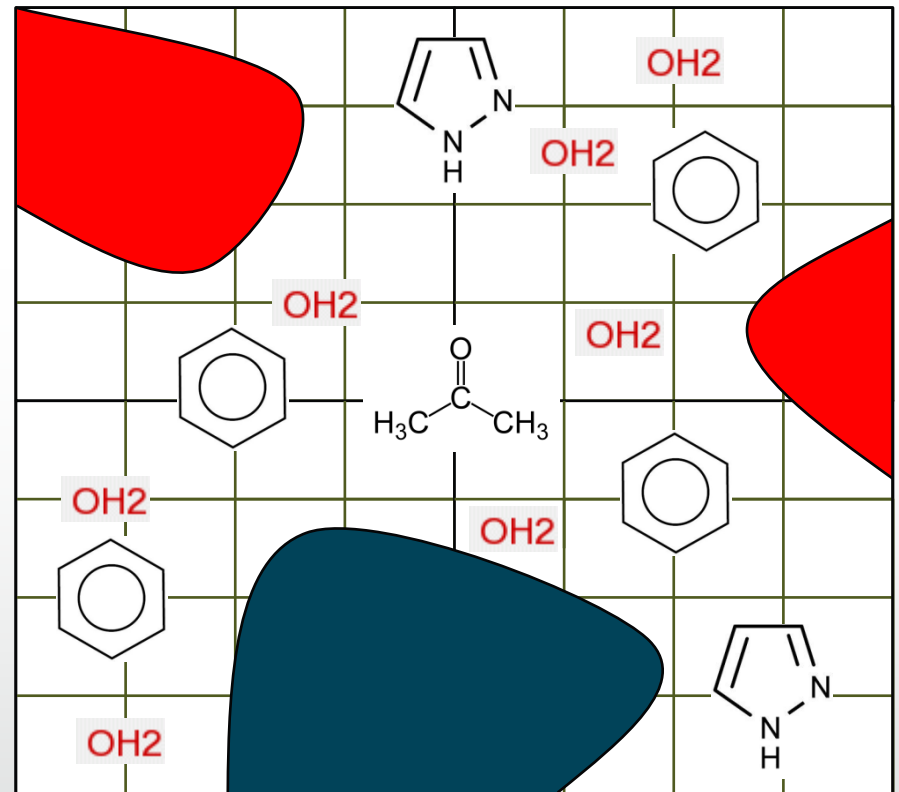
- Currently available (a selection):
  - Fragment docking (*J. Comput. Aided. Mol. Des.* **2009**, 23, 527)
  - FTMAP (*Bioinformatics* **2009**, 25, 621)
  - MCSS (*J. Comput. Aided. Mol. Des.* **2009**, 23, 475)
  - 3D-RISM (*J. Am. Chem. Soc.* **2009**, 131, 12430)
  - GCMC (*J. Chem. Inf. Model.* **2009**, 49, 934–943)
  - ...

# Computational methods

- What's desirable in a method?
  - Locate fragments based on free energy
  - Protein flexibility
  - “Accurate” energy functions
  - Allow competition between fragments, and most importantly, water
  - Rank fragment binding based on free energy
- JAWS method can do this for water – can it be extended to fragment mixtures?

# JAWS: From water to fragments

- JAWS method successful for a variety of different protein systems
- Can we apply this to fragments?
- Fragments can compete against water and each other



# Inclusion of desolvation penalty

- Mixing a pot of fragments and competing only tells half the story – JAWS I
  - Fragment might prefer to stay in the bulk
- Introduce a bias based upon decoupling the fragment in water
- Add this onto the potential energy function



