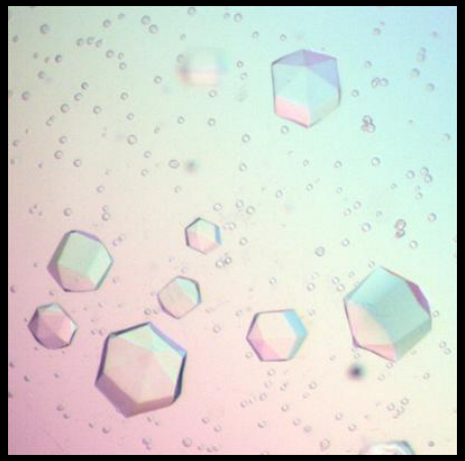
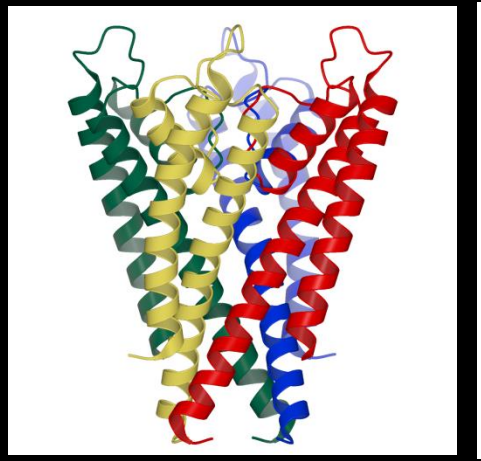
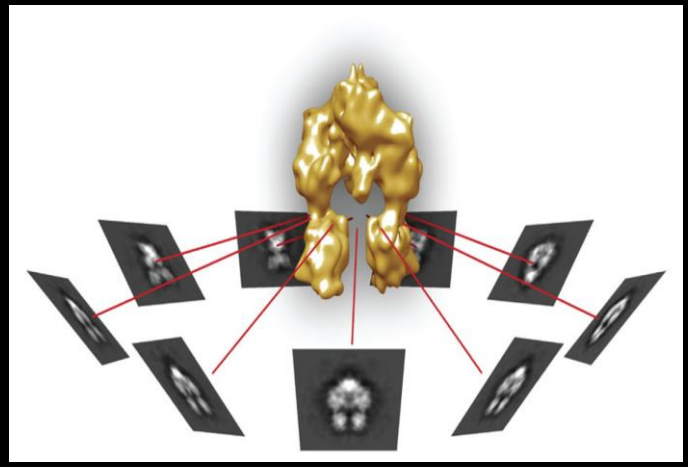




Structural Data:
Introduction to X-ray Crystallography & Cryo-EM

Jesse Rinehart, PhD

Biomedical Data Science: Mining & Modeling

CBB 752, Spring 2019

Cellular & Molecular Physiology
Yale University School of Medicine

YALE SYSTEMS BIOLOGY INSTITUTE

Thank you to **Yong Xiong** and **Fred Sigworth** for contributions to this lecture

Yale Structure Courses:

C&MP 710b/MB&B 710b4, Electron Cryo-Microscopy for Protein Structure Determination

MB&B 720a, Macromolecular Structure and Biophysical Analysis

MB&B 721b, Macromolecular Interactions and Dynamic Properties

MB&B 760b3: Principles of Macromolecular Crystallography

MB&B 761b4: X-ray Crystallography Workshop

Pharmacology 529b: Structural Pharmacology

Additional Resources:

"Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models"
by Gale Rhodes (Third Edition, 2006 Elsevier/Academic Press)

"Crystallography 101" http://www.ruppweb.org/Xray/101index.html

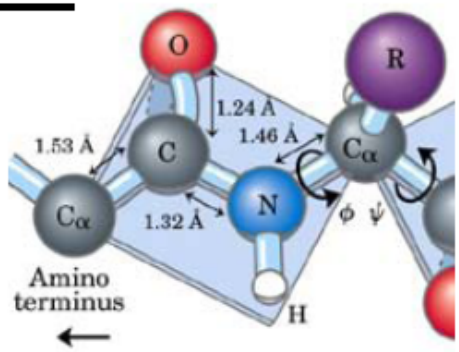
"Single particle electron cryomicroscopy: trends, issues and future perspective."
Vinothkumar KR, Henderson R. Q Rev Biophys. 2016 pubmed:27658821

“Just as we see objects around us by interpreting the light reflected from them, x-ray crystallographers "see" molecules by interpreting x-rays diffracted from them.”
- Gale Rhodes

- There's a limit to how small an object can be seen under a light microscope.
- The diffraction limit: you can not image things that are much smaller than the wavelength of the light you are using.
- The wavelength for visible light is measured in hundreds of nanometers, while atoms are separated by distances of the order of 0.1nm, or 1Å.

We need to use x-rays to resolve atomic features.

Distances between atoms are small:
Lab x-ray sources use CuKα radiation. Wavelength = 1.54 Å.
Synchrotron radiation wavelengths in the range 0.5 Å - 2.5 Å.



Yong Xiong

The 2014 Nobel Prize in Chemistry: Eric Betzig, W.E. Moerner, and Stefan Hell
"The development of super-resolved fluorescence microscopy"

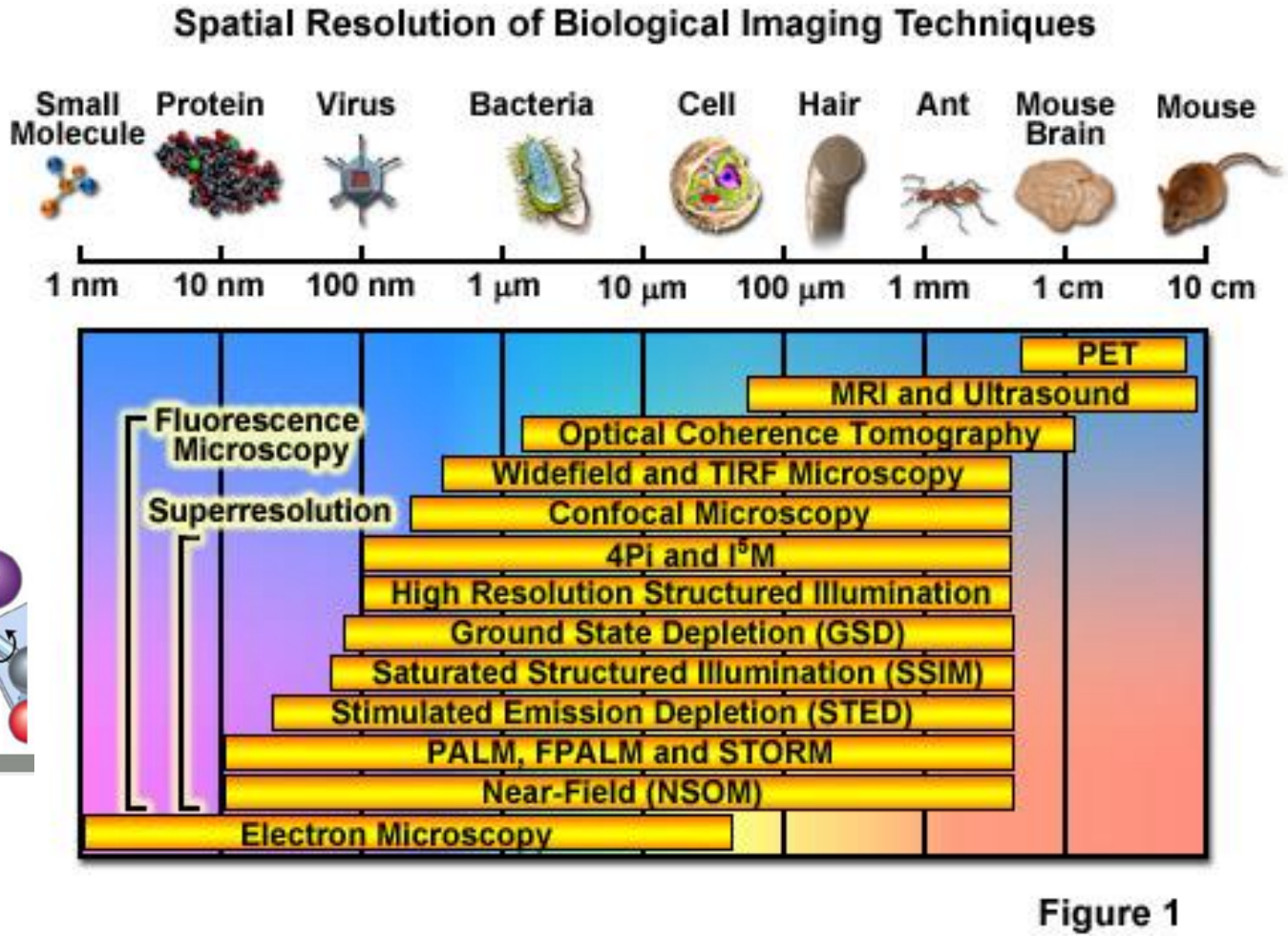
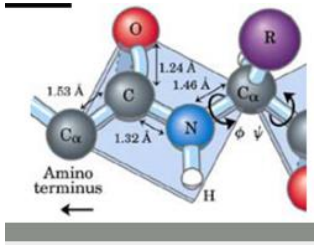


Figure 1

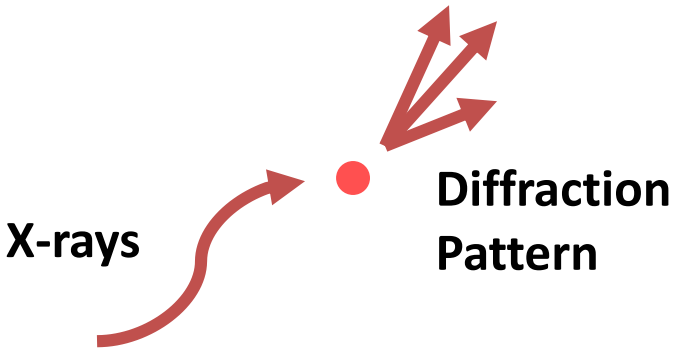
1Å = 0.1nm



http://zeiss-campus.magnet.fsu.edu/articles/superresolution/introduction.html

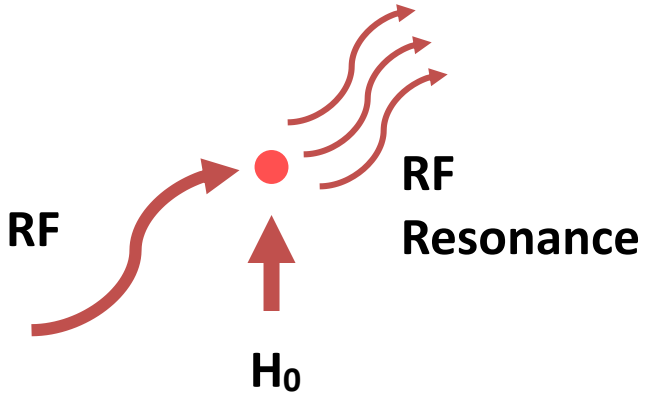
Experimental Determination of Atomic Resolution Structures

