

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/9042194>

The phase problem

Article in *Acta Crystallographica Section D Biological Crystallography* · December 2003

DOI: 10.1107/S09074444903017815 · Source: PubMed

CITATIONS

71

READS

51

1 author:



Garry L Taylor

University of St Andrews

129 PUBLICATIONS 5,145 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Molecular Visualization [View project](#)

All content following this page was uploaded by [Garry L Taylor](#) on 15 March 2015.

The user has requested enhancement of the downloaded file. All in-text references [underlined in blue](#) are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

The Phase Problem

Garry Taylor (2003) *Acta
Crystallographica* Section D**59**, 1881-
1890.

Paper Overview

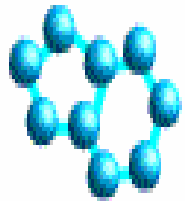
- What Phase Problem?
- How phases are derived from some prior knowledge of structure
- Overview of Phasing methods:
 - Direct, Molecular Replacement, and Heavy Atom Isomorphous Replacement

Crystallography Jargon

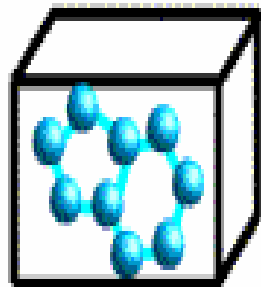
- Isomorphism: A close similarity in the crystalline structure of two or more substances of similar chemical composition
- Same unit cell, same orientation of protein in the cell

Crystal Packing

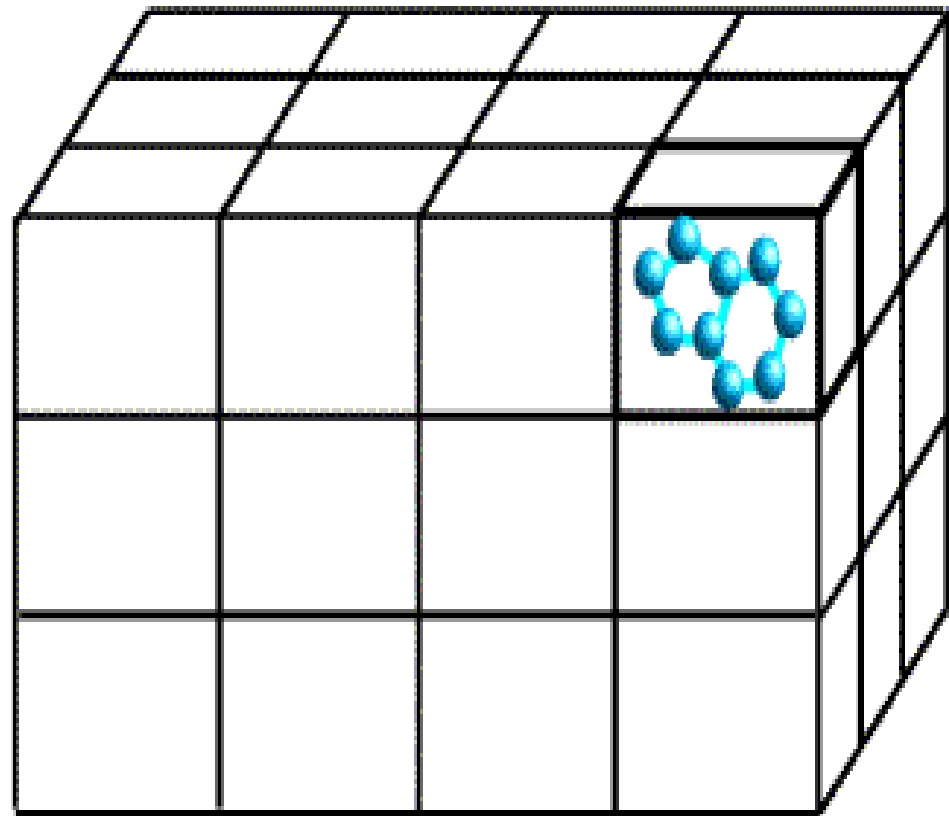
molecule



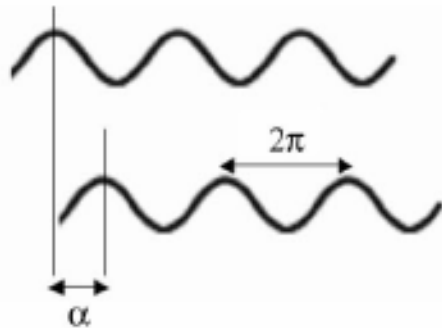
unit cell



crystal



Constructive and Destructive Interference

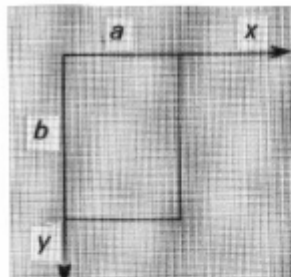


Wave I

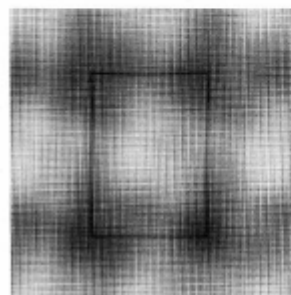
Consider 2 waves with different phase angles (α)

Wave II

(a)



Wave I

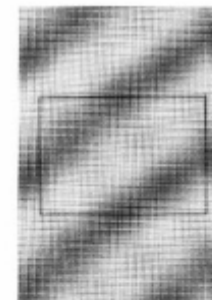
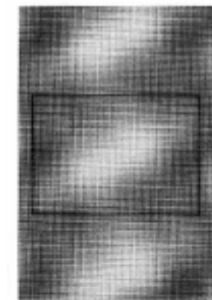


Waves I + II

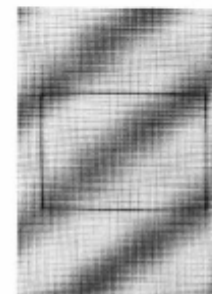
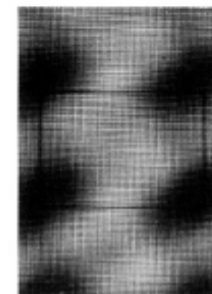
Now add a 3rd wave with a different phase angle

Wave III

Waves I + II + III



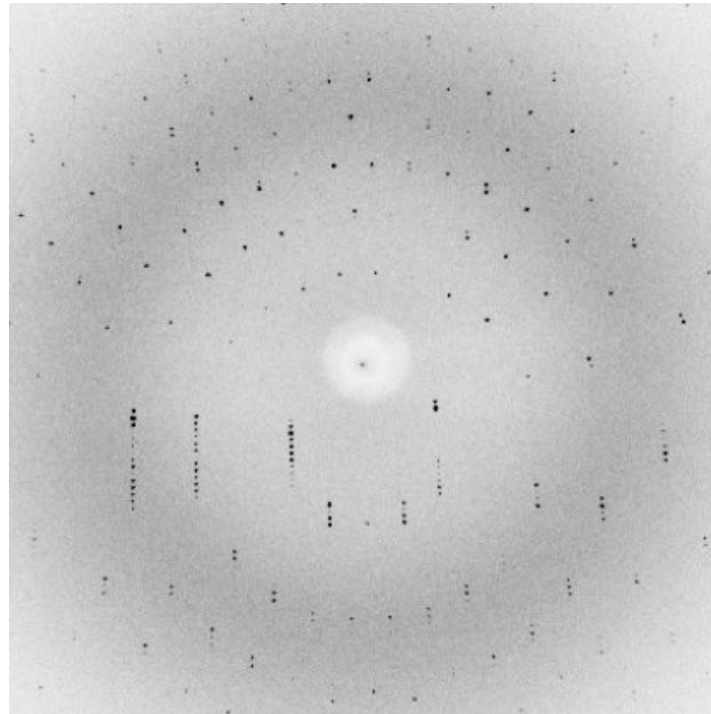
$\alpha = 0$



$\alpha = \pi$

Diffraction Pattern

- If the waves add up in phase in some directions, they have to cancel out in a lot of other directions. That is why the diffraction pattern from a crystal is an array of spots



Diffraction Experiments

- Measure the intensities of waves scattered from planes (hkl) within the crystal in order to calculate the electron density at position (xyz) in the unit cell

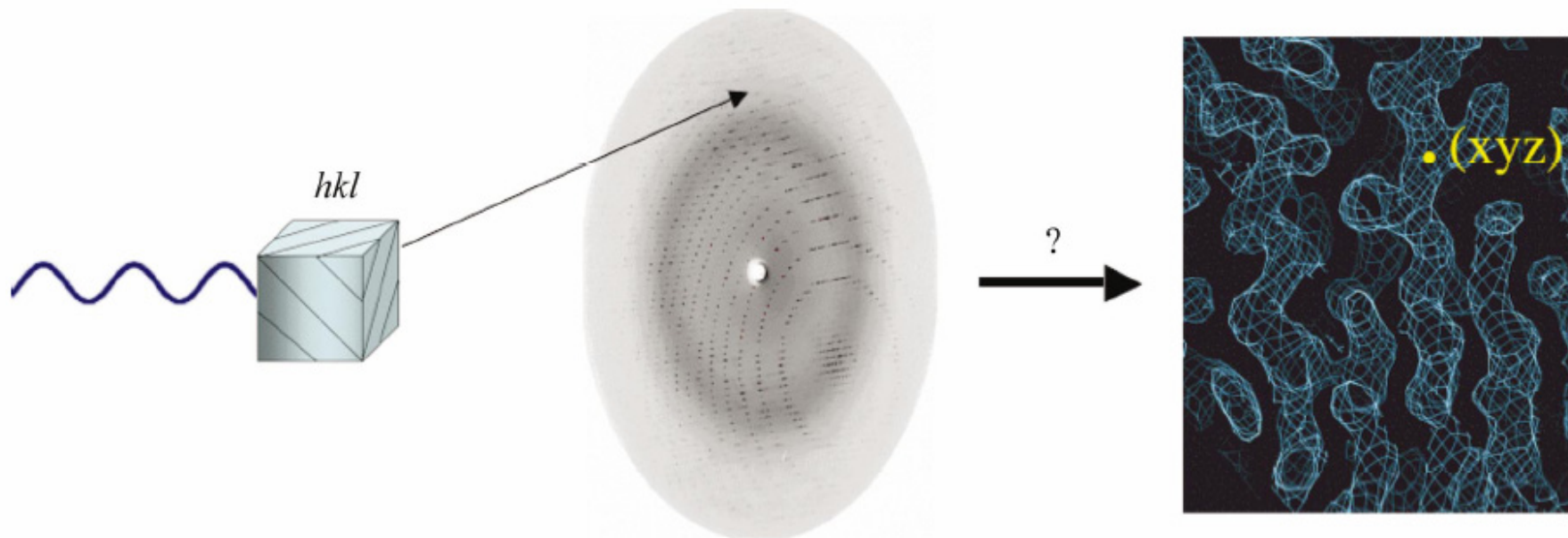


Figure 1
The diffraction experiment.

- **Perform the summation over all hkl planes**

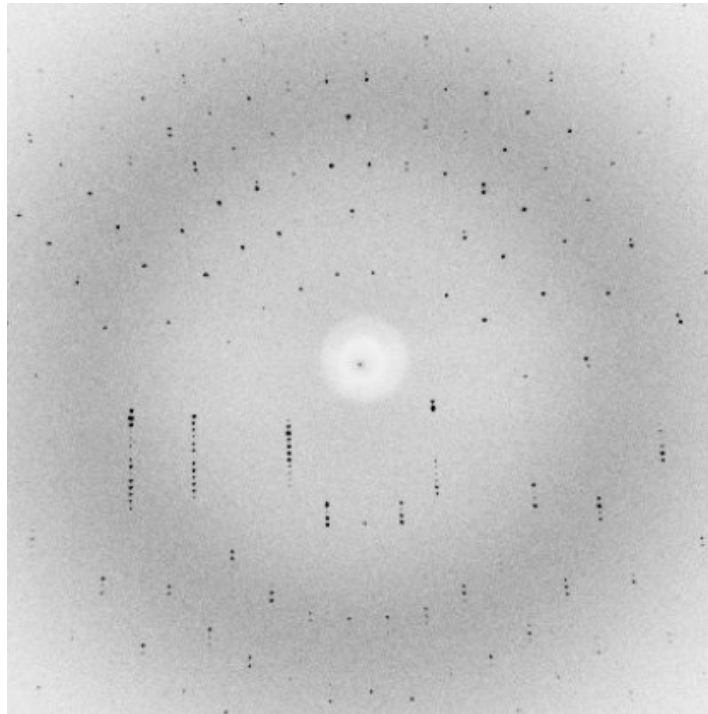
Calculation of Electron Density: Fourier Transform

- Electron density at (xyz) = the sum of contributions to the point (xyz) of waves scattered from the plane (hkl) whose amplitude depends on the number of electrons in the plane, added with the correct relative phase relationship.

