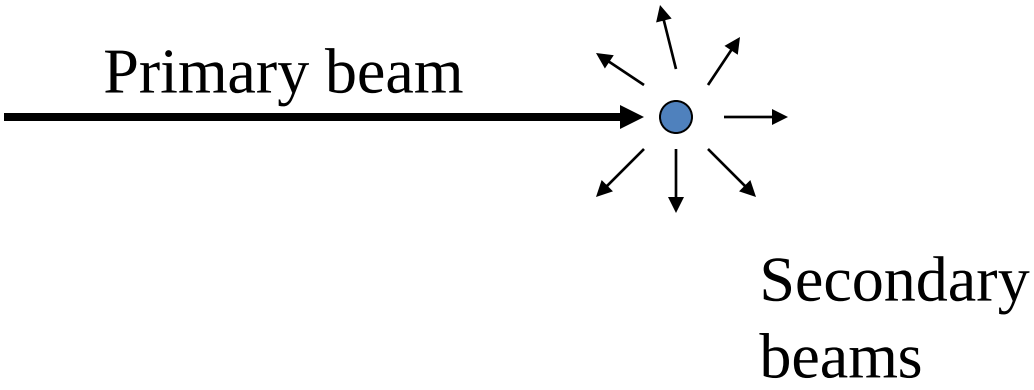
Part 1
X-ray Crystallography

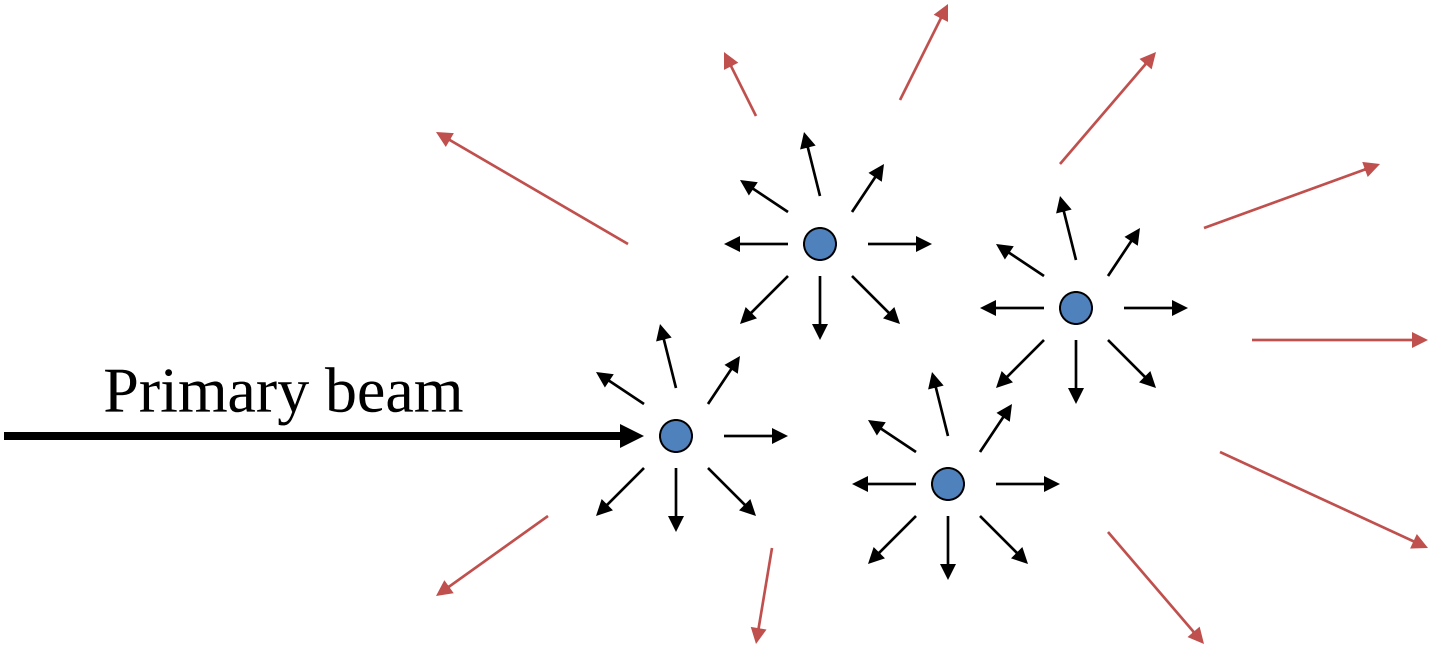
What happens to electron when it is hit by x-rays?

1. The electron starts vibrating with the same frequency as the x-ray beam
2. As a result, secondary beams will be scattered in all directions



Scattering from a molecule

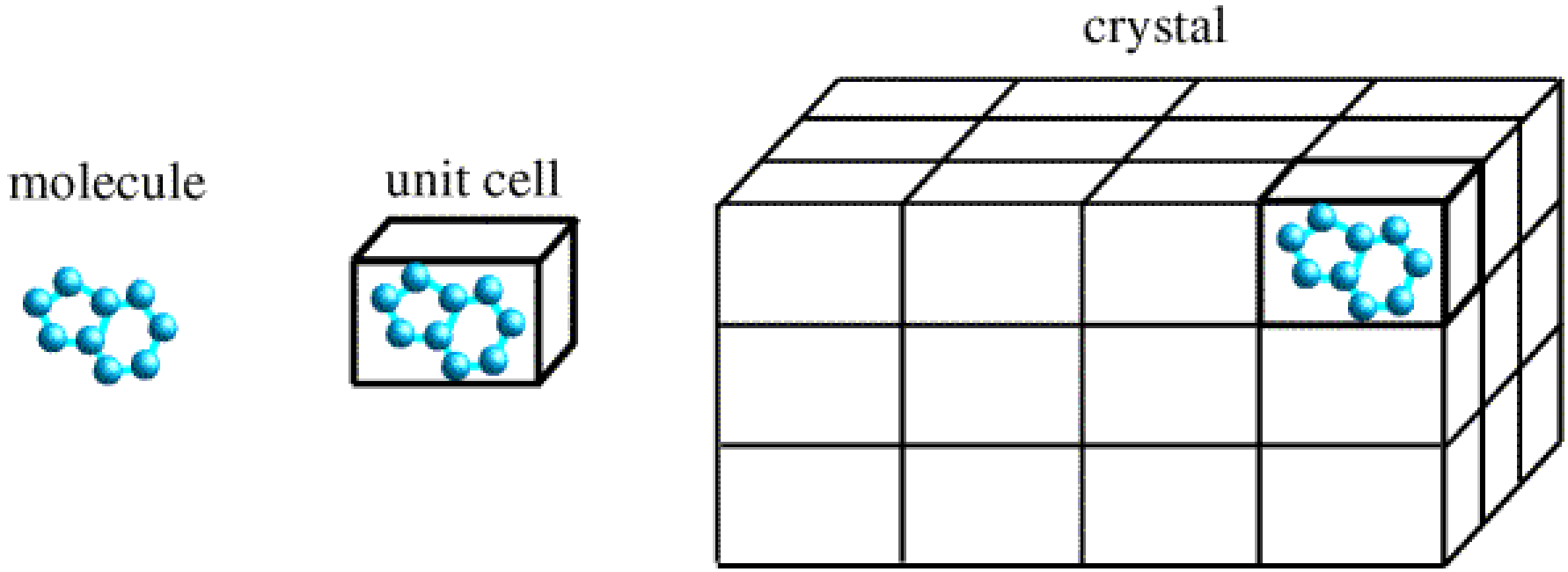
1. Molecule is composed of many electrons
2. Each electron will scatter secondary radiation uppon exposure to x-rays
3. The scattered secondary beams will interact and cause interference
4. The scattering from a molecule is dependent upon the number of and distances between electrons
5. In other words, scattering from molecule is dependent on its atomic structure
6. If we would know the amplitudes and phases of scattered molecule, we could calculate the structure of molecule...

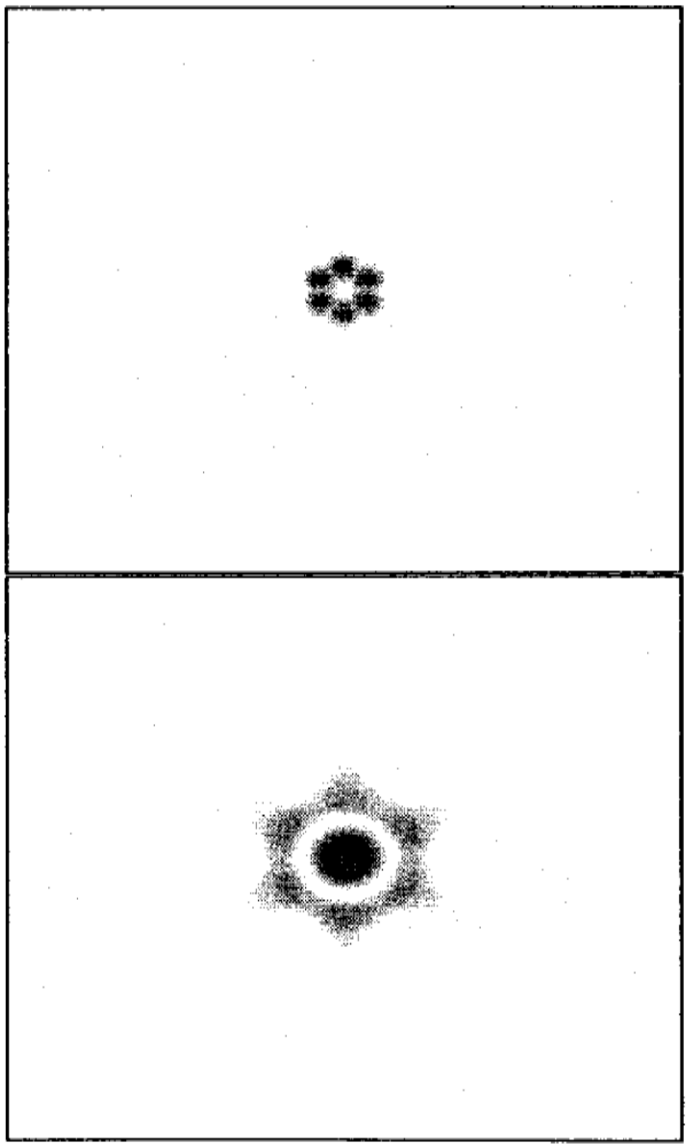


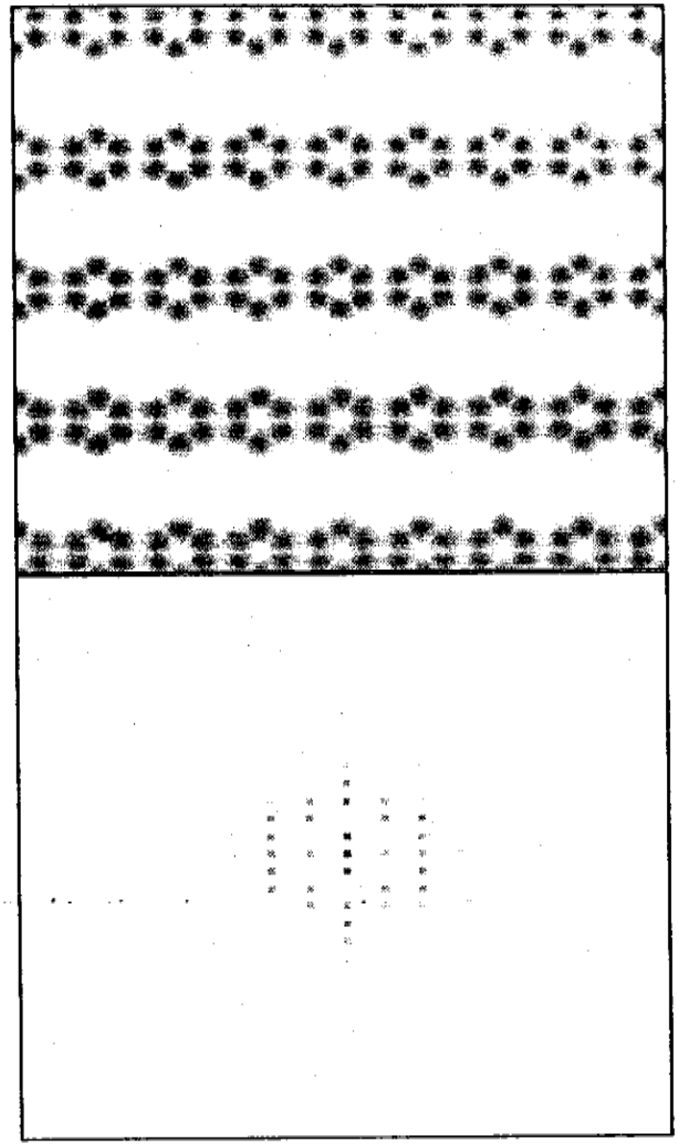
In practice...

- Scattering from a single molecule is far too weak to be observed but possible.
- If molecules are all oriented in the same way (like in crystal), the scattering from individual molecules will be multiplied in certain directions

The unit cell and a crystal



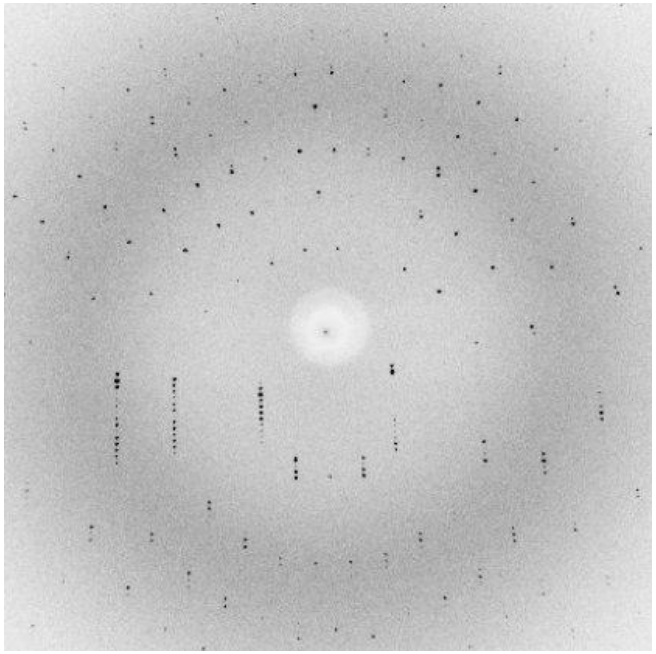


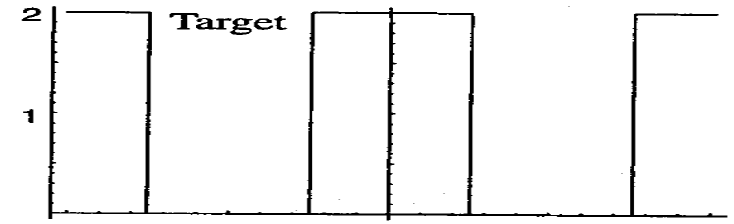


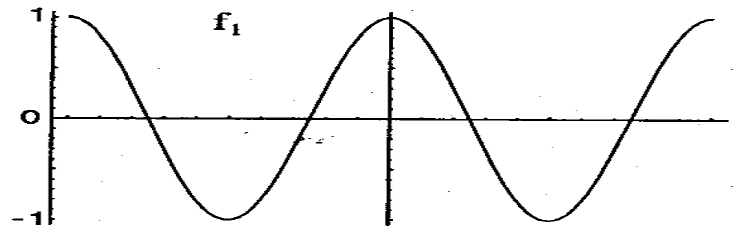
The electron density equation

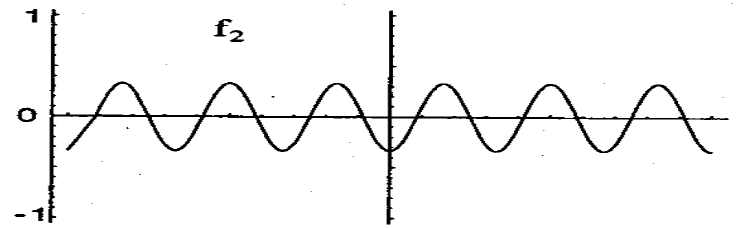
$$\rho(xyz) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F(hkl)\exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)]$$

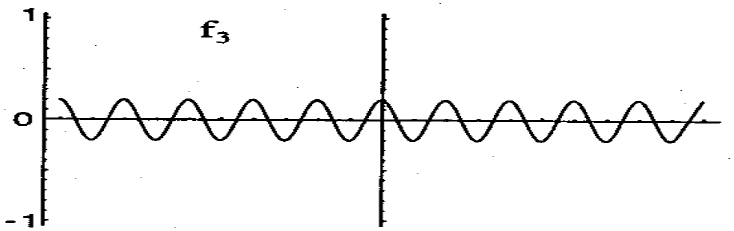
- h,k,l – indices of reflections
- xyz – coordinates
- F – amplitude of reflections
- α – phase of reflections
- V- unit cell volume