X-ray



- Direct detection of atom positions
- Crystals

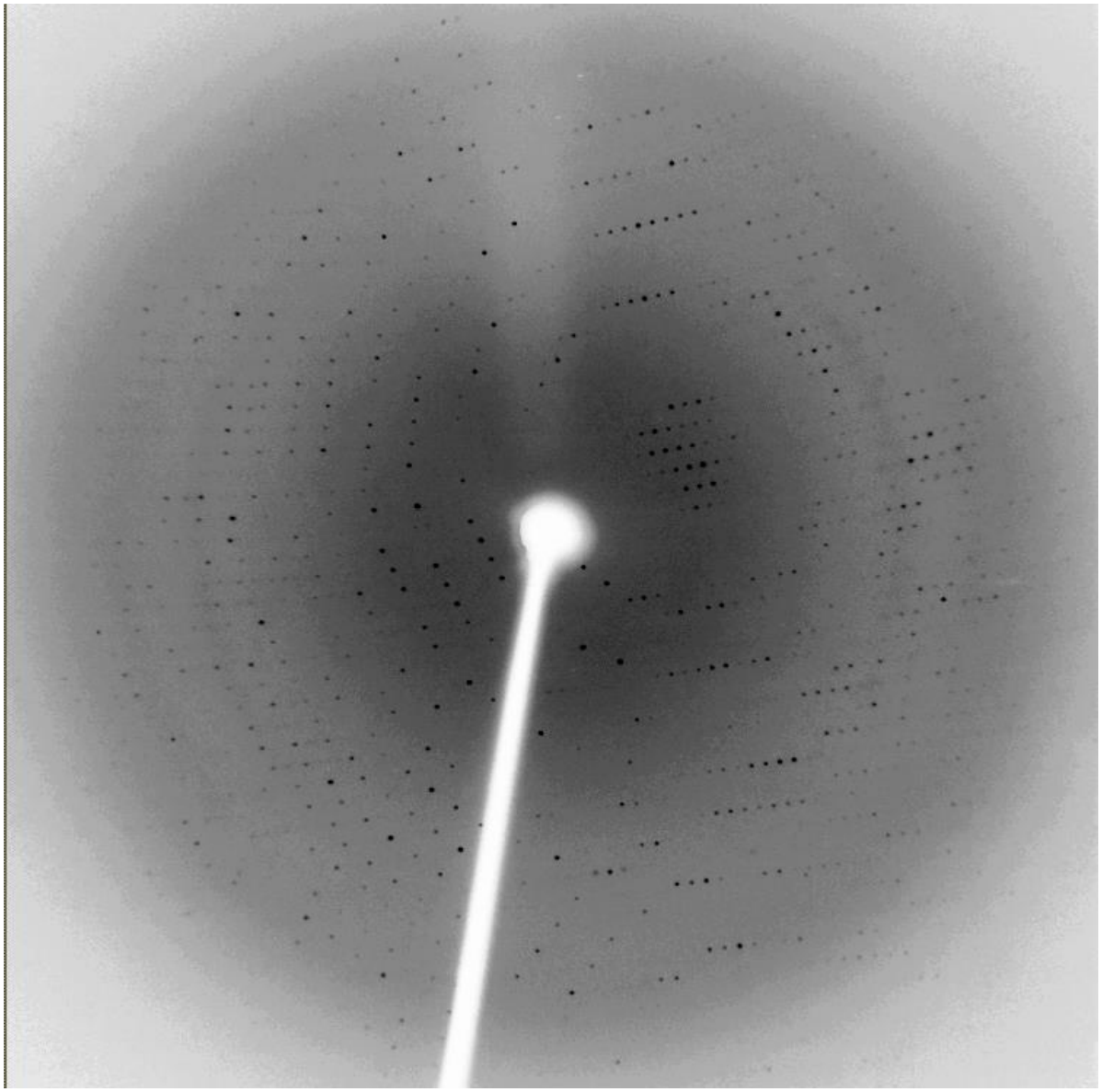
NMR



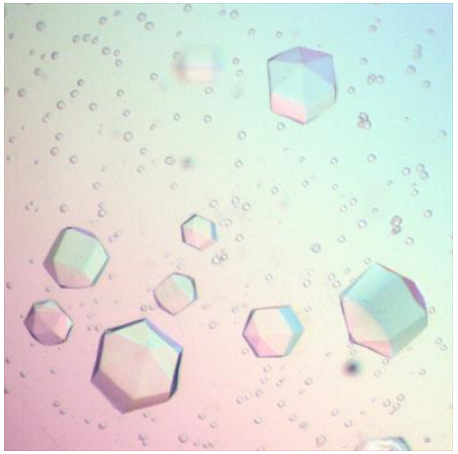
- Indirect detection of H-H distances
- In solution

Other methods for determining protein structures:
-EM (Electron Microscopy), Cryo-EM, ESR/Fluorescence

Image of X-ray diffraction of a protein crystal



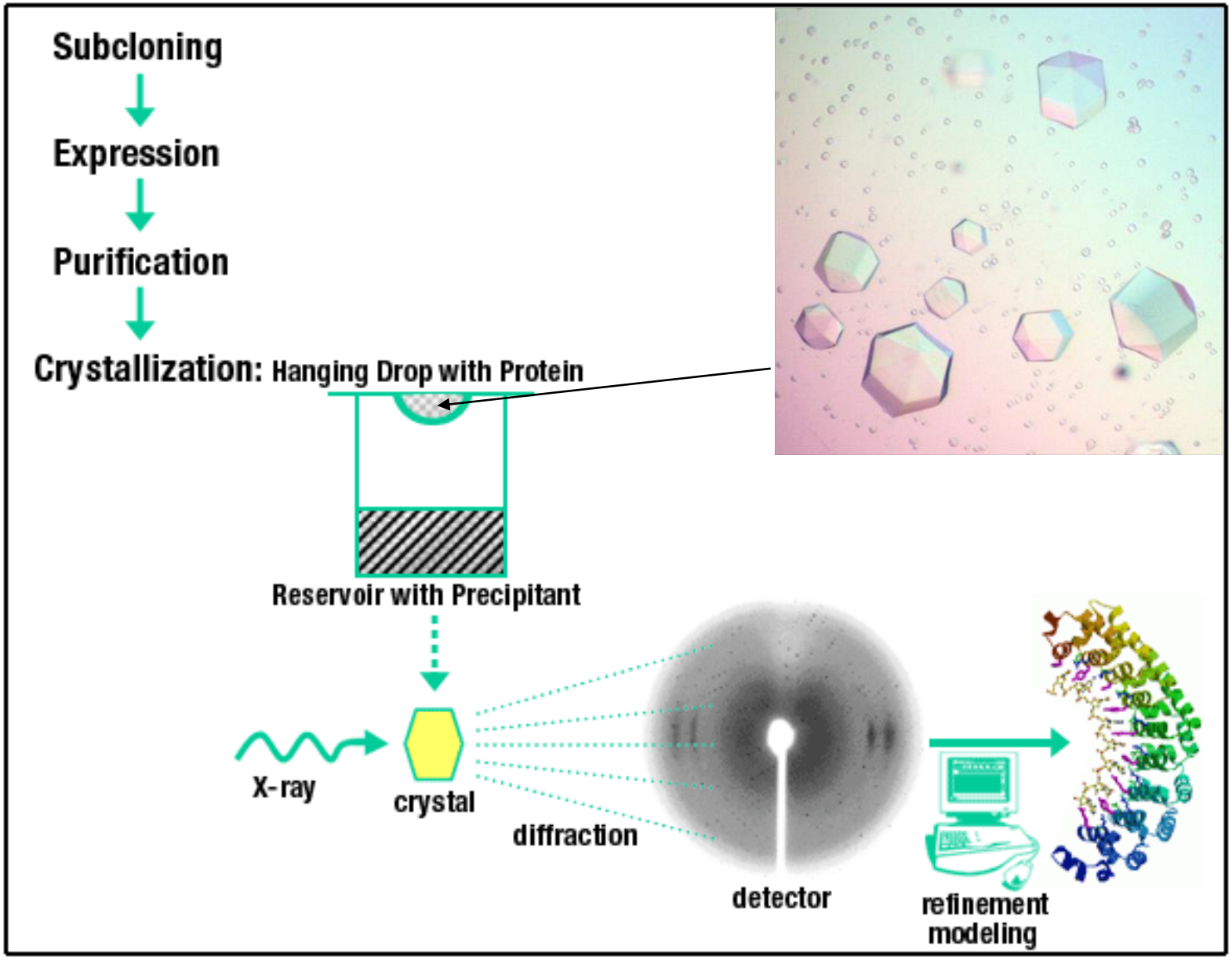
Why Crystals?



X-rays are scattered by electrons, too weak to record scattering from a single molecule. Crystals are therefore used because they present many molecules (N) in exactly the same orientation. The scattering from each of the N molecules interferes constructively to give a measurable diffraction pattern (enhanced ~N^2 fold).

