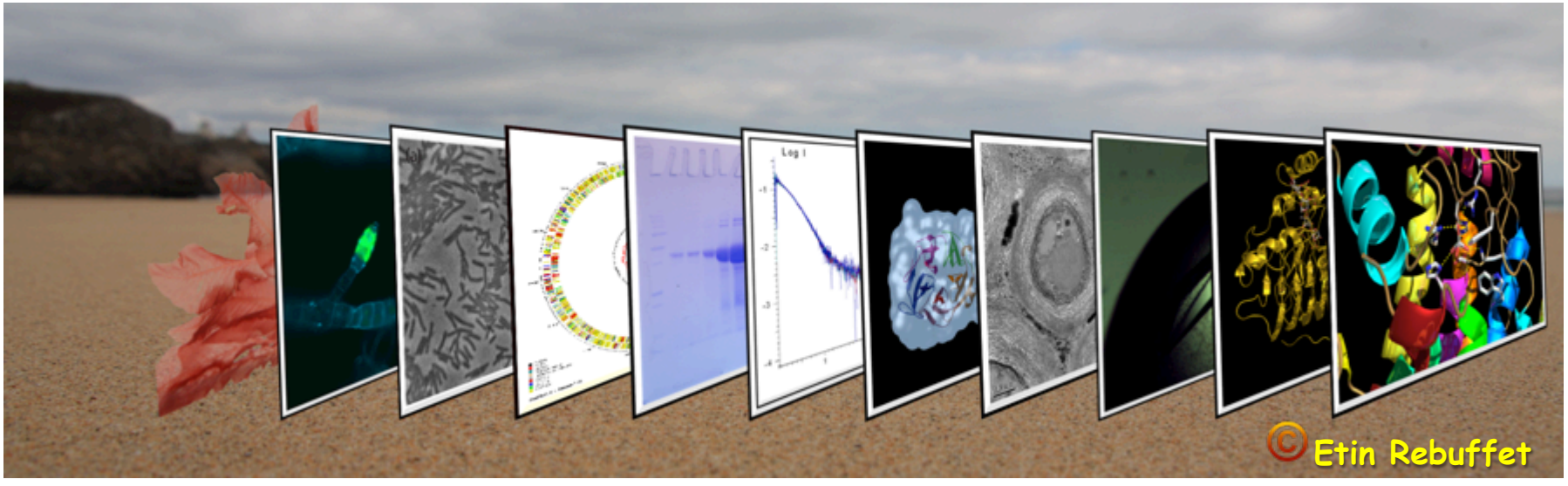
Mirjam Czjzek Research Director of CNRS

Station Biologique de Roscoff
(CNRS-Sorbonne University Paris VI)

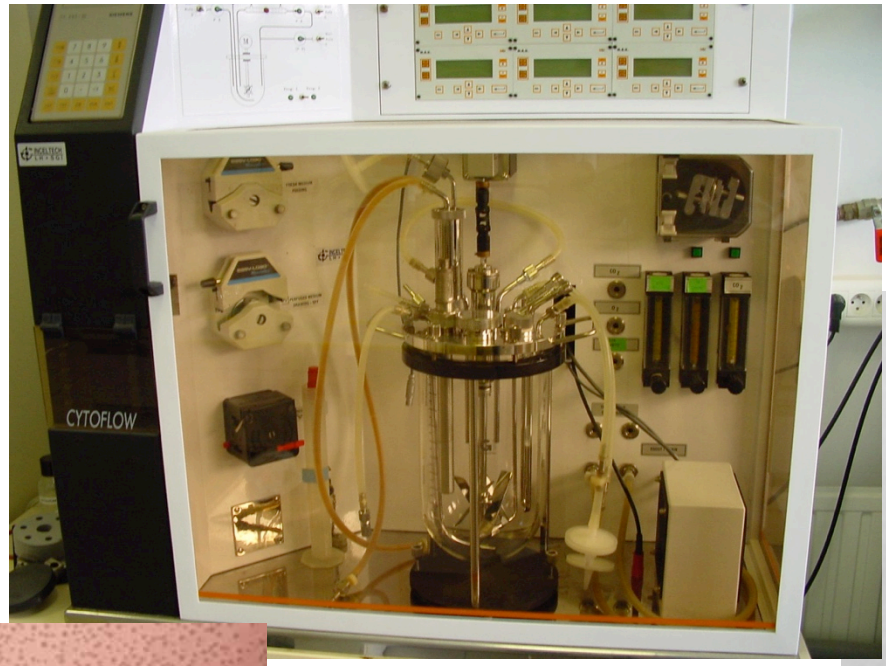
czjzek@sb-roscoff.fr

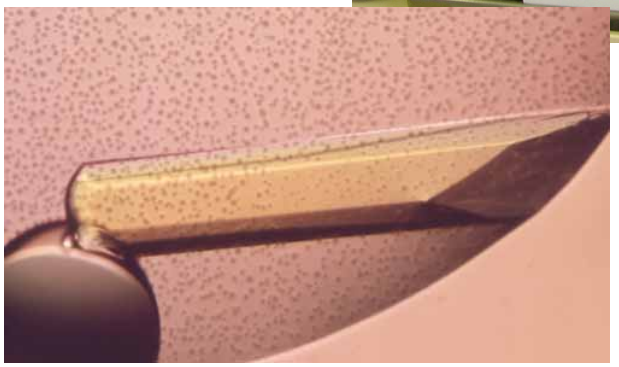
Marine Glycobiology

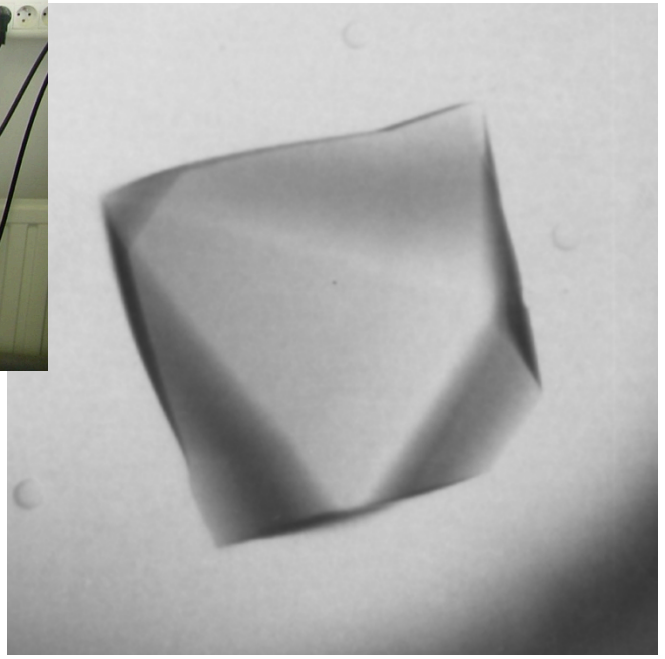
The algal associated « microbiome » : the metabolism of macro-algal polysaccharides



Sample preparation and crystallogenesis



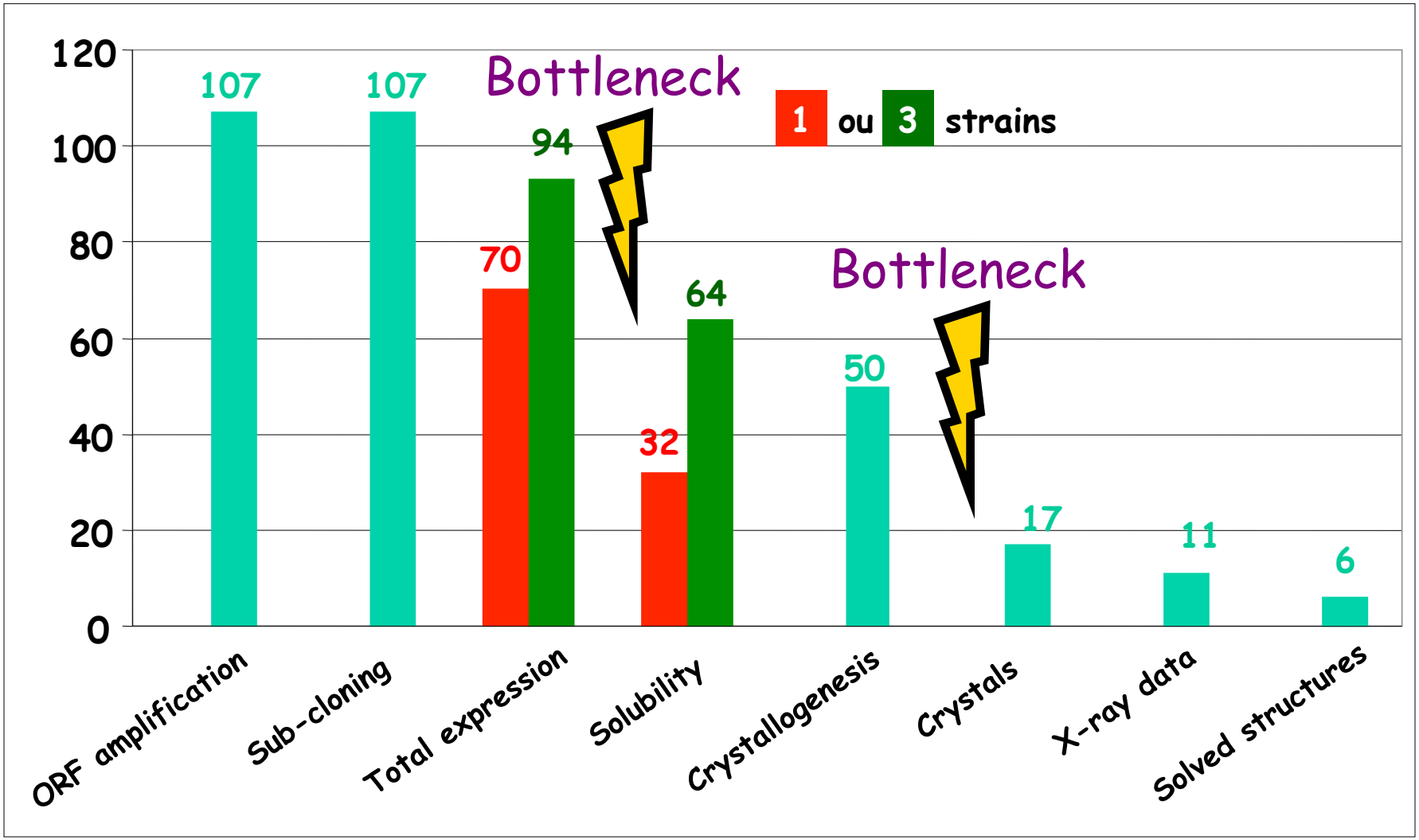




Critical steps to success....

- Overexpression of the protein/complex
- Purification
- Grow crystals….
- **Collect data**
  in the laboratory
  at the synchrotron (ESRF, SOLEIL,…)
- Solve the phase problem
- Analyse the electron density and construct the model
- analyse, compare and publish ……..

Structural Genomics Projet at AFMB : june 2001 - may 2002



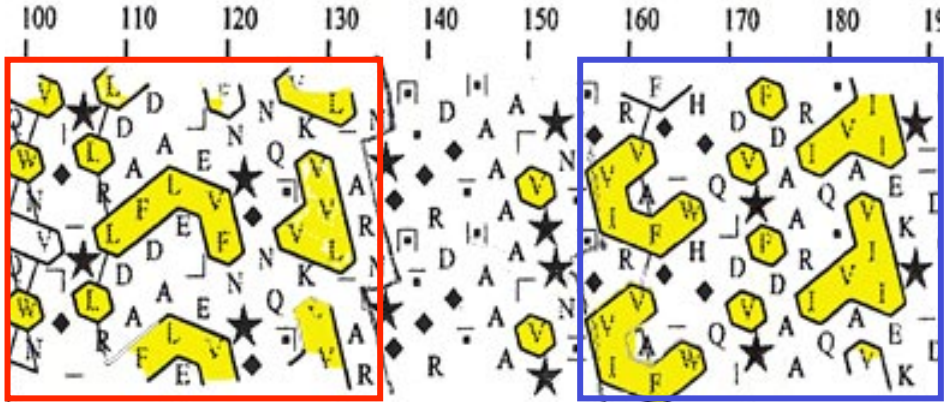
courtesy by R. Vincentelli & A. Gruez

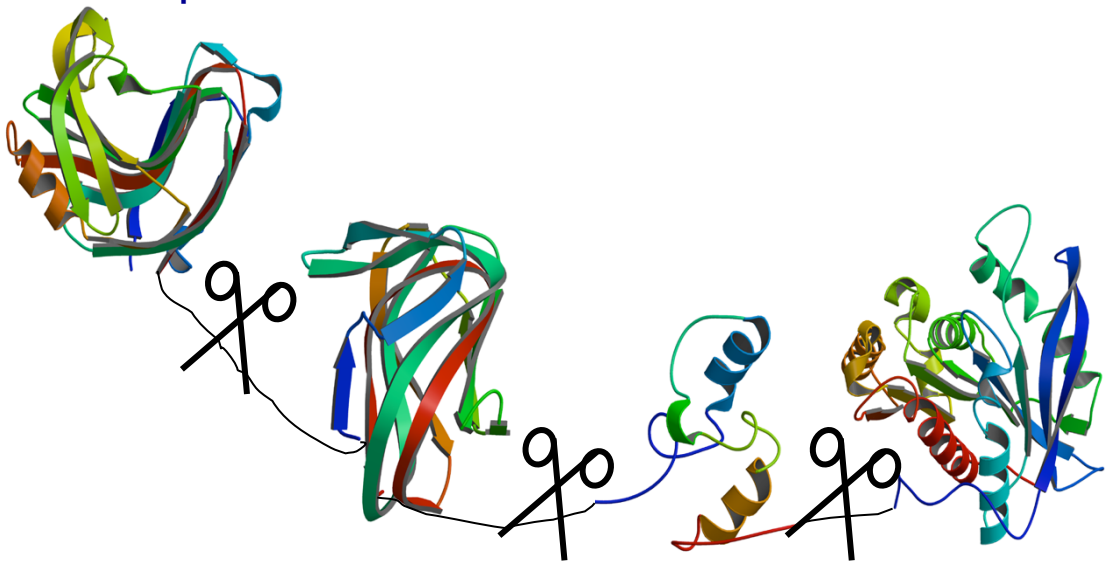
sequence analysis and various constructions of multimodular proteins

Protein families

- globular proteins
- fibrous proteins
- membrane proteins
- intrinsically disordered proteins
- multi-modular proteins
- multi-protein complexes

Bioinformatic analyses





xylanase CBM-6 dockerin carbohydrate esterase

Protein quality already conditioned by over-expression

recombinant or endogenous

bacterial expression

advantage "easy", in general
great over-expression,
low protease activity,
no post-translational
modifications

disadvantage protein solubility, lack
of post-translational
modifications,
yield, abundancy

eukaryotic expression

advantage protein solubility
post-translational
modifications

disadvantage expensive, low yield,
proteases, time consuming

protein isolated from native source

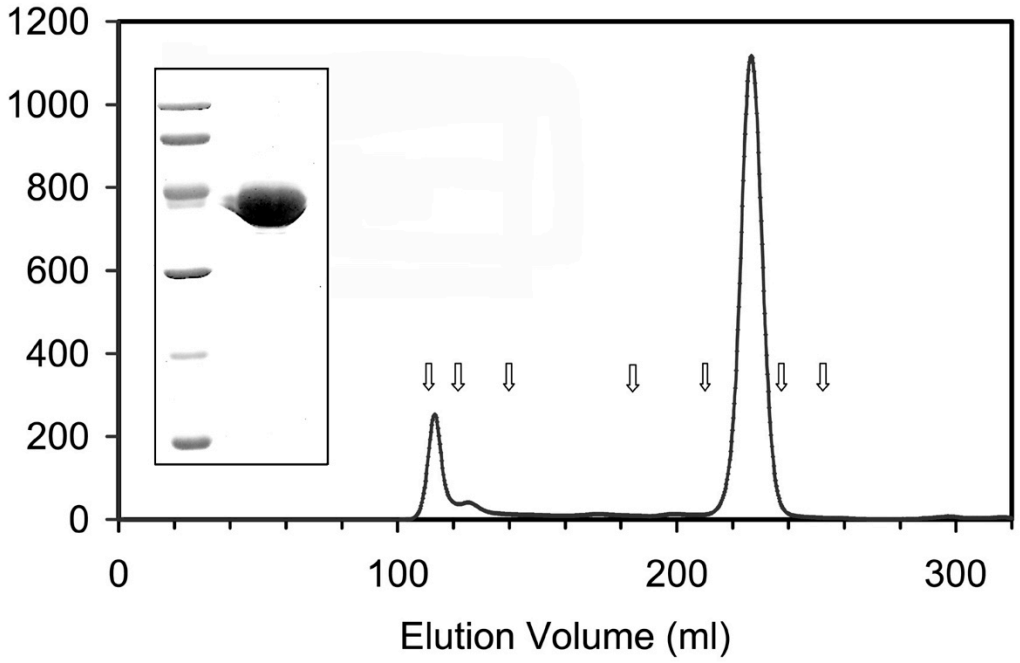
advantage protein solubility,
authenticity
(cofactor, ions, ligand,
oligomeric state)

disadvantage expense/effort,
yield, abundancy

Importance of the purity of the protein

Is it pure, could it be glycosylated?
Heterogenity will inhibit crystal growth
Do a mass spectrum

