Bioinformatics III

Analysis and prediction of 3D macromolecule structures

Experimental methods for 3D structure determination

Saulius Gražulis
2024 m.

Methods to obtain structural information

- **X-ray neutron crystallography**
- **NMR**
- **CryoEM (Cryo Electron microscopy, multiple particle analysis, electron tomography)**
- Electron diffraction
- EM (Electron microscopy, negative staining)
- SAXS/SANS – Small Angle X-ray/Neutron Scattering
- EXAFS, XANES
- …

CryoEM

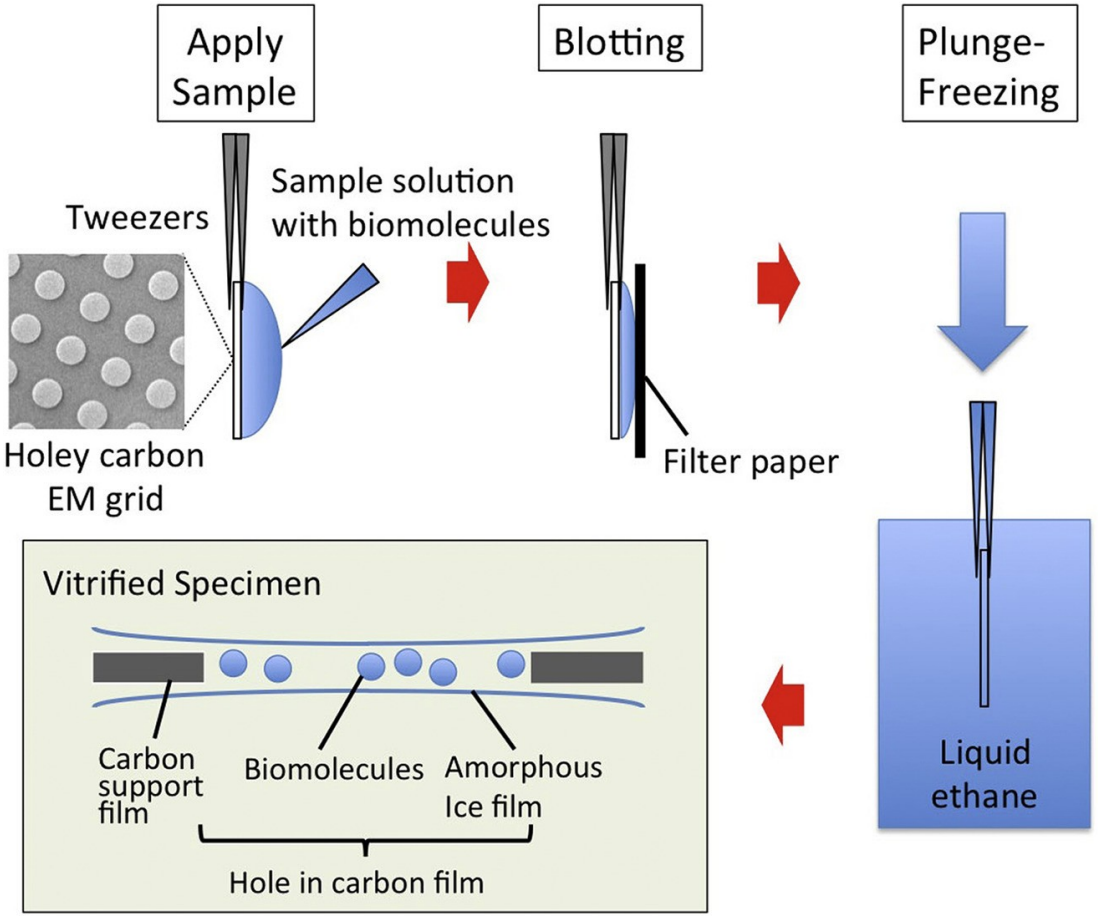
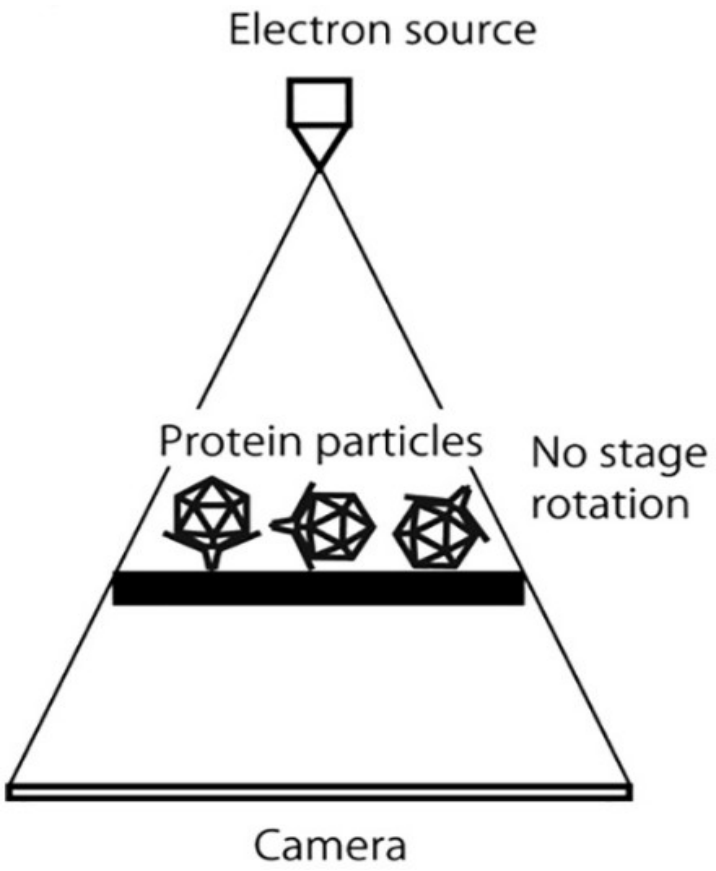
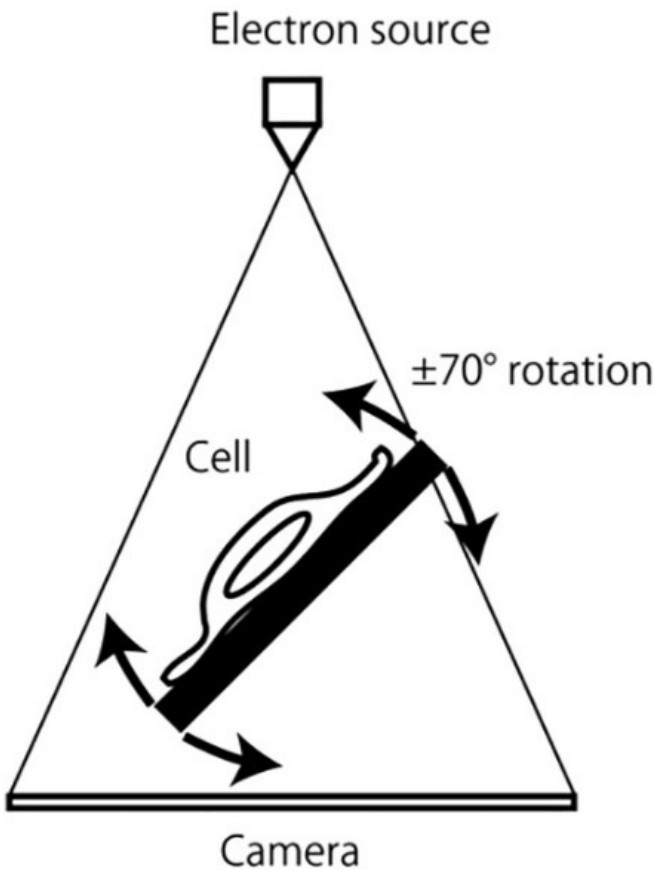