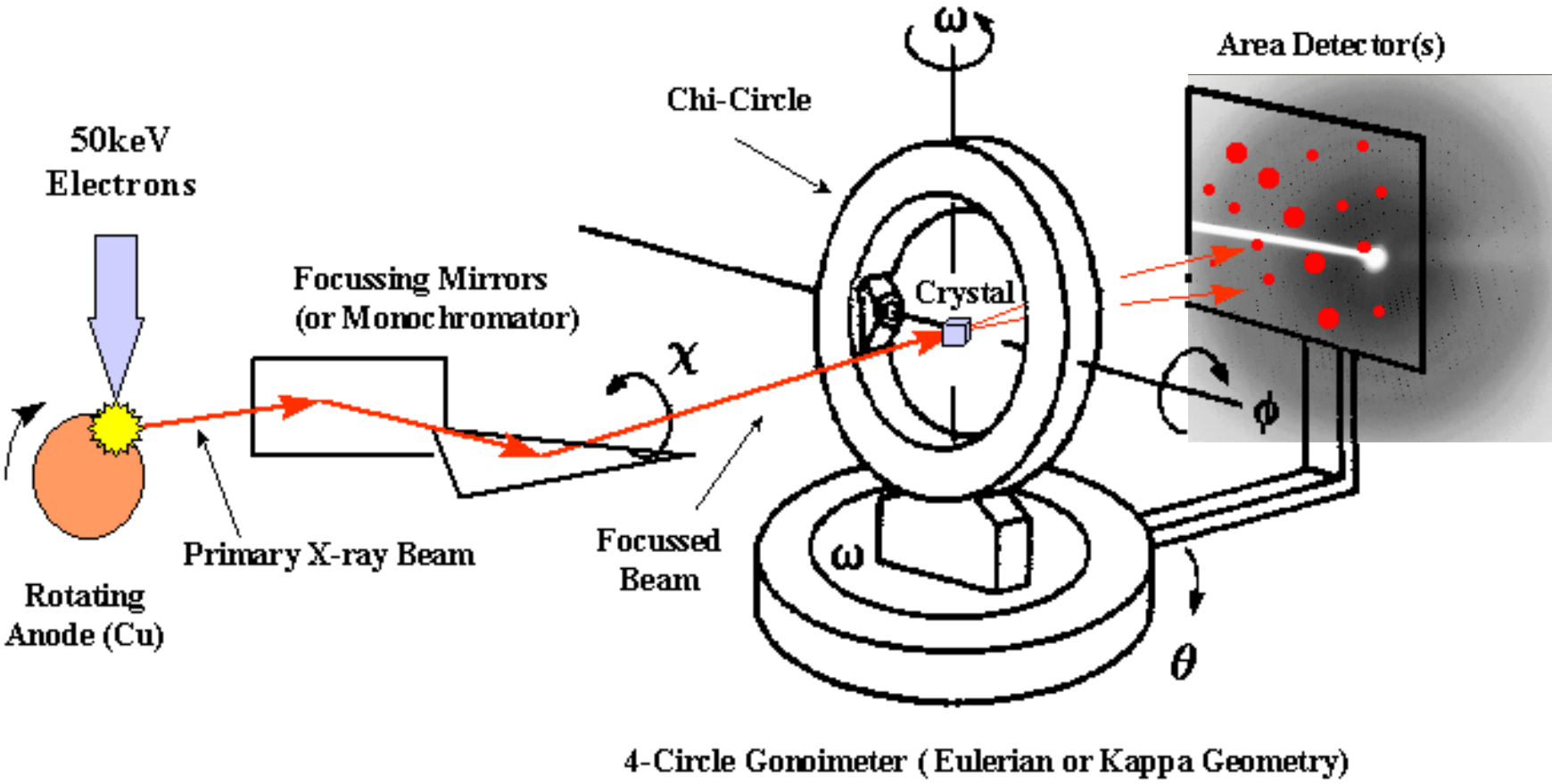
Yong Xiong

Determination of Protein Crystal Structure



http://www.noble.org/PlantBio/Wang/crystallography.html

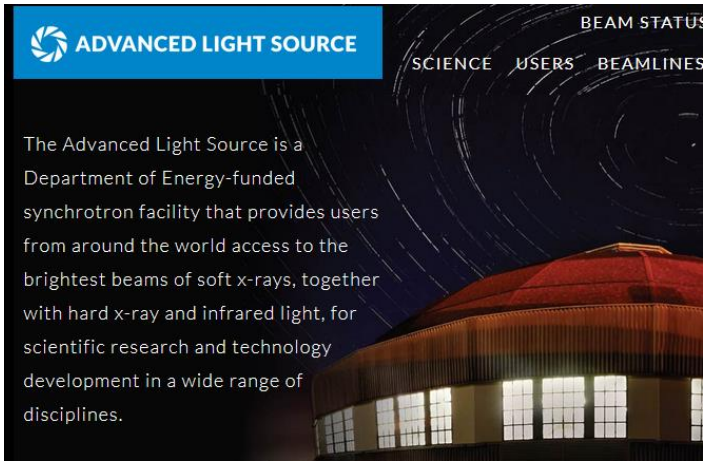
Data Collection



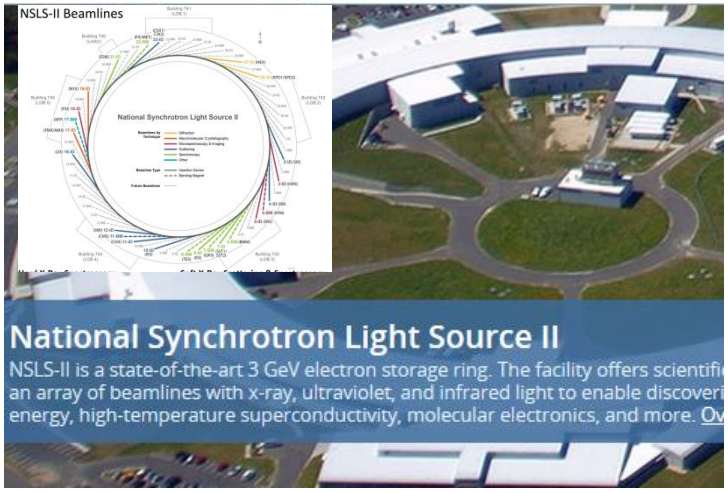
Synchrotron X-ray Sources are the method of choice

Lab x-ray sources @ 1.54 Å compared to Synchrotron X-ray @ 0.5 Å - 2.5 Å.

APS Chicago



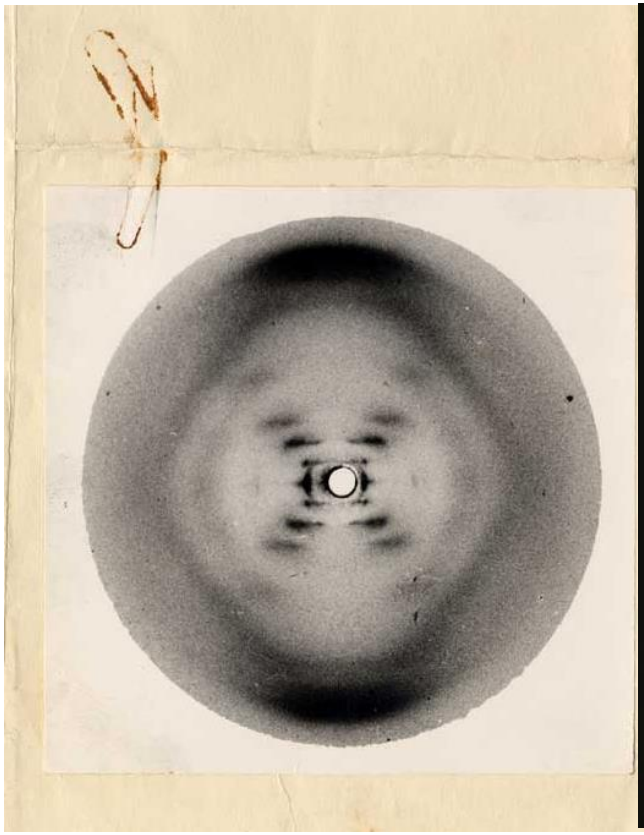
ALS Berkeley

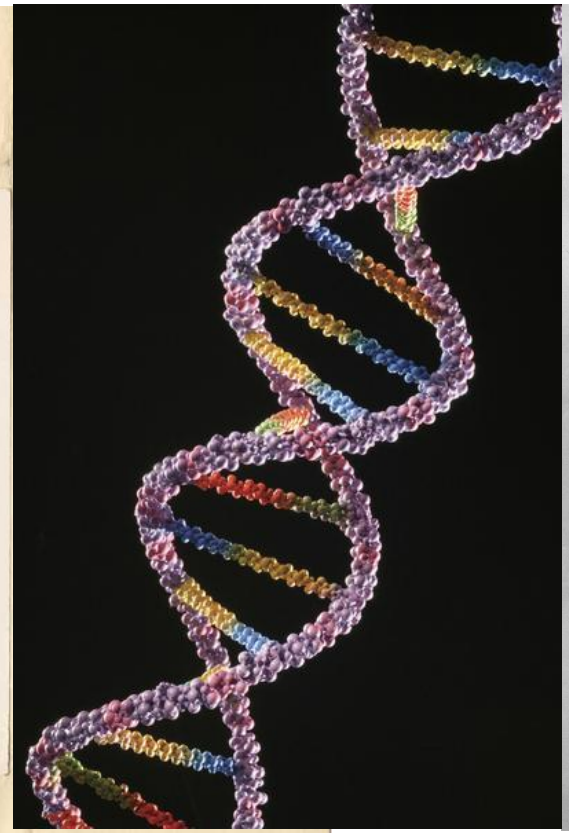


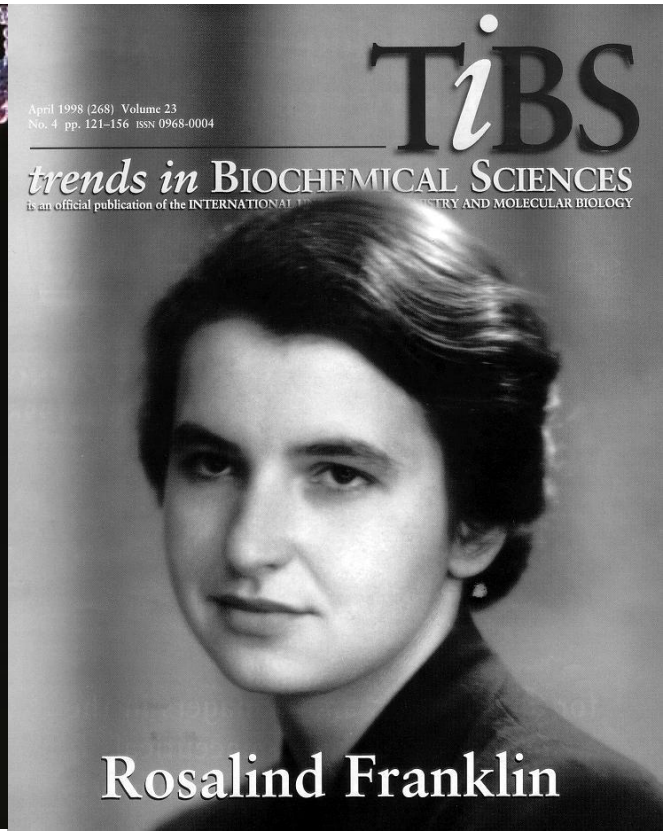
NSLS-II Brookhaven

CHESS Ithaca

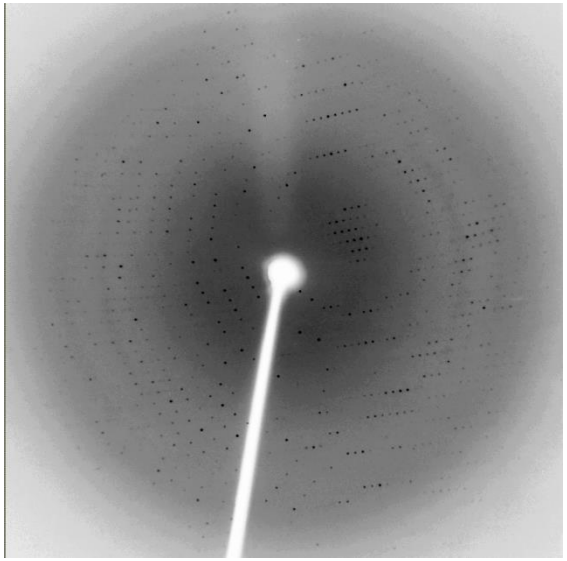
Most famous X-ray diffraction pattern







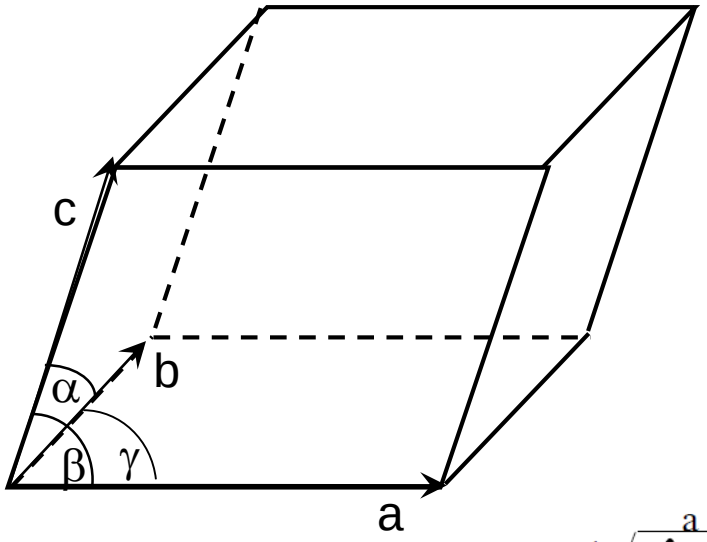
The information we get from a single diffraction experiment



