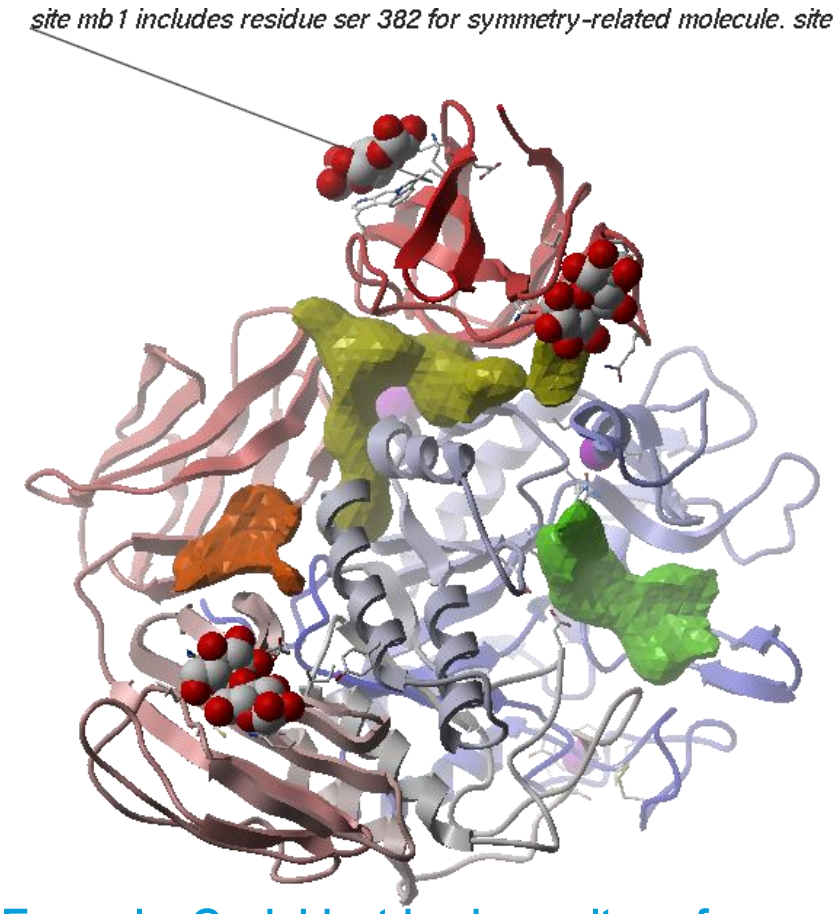
Refinement of a pdb-structure and Convert

Some Guidelines:

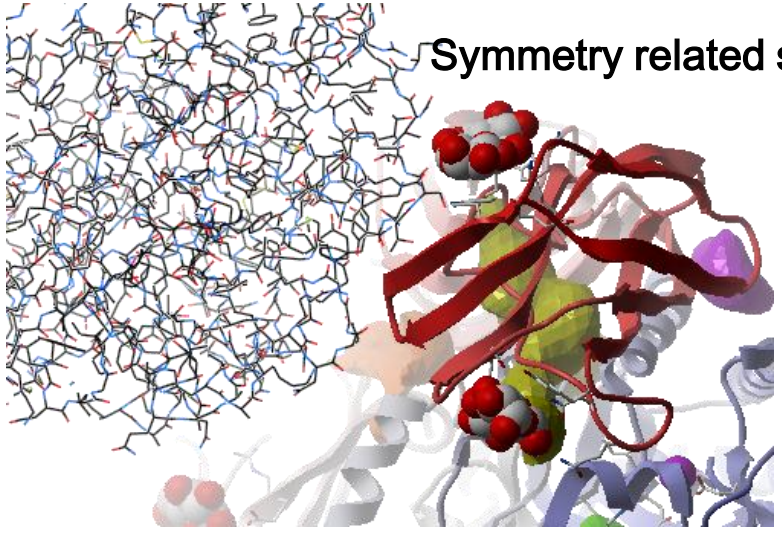
1. Search for a pdb with the closest sequence to your protein of interest.
2. Choose the most suitable entry (or several entries).
3. Convert and resolve errors and uncertainties.