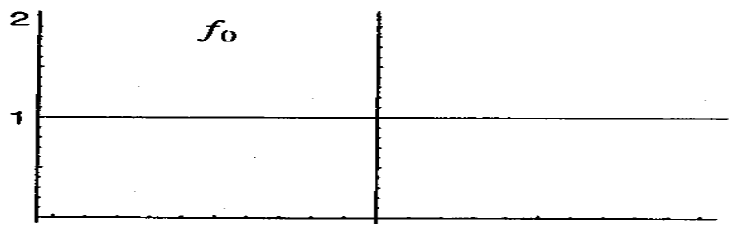


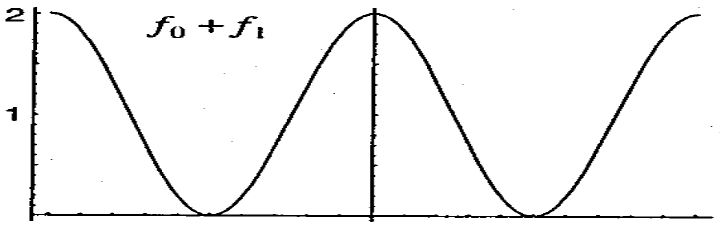

2
1
Target

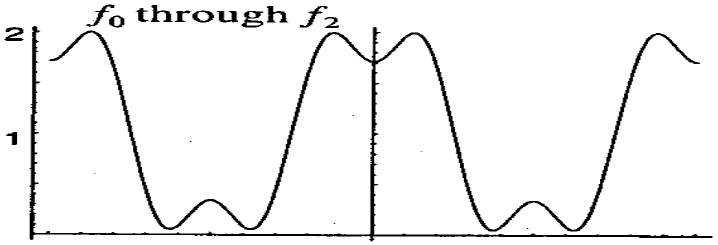

1
0
-1
f1

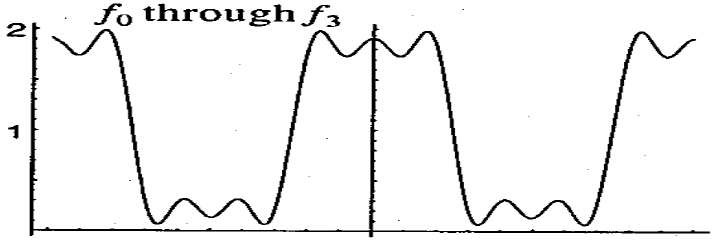

1
0
-1
f2

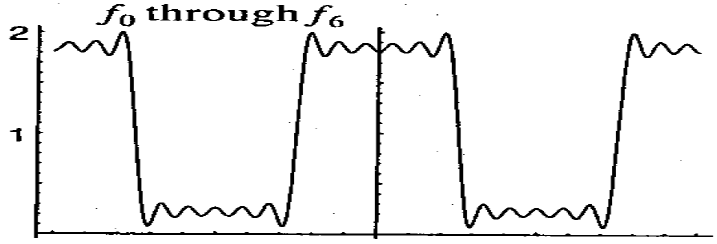

1
0
-1
f3


2
1
f0


2
1
f0 + f1

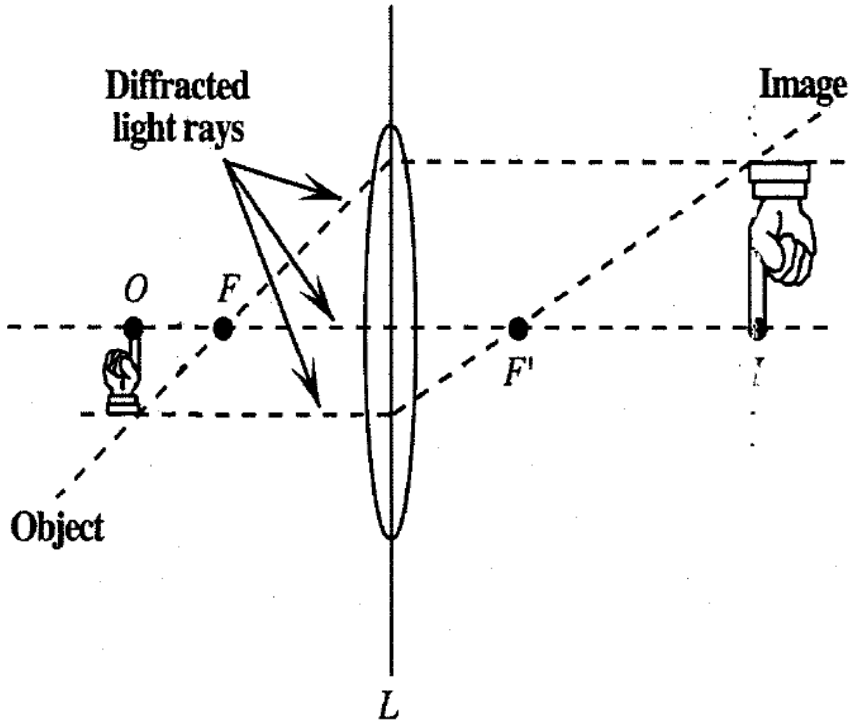

2
1
f0 through f2

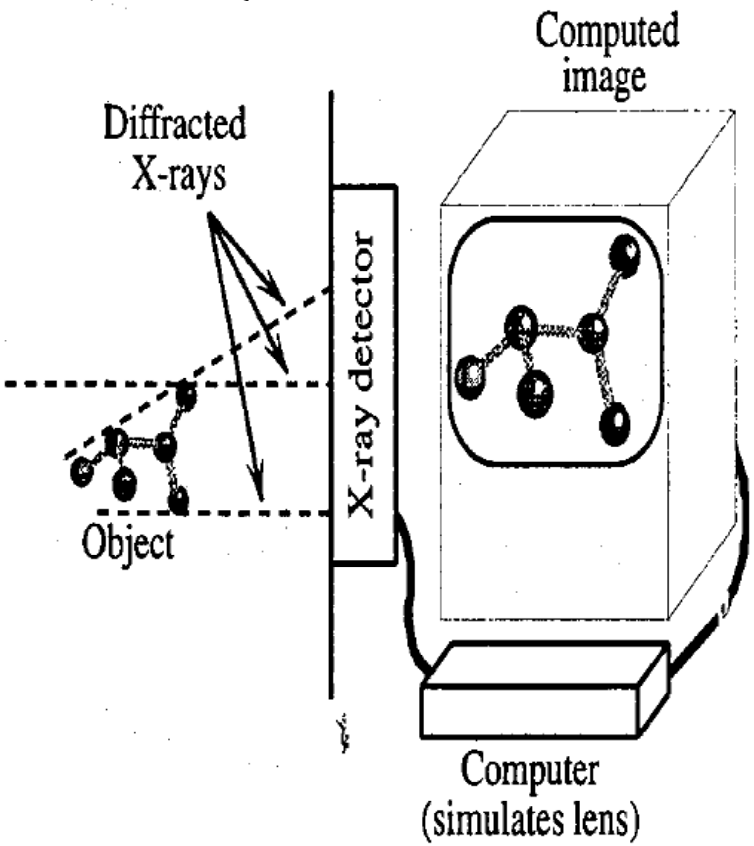

2
1
f0 through f3


2
1
f0 through f6

The Phase Problem

- With detector you can measure only the intensity of reflections
- The information about phases is lost – there is no such thing as “phase meter” because we have no lenses for x-rays.
- This means, you must obtain phase information in some other way
- For small molecules (<100 atoms), direct methods exist. This means, that you can calculate phases from amplitudes without any extra information.
- Proteins are far too big to use direct methods, so other tools are developed


Diffracted
light rays
Image
O
F
F'
Object
L


Computed
image
Diffracted
X-rays
X-ray detector
Object
Computer
(simulates lens)

Isomorphous replacement

- By introducing heavy atoms in protein crystal (by soaking), the diffraction pattern can be altered
- It is possible to determine positions of heavy atoms and from them the phases
- One must use at least 2 different heavy atom soaks
- Problems:
- 1) Unit cell dimensions of crystal might change upon soaking
- 2) Crystal might get destroyed upon soaking and not diffract at all
- 3) Heavy atom ions might not bind in well defined places

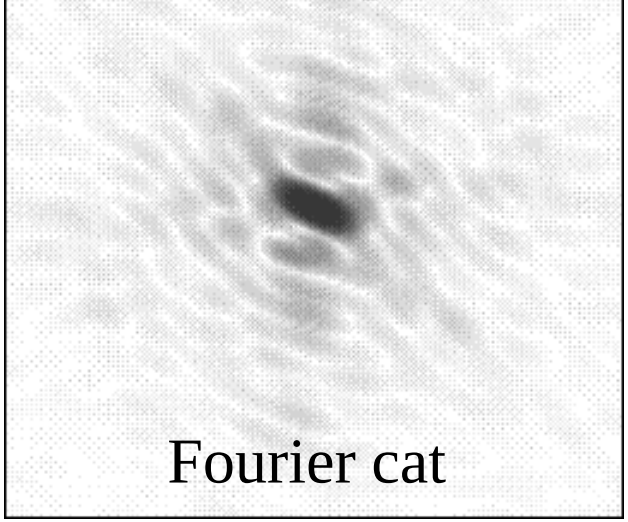
Molecular replacement

- Currently the most common technique
- Applicable only if a similar structure already exists (at least 25% sequence identity)
- The phases of known structure are combined with intensities of unknown
- Before that, the known model has to be *in silico* placed in an artifical unit cell in the same orientation and translation from origin as in the structure of interest
- For this, rotation and translation functions exist

Problems:

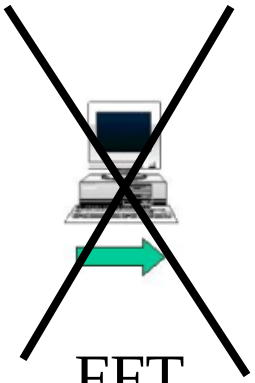
- May not work, if unknown structure is less than 30 % idendical to the known structure
- Model bias – what's that?

Observed amplitudes



Fourier cat

Phases unknown!



FFT

Unknown structure

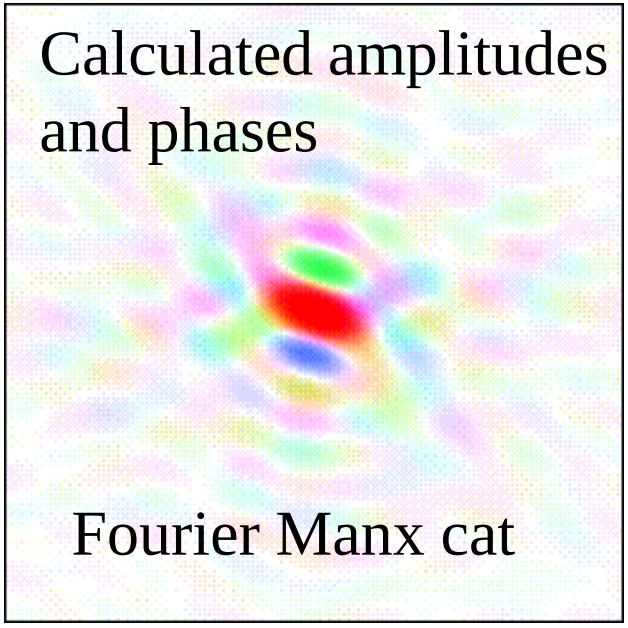
Cat

Known structure

Manx cat

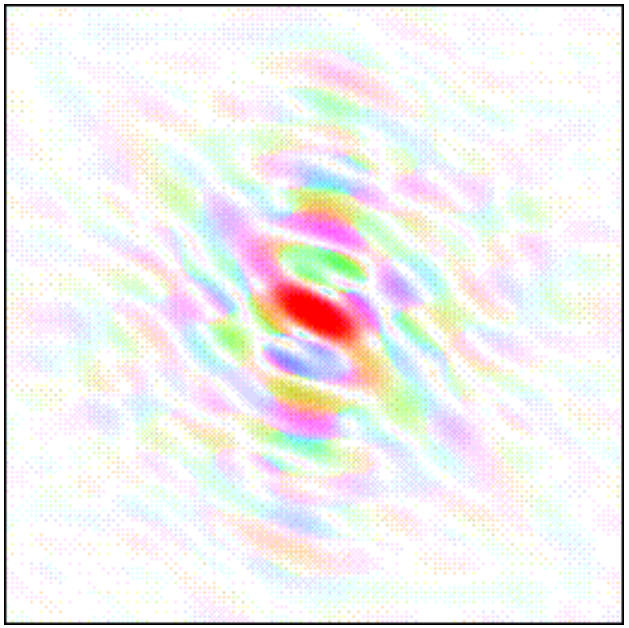
FFT

Calculated amplitudes and phases



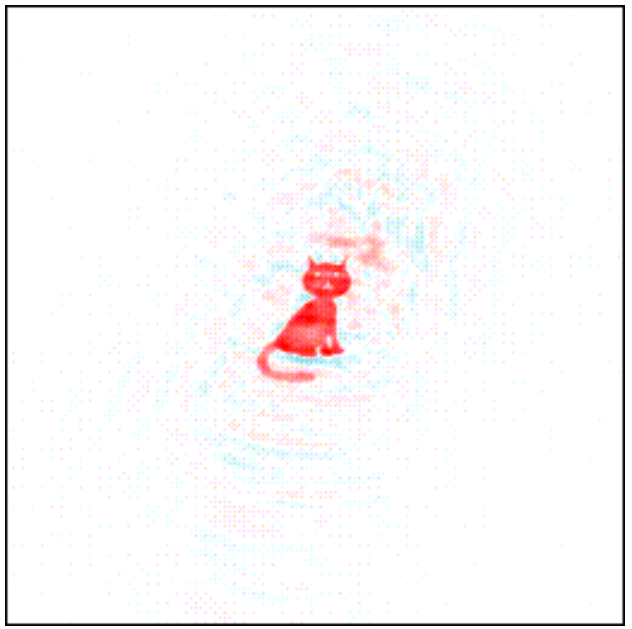
Fourier Manx cat

Observed amplitudes,
calculated phases



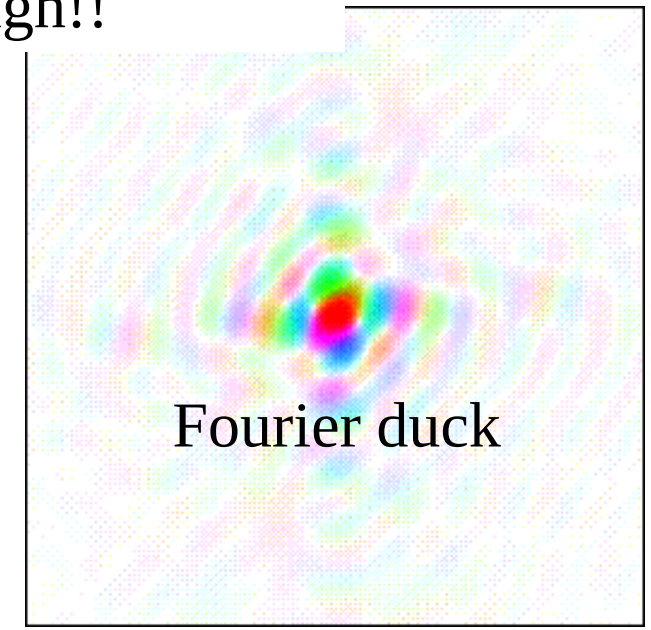
FFT

The tail becomes visible!



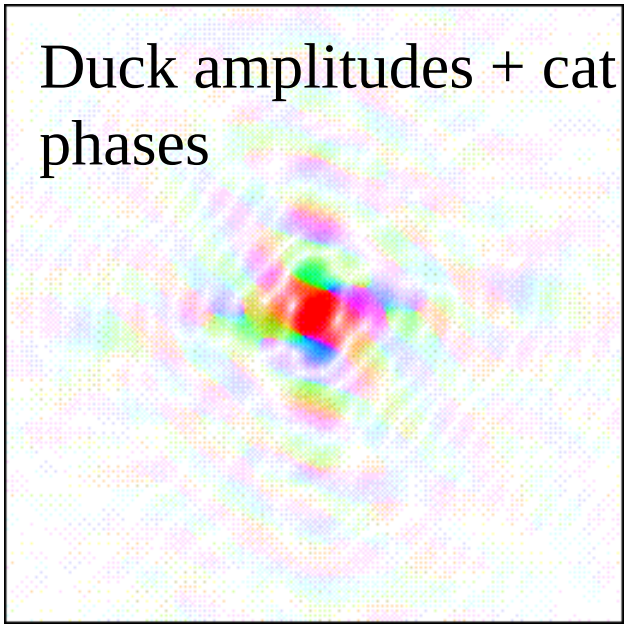
Be aware – this happens, if structures are not similar enough!!

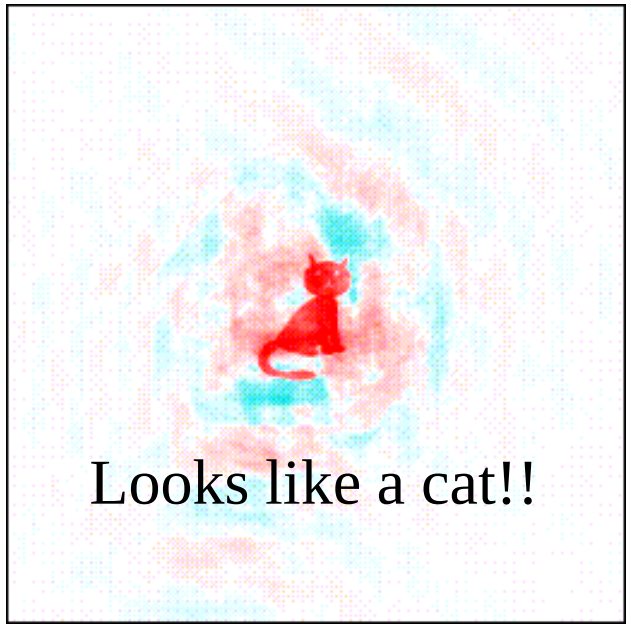
Duck



Fourier duck

Duck amplitudes + cat phases

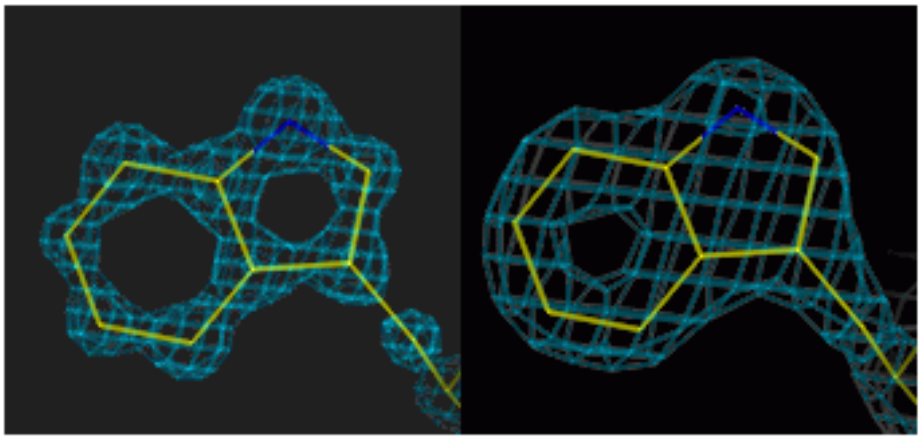




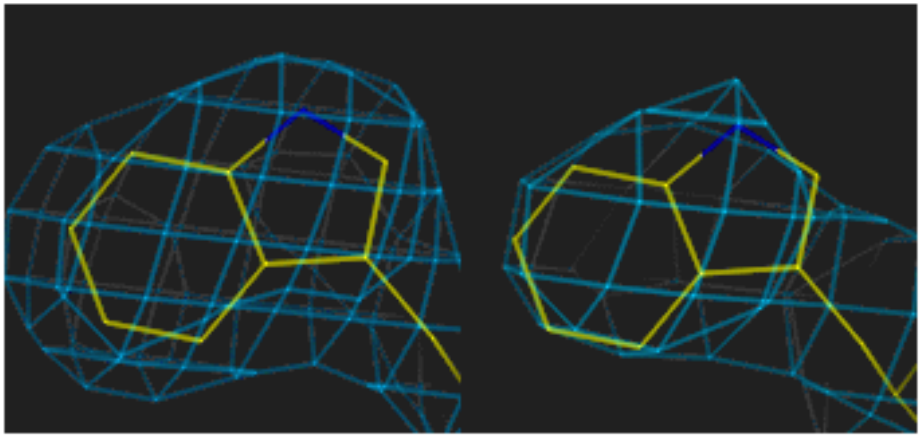
Looks like a cat!!

Model building

- Fitting of protein sequence in the electron density
- Easy in molecular replacement
- More difficult if no initial model is available
- Unambiquous if resolution is high enough (better than 3.0 Å)
- Can be automated, if resolution is close to 2Å or better



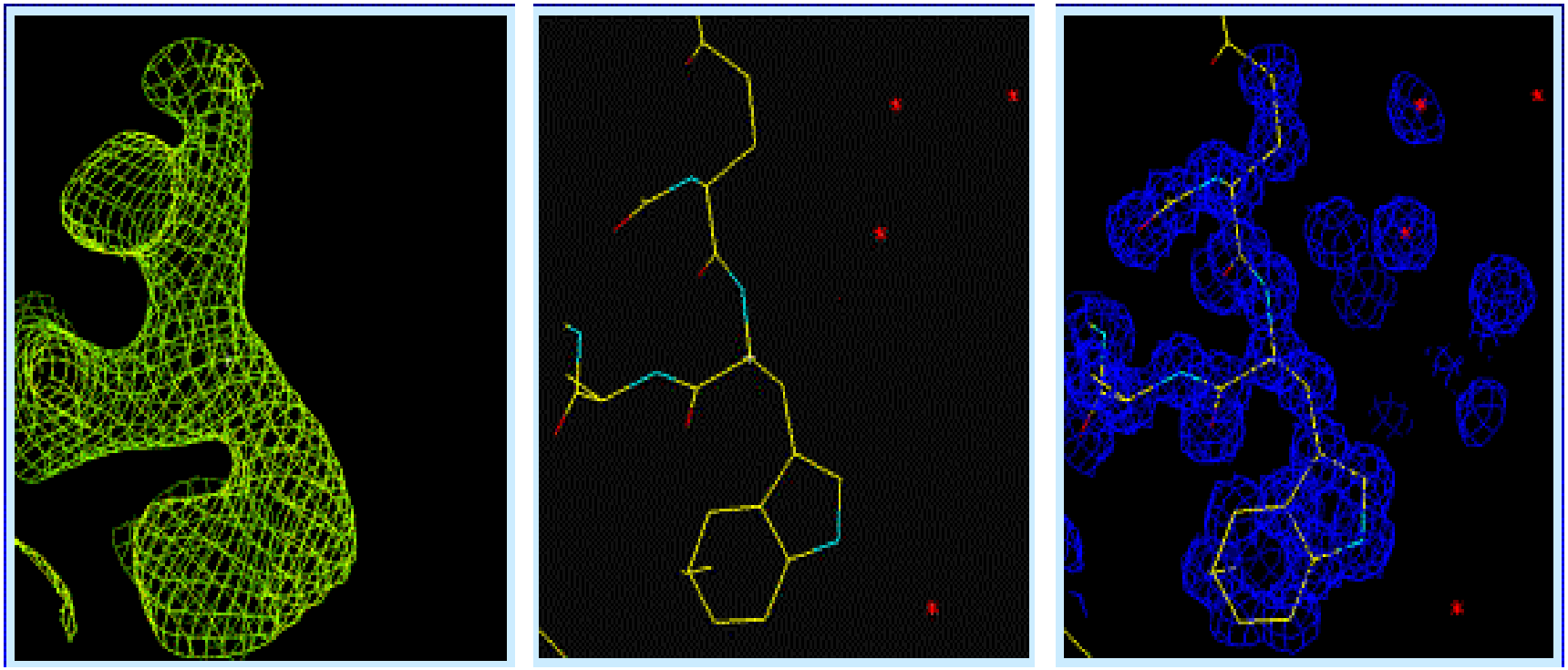
1.0Å 2.5Å



3.0Å 4.0Å

Refinement

- Automated improvement of the model, so it explains the observed data better
- The phases get improved as well, so the electron density maps get better



Validation

- Assesment of the final(?) model quality
- How the geometry of amino acids look like? (Ramachandran plot)
- Are non-covalently atoms far enough from each other? (no atom bumps)
- Are residues “happy” in their environment? (hydrophobic in core, polar on surface)
- Are the hydrogen donors/acceptors satisfied?

