Introduction to X-ray crystallography

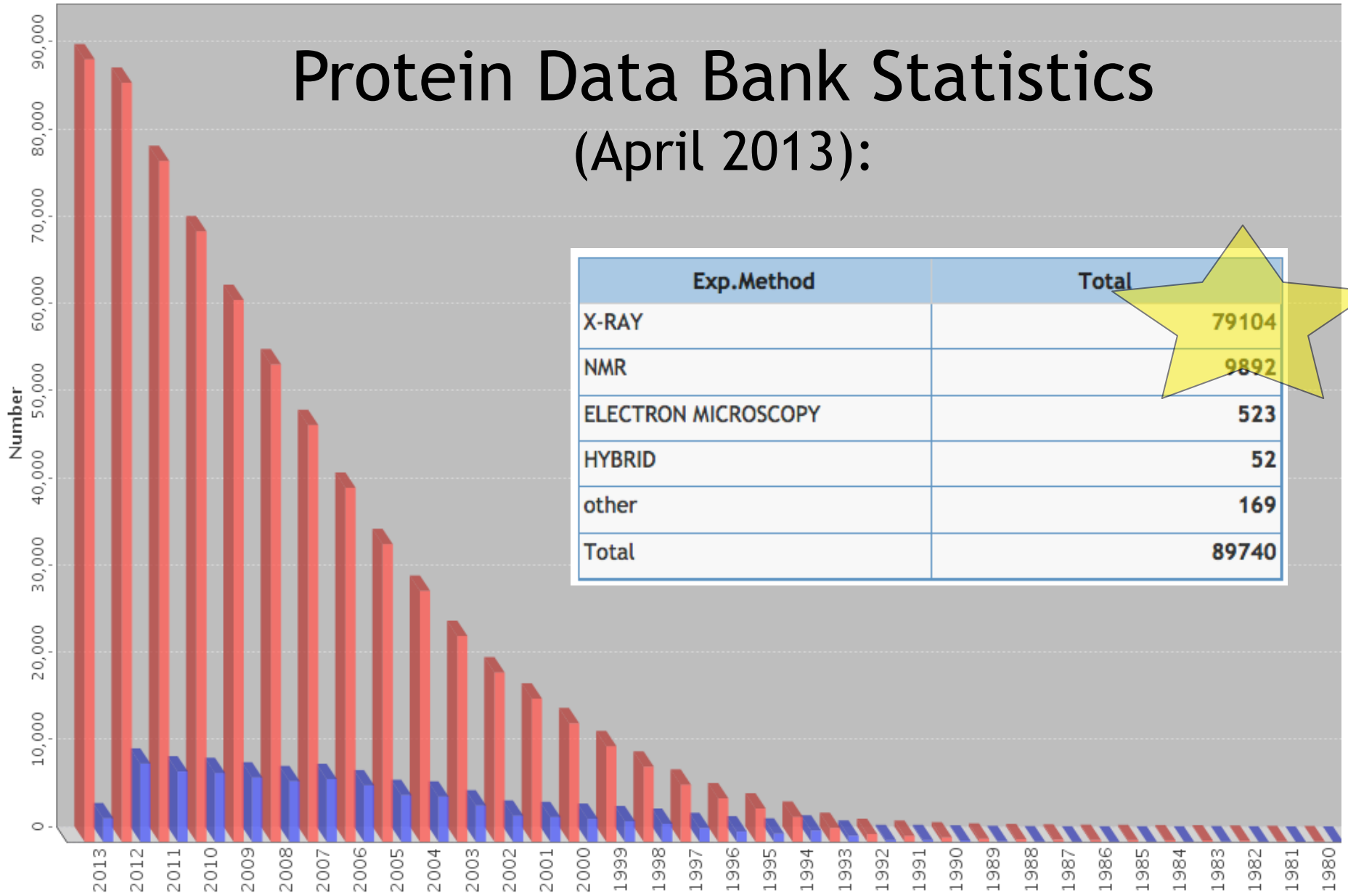
Jan Löwe
MRC Laboratory of Molecular Biology

Cambridge 23rd April 2013

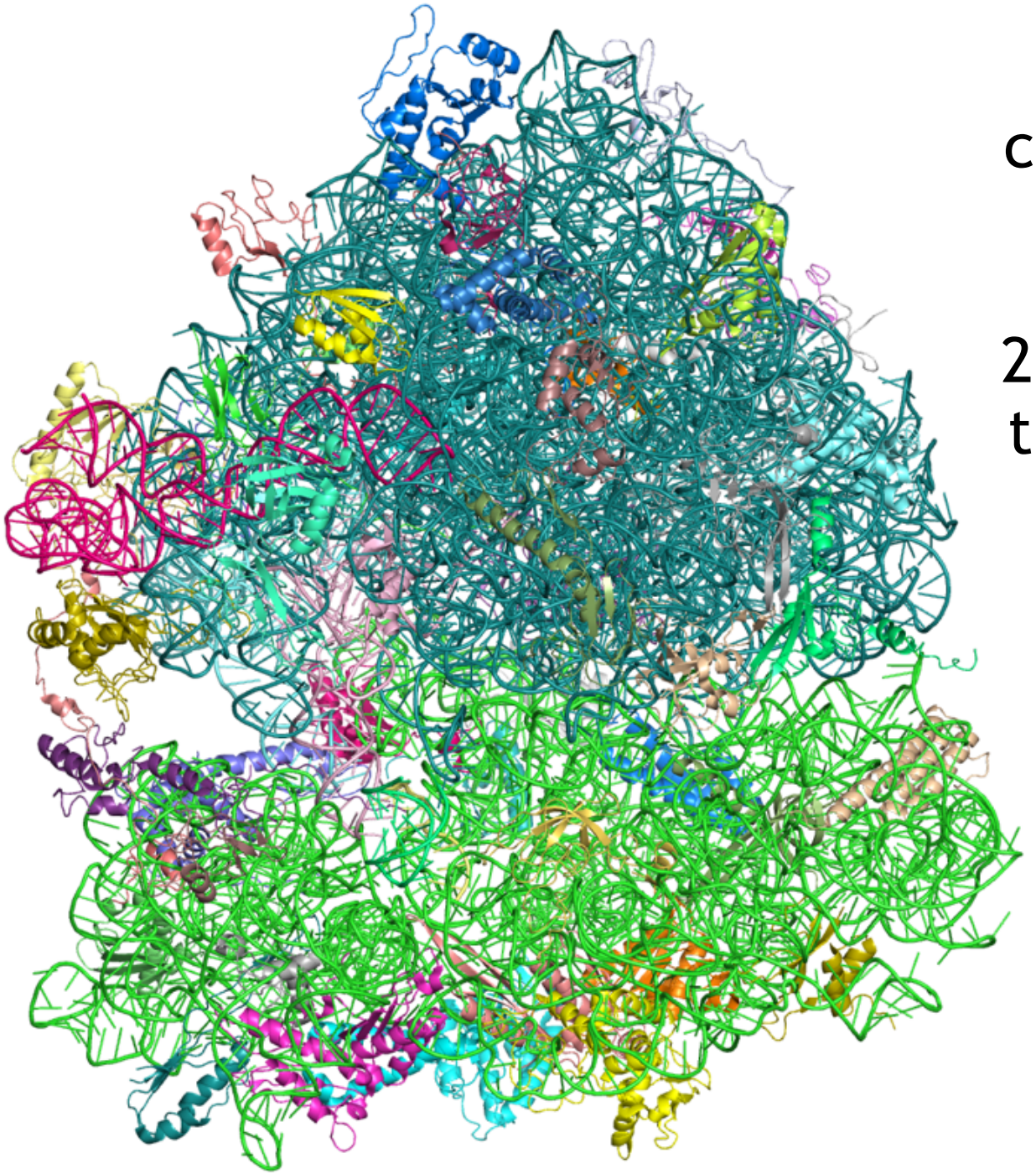
Sizes of biological structures

(universe ?)	telescope
planet	telescope
ecosystem	lens
organism	eye
cell	light microscope
molecule	X-ray, EM, NMR
atom	X-ray, (EM), NMR
(sub-atomic ?)	accelerator

Protein Data Bank Statistics
(April 2013):



Exp.Method	Total
X-RAY	79104
NMR	9892
ELECTRON MICROSCOPY	523
HYBRID	52
other	169
Total	89740



This is why crystallography is so good:

2.8 Å structure of the 70S ribosome

Selmer 2006

Complementarity of Methods

X-ray crystallography - highest resolution and reliability structures

NMR - enables widely varying solution conditions; characterisation of motions and dynamic, weakly interacting systems, molecules with no ordered structure

Electron microscopy - large molecules, fibres, organelles, cells

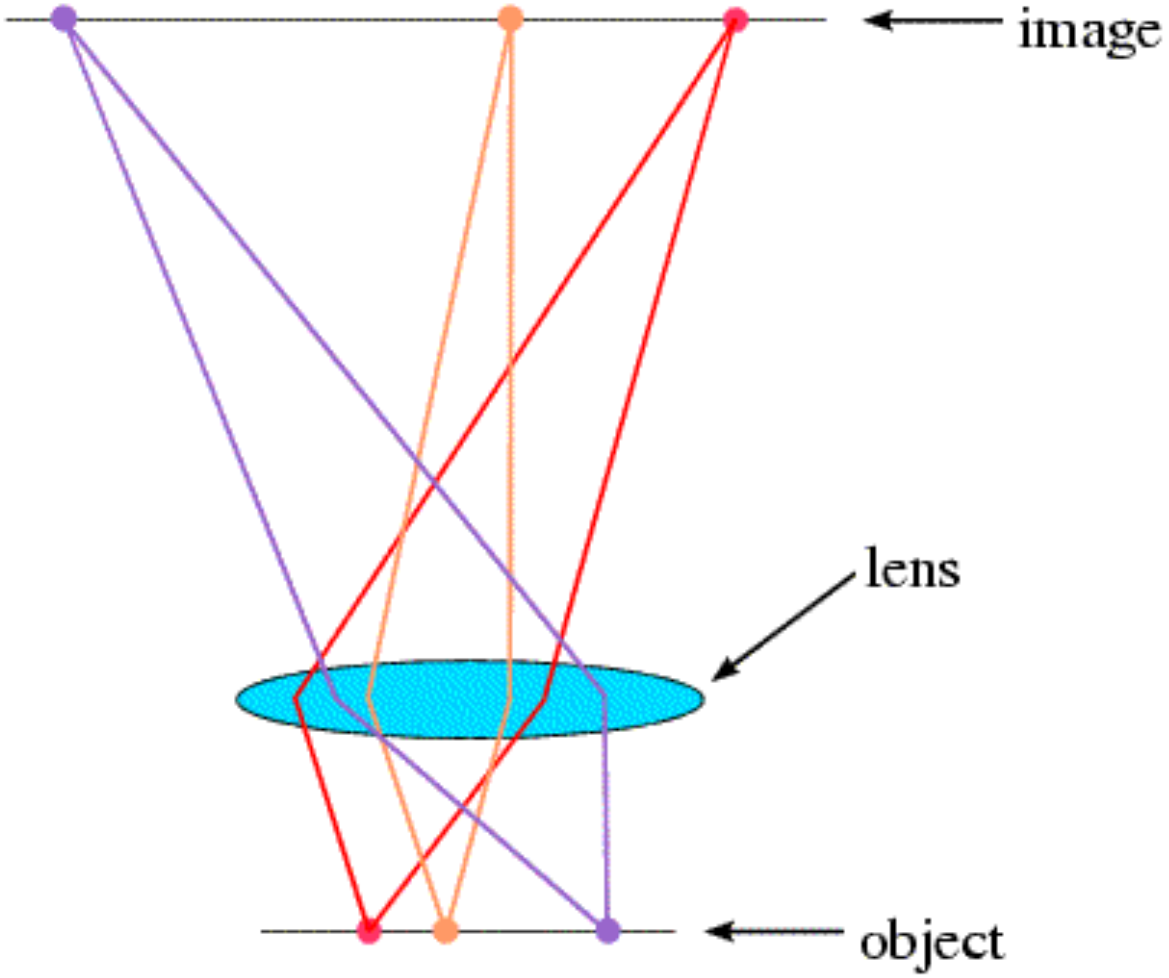
How to build an ‘atomic microscope’

For small objects (molecules) need short wavelength waves

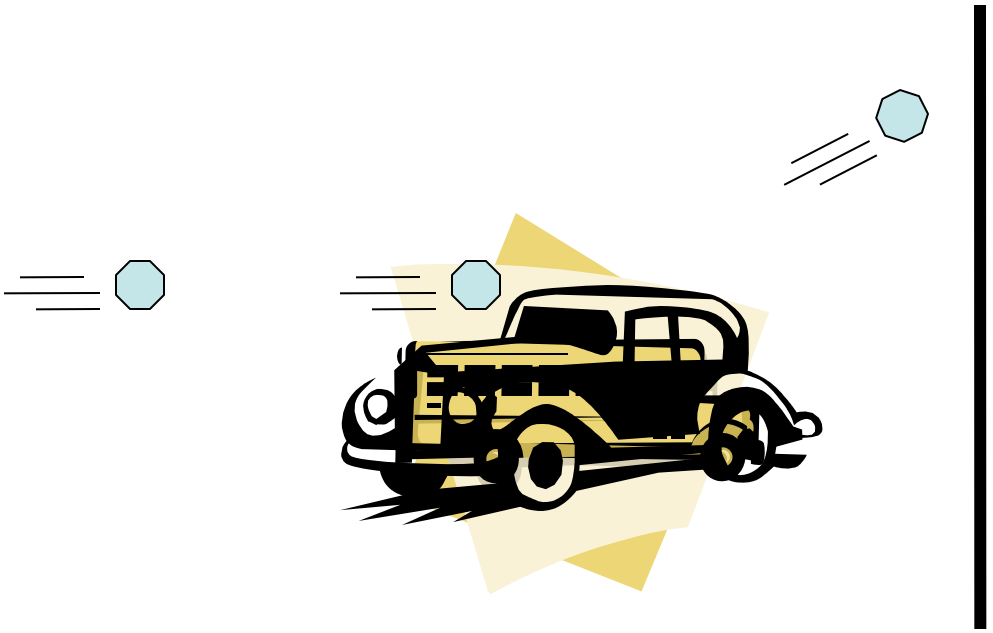
Inter-atomic distances are 1.2 – 2.0 Å so suitable wavelengths are < 2 Å (0.2 nm)

