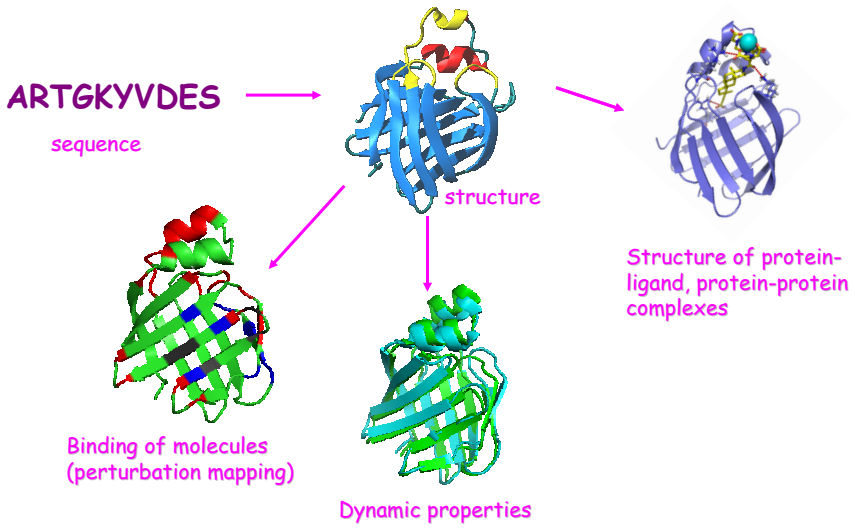
Protein isotopic enrichment for NMR studies

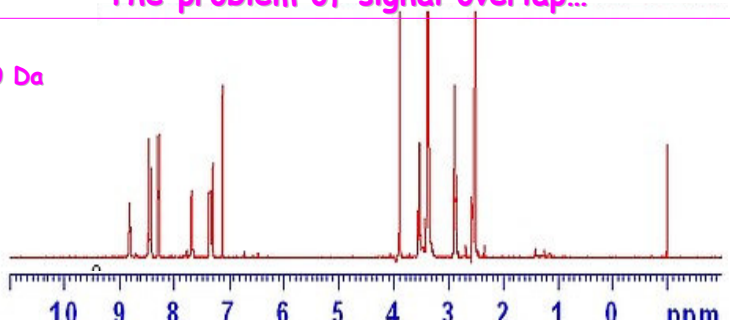
Protein NMR studies



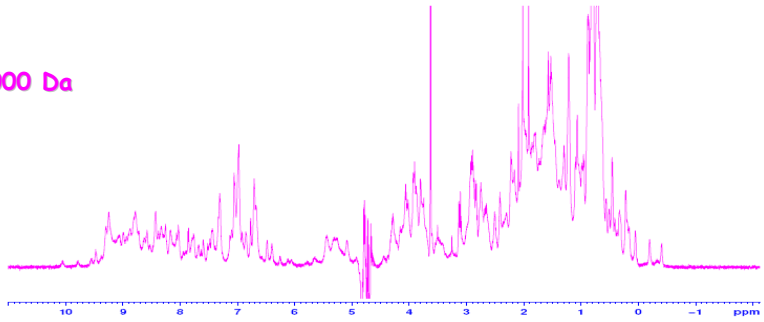
We need to analyze very carefully very complex spectra

The problem of signal overlap…

~300 Da

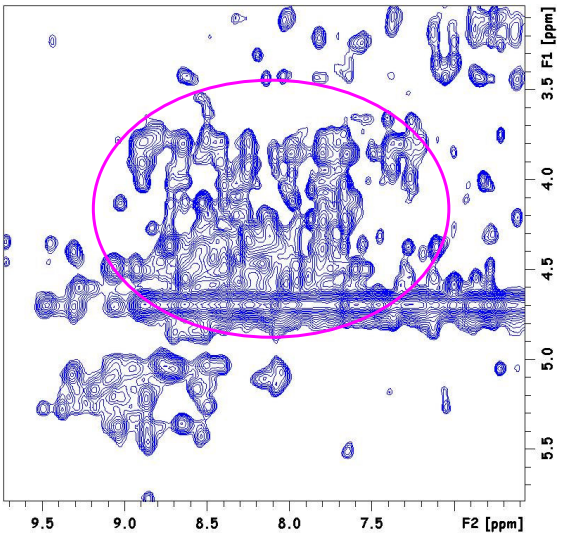


~15000 Da



How can we solve this problem?

