

**APPENDIX 2: Items for
Independent Consideration
(Recall from BCH 408)**

Least Squares Structural Refinement

- Among the most common computational methods designed to minimize the agreement between the model and an ideal target geometry (derived from a library derived from high-resolution structures).
- Parameters of the minimization include:
 - 1). Agreement of the model position with the electron density
 - 2). Bond lengths
 - 3). Bond angles
 - 4). Bond (group) planarity
 - 5). Chiral Centers (or volume)
 - 6). Pseudorotation angles of puckered rings
 - 7). Torsional Angles
 - 8). Non-bonded contacts
- Minimization should reduce the R_{factor} and R_{free} .

Structural Refinement

A typical minimization function may have the following terms:

$$Q = \sum_{hkl} w_{hkl} [(|F_o| - |F_c|)]^2 + \sum_{j\text{-dist}} w_D [(|d_j^{ideal}| - |d_j^{model}|)]^2 +$$

$$\sum_k \sum_i w_P [(m_k \cdot r_{i,k} - d_k)]^2 +$$

planes co-planar atoms

$$\sum_l w_C [(V_l^{ideal} - V_l^{model})]^2 +$$

chiral V

$$\sum_m w_N [(|d_m^{min}| - |d_m^{model}|)]^2 +$$

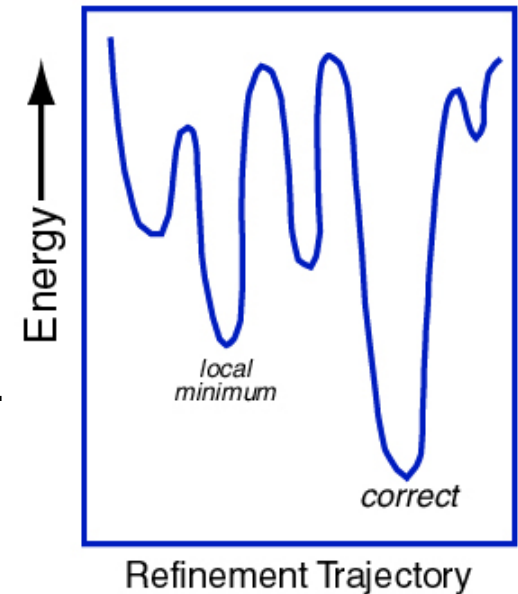
non-bonded

$$\sum_t w_t [(|X_t^{ideal}| - |X_t^{model}|)]^2 +$$

torsion angles

Radius of Convergence in Refinement

- The greatest (XYZ) distance that refinement will traverse to converge upon a correct structural solution is the *radius of convergence*.
- A theoretical value is $d_{\min}(\text{\AA})/4$ for conventional least squares, but is much greater for energy-related methods (several \AA 's distance).
- The method of least squares assumes the structure under refinement is not far from its correct atomic position.
- There may be numerous local energy *mimima* on the path to the lowest energy (correct) structure.
- The least squares method may not escape these local 'structural' wells and refinement may stall.
- This may be corrected by manual intervention by the crystallographer, and/or by use of *energy minimization* or *molecular dynamics algorithms*.



