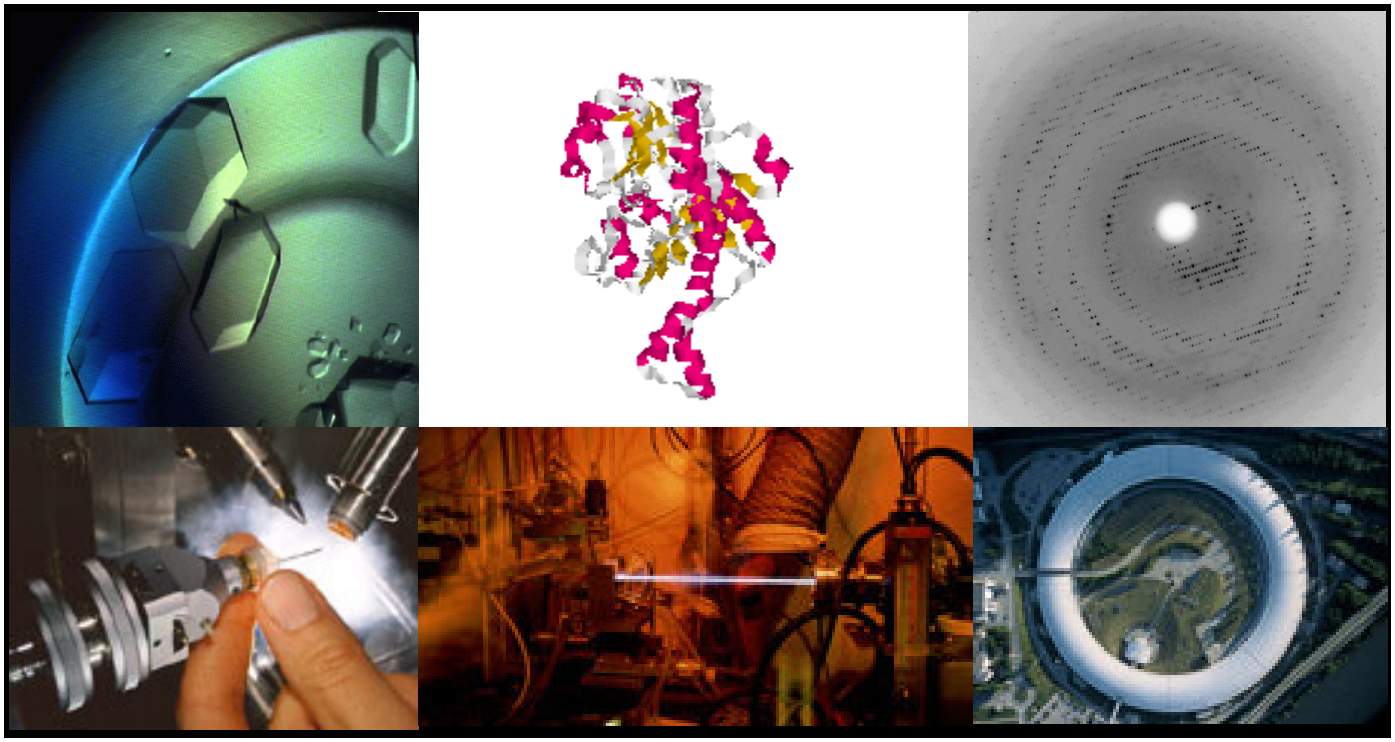
XIV School on Synchrotron Radiation:
Fundamentals, Methods and Applications
Muggia, Italy / 18-29 September 2017
SILS Società Italiana di Luce di Sincrotrone
Elettra Sincrotrone Trieste

Synchrotron Radiation and Biocrystallography



Marco Nardini
Dept. of Biosciences
University of Milano
marco.nardini@unimi.it

Dipartimento di Bioscienze

Biocrystallography

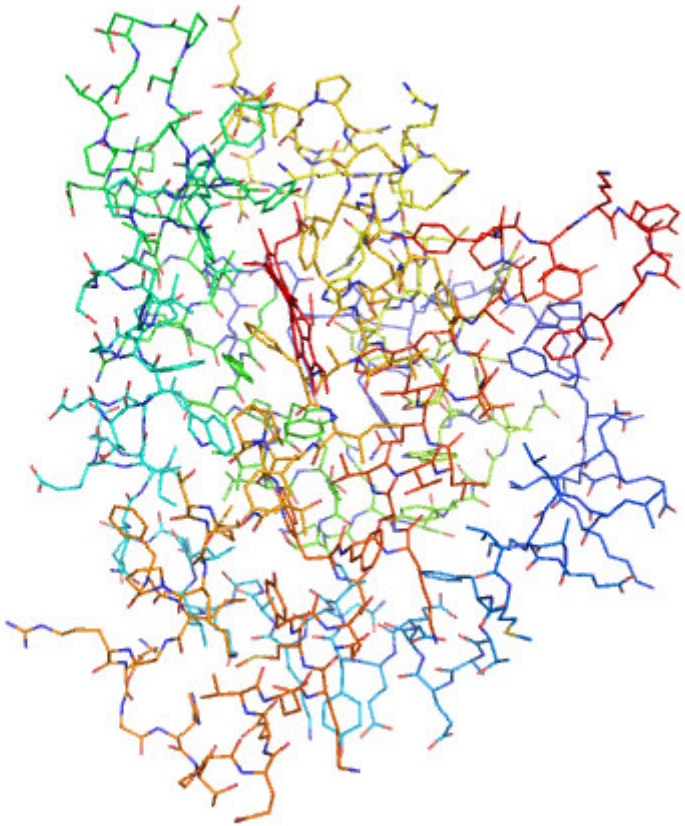
X-ray crystallography is the science of determining the arrangement of atoms within a crystal from the manner in which X-rays are deflected by the crystal

Aim:

3D structure determination of biological macromolecules at atomic resolution (x, y, z positions for each atom of the macromolecule), but ...

Strictly speaking, X-ray crystallography measures only the **density of electrons** within the crystal, from which the atomic positions can be inferred

Protein 3D structure



Biocrystallography

Object: **real system** (not a model system)

proteins, DNA, RNA, and their complexes (virus, ribosome,…)

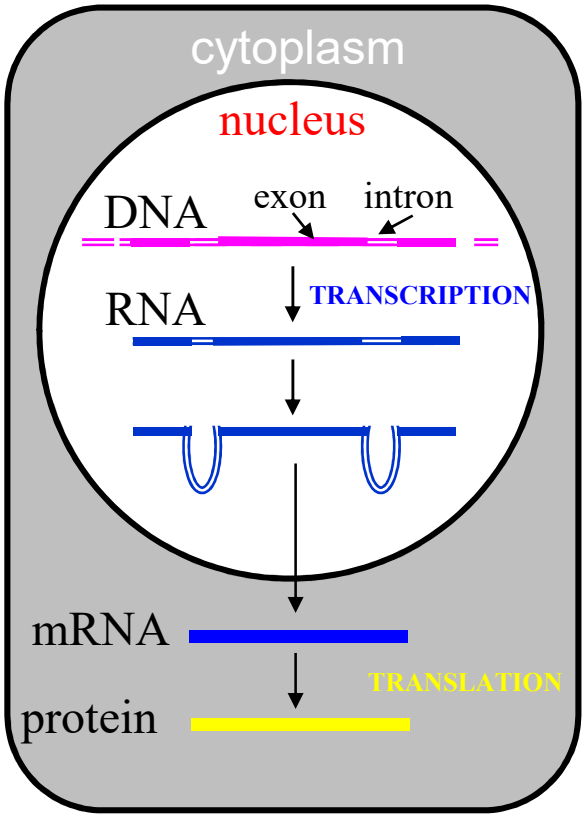
- the specimen should not be damaged during the experiment (the **sample** is **X-ray sensitive**)

- the smallest proteins have well **over 1000 atoms** and the largest proteins may have **between 10000 and 100000 atoms**

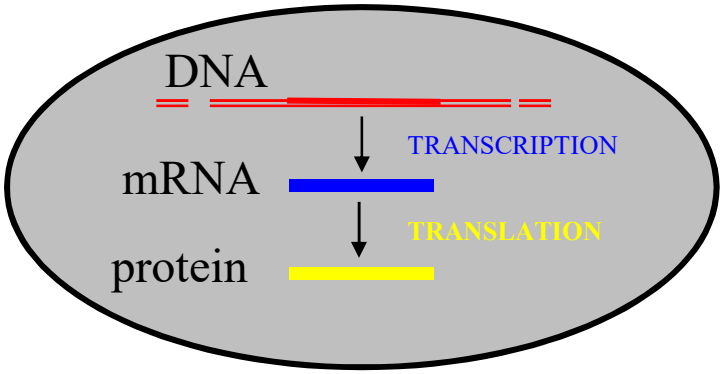
Ribosome: 19198 protein atoms di proteina, 32470 RNA atoms
(>50000 atoms)

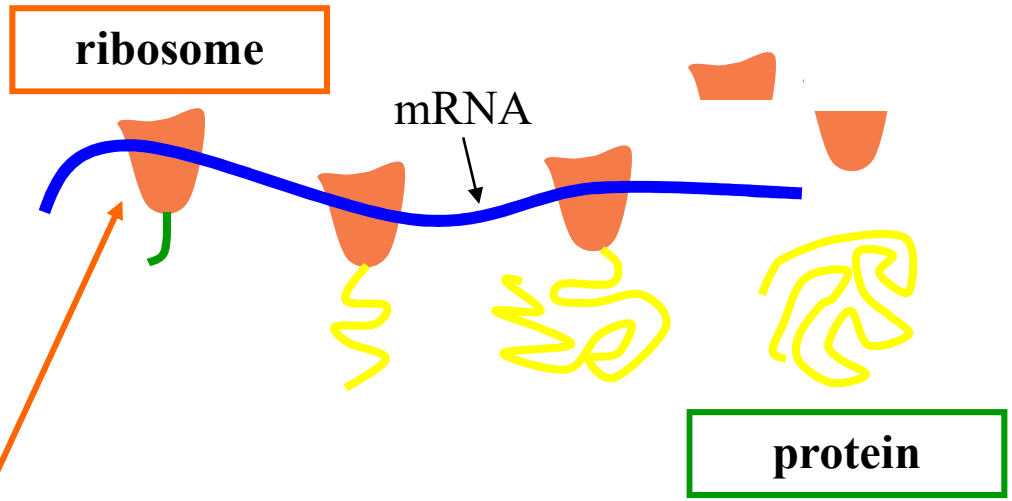
DNA, RNA & Proteins

Eukaryotes



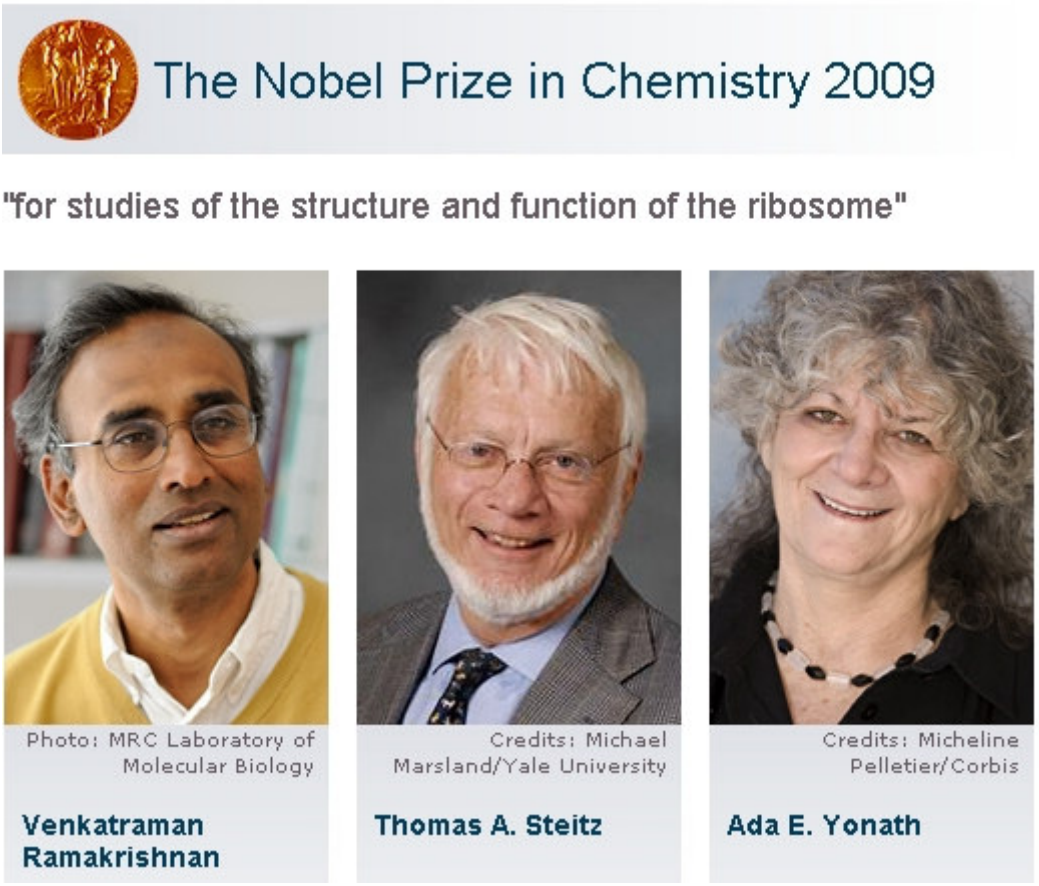
Prokaryotes





Ribosome: 19198 protein atoms, 32470 RNA atoms (**>50000 atoms**)

The structure of the Ribosome



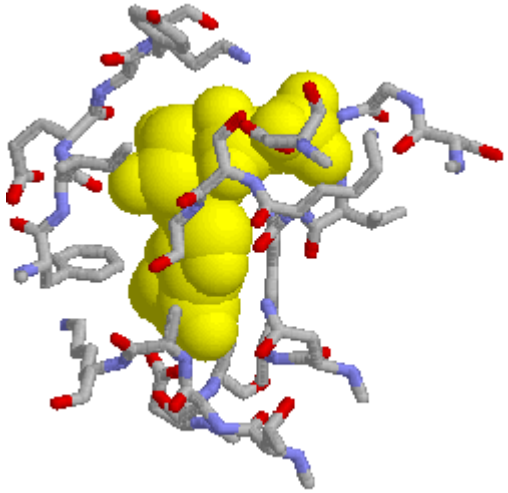
Examples of protein targets:

- **Catalytic reactions:** enzymes that catalyze (i.e. accelerate) biochemical reactions, and are vital for metabolism
- **Structural or mechanical functions**: actin and myosin in muscle, collagen in skin and bones, keratin in hair, fibrin in silk
- **Transcription factors:** protein/DNA interaction
- **Immune responses:** antibodies used to identify and neutralize foreign objects, such as bacteria and viruses. Search for antigens.
- **Cell signaling:** receptors
- **Molecular transport**: carriers for small molecules and/or ions (hemoglobin)
- **Membrane channelling:** membrane proteins control the flow of small molecule (i.e. ions) through cell membranes and organelles

Application of Biocrystallography

3D structures of macromolecules allow us to understand **biological processes and interactions** at atomic resolution (i.e. how a particular macromolecule accomplishes its various functions)

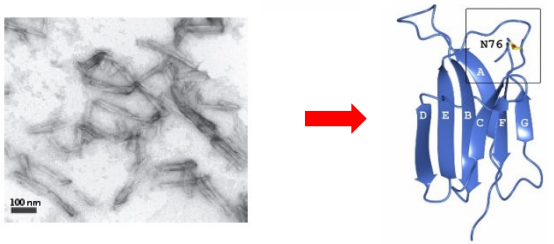
- macromolecule to macromolecule interactions
- macromolecule to small molecules (substrates, cofactors, inhibitors, ions …) interactions
- structural-functional studies on enzymes
- **rational drug design** (how drug lead compounds interact with their protein targets)
- biotechnological applications



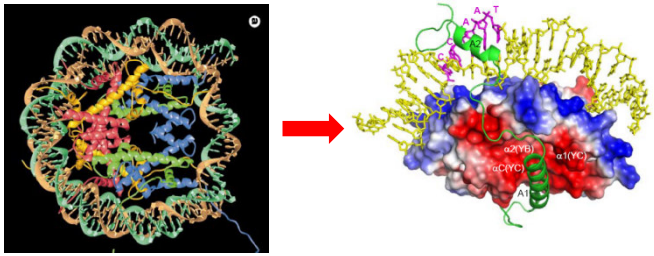
Application of Biocrystallography @ UNIMI

Università degli Studi di Milano
DBS Dipartimento di Bioscienze

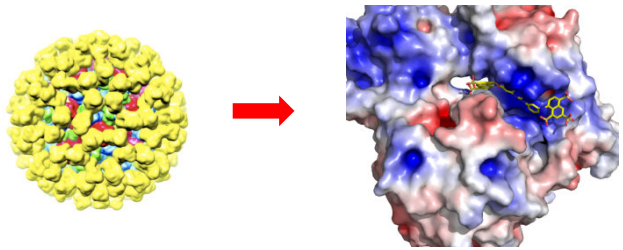
- **Molecular bases of protein mis-folding disease**: engineering and biophysics of pathogenic proteins involved in degenerative diseases



- **Transcription factors in chromatin regulation and dynamics:** structural principles and protein/DNA complexes



- **Drug design/discovery (antivirals):** targeting flavivirus replication machinery



- **Structural vaccinology:** designing epitopes based on 3D structure of antigens from bacterial pathogens

Application of Biocrystallography

Good news !!!!