Acquiring two dimensional spectra…



Again we can have signal overlap!!!

The limit in size for protein homonuclear NMR is about 100 amino acids…

NMR active nuclei

Isotope	Spin (I)	Natural abundance	Magnetogyric ratio $g/10^7$ rad $T^{-1}s^{-1}$	NMR frequency MHz (2.3 T magnet)
^{1}H	1/2	99.985 %	26.7519	100.000000
^{2}H	1	0.015	4.1066	15.351
^{13}C	1/2	1.108	6.7283	25.145
^{14}N	1	99.63	1.9338	7.228
^{15}N	1/2	0.37	-2.712	10.136783
^{17}O	5/2	0.037	-3.6279	13.561
^{19}F	1/2	100	25.181	94.094003
^{23}Na	3/2	100	7.08013	26.466
^{31}P	1/2	100	10.841	40.480737
^{113}Cd	1/2	12.26	-5.9550	22.193173

Isotopic enrichment

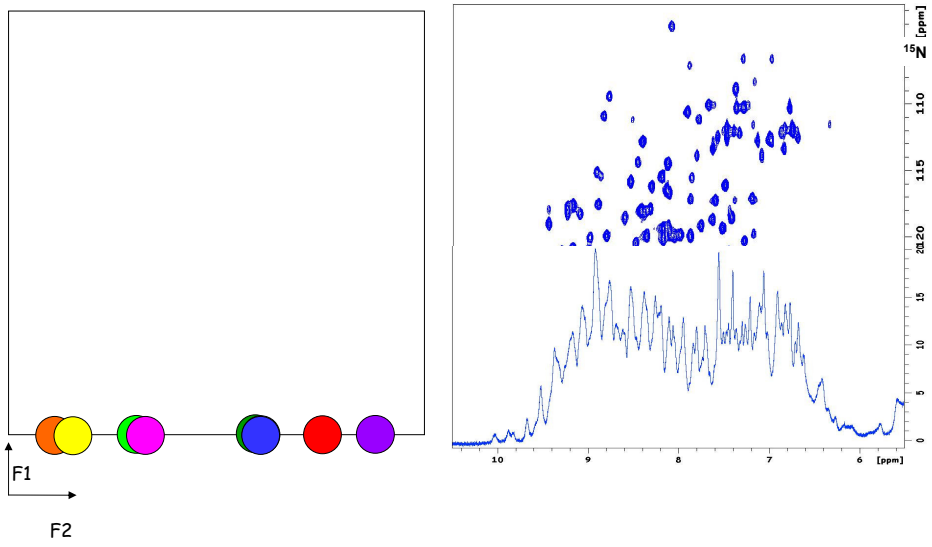
^{1}H (99.98%)
^{13}C (1.108%)
^{15}N (0.37%)

We have additional information in NMR active nuclei in the biomolecules, but there is the limit of the natural abundance.

It is possible to substitute inactive with active nuclei through the isotopic enrichment.

The active nuclei can be used to transfer the magnetization through covalent bonds using heternuclear J coupling

Heteronuclear multidimensional spectra



What it is needed for NMR purposes

Overexpression of the protein to be analyzed (mg of the sample!) → The biological material can be difficult to find

Isotopically enriched samples → Impossible to obtain *in vivo*

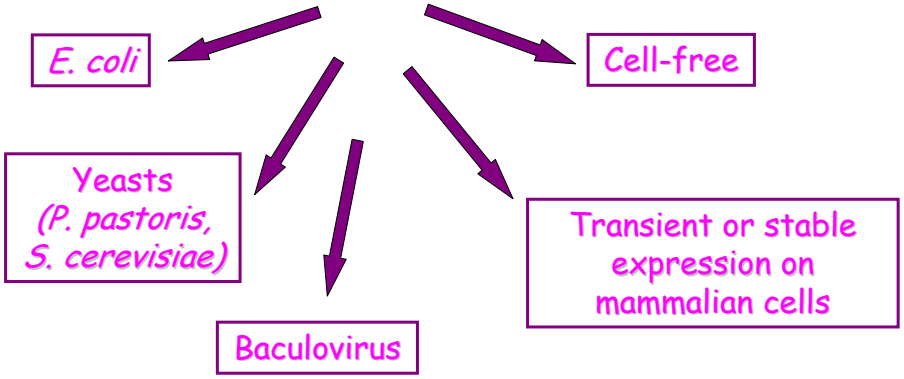
The steps to obtain an NMR sample

- *Cloning*
- *Mutagenesis*
- *Protein expression (unenriched and enriched)*
- *Purification*
- *Final steps and NMR sample preparation*

