Interpretation of electron density

March 5, 2014

Electron Density

- Once you have determined the phases then use this information to calculate electron density
- Then build into electron density.
- Atomic resolution requires about 1-1.2Å resolution
- Most protein crystals do not diffract to this resolution
- For protein and DNA we do know the structure of the individual amino acids and nucleotides
- We can approximate their position in the electron density

Calculation of Electron Density

- Structure factor amplitudes calculated from the model, $|F_{calc}|$ (F_C), and the measured structure factors from the diffraction patterns, $|F_{obs}|$ (F_O).
- Two types of maps are calculated $2F_O$-F_C and F_O-F_c maps
- Electron density calculation
- $2F_O$-F_C

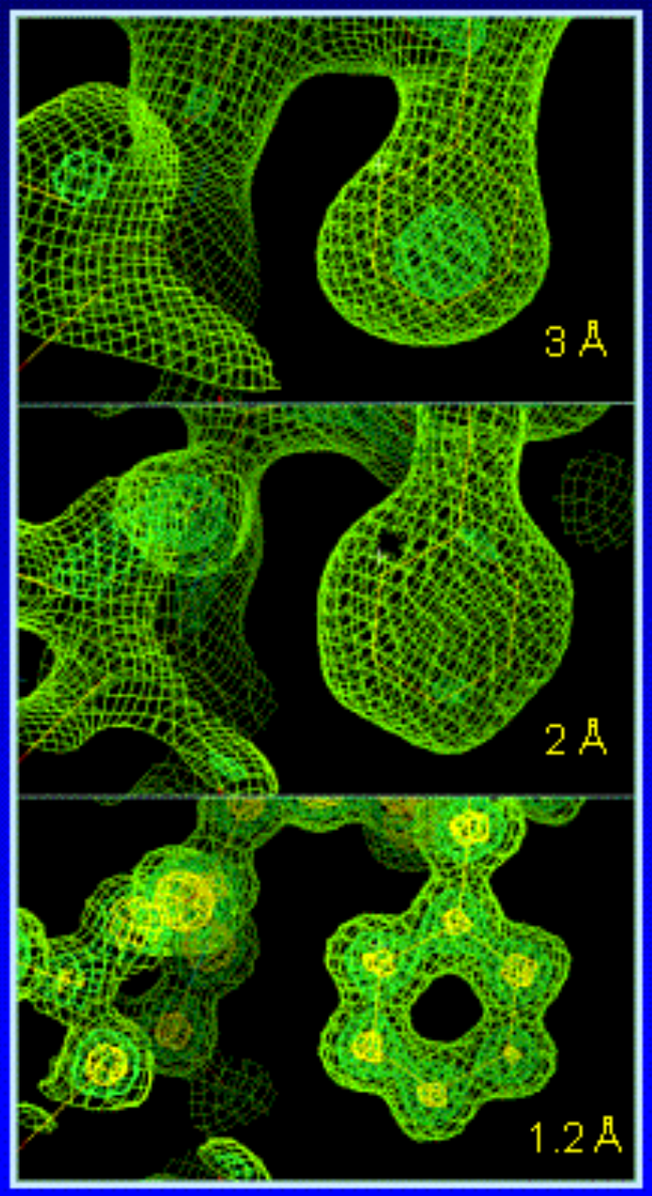
$$\rho(x, y, z) = \frac{1}{V} \sum_{hkl} (2|F_{obs}| - |F_{calculated}|) \exp[-2\pi i(hx + ky + lz) + i\alpha_{calculated}]$$