Analyze the pattern of the reflections

(a) space group of the crystal

(b) unit cell dimensions

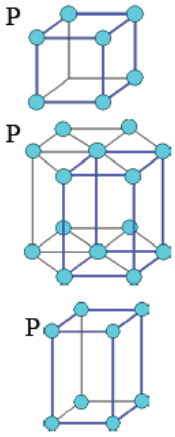


Cubic
a = b = c,
α = β = γ = 90°

Hexagonal
a = b ≠ c,
α = β = 90°, γ = 120°

Trigonal
a = b ≠ c,
α = β = 90°, γ = 120°

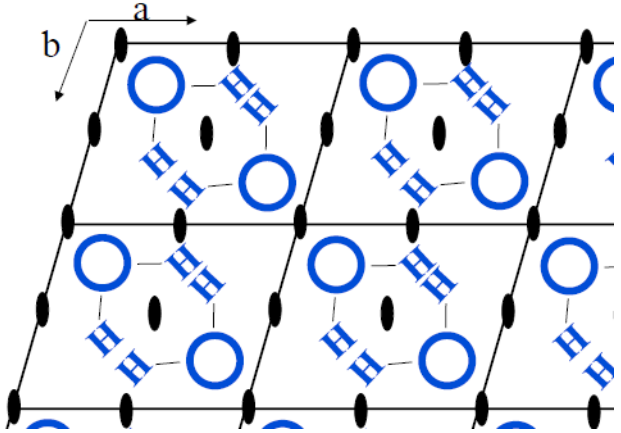
Tetragonal
a = b ≠ c,
α = β = γ = 90°

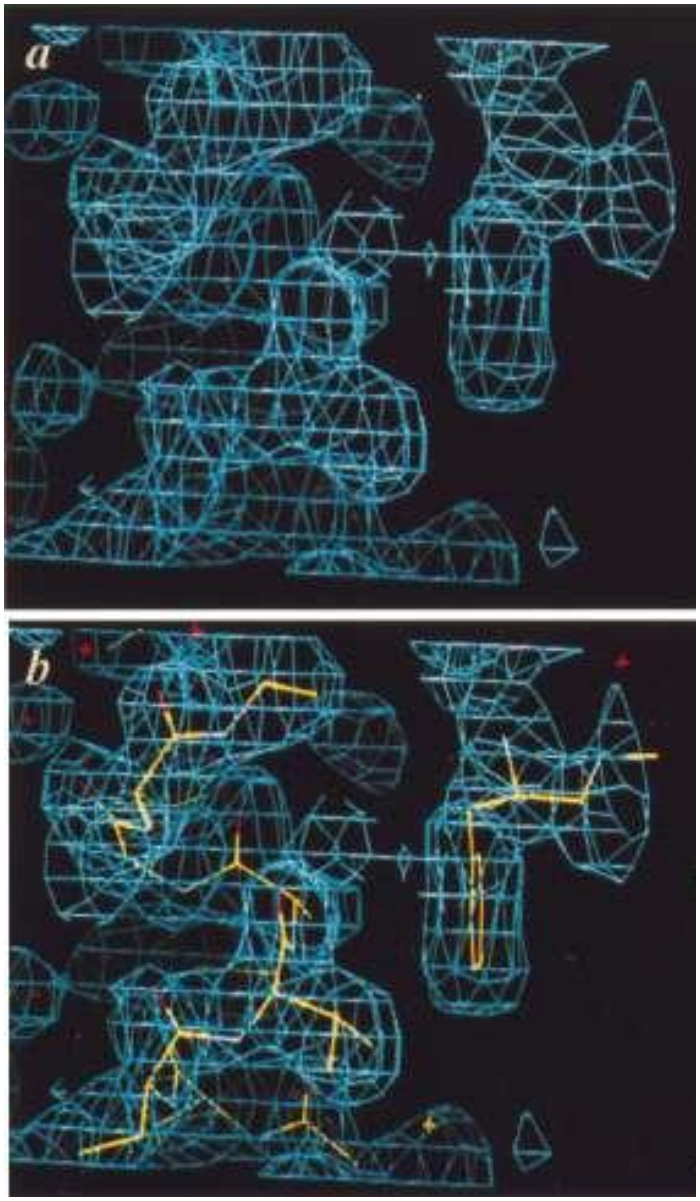


How to understand symmetry?

Crystal = lattice + unit cell content

(asymmetric units (asu) content)

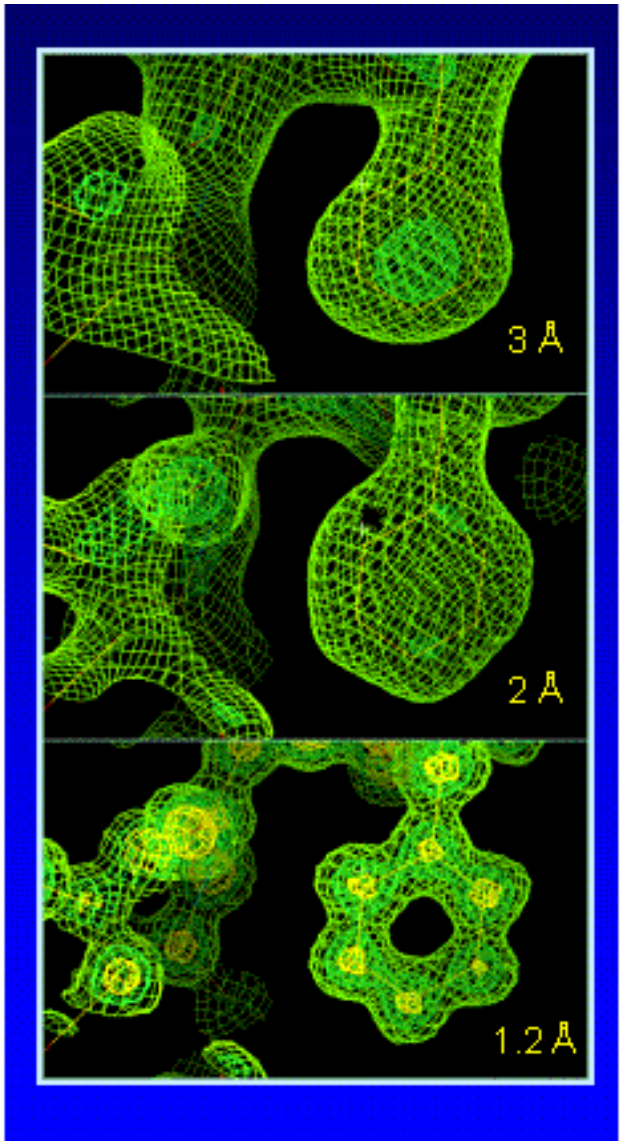


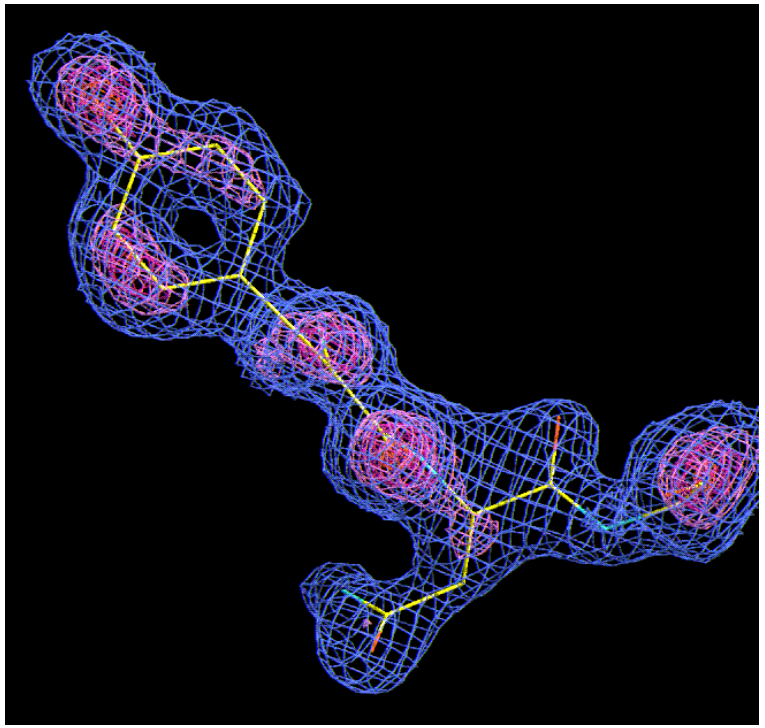


Electron density map

Building a structure model

The importance of resolution





Experimental electron density map created from multi-wavelength data collected at SSRL beam line 1-5 on a Gold derivative of tetanus C fragment.

Example of high quality Experimental data where very little refinement has been applied to fit a tyrosine into the density map.

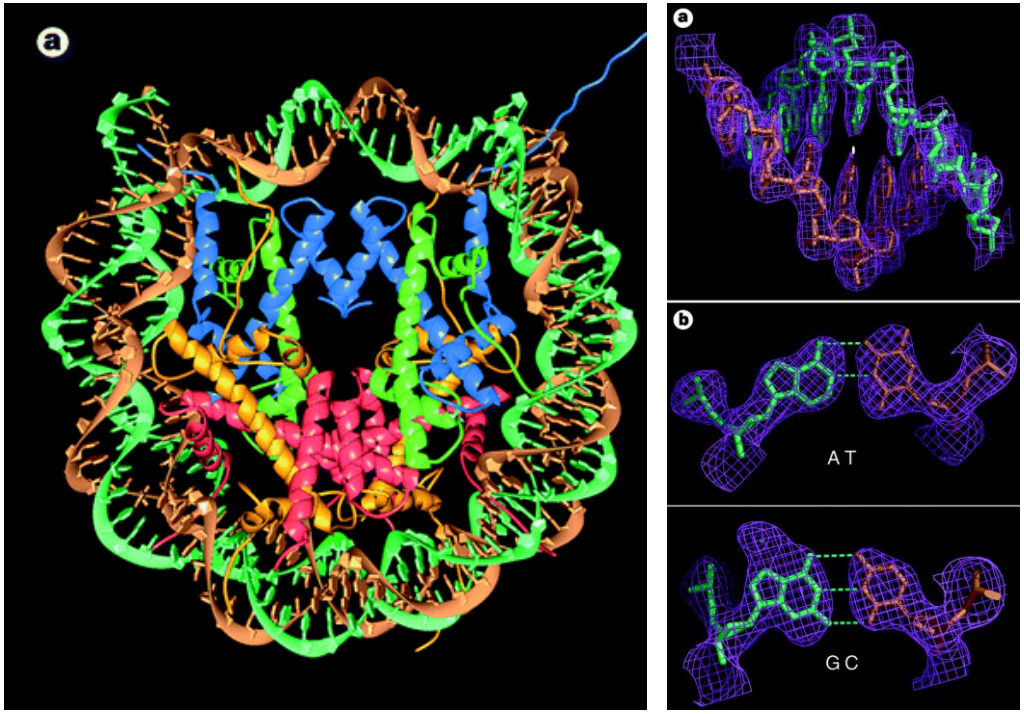
http://www.ruppweb.org/Xray/101index.html

Crystal structure of the nucleosome core particle at 2.8 Å resolution

Karolin Luger, Armin W. Mäder, Robin K. Richmond, David F. Sargent & Timothy J. Richmond

Institut für Molekularbiologie und Biophysik ETHZ, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

The X-ray crystal structure of the nucleosome core particle of chromatin shows in atomic detail how the histone protein octamer is assembled and how 146 base pairs of DNA are organized into a superhelix around it. Both histone/histone and histone/DNA interactions depend on the histone fold domains and additional, well ordered structure elements extending from this motif. Histone amino-terminal tails pass over and between the gyres of the DNA superhelix to contact neighbouring particles. The lack of uniformity between multiple histone/DNA-binding sites causes the DNA to deviate from ideal superhelix geometry.