- **No limitation in Mw** of the sample

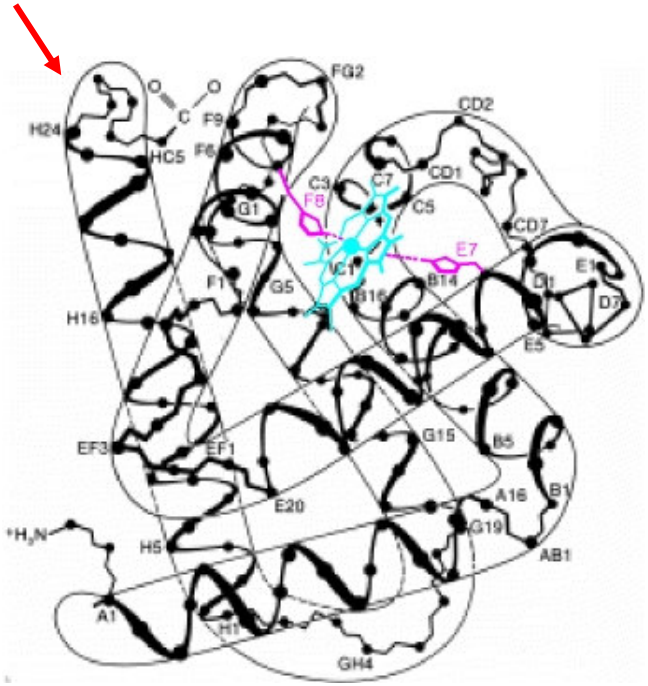
  (instead NMR < 30 kDa, cryoEM > 200 kDa)

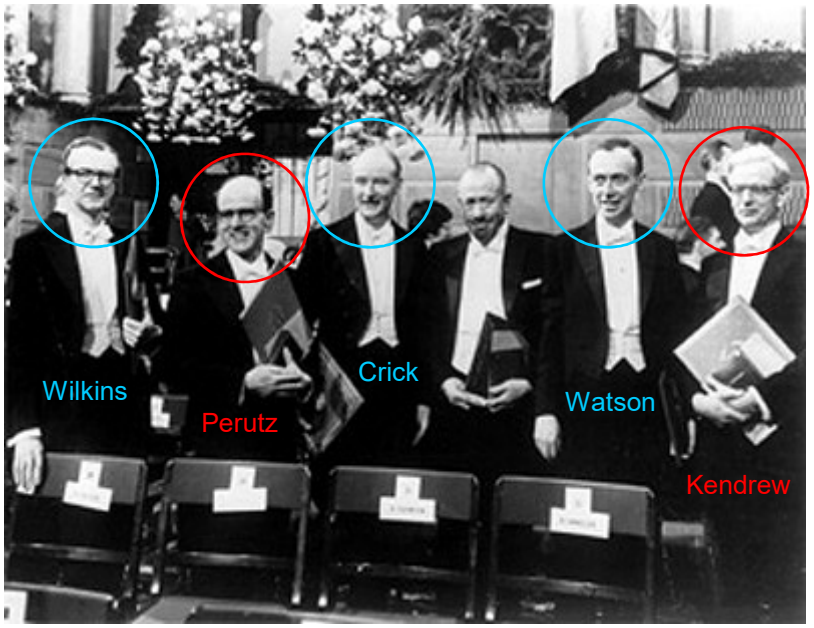
Bad news ¡¡¡¡

- good diffracting **crystals are needed**

Past, present and future

The first solved protein crystal structure was Sperm Whale **myoglobin** (1958)



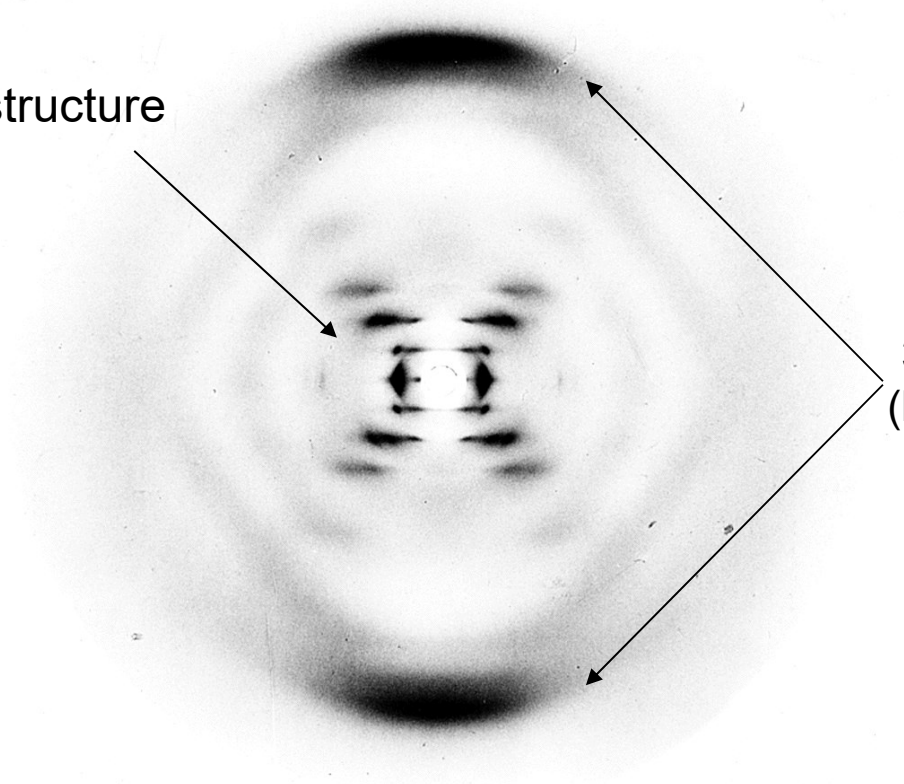


Max Perutz and Sir **John Cowdery Kendrew**, awarded the Nobel Prize in Chemistry in 1962 for their structural studies on globular proteins (hemoglobin and myoglobin, respectively)

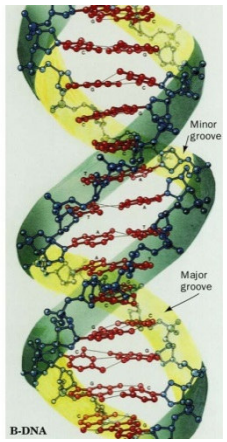
Past, present and future

ROSALIND ELSIE
FRANKLIN

helical nature of the DNA structure



X-ray diffraction image (n°51) from DNA fibers (1953)



Past, present and future

Biocrystallography is, to date, the most prolific discipline within the area of structural biology

PDB Current Holdings Breakdown

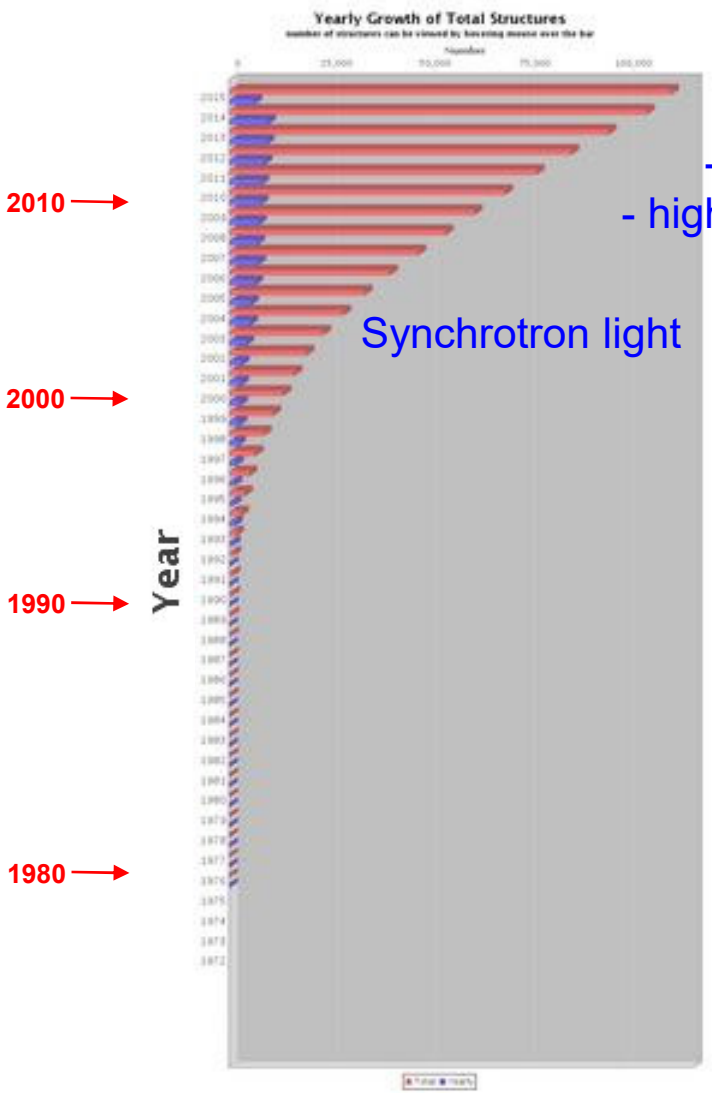
Exp.Method	Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
X-RAY	111813	1878	5711	4	119406
NMR	10488	1223	245	8	11964
ELECTRON MICROSCOPY	1235	30	432	0	1697
HYBRID	102	3	2	1	108
other	199	4	6	13	222
Total	123837	3138	6396	26	133397

(Click on any number to retrieve the results from that category.)
109195 structures in the PDB have a structure factor file.
9304 structures in the PDB have an NMR restraint file.
3056 structures in the PDB have a chemical shifts file.
1700 structures in the PDB have a 3DEM map file.

PDB PROTEIN DATA BANK

out of the **133397** 3D structures solved, X-ray crystallography is responsible for **119406** (**89.5%**)

Past, present and future

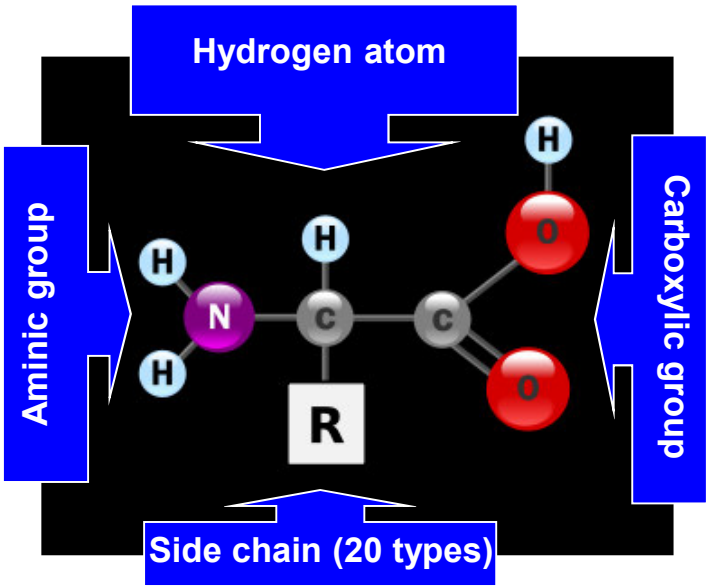


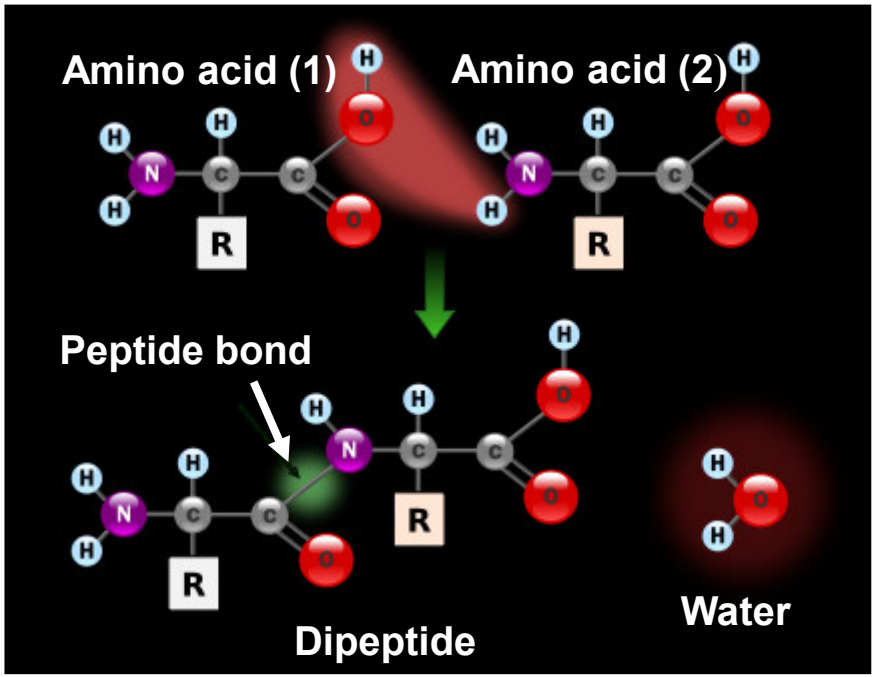
Structural genomics
- genome wide scale
- high throughput approaches

Synchrotron light

Protein structure

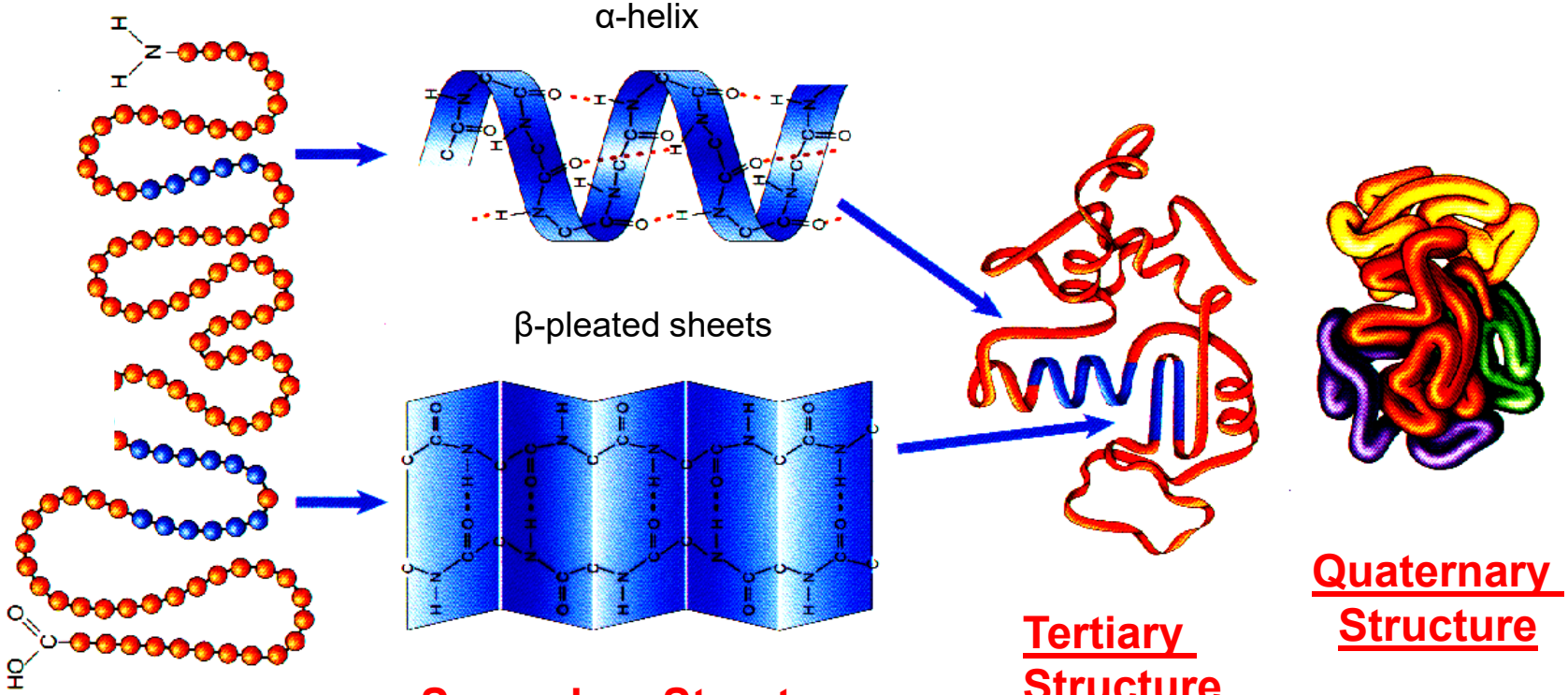
A protein forms via the **condensation of amino acids** to form a chain of amino acid "residues" linked by peptide bonds.





1	Alanine	Ala	A	$R = CH_3$
2	Aspartate	Asp	D	$R = CH_2\text{-}COO^-$
3	Cysteine	Cys	C	$R = CH_2\text{-}SH$
4	Glycine	Gly	G	R = H
.	.	.	.	.
.	.	.	.	.
20	Valine	Val	V	$R = CH\text{-}CH_3\text{-}CH_3$

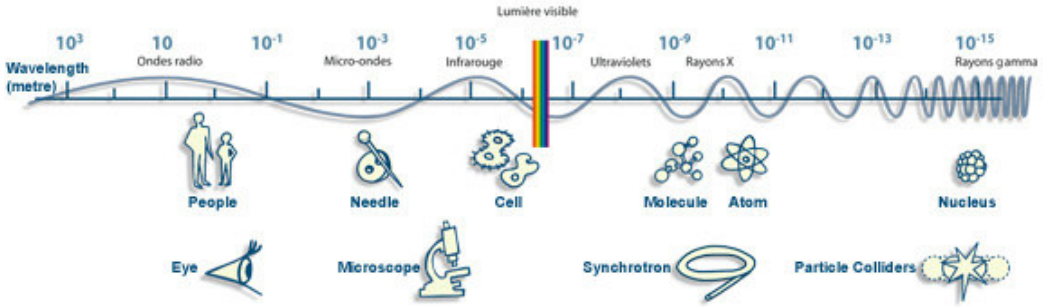
Protein structure

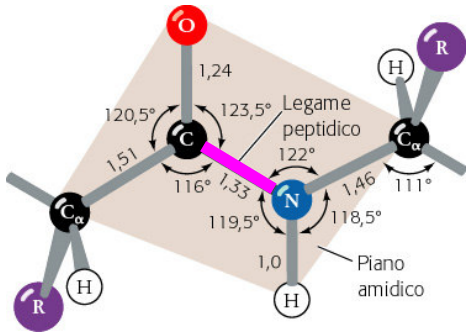


MPRPLVALLDGRD · · · · · · ETVEMPILKDVR

Why X-rays ?

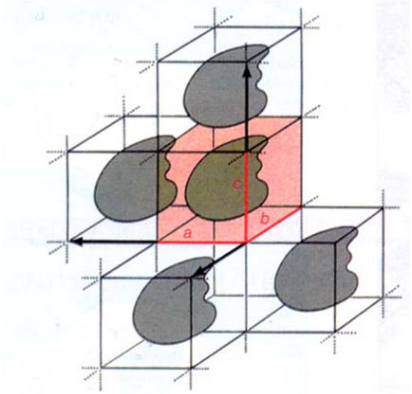
The wavelength of a X-rays is roughly 1 Å, which is on the scale of a single atom, and it allows to have **sufficient resolution to determine the atomic positions**



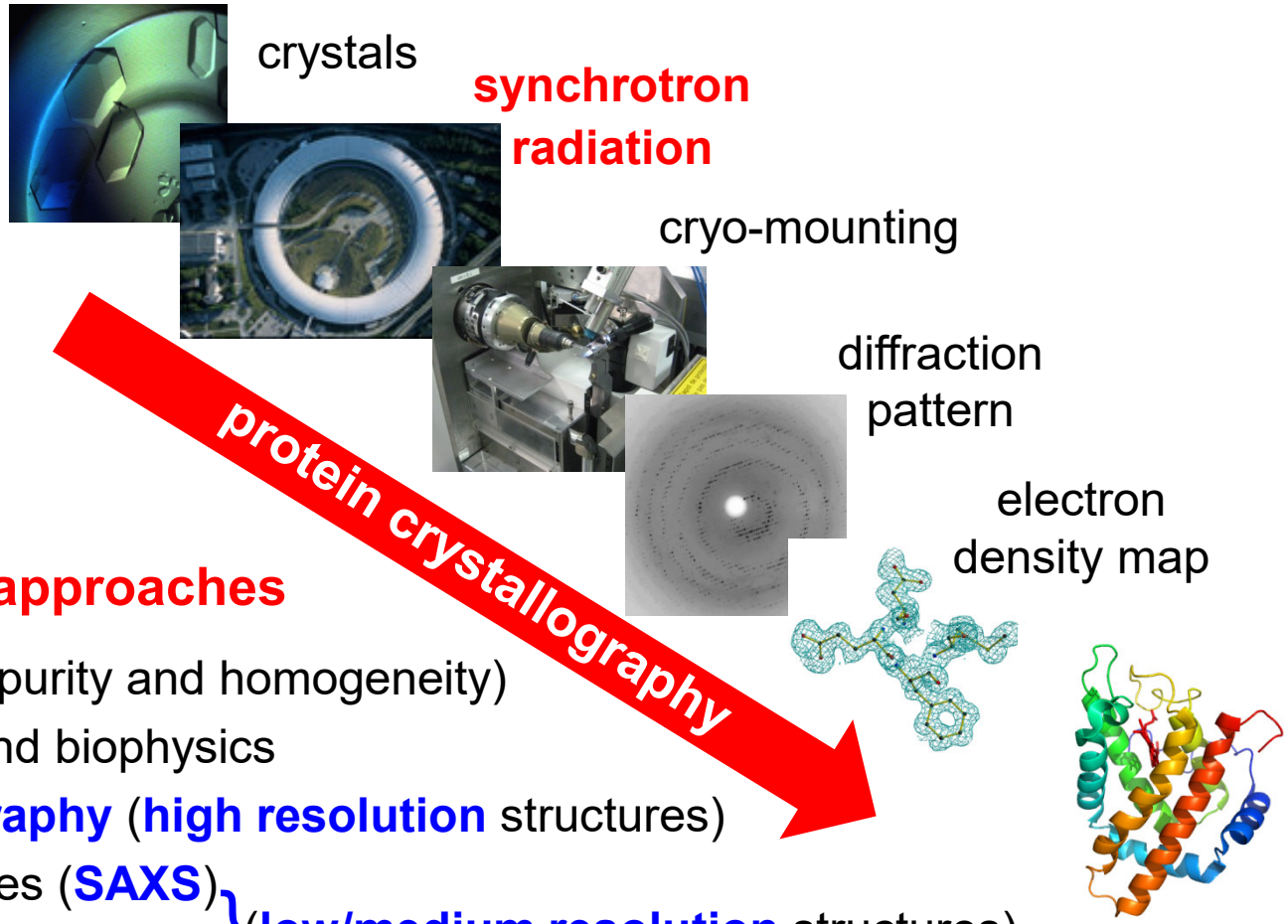


Why crystals ?

