School on Synchrotron and Free-Electron-Laser Methods for Multidisciplinary Applications

ICTP

7-18 May 2018
Trieste, Italy

Protein X-ray Crystallography : *Basic Aspects*

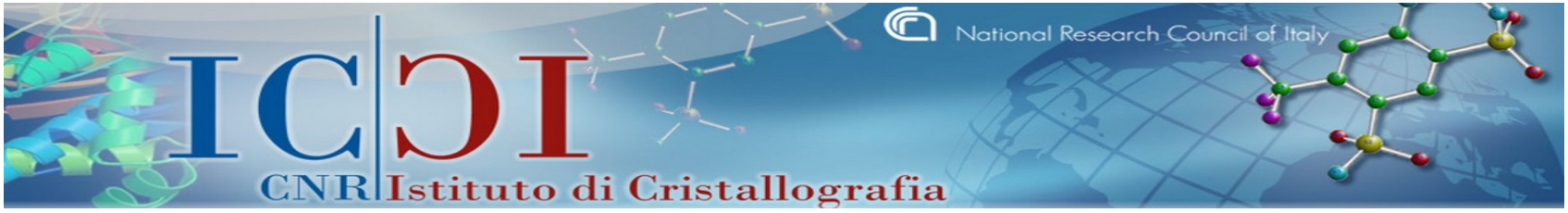
Doriano Lamba

Istituto di Cristallografia, C.N.R – Trieste Outstation

Structural Biology Laboratory

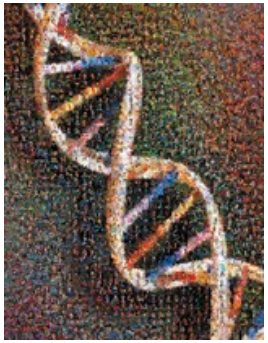
Area Science Park – Basovizza, S.S. 14, Km 163,5 I-34149 Trieste

E-mail: doriano.lamba@ic.cnr.it

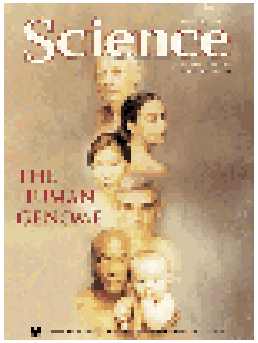


From Molecular Genetics to Structural Biology
via X-ray Crystallography

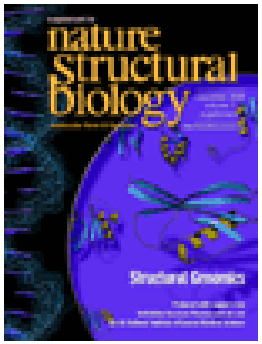
Nature **171**: 4356
25 April 1953



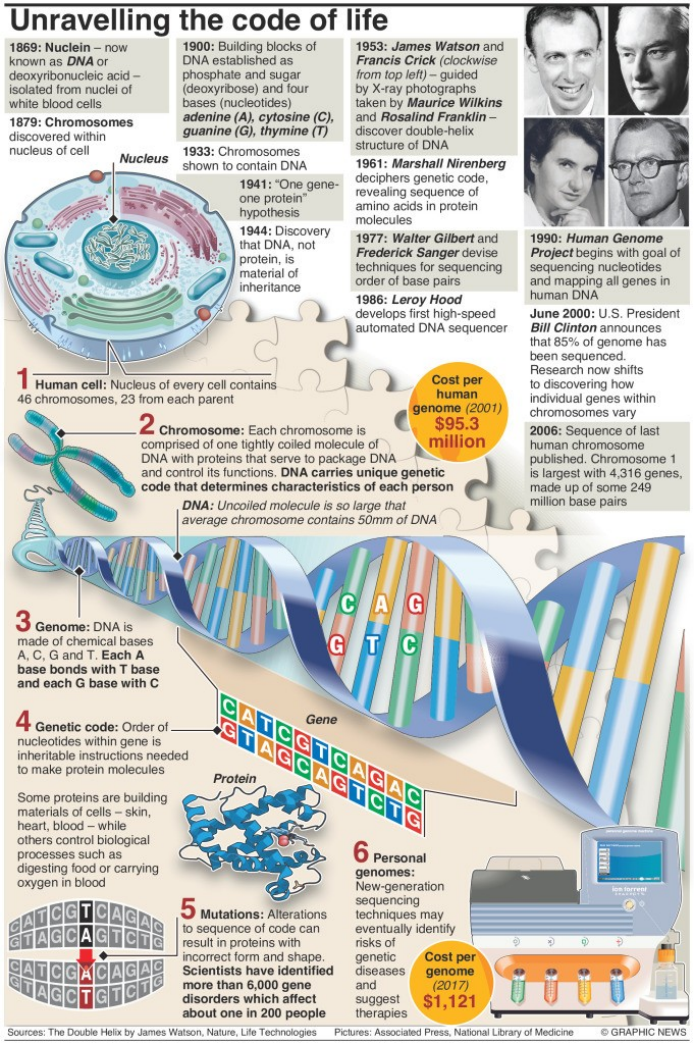
Nature **409**: 6822
15 February 2001



Science **291**: 5507
16 February 2001



Nature Structural Biology **7**
November 2000



"***Life happens too fast for you ever to think about it***"
K. Vonnegut

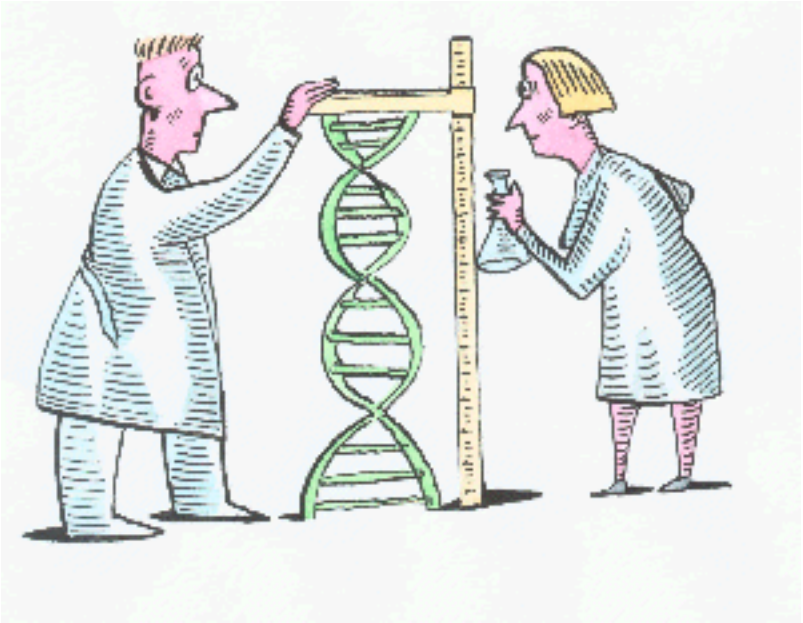
A Tour into the 3D secrets of the Biological Matter
Starring Guides: the Nobel prizes!!!

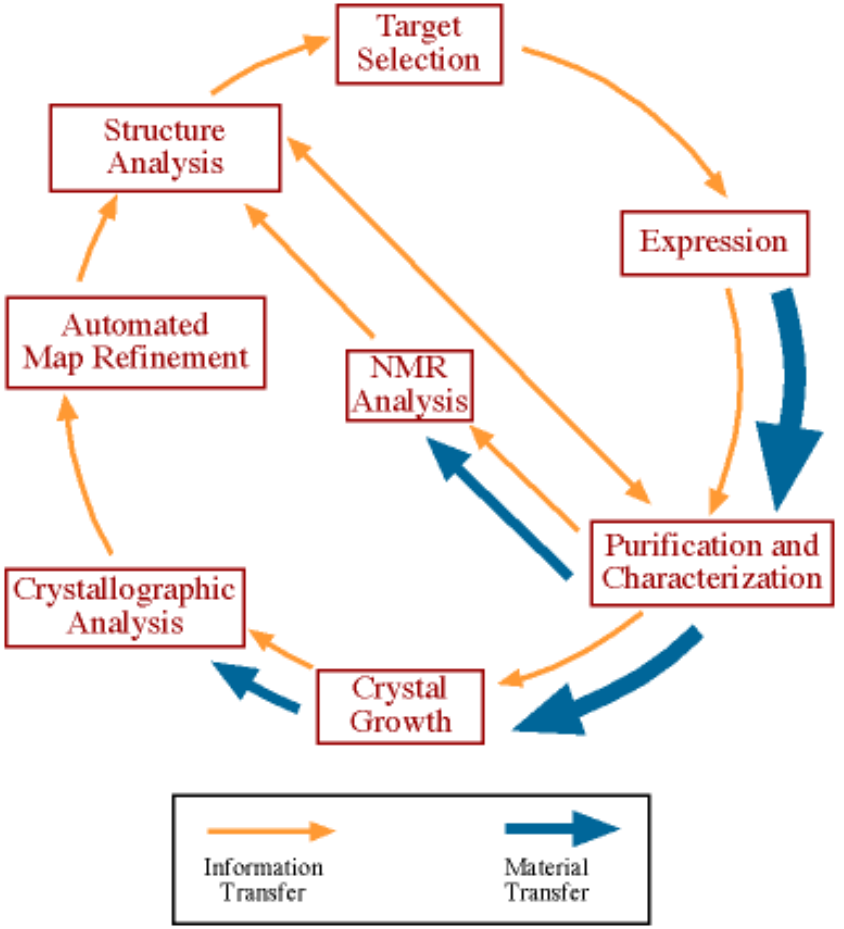
Recipient	Year	Discipline	Nationality	Awarded
Wilhelm Conrad Röntgen	1901	Physics	Germany	In recognition of the extraordinary services he has rendered by the discovery of the remarkable rays subsequently named after him
Max von Laue	1914	Physics	Germany	For discovery of the diffraction of X-rays by crystals
William Henry Bragg	1915	Physics	UK	For their services in the analysis of crystal structure by means of X-rays
William Lawrence Bragg	1915	Physics	UK	
Peter Debye	1936	Chemistry	Germany	For his contributions to our knowledge of molecular structure through his investigations on dipole moments and on the diffraction of X-rays and electrons in gases
Clinton Joseph Davisson	1937	Physics	USA	For their experimental discovery of the diffraction of electrons by crystals
George Paget Thomson	1937	Physics	UK	
James Batcheller Sumner	1946	Chemistry	USA	For his discovery that enzymes can be crystallized
John Howard Northrop	1946	Chemistry	USA	For their preparation of enzymes and virus proteins in a pure form
Wendell Meredith Stanley	1946	Chemistry	USA	
Linus Pauling	1954	Chemistry	USA	For his research into the nature of the chemical bond and its application to the determination of the structure of complex substances
John Kendrew	1962	Chemistry	USA	For their studies of the structures of globular proteins
Max Perutz	1962	Chemistry	UK	
Francis Crick	1962	Medicine	UK	For their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material
James Watson	1962	Medicine	UK	
Maurice Wilkins	1962	Medicine	UK	
Dorothy Hodgkin	1964	Chemistry	UK	For her determinations by X-ray techniques of the structures of important biochemical substances
E.Betzig, S.W Hell, E Moerner 2014		Chemistry		*"for the development of super-resolved fluorescence microscopy".*
J. Frank, R.Henderson, J. Dubochet 2017		Chemistry		*"for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".*
Aaron Klug	1982	Chemistry	UK	For his development of crystallographic electron microscopy and his structural determination of biologically important nucleic acid–protein complexes
Herbert Hauptman	1985	Chemistry	USA	For their outstanding achievements in the development of direct methods for the determination of crystal structures
Jerome Karle	1985	Chemistry	USA	
Johann Deisenhofer	1988	Chemistry	Germany	For the determination of the three-dimensional structure of a photosynthetic reaction center
Robert Huber	1988	Chemistry	Germany	
Hartmut Michel	1988	Chemistry	Germany	
Pierre-Gilles de Gennes	1991	Physics	France	For discovering that methods developed for studying order phenomena in simple systems can be generalized to more complex forms of matter, in particular to liquid crystals and polymers
Georges Charpak	1992	Physics	France	For his invention and development of particle detectors, in particular the multiwire proportional chamber
Bertam Brockhouse	1994	Physics	Canada	For the development of neutron spectroscopy
Clifford Shull	1994	Physics	USA	For the development of the neutron diffraction technique
John Walker	1997	Chemistry	UK	For the determination of the enzymatic mechanism underlying the synthesis of ATP
Roderick MacKinnon	2003	Chemistry	USA	For structural and mechanistic studies of ion channels
Roger Kornberg	2006	Chemistry	USA	For his studies of the molecular basis of eukaryotic transcription
Venkatraman Ramakrishnan	2009	Chemistry	UK	For studies of the structure and function of the ribosome
Thomas Steitz	2009	Chemistry	USA	
Ada Yonath	2009	Chemistry	Israel	
Andre Geim	2010	Physics	UK	For groundbreaking experiments regarding the two-dimensional material graphene
Konstantin Novoselov	2010	Physics	UK	
Dan Shechtman	2011	Chemistry	Israel	For the discovery of quasicrystals
Robert Lefkowitz	2012	Chemistry	USA	For studies of G-protein coupled receptors
Brian Kobilka	2012	Chemistry	USA	
Martin Karplus	2013	Chemistry	USA	For the development of multiscale models for complex chemical systems
Michael Levitt	2013	Chemistry	USA	
Arieh Warshel	2013	Chemistry	USA	

From Molecular Genetics to Structural Biology
via X-ray Crystallography

Large Scale Determination and Analysis Of 3D Structures

High-Throughput!!!





Brookhaven Protein Database

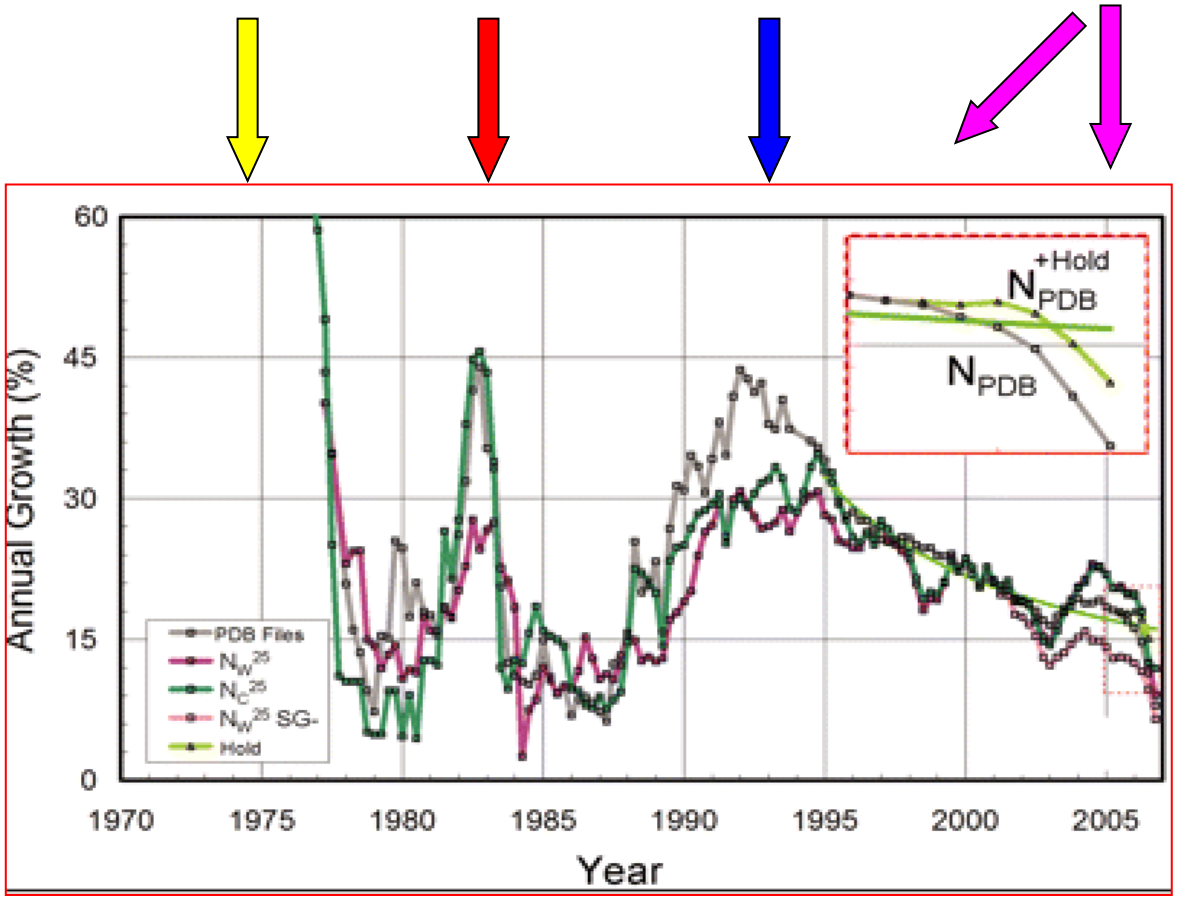
Structural data growth since the inception of the PDB (1972)

Initial explosion of Protein X-ray Crystallography

Digital VAX 780 Computers

Synchrotron Radiation Cryocooling

Protein Structure Initiative

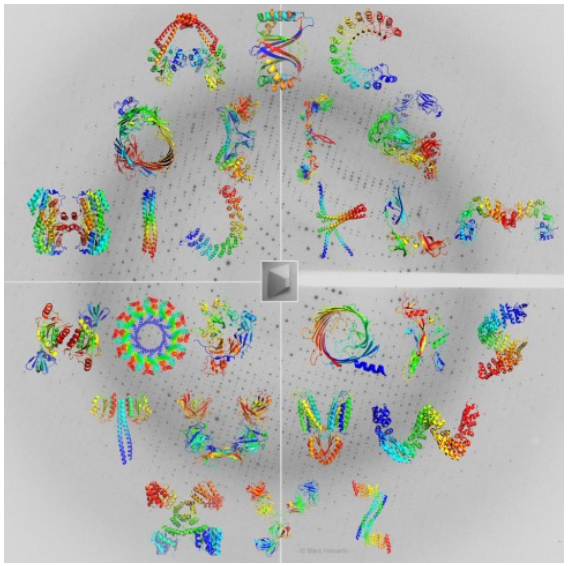


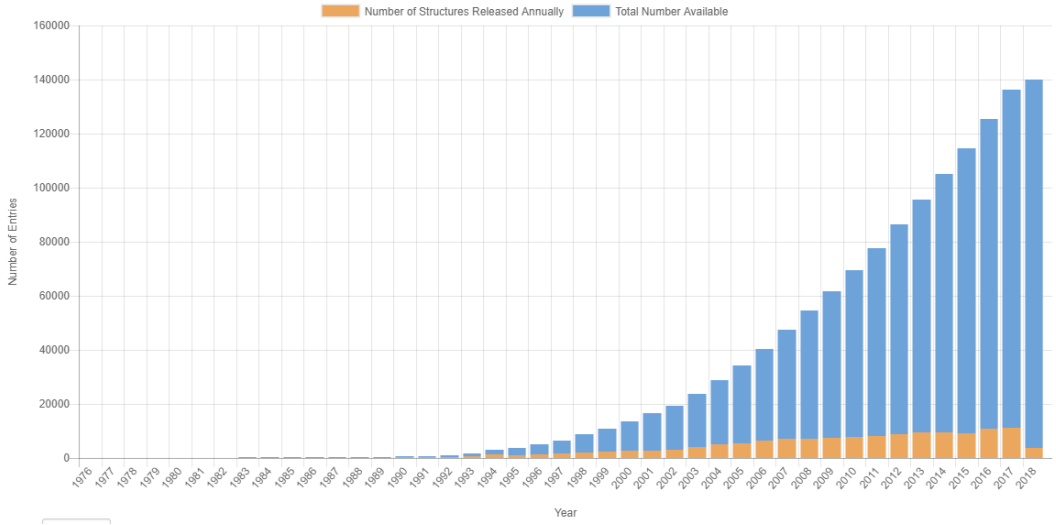
M. Levitt, Proc. Natl. Acad. Sci. USA (2007) 104, 3183

Brookhaven Protein Database (2018)

PDB Current Holdings Breakdown

Exp.Method	Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
X-RAY	117483	1923	6011	4	125421
NMR	10704	1243	249	8	12204
ELECTRON MICROSCOPY	1546	31	542	0	2119
HYBRID	120	4	2	1	127
other	215	4	6	13	238
Total	130068	3205	6810	26	140109

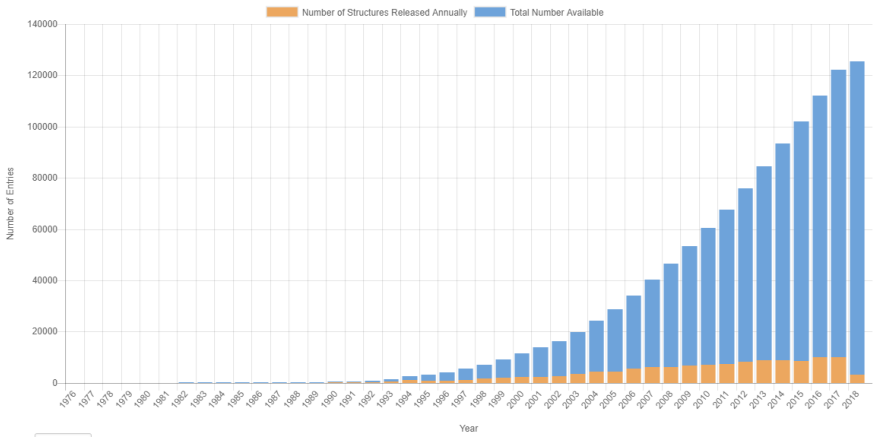




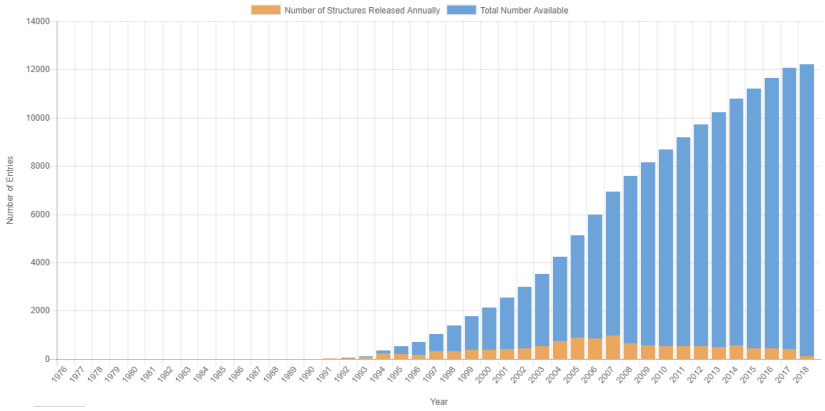
Overall Growth of Released Structures Per Year

Last database update: May 8th 2018

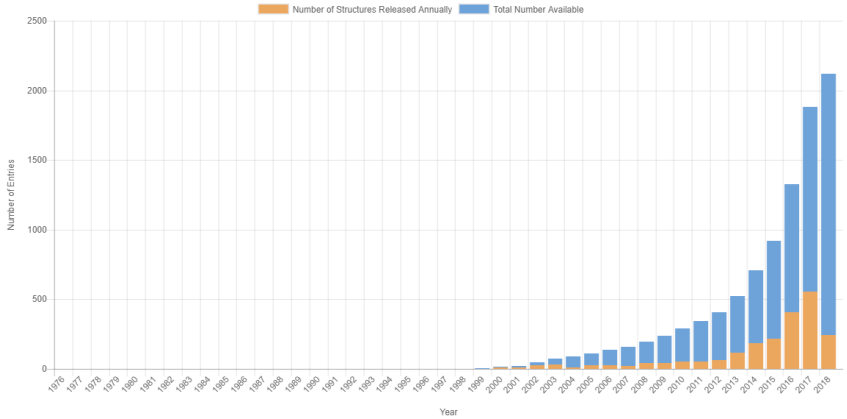
Brookhaven Protein Database



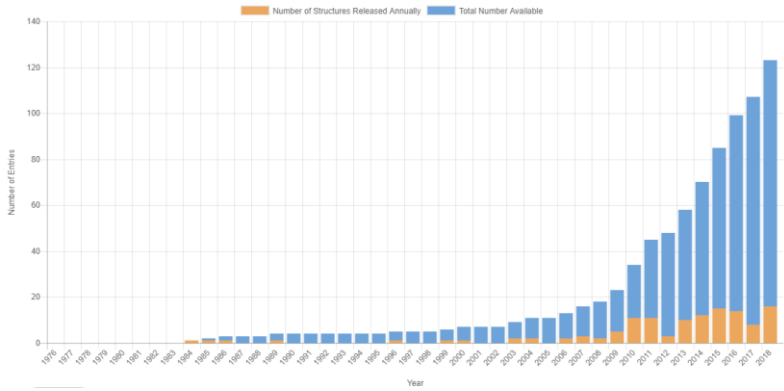
Growth of Structures from **X-ray** Experiments Released per Year



Growth of Structures from **NMR** Experiments Released per Year

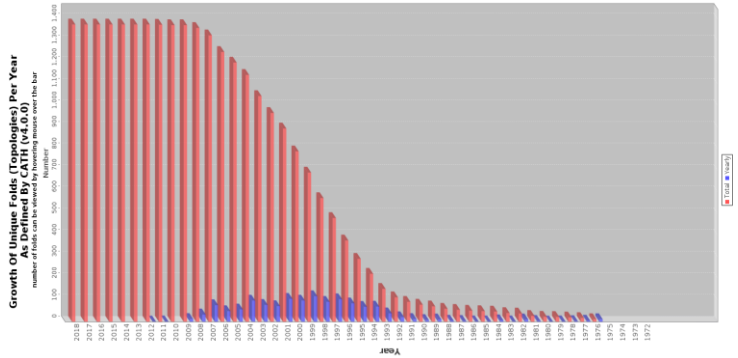


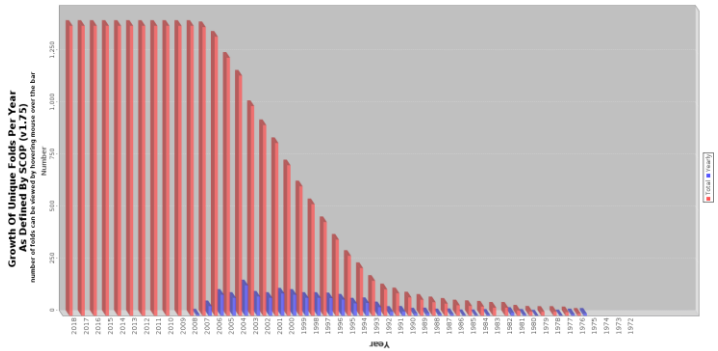
Growth of Structures from 3D **EM** Experiments Released per Year



Growth of Structures from **Hybrid** Methods Released per Year

Brookhaven Protein Database





Structural Biology Scales Down

The USA is winding down a $1 billion project to churn out protein structures

Science (2014) 343:1073-1075

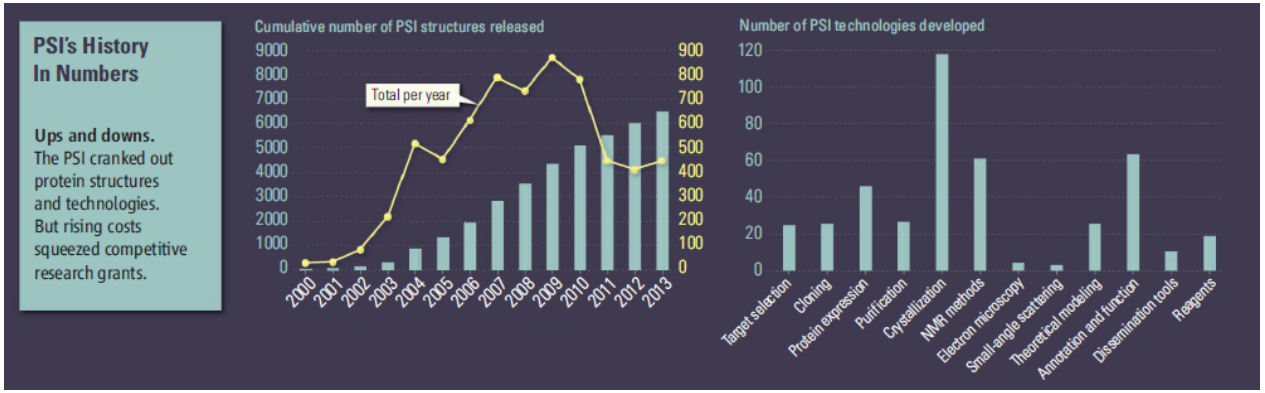
Crystallography at 100

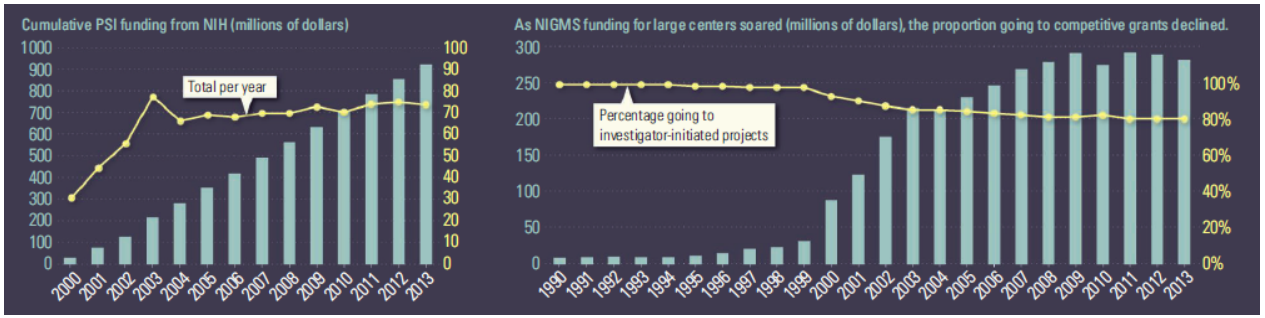
2014

international year of crystallography

United Nations Educational, Scientific and Cultural Organization · International Union of Crystallography

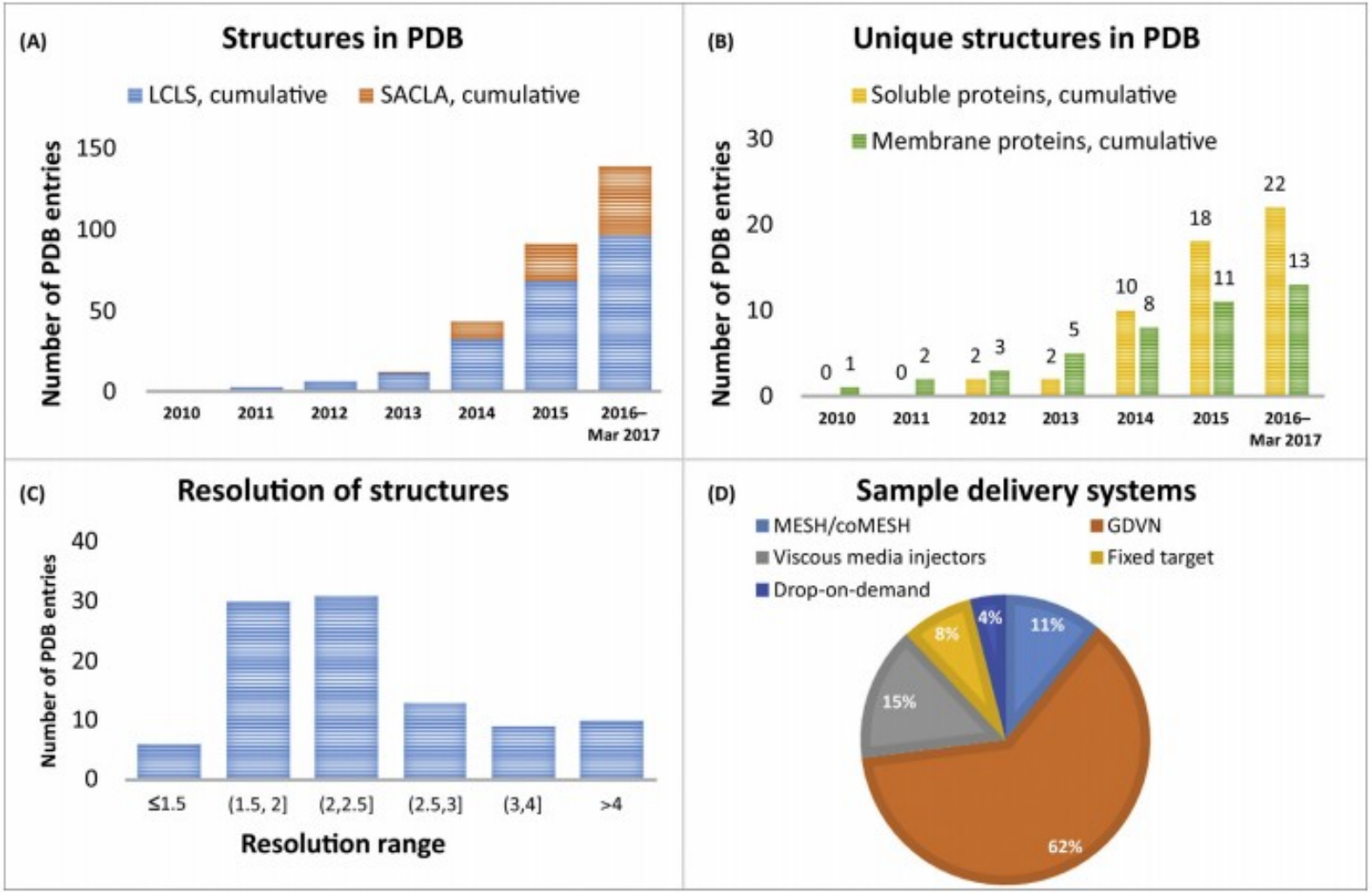
Partners for the International Year of Crystallography 2014



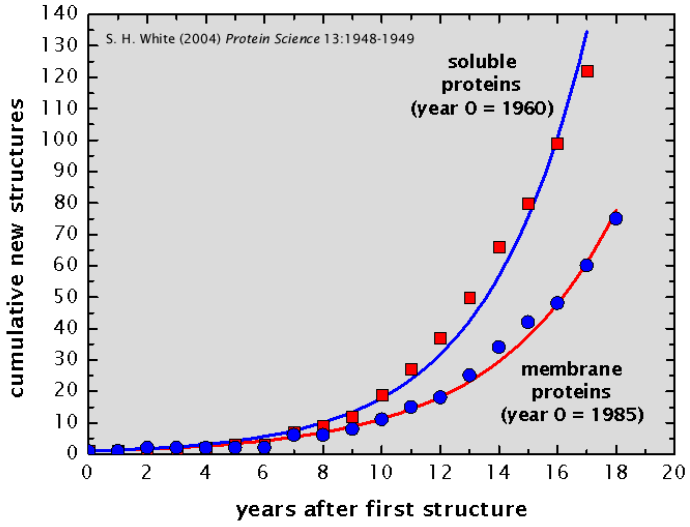


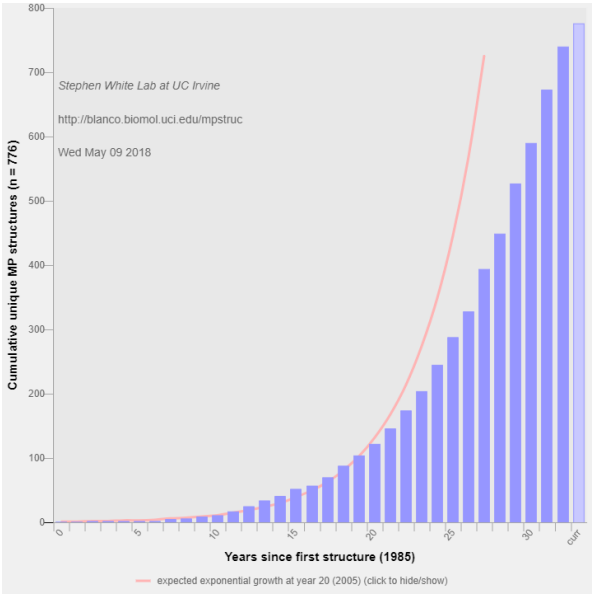
Statistics on the Macromolecular Structures Solved Using X-FELs To Date

Trends in Biochemical Sciences (2017) 42:749–762



Membranes Proteins Structures





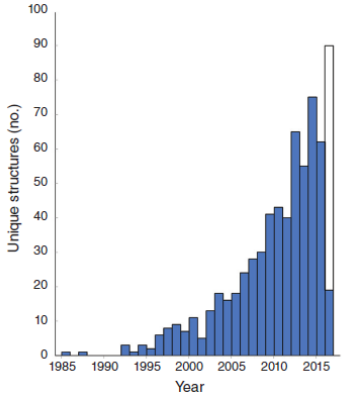
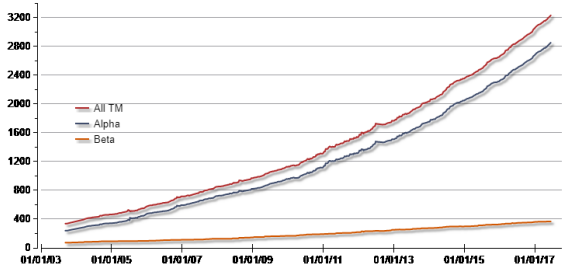


Figure 1 Growth in membrane-protein structures. Annual numbers of unique structures are plotted as a function of time. The numbers of unique structures are taken as those defined and recorded as of 17 March 2016 in the White database of membrane proteins of known structure (http://blanco.biomol.uci.edu/mpstruc/). The open portion of the bar for 2016 projects production continuing at the same rate as that of the year to date.

