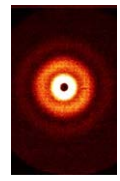
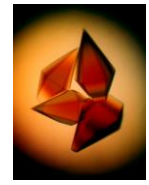
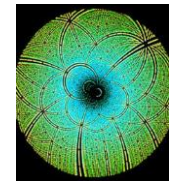
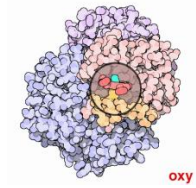
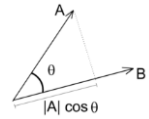







# Principles of BioX-ray (diffusion and diffraction)

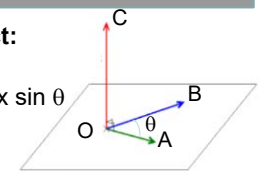
Patrice Gouet, Université de Lyon

## Bases in X-rays: mathematical tools

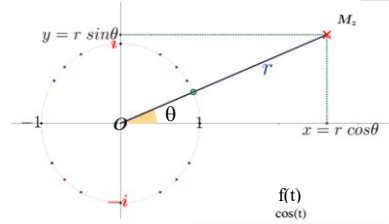
**Scalar product:**
$\overrightarrow{OA} . \overrightarrow{OB}$ = OA x OB x cos θ




**Vector product:**
$\overrightarrow{OC} = \overrightarrow{OA} \wedge \overrightarrow{OB}$
OC= OA x OB x sin θ

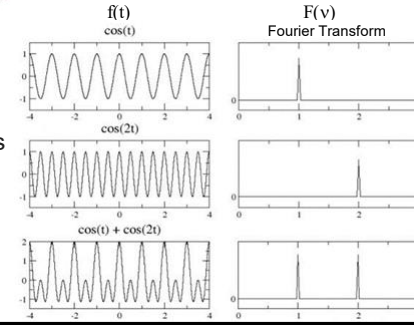



**Representing vectors with complex number:**
$\vec{r}$ = r cos θ + i r sin θ = r $e^{i\theta}$




**Fourier Transform :**
The **Fourier transform** (**FT**) decomposes a function of time into its constituent frequencies

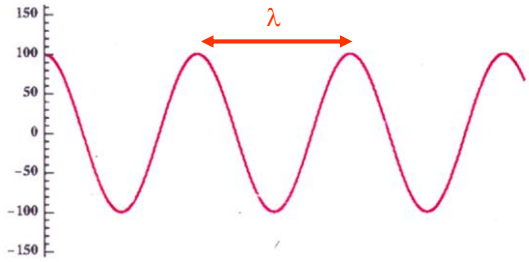
**F(v) =** $\int_{-\infty}^{\infty} \boldsymbol{f(t)} e^{-i2\pi vt} \boldsymbol{dt}$



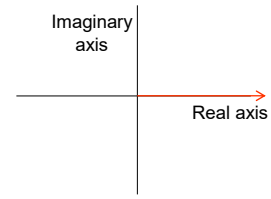
## Bases in diffraction: monochromatic waves

Energy: $E=hc/\lambda$ ; Equation: $F(x)=A\cos(2\pi x/\lambda+\phi)$ where A is the peak amplitude and $\phi$ the phase ; structure factor $\mathbf{F}=Ae^{i\phi}$ ; intensity $I \propto \mathbf{F}\mathbf{F}^*=A^2$

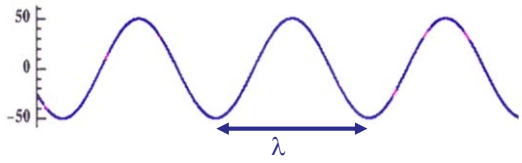
$F1=100\cos(2\pi x/\lambda+0)$



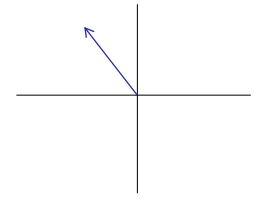

$\mathbf{F1}=100\,e^{i0}$




$F2=50\cos(2\pi x/\lambda+2\pi/3)$



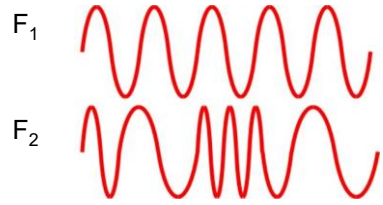

$\mathbf{F2}=50\,e^{i2\pi/3}$



## Bases in diffraction: adding waves

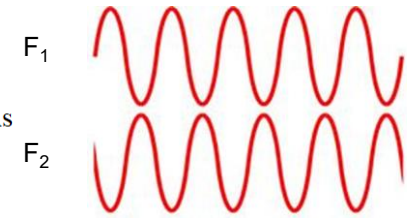
- **Incoherent waves**, intensities are added

$$I_{\text{total}} \propto |F_1|^2 + |F_2|^2 = I_1 + I_2$$



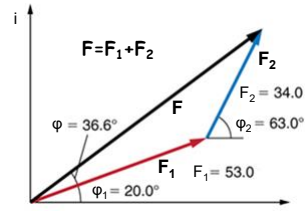

- **Coherent waves** (difference in phase remains constant) structure factors are added and **interferences** are observed

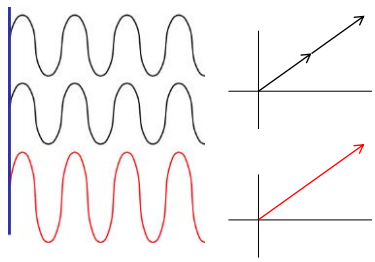
$$I_{\text{total}} \propto |\vec{F_1} + \vec{F_2}|^2 \propto |F_1|^2 + |F_2|^2 + \text{cross terms} \neq I_1 + I_2$$



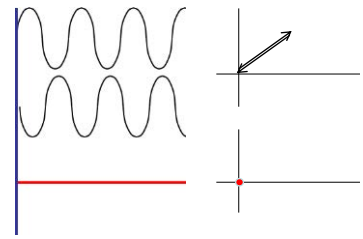
## Bases in diffraction: interferences

- Waves with same λ and spatially close to one other
- Structure factors are added $Fe^{i\varphi}= F_1e^{i\varphi 1}+ F_2e^{i\varphi 2}$ and interferences are observed



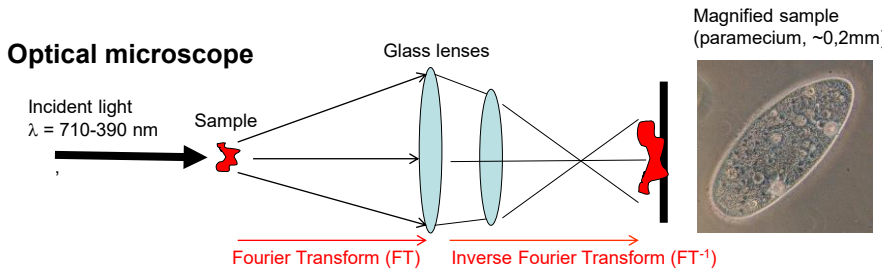
**Constructive interference** occurs when the phase difference between the waves is equal to **n 2π** and the path difference to **n λ**



**Destructive interference** occurs when the phase difference is equal to **π+ n 2π** and the path difference to **λ/2+n λ**