X-ray crystallography requires a crystal to amplify the signal (**10^{15}-10^{16} identical molecules**); the periodicity of the electron density is used to diffract the X-rays with manageable measurement error



Crystallographic experiment



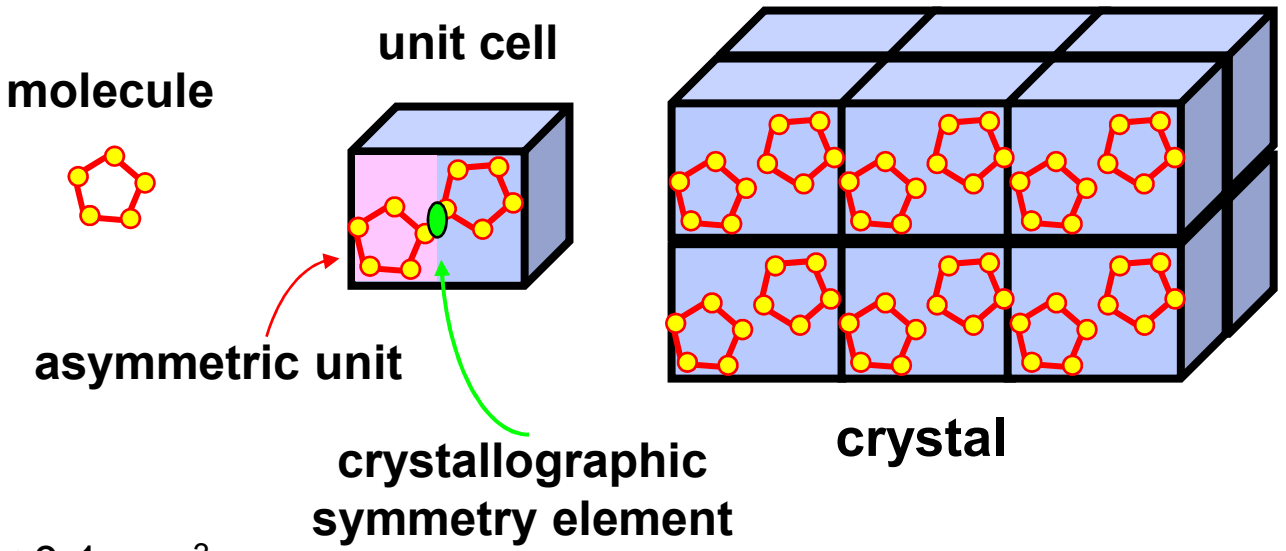
The experimental approaches

- Protein production (purity and homogeneity)
- Protein chemistry and biophysics
- **Protein crystallography** (**high resolution** structures)
- Solution X-ray studies (**SAXS**)
- Single particle **Cryo-EM**

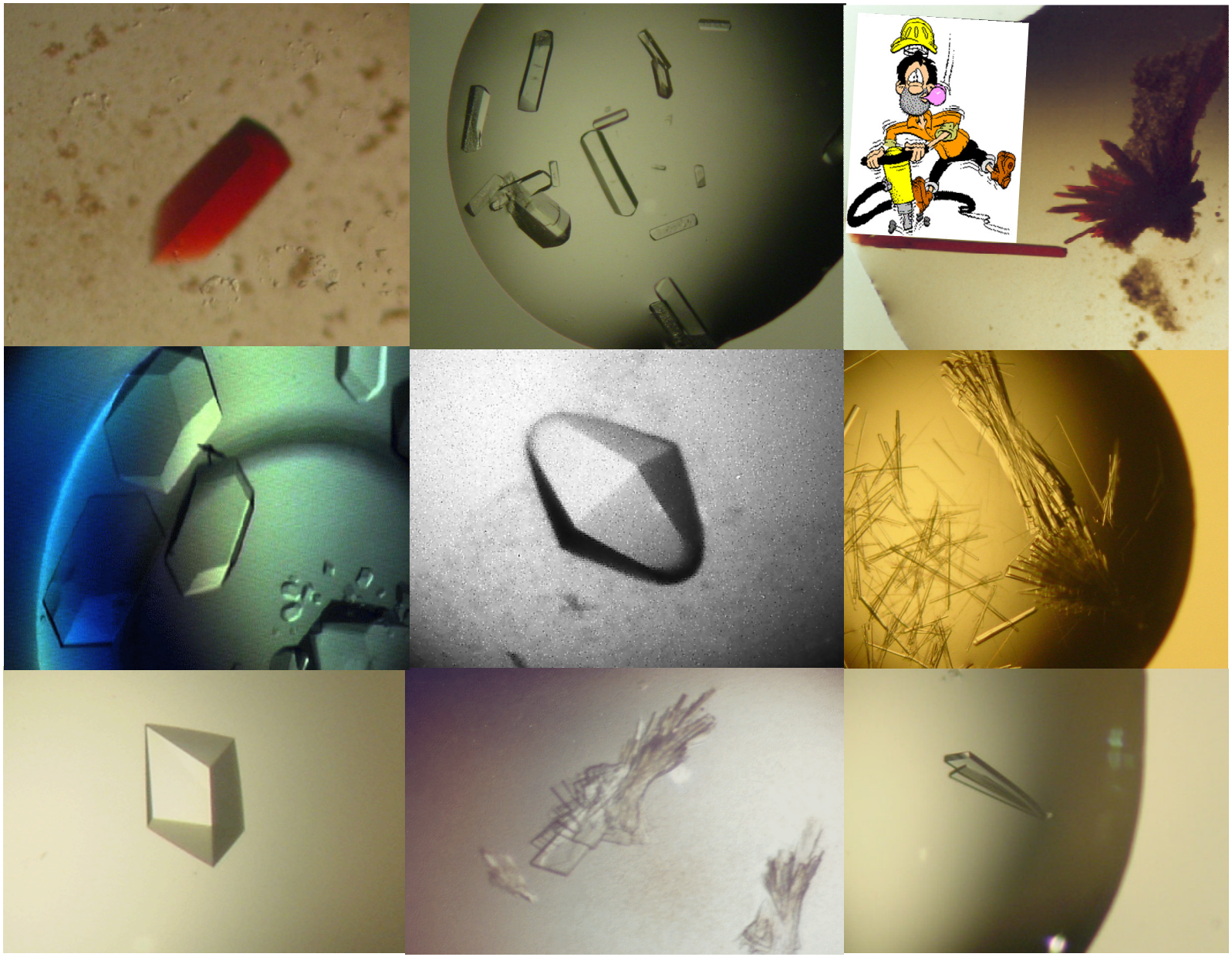
} (**low/medium resolution** structures)

- Computational biology simulations (*in silico* modelling)

Protein Crystal



- volume ≤ 0.1 mm^3
- crystal lattice periodicity > 100 Å
- solvent content 30% - 80% v/v
- mechanic fragility ($E_{stab.}$ < 10 kcal/mol)
- non-covalent interactions (surface a.a. residues)



FIRST BIG PROBLEM !!

How to crystallize a protein

(the “bottleneck” of the procedure…)

Crystallization

- protein crystallization is mainly **a trial-and-error procedure** in which the **protein is slowly precipitated** from its solution (to avoid formation of useless dust or amorphous gel)

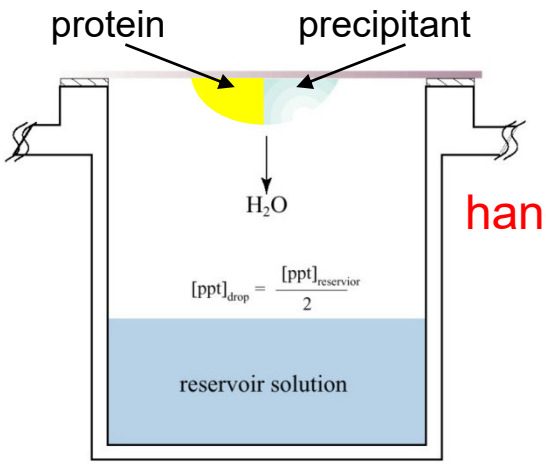
- crystal growth in solution is a multiparameter process involving three basic steps: **nucleation** (possibly having only 100 molecules), **growth**, and **cessation of growth**

- it is extremely difficult to predict good conditions for nucleation or growth of well-ordered crystals. **In practice**, favorable conditions are identified by **screening** (hundreds, even thousands, of solution conditions are generally tried)

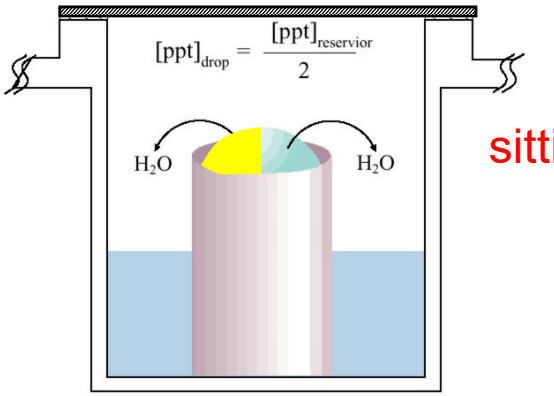
- large amounts (**milligrams**) of highly pure protein are required (due to high concentration of the molecule(s) to be crystallized)

Crystallization

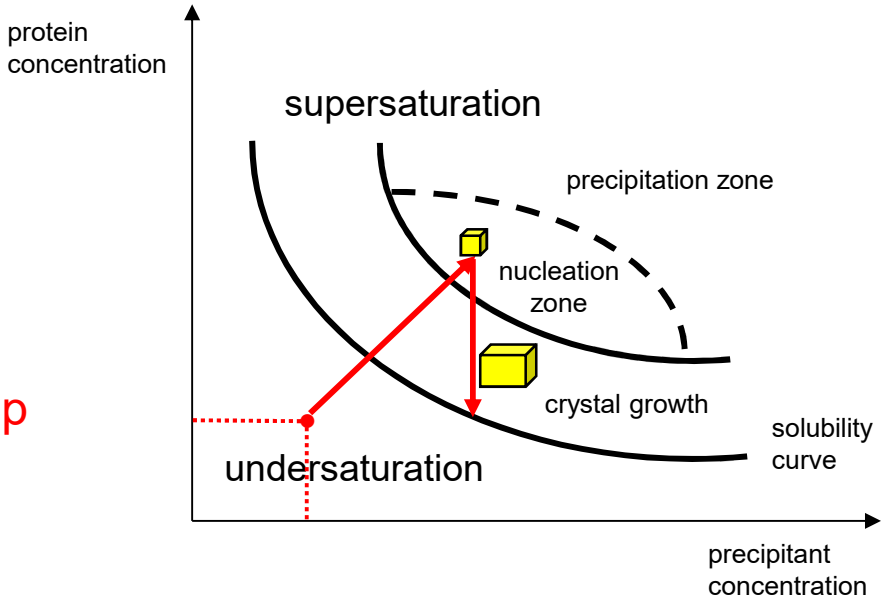
Vapor diffusion method



hanging drop



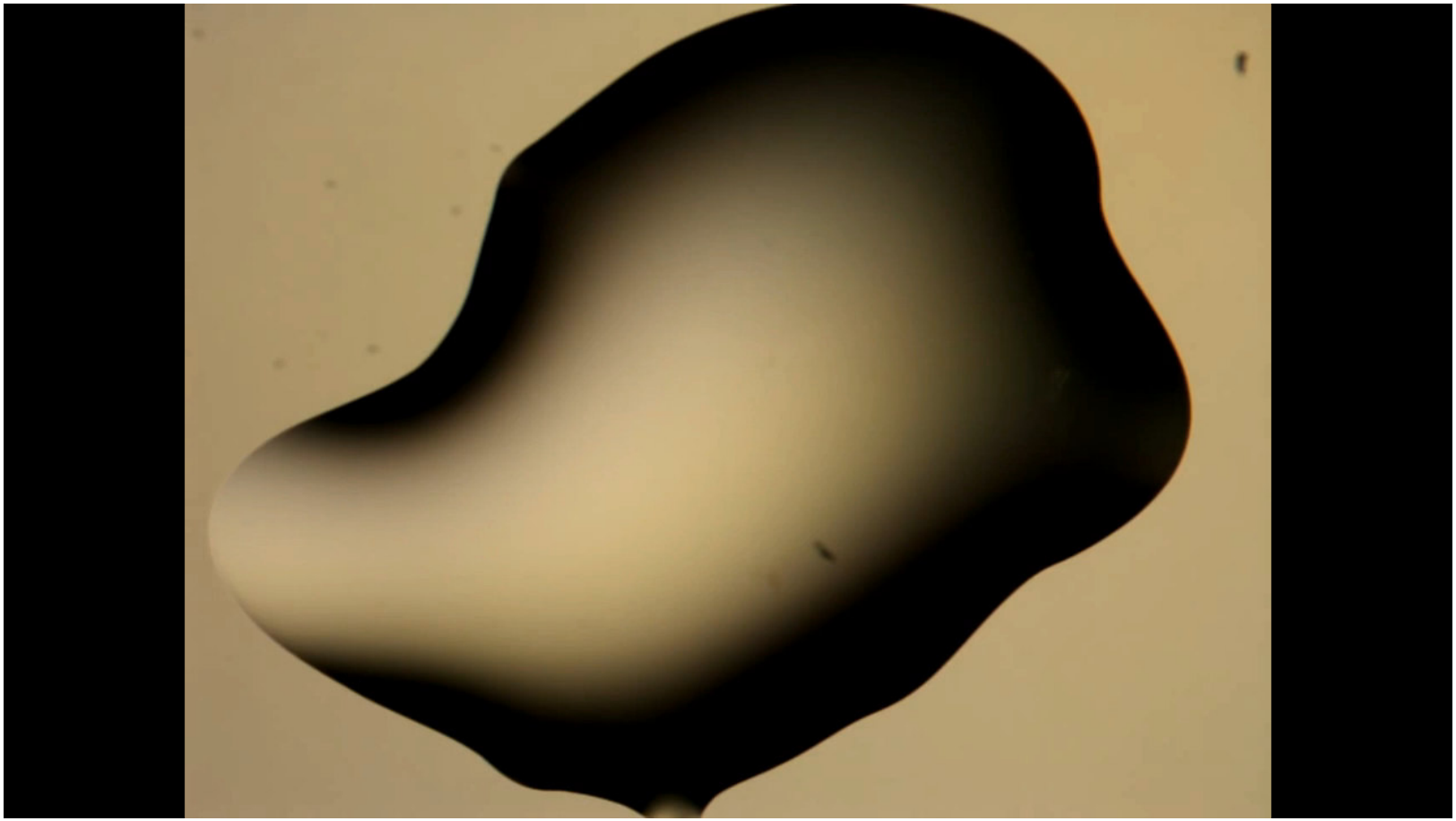
sitting drop



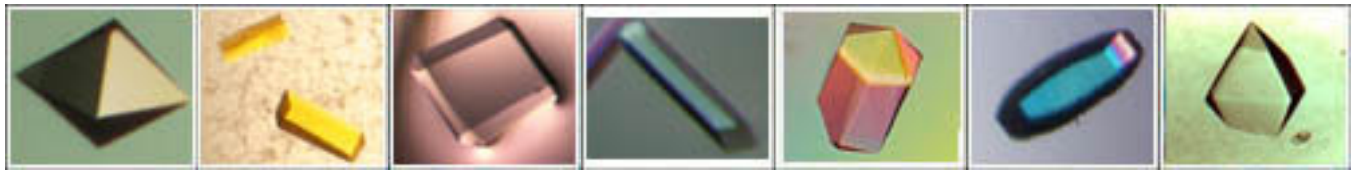
Physical-chemical parameters

- ionic strength
- buffer, pH
- temperature
- precipitant & protein concentrations
- dielectric constant of the solvent

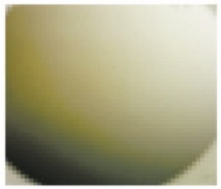
Crystallization (real life)

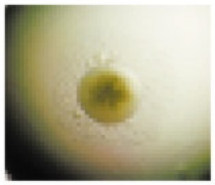


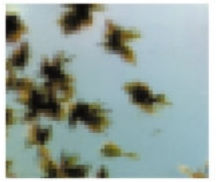
Final optimized results



Initial crystallization screening results


Clear Drop


Precipitate/Phase


Needle Cluster


Skin/Precipitate


Quasi Crystals


Plates


Precipitate


Microcrystals

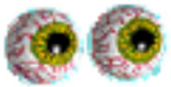

Rod Cluster

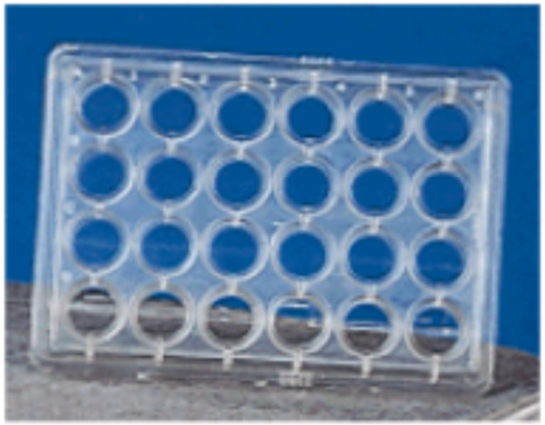
Crystallographer's fortune cookie (1)

Set up trials

…just remember:

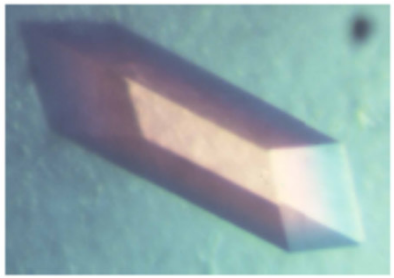
A watched crystal never grows

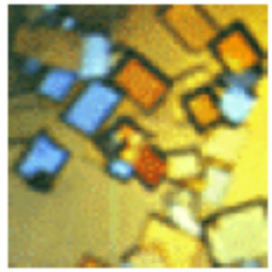




Crystallographer's fortune cookie (2)

If you see a crystal:



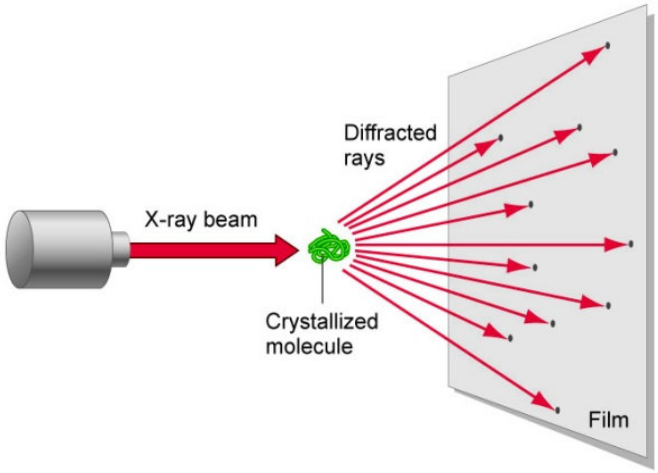


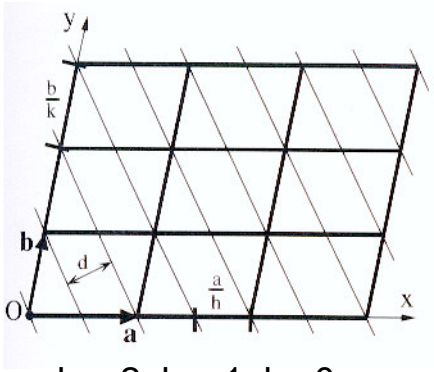
don't go running down the corridor screaming "URRA!!!" until...

you know your crystal isn't salt...

