

X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography (after NMR)

- Crystal Growth – Materials / Methods
- Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit
- X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron
- Theory of Diffraction – Bragg's Law / Reciprocal Space
- Data Collection – Methods / Detectors / Structure Factors
- Structure Solution – Phase Problem: MIR / MR / MAD
- Refinements and Models
- Analysis and presentation of results

PDB Holdings List: 5-Apr-2005

		Molecule Type					
		Protein, Peptides, and Viruses	Protein/Nucleic Acid Complexes	Nucleic Acids	Carbohydrates	Total	
Exp.	X-ray Diffraction and other	23542	1147	770	11	25370	
Techn.	NMR	3721	111	646	2	4680	
	Total	27663	1258	1427	13	30361	

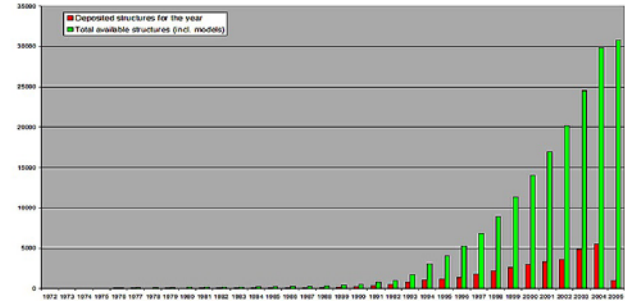
Please note that theoretical models have been removed, effective July 02, 2003, as per [CDS policy](#)

35,917 4/6/06

42,627 4/4/07

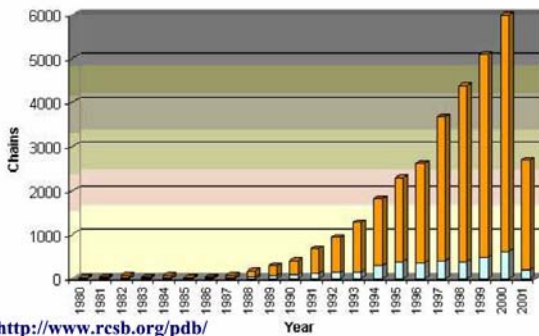
16253 Structure Factor Files
2424 NMR Restraint Files

PDB Content Growth



Protein Data Bank - 17,679 Structures - March 2002

X-ray - 14595
NMR - 2706 "Old" Folds
Theory - 378 "New" Folds



<http://www.rcsb.org/pdb/>

Analyze – structure (Ramachandran Plot) and biochemistry

Publish in leading biochemical or structural biology journal

Contribute results (coordinates, etc.) to PDB

Data Mining

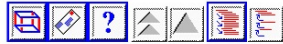
Visualization programs (Cn3D / RasMol / SwissPDBV / etc)

SCOP – Structural Classification of Proteins

CATH – Classification / Arch / Topology

SCOP Structural Classification of Proteins

Structural Classification of Proteins



Root: scop

Classes:

1. All alpha proteins (151) [\[?\]](#) [\[A\]](#) [\[C\]](#)
2. All beta proteins (111) [\[?\]](#) [\[A\]](#) [\[C\]](#)
3. Alpha and beta proteins (α/β) (117) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Mainly parallel beta sheets (beta-alpha-beta units)
4. Alpha and beta proteins ($\alpha+\beta$) (212) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Mainly antiparallel beta sheets (segregated alpha and beta regions)
5. Multi-domain proteins (alpha and beta) (39) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Folds consisting of two or more domains belonging to different classes
6. Membranes and cell surface proteins and peptides (12) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Does not include proteins in the immune system
7. Small proteins (29) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Usually dominated by metal ligand, heme, and/or disulfide bridges
8. Coiled coil proteins (5) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Not a true class
9. Low resolution protein structures (17) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Not a true class
10. Peptides (9) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Peptides and fragments. Not a true class
11. Designed proteins (36) [\[?\]](#) [\[A\]](#) [\[C\]](#)
Experimental structures of proteins with essentially non-natural sequences. Not a true class