## Bases in diffraction: optical microscopy and X-ray diffraction

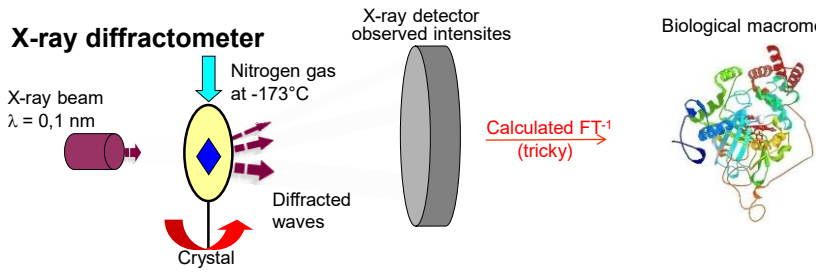
**Optical microscope**




Magnified sample (paramecium, ~0,2mm)



**X-ray diffractometer**



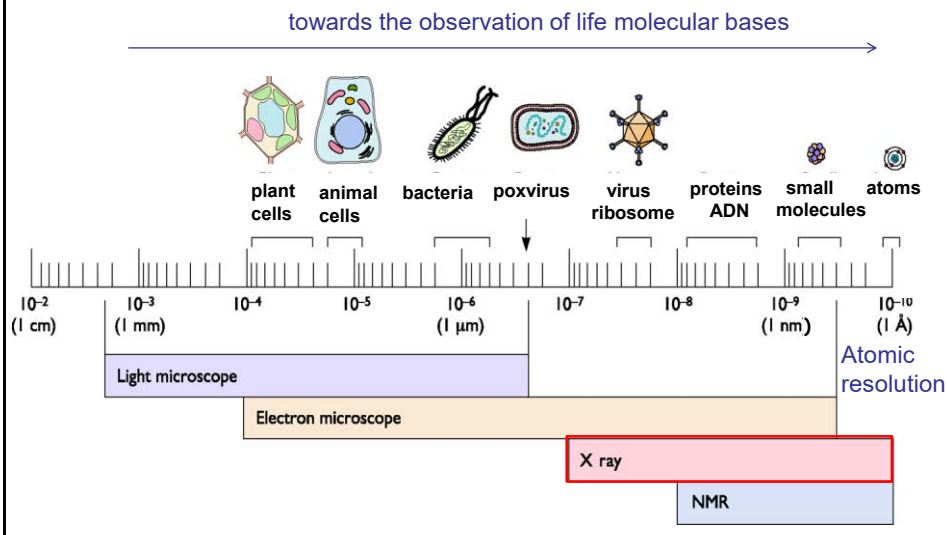

Biological macromolecule



Hard X-rays (λ< 2 Å) are high energy waves with a wavelength comparative to interatomic spacing. Once scattered, no physical way to change their trajectories to directly observe a magnified image

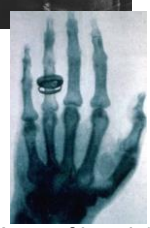
→ **Phasing problem** in SAXS and X-ray crystallography ($I_{obs}= A^2$ measured, φ lost)

## X-rays in Structural Biology



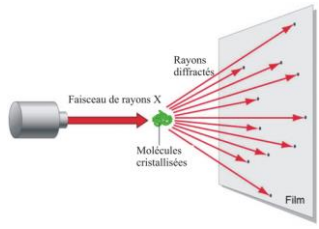

## Invention of X-ray crystallography

1901, Nobel Prize in physics for Wilhelm Röntgen for the **discovery of X-rays** and its use in **medicine**

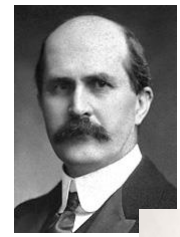


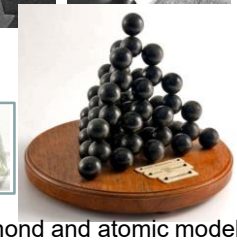
X-ray of hand,1896

1914, Nobel Prize in Physics for Max Von Laue for the discovery of **X-ray diffraction** by crystals




1915, Nobel Prize in Physics for Sir William Henry Bragg and his son, William Lawrence Bragg, for their work of **analyzing crystal structures** using X-rays



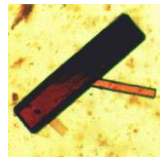


Diamond and atomic model of its structure,1913

## 1946-1962 Development of X-ray bio-crystallography

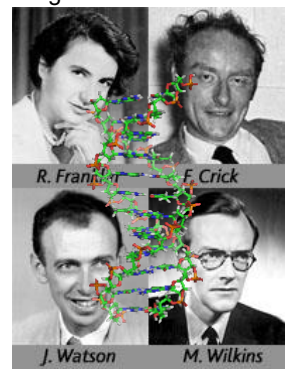
**1946,** Nobel Prize in Chemistry for the discovery of **protein crystallization**

*J. Summer, W. Stanley, J. Northrop*

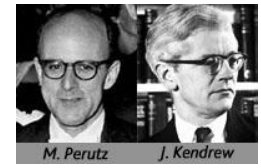


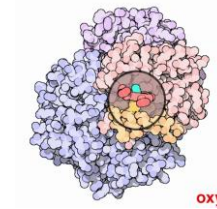
Protein crystal

**1962,** Nobel Prize in medicine for the molecular structure **of nucleic acids** and their function of transferring information in living matter



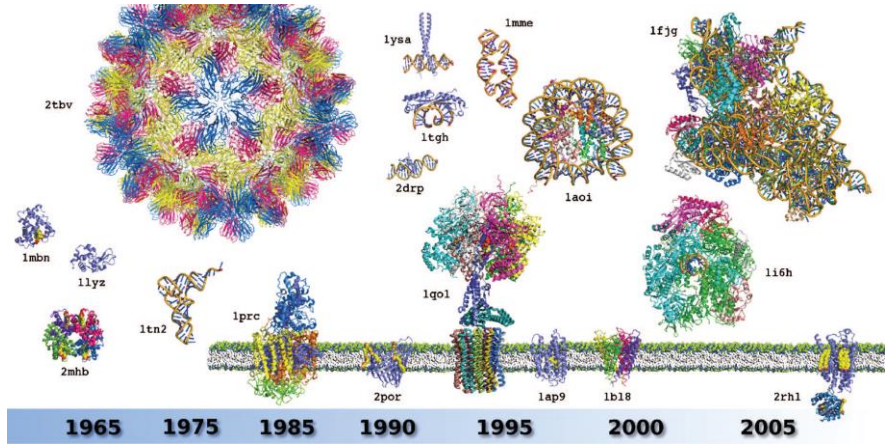

**1962,** Nobel Prize in chemistry for the structure of hemoglobin and globular proteins




→ The structures of DNA and hemoglobin are two essential discoveries that have used crystallography and X-ray diffraction.