Is it correctly folded?
monodispers? **monomeric**?
activity test
do CD or 1D-NMR
and/or DLS, SAXS,
ultracentrifugation

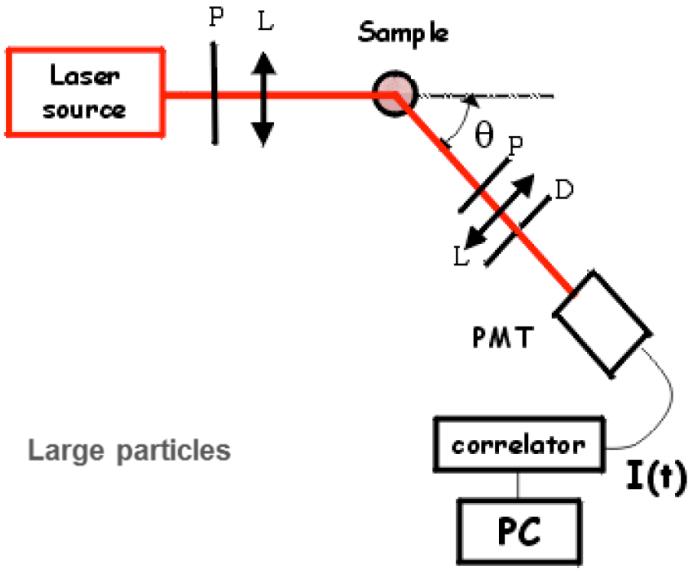


Is it freshly prepared?
The protein may degrade/denature with time

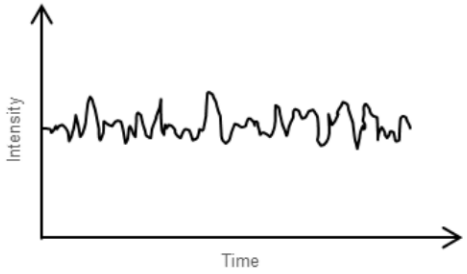
Is it stable?
define storage temperature, add salt, glycerol, reducing agent…

Methods for “quality control”

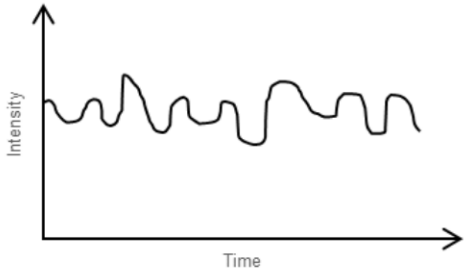
DLS



Small particles

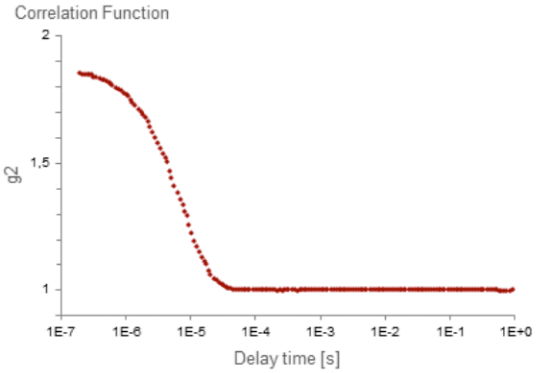


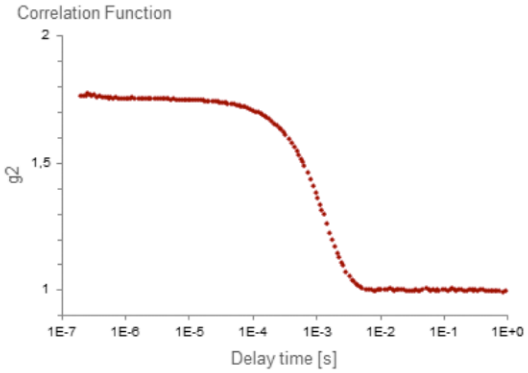
Large particles

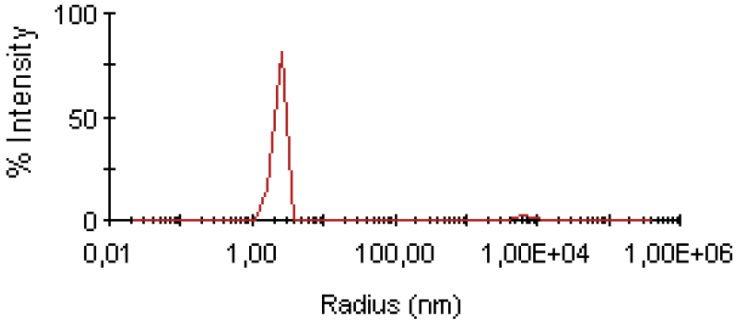


gives measure of

R_h hydrodynamic radius



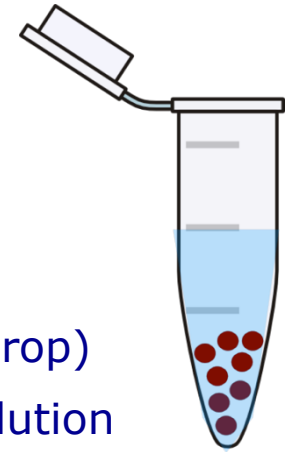




solubility of protein

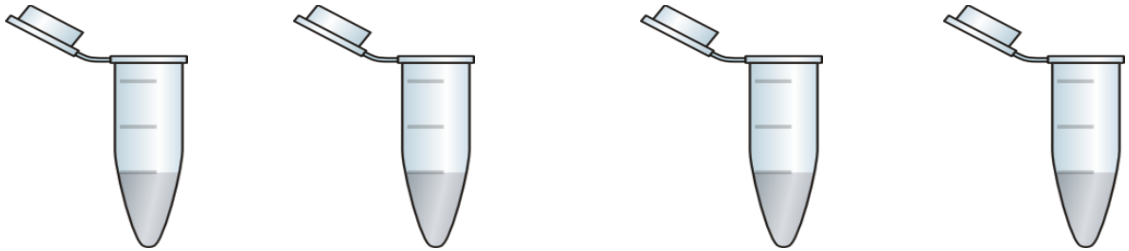
or the standard lab buffer is not always the best...

- Protein concentration step is critical



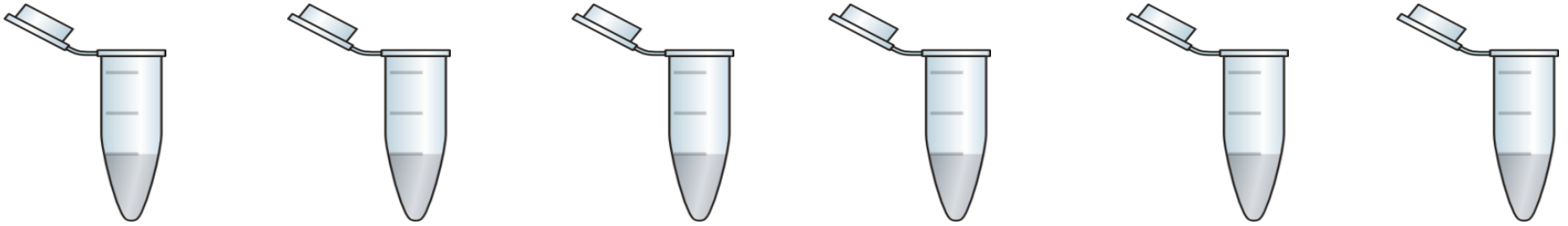
Solubility screen

1) precipitate your protein (dialysis or PEG)
2) transfer flocculate precipitate into several tubes
3) add different buffer/salt solutions
4) measure the re-solubilized protein quantity (nanodrop)
5) next time you do Gel-filtration use this buffer for elution



1) Na- MES pH 5.6 | PIPES pH 6.5 | HEPES pH 7.5 | TAPS pH 8.5

nanodrop
DLS



2) NH_4Cl | NaCl | KCl | LiCl | $MgCl_2$ | $CaCl_2$

Collins et al Acta Cryst D 2004

Principles of crystallisation

starting phase		Crystal	small molecules	Proteins
solid phase	->	solid phase	+	-
liquid phase	->	solid phase	+	-
vapor phase	->	solid phase	+	-
solution	->	solid phase	+	+

solvent categories : (ε = dielectric constant)

1. dipolar protic solvents have $\varepsilon > 15$
2. dipolar aprotic solvents have $\varepsilon > 15$
3. non-polar and necessarily aprotic solvents have $\varepsilon < 15$

Proteins $\Longrightarrow$ solvent based on H_2O (category 1, $\varepsilon = 80$)

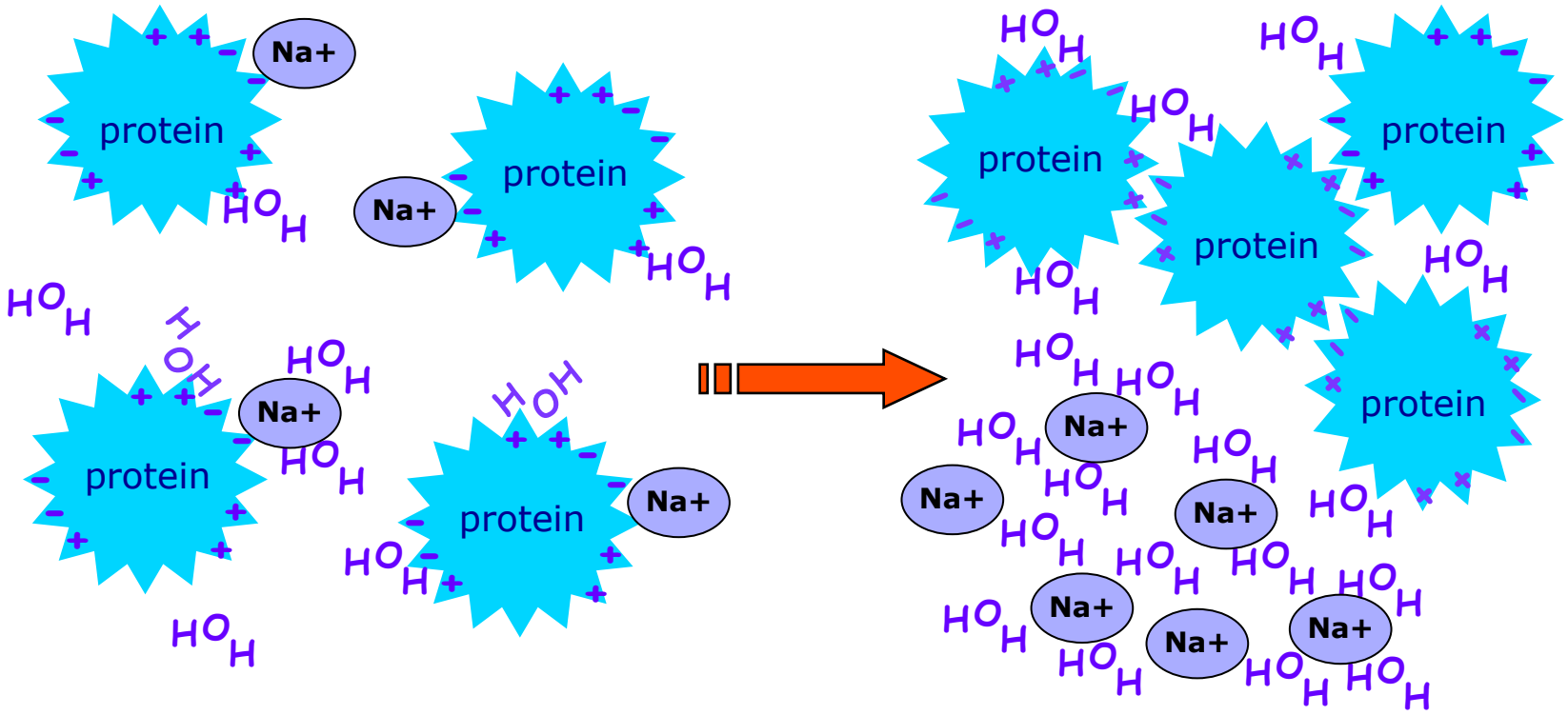
+ additives that influence the polarity or the ε value of the solvent.

Interactions between solute and solvent

1. Ion-dipole
2. Dipole-dipole
3. hydrogen bonds
4. VDW interactions

Proteins can be considered as electrolytes (charged particles)

=> solubility decreases with size!



Aim of using precipitants

1. Interrupt the hydratation sphere of proteins
2. Decrease ε value of the medium to decrease effect of electrostatic screen between molecules.
3. Induce phase separation
4. establish a new free energy minimum (increase attractive charges and decrease repulsive charges).

Types of precipitants used

Salts :

competition for hydratation sphere

Solvents :

decrease the dielectric constant

Polymers (PEG) :

exclusion, induce phase separation

The most often used salts

Sulfate:	Sodium	Ammonium	Magnesium	Lithium
Citrate :	Sodium	Ammonium		
Chloride :	Sodium	Ammonium	Potassium	Calcium
Acetate :	Sodium	Ammonium		
Formiate :	Sodium		Lithium	
Phosphate :	Sodium	Ammonium	Potassium	

The most often used solvents

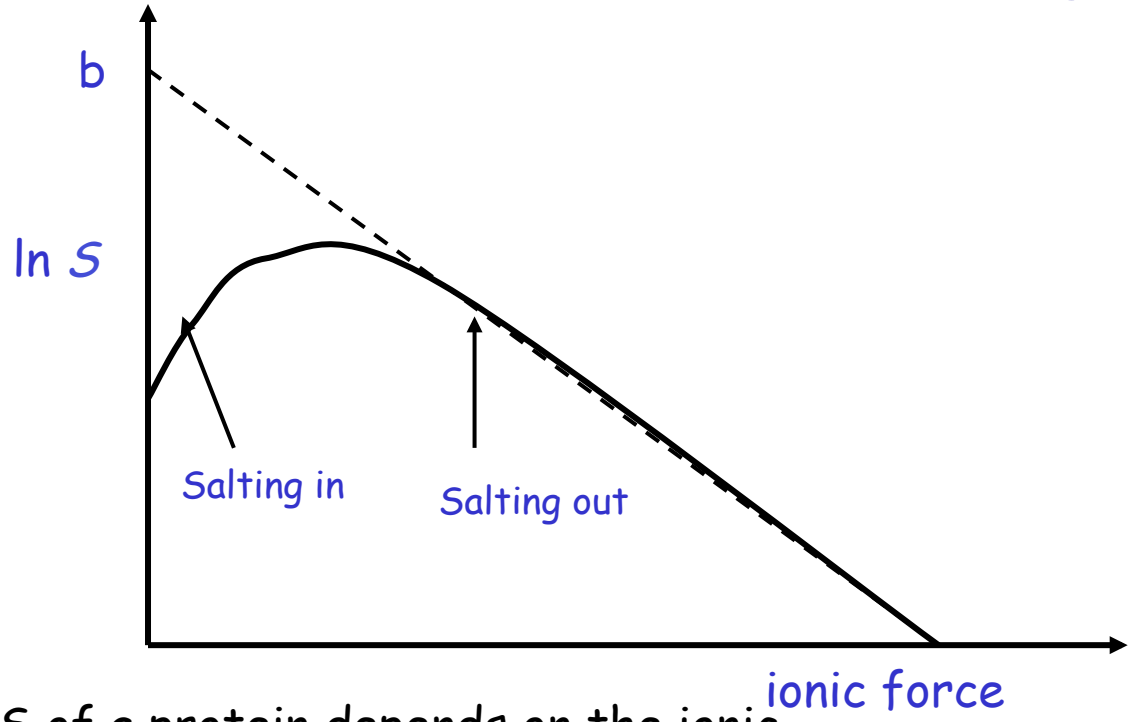
methanol; ethanol; isopropanol; 1,3-propanediol, butanol; 2-methyl-2,4-pentanediol (MPD); 2,5-hexanol; dioxane;...

