BCHS 6229
Protein Structure and Function

Lecture 8 (Nov 3, 2011)

Methods for Protein Structure-Function Studies (I)

X-ray Crystallography (I)

X-ray crystallography

Exciting time for structural studies by X-ray crystallography: molecular cloning, recombinant protein expression, synchrotron radiation, and software advancement have revolutionized structure determination - a much faster and often ‘routine process’.

Recommended Books on Crystallography:

- Crystallography made crystal clear, by Gale Rhodes (2nd *Ed.,* 2000).
- Principles of Protein X-ray Crystallography, by Jan Drenth (Basics, lots of equations).
- Outline of Crystallography for Biologists, by David Blow.
- Crystal Structure Analysis, Jenny Glusker and Kenneth Trueblood, 2nd *Ed*, 1985) (Good description of diffraction theory and good glossary).
- Protein Crystallography, by Blundell and Johnson (Academic Press, 1976) (For people who want in depth coverage. Most complete, but may be out of date for some sections).

Introduction

1. Why X-ray crystallography?

2. Crystallography as a tool to determine 3-D structures.
   The most common technique to determine 3-D structures.
   Crystal structures and NMR structures are equivalent.
   X-ray crystal structures are compatible with other structural evidence
   X-ray structures give structural support to biological data.

3. What is a crystal?
   It is an orderly 3-D array of molecules, held together by non-covalent interaction.
   It consists of unit cells arranged in a lattice.
   Crystals are ~50% water (Proteins are still active in crystals).

4. Proteins form crystals: a miracle
   How to grow protein crystals?

5. Protein crystals diffract X-rays: the diffraction pattern

6. Electron density: the end product of X-ray crystallography

X-ray Crystallography and NMR

	X-ray Crystallography	NMR
Data Collection	Fast	Usually slower, requiring a lot more time on data collection and structure determination
Target Molecules	No theoretical size limit	Limited by size
Accuracy	Accurate distance information: ideal for studying macromolecular interactions and enzyme catalysis	Not as accurate
Molecule state	Sometimes regions of the molecule may be influenced by crystal packing	Solution state
Pre-requisite	Need well-diffracting crystals	No need to grow crystals
Outcome	Static picture	Dynamic information such as time scales of movement under “native” condition

A 8-step program for protein structure determination by x-ray crystallography

1. Produce (monodisperse) protein either alone or as relevant complexes
2. Grow and characterize crystals
3. Collect X-ray diffraction data
4. Solve the phase problem either experimentally or computationally
5. Build an atomic model using the electron density map
6. Iteratively refine and rebuild the structural model
7. Analyze and validate (How do you know if a crystal structure is correct?)
8. Develop structure based hypotheses (interpret biological implications)

1. Produce (monodisperse) protein either alone or as relevant complexes

Methods to determine protein purity, heterogeneity, and monodispersity

SDS-PAGE

Size exclusion chromatography

Dynamic light scattering

CD spectroscopy www.ruppweb.org/cd/cdtutorial.htm

Other biophysical methods

2. Grow and characterize crystals

A. Overview of X-ray crystallography

Why crystallography?

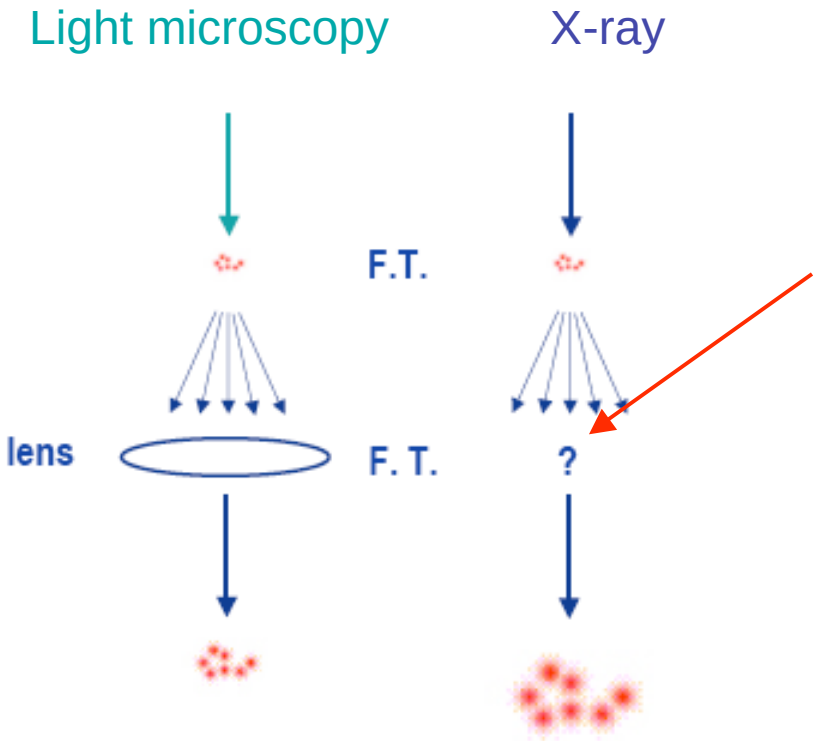
Ideally: Microscope

Two Problems: lens, single molecule

X-ray crystallography may be considered a form of microscopy:

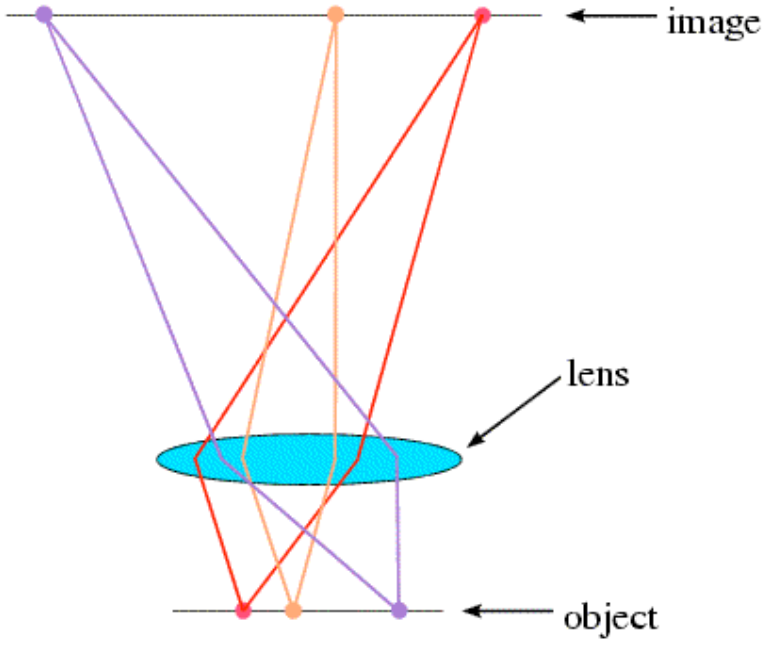
The Fourier transform (F.T.) of the F.T. of an object is itself.

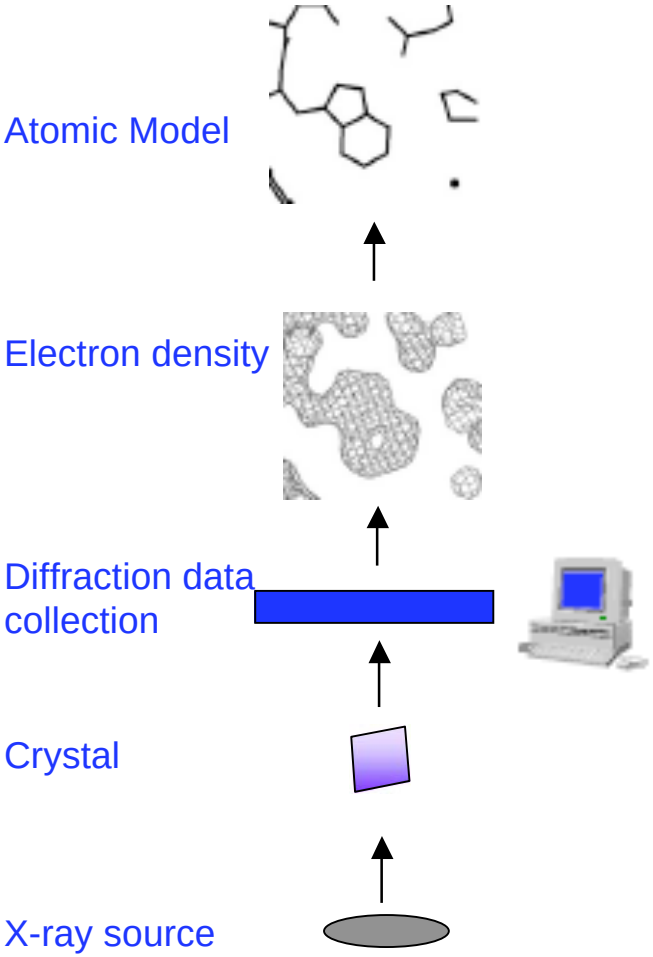
The F.T. of an object is its diffraction or scattering; The F.T. of its diffraction is the object.



"X-ray microscopy" is impossible, since there is no X-ray lens that can recombine the scattered rays to generate an enlarged image (x-ray cannot be focused by lens)

Wavelength:	400-700 nm	~1Å
Resolution:	up to fractions of a μm	~Å; "atomic" resolution





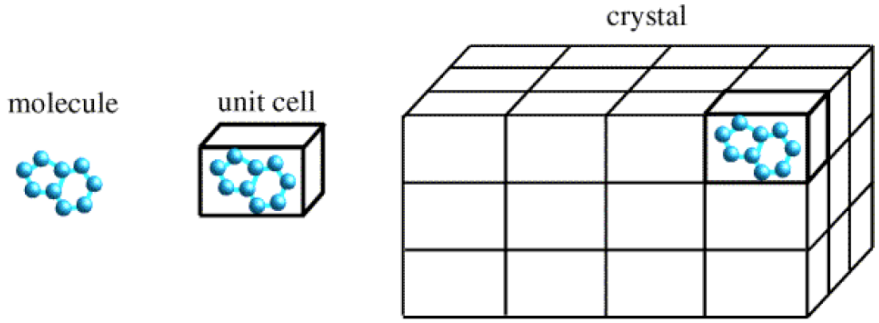
In X-ray crystallography, “Computer” simulates “lens”

Why do we need crystals for X-ray crystallography?

Diffraction from a single molecule is too weak to measure

Crystals INTENSIFY and SIMPLIFY a diffraction pattern:
Selective constructive interference or sampling
at ‘reflection’ points or lattice points.

A crystal of 100μm x 100μm x 100μm could contain 10^{12} identical “unit cells” or molecules of 100Å x 100Å x 100Å.



A Crystal is an Amplifier of X-ray scattering

X-ray diffraction from Crystals

- Diffracted x-rays are emitted from collisions with electrons.

