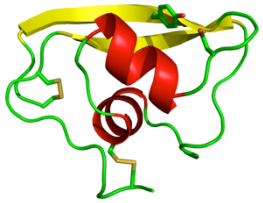


Central concepts in (bio-)crystallography

3. Patterson function and method

Jeroen Mesters
UNIVERSITÄT ZU LÜBECK
STIFTUNGSUNIVERSITÄT
SEIT 2015

Electron density

Fourier pair of equations:

$$\underline{\mathbf{F}}_{hkl} = V \int_{xyz} \rho(xyz) \exp^{[2\pi i\cdot(hx + ky + lz)]} dxdydz$$

$$\rho(xyz) = 1/V \sum_{hkl} |\underline{\mathbf{F}}_{hkl}| \exp^{[-2\pi i\cdot(hx + ky + lz) + i\alpha(hkl)]}$$

For $\underline{\mathbf{F}}_{hkl}$ and $\underline{\mathbf{F}}_{\underline{hkl}}$ the amplitudes are the same but the phase angles are of opposite value: so-called Friedel or Bijvoet pairs. The electron density equation reduces to:

$$\rho(xyz) = {}^{1}/_{V}|\underline{\mathbf{F}}_{000}| + {}^{2}/_{V} \sum\nolimits_{+hkl} |\underline{\mathbf{F}}_{hkl}| \cos[2\pi\cdot(hx + ky + lz) - \alpha_{hkl}]$$

* Note that $e^{ix} = \cos(x) + i\sin(x)$.

Crystallographic Phase Problem

$$\rho(xyz) = {}^{1}/_{V}|\underline{\mathbf{F}}_{000}| + {}^{2}/_{V}\Sigma_{+h}\Sigma_{k}\Sigma_{l}\,|\underline{\mathbf{F}}_{hkl}|\cos[2\pi\cdot(hx+ky+lz)-\alpha_{hkl}]$$

How to obtain information about α?

1) Isomorphous crystal structure (difference-Fourier synthesis)
2) Molecular replacement technique (MR)
3) Multiple isomorphous replacement (MIR)
4) Multi-wavelength anomalous diffraction (MAD)

) Patterson

5) Ab initio

Patterson map

- We can still calculate a Fourier summation with intensities as coefficients and all α_{hkl} equal to zero, *i.e.* a Patterson map.

- $P(uvw) = 1/V \cdot \sum_{hkl} |F_{hkl}|^2 \cdot \exp^{[-2\pi i \cdot (hu+kv+lw)]}$
  $= 2/V \cdot \sum_{+hkl} |F_{hkl}|^2 \cdot \cos[2\pi \cdot (hu+kv+lw)]$

  $P(\underline{\mathbf{u}}) = 2/V \cdot \sum_{S} |F_{\underline{\mathbf{S}}}|^2 \cdot \cos[2\pi \cdot \underline{\mathbf{u}} \cdot \underline{\mathbf{S}}]$

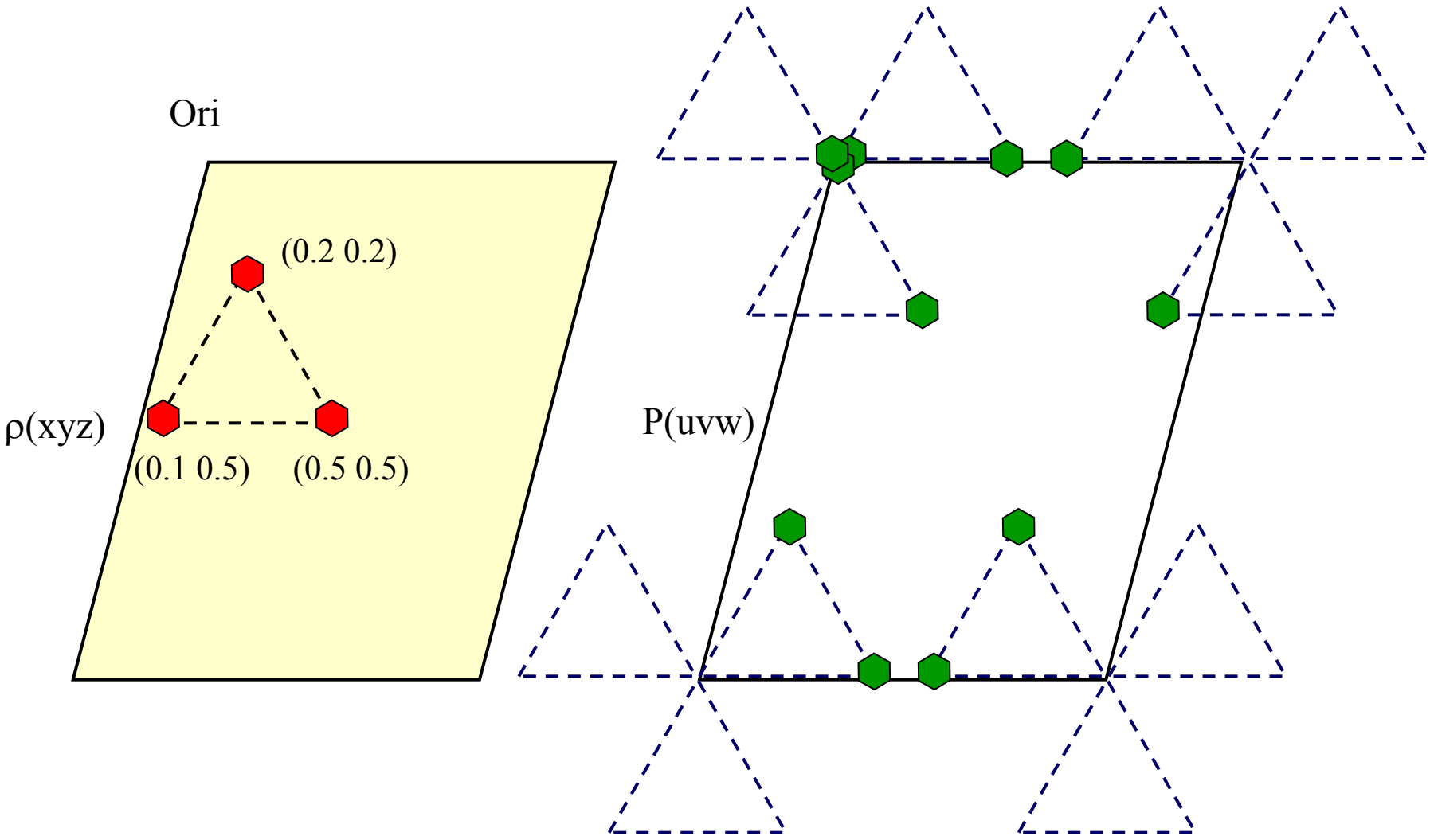
  is equivalent to

  $= \int_{\mathbf{v}} \rho(xyz) \cdot \rho(x+u\ y+v\ z+w)\ dV$

  $= \int_{\mathbf{r}} \rho(\underline{\mathbf{r}}) \cdot \rho(\underline{\mathbf{r}}+\underline{\mathbf{u}})\ dV$

* Note that $e^{ix} = \cos(x) + i\sin(x)$.

$\int_V \rho(x,y,z)\cdot\rho(x+u,y+v,z+w)dV$



Put each corner of the triangle (atoms of the structure) at the origin!

Calculate structure

1. In order to solve the 2D-structure, we must somehow combine the 6 interatomic vectors [$\mathbf{x}_1$-$\mathbf{x}_2$, $\mathbf{x}_1$-$\mathbf{x}_3$, $\mathbf{x}_2$-$\mathbf{x}_1$, $\mathbf{x}_2$-$\mathbf{x}_3$, $\mathbf{x}_3$-$\mathbf{x}_1$ and $\mathbf{x}_3$-$\mathbf{x}_2$] with the coordinates of the peaks in the map [(0.3 0.3) , (0.9 0.3) , (0.4 1.0) , (0.6 1.0) , (0.1 0.7) and (0.7 0.7)].

