

---

---

# INTRODUCTION

SCHRÖDINGER.

# Introduction

---

- Ligand based drug-design (LBDD)
  - As its name suggests LBDD focuses solely on the structure of the ligands.
    - Often LBDD is the only possibility, many important target classes have little or no structural data available.
    - Nonetheless, with care, powerful and useful models can be developed which can provide guidelines to synthetic-chemists.
- Structure based drug-design (SBDD)
  - In SBDD we also get to consider the structure and influence of the target protein.
    - We have to be very careful in ensuring that the protein structure we're using is reasonable for the job in hand.
    - Of course if this condition is met, the protein structure provides us with a host of valuable information.
- Joint approaches
  - Naturally there is no reason to favour one approach over the other.
    - If structural data is available we should use it, but the techniques of LBDD still remain useful.

---

---

# **SELECTING CRYSTAL STRUCTURES**

# Selecting a Useful Crystal Structure

---

- All SBDD calculations use a crystal structure as their foundation.
  - Selecting the correct structure(s) to use determines the ultimate quality of our results.
- There are many things to look for:
  - The quality of the crystal structure.
  - Potential for movement:
    - Backbone conformational flexibility
    - Sidechain conformational flexibility
  - Relevance of the bound ligand.

---

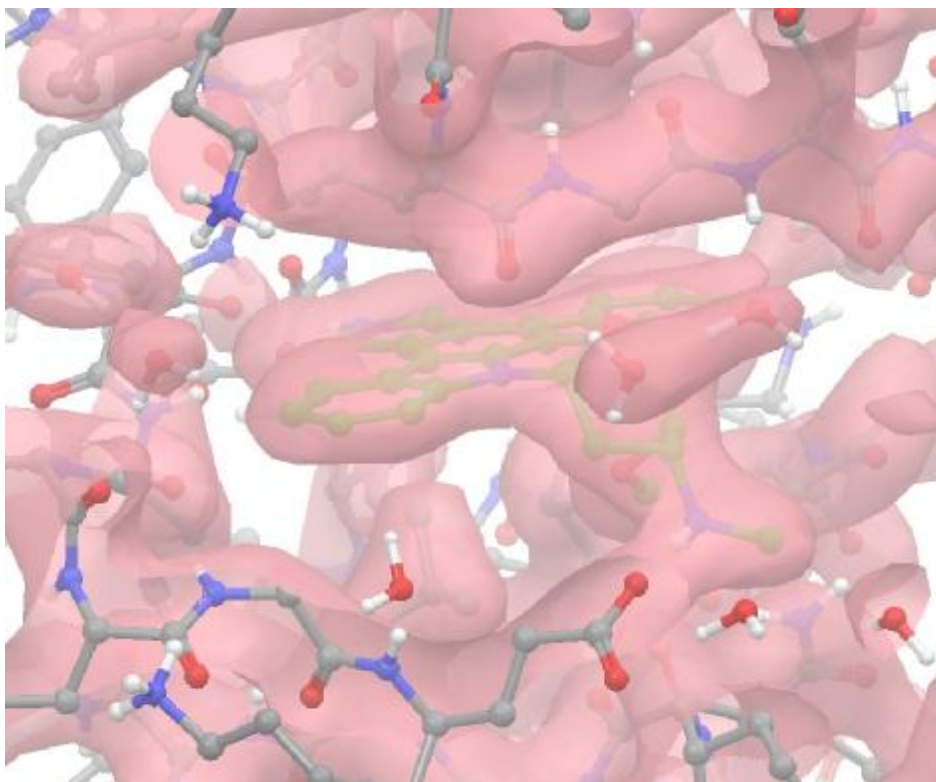
---

# ASSESSING CRYSTAL STRUCTURE QUALITY

# Crystallography is an Experimental Technique

---

- Crystal structures are derived by fitting atoms to the experimental electron density.
  - This requires some interpretation:

