Haupt/Masterstudiengang Physik
Methoden moderner Röntgenphysik II:
Streuung und Abbildung
SS 2015

Biology I: Structural Biology and Macromolecular Crystallography (MX)

Thomas R. Schneider, EMBL Hamburg
25/6/2015

thomas.schneider@embl-hamburg.de

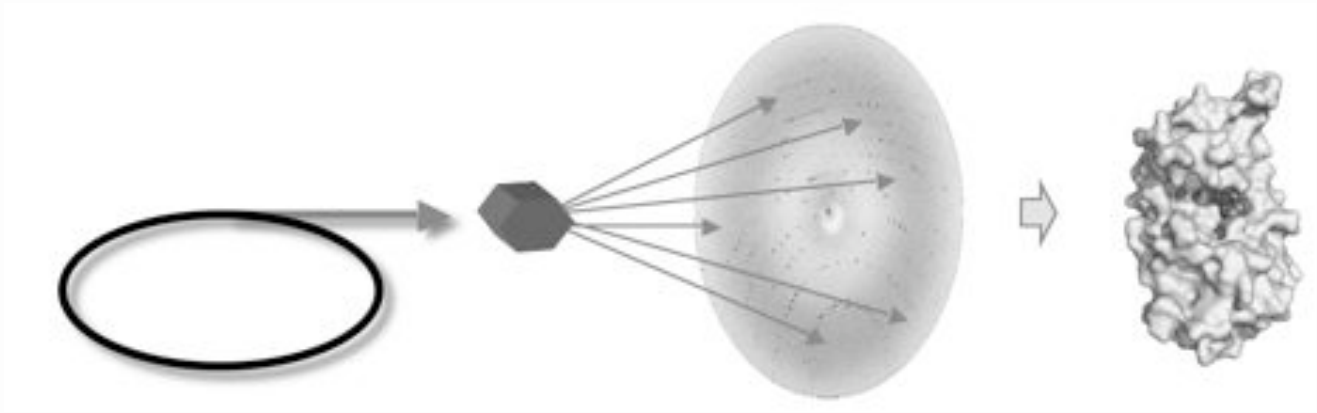
EMBL

EMBL

- Funded 1974
- Inter-governmental organization
- 21 member states (+ Czech Rep)
- Two associate states (+Argentina)
- Missions
  - Basic Research in Molecular Biology
  - Technology and Instrumentation
  - Facilities and Services
  - Teaching and Training
  - Technology Transfer
- 5 sites
- 1400 employees
  from 60 nations
- Annual Budget ca. 85 M€

Biology I to IV

- Structural Biology and Macromolecular Crystallography (MX)
- MX – The method
- MX – Collection and processing of diffraction data
- MX – Building models and the future



Today

- What is a protein?
- The central dogma of molecular biology
- Milestones in macromolecular crystallography
  - The first structures
  - Recombinant production of proteins
  - Use of synchrotron radiation
  - Detector technologies
  - The protein data bank
  - Cryogenic sample cooling
- Crystallographic Workflow

Look at Abl kinase

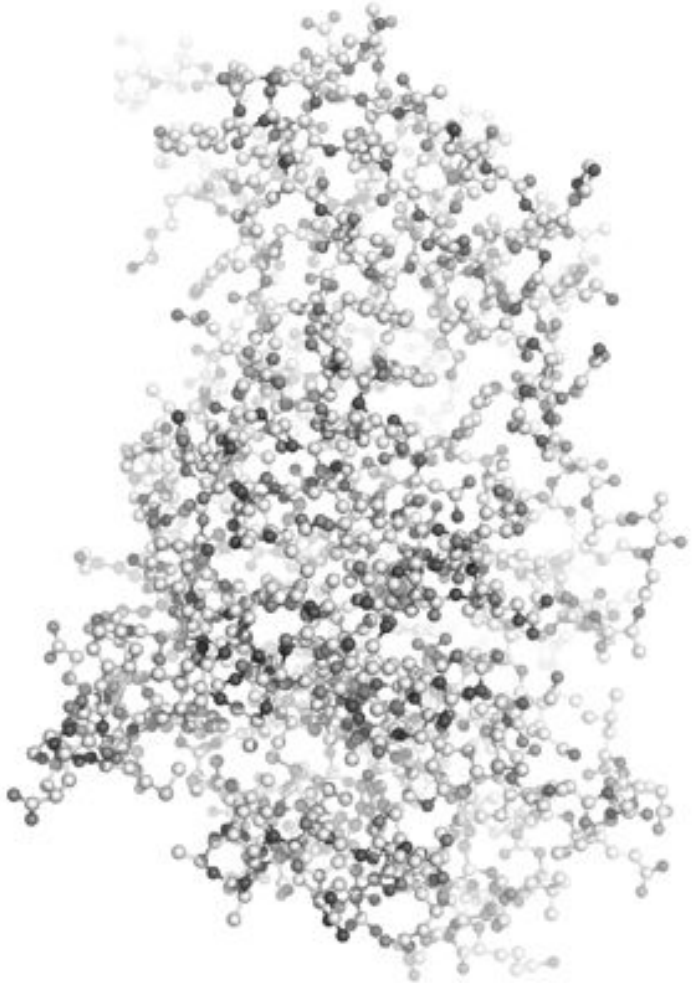
- 2148 atoms per molecule, 275 amino acids
- A protein consists of chained amino acids (20 types)
- Backbone vs. Sidechain
- alpha-helices
- beta-sheets
- Schematic 'cartoon' representations of proteins
- Surface representation of proteins (with properties mapped)
- Interactions with ligands can be studied
- Gleevec / Imantinib

- Folding (Stretch out 275 amino acids -> 3.8 cm x 275 = 10.41 m, compare to 53 A / cm)

Downloads:
molecule viewer: http://pymol.org/educational/
pdb-model: http://www.rcsb.org/pdb/explore/explore.do?structureId=1FPU
viewer script: see course website

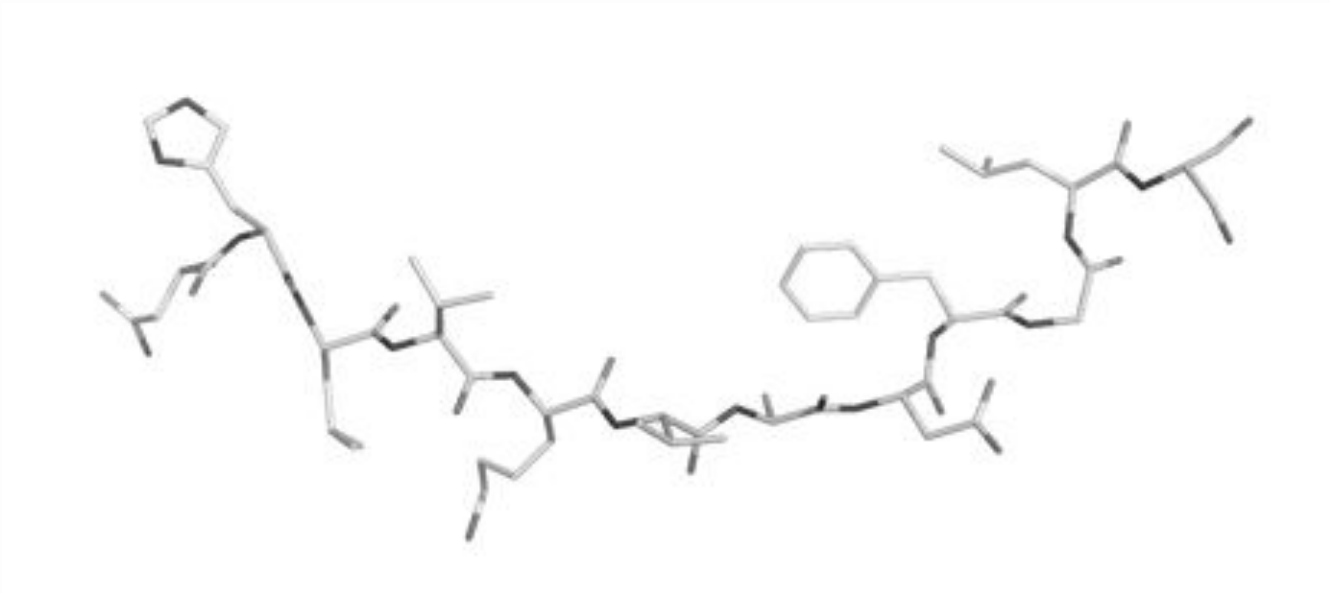
A protein molecule

- Looking at a collection of >2000 atoms can be rather confusing:



Polypeptide chain

- Proteins are heteropolymers of amino-acids:

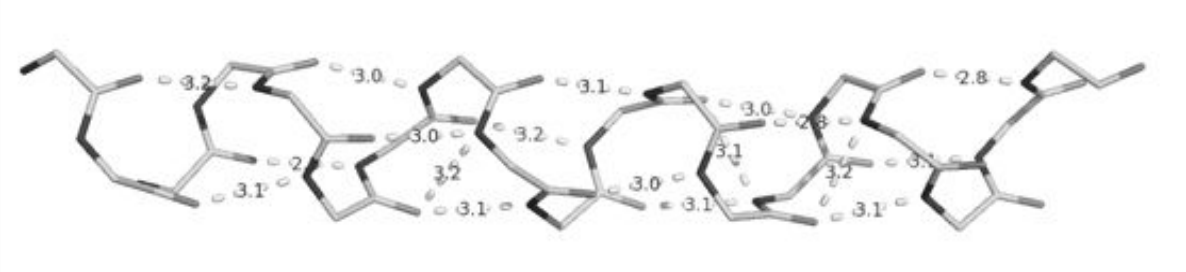


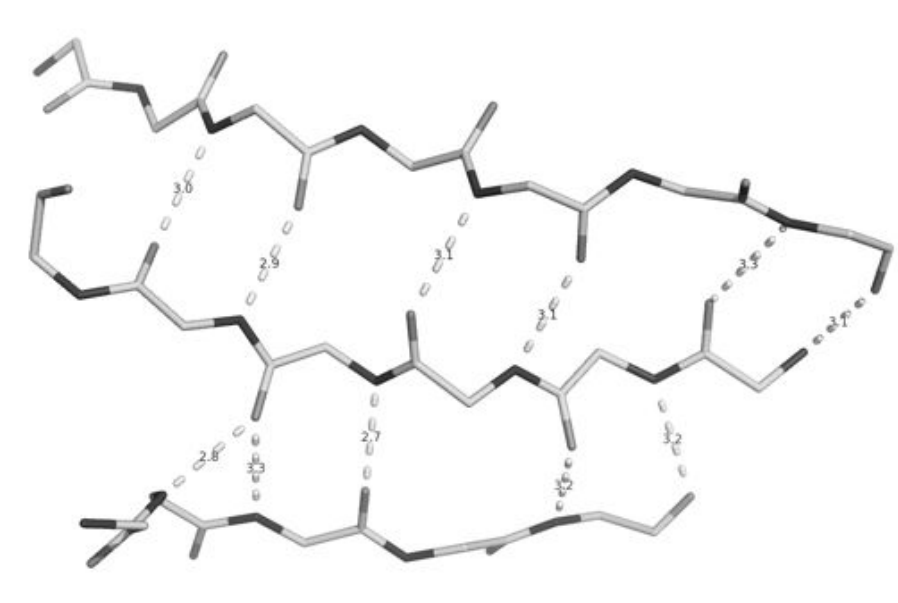
Protein Backbone

- Looking at the backbone only allows to follow the polypeptide chain across the protein

Protein conformation

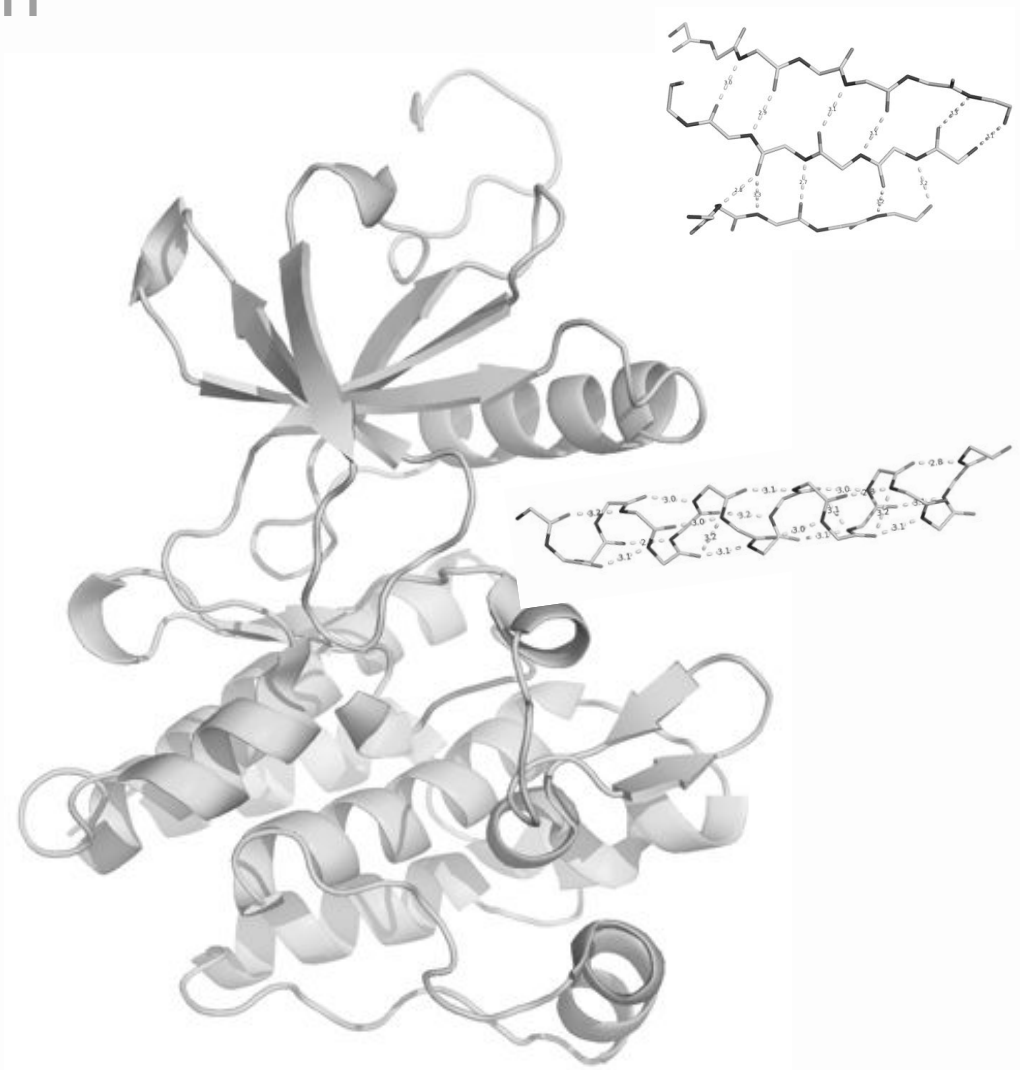
- alpha helices and beta sheets contain repeating pattern / hydrogen bonds:





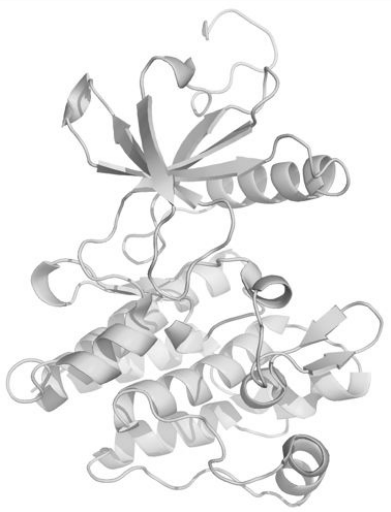
Protein conformation

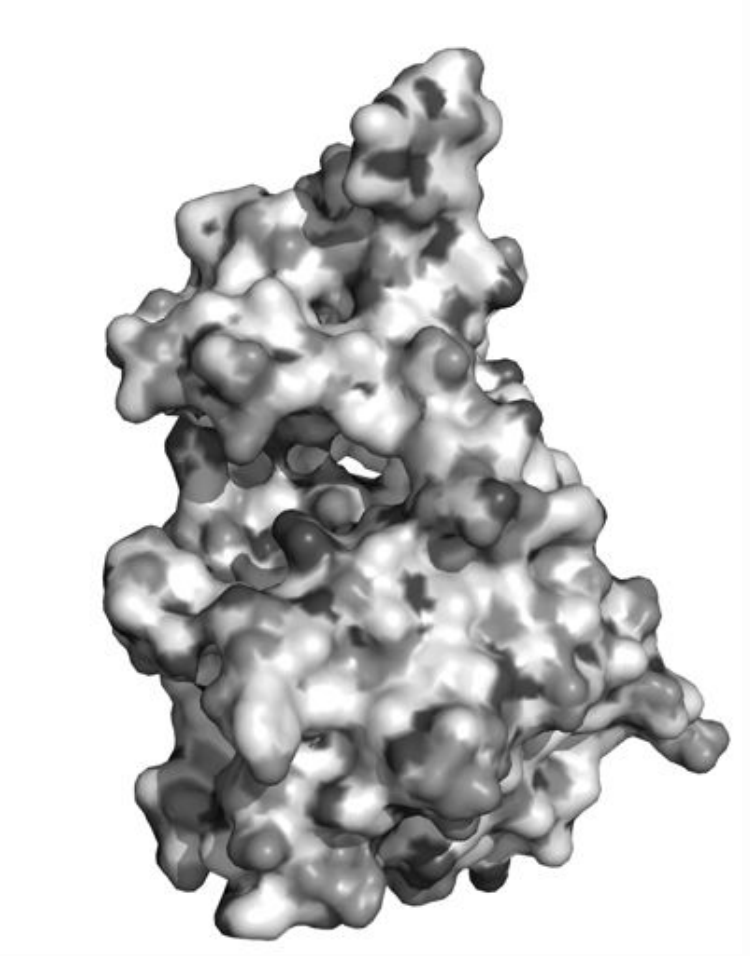
- Schematic representation of helices and sheets illustrates the 'fold'



Protein surfaces

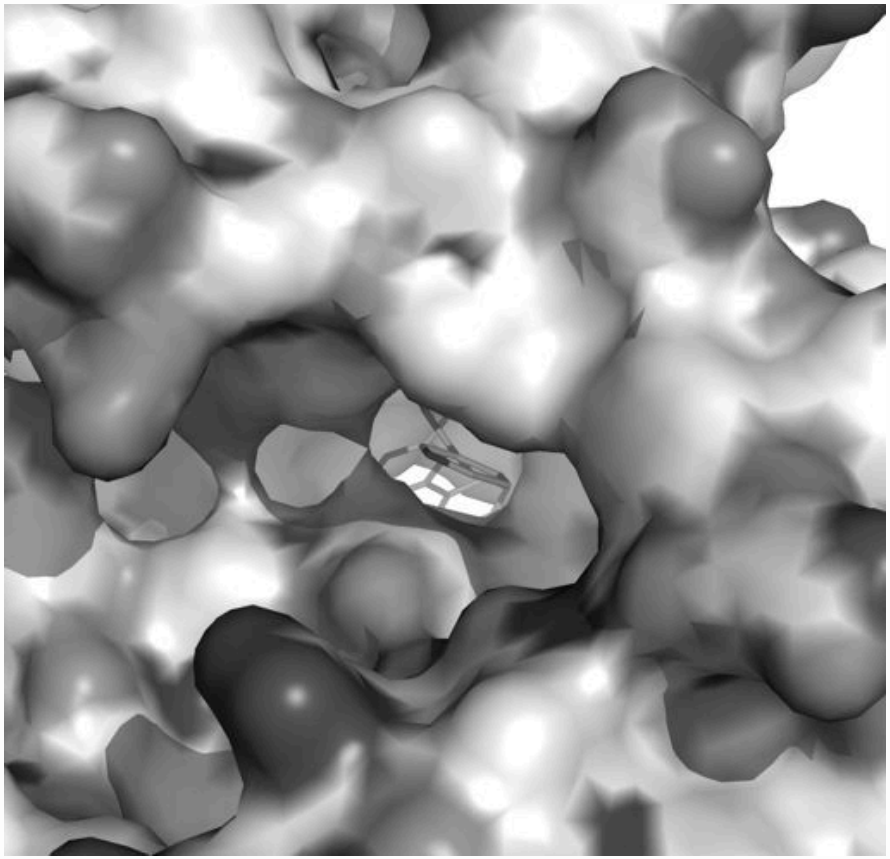
- What other molecules 'see' is a molecular surface.
- Active sites are often found in cavities





Active sites

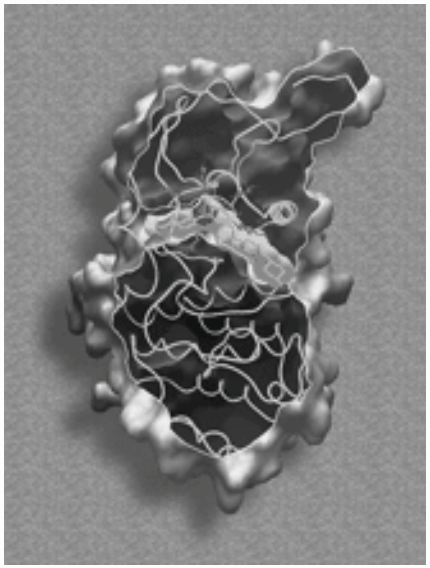
- Knowledge of the active site geometry allows to design ligands binding to the active sites thus acting as drugs.
- Here: Gleevec binding Abl-Kinase



Crystal structures can be used to understand the action of drugs

The drug

Mortality rate reduced by 80% for patients where interferon did not work

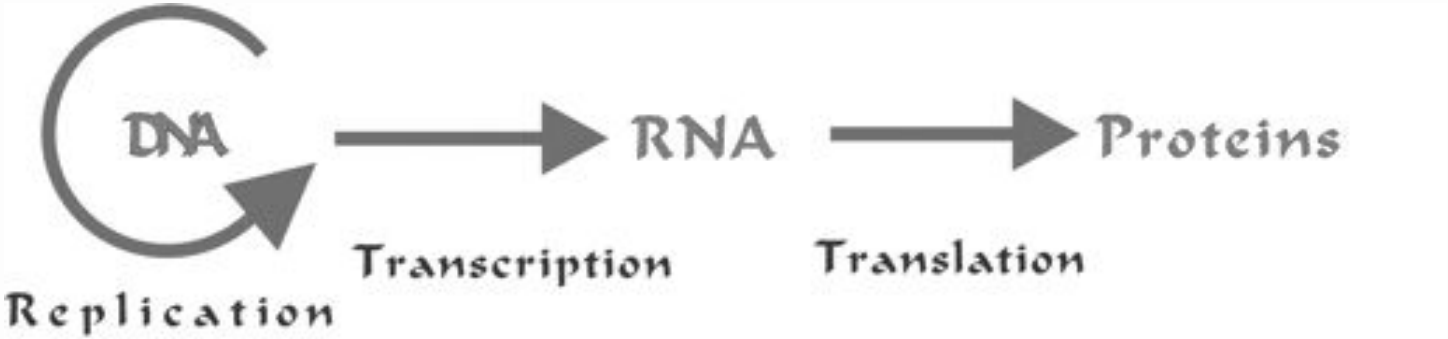


Crystal structure of the drug bound to its target Abl-Kinase

red shows resistant mutation

The Central Dogma of Molecular Biology

- 'DNA makes RNA makes protein' (Nirenberg 2010)

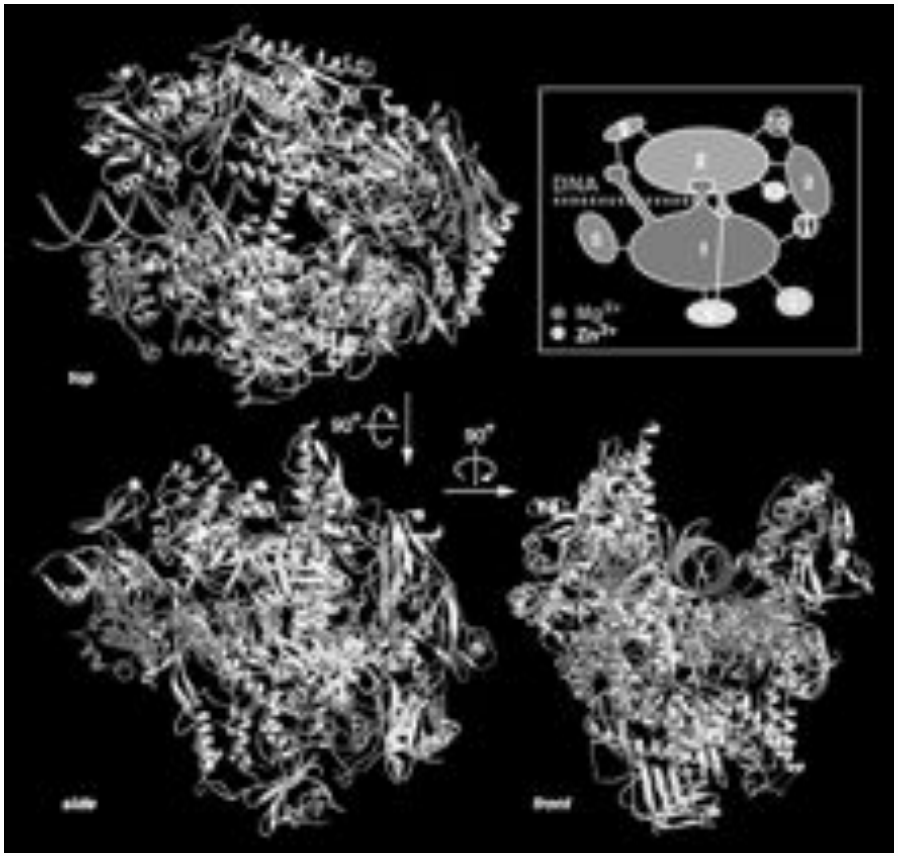


http://library.thinkquest.org/C0122429/intro/genetics.htm

- Transcription: 'Umschreiben' of DNA (double-stranded, protected) to RNA (single-stranded, code accessible).
- Translation: Following the sequence of triplet-codons in the RNA molecule, a protein is assembled from amino acids.

Crystal Structures support Basic Biology

- RNA Polymerase II – Structural Basis of Transcription
- 12 proteins
- ~30000 atoms
- Nobel Prize 2006 to Roger Kornberg



http://www.youtube.com/watch?v=6QMPU9nuQso
http://www.lmb.uni-muenchen.de/cramer/pr-materials/index.htm

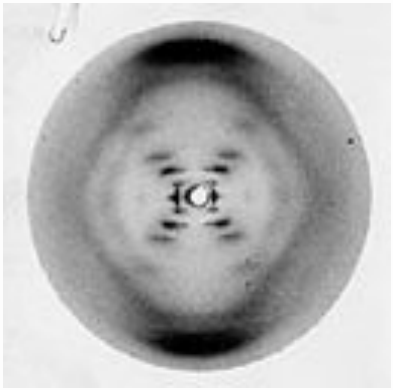
Crystal Structures support Basic Biology

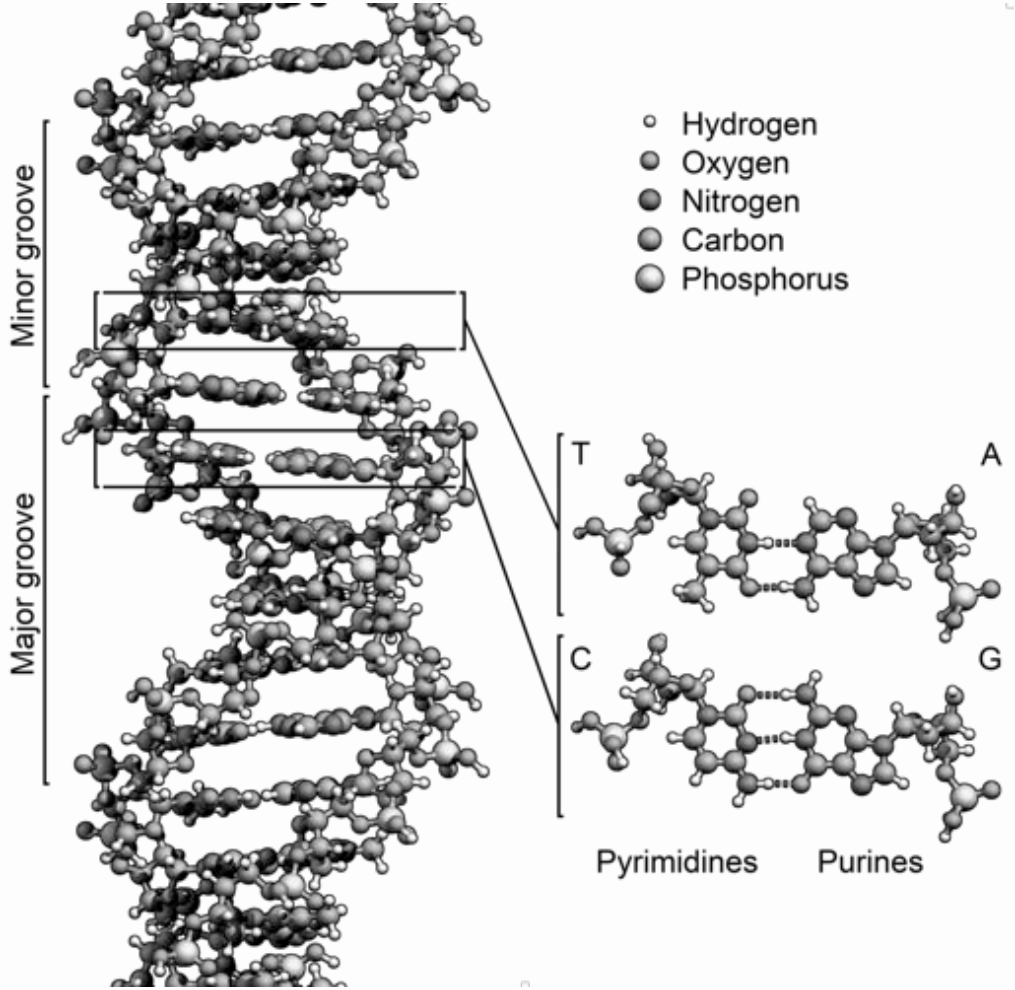
- Ribosome – Structural Basis of Translation
- 3 stretches of RNA, 52 proteins
- ~37000 RNA atoms, ~22000 protein atoms
- Nobel Price 2009 to Steitz, Yonath, Ramakrishnan

http://www.youtube.com/watch?v=Jml8CFBWcDs

Some milestones

DNA (1953)