2. Find a set of three representative equations and choose three peaks (sum equal to 0 or n, remember periodicity):

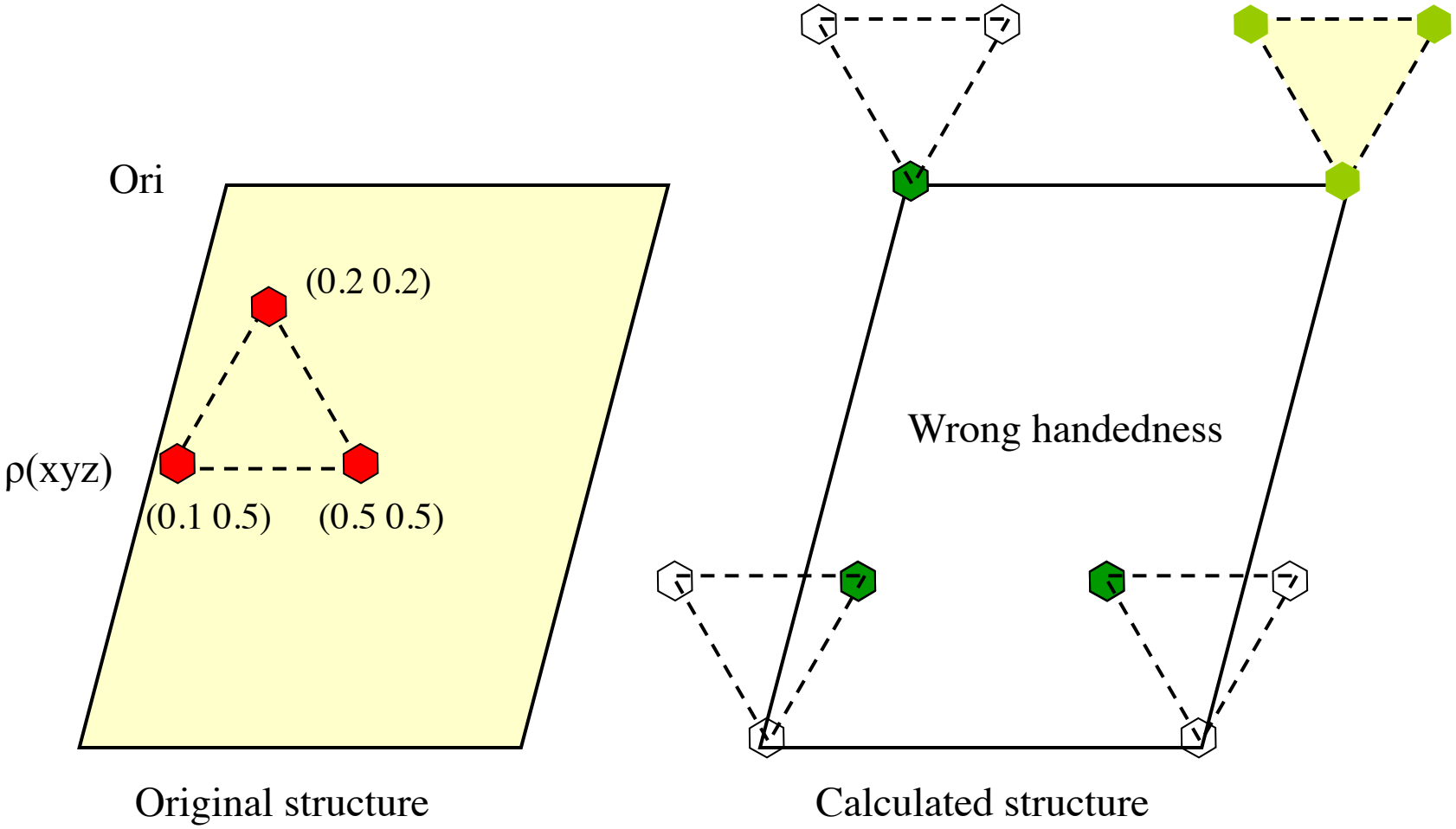
$$\begin{array}{lcl} \mathbf{x}_1\text{-}\mathbf{x}_2 & = & (0.9\ 0.3) \\ \mathbf{x}_2\text{-}\mathbf{x}_3 & = & (0.4\ 1.0) \\ \mathbf{x}_3\text{-}\mathbf{x}_1 & = & (0.7\ 0.7) \quad + \\ \hline 0 & & 2 \quad 2 \end{array}$$

Can be solved if we choose the origin! x_1 = (0 0)

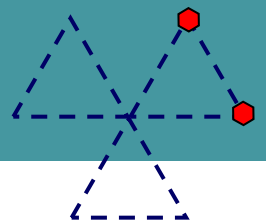
3. With x_1 (0 0) it follows that x_2 (0.1 0.7). With x_2 (0.1 0.7) it follows that x_3 (0.7 0.7). Finally, with x_3 (0.7 0.7) it follows that x_1 (0 0)! Indeed, our starting assumption. Remember periodicity: $0 \equiv 1$ or $-0.3 \equiv 0.7$!

Calculate structure

$$\int_V \varrho(x,y,z)\varrho(x+u,y+v,z+w)dV$$



Consequences



- Only inter-atomic vectors in real space show up as peaks in Patterson space since $P(uvw) = \int_v \rho(x,y,z)\cdot\rho(x+u,y+v,z+w)dV$
- The map of a real unit cell with N atoms will contain N^2-N peaks in Patterson space outside the origin. Origin will contain N peaks.
- Patterson maps contain an additional symmetry element (centro-symmetry).
- In simple structures with a limited number of atoms, the atomic positions can be derived fairly straightforward.
- You can obtain the inverted structure, going from L- to D-amino acids (due to centro-symmetry).
- The (correct) structure we obtain is exactly the same one but often with a different origin, *e.g.* structures with the same orientation but different origin give the same Patterson map.

Harker sections

Space group $P2_1$ (along y axis)
with operator 1 = (x, y, z) and operator 2 = (-x, y +1/2, -z)
and
vector 1-2 = (2x, -1/2, 2z) and vector 2-1 = (-2x, 1/2, -2z)

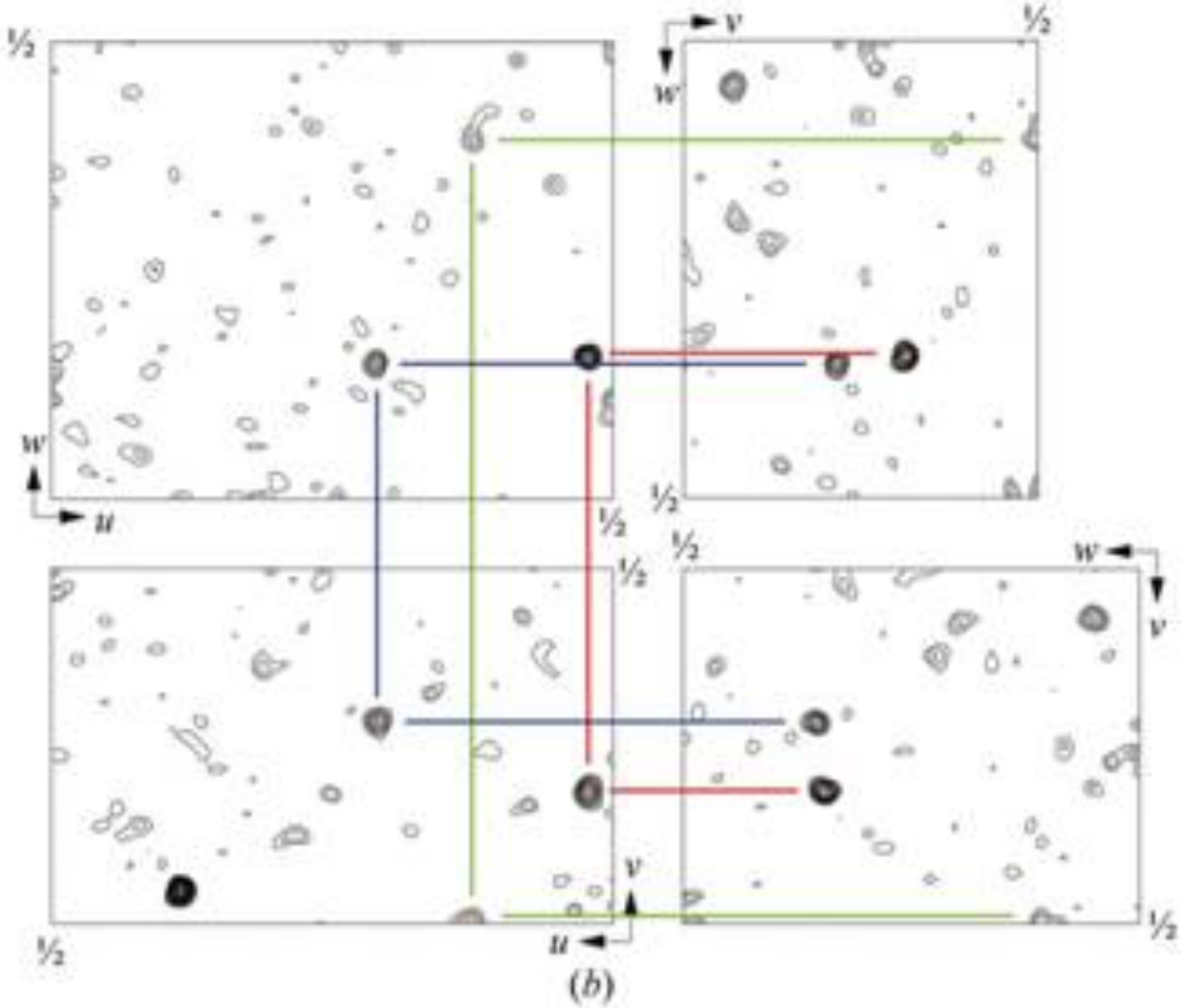
Subtracting operator 2 from operator 1, will give a peak at (u, v, w) = (2x, 0.5, 2z) and the opposite direction yields the centro-symmetric peak, (u, v, w) = (-2x, 0.5, -2z).