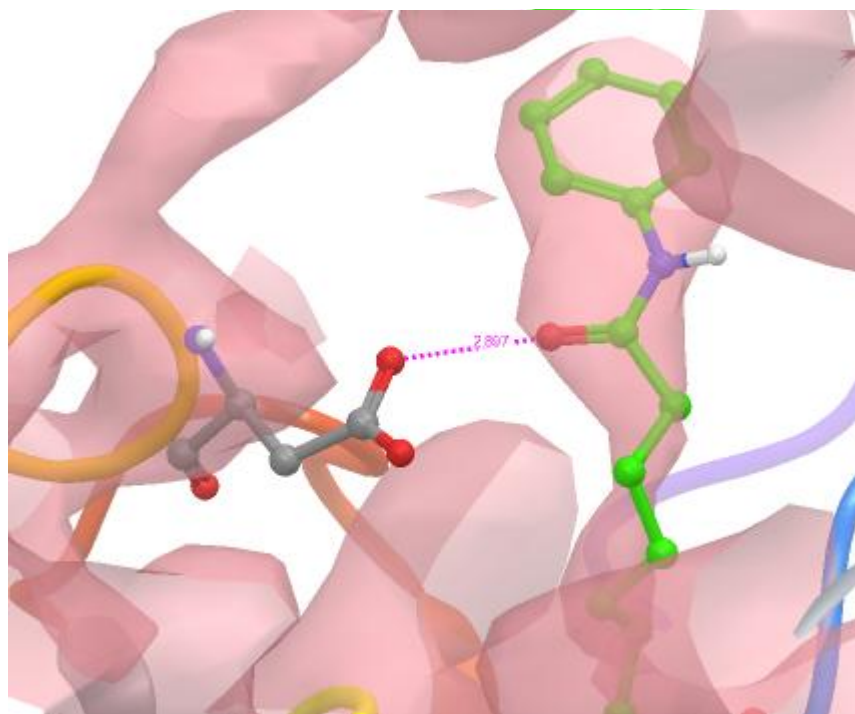


1OKY PDK-1/Staurosporine

# The Importance of Visual Inspection

---

- Crystallography tools exist to assist with the interpretation of electron density.
  - But things often go wrong.
    - When they do, they can ruin a whole series of perfectly good calculations.
- A visual inspection is the primary defence against this sort of issue.



1T69 HDAC8/SAHA

SCHRÖDINGER.

# Electron Density Tools in the Schrödinger Suite

---

- PrimeX is a complete crystallographic analysis and refinement package.
  - A comprehensive discussion of its capabilities is outside the scope of this talk.
- Perhaps its most relevant feature however, is that it is licensed **free of charge** to academic research institutions.
  - Speak to your academic account manager.



---

---

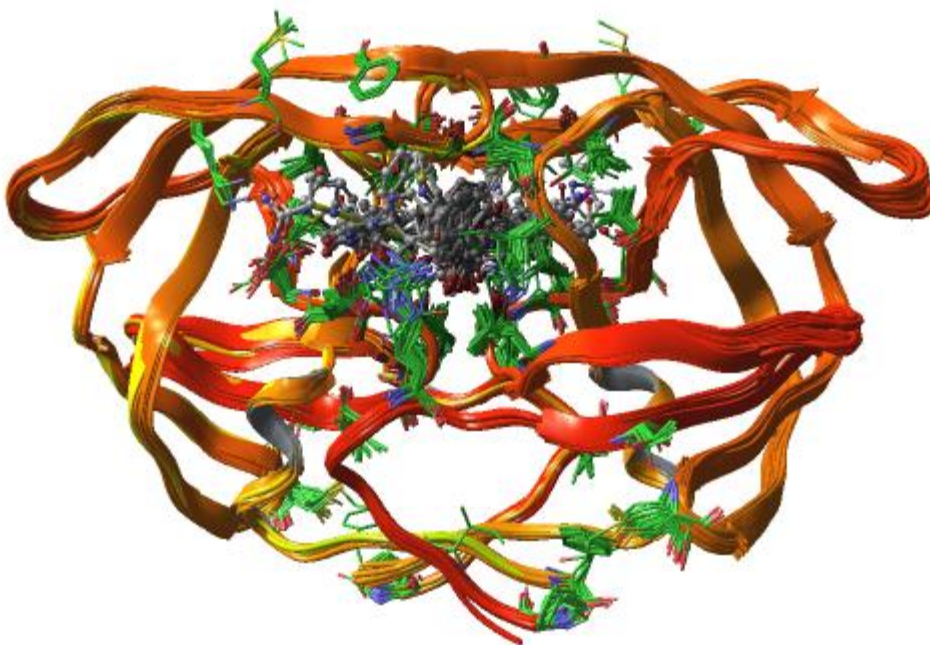
# THE POTENTIAL FOR PROTEIN FLEXIBILITY

SCHRÖDINGER.

# How Many Protein Structures Do You Need?

---

- Looking at a single crystal structure can give a very misleading impression of a protein.
  - All we're looking at is a single 'snapshot' of the protein.
- In reality things are often far more dynamic.
  - Having access to multiple crystal structures can really help.



Kinase bound to various  
ligands. Note. Motion is not  
restricted to side chains only.