- The molecules within a crystal are aligned so that their (continuous) diffraction patterns are in phase only at discrete positions that depend on the internal dimensions and symmetry of the crystals.

- The resulting x-ray diffraction pattern is recorded on a 2D detector as discrete data points known as "reflections."We can consider the diffraction pattern to arise from x-rays "reflecting"off of discrete planes (Bragg planes) within the crystal. These planes consisting of equivalent atoms of the structure aligned with the incident x-ray beam.

- The positions of x-ray reflections depend upon crystal parameters and the wavelength of the incident x-rays. The relative intensities of the reflections contain information about the structure of the molecule.

Electrons scatter X-rays:

- Therefore, what X-ray diffraction could give is “electron density” of the molecule.
- Scattering intensities are roughly proportional to the square of number of electrons.
- Hydrogen atoms are usually invisible in an electron density map.
- “Heavy” atoms, referring to those with a lot more electrons than C, N and O atoms that are the major components of macromolecules, scatter X-rays much more strongly.

Fourier Transforms and the Phase Problem

- The x-ray diffraction pattern is related to the scattering object by a mathematical operation known as a **Fourier Transform**.

$$\mathbf{F}(\mathbf{s}) = \int_{space} \rho(\mathbf{r}) \exp(2\pi i \mathbf{s} \cdot \mathbf{r}) d\mathbf{r}$$

- The objective lens of a light microscope performs the same function as the Fourier transform used in x-ray crystallography.
- Fourier transforms can be inverted–the Fourier transform of the diffraction pattern will reveal the structure of the scattering object.

- The problem with phases:We need to know the amplitudes of the diffracted x-rays as well as their relative phases to compute the Fourier transform. The x-ray experiment measures only the amplitudes (reflection intensities are the square of the Fourier amplitudes).

The Interpretation of X-ray Diffraction Data

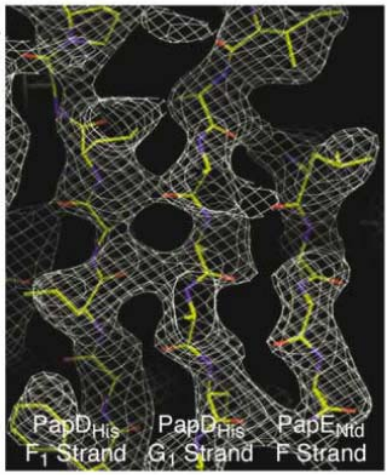
- Our (un)certainty about an x-ray structure is directly related to the quality of the x-ray diffraction data.

- The electron density revealed by the Fourier transform of the diffraction data (actually, the square root of the intensities) has resolution-dependent features.

- Errors in measurement of the reflection intensities and the phases degrade the quality of the “x-ray image”of the molecule.

The Interpretation of X-ray Diffraction Data

A crystallographic model of a protein is built on the basis of the shape of the electron density, the known amino acid sequence, standard chemical constraints/restraints for polypeptides (bond angles and lengths, allowed torsions, etc.), and the agreement between the measured x-ray data and the diffraction pattern calculated from the model (the R-factor).



Electron density superimposed with a molecular model

Judging the Quality of X-ray Structures

X-ray Data Quality

- R_{sym}–the error in measured intensities of equivalent reflections (typically ranging from 3% at low resolution to 35% at the high resolution limit).

- Resolution, signal-to-noise ratio (I/σ > 3-4 for useful data)

Judging the Quality of X-ray Structures

Crystallographic Model quality

- R_{cryst}–the error in agreement between the model and experimental structure factor amplitudes (typically ranging from16% (high resolution structure) to 28% (lower resolution).
- Free R-factor (R_{free}) –a crystallographic R-factor calculated from a small set (5-10%) of reflections that are reserved and not used during model refinement (R_{free} is typically larger (+ 2-4%) than R_{cryst}). Over-refinement causes an artificial decrease in R_{cryst} with little or no change in R_{free}.
- Agreement between the model and known structures.
  - Ramachandran plot.
  - Deviation from standard geometry (bond angles, lengths, etc.).

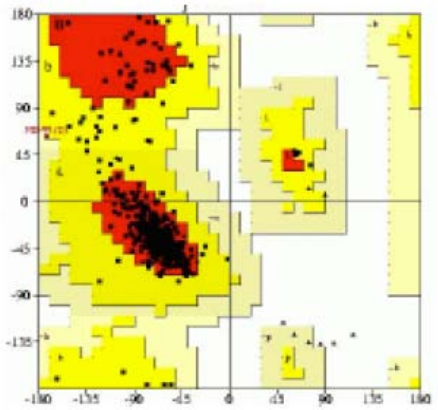


Table 1: a standard for X-ray Publication

Table I Data collection, phasing and refinement statistics

Data set	SeMet-labeled FapR			$FapR_{\Delta 67}$	$FapR_{\Delta 43}$–malonyl-CoA
Data collection					
Resolution (Å)[a]		40–3.5 (3.69–3.5)		60–2.5 (2.64–2.5)	63.2–3.1 (3.27–3.1)
Wavelength (Å)	0.9791	0.9793	0.9755	1.072	0.9794
Measured reflections	22 839	22 837	23 045	74 467	87 334
Multiplicity[a]	5.8 (5.8)	5.8 (5.7)	5.8 (5.8)	5.0 (4.3)	6.9 (7.2)
Completeness (%)[a]	99.4 (99.4)	99.5 (99.7)	99.3 (99.2)	80.5 (50.0)	100 (100)
R_{sym} (%)[a,b]	9.3 (27.5)	9.7 (32.6)	11.5 (40.3)	8.3 (28.7)	7.9 (31.0)
$\langle I/\sigma \rangle$[a]	13.5 (4.9)	12.8 (4.2)	11.3 (3.3)	13.7 (4.8)	19.3 (6.0)
Refinement					
Resolution (Å)				30–2.5	63.2–3.1
R_{cryst}[c] (No. refs)				0.221 (14035)	0.187 (11604)
R_{free}[c] (No. refs)				0.267 (790)	0.227 (941)
R.m.s. bonds (Å)				0.019	0.02
R.m.s. angles (deg)				1.72	2.14
Protein atoms				2964	2238
Water molecules				8	10
Ligand atoms				—	64

[a]Values in parentheses apply to the high resolution shell.
[b]$R_{sym} = \sum_{hkl} \sum_i |I(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i |I(hkl)|$.
[c]$R = \sum_{hkl} |F(h)obs - F(h)calc| / \sum_{hkl} |F(h)obs|$.
R_{cryst} and R_{free} were calculated from the working and test reflection sets, respectively.

Determining phases becomes the most important part of macromolecular crystallography.

Solving a structure - obtain the phase!

B. Crystallization: theory and practice

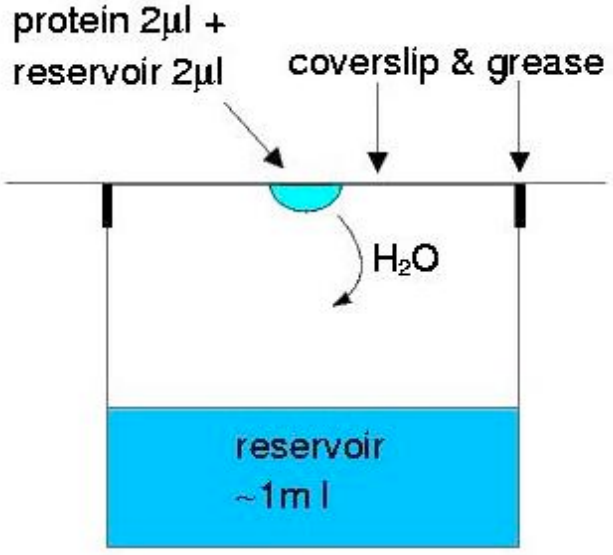
1.Theory
2. Protein
3. Screening
4. Observation
5. Optimization
6. Shooting