Image taken from: (Murata & Wolf, 2018) CC BY

SPA & CryoET



SPA (Single Particle Analysis)



ET (Electron Tomography)

Paveiksliukas iš: (Murata & Wolf, 2018) CC BY

CryoEM electron density

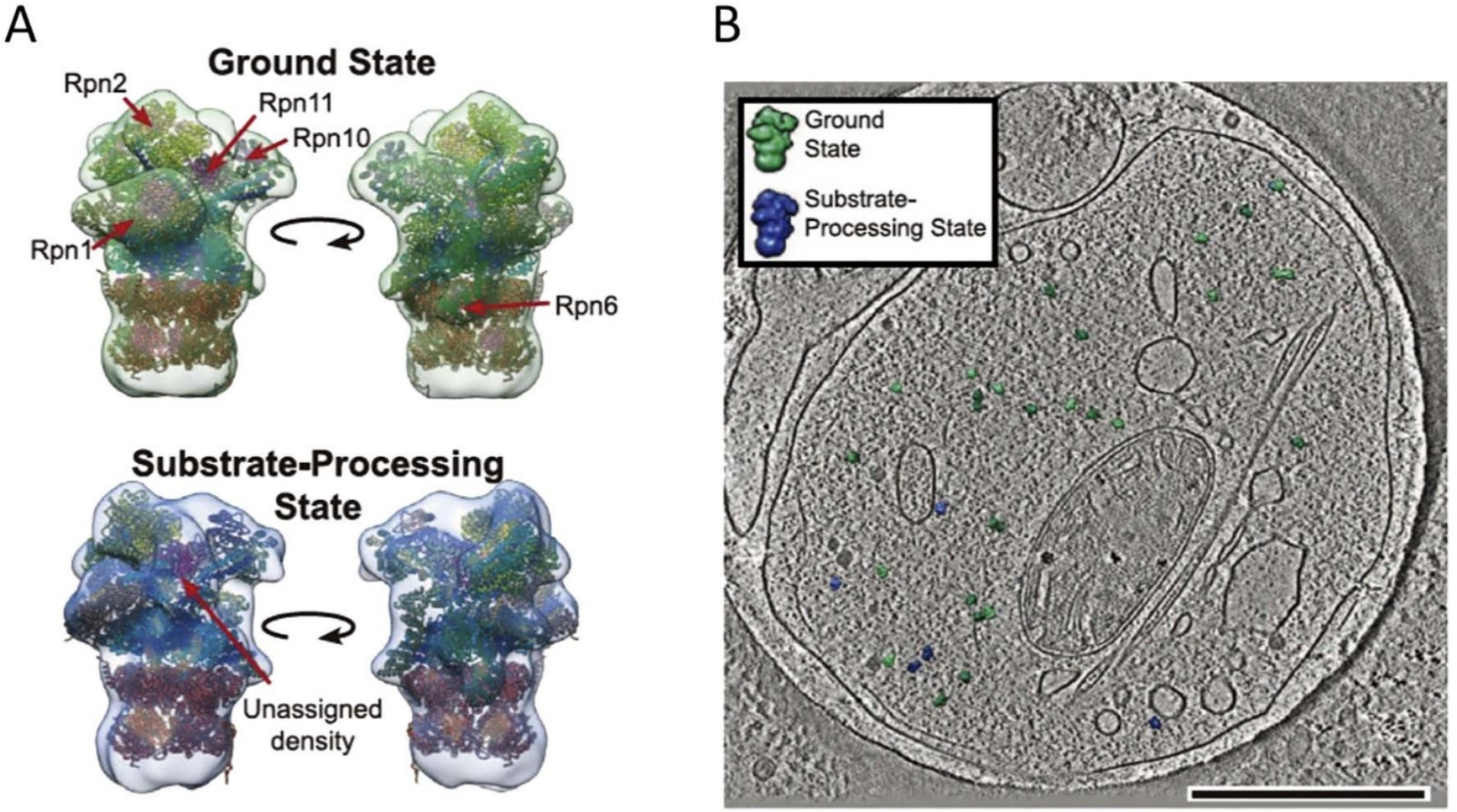
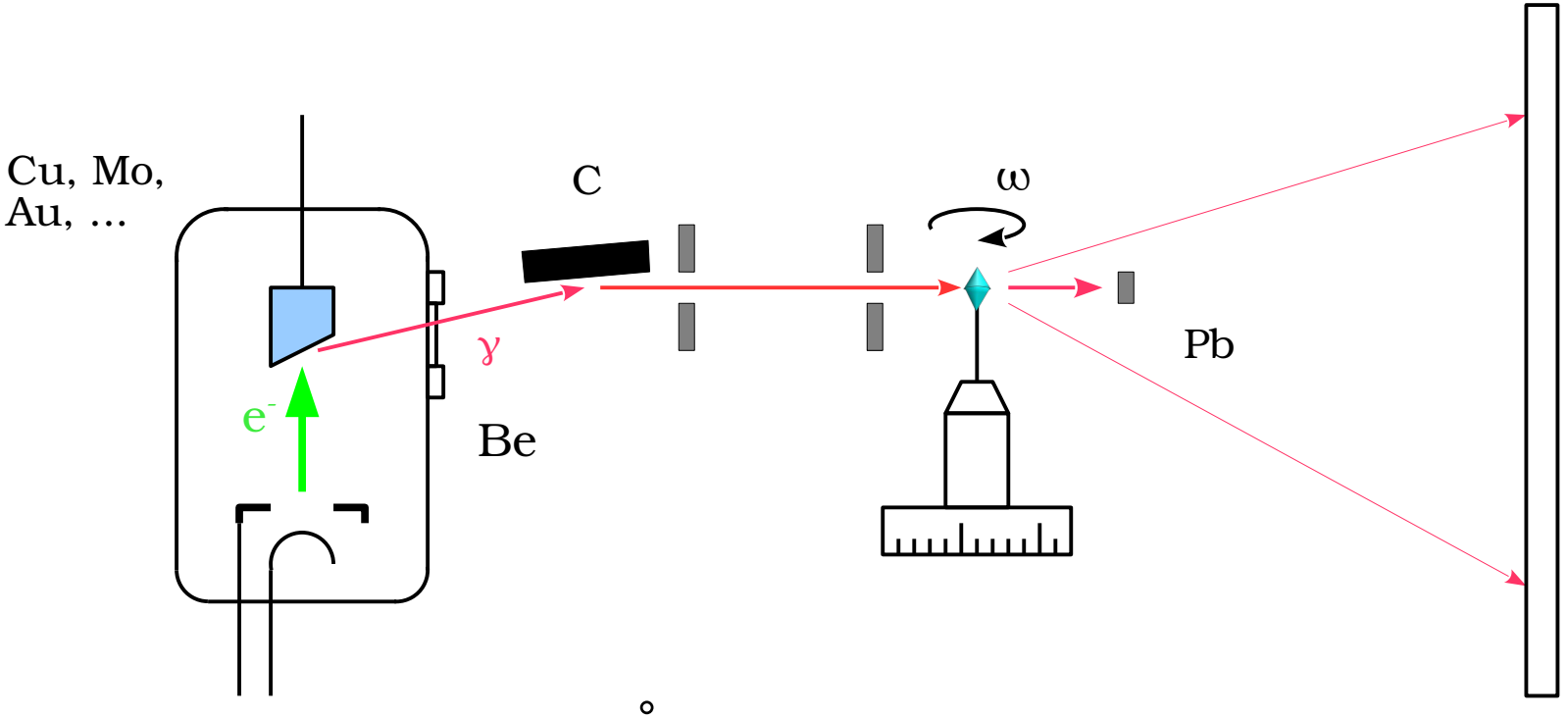


