

# X-ray crystallography

## Sources:

1. The Elements of Physical Chemistry by Peter Atkins (including some nice color slides incorporated into these lectures)
2. Physical Chemistry by Tinoco, Sauer, Wang and Puglisi (a more complete treatment than Chang)
3. Physical Chemistry by Chang
4. Internet:
  - a. Crystallography 101 at <http://www.ccp14.ac.uk/ccp/web-mirrors/llnrupp/Xray/101index.html>. A nice introduction at a fairly detailed level.
  - b. An advance crystallography course in PowerPoint at: <http://isites.bio.rpi.edu/bystrc/pub/Biophysics/> Add lecture.n.ppt to this URL (n = 1-6) to get 6 technically detailed lectures.
  - c. Protein Data Bank (PDB) at URL <http://www.rcsb.org/pdb/>

# What is X-ray Crystallography ?

X-ray crystallography is an experimental technique that exploits the fact that X-rays are **diffracted** by **crystals**. It is not an imaging technique.

1. X-rays have an appropriate wavelength (in the Ångström range,  $\sim 10^{-8}$  cm) to be scattered by the electron cloud of an atom of comparable size.
2. A diffraction pattern is obtained from X-rays “reflected” (scattered) off the periodic assembly of molecules or atoms in the crystal.
3. The electron density, and hence the positions of atoms, can be reconstructed from the diffraction pattern.
4. The diffraction data show the **intensities** arising from constructive and destructive interference of the scattered wave fronts, but do not show the **phases** of the wave fronts. Since both types of information are needed, this deficiency is known as “**the phase problem**” in crystallography.

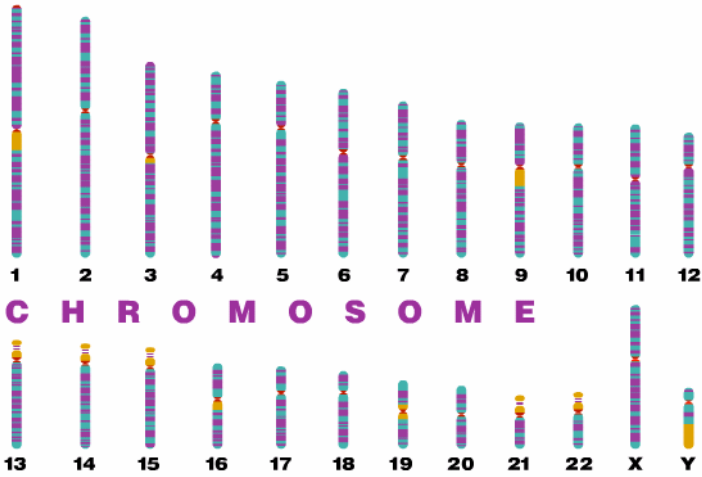
## Solving the “phase problem”

5. In order to complete the reconstruction , additional phase information must be extracted either from the diffraction data itself (possible for small molecules) or by supplementing diffraction experiments with additional information.
6. This involves adding atoms that are easily identified to the crystal, which are a part of the same crystal lattice. The phases of these atoms can be solved, because they can be treated as a “simple” crystal.
7. Techniques for this latter approach are Multiple Isomorphous Replacement (**MIR**), and Multiwavelength Anomalous Diffraction (**MAD**). An alternative approach to the phase problem is use of multiple X-ray wavelengths, or **Laue diffractometry**.
8. Once the phase problem has been solved, a model is then progressively built into the experimental electron density, refined against the data. The result is a quite accurate molecular structure.

# Why do we need information about macromolecular structure?

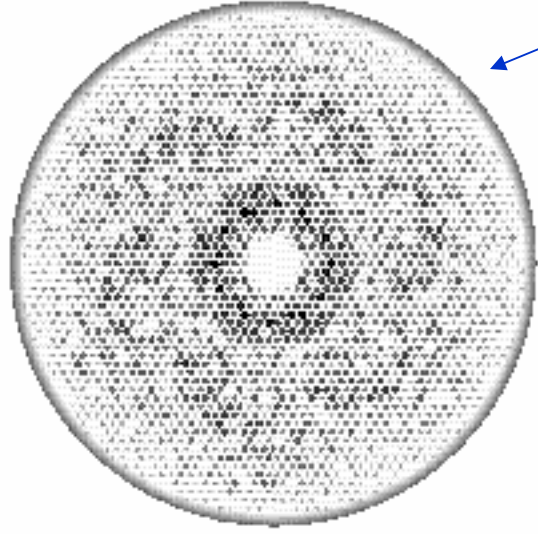
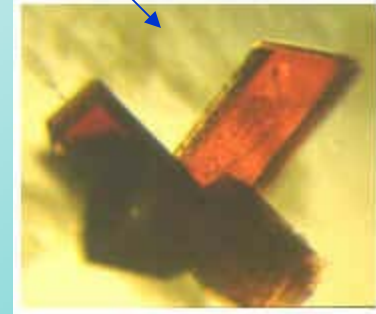
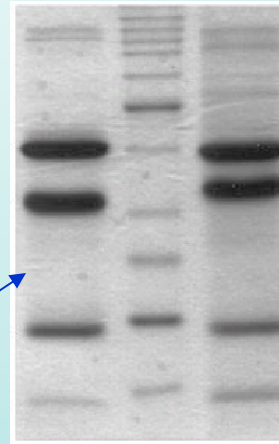
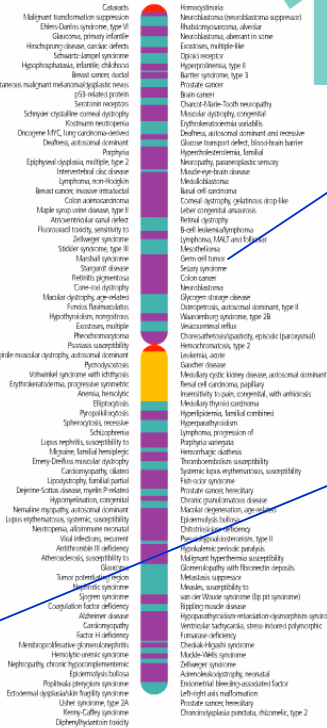
1. Understanding mechanism
  - a. Binding of substrates, products, inhibitors
  - b. Pathways for electron and proton, etc. transfer
  - c. Pathways for transport of metabolites, H<sub>2</sub>O, ions, etc.
  - d. Conformation changes important to mechanism
2. Subunit interactions
3. Protein-protein interactions
4. Protein nucleic acid interactions
5. Rational drug design
6. General philosophical interest

# Human genome project – summary of insights

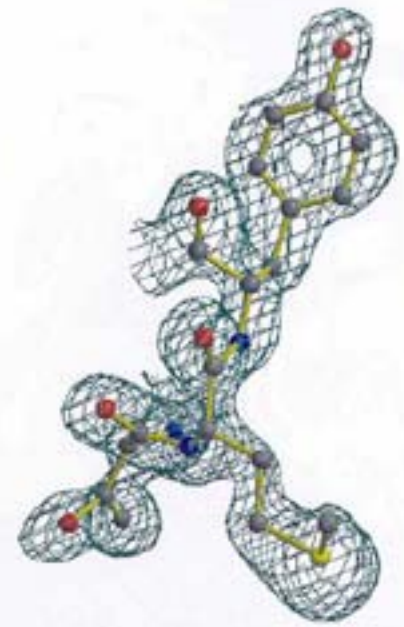


[Link to clickable chromosomes](#)

263 million bases



30-35,000 different genes



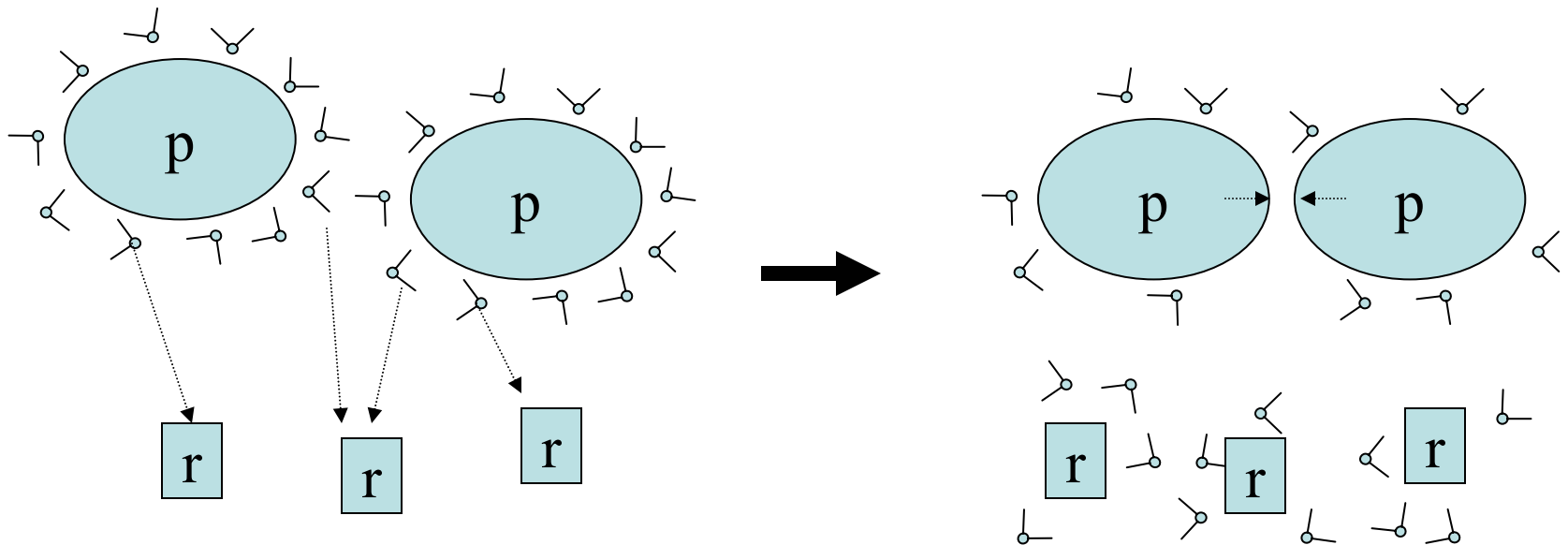
## First step – crystallization of protein

General idea, - start with protein at high concentration, and then increase concentration by removal of solvent (water). Precipitant added to aid in redistribution of water. Aim for metastable conditions and slow growth of crystals for best quality

<http://www.ccp14.ac.uk/ccp/web-mirrors/llnlrupp/Xray/101index.html>

### Crystallography 101

p = protein; r (precipitant) = EtOH,  $(\text{NH}_4)_2\text{SO}_4$ , methylpentanediol, polyethylene glycol, etc

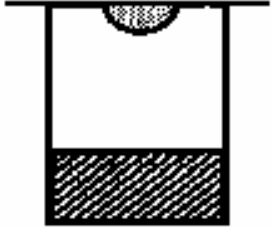


vapor diffusion

a Linbro plate

[Table of crystallization protocols](#)

