

X-ray crystallography

CS/CME/BioE/Biophys/BMI 279

Oct. 29, 2020

Ron Dror

No class on Tuesday (election day)

Outline

- Overview of x-ray crystallography
- Crystals
- Electron density
- Diffraction patterns
- The computational problem: determining structure from the diffraction pattern

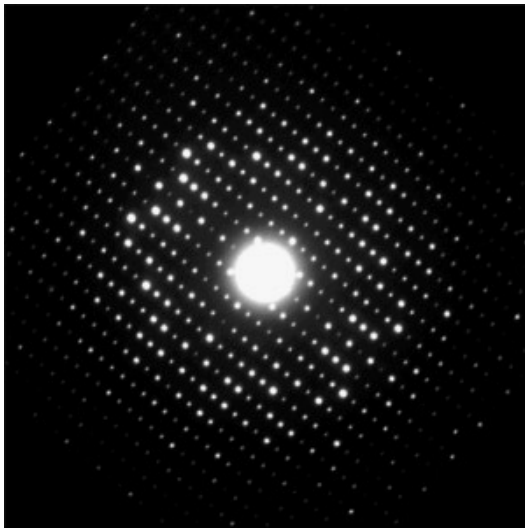
Overview of x-ray crystallography

X-ray crystallography is the most common way to determine 3D molecular structures

- Nearly 90% of the structures in the PDB were determined through x-ray crystallography
- X-ray crystallography is also frequently used to determine structures of small molecules (including drugs)
- Why are we covering it in this course?
 - So you know where biomolecular structures come from
 - Because determining a structure this way involves solving a challenging computational problem

The basic idea

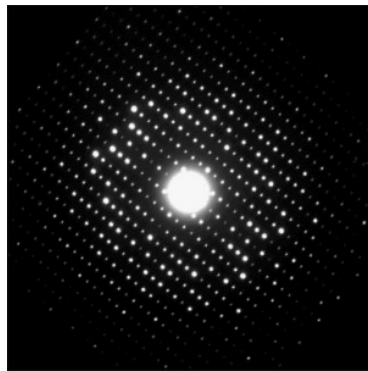
- Get the molecule whose structure you want to determine to form a crystal
- Shine an intense beam of x-rays through the crystal, giving rise to a “diffraction pattern” (a pattern of spots of varying brightnesses)



<http://lacasadeloscristales.trianatech.com/wp-content/uploads/2014/09/image005-300x300.jpg>

The basic idea

- From that pattern, infer the 3D structure of the molecule
 - In fact, one uses multiple images, with the x-rays shining through the crystal at different angles
- This is a challenging computational problem!
- It turns out the diffraction pattern is closely related to the *Fourier transform* of the electron density of the molecule that was crystallized
 - Before we even worry about what that means, let's go back and discuss what a crystal is and what electron density is



<http://lacasadeloscristales.trianatech.com/wp-content/uploads/2014/09/image005-300x300.jpg>

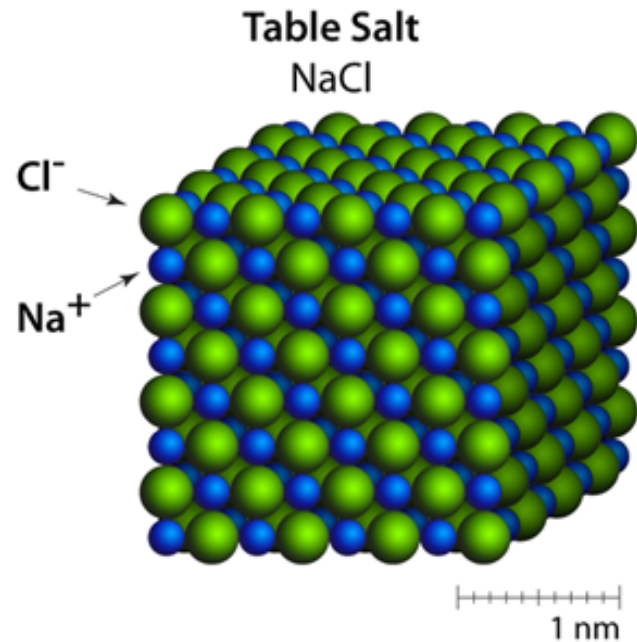
Crystals

What's a crystal?