The most often used PEGs

PEG 400 – 20 000 in Mw and some methylether derivatives.

Precipitants		Additive
0.5 – 2 Molar		10-100mM
	or	
20 – 70%		1-5%

Precipitation with salts (salting out)



The solubility S of a protein depends on the ionic force **I** of its environnement:

$$I = 1/2 \sum c_i z_i^2$$

z = number of charges

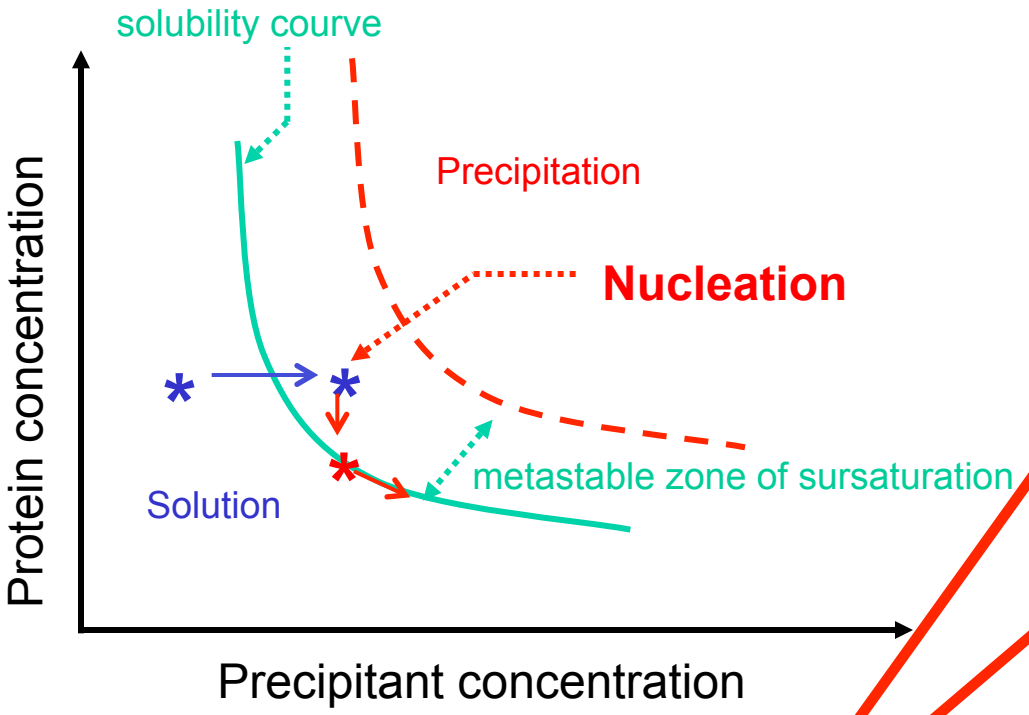
anions: $SO_4^{2-} > HPO_4^{2-} > CH_3COO^- > Cl^- > Br^- > NO_3^- > I^- > SCN^-$

cations: $Mg^{2+} > Ca^{2+} > Ba^{2+} > Li^+ > Na^+ > K^+ > NH_4^+$

Ionic force increases

depends on charge and ionic radius

The principals of the vapor diffusion method



Parameters that can be varied

Nature of precipitant
salt
polymers (PEG)
organic solvents

Ionic force
salt concentration
variation of pH
different buffers

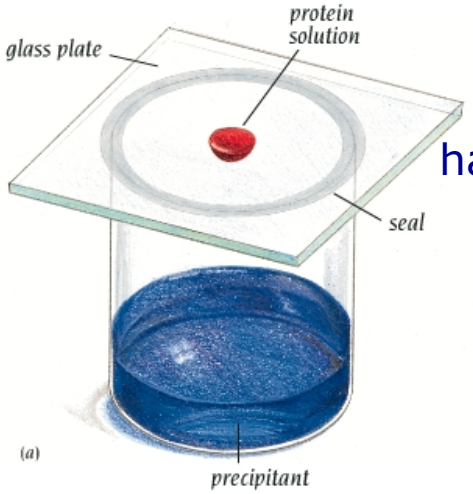
Additives
nature of cation/anion
chaotrop
solvent(alcohol,detergent)
substrates, cofactors

Others : concentration of protein, temperature, purity of protein…

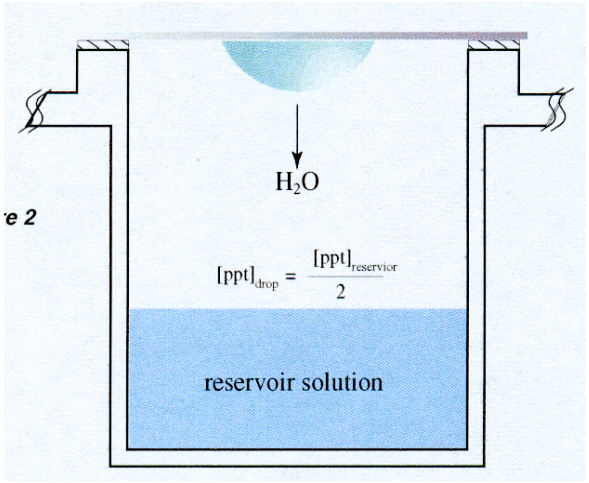
Screening crystallisation conditions

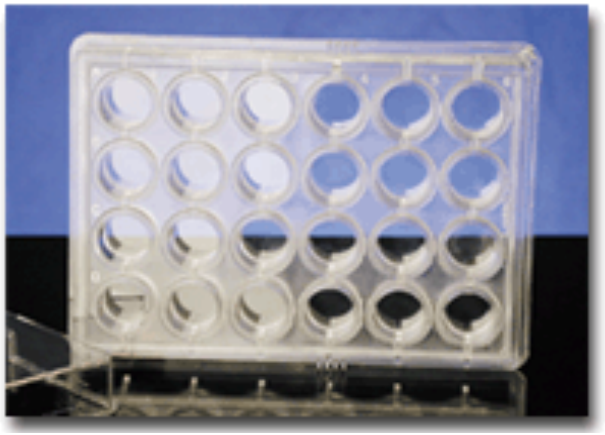
- most commercial kits are based on empirical analyses of a large number of example proteins (incomplete factorial)
- techniques at low scale are used to diminish the quantity of protein necessary

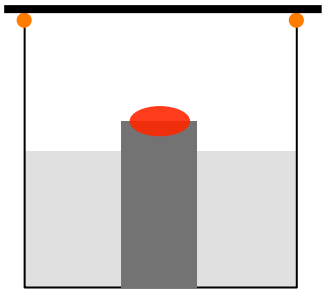
Vapor diffusion with hanging or sitting drop



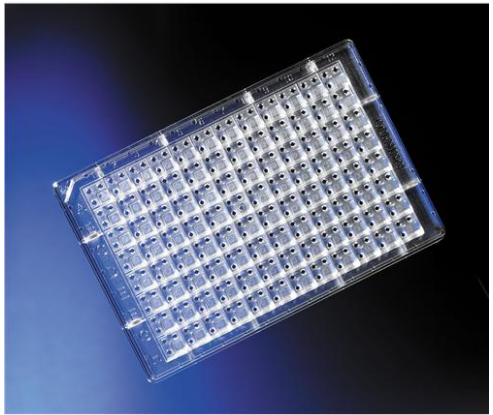
hanging



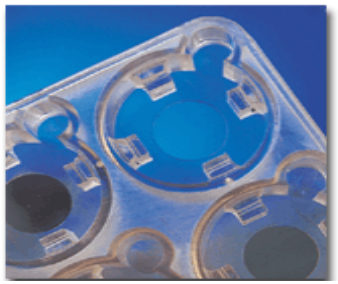


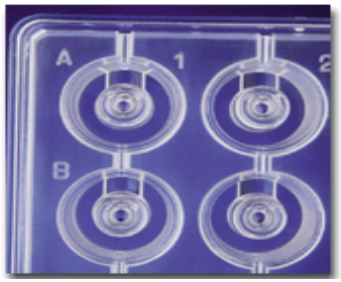


sitting

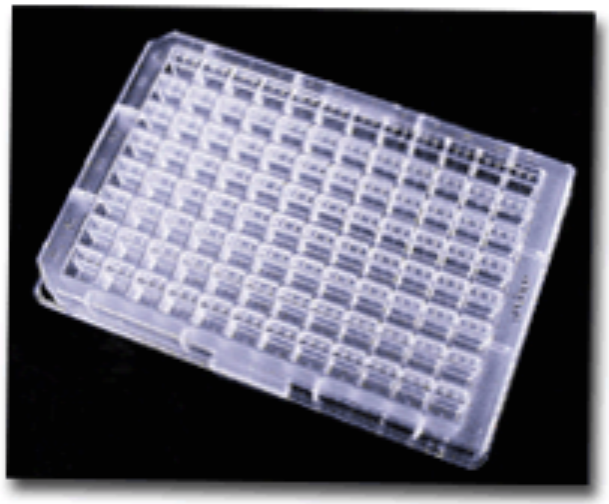


Other techniques



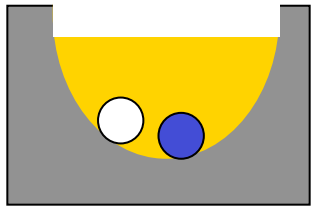


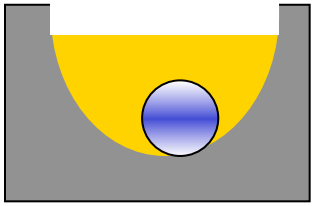
3 drops for one well



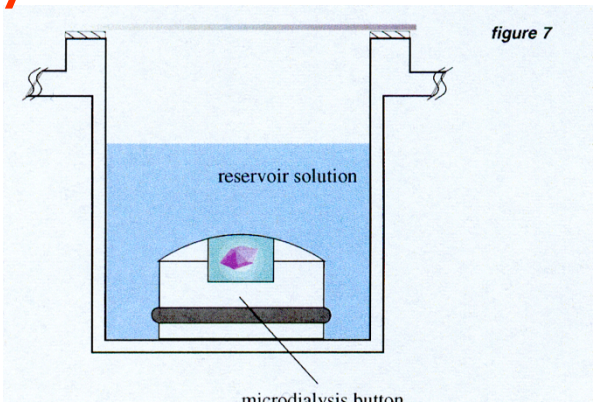
High Throughput
96 Well Crystallization!

Batch

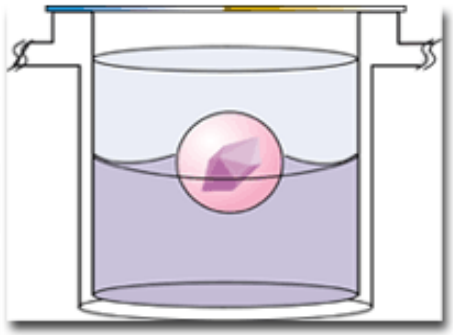




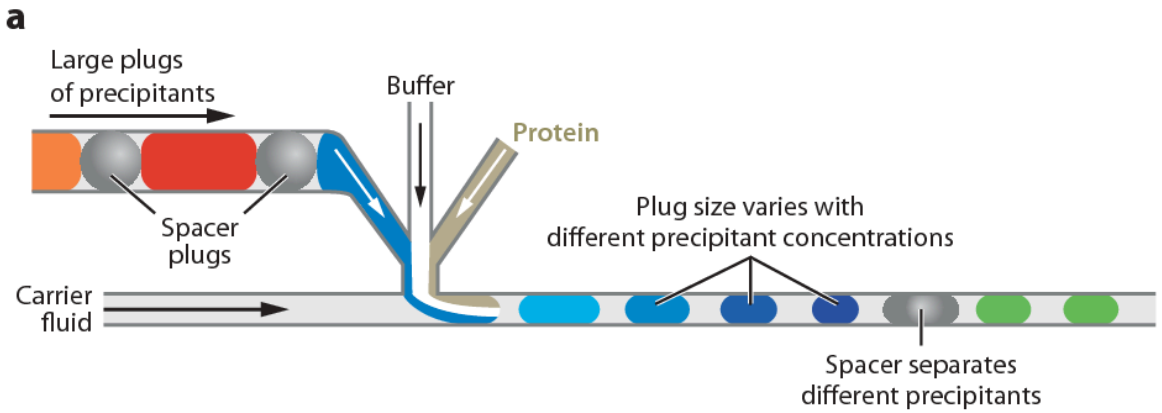
Dialysis

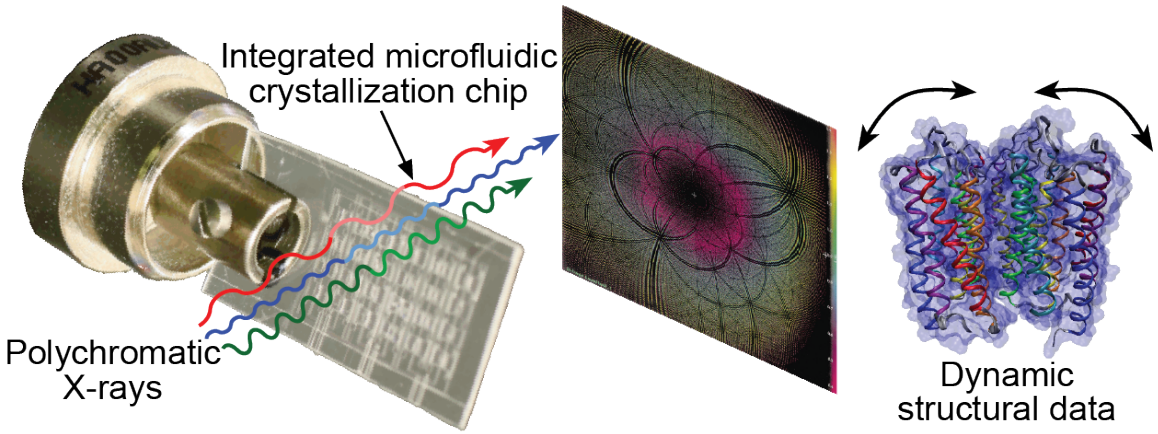


Crystallisation without contact to solid :

