IISc-Alumni Association- M.Vijayan Lecture.

Cross Fertilisation: How Crystallography has built on Mathematics.

Eleanor Dodson
17th August
2017

Nostalgia time:

I cannot resist revisiting old photographs!

Oxford 1968?

Wedding Day – July 1969

DCH building first insulin model

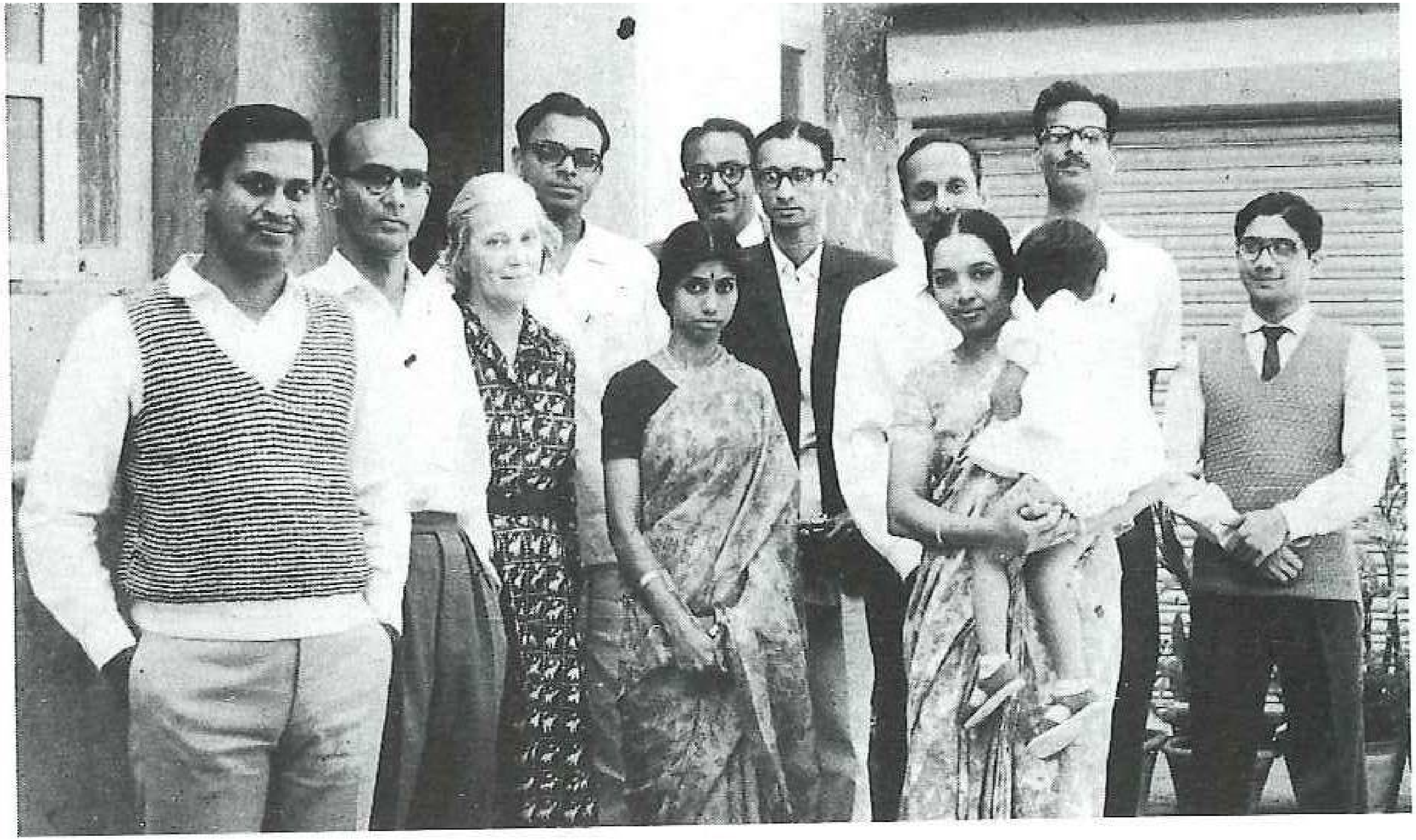
Vijayan lost in shadows?

Oxford lab - 1969 Britishx5, Indian, NZ, Aust.

David Phillips, Dorothy, Tony North, Thomas, Vijaya, Tom Blundell, Ted Baker, EJD

Bangalore – 1970?

Bangalore 1970s



Bangalore, 1970s: Vijayan and Kalyani are to the right of Dorothy; Siv Ramaseshan stands behind his wife Kausalya.

(Courtesy of the Hodgkin family)

Kalyani, Vijayan, Tom Blundell -1975?

Bangalore – 1979-80. Vijanan Davi Kalyani

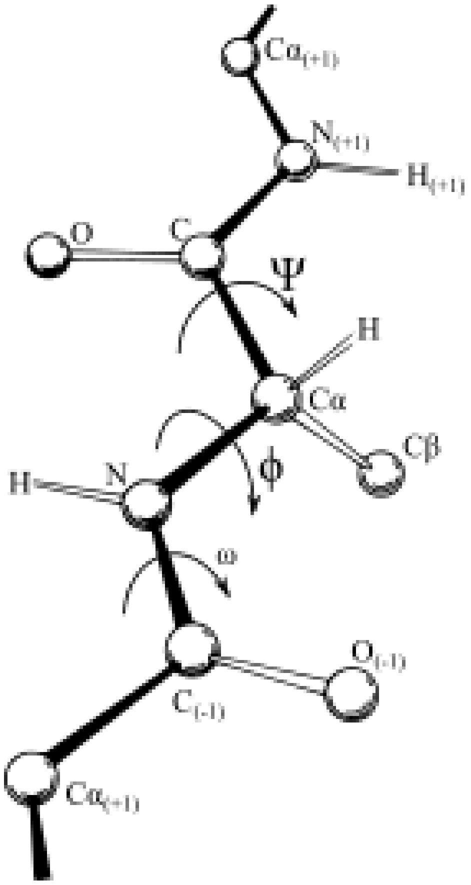
Bangalore 2007: Guy's 70th Birthday Cake

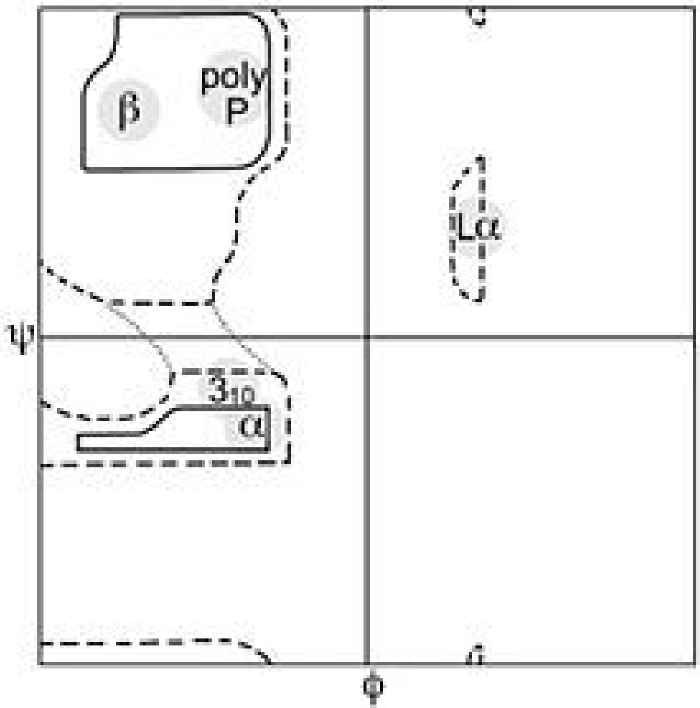
Workshop 2008

DCH & Siv Ramasechan – 1965
our Anomalous Dispersion Guru

First anomalous data measurements done in Oxford ~ 1960 on B12 derivative by K. Venkatesan

Validation:
More Thought Experiments:
G.N. Ramachandran





Mathematics in the Service of Science?

Michael Mosley -BBC 2010

"Knowledge leaps forward when brilliant experiments are analysed by independent minds."

He was referring to the discovery that Mars moves in an elliptical orbit around the sun – established by Johannes Kepler -1599 using the observations of Tycho Brahe.

Mathematics in the Service of Crystallography?

Structure Solution exploits:

Fundamental Mathematical concepts

(Some elegant, but restricted to 3D for structure, and 2D for diffraction)

Technical Mathematics and Computation

(messy and very *easy to get wrong!)*

Statistics

(Much needed but messy too)

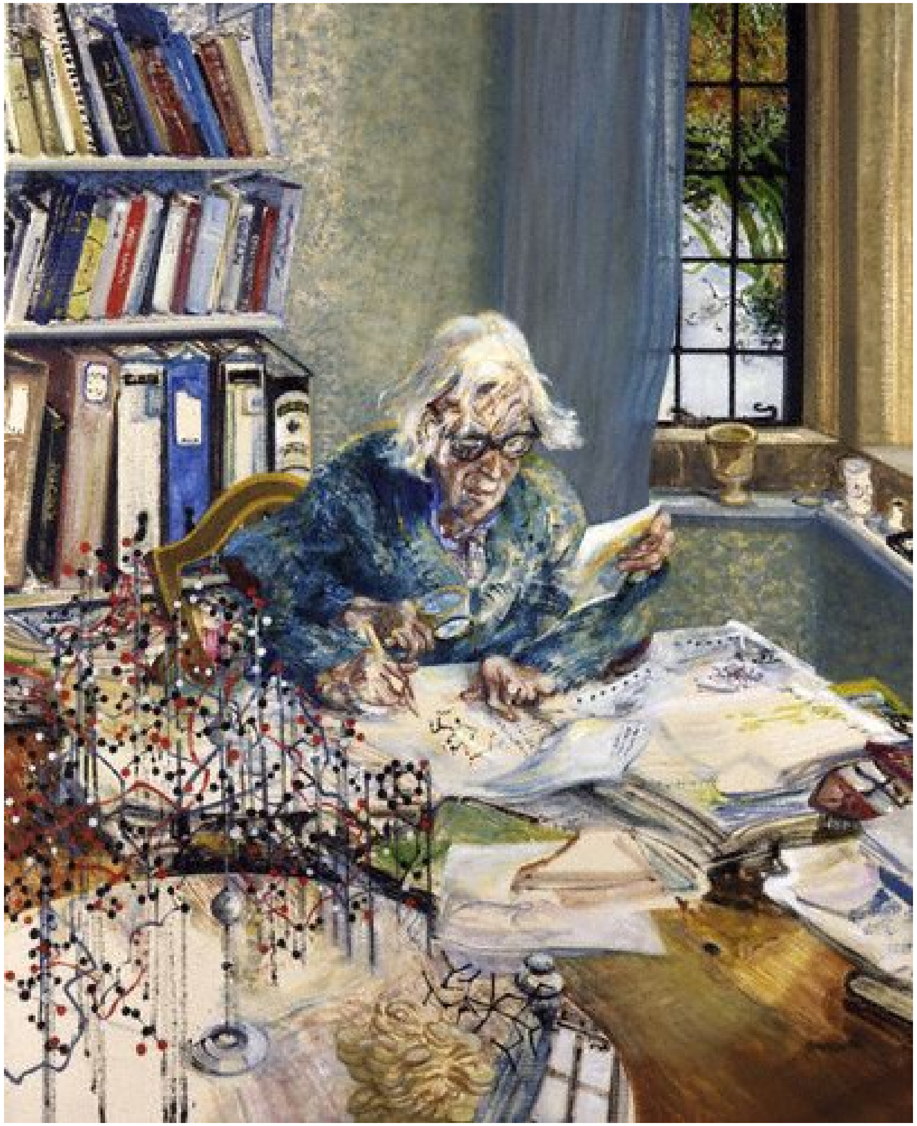
Why crystallography?

Quote: DCH:
I also first learnt at the same time about biochemistry which provided me with the molecules it seemed most desirable to 'see'.

…[The] great advantage of X-ray analysis as a method of chemical structure analysis is its power to show some totally unexpected and surprising structure and to do so with absolute certainty…. (Nobel Lecture 1965)

On many occasions Dorothy after a satisfying examination of a map would exclaim “There is nothing like seeing the atoms!”

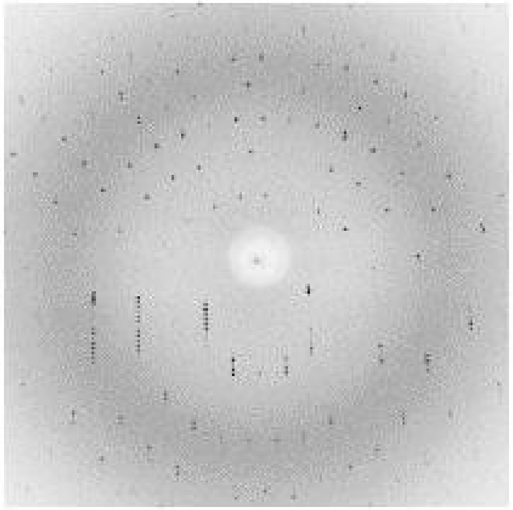
Better Title: The Joys of "Seeing" and How to achieve it.