- Under certain conditions, molecules line up into a regular grid (a “lattice”).
 - Example: table salt



<http://www.bigfoto.com/miscellaneous/photos-16/salt-crystals-94jf.jpg>



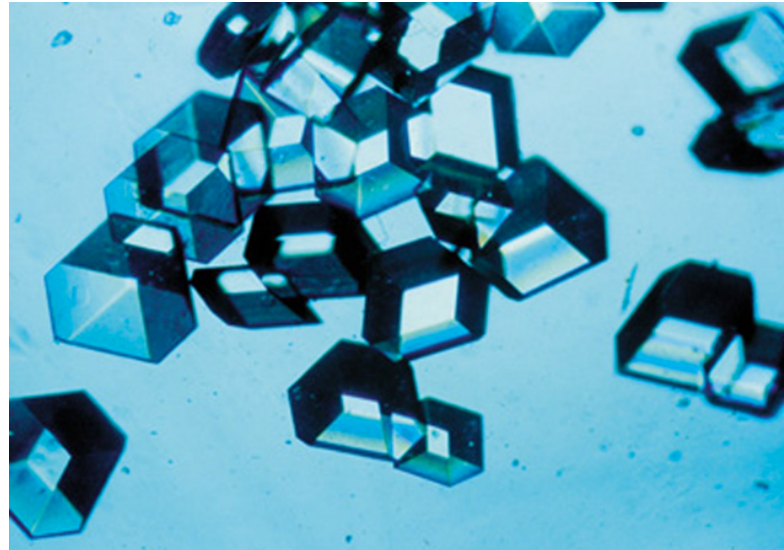
<http://www.atomsinmotion.com/book/chapter4/rockSalt.png> 9

Macromolecules can also form crystals

Macromolecules including protein, DNA, RNA, etc.

- Under certain conditions, proteins and other macromolecules will pack into a regular grid (a lattice)

Insulin crystals

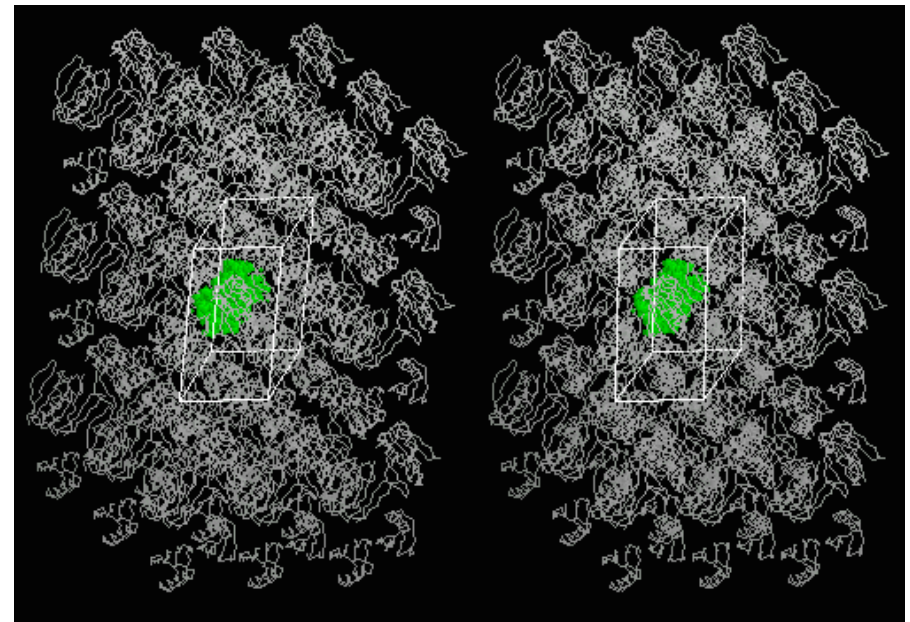
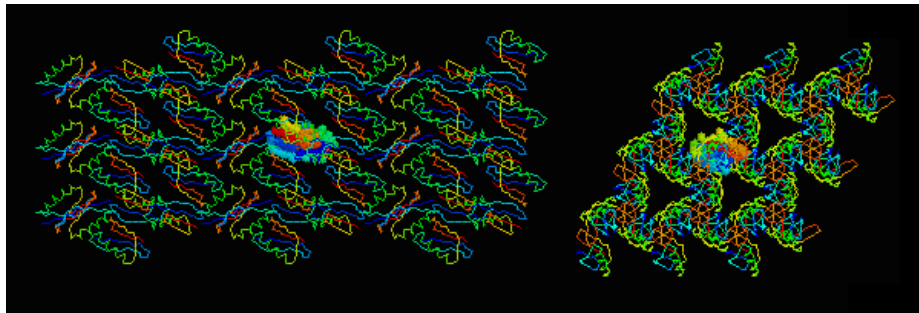


