

for Crystallography (up to Mid-sem)

: Text Book by Jan Drenth
by Johnson & Stolt (for symmetry) - 18 pages
by Bergfors. (for Crystallization only)

Intro to Prot. Structure by Branden & Tooze (Structures & fold only)

After Mid Sem:

Prot. Engg & Design by Park & Cochran,

Prot. Engg & Design by Paul R. Carey

Journal Papers : (4)

Mid Sem = 30

End Sem = 50

Internal Assessment = 20

Attendance
Quizzes
Class Test
Assignment
Board work

M

A.

R

K

T

N

G

Protein Structures

atom-atom
distance

3D Structures

- ① single crystal X-ray crystallography
- ② NMR Spectroscopy
- ③ Cryo-electron microscopy
(single particle + tomography)
- ④ Modelling

purified protein

Recombinant
or
wild type

Fold/Local Structure

- ① CD (Circular dichroism) & fluorescence
- ② FTIR (Fourier Transform IR spec)
- ③ Raman Spectroscopy
- ④ Analytical ultracentrifugation
- ⑤ Mass spectrometry

Bottlenecks:

- ① high amt. of protein $\approx 10 \text{ mg/ml}$ (10 mg of protein)
- ② Crystals \rightarrow (in God's hand :))
- ③ Diffraction Limit \rightarrow availability \rightarrow at least 3\AA $\rightarrow 2\text{\AA}$ is better than 3\AA
- ④ Phasing \rightarrow program Resolution

NMR Spectroscopy Bottlenecks:

- ① Solution state
- ② Protein should be labelled with C¹³, N¹⁵.
- ③ Strong field ~ 600 MHz magnet
- ④ Limited molecular wt. ~ 15-20 kDa (in monomeric)
- ⑤ Average structure → since molecules are in Brownian Motion.

Cryo-EM Bottlenecks:

- ① Big assembly like ribosome
 - proteasome
 - secretory system, etc.
- ② Small molecule like 30-60 kDa (or) <100 kDa challenging
- ③ Tomography → (within the cell but at ~8 Å level)

Modelling:

↓
Alpha Fold

using known structures & training set
and use of AI/ML

but, ① Can you predict the multimeric state? → Not yet!

Circular Dichroism (CD):

- only secondary structure (βS)
- % prediction

✓ whether properly folded

same as ✓ whether there will be any change is αs upon binding to ligands

FTIR

Raman

Analytical ultracentrifuge:

- Multimeric state
- Binding site identification

Fluorescence:

Quenching study → location of particular domain or aa
[Whether in the core or outside (protruding)]

Purified protein

homogeneous purification

conformational homogeneity

- through
- ① SDS-PAGE
 - ② Western Blotting
 - ③ Molecular weight
 - ④ Assay for the activity

If DNA bound -

- ① How to confirm whether DNA is bound to every prot. molecule?
- ② Whether they are all identical DNA?
- ③ OD_{260}/OD_{280} ratio.
- ④ PEI precipitation, etc.

↓
polyethylene imine → to ppt DNA only

Recombinant expression system

- ① Bacterial
- ② Yeast
- ③ Baculovirus
- ④ Mammalian cell culture

Vectors:

① Bacterial vectors

His-tagged → Histidine (6X)

GST-tagged → Glutathione S-transferase

Thr-tagged → Thiorredoxin

MBP-tagged → Maltose binding
(or) others

Q. Should we remove the tag? Then how? What are the advantages / disadvantages?

Wild type (not through recom")

- ① heart muscle → Recent paper in NATURE; actin & myosin
- ② membrane ⇒ how lub-dub, sounds is generated in the heart
- ③ agglomerated form

Dutta Debabrata
Roger Craig

9/11/23

Crystallization

programmed way of precipitation

random process

happens from supersaturation

Unsaturation

✓ Brownian motion → molecules do not come close to interact

e.g. $\text{Na}^+ \text{Cl}^-$ no defects no crystallization

✓ small molecule → no conformational variations

big molecule → conformations → have to be purified in one conformation

conformation /
defects

interactive forces → crystal packing

✓ crystal defects Schottky
frenkel

e.g. Cys might get oxidized

mixt. of normal Cys & oxidized Cys-containing proteins

surface corrosion

✓ crystals can be grown only to a particular size.

due to presence
of crystal defects

✓ crystallization in enantiomeric forms:

both D and L forms w.r.t.
not $\frac{1}{2}$ of plane polarized light

✓ Pasteur → tartaric acid
crystals

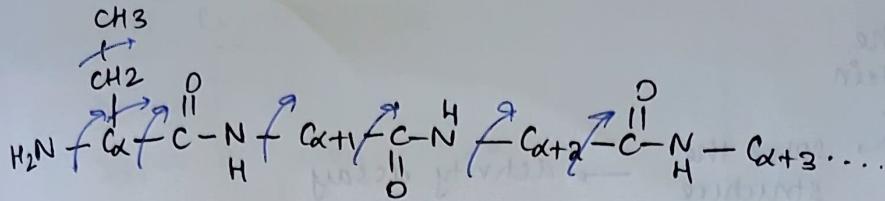
(D/L) → pasteur saw crystals
of 2 different
morphologies

not a problem

for proteins as
naturally all

a.a. are L form

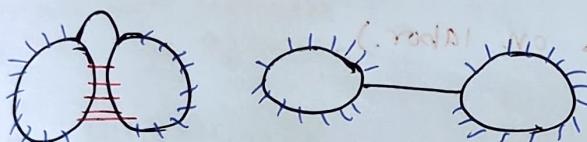
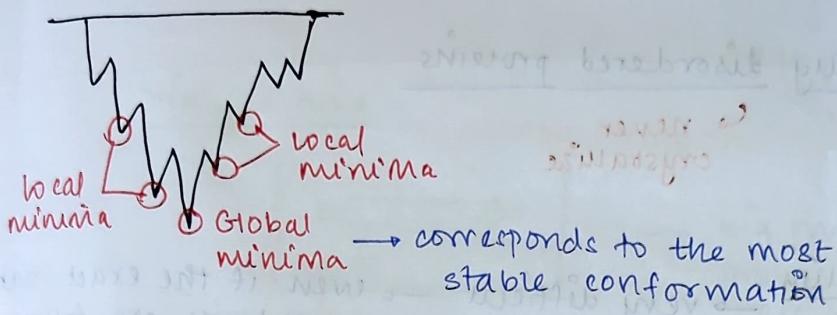
presence of both might
hamper crystallization;
sometimes facilitate.



→ possible rotations

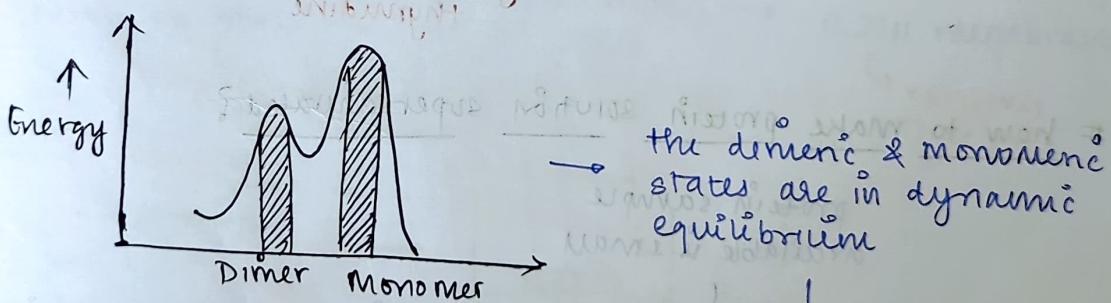
so protein can have enormous no. of conformations.

but only one of these is stable & biologically active



- available interactions
- blocked interactions

closed conform → open conform
(no. of available interactions → with the purification matrix resin)



the dimeric & monomeric states are in dynamic equilibrium

very difficult to isolate dimeric / monomeric state

In that case, we need to know the exact conditions which would shift the equilibrium to a particular state

✓ purify the protein

↳ solve the structure → activity assay

↓
mutate residues → activity assay → if activity is affected, we would know that our original structure may be the correct one

✓ an active protein in-vitro may

✓ intrinsically disordered proteins

↳ never crystallize

reproducibility of crystals

→ very difficult → even if the exact same procedures are followed

(depends on labor.)

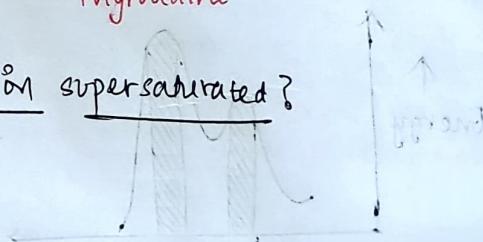
azidothymidine → 2'-OH group is replaced by azide (N_3^-)
(HIV drug)

incorporated in DNA
preferably over normal thymidine

DNA synthesis stopped

how to make protein solution supersaturated?

protein sample available is small



↳ if we play with the medium → supersaturation possible at low conc.