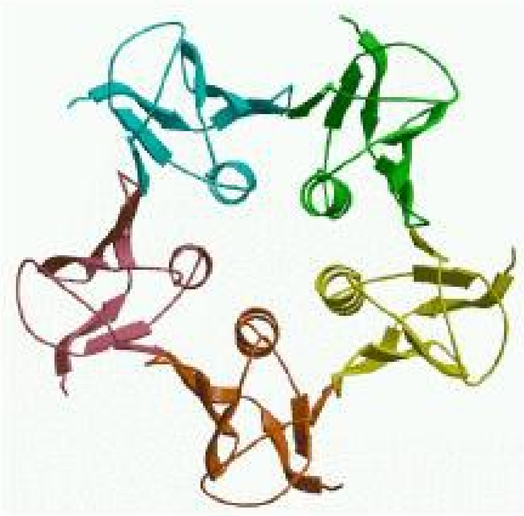
Crucial steps in X-ray structure solution

Crystal

Diffraction



Interpretation



Bayes Theorem (1763)

This underlies all scientific thinking

$P(x \mid O) = P(x)\, P(O \mid x) \,/\, P(O)$
$P(x)$ – *prior* probability of a model parameter x.
$P(O)$ – probability of a observation O.

Restate: The best model is the one with the highest probability, given a set of observations and some prior knowledge of a likely model.

P(Model;Obsns) =
P(Model) P(Obsns;Model) / P(Obsns)

How to express those probabilities?

Part 1: Early crystallography

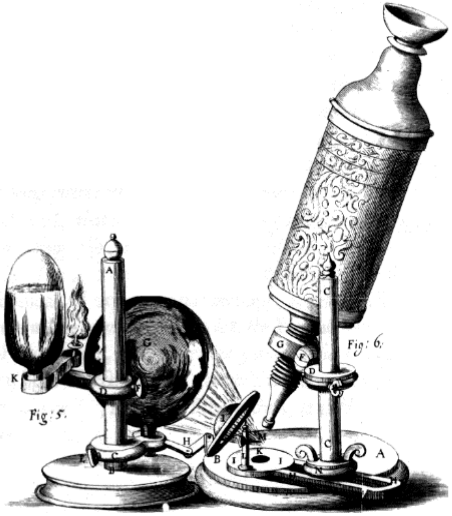
All before any diffraction observed

Pure Thought Experiments..

Morphology:

Based on observing crystals using microscopes

Scientists entranced by their beauty.
Hooke's microscope ~ 1700



Some Crystallographic Time lines (1)

Johannes Kepler (1571-1630)
Kepler's essay, "On the Six-Cornered Snowflake," is considered to be the first work that analyzes and describes the structure of crystal symmetry.

Rene Hauy (1743-1826)
The geometrical law of crystallization associated with his name.

Some Crystallographic Time lines(2)

Christian Weiss (1780-1856) Crystal systems

He stressed the significance of direction in crystals, considering crystallographic axes to be a possible basis for classification of crystals. He is credited for introducing the categorization schema of crystal systems, and has a basic law of crystallography named after him called the "Weiss zone law".

Some Crystallographic Time lines (3)

William Miller (1801-1880) – crystal faces
Miller indices are named after William Hallowes Miller, the method having been described in his Treatise on Crystallography (1839).

Deductions from Observations and measurements of Crystal faces & normals

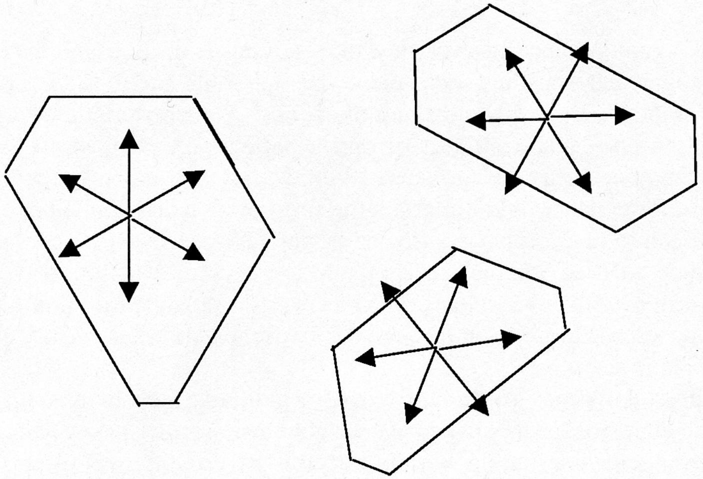


Figure 15.2 Three platy crystals, all showing a basic hexagonal symmetry.

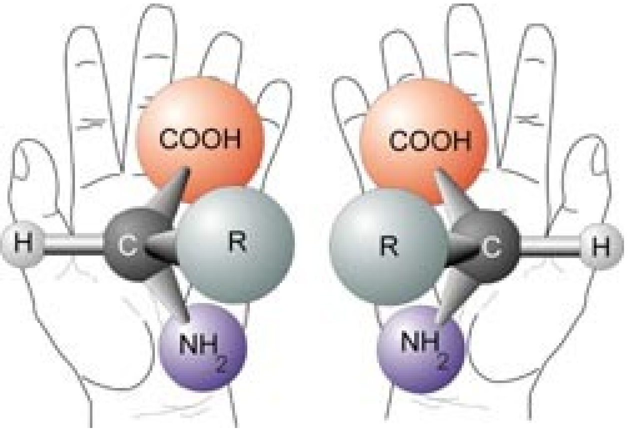
The normals intersect in conserved and simple angles for all crystal fragments. There must be a repeating lattice. (Miller)

A triumph
Pasteur 1853

Chemistry derived from morphology

Mirrored crystals had different properties

Figure 3 Large crystals of sodium ammonium tartrate prepared by the seeding method. Left, (–)-enantiomer; right, (+)-enantiomer: the pictures are taken for the crystals prepared about 25 years ago and stored as described in the text.



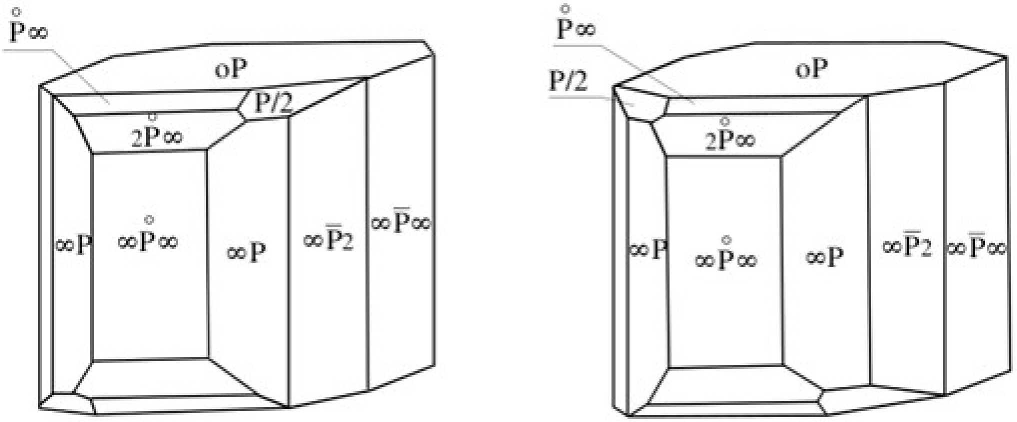


Figure 1 Sketches of the hemihedral crystals of sodium ammonium tartrate according to Kekulé.

Citation by the Royal Society on the Occasion of the Awarding of the Rumford Medal to Louis Pasteur December 1857:

The President, Lord Wrottesley:

The Rumford Medal has been awarded to M. Pasteur for his discovery of the nature of tartaric and para-tartaric acid, and their relation to polarized light.

Chemists had long been acquainted with a peculiar acid, .. which had the same composition as tartaric acid, .. and resembled its properties in a very remarkable manner, yet the two acids were not identical,

In a careful scrutiny of their crystalline forms M. Pasteur recognized the almost universal presence of hemihedral faces, of such a character that the two hemihedral forms were "dissymmetric ", that is, could not be superposed on each other, but each could be superposed on the image of the other in a mirror.

Some Crystallographic Time lines (4)

Auguste Bravais (1811-1863) – the conception of Bravais lattices, and the formulation of Bravais law. (1850)

32 Crystal Classes

Deduced: 7 Crystal systems

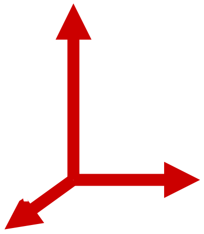
System		Angles	Axes
Triclinic		$\alpha \neq \beta \neq \gamma \neq 90°$	$a \neq b \neq c$
Monoclinic		$\alpha = \gamma = 90° \neq \beta$	$a \neq b \neq c$
Orthorhombic		$\alpha = \beta = \gamma = 90°$	$a \neq b \neq c$
Tetragonal		$\alpha = \beta = \gamma = 90°$	$a = b \neq c$
Hexagonal		$\alpha = \beta = 90°, \gamma = 120°$	$a = b \neq c$
Rhombohedral	(R)	$\alpha = \beta = \gamma$	$a = b = c$
	(H)	$\alpha = \beta = 90°, \gamma = 120°$	$a = b \neq c$
Cubic		$\alpha = \beta = \gamma = 90°$	$a = b = c$

Due to **symmetry requirements** some unit cells may not be primitive:

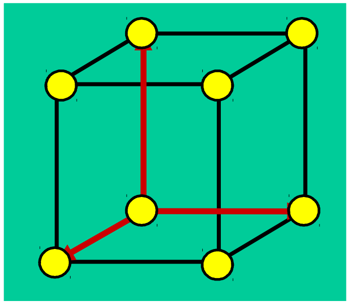
Only 14 different combinations of ***a***, ***b***, ***c*** and α, β, γ can exist = **14 Bravais' lattices**

Therefore we can have:

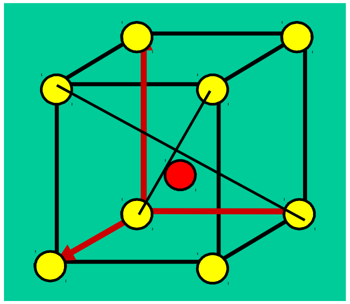
- **P - primitive**
- **I – body centred**
- **A, B, C – face centred**
- **F – all-face centred unit cells**