Image taken from: (Murata & Wolf, 2018) CC BY

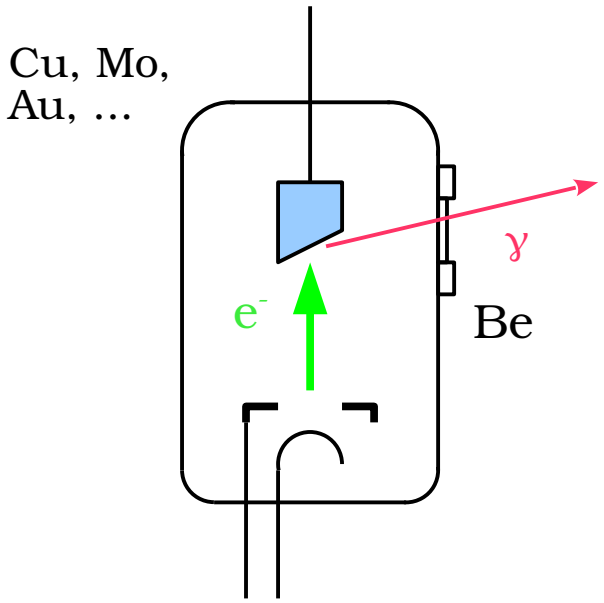
Diffraction experiment



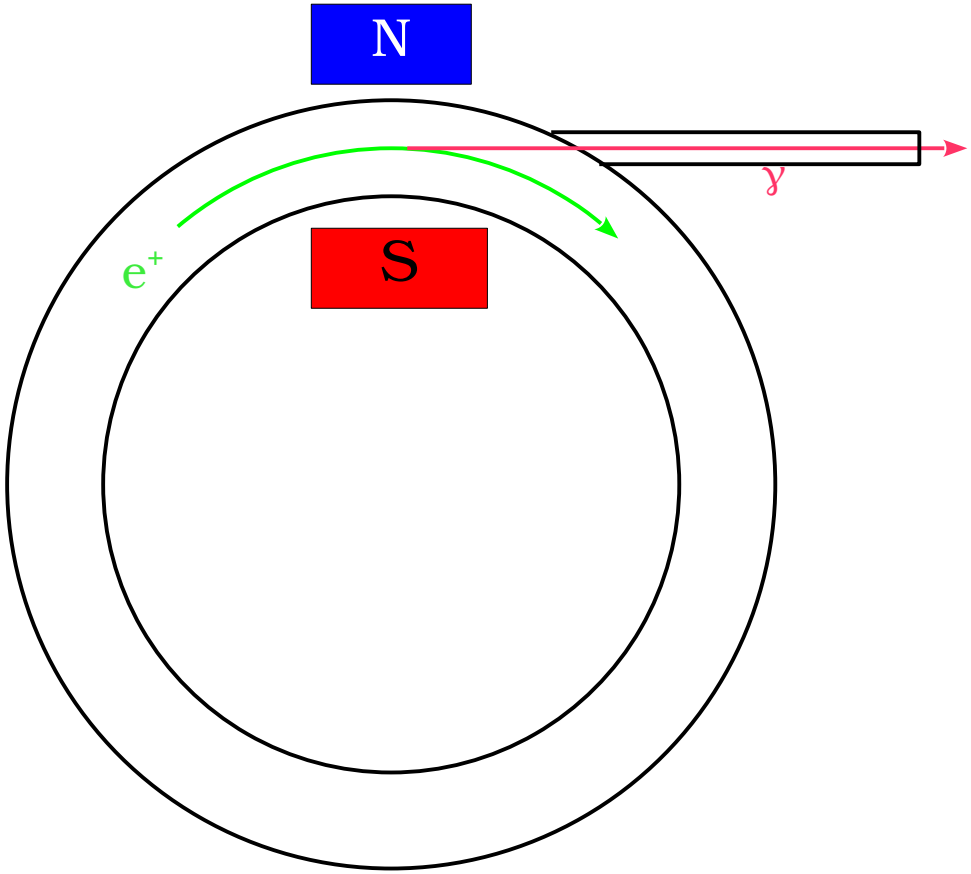
$\lambda \simeq 1\text{Å}$

X-ray sources

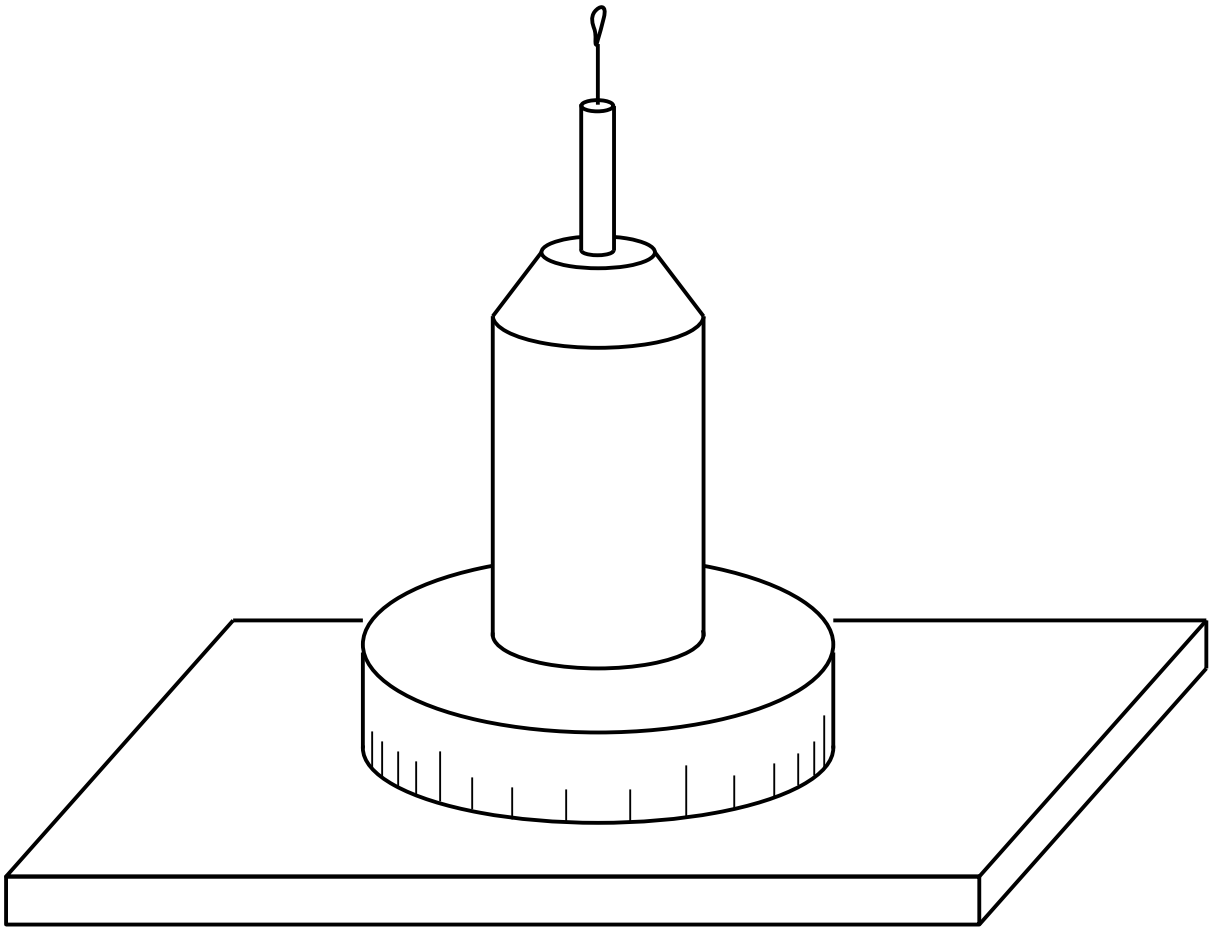
X-ray tubes



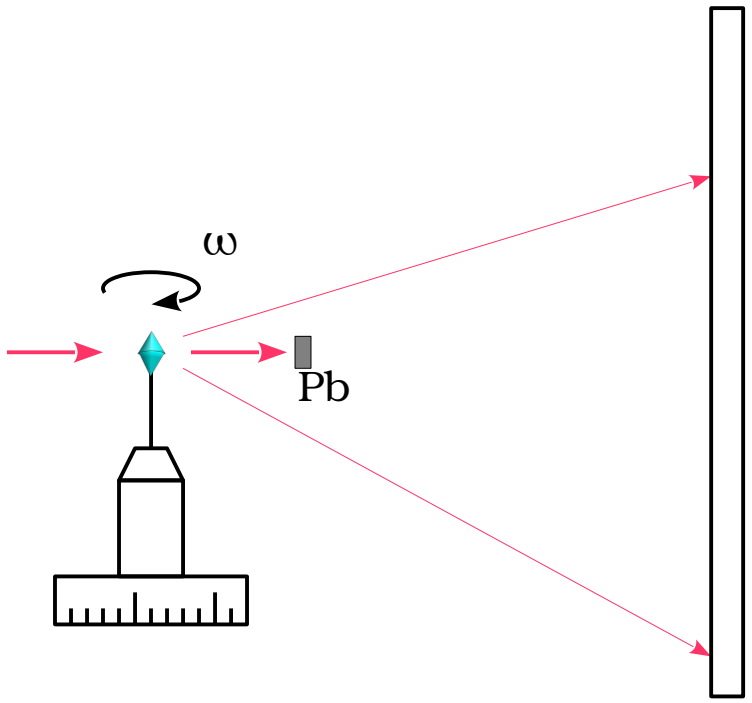
Syncrotrons (particle accelerators)

