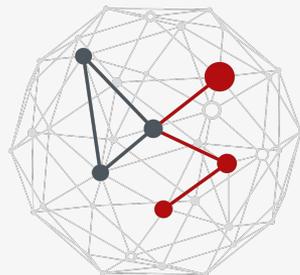


1222-2022  
800  
ANNI



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

 DIPARTIMENTO  
MATEMATICA



**DATA SCIENCE**  
UNIVERSITY OF PADOVA

# STRUCTURAL DATA

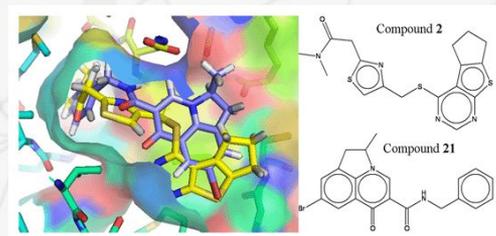
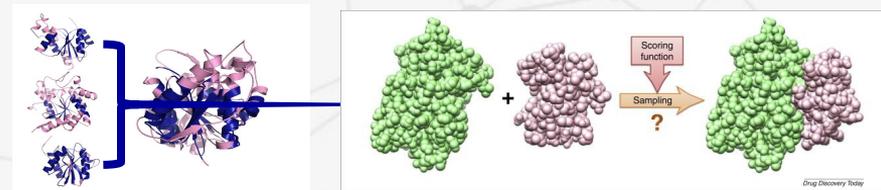
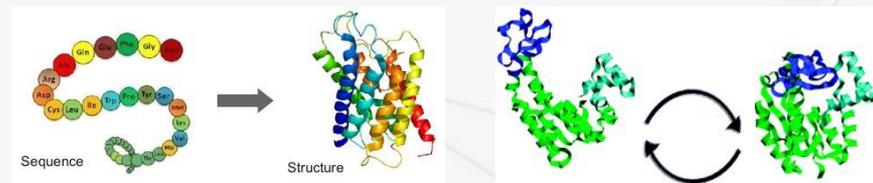
Master of Science in Data Science

Damiano Piovesan

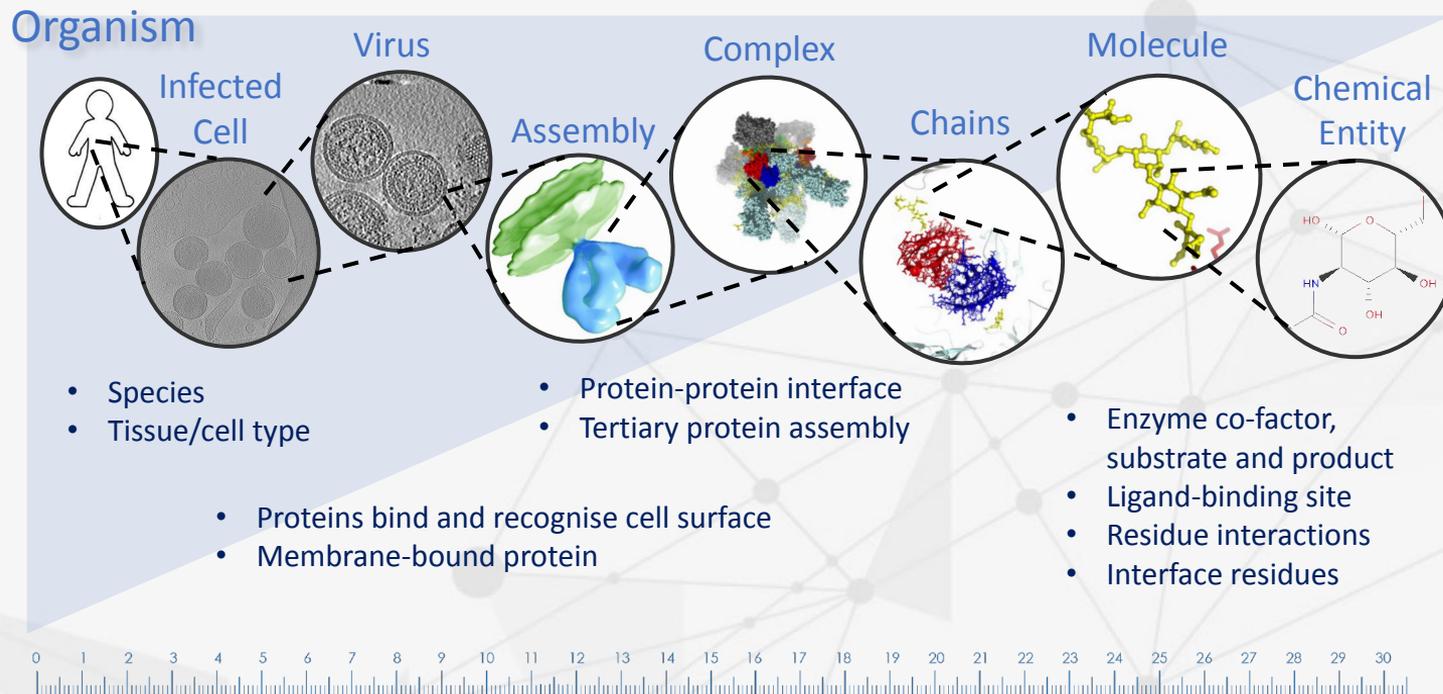


# Uses of structural data

- Structure prediction
- Structure comparison and evolutionary classification
- Prediction of protein interactions and assemblies
- Protein dynamics
- Mechanism of interaction with ligands
  - Metabolites
  - Drug compounds
  - DNA