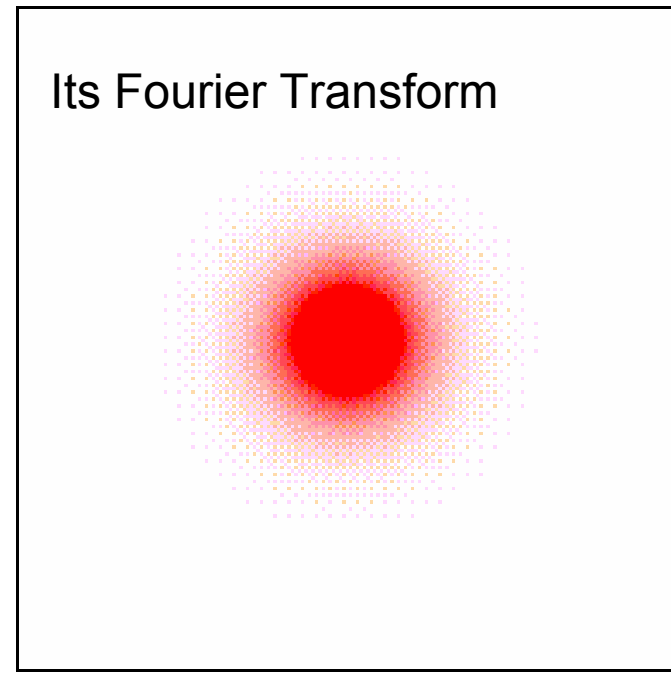
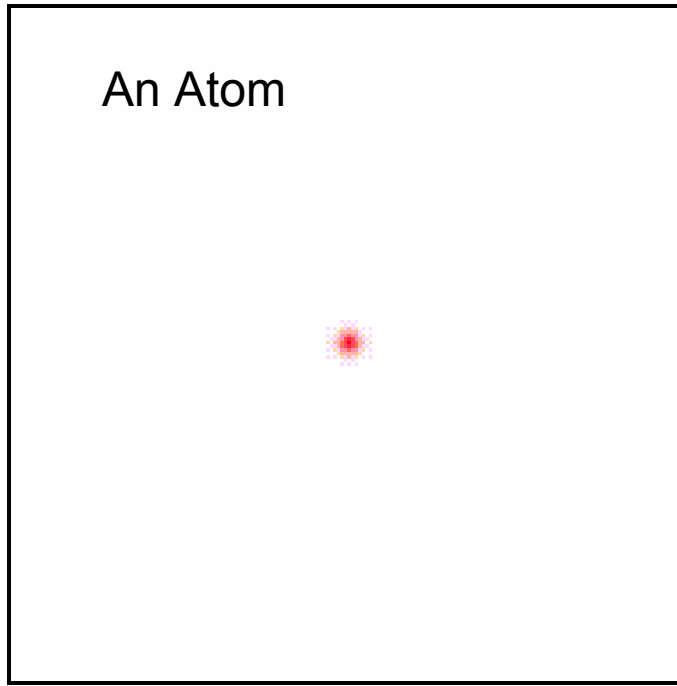
$$\rho(xyz) = 1/V \sum |F_{hkl}| \exp(i\alpha_{hkl}) \exp(-2\pi i h x + k y + l z)$$

Density (xyz)	=	1/Volume of unit cell	Structure Factor Amplitude (hkl)	α_{hkl} Phase associated with the structure factor amplitude	Wave scattered from the plane hkl
------------------	---	-----------------------------	---	--	--

Diffraction Patterns are the Fourier Transform of Whatever Object You're Shooting



$$\rho(xyz) = 1/V \sum |F_{hkl}| \exp(i\alpha_{hkl}) \exp(-2\pi ihx + ky + lz)$$

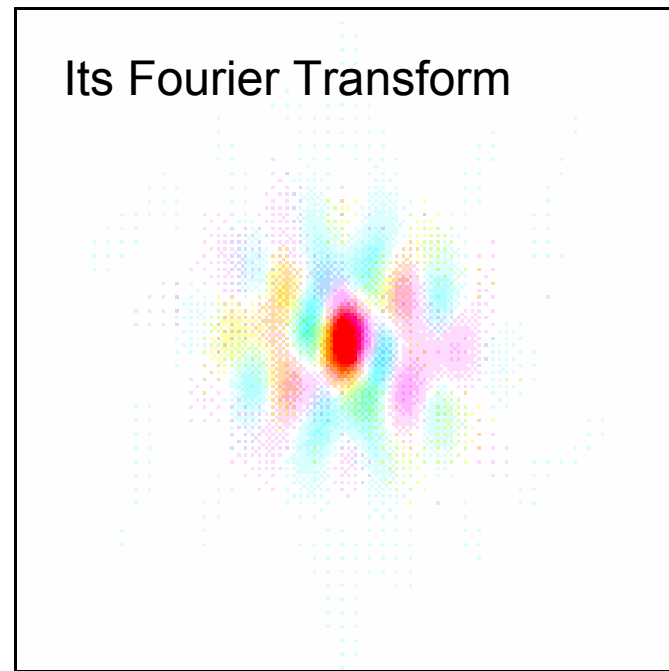
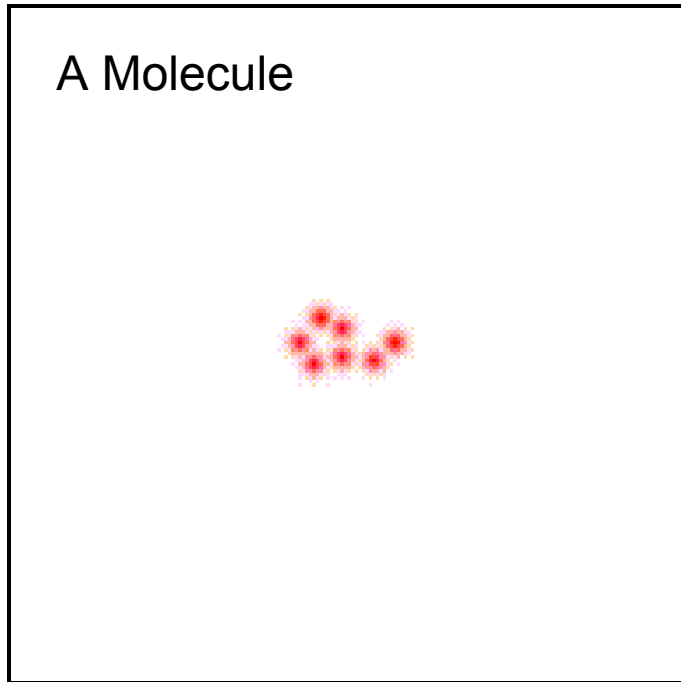


The atom is a sharp feature, whereas its transform is a broad smooth function

The Next Few Figures were taken from:

Kevin Cowtans Book of Fourier Transform:

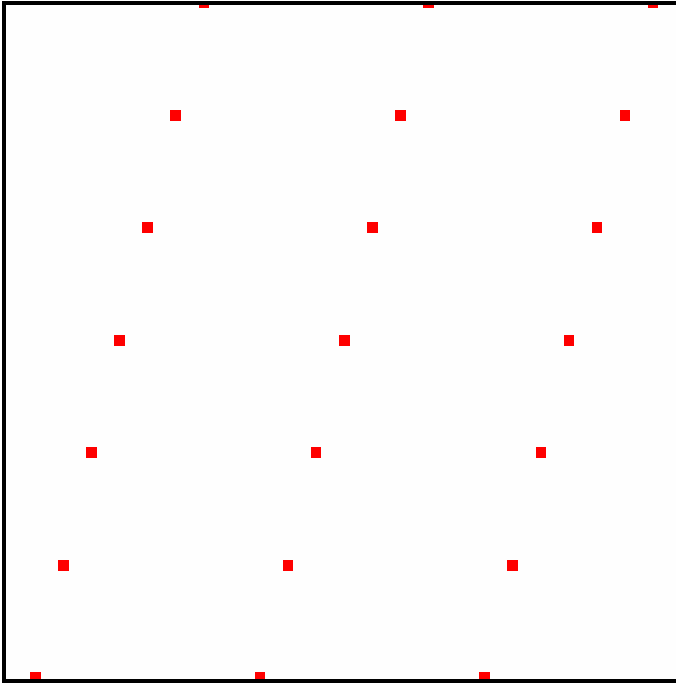
<http://www.ytbl.york.ac.uk/~cowtan/fourier/crys1.html>



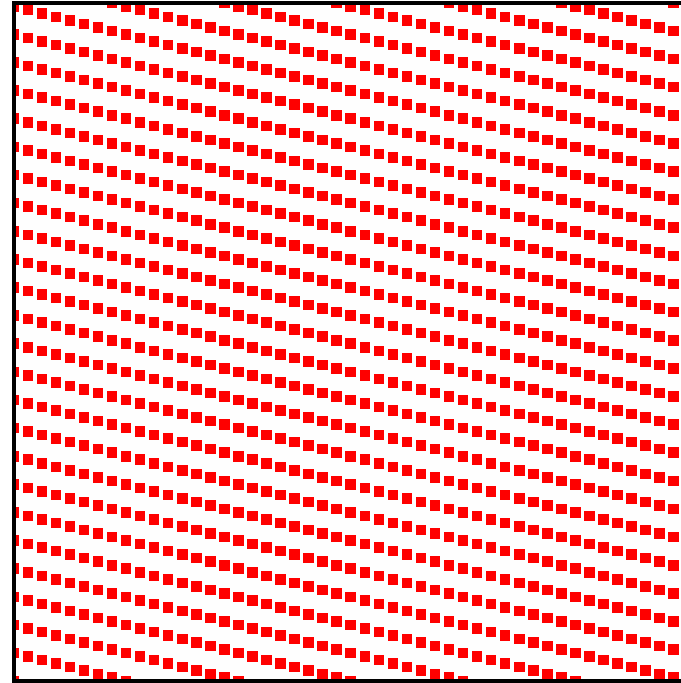
Transform shows some detail, but the overall shape is still that of the atomic transform

The molecule as the convolution of the *point atom structure* and the *atomic shape*. Thus, its transform is the product of the *point atom transform* and the *atomic transform*

A Grid Plane

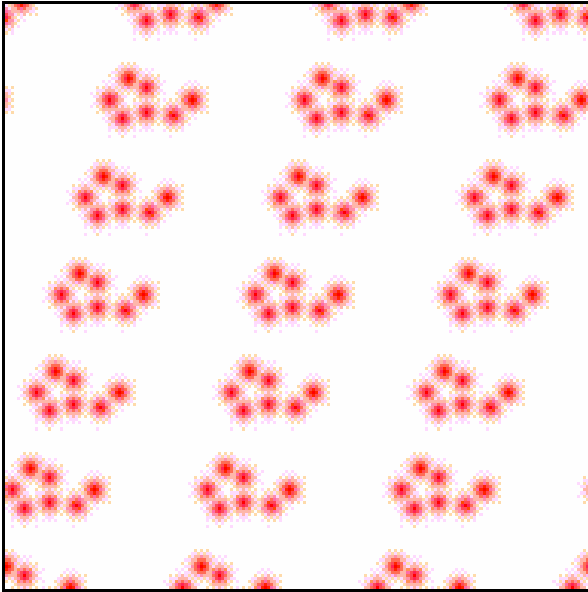


Its Fourier Transform

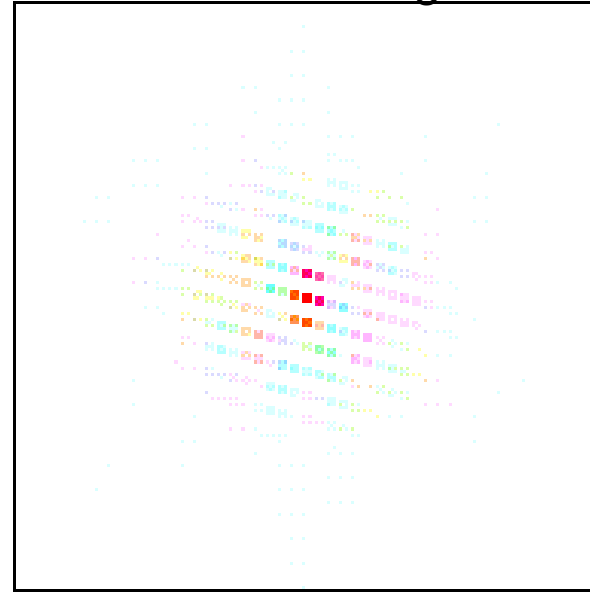


Note that the Fourier transform of a grid is a grid with reciprocal *directions* and *spacings*. This is the origin of the reciprocal lattice

A Molecule within a grid



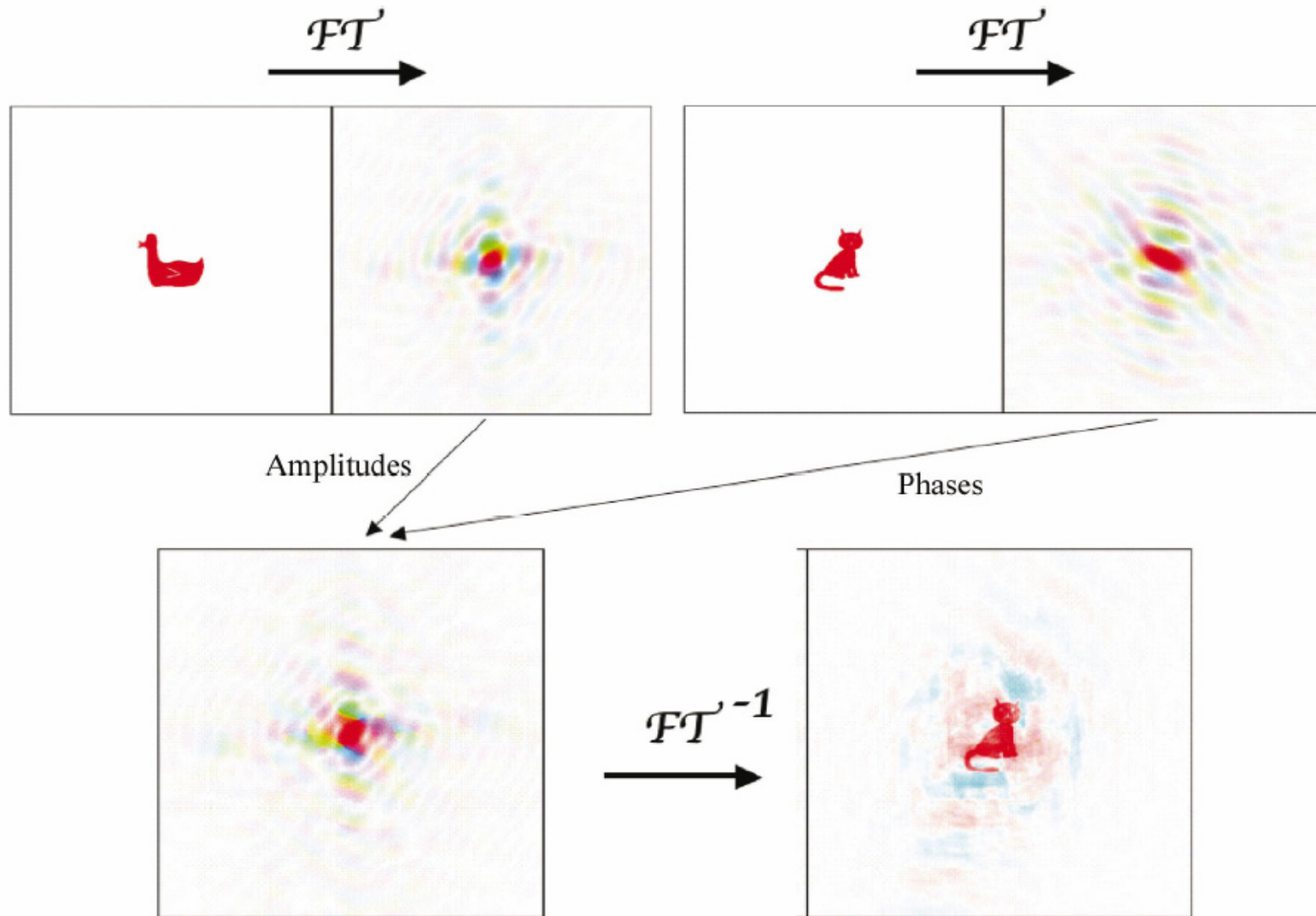
Fourier Transform of a Molecule within a grid