The steps to obtain an NMR sample

Cloning

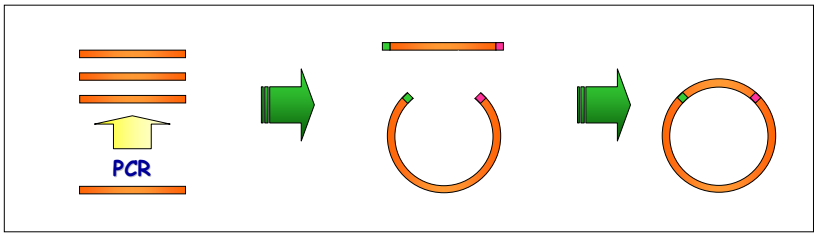
How to choose the expression host



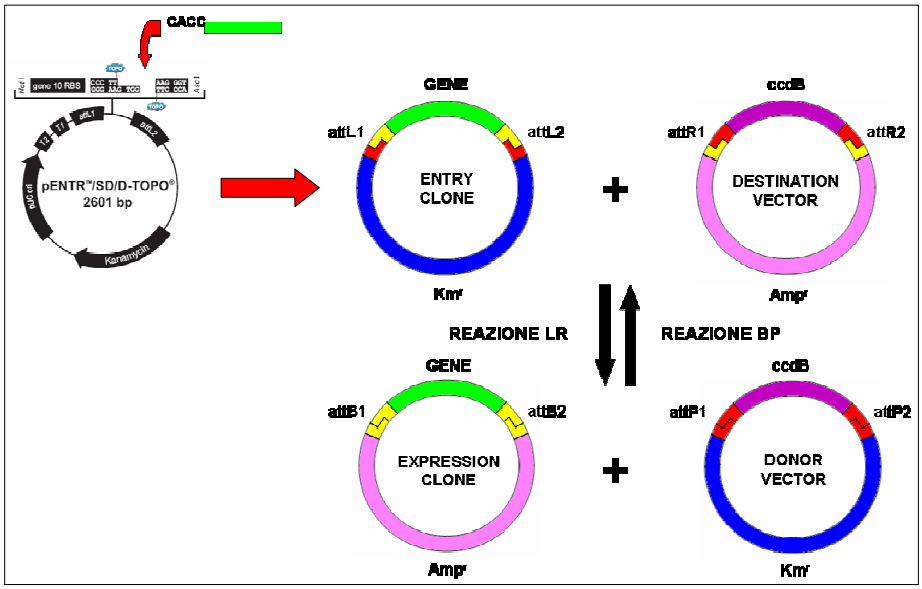
We have to pay attention to the costs of the culture media!

Cloning method: restriction-ligation

Amplification of the gene of interest → Digestion of the amplified gene and of the plsmid using rstriction enzymes → Ligation



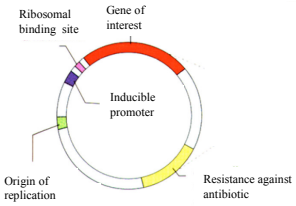
Cloning method: Gateway



How to choose the plasmid

Depending on how we want to engineer the protein construct

- C or N terminal tag for protein purification (his-tag, GST, etc…)
- Cleavage sites to get rid of the tags (FactorXa, thrombin, TEV)
- choice of the tag (His-tag, fusion proteins, etc…)
- signal peptides for cellular localization