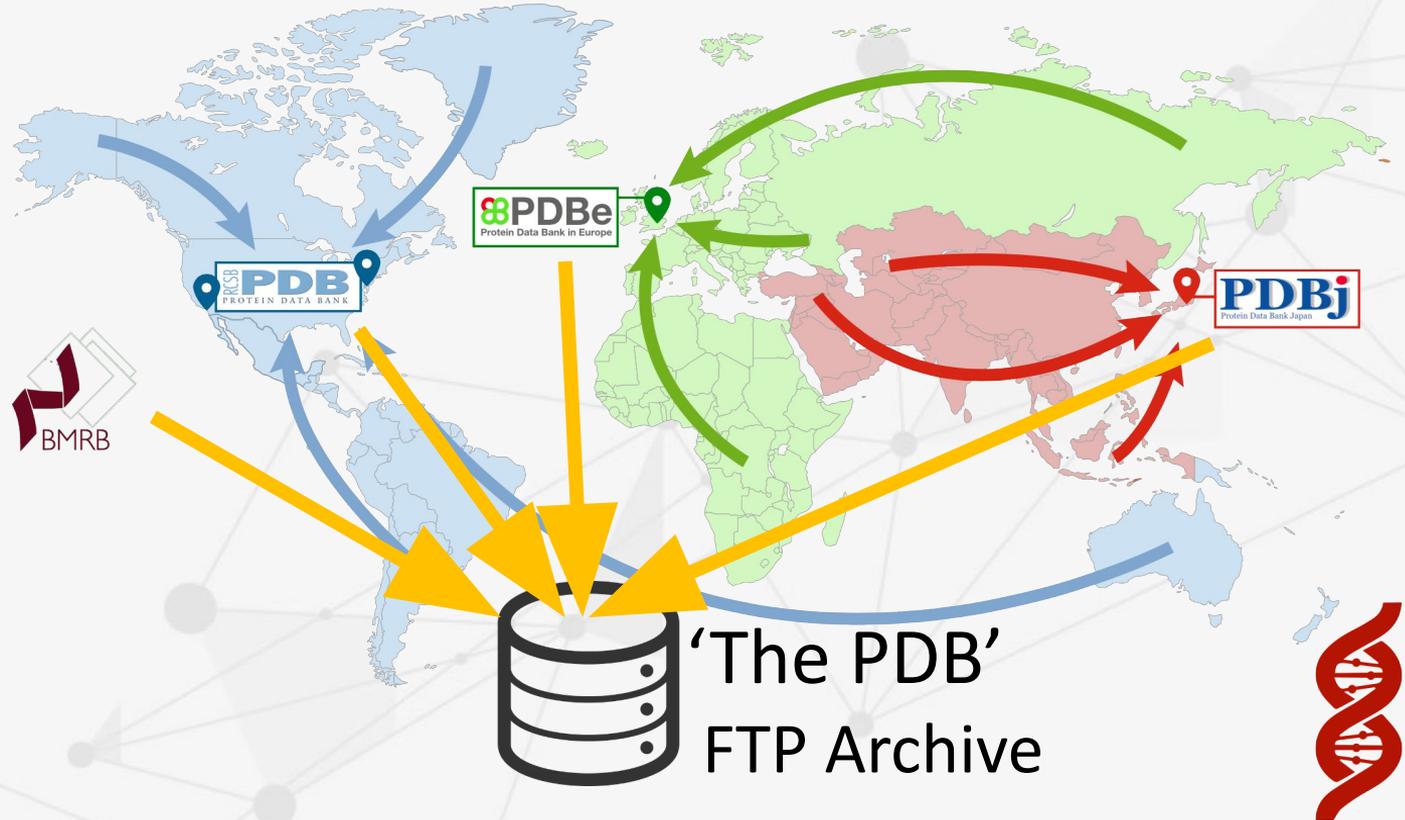
# (Macro)molecular structure data



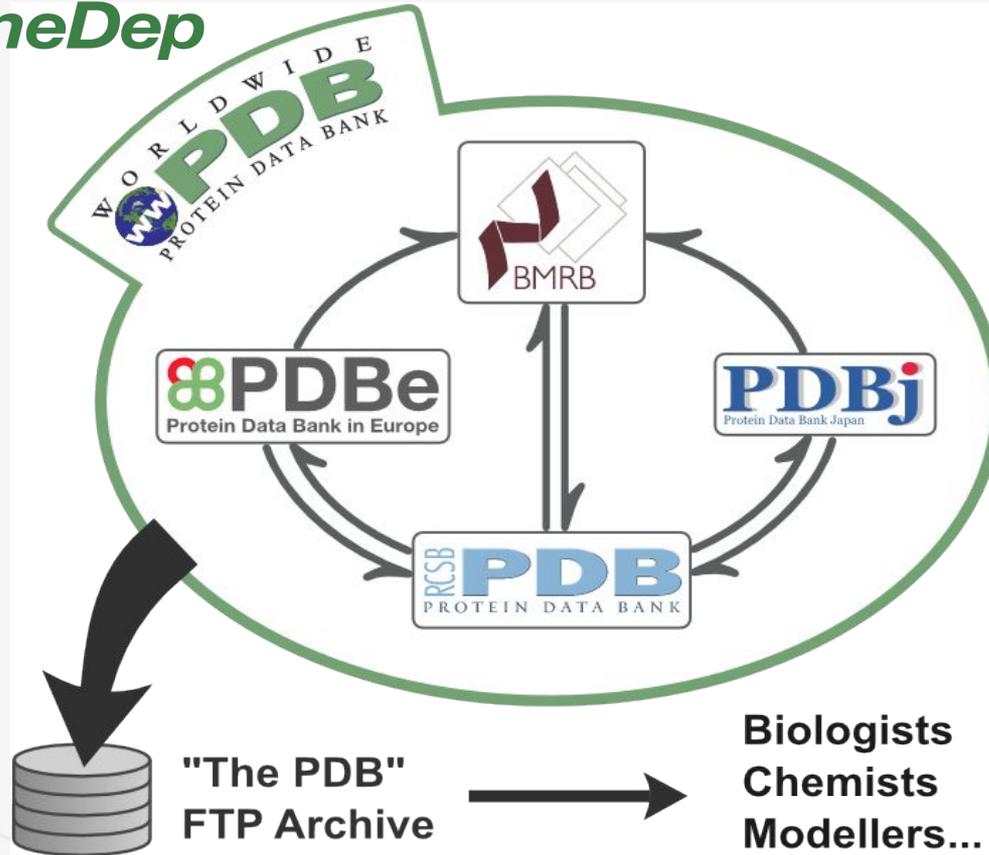
# The Protein Data Bank (PDB)

The PDB website  
(US repository)

<https://www.rcsb.org/>

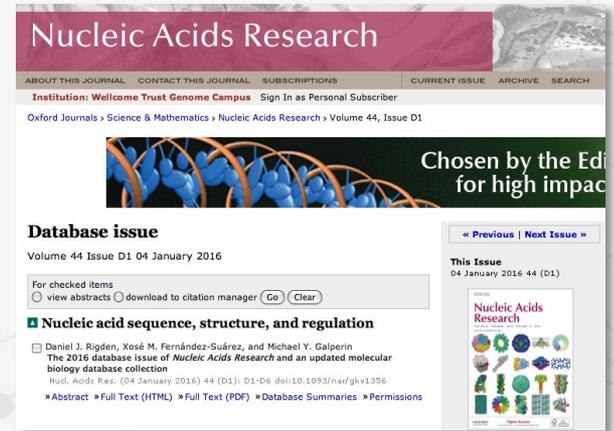


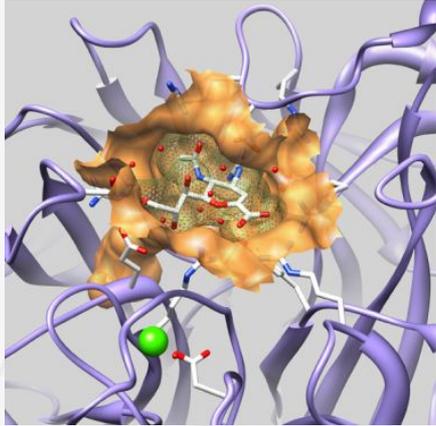
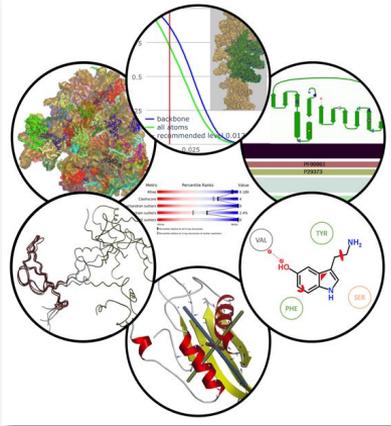
# wwPDB Deposition (OneDep)



# Many PDB-derived databases/resources!

- Since 2011, >25% of new databases described in annual **NAR** Database issues used PDB data (119 of 452)
- >200 databases use PDB data (2016)
  - 123 structure databases
  - 49 sequence databases
  - 22 metabolic and signalling pathways databases



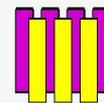
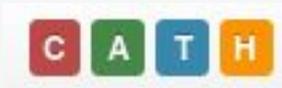


Mutation X  
disrupts  
function of  
enzyme Y, which  
causes disease Z

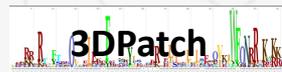
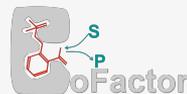
“3D-coordinates by themselves just specify shape and are not necessarily of intrinsic biological value, unless they can be related to other information”

*Integrative database analysis in structural genomics, Mark Gerstein, Nature Structural Biology 7, 960 , 2000*





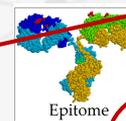
Arpeggio



Maintained  
at UNIPD



AKID



Phyre2



M-CSA



Active  
collaboration  
with UNIPD

3DLigandSite



DynaMine

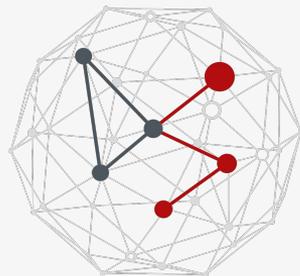


1222-2022  
800  
ANNI



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

 DIPARTIMENTO  
MATEMATICA



**DATA SCIENCE**  
UNIVERSITY OF PADOVA

# X-RAY CRYSTALLOGRAPHY

Master of Science in Data Science

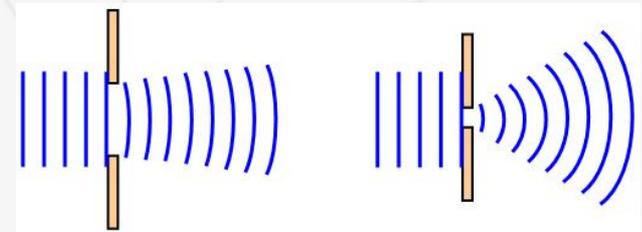
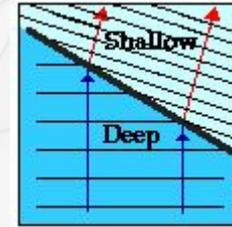
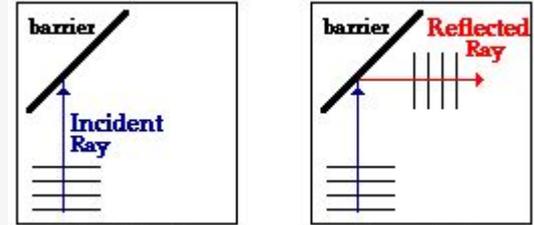
Damiano Piovesan



# Reflection, refraction, diffraction

Boundary behaviors of waves associated with the bending of the path of a wave

- **Reflection**, the incident angle equals the reflection angle
- **Refraction**, change in the direction and change in speed and wavelength, when a wave bounce off a barrier
- **Diffraction**, change in direction of waves as they pass through an opening or around a barrier