## « Protein Data Bank » >150,000 atomic structures

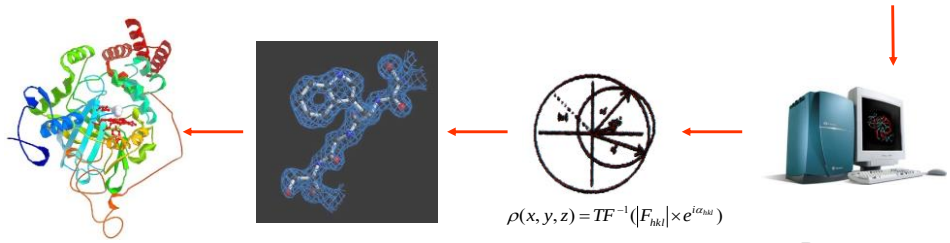



*Why water boils at 100ºC and methane at -161ºC, why blood is red and grass is green, why diamond is hard and wax is soft… The answers to all these problems have come from structural analysis.*

***Max Perutz**, July 1996, Churchill College, Cambridge*

## Structural studies by X-ray bio-crystallography

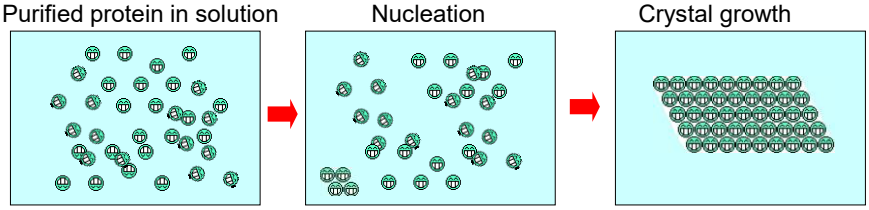
**1. Crystallogenesis** **2. Mounting crystal** **3. X-ray data collection**



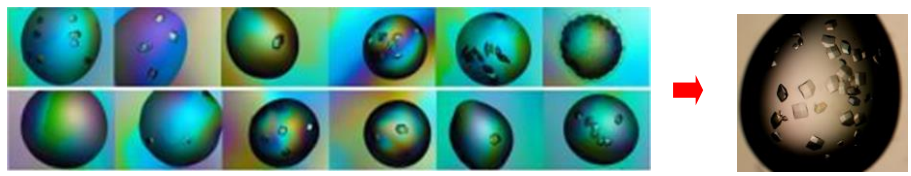

**7. Structure analysis** **6. Electron density maps and structural refinement** **5. Data phasing** **4. Data treatment**

## 1. Crystallogenesis

Purified protein in solution — Nucleation — Crystal growth



Crystals generally obtained by an evaporation method (**slow concentration**) in an aqueous medium that promotes crystallization (salt, pH ...)



Protein crystals are often **fragile** and **difficult** to obtain

Thousands of trials may be required to obtain crystals of sufficient size 1mm to 10µm that grow in 1h to 1year

## 2. Using a synchrotron light source for X-rays

**Principle**:

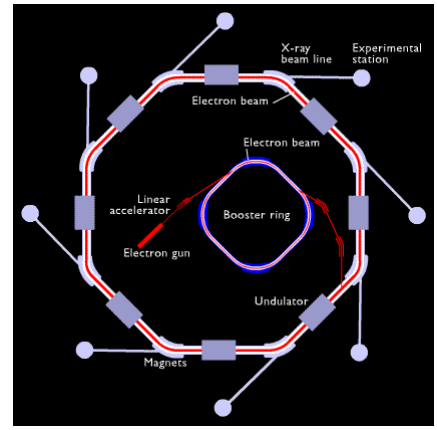
- Electrons are accelerated to high speeds to achieve a final energy that is typically in the gigaelectronvolt range

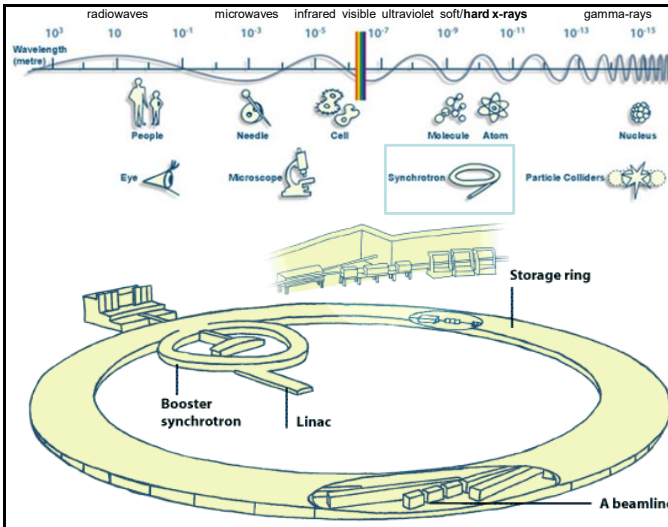
- Electromagnets inserted along the storage ring bend their trajectory

- Electromagnetic radiation (X-rays) emitted during the change of trajectory

- A wavelength of 0.1 nm = 1 Å is often selected in biocrystallography (monochromator)

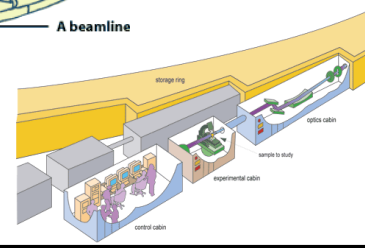
**Examples**: ESRF-Grenoble, SOLEIL-Saclay



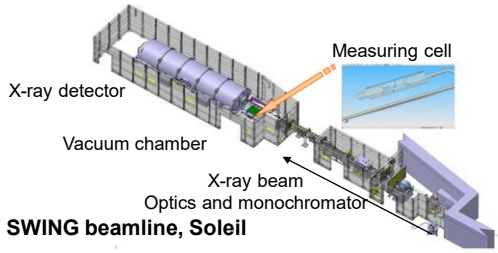




| | |
|---|---|
| **Circumference of the ring** | 844 m |
| **Energy** | 6 Gev |
| **Number of hutchs** | 40 Beamlines |

ESRF Grenoble



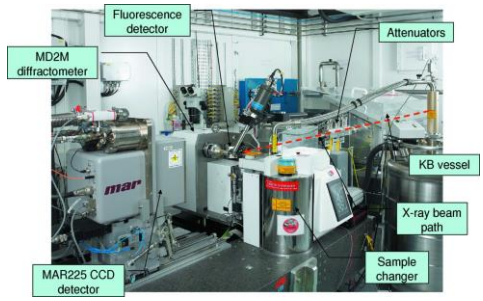
## 2. Two types of beamline to study biological macromolecules




**SWING beamline, Soleil**

**SAXS (Small angle X-ray Scattering) beamline for macromolecules in solution**

→ Talk on Sunday, TP on Thursday

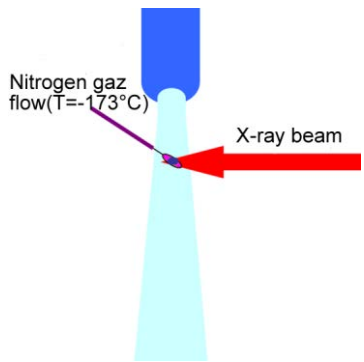



**MX beamline, ESRF**

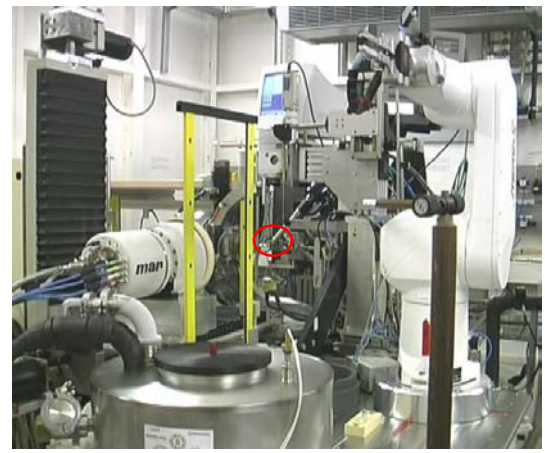
**Macromolecular crystallography beamline**