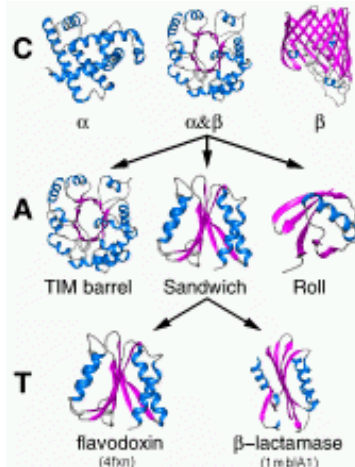
CATH - Protein Structure Classification

CATH is a novel hierarchical classification of protein domain structures, which clusters proteins at four major levels: Class (C), Architecture (A), Topology (T), and Homologous (H) Superfamily

Class, derived from secondary structure content, is assigned for more than 90% of protein structures automatically.

Architecture, which describes the gross orientation of secondary structures, independent of connectivities, is currently assigned manually. The **topology** level clusters structures according to their topological connections and numbers of secondary structures. The **homologous superfamilies** cluster proteins with highly similar structures and functions. The assignments of structures to topology families and homologous superfamilies are made by sequence and structure comparisons.

CATH



X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

1. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

2. Image Formation

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

3. X-Ray Crystallography

a) Crystal Growth – Materials / Methods

b) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

c) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

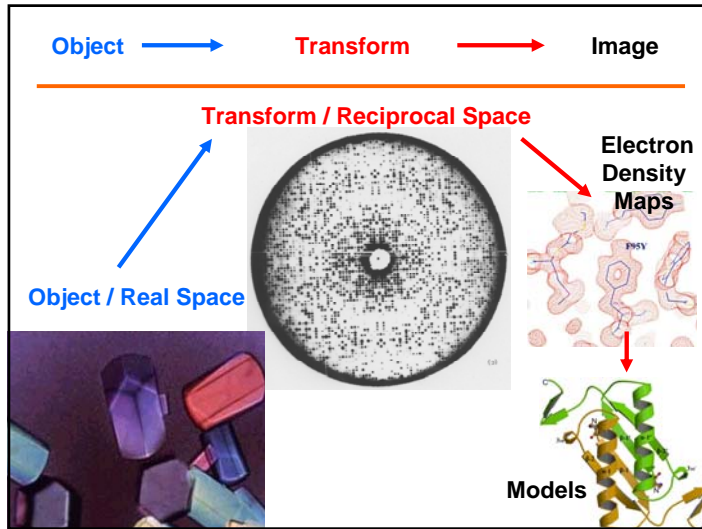
d) Theory of Diffraction – Bragg's Law / Reciprocal Space