http://www-structmed.cimr.cam.ac.uk/Course/Crystals/intro.html

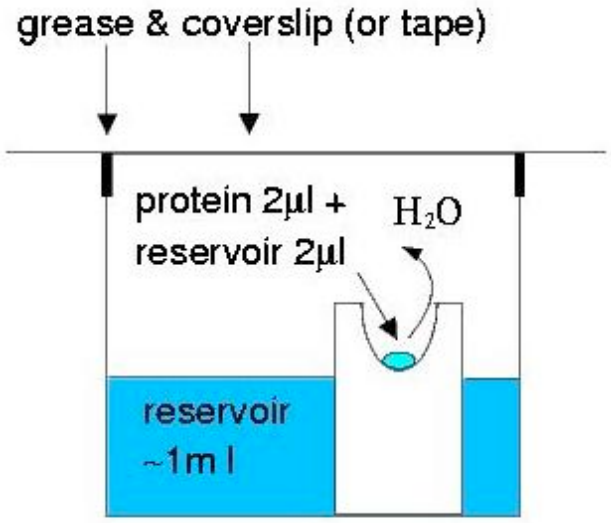
1.Theory

1. Methods for protein crystallization

Vapor diffusion experiments



Hanging drop



Sitting drop

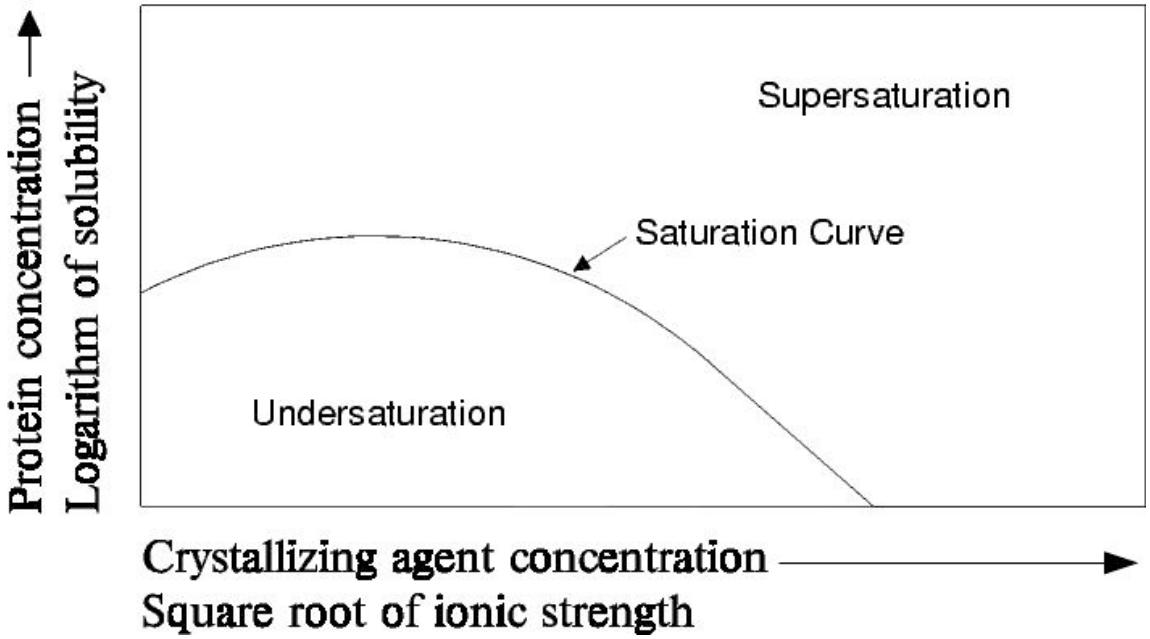
Batch Experiments

Dialysis

1.Theory

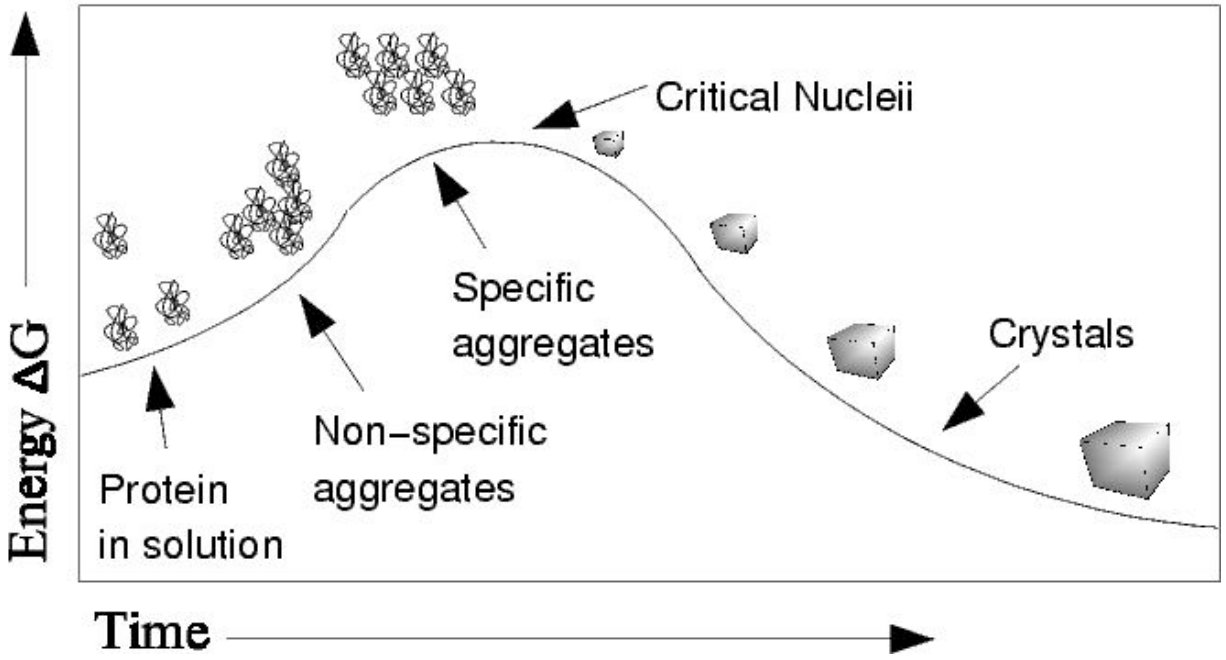
2. Phase diagrams

The solubility of proteins can be represented in phase diagrams



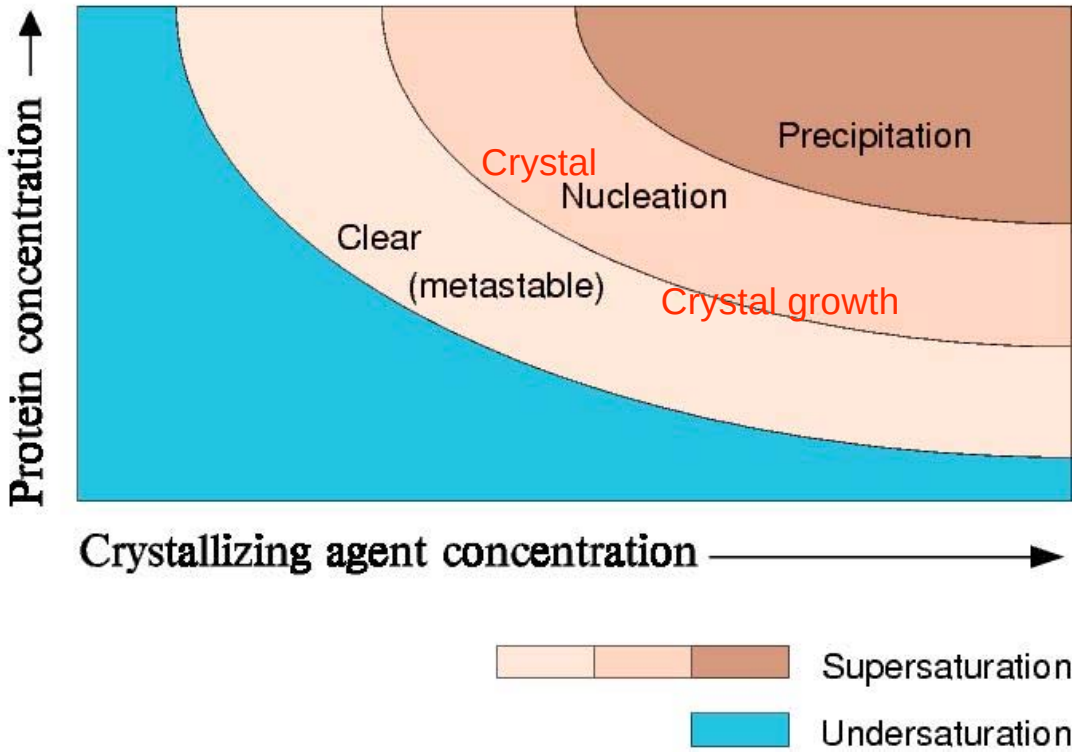
Phase Diagram for a typical protein

There is an energy barrier to crystallization



Energy Diagram for Crystallization

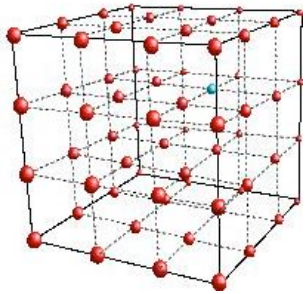
The probability of nucleation increases with increasing supersaturation



Phase Diagram showing zones for crystal nucleation, growth and precipitation.

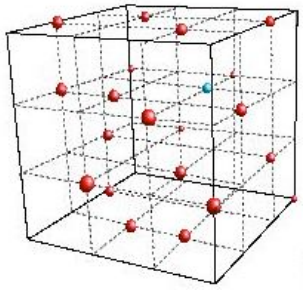
1.Theory

3. Screening: searching phase space

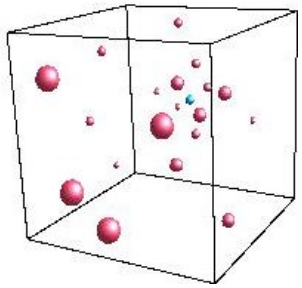


Full factorial

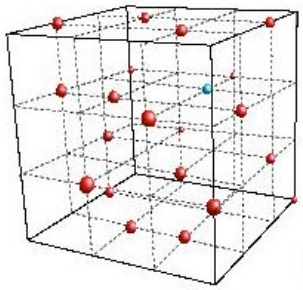
all elements of the matrix of parameters are sampled



Incomplete factorial



Random



Sparse matrix

an intentional bias towards combinations of conditions that have worked previously

It may be possible that the best packed regular or crystalline arrangement is more open (or, more correctly, has a higher free energy) than irregular arrangements. In this case the protein would never crystallize.

Blundell & Johnson, Protein Crystallography (1976)

2. Protein

Is it pure?

Is it folded?

Is it fresh?

Is it monodispersed?

Does your protein need to be kept reduced?
Does your protein need the addition of something (e.g. salt) to stay in solution?
Is your protein stable at room temperature?
Does your protein break down rapidly?
Has anything similar been crystallized before?

And now for something completely different...

Ligand-protein complex

Different species

Deglycosylation

Know your protein
Crystallographer's psychiatrist

3. Screening

Commercially Available Screens

Hampton research (Sparse matrix sampling)

Crystal Screen I, Crystal Screen II

Grid screens for popular precipitants: $(NH_4)_2SO_4$, PEG6K, PEG/LiCl, NaCl, PEG4K/Ion

Molecular dimensions

Emerald BioStructures

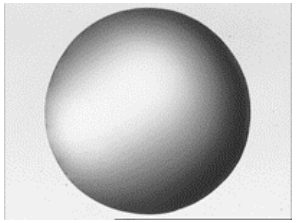
Jena BioScience

Let some else do the experiment

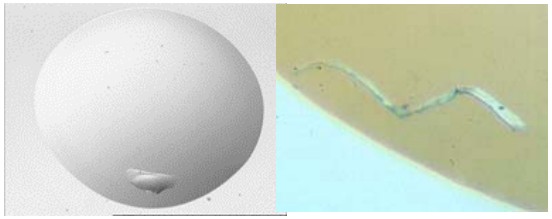
http://www.decode.com/Services/Structural-Biology.php