→ Talks and TP on Sunday

## 2. Crystal fishing and mounting on a goniometer head

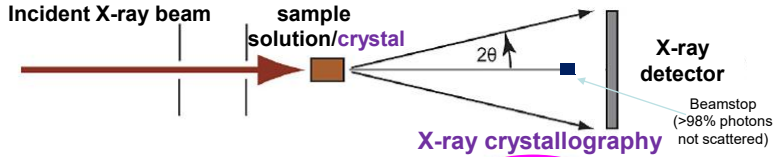



Crystal **fishing** with a cryo-loop and **freezing** to preserve the crystal (dehydration, radiation damage)



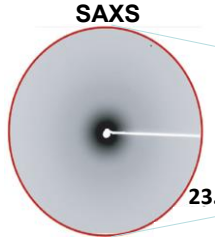
Automated crystal mounting at the ESRF beamline **FIP BM30A**, Grenoble

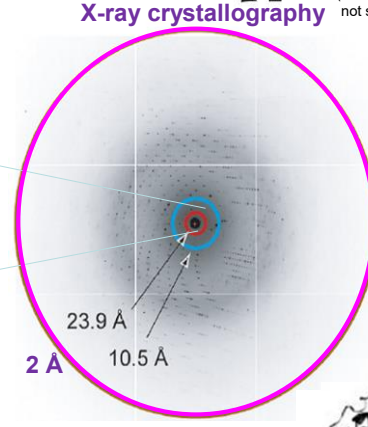
## 3. Small Angle X-ray Scattering *vs* X-ray crystallography



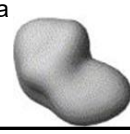

X-ray crystallography

SAXS




Observed intensities I(2θ)
→ *ab initio* calculation of a molecular envelope

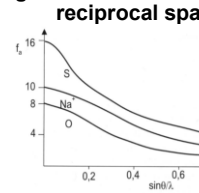


I(hkl) observed to high resolution
→ data phasing and atomic structure refinement

**elastic X-ray scattering**

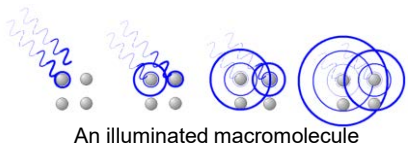
**real space** → **reciprocal space**

An illuminated atom →

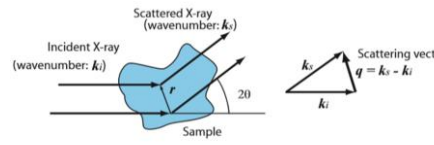



FT of electron density

diffusion factor
= $Z_e$ at θ = 0



An illuminated macromolecule →




$\rho(r)$: electron density $q = |q| = 4\pi \sin\theta/\lambda$

$$\mathbf{F(q)} = \int_V \rho(r) \exp(-i\mathbf{q}\cdot\mathbf{r}) dr \qquad I(q) = |F(q)|^2$$

FT of electron density

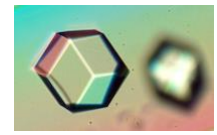
**SAXS**: N distant particles in solution adopting all orientations / X-ray beam. →

Incoherent waves from particles
$I(q) = \mathbf{N} \langle I(q) \rangle_\Omega$

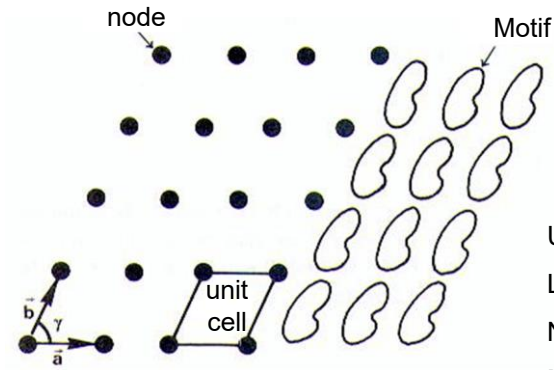
**X-ray crystallography**: Periodic assembly of N unit cells containing N' particles each →

Coherent waves from unit cells
$I(h,k,l) = \mathbf{N^2}\, I_{N'}(h,k,l)$
hkl indicate directions of constructive interferences

## 3. The case of X-ray protein crystallography



- A **crystal structure** describes the ordered arrangement of molecules in a crystalline material
- The smallest group of particles in the material that constitutes the pattern repeated by translation is the **unit cell**




Crystal lattice

Unit cell = «maille»

Lattice = «réseau»

Node = «nœud»

Motif: can be >1 in the unit cell

## Convolution

Convolution: take one function, f(**r**), and put it down at every point of a second function, g(**r**)

$$f(\mathbf{r}) \otimes g(\mathbf{r}).$$

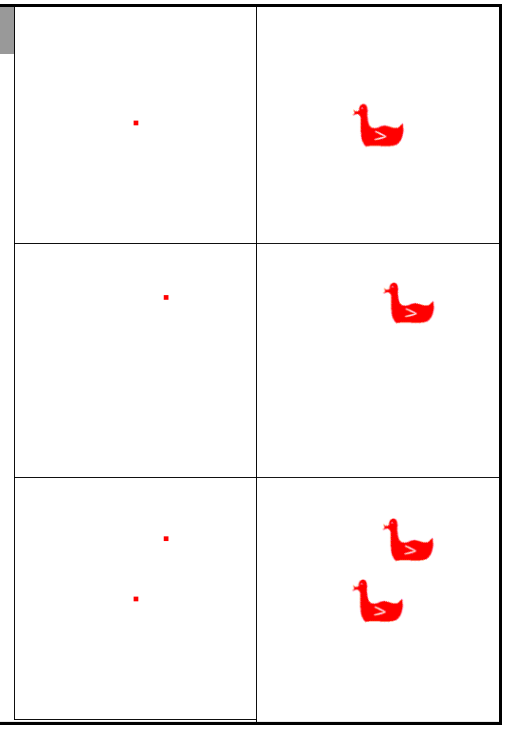
Here ⊗ is the convolution operator.

