

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

Jesse Rinehart, PhD

Biomedical Data Science: Mining & Modeling
CBB 752, Spring 2018



Cellular & Molecular Physiology
Yale University School of Medicine



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

Fred J. Sigworth, C.V. Sindelar

MB&B 720a, Macromolecular Structure and Biophysical Analysis

Yong Xiong , Andrew Miranker, Anna Marie Pyle

MB&B 721b, Macromolecular Interactions and Dynamic Properties

Anna Pyle, Donald Engelman, Elizabeth Rhoades, Hongwei Wang

MB&B 760b3: Principles of Macromolecular Crystallography

Thomas Steitz

MB&B 761b4: X-ray Crystallography Workshop

Yong Xiong, Yorgo Modis, and staff

Pharmacology 529b: Structural Pharmacology

Ya Ha, Titus Boggon

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

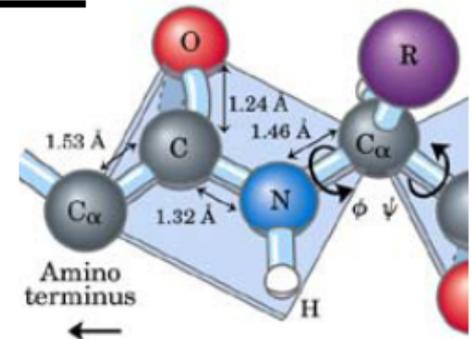
- There's a limit to how small an object can be seen under a light microscope.
- The diffraction limit: you can not image things that are much smaller than the wavelength of the light you are using.
- The wavelength for visible light is measured in hundreds of nanometers, while atoms are separated by 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 $\text{CuK}\alpha$ radiation. Wavelength = 1.54 Å.

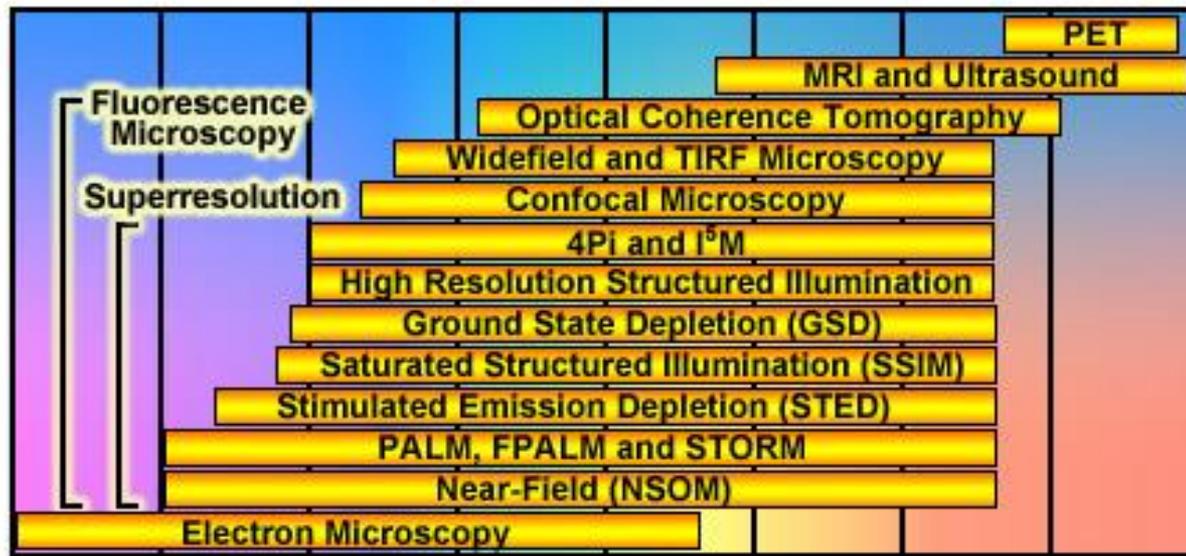
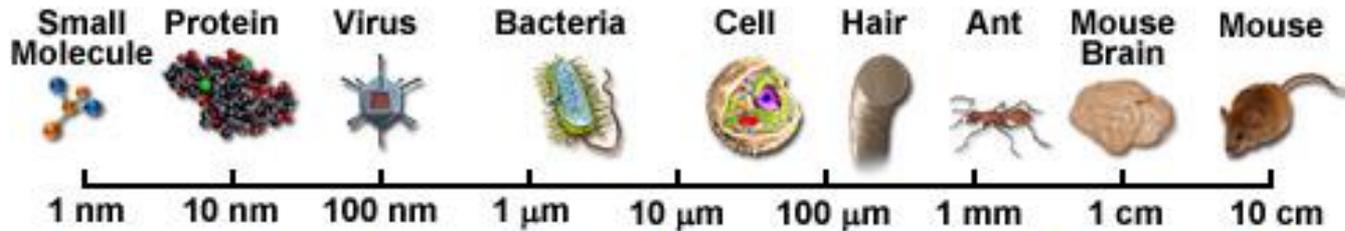
Synchrotron radiation wavelengths in the range 0.5 Å - 2.5 Å.



Yong Xiong

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

Spatial Resolution of Biological Imaging Techniques



$1 \text{ \AA} = 0.1 \text{ nm}$

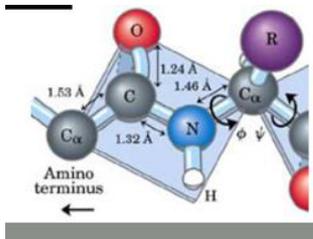
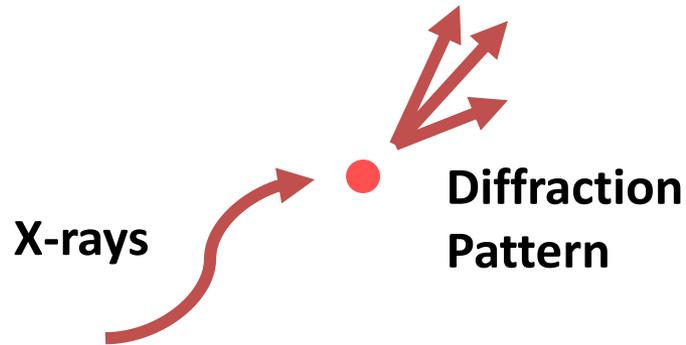


Figure 1

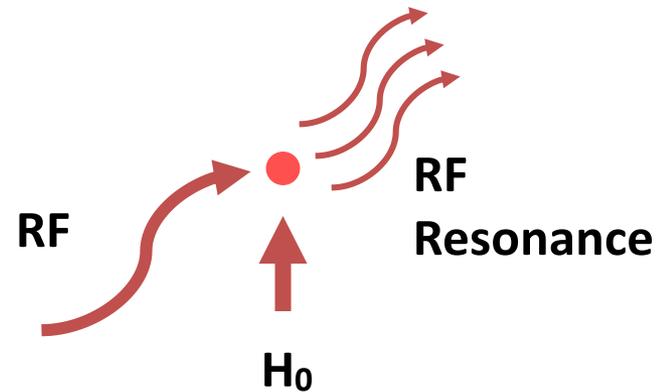
Experimental Determination of Atomic Resolution Structures