- X-ray with up to 2.5-2.8A resolution is preferable over NMR
- NMR or homology models are only dockable by skillful operators
- Forget electron microscopy
- X-ray Resolution < 2.2 A is preferable.
  (Structures with resolution > 2.3 A may have up to 30% peptide flips, the maps are not self-refinable)
- Analyze symmetry if the pocket might be at the interface
- Analyze relative b-factors. B > 100. are not credible
- Pay attention to occupancies.
- Analyze alternative positions
- Check orientations of His, Asn, Gln
- Check protonation states of Glu, Asp, His
- Analyze stongly bound water molecules, ions and co-factors .

Preparations: symmetry



Example: Cycloldextrin glycosyltransferase
Entry: 1cdg, Res. 2.0A (Docking Rmsd without symmetry: 9.76)
More examples: transthyretin 1f41 (thyroid hormone binds at the dimer interface)



Problem: the true pocket is formed by chains which are not explicitly present in a pdb entry.

Goal: Find all molecules/subunits or chains involved in the interaction with the ligand.

Warning signs: ICM pocket finder does not show pocket density; Binding site is obviously exposed

Recovery: generate symmetry related subunits (View/Cryst.Cell)

Preparations: occupancies, b-factors and alternatives

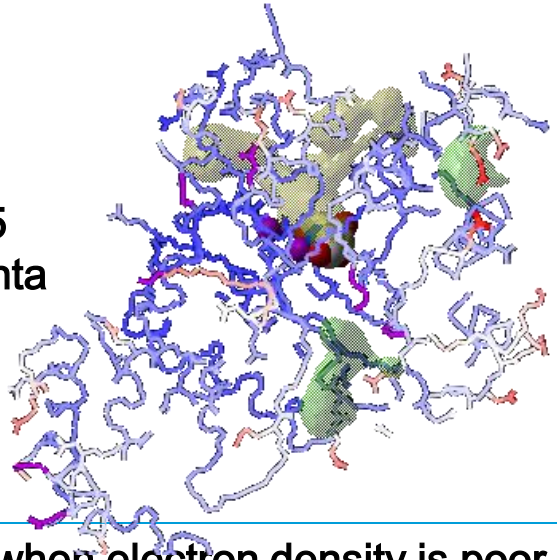
Glossary:

B-factor (or temperature factor):
mean-square displacement of atom from its position in the model.
$Bi = 79*\langle u^2 \rangle$ (B of 80 means 1A dev.)
Normal range: 5. – 50. A^2.

Occupancy:
A fraction of atomic density at a given center. It there are two equally occupied conformers, both will have occupancies of 0.5
Normal value: 1. Range: 0.-1.

Alternatives:
If two or more alternative conformations for the same atom or group are discernable in the density, several alternative sets of coordinates are deposited.

Occupancies <= 0.5 are shown in magenta
High b-factors are colored red



Problem: sometimes, when electron density is poor and/or ambiguous, crystallographers make things up (or just deposit an arbitrary conformation from a refinement program)

