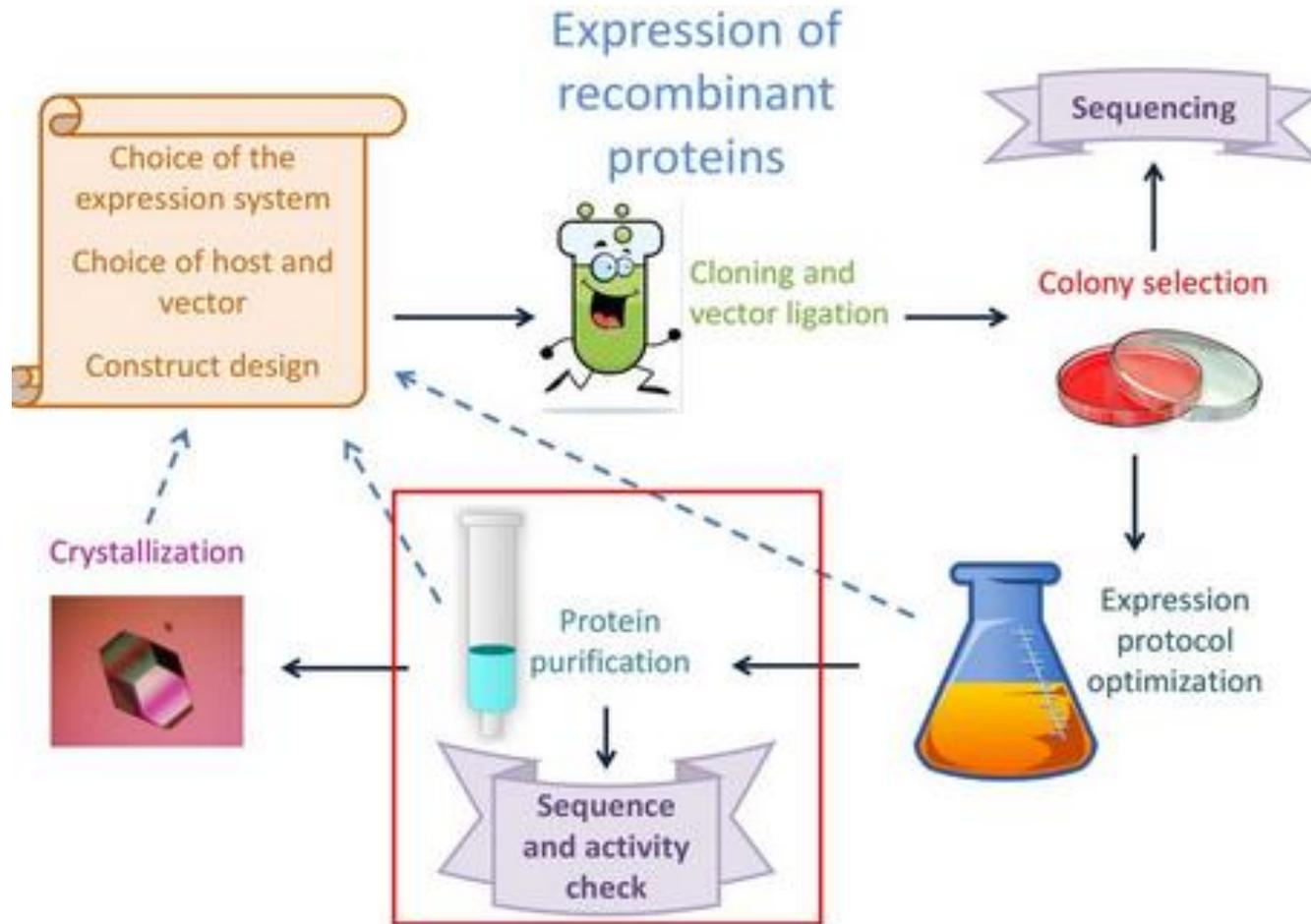


Proteine ricombinanti per studi strutturali: espressione, purificazione e ottimizzazione

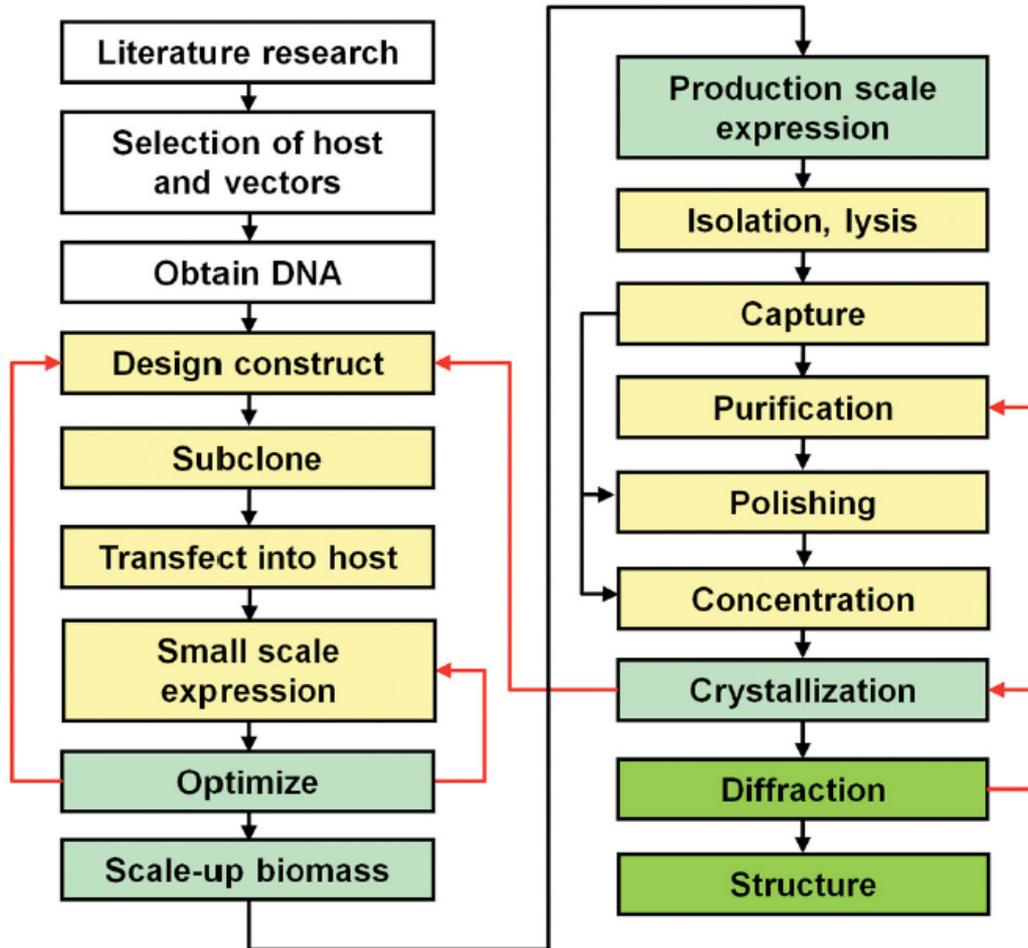
Sonia Covaceuszach

Istituto di Cristallografia - C.N.R.
Area Science Park - Basovizza
sonia.covaceuszach@ic.cnr.it

Espressione/purificazione di proteine ricombinanti per studi strutturali



Espressione/purificazione di proteine ricombinanti per studi strutturali flowchart



Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica

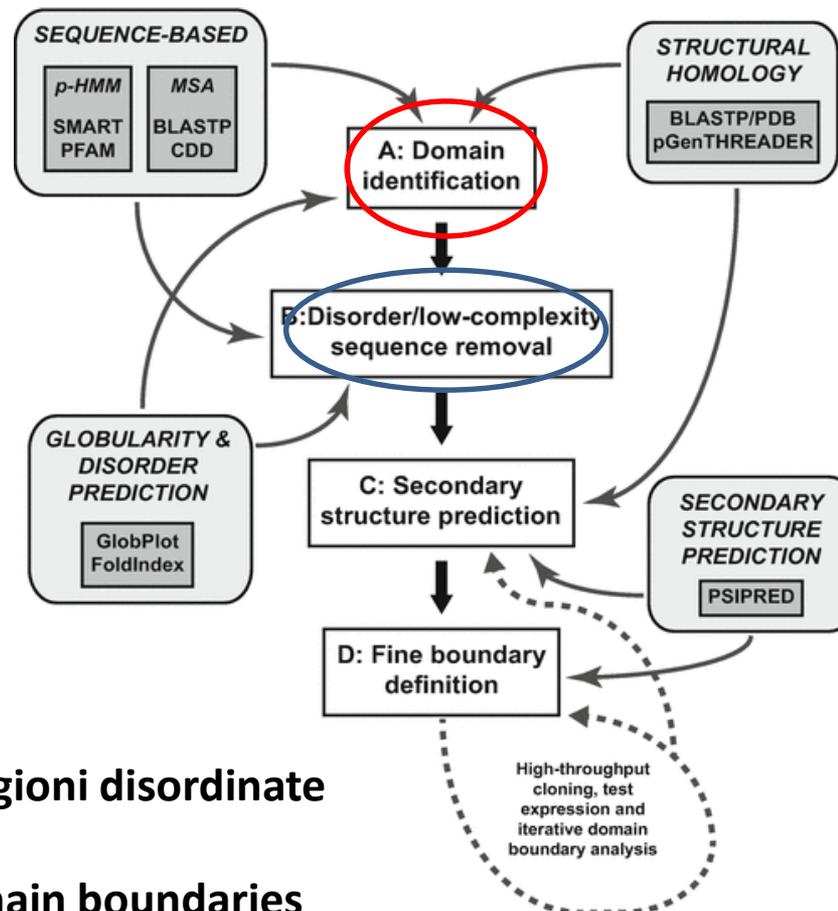
1) Analisi della sequenza (Codon usage, Siti di proteolisi, sequenze di secrezione)

		Seond letter					
		U	C	A	G		
U	UUU] Phe	UCU] Ser	UAU] Tyr	UGU] Cys	U C A G		
	UUC]	UCC]	UAC]	UGC]			
	UUA] Leu	UCA]	UAA Stop	UGA Stop			
	UUG]	UCG]	UAG Stop	UGG Trp			
C	CUU]	CCU] Pro	CAU] His	CGU]	U C A G		
	CUC] Leu	CCC]	CAC]	CGC] Arg			
	CUA]	CCA]	CAA] Gin	CGA]			
	CUG]	CCG]	CAG]	CGG]			
A	AUU] Ile	ACU] Thr	AAU] Asn	AGU] Ser	U C A G		
	AUC]	ACC]	AAC]	AGC]			
	AUA]	ACA]	AAA] Lys	AGA] Arg			
	AUG Met	ACG]	AAG]	AGG]			
G	GUU]	GCU] Ala	GAU] Asp	GGU]	U C A G		
	GUC] Val	GCC]	GAC]	GGC]			
	GUA]	GCA]	GAA] Glu	GGA] Gly			
	GUG]	GCG]	GAG]	GGG]			

Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica



2) Definizione delle regioni disordinate

3) Definizione dei domain boundaries

Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: A) disorder prediction:

1) FOLDINDEX: <https://fold.proteopedia.org/cgi-bin/index>

$$I = 2.785 \langle H \rangle - | \langle R \rangle | - 1.151$$

con $| \langle R \rangle |$ carica netta media
 $\langle H \rangle$ idrofobicità media

Summary:

Number Disordered Regions: 4
Longest Disordered Region: 72
Number Disordered Residues: 147
Predicted disorder segment: [260]-[264] length: 5 score: -0.01 ± 0.01
Predicted disorder segment: [286]-[293] length: 8 score: -0.02 ± 0.02
Predicted disorder segment: [389]-[460] length: 72 score: -0.25 ± 0.10
Predicted disorder segment: [462]-[523] length: 62 score: -0.08 ± 0.05

```
1 HWPSEPSEAV RDWENQLEAS MHSVLSDLHE AVPTVWGIPD GTAVVGRSFR
51 VTIPTDLIAS SGDIKVSAA GKEALPSWLH WDSQSHTLEG LPLDTDKGVH
101 YISVSATRLG ANGSHIPQTS SVF SIEVYPE DHSELQSWRT ASPDPGEVVS
151 SACAADEPVT VLTVILDADL TKMTPKQRID LLHRMRSFSE VELHNMKLVF
201 VVNNRLFDMs AFMAGPGNAK KVVENGALLS WKLGC SLNQN SVPDIHGVEA
251 PAREGAMSAQ LGYPVVGWHI ANKKPPLPKR VRRQIHATPT PVT AIGPPTT
301 AIQEPPSRIV PTPTSPAIA PTE TMAPPVR DPVPGKPTVT IRTRGAI IQT
351 PTLGPIQPTR VSEAGTTVPG QIRPTMTIPG YVEPTAVATP PTTT TKKPRV
401 STPKPATPST DSTTTTTRRP TTKPRTPRPV PRVTTKVSIT RLETASPPTR
451 IRTTTSGVPR GGEPNQRPEL KNHIDRVD AW VGTYFEVKIP SDTFYDHEDT
501 TTDK LK L TLK LREQQLVGEK SWVQFNSNSQ LMYGLP DSSH V GKHEYFMHA
551 TDKGGLSAVD AFEIHVHRRP QGDRAPARFK AKFVGD PALV LNDIHKKIAL
601 VKKLAFAFGD RNCSTITLQN ITRG
```

(Predicted disordered segment)

Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: A) disorder prediction:

2) GLOBPLOT: <http://globplot.embl.de/>

PROPENSIONE al disordine di ogni aa

$$P = RC - SS$$

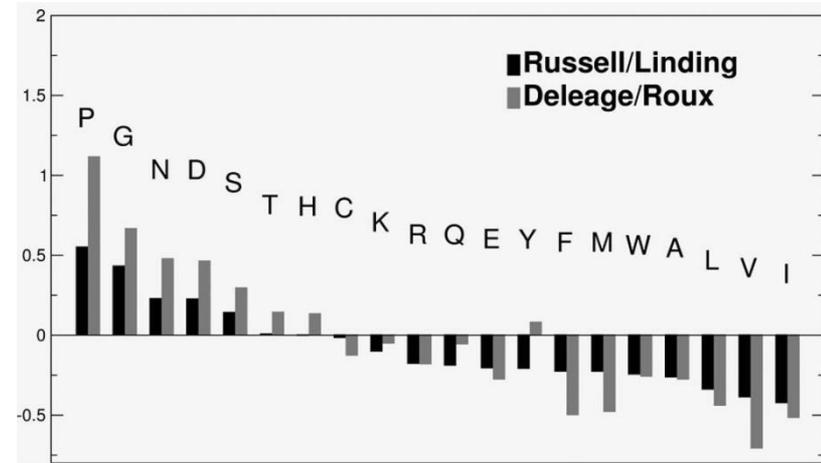
RC random coil (loops, turns)

SS strutture secondarie regolari

Calcolata su un insieme non ridondante di proteine da ciascun rappresentante di ogni superfamiglia

nel database SCOP (Structural Classification of Proteins):

2022-06-29 include **72,544** domini non ridondanti in **861,631** strutture.



Disordered by Russell/Linding definition

```
>none_Disorder 2-6, 88-93, 111-117, 137-149, 237-244, 290-338, 352-377, 382-431, 449-469, 531-540, 616-624
hWPSEpSeav rdwenqleas mhsvlsdlhe avptvvgipd gtavvgrsfr vtiptdliaS sgdiiKvsaa gkealpswlh wdsqshtLEG LPLdtDKgvh yisvsatrlg ANgSHIPqts svfsieVype dhSelqSVRT
ASPDpGEVVs sacaaDepvt vltvildadl tkmtpkqrid llhrmrSfse velhnmklvp vvnrlfdms afmagpgnak kvvengalls wklGCSlnQN SVPDIhgvea pareGamsaq lgyppvGwhi ankkpplpkr
vrrqihatpT PVTaIGpPTt AIQePPSRIV PPTSPaIAP PTeTMappvR DpVpGkptvt irtrgaiIqt pTLGPIQpTR VSEAGttVpG QIRpTmtipg yVEPTAVATP PTTtTKKpRV STPKPATpST DSttttTRRP
TKKpRtPRpV Prvttkvsit rletaspPTr IRtTtSGVpR GGePnQrPeL knhidrvdaw vGtyfeVkip sdtfydhedt ttdklkltlk lreqqlvgek swvqfnSnsq lMYGLpDSSh vGkheyfMha tdkGGLsAVD
afeihVhrrp qGdraparfk akfvGdpalv lndihkKial vkklafafgd rncstITLQN ITRG
```

Potential globular domains (GlobDoms) by Russell/Linding definition

```
>none_GlobDoms 1-289, 470-622
HWPSEpSEAV RDWENqLEAS MHSVLSDLHE AVPTVVGIPD GTAVVGRSFR VTIPTDLIAS SGDIIKVSAA GKEALPSWLH WDSQSHTLEG LPLDTDKGVH YISVSATRLG ANGSHIPQTS SVFSIEVYPE DHSELQSVRT
ASPDpGEVVS SACAADEPVT VLTVILDADL TKMTPKQRID LLHRMRSFSE VELHNMKLVP VVNNRFLDMS AFMAGPGNAK KVVENGALLS WKLGCsLNQN SVPDIHGVEA PAREGAMSAQ LGYPVVGWHI ANKKPPLPKR
VRRQIHATPT PVTaigpPTt aiqepPsriv ptptspaiap ptetmappvr dpvpgkptvt irtrgaiIqt ptlgpiqptr vseagttvpg qirptmtipg yveptavatp pttttkkprv stpkpatpst dstttttrrp
tkkprtprpv prvttkvsit rletaspPTr irtttsGvpr ggepnqrPeL KNHIDRVDAW VGTyFEVKIP SDtFYDHEDT TTDKLKLTLK LREQQLVGEK SWVQFNsNSQ lMYGLpDSSh VGKHEYFMHA TDKGGLSAVD
AFEIHVHRRP QGDRAPARFK AKFVGDPAIV LNDIHKKIAL VKKLAFAFGD RNCSTITLQN ITRG
```

Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

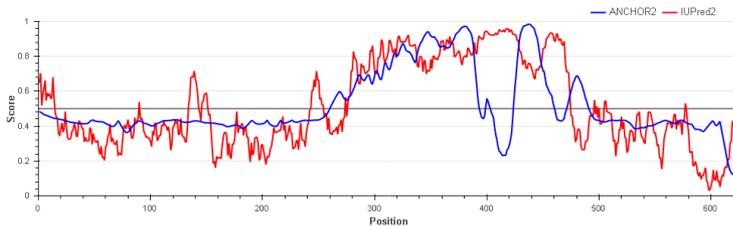
Analisi bioinformatica: A) disorder prediction:

3) <https://iupred2a.elte.hu/>

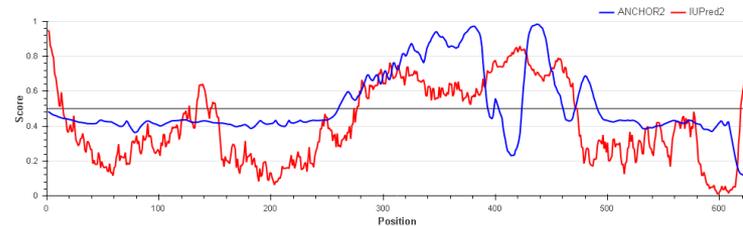
stima l'energia totale di interazione tra singole coppie di aa

(IPOTESI: sequenze disordinate non formano sufficienti interazioni stabilizzanti fra aa)

long disorder

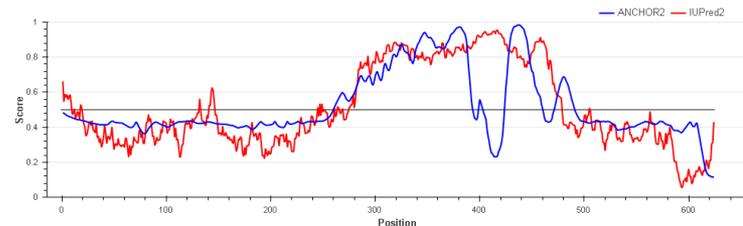


short disorder



structured domains

```
Number of globular domains: 2
  globular domain 1. 21-257
  globular domain 2. 480-623
hwpsepseav rdwenqleas MHSVLSDLHE AVPTVVGIPD GTAVVGRSFR VTIPTDLIAS
SGDIKVSAA GKEALPSWLH WDSQSHTLEG LPLDTDKGVH YISVSATRLG ANGSHIPQTS
SVFSIEVYPE DHSELQSVRT ASPDPGEVVS SACAADEPVT VLTVILDADL TKMTPKORID
LLHRMRSFSE VELHNMKLVV VVNNRFLDMS AFMAGPGNAK KVVENGALLS WKLGC SLNQN
SVPDIHGVEA PAREGAMsaq lgyppvvgwhi ankpplpkr vrrqihatpt pvtaigp ptt
aiqepssriv ptptspaiap ptetmappvr dpvpgkptvt intrgaiqt ptlqpiqptr
vseagttvpg qirptmtipg yveptavatp pttttkkprv stpkpatpst dstttttrrp
tkkprtprpv prvttkvsit rletasppt irtttsgvpr ggepnqrpel knhidrvdaW
VGTYFEVKIP SDTFYDHEDT TTDKLLKTLK LREQQLVGEK SWVQFNNSNQ LMYGLPSSSH
VGKHEYFMHA TDKGGLSAVD AFEIHVHRRP QGDRAPARFK AKFVGD PALV LNDIHKKIAL
VKKLAFAPGD RNCSTITLQN ITRg
```



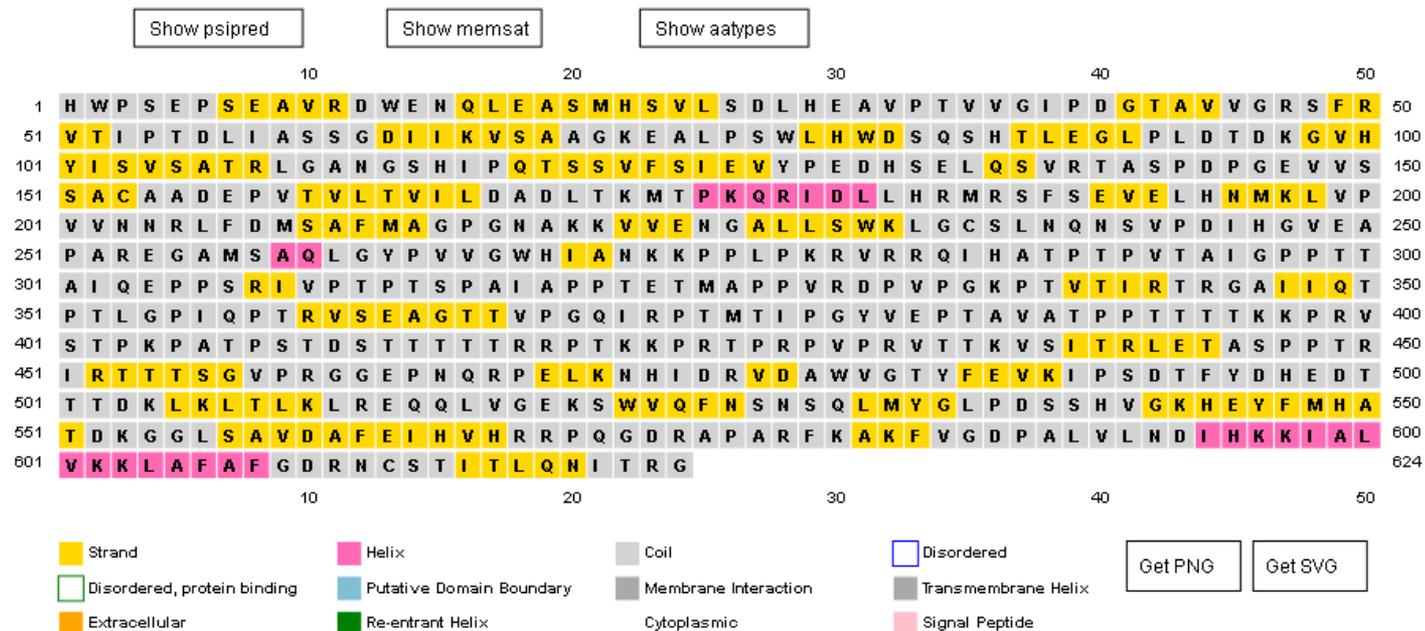
Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: B) Predizione della struttura secondaria

- <http://bioinf.cs.ucl.ac.uk/psipred/>

Sequence Plot



- http://www.compbio.dundee.ac.uk/jpred/index_up.html

Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: C) Allineamento con omologhi

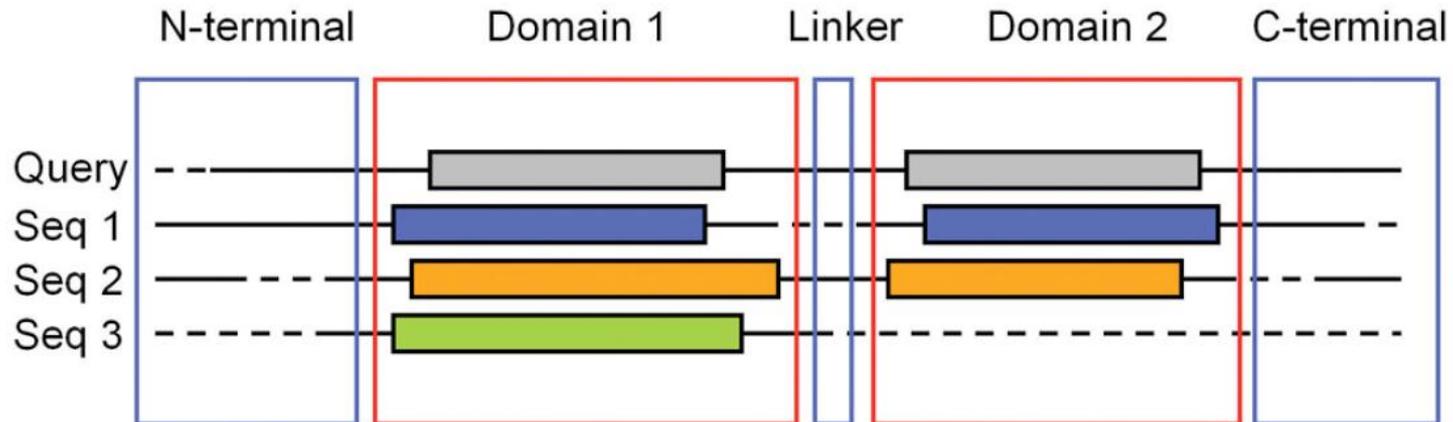
1) basato su struttura I

MSA:

- https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome
- https://www.ncbi.nlm.nih.gov/Structure/cdd/docs/cdd_search.html (Conserved Domains Database)

MSA+p-HMM (Hidden Markov Model):

- <https://pfam.xfam.org/>
- http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1 (Simple Modular Architecture Research Tool)