Raman, P., Cherezov, V., Caffrey, M. The Membrane Protein Data Bank. *Cell Mol. Life Sci.* (2006) 63, 36



PDBTM version: June 16 2017
Number of transmembrane proteins: 3227
(alpha: 2848 , beta: 366)

Last database update**: May 8th 2018**

Unique proteins in database = 776

Number of coordinate files in database = 2520

Integral membrane proteins whose structures have been determined by ɔlution-state NMR, oriented solid-state NMR or magic-angle spinning solid-state NMR

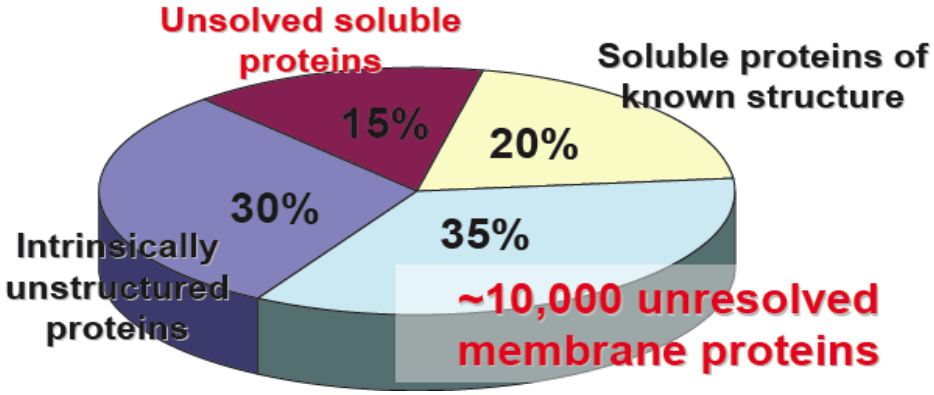
Unique proteins in database : 186; **Number of coordinate files in database : 294**

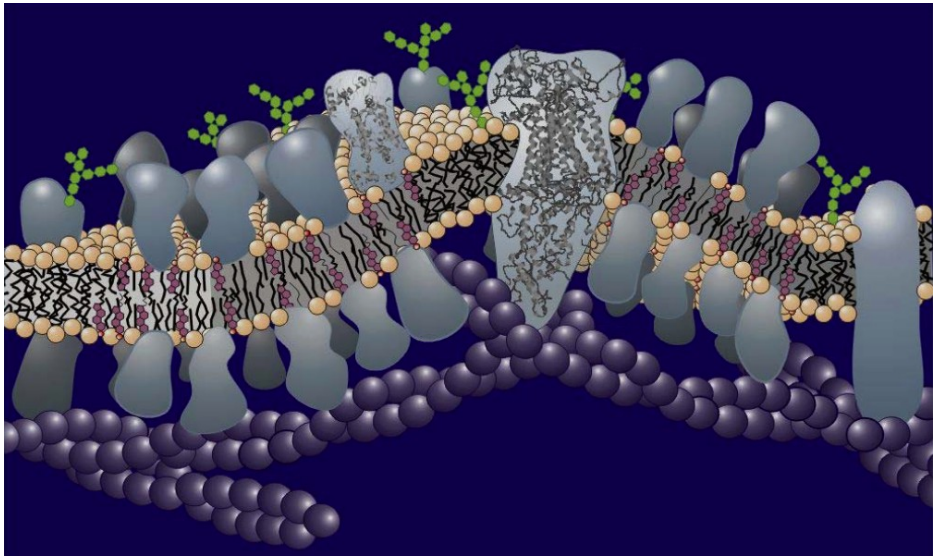
Last database update: April 8th 2015

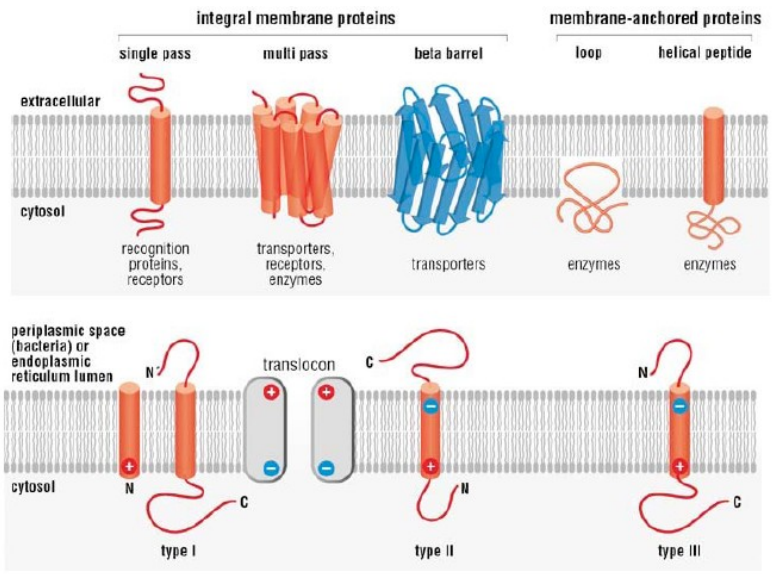
Membrane Strutural Biology is lagging 20-25 years behind the study of soluble proteins

Membranes Proteins Structures

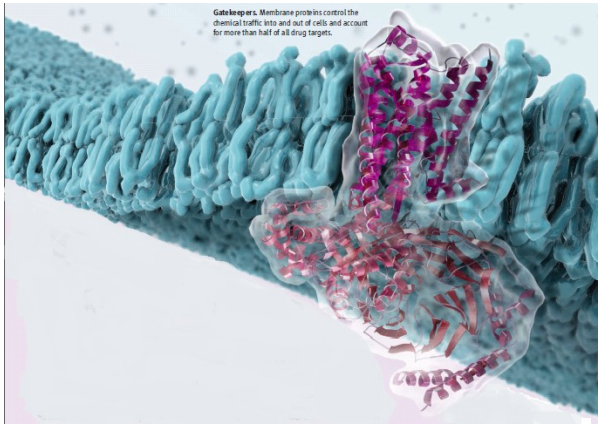
Human genome: ~30,000 proteins

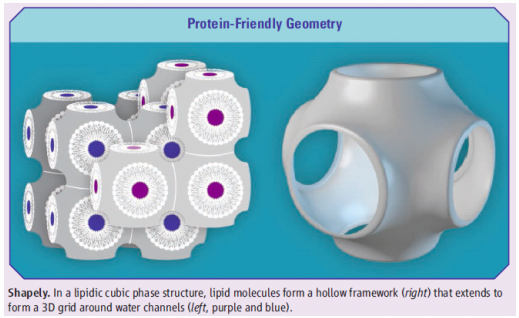


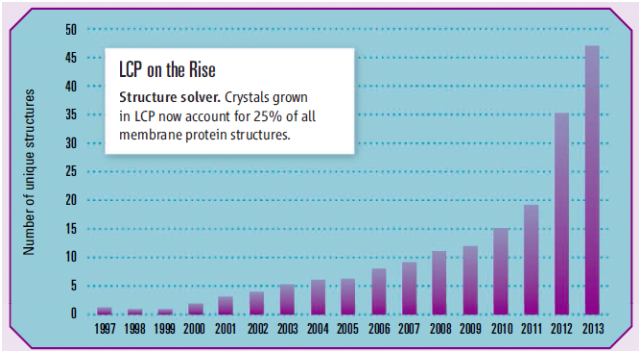


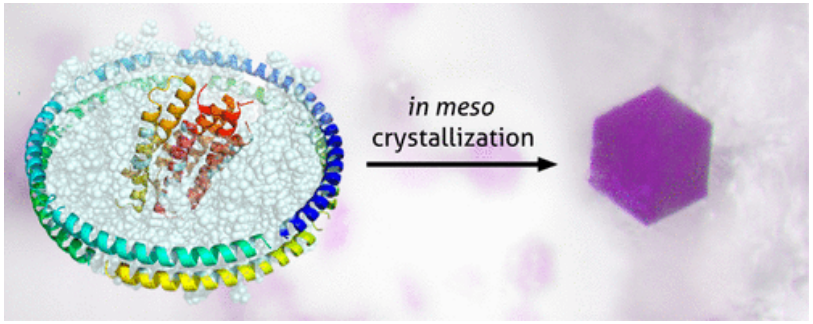


Membrane Protein Structure

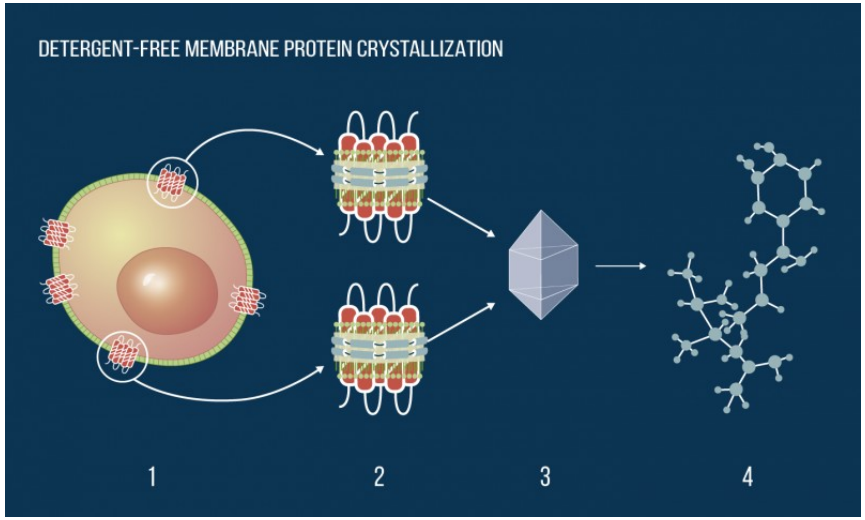






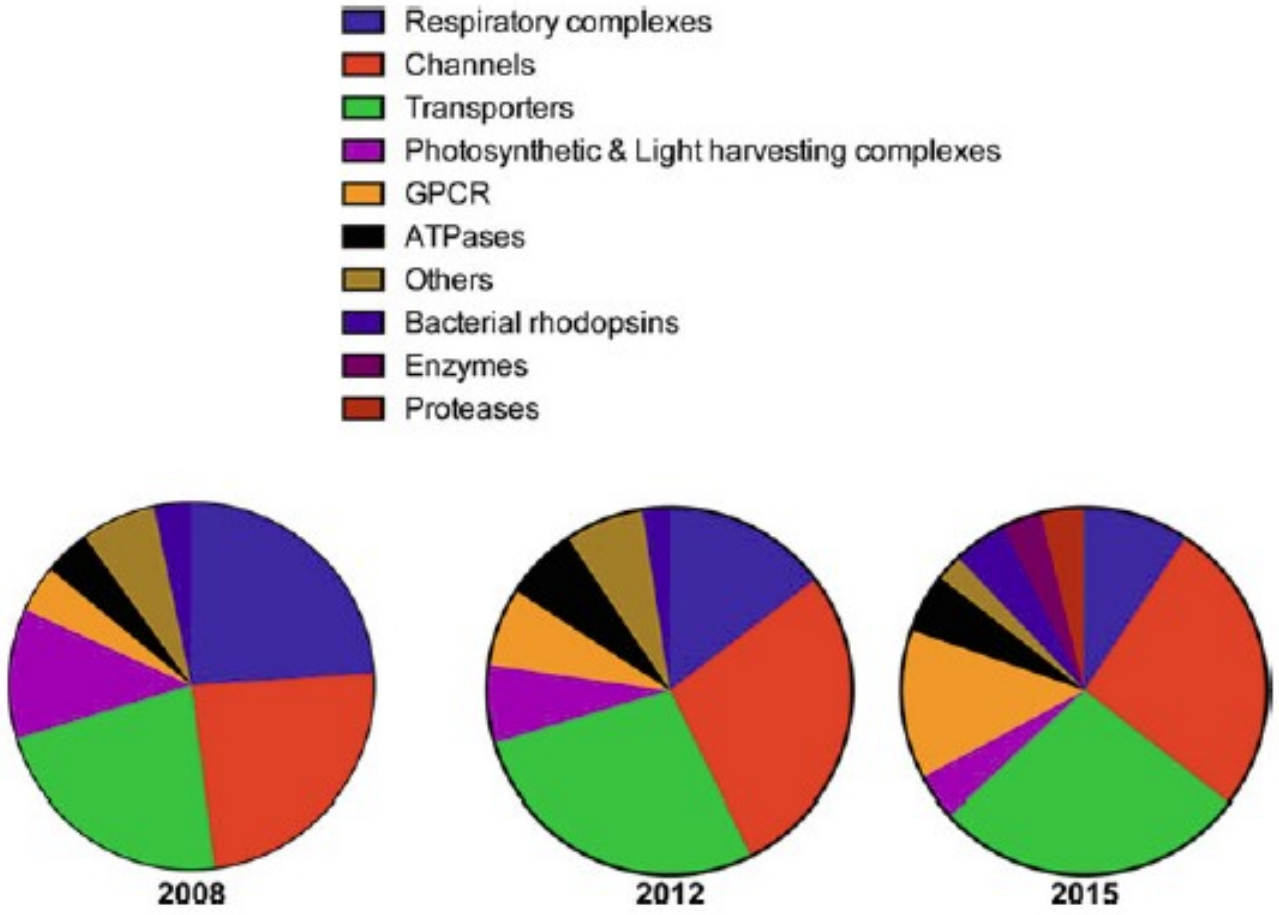


Membrane-like nanodiscs (ND) have become an important tool for the cell-free expression, solubilization, folding, and in vitro structural and functional studies of membrane proteins (MPs). Direct crystallization of MPs embedded in NDs would be of high importance for structural biology



Membrane proteins in a cell membrane (1). The same proteins (2) embedded in nanodiscs and stabilized in a water solution by a specially engineered protein belt. A protein crystal (3). Molecular structure (4) obtained by X-ray crystallography.

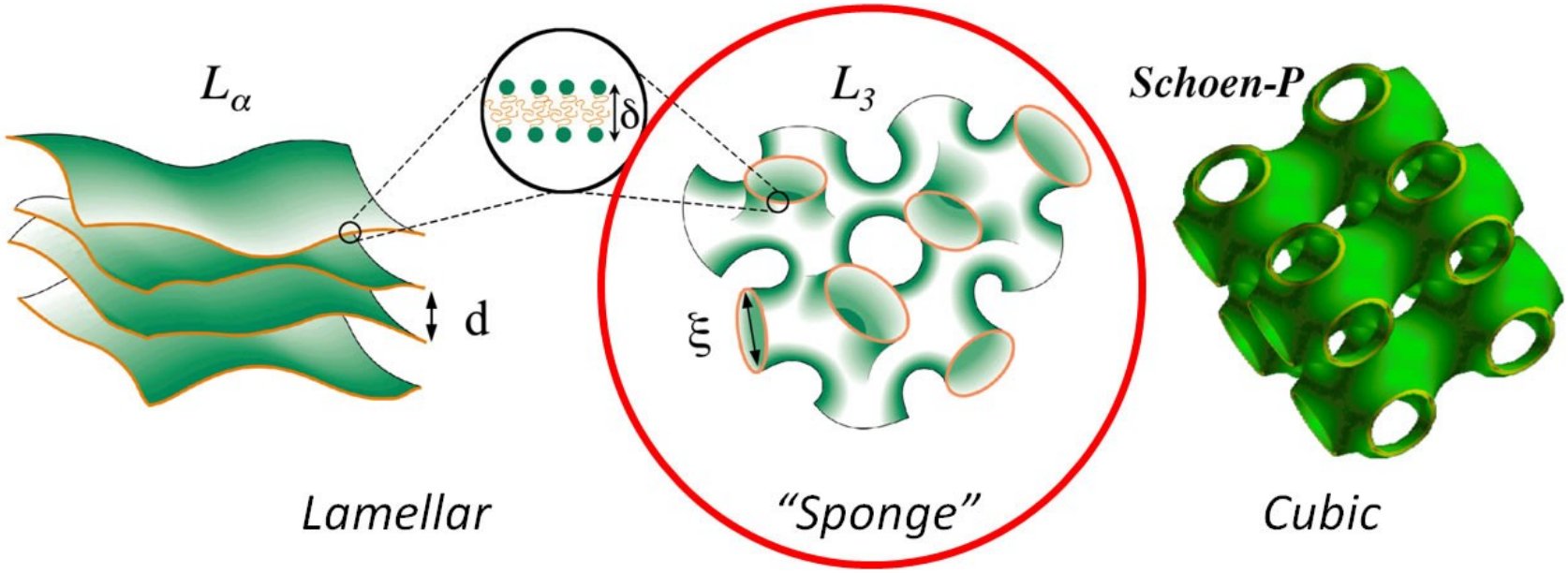
Membrane Protein Structure



Number and types of alpha helical membrane protein structures reported from 2008 to 2015.
Pie charts showing the change in the proportion of structures belonging to each family group from 2008, 2012 and 2015. Membrane proteins were broken down into the following families:
Respiratory complexes (*blue), GPCRs* (*orange), ATPases (black), Bacterial Rhodopsins (dark blue), Enzymes (purple), Proteases (dark red) and where* a protein didn't fit in these categories, others (*brown)*

Serial Femtosecond Crystallography

A microsized jet required to deliver microcrystals



Lipidic Sponge Phase is made by mixing monoolein and water with:

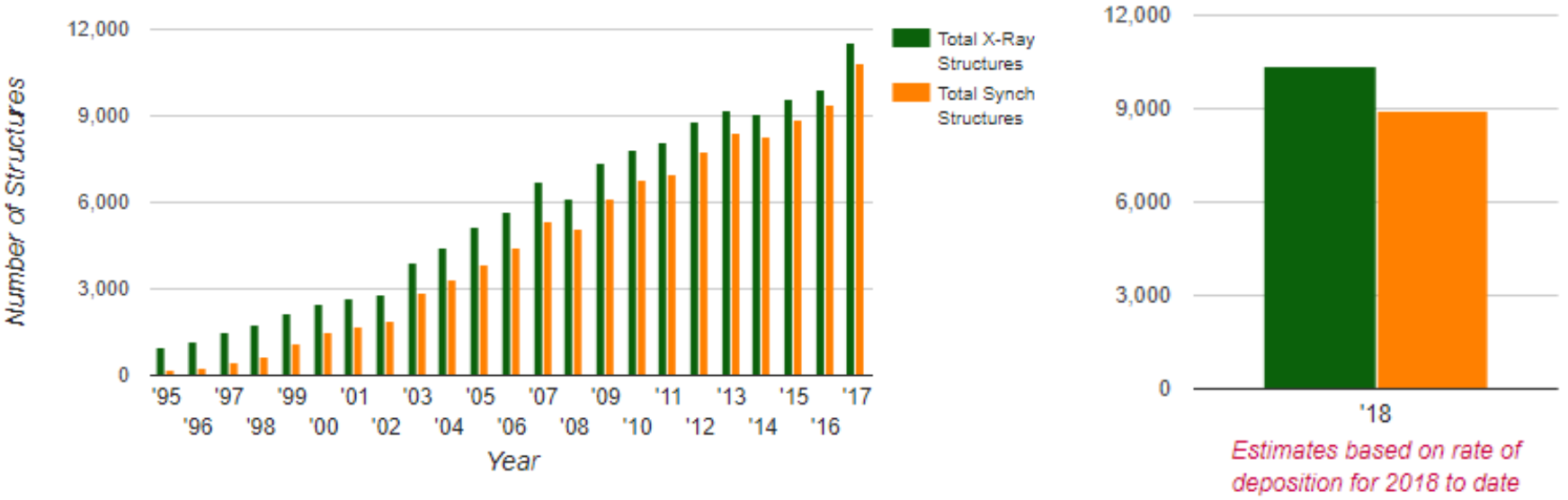
Jeffamine or PEG

that swells the cubic phase into a liquid phase

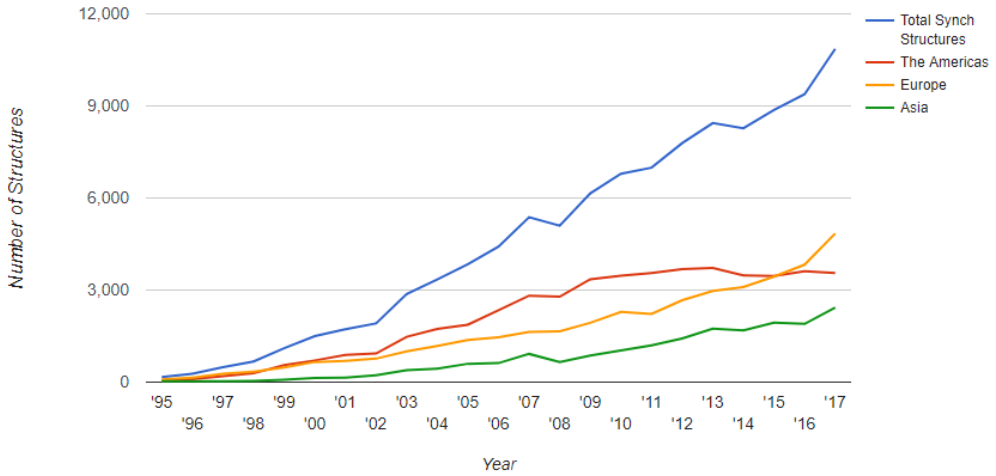
It can be adapted to serial femtosecond crystallography

Lipidic Cubic Phase, because of its semisolid nature does not readily form a micrometer sized jet required to deliver microcrystals

Synchrotron Structures Deposited in the PDB vs all PDB Deposited Structures



Regional synchrotron structures deposited in the PDB vs all synchrotron structures



http://biosync.sbkb.org/
Update May 8th 2018

James Batcheller Sumner (1887-1955)

Nobel prize in Chemistry 1946 *for his discovery that enzymes can be crystallized*

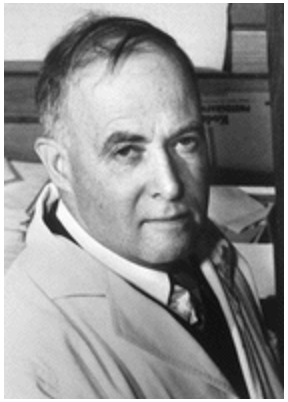


Fig. 1. Crystalline pepsin (x 90). Northrop, 1930.

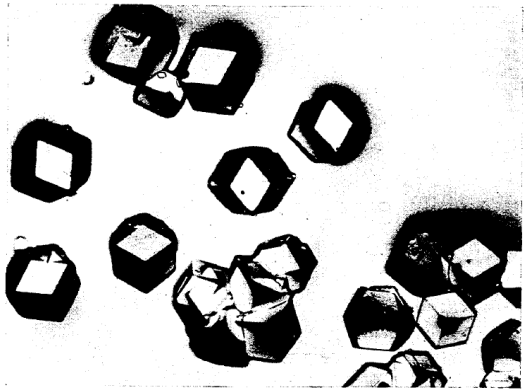


Fig. 12. Crystals of tomato bushy stunt virus (x 224). (From W. M. Stanley, J. *Biol. Chem., 135 (1940) 437.)*

John Howard Northrop (1891-1987), Wendell Meredith Stanley (1904-1971)

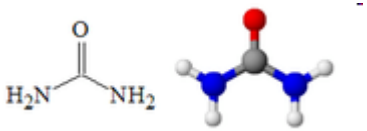
Nobel prize in Chemistry 1946 *for their preparation of enzymes* (pepsin from porcine gastric mucose) *and virus proteins* (Tomato bushy stunt virus*) in a pure form*

James Batcheller Sumner (1887-1955)

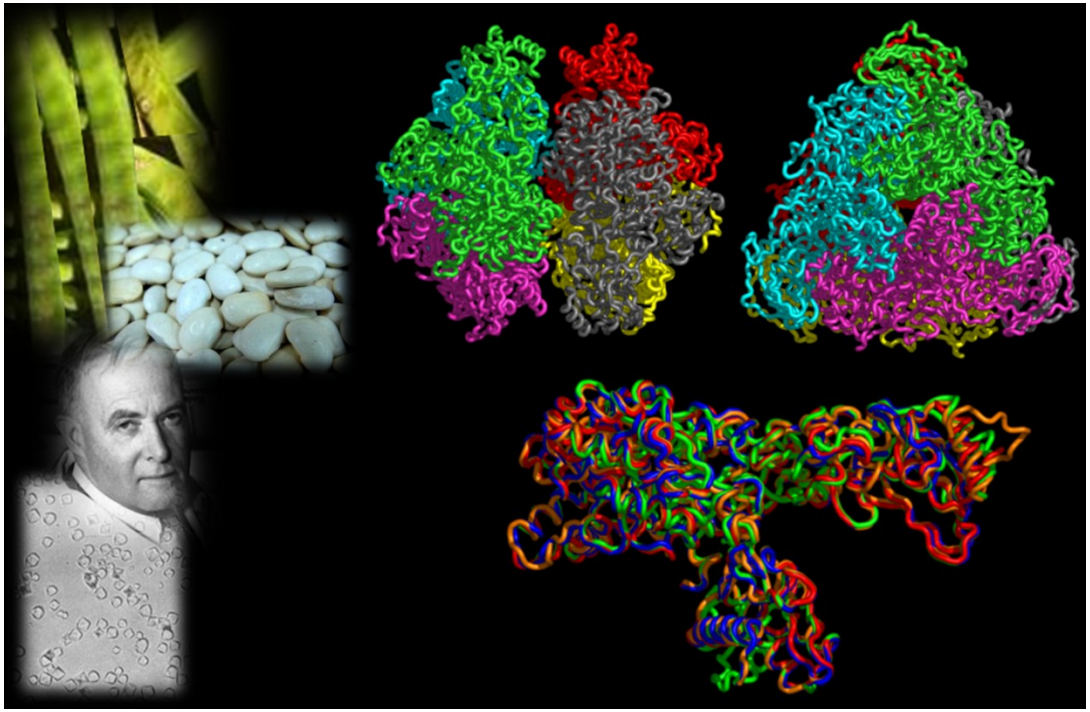
Nobel prize in Chemistry 1946 *for his discovery that enzymes can be crystallized*.

J.B. Sumner Isolation and crystallization of the enzyme urease. J. Biol. Chem. (**1926**) 69, 435

Canavalia ensiformis



Urea

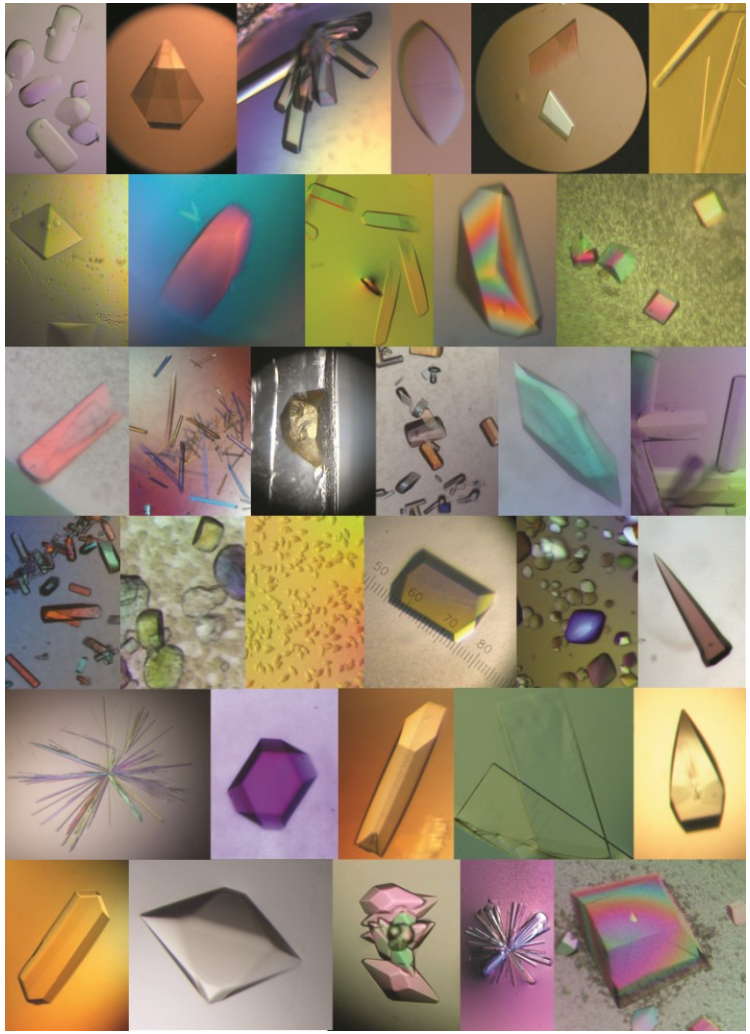


The enzyme urease catalyzes the reaction of urea with water to produce carbon dioxide and ammonia. In water, without the enzyme, the reaction proceeds with a first-order rate constant of 4.15 s^{-1} at 100°C. In the presence of the enzyme in water, the reaction proceeds with a rate constant of 3.4 s^{-1} at 21° C.

Crystal structure of the first plant urease from jack bean: 83 years of journey from its first crystal to molecular structure

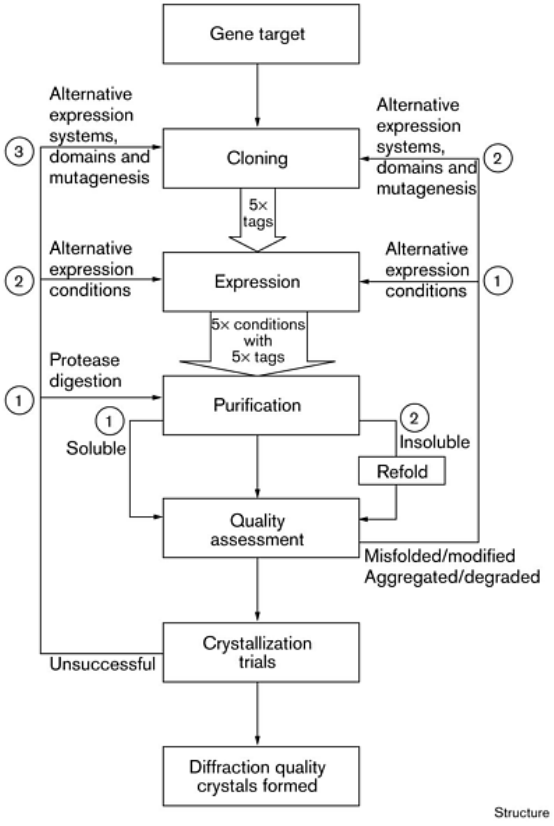
Balasubramanian A, Ponnuraj K., J Mol Biol (2010) Jul 16 400, 274

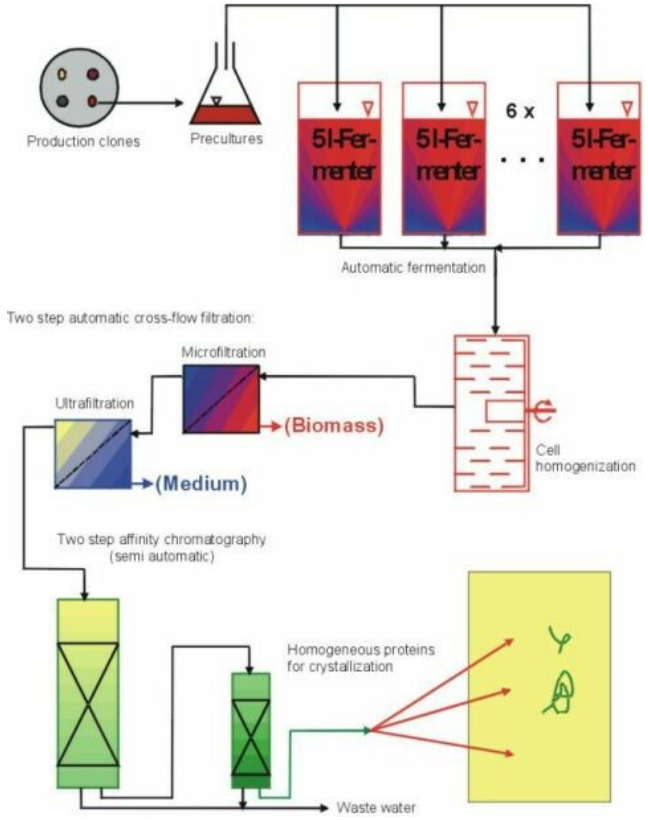
Crystallization is an Art



Crystallography is a Science

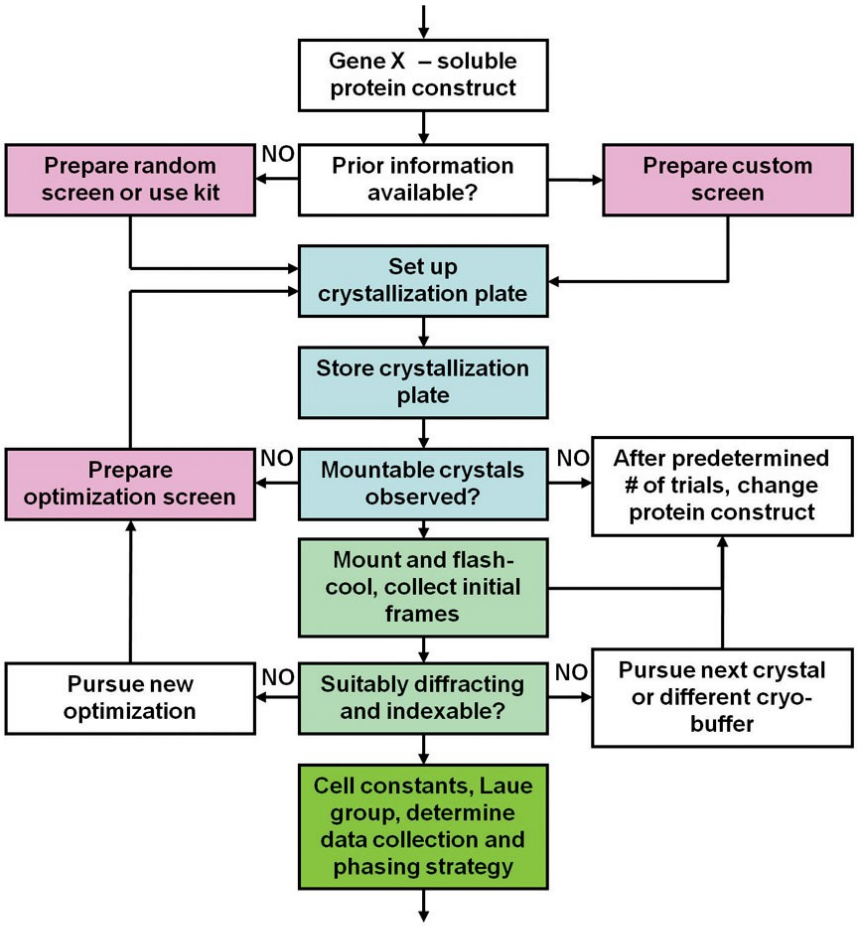
Overview of the steps involved in high-throughput protein production for structural studies.

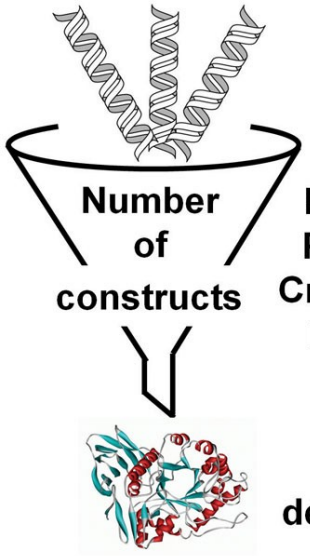




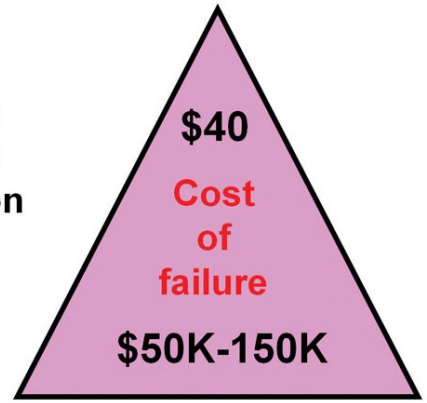
High-throughput Gene Cloning, Protein Expression, Refolding and Production:
Feeding The Crystallographers And NMR Spectroscopists

Workflow of a Crystallization Project





Cloning
Expression
Purification
Crystallization
Harvesting
Data collection
Structure determination



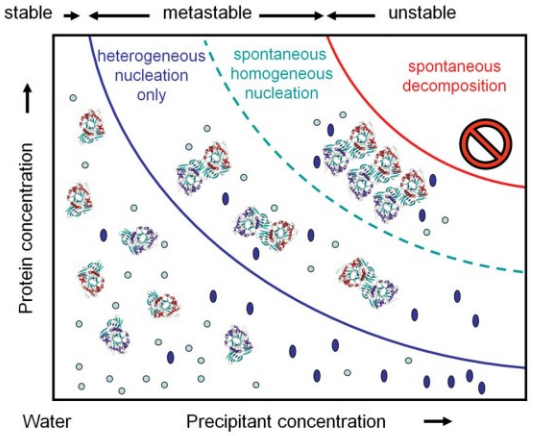
According to NIH, the total average cost of a PX in 2006 amounted to 100K $.

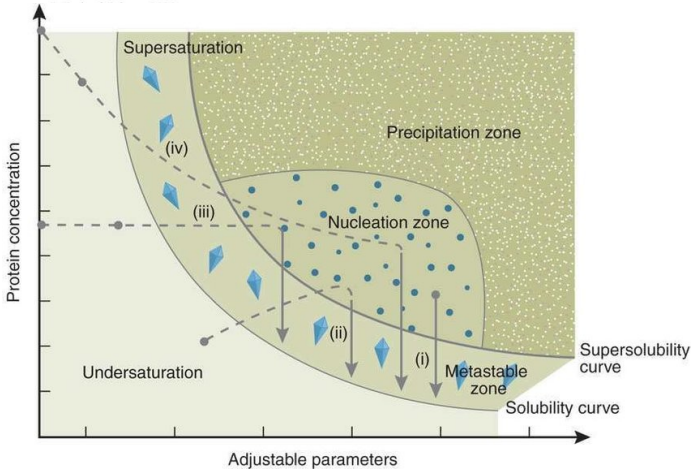
A protein crystal of 0.036 mm^3 (ca. 0.43 mg) is *ca*. 3-4 times more expensive of a quality diamond (5 carats) !!!

Protein Crystallization: Physical and energetic principles

What properties confer crystallizability on a protein or protein complex?

- **The protein must have a surface that confers adequate solubility to reach supersaturation levels required for nucleation;**
- **The surface must contain patches with structured water solvent, allowing for the ordering of nascent nuclei by mediating thermodynamically viable intermolecular contacts;**
- **There should be few, if any, unstructured elements that elevate the entropic cost of crystallization, such as intrinsically disordered N- and C-termini, long partly or wholly disordered loops, or flexible carbohydrate moieties due to posttranslational modifications.**
- **Other properties, such as the Gravy index and pI may be indirectly, but positively correlated with crystallizability**

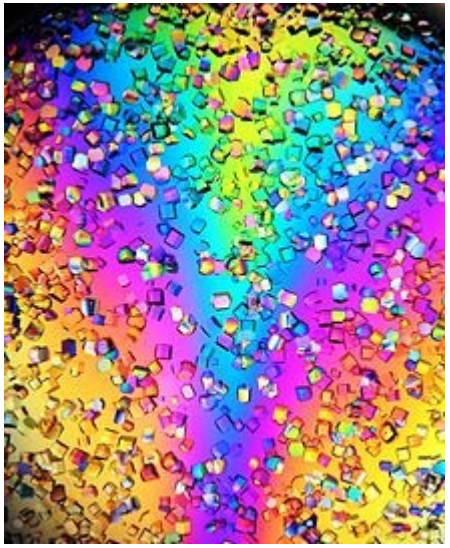




Higher supersaturation is necessary for spontaneous formation of stable crystallization nuclei (homogeneous formation), while at low supersaturation nucleation requires external seeds in the form of microcrystals or other particulate matter (heterogeneous nucleation)

Protein Crystals

- Composed of ~50% solvent on average (can range from 25 to 90%)
- Similar to ordered gels, with large pores
- Small molecules and solvent can diffuse freely into these pores
- Molecules form a relatively small number of bonds with their neighbours
- Mechanically fragile, soft and easy to crush
- Sensitive to dehydration
- Sensitive to temperature changes
- Diffract X-rays poorly
- Sensitive to radiation damage



Parameters Affecting Protein Crystallization

Biochemical and biophysical parameters:

- Sensitivity of the protein against physical parameters like pH, temperature etc.
- Binding of ligands like inhibitors, cofactors, metal ions etc
- Specific additives like reducing agents, detergents etc.