- F_O-F_C

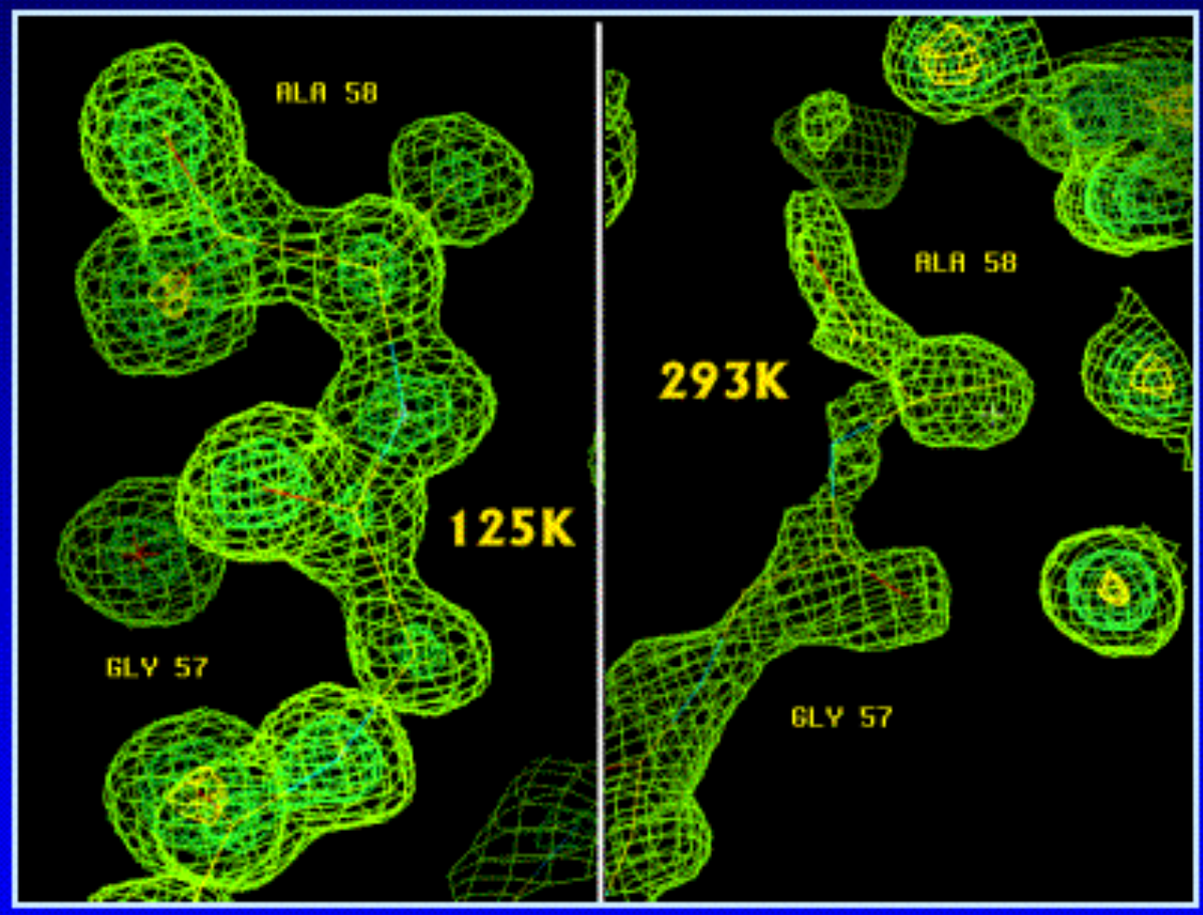
$$\Delta\rho(x, y, z) = \frac{1}{V} \sum_{hkl} (|F_{obs}| - |F_{calculated}|) \exp[-2\pi i(hx + ky + lz) + i\alpha_{calculated}]$$