---

---

# PROTEIN PREPARATION

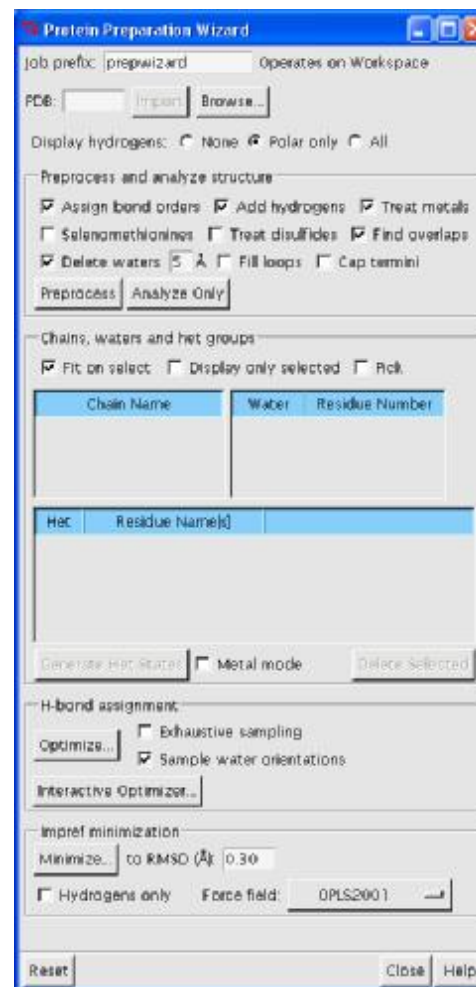
# The Importance Of Protein Preparation

---

- Almost all protein structures require some sort of remediation before they can be used in drug-design.
  - Protonation.
    - Most structures come from X-ray crystallography. As protons don't show up well in X-ray experiments they are normally missing from structures and need to be added.
  - Missing side-chains.
    - Any side-chain which is too mobile will not diffract well and will not be visible in the electron density. Simply ignoring this side-chain may not be a good idea as any ligand may well interact with the side-chain and cause it to adopt a fixed position.
  - Missing loops.
    - Similar to the above situation. However in this case whole residues can be missing from the final structure.
  - Counterions/random small molecules/waters.
    - The crystallisation media will often contain other counterions and small molecules along with water. These frequently show up in the final structure. Sometimes these species reveal important information (particularly water molecules), but in many cases they need to be removed.
  - Bonding/ionisation/tautomerisation.
    - Crystallography only provides the atomic coordinates of the structure. The bonding information needs to be added manually. For standard amino-acids this is trivial, however other species such as ligands and cofactors will need to be edited manually. Related to this the ionisation state and tautomerisation state of any non-standard species present will need to be assigned.

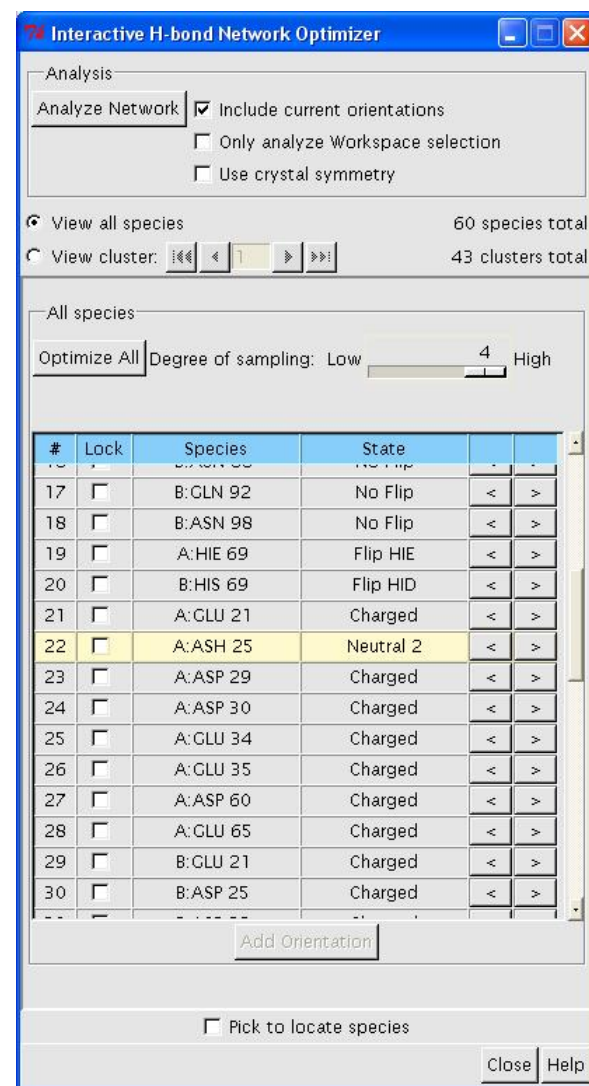
# Automated Protein Preparation in the Schrödinger Suite (PPrep)

- The PPrep wizard runs through the various stages required to remediate a protein structure.
  - The wizard is worked through from top to bottom.
- Most of the options are fairly straightforward.
  - The ‘Impref minimization’ option at the bottom is used to remove any ‘strain’ in the crystal structure.
    - It is a highly restrained optimisation procedure which is designed to correct the bond lengths/angles and interatomic distances to something which is more reasonable for the OPLS2001/5 force field.



