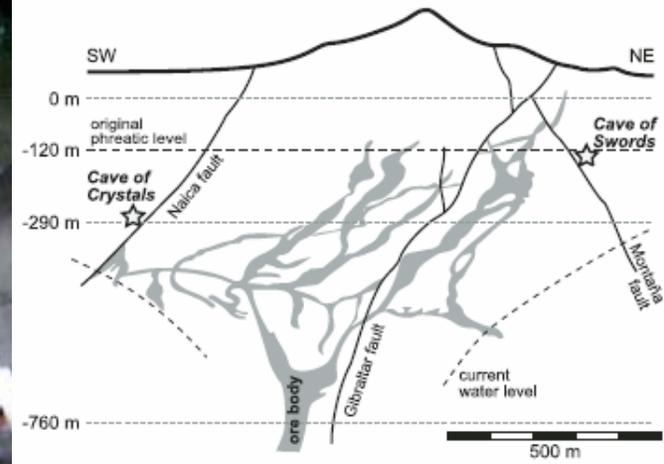


Niaca Mine, Chihuahua, Mexico



X-Ray Crystallography

“If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist.” – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography (after NMR)

a) Crystal Growth – Materials / Methods

b) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

c) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

d) Theory of Diffraction – Bragg's Law / Reciprocal Space

e) Data Collection – Methods / Detectors / Structure Factors

f) Structure Solution – Phase Problem: MIR / MR / MAD

h) Refinement, Analysis and Presentation of Results

i) Use of Difference Fourier

Object

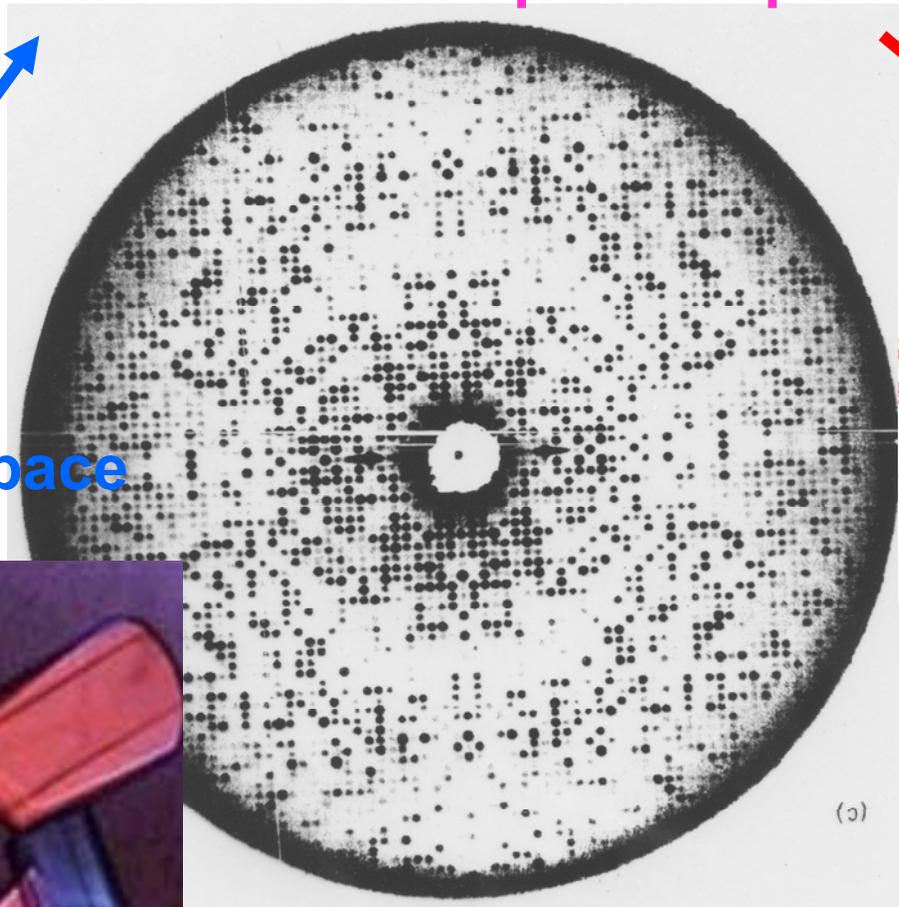


Transform



Image

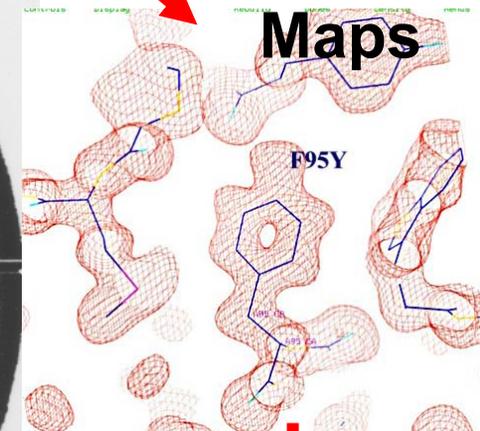
Transform / Reciprocal Space



Object / Real Space

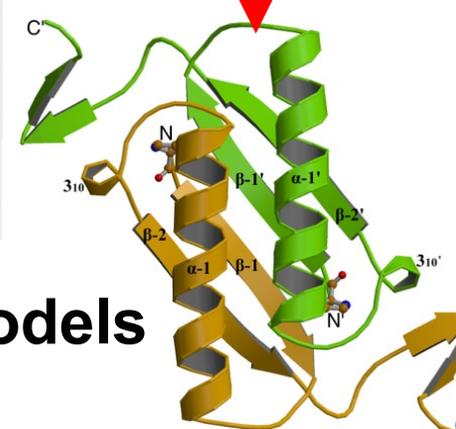


Electron Density Maps



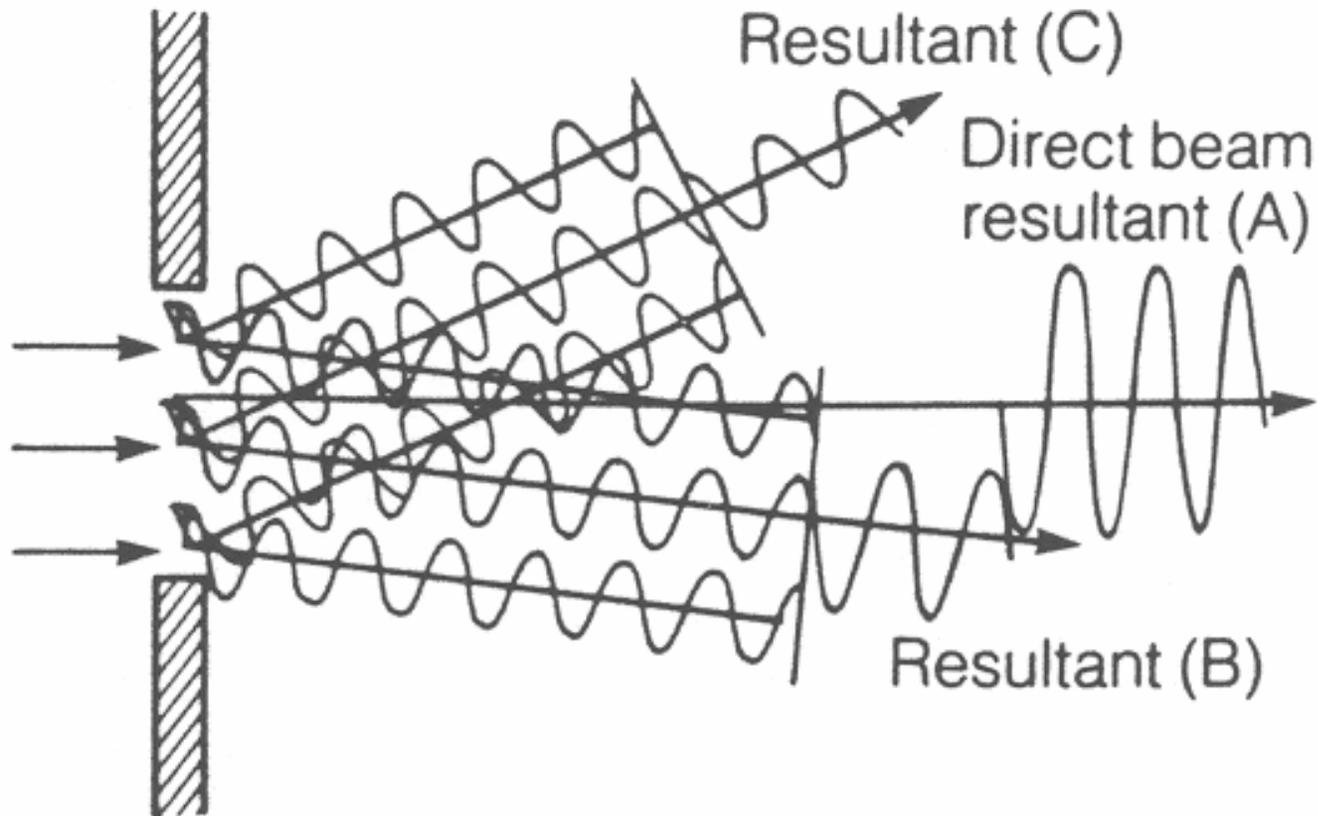
(c)

Models



Single Hole Scattering Experiment

Transforms / Reciprocal Space



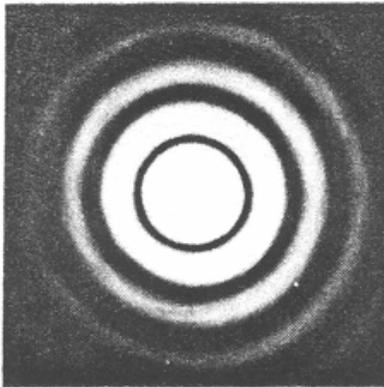
(b)

Single Hole Scattering Experiment

Transforms / Reciprocal Space

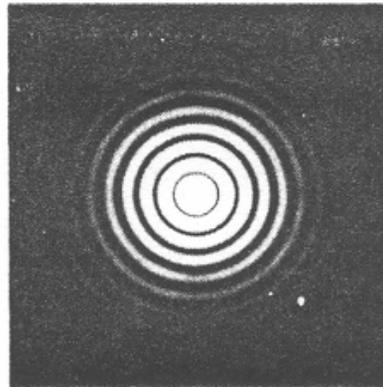
Different size holes

(a) ●



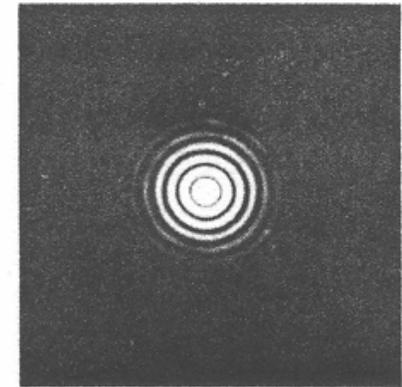
(a)

(b) ●



(b)

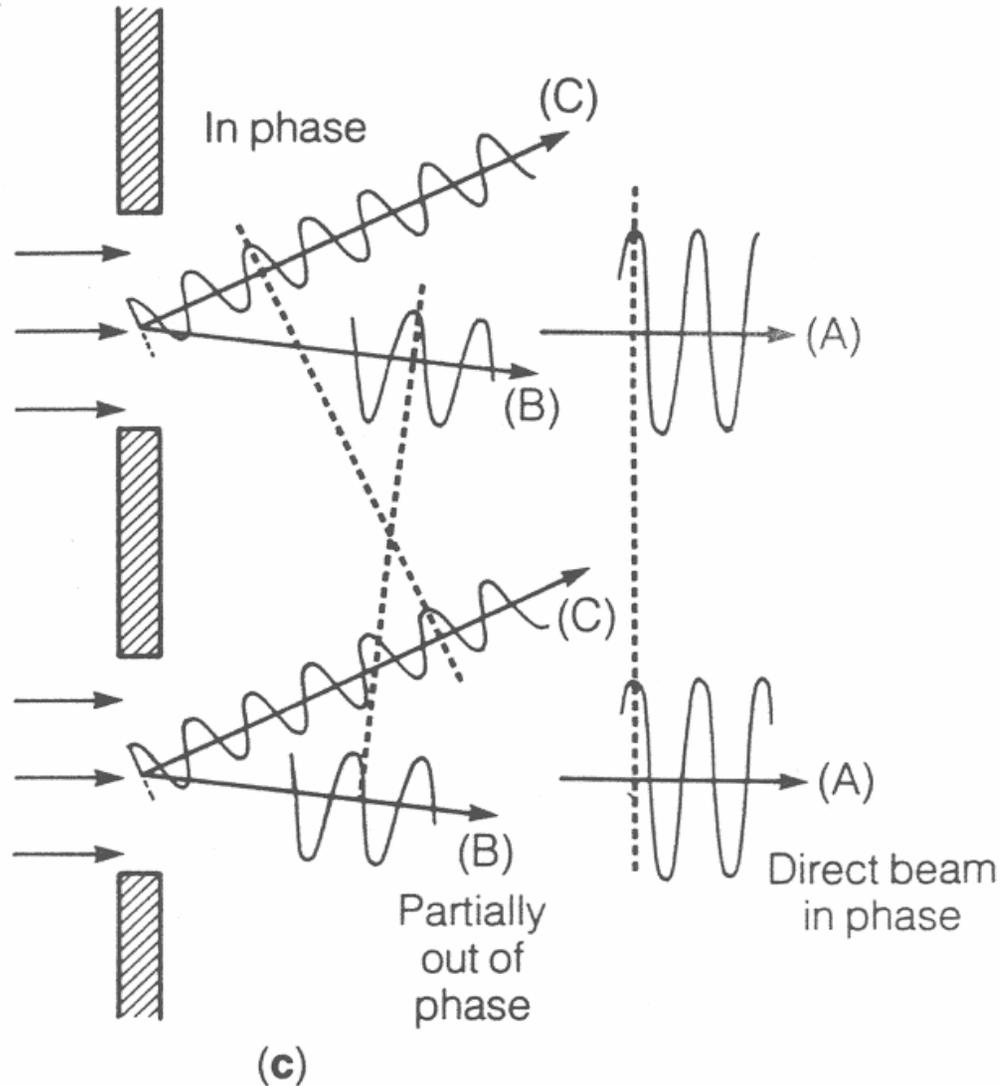
(c) ●



(c)

Effect of Multiple "Scatterers"

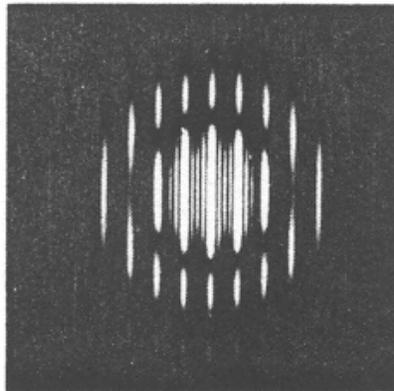
Transforms / Reciprocal Space



Transforms / Reciprocal Space

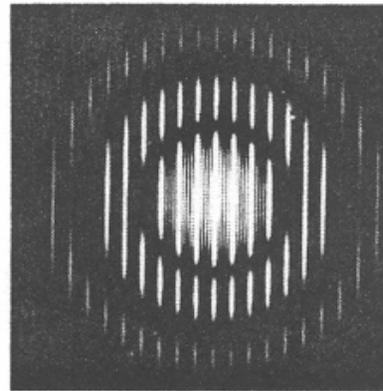
Five horizontal holes
with various spacings

(j) ● ● ● ● ●



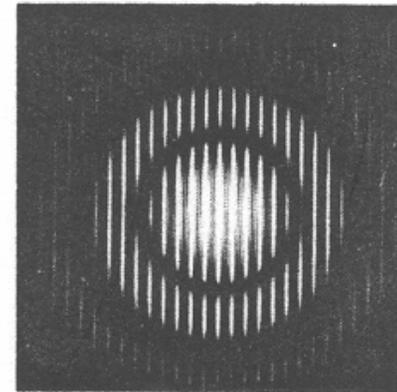
(j)

(k) ● ● ● ● ●



(k)

(l) ● ● ● ● ●

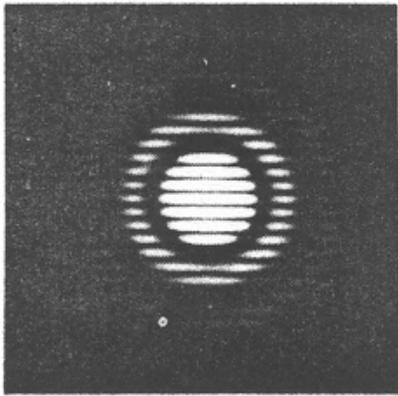


(l)

Transforms / Reciprocal Space

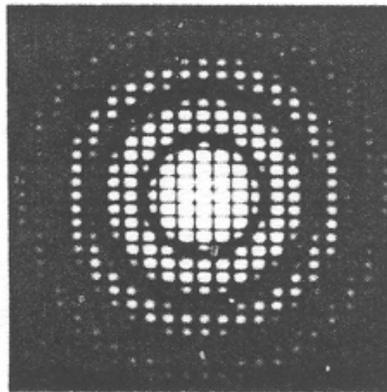
Vertical holes and nets of holes

(g)