Advanced Methods in Modern Biomolecular Crystallography

Importance of resolution

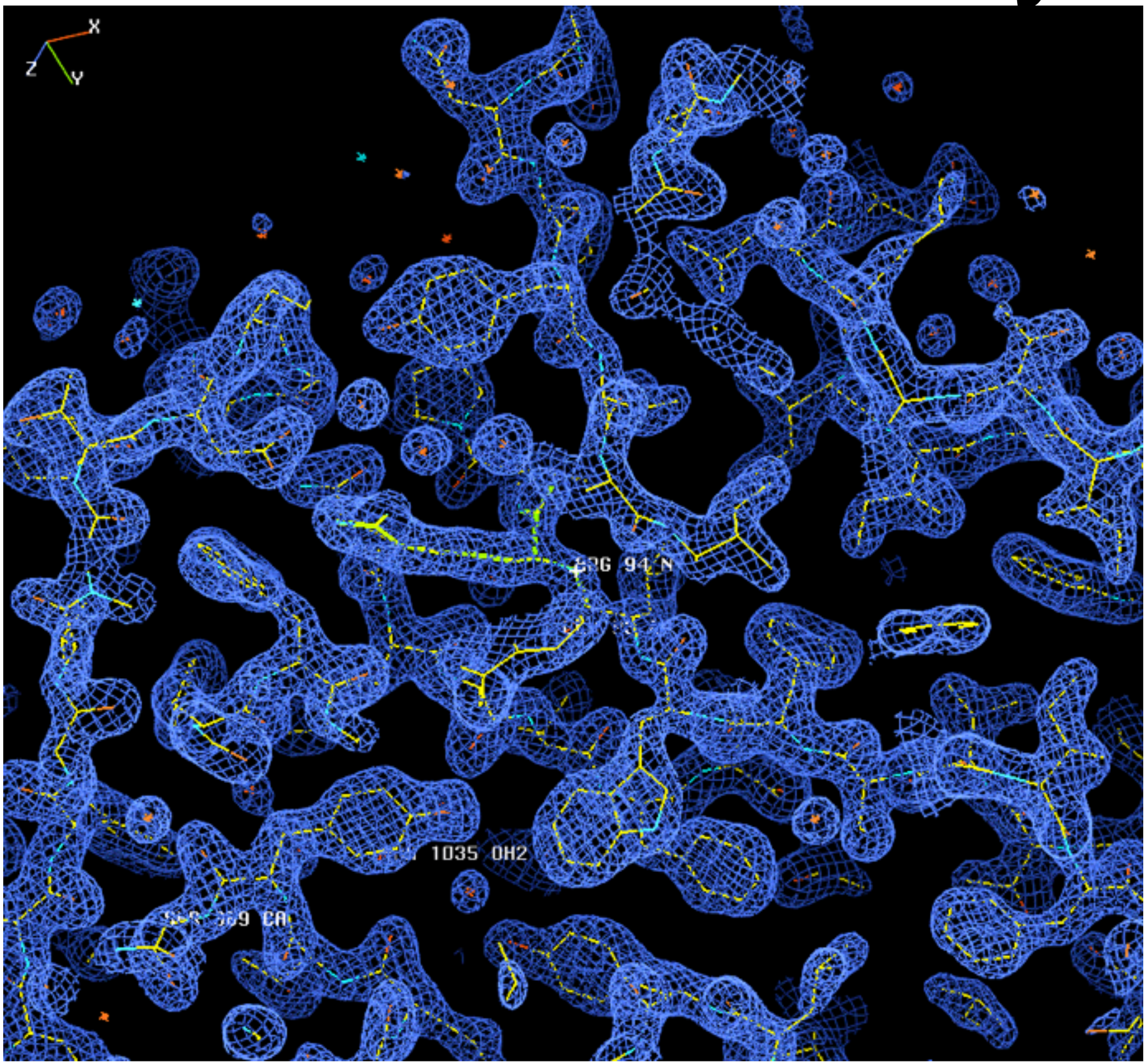


Reduced disorder at low temperature



Dramatic improvements in the overall structure are likely to result from better definition of disordered regions regardless of resolution

Electron Density



Model Building

- The ultimate goal is to produce the best model that fits the data. Remember that this is a model with bias from the person solving the structure.

- Steps (1) phasing the structure factors; (2) calculating an electron density distribution; (3) fitting an atomic model to the electron density; (4) refining the model with respect to the initial experimental observations and structural chemistry; (5) deposition of the structure into the PBD

