Physics of Life

PHYS-468

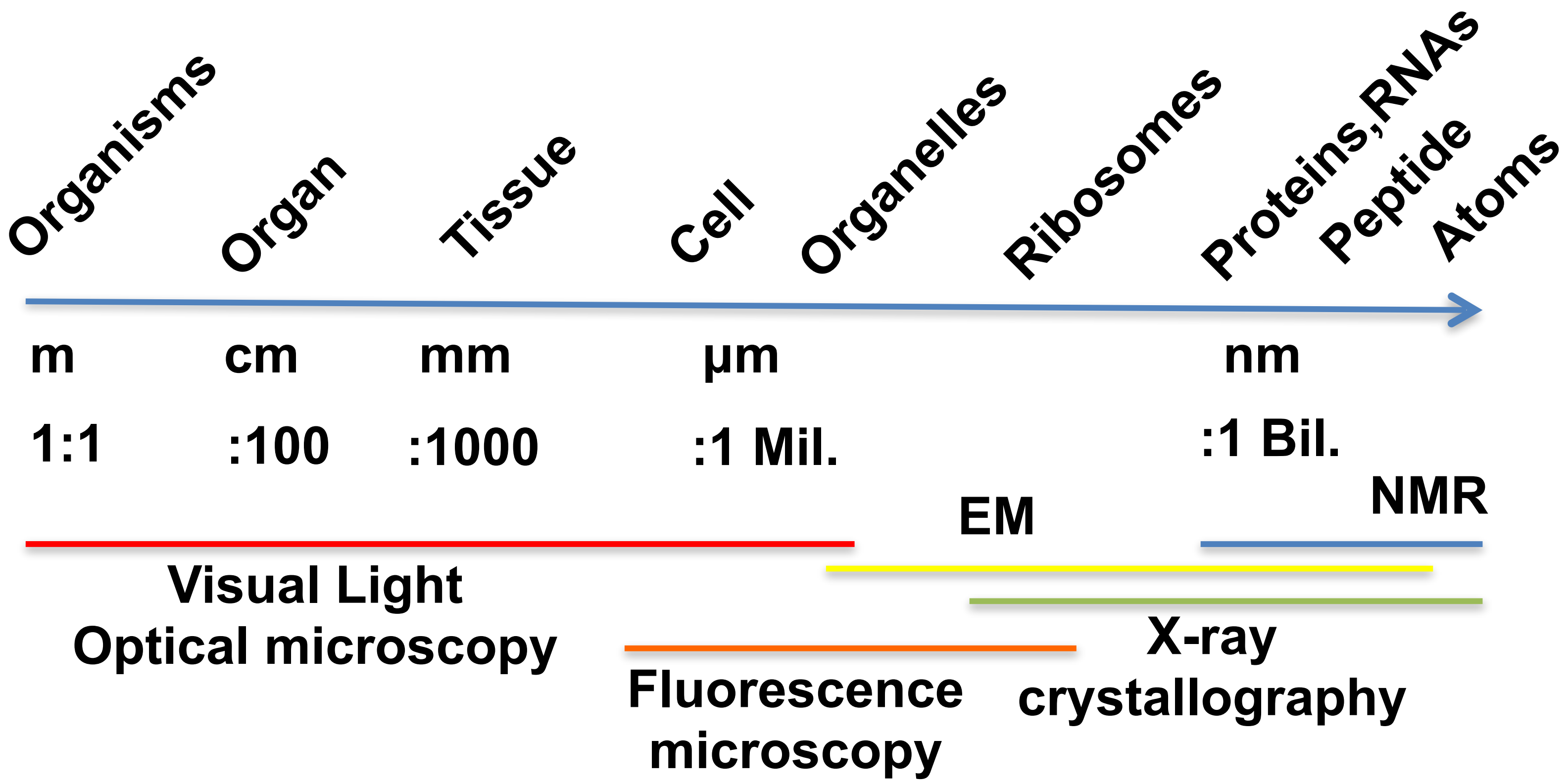
X-Ray Diffraction

Henning Stahlberg,
LBEM, IPHYS, SB, EPFL

Syllabus X-ray Crystallography

- Dimensions of proteins, nucleic acids and wavelengths
- 100 years of crystallography
- X-ray diffraction vs. microscopy
- X-ray crystallography workflow
- Crystallization
  - properties of biological crystals
  - crystallization principles and techniques
  - crystallization artifacts
- X-ray data collection
  - properties of X-rays
  - generation and detection of X-rays
  - radiation damage
- Diffraction
  - Single crystal diffraction
  - Bragg’s law
  - Information content of position and intensity of reflections
  - Resolution
  - Images to amplitudes
- From amplitudes to model
  - Solving the phase problem (heavy atom phasing and molecular replacement)
  - Calculation of electron density
  - From electron density to model
  - Resolution and molecular detail
  - Deposition and validation in Protein Data Bank
- Opportunities and limitations

Techniques in Structural Biology



Organisms
Organ
Tissue
Cell
Organelles
Ribosomes
Proteins,RNAs
Peptide
Atoms

m
cm
mm
µm
nm

1:1
:100
:1000
:1 Mil.
:1 Bil.

Visual Light
Optical microscopy

Fluorescence microscopy

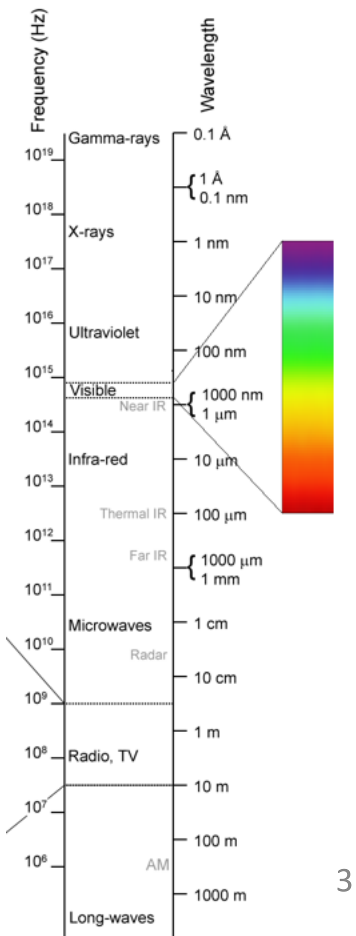
EM

X-ray crystallography

NMR

Wavelength used in:

- Light Microscopy: 300 nm – 800 nm
- X-ray crystallography: 0.1-0.2 nm (1-2Å)
- Electron Microscopy: 0.002-0.003 nm



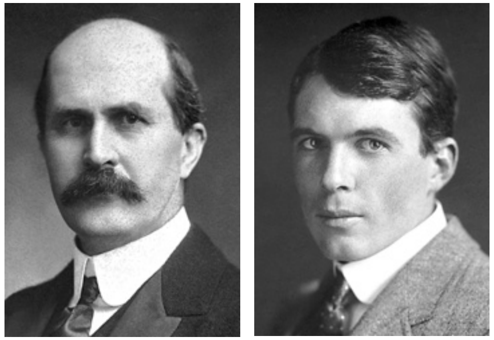
100 years of X-ray Crystallography

1901 Physics: Wilhelm Conrad Röntgen

Discovery of X-rays

1914 Physics: Max von Laue

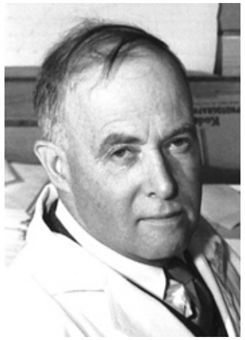
Diffraction of X-rays by crystals



1915 Physics: William Henry Bragg,
William Lawrence Bragg

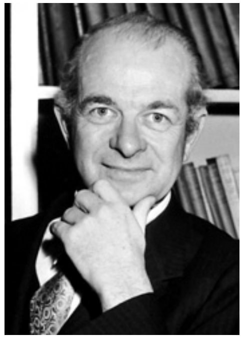
Use of X-rays to determine crystal structures

100 years of Crystallography



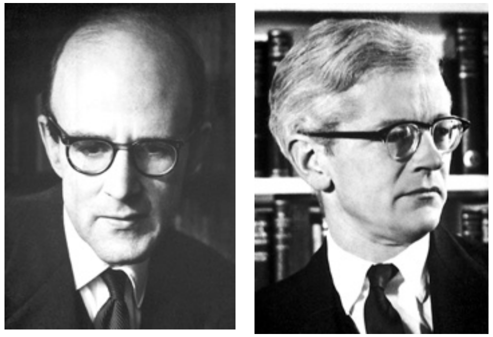
1946 Chemistry: James Sumner

For his discovery that enzymes can be crystallised (urease)



1954 Chemistry: Linus C. Pauling

For his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances



1962 Chemistry: Max F. Perutz, John Kendrew

For their studies of the structures of globular proteins (Haemoglobin, Chymotrypsin; Myoglobin)

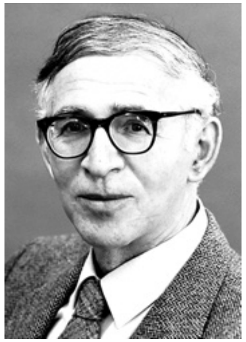
1962 Medicine:
Francis Crick, James Watson, Maurice Wilkins

The helical structure of DNA

100 years of Crystallography

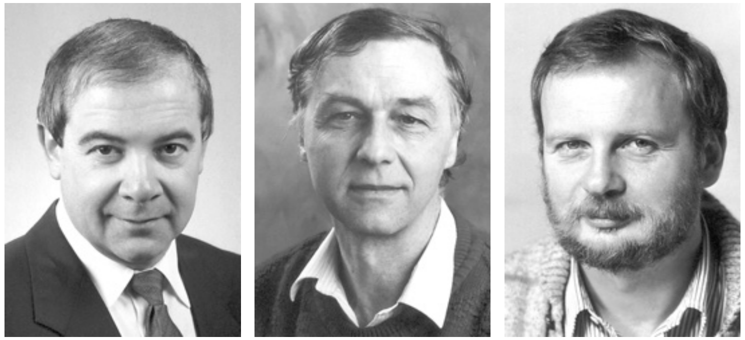
1972 Chemistry: Christian B. Anfinsen

Folding of protein chains: Sequence determines Structure (Ribonuclease)



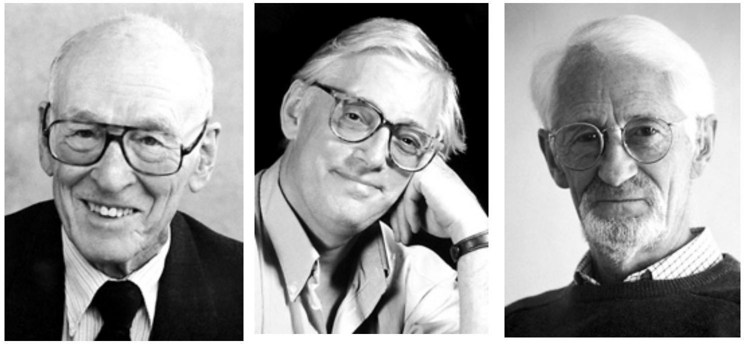
1982 Chemistry: Aaron Klug

Development of crystallographic electron microscopy, discovery of the structure of biologically important nucleic acid-protein complexes (TMV , Nucleosom)



1988 Chemistry: Johann Deisenhofer, Robert Huber, Hartmut Michel

For the determination of the three-dimensional structure of a photosynthetic reaction centre



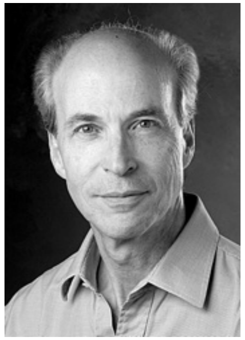
1997 Chemistry: Paul Boyer, John E Walker, Jens C. Skou

Elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate (ATP) and discovery of an ion-transporting enzyme

100 years of Crystallography

2003 Chemistry: Peter Agre, Roderick MacKinnon

Aquaporins; Potassium channels



2006 Chemistry: Roger D. Kornberg

Studies of the molecular basis of eukaryotic transcription

2009 Chemistry:
Venki Ramakrishnan, Tom Steitz, Ada Yonath

Studies of the structure and function of the ribosome

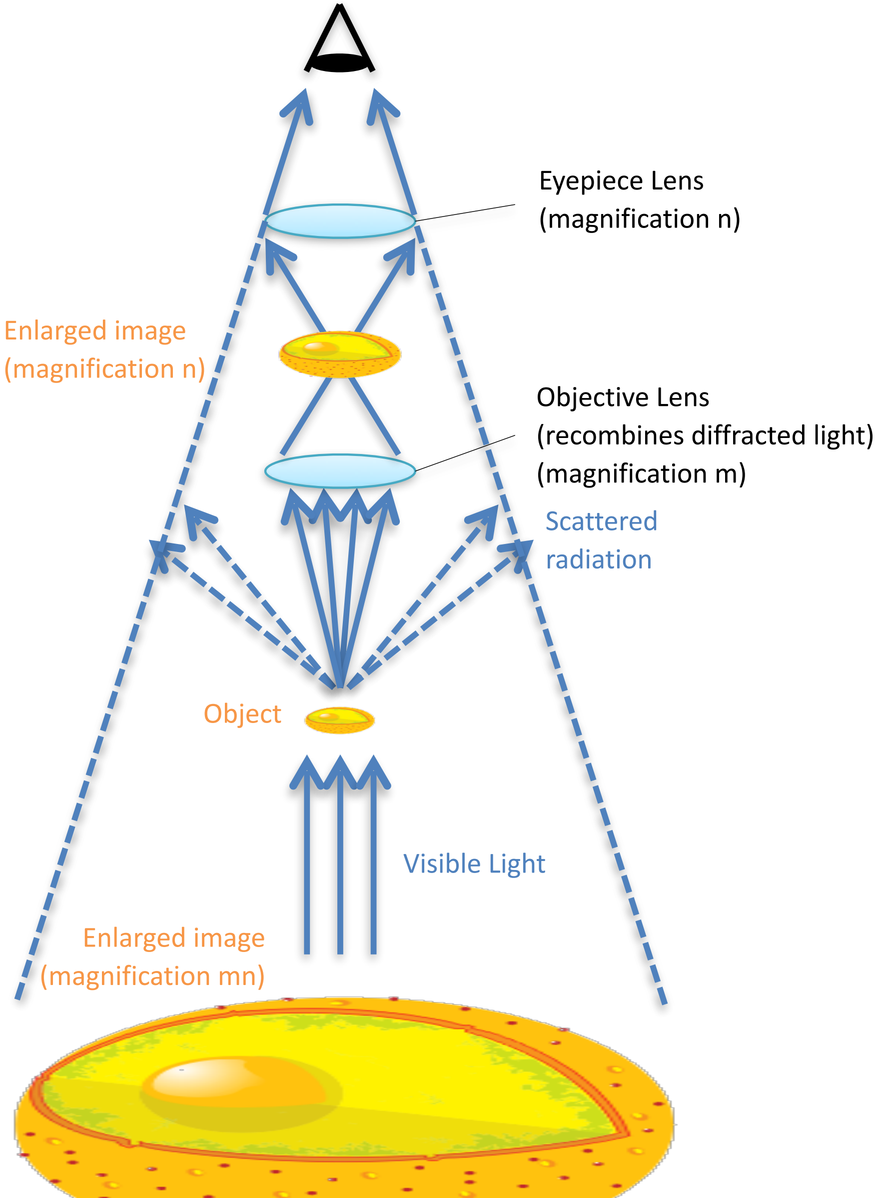
2012 Chemistry:
Robert J. Lefkowitz, Brian Kobilka

Studies of G-protein coupled receptors

X-ray diffraction vs. microscopy

- In XRD, X-rays are not refocussed but only diffracted.