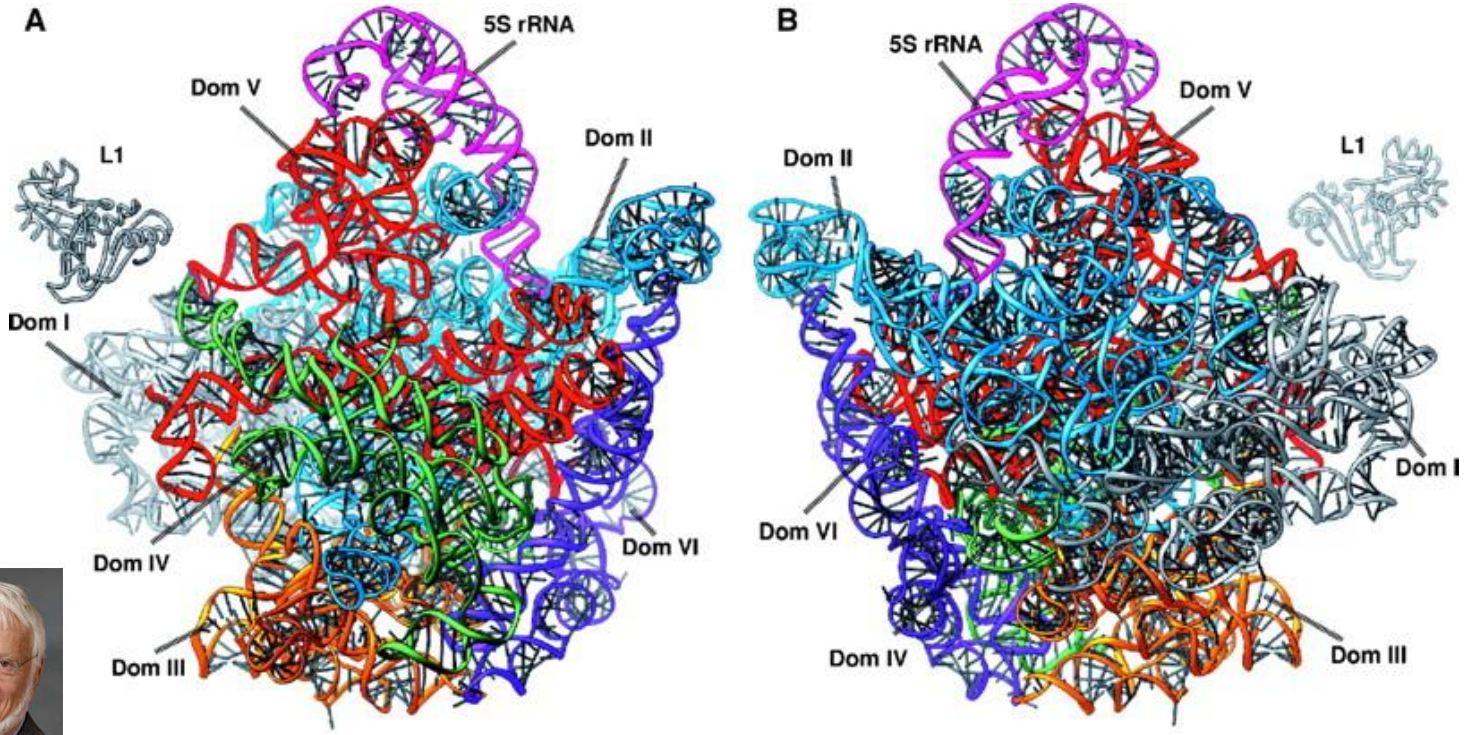


Science
AAAS

RESEARCH ARTICLES

The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution

Nenad Ban,[1]* Poul Nissen,[1]* Jeffrey Hansen,[1] Peter B. Moore,[1,2]
Thomas A. Steitz[1,2,3]†

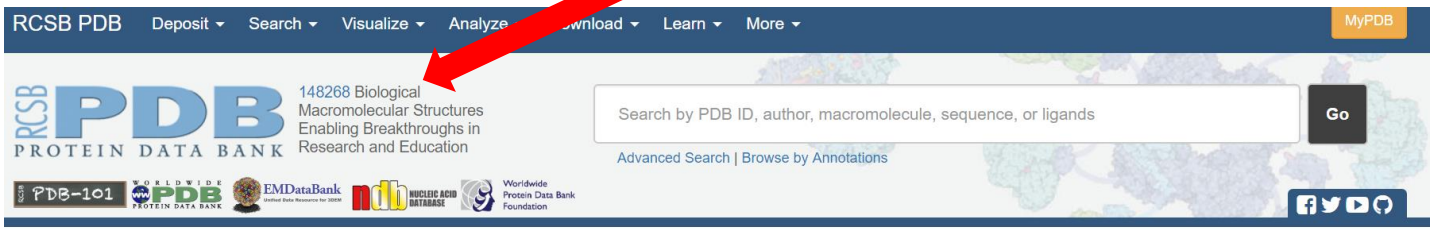


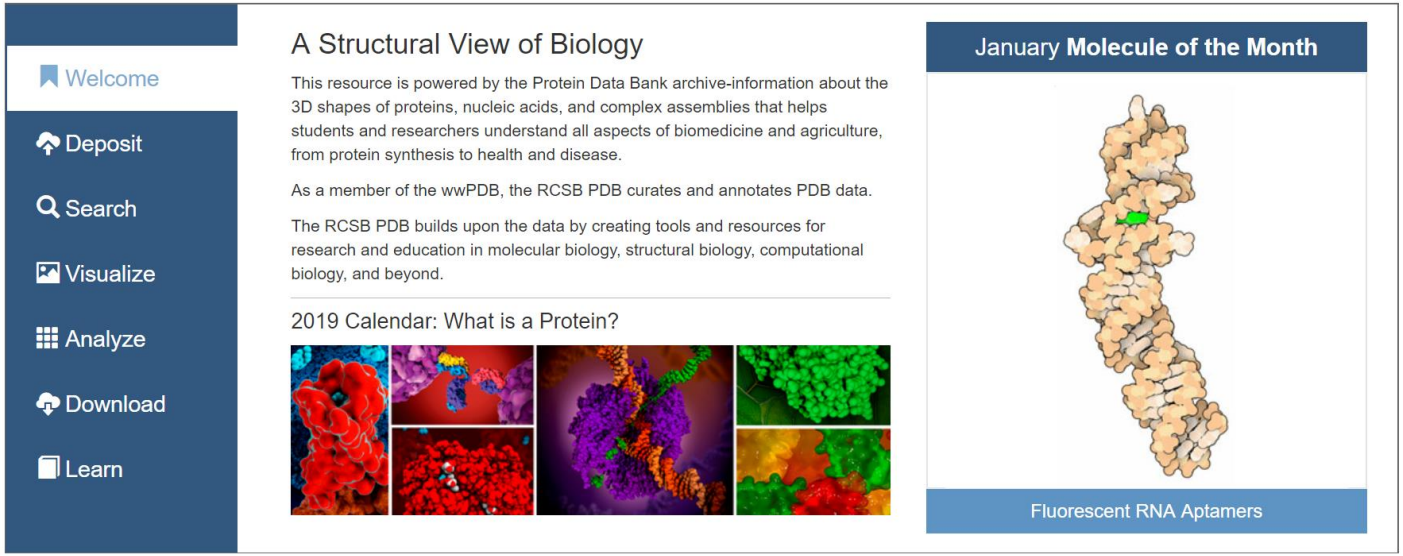
Thomas Steitz shared 2009 Nobel Prize in Chemistry for this structure

Structure Databases

- **Where does protein structural information reside?**
  - **PDB:**
    - http://www.rcsb.org/pdb/
  - **MMDB:**
    - http://www.ncbi.nlm.nih.gov/Structure/
  - **FSSP:**
    - http://www.ebi.ac.uk/dali/fssp/
  - **SCOP:**
    - http://scop.mrc-lmb.cam.ac.uk/scop/
  - **CATH:**
    - http://www.biochem.ucl.ac.uk/bsm/cath_new/

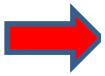
```
# of PDB structures
  2017:    117,184
  2018:    137,178
  2019:    148,268
```





https://pdb101.rcsb.org/learn/videos/what-is-a-protein-video

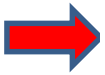
PDB: What species are the structures from?

human 

ORGANISM

Homo sapiens (42668)
Escherichia coli (9294)
Mus musculus (6313)
Saccharomyces cerevisiae (4133)
synthetic construct (3707)
Rattus norvegicus (2988)
Bos taurus (2852)
Other (77188)

Which methods?

X-ray 

EXPERIMENTAL METHOD

X-ray (132583) Resolution range 15 - 0.48 Å
Solution NMR (12391)
Electron Microscopy (2783) Resolution range 70 - 1.8 Å
Hybrid (138)
Electron Crystallography (112)
Solid-State NMR (101)
Neutron Diffraction (66)
Fiber Diffraction (38)
Solution Scattering (32)
Other (24)

http://www.rcsb.org/pdb/home/home.do

PDB X-ray Structures:

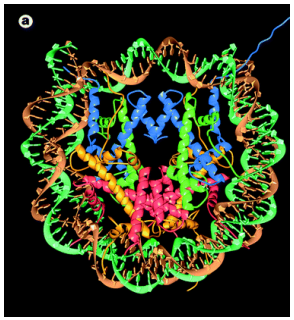
http://www.rcsb.org/pdb/results/results.do?outformat=&qrid=1B04C26E&tabtoshow=Current

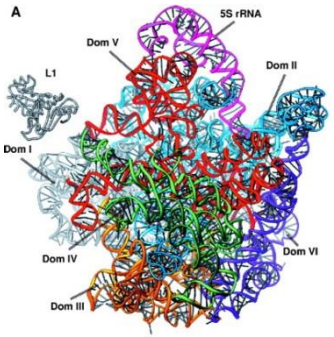
ORGANISM

Homo sapiens (37692)
Escherichia coli (8330)
Mus musculus (5352)
Saccharomyces cerevisiae (3437)
synthetic construct (3305)
Rattus norvegicus (2623)
Bos taurus (2570)
Other (reached drill-down ... (71122)

POLYMER TYPE

Protein (124178)
Mixed (6508)
DNA (1074)
RNA (819)





MEMBRANE PROTEINS

ALPHA-HELICAL (3071)
BETA-BARREL (914)
MONOTOPIC MEMBRANE PROTEINS (486)

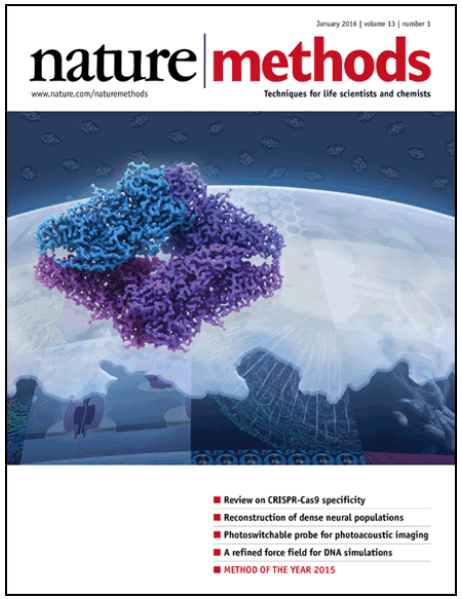
Small % of the total x-ray data

Tools for Viewing Structures

- **Jmol**
  - **http://jmol.sourceforge.net**
- **PyMOL**
  - **http://pymol.sourceforge.net**
- **Swiss PDB viewer**
  - **http://www.expasy.ch/spdbv**
- **Mage/KiNG**
  - **http://kinemage.biochem.duke.edu/software/mage.php**
  - **http://kinemage.biochem.duke.edu/software/king.php**
- **Rasmol**
  - **http://www.umass.edu/microbio/rasmol/**

Cryo-EM for biomolecular structures

2015 Method of the Year: Single-particle Cryo-EM



METHOD OF THE YEAR 2015

At *Nature Methods* we are ringing in a new year with our celebration of single-particle cryo-electron microscopy (cryo-EM) as our Method of the Year 2015. Cryo-EM has its roots in work first performed in the 1960s. It has steadily progressed over the past few decades as a medium-resolution structural technique for obtaining information about macromolecular samples that resist analysis by X-ray crystallography. But very recent technical advances, especially the development of direct-detection cameras, have enabled the field to achieve impressive leaps in resolution—even reaching the near-atomic realm of X-ray crystallography—and, by extension, biological applicability. An Editorial, News Feature, Primer, Historical Commentary and Commentary discuss how cryo-EM works, what it is used for, how the field began, why now is such an exhilarating time, and where the field is going in the future. We also cast our predictions about methods with exciting potential in our Methods to Watch section.