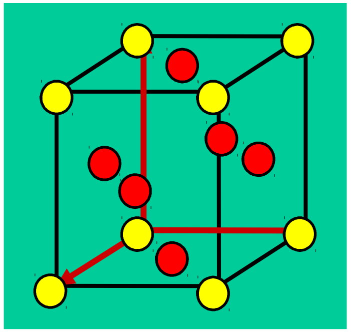
P



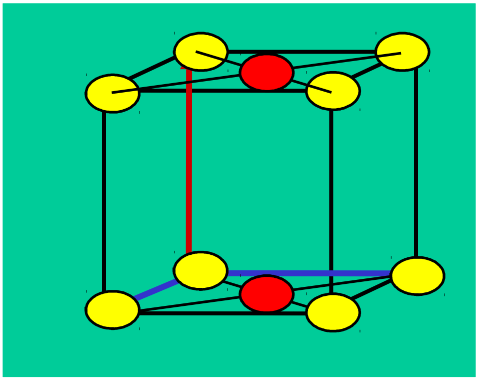
I



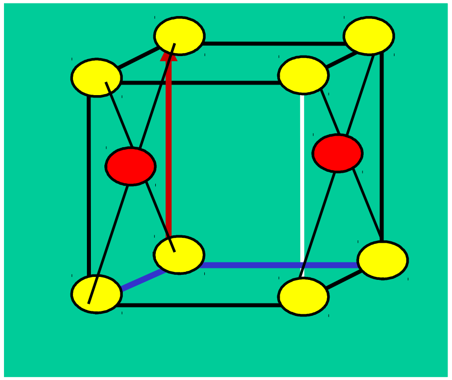
F



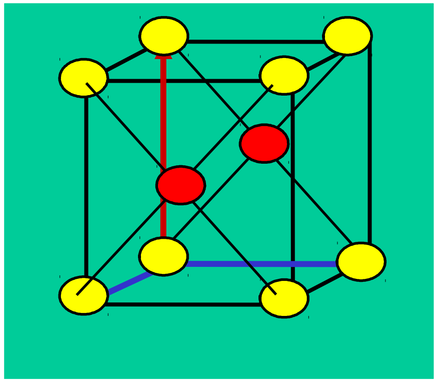
C



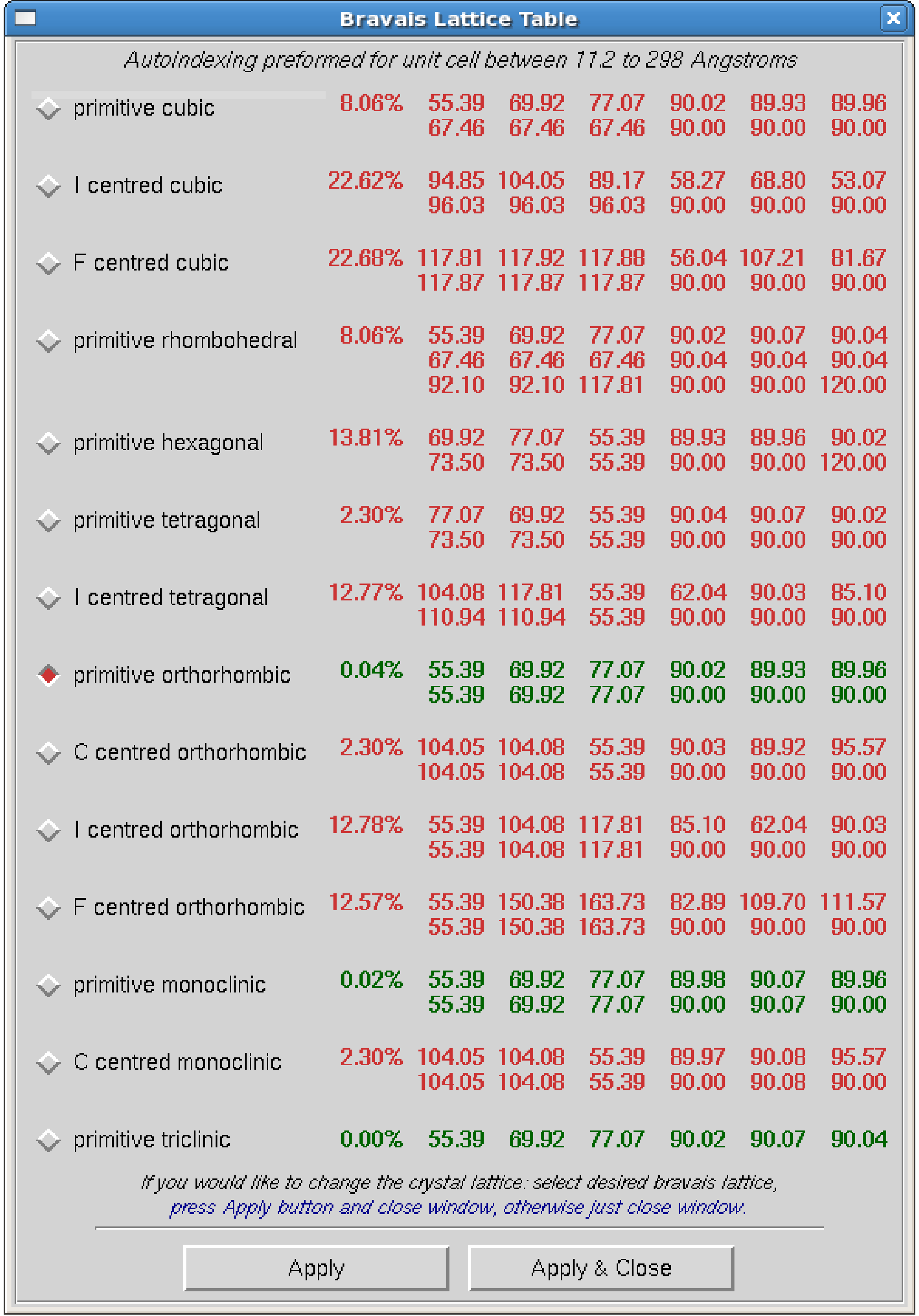
B



A



Integration software tabulates these



Autoindexing preformed for unit cell between 11.2 to 298 Angstroms

◇	primitive cubic	8.06%	55.39 67.46	69.92 67.46	77.07 67.46	90.02 90.00	89.93 90.00	89.96 90.00
◇	I centred cubic	22.62%	94.85 96.03	104.05 96.03	89.17 96.03	58.27 90.00	68.80 90.00	53.07 90.00
◇	F centred cubic	22.68%	117.81 117.87	117.92 117.87	117.88 117.87	56.04 90.00	107.21 90.00	81.67 90.00
◇	primitive rhombohedral	8.06%	55.39 67.46 92.10	69.92 67.46 92.10	77.07 67.46 117.81	90.02 90.04 90.00	90.07 90.04 90.00	90.04 90.04 120.00
◇	primitive hexagonal	13.81%	69.92 73.50	77.07 73.50	55.39 55.39	89.93 90.00	89.96 90.00	90.02 120.00
◇	primitive tetragonal	2.30%	77.07 73.50	69.92 73.50	55.39 55.39	90.04 90.00	90.07 90.00	90.02 90.00
◇	I centred tetragonal	12.77%	104.08 110.94	117.81 110.94	55.39 55.39	62.04 90.00	90.03 90.00	85.10 90.00
◆	primitive orthorhombic	0.04%	55.39 55.39	69.92 69.92	77.07 77.07	90.02 90.00	89.93 90.00	89.96 90.00
◇	C centred orthorhombic	2.30%	104.05 104.05	104.08 104.08	55.39 55.39	90.03 90.00	89.92 90.00	95.57 90.00
◇	I centred orthorhombic	12.78%	55.39 55.39	104.08 104.08	117.81 117.81	85.10 90.00	62.04 90.00	90.03 90.00
◇	F centred orthorhombic	12.57%	55.39 55.39	150.38 150.38	163.73 163.73	82.89 90.00	109.70 90.00	111.57 90.00
◇	primitive monoclinic	0.02%	55.39 55.39	69.92 69.92	77.07 77.07	89.98 90.00	90.07 90.07	89.96 90.00
◇	C centred monoclinic	2.30%	104.05 104.05	104.08 104.08	55.39 55.39	89.97 90.00	90.08 90.08	95.57 90.00
◇	primitive triclinic	0.00%	55.39	69.92	77.07	90.02	90.07	90.04

If you would like to change the crystal lattice: select desired bravais lattice,
press Apply button and close window, otherwise just close window.

Some Crystallographic Time lines (5)

Fyodorov (1853-1919) His best-known result is his derivation of the 230 symmetry space groups. (1891)

Schoenflies (1853 –1928), a German mathematician known for his contributions to the application of group theory to crystallography.

Only 230 Spacegroups are possible

All operators must generate a repeating lattice with straight edges.

This allows:
Rotation Operators 2 3 4 or 6
Mirrored Structures
Screw axes and Glide planes

Derivation is an example of Finite Groups.

Bernal's MSc thesis to redo this!

Systemised by Kathleen Lonsdale
Int. Tab. vol A: spacegroup $P3_1$

Coordinates of the general position:

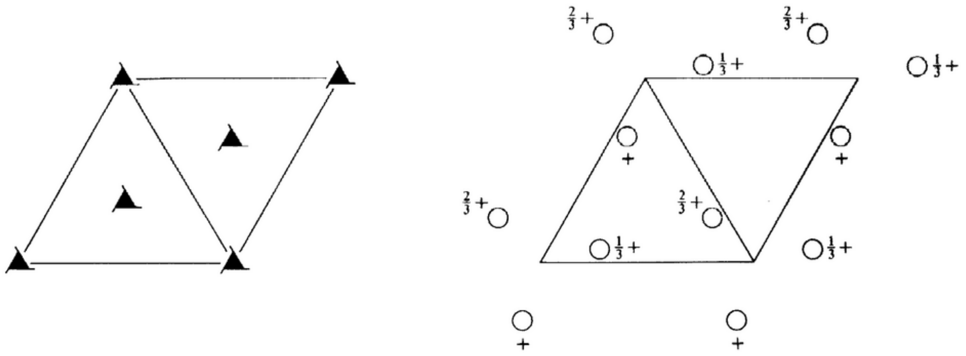
m=1: x, y, z

m=2: -y, x-y, z+1/3

m=3: -x+y, -x, z+2/3

$P3_1$ C_3^2 3 Trigonal

No. 144 $P3_1$ Patterson symmetry $P\bar{3}$



Origin on 3_1

Asymmetric unit $0 \le x \le 1$; $0 \le y \le 1$; $0 \le z \le \frac{1}{3}$

Vertices $0,0,0$ $1,0,0$ $1,1,0$ $0,1,0$

$0,0,\frac{1}{3}$ $1,0,\frac{1}{3}$ $1,1,\frac{1}{3}$ $0,1,\frac{1}{3}$

Symmetry generators

Symmetry operations

(1) 1 (2) $3^+(0,0,\frac{1}{3})$ $0,0,z$ (3) $3^-(0,0,\frac{2}{3})$ $0,0,z$

Generators selected (1); $t(1,0,0)$; $t(0,1,0)$; $t(0,0,1)$; (2)

Positions

Multiplicity, Wyckoff letter, Site symmetry			Coordinates		
3	*a*	1	(1) x,y,z	(2) $\bar{y},x-y,z+\frac{1}{3}$	(3) $\bar{x}+y,\bar{x},z+\frac{2}{3}$

Systematic absences:

Reflection conditions

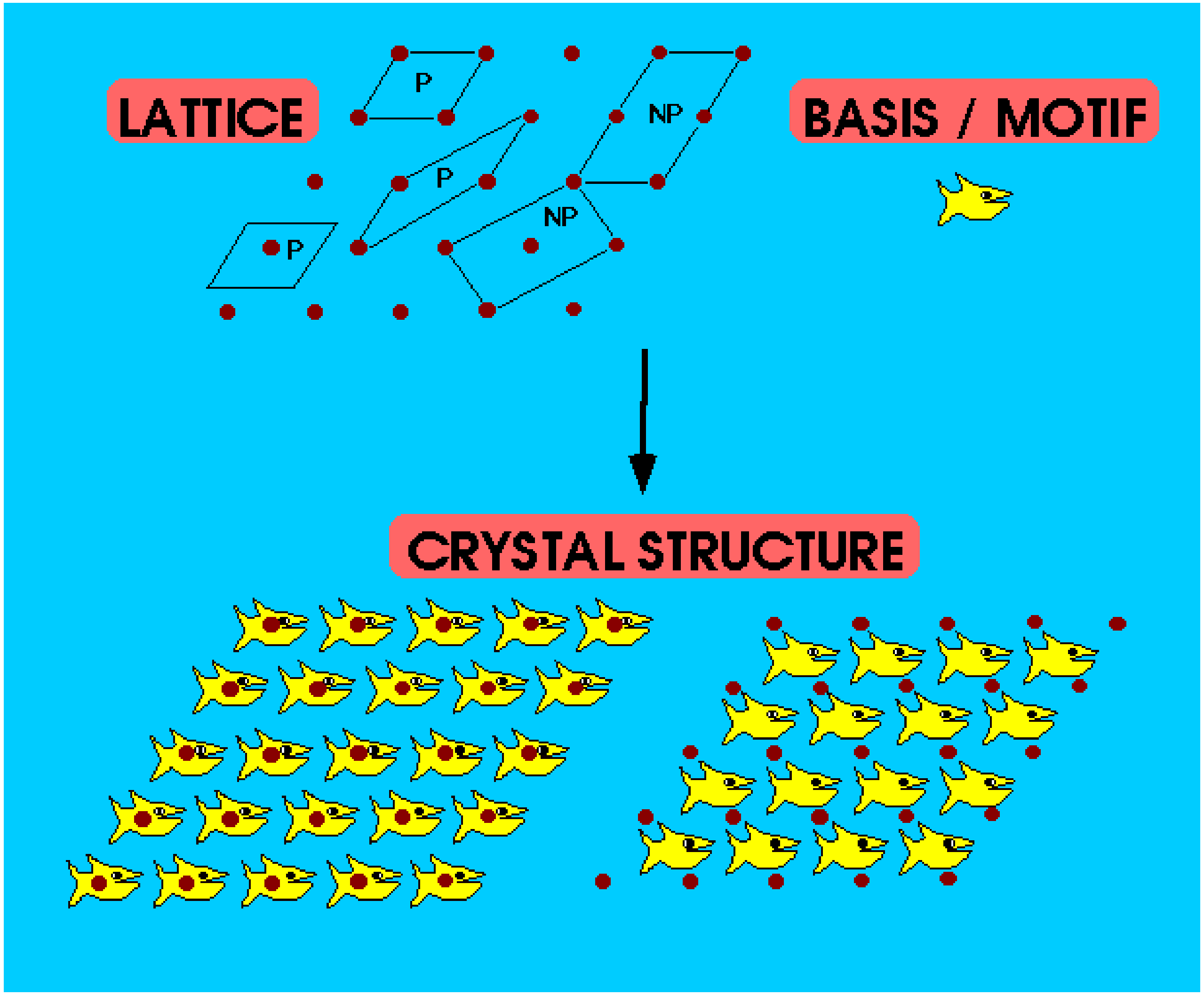
General:

$000l$: $l = 3n$

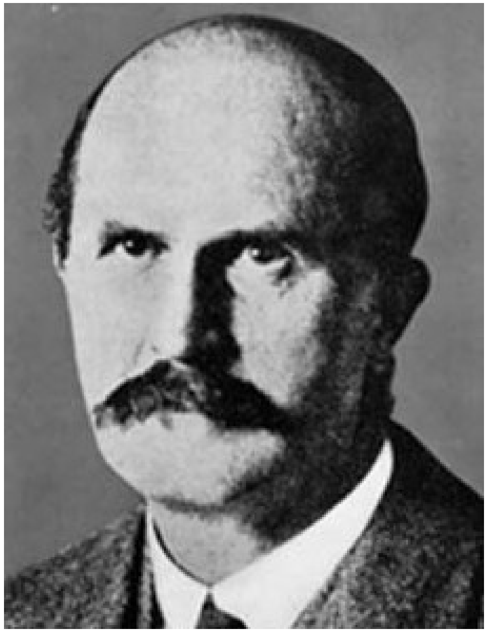
Symmetry of special projections

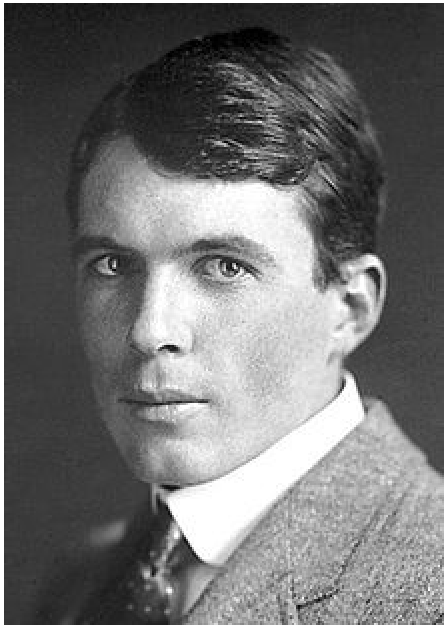
Along [001] *p*3	Along [100] *p*1	Along [210] *p*1
$\mathbf{a}' = \mathbf{a}$ $\mathbf{b}' = \mathbf{b}$	$\mathbf{a}' = \frac{1}{2}(\mathbf{a}+2\mathbf{b})$ $\mathbf{b}' = \mathbf{c}$	$\mathbf{a}' = \frac{1}{2}\mathbf{b}$ $\mathbf{b}' = \mathbf{c}$
Origin at $0,0,z$	Origin at $x,0,0$	Origin at $x,\frac{1}{2}x,0$

Still Technical Problems:
Which unit cell + lattice for crystal?



Part 2: Pioneers of Diffraction

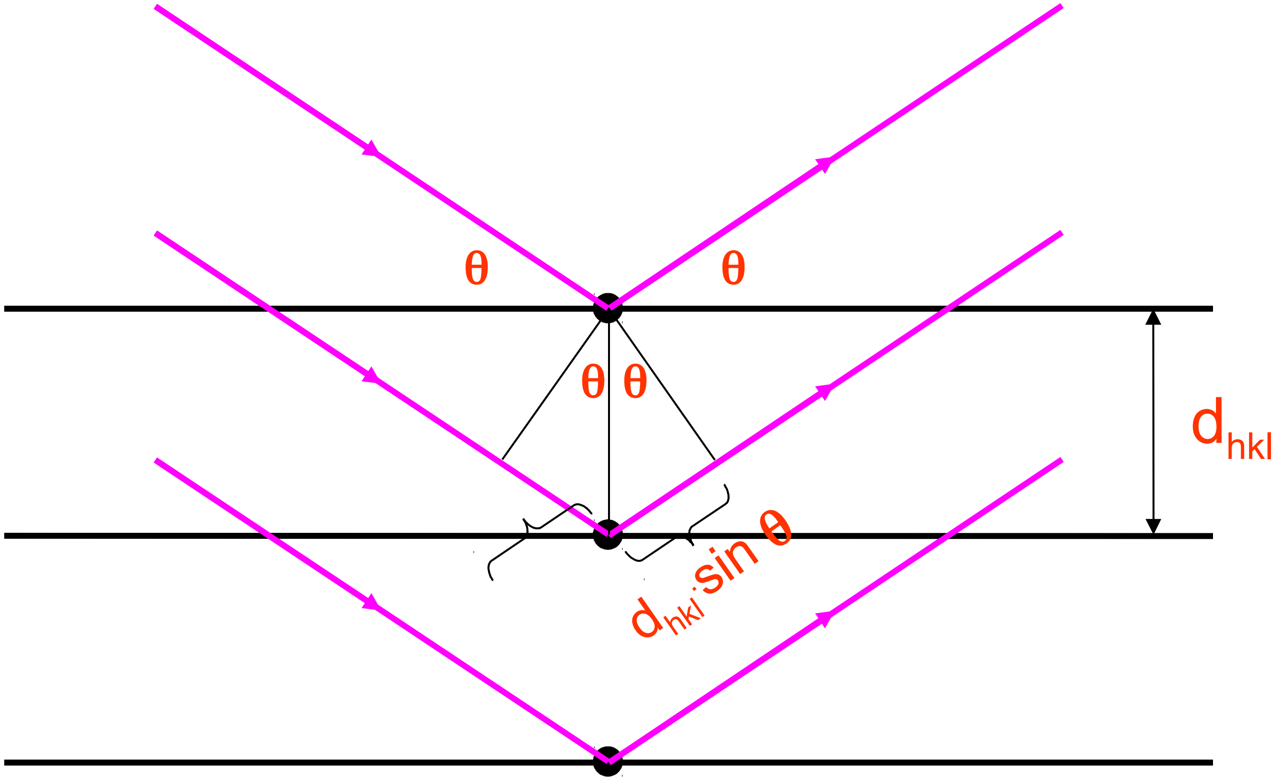




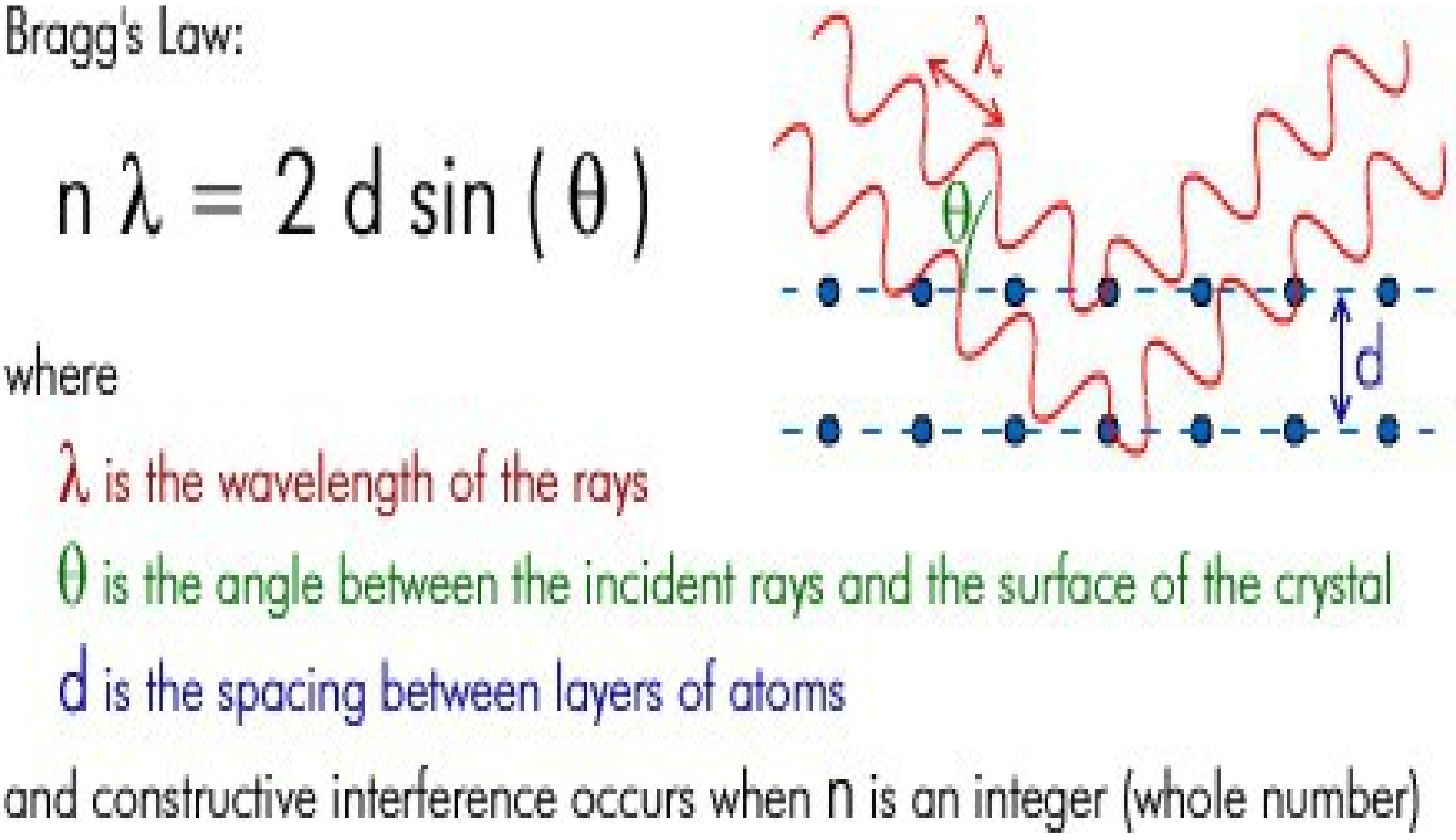
Laue – First saw
diffraction 1912

Braggs – Father and son
Braggs' Law 1912 explained it

Bragg's law relates diffraction angle θ, interplanar spacing d and wavelength λ



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



Bragg's Law:

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

where

λ is the wavelength of the rays

θ is the angle between the incident rays and the surface of the crystal

d is the spacing between layers of atoms

and constructive interference occurs when n is an integer (whole number)

Black dots represent single atoms (lattice points) repeating at intervals related to the unit cell edges.

The Structure Factor Generated by a “Real” crystal

The lattice could be defined by the black or red or blue atoms. Same lattice diffraction properties – different scattering centres so different phases when the wave hits the detector.

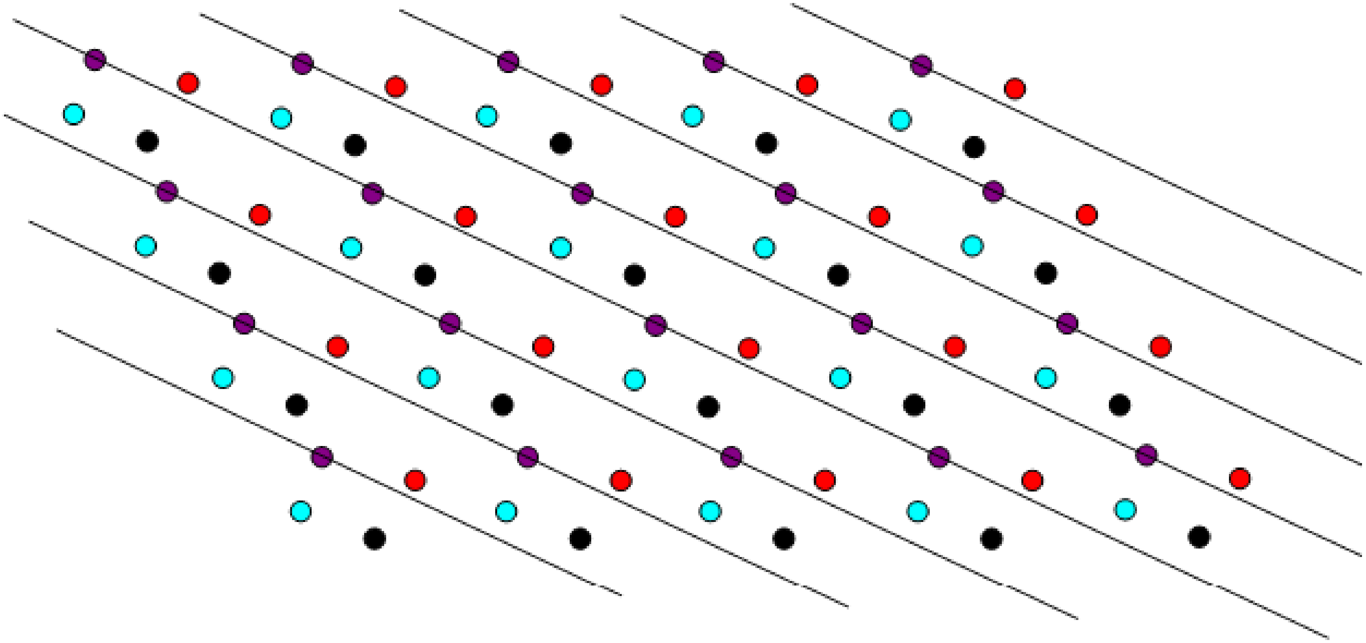


Figure 18.13 A hypothetical two-dimensional crystal showing richly populated lines

Different scattering planes would generate a different structure factor.