Light Microscope


Eyepiece Lens
(magnification n)

Enlarged image
(magnification n)

Objective Lens
(recombines diffracted light)
(magnification m)

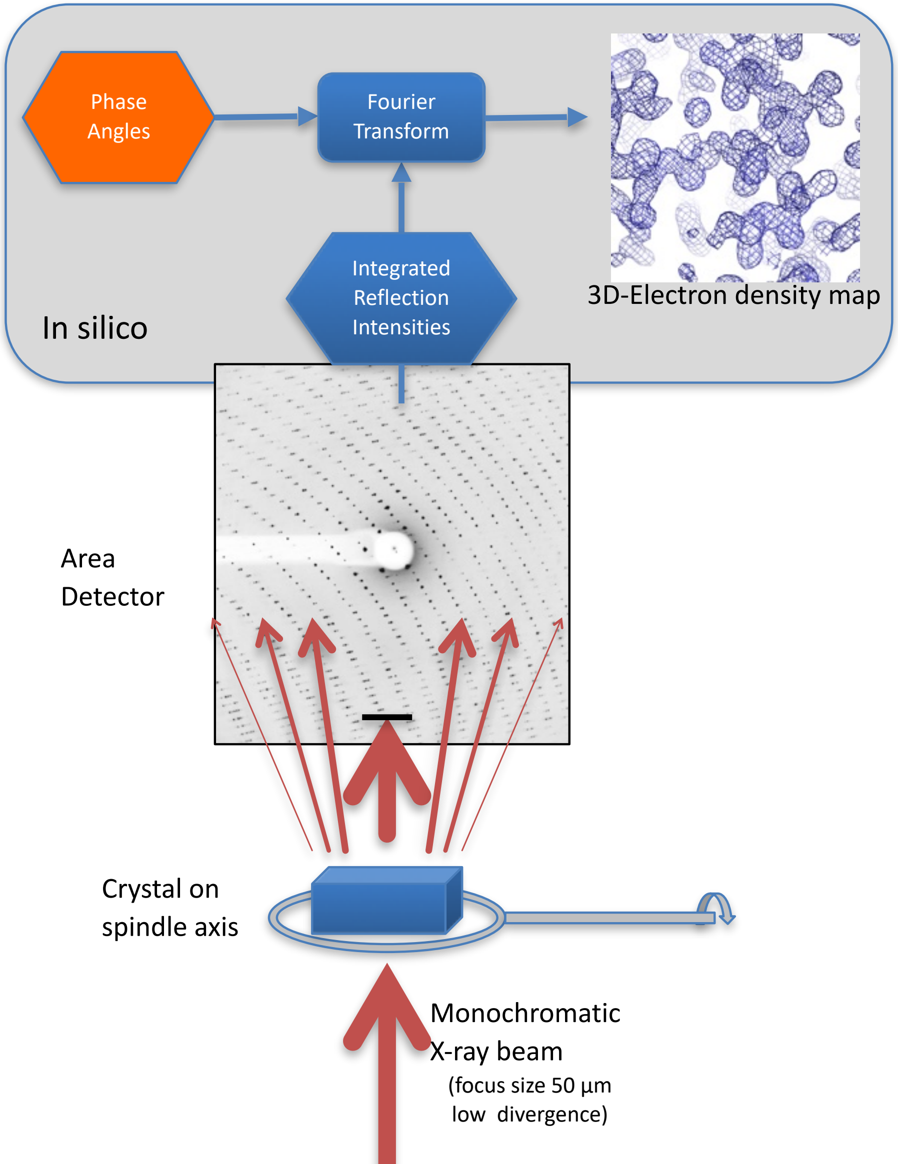
Scattered
radiation

Object

Visible Light

Enlarged image
(magnification mn)

X-ray Crystallography


Phase
Angles

Fourier
Transform

3D-Electron density map

In silico

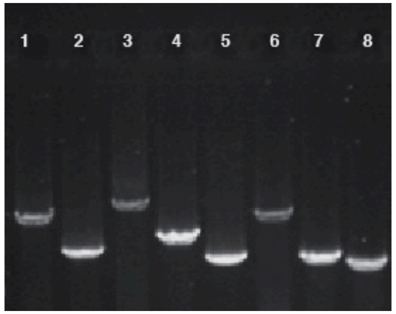
Integrated
Reflection
Intensities

Area
Detector

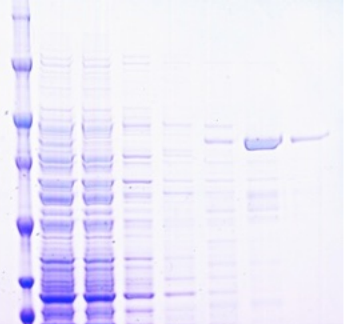
Crystal on
spindle axis

Monochromatic
X-ray beam
(focus size 50 µm
low divergence)