http://science.nasa.gov/media/medialibrary/1999/09/10/msad20sep99_1_resources/9901879.jpg

Macromolecules can also form crystals

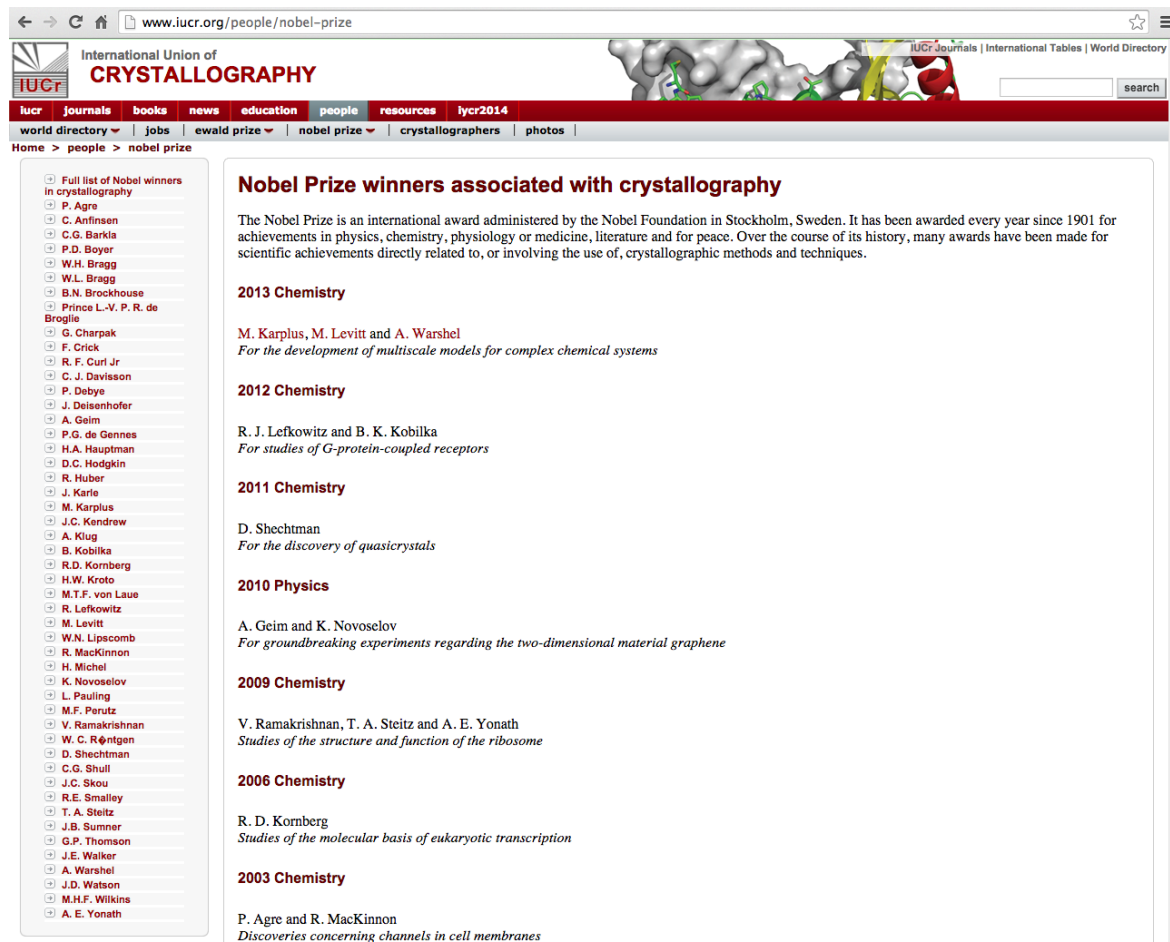
Under certain conditions, entire proteins and even protein complexes will pack into a regular grid (a lattice)