A crystal is a convolution of one function (a motif) with another (a lattice)

**1.** Motif is any object; *e.g.* a protein molecule, duck etc.

**2.** Lattice is an array of regularly spaced mathematical points

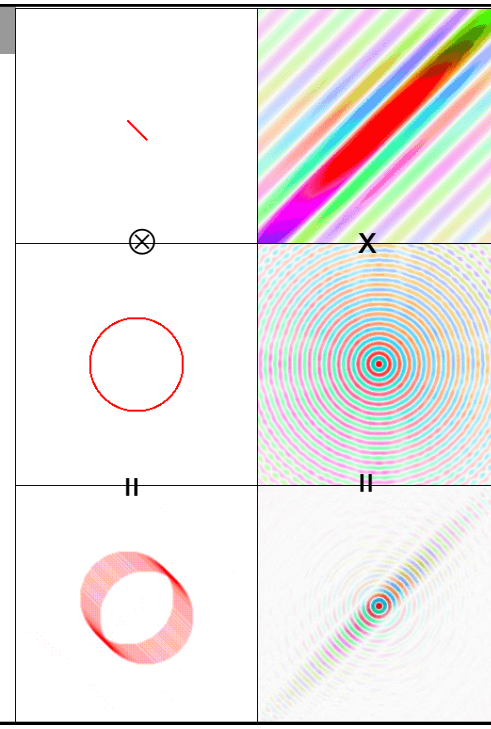
**Lattice ⊗ Motif = Crystal**



## Convolution Theorem

The Fourier transform of one function **convoluted** with another is the same as the Fourier transform of the first **multiplied** by the Fourier transform of the second

**FT** [f(**r**) ⊗ g(**r**)] = **FT**(f(**r**)) x **FT**(g(**r**))




## Diffusion and diffraction

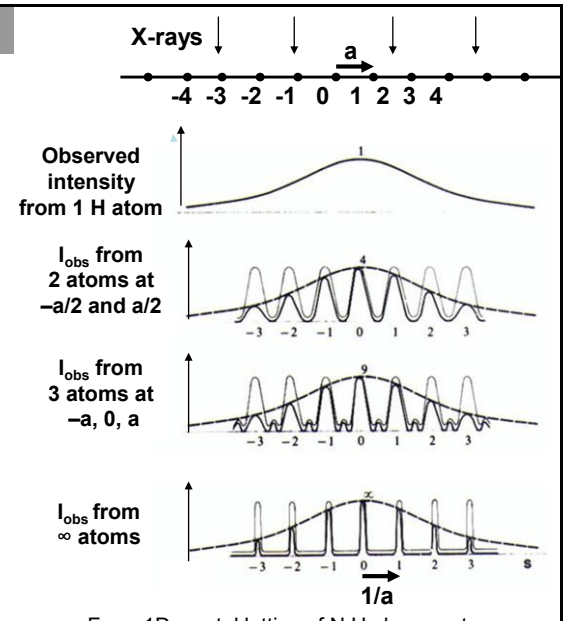
**Fourier Transform of a Crystal**

**FT** (motif): a continuous function with no sharp discontinuities

**FT** (lattice): another lattice with spacings which are reciprocal of those of the original lattice

FT(motif ⊗ lattice) = FT(motif) x FT (lattice)

**from Structural Biology, J. Janin and M. Delepierre**

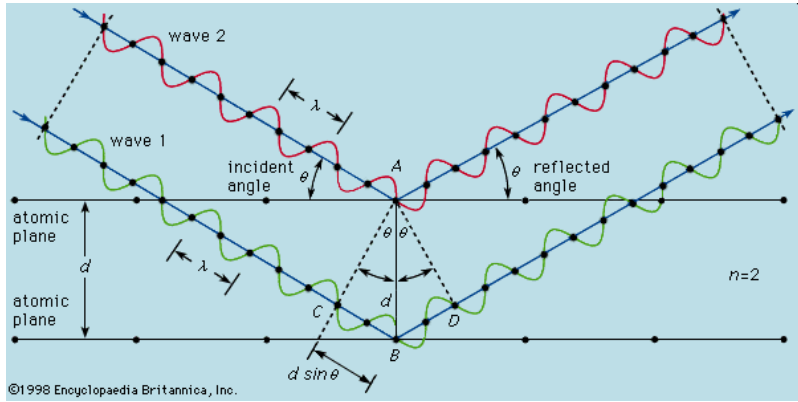



For a 1D crystal lattice of N Hydrogen atoms

$I_{obs}(h) = N^2 I_{obs\,1H'}(h)$

where h indicate directions of constructive interferences

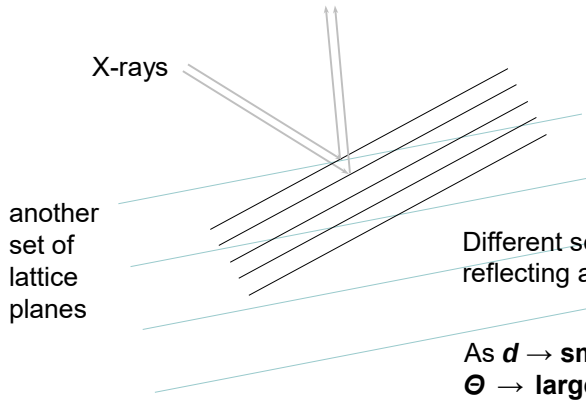
## The Bragg's law




- For a crystal, waves are scattered from atomic planes separated by *d*
- The path difference between two waves undergoing interference is given by 2*d*sin *θ*, where *θ* is the scattering angle.
- The effect of constructive or destructive interference intensifies in successive crystallographic planes
- This leads to Bragg's law, which describes the condition on *θ* for the **constructive interference** to be at its **strongest**: **2*d* sinθ = n λ**

## The notion of resolution in X-ray diffraction

**Braggs' law** **λ = 2*d* sin θ**




Waves scattered at a **high angle** contain **high resolution** details on the illuminated structure

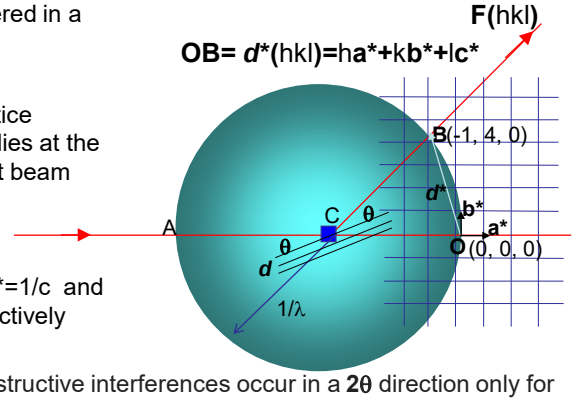
## The Ewald construction

Let's imagine that a crystal centered in a sphere of radius 1/λ

The origin **O** of the reciprocal lattice defined by the vectors **a*, b*, c*** lies at the intersection between the incident beam and the sphere surface

**a*=b∧c/V ; b*=c∧a/V ; c*=a∧b/V**

If α= β= γ=90°, a*=1/a, b*=1/b, c*=1/c and **a**, **b**, **c** and **a***, **b***, **c*** are // respectively



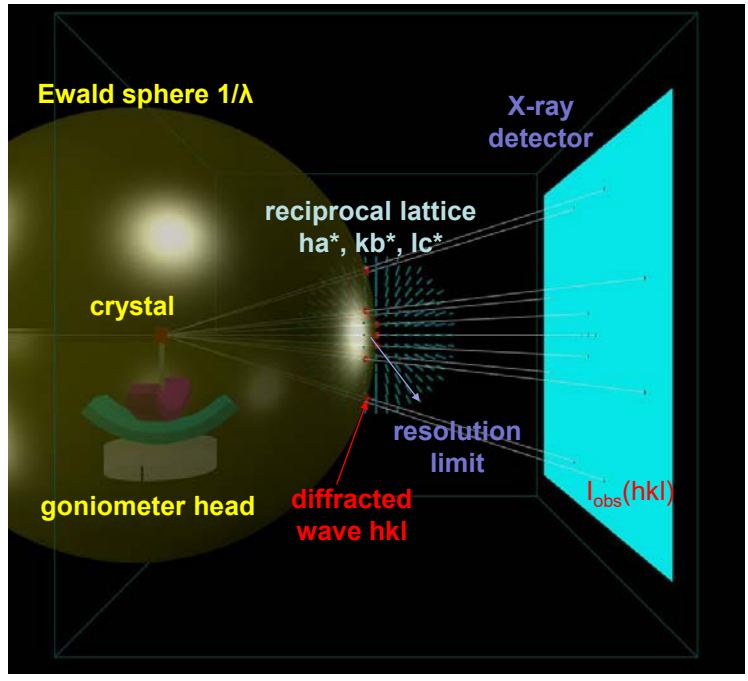

With respect to Bragg's law, constructive interferences occur in a **2θ** direction only for reciprocal lattice points ***d**(hkl)** that intersects with the Ewald sphere.

