Macromolecular crystallography

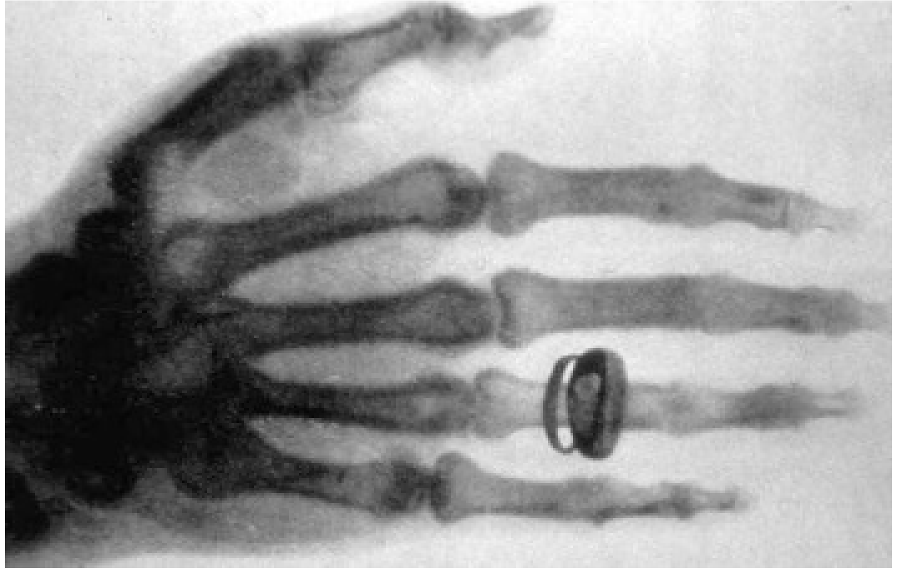
Pavel Plevka

- Development of crystallography
- Waves and radiation
- Diffraction
- Solution of phase problem
- Model building and structure validation

WILHELM CONRAD RÖNTGEN (1845-1923)

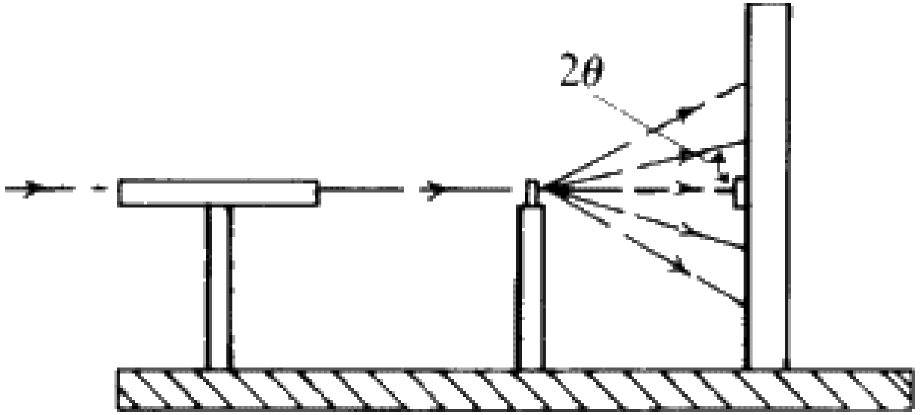
- **1901 Nobel Laureate in Physics**

  discovery of the remarkable rays subsequently named after him.

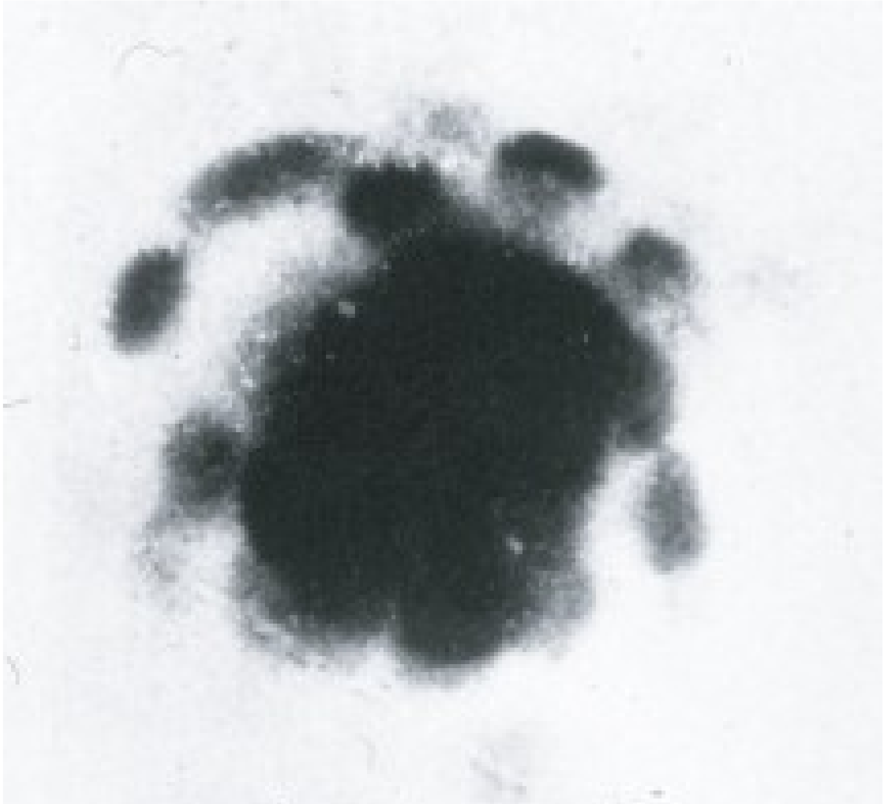


MAX VON LAUE
(1879-1960)

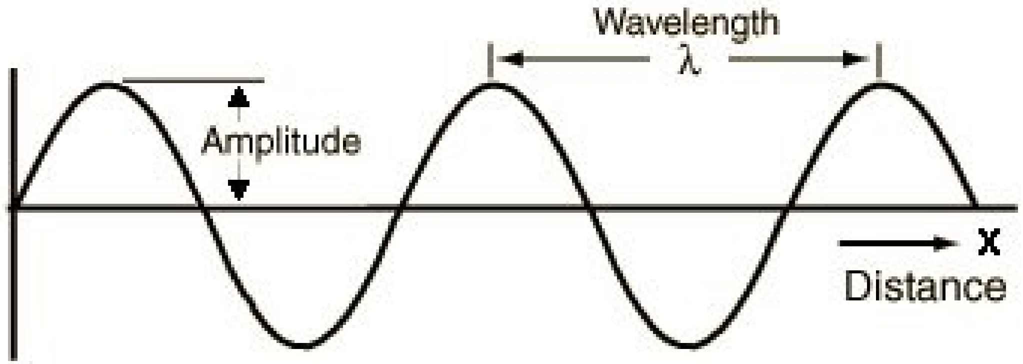
- **1914 Nobel Laureate in Physics**
  for his discovery of the diffraction of X-rays by crystals

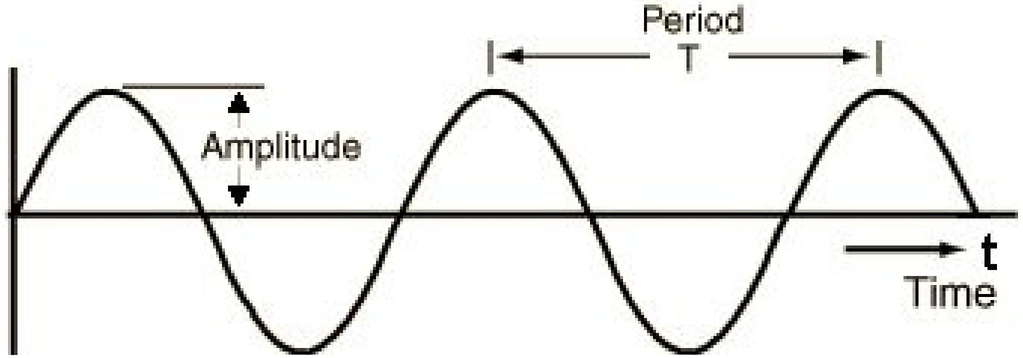


Friedrich and Knipping

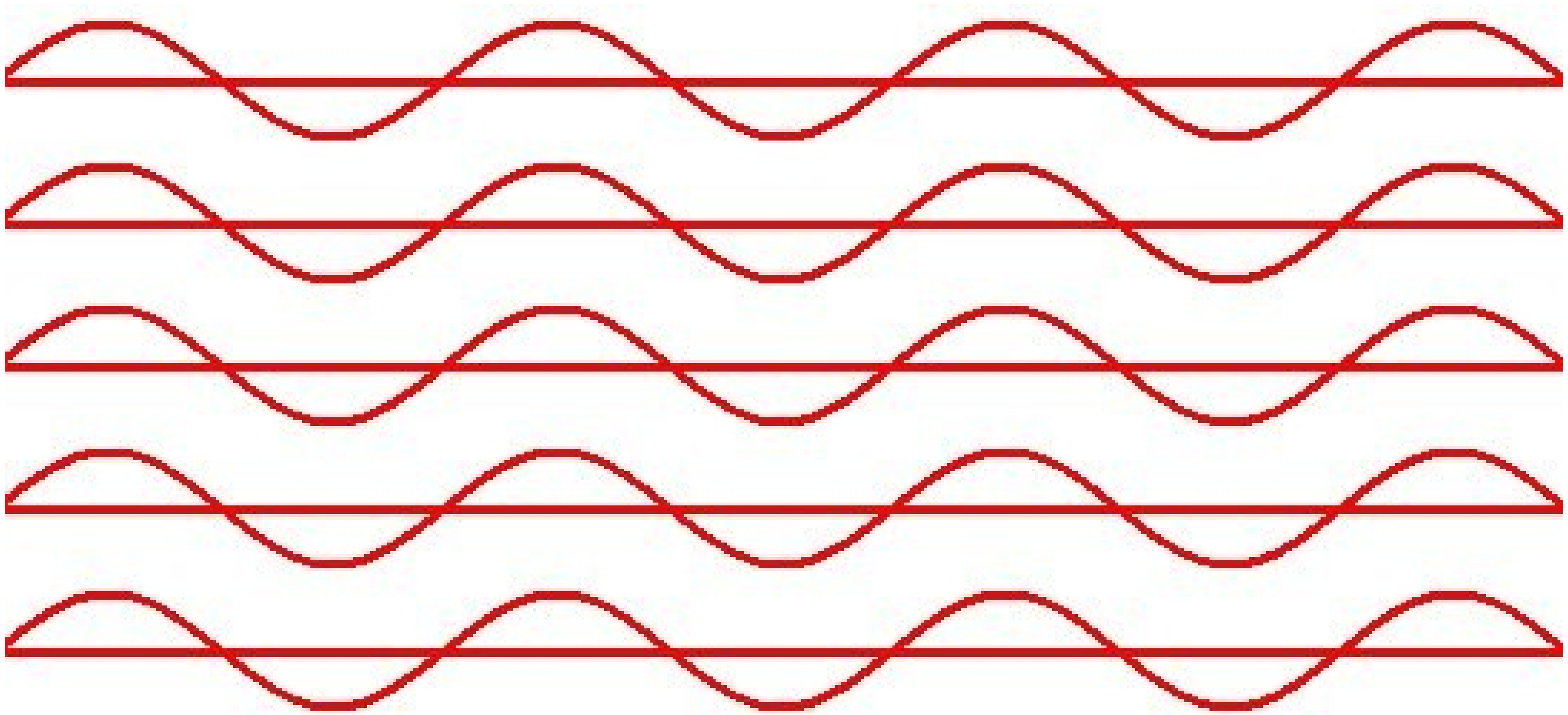


Waves

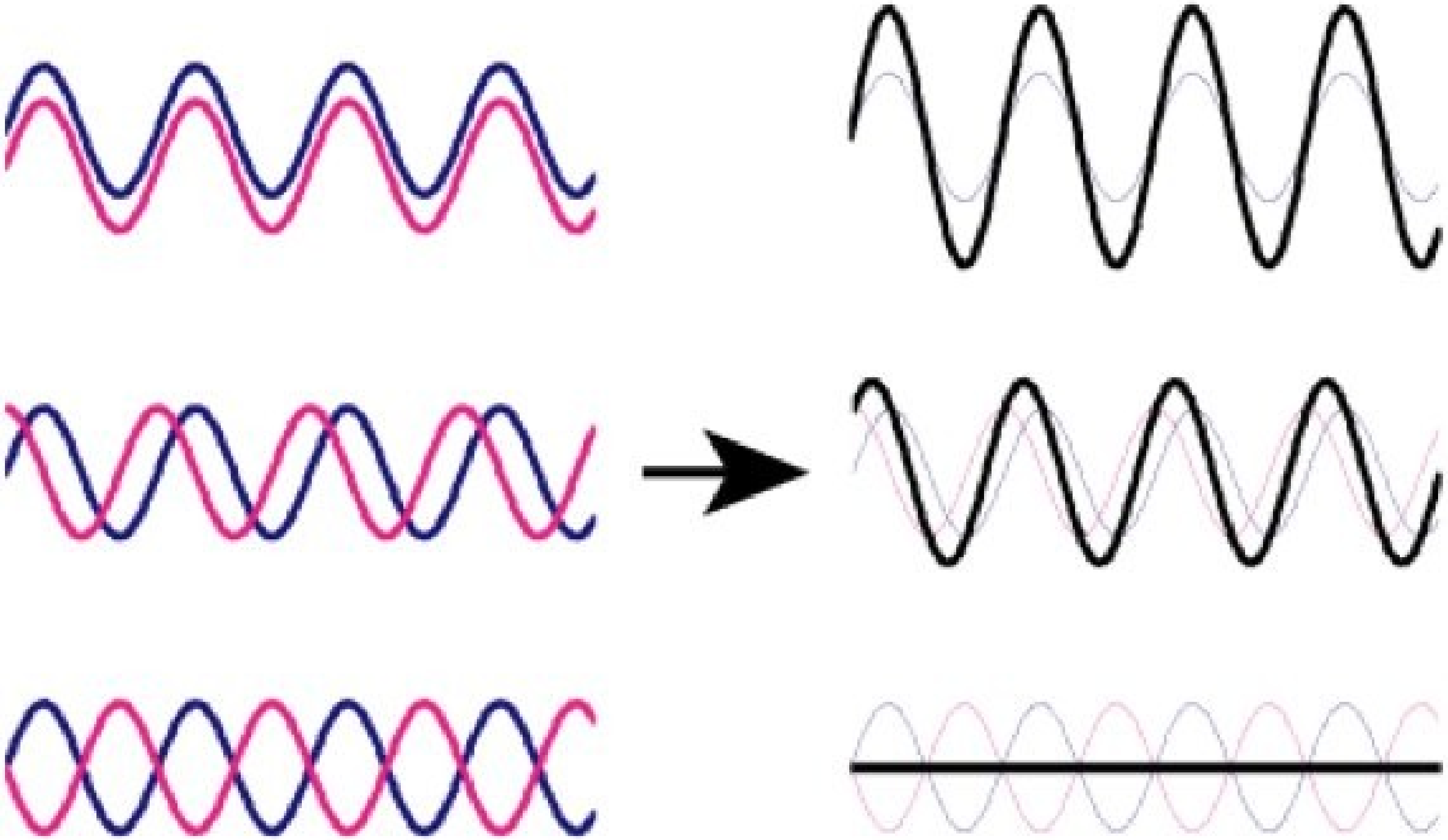

Wavelength
λ
Amplitude
x
Distance


Period
T
Amplitude
t
Time

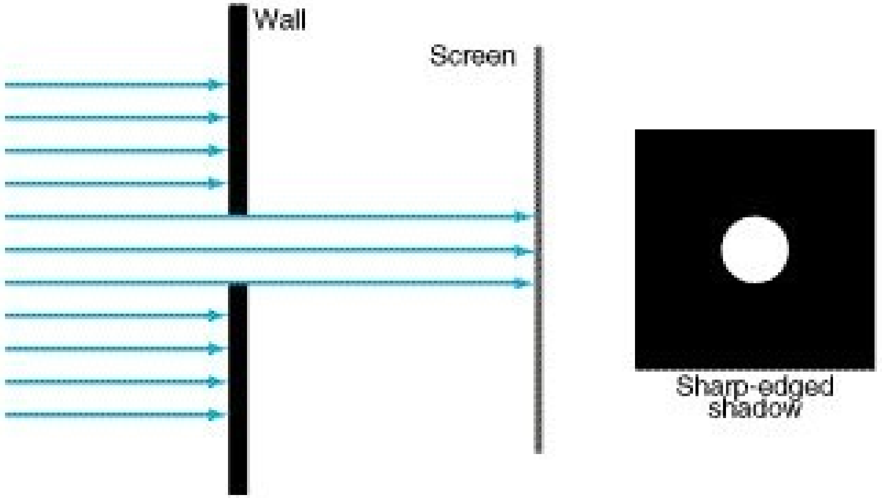
Coherent beam

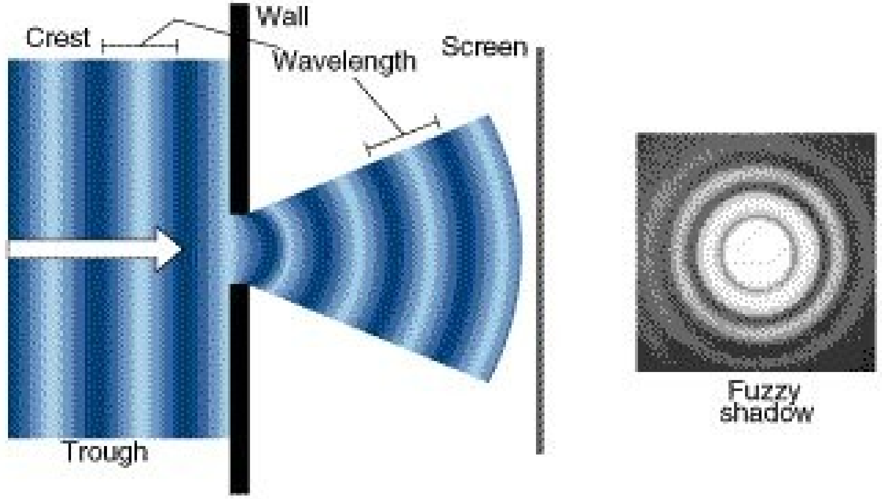


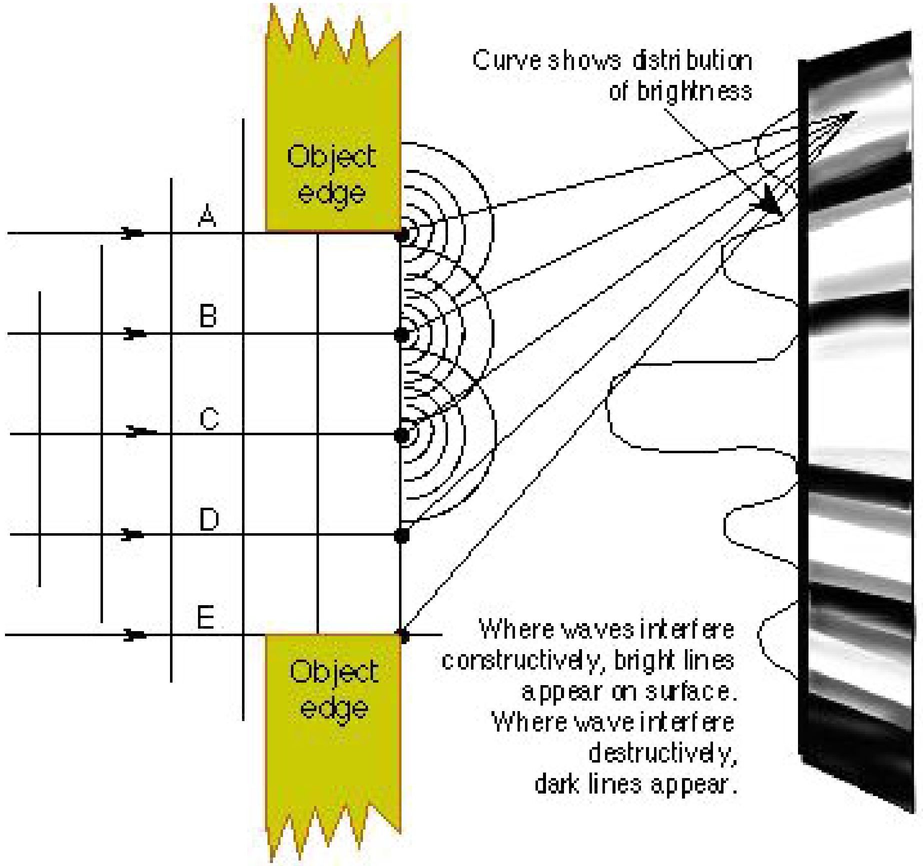
Addition of waves



Particles & waves



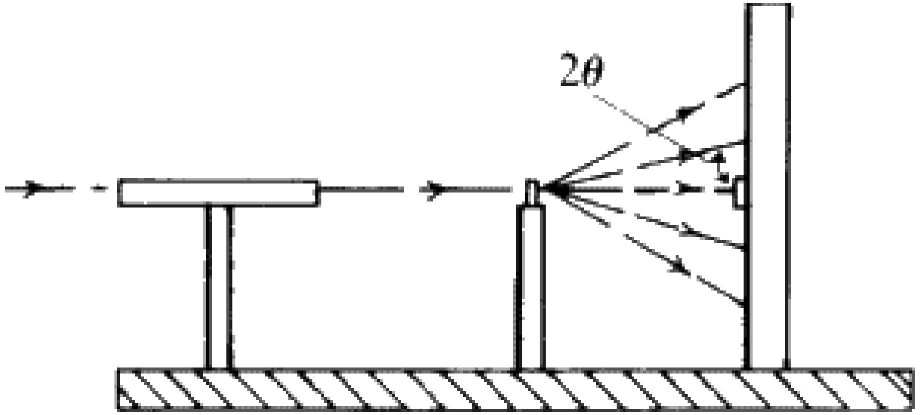




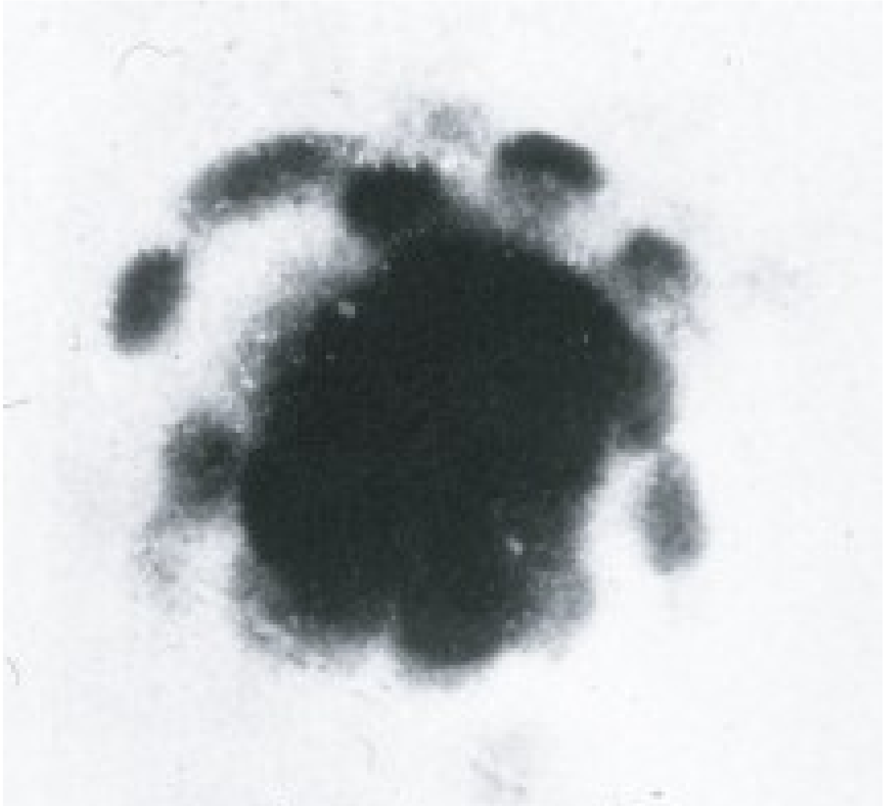
MAX VON LAUE
(1879-1960)