Part 2
NMR Structure Determination

NMR Interactions

NOE

- a through space correlation (<5Å)
- distance constraint

Coupling Constant (J)

- through bond correlation
- dihedral angle constraint

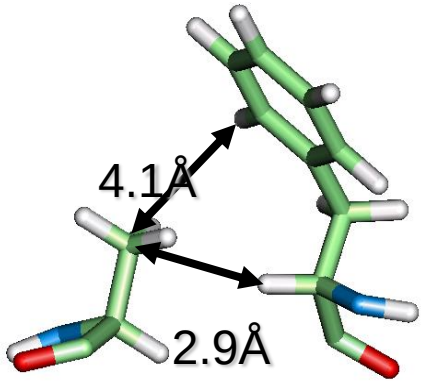
Chemical Shift (δ)

- very sensitive to local changes in environment
- dihedral angle constraint

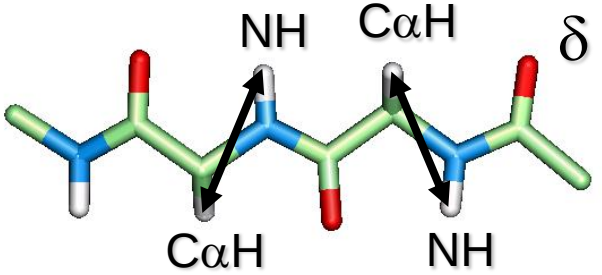
Dipolar coupling constants (D)

- bond vector orientation relative to magnetic field
- alignment with bicelles or viruses

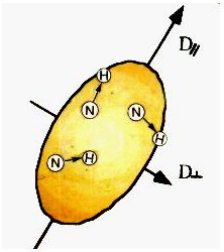
NOE

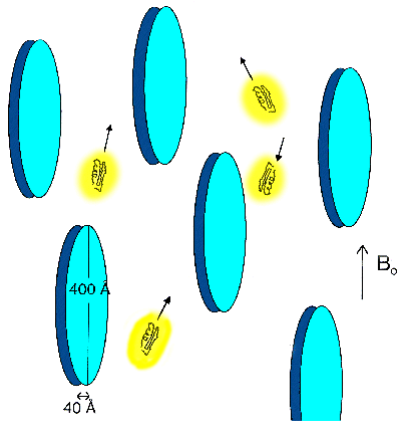


J

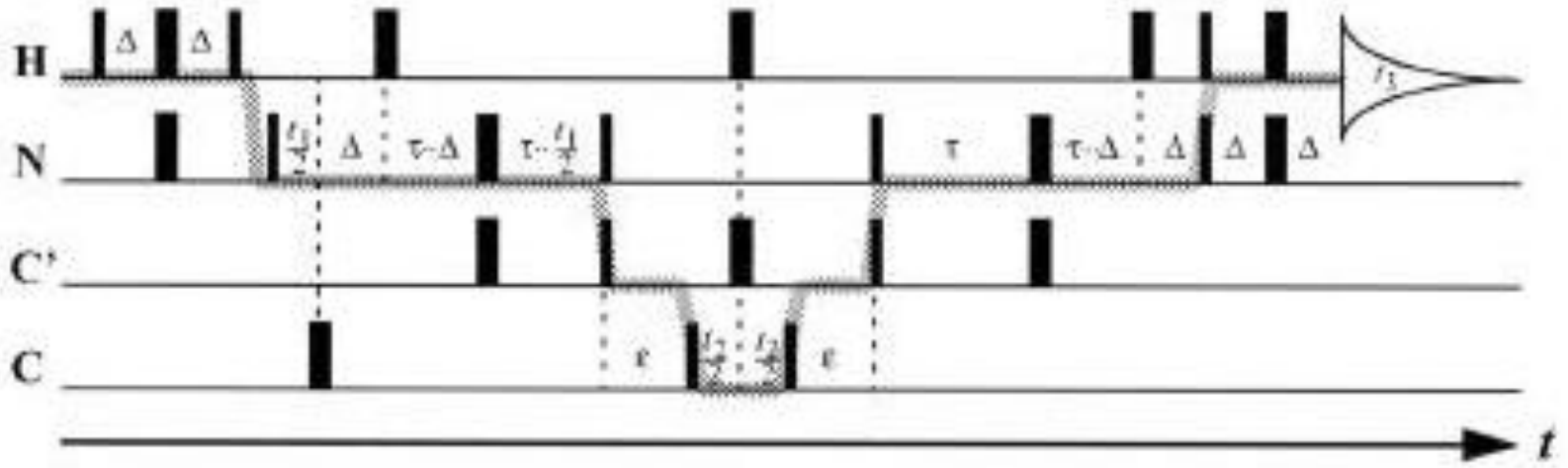


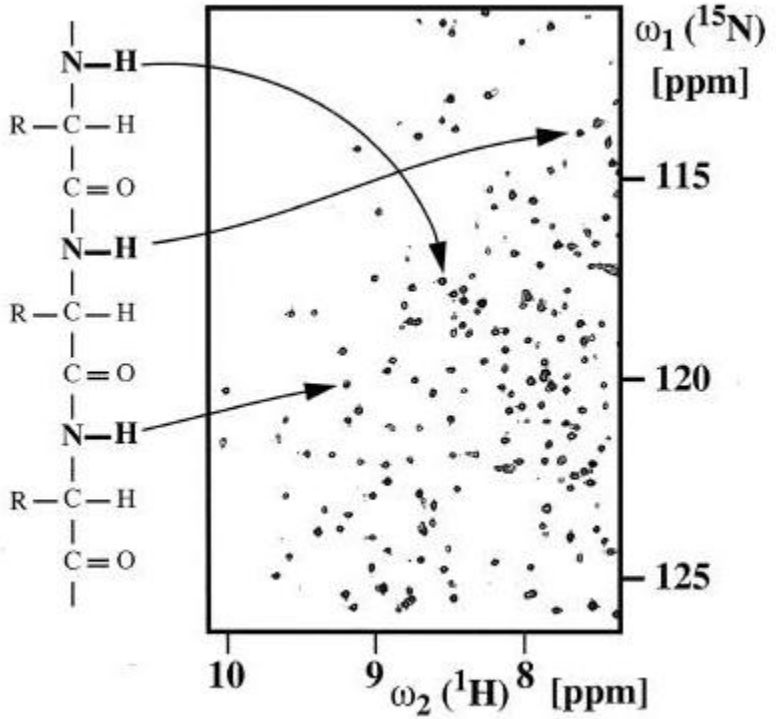
D



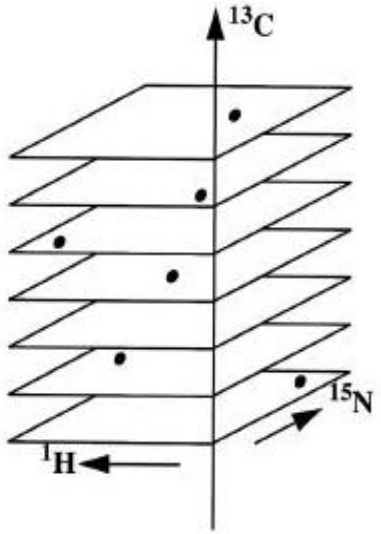


COSY- COrrelation SpectroscopY





A



B

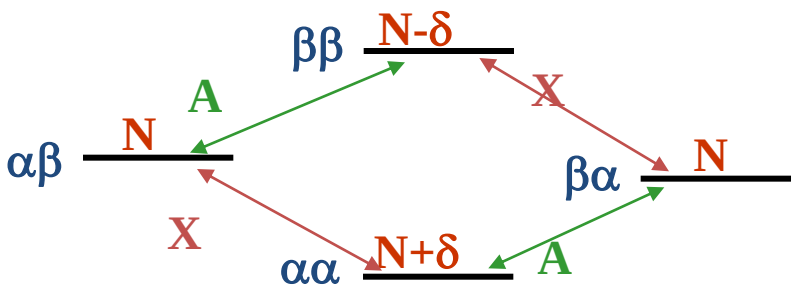
NOE- Nuclear Overhauser Effect

Nuclear Overhauser Effect (NOE, h) – the change in intensity of an NMR resonance when the transition of another is perturbed, usually by saturation.

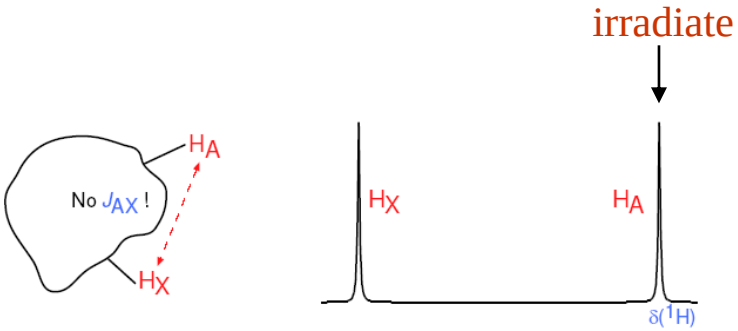
$$\eta_i = (I\text{-}I_o)/I_o$$

where I_o is thermal equilibrium intensity

Saturation – elimination of a population difference between transitions (irradiating one transition with a weak RF field)

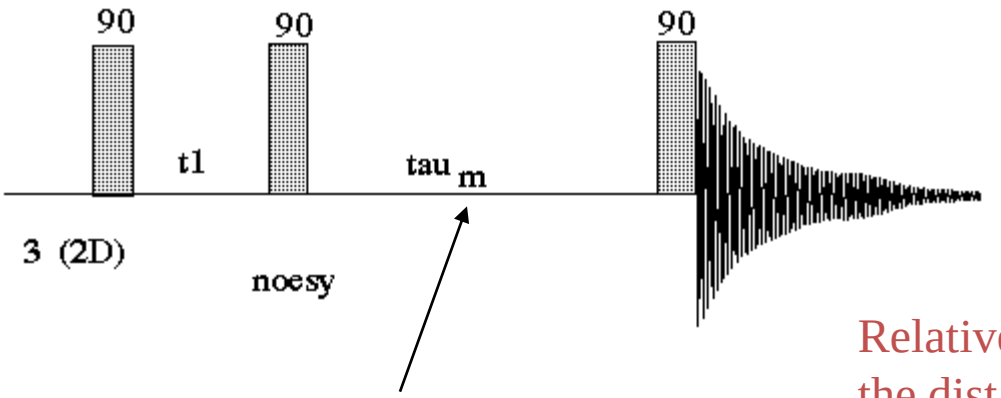


Populations and energy levels of a homonuclear AX system (large chemical shift difference)



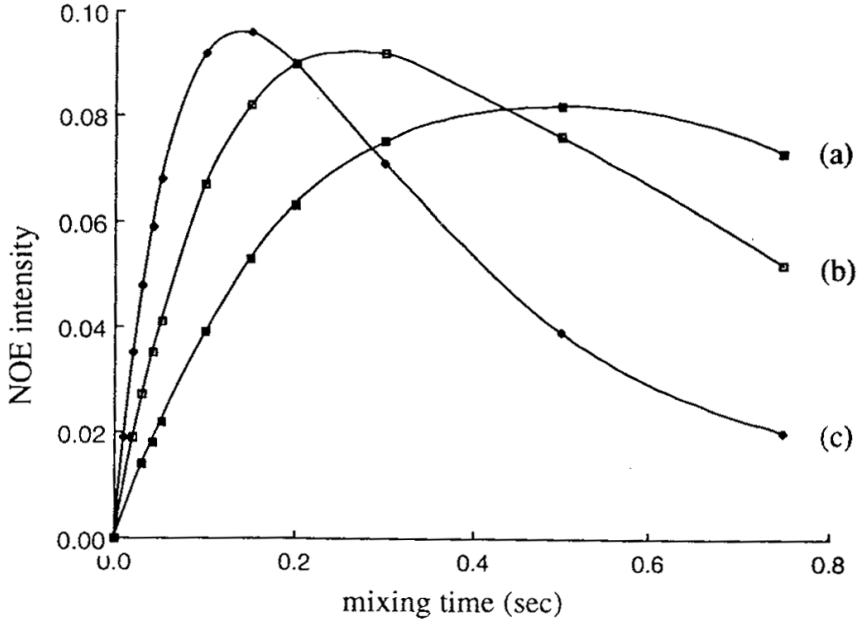
Observed signals only occur from single-quantum transitions

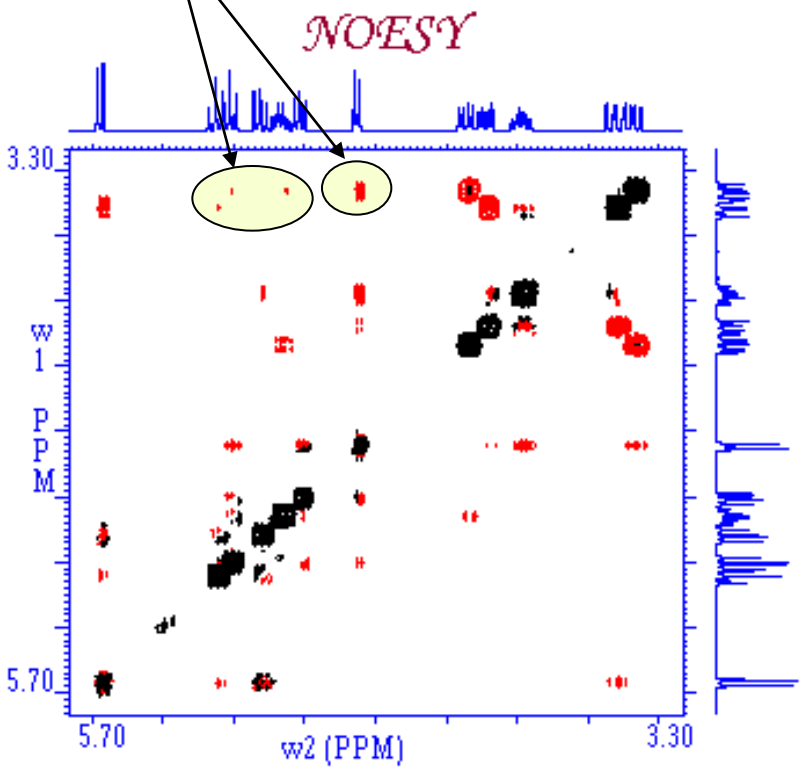
2D NOESY (Nuclear Overhauser Effect)



Relative magnitude of the cross-peak is related to the distance ($1/r^6$) between the protons (≥ 5Å).

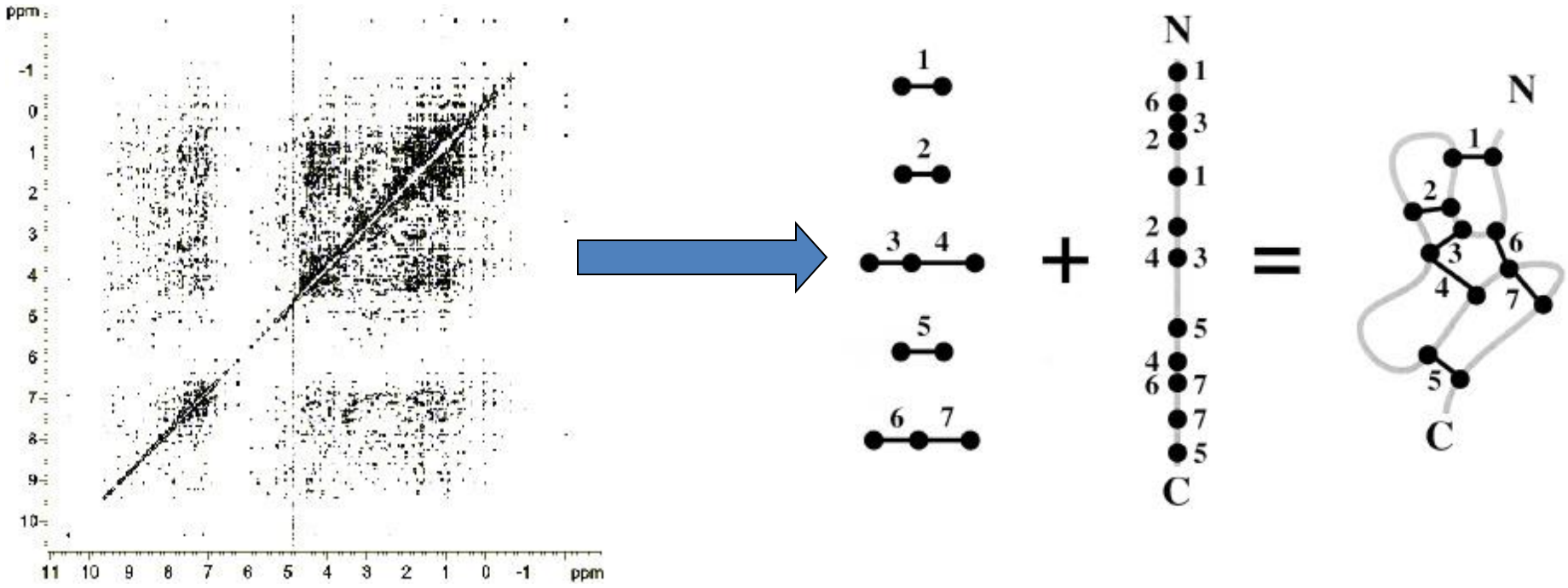
NOE is a relaxation factor that builds-up during
The “mixing-time (tm)