Physico-chemical parameters:

- Concentration of protein and precipitants
- Temperature, pH, pressure, electric and magnetic fields,
- Surface exposed to the air, interface effects
- Time to reach supersaturation
- Ionic strength and purity of the chemicals used
- Density and viscosity effects, speed of diffusion and convection

Biological parameters:

- amount of protein available
- different biological sources of proteins (thermophiles vs halophiles, mesophiles; growing vs. stationary phase)
- contamination (fungi or bacteria)

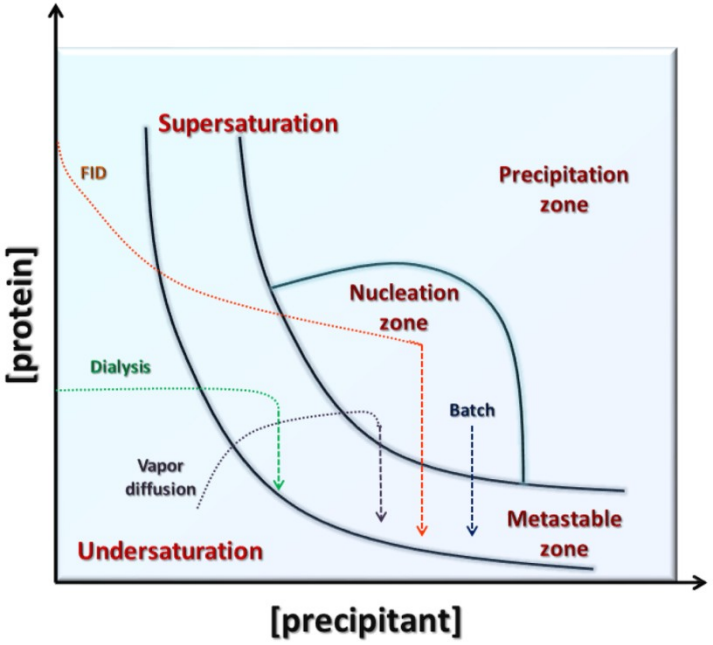
Purity of protein:

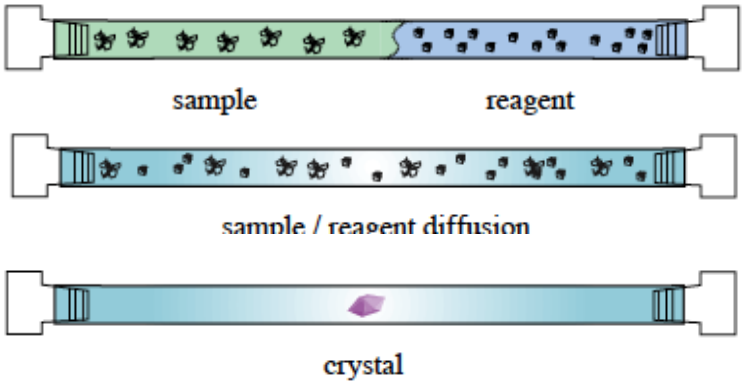
- macromolecular contaminants
- batch differences
- microheterogeneities

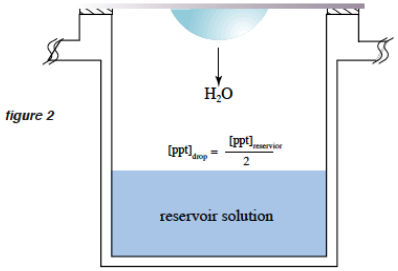
Sources of microheterogeneity:

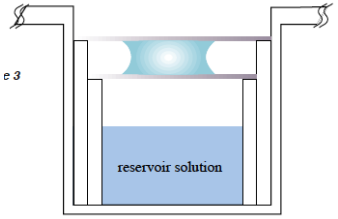
- Variation in primary structure (genetic variations, degradations)
- Variation in secondary structure (errors in folding or partial unfolding)
- Variation in tertiary structure (flexible domains, aggregation, conformer equilibrium)
- Variation in quaternary structure (oligomerization)
- Variation in post-translational modifications (e.g. glycosylation)
- Partial binding of ligands
- Fragmentation (i.e. hydrolysis)
- Partial oxidation (e.g. sulphydryl groups in proteins)
- Ageing (e.g. Deamidation, isomerization, racemization of Asn to isoAsp are key steps in crystallization)

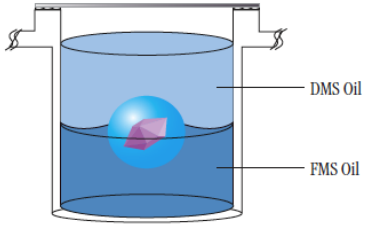
Protein Crystal Growth Techniques

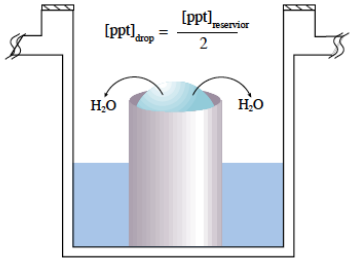


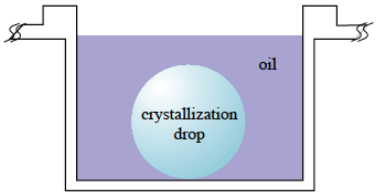


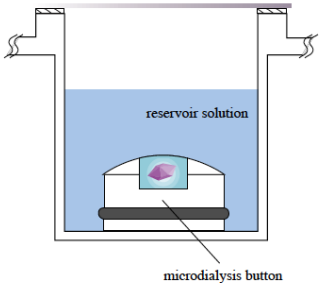


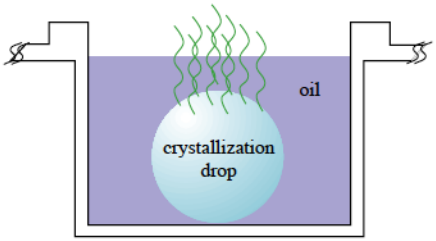




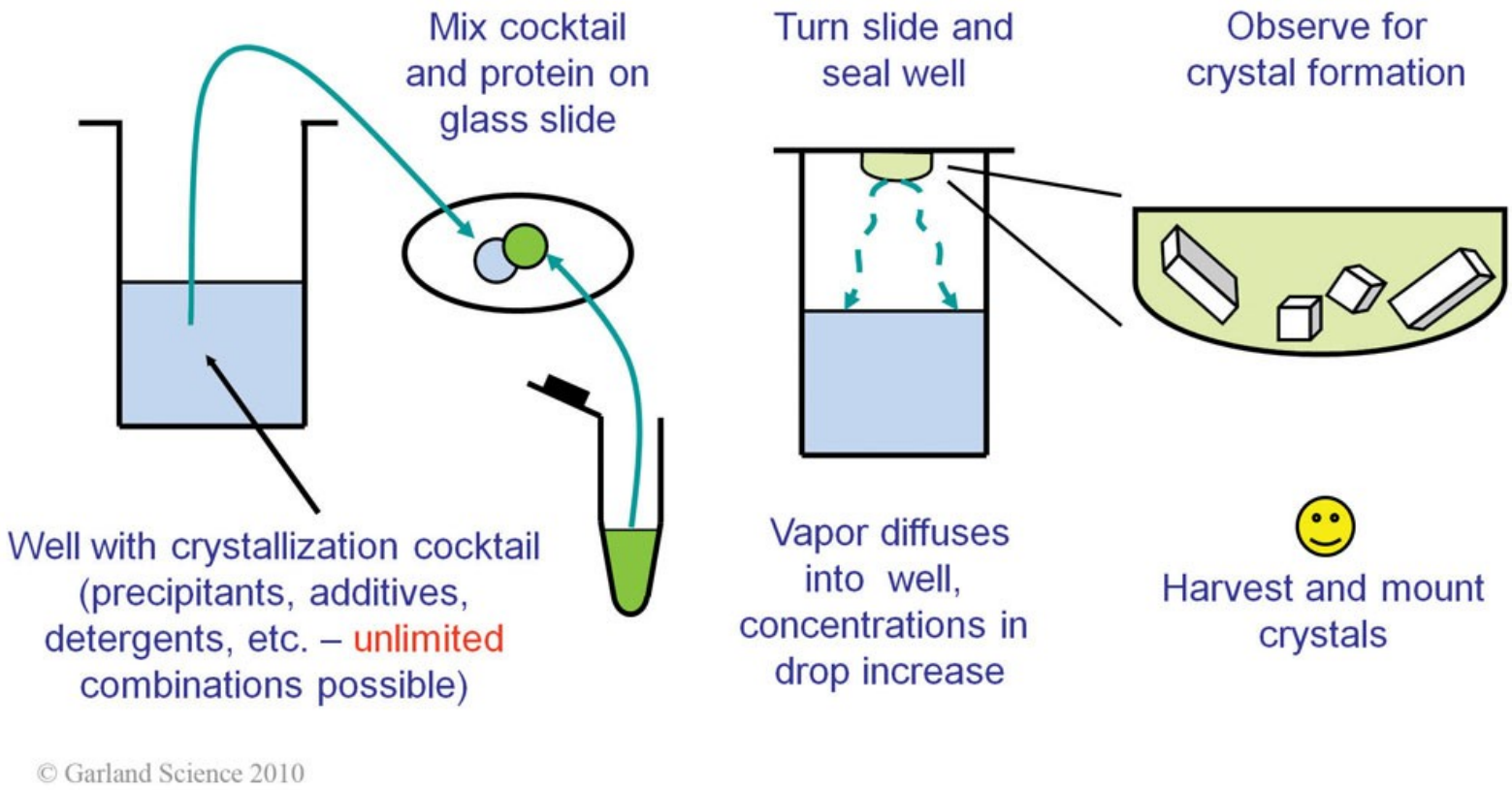








Basic hanging-drop vapor diffusion



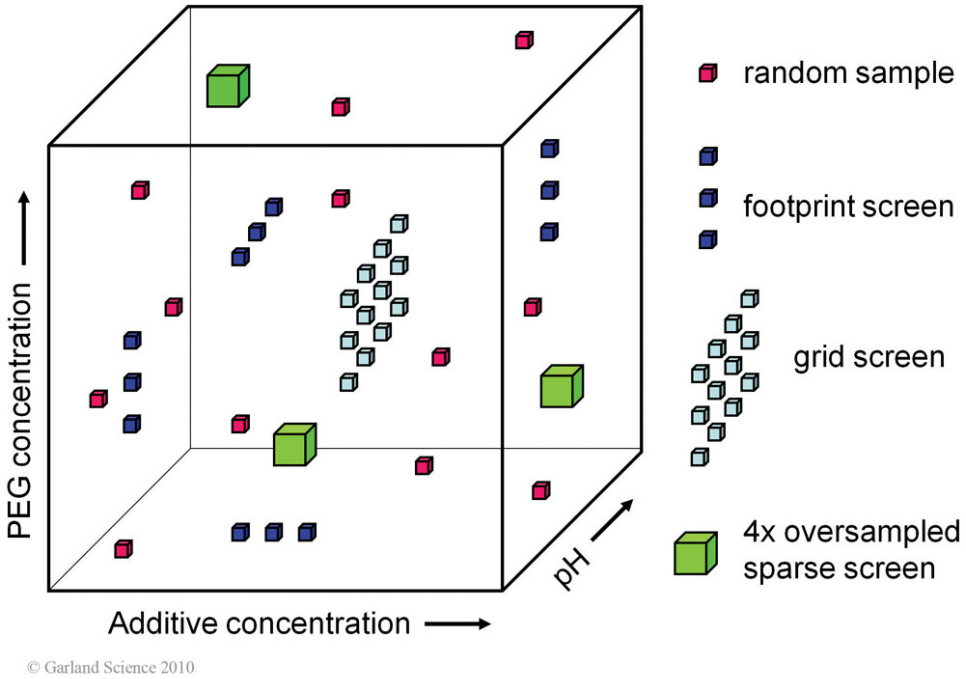
In order to form crystals, the protein solution must become supersaturated. In the supersaturated, thermodinamically metastable state, nucleation can occur and crystals may form while the solution equilibrates.

The most common technique for protein crystal growth is by vapor diffusion, where water vapor equilibrates from a drop containing protein and a precipitant into a larger reservoir with higher precipitant concentration.

Crystal formation is an **inherently unlikely process**, and many trials may be necessary to obtain well diffracting crystals.

Crystallization as a multivariate sampling/optimization problem

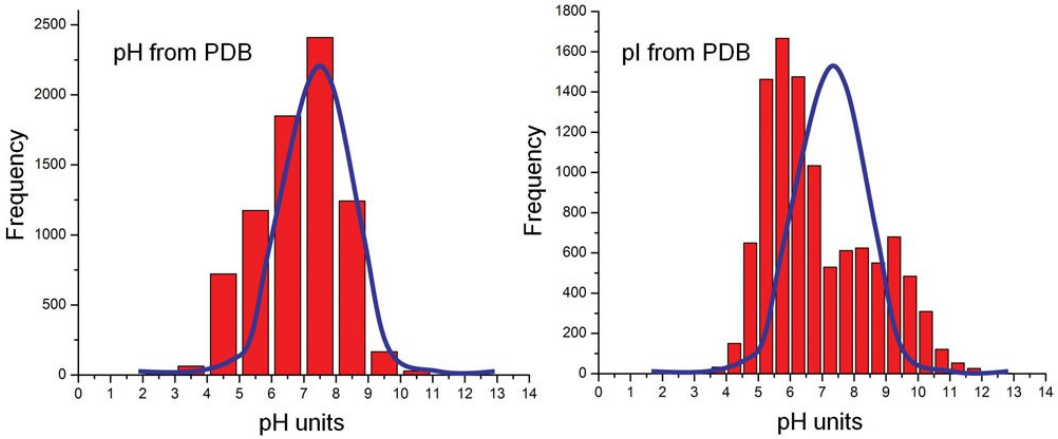
- Factorial designs:full & partial
- Sparse matrix sampling
- Random sampling
- Grid screens
- Footprint screen



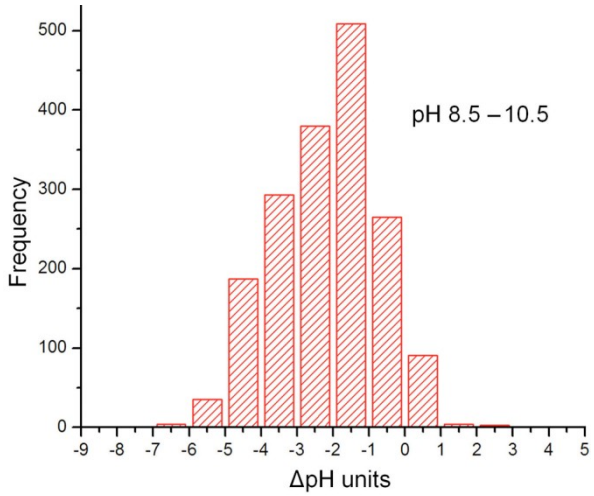
Random sampling and other large-scale trials have shown that if no promising results are obtained after about 300 trials it is likely that the protein is a difficult case for crystallization.
Consider other proteins constructs, orthologs, or protein engineering.
Accept that the chance of obtaining diffracting crystals of a protein whitout any additional procedural adjstments or protein modifications is only 10-20%.
Whether crystallization will succeed or not is already predetermined by the protein construct itself.

If the protein cannot crystallize, it will not, no matter how many crystallization trials are performed!!!

Distribution of crystallization pH and distribution of proteins

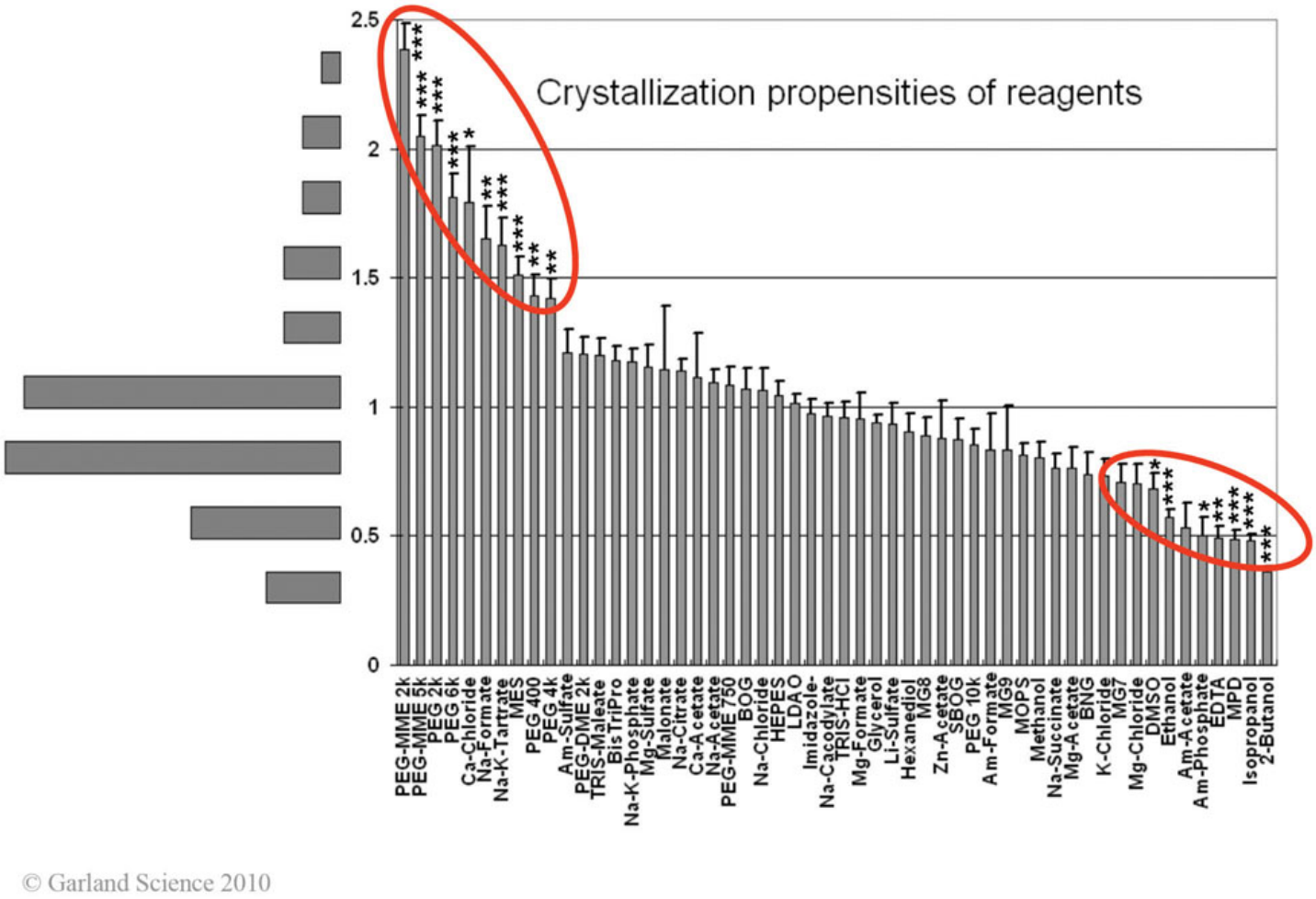


Discrete crystallization pH frequency distribution for pI between 8.5 and 10.5



The optimal pH is ca. 1.5 pH units below the pI.

Crystallization propensities for 50 common reagents



ca. 500.000 experiments, but soluble, bacterially expressed prokaryotic proteins!

Shortcoming: only 50 reagents have been analyzed and synergistic effects are not considered!

Crystal Screen

Reagent Formulation

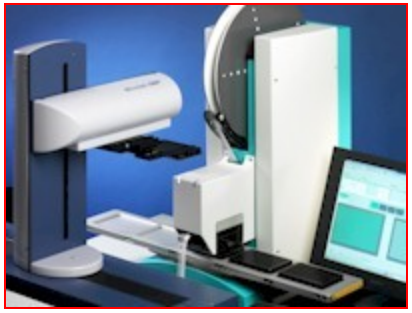
Tube Number	Salt	Buffer †	Precipitant
1.	0.02 M Calcium Chloride dihydrate	0.1 M Sodium Acetate trihydrate pH 4.6	30% v/v 2-Methyl-2,4-pentanediol
2.	None	None	0.4 M Potassium Sodium Tartrate tetrahydrate
3.	None	None	0.4 M mono-Ammonium dihydrogen Phosphate
4.	None	0.1 M Tris Hydrochloride pH 8.5	2.0 M Ammonium Sulfate
5.	0.2 M tri-Sodium Citrate dihydrate	0.1 M HEPES - Na pH 7.5	30% v/v 2-Methyl-2,4-pentanediol
6.	0.2 M Magnesium Chloride hexahydrate	0.1 M Tris Hydrochloride pH 8.5	30% w/v Polyethylene Glycol 4000
7.	None	0.1 M Sodium Cacodylate pH 6.5	1.4 M Sodium Acetate trihydrate
8.	0.2 M tri-Sodium Citrate dihydrate	0.1 M Sodium Cacodylate pH 6.5	30% v/v iso-Propanol
9.	0.2 M Ammonium Acetate	0.1 M tri-Sodium Citrate dihydrate pH 5.6	30% w/v Polyethylene Glycol 4000
10.	0.2 M Ammonium Acetate	0.1 M Sodium Acetate trihydrate pH 4.6	30% w/v Polyethylene Glycol 4000
11.	None	0.1 M tri-Sodium Citrate dihydrate pH 5.6	1.0 M mono-Ammonium dihydrogen Phosphate
12.	0.2 M Magnesium Chloride hexahydrate	0.1 M HEPES - Na pH 7.5	30% v/v iso-Propanol
13.	0.2 M tri-Sodium Citrate dihydrate	0.1 M Tris Hydrochloride pH 8.5	30% v/v Polyethylene Glycol 400
14.	0.2 M Calcium Chloride dihydrate	0.1 M HEPES - Na pH 7.5	28% v/v Polyethylene Glycol 400
15.	0.2 M Ammonium Sulfate	0.1 M Sodium Cacodylate pH 6.5	30% w/v Polyethylene Glycol 8000
16.	None	0.1 M HEPES - Na pH 7.5	1.5 M Lithium Sulfate monohydrate
17.	0.2 M Lithium Sulfate monohydrate	0.1 M Tris Hydrochloride pH 8.5	30% Polyethylene Glycol 4000
18.	0.2 M Magnesium Acetate tetrahydrate	0.1 M Sodium Cacodylate pH 6.5	20% Polyethylene Glycol 8000
19.	0.2 M Ammonium Acetate	0.1 M Tris Hydrochloride pH 8.5	30% v/v iso-Propanol
20.	0.2 M Ammonium Sulfate	0.1 M Sodium Acetate trihydrate pH 4.6	25% w/v Polyethylene Glycol 4000
21.	0.2 M Magnesium Acetate tetrahydrate	0.1 M Sodium Cacodylate pH 6.5	30% v/v 2-Methyl-2,4-pentanediol
22.	0.2 M Sodium Acetate trihydrate	0.1 M Tris Hydrochloride pH 8.5	30% w/v Polyethylene Glycol 4000
23.	0.2 M Magnesium chloride hexahydrate	0.1 M HEPES - Na pH 7.5	30% v/v Polyethylene Glycol 400
24.	0.2 M Calcium Chloride dihydrate	0.1 M Sodium Acetate trihydrate pH 4.6	20% v/v iso-Propanol
25.	None	0.1 M Imidazole pH 6.5	1.0 M Sodium Acetate trihydrate
26.	0.2 M Ammonium Acetate	0.1 M tri-Sodium Citrate dihydrate pH 5.6	30 % v/v 2-Methyl-2,4-pentanediol
27.	0.2 M tri-Sodium Citrate dihydrate	0.1 M HEPES - Na pH 7.5	20% v/v iso-Propanol
28.	0.2 M Sodium Acetate trihydrate	0.1 M Sodium Cacodylate pH 6.5	30% w/v Polyethylene Glycol 8000
29.	None	0.1 M HEPES - Na pH 7.5	0.8 M Potassium Sodium Tartrate tetrahydrate
30.	0.2 M Ammonium Sulfate	None	30% w/v Polyethylene Glycol 8000

Humans (PhDs, PostDocs: Low Throughput Crystallization

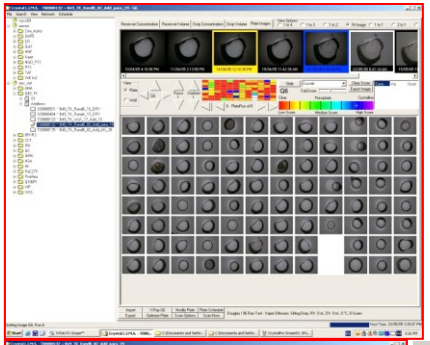
1-5 µL scale

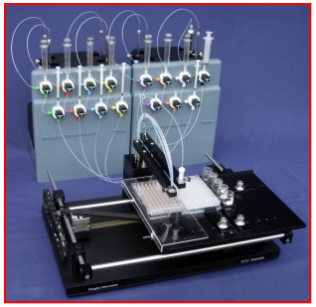
Robotics: High Throughput Crystallization

TECAN

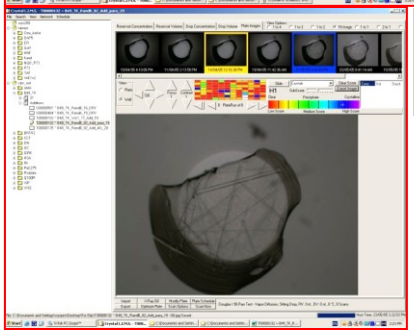


TTP LABTECH





Fluidigm TOPAZ



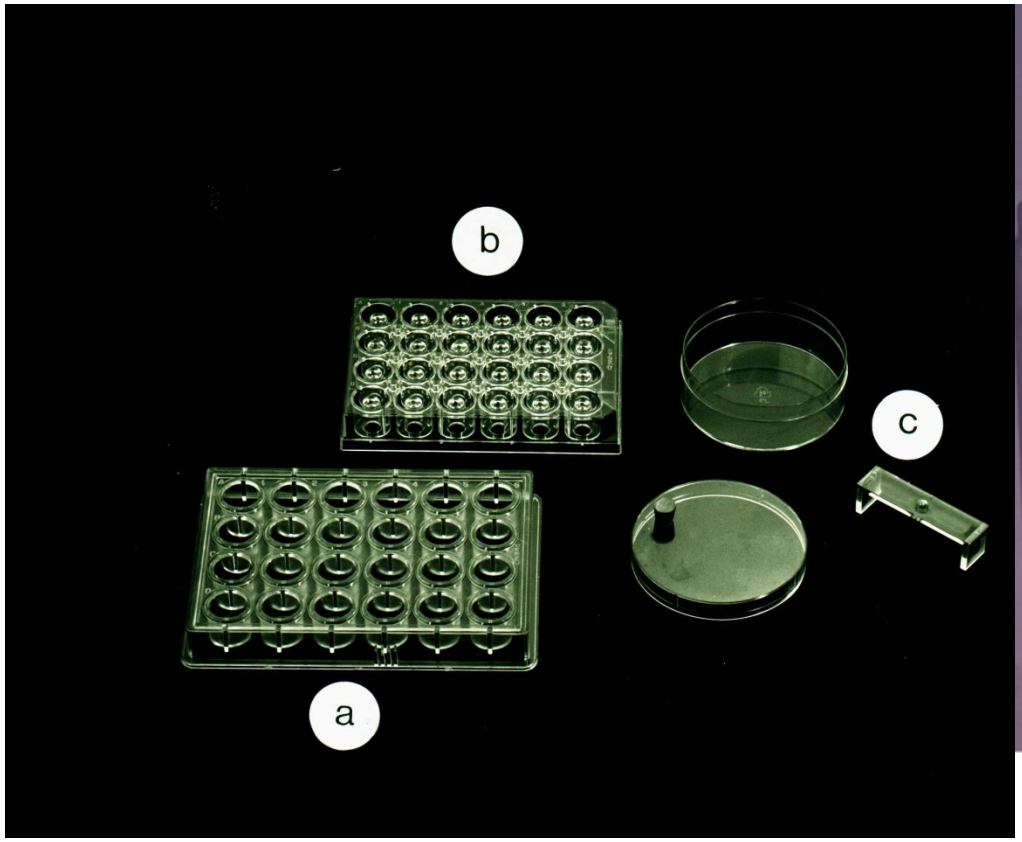
TriTek

Douglas Instruments
Success in protein crystallization

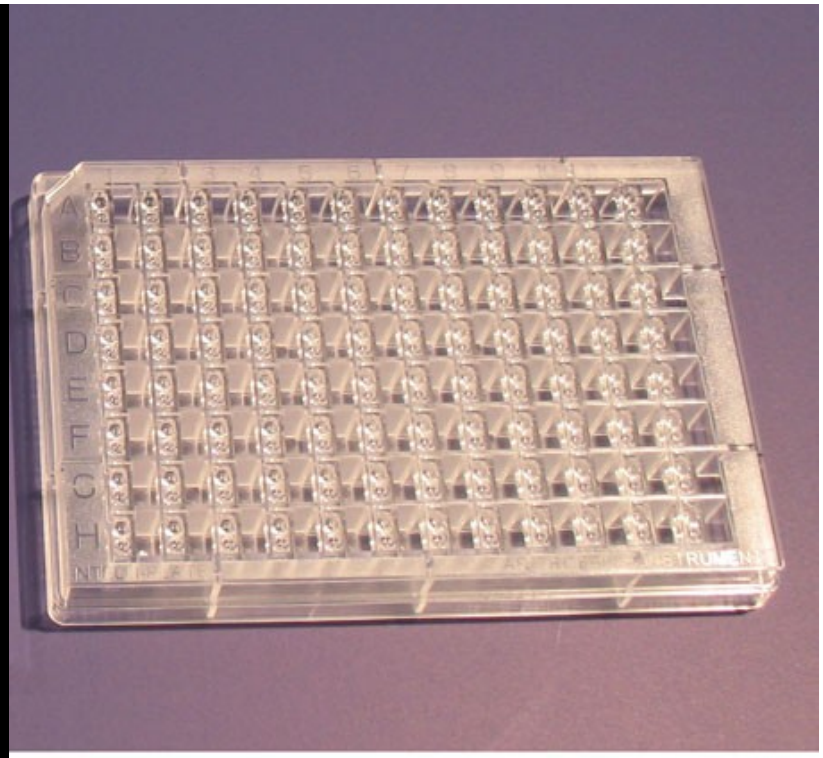
DISCOVERY PARTNERS INTERNATIONAL

10-100 nL scale

Crystallization plates: vapour diffusion



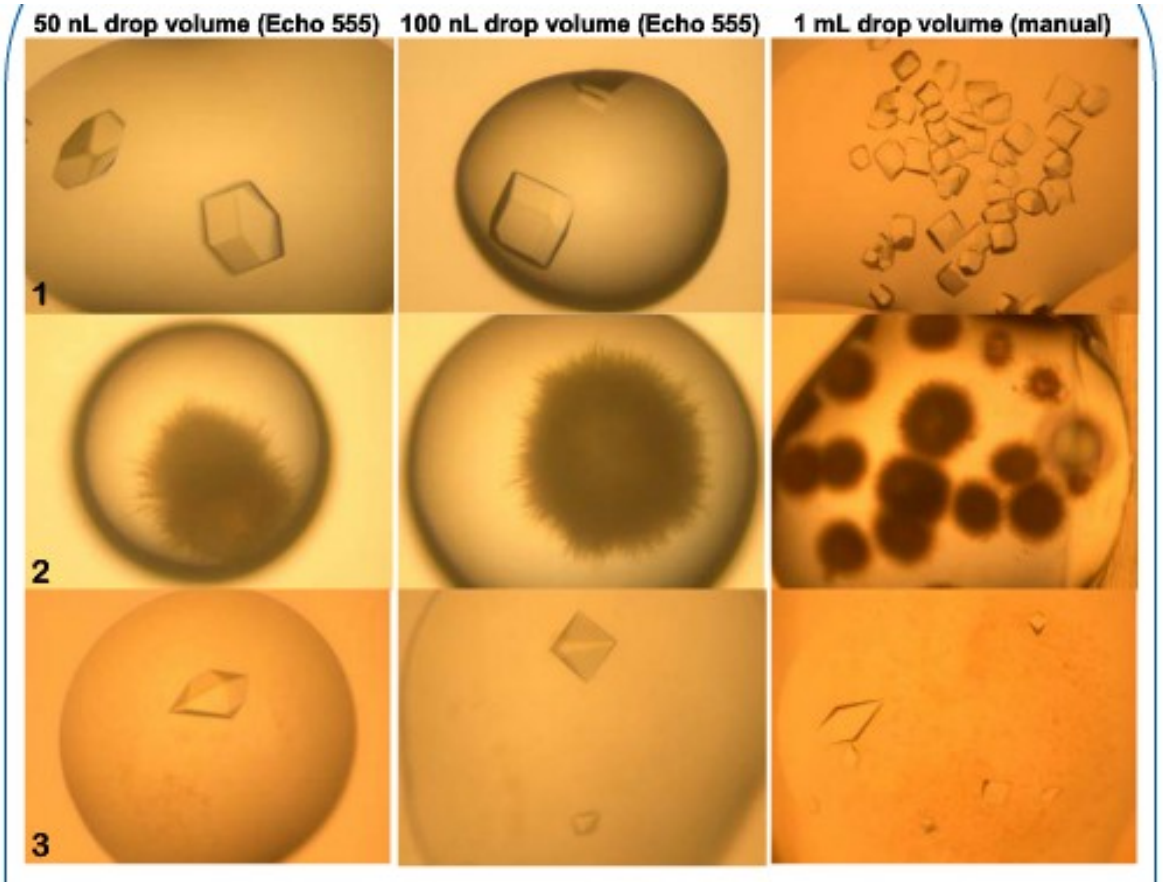
Linbro platesoriginally designed for cell culture



Intelli-plate:96-well sitting-drop plate

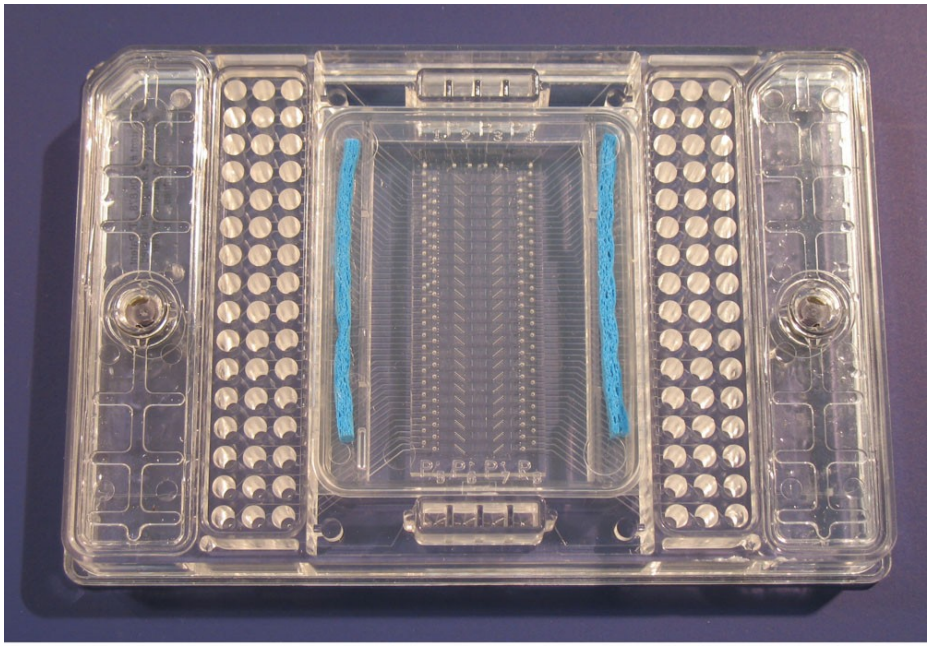
Robotics: High Throughput Crystallization - State-of-the-art

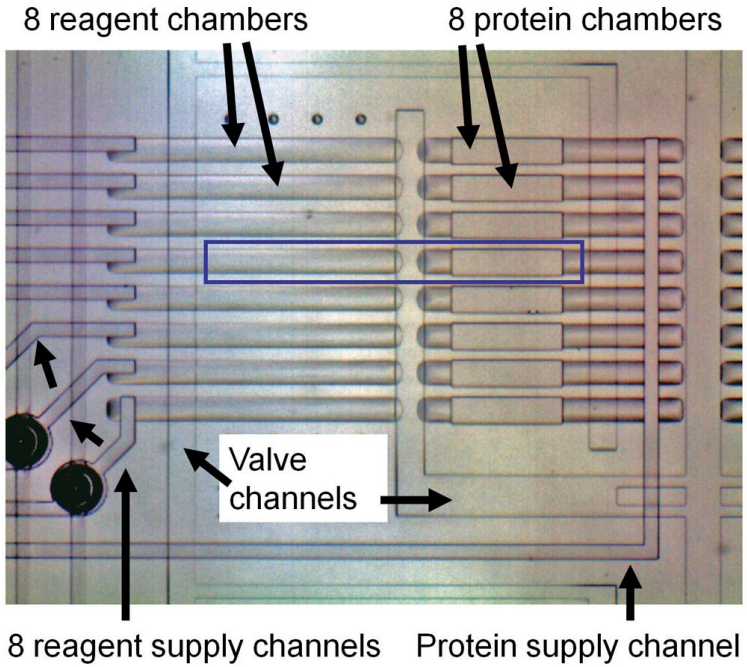
LABCYTE™



Echo-nano dispensing: **down to 2.5 nL**!!!