X-ray



- Direct detection of atom positions
- Crystals

NMR

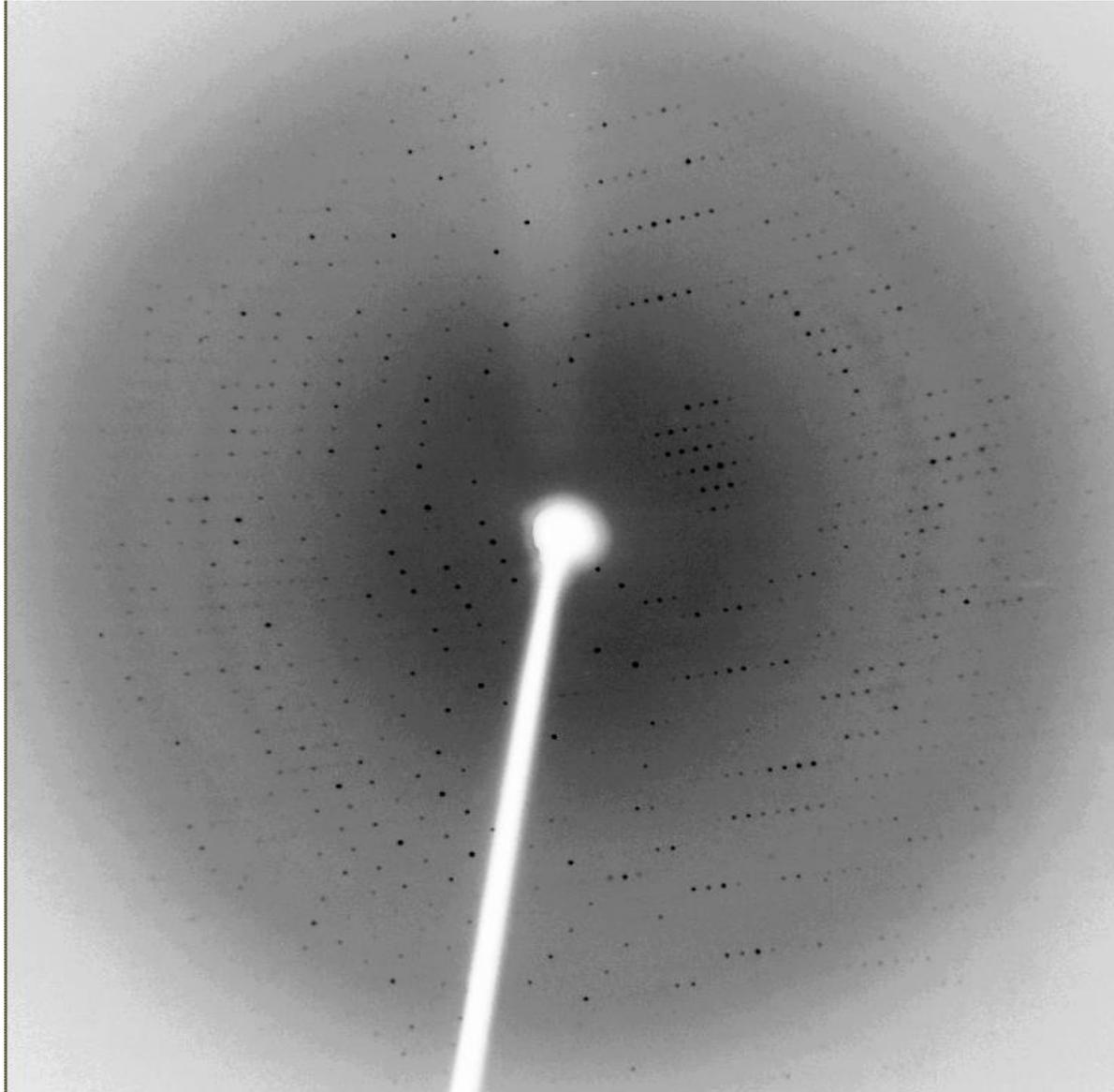


- Indirect detection of H-H distances
- In solution

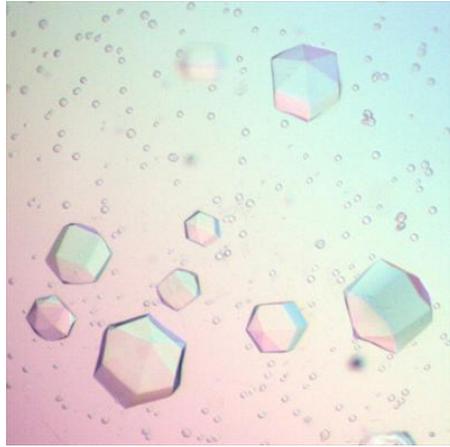
Other methods for determining protein structures:

-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 $\sim N^2$ fold).