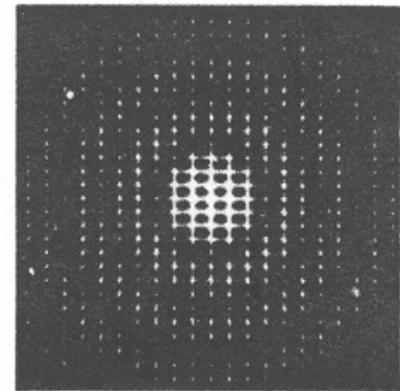
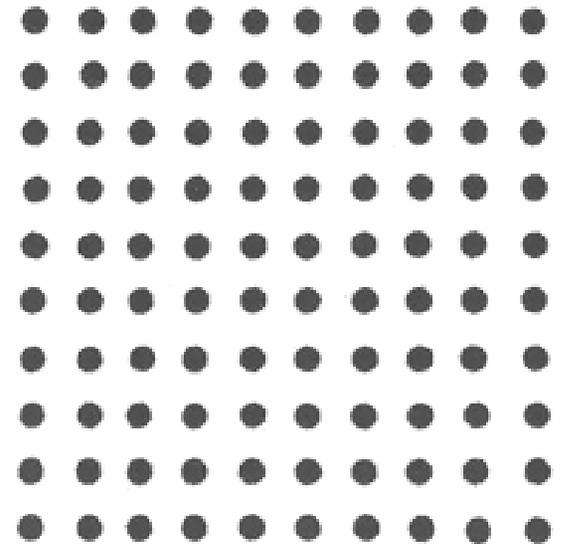
(g)

(h)



(h)

(i)

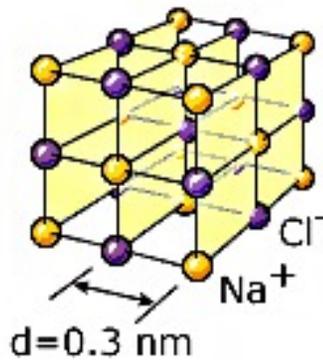


(i)

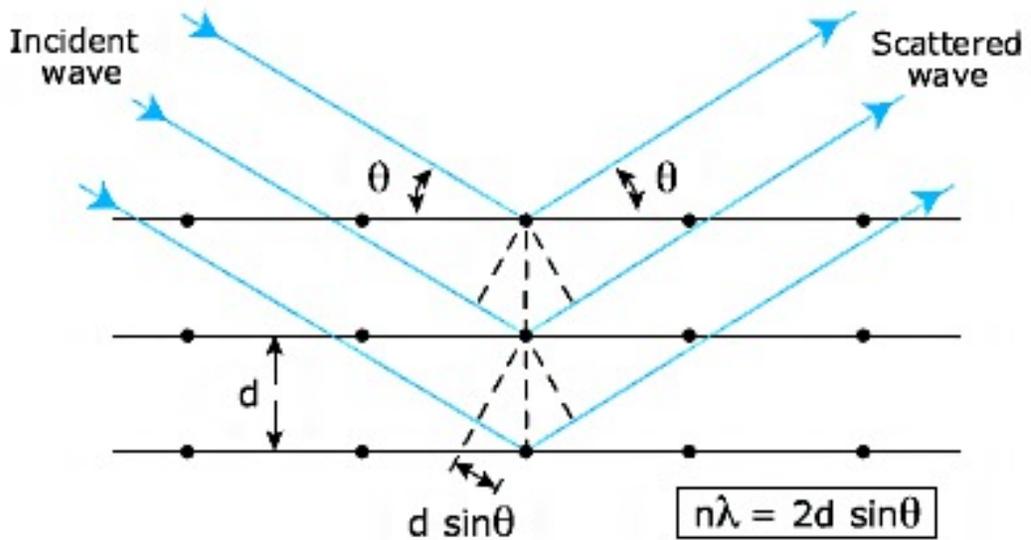


More About the Bragg Formula

X-rays scattered from different layers of atoms can interfere with each other. The interference depends on the wavelength of the X-ray and on the distance between the atom layers. An X-ray with well-known wavelength can be used to explore the structure of the crystal. For a well-known crystal, the X-ray properties can be examined.



Crystal planes, in NaCl, ordinary salt. Other planes are also possible.



X-ray scattering from three crystal planes, separated by the distance d . For constructive interference in a direction θ the path difference must be an even number of wavelengths.

Related Laureates



The Nobel Prize in Physics 1915 - Sir William Henry Bragg »



The Nobel Prize in Physics 1915 - William Lawrence Bragg »

Diffraction: Scattering from (two) “atoms”

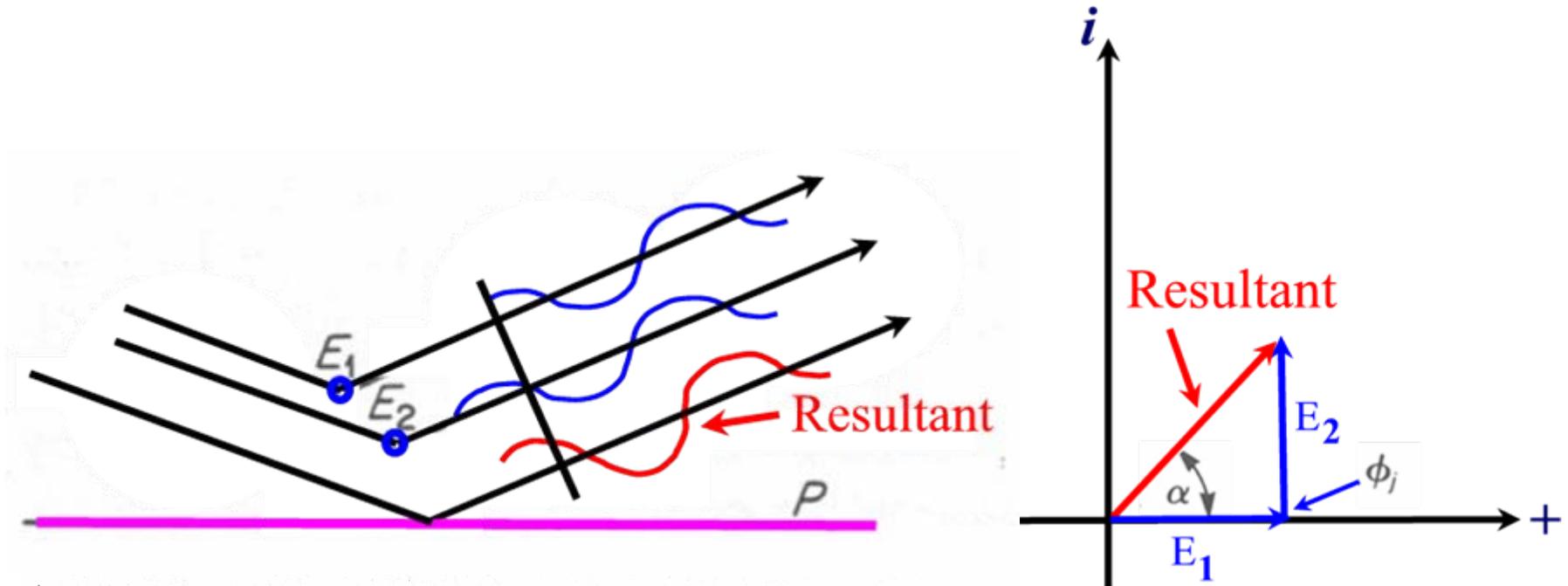


Figure 2.10. Diffraction from E_1 and E_2 as if reflected from plane P .

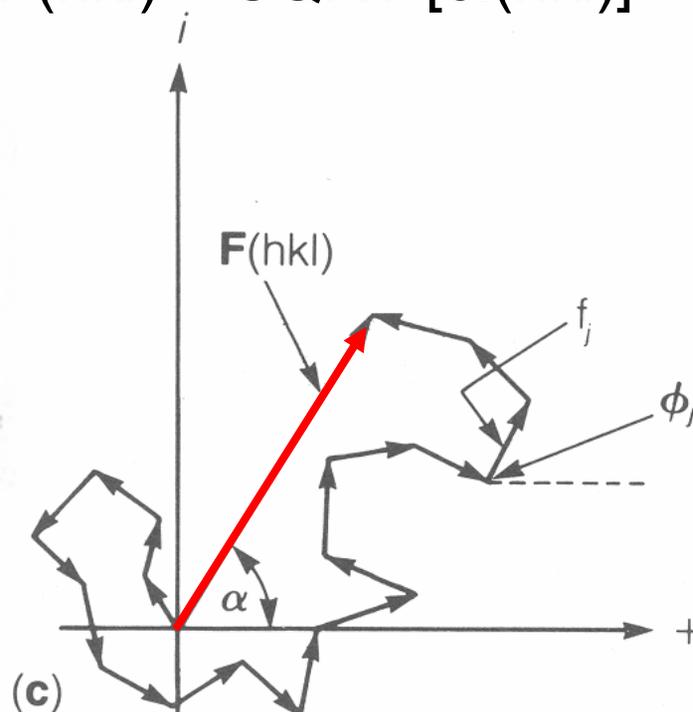
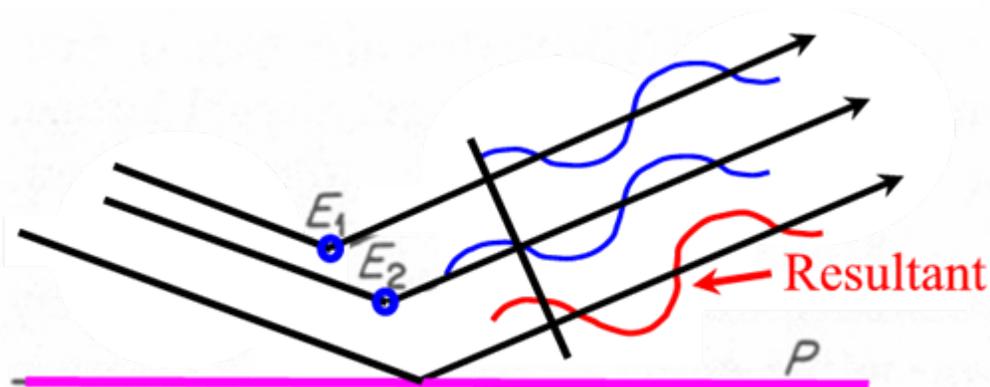
Scattering from “many atoms”

$$\mathbf{F}(hkl) = F(hkl)e^{i\alpha(hkl)} = \sum_{j=1}^{N'} \mathbf{f}_j(hkl) = \sum_{j=1}^{N'} f_j(hkl)e^{i\phi_j(hkl)}$$

← Calculated

← Experimental

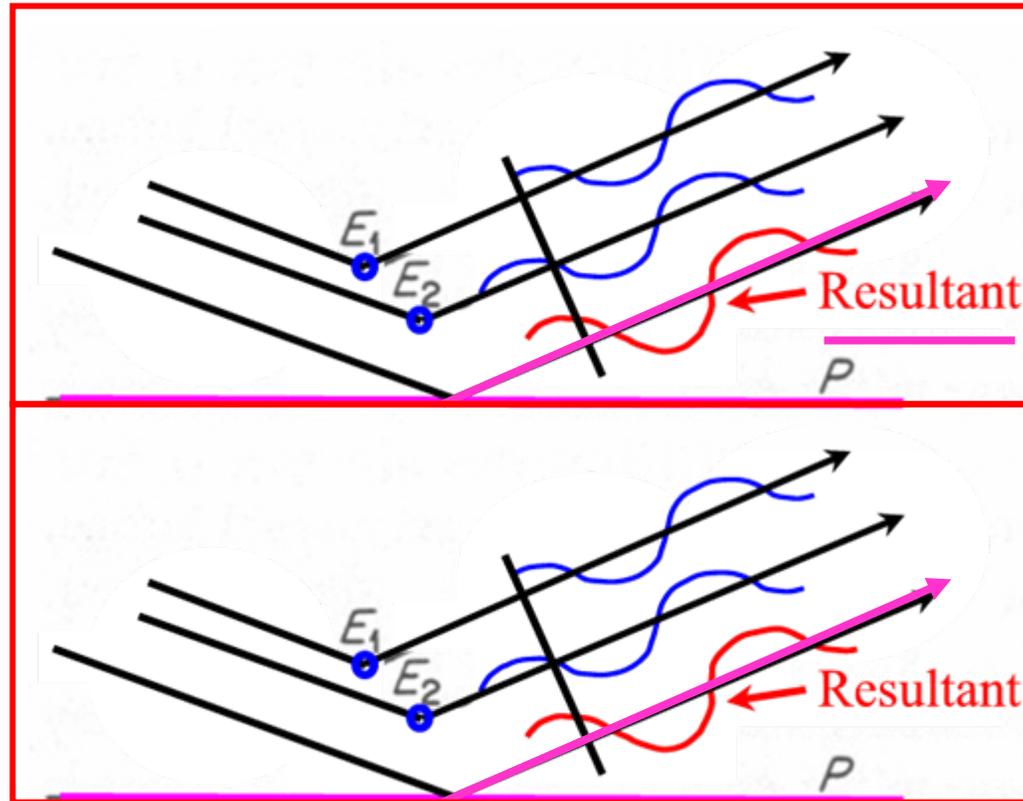
$$F(hkl) = \text{SQRT} [cI(hkl)]$$



The structure factor for a reflection may be thought of as the vector sum of the x-ray scattering contributions from many atoms.

Each of the j contributions may be represented as a vector in the complex plane, with amplitude f_j and phase ϕ_j .

Scattering from “atoms in two unit cells”

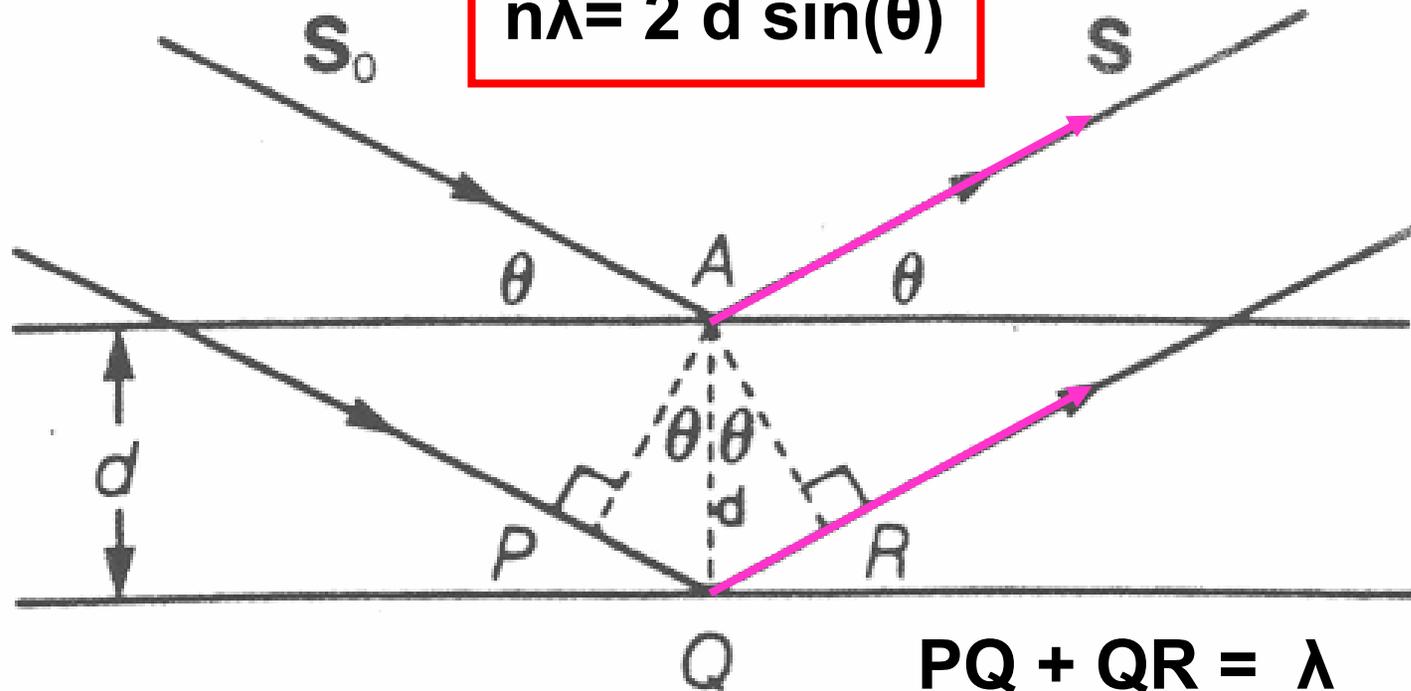


Crystals: Scattering from “planes”

Resultant scattering of resultant scattering!