e) Data Collection – Methods / Detectors / Structure Factors

f) Structure Solution – Phase Problem: MIR / MR / MAD

h) Refinements and Models

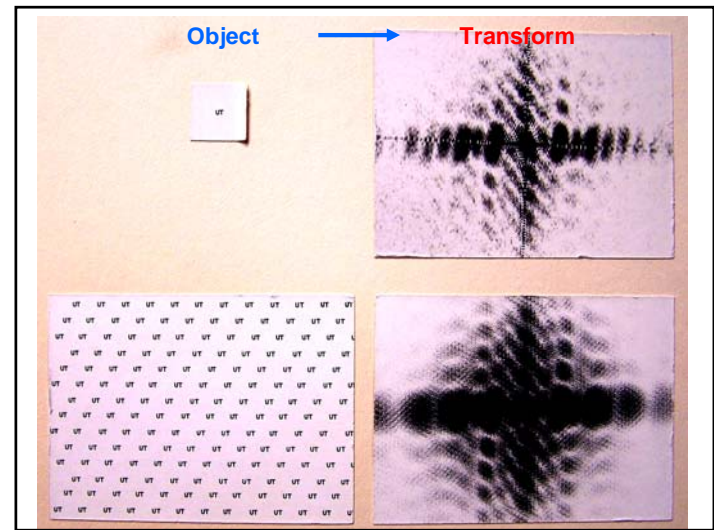
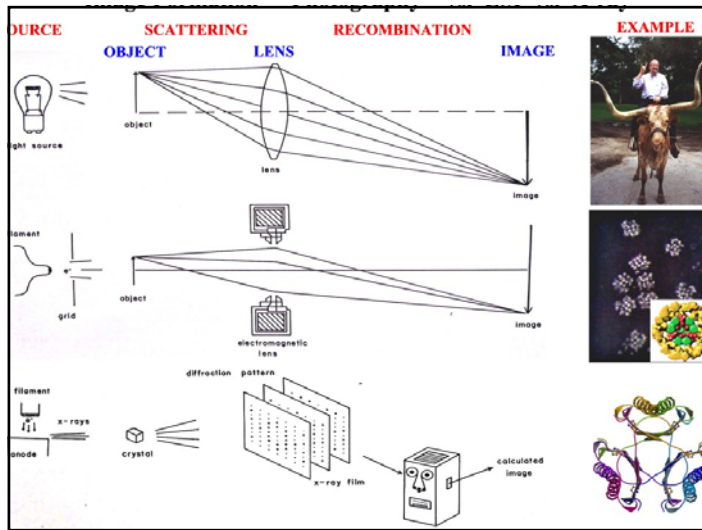
i) Analysis and presentation of results

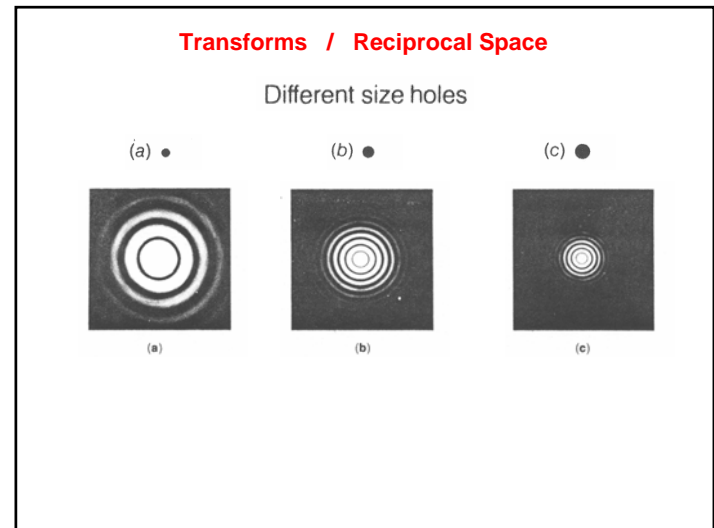
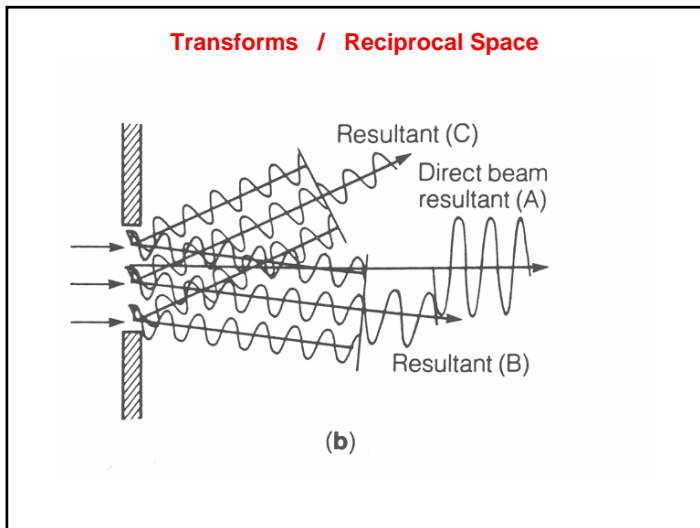
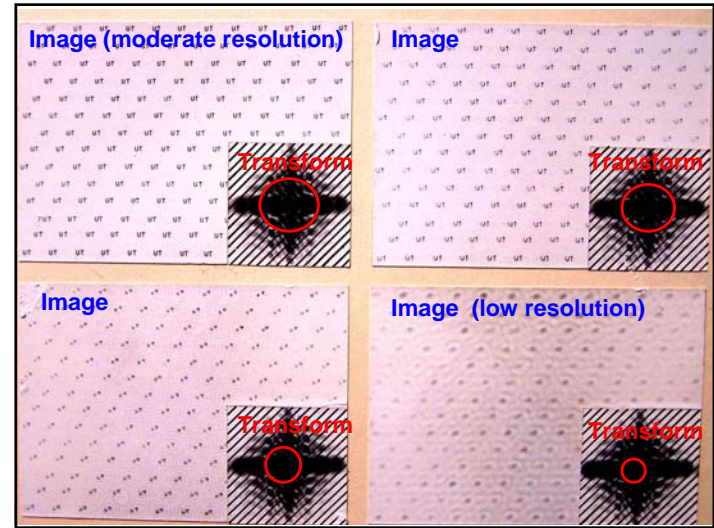
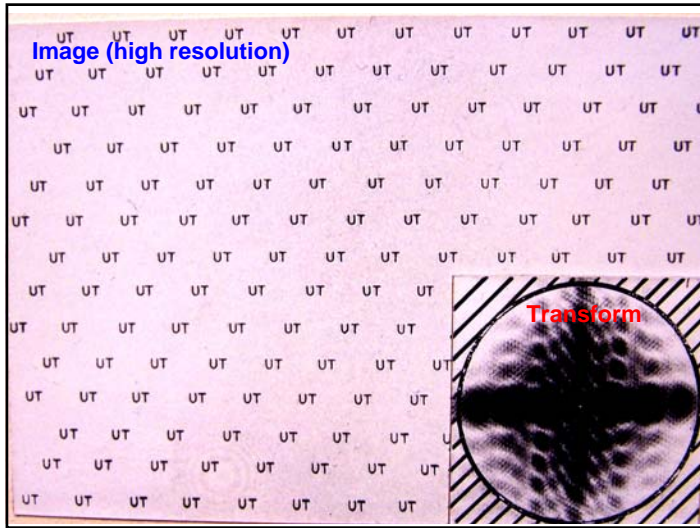