Finally, we build up a crystal by convoluting the *molecule* with the *grid*. The result is a crystal structure

The Fourier transform of the crystal is thus the product of the *molecular transform* and the *reciprocal lattice*. This is the *diffraction pattern*

The Importance of Phase in Carrying Information



Recovering the Phases

- There is no formal relationship between the amplitudes and phases; only via the molecular structure of electron density
- If we can assume some prior knowledge of the electron density or structure, this can lead to values for the phases
- Stated another way: phase values can only be discovered through some prior knowledge of the structure

Direct Methods

- Are based on the positivity and atomicity of electron density that leads to phase relationships between the normalized structure factors:

$$\alpha_{-\mathbf{h}} + \alpha_{\mathbf{h}'} + \alpha_{\mathbf{h}-\mathbf{h}'} \simeq 0,$$

$$\tan \alpha_{\mathbf{h}} = \frac{\langle E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} \sin(\alpha_{\mathbf{h}'} + \alpha_{\mathbf{h}-\mathbf{h}'}) \rangle_{\mathbf{h}'}}{\langle E_{\mathbf{h}'} E_{\mathbf{h}-\mathbf{h}'} \cos(\alpha_{\mathbf{h}'} + \alpha_{\mathbf{h}-\mathbf{h}'}) \rangle_{\mathbf{h}'}}$$

- Once the phases of some reflections are known, or can be given a variety of starting values, then the phases of other reflections can be bootstrapped
- *ab initio* phase determination are limited to $<1.2\text{\AA}$
- Are used in Shake-and-Bake, SHELXD, SHARP to find the heavy atom substructure

Molecular Replacement (MR)

- Requires a homology model
- Sequence identity >25% and <2.0Å RMSD α C atoms (are exceptions)
- Uses Patterson methods to Rotate and Translate model to the unknown structure

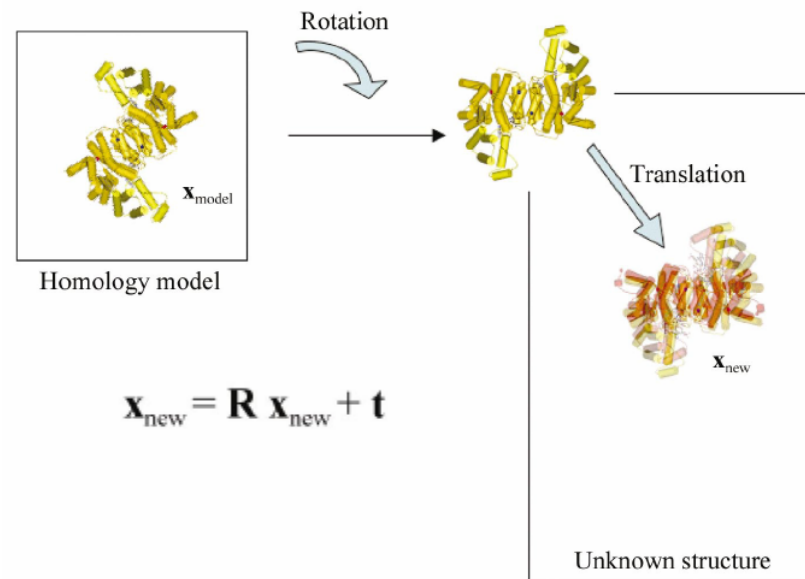
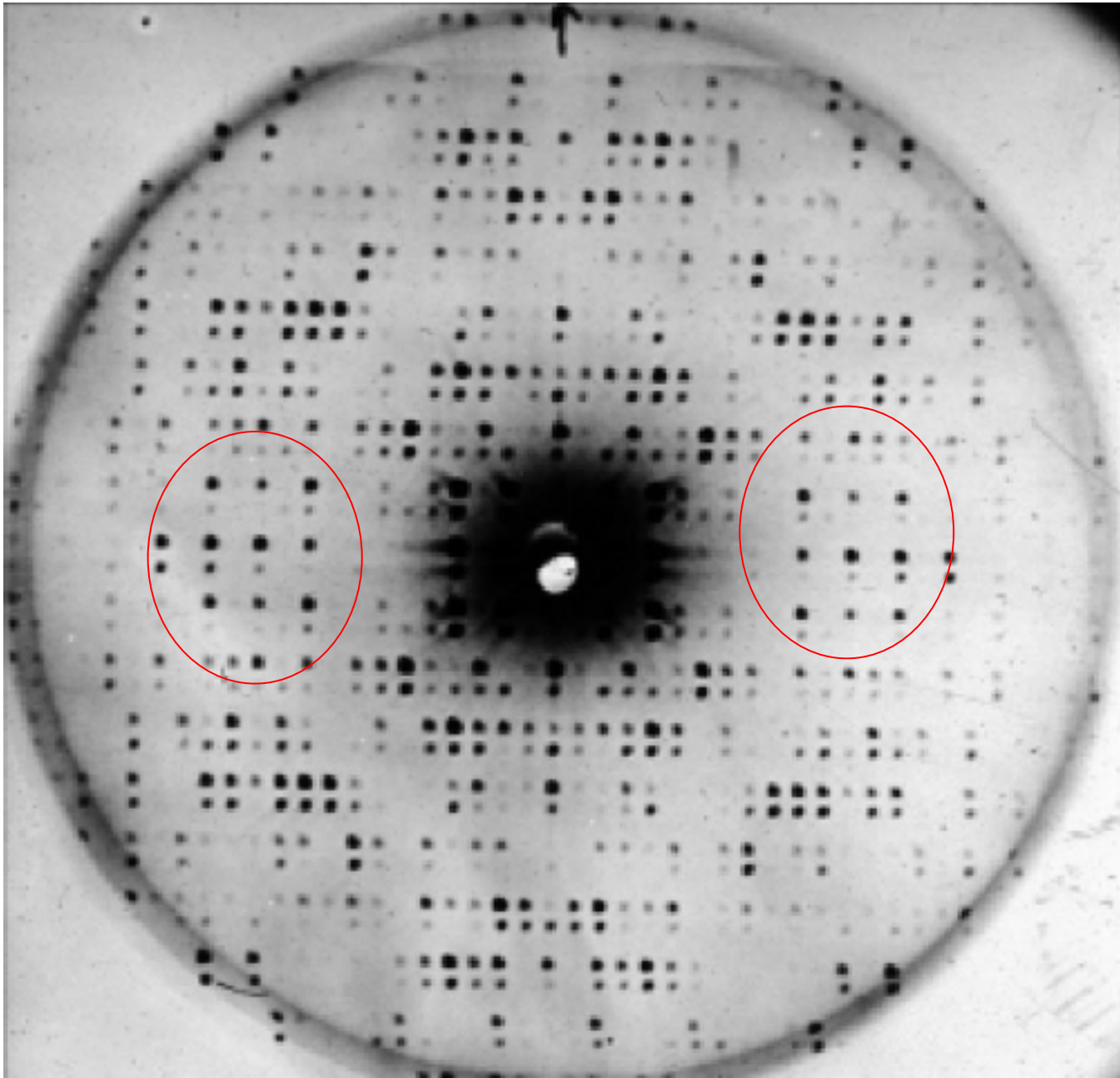


Figure 4
The process of molecular replacement.

Isomorphous Replacement

- Soaking protein crystals in heavy-atom solutions to create isomorphous HA derivatives
- Gives measurable intensity changes which could be used to deduce the positions of heavy atoms

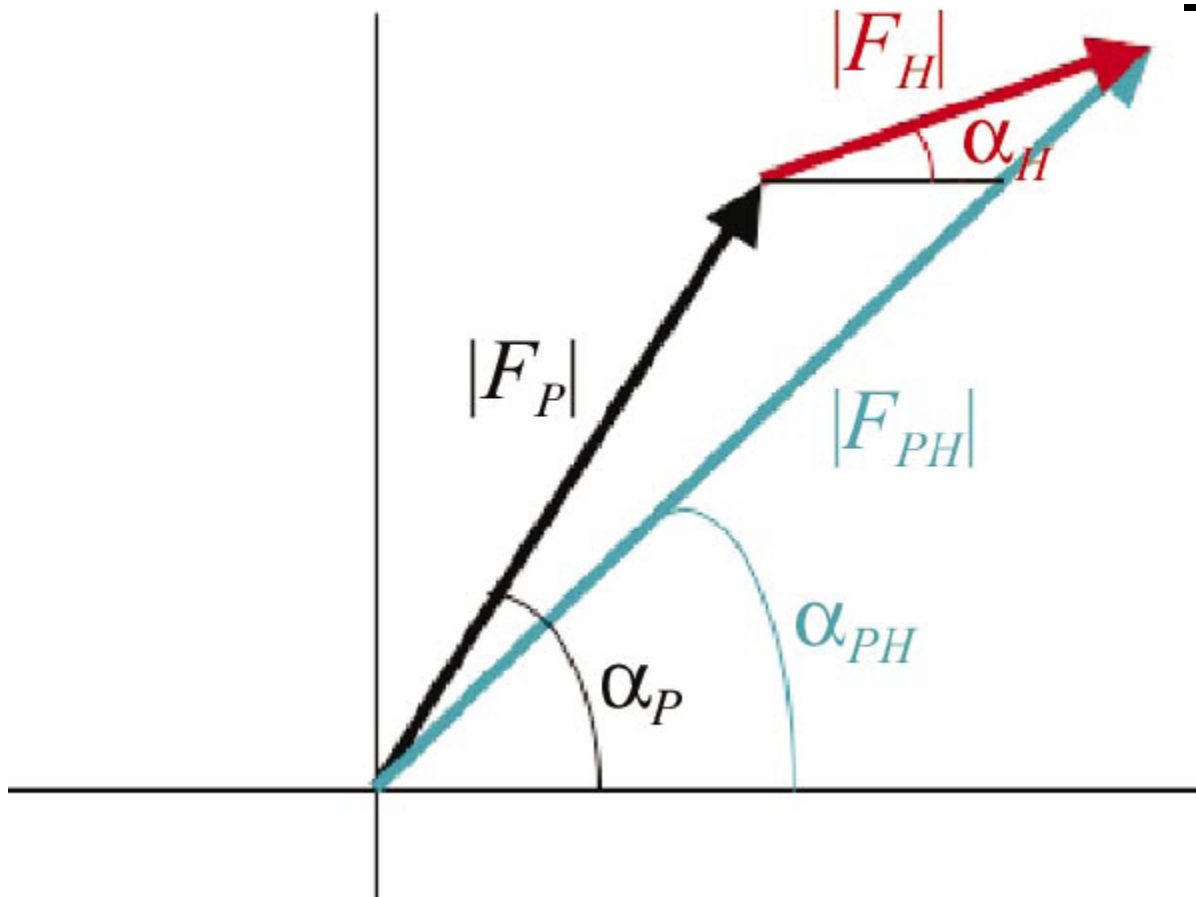


**2 Protein
Diffraction
patterns
superimposed
and shifted
vertically relative
to one another**

Single Isomorphous Replacement (SIR)

- The isomorphous difference, $|F_H| \simeq |F_{PH}| - |F_P|$ can be used as an estimate of the HA structure factor amplitude to determine the HA positions using Patterson or direct methods.
- HA parameters (xyz positions, occupancies and Debye-Waller thermal factors(B)) can be refined to calculate a more accurate $|F_H|$ and its corresponding phase α_H