Hanging Drop  
with Protein



Reservoir with  
Precipitant



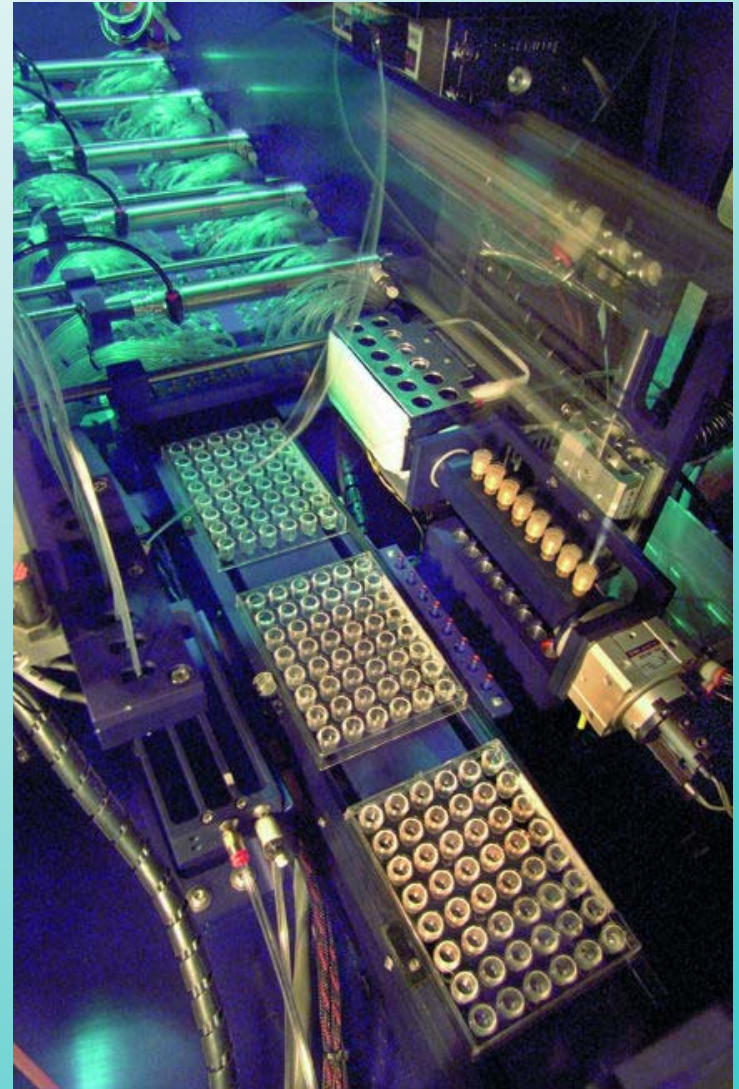
The most common setup to grow protein crystals is by the hanging drop technique : A few microliters of protein solution are mixed with an about equal amount of reservoir solution containing the precipitants. A drop of this mixture is put on a glass slide which covers the reservoir. As the protein/precipitant mixture in the drop is less concentrated than the reservoir solution (remember: we mixed the protein solution with the reservoir solution about 1:1), water evaporates from the drop into the reservoir. As a result the concentration of both protein and precipitant in the drop slowly increases, and crystals may form.

The Hanging Drop approach is commonly done in batches, using a Linbro plate, and a matrix of crystallization conditions.

With the genomic revolution a *fait accompli*, we are stuck with the coding (and hence the amino acid sequences) of millions of proteins. For any genome, we have no idea of the function of about a third of these. As our understanding of the functional significance of these improves, the need to get structures will become increasingly important.

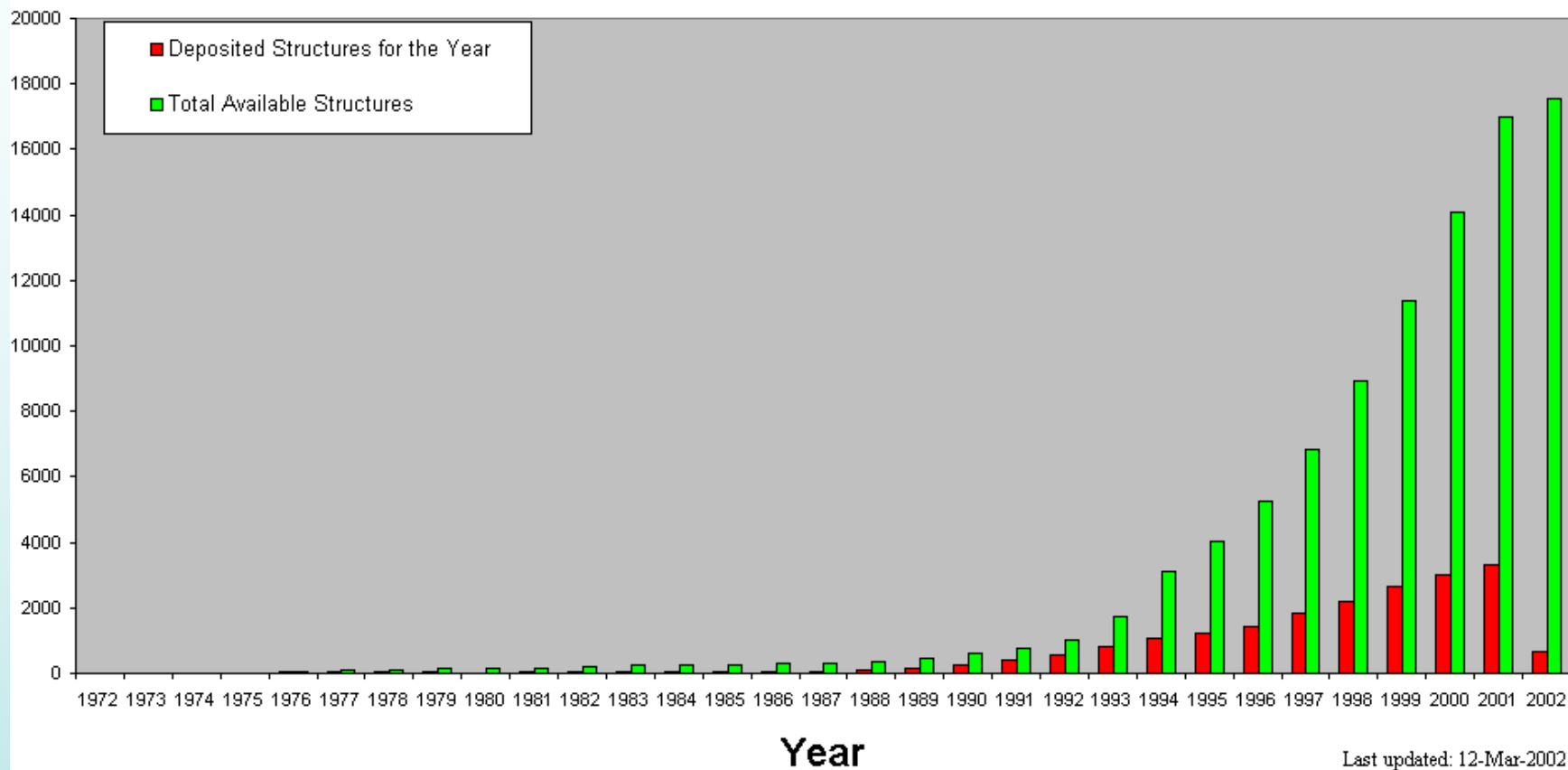
Currently, the [Protein Data Bank](#) (PDB) contains 17737 structures, but many of these are “repeats” (higher resolution structures than previous, same protein from different organisms, mutant versions of same protein, etc.). The unique structures currently available represent a very small fraction of those we would like to have.

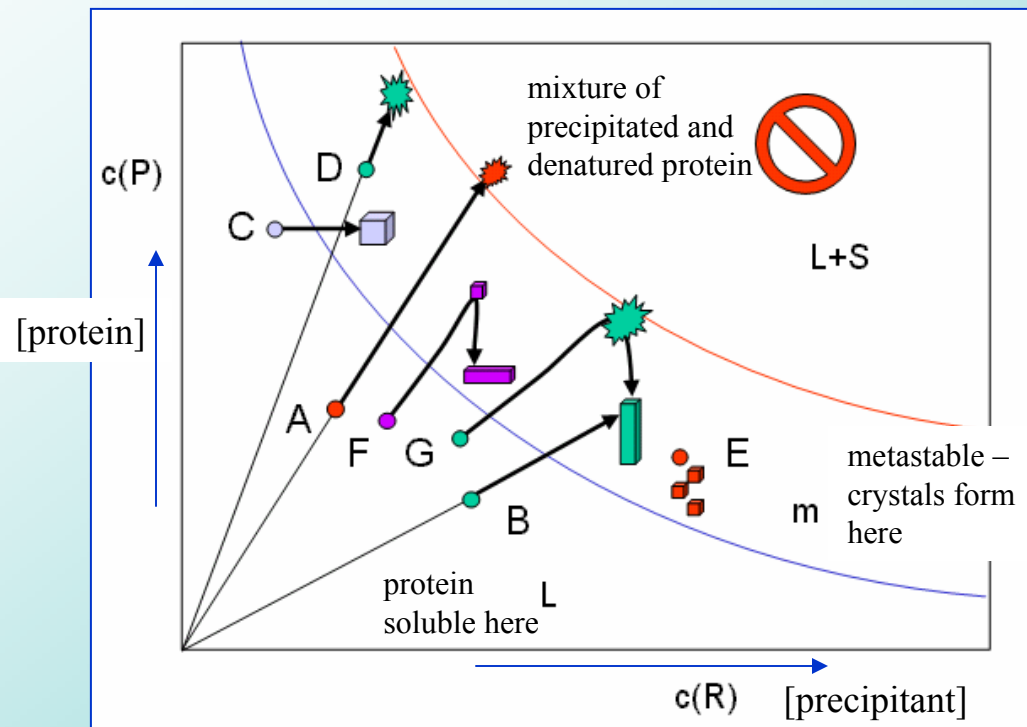
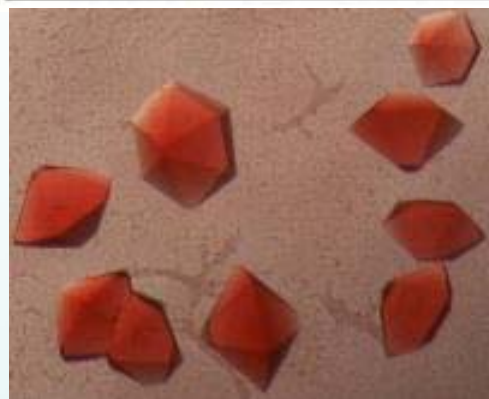
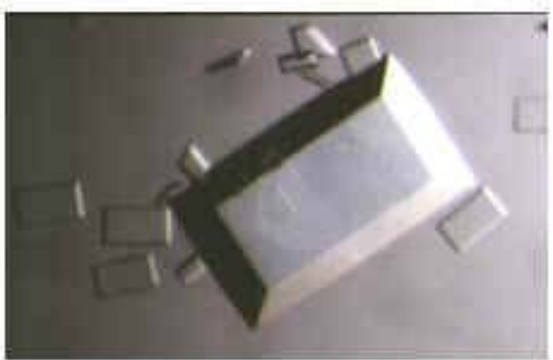
In order to increase throughput, the exploration of growth conditions for crystals can be highly automated (see right).