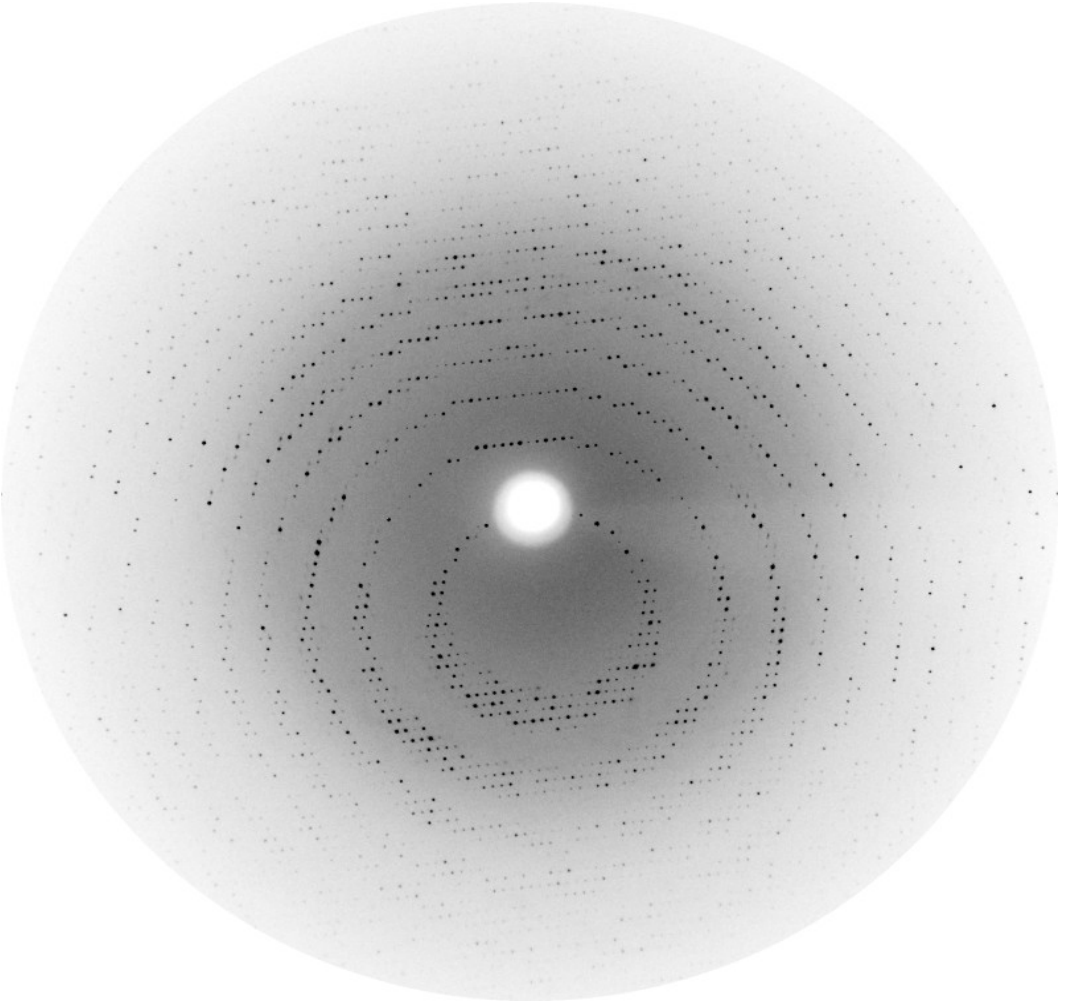


Goniometer head



A diffraction image





Properties of reflections

- Each reflections has:
  - ***an amplitude*** (of the electric field);
  - and ***phase*** (delay with respect to other reflections);
- Both of these parameters depend on a ***structure factor***.

Properties of reflections

- The reflection **positions** (i.e. reflection angles) depend only on the lattice parameters;
- The intensities of the reflections depend on the **conetnts** of the unit cell.

A complete dataset

When a crystal is rotated, some reflections appear, and some others vanish.