Robotics: High Throughput Crystallization - State-of-the-art



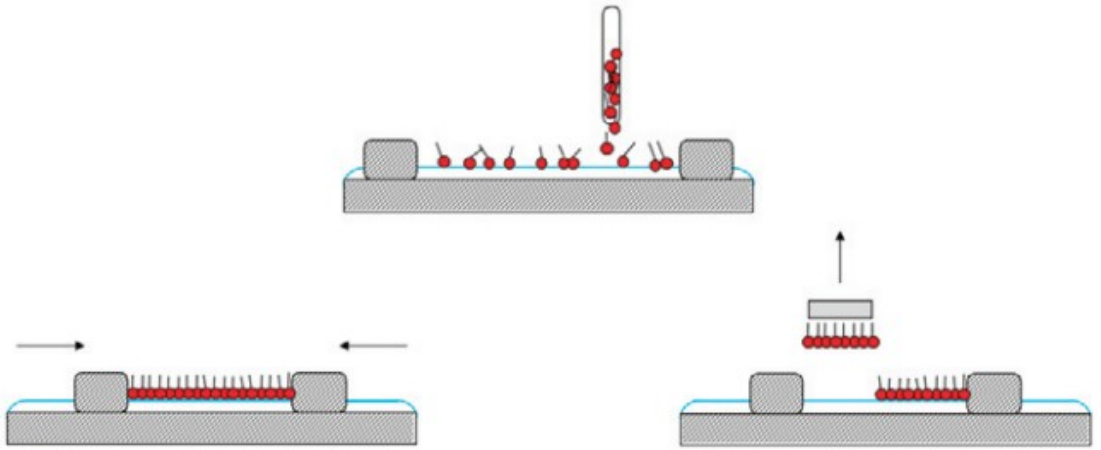


Free-interface diffusion chambers

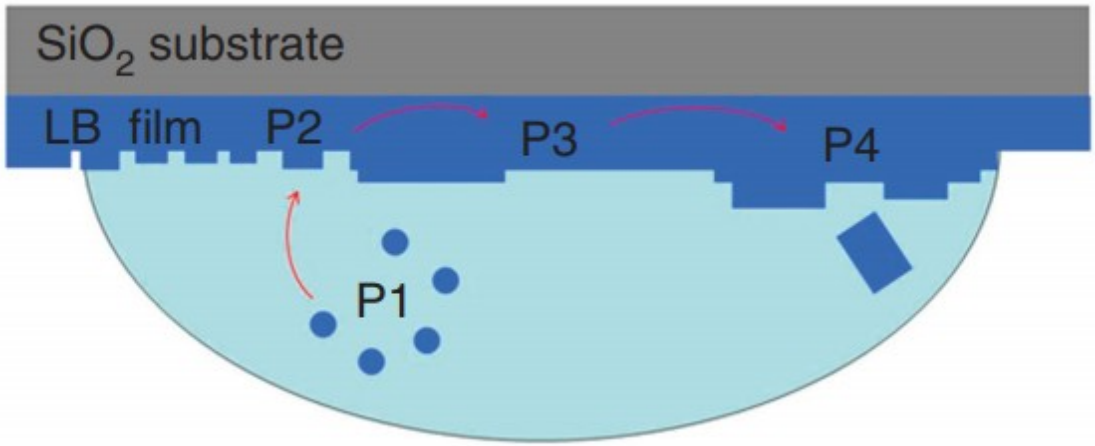
A microfluidic free-interface screening chip

Langmuir–Blodgett nanotemplates for protein crystallography

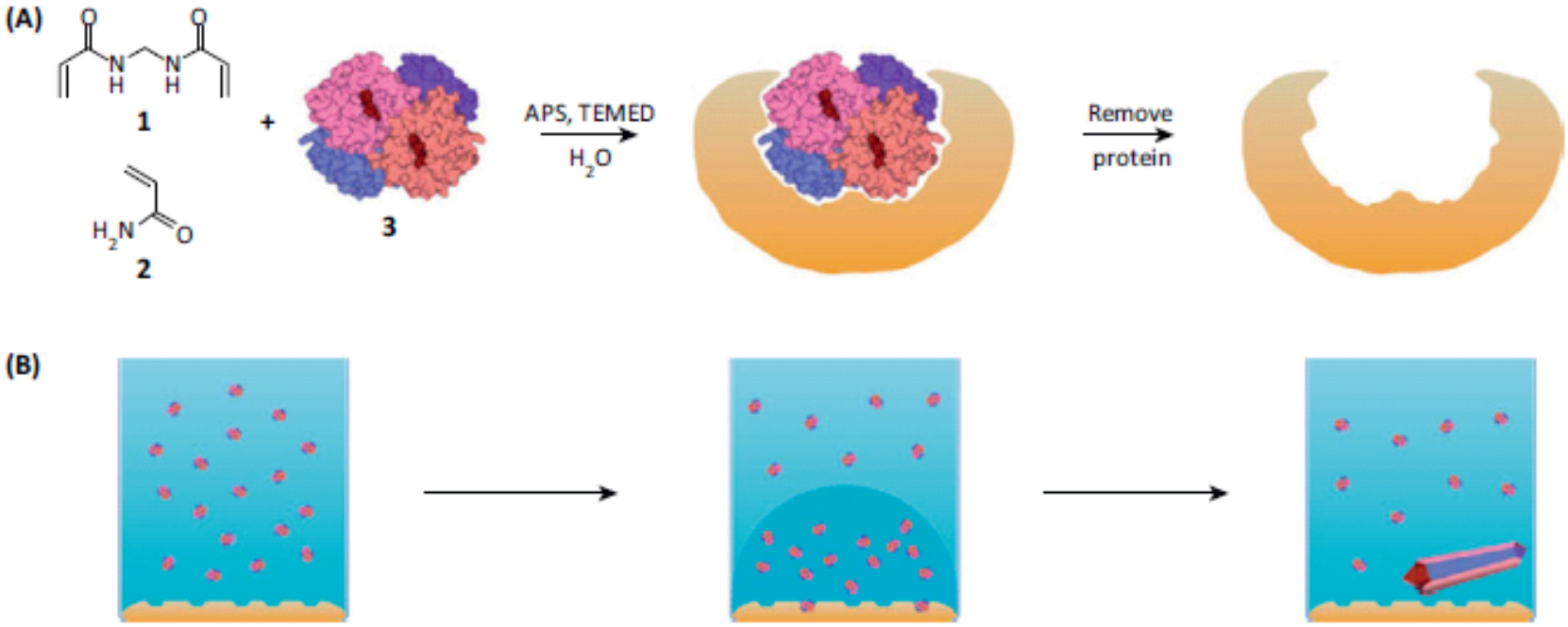
Nature Protocols (2017) **12**:2570–2589



Formation of Languimir Blodgett protein nanotemplate

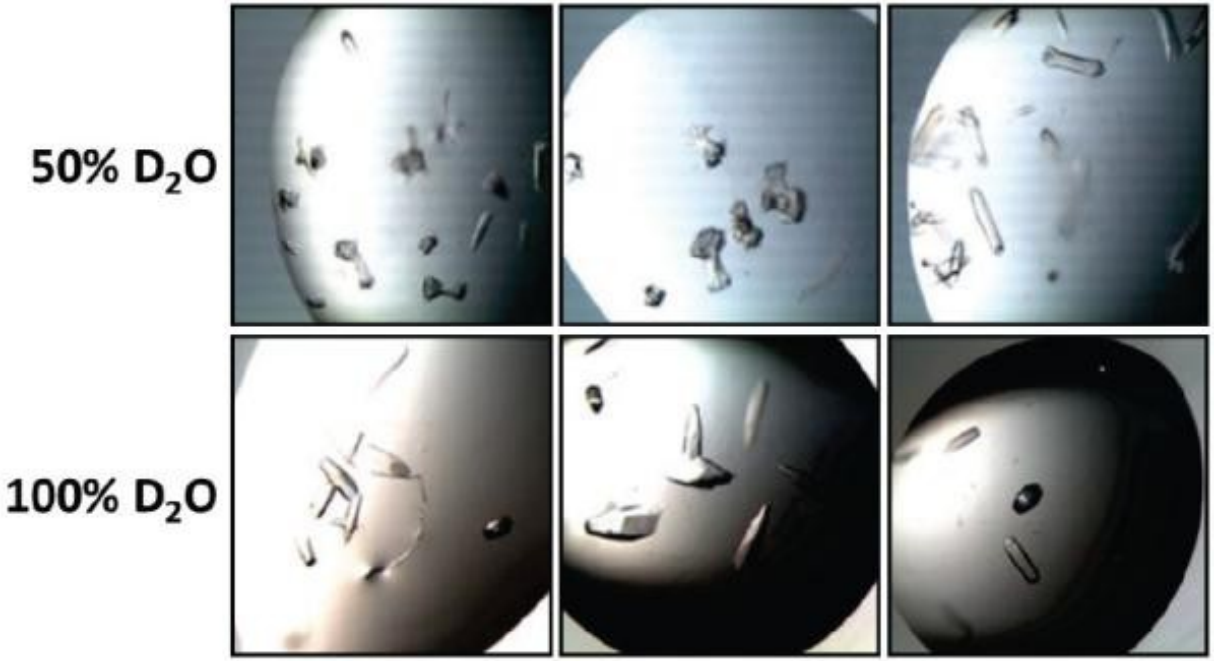


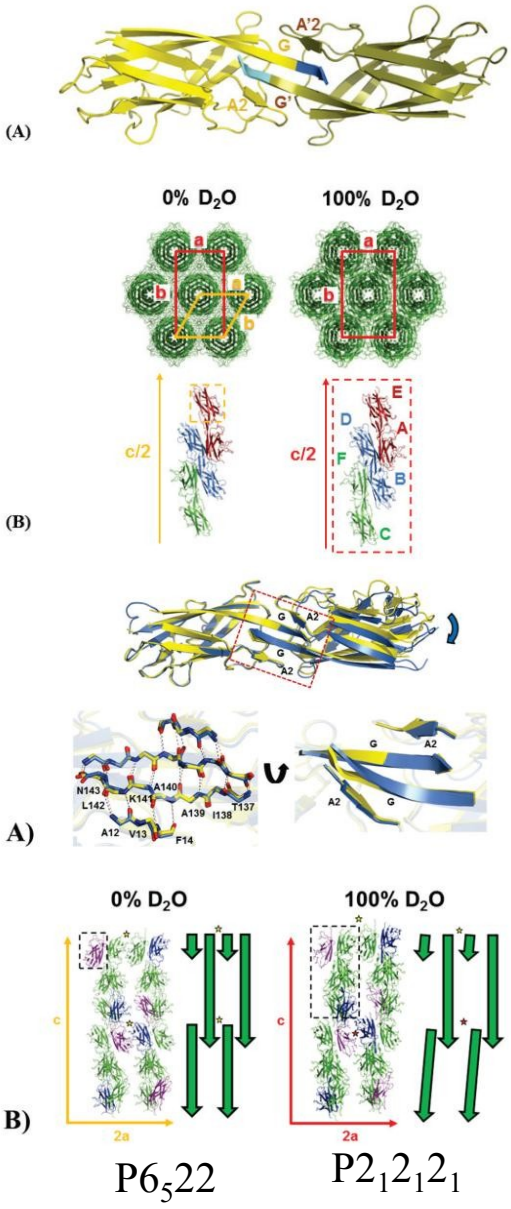
Protein crystallization facilitated by molecularly imprinted polymers



An imprinted hydrogel is formed when acrylamide (2) is polymerized with a crosslinker (methylene-bis-acrylamide,(1) in the presence of a protein template.
Polymerisation in water with ammonium persulfate (APS) and tetramethylethylenediamine (TEMED) gives rise to the hydrogel with embedded protein molecules.
Removal of the template protein results in the formation of the imprints, cavities complementary in size and shape to the protein templates. [E.Saridakis, N. E. Chayen (2013) Trends in Biotech. **31**, 515]

Protein Crystallization in Deuterated Water

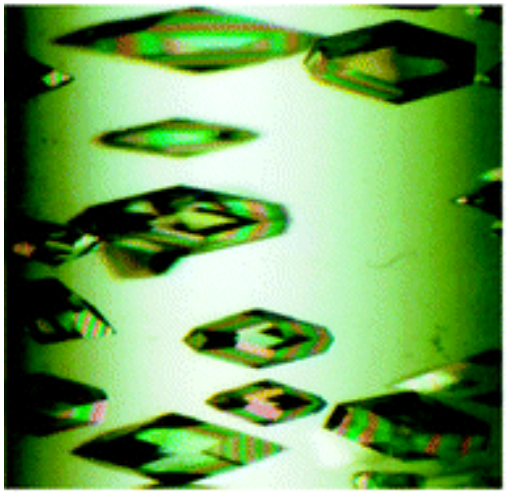


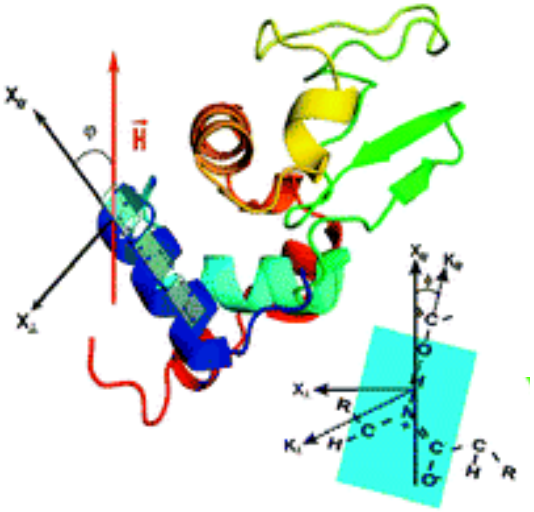


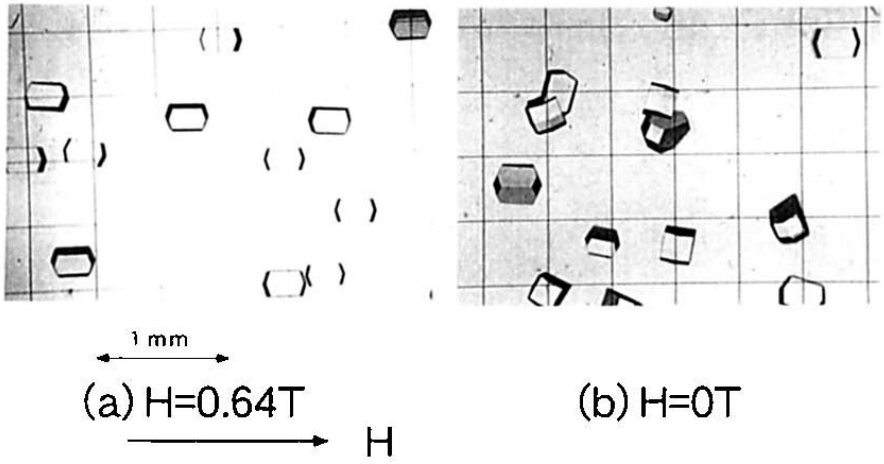
***Salmonella enteritidis* fimbriae 14 pilin SefD**

Biochem Biophys Res Commun. **2012** *421:* 208–213

Crystallization under magnetic field



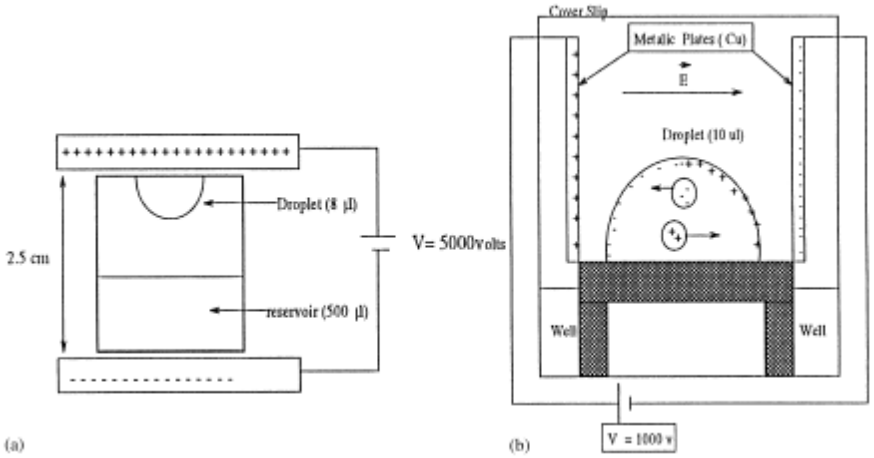




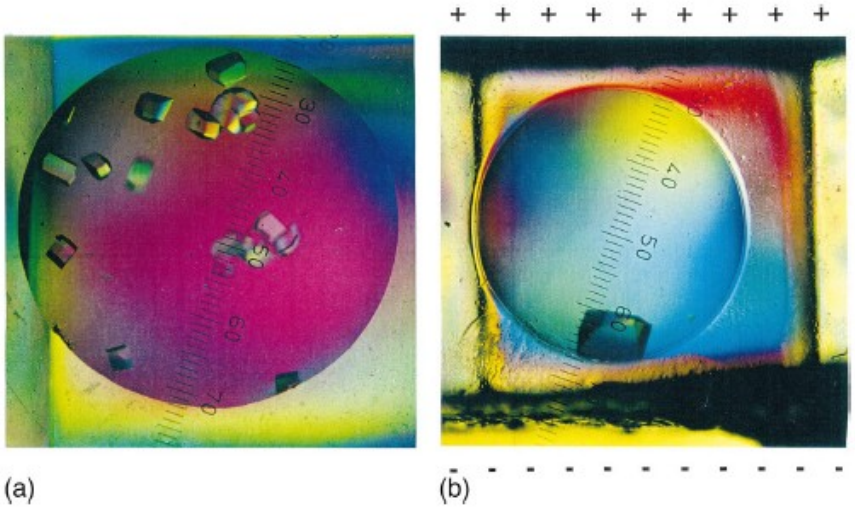
Lysozyme

Cryst. Growth Des., **2009**, *9* : 2610–2615

Crystallization under electric field



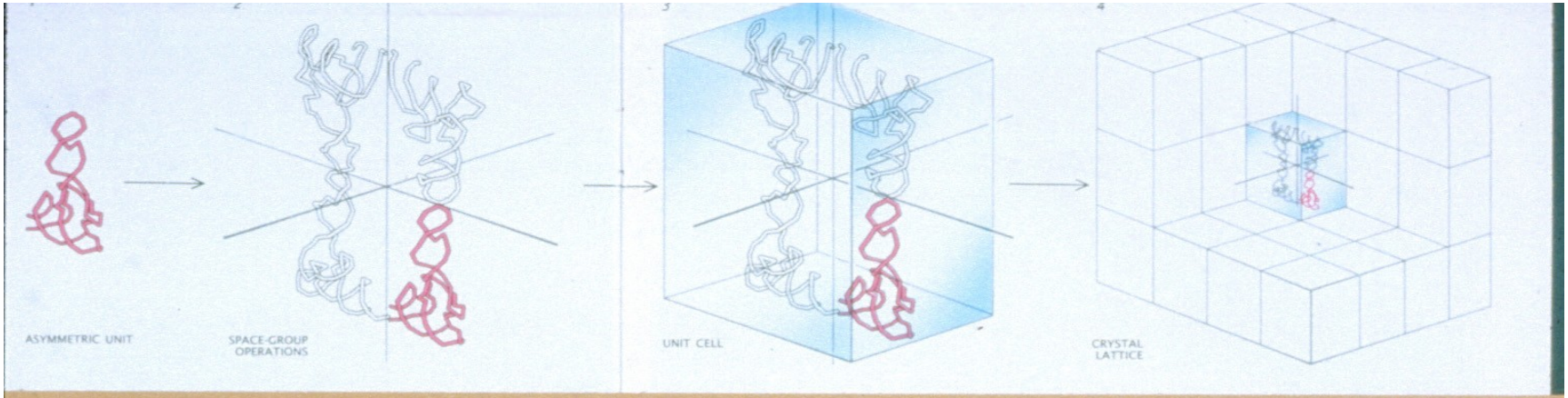
The two crystallization devices used for crystallization under an external electric field:
(a) the whole Linbro plate is placed between two metal plates connected to a power supply,
(b) the droplet is directly placed between two metal plates connected to a power supply.



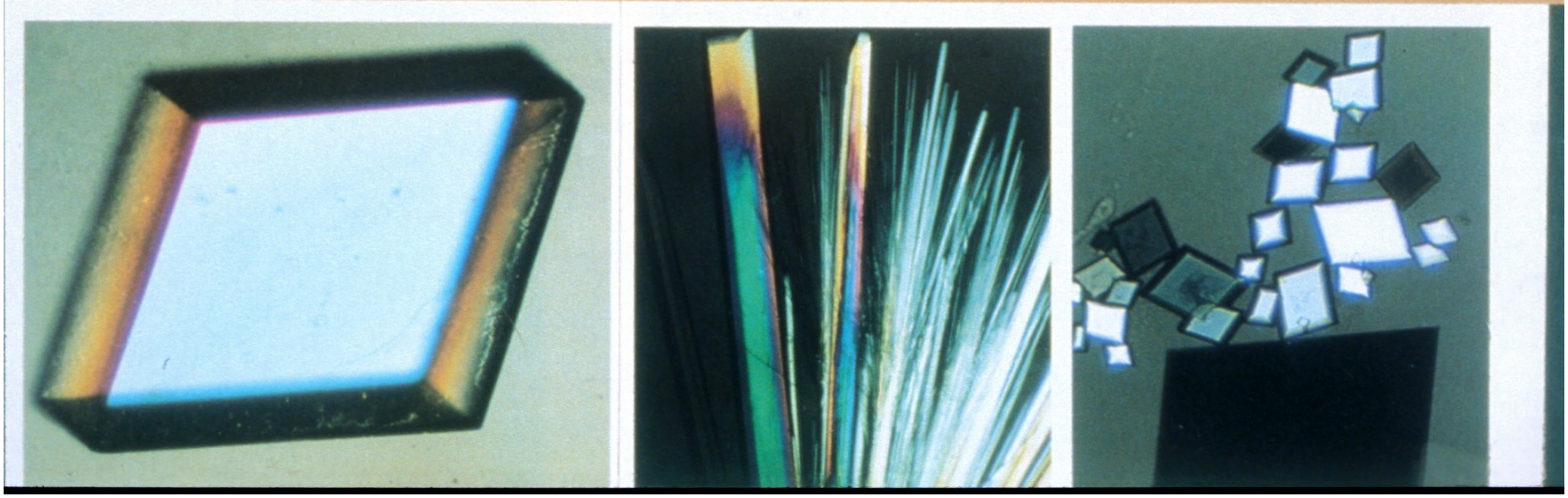
Crystals formation in the drop:
(a), without electric field many crystals grow randomly inside the droplet and
(b) under electric field, and in the same physico-chemical conditions. Crystallization then occurred near the cathode due to the positive charge of lysozyme at the pH of the experiment, with fewer but larger crystals under the electric field

Progress in Biophysics and Molecular Biology (2009) 101:38-44

Basic principles of X-ray crystallography

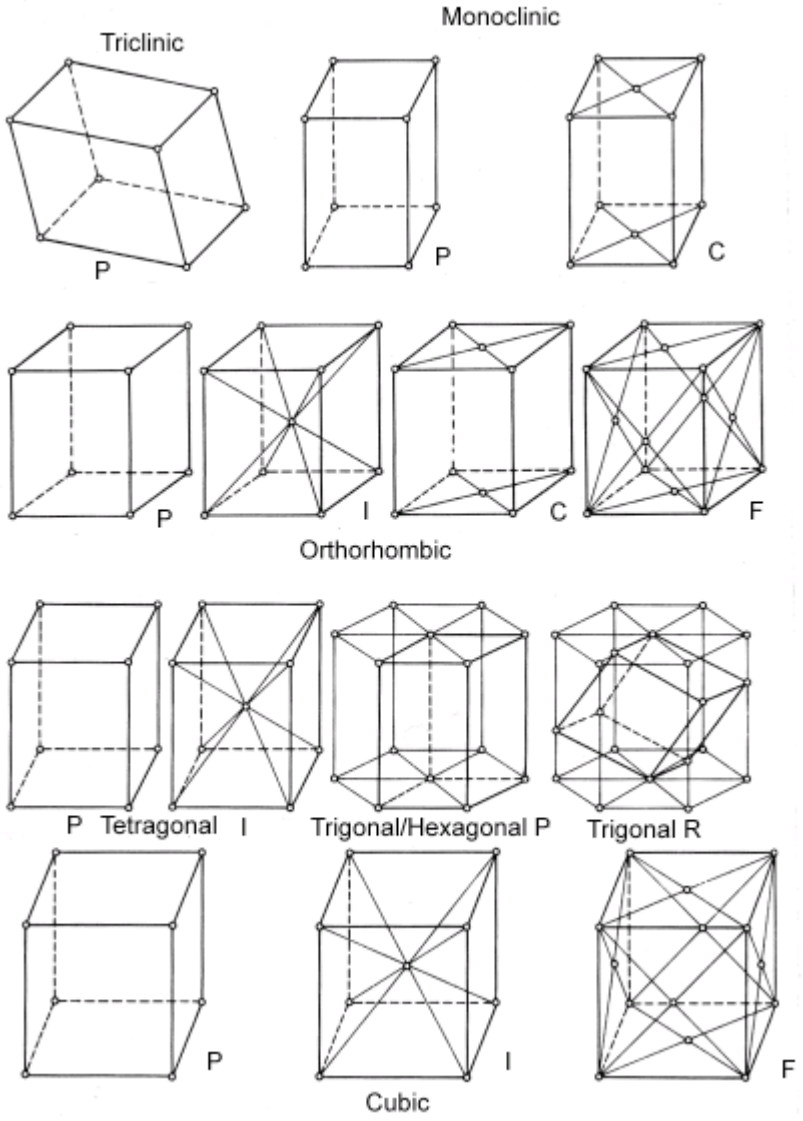


Unit Cell and Crystals



14 Bravais lattices

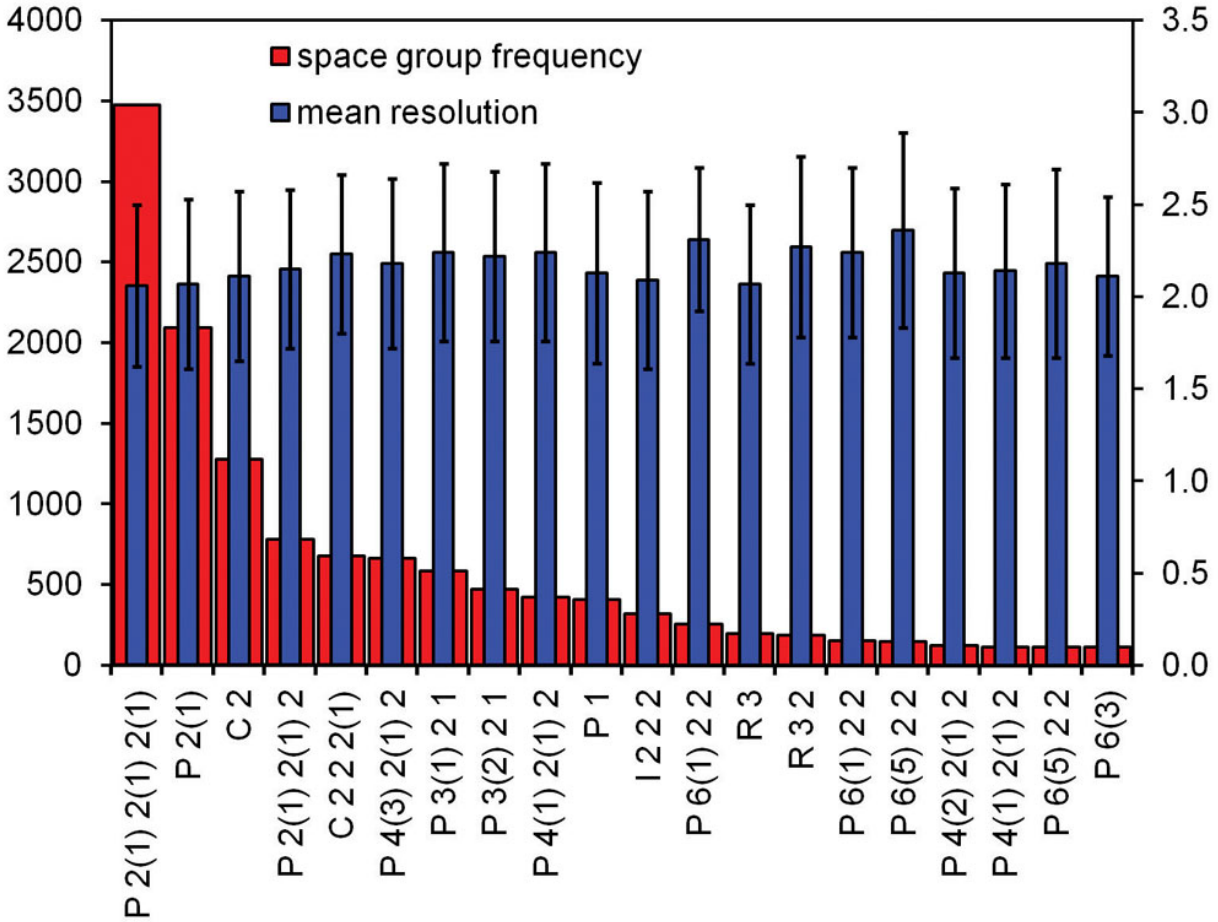
Type of symmetry element	Written symbol	Graphical symbol
Centre of symmetry	$\bar{1}$	o
		Perpendicular to paper / In plane of paper
Mirror plane	m	
Glide planes	$a\ b\ c$	glide in plane of paper / arrow shows glide direction
		..... glide out of plane of paper
	n	
Rotation axes	2	
	3	
	4	
	6	
Screw axes	2_1	
	$3_1, 3_2$	
	$4_1, 4_2, 4_3$	
	$6_1, 6_2, 6_3, 6_4, 6_5$	
Inversion axes	$\bar{3}$	
	$\bar{4}$	
	$\bar{6}$	



The 65 Chiral Space Groups

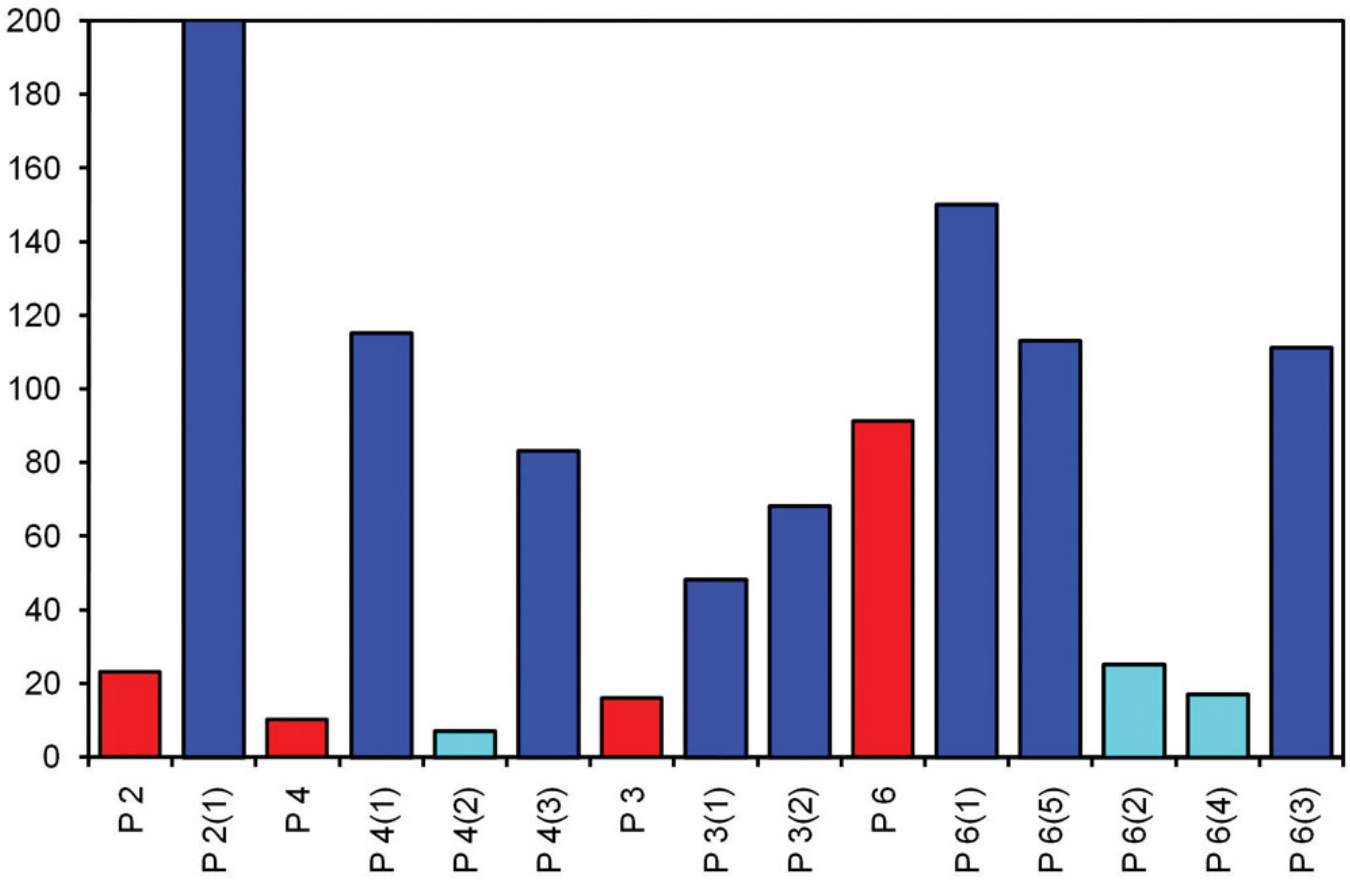
TRICLINIC				
P 1				
MONOCLINIC				
P 2	P 21	C 2		
ORTHORHOMBIC				
P 2 2 2	P 2 2 21	P 21 21 2	P 21 21 21	C 2 2 21
C 2 2 2	F 2 2 2	I 2 2 2	I 21 21 21	
TETRAGONAL				
P 4	P 41	P 42	P 43	I 4
I 41	P 4 2 2			
P 4 21 2	P 41 2 2	P 41 21 2	P 42 2 2	P 42 21 2
P 43 2 2	P 43 21 2	I 4 2 2	I 41 2 2	
TRIGONAL				
P 3	P 31	P 32	R 3	
P 3 1 2	P 3 2 1	P 31 1 2	P 31 2 1	
P 32 1 2	P 32 2 1	R 3 2		
HEXAGONAL				
P 6	P 61	P 65	P 62	P 64
P 63	P 6 2 2			
P 61 2 2	P 65 2 2	P 62 2 2	P 64 2 2	P 63 2 2
CUBIC				
P 2 3	F 2 3	I 2 3	P 21 3	I 21 3
P 4 3 2	P 42 3 2	F 4 3 2		
F 41 3 2	I 4 3 2	P 43 3 2	P 41 3 2	I 41 3 2

Space Groups Preferences – Protein Crystals



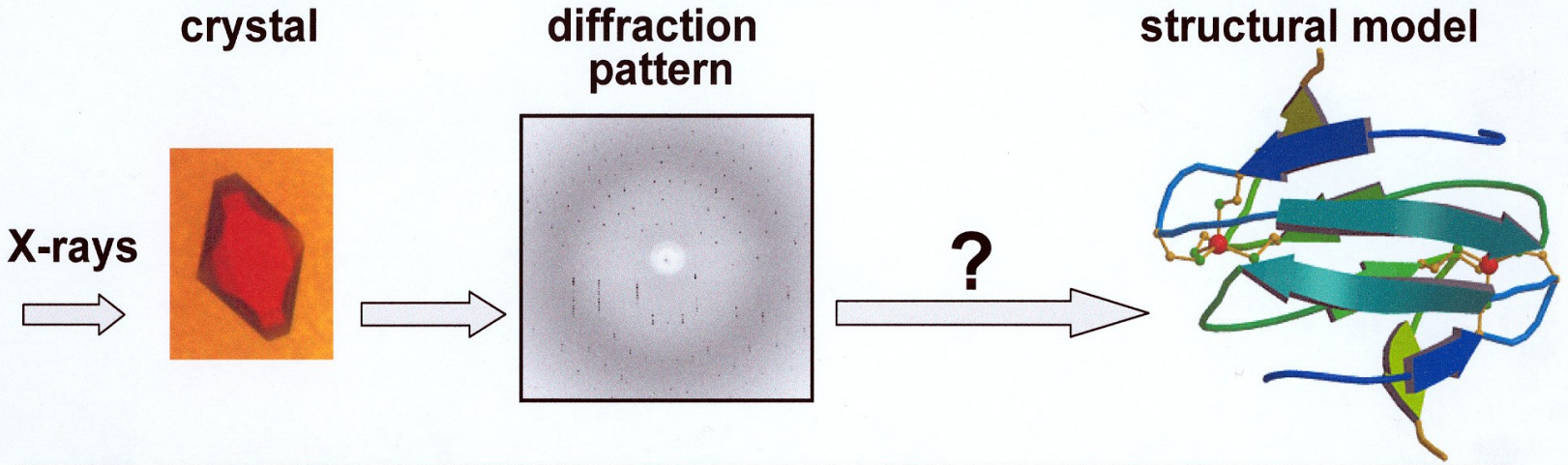
Nearly 90% of all structures crystallize in one of these 20 space groups.
Note that no correlation exists between the observed maximum resolution and the frequency of a space group.