In magnitude *d**(hkl) = 1/*d*(hkl) where *d* is the distance between the reflecting planes

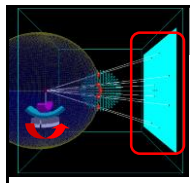
The structure factor **F(**hkl**)** of the diffracted wave is equal to the **Fourier Transform** of the crystal assembly:

$$F_{hkl} \exp i\phi_{hkl} = \sum_{j=1}^{M} f_j [\exp i2\pi(lx_j + ky_j + lz_j)]$$

where M is the number of atoms in the unit cell



## 3. Data collection by oscillation steps

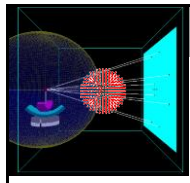


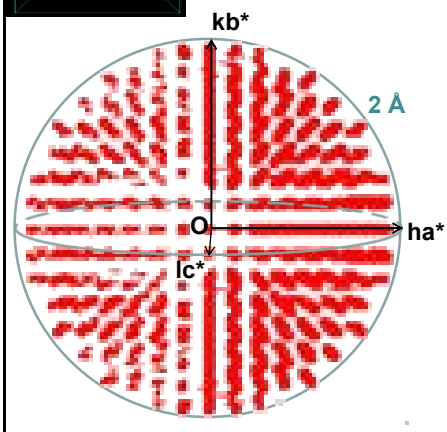



0° → 1° `lyso_001.mar2000` 1° → 2° `lyso_002.mar2000` 2° → 3° `lyso_003.mar2000` ……..179°→ 180°

Data collected by **oscillation steps** and **detector erased** after each oscillation to avoid superposed Bragg's spots. Recent digitized X-ray detectors with short readout time (1ms) allow fast data collection with small oscillation (0.1°)

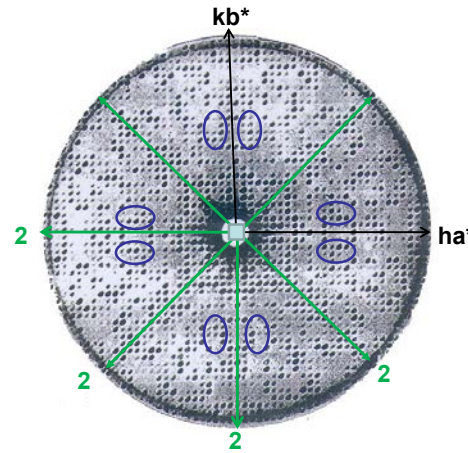
## Data processing (more explanations on Sunday)






~100% of the $I_{obs}$(hkl) collected in total (0°-1°, 1°-2°, …,179°-180°) to the resolution limit (2 Å)

**Symmetries of I(hkl) respect the Laue class of the crystal**



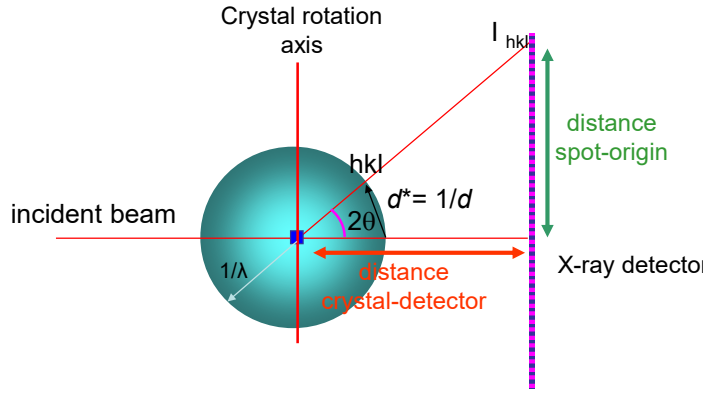

**hk0 plane**

Crystal in $P4_32_12$ X-ray I(hkl) in 422

## Resolution circles

The resolution of each diffracted wave I(hkl) can be estimated by the Bragg relation : $d = \lambda / 2 \sin\theta$ and ***d* = Resolution = *1/d****

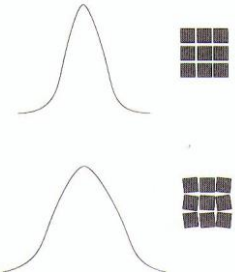
Resolution ≈ λ(distance crystal-detector)/(distance spot-origin)

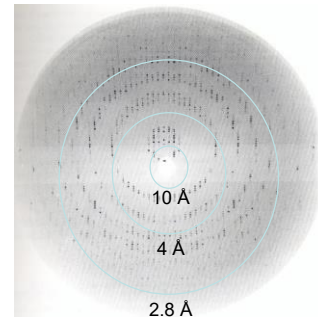



Before data collection, the user **optimizes** the **crystal-detector distance** in order to collect the best possible dataset (well separated spots) to the highest possible resolution (weakest Bragg spots on detector edges)

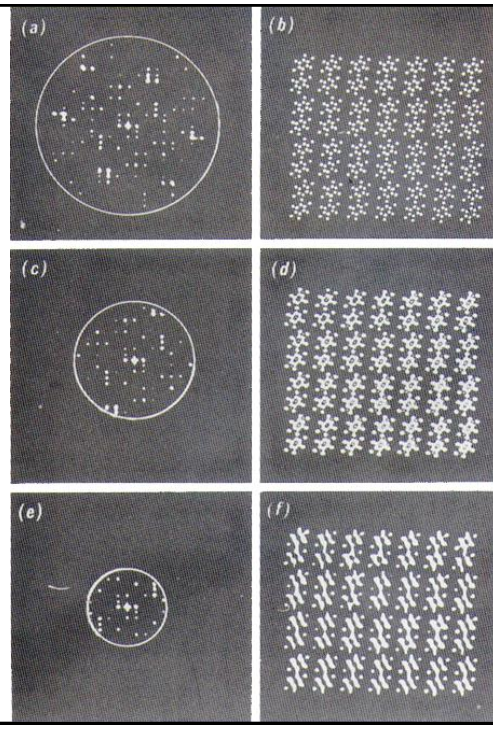
## Resolution and mosaicity

- Macromolecular crystals are imperfect and can be described as a **mosaic** of small blocks.
- The effect on the diffraction pattern is to broaden the diffraction spot profiles and **mosaicity** is defined as the full width at half maximum of diffraction peaks
- A consequence of this mosaicity is the rapid weakening of Bragg spots at high 2θ.
- A low mosaicity crystal generally diffracts to high resolution.
- For macromolecular crystals:
  resolution limit > 3.5Å, is low
  3.5Å > resolution limit > 1.5Å, is medium
  1.5Å > resolution, is high