Determination of Protein Crystal Structure

Subcloning



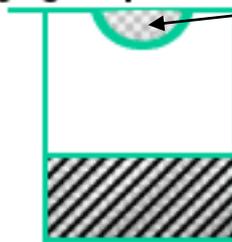
Expression



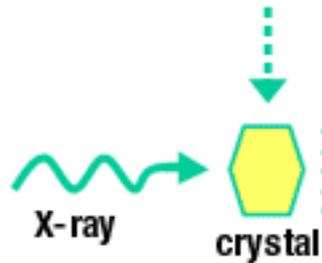
Purification



Crystallization: Hanging Drop with Protein



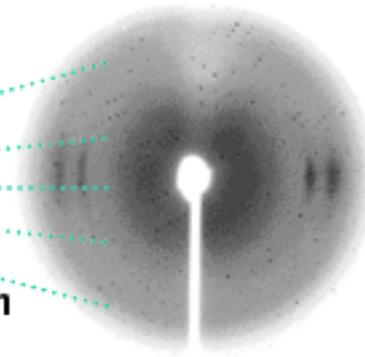
Reservoir with Precipitant



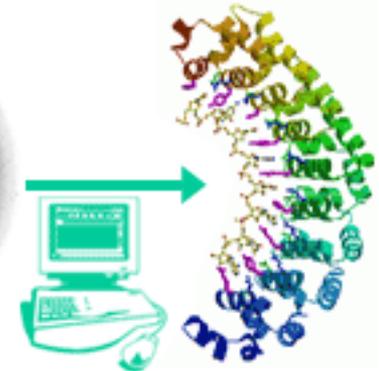
X-ray

crystal

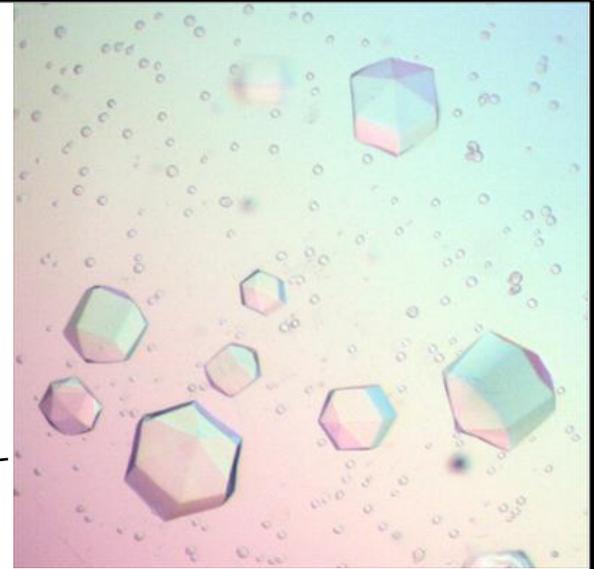
diffraction



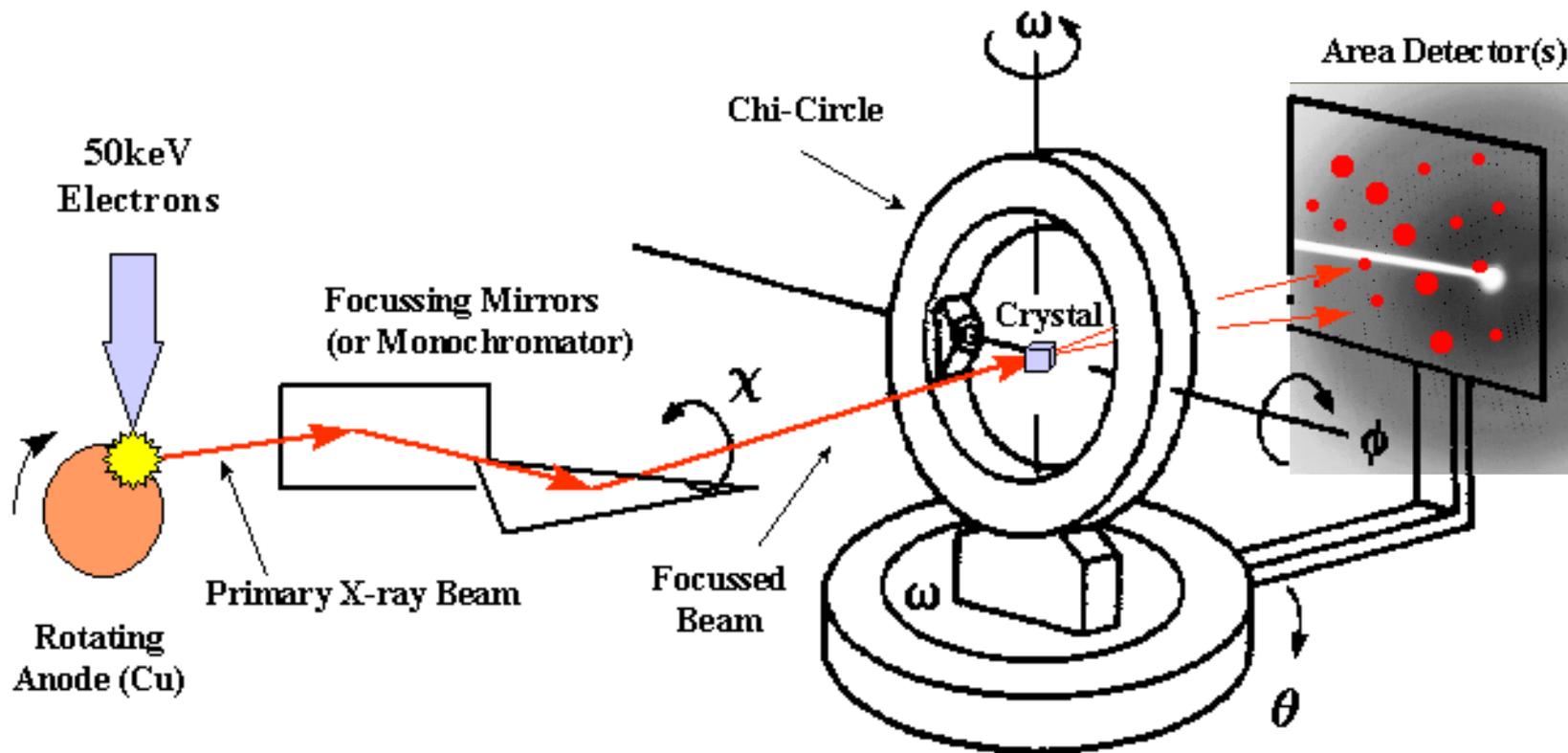
detector



refinement
modeling



Data Collection



4-Circle Goniometer (Eulerian or Kappa Geometry)

Synchrotron X-ray Sources are the method of choice

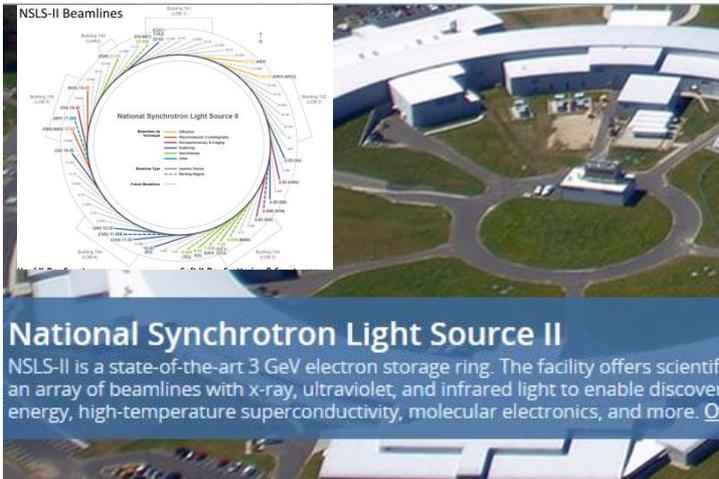
Lab x-ray sources @ 1.54 Å compared to Synchrotron X-ray @ 0.5 Å - 2.5 Å.



APS Chicago



ALS Berkeley

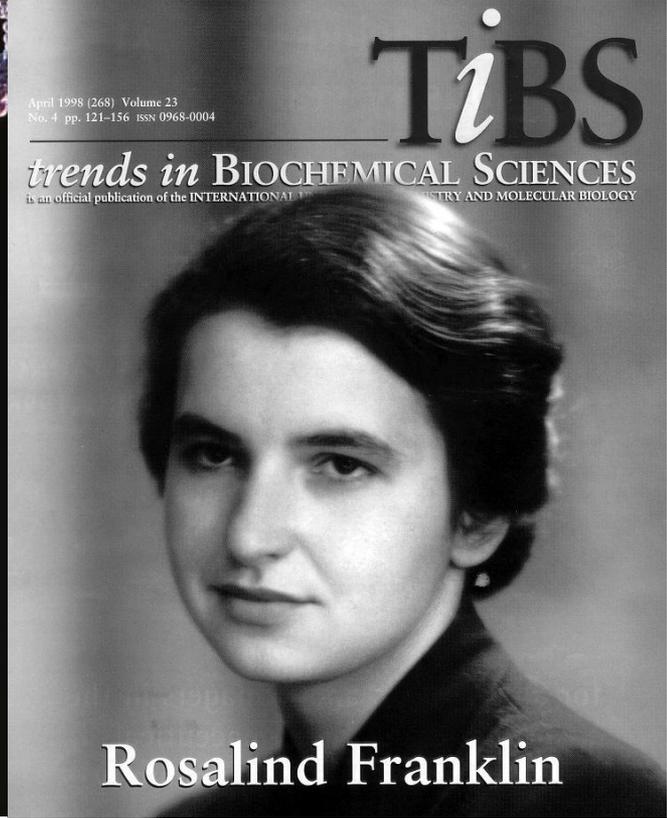
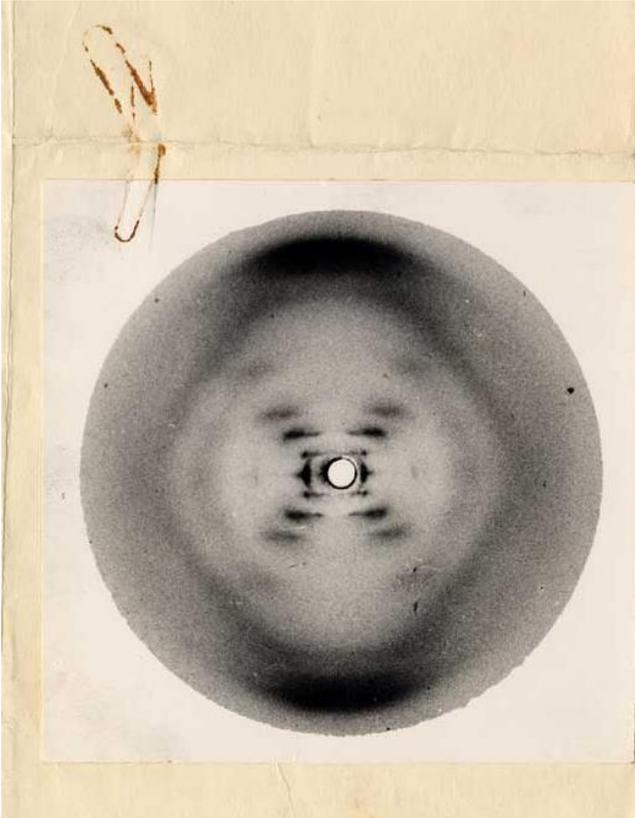