Notice that the symmetry operator that takes all values of y and translates them half a cell along y makes it such that all non-origin peaks generated by crystallographic symmetry in $P2_1$ Patterson maps are at the v = 0.5 section.

These special symmetry related sections are called **Harker sections**.

These Harker sections are always the first sections examined in Patterson solutions.

Harker sections ; $P2_12_12_1$



x,y,z
-x+0.5,-yz+0.5
-x, y+0.5,-z+0.5
x+0.5,-y+0.5,-z

http://research.uni-leipzig.de/straeter/practical/NTPDase2.html

Simple Fourier Synthesis

Homologous structures (variants and/or complexed with small compound) can be calculated straightforward only if space group and unit cell dimensions are equivalent (so-called **isomorphous**!).

PDB file (atomic coordinates, *e.g.* electron density distribution)

$$\underline{\mathbf{F}}(hkl) = V \int_x \int_y \int_z \rho_{model}(xyz) \exp^{[2\pi i \cdot (hx + ky + lz)]} dxdydz$$

⇓⇓⇓ Fourier analysis

$$|\underline{\mathbf{F}}_{model}(hkl)| \text{ and } \alpha_{model}(hkl)$$

⇓⇓⇓

Combine $\alpha_{model}(hkl)$ with $|\underline{\mathbf{F}}_{observed}(hkl)|$

⇓⇓⇓ Fourier synthesis

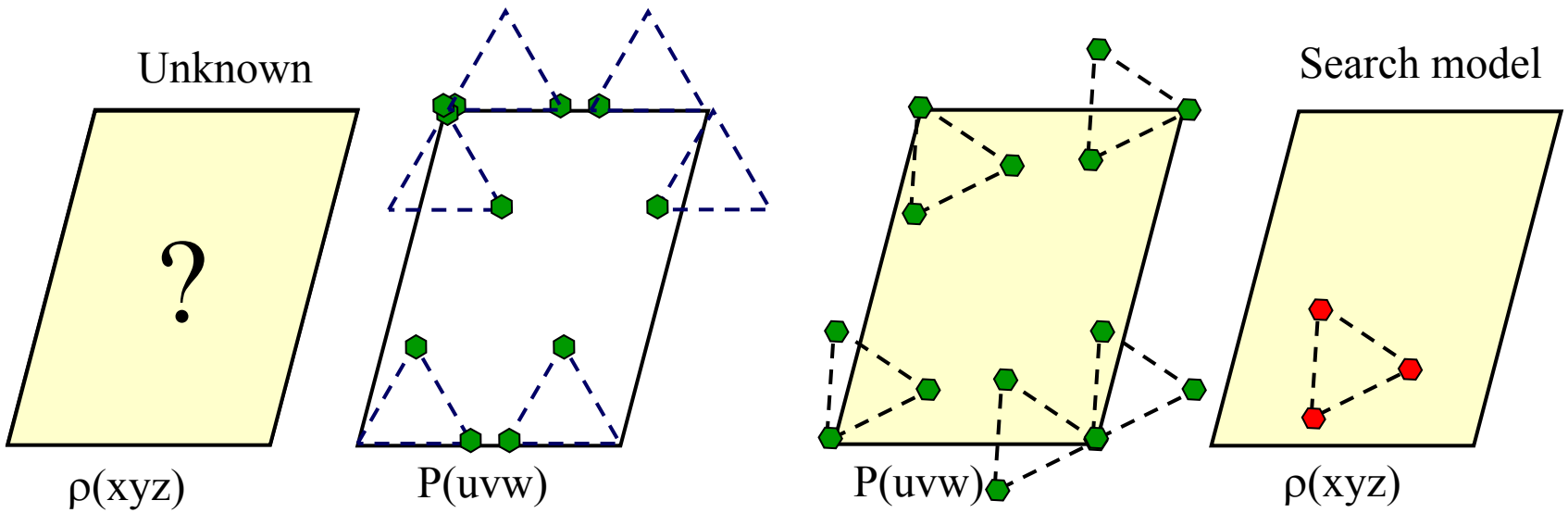
$$\rho(xyz) = {}^{1}/_{V}|\underline{\mathbf{F}}_{000}| + {}^{2}/_{V}\Sigma_{+hkl} |\underline{\mathbf{F}}_{o}(hkl)| \cos[2\pi \cdot (hx+ky+lz) - \alpha_{m}(hkl)]$$

Molecular replacement

What if we do not have an isomorphous structure? Does it help if a homologous structure were available?

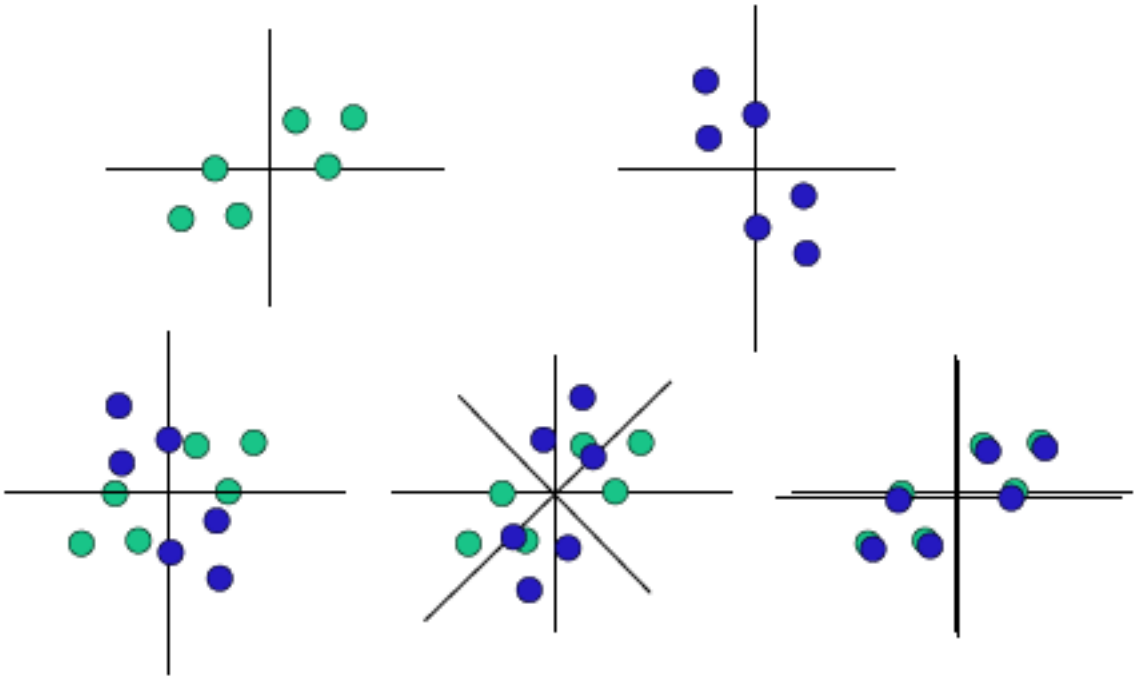
Yes! We can apply the „molecular replacement“ technique that is based on Patterson approach. In this way we create an isomorphous structure ourselves.