## Resolution effect

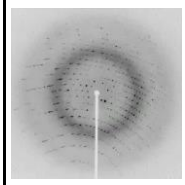



Effect of the shape of a lattice on the appearance of the transform and *vice versa*

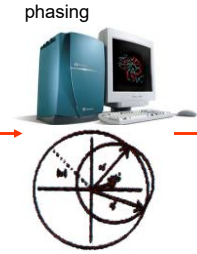
**From Protein Crystallography, T.L. Blundell and L.N. Johnson**

## 4. Data phasing, 5. Electron density map, 6. Building 3D-structure and crystallographic refinement

Bragg's spots $I_{obs}$

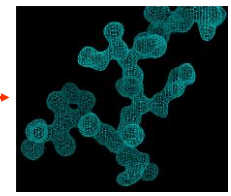


Data treatment and phasing

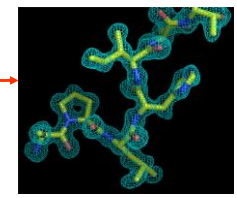


$$\rho(x,y,z) = TF^{-1}(|F_{hkl}| \times e^{i\alpha_{hkl}})$$

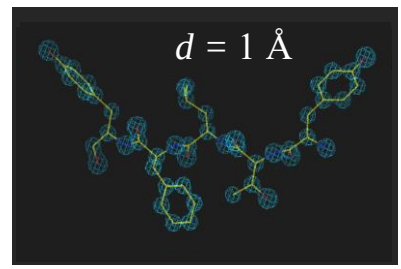
Resulting electon density map



Structure Building (C yellow, N blue, O red)



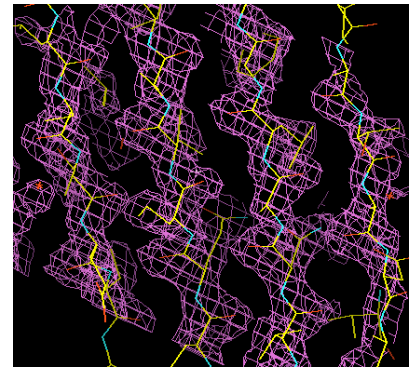
- The **resolution** of the electron density map depends on the signal-to-noise ratio of each observed intensity $I_{obs}$(hkl)
- So the brightness of the **X-ray source**, the size and the periodicity of the **crystal** and the sensitivity of the X-ray **detector**
- Better is the resolution of the electron density map, more accurate is the atomic model

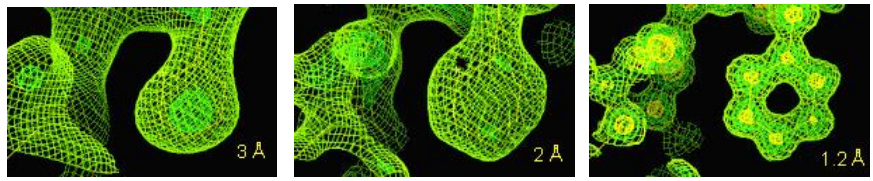



Chaîne de 5 résidus d'acides aminés

## Resolution of electron density maps in biocrystallography

Medium resolution 3Å data is good enough to see the backbone with space in between.





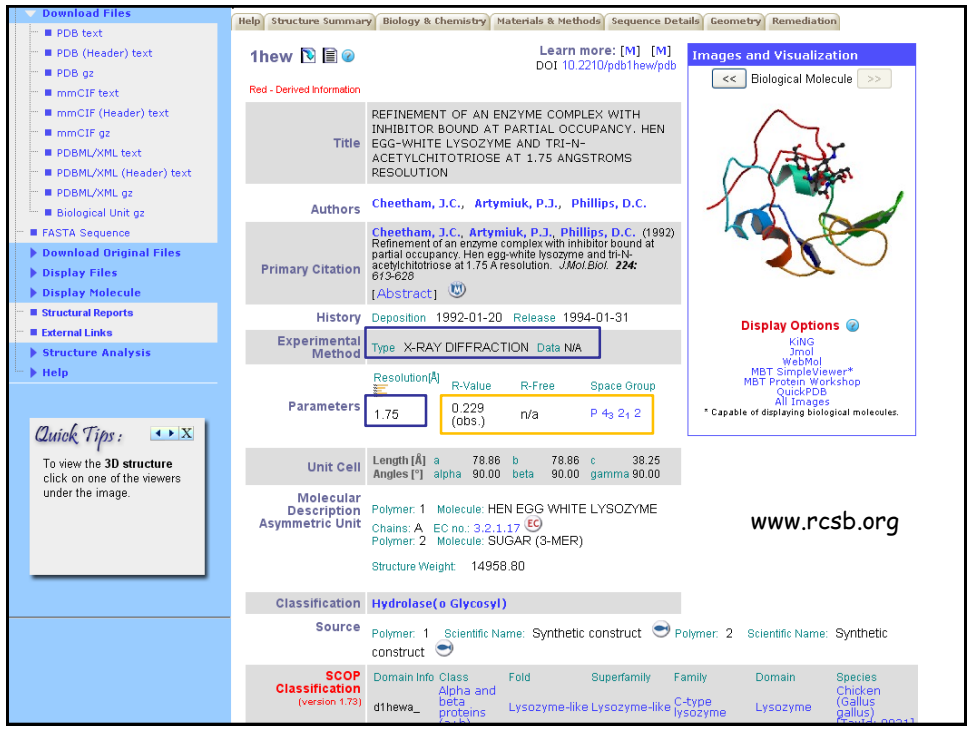

Electron density of a phenylalanine residue in a protein crystal structure determined from 3 Å to 1.2 Å resolution (error in position from 0.3 to 0.1 Å)