NSLS-II Brookhaven

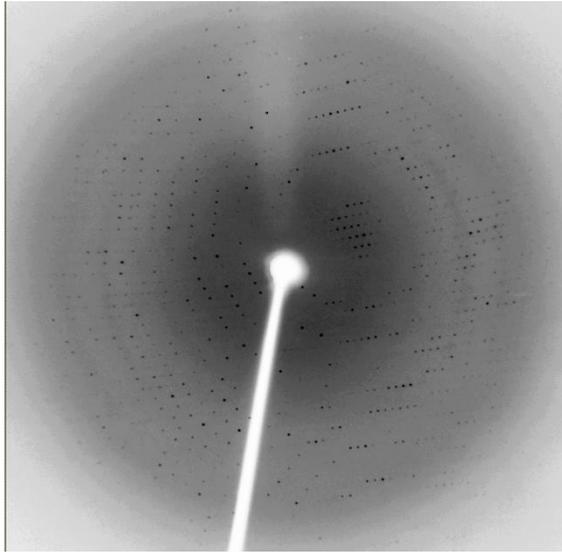


CHESS Ithaca

Most famous X-ray diffraction pattern



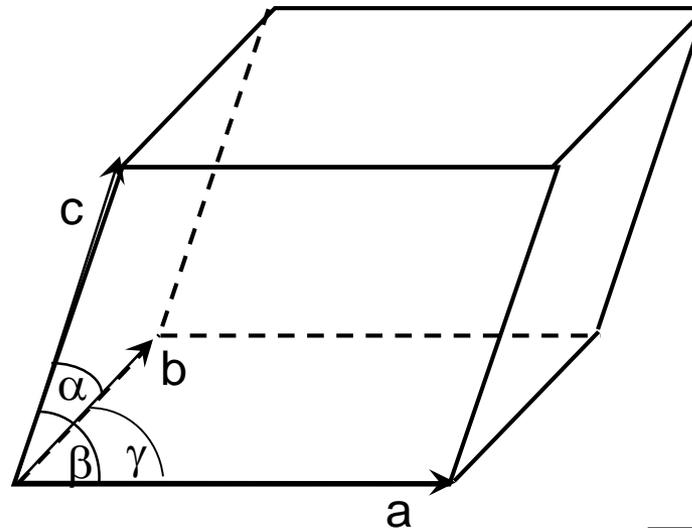
The information we get from a single diffraction experiment



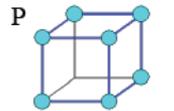
Analyze the pattern
of the reflections



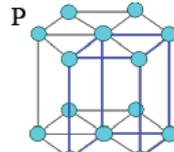
- (a) space group of the crystal
- (b) unit cell dimensions



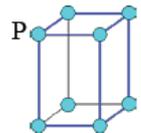
Cubic
 $a = b = c$,
 $\alpha = \beta = \gamma = 90^\circ$



Hexagonal
 $a = b \neq c$,
 $\alpha = \beta = 90^\circ, \gamma = 120^\circ$



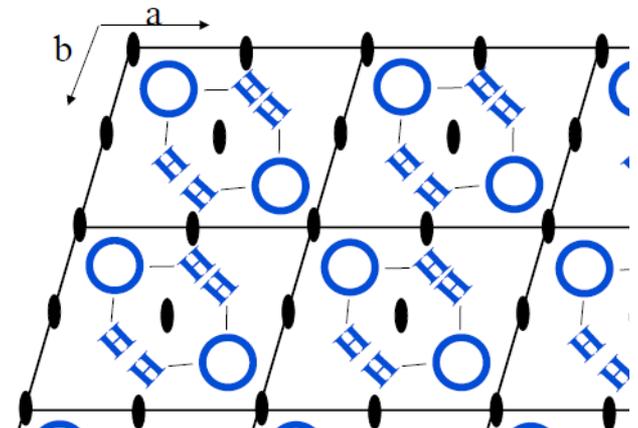
Trigonal
 $a = b \neq c$,
 $\alpha = \beta = 90^\circ, \gamma = 120^\circ$

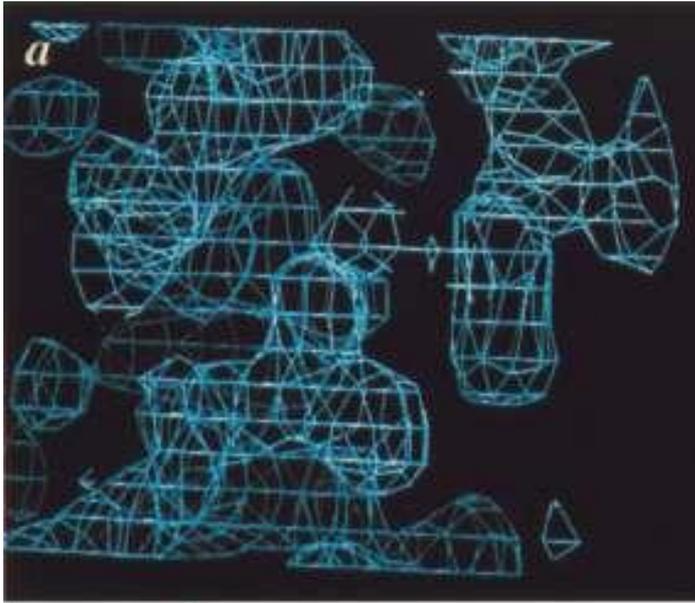


Tetragonal
 $a = b \neq c$,
 $\alpha = \beta = \gamma = 90^\circ$

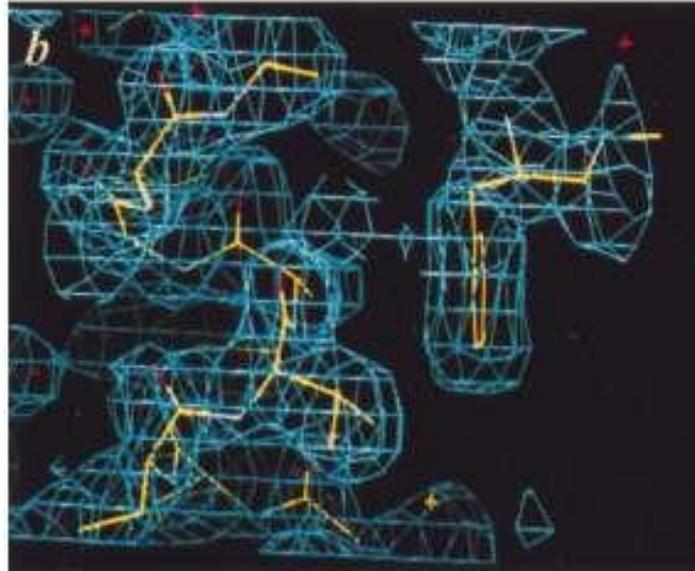
How to understand symmetry?

Crystal = lattice + unit cell content
 (asymmetric units (asu) content)