## Growth of the Protein Data Bank since first structures



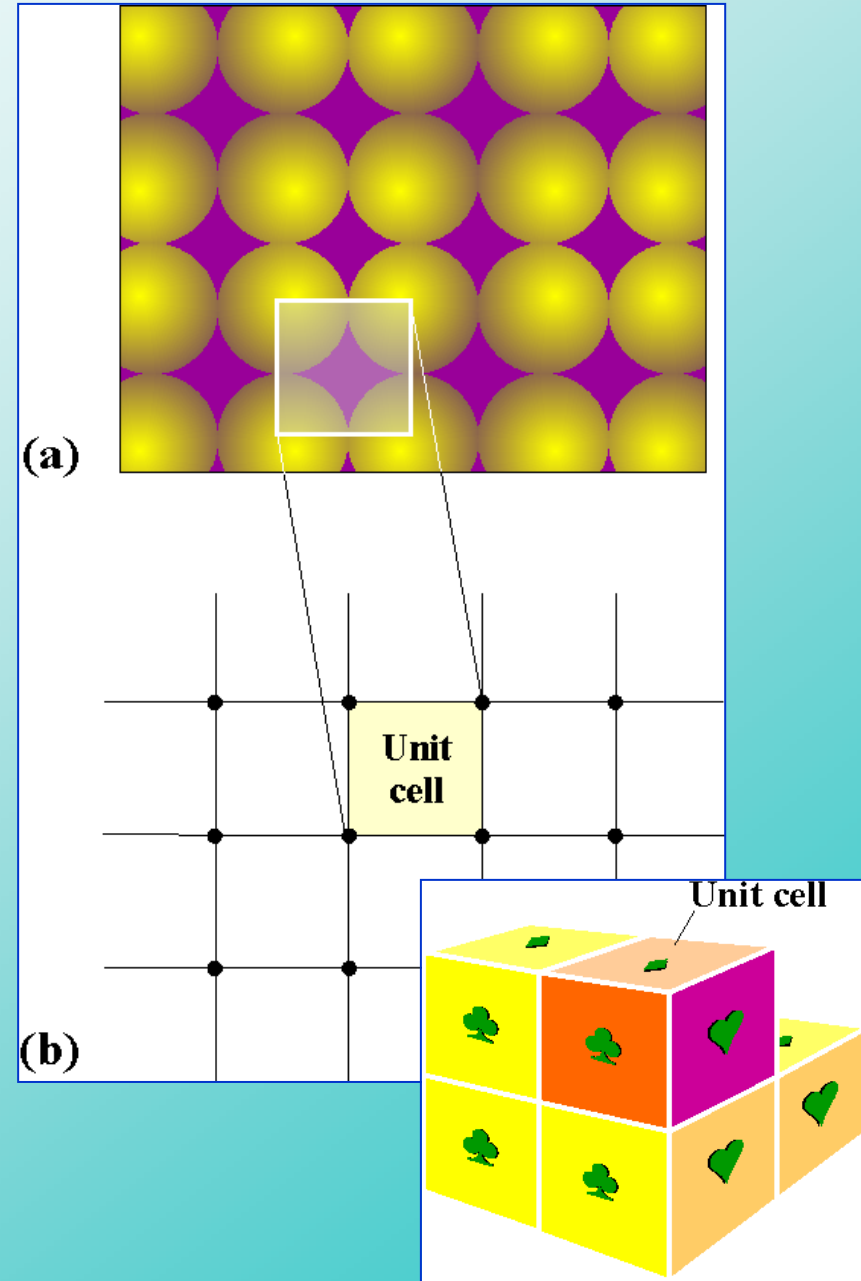


The diagram above shows some pathways of crystallization experiments through a phase diagram, with initial conditions represented as dots. In a vapor diffusion (VD) experiment, the ratio of protein to precipitant remains constant, and we actually move away from the water corner (see arrows starting at origin). The extension of the path has to go through the origin for VD experiments. A, B, D, G and F represent schematic VD experiments, although with different results. C describes the path for a dialysis experiment against higher precipitant concentration. It is important to note, that the precipitate endpoint of A just at the border of the spinodal region (⊘) is strikingly different from an endpoint inside the metastable region (D,G).

## Step 2 – understanding diffraction

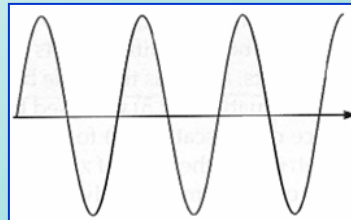
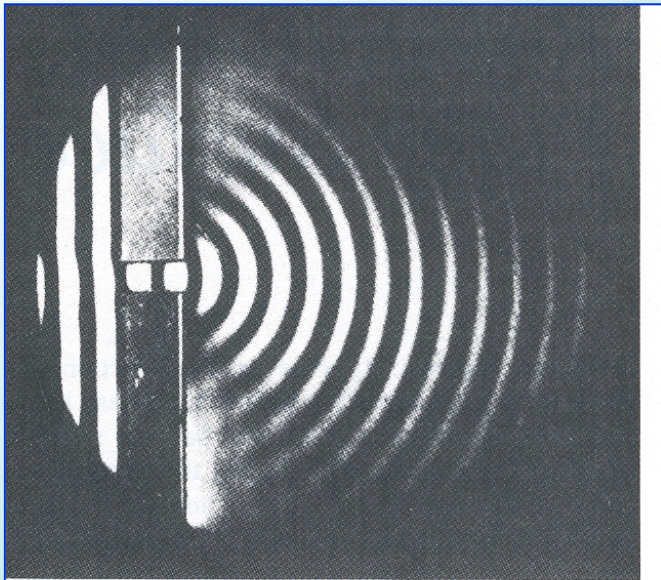
Crystals are formed by regular alignment of molecular constituents. The molecules are arranged in a repeat pattern. The repeat is made of copies of the **unit cell**. The unit cell is defined as the minimal repeating unit from which the crystal can be generated by simple rotations and translations. Note, this is **not necessarily a single copy of the constituent molecule**, - the unit cell has to provide the dimensional information that goes to determination of the crystal form.

The diagram (right, top) shows a two-dimensional lattice containing a two dimensional unit cell. The inset shows a 3-D version to emphasize the way in which the dimensional repeat, packing, orientation, etc., determine the crystal form.



**Newton**'s theory of light was "corpuscular"; he believed that light must be made of particles, because it didn't bend around corners in the way that waves were observed to do.

**Huygens** in contrast believed that "...an expanding sphere of light behaves as if each point on the wave front were a new source of radiation of the same frequency and phase."



Cross-section  
of a wave  
front

Huygens' principle  
illustrated by water waves

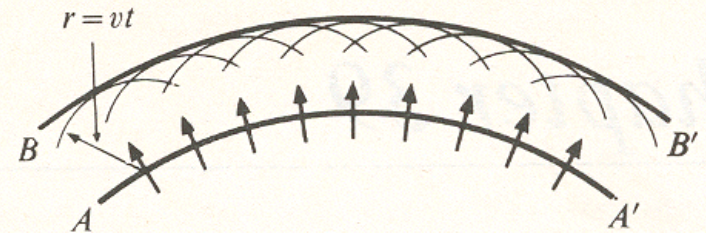
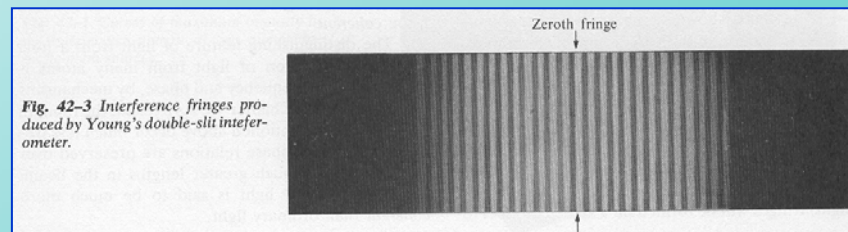
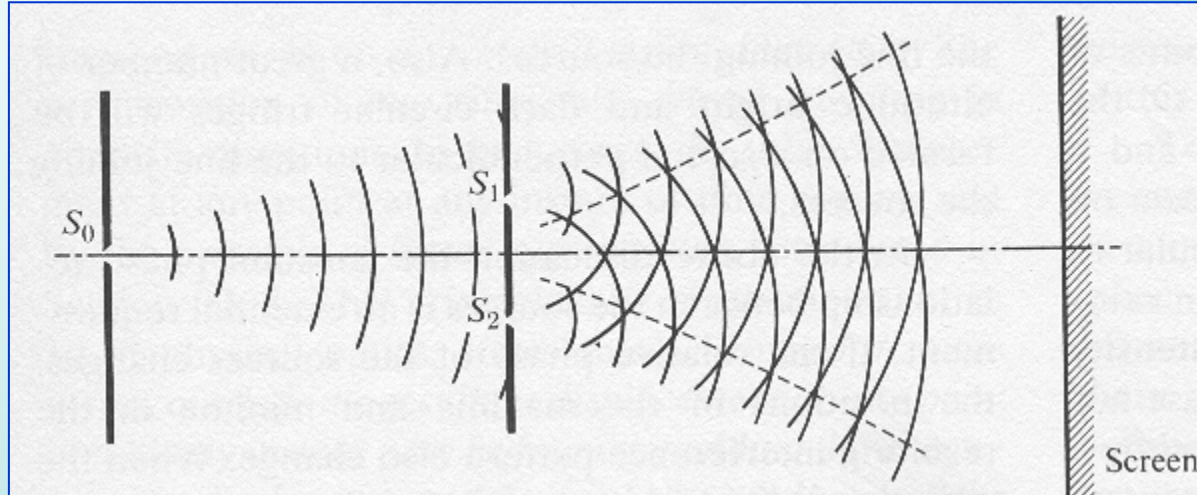


Fig. 39-1 Huygens' principle.

# Diffraction

Young's experiment with light shining through a set of slits showed **diffraction**. This was as expected from Huygens' hypothesis, and showed that light did bend around corners. Disproved Newton's corpuscular hypothesis.