- electromagnetic radiation : X-rays
- electrons
- neutrons

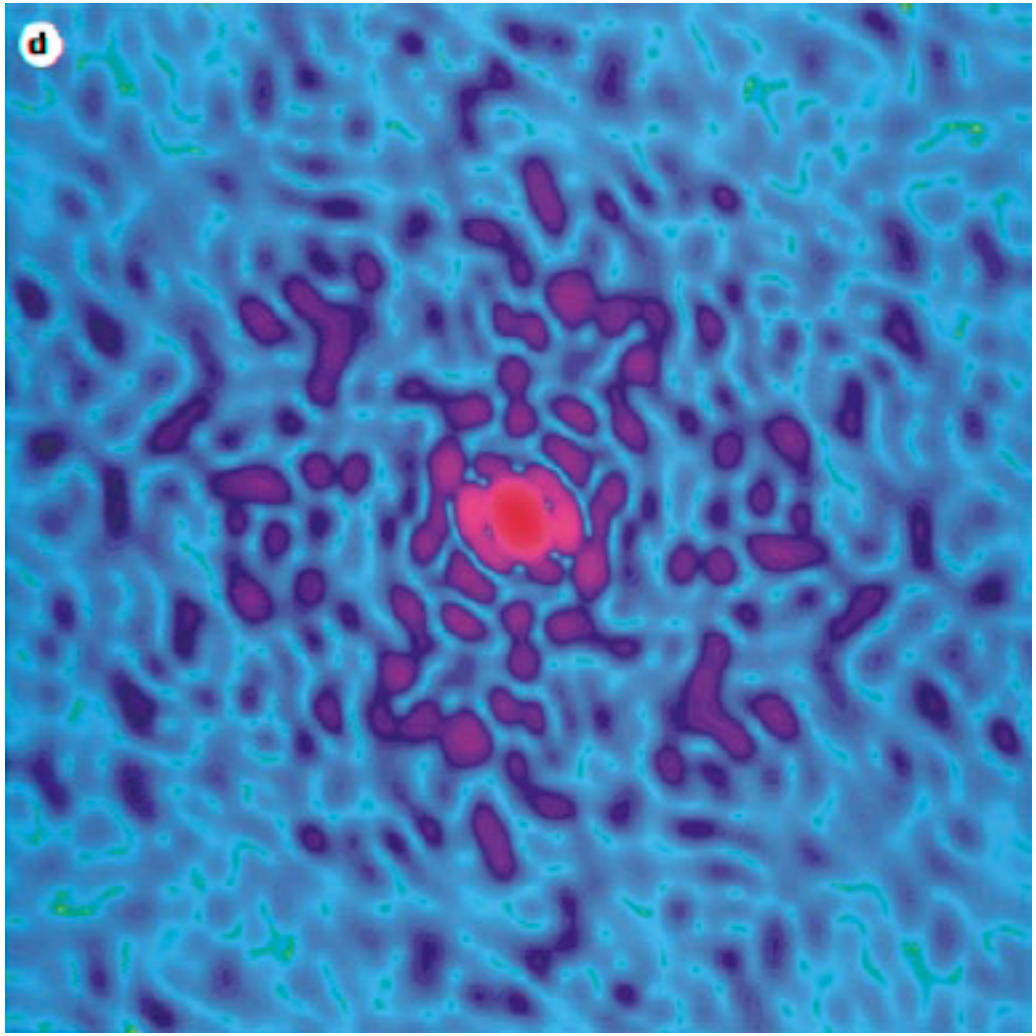
Unfortunately, no suitable lenses exist for X-rays



Imaging without lenses



Section through the *Fourier* transform of a protein

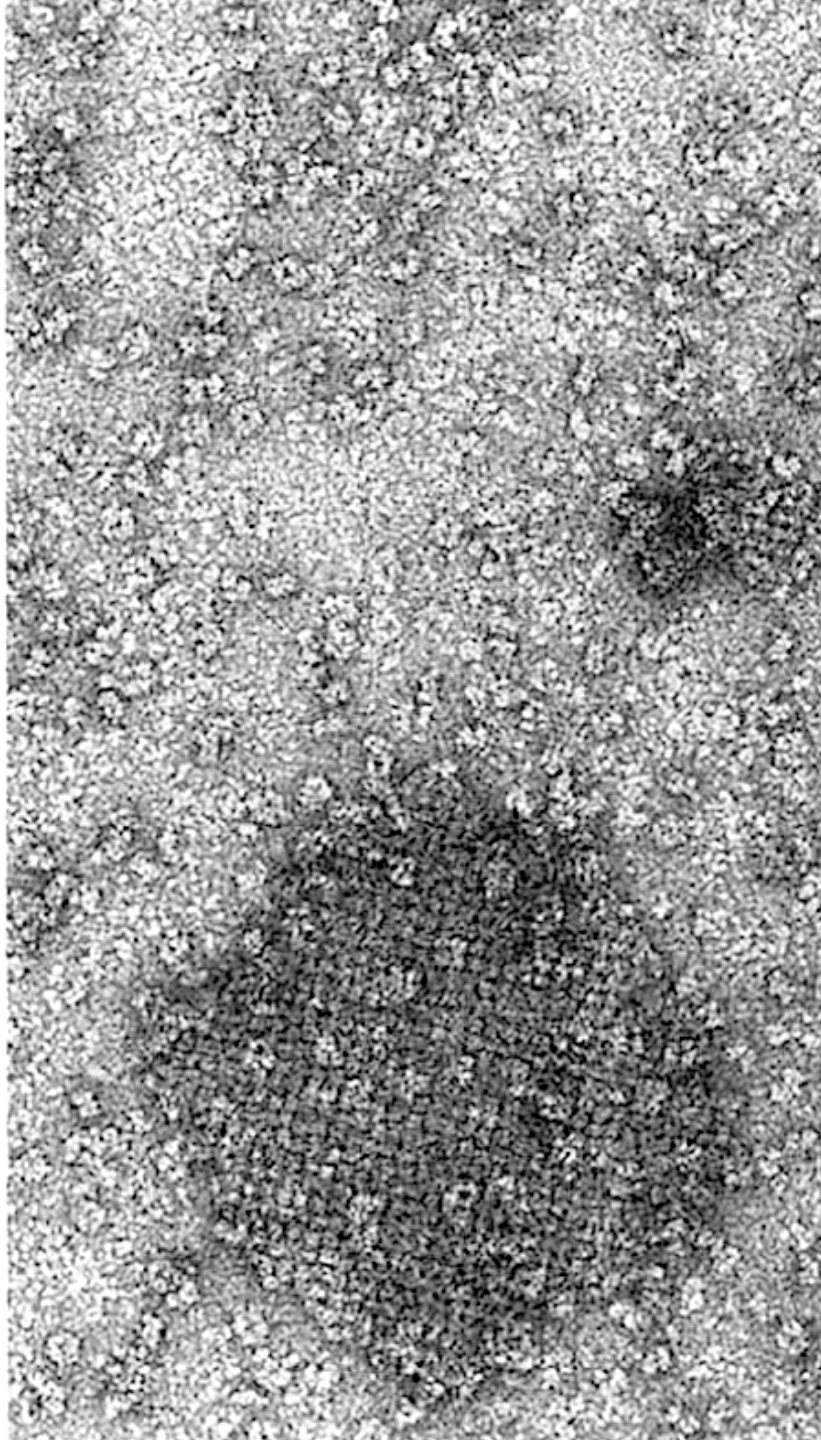


Diffraction from a single molecule is undetectable

To amplify the signal, we make arrays of molecules oriented identically

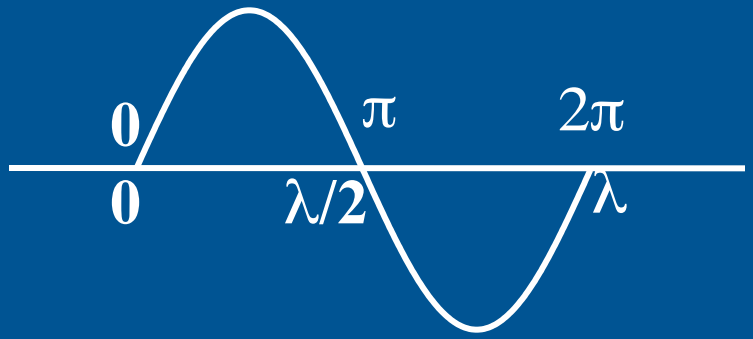
These arrays are crystals

A crystal 100 μm^3 contains 10^{12} molecules each 100 Å across, so amplifies signal by 10^{12}



In order to understand lenseless reconstruction, the wave nature of X-ray photons needs to be considered