Measuring the Unit Cell Dimensions & Space Group Determination

$$a = \Delta a^* / d \lambda$$

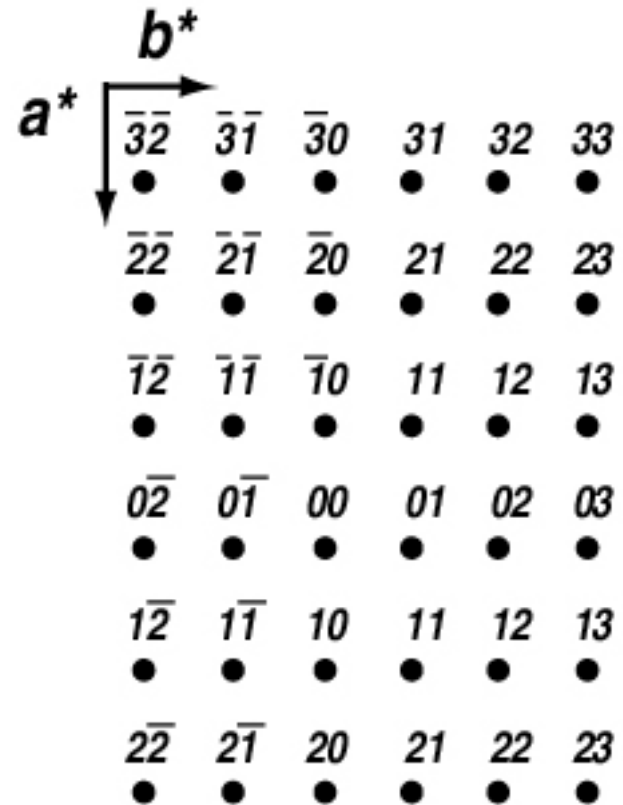
Δa^* = the distance between reflections

d = the crystal to detector distance

λ = the incident X-ray wavelength

a = the unit cell length in direct space

- This description is valid for $\alpha = \beta = \gamma = 90^\circ$.
- The volume can be calculated to tell what fraction of the unit cell is occupied by solvent.



Structure Quality Indicators

R-factor -- the best indicators that a model reflects the actual experimental data is the crystallographic R-factor (<25%) and the R-free (not more than 8% higher than the R-factor).

Resolution – the resolution of crystallography means we can resolve planes of atoms that are 2 Å apart. This resolution is sufficient to see carbonyl oxygens and waters. Higher resolution means a better model – low error.

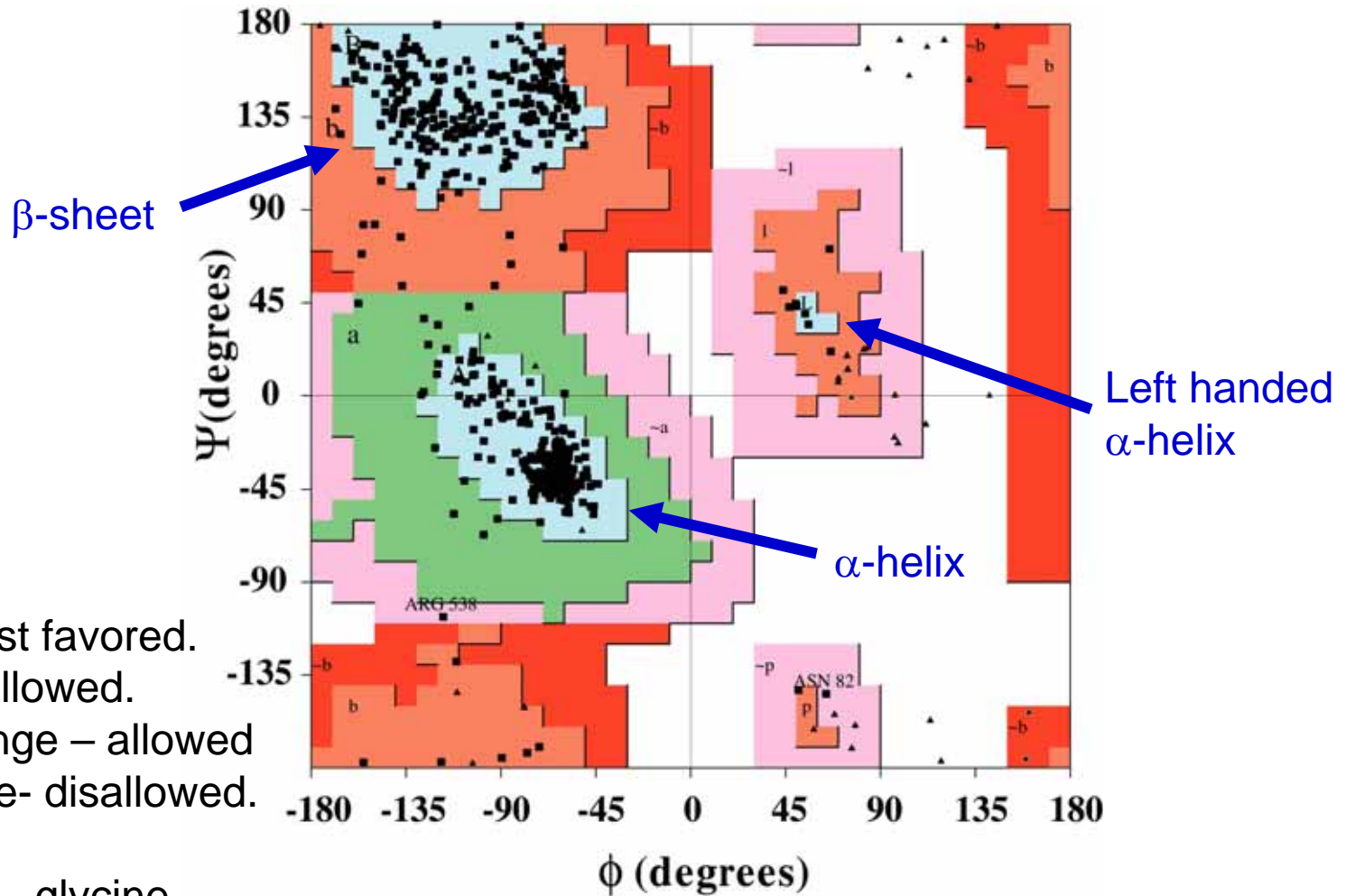
Geometry – the model must have reasonable bonds lengths, bond angles and overall geometric agreement compared to other well-defined structures. Bonds <0.01 Å and angles < 2.0° rms difference from ideal values.