Bragg Equation

$$n\lambda = 2d \sin(\theta)$$



→ Scattering will only be “observed” at discrete **Bragg angles** (θ)

The spacings of the **Bragg reflections** → **Lattice Constants**

Bragg Planes

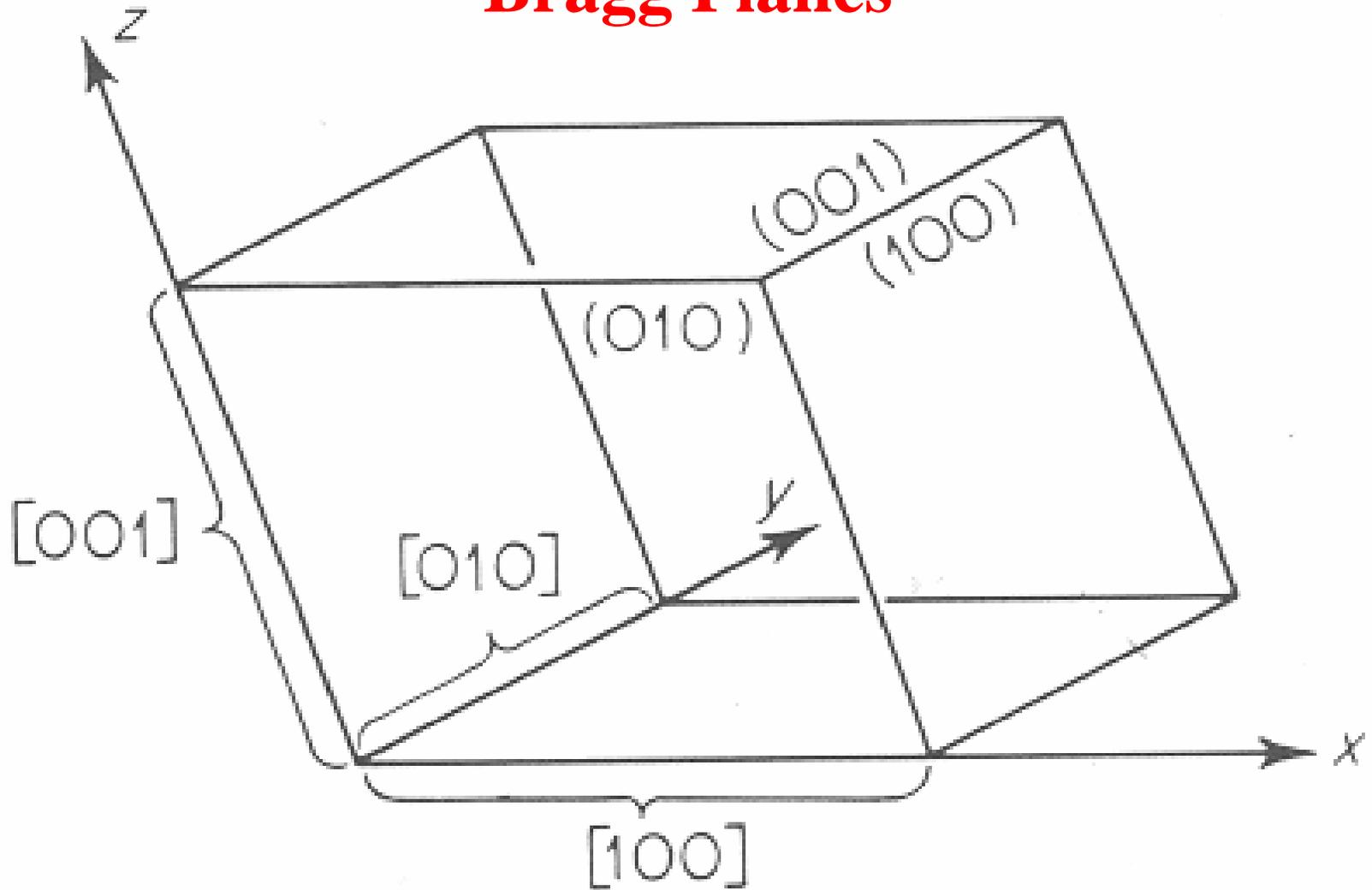


Figure 2.7. Unit cell showing bounding planes and edges.

Name that Bragg "plane"

1 1 0

1 3 0

2 -1 0
(or -2 1 0)

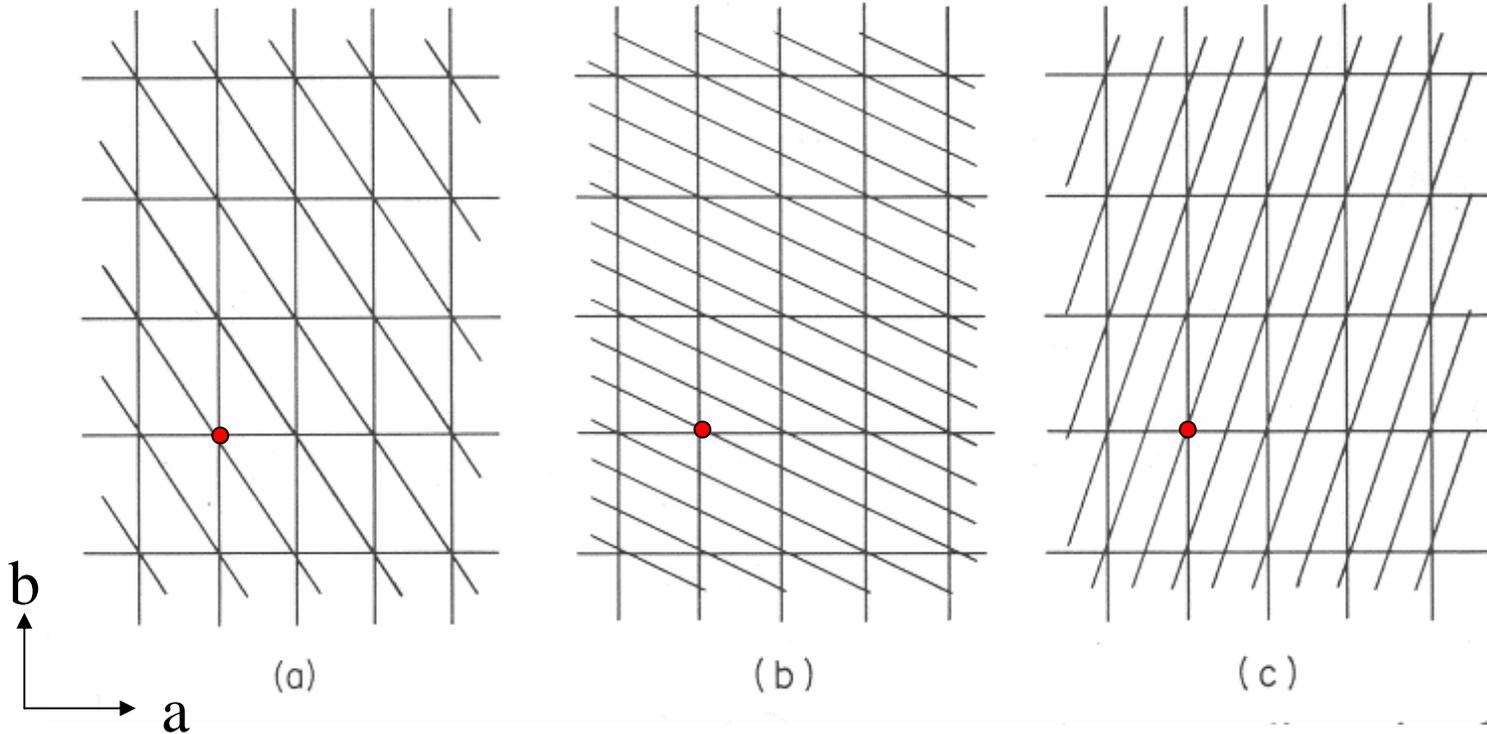


Figure 2.5. Three families of lattice "planes" in a two-dimensional lattice.

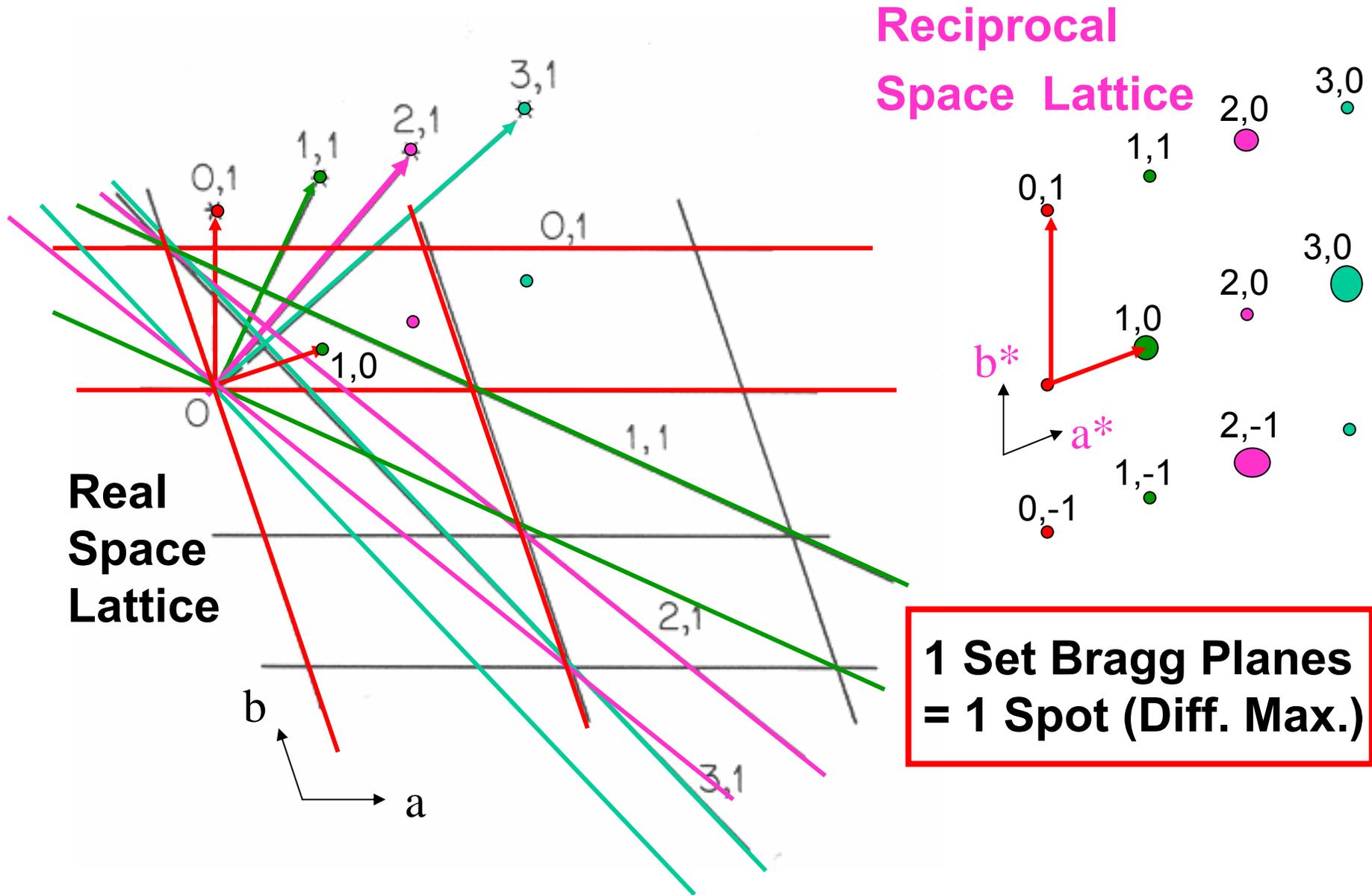
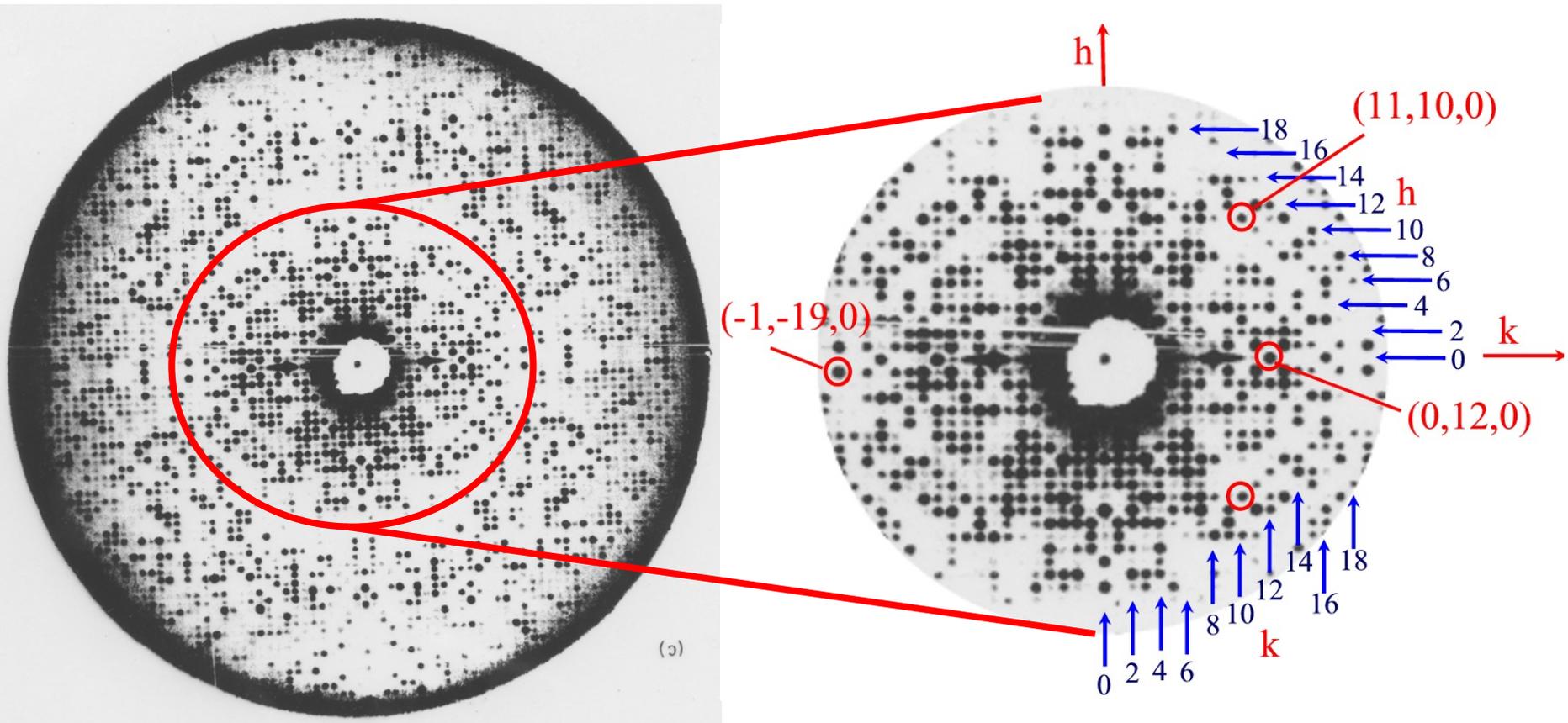


Figure 2.11. **Planes in direct space** represented by **points in reciprocal space**.

Electron Density Function

$$\rho(X, Y, Z) = \frac{1}{V} \sum_h \sum_k \sum_l \underline{F(hkl)} \underline{\exp[i\alpha(hkl)]} \exp[-2\pi i(hX + kY + lZ)]$$



Measure thousands of **Amplitudes** - $[F_{hkl}]$'s - ?? How do we obtain **Phases** α_{hkl} ??