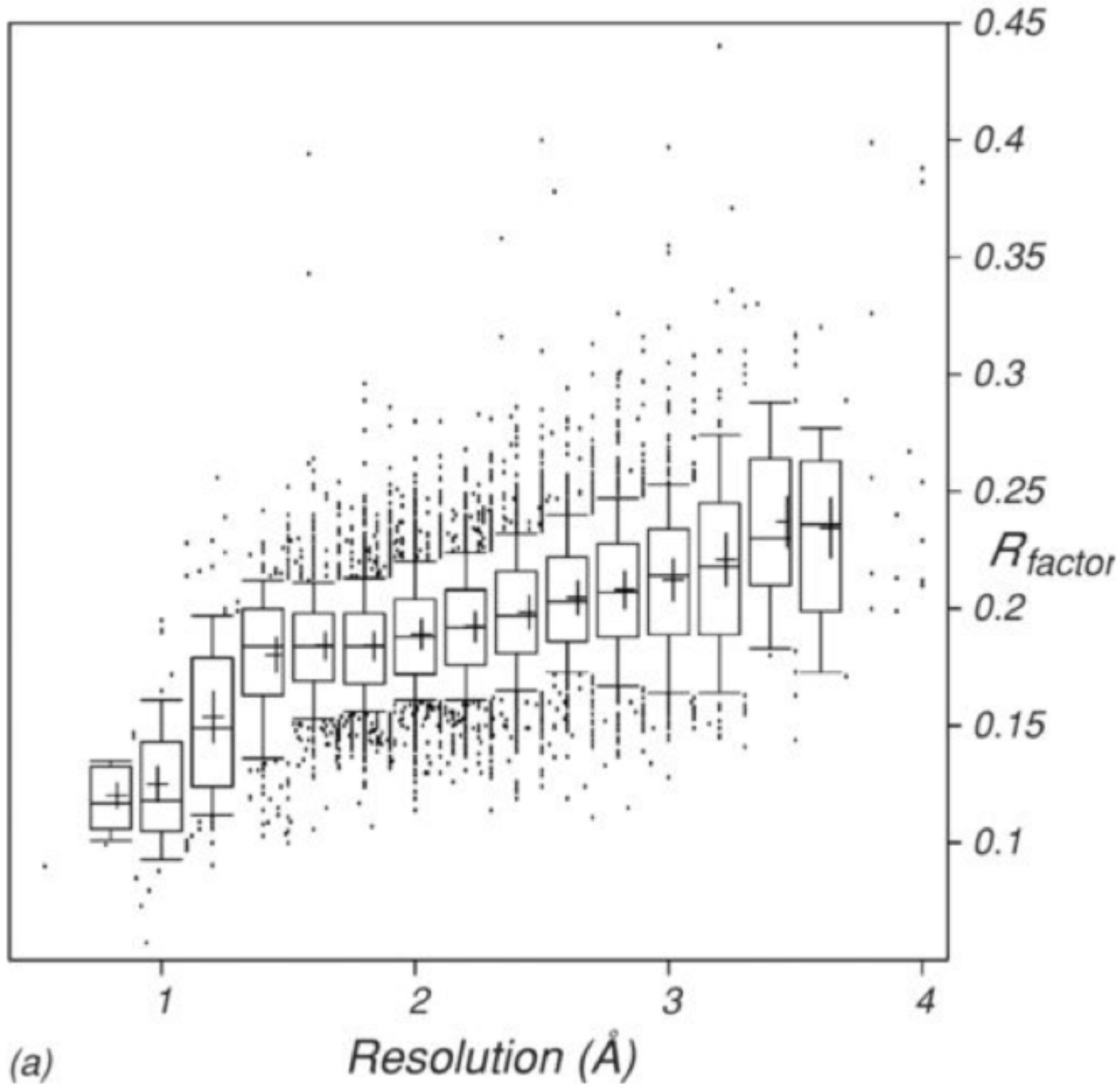
R_{factor} and R_{free}

- Structure factor amplitudes calculated from the model, $|F_{calc}|$, can be compared with the observations, $|F_{obs}|$
- R_{factor} and R_{free} are the tradition method for assessing how well the model fits the data.

