Introduction:
New methods for solving tough crystal structures

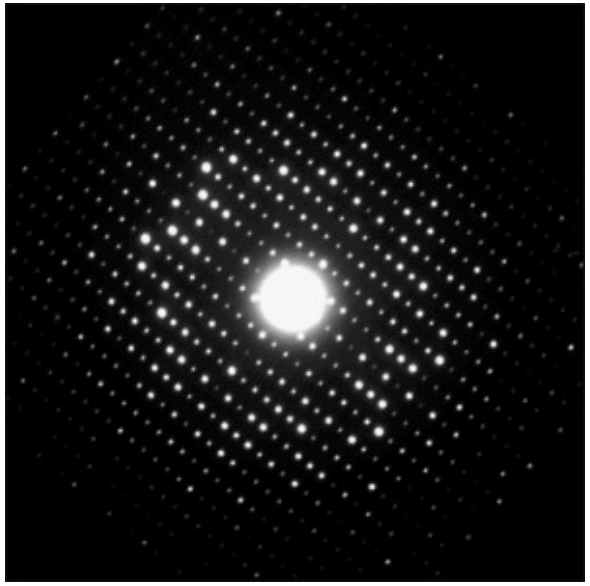
CS/CME/Biophys/BMI 371

Feb. 1, 2017

Ron Dror

X-ray crystallography

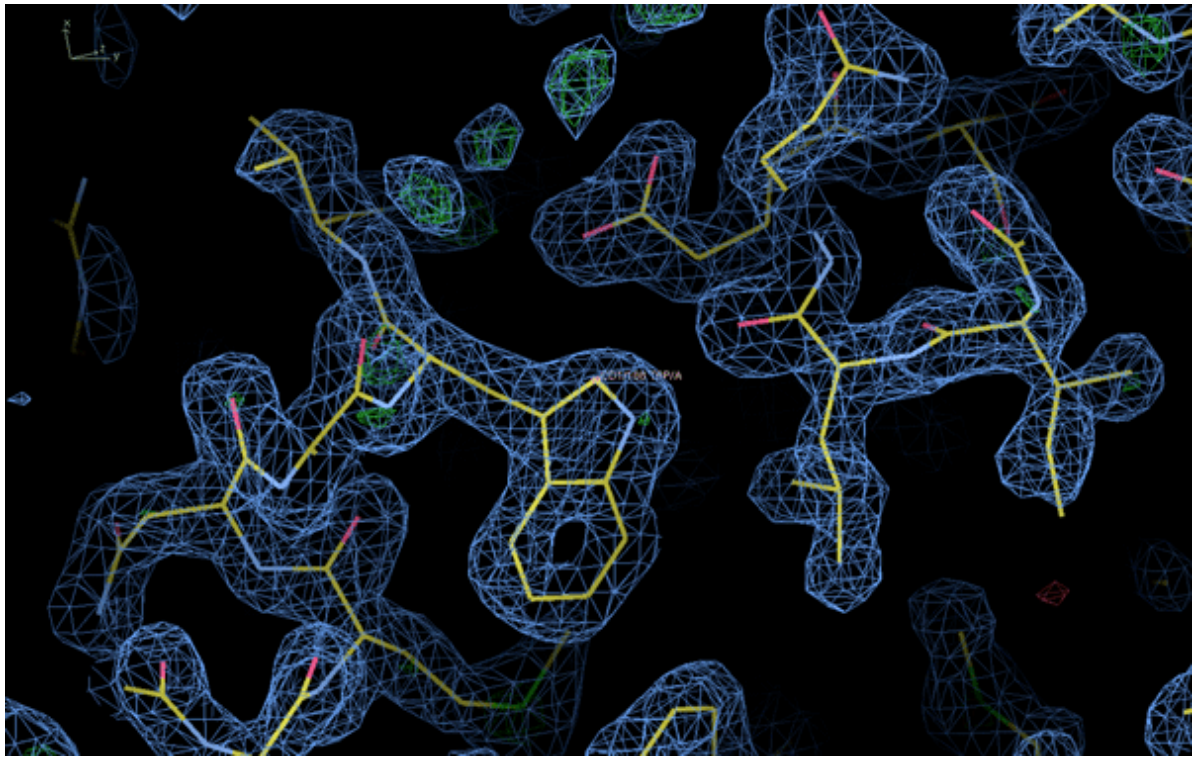
- The most common method to determine structures of proteins and other biomolecules
- Basic idea: Get the molecule whose structure you want to determine to form a crystal
- Then shine an intense beam of x-rays through the crystal, giving rise to a “diffraction pattern” (a pattern of spots of varying brightnesses)



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

The diffraction pattern of a crystal depends on its electron density

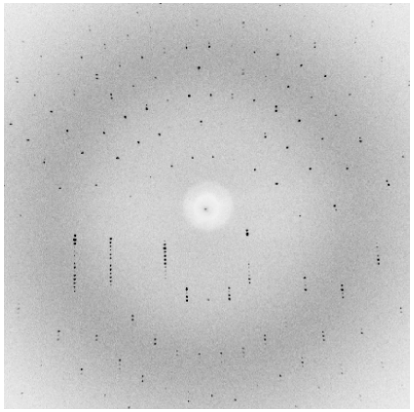
- The *electron density* corresponding to the 3D structure of a molecule gives the probability of finding an electron at each point in space
- X-rays bounce off electrons they hit



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

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 Fourier transform gives a magnitude and a phase (shift) for each sinusoid, but it's only practical to measure the magnitude, not the phase
    - Brightness of the spot gives the magnitude



- To recover the electron density from the diffraction pattern, one needs to estimate the phases
  - In some cases this is done experimentally
  - In most cases it's done computationally (at least in part)

Problem 1: Given electron density, estimate atom positions

- This is relevant when phases have been determined experimentally, so that we can easily compute an electron density
- Now the challenge is to figure out where each atom is located
- For high-resolution structures, this is fairly straightforward
- For low-resolution structures, it's harder.
  - This is the subject of one of Monday's paper
  - The paper uses a method based on *probabilistic graphical models*, a machine learning technique
    - Daphne Koller (Stanford) is a pioneer in this area

Problem 2: Given diffraction pattern (with unknown phases), estimate atom positions

- If one doesn't know the phases (or only knows them approximately), one generally estimates the phases and atom positions simultaneously
- One usually does this by solving an optimization problem where one tries to pick a structure that:
  - Would give a diffraction pattern similar to that observed experimentally **AND**
  - Has low energy according to a force field roughly similar to those used for molecular dynamics (e.g., two carbons bonded to one another should be approximately a certain distance apart)
- Monday's second paper introduces an additional term to this optimization, which favors structures similar to a computational model structure based on structures of related proteins
  - This makes it possible to solve structures from low-resolution diffraction patterns
  - From Axel Brunger and Michael Levitt (Stanford)

Background material

- Crystallography slides from CS/CME/BioE/ Biophys/BMI 279:
  - http://web.stanford.edu/class/cs279/lectures/lecture12.pdf
- Probabilistic Graphical Models course on Coursera:
  - https://www.coursera.org/specializations/probabilistic-graphical-models
  - by Daphne Koller, co-founder of Coursera