How to choose the bacterial strain

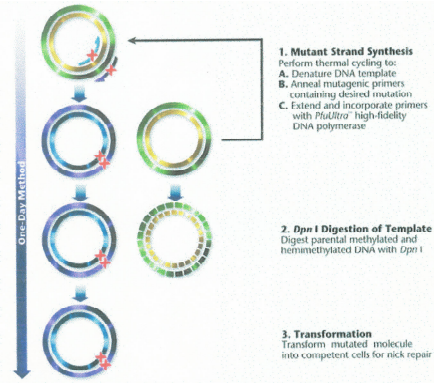
It is good idea to try different *E. coli* strains to test their ability to survive in the growth media and to express, with a good yield, the protein of interest.
The strains more often used are
BL21DE3, BL21DE3pLysS, Rosetta, Codon plus, Origami, SG…

BL21 strain lacks two proteases
pLysS aviods leaky expression i.e. expression without induction
Rosetta and Codon plus strains are designed to enhance the expression of eukaryotic proteins containing rare codons
Origami strain greatly enhances disulfide bond formation in the cytoplasm

The steps to obtain an NMR sample

Mutagenesis

Analysis → Structural
Analysis → Functional



The mutation can be done:

-Substituting the residue with an Ala

-Substituting the residue with one conserved in the family

The steps to obtain an NMR sample

Protein expression (unenriched and enriched)

Optimization of the growing conditions

Growth media

rich medium (i.e. LB)
minimal medium (i.e. M9) → ENRICHED PROTEIN

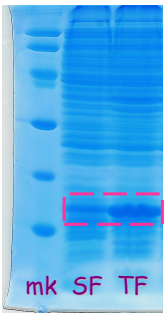
Induction conditions

temperature
time
amount of IPTG → HIGH YIELD OR PROTEIN SOLUBILITY?

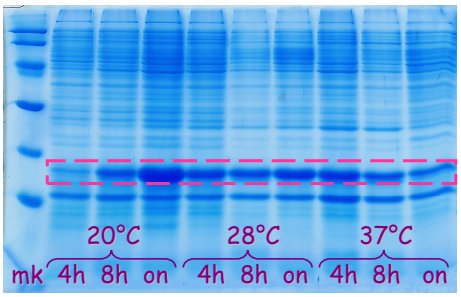
Cell lysis → Isolation of the soluble/unsoluble fraction (refolding)

The steps to obtain an NMR sample

An example



37°C



(soluble fractions)

SF soluble fraction
TF total fraction

Isotopic enrichment procedures

The most useful isotopes (^{13}C, ^{15}N) to study proteins by NMR are not naturally abundant

Uniform → the bacteria are grown in a medium containing only sources of the isotope needed. All the atoms of the protein are enriched.

Residue-specific → we enrich only the atoms of the "interesting" residues of the protein

^{13}C e ^{15}N have high costs, this implies that each step for sample preparation has to be optimized to obtain as much protein as possible!

Isotopic enrichment procedures

Uniform enrichment

We grow the *E. coli* cells in a minimal medium (i.e. M9) containing selected labelled nutrients (e.g. $^{15}NH_4Cl$, ^{13}C-glucose) or in a labelled ready-to-use medium (bacterial or algal hydrolysate).

Bacterial growth is generally higher in ready-to-use media

Minimal media are generally less expensive

Minimal media

It is a growing broth prepared in the lab starting from simple reagents.
It is very important to control the carbon and nitrogen sources

Isotopic enrichment procedures

Minimal medium composition

Carbon source: glucose, glycerol, acetate, succinate o methanol

Nitrogen source : NH_4Cl o $(NH_4)_2SO_4$

Salts: NaCl/KCl, $MgSO_4$, $CaCl_2$

Buffer solution: generally phosphate at pH 7.5

When we use ^{13}C e ^{15}N the isotopically enriched sources are reduced at minimal possible level to reduce the costs!!!

Classic protocol → the culture is grown and induced directly in the enriched minimal medium

Mixed protocol → the culture is grown in unlabeled rich medium (i.e. LB) and, right before induction, the cells are harvested, washed and resuspended in the enriched minimal medium

Isotopic enrichment procedures

Residue-specific enrichment

It is achieved using *E. coli* auxotrophic strains which are grown in the presence of selected labeled aminoacids or using more sophisticated techniques such as cell-free.