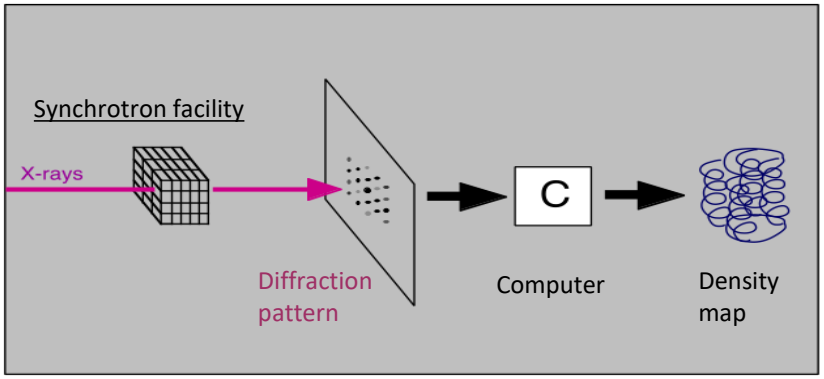
Special feature starts on p19

2017 Nobel Prize in Chemistry

"for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution"

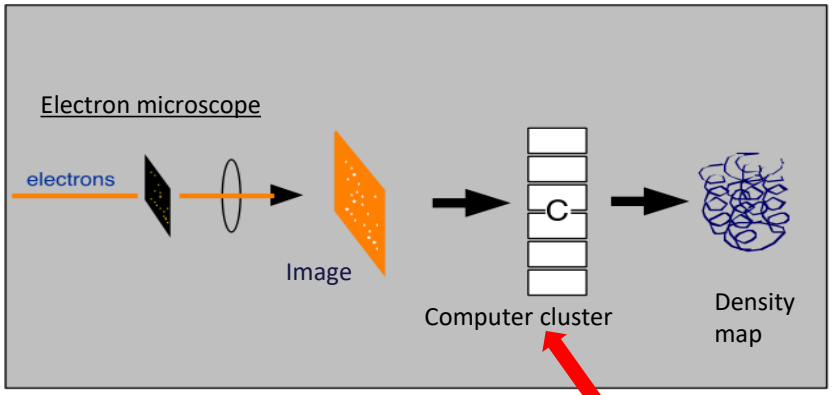
Jacques Dubochet (University of Lausanne, Switzerland)
Joachim Frank (Columbia University, New York, USA)
Richard Henderson (MRC Laboratory of Molecular Biology, Cambridge, UK)

Two methods for structure determination



X-ray crystallography

Well-established (since 1960s)
Requires well-ordered crystals
>10^{12} copies of protein



Single-particle cryo-EM

Recent (1990s-present)
No crystals required!
~10^{5} copies of protein

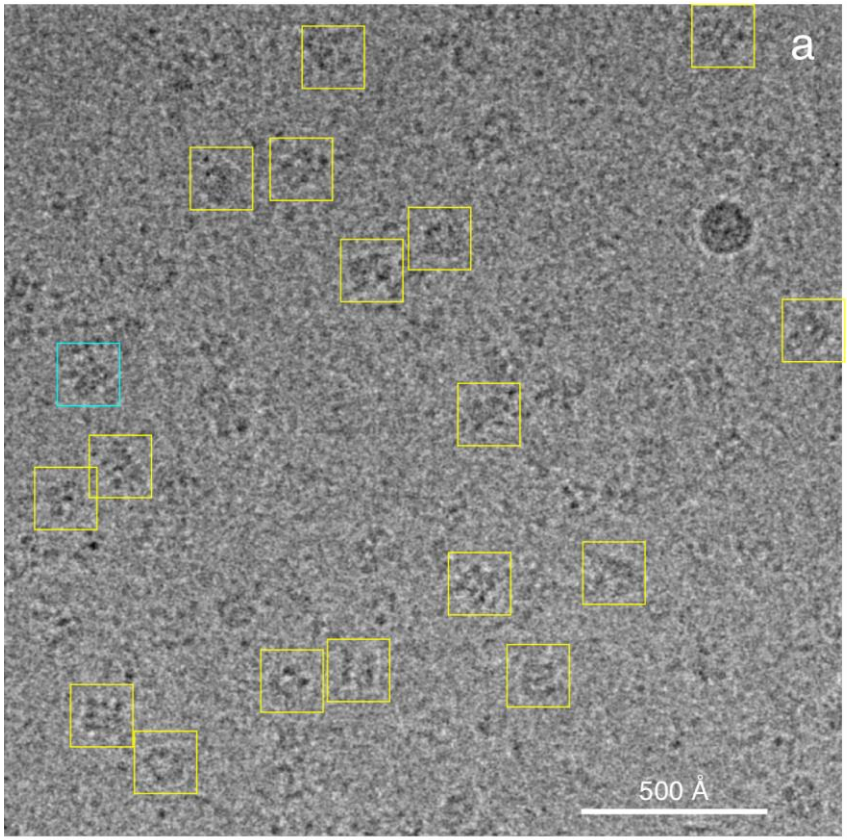
Some gaming PCs can now replace the cluster ☺

Fred Sigworth

The Cryo-EM specimen gives only a phase contrast image

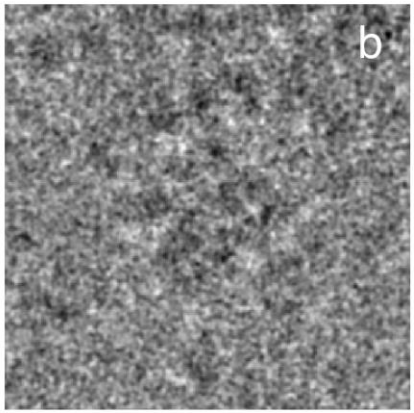
- A constellation of images and data processing are essential.

1/4 of a micrograph, showing some particles

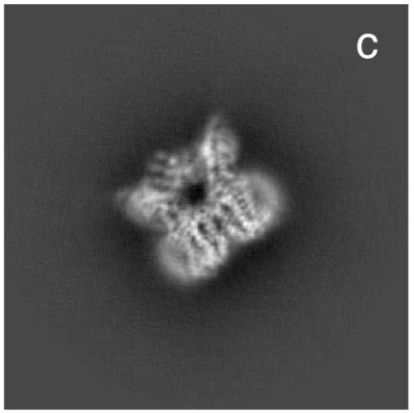


Y. Cheng and D. Julius lab. Nature 2013

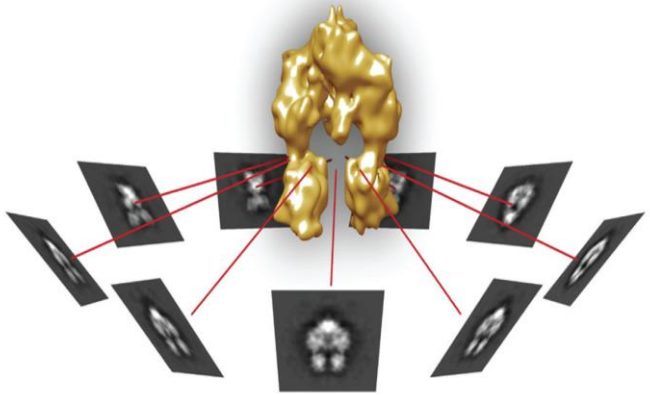
Image



Projection



- orientation assignment and averaging
- 3D reconstruction



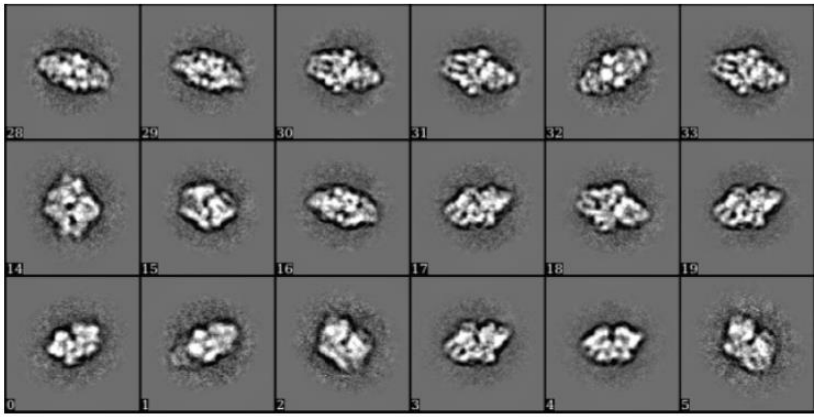
Fred Sigworth

A landmark study for high-resolution single-particle structures

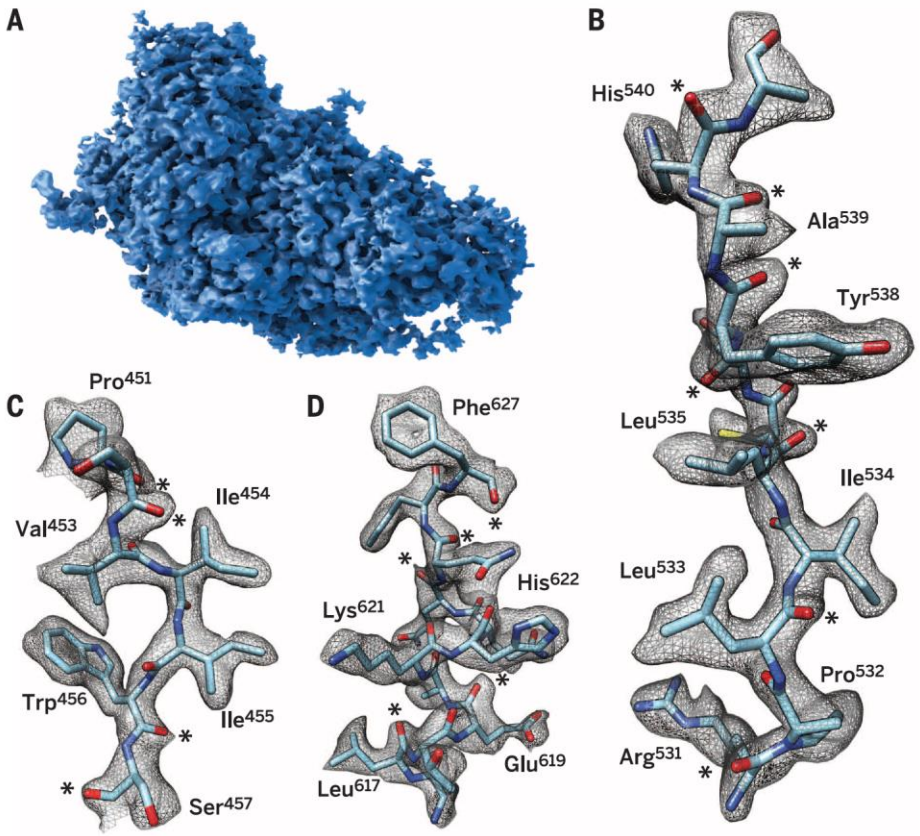
2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor