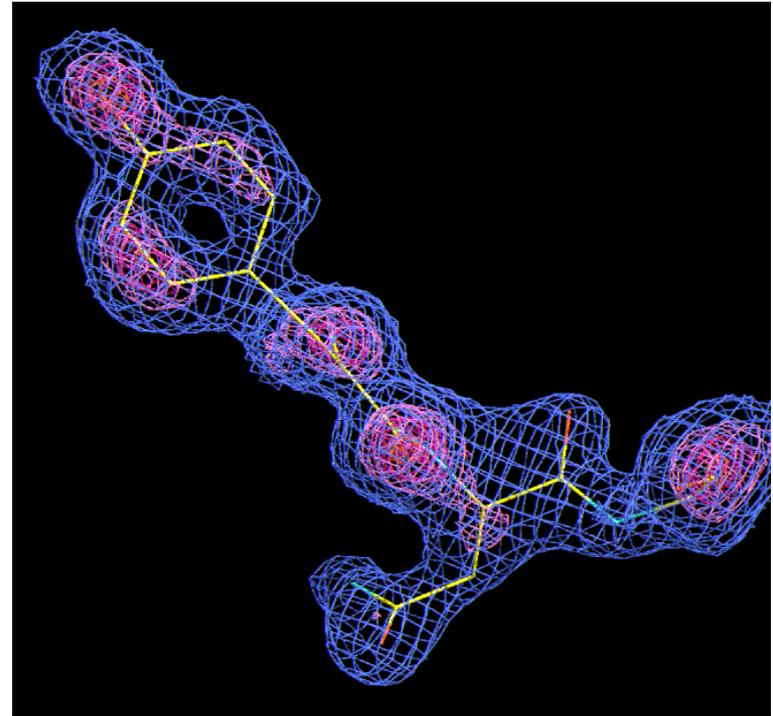
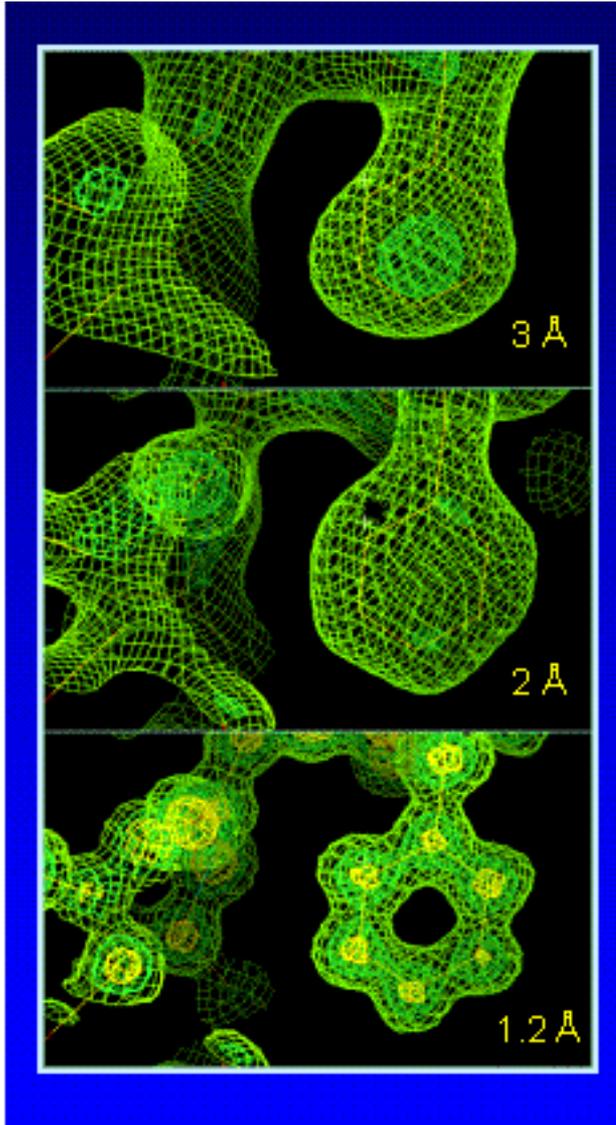
Electron density map



Building a structure model

- © 2006
- Academic Press

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.

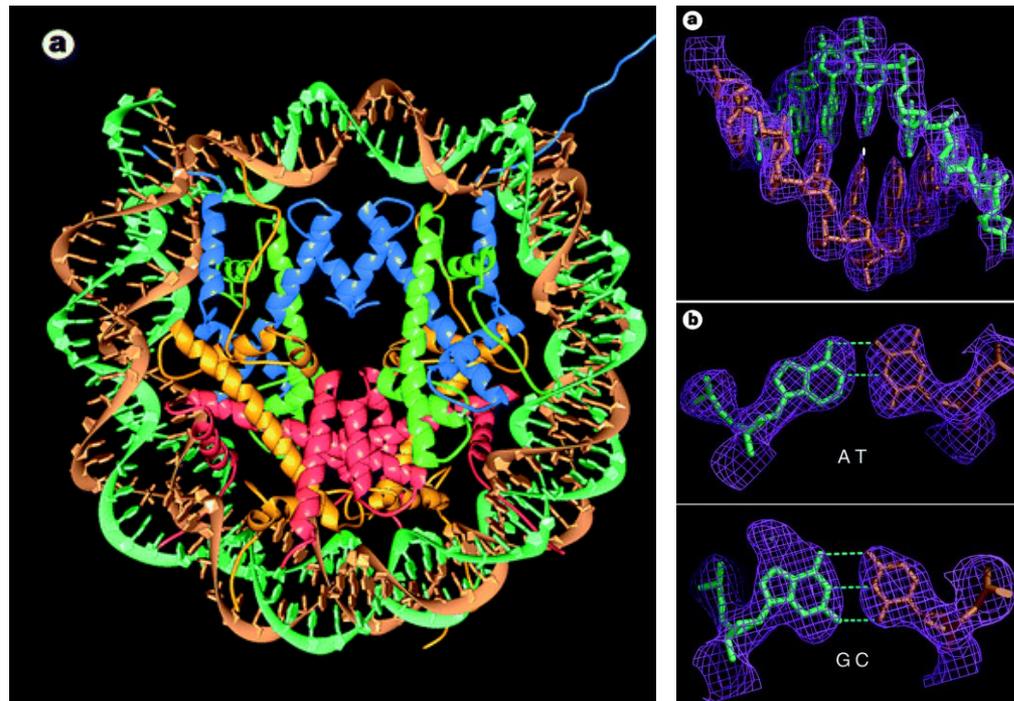
Example of high quality Experimental data where very little refinement has been applied to fit a tyrosine into the density map.

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

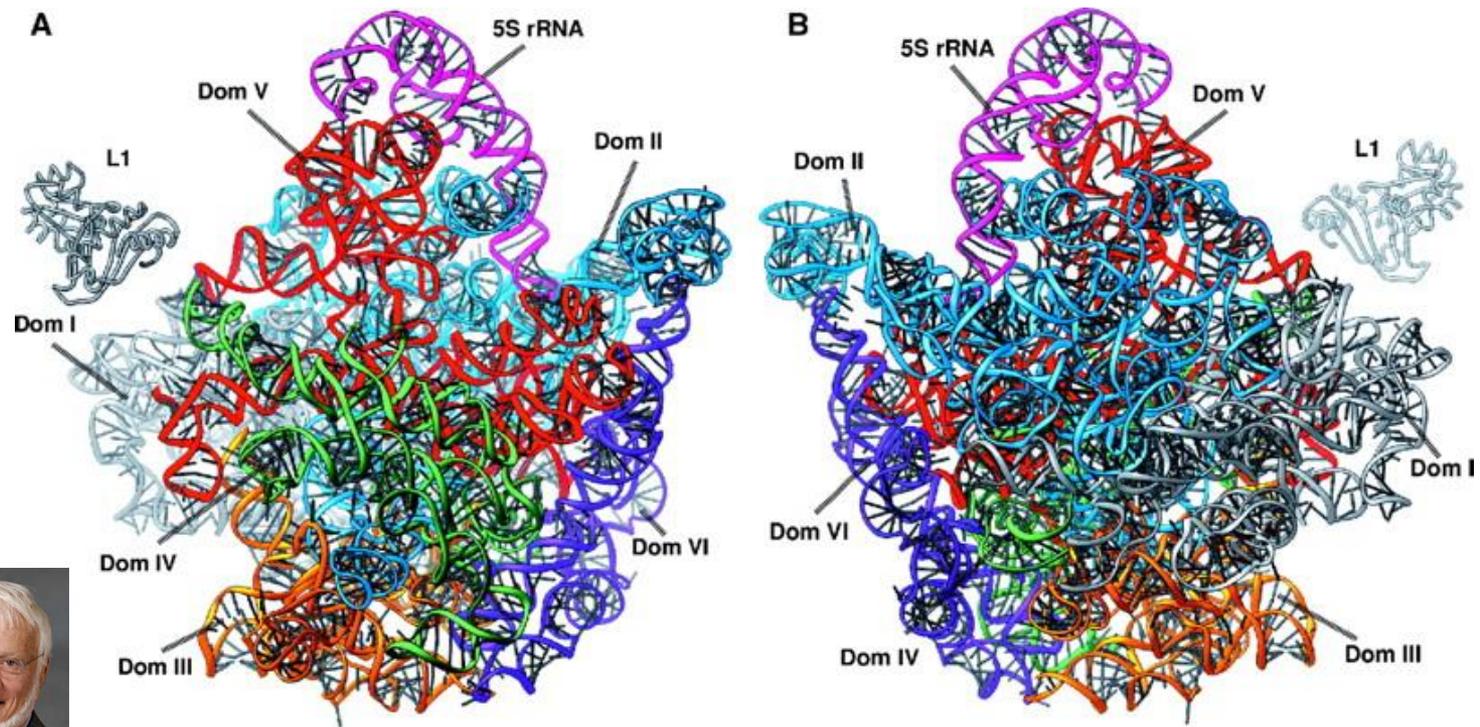
Institut für Molekularbiologie und Biophysik ETHZ, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

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.