Molecular replacement: Rotation



- Superimpose (rotate and translate) the search model.
- Calculate $F_c(hkl)$ and $\alpha_c(hkl)$ from the superimposed structure.
- Combine $\alpha_c(hkl)$ and $F_o(hkl)$, calculate $\rho(xyz)$ and start model refinement cycles.

Molecular replacement: Rotation



Molecular replacement: Rotation

- An overlap function R of P($\underline{\mathbf{u}}$) with the rotated version, $P_r(\underline{\mathbf{u}}_r)$, of the same crystal lattice (self-rotation function) or a different crystal lattice (cross-rotation function) is defined as

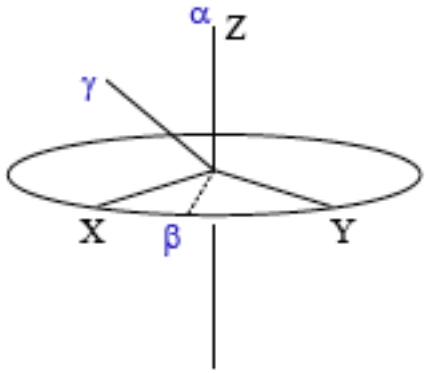
$$R(\alpha,\beta,\gamma)= \int_{\mathbf{u}} P(\underline{\mathbf{u}})\cdot P_r(\underline{\mathbf{u}}_r)dv$$

- The function R depends on the rotation angles and will be maximal for the correct overlap!

Eularian angles α, β and γ

Rossmann & Blow, 1962:

1. rot. by α around Z
2. rot. by β around new X'
3. rot. by γ around new Z'



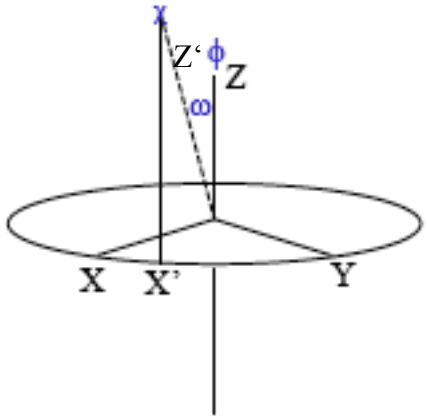
Machin, 1985:

2. rot. by β around new Y'

Symmetry of rotation function clearly shows up.

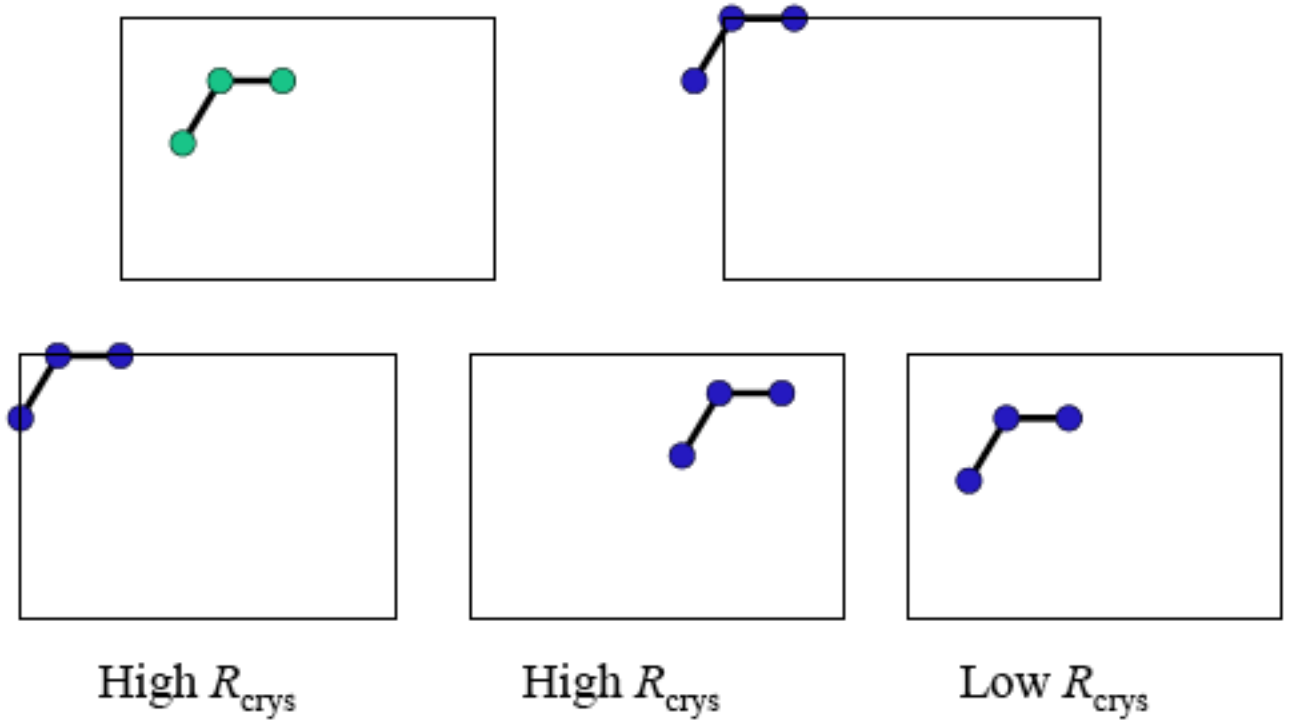
Polar angles χ, ω and ϕ

1. rot. by ϕ (phi) around Z
2. ω (omega) is angle between ZOZ ‘ in plane ZOX’
3. rot. by χ (chi) around new Z’



Convenient for self-rotation function (χ can be restricted to 120º or 180º).

Molecular replacement: Rotation



$$\underline{\mathbf{F}}_{hkl} = \sqrt{I_{hkl}}$$

$$\underline{\mathbf{F}}_{hkl} = V \int_{xyz} \rho(xyz) \exp^{[2\pi i \cdot (hx + ky + lz)]} dv$$

Molecular replacement: Translation

After the rotation, we must perform a translation by simple trial and error. The known molecule is moved through the asymmetric unit and structure factors are calculated and compared with the observed structure factors (Fobs) by calculating the R-factor

$$R = \Sigma_{hkl} (| F_{obs} | - k \cdot | F_{cal} |) / \Sigma_{hkl} | F_{obs} | \qquad (k = \text{scaling factor})$$

or correlation coefficient

$$C = \{ \Sigma_{hkl} (| F_{obs} |^2 - < F_{obs} >^2) \cdot (| F_{cal} |^2 - < F_{cal} >^2) \} / \{ \Sigma_{hkl} (| F_{obs} |^2 - < F_{obs} >^2)^2 \cdot (| F_{cal} |^2 - < F_{cal} >^2)^2 \}^{-1/2}$$

with C being scaling insensitive.

Molecular replacement: Consequences

- Molecular replacement is a rotation and translation problem, *e.g.* maximization of overlap function $R(\alpha\beta\gamma)$ and minimization of R_{cryst} :
  1) $R(\alpha\beta\gamma) = \int_{\mathbf{u}} P(\underline{\mathbf{u}}) \cdot P_r(\underline{\mathbf{u}}_r)\, dv$ (self- or cross-rotation)
  2) $R_{cryst} = \sum_{hkl} (|F_{obs}| - k \cdot |F_{cal}|) / \sum_{hkl} |F_{obs}|$ (translation).
- Contains inter- and intra-molecular atom difference vectors. Best to restrict the length of vector $\underline{\mathbf{u}}$ to 2/3 to 3/4 of the molecule diameter.
- Low resolution data are rather insensitive to rotation and can be left out and high resolution data are more sensitive for the model. Best range between 3.5 and 6-8 Å.
- Beware, bias originating from the search model is introduced into the newly calculated structure!

Multiple Isomorphous Replacement

Prepare isomorphous heavy atom containing derivatives of the protein in the crystalline state. Collect a full dataset of the derivative crystal.

Calculate $P(uvw) = 1/V \cdot \Sigma_{hkl} (|F_{PH}| - |F_P|)^2 \cdot \cos[2\pi \cdot (hu+kv+lw)]$ and determine ($|F_H|$) and phases (α_H).