Multiple views of the crystal formed by an immunoglobulin-binding domain
(PDB entry 1PGB)



Caveats

- Getting macromolecules to form crystals can be hard
 - Crystallographers have sometimes worked for decades to get good crystals of a particular protein



The screenshot shows the IUCr website page titled "Nobel Prize winners associated with crystallography". The page features a navigation menu at the top with categories like "iucr", "journals", "books", "news", "education", "people", "resources", and "iucr2014". Below the navigation, there is a search bar and a list of links including "world directory", "jobs", "ewald prize", "nobel prize", "crystallographers", and "photos". The main content area is titled "Nobel Prize winners associated with crystallography" and includes a brief introduction to the Nobel Prize. The page lists winners from 2003 to 2013, with their names and the specific achievements for which they were awarded. A sidebar on the left provides a full list of Nobel winners in crystallography, including names like P. Agre, C. Anfinsen, C.G. Barkla, P.D. Boyer, W.H. Bragg, W.L. Bragg, B.N. Brockhouse, Prince L.-V. P. R. de Broglie, G. Charpak, F. Crick, R. F. Curl Jr, C. J. Davison, P. Debye, J. Deisenhofer, A. Geim, P.G. de Gennes, H.A. Hauptman, D.C. Hodgkin, R. Huber, J. Karle, M. Karplus, J.C. Kendrew, A. Klug, B. Kobilka, R.D. Kornberg, H.W. Kroto, M.T.F. von Laue, R. Lefkowitz, M. Levitt, W.N. Lipscomb, R. MacKinnon, H. Michel, K. Novoselov, L. Pauling, M.F. Perutz, V. Ramakrishnan, W. G. Röntgen, D. Shechtman, C.G. Shull, J.C. Skou, R.E. Smalley, T. A. Steltz, J.B. Sumner, G.P. Thomson, J.E. Walker, A. Warshel, J.D. Watson, M.H.F. Wilkins, and A. E. Yonath.

International Union of
CRYSTALLOGRAPHY

Home > people > nobel prize

Nobel Prize winners associated with crystallography

The Nobel Prize is an international award administered by the Nobel Foundation in Stockholm, Sweden. It has been awarded every year since 1901 for achievements in physics, chemistry, physiology or medicine, literature and for peace. Over the course of its history, many awards have been made for scientific achievements directly related to, or involving the use of, crystallographic methods and techniques.