# Preparing a Protein Structure

- For the most part the default options in PPrep can be applied.
  - However many systems require some manual intervention.
    - Catalytic residues frequently have unusual protonation states for example.
  - PPrep gives full control over this.



---

---

# GETTING A 'FEEL' FOR YOUR TARGET PROTEIN

SCHRÖDINGER.

# Grid Based Calculations

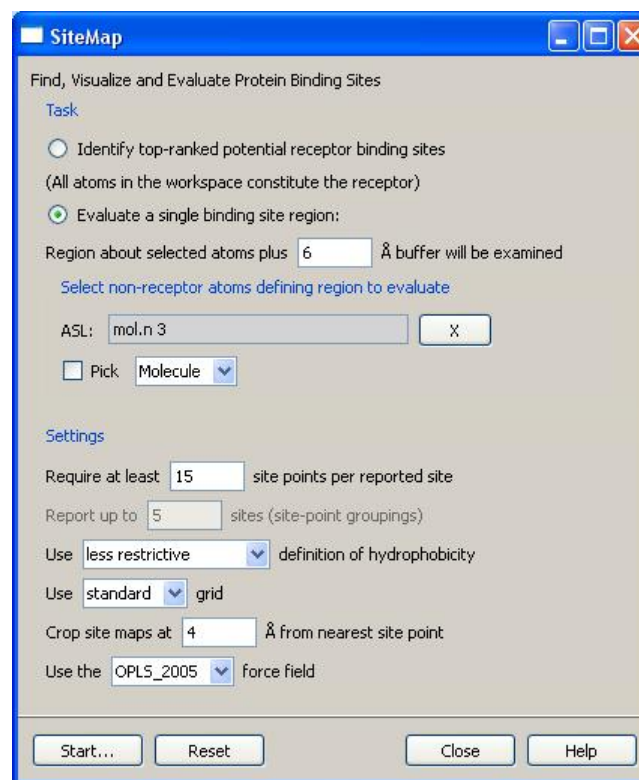
---

- Obviously the first thing to do when working on a new target is to sit and look at the target protein.
  - Search for clear regions where interactions can be made.
    - It is always tempting to consider hydrogen-bonds as being critical, largely because the software highlights these interactions.
    - However hydrophobic interactions are equally, if not more, important.
      - Hydrophobic interactions are however harder to spot. Grid based calculations make this easier.
- Grid based calculations work by evaluating the energies of various ‘probe atoms’ on a grid surrounding the protein.
  - Regions where these probes have particularly (un)favourable energies highlight the areas where certain interactions are important.



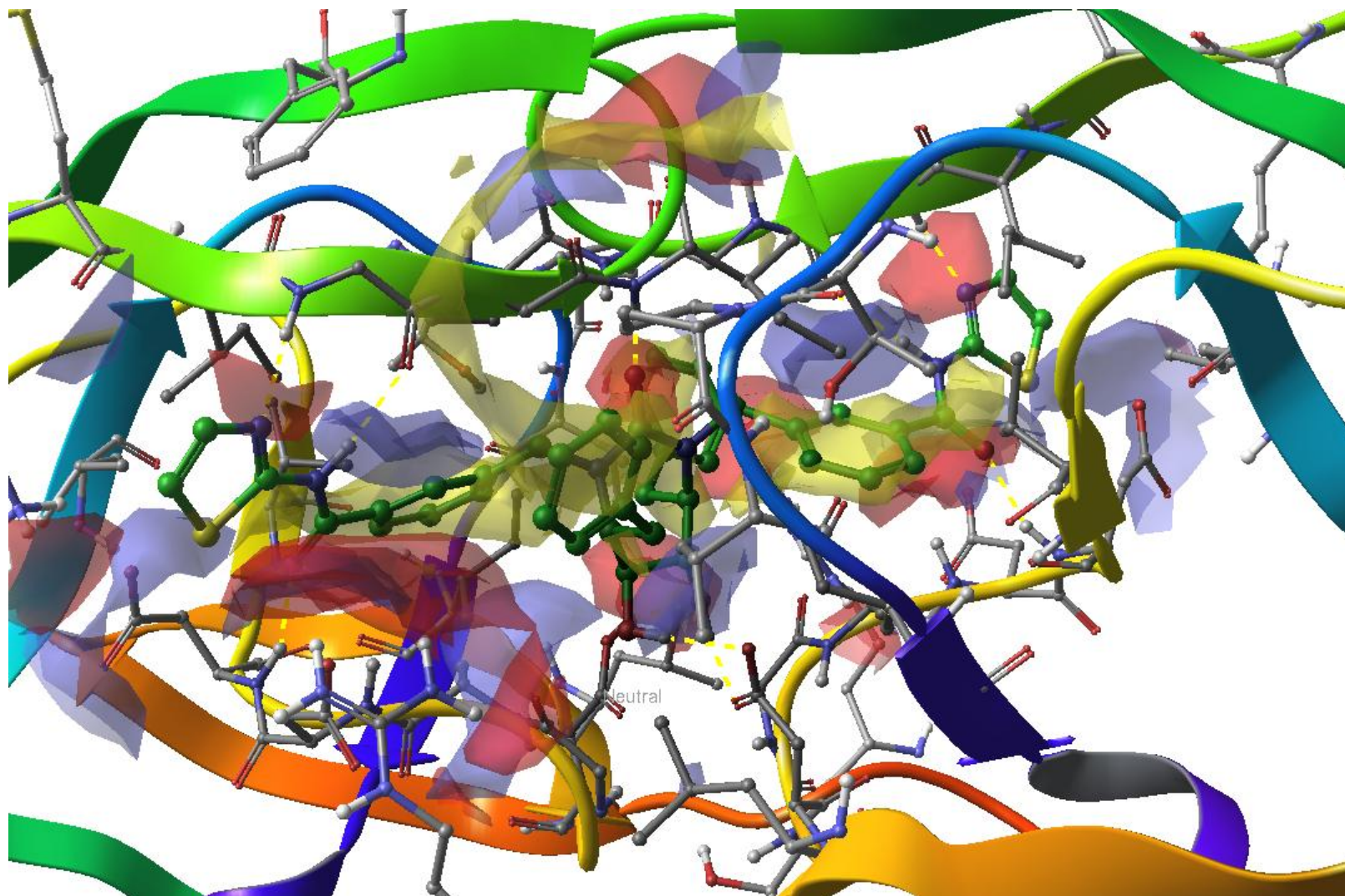
# Grid Based Calculations Within the Schrödinger Suite (SiteMap)