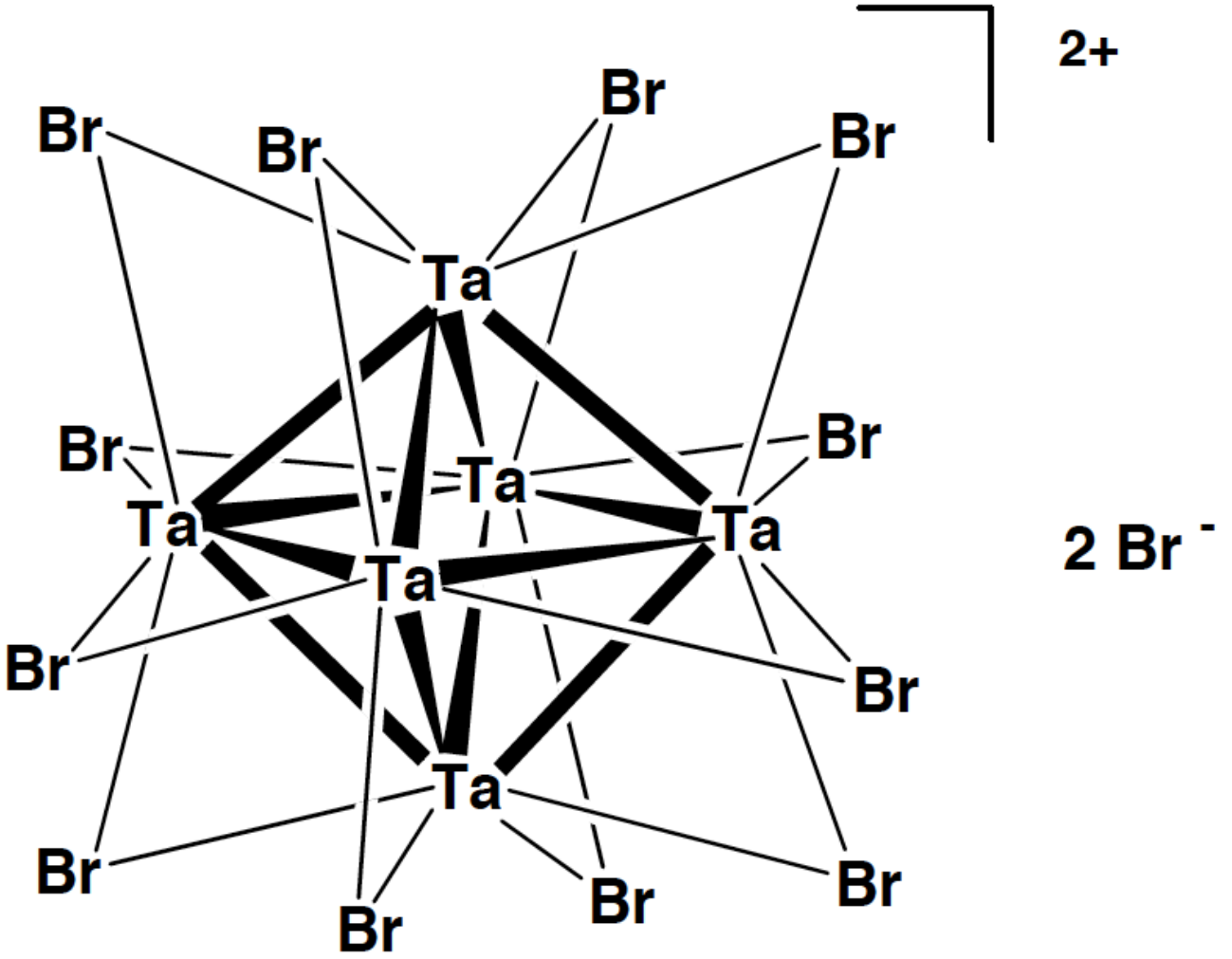
Can one heavy atom change the protein intensities? In principle yes: suppose we have a 42 kDa protein with 3000 non-hydrogen atoms

$\langle|F_P|\rangle \approx 7 \cdot \sqrt{3000} \approx 383$ (Average electrons C, N, O = 7)

$\langle|F_H|\rangle = 80$ for Hg

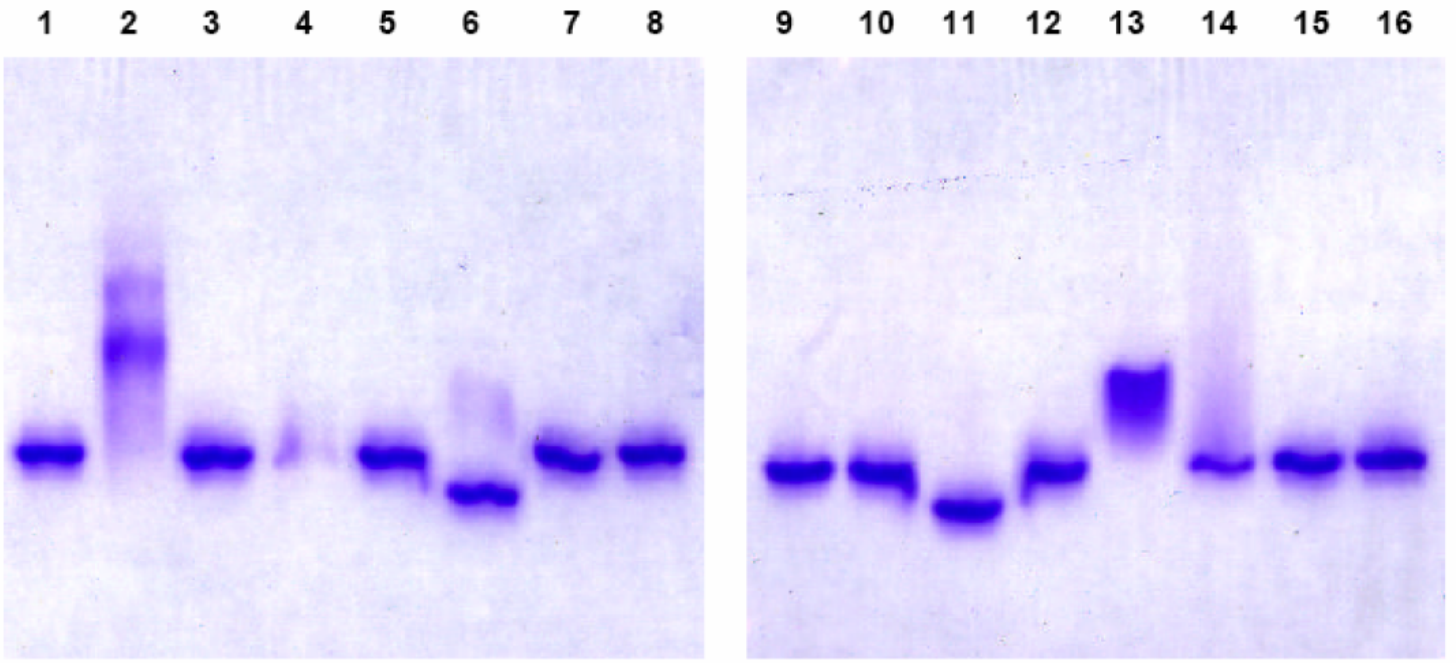
$\langle|F_H|\rangle / \langle|F_P|\rangle \approx 1/4$ although Mr(heavy)/Mr(protein) $\approx 0.21\%$