"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

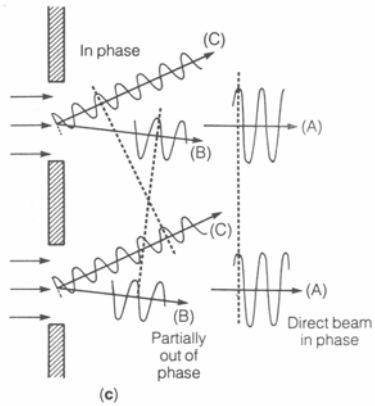
- Light Photography
 $\lambda \sim 400 - 700 \text{ nm}$
- Electron Microscopy
 $\lambda \sim 0.001 - 0.1 \text{ nm}$
- X-Ray or NMR
 $\lambda \sim 0.1 \text{ nm}$

Image Formation
Abbe (~1873):
Limit Res. $\sim \lambda/2$

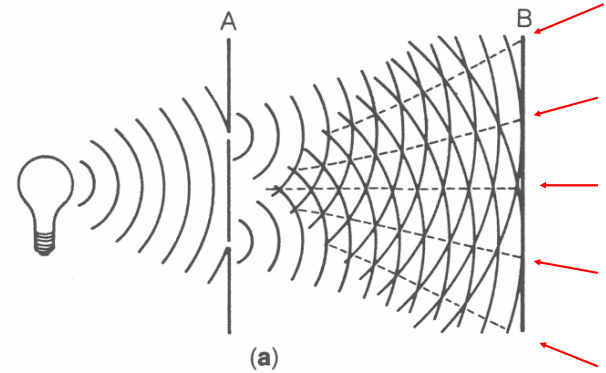




Transforms / Reciprocal Space

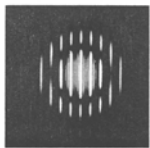


Transforms / Reciprocal Space

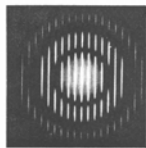


Transforms / Reciprocal Space

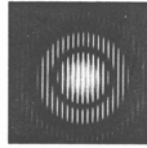
Five horizontal holes
with various spacings