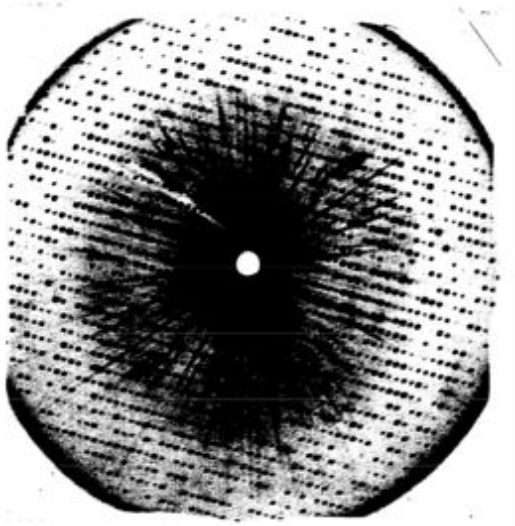
Watson & Crick (1953) Nature 172:137

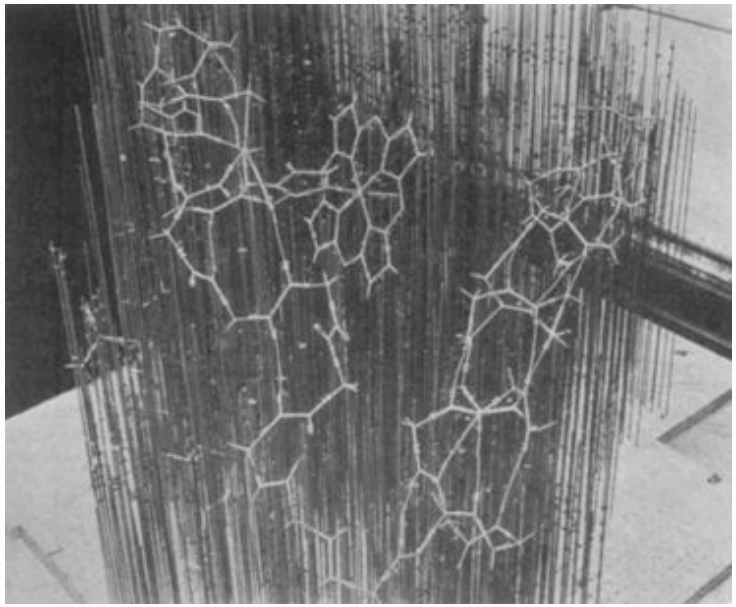
http://www.nobelprize.org/nobel_prizes/medicine/laureates/1962/

http://commons.wikimedia.org/wiki/File:DNA_Structure%2BKey%2BLabelled.pn_NoBB.png

Myoglobin (1960)

- Started in 1954
- Structure at 2 A resolution published in 1960
- Key-step: choose sperm-whale as source





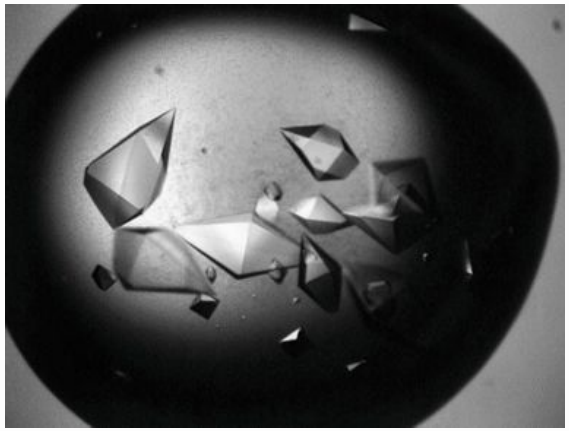
KENDREW, J. C. et al. (1960) Nature, 185, 422–427.

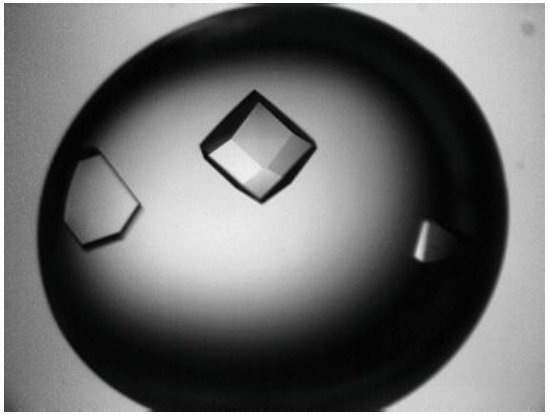
http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1962/kendrew-lecture.html

pdb: 1MBN

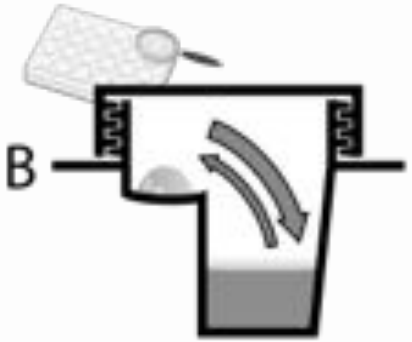

Crystals
Are difficult to grow
Are difficult to reproduce
Are often very small (microns)
Are often inhomogeneous
Are mechanically fragile
Are radiation sensitive
Here: Cueva de los Cristales, Chihuahua region, Mexico

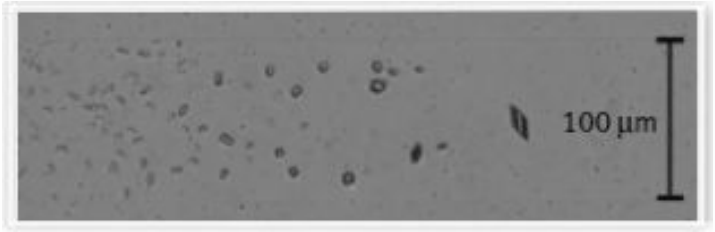
Finding Conditions

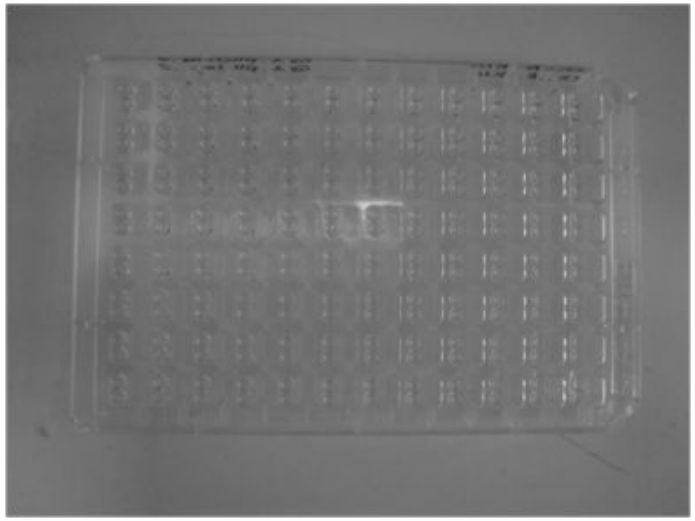




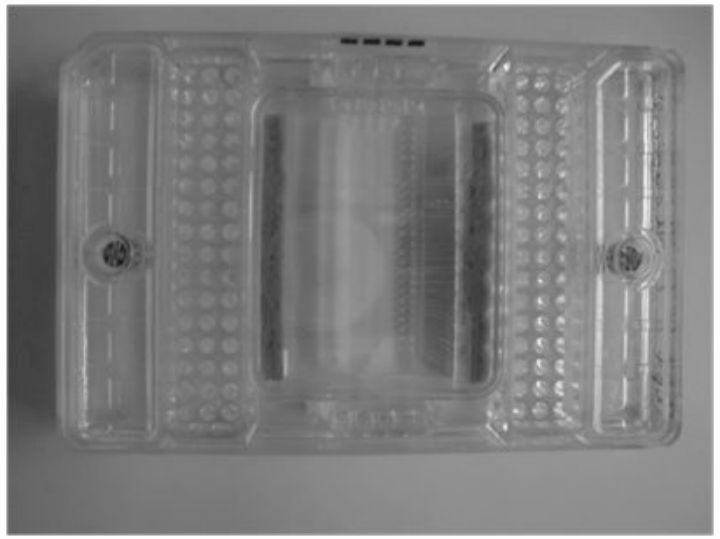
Small Volume HTP Crystallization







96 x 50-200 nl

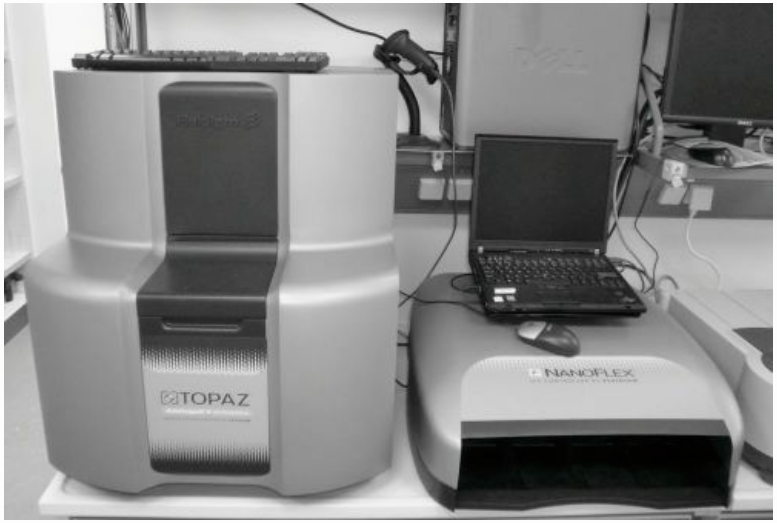


96 x 10 nl

High-Throughput Crystallization at EMBL Hamburg

- Miniaturization
  - 200nl/exp for vapour diffusion
  - 10nl/exp for counter diffusion

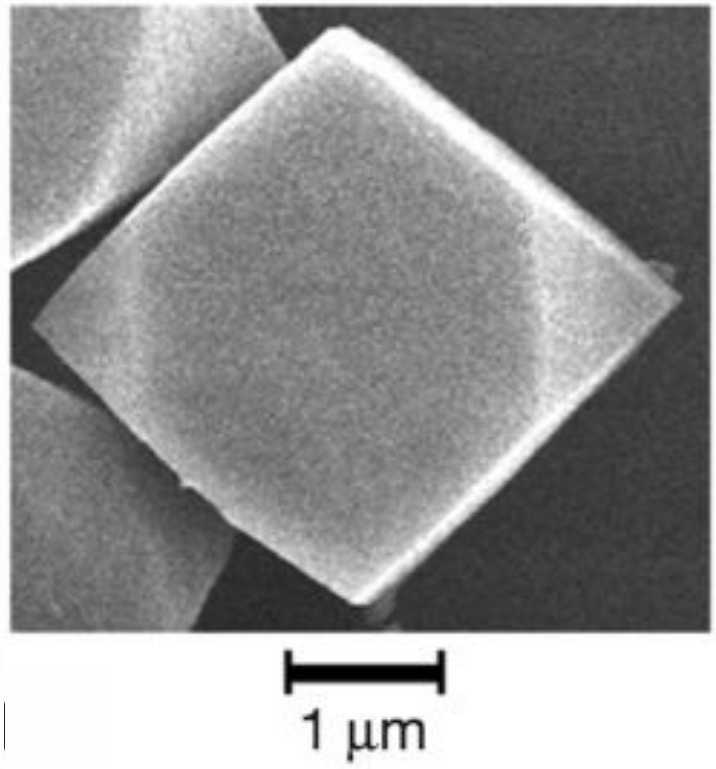
- Reference:
  - J. Müller-Dieckmann (2006) Acta Cryst. D64:1146-1152.