- Protein conc. (10 mg/mL) → how to concentrate?
- Precipitating agent
- Buffer (pH variation)
- Ions (Mg^{2+} , Zn^{2+} , ...)
- Ligands
- Temperature (4°C , 16°C , 25°C)

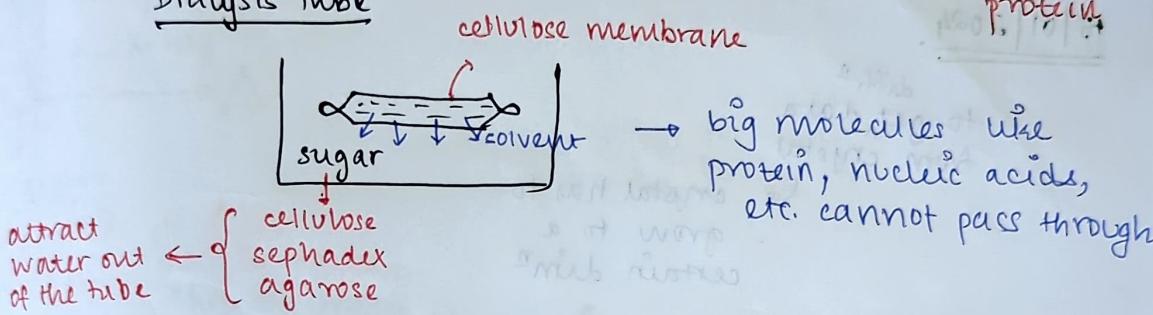
Method of Crystallization:

- ① Dialysis
- ② Vapor Diffusion



✓ we have (say) 1 mg / 1 L protein → we want 1 mg / mL → concentrate ↓ heat? → may denature protein

Dialysis tube



Ultrafiltration

molecular wt. cut-off membrane

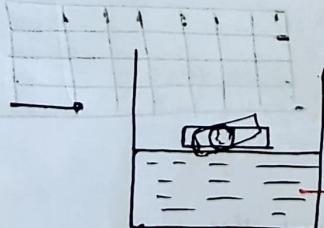
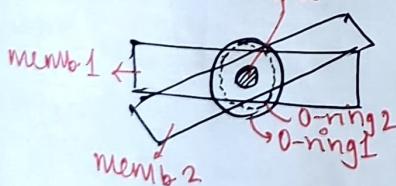
removal of dust particles / particulate matter

if precipitation occurs on dust, → cryst. stops

then pass through 0.22 μ membrane

removes bacteria

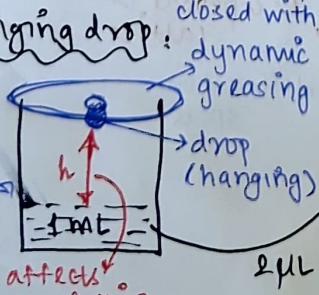
Microdialysis



Vapor diffusion:

① Hanging drop:

(+) diff. compositions
(mother liquor)



cover slip closed with dynamic greasing

drop (hanging)

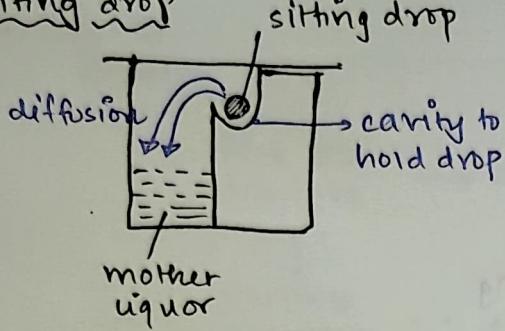
2 μL

protein solution

cover slip treated with silane

hydrophobic → drop does not spread

② Sitting drop



Seeding ← Micro
Macro

15/01/2024

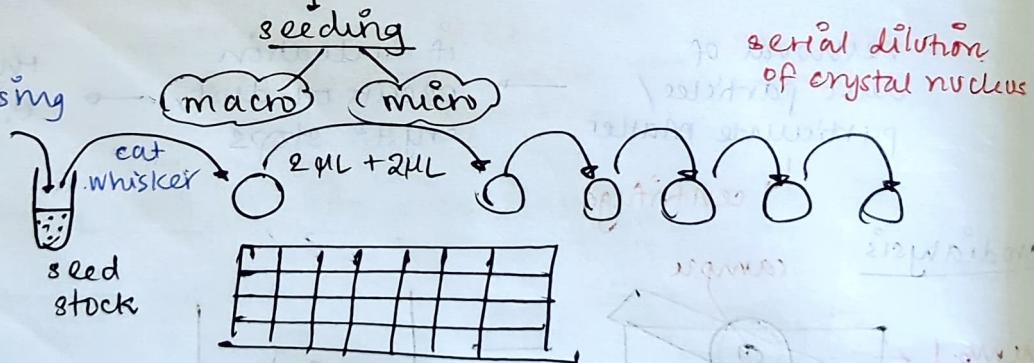
diff²
to get data
from crystal

↳ crystal has to
grow to a
certain dimⁿ

nuclear point → starting point

if crystal does
not grow → two methods

Micro: crush using
glass rod.
broken into
small points/
crystals
↳ nucleus



Macro:

no crushing → take out the
crystal and
put in a new
solution

a, b, c
Real lattice

Reciprocal lattice

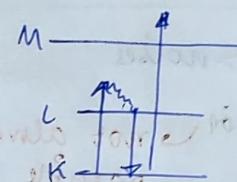
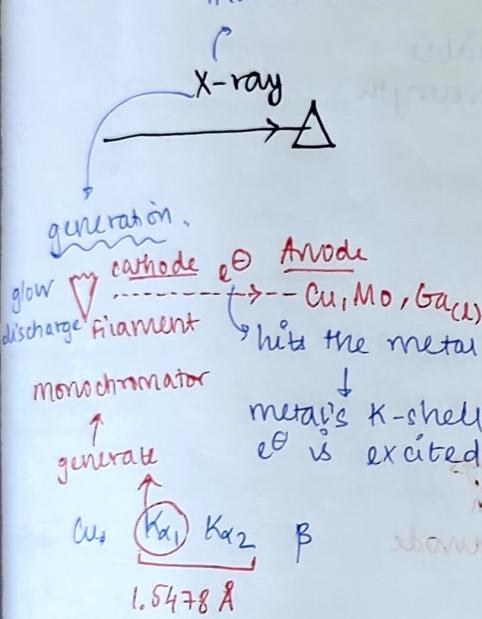
$$\frac{1}{a}, \frac{1}{b}, \frac{1}{c}$$

x-ray

crystals in
a particular
arrangement

diffraction
spots

- ✓ if in the real lattice distance is more \rightarrow in reciprocal lattice, less divergent
- ✓ detector \rightarrow must be nearer to the crystal \rightarrow better resolution
- ✓ any diffraction spot \rightarrow intensity $\propto A^2$ amplitude
- but phase is lost \rightarrow ultimately needed to determine position of constituents
- specific λ
- monochromatic



① sealed tube \rightarrow vacuum
X-ray intensity low
cathode & anode fixed
constantly hit by e^-
surface corrodes \rightarrow hole \rightarrow hole edges obstruct diffraction
intensity gradually +

water - very bad diffractor

Resolution
(say) of 2 \AA means atoms at a 2 \AA distance can be separated out

not so good for protein crystals

very good for small molecules, salts

if we do X-ray crystallography

we do not get any intensity

2 reasons

too many water molecules bound

crystal is not aligned properly with X-ray

② Rotating Anode X-ray generator

→ intensity is better compared to sealed-tube generator.

one machine → one type of wavelength
 Cu - 1.54 \AA
 Mo -
 $\text{Ga}(\text{e}) \rightarrow 2 \text{ \AA}$

solving the phase: → changing the wavelength

(anomalous dispersion)

↓ variable wavelength

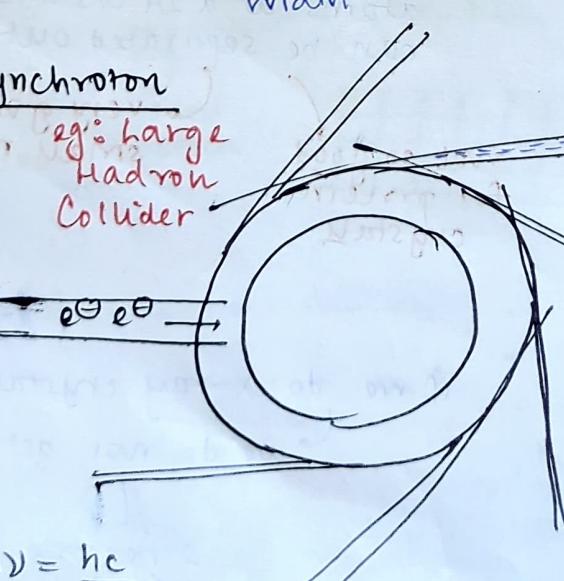
Se
 Hg - 0.98 \AA
 Pt - $\varnothing 1.5 \text{ \AA}$
 $\Omega = 2\text{R}$

beam width of 70μ
 50μ → use 10μ beam width → + noise

not always possible in rotating anode

synchrotron
 2-6 giga eV
 \approx large Hadron Collider

undulators
 magnets
 EM lenses



$$E = h\nu = \frac{hc}{\lambda}$$

varying energy, we tune the λ

variable wavelength generation

intensity is very high

data from even small crystals

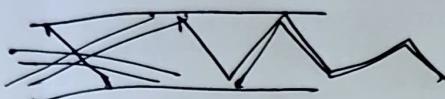


Hamburg city
 largest beam

PETRA

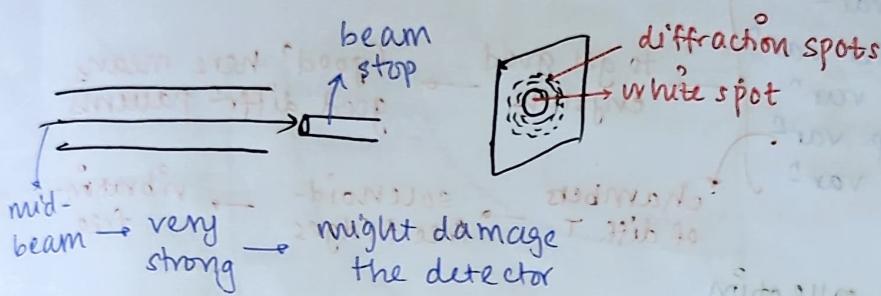
high energy x-rays → hit the crystal → at the pt. of hitting → lot of heat generation

water mol.
evaporate



intensity
adds up

∴ noise minimizⁿ → job of the engineer → yields good data.