- **1914 Nobel Laureate in Physics**

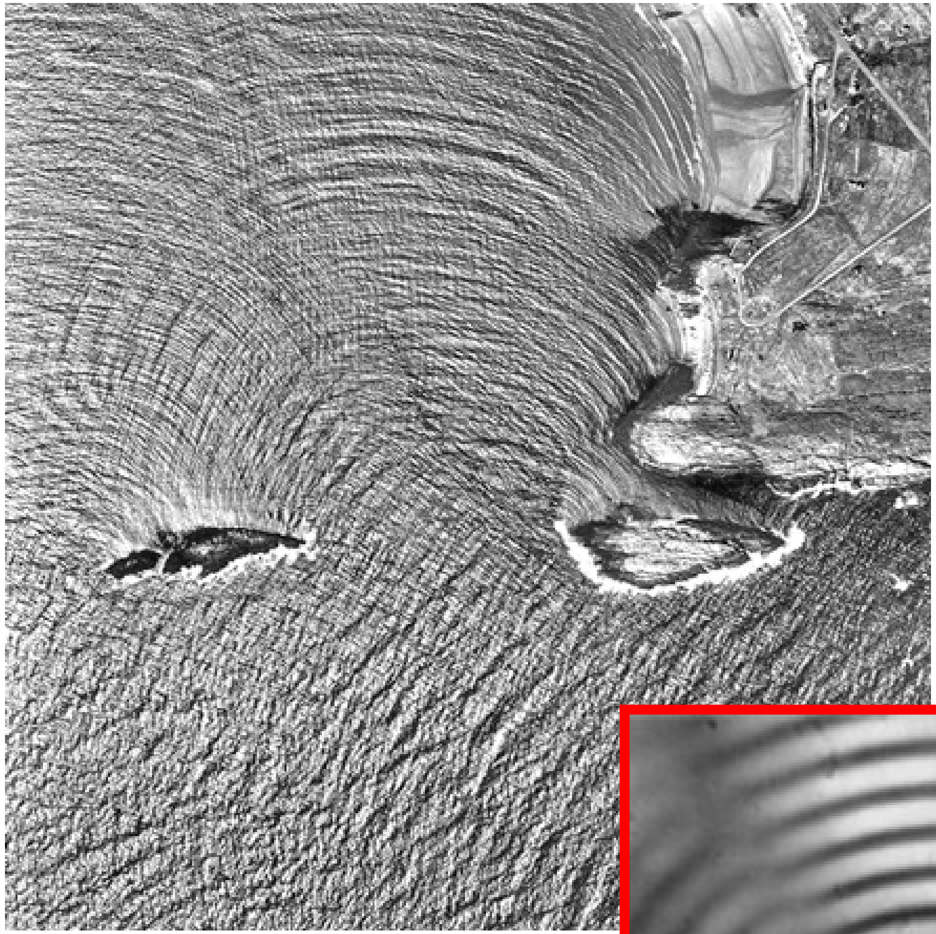
  for his discovery of the diffraction of X-rays by crystals

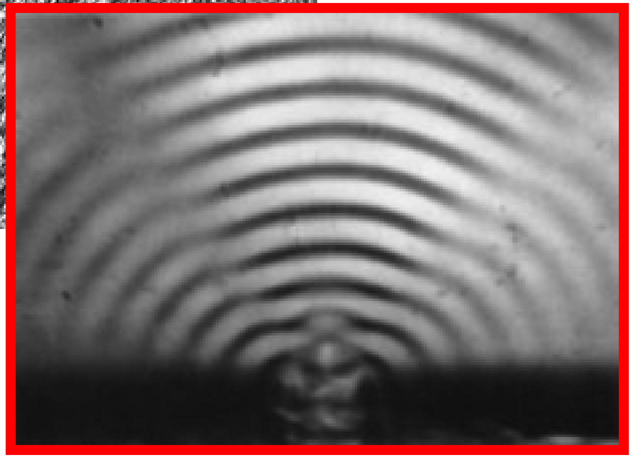


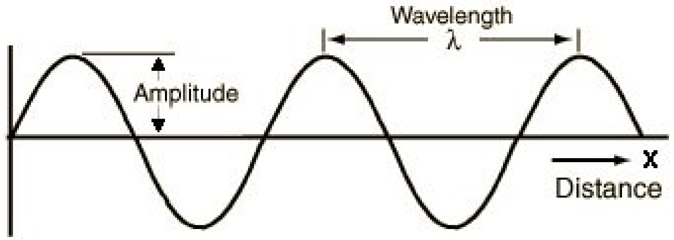
Friedrich and Knipping



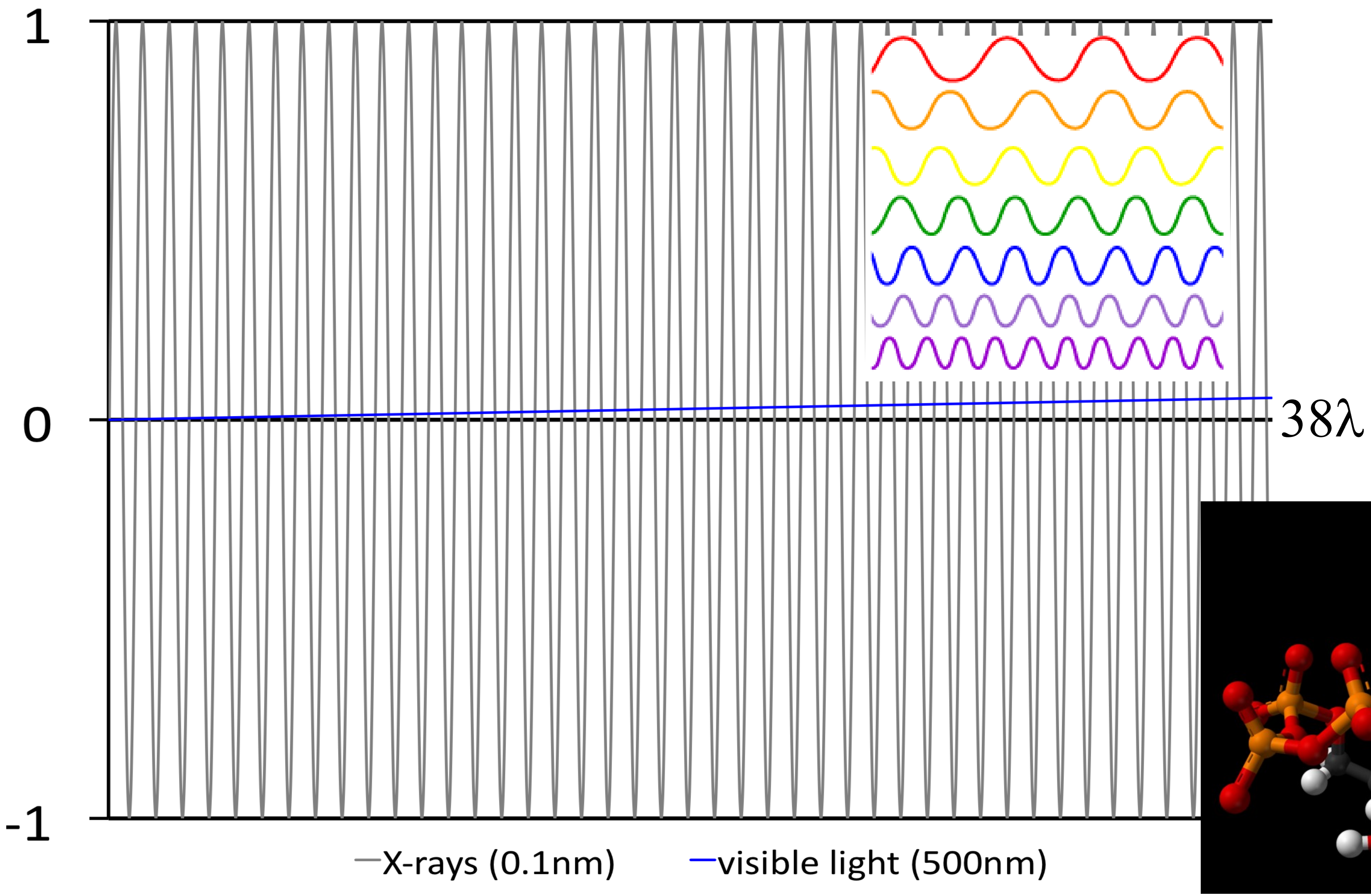
Wavelength and diffraction

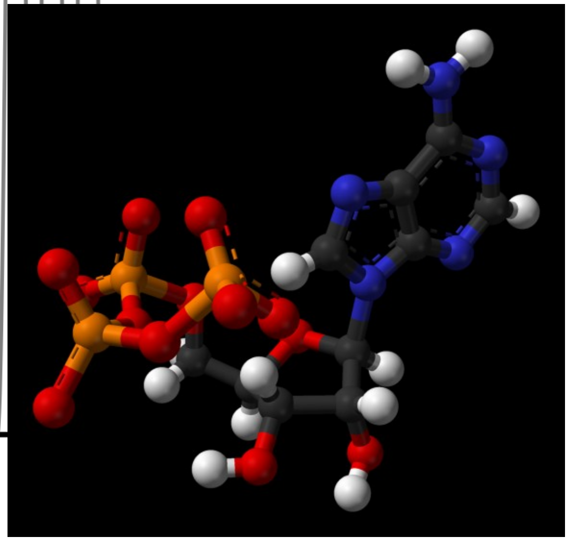


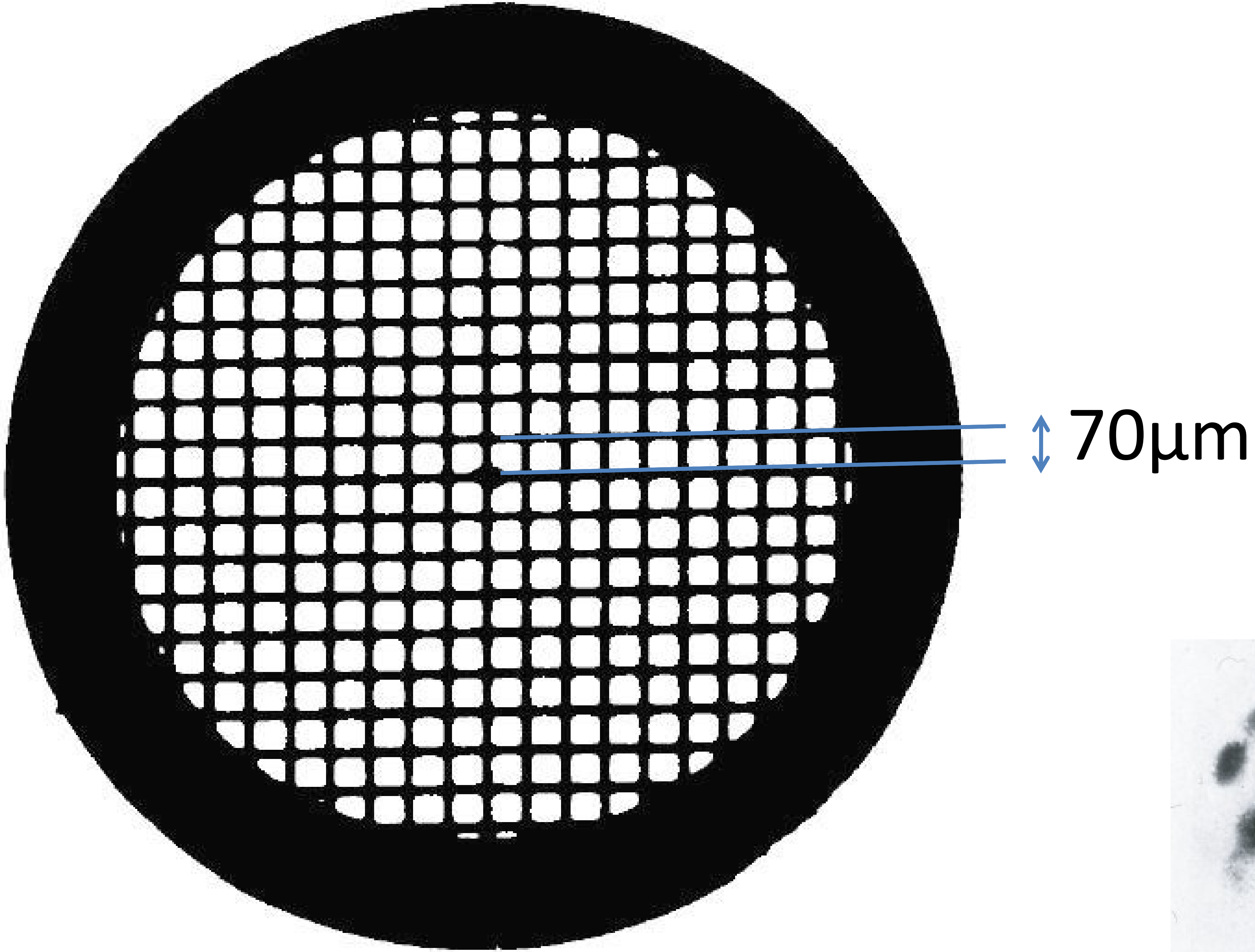


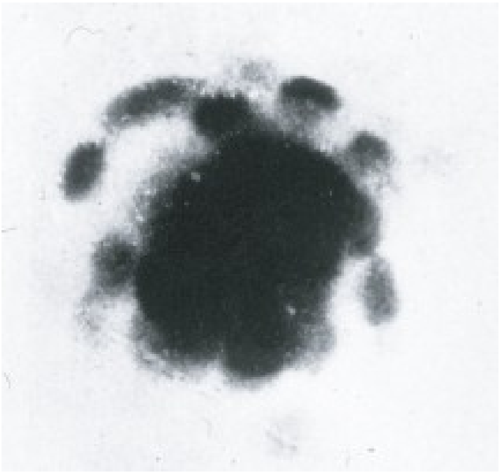


Wavelength comparison of X-rays and visible light






70µm

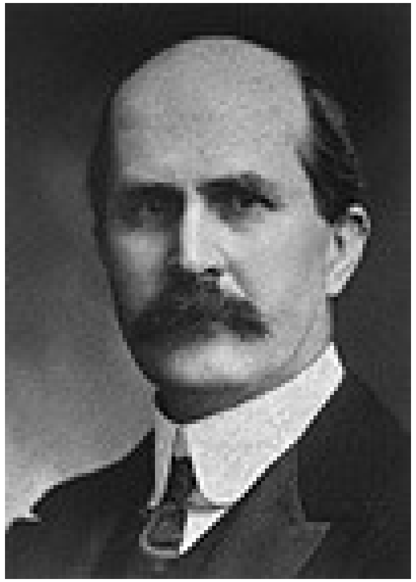


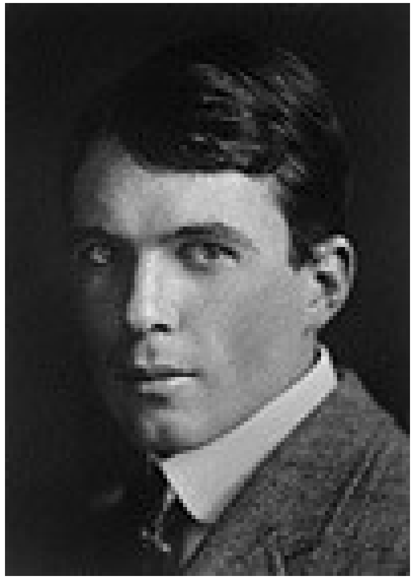
WILLIAM HENRY BRAGG (1862-1942)
WILLIAM LAWRENCE BRAGG (1890-1971)

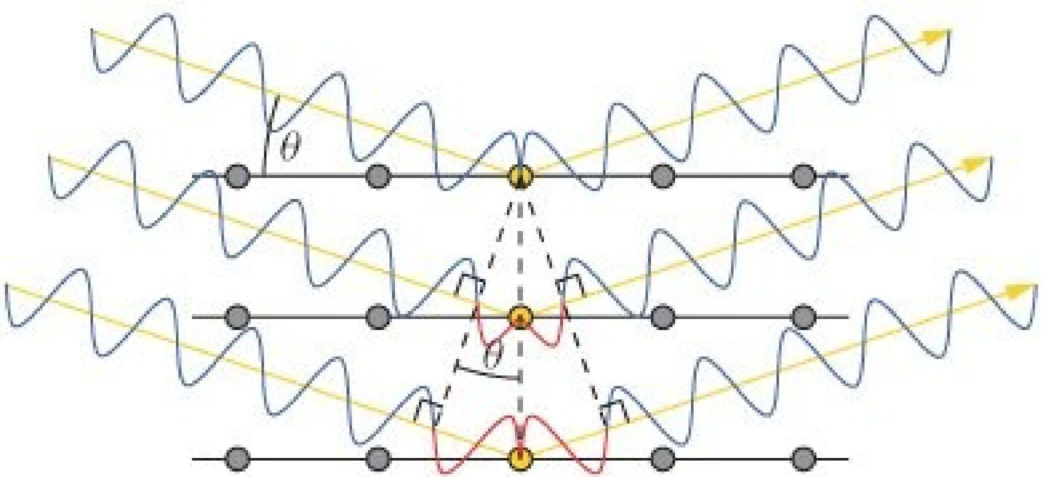
- **1915 Nobel Laureates in Physics**

  for the analysis of crystal structure by means of X-rays

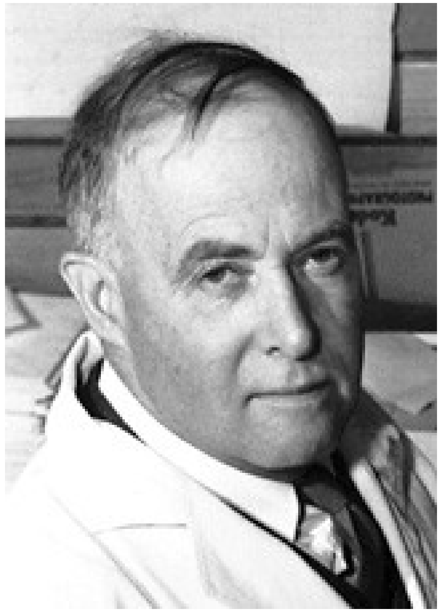
$$n\lambda = 2d \sin\theta$$



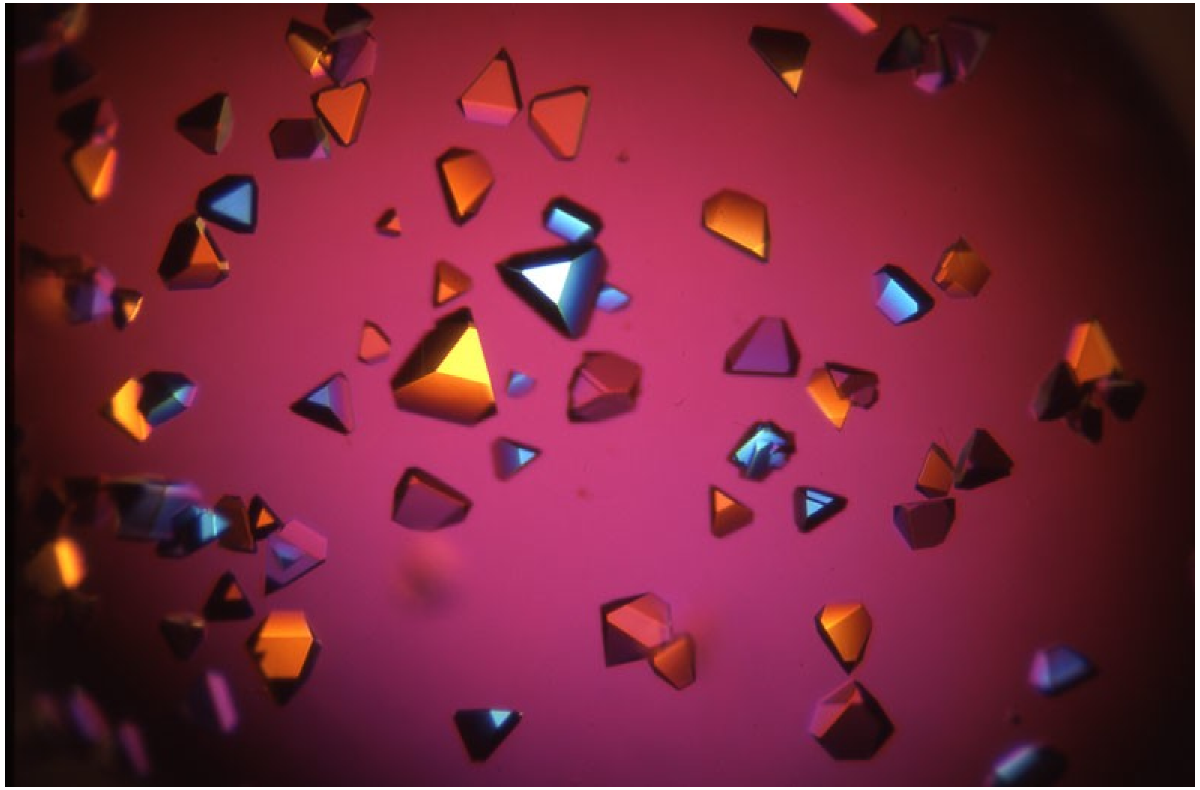




James Batcheller Sumner
(1879-1960)



- **1946 Nobel Laureate in Chemistry**
  for his discovery that enzymes can be crystallized

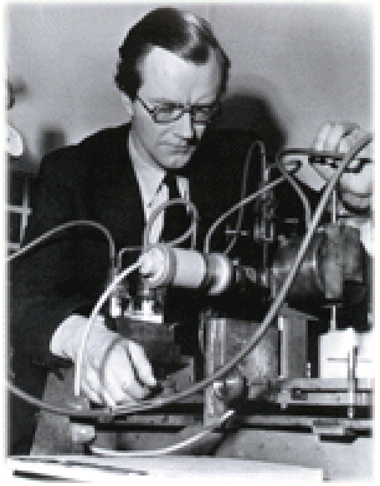


FRANCIS HARRY COMPTON CRICK (1916-2004)
JAMES DEWEY WATSON (1928)
MAURICE HUGH FREDERICK WILKINS (1916-2004)

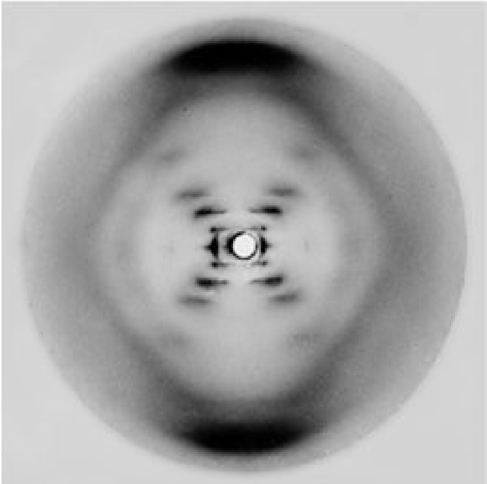
- **1962 Nobel Laureates in Physiology and Medicine**

for their discoveries concerning the molecular structure of nuclear acids and its significance for information transfer in living material.