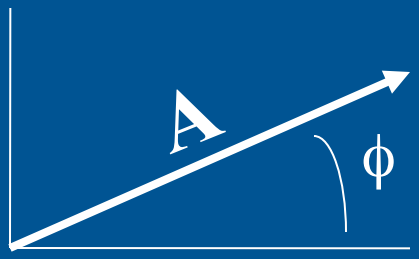
Waves



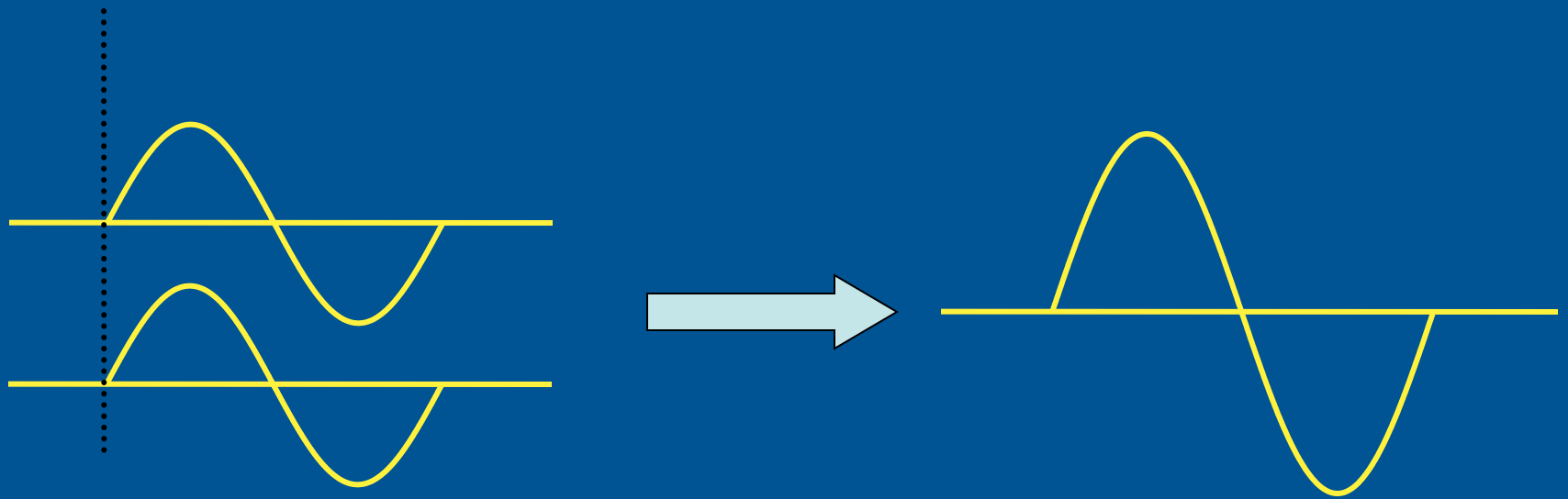
Phase (ϕ)
Displacement

$$A \sin\{2\pi(x/\lambda - \upsilon t)\}$$

λ = wavelength
υ = frequency
A = amplitude



Superposition of Waves

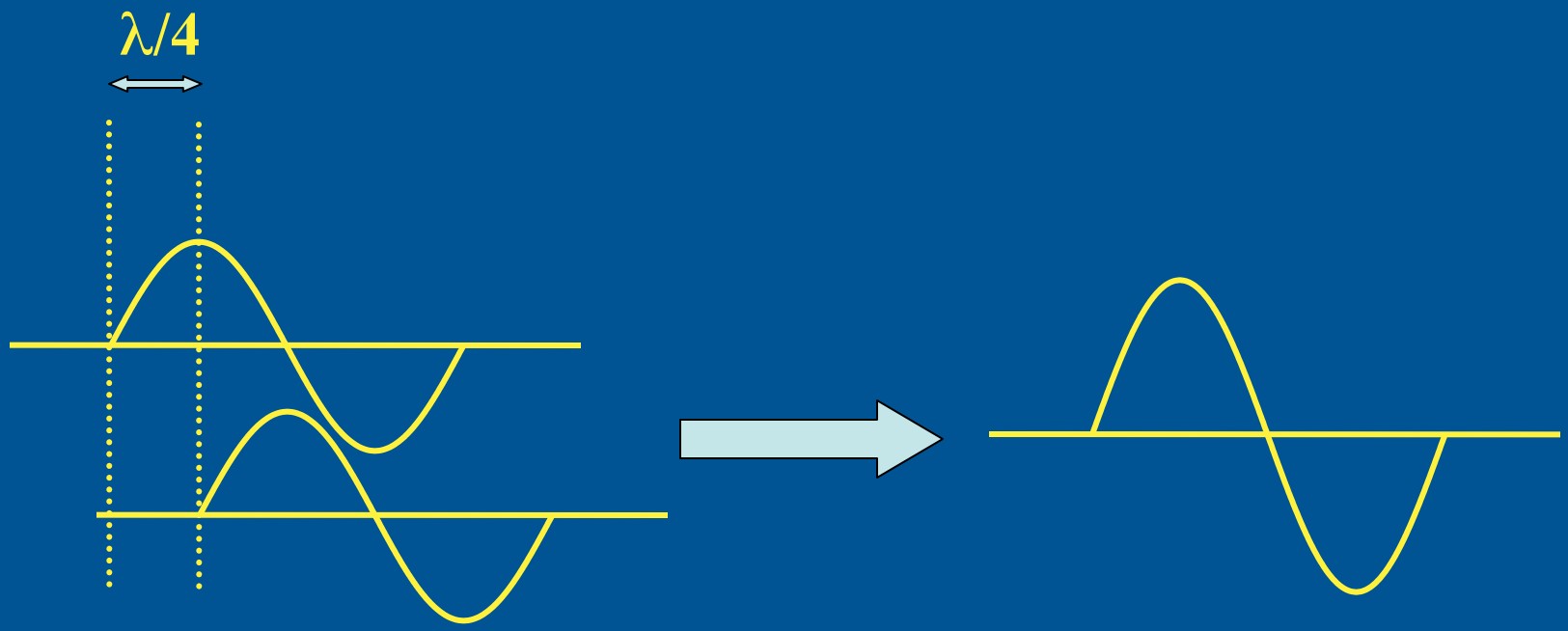


amplitude = A

amplitude = 2A

Constructive interference

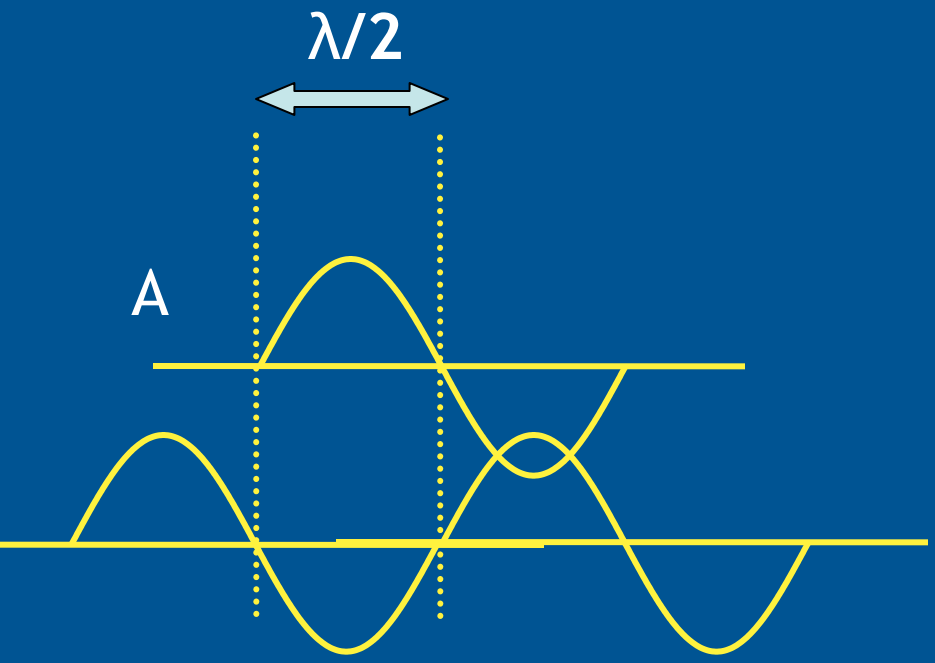
Superposition of Waves



amplitude = A

amplitude = 1.4A

Superposition of Waves

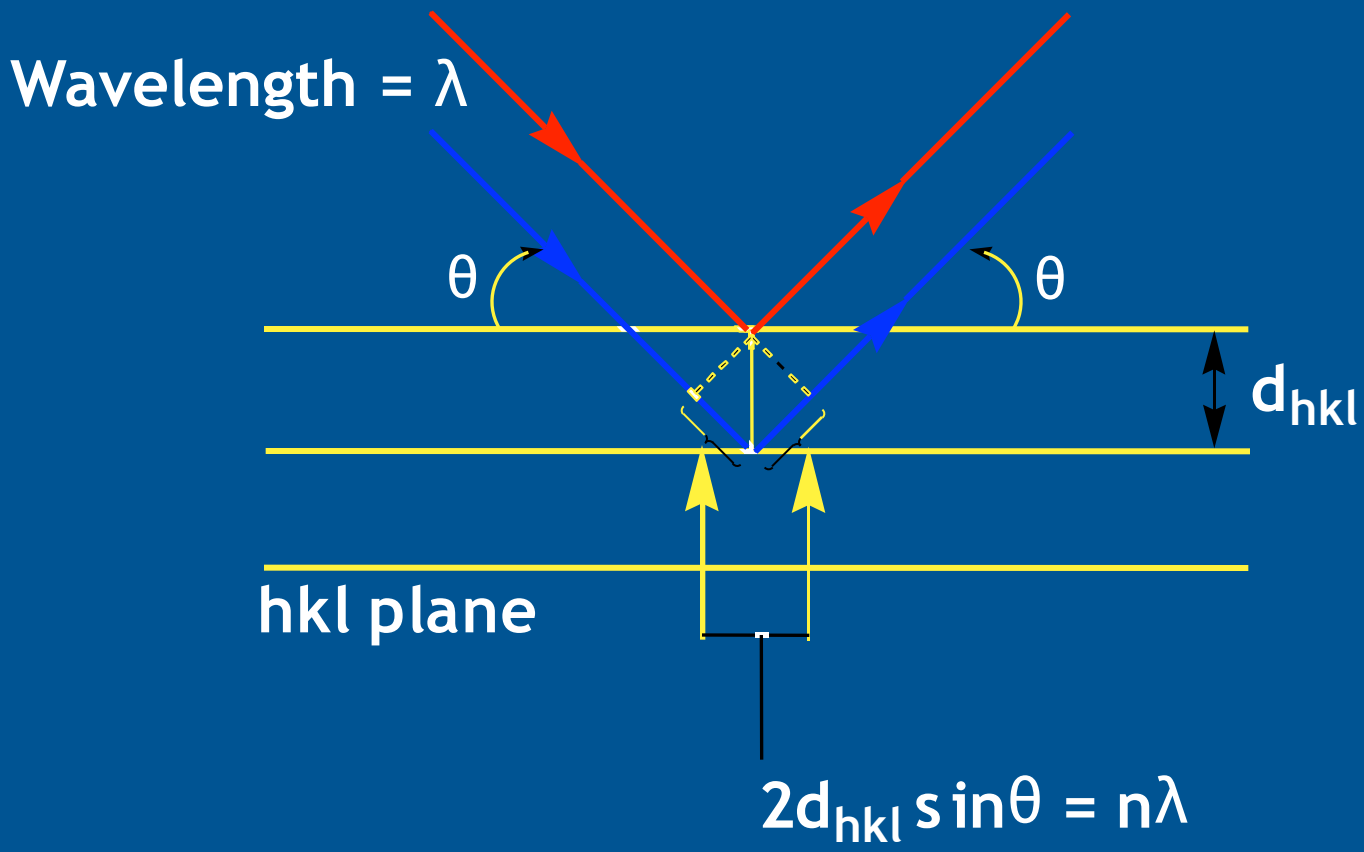


amplitude = A

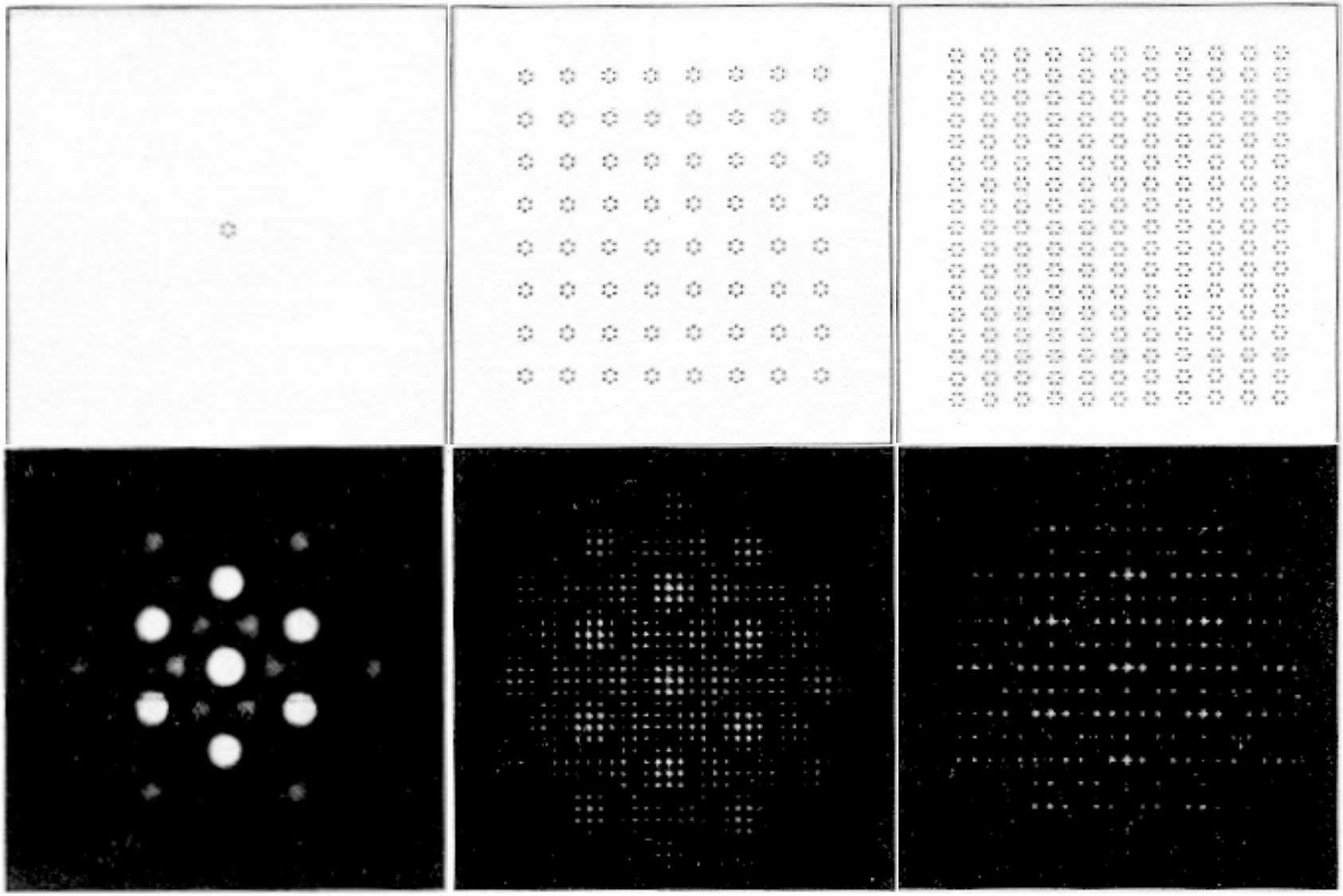
amplitude = 0

Destructive interference

Bragg's law



‘Convolution’ of the contents and the lattice



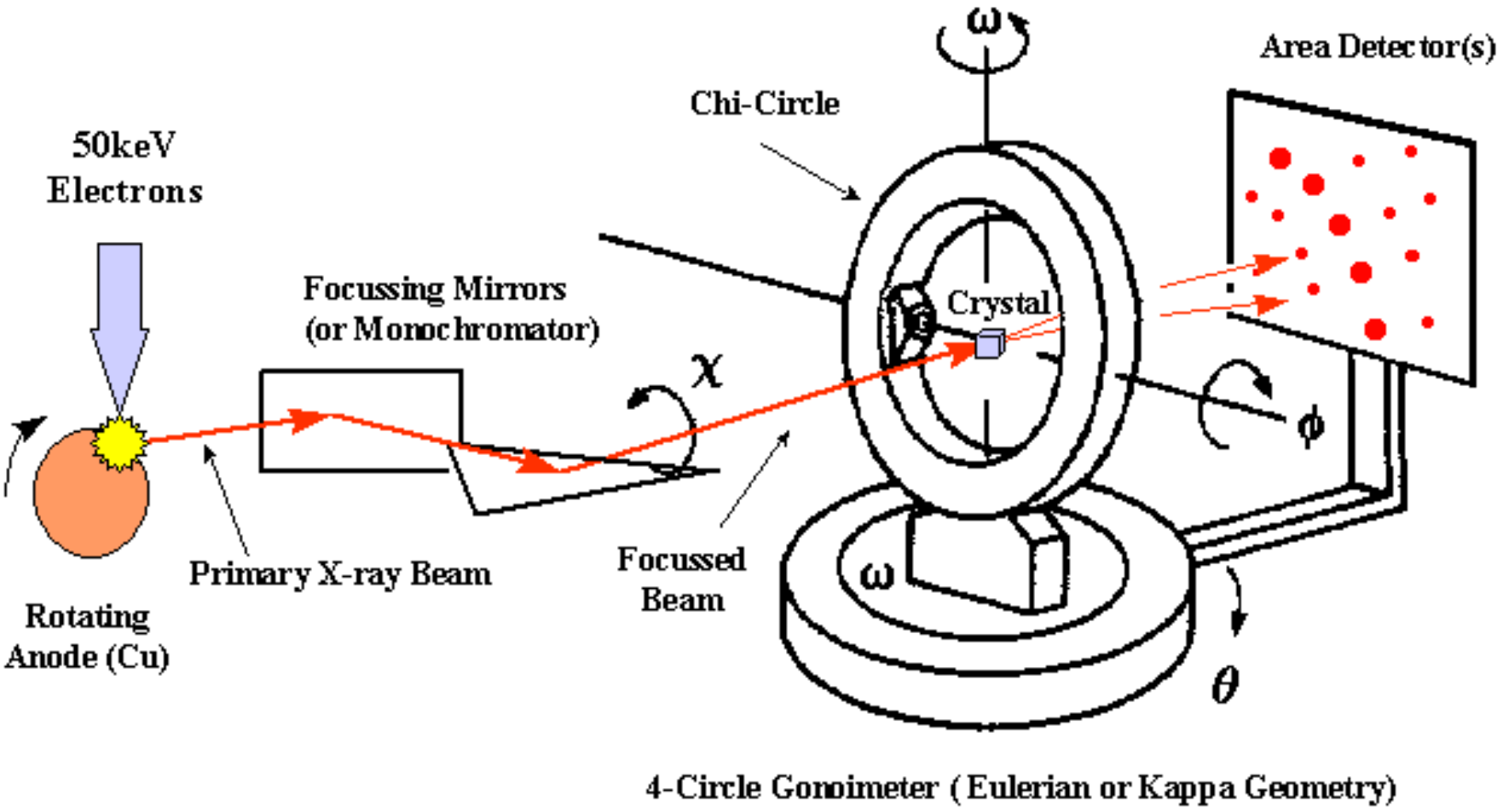
So how does one go about solving a crystal structure?

formulate question	very hard!
make sample	cloning, expression, purification
make crystal	screening, optimisation
collect diffraction data	synchrotron, integration, scaling
solve phase problem	MR, SIRAS, MIRAS, SAD, MAD, hybrid
build model	manual or autotracing
refine model	agreement of model and data
interpret model	very hard! back to top?

... and: might fail at any step!

How to obtain diffraction data from crystals ?