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

Protein 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
2016:	115,559
2017:	117,184
2018:	137,178

RCSB PDB PROTEIN DATA BANK 137178 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligands Go

Advanced Search | Browse by Annotations

PDB-101 WORLDWIDE PDB PROTEIN DATA BANK EMDatabank NUCLEIC ACID DATABASE Worldwide Protein Data Bank Foundation

Welcome

- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

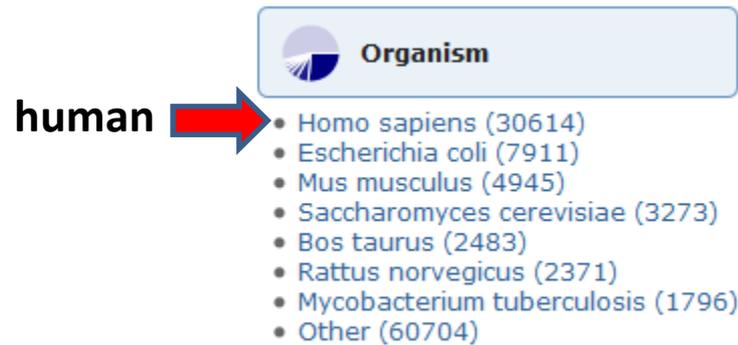
New Video: What is a Protein?

VIDEO WHAT IS A PROTEIN?

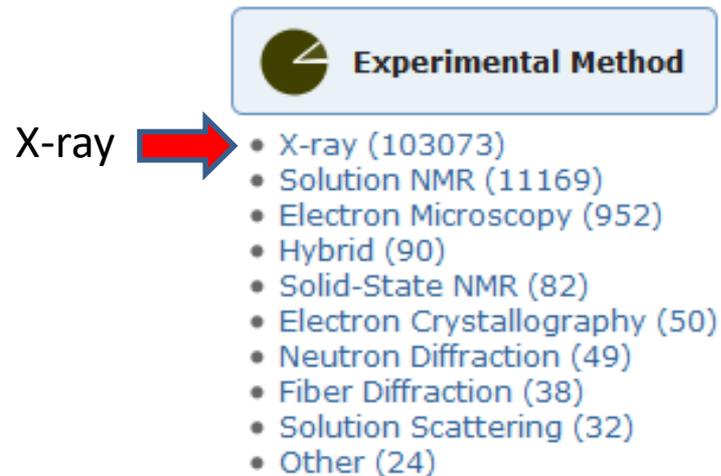
January Molecule of the Month

Opioid Receptors

PDB: What species are the structures from?



Which methods?



PDB X-ray Structures:

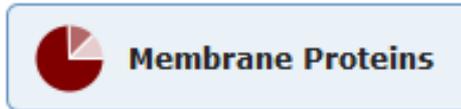
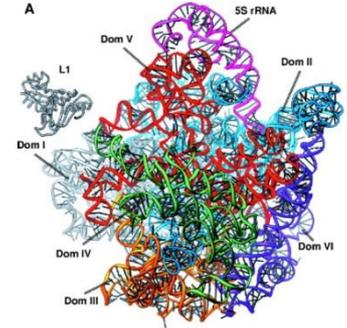
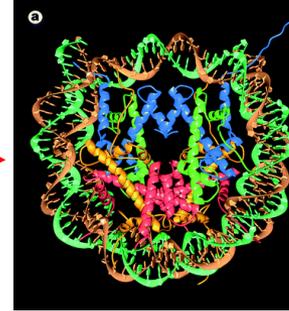
<http://www.rcsb.org/pdb/results/results.do?outformat=&qrid=1B04C26E&tabtoshow=Current>



- Homo sapiens (26612)
- Escherichia coli (7185)
- Mus musculus (4184)
- Saccharomyces cerevisiae (2858)
- Bos taurus (2275)
- Rattus norvegicus (2113)
- Mycobacterium tuberculosis (1745)
- Other (55853)



- Protein (96499)
- Mixed (4959)
- DNA (973)
- RNA (638)



- ALPHA-HELICAL (2100)
- MONOTOPIC MEMBRANE PROTEINS (358)
- BETA-BARREL (352)

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"

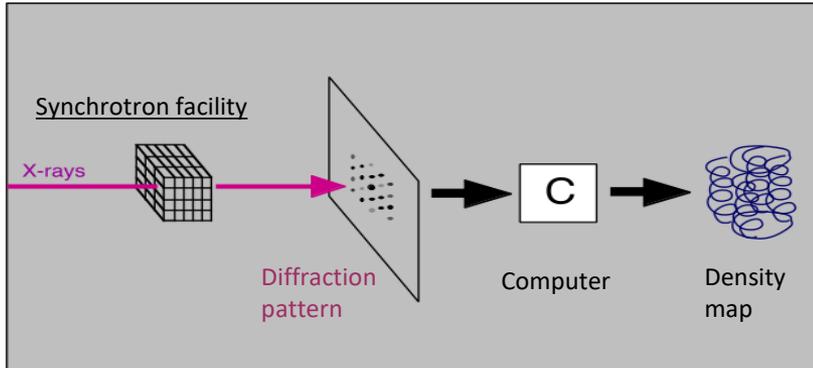


Jacques Dubochet (University of Lausanne, Switzerland)

Joachim Frank (Columbia University, New York, USA)

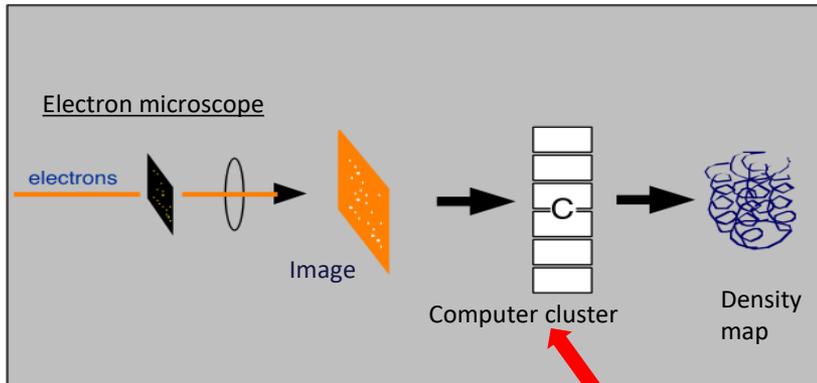
Richard Henderson (MRC Laboratory of Molecular Biology, Cambridge, UK)

Two methods for structure determination



X-ray crystallography

Well-established (since 1960s)
Requires well-ordered crystals
>10¹² copies of protein