16/11/24

Microarray Datasets

genome → ORF → protein + signal peptide → transport purpose only

- ① initial bioinformatics
- ② cloning
- ③ exprⁿ & purificⁿ
- ④ crystallization
- ⑤ Data collection
- ⑥ Structure determinⁿ
- ⑦ Validation analysis
- ⑧ BioInfo → VLS, SGD

to identify protein function

Initial bioinformatics

to identify unstructured parts

hydrophobic parts
if present → protein not in the cytosol

not in active form ← precipitates

→ making proteins soluble → challenge

e.g.: insulin → removing first few a.a.
does not affect function
but makes protein soluble

expressing human
proteins in bacteria

post-transl. modifications

codon bias

tRNA req. is
not present
in bacteria

thus, before cloning, we need
to optimize the codons.

cloning

exprⁿ & purificⁿ

crystallization

1. pH varⁿ
2. buffer varⁿ
3. salt varⁿ
4. temp. varⁿ
5. ppt. varⁿ

to get good
crystals

'good' here means
good diffraction patterns

chambers
of diff. T

solenoid-
free fridges

vibration-
free

Data collection

Structure

determⁿ
by various
methods

eg: X-ray
crystallography

validation

analysis

Bioinformatics

if the structure
follows proper
stereochemistry

whether this
kind of fold
is previously
known

if yes what is
the function

Drug development

of target protein

function

#

Target Gene

soluble protein

crystal

Data

Phasing
structure

interⁿ b/w protein
& drugs

Phase I
Phase II
Phase III
Phase IV

human trials

chimpanzee

monkeys

mice

animal
models

Structure-
based drug
design

in-silico cell
test → culture

mammalian
cells

X-ray crystallography → 2 Å level → best method to determine interⁿ at the atomic level

NMR → avg. structure only

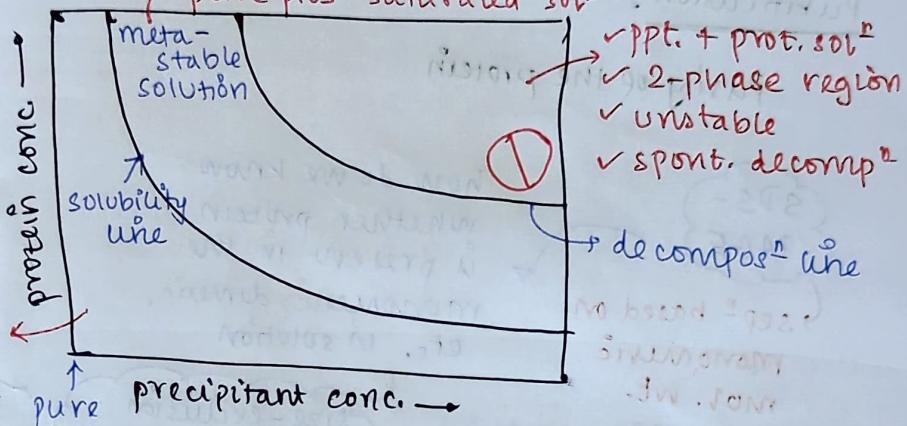
cryo-EM → 8 Å level.

-10 to -15 kcal/mol → protein stabilizes in this range

✓ personalized drug

* Principles of Protein Crystallization

will eventually separate into protein (maybe in form of crystals), pure plus saturated solⁿ



- ✓ clear protein solⁿ
- ✓ single phase
- ✓ stable

✓ precipitants: → ammonium sulfate → (mol bio grade chemicals)

① salt: AS, Na₂SO₄, phosphates

② Polyethylene Glycols (PEGs)

PEG2000	liquid	wt. % as polymers
400		
5500		

PEG-3350	solid	(2.1)
4K		
8K		

20K

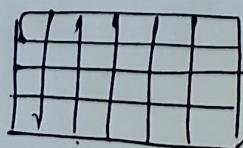
③ MPDs - methyl pentadioic

"we have many choices"

✓ buffers:

✓ ions: Zn²⁺, Mg²⁺, Mn²⁺, Cu²⁺

✓ matrix:



4 x 6
= 24
wells
(or)

8 x 12
= 96
wells

SPARSE MATRIX

Hampton Research
 Jena
 Molecular Dimension

diff. companies which
 produce diff. choices
 of the above components

- ✓ automated crystalliz^z using robots
- ✓ nanofluidics-based (UC Berkeley)
 - ↳ arbitrary mixing
 - ↳ UV scan (280nm)

Purification:

purity of the protein

SDS-PAGE

sepⁿ based on
monomeric
mol. wt.

how do we know
whether protein
is present in the
monomeric, dimeric,
etc. in solution

size-exclusion
chromatography

- (or) Gel filtration
- (or) Gel permeation

(if dimeric)

Dynamic
Light Scattering (DLS)

protein molecules in
solⁿ as particles

send laser → diffraction

hydrodynamic
radius

(globular protein
with water mol.)

monomeric

fit for crystalliz^z

monomodal

poly-modal

no fit for
crystalliz^z

Freshness :

↳ stored over time → degradation.

↳ Lyophilizer → drying machine
by vacuum

↳ dry → dry → powder

some water molecules

are structural H₂O molecules

→ on cont. drying → there also get evaporated

e.g. Lysozyme, BSA, HSA

powder form
but the structure
is lost

but not all protein
on addⁿ of water,
Lysozyme regained
structure

losing structural H₂O molecules
does not mean they will be
gained in the same order

20-30% glycerol → -20°C
freezer → proteins remain
in solⁿ form;
does not freeze

cryoprotectant

Buffers, salts & additives in protein stock.

✓ if we get mountable
protein crystal in drop

Qualitative test
add Coomassie Blue → proteins take up blue

✓ diffraction spot

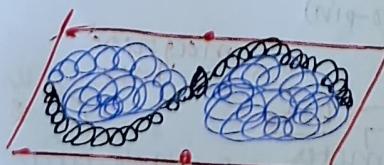
very close → mostly a protein crystal

far away → most likely a salt crystal

blue wmpⁿ
crystals present

by analysis,
we can tell

Crystal packing.



unit cell

repeats in all dirⁿ to produce the crystal.

① Batch Crystallization:

- ② Vapor diffusion → Hanging drop
- ③ " " → Silting "
- ④ Dialysis

Analyzing crystallizⁿ outcomes

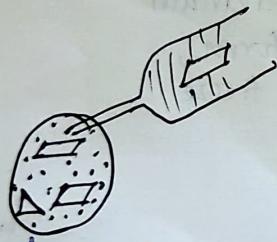
observe under stereomicroscope

polarizer + analyzer

how do we know it is a crystal or some amorphous substance?
on every 90° rot, colour changes! → except cubic crystals
otherwise, $\alpha = \beta = \gamma = 90^\circ$, crystal decays

crystal must always be in mother liquor

Mounting crystals



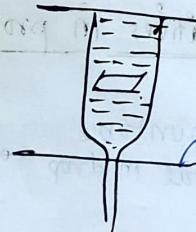
if more than one type of crystal present, diff. to pick out one