Tantalumbromide cluster



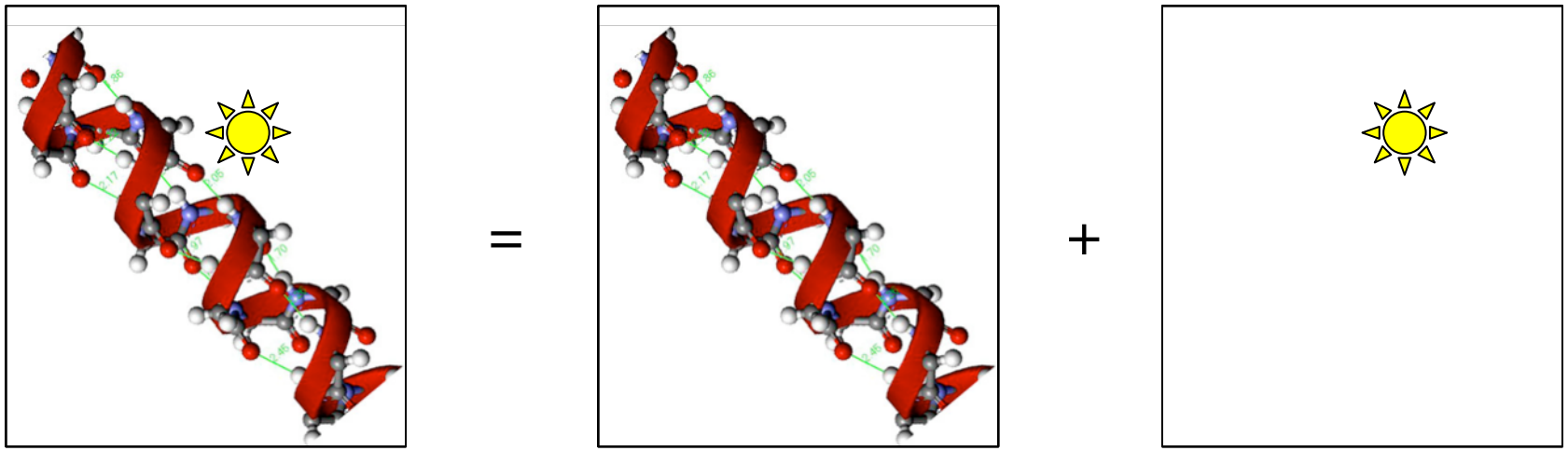
Heavy enough for ribosomes (2.5 MDa)

Native Band Gel-Shift Assay



1. KdsB Control
2. KdsB + K_2PtCl_4
3. KdsB + $KAu(CN)_2$
4. KdsB + $UO_2(C_2H_3O_2)$
5. KdsB + $(CH_3)_3Pb(C_2H_3O_2)$
6. KdsB + para-chloro mercuri phenyl sulfonic acid
7. KdsB + Thimerosal $CgHg_2O_2S$-Na
8. KdsB Control
9. KdsB Control
10. KdsB + K-perrhenate KO_4Re
11. KdsB + p-hydroxymercuri benzoate
12. KdsB + phenyl mercuric acetate $C_8H_8HgO_2$
13. KdsB + $HgCl_2$
14. KdsB + $SmCl_3$
15. KdsB + Pb acetate $Pb(C_2H_3O_2)_2$
16. KdsB Control

$\underline{F}_{PH} = \underline{F}_{P} + \underline{F}_{H}$



$$\underline{\mathbf{F}}_{hkl} = \int_x \int_y \int_z \{\varrho_{ph}(x,y,z)\} \cdot \exp^{[2\pi i \cdot (hx+ky+lz)]} \cdot dxdydz$$
$$= \int_x \int_y \int_z \{\varrho_{p}(x,y,z) + \varrho_{h}(x,y,z)\} \cdot \exp^{[2\pi i \cdot (hx+ky+lz)]} \cdot dxdydz$$

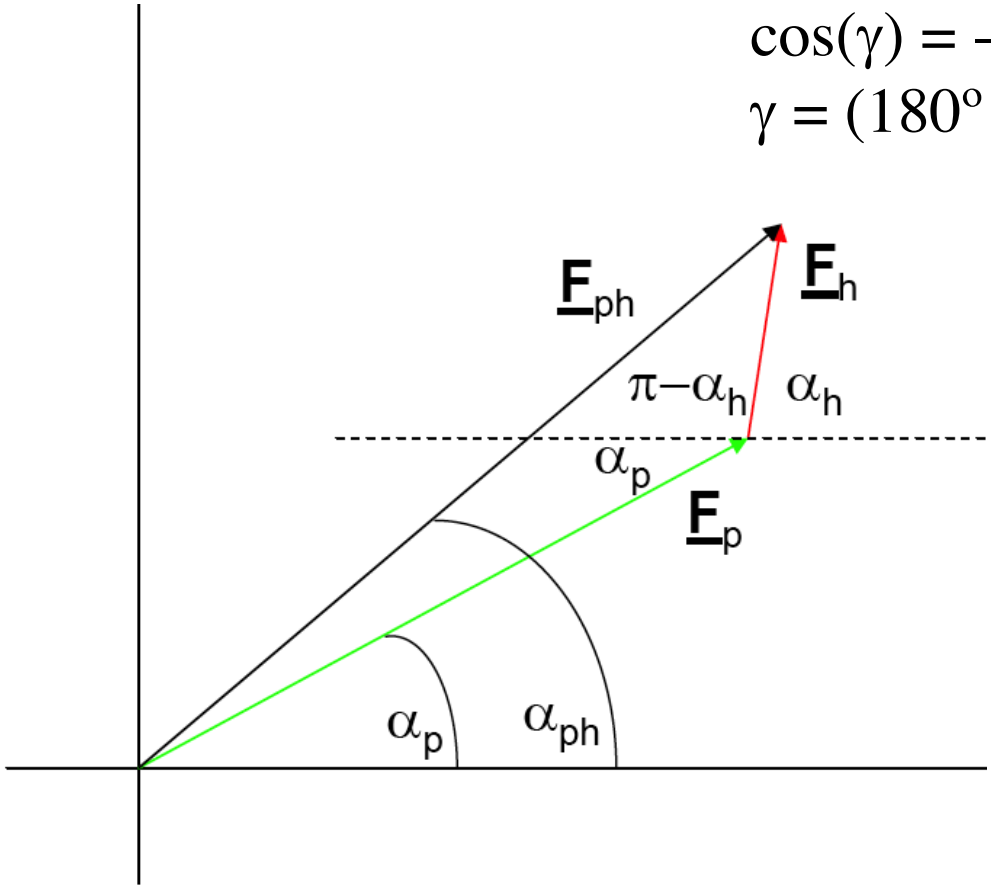
$$P(uvw) = 1/V \cdot \sum_{hkl} (|F_{ph}| - |F_{p}|)^2 \cdot \exp^{[-2\pi i \cdot (hu+kv+lw)]}$$

$\underline{F}_{PH} = \underline{F}_P + \underline{F}_H$

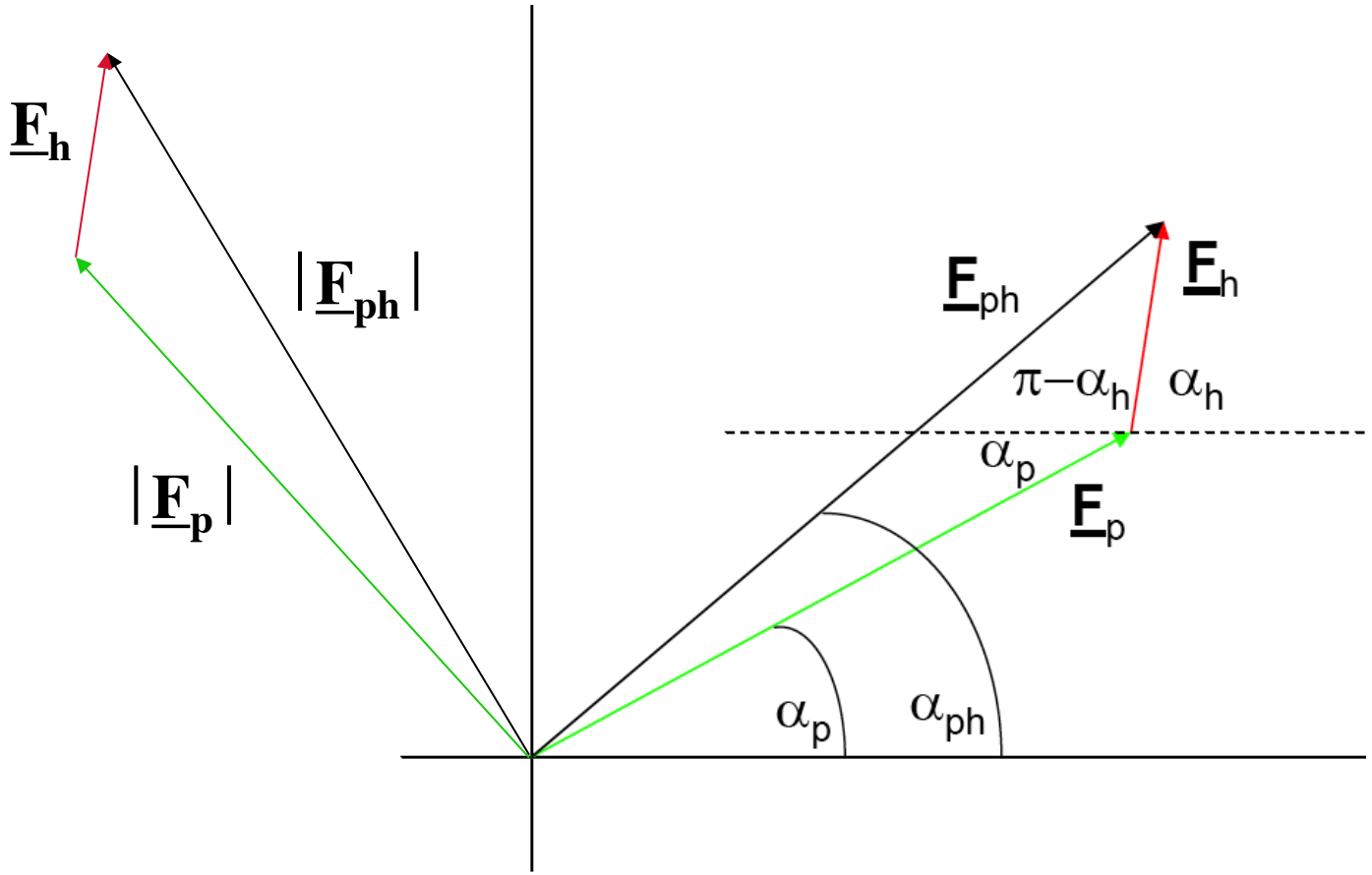
Cosine rule $c^2 = a^2 + b^2 - 2ab \cdot \cos(\gamma)$