NMR Structure Determination

Going from NOESY Data to a structure



2D NOESY Spectra at 900 MHz

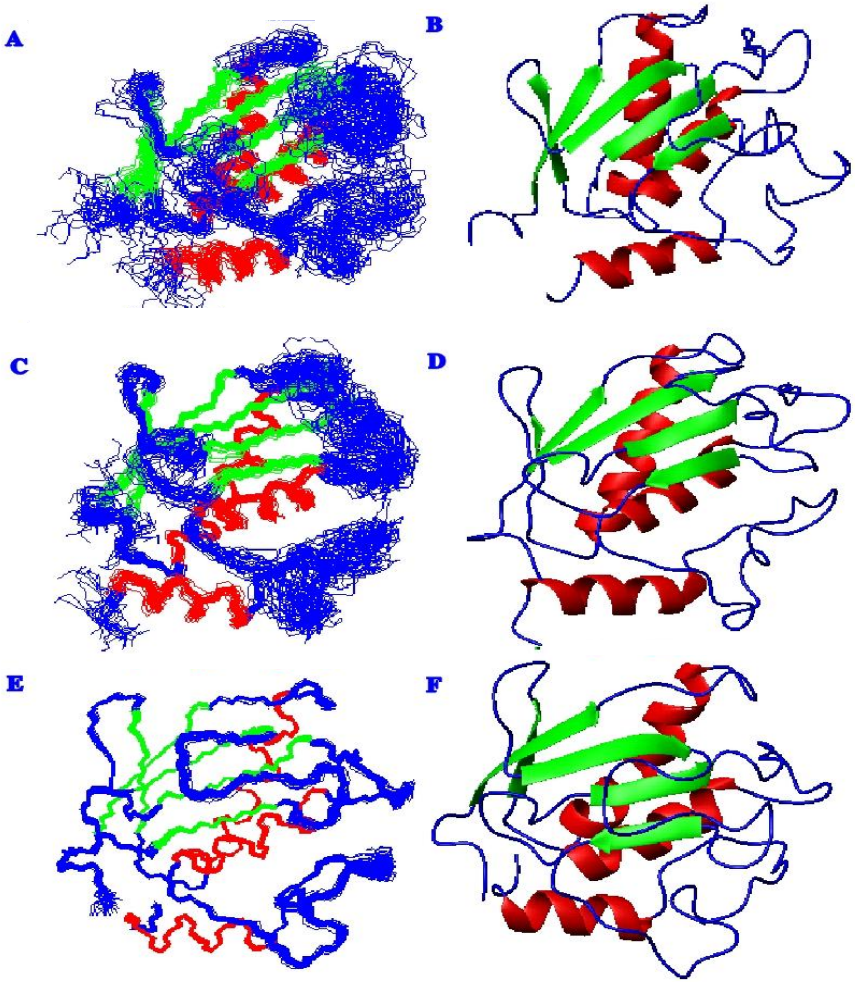
Once we *complete the Assignment of All the Remaining NOEs in an Iterative Process to Obtain the Structure*

Protein Structure Refinement

Iterative Cycle

Distance Constraints Assignments
Stereospecific Assignments
Torsion–Angle Assignments

3D Structure Determination

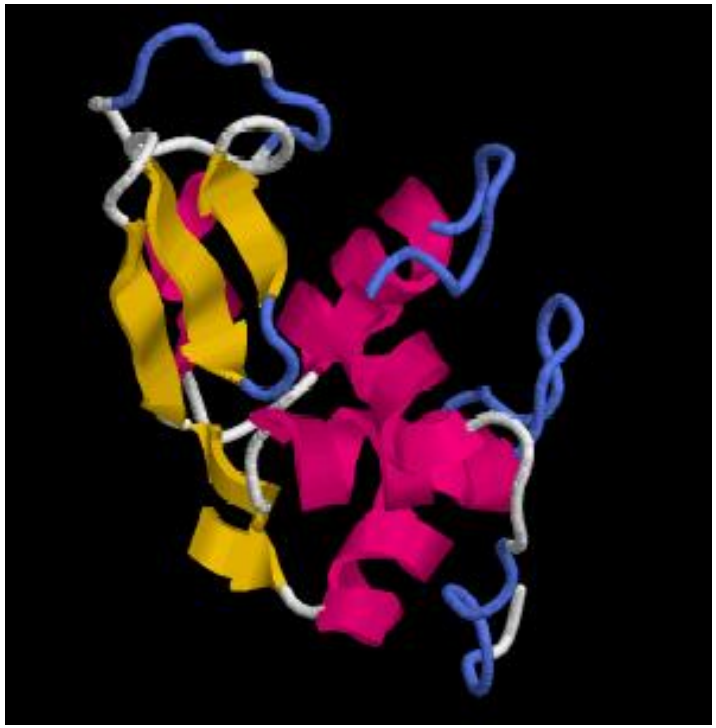


More Constraints the Better the Structure

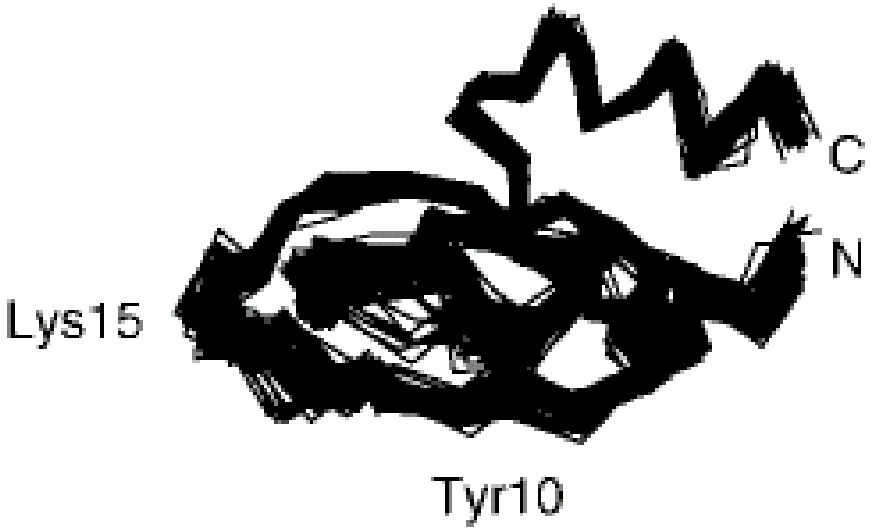
NMR Structure Determination

Two Very Important Facts to Remember

- NOEs Reflect the *Average* Distance
- Protein Structures Are *Dynamic*



We visualize protein structures as a static image



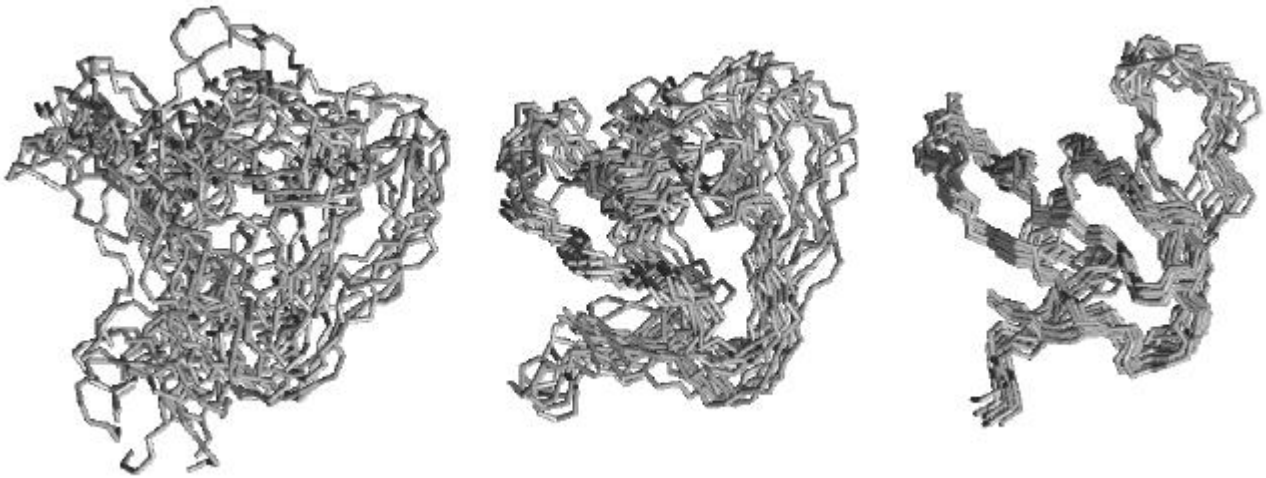
In reality, protein undergoes wide-ranges of motions (snapshots of 100 BPTI conformations)

J. Mol. Biol. (1999) 285, 727±74

Quality of NMR structures

Improving the Quality of NMR Structures

- Stereospecific Assignments
  - Making stereospecific assignments increase the relative number of distance constraints while also tightening the upper bounds of the constraints
  - There is a direct correlation between the quality of the NMR structure and the number of distance constraints
    - more constraints → higher the precision of the structure



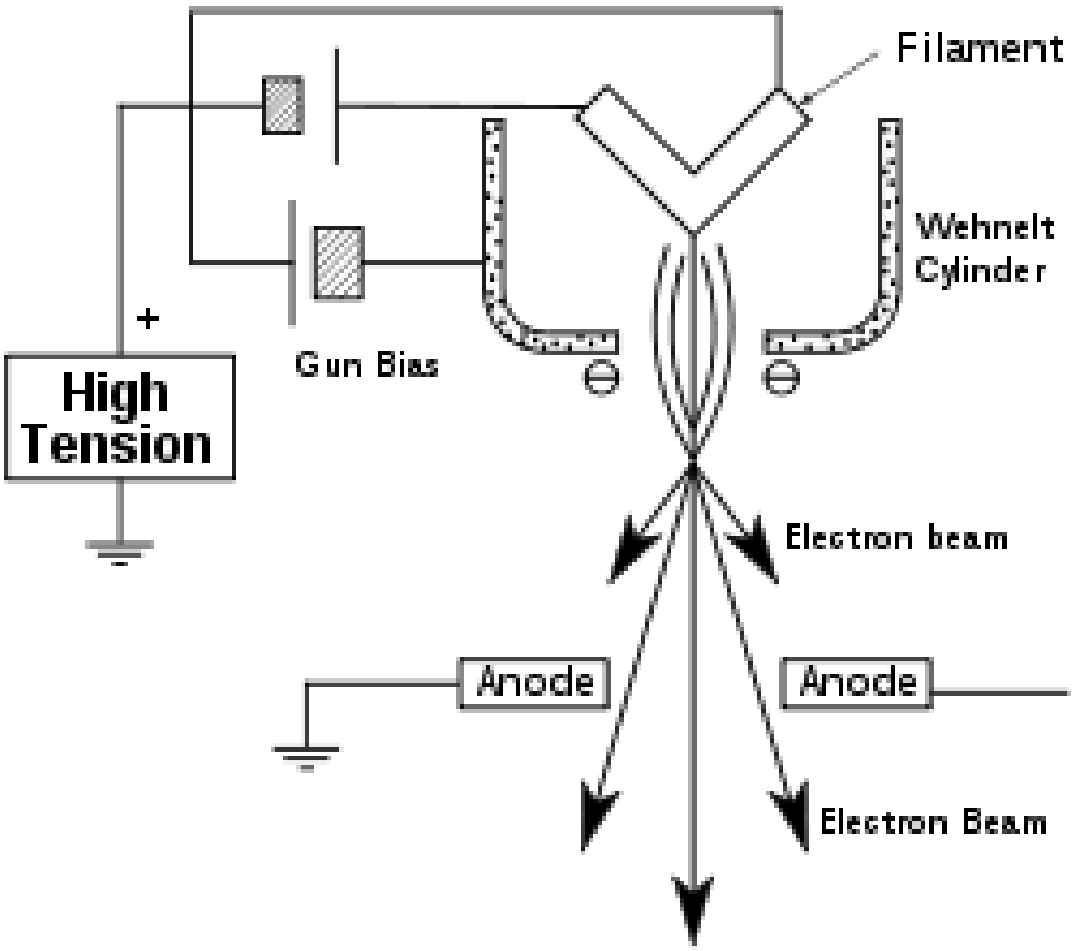
Increasing Number of NOE Based Constraints

Part 3
Structure Determination by Electron Microscopy

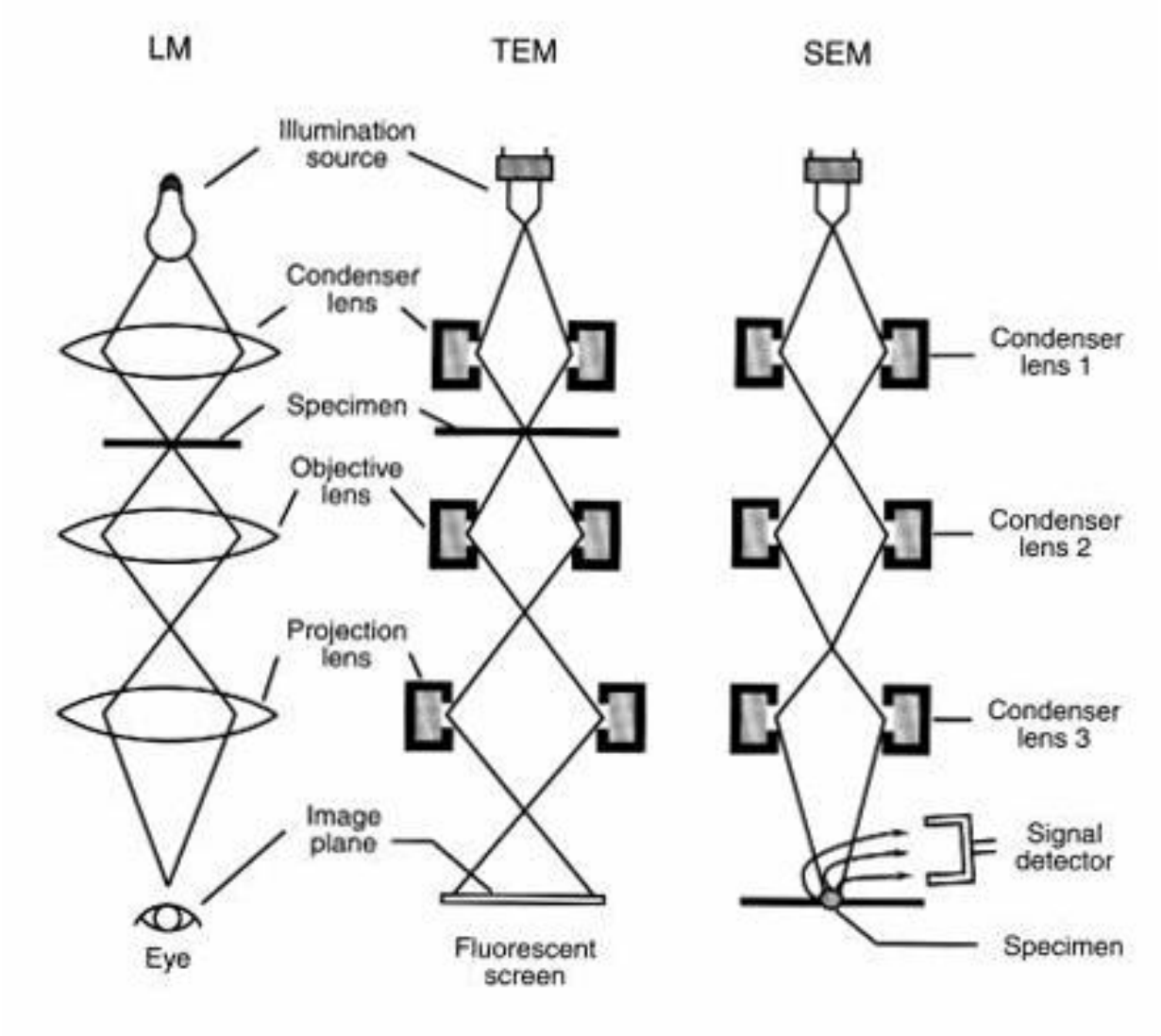
References and other useful material

- Texts
  - Biophysical Electron Microscopy: Basic Concepts and Modern Techniques by U. Valdre (Editor), Peter W. Hawkes (Editor)
  - Three-Dimensional Electron Microscopy of Macromolecular Assemblies by Joachim Frank
  - Negative Staining and Cryoelectron Microscopy: The Thin Film Techniques by Robin J. Harris, James R. Harris
- Reviews
  - Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* **28**, 171-93 (1995).
  - Glaeser, R. M. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J Struct Biol* **128**, 3-14 (1999).
  - Stowell, M. H., Miyazawa, A. & Unwin, N. Macromolecular structure determination by electron microscopy: new advances and recent results. *Curr Opin Struct Biol* **8**, 595-600 (1998).
- Web
  - http://ncmi.bcm.tmc.edu/%7Estevel/spintro/siframes.htm
  - http://cryoem.berkeley.edu/~nieder/em_for_dummies/

Thermionic emission (Shottkey effect)



Electron lenses and “optical path”



Why use electrons….Part 1

?

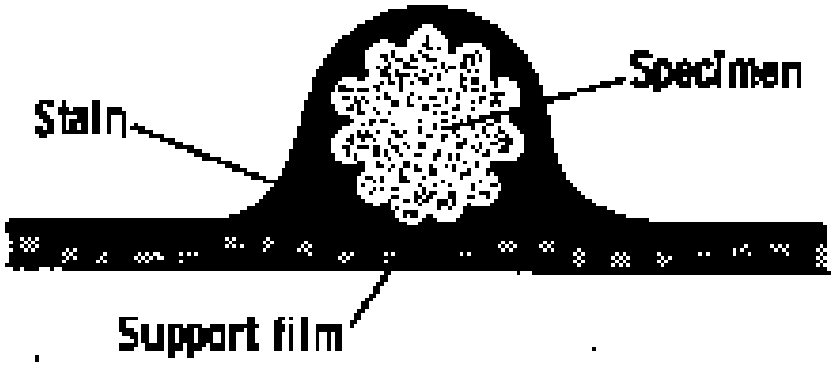
Why use electrons…Part 2

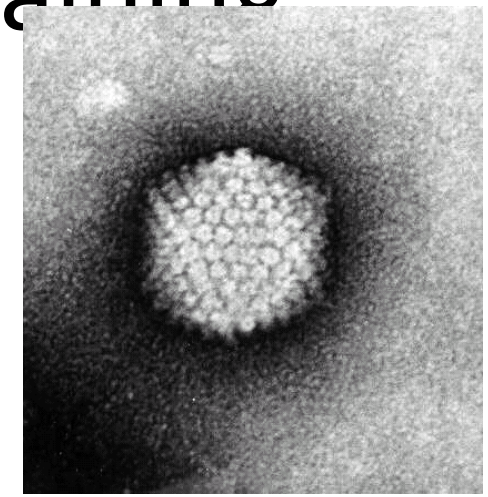
	Electrons 80-500keV	X-rays 1.5 A
Ratio inelastic/elastic	3	10
Mechanism of damage	2^{nd} e^- emission	Photoelectric e^- emission
Energy per inelastic event	20 eV	8 keV
Energy per elastic event	60 eV	80 eV
Energy relative to electrons		
inelastic (Compton)	1	400
elastic (Rayleigh)	1	1000

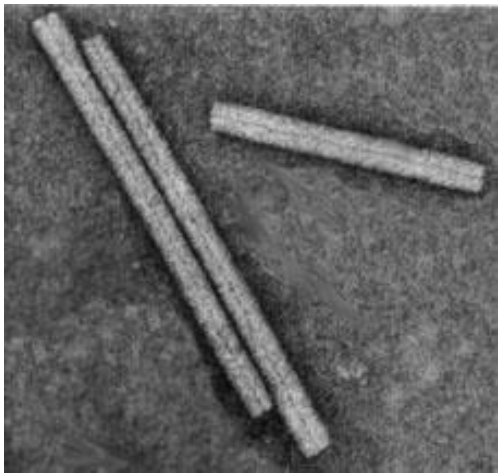
Negative Stain and Cryo

- Negative stain (usually 0.5% uranyl acetate)
  - Easy to prepare
  - Good contrast
  - Preservation
  - Sample distortion
  - Resolution limited to about 20 angstroms
- Cryo
  - Difficult sample prep
  - Low contrast
  - Best preservation and therefore resolution