Lindemann Glass tube



crystal should not float much

↳ anomaly in diffraction

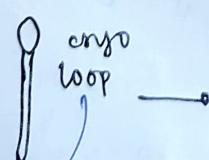


ice cold water tissue paper
↳ absorb all the heat coming

Cryo-Mounting

↳ cryogenically freeze the crystal → Flash frozen

X-ray is fixed



choose according to crystal size

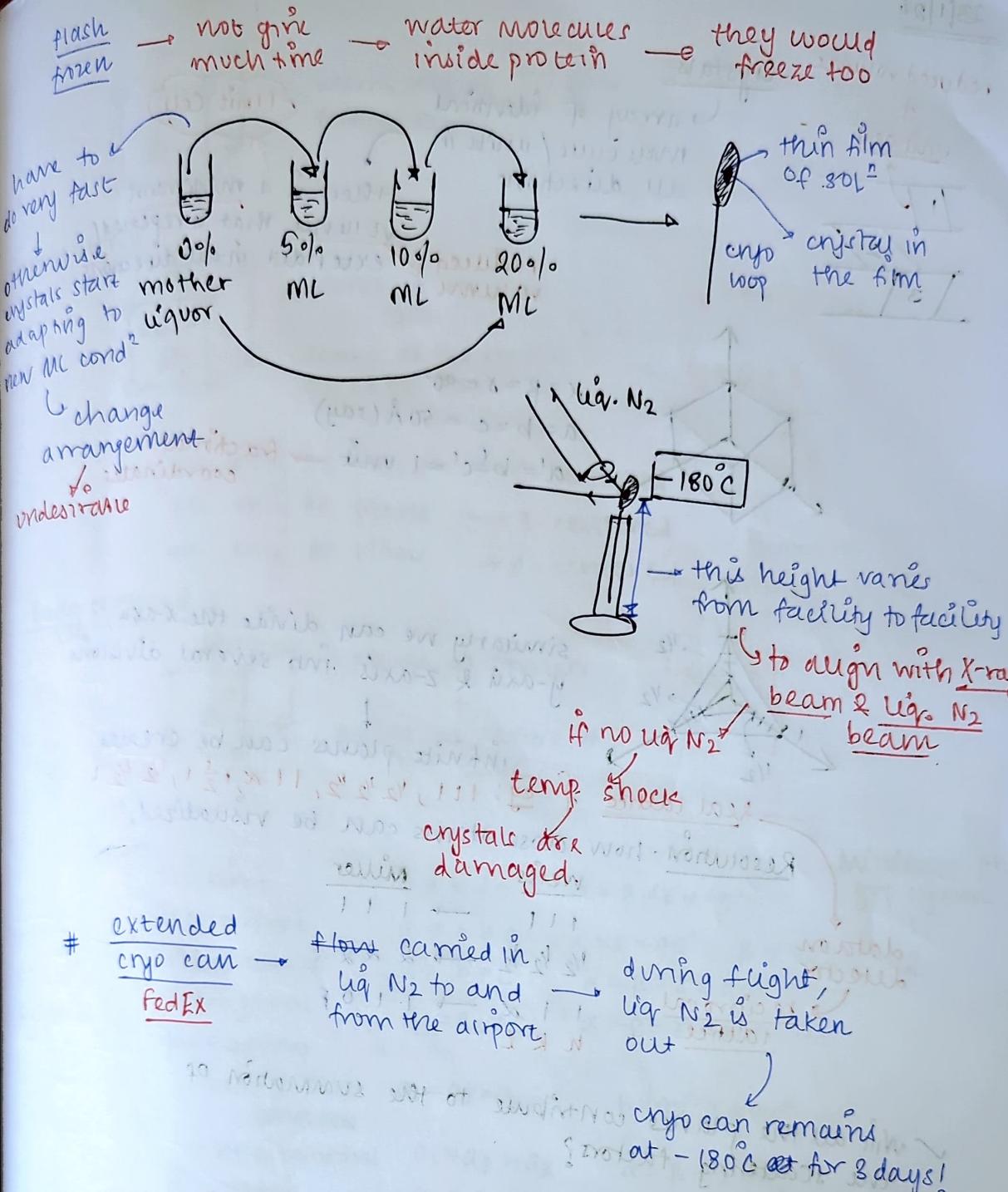


Goniometer

magnetic base

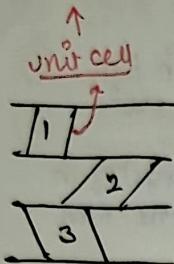
used to rotate & obtain diff. crystal faces

single-circle goniometer



23/1/24

reduced volume



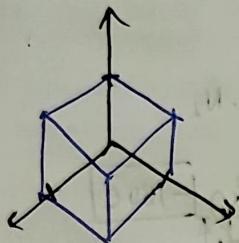
Crystals

array of identical molecules / units in all directions

reduced volume

Unit Cell

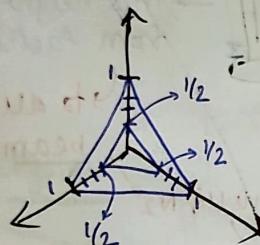
there is a minimum volume that is repeated extended in all directions



$$\alpha = \beta = \gamma = 90^\circ$$

$$a = b = c = 50 \text{ \AA} (\text{say})$$

$$a' = b' = c' = 1 \text{ unit} \rightarrow \text{fractional coordinates}$$



real lattice

Resolution: how close atoms can be visualized,

data on detector

reciprocal lattice

Weiss

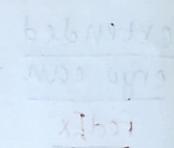
$$111 \rightarrow 111$$

$$\frac{1}{2} \frac{1}{2} \frac{1}{2} \rightarrow 222$$

$$11\alpha \rightarrow 110$$

$$h k L$$

Miller



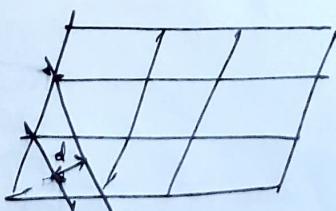
Will all the planes contribute to the summation of the scattering factors?

3 Laws:

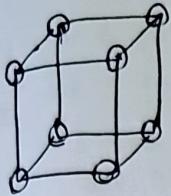
Lau's law \rightarrow mathematical deriv²
 Bragg's law \rightarrow geometrical deriv²
 Ewald's law \rightarrow in spherical coord.
 } but they say the same thing

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

resolution
minimum distance that is measured b/w \parallel^2 planes

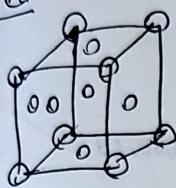


Primitive



atoms only at the corners → each corner is shared by 8 → $\frac{1}{8} \times 8 = 1$ unit cells

face centered



atoms at corners +

atoms at the centre of each face

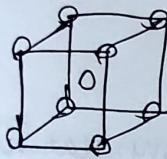
$$\frac{1}{8} \times 8 = 1$$

$$\frac{1}{2} \times 6 = 3$$

$$\underline{\underline{4}}$$

- on only AB plane → C-centered
- on only AC plane → B-centered
- on only BC plane → A-centered

Body centered



atoms at corners +

atoms at centre of the cube

$$\frac{1}{8} \times 8 = 1$$

$$\frac{1}{2} = 1$$

$$\underline{\underline{2}}$$

- + lattice systems
- Triclinic
 - Monoclinic
 - Orthorhombic
 - Tetragonal
 - Hexagonal
 - hexagonal / rhombohedral
 - isometric

$$a \neq b \neq c \quad \alpha \neq \beta \neq \gamma \neq 90^\circ$$

$$a \neq b \neq c \quad \alpha = \gamma = 90^\circ; \beta \neq 90^\circ$$

$$a_1 = a_2 \neq c \quad \alpha = \beta = \gamma = 90^\circ$$

$$a_1 = a_2 = a_3 \neq c \quad \beta = 90^\circ; \gamma = 120^\circ$$

$$a_1 = a_2 = a_3 \quad \alpha = \beta = \gamma \neq 90^\circ$$

$$a_1 = a_2 = a_3 \quad \alpha = \beta = \gamma = 90^\circ$$

Minimum pt. grp symm.

1

2

Bravais lattice = 14 } combining P, FCC, BCC with the 7 systems

Symmetry

① Rotation:

1 fold = 360° → basically no symmetry

crystallographically 2 fold = 180° → 2 molecules

through an axis

generate identical molecule

$$\checkmark 3 \text{ fold} = 120^\circ$$

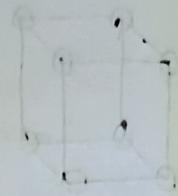
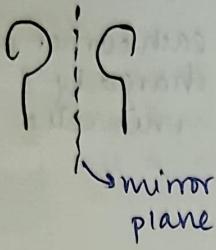
$$\checkmark 4 \text{ fold} = 90^\circ$$

$$5 \text{ fold} = x \rightarrow \text{mathematically not derivable}$$

$$\checkmark 6 \text{ fold} = 60^\circ$$

7 fold X

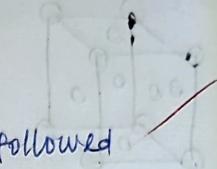
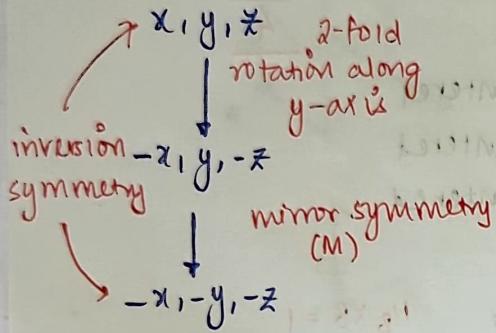
② Mirror:



③ Inversion:

✓ mirror plane \perp to rotⁿ axis

✓ 2-fold rotation followed by mirroring

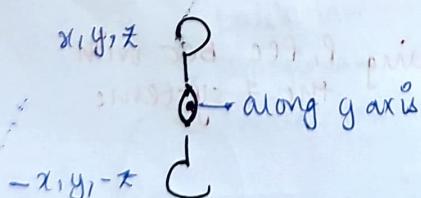
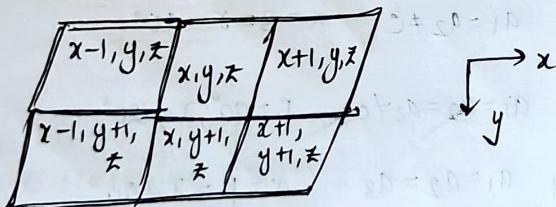


Point groups:

✓ groups consisting of possible combinations of symmetry elements.