- SiteMap is used to set up, perform and analyse the grid based calculations.
  - Set up is via a very simple GUI.
- The primary output is a set of surfaces showing where various interaction types are preferred in the context of the active site.



# SiteMap Results: An Example

---



SCHRÖDINGER.

---

---

# DOCKING

# An Overview of Docking

---

- The aim of any docking program is to locate the translation, rotation and conformation of a given ligand which maximises the interactions of that ligand with the protein.
  - In this light all docking programs are simply search algorithms.
- The number of interactions made by the ligand to the protein is measured by the “scoring function”.
  - Hopefully the optimum value of the scoring function co-incides with the correct (crystallographically observed) pose.
  - Failing that, the correct pose should be in the top 5-10 results.
  - Considerable work goes into the creation of scoring functions.

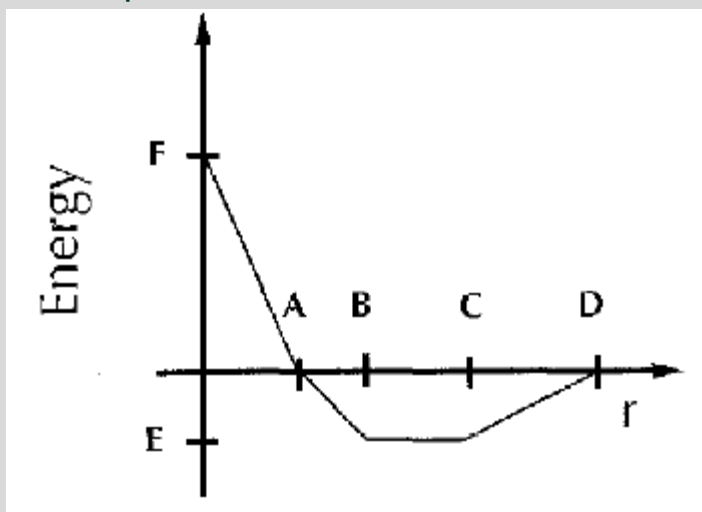
# An Overview of Docking

---

- The general search approach of docking programs has both strengths and weaknesses:
  - On the good side, the search carried out by the docking program can be entirely unbiased.
    - This can throw up interesting ‘alternative’ binding modes for a molecule which may be highly relevant.
  - On the bad side the search algorithm will almost always come up with a result.
    - As the only chemical sense that the docking program has is the scoring function, these results can frequently contain junk.
      - Filtering out this junk can be very time consuming.

# Additional Information: Scoring Functions

- Docking algorithms may evaluate hundreds or thousands of different ligand-protein poses.
  - Each of these poses is passed through the scoring function to determine whether it represents a plausible binding mode.
- Because the scoring function is invoked so many times, it is vital that it does not require significant computation.
  - This generally rules out a full force-field based calculation of the pose energy.
- Perhaps the simplest of all scoring functions is the Piecewise-Linear-Potential model (PLP).
  - This uses only a few atom types (donor, acceptor, both or hydrophobe) and the distances between pairs of atoms to determine the pair interaction energy according to the following simple linear model:



The parameters A-F are varied according to the atom types which are interacting, two classifications are used, *steric* or *hydrogen bonding*.