(j)



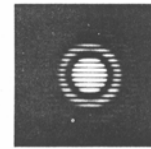
(k)



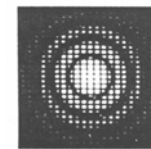
(l)

Transforms / Reciprocal Space

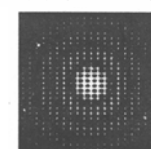
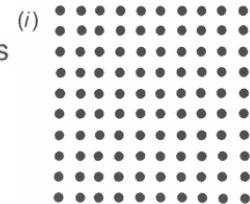
Vertical holes and nets of holes



(g)



(h)



(i)


Kevin Cowtan's Picture Book of Fourier Transforms - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: <http://www.yybl.york.ac.uk/~cowtan/fourier/fourier.html>

Kevin Cowtan's Book of Fourier



This is a book of pictorial 2-d Fourier Transforms. These are particularly relevant to my own field of X-ray crystallography, but should be of interest to anyone involved in signal processing or frequency domain calculations.

Contents: <http://www.yybl.york.ac.uk/~cowtan/fourier/fourier.html>

- [Introduction](#)
- [Book of Crystallography](#)
- [Duck Tales](#) and missing data
- [A Little Animal Mating](#) and cross phasing
- [A Tale of Two Cats](#) and image restoration
- [Animal Liberation](#) and free-sets
- [The Gallery](#) Other interesting pictures

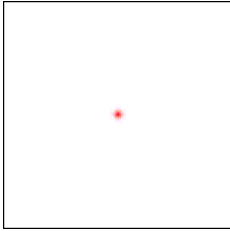
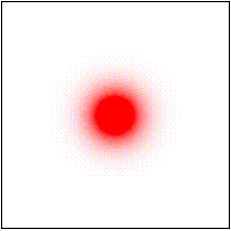
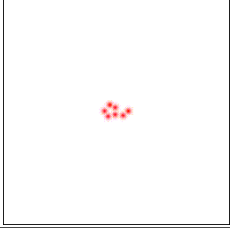
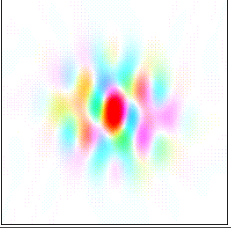
Other topics:

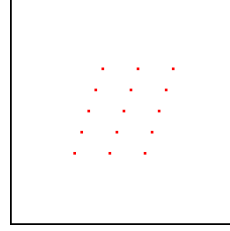
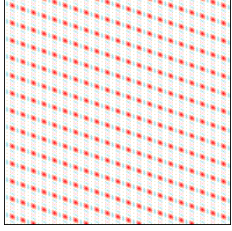
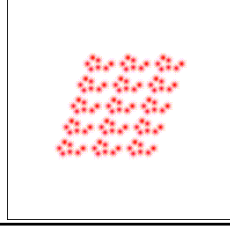
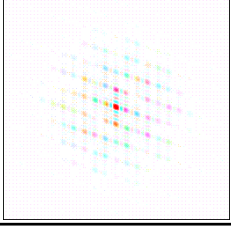
[The Interactive Structure Factor Tutorial](#): Learn about structure factors and maps.