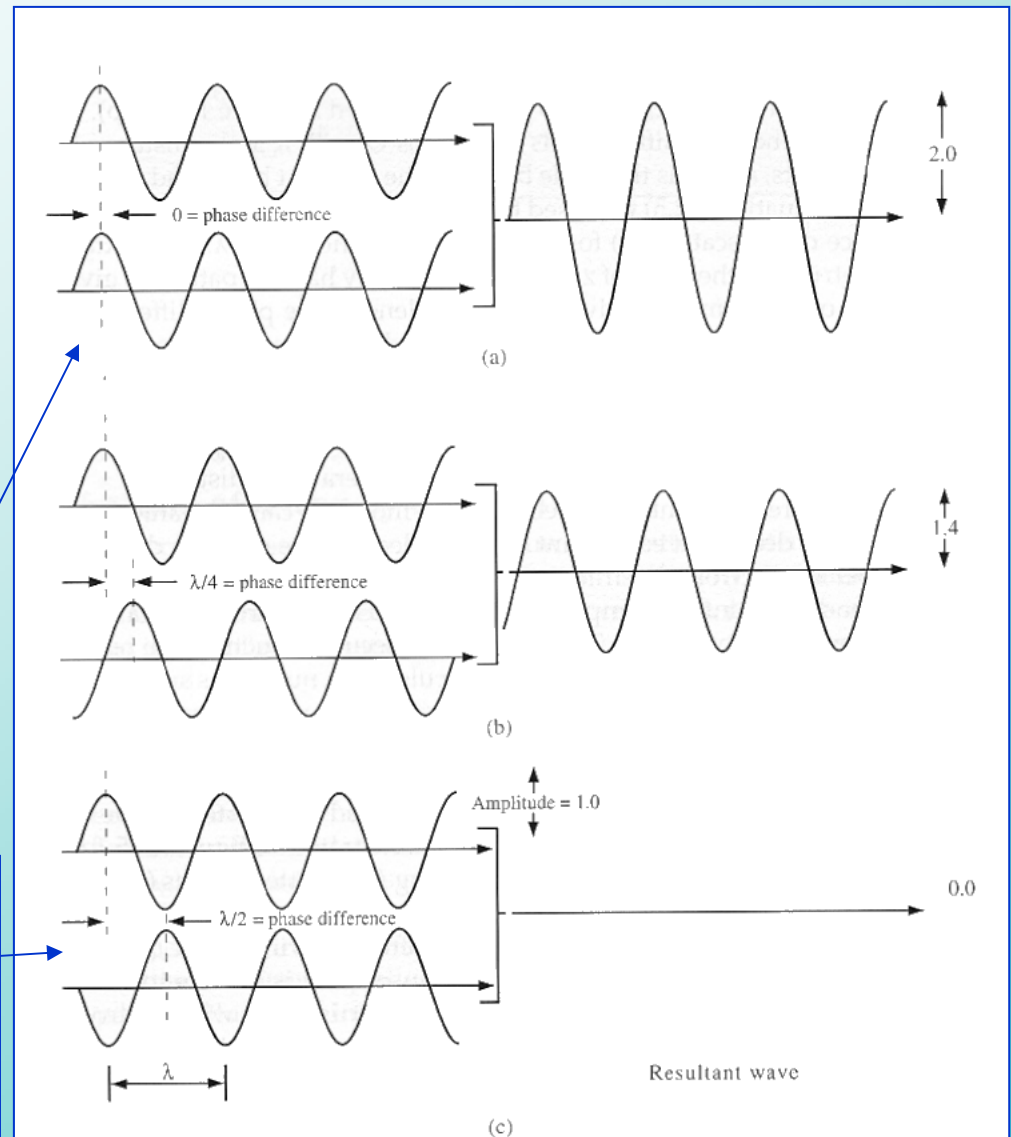
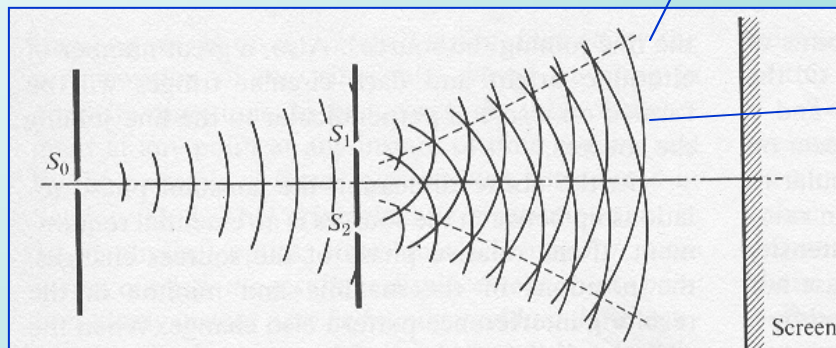


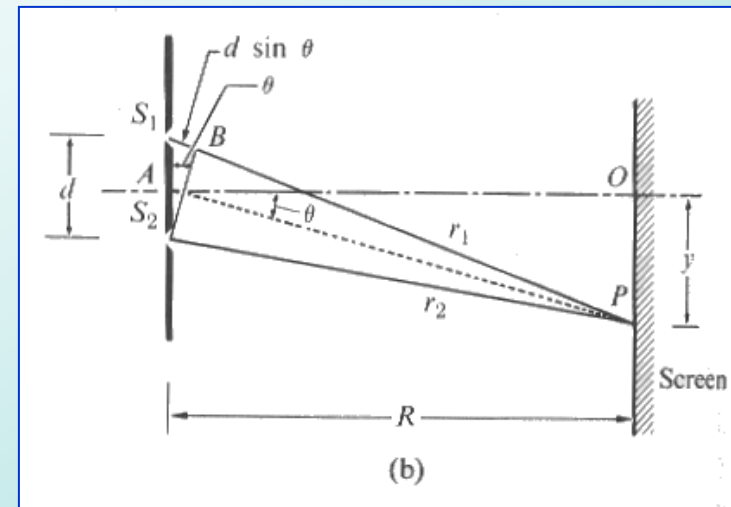
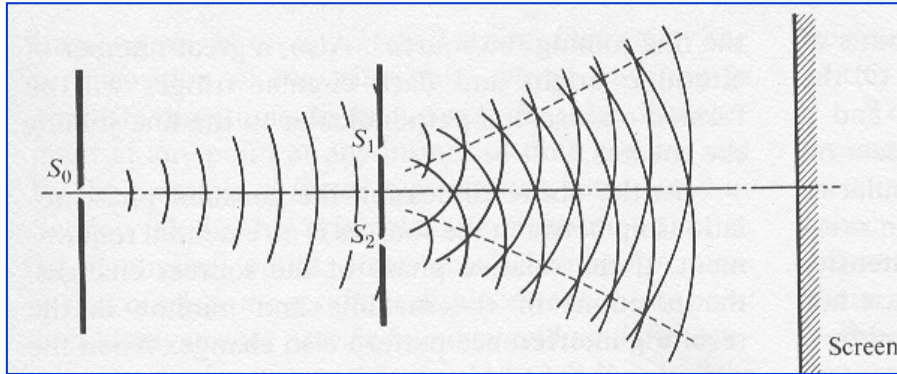
## Interference between rays in and out of phase.

Top: Rays in phase – positive (constructive) interference.

Middle: Rays partly out of phase – weak constructive interference.

Bottom: Rays out of phase – negative (destructive) interference.





The wavelets diverge from each slit as **coherent sources** (same phase, frequency and orientation). The interference fringes occur at those points at which the wavelets from one slit are in phase with the wavelets from the other, - this occurs when the **path difference** of the two rays is **some integral number of wavelengths,  $m$** . The right-hand diagram explains how to find the path difference. If the distance  $d$  between the slits is small compared to the distance away of the screen,  $R$ , the arc  $S_2B$  can be treated as a straight line. Then the angle made by points  $S_1S_2B$  and  $OAP$  are the same value ( $\theta$ ). The line  $AO$  is along the optical axis. The difference in path lengths is then  $d \sin \theta$ , giving:

$$d \sin \theta = m\lambda, \text{ or } d = \frac{m\lambda}{\sin \theta}, \text{ or } \sin \theta = \frac{m\lambda}{d}$$

The *zeroth fringe* lies on the optical axis. The distance  $y$  from this fringe line to the next is given by  $y = R \tan \theta$ , but for small angles,  $\tan \theta \approx \sin \theta$ , so substitution gives us:

$$y = R \frac{\lambda}{d}, \text{ or for the } m\text{th fringe, } y_m = R \frac{m\lambda}{d}, \text{ or } d = R \frac{m\lambda}{y_m}$$

We can obtain information about the distance between the slits from  $R$  and  $y$ , if we know  $\lambda$ .

In X-ray crystallography, we use this principle to find the distances between all the atoms.

The general idea is as follows:

Atoms in a particular plane will align so as to give a diffraction pattern at a particular angle of incidence of the X-ray beam. The reflected beam from each atom acts like the slit in Young's optical diffraction experiment.

Each atom in the selected plane acts as a **coherent point source**, and the wave fronts interfere to provide a diffraction pattern.

Because we have coherent sources in three dimensions, we get an interference pattern consisting of spots instead of lines. However, the method for finding the distances follows the same general idea as that developed in the previous slide.

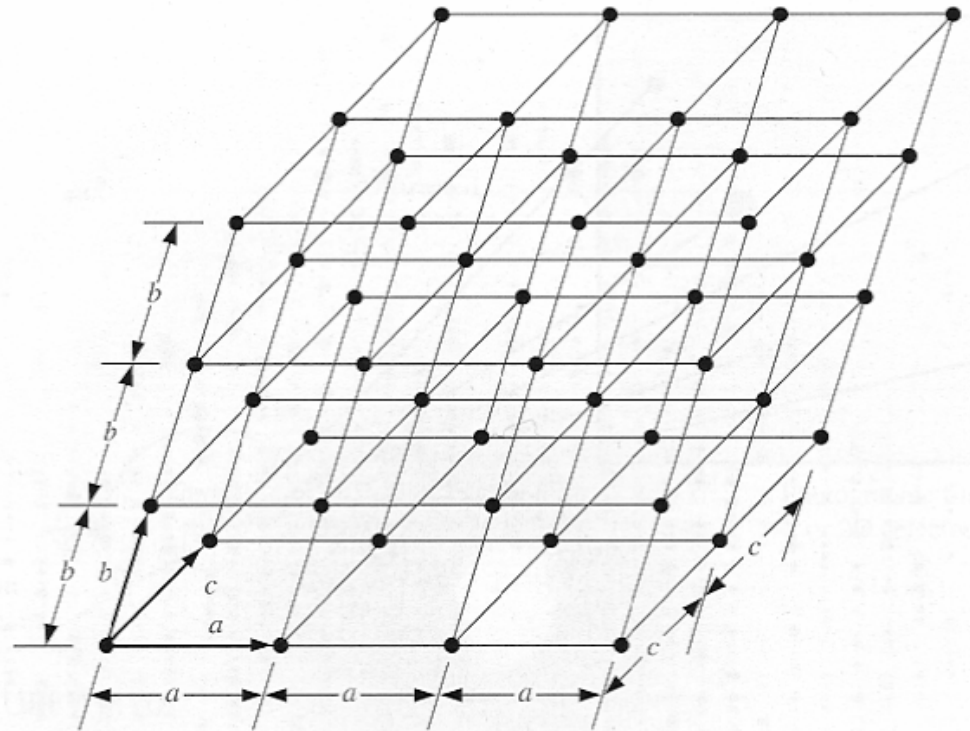
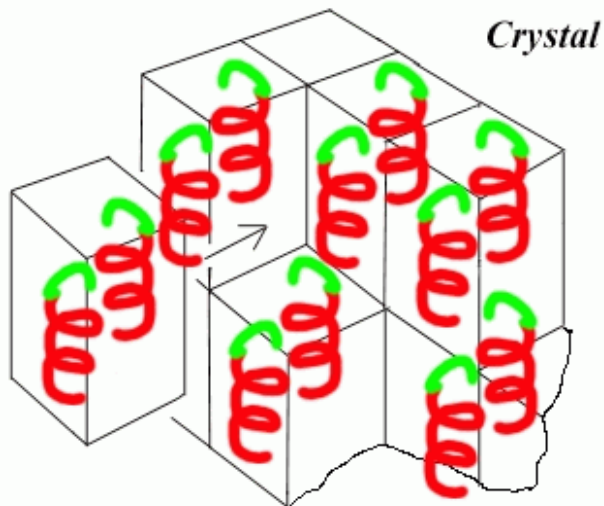
Before looking at the method, we need to understand crystals better.



