Outline lecture 1.

- Main steps in protein crystallography.
- Protein preparation.
- Protein crystals symmetry.
- Protein crystallyzation: methods and phase diagrams.
- Screening and optimization.
- Protein-ligand crystallization.
- Crystal mounting

STRUCTURAL BIOLOGY

Study of the 3-D structure of biological macromolecules (proteins, nucleic acids) giving foundamental informations on the structure-function relationship

Methods:

- X-ray crystallography

- NMR

Advantages of NMR	Disadvantages of NMR	Advantages of X-ray	Disadvantages of X-ray
1. several types of information from lots of types of experiments	1. we have lots of atoms and a lot of extracted data from a system.	1. We can examine also by this way the solvent effect since from different solvents the same protein may crystallize into different crystalloid form.	1. the crystal structure is necessary only that proteins which can be crystallized are examinable
2. they obtain angles, distances, coupling constants, chemical shifts, rate constants etc. These are really molecular parameters which could be examined more with computers and molecular modeling procedures.	2. This is good for the more accurate determination of the structure, but not for the availability of higher molecular masses	2. So we are able to force the protein to an other form of crystallization by the change of its solvent.	2. we cannot examine solutions and the behavior of the molecules in solution
3. if we have enough strength of the magnetic field (the resolution is the function of that) than we can handle all of the atoms "personally"	3. the resolving power of NMR is less than some other type of experiments (e.g.: X-ray crystallography) since the information got from the same material is much more complex	3.we could get the whole 3D structure by the systematic analysis of a good crystallized material	3. This happens when we try to examine powders, gases
4. With a suitable computer apparatus we can calculate the whole 3D structure	4. the highest molecular mass which was examined successfully is just a 64kDa protein- complex		4. studying of motions are not available
5. There are lots of possibilities to collect different data- sets from different types of experiments for the ability to resolve the uncertanities of one type of measurements	5. there are lots of cases when from a given data-set - a given type of experiment - we could predict two or more possible conformations, too		5. we can get only one parameter-set so we are able to observe only one conformation
6. the motion of the segments (domains) can be examined	6. unfortunately we are just able to determine the degree of probability of being of the protein segment in the given conformation		6.there is no possibility to examine small parts in the molecule
7. this method is capable to lead us for the observation of the chemical kinetics	7. The cost of the experimental implementation is increasing with the higher strength and the complexity of the determination		7. There is no chance for direct determination of secondary structures and especially domain movements (big disadvantage against the NMR)
8.(activation-)thermodinamic (and certainly kinetic) data could be determined from a well-prepared (dynamic-)NMR experiment			8. the hydrogen in the molecules are not examinable since it has only one electron
9. we can investigate the influence of the dielectric constant, the polarity and any other properties of the solvent or some added material			

Sequence = specific folding

A Sequence of Bases in DNA...

Triplets of bases read from one strand

Is Translated to a Sequence of Amino Acids in a Protein... Which Folds Spontaneously to a Precise Three-Dimensional Structure





Main steps in protein crystallography.

- 1. Protein purification
- 2. Protein crystallization
- 3. Testing crystals
- 4. Data collection
- 5. Data processing
- 6. Structure solution



- Obtaining single crystals that diffract to high resolution remains the primary bottleneck of protein crystallography!!!





Colony of cells each containing copies of the same recombinant plasmid

- <u>Preparing protein for crystal trails usually</u> <u>requires a large amount of highly purified</u> protein.

- Using recombinant DNA technique, it is now possible to instruct a variety of cells and organisms to make a large amount of almost any protein chosen by investigator.

- Not only can specific proteins be expressed in large quantities also recombinanat proteins can be modified in ways that make the task of the crystallographer simpler and in some cases, dramatically improve the quality of the resulting crystals.

Expression systems. 1.E. coli 2.Yeast 3.Insect Cell 4.Animal Cell



The goal of a protein separation is to obtain a protein in a pure, active form with minimum steps and within shortest time. Affinity choromatography can allow efficient purification of fusion protein or proteins with well defined ligand-binding domains.