X-ray Crystallography

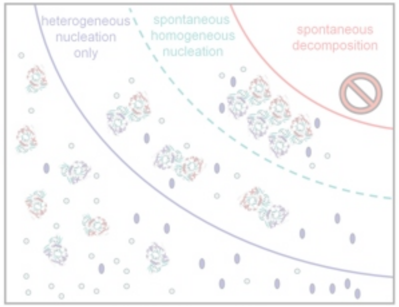


Molecular Cloning

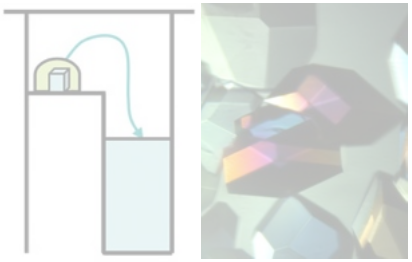


Protein Purification

Protein Production
&
Characterization

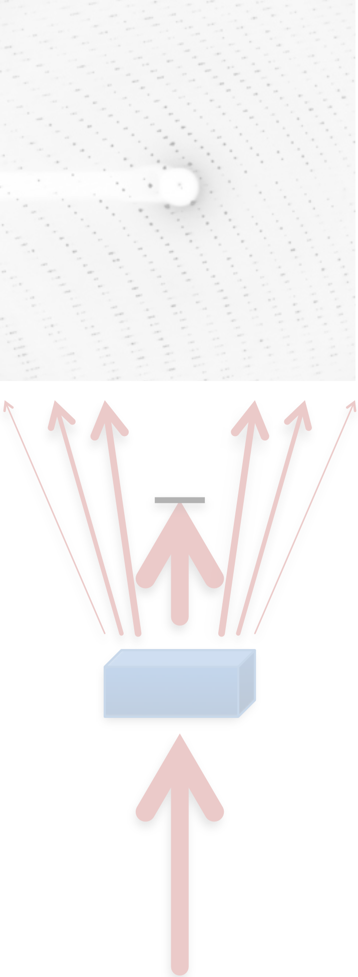


Phase Diagram

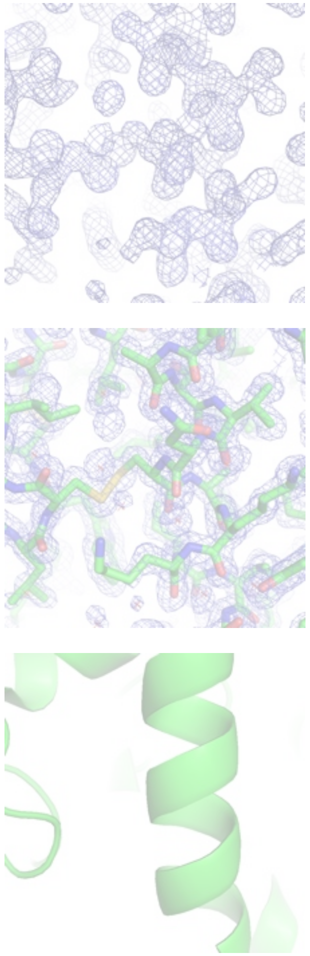


Crystal Growth

Crystallization

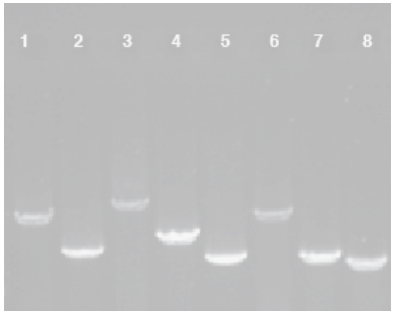


Diffraction
Experiment

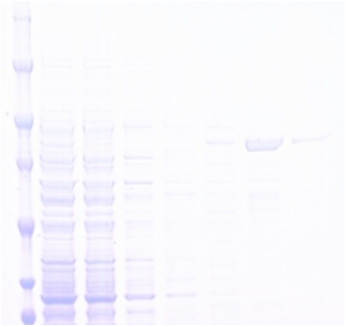


Structure
Determination
&
Model building

X-ray Crystallography

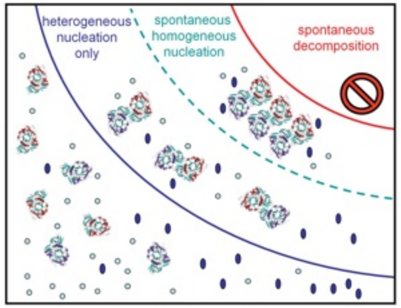


Molecular Cloning

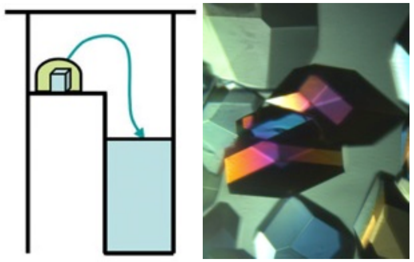


Protein Purification

Protein Production
&
Characterization

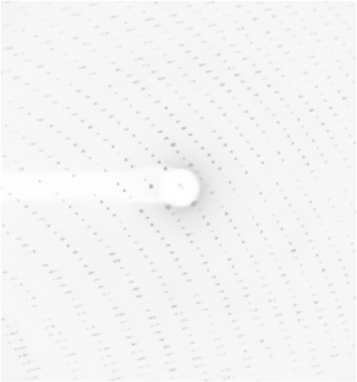


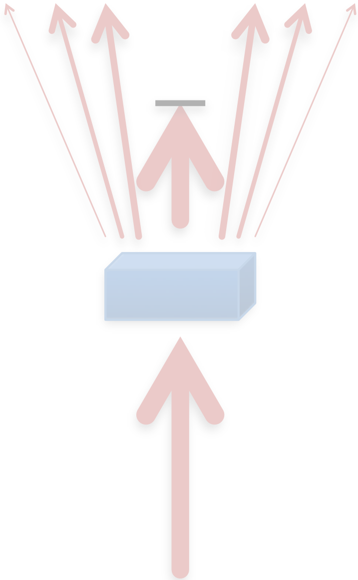
Phase Diagram



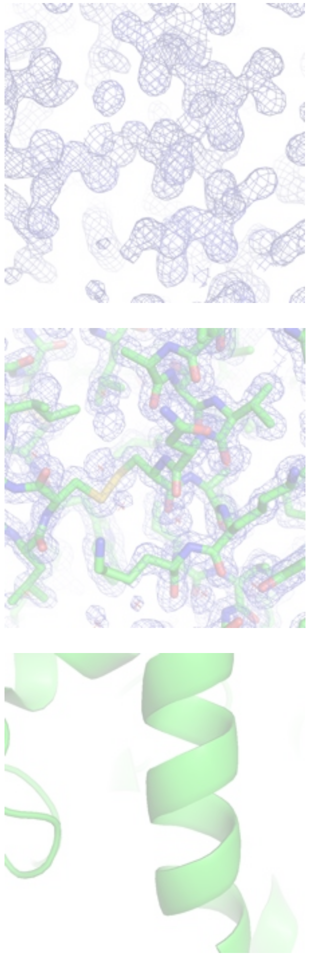
Crystal Growth

Crystallization



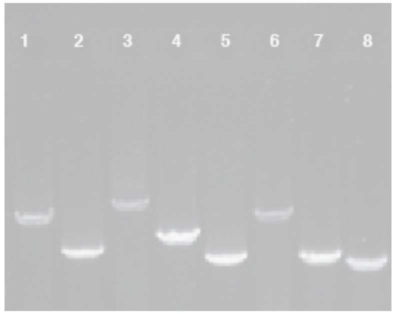


Diffraction
Experiment

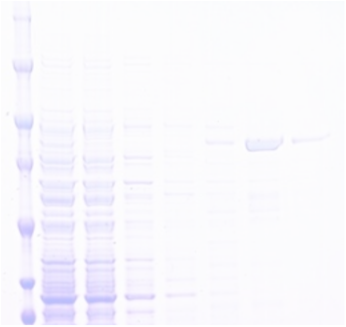


Structure
Determination
&
Model building

X-ray Crystallography

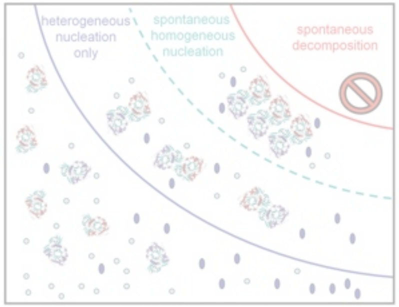


Molecular Cloning

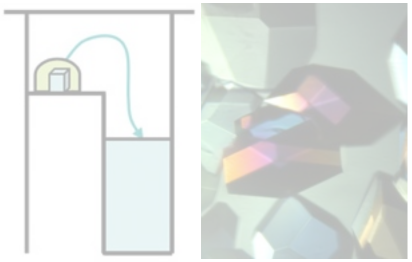


Protein Purification

Protein Production
&
Characterization

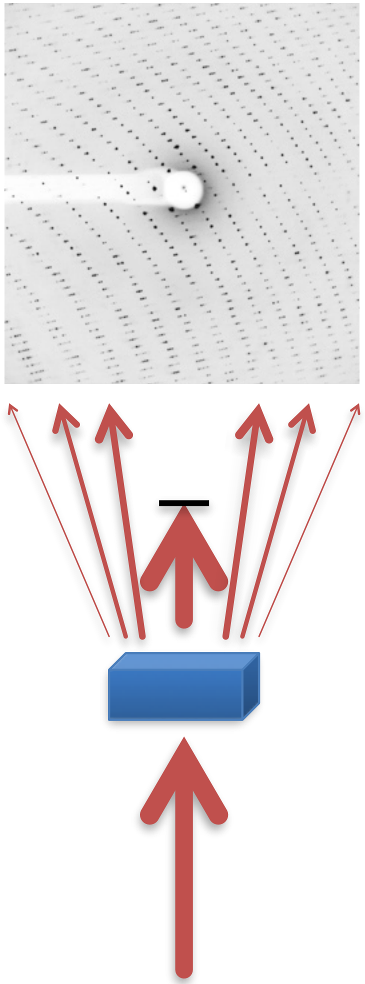


Phase Diagram

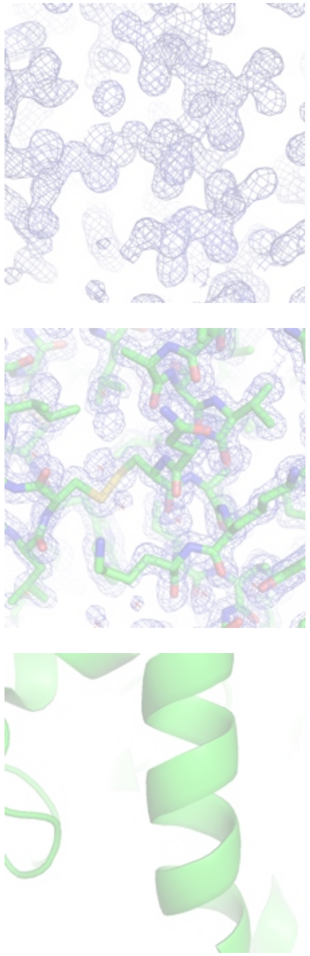


Crystal Growth

Crystallization

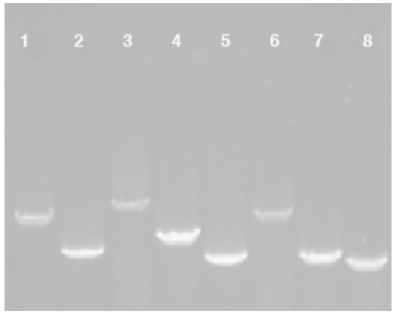


Diffraction
Experiment

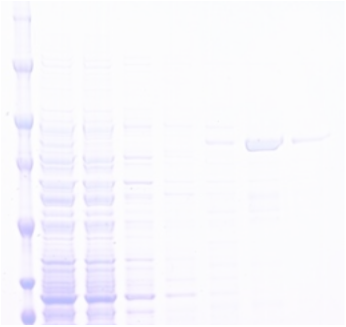


Structure
Determination
&
Model building

X-ray Crystallography

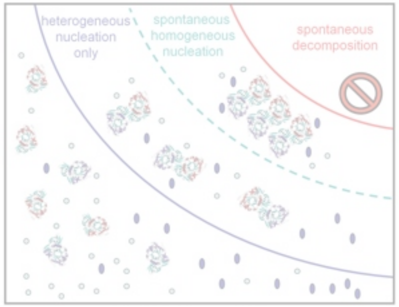


Molecular Cloning

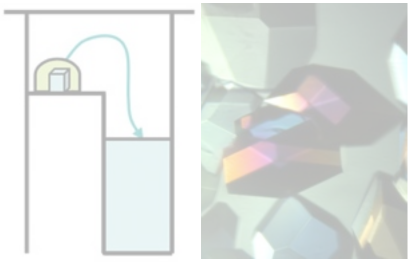


Protein Purification

Protein Production
&
Characterization

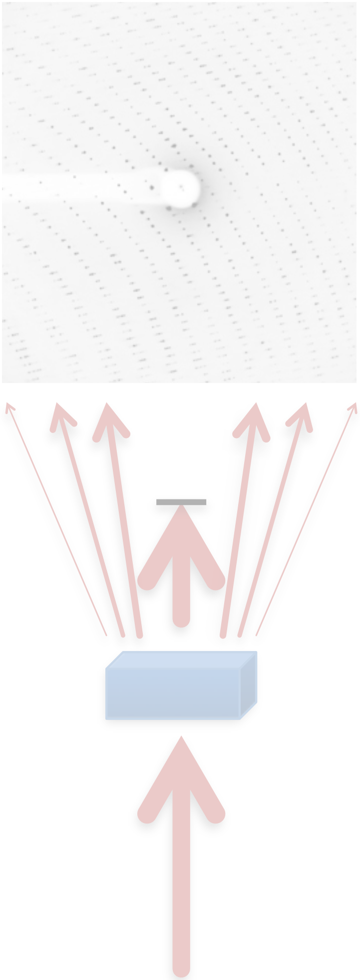


Phase Diagram

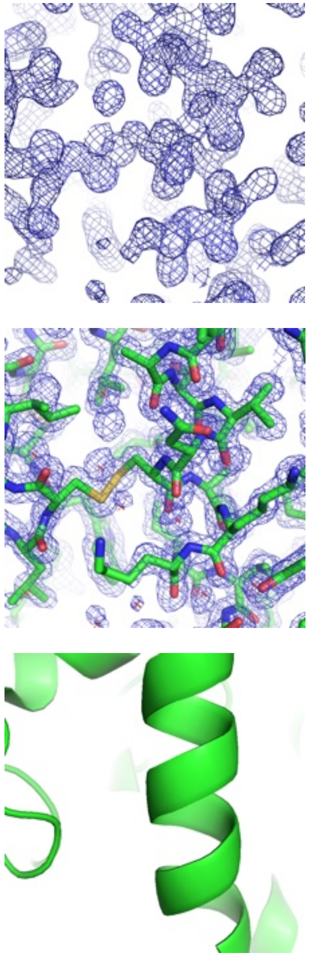


Crystal Growth

Crystallization

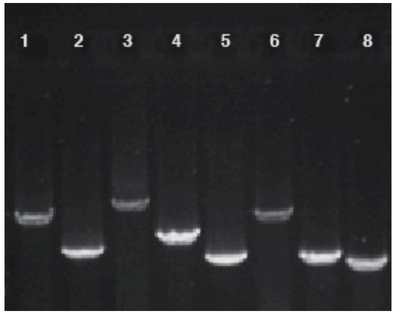


Diffraction
Experiment

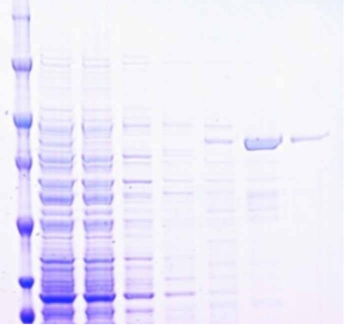


Structure
Determination
&
Model building

X-ray Crystallography

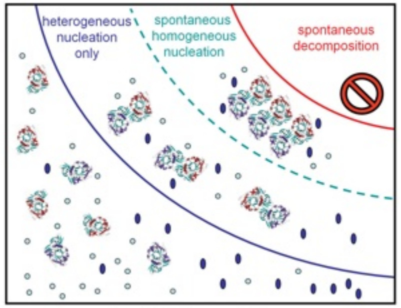


Molecular Cloning

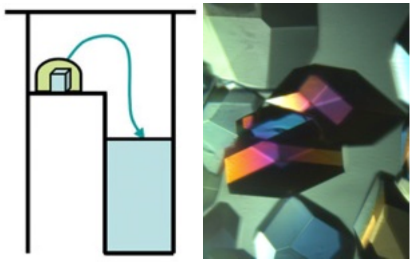


Protein Purification

Protein Production
&
Characterization

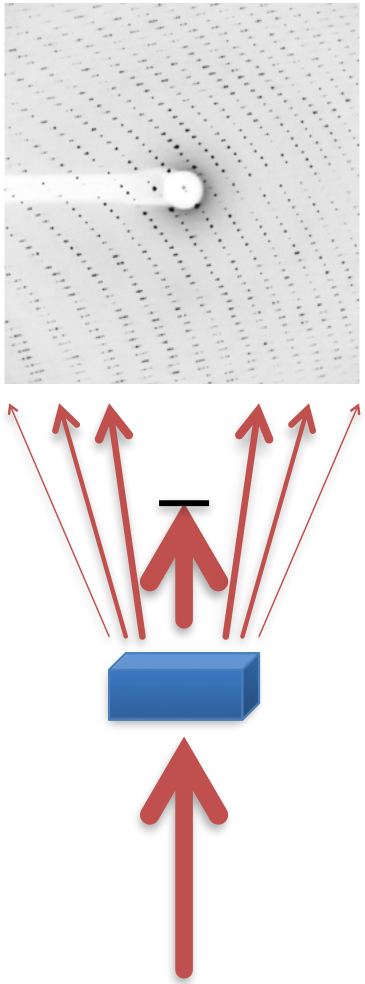


Phase Diagram

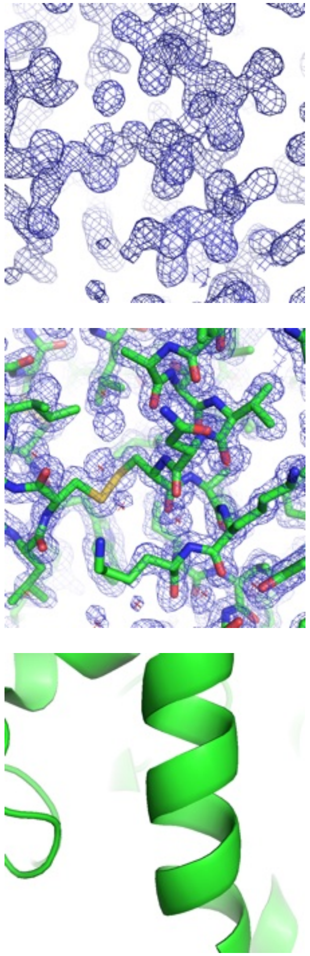


Crystal Growth

Crystallization



Diffraction
Experiment

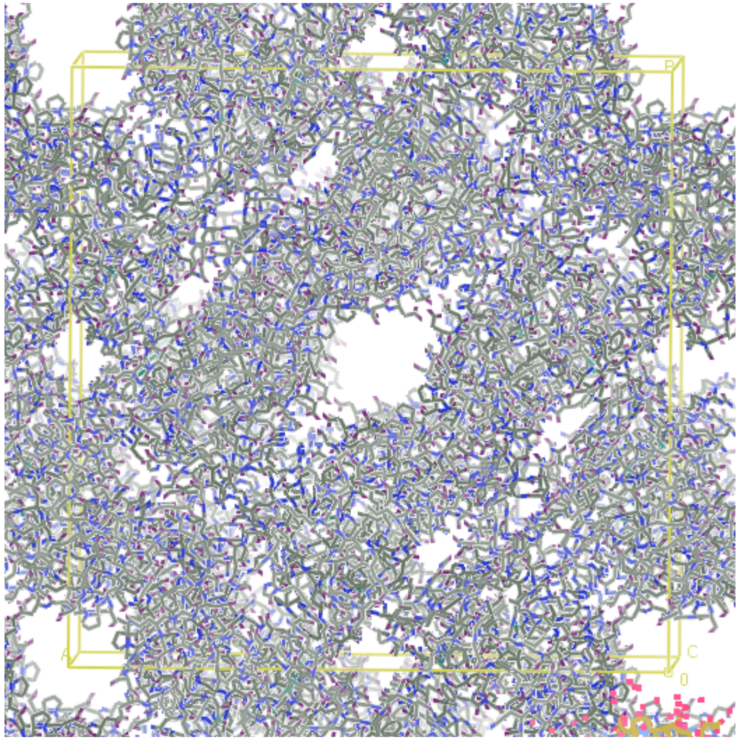


Structure
Determination
&
Model building

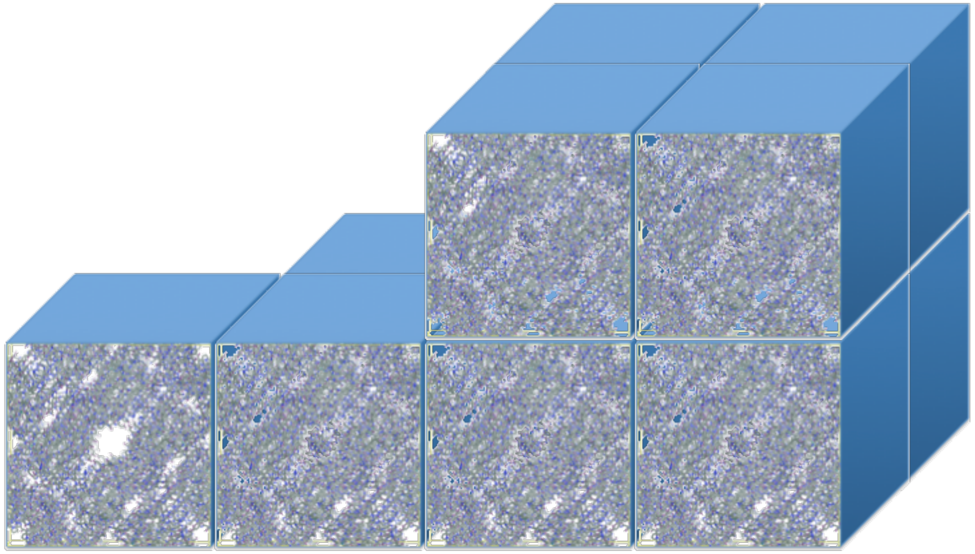
- Crystallization
  - What is a crystal ?
  - Properties of biological crystals
  - Crystallization principles
  - Crystallization techniques
  - Automation

Biomacromolecular Crystals

- Regular arrangement of multiple copies of a structure in 3D
- Build by translations of a basic repeat unit: **Unit cell**
- The unit cell may comprise additional symmetry elements
- The smallest element representative of the crystal is the “**asymmetric unit**”

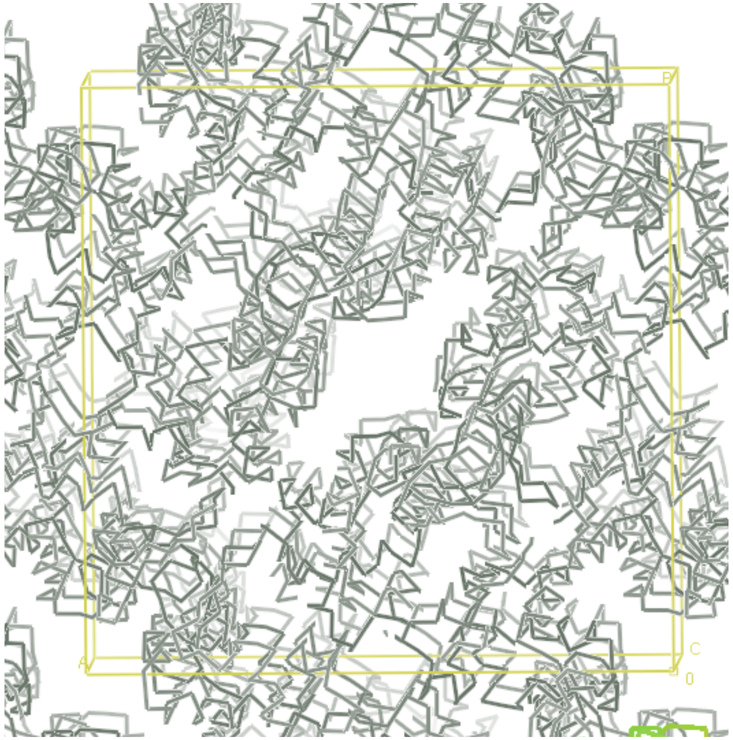


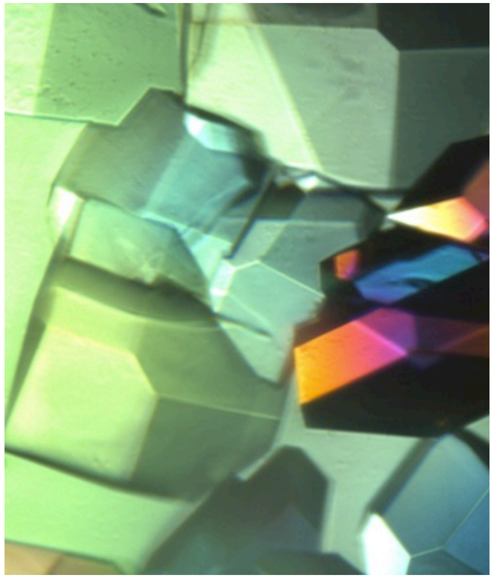
Crystallography depends on crystals !



Biomacromolecular Crystals

- Typical dimensions of protein crystals are 20 – 500 µm
- May adopt many kinds of shapes
- Are often birefringent
- Are very sensitive (dehydration, mechanical stress)
- Protein crystals contain approx. 50% solvent

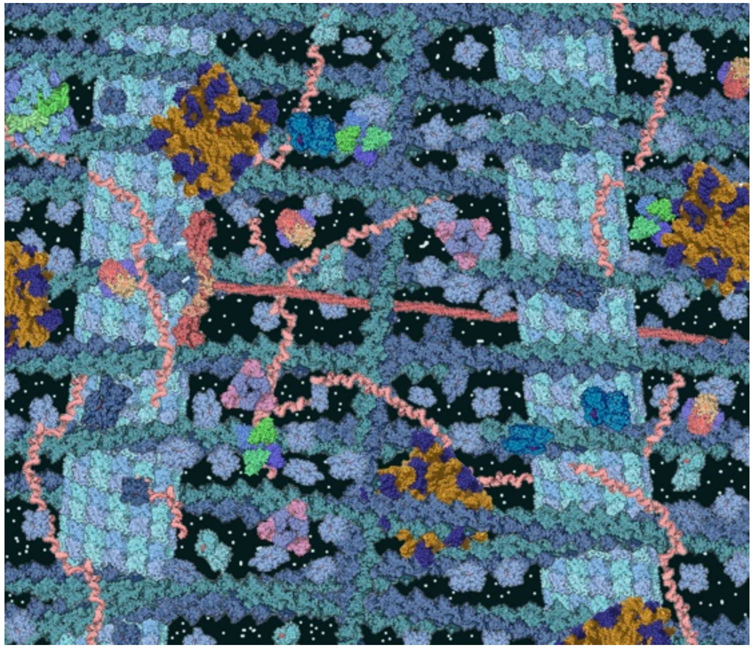




Ribbon diagram and unit cell of Lysozym (pdb 3p64, spacegroup $P4_32_12$)

Crystallization Artifacts ?

- Crystals are typically grown under native conditions
- Protein concentrations inside crystals: 200-800 mg/ml
  - Cytosol
    - 20-30% macromolecules
    - >200 mg/ml macromolecules

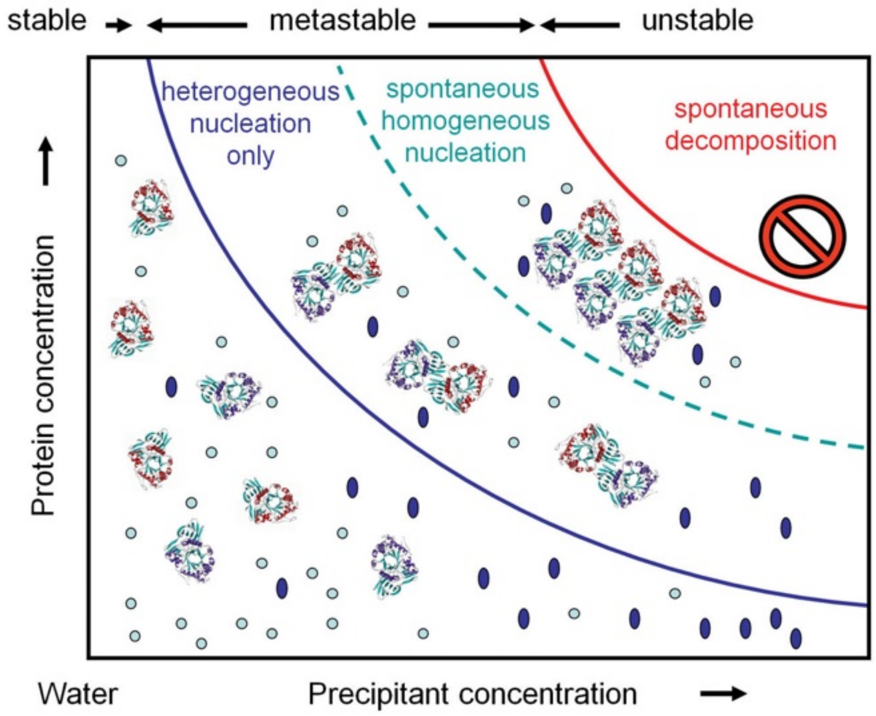


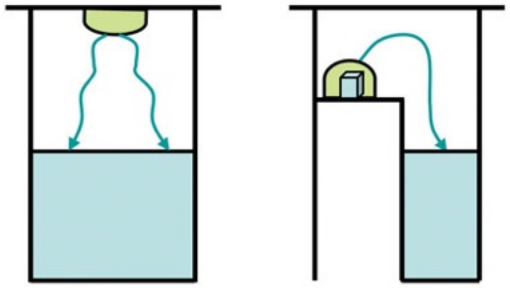
Goodsell DS (June 1991). "Inside a living cell". Trends Biochem. Sci. 16 (6): 203–6.