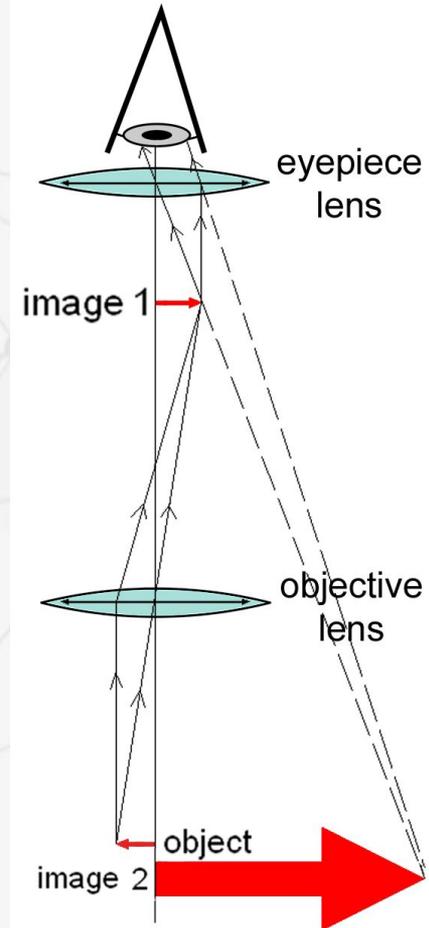


# Light microscopy

Μικρός (micros): small

Σκοπεῖν (skopein): to observe

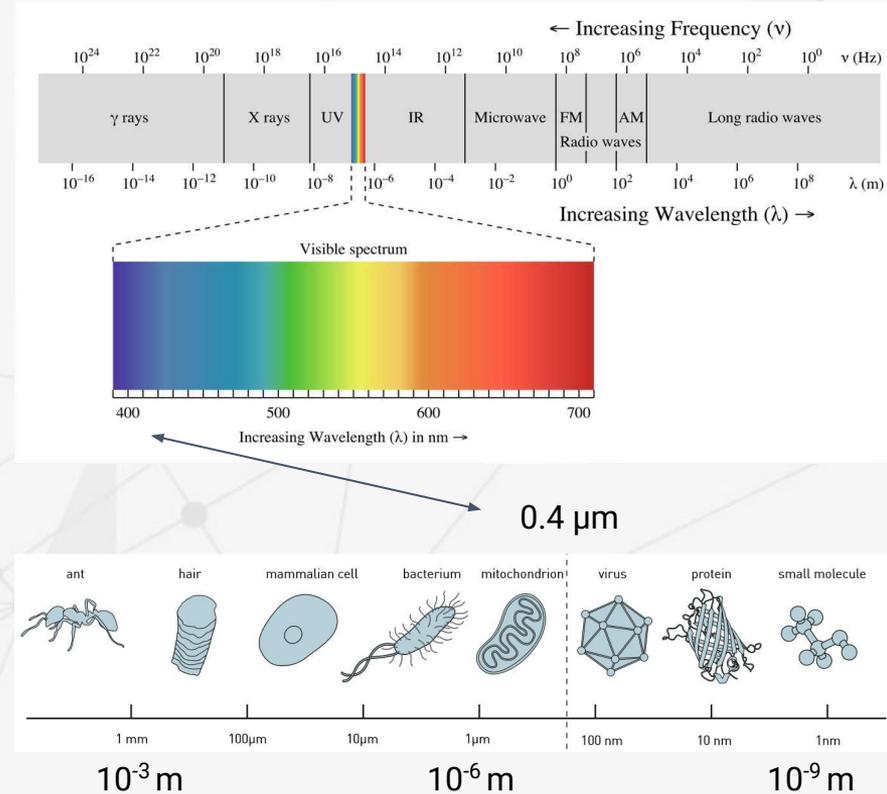
- A **lens** is a transmissive optical device that focuses or disperses a light beam by means of **refraction** (change in direction)
- The **objective lens** focuses a real image of the object inside the microscope (image 1)
- The **eyepiece lens** gives an enlarged inverted virtual image of the object (image 2)



*Compound microscope*

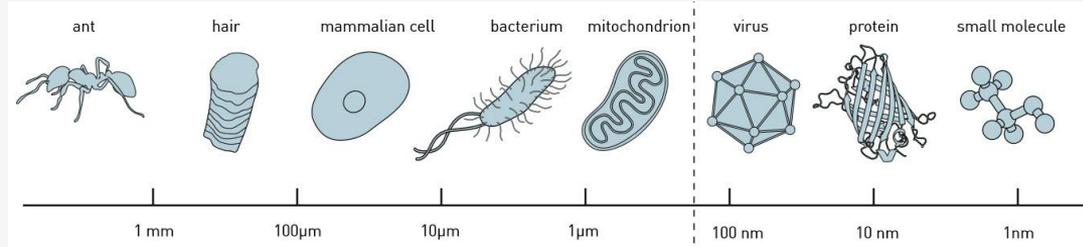
# Can we see molecules?

- Atoms or molecules exposed to light absorb light energy (photons) and **re-emit light in different directions (scattering)** with different intensity
- You can only resolve (see) the objects by looking at the **reflected waves**
- **Diffraction** limits resolution of objects with the size less than half wavelength



# Minimum resolvable distance

$$d = \frac{\lambda}{2n \sin \theta} = \frac{\lambda}{2NA}$$

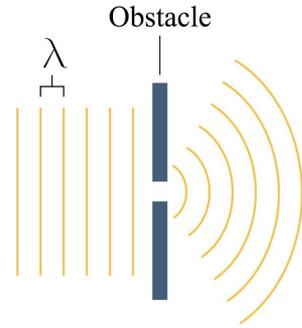


- **Numerical aperture (NA)**, property of the lens. Max 0.95 with air and 1.5 with oil
- **Light wavelength ( $\lambda$ )**, depends on the refractive materials of the lens
- In practice the lowest diffraction with conventional lenses is at about **200 nm**
- To see molecules smaller than 200nm we need to use **X-ray**



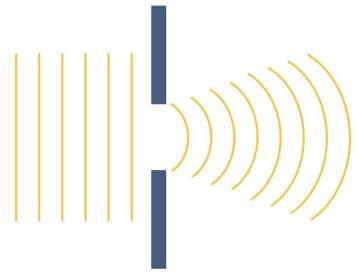
# Diffraction

## Single Slit



More diffraction when the length of the opening is smaller than the wavelength

$\lambda$  = Wavelength



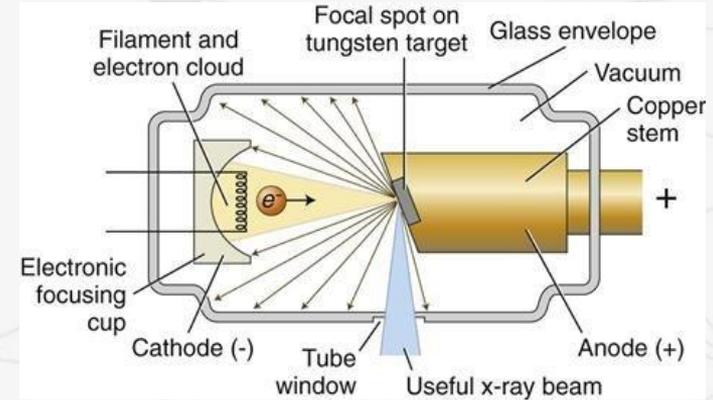
Less diffraction when the length of the opening is larger than the wavelength



# Producing X-ray

- A X-ray tube converts electrical input into X-rays
- The cathode/anode pair generates an electrical current, the **beam**
- The **high voltage accelerate the electrons** (30 to 150 kV)
- Electrons collide with the anode material and accelerate other electrons, ions and nuclei
- About **1% of the energy** is emitted/radiated as **X-rays**

## X-ray tube



Wilhelm Conrad Röntgen

Nobel 1901 for discovering X-ray

*(photoelectric effect of different elements depends on  $Z$  - atomic number)*



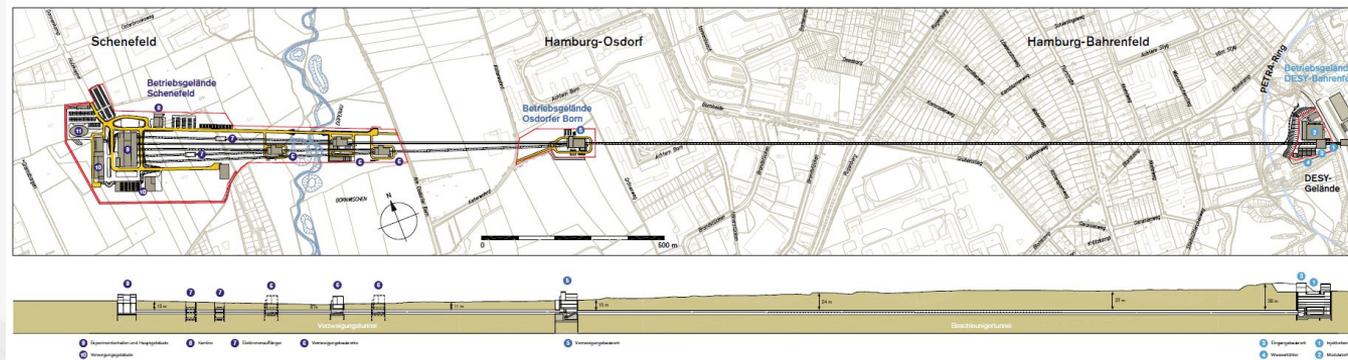
# Producing X-ray 2.0

- European **Synchrotron Radiation Facility, Grenoble**
- **844 m** circumference, does not fit on your desk