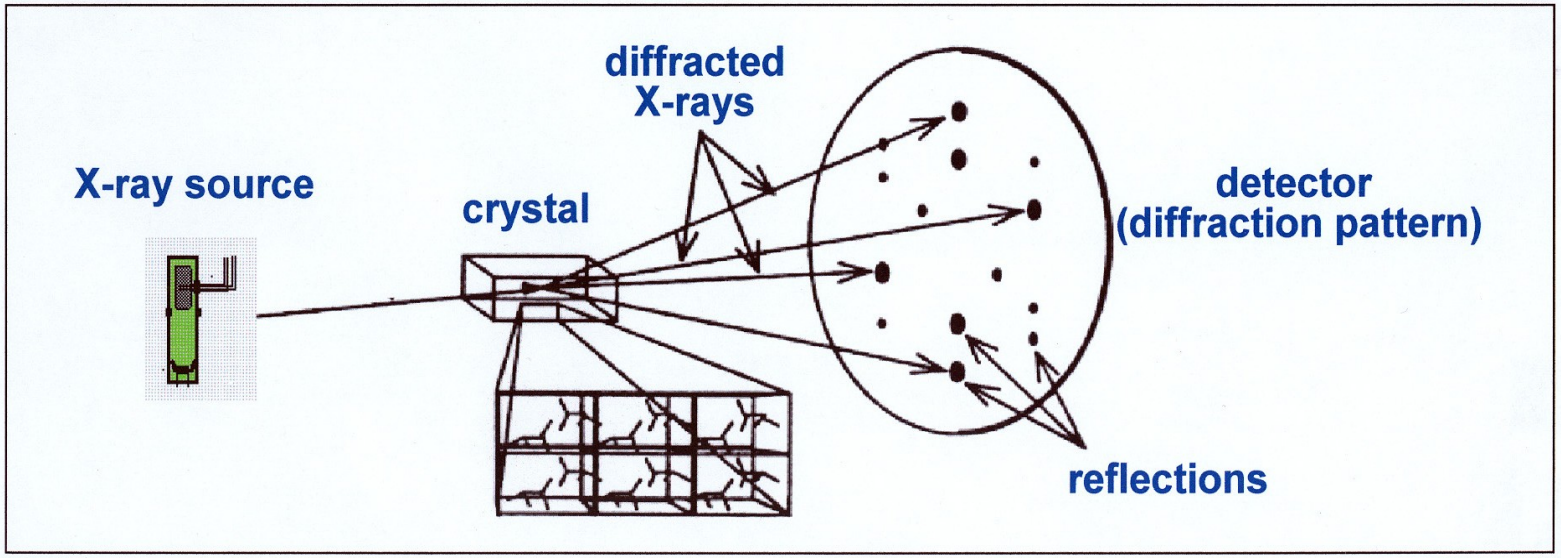
Space Groups Preferences – Protein Crystals



Plain rotation axes *vs* screw axis

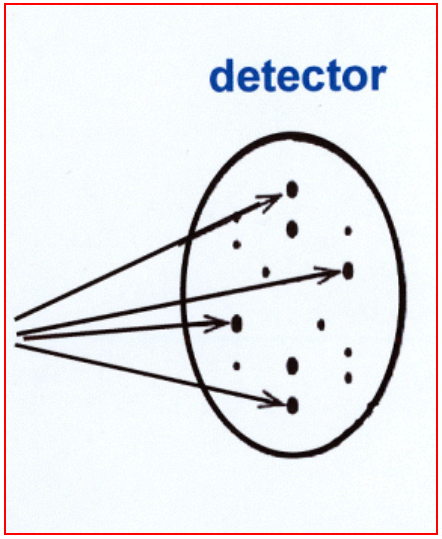
Basic principles of X-ray crystallography

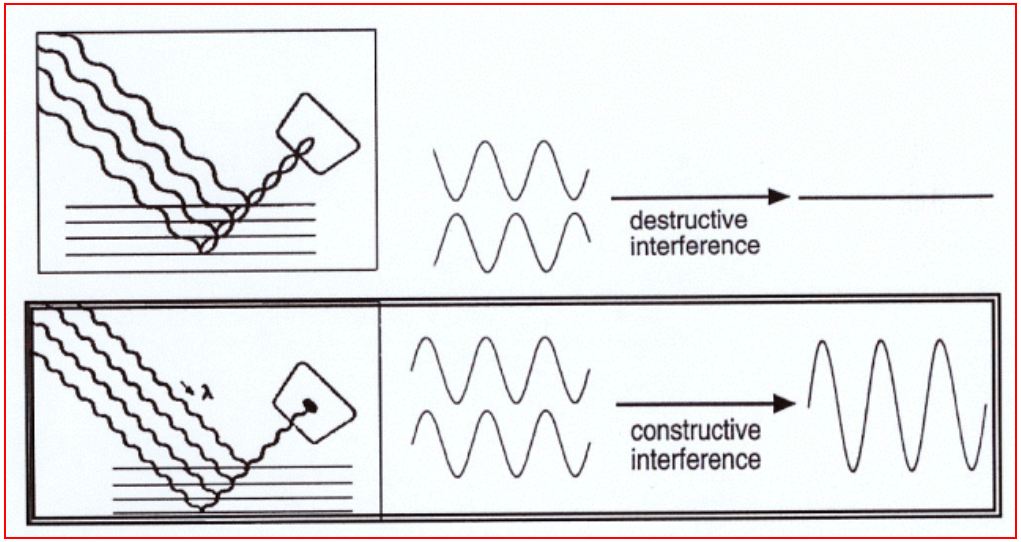


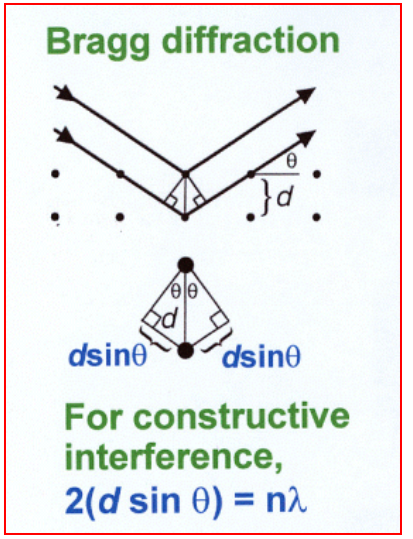


Basic principles of X-ray crystallography

X-ray diffraction



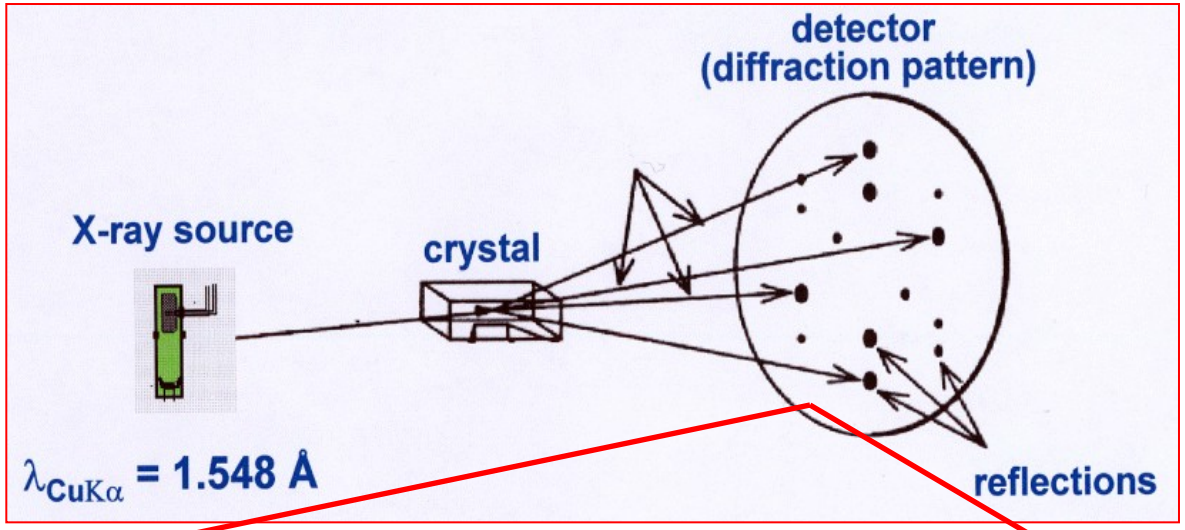




What When Why

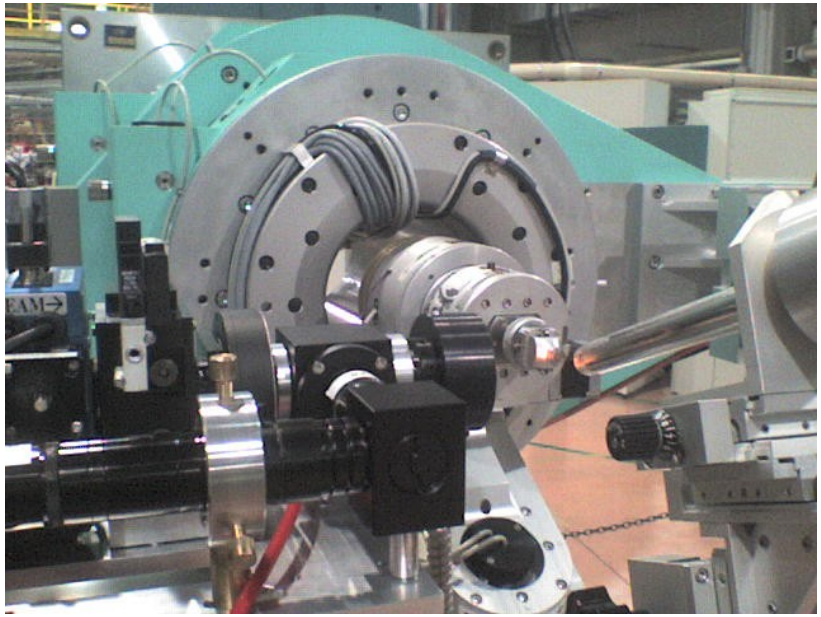
Basic principles of X-ray crystallography

X-ray diffraction



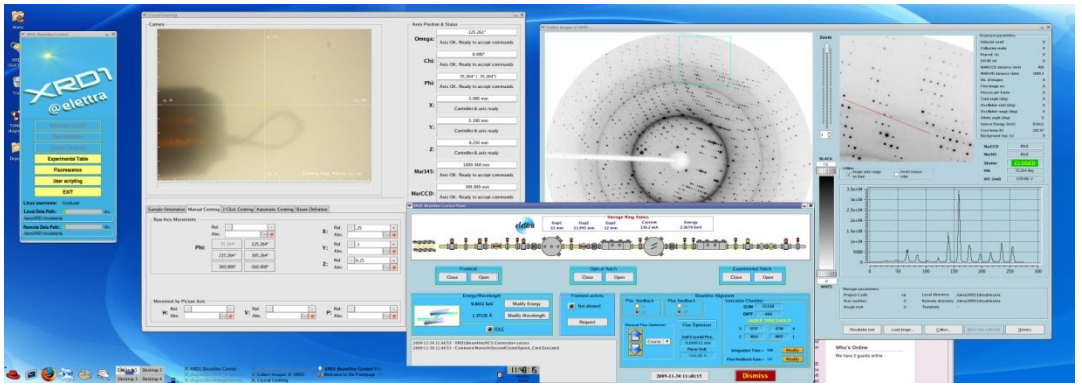
Basic principles of X-ray crystallography

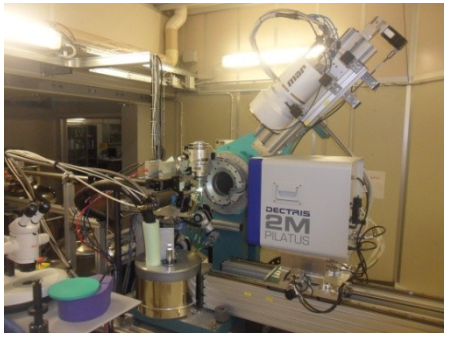
X-ray diffraction

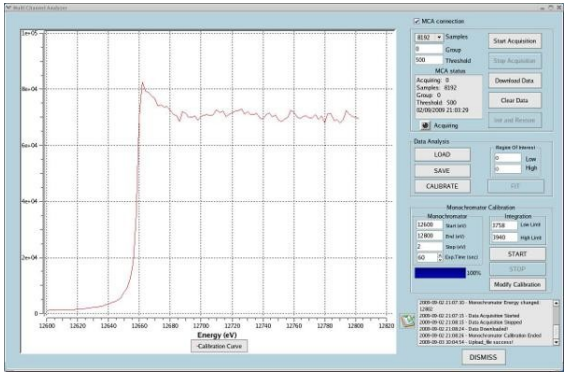


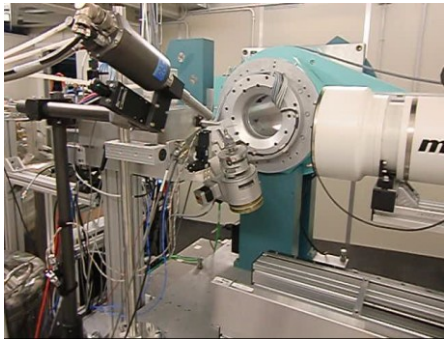
Experimental hutch @ XRD-1 ELETTRA,Trieste

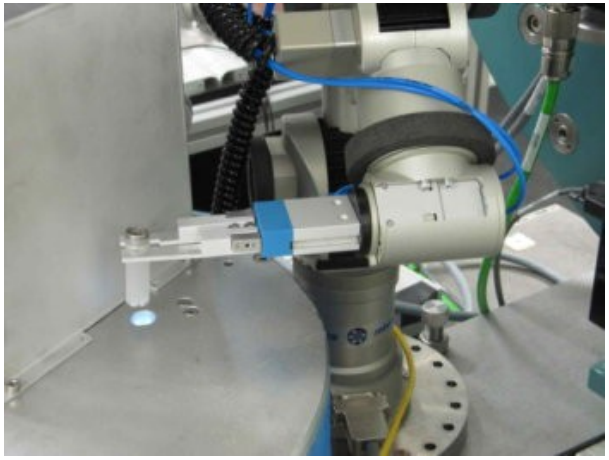
Experimental hutch @ XRD-1 ELETTRA,Trieste

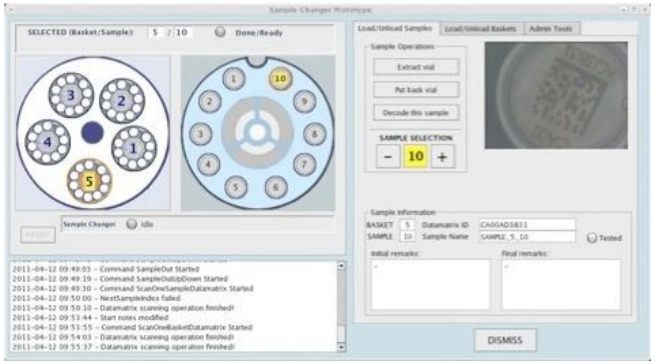






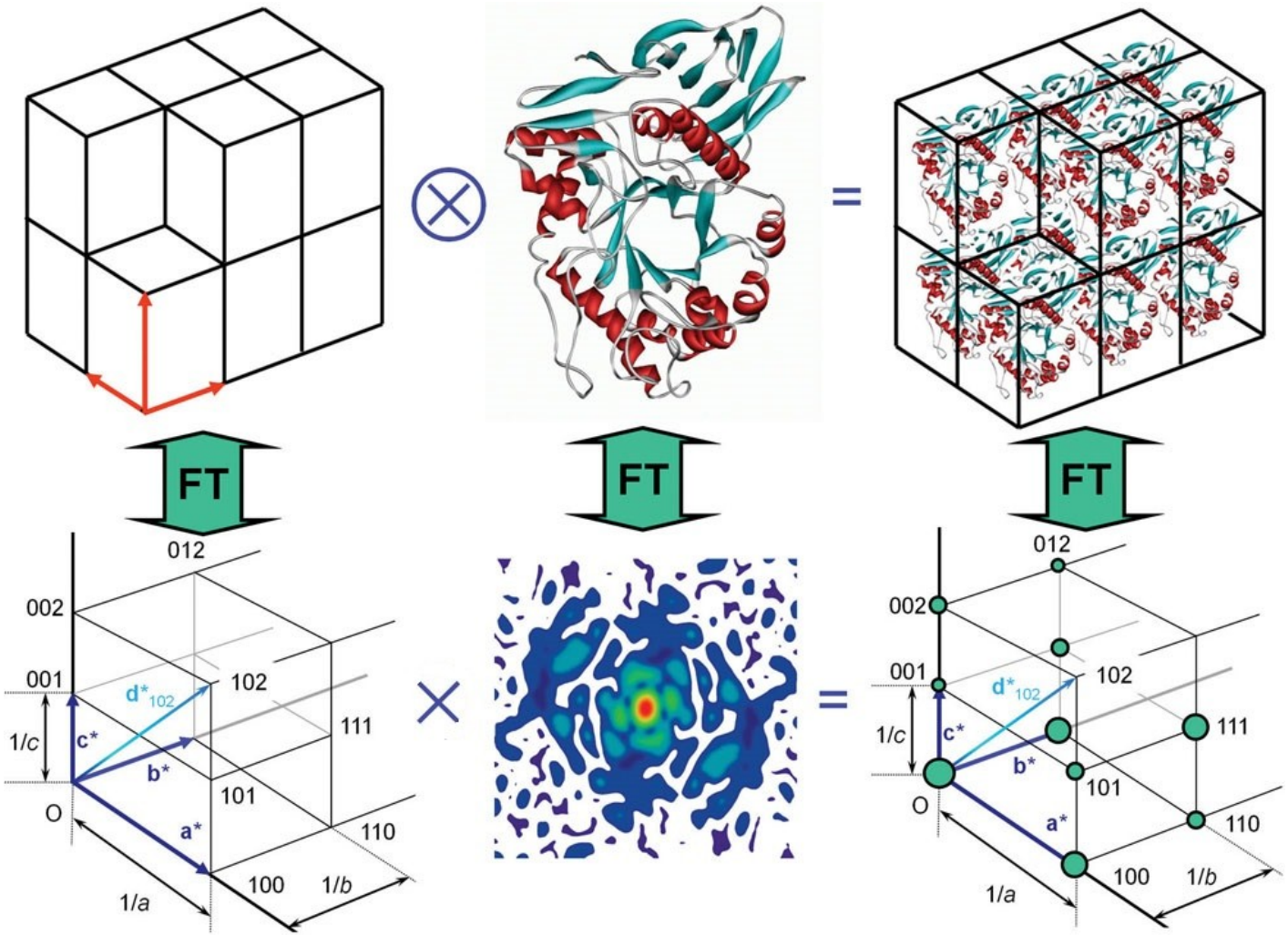






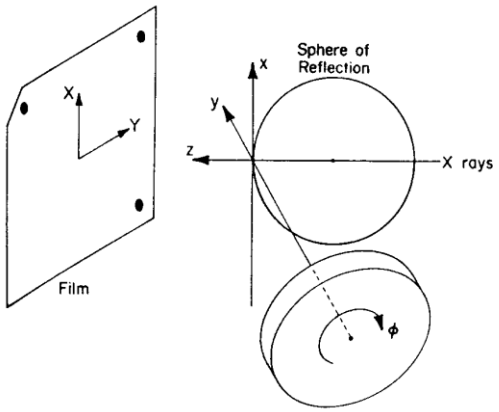
In house developed
Sample Changer
In operation since 2012

Convolutions in Real and Reciprocal Space



Data Collection – Rotation Method

- The crystal is rotated around a generic crystallographic direction so that a certain number of r.l.p.s are brought into diffraction conditions.
- Advantages: Fast
- Disadvantages: Distorted Image of r.l.p planes.
- **Rotation Method is the Standard method of data collection for macromolecules.**



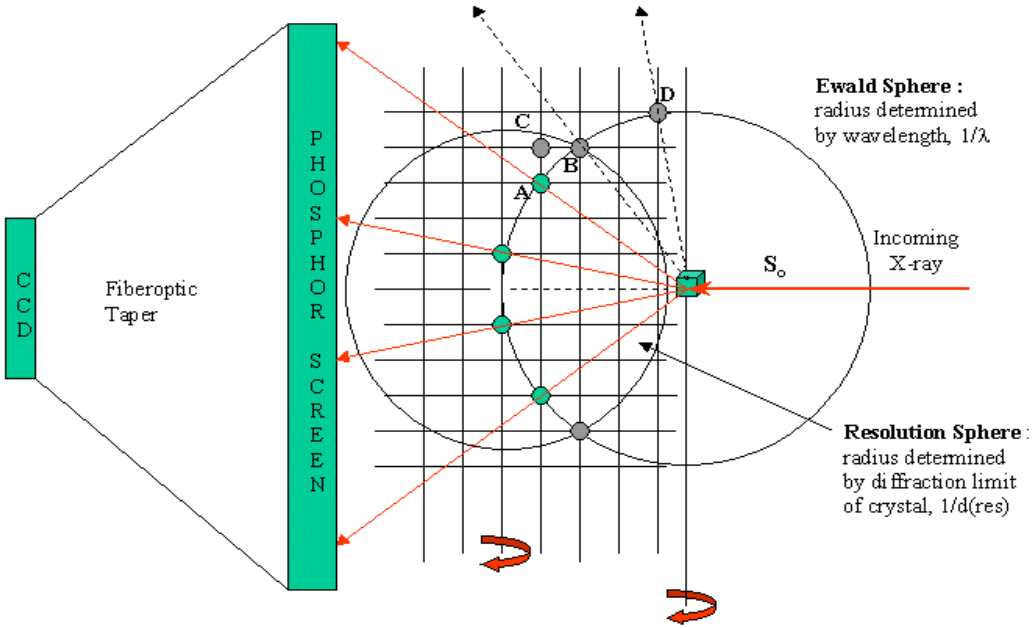
•The simplest setup consist of:

–monochromatic X-ray beam

–single axis goniometer orthogonal to the incident beam

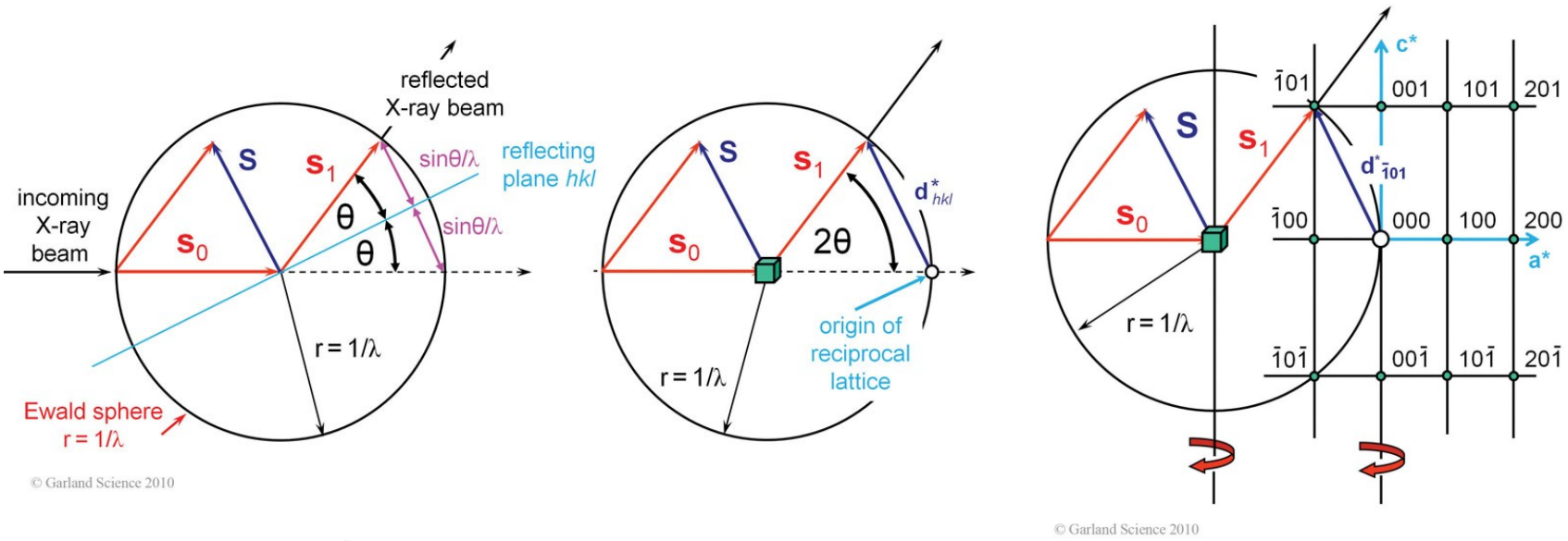
–flat bidimensional detector parallel to the rotation axis. The detector is normally orthogonal to the incident beam, but not necessarily.

Rotation Method - Geometry



- The crystal is rotated around the goniometer axis of a certain angle so that several rlps will be in diffraction condition during the rotation. The diffracted X-rays are collected from the detector behind the crystal, without any intercepting screen between them.
- The rotation is repeated for contiguous angles until, given the orientation of the crystal, the independent part of the reciprocal lattice is completely scanned.

Diffraction Geometry: Edwald construction and reflection condition

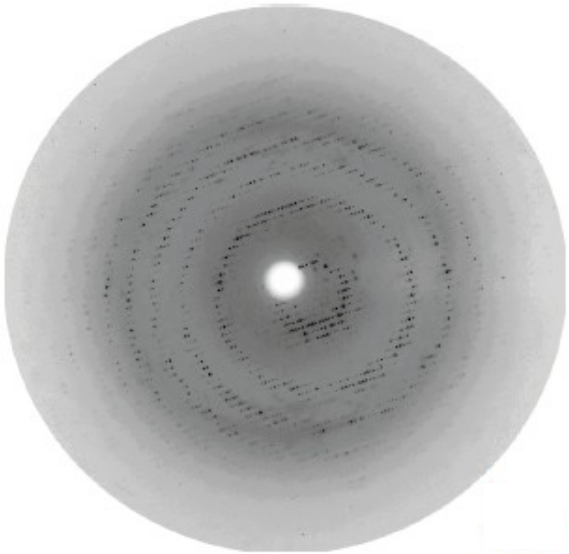


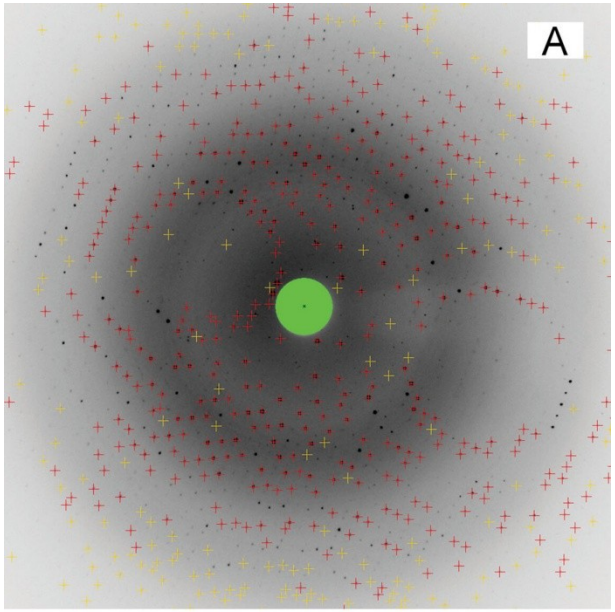
$$\text{Bragg's Law}: n\lambda = 2d_{hkl}\sin\theta_{hkl}.$$

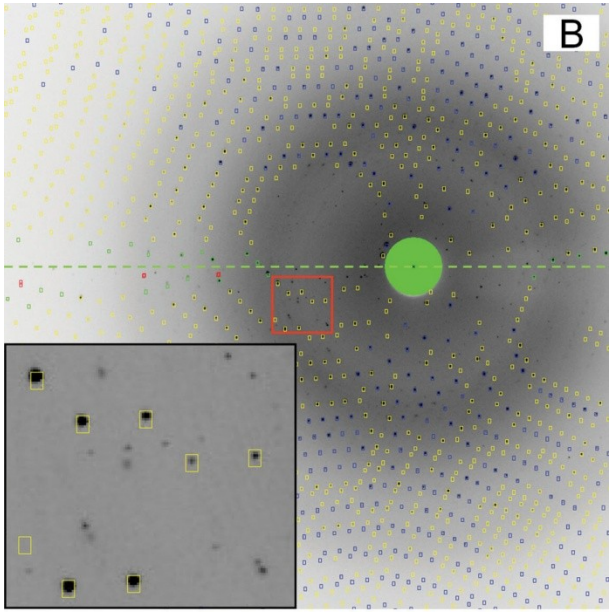
The geometric reflection condition for the set of lattice planes *hkl* is fullfilled when its reciprocal lattice point *hkl* lies on the Edwald sphere.

As few reciprocal lattice vectors will fulfill the diffraction condition of lying on the Ewald sphere in each specific crystal orientation, **rotation of the crystal** – **and thus concurrent rotation of its reciprocal space** – **will bring more reciprocal lattice points to intersect the Ewald sphere whereupon the diffraction conditions are fulfilled**

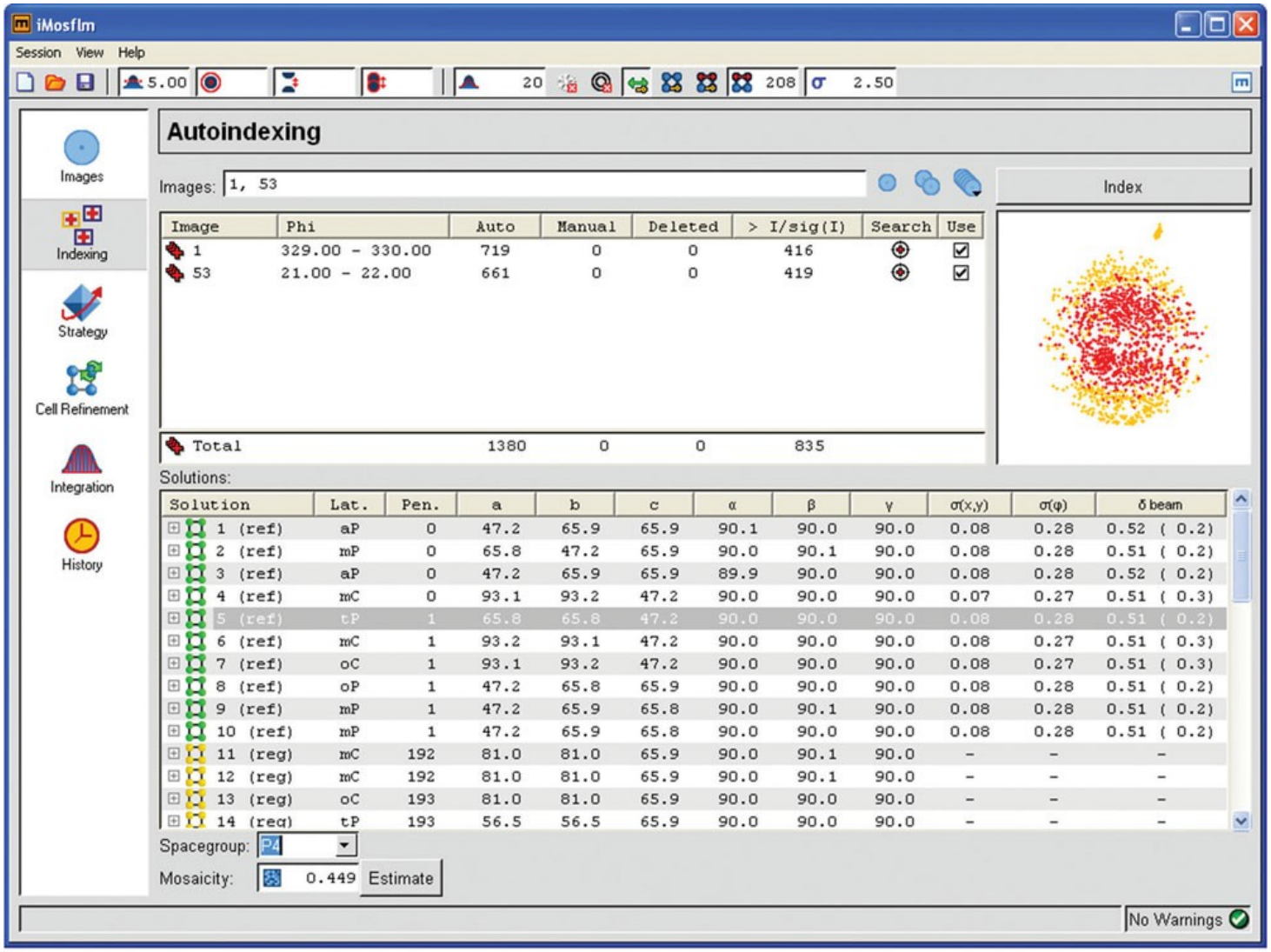
Spot finding and Indexing of images



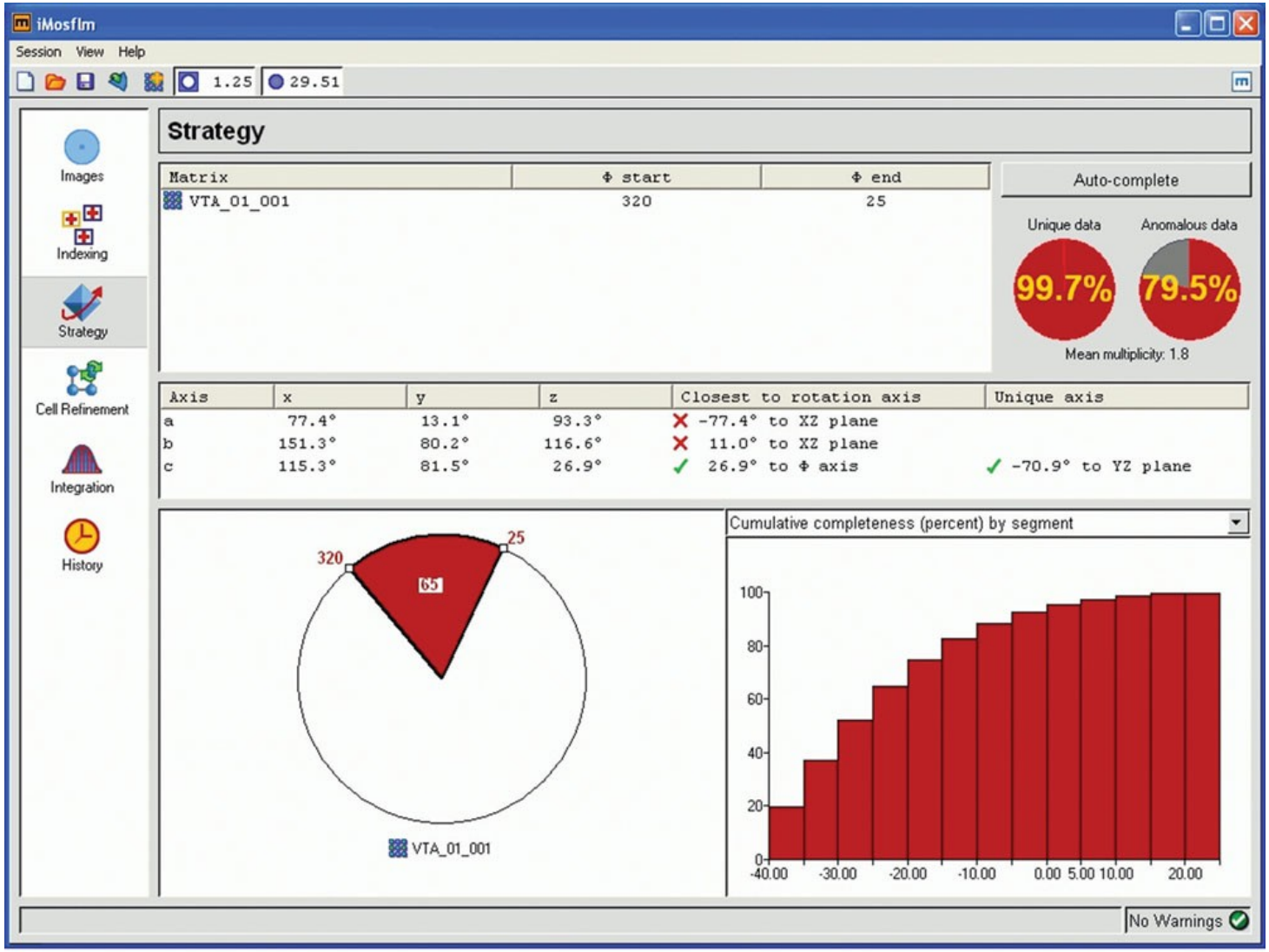




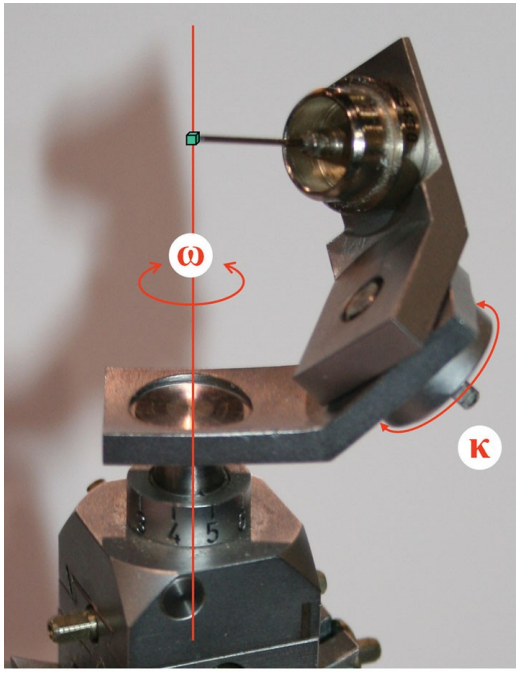
Indexing panel of iMOSFLM

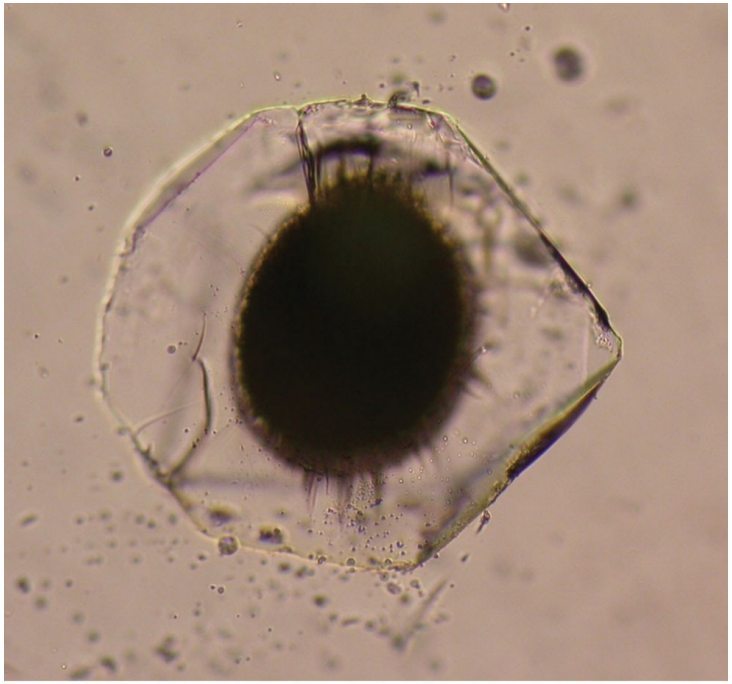


Strategy panel of iMOSFLM

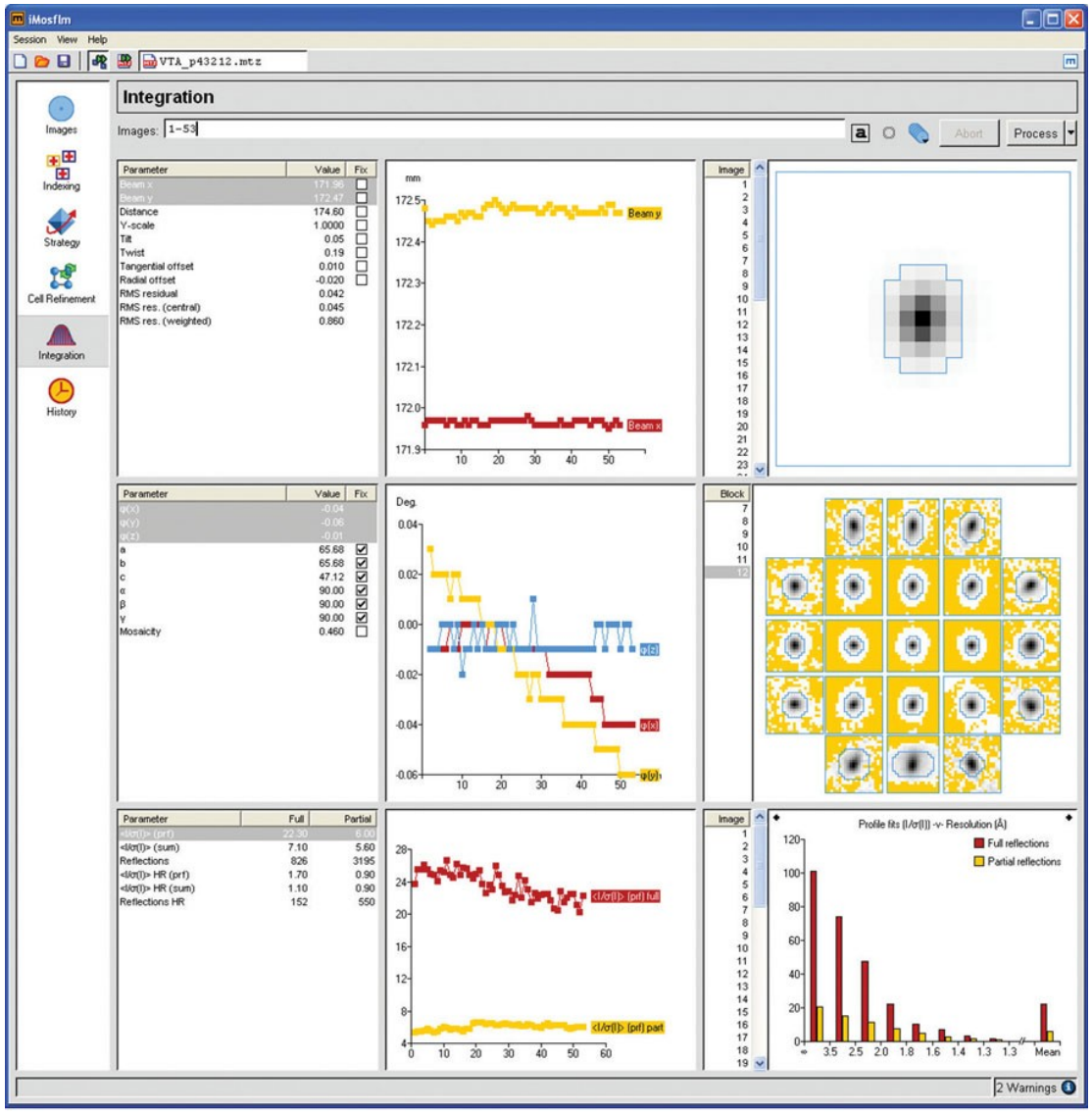


The importance of designing a correct Data Collection Strategy



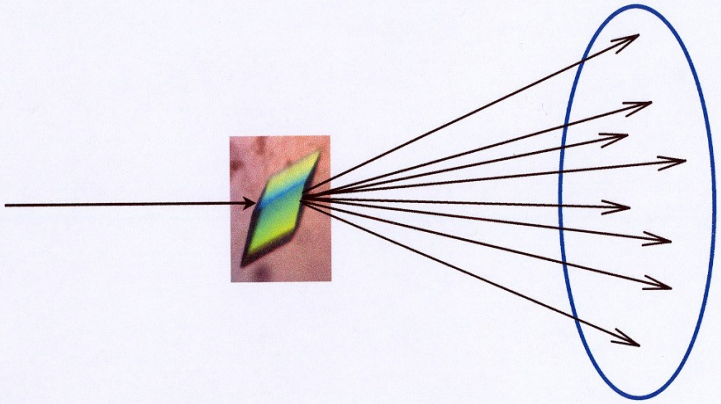


Integration panel of iMOSFLM

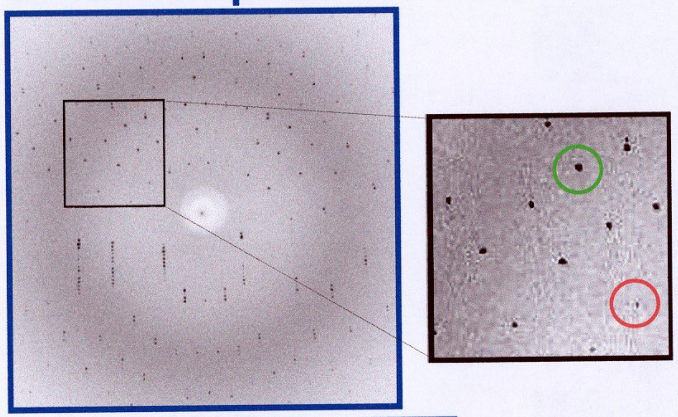


The information from a single diffraction experiment…

(...lacks the phase angle needed for the electron density reconstruction)



diffraction pattern



The reflections are indexed (assignment of the reciprocal cell indices *hkl*) to obtain a long list of intensities, I_{hkl} and associated experimental error σ_{hkl} for the measured reflections