Negative staining





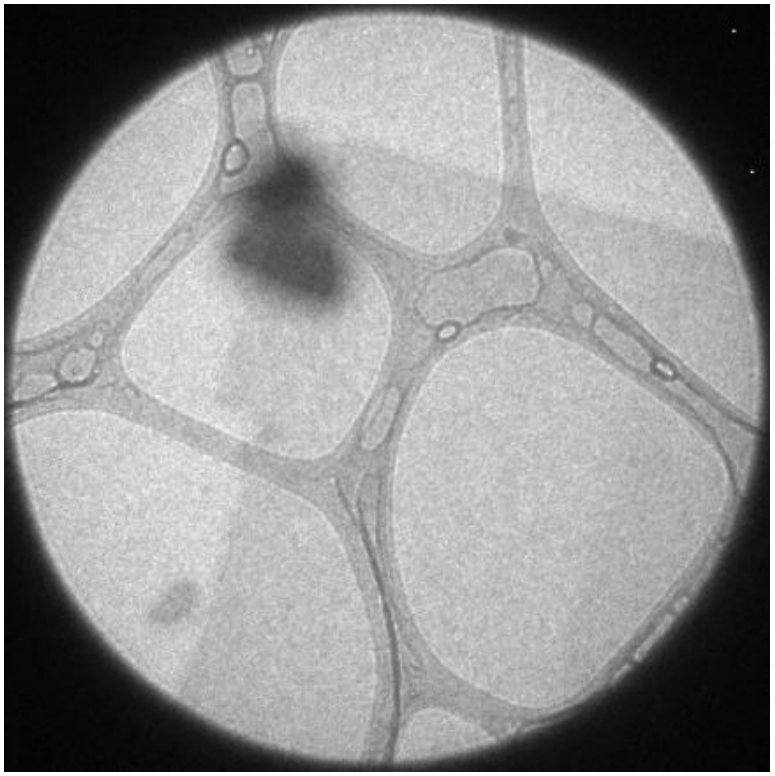


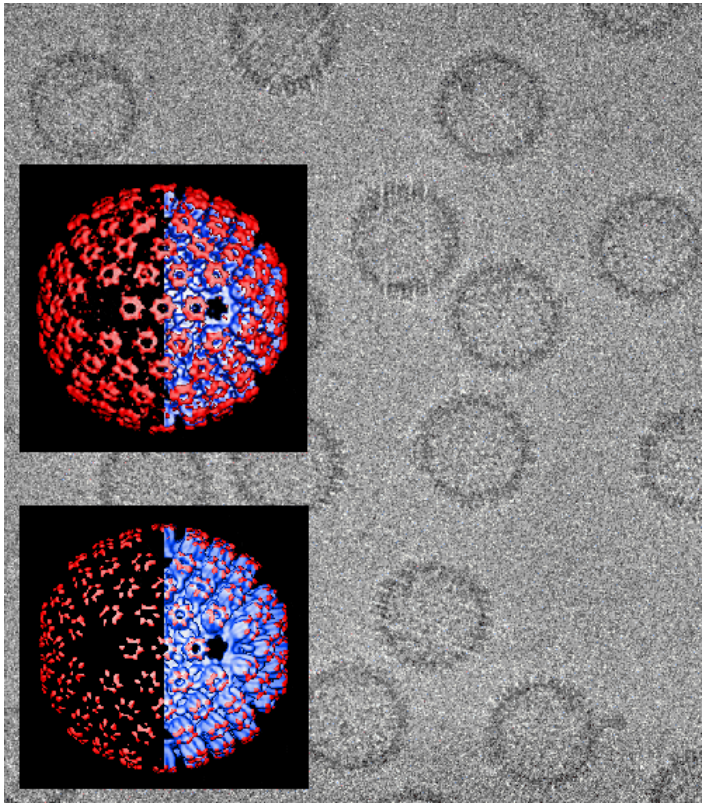
Bob Horne (Cambridge)

Cryo prep using holey film

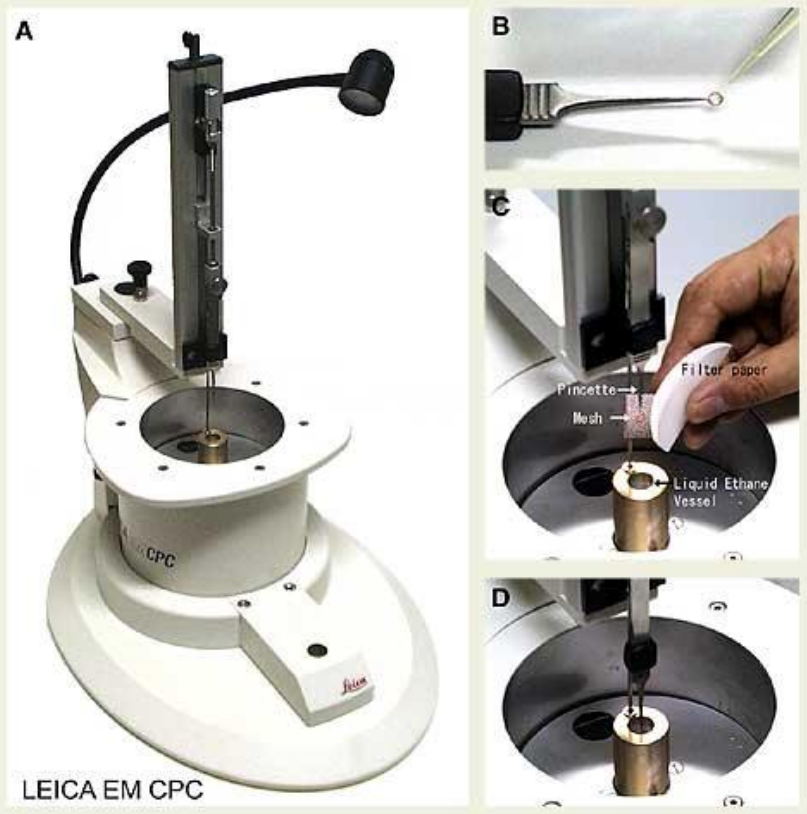
H. Fernandez-moran
B. Glaeser
K. Taylor
J. Dubochet

Aaron Klug

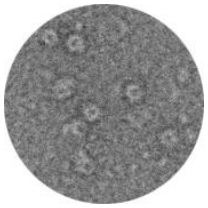




Flash freeze in liquid ethane

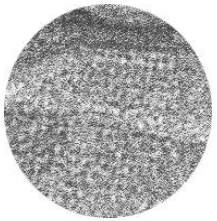


Samples

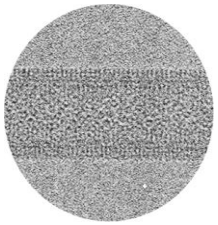


- Single Particles (Proteins, Ribosome)
  - No crystallization
  - Weak amplitude, no diffraction, alignment ambiguity, particle flexibility
  - ~7 angstroms

- Fibers and filaments (tubulin, collagen)
  - No crystallization, 2D distortion corrections, phase restrictions
  - Weak amplitude, no diffraction
  - ~9 angstroms



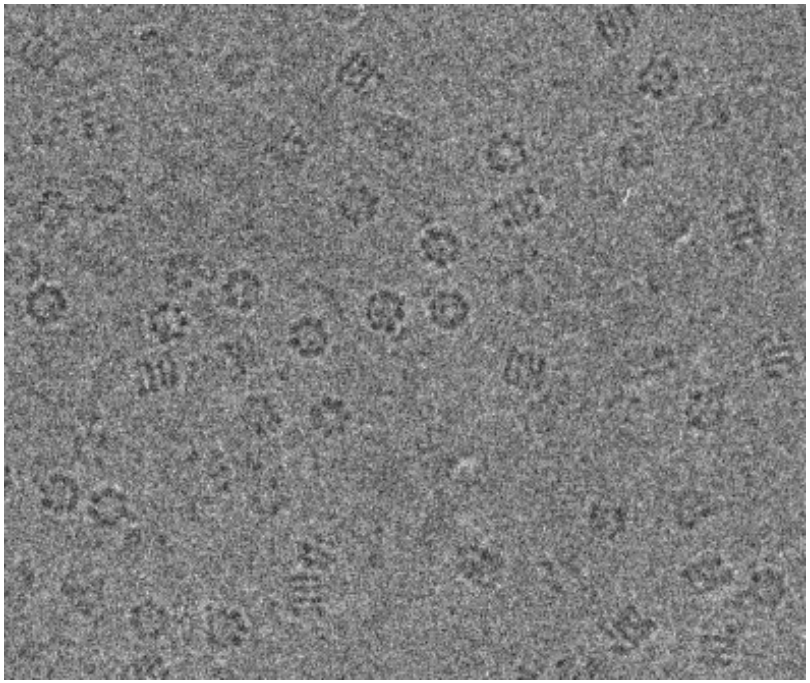
- 2D crystals (BR, AQP, LHCII)
  - Diffraction amplitudes, 2D distortion corrections, crystallographic methods
  - Crystallization, many tilts required, anisotropic data
  - ~3 angstroms



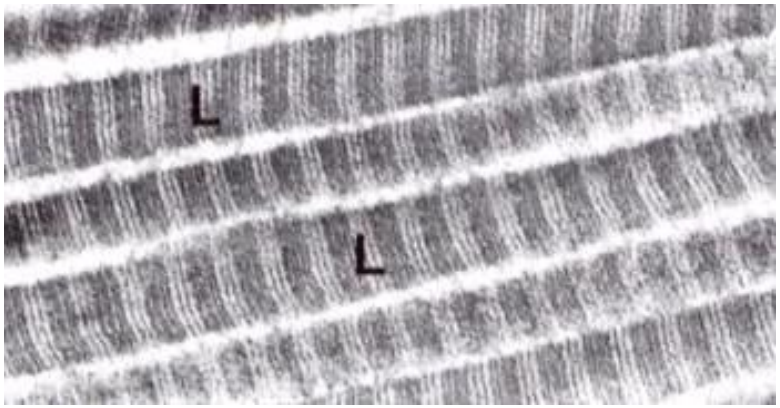
- Tubular crystals (AchR, Ca^{++}-ATPase)
  - Crystallization, No diffraction
  - Isotropic data, 3D distortion corrections, phase restrictions
  - ~5 angstroms

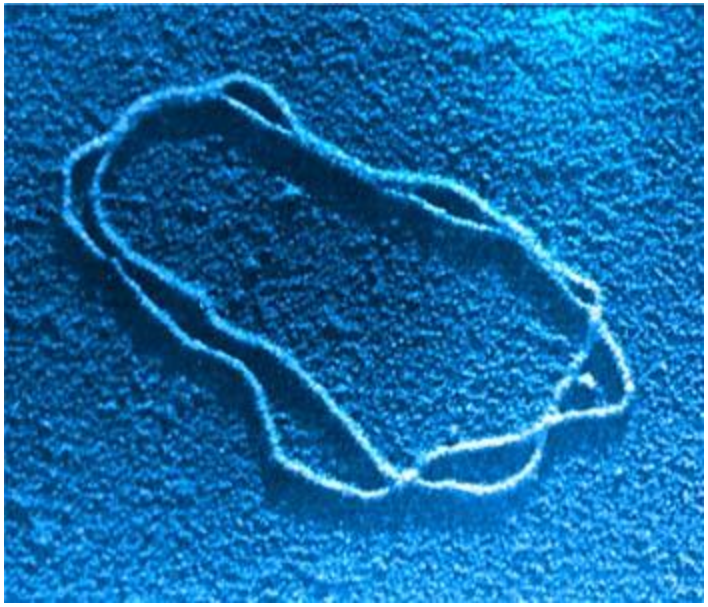
Single particles

- Applicable to any protein or protein complex > 50kD
- Most common sample
- Number of software suites available
- Resolution ~9A (<7 with symmetry)



Fibers and filaments

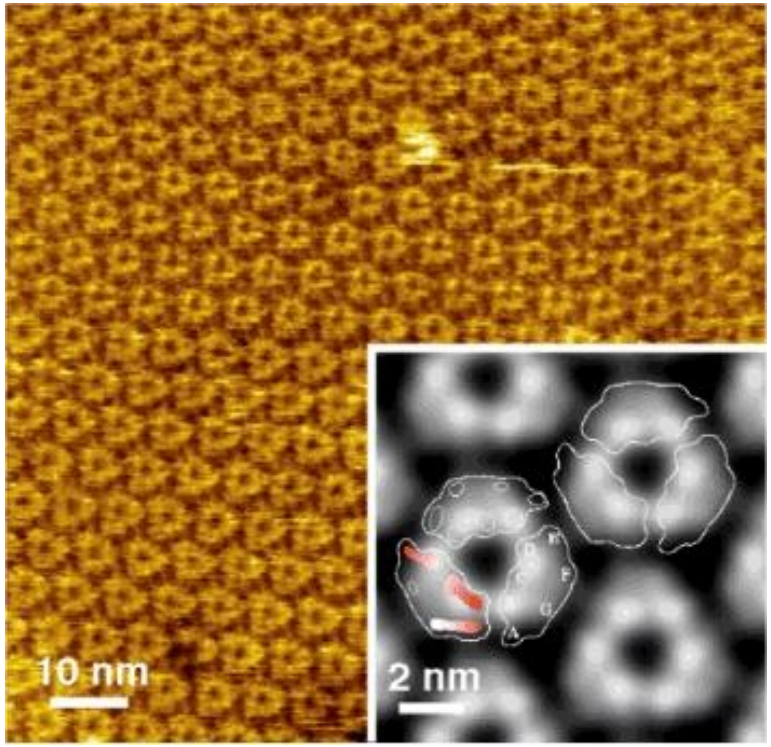


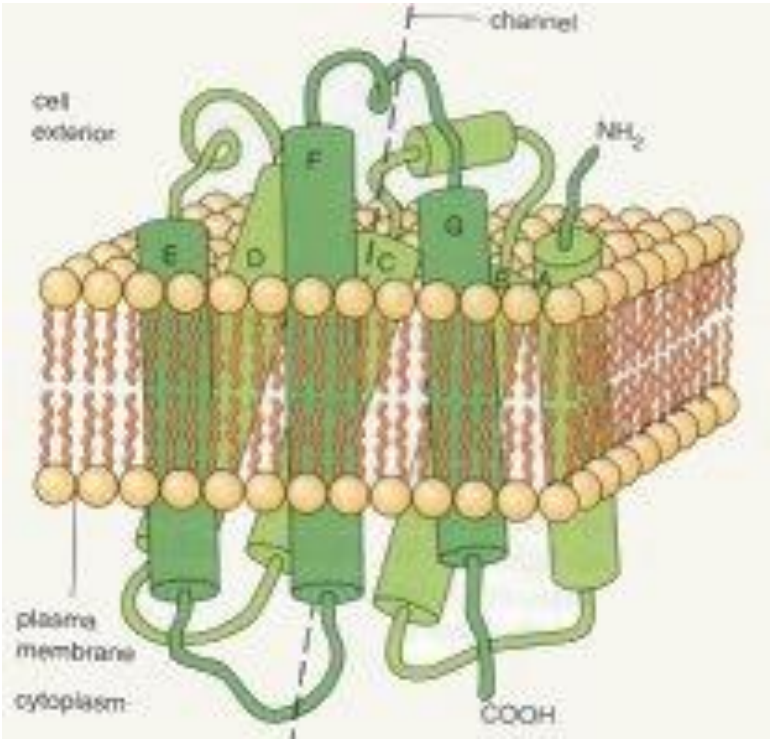


DNA, collagen, etc

2D Xtals

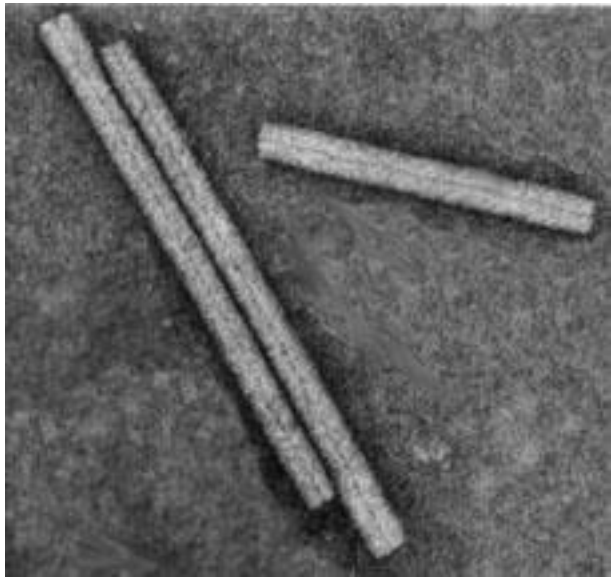
Henderson and Unwin



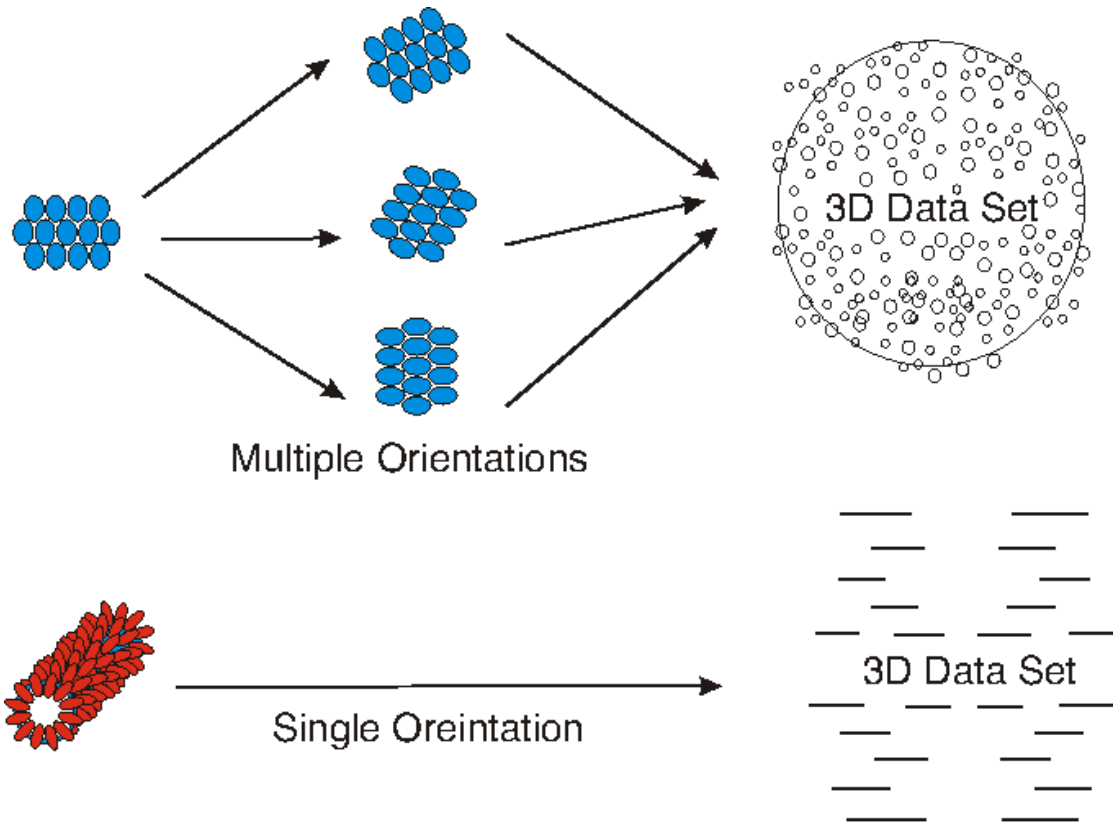


Tubular crystals

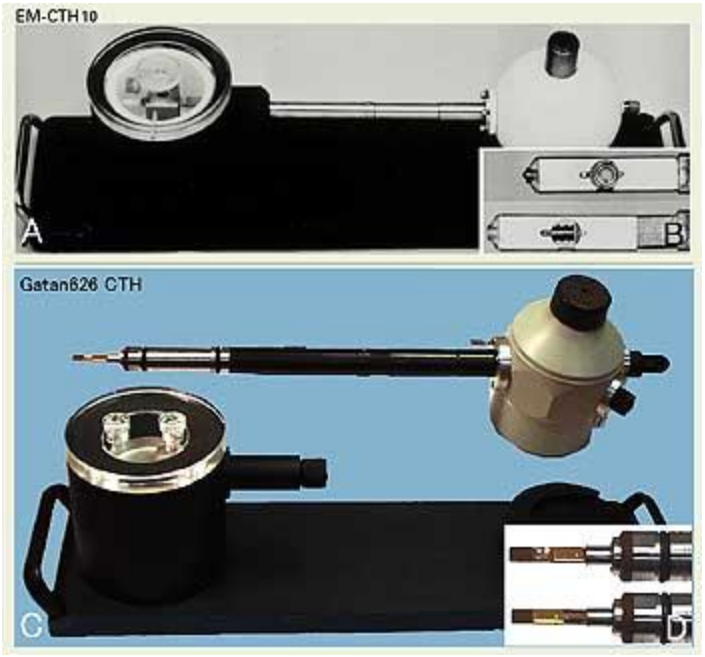
Rolled 2D xtal



Tubular xtal versus 2D or 3D xtal



Data collection



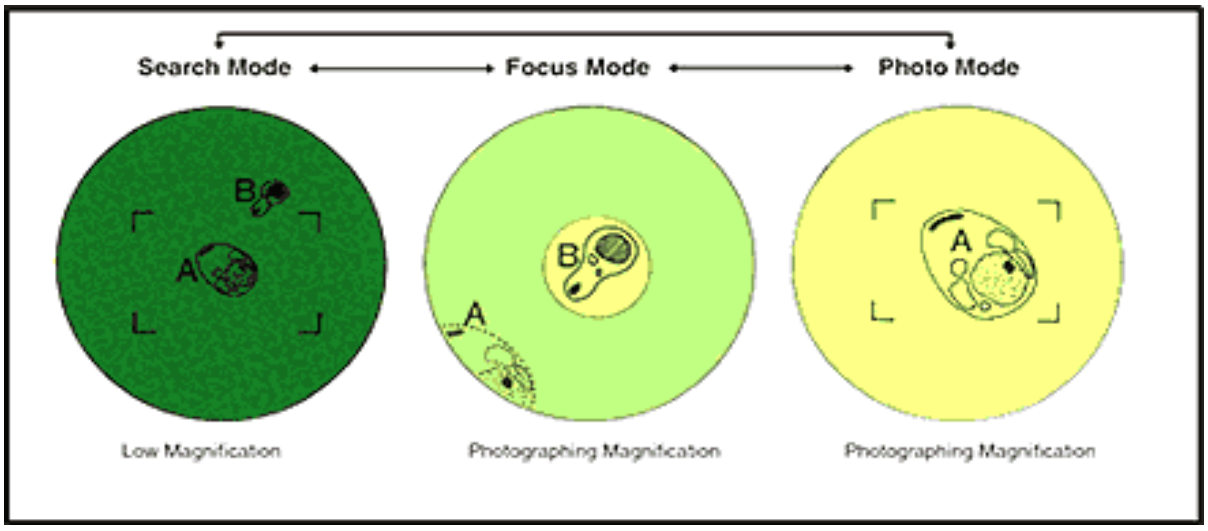
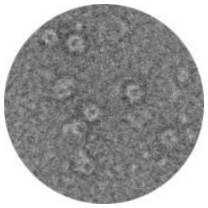