Histidine-tagged proteins





USEFUL TECHNIQUES FOR CRYSTALLOGRAPHERS

The most important thing you will ever put in your crystallization trials is your protein. Some basic things to check before you start are:

•Is it pure?

SDS-PAGE with Silver staining

•Is it homogenous?

Absence of post-translation modifications with mass spectroscopy

•Is it folded?

Circular dichroism

•Is it fresh?

•Is it monodispersed?

Gel filtration and light scattering

•Does your protein need to be kept reduced?

Gel filtration in presence and absence of reducing agent

•Does your protein need the addition of something (eg salt) to stay in solution?

•Is your protein stable at room temperature?

•Does your protein break down rapidly?

•Has anything similar been crystallized before?

Check the PDB and look in the header records for crystallization details

If you are getting nowhere with your protein, try crystallizing something else.

• *Ligand-protein complex.* Does your protein bind a ligand? If it does, it can help because the binding of a ligand

- is likely to order the region of the protein that binds ligand
- may bring two subdomains together and reduce flexibility
 - will change the surface properties of the protein
 - may cause a conformational change in your protein
- *Different constructs*. Heterogeneous tertiary/quaternary structure can hinder crystallization
- Does your protein break down to a stable proteolytic fragment (either "spontaneously" or with the help of an added protease)?
- \bullet Does the homology of your protein to others in the same family drop off at the N- and C-termini?
- Does your protein have a domain structure?
 - Check the <u>Pfam</u> database
 - Check the <u>ProDom</u> database
- •Does your protein have low complexity regions?
- *Different species.* Sometimes, a point mutation is all that is required to stop/start a protein crystallizing. Working on a different species is the easiest way to get a collection of point mutations that do not affect function.
- *Deglycosylation*. If your protein is glycosylated the floppy and heterogeneous carbohydrates may be interfering with the crystallisation. Try <u>enzymatic deglycoylation</u>.

Crystallizability is inversely proportional to biological interest. *Murphy's crystallization law*

2. Protein crystallization

The solubility of proteins can be represented in **phase diagrams**.

The phase diagram plots the solubility curve of a protein.

- •the horizonal axis shows the parameter being varied (usually precipitant concentration)
- •the vertical axis shows the protein concentration.



•Saturation occurs when the rate of loss and gain of both the solid and solution phases of the protein are equal, and the system is in equilibrium.

•*Salting-out* is seen on the right hand side of the diagram where there is a reduction in protein solubility as the concentration of salt increases

•*Salting-in* is seen on the left hand side of the diagram where there is an increase in protein solubility as the concentration of salt increases.



In the crystal lattice, all the unit cells have the sameshape, size and content.

In three dimensions, the unit cell is the smallest portion of volume of the lattice which, translated parallel to itself, reconstructs the entire crystal.



Now one could go ahead and replace the lattice points by more complex objects (called basis), e.g. a group of atoms, a molecule. This generates a structure that is referred to as a crystal:



A crystal is defined as a lattice with a basis added to each lattice site. Usually the basis consists of an atom, a group of atoms or a molecule.

The asymmetric unit and the unit cell



Unit cell:

characterized by edge length and angles can contain several molecules is repeated multiple times along the axis of the crystal

Assymetric Unitminimal part of the unit cell which is related to other parts by defined symmetry operations





The asymmetric unit

The asymmetric unit (shaded duck) is the smallest unit of structure that can generate the whole crystal after application of the crystal symmetry.



Protein crystals symmetry.





The unit cell is outlined by three cell axes, **a**, **b**, and **c** along the x, y and z directions respectively. The use of bold font indicates that the cell axes are vectors, with length and direction. The unit cell is often described by the length of the axes and the angles between them: $a = |\mathbf{a}|, b = |\mathbf{b}|, c = |\mathbf{c}|$

 α = angle between **b** and **c**; β = angle between **a** and **c**; γ = angle between **a** and **b**

Symmetry operators in crystals



A **point group** is a group of geometric symmetries (isometries) leaving a point fixed (origin).

Asymmetric unit: the minimum portion of the unit cell that can generate all the unit cell by applying the simmetry operations.

The **space groups** derive from point groups with the addition of translation operations.

A space group contains all the simmetry operations of the atoms in a crystal.