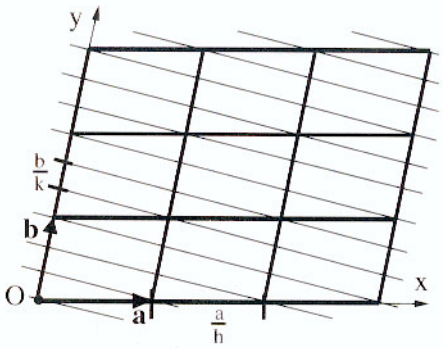
and that it diffracts

Data collection

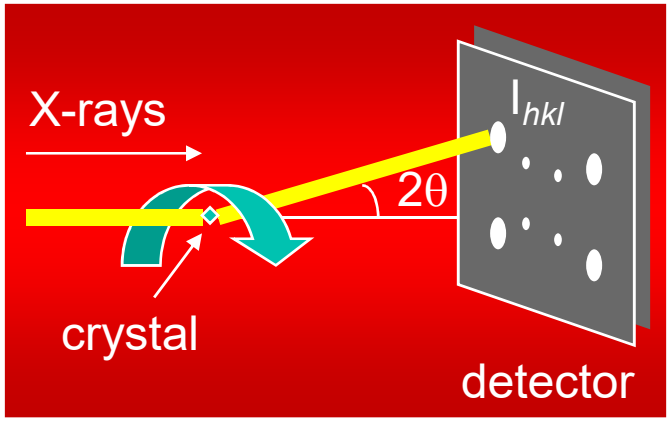




h = 2, k = 1, l = 0



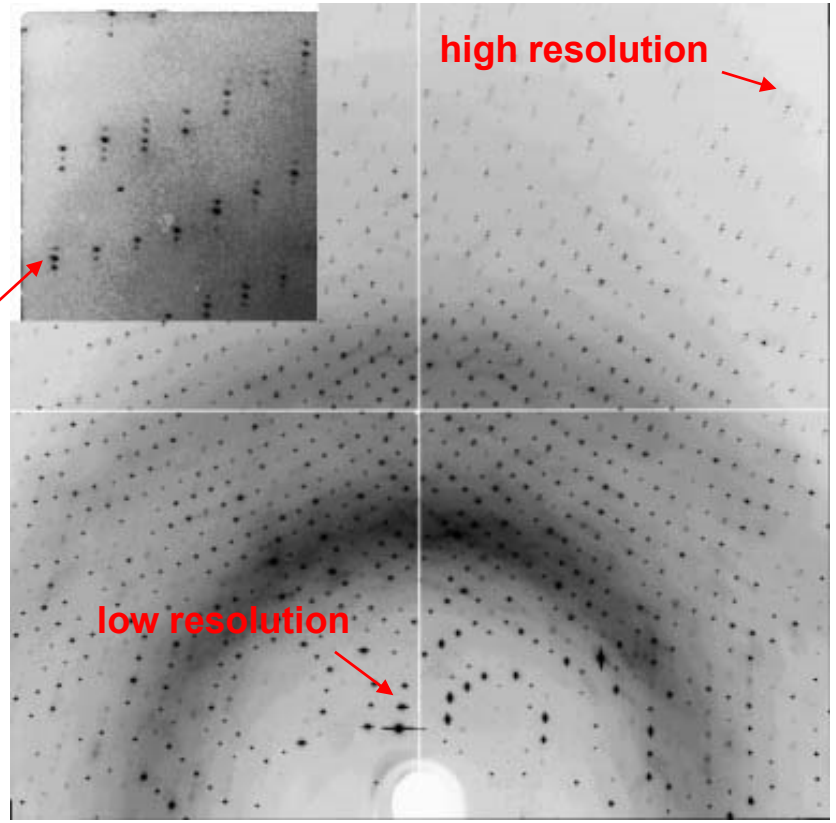
h = 1, k = 3, l = 0



$2d_{hkl} \sin\theta = n\lambda$ **(Bragg's equation)**

Data collection

h	k	l	I	σ
0	0	18	5377.7	426.7
0	0	30	87315.1	7080.9
0	0	39	79150.9	5678.3
0	0	42	88255.3	6544.6
0	0	45	14582.6	1511.1
0	0	48	8125.2	596.7
0	0	51	46929.6	3740.0
0	0	54	79917.3	8107.1
...				
etc				



Why Synchrotron radiation ?

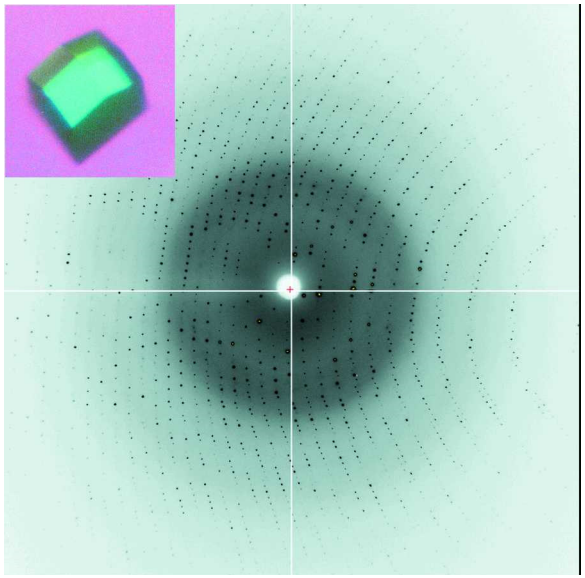
Advantages over laboratory based X-ray sources:

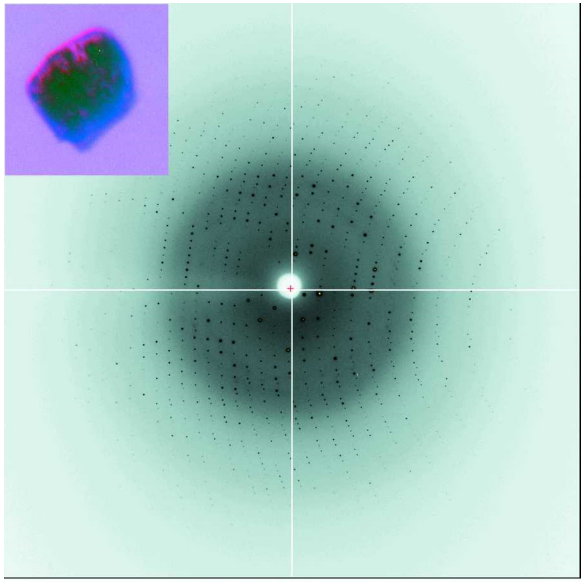
- intensity typically >10,000 times that emitted by rotating anodes
- plane polarized X-rays, with emission concentrated in a small forward pointing (flattened cone emission instead of rotating anode spherically emission)
- X-rays selected out of a relatively **wide range of wavelengths** to optimise the experiment around the sample properties (i.e. l tuned to exploit the **absorption properties** of heavier chemical elements naturally present or added to the crystal).
- **cryo-cooling** of the sample is required

Radiation damage

The energy range of X-rays used for diffraction (6 - 15 keV) is a severely ionizing radiation

⇒ formation of reactive **radicals** in the sample, which rapidly destroy any protein crystal, particularly at dose rates experienced at synchrotrons





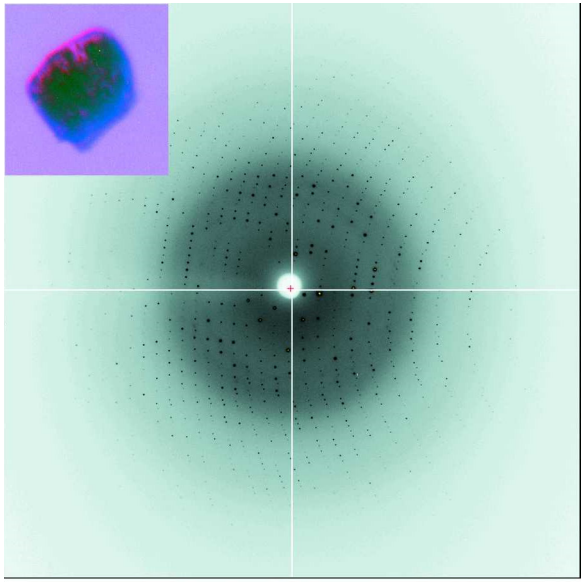
Radiation damage

The energy range of X-rays used for diffraction (6 - 15 keV) is a severely ionizing radiation

⇒ formation of reactive **radicals** in the sample, which rapidly destroy any protein crystal, particularly at dose rates experienced at synchrotrons

Symptoms:

- increase in unit cell parameters
- decrease of intensity and resolution
- non isomorphism within a data series
- site-specific damages (disulphide bond breakage, decarboxylation of acidic residues, reduction of metal centers, …)



Cryo-crystallography at the synchrotron

- an efficient way to suppress radiation damage by slowing down the kinetics of the radical reactions is **cryogenic cooling**

⇒ flash-cooling crystals to liquid nitrogen temperatures, either in cold nitrogen gas streams or directly into liquid nitrogen

- to prevent the formation of crystalline ice during flash-cooling of the crystals, **cryoprotectants** are necessary

⇒ ethylene glycol (the anti-freeze in automobile radiators), glycerol, higher alcohols, … etc

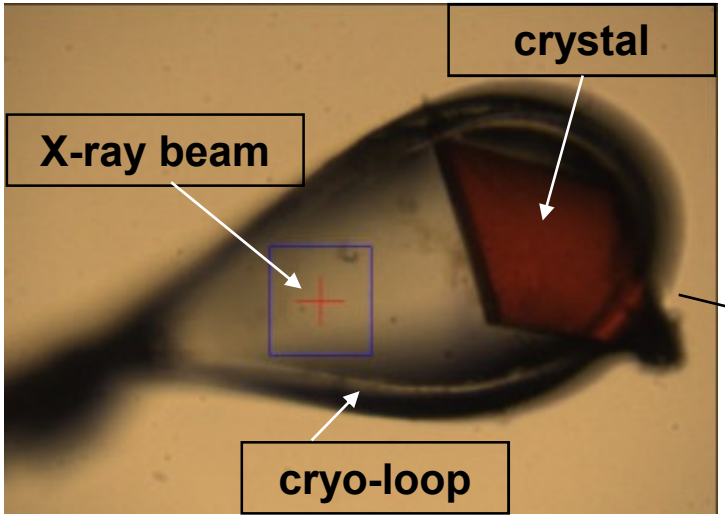
A note on nomenclature:

in our cooling experiments we want to avoid the formation of an ice phase, so we **cool** our crystals, we do **not freeze** them

Cryo-crystallography at the synchrotron

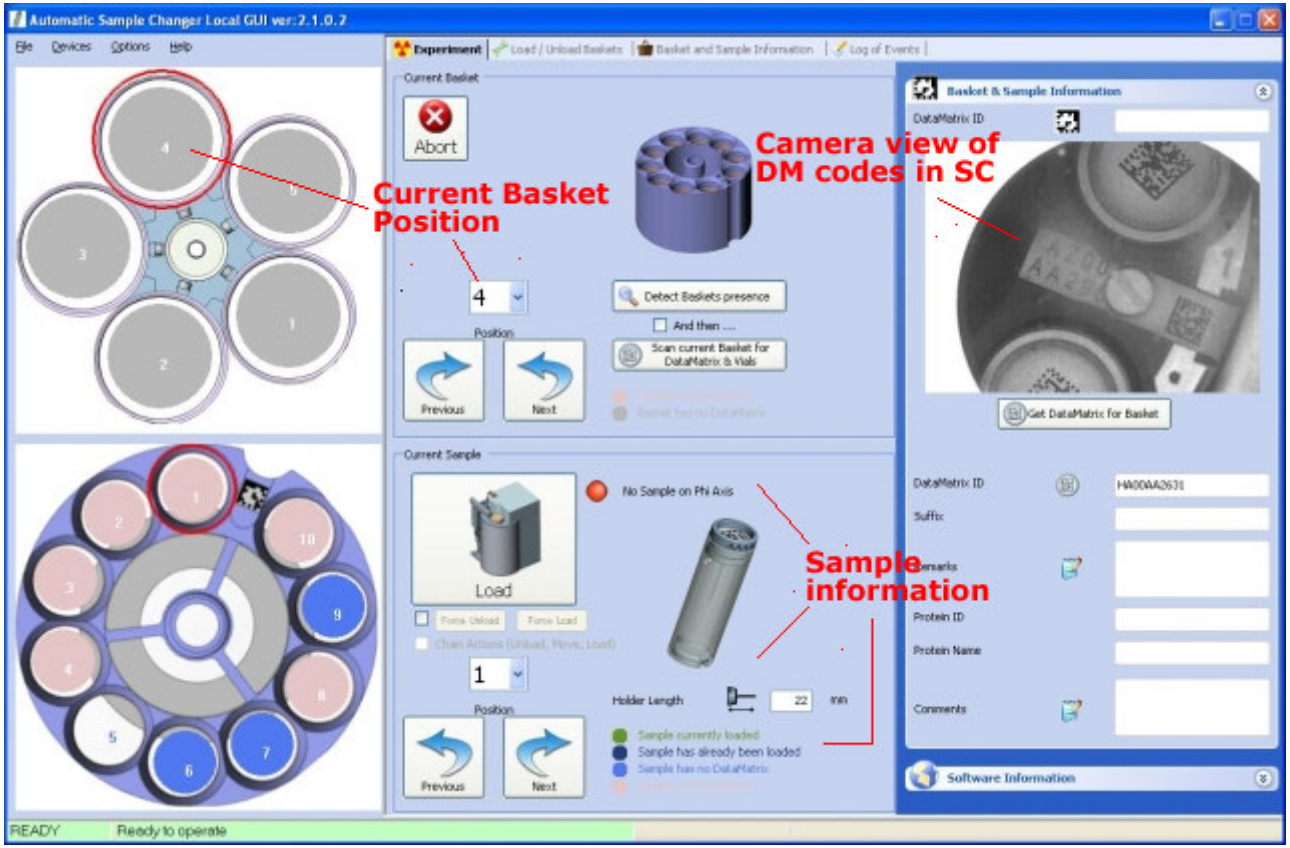
Cryo-mounting:

the crystal is removed from the crystallization drop using cryo-loops and briefly dipped into a cryoprotectant before being immersed into liquid nitrogen