Diffraction indexed as (h k l) in “reciprocal space” for convenience

Define convenient axes **a*, b* and c*** such that

a* is perpendicular to **b** and **c**
b* is perpendicular to **a** and **c**
c* is perpendicular to **a** and **b**

For an orthorhombic crystal
length of a* proportional to 1/(length of a)

Total diffraction is a sum from all the atoms

Algebraic representation of Waves with differing phases can be written in various ways.

$$\mathbf{F} = \sum_{j=1}^{N} f_j \exp\left(2\pi i \mathbf{h}\cdot\mathbf{x}_j\right) = A + iB$$

$F_{hkl}\exp(i\Phi_{hkl})$

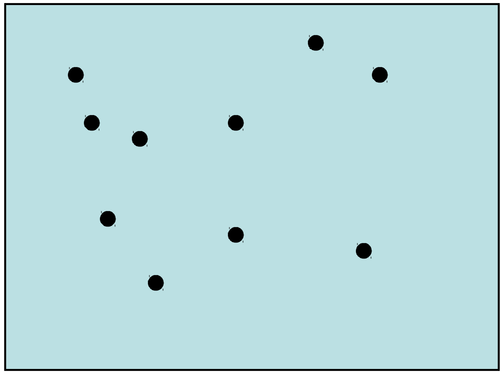
$$A = \sum_{j=1}^{N} f_j \cos\left(2\pi \mathbf{h}\cdot\mathbf{x}_j\right)$$

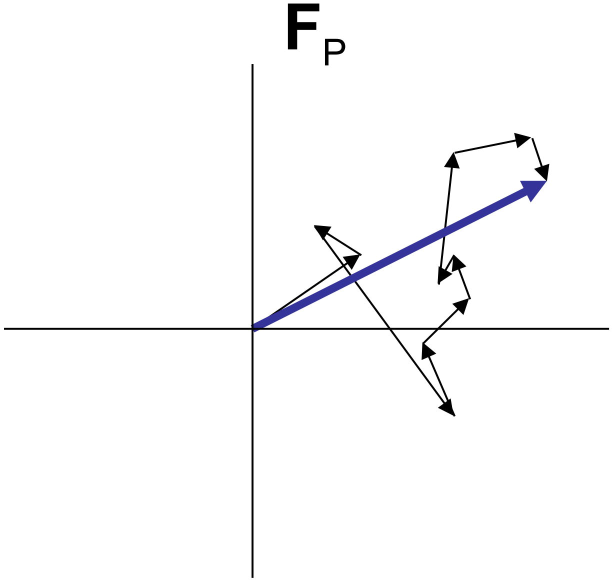
$\mathbf{h} = ha^* + kb^* + lc^*$

$\boldsymbol{x} = x.a + y.b + z.c$

$$B = \sum_{j=1}^{N} f_j \sin\left(2\pi \mathbf{h}\cdot\mathbf{x}_j\right)$$

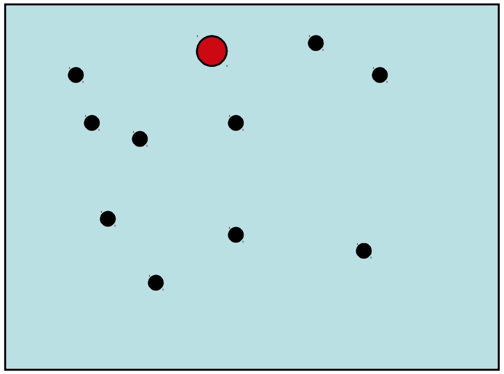
Or as vectors in 2D - Structure Factor

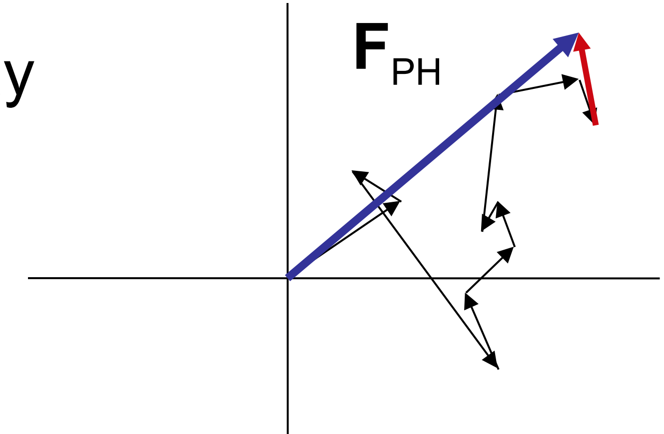




$F = |F| \exp(i\,\Phi)$

Adding one (or more) atoms in known positions changes the structure factor in a known way

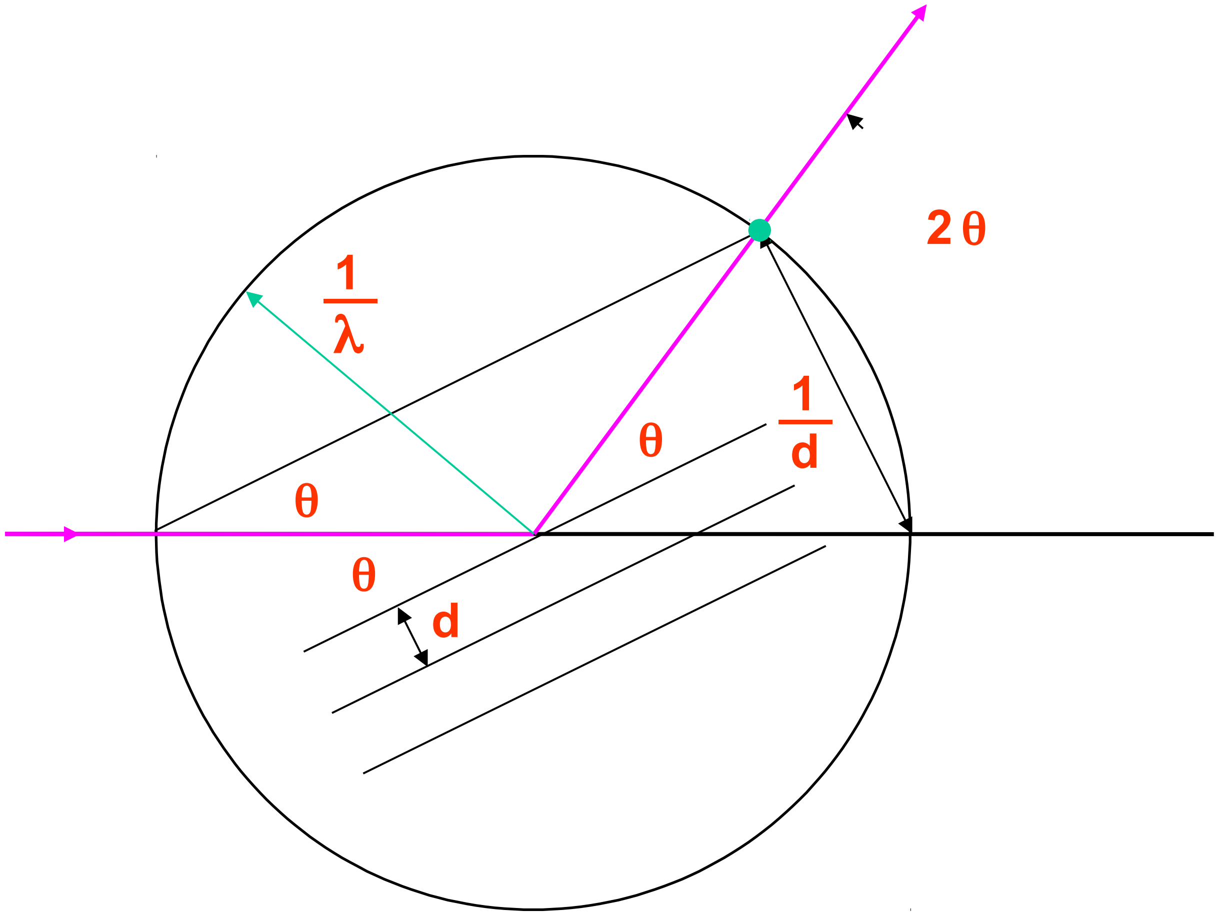




$F = |F| \exp(i\,\Phi p + \Phi new)$

An aside: Ewald's sphere construction illustrates Bragg's law in three dimensions

Ewald and Bragg on opposite sides in WW1



$$\sin\theta = \frac{\frac{1}{d}}{\frac{2}{\lambda}}$$

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

William L. Bragg realised that **IF** he measured and indexed diffraction for each Fobs(hkl) from a crystal

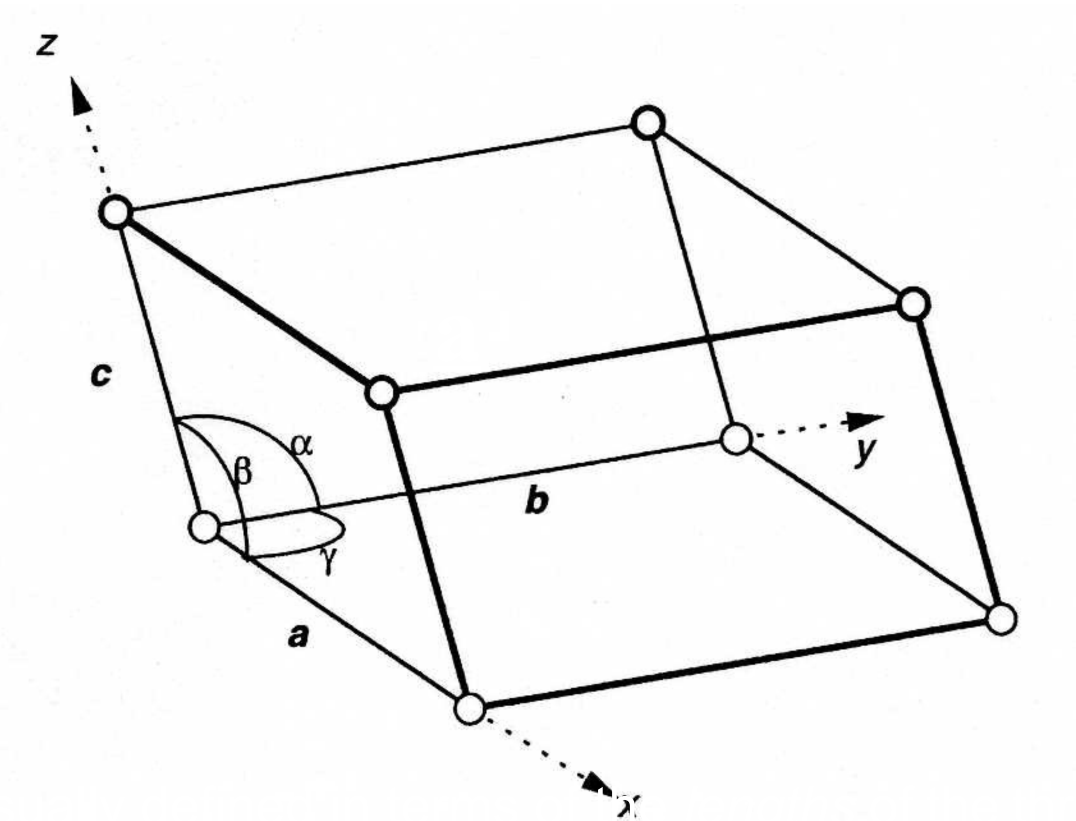
and **IF** he knew the relative position of all the atoms in the crystal

and **IF** he calculated the structure factors Fcalc(hkl) from these positions for each reflection

then he should find
Fcalc (hkl) had the same magnitude as ***Fobs(hkl)***

He positioned atoms according to symmetry and known chemical considerations for sodium chloride, zinc sulphide and diamond and thus determined the very first crystal structures.

An aside: Is crystallographic maths simple??
Yes – only 3 dimensions, but ..