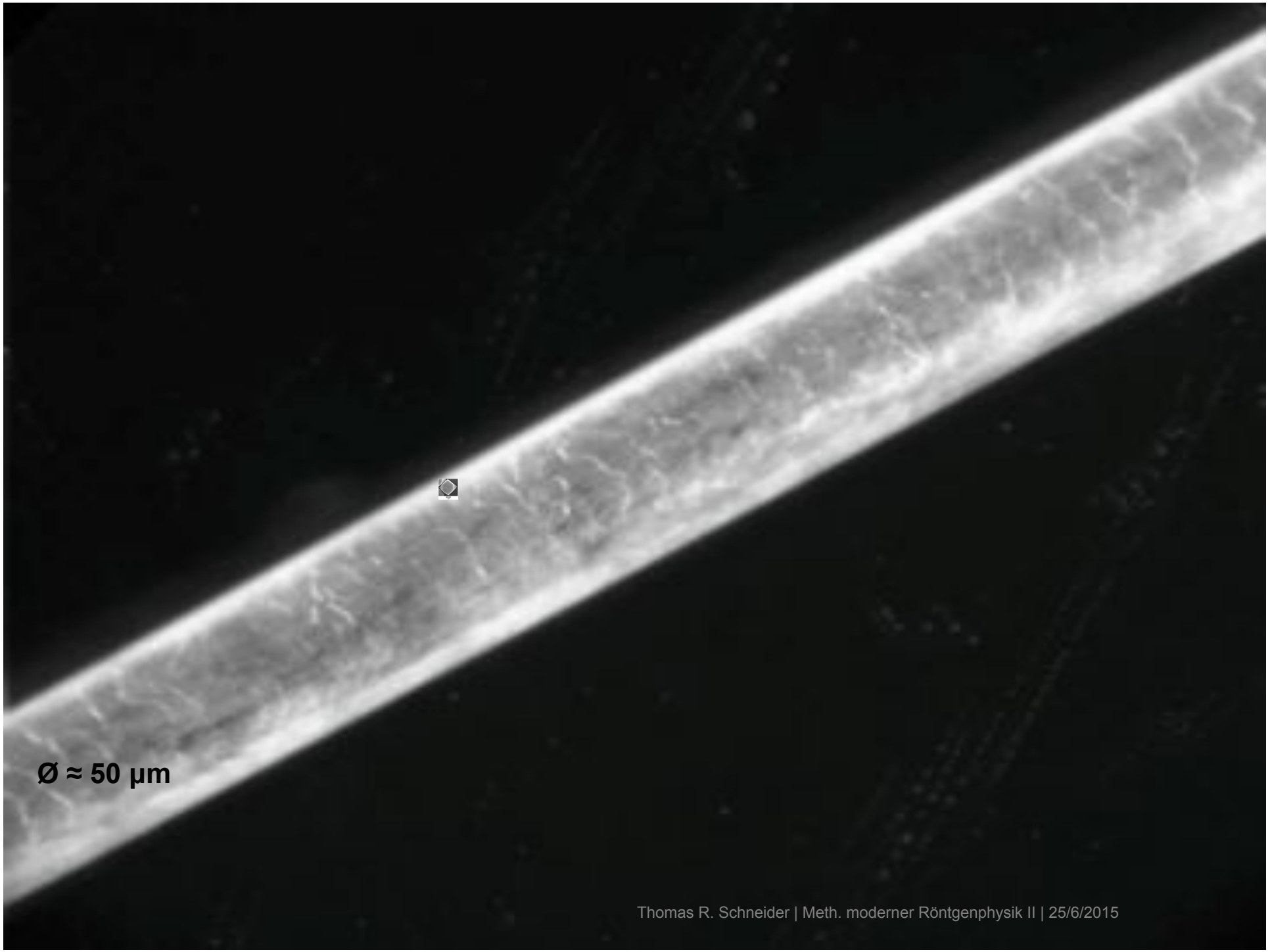
Jochen Müller-Dieckmann
Xandra Kreplin

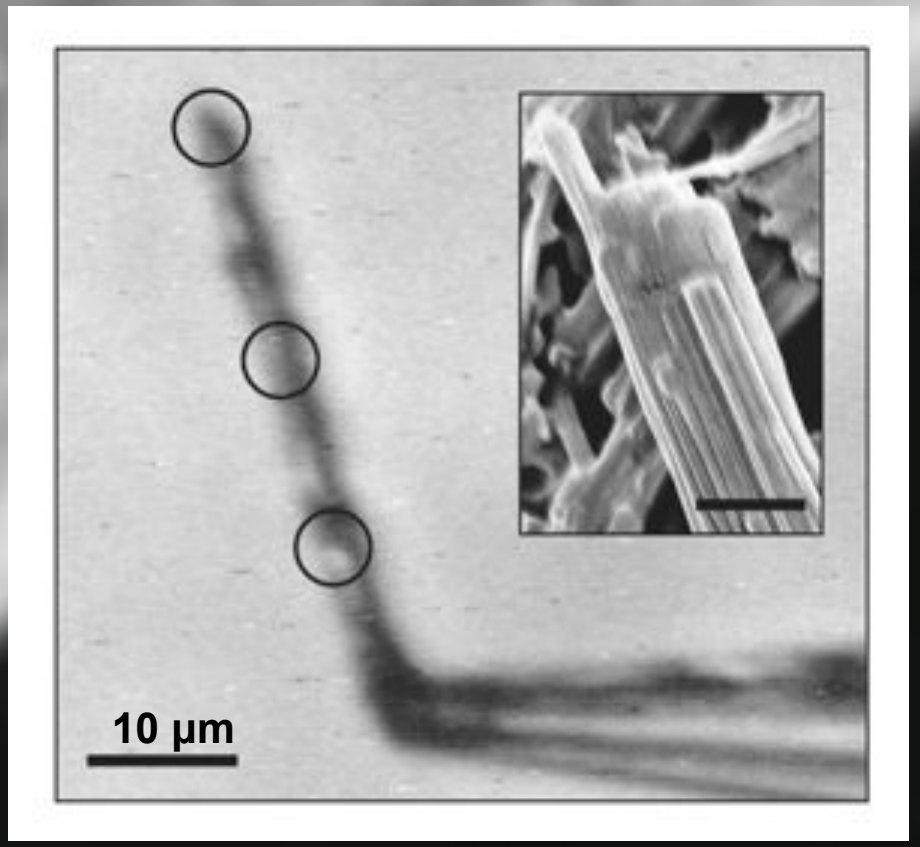
Small Crystals

- Crystal dimensions of 1-10 micron are not rare.
- Often these crystals are of high quality
- Small and parallel beams needed.

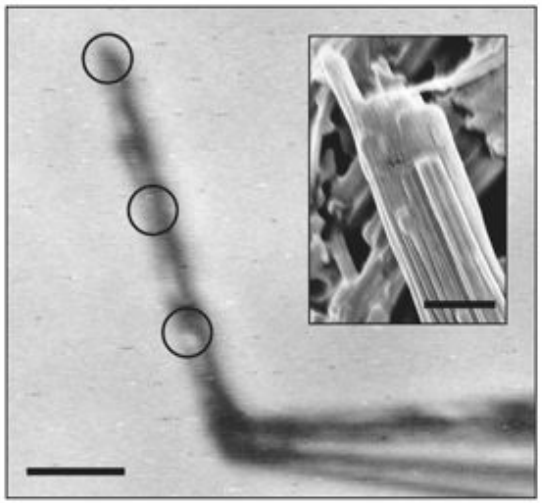


Coulibaly et al. The molecular organization of cypovirus polyhedra. Nature (2007) 446: 97-101

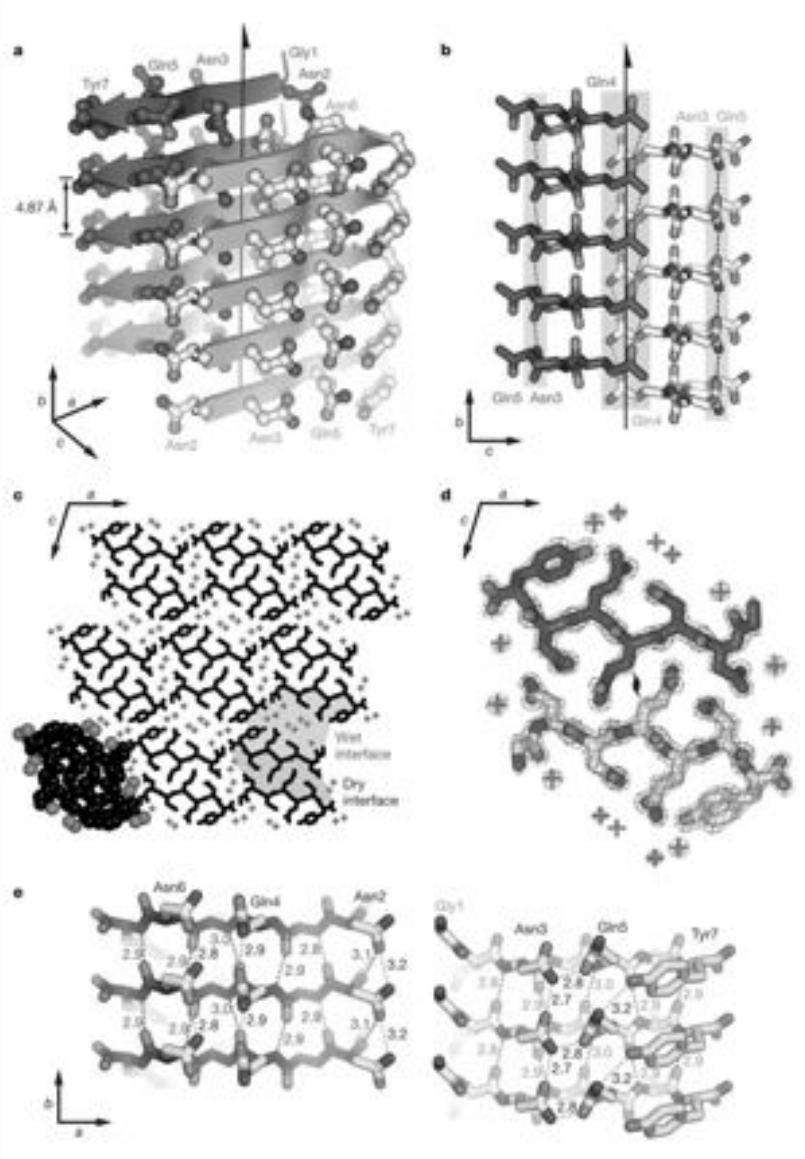

Ø ≈ 50 µm


10 µm

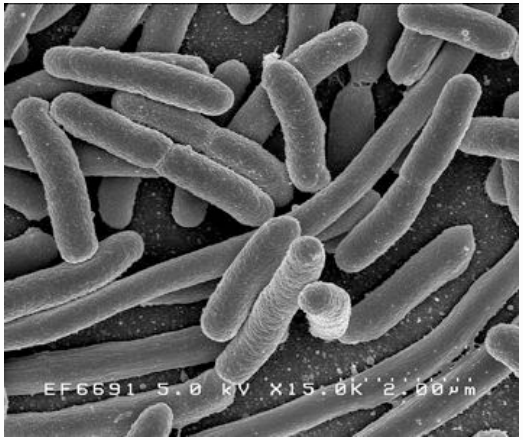
Amyloid fibrils



Nelson et al.Structure of the cross-beta spine of amyloid-like fibrils. Nature (2005) 435: 773-8

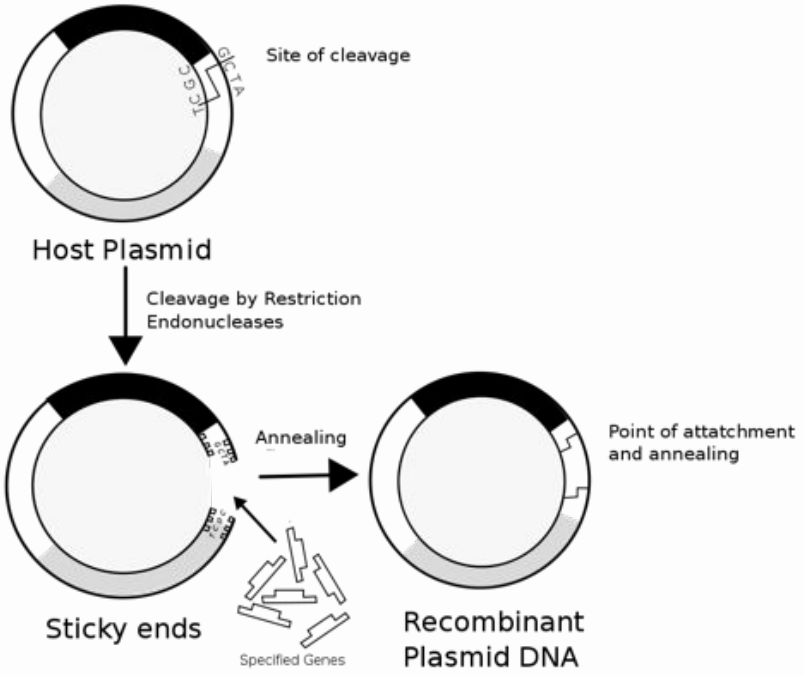


Recombinant production of proteins



- Insert DNA from other organism into bacteria (Lobban?, 1972) using enzymes as tools for the manipulation of DNA
- 1982 synthetic human insulin (Genentech & Eli Lilly) entered the market for diabetes therapy.

- Recombinant production (and 'overexpression') of protein molecules is crucial for macromolecular crystallography as large amounts of material are needed to produce crystals.



Use of synchrotron radiation in Biology

Synchrotron Radiation as a Source for X-ray Diffraction

G. ROSENBAUM & K. C. HOLMES
Max-Planck-Institut für Medizinische Forschung, Heidelberg
J. WITZ
Laboratoire des Virus des Plantes, Institut de Botanique de la Faculté des Sciences de Strasbourg, Strasbourg

Some preliminary results have been obtained with synchrotron radiation from the 7.5 GeV electron synchrotron Deutsches Elektronen - Synchrotron (DESY) in Hamburg as a source for X-ray diffraction.

When an electron is accelerated it emits radiation. At the very high energies used in DESY, the emitted radiation is confined to a narrow cone about the instantaneous direction of motion of the electron. Thus the synchrotron radiates tangentially. Synchrotron radiation is polychromatic, with a peak in the X-ray region for an electron energy of 7.5 GeV (see ref. 1 for the original theoretical description and refs. 2–4 for experimental details).

The DESY synchrotron uses bursts of 50 pulses/s and each 10 ms pulse contains 6×10^{10} electrons (10 mA average beam current). The injection energy is relatively low and the electrons are accelerated up to 7.5 GeV in the 10 ms.

Most of the X-radiation is emitted during the last 3 ms of each pulse: little radiation is produced at the lower electron energies, and so the time averaged intensity at 1.5 Å is about 20% of the peak value.

Table 1 Data for Quartz Monochromator in Synchrotron Radiation Beam

Synchrotron	7.5 GeV, 10 mA beam current
Electron beam diameter	approximately 6 mm (= effective X-ray source diameter)
Distance	15 m from synchrotron to monochromator
Cross-fire of the incident beam	approximately 10^{-3} rad
Polarization	85% at 1.5 Å in the eighth ms of the cycle, polarized in the plane of the synchrotron
Be-window	0.5 mm (96 mg cm^{-2})
Crystal	quartz cut at $\alpha = 8°$ 30′ to the 1011 planes, dimensions 45 × 13 × 0.3 mm^3
Bender	pins: outer pair 40.5 mm inner pair 39.5 mm radius of curvature of crystal, 9 m
Wavelength	1.53 Å (θ = 13° 15′)
Wavelength spread	$\Delta\lambda = 3 \times 10^{-3}$ Å (due to deviation from Johann focusing and to finite source size)
Focus	1.5 m from crystal line focus 180 μm wide
Angular aperture of reflected beam	horizontal: 2 mrad (convergence) vertical: 3–4 mrad (divergence)
Measured flux in line focus	1.8×10^9 photons s^{-1} mm^{-1} (of focal length) (at the eighth ms of the cycle)

Table 2 Biological Applications

Specimen	Elliott fine-focus X-ray tube*	DESY synchrotron with Berreman point-focusing monochromator †
Single crystal	Standard collimator 0.5 mm diameter	
$a = 0.5$ mm $b = 0.5$ mm $L = 7.5$ cm	$A = 12.5$ cm $d = 0.7$ mm $P = 10^9$ photons s^{-1} $I = 2 \times 10^6$ photons s^{-1} mm^{-2}	$D = 1$ m $d = 120$ μm $P = 4 \times 10^9$ photons s^{-1} $I = 2.5 \times 10^{11}$ photons s^{-1} mm^{-2}
Tobacco mosaic virus gel	Double-crystal focusing monochromator ‡	
$a = 0.6$ mm $b = 1$ mm $L = 12$ cm	$d = 80$ μm $P = 10^7$ photons s^{-1} $I = 2 \times 10^4$ photons s^{-1} mm^{-2}	$D = 0.8$ m $d = 100$ μm $P = 3 \times 10^8$ photons s^{-1} $I = 3 \times 10^{11}$ photons s^{-1} mm^{-2}
Insect muscle	Double-crystal focusing monochromator ‡	
$a = 3$ mm $b = 0.3$ mm $L = 40$ cm	$d = 100$ μm $P = 5 \times 10^7$ photons s^{-1} $I = 3 \times 10^5$ photons s^{-1} mm^{-2}	$D = 1.5$ (3) m $d = 180$ (350) μm $P = 5 \times 10^8$ (2×10^9) photons s^{-1} $I = 1.5 \times 10^{10}$ photons s^{-1} mm^{-2}

a, Width of specimen; b, height of specimen; L, specimen film distance; A, anode specimen distance; D, focal length, that is, monochromator film distance; d, spot or focus diameter on film; P, X-ray power reaching the specimen; and I, flux density at the focus.