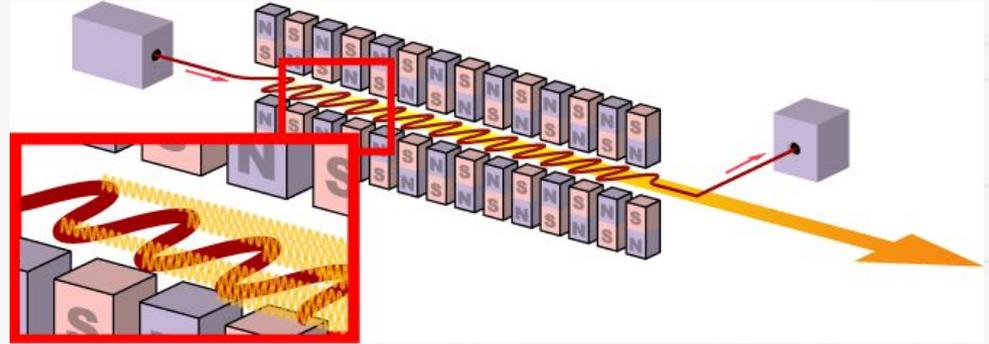
# Producing X-ray 3.0

- European **X-Ray Free-Electron Laser** (XFEL), 2017
- **3.4 Km**, Germany (Hamburg → Schenefeld)
- Quality of radiation:  **$10^{35}$**  brilliance (FEL) Vs  **$10^{24}$**  brilliance (Synchrotron)

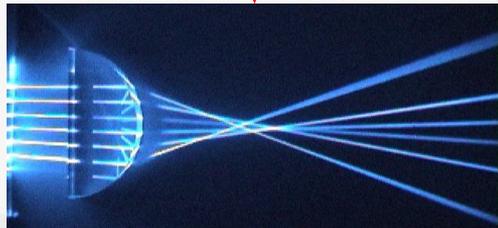
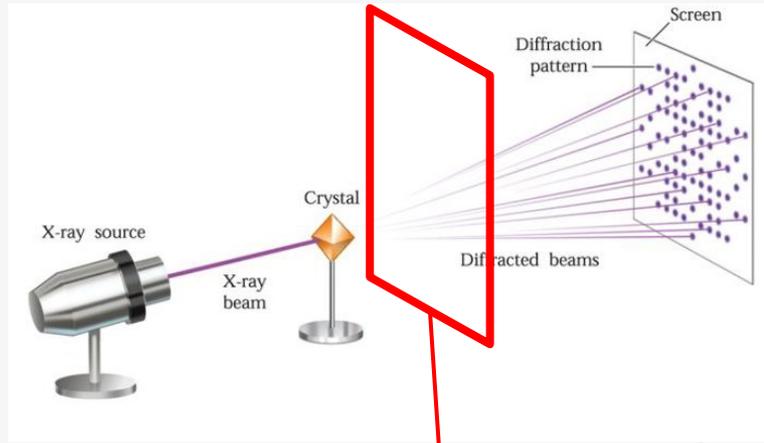


# Free-electron laser (FEL)

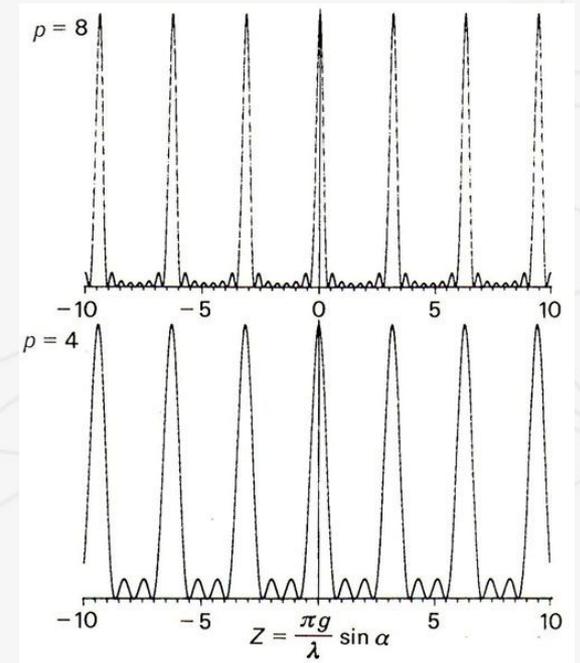
- **Very-high-speed electrons** moving freely through a magnetic structure (**undulator**)
- The **transverse acceleration** of the electrons across this path **release photons (synchrotron radiation)**
- **Monochromatic photons** (single frequency), but **incoherent** (not in phase)
- **Standing waves** (in phase) are provided by mirrors at each end of the undulator that create an optical cavity
- Produced radion is **very intense** and **concentrated in narrow energy bands** in the spectrum (**X-ray**). It is also **collimated** on the orbit plane of the electrons



# Geometrical optic Vs scatter spectra



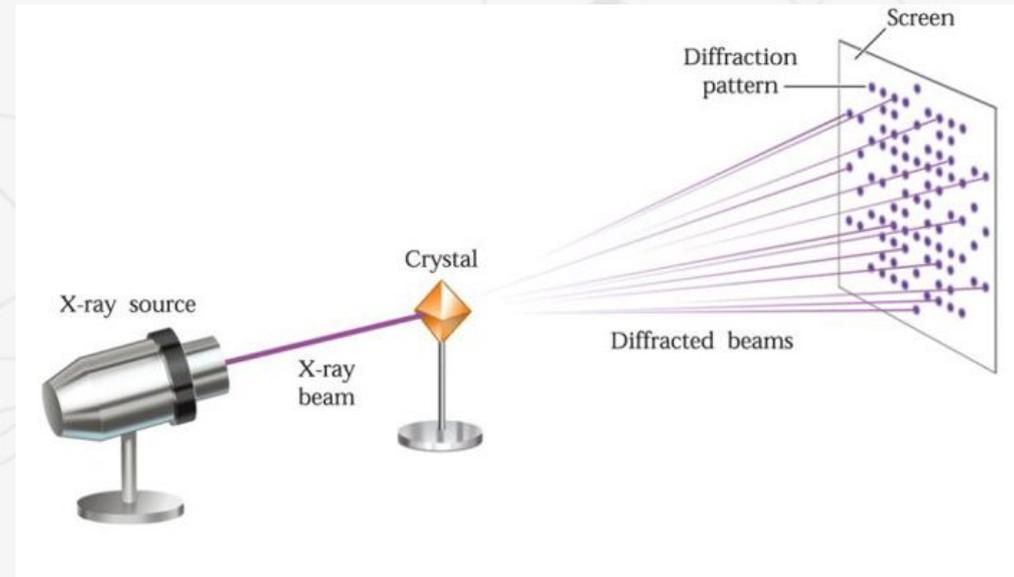
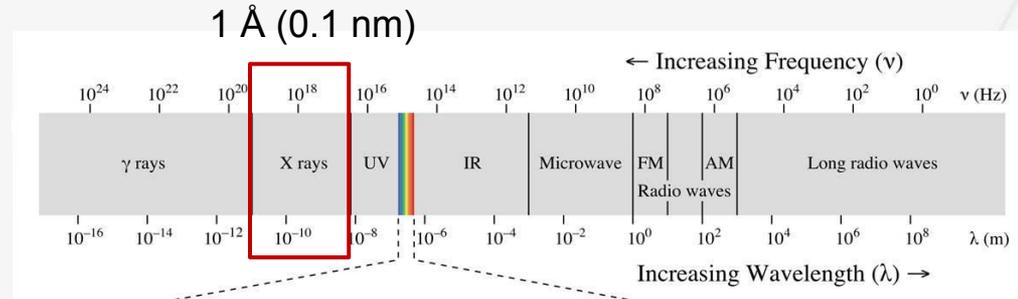
What you would like



Instead you get

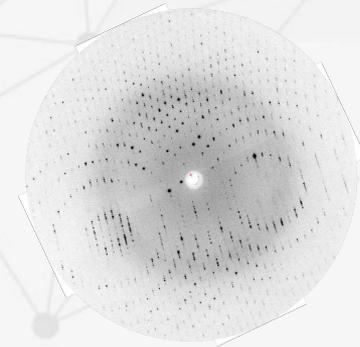
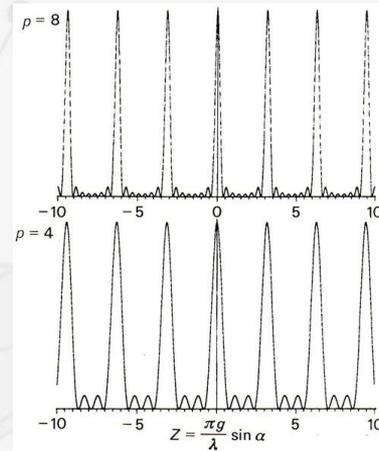
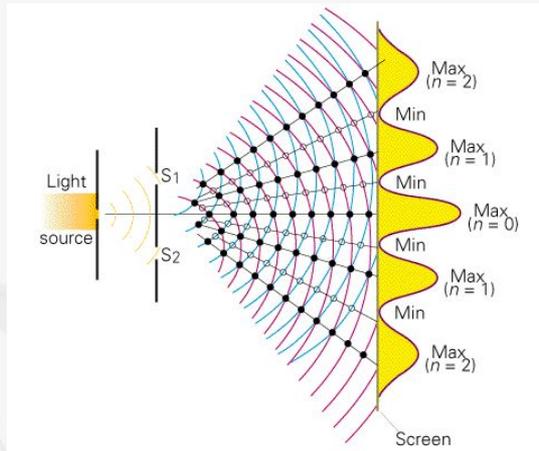
# No lenses for X-rays