A more typical crystal lattice, in 3-D.

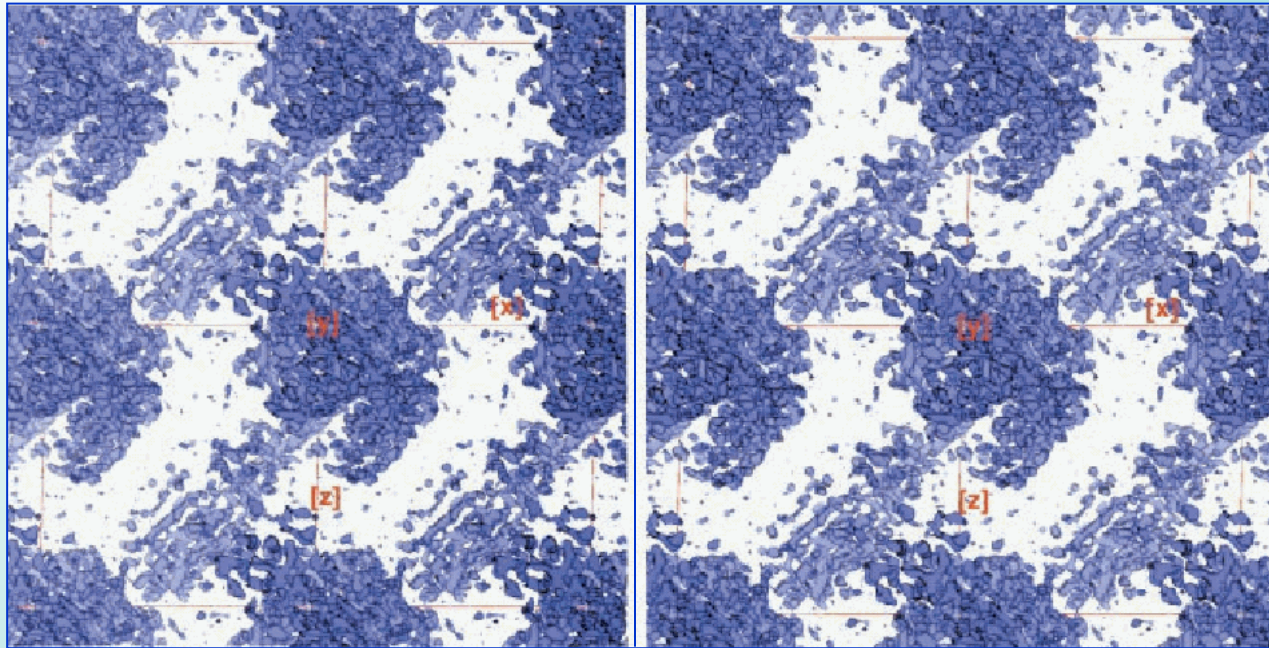
Note that the angles of the unit cell are not necessarily right angles. The dimensions of the unit cell are generally shown as  $a$ ,  $b$ , and  $c$

### Building a unit cell with a protein



How does this relate to diffraction? When X-rays strike a crystal each atom acts to scatter the rays, acting as a mirror. The “reflections” act like the slits of our optical diffraction example, - each is a *coherent source*. In each repeat of the crystal, the distance  $d$  of a particular atom acts like a set of slits, giving a unique diffraction pattern.

Crystal packing in a large membrane protein – the yeast mitochondrial  $F_1F_0$  ATP synthase



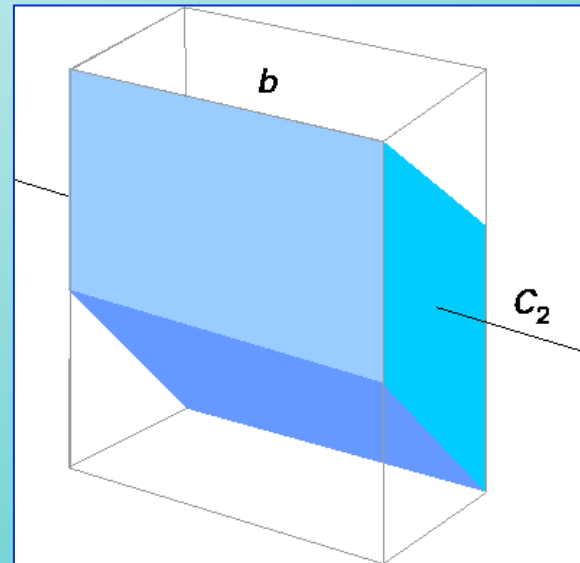
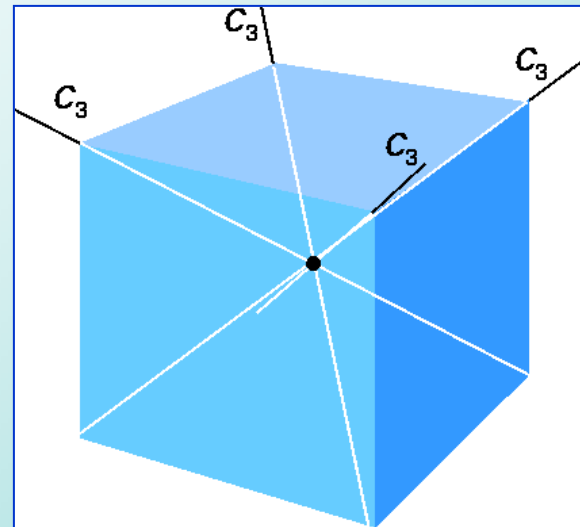
Stereo pair for crossed-eye viewing

Crystals come in a limited number of forms. These can be classified according to the axes of symmetry in each form. For example, the top crystal here has four axes along which it can be rotated about a 3-fold symmetry,  $C_3$ . It has cubic symmetry.

What is three-fold symmetry? It means that, as we rotate the crystal about the full  $360^\circ$  along one of these axes, there are three points (every  $120^\circ$ ) at which it looks the same.

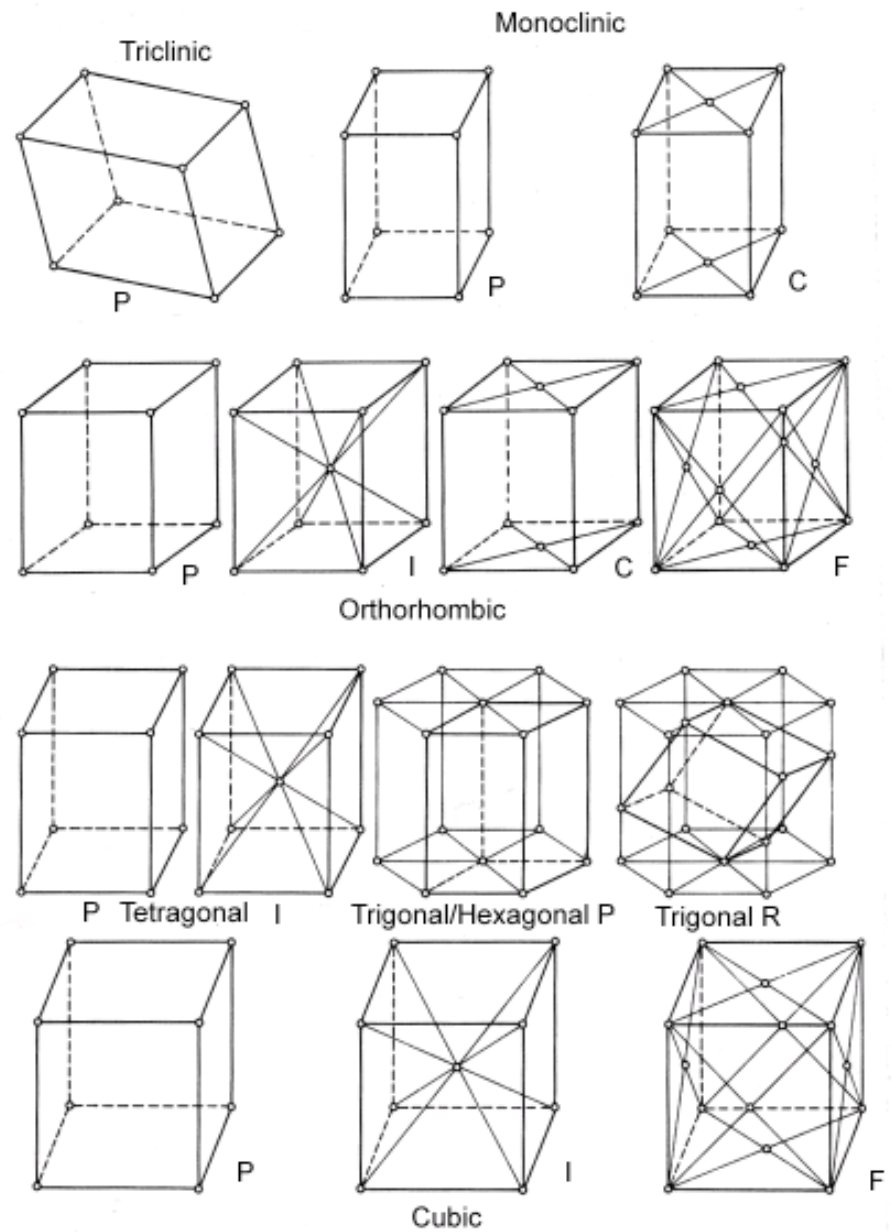
The bottom crystal is monoclinic, and has one axis of two-fold symmetry,  $C_2$ .

The crystal form reflects the dimensions of the unit cell.



Unique crystal forms – the fourteen Bravais lattices. The letter *P* denotes a primitive unit cell; *I* a body-centered unit cell; *F* a face-centered unit cell; and *C* (or *A* or *B*) a unit cell with lattice points on two opposite faces.

The shaded areas are the unit cell. Where the angles are not 90°, the dotted lines show the equivalent faced right-angled volume.