* Loaded with 40 kV, 50 mA into a 0.2×2 mm^2 electron focus at the anode in the first case, and 40 kV, 15 mA into a 0.14×0.7 mm^2 focus in the other two cases. This set is the most powerful fine-focus X-ray tube currently available.

† The setting of this Johann-type[5] monochromator is optimized for each type of specimen.

‡ Conditions of the synchrotron are as in Table 1, computed for 1.5 Å radiation.

We have evaluated the spectral luminance (that is, the power in photons per second radiated per unit area, solid angle, and wavelength interval) of both the synchrotron and a fine-focus rotating anode X-ray tube (see Table 2). The values are 2×10^{11} (time averaged) and 3×10^{10} photons s^{-1} sterad^{-1} cm^{-2} Å^{-1} respectively at 1.54 Å, showing clearly that the synchrotron is, relative to present X-ray tubes, a very bright source. The actual advantage to be gained in a diffraction experiment depends critically on the optical system necessary to focus and monochromate the radiation. Three types of focusing mono-

© 1971 Nature Publishing Group

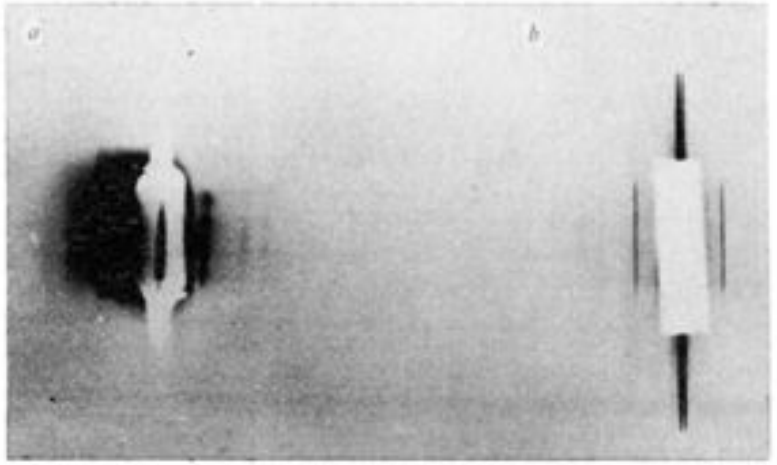


Fig. 3 Equatorial reflexions from dorsolongitudinal flight muscle of *Lethocerus maximus* recorded with: *a*, monochromated synchrotron radiation; electron energy 5 GeV, beam current 8 mA, exposure time 15 min, specimen film distance 40 cm; note the parasitic scattering on the left of the backstop arising from fluorescence from the monochromator holder; *b*, Elliott fine-focus rotating anode tube at 40 kV, 15 mA, exposure time 1 h, specimen film distance 36 cm. The strong line is the 20 reflexion ($d = 231$ Å); the weak lines are the 21, 31 and 32 reflexions.

Use of synchrotron radiation in MX

Applications of synchrotron radiation to protein crystallography: Preliminary results

(x-ray diffraction/anomalous dispersion/rubredoxin/azurin/nerve growth factor/glutaminase-asparaginase)

JAMES C. PHILLIPS, ALEXANDER WLODAWER, MARGUERITE M. YEVITZ, AND KEITH O. HODGSON*

Department of Chemistry and Stanford Synchrotron Radiation Project, Stanford University, Stanford, California 94305

Communicated by Richard H. Holm, October 23, 1975

ABSTRACT X-ray diffraction photographs of protein single crystals have been obtained using synchrotron radiation produced by an electron-positron storage ring. The diffracted intensities observed with this unconventional source are a factor of at least 60 greater than those obtained with a sealed x-ray tube using the same crystal and instrumental parameters. Diffraction data have been collected by the precession method to higher resolution and using smaller protein crystals than would have been possible with a conventional source. The crystal decay rate in the synchrotron beam for several proteins appears to be substantially less than that observed with Ni-filtered Cu radiation. The tunable nature of the source (which allows selective optimization of anomalous contributions to the scattering factors) and the low angular divergence of the beam make the source very useful for single crystal protein diffraction studies.

The use of synchrotron radiation as a source for single crystal x-ray diffraction studies has recently been the subject of considerable discussion and controversy. In contrast to contation of synchrotron radiation produced by the SPEAR electron-positron storage ring at the Stanford Linear Accelerator Center, Stanford, Calif.

Synchrotron radiation is emitted tangentially to the instantaneous path of the charged particles as they are main-

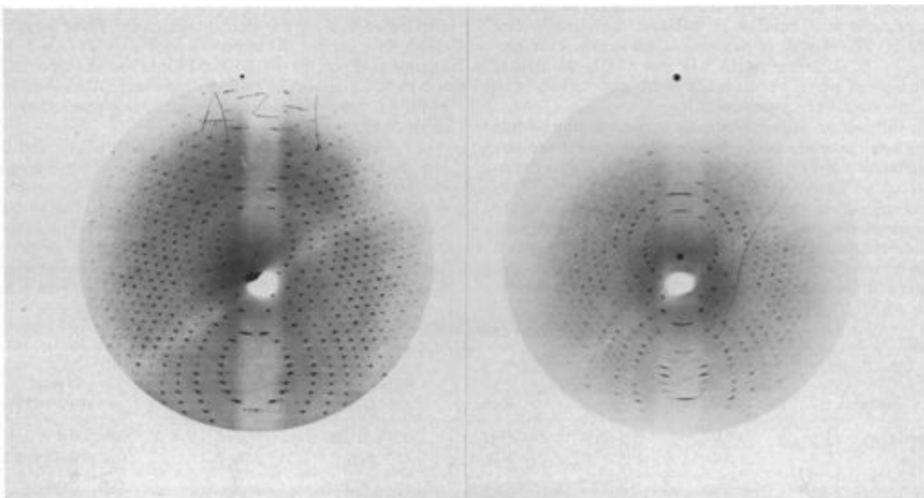


FIG. 3. Cone-axis oscillation photographs of the same azurin crystal. Precession angle 6.5°, oscillation angle 20°. (left) Synchrotron source, E = 3.7 GeV, I = 40 mA, only electrons present, λ = 1.540 Å, exposure time 10 min. (right) Philips fine-focus sealed Cu anode tube, operated at 40 kV, 30 mA, exposure time 6 hr, Ni filter.

Detectors

Technology	When	Readout	Remark
X-ray film	1940	30 min	grain size < 10 µm
Image Plates	1990	1.5 min	pixel: 150 x 150 μm^2 + PSF
CCD detectors	1997	1 sec	pixel: 80 x 80 μm^2 + PSF
Pixel-Array Detectors (2010)	2010	3 msec	pixel: 173 x 173 μm^2 sharp 6 MPixel
Pixel Array Detectors (2014)	2014	3 µsec	pixel: 75 x 75 µm 16 MPixel

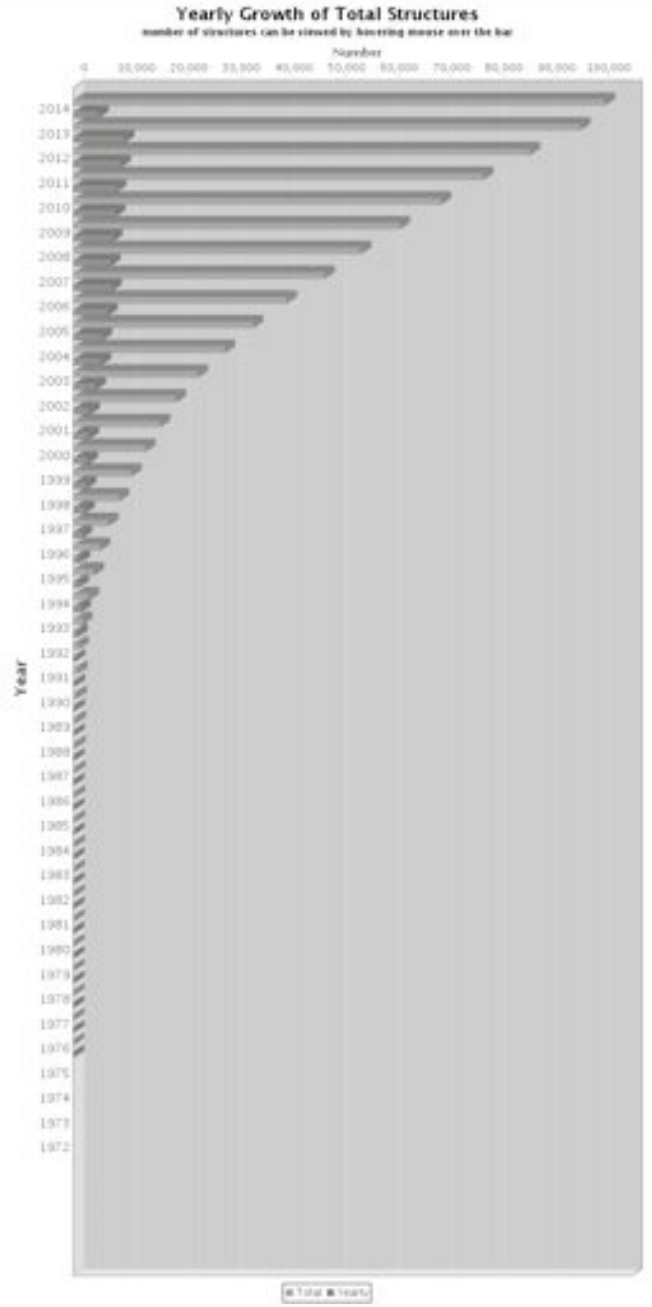
- PSF = Point spread function

The Protein Data Bank

- Started as a grass-root movement in the 70's
- As of Tuesday Jun 24, 2014 17:00 PDT, **101207 structures** are in the protein data bank.
- Out of these 89670 were determined by X-ray crystallography

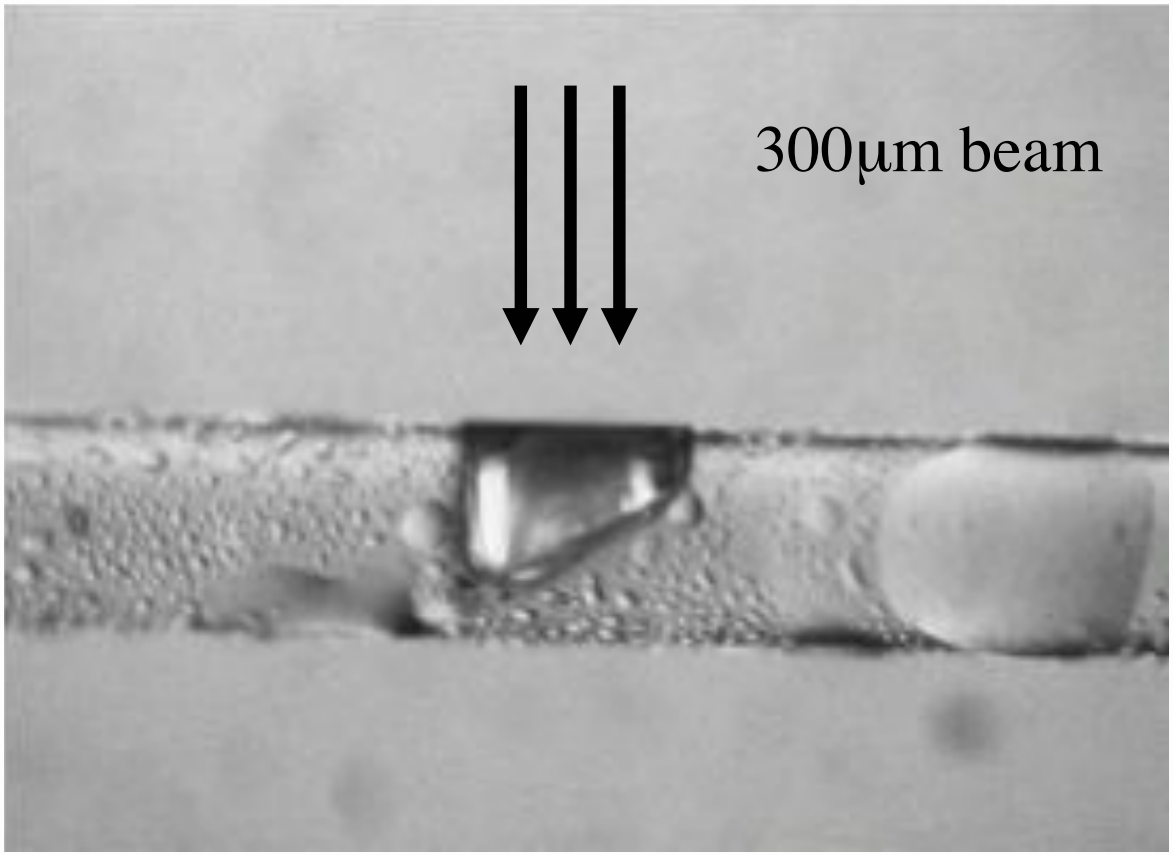
- 1976:13
- 1980:69
- 1990: 507
- 2000: 13596
- 2010: 70013
- 2014: 100000

- A large fraction of structures determined today are determined by 'Molecular Replacement'



http://www.rcsb.org/pdb/statistics/contentGrowthChart.do?content=total&seqid=100

Radiation Damage

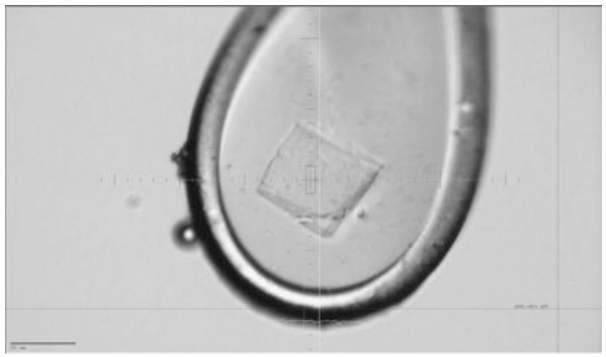


Garman & Schneider (1997) J. Appl. Cryst. 30:211

Data collection at 100 K

- Mounting protein-crystals in a free-standing film revolutionized the field.