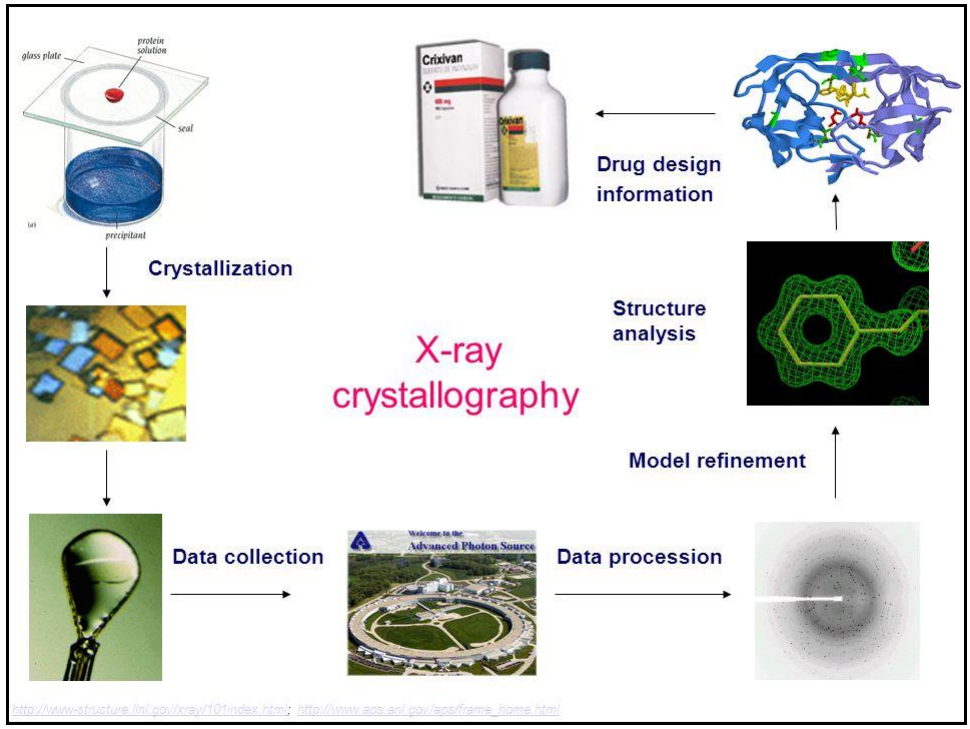
```
HEADER    HYDROLASE(O-GLYCOSYL)                   20-JAN-92   1HEW
COMPND    2 MOLECULE: HEN EGG WHITE LYSOZYME;
JRNL        AUTH   J.C.CHEETHAM,P.J.ARTYMIUK,D.C.PHILLIPS
REMARK    2 RESOLUTION.    1.75 ANGSTROMS.
.....................................................................
DBREF  1HEW A    1   129  UNP    P00698   LYC_CHICK       19    147
SEQRES   1 A  129  LYS VAL PHE GLY ARG CYS GLU LEU ALA ALA ALA MET LYS
SEQRES   2 A  129  ARG HIS GLY LEU ASP ASN TYR ARG GLY TYR SER LEU GLY
.....................................................................
HET    NAG  A 201      15
HET    NAG  A 202      14
HET    NAG  A 203      14
HETNAM     NAG N-ACETYL-D-GLUCOSAMINE
CRYST1   78.860   78.860   38.250  90.00  90.00  90.00 P 43 21 2      8
ATOM      1  N   LYS A   1       3.398   9.981  10.408  1.00 30.48           N
ATOM      2  CA  LYS A   1       2.459  10.365   9.364  1.00 28.03           C
ATOM      3  C   LYS A   1       2.458  11.880   9.149  1.00 21.93           C
ATOM      4  O   LYS A   1       2.481  12.672  10.100  1.00 14.10           O
ATOM      5  CB  LYS A   1       1.026   9.935   9.695  1.00 30.54           C
ATOM      6  CG  LYS A   1       0.028  10.169   8.558  1.00 37.93           C
ATOM      7  CD  LYS A   1      -1.415  10.089   9.048  1.00 33.23           C
ATOM      8  CE  LYS A   1      -2.357  10.822   8.082  1.00 32.17           C
ATOM      9  NZ  LYS A   1      -3.661  10.090   8.025  1.00 31.92           N
ATOM     10  N   VAL A   2       2.429  12.232   7.880  1.00 17.30           N
ATOM     11  CA  VAL A   2       2.395  13.653   7.465  1.00 14.47           C
.....................................................................
ATOM   1000  CD2 LEU A 129     -13.441  19.891   8.982  1.00 29.73           C
ATOM   1001  OXT LEU A 129     -17.993  19.662   8.407  1.00 31.81           O
TER    1002      LEU A 129
HETATM 1003  C1  NAG A 201       5.991  25.237  25.980  1.00 32.10           C
HETATM 1004  C2  NAG A 201       4.850  24.302  26.455  1.00 29.05           C
HETATM 1005  C3  NAG A 201       4.046  24.991  27.538  1.00 14.31           C
HETATM 1006  C4  NAG A 201       5.038  25.548  28.618  1.00 41.63           C
.....................................................................
HETATM 1046  O   HOH A 204     -16.295  29.471   0.511  1.00 18.64           O
HETATM 1047  O   HOH A 205      -1.660  14.995   1.659  1.00 45.86           O
END
```

positions X, Y, Z in Å (from X-rays, NMR or cryoEM experiments)

occupancy

temperature factor in $Å^2$ $= 8\pi^2 \langle U^2 \rangle$ (can be refined from X-rays if medium-high resolution)


Download Files
PDB text
PDB (Header) text
PDB gz
mmCIF text
mmCIF (Header) text
mmCIF gz
PDBML/XML text
PDBML/XML (Header) text
PDBML/XML gz
Biological Unit gz
FASTA Sequence
Download Original Files
Display Files
Display Molecule
Structural Reports
External Links
Structure Analysis
Help
Quick Tips :
To view the 3D structure click on one of the viewers under the image.
Help
Structure Summary
Biology & Chemistry
Materials & Methods
Sequence Details
Geometry
Remediation
1hew
Learn more: [M] [M]
DOI 10.2210/pdb1hew/pdb
Red - Derived Information
Title
REFINEMENT OF AN ENZYME COMPLEX WITH INHIBITOR BOUND AT PARTIAL OCCUPANCY. HEN EGG-WHITE LYSOZYME AND TRI-N-ACETYLCHITOTRIOSE AT 1.75 ANGSTROMS RESOLUTION
Authors
Cheetham, J.C., Artymiuk, P.J., Phillips, D.C.
Primary Citation
Cheetham, J.C., Artymiuk, P.J., Phillips, D.C. (1992) Refinement of an enzyme complex with inhibitor bound at partial occupancy. Hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 A resolution. J.Mol.Biol. 224: 613-628
[Abstract]
History
Deposition 1992-01-20 Release 1994-01-31
Experimental Method
Type X-RAY DIFFRACTION Data N/A
Parameters
Resolution(Å)
R-Value
R-Free
Space Group
1.75
0.229 (obs.)
n/a
P 43 21 2
Unit Cell
Length [Å] a 78.86 b 78.86 c 38.25
Angles [°] alpha 90.00 beta 90.00 gamma 90.00
Molecular Description Asymmetric Unit
Polymer: 1 Molecule: HEN EGG WHITE LYSOZYME
Chains: A EC no.: 3.2.1.17
Polymer: 2 Molecule: SUGAR (3-MER)
Structure Weight: 14958.80
Classification
Hydrolase(o Glycosyl)
Source
Polymer: 1 Scientific Name: Synthetic construct
Polymer: 2 Scientific Name: Synthetic construct
SCOP Classification (version 1.73)
Domain Info
Class
Fold
Superfamily
Family
Domain
Species
d1hewa_
Alpha and beta proteins
Lysozyme-like
Lysozyme-like
C-type lysozyme
Lysozyme
Chicken (Gallus gallus)
Images and Visualization
<< Biological Molecule >>
Display Options
KiNG
Jmol
WebMol
MBT SimpleViewer*
MBT Protein Workshop
QuickPDB
All Images
* Capable of displaying biological molecules.
www.rcsb.org


glass plate
protein solution
seal
precipitant
Crystallization
Crixivan
Drug design information
Structure analysis
X-ray crystallography
Model refinement
Data collection
Welcome to the Advanced Photon Source
Data procession
http://www-structure.llnl.gov/xray/101index.html; http://www.aps.anl.gov/aps/frame_home.html