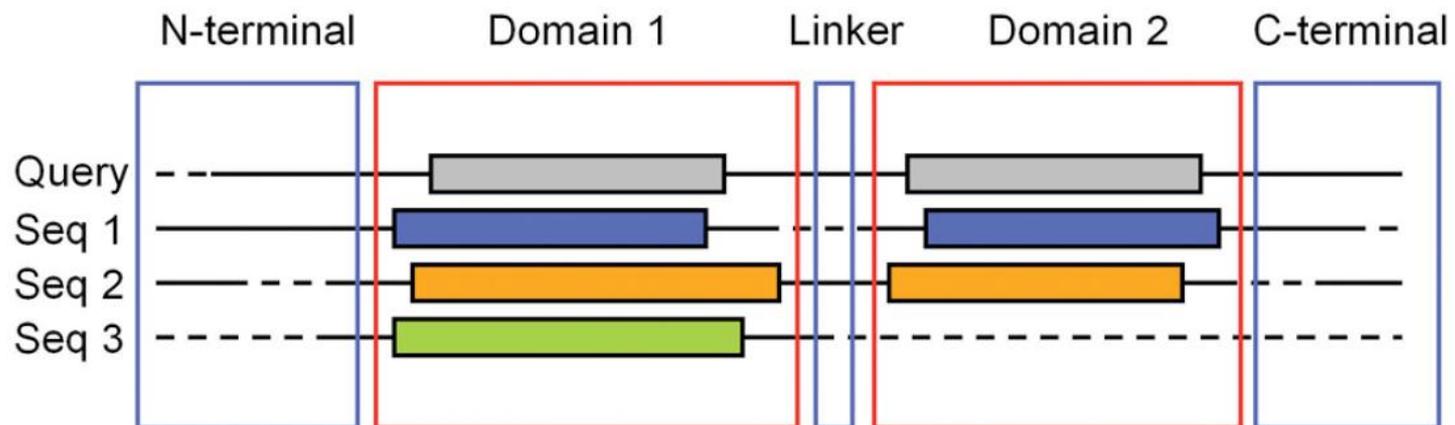
Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: C) Allineamento con omologhi

2) basato su struttura III

- <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins> (**Database PDB**)
- <http://bioinf.cs.ucl.ac.uk/psipred/> opzione: **pDomTHREADER** (Protein Domain Fold Recognition)



Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Analisi bioinformatica: C) Allineamento con omologhi

3) basato su approcci ibridi

<https://www.ics.uci.edu/~baldig/dompro.html> (**Protein Domain Prediction**)

Homology, Position Specific Scoring Matrix (PSSM), Secondary Structure, Solvent Accessibility

<http://bioinf.cs.ucl.ac.uk/psipred/> opzione **DomPred** (**Domain Prediction**)

PSSM, Fold Recognition, Secondary Structure

Espressione di proteine ricombinanti per studi strutturali

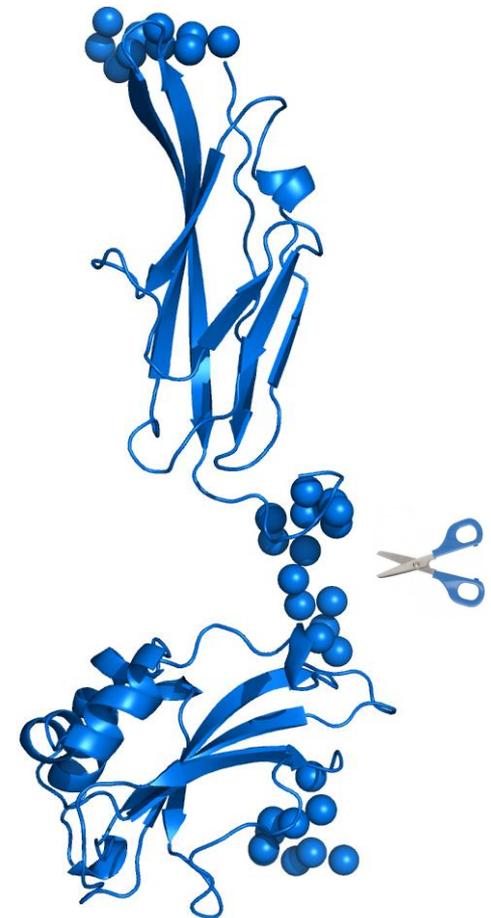
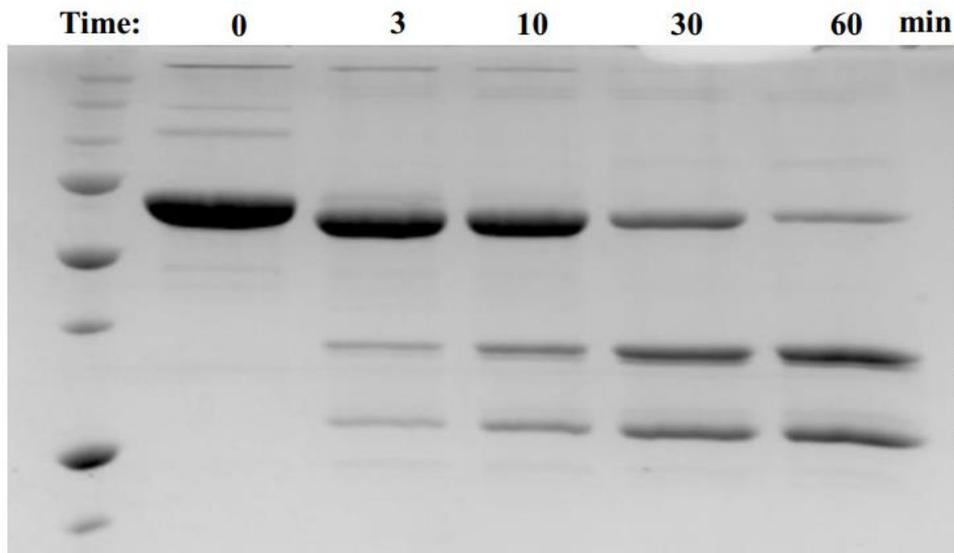
Design del costrutto

Metodi sperimentali: D)

Definizione dei domain boundaries

1. Limited Proteolysis - Mass Spectrometry (LP-MS)

https://web.expasy.org/peptide_cutter/



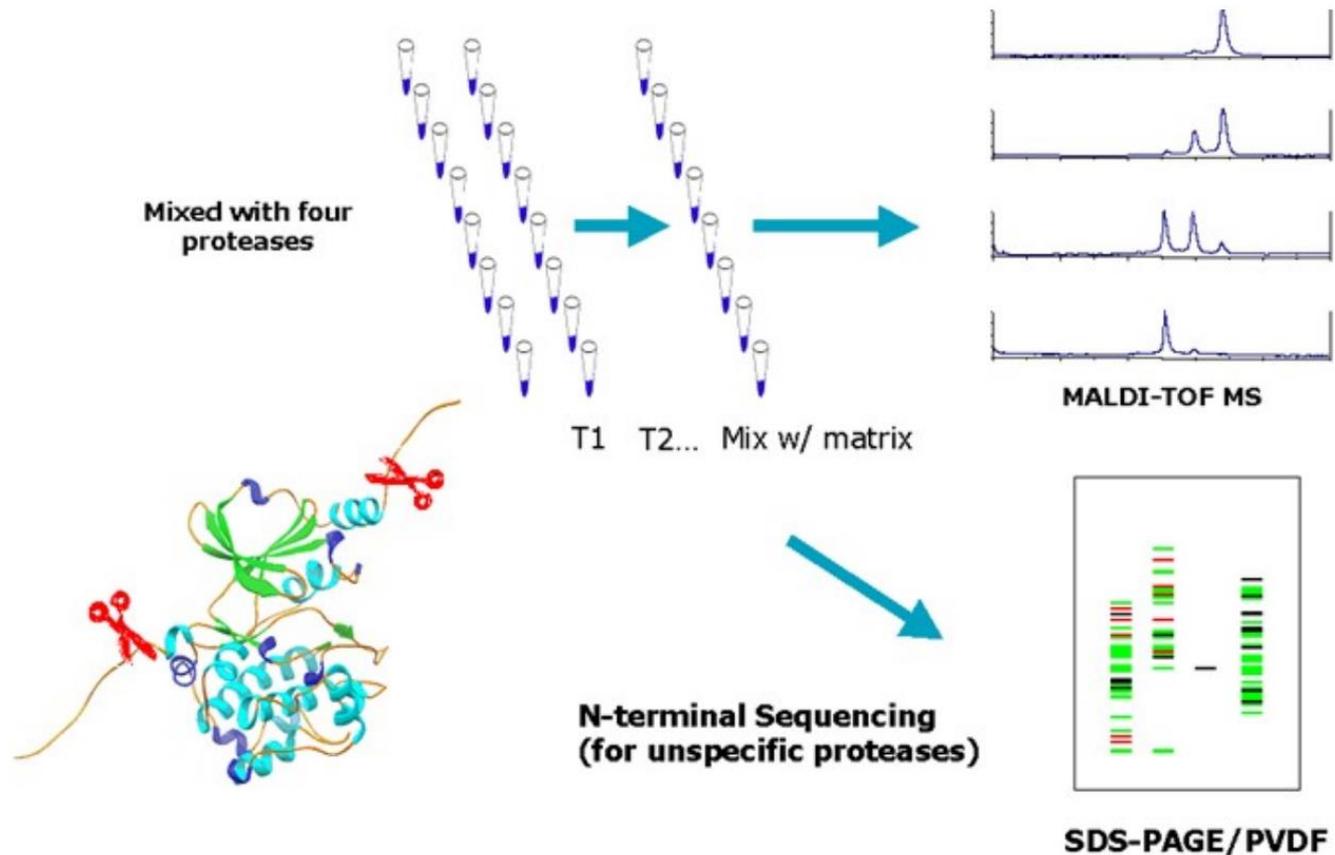
Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Metodi sperimentali: D)

Definizione dei domain boundaries

1. Limited Proteolysis Mass Spectrometry (LP-MS)



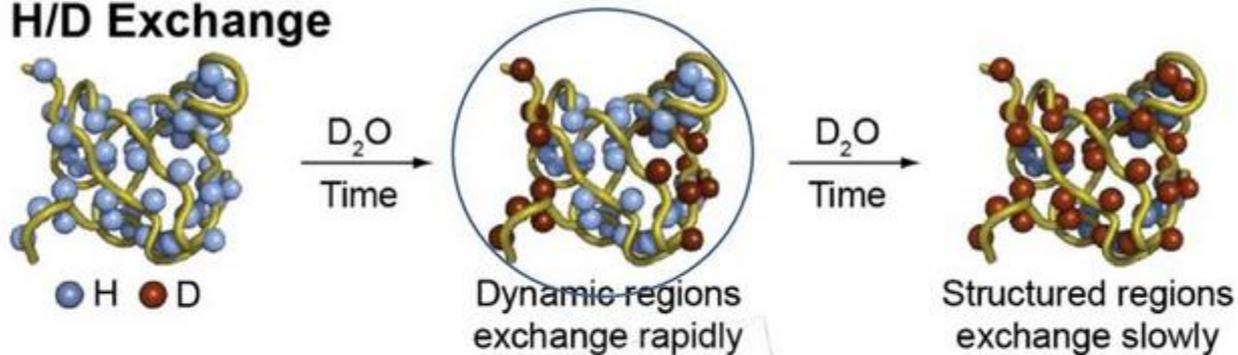
Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

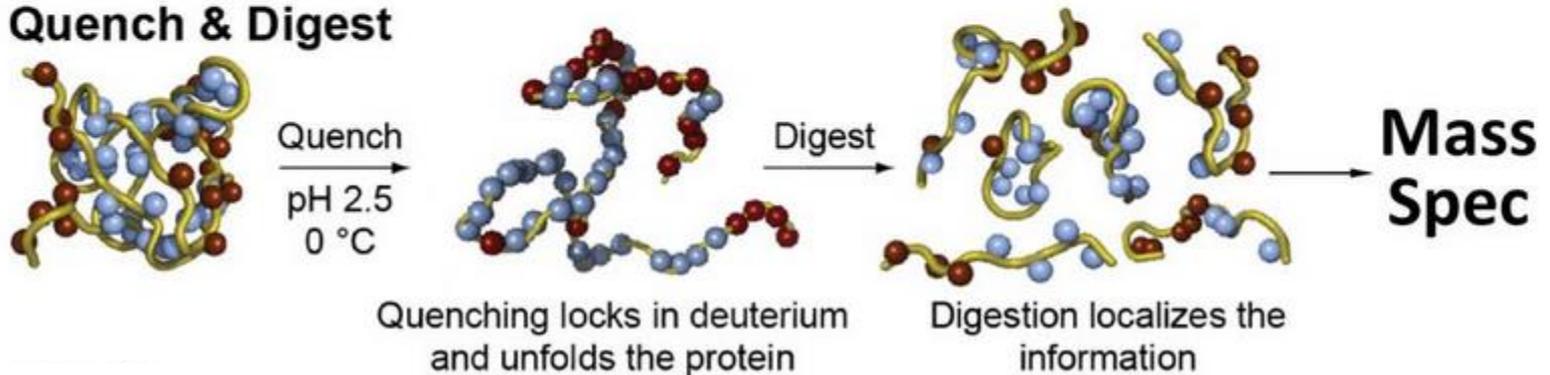
Metodi sperimentali: D) Definizione dei domain boundaries

2) Amidic Hydrogen/Deuterium Exchange (DX-MS)

H/D Exchange



Quench & Digest



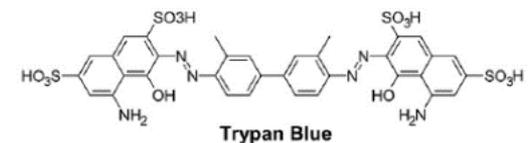
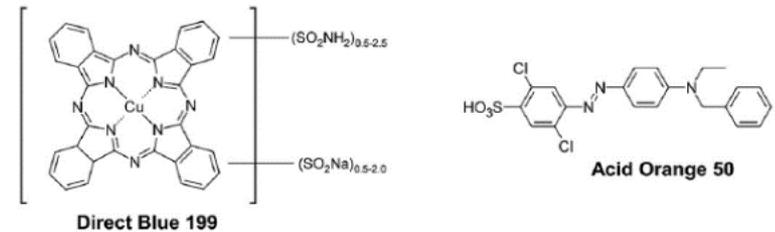
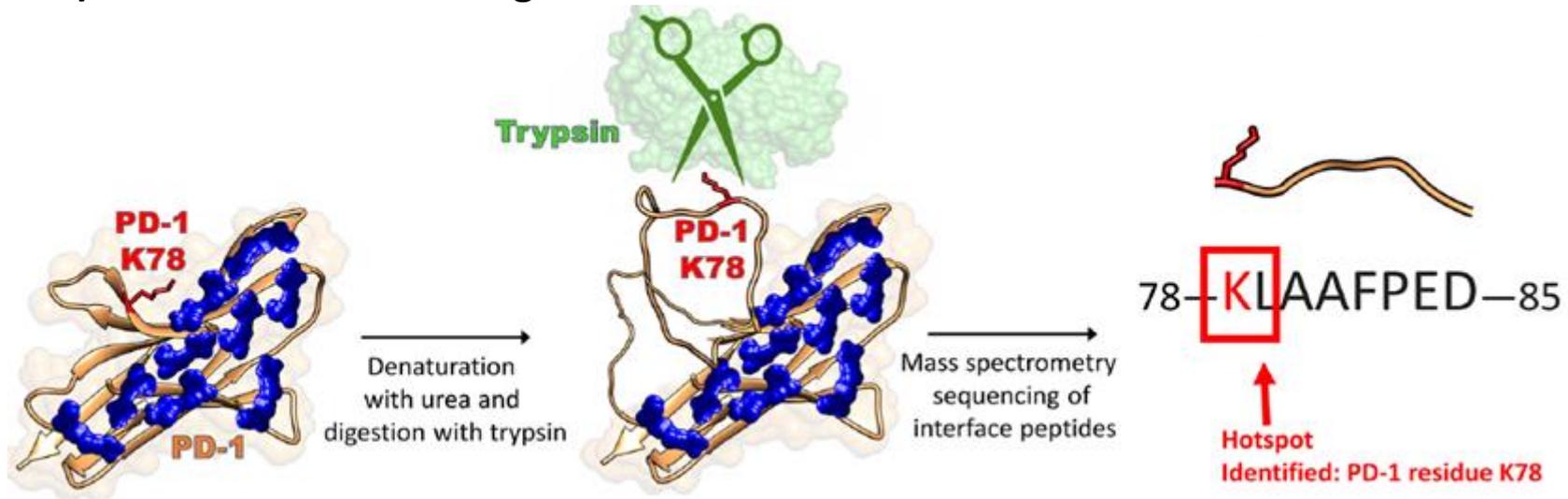
Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Metodi sperimentali: D)

Definizione dei domain boundaries

3) Protein Painting

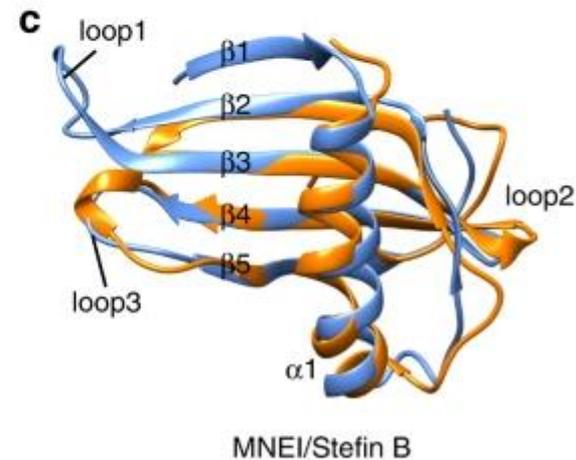
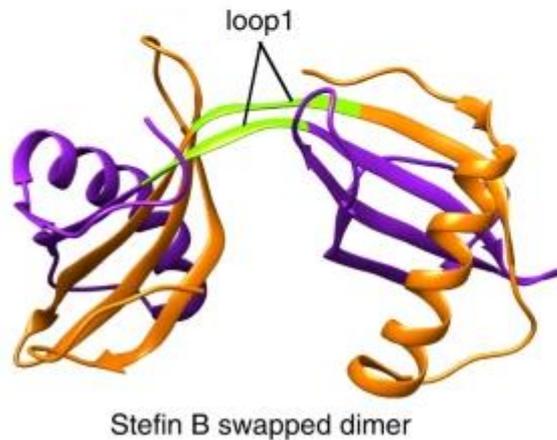
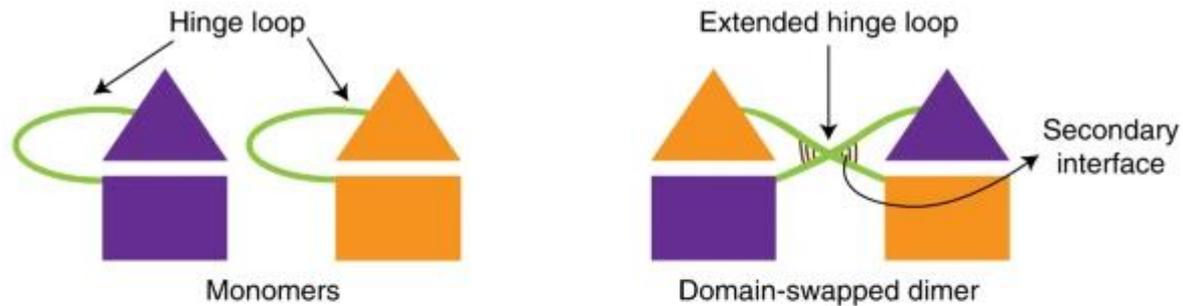


Espressione di proteine ricombinanti per studi strutturali

Design del costrutto

Espressione di singoli domini: CAVEAT

Domain swapping



Domain swapped dimer:

Folding non nativo

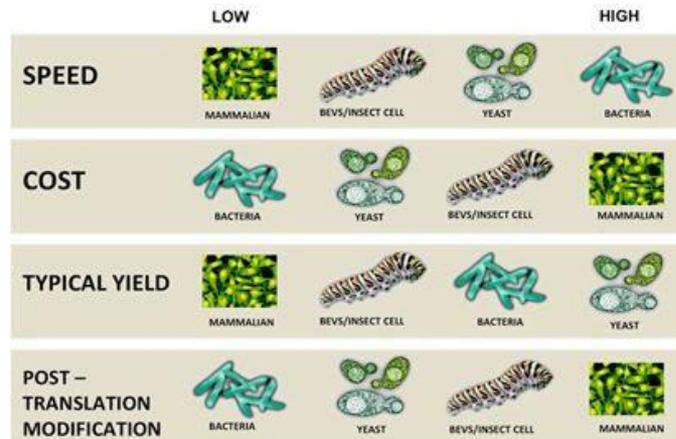
Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite

Expression system	Most common application	Advantages	Challenges
Mammalian 	<ul style="list-style-type: none">• Functional assays• Structural analysis• Antibody production• Expression of complex proteins• Protein interactions• Virus production	<ul style="list-style-type: none">• Highest-level protein processing• Can produce proteins either transiently, or by stable expression• Robust optimized transient systems for rapid, ultrahigh-yield protein production	<ul style="list-style-type: none">• Gram-per-liter yields only possible in suspension cultures• More demanding culture conditions
Insect 	<ul style="list-style-type: none">• Functional assays• Structural analysis• Expression of intracellular proteins• Expression of protein complexes• Virus production	<ul style="list-style-type: none">• Similar to mammalian protein processing• Can be used in static or suspension culture	<ul style="list-style-type: none">• More demanding culture conditions than prokaryotic systems• Production of recombinant baculovirus vectors is time consuming
Yeast 	<ul style="list-style-type: none">• Structural analysis• Antibody generation• Functional analysis• Protein interactions	<ul style="list-style-type: none">• Eukaryotic protein processing• Scalable up to fermentation (grams per liter)• Simple media requirements	<ul style="list-style-type: none">• Fermentation required for very high yields• Growth conditions may require optimization
Bacterial 	<ul style="list-style-type: none">• Structural analysis• Antibody generation• Functional assays• Protein interactions	<ul style="list-style-type: none">• Scalable• Low cost• Simple culture conditions	<ul style="list-style-type: none">• Protein solubility• May require protein-specific optimization• May be difficult to express some mammalian proteins
Cell-free 	<ul style="list-style-type: none">• Toxic proteins• Incorporation of unnatural label or amino acids• Functional assays• Protein interactions• Translational inhibitor screening	<ul style="list-style-type: none">• Open system; able to add unnatural components• Fast expression• Simple format	<ul style="list-style-type: none">• Scaling above multimilligram quantities may not be costly

Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite

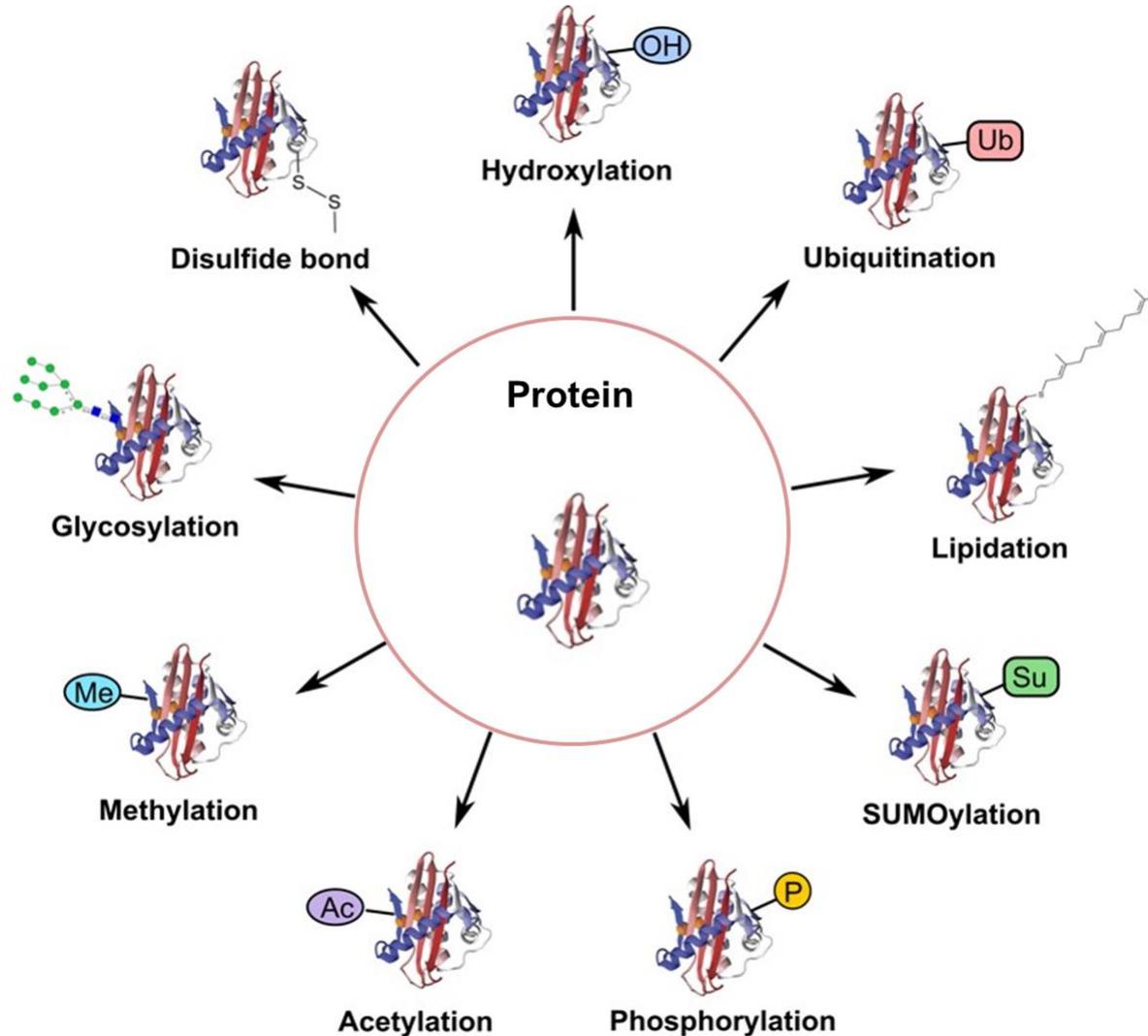


Characteristics	<i>Escherichia coli</i>	<i>Pichia pastoris</i>	CHO cell
Doubling time	30 min	60-120 min	24 hr
Cost of growth medium	Low	Low	High
Complexity of growth medium	Minimum	Minimum	Complex
Expression level	High	Low to high	Low to moderate
Extracellular expression	Secretion to periplasm	Secretion to medium	Secretion to medium
Protein folding	Refolding usually required	Refolding may be required	Proper folding
N-linked glycosylation	None	High mannose	Complex
O-linked glycosylation	No	Yes	Yes
Phosphorylation & acetylation	No	Yes	Yes
Drawback	Accumulation of LPS	Codon bias	Contamination with animal viruses

Espressione di proteine ricombinanti per studi strutturali

Scelta della cellula ospite:

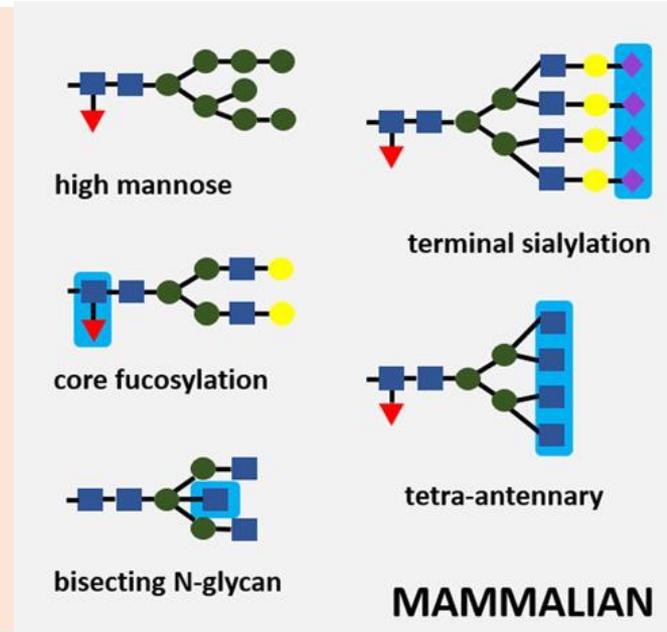
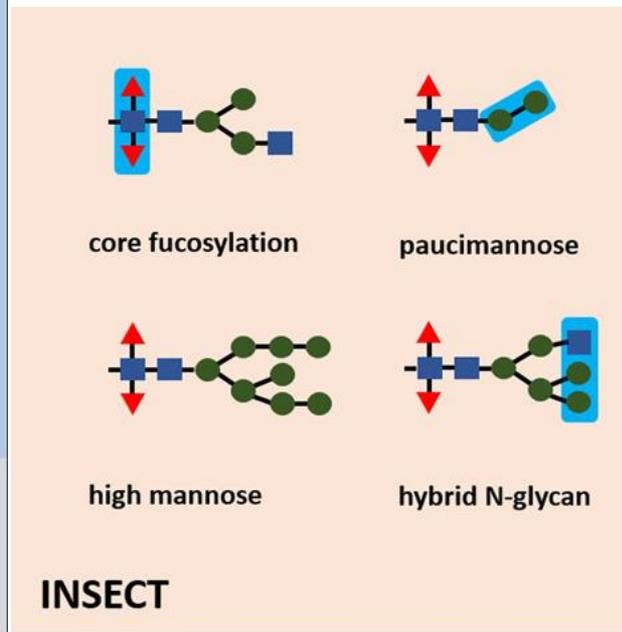
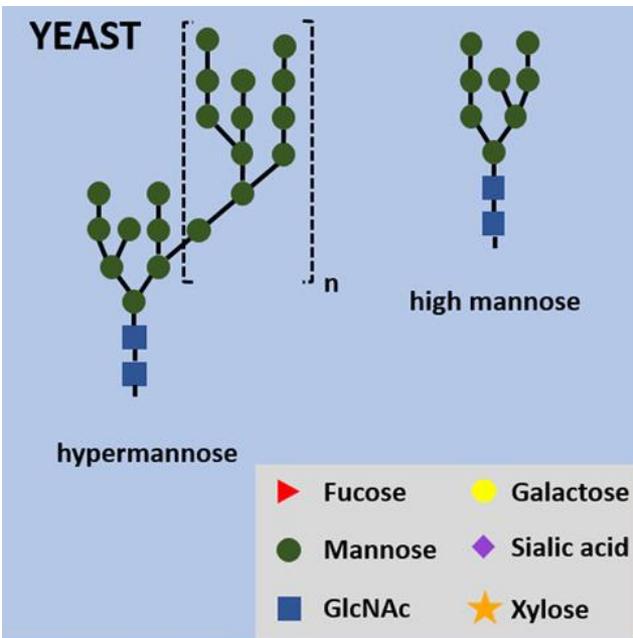
Modifiche post traduzionali



Espressione di proteine ricombinanti per studi strutturali

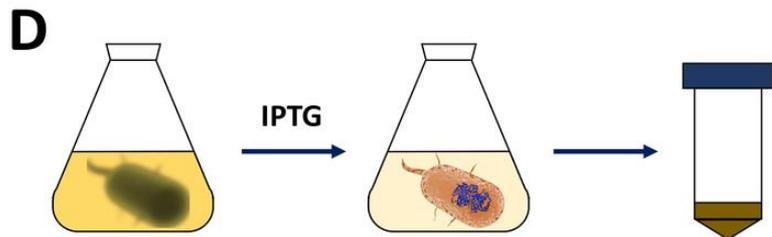
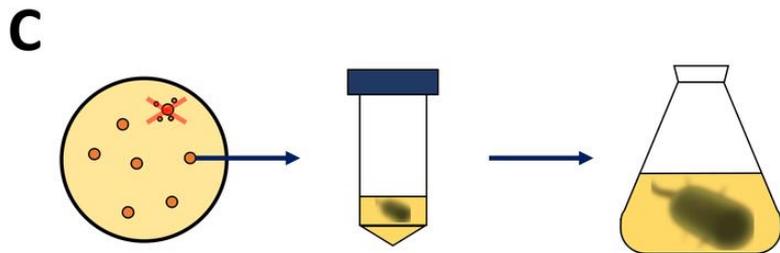
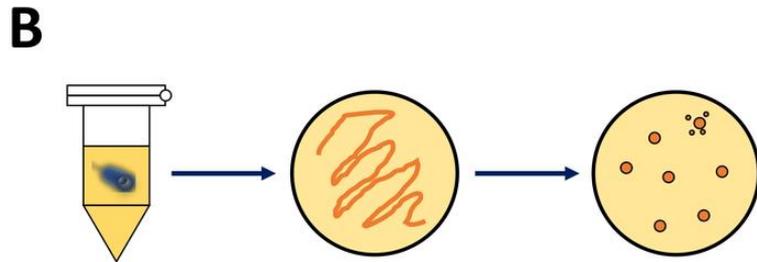
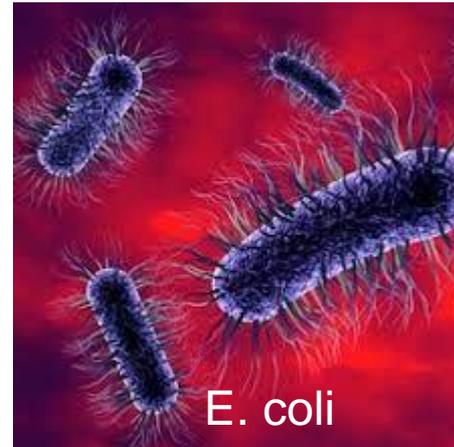
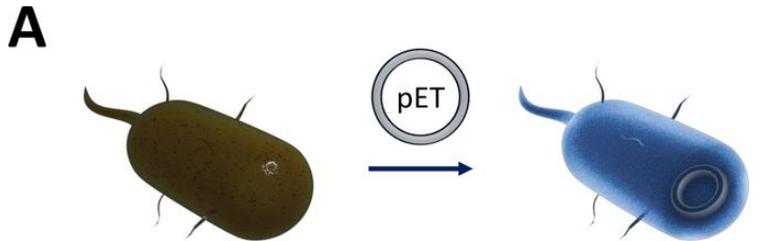
Scelta della cellula ospite:

Modifiche post traduzionali



Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite: Procarioti

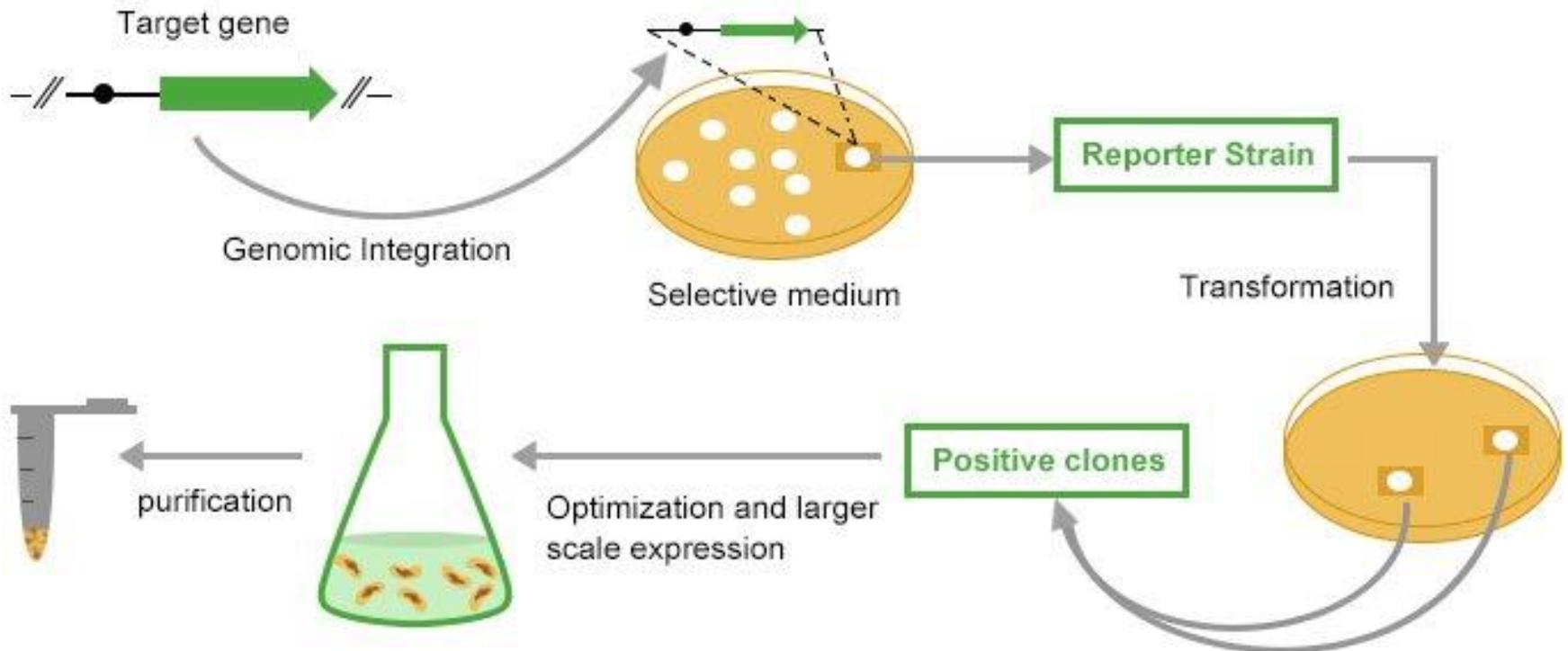
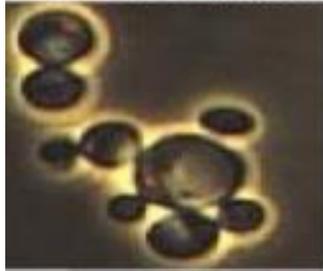


STRAINS	VECTORS
BL21(DE3) BL21(DE3) pLysS	pET plasmids T7 vector suite
TUNABILITY Lemo21/Tuner/AI	TUNABILITY pBAD series of vectors and derivatives
CODON BIAS CodonPlus/Rosetta	
TOXIC/MB PROTEINS C41/C43 (DE3)	CO-EXPRESSION pDuet vectors
DISULFIDE BONDS Origami/SHuffle CyDisCo	FOLDING ASSISTANCE Chaperone co-expression

Espressione di proteine ricombinanti per studi strutturali

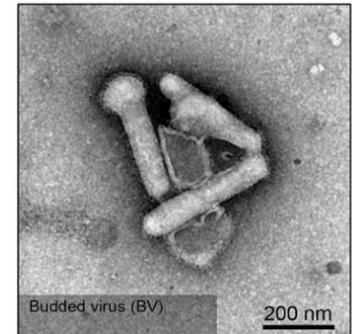
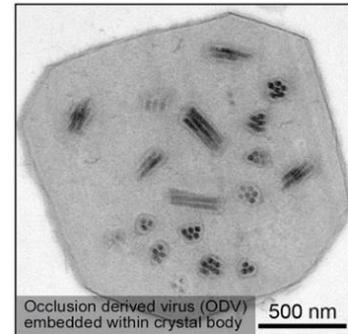
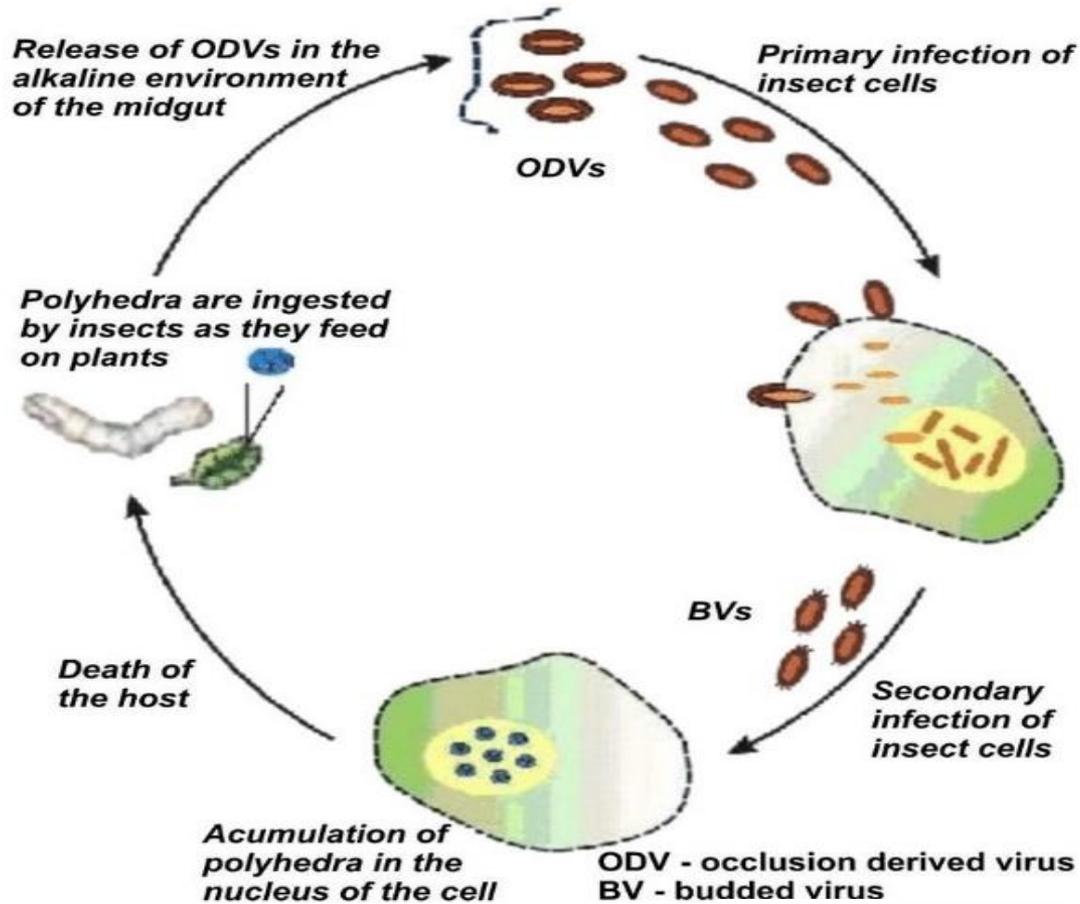
Scelta del sistema ospite: Lievito

Yeasts strains



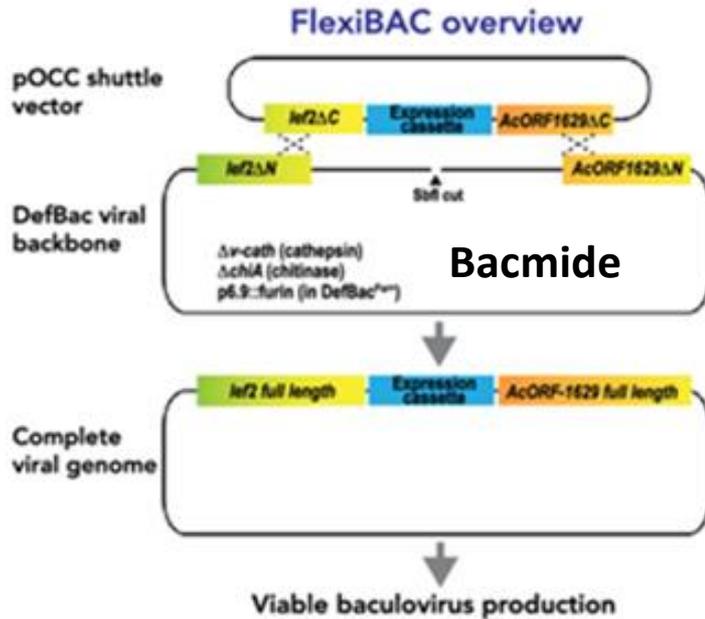
Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite: Baculovirus



Espressione di proteine ricombinanti per studi strutturali

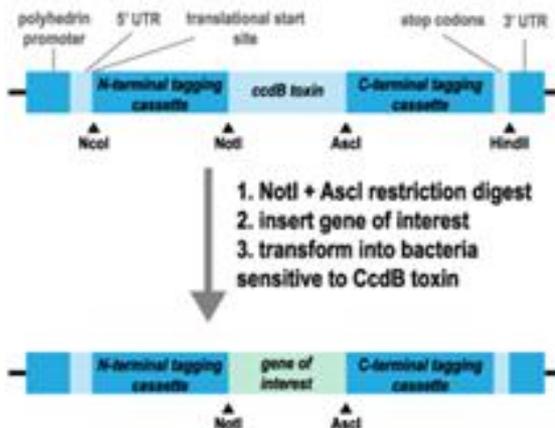
Scelta del sistema ospite: Baculovirus



Baculovirus Genome

- Deletion of *cathepsin*
- Deletion of *chitinase*
- Double deletion of *cathepsin* and *chitinase*
- Triple deletion of *p26*, *p10* and *p74*
- BmNPV *fp25k* replaced by AcMNPV *fp25k*

Expression cassette overview

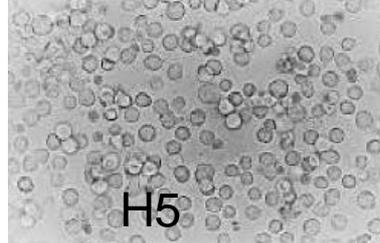
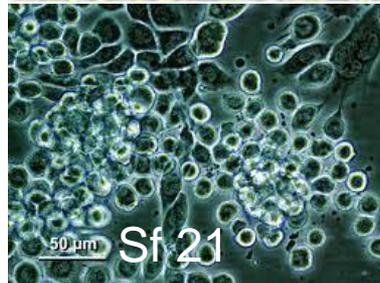
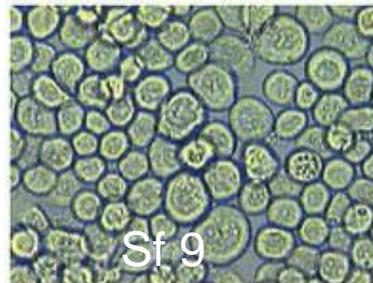


Signal Peptides

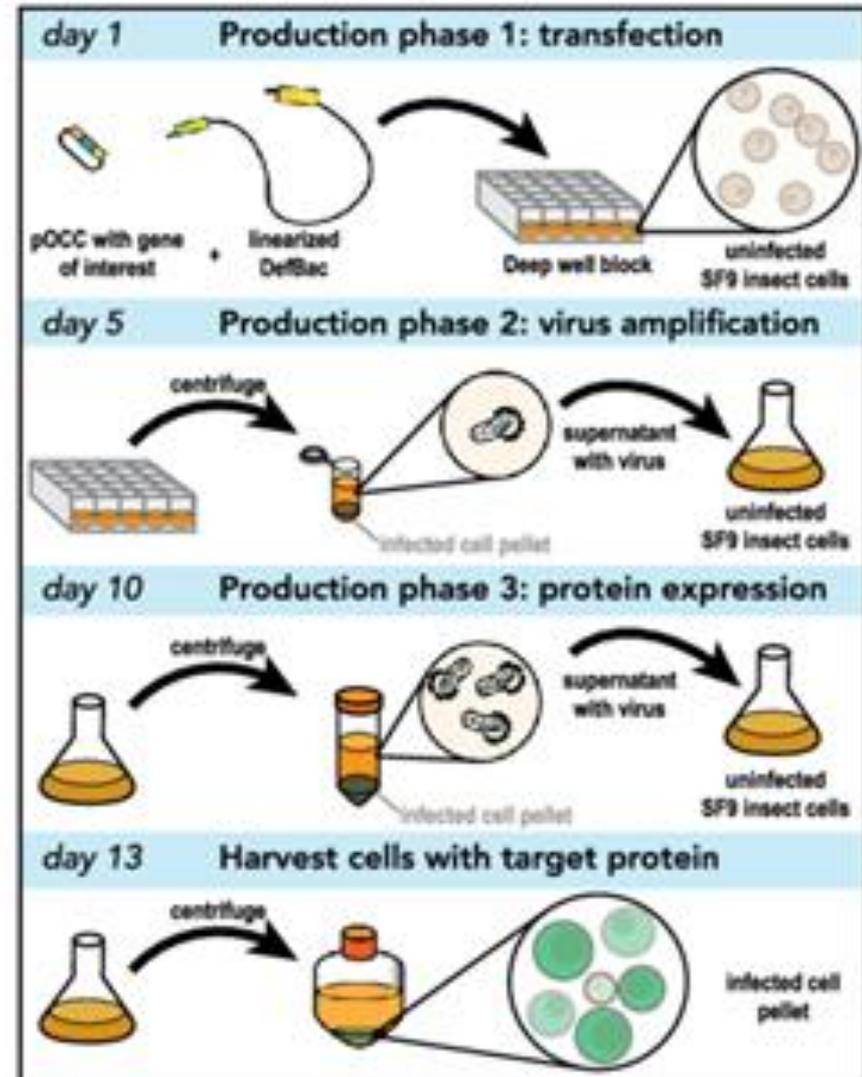
- honeybee melittin signal peptide
- cecropin B signal peptide
- Staphylococcal* protein A signal peptide
- human azurocidin signal peptide
- B.mori bombyxin(*bx*) signal peptide
- B.mori prophenoloxidase-activating enzyme(*ppae*) signal peptide

Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite: Baculovirus

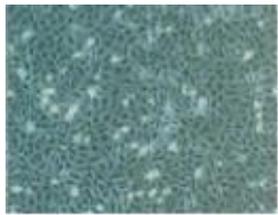


Protein expression timeline



Espressione di proteine ricombinanti per studi strutturali

Scelta del sistema ospite: mammalian cells



Mammalian cells



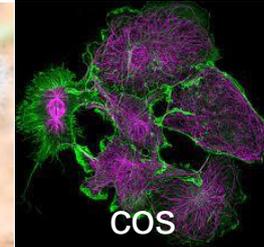
Chinese hamster



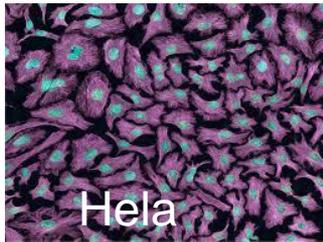
CHO



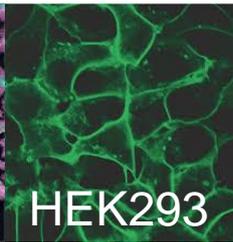
Monkey



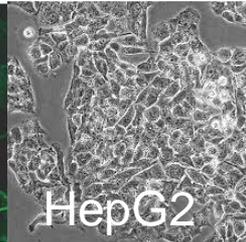
COS



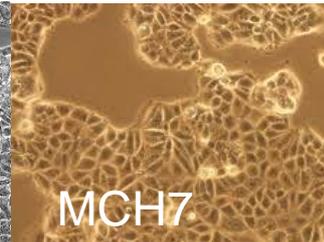
HeLa



HEK293

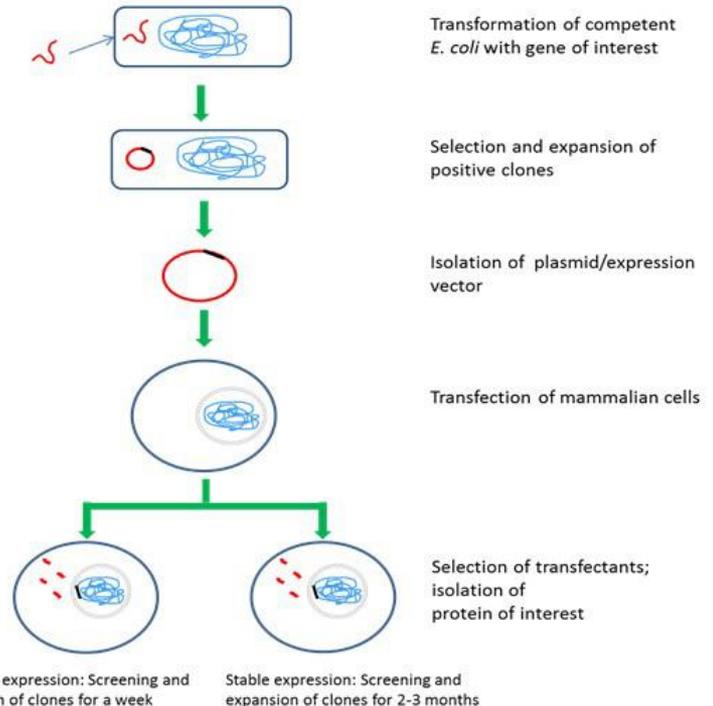


HepG2



MGH7

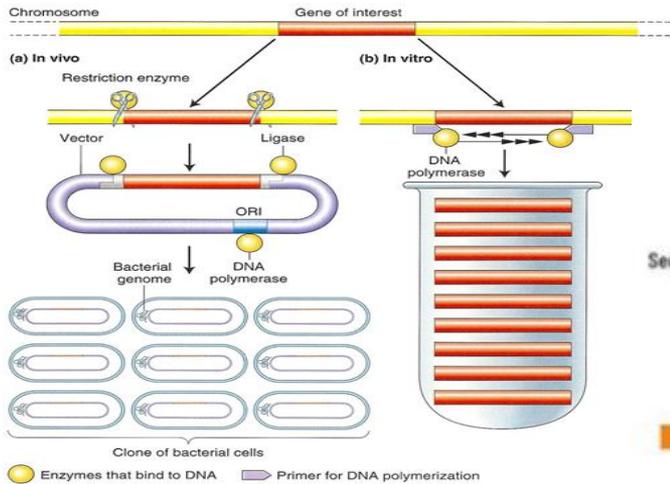
HEK293	Human embryonic kidney (high transfection efficiency)
HeLa	Human cervical carcinoma (historical, low RNase)
CHO	Chinese hamster ovary (hardy, diploid DNA content, mutants)
Cos	Monkey cells with SV40 replication proteins (-> high transgene copies)
3T3	Mouse or human exhibiting ~regulated (normal-like) growth
+ various others, many differentiated to different degrees, e.g.:	
BHK	Baby hamster kidney
HepG2	Human hepatoma
GH3	Rat pituitary cells
PC12	Mouse neuronal-like tumor cells
MCF7	Human breast cancer
HT1080	Human with near diploid karyotype
IPS	induced pluripotent stem cells
and:	
<u>Primary</u> cells cultured with a limited lifetime (frozen stocks available)	
E.g.,	MEF = mouse embryonic fibroblasts, HDF = Human diploid fibroblasts



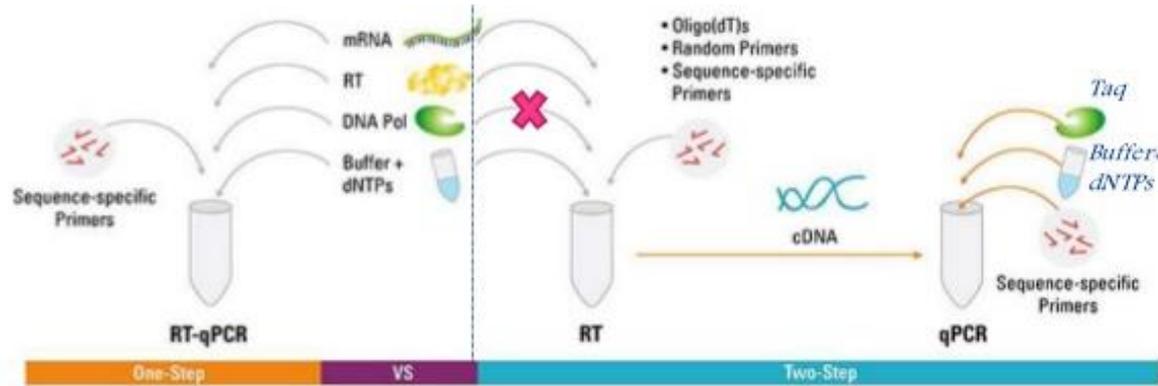
Espressione di proteine ricombinanti per studi strutturali

cDNA isolation:

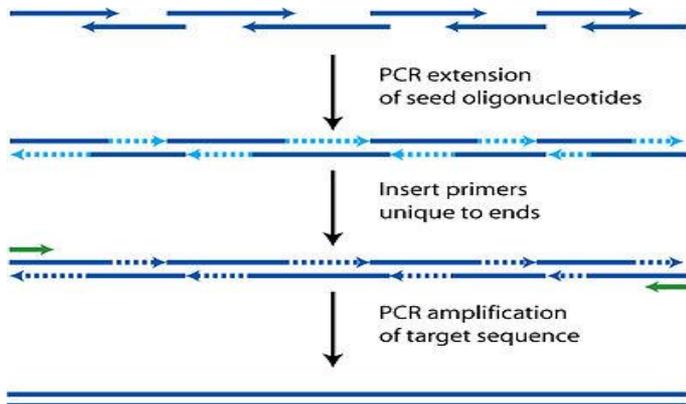
genomic fishing



fishing from mRNA



strategia ibrida sintesi/PCR (overlap assembly PCR)



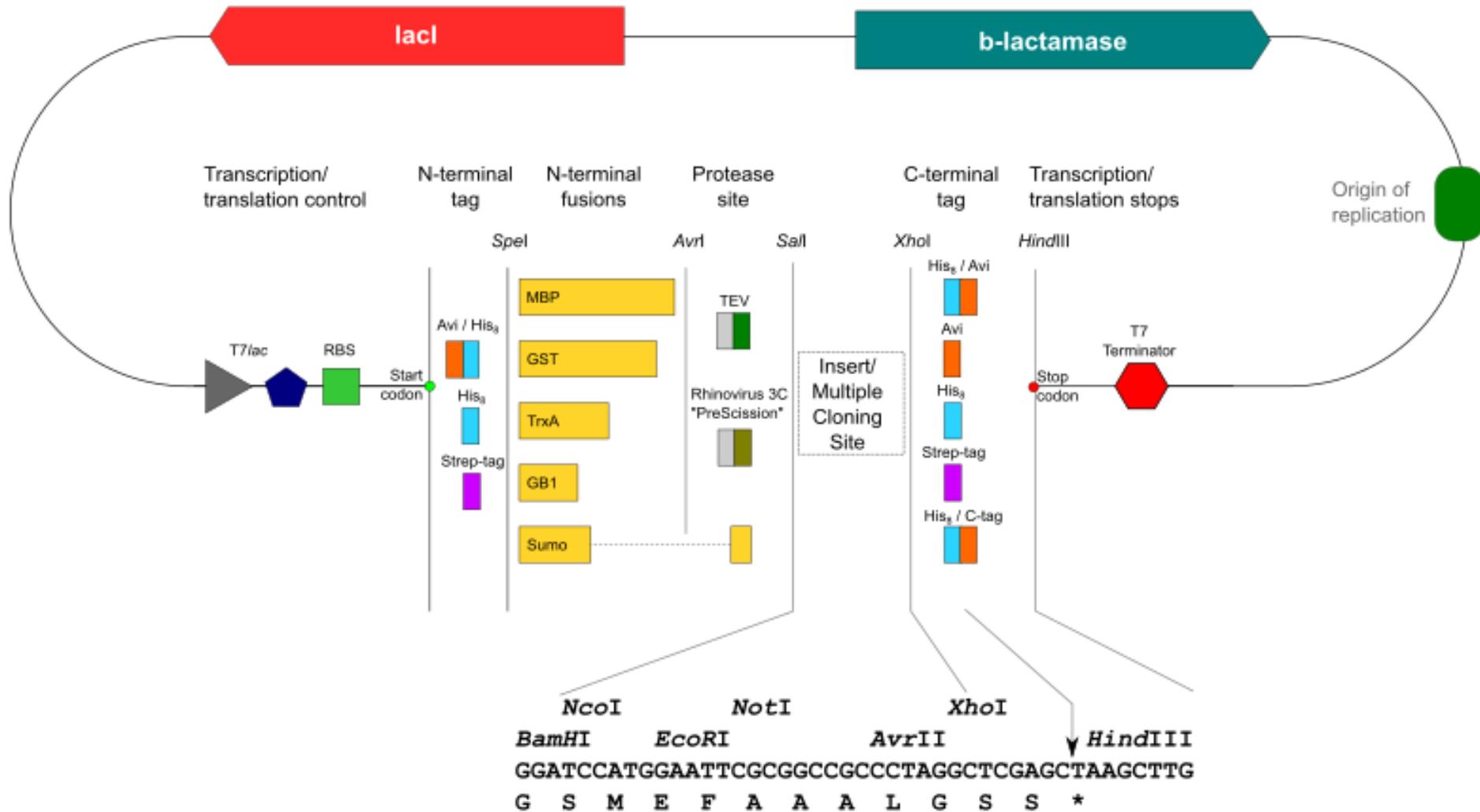
sintesi genica



Espressione di proteine ricombinanti per studi strutturali

Sceita del vettore :

elementi principali

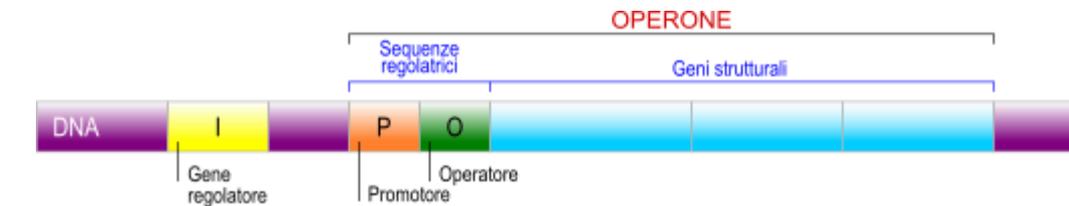


Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore : PROMOTORE

- Ospite
- Forza
- Costitutivo/inducibile
- Leakiness
- Lag time

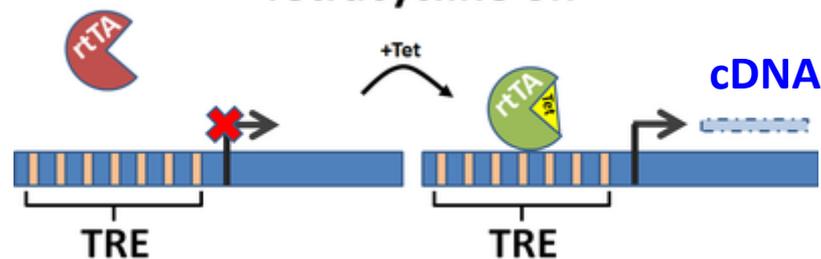
In base a



Positive inducible



Tetracycline on



Espressione di proteine ricombinanti per studi strutturali

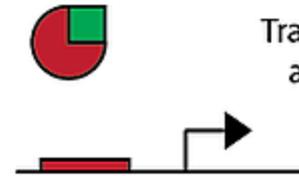
Scelta del vettore : PROMOTORE

Negative inducible

repressor

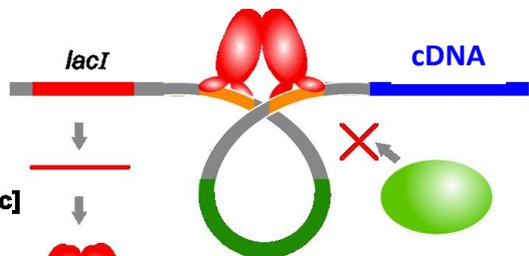


Inducer addition

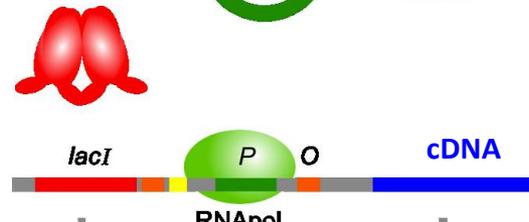


Transcription activated

pLac promoter

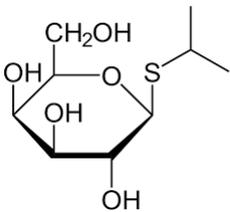


low [Lac]

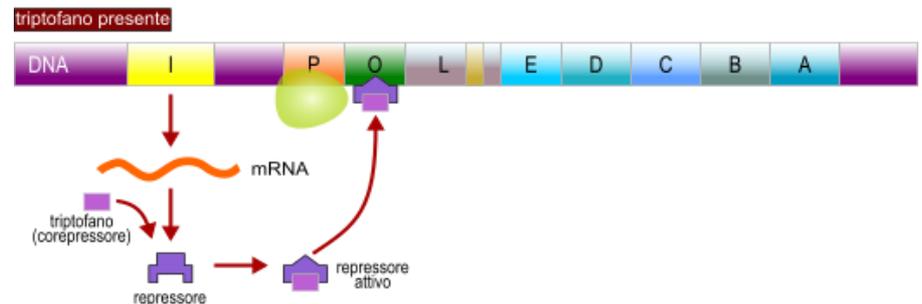
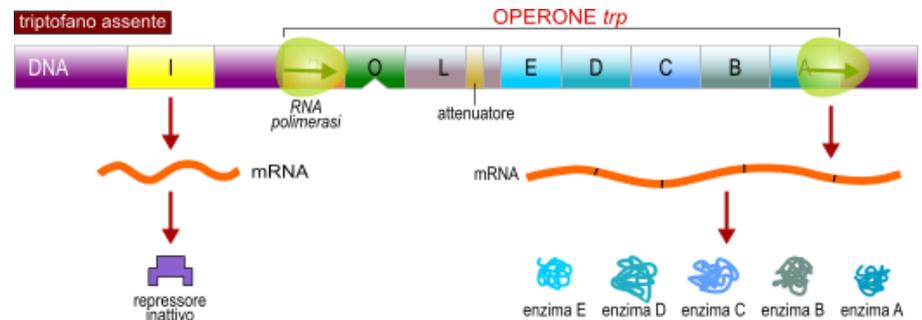


high [Lac]

protein



ptrp promoter



Espressione di proteine ricombinanti per studi strutturali

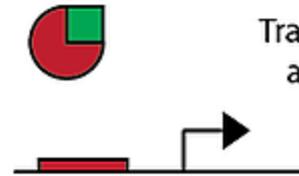
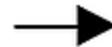
Scelta del vettore : PROMOTORE

Negative inducible

repressor

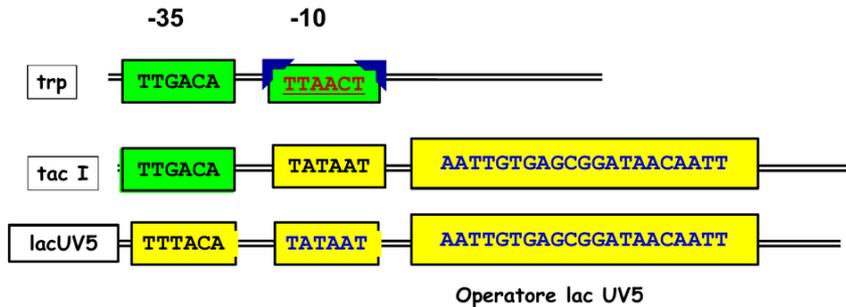


Inducer addition



Transcription activated

ptac promoter



INDUZIONE: IPTG

RESA: MEDIO-ALTA (5x rispetto ai parentali)

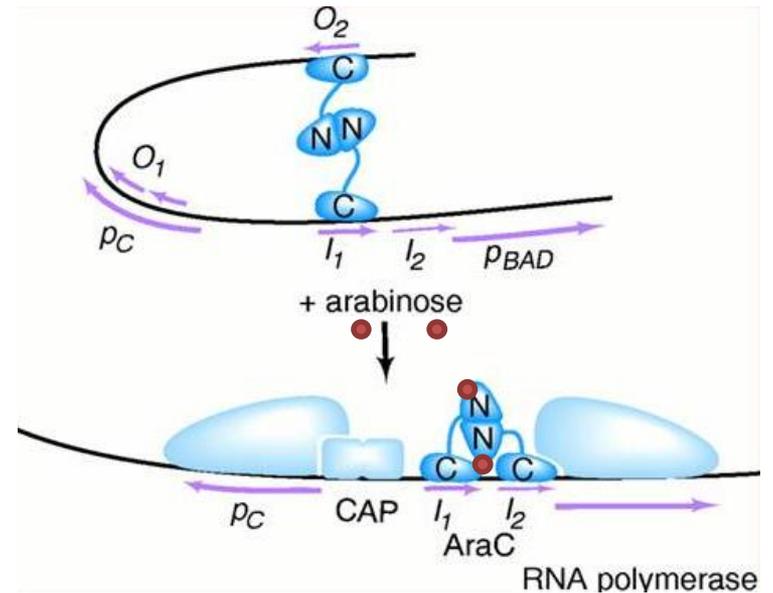
REGOLAZIONE: POSSIBILE (LAC I)

ESPRESSIONE BASALE: ALTA

Ottiene una buona resa anche se inferiore a quella del T7

La sua forza può essere un problema con prodotti tossici e proteine di membrana

pBAD promoter



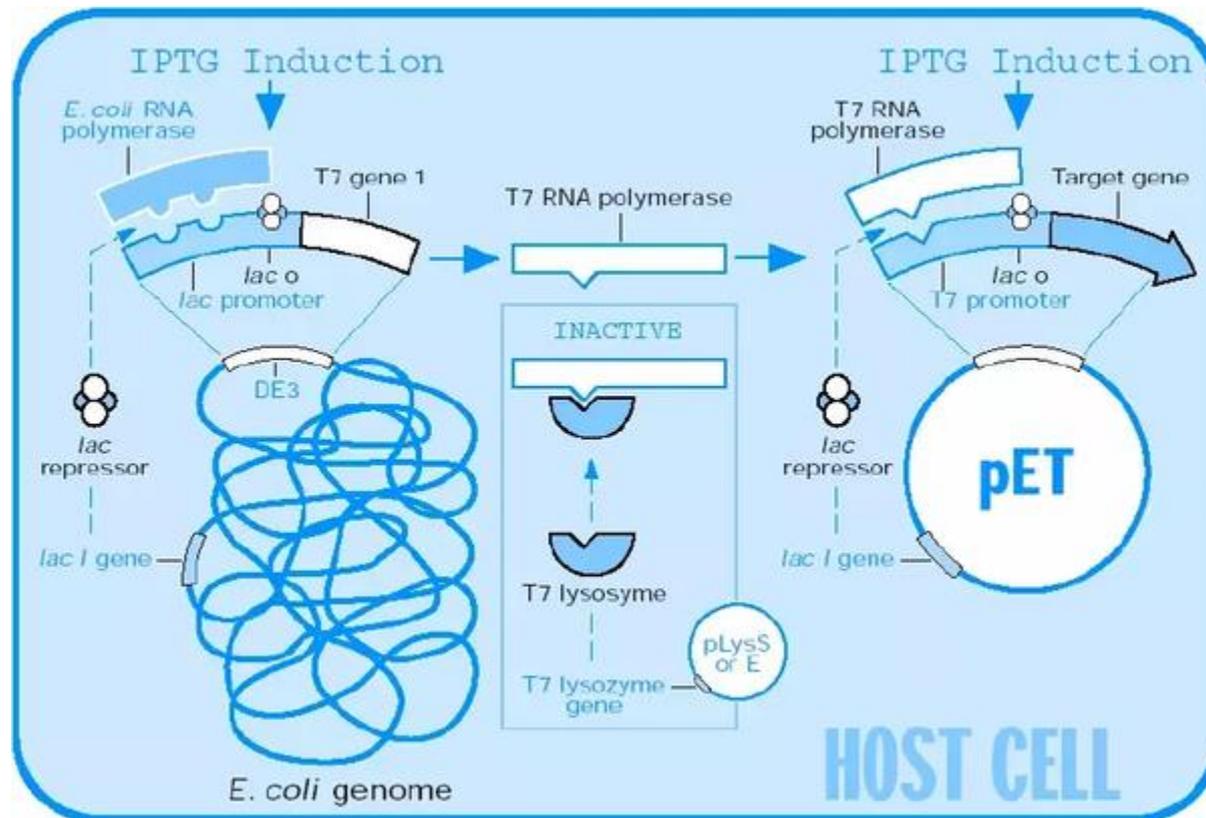
Espressione di proteine ricombinanti per studi strutturali

promotori batterici inducibili

Promoter	Description	Expression	Additional considerations
T7	Promoter from T7 bacteriophage	Constitutive, but requires T7 RNA polymerase.	
T7lac	Promoter from T7 bacteriophage plus lac operators	Negligible basal expression. Requires T7 RNA polymerase, which is also controlled by lac operator.	Very tightly regulated by the lac operators, tunable.
Sp6	Promoter from Sp6 bacteriophage	Constitutive, but requires SP6 RNA polymerase.	SP6 polymerase has a high processivity.
araBAD	Promoter of the arabinose metabolic operon	Inducible by arabinose repression in the presence of glucose or by competitive binding of the anti-inducer fucose	Weaker. Low basal expression; not tunable
trp	Promoter from E. coli tryptophan operon	Repressible with high levels of cellular trp.	
lac	Promoter from lac operon	Constitutive in the absence of lac repressor (lacI or lacIq). Can be induced by IPTG or lactose.	Leaky promoter weak expression tunable
Ptac	Hybrid promoter of lac and trp	Regulated like the lac promoter	Contains -35 region from trpB and -10 region from lac. Very tight regulation. tunable
pL	Promoter from bacteriophage lambda	temperature regulable	Often paired with the temperature sensitive cl857 repressor.
T3	Promoter from T3 bacteriophage	Constitutive requires T3 RNA polymerase	

Espressione di proteine ricombinanti per studi strutturali sistemi batterici

DE3 system



**Espressione inducibile
Autoinduzione**

Espressione di proteine ricombinanti per studi strutturali

Autoinduzione

In alcuni media (mix glucosio e lattosio) è possibile indurre l'espressione quando i batteri raggiungono la saturazione.



Available online at www.sciencedirect.com



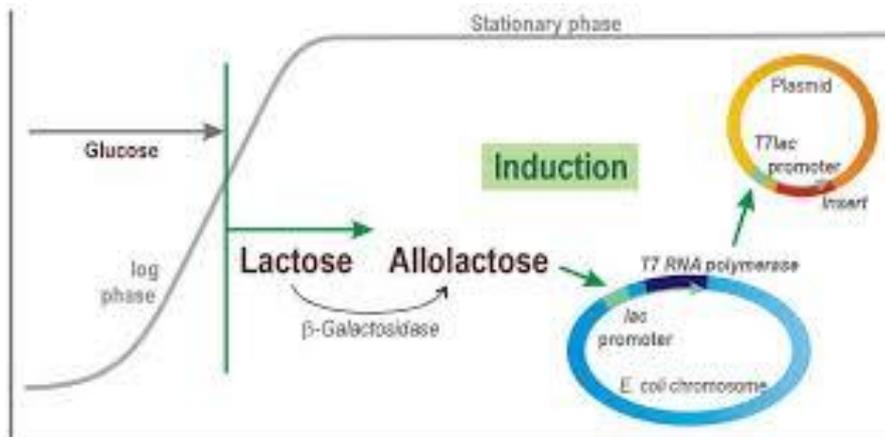
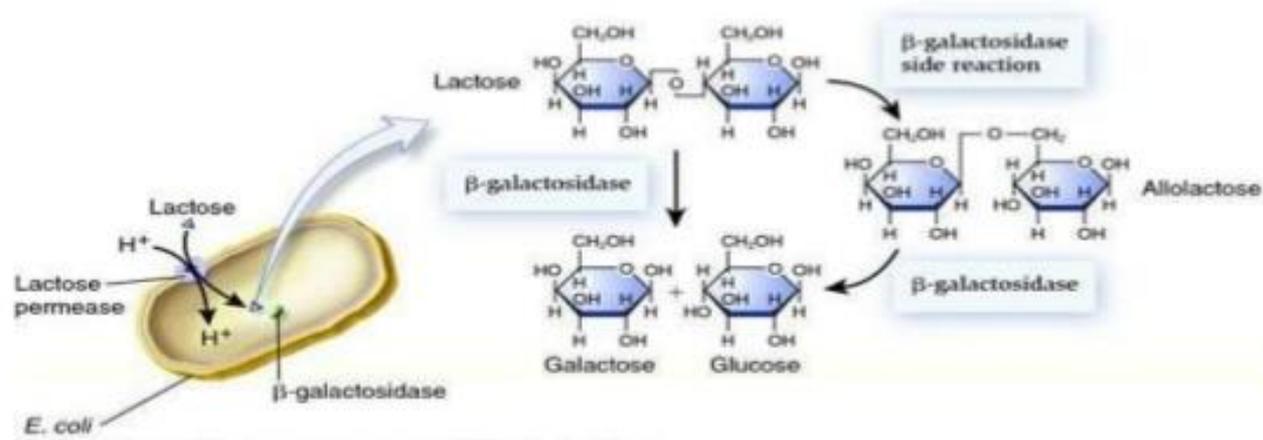
Protein Expression and Purification 41 (2005) 207–234

Protein
Expression
& Purification

www.elsevier.com/locate/yprep

Protein production by auto-induction in high-density shaking cultures

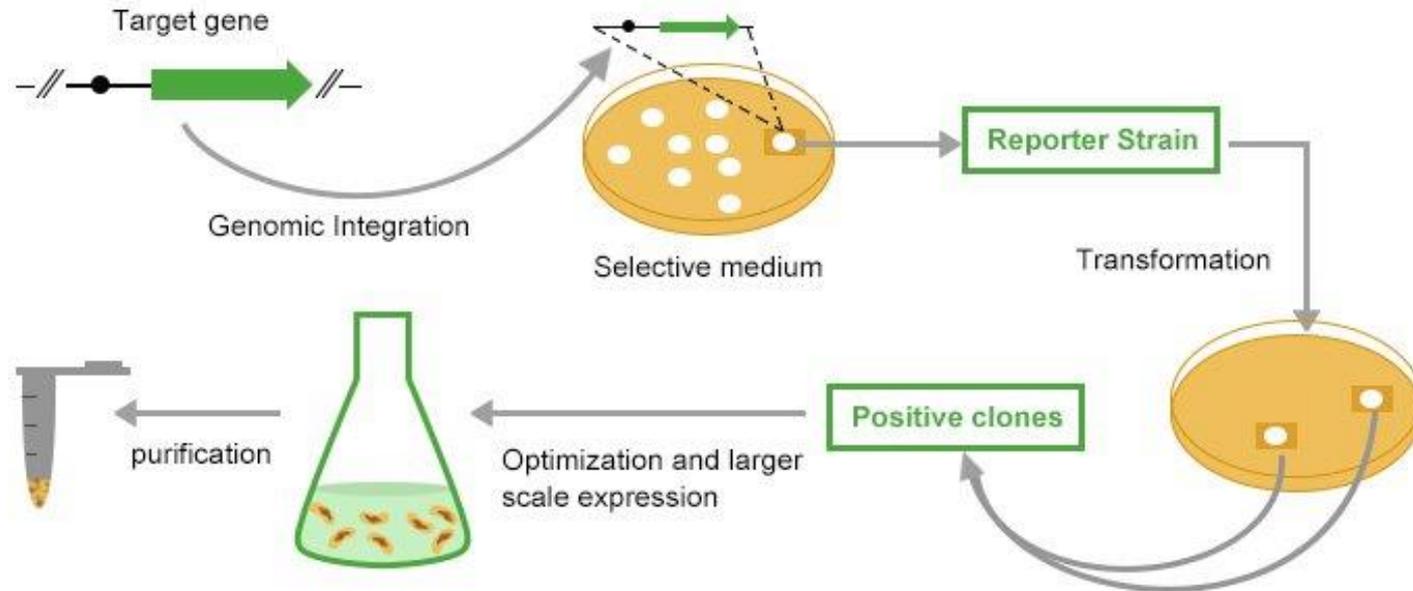
F. William Studier*



Si possono raggiungere da 2 a 3X OD600nm

Espressione di proteine ricombinanti per studi strutturali

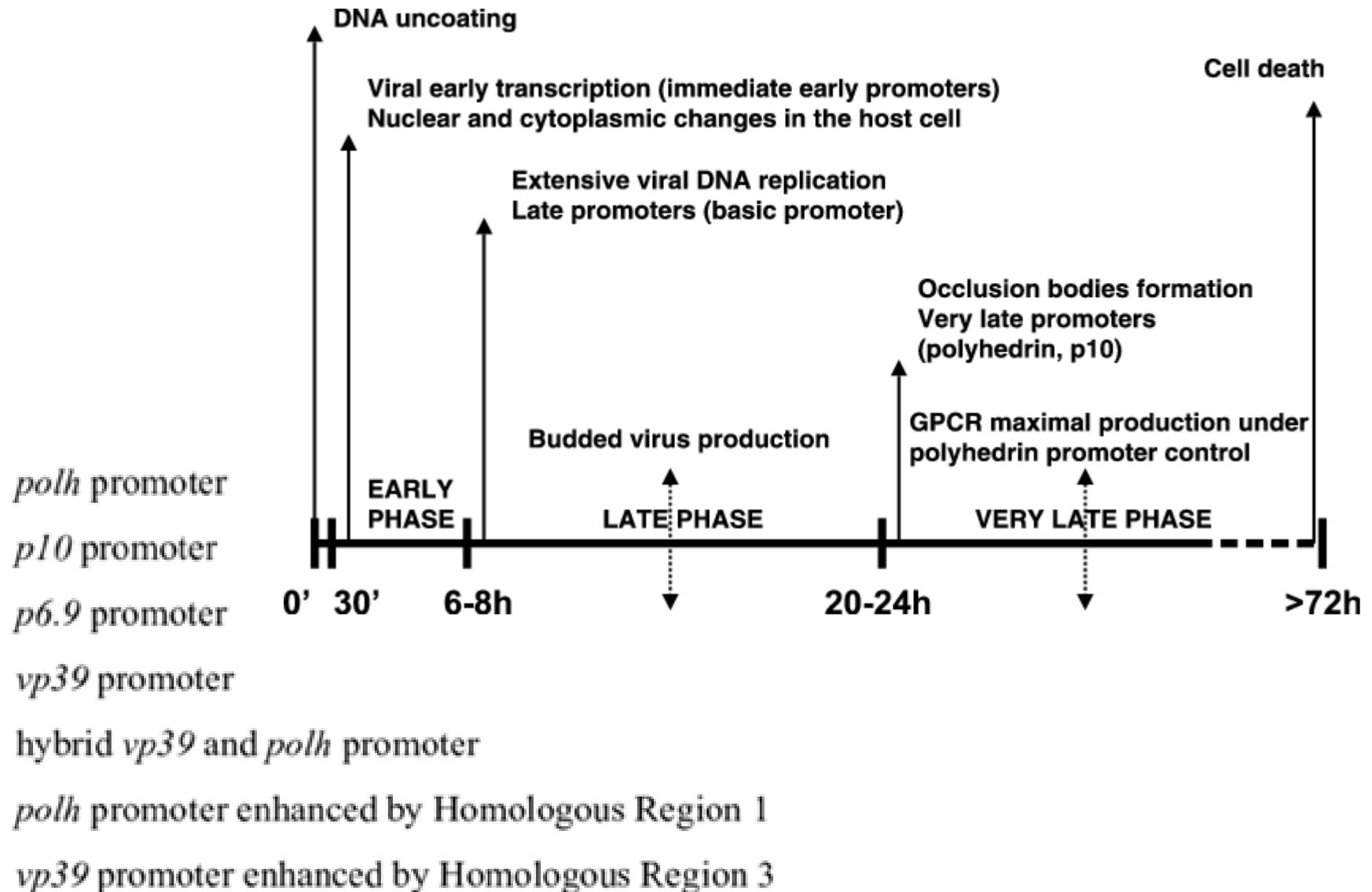
promotori di Lievito



Promoter	Description	Expression	Additional considerations
TEF1	from transcription elongation factor	Constitutive	Analogous to mammalian EF1a promoter.
GDS	Strong from glyceraldehyde 3-phosphate dehydrogenase	Constitutive	Very strong
GAL1, 10	adjacent, divergently transcribed promoters	Inducible with galactose; repressible with glucose	Used independently or together. Regulated by GAL4, 80.
ADH1	From alcohol dehydrogenase I	Repressed by ethanol	Full length version is strong. Truncated constitutive with lower expression.