- There is no way to re-focus X-rays
- All we get is the **diffracted** wave packets (**scattering**)
- The crystalline structure causes a beam of incident X-rays to diffract into many **specific directions**



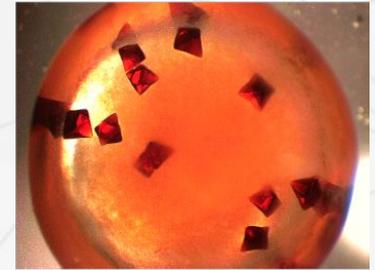
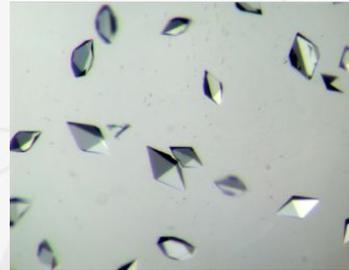
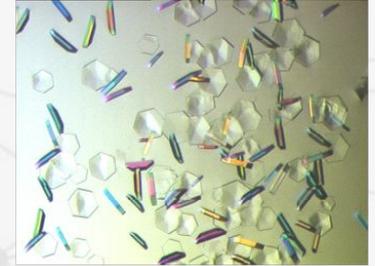
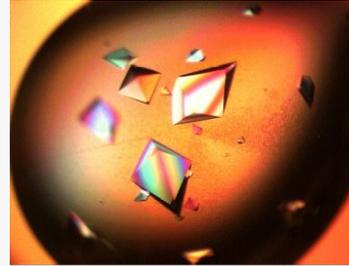
# The role of crystals

- Diffraction from a single molecule is **undetectable**
- The grating of the crystal lattice generates **interference** of the diffracted X-ray photons
- More slits = sharper maxima
- Diffraction on a crystal lattice results in a **Discrete** and **Amplified** diffraction pattern



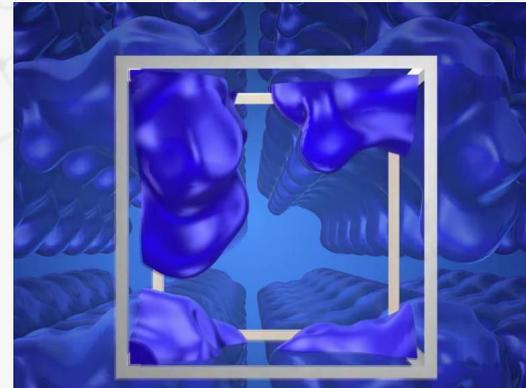
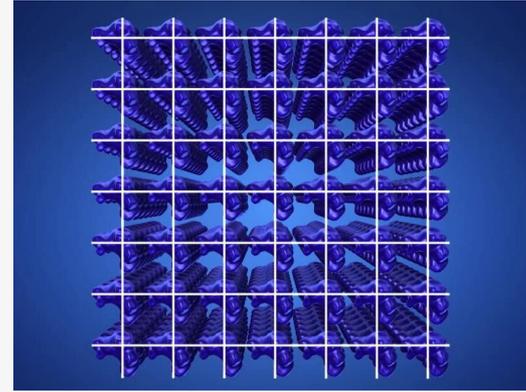
# Crystals

- Crystals are grown with chemical techniques
- Most of them are generated by protein **precipitation** due to **high concentration**
- Bacteria are modified to **overexpress** the protein and generate high concentration



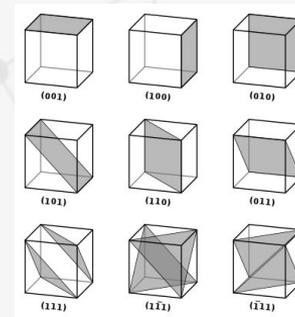
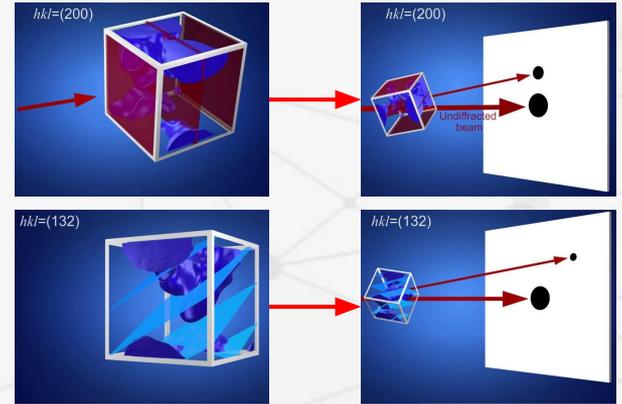
# Crystal lattice

- In **ideal crystals**, proteins are arranged in a **perfect array** and all are in the **same orientation**
- The organization of **proteins in a crystal** can be described by a **crystal lattice** that highlights its repetitive structure
- The unity of the lattice is the **unit cell** which is **translationally repeated** in three dimensions

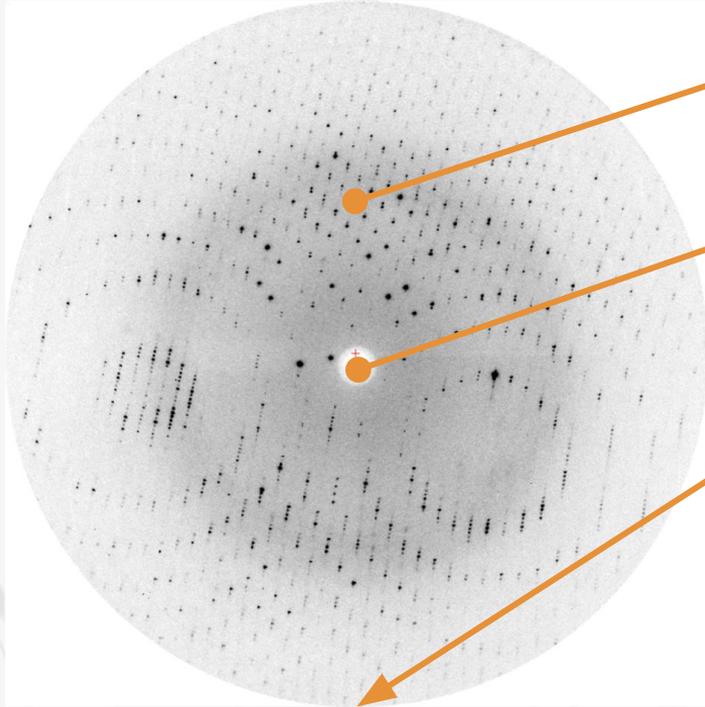


# Diffraction patterns

- The unit cell can be divided in **planes** that diffract X-rays in different directions
- Each set of **planes produce a point** in the diffraction screen
- The more packed the planes, the greater the **distance of the point** from the center
- **Planes names** are based on how many times they divide the unit cell
- **Miller indexes**, how many equal parts the unit cell is divided by the lattice plane



# A diffraction pattern

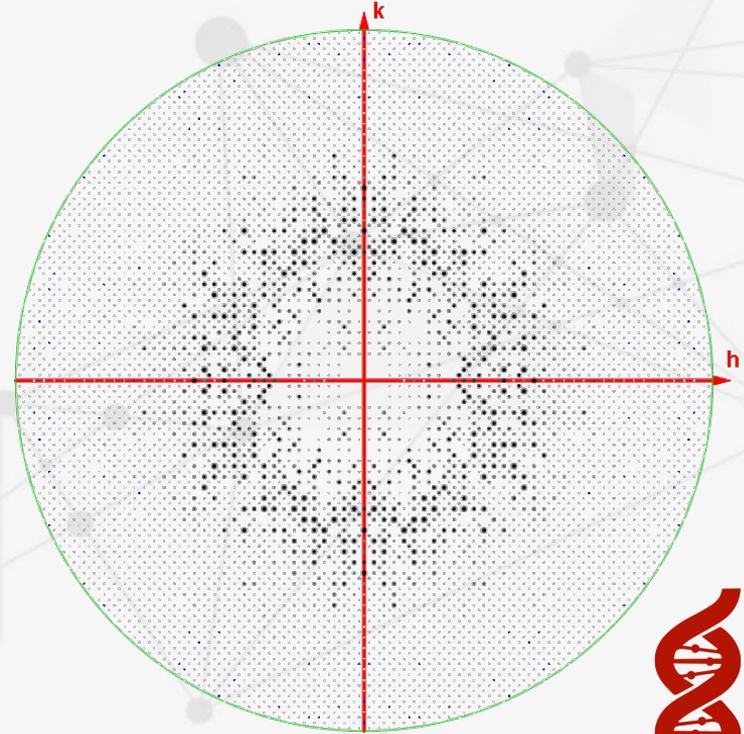


- **Reflections** distinct intensity maxima appear in a regular arrangement
- **Beam Stop Shadow** most X-rays pass the crystal without interference and are stopped by a lead beam stop
- **Resolution** the further away from the center a reflection appears, the more detailed is the structural information encoded