**Table 15.3** The essential symmetries of the seven crystal systems

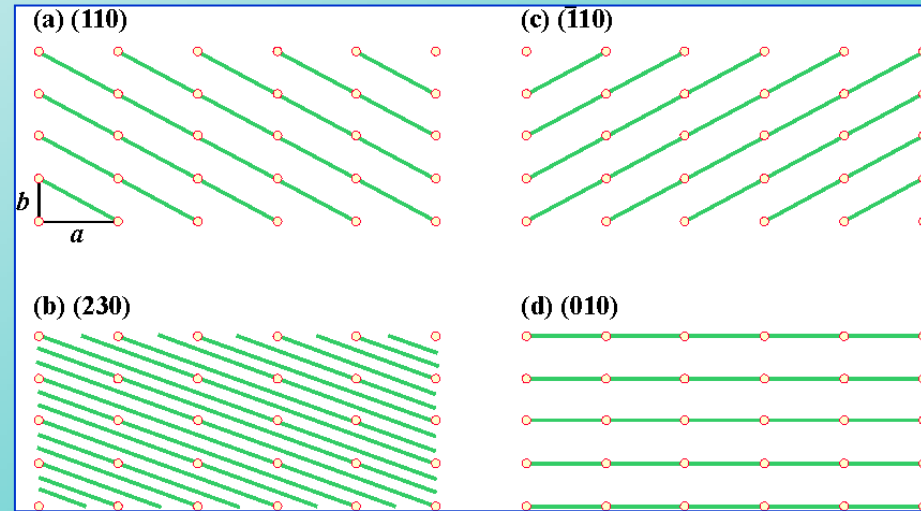
The systems	Essential symmetries
Triclinic	None
Monoclinic	One twofold axis
Orthorhombic	Three perpendicular twofold axes
Rhombohedral	One threefold axis
Tetragonal	One fourfold axis
Hexagonal	One sixfold axis
Cubic	Four threefold axes in a tetrahedral arrangement

# Planes in the crystal

For any repeating two dimensional lattice, we can draw lines through the points which have various slopes. In a 3-D lattice, these lines are projected in the third dimension to form planes. For a molecular crystal, **each atom is in a regular repeating pattern**, so we could draw planes through the sub-lattice for each atom. The distance between the planes for different atoms would then tell us the structure of the lattice, and hence the structure of our molecule.

The general principle of crystallography is to determine the structure by looking at these repeat distances.

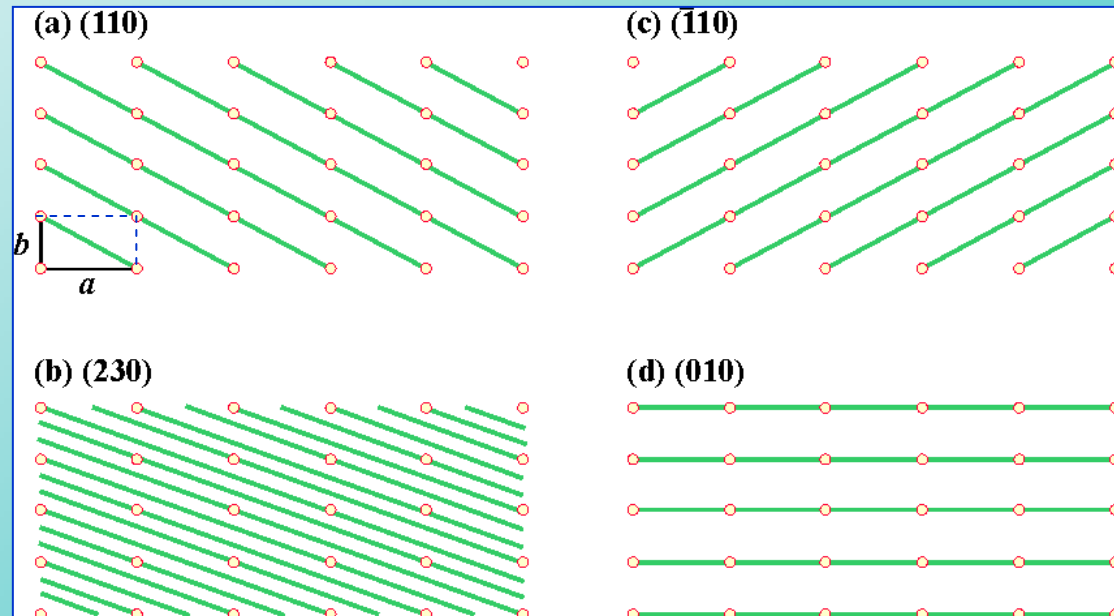
This is achieved by taking advantage of the diffraction of X-rays (or accelerated electrons or neutrons), all of which have wavelengths in the range  $\sim 1 \text{ \AA}$  appropriate for molecular bond distances. X-rays are “reflected” off the electron clouds around atoms and bonds, and form diffraction patterns that depend on angle of the X-rays with the planes, and on a coincidence between the repeat distance and the wavelength of the X-rays.



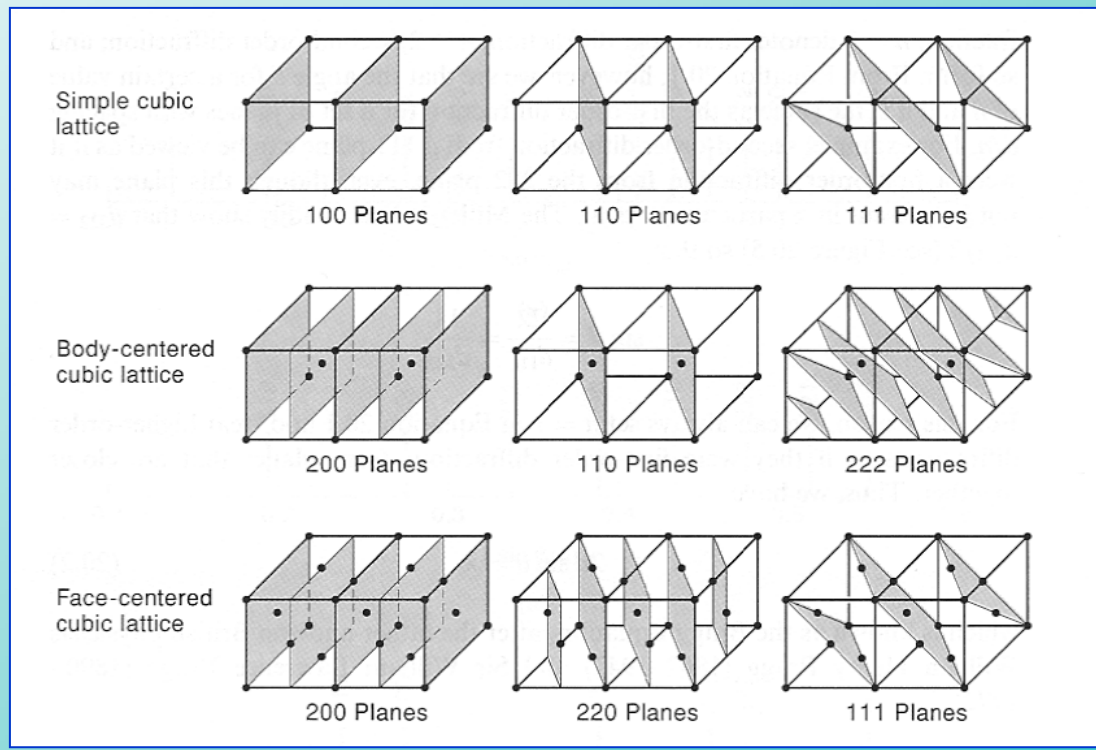
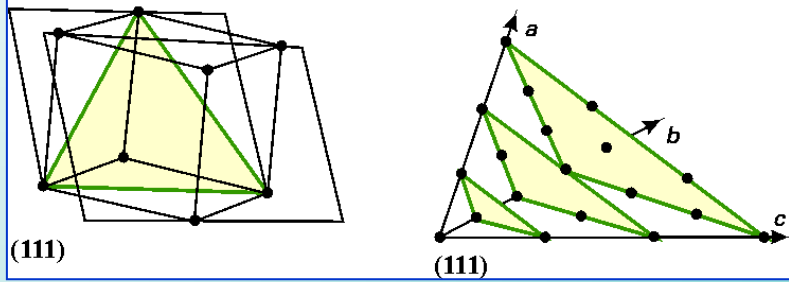
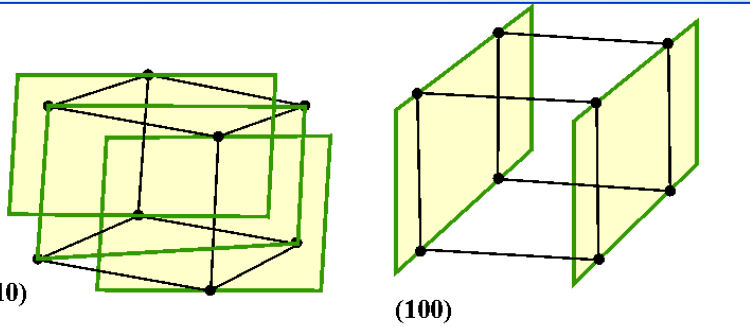
## Planes and cells – Miller indices

In a 2-D lattice, the planes are represented by parallel lines connecting the points. The slope of the line is given by the number of units in the  $a$  and  $b$  axes, where  $a$  and  $b$  define the unit cell. In (a) below, we have a slope defined by  $1a, 1b$ ; in (b), the slope is  $3a, 2b$ ; in (c),  $-1a, 1b$ . The description of lines in (d) are problematic, because they never intercept the  $a$  axis, since they are parallel to it. Similarly, if the lines represent planes, with the third dimension perpendicular to the  $ab$  plane, then all the planes would be parallel to the  $c$  axis. In principle, we could define the planes in terms of the intercepts, so (a) would be  $1, 1, \infty$ ; (b) would be  $3, 2, \infty$ ; etc.

In order to avoid the infinity term, it is more convenient to take reciprocals of these values. In order to avoid fractional numbers, the result is multiplied through to give the lowest integer set. What we get then are the **Miller indices** for the planes. These are named the  $h$ ,  $k$ , and  $l$  values, and are shown without commas, as in the Fig.



Examples of planes through different crystal lattices. The planes on the right are for cubic lattices. Some of these on the right are not right angled. The planes are shaded.