James Watson
and Francis Crick

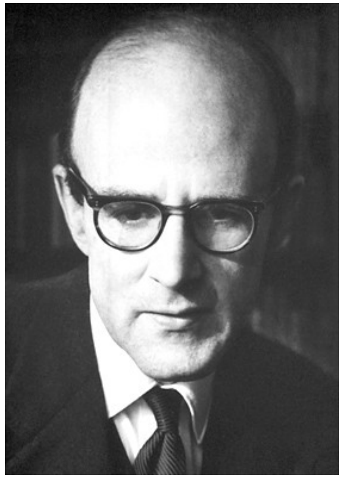


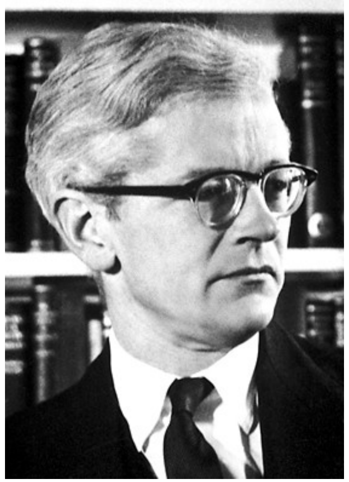
Maurice Wilkins



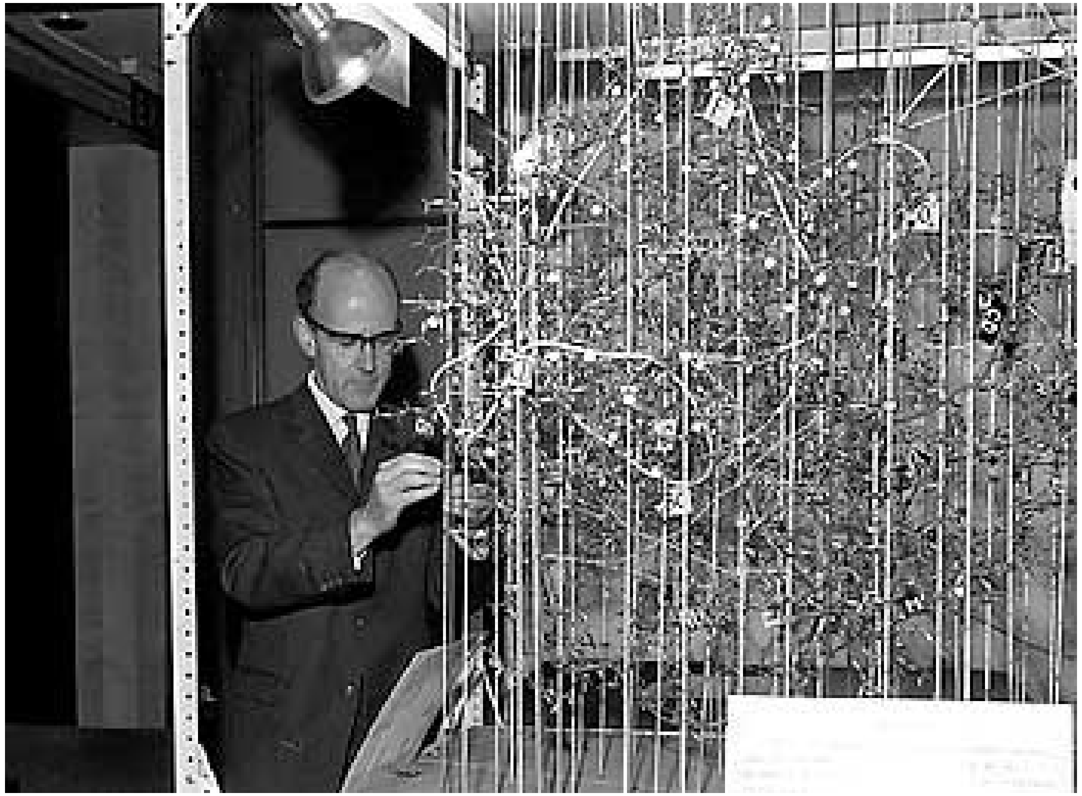
Rosalind Franklin

Max Ferdinand Perutz (1914 – 2002)
John Cowdery Kendrew (1917 – 1997)

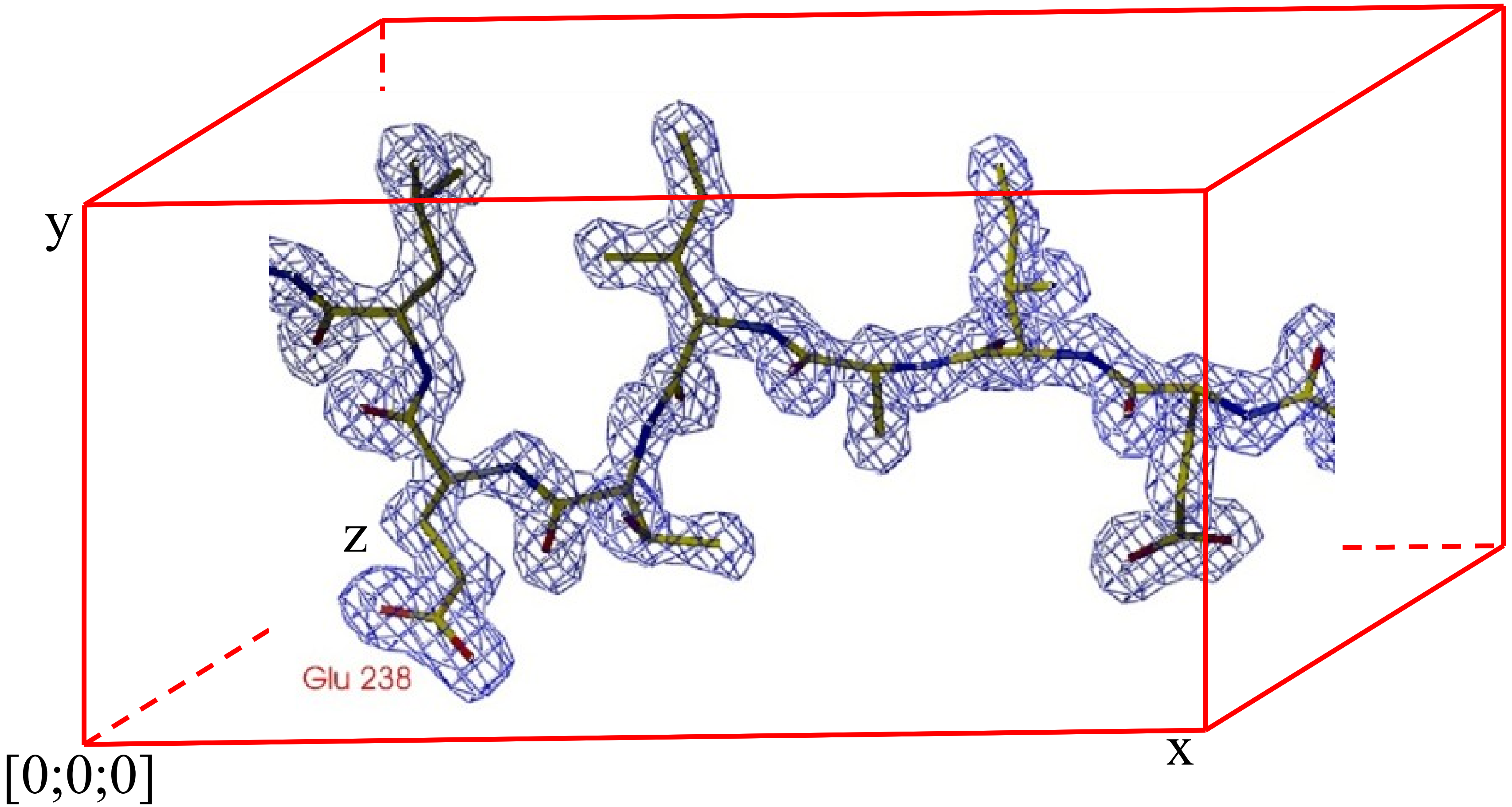




- **1962 Nobel Laureates in Physics**
  for their studies of the structures of globular proteins

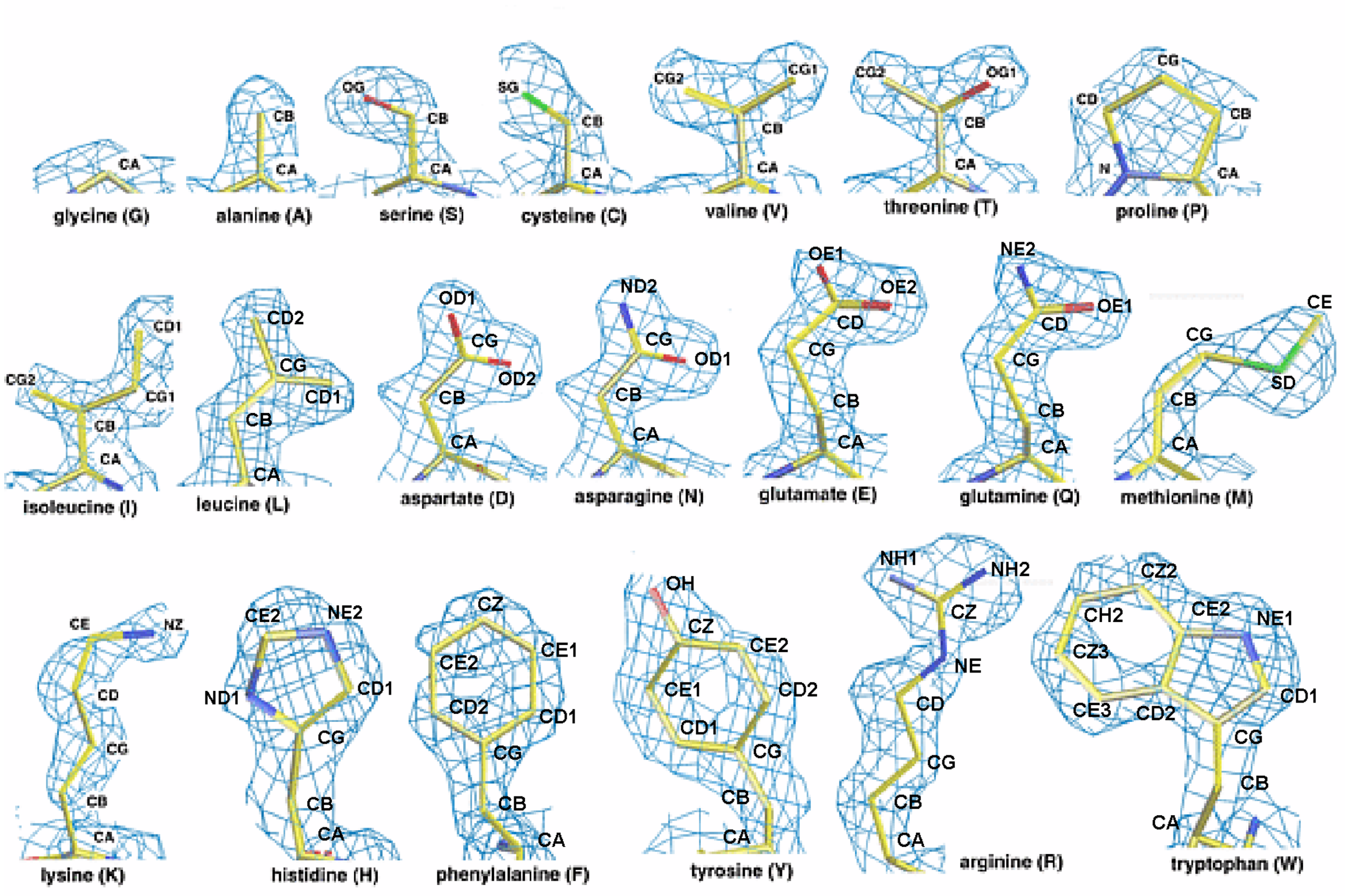


Information from X-ray diffraction experiment

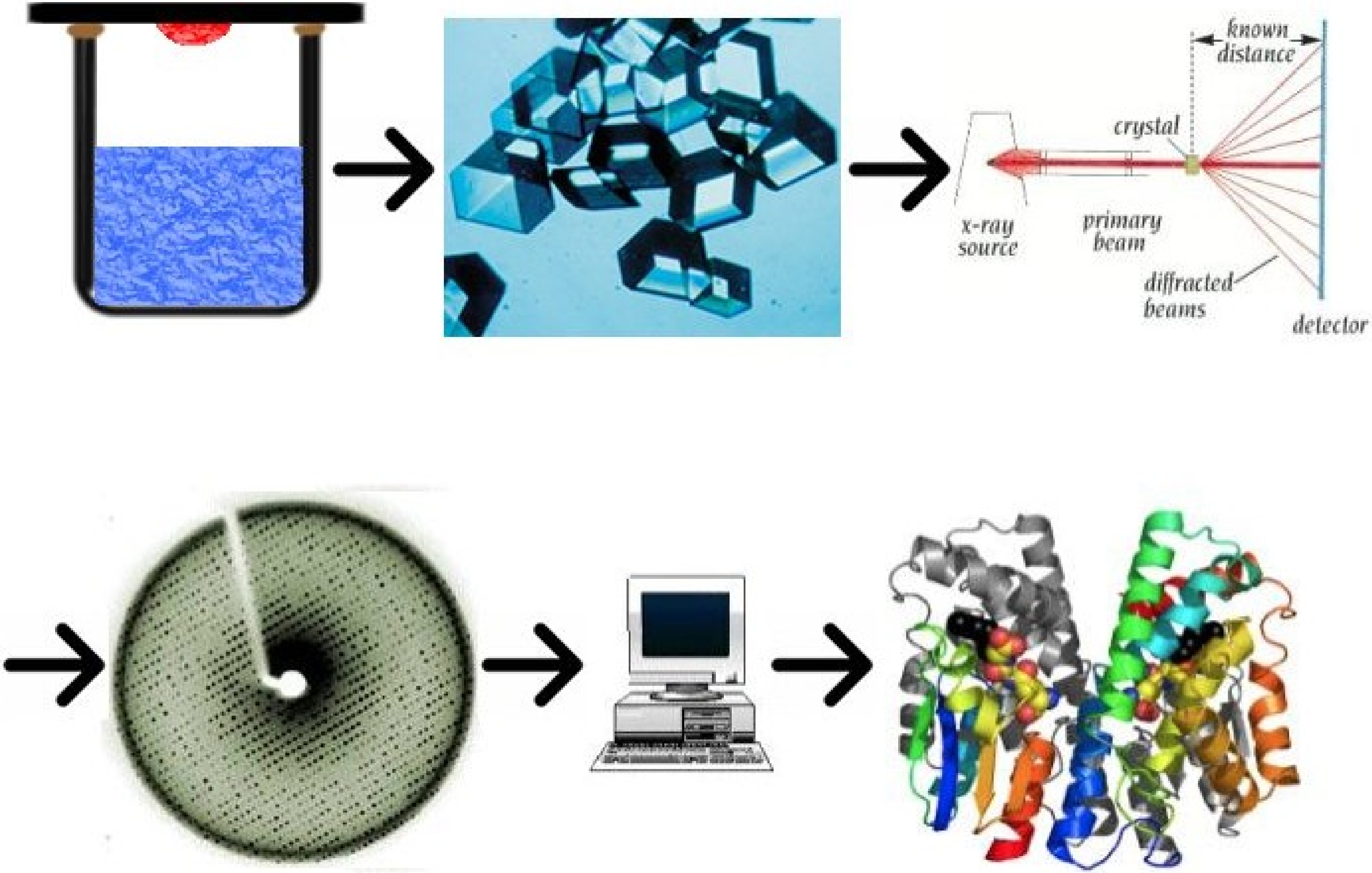


$$\rho(x\,y\,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}\left|F(h\,k\,l)\right|\exp\left[-2\pi i(hx+ky+lz)+i\alpha(h\,k\,l)\right]$$

Representative electron density for amino acid side chains

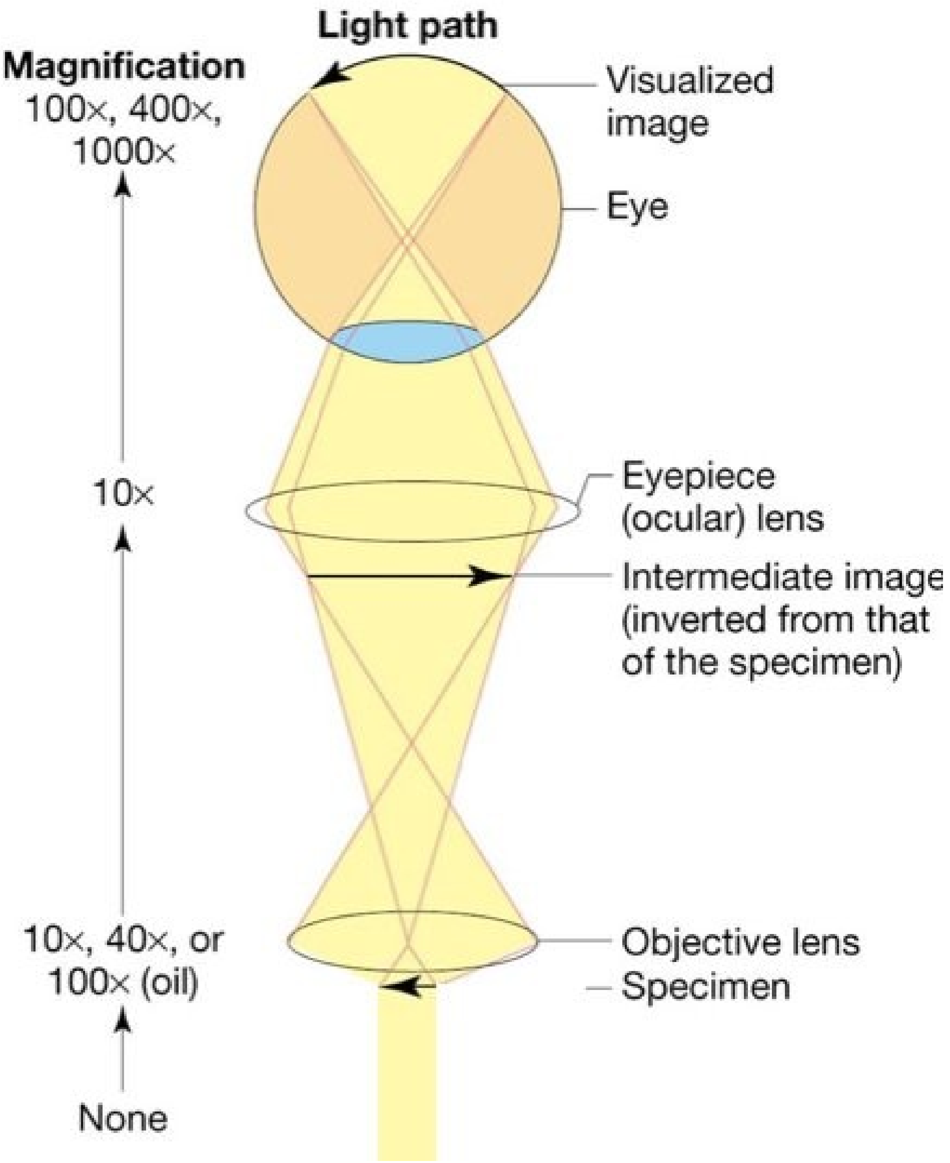


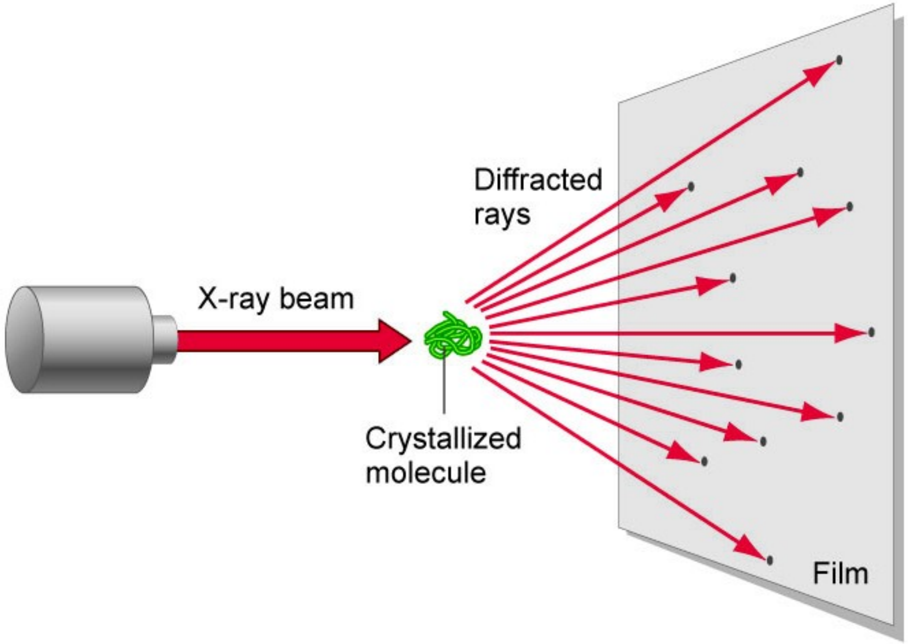
Electron density maps calculated at 1.5 Angstrom resolution.