## Additional Information: Scoring Functions

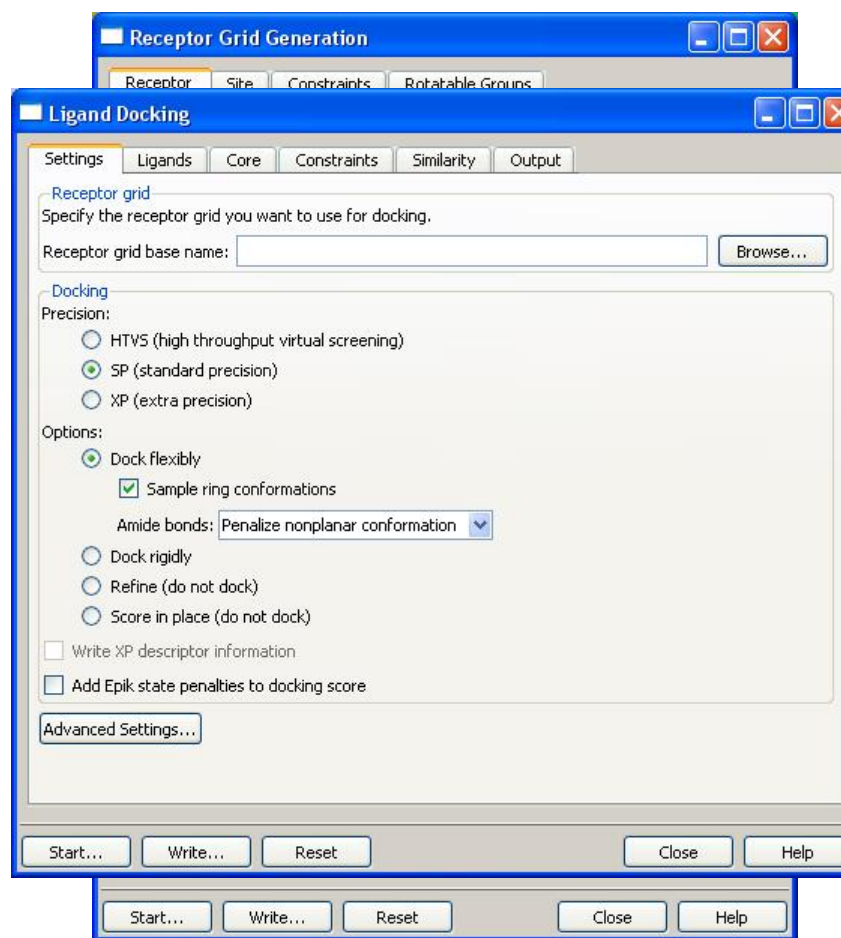
---

- With a limited number of atom types and interactions it is possible to pre-calculate all possible interactions within a protein cavity.
  - These pre-calculated values can then be stored on a 3D-grid.
  - Evaluating a docked ligand pose can then be a simple matter of looking up the precalculated energy from the correct region of the docking grid.
    - This look-up table approach massively speeds up the process of docking.
- That said, most scoring functions in widespread use are somewhat more complex than the PLP function described here.
  - Common examples are Chemscore and XScore, although these have each spawned their own variants.
    - These variants include a host of additional atom types and additional energy terms and penalties.
      - Examples of such terms include penalties for burying charges in hydrophobic regions, terms to include a measure of ligand strain in to the scoring function and terms which judge the quality of hydrophobic packing in the cavity.
  - Adding new terms to scoring functions is still an active area of research.
    - In all cases there is still a strong desire to keep the calculations simple enough to enable precalculation on a grid.



# Docking Solutions in the Schrödinger Suite (Glide)

- Glide is Schrödinger's primary docking tool.
  - It is based around a custom scoring function 'GlideScore'.
    - Considerable validation has been performed on GlideScore and generally it is considered one of the better scoring functions available.
      - GlideScore is a derivative of ChemScore.
- Most aspects of a docking calculation can be configured via the Glide GUIs.