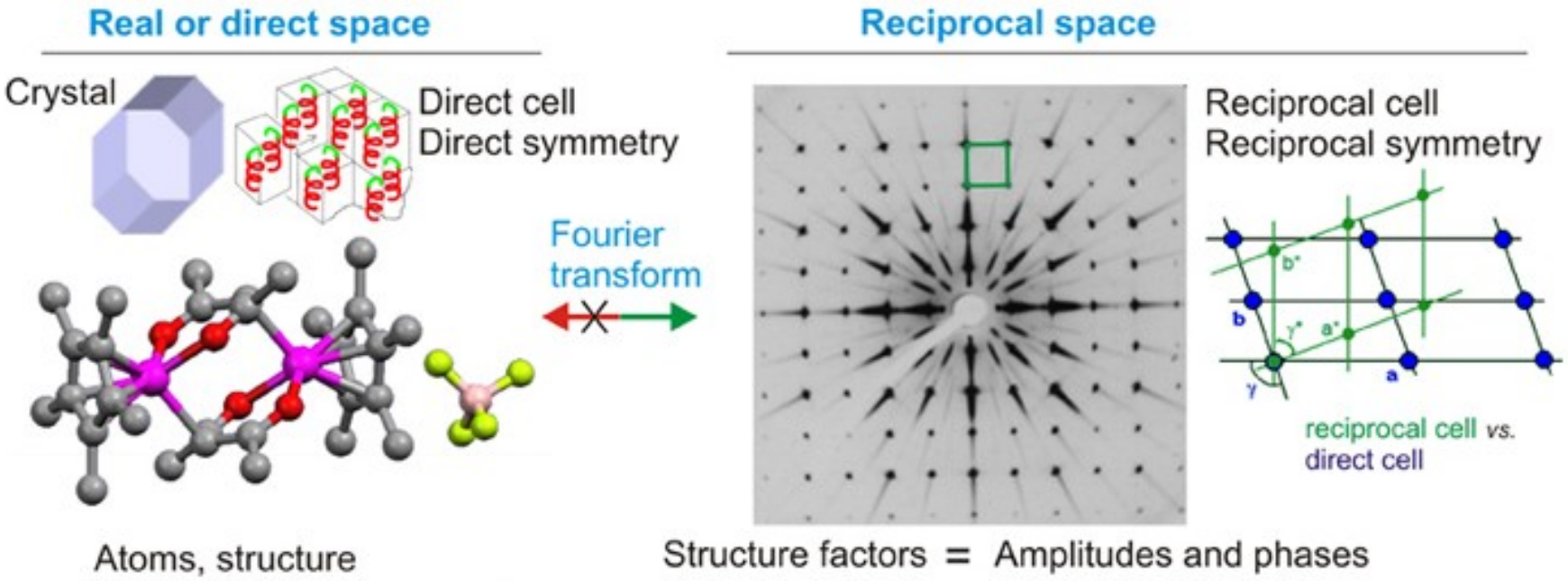
h	k	l	I_{hkl}	σ_{hkl}
8	2	-6	31.86	9.39
8	2	-5	15.79	9.09
8	2	-4	1002.60	20.13
8	2	-2	160.94	6.17
8	2	-1	233.85	6.67
8	2	1	41.78	4.81
8	2	2	309.41	7.83
8	2	3	97.07	5.24
8	2	4	14.93	6.18
8	2	5	6.19	8.37
8	2	6	14.06	7.99

Data Collection – Scaling and Merging

- All the integrated spots from different images are put on a common scale.
- The partial reflections are summed together.
- Better estimate of the unit cell parameters are determined (Post-refinement)
- Symmetry related reflections are merged together giving a final dataset useful for the subsequent structure solution and refinement.
- Several figures of merit are issued at the end of the scaling and merging stage, giving an idea of the goodness of the dataset.

$$R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i(hkl) - \langle I(hkl) \rangle \mid / \Sigma_{hkl} \mid \langle I(hkl) \rangle \mid$$

The Phase Problem



$$\rho(xyz) = \frac{1}{V} \sum_{hkl}^{+\infty} |F(hkl)| \, . \, e^{-2\pi i[hx+ky+lz-\phi(hkl)]}$$

Amplitudes

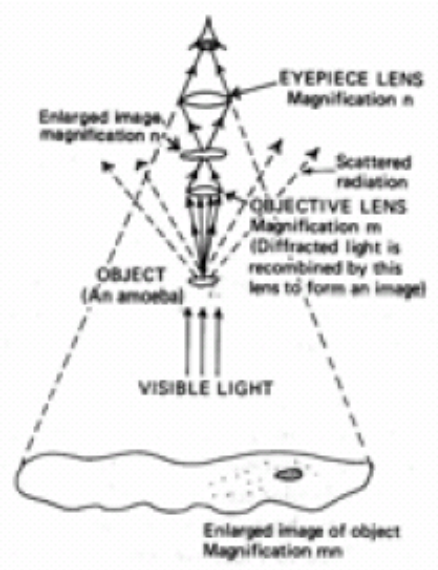
Phases?

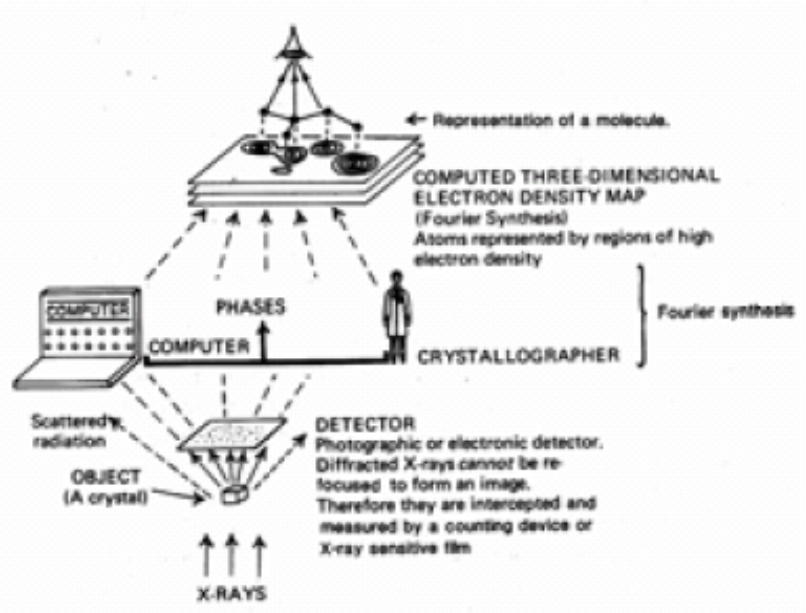
Real or direct space

Reciprocal space

The Phase Problem

An Analogy



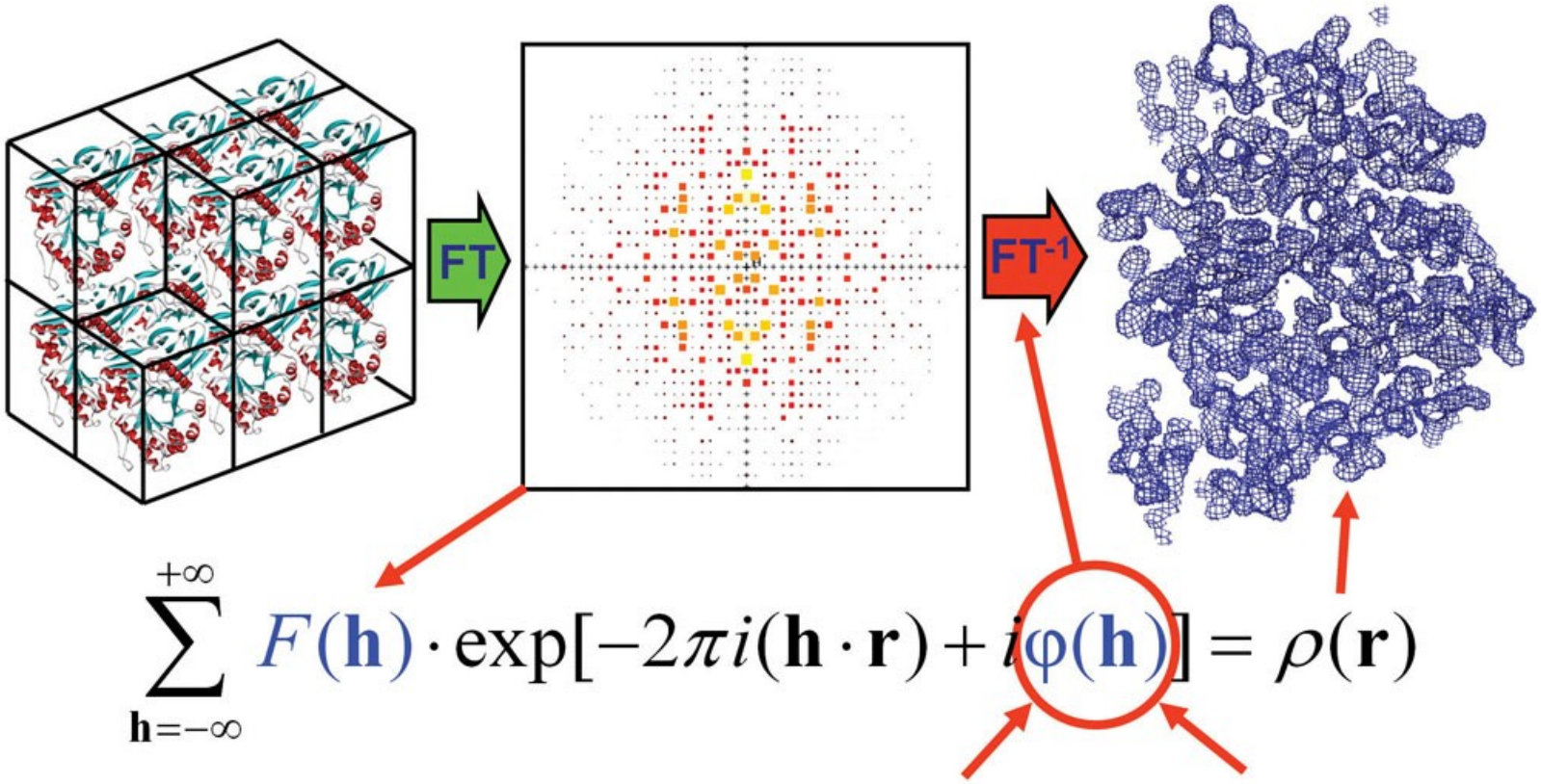


Light microscopy

X-ray crystallography

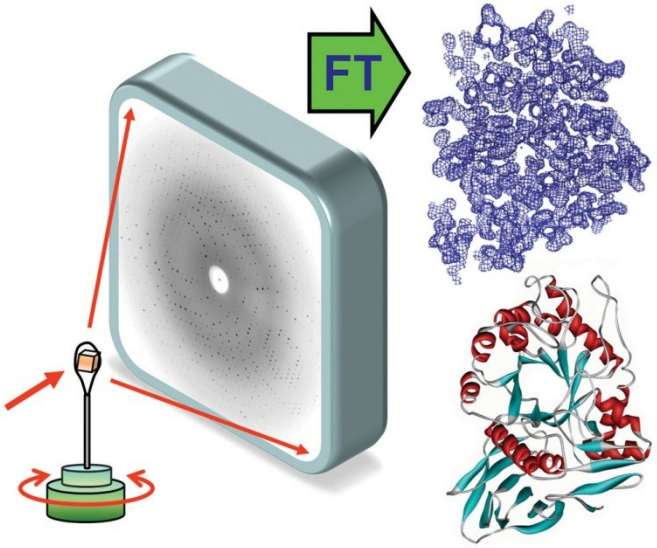
J.P. Glusker & K.N. Trueblood, Crystal Structure Analysis: A Primer. (Oxford, 1985.) pp. 4-5

The Phase Problem

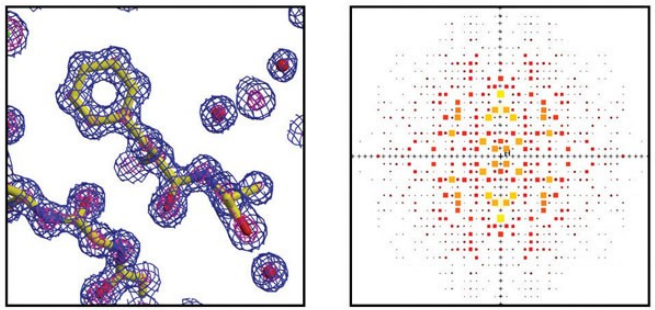


$$\sum_{\mathbf{h}=-\infty}^{+\infty} F(\mathbf{h}) \cdot \exp[-2\pi i(\mathbf{h} \cdot \mathbf{r}) + i\varphi(\mathbf{h})] = \rho(\mathbf{r})$$

The Phase Problem



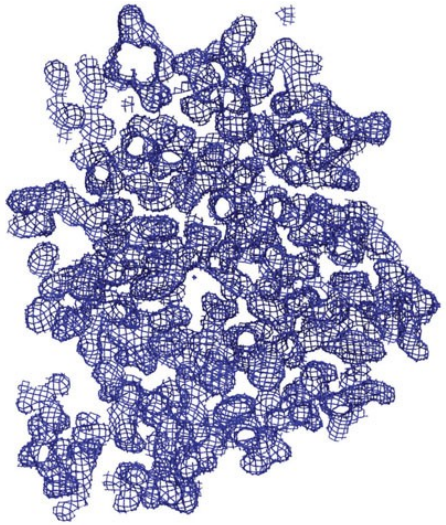
$$\rho(x,y,z) = \frac{1}{V}\sum_{-h}^{h}\sum_{-k}^{k}\sum_{-l}^{l} F_{hkl} \cdot \exp[-2\pi i(hx + ky + lz - \alpha_{hkl})]$$

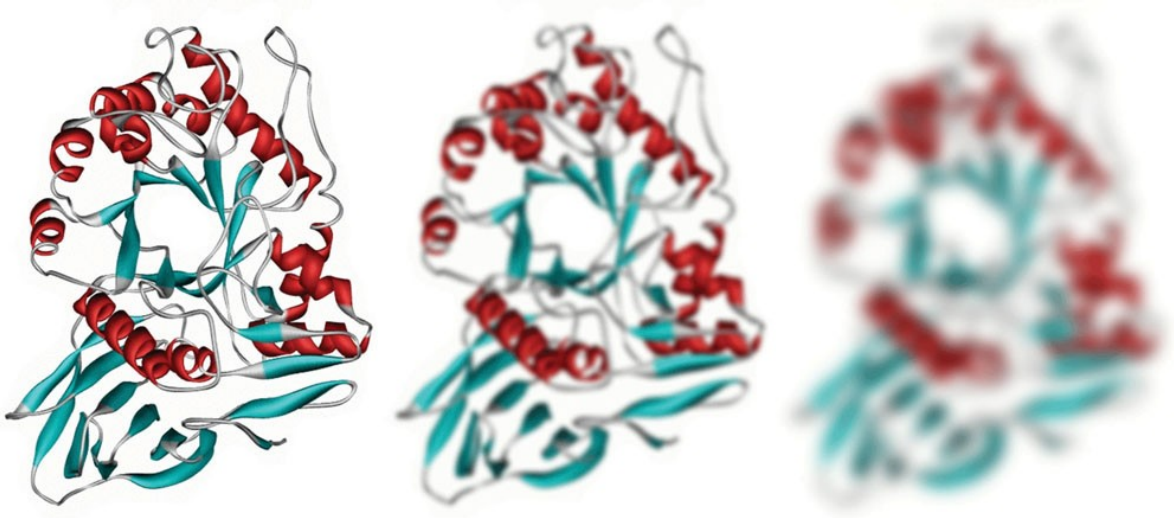


The crystallographic phase problem

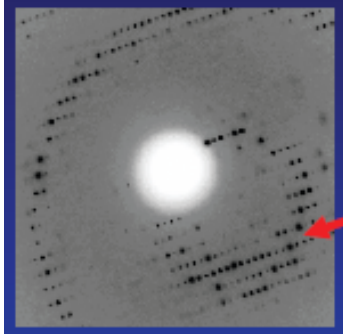
The reconstruction of electron density

h	k	l	F(hkl)	φ(hkl)
2	0	0	228.0	180.0
1	0	1	10.4	90.0
2	0	1	901.8	270.0
1	1	1	367.0	332.1
1	2	3	149.3	37.8
8	9	1	97.9	255.1
7	7	2	111.5	139.7





The Phase Problem



$$I_{hkl} \propto |\mathbf{F}_{hkl}|^2$$

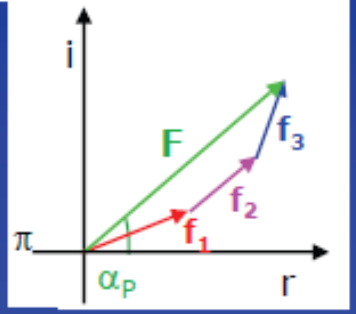
Atomic property

Structural property (position)

$$\mathbf{F}_{hkl} = \sum_{j=1}^{N\ atomi} f_j \exp\left[2\pi i\,(hx_j + ky_j + lz_j)\right] = |\mathbf{F}_{hkl}| \exp\,(i\alpha_{hkl})$$

Structure factor

$$\mathbf{F}_{hkl} = \int_{Vcell} \rho(x,y,z) \exp\left[2\pi i(hx+ky+lz)\right] dV$$



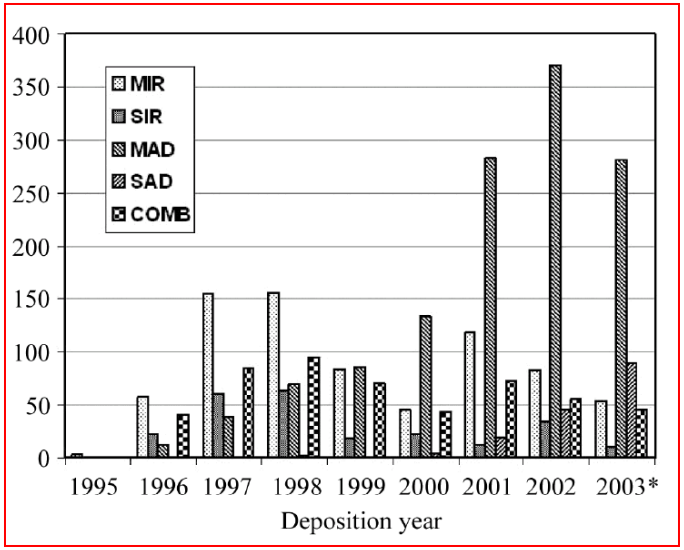
Electron density

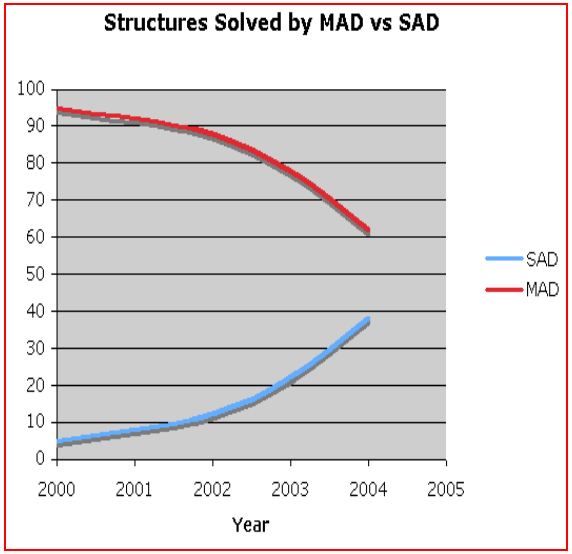
$$\rho(x,y,z) = \frac{1}{V} \sum_{hkl} |\mathbf{F}_{hkl}| \exp\left[-2\pi i(hx+ky+lz-\alpha'_{hkl})\right]$$

not measured

Experimental Phasing Methods in Protein X-ray Crystallography

Methods	Structures	%
Molecular replacement	7548	59.2
Experimental phasing	2956	23.2
Miscellaneous	2128	16.7
Direct methods	119	0.9
Experimental phasing		
Single	2445	
MIR	757	31.0
SIR	248	10.1
MAD	1276	52.2
SAD	164	6.7
Combinations	511	
MIRAS	206	40.3
SIRAS	165	32.3
MADIR	46	9.0
MR + AS	26	5.1
MR + IR	59	11.6
MR + AS + IR	9	1.8
Miscellaneous		
Fourier	1174	55.2
Refine	158	7.4
Model	94	4.4
Other	673	31.6
Miscellaneous	29	1.4



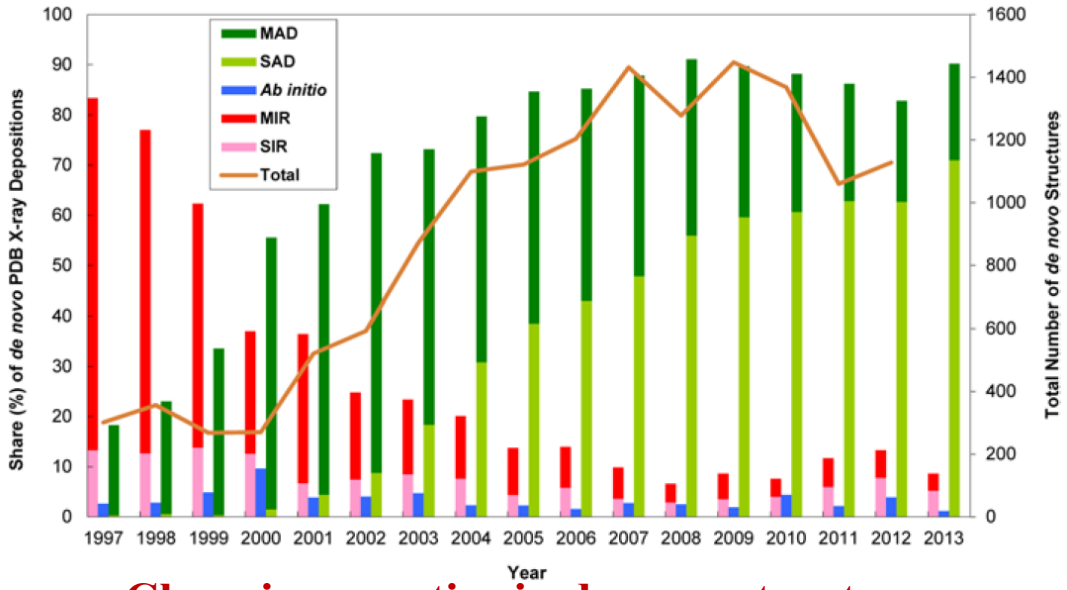


Reported X-ray crystallography methods in PDB depositions (as of 31 December 2003)

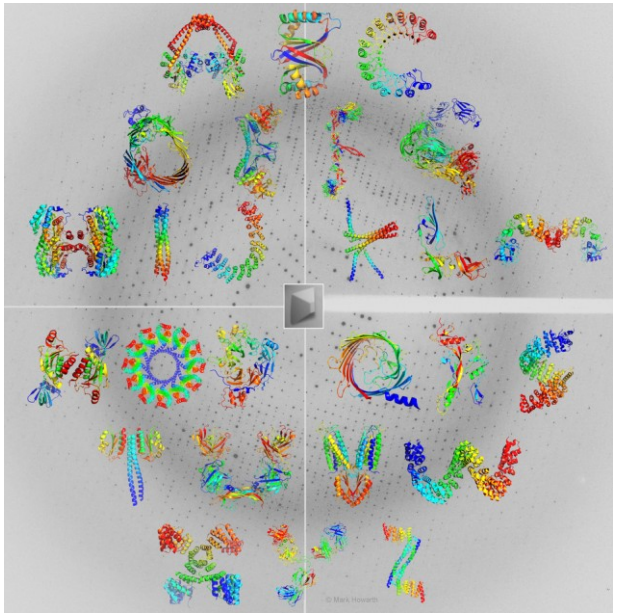
The numbers of reported experimental phasing methods MIR, SIR, MAD,SAD and their combinations over the deposition years

J. Jiang & R.M. Sweet, *J. Synchr. Rad*. (2004) 11, 319

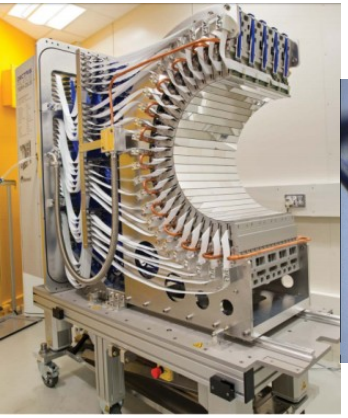
Evolution of diffraction methods for solving macromolecular crystal structures

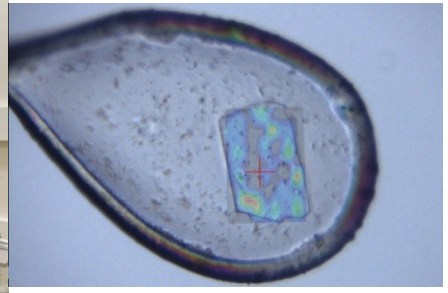


Changing practice in de novo structure determination



Say it with proteins: An *alphabet* of crystal structures





❖**Long Wavelength Macromolecular Crystallography Beamline** takes shape at Diamond : *in vacuo* X-Ray detector Pilatus 12M

Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection.

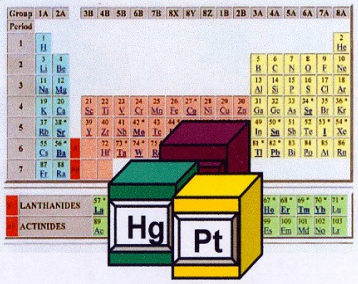
❖**In cristallo optical spectroscopy:** synergies between X-Ray Crystallography and Single Crystal UV-Vis, Raman Microspectrophotometry .

❖**FedEX Crystallography**; **Remote data collection**

❖**An Integrated Structural Biology Infrastructure for Europe**

Quartery Reviews of Biophysics (2014) 47: 49; Acta Cryst. (2016). **D72**:430

Methods to obtain initial experimental phases



1- Multiple Isomorphous Replacement (MIR)

- At least one native and 2 derivative crystals must be available.
- Non-isomorphism limits phases to lower resolution.

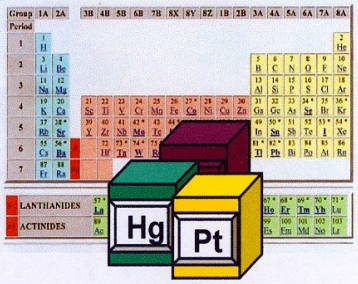
2- Multiple Wavelength Anomalous Dispersion (MAD)

- Several data sets collected at different wavelengths (needs a tunable synchrotron radiation source)
- But an anomalous scatterer must be present.
- Specially useful in metalloproteins which have intrinsic metals (e.g., Fe, Co, Zn, W)
- Used in the case of Se-methionine substituted proteins

3- Patterson Search (PS) or Molecular Replacement (MR)

- An homologous structure must be available (search model)

Methods to obtain initial experimental phases



1- Multiple Isomorphous Replacement (MIR)

- At least one native and 2 derivative crystals must be available.
- Non-isomorphism limits phases to lower resolution.

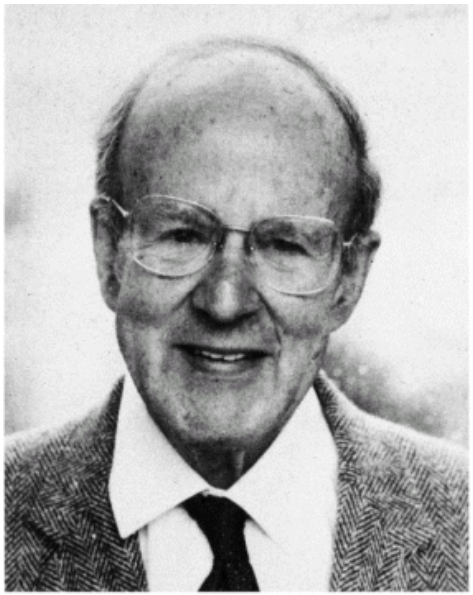
2- Multiple Wavelength Anomalous Dispersion (MAD)

- Several data sets collected at different wavelengths (needs a tunable synchrotron radiation source)
- But an anomalous scatterer must be present.
- Specially useful in metalloproteins which have intrinsic metals (e.g., Fe, Co, Zn, W)
- Used in the case of Se-methionine substituted proteins

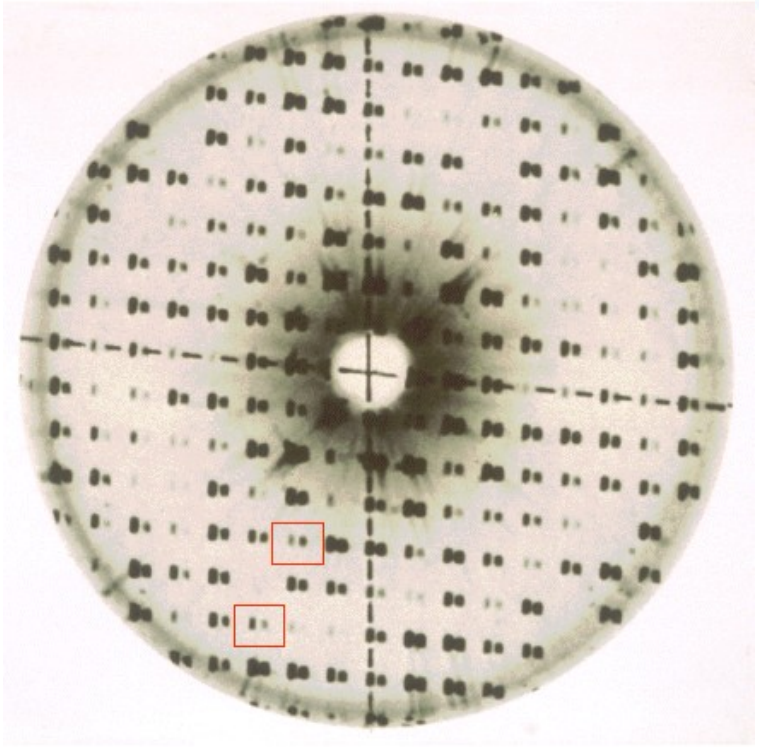
3- Patterson Search (PS) or Molecular Replacement (MR)

- An homologous structure must be available (search model)

Single-Multiple Isomorphous Replacement (SIR-MIR)



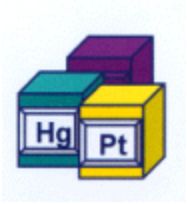
Max F. Perutz (1914-2002)
MRC Mol. Biol. Lab.
Cambridge, UK
Nobel Laureate in Chemistry, 1962



Native and Heavy Atom Derivatized Protein Crystal

Basic principle: Binding of heavy atoms to the macromolecules ***does not*** change its structure (*Isomorphism*)

Single-Multiple Isomorphous Replacement (SIR-MIR)

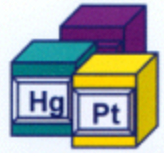


- The presence of the heavy atom(s) introduces differences to the diffraction pattern with respect to the diffraction pattern of the native crystal: The **differences** are in the ***intensities*** of the diffracted X-rays
- When an heavy atom binds ***isomorphously***, then the difference between the two samples are due ***only*** to the presence of the heavy atoms(s)

Frequently used heavy atoms

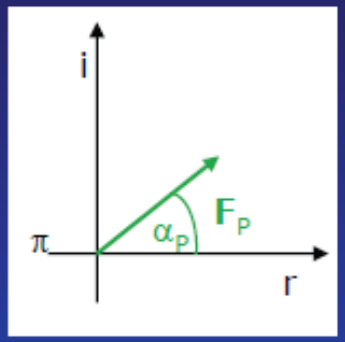
- Pt, Au, Hg, Pb, Th, U, Re, Os, Ir,
- Pd, Ag (small atomic number) - for small proteins
- J, iodinated tyrosine, modified nucleic acid bases (J, Br)
- Lanthanides (La-Lu) - can substitute Mg^{+2} or Ca^{+2}
- Noble gasses (Xe, Kr)
- Cryo halides (NaBr, KI)

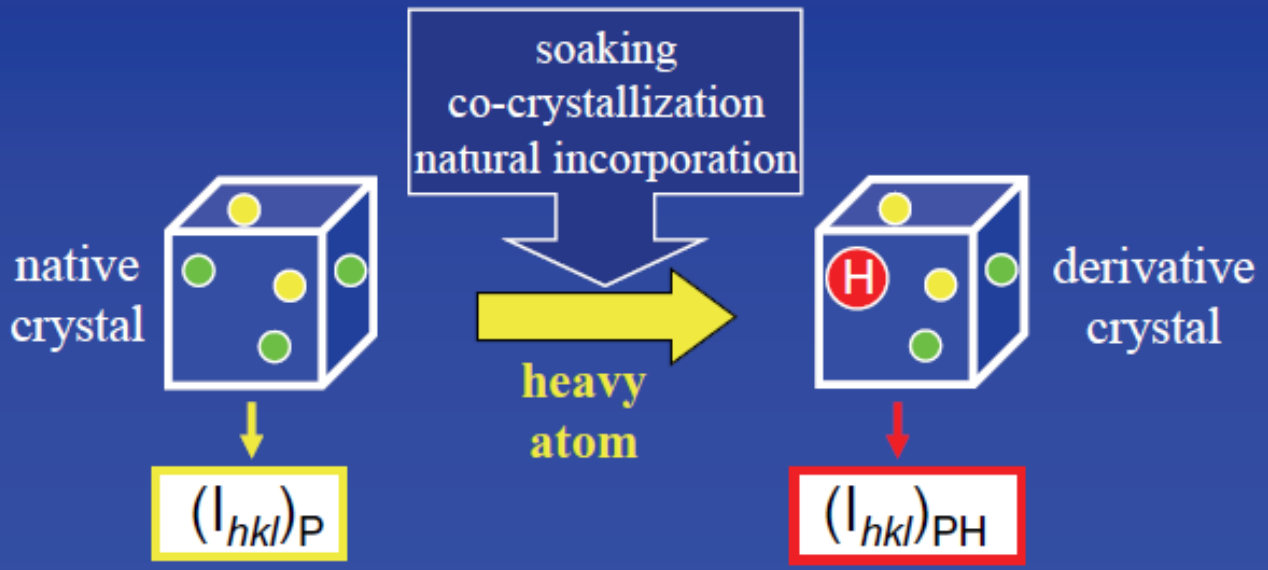
Single-Multiple Isomorphous Replacement (SIR-MIR)



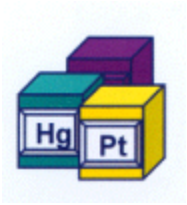
Phasing by heavy atoms:

$$\mathbf{F}_{hkl} = \sum_{j=1}^{N\ atoms} f_j \exp\left[2\pi i\,(hx_j + ky_j + lz_j)\right] = |\mathbf{F}_{hkl}|\exp\left(i\alpha_{hkl}\right)$$



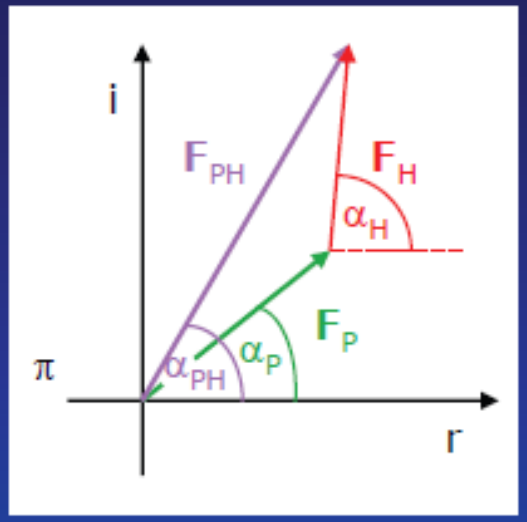


Single-Multiple Isomorphous Replacement (SIR-MIR)



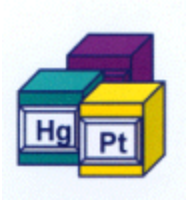
Phasing by heavy atoms: MIR

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_H$$



- $|F_P|$, $|F_{PH}|$ (experimentally measured: $I \propto |F|^2$)
- $|F_H|$, α_H (calculated from heavy atom position ⇔ **Patterson synthesis**)
- isomorphism between native and derivative crystals is required

Single-Multiple Isomorphous Replacement (SIR-MIR)



Determination of heavy atom positions: Patterson Map

The Patterson function P (uvw) is essentially the Fourier Transform of the **Intensities** rather than the **Structure Factors**

A.L. Patterson "A Fourier Series Method for the Determination of the Components of Interatomic Distances in Crystals" Phys. Rev. (1934) **46**, 372

$$\rho(xyz) = \frac{1}{V} \sum_{\substack{hkl \\ -\infty}}^{+\infty} |F(hkl)| \,.\, e^{-2\pi i[hx+ky+lz-\phi(hkl)]}$$

$$P(uvw) = \frac{1}{V} \sum_{\substack{hkl \\ -\infty}}^{+\infty} |F(hkl)|^2 \,.\cos 2\pi\,[hu + kv + lw]$$

$$|F(hkl)|^2 = \frac{I(hkl)}{K.A.L.p}$$

K is a scale factor, **A** is the absorption factor, **L** is the Lorentz factor, and **p** represents the polarization factor

It can always be calculated from the experimental diffraction data (no phase information is needed!!!