2013 Chemistry

M. Karplus, M. Levitt and A. Warshel
For the development of multiscale models for complex chemical systems

2012 Chemistry

R. J. Lefkowitz and B. K. Kobilka
For studies of G-protein-coupled receptors

2011 Chemistry

D. Shechtman
For the discovery of quasicrystals

2010 Physics

A. Geim and K. Novoselov
For groundbreaking experiments regarding the two-dimensional material graphene

2009 Chemistry

V. Ramakrishnan, T. A. Steitz and A. E. Yonath
Studies of the structure and function of the ribosome

2006 Chemistry

R. D. Kornberg
Studies of the molecular basis of eukaryotic transcription

2003 Chemistry

P. Agre and R. MacKinnon
Discoveries concerning channels in cell membranes

- Full list of Nobel winners in crystallography
- P. Agre
- C. Anfinsen
- C.G. Barkla
- P.D. Boyer
- W.H. Bragg
- W.L. Bragg
- B.N. Brockhouse
- Prince L.-V. P. R. de Broglie
- G. Charpak
- F. Crick
- R. F. Curl Jr
- C. J. Davison
- P. Debye
- J. Deisenhofer
- A. Geim
- P.G. de Gennes
- H.A. Hauptman
- D.C. Hodgkin
- R. Huber
- J. Karle
- M. Karplus
- J.C. Kendrew
- A. Klug
- B. Kobilka
- R.D. Kornberg
- H.W. Kroto
- M.T.F. von Laue
- R. Lefkowitz
- M. Levitt
- W.N. Lipscomb
- R. MacKinnon
- H. Michel
- K. Novoselov
- L. Pauling
- M.F. Perutz
- V. Ramakrishnan
- W. G. Röntgen
- D. Shechtman
- C.G. Shull
- J.C. Skou
- R.E. Smalley
- T. A. Steltz
- J.B. Sumner
- G.P. Thomson
- J.E. Walker
- A. Warshel
- J.D. Watson
- M.H.F. Wilkins
- A. E. Yonath

Caveats

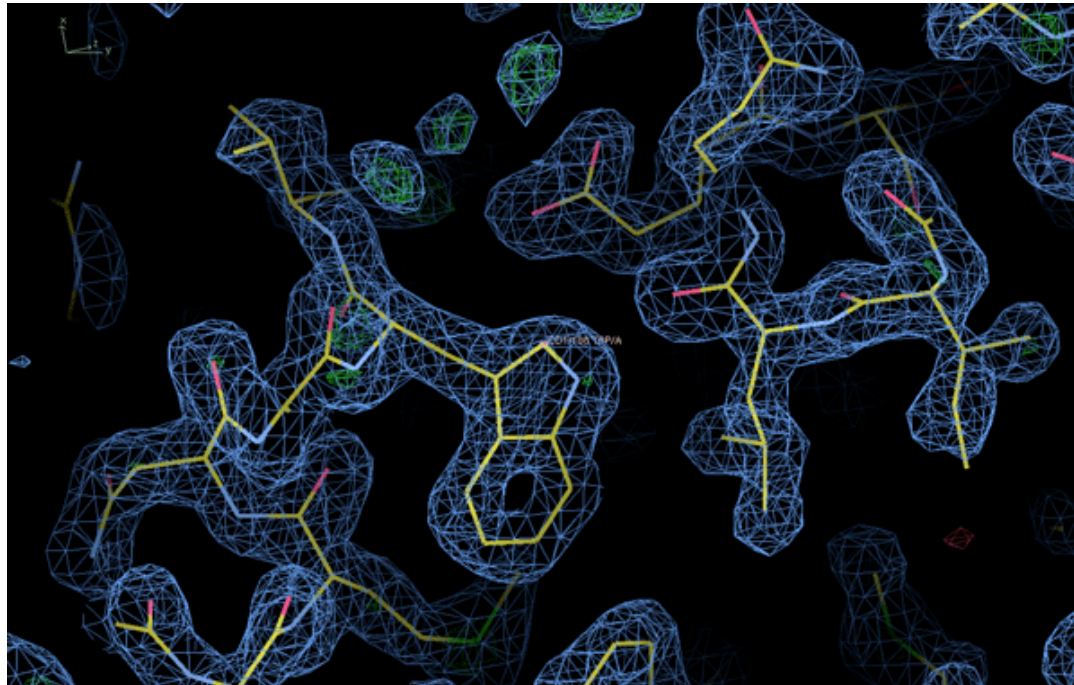
- Sometimes a molecule will adopt a different structure in a crystal than it does in its natural environment
- Crystallography gives you a static snapshot of a molecule's structure Remember molecule is not static!
 - Usually (but not always) this snapshot corresponds to the molecule's “average” structure