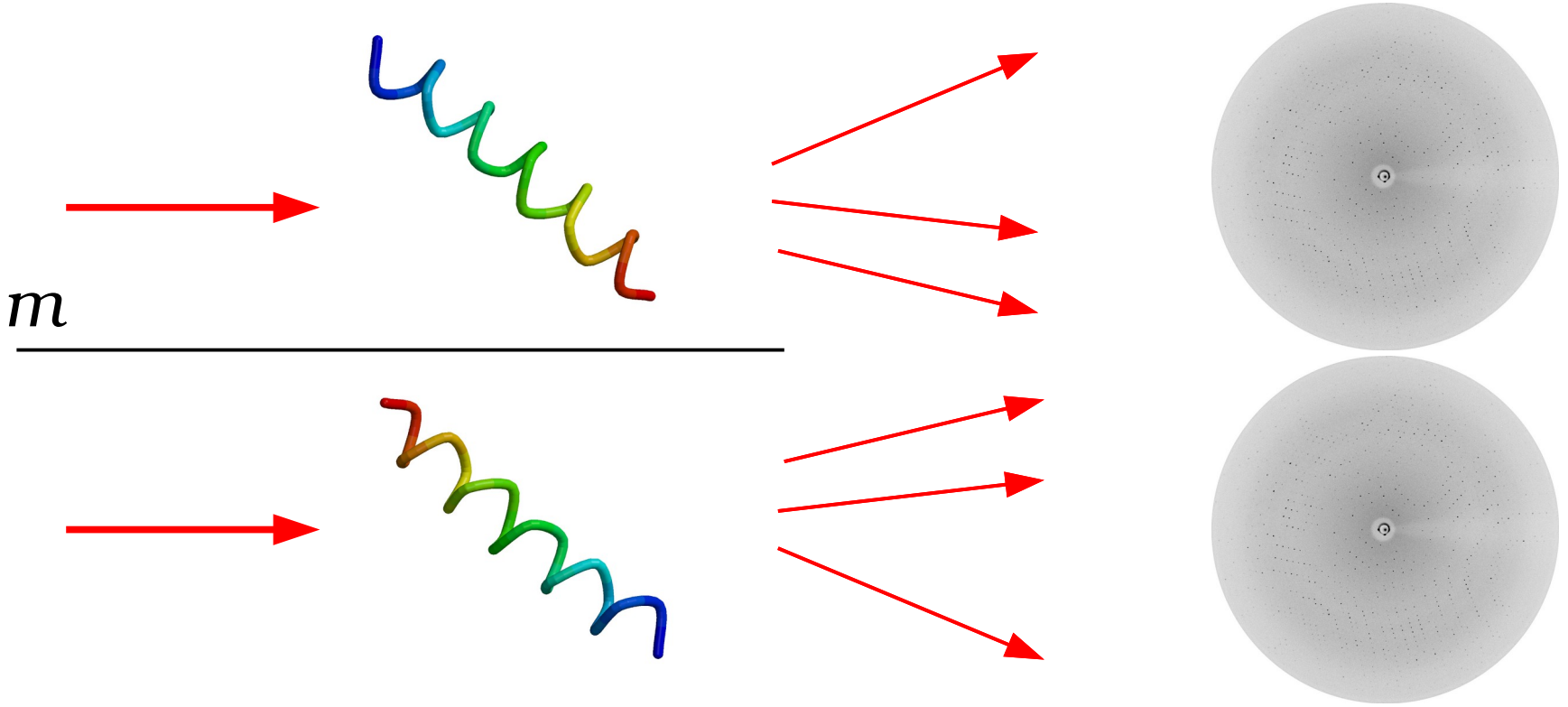
After max. 360° rotation reflections start repeating themselves (maybe sooner, especially if the crystal is highly symmetric).

Properties of a dataset

- Resolution
- Completeness
- Multiplicity
- Signal-to-noise ratio (R_{merge}, I/σ_I)

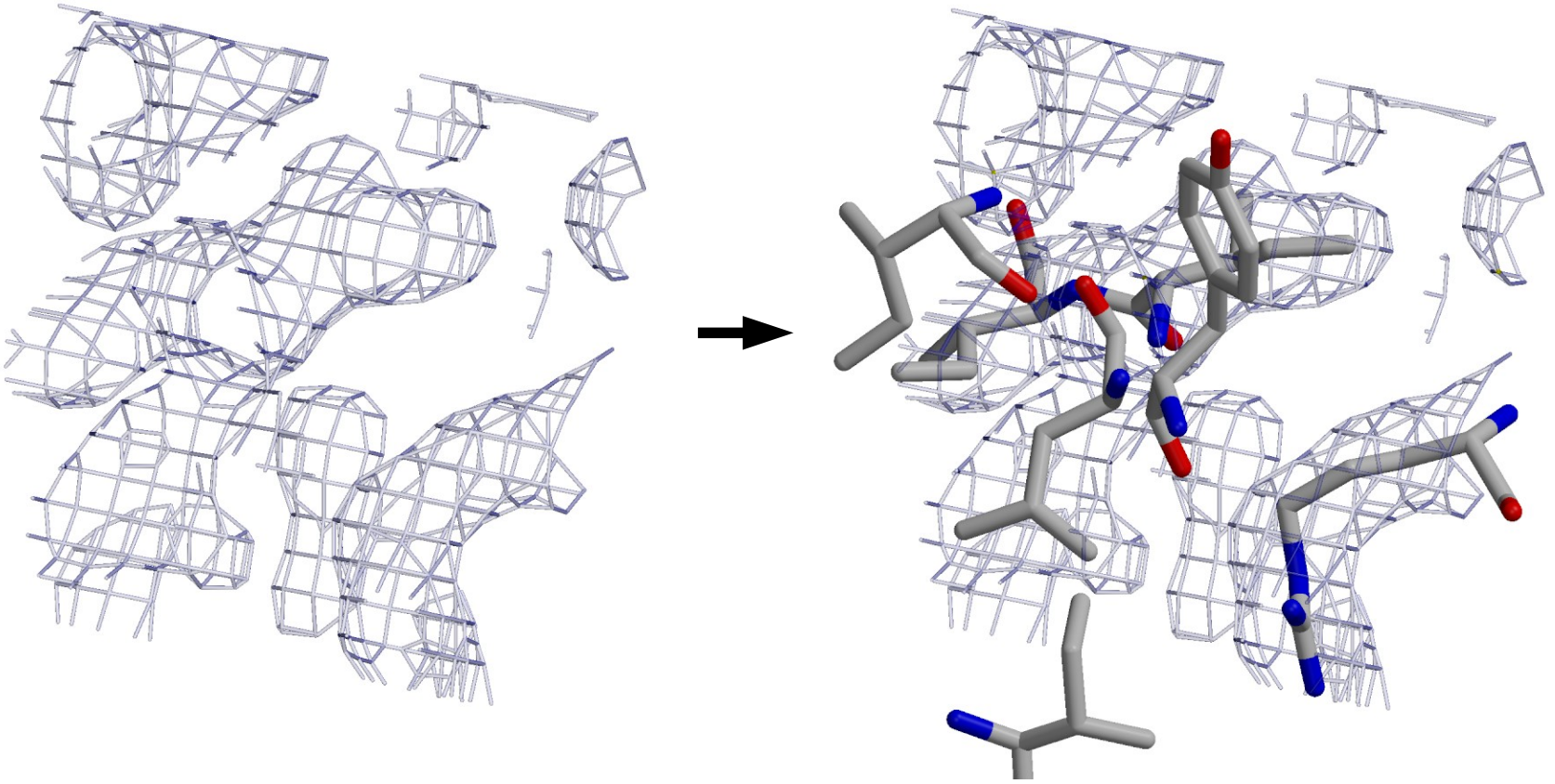
A model might not be unique!