4. Observation

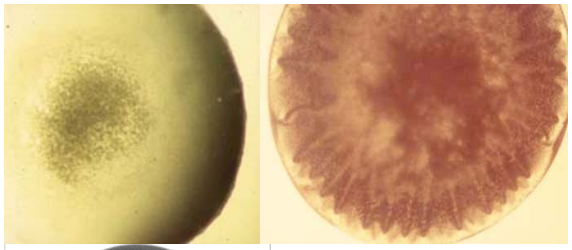
Clear drop



Matter

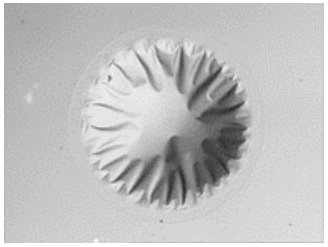


Precipitate



Gels

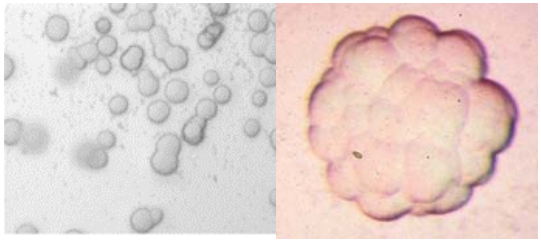
Skin



Phase Separation

Oils

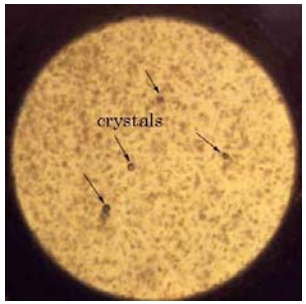
Spherulites

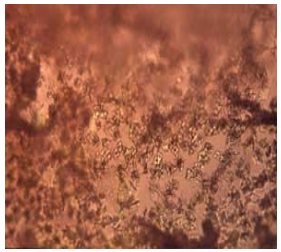


Crystals

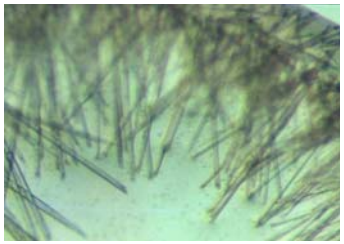
Crystals

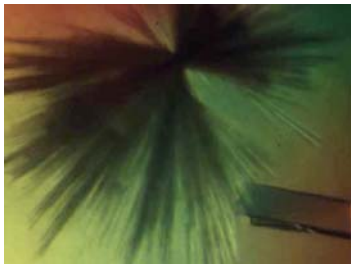
Microcrystalline precipitate

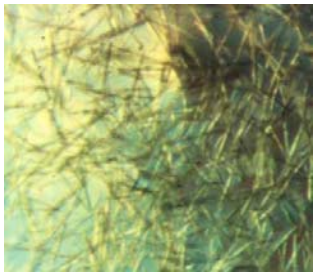




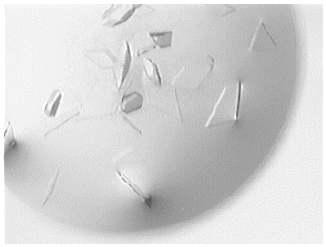
1D - needles



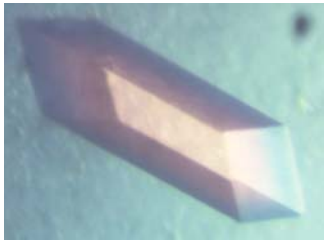


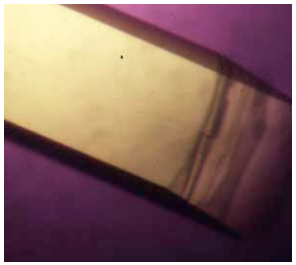


2D - plates



3D - boulders





If you see a crystal…

Is it salt or is it protein?

If you see a crystal a year or two later…

If you see a crystal, don't go running down the corridor screaming "EUREKA!!!" until...

you know your crystal isn't salt...

and that it diffracts.

5. Optimization

Screening around the conditions

Precipitate concentration
pH
Protein concentration
Temperature
Method
Protein

Additives

glycerol (1-25%), ethanol, dioxan, divalent cations, a detergent

Oil in the reservoir

Seeding

6. Shooting

Diffraction

size

order

The first shot

mounting crystals

Poor Crystals

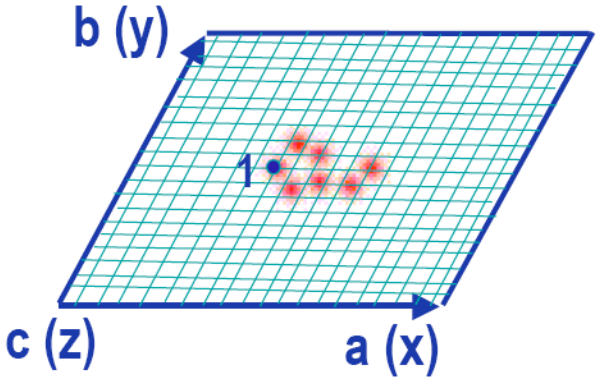
Crystals for cryo-crystallography

C. Basic concepts: reciprocal lattice, symmetry

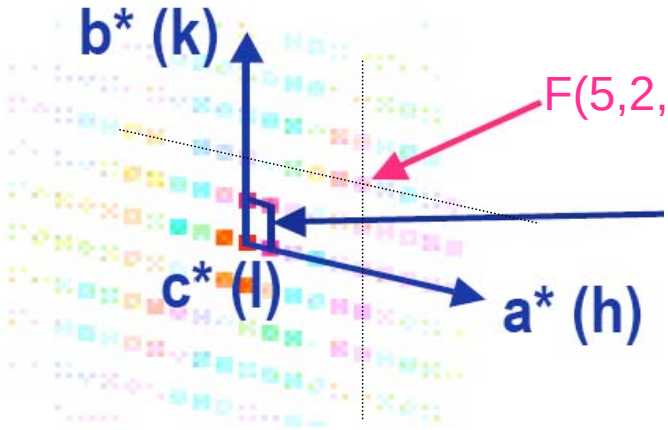
Before we can talk about how to solve the phase problem, we need to talk about symmetry in reciprocal space.

Symmetry in the crystal (real space) influences the symmetry in reciprocal space. A twofold in real space causes a twofold in reciprocal space, a mirror, causes a mirror, a fourfold a fourfold, etc.

In real space, the unit cell content is represented by
electron density $\rho(x,y,z)$ (x=0-1; y=0-1; z=0-1) at a 3-D grid;
e.g. electron density at $\rho(0.35,0.5,0.0) = 6.0$.
or by atomic positions, (x_i,y_i,z_i), i=1,7;
e.g. atom 1 is at $x_1=0.35$, $y_1=0.5$, $z_1=0.0$



3Å diffraction; grid size of 1Å



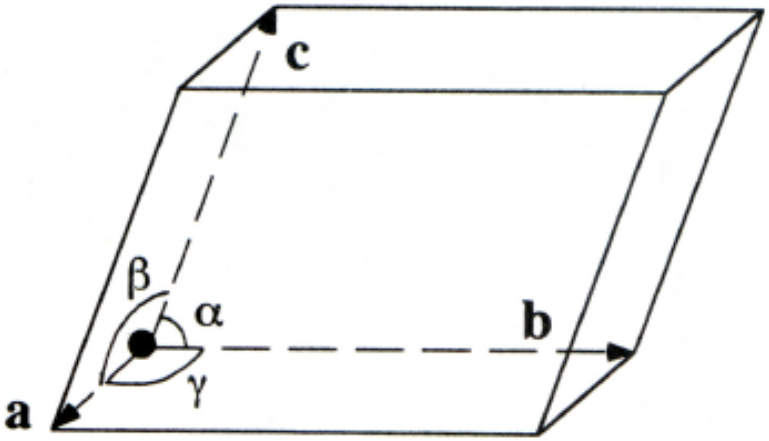
Reciprocal unit cell: the larger the real space unit cell, the smaller the reciprocal unit cell.

In reciprocal space, each diffraction spot (or 'reflection') is represented by its structure factor F(h,k,l). With an amplitude |F| and a phase α.
For example, reflection h=5, k=2, l=0 has an amplitude (color saturation) and a phase of ~60°.

A 3D unit cell:

Defined by three unit cell length: a, b, c,

and three unit cell angles: α, β, γ



Relationship between real and reciprocal unit cells

$$\mathbf{a}^* = \frac{\mathbf{b} \times \mathbf{c}}{V_c}, \quad \mathbf{b}^* = \frac{\mathbf{c} \times \mathbf{a}}{V_c}, \quad \mathbf{c}^* = \frac{\mathbf{a} \times \mathbf{b}}{V_c}$$

with $V_c = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = 1/V_c^*$.

Table 4.1. The Relationship between the Axes and Angles in the Direct and the Reciprocal Lattice in a Triclinic Space Group

$$a^* = \frac{bc \sin\alpha}{V} \qquad a = \frac{b^*c^* \sin\alpha^*}{V^*}$$

$$b^* = \frac{ac \sin\beta}{V} \qquad b = \frac{a^*c^* \sin\beta^*}{V^*}$$

$$c^* = \frac{ab \sin\gamma}{V} \qquad c = \frac{a^*b^* \sin\gamma^*}{V^*}$$

$$V = \frac{1}{V^*} = abc\sqrt{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma}$$

$$V^* = \frac{1}{V} = a^*b^*c^*\sqrt{1 - \cos^2\alpha^* - \cos^2\beta^* - \cos^2\gamma^* + 2\cos\alpha^*\cos\beta^*\cos\gamma^*}$$

$$\cos\alpha^* = \frac{\cos\beta\cos\gamma - \cos\alpha}{\sin\beta\sin\gamma} \qquad \cos\alpha = \frac{\cos\beta^*\cos\gamma^* - \cos\alpha^*}{\sin\beta^*\sin\gamma^*}$$

$$\cos\beta^* = \frac{\cos\alpha\cos\gamma - \cos\beta}{\sin\alpha\sin\gamma} \qquad \cos\beta = \frac{\cos\alpha^*\cos\gamma^* - \cos\beta^*}{\sin\alpha^*\sin\gamma^*}$$

$$\cos\gamma^* = \frac{\cos\alpha\cos\beta - \cos\gamma}{\sin\alpha\sin\beta} \qquad \cos\gamma = \frac{\cos\alpha^*\cos\beta^* - \cos\gamma^*}{\sin\alpha^*\sin\beta^*}$$

In most cases, the relationship between real and reciprocal unit cells is much simpler.

For example, if $\alpha = \gamma = 90°$

$$a^* = \frac{1}{a\sin\beta},\ b^* = \frac{1}{b},\ c^* = \frac{1}{c\sin\beta}$$

If $\alpha = \beta = \gamma = 90°$

$$a^* = \frac{1}{a},\ b^* = \frac{1}{b},\ c^* = \frac{1}{c}$$

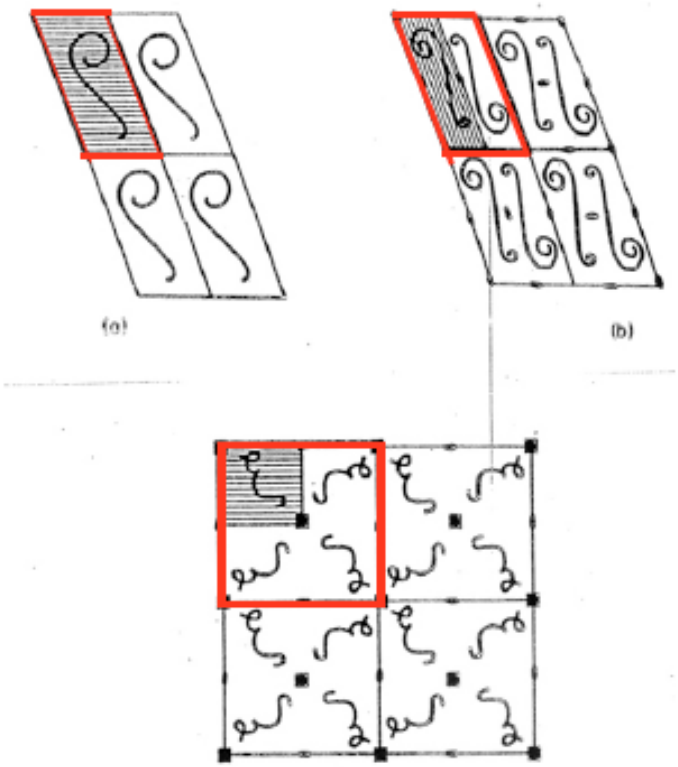
Crystal Symmetry and Space Group

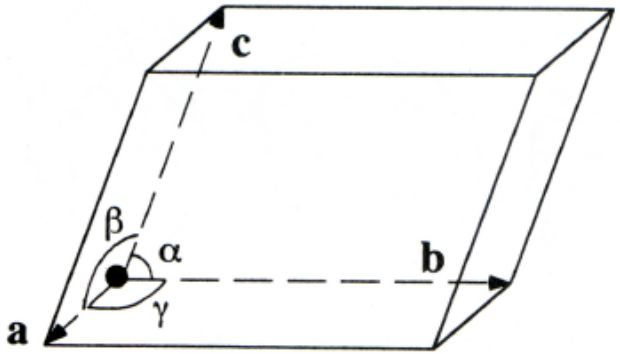
I. Definition

- Space lattice: an arrangement of points such that each point is in exactly the same environment and in the same orientation as every other point.