## Relation between Miller indices and unit cell

In a diffraction experiment, the Miller indices and unit cell dimensions can be read directly from the diffraction pattern. It is therefore important to know the relationship between the indices and the distance between planes.

$$\sin \phi = \frac{d}{(a/h)} = \frac{hd}{a}$$

$$\cos \phi = \frac{d}{(b/k)} = \frac{kd}{b}$$

Since  $\sin^2 \phi + \cos^2 \phi = 1$

$$\frac{h^2 d^2}{a^2} + \frac{k^2 d^2}{b^2} = 1$$

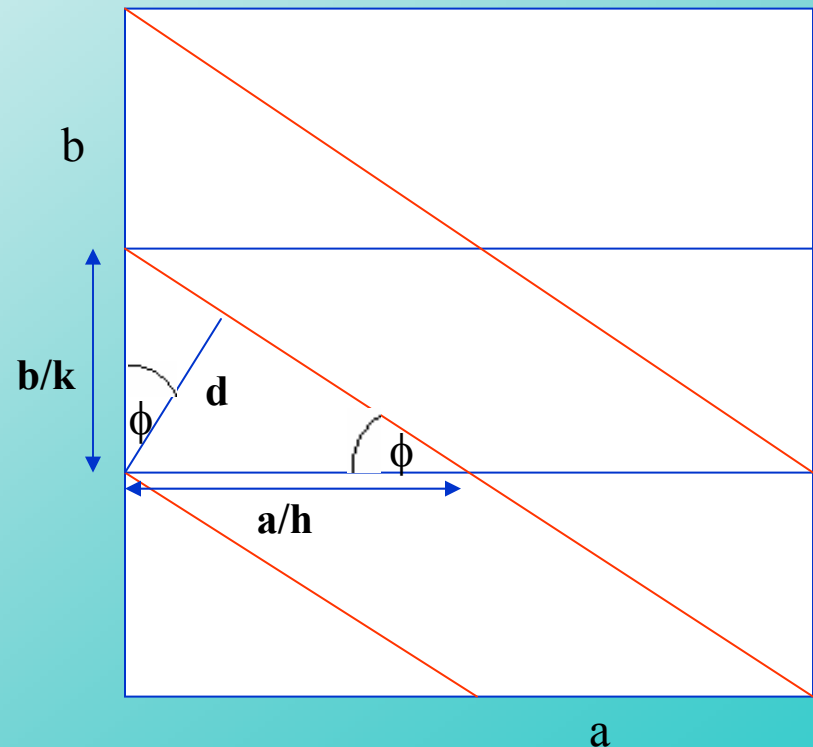
or, on rearrangement

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2}$$

In three dimensions, this becomes

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}$$

In diagram below,  $a$  and  $b$  are dimensions of the unit cell, and the red lines are planes through the lattice, spaced by distance  $d$ .





# Derivation of Bragg's Law

The angle of incidence (the glancing angle),  $\theta$ , and the condition for positive interference are related through the geometry shown in the schemes below.

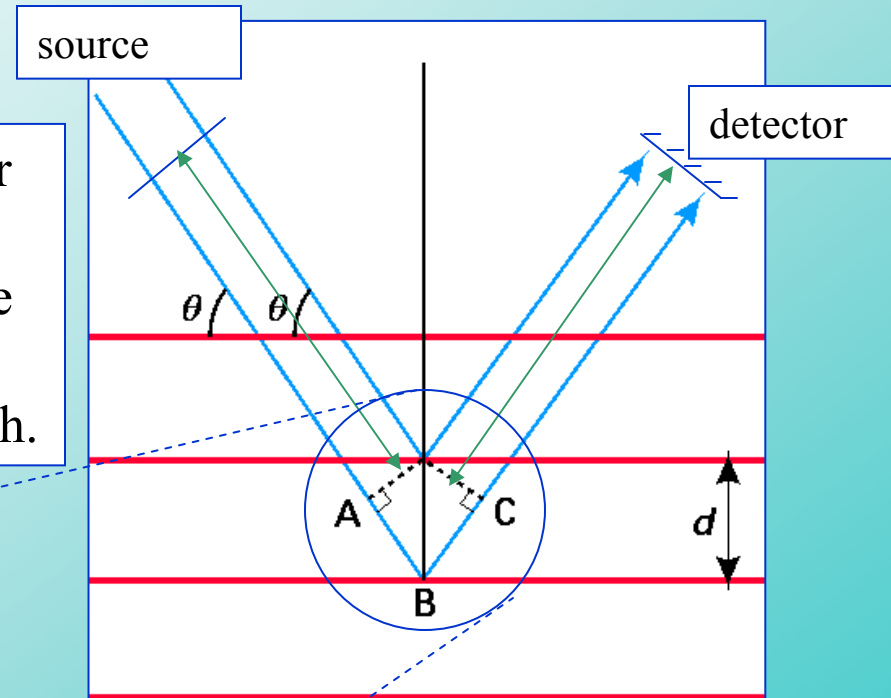
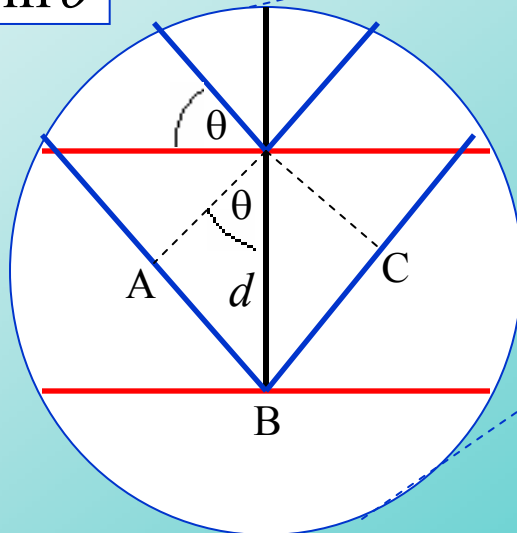
$$\sin \theta = \frac{AB}{d} = \frac{BC}{d}$$

$$AB + BC = 2d \sin \theta$$

$AB + BC$  gives the path length difference for the two rays shown, which are separated by the unit cell dimension. Positive interference occurs when the path length difference is equal to some integer multiple of wavelength.

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

The positive interference gives rise to the diffraction pattern from which the structure can be calculated.



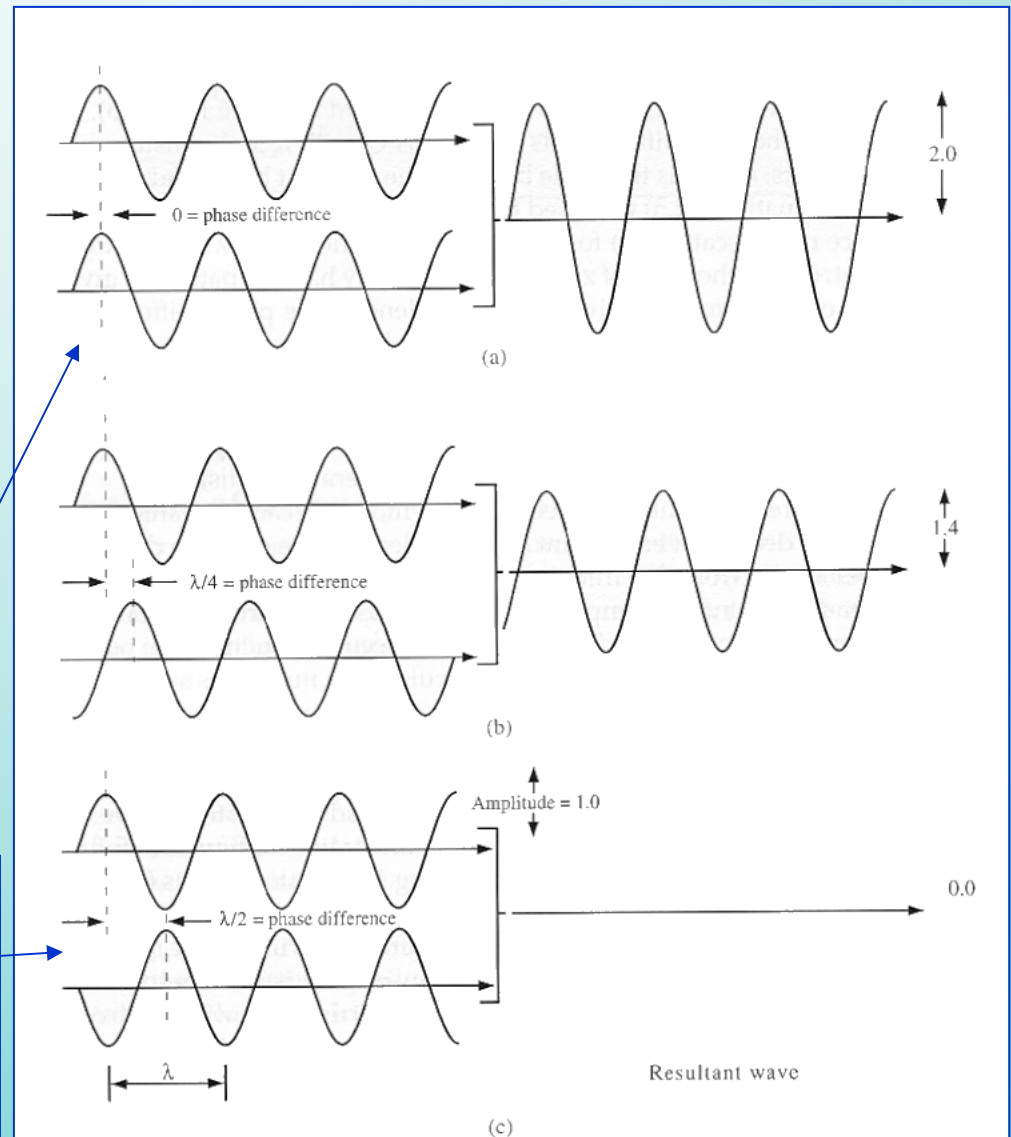
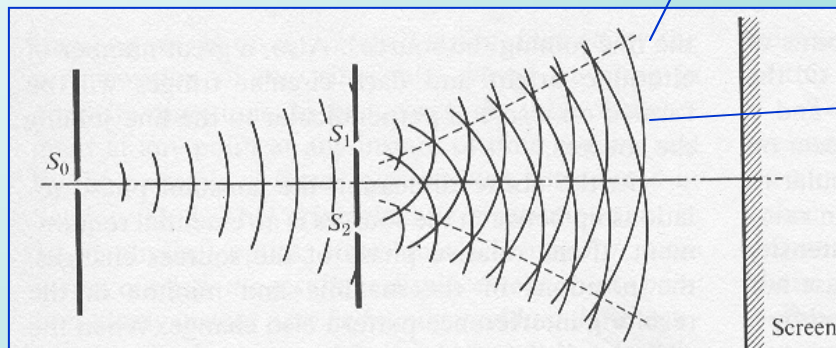
[Interactive Bragg's Law](http://www.journey.sunysb.edu/ProjectJava/Bragg/home.html)

## Interference between rays in and out of phase.

Top: Rays in phase – positive (constructive) interference.

Middle: Rays partly out of phase – weak constructive interference.