✓ 32 in total.

Translation along an axis:



if, $-x, y+\frac{1}{2}, -z$ screw axis \rightarrow translation along the rotⁿ axis

$$120^\circ \text{ rot}^n \rightarrow y + \frac{1}{3}$$

$$120^\circ \text{ rot}^n \rightarrow y + \frac{2}{3}$$

$$120^\circ \text{ rot}^n \rightarrow y + \frac{3}{3} = y + 1$$

21 → 2 fold rotation with $\frac{1}{2}$ translation

31 → 3 fold rotation with $\frac{1}{3}$ translⁿ

32 → 3 fold rotation with $\frac{2}{3}$ translⁿ

33 → 3 fold rotⁿ with no translation: $y; y + \frac{2}{3}; y + \frac{4}{3} = y + 1 + \frac{1}{3}; y + \frac{6}{3} = y + 1 + 1$

so we need 2 unit cells

$$= y + 2$$

to complete a full

symmetric rotation

4

41

42

43

6

61

62

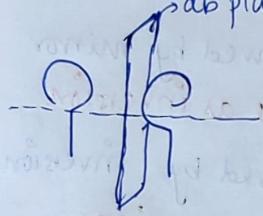
63

64

65

Mirror:

ab plane



Gliding after mirroring, translation by $\frac{1}{2}$

translation along a → a-glide

translation along b → b-glide

along ab face diagonal → n-glide

along diagonal of the unit cell → d-glide



230 space groups.

We check for symmetry to reduce the data points and at the same time understand the structure more easily.

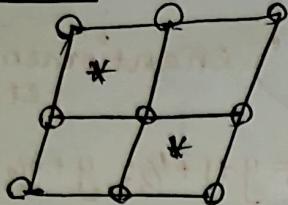
an organic molecule
crystallizes in any one
of the 230 space groups

Proteins → L-amino acids only

- mirror not possible
- inversions not possible

65 possible
space groups
for proteins.

Triclinic



x	0-1/2	0-1	0-1
y	0-1 (or)	0-1/2 (or)	0-1
z	0-1	0-1	0-1/2

half volume of unit cell.

- ✓ application of symmetry elements on one molecule generates other molecule of the unit cell.

primitive $P\bar{I}$ → 1 fold symm. (no symm.) → triclinic system
 primitive $P\bar{I}$ → 1 fold symm. + M (inversion pt. present) → triclinic system → not possible in proteins

generates the other volume ← symmetry element
 proteins show only this

- ✓ thus, half of the crystal/unit cell give information about the entire crystal/unit cell.

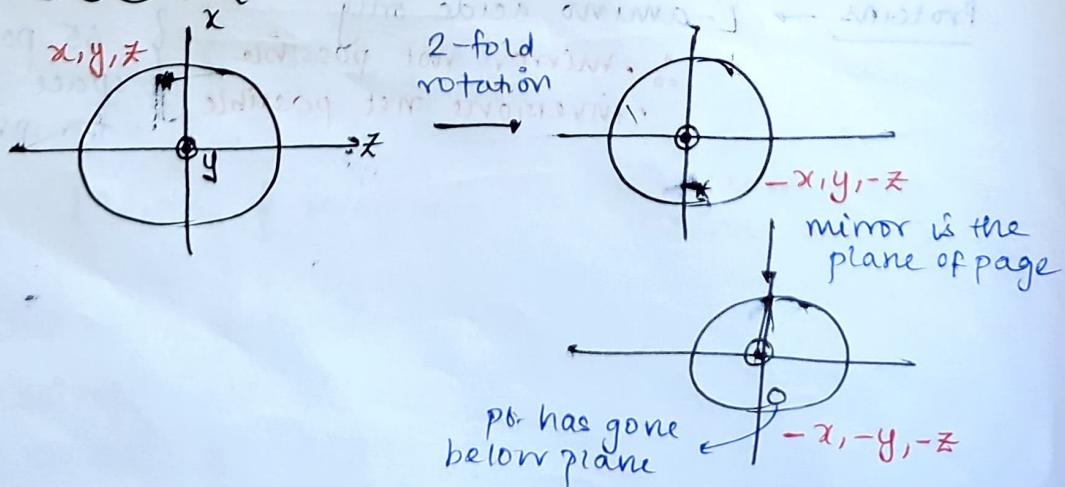
Roto-reflection → rotation followed by minor
 Lo same as inversion

Roto-inversion → rotation followed by inversion
 Lo same as minor

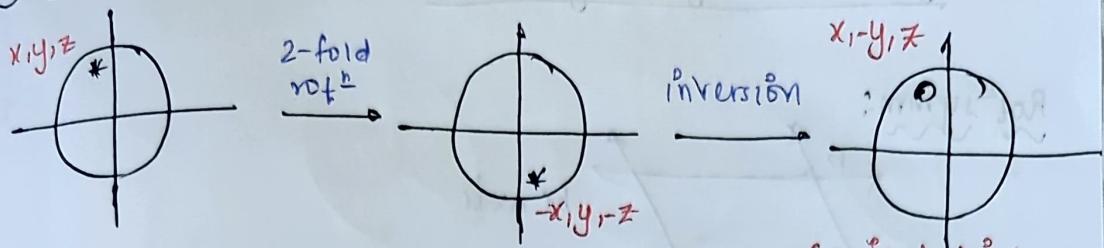


Stereographic Projection

① Roto-reflection $\rightarrow \bar{2}\bar{1}\bar{1}$

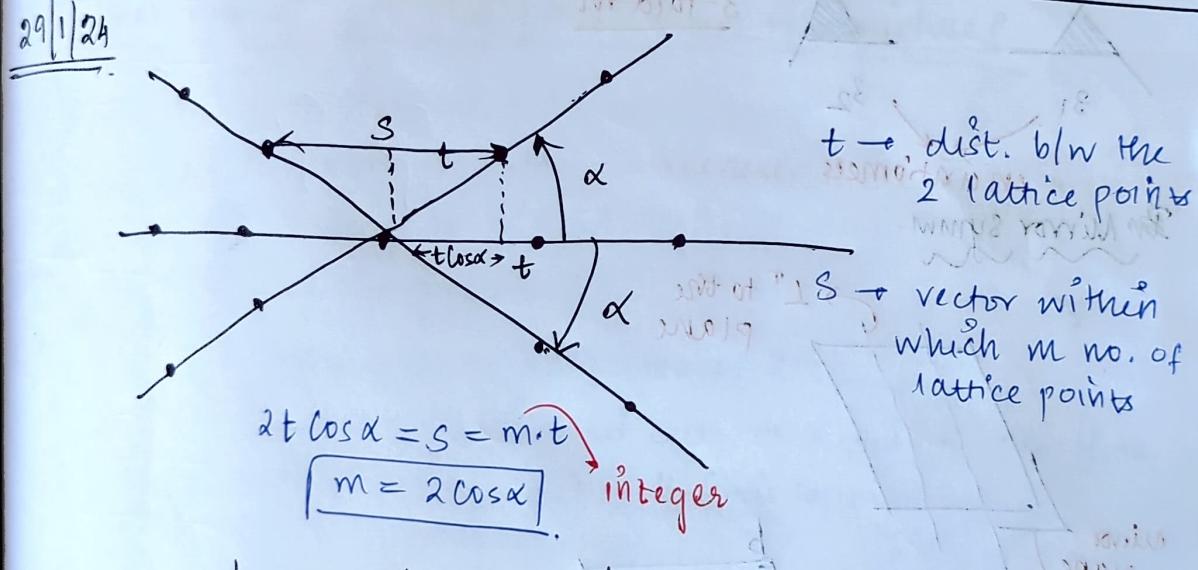


② Roto-inversion: \rightarrow 2M



(as if initial object has been mirrored).

Why is 5-fold symmetry not possible?



s → vector within which m no. of lattice points

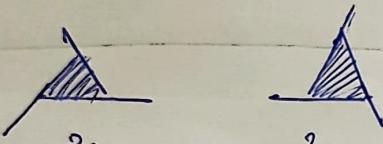
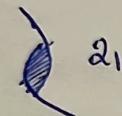
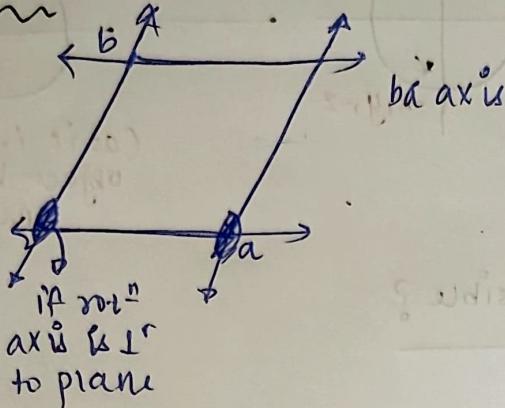
m	$\cos \alpha$	α	axis
2	1	0	2π
1	$1/2$	$\pi/3$	1
0	0	$5\pi/3$	6
-1	$-1/2$	$\pi/2$	4
-2	-1	$2\pi/3$	3
		$-\pi$	2
		$4\pi/3$	

Absence of 5-fold crystal symmetry.