- Unit cell: basic parallelepiped shaped block from which the whole volume of the crystal may be built by regular assembly of these blocks.

- Asymmetric unit: basic repeating object which is related to all the other identical objects in the unit cell by the operation of the symmetry elements (*i.e.* the unique part of the unit cell with no crystal symmetry, which is the part that one needs to determine).





The rotational symmetry elements (a) 1, (b) 2, and (c) 4. A single unit cell is outlined by red lines. Other unit cell are shown with single lines (in black). The asymmetric unit is shown shaded. (After Kestleman, 1968).

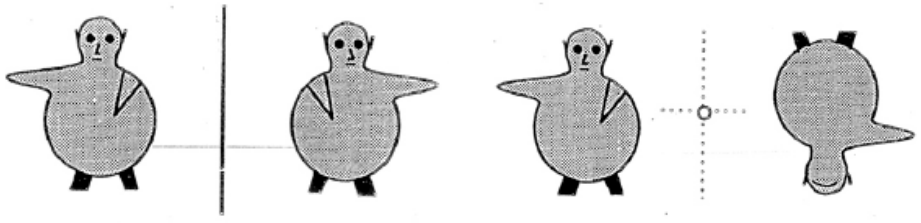
II. Symmetry operations

A. There are 3 basic types of symmetry operations:
   1. Plane symmetry: mirror or glide
   2. Rotational symmetry: 2-, 3-, 4-, 6- fold.
   3. Inversion about a point

   However, biological molecules (proteins DNA) are enantiomorphic and therefore ONLY rotation axis does occur.

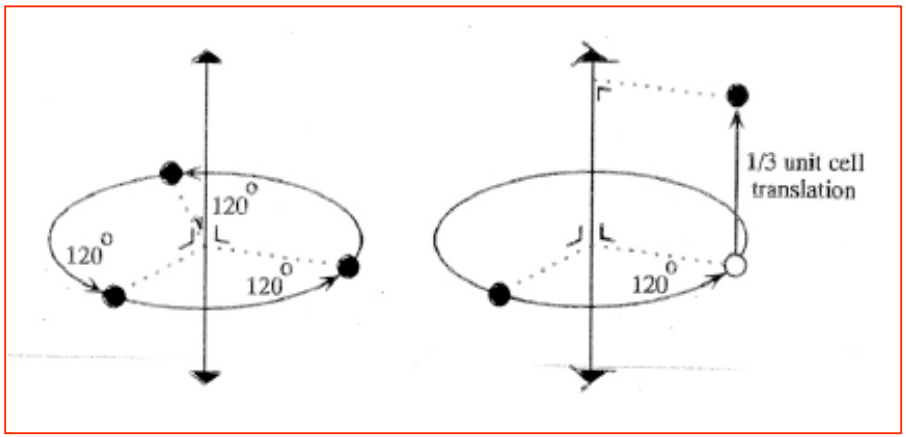
   An x-fold rotation axis rotates an object 360°/x about the axis.

B. An additional sub-symmetry operator: the screw axis.



mirror plane

Inversion center



A 3-fold axis

A 3-fold screw axis

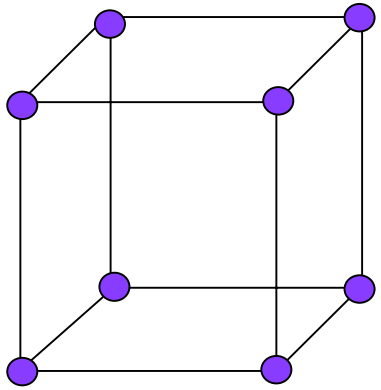
Asymmetric unit does not necessarily have anything to do with the functional unit of molecular assembly.

The symmetry of functional macromolecular complexes in solution is sometimes important to understanding their functions. Care should be taken to distinguish the crystallographic asymmetric unit from the functional unit, which the Protein Data Bank has dubbed the "biologically functional molecule."

For example, the functional unit of mammalian hemoglobin is a complex of four subunits, two each of two slightly different polypeptides, called α and β. We say that hemoglobin functions as an $\alpha 2 \beta 2$ tetramer (quaternary structure). In some hemoglobin crystals, the twofold rotational symmetry axis of the tetramer corresponds to a unit-cell symmetry axis, and the asymmetric unit is a single $\alpha\beta$ dimer. In other cases, the crystallographic asymmetric unit may contain more than one $\alpha 2 \beta 2$ biological unit.

III. The lattice points

By definition, a lattice point has the same environ as all other lattice points.
The easiest way to obtain such a set of points is to place each point at the corner of the unit cell. Such a lattice is called “primitive”



A primitive Lattice

Bravais Lattices

Symbol	Type	Description
P	Primitive	Lattice points only at corners of the unit cell
A, B, or C	Face-centered	Lattice points at center of one face of the unit cell
F	All face-centered	Lattice points at centers of all faces of the unit cell
I	Body-centered	Lattice points at body center of cell
R	Rhombohedral	Lattice points only at corners of cell with 3-fold axis along one body diagonal

P cells contain 1 lattice point

C cells contain 2 lattice points

F cells contain 4 lattice points

R cells contain 3 lattice points

IV. The seven crystal systems and their unit cell requirements

Rotation axis and combinations of rotation axes define seven types of unit cells (or crystal systems)

TABLE 4.I
The seven crystal systems

Name	Possible Bravais Lattices	Axes of symmetry	Lattice	
Triclinic	P	No axes of symmetry	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$
Monoclinic	P,C	1 dyad axis (parallel to b)	$a \neq b \neq c$	$\alpha = \gamma = 90^\circ \neq \beta$
Orthorhombic	P,C,I,F	3 dyad axes mutually orthogonal	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Tetragonal	P,I	1 tetrad axis (parallel to c)	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Trigonal	P (or R)	1 triad axis (parallel to c)	$a = b \neq c$	$\alpha = \beta = 90^\circ, \gamma = 120^\circ$
			$a = b = c$	$\alpha = \beta = \gamma < 120^\circ, \neq 90^\circ)$
Hexagonal	P	1 hexad axis (parallel to c)	$a = b \neq c$	$\alpha = \beta = 90^\circ, \gamma = 120^\circ$
Cubic	P,I,F	4 triad axes (along the diagonals of the cube)	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

V. The 14 Bravais lattices

<u>The 14 space lattices</u>

P: primitive;

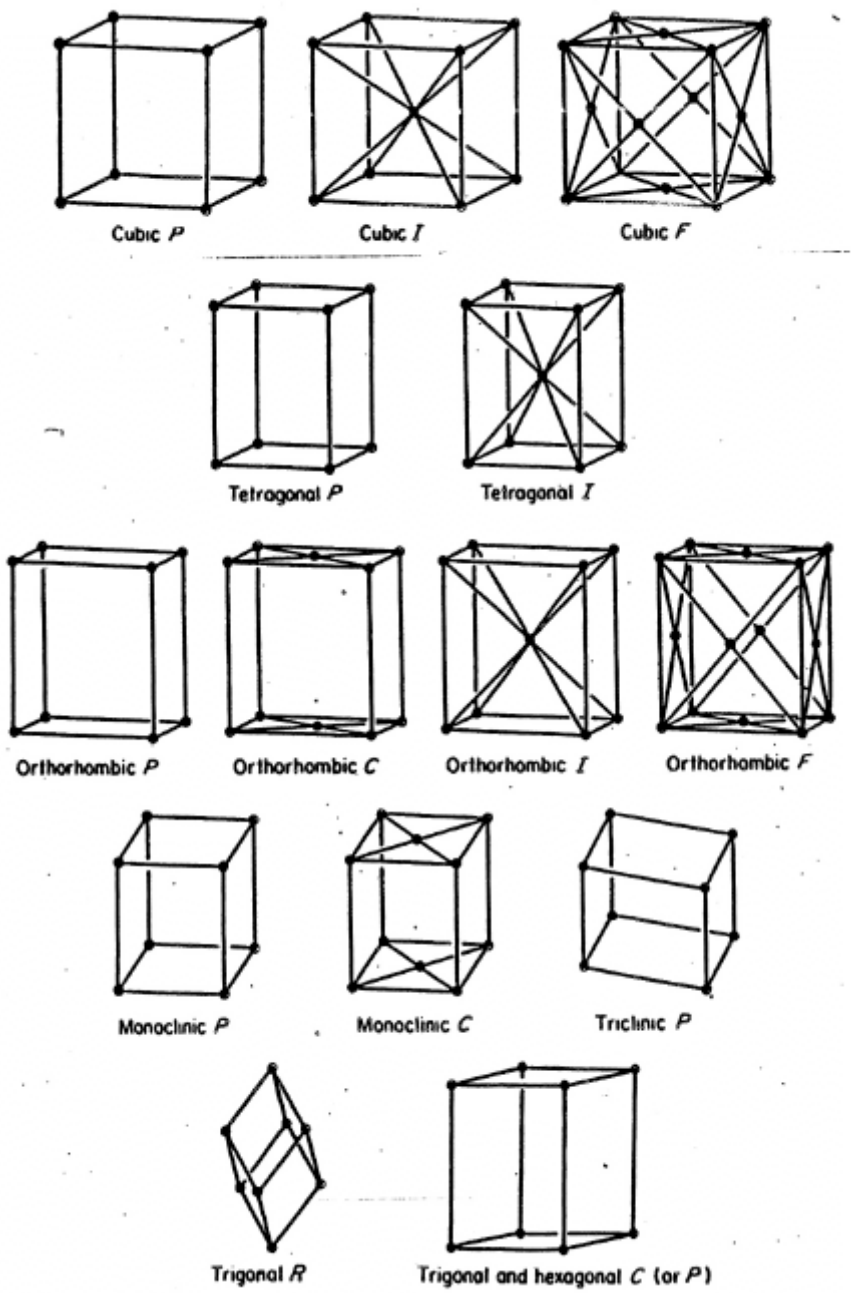
Non-primitive--

C: face-centered on the AB face;