Espressione di proteine ricombinanti per studi strutturali

promotori di Baculovirus

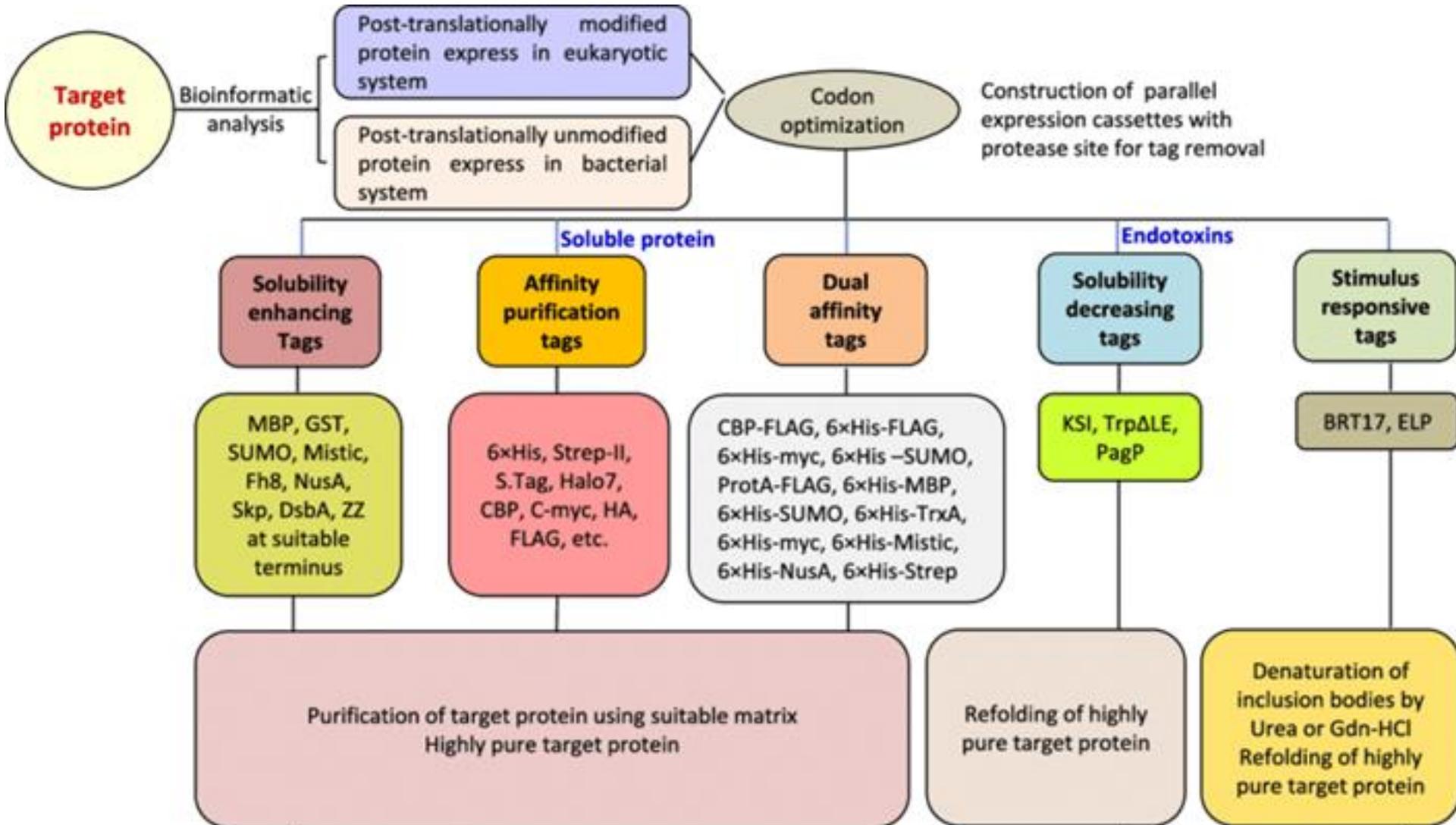


Espressione di proteine ricombinanti per studi strutturali promotori di mammalian cells

Promoter	origin	Expression	comments
CMV	Strong from the human cytomegalovirus	Constitutive	May contain an enhancer. Silenced in some cell types.
SV40	from the simian vacuolating virus 40	Constitutive	May include an enhancer.
EF1a	Strong from human elongation factor 1 α	Constitutive	High expression
PGK1 (human or mouse)	from phosphoglycerate kinase gene.	Constitutive	vary by cell type. Resist downregulation due to methylation or deacetylation.
Ubc	from the human ubiquitin C gene	Constitutive	Ubiquitous
human beta actin	from beta actin gene	Constitutive	Ubiquitous
TRE	Tetracycline response element promoter	Inducible	minimal promoter with low basal activity and several tetracycline operators.

Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: TAGs & cleavage sites



Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: TAGs

Tag	Protein source	Residue size (aa/-kDa)	Tag placement	Source organism	Ligand/Matrix	Purification condition	Suitable expression system	Advantages	Limitations
Solubility enhancing tags									
MBP	Maltose binding protein	396/43	N- or C-terminus	<i>E. coli</i>	Cross linked amylose	10 mM Maltose	<i>E. coli</i>	Enhanced solubility of target, do not leave extraneous amino acids	Can decrease expression yield
GST	Glutathione-S-transferase	211/26	N- or C-terminus	<i>Schistosoma japonicum</i>	Glutathione	10- 20 mM reduced Glutathione	Bacteria, Yeast, insect, mammalian cells	Enhanced solubility of target, Highly specific binding	Can leach from column
SUMO	Small ubiquitin modifier	100/18	N-terminus	<i>H. sapiens</i>	NA	N/A	Bacteria, Yeast, Insect, Mammalian cells	Enhanced target solubility, unique tag removal method	Do not aid in purification
TrxA	Thioredoxin A	109/12	N- or C-terminus	<i>E. coli</i>	Phenylarsinine oxide	5-1000 mM β -mercapto-ethanol	<i>E. coli</i>	Helps in crystallization, Aid in refolding of target protein	Do not aid in purification
NusA	N-utilization substance A	495/55	N- or C-terminus	<i>E. coli</i>	N/A	N/A	Bacteria	Enhanced solubility	Large size decreases yield
Mistic	Mistic tag	110/13	N- or C-terminus	<i>Bacillus subtilis</i>	NA	N/A	Bacteria	Commendable tag for membrane proteins	Do not aid in purification and crystallization
HaloTag7	Haloalkane dehalogenase	~300/34	N- or C-terminus	<i>Rhodococcus/ engineered</i>	Chloroalkane coupled agarose Halolink™ resin	In-column proteolytic digestion	Insect cells, Mammalian cells	Increased expression, solubility, purification yield and high purity, in-vivo imaging and multiplexing	Application dependent ligand requirement; Functional inactivation of target protein
Fh8	Ca ²⁺ binding recombinant protein	69/8.0	N-terminus	<i>Faciola hepatica</i>	Phenyl sepharose	N/A	<i>E. coli</i>	Enhanced expression & solubility, aid in purification on HIC	
DsbA	Protein disulphide isomerase A	209/-21	N-terminus	<i>E. coli</i>	NA	N/A	<i>E. coli</i>	Enhanced solubility, native disulfide bond formation in target	
Skp	Small <i>E. coli</i> chaperone	162/17.9	N-terminus	<i>E. coli</i>	N/A	N/A	<i>E. coli</i>	Enhanced solubility	Do not aid in purification
GB1	IgG domain B1 of Streptococcus Protein G	56/7.5	N-terminus	<i>Streptococcus aureus</i>	N/A	N/A	Bacteria	Enhanced refolding	
ZZ domain	IgG binding ZZ domain of protein A	/31	N- or C-terminus	<i>Staphylococcus aureus</i>	N/A	N/A	Bacteria, Yeast	Enhanced refolding	

Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: TAGs

Tag	Protein source	Residue size (aa/ -kDa)	Tag placement	Source organism	Ligand/Matrix	Purification condition	Suitable expression system	Advantages	Limitations
Stimulus responsive tags									
BRT17	β -roll repeat of GGAGNDTLY	153/14.7	N-terminus	Synthetic	N/A	Precipitation in presence of 25 -75 mM Ca ⁺⁺	Bacteria	Selective purification	Refolding of precipitated target
ELP	Elastin-like peptides	550 (110 repeats)/ -25	N- or C-terminus	tandem repeats of VPGXG sequence	Reverse phase transition	Precipitation by temperature change and/ or > 1.5 M NaCl	Bacteria	Cost effective than conventional purification	Refolding of precipitated target
Solubility decreasing (inclusion body targeting) tags									
KSI	Ketosteroid isomerase	125/13	N-terminus	<i>Pseudomonas testosterone</i>	N/A	N/A	<i>E. coli</i>	In column purification, refolding and sorting	Tag removal uses CNBr
Trp Δ LE	<i>Trp L</i> operon	17/27	N- or C-terminus	<i>E. coli</i>	N/A	N/A	<i>E. coli</i>	High expression of toxic and membranes proteins	Limited Refolding efficiency
PagP	β -barrel outer membrane protein			<i>Pseudomonas aeruginosa</i>	N/A	N/A	Bacteria	Easier inclusion body solubilisation of toxic target, commercialized	Refolding efficiency

Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: TAGs

Tag	Protein source	Residue size (aa/ -kDa)	Tag placement	Source organism	Ligand/Matrix	Purification condition	Suitable expression system	Advantages	Limitations
Affinity Purification tags									
Poly Arg	N/A	5-6/0.8	C terminus	Synthetic	Cation exchange resin	Alkaline NaCl	Bacteria	General cation exchange resin for purification	Cryptic cleavage, altered tertiary structure of target
Poly His	N/A	6-10/0.84-1.4	N- or C terminus	Synthetic	Ni ²⁺ -NTA agarose	20-250 mM Imidazole/low pH	Bacteria	Minimal interference with target protein, well defined protocol and availability	Co-purification of contaminants
Strep II Tag	Bacterial peptide	8/1.06	N- or C terminus	Bacteria	Engineered streptavidin (Strep-tactin)	2-25 mM Desthiobiotin	Bacteria, Yeast, Plants, Mammalian, Insect cells	Immunodetection, specificity and versatility	
S•Tag	Bovine pancreatic RNase A	15/1.75	N-/C- or internal	Bovine	S•protein (21-124 aa of)	3 M GITC; 0.2 M Potassium citrate	Bacteria	Unlikely interference with protein folding	Compromised purification
Calmodulin Binding Protein	Skeletal muscle myosin light chain kinase	26/2.5	N- or C terminus	<i>H. sapiens</i>	Calmodulin	EGTA and >1.5 M NaCl	Bacteria	High specificity, purification yield, native elution	Unsuitable for eukaryotic systems
T ₇ Tag		11/1.2	N-/C- or internal	Bacteriophage T ₇	Anti-T ₇ epitope monoclonal antibody	Low pH	Bacteria	Increased expression	Inclusion body formation, refolding efficiency of purified target is compromised
Chitin binding Protein	Chitinase A	51/6.0	N- or C-terminus	<i>Bacillus circulans</i>	Chitin	β-mercaptoethanol or DTT	Bacteria	High purity	Used with auto-cleavable intein tag, target, Target lose first methionine after tag removal

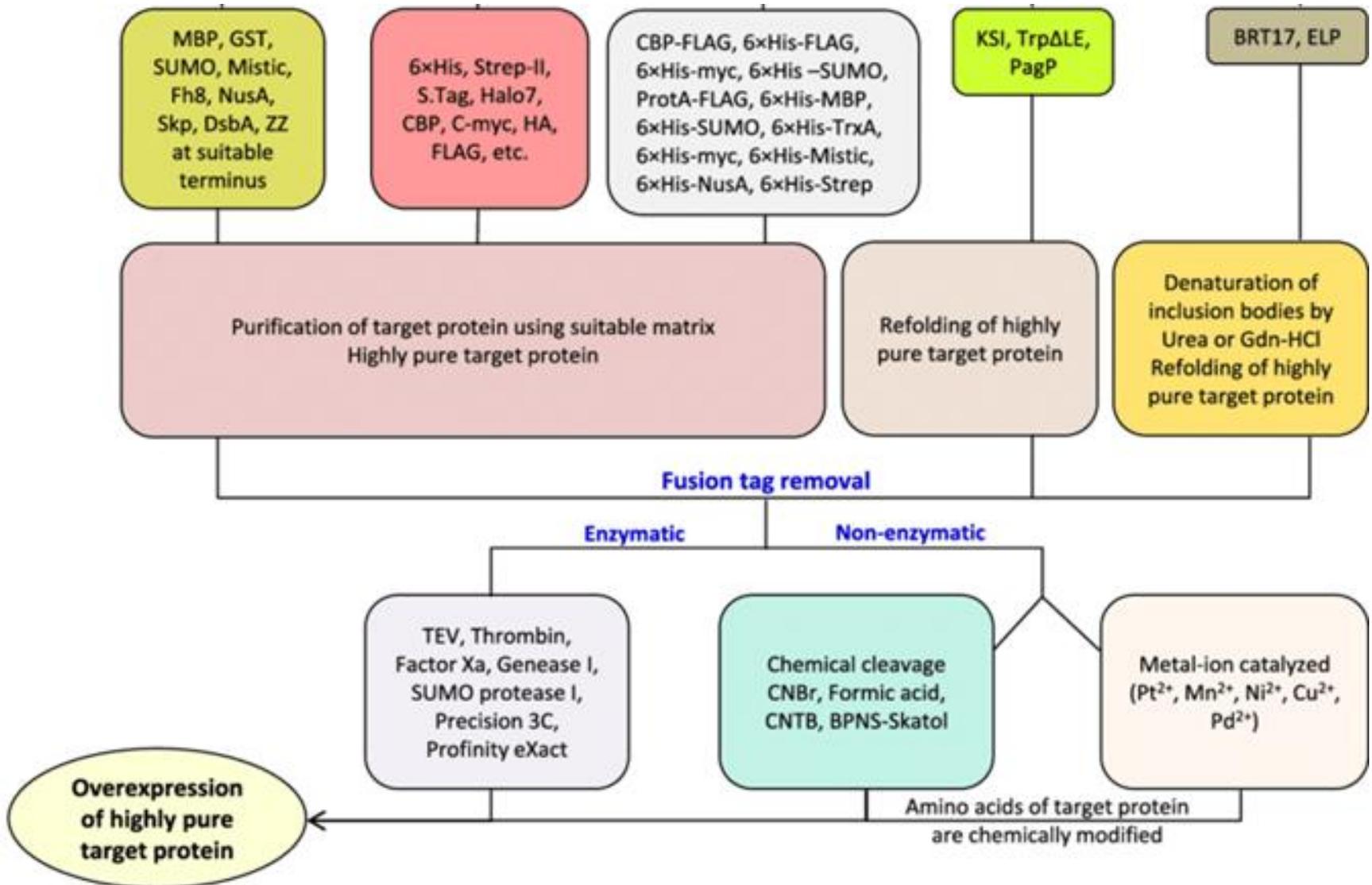
Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: TAGs

Tag	Protein source	Residue size (aa/-kDa)	Tag placement	Source organism	Ligand/Matrix	Purification condition	Suitable expression system	Advantages	Limitations
Affinity epitope tags									
Rho1D4	Last 9 aa of C-terminus of Bovine rhodopsin	9/1.0	C terminus	Synthetic	Monoclonal anti-Rho1D4	Low pH	Bacteria, Yeast, Mammalian cell lines	High specificity, membrane protein purification, Immunolocalization	
c-Myc	Human c-Myc protooncogene	11/1.2	N-/C- or internal	<i>H. sapiens</i>	Anti-Myc antibody (Myc1-9E10) coupled on agarose	Low pH	Mammalian cell lines	Phenomenal purity	Expensive due to limited reuse
PDZ	Ina D	80-90/-9-10	N-/C- or internal	<i>H. sapiens</i>	Immobilized C-terminal PDZ ligand sequence	DTT or β -mercaptoethanol	Insect, Mammalian	Specificity and high affinity	
HA	Human influenza virus hemagglutinin protein	9/1.0	N-/C- or internal	Human influenza virus	Monoclonal antibody immobilized on agarose	Low pH or HA peptide	Mammalian cell lines	Useful in detection, isolation and purification	Affects cellular trafficking and folding
FLAG [®]	Engineered	8/1.01	N- or C-terminus	Synthetic	Anti-FLAG antibody immobilized on agarose	2-5 mM EDTA	Bacteria, Yeast, Insect, Mammalian	Minimal impact on target	Expensive & limited reuse

Espressione di proteine ricombinanti per studi strutturali

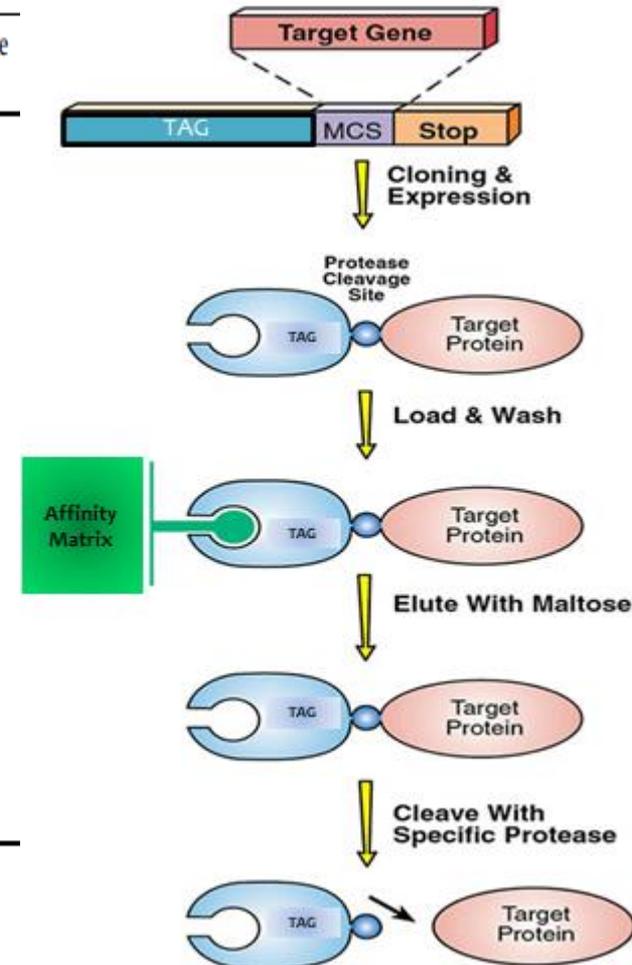
Scelta del vettore: siti di taglio per Proteasi



Espressione di proteine ricombinanti per studi strutturali

Scelta del vettore: siti di taglio per Proteasi

Protease	Cleavage site	Location	Residual aa	pH range	Chaotropic sensitivity	Salt sensitivity	Enzyme/substrate ratio
TEV, AcTEV, ProTEV	ENLYFQ* (G/S)	N	G/S	5.5	2 M urea	<0.1 M	1-3% (wt/wt)
		C	ExxYxQ	-8.5			
Thrombin (fIIa)	LVPR*GS	N	G	5-10	<0.1 M urea	<0.1 M	1-10% (wt/wt)
		C	LVPR				
Factor Xa (fXa)	IEGR* x x ≠ R, P	N	None	6-9	<0.1 M urea	<0.15 M	1-10% (wt/wt)
		C	IEGR				
Genesee I							
Enterokinase	DDDDK* x	N	None	7-8			0.1% (wt/wt)
SUMO proteases	C-terminus of SUMO* xaa- GG/yaa	N only	None	6-9	<0.1 M GnCl 2 M urea	<0.5 M	0.1% (wt/wt)
		C	DDDDK				
PreScission (human rhinovirus 3C protease)	LEVLFQ* GP	N					
TVMV (Tobacco vein mottling virus protease)	ETVRFQG* S	N	GP				
Immobilized subtilisin (BPN); Profinity eXact system	C-terminus of propeptide* x EEDKLFAL	N only	None	7.2	2 M urea		1:1 (mol: mol)

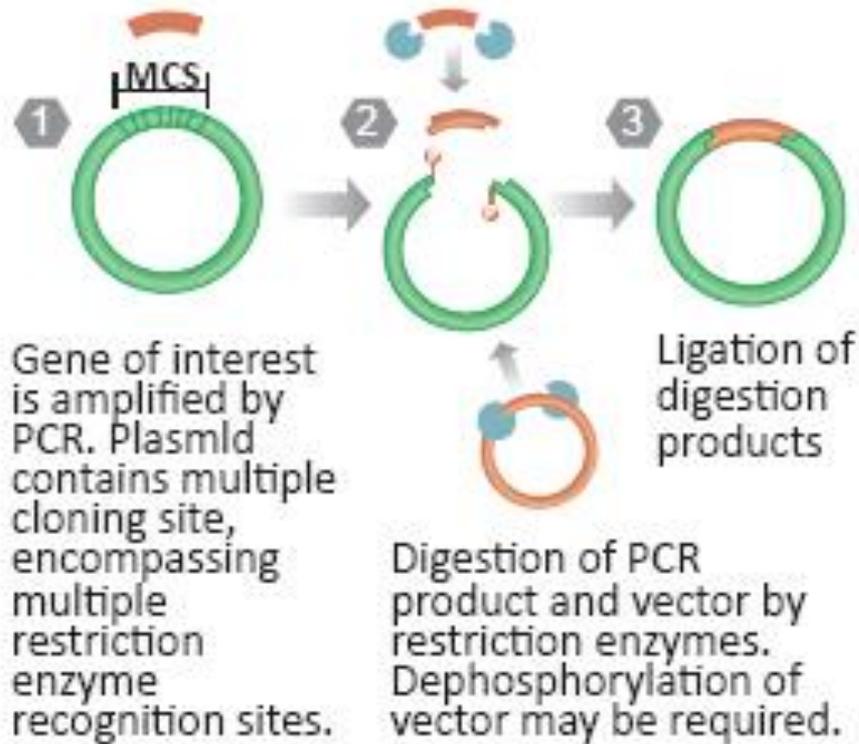


Espressione di proteine ricombinanti per studi strutturali

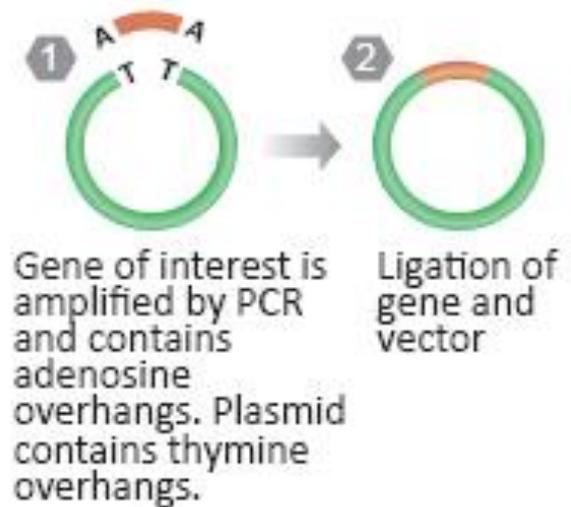
Clonaggio:

Ligation-dependent

Clonaggio per restrizione



TA cloning



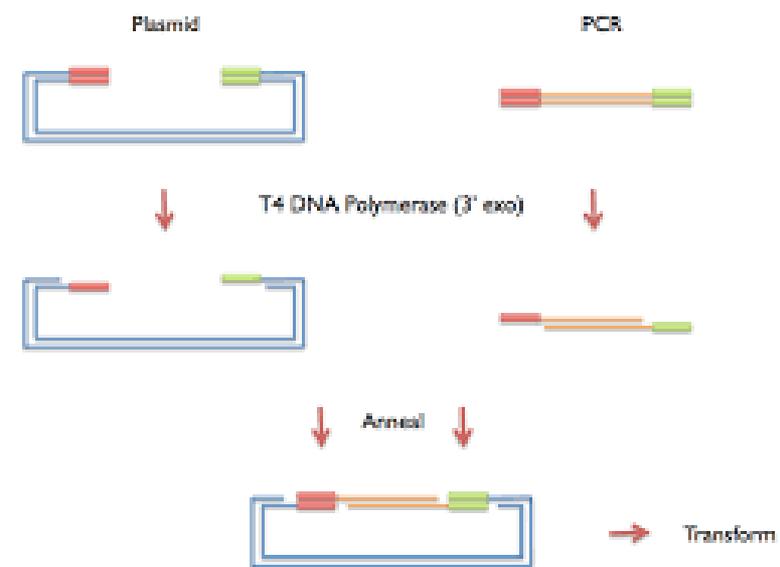
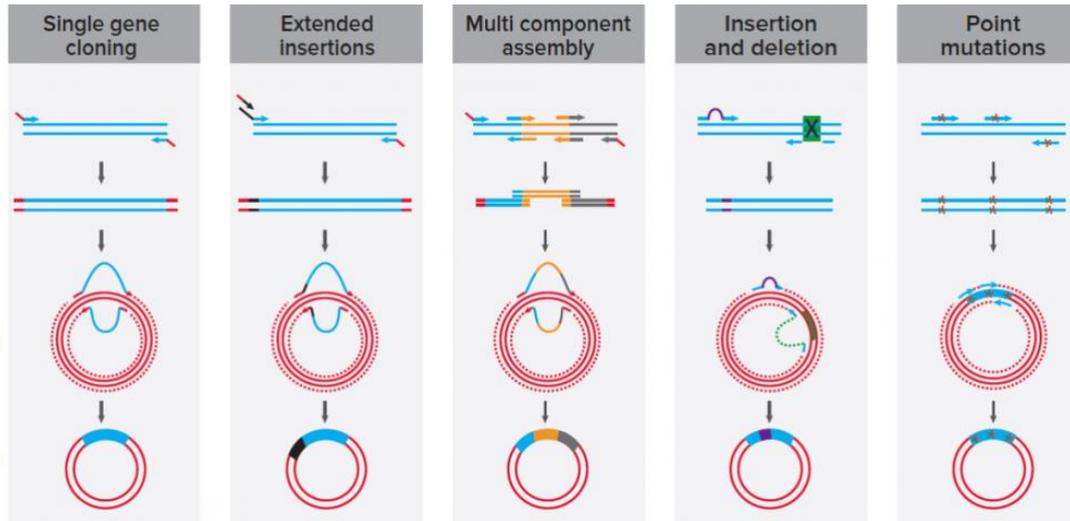
Espressione di proteine ricombinanti per studi strutturali

Clonaggio:

Ligation-independent

RF-cloning

LI- cloning (LIC)