Most general representation of crystal cell –
note potential nightmare for mathematics.
no need for equal sides or right angles

Converting coordinates is messy

The default PDB **orthogonal coordinates X, Y, Z use the following transformation from the fractional coordinates x, y, z:**

$$\begin{aligned} X &= a\,x + (b\cos\gamma)\,y + (c\cos\beta)\,z \\ Y &= 0 + (b\sin\gamma)\,y + (-c\sin\beta\cos\alpha^*)\,z \\ Z &= 0 + 0 + (c\sin\beta\sin\alpha^*)\,z \end{aligned}$$

where $\cos\alpha^* = (\cos\beta\cos\gamma - \cos\alpha) / (\sin\beta\sin\gamma)$

and $\sin\alpha^* = \sqrt{(1 - \cos^2\alpha^*)}$

For the X, Y, Z $\longrightarrow$ x, y, z transformation the inverse matrix is used.

$$\begin{aligned} z &= Z / (c\sin\beta\sin\alpha^*) \\ y &= (Y - (-c\sin\beta\cos\alpha^*)\,z) / (b\sin\gamma) \\ x &= (X - (b\cos\gamma)\,y - (c\cos\beta)\,z) / a \end{aligned}$$

Fourier Series
Calculating electron density

Any continuous differentiable function f of the real variable x which is defined over a finite interval can be expanded as a summation of cos and sin terms

It can be represented by (under certain conditions, a Fourier series of the form:

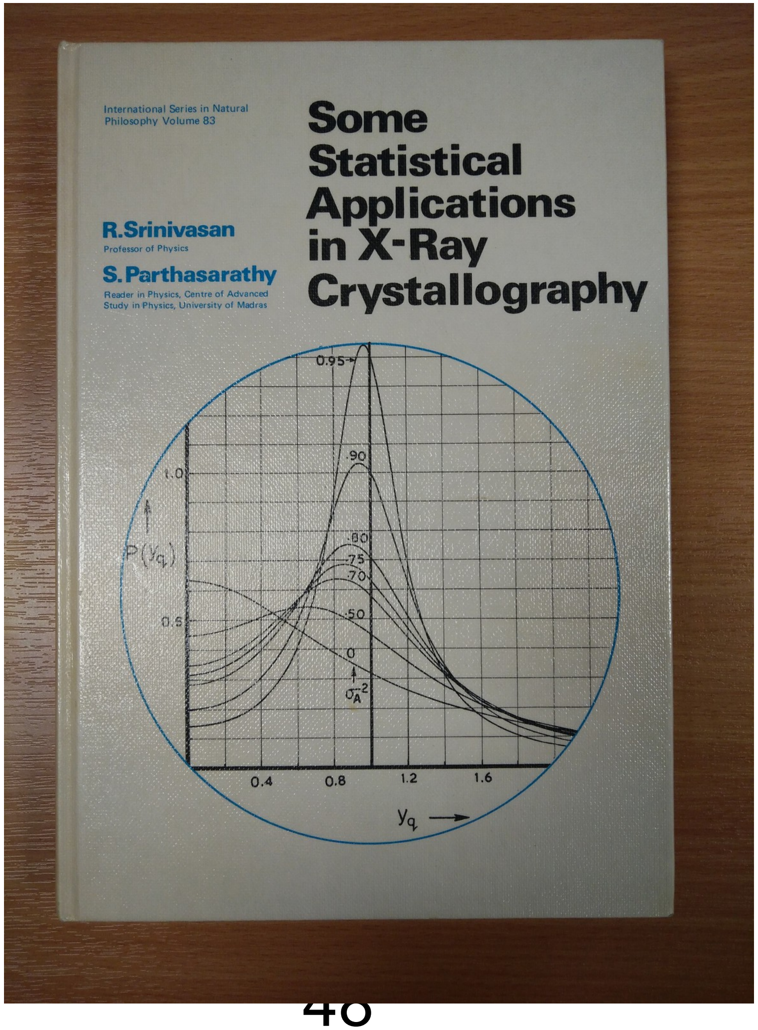
$F(x) = a_0 + \Sigma_k\ a_k \cos[(\pi kx)/l] + b_k \sin[(\pi kx)/l]$

It is obviously designed for crystallography

Fourier Series
Calculating electron density

Favourite References:
"Fourier methods in crystallography",
R. Srinivasan & G.N. Ramachandran.
(Lost, stolen or strayed..)

"Some statistical applications in X-ray crystallography".
R. Srinivasan & S. Parthasarathy.



International Series in Natural Philosophy Volume 83

Some Statistical Applications in X-Ray Crystallography

R.Srinivasan
Professor of Physics

S.Parthasarathy
Reader in Physics, Centre of Advanced Study in Physics, University of Madras

Electron Density

Exact Φ_{hkl} and accurate measurements gives perfect map

All atoms will be visible

$$\rho(x_i,y_i,z_i) = \sum f_{hkl} \exp(i\Phi_{hkl}) \cos(2\pi(hx_i+ky_i+lz_i))$$

Partial or inaccurate Φ_{hkl} gives flawed map

$$\rho(x_i,y_i,z_i) = \sum F_{hkl} \exp(i\Phi_{hkl}) \cos(2\pi(hx_i+ky_i+lz_i))$$

Here partial structure visible.
missing atoms will show up at < half weight

Peak height is a function of the ratio of known:unknown

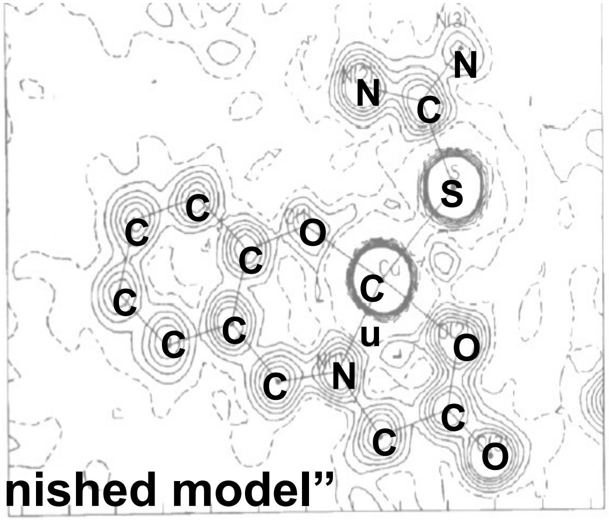
The importance of the difference electron density maps

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l (\underbrace{|F_{obs}|}_{\text{Real crystal}} - \underbrace{|F_{calc}|}_{\text{model}}) \exp[-2\pi i(h\underbrace{x_j + ky_j + lz_j}_{\text{model}})]$$

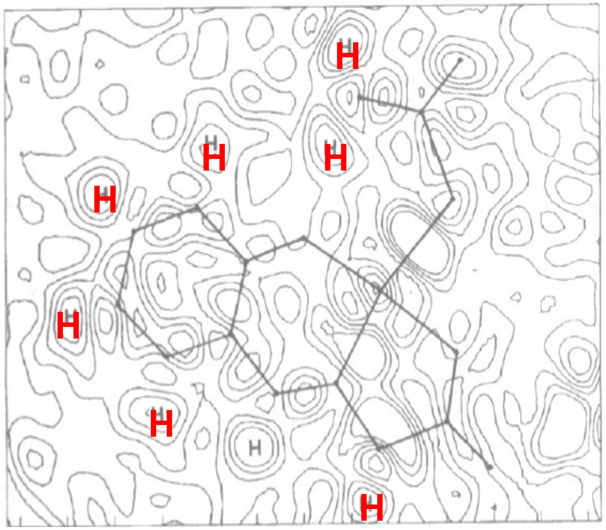
The ($|F_{obs}|$ - $|F_{calc}|$) should show only the differences between the actual and observed model: measured intensity, *I*, gives $|F_{obs}|$, the model that has been built gives $|F_{calc}|$.

The difference electron density(ed) map should show maxima (+ed) where the atoms are 'missing' in our model.

The difference ed map should show minima (negative ed) if the atom in the model was modelled in a wrong place, *i.e.* is not where it should be



"Finished model"



Difference ed map showing positions of missing hydrogen atoms

Examples from DCH work Cholesterol Iodide Maps 1943

This was the first structure where a crystallographer “saw” atoms..

There are 27 light atoms and one Iodine which provided 75% of the scattering.

Cholesteryl Iodide B

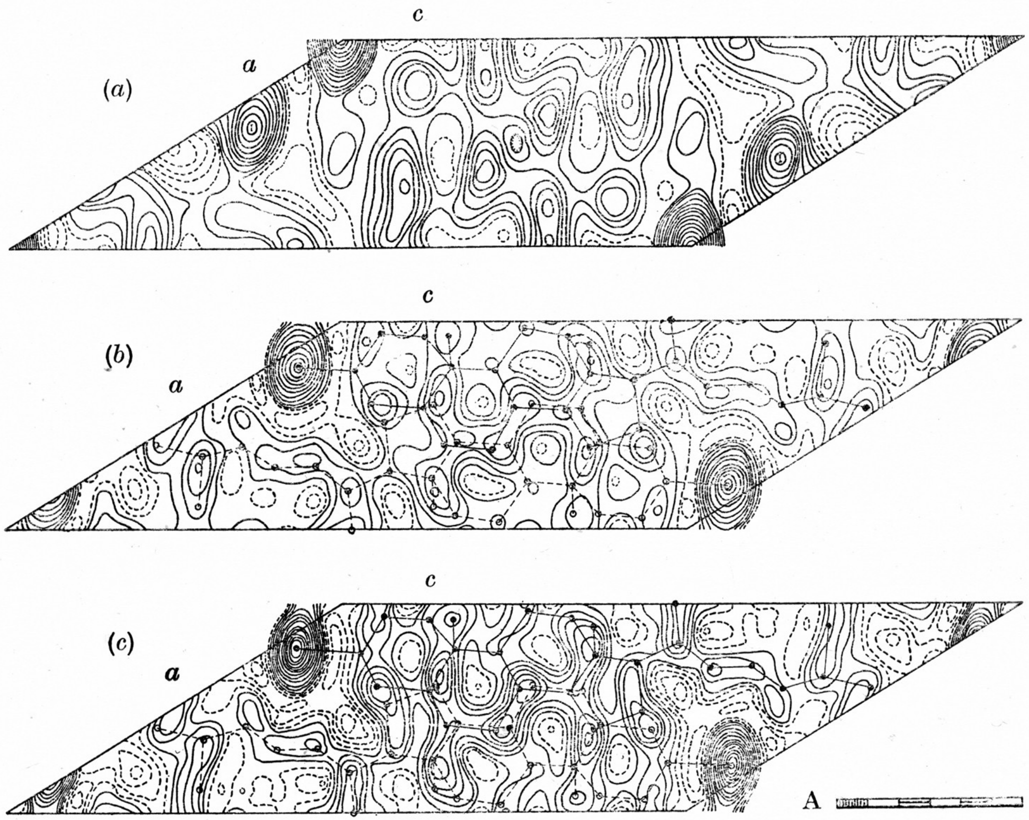


FIGURE 3. Cholesteryl iodide *B*. (*a*) Patterson projection (Pxz), on (010). The peaks *I*, *I* are due to the iodine-iodine vectors. (*b*) First projection of electron density, (ρxz), on (010). (*c*) Projection of electron density on (010) calculated, using intensities corrected for absorption and signs corrected for the carbon atom positions. Contours at rather closer intervals than (*b*).