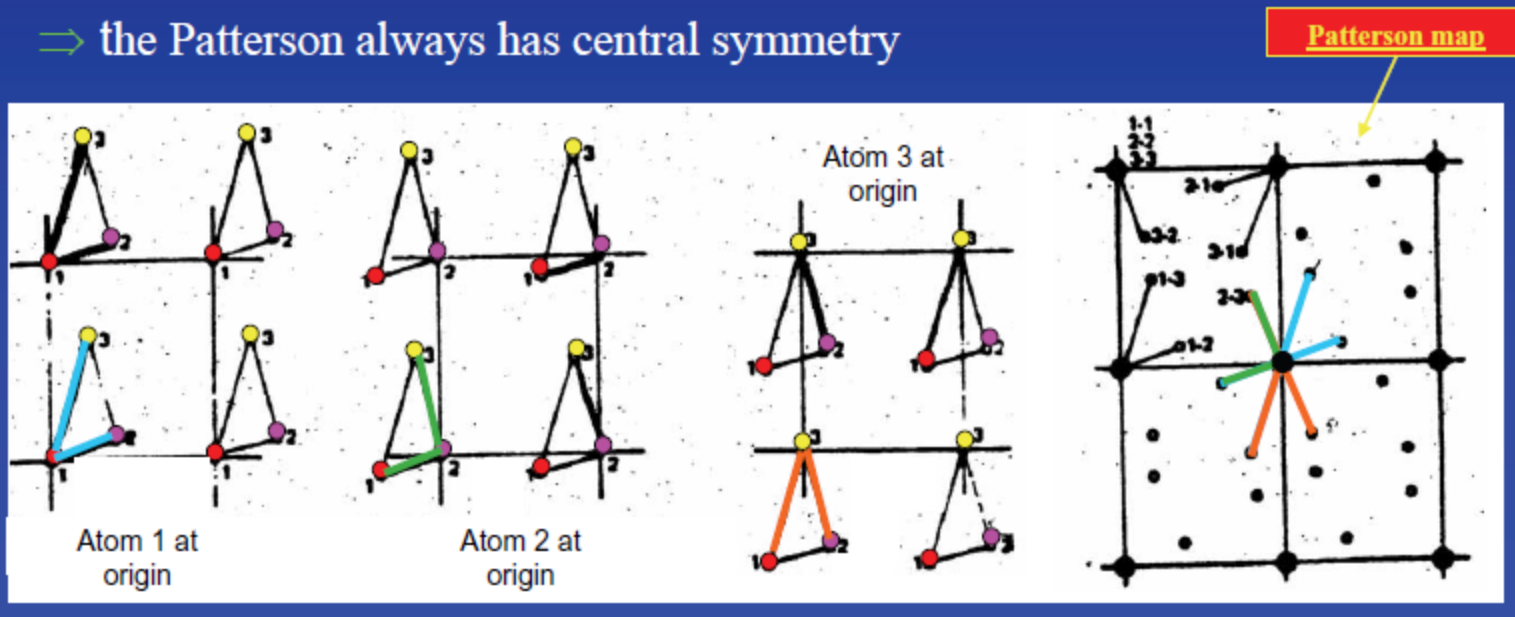
Single-Multiple Isomorphous Replacement (SIR-MIR)

Patterson synthesis:

⇒ a Patterson map of N points will have N(N−1) peaks, excluding the central peak (N peaks at the origin) and any overlap.

⇒ the **Patterson peaks are the interatomic distances** weighted by the product of the number of electrons in the atoms concerned ($\propto Z_A \times Z_B$)

⇒ the Patterson always has central symmetry



Single-Multiple Isomorphous Replacement (SIR-MIR)

Patterson synthesis:

To obtain the Patterson function solely for the heavy atoms in derivative crystals, we construct a **difference Patterson function**

Difference Patterson function

$$\Delta\rho^{P}(u, v, w) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k} \sum_{l} \Delta F_{hkl}^{2} \exp\left[-2\pi i(hu+kv+lw)\right]$$

where $\Delta F_{hkl} = (|F_{PH}| - |F_{P}|)_{hkl}$

Deconvolution of the difference Patterson function allows the calculation of heavy atom positions in the crystal unit cell $\Rightarrow |F_H|, \alpha_H$

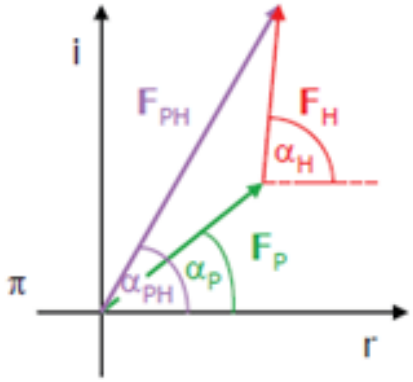
Single-Multiple Isomorphous Replacement (SIR-MIR)

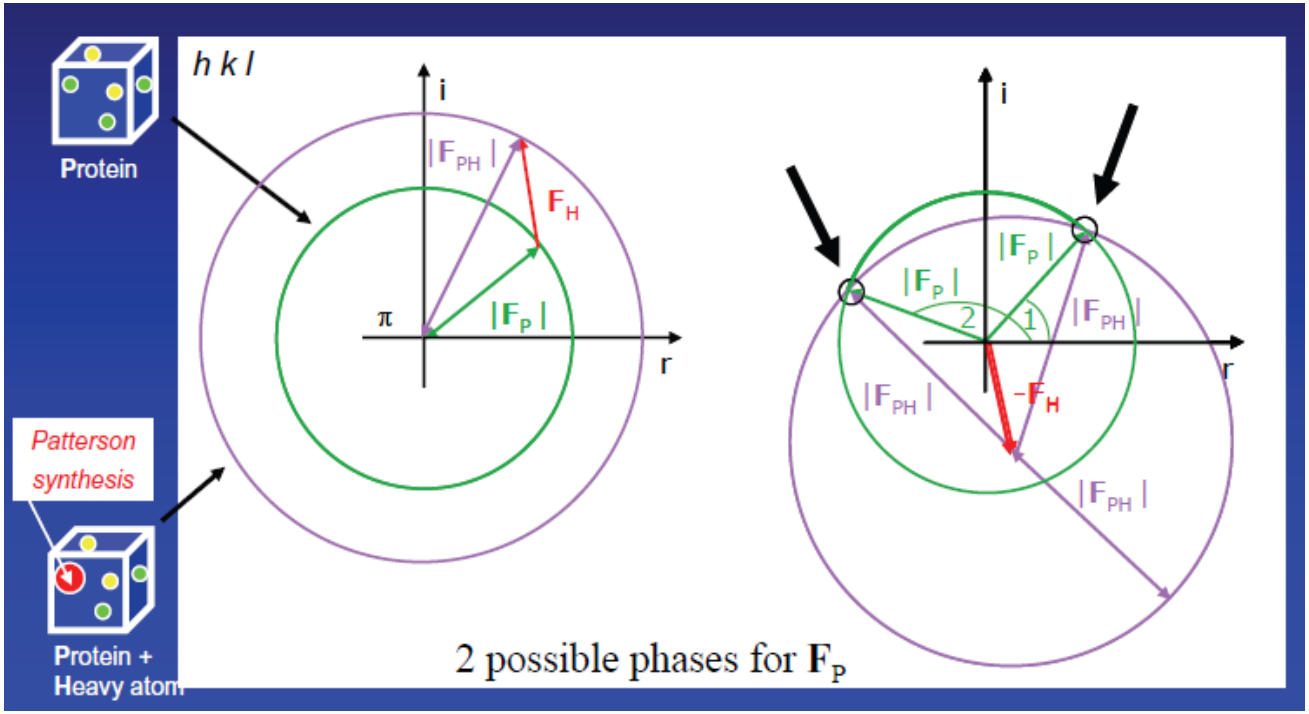
$F_{ph} = F_p + F_h$

$F_{ph}^2 = F_p^2 + F_h^2 + 2F_pF_h\cos(\alpha_p - \alpha_h)$

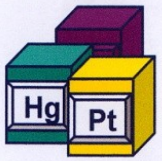
- amplitudes: F_p^2, F_h^2, F_{ph}^2
- phases: α_p, α_h

- $\alpha_p = \alpha_h +/- \arccos[(F_{ph}^2 - F_p^2 - F_h^2)/2F_pF_h]$
- $\alpha_p = \alpha_h +/- \alpha'$
- The sign of α' is not determined with a single derivative





1- Multiple Isomorphous Replacement (MIR)

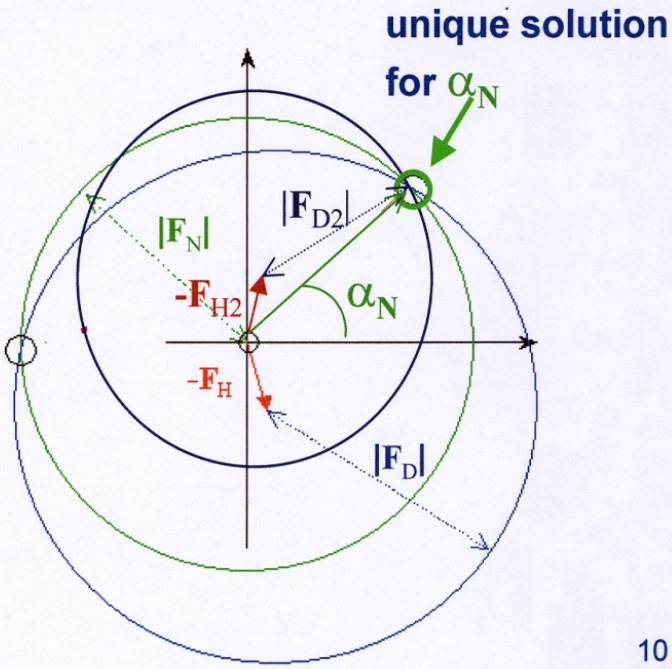


We know only the magnitudes $|F_D|$ (derivative) and $|F_N|$ (native protein), which can be represented in the complex plane as a circle of radius $|F_D|$ and $|F_N|$ respectively.

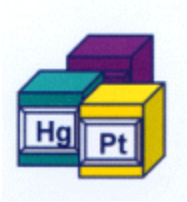
- F_H (amplitude and phase) can be calculated from the known heavy atom positions ($x_i\ y_i\ z_i$) with, $F_H = \sum f_i \exp 2\pi i\,(hx_i+ky_i+lz_i)$

The *phase ambiguity* is overcome with a second derivative F_{H2}

(at a different position from the first)

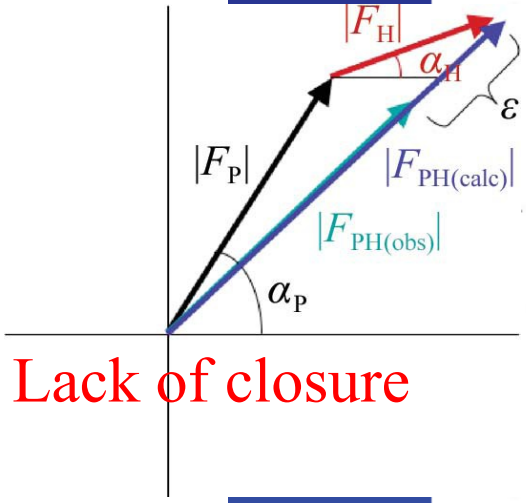


Single-Multiple Isomorphous Replacement (SIR-MIR)

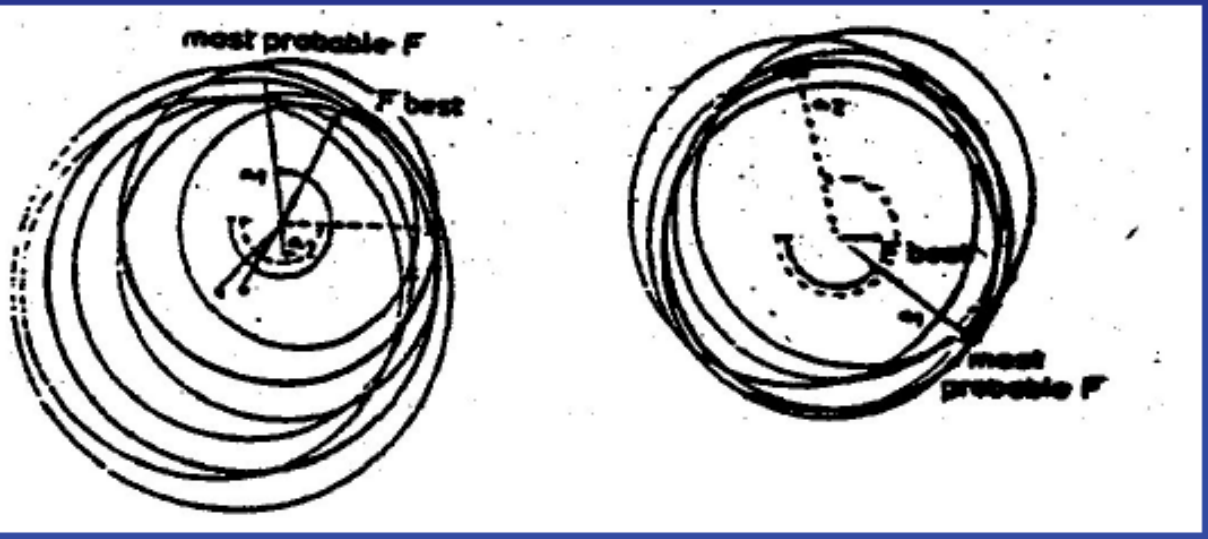


Real case:

- experimental errors in $|F_{PH}|$ and $|F_P|$
- poor isomorphism between native and derivative crystals
- errors in heavy atom localization (Patterson "noise")



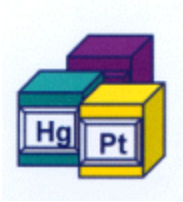
Lack of closure



$$\rho(x,y,z) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k} \sum_{1} m \left[|F_{hkl}| \, e^{-\alpha(best)_{hkl}} \right] e^{-2\pi i(hx+ky+lz)}$$

m = "figure of merit"
$0 \le m \le 1$

Single-Multiple Isomorphous Replacement (SIR-MIR)

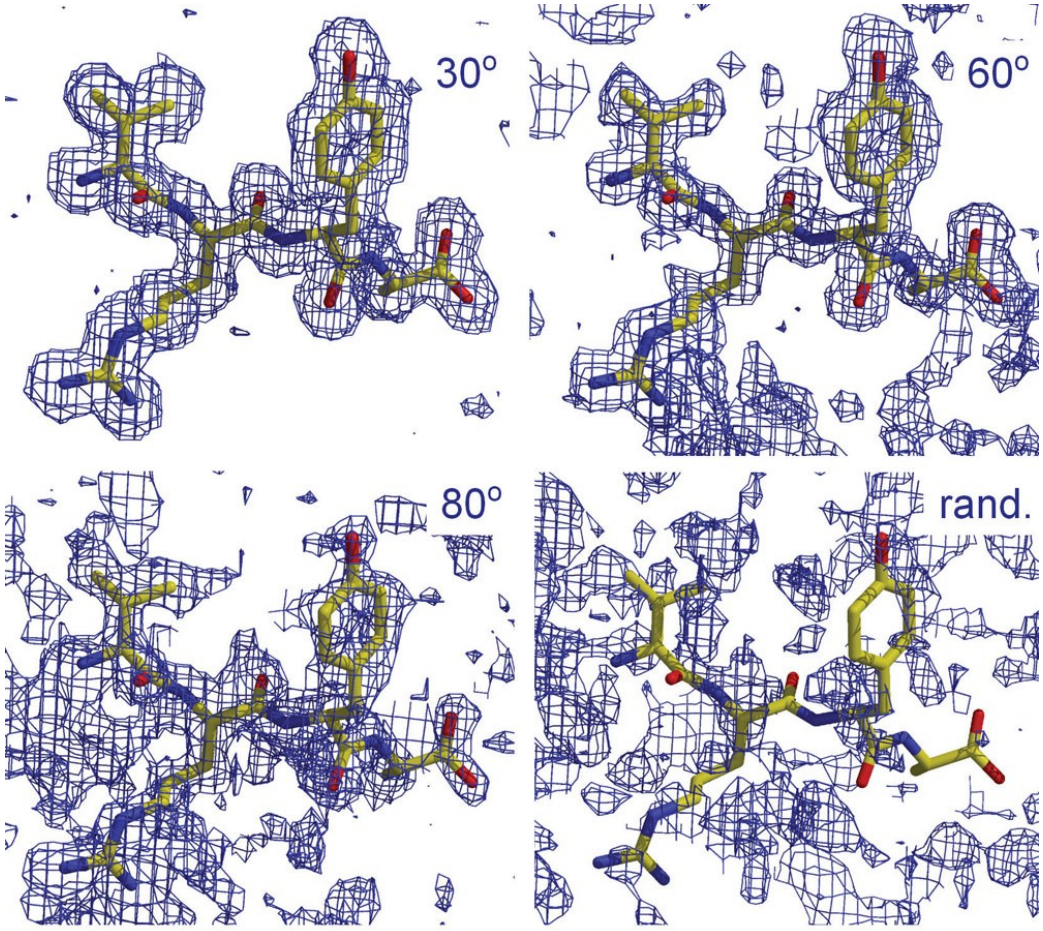


Having determined α_p for each diffracted beam hkl we can calculate the electron density map

$$\rho(xyz) = (1/V_c)\Sigma\Sigma\Sigma F_p(hkl)\exp[-2\pi i(hx + ky + lz)]$$

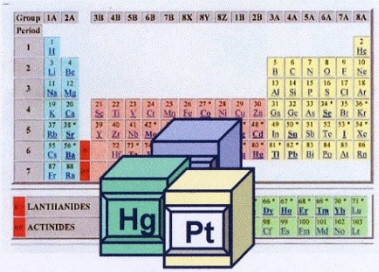
$$F_p = |F_p| \exp(i\alpha_p) = |F_p| \cos\alpha_p + i\,|F_p| \sin\alpha_p$$

Single-Multiple Isomorphous Replacement (SIR-MIR)



Influence of the phase error on electron density map reconstruction

Methods to obtain initial experimental phases

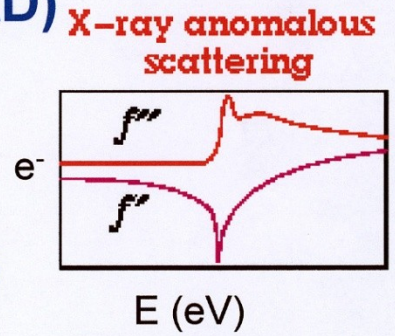


1- Multiple Isomorphous Replacement (MIR)

- At least one native and 2 derivative crystals must be available.
- Non-isomorphism limits phases to lower resolution.

2- Multiple Wavelength Anomalous Dispersion (MAD)

- **An anomalous scatterer* must be present.**
- **Specially useful in metalloproteins which have intrinsic metals (e.g., Fe, Co, Zn, W)**
- **Used in the case of Se-methionine substituted proteins**
- **Several data sets collected at different wavelengths (needs a tunable synchrotron radiation source)**



***atom which has an absorption edge at a chosen λ**

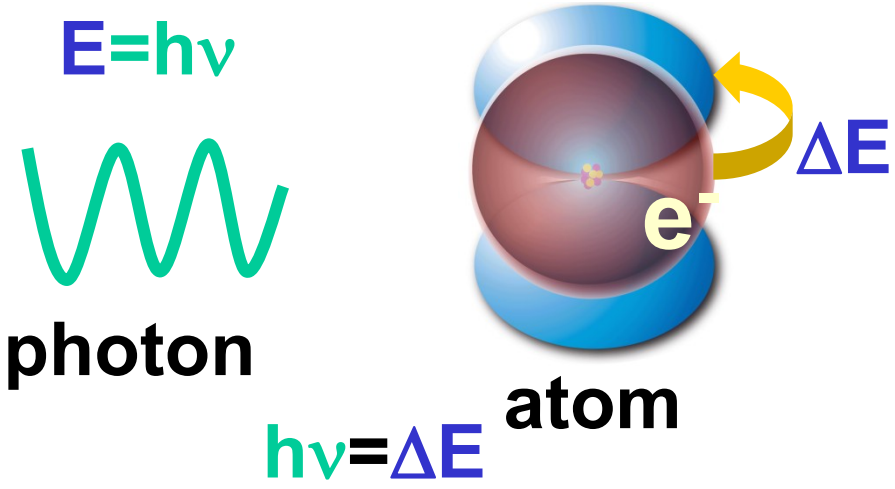
3- Patterson Search (PS) or Molecular Replacement (MR)

- An homologous structure must be available (search model)

Multiple Wavelength Anomalous Dispersion

Anomalous scattering

Anomalous scattering arises when the incident radiation has energy sufficient to promote an electronic transition in the scattering atom

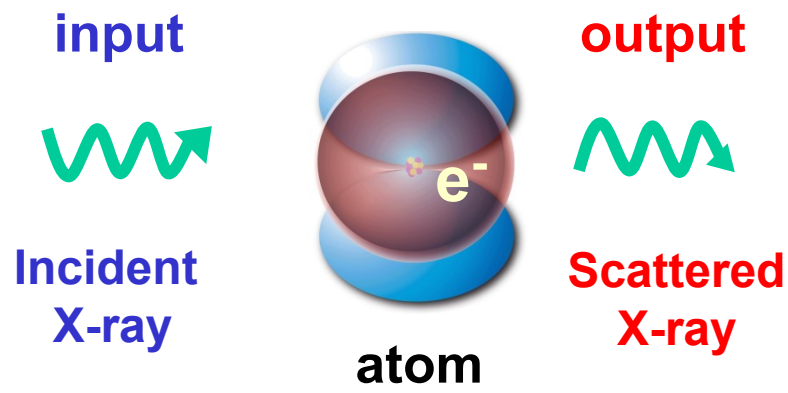


This condition depends on:

- the Energy of the incident photon
- the quantum properties of scattering atom.

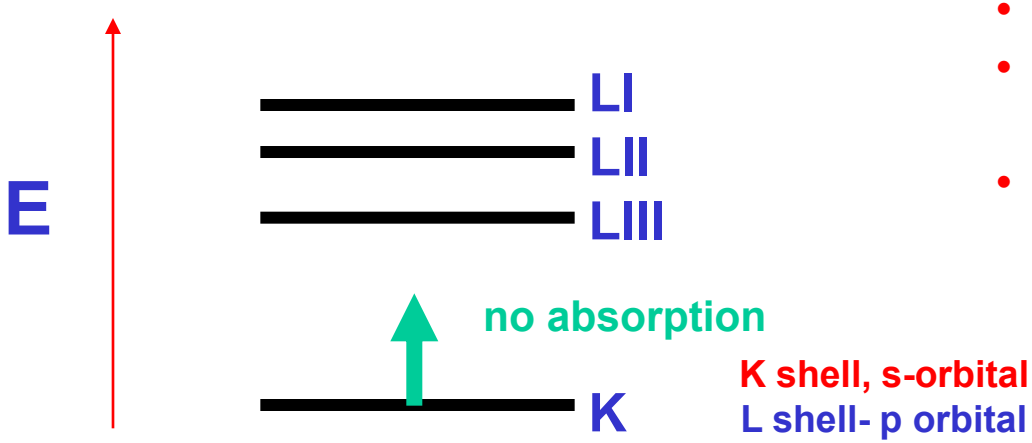
Multiple Wavelength Anomalous Dispersion

Normal scattering



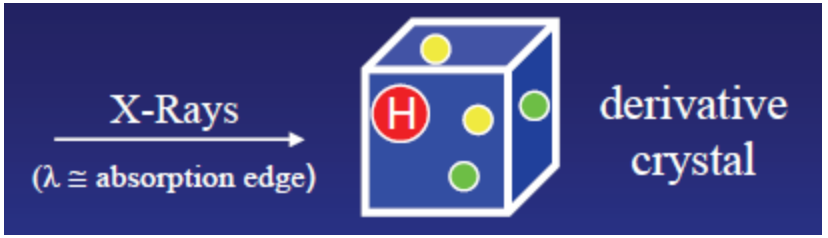
When the X-ray wavelength is far from the absorption edge, X-ray scattered normally

- The energy of the incident X-ray is far from the energy needed to promote an electronic transition.
- No absorption occurs.
- Incident photon bounces off the electrons elastically.
- Scattered X-ray has no phase shift. Thompson scattering



Multiple wavelength Anomalous Dispersion

Anomalous scattering

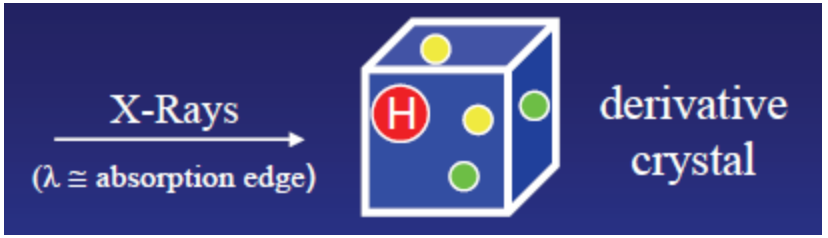


When the X-ray wavelength is near the absorption edge something *peculiar* happens

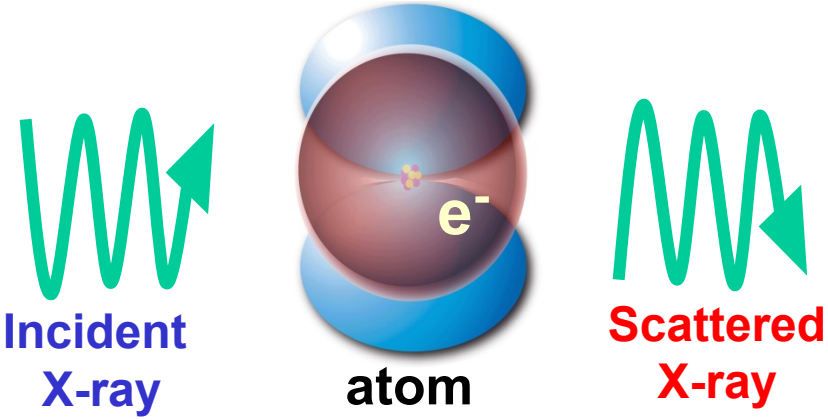
e^-

Multiple wavelength Anomalous Dispersion

Anomalous scattering

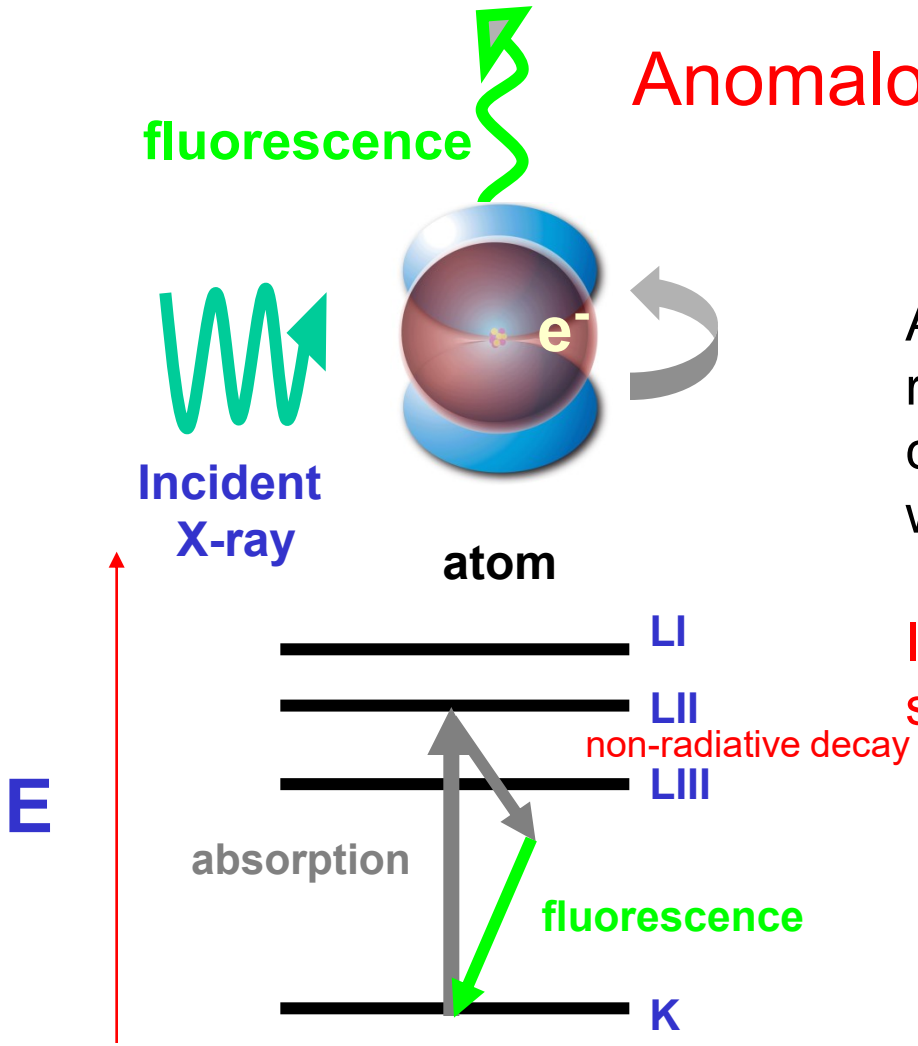


When the X-ray wavelength is near the absorption edge something *peculiar* happens



One component of the incident X-rays is scattered normally.
No phase delay.

Multiple Wavelength Anomalous Dispersion



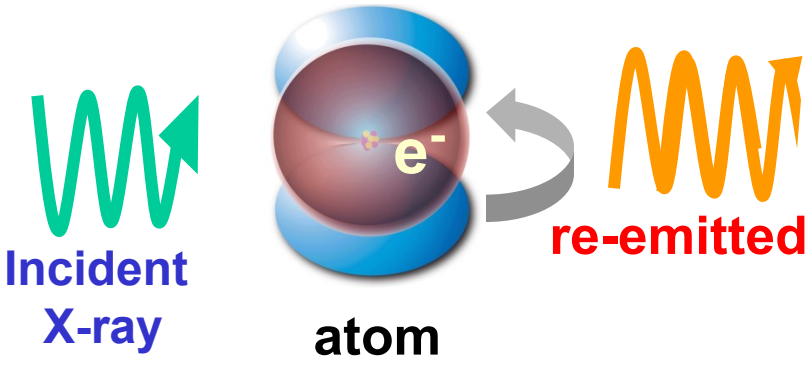
Anomalous scattering

A second component of the incident X-rays is absorbed and the energy is converted into fluorescence of a longer wavelength.

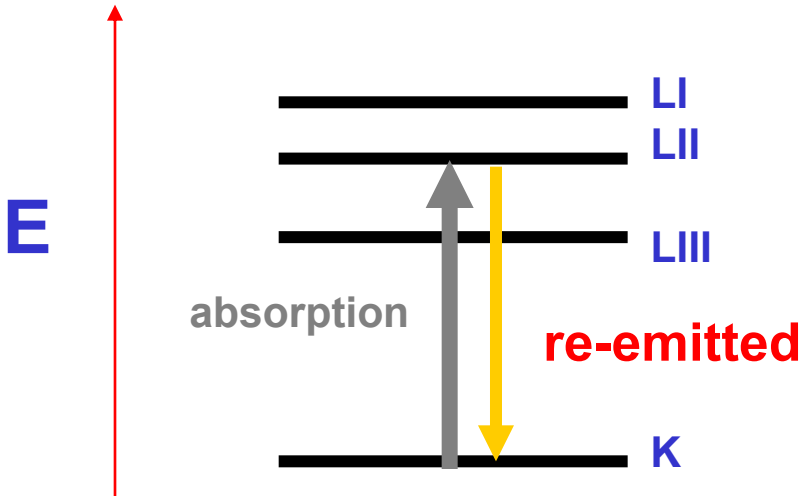
It diminishes the strength of the normal scattering (180° phase shift)

Multiple Wavelength Anomalous Dispersion

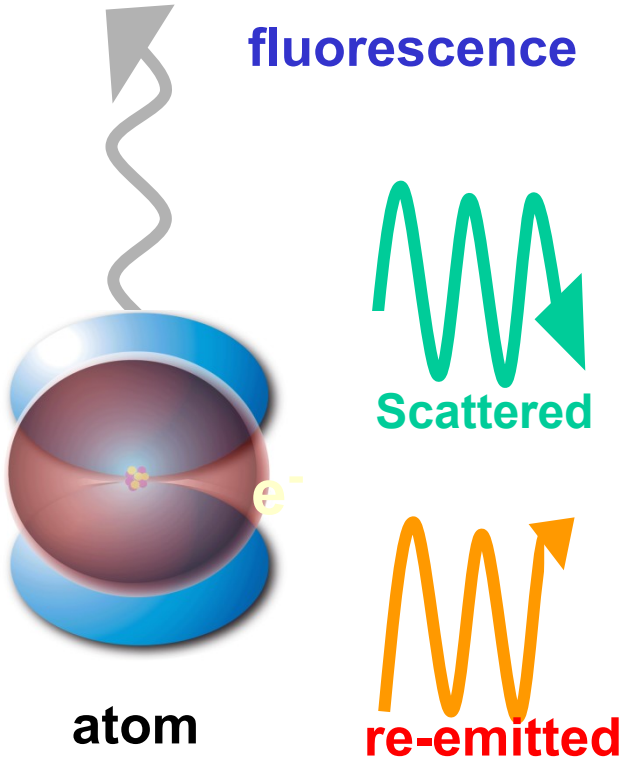
Anomalous scattering



A third component of the incident x-rays is absorbed and re-emitted at the same wavelength but its phase is delayed by 90°



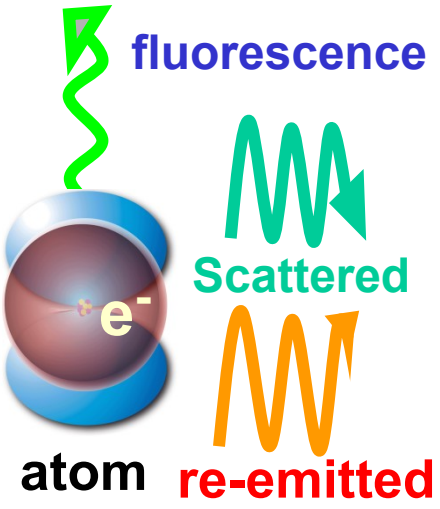
Multiple wavelength Anomalous Dispersion



Under anomalous scattering conditions, the scattered radiation has 3 components:

- a component that is scattered **normally** (no phase delay),
- a component that is absorbed and lost in **fluorescence** (180°, phase shift) of longer wavelength
- a component that is absorbed and **re-emitted** at the same wavelength (90° phase delay)

Multiple wavelength Anomalous Dispersion



Anomalous scattering

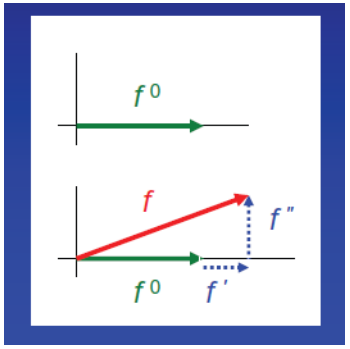
The atomic scattering factor for this atom must also have 3 components.

The sum of the three components of the scattering factor makes it a complex number

(1) (2) (3)

normal scattering (real) + absorption correction (real) + anomalous correction (imaginary) (complex) = $\mathbf{f_H(hkl)}$

(no phase shift) *(180° phase shift)* *(-90° phase shift)*



Multiple wavelength Anomalous Dispersion

Phasing by anomalous scattering: MAD

- Presence of heavy atoms with strong anomalous signal (soaking or co-crystallizzation, "natural" incorporation)
- Only one crystal is required
- Three data sets ⇒ (three different wavelengths)
- Cryogenic conditions
- Tunable radiation
- Wavelengths are carefully chosen to optimize the difference in intensity Bijvoet pairs and between the diffraction at selected wavelengths

Multiple wavelength Anomalous Dispersion

Anomalous scattering

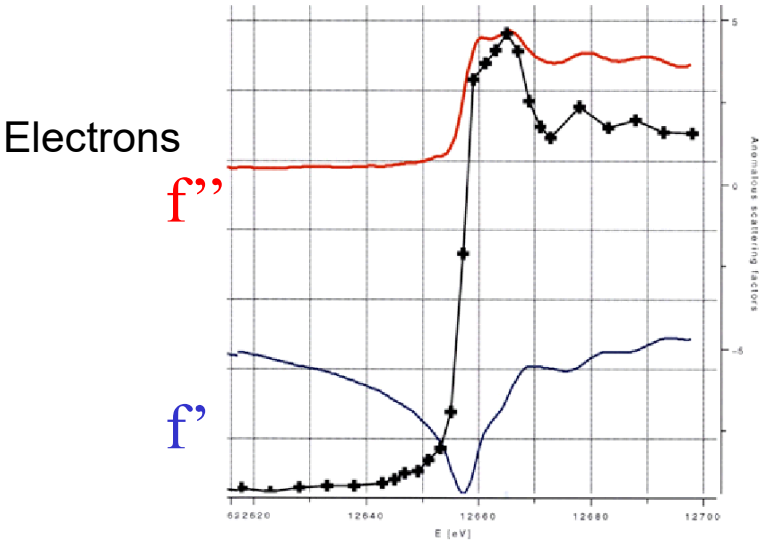
The atomic scattering factor at any λ may be written:

$$f(\lambda) = f_0 - \delta f'(\lambda) + i f''(\lambda)$$

f_0 is the conventional atomic scattering factor for λ far from an absorbtion edge

$\delta f'(\lambda)$ is the amount by which f_0 is reduced at wavelength λ (Absorption/Dispersion)

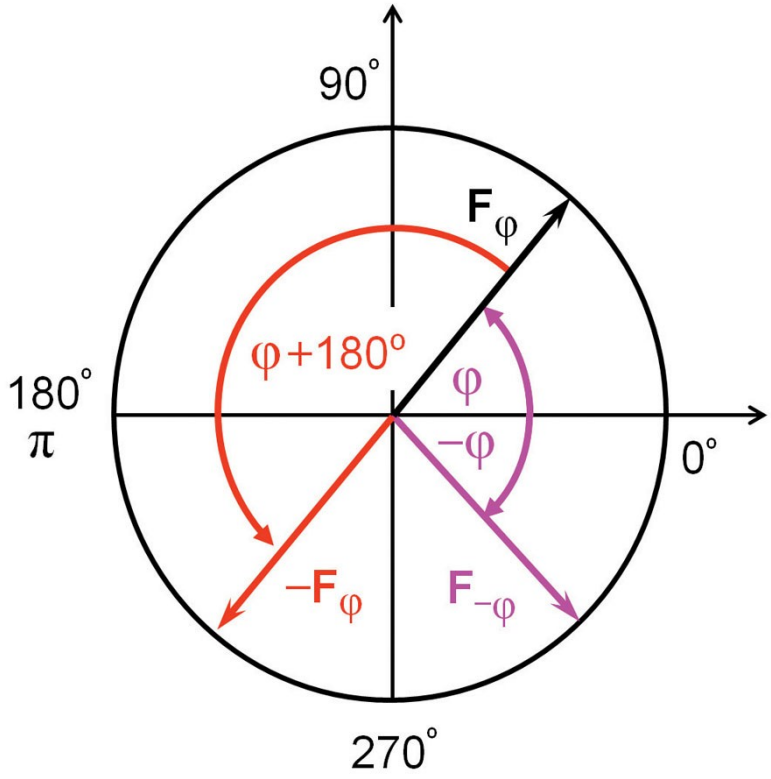
$f''(\lambda)$ is the amount of 'anomalous scattering', the out-of-phase component of the scattering at this λ



Edge	E (keV)	λ (Å)
K	12.6578	0.9795
L-I	1.6539	7.4965
L-II	1.4762	8.3989
L-III	1.4358	8.6352

Element **Se** (Z = 34) Atomic mass: 78.9600 Energy

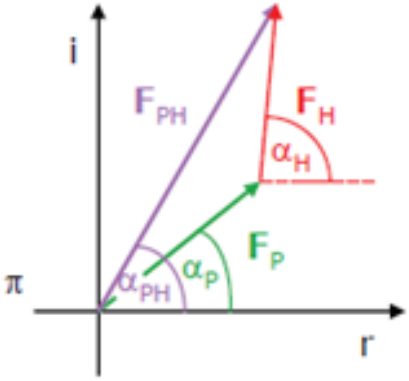
Conjugate versus negative structure factors

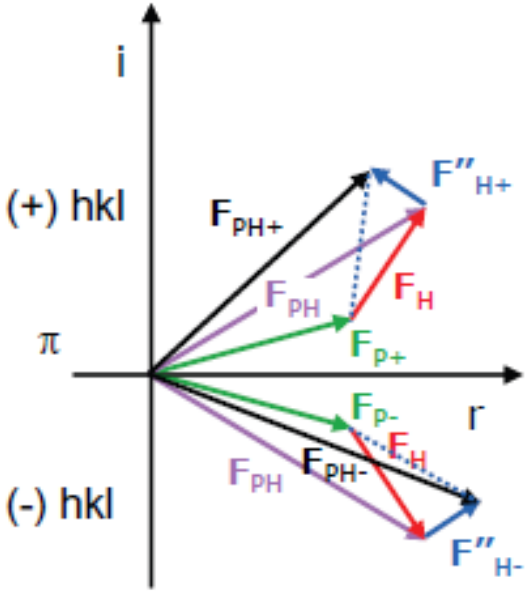


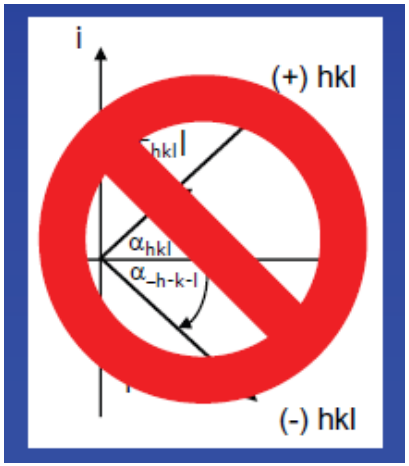
For a given structure factor **F** with phase φ , the conjugate or "inverse" structure factor **F**(- φ) has conjugate phase –φ, while negative structure factor –**F** have phase φ +π.