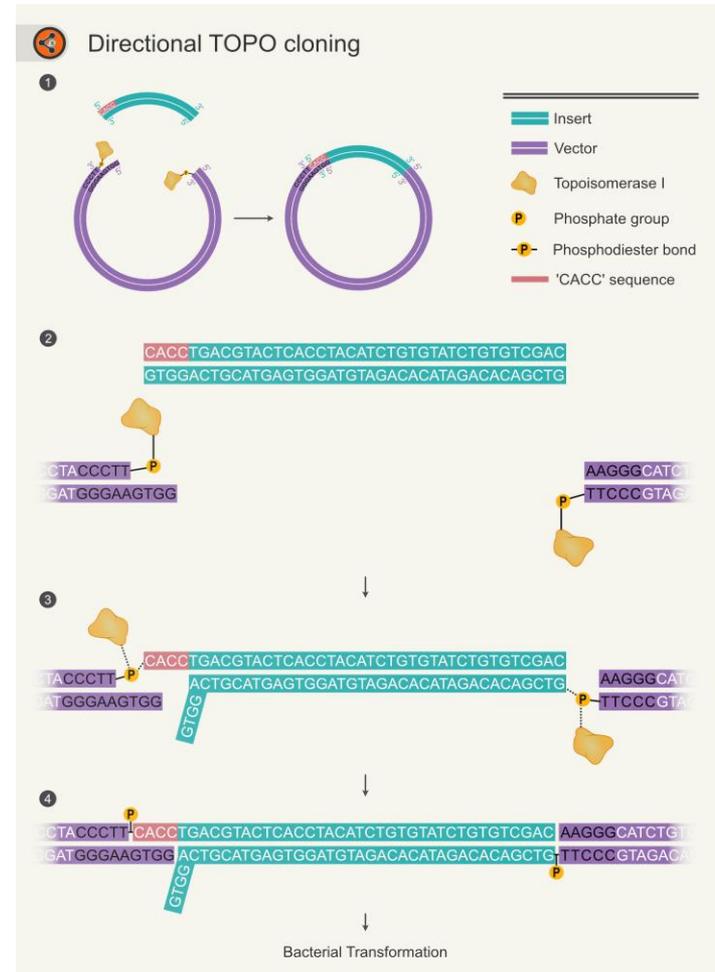
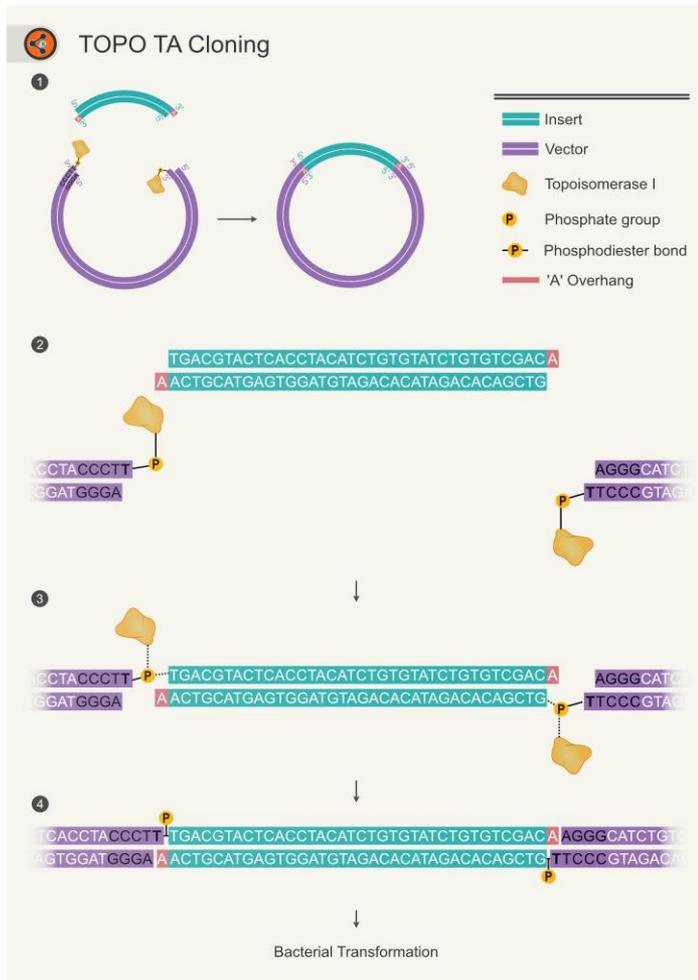
Espressione di proteine ricombinanti per studi strutturali

Clonaggio:

Ligation-independent

TOPO cloning:

Vaccinia Topoisomerasi I (InVitrogen)



Espressione di proteine ricombinanti per studi strutturali

Clonaggio:

Ligation-independent

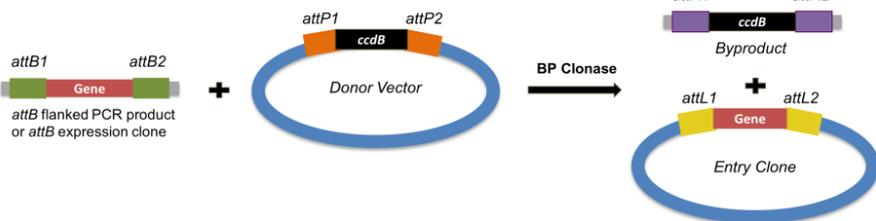
Gateway cloning:

fago λ – attB/P (Life technologies)

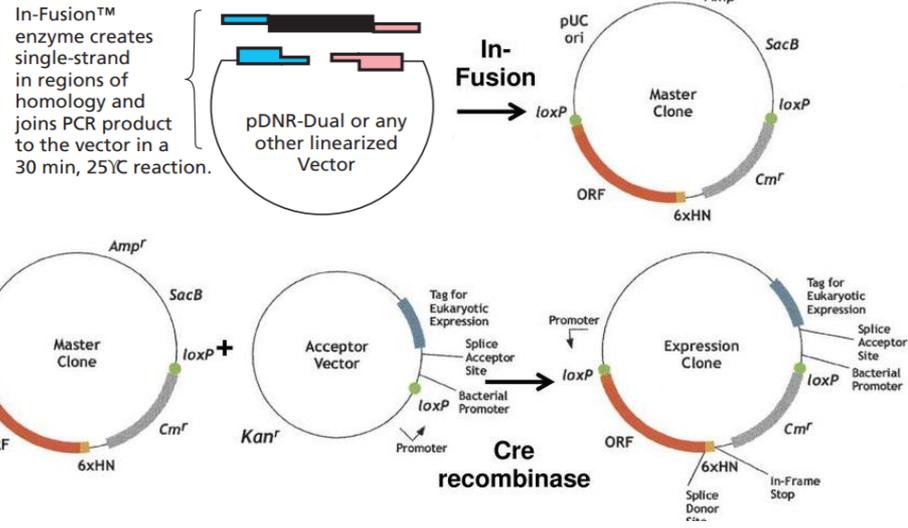
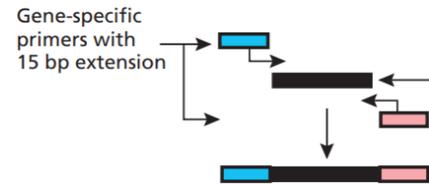
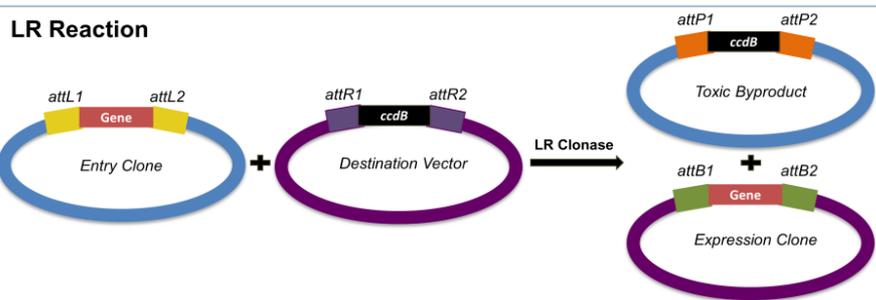
In Fusion-Creator cloning

fago P1 Cre-LoxP (Clontech)

BP Reaction



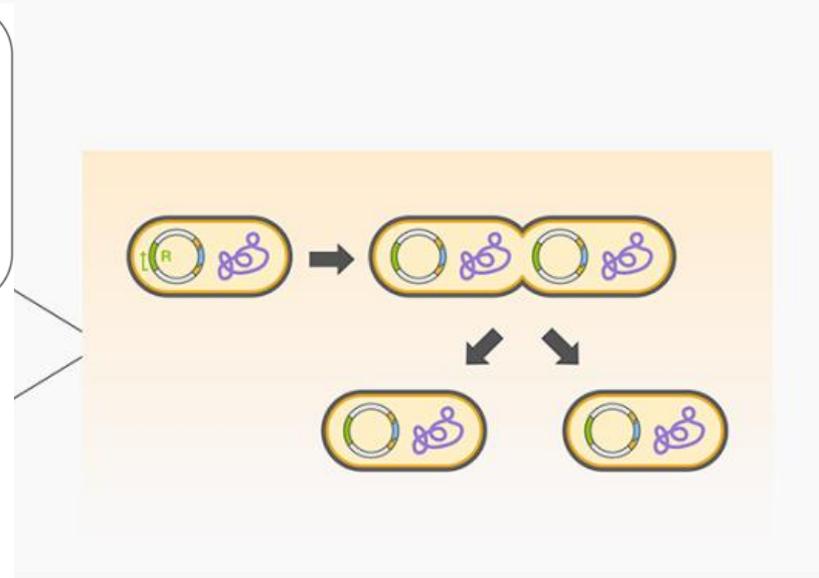
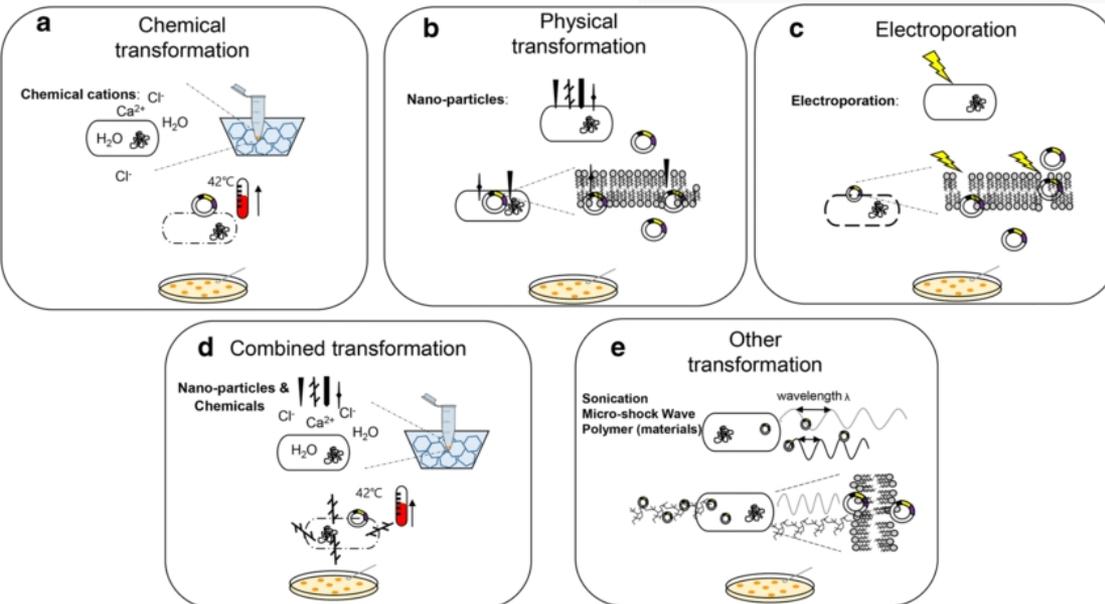
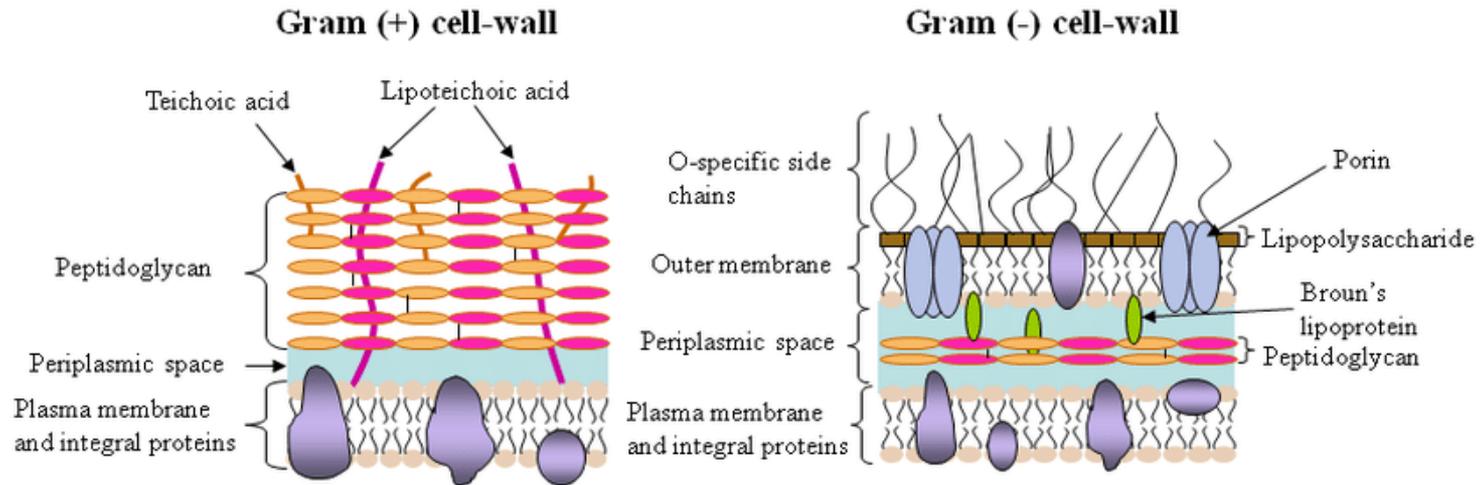
LR Reaction



Espressione di proteine ricombinanti per studi strutturali

Inserimento nella cellula ospite:

Trasformazione: Batteri

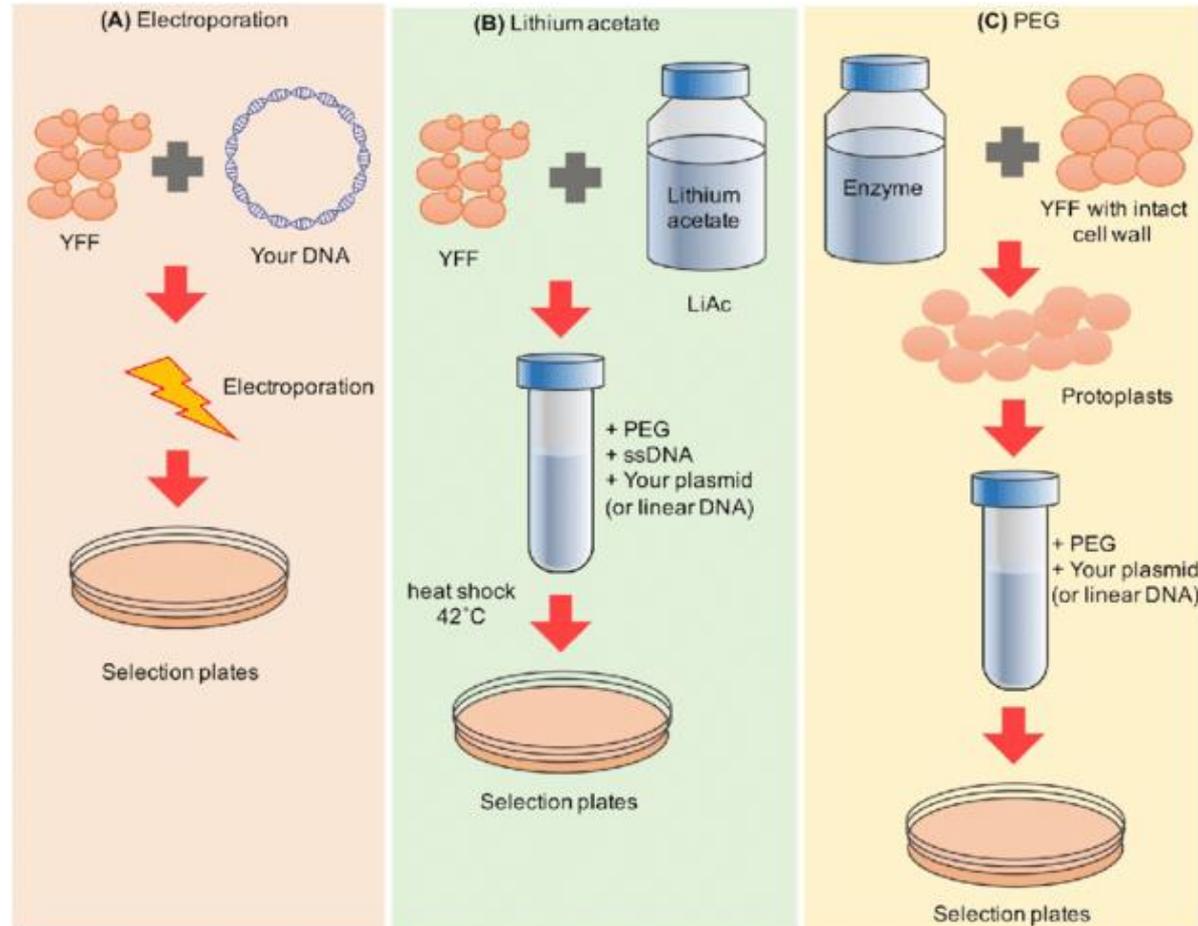


Espressione di proteine ricombinanti per studi strutturali

Inserimento nella cellula ospite:

Trasformazione: Lievito

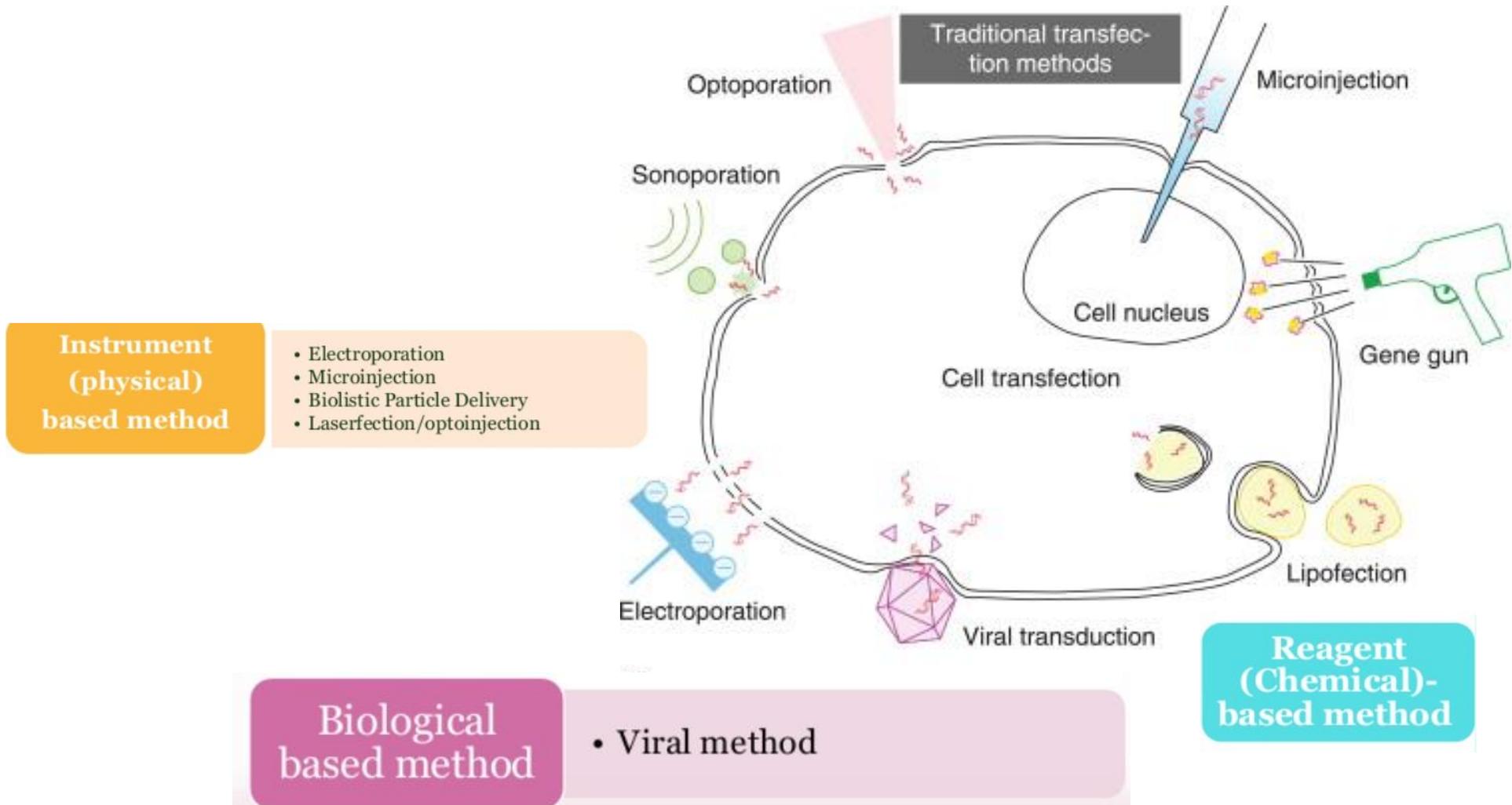
Yeast Cell Wall



Espressione di proteine ricombinanti per studi strutturali

Inserimento nella cellula ospite:

Trasfezione/Trasduzione: Cellule insetto/mammifero

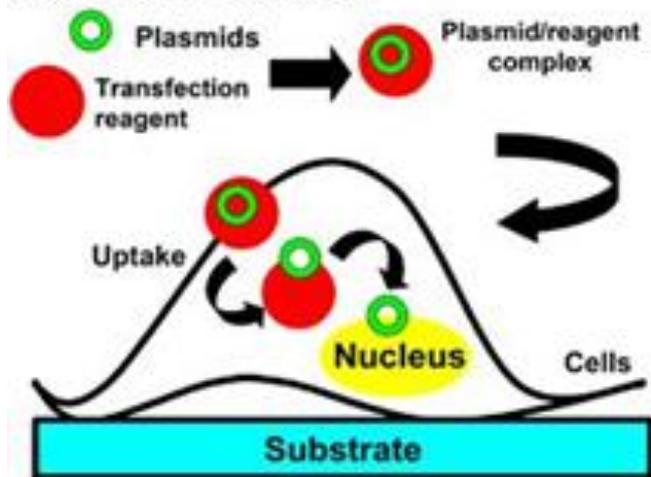


Espressione di proteine ricombinanti per studi strutturali

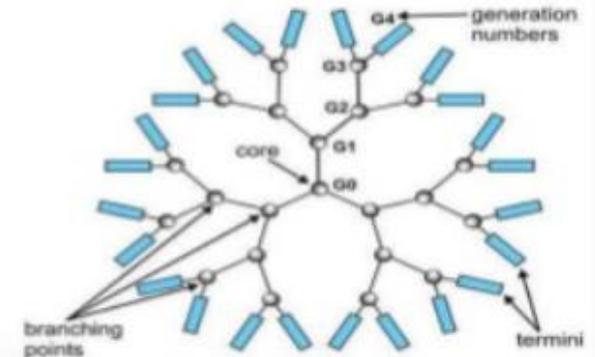
Inserimento nella cellula ospite:

Cellule insetto/mammifero: Trasfezione

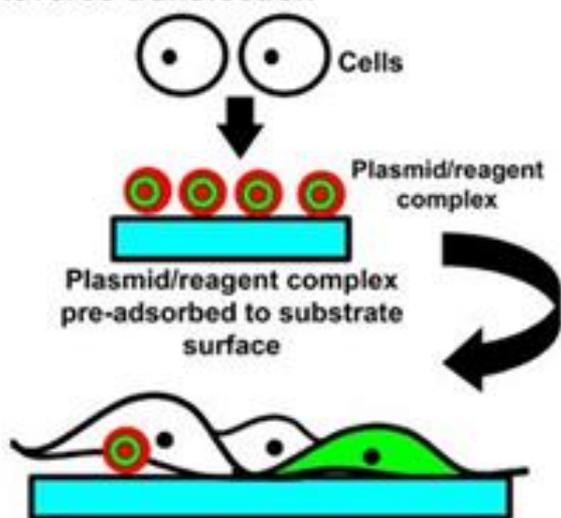
Non-viral transfection



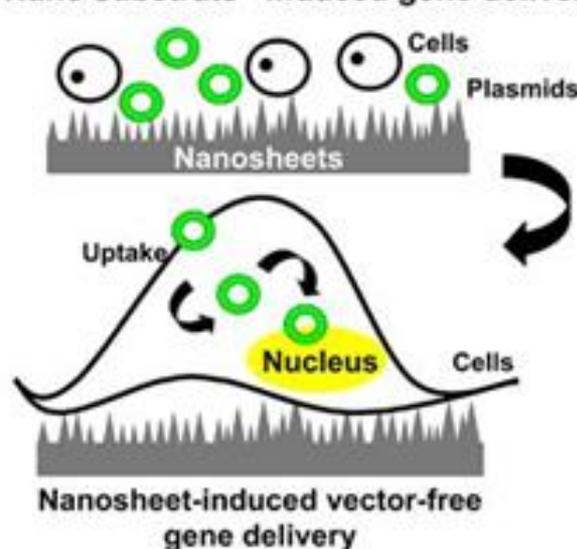
- Cationic Lipids : Liposome/Lipoplexes
- Calcium phosphate
- Cationic polymer
- DEAE-Dextran
- Magnet-mediated transfection
- Activated dendrimers



Reverse transfection



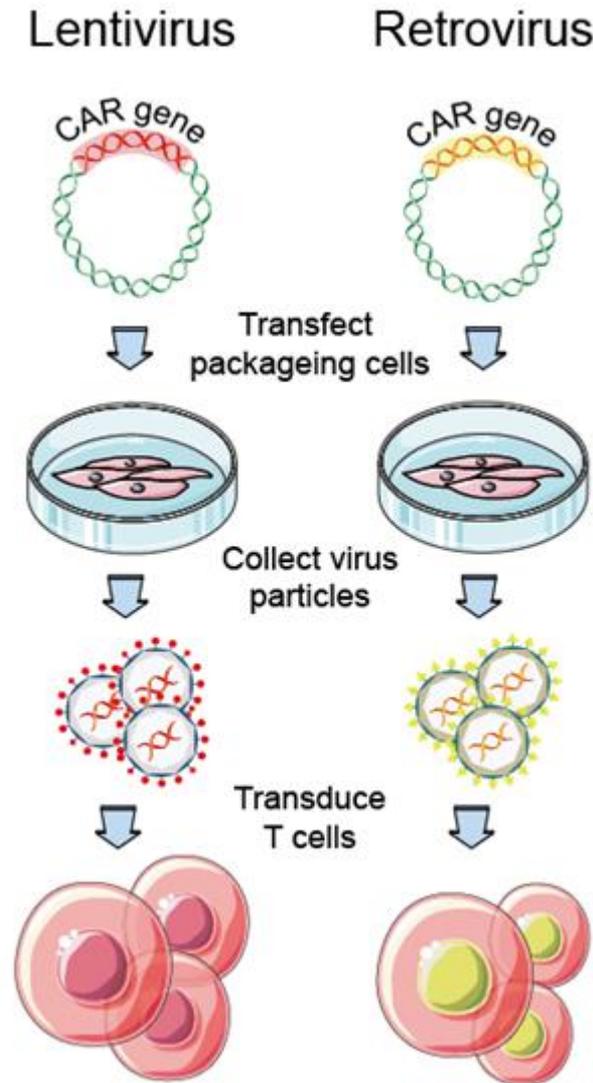
"Nano substrate"-induced gene delivery



Espressione di proteine ricombinanti per studi strutturali

Inserimento nella cellula ospite:

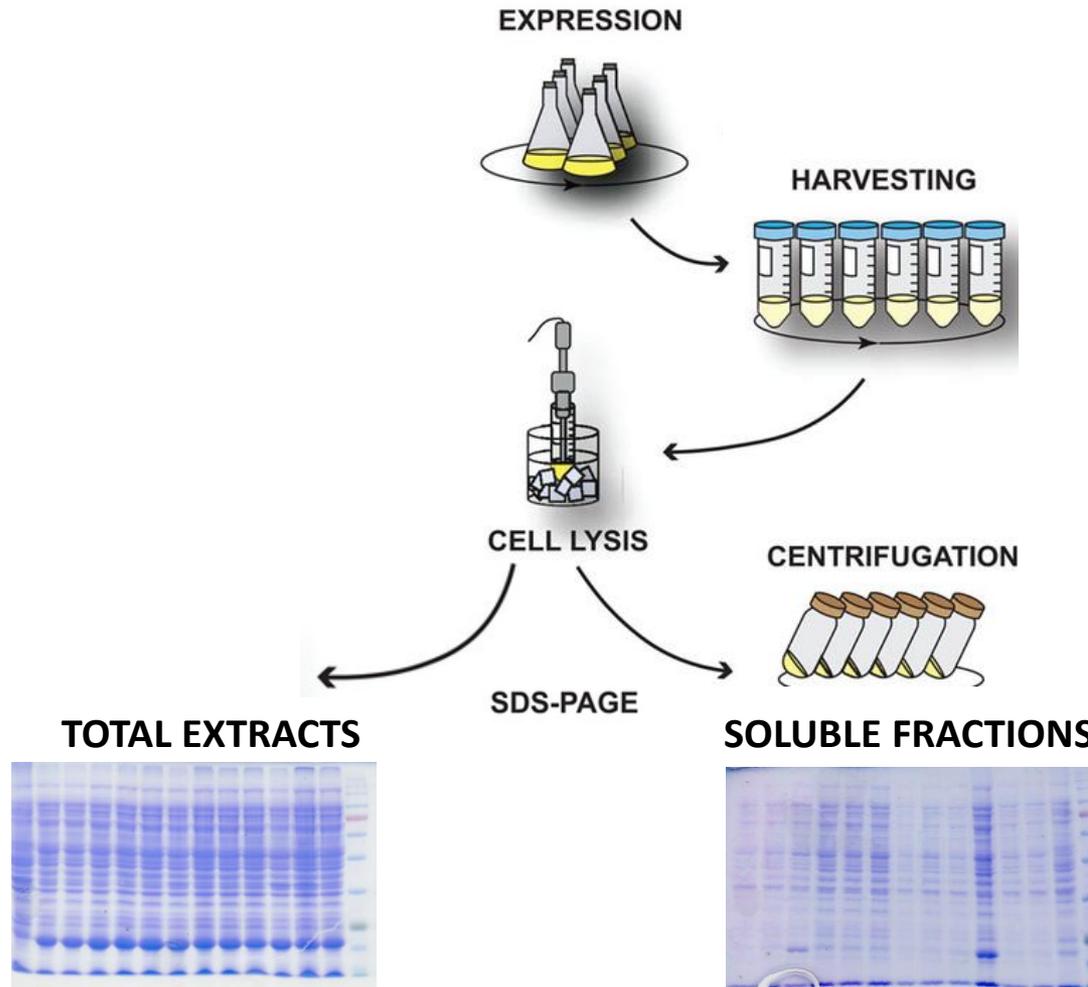
Cellule insetto/mammifero: Trasduzione



Espressione/purificazione di proteine ricombinanti per studi strutturali

Small scale inducible expression:

To test/optimize different strains/ expression conditions ($^{\circ}\text{T}$ /[inducer]/media/time)



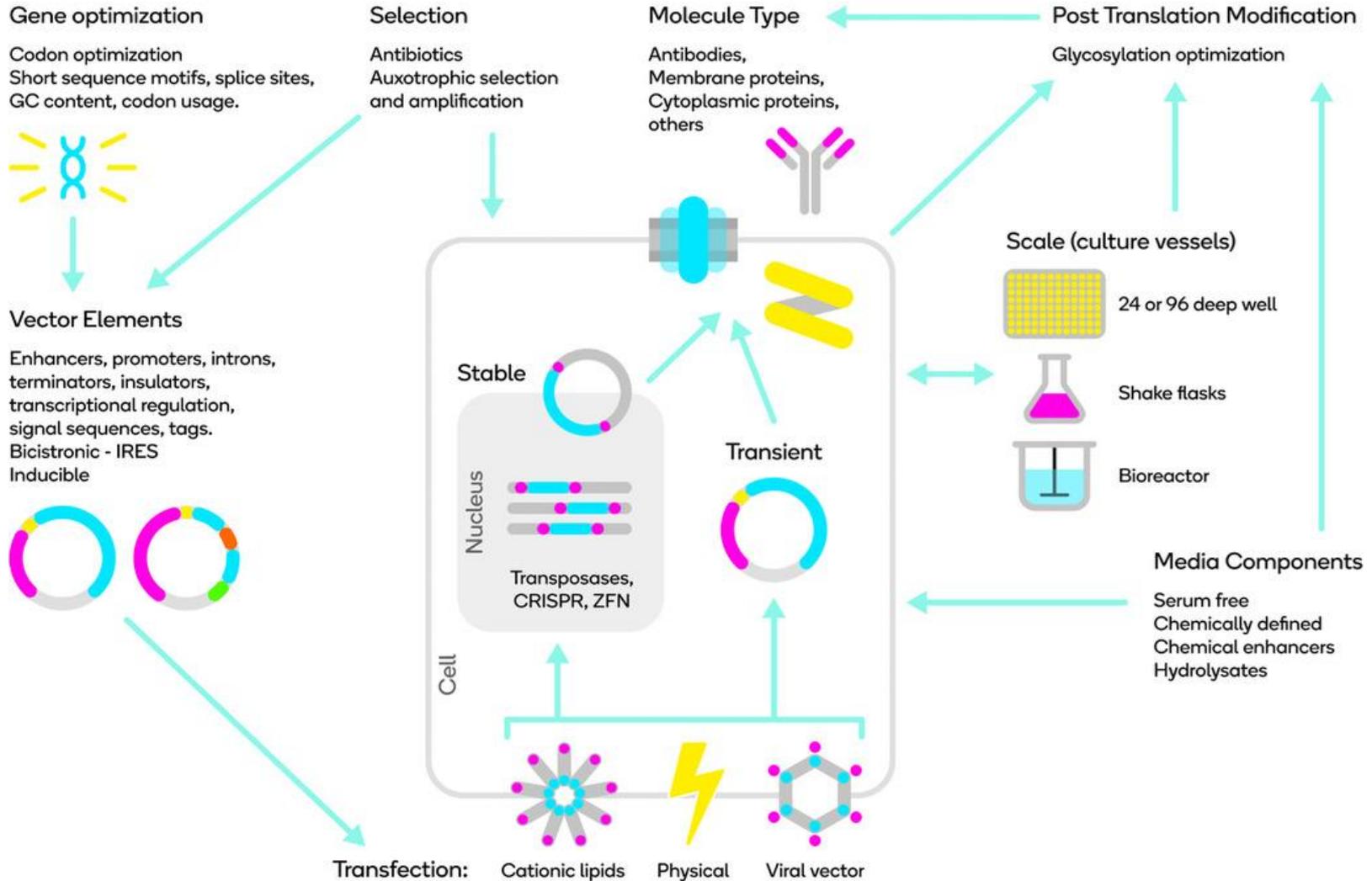
Large scale expression:

To adjust optimized expression conditions to larger scales

Espressione/purificazione di proteine ricombinanti per studi strutturali

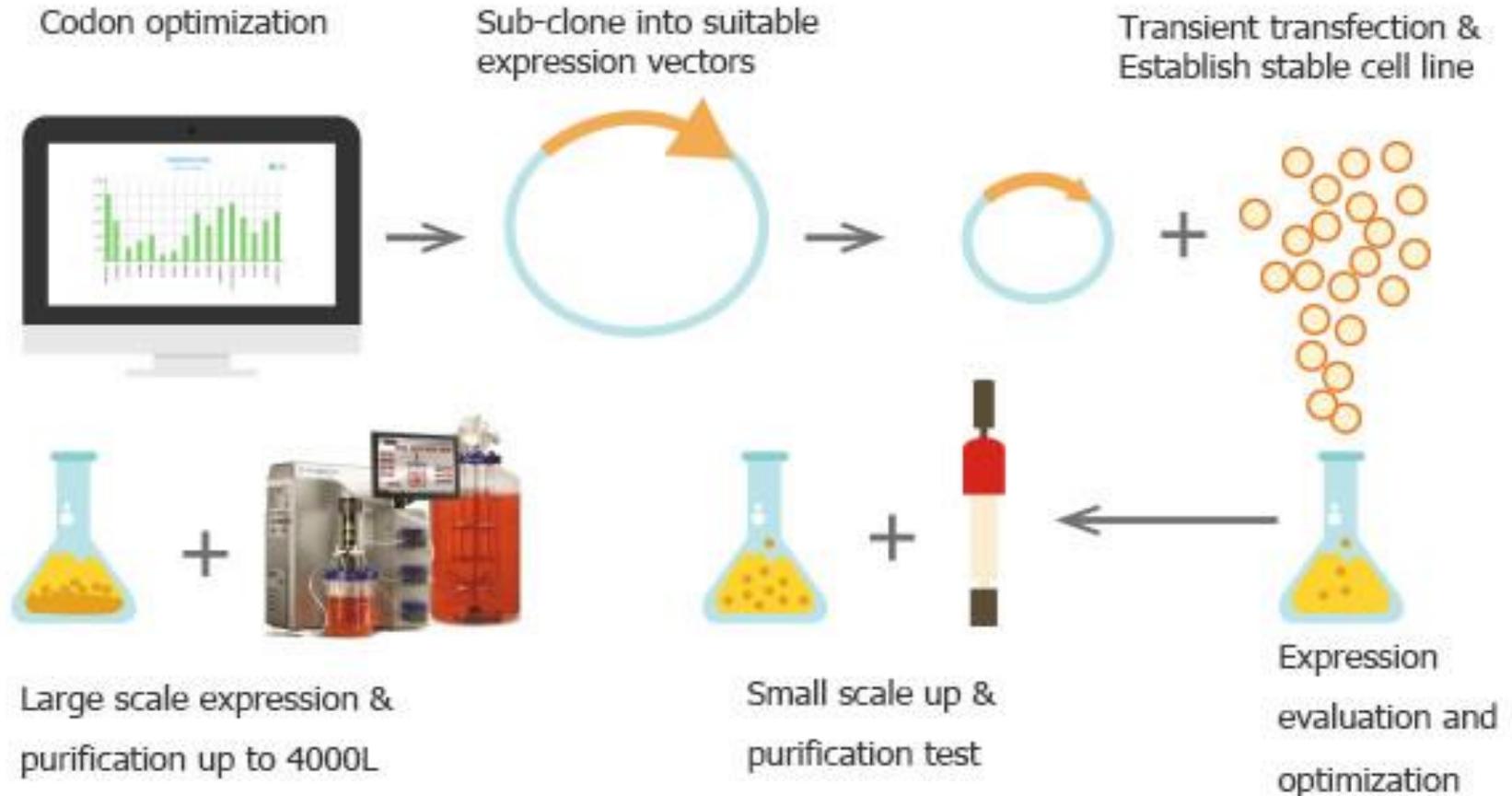
Small scale inducible expression:

To test/optimize different strains/ expression conditions



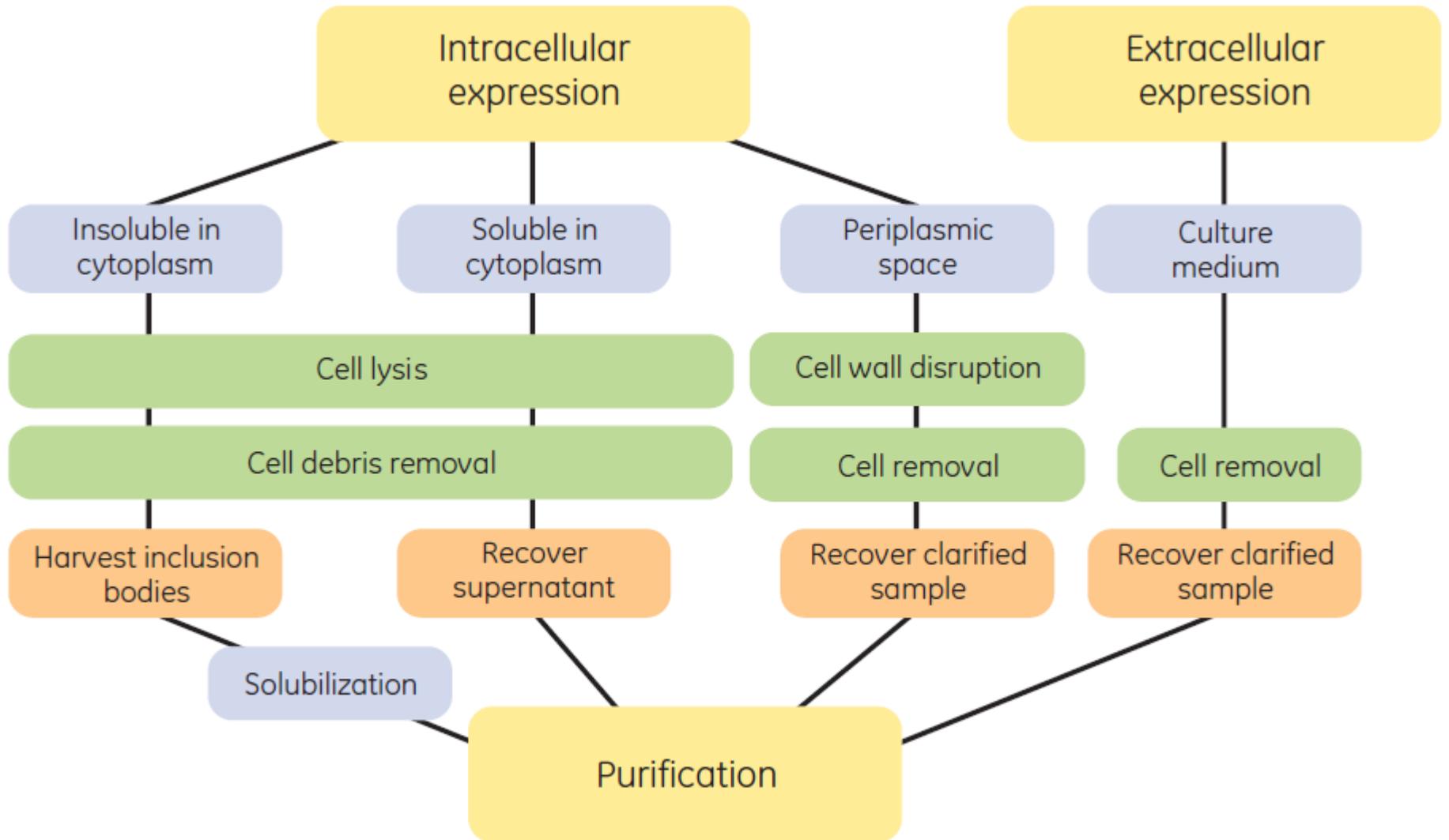
Espressione/purificazione di proteine ricombinanti per studi strutturali

Large scale expression: To adjust optimized expression conditions to larger scales



Espressione di proteine ricombinanti per studi strutturali

expression flowchart



Espressione di proteine ricombinanti per studi strutturali

collection & lysis

CENTRIFUGAZIONE

SURNATANTE per il recupero di proteine secrete

PELLET: **LISI CELLULARE** per recuperare proteine

Periplasmatiche: lisozima/cold osmotic shock/ MgCl₂
citoplasmatiche / di membrana o proteine in corpi di inclusione
lisato (proteina target, insieme a proteine solubili dell'ospite,
membrane e proteine di membrana, organelli, DNA, matrice
citosolica della cellula)



Cellule di **INSETTO/MAMMIFERO**: soluzione ipotonica (Shock osmotico)

BATTERI E LIEVITI: più difficile

Metodo enzimatico
(lisozima)

Sonicazione
(alta pressione localizzata)

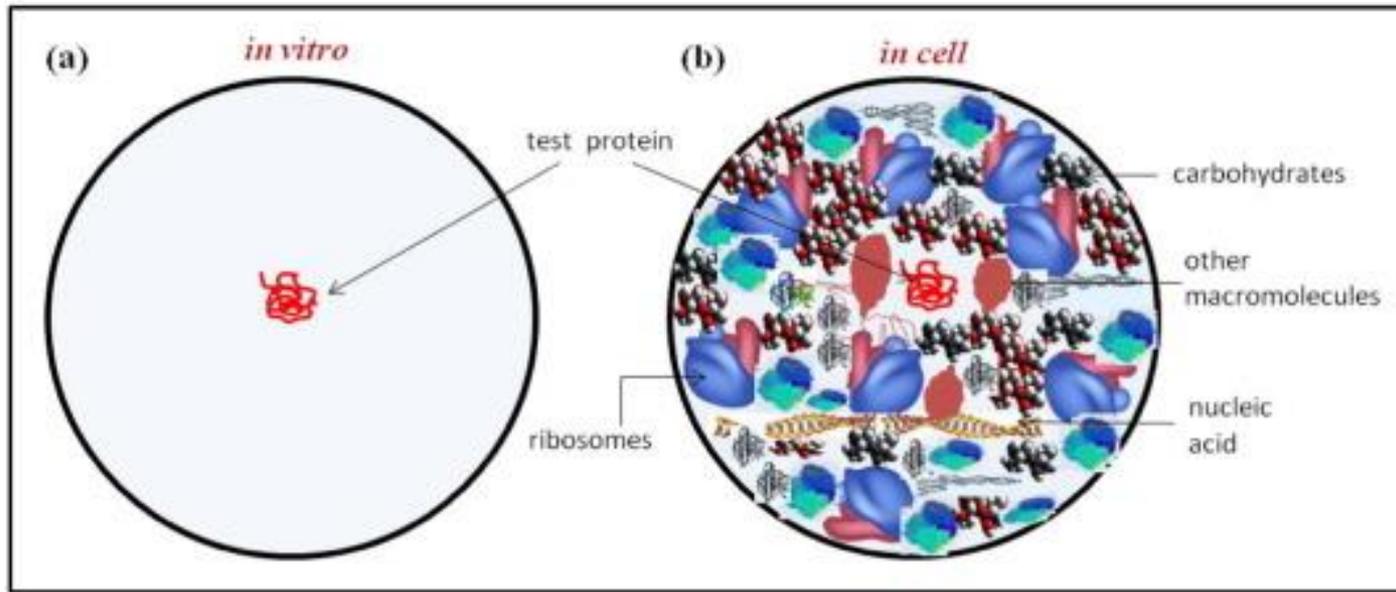
French press & Microfluidizer
(valvola stretta ad alta pressione)

Bead Mill
(microsfere e agitazione
ad alta frequenza)

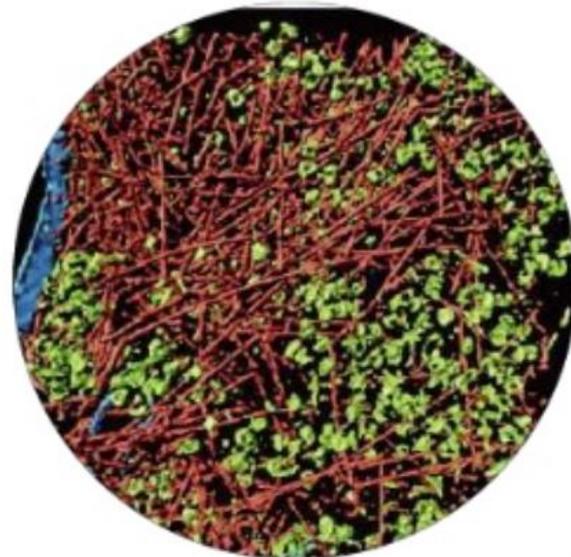
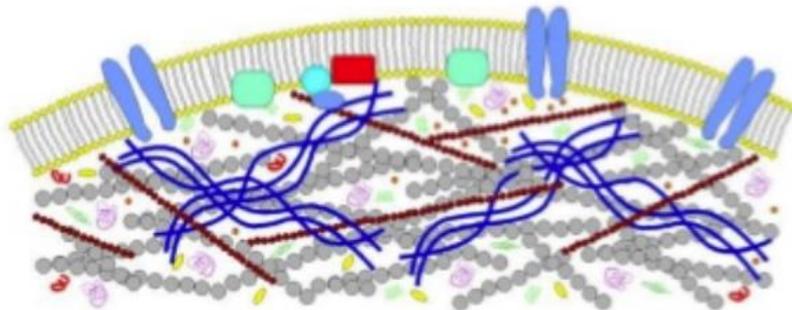


Espressione di proteine ricombinanti per studi strutturali

Folding in vivo:



cytoplasm crowding
Confinement

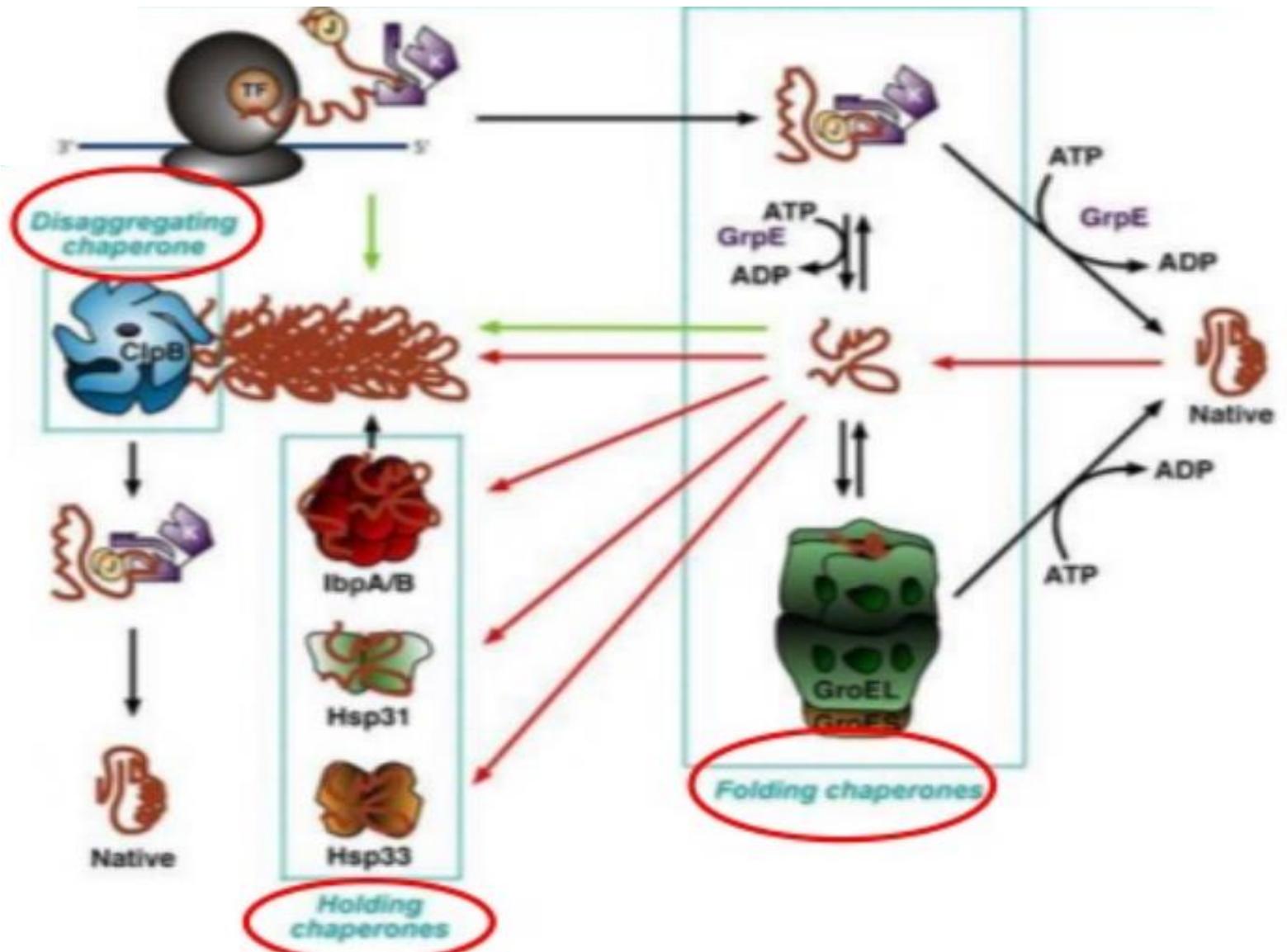


Espressione di proteine ricombinanti per studi strutturali

Folding in vivo:

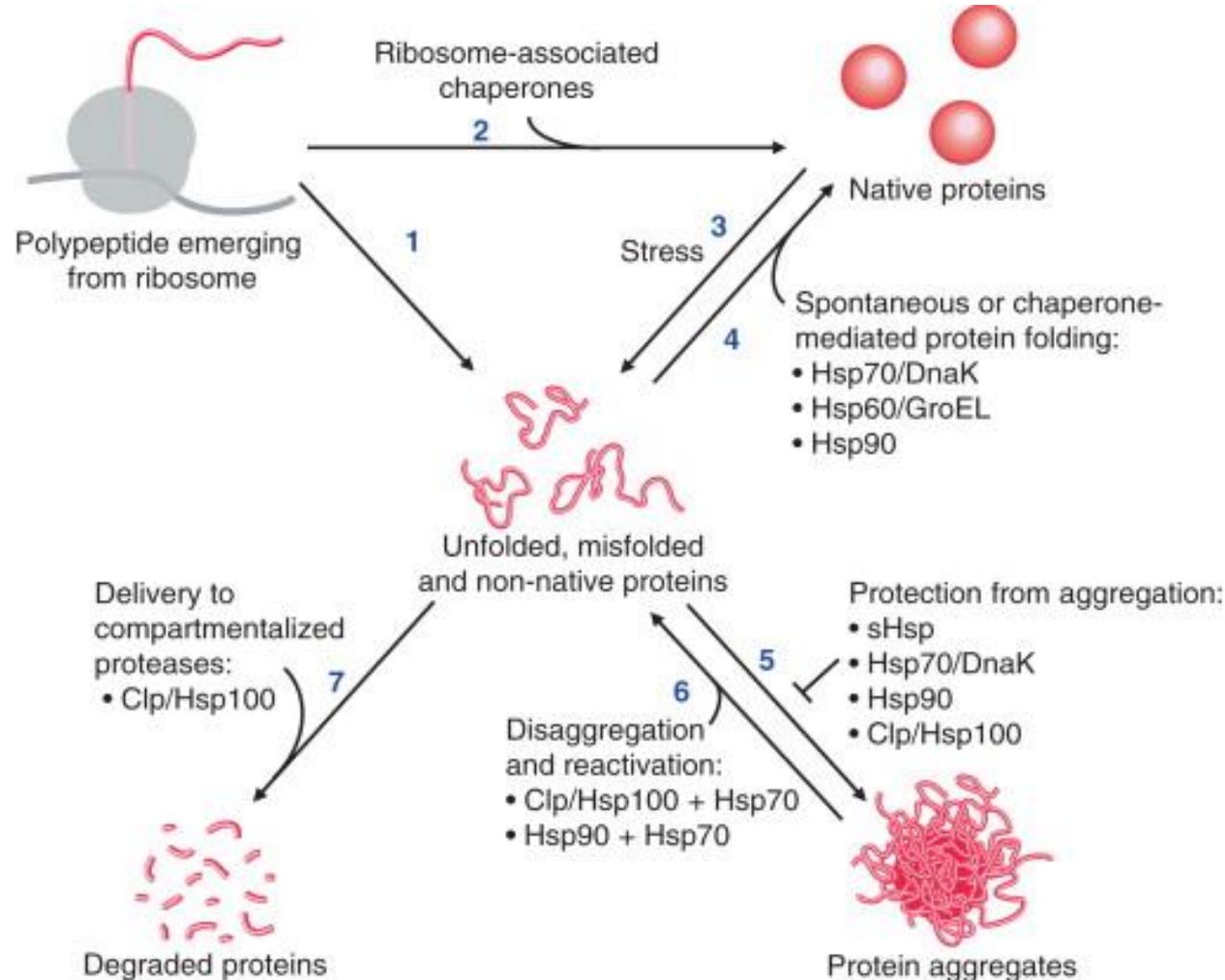
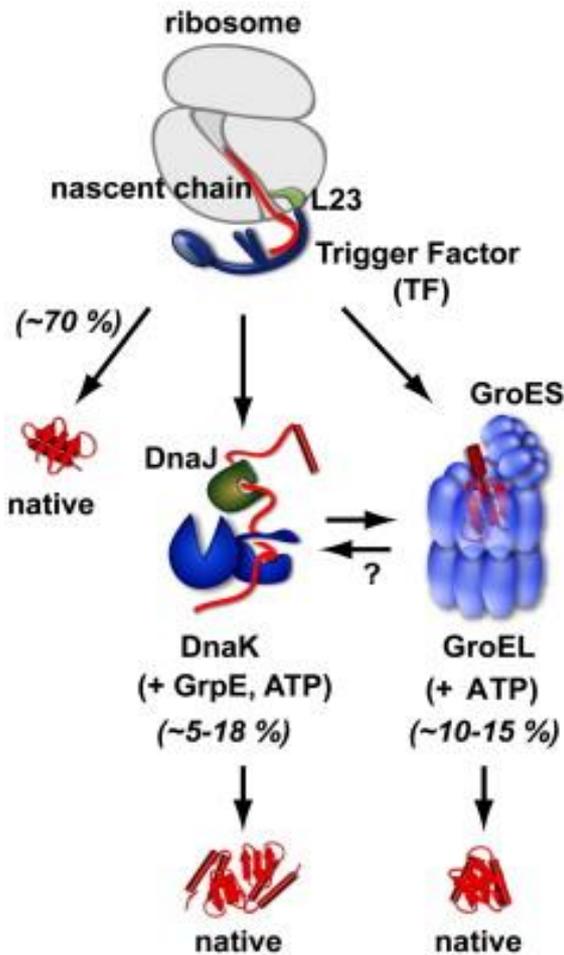
Espressione di proteine ricombinanti per studi strutturali