$$\text{Bias} = \text{PMF}_{\text{old}} - \text{PMF}_{\text{new}}$$

# Further modifications

- Another issue which needs to be considered is the standard state

- 55 M vs. 1 M

- Included via a volume correction term



$$\text{Bias} = \theta_i (-kT \ln (V^{\text{sim}}/V^0))$$

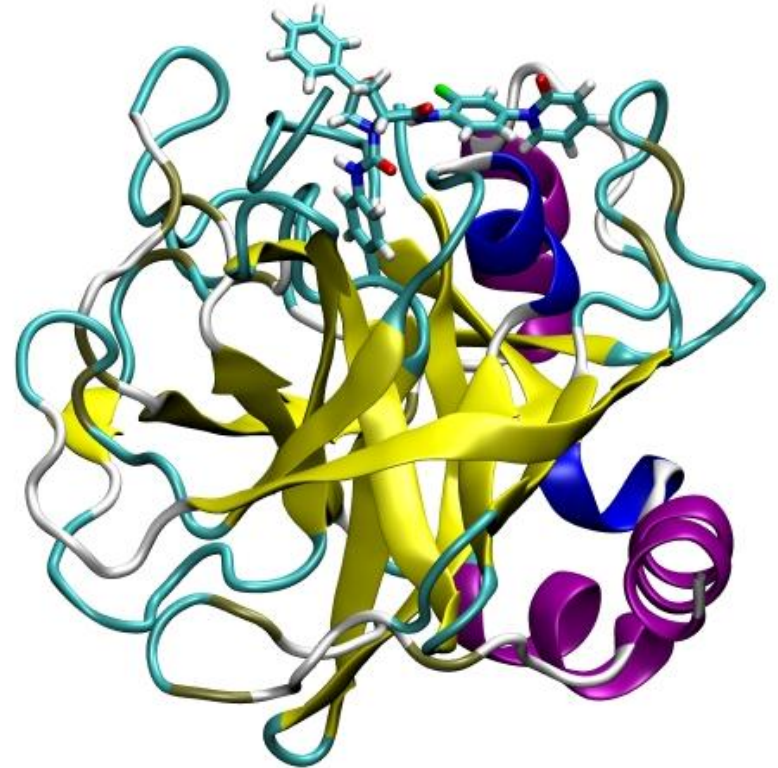
- Finally, we introduce a softcore potential for the fragments

- Allows molecules to pass over each other when 'off'