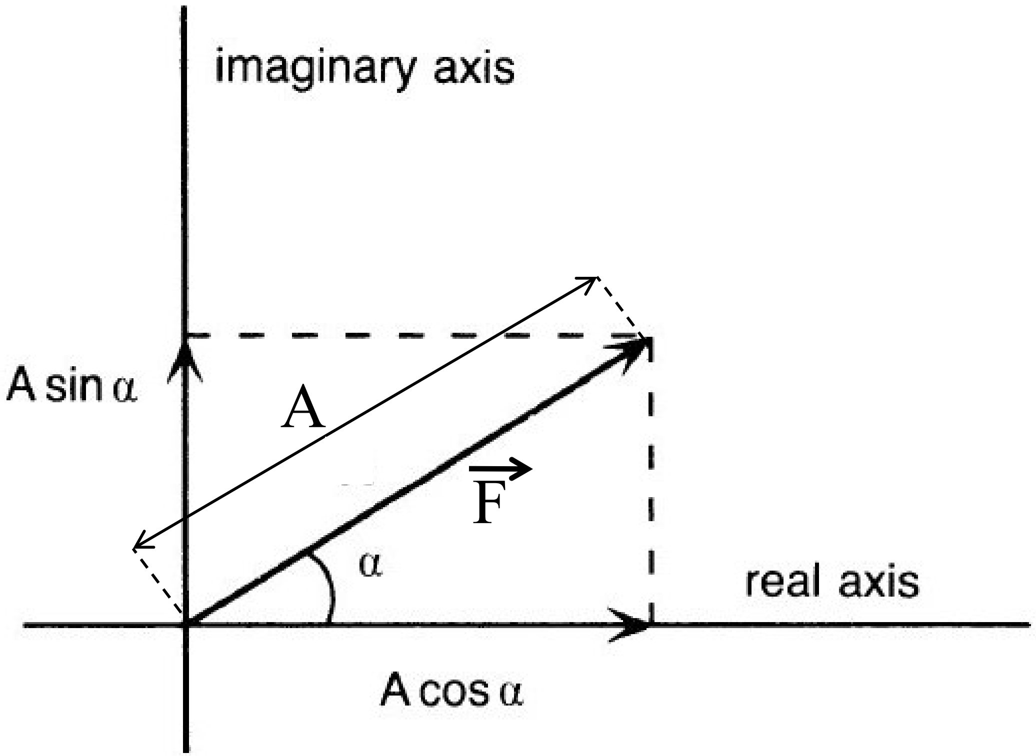
known
distance
crystal
x-ray
source
primary
beam
diffracted
beams
detector

Comparison of microscope and diffraction





Wave as a vector



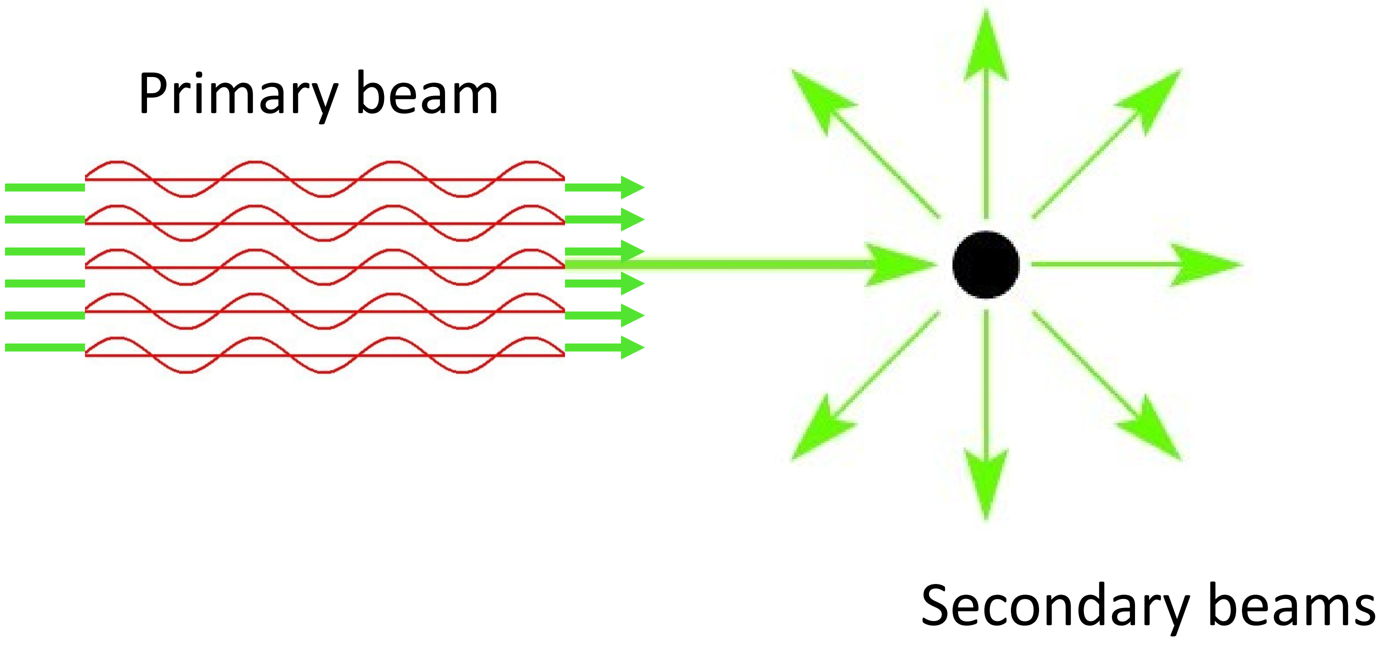
A- wave amplitude

α- wave phase

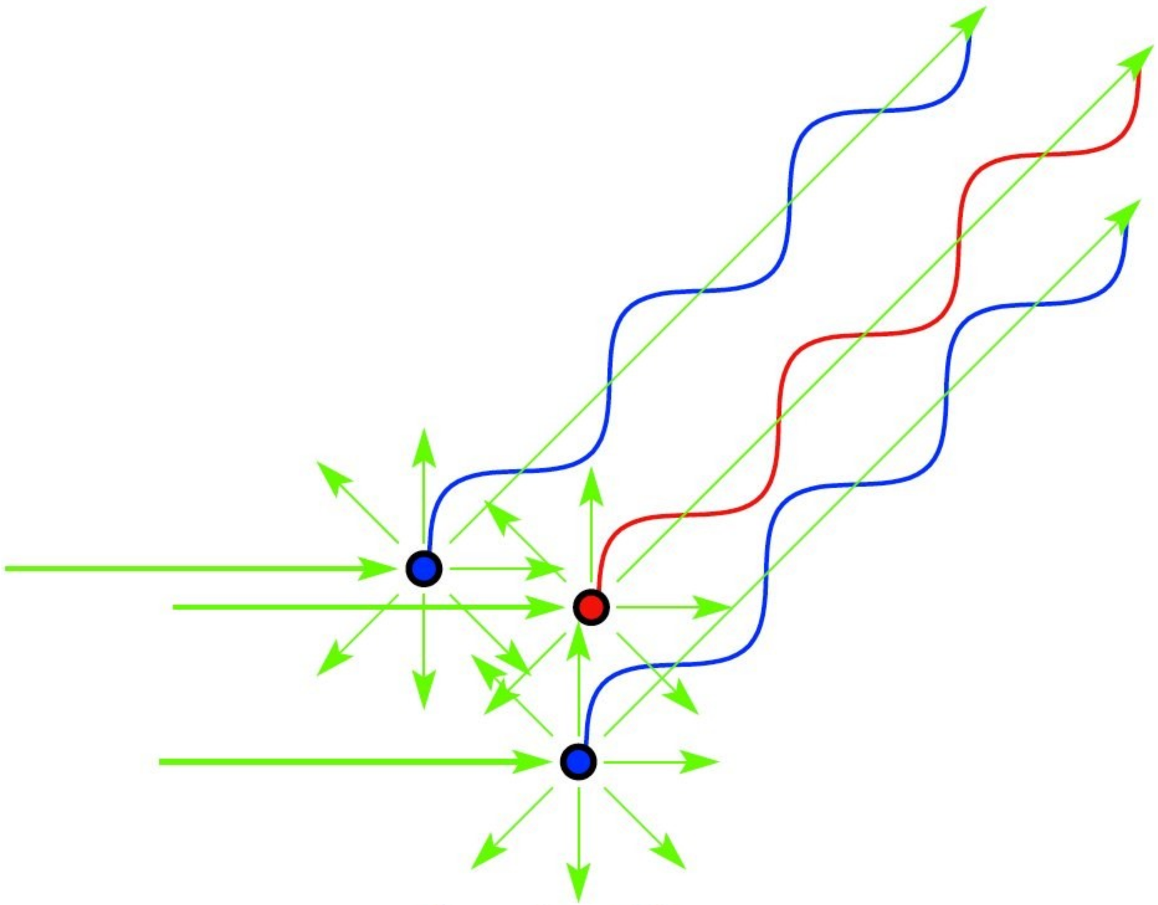
$\vec{F}=A\cos\alpha+iA\sin\alpha$

$\vec{F}=A\exp(i\alpha)$

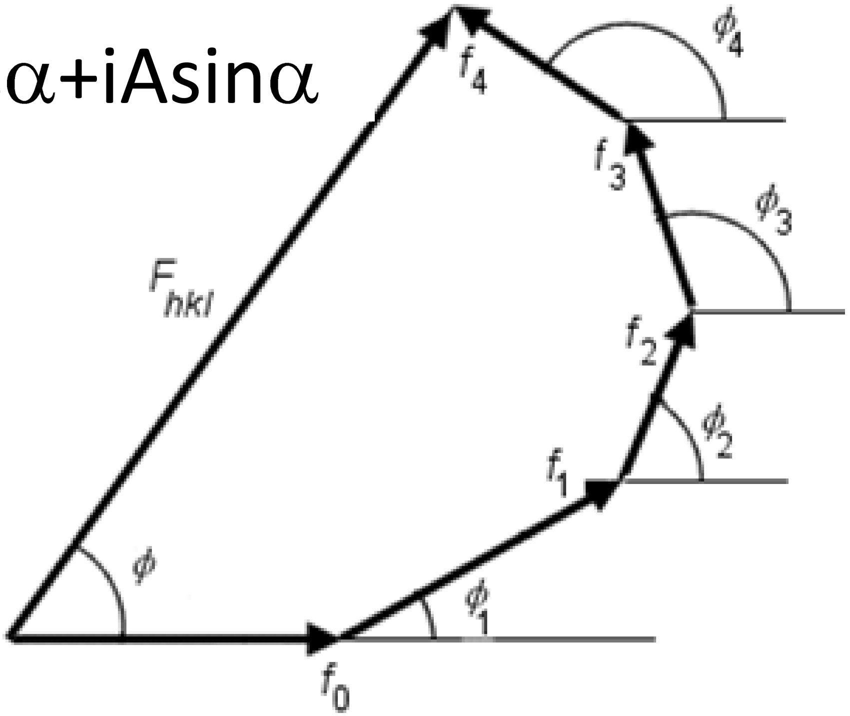
X-rays scatter from electrons in all directions



Adition of waves



F=Acosα+iAsinα



$$F(h\,k\,l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\,y\,z) \exp[2\pi i(hx + ky + lz)]\,dx\,dy\,dz$$

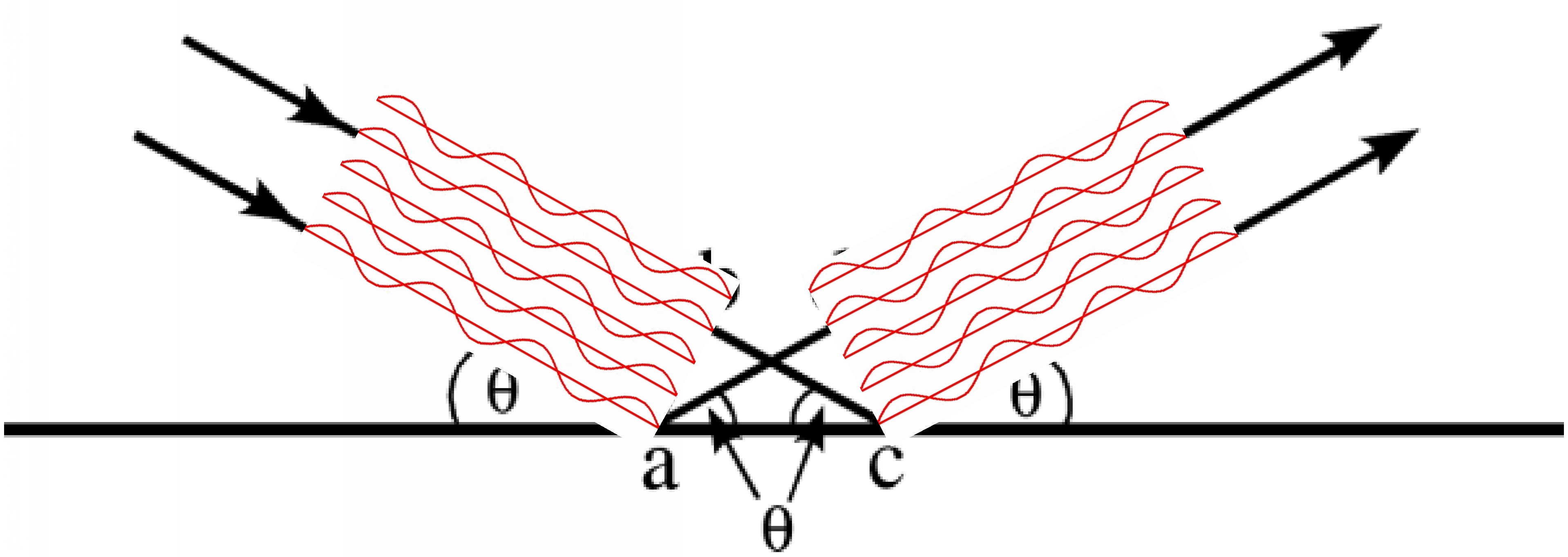
$$F(h\,k\,l) = |F(h\,k\,l)|e^{i\alpha(h\,k\,l)}$$

- Scattering from a single molecule is not detectable

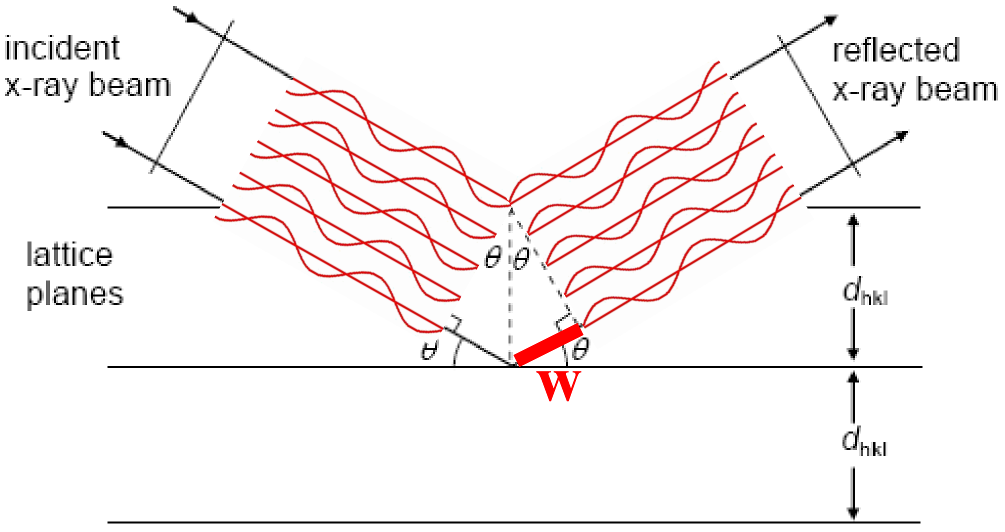
- If molecules are all oriented in the same way, the scattering from individual molecules will add in certain directions

    – Which directions?

There is no path and PHASE DIFFERENCE when rays reflect from a plane

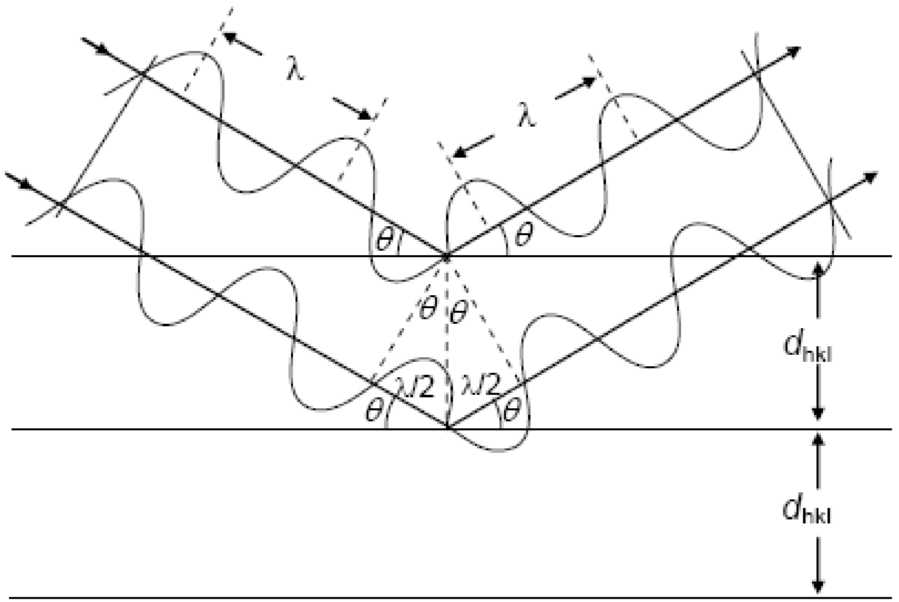


There is NO PHASE DIFFERENCE if the path differences are equal to whole number multiplies of wavelength.



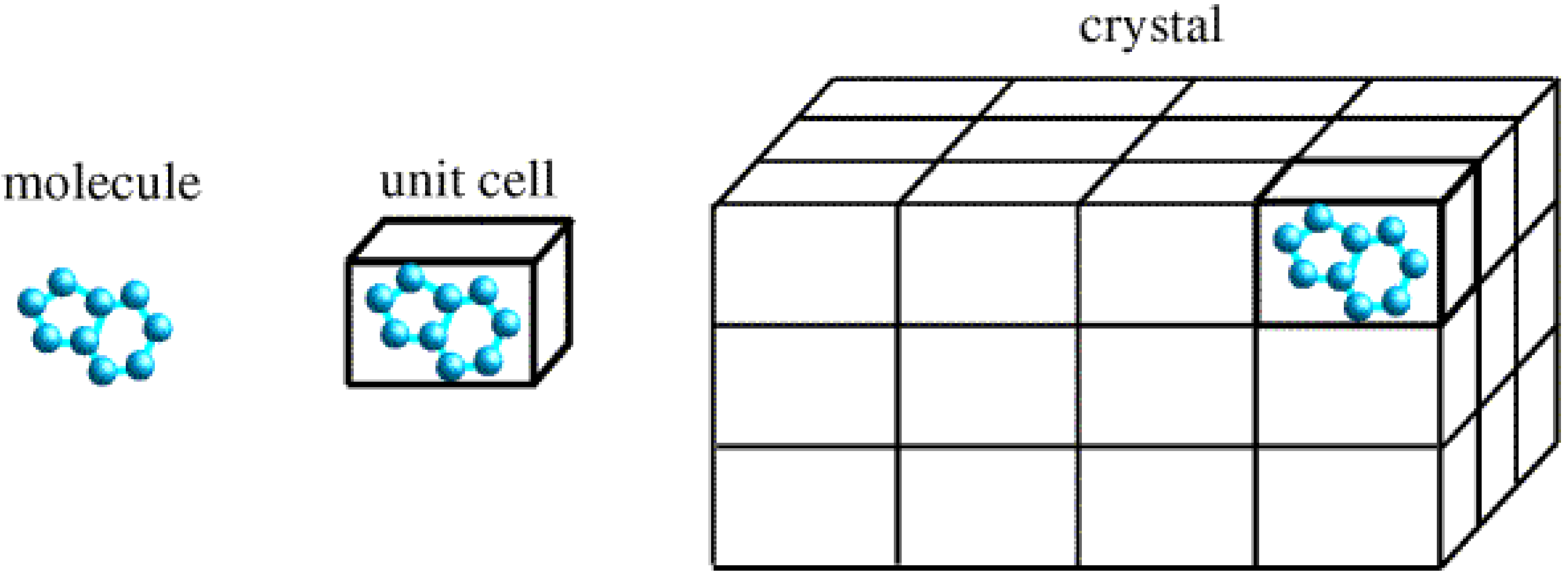
Bragg's law:

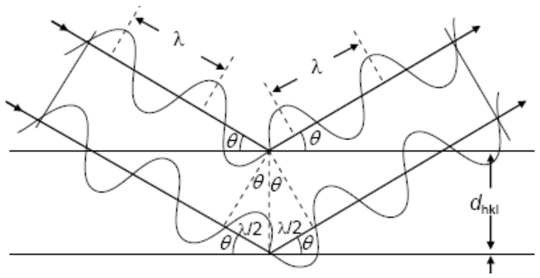
$$n\lambda = 2d \sin\theta$$

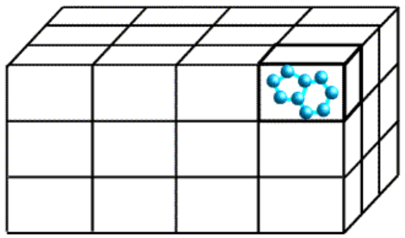


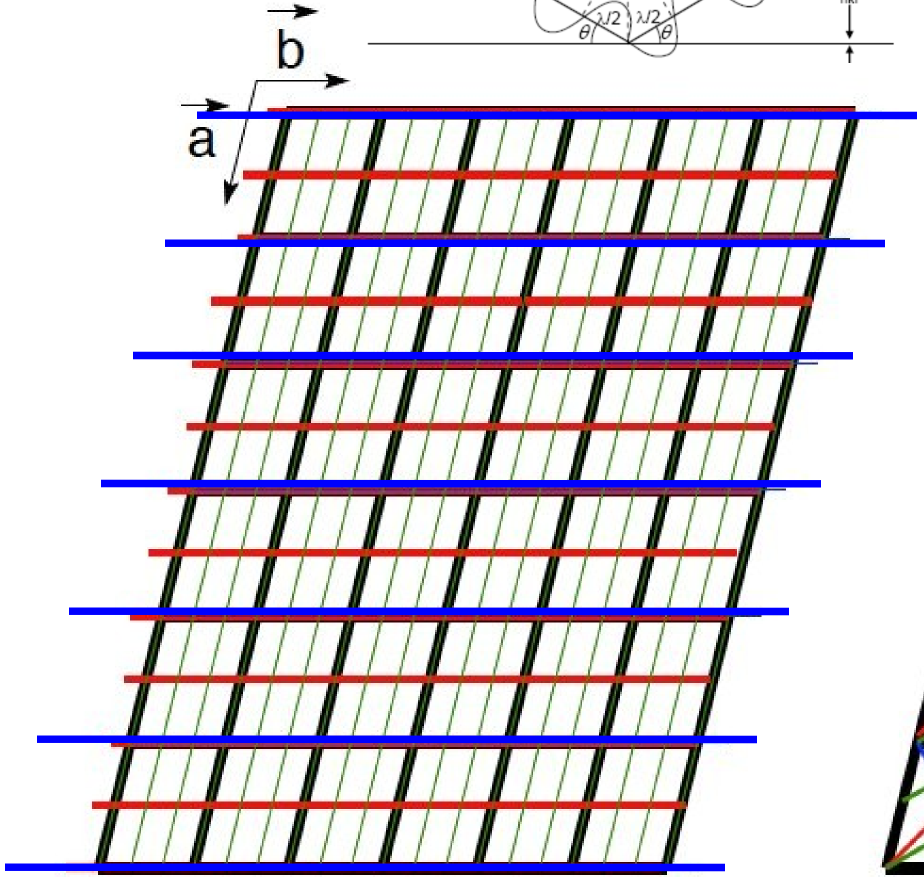
$\sin\theta =$ **w**$/d$

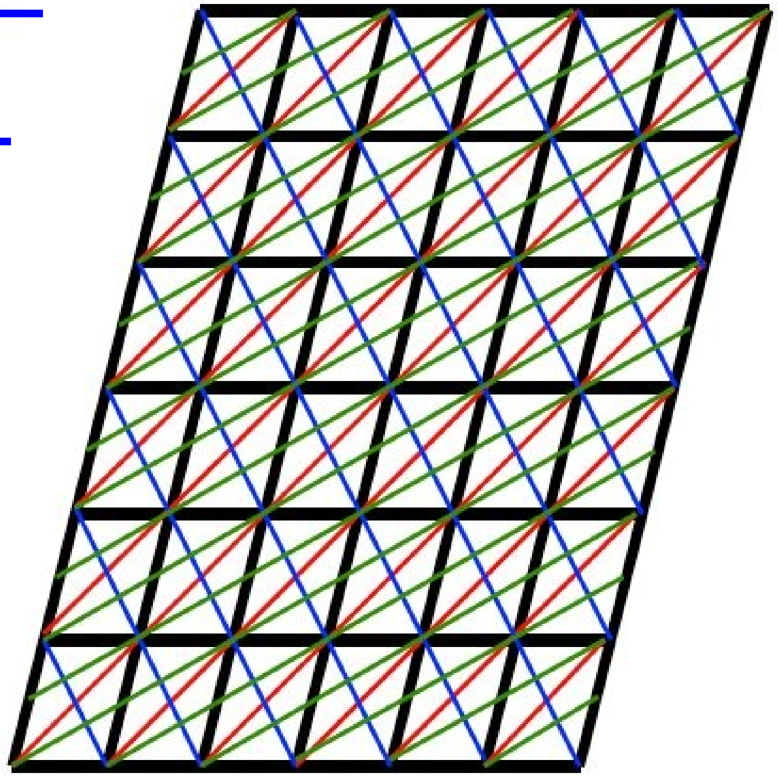
2**w** $= n\lambda$

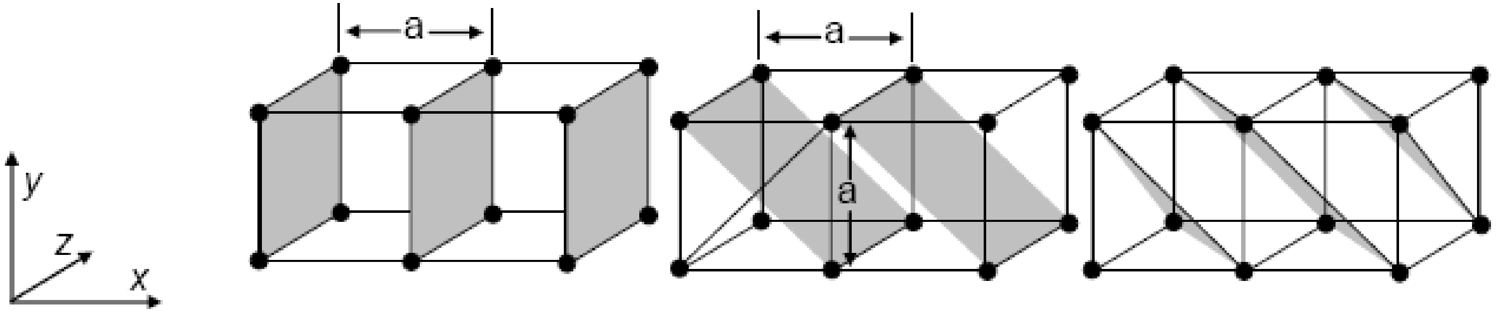

molecule
unit cell
crystal





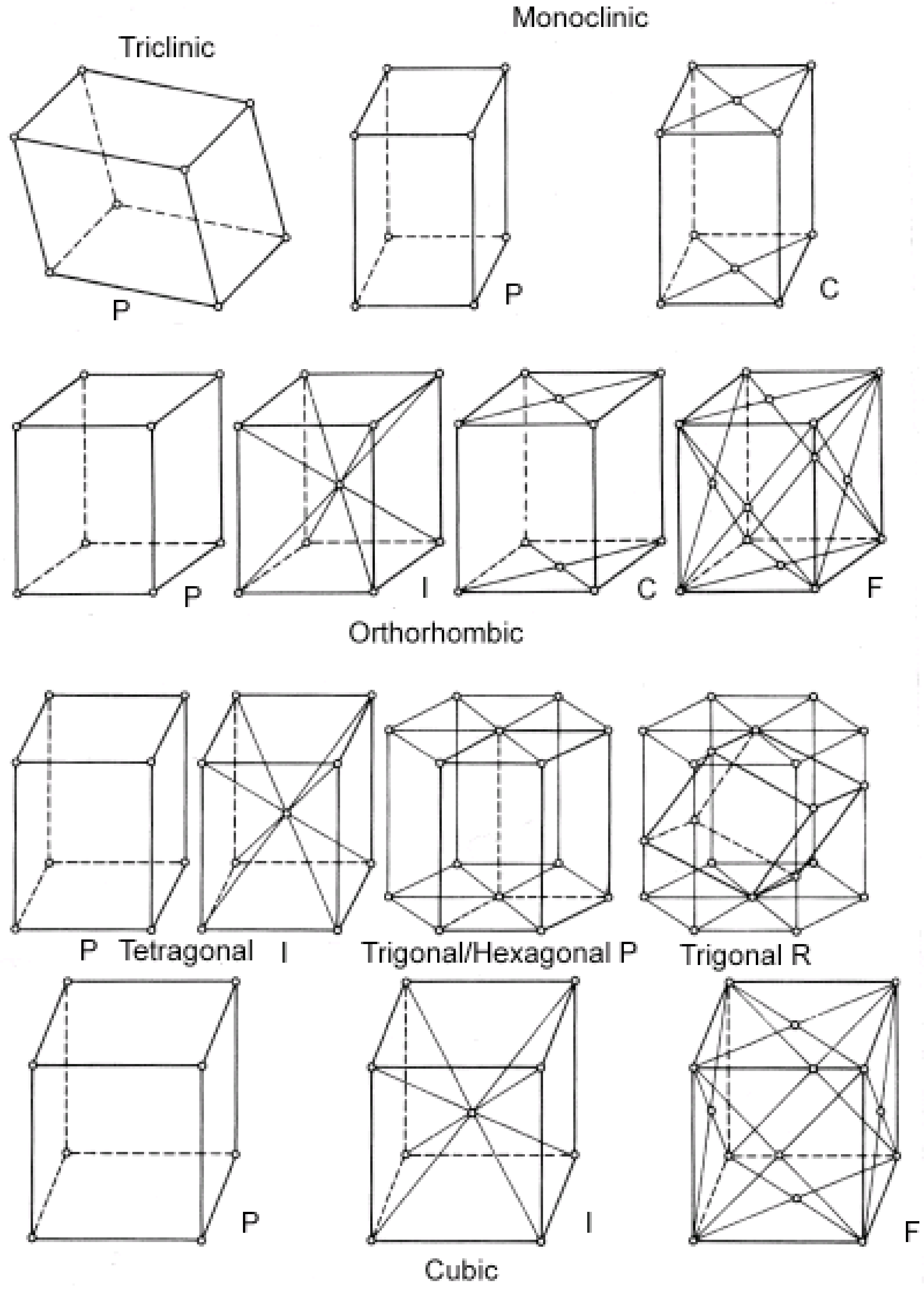




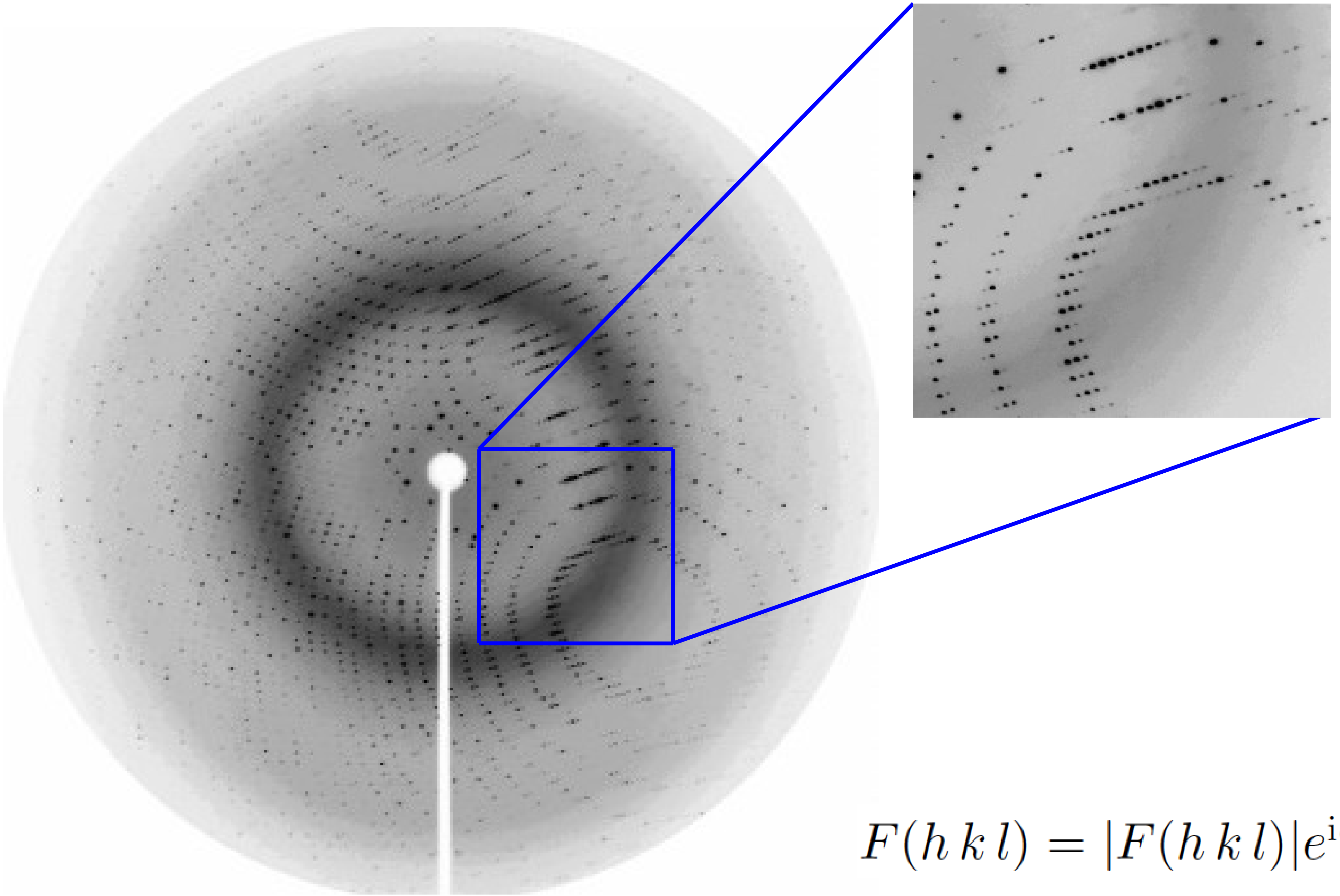


(h, k, l)

14 Bravais Lattices

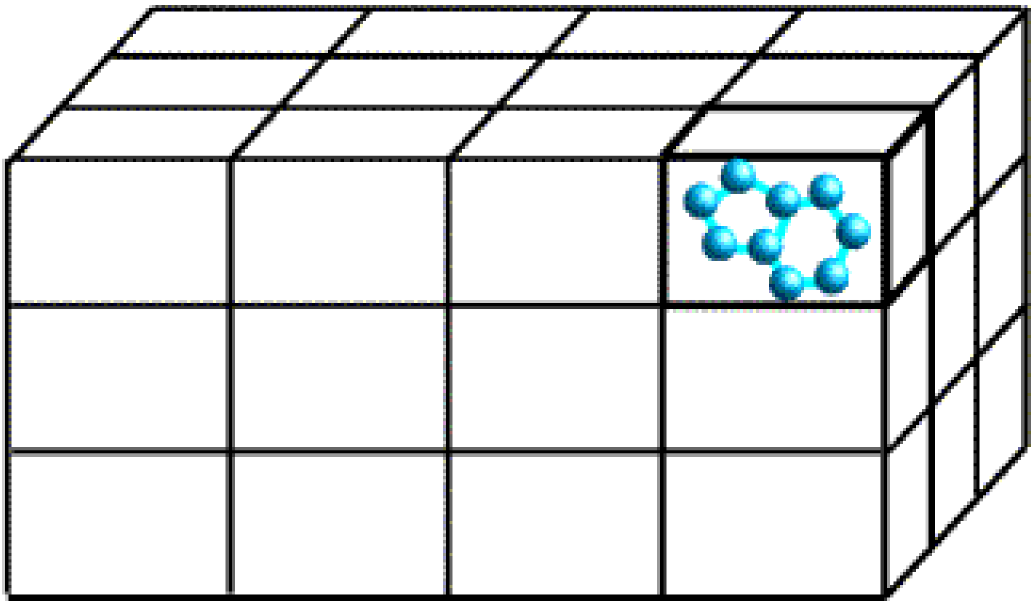


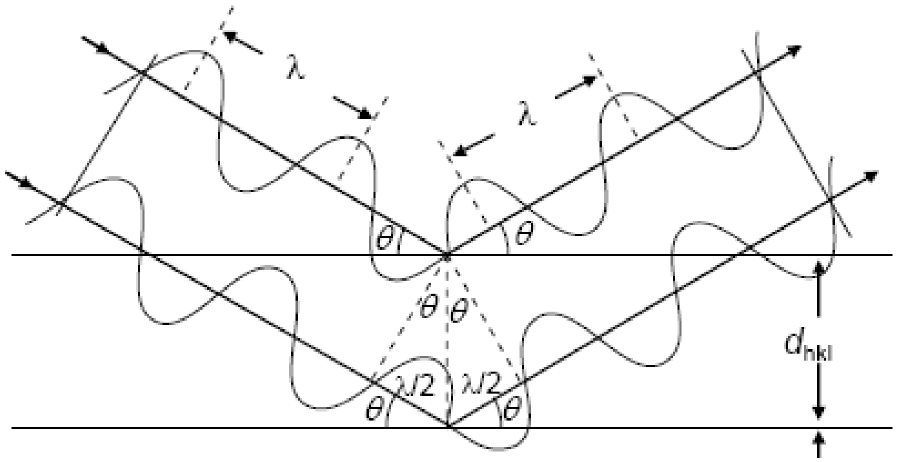
Diffraction pattern from a protein crystal



$$F(h\,k\,l) = |F(h\,k\,l)|e^{i\alpha(h\,k\,l)}$$

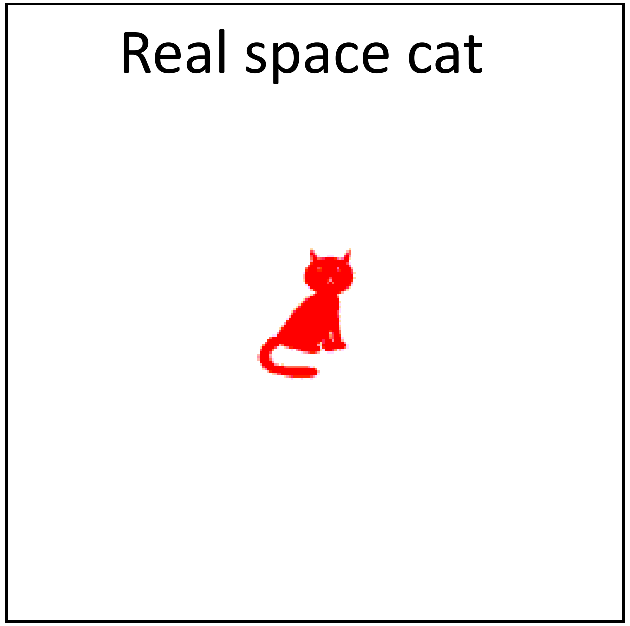
crystal



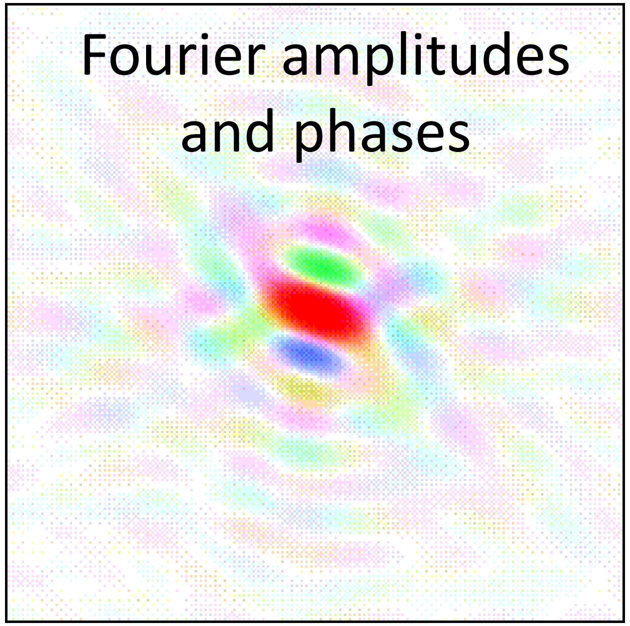


$n\lambda = 2d \sin\theta$

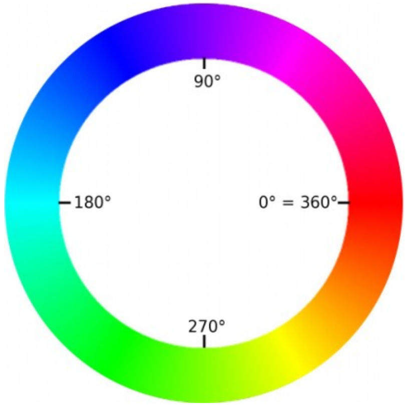
$$F(h\,k\,l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\,y\,z) \exp[2\pi i(hx + ky + lz)]\, dx\, dy\, dz$$