Teng, T. Y. (1990)
J. Appl. Cryst 23: 387-391



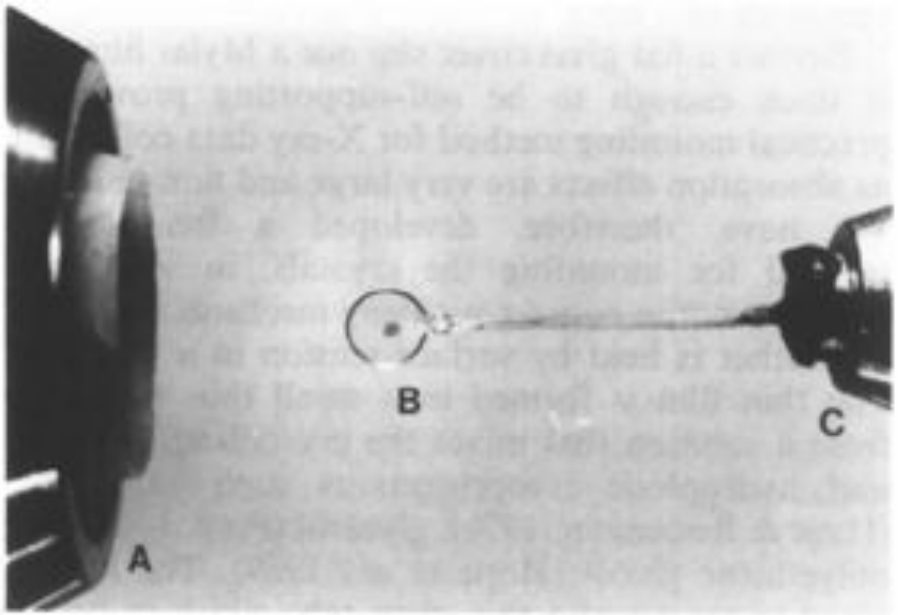
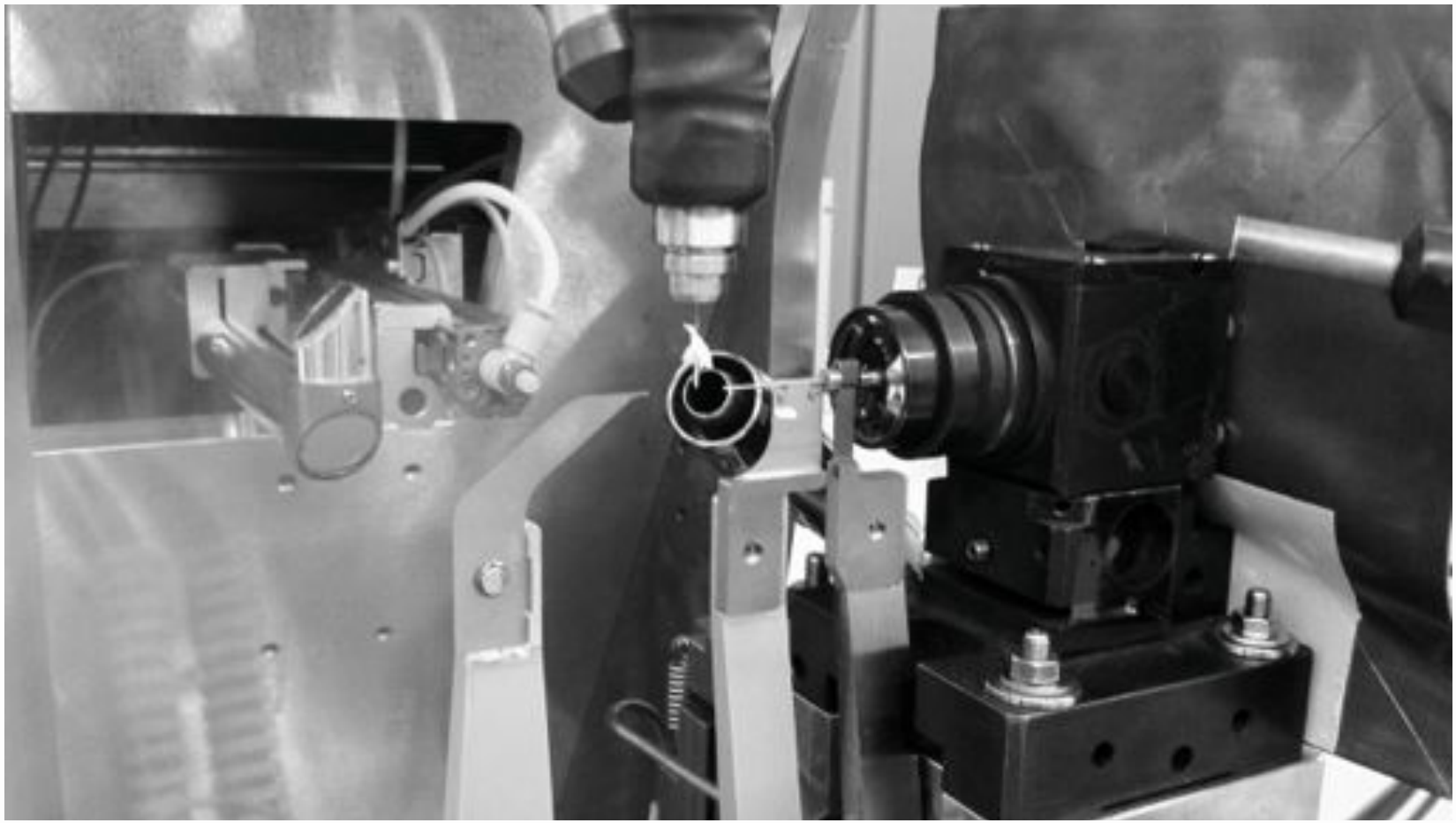


Fig. 2. The loop mounted on an oscillation camera used at CHESS: *A* cold nitrogen nozzle; *B* the loop with a frozen crystal; *C* cold-gas-stream reflector mounted on the goniometer head. The cold nitrogen nozzle (*A*) of a transfer line is 1/4 in in diameter which delivers a gas stream at from 80 to 230 K. Within a cone-shaped working volume of 65 mm^3, the temperature gradient is less than 4 K, at an initial exit temperature of 85 K. Fog or ice formation around the nozzle, crystal and goniometer head is avoided by a coaxial warm and dry nitrogen stream that surrounds the cold stream, and by a built-in heater on the base of the cold-gas-stream reflector (*C*).

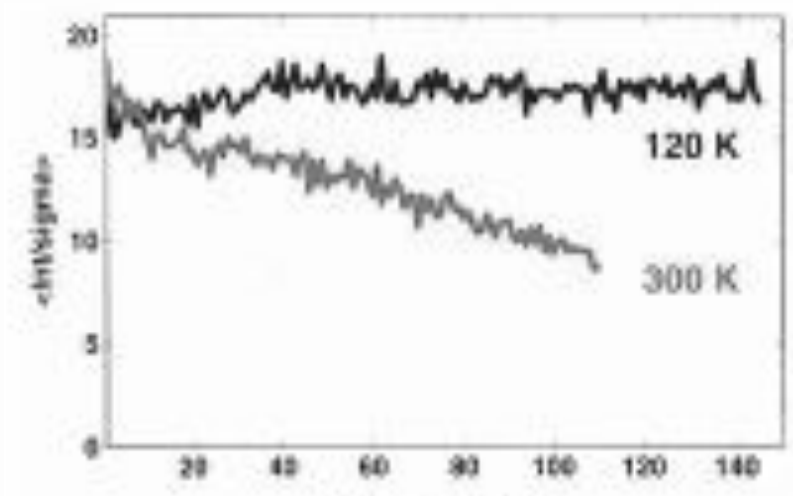
HPGonioV: Sample environment

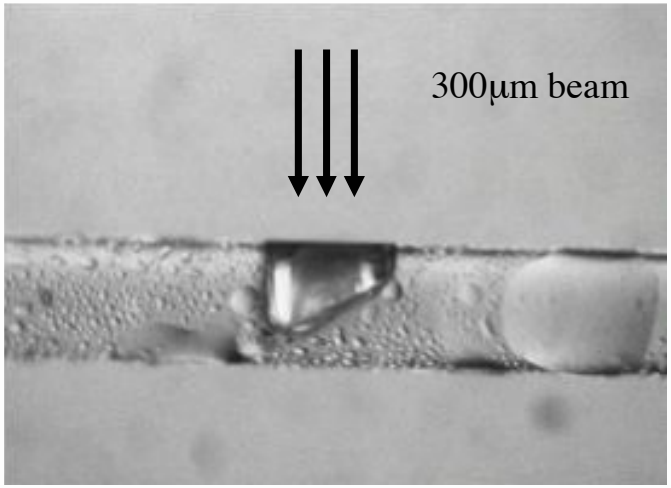


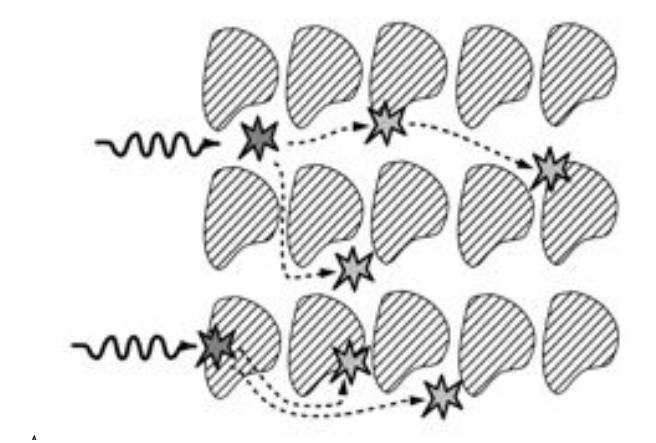
- 400 micron thick blade.
- Motorized adjustable distance to sample: 7-36 mm.

Radiation Damage

- Data collection at 100 K significantly extends crystal lifetime (factor of 50)

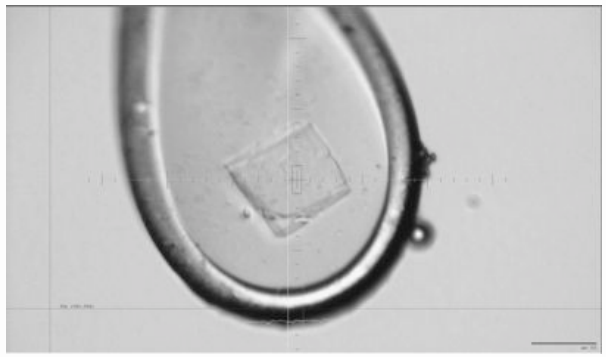


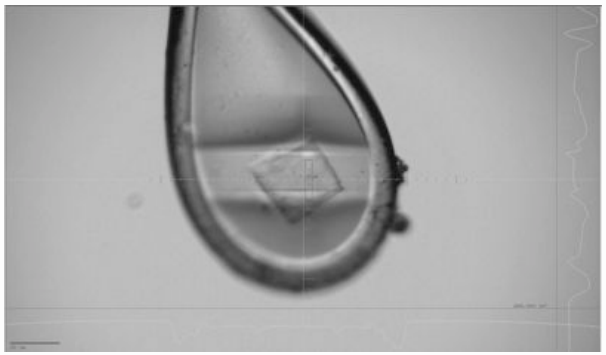




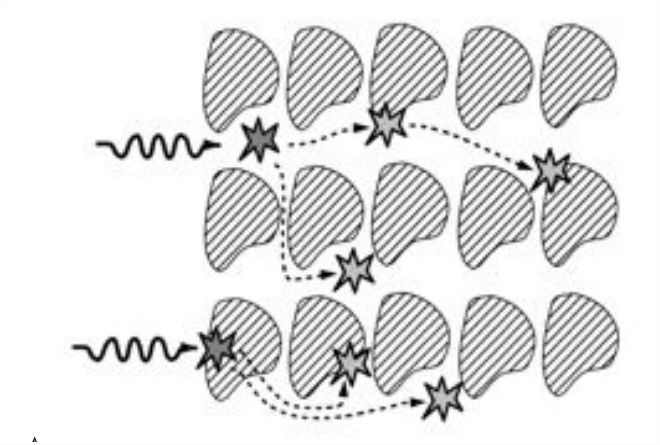
Garman & Schneider (1997) J. Appl. Cryst. 30:211

Radiation damage at 100 K



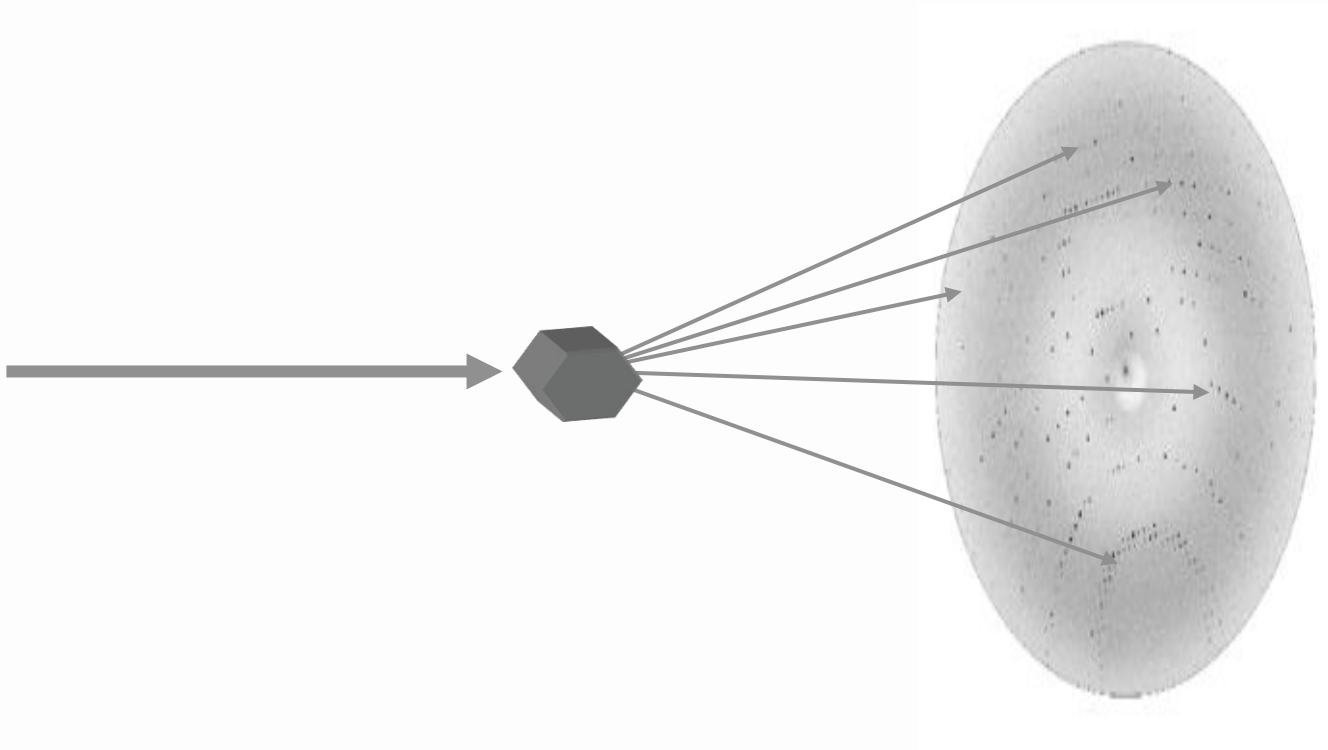


- On modern synchrotron beamlines the lifetime of protein crystals at 100k is on the order of seconds to minutes.
- Going back to Room temperature, Owen et al. (2012) Acta Cryst. D68:810 have shown that in the first 200 msec of a room temperature experiment, one can outrun hydroxyl radicals.