Image recording

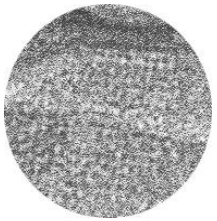
- Film
  - High density content (~20kx16k pixels)
  - Slow (development time, drying)
  - Requires digitization (scanning takes hours)
- CCD
  - Low density content (4kx4k pixels)
  - Fast (ms to sec)
  - Direct digital

Processing data

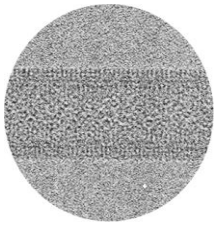


- Single Particles (Proteins, Ribosome)
  - Pick particles
  - Align
  - Classify, average and reconstruction

- Fibers and filaments (tubulin, collagen)
  - Pick segments determine symmetry
  - Align/rotate
  - Average



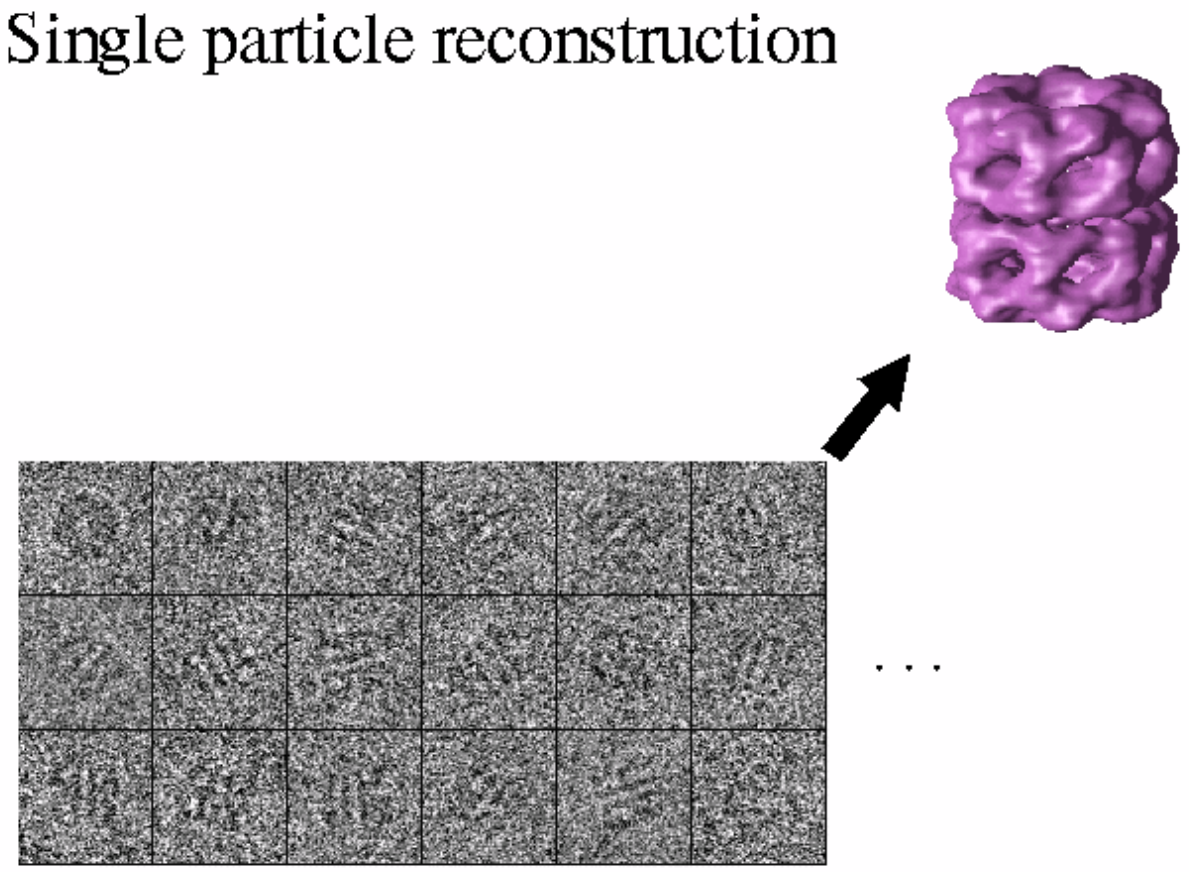
- 2D crystals (BR, AQP, LHCII)
  - Process images to achieve phases
  - Process diffraction data for amplitudes
  - Combine and refine as in X-ray



- Tubular crystals (AchR, Ca^{++}-ATPase)
  - Determine tube symmetry
  - Pick segments and distortion correction
  - Average and sum segments

Data processing 1: single particle

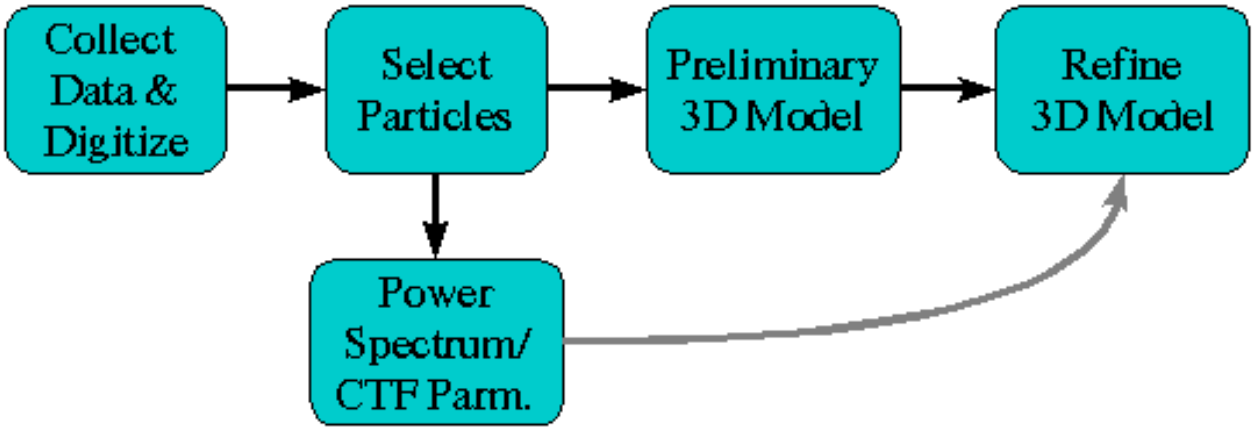
Mostly swiped from Steve Ludtke's web site **http://ncmi.bcm.tmc.edu/~stevel/EMAN/doc/**



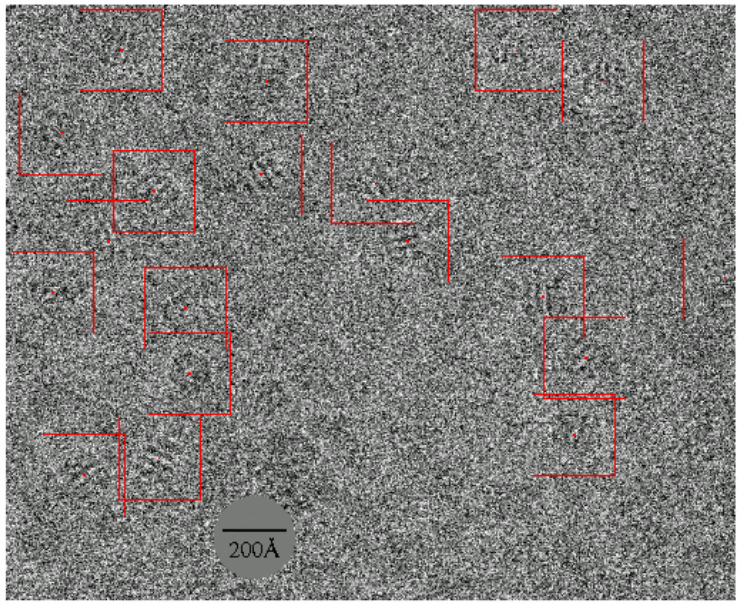
Software

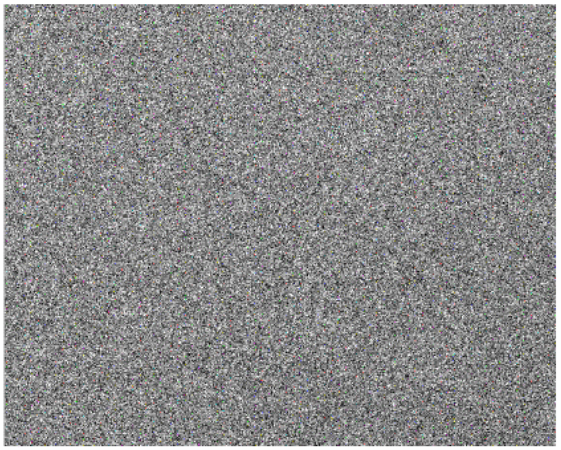
- Spider
  - http://www.wadsworth.org/spider_doc/spider/docs/
  - $1500 + $300/yr for updates, with source
- Imagic
  - http://www.imagescience.de/imagic/welcome.htm
  - Commercial package, ~ $6000/yr
- EMAN
  - http://ncmi.bcm.tmc.edu/~ stevel/EMAN/doc
  - Free, complete with C++ source

The Reconstruction Process

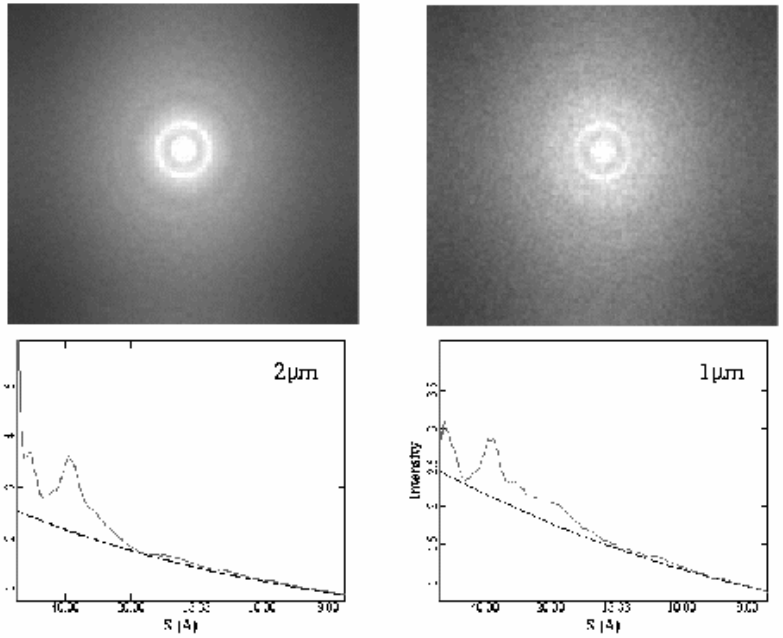


Pick particles (manual or semiauto)

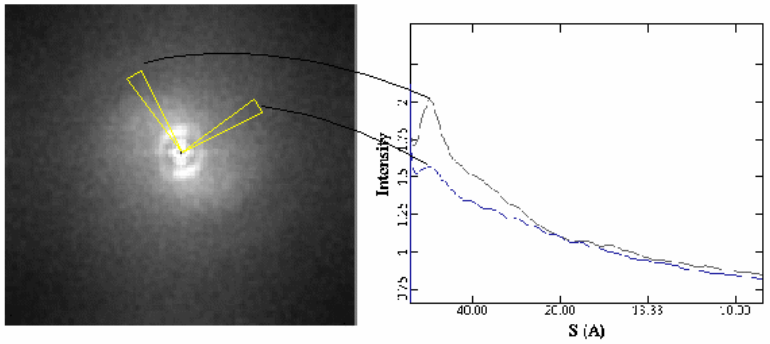




Evaluate Particles



~~Astigmatism &/or Drift~~



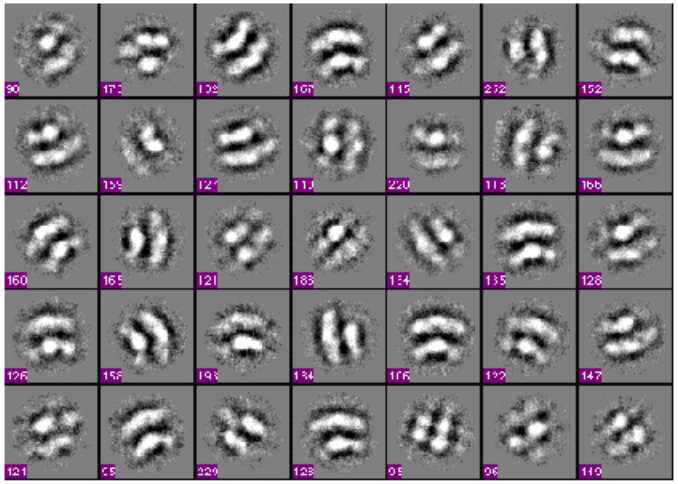
Looking for astigmatism, drift, charging etc….

Now on to the first model

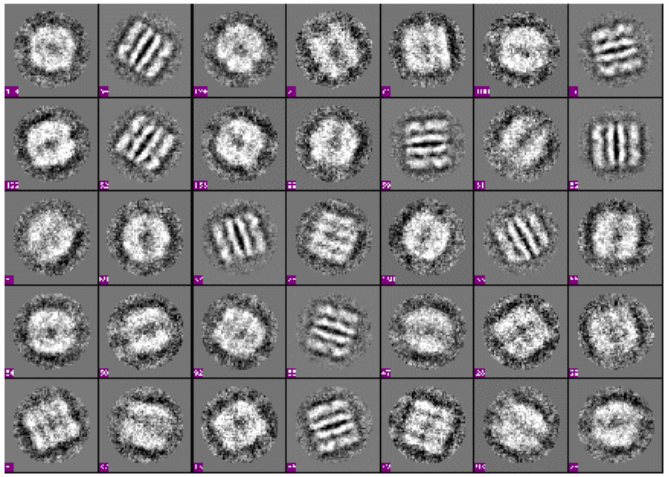
- First rule of thumb…be cautious…
- How to classify particles
  - Reference free classification and alignment
  - MSA
- Application of symmetry
- Random conical tilt

Reference free classification MSA

FAS
reference free class averages

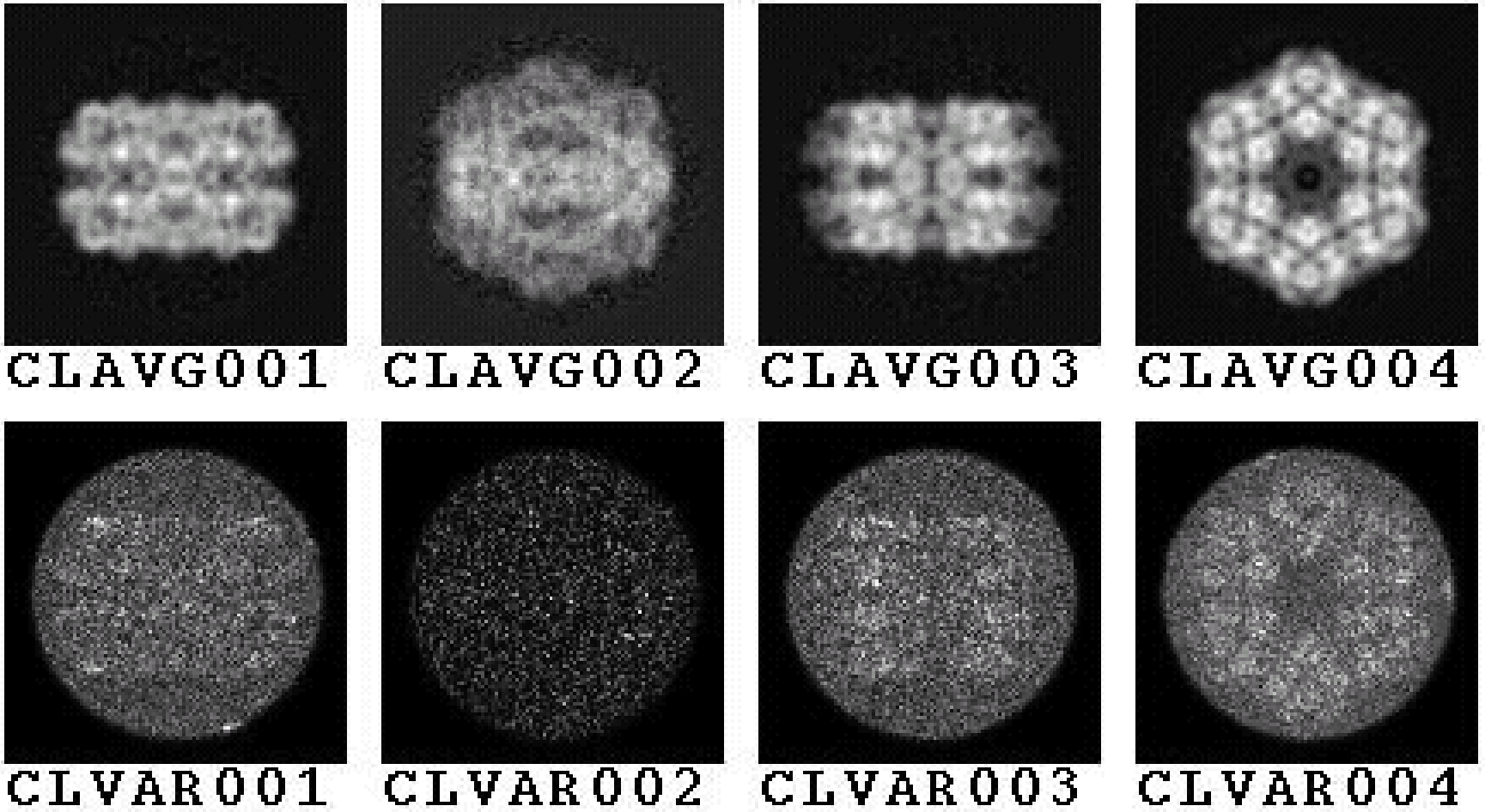


GroEL
reference free class averages



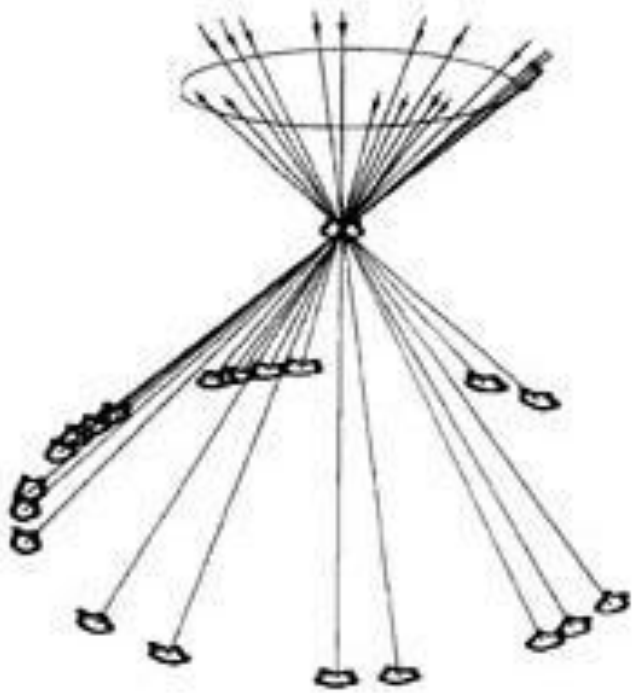
Can we tell the symmetry a priori???