Argand Diagram for SIR



$$|F_H| \simeq |F_{PH}| - |F_P|$$

Amplitudes of a reflection

$|F_P|$ Native Crystal

$|F_{PH}|$ HA Derivative Crystal

$|F_H|$ Isomorphous Difference

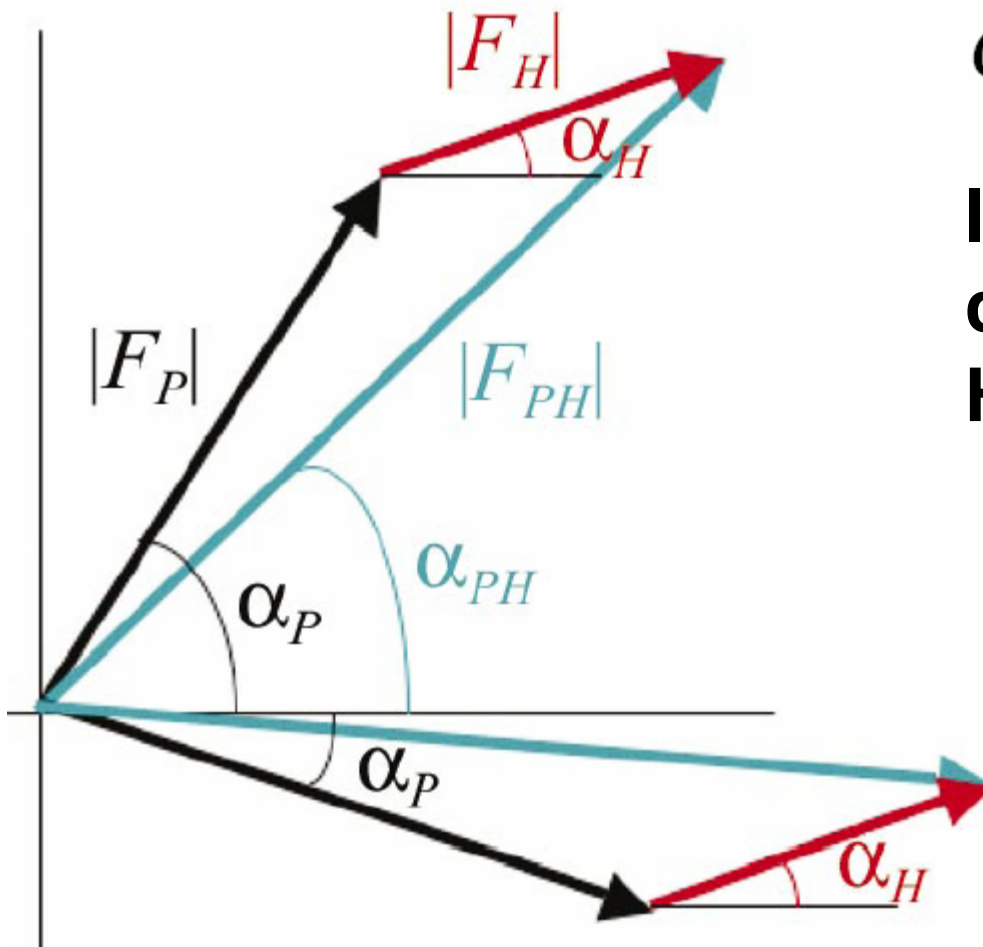
Phases of a reflection

α_P Native Crystal

α_{PH} HA Derivative Crystal

α_H Isomorphous Difference

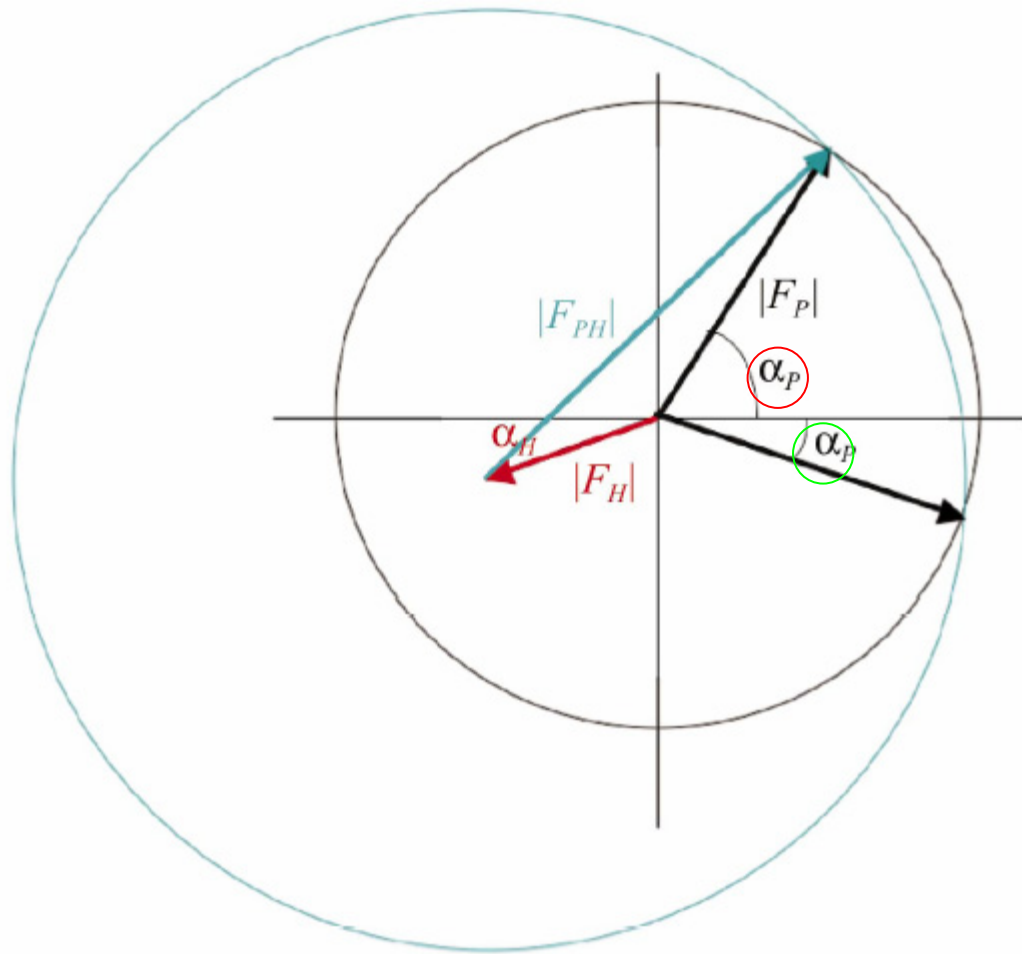
Cosine Rule: 2 Solutions



α_H Isomorphous
Difference

**Is symmetrically
distributed about the
HA phase**

Harker Construction for SIR

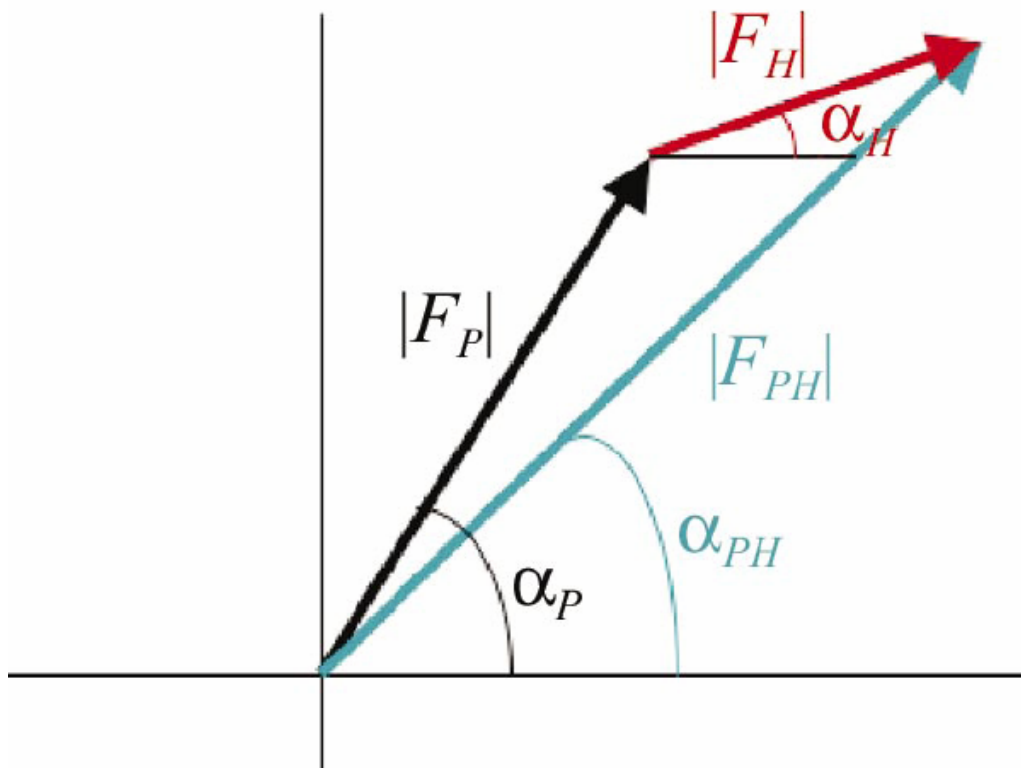


The 2 possible phase solutions occur where the circles intersect

The problem arises as to which phase to choose \circ or \circ ?

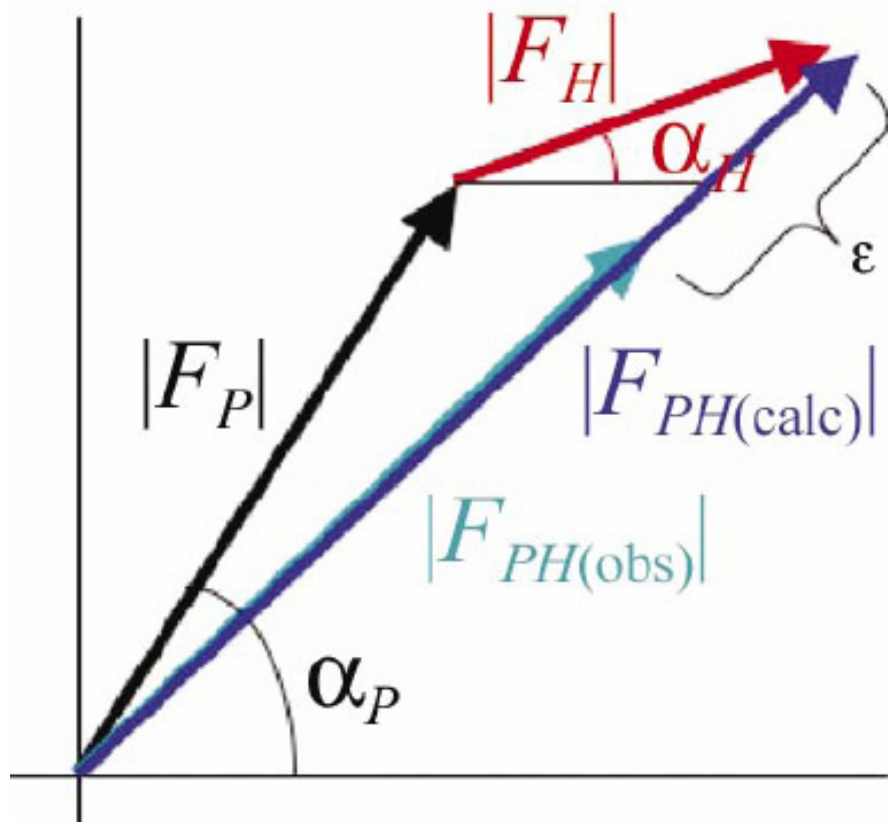
This requires a consideration of phase probability

Reality is that the ▲ Rarely Closes



- Due to Errors in measurement of Structure Factors, HA positions, and their occupancies

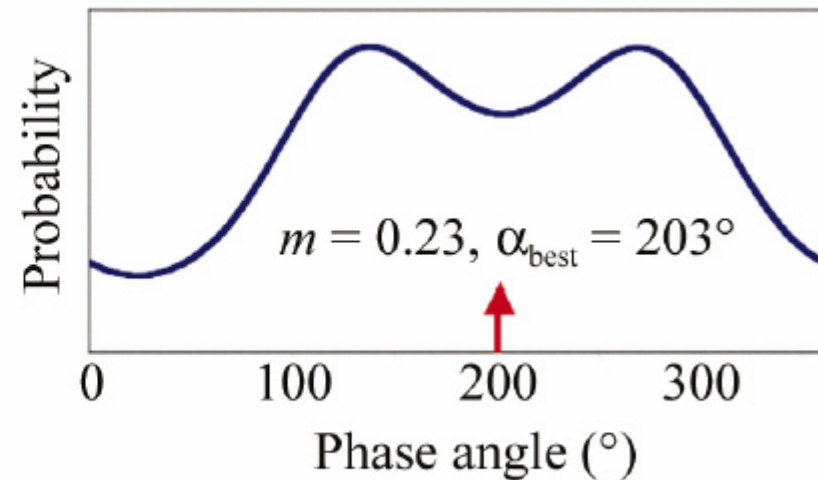
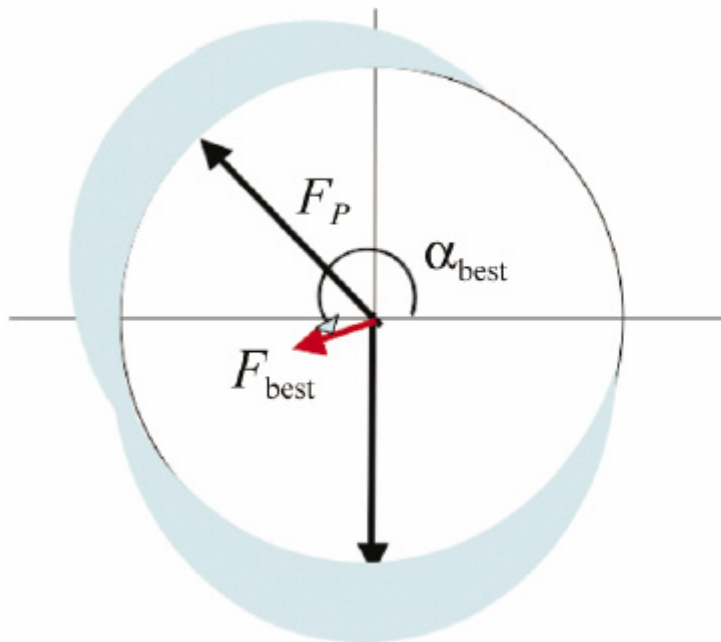
Phase Probability