→ **Phase Problem**

X-Ray Crystallography

“If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist.” – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography (after NMR)

a) Crystal Growth – Materials / Methods

b) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

c) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

d) Theory of Diffraction – Bragg's Law / Reciprocal Space

e) Data Collection – Methods / Detectors / Structure Factors

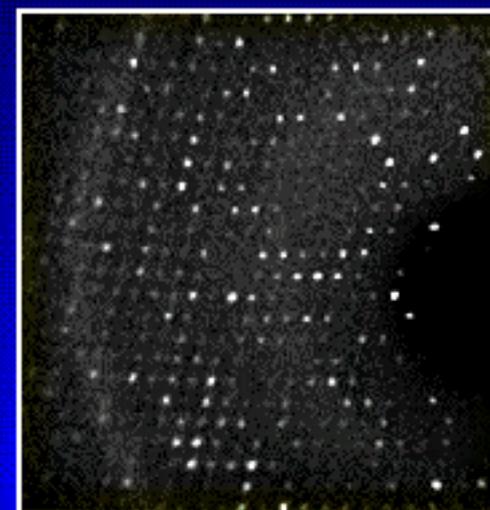
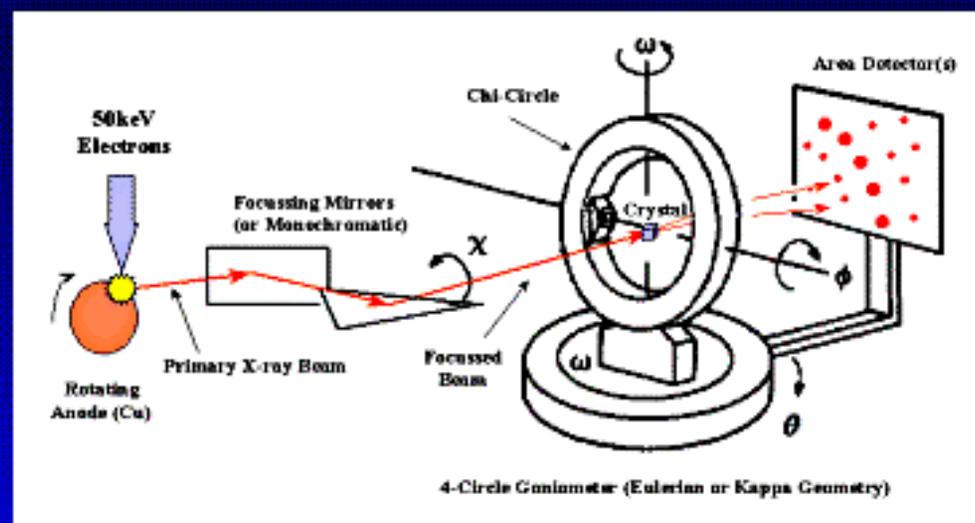
f) Structure Solution – Phase Problem: MIR / MR / MAD

h) Refinement, Analysis and Presentation of Results

i) Use of Difference Fourier

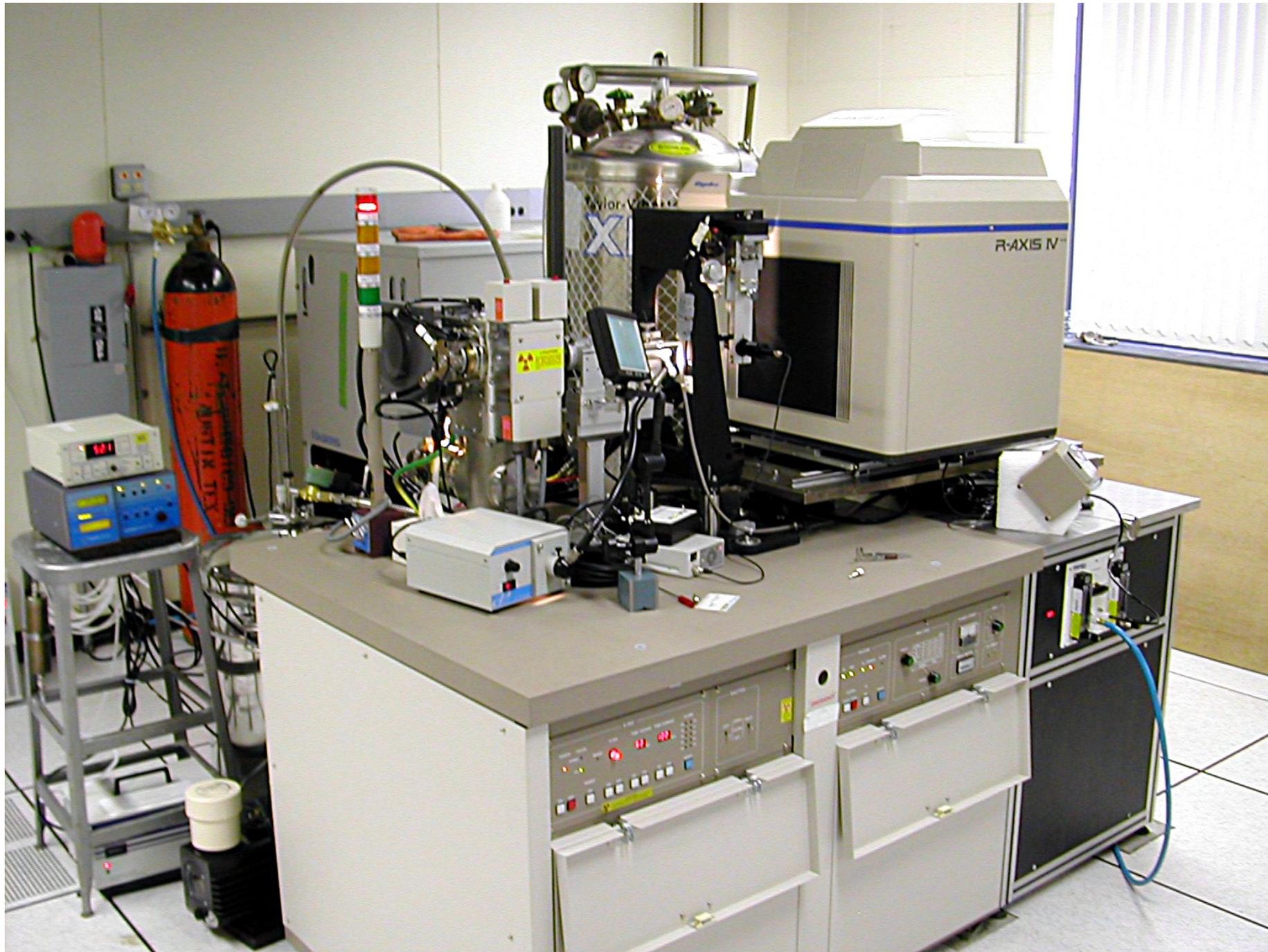


The information we get from a single diffraction experiment.....



The reflections are indexed (consistent assignment of reciprocal cell indices h,k,l) and all we get for the money is a long list of intensities from several ten thousand reflections

2	10	1	326.58
3	10	1	1644.72
4	10	1	3228.45
5	10	1	1279.83
6	10	1	320.48
7	10	1	775.63
8	10	1	1344.55
9	10	1	431.73
10	10	1	1760.14
11	10	1	709.18
12	10	1	20.37
13	10	1	408.72
14	10	1	51.36
15	10	1	114.72
16	10	1	776.26
17	10	1	87.57
18	10	1	30.93
0	11	1	99.30
1	11	1	2258.68
2	11	1	770.18

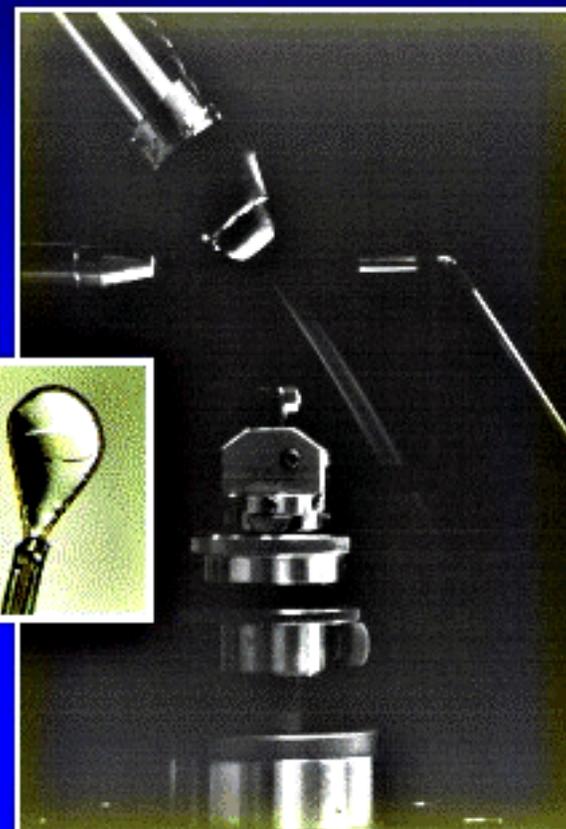




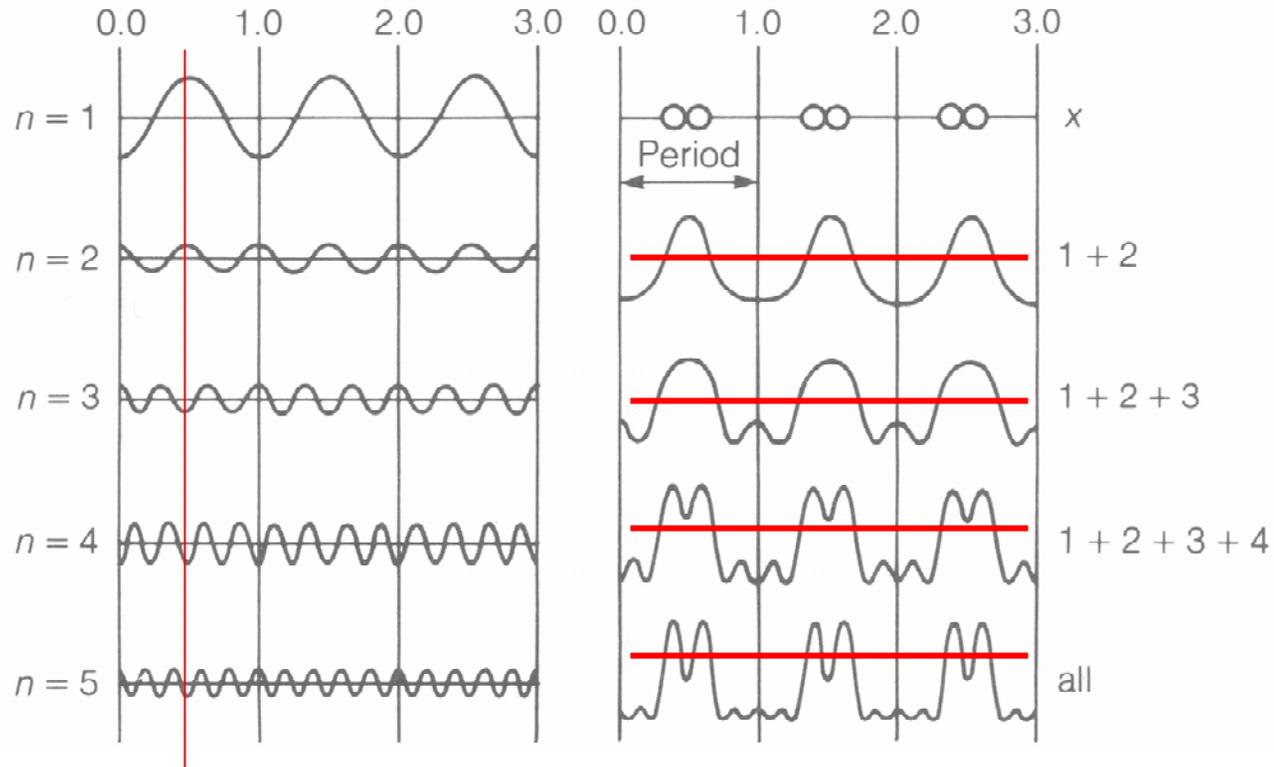
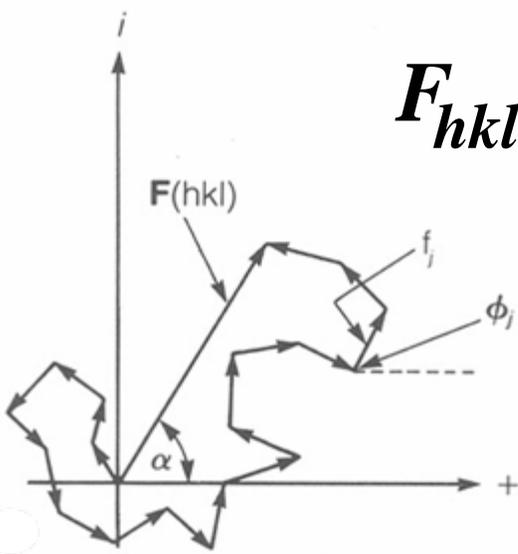
Cryo-cooling efficiently improves data quality



- Crystals are rapidly cooled (**NOT FROZEN**) to near liquid nitrogen temperature
- Reduced thermal vibrations
- **Increased resolution**
- Reduced disorder
- **Eliminated radiation damage**
- No merging and scaling errors

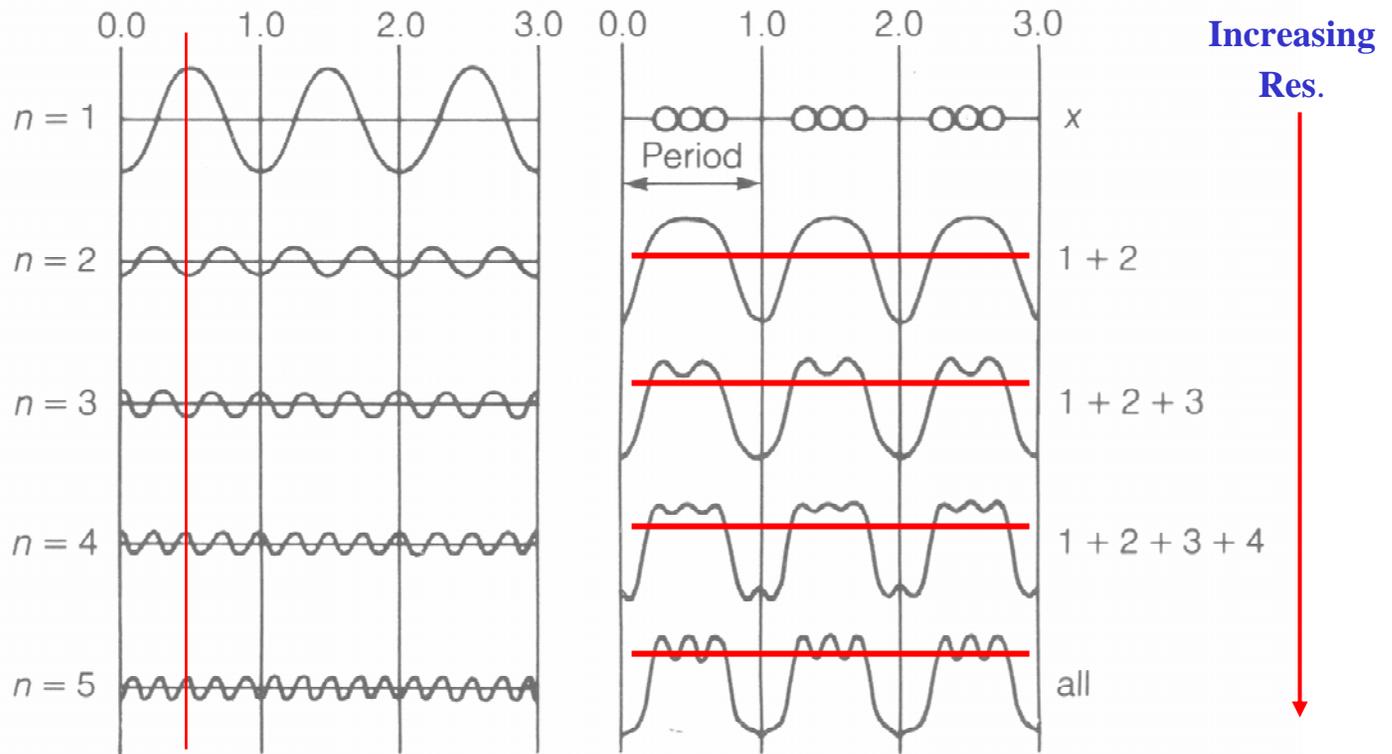


F_{hkl} : Amplitudes and Phases



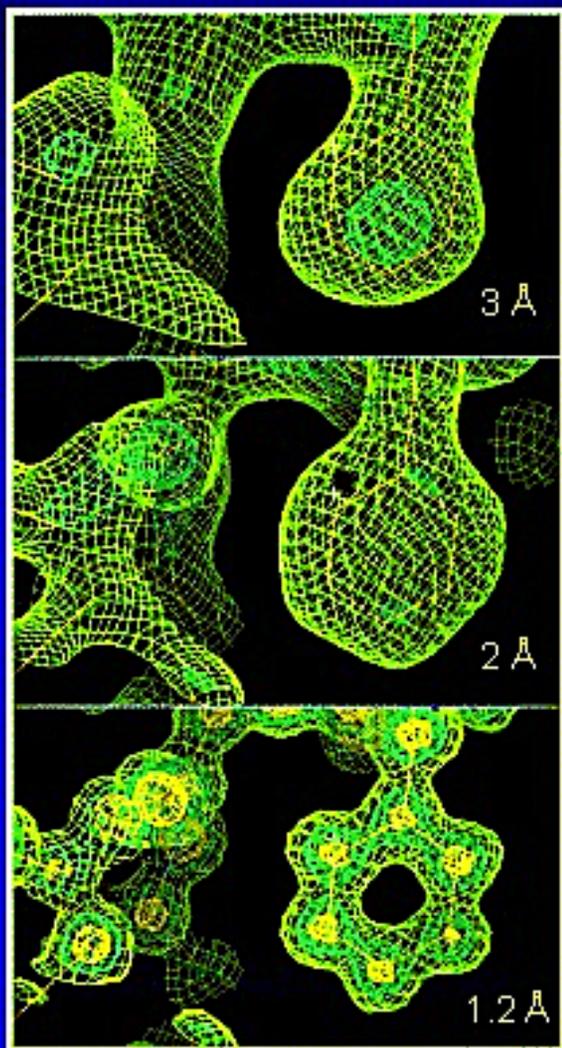
Representation of the electron density of a one-dimensional "crystal" by a superposition of waves. The crystal is formed by a periodic repetition of a diatomic molecule, as shown at the top of the right-hand column. The component waves, each with proper phase and amplitude, are on the left. The curves on the right show the successive superposition of the five waves on the left. (From Waser, 1968.)