Cryo-crystallography at the synchrotron

Sample changer

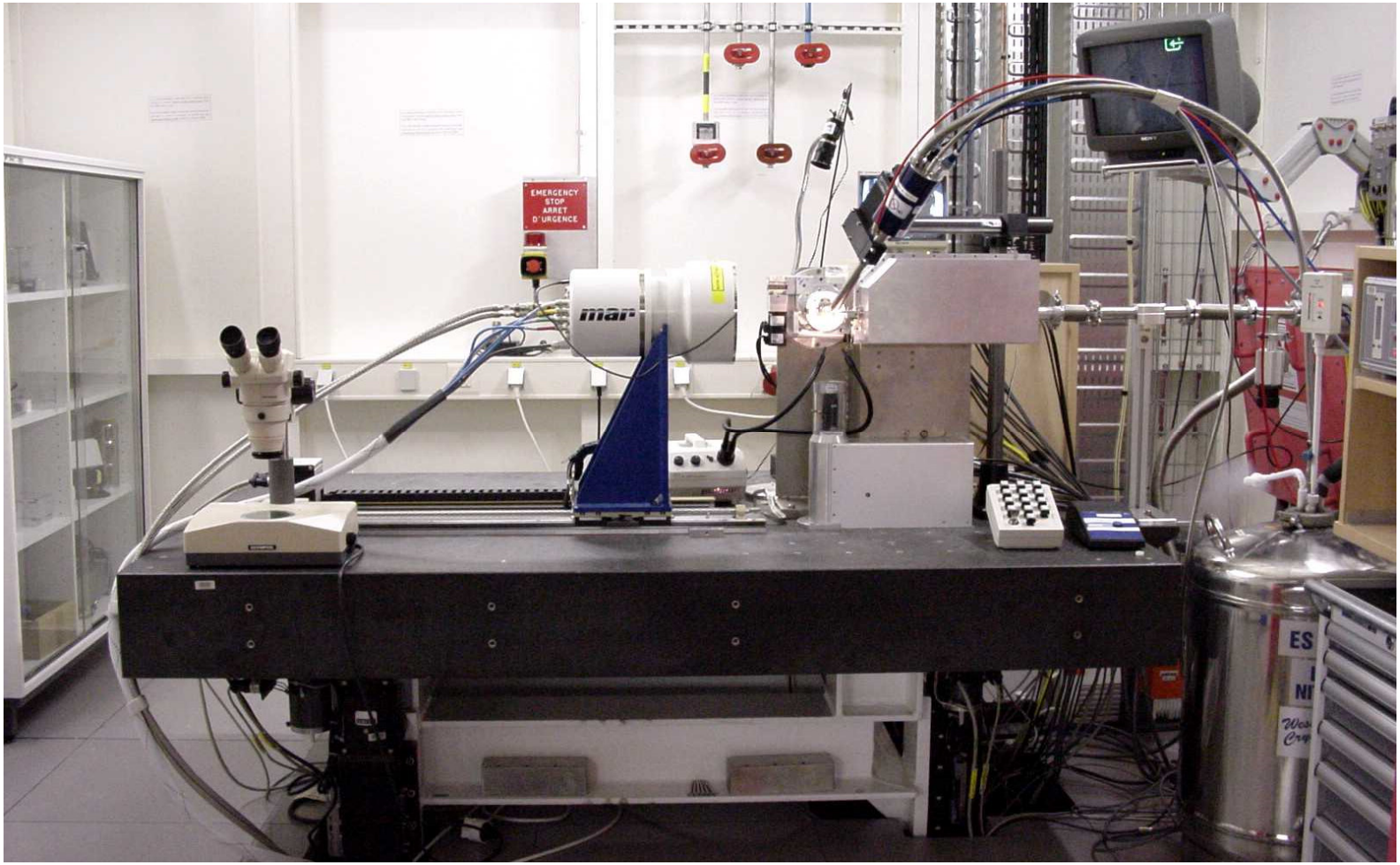


Cryo-crystallography at the synchrotron

Crystal mounting

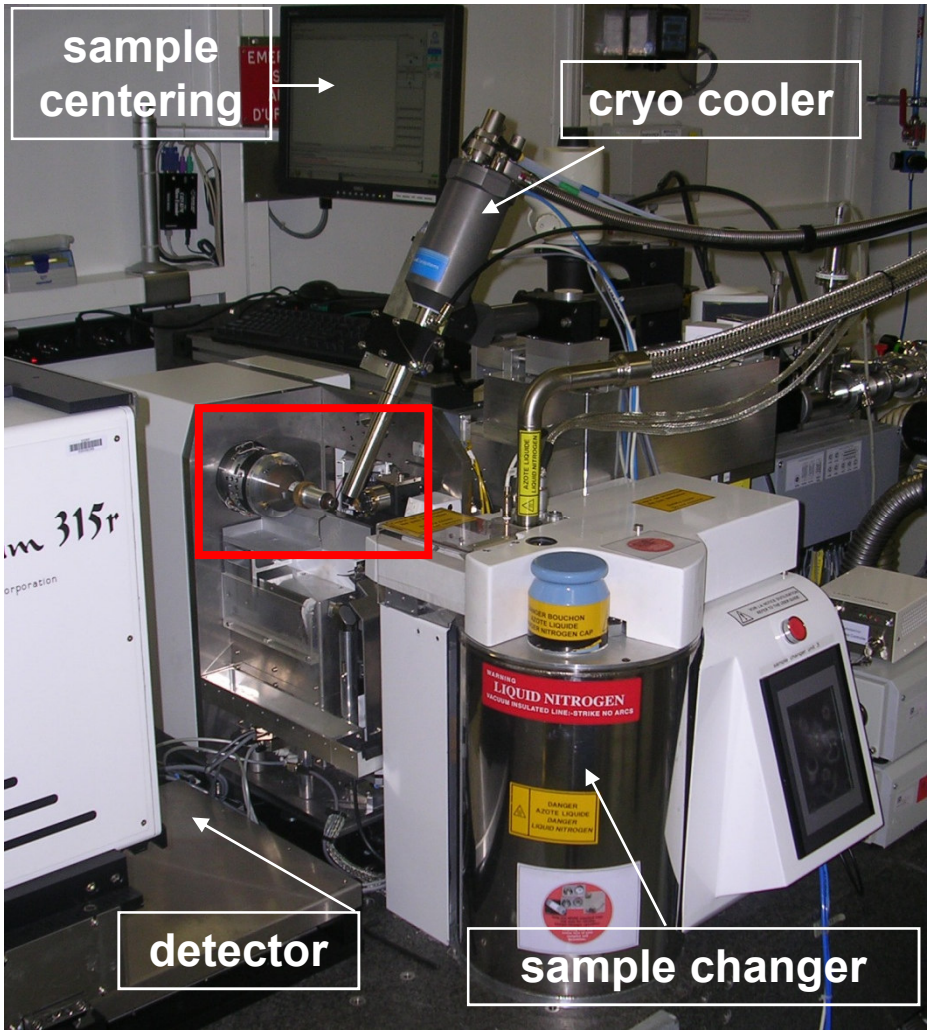
Experimental setup at synchrotron

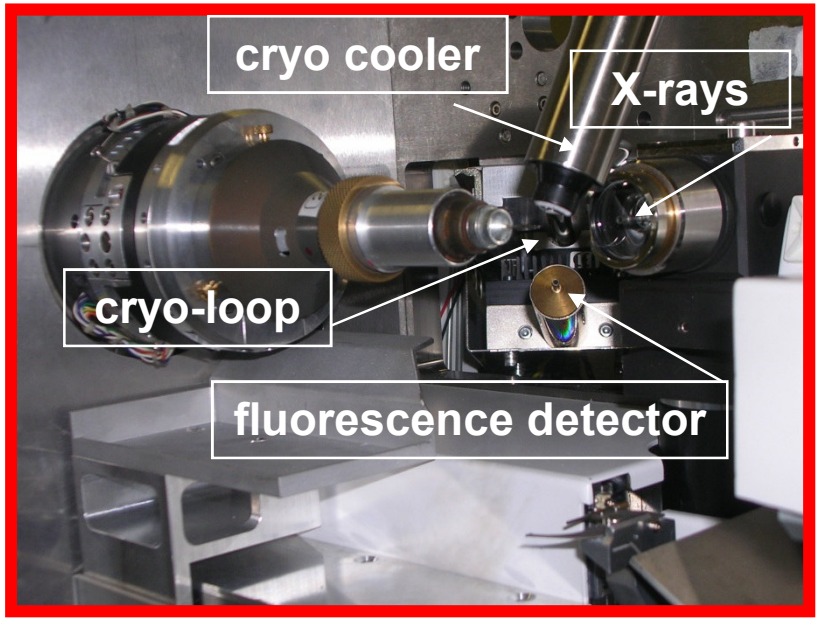
ESRF ID14-2

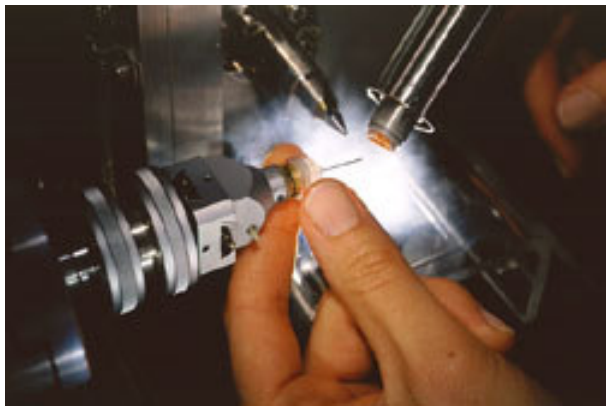


Experimental setup at synchrotron

ESRF ID23-1

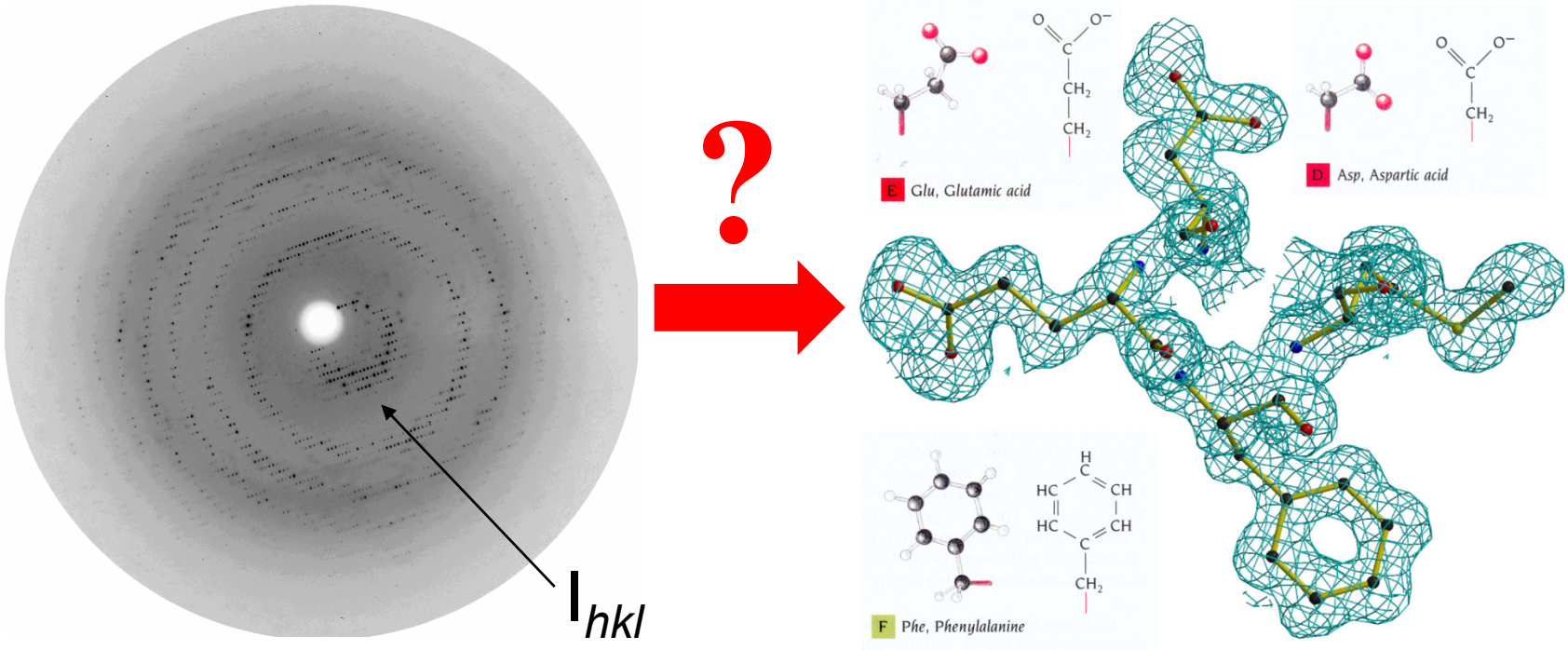




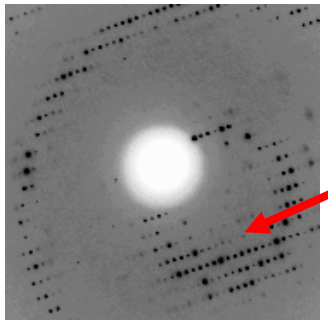


SECOND BIG PROBLEM !!

Phasing



The "phase problem"



$$I_{hkl} \propto I_0 \frac{V_{xtal}}{V_{Cell}^2} |F_{hkl}|^2$$

atomic property

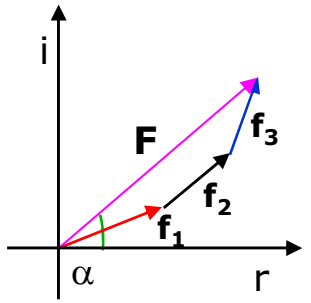
structural property (position)

$$F_{hkl} = \sum_{j=1}^{N\ atoms} f_j \exp\left[2\pi i\,(hx_j + ky_j + lz_j)\right] = |F_{hkl}| \exp(i\alpha_{hkl})$$

structure factor

$$F_{hkl} = \int_{Vcell} \rho(x,y,z) \exp\left[2\pi i(hx+ky+lz)\right] dV$$

Fourier transformation



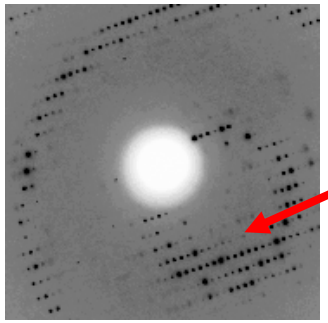
not measured

electron density

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

inverse Fourier transformation

The "phase problem"



$$I_{hkl} \propto I_0 \frac{V_{xtal}}{V_{Cell}^2} |F_{hkl}|^2$$

atomic property

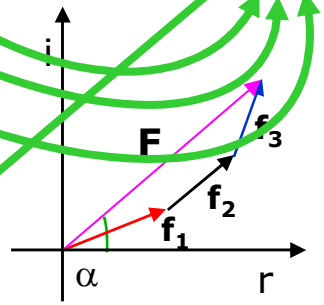
structural property (position)

$$F_{hkl} = \sum_{j=1}^{N\ atoms} f_j \exp\left[2\pi i\,(hx_j + ky_j + lz_j)\right] = |F_{hkl}| \exp(i\alpha_{hkl})$$

structure factor

$$F_{hkl} = \int_{Vcell} \rho(x,y,z) \exp\left[2\pi i(hx+ky+lz)\right] dV$$

Fourier transformation



not measured

electron density

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

inverse Fourier transformation

The "phase problem"

in X-ray crystallography, there are several ways to recover the lost phases:

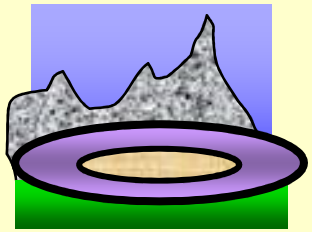
- **Molecular Replacement** (**MR**) **method**
  (synchrotron radiation not required)

- **Heavy atom method**

  - **Multiple Isomorphous Replacement** (**MIR**)
    (synchrotron radiation not required)

  - **Single (or Multiple) Isomorphous Replacement with anomalous scattering** (**SIRAS** or **MIRAS**)
    (synchrotron radiation required)

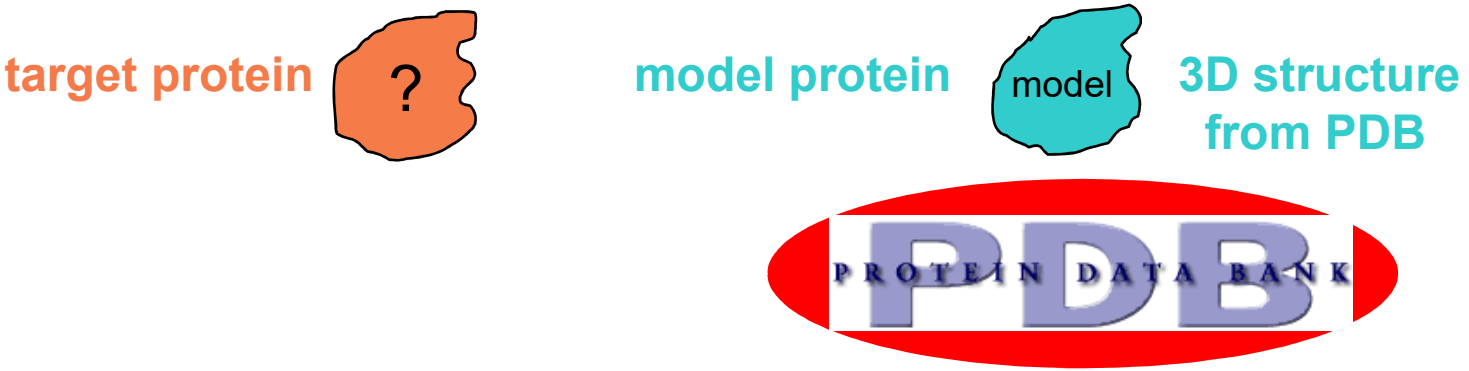
  - **Multiple wavelength Anomalous Diffraction** (**MAD**)
    (synchrotron radiation required)



Synchrotron

Molecular Replacement method

- a **good model** for the protein of unknown structure is needed



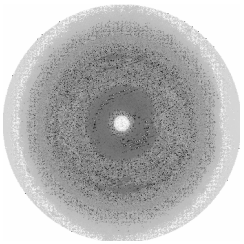

```
Protein .....YKTQAGKTVDYINAAIGG----SADGAGL....
Model   .....YQTQASKTVDYITAALAGSRNVSADAAGL....
              * *** ****** **  *    *** ***
```

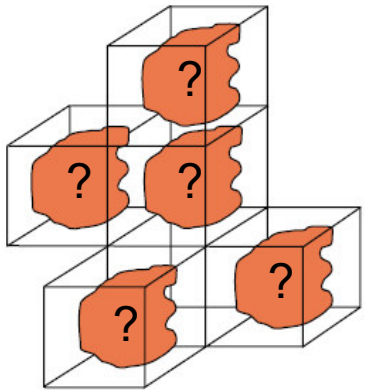
≥ 30% sequence identity ➡ good 3D structure homology

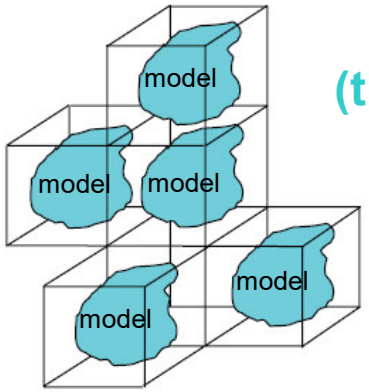
Molecular Replacement method

- the “calculated” phases can be obtained by simulating the molecule's packing in the crystal (using the model protein)

target protein (real crystal)







model protein (theoretical crystal)

→ (x_j, y_j, z_j) for every atom j in the crystal

$$I^{obs}_{hkl} \propto |F^{obs}_{hkl}|^2$$

same unit cell and crystallographic symmetry

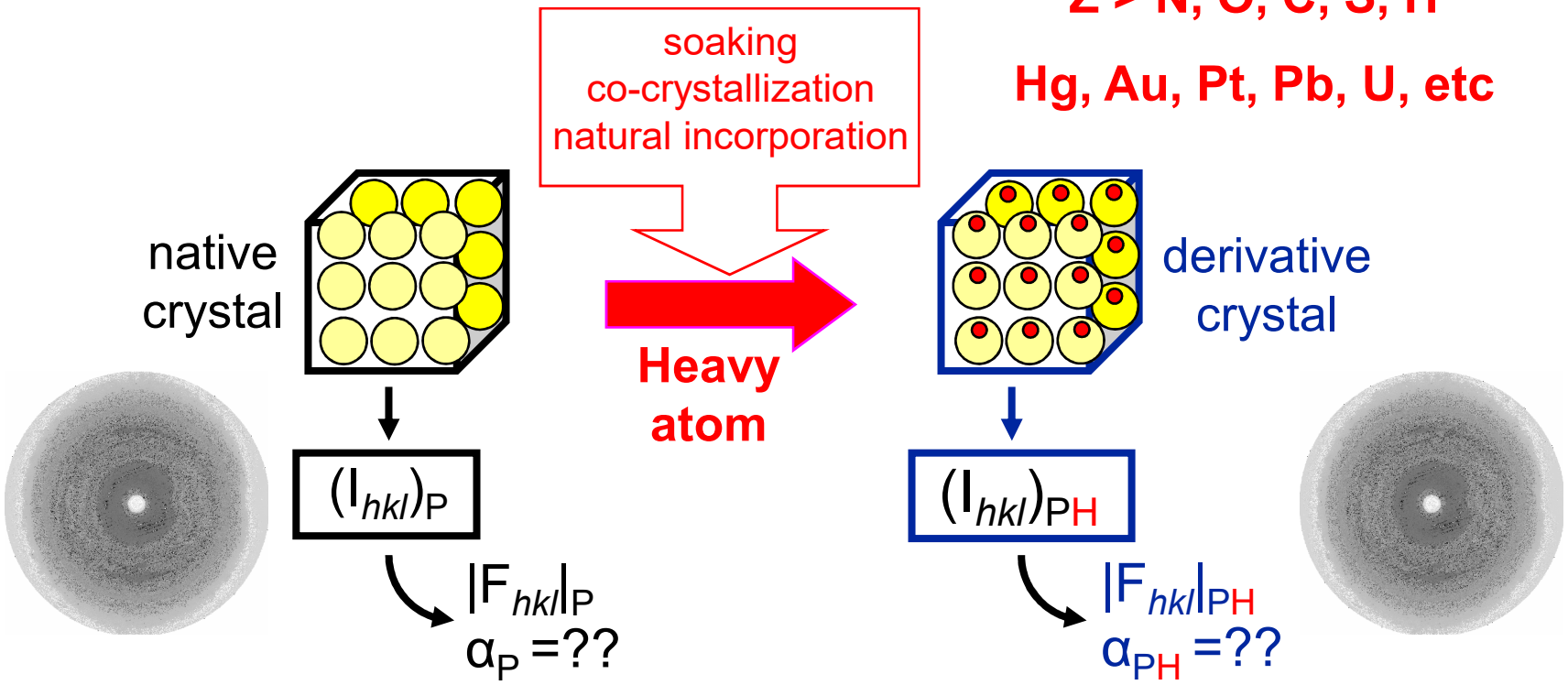
$$F^{calc}_{hkl} = \sum_{j=1}^{N\ atoms} f_j \exp\left[2\pi i\,(hx_j + ky_j + lz_j)\right] = |F^{calc}_{hkl}| \exp\left(i\alpha^{calc}_{hkl}\right)$$

$$\rho(x,y,z) = \frac{1}{V} \sum_{hkl} \left(|F^{obs}_{hkl}|\, e^{i\alpha^{calc}_{hkl}} \right) e^{-2\pi i(hx+ky+lz)}$$

Heavy atom method

Z > N, O, C, S, H

Hg, Au, Pt, Pb, U, etc

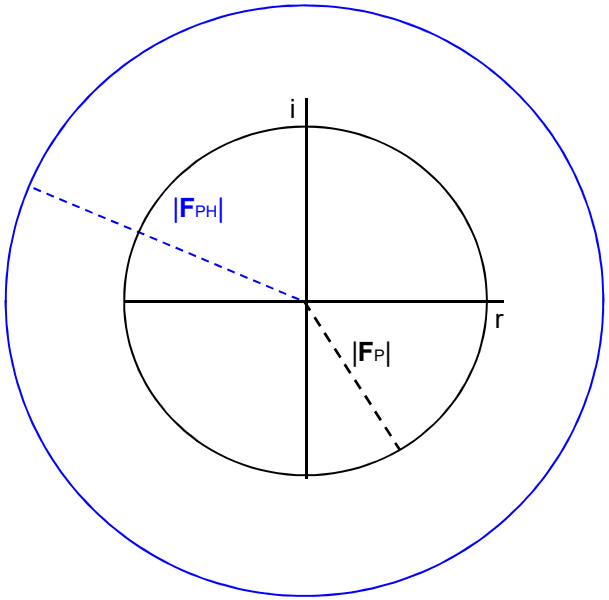


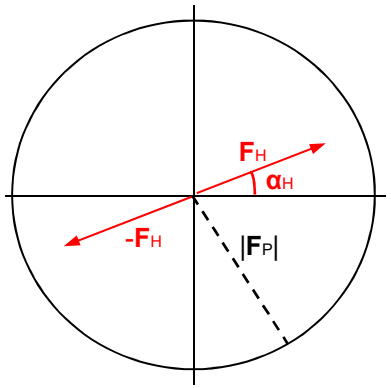
- isomorphism between native and derivative crystals is required
- comparison of the diffraction pattern of the two crystals

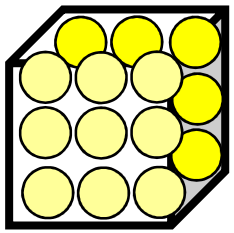
➡ identification of the heavy atom positions within the unit cell