Not all rotational symmetries are allowed in a crystal

The only rotational symmetries possible in a crystal lattice are 2, 3, 4 and 6, because it is not possible to fill space with other symmetries.



(I am excluding "quasi-crystals" which can show 5-fold symmetry, discovered by Dan Schechtman in 1982, Nobel prize in Chemistry 2011)

Note this restriction does not apply to molecular symmetry, for example C-reactive protein has 5-fold symmetry, GroEL has 7-fold etc

Crystal systems

There are restrictions in simmetry that have to be applied to the axes and angles of the unit cell. This constrains the simmetry of the lattice and gives rise to the 7 crystal systems.

Crystals can be described by shape of unit cell (called lattice type) and symmetry type of asymmetric unit

65 different space groups possible for natural biomolecules

Example P2₁2₁2₁



Triclinic	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$
Monoclinic	$a \neq b \neq c$ $\alpha = \beta = 90^{\circ} \neq \gamma$ (1st setting) $\alpha = \gamma = 90^{\circ} \neq \beta$ (2nd setting)
Orthorhombic	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90^{\circ}$
Tetragonal	$a = b \neq c$ $\alpha = \beta = \gamma = 90^{\circ}$
Cubic	a = b = c $\alpha = \beta = \gamma = 90^{\circ}$
Hexagonal	$a = b \neq c$ $\alpha = \beta = 90^{\circ}; \gamma = 120^{\circ}$
Trigonal (Rhombohedral)	same as hexagonal (a = b = c; $\alpha = \beta = \gamma$)

Crystal systems and Bravais lattices



The French crystallographer Auguste Bravais established that in three-dimensional space only fourteen different lattices may be constructed.

The Bravais lattices are of three different types. A **primitive lattice** has only a lattice point at each corner of the three-dimensional unit cell. A **body-centered lattice** contains not only lattice points at each corner of the unit cell but also contains a lattice point at the center of the three-dimensional unit cell. A **face-centered lattice** possesses not only lattice points at the corners of the unit cell but also at either the centers of just one pair of faces or else at the centers of all three pairs of faces.

A space group is a representation of the ways that the macroscopic and microscopic symmetry elements (operations) can be self-consistently arranged in space. There are 230 unique manners in which this can be done and, thus, 230 space groups.



-Not all 230 space groups are allowed for protein crystals.

- Mirror planes and inversion centers would change the symmetry of the aminoacids from Lto D-aminoacids, never found in proteins.

- Rotation and screw axis are allowed.

2 Rotation Axis (ZINJAH)



4 Rotation Axis (FOYTAO)



6 Rotation Axis (GIKDOT)







Bovine galectin

Pea lectin

Concanavalin A





Serum amyloid P-component

Limulus CRP

2₁ Screw Axis (ABEBIS)





3₁ Screw Axis (AMBZPH)













Protein crystals





Protein crystallization

There is an energy barrier to crystallization.

In order for a protein to crystallize it must overcome an energy barrier analogous to that for conventional chemical reactions.



Methods of Protein Crystallization

Vapour Diffusion

In a vapour diffusion experiment, small volumes of precipitant and protein mixed together and the drop equilibrated against a larger reservoir of solution containing precipitant or another dehydrating agent.

- Drop equalizes with reservoir
- Volume of drop slowly decreases
- Protein concentration slowly increases





Crystallization tools

24- and 96-well plates.