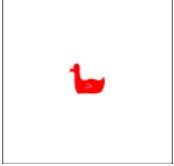
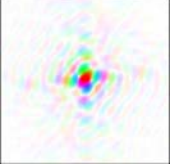

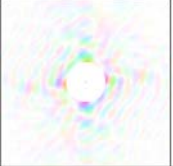
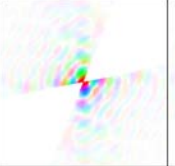

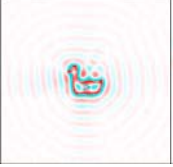
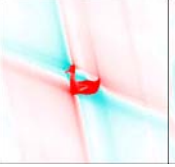
An introduction to crystallographic [Fourier transforms](#). The mathematical link between [Scattering theory](#) and Fourier theory. An explanation of the [contribution theorem](#).

Teaching materials elsewhere

Document: Done

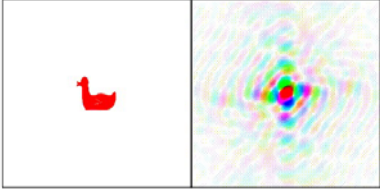
Object / Real Space	Transform / Reciprocal Space
	
	

Object / Real Space	Transform / Reciprocal Space
	
	

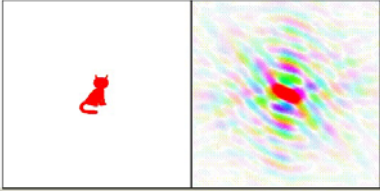
Objects – Transforms and Image Formation	A Duck	Transform of a Duck
		
		
		

Kevin Cowtan's Book of Fourier
<http://www.yesbl.york.ac.uk/~cowtan/fourier/fourier.html>

Here is our old friend, the Fourier Duck, and his Fourier transform:

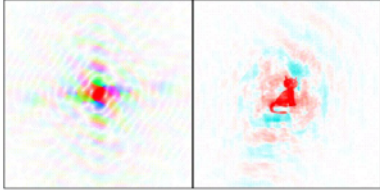


And here is a new friend, the Fourier Cat and his Fourier transform:

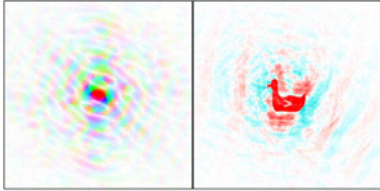


Kevin Cowtan's Book of Fourier
<http://www.yesbl.york.ac.uk/~cowtan/fourier/fourier.html>

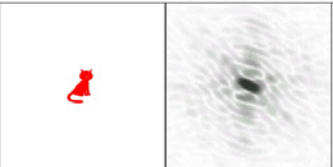
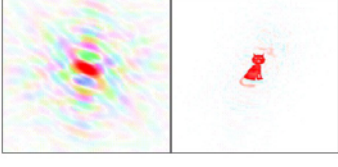
Duck Transform Amplitudes + Cat Phases



Cat Transform Amplitudes + Duck Phases

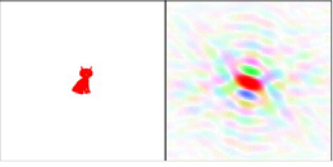
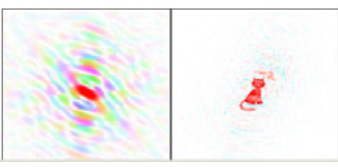


Kevin Cowtan's Book of Fourier
<http://www.yesbl.york.ac.uk/~cowtan/fourier/fourier.html>

a) Cat - Cat Transform (Amplitudes only)
 b) Manx (tailless) Cat - Manx Transform

c) Cat Amplitudes + Manx Phases
 d) $[2x(\text{Cat Amplitudes}) - \text{Manx Amplitudes}] + \text{Manx Phases}$

X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

- Protein Data Bank (PDB)**
 Data mining and Protein Structure Analysis Tools
- Image Formation**
 Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)
- X-Ray Crystallography (after NMR)**
 - Crystal Growth – Materials / Methods
 - Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit
 - X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron
 - Theory of Diffraction – Bragg's Law / Reciprocal Space
 - Data Collection – Methods / Detectors / Structure Factors
 - Structure Solution – Phase Problem: MIR / MR / MAD
 - Refinements and Models
 - Analysis and presentation of results