The general crystallographic setup



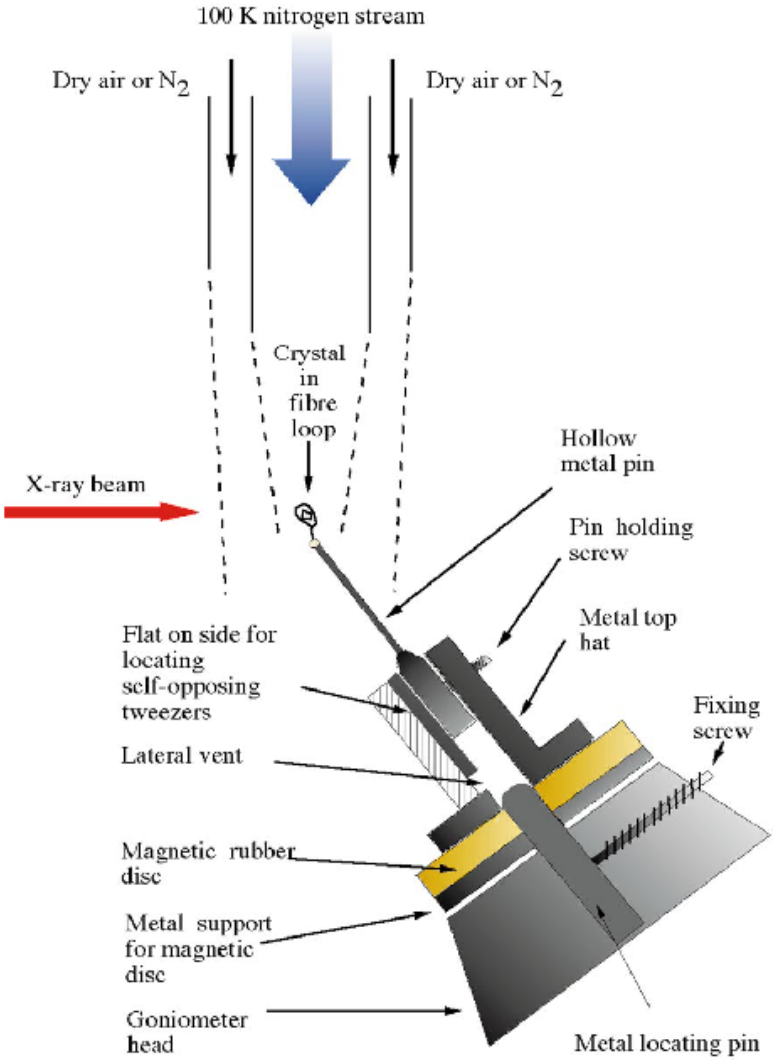
Diffraction hardware

- **X-rays**
  - Laboratory source, rotating anode
    - Electron beam hitting Cu target
    - Fixed wavelength 1.5418Å
  - Synchrotron
    - Circular accelerator + Undulator
    - Much more intense
    - Narrow beam
    - Wavelength tunable
  - Free Electron Laser
    - Linear accelerator + Undulator
    - Even more intense
    - Femtosecond pulses
- **Detector, to produce digital image**
  - Image plate (phosphorimager)
  - CCD (photon coupled)
  - Solid state (direct)

Automated sample loading needed for screening many crystals

at synchrotron and in LMB

Crystals are usually frozen to protect from radiation damage



Each image represents a slice through the ‘reciprocal lattice’

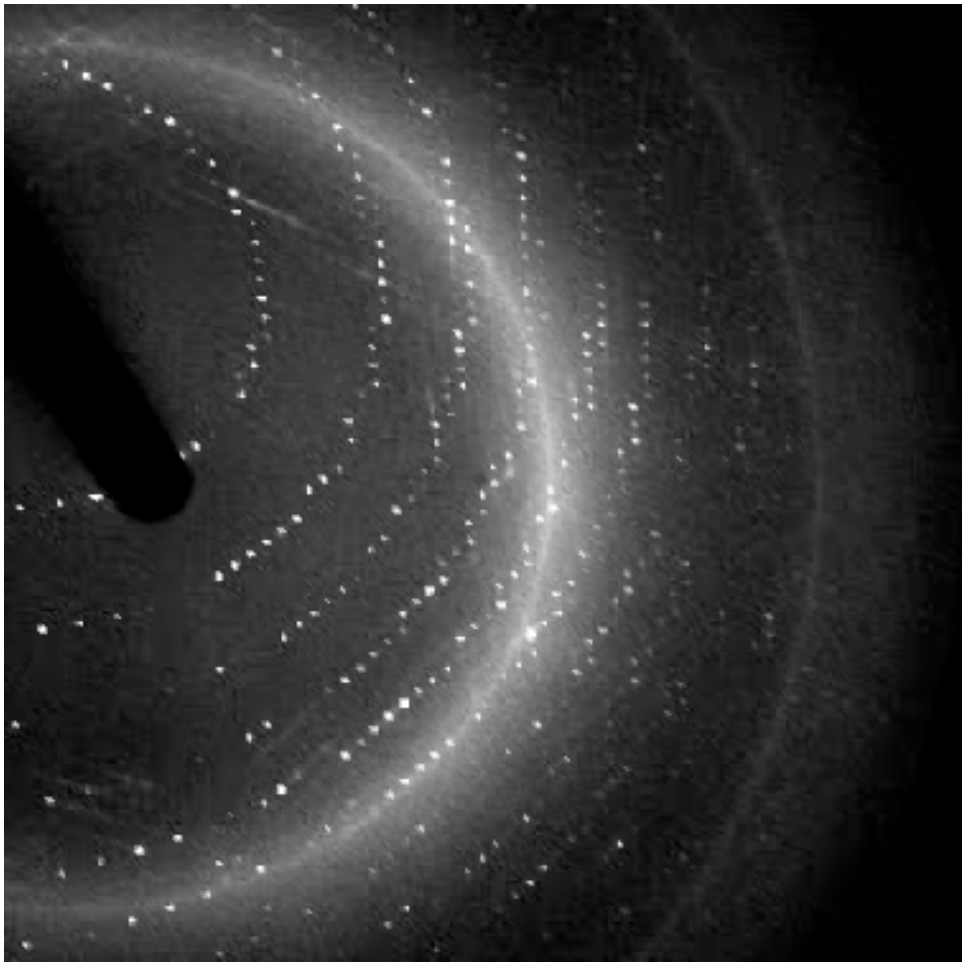
A piece of lead blocks the direct beam

The greater the distance form the centre, the higher the resolution

The amplitudes of the scattered waves are recovered from the intensity of each spot

Each image represents a slice through the ‘reciprocal lattice’

A piece of lead blocks the direct beam

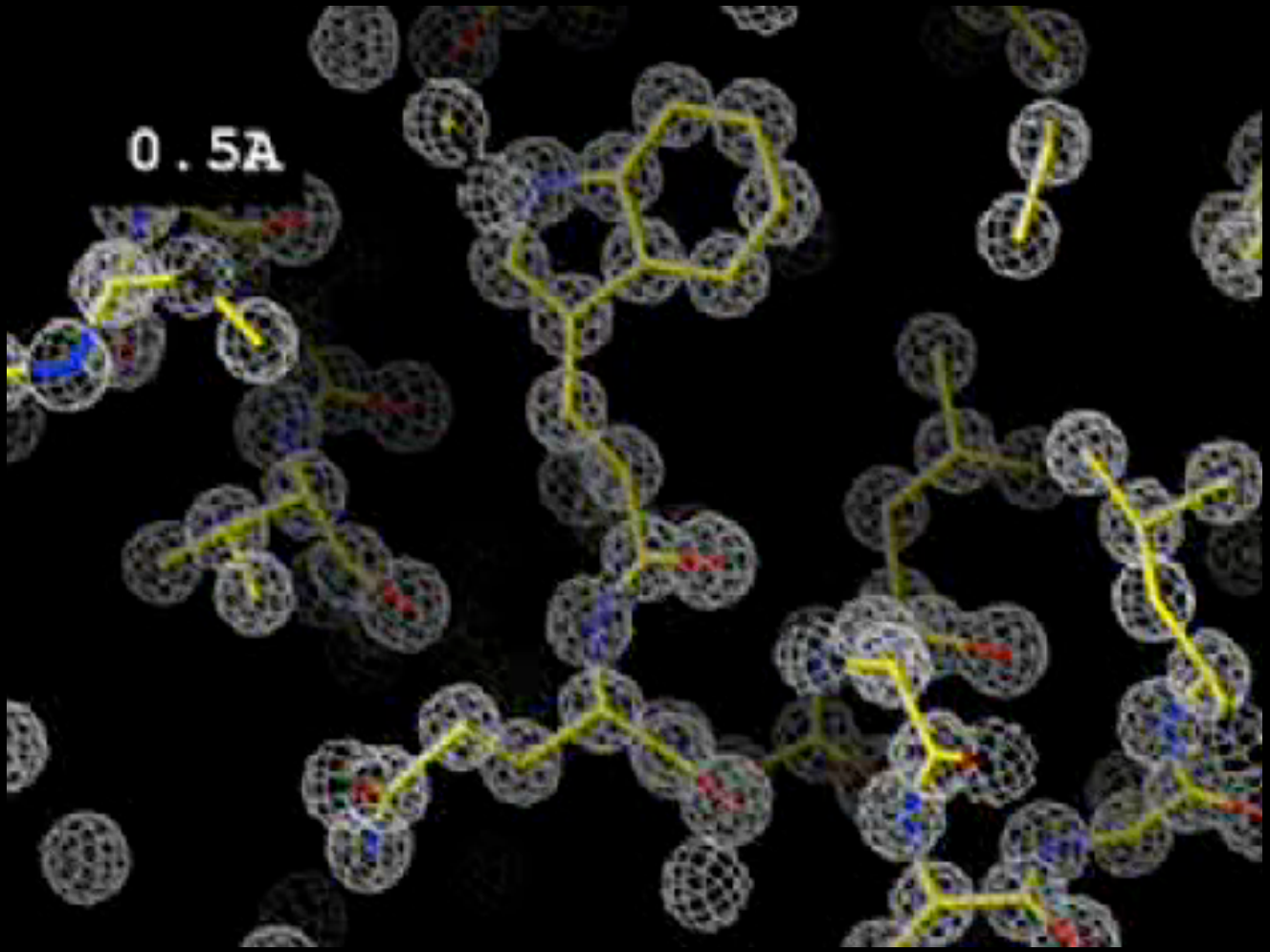


The greater the distance form the centre, the higher the resolution

The amplitudes of the scattered waves are recovered from the intensity of each spot

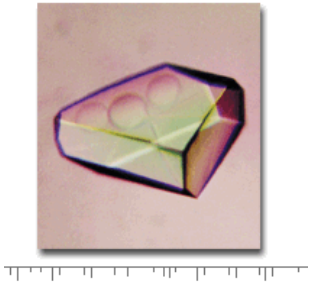
The effect of resolution on structural details

The effect of resolution on structural details

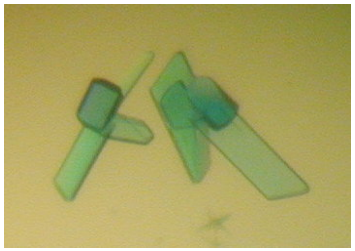


How to obtain crystals ?

Protein crystals are about half water, very fragile and grown from super-saturated solutions

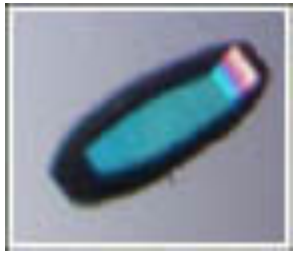


complex of a Myosin V Light Chain with a Heavy Chain fragment

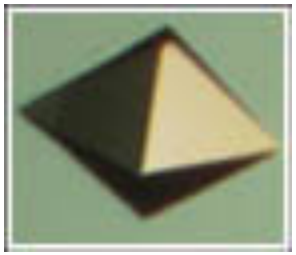


azurin

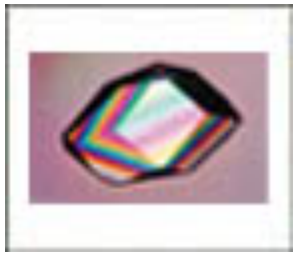
Hex1 protein involved in cell wound healing



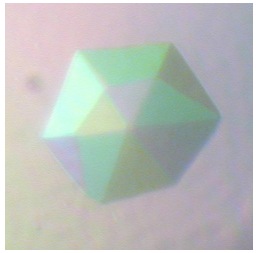
Crystal of "Cobra Venom Factor" isolated and purified from venom of indian cobra (Naja naja sagittifera).



Crystal of a phospholipase A2 monomer isolated from Indian Cobra (Naja naja sagittifera).