Robot





Low molecular weight PEG based

Data Sheet



🧯 🗢 Jena Bioscience

JBScreen Classic 1 Cat.-No: CS-101L

Number	Precipitant 1	Precipitant 2	Buffer	pH	Additive
A 1	15 % w/v PEG 400	None	100 mM Sodium Acetate	4.6	100 mM Calcium Chloride
A 2	15 % w/v PEG 400	None	100 mM MES Sodium Salt	6.5	None
A 3	15 % w/v PEG 400	None	100 mM HEPES Sodium Salt	7.5	200 mM Magnesium Chloride
A 4	15 % w/v PEG 400	None	100 mM Tris-HCl	8.5	200 mM Sodium Citrate
A 5	25 % w/v PEG 400	None	100 mM Sodium Acetate	4.6	100 mM Magnesium Chloride
A 6	25 % w/v PEG 400	None	100 mM Tris-HCl	8.5	200 mM Lithium Sulfate
B 1	28 % w/v PEG 400	None	100 mM HEPES Sodium Salt	7.5	200 mM Calcium Chloride
B 2	30 % w/v PEG 400	None	100 mM Sodium Acetate	4.6	100 mM Calcium Chloride
B 3	30 % w/v PEG 400	None	100 mM MES Sodium Salt	6.5	100 mM Sodium Acetate
B 4	30 % w/v PEG 400	None	100 mM MES Sodium Salt	6.5	100 mM Magnesium Chloride
B 5	30 % w/v PEG 400	None	100 mM HEPES Sodium Salt	7.5	200 mM Magnesium Chloride
B 6	30 % w/v PEG 400	None	100 mM Tris-HCl	8.5	200 mM Calcium Citrate
C 1	30 % w/v PEG 550 MME	None	100 mM Bicine	9.0	100 mM Sodium Chloride
2	28 % w/v PEG 550 MME	None	100 mM MES Sodium Salt	6.5	10 mM Zinc Sulfate
с з	25 % w/v PEG 1000	None	100 mM HEPES Sodium Salt	7.5	None
C 4	30 % w/v PEG 1000	None	100 mM Tris-HCl	8.5	None
C 5	15 % w/v PEG 1500	None	None		None
6 0	20 % w/v PEG 1500	None	100 mM HEPES Sodium Salt	7.5	None
D 1	30 % w/v PEG 1500	None	None		None
D 2	20 % w/v PEG 2000 MME	None	100 mM Tris-HCl	8.5	10 mM Nickel (II) Chloride
D 3	25 % w/v PEG 2000 MME	None	None		None
D 4	30 % w/v PEG 2000 MME	None	100 mM MES Sodium Salt	6.5	100 mM Sodium Acetate
D 5	20 % w/v PEG 3000	None	100 mM HEPES Sodium Salt	7.5	200 mM Sodium Acetate
D 6	30 % w/v PEG 3000	None	100 mM Tris-HCl	8.5	200 mM Lithium Sulfate

Jena Bioscience GmbH | Löbstedter Str. 80 | 07749 Jena, Germany | Tel.:+49-3641-6285 000 | Fax:+49-3641-6285 100 http://www.jenabioscience.com

Last update: May 08, 200

High molecular weight PEG based

Data Sheet



🧧 💿 Jena Bioscience

٠

JBScreen Classic 5 Cat.-No: CS-105L

Number	Precipitant 1	Precipitant 2	Buffer	pH	Additive 1	Additive 2
A 1	12 % w/v PEG 8000	5 % w/v Glycerol	None		100 mM Potassium Chloride	None
A 2	12 % w/v PEG 8000	10 % w/v Glycerol	None		500 mM Potassium Chloride	None
A 3	15 % w/v PEG 8000	None	None		200 mM Ammonium Sulfate	None
A 4	15 % w/v PEG 8000	500 mM Lithium Sulfate	None		None	None
A 5	15 % w/v PEG 8000	None	100 mM MES Sodium Salt	6.5	200 mM Sodium Acetate	None
A 6	15 % w/v PEG 8000	None	None		50 mM Ammonium Sulfate	100 mM Sodium Citrate
B 1	18 % w/v PEG 8000	None	100 mM HEPES Sodium Salt	7.5	200 mM Calcium Acetate	None
B 2	18 % w/v PEG 8000	2 % w/v 2-Propanol	100 mM HEPES Sodium Salt	7.5	100 mM Sodium Acetate	None
B 3	18 % w/v PEG 8000	None	100 mM Tris-HCI	8.5	200 mM Lithium Sulfate	None
B 4	20 % w/v PEG 8000	None	100 mM HEPES Sodium Salt	7.5	None	None
B 5	20 % w/v PEG 8000	None	100 mM MES Sodium Salt	6.5	200 mM Magnesium Acetate	None
B 6	20 % w/v PEG 8000	None	100 mM CHES	9.5	None	None
0 1	22 % w/v PEG 8000	None	100 mM MES Sodium Salt	6.5	200 mM Ammonium Sulfate	None
C 2	25 % w/v PEG 8000	None	None		200 mM Lithium Chloride	None
C 3	30 % w/v PEG 8000	None	None		200 mM Ammonium Sulfate	None
C 4	8 % w/v PEG 10000	None	100 mM Sodium Acetate	4.6	None	None
C 5	14 % w/v PEG 10000	None	100 mM Imidazole·HCI	8.0	None	None
C 6	16 % w/v PEG 10000	None	100 mM Tris-HCI	8.5	None	None
D 1	18 % w/v PEG 10000	20 % w/v Glycerol	100 mM Tris-HCI	8.5	100 mM Sodium Chloride	None
D 2	20 % w/v PEG 10000	None	100 mM HEPES Sodium Salt	7.5	None	None
D 3	30 % w/v PEG 10000	None	100 mM Tris-HCI	8.5	None	None
D 4	10 % w/v PEG 20000	None	100 mM MES Sodium Salt	6.5	None	None
D 5	17 % w/v PEG 20000	None	100 mM Tris-HCI	8.5	100 mM Magnesium Chloride	None
D 6	20 % w/v PEG 20000	None	None		None	None