Alberto Bartesaghi,[1]* Alan Merk,[1]* Soojay Banerjee,[1] Doreen Matthies,[1] Xiongwu Wu,[2] Jacqueline L. S. Milne,[1] Sriram Subramaniam[1]†

Science 2015



2D class averages



Fred Sigworth

New Technologies, Automation, & Computation are accelerating the field

Krios at National University of Singapore

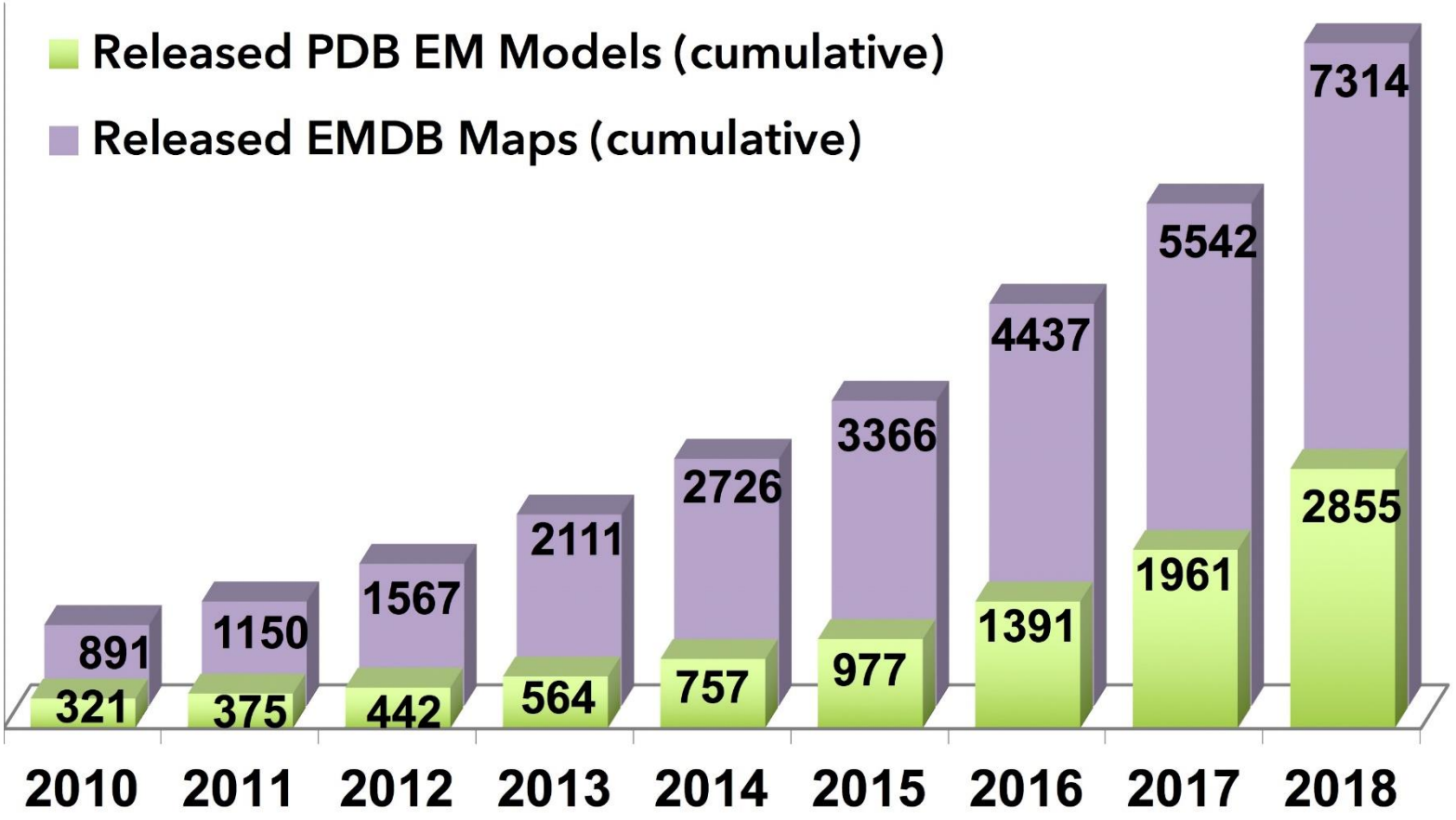
Control room at Scripps Research Institute, La Jolla

Krios TEM installation on Yale's West Campus.

Fred Sigworth

EMDB statistics

https://www.ebi.ac.uk/pdbe/emdb/



Source: http://www.emdatabank.org/statistics.html

Cryo-EM: membrane proteins, protein complexes, proteins difficult to crystalize

Science

RESEARCH ARTICLES

Cite as: R. Zhou *et al.*, *Science* 10.1126/science.aaw0930 (2019).

Recognition of the amyloid precursor protein by human γ-secretase

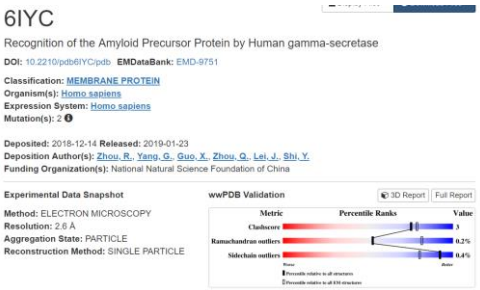
Rui Zhou[1]*, Guanghui Yang[1]*, Xuefei Guo[1], Qiang Zhou[2,3], Jianlin Lei[1,4], Yigong Shi[1,2]†

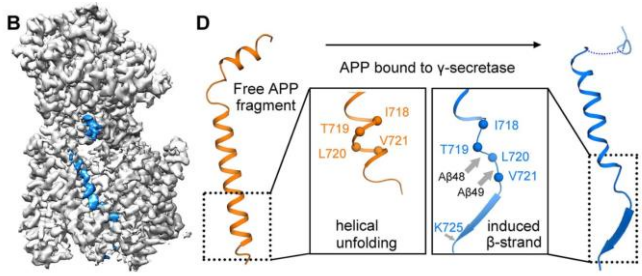
[1]Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China. [2]Institute of Biology, Westlake Institute for Advanced Study, Westlake University, 18 Shilongshan Road, Xihu District, Hangzhou 310024, Zhejiang Province, China. [3]School of Life Sciences, Westlake University, 18 Shilongshan Road, Xihu District, Hangzhou 310024, Zhejiang Province, China. [4]Technology Center for Protein Sciences, Ministry of Education Key Laboratory of Protein Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China.

*These authors contributed equally to this work.

†Corresponding author. Email: shi-lab@tsinghua.edu.cn

Cleavage of amyloid precursor protein (APP) by the intramembrane protease γ-secretase is linked to Alzheimer's disease. We report an atomic structure of human γ-secretase in complex with a transmembrane APP fragment at 2.6-Å resolution. The transmembrane helix (TM) of APP closely interacts with five surrounding TMs of PS1 (the catalytic subunit of γ-secretase). A hybrid β-sheet, which is formed by a β-strand from APP and two β-strands from PS1, guides γ-secretase to the scissile peptide bond of APP between its TM and β-strand. Residues at the interface between PS1 and APP are heavily targeted by recurring mutations from AD patients. This structure, together with that of γ-secretase bound to Notch, reveal contrasting features of substrate binding, which may be exploited toward design of substrate-specific inhibitors.





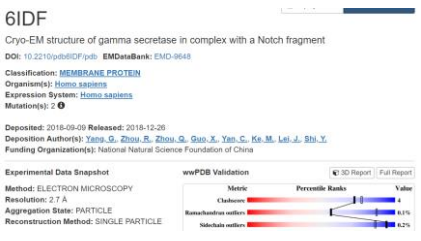
ARTICLE

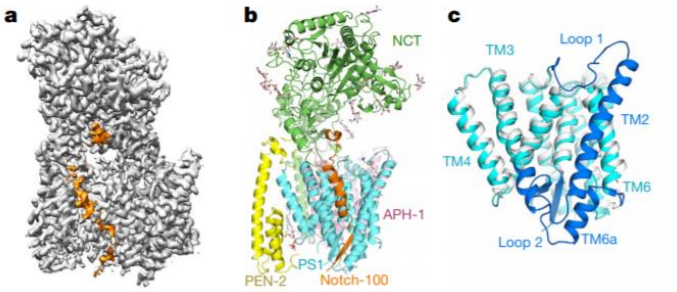
https://doi.org/10.1038/s41586-018-0813-8

Structural basis of Notch recognition by human γ-secretase

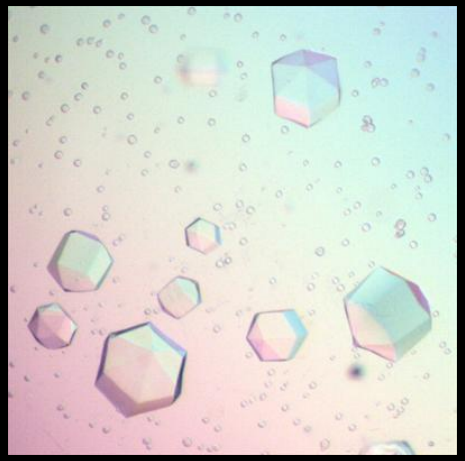
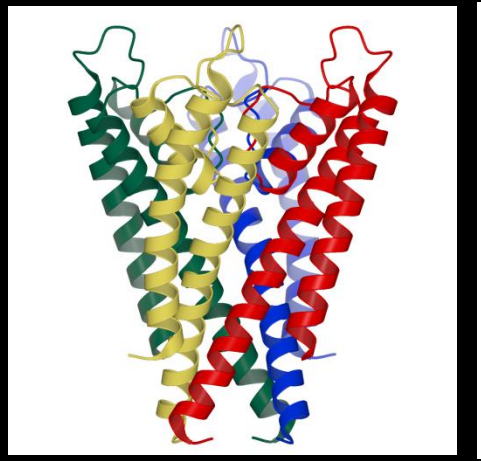
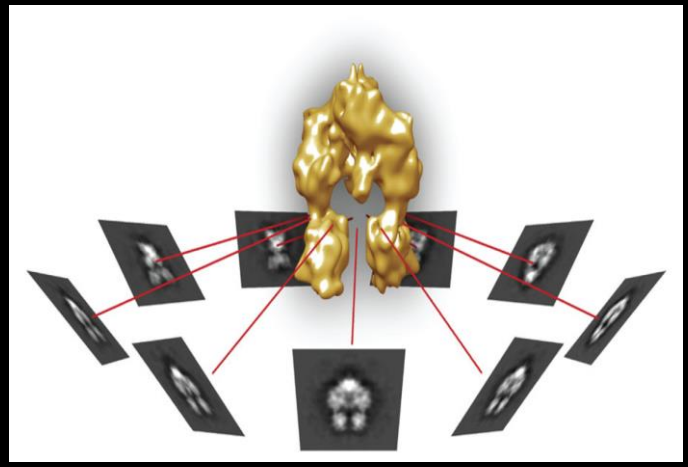
Guanghui Yang[1,4], Rui Zhou[1,4], Qiang Zhou[1,2], Xuefei Guo[1], Chuangye Yan[1], Meng Ke[1], Jianlin Lei[1,3] & Yigong Shi[1,2]*

Aberrant cleavage of Notch by γ-secretase leads to several types of cancer, but how γ-secretase recognizes its substrate remains unknown. Here we report the cryo-electron microscopy structure of human γ-secretase in complex with a Notch fragment at a resolution of 2.7 Å. The transmembrane helix of Notch is surrounded by three transmembrane domains of PS1, and the carboxyl-terminal β-strand of the Notch fragment forms a β-sheet with two substrate-induced β-strands of PS1 on the intracellular side. Formation of the hybrid β-sheet is essential for substrate cleavage, which occurs at the carboxyl-terminal end of the Notch transmembrane helix. PS1 undergoes pronounced conformational rearrangement upon substrate binding. These features reveal the structural basis of Notch recognition and have implications for the recruitment of the amyloid precursor protein by γ-secretase.