F_{hkl} : Amplitudes and Phases

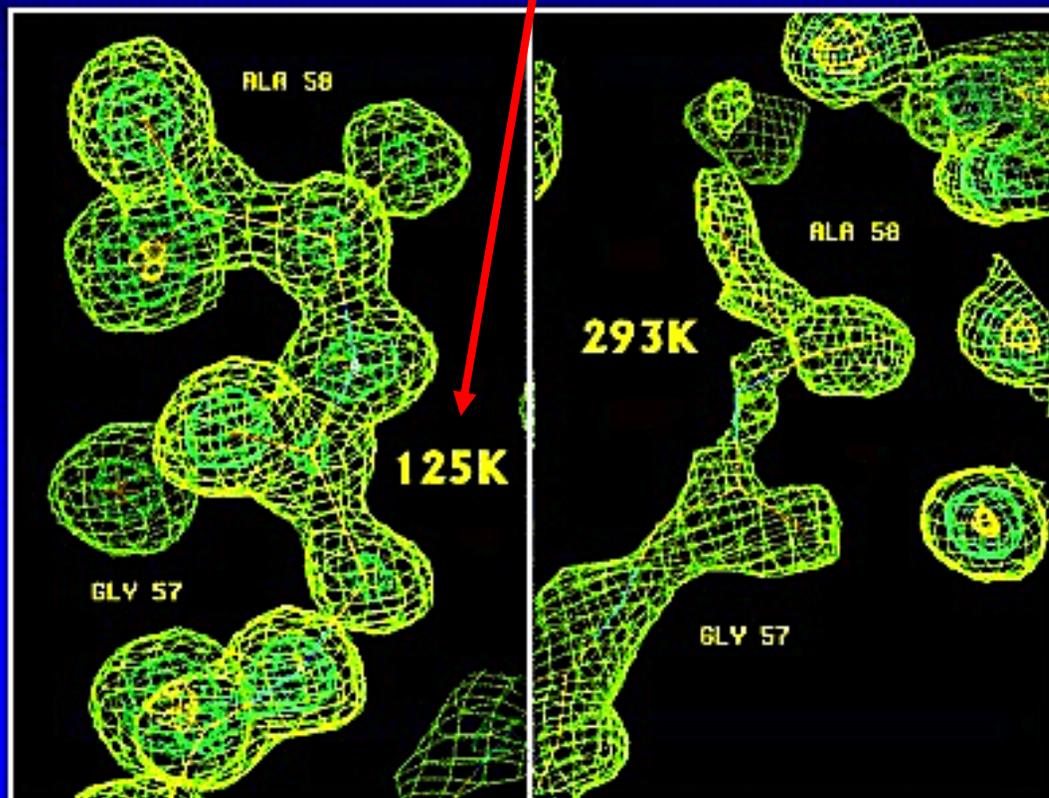


Representation of another one-dimensional crystal, this one containing a triatomic molecule. Note that this crystal is built up from the same waves as the crystal of (a) ; only the amplitudes and phases have been changed. (From Waser, 1968.)

Effect of Resolution



Reduced Disorder at Lower Temperatures



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

X-Ray Crystallography

“If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist.” – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography (after NMR)

a) Crystal Growth – Materials / Methods

b) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

c) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

d) Theory of Diffraction – Bragg's Law / Reciprocal Space

e) Data Collection – Methods / Detectors / Structure Factors

f) Structure Solution – Phase Problem: MIR / MR / MAD

h) Refinement, Analysis and Presentation of Results

i) Use of Difference Fourier

Solving the Phase Problem

1. **MIR:** Multiple Isomorphous Replacement (Heavy Atom)
2. **MR:** Molecular Replacement
3. **MAD:** multiwavelength anomolous dispersion

Use of Heavy Metal Ions for Phasing by MIR Methods

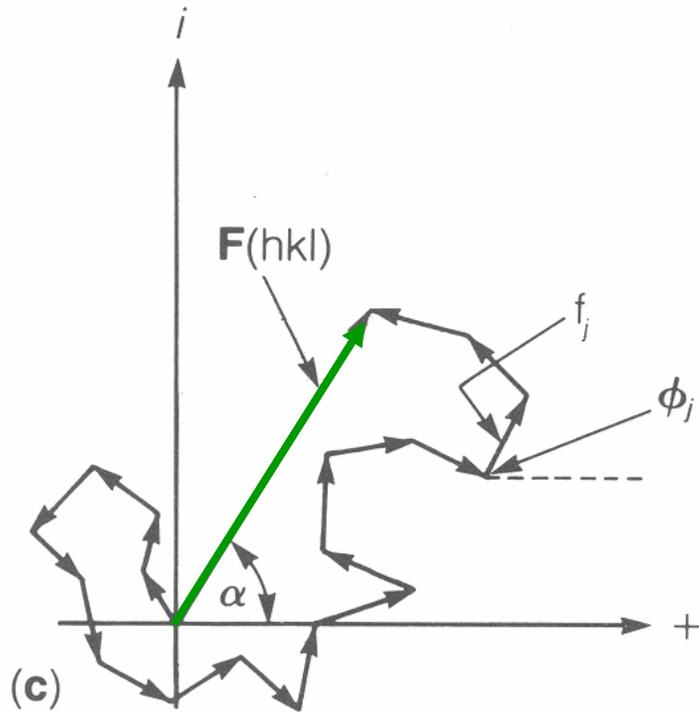


Native Phosphorylase

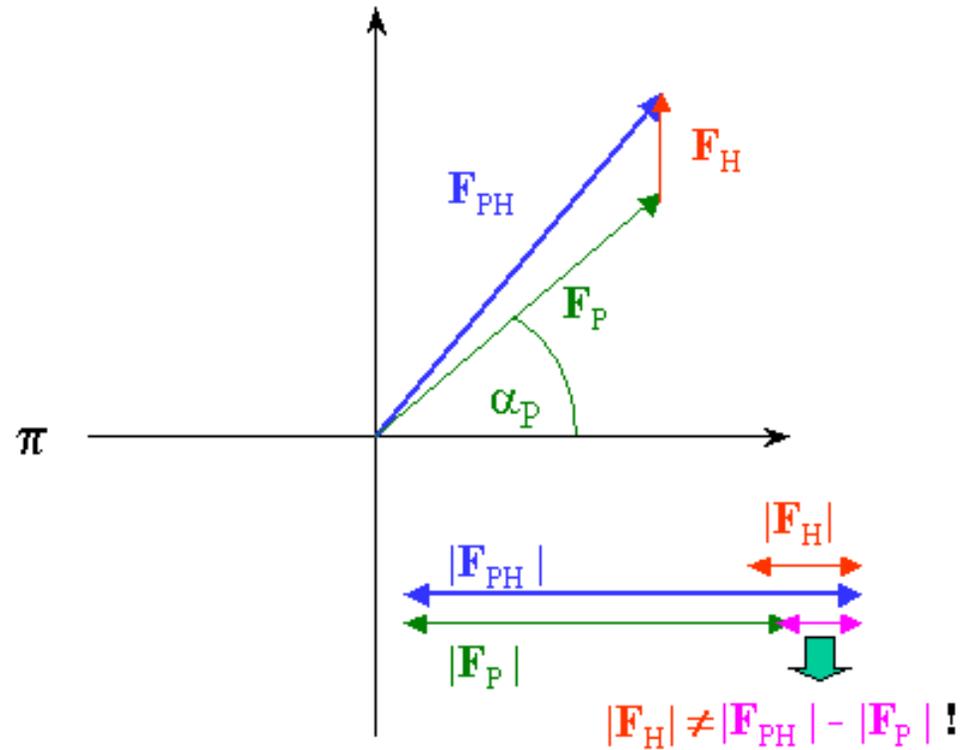


Phosphorylase + Ethyl
Hg thiosalicylate

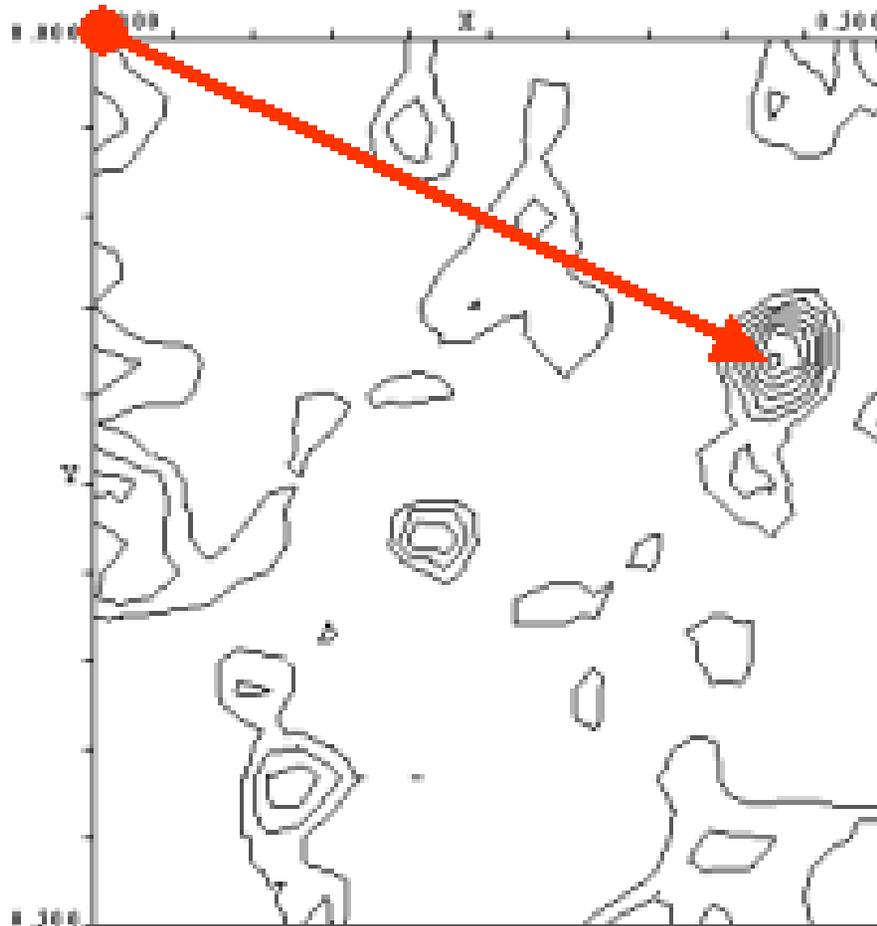
Effect of adding 1 “heavy” atom with lots of electrons!



$$F_{PH} = F_P + F_H$$



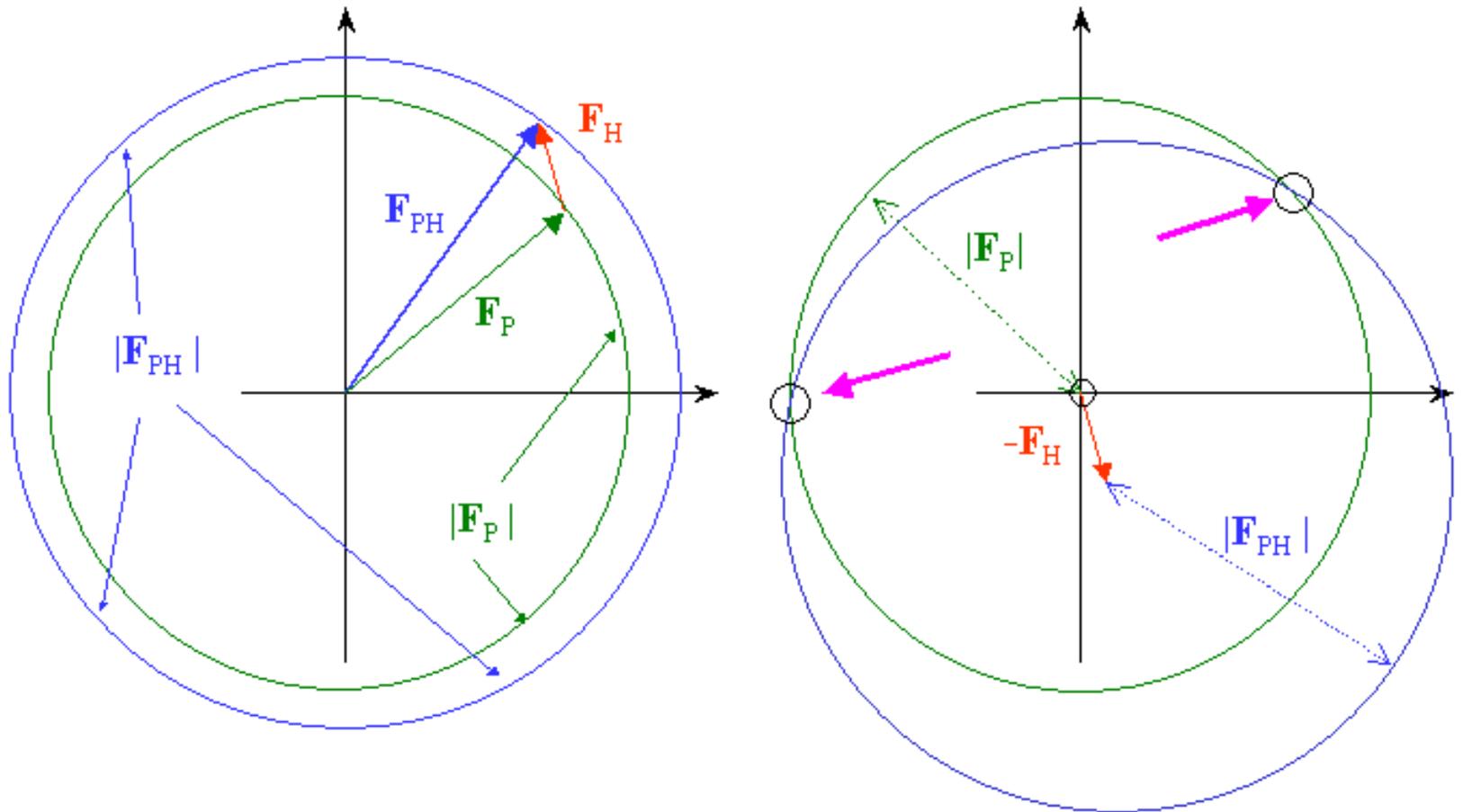
Difference Patterson Map $(F_{hkl})^2$



Multiple Isomorphous Replacement (MIR) method

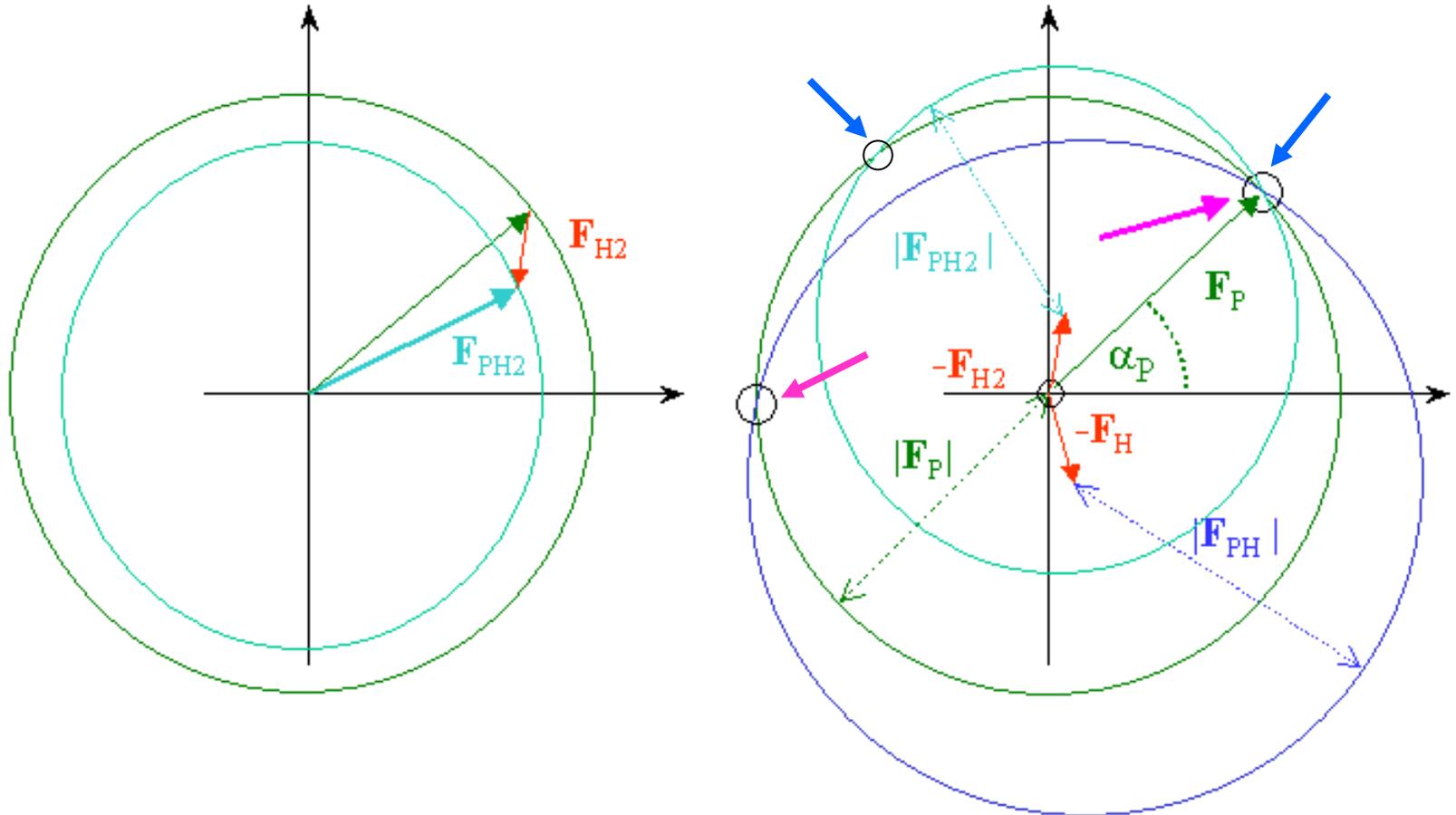
$$F_{PH} = F_P + F_H$$

$$F_P = F_{PH} - F_H$$



Multiple Isomorphous Replacement (MIR) method

$$F_P = F_{PH} - F_H$$



Solving the phase problem by **“Molecular Replacement”**.