➡ calculation of approximate initial phases α_P (for every hkl)

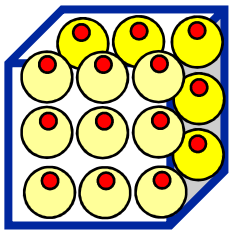
Heavy atom method: MIR







native crystal



derivative crystal

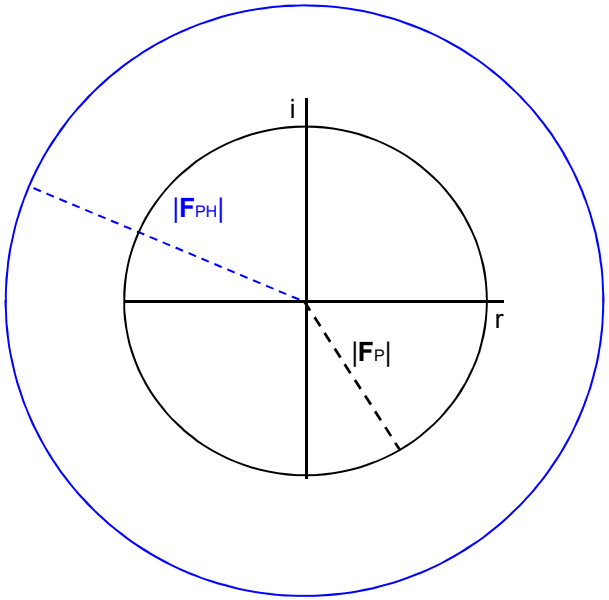
heavy atom positions from Patterson synthesis $\Rightarrow |\mathbf{F}_H|$ and α_H

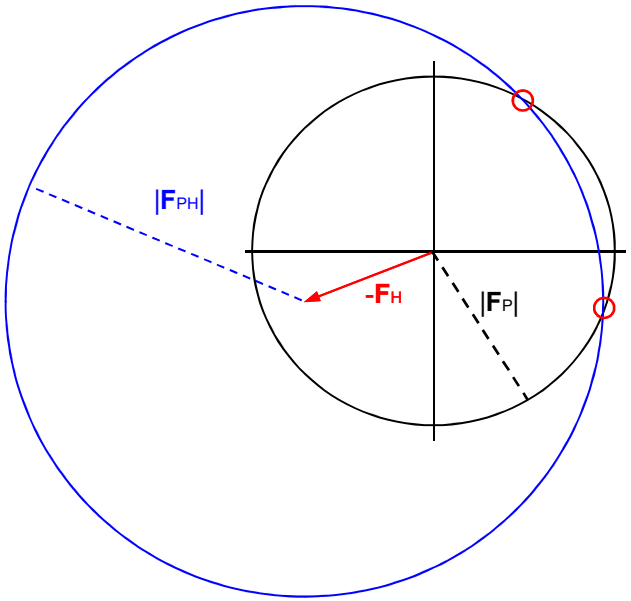
N heavy atoms

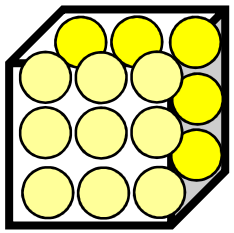
$$\mathbf{F}_H = \sum_{j=1} f_j \exp\left[2\pi i \left(hx_j + ky_j + lz_j\right)\right] =$$

$$= |\mathbf{F}_H| \exp\left(i\alpha_H\right) \quad \text{for every } hkl$$

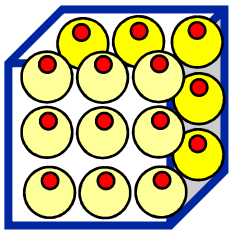
Heavy atom method: MIR





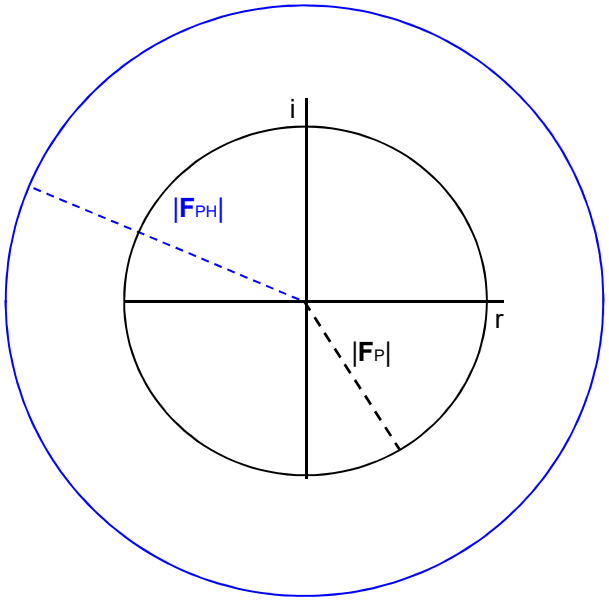


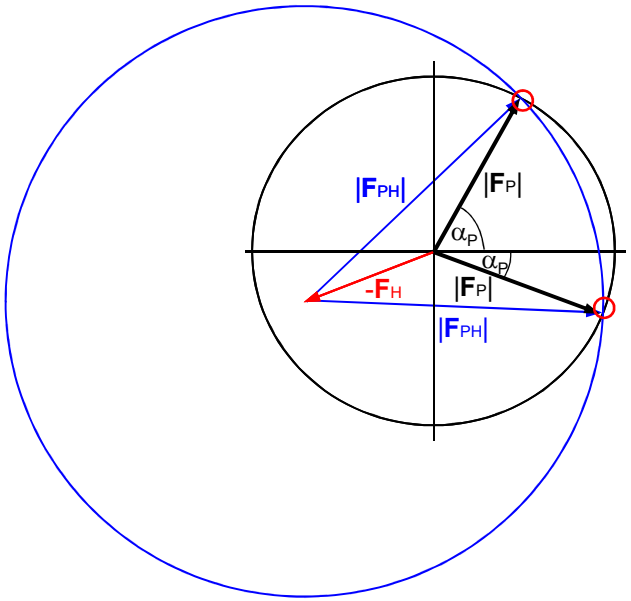
native crystal

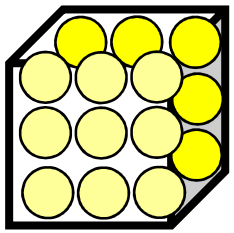


derivative crystal

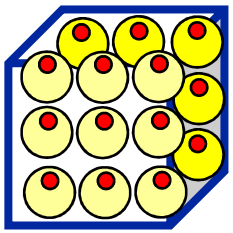
Heavy atom method: MIR







native crystal

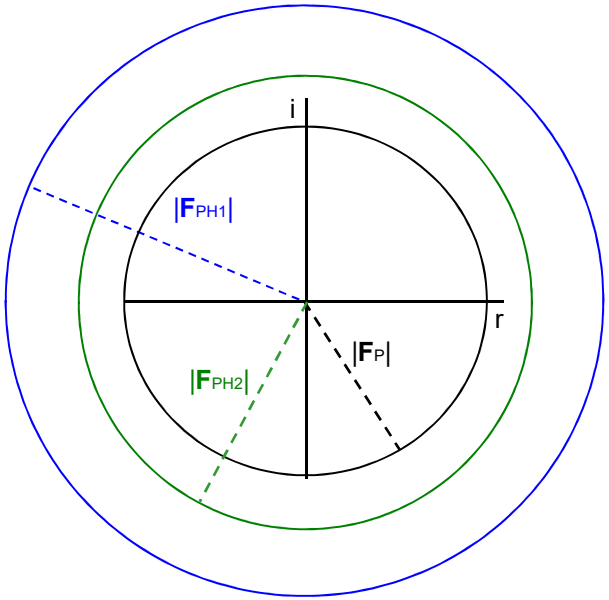


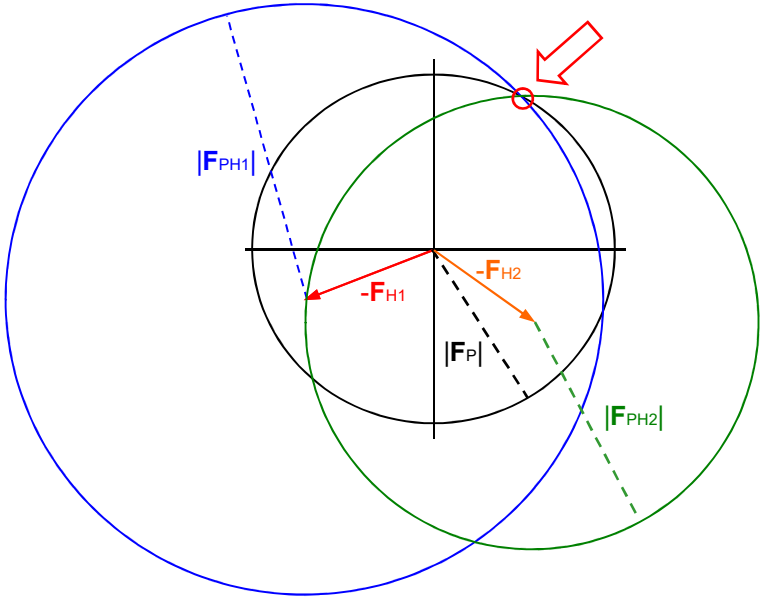
derivative crystal

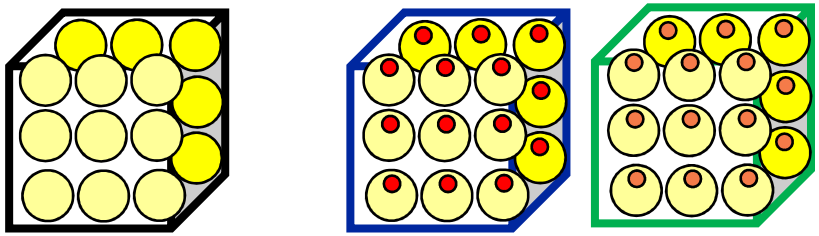
$$F_{PH} = F_P - (-F_H)$$

2 possible phases for F_P

Heavy atom method: MIR







native crystal

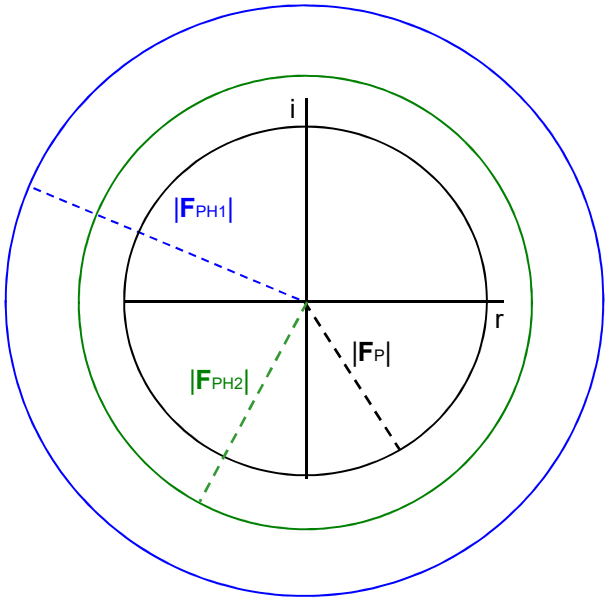
derivative crystal 1

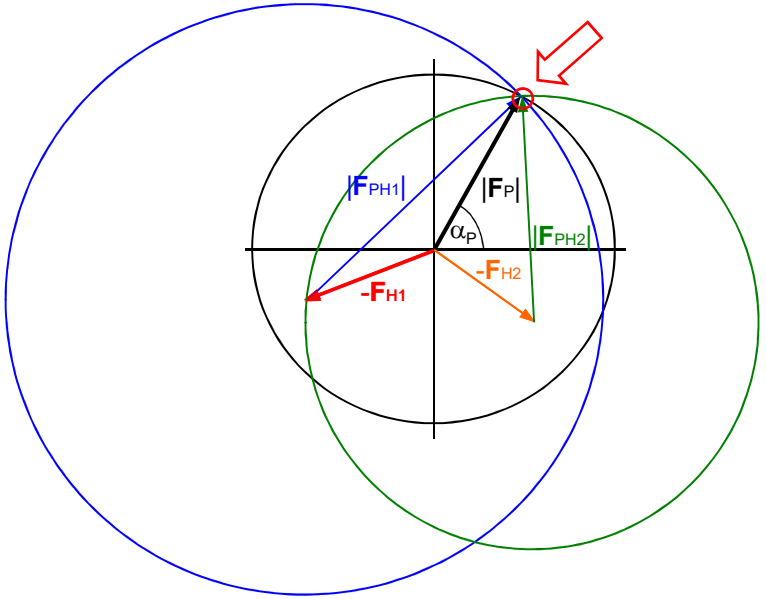
derivative crystal 2

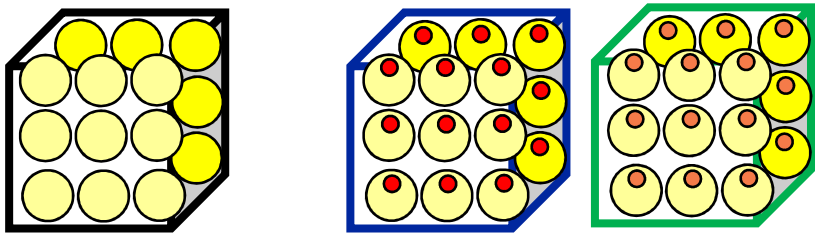
heavy atom position from Patterson synthesis $\Rightarrow |F_{H1}|$ and α_{H1}

$\Rightarrow |F_{H2}|$ and α_{H2}

Heavy atom method: MIR







native crystal

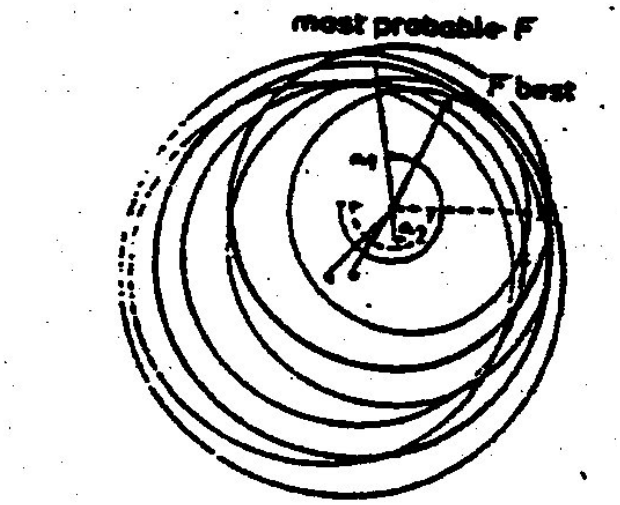
derivative crystal 1

derivative crystal 2

Heavy atom method: MIR

real case:

- experimental errors in $|\mathbf{F}_{PH}|$ and $|\mathbf{F}_{P}|$
- poor isomorphism between native and derivative crystals
- errors in heavy atom localization (Patterson "noise", $\mathbf{F}_{H}$)



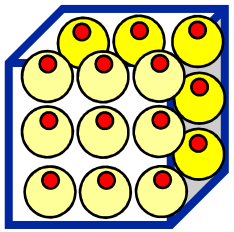
m = "figure of merit"
$0 \leq m \leq 1$

$$\rho(x,y,z) = 1/V \sum_{hkl} m_{hkl} |\mathbf{F}_{hkl}| \exp [i\alpha(best)_{hkl}] \exp [-2\pi i(hx+ky+lz)]$$

Anomalous scattering

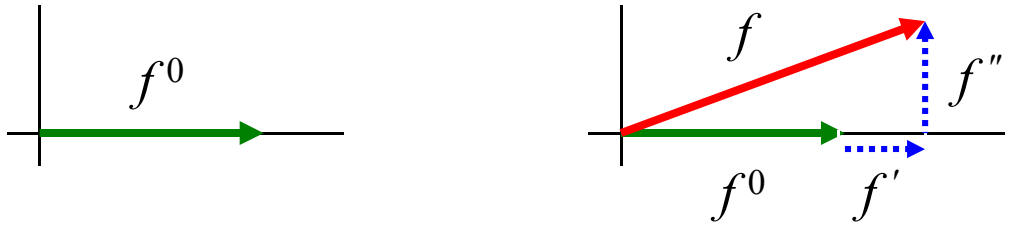
X-rays

($\lambda \cong$ **absorption edge**)



derivative crystal

- in this technique, atoms' inner electrons absorb X-rays of particular wavelengths, and reemit the X-rays after a delay, inducing a **phase shift** in all of the reflections, known as the **anomalous dispersion effect**



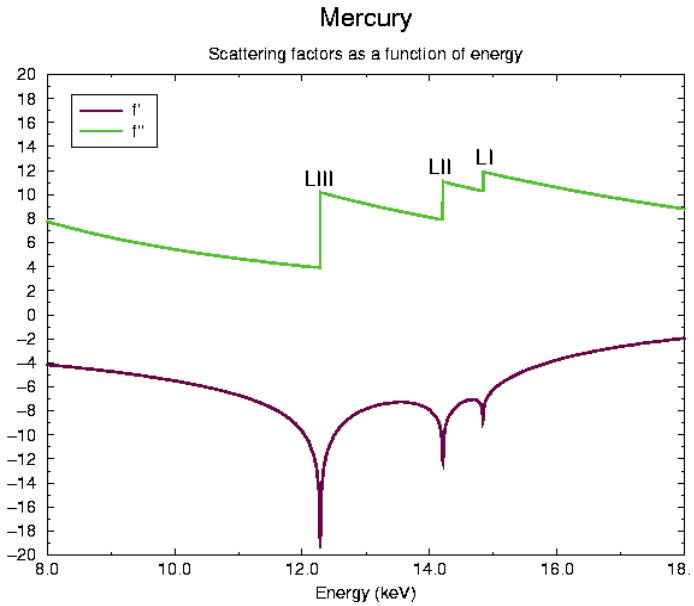
$$f_{anomalous}(\theta,\lambda) = f^0(\theta) + f'(\lambda) + i\, f''(\lambda)$$

Anomalous scattering

$$f_{anomalous}(\theta,\lambda) = f^0(\theta) + f'(\lambda) + i\, f''(\lambda)$$

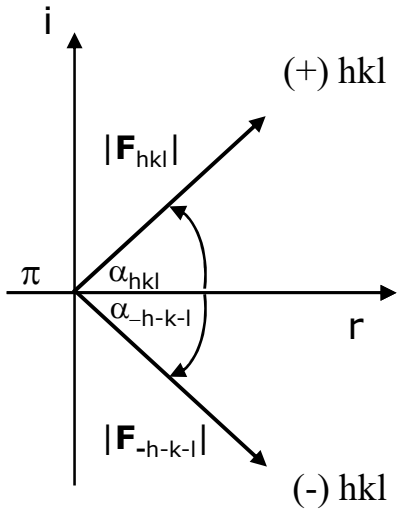
$$f''(E) = \frac{mc}{4\pi e^2 \hbar} E\, \mu_a(E)$$

$$f'(E) = \frac{2}{\pi} \int \frac{E'\, f''(E')}{(E^2 - E'^2)}\, dE'$$



The Friedel's law

$$I_{hkl} = I_{-h-k-l}\ ,\ |\mathbf{F}_{hkl}| = |\mathbf{F}_{-h-k-l}| \text{ and } \alpha_{hkl} = -\alpha_{-h-k-l}$$