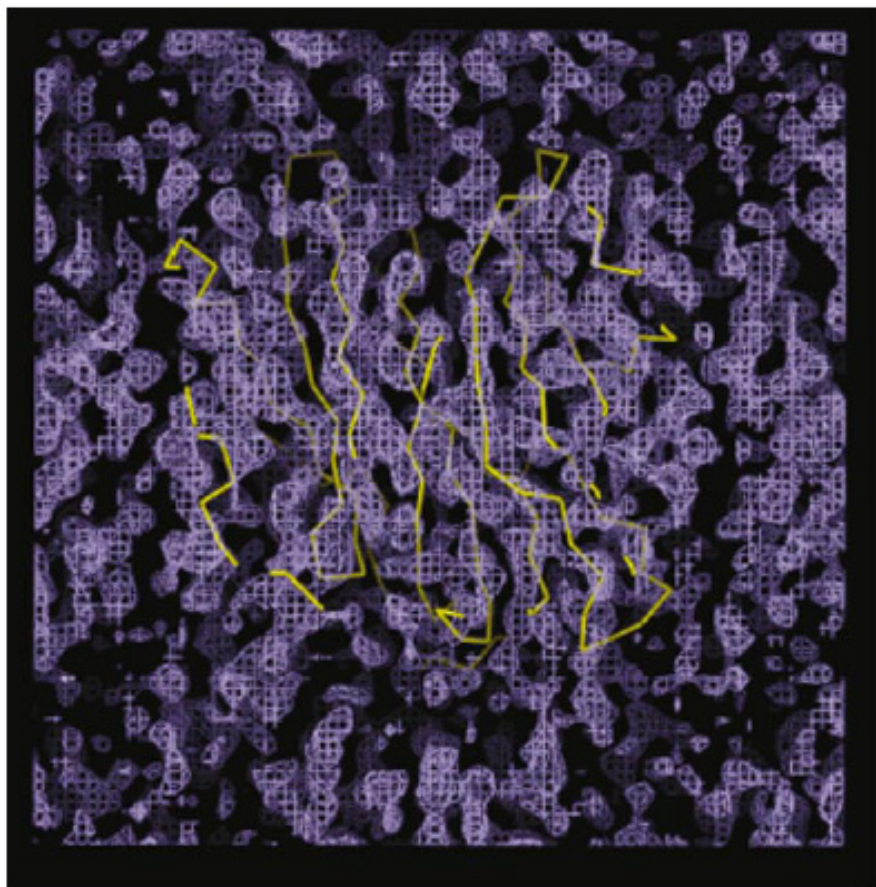
- Lack of Closure (ε)
- If we assume all errors are due to $F_{PH(\text{calc})}$ and that errors follow Gaussian distribution then:

$$\begin{aligned} \varepsilon &= |F_{PH(\text{obs})}| - |F_{PH(\text{calc})}| \\ &= |F_{PH(\text{obs})}| - | [|F_P| \exp(i\alpha_P) + |F_H| \exp(i\alpha_H)] | \end{aligned}$$

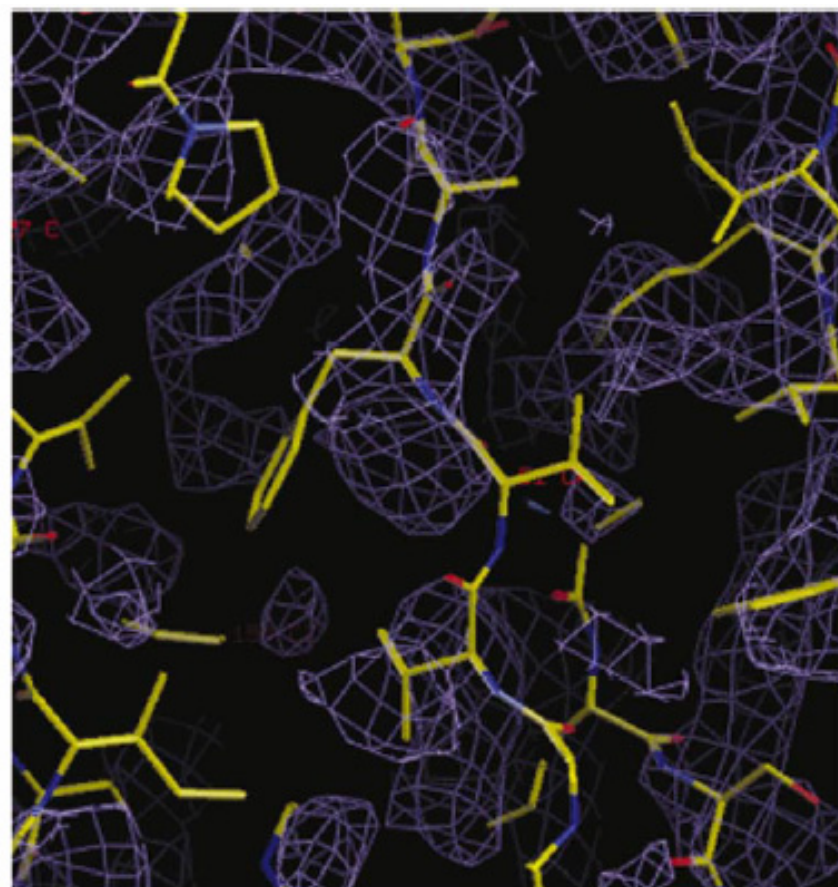
Phase Probability for 1 reflection in an SIR experiment



The problem is that a weighted amplitude representing the centroid of the phase distribution results in the least error



(a)



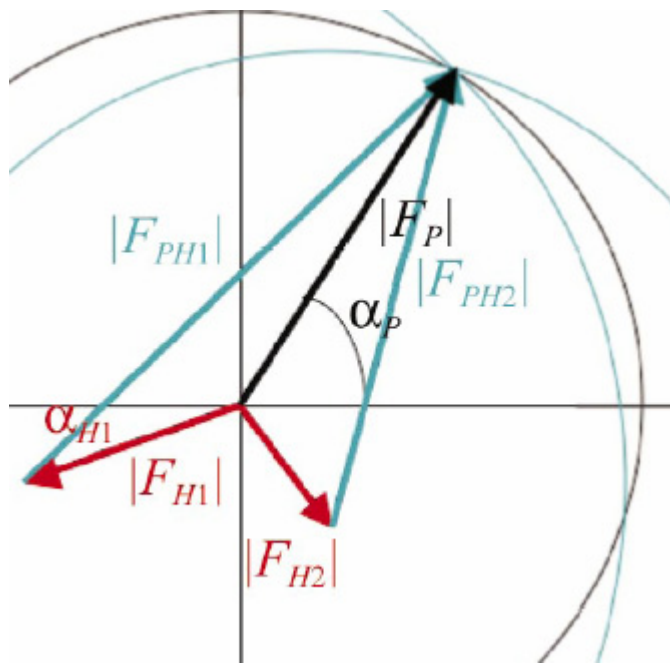
(b)

Figure 11

(a) A 2.6 Å SIR electron-density map with the final α -carbon trace of the structure superimposed. $\rho(\mathbf{x}) = (1/V) \sum m |F_P| \exp(i\alpha_{\text{best}}) \exp(-2\pi i \mathbf{h} \cdot \mathbf{x})$.
 (b) A small section of the map with the final structure superimposed.

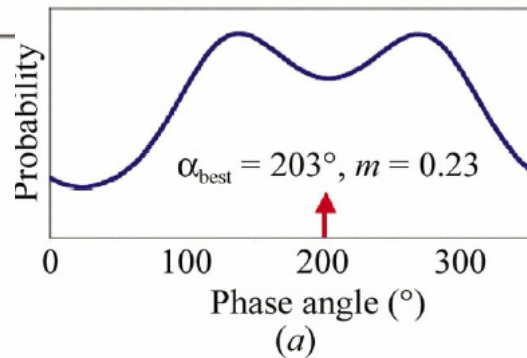
Multiple Isomorphous Replacement (MIR)

- Phase Probability using MIR (more than one HA) to break the phase ambiguity

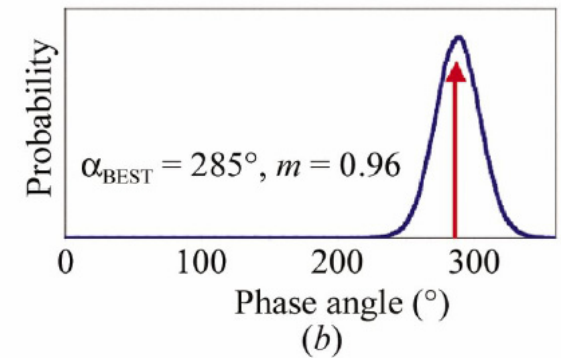


2 HA Derivatives

1 HA Derivative,
Low Error, Low FOM



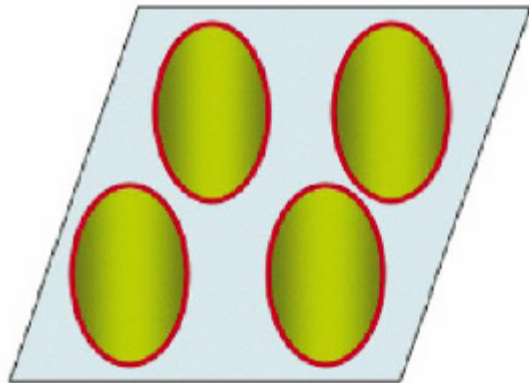
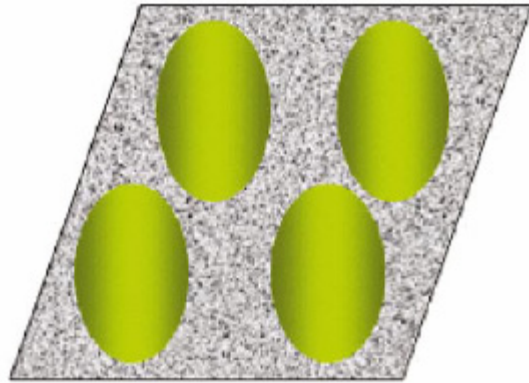
3 HA Derivatives,
Unimodal Distribution,
High FOM



Phase Improvement

- Experimental Phases are the starting point
- Improvement based on some prior knowledge
- Methods Include:
 - » **Solvent Flattening**
 - » **Histogram Matching**
 - » **Non-Crystallographic Symmetry (NCS) averaging**

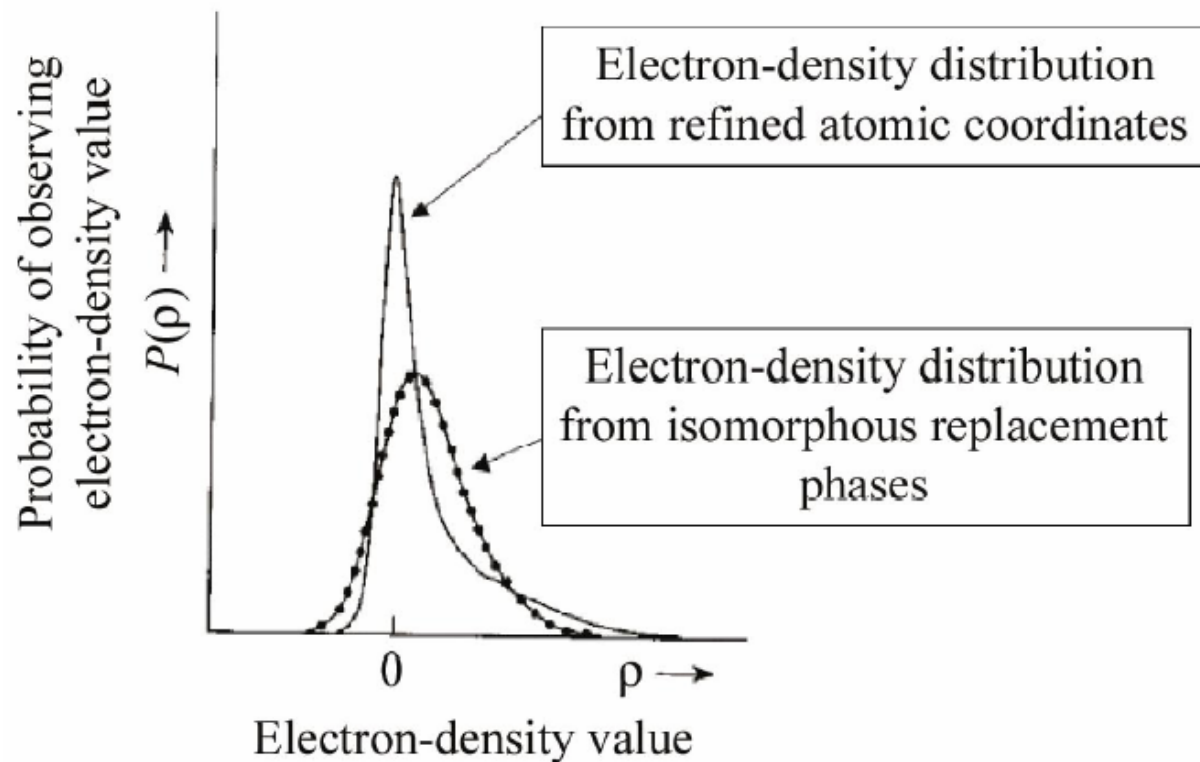
Solvent Flattening



- Sets values of electron density in the solvent regions to a typical value of $0.33 \text{ e}\text{\AA}^{-3}$, in contrast to typical protein electron density of $0.43 \text{ e}\text{\AA}^{-3}$.