- Protein crystals are stabilized by weak forces
  - 3-12 (mostly weak) contacts per molecule
  - Interface area of 200-500 Å^2 (0.5-5% of total surface)
  - Many proteins crystallized in different crystal forms
  - Physiological binding events induce structural variations
  - In general in agreement with solution techniques: SAXS, NMR
  - Many proteins have proven to be “active” in crystalline state

Crystallization

- Crystallization is not predictable
- Systematic screening of a large range of favorable conditions
- Reduction of protein solubility by precipitant agents
- Two step process: Nucleation & Crystal Growth
- Equilibration of protein/precipitant vs. precipitant solution



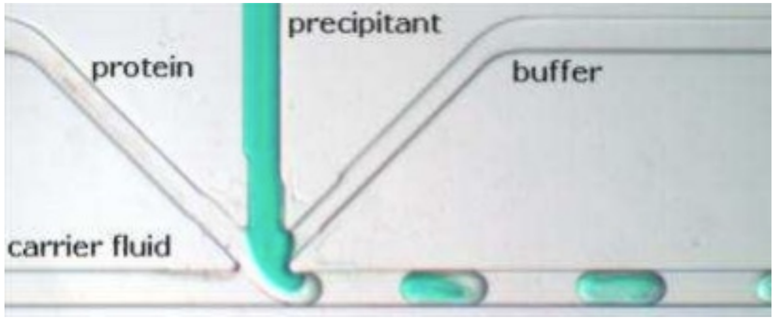


The classic: hanging-drop vapor diffusion

The variant: sitting-drop vapor diffusion

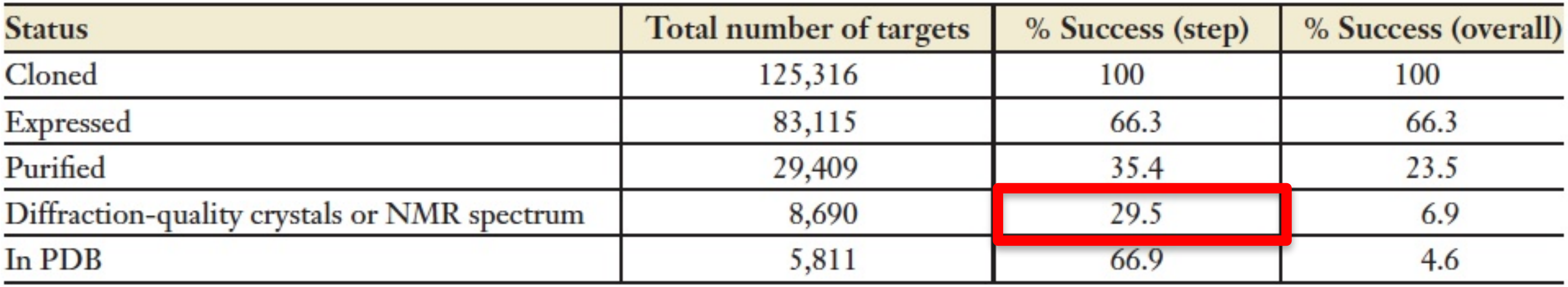
Crystallization

- Classic manual pipetting: min. 1 µl drops
  - 100 µl of 10 mg/ml = 1 mg protein for 96 conditions
- Robotic nanoliter crystallization: min. 100 nl drops
  - 100 µl of 10 mg/ml = 1mg protein for 960 conditions
- Microfluidic Chip Crystallization: < 20 nl plugs
  - 16 µl of 10 mg/ml = 0.16 mg protein for 800 conditions



www.emeraldbioscience.com

Success Rate of Crystallization



Status	Total number of targets	% Success (step)	% Success (overall)
Cloned	125,316	100	100
Expressed	83,115	66.3	66.3
Purified	29,409	35.4	23.5
Diffraction-quality crystals or NMR spectrum	8,690	29.5	6.9
In PDB	5,811	66.9	4.6

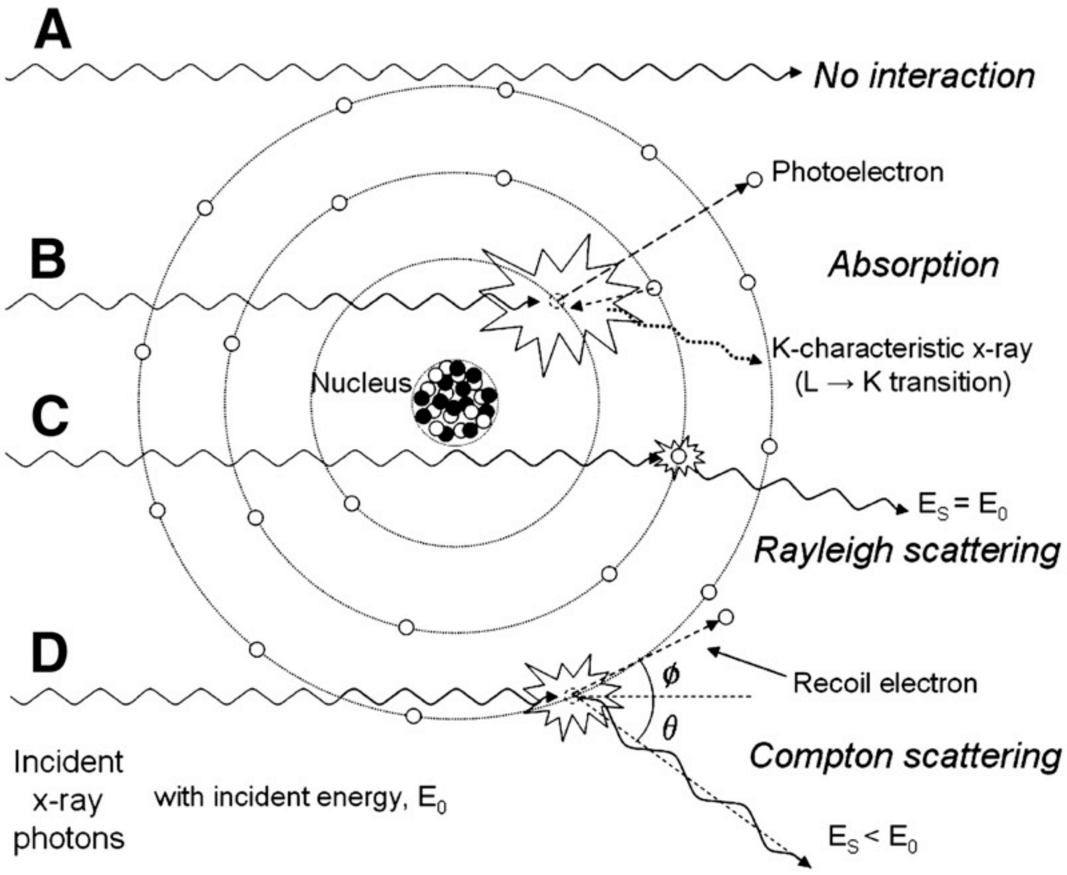
Data from 2009, Success rates in Structural Genomics
Annu. Rev. Biophys. 2009. 38:371–83

- Crystallization is a key bottleneck for all targets
- Success rates for human, multi-domain or membrane proteins even lower
- Escape Strategy: Screening homologues

- X-ray Data Collection
  - Properties of X-rays
  - Generation and detection of X-rays
  - X-ray induced radiation damage

(Hard) X-rays

- X-ray interaction with matter:



Interaction of 12 keV X-rays when passing through 0.1 mm soft tissue

	Interaction	
A	Transmission	96.6%
B	Photoabsorption	3%
C	**Rayleigh Scattering**	0.2%
D	Compton Scattering	0.2%

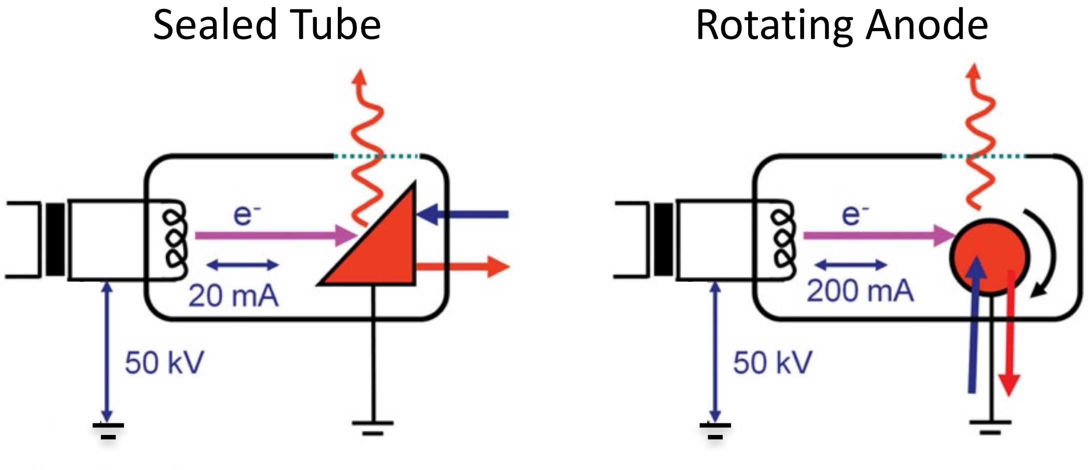
No secondary scattering !

Adapted from J. Nucl. Med. Technol. March 1, 2005 vol. 33 no. 1 3-18

Classic X-ray Generation and Detection

Generation of X-rays

- Cathode ray tubes: Thermally extracted electrons hit anode target, photoabsorbtion leads to emission of characteristic X-rays
- Rotating Anode: heating problem overcome by rotating anode



Sealed Tube

Rotating Anode

© Garland Science 2010

www.rigaku.com

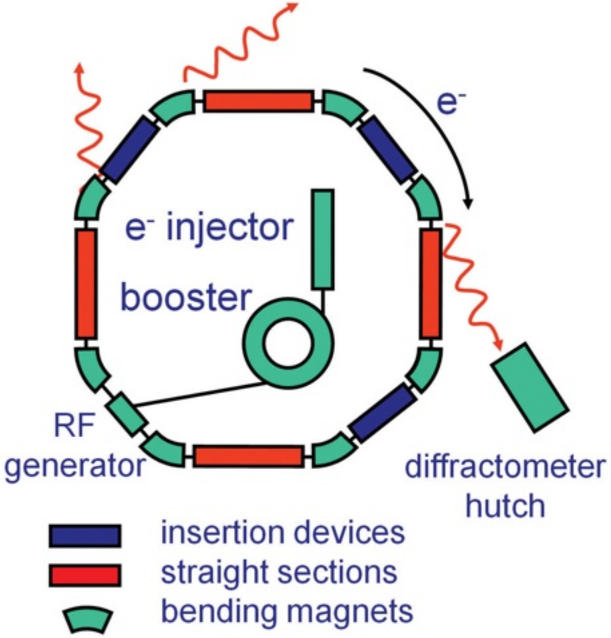
Detection

- X-ray sensitive film
- Image Plates: Phosphorescent Screens

Modern X-ray Generation and Detection

- Particle accelerators specialized in production of synchrotron radiation
- Largest research instruments in the biosciences
- Operated at international facilities
- Highly brilliant, tunable wavelength

3. Generation
Synchrotron Source



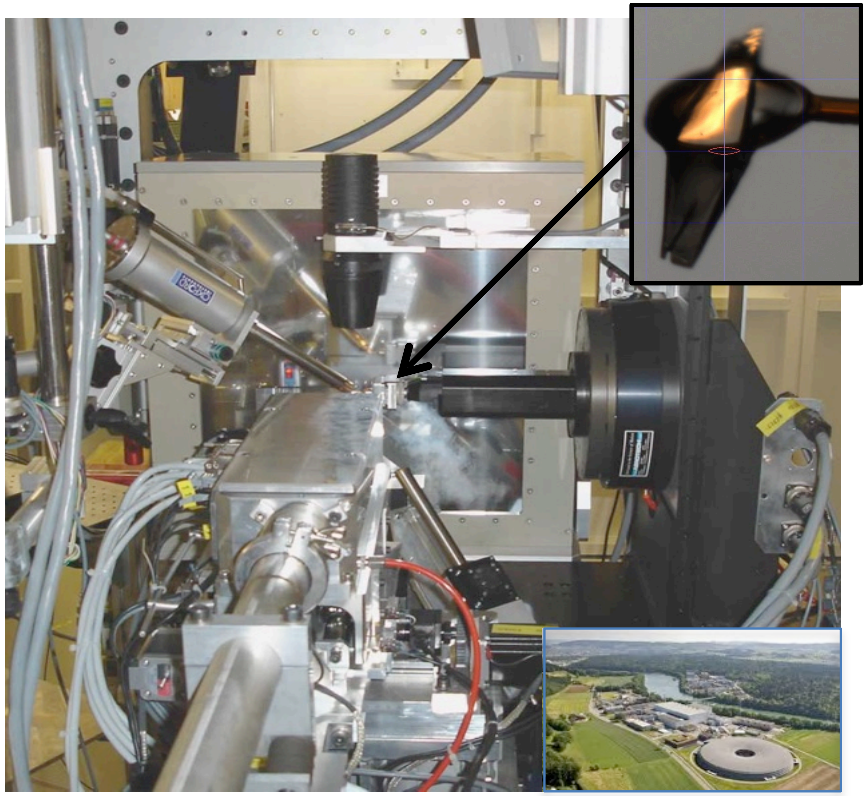
© Garland Science 2010

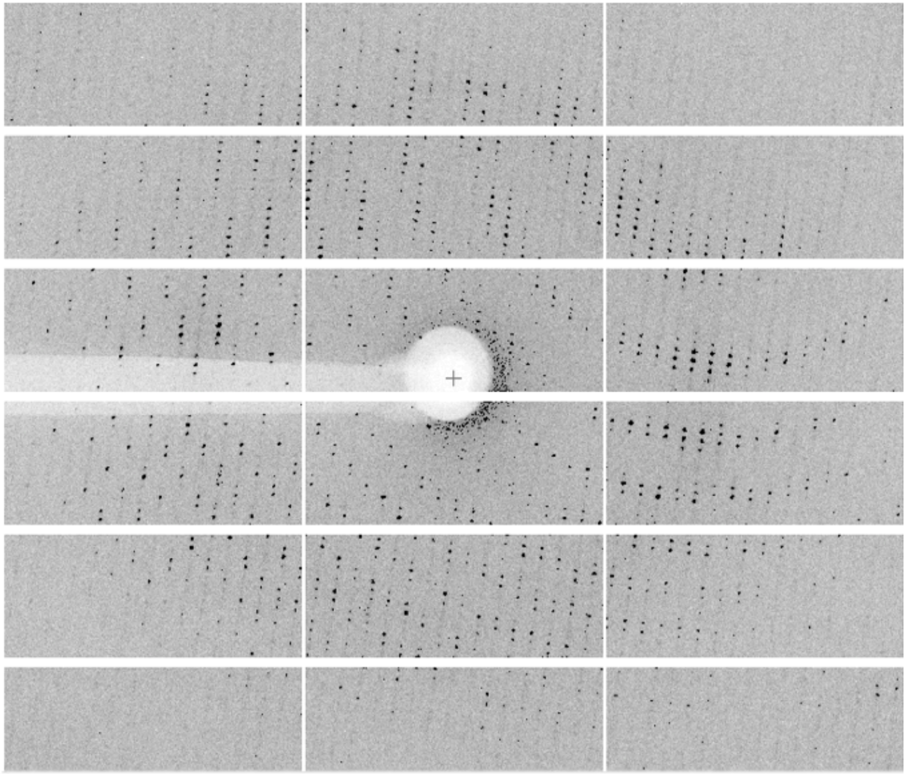
Swiss Light Source, PSI, Villigen (AG)

www.psi.ch

Modern X-ray Generation and Detection

- Cryo Data collection at 100 K in nitrogen gas **to reduce radiation damage**
- Photon-counting Silicon-pixel detectors