Patterson showing Iodine vector

First Projection Fourier map

Projection Fourier map with improved data

Penicillin -1942-1946

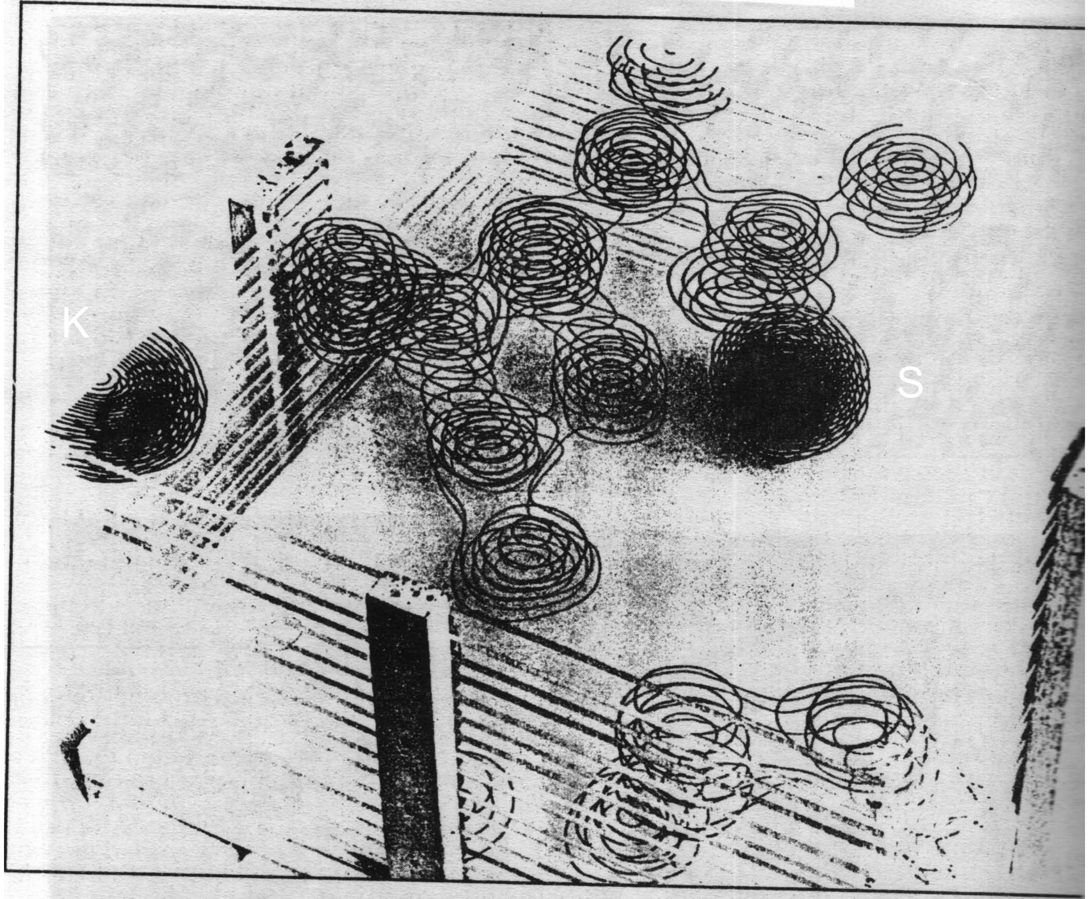
Same techniques, but bigger and more desperate need. (22 atoms + Rb) Chemistry not fully understood

…[The] great advantage of X-ray analysis as a method of chemical structure analysis is its power to show some totally unexpected and surprising structure and to do so with absolute certainty

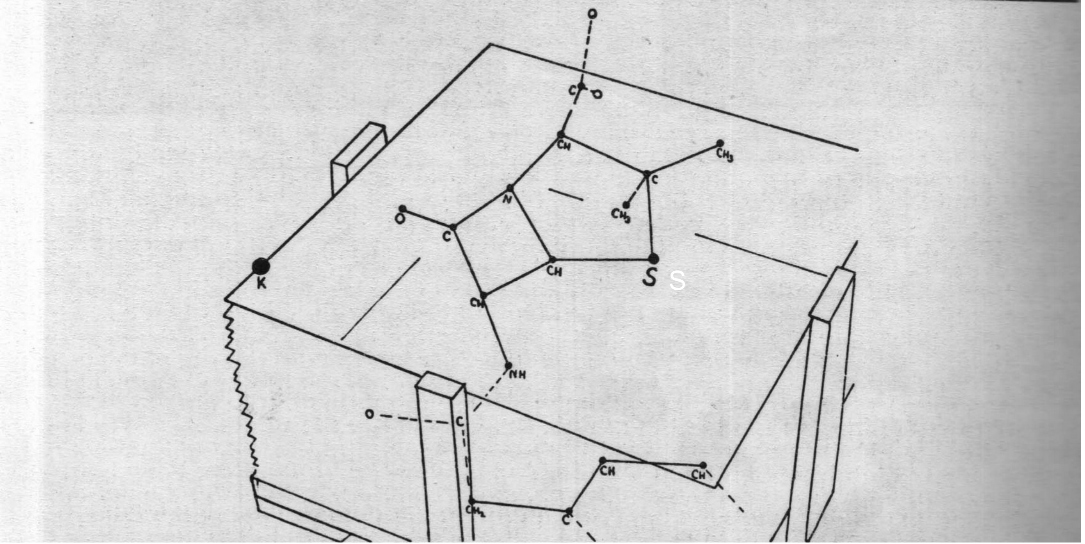
Penicillin was such a problem. Our urgent reason for doing the crystal structure was to try and plan its synthesis. ….

(Dorothy Hodgkin – Nobel Lecture 1965)

The 3-dimensional penicillin G map calculated in 1944. This unequivocally determined its chemical structure. (The hand however is wrong.)



The chemical interpretation of the electron density map. The four membered beta-lactam ring, the centre of the controversy.

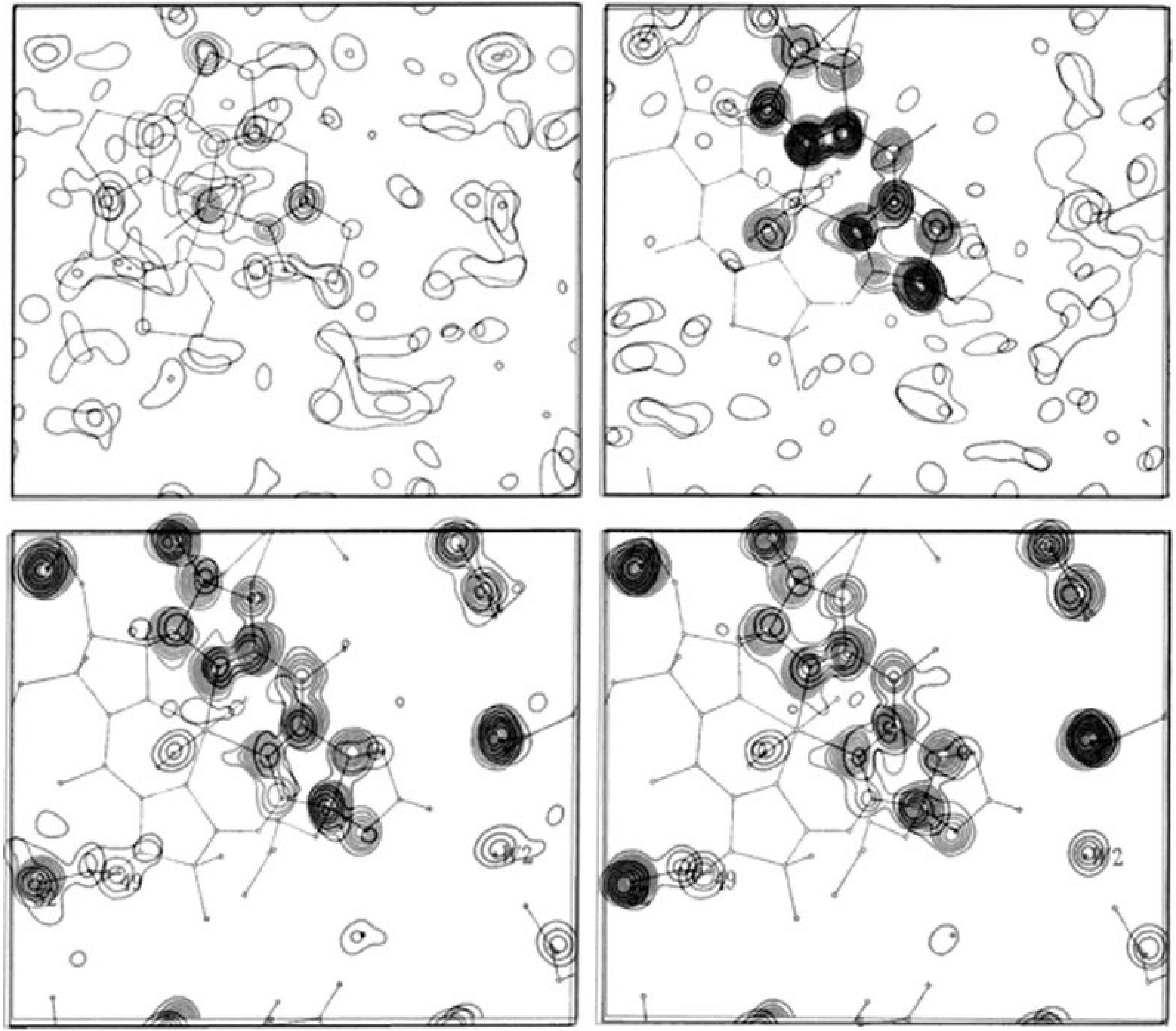


No successful cheap synthesis was achieved.

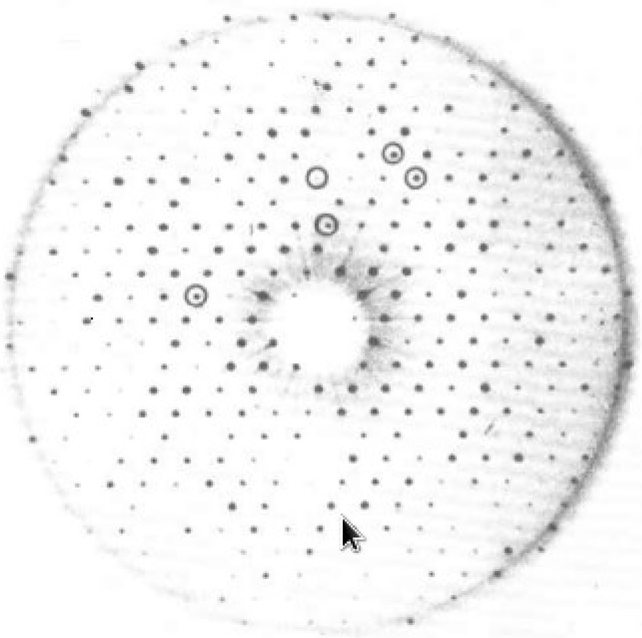
Vitamin B12 Electron Density

Much bigger (112 atoms)..
A corrin ring with a Co, but Co scattering only 13%

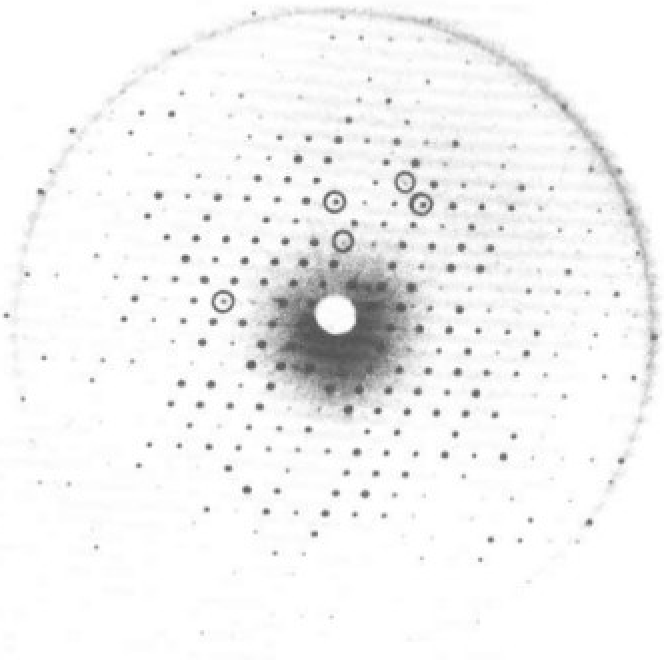
Uses technology.. Calculators then COMPUTERS!!


W2
49
52
W2
49

Practicalities of Protein Crystallography: Insulin Diffraction – 1960s (weak!)



HUMAN 2 Zn INSULIN (NOVO) Okl u = 18°
SEMISYNTHESIS FROM PIG INSULIN



PIG 2 Zn INSULIN Okl u = 21°

Insulin -1936-1969

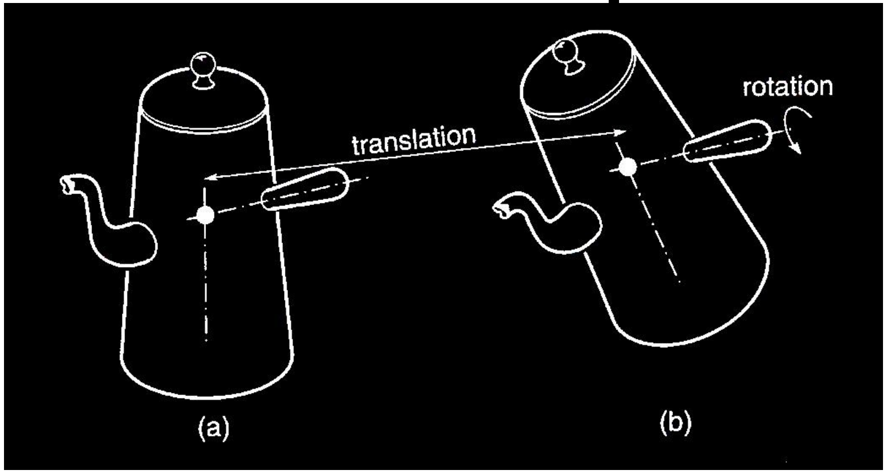
Bigger (800 atoms), protein = poorer data.