PMID: 30630874; PMID: 30598546; PMID: 25918421

Structural Data:
Introduction to X-ray Crystallography & Cryo-EM

Jesse Rinehart, PhD

Biomedical Data Science: Mining & Modeling

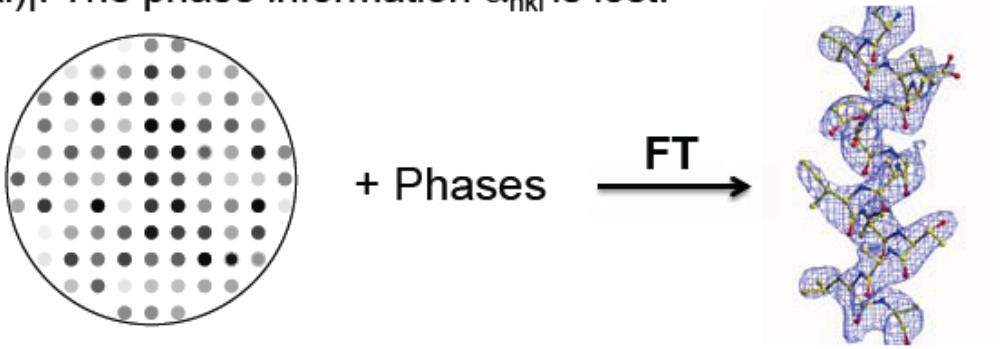
CBB 752, Spring 2019

Cellular & Molecular Physiology
Yale University School of Medicine

YALE SYSTEMS BIOLOGY INSTITUTE

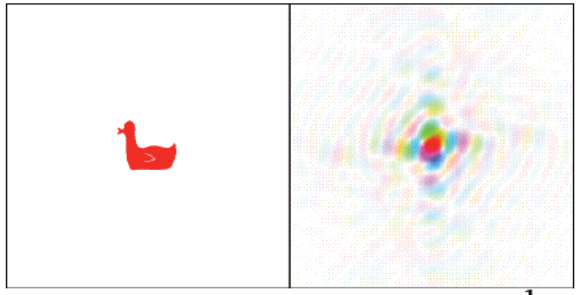
Appendix

The phase problem: F(hkl) is a complex vector. Measured diffraction data give the amplitude |F(hkl)|. The phase information α_{hkl} is lost!

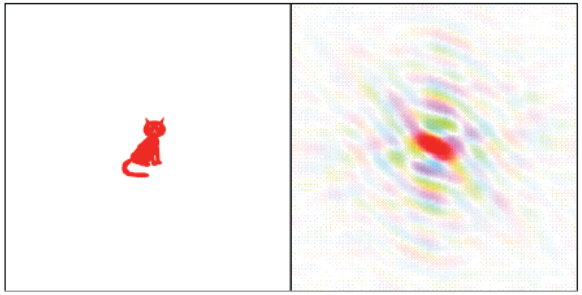


How important are amplitude and phase?

Fourier Duck and his Fourier transform
Phase is color coded

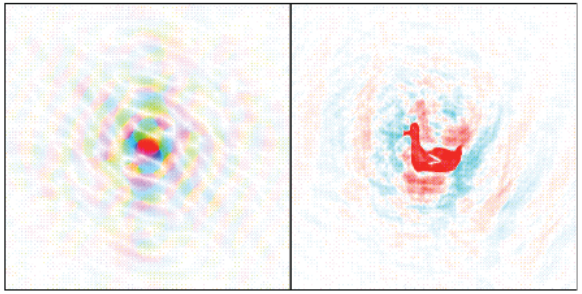


Fourier Cat and his Fourier transform
Phase is color coded

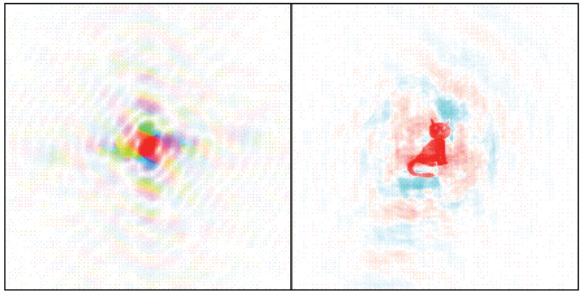


$$\rho(xyz) = \frac{1}{V}\sum_{hkl}|F(hkl)|e^{-2\pi i(hx+ky+lz)+i\alpha_{hkl}}$$

Duck phase and Cat amplitude



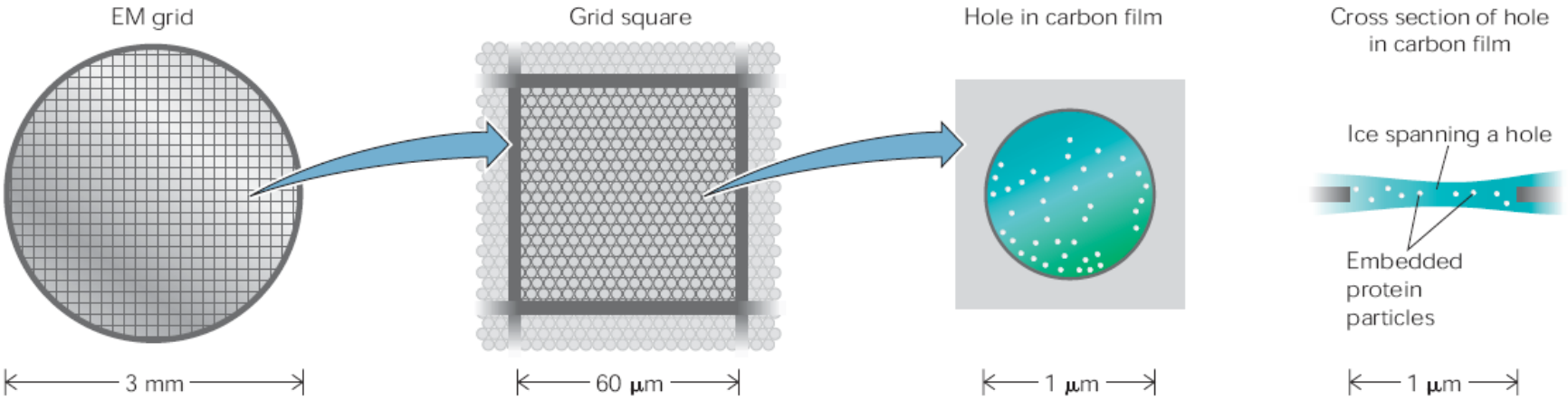
Cat phase and Duck amplitude



http://www.ruppweb.org/Xray/Phasing/Phasingt.html

Yong Xiong

In a cryo-EM specimen, the fast-frozen sample is supported by a perforated carbon film



Adhering to a standard 3-mm electron microscope grid is a carbon film of ~500 Å thickness perforated with holes 1–2 μm in diameter. The carbon film supports a 1,000-Å layer of buffer, in which the particles of interest are embedded. This layer is rapidly frozen in liquid ethane to form vitreous ice. The specimen is maintained continuously below –160°C during storage and also during imaging in the electron microscope to prevent the formation of ice crystals.

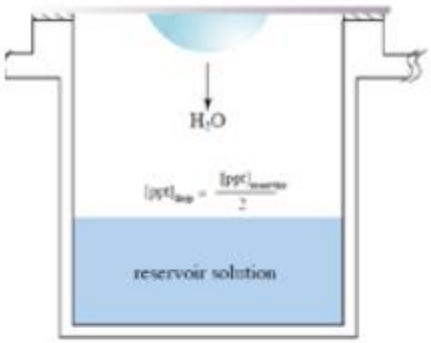
PDB explore

http://www.rcsb.org/pdb/home/home.do

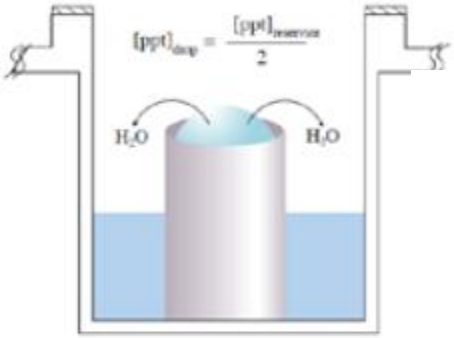
Some Crystallization Methods:

Vapor diffusion

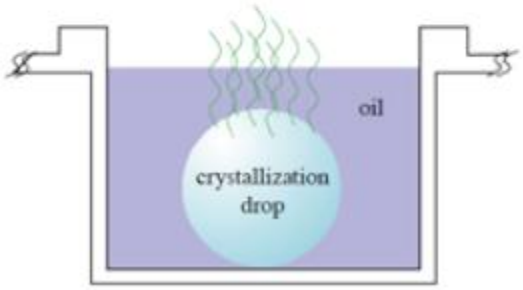
Hanging-drop



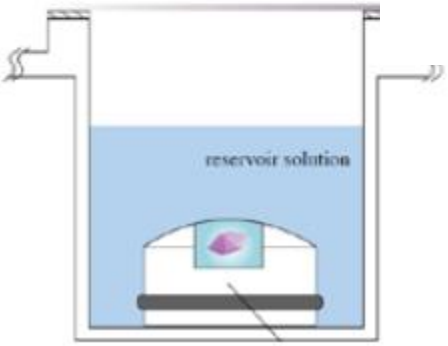
Sitting-drop

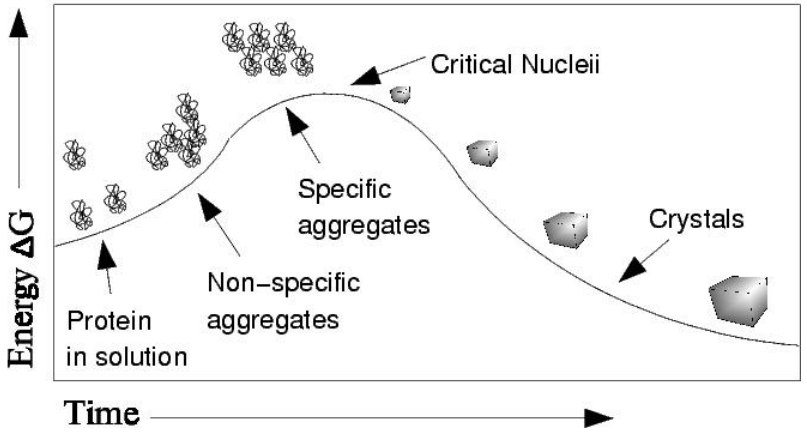


Batch:
micro batch under oil



Dialysis





Yong Xiong