Jena Blossience GmbH | Löbstedter Str. 80 | 07749 Jena, Germany | Tel.:+49-3641-6285 000 | Fax:+49-3641-6285 100 http://www.jenabloscience.com tast

Last update: May 08, 2009

Ammonium sulfate based

Data Sheet



Jena Bioscience

JBScreen Classic 6 Cat.-No: CS-106L

Number	Precipitant 1	Precipitant 2	Buffer pH		Additive
A 1	500 mM Ammonium Sulfate	1.0 M Lithium Sulfate	None		100 mM Sodium Citrate
A 2	1.0 M Ammonium Sulfate	None	None		None
A 3	1.0 M Ammonium Sulfate	None	100 mM Sodium Acetate 4.6		None
A 4	1.0 M Ammonium Sulfate	2 % w/v PEG 400	100 mM HEPES Sodium Salt	7.5	None
A 5	1.0 M Ammonium Sulfate	None	100 mM Tris-HCI	8.5	None
A 6	1.2 M Ammonium Sulfate	3 % w/v 2-Propanol	None		50 mM Sodium Citrate
B 1	1.5 M Ammonium Sulfate	15 % w/v Glycerol	100 mM Tris-HCI	8.5	None
B 2	1.6 M Ammonium Sulfate	500 mM Lithium Chloride	None		None
B 3	1.6 M Ammonium Sulfate	1.0 M Lithium Sulfate	None		None
B 4	1.6 M Ammonium Sulfate	None	100 mM HEPES Sodium Salt	7.5	200 mM Sodium Chloride
B 5	1.6 M Ammonium Sulfate	2 % w/v PEG 1000	100 mM HEPES Sodium Salt	7.5	None
B 6	1.8 M Ammonium Sulfate	None	100 mM MES Sodium Salt	6.5	None
C 1	2.0 M Ammonium Sulfate	2.0 M Sodium Chloride	None	None	
C 2	2.0 M Ammonium Sulfate	None	100 mM Sodium Acetate	4.6	None
C 3	2.0 M Ammonium Sulfate	5 % w/v PEG 400	100 mM MES Sodium Salt	6.5	None
C 4	2.0 M Ammonium Sulfate	None	100 mM Tris-HCI	8.5	None
C 5	2.2 M Ammonium Sulfate	None	None		None
C 6	2.2 M Ammonium Sulfate	20 % w/v Glycerol	None		None
D 1	2.4 M Ammonium Sulfate	None	None		100 mM Sodium Citrate
D 2	3.0 M Ammonium Sulfate	1 % w/v MPD	None		None
D 3	3.0 M Ammonium Sulfate	10 % w/v Glycerol	None		None
D 4	3.5 M Ammonium Sulfate	None	100 mM HEPES Sodium Salt	7.5	None
D 5	3.5 M Ammonium Sulfate	1 % w/v MPD	100 mM MES Sodium Salt	6.5	None
D 6	3.5 M Ammonium Sulfate	None	None		None