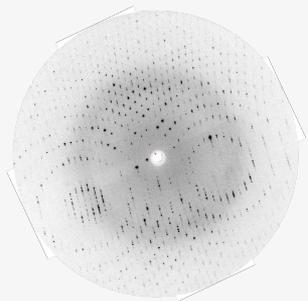


# Reciprocal lattice

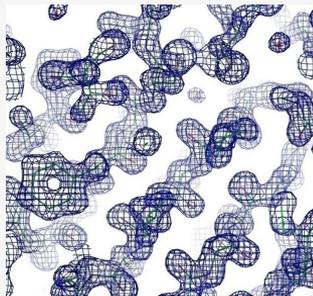
- The **lattice of diffraction maxima** is related to the **crystal lattice** but its dimensions are **reciprocal**
- **Large crystals** cause **short distances** between diffraction spots and vice versa
- Coordinates in the reciprocal space are the **Miller indexes (hkl)**



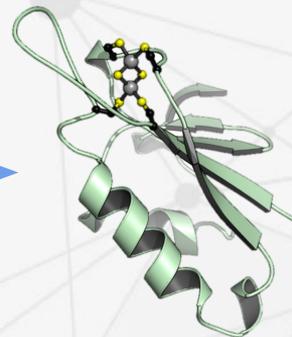
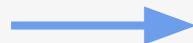
# Solving a crystal structure



X-ray diffraction  
pattern



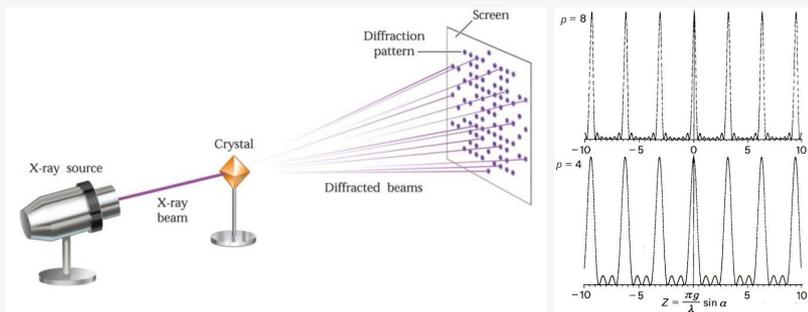
Electron density map



Atomic model

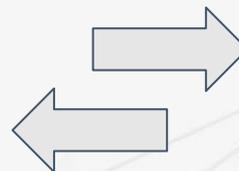


# A mathematical lens

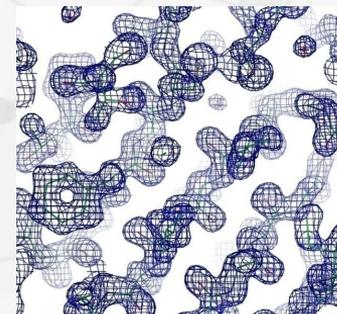


Reciprocal space

Inverse Fourier Transform  
( $FT^{-1}$ )



Fourier Transform  
( $FT$ )