Example:

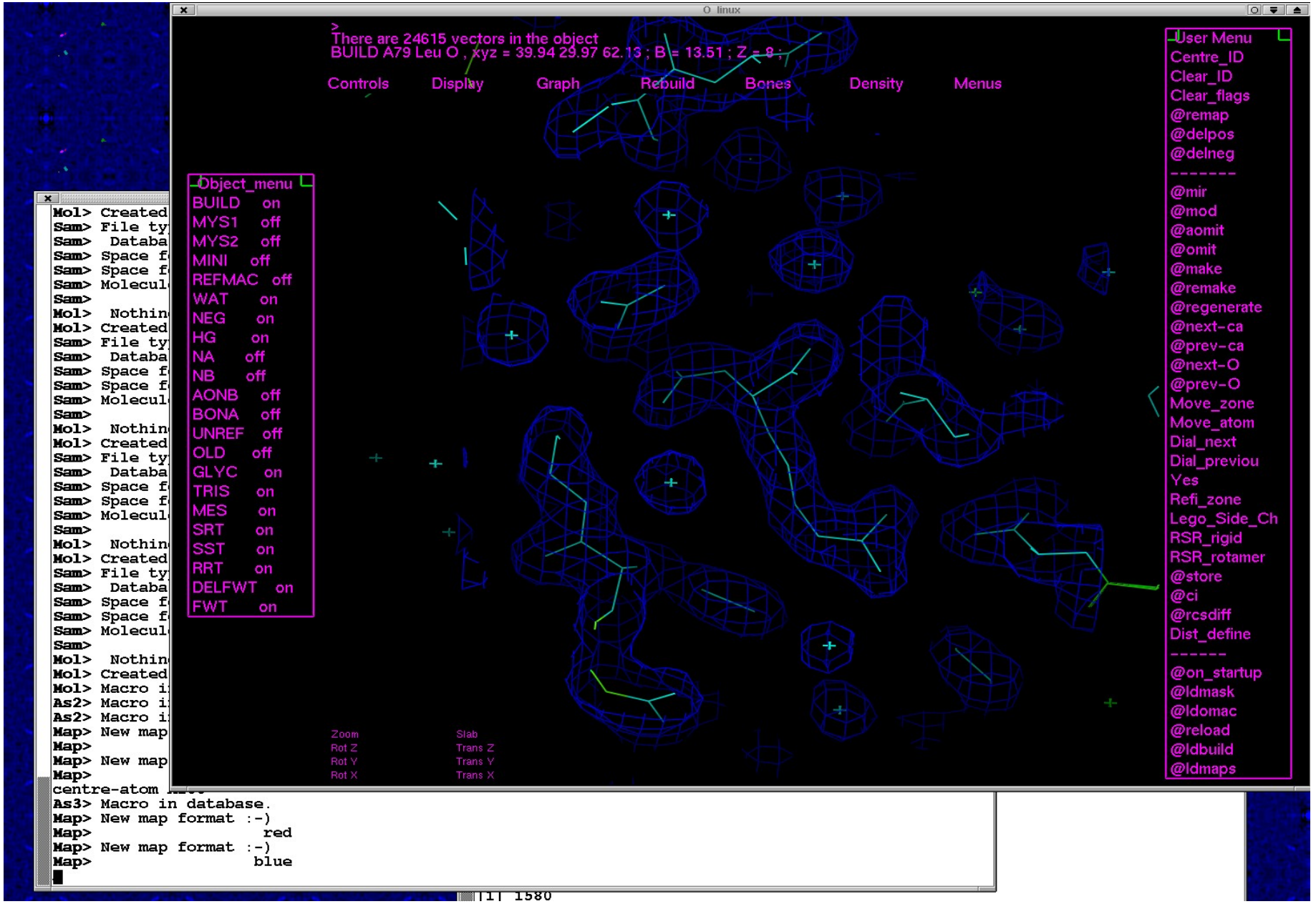


A mirror image of a structure has (nearly) identical diffraction pattern (identical if there is no anomalous scattering)