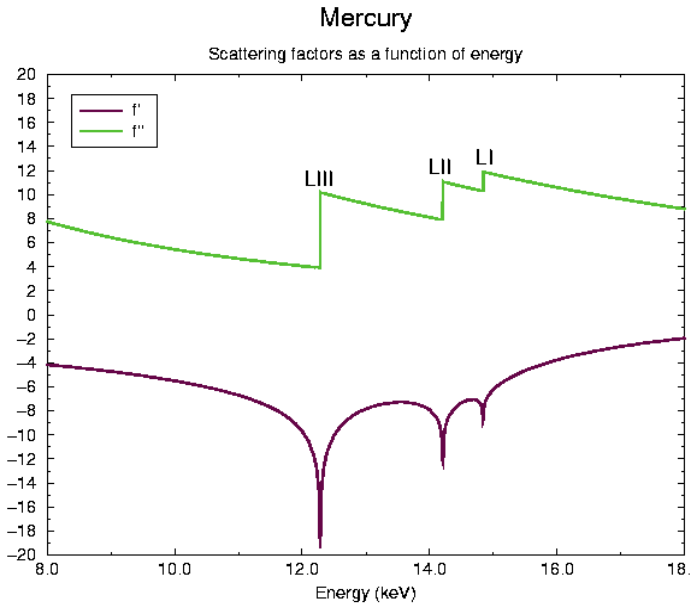


Anomalous scattering

$$f_{anomalous}(\theta,\lambda) = f^0(\theta) + f'(\lambda) + i\, f''(\lambda)$$

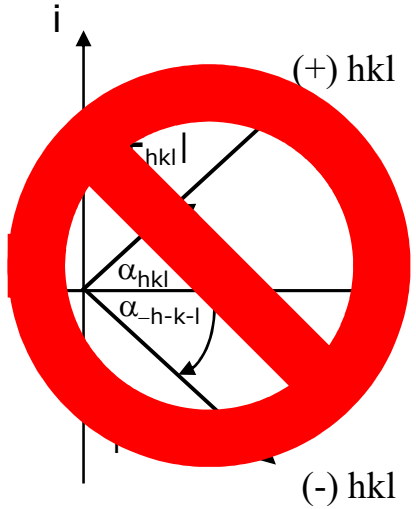
$$f''(E) = \frac{mc}{4\pi e^2 \hbar} E\, \mu_a(E)$$

$$f'(E) = \frac{2}{\pi} \int \frac{E'\, f''(E')}{(E^2 - E'^2)}\, dE'$$



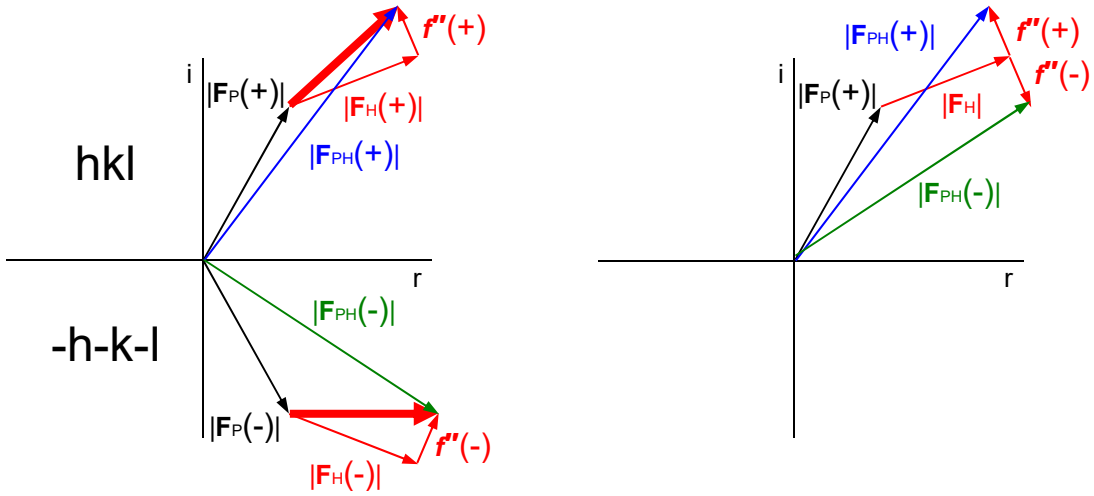
The Friedel's law

$$I_{hkl} = I_{-h-k-l}\ ,\ |F_{hkl}| = |F_{-h-k-l}| \text{ and } \alpha_{hkl} = -\alpha_{-h-k-l}$$



Anomalous scattering

$$f_{anomalous}\,(\theta,\lambda) = f^{0}(\theta) + f'(\lambda) + i\,f''(\lambda)$$

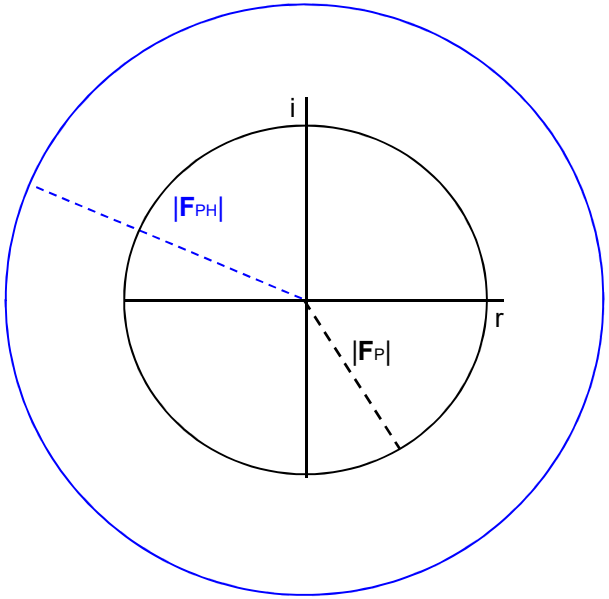


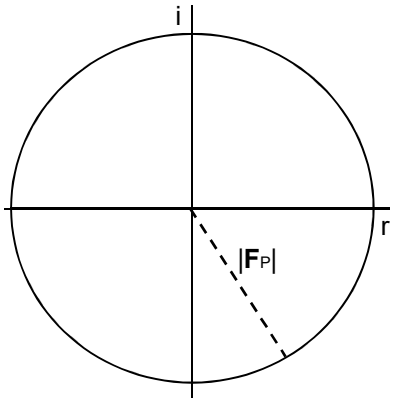
$$|F_{PH+}| \neq |F_{PH-}| \text{ and } \alpha_{PH+} \neq -\alpha_{PH-}$$

F real contribution from anomalous scattering atoms

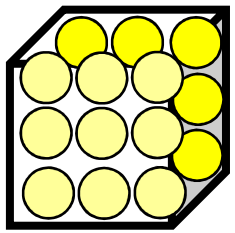
f'''(+) and ***f'''***(-) imaginary contribution from anomalous scattering atoms

Heavy atom method with anomalous scattering: SIRAS

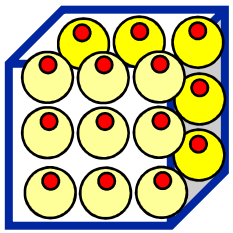




(λ ≅ **absorption edge of the heavy atom**)

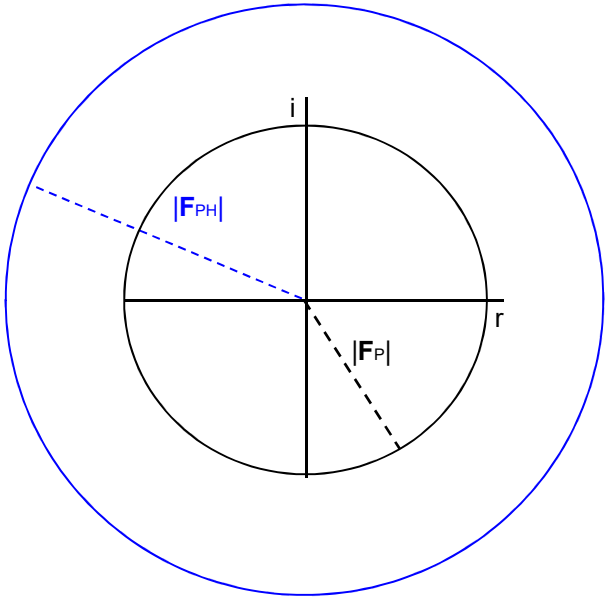


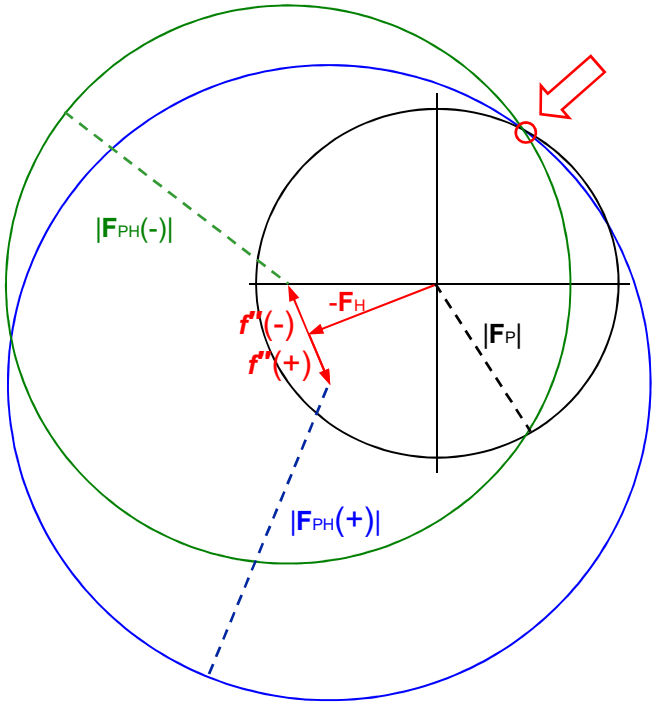
native
crystal



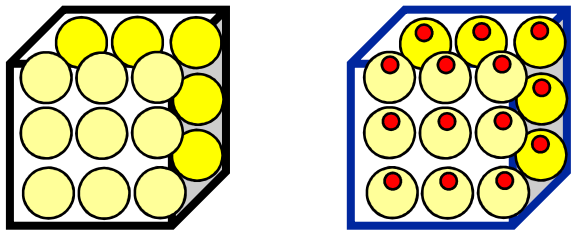
derivative
crystal 1

Heavy atom method with anomalous scattering: SIRAS





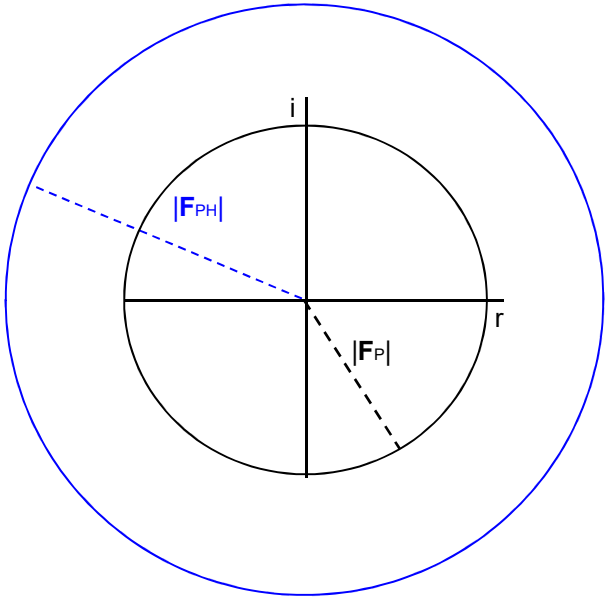
(λ ≅ **absorption edge of the heavy atom**)

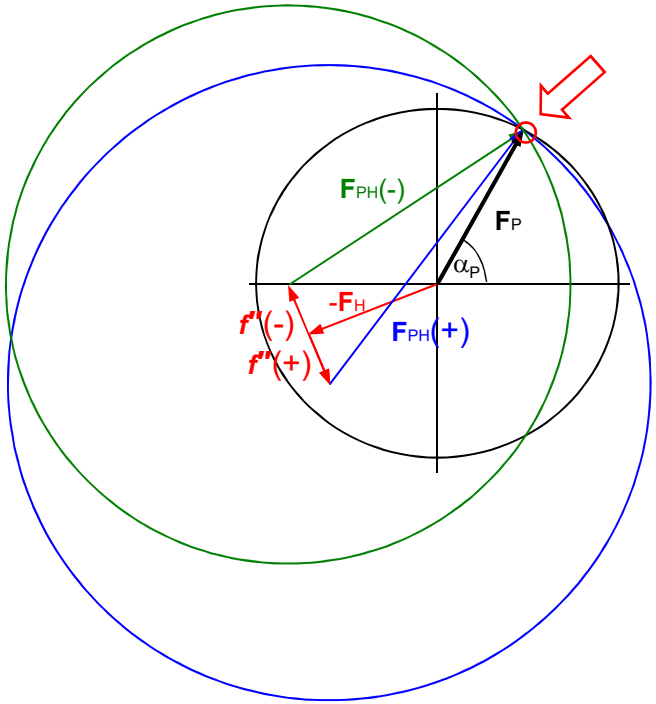


native crystal

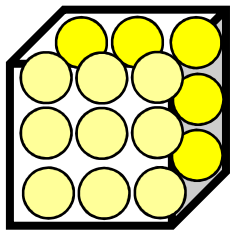
derivative crystal 1

Heavy atom method with anomalous scattering: SIRAS

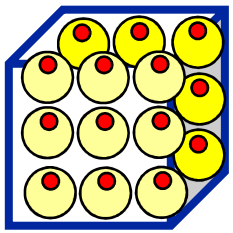




(λ ≅ absorption edge of the heavy atom)



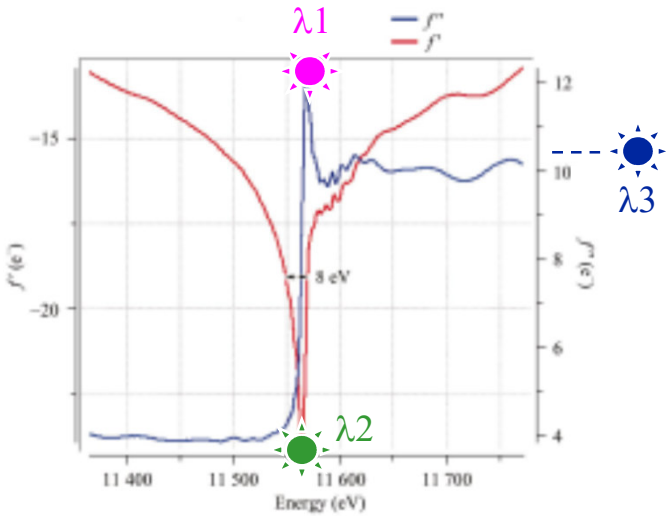
native crystal



derivative crystal 1

Heavy atom method with anomalous scattering: MAD

$$f_{anomalous}(\theta,\lambda) = f^{0}(\theta) + f'(\lambda) + i\, f''(\lambda)$$



λ1) *f* '' has its maximum (maximum of difference between Bijvoet pairs)

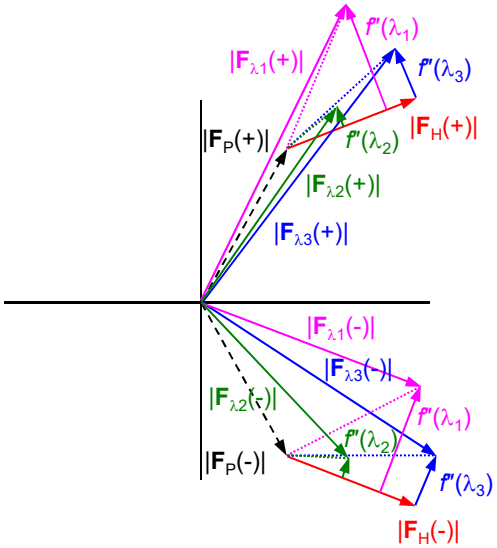
λ2) f' has its minimum (maximum of dispersive difference)

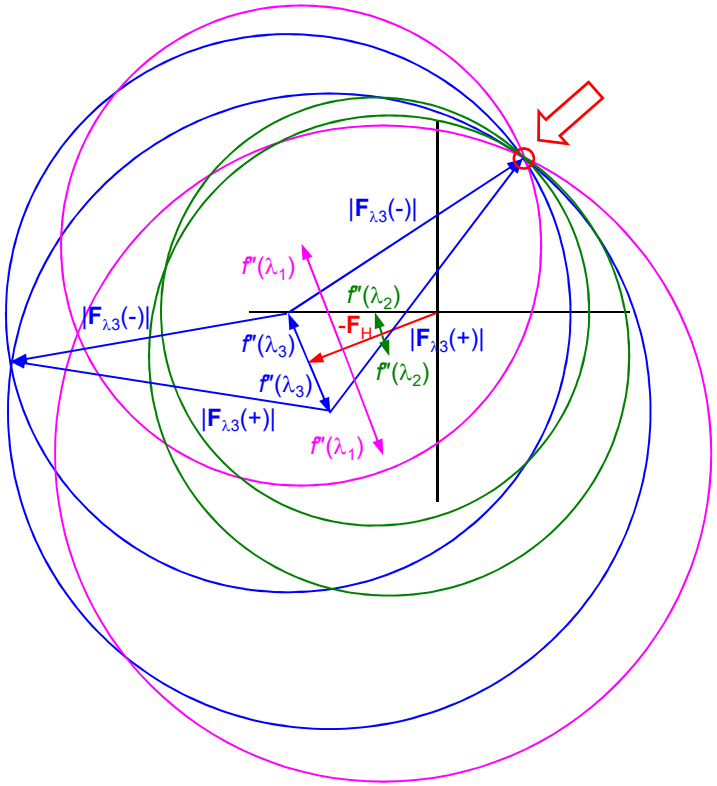
$$\Delta F_{\Delta\lambda}(hkl) = \overline{|\mathbf{F}^{\lambda 2}(hkl)|} - \overline{|\mathbf{F}^{\lambda 3}(hkl)|}$$

$$\text{where } \overline{|\mathbf{F}^{\lambda}(hkl)|} = \frac{|\mathbf{F}^{\lambda}(hkl)| + |\mathbf{F}^{\lambda}(-h-k-l)|}{2}$$

λ3) "*remote*", where *f* ' e *f* '' are small (almost no anomalous scattering) (>1000 eV from λ1)

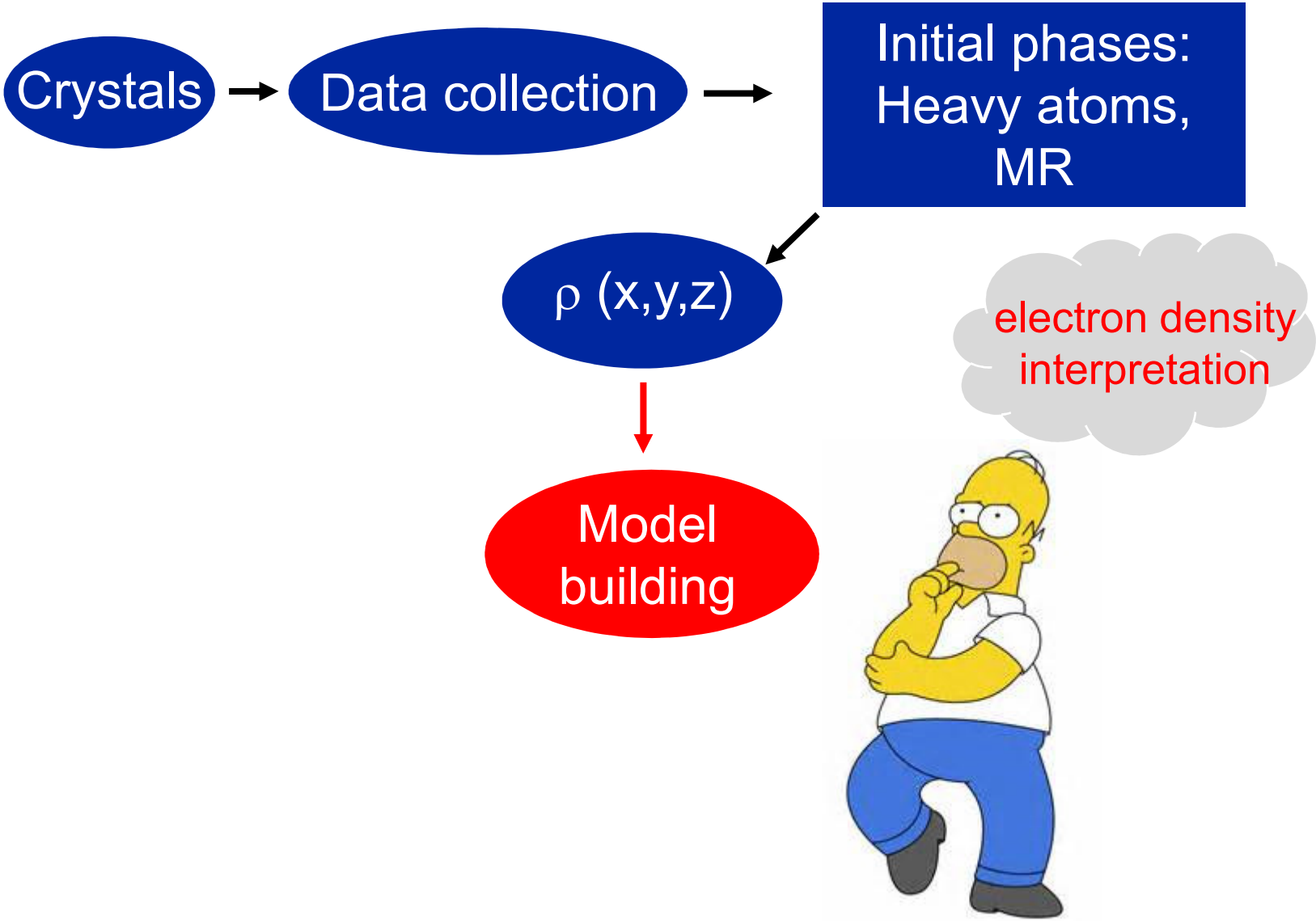
Heavy atom method with anomalous scattering: MAD