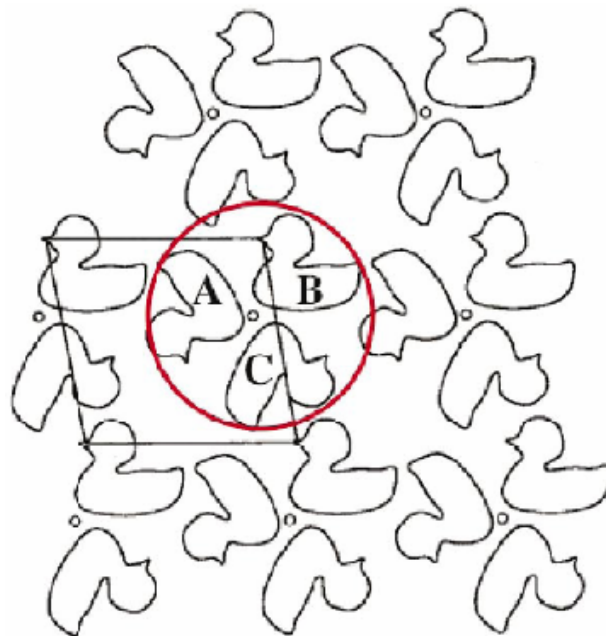
Histogram Matching

- Alters the values of electron density points to concur with an expected distribution of values



Non-Crystallographic Symmetry (NCS) Averaging

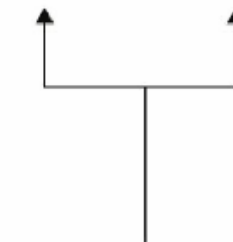
- Imposes equivalence on electron density values when more than one copy of a molecule is present in the asymmetric unit.



(c)

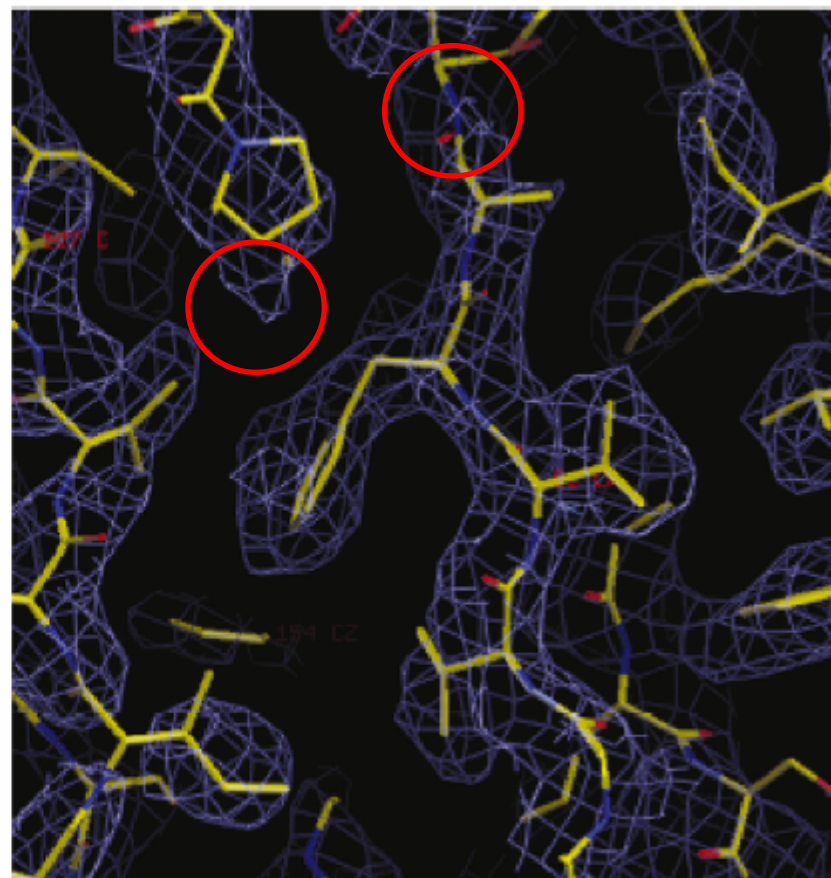
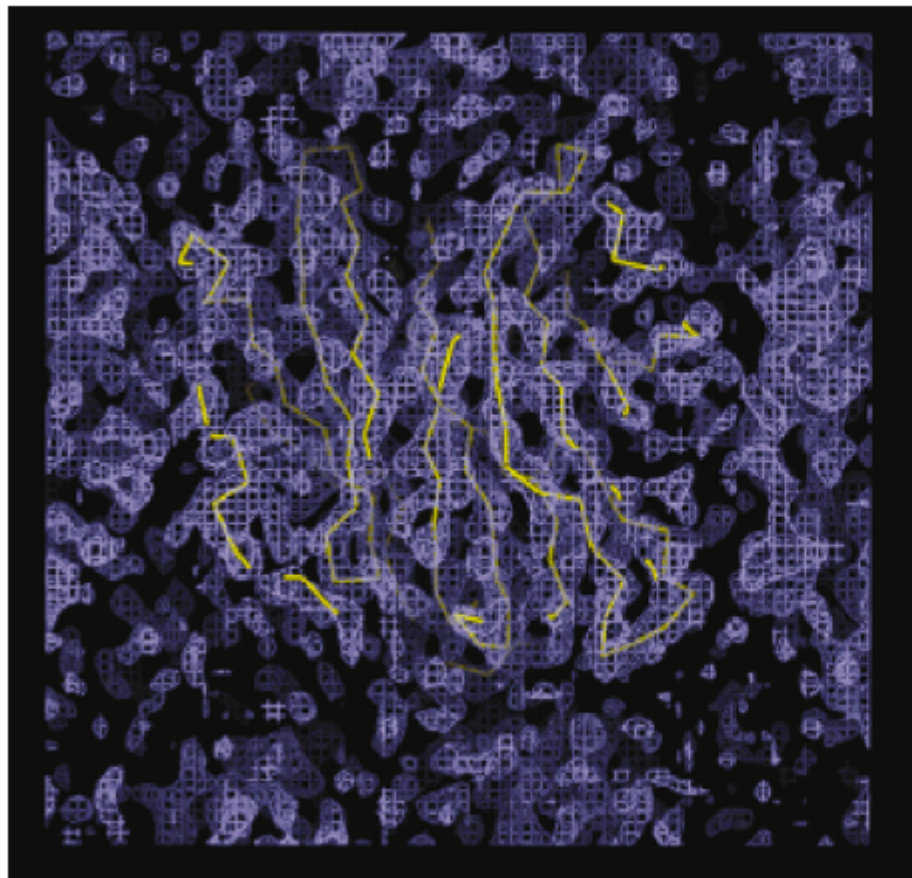
$$\rho(\mathbf{x}_B) = \mathbf{R}_B \rho(\mathbf{x}_A) + \mathbf{t}_B$$

$$\rho(\mathbf{x}_C) = \mathbf{R}_C \rho(\mathbf{x}_A) + \mathbf{t}_C$$

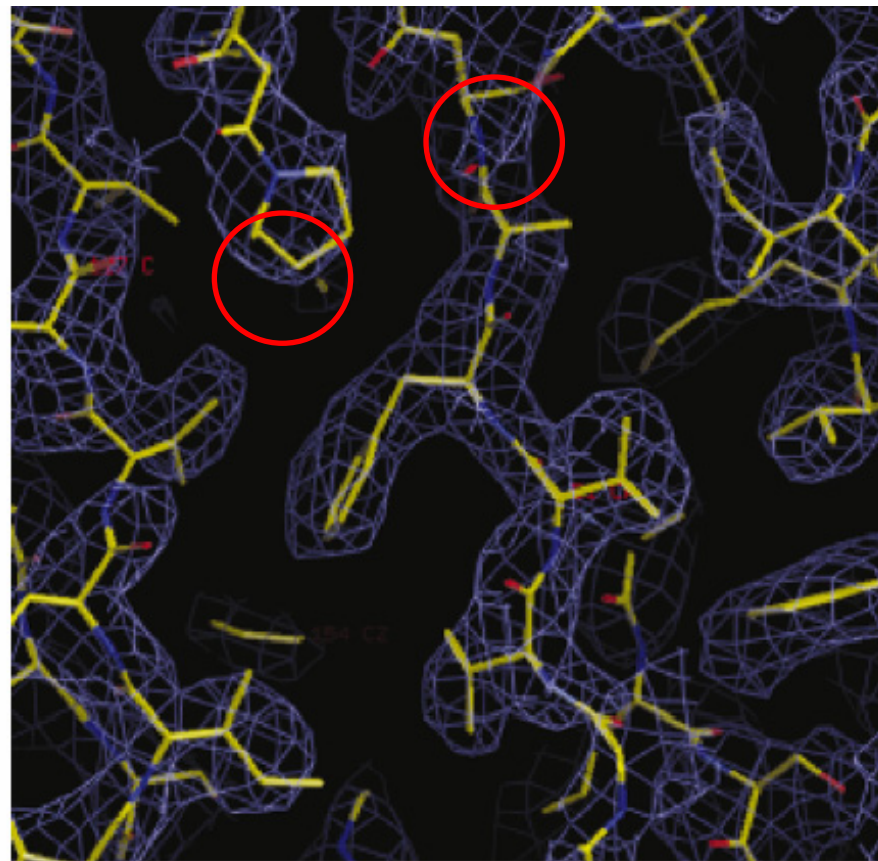
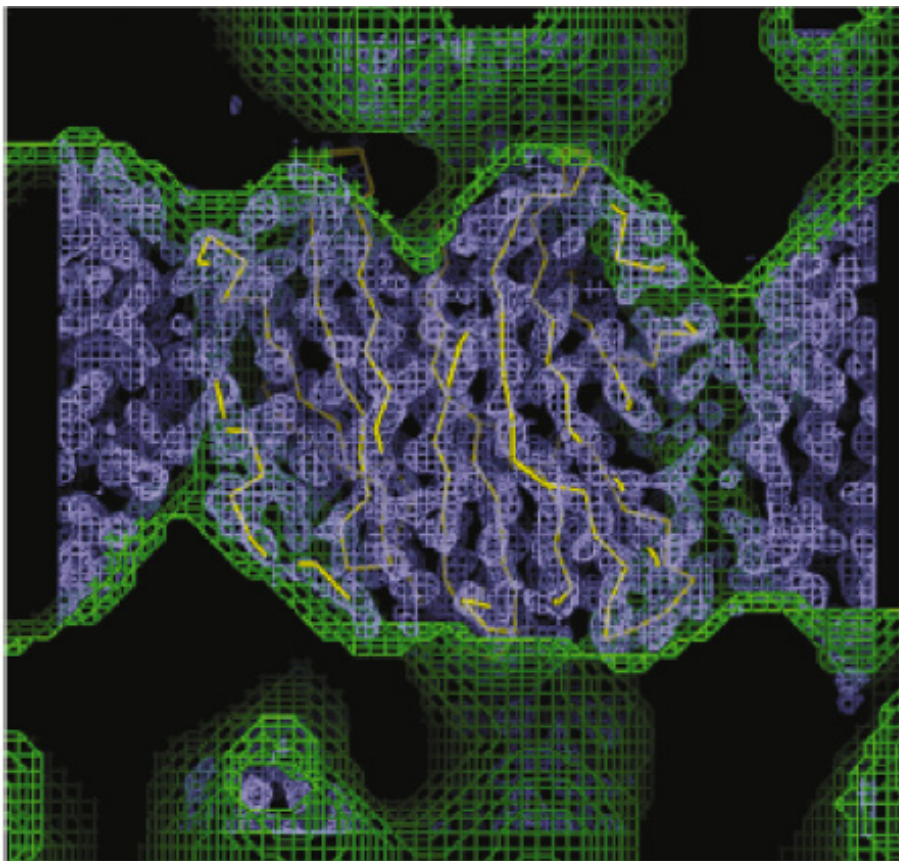


NCS symmetry
operators

Electron Density before Solvent Flattening and Histogram Matching in DM



Electron Density after Solvent Flattening and Histogram Matching in DM



- **Phase improvement by DM is an iterative process**
- **Which is Why often “Less is More”**