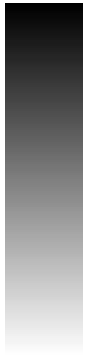
Fourier transform

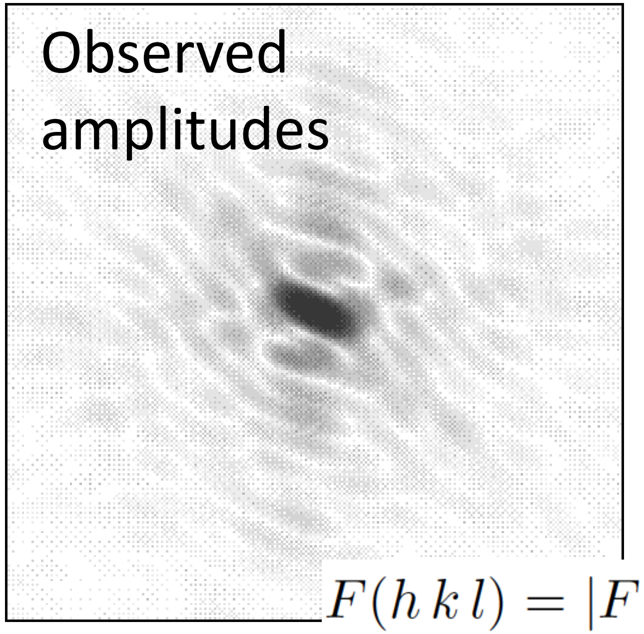


Circular rainbow scale of phases



Linear intensity scale of amplitude size

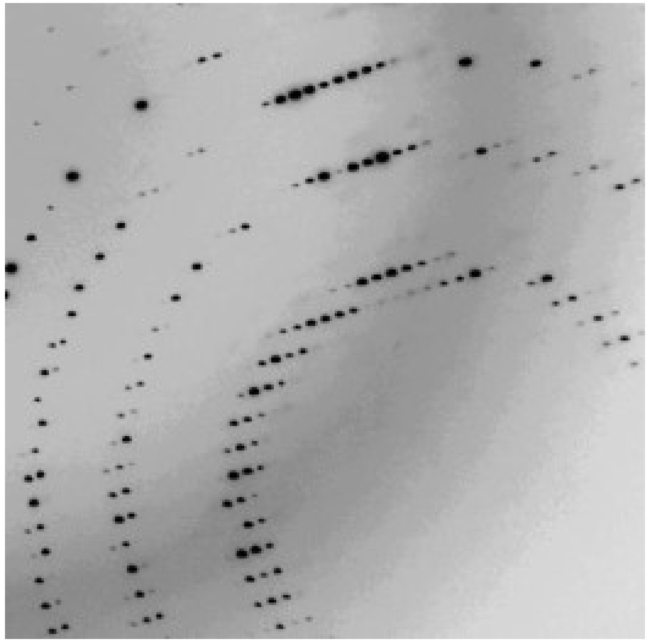


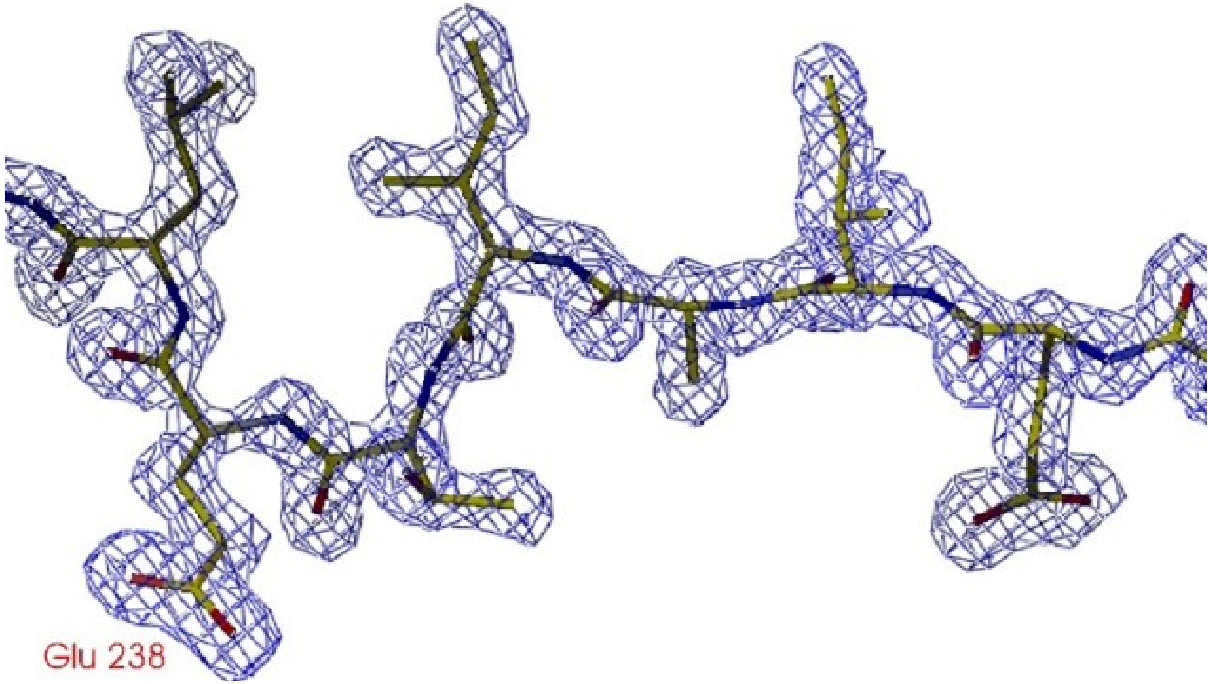


$$F(h\,k\,l) = V \int_{x=0}^{1}\int_{y=0}^{1}\int_{z=0}^{1} \rho(x\,y\,z)\exp[2\pi i(hx+ky+lz)]\,dx\,dy\,dz$$

$$F(h\,k\,l) = |F(h\,k\,l)|e^{i\alpha(h\,k\,l)}$$

Electron density equation + PHASE PROBLEM





$$F(h\,k\,l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x\,y\,z) \exp[2\pi i(hx + ky + lz)]\, dx\, dy\, dz$$

$$\rho(x\,y\,z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} \left| F(h\,k\,l) \right| \exp[-2\pi i(hx + ky + lz) + i\alpha(h\,k\,l)]$$

Solving the phase problem by:

Molecular replacement

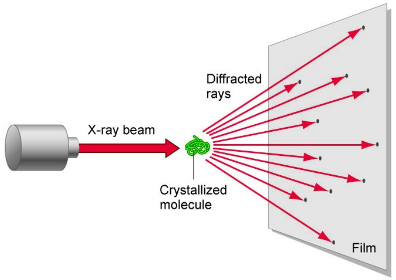
1. source of initial phases is a model
2. the model is oriented and positioned to obtain the best agreement with the x-ray data
3. phases are calculated from the model
4. The calculated phases are combined with the experimental data

Molecular **R**eplacement was invented by **M**ichael **R**ossmann

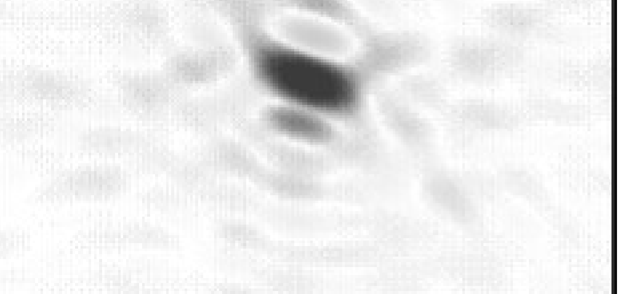
Unknown structure, unknown orientation

Cat

Diffraction experiment

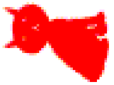


Observed amplitudes Phases unknown!



Fourier cat

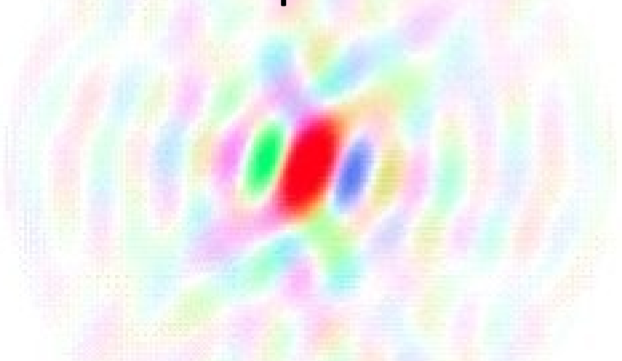
Known structure



Manx cat

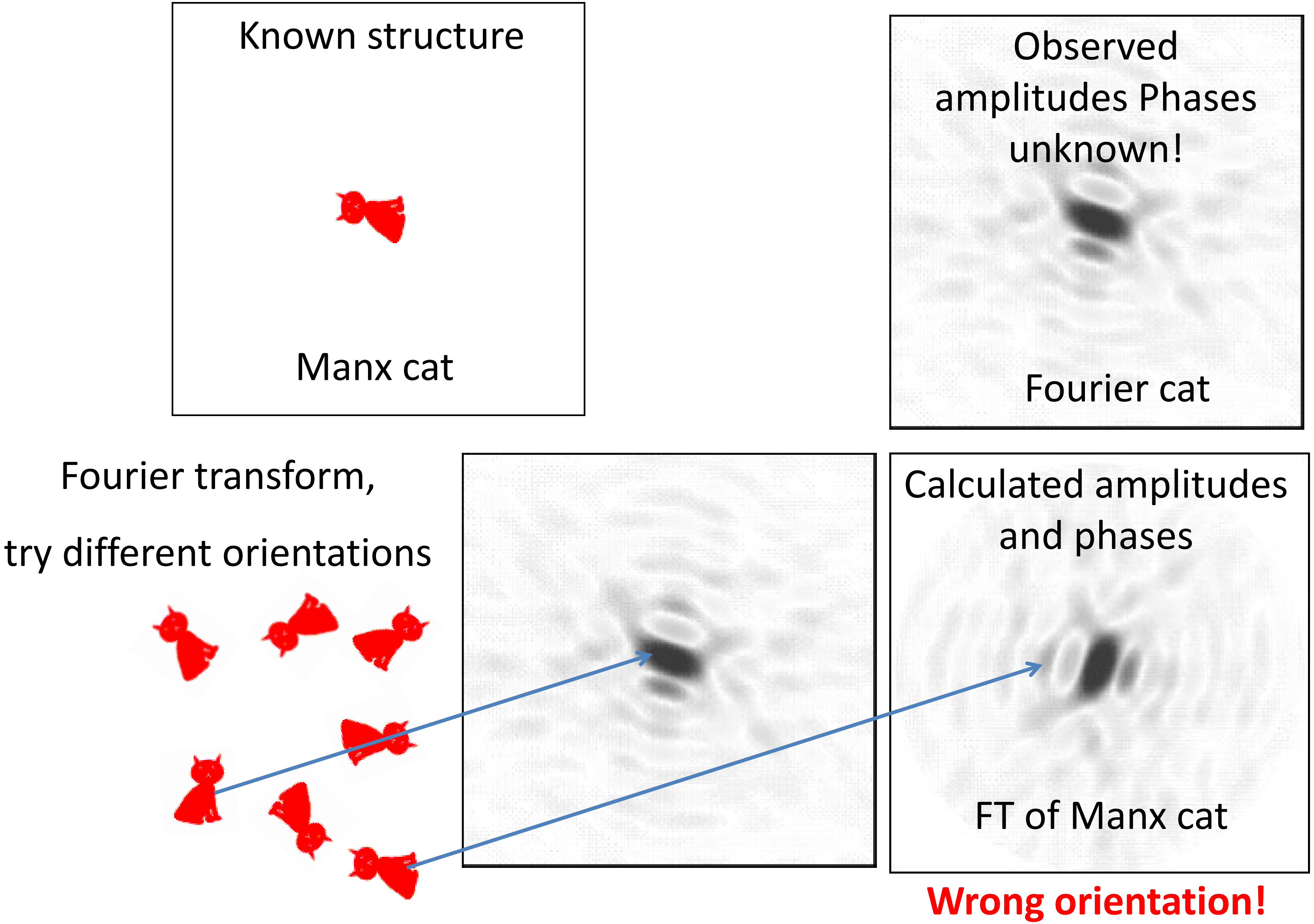
Fourier transform

Calculated amplitudes and phases

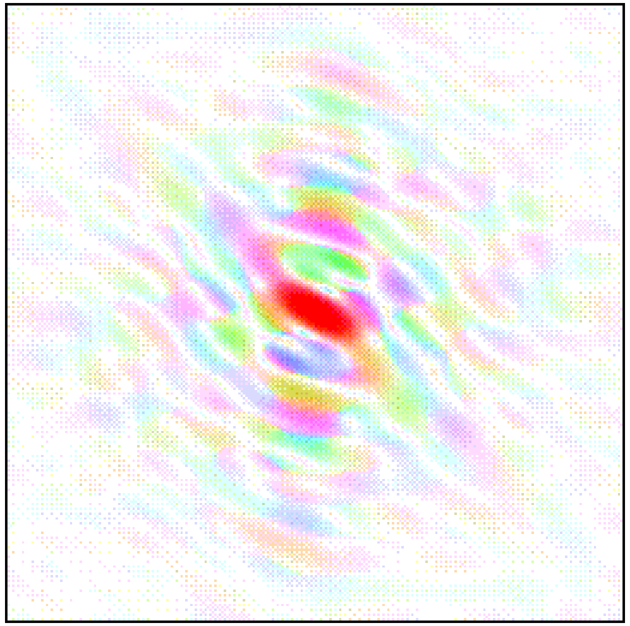


FT of Manx cat

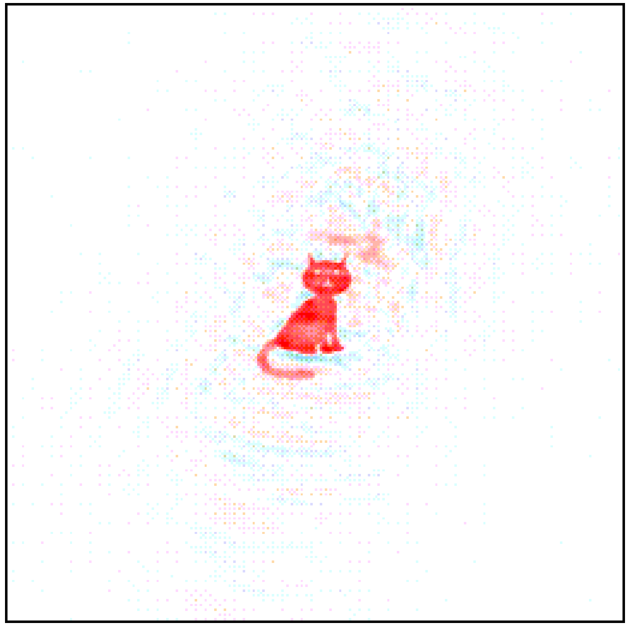
Wrong orientation!


Known structure
Manx cat
Observed amplitudes Phases unknown!
Fourier cat
Fourier transform,
try different orientations
Calculated amplitudes and phases
FT of Manx cat
Wrong orientation!

Observed **amplitudes** (tailed cat), calculated **phases** (Manx cat)

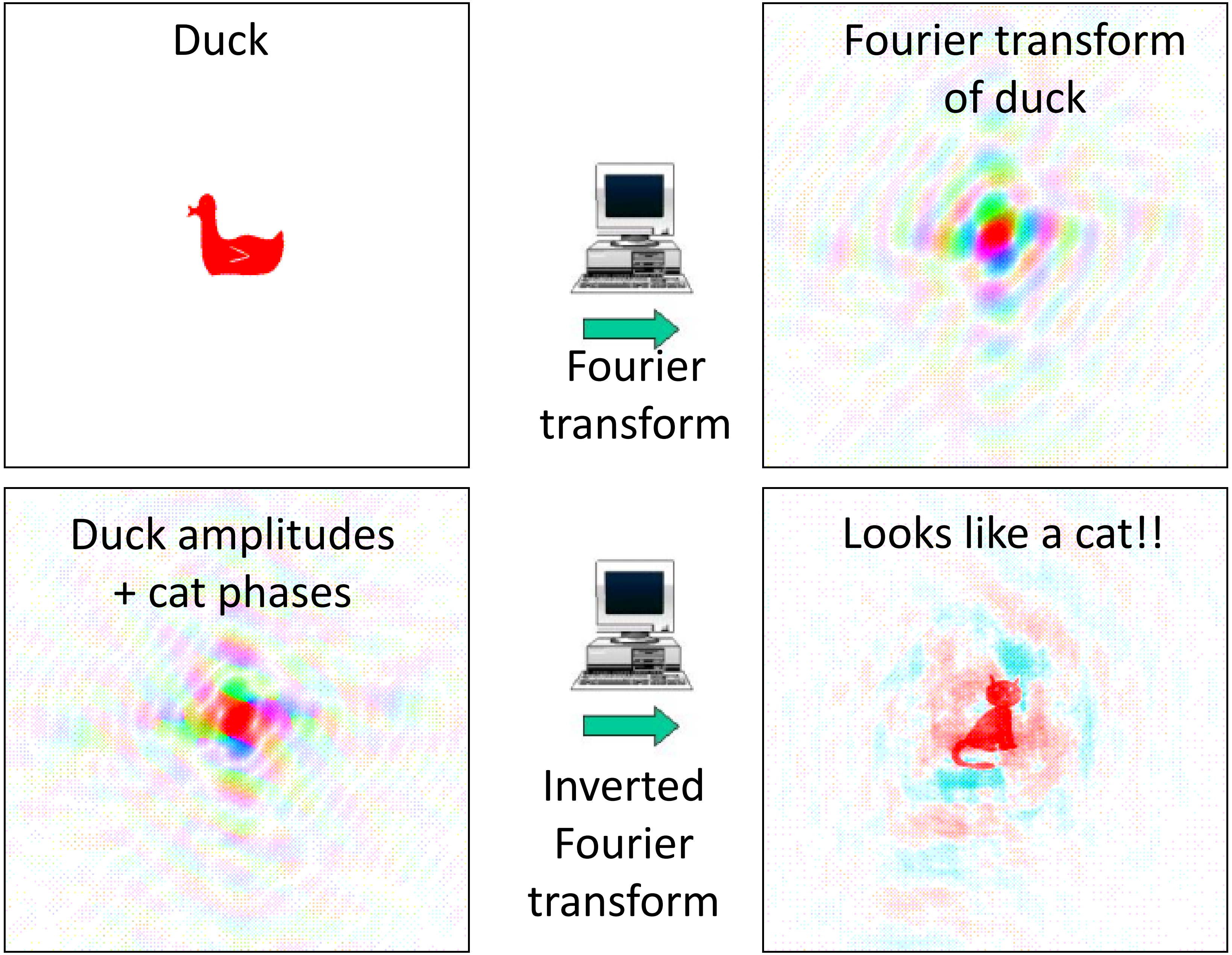


Inverted Fourier transform



Even the tail becomes visible!

Model Bias


Duck
Fourier transform
Fourier transform
of duck
Duck amplitudes
+ cat phases
Inverted
Fourier
transform
Looks like a cat!!

Solving the phase problem by:

Multiple/Single Isomorphous Replacement (MIR/SIR)

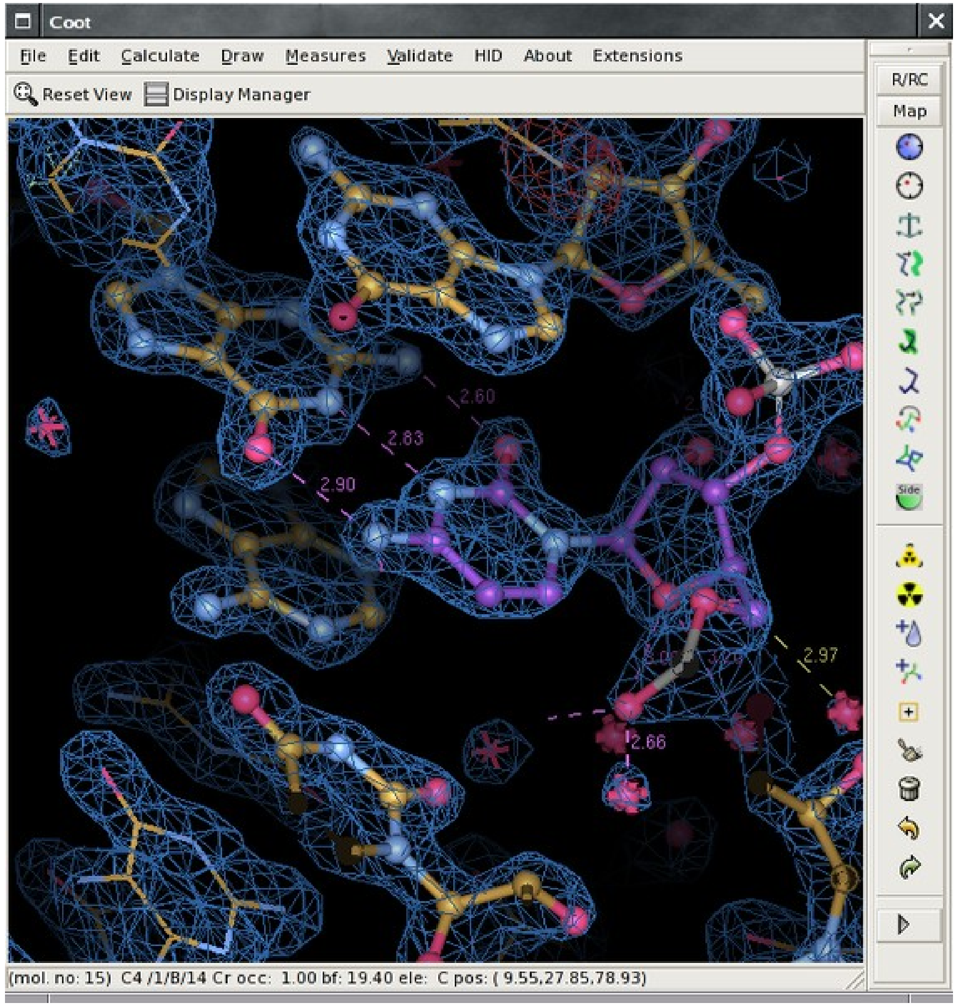
- source of phases – intensity differences between data from native and derivative (heavy atom containing) crystals
- Positions of heavy atoms identified from isomorphous difference Patterson maps

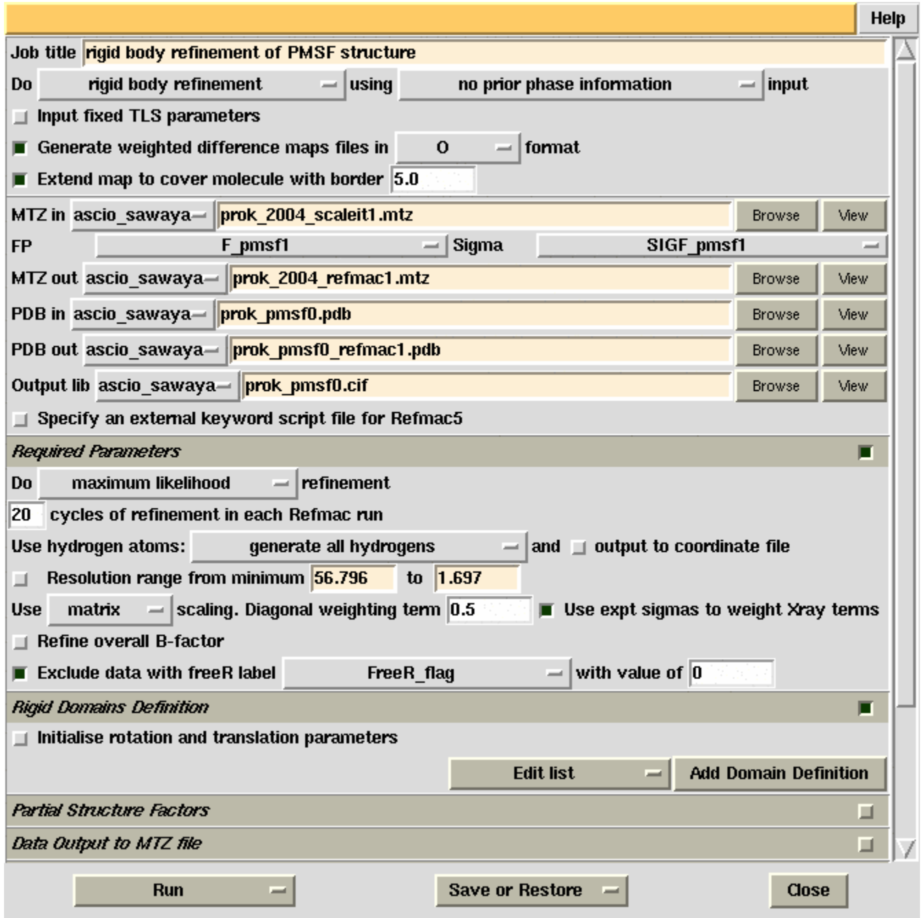
Solving the phase problem by:

Multiple/Single-wavelength anomalous diffraction (MAD/SAD)

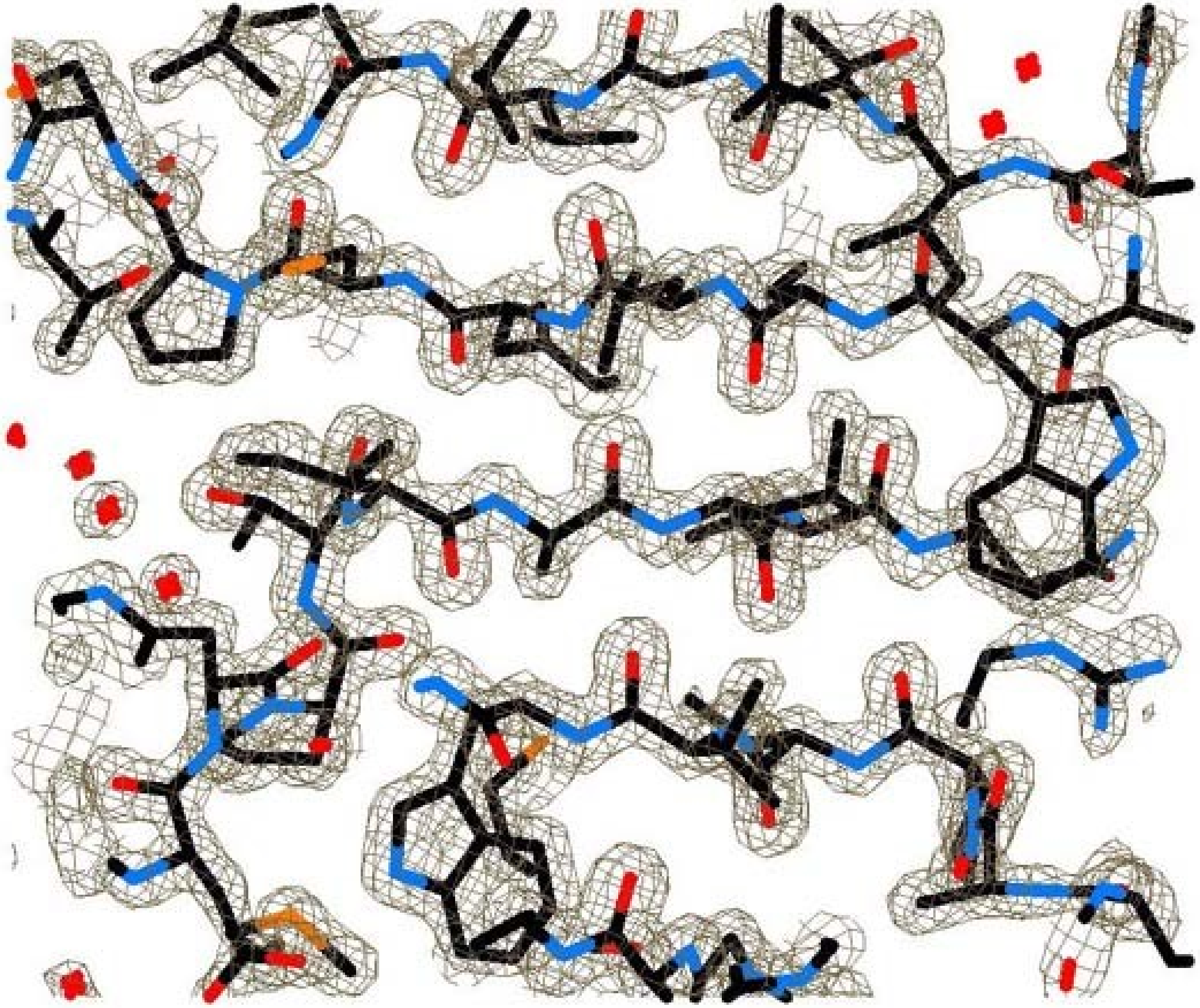
- source of phases – intensity differences between structure factors due to the presence of atom that specifically interacts with X-rays of a given wavelength
- Positions of heavy atoms identified from anomalous difference Patterson maps

Model building & refinement

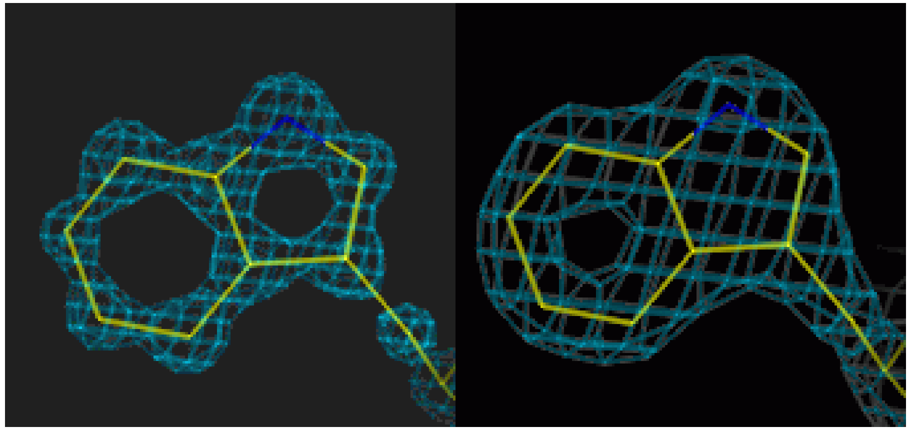




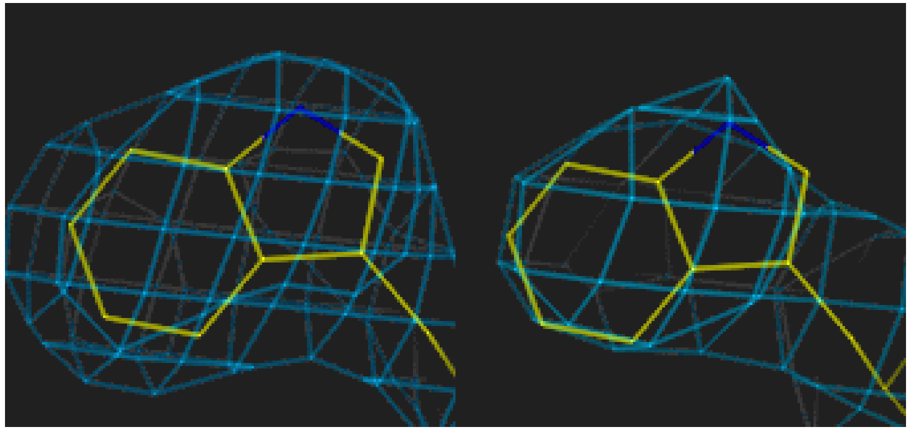
Model building & refinement



Model building & resolution



1.0Å 2.5Å



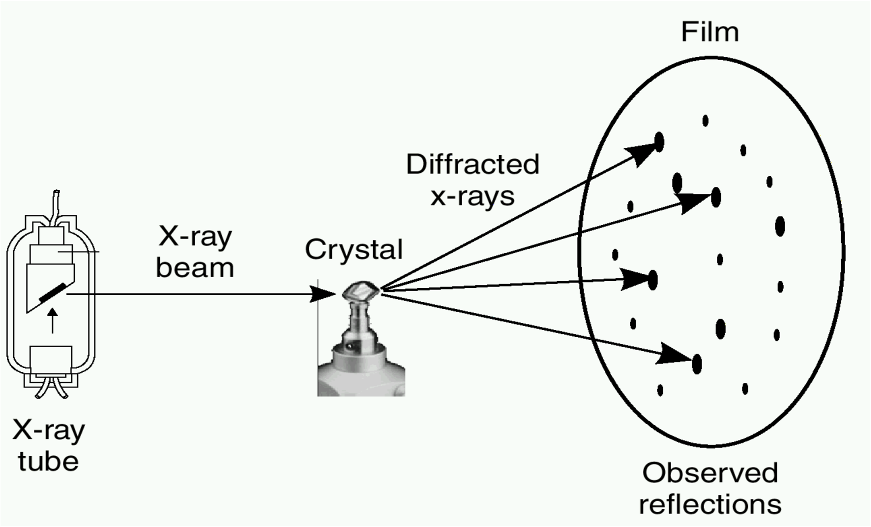
3.0Å 4.0Å

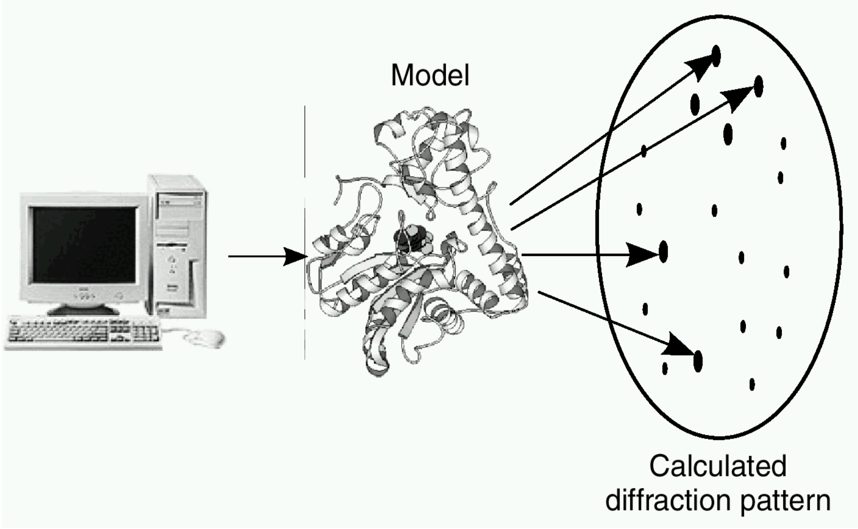
Validation

Assesment of model quality:

- Is the model in agreement with experimwntal data?
- How the geometry of amino acids look like?
- Are atoms far / close enough from each other?
- Are residues “happy” in their environment?
- Are the hydrogen donors/acceptors satisfied?

R-factor, R_{free} factor





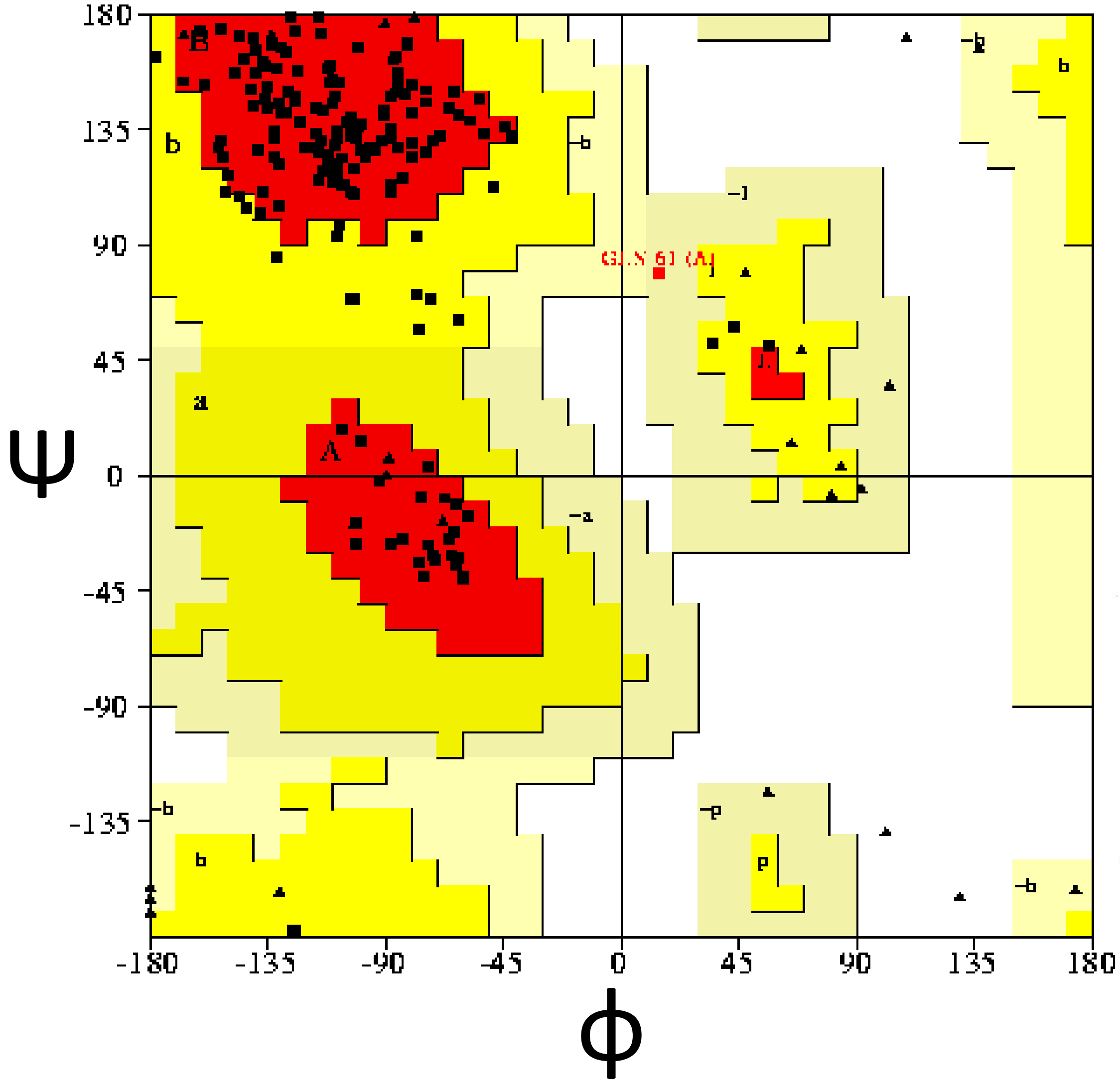
R-factor

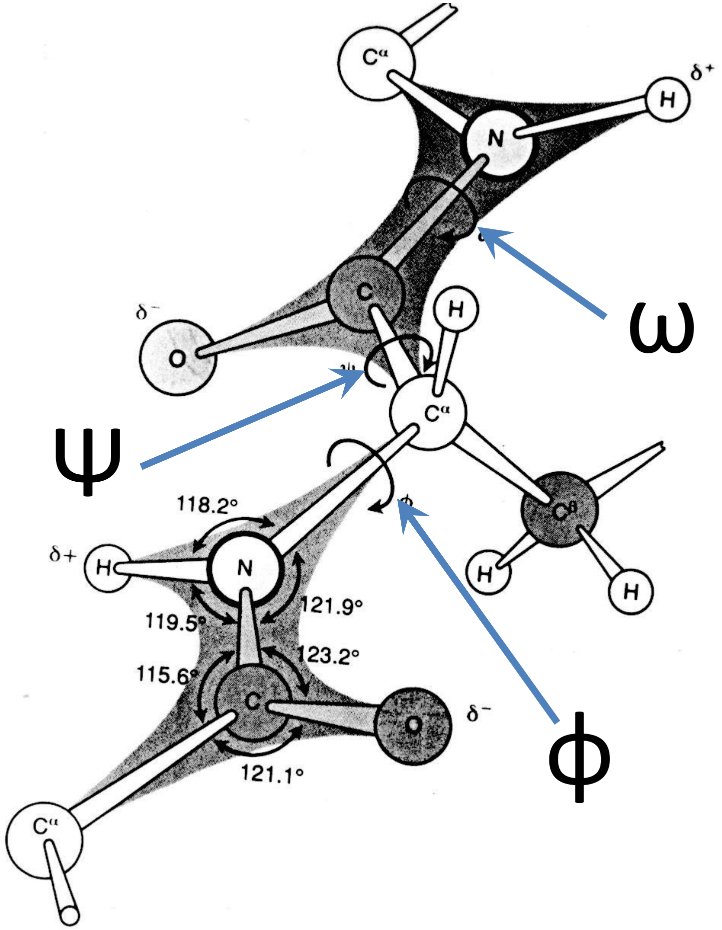
$$R = \frac{\sum_{hkl} ||F_{\text{obs}}| - k|F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$$

R_{free} factor

$$R_{\text{free}} = \frac{\sum_{hkl \subset T} ||F_{\text{obs}}| - k|F_{\text{calc}}||}{\sum_{hkl \subset T} |F_{\text{obs}}|}$$

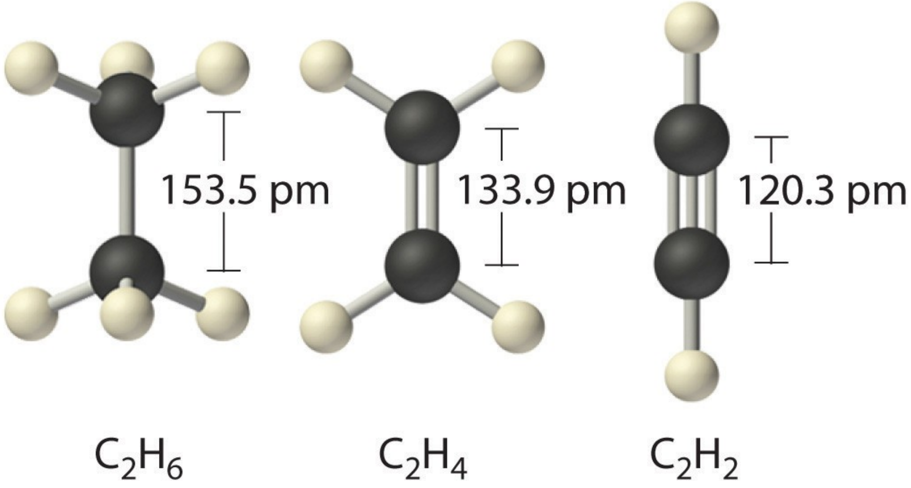
Ramachandran plot



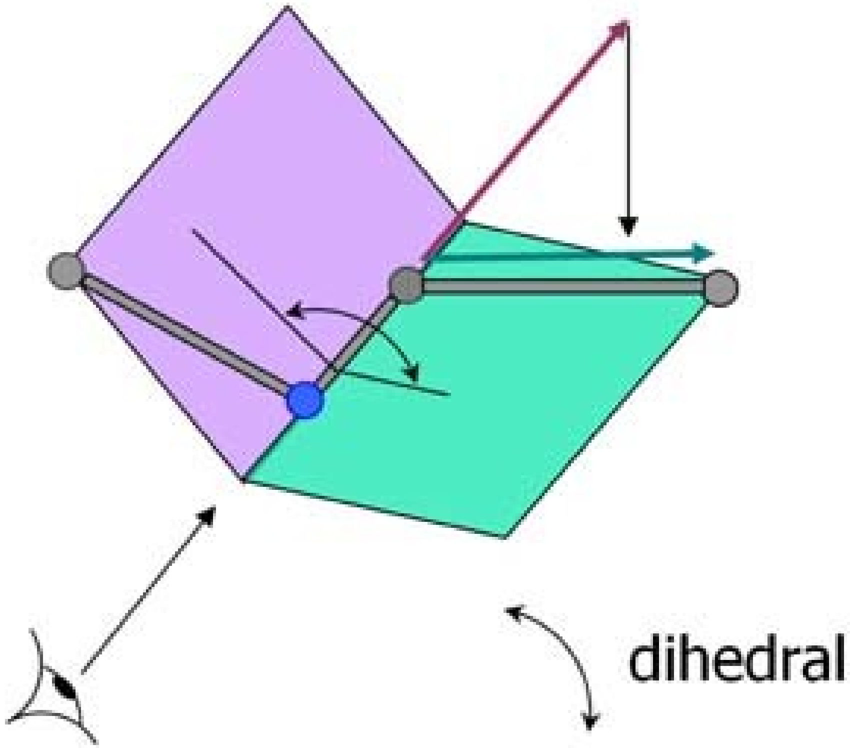


Geometry and stereochemistry

Bond lengths

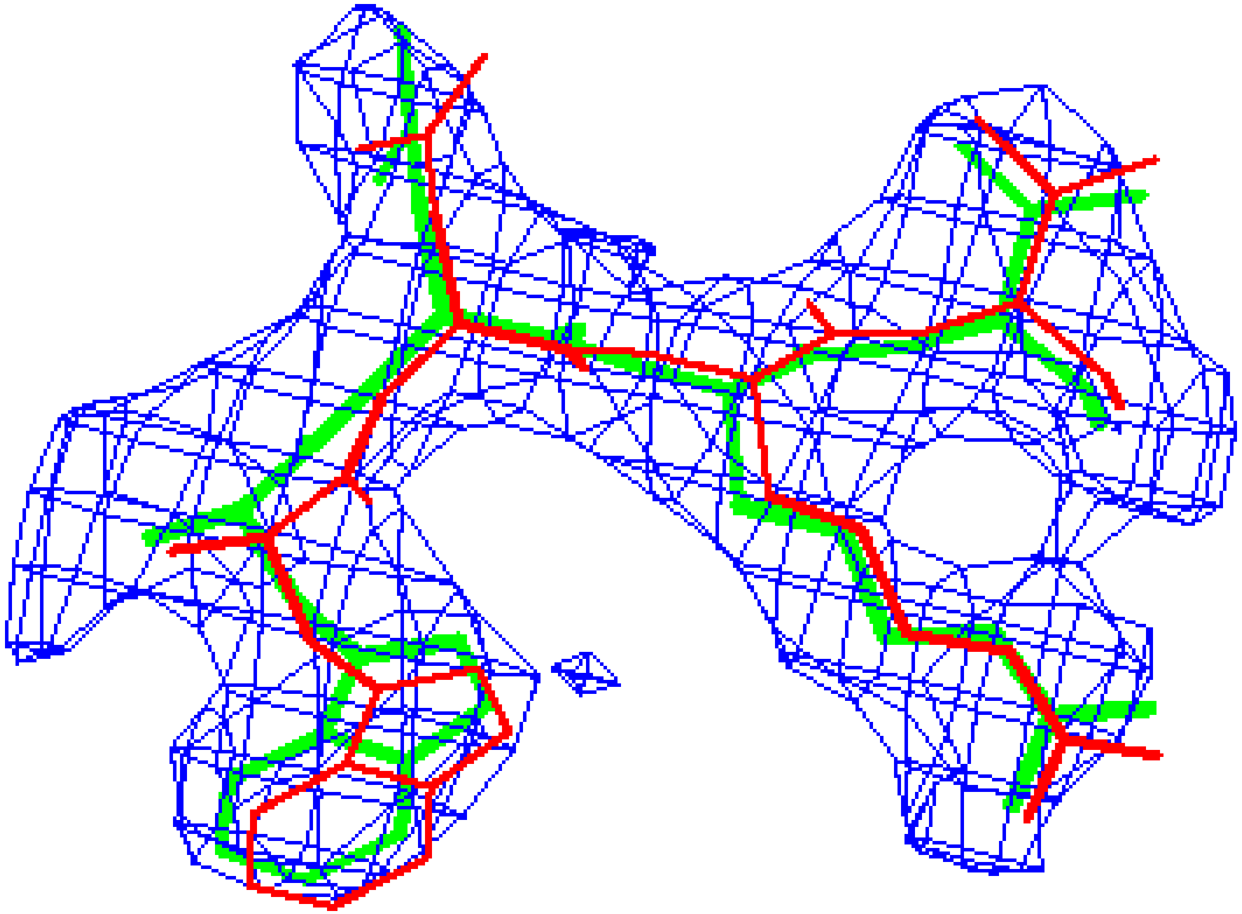


Dihedral angles



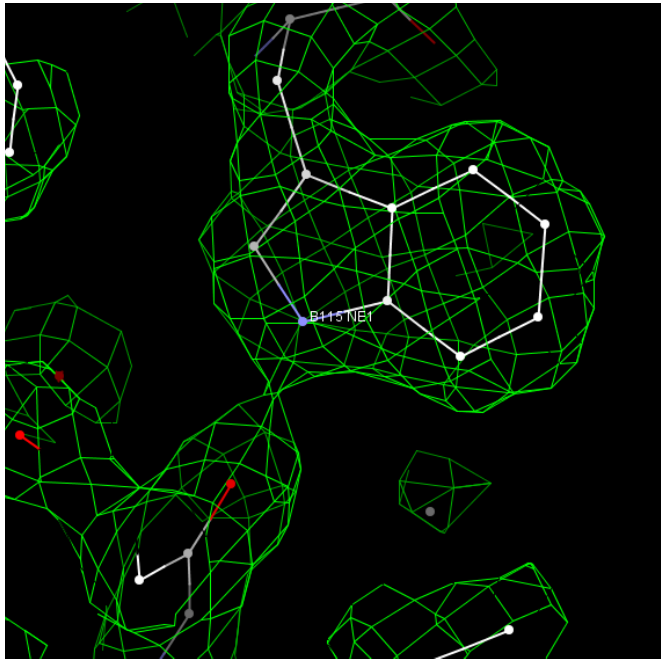
$$\mathrm{RMSD} = \sqrt{\frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n}}.$$