B-factors (temperature factors) – an indicator of thermal vibration or disorder (static or conformational). Scale of 1 to 400 Å² with good values 20 to 35 Å².

Luzatti Plots – provide an estimate of the model coordinate error. Values depend on resolution. Typical values for good structures are 0.15 to 3 Å error.

Ramachandran Plots – define whether or not the main chain dihedral angles fall into spatially allowed conformational regions. >99% in allowed region with 90% or more in most favored regions.

The Ramachandran Plot



Cyan- most favored.
Green – allowed.
Light Orange – allowed
Pink, white- disallowed.

Triangles – glycine.

Other Methods: Fourier Analysis

- Coordinates (XYZ) of a known structure can phase a new set of experimental amplitudes (F_{obs}) provided the model and unknown structures are isomorphous.
- Atomic coordinates, unit cell, and space group are nearly identical.
- The new electron density map may be used to identify the location of ligand binding such as an enzyme substrate or metal.
- The position of the new ligand or atom can be found by *Fourier Analysis* in which the amplitudes for the electron density map are:

$$\rho(X, Y, Z) = (1/V) \sum_h \sum_k \sum_l F_{hkl}^{\text{new}} e^{2\pi i \alpha_{hkl} \text{model}} e^{-2\pi i (hX_j + kY_j + lZ_j)}$$

- During refinement the ligand is modeled and any other changes corrected so that the final phases come from the newly refined model.
- Fourier analysis can be useful to locate the positions of heavy atom derivatives once the initial heavy atom site(s) are known.

Fourier Coefficients for Model Building into Electron Density Maps

- *Fourier Synthesis* can be calculated with numerous coefficients (amplitudes).
- During the process of model building, the electron density equation is often calculated with coefficients:

$(3F_o - 2F_c)$ – Used when a significant portion of the model is missing.

$(2F_o - F_c)$ – Used to optimize the agreement between observed data and that from the model. (Most frequently used map).

$(F_o - F_c)$ – designed to check for negative or positive deficiencies due to parts of the model that are not observed in the experimental data (hole), or parts of the experimental data that cannot be accounted for by the model (peak).

- Model phases (α_{calc}) are used with these coefficients in the later stages.

Anomalous Differences as Fourier Coefficients

- The AS signal is measured as the difference between $\mathbf{F}_H^{\lambda n+}$ and $\mathbf{F}_H^{\lambda n-}$ reflections.
- These are collected systematically as hkl and $-h-k-l$.
- A non-AS data set $\mathbf{F}_H^{\lambda 1+}$ is not required for the phase calculation.
- A heavy atom position must be known from difference Patterson methods to converge on a unique solution for ϕhkl .
- The differences can be phased by light atoms (XYZ) from a known structure to identify an element (α is corrected to $\alpha-90^\circ$).
- The differences can be used in a Patterson Synthesis to locate a heavy atoms (Single Anomalous Dispersion or SAD).