Heavy atom method with anomalous scattering: MAD

- presence of heavy atoms with strong anomalous signal (soaking or co-crystallizzation, **"natural" incorporation**)
- only **one crystal** is required
- **3 data sets** at 3 different wavelengths
- cryogenic conditions
- tunable radiation
- wavelengths are **carefully chosen** to optimize the difference in intensity of Bijvoet pairs and between the diffraction at selected wavelengths


Crystals
Data collection
Initial phases: Heavy atoms, MR
ρ (x,y,z)
electron density interpretation
Model building

The importance of the resolution

$$\rho(x,y,z) = \frac{1}{V}\sum_{hkl} |\mathbf{F}_{hkl}| \exp\,[-2\pi i(hx+ky+lz-\alpha'_{hkl})]$$

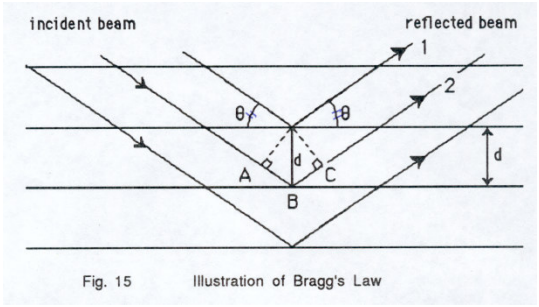
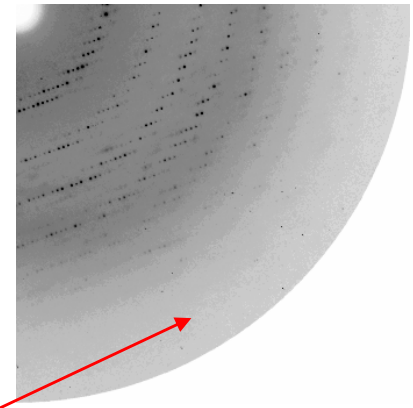


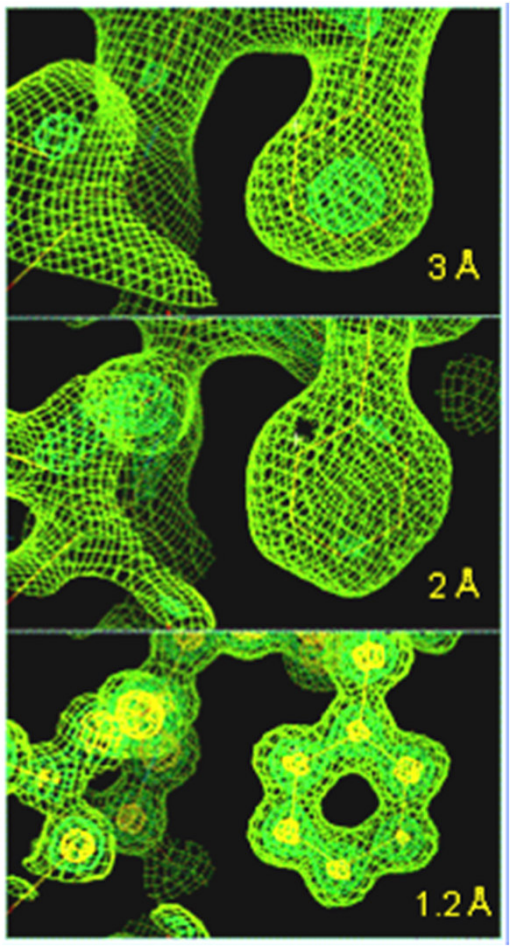
Fig. 15 Illustration of Bragg's Law



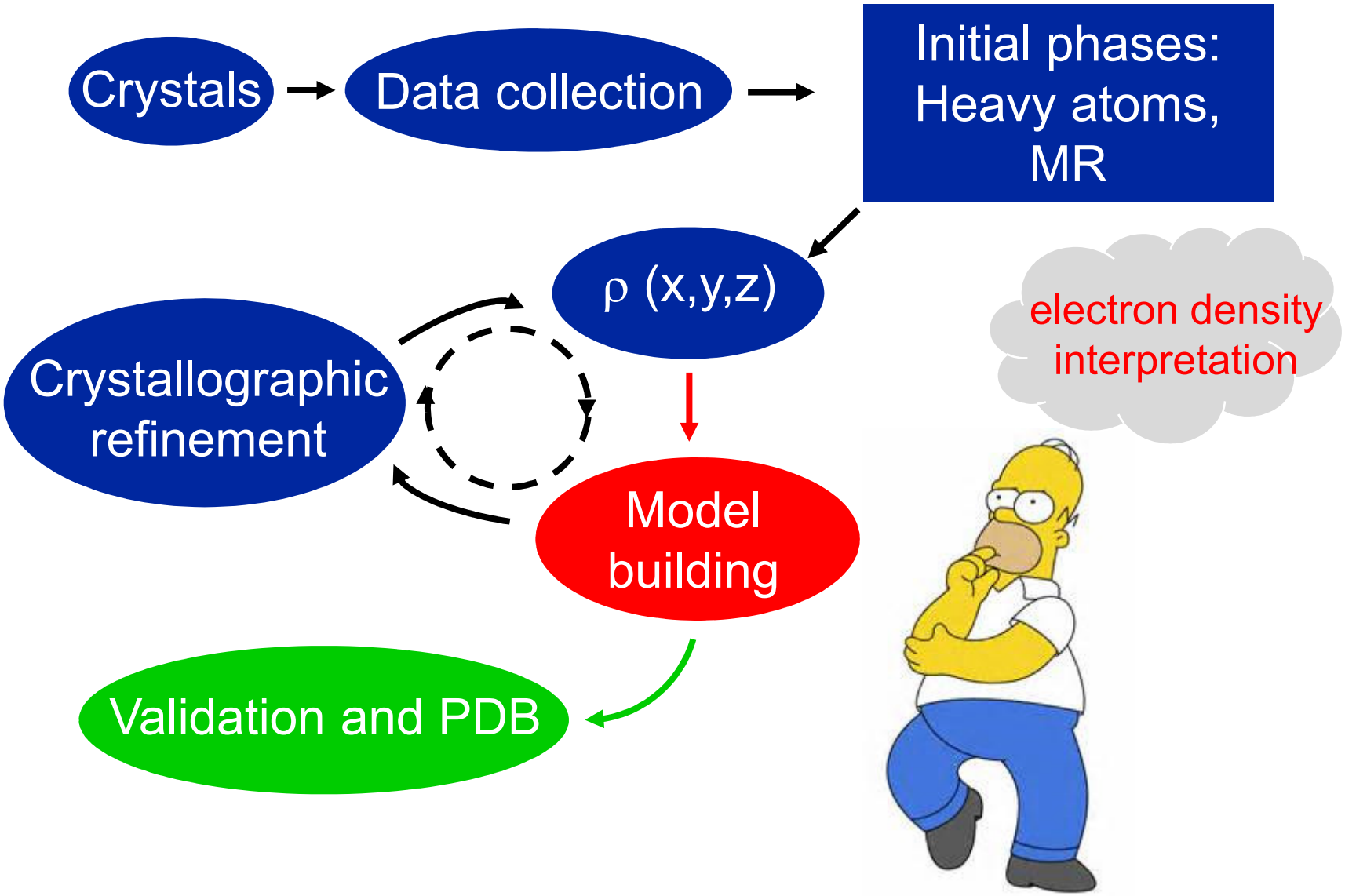
$2d_{hkl}\sin\theta = n\lambda$

$d_{hkl}^{min} = n\lambda / 2\sin\theta^{max}$

d_{hkl}^{min} = resolution of the diffraction data



high resolution (d_{hkl}^{min} small) ⇒ high hkl ⇒ high frequences in the Fourier synthesis ⇒ well interpretable electron density map


Crystals
Data collection
Initial phases: Heavy atoms, MR
ρ (x,y,z)
electron density interpretation
Crystallographic refinement
Model building
Validation and PDB

Protein Data Bank

file.pdb

```
CRYST1   39.550   74.570   66.560  90.00  99.94  90.00 P 1 21 1       4
ATOM      1  CB  SER     1       30.854  -6.329  36.118  1.00 41.46   6
ATOM      2  OG  SER     1       31.600  -7.531  36.190  1.00 44.54   8
ATOM      3  C   SER     1       30.991  -3.833  36.183  1.00 40.24   6
ATOM      4  O   SER     1       30.868  -2.883  36.961  1.00 40.08   8
ATOM      5  N   SER     1       31.848  -5.217  38.114  1.00 39.55   7
ATOM      6  CA  SER     1       31.668  -5.130  36.630  1.00 40.83   6
ATOM      7  N   THR     2       30.605  -3.796  34.906  1.00 39.92   7
ATOM      8  CA  THR     2       29.920  -2.658  34.286  1.00 38.97   6
ATOM      9  CB  THR     2       30.734  -2.067  33.086  1.00 39.67   6
ATOM     10  OG1 THR     2       31.045  -3.101  32.139  1.00 38.83   8
ATOM     11  CG2 THR     2       32.020  -1.403  33.566  1.00 38.83   6
ATOM     12  C   THR     2       28.542  -3.116  33.777  1.00 38.40   6
ATOM     13  O   THR     2       27.815  -2.341  33.141  1.00 38.79   8
.
.
.
ATOM   2166  HOH WAT   401       20.048   3.400  49.038  1.00 31.45   8
ATOM   2167  HOH WAT   402       -9.403   0.553  32.633  1.00 44.09   8
ATOM   2168  HOH WAT   403        3.928  -6.370  32.724  1.00 21.48   8
END
```

x y z occ B-factors

Protein Data Bank

file.pdb

PROTEIN DATA BANK

unit cell parameters

```
CRYST1   39.550   74.570   66.560  90.00  99.94  90.00 P 1 21 1         4
ATOM      1  CB  SER     1      30.854  -6.329  36.118  1.00 41.46      6
ATOM      2  OG  SER     1      31.600  -7.531  36.190  1.00 44.54      8
ATOM      3  C   SER     1      30.991  -3.833  36.183  1.00 40.24      6
ATOM      4  O   SER     1      30.868  -2.883  36.961  1.00 40.08      8
ATOM      5  N   SER     1      31.848  -5.217  38.114  1.00 39.55      7
ATOM      6  CA  SER     1      31.668  -5.130  36.630  1.00 40.83      6
ATOM      7  N   THR     2      30.605  -3.796  34.906  1.00 39.92      7
ATOM      8  CA  THR     2      29.920  -2.658  34.286  1.00 38.97      6
ATOM      9  CB  THR     2      30.734  -2.067  33.086  1.00 39.67      6
ATOM     10  OG1 THR     2      31.045  -3.101  32.139  1.00 38.83      8
ATOM     11  CG2 THR     2      32.020  -1.403  33.566  1.00 38.83      6
ATOM     12  C   THR     2      28.542  -3.116  33.777  1.00 38.40      6
ATOM     13  O   THR     2      27.815  -2.341  33.141  1.00 38.79      8
.
.
.
ATOM   2166  HOH WAT   401      20.048   3.400  49.038  1.00 31.45      8
ATOM   2167  HOH WAT   402      -9.403   0.553  32.633  1.00 44.09      8
ATOM   2168  HOH WAT   403       3.928  -6.370  32.724  1.00 21.48      8
END
```

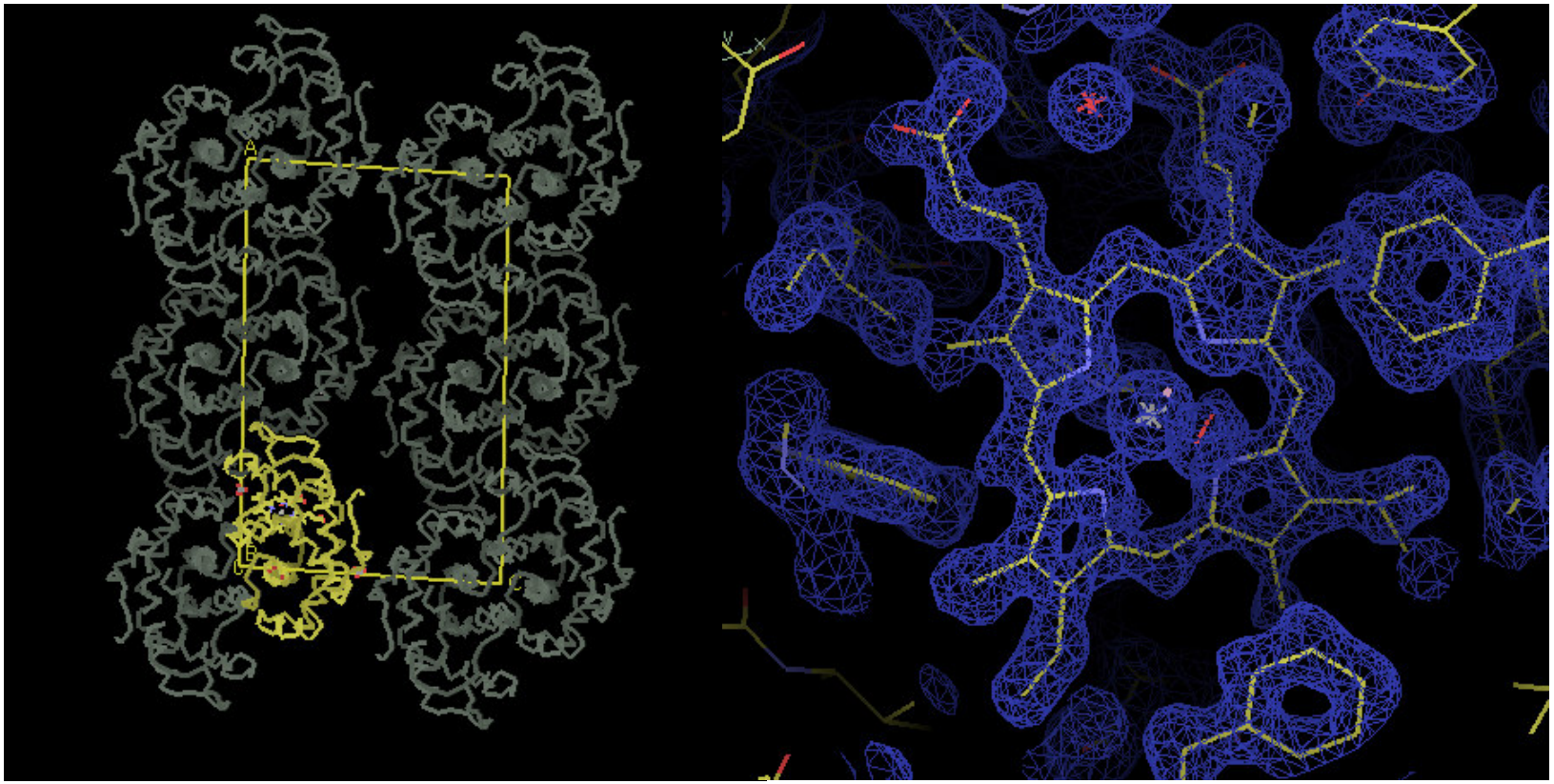
side chain

main chain

ligands

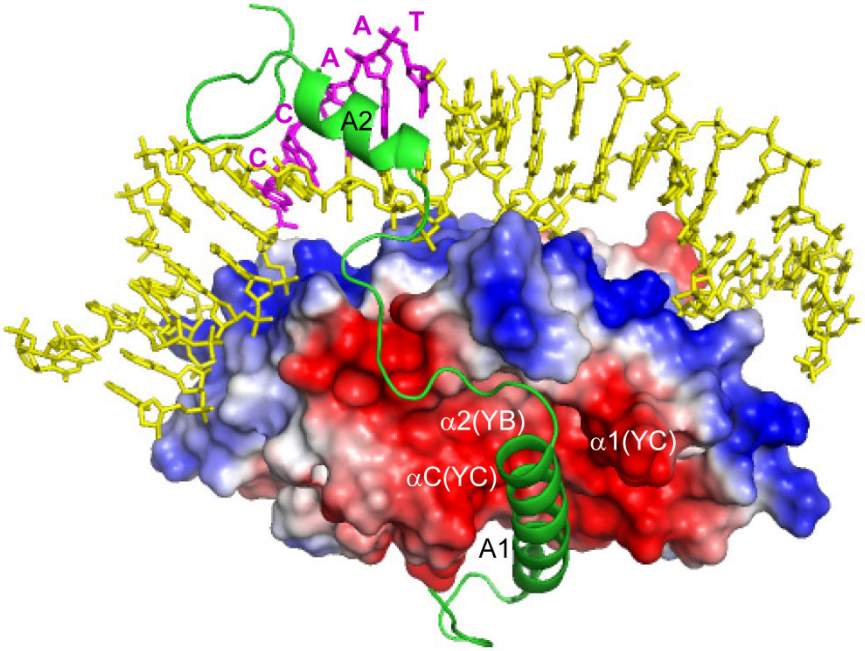
x y z occ B-factors

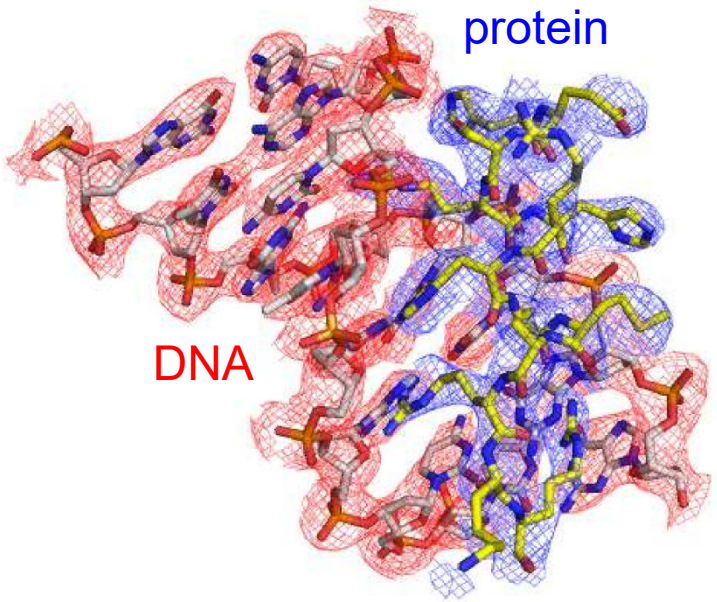
Example: X-ray structure of Protoglobin at 1.3 Å resolution



Nardini et al., EMBO Rep. 9, 157-163 (2008)

Example: X-ray structure of NF-Y/DNA at 3.1 Å resolution





Nardini et al., Cell 152, 132-143 (2013)