If an approximate model of the protein structure is known in advance, approximate phases can be guessed, and the unknown parts of the structure can be calculated in an iterative procedure.

No heavy atom derivative required.

BUT – need starting model and orientation (rotation and translation)

For example, molecular replacement can be used to determine the structure of an **complex with inhibitor** bound to an enzyme active site, if the structure of the enzyme itself is already known. Also, MR is often used to solve the structures of **closely related proteins** in a superfamily.

"Multiwavelength Anomalous Dispersion"

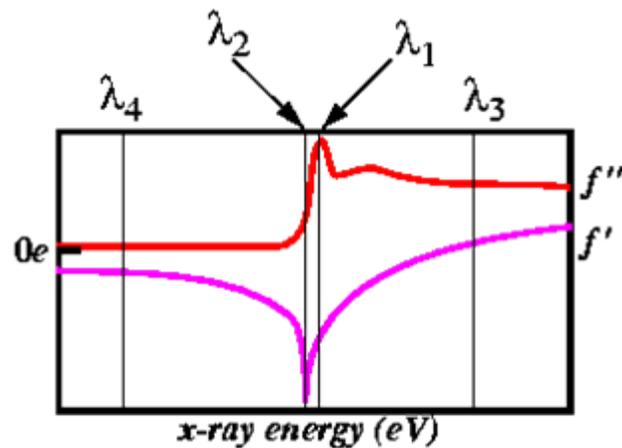
(MAD) methods

Additional information used in calculating phases can be obtained if x-ray diffraction intensities can be measured at **wavelengths near the absorption edge** of the heavy atom derivative.

A **tunable x-ray source** is required (provided by a **synchrotron**). In a synchrotron, accelerated electrons traveling near the speed of light emit intense x-rays.

a) often only a single heavy atom derivative is required to solve a structure (**selenomethionine**).

b) it is possible to solve structure of higher molecular weight molecules (such as the ribosome, at MW = 2,500,000).



What should they be?

- The largest signal will come from choosing the wavelength with maximal f'' (λ_1 in the figure above).
- The second wavelength is usually chosen to have maximal $|f''|$ (λ_2 in the figure above). Note that (1 and 2) are very close together, requiring great precision in setting up the apparatus which controls wavelength during data collection.
- Additional wavelengths (3 and 4) are chosen at points remote from the absorption edge. The available signal increasing slowly as the distance from the first two wavelengths increases. However the diffraction conditions (crystal absorption and diffracting power, diffraction geometry, etc) become more disparate as the distance increases. The choice usually comes down to the practical limitations imposed by the particular beamline apparatus being used. Typically λ_3 and λ_4 are between 100eV and 1000eV from the absorption edge.

X-Ray Crystallography

“If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist.” – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography (after NMR)

a) Crystal Growth – Materials / Methods

b) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

c) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

d) Theory of Diffraction – Bragg’s Law / Reciprocal Space

e) Data Collection – Methods / Detectors / Structure Factors

f) Structure Solution – Phase Problem: MIR / MR / MAD

h) Refinement, Analysis and Presentation of Results

i) Use of Difference Fourier

Least-Squares Refinement

$$\sum_{r=1}^m w_r \left(\frac{\partial |kF_{c,r}|}{\partial p_1} \right)^2 \Delta p_1 + \sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_1} \frac{\partial |kF_{c,r}|}{\partial p_2} \Delta p_2 + \dots$$

$$+ \sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_1} \frac{\partial |kF_{c,r}|}{\partial p_n} \Delta p_n = \sum_{r=1}^m w_r \Delta F_r \frac{\partial |kF_{c,r}|}{\partial p_1}$$

$$\sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_2} \frac{\partial |kF_{c,r}|}{\partial p_1} \Delta p_1 + \sum_{r=1}^m \left(\frac{\partial |kF_{c,r}|}{\partial p_2} \right)^2 \Delta p_2 + \dots$$

$$+ \sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_2} \frac{\partial |kF_{c,r}|}{\partial p_n} \Delta p_n = \sum_{r=1}^m w_r \Delta F_r \frac{\partial |kF_{c,r}|}{\partial p_2}$$

⋮

$$\sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_n} \frac{\partial |kF_{c,r}|}{\partial p_1} \Delta p_1 + \sum_{r=1}^m w_r \frac{\partial |kF_{c,r}|}{\partial p_n} \frac{\partial |kF_{c,r}|}{\partial p_2} \Delta p_2 + \dots$$

$$+ \sum_{r=1}^m w_r \left(\frac{\partial |kF_{c,r}|}{\partial p_n} \right)^2 \Delta p_n = \sum_{r=1}^m w_r \Delta F_r \frac{\partial |kF_{c,r}|}{\partial p_n}$$

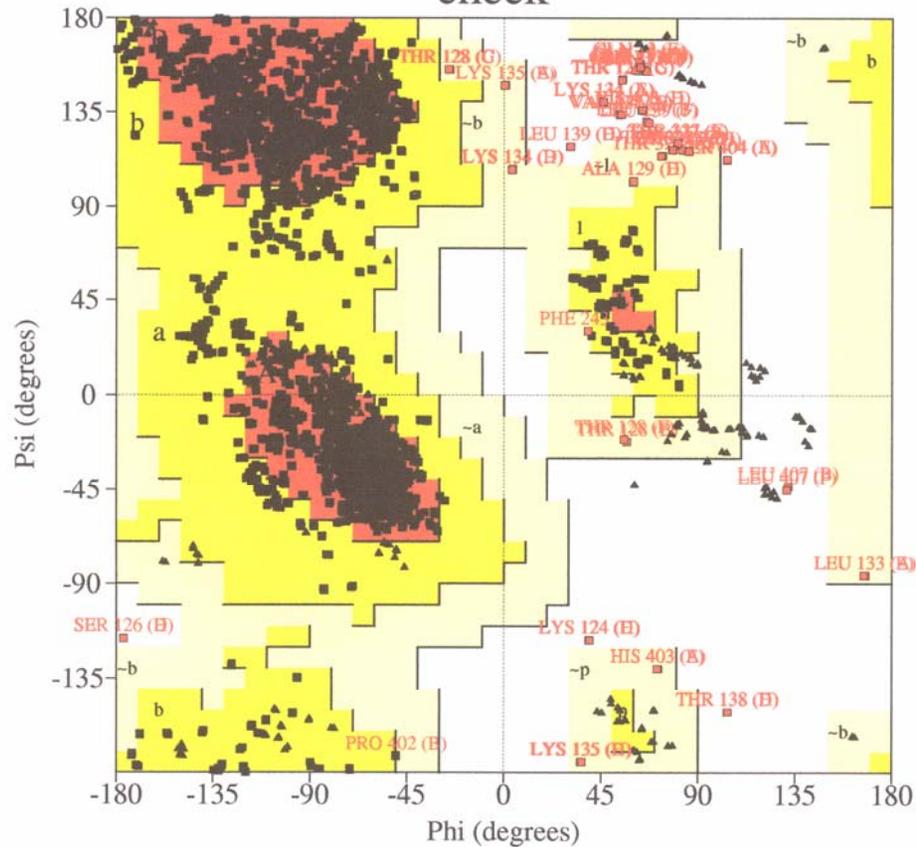
Energy Refinement (Simulated Annealing)

$$E_{TOTAL} = E_{EMPIRICAL} + E_{EFFECTIVE}$$

$$E_{EFFECTIVE} = E_{XREF} + E_{NOE} + E_{HARM} + \\ E_{CDIH} + E_{NCS} + E_{DG} + E_{RELA} + E_{PLAN}$$

$$E_{EMPIRICAL} = \sum_{p=1}^N [w^p_{BOND} E_{BOND} + w^p_{ANGL} E_{ANGL} + \\ w^p_{DIHE} E_{DIHE} + w^p_{IMPR} E_{IMPR} + \\ w^p_{VDW} E_{VDW} + w^p_{ELEC} E_{ELEC} + \\ w^p_{PVDW} E_{PVDW} + w^p_{PELE} E_{PELE} + \\ w^p_{HBON} E_{HBON}].$$

Ramachandran Plot check



Plot statistics

Residues in most favoured regions [A,B,L]	3149	85.9%
Residues in additionally allowed regions [a,b,l,p]	460	12.6%
Residues in generously allowed regions [-a,-b,-l,-p]	24	0.7%
Residues in disallowed regions	31	0.8%
Number of non-glycine and non-proline residues	3664	100.0%
Number of end-residues (excl. Gly and Pro)	228	
Number of glycine residues (shown as triangles)	312	
Number of proline residues	168	
Total number of residues	4372	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Crystal Structure of *M. tuberculosis* Alanine Racemase

Table 1: Data Collection and Processing Statistics for the MAD and Native Data Sets of Alr_{Mtb}

	MAD 1	MAD 2	MAD 3	MAD 4	native
λ (Å)	0.9788	0.9790	0.9562	0.9809	0.9160
resolution (Å)		2.20			1.80
mosaicity		0.50			0.65
no. of reflections observed $> 1\sigma$	432376	446744	431524	336135	779600
no. of unique reflections $> 1\sigma$	35817	37506	36020	36242	67592
R_{merge}^a (%)	6.9	6.4	5.1	3.7	6.0 (67.2)
completeness (%)	91.8	95.8	92.1	92.1	99.3 (95.6)
$\langle I/\sigma \rangle$	30.3	34.3	41.6	50.9	34.5 (2.6)

$$^a R_{\text{merge}} = \sum |I_{\text{obs}} - I_{\text{avg}}| / \sum |I_{\text{avg}}|$$

Biochemistry 2005, 44, 1471–1481

Table 2: Final Refinement Statistics for Alr_{Mtb} at 1.9 Å Resolution

R factor ^a (%)	20.4
R_{free} (%) (for 1747 reflections)	25.4
average B factor (Å ²) ^b	
main chain	25.5
side chain	31.5
PLP	21.9
waters	32.4
rms deviations	
bond lengths (Å)	0.006
bond angles (deg)	1.9
no. of reflections $> 2\sigma$	55001
no. of residues	722
no. of protein atoms	5360
no. of PLP atoms	30
no. of water molecules	350

$$^a R\text{-factor} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum |F_{\text{obs}}|. \quad ^b \text{All isotropic model.}$$

The 1.9 Å Crystal Structure of Alanine Racemase from *Mycobacterium tuberculosis* Contains a Conserved Entryway into the Active Site^{1,‡}

Pierre LeMagueres,[§] Hookang Im,[§] Jerry Ebalunode,[§] Ulrich Strych,[§] Michael J. Benedik,[¶] James M. Briggs,[§]
Harold Kohn,[‡] and Kurt L. Krause^{*,§,¶}

Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, Department of Biology,
Texas A&M University, College Station, Texas 77843-3258, Division of Medicinal Chemistry and Natural Products,
School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360, and
Section of Infectious Diseases, Department of Medicine, Baylor College of Medicine, Houston, Texas 77030

Received June 27, 2004; Revised Manuscript Received October 22, 2004

1471

Analyze – structure (Ramachandran Plot) and biochemistry

Publish in leading biochemical or structural biology journal

Contribute results (coordinates, etc.) to PDB

Data Mining

Visualization programs (Cn3D / RasMol / SwissPDBV / etc)

SCOP – Structural Classification of Proteins

CATH – Classification / Arch / Topology

Difference Fourier

Obs. $\rho_o(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{o,hkl} e^{-2\pi i(hx+ky+lz)} + R$

Calc. $\rho_c(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{c,hkl} e^{-2\pi i(hx+ky+lz)} + R'$

$$\rho_o(x, y, z) - \rho_c(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l (F_o - F_c)_{hkl} e^{-2\pi i(hx+ky+lz)} + R - R'$$

$$\rho_o - \rho_c = \frac{1}{V} \sum_h \sum_k \sum_l \Delta F_{hkl} e^{-2\pi i(hx+ky+lz)}$$



Back



Forward



Reload



Home



Search



Netscape



Print



Security



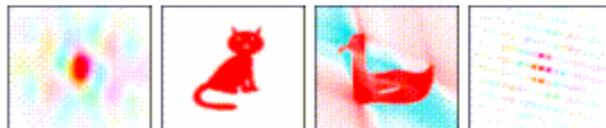
Shop



Stop



Kevin Cowtan's Book of Fourier



This is a book of pictorial 2-d Fourier Transforms. These are particularly relevant to my own field of *X-ray crystallography*, but should be of interest to anyone involved in signal processing or frequency domain calculations.

Contents:

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

- [Introduction](#)
- [Book of Crystallography](#)
- [Duck Tales](#) and missing data.
- [A little Animal Magic](#) and cross phasing.
- [A Tail of Two Cats](#) and image restoration.
- [Animal Liberation](#) and free-sets.

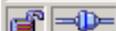
- [The Gallery](#). Other interesting pictures.

Other topics:

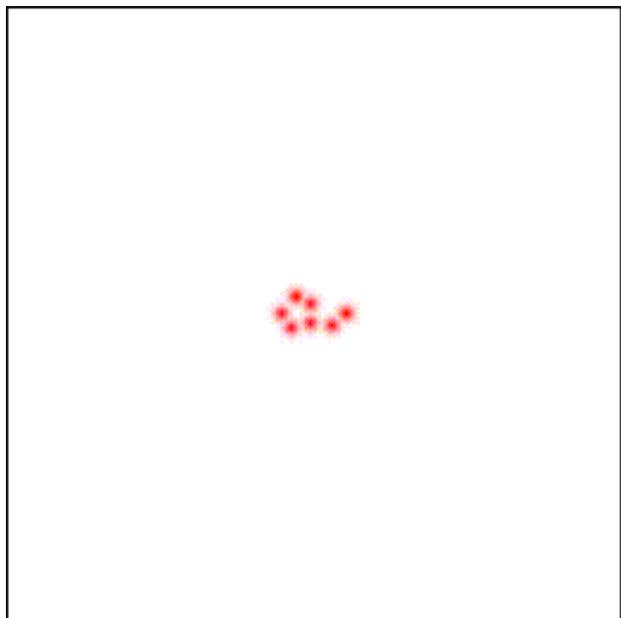
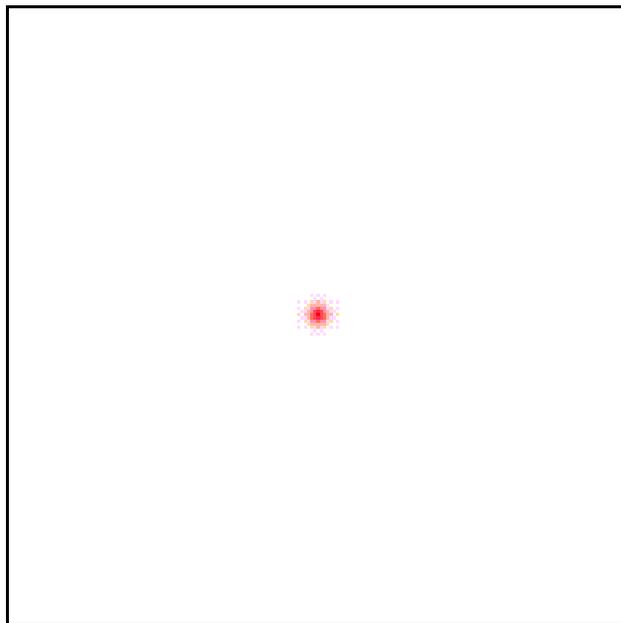
[The Interactive Structure Factor Tutorial](#): Learn about structure factors and maps.

An introduction to crystallographic [Fourier transforms](#). The mathematical link between [Scattering theory](#) and Fourier theory.
An explanation of the [convolution theorem](#).