Electron density

Electron density of a molecule

- The *electron density* corresponding to the 3D structure of a molecule gives the probability of finding an electron at each point in space
- X-rays bounce off electrons they hit

The blue mesh represents the electron density (which corresponds to the probability of finding electrons in that area)



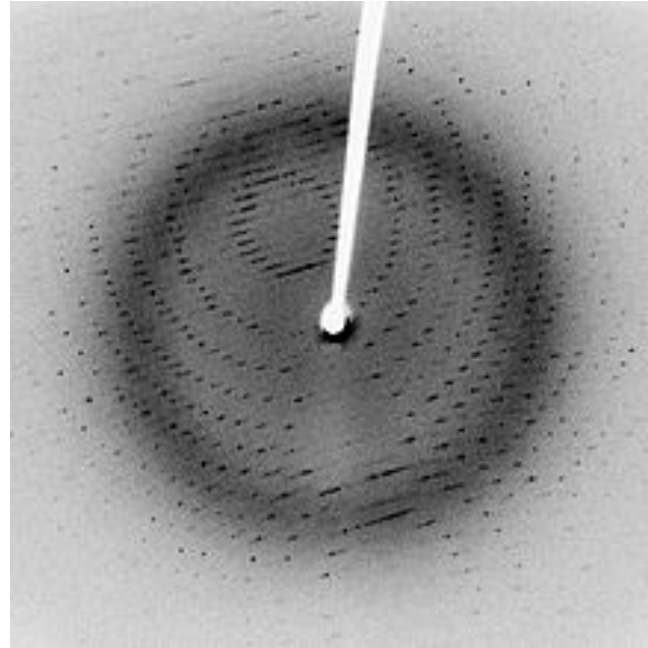
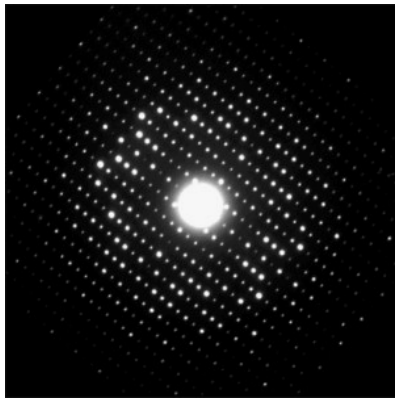
http://www.lynceantech.com/images/electron_density_map.png

Diffraction patterns

Diffraction patterns

- When you shine a light beam through a crystal, you get a distinctive pattern of bright spots called a diffraction pattern

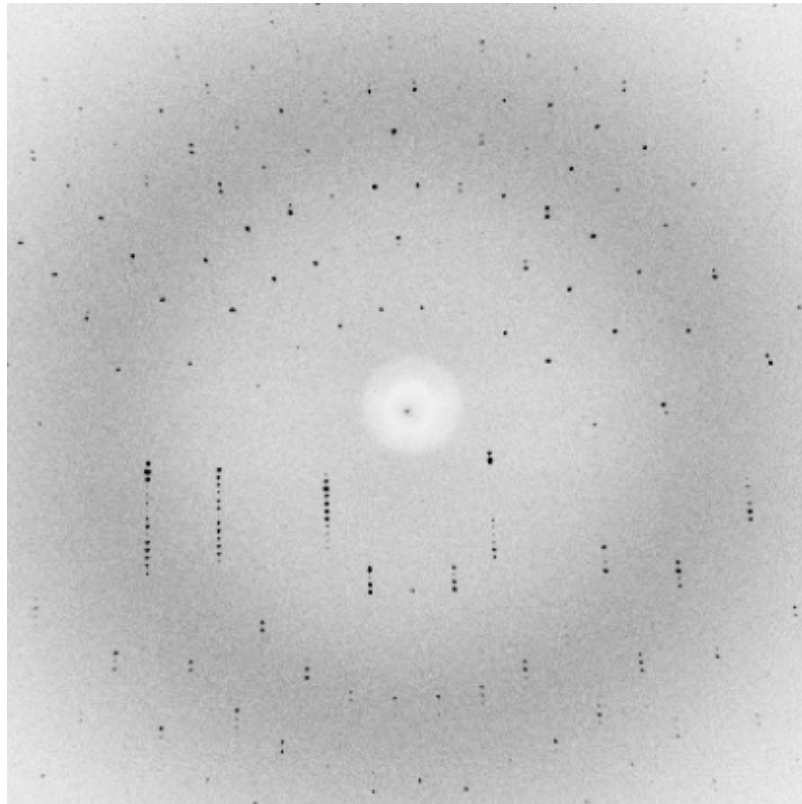
X-ray is used in crystallography because the wavelength of X-ray is roughly the same as the radius of an atom (on the order of $1\text{\AA}/10^{-10}\text{m}$)