Single-particle cryo-EM

Recent (1990s-present)
No crystals required!
~10⁵ copies of protein

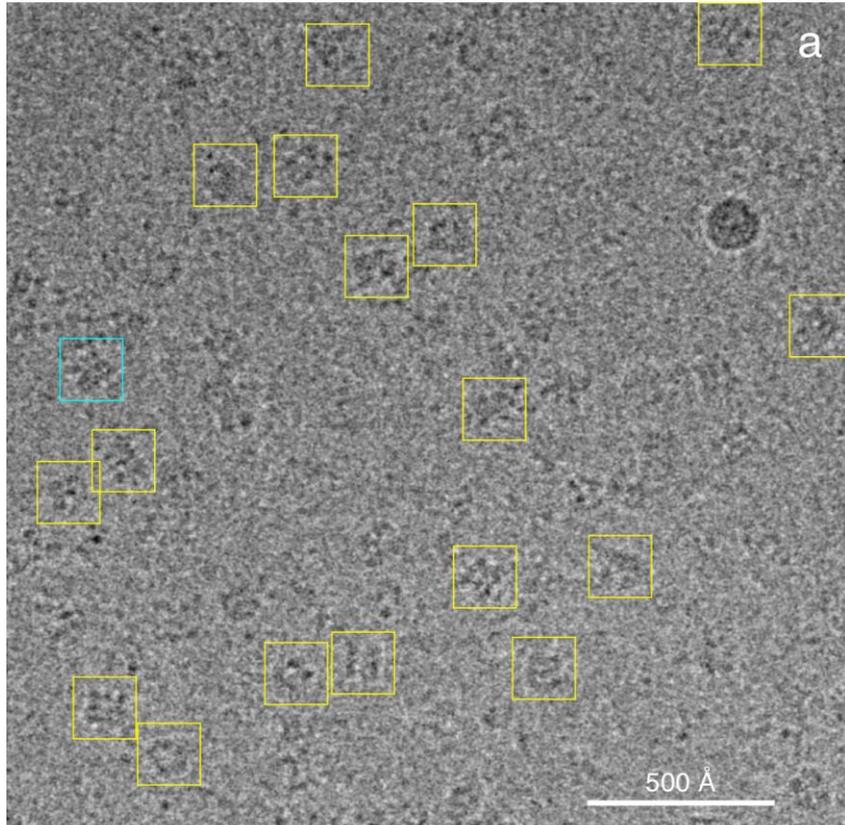
Some gaming PCs
can now replace
the cluster 😊



The Cryo-EM specimen gives only a phase contrast image

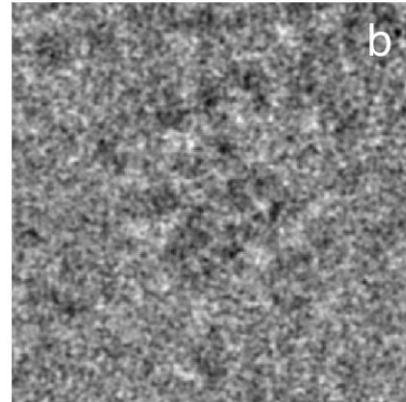
- A constellation of images and data processing are essential.

1/4 of a micrograph, showing some particles

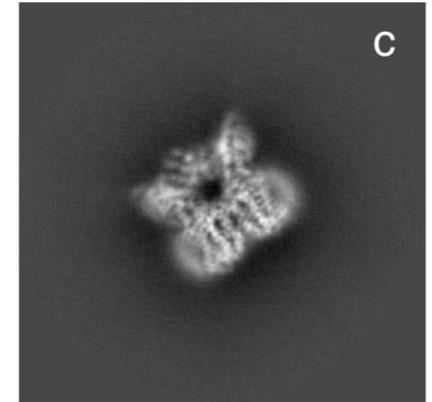


Y. Cheng and D. Julius lab. Nature 2013

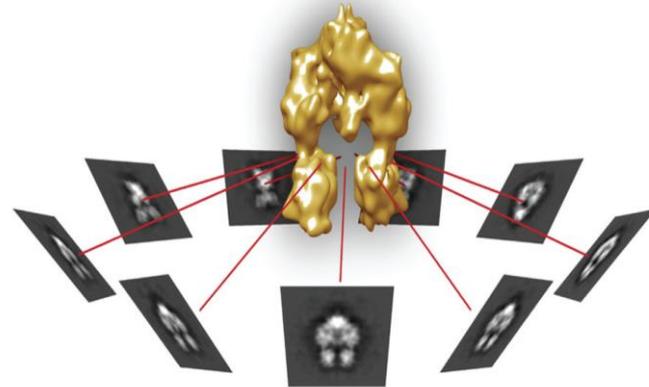
Image



Projection



- orientation assignment and averaging
- 3D reconstruction



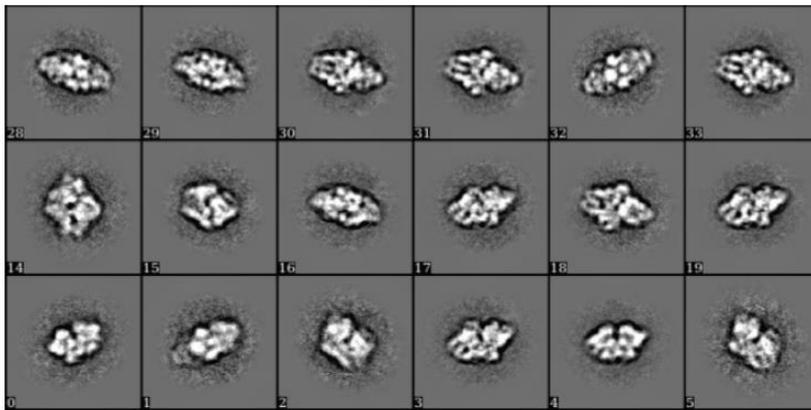
Fred Sigworth

A landmark study for high-resolution single-particle structures

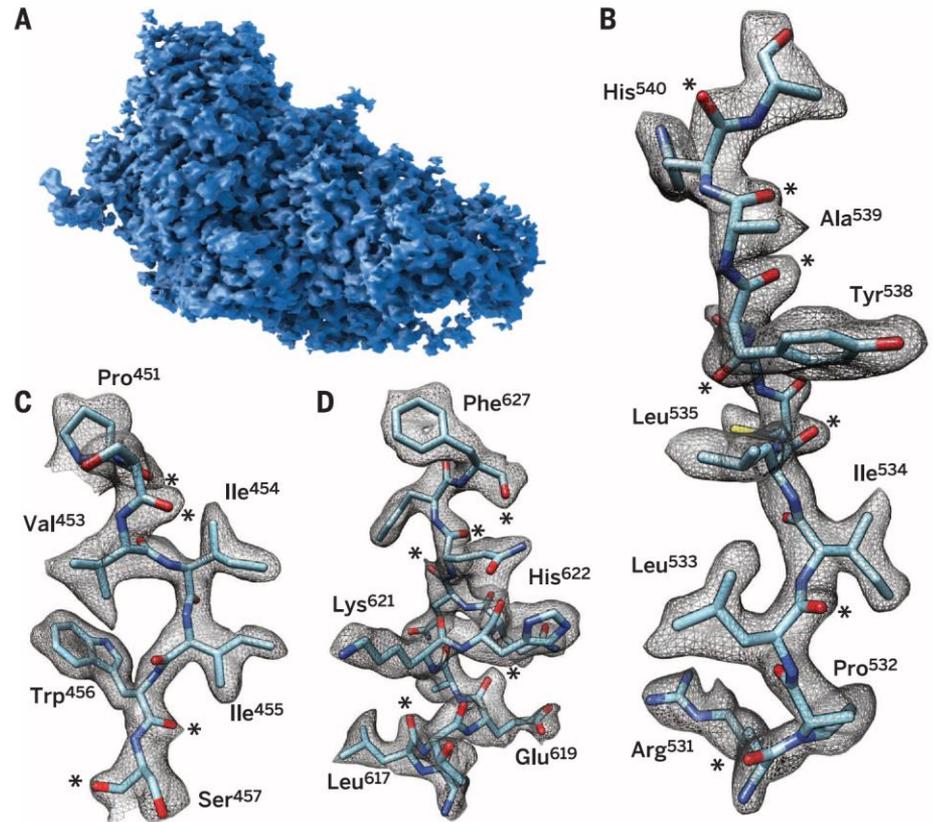
2.2 Å resolution cryo-EM structure of β -galactosidase in complex with a cell-permeant inhibitor