Tetragonal Lysozyme

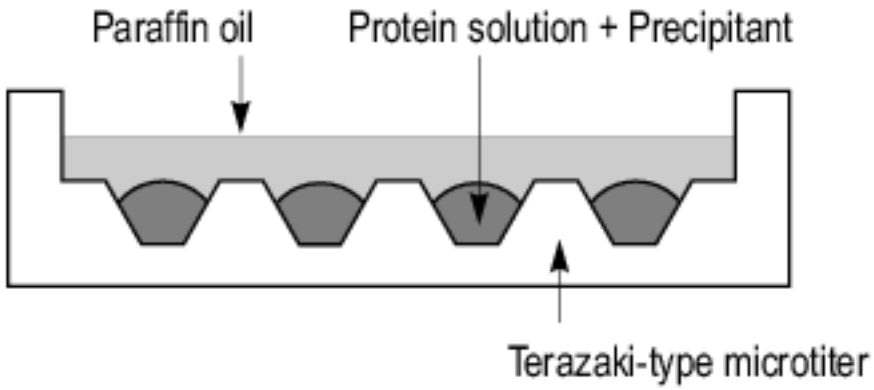


thioredoxin

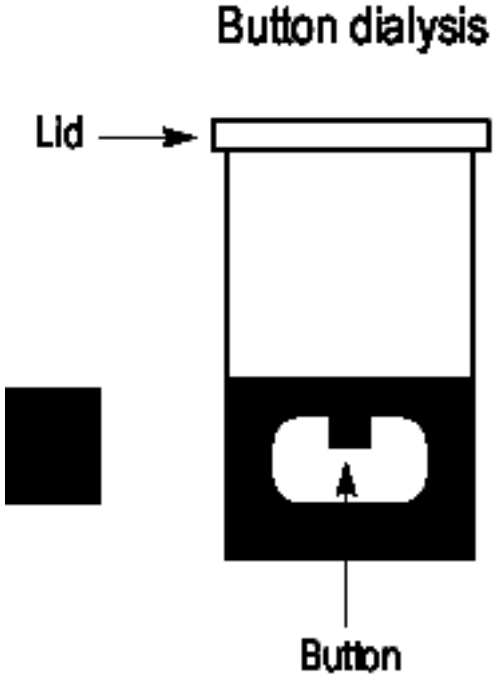
HIV-1 Integrase Core

Protein Crystallisation Techniques

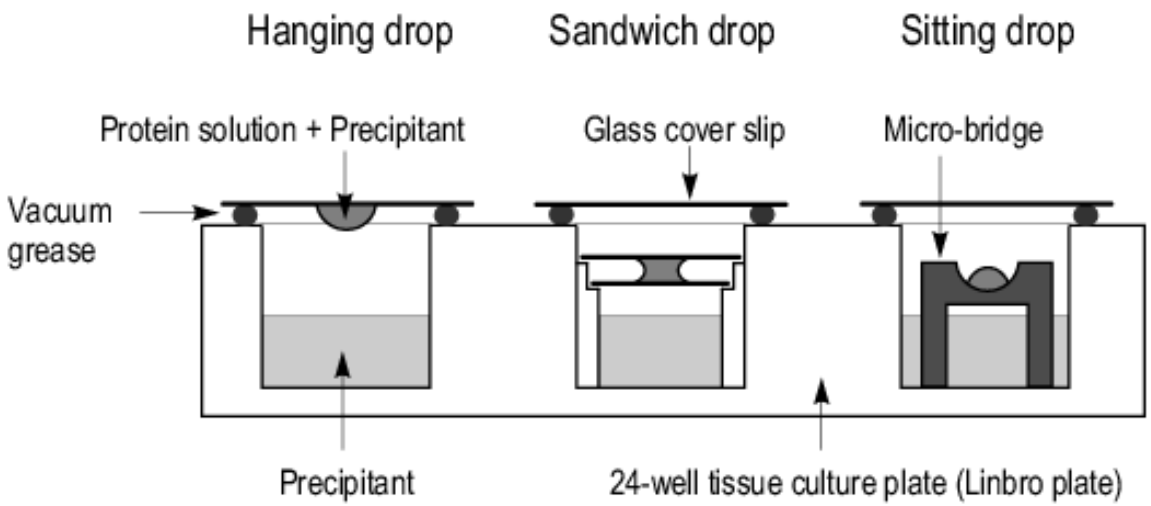
Micro batch

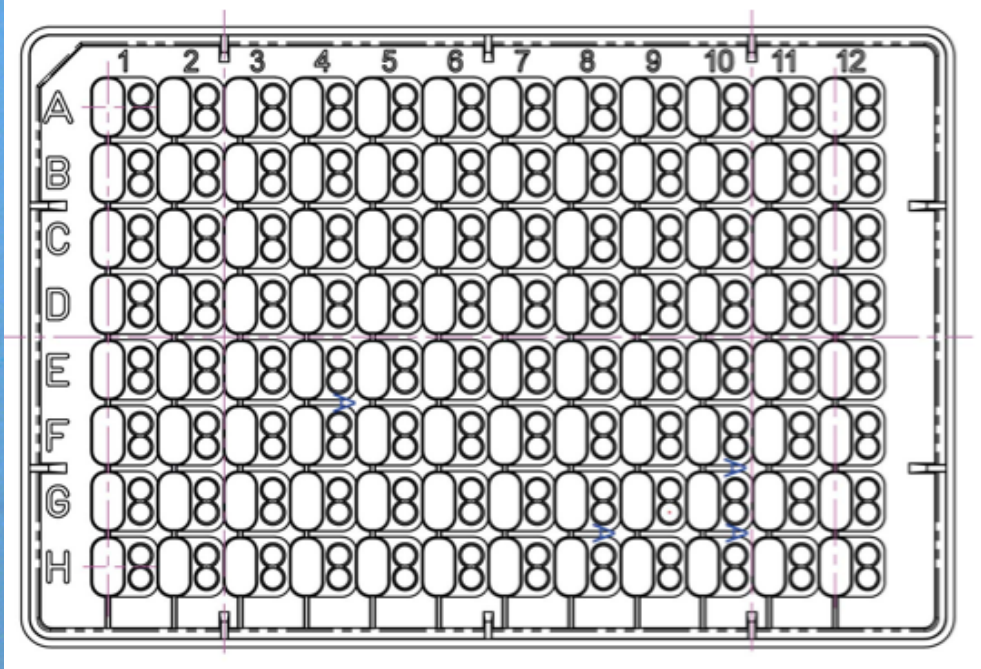


Dialysis



Vapour diffusion

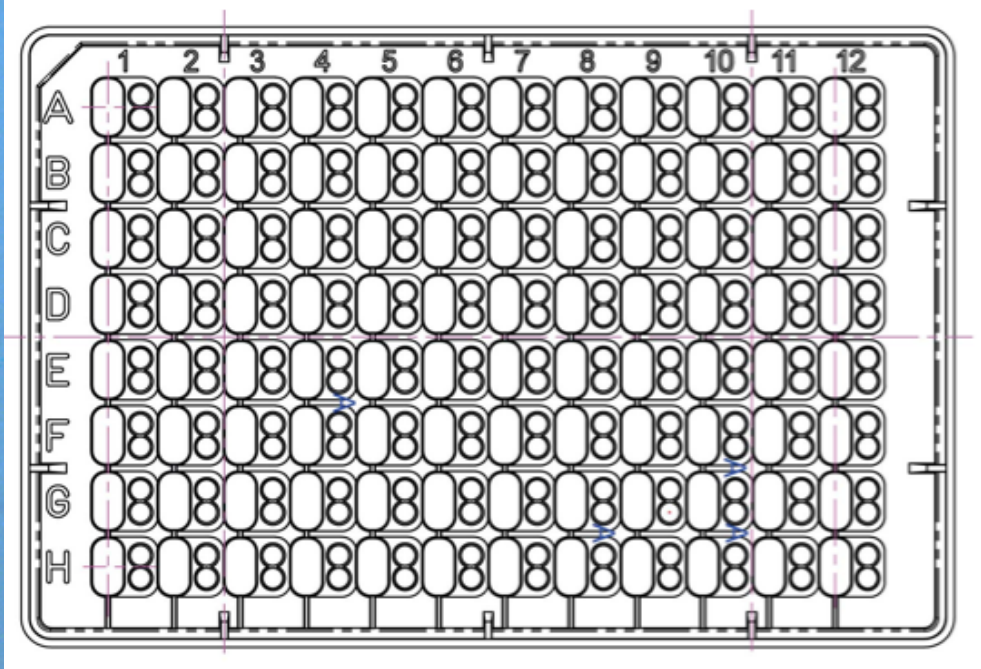



1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H

MRC Laboratory of Molecular Biology

Routine nanolitre protein crystallisation

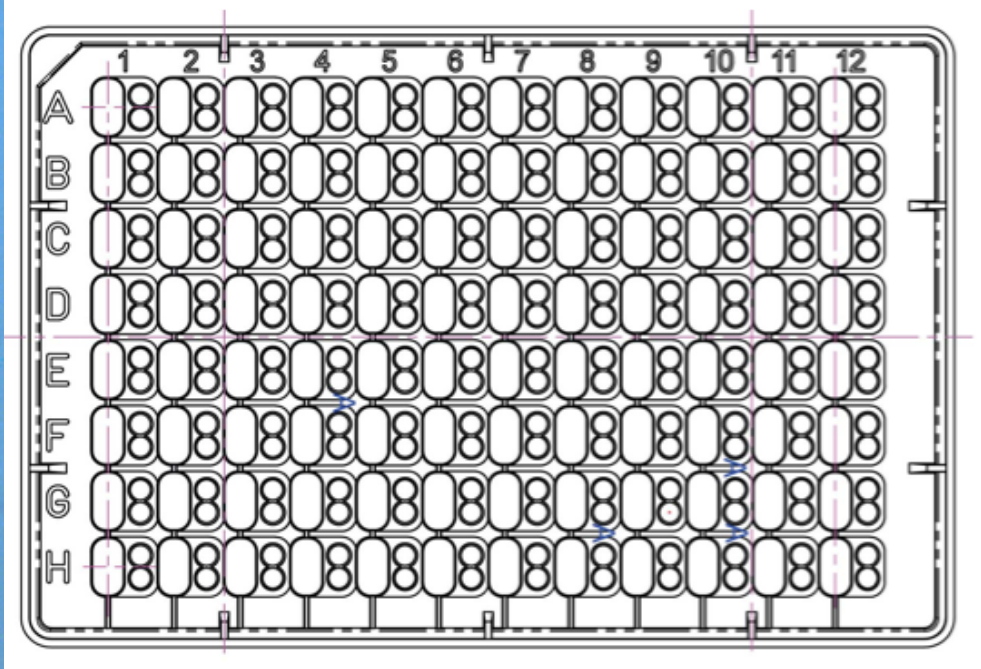
Fabrice Gorrec, Olga Perisic, Jan Löwe



MRC Laboratory of Molecular Biology

Routine nanolitre protein crystallisation

Fabrice Gorrec, Olga Perisic, Jan Löwe

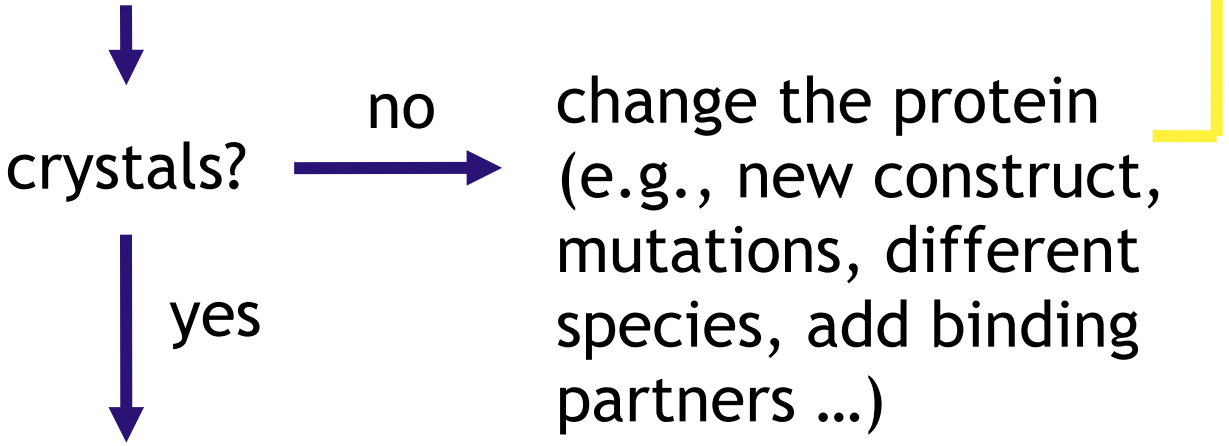


Typical LMB crystallisation strategy

large (1824 conditions) “sparse” matrix screen

1 hour and about 4 mg protein

3 screening robots



optimise conditions (e.g. pH, salt, PEG grids general and specific additives)

3 gridding robots

Protein for X-ray crystallography

- typically about 4 mg for a 1800 condition screen
- highest purity achievable. Affinity, ion exchange and gel filtration chromatography is a common protocol.
- affinity tags should be minimal or cleavable with specific proteases
- typically requires several (very many) constructs

typically requires several constructs

Group	#constructs	#crystals
A	10	0
B	5	0
C	5	0
D	30	8
E	1	1
F	10	2
G	15	2
F	10	3
G	20	2
H	12	2
I	10	8
J	6	1
K	10	0
L	6	4
M	40	4
N	30	4

220 41 = 5-10 constructs per success

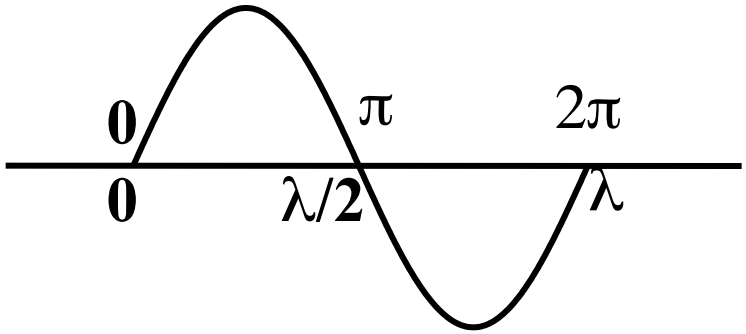
The Phase Problem

Diffraction amplitudes are recovered
from the measured intensity of the spots

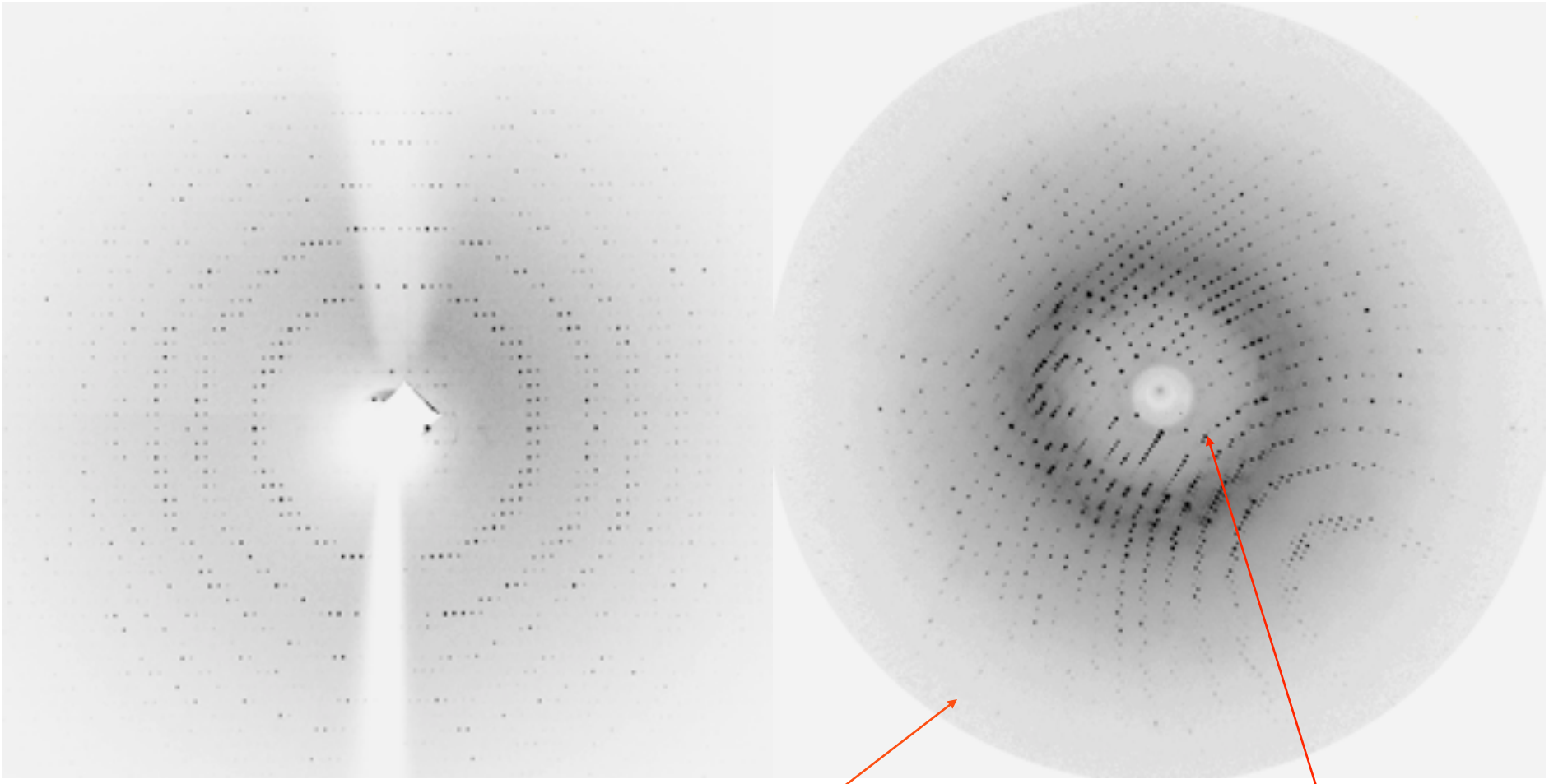
In order to reconstruct an image, we need to know the amplitudes *and phases* of the diffracted waves