$\cos(\gamma) = - (c^2 - a^2 - b^2) / 2ab$

$\gamma = (180^o - \alpha_H) + \alpha_P$

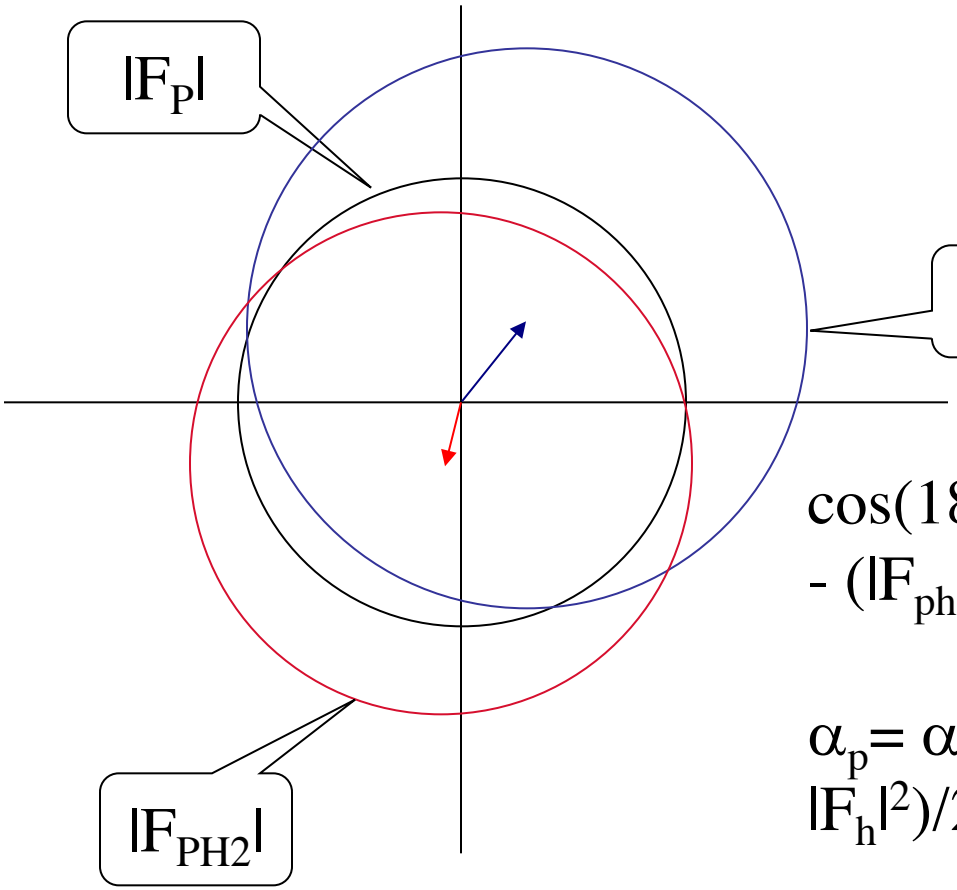


$\underline{F}_{PH} = \underline{F}_P + \underline{F}_H$; 2 solutions



Vector $\underline{F}_h$ can be determined by Patterson map analysis

$\underline{F}_{PH} = \underline{F}_P + \underline{F}_H$; Harker construction



$$\underline{\mathbf{F}}_{PH} = \underline{\mathbf{F}}_P + \underline{\mathbf{F}}_H$$

$\cos(180^o + \alpha_p - \alpha_h) = - \cos(\alpha_p - \alpha_h) =$
$- (|F_{ph}|^2 - |F_p|^2 - |F_h|^2) / 2 \cdot |F_p| \cdot |F_h|$

$\alpha_p = \alpha_h + \text{arc cos}(|F_{ph}|^2 - |F_p|^2 - |F_h|^2)/2(|F_p||F_h|)$

Two solutions for one derivative!

More derivatives will solve this ambiguous result.

Difference Patterson

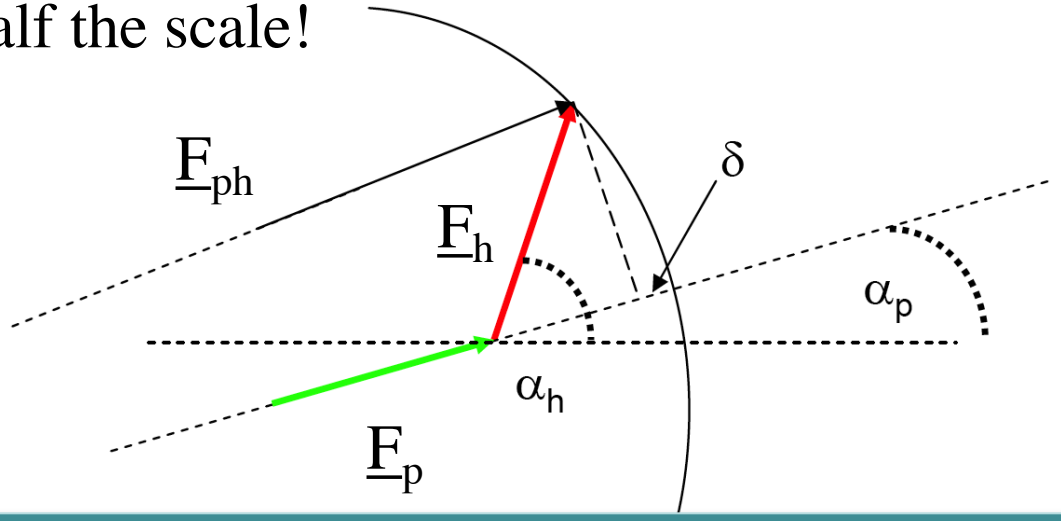
$P(uvw) = 1/V \cdot \sum_{hkl} |F_{hkl}| \cdot \exp^{[-2\pi i \cdot (hu+kv+lw)]}$

$P(uvw) = 1/V \cdot \sum_{hkl} (|F_{ph}| - |F_p|)^2 \cdot \cos[2\pi \cdot (hu+kv+lw)]$

$|F_{ph}| = |F_p| + |F_h| \cdot \cos(\alpha_h - \alpha_p) + \delta$ but if $|Fh| << |Fp|$ than $\delta \approx 0$

(δ for difference, in crystallography also termed ε as we shall see later)

With $(|F_{ph}| - |F_p|)^2 = |F_h|^2 \cdot \cos(\alpha_h - \alpha_p)$ and $\langle \cos^2(\alpha) \rangle \approx 0.5$
it can be seen that the difference Patterson is nothing else than a pure heavy atom Patterson on half the scale!