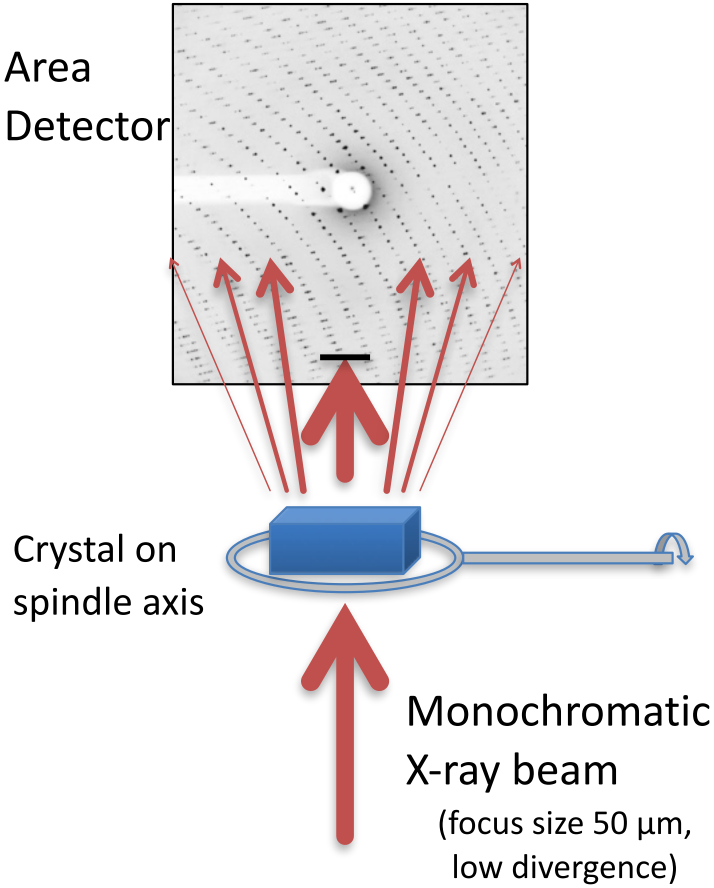
- Data collection times:

Inhouse / Image Plate detector (2 d):	2880 min
Second generation synchrotron / CCD detector (1 h):	60 min
Third generation synchrotron / Pixel-counting detector:	1 min

Data Collection

Rotation Method

- Crystal mounted in unknown, arbitrary orientation
- Crystal rotates (e.g 0.5°/image) around an axis perpendicular to the incident beam
- One image taken during each rotation, e.g. 360 images for 360 successive rotations
- Use of monochromatic X-rays



$$I_{hkl} = I_0(\lambda^3/\omega)(V_x \cdot L \cdot p \cdot A / V^2)|F_{hkl}|^2$$

I_0 = incident beam intensity

λ = wavelength of radiation

ω = rotation velocity of the crystal

V_x = volume of the crystal

L = Lorentz factor

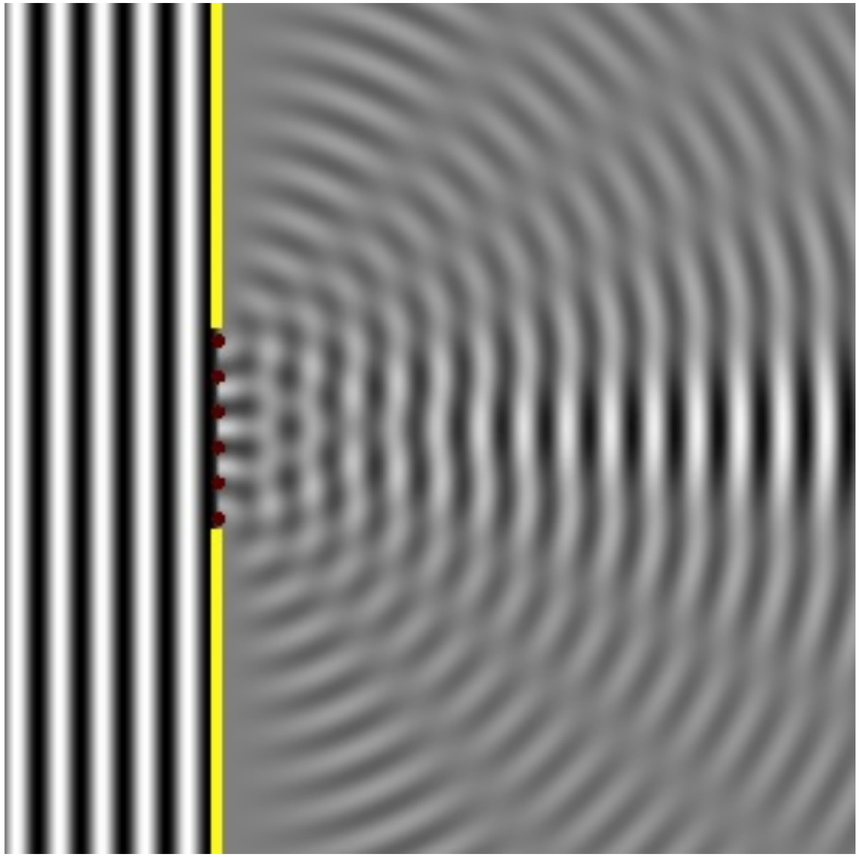
p = the polarization factor

A = absorption factor

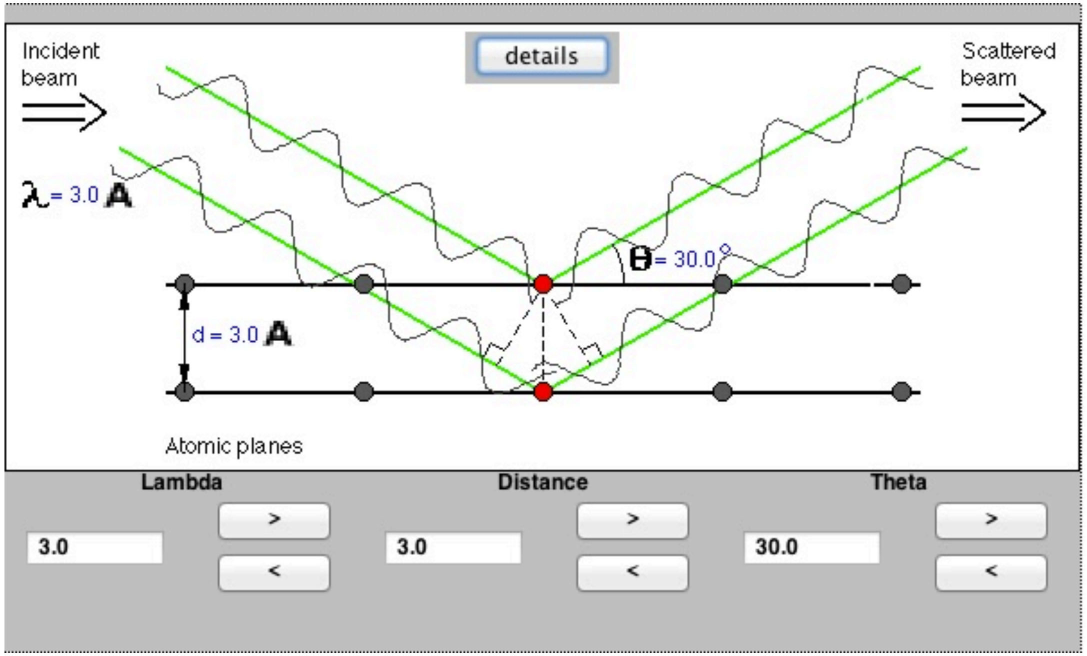
V = volume of the unit cell

$|F_{hkl}|$ = the observed structure factor

Bragg’s Law



By Lookang many thanks to Fu-Kwun Hwang and author of Easy Java Simulation = Francisco Esquembre (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons



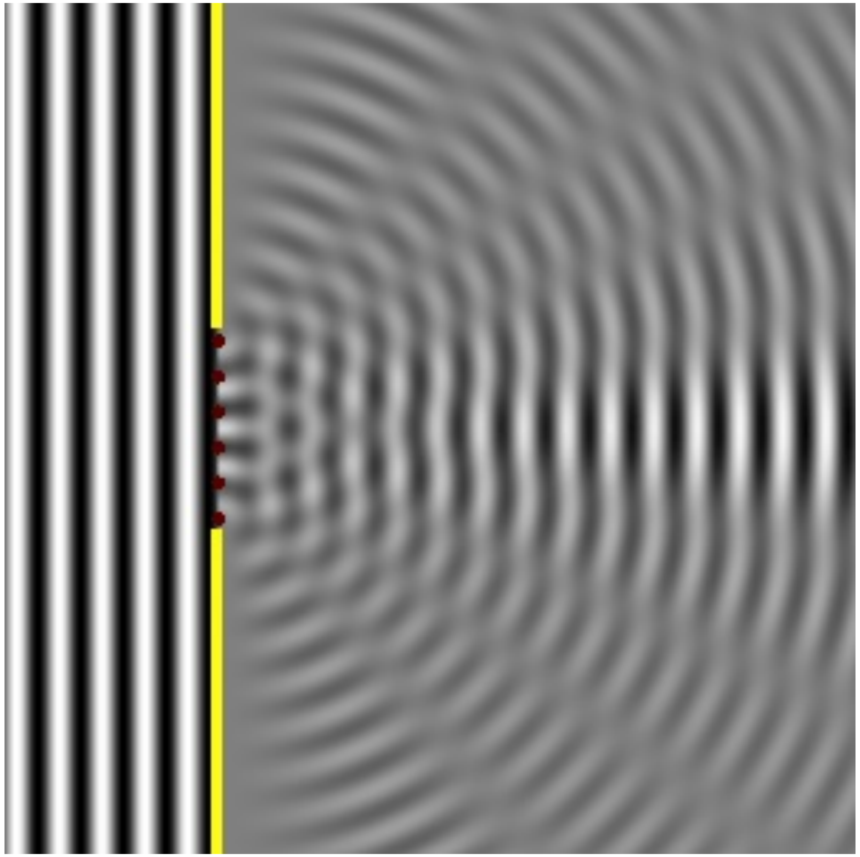
Lambda 3.0

Distance 3.0

Theta 30.0

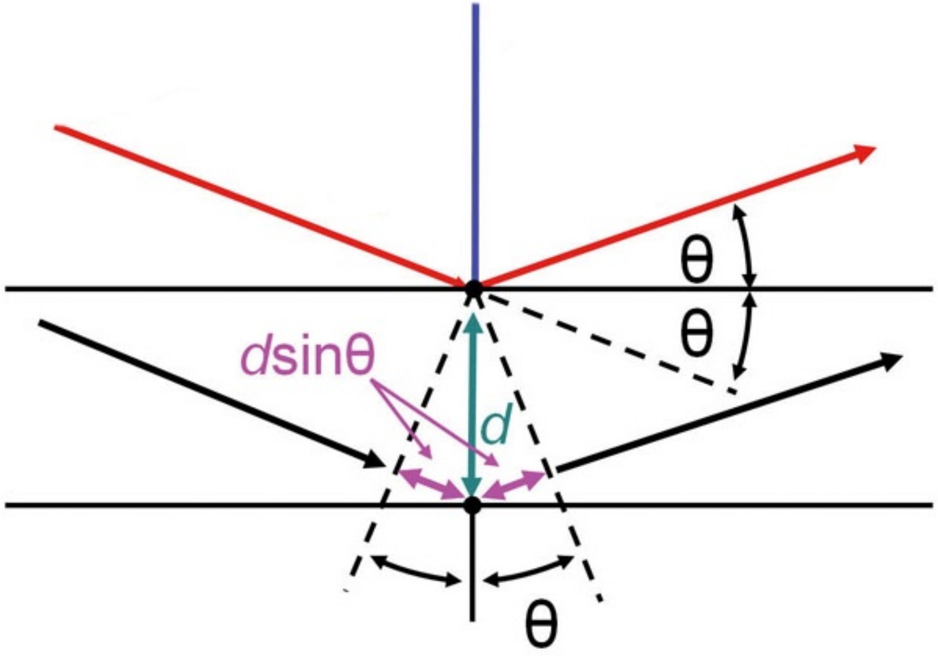
http://skuld.bmsc.washington.edu/people/merritt/bc530/bragg/index.html

Bragg's Law



By Lookang many thanks to Fu-Kwun Hwang and author of Easy Java Simulation = Francisco Esquembre (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons

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



B. Rupp:Biomolecular Crystallography, Garland Press 2009

Single crystal diffraction

Position of reflections

Determined by parameters and orientation of the crystal lattice

Information about unit cell dimensions, crystal orientation relative to lab coordinate system

Intensity of reflections

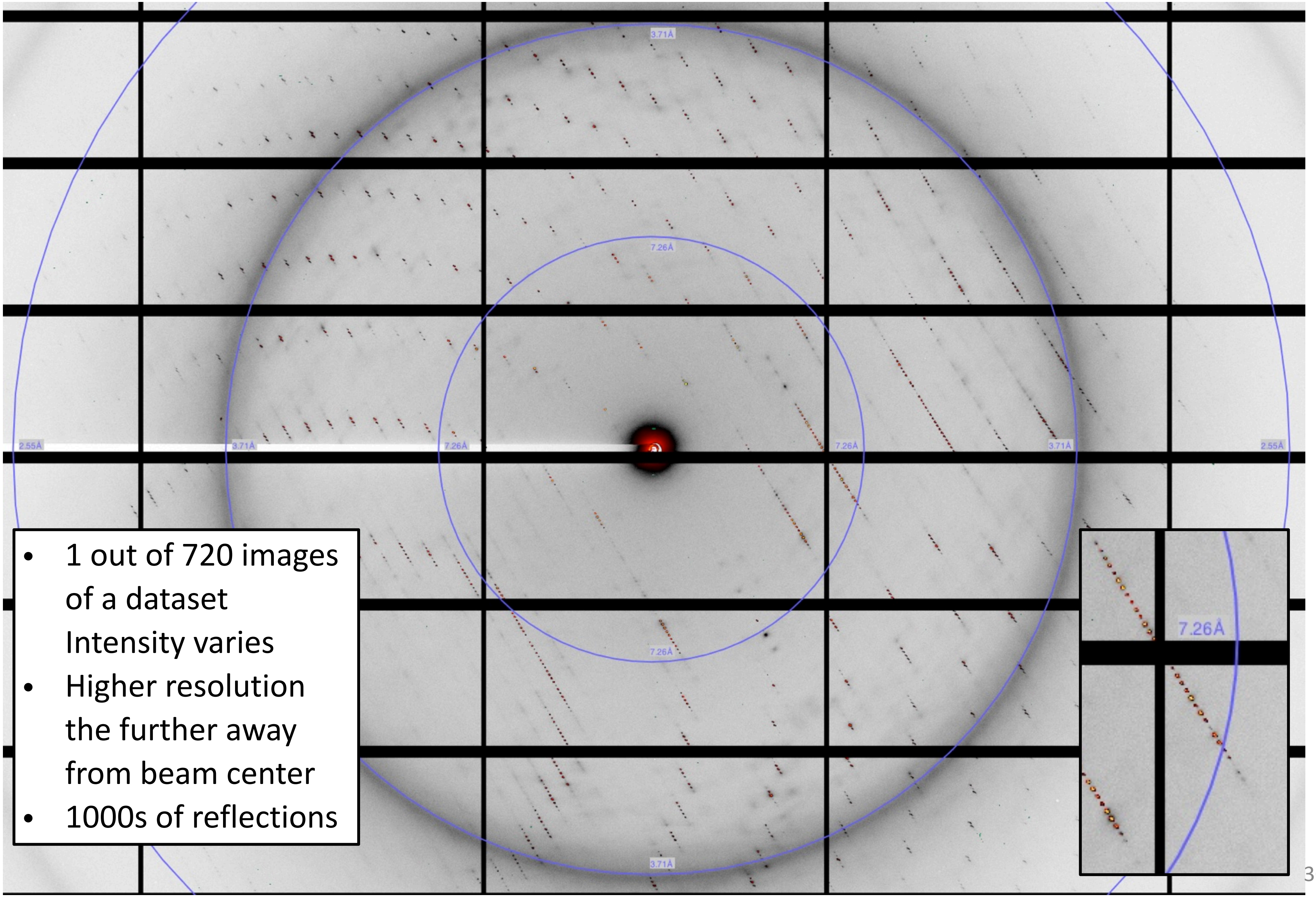
Determined by relative positions of atoms (electron density distribution) in the unit cell

Information about the **STRUCTURE** of the scattering object

Only intensities are recorded.