*Biophys. J.* **1997**, 72, 1047

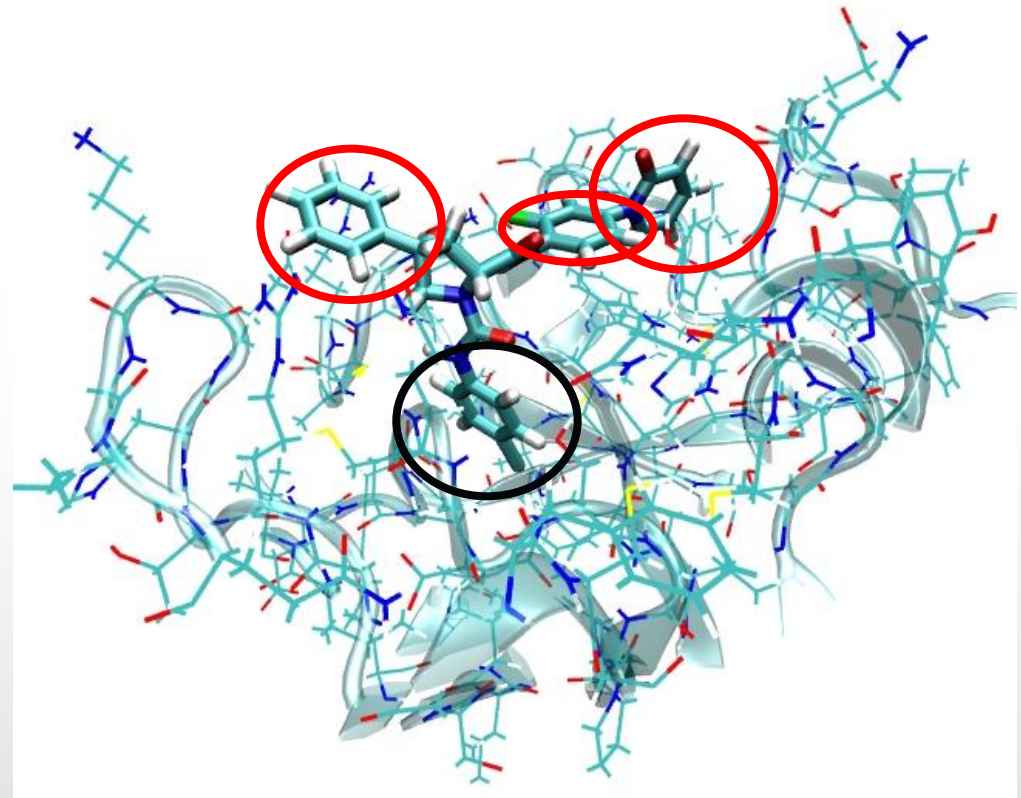
# Factor Xa

- Activated form of Factor X
- Plays a key role in the coagulation cascade
- Inhibition of Factor Xa results in anti-coagulation
  - Treatment of thrombosis



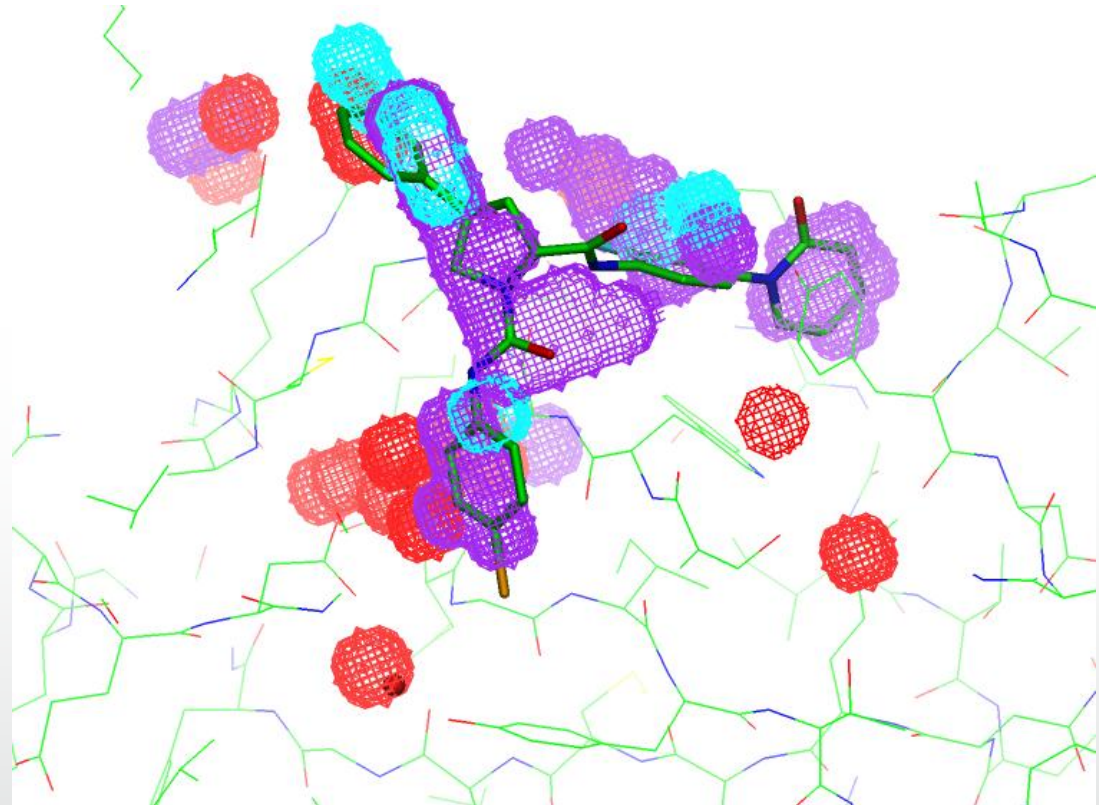
# Factor Xa binding pocket

- Buried phenyl pocket
- 4 aromatic regions in total
- Several polar regions scattered around the pocket