✓ Plane which satisfies Bragg's law, will produce diff're spots.

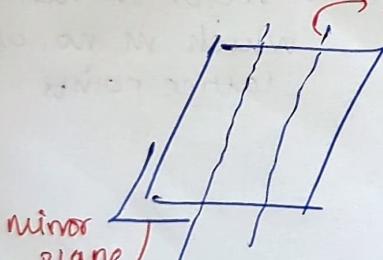
Symbols of Symmetry elements

Rotⁿ symms:



3-fold rotⁿ

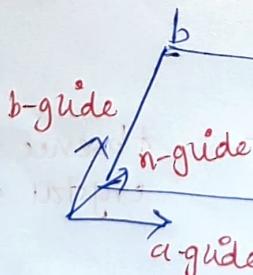
Mirror Symm.



⊥ to the plane

mirror plane parallel to the plane

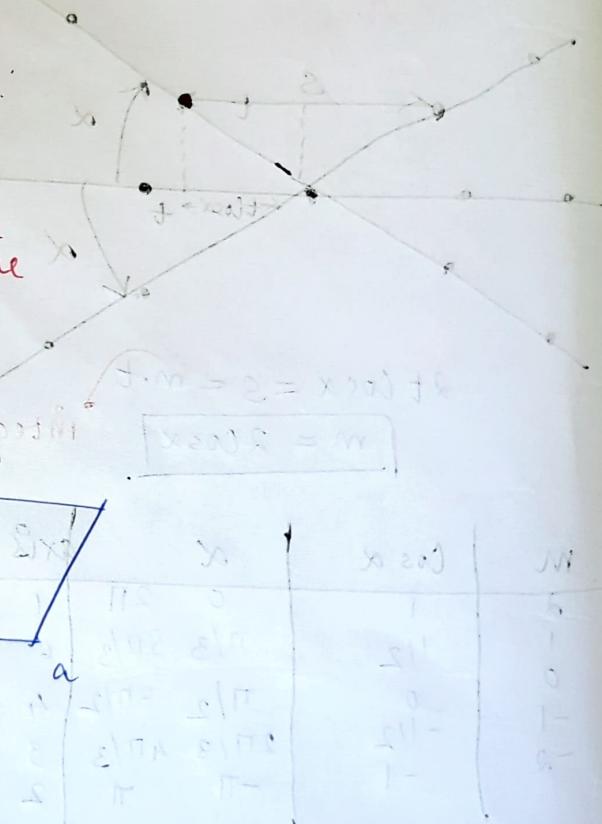
mirror plane parallel to the plane



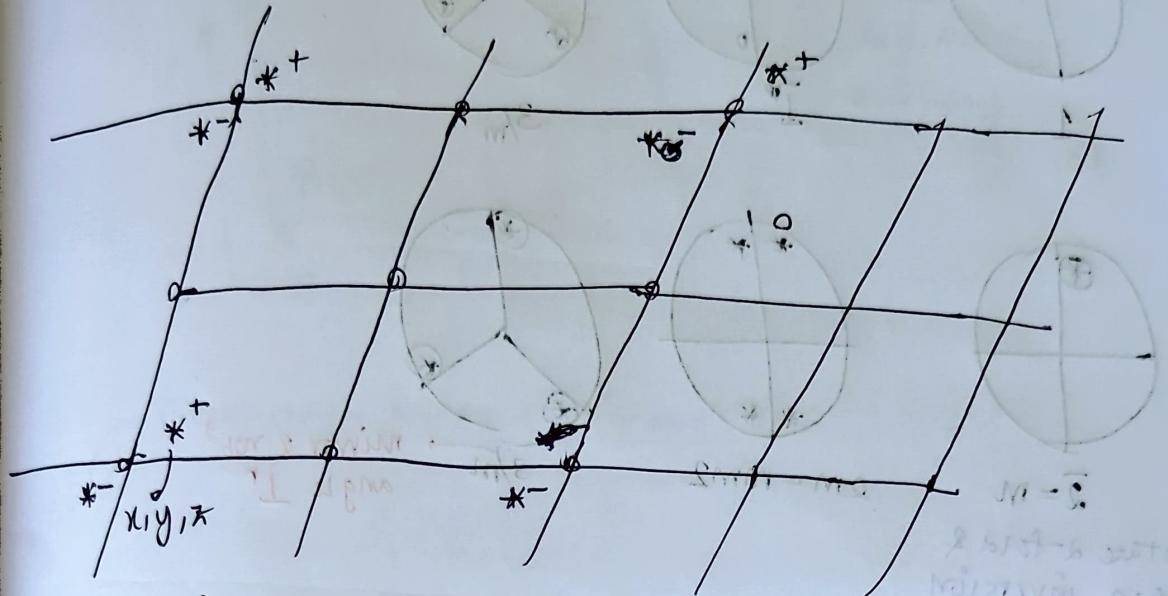
Inversion Symmetry.



inversion line and repeat relation with m₁



$\bar{P}\bar{I}$ → triclinic system with
an inversion point. → 1 fold rotation
followed by mirror.



what information can you extract from this picture?

- ① space group = $\bar{P}\bar{I}$
- ② Centrosymmetric → whenever there is an inversion point.
- ③ belongs to Triclinic system
- ④ point group = \bar{I}
- ⑤ Space group no = 2 (out of 330).
- ⑥ 2 molecules in unit cell → $x, y, z \& -x, -y, -z$
- ⑦ asymmetric unit = $\frac{1}{2}$ (unit cell).
(A.U.)

space group $\bar{P}\bar{I}$:

Triclinic

Pt. Group = 1

No. of molecules = 1

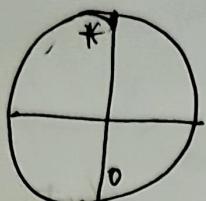
A unitic sp. group

A.U. = full unit cell

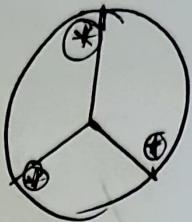
Stereographic Projection:



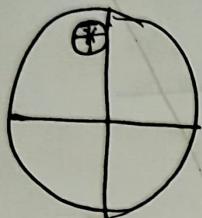
1



2



3/m

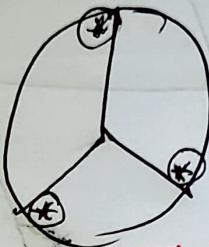


$\bar{2} = m$

rotate 2-fold &
then inversion



$2m = mm2$



$3/\bar{m}$

mirror & rot¹
angle 1°



$\bar{6} = 3/m$

$P_{21/c}$

draw equivalent = in quinq. unq.
symmetries

(N.B. time) $Q^1 = \text{unit interval}$ 2π

(N.A.)

$\bar{1} = \text{quinc. angle}$

$\bar{3} = \text{quinc. angle}$

$\bar{6} = \text{quinc. angle}$

$\bar{1} = \text{quinc. angle}$

$\bar{3} = \text{quinc. angle}$

$\bar{6} = \text{quinc. angle}$

$\bar{1} = \text{quinc. angle}$

$\bar{3} = \text{quinc. angle}$

$\bar{6} = \text{quinc. angle}$

$\bar{1} = \text{quinc. angle}$

$\bar{3} = \text{quinc. angle}$

$\bar{6} = \text{quinc. angle}$

$\bar{1} = \text{quinc. angle}$

$\bar{3} = \text{quinc. angle}$

$\bar{6} = \text{quinc. angle}$

$\bar{1} = \text{quinc. angle}$

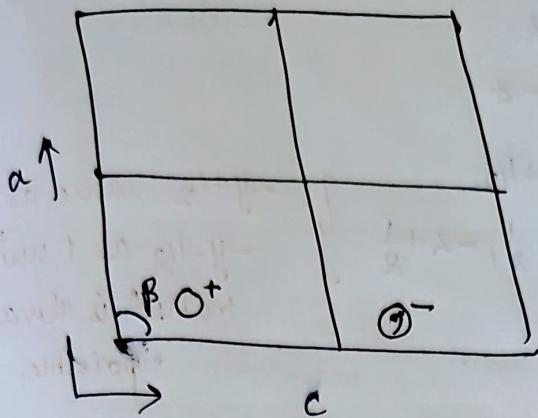
30/1/24

Pc

P1c1

a m

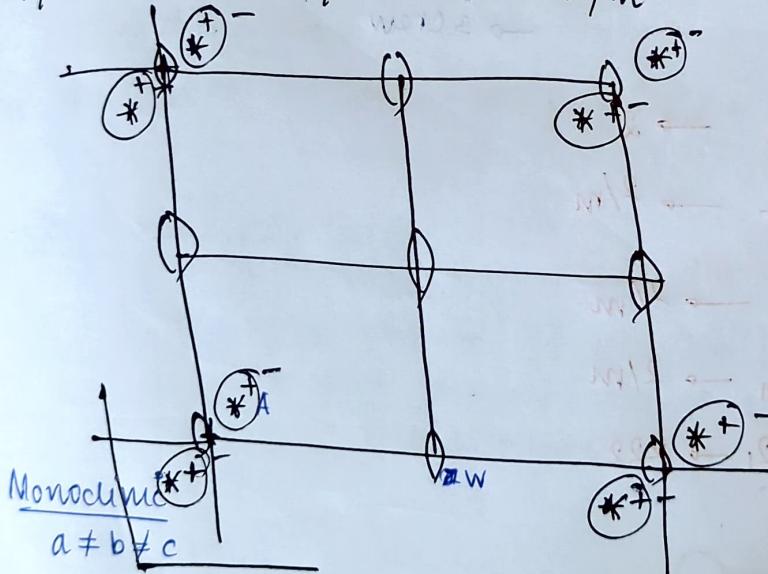
No. 7



- ① x, y, z
- ② $x, \bar{y}, z + \frac{1}{2}$

\curvearrowleft C-centered \rightarrow middle of ab plane
Cc

P2/m P1 2/m 1 2/m No. 10



if we take 2-fold rotⁿ of A & translate by an unit, we automatically generate the w axis.