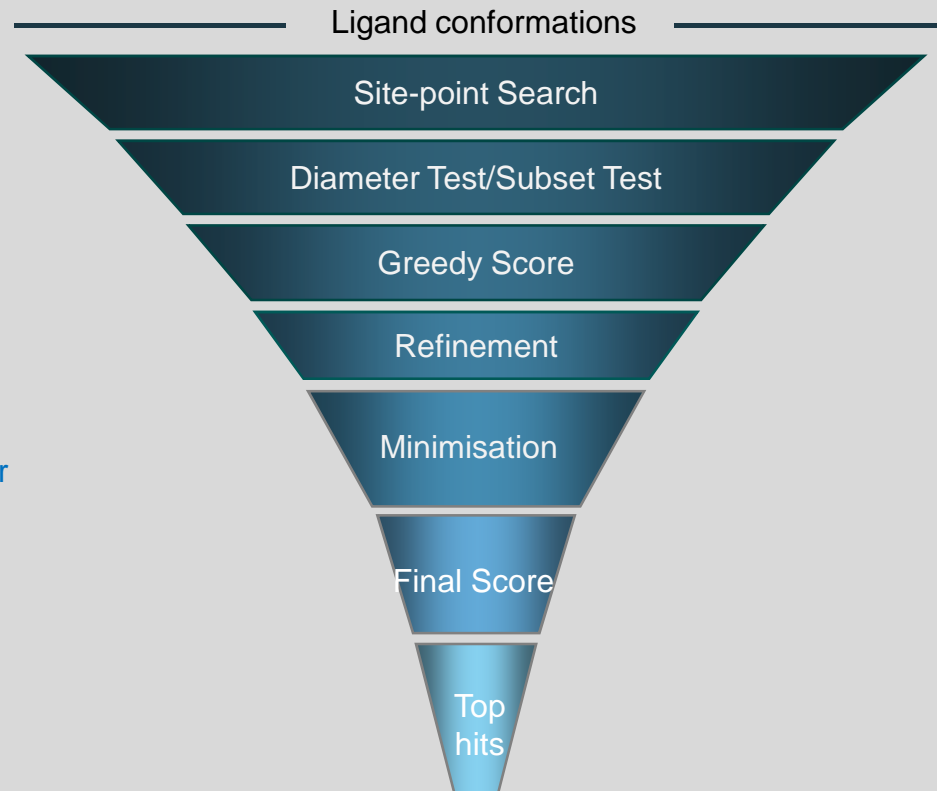
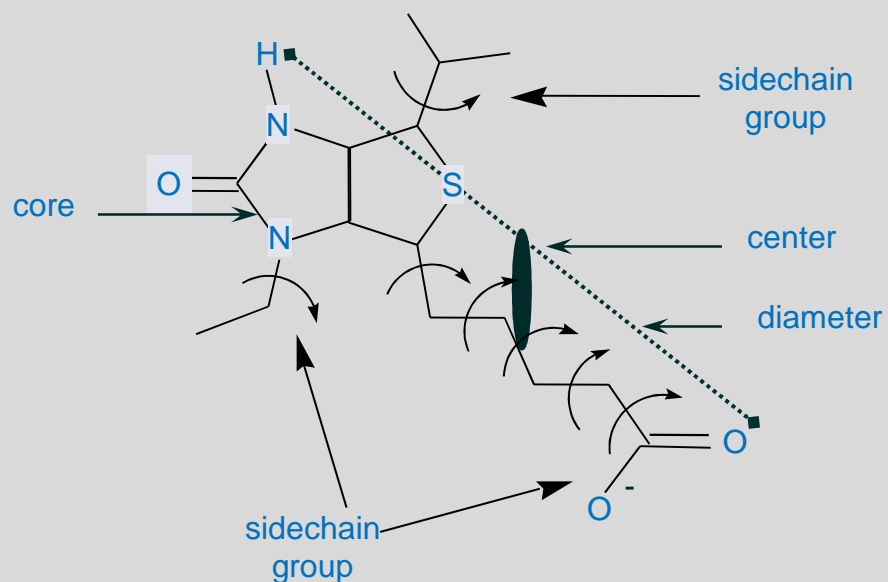
## Additional Information: Glide Implementation Details

---

- The standard Glide precision mode 'SP' provides a useful basis for discussing the overall strategy taken by Glide to perform a protein-ligand docking:
- In the first stage a series of ligand conformations are generated by ConfGen:
  - ConfGen is Schrödinger's small molecule conformation generator.
    - ConfGen is based around the OPLS force-field.
      - However the various terms within the force-field are rebalanced to reflect the energetics of a bound ligand vs. those of a free ligand.
        - » For example open ligand conformations are favoured.
        - » Charge-charge interactions in the ligand are reduced.
      - These tweaks to the force-field enable fewer conformations to be considered and speed up docking.
        - » Nonetheless it is often necessary to augment Glide's default conformational scan with an additional scan which considers more diverse conformations.

# Additional Information: Glide Implementational Details

- The conformations then enter the 'Glide Filter'
  - This is a bunch of hierarchical filters which are used to rapidly eliminate poses of the ligand which cannot correspond to a well-docked solution.



SCHRÖDINGER.