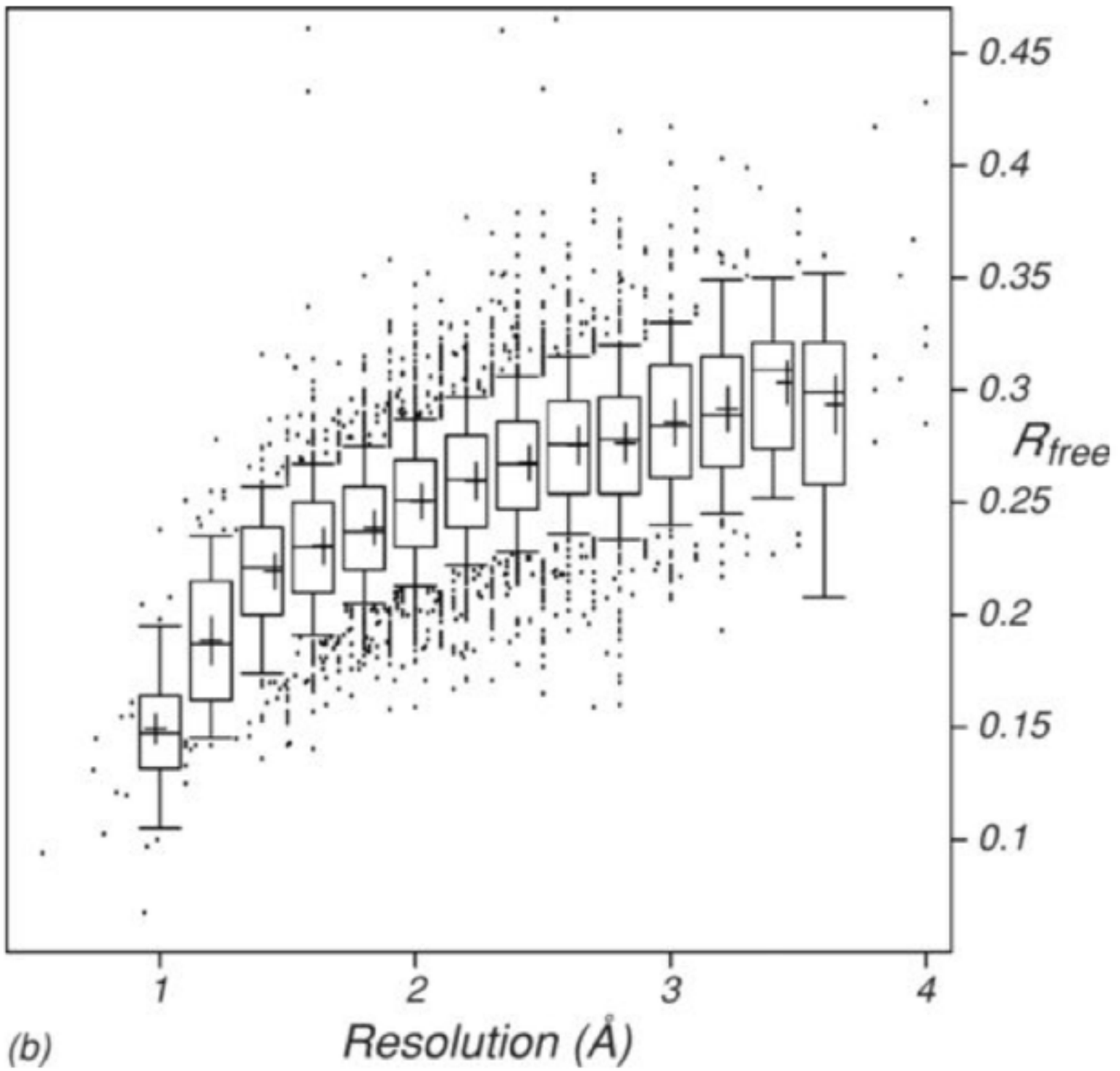
$$R_{\text{factor}} = \frac{\sum_{hkl} ||F_{\text{obs}}(hkl)| - K|F_{\text{calc}}(hkl)||}{\sum_{hkl} |F_{\text{obs}}(hkl)|}$$

-
- Random distribution of atoms placed at random in the unit cell would have an R_{factor} of 59% for *acentric* and 89% for *centric* (where the phase is 0 or π) reflections
- It is possible to *overfit* the model which results in an artificial lowering of R_{factor}
- It is possible to fit too many parameters for the number of observations present (Model bias)
- R_{free} was designed to prevent overfitting
- R_{free} is the same calculation with structure factor amplitudes that are not included in the refinement (5-10% of the reflections are removed).
- R_{free} validates the extent to which the model explains the diffraction data.

Final R_{factor} of Structures



Final R_{free} of Structures



Refinement

- Phases have a strong influence on map quality.
- Need to be careful about phase biased maps due to just the model.
- refinement procedure is to optimize the agreement between an atomic model with the observed diffraction data and chemical restraints.
- E_{chem} describes empirical information about the chemical interactions
- $E_{X\text{-}ray}$ term describes the difference between the observed and calculated diffraction data and can be understood as the experimental potential energy.
- $W_{X\text{-}ray}$ is a weight that controls the relative contributions of the E_{chem} and $E_{X\text{-}ray}$ terms.

$$E = E_{chem} + W_{X\text{-}ray}.E_{X\text{-}ray}$$

Refinement Procedures

Types of refinement

- *Solvent flattening* - a molecular envelope, or mask, can be created around protein and any density outside that mask is set to a uniform level

- *Gradient descent refinement methods* -refine the entire structure (rigid body) or each individual atom within the structure (positional refinement)

- *B factor refinement* - How much an atom contributes to the diffraction pattern is related to its temperature factor (B factor).
- $B = 8\pi^2 \langle u^2 \rangle$ where $\langle u^2 \rangle$ is the atom's mean square displacement relative to the average position

- *Simulated annealing*- allows for more sampling of conformational space. Simulated annealing can help overcome a local minima during refinement. based on molecular dynamics. Kinetic energy is initially added by increasing the temperature in order to permit the model to sample many alternative conformations and then slowly cooled to conformation more energetically compatible with the X-ray diffraction data.