Multiple wavelength Anomalous Dispersion

Anomalous scattering







The Friedel's law

$I_{hkl} = I_{-h-k-l}$, $F_{hkl} = F_{-h-k-l}$ and $\alpha_{hkl} = -\alpha_{-h-k-l}$ is not valid

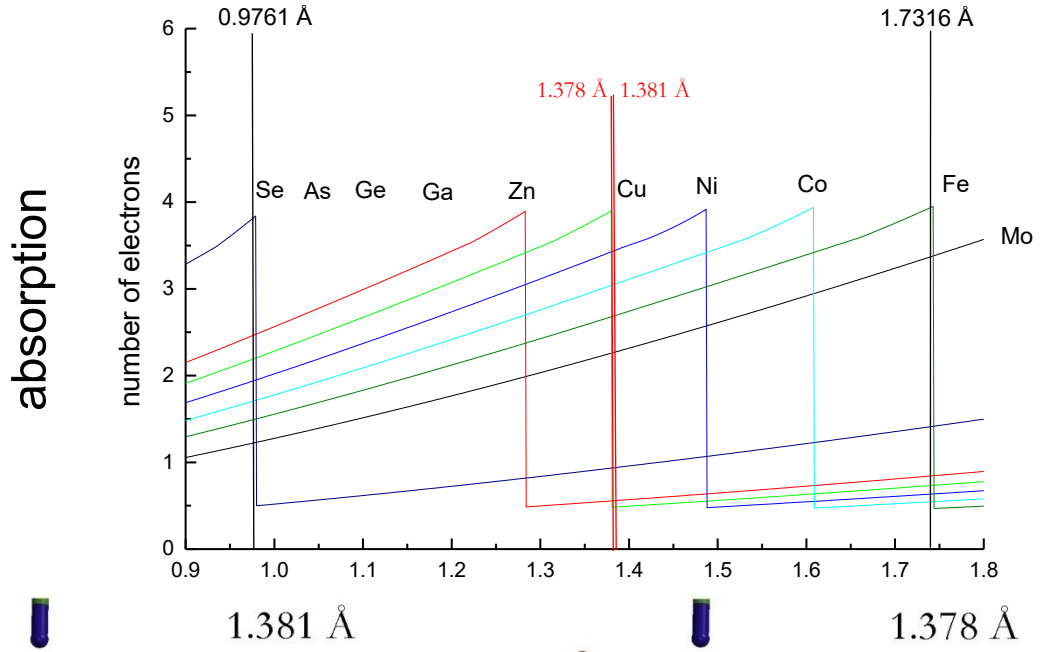
F_H real contribution from anomalous scattering atoms

F''_{H+} and F''_{H-} imaginary contribution from anomalous scattering atoms

Multiple wavelength Anomalous Dispersion: An analytical method

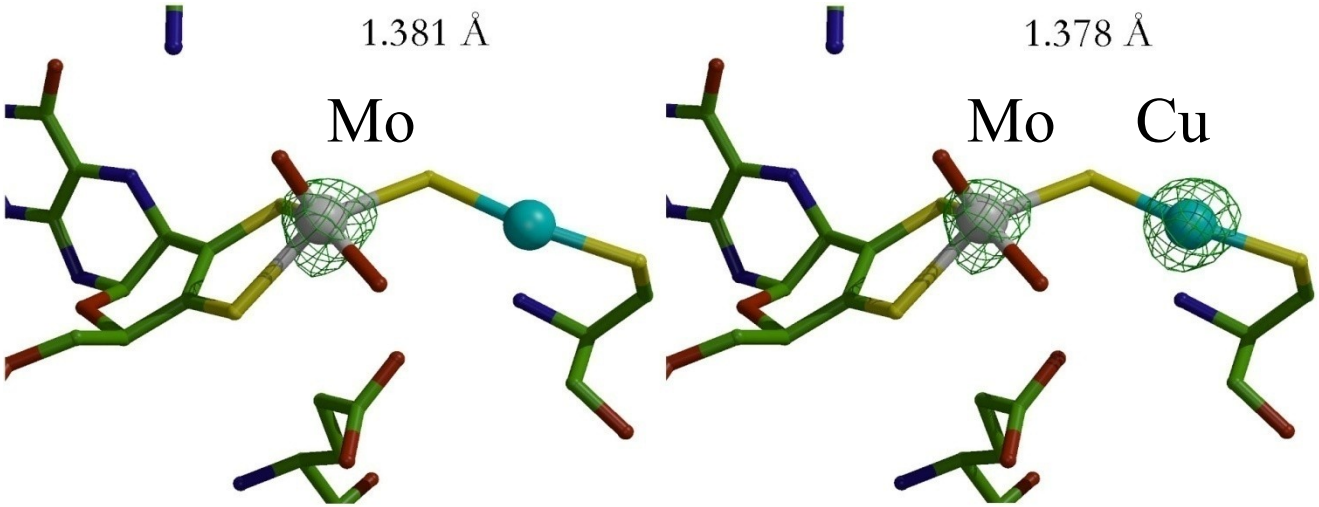
Mo

Edge	keV	Å
K	19.9995	0.6199
L-I	2.8655	4.3268
L-II	2.6251	4.7230
L-III	2.5202	4.9196



Cu

Edge	keV	Å
K	8.9789	1.3808
L-I	1.0961	11.3114



Anomalous Difference Fourier maps: Determining the element specific anomalous absorption edge

Multiple wavelength Anomalous Dispersion

Top 5 reasons to use Anomalous Scattering in solving a macromolecular structure

How many crystallization plates does it take to find a decent heavy atom derivative?

- ✓ One less heavy atom derivative to find. Breaking the phase ambiguity by isomorphous replacement requires 2 derivatives. But SIR+AS requires only one. The search for heavy atom derivatives can be time consuming.
- ✓ Perfect isomorphism. Heavy atom derivatives are often non-isomorphous with native crystals, and this degrades the phase accuracy. But, anomalous differences are taken within the same xtal. Good phasing power.
- ✓ Can take advantage of seleno-methionine derivatives which produce a good anomalous signal but very little isomorphous differences. Derivatization involves less trial and error. Some proteins just don't have the reactive groups necessary to bind heavy atoms.
- ✓ Determines the correct handedness of the electron density (x,y,z) vs. (-x,-y,-z).
- ✓ You get to go on a synchrotron trip. Fun!

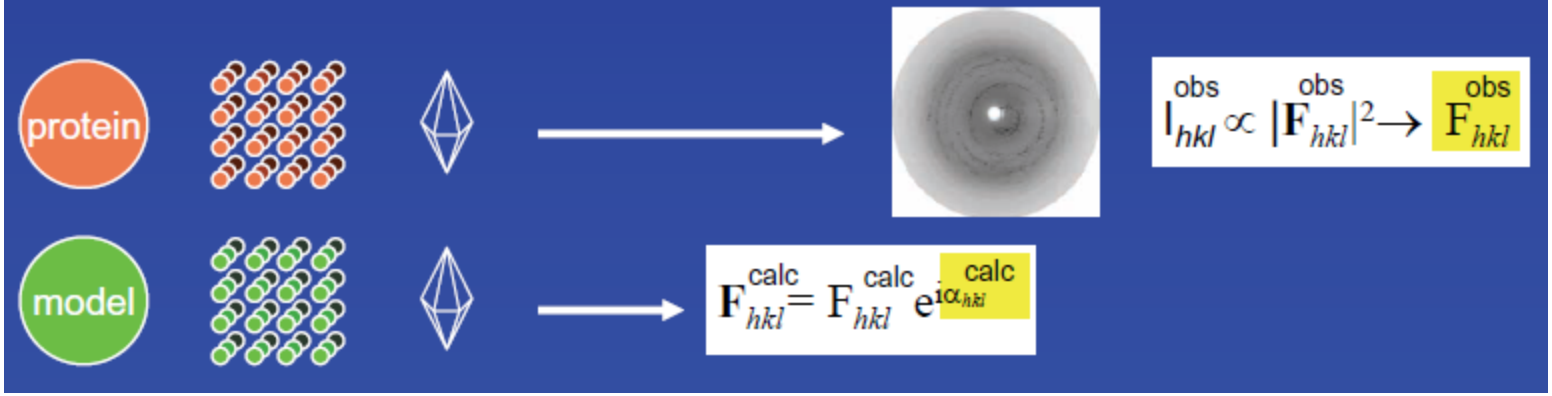
Molecular Replacement

A **good model** for the protein of unknown structure is needed ($\geq$ 30% sequence identity).

Model's phases are grafted onto the intensities which are experimentally determined.

$$\rho(x,y,z) = \frac{1}{V} \sum_{hkl} \left(|F_{hkl}^{obs}| \, e^{i\alpha_{hkl}^{calc}} \right) e^{-2\pi i(hx+ky+lz)}$$

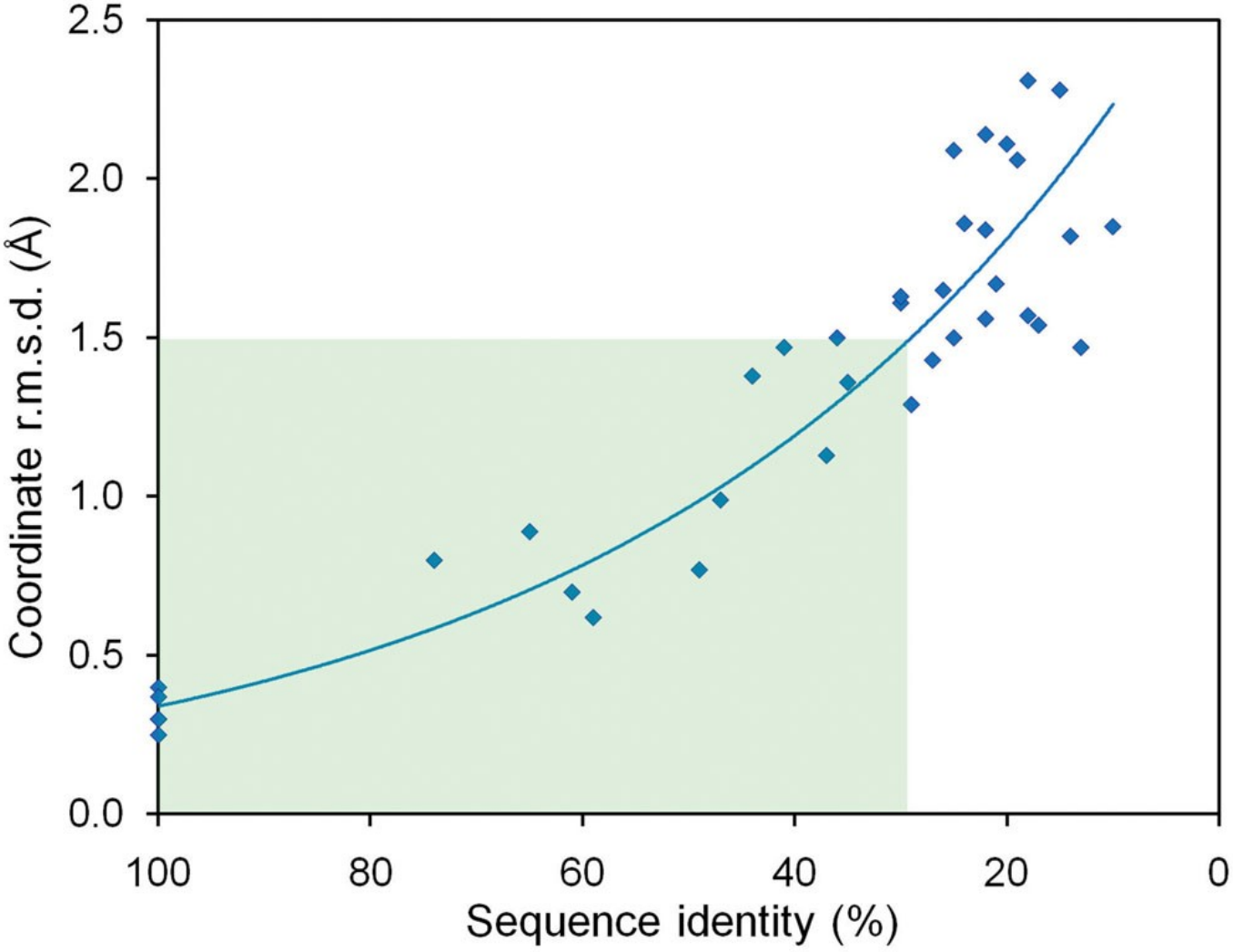
The "calculated" phases can be obtained by simulating the molecule's packing in the crystal (using the model protein) and obtaining theoretical phases.



$$I_{hkl}^{obs} \propto |F_{hkl}^{obs}|^2 \rightarrow F_{hkl}^{obs}$$

$$F_{hkl}^{calc} = F_{hkl}^{calc} e^{i\alpha_{hkl}^{calc}}$$

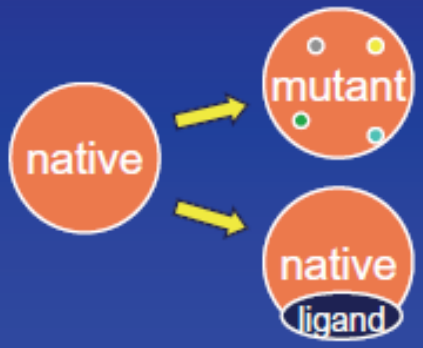
Molecular Replacement:
Model template



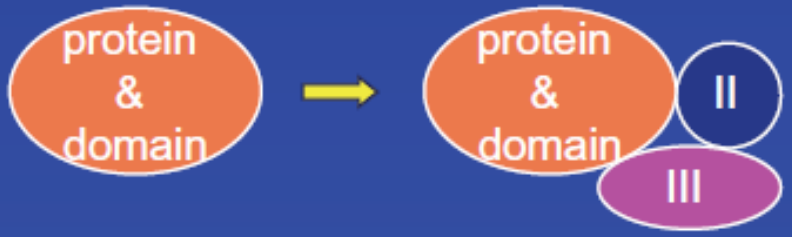
Molecular Replacement

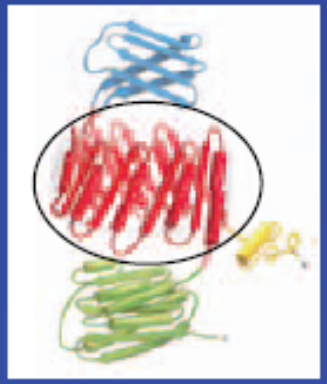
- homologous protein (from PDB)

- structural studies on mutant proteins and complexes (protein + ligands)



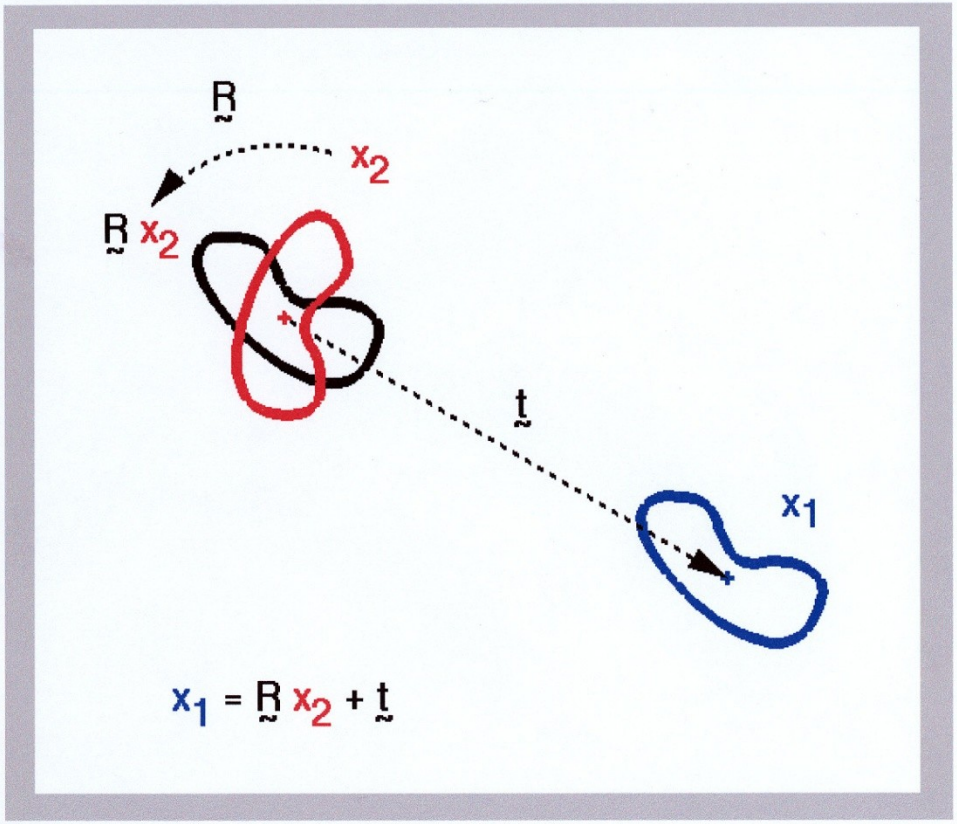
- multi-domain protein multi-protein complex





3- Patterson Search (Molecular Replacement)

- Phases are calculated using a homologous structure (phasing model)
- Phases are calculated once the position of the search model in the unit cell has been determined



X_1 – unknown or target

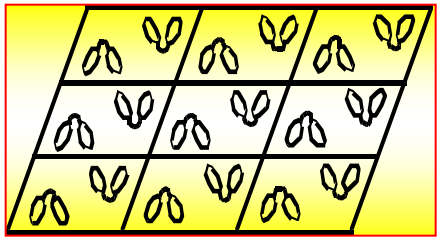
X_2 - known model

R - rotation matrix (θ_1, θ_2, θ_3)

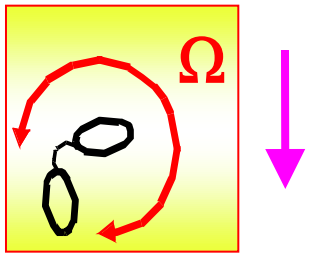
t - translation vector (t_x, t_y, t_z)

Molecular Replacement

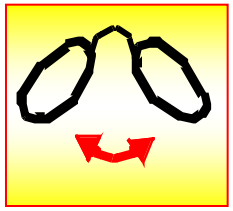
Crystal



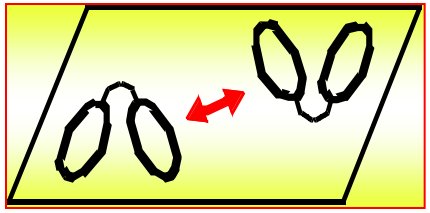
1.RotationSearch

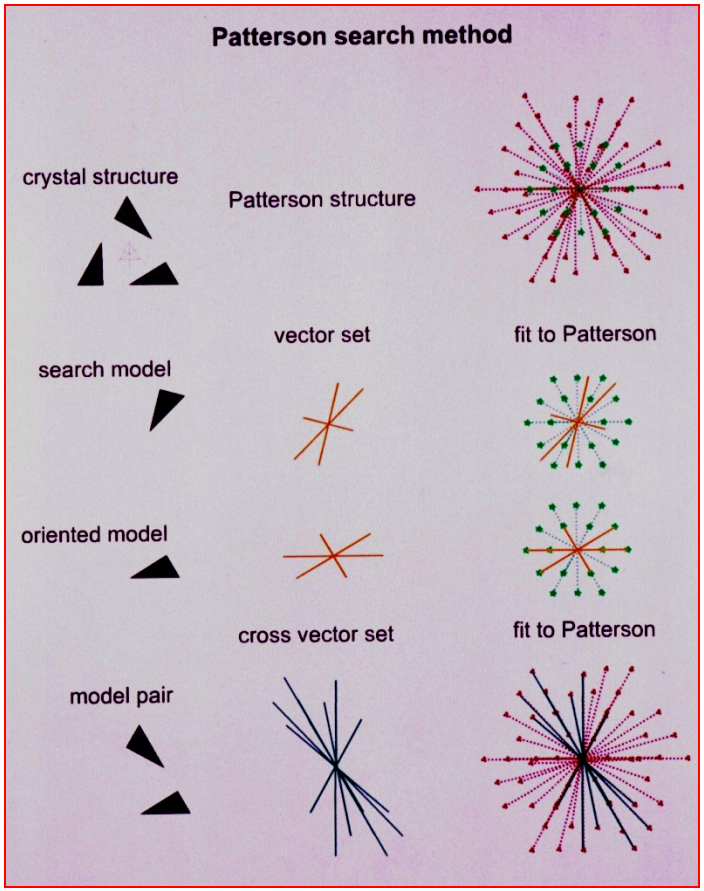


2. PC Refinement



3. Translation Search

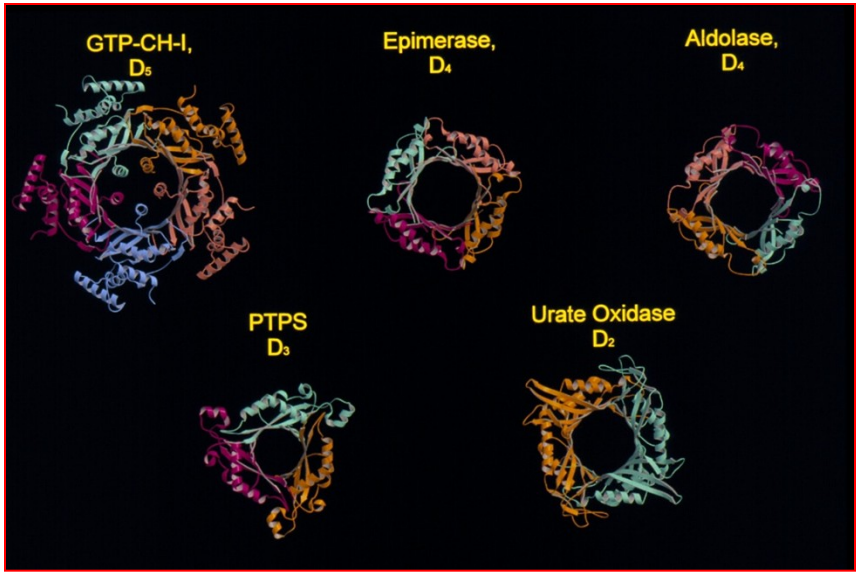




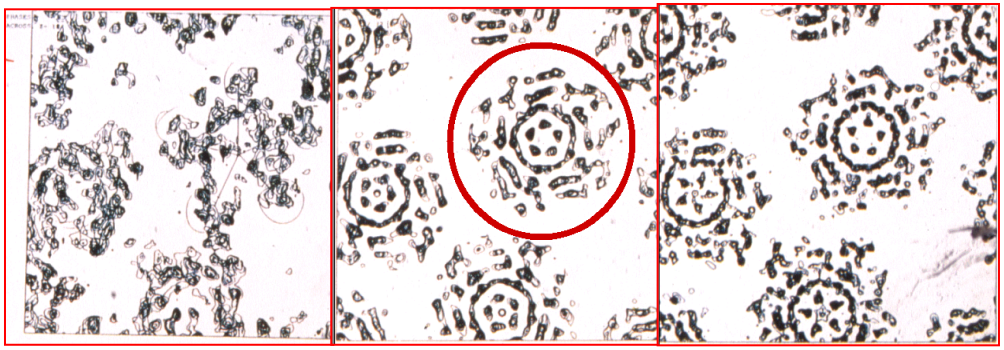
Methods to improve initial experimental phases

Non-crystallographic symmetry

Symmetrical oligomers



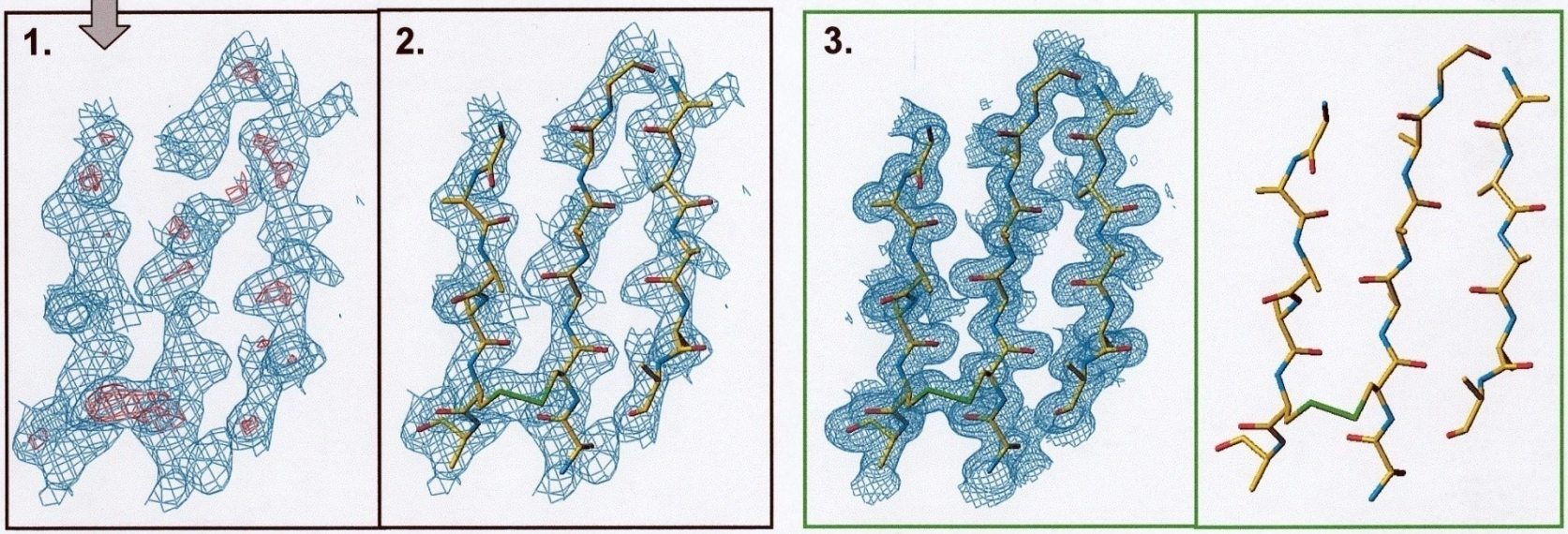
Cyclic Averaging



After obtaining initial experimental phases (α_{hkl}) a structural model can be built

A protein model (our idea of what is right) is fitted into the initial experimental electron density map ρ_{xyz} :

$$\rho_{(x,y,z)} = \frac{1}{V}\sum_h\sum_k\sum_l |F_{(h,k,l)}| \exp[-2\pi.i(hx+ky+lz-\alpha_{(h,k,l)}]$$



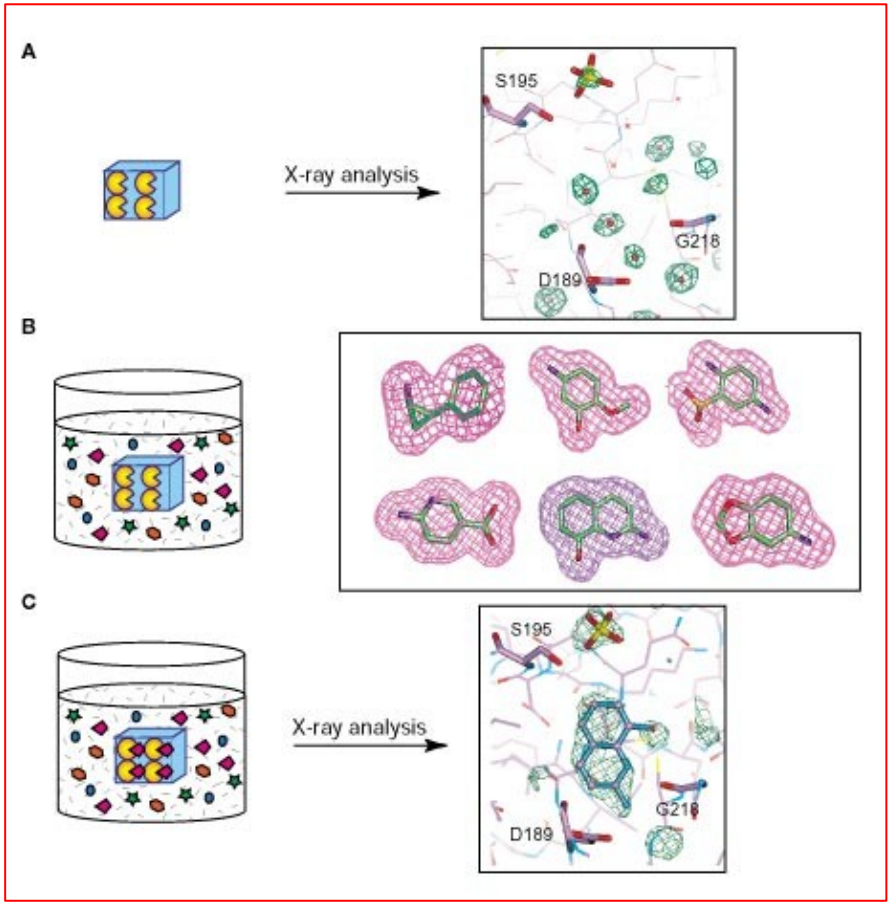
1. What we "see" contour map of the e⁻ density

2. We build a model (that makes sense)

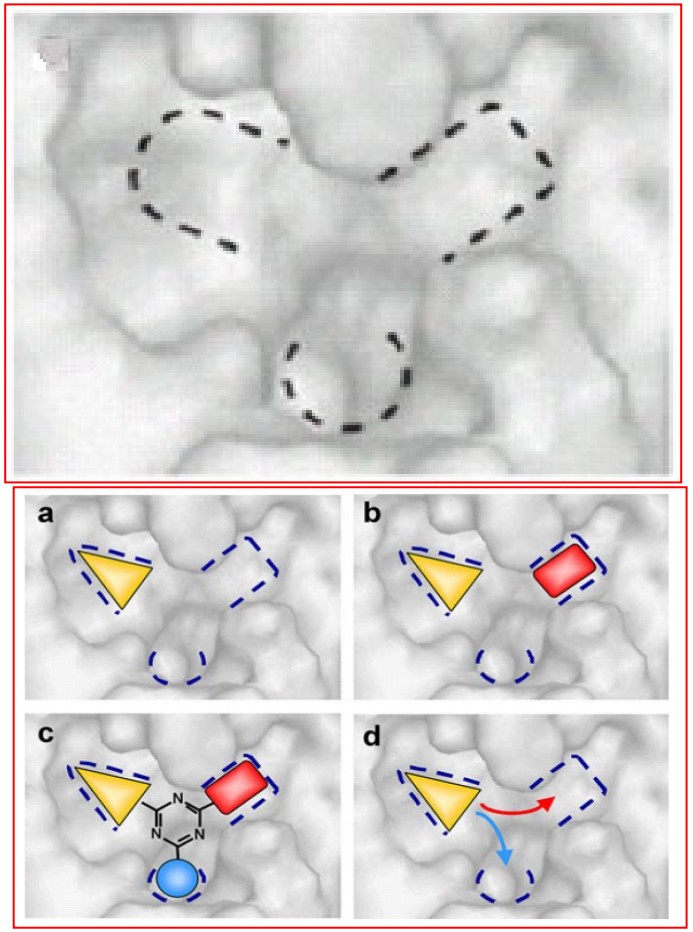
3. The final refined model (better phases)

Model (what is published)

From Molecules to Medicines: Integrating Protein X-ray Crystallography in Drug Discovery



Crystallographic screening: Cocktail Soaking



Binding site comprising three possible binding pockets

Nat. Biotech. (2000) 18, 1105

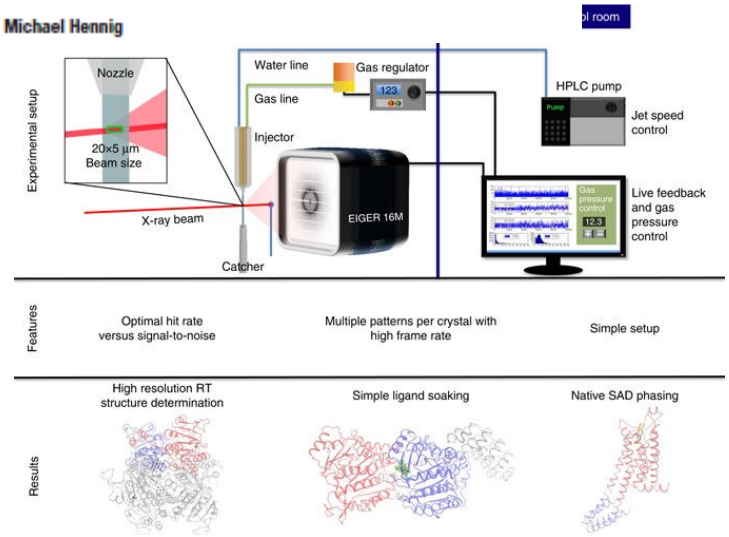
Essays in Biochemistry (2017) 61 529–542
https://doi.org/10.1042/EBC20170031

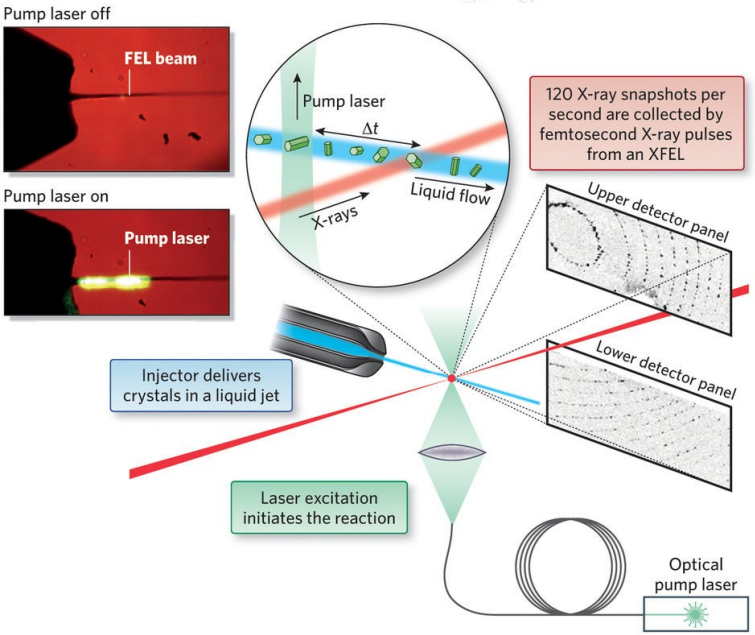
PORTLAND PRESS

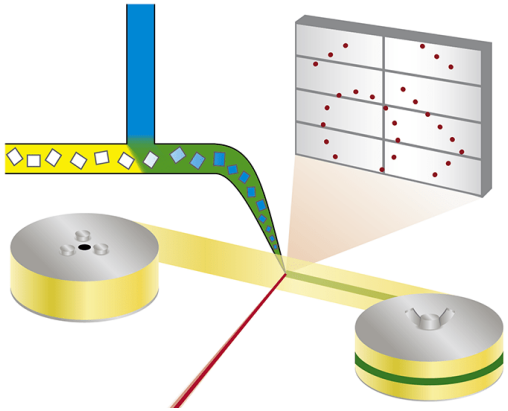
Review Article

X-ray free electron laser: opportunities for drug discovery

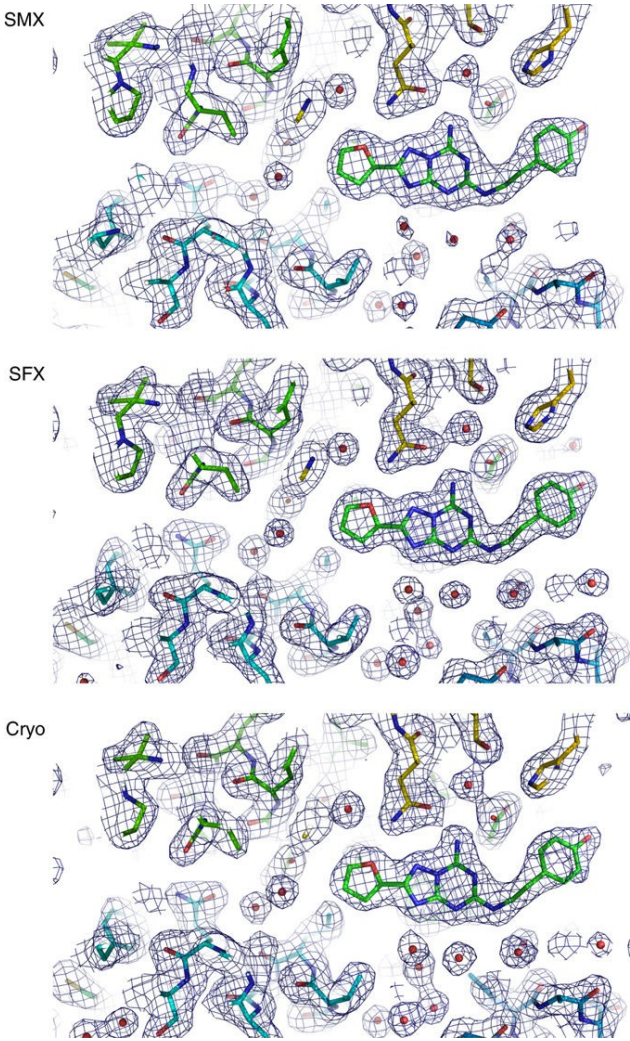
Robert K.Y. Cheng, Rafael Abela and Michael Hennig







IUCrJ (2017) **4**, 769–777



Serial Millisecond Crystallography for routine room-temperature structure determination at synchrotrons.

Mix-and-diffuse : Serial Synchrotron Crystallography – Drug Discovery