- m
- ① x, y, z
 - ② $-x, y, -z$
 - ③ $x, -y, z$
 - ④ $-x, -y, -z$

P₂₁ → 2 fold screw axis

Monoclinic

$$\begin{array}{c} x, y, z \\ -x, y + \frac{1}{2}, -z \\ x, -y, z + \frac{1}{2} \\ -x, -y + \frac{1}{2}, -z + \frac{1}{2} \end{array}$$

c

$\underline{\underline{P_2}_1}$

-y + 1/2 same as
-y - 1/2 as 1 unit
transl. is always
possible

$\frac{P_2_1}{a}$ → glide along a axis

rotation axis is $\frac{1}{4}$ th away from the origin along the glide axis.

Mirror plane is $\frac{1}{4}$ th away from the origin in the dir. of the rot. axis

✓ when mentioning point group → do not mention

→ glide

→ screw

e.g.:

P₂₁

→ 2

P₂₁

→ 2/m

P₂₁

→ 2/m

P₂/m

→ 2/m

P₂2₁

→ 222

Matthews Coefficient:

water
30% content < 80%

if a protein is hexameric
& crystallizes as such

the hexamer is considered
as an unit & symmetry
elements apply.

if the hexamer shows 6-fold
crystallographic symmetry, one
monomeric unit is considered
as an unit.

✓ $> 3.4 \text{ \AA}$ - independent atom

✓ $2.5 \text{ \AA} < \text{H-bonding dist.} < 3.4 \text{ \AA}$

✓ $1.2 - 1.6 \text{ \AA}$ - Covalent

Volume of unit cell:

$$V_{\text{cell}} = abc [1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2 \cos\alpha \cos\beta \cos\gamma]^{1/2}$$

When collecting data

↪ since, centrosymmetric
crystal rotated 180° (or) 360° (or) 720°
and data collected / frames collected
every $1/2^\circ$ increment.

Ghost spot

↪ twin crystals
 0.02 \AA deviation

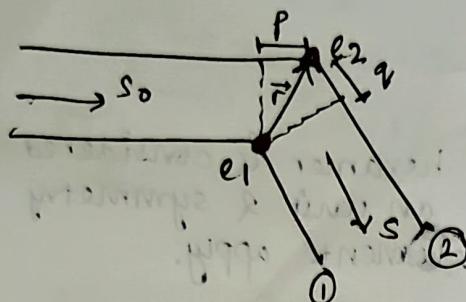
Beam Stopper

if X-ray goes directly
to detector, the detector
might be damaged.

atomic scattering

(f) structure factor

contribution of all the e^- lying on that plane



vectorial form:

of Bragg's Law:

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

The wave along electron e_2 is lagging behind in phase.

Scattering by an atom

$\rho(r) \rightarrow e^-$ density at dist.
 $'r'$ from nucleus

$$f = \int_{-r}^{\infty} \rho(r) e^{(2\pi i \vec{r} \cdot \vec{s})} dr$$

$$f = \left[\rho(r) e^{(2\pi i \vec{r} \cdot \vec{s})} + \rho(r) e^{-2\pi i \vec{r} \cdot \vec{s}} \right] dr$$

$$f = 2 \int \rho(r) \cos[2\pi \vec{r} \cdot \vec{s}] dr$$

high-angle scattering
wide-angle

30°
2-5°

structure
determin

in small angle X-ray scattering
(SAXS)

we can determine
the topology of
the protein

Scattering by a Unit Cell

$$\vec{f}_j = |f_j| e^{(2\pi i \vec{r}_j \cdot \vec{s})}$$

$$\vec{F}(\vec{s}) = \sum_{j=1}^m |f_j| e^{2\pi i \vec{r}_j \cdot \vec{s}}$$

structure factor

depends on the arrangement of atoms in the unit cell.

Scattering by a Crystal

consist of many unit cells → translate in all directions

$$F(\vec{s}) \propto e^{(2\pi i \vec{a} \cdot \vec{s})}$$

$$\begin{aligned} \vec{a} \cdot \vec{s} &= h \\ \vec{b} \cdot \vec{s} &= k \\ \vec{c} \cdot \vec{s} &= l \end{aligned}$$

only then we will get the value of $K(\vec{s})$

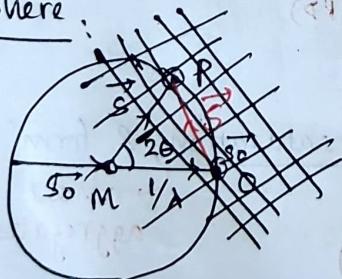
Lau conditions

otherwise zero.

Real lattice → in the crystal → Weiss indices: $a, b, c, \alpha, \beta, \gamma$

Reciprocal lattice → in the data (or) detector → Miller's indices: a^*, b^*, c^* , $\alpha^*, \beta^*, \gamma^*$

Ewald Sphere

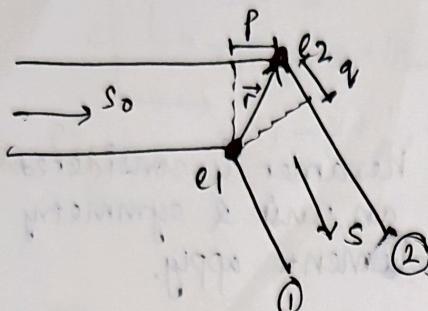


Lau → Bragg → Ewald

(+ wave vector and phase
+ wave intensity)

atomic scattering
f → structure factor

contribution of all the e^- lying on that plane



vectorial form

of Bragg's Law

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

the wave along electron e_2 is lagging behind in phase

Scattering by an atom

$\rho(r) \rightarrow e^-$ density at dist.
'r' from nucleus

$$f = \int_{-r}^r \rho(r) e^{(2\pi i \vec{F} \cdot \vec{s})} dr$$

$$f = \int_{-r}^r \left[\rho(r) e^{(2\pi i \vec{F} \cdot \vec{s})} + i \rho(r) e^{-2\pi i \vec{F} \cdot \vec{s}} \right] dr$$

$$f = 2 \int_{-r}^r \rho(r) \cos [2\pi \vec{F} \cdot \vec{s}] dr$$

high-angle/ scattering
wide-angle

30°
2-5°

structure
determin.

Small angle X-ray scattering
(SAXS)

→ we can determine
the topology of
the protein

Scattering by a Unit Cell:

$$\vec{f}_j = |f_j| e^{(2\pi i \vec{r}_j \cdot \vec{s})}$$

$$\vec{F}(\vec{s}) = \sum_{j=1}^n |f_j| e^{(2\pi i \vec{r}_j \cdot \vec{s})}$$

structure factor

depends on the arrangement of atoms in the unit cell.

Scattering by a Crystal:

consist of many unit cells → translate in all directions

$$F(\vec{s}) \propto e^{(2\pi i \vec{a} \cdot \vec{s})}$$

$$\left. \begin{array}{l} \vec{a} \cdot \vec{s} = h \\ \vec{b} \cdot \vec{s} = k \\ \vec{c} \cdot \vec{s} = l \end{array} \right\} \rightarrow \text{only then we will get the value of } F(\vec{s})$$

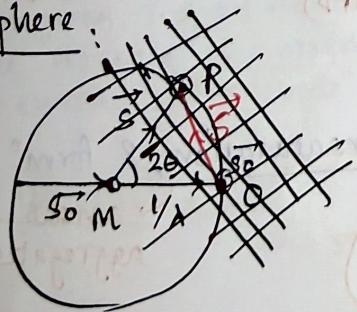
Lau conditions

otherwise zero.

Real lattice → in the crystal → Weiss indices: $a, b, c, \alpha, \beta, \gamma$

Reciprocal lattice → in the data (or) detector → Miller's indices: a^*, b^*, c^* , $\alpha^*, \beta^*, \gamma^*$

Ewald Sphere:



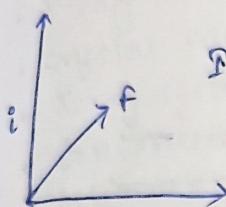
Lau → Bragg → Ewald

① $\frac{121}{a}$ → derive the coordinates.