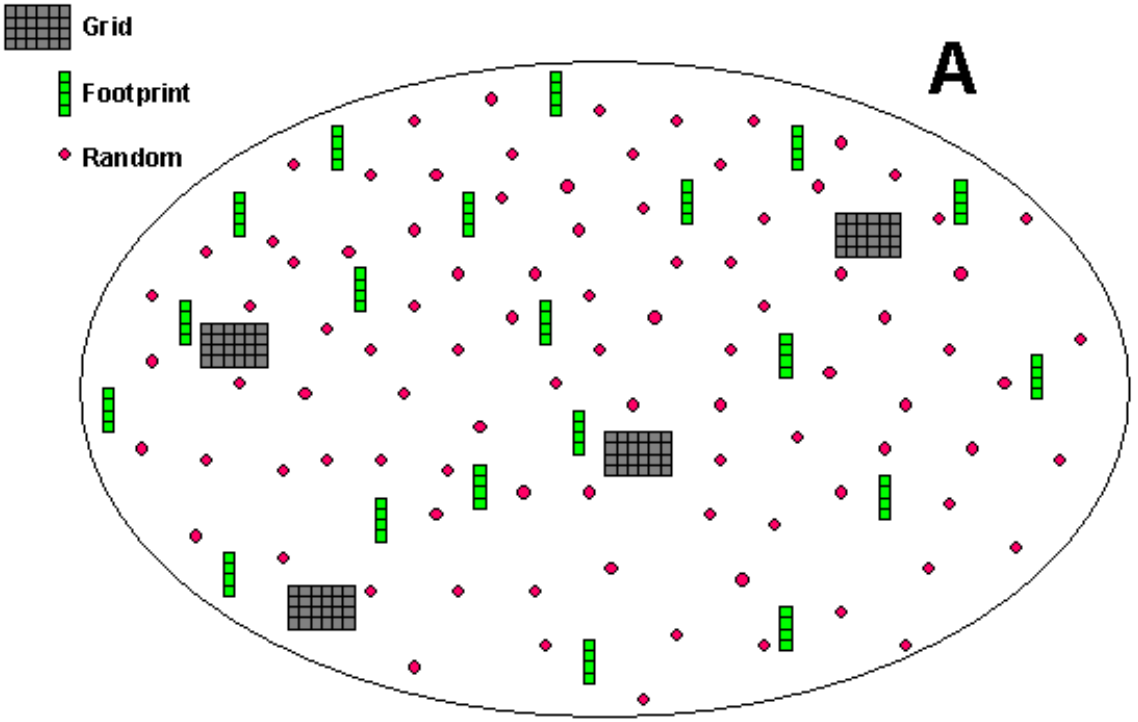


future methods : microfluidics





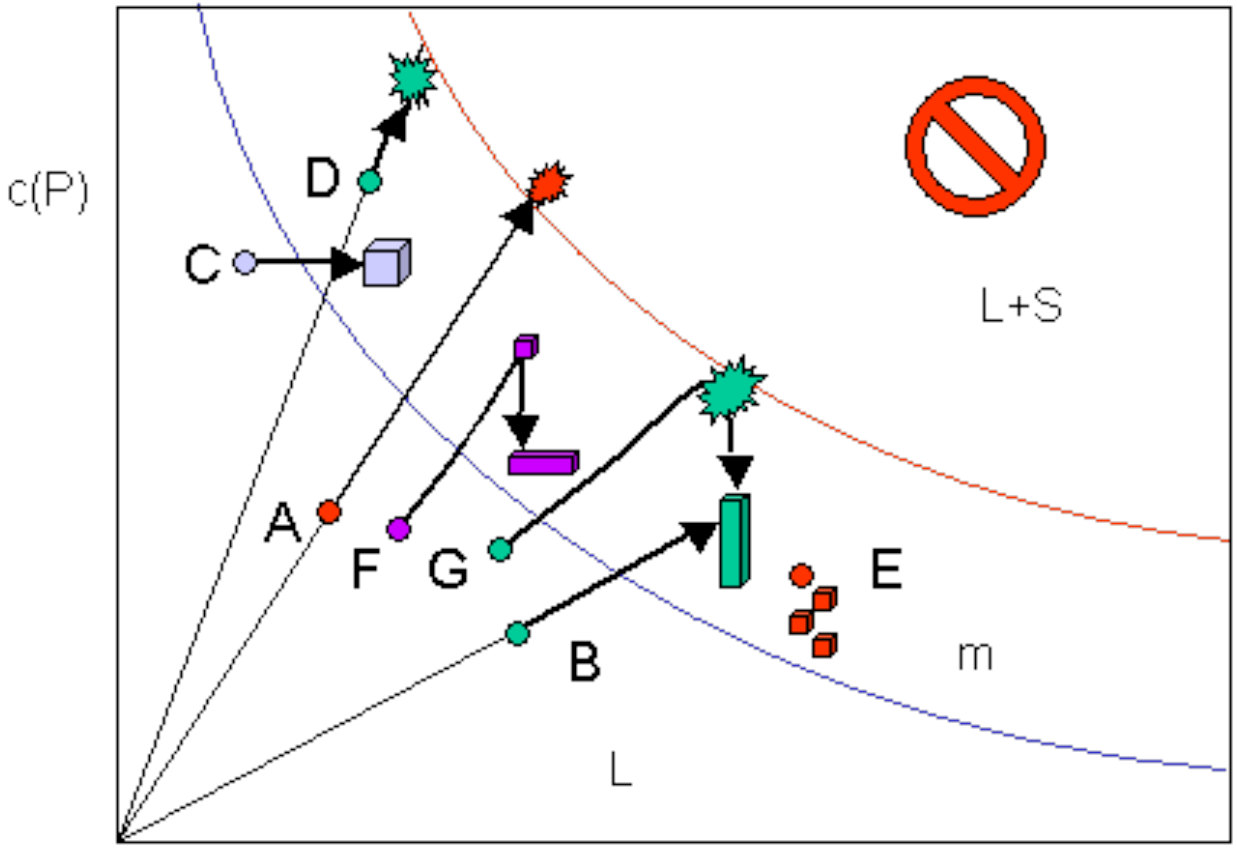
Strategies for crystallisation



Incomplete factorial; random screening (>< random)

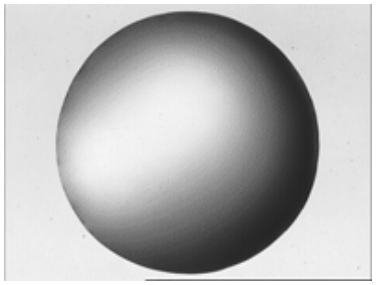
Variation of a single parameter (footprint)

Screening in ‘2 dimensions’; optimisation (grid)

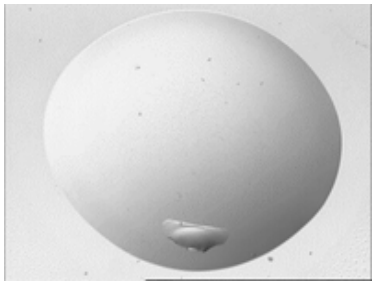


A : kinetics too fast
B : good vapor diffusion condition
C : dialysis
D : open batch (too fast)
E : batch
F : two crystal forms because of variation of concentration
G : decrease in concentration due to precipitation→ crystals

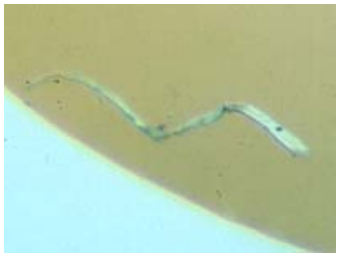
Analysis by observation



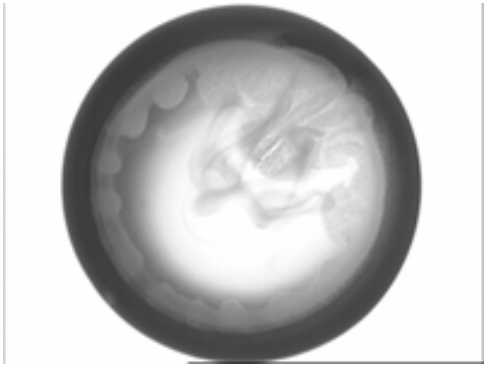
Clear



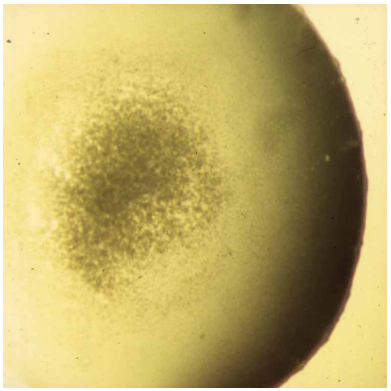
Glas fragment

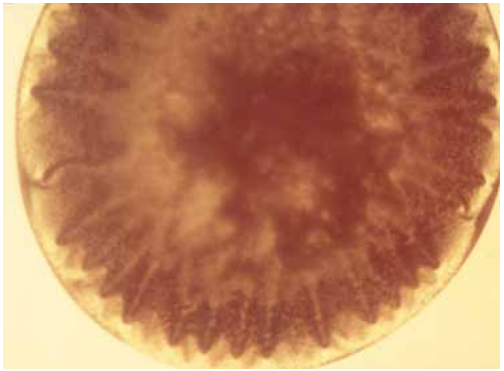


Dust

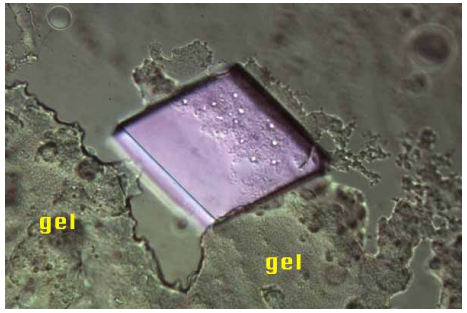


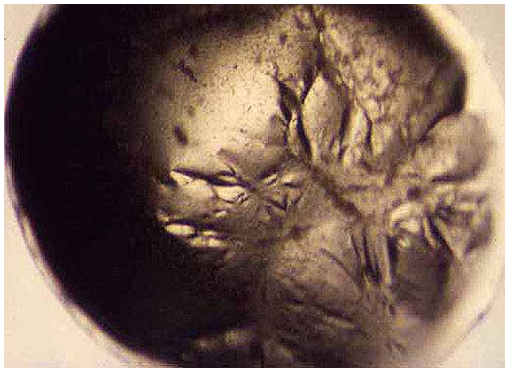
Gel

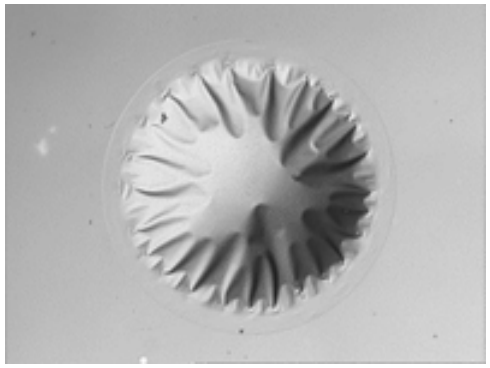




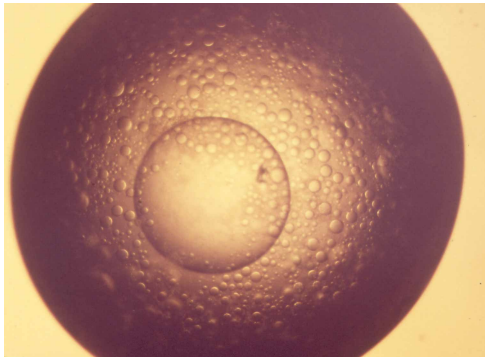
Precipitate



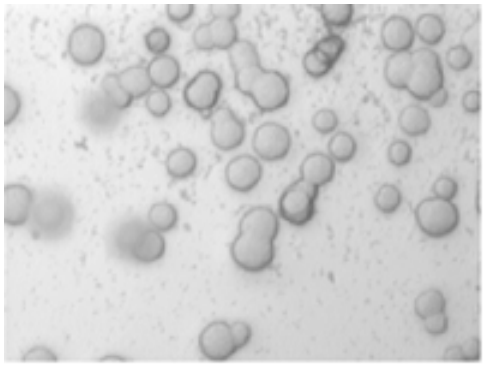
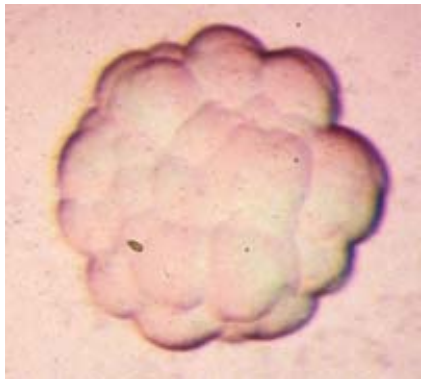
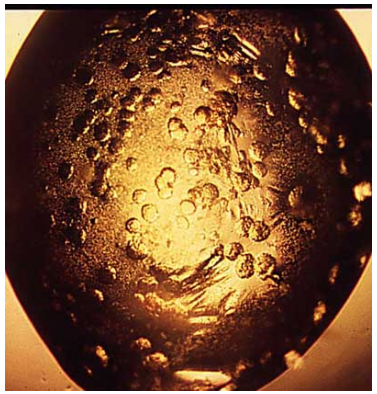




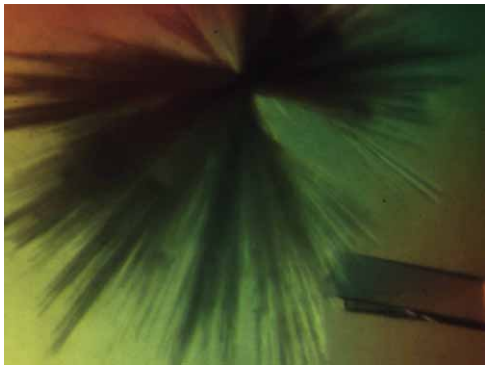
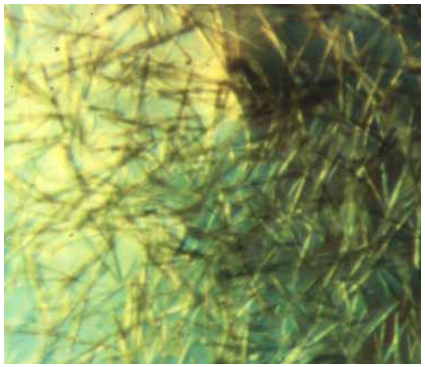
Skin of precipitant (PEG, $(NH_4)_2SO_4$