Folding in vivo:



Espressione di proteine ricombinanti per studi strutturali

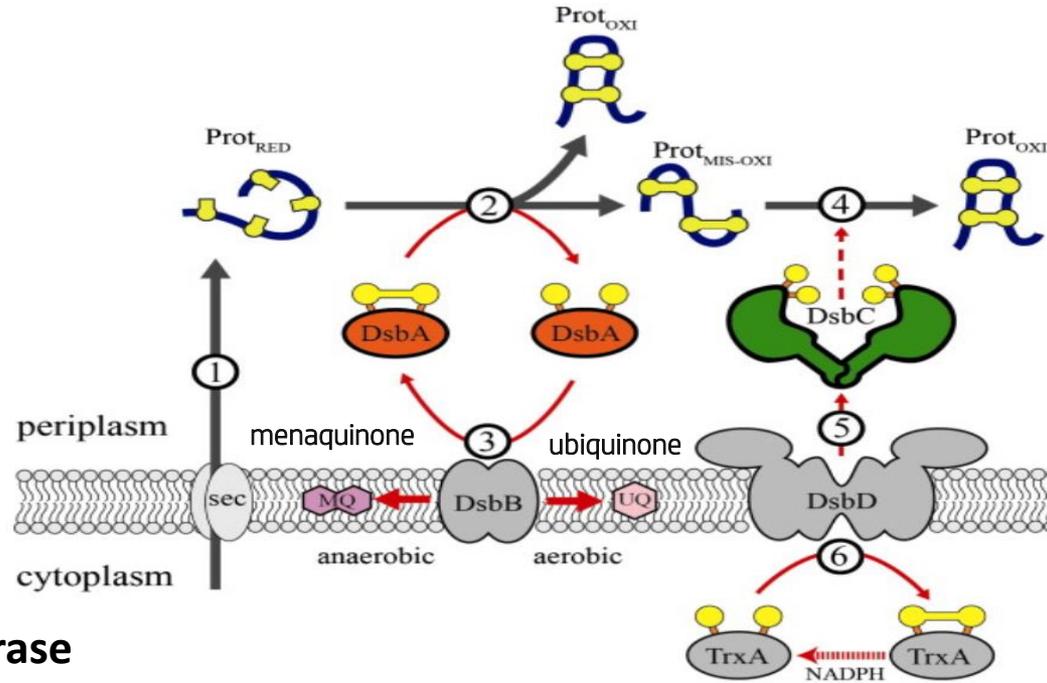
Folding in vivo:



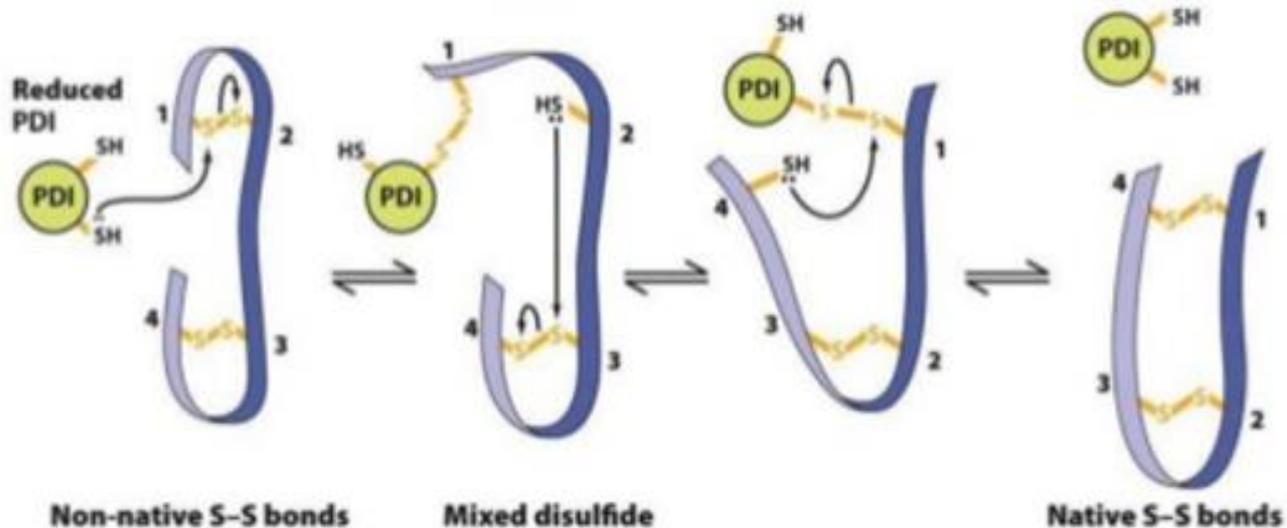
Espressione di proteine ricombinanti per studi strutturali

Folding in vivo: ponti di solfuro

Ruolo delle Dsb A/C



Protein Disulfide Isomerase



Espressione di proteine ricombinanti per studi strutturali

Espressione in forma insolubile e refolding:

L'over-espressione satura il sistema enzimatico:

aggregazione/accumulo di proteine misfolded in corpi di inclusione:

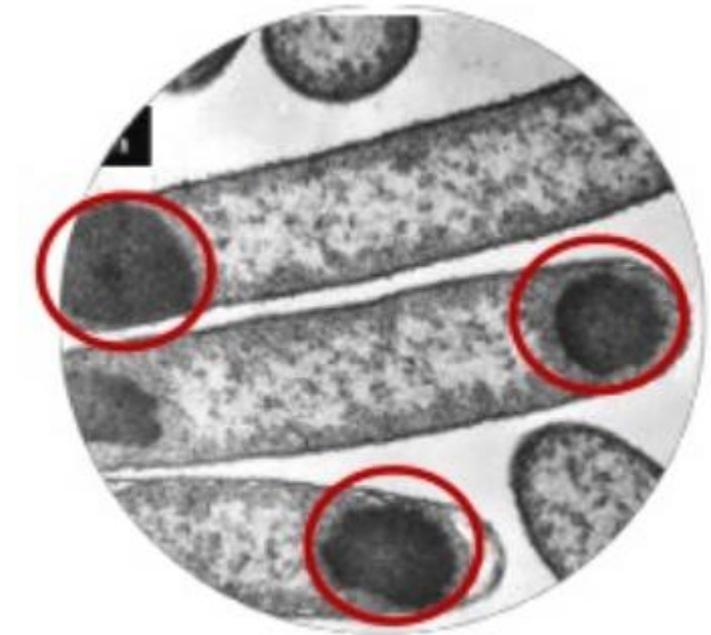
Vantaggi:

Protezione da proteasi citoplasmatiche

Elevate rese

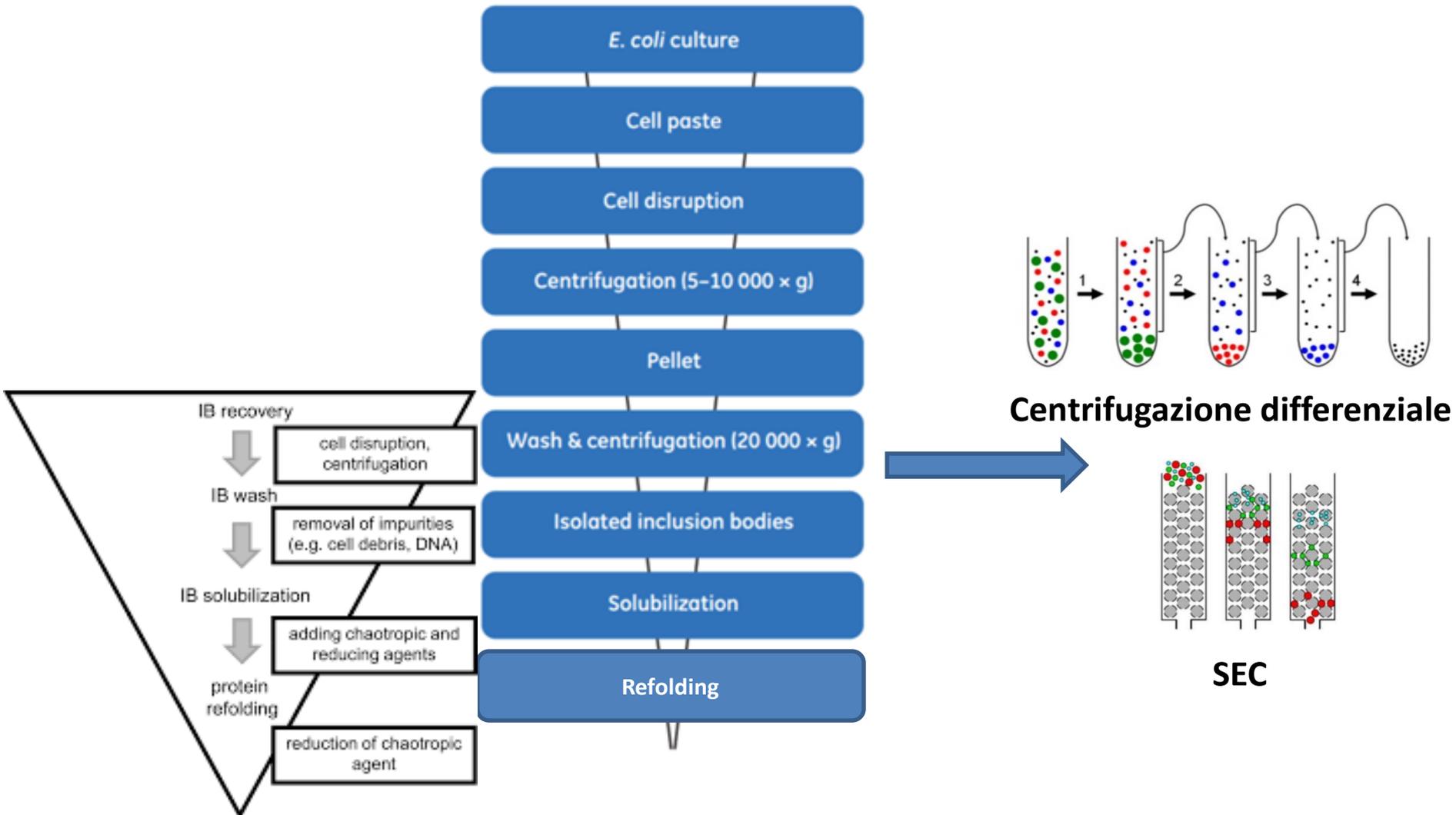
Elevata purezza

Produzione di proteine tossiche



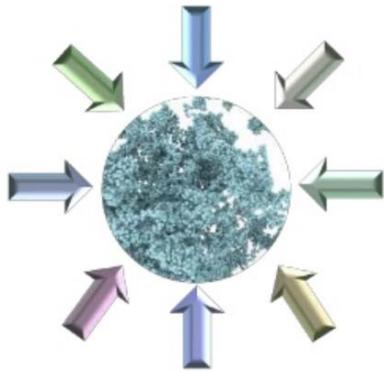
Espressione di proteine ricombinanti per studi strutturali

Espressione in forma insolubile e refolding:

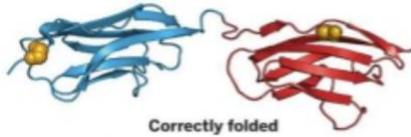


Espressione di proteine ricombinanti per studi strutturali

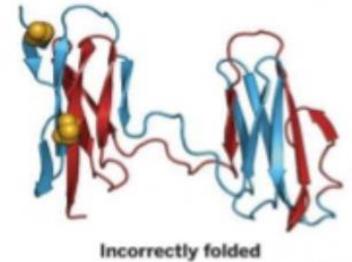
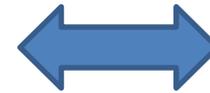
Espressione in forma insolubile e refolding:



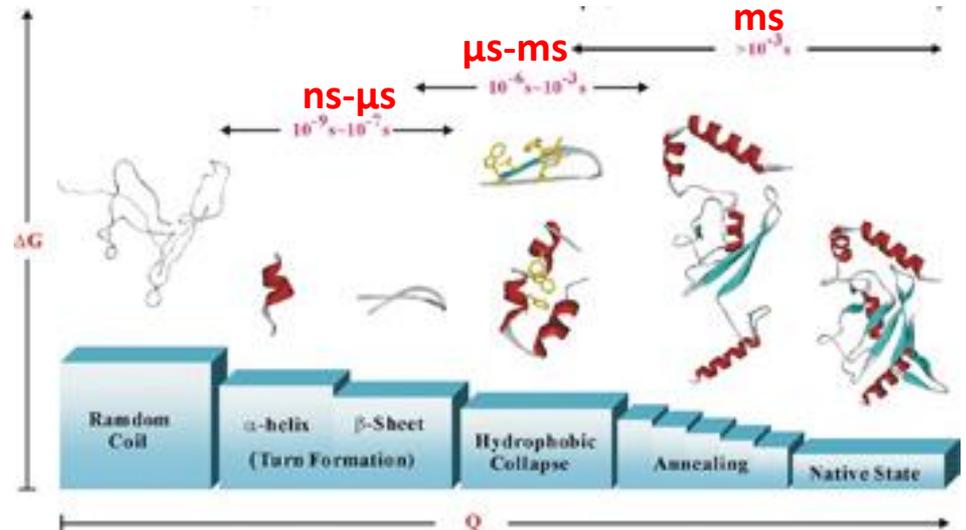
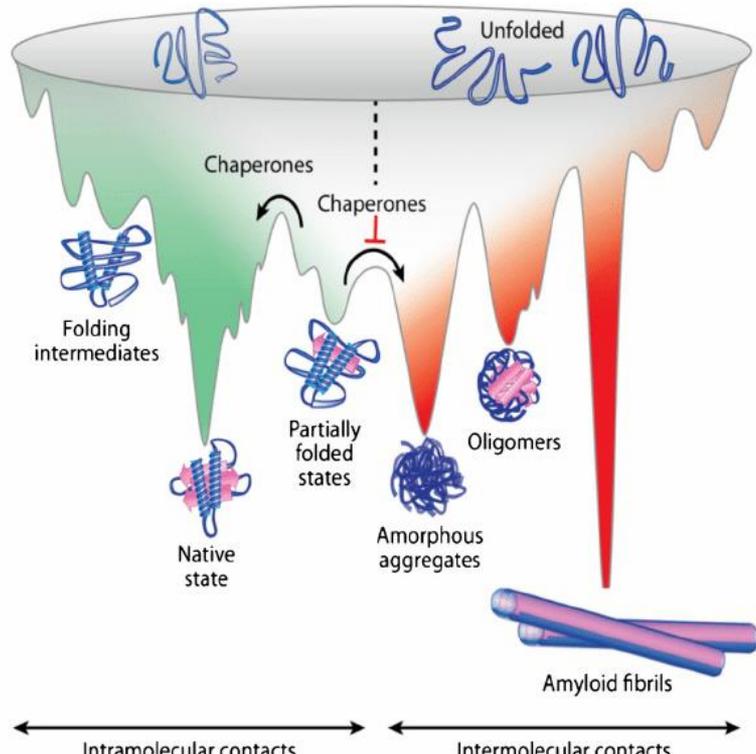
Aggregazione



Refolding

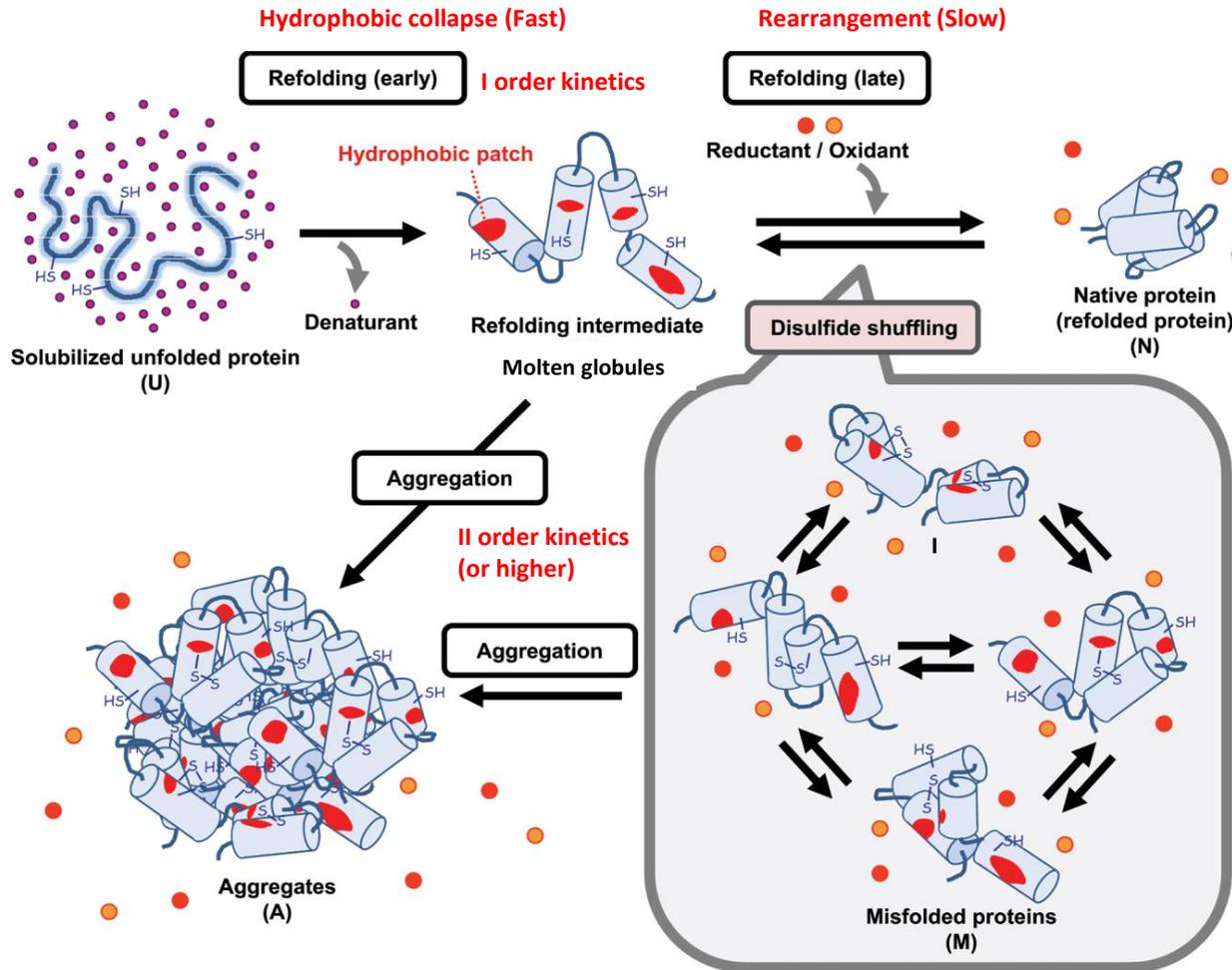


Misfolding



Espressione di proteine ricombinanti per studi strutturali

Espressione in forma insolubile e refolding:



Driving forces: interazioni idrofobiche > WdW > Pro isomerization > ponti di solfuro > legami H

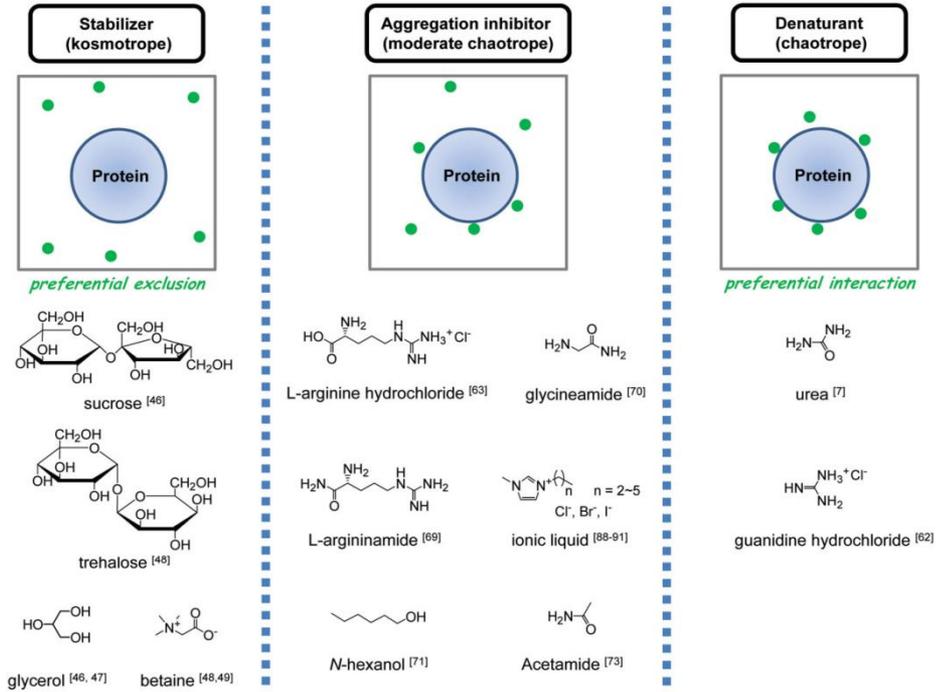
Espressione di proteine ricombinanti per studi strutturali

Metodi di refolding: il tampone di refolding

Da ottimizzare vari parametri (T, t, pH, forza ionica)

Impiego di additivi

Additive	Concentration	Effect
CHAPS	30 mM	Detergent
EDTA	20 mM	Chelate
Glycerol	10–50%	Stabilizer
Guanidine HCl	0.1–1 M	Chaotroph
L-arginine	0.4–0.5 M	Stabilizer
Laroylsarcosin	Up to 4 M	Detergent
MgCl ₂ /CaCl ₂	2–10 mM	Cation
NaCl/Ammonium sulfate	0.2–0.5 M	Salt
Non-detergent sulfo betain	Up to 1 M	Solublizer
PEG 3350 and cyclodextrin	Up to 0.5% W/V	Osmolyte
SDS	0.1%	Detergent
Sodium citrate/sulfate	0.2–0.5 M	Salt
Sucrose/glucose	Up to 0.75 M	Stabilizer
TMAO	1–3 M	Osmolyte
Tris	0.4–1 M	Buffer
Triton X-100	0.1–1%	Detergent
Tween-80	0.01%	Detergent
Glycine	Up to 1 M	Osmolyte
Proline	Up to 1 M	Osmolyte
Urea	0.1–2 M	Chaotroph



Artificial chaperones

Detergents: Triton X-100, CHAPS, Sarkosyl, SDS and CTAB

Detergents with cycloamylose or cyclodextran

Non-detergent zwitterionic agents such as Non-detergent sulfobetaines (NDSB), substituted pyridines and pyrroles, and substituted amino-cyclohexanes

Espressione di proteine ricombinanti per studi strutturali

Metodi di refolding: il tampone di refolding

Impiego di additivi proteici

Chaperone mediated (soluble or immobilized and reusable system format)

ATP dependent: GroEL-GroES, DnaK, DnaJ, GrpE minichaperones

ATP independent: peptidyl-prolyl isomerase

DsbA, Trx and GST

Antibody mediated

specific for conformational epitopes

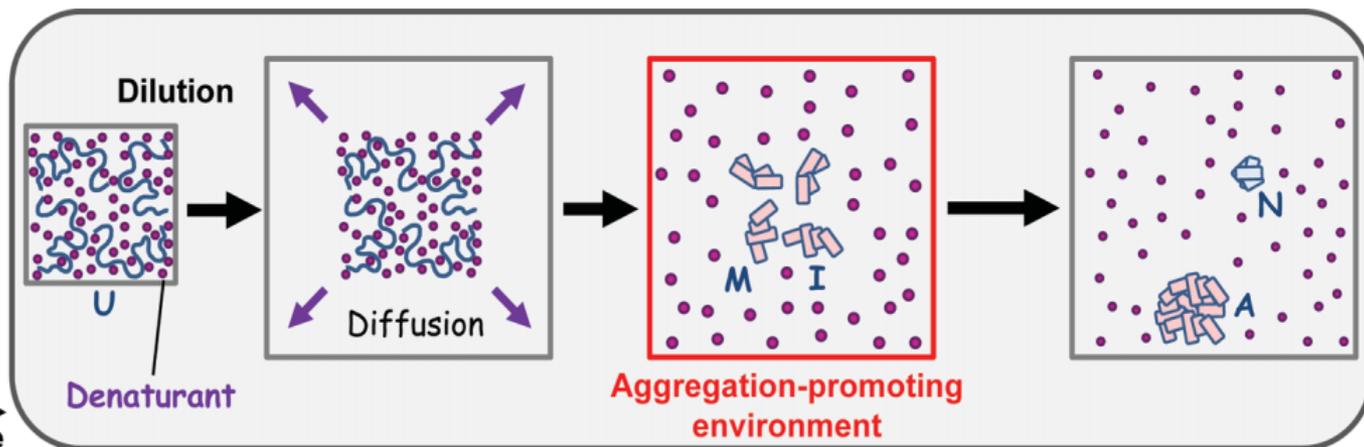