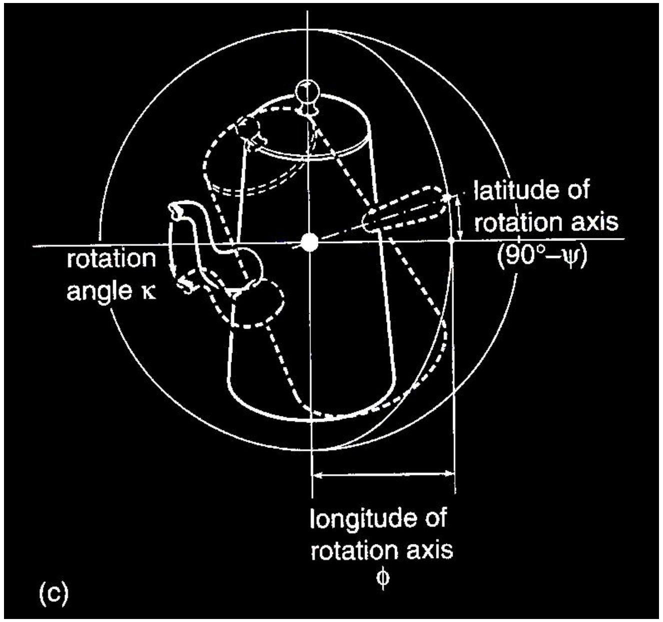
The Patterson/ Fourier techniques could not work. Needed other phasing ideas.

Isomorphous replacement using **anomalous signal** from heavy atoms.

Wild idea – maybe **molecular averaging** could help? My first task..

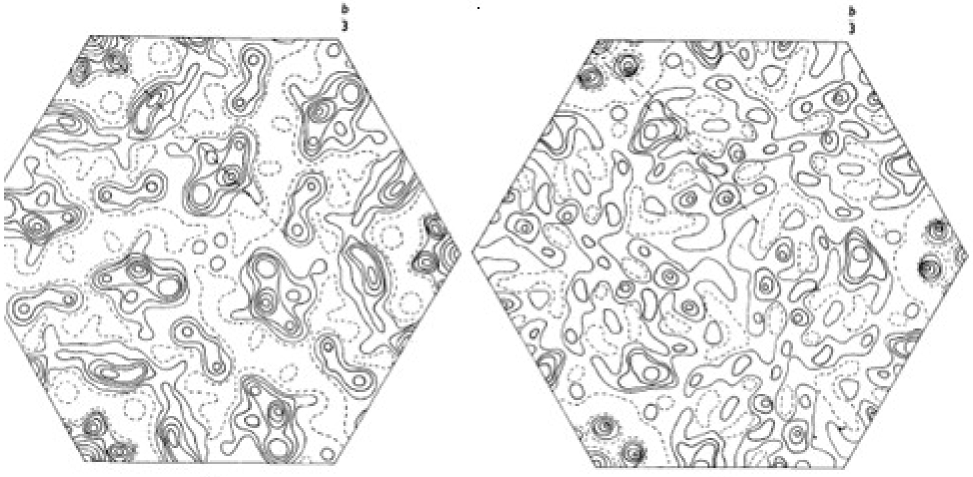
Molecular Replacement
Another Patterson Technique





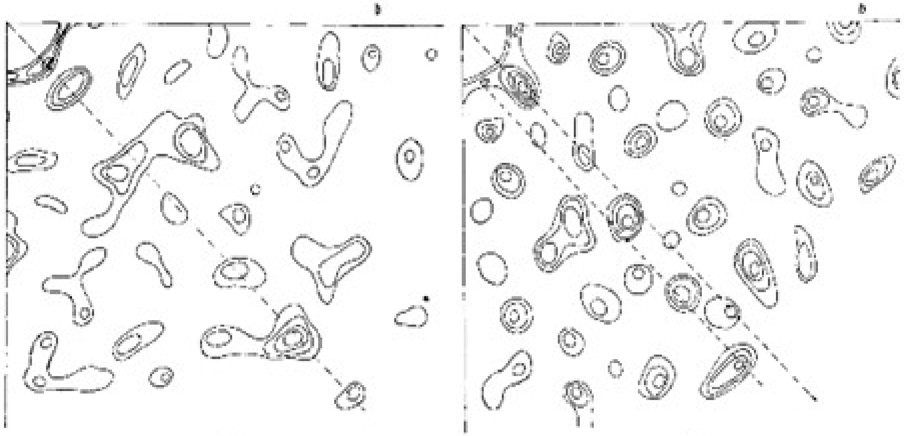
We need to understand the relationship between different copies of the same molecule in the unit cell.
Each molecule will generate similar Patterson features, but in different orientations.
Insulin was known to have 2 molecules in the crystal, so David Blow, Michael Rossmann, Marjorie Harding and I looked for patterns.
So many mistakes...
So much fun..

Insulin 2Zn and 4Zn Patterson showing 2-fold axis between molecules



(a) (b)

Fig. 12



(c) (d)

Fig. 12. Comparison of the Patterson projections of 2 Zn and 4 Zn insulin.

(a) and (b) Pattern unit corresponding with Figs 8(b) and 9(b) (unsharpened).

(c) and (d) Enlarged portions of asymmetric unit of the sharpened Patterson projections. The mirror plane passing through the origin is shown in both maps; the displaced dotted line in the 4 Zn map corresponds with a value of p of 0·8 Å.

Heavy-atom methods

- Isomorphous

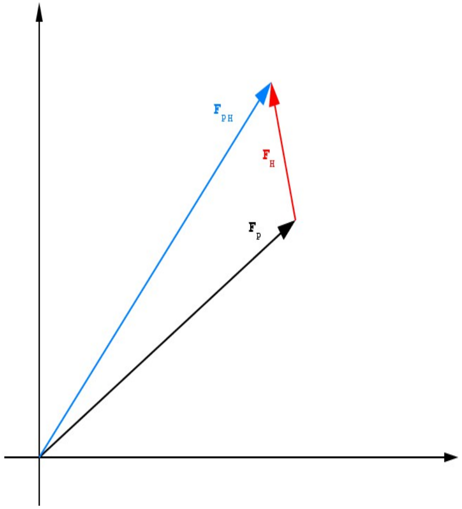
  differences between
  native and derivative data

- Anomalous

  differences between Friedel-related
  reflections, $|F(h)| \neq |F(-h)|$

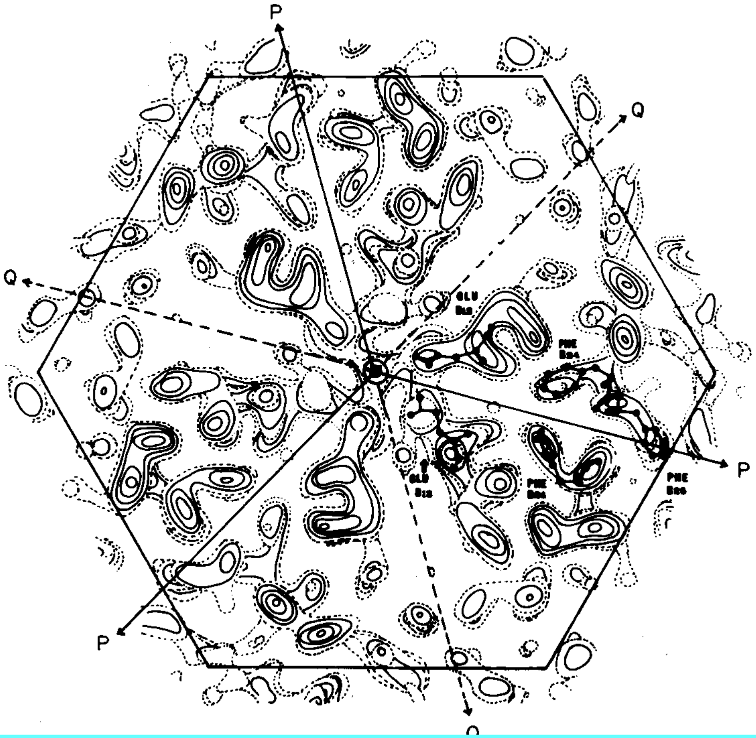
- All possible combinations

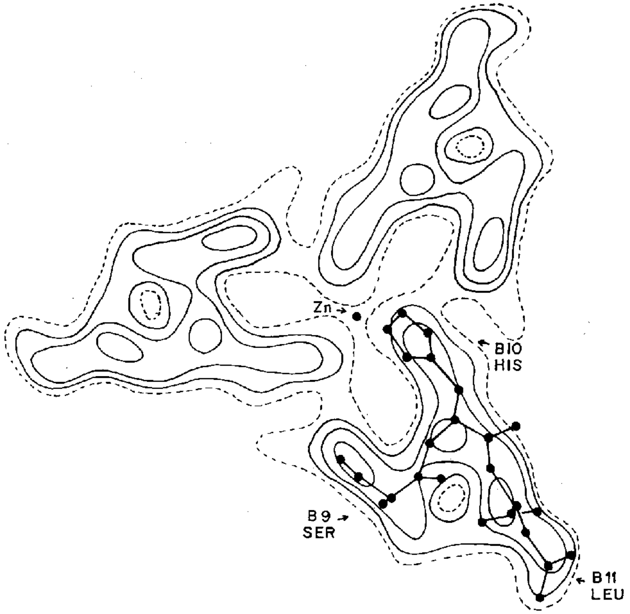
  SIR, MIR, SIRAS, MIRAS, SAD, MAD



DCH & Siv Ramasechan – 1965 our Anomalous Dispersion Guru

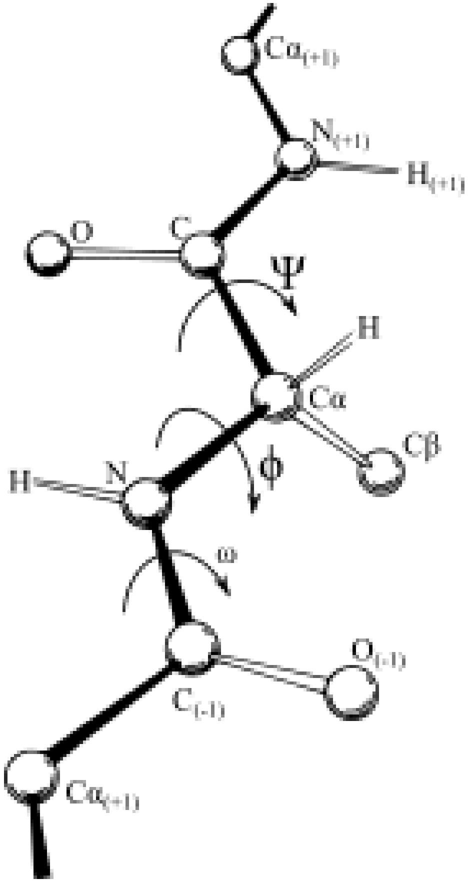
Electron density of the 2Zn insulin hexamer At 2.8 Å resolution

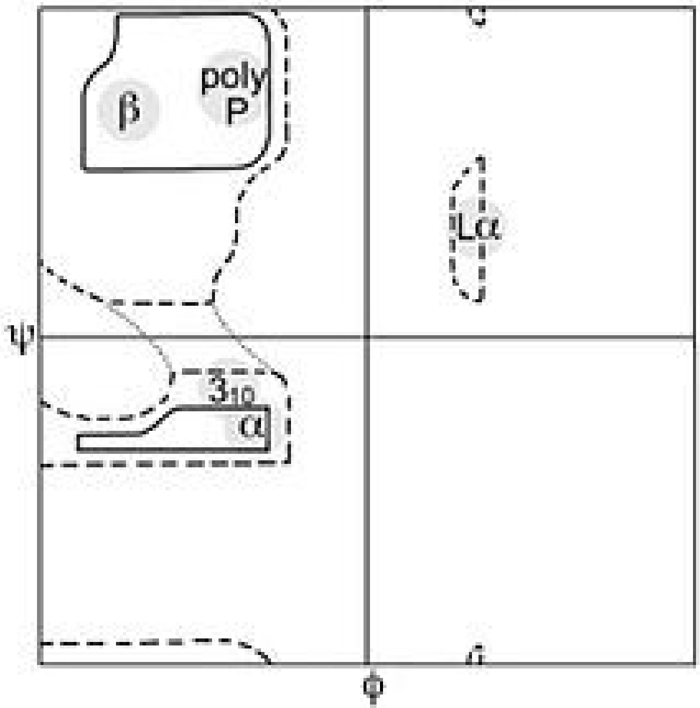




Validation:

More Thought Experiments:
G.N. Ramachandran





Happy Mapping

Acknowledgements

All the people I have worked with.
Particularly those who are or have been here..
Sivaraj, Vijayan, Tom Blundell, Heather & Ted Baker

Pictures from York laboratory, Hodgkin family , friends