Jena Bloscience GmbH | Löbstedter Str. 80 | 07749 Jena, Germany | Tel.:+49-3641-6285 000 | Fax:+49-3641-6285 100 http://www.jenabloscience.com tast

Last update: May 08, 2009

Variables to consider:

-Temperature (usually 4 and 20 °C)

- Protein concentration (can vary from 2 mg/ml to 60 mg/ml)
- Protein buffer (pH, ionic strength, detergent, reucing agents, etc...)
- Method of crystallization
- Volume ratio protein:reservoir

Observation

Clear Drop









Precipitate





Phase Separation



Microcrystalline Precipitate





Needles





Plates











Optimization

When you get a crystal in a screening, an optimization is usually required to increase size and quality of the crystals.

Screening around the conditions

Screening around the conditions that give microcrystals is done by varying

- Precipitant concentration
- р рН
- Protein concentration
- Temperature
- Method

Additives and detergent screening

Popular additives are:glycerol, which may stop nucleation and may give you fewer, larger crystals, and has the advantage of doubling as a cryo-protectant.

Oil on the reservoir

Diffusion can be slowed down by adding a layer of oil on the top of the hanging drop well The proportion of light silicone oil to heavy paraffin oil must be mixed in different proportions to get an appropriate rate of diffusion

Streak seeding

- Make up a new drop with lower protein or precipitant
- Touch a whisker to the seed crystal
- Draw a line in the equilibrated drop
- Wait











Crystallization of protein-ligand complexes

Strategies for obtaining crystals of protein-ligand complexes:

- 1. Co-crystallization:
- (i) co-expression of the protein with the ligands of interest,
- (ii) use of the ligands during protein purification,
- (iii) the ligand is added before crystallization trials

- 2. Soaking:
- Crystals are often put into stabilization buffers before they are immersed in the ligand solution. These buffers may contain increased concentrations of the precipitant(s) and a stepwise/gradual increase in reagent concentration or the introduction of a cryoprotectant may be required so that the crystals are not damaged.

Crystallization of protein-ligand complexes





Ghosh et al., Nature 2009

Crystal mounting

- Crystals should be kept in the saturated vapour of their mother liquor or at a sufficiently low temperature to prevent evaporation of the solvent (40-60% of the crystal volume is occupied by solvent).

- Flash freezing or shock cooling (use of cryoprotectant such as glycerol, MPD, ethyleneglycol and low molecular weight PEG) in liquid nitrogen or in a stream of cold nitrogen gas at a temperature range of 100-120°K.

Capillary

Loops







Cryo-mounting

Cryoprotection is effectively accomplished during harvesting, when the crystals are scooped up from the drop in <u>cryo-loops</u> and briefly swept though a cryoprotectant before being dipped into liquid nitrogen. Common cryoprotectants are ethylene glycol (the anti-freeze in automobile radiators), glycerol, higher alcohols, polyethylene glycols (PEGs), or high concentration solutions of sucrose or salts. Once the protein crystals are flash-cooled and stored in pucks in a liquid nitrogen dewar (dry shipper), they can be safely sent to a synchrotron for data collection.

Radiation damage.

Proteins are sensitive to X-ray radiation damage. The energy range of X-rays used for diffraction is in the 6 to 15 keV range, which is in fact severely ionizing radiation. The ionizing absorption events create radicals, which rapidly destroy any protein crystal, particularly at dose rates experienced at synchrotrons. An efficient way to suppress radiation damage by slowing down the kinetics of the radical reactions is cryogenic cooling. Rapidly quenching or flash-cooling crystals to liquid nitrogen temperatures, either in cold nitrogen gas streams or directly into liquid nitrogen, will strongly reduce radiation damage. To prevent the formation of crystalline ice during flash-cooling of the crystals, cryoprotectants, present in the mother liqor or added to the mother liquor, are necessary.

Benefits of Cryo-Cooling

Many factors contribute to improvements in data quality during cryo-protection:

- reduced thermal vibrations,
- enhanced signal-to-noise ratio,
- reduced conformational disorder,
- a higher limiting resolution.