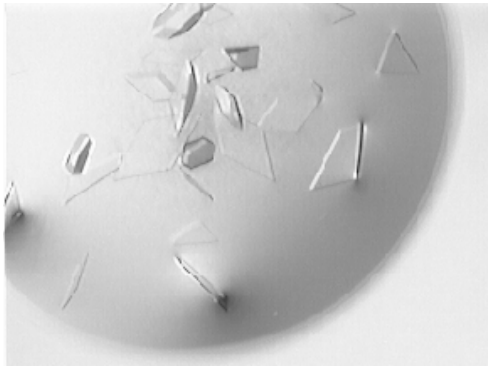
Phase separation

Spherolites

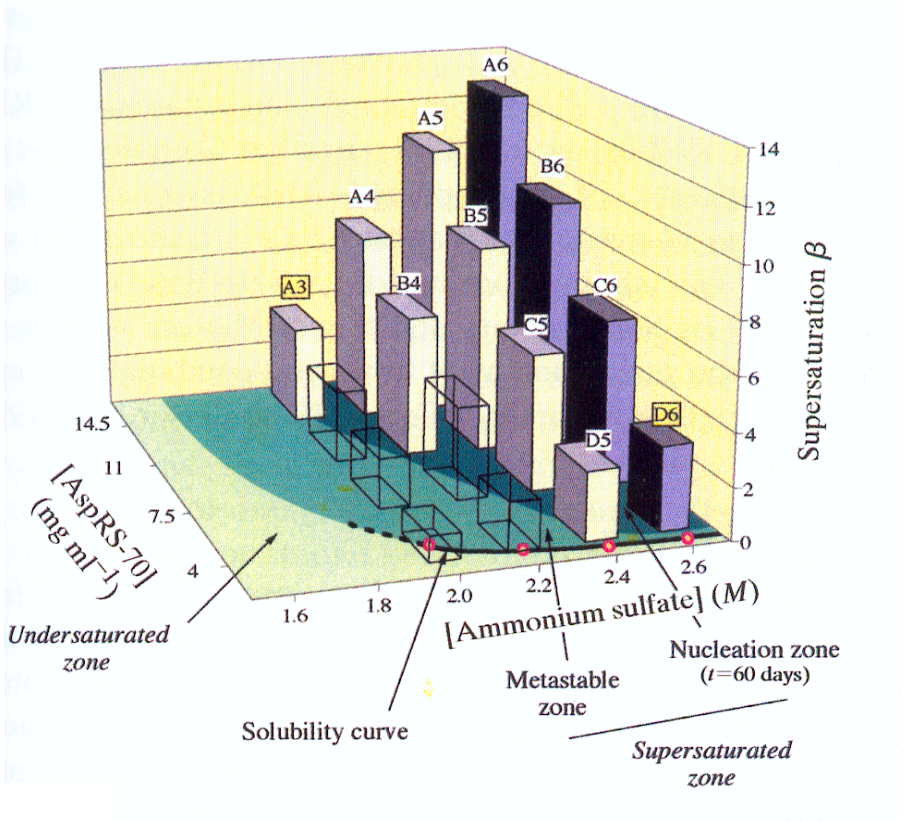



Needles, ‘hair’, urchin

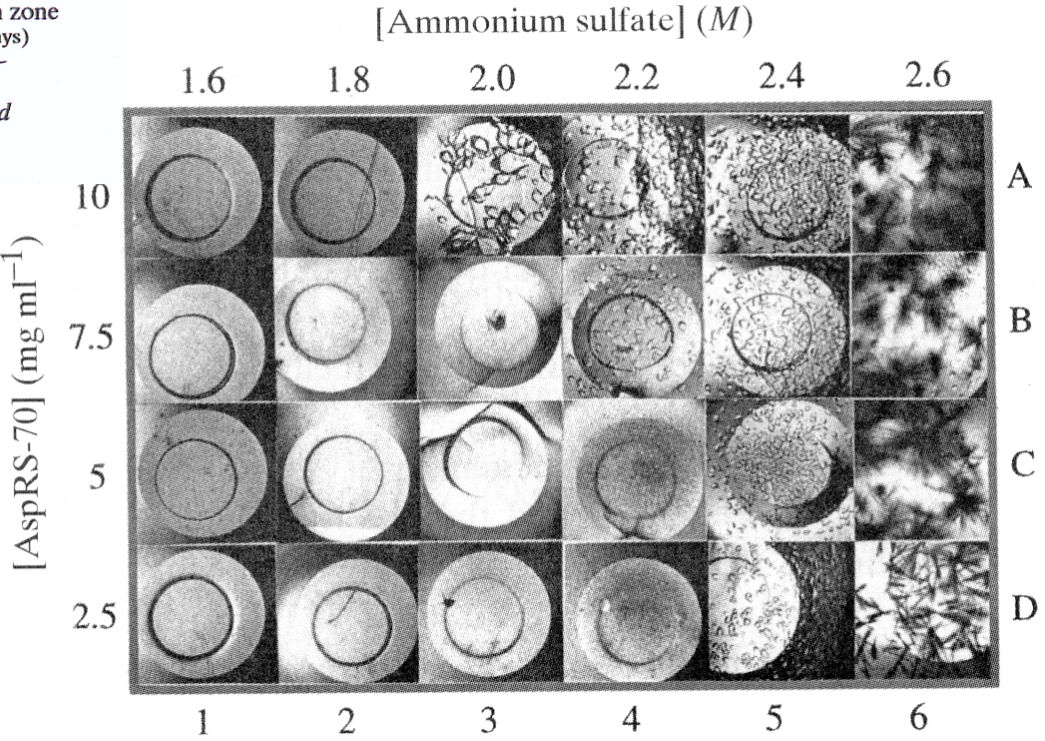




Plates

Optimisation



Variation of pH :
Proteins often crystallize at a pH close to their pI



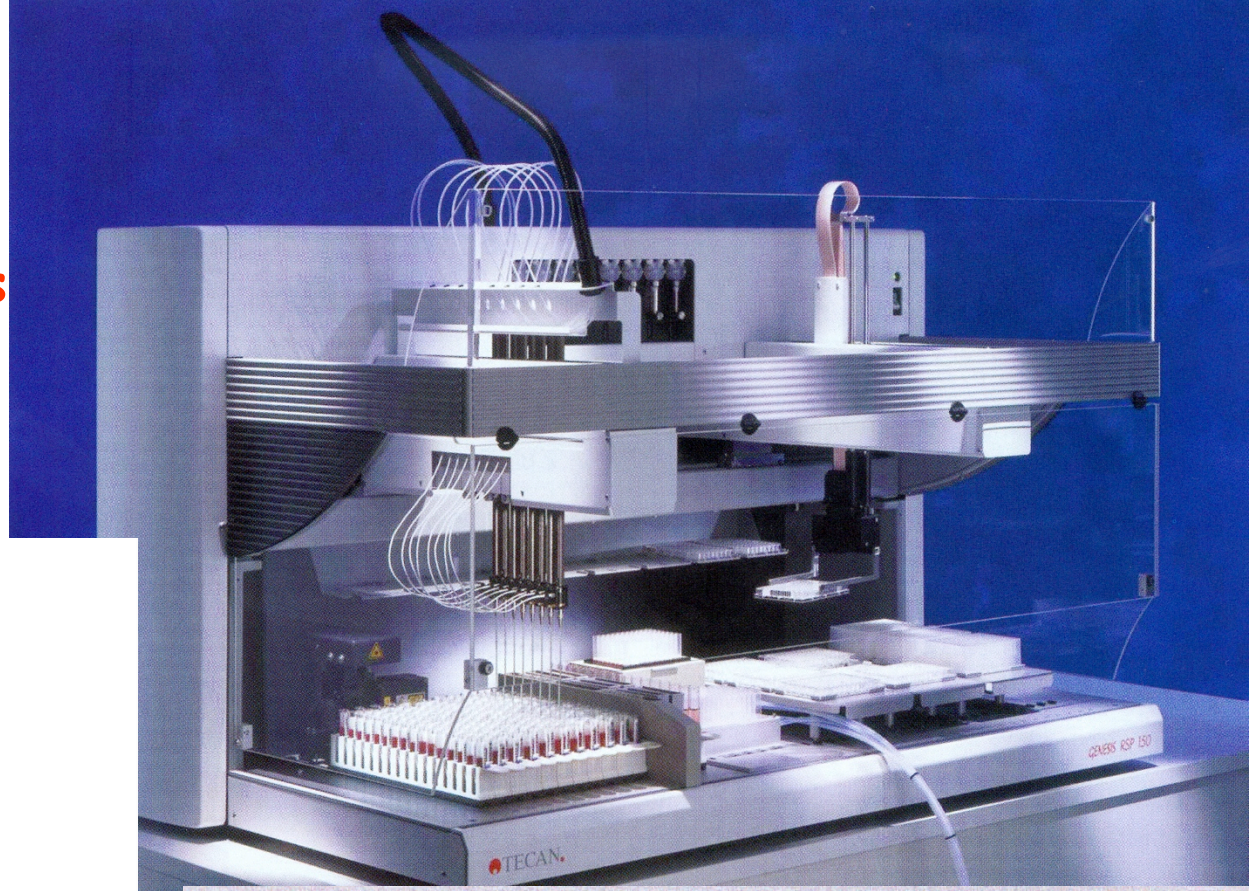
Automatisation

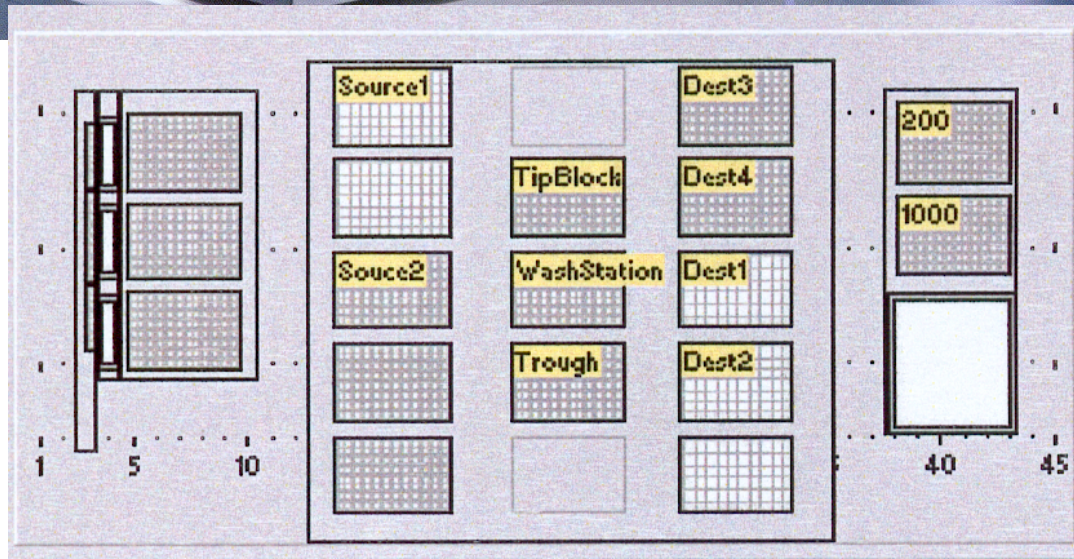
➔

One can perform much
More conditions with less
sample quantitiy

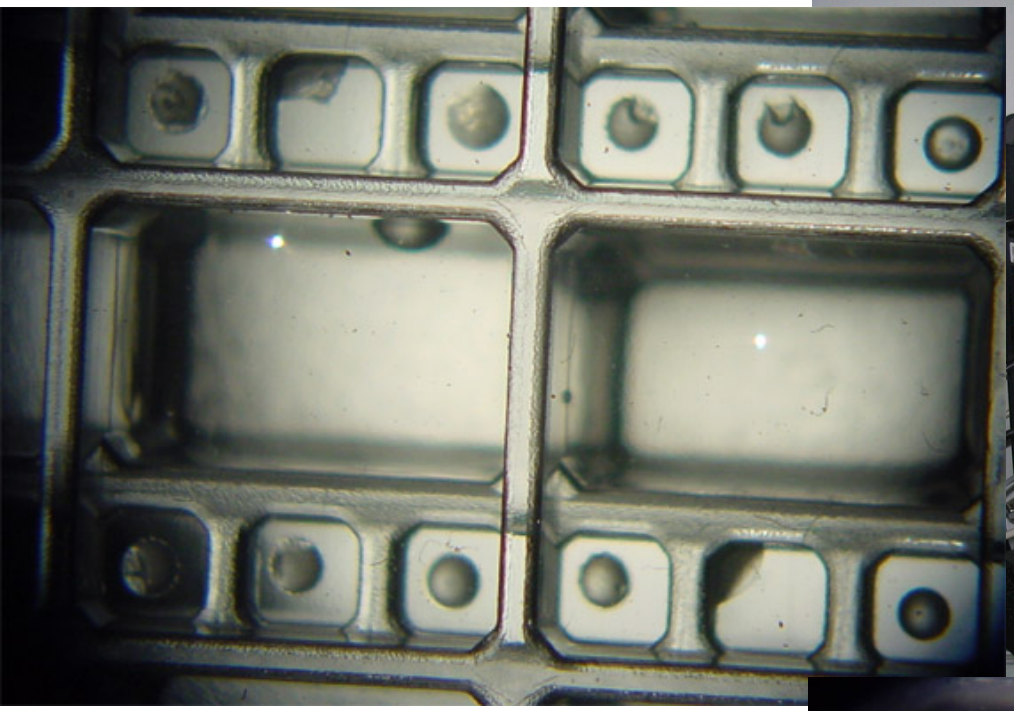
➔

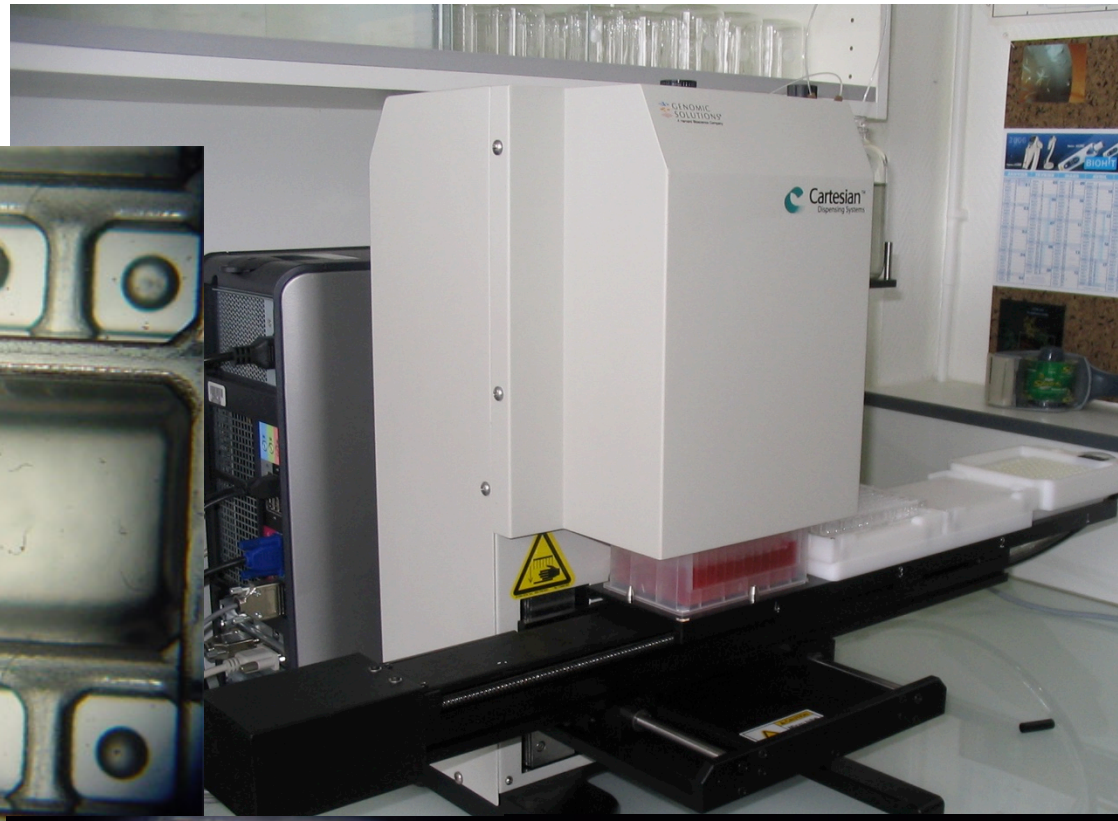
nanodrops





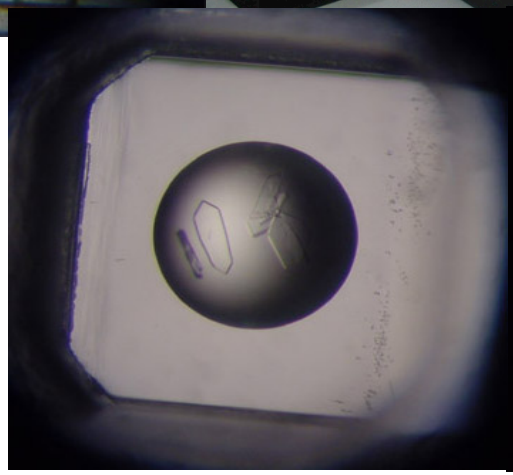
nano-drop dispensing robots

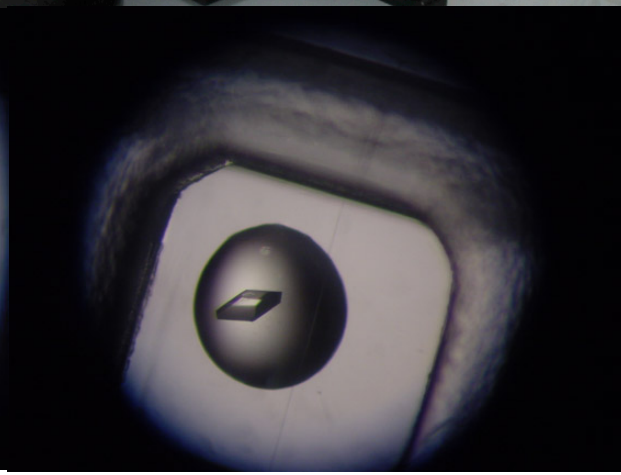




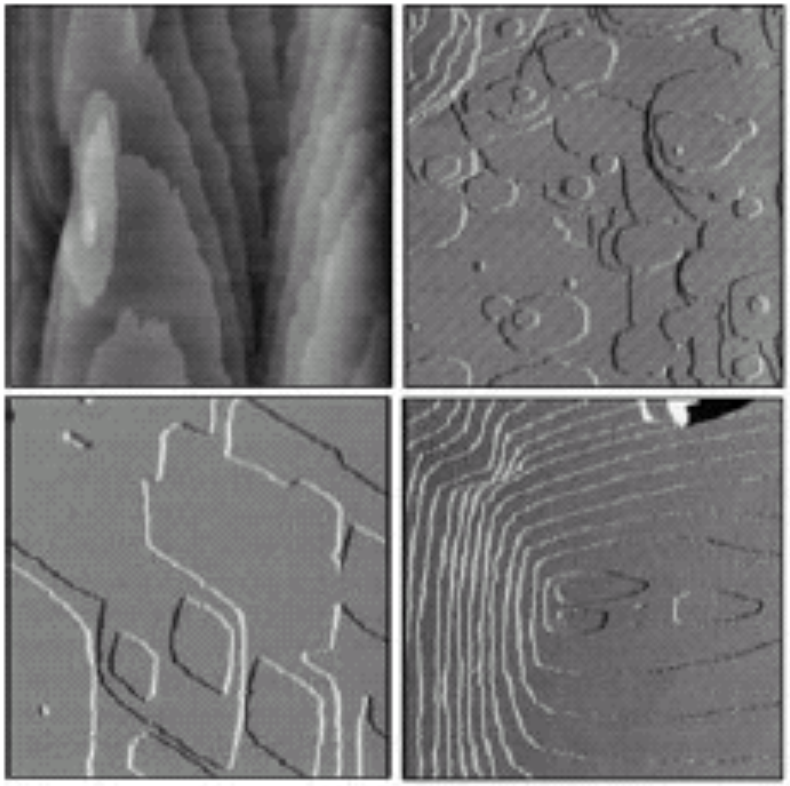
drops of 100-200 nl

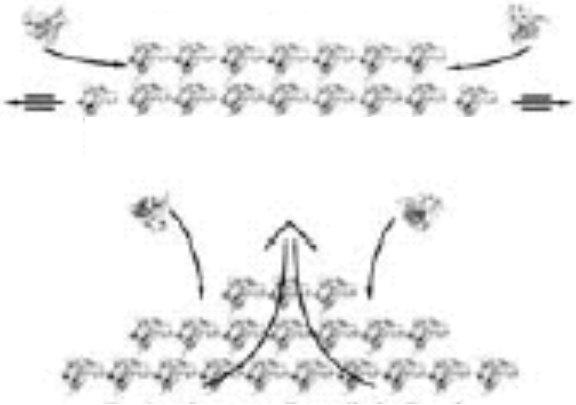
You can test 10 times more conditions with same amount of protein ~ 30 to 50µl

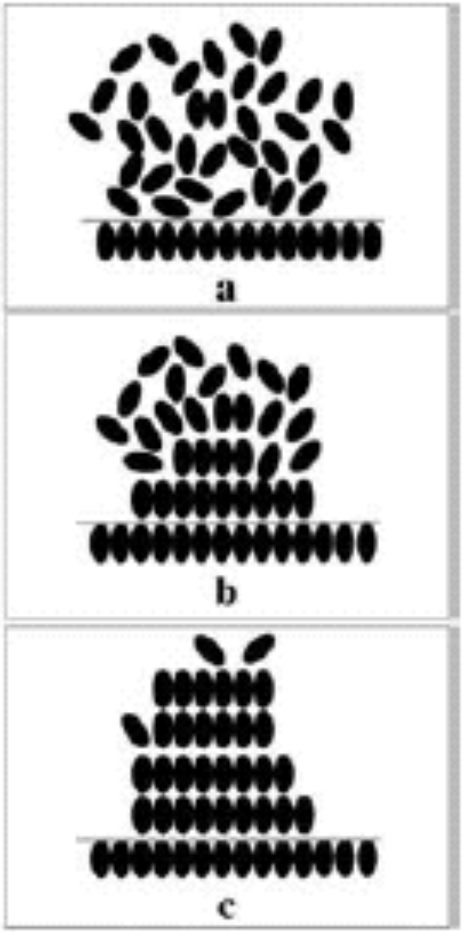




Crystal growth

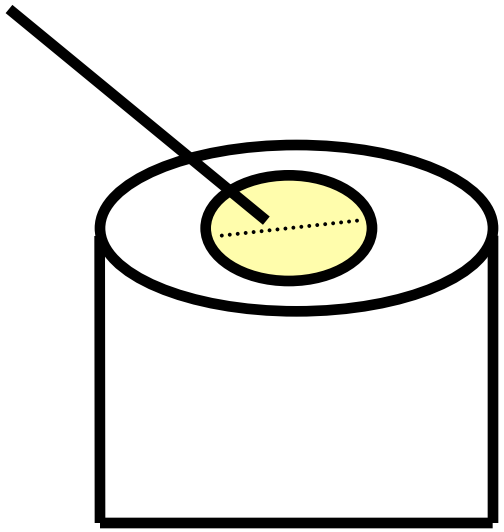






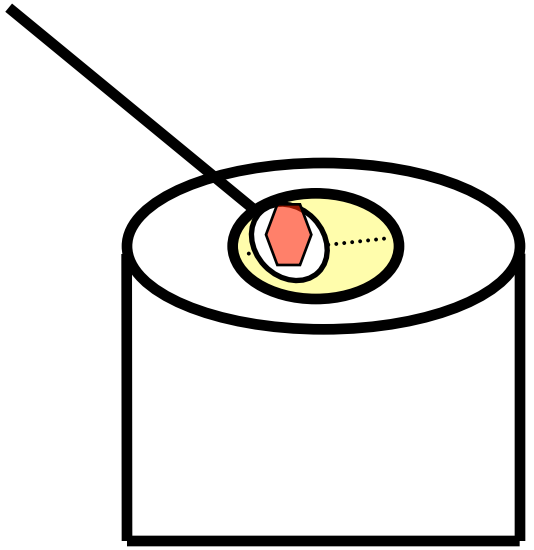
McPherson et al., 2000

Seeding
when nucleation is the limiting step



Micro seeding :

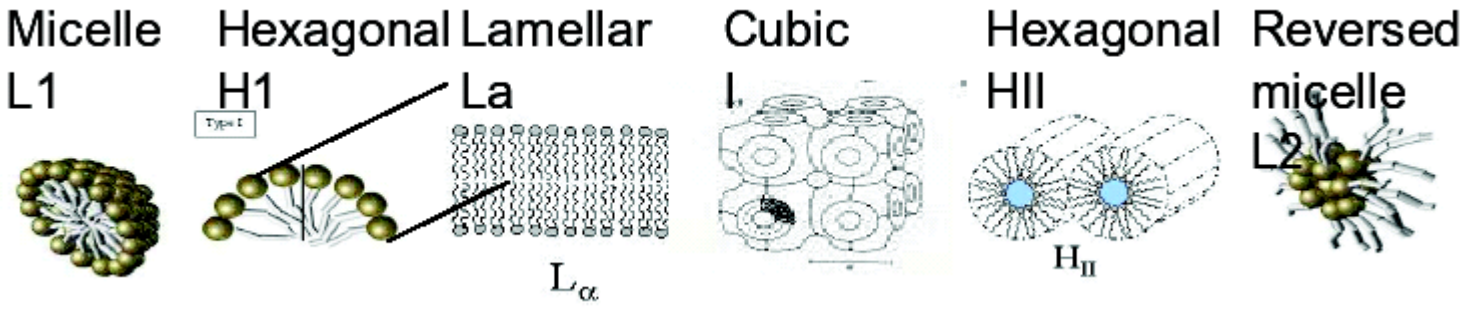