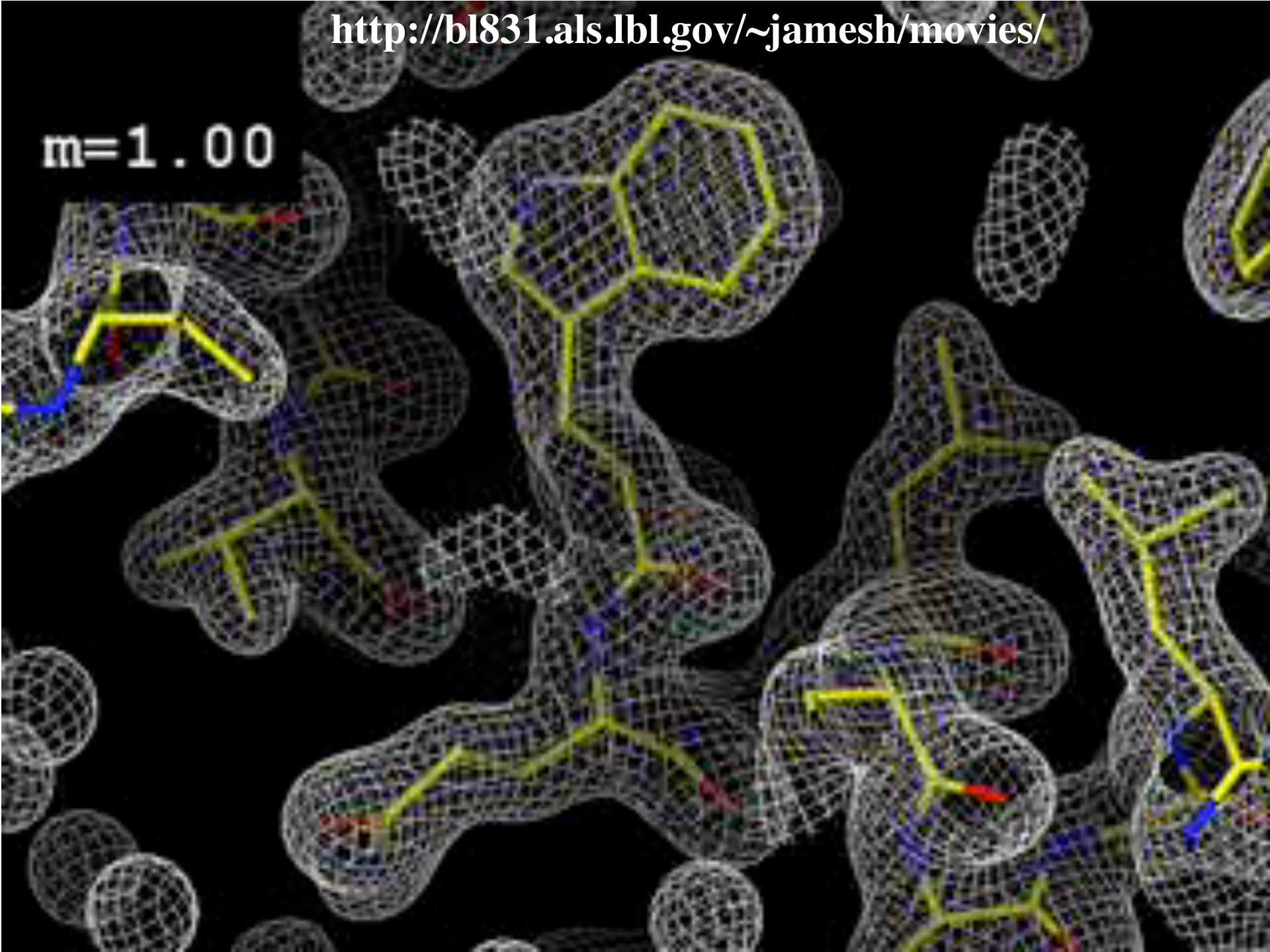
$\underline{F}_{PH} = \underline{F}_{P} + \underline{F}_{H}$; Lack of closure

- In practice when you have found the phases experimentally there is some mismatch:
- The mismatch is called the lack of closure & is given the symbol ε.
- From the mismatch you can estimate the phasing power

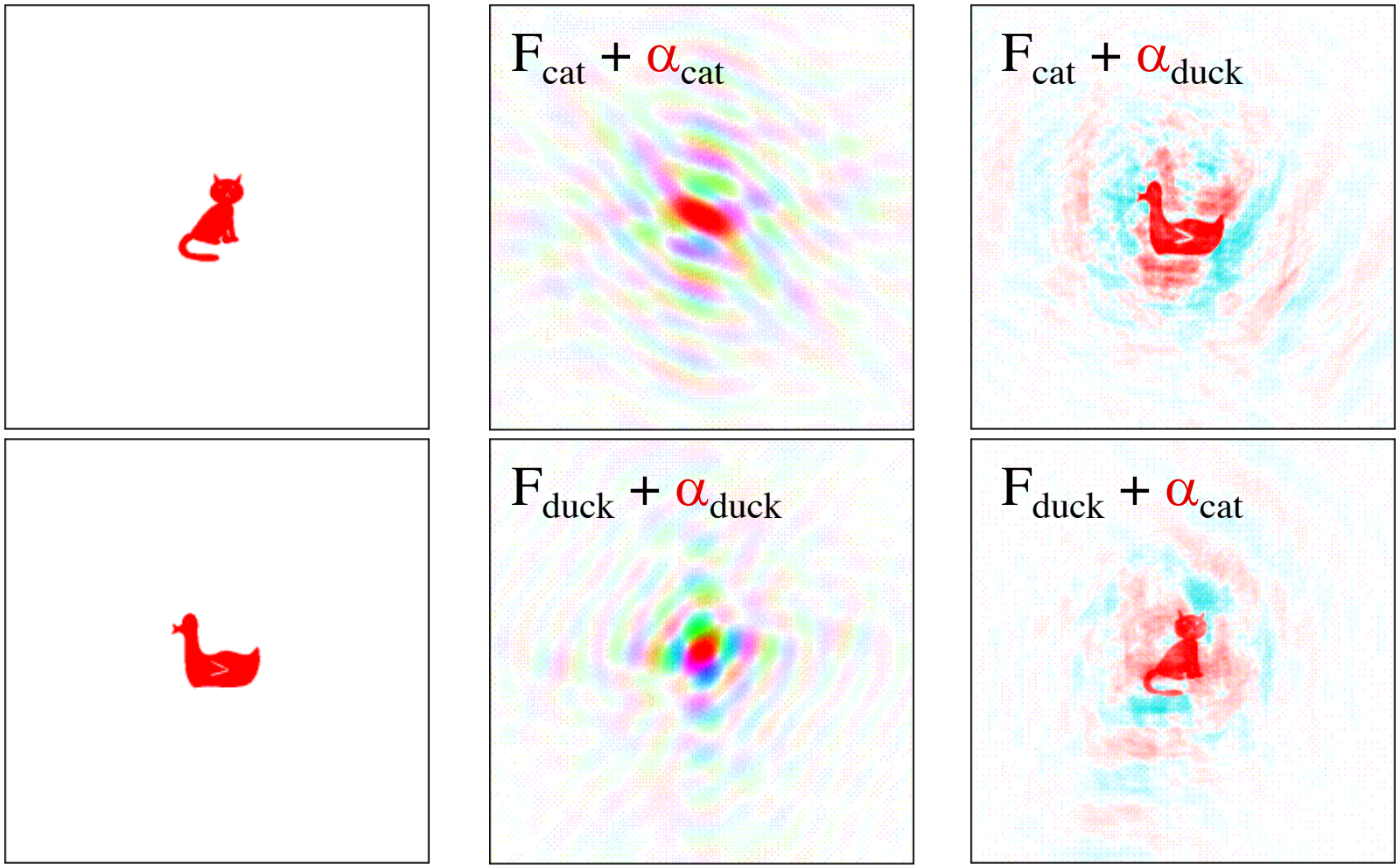
Phasing Power = $\sqrt{(\Sigma |F_H|^2 / \Sigma \varepsilon^2)}$

- A phasing power of 4 is excellent but rare
- A value between 1 & 2 is acceptable & means that the scattering of the heavy-atom is larger than the lack of closure.

$$\varepsilon = |F_H(h) + F_P(h)| - |F_{PH}(h)|$$


http://bl831.als.lbl.gov/~jamesh/movies/
m=1.00

|F(hkl)| versus α(hkl)



http://www.ysbl.york.ac.uk/~cowtan/fourier/duck1.html