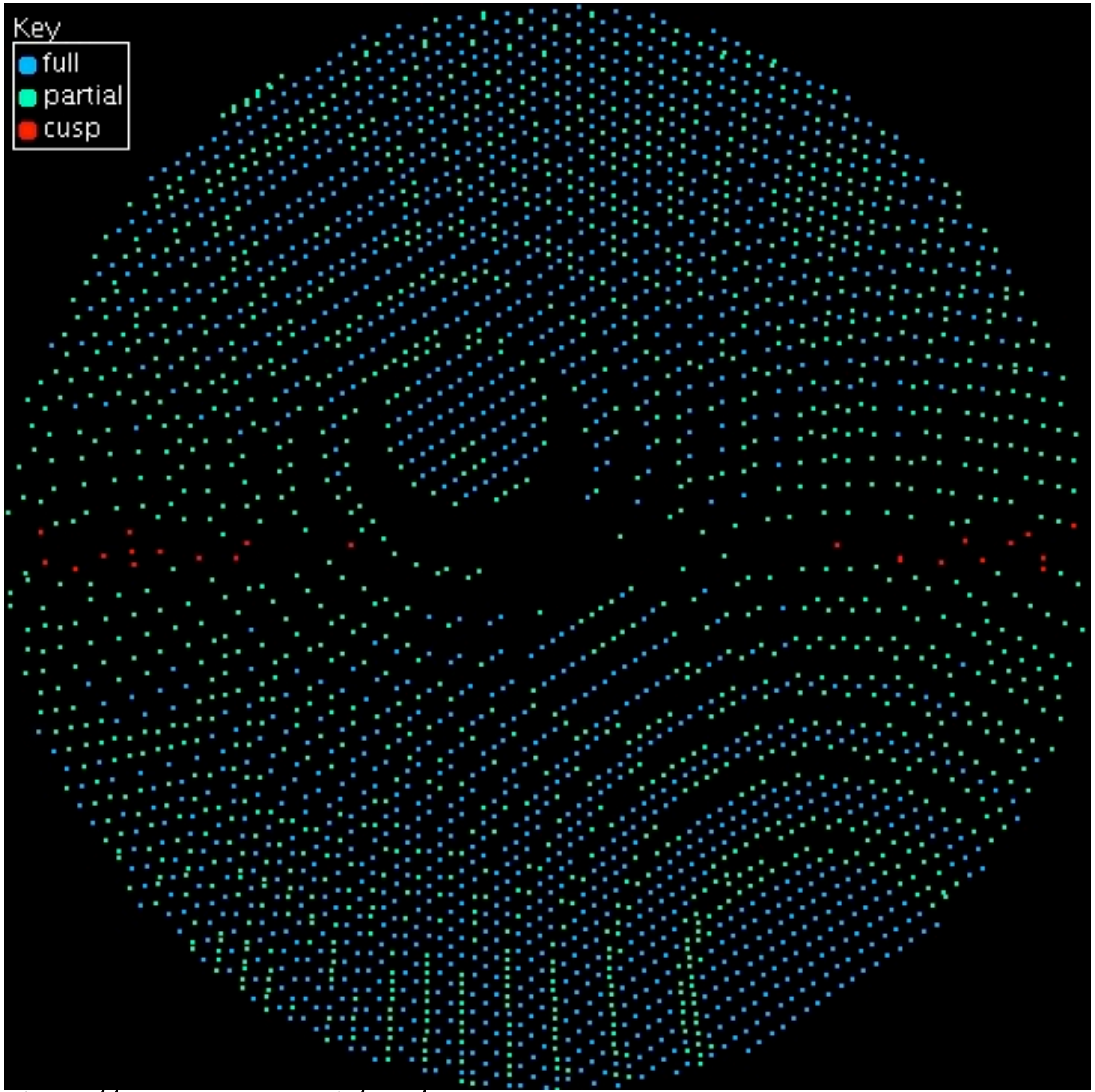
Part of the information required to restore the electron density distribution is contained in the missing phases !

Diffraction Pattern



- 1 out of 720 images of a dataset
  Intensity varies
- Higher resolution the further away from beam center
- 1000s of reflections

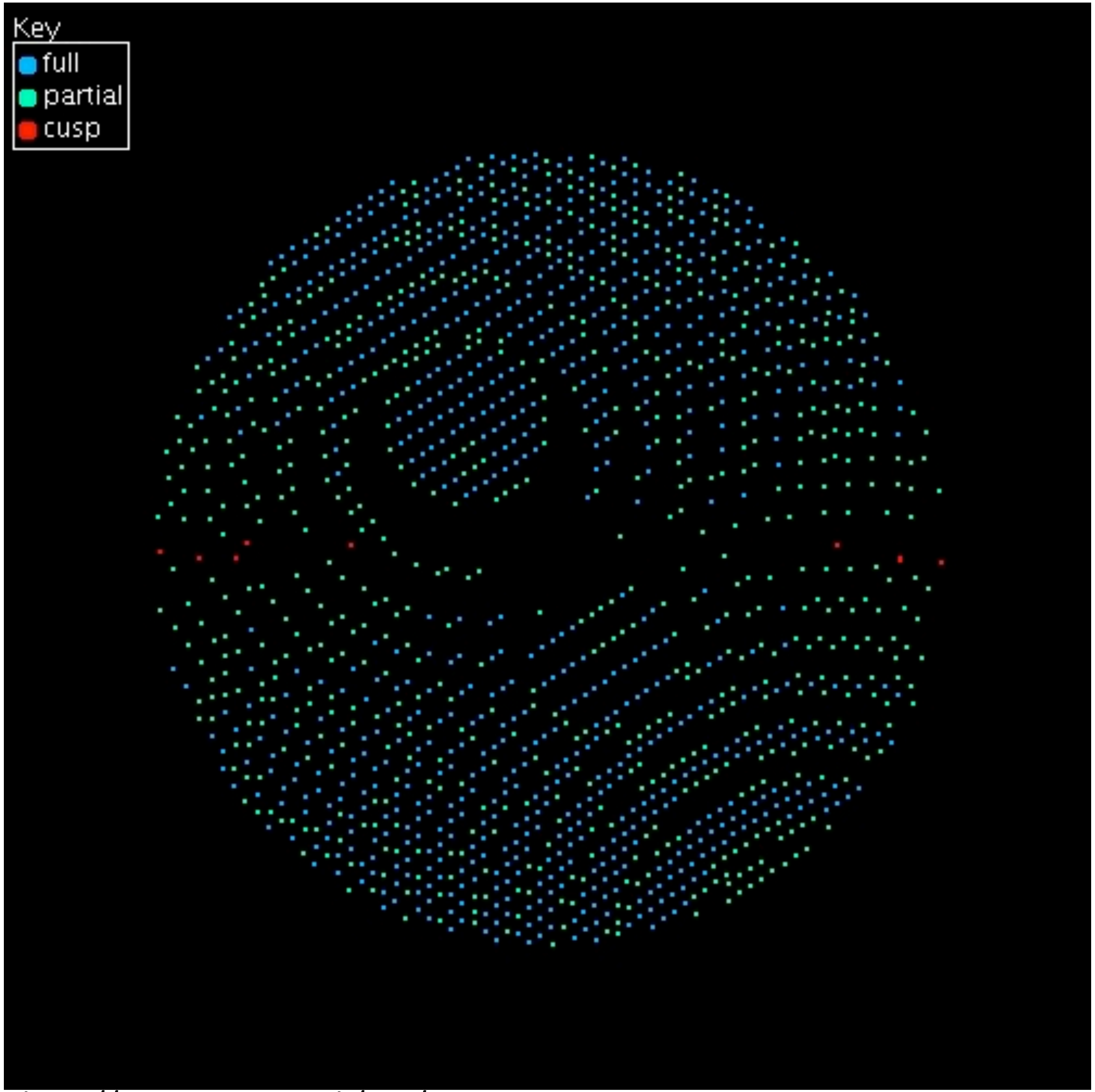
Diffraction Pattern: Resolution



http://www.ccp4.ac.uk/jwc/PXSim

Protein Crystal
Space group: C2
Unit cell parameters:
a=96.2Å
b=85.8Å
c=81.9Å
α=90°
β=120.5°
γ=90°
Resolution: 1.75Å

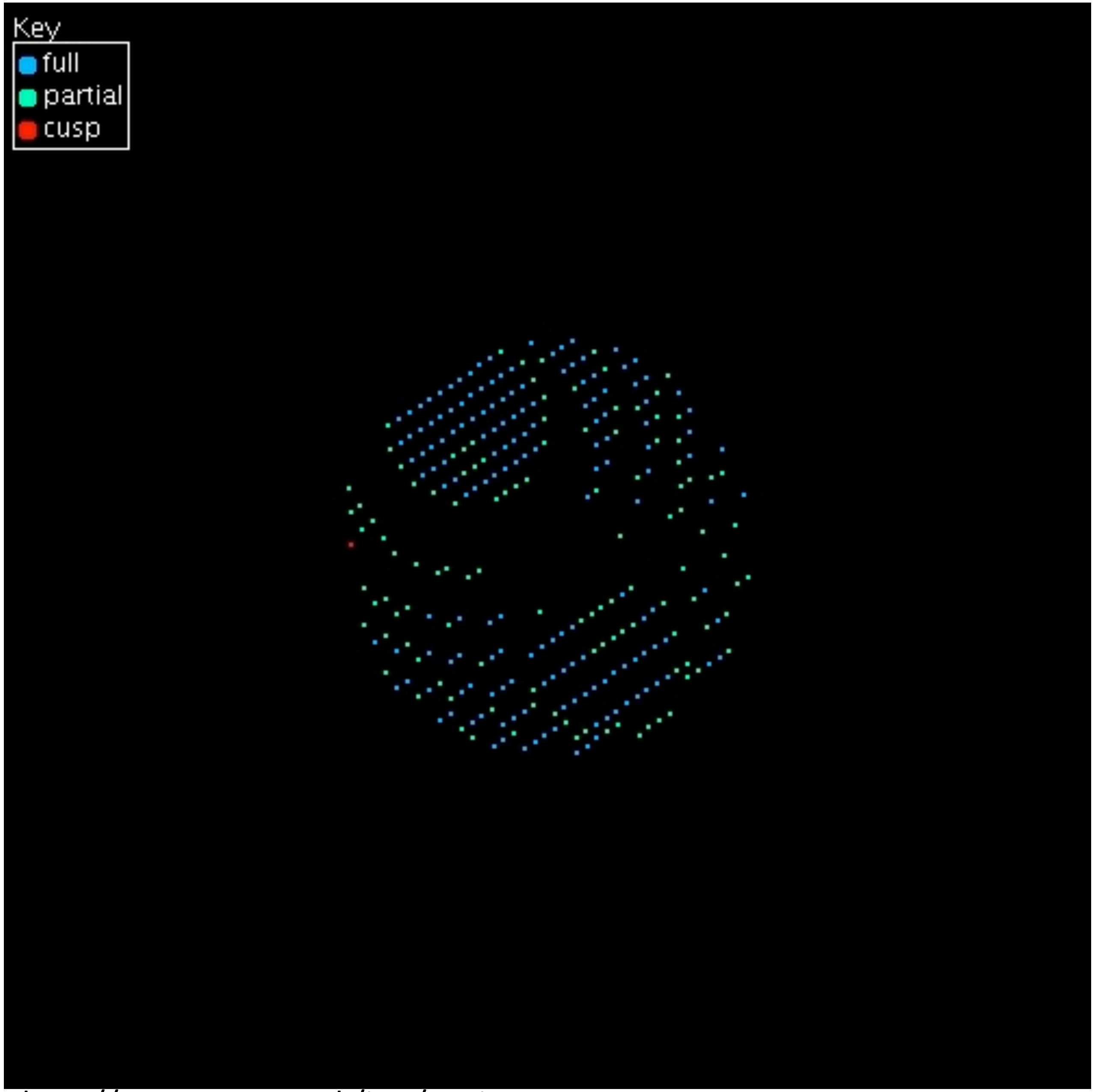
Diffraction Pattern: Resolution



http://www.ccp4.ac.uk/jwc/PXSim

Protein Crystal
Space group: C2
Unit cell parameters:
a=96.2Å
b=85.8Å
c=81.9Å
α=90°
β=120.5°
γ=90°
Resolution: 2.50Å

Diffraction Pattern: Resolution



http://www.ccp4.ac.uk/jwc/PXSim

Protein Crystal

Space group: C2

Unit cell parameters:

a=96.2Å

b=85.8Å

c=81.9Å

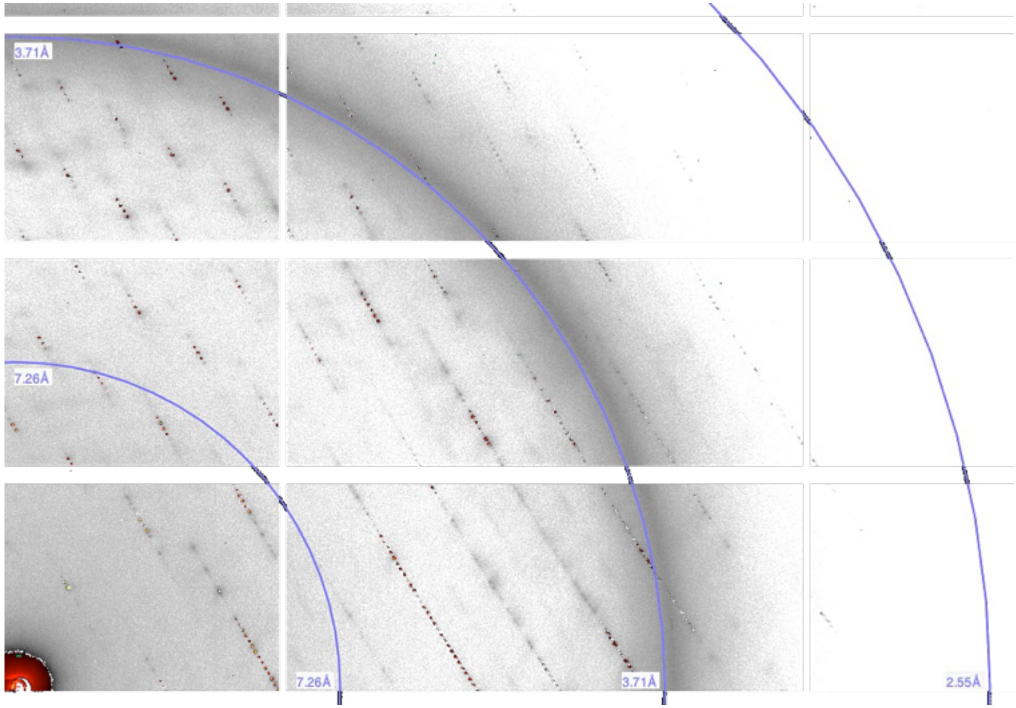
α=90°

β=120.5°

γ=90°

Resolution: 4.50Å

Images to Amplitudes



Applying detector corrections

Definition of Background

Finding Spots

Autoindexing:

- Determination of Bravais Lattice and Crystal orientation

Initial refinement:

- Refine cell constants and geometric parameters

Integration:

- Integrate all individual reflections

Correction/Post-refinement:

- Refine all processing parameters

Intensities to Amplitudes

- Structure Factor Amplitude is square root of integrated intensity
- Background Subtraction may result in negative intensities
- Treated by fitting to known intensity distributions

$$I_{hkl} = \left| F_{hkl} \right|^2$$

- From amplitudes to model
  - Solve phase problem
  - Fourier transformation of amplitudes & phases (complex structure factor F_{hkl})
  - Semi-automated model building
  - (Iterative) phase and model improvement
  - Validation & deposition

What is the **phase problem**?