The latter approach allows to simplify problems due to protein size. It is possible to focus on specific regions of the protein

Auxotrophic strains → we need a specific strain for each residue

Cell-free → more flexible when using labeled aminoacids

The steps to obtain an NMR sample

Purification

The isolation of the protein is done through several cromatography steps.

The separation is based on one of the following physical properties:

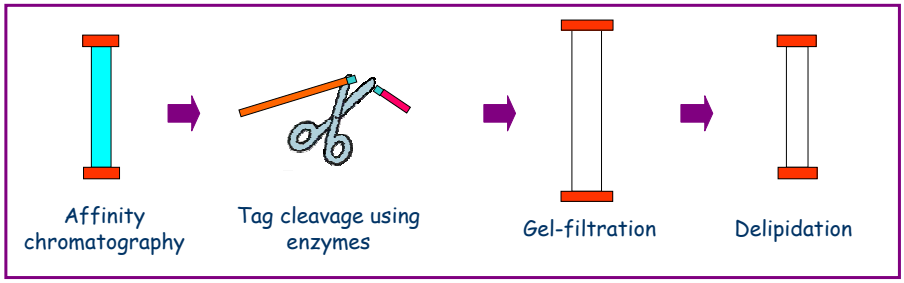
Solubility ⟶ i.e. selective precipitation

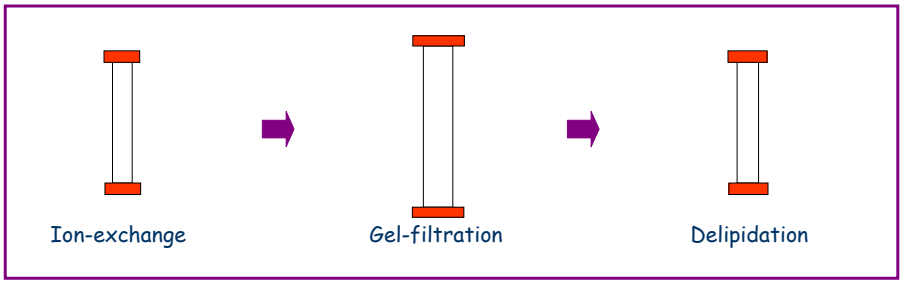
Charge ⟶ i.e. ion exchange

Affinity ⟶ i.e. affinity chormatography

Molecular weight ⟶ i.e. gel filtration

The steps to obtain an NMR sample





The steps to obtain an NMR sample

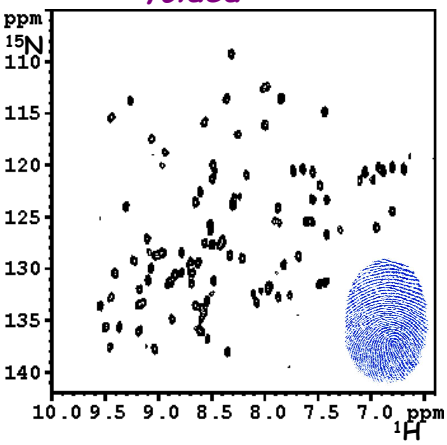
Final steps and NMR sample preparation

- (Refolding)
- Deglycosylation (not if expressed in *E. coli)*
- Delipidation
- Buffer exchange ⟶ Choose buffers suitable for NMR
- Concetrate sample ⟶ [P] in the mM range
- Complexes formation
- Storing ⟶ i.e. freezing, liophilization,...
- Check protein identity and fold ⟶ i.e. mass spectrometry, NMR analysis

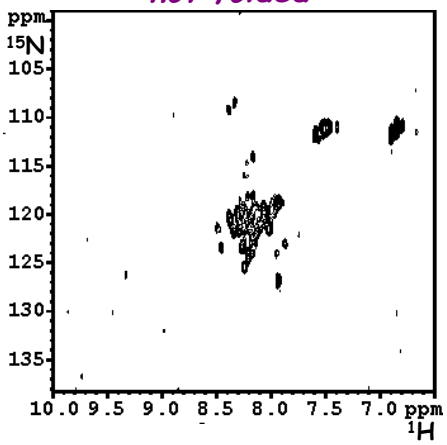
Preliminary characterization using NMR

An easy-to-run spectrum lasting 10 minutes is enough to check protein folding!

folded



not folded



Even proteins have "fingerprints"!