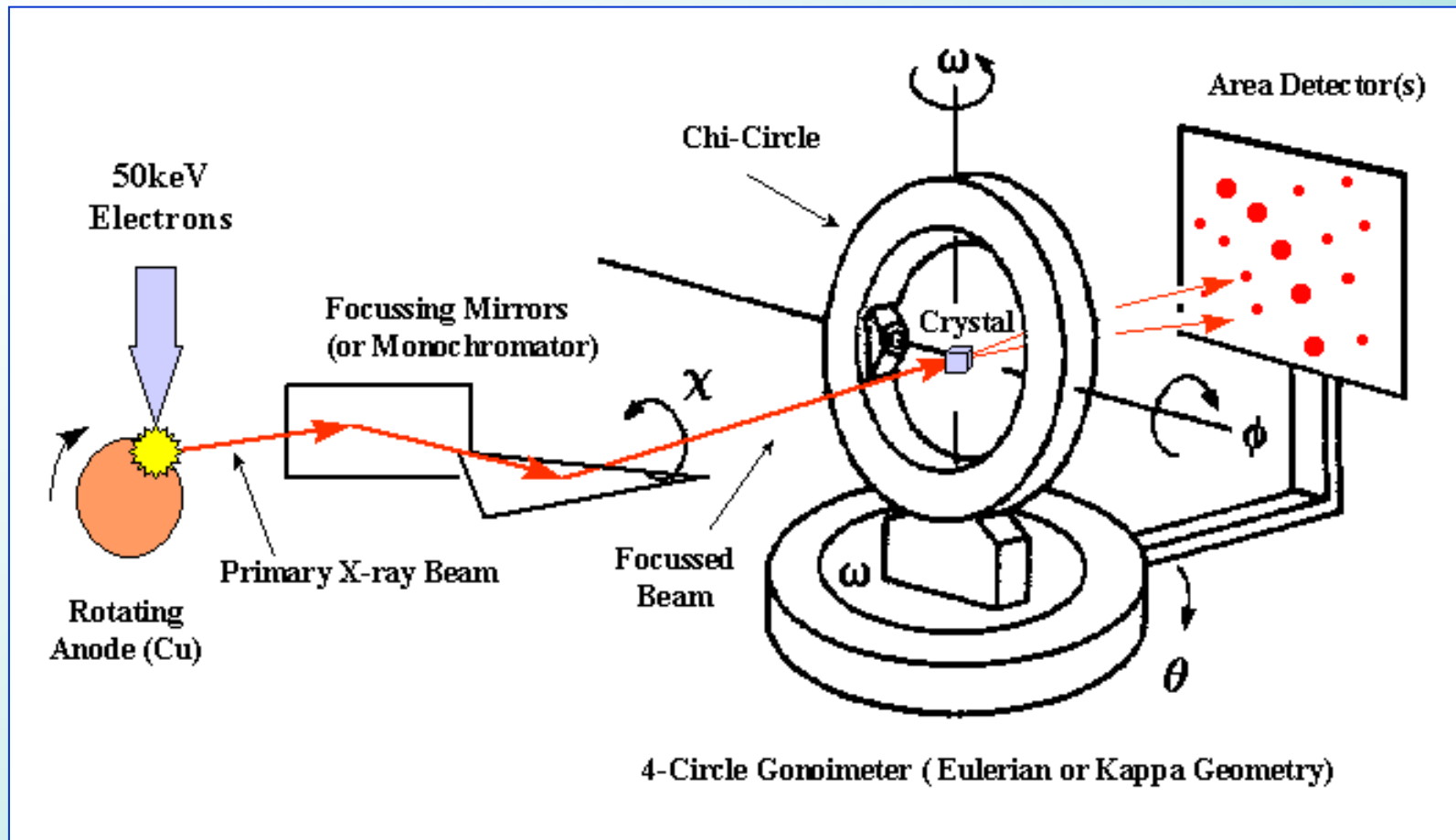
Bottom: Rays out of phase – negative (destructive) interference.



## Take home messages so far

1. Need to grow crystals in order to use X-ray crystallography.
2. The crystal is made up from repeat units, each of which is a unit cell.
3. The form of the crystal is determined by symmetry. There are 14 types of crystal (the Bravais lattices), falling into seven different crystal systems.
4. Diffraction occurs when X-rays reflected from different planes of the crystal lattice interfere.
5. The relation between the planes and the unit cell is determined by the Miller indices.
6. Diffraction spots occur when the distance between planes is such that the difference in path length between two rays is an integer multiple of the wavelength.
7. The difference in path length is dependent on the glancing angle and the separation distance between the planes, as defined by Bragg's Law.
8. The unit cell dimensions, and the Miller indices of each plane, can be read directly from the diffraction pattern.

## Step 3 – the diffractometer

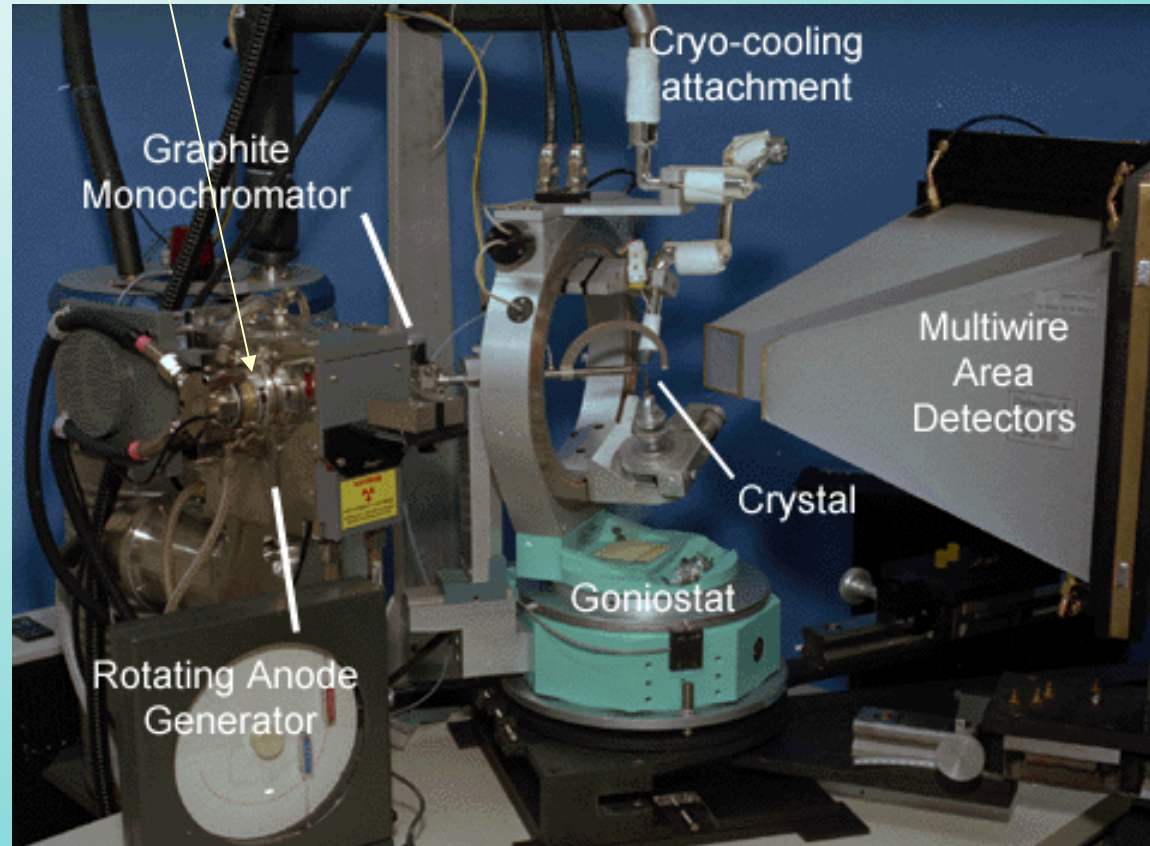
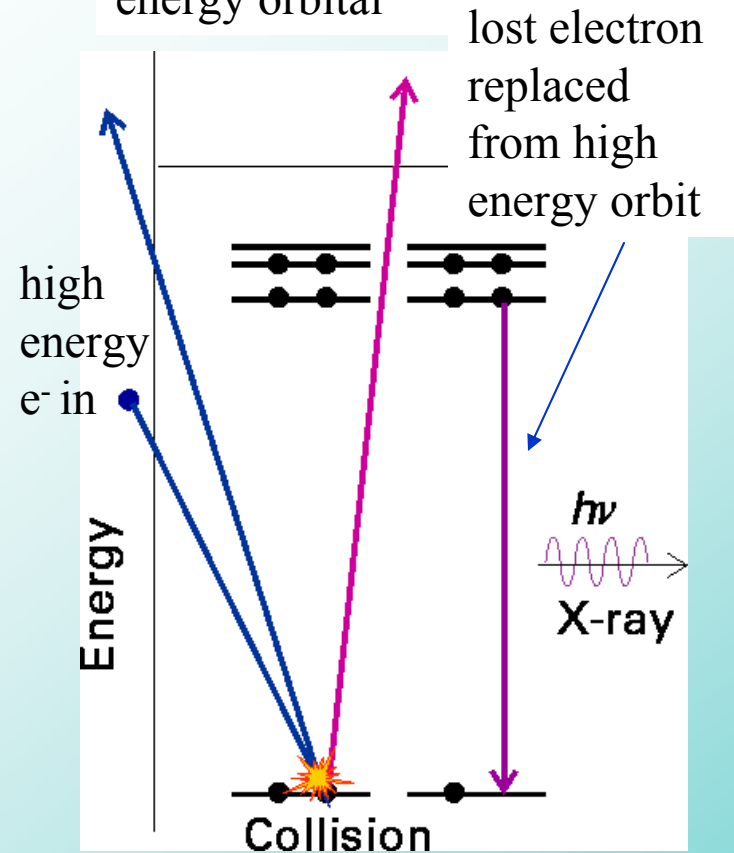


The diagram shows the main elements of a diffractometer. We will look at some of these in greater detail.

How are X-rays generated? On a laboratory scale, X-rays can be generated by bombarding a metal surface with high-energy electrons. The diagram on the left shows the principle. The photo shows a rotating anode X-ray generator. Rotation of the anode (a Cu metal surface) allows continuous operation without depletion.

$e^-$  ejected from low energy orbital

X-ray source



## The Advanced Photon Source (APS) synchrotron ring at Argonne.



Electrons are accelerated in the synchrotron ring, and then force to turn corners, through use of undulators and wigglers (extra magnets). As they do so, they emit photons to get rid of the excess energy. The APS is designed to optimize X-ray generation.

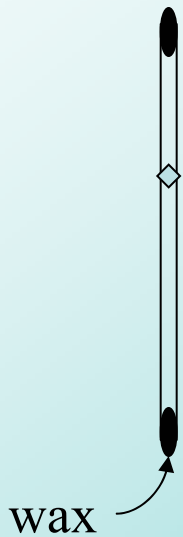
A diffractometry station at a modern synchrotron facility. The detector shown contains a large CCD array, which records the diffraction pattern produced on a phosphor screen as the X-rays excite it.



# Mounting

If not freezing

mount in thin-walled glass capillary tube



If freezing

mount on a film of solvent in a wire loop fixed to a metal capillary\*



\*freeze immediately!

Attach capillary to goniometer head



eucentric goniometer head  
(Nonius)