X-ray detectors are sensitive to the amplitude only

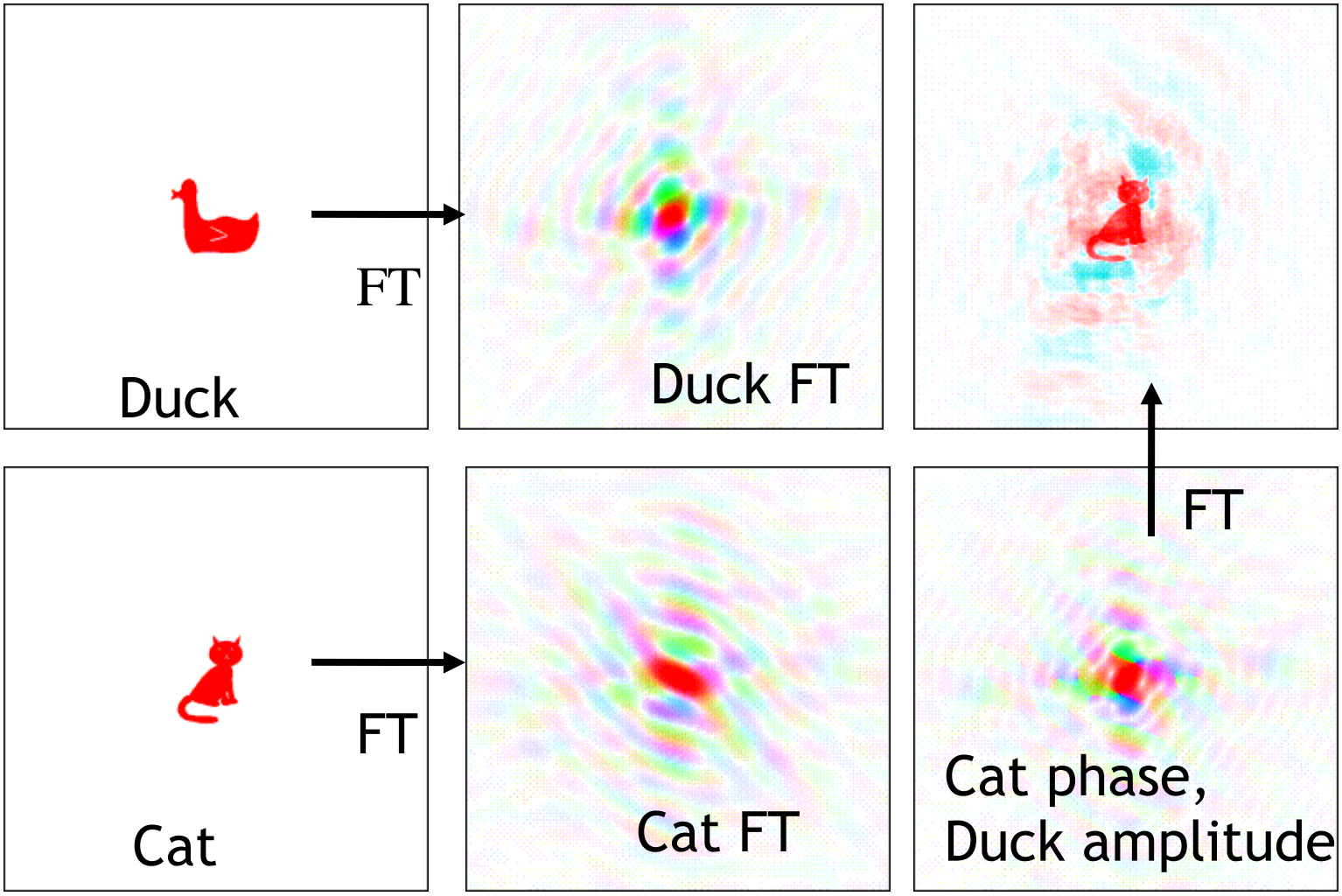
Recovering the lost phases is the "phase problem" of X-ray crystallography



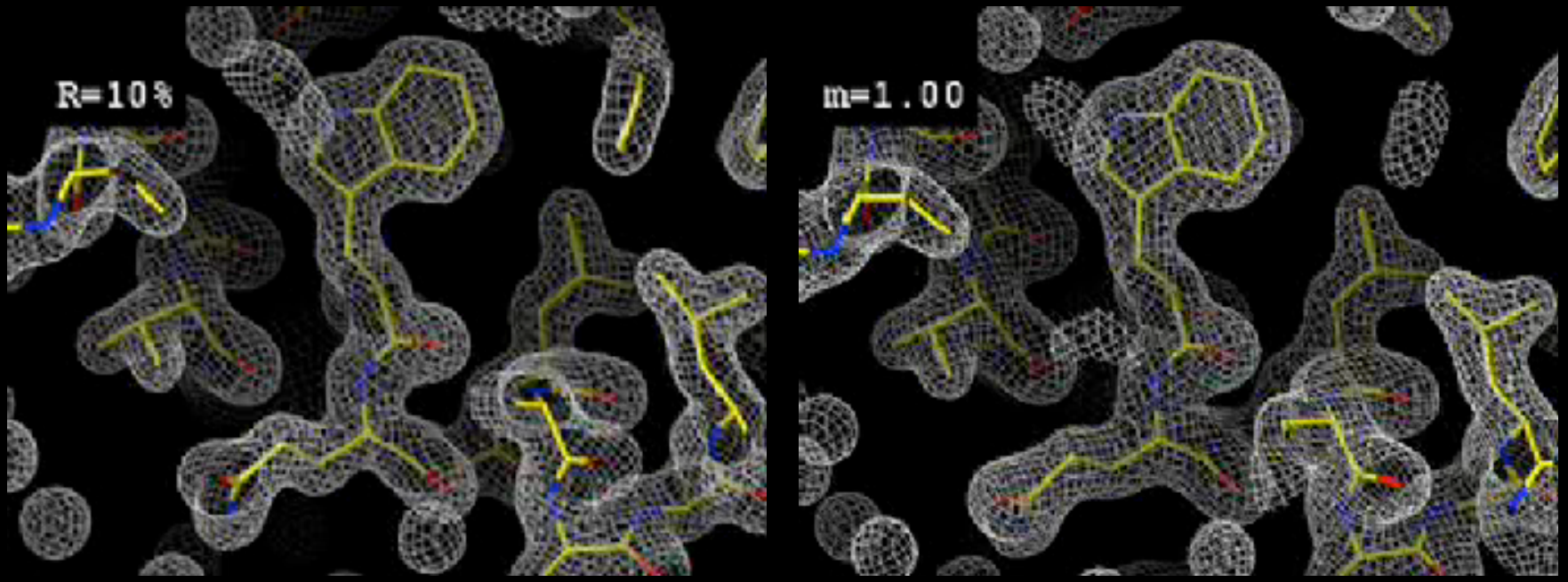
Diffraction patterns



Phases are very important

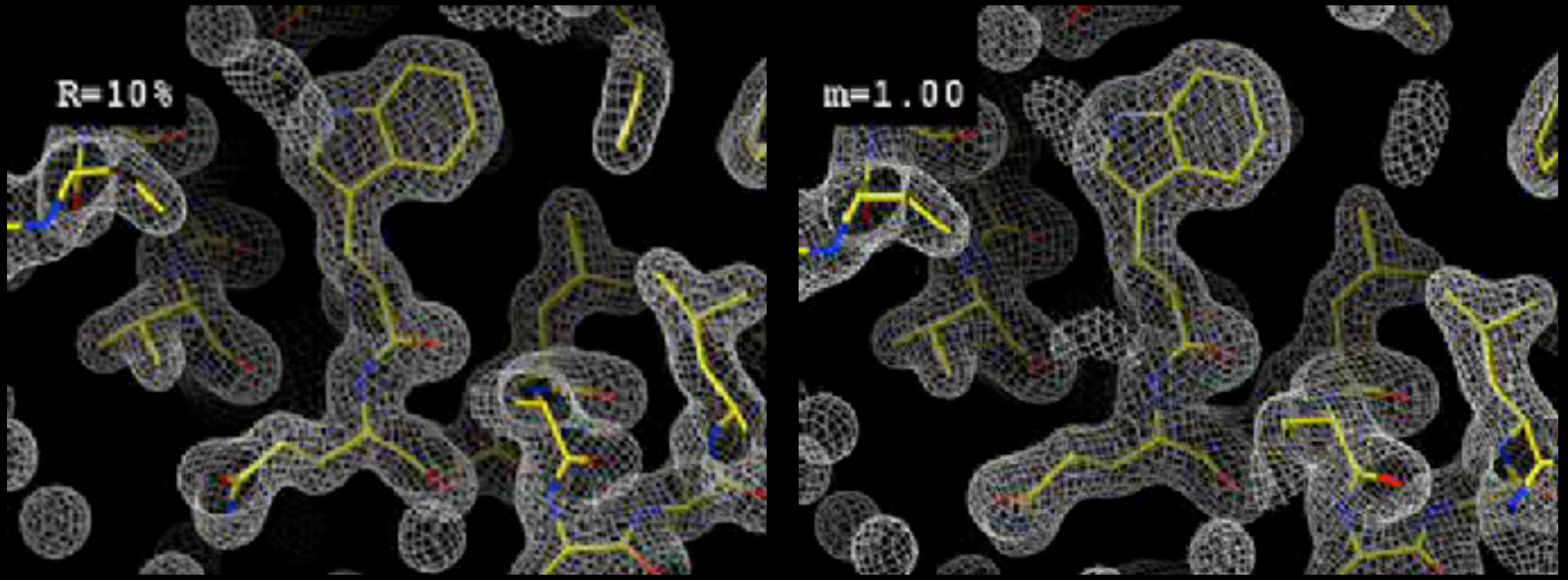


The phases are more important than the amplitudes



The phases are more important than the amplitudes

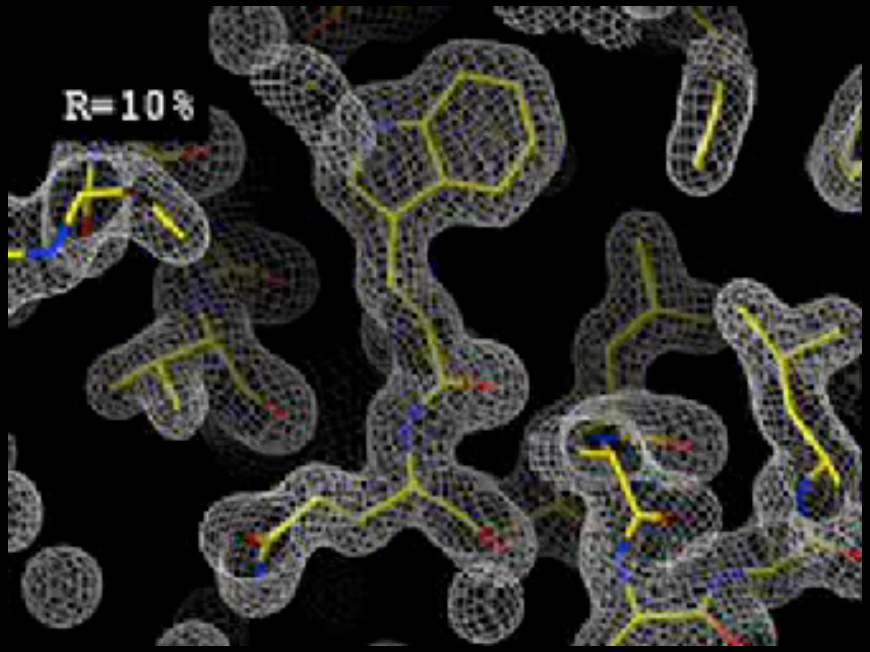
Influence of intensities

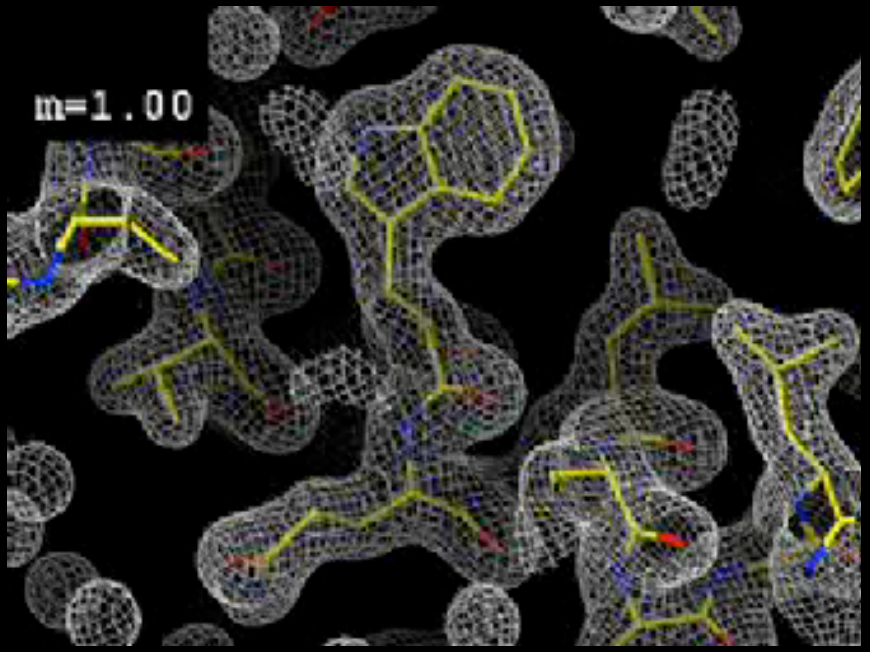


The phases are more important than the amplitudes

Influence of intensities

Influence of phases





The phases are more important than the amplitudes

The crystallographic phase problem is solved via:

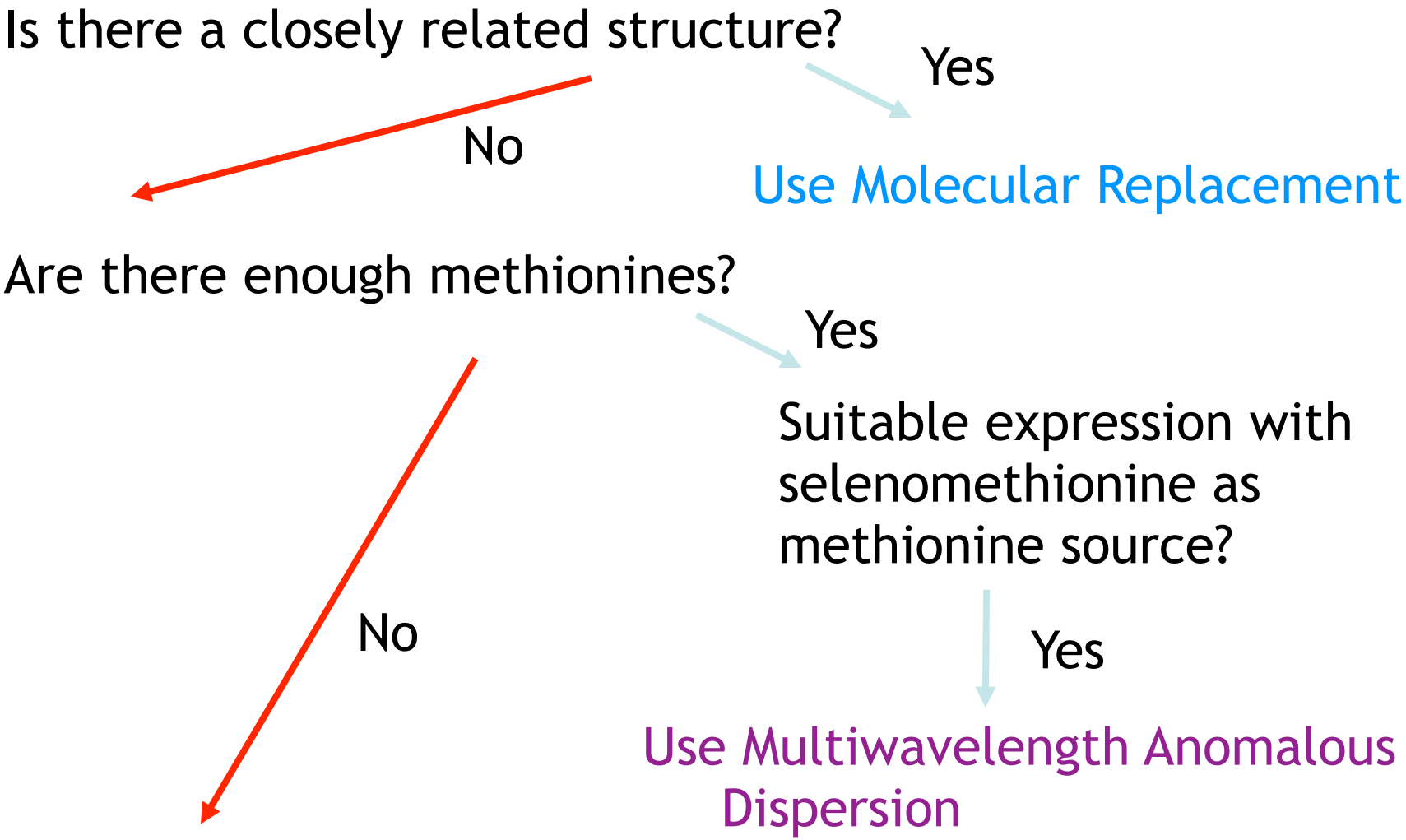
Isomorphous replacement (SIR/MIR)

Anomalous scattering (SAD)

Multiple wavelength anomalous dispersion (MAD)

Molecular replacement (MR)

Difference Fourier methods


Is there a closely related structure?
Yes
No
Use Molecular Replacement
Are there enough methionines?
Yes
Suitable expression with selenomethionine as methionine source?
No
Yes
Use Multiwavelength Anomalous Dispersion
Can I soak heavy atoms into my protein crystals?
Use Multiple Isomorphous Replacement.

New computational methods make locating many heavy atom and Se sites very easy

The current record is 140 Se sites located automatically by SnB

This has made possible the use of very “non-specific” heavy-atom containing reagents (Br, I, Gd-chelates, Xe, Kr…)

Anomalous diffraction
requires tuneable X-rays:

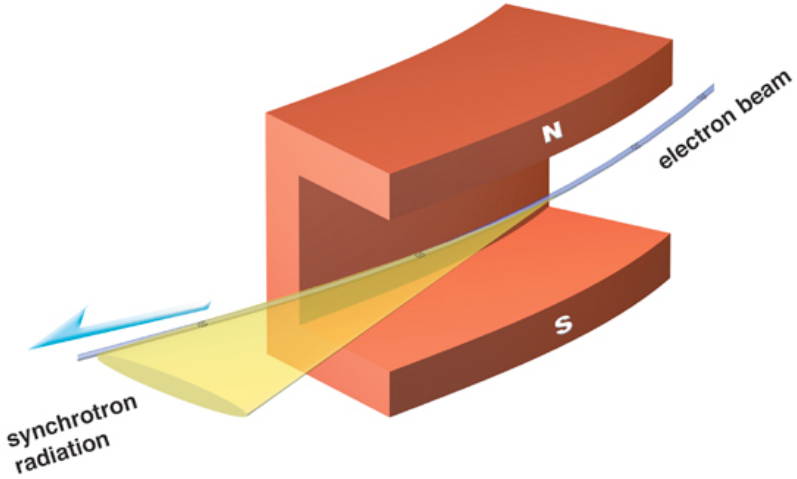
selectable wavelength (~ 0.6 - 2 Å)

Ernest Orlando Lawrence (1901-1958)

Nobel prize in Physics 1939 for the invention and development of the cyclotron and for results obtained with it, especially with regard to artificial radioactive elements

Diamond (Oxford)

ESRF (Grenoble)



synchrotron radiation

bending magnet

undulator

Modern *de-novo* structure determination can be done in a few hours (minutes)

Collect diffraction data for a crystal with anomalous scatters at 1-3 wavelengths

From the data set with the strongest anomalous signal use direct methods to find the locations of the anomalous scatters

Use all of the wavelengths and then calculate the phases for the crystal

Calculate solvent-flattened electron density map

Automatically interpret the map to give a model

but ...

obtaining diffraction quality crystals is normally rate limiting (and can take years)

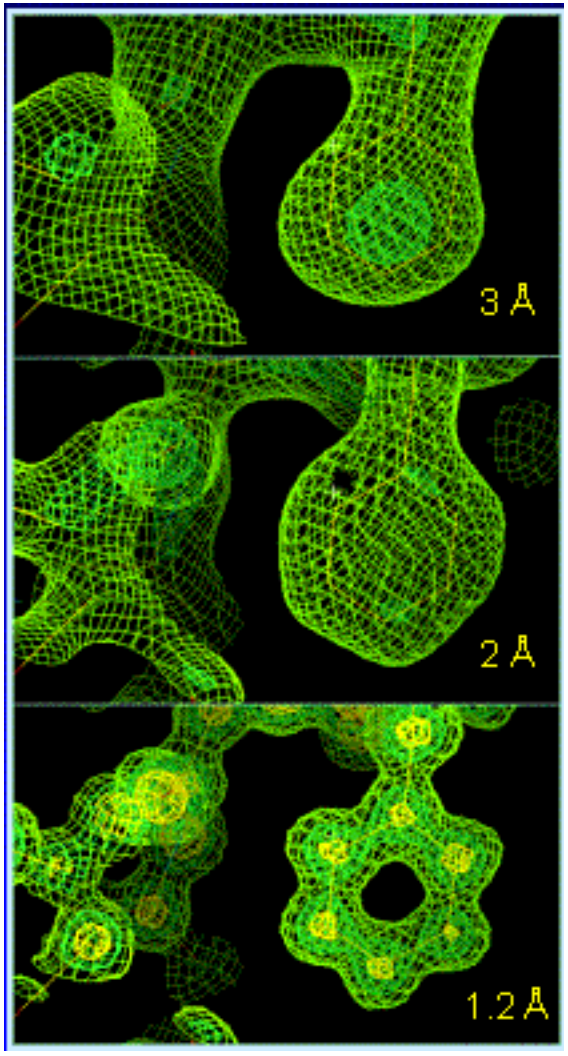
Primary and secondary data:

electron density maps and atomic ‘models’

Electron-density maps

- X-rays are scattered by electrons, so this is what we see
- Calculate phases from heavy-atom positions and observed amplitudes
- Improve phases by solvent-flattening or averaging
- Interpret map by building model (autotrace or hand building)
- Improve model by refinement and rebuilding

Resolution

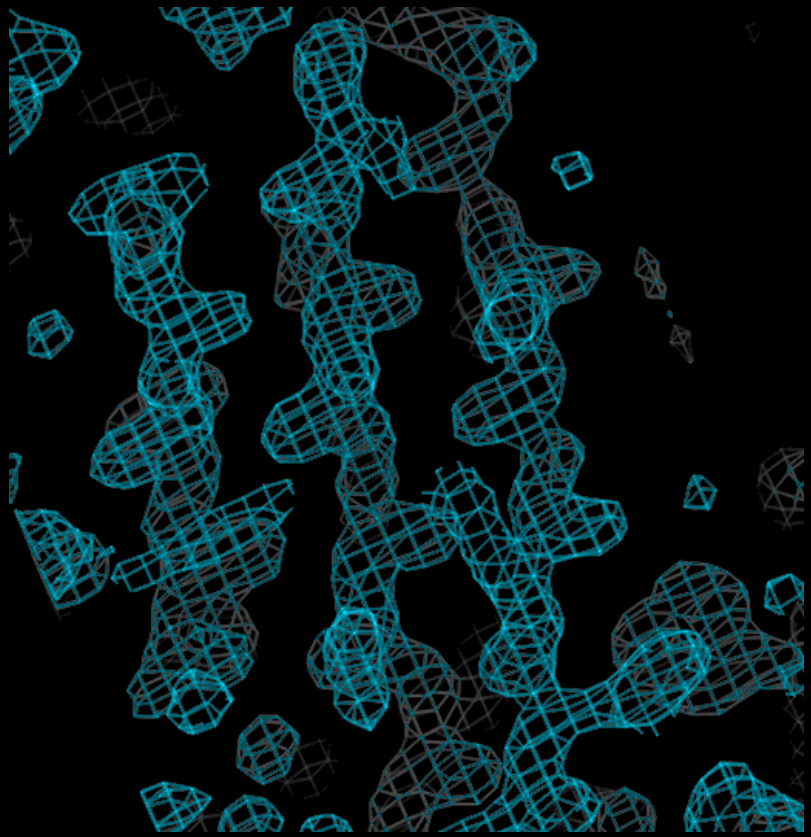


Information about the finest details comes from the highest-angle part of the diffraction pattern.

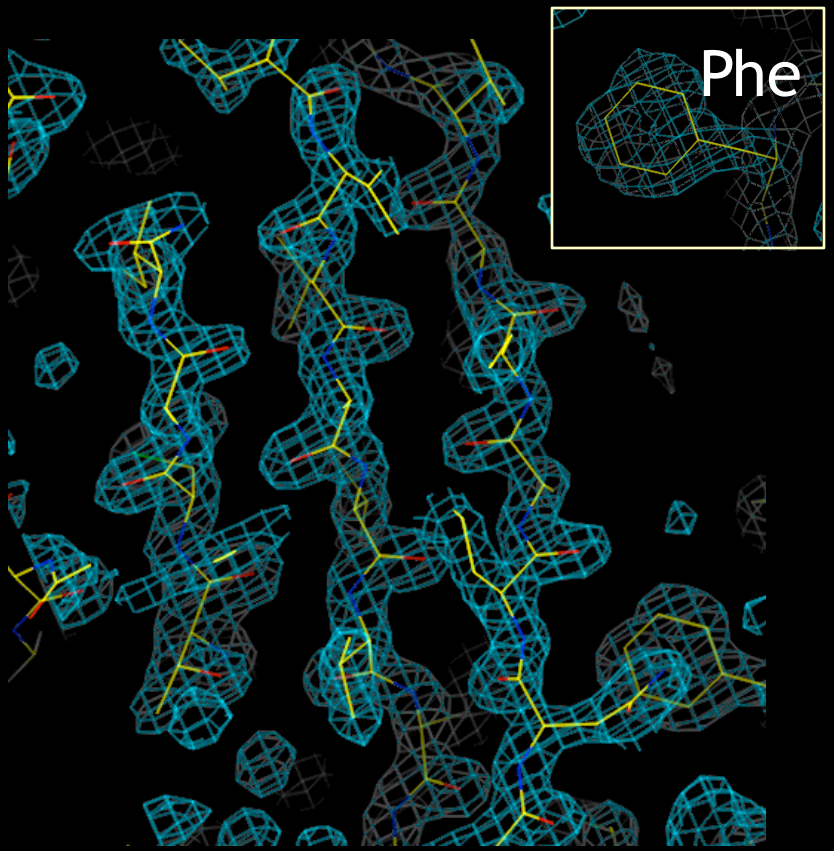
Fine details can only be observed if the crystals are well-ordered, as any difference between molecules, or motion during the experiment, blurs the image.