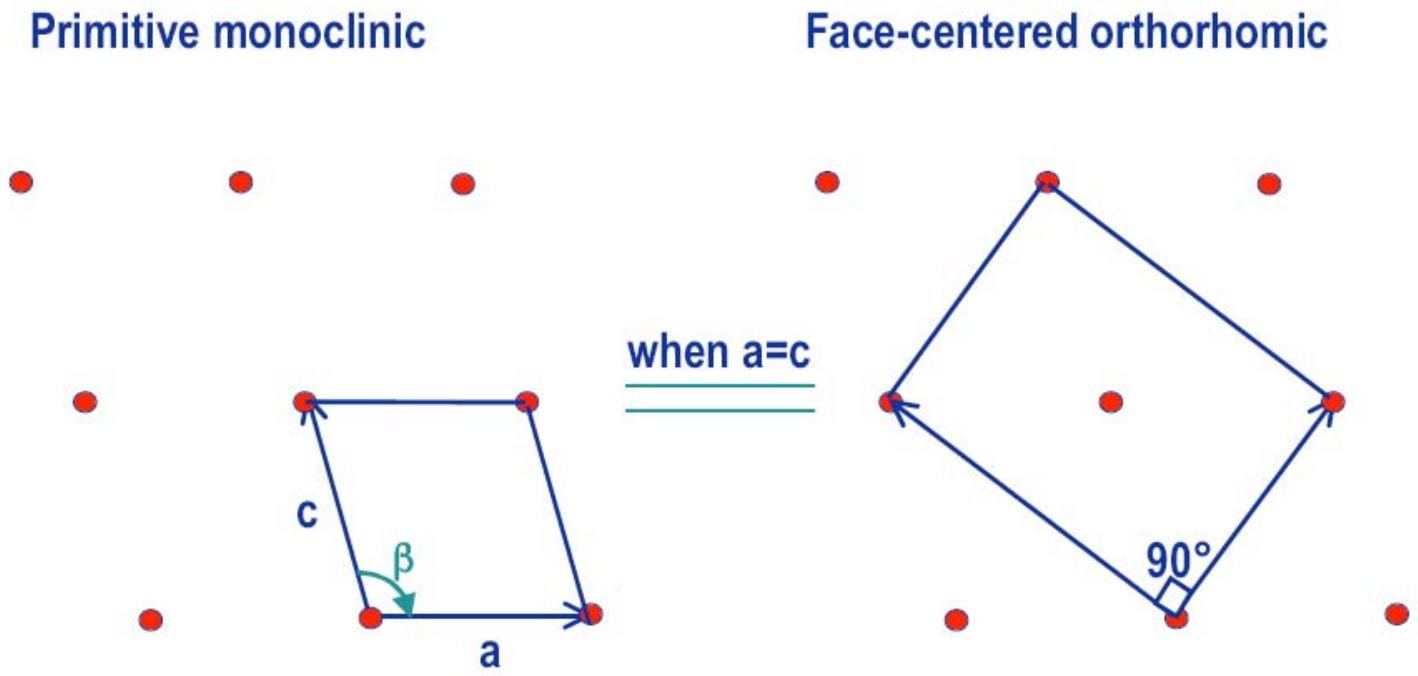
F: face-centered on all the faces;

I: body-centered.

Any diffraction may be interpreted by a primitive lattice; however, you may get higher symmetry when choosing a non-primitive lattice to decrease the 'asymmetric unit'.



A simple case on how a primitive unit cell can be converted to a non-primitive unit cell



VI. 230 space groups

By combining all known symmetry operators, 230 space groups can be defined.
However, because proteins are enantiomorphic, only 65 space groups occur in protein crystals.

TABLE 4.II
Enantiomorphic space groups

System	Class	Space group symbols
Triclinic	1	$P1$
Monoclinic	2	$P2, P2_1, C2$
Orthorhombic	222	$C222, P222, P2_12_12_1, P2_12_12, P222_1, C222_1, F222, I222, I2_12_12_1$
Tetragonal	4	$P4, P4_1, P4_2, P4_3, I4, I4_1$
	422	$P422, P42_12, P4_122, P4_12_12, P4_222, P4_22_12, P4_32_12, P4_322, I422, I4_122$
Trigonal	3	$P3, P3_1, P3_2, R3$
	32	$P312, P321, P3_121, P3_112, P3_212, P3_221, R32$
Hexagonal	6	$P6, P6_5, P6_4, P6_3, P6_2, P6_1$
	622	$P622, P6_122, P6_222, P6_322, P6_422, P6_522$
Cubic	23	$P23, F23, I23, P2_13, I2_13$
	432	$P432, P4_132, P4_232, P4_332, F432, F4_132, I432, I4_132$

How to determine the space group of a particular crystal?

Due to the effect of crystal unit cell and symmetry on the diffraction pattern, the space group of the crystal can be determined from its diffraction pattern.
Cell dimensions may be deduced from a single diffraction pattern. These will SUGGEST possible space groups.

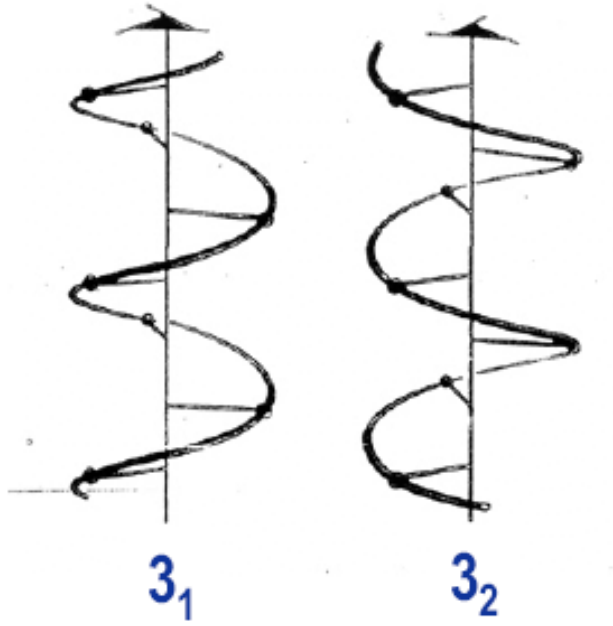
Symmetry of the crystal may be determined by comparing the amplitudes (or intensities) of potentially 'symmetry-related/equivalent' reflections. Symmetry DETERMINES the space group.

If you have a 'handed' space group, the enantiomorph (hand) of the space group can not be determined from the diffraction pattern, e.g. 3_1 versus 3_2; 6_1 versus 6_5; 6_2 versus 6_4; 4_1 versus 4_3.

Handedness of screw axes

For example, the 3_1 and the 3_2 axes are indistinguishable from a diffraction pattern.

The exact hand of a space group can only be determined during phase determination and/or from the ‘expected’ hand of a chiral molecule.



VII. Why bother with space groups?

- Reduce the size of the asymmetric unit (a unique part of the crystal).

- Simplify data collection by reducing the number of unique reflections to be measured.

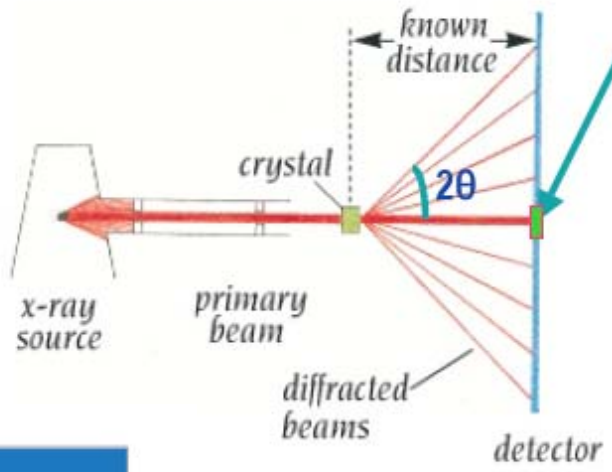
- Or increase the redundancy of the data.

- Symmetry may be useful for phase determination.

- International Table (Volume I) is a ‘bible’ for crystal symmetry.

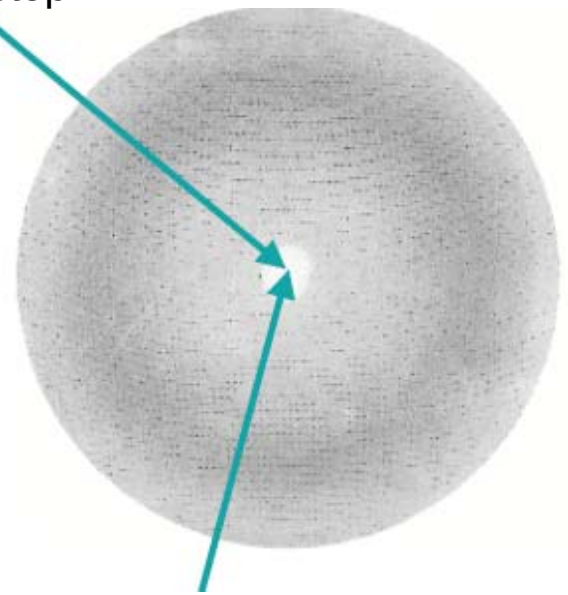
3. Diffraction data collection and crystal characterization