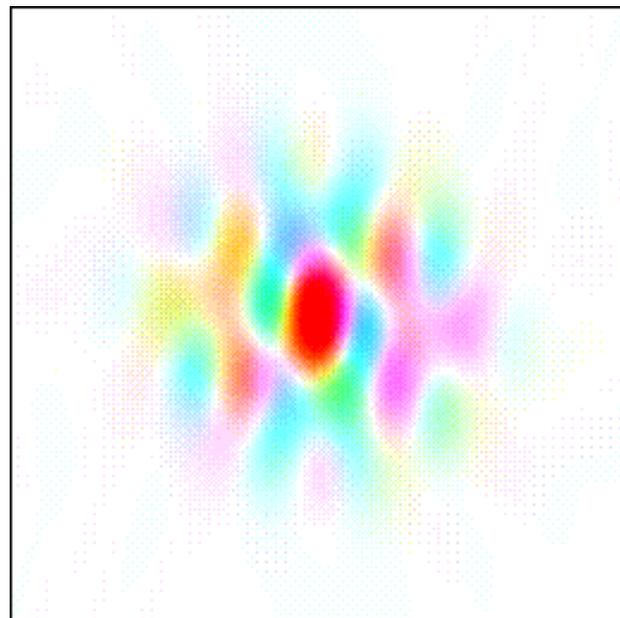
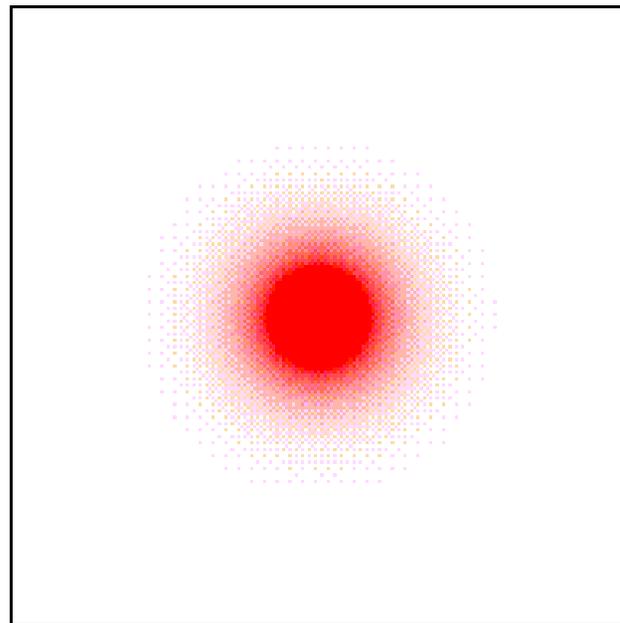
Teaching materials elsewhere



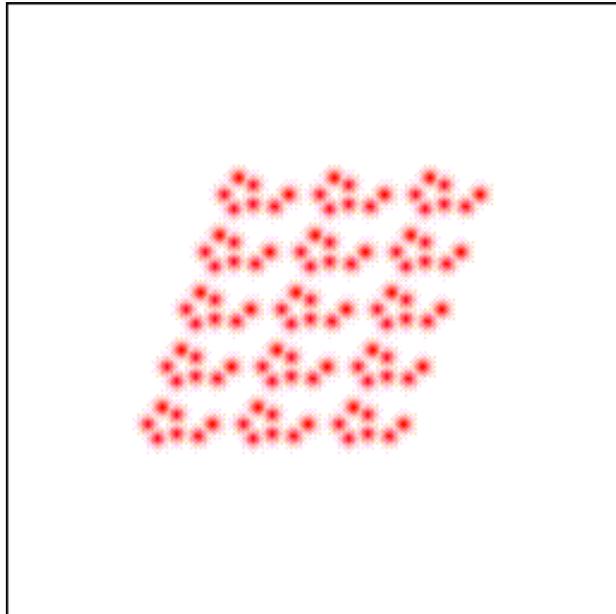
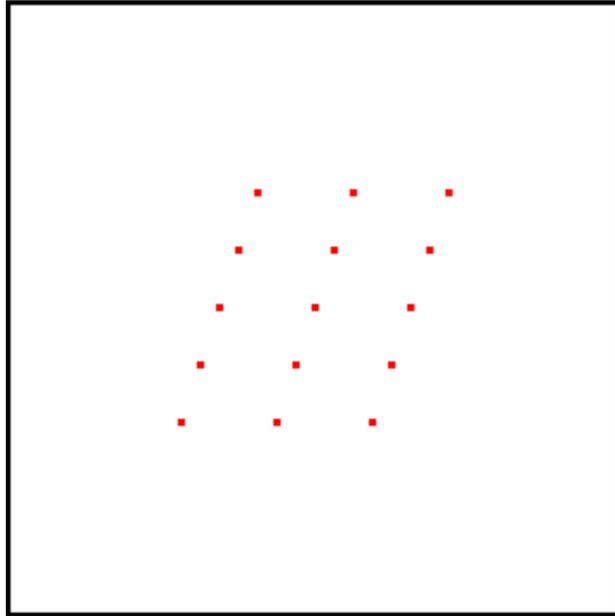
Object / Real Space



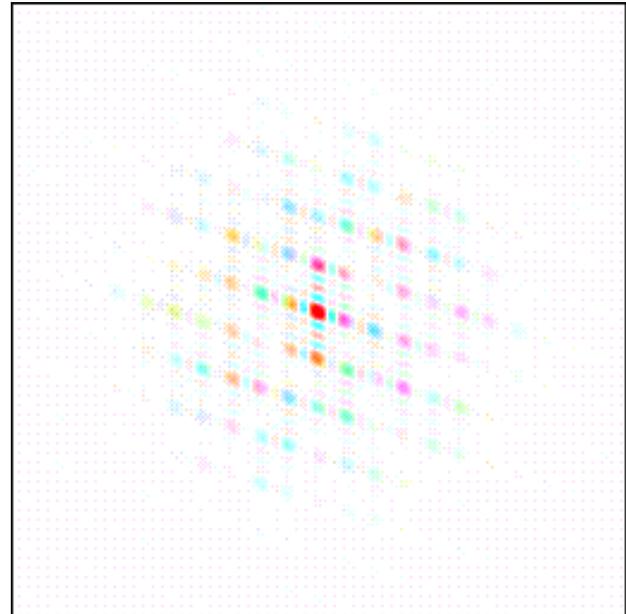
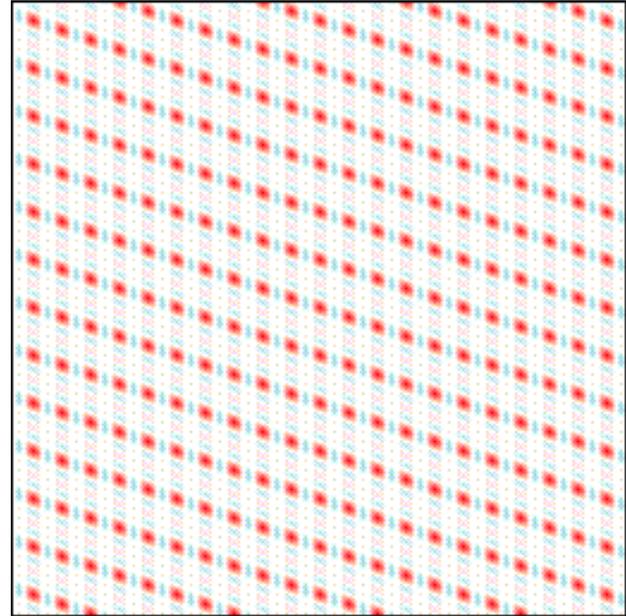
Transform / Reciprocal Space



Object / Real Space



Transform / Reciprocal Space

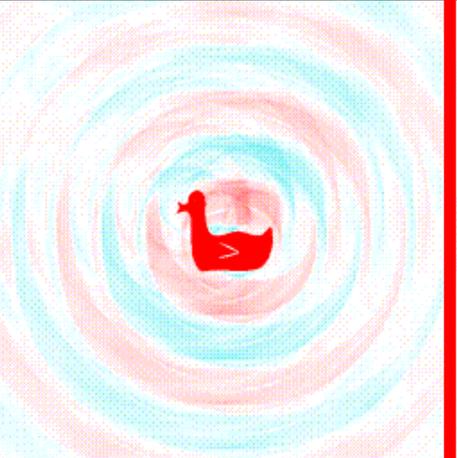
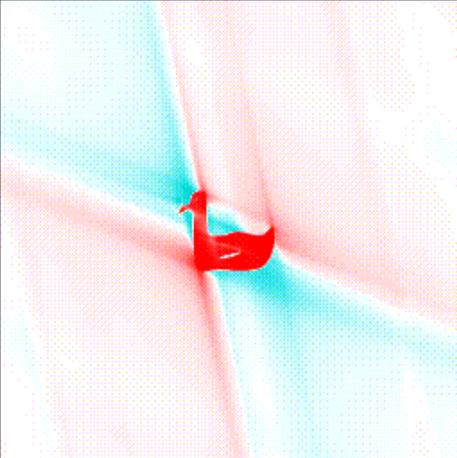
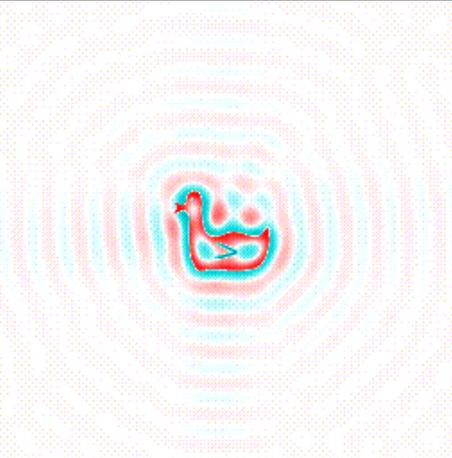
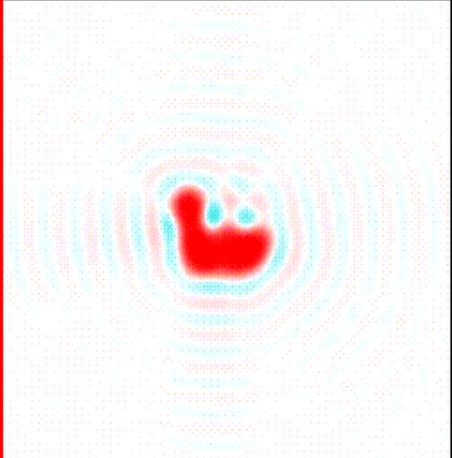
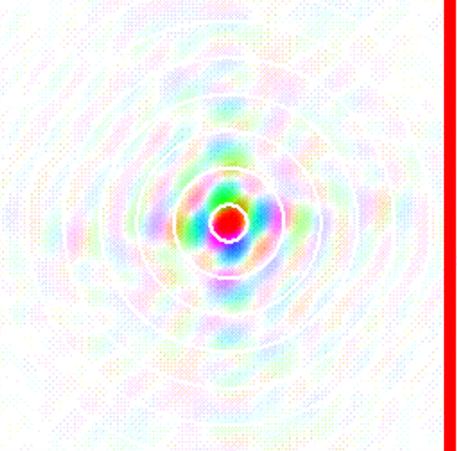
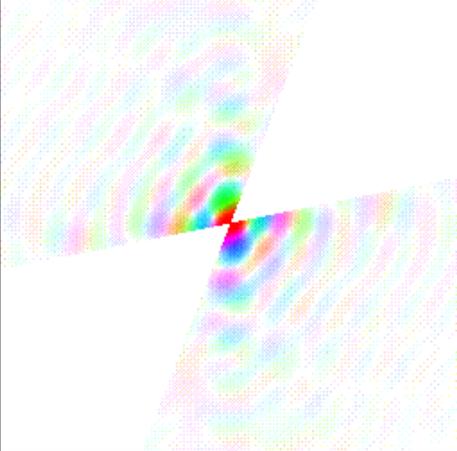
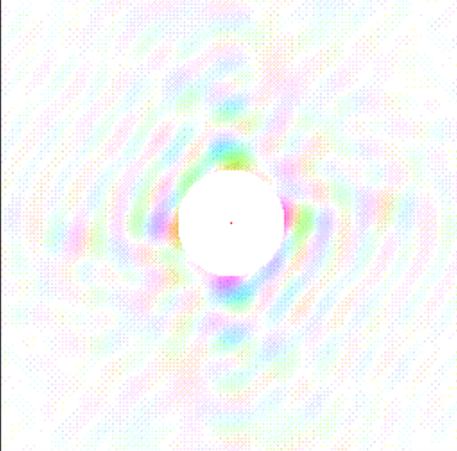
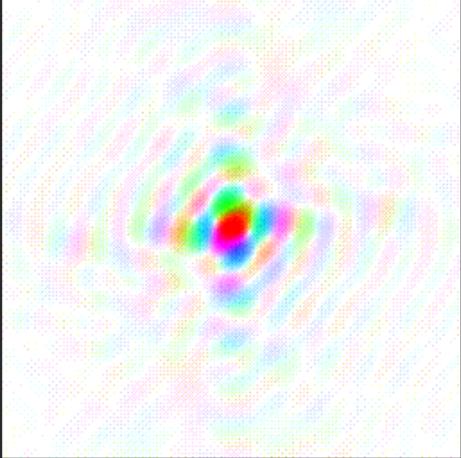


Objects – Transforms and Image Formation

A Duck



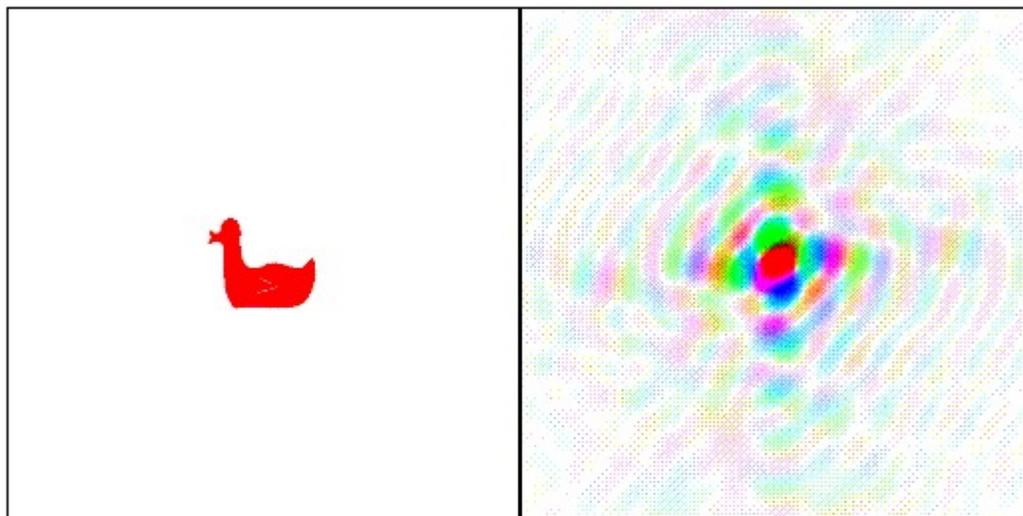
Transform
of a Duck



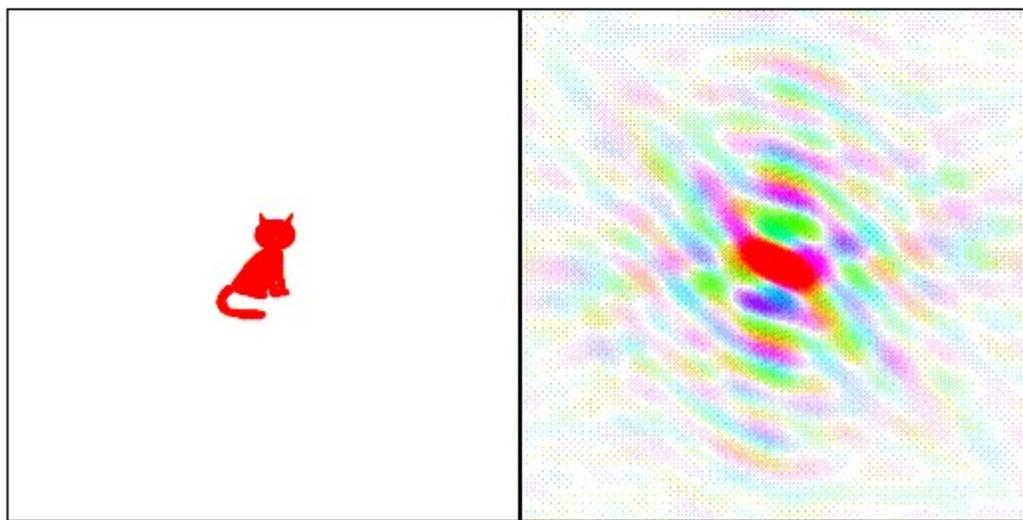
Kevin Cowtan's Book of Fourier

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

Here is our old friend; the Fourier Duck, and his Fourier transform:



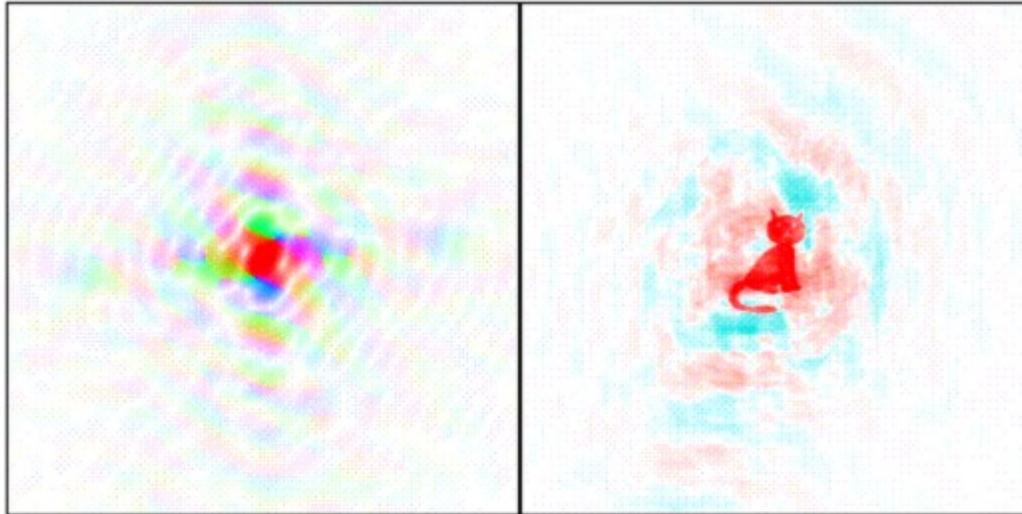
And here is a new friend; the Fourier Cat and *his* Fourier transform:



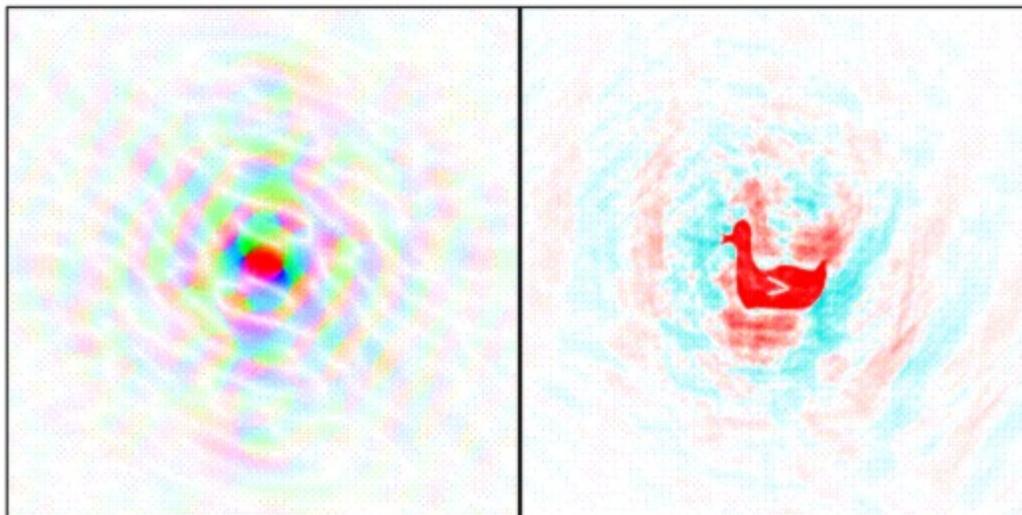
Kevin Cowtan's Book of Fourier

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

Duck Transform Amplitudes + Cat Phases

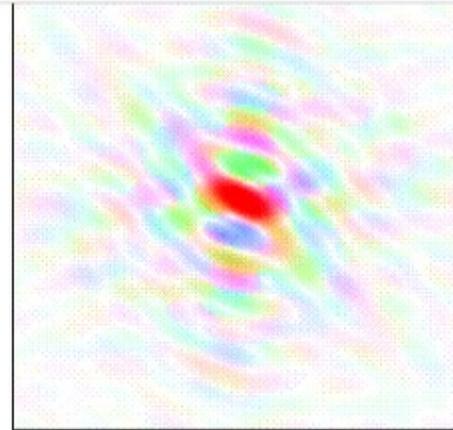
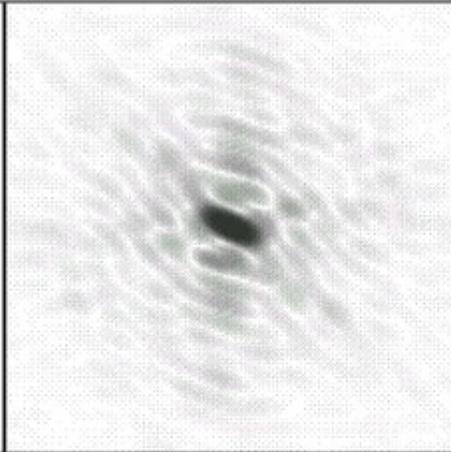


Cat Transform Amplitudes + Duck Phases



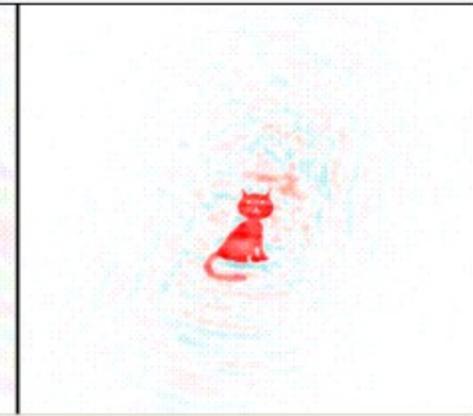
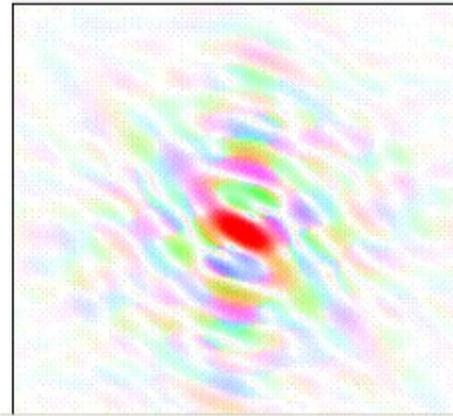
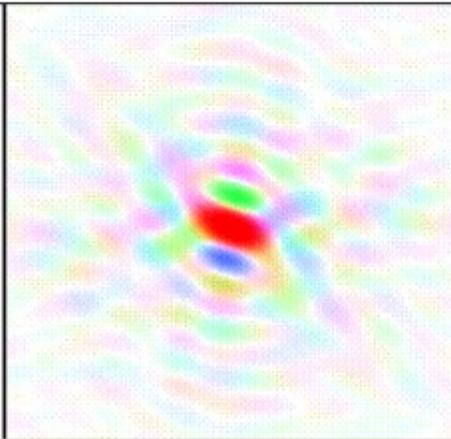
Kevin Cowtan's Book of Fourier

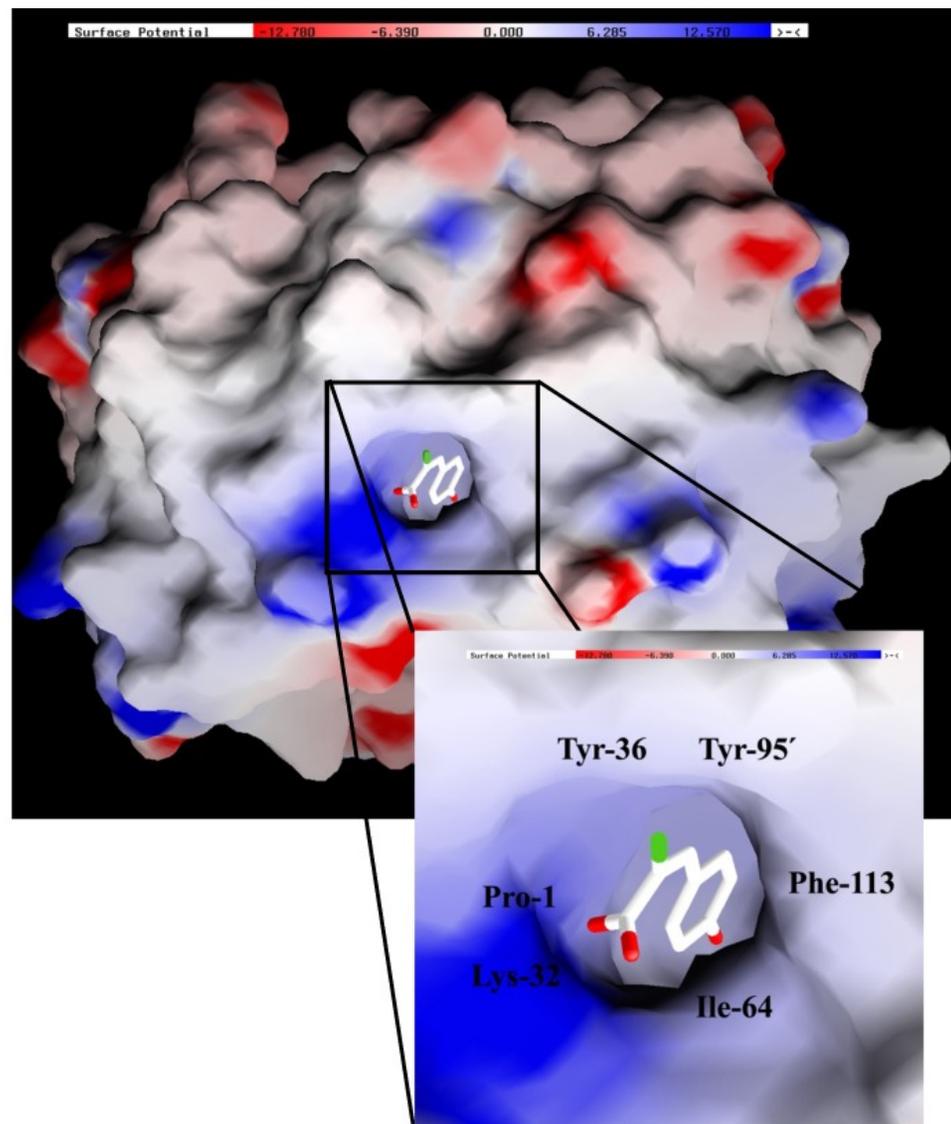
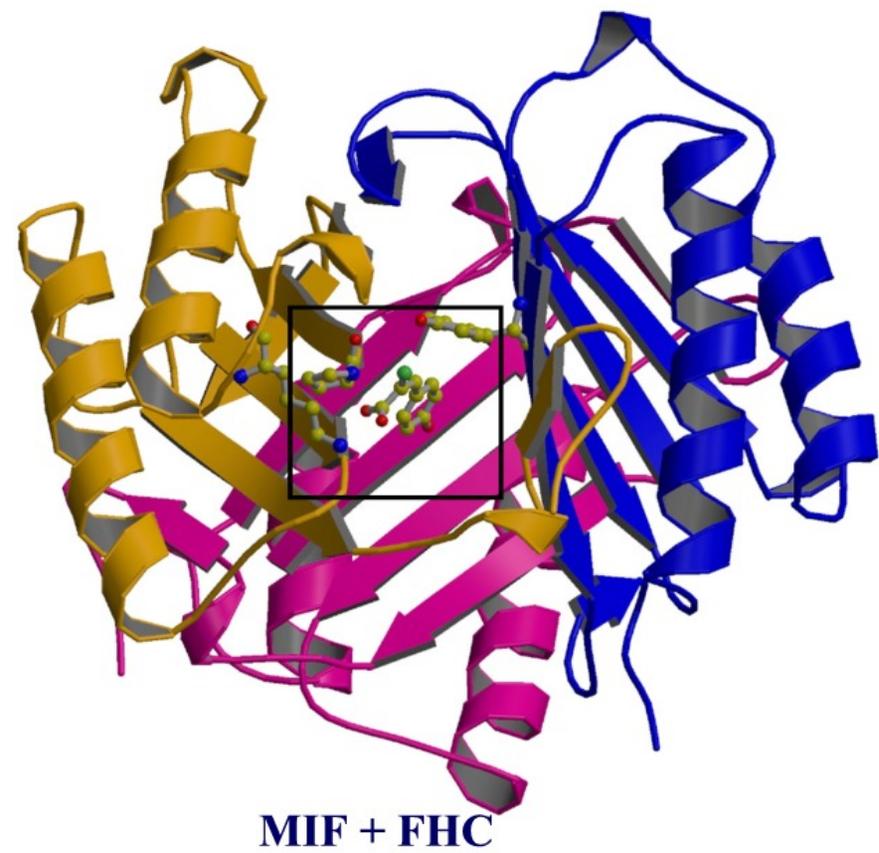
<http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html>



a) Cat - Cat Transform (Amplitudes only)
b) Manx (tailless) Cat - Manx Transform

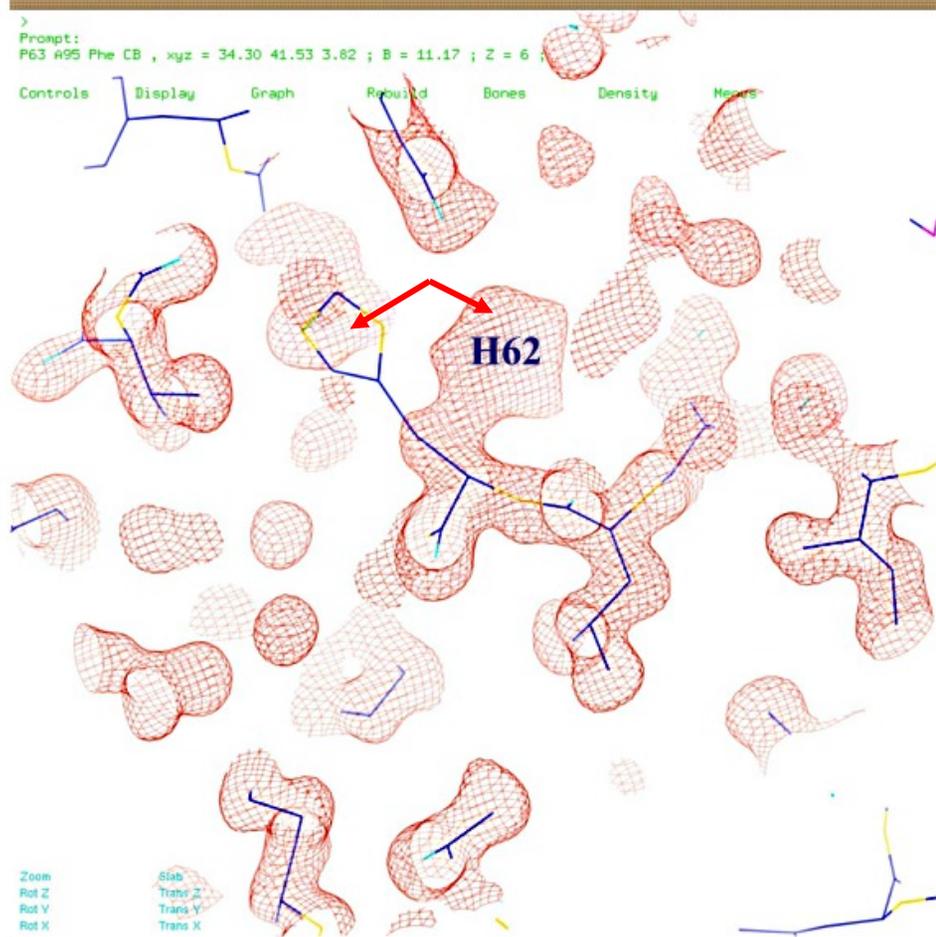
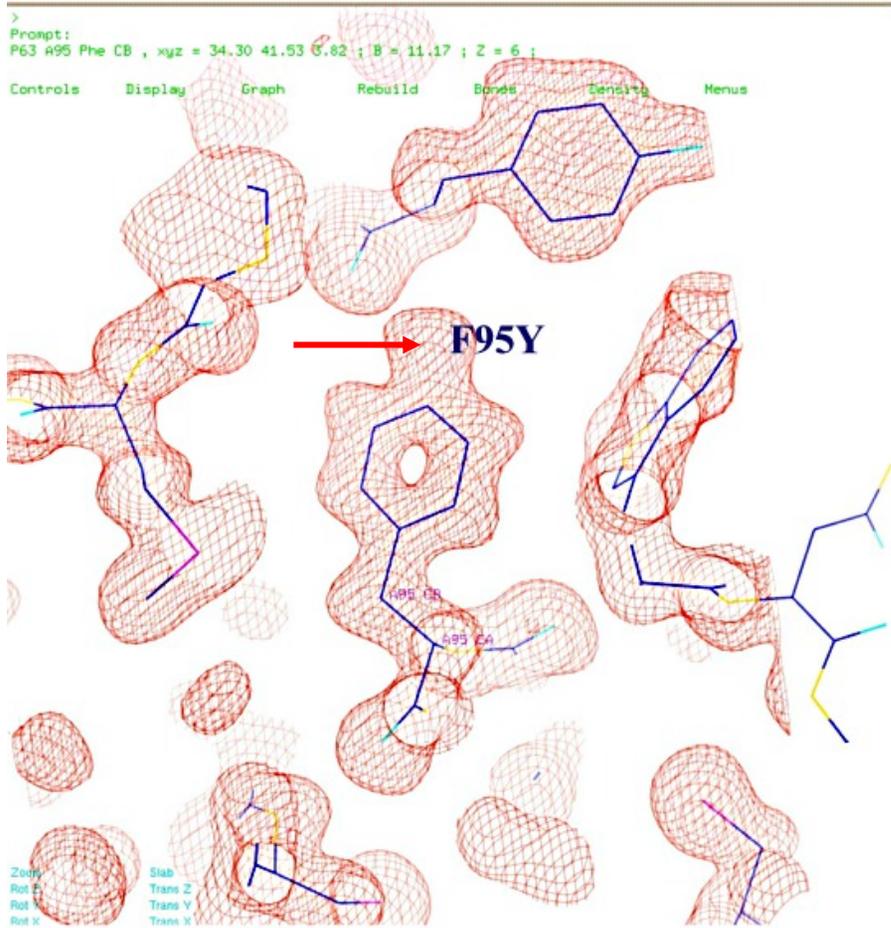
c) Cat Amplitudes + Manx Phases
d) [2x(Cat Amplitudes) - Manx Amplitudes] + Manx Phases





Examples of Difference Fourier

MIF - 1.5Å 2Fo-Fc



X-Ray Crystallography

Quiz questions:

1. Crystal Growth – Materials / Methods

What is the single most important factor that determines crystal growth?

What are the two most common precipitating agents for growing protein crystals?

2. Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

Identify the unit cell, asymmetric unit and symmetry present in the pattern shown.



3. X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

What is responsible for “characteristic” X-rays?

What are the major advantages of using synchrotron radiation?

4. Theory of Diffraction – Bragg’s Law / Reciprocal Space

When collecting an X-ray data set, what is being measured and how is that data useful?

5. Phasing and Refinement

Identify the meaning of the terms: MIR, MR, MAD, Difference Map, Simulated Annealing