Low resolution images are less precise, and have more gross errors due to misinterpretation

Electron density maps

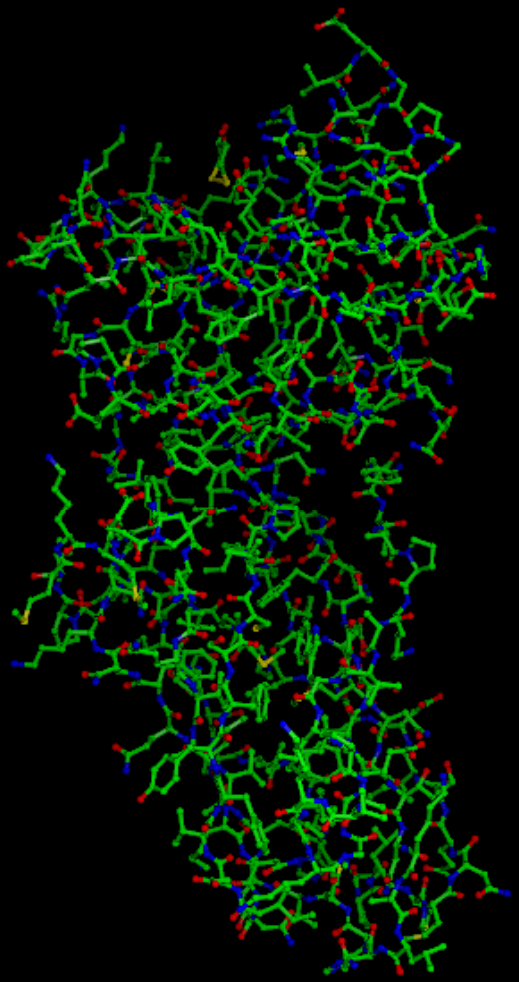


Map at 1.8Å resolution

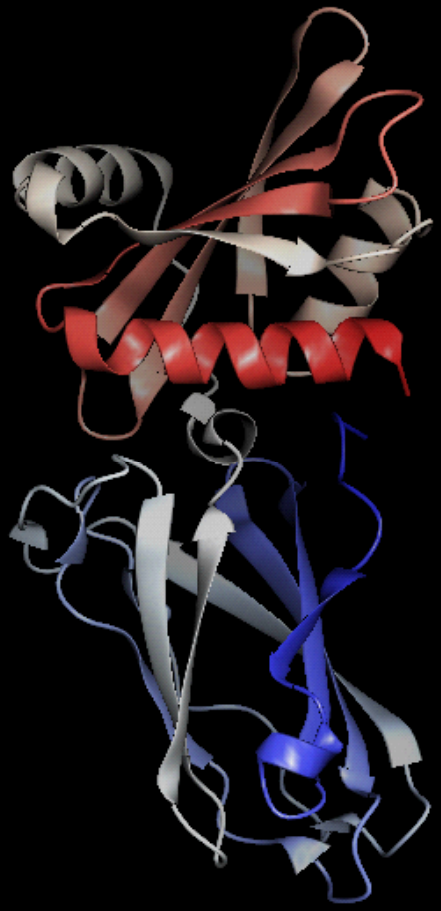


Automatic tracing of map

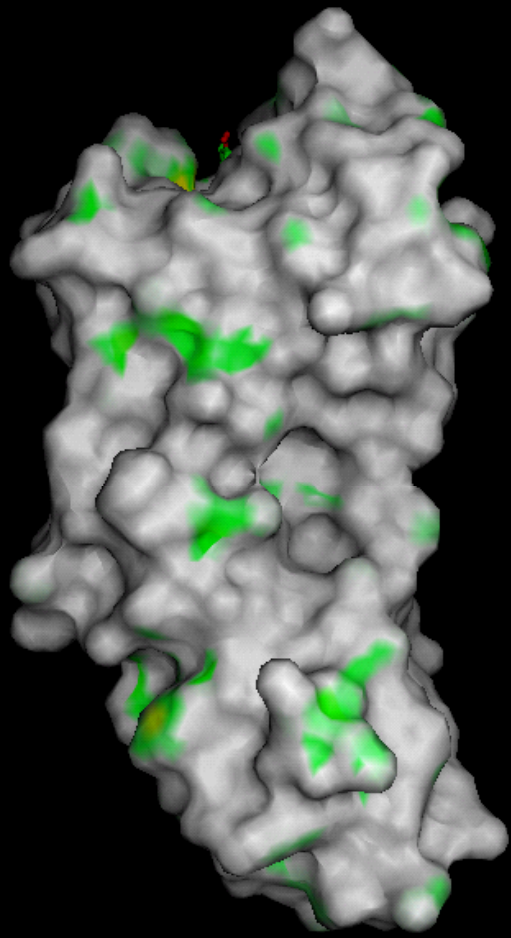
Representations



All atoms



Ribbons



Surface

... when drawing conclusions, always consult the primary data:

diffraction data
electron density maps
crystal packing

New developments

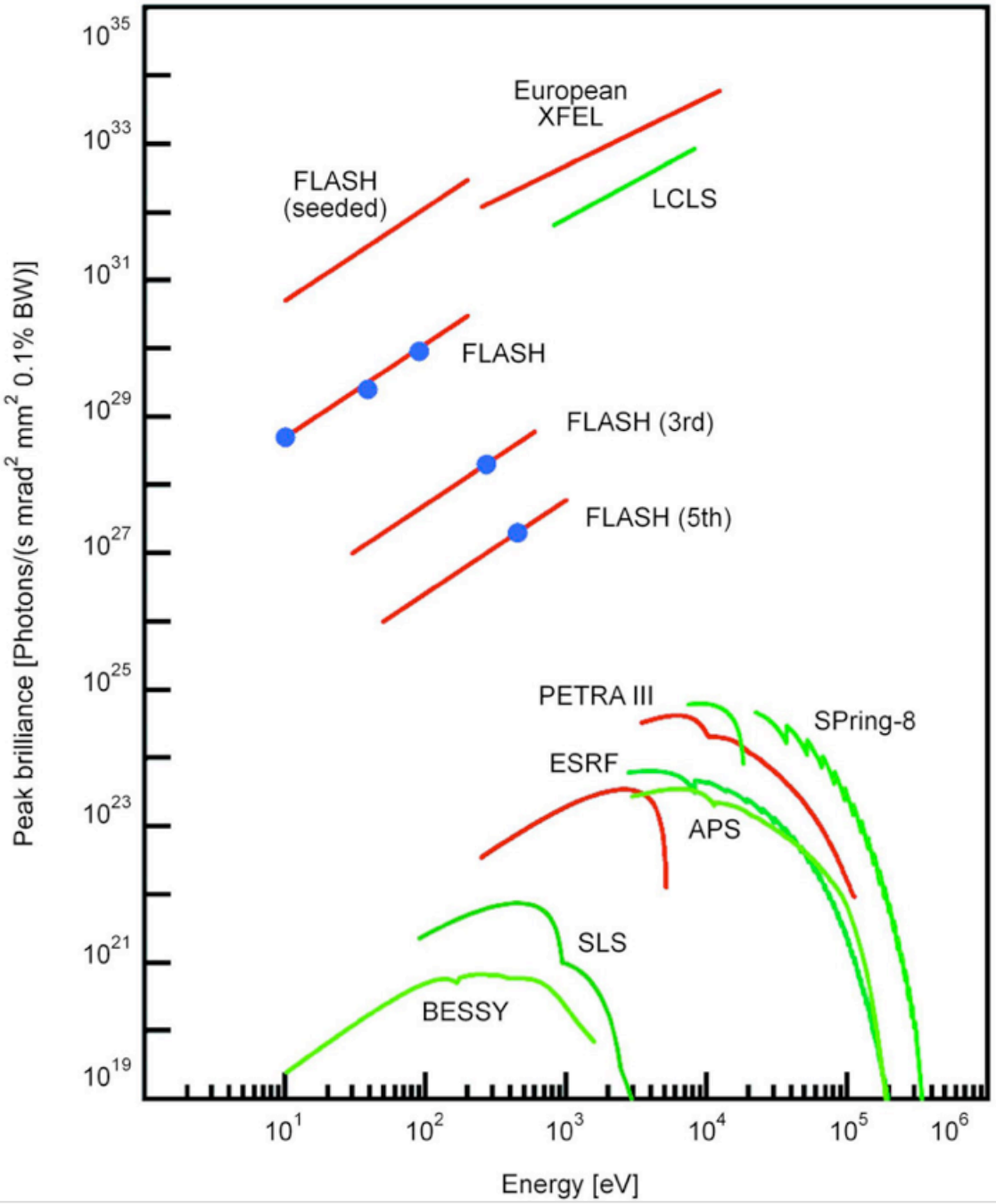
The future of crystallography may be to abandon crystals

Damage is the essential limitation for both X-ray and EM

With crystalline samples, as long as only a small fraction of molecules are damaged, the crystal is useful

With single molecules or small clusters, limitations are severe

X-ray structures of non-crystalline samples will require much more intense X-ray sources.

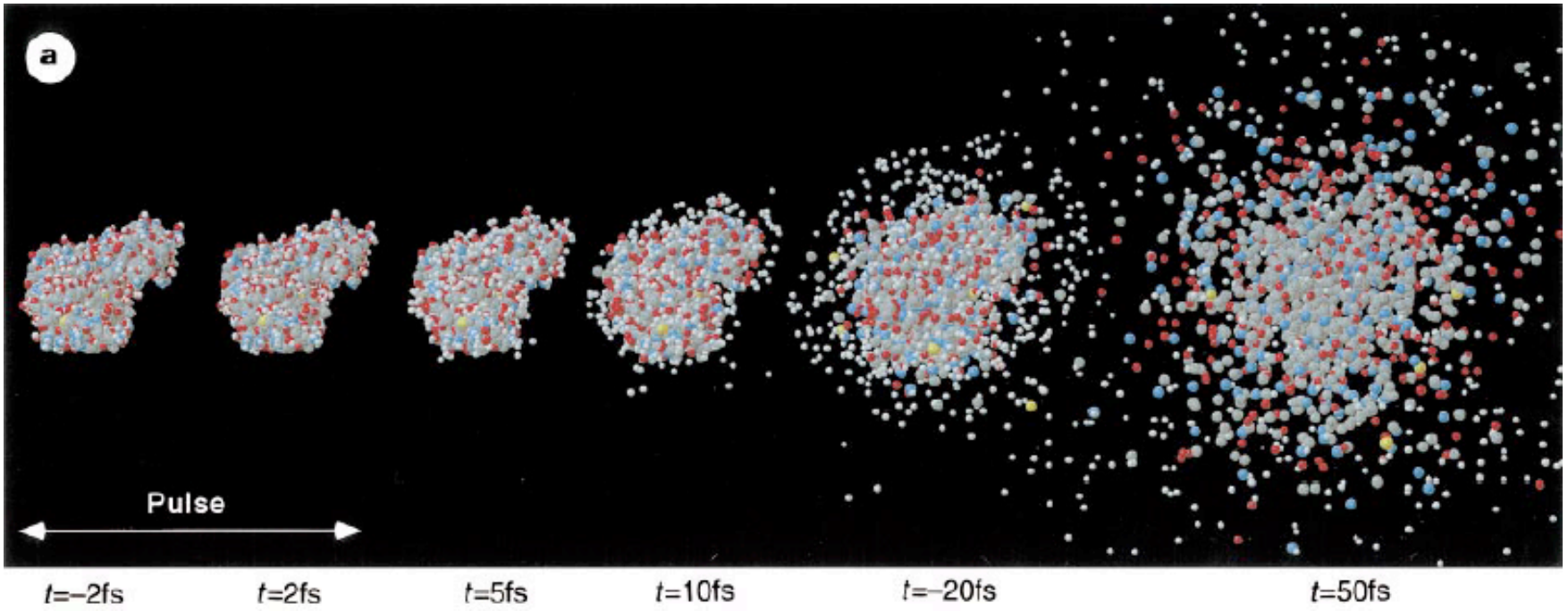


synchrotrons and FELs

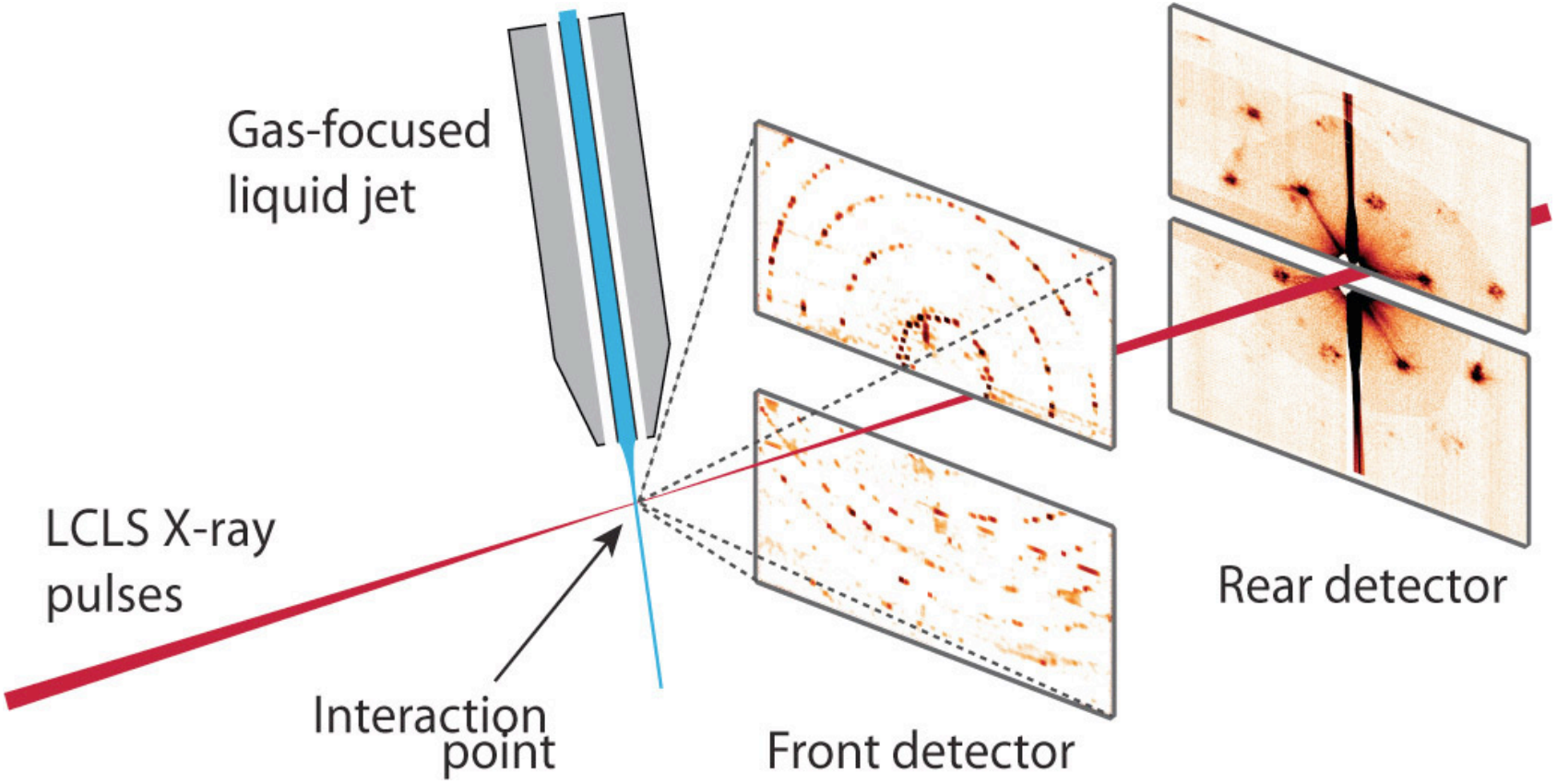
Speed is the answer

Single particle analysis using free electron lasers might be useful

If data are collected faster than damage can occur



Serial Crystallography



John Spence et al.