Placing Ions and Solvent

- Often ions from the crystallization condition or water molecules are found associated with the model

- These can be placed into the model.

- Difficult to say for certain what the exact nature of the ion can be. Usually depends on the environment and coordination.

Agreement between the model and chemistry

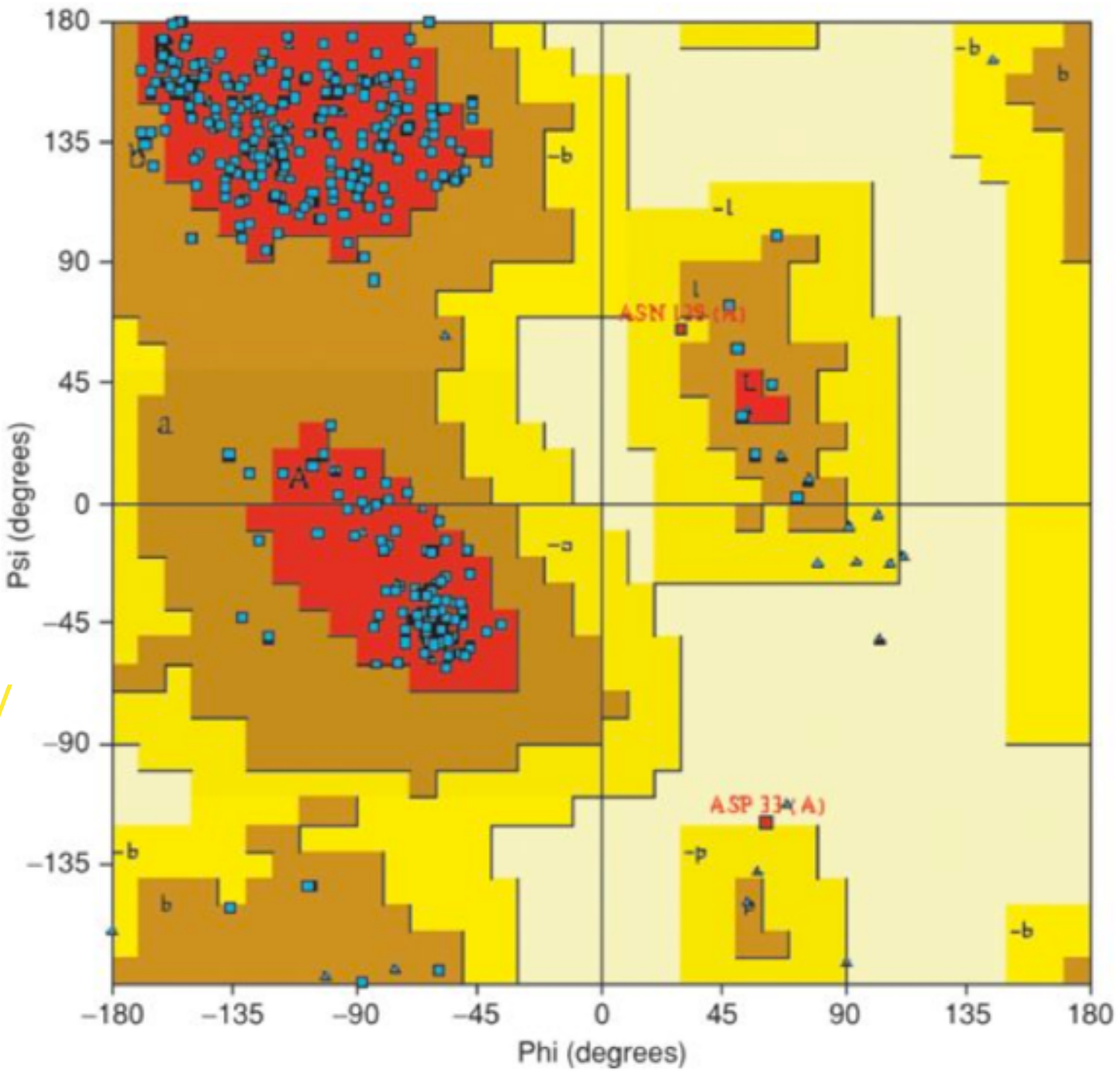
- Since most crystals do not diffraction to atomic resolution, it is critical to ensure that proper geometry is maintained.
- Check that the bond lengths and angles
- Ensure backbone conformations are in allowed regions of the Ramachadran plot

Ramachandran Plot

Favored

Allowed

Generously allowed



Outliners