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



New Technologies, Automation, & Computation are accelerating the field



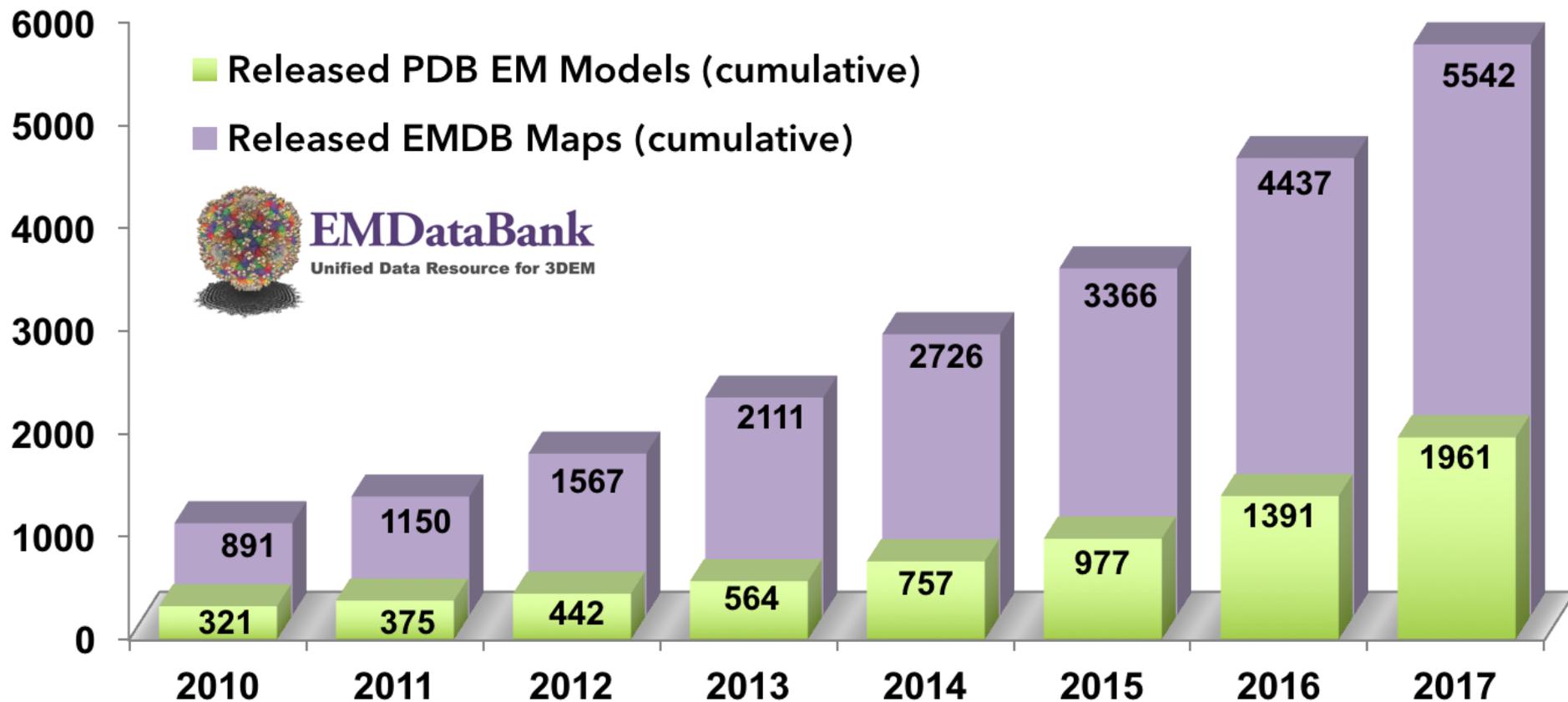
Krios at National University of Singapore



Control room at Scripps Research Institute, La Jolla



Krios TEM installation on Yale's West Campus.



Jan 2017: 4,492 total maps
3,614 (~77% single particle)

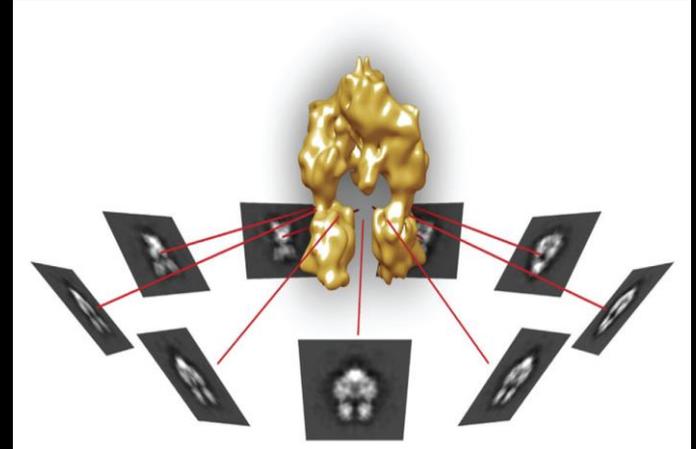
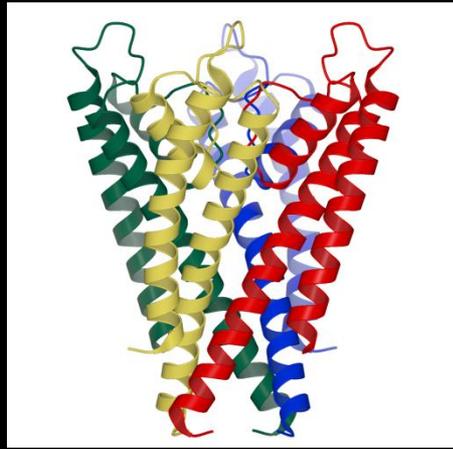
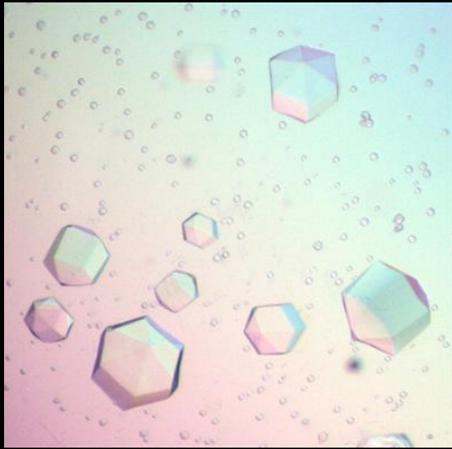
EM method	Component type
Single-particle (3614 / 3614)	Protein (2603 / 2603)
	Virus (761 / 761)
	Nucleic acid (415 / 415)
	Ligand (274 / 274)
	Prokaryotic ribosome (264 / 264)
	Eukaryotic ribosome (217 / 217)
	Cell component (113 / 113)
	EM label (2 / 2)



Less than 5Å
resolution

Filter results ?
748 entries

EM method	Component type
Single-particle (661 / 661)	Protein (601 / 601)
Helical (63 / 63)	Ligand (178 / 178)
2D crystallography (21 / 21)	Nucleic acid (140 / 140)
Subtomogram averaging (3 / 3)	Virus (134 / 134)
	Eukaryotic ribosome (69 / 69)
	Cell component (44 / 44)
	Prokaryotic ribosome (22 / 22)



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

Jesse Rinehart, PhD

Biomedical Data Science: Mining & Modeling
CBB 752, Spring 2018



Cellular & Molecular Physiology
Yale University School of Medicine