Goal: Identify fantasy atoms/groups

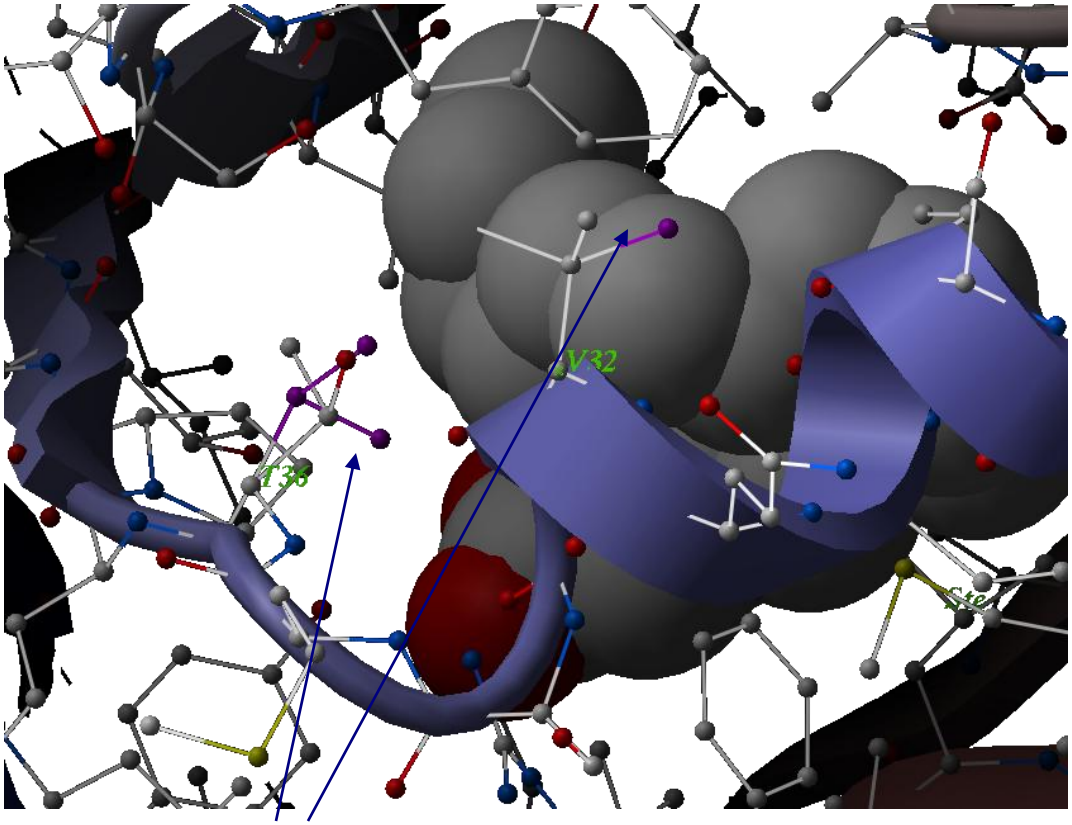
Warning signs: occupancies less than 0.5, b-factors larger than 60-80 A^2.
Tool: Color/label pocket atoms by occupancies/b-factors.

Recovery: Choose another entry, or refine with a ligand, or perform restrained minimization. Choose one of alternatives, or create alternative models

Preparations: occupancies, b-factors and alternatives. Example.

This is a very high resolution structure. For some key residues two alternative conformations are provided.