MSA…….variance….$(SD)^2$

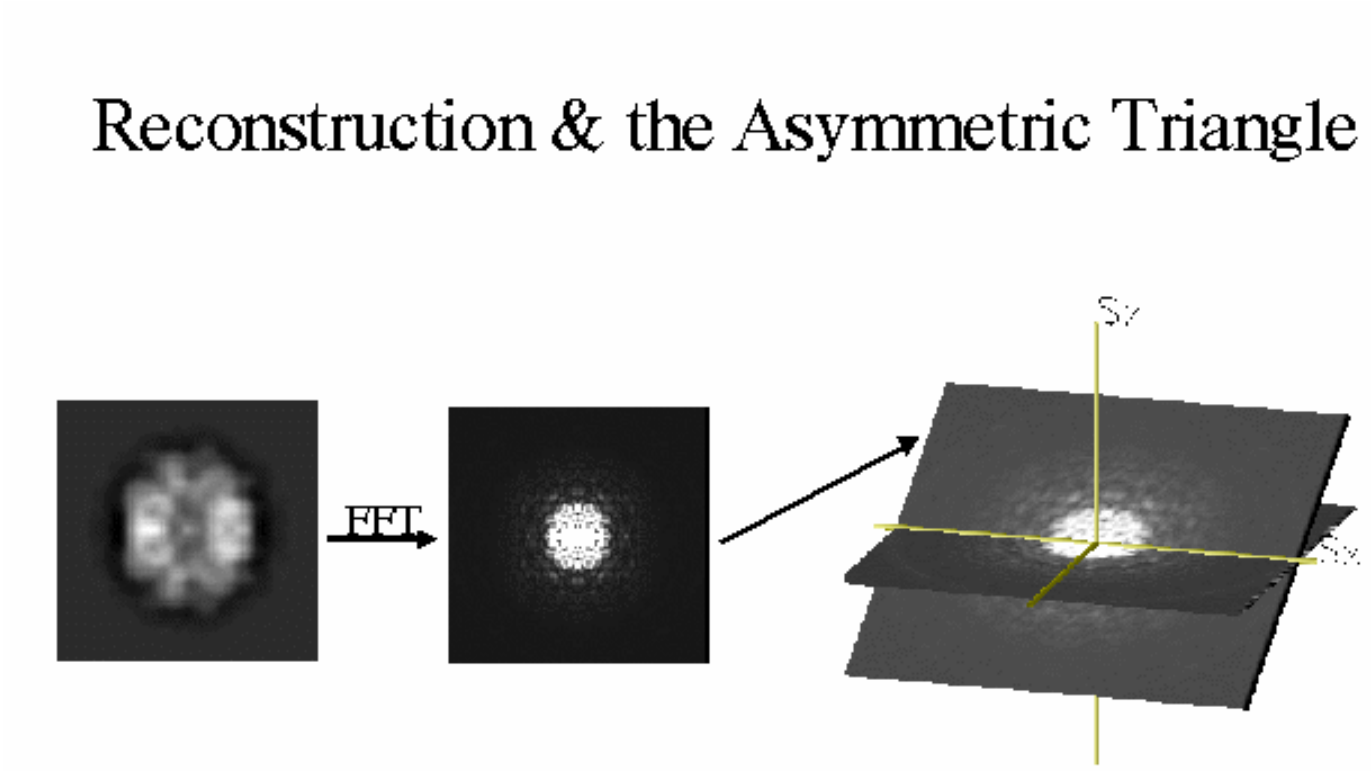


Random conical tilt

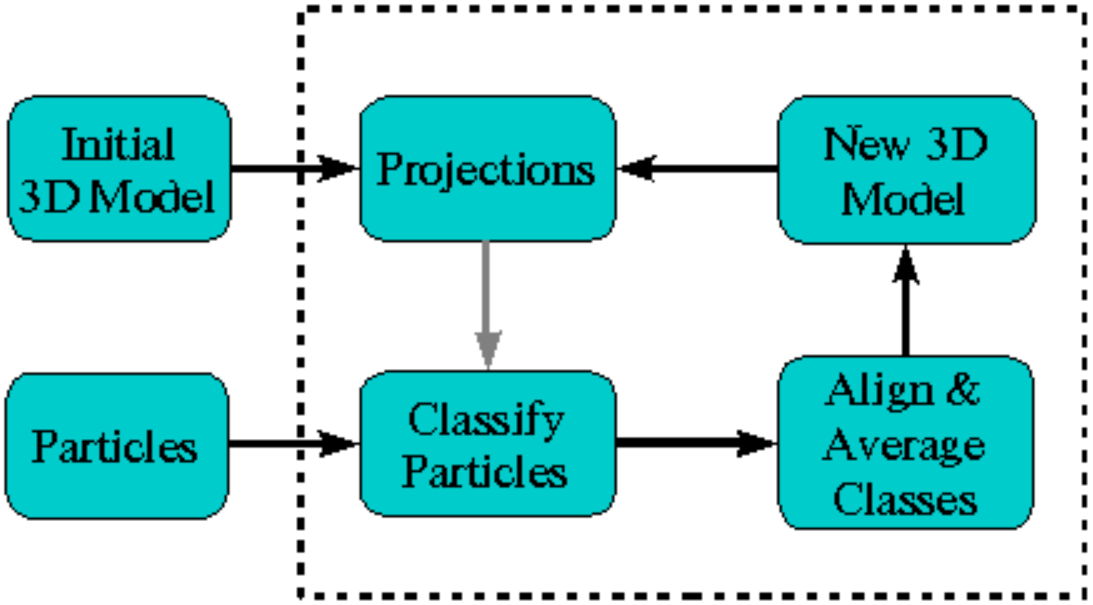
- Image pairs taken of the same sample with an angular tilt applied between them
- Determine particle pairs and construct reference model



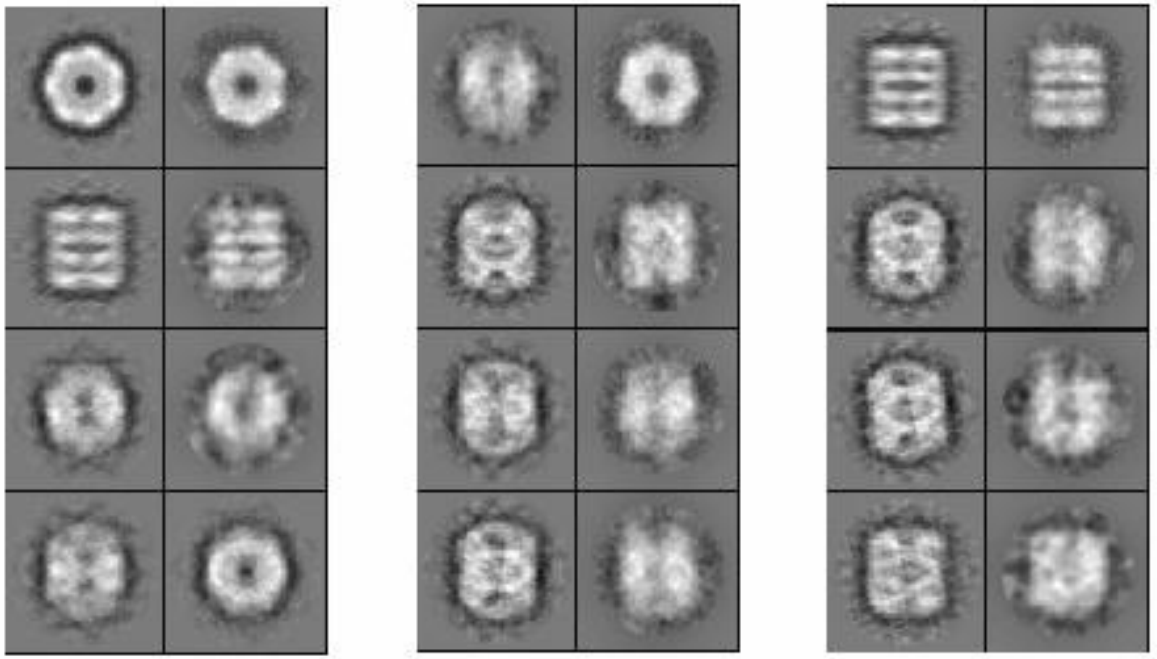
Use of common lines to align different orientations



Refinement - EMAN

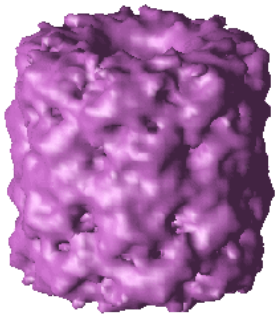


Class Averages

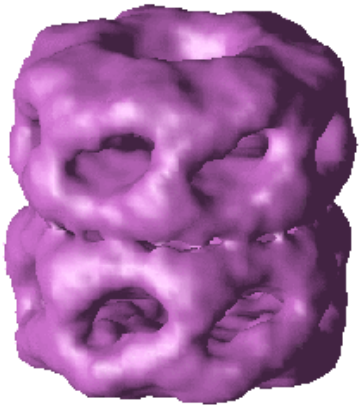


Multiple rounds of refinement…

Iteration 1

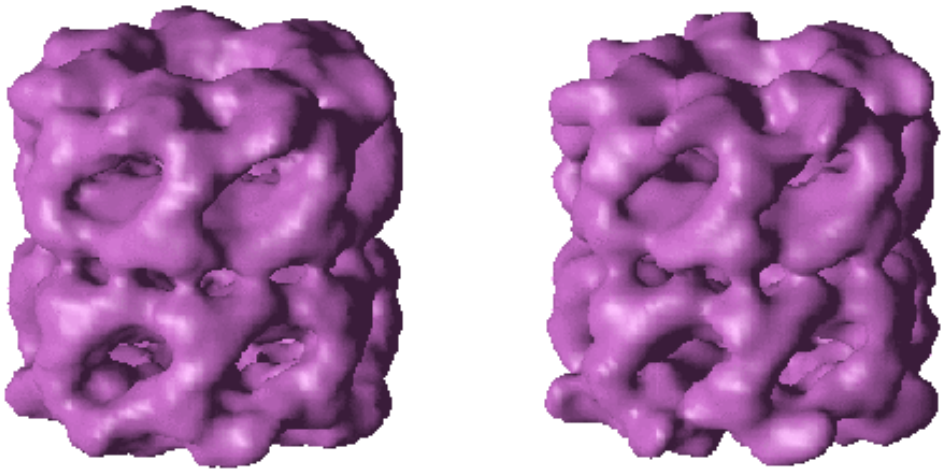


Iteration 3

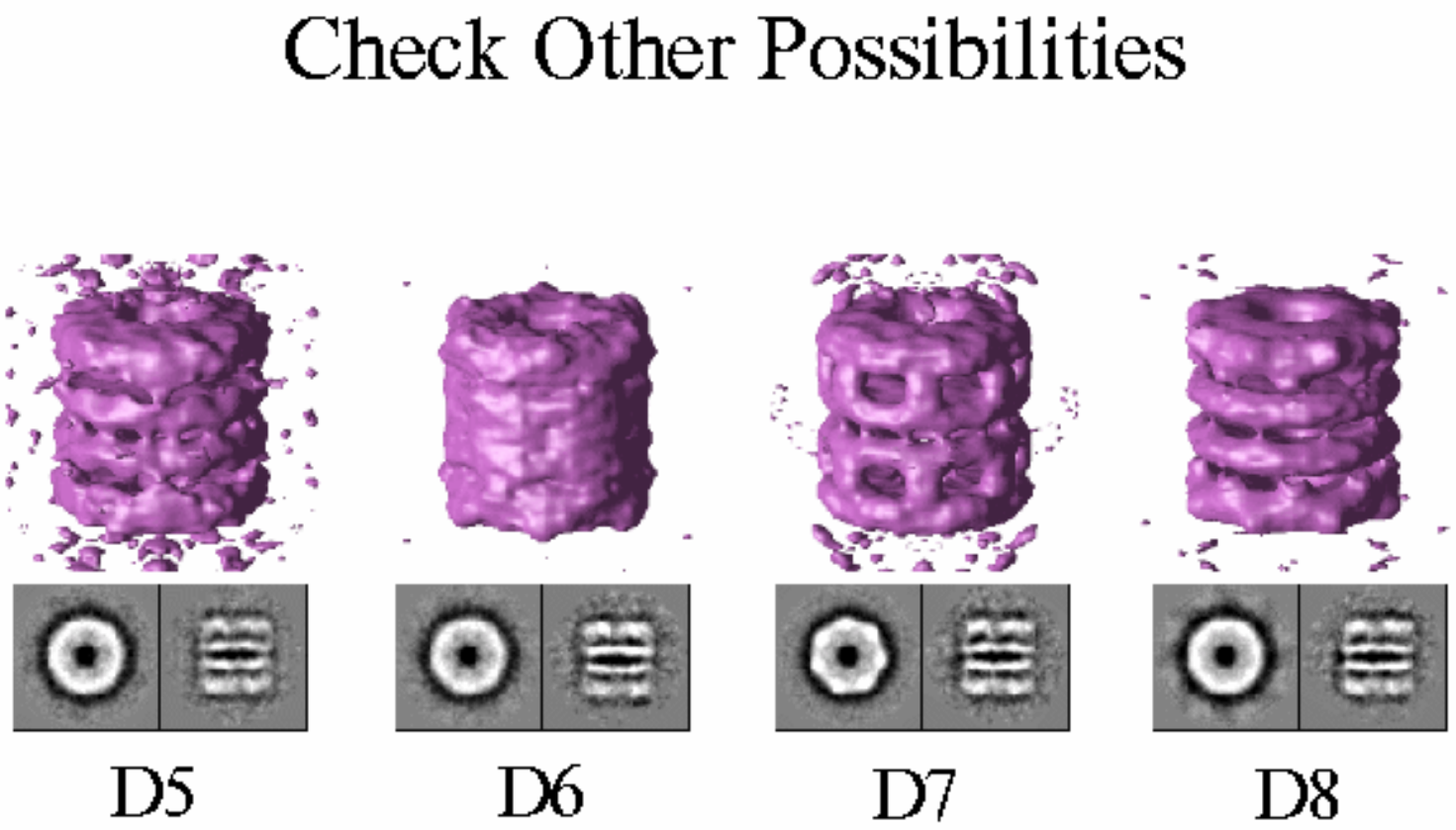


Convergence when no improvement in the alignment statistics

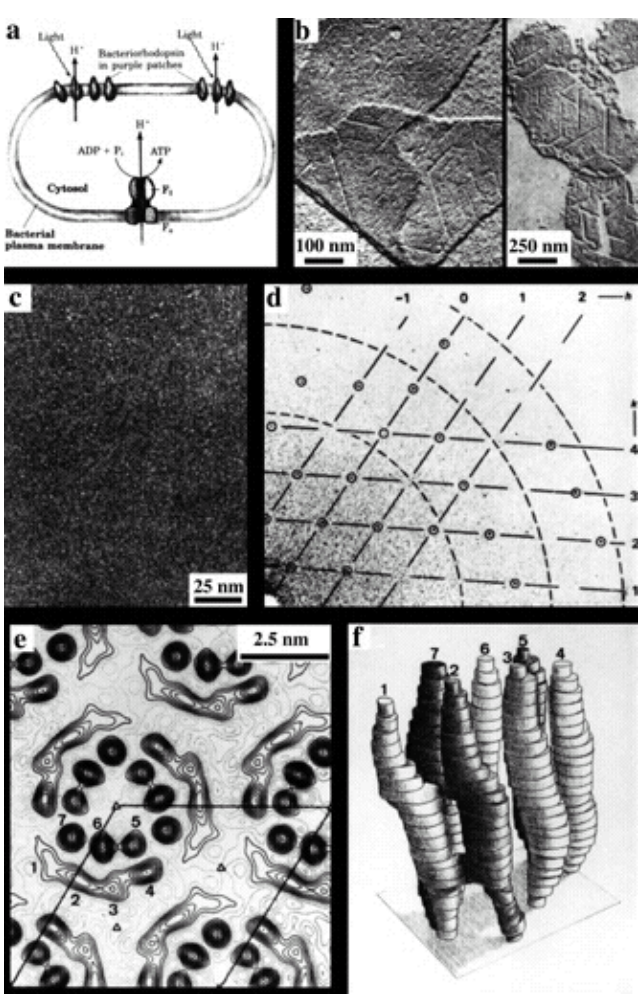
Final vs. x- ray

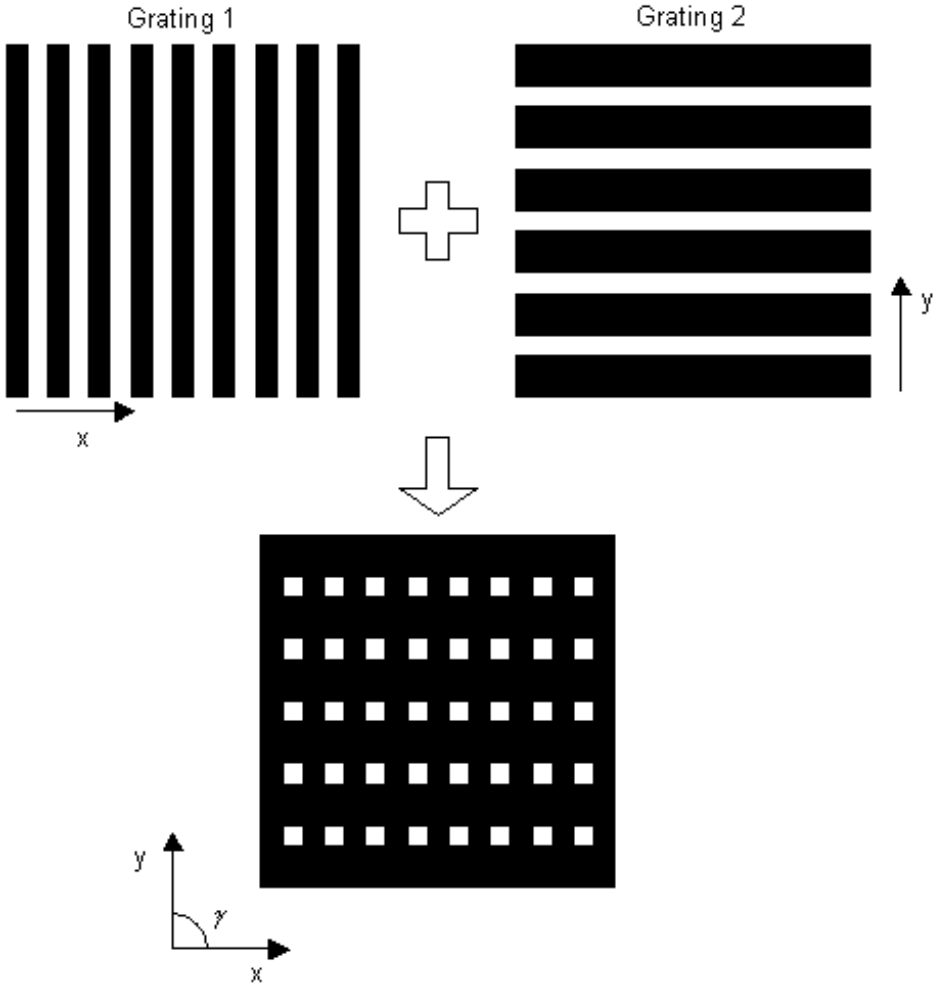


Try different symmetries



Data processing 2: 2D xtal





Why lattice lines?