We streak through the drop with a cat/rabbit/horse hair → the locally provoked perturbation will enhance nucleation



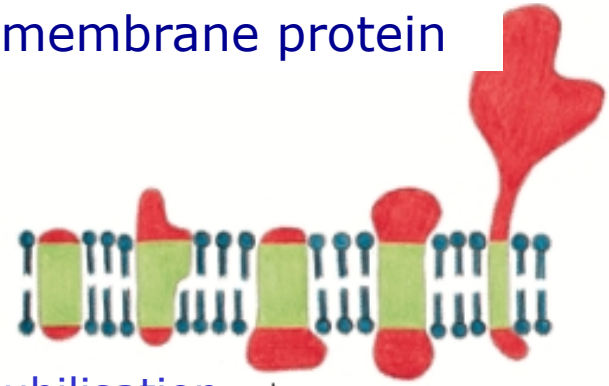
Macro seeding :

A tiny crystal or peice of crystal is transfered into the drop the concentration should by slighly below the spontaneous nucleation limit → the deposited crystal will grow

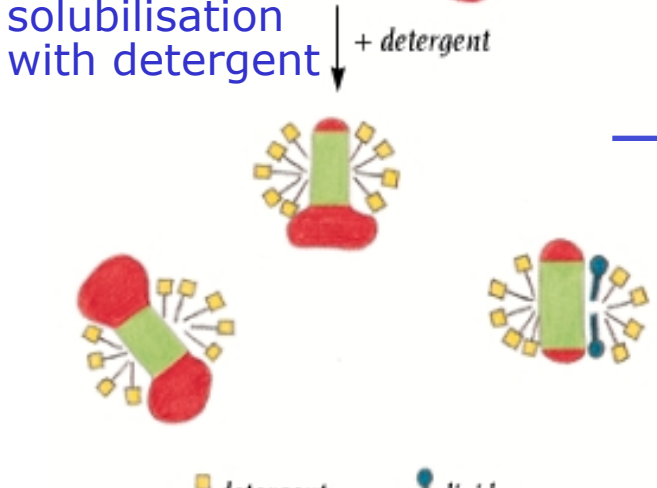
Crystallisation of membrane proteins

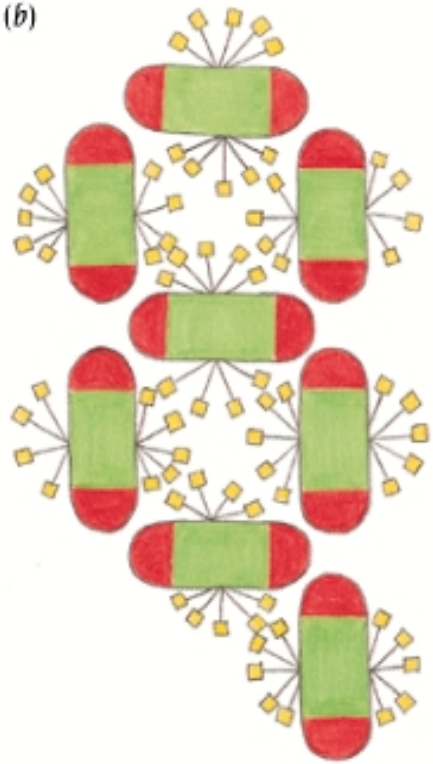


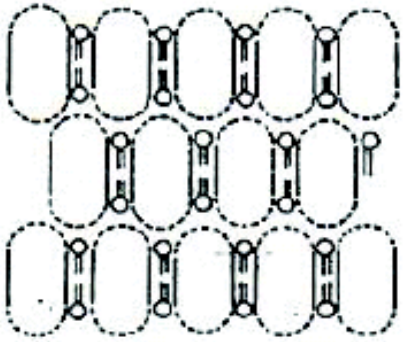
membrane protein



solubilisation with detergent







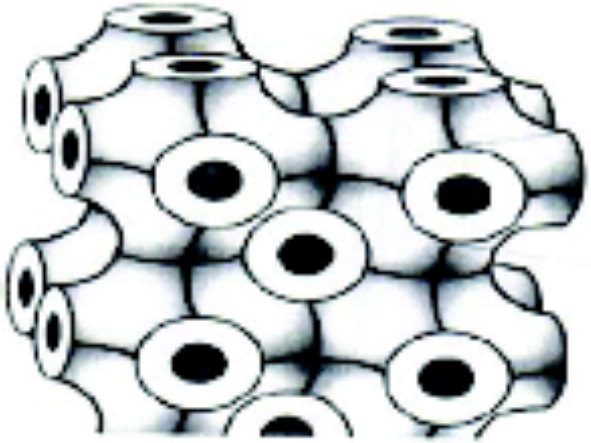
The length of the hydrophobic tail of the detergent must be adapted to the size of the membrane protein molecule

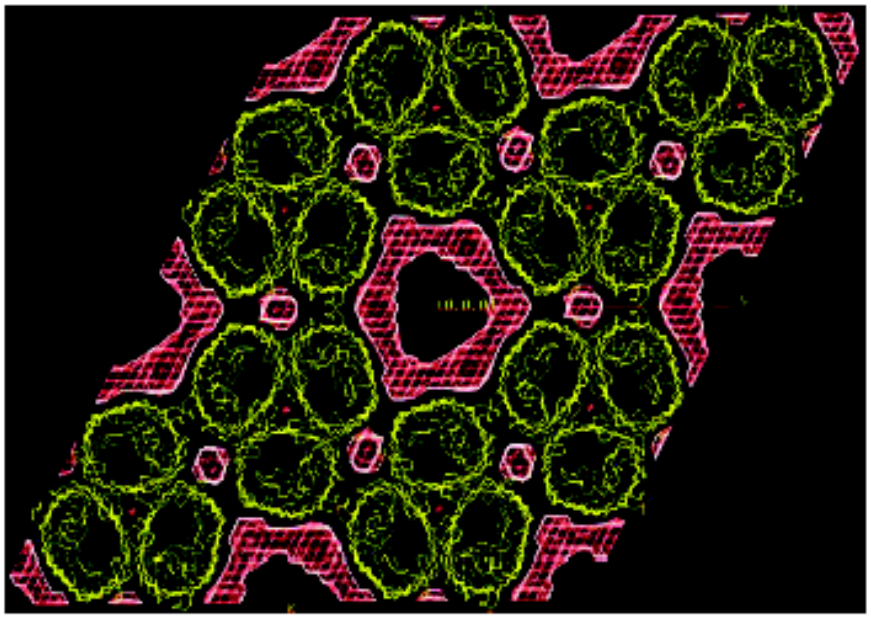
incorporation of detergent/lipid molecules into the protein crystal lattice

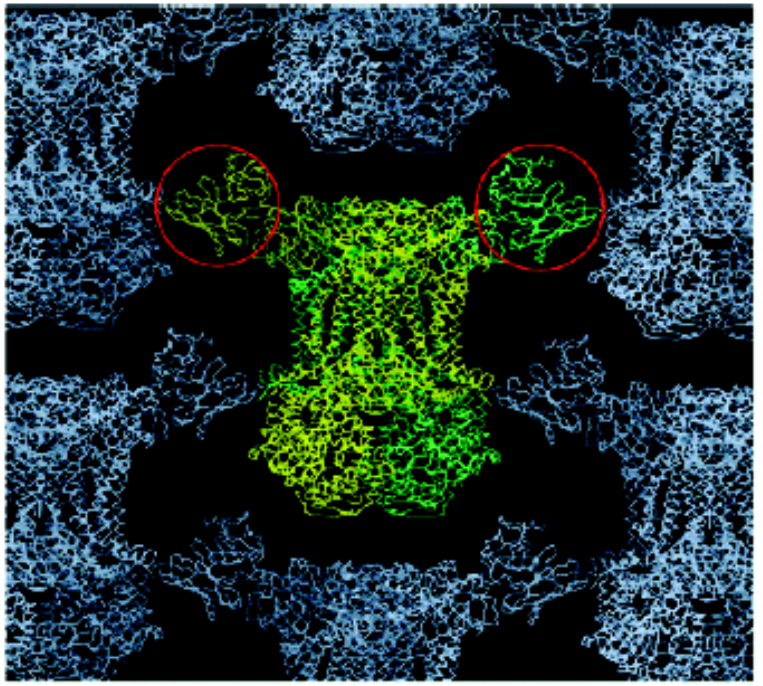
Crystallisation of membrane proteins

lipid cubic phases :

For some specific mixtures lipid/H_2O one can obtain particular phases that form continuous channels of water and lipid that facilitate the crystallisation of membrane proteins