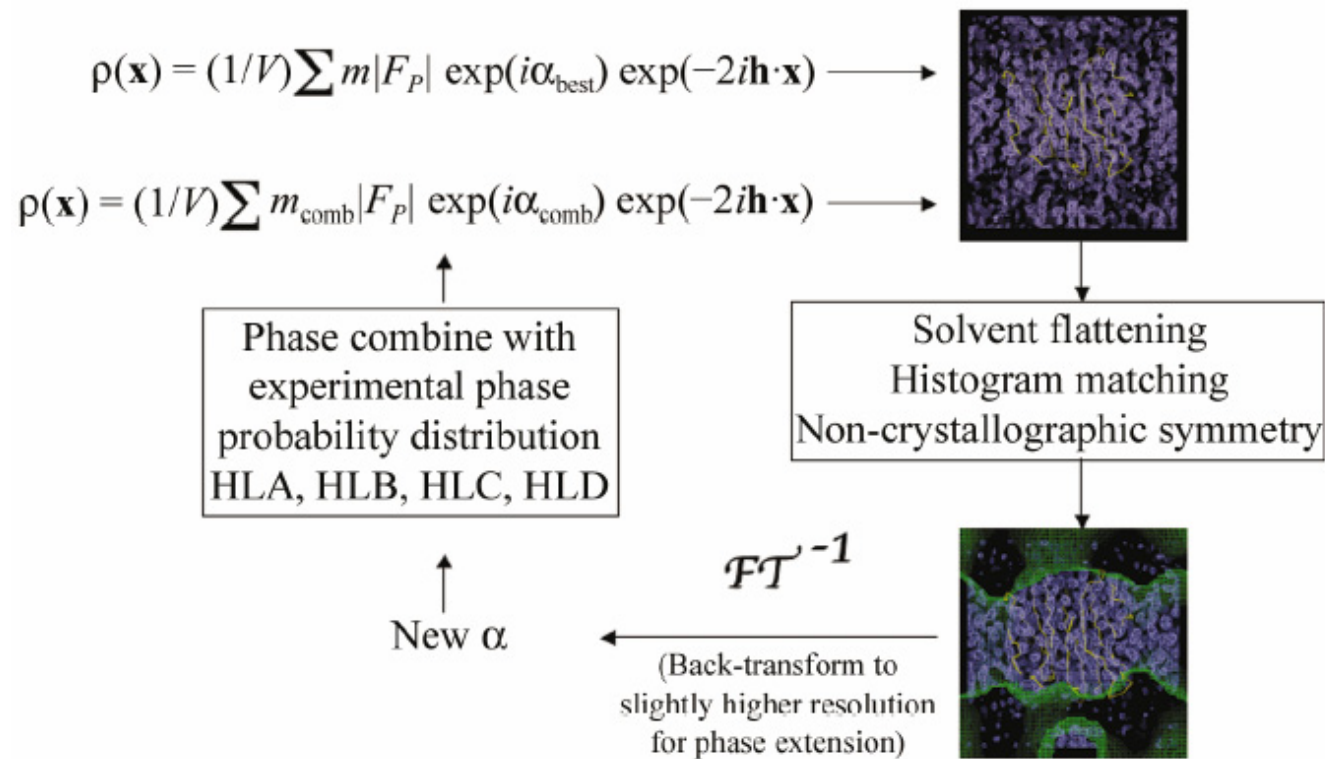


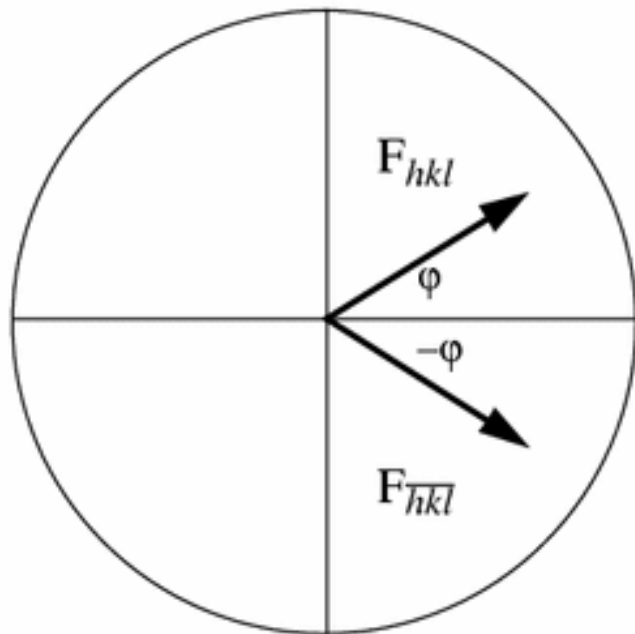
Figure 15
Phase improvement by density modification.

Anomalous Scattering

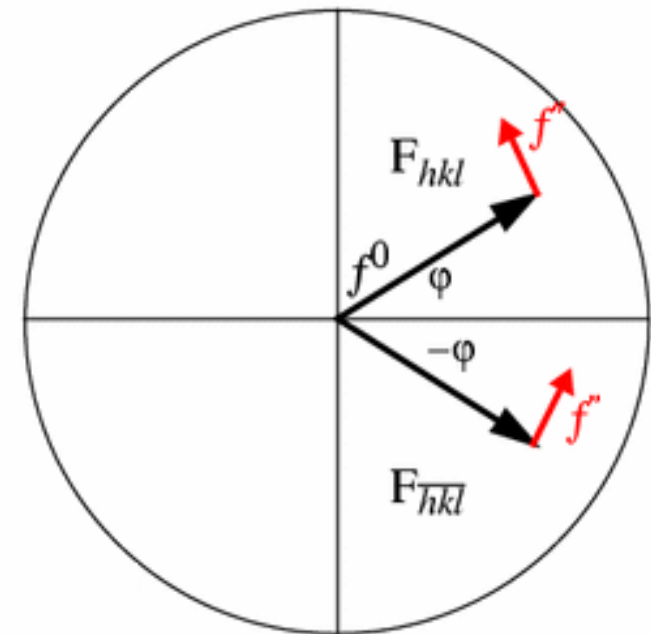
- Atomic Scattering factors are dependent on the Bragg angle and two on wavelength
- One at the absorption edge when the X-ray photon energy is sufficient to promote an electron from an inner shell, leading to a breakdown in Friedel's law
- This breakdown gives rise to anomalous differences that can be used to locate the anomalous scatters

Friedel's Law

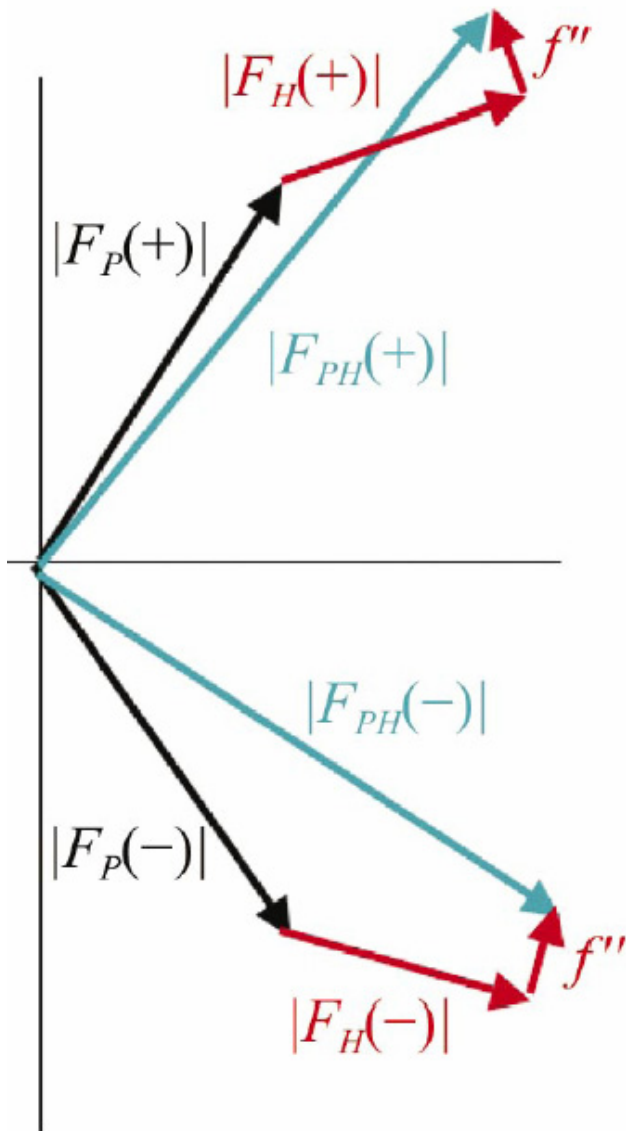
- Members of a Friedel pair have equal amplitude and opposite phase: $|F_{hkl}| = |F_{\bar{h}\bar{k}\bar{l}}|$ $\varphi_{hkl} = -\varphi_{\bar{h}\bar{k}\bar{l}}$
- If all atoms scatter equally, then the amplitudes remain equal but the phase relationship no longer holds. This is because the f term is always positive.



*Equal atom case:
all atoms have the same
scattering behaviour*



Breakdown of Friedel's Law



When an anomalous scatterer is present:

$$f(\theta, \lambda) = f_0(\theta) + f'(\lambda) + if''(\lambda)$$

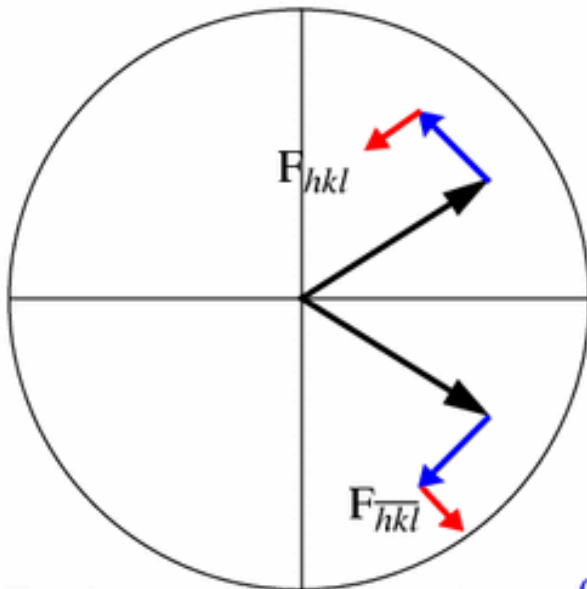
$$F_{hkl} \neq F_{\bar{h}\bar{k}\bar{l}}$$

And the Bijvoet Difference:

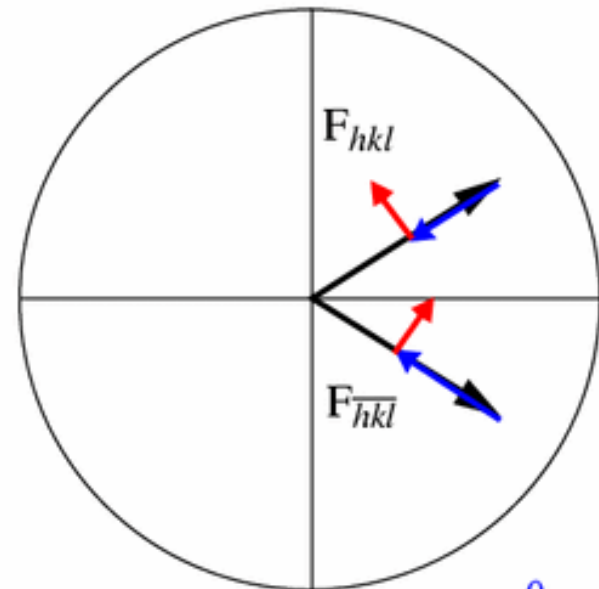
$$\Delta F^{\pm} = |F_{PH}(+)| - |F_{PH}(-)|$$

Breaking Friedel's Law

- The contribution of f^0 will be greatest at $\pi/2$ and smallest when f^0 is zero .

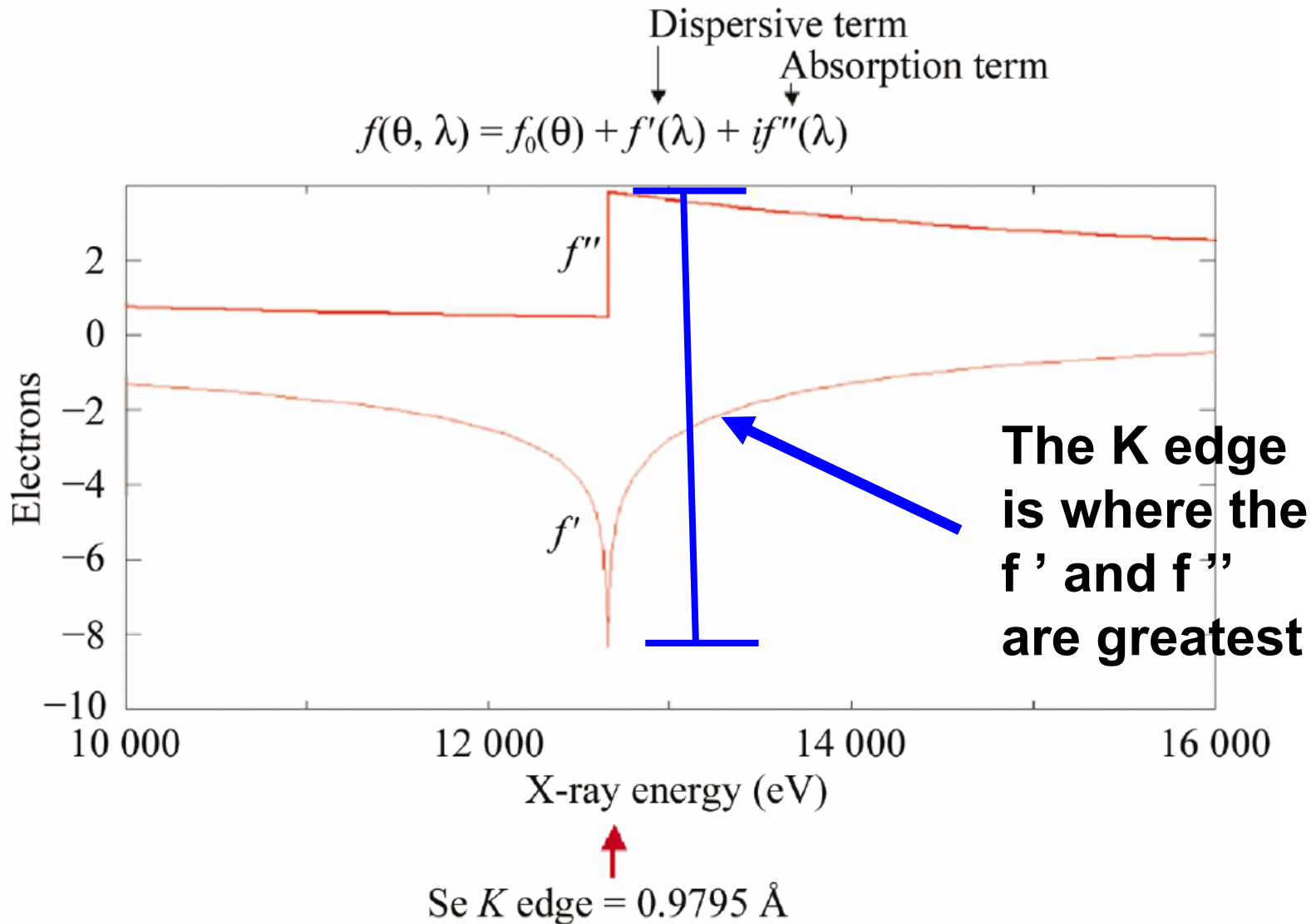


ΔF is largest when the contribution f^0 from the anomalous scattering atoms is $\pi/2$ out of phase with that of the non-anomalous atoms