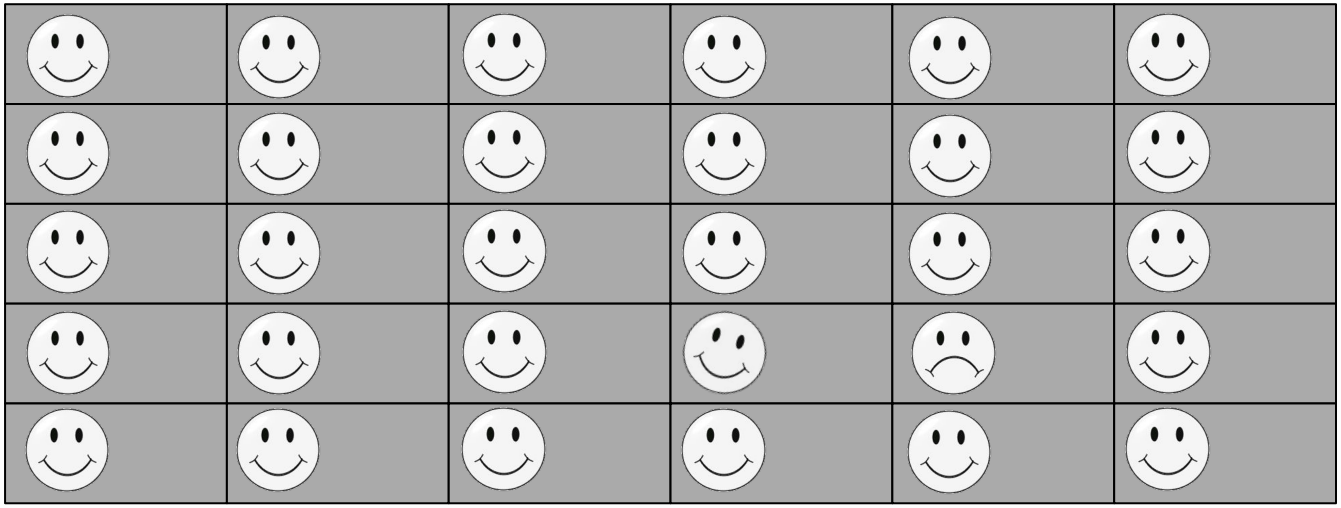


Crystal Structure Determination

Diffraction from a Crystal

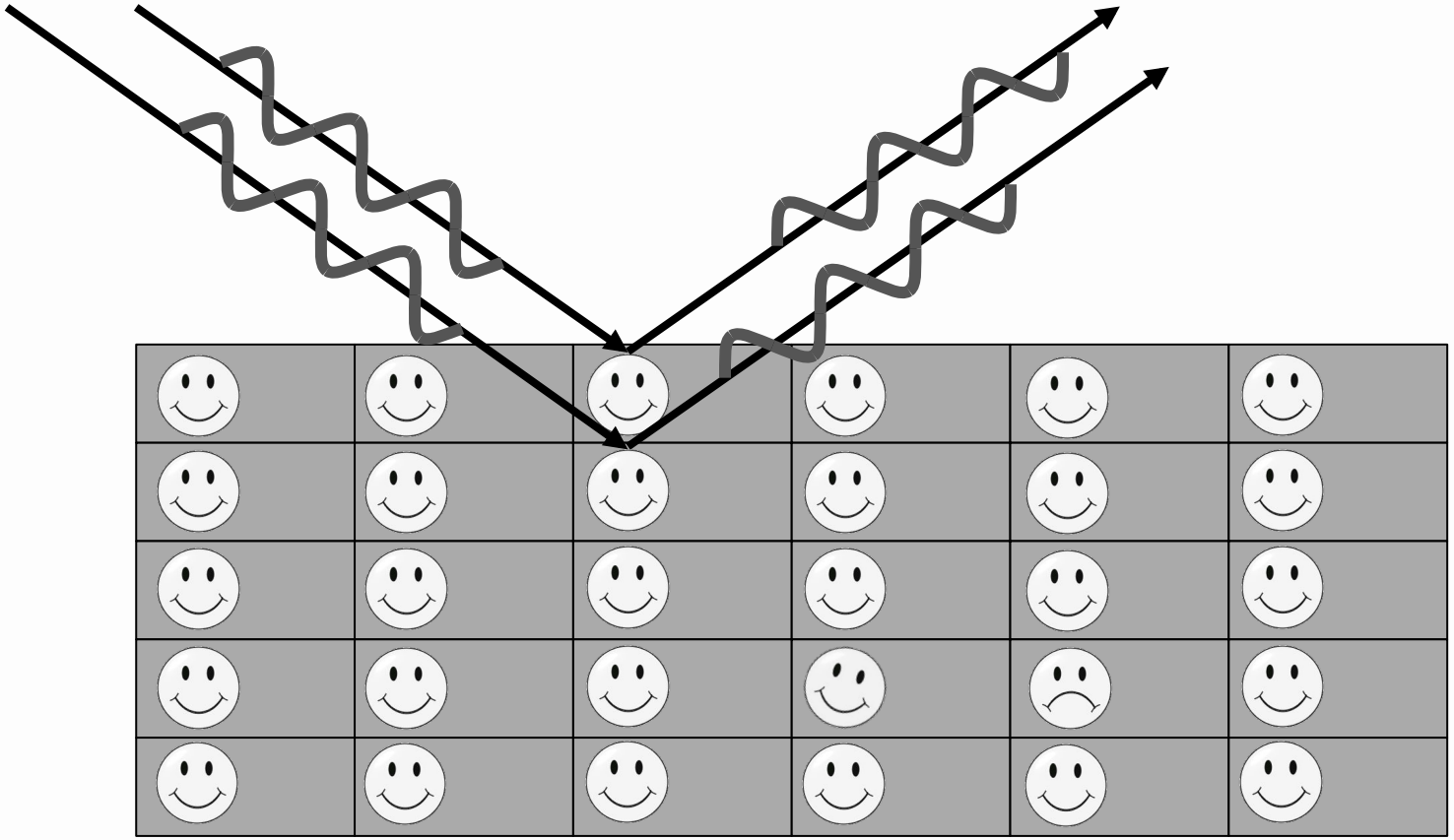


Inside a crystal



- A crystal consists of repeating units, the crystallographic unit cells.
- Each unit cells has 'the same' content, i.e. the same molecules in the same conformation and in the same orientation
- In real crystals, there is always some amount of 'disorder'

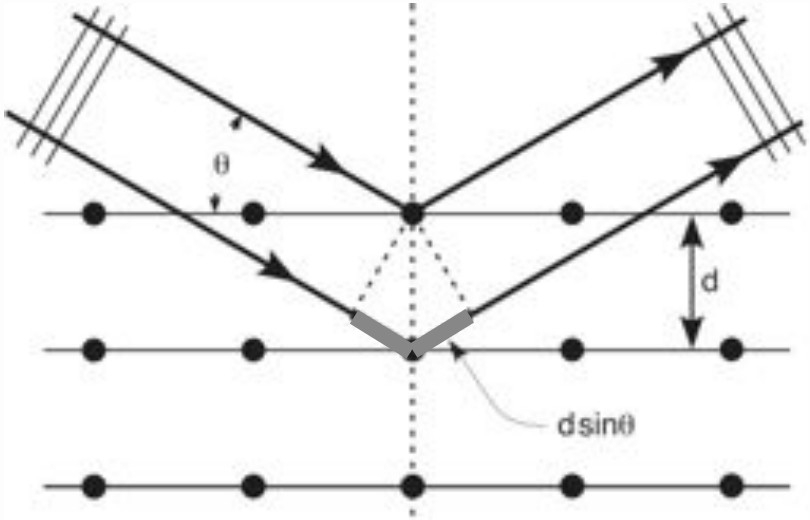
Diffraction from a crystal



- When electromagnetic waves are interacting with a periodic structure, interference effects will occur.

Diffracting planes

http://en.wikipedia.org/wiki/X-ray_crystallography

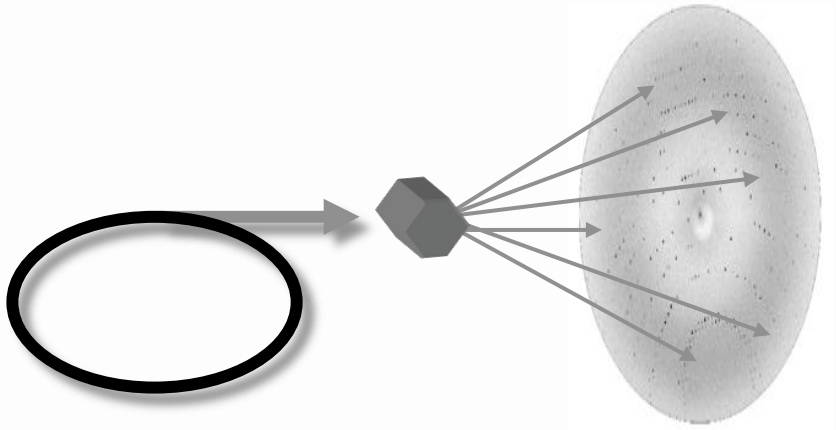


Constructive interference occurs when Bragg's law is fullfilled:

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

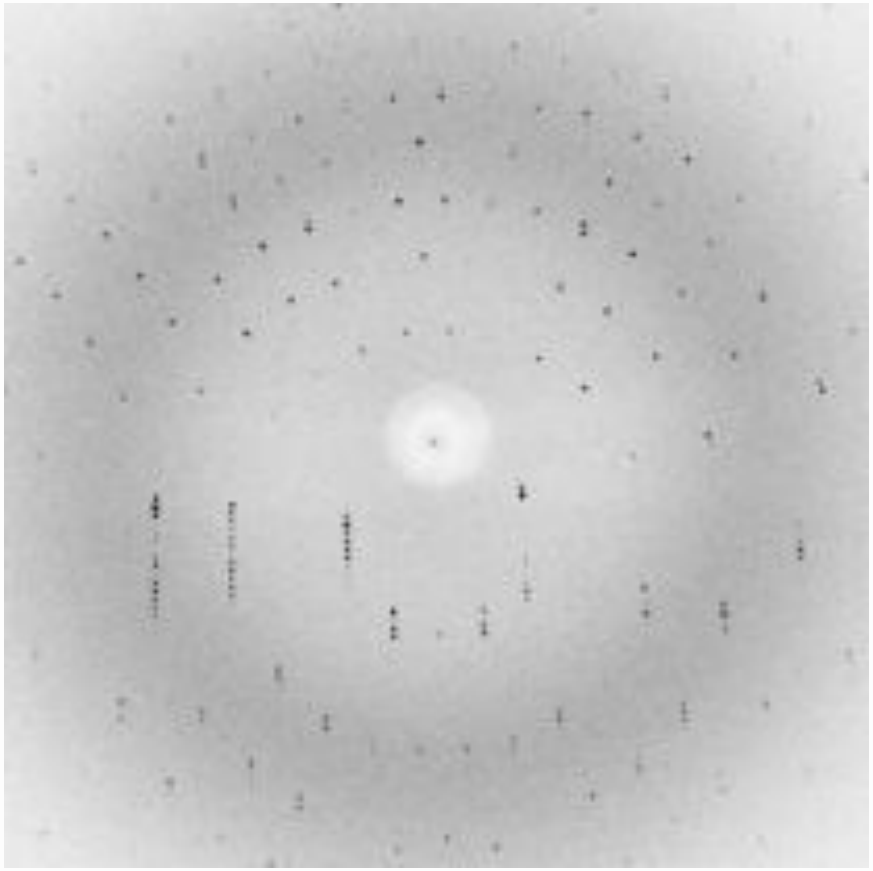
Nobel prize for physics 1914 to Max von Laue

Nobel prize for physics 1915 to William and Lawrence Bragg



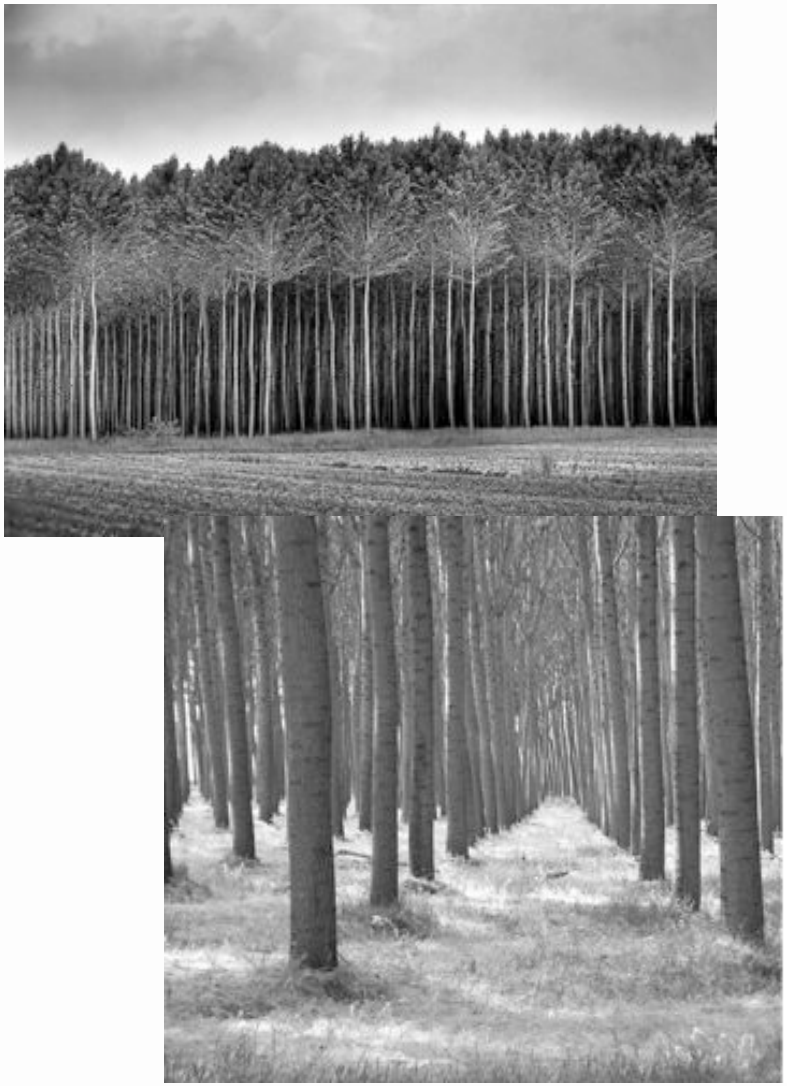
Diffraction Data

- The diffraction pattern changes when the crystal is rotated.
- By rotating the crystal, different Bragg-planes are brought into their diffracting position