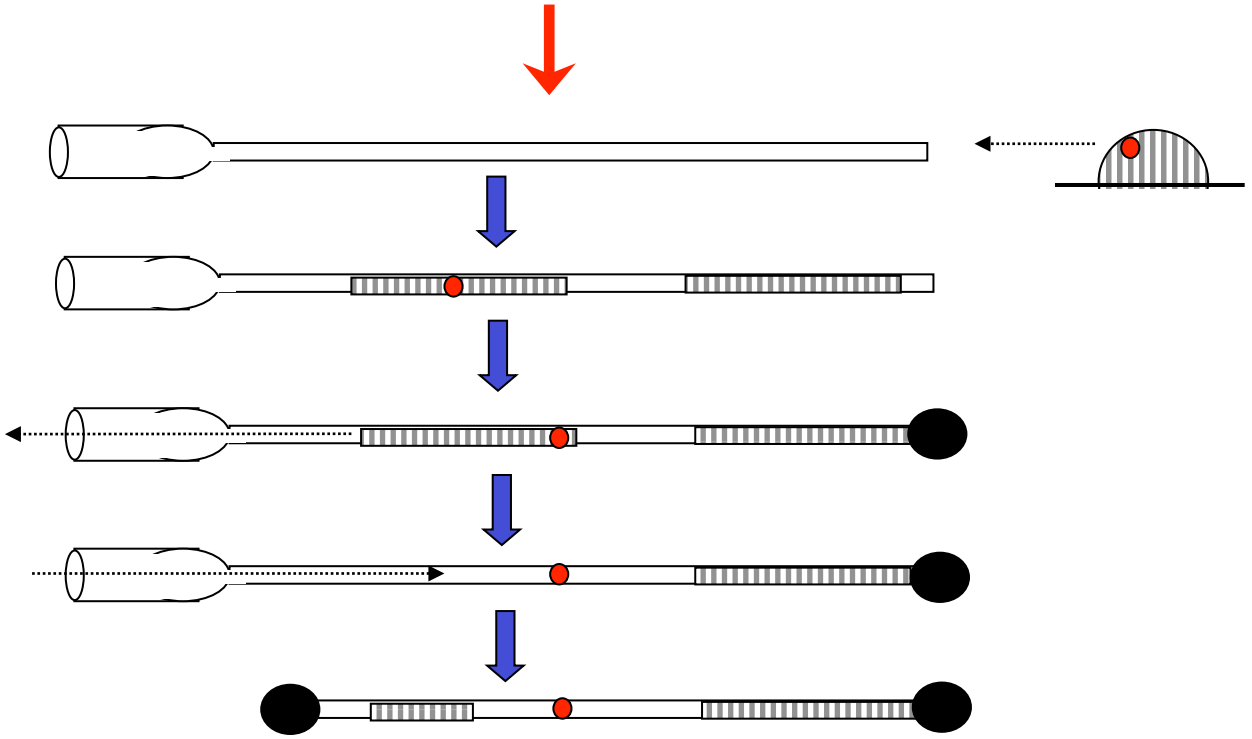




“fishing” crystals (1)

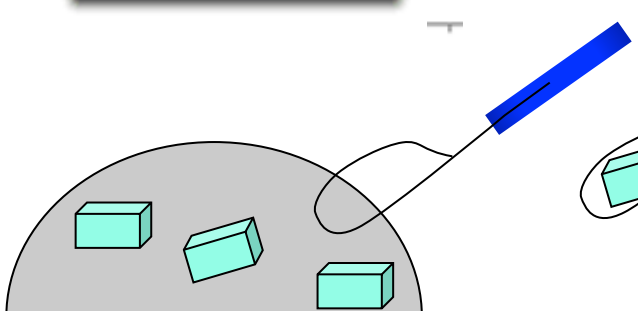
- the crystals are fragil, they have to stay in a hydrated atmosphere.
- Radiation dammage.
- or no solution found for cryoprotection.

Mounting in a capillary

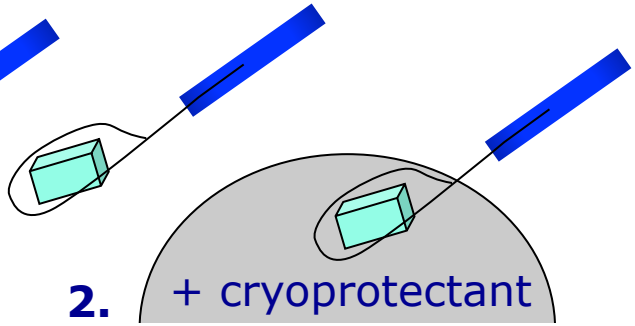


“fishing” crystals (2)

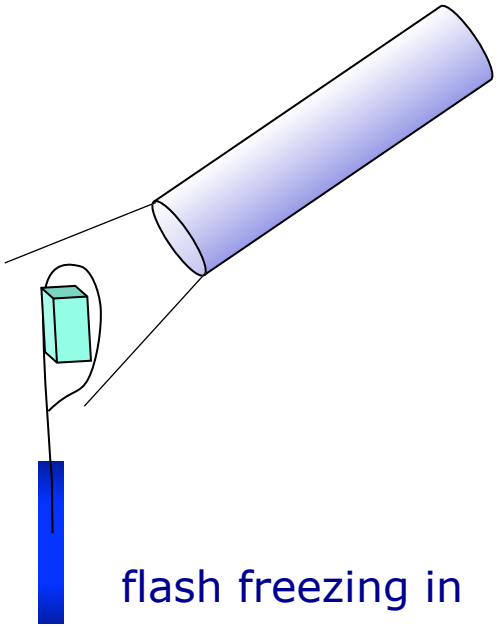
Mounting in a cryoloop



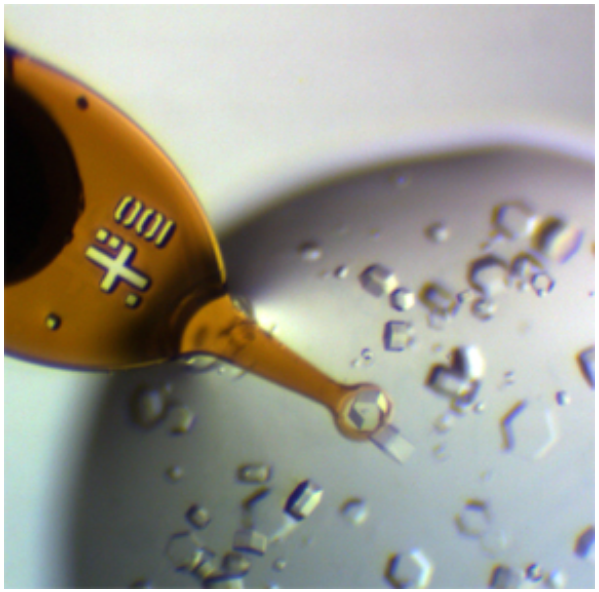
1.



2. + cryoprotectant



3. flash freezing in cryostream



+ glycerol

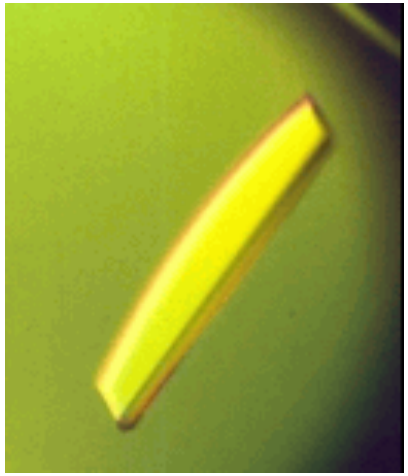
ethylenglycol

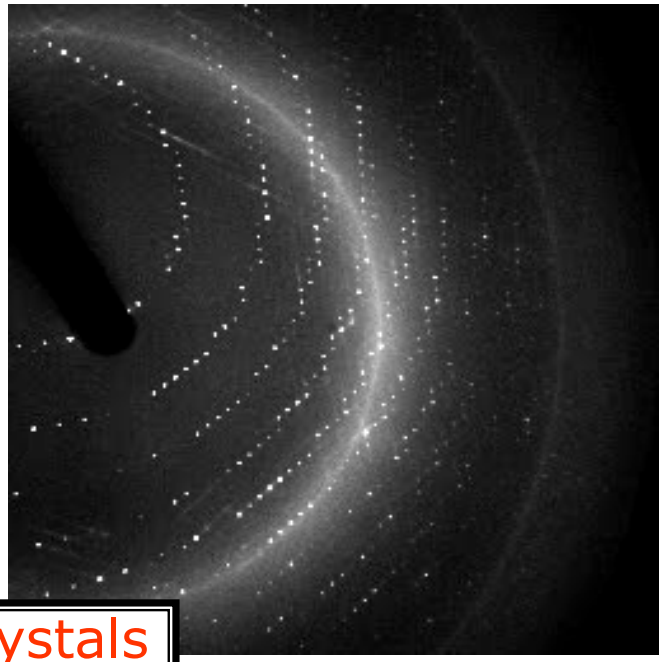
MPD

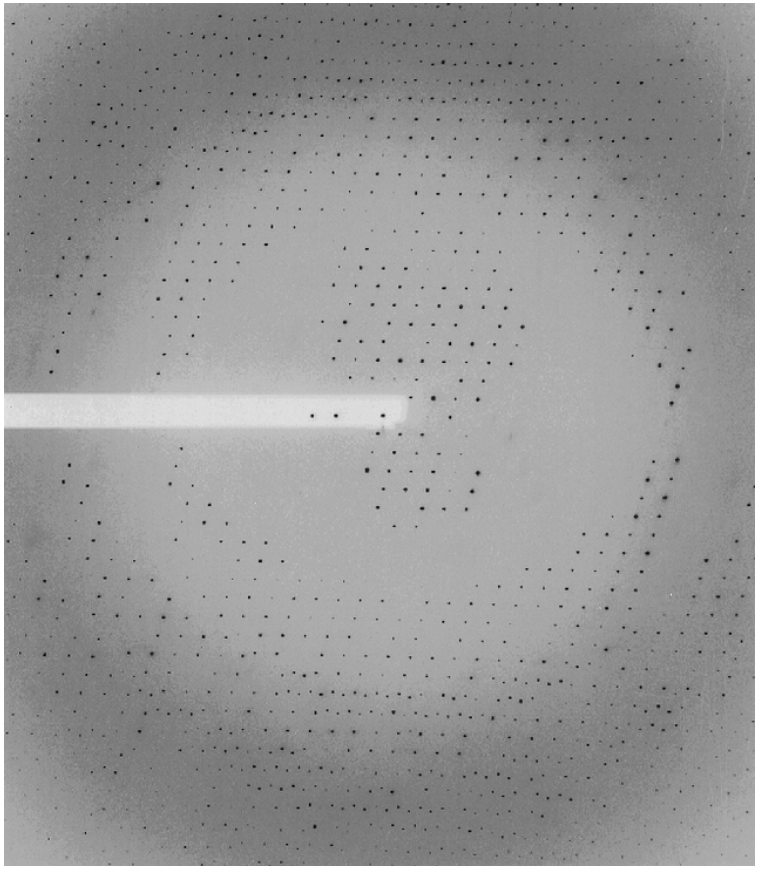
PEG 100-400

....

X-ray Diffraction



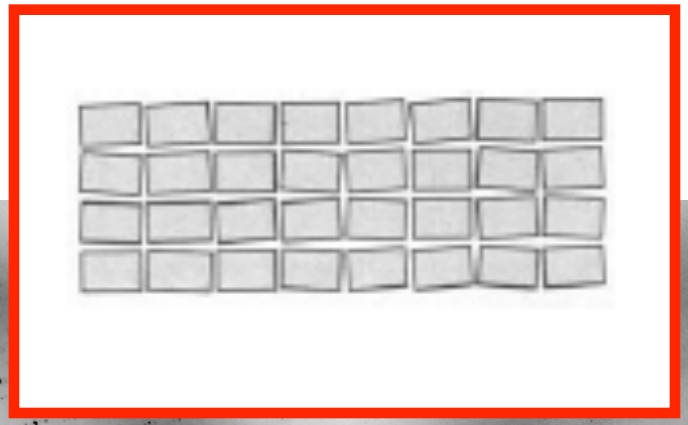


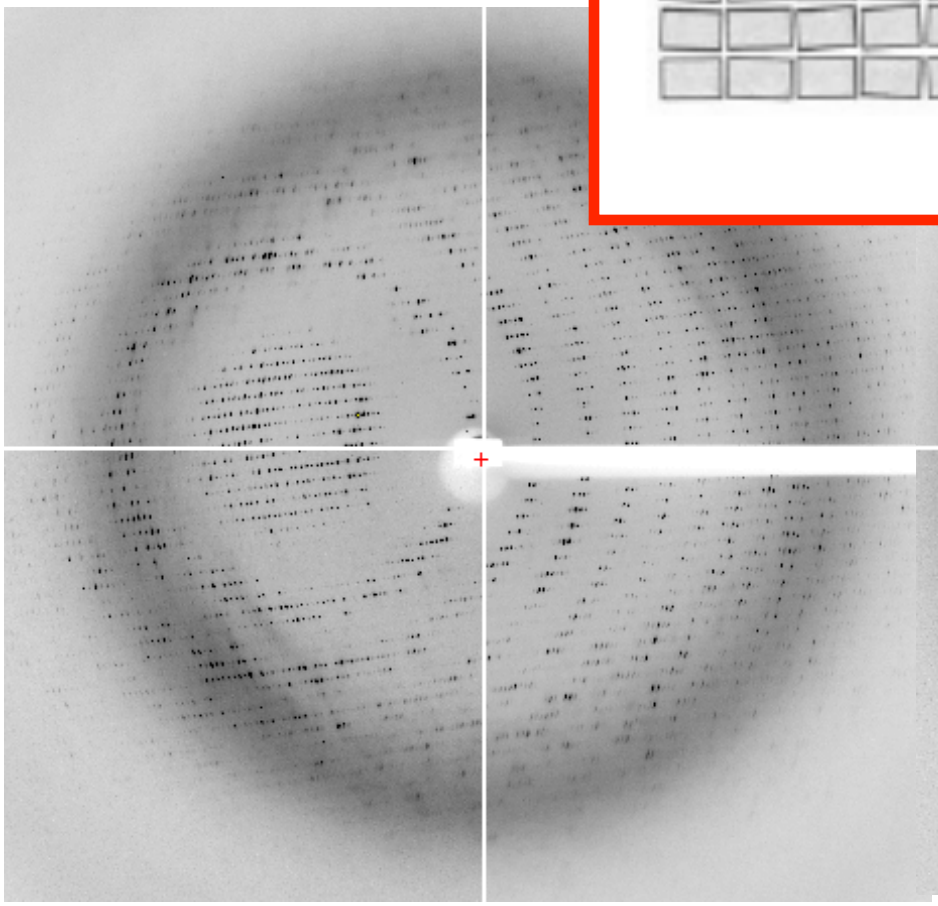


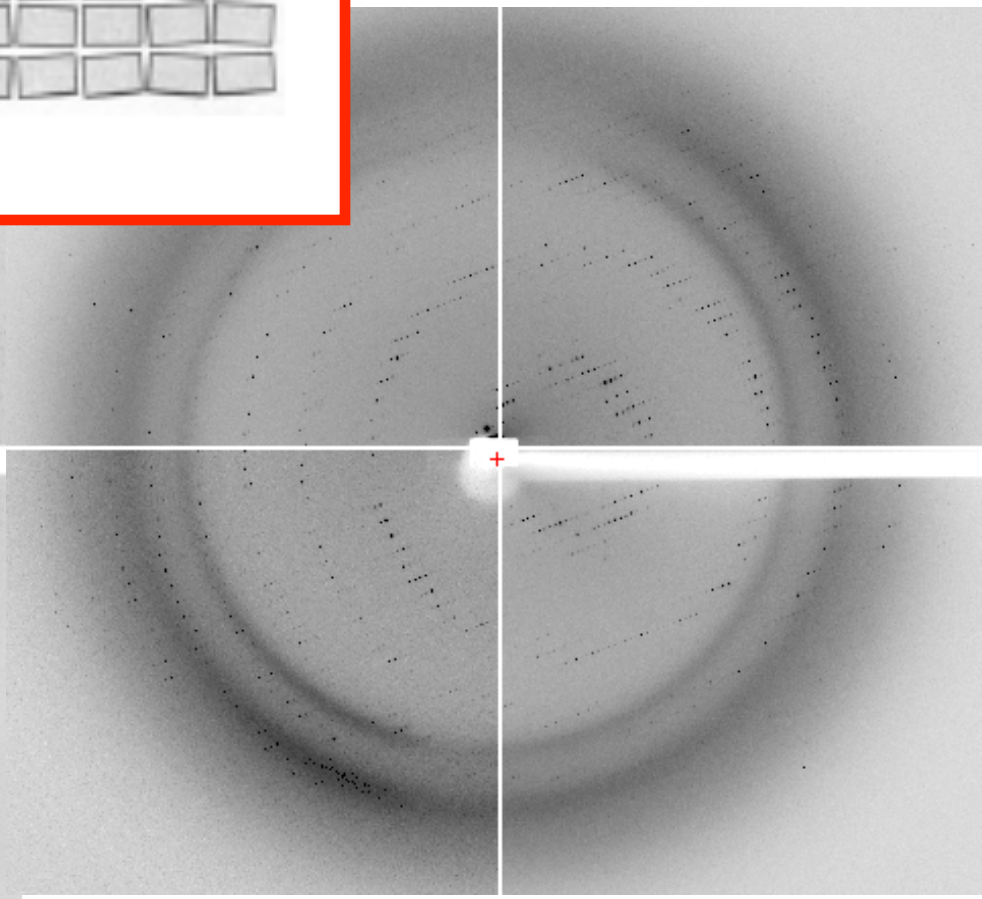
Are the crystals of good quality?