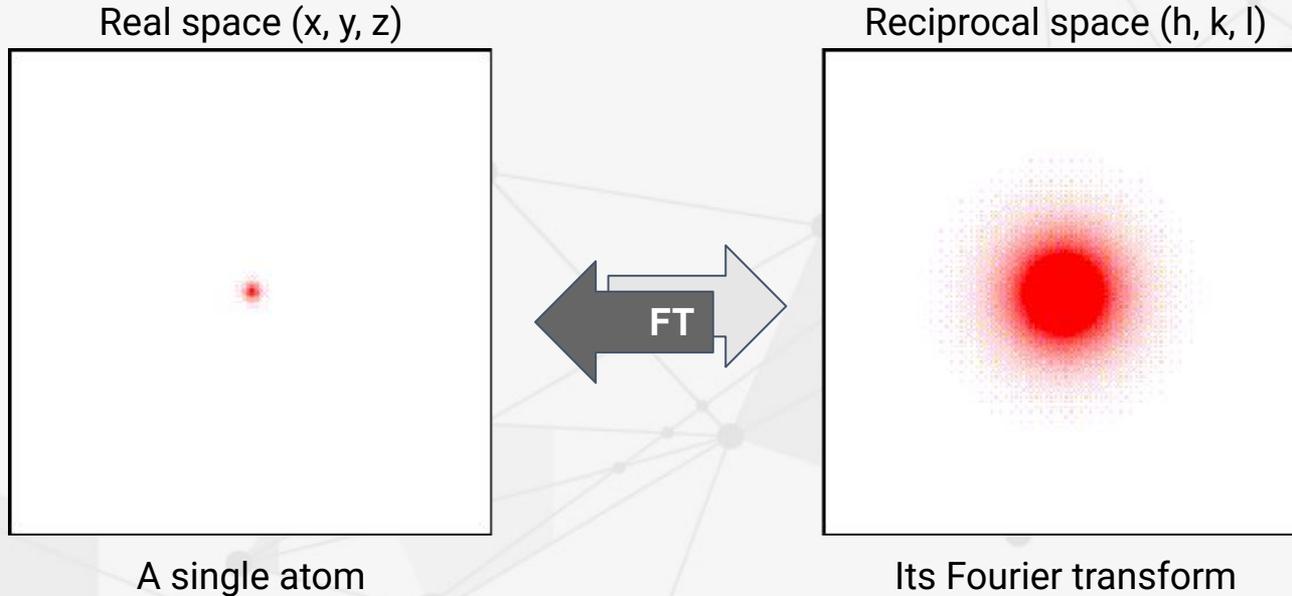


Electron density map



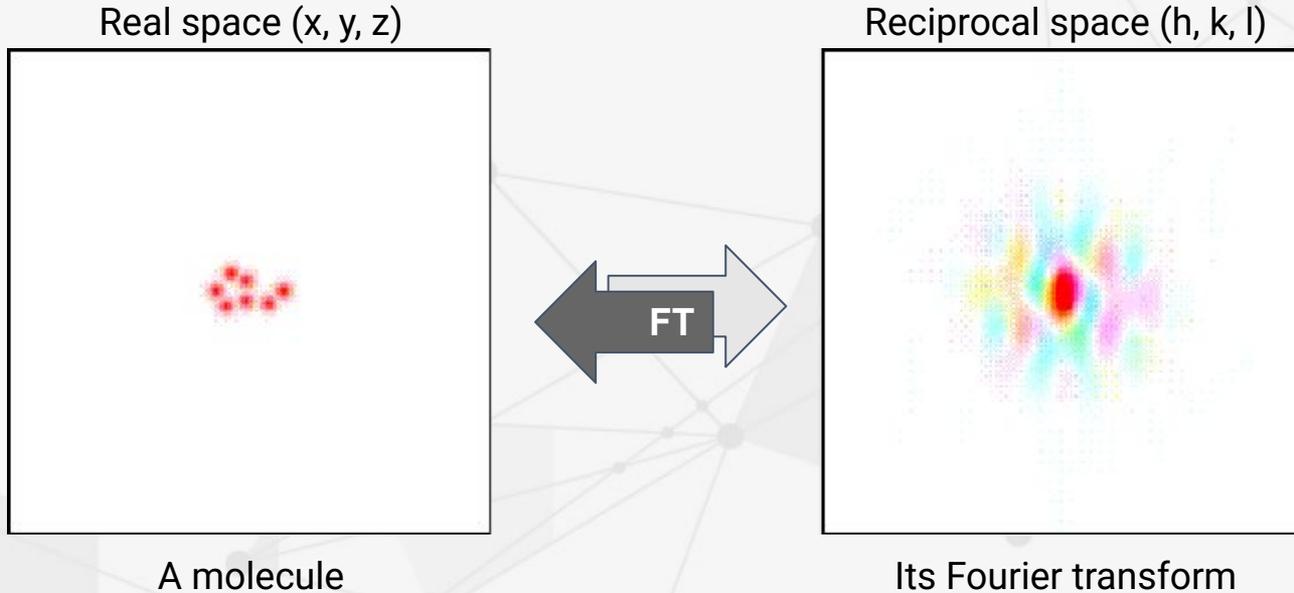
# Fourier transform

The Fourier transform reversibly connects two **equivalent** representations



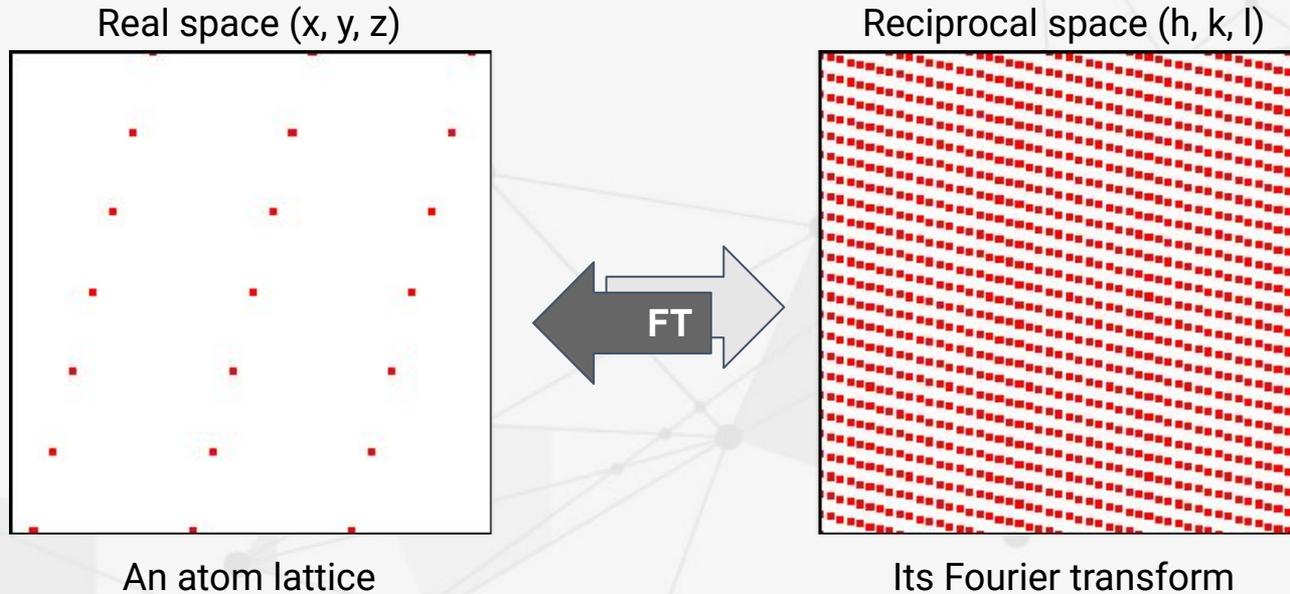
# Fourier transform

The Fourier transform reversibly connects two **equivalent** representations



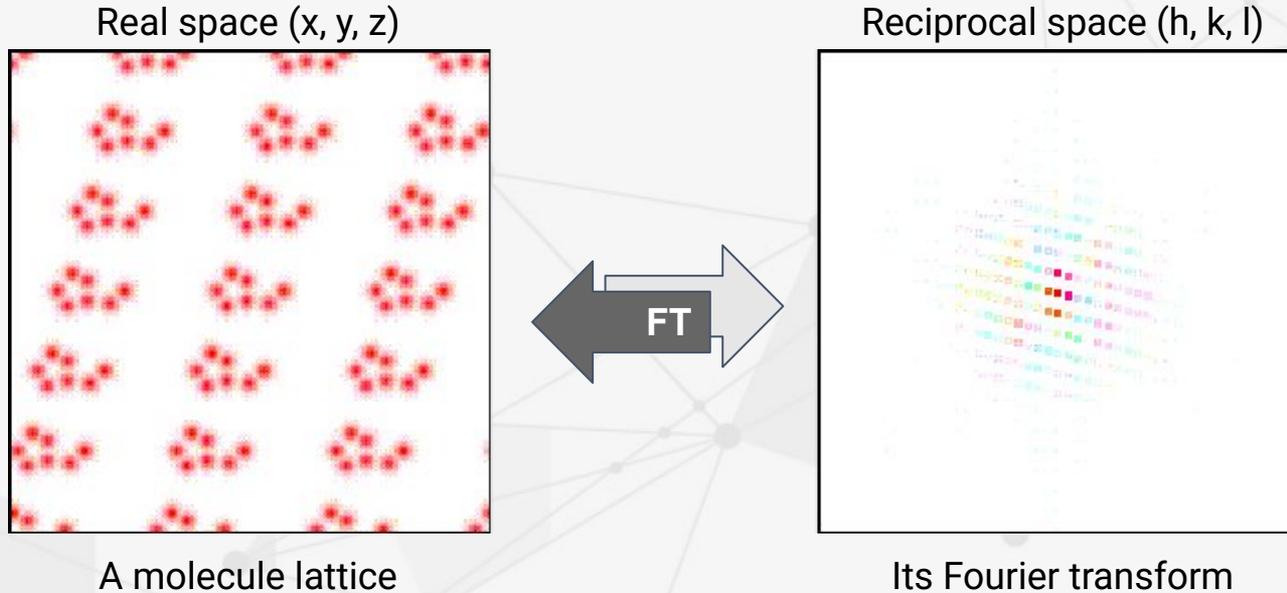
# Fourier transform

The Fourier transform reversibly connects two **equivalent** representations



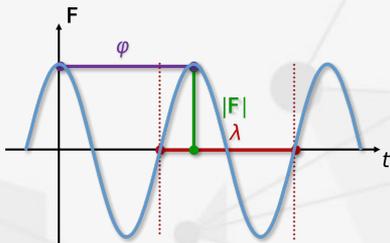
# Fourier transform

The Fourier transform reversibly connects two **equivalent** representations

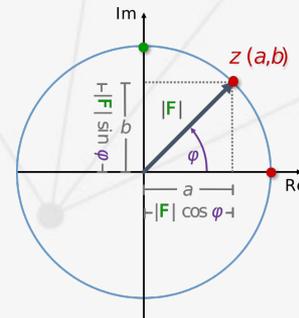


# Wave functions and diffraction plots

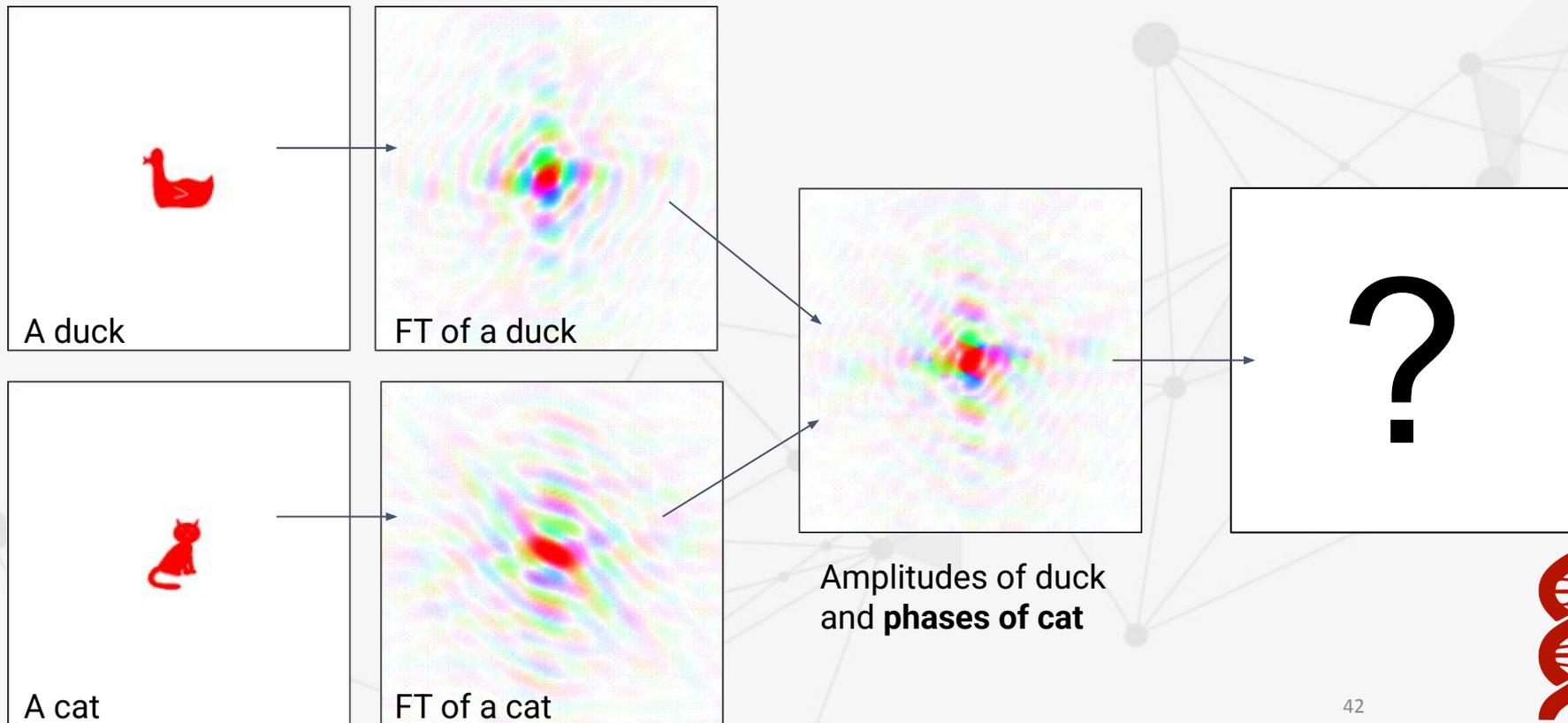
- Diffraction experiments use **monochromatic X-ray** radiation,  **$\lambda$  is constant and known**
- Every wave function is **fully described** by its **amplitude** and its **phase angle**
- In a diffraction pattern the **intensity of reflections** represents the squared value of  **$|F|$**  (structure factor)  $\rightarrow I \approx |F|^2$
- Information about the **phase is missing and must be inferred**