A diffraction experiment setup



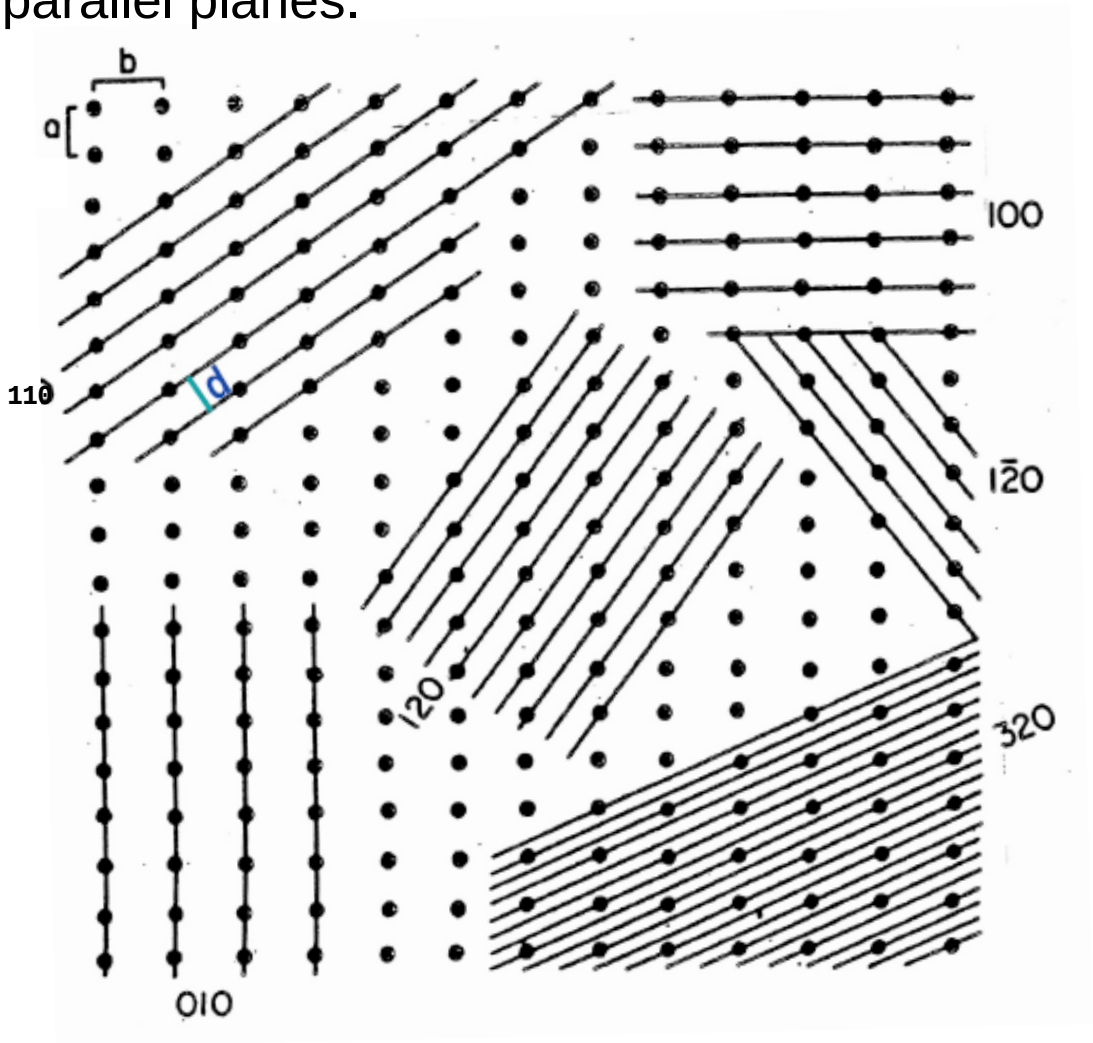
θ: diffraction angle

Beam stop



Origin of the reciprocal space

Bragg showed that diffraction spots emerged from a crystal can be treated as if they were ‘reflections’ from sets of parallel lattice planes. 3 indices, hkl, identify a set of equivalent, parallel planes.



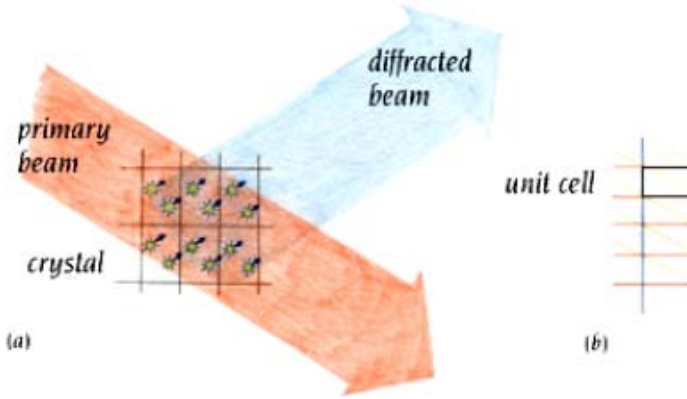
d is known as “resolution”.

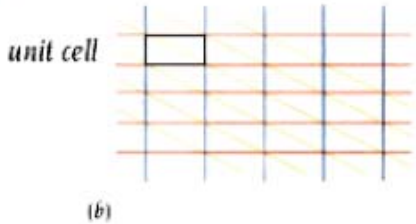
A lattice plane is indexed (hkl) when it makes intercepts a/h, b/k and c/l with the edges of the unit cell. [$h = 1/n_a$, where n_a= the number of times the a-axis is divided etc.]

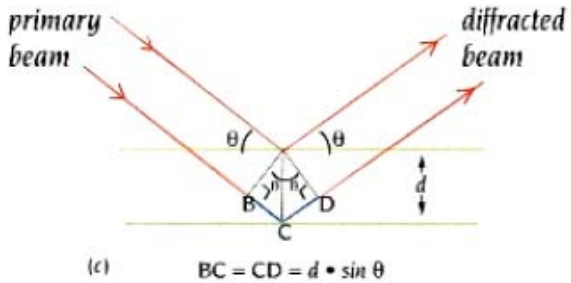
The larger the indices, the smaller the inter-plane distances (d).

Bragg's law: $2d\sin\theta = \lambda$

When this condition is satisfied, a plane with a spacing of d will 'reflect' X-rays with an incident angle of θ, due to constructive interference.







Most of the scattered x-rays cancel out (destructive), but certain directions (blue, constructive), they reinforce each other and add up to a diffracted beam.

Path difference: $2d\sin\theta$

d: resolution

θ : scattering angle

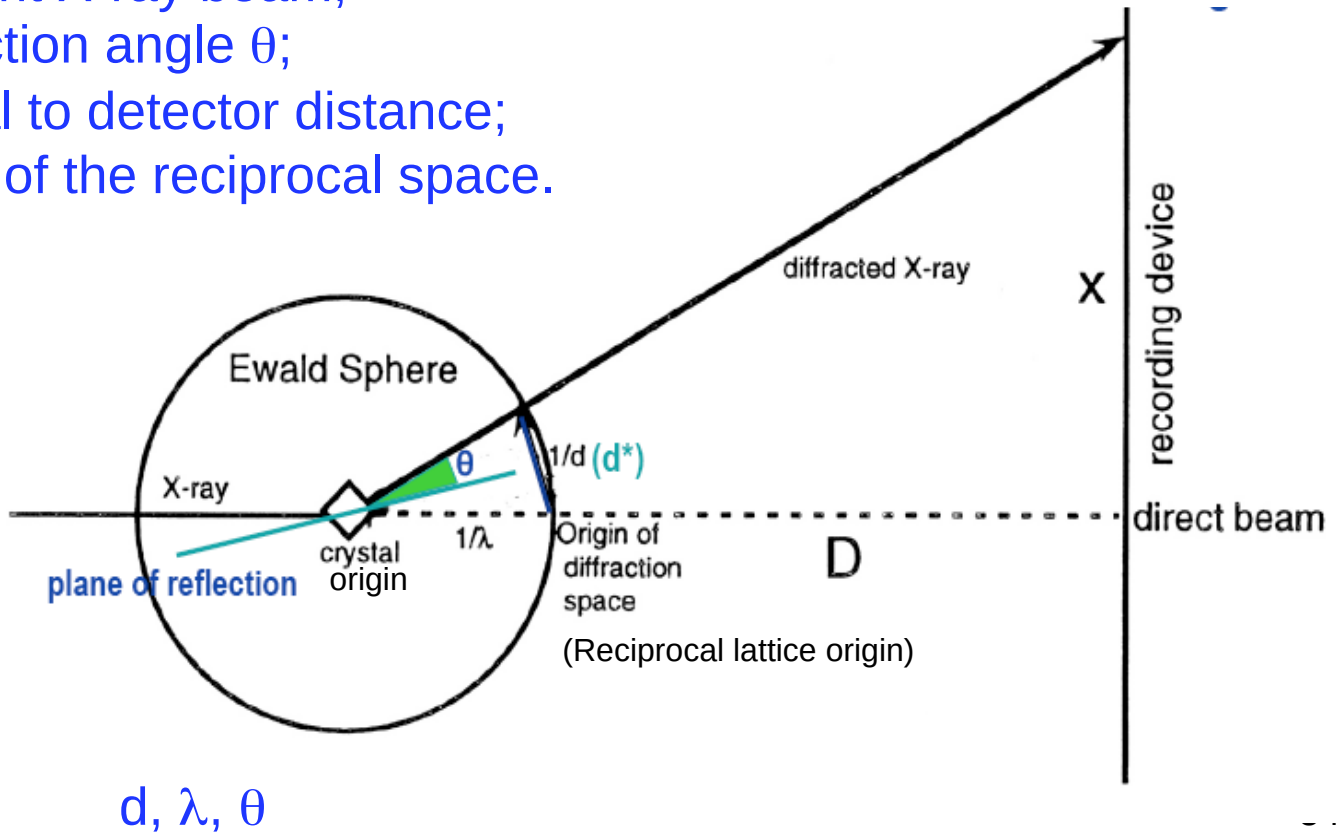
Ewald construction:

Any point on the Ewald sphere follows the Bragg's law:

$2d\sin\theta = \lambda$

Ewald sphere:

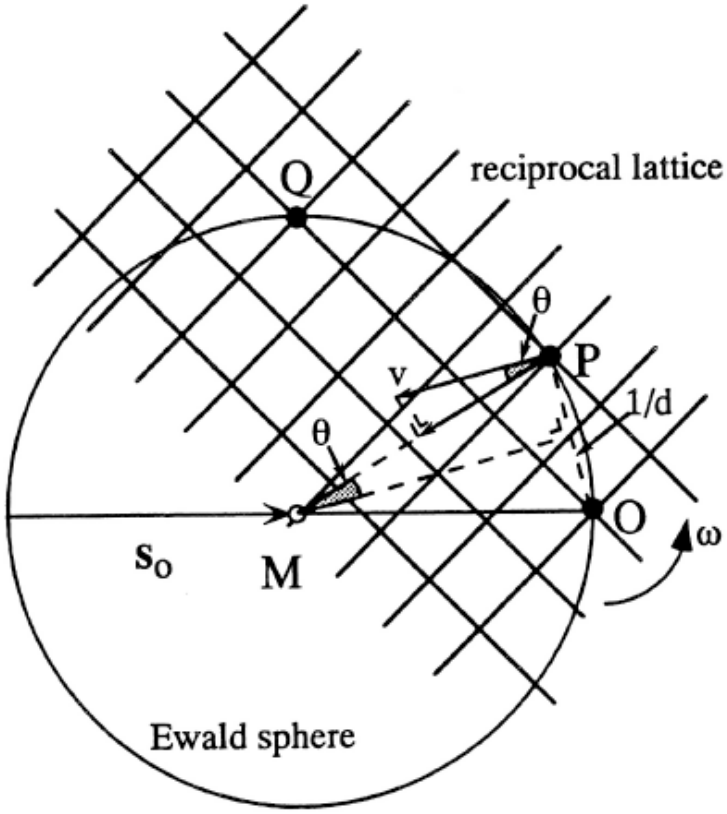
center of the sphere - crystal;
radius -$1/\lambda$;
incident X-ray beam;
diffraction angle θ;
crystal to detector distance;
origin of the reciprocal space.



d, λ, θ

At a certain orientation of the crystal, only a small number of reflections are in ‘diffraction condition’, i.e. those lattice points (P, Q) that intercept with the Ewald sphere.