# Utilising JAWS in FBDD - fXa

- PDB : 2W3K.pdb
  - T-shaped ligand
  - Polar S1-pocket
- Three fragments: **Benzene**, **Pyrazole**, **Water**
- Observe a striking preference for **pyrazole** over **benzene** in S1 pocket
- Crystallographic water

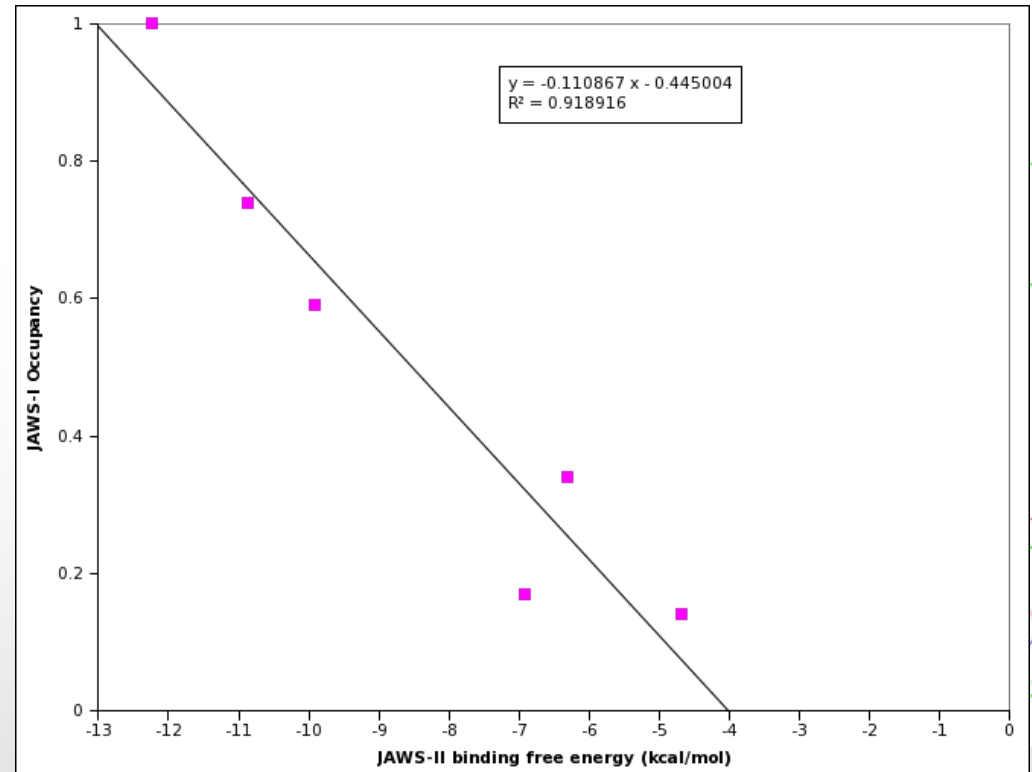


Increasing probability 



# What do these maps mean?

- Observe the major binding regions of each fragment
- More populated regions indicate favourable binding regions
  - Indication of the binding free energy of the fragment
- Guide inhibitor design?



# Issues....

- Much more validation and testing required
- Can we capture the correlated behaviour of the fragments better?
- What about the orientations of the fragments?
- Can we be more quantitative? Perhaps run JAWS II on the most populated fragment sites to get a direct estimate of the binding free energy?
- Issues of sampling and a combinatorial explosion as the number of fragments increases
  - Groups of fragments to explore particular chemistries?

# Conclusions

- Methods developed can locate water molecules and fragments and give a rough indication of the binding free energy
  - Allows competition with solvent, protein flexibility, desolvation etc.
- Need to test fragment-JAWS against real life assay data and run at correct concentration (~ 50 mM)
- Need to find a more reliable way of energetically scoring the fragment hits

# Acknowledgments

- Michael Bodnarchuk
- Dr Russell Viner
  
- Syngenta, BBSRC