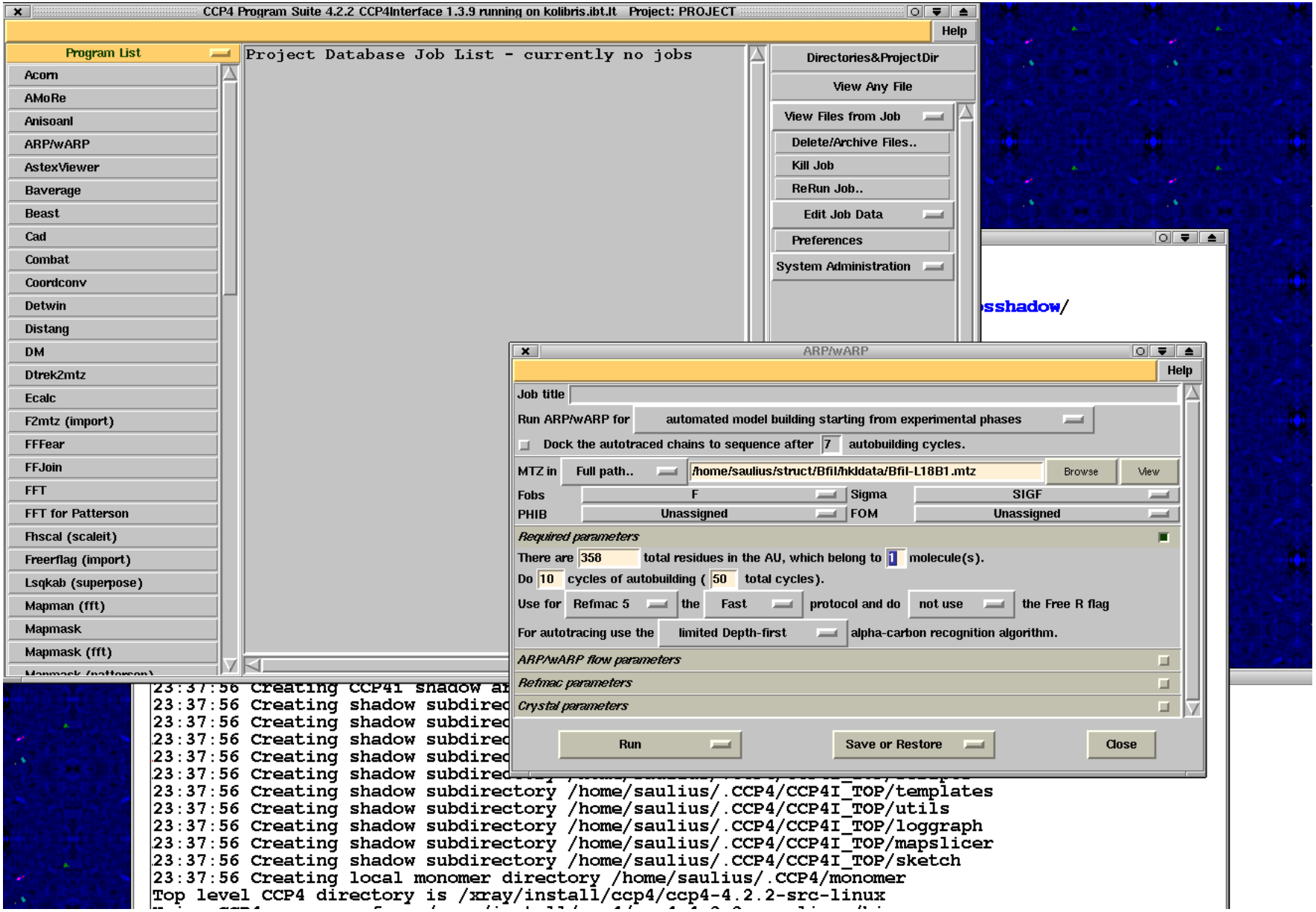
Electron density to model



Rankinis modelio konstravimas



Automatiniai modelio konstravimo įrankiai



Prior information is essential!

- Number of molecules in the asymmetric unit
- Protein size and sequence
- Structures of amino acids ;-)
- Composition of crystallisation conditions (ions, buffers)
- Protein production and purification method

Structure refinement

- Lest squares method
- Maximum likelihood methods
- Bayesian methods

Refinement criteria

Crystallographic R-factor:

$$R_{cryst} = \sum_{h\,k\,l} |F_{hkl}^{obs} - F_{hkl}^{calc}| / \sum_{h\,k\,l} |F_{hkl}^{obs}|$$

R-free:

hkl → ~10% (test set)

$$R_{free} = \sum_{h\,k\,l \in Test\ set} \ldots$$

hkl → ~90% (working set)

$$R_{cryst} = \sum_{h\,k\,l \in Working\ set} \ldots$$

Parameters

- Aromic coordinates
- Thermal displacement parameter(s) (B-factors)
- Occupancies

- Important to know: the observation-to-parameter ratio

Structure quality indicators

- Resolution (Å)
- R factor (full, crystallographic), R_{cryst}
- „Free“ R factor, R_{free}
- R_{free} - R_{cryst}, R_{cryst}/R_{free}
- R_{cryst} and R_{free} in the outer resolution shell
- Average error of bond lengths and angles

Non-polymeric molecules

- Solvent (water!)
- Ions
- Ligands, cofactors
- Modified amino acids and/or DNA/RNA bases

Evidence of chaos

- Zero occupancies
- Huge B-factors
- Residues without the side chains

Resources in the Net

Protein Data Bank (PDB)
http://www.rcsb.org/pdb/

Macromolecular Structure Database
http://www.ebi.ac.uk/msd/

International Union of Crystallography
http://www.iucr.org/

Crystallography Open Database (COD)
http://www.crystallography.net/