Note that the bright spots are sometimes pictured in light/white shades (left) and sometimes in dark/black shades (right)

Diffraction patterns

- This pattern is actually three dimensional.
 - If you move the imaging plane (or rotate the crystal), you see different parts of it

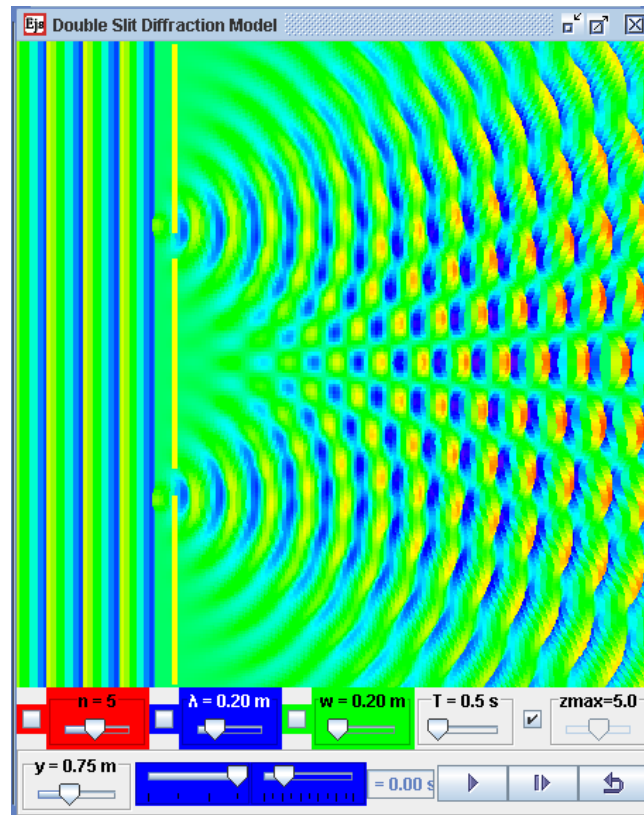


What causes diffraction patterns?

- Short answer: interference of light
 - The bright spots are places where light interferes constructively. Elsewhere it tends to interfere destructively (cancel out).

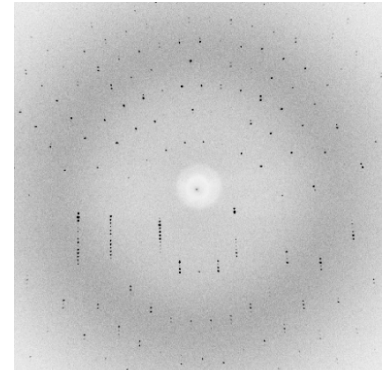
You're not responsible for this

<http://weelookang.blogspot.com/2011/10/ejs-open-source-double-slit-diffraction.html>