- Only amplitudes ($\left|F_{hkl}\right| = \sqrt{I_{hkl}}$) are measured

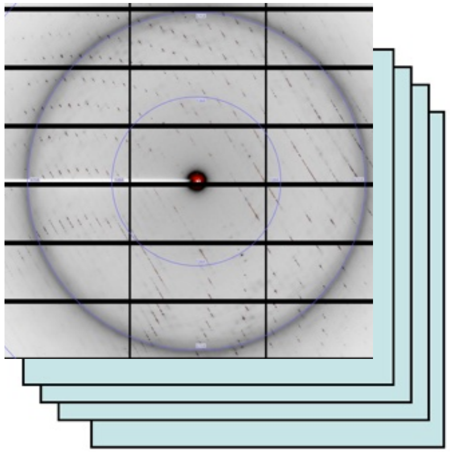
- Structure factor F_{hkl} is a complex number

$$F_{hkl} = \left|F_{hkl}\right| e^{i\varphi_{hkl}} = \left|F_{hkl}\right| (\cos\varphi_{hkl} + i\sin\varphi_{hkl})$$

- Phases (φ_{hkl}) are required for electron density calculation by the Fourier transform

$$\rho(x, y, z) = 1/V \sum_{hkl} \left|F_{hkl}\right| e^{-i[2\pi(hx+ky+lz)-\varphi_{hkl}])}$$

Electron density calculation



Integrated reflection intensities

Phase of individual reflections

(have to be determined indirectly)

PHASE PROBLEM OF CRYSTALLOGRAPHY

Fourier transformation

Electron density distribution

Depending on unit cell size and resolution one dataset may contain between a few thousand and millions of reflections

- Solutions to the Phase Problem:
  - **Heavy atom phasing:** introduce heavy atoms, calculate phases from difference pattern
  - **Molecular Replacement**: calculate phases from known similar, model (sub)structure
  - **‘Direct methods’:** for small structures at high resolution (Nobel Prize Hauptmann & Karle, 1985)

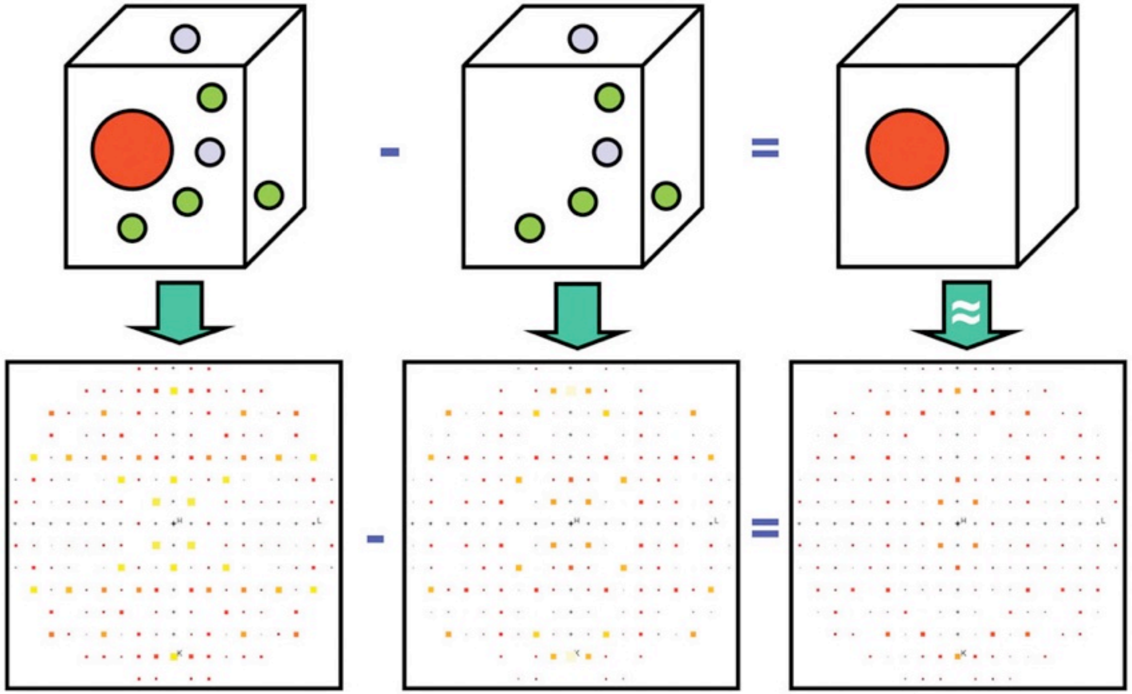
Heavy Atom Phasing

1. Introduce few heavy atoms (strongly interacting with X-rays) into crystal
    - Most common method: *In vivo* selenomethionine incorporation
2. Measure diffraction patterns of derivatized and normal crystals

   (or the anomalous signal of a single crystal)

3. Use difference data to solve small structure *ab initio*

4. Allows phase estimation for non-derivatized crystal



© Garland Science 2010

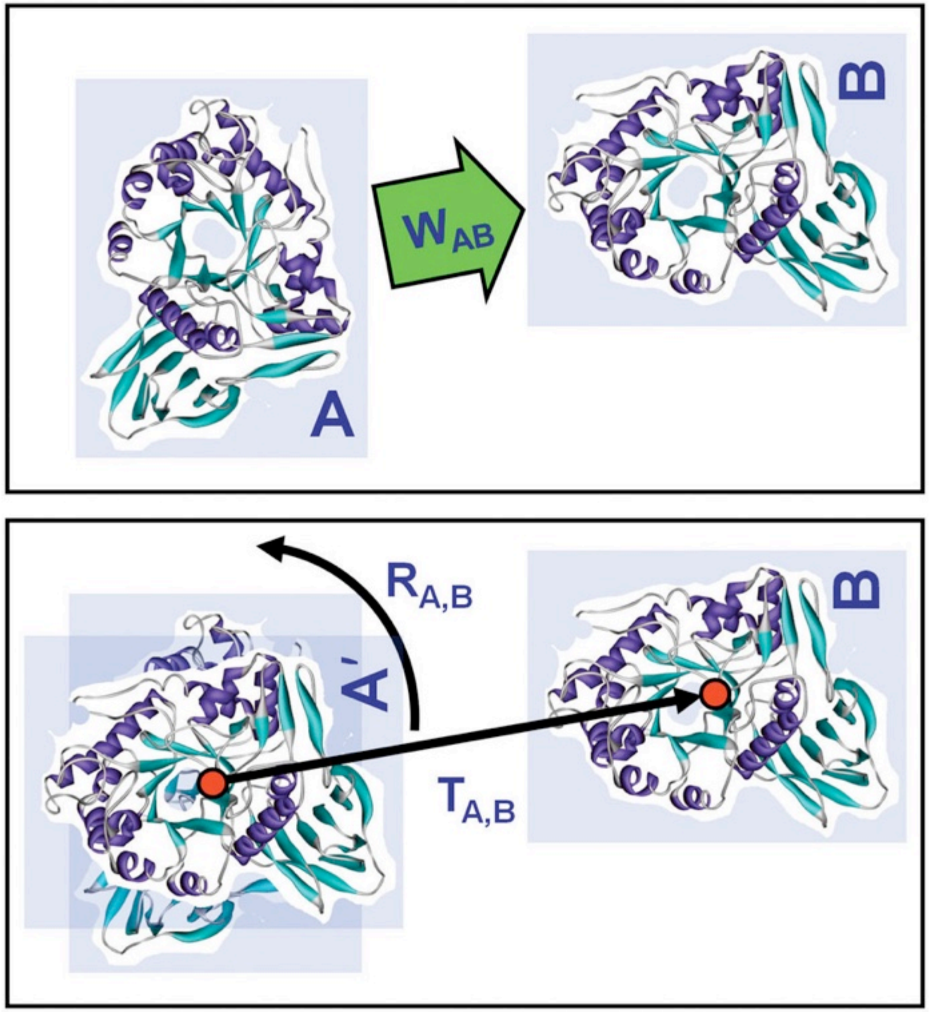
Molecular Replacement

1. Place a known structure similar to the crystallized one in the unit cell
2. Calculate phases from placed model
3. Use phases of model and measured amplitudes to calculate map
4. Adjust model (position + chemical structure) based on difference map

Searching correct model position is a 6-dimensional search!

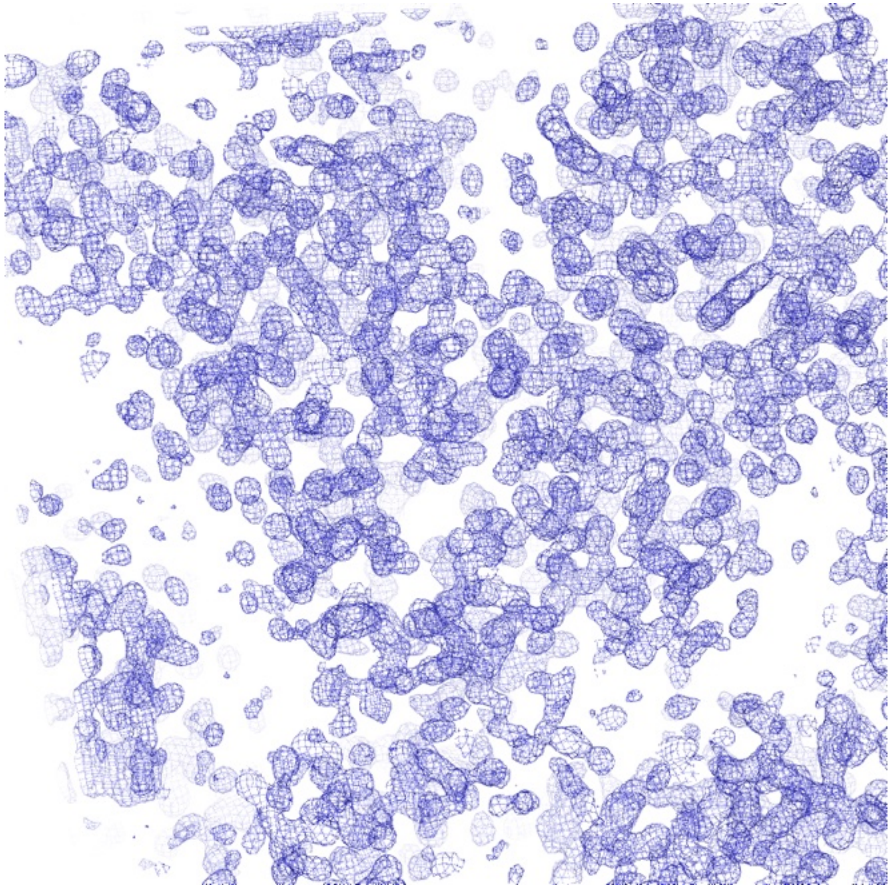
Three dimensions each for:

1. Rotation
2. Translation



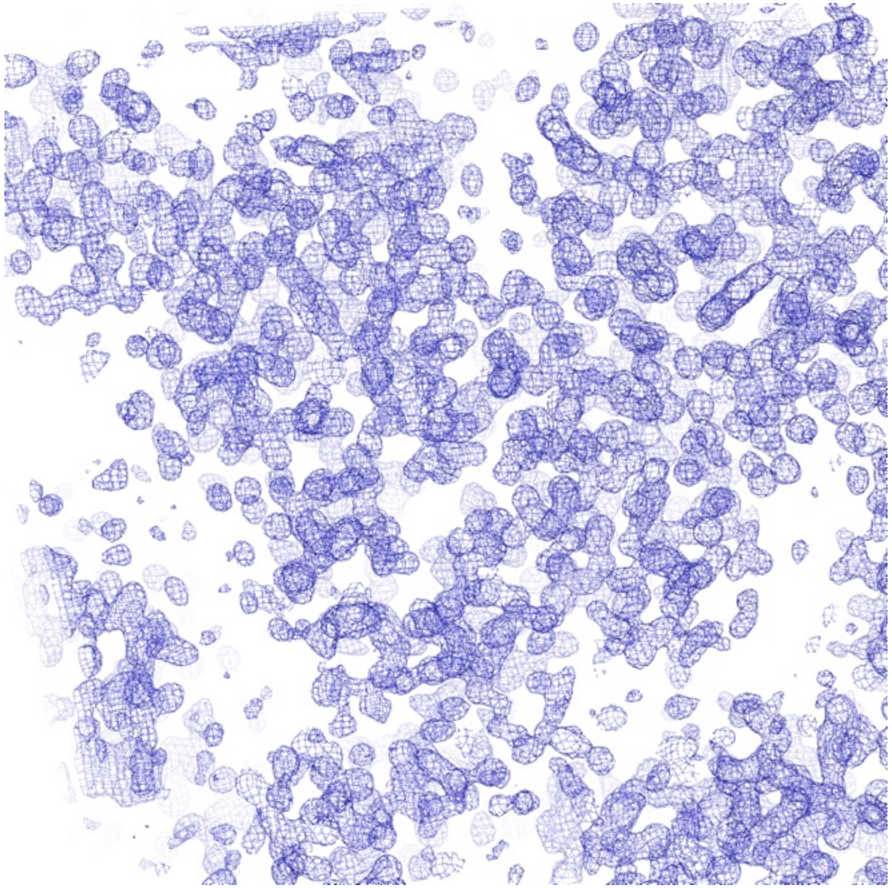
© Garland Science 2010

From electron density to atomic model

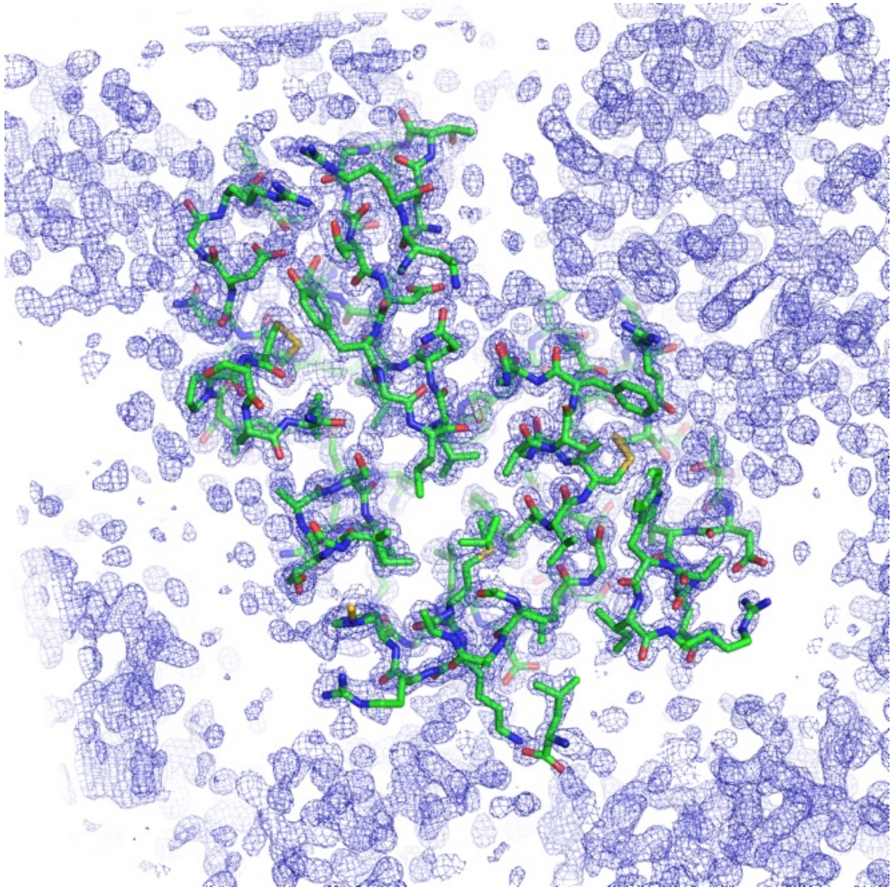


Electron density map
(Experimental Result)

From electron density to atomic model

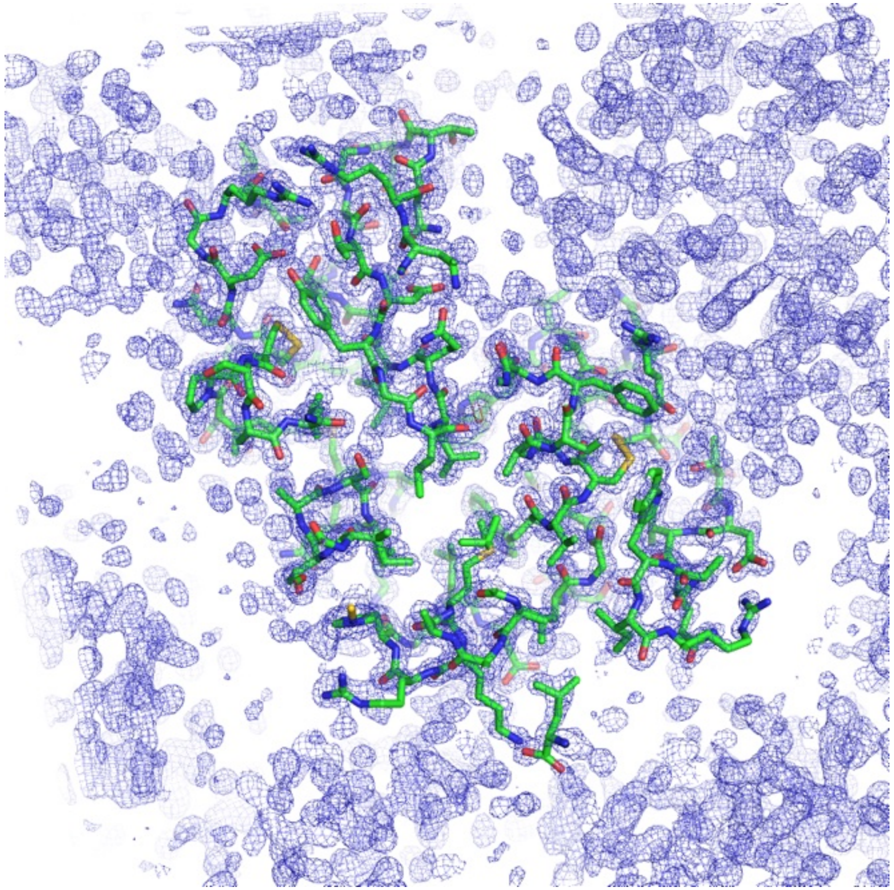


Electron density map
(Experimental Result)