ΔF is smallest (zero) when the f^0 contribution is in phase with that of the non-anomalous atoms

Variations in Anomalous Scattering at the K (peak) edge of Selenium



SIRAS and MIRAS

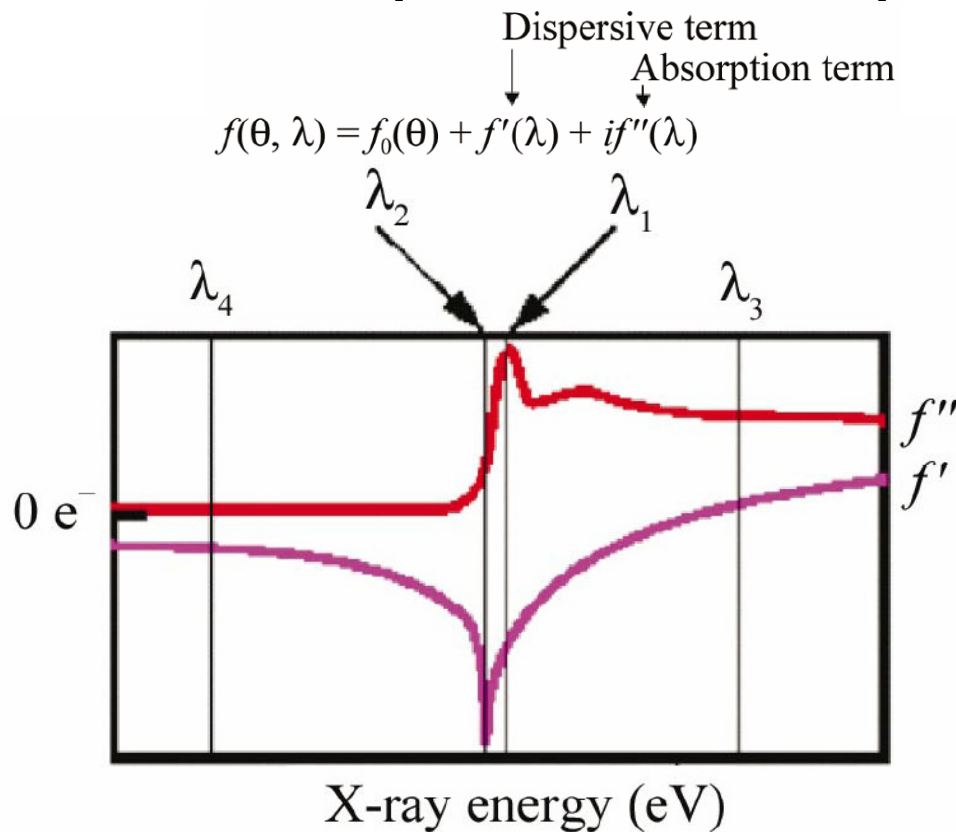
- Single Isomorphous Replacement with Anomalous Scattering (SIRAS) uses anomalous scattering to break the phase ambiguity in a single isomorphous replacement experiment.
- Multiple Isomorphous Replacement with Anomalous Scattering (MIRAS) uses multiple anomalous scatterers to break the phase ambiguity in a multiple isomorphous replacement experiment.
- Both of these use the anomalous or Bijvoet difference in the same way as the isomorphous difference in Patterson or direct methods to locate the anomalous scatterers.

SIRAS-MIRAS Limitations

- Non-isomorphism between crystals: unit cell changes, reorientation of the protein, conformational changes, changes in salt and solvent ions
- Problems in locating all the heavy atoms
- Problems in refining heavy atom positions
- Occupancies and thermal parameters
- Errors in intensity measurements

Multiwavelength Anomalous Diffraction (MAD)

- Overcomes non-isomorphism problems
- Collect 3λ , in order to maximize the absorption and dispersive effects



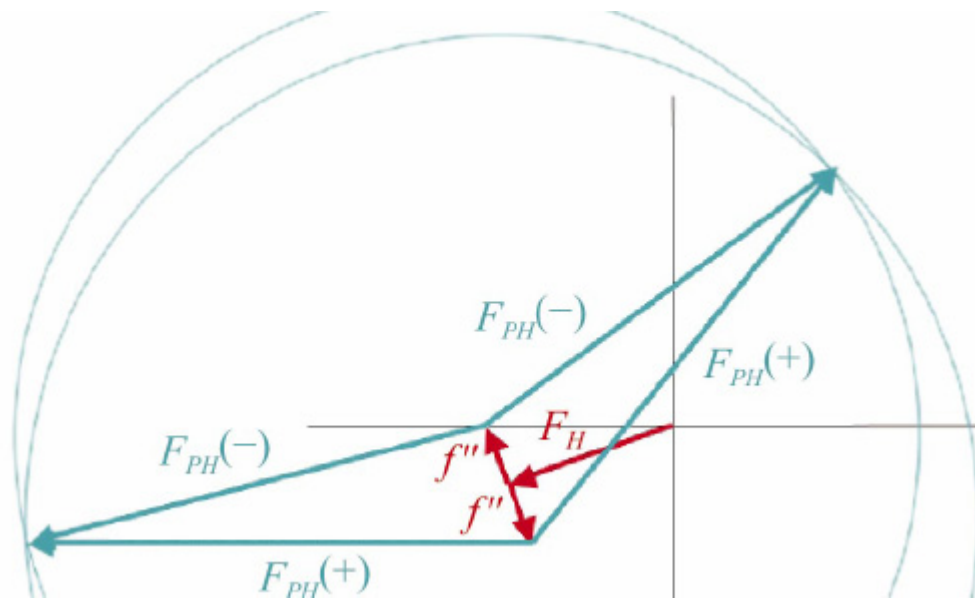
f'' Peak (λ_1)

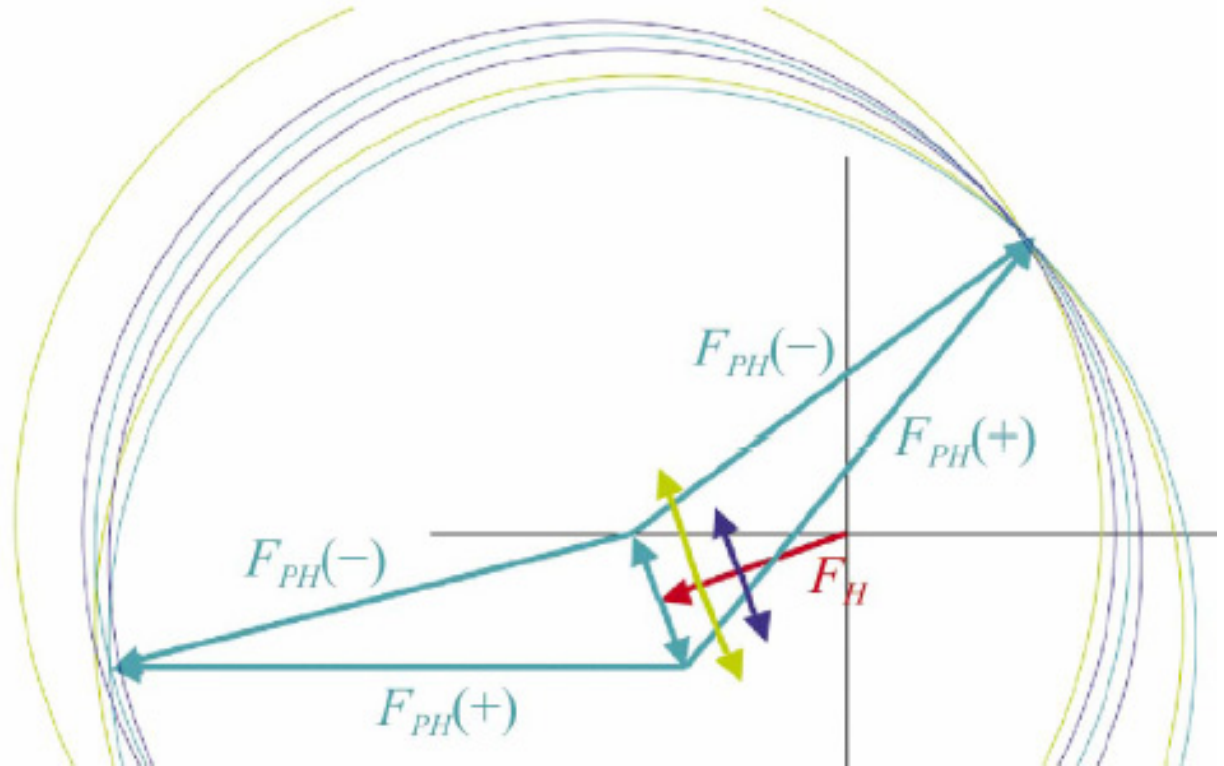
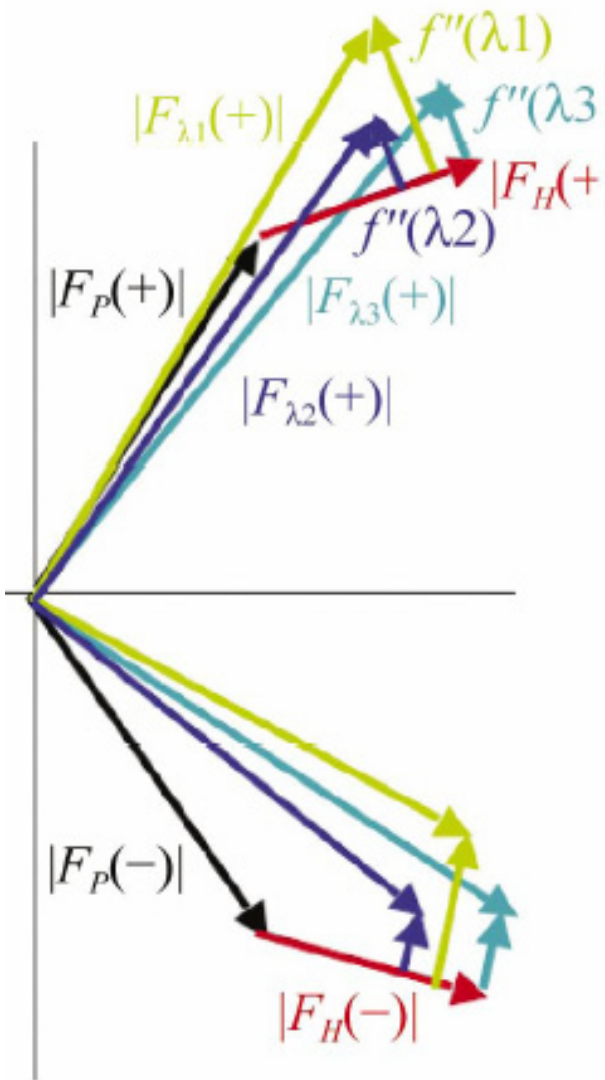
f'' Inf (λ_2), where f' is at its minimum

Remote (λ_3) or (λ_4), the “soft” and “hard” remotes typically 100 to 1000eV from K

Single-Wavelength Anomalous Diffraction

- Typically at the peak, and to use density-modification protocols to break the phase ambiguity. See Dodson 2003



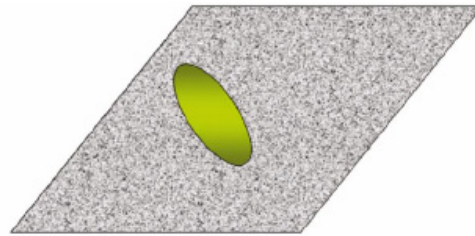


F_P is not measured so use one of its λ s, λ_3 as 'native'

Cross-Crystal Averaging

- Cases where MAD or SAD cannot be used
- Low-resolution phases, high-resolution data for another crystal form
- Involves mapping the electron density from the one unit cell into the other.
- “one can bootstrap the phases to high resolution”

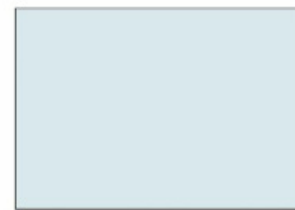
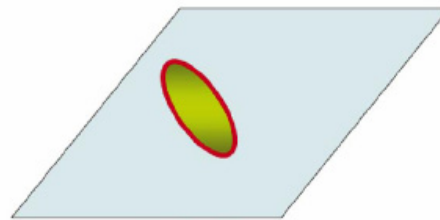
Cross Crystal Averaging



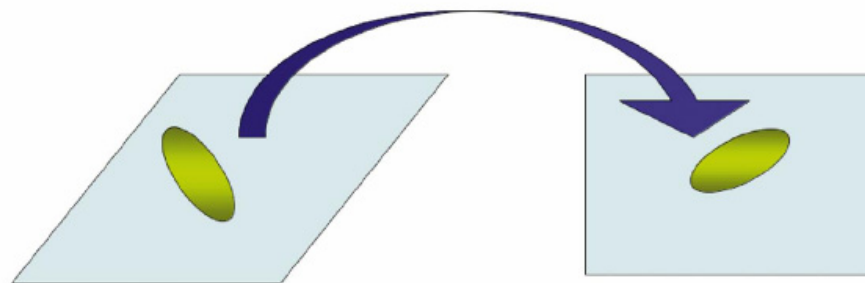
4 Å data and phases
from this crystal form



2 Å native data, no phases
from this crystal form



Cut out density of one
molecule, flatten solvent $\rightarrow \mathcal{FT}^{-1} \rightarrow |F_{\text{model}}| \rightarrow \text{MR}$



Find \mathbf{R} and \mathbf{t} that transform the molecule from A to B
Cross-crystal average and phase extend (*DMMULTI*)



"The central importance of this is it can enable us to design more effective drugs for prevention and treatment of disease."

From the Shake and Bake Website:
<http://www.psc.edu/science/Hauptman/Hauptman.html>