Relationship between diffraction pattern and electron density

- It turns out that the diffraction pattern is the *Fourier transform* of the electron density
 - Both the electron density and the diffraction pattern are functions of three dimensions (i.e., defined at every point in a 3D volume)
 - Each bright spot in the diffraction pattern corresponds to one sinusoidal component of the electron density The brightness of the spot gives the magnitude of the sinusoid
 - The Fourier transform gives a magnitude and a phase for each sinusoid, but it's only practical to measure the magnitude, not the phase
 - Brightness of the spot gives the magnitude



You need to understand this relationship, but not exactly *why* it holds

The computational problem: determining structure from the diffraction pattern

The challenge

- Given a diffraction pattern, determine the electron density and/or the position of each atom
- If we had a magnitude and a phase associated with each spot in the diffraction pattern—and thus with each 3D sinusoid—then we could just sum up appropriately scaled and shifted 3D sinusoids to recover the electron density
- But we don't have the phases
 - This makes the problem “underdetermined”—in principle, multiple electron densities could give rise to the same set of diffraction pattern magnitudes
 - **But the vast majority of those won't correspond to reasonable 3D structures of the protein**

General approach to solution

- **Step 1: *Initial phasing*** Solving for the shift of each of the sinusoids
 - Come up with an approximate solution for the structure (and thus an approximate set of phases)
- **Step 2: *Phase refinement***
 - Then consider perturbations to the structure
 - Search for perturbations that improve the fit to the experimental data (the diffraction pattern)

Initial phasing

- The most common method for initial phasing is *molecular replacement*
 - Start with a computational model of the protein structure (often the structure of a homologous protein)
 - Search over the possible ways that a protein with this structure could be packed into a crystal, and find the one that gives the best fit to the data
 - If one can't build a good computational model of the protein, then one can try various experimental methods to help determine phases
 - Example: *isomorphous replacement*, where one replaces several atoms of the protein with heavier atoms (usually metals), and then uses the *change* in the diffraction pattern to solve for the phases
- Mercury, lead, uranium are often used for isomorphous replacement

Phase refinement

- Once we have an initial model, we can search for perturbations to that model that improve the fit to the experimental data
 - This is usually done through a Monte Carlo search (via simulated annealing)
 - One usually restrains the search to “realistic” molecular structures using a molecular mechanics force field Including a term that corresponds to agreement with the diffraction pattern
 - This dramatically improves the accuracy of the results
 - The idea was introduced by Axel Brunger, now on the Stanford faculty

Phase refinement

- A major challenge in the phase refinement process is to avoid overfitting—i.e., fitting to the noise in the experimental measurements
- To avoid this, one generally ignores a small subset of the experimental data during the refinement process, then sees how well one can predict it at the end
 - Just like cross-validation in machine learning
 - This idea also came from Brunger (who introduced the term R_{free} to quantify the error in the prediction)

Lower R_{free} is better (i.e. your prediction is more accurate)

Computational methods continue to improve

- Although the phasing problem is decades old, researchers are still inventing better solutions

nature

Vol 464 | 22 April 2010 | doi:10.1038/nature08892

Super-resolution biomolecular crystallography with low-resolution data

Gunnar F. Schröder^{1,2}, Michael Levitt² & Axel T. Brunger^{2,3,4,5,6}

NOVEMBER 2013 | **NATURE METHODS**

Improved low-resolution crystallographic refinement with Phenix and Rosetta

Frank DiMaio^{1,6}, Nathaniel Echols^{2,6}, Jeffrey J Headd², Thomas C Terwilliger³, Paul D Adams^{2,4} & David Baker^{1,5}

A few additional notes

- Protein crystals contain water
 - Often half the crystal is water (filling all the empty spaces between copies of the protein)
 - Usually only a few water molecules are visible in the structure, because the rest are too mobile
- One usually can't determine hydrogen positions by x-ray crystallography
 - But one can model them in computationally
- Some high-profile, published crystal structures have turned out to be completely incorrect, due to computational problems/errors