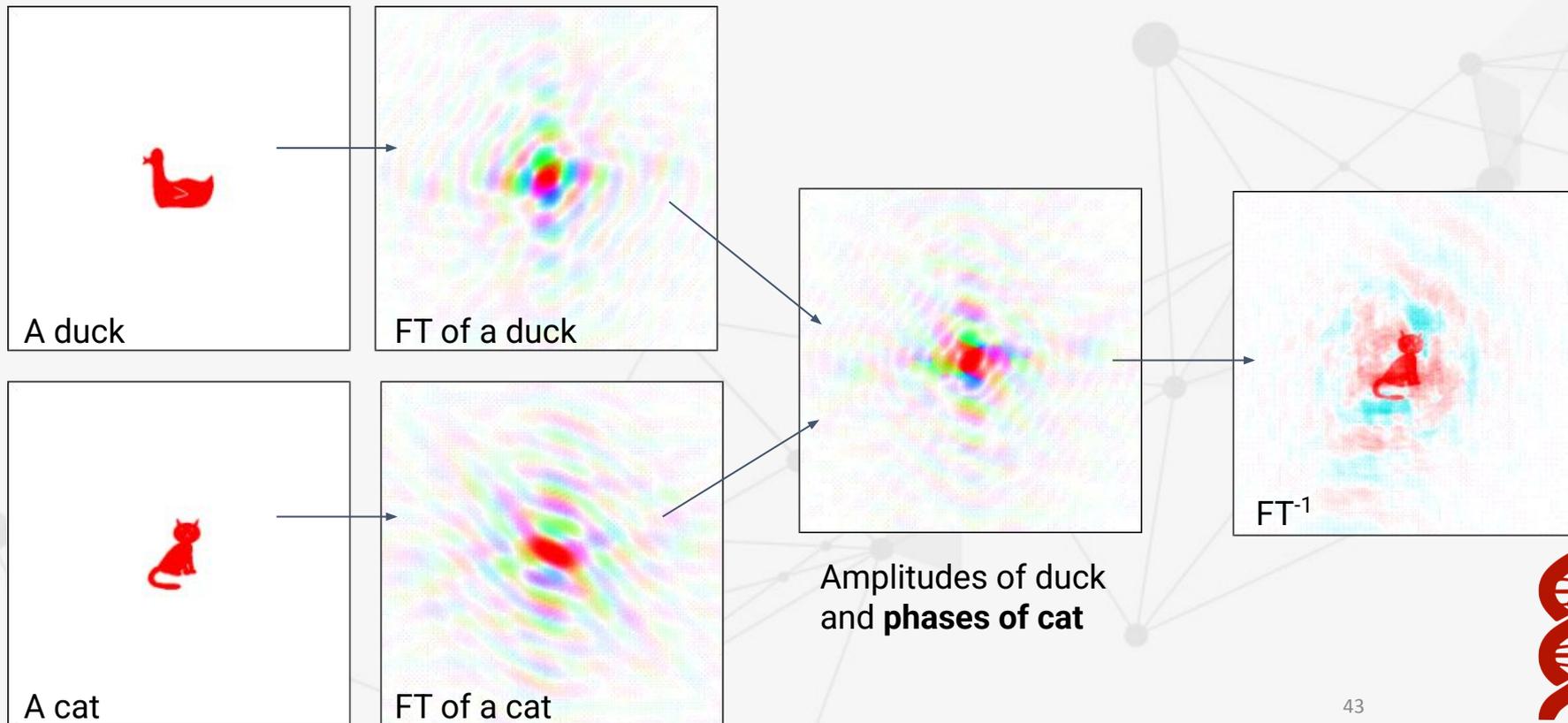
- **Wavelength  $\lambda$**
- **Amplitude  $|F|$**
- **Phase angle  $\phi$**



# Are phases so important?



# Are phases so important?



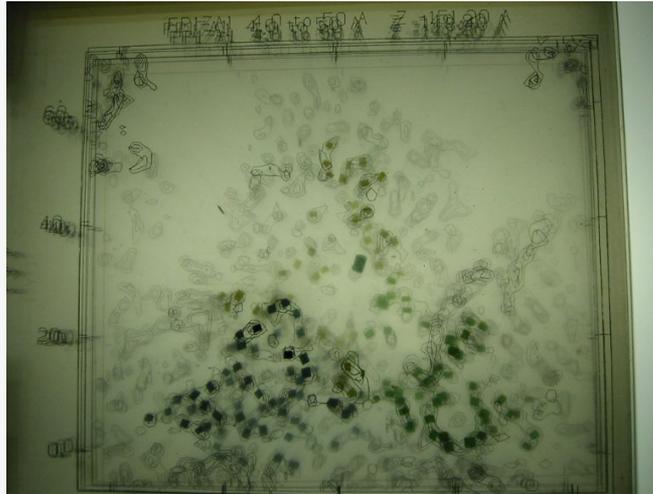
# How to solve the Phase Problem

- Based on a known structure
  - **Molecular replacement**
- *De novo* structure solution
  - **Substructure Analysis**
  - **Direct methods**

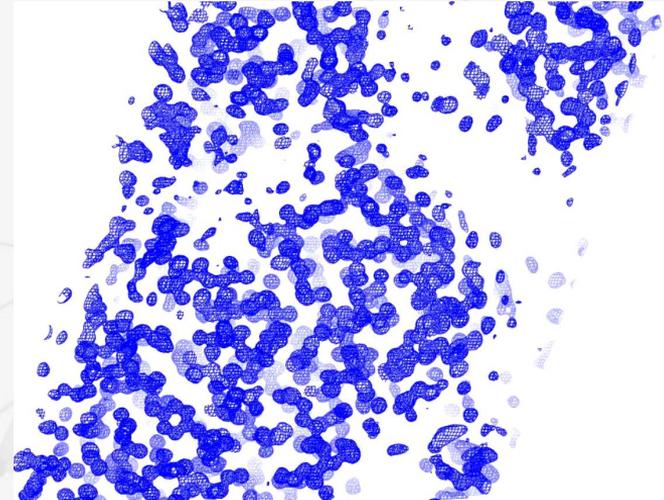


# The electron density (ED) map

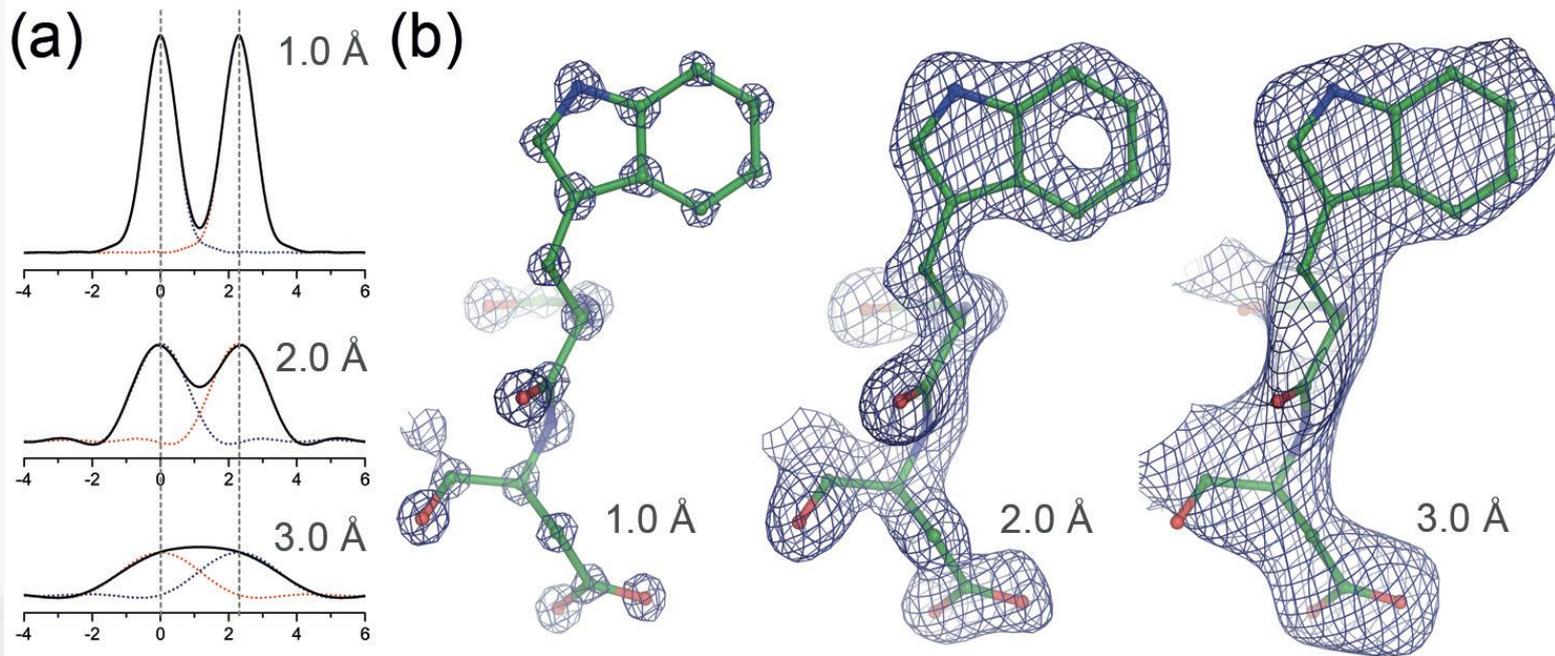
In the pre-computer era



Today



# Resolution of X-ray crystallography



# PDB guide

<https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/introduction>