http://www-structmed.cimr.cam.ac.uk/Course/Basic_diffraction/data_animation.html

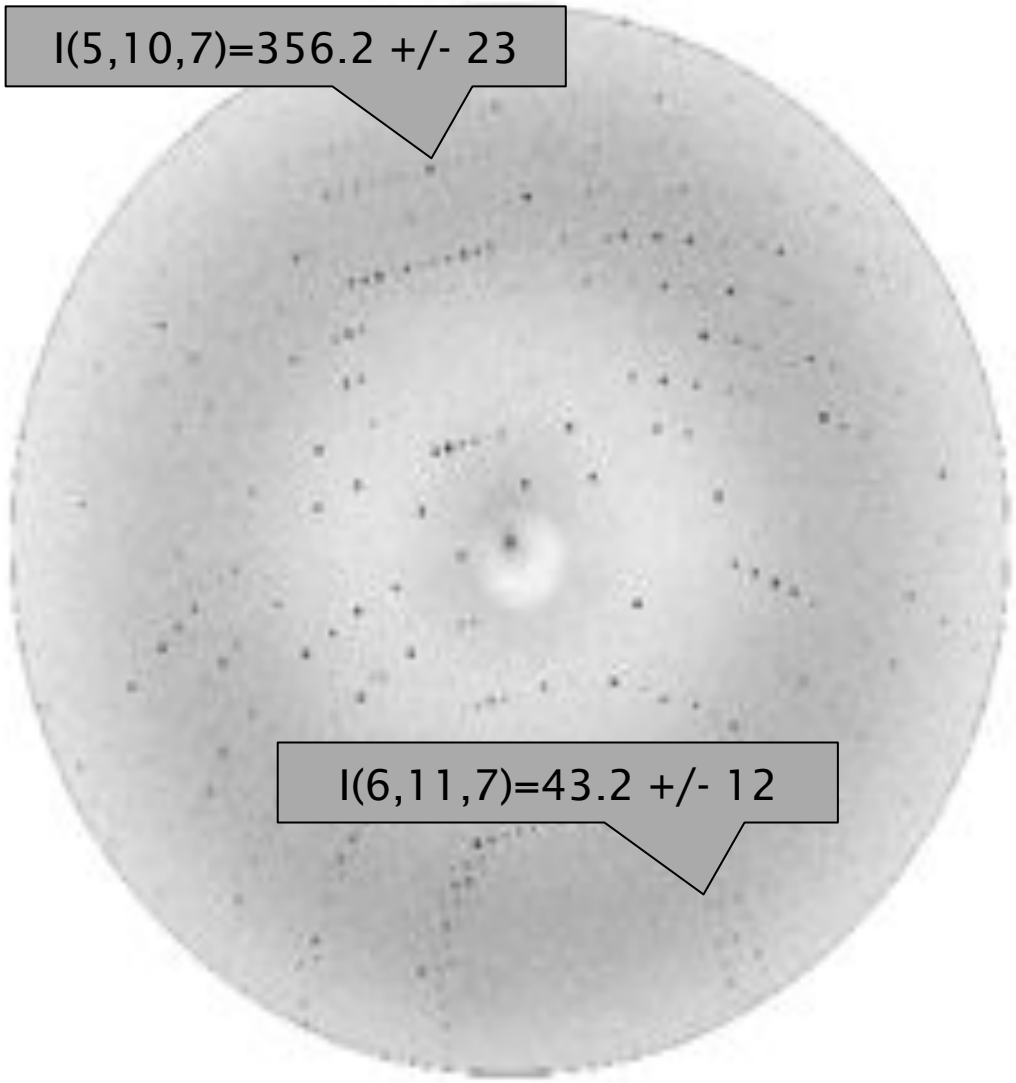
Diffraction and forests ...



http://flickr.com/photos/rossogiallobianco/2486114038

Indices and Structure factor amplitudes

- Every diffraction spot is marked by an index hkl
- For every diffraction spot an Intensity I is measured.
- The result of the experiment is an indexed set of I's
- Diffraction at low 2theta is stronger than at high 2theta



Diffraction data

h	k	l	I	sig(I)
0	20	35	4980.5	122.6
0	20	36	6906.6	216.6
0	20	37	8302.3	231.7
0	20	38	3209.5	89.3
0	20	39	459.6	22.1
0	20	40	1017.4	33.8
0	20	41	-5.6	18.3
0	20	42	33.8	15.6
0	20	43	4545.7	133.4
0	20	44	210.5	19.2
0	20	45	808.8	29.1

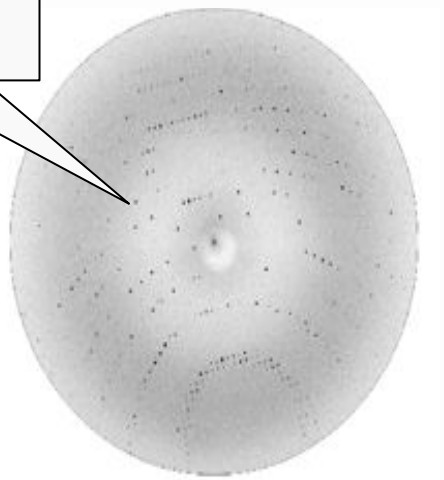
Structure Factor Amplitudes

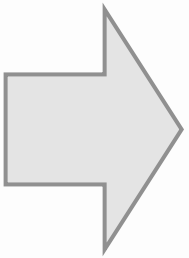
- For formal reasons, the measured I's are usually converted to 'Structure Factor Amplitudes' F by:

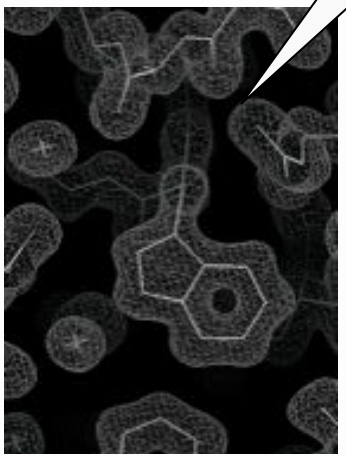
$$F = \mathrm{sqrt}(I).$$

Calculating Electron density

Intensity I_{hkl} for some set of planes with Miller indices hkl







Electron density for all points in the unit cell

$$\rho_{xyz} = \sum_{hkl} \left| F_{hkl} \right| e^{-i\varphi_{hkl}} e^{-2\pi i(hx+ky+lz)}$$

Electron density at some point *xyz* in space

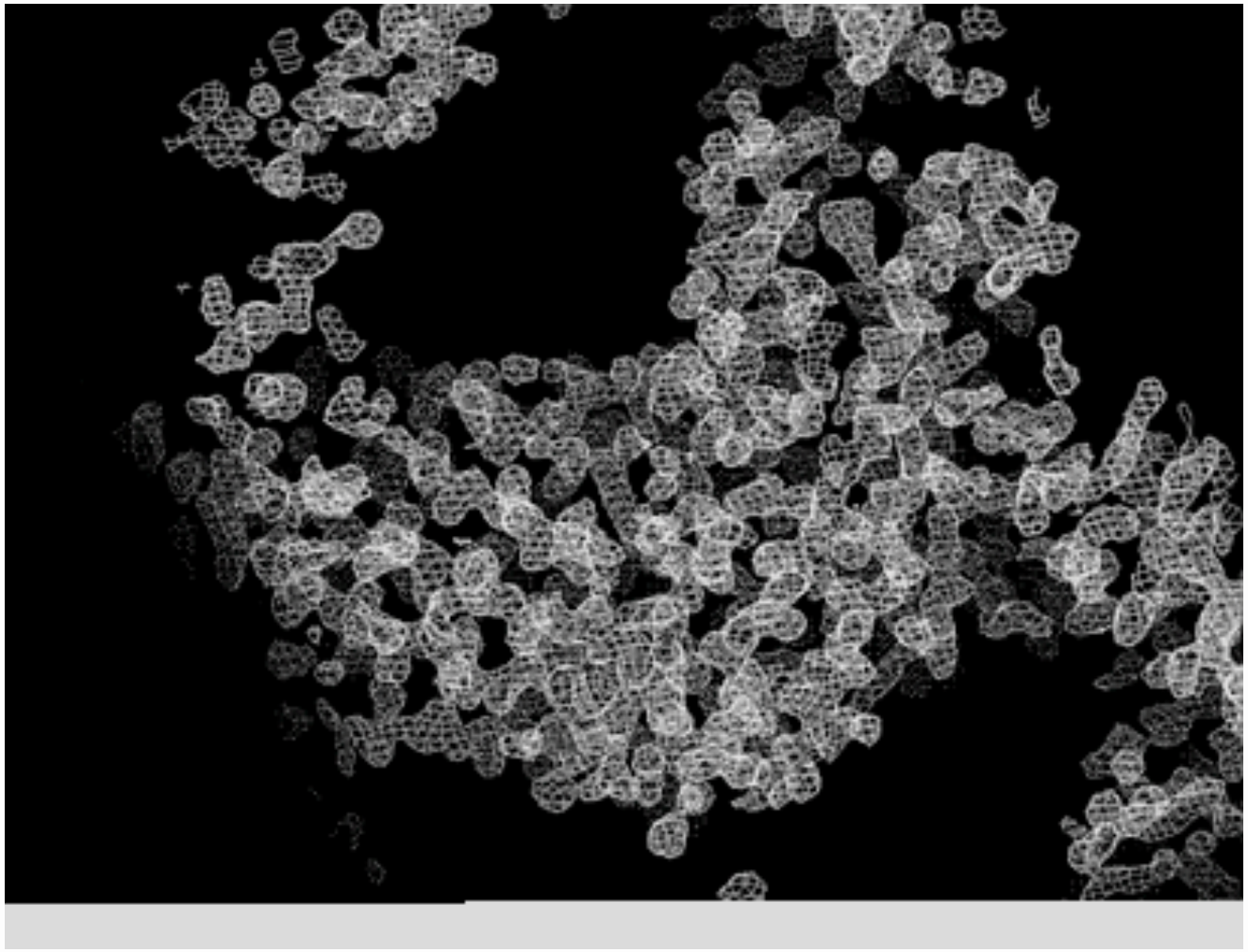
Sum over all Bragg reflections

Structure Factor Amplitude for each reflection hkl. $F_{hkl} \approx \sqrt{I_{hkl}}$

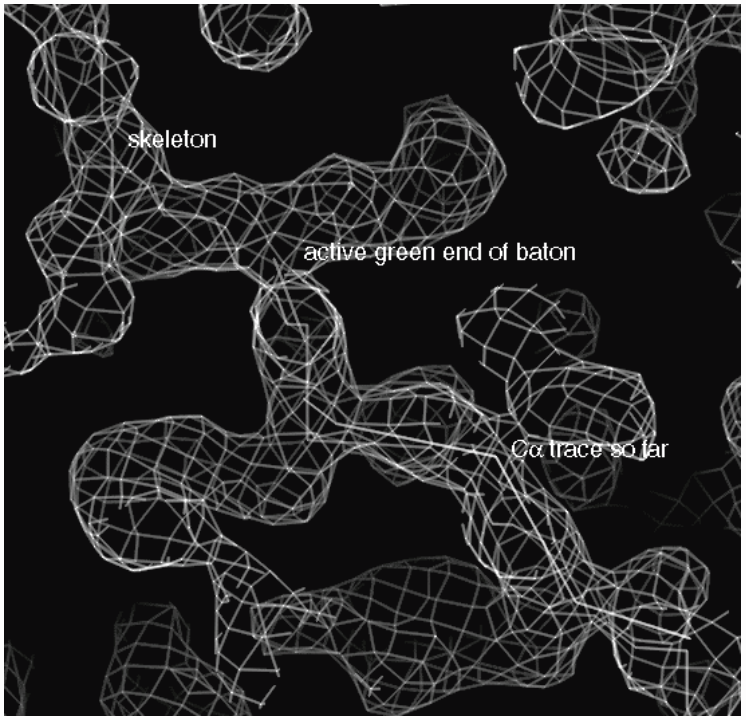
Structure Factor Phase for each reflection hkl.

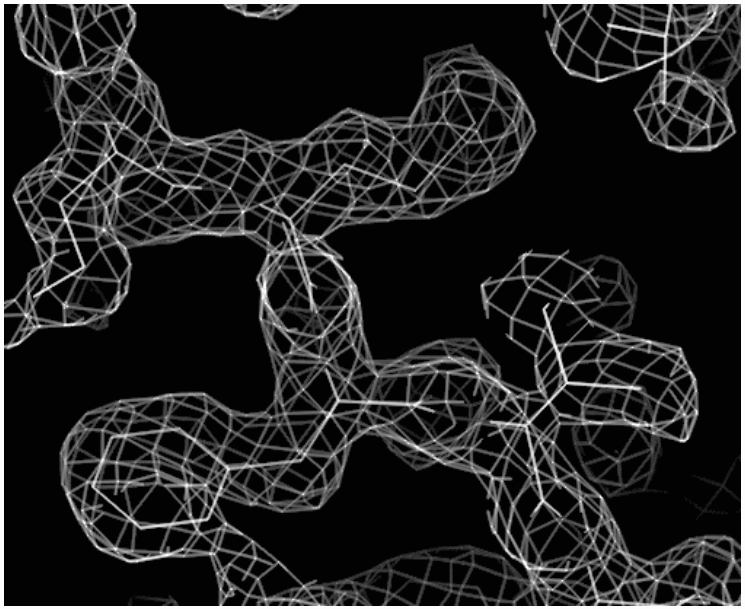
Phase shift depending on hkl and position in space

Typical initial electron density map



Interpretation of the electron density map





A 'pdb'-file (www.rcsb.org)

```
HEADER    TRANSFERASE                             06-MAR-03   1OPJ
TITLE     STRUCTURAL BASIS FOR THE AUTO-INHIBITION OF C-ABL TYROSINE
TITLE    2 KINASE
.                                    X       Y       Z        B
.
ATOM      1  N   ALA A 243      20.064  -2.529  43.315  1.00 51.64           N
ATOM      2  CA  ALA A 243      19.658  -1.370  42.459  1.00 51.36           C
ATOM      3  C   ALA A 243      20.832  -0.838  41.643  1.00 50.91           C
ATOM      4  O   ALA A 243      20.650   0.010  40.776  1.00 50.77           O
ATOM      5  CB  ALA A 243      19.092  -0.263  43.322  1.00 50.20           C
ATOM      6  N   MET A 244      22.030  -1.350  41.906  1.00 50.62           N
ATOM      7  CA  MET A 244      23.218  -0.888  41.195  1.00 50.73           C
ATOM      8  C   MET A 244      23.537  -1.709  39.950  1.00 50.43           C
ATOM      9  O   MET A 244      24.554  -1.488  39.298  1.00 50.33           O
ATOM     10  CB  MET A 244      24.420  -0.882  42.146  1.00 51.32           C
ATOM     11  CG  MET A 244      24.215   0.020  43.361  1.00 52.32           C
ATOM     12  SD  MET A 244      25.597   0.070  44.515  1.00 54.49           S
ATOM     13  CE  MET A 244      26.730   1.130  43.627  1.00 53.37           C
ATOM     14  N   ASP A 245      22.662  -2.653  39.620  1.00 49.87           N
ATOM     15  CA  ASP A 245      22.857  -3.505  38.446  1.00 49.28           C
ATOM     16  C   ASP A 245      22.115  -2.895  37.257  1.00 47.55           C
```