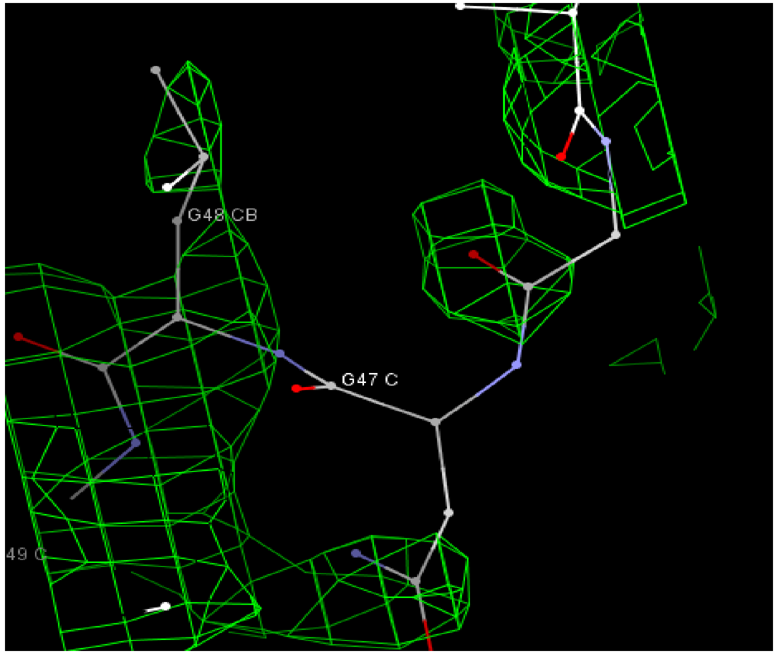
Real-space fit



Data deposition

- Protein Data Bank (PDB)
- Some structures are wrong!





Summary

1. X-rays have suitable wavelength for study of molecular structures

2. Crystals allow measurement of useful diffraction data because they diffract strongly in certain directions

3. Our goal is to obtain three-dimensional distribution of electron density, because it shows the shape of a molecule

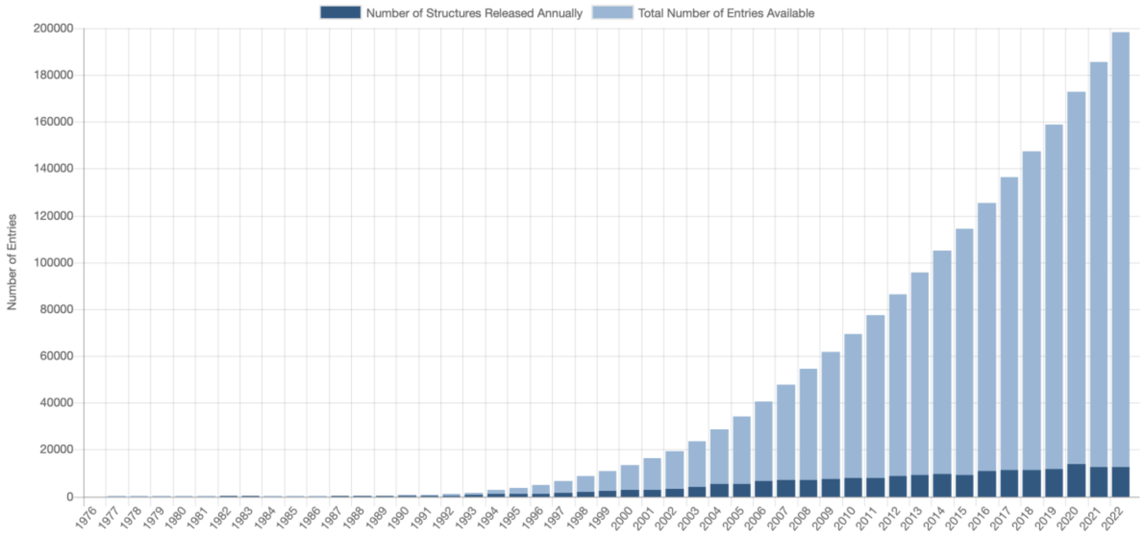
4. Diffraction experiments provide only amplitudes of structure factors => Phase problem

5. Solution of the phase problem:
   - Molecular replacement
   - Isomorphous replacement
   - Anomalous diffraction

6. Model building, refinement, validation, deposition

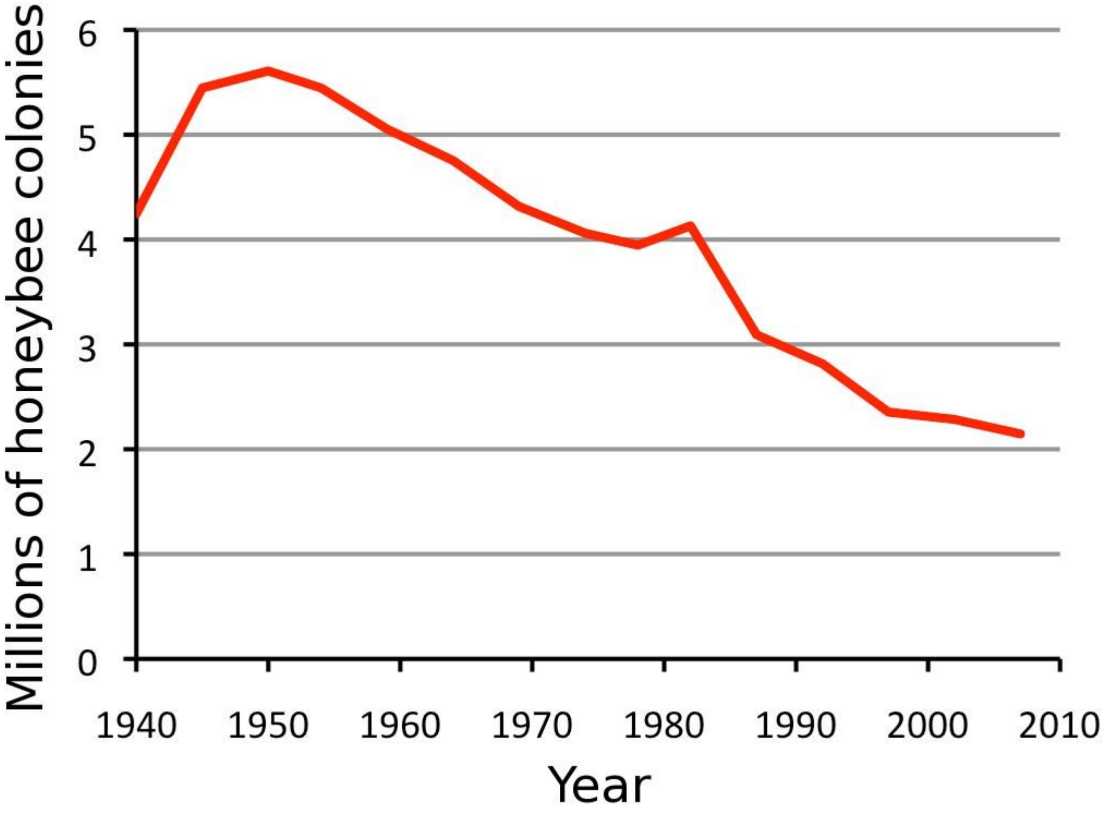
X-ray crystallography

- First method to determine structure of molecules with atomic resolution
- As of November, 2022 there were almost 200,000 structures available from Protein Data Bank (170,000 by X-ray)
- Macromolecular structures are crucial for our understanding of life at the molecular level
- 28 Nobel prizes



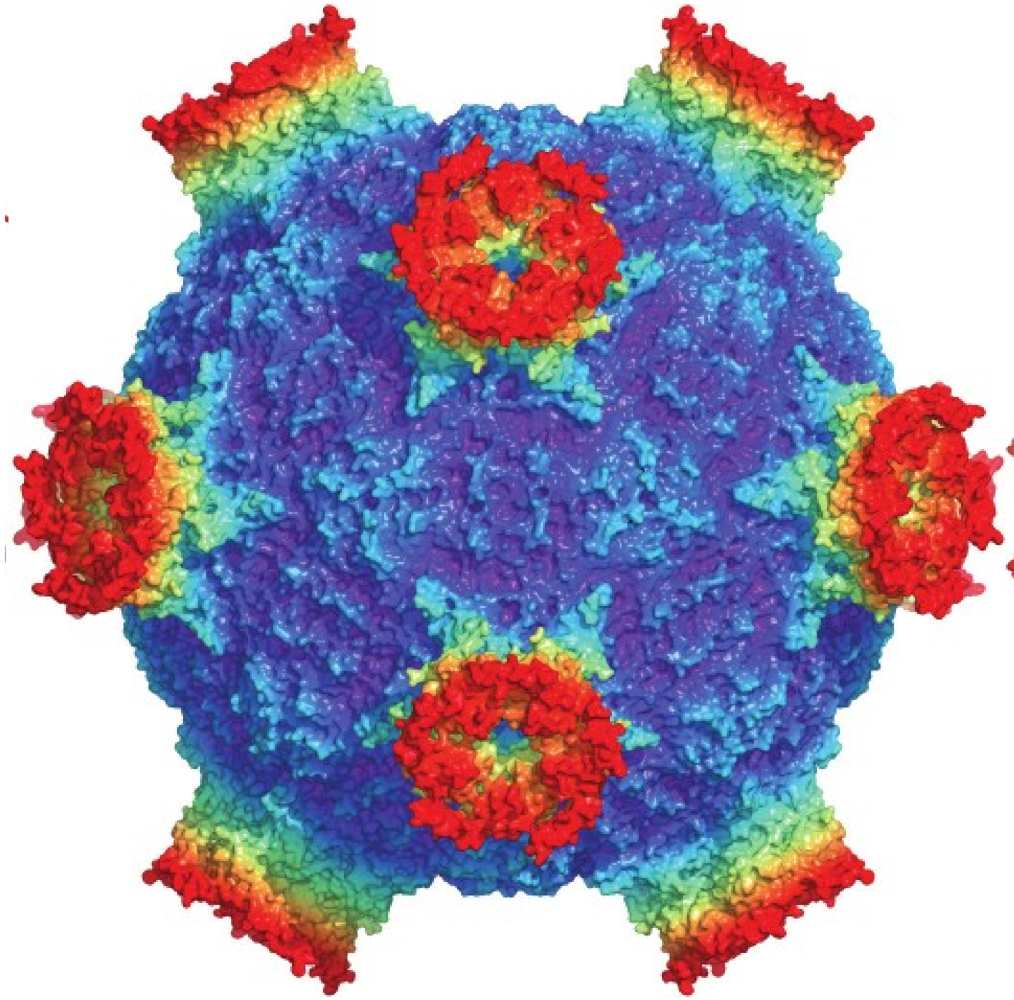
Deformed wing virus

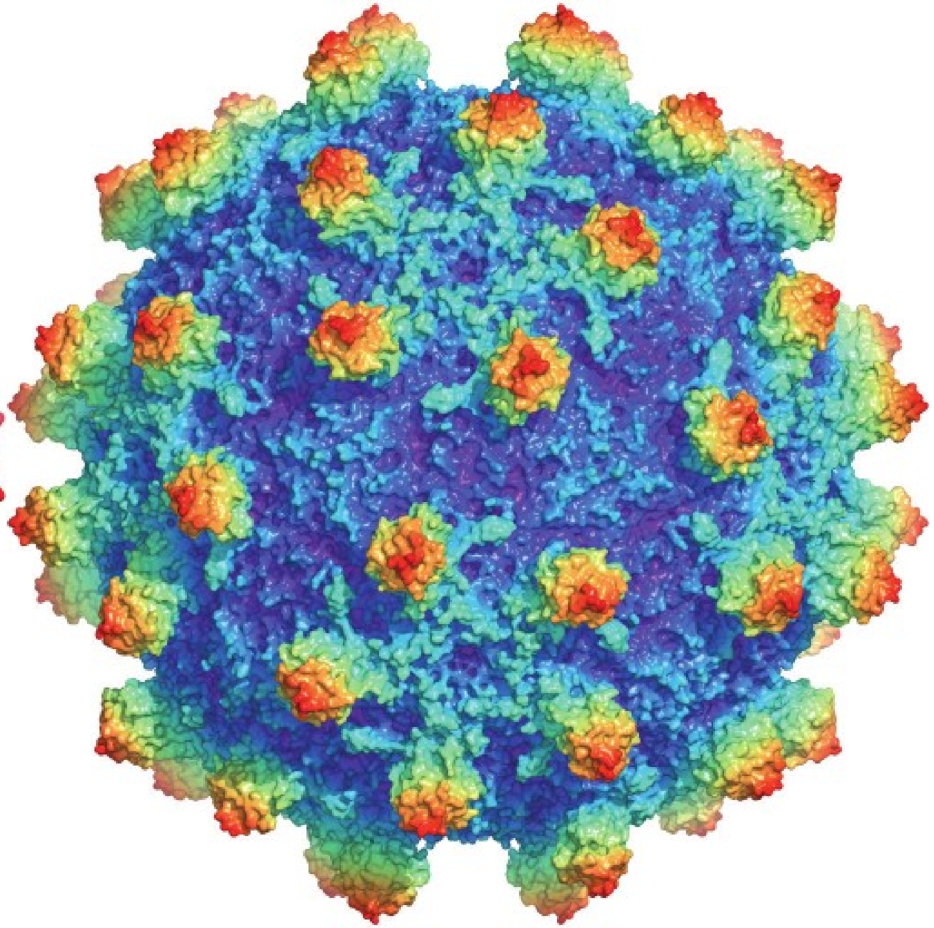
Honeybee colonies in US

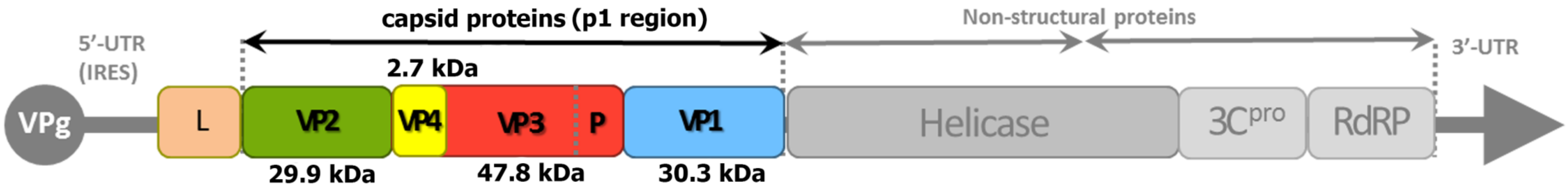


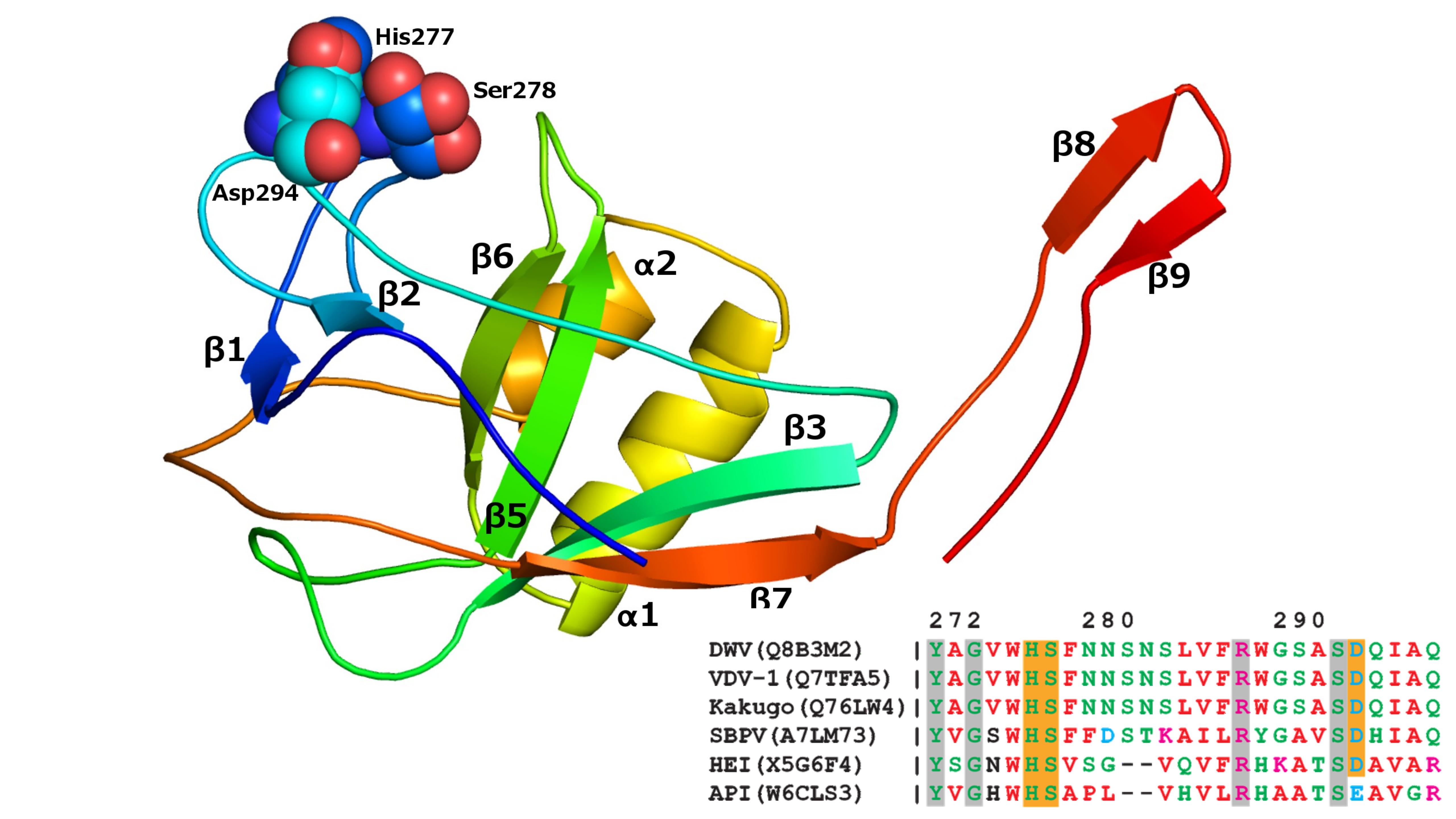
DWV infected pupae

Deformed wing virus








His277
Ser278
Asp294
β1
β2
β3
β5
β6
α1
α2
β7
β8
β9
272
280
290
DWV(Q8B3M2) | Y A G V W H S F N N S N S L V F R W G S A S D Q I A Q
VDV-1(Q7TFA5) | Y A G V W H S F N N S N S L V F R W G S A S D Q I A Q
Kakugo(Q76LW4) | Y A G V W H S F N N S N S L V F R W G S A S D Q I A Q
SBPV(A7LM73) | Y V G S W H S F F D S T K A I L R Y G A V S D H I A Q
HEI(X5G6F4) | Y S G N W H S V S G - - V Q V F R H K A T S D A V A R
API(W6CLS3) | Y V G H W H S A P L - - V H V L R H A A T S E A V G R

Thank you!

1.) Jakou část strukturního faktoru můžeme změřit v difrakčním experimentu:

a) amplitudu (ve formě intensity)

b) fázi

2.) Nejčastější metoda pro získání fází je:

a) molekulární nahrazení (molecular replacement)

b) isomorfní nahrazení

c) anomální diffrakce

3.) Ramachandran plot ukazuje:

a.) distribuci úhlů v hlavním řetězci proteinu

b.) vzdálenosti mezi atomy

c.) konformace postranních řetězců aminokyselin

1. Rentgenové paprsky se používají ke studiu makromolekulárních struktur protože:
A.) Mají vlnovou délku podobnou meziatomovým vzdálenostem.
B.) Jako jediné elektromagnetické záření interagují s biologickým materiálem.
C.) Byly objeveny v době intenzivního zájmu o strukturu makromolekul a z historických důvodů se používají dodnes.

2. To, že makromolekuly tvoří krystaly znamená že:
A.) Mají enzymatickou aktivitu
B.) Jsou součástí kostry buňky (cytoskeletu)
C.) Mají stabilní strukturu.

3. Mapa elektronové hustoty, která je výsledkem rentgenové analýzy krystalů:
A.) Ukazuje tvar molekul, které tvoří krystal
B.) Má vždy bílou barvu
C.) Ukazuje tvar molekuly po denaturaci