# Additional Information: Glide Implementation Details

---

- Site-point Search
  - The binding site is divided into a 2Å-grid.
    - The ligand is placed at these points and a rapid scan determines if there is any possibility that the ligand could fit in the binding site at this point.
- Diameter/Subset Test
  - Various orientations of the ligand are considered at each of the surviving site points.
    - The diameter of the ligand is used as a measure to determine whether the ligand can possibly fit in this orientation.
      - If there are too many clashes predicted, the orientation is discarded.
  - Then a subset of 'interacting' atoms on the surface are selected and scored at the surviving orientations.
    - If these fail to score reasonably the orientation is discarded.
- Greedy Scoring
  - The remaining poses are then scored using ChemScore.
    - The calculation takes place on a grid.
    - As all of the tests thus far have had quite a low resolution, this scoring step is slightly modified.
      - An atom receives the highest score it can based on the current grid point it occupies, or the maximum value of the surrounding grid points.
        - » It is this last aspect which makes this a 'greedy' score.
- Refinement
  - Only the best 100-400 poses are then enter the refinement stage.
    - Here the pose is optimised, using precalculated grids storing the OPLS vdW and Coulombic energies.
    - Where required, new conformations of ligand side-chains are generated in an attempt to improve the score.
- Final Scoring
  - The best poses following refinement are then scored using the full  $E_{\text{model}}$  and GlideScore equations.
    - $E_{\text{model}}$ : Used to rank the poses.
    - GlideScore: Used to rank ligands.

## Additional Information: Glide Implementation Details

---

- The value of `GlideScore` is determined as follows:

$$\text{GlideScore} = 0.065E_{\text{coul}} + 0.130E_{\text{vdW}} + E_{\text{Lipo}} + E_{\text{HBond}} + E_{\text{Metal}} + P_{\text{BuryP}} + P_{\text{RotB}} + \text{Site}$$

- The  $P_{\text{BuryP}}$  is a penalty term for burying polar functionality in a hydrophobic environment.
- The  $P_{\text{RotB}}$  is a penalty term for freezing rotatable bonds.
- The `Site` term rewards polar, but non-hydrogen bonding interactions in the site.

# Additional Information: Glide Implementation Details

---

- Glide also has two alternate docking modes:
  - Glide HTVS:
    - Designed for screening millions of molecules.
    - Features:
      - Sampling:
        - » Same sampling algorithm as SP with reduced funnel
        - » Reduced number of poses generated in ConfGen
        - » Maximum of 20 poses carried forward to refinement (vs. 400 in SP)
      - Truncated refinement
      - Fewer minimization steps cf. full minimization with SP
      - No full force field minimization
      - Scoring is based on a truncated scoring function.
  - Glide XP:
    - Designed for detailed docking of a few molecules.
    - Features:
      - Similar to Glide SP
        - » The funnel of filters is wider to allow a greater diversity of poses to be considered.
      - Ligand reconstruction algorithm differs
        - » Good scoring anchors are used as a starting point for ligand reconstruction.
        - » Enhanced sampling is performed on each ligand sidechain independently
      - XP GlideScore is then calculated including the various penalties and rewards.
        - » Side chains which suffer a penalty are then resampled at a very fine resolution to avoid the penalty.
      - Full minimization is applied prior to final rescoring with the XP GlideScore.

# Glide Publications

---

- Glide is a widely used and well-respected docking application:
  - Some recent publications detailing its use are:

Podvinec, M.; Lim, S. P.; Schmidt, T.; Scarsi, M.; Wen, D.; Sonntag, L.; Sanschagrin, P.; Shenkin, P. S.; Schwede, T., "Novel Inhibitors of Dengue Virus Methyltransferase: Discovery by in Vitro-Driven Virtual Screening on a Desktop Computer Grid," *J. Med. Chem.*, **2010**, *53*, 1483–1495

Ravindranathan, K. P.; Mandiyan, V.; Ekkati, A. R.; Bae, J. H.; Schlessinger, J.; Jorgensen, W. L., "Discovery of Novel Fibroblast Growth Factor Receptor 1 Kinase Inhibitors by Structure-Based Virtual Screening," *J. Med. Chem.*, **2010**, *53*, 1662–1672

Agostino, M.; Jene, C.; Byle, T.; Ramsland, P. A.; Yuriev, E., "Molecular Docking of Carbohydrate Ligands to Antibodies: Structural Validation against Crystal Structures," *J. Chem. Inf. Model.*, **2009**, *49*, 2749-2760.

Kawatkar, S.; Wang, H.; Czerminski, R.; Joseph-McCarthy, D., "Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide," *J. Comput. Aided Mol. Des.*, **2009**, *23*, 527-539.

Pierce, A.C.; Jacobs, M.; and Stuver-Moody, C., "Docking Study Yields Four Novel Inhibitors of the Protooncogene Pim-1 Kinase," *J. Med. Chem.*, **2008**, *51*, 1972–1975

For a full list visit: <http://www.schrodinger.com/productpublications/14/5/>

---

**NOW FOR SOME HANDS ON...**

SCHRÖDINGER.