For our needs they're sufficiently 'good' when they diffract X-rays

Mosaicity…

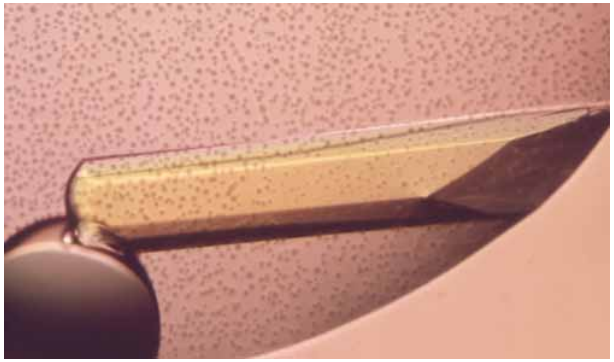


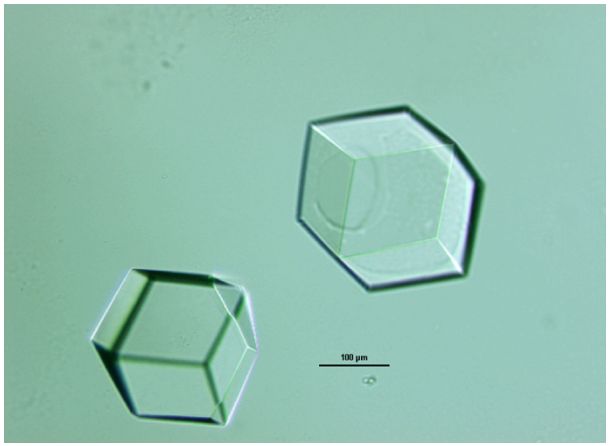


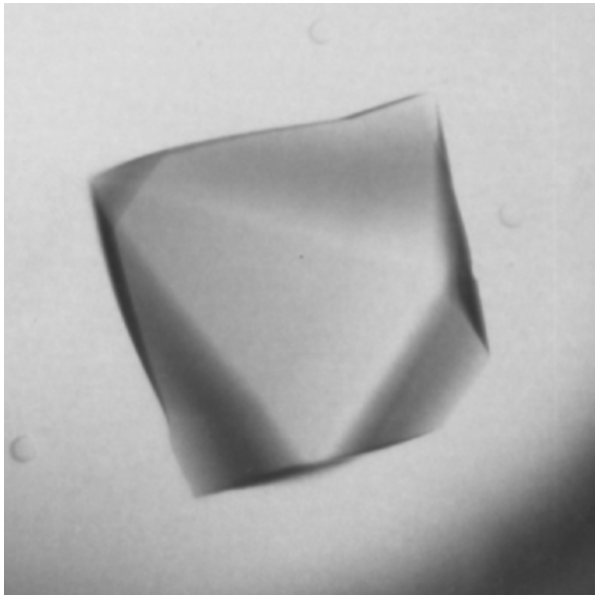


High…

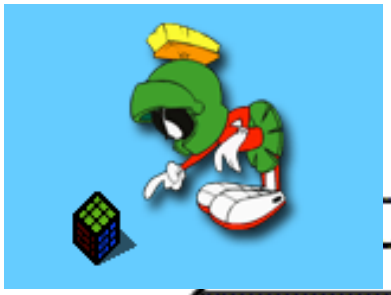
Low…

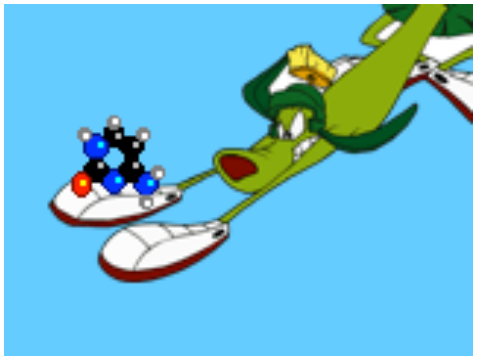



100 µm

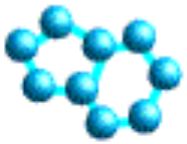


What is a crystal ?

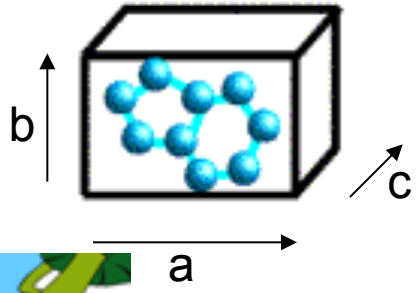




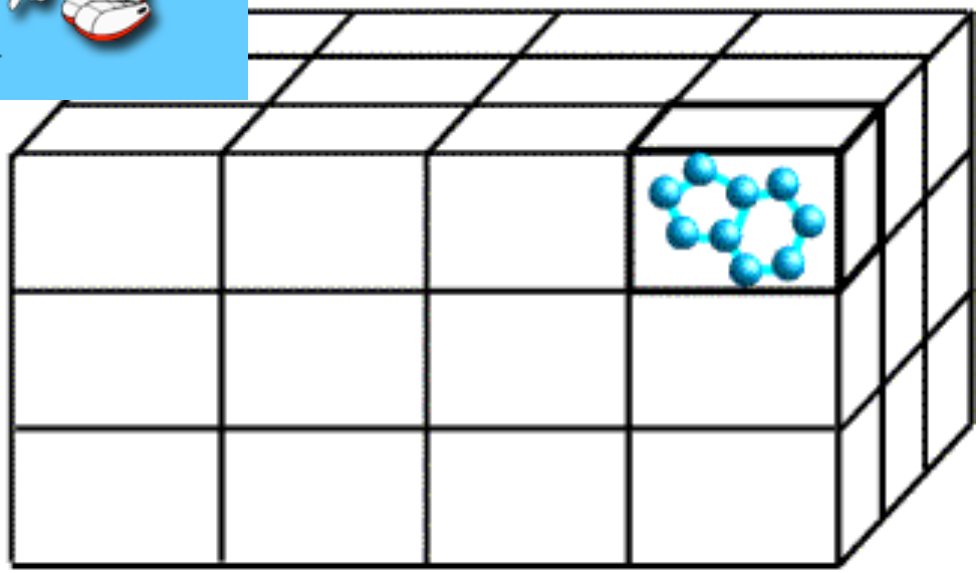
molecule



unit cell

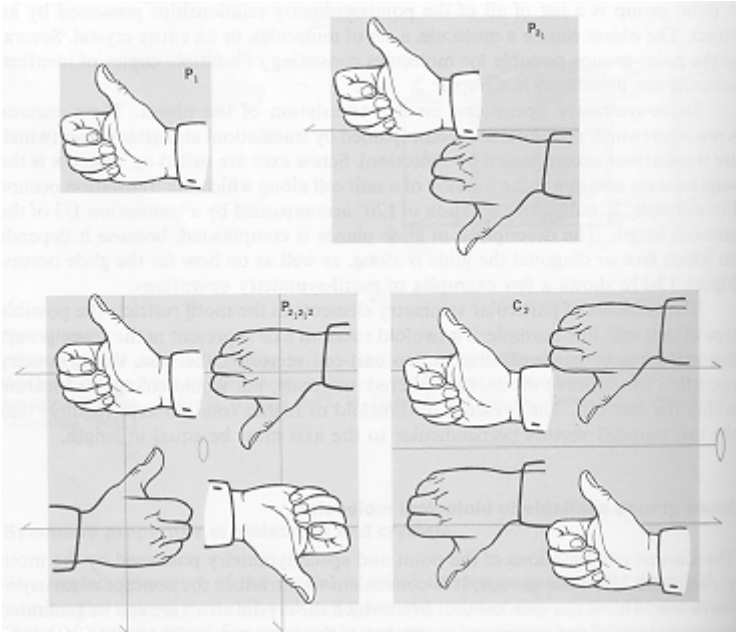


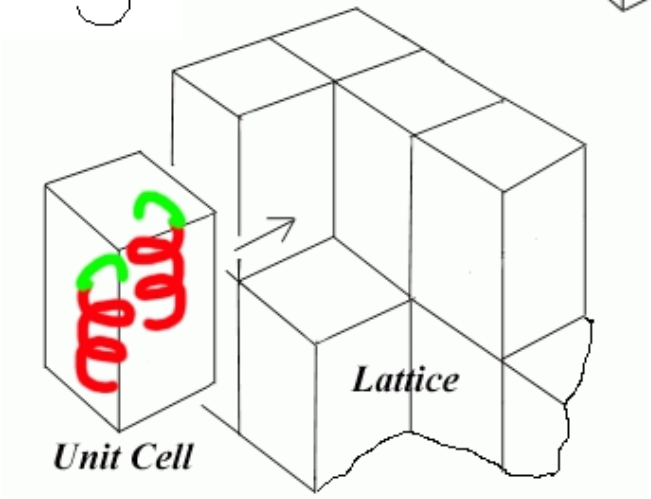
crystal

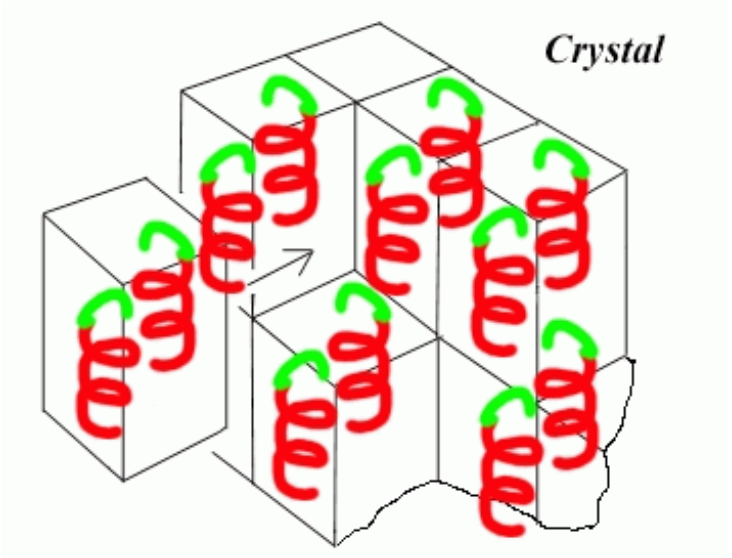


Symmetry

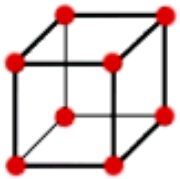
- biological macromolecules are chiral. There is no natural ”mirror” image. Only a limited number of operations are possible in the crystal (65 of 230 space groupes are compatible with biological active molecules)



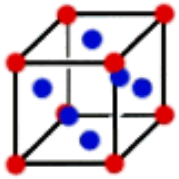




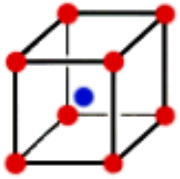
The 14 Bravais Lattices



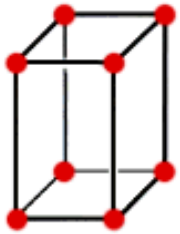
Simple cubic



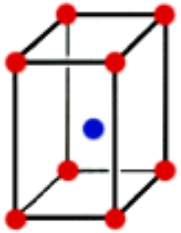
Face-centered cubic



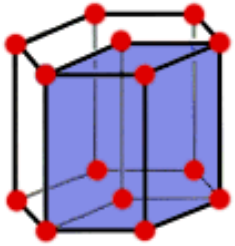
Body-centered cubic



Simple tetragonal



Body-centered tetragonal



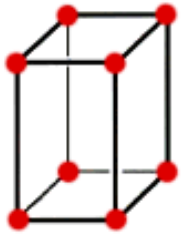
Hexagonal

P = primitive

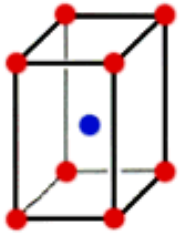
F = Face centered

I = Body centered

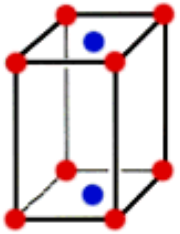
C= C-face centered



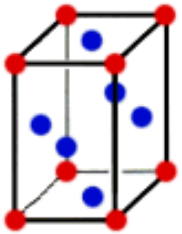
Simple orthorhombic



Body-centered orthorhombic

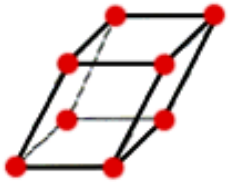


Base-centered orthorhombic

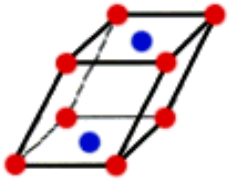


Face-centered orthorhombic

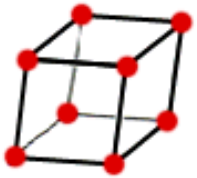
Rhombohedral



Simple Monoclinic

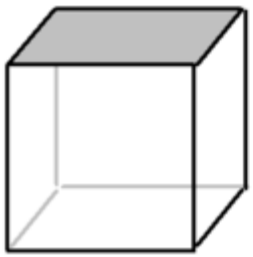


Base-centered monoclinic

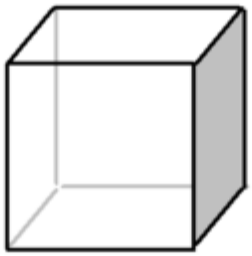


Triclinic

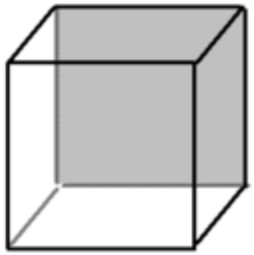
Lattice planes and Miller indices



(001)

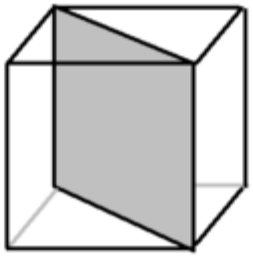


(100)

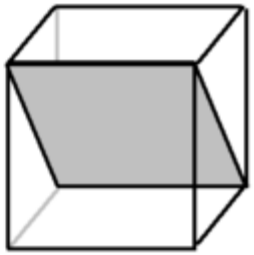


(010)

(101)

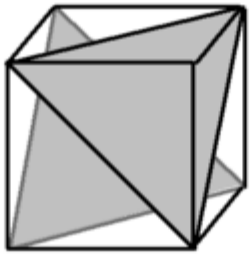


(110)

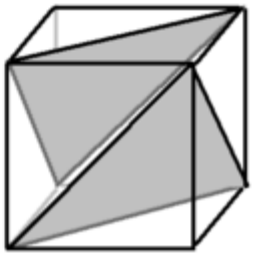


(011)

(111)



$(1\bar{1}1)$



$(\bar{1}11)$