②

6[2/24]

$$F = a + ib$$



$$S(xyz) = \frac{1}{\sqrt{hkl}} \sum_{hkl} |F(hkl)| \exp[-2\pi i(hx+ky+lz) + i\alpha(hkl)]$$

conjugate

$$\begin{aligned} I &= F \times F^* \\ &= (a+ib)(a-ib) \\ &= a^2+b^2 \end{aligned}$$

structure factor
amplitude

final eqn. to calculate e^- density from observed structure factor.

no yet known

have to be determined ← lost during crystallization [Phase]

Phase determination:

small molecules → statistical method of phase determⁿ

①

Direct method

- only applicable for:
1) high resolⁿ datasets
2) small molecules
eg: insulin → 5.6 kDa

②

Heavy Atom Derivatives

→ Isomorphous replacement

we have to put some heavy atoms in the crystal but the dimensions of crystals must not change after soaking in heavy atoms

salts, oxides, hydroxides

Hg, Sm, Pb, I₂, U, W, Pt

in reality → some changes do take place

- dataset
- i) w/o heavy atom soaking
 - ii) with heavy atom soaking
- compare the change

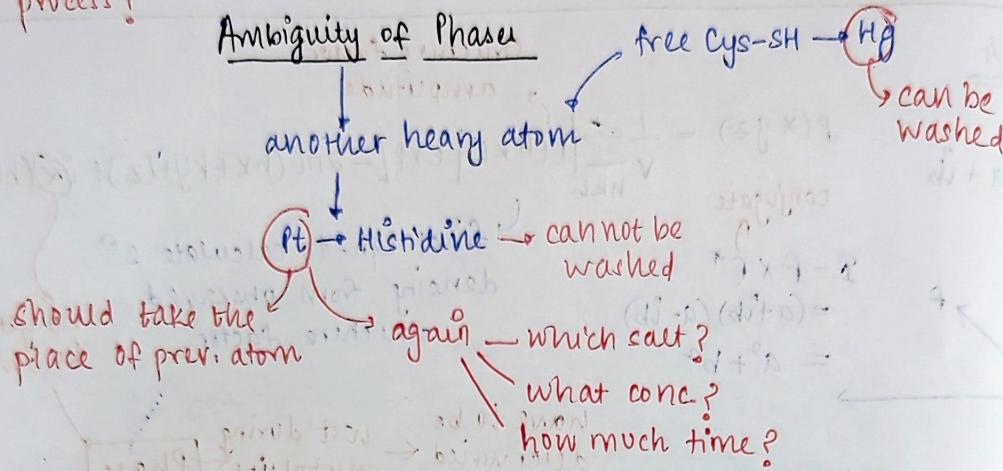
Drawbacks:

- It is not guaranteed that heavy atoms will enter the crystal.
It depends on the space & environmental conditions
- If free Cys present → SH → Hg has affinity
- It is heavily trial-and-error. We do not know which metal to chose.



All are
rat-and-
error
+ science
random
process!

- Which derivative of heavy atom (salt/oxide/hydroxide) should crystal be soaked in?
- What concentration?
- How much time → optimise → but crystal size will vary.



Tyr → derivatize with I_2

Cys/S → replaced by Se
Met → but selenocysteine/selenomethionine are very toxic.

✓ every heavy atom → anomalous character → at certain wavelengths.

H.A.-derivative:

- ① SIR single isomorphous replacement + AS = SIRAS
- ② MIR multiple isomorphous replacement + AS = MIRAS
- ③ MAD multi-anomalous dispersion

Anomalous Scattering (AS)

$4 \times 360^\circ$ → multiple readings of the same sample
 ≈ 30 min time

single wavelength wavelength at which AS occurs

✓ IR method is troublesome & uncertain!

② Molecular replacement

If we know similar to protein this knowledge a structure of our interest can be used to get phase data of new molecule.

- scatt. by e^-
 scatt. by atom
 scatt. by unit cell
 scatt. by crystal
 structure factor defⁿ = unit cell }
 e^- density from structure factor } crystal
 phase problem
 methods of phase determ.¹

in terms of summⁿ of atomic factors.

Lau's cond².

- synchronization vs in-house source
 X-ray source
 crystal mounting
 X-ray system diagram
 Gonionometer head
 Gonionometer colimator

\vec{F}_P → steric factor for native protein

\vec{F}_{PH} → steric factor for native protein
soaked in heavy atom

$$\vec{F}_{PH} = \vec{F}_P + \vec{F}_H \rightarrow \text{true for centrosymm. space group}$$

$$F_P + F_H$$

$$F_P - F_H$$

$$\vec{F}_{PH_1} = \vec{F}_P + \vec{F}_{H1}$$

$$\vec{F}_{PH_2} = \vec{F}_P + \vec{F}_{H2}$$

① Determine the heavy atom positions

Patterson function → determining heavy atom position

Intensity (I)

Difference Patterson:

subtract native prot. from the soaked prot. → we get contribⁿ of only heavy atoms

phase is not important → as works on intensities

P21

Harker planes Harker lines

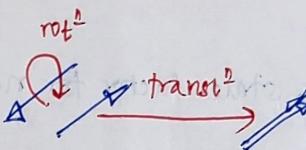
	x y z	-x, y + 1/2, -z
x y z	0 0 0	$\frac{2x}{u}, \frac{1}{2}, \frac{2z}{w}$
-x, y + 1/2, -z	-2x, $\frac{1}{2}$, -2z	0 0 0

$$\begin{array}{c|c|c} u & v & w \\ \hline 2x = 0.32 & 0.5 & 2z = 0.58 \end{array} \quad \left. \begin{array}{l} x = 0.16 \\ z = 0.29 \end{array} \right\} \text{now we can find } y \text{ at any pos.}$$

② Molecular replacement

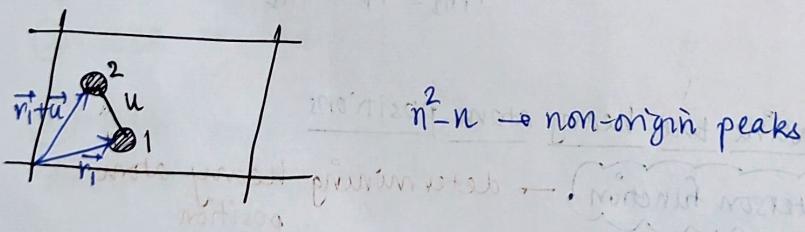
① Translational

② Rotational



Patterson funcⁿ works on intensity
centrosymm. funcⁿ

$$P(\vec{u}) = \int p(\vec{r}_1) \times p(\vec{r}_1 + \vec{u}) dV. \quad \begin{array}{l} \text{position } \vec{r}_1, \\ \text{translated by } \vec{u}. \end{array}$$



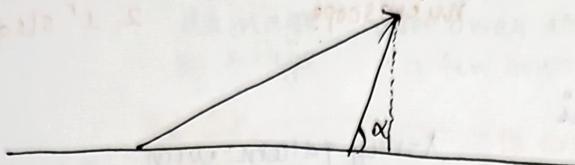
Wilson plot → related to the quality of the data

Effect of heavy atoms on X-ray intensities

$$\vec{F}_{PH} = \vec{f}_P + \vec{F}_H$$

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

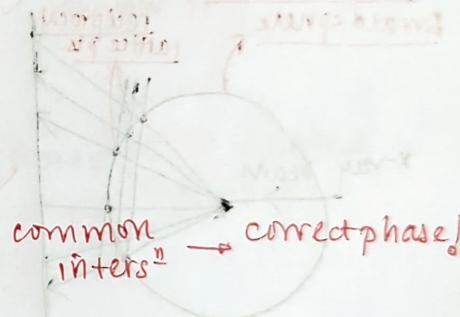
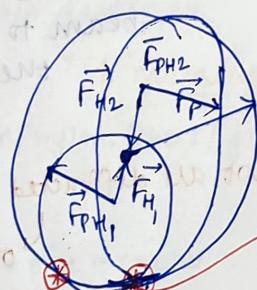
$\alpha_{PH} - \alpha_H$ } ambiguity of phases.
 $\alpha_{PH} + \alpha_H$



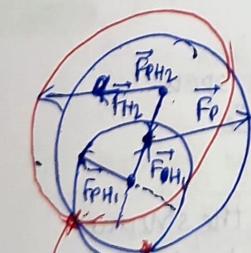
How to remove ambiguity of phases :

Harker Phase Diagram

for each reflⁿ → draw phase circle $\frac{\vec{F}_P}{\vec{F}_{PH}}$ } intersect/cut at two points



1 right 1 wrong.



Friedel (or) Bijvoet (pr: Bifet) pair

$F(h k l) \neq F(\bar{h} \bar{k} \bar{l})$ → if there is anomalous scattering.

at anomalous λ : f^+ f^-

$$\Delta F = F^+ - F^-$$

$f^+ = f^-$ for C, N, O

for heavy atoms → stronger dipole → $f^+ \neq f^-$ diff.