Recovery:
Choose one alternative or generate several separate docking models

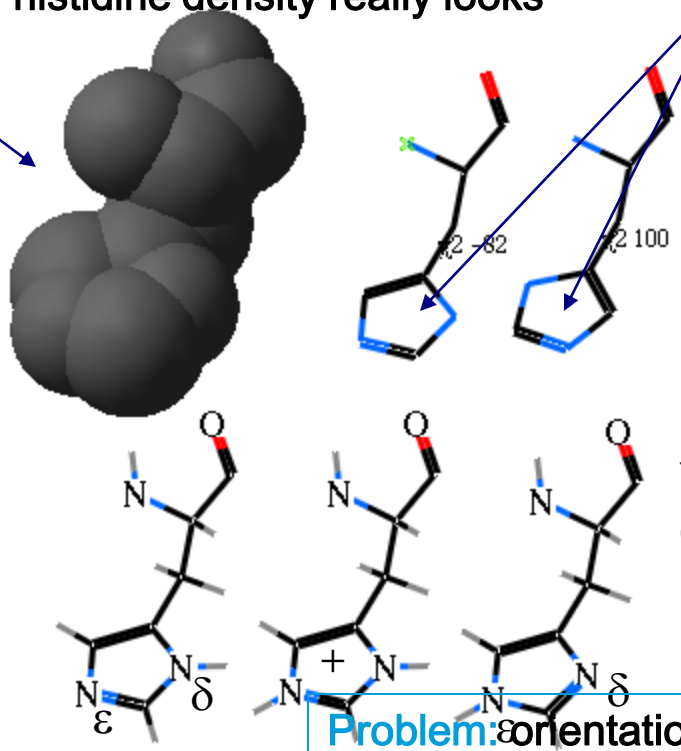


Alternative positions for Thr and Val32

Entry: 1hmt.
Res. 1.4
Fatty Acid Binding Protein with stearic acid

Preparations: fixing histidines

That is how histidine density really looks



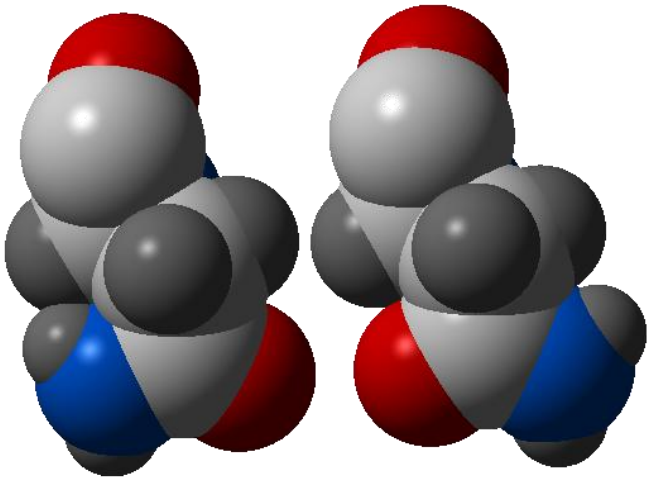
Orientation at the heavy atom level
We need to discriminate between
These two conformations

Often the xi2 angle needs to be
Corrected by 180 degrees.

Uncertainly at the protonation level
You need to decide which of the three conformations is correct for each important location. The charged conformation is rare.

Problem: orientations and protonation states of histidines are frequently wrong in pdb entries and need to be fixed to ensure correct docking results.
Placement principle: maximization of hydrogen bonds and other interactions with the rest of the protein and/or with the ligand.
Recovery: ICM procedure optimizeHisProAsnGln finds the best orientation and protonation state

Preparations: determining orientations of Gln, Asn, side chains



Orientation **at the heavy atom level**
The two conformations shown give similar electron density.
We need to discriminate between these two conformations of the **Asn** side chains. The same ambiguity needs to be resolved for the xi3 angle of **Gln**

Background: **xi2 in asparagines and xi3 in glutamines are frequently wrong or undefined and need to be corrected ensure correct docking.**

Placement principle: **maximization of hydrogen bonds and other interactions with the rest of the protein and/or with the ligand.**

Recovery: **ICM optimizeHisProAsnGln procedure.**

Preparations: do I need to uncharge Asp, Glu, Lys and Arg?

Definitions: DERK is Asp (D), Glu (E), Arg(R) or Lys (K)
Facts: pKs: His 6.0, Cys 8.3 Glu 4.2, Asp 3.9

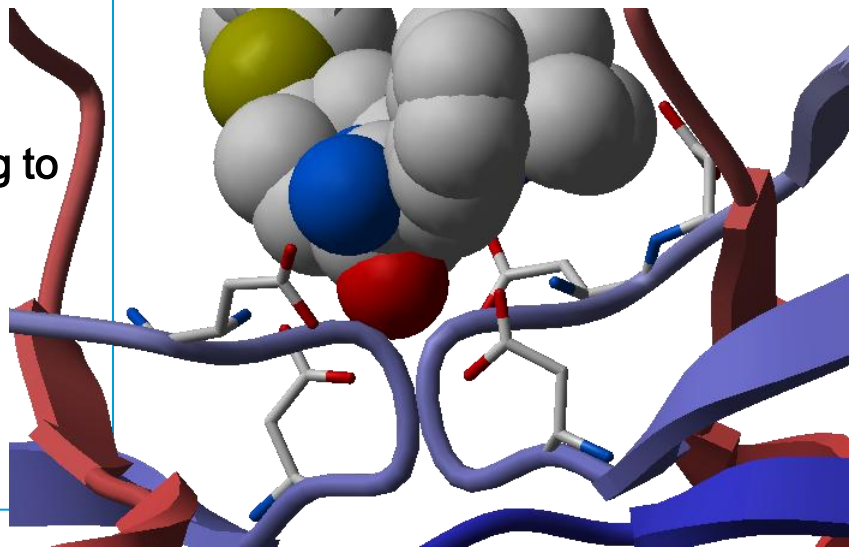
General recommendation: keep the DERK residues charged.