Z dimension has an effective real space D of infinity
Hence in reciprocal space the lattice spacing is 0

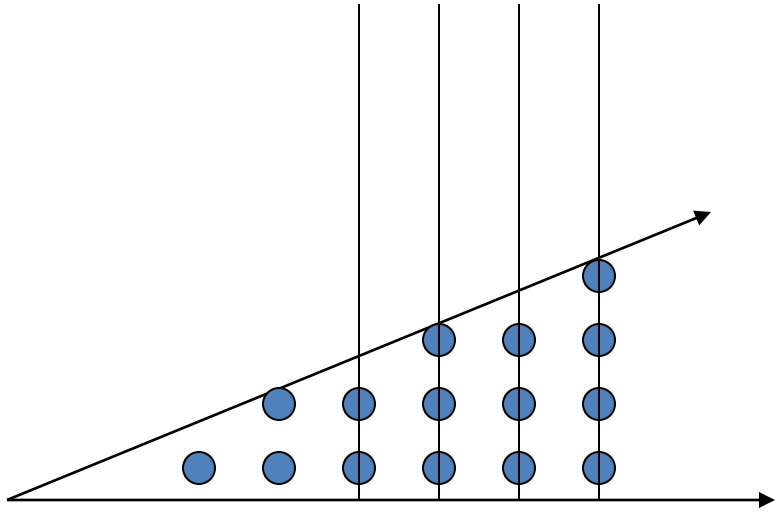
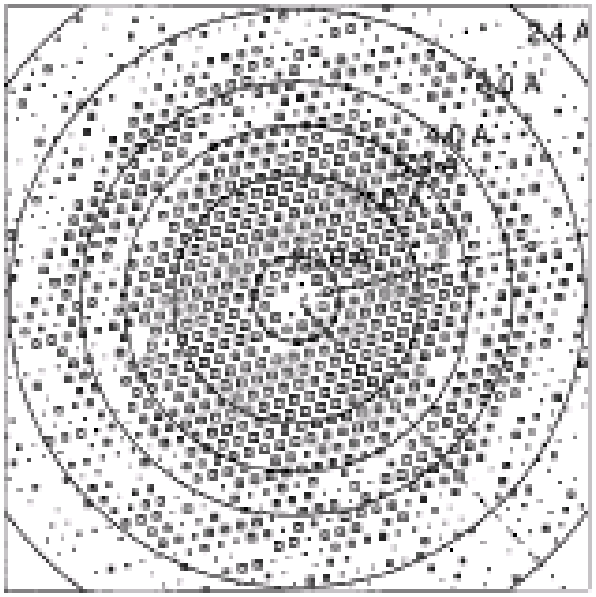
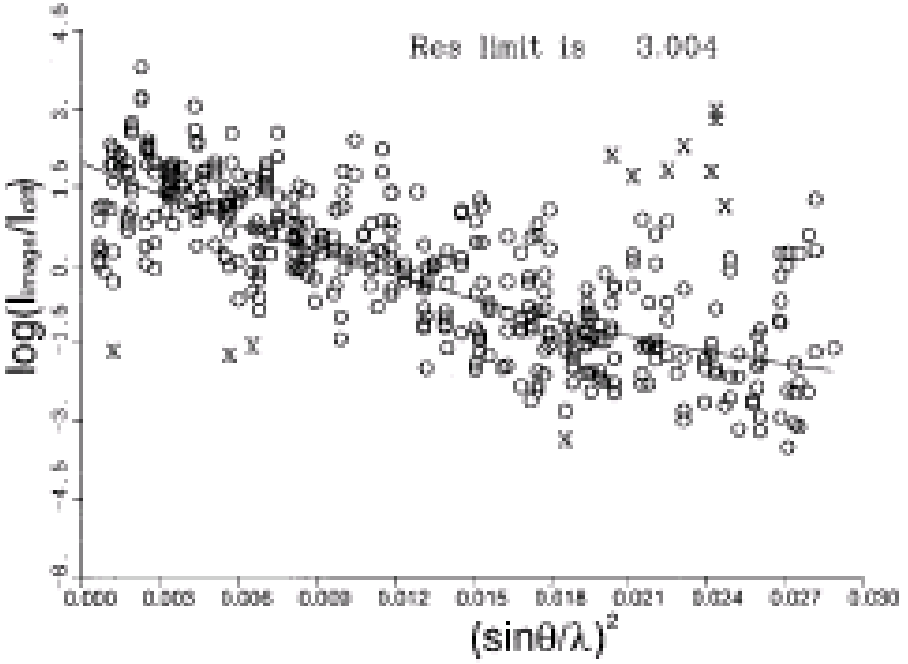


Table 1. Electron crystallographic table

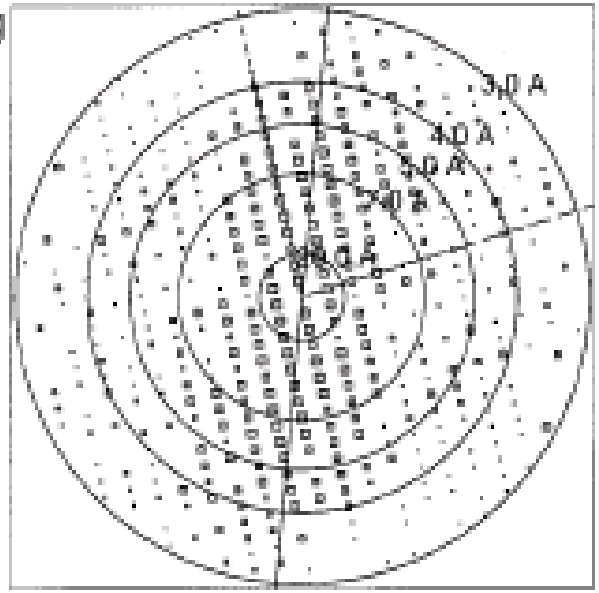
A. *Crystallographic parameters*	
Space group (layer group)	*P*3 (p3)
Lattice constants	$a = b = 62.45$ Å, $\gamma = 120^\circ$
	$c = 100$ Å (assumed in refinement)
Thickness	70 Å (used in LATLINE)
B. *Electron diffraction (amplitude information)*	
No. of diffraction patterns	339
Resolution limit used (Å)	3.0
Maximum tilt angle (deg.)	70.6
No. of observed reflections	110,812
Friedel *R*-factor (%)	17.6
C. *Electron microscopy (phase information)*	
No. of images	181
Resolution limit used (Å)	3.0
Maximum tilt angle (deg.)	61.2
No. of observed reflections	25,225
D. *Merged data*	
Resolution (Å)	3.0
No. of unique reflections	6892 (with amplitudes and phases)
Merging *R*-factor (%)	31.3
Phase residual (deg.)	46.8
Completeness (%)	78.4
E. *Refinement*	
No. of degrees of freedom	6672 (without hydrogen atoms)
R-factor (%)	23.7
Free *R*-factor (%)	33.0
Phase residual (deg.)	54.4
Free phase residual (deg.)	63.3

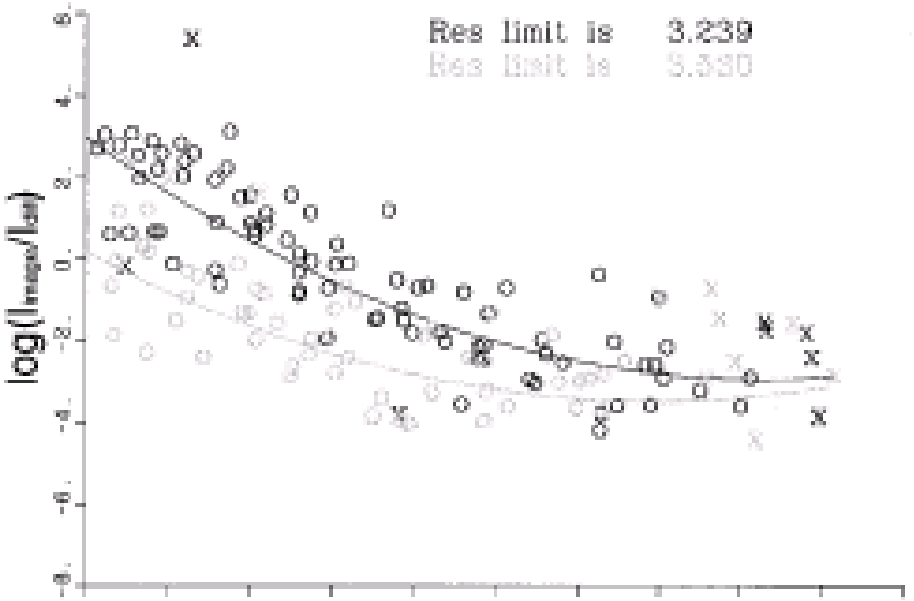
0 deg


2.4 A
3.0 A
4.0 A
6.0 A

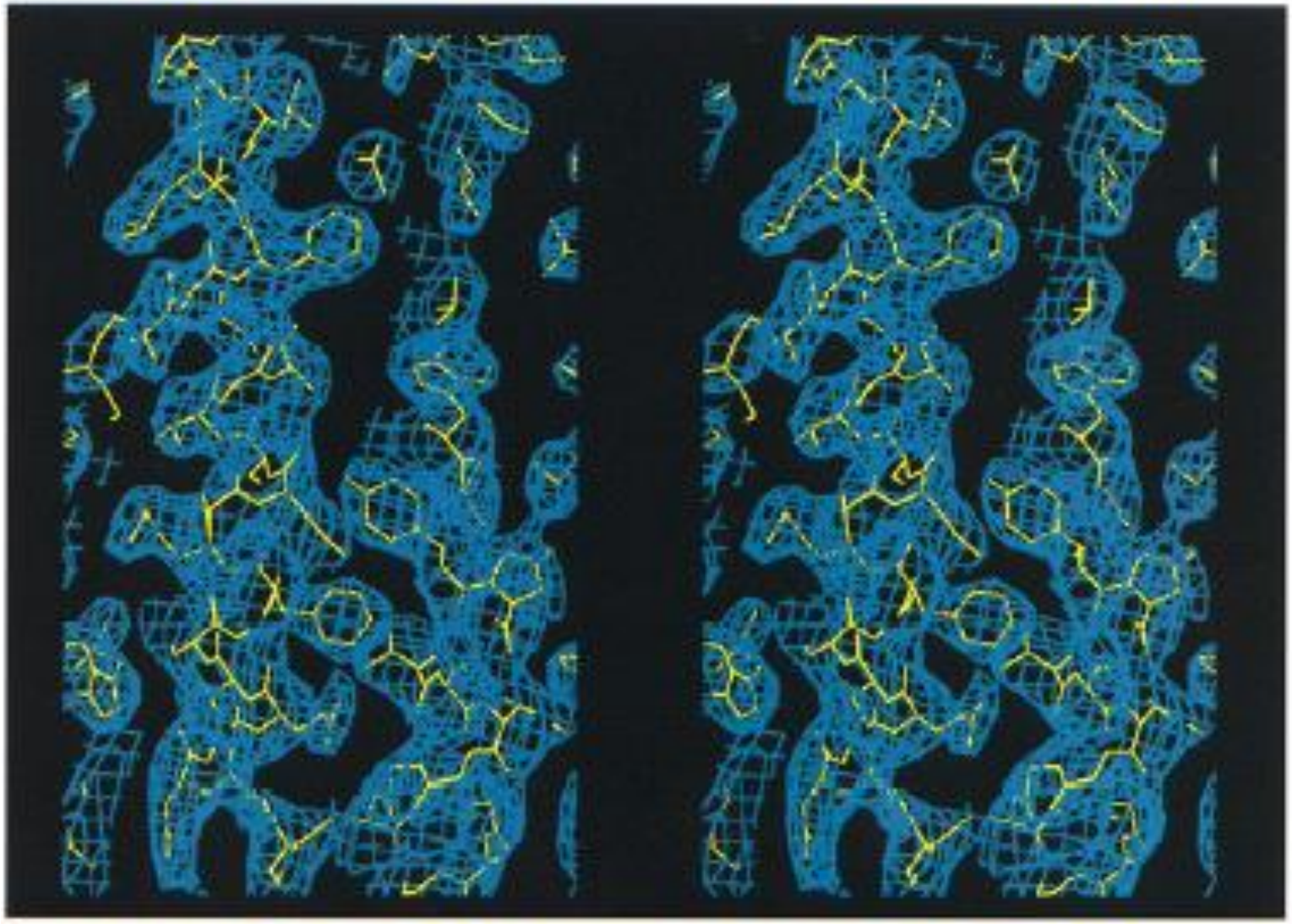

Res limit is 3.004
log(I_meas/I_calc)
(sinθ/λ)²

60 deg


3.0 A
4.0 A
6.0 A


Res limit is 3.239
Res limit is 3.330
log(I_meas/I_def)

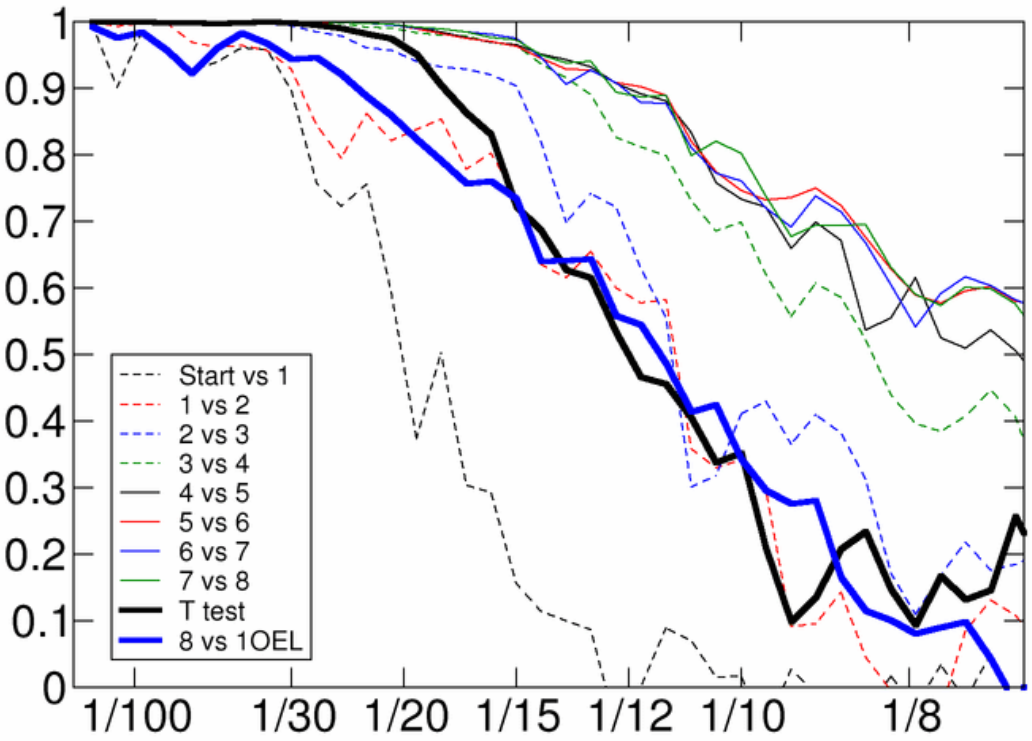
A well refined EM map



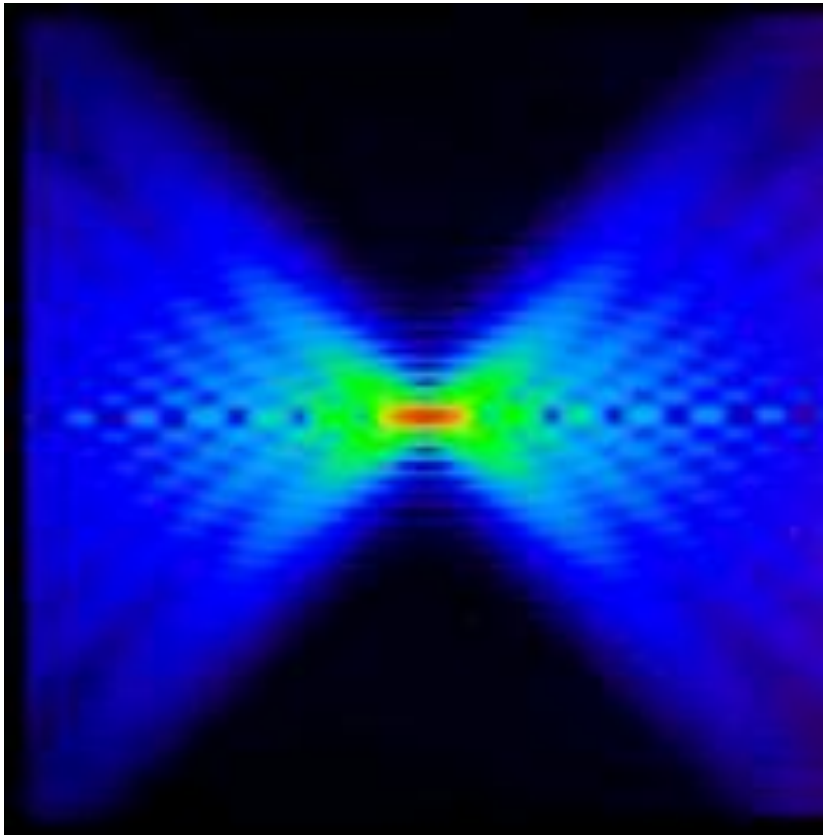
Resolution and Resolvability

- Single particles, filaments, tubes
  - FSC
    - Which criteria to use (0.5 or 3 sigma)
- 2D xtals diffraction (like X-ray)
  - But anisotropy or point spread function

FSC


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1/100
1/30
1/20
1/15
1/12
1/10
1/8
Start vs 1
1 vs 2
2 vs 3
3 vs 4
4 vs 5
5 vs 6
6 vs 7
7 vs 8
T test
8 vs 1OEL

Point spread function



Resolution vs Resolvability

- Resolution is a calculated value
  - FSC or measured amplitudes above a certain sigma value.
- Resolvability is a perceived value
  - What can I see in the map
    - Is a 4 angstroms map really 4 angstroms if one cannot discern beta sheet structure?