Therefore, the crystal has to be rotated to collect all the reflections, *e.g.* one may need 90° for a certain crystal, which could mean 90 images with 1° oscillation each.



Data Collection: measuring positions (h,k,l) and intensities (*I*) of reflections in the diffraction pattern produced by the macromolecular crystal.

Capillary mounting

for crystal characterization and data collection above 0°C.

Cryo-Crystal mounting

Soak the crystal in a cryo-protectant solution, then collect data around 100K. (Commonly used cryo-protectant: glycerol, ethylene glycol, PEG400, sugars, MPD, etc)

Each diffraction image contains diffraction from a small angular sweep (oscillation). Rotate the crystal until a complete 3-D diffraction is completed and collected.

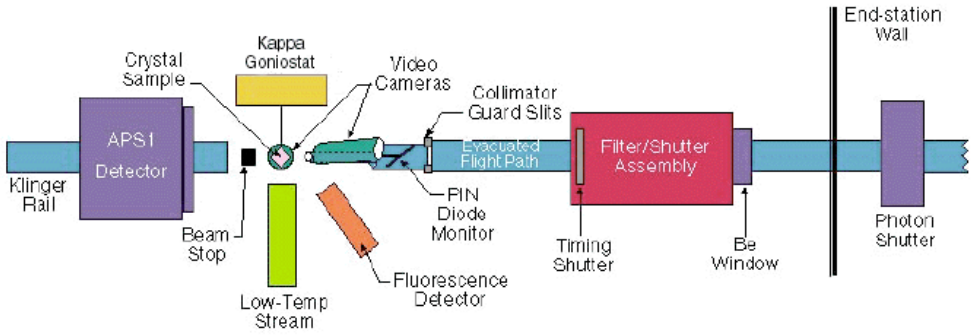
These diffraction images are recorded by imaging plates or CCD detectors as digital images.

Beam sources: collect X-ray diffraction data

- Initiate experiments using home-source x-ray generator and detector
- Determine liquid nitrogen cryo-protection conditions to reduce crystal decay
- While home x-rays are sufficient for some questions, synchrotron radiation is preferred
- Anywhere from one to **hundreds** of crystals and diffraction experiments may be required

A “home” X-ray source

A Synchrotron X-ray source



Argonne National Laboratory Structural Biology Center beamlineID19

at the Advanced Photon Source http://www.sbc.anl.gov

Crystal characterization

#Diffraction resolution limit. $d = \frac{\lambda}{2\sin\theta}$

will determine what kind of detail you may see from the final structure. Usually 3.0Å or beyond is OK, and 2.0Å or beyond is good. Note that the resolution limit of most macromolecular crystals does not give "atomic" resolution, *i.e*. electron density does not resolve into atoms; but the fitting of "atomic models" into an electron density leads to "atomic" resolution.

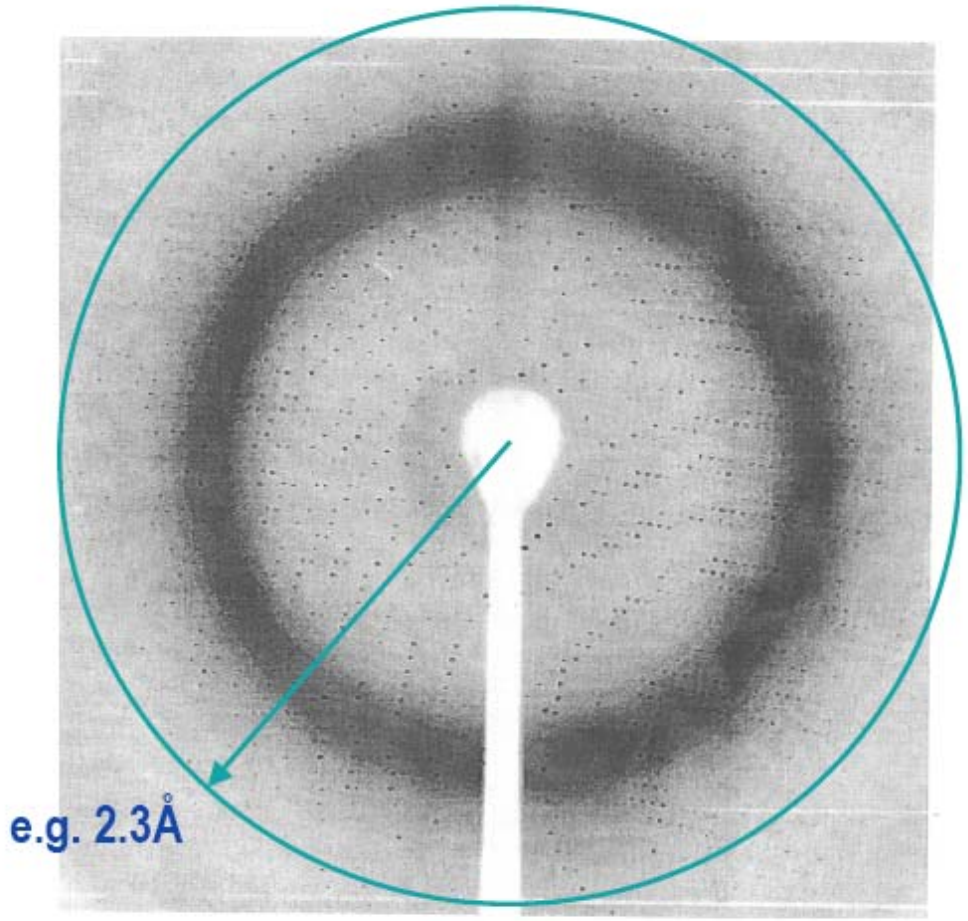
Cell dimensions.

Mosaicity: usually 0.3-1.0°. Indication of variability between different unit cells in the crystal. High mosaicity increases the width of reflections, increases the number of reflections on each image and may cause reflection overlap, leading to problems in data collection.

Space group.

Data quality: R factor.

Resolution limits (mostly between 3-1.5Å)



Mosaicity increases the width of reflections

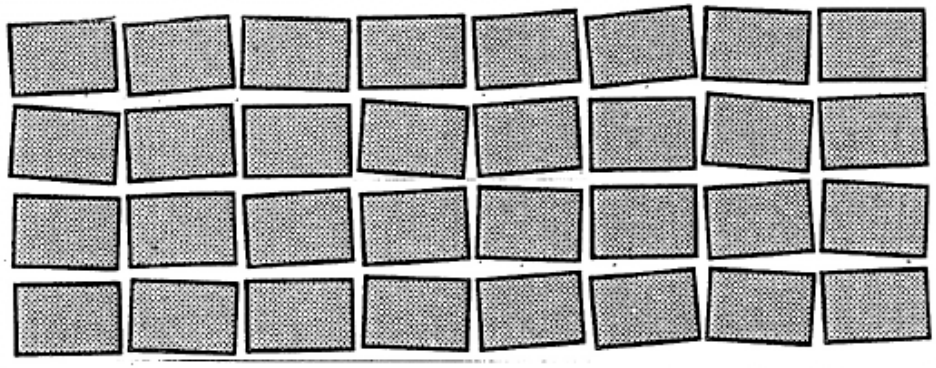


Figure 4.29. Most crystals are imperfect and can be regarded as being composed of small mosaic blocks.

'normal' mosaicity of protein crystals: approximately 0.2-1.0°

Data processing: a typical output

Space group: P212121

Shell Lower limit	Upper limit	% of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
99.00	4.09	1.0	2.7	4.1	5.1	6.2	9.6	19.2	80.4	99.6
4.09	3.25	1.0	2.8	4.1	6.0	8.3	14.4	28.0	71.9	99.9
3.25	2.84	1.2	4.3	6.4	8.8	13.7	24.3	45.3	54.7	100.0
2.84	2.58	1.4	7.0	12.1	16.0	24.0	39.7	64.8	35.0	99.8
2.58	2.39	1.5	6.0	12.8	19.4	29.9	48.2	76.7	23.1	99.8
2.39	2.25	2.8	10.9	18.4	26.7	37.6	57.4	82.1	17.5	99.5
2.25	2.14	2.4	11.0	20.5	27.9	40.5	62.7	87.4	12.3	99.7
2.14	2.05	5.0	15.6	27.1	37.8	51.5	73.7	92.3	6.6	98.9
2.05	1.97	5.2	19.9	32.9	45.2	60.4	83.0	96.5	1.8	98.4
1.97	1.90	6.3	25.7	43.6	56.0	70.7	88.4	94.6	0.4	95.0
All hkl		2.7	10.4	17.9	24.5	33.8	49.4	67.9	31.2	99.1

Completeness

Summary of reflections intensities and R-factors by shells
R linear = SUM (ABS(I - <I>)) / SUM (I)
R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)
Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))
In all sums single measurements are excluded

Shell Lower limit	Upper limit	Average I	Average error	Norm. stat.	Chi**2	Linear R-fac	Square R-fac
99.00	4.09	5589.5	156.6	143.1	0.734	0.038	0.064
4.09	3.25	4391.0	131.0	109.4	0.693	0.040	0.053
3.25	2.84	2118.1	82.2	73.3	0.607	0.050	0.166
2.84	2.58	1164.9	64.3	60.1	0.583	0.065	0.073
2.58	2.39	839.3	61.9	59.5	0.572	0.083	0.089
2.39	2.25	695.7	64.8	61.7	0.524	0.100	0.109
2.25	2.14	643.9	68.2	65.5	0.519	0.115	0.178
2.14	2.05	496.0	71.9	70.0	0.436	0.135	0.149
2.05	1.97	394.6	79.6	78.3	0.421	0.170	0.182
1.97	1.90	296.9	91.6	90.8	0.343	0.202	0.232
All reflections		1729.2	88.1	81.9	0.562	0.056	0.077

R factor

Intensities of systematic absences

h	k	l	Intensity	Sigma	I/Sigma
0	0	3	-2.9	7.5	-0.4
0	3	0	3.9	17.4	0.2
3	0	0	8.7	9.1	1.0

systematic absences

$$R_{sym}(I) = \frac{\sum_{hkl}\sum_{i}\left|I_i(hkl) - \overline{I(hkl)}\right|}{\sum_{hkl}\sum_{i} I_i(hkl)}$$

R_{sym} of diffraction data: comparison among intensities of symmetry-related reflections.

Crystal characterization

Possible solvent content

Since many proteins have globular shapes, the expected solvent content is ~50%.

A easy way to calculate solvent content is by using V_m (Matthews coeff.).

$$V_m = \frac{V}{MW \bullet N \bullet n}$$

V: volume of unit cell;

MW: molecular weight of your protein in Daltons;

N: number of asymmetric unit;

n: possible # of molecules per asymmetric unit.

Normal range of V_m is roughly 1.7-3.5 $Å^3$/Dalton.

% of solvent = $1-1.23/V_m$ (normally 30-70%)