Problem: while in most cases DERKs are charged, in some special cases ED need to be uncharged or His needs to be charged.

Warning signs:
a DERK is buried and NOT involved in a salt bridge;
Several DERKs of the same kind/charge are pointing to the same space.

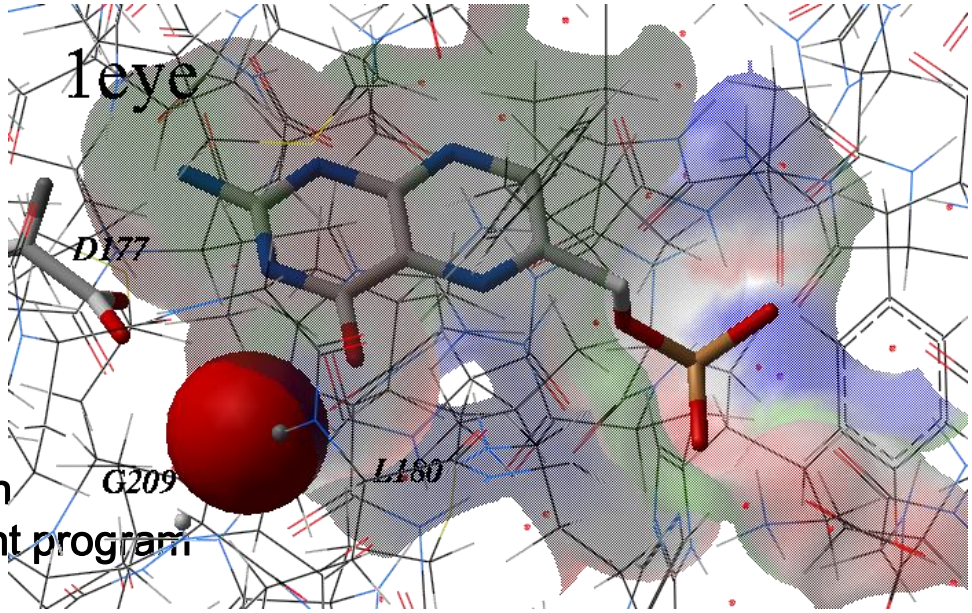
Example: HIV protease. 1ida. Asp 25 and 25' are protonated.

Recovery: Modify them to the uncharged forms.



Preparations: which waters to keep?

Example: 1eye dihydropteroate synthase, anti-mycobacterial/TB target. It binds to the buried Asp177 and improves electrostatic desolvation by ~10 units.



Definition: crystallographic water: an oxygen placed by a crystallographer or a refinement program to a blob of electron density.

General recommendation: get rid of all water molecules,
Keep only water molecules with three or four hydrogen bonds with the protein or ligand atoms.

Reason: keeping inappropriate water(s) will prevect correct docking, while dropping good waters is usually tolerated.
However some tightly bound water molecules help docking and scoring and prevet from erroneous placement of H-bond-rich ligand groups in water sites.

Recovery: Find interface waters with 3 or more protein/ligand neighbors and include them into your model.

Preparations: cofactors and metals?

Problem: metals may be required to dock a charged native ligand (e.g. ATP is charged and requires 2 Mn++ ions.)

However, to the metals may not always be necessary for docking of neutral drugs.

Example: a kinase domain. 1atp