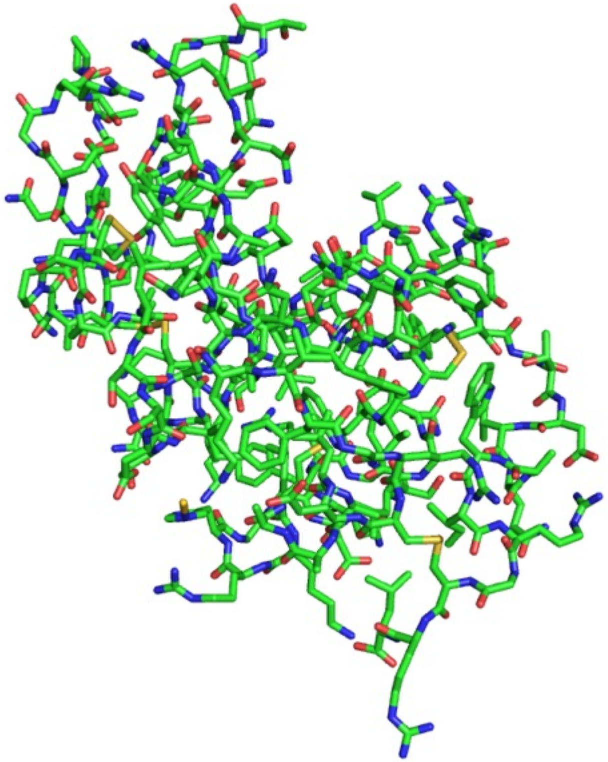


Interpretation as atomic structure
(Model building)

From electron density to atomic model

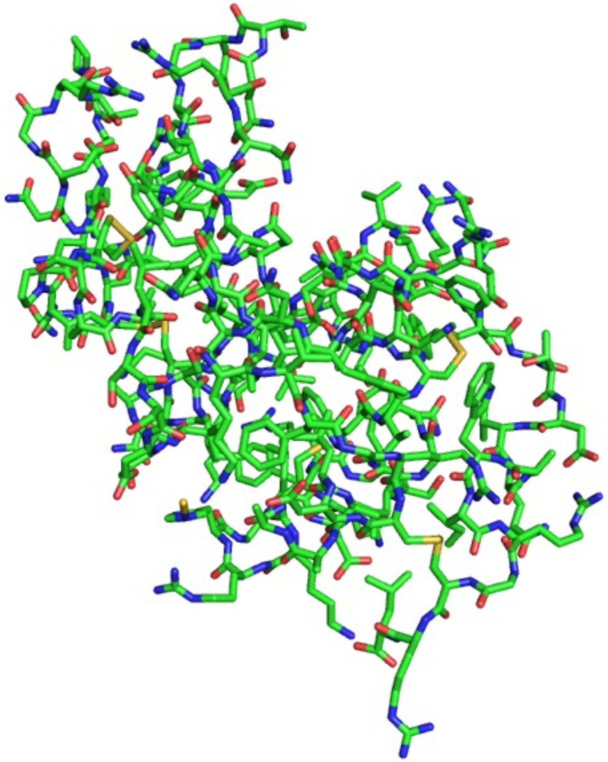


Interpretation as atomic structure
(Model building)

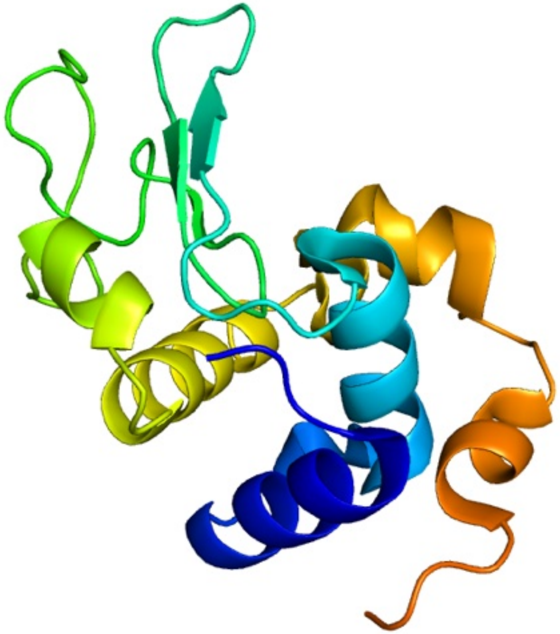


Atomic Model

From electron density to atomic model

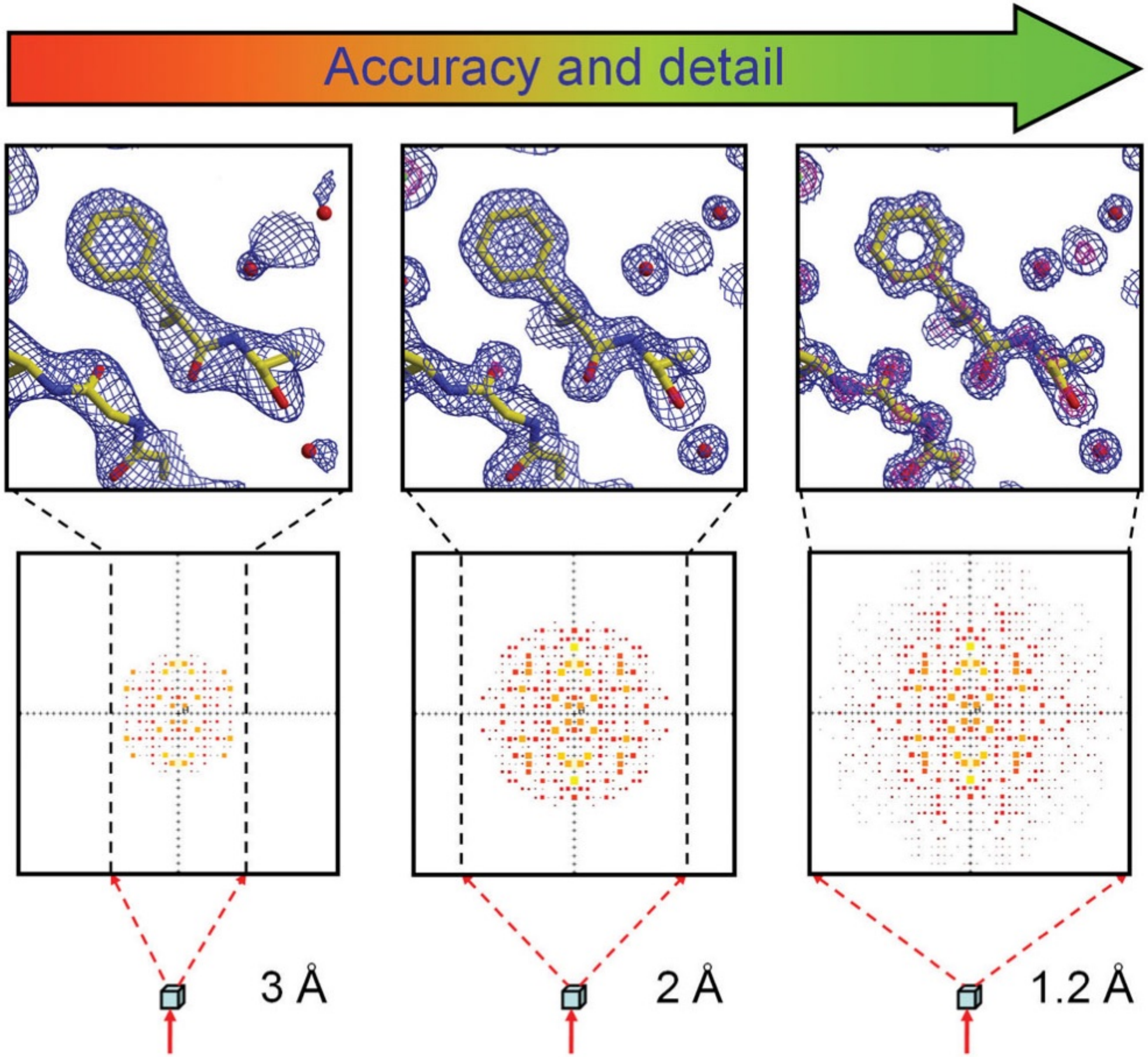


Atomic Model



Simplified cartoon representation

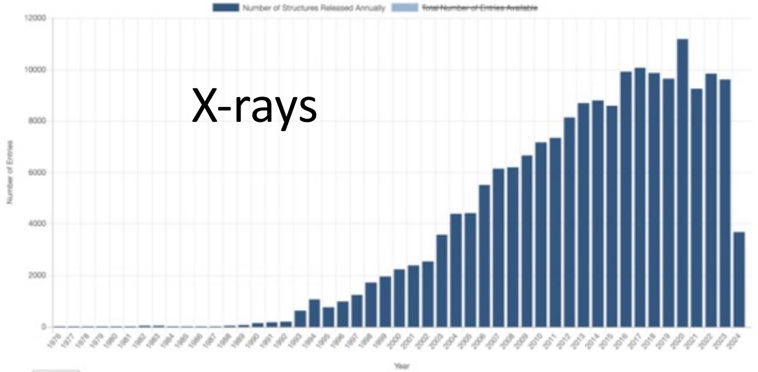
Resolution

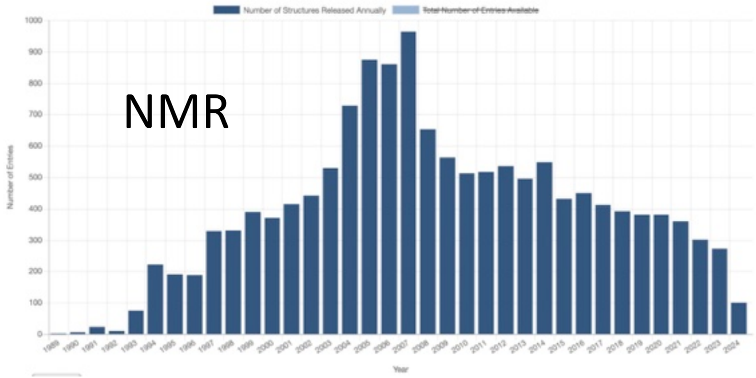


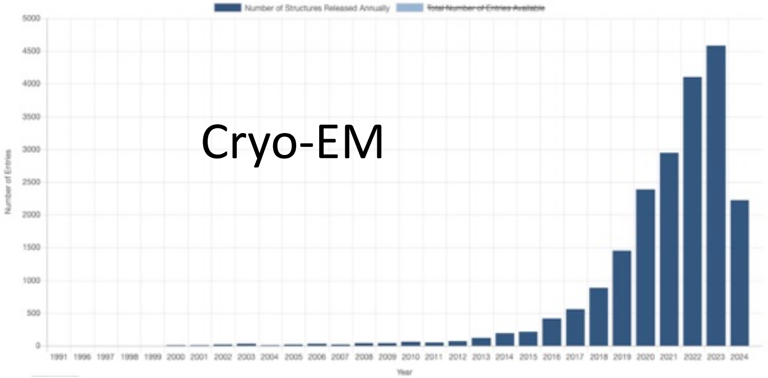
© Garland Science 2010

Deposition & Validation

- Processed Diffraction data and Models deposited in PDB
- Central repositories for all diffraction data
- Anonymous access free of cost for everyone
- Automatically and manually checked for errors
- Extensive validation
- Provides information for scientific reviewers







from www.rcsb.org , 27.5.2024

Opportunities & Limitations

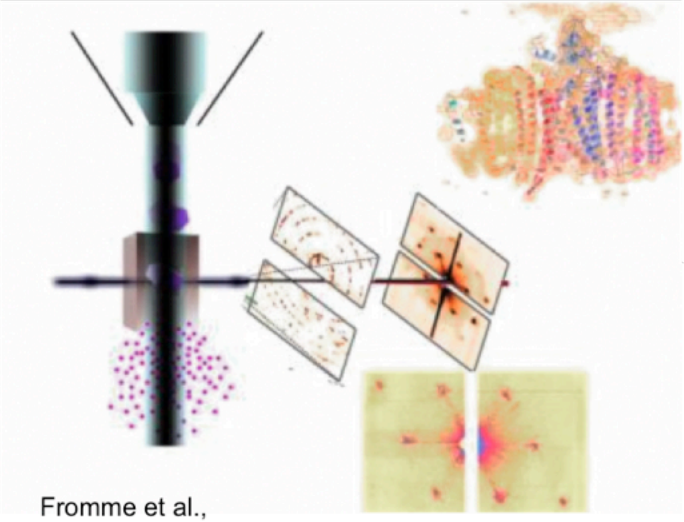
- All sizes: 50 Da-100 MDa
- Atomic Resolution
- Very fast for simple problems
- Robust technique
- High degree of automation
- No labeling required

- “Large” crystals required
- mg amounts of protein
- Radiation damage limits resolution for small crystals
- Provides static picture
- Highly disordered regions or proteins not accessible

Most successful method for structure determination at high resolution

The next step: SwissFEL (started end 2017)

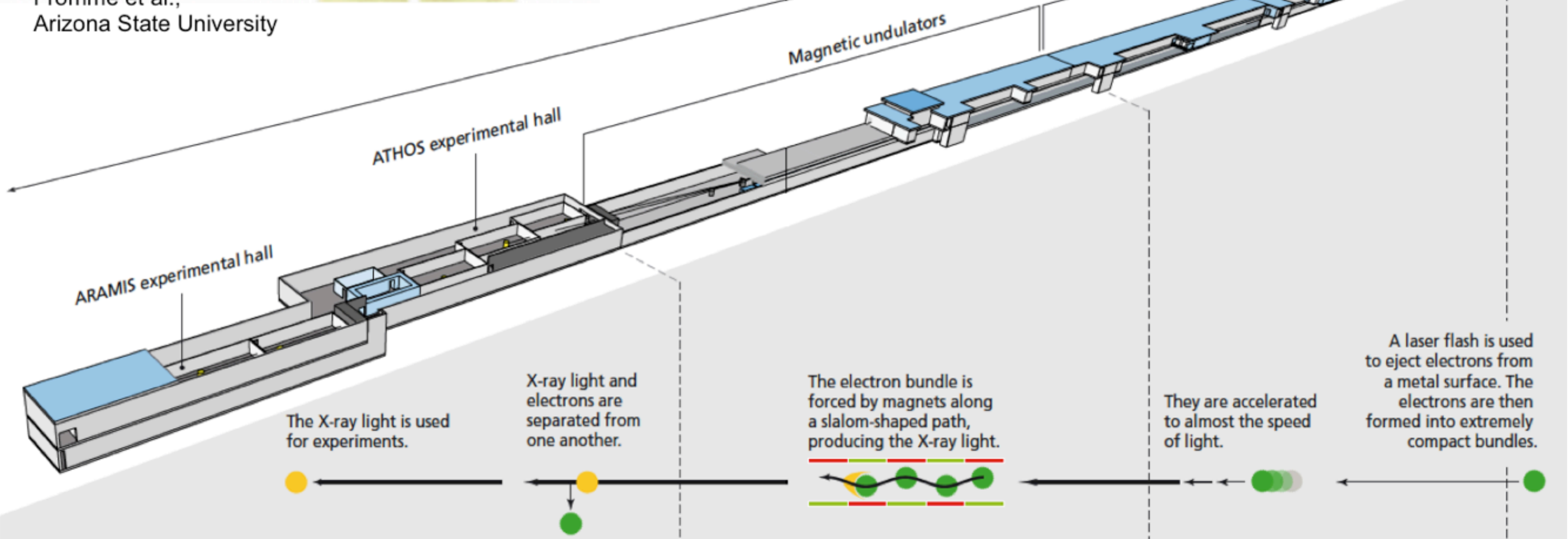
- Nano-crystals in solution at RT
- Diffract before Destroy
- Intracellular crystals
- fs-time resolved pump-probe experiments



Fromme et al.,
Arizona State University

PAUL SCHERRER INSTITUT

PSI



The X-ray light is used for experiments.

X-ray light and electrons are separated from one another.

The electron bundle is forced by magnets along a slalom-shaped path, producing the X-ray light.

They are accelerated to almost the speed of light.

A laser flash is used to eject electrons from a metal surface. The electrons are then formed into extremely compact bundles.

Syllabus X-ray Crystallography

- Dimensions of proteins, nucleic acids and wavelengths
- 100 years of crystallography
- X-ray diffraction vs. microscopy
- X-ray crystallography workflow
- Crystallization
  - properties of biological crystals
  - crystallization principles and techniques
  - crystallization artifacts
- X-ray data collection
  - properties of X-rays
  - generation and detection of X-rays
  - radiation damage
- Diffraction
  - Single crystal diffraction
  - Bragg’s law
  - Information content of position and intensity of reflections
  - Resolution
  - Images to amplitudes
- From amplitudes to model
  - Solving the phase problem (heavy atom phasing and molecular replacement)
  - Calculation of electron density
  - From electron density to model
  - Resolution and molecular detail
  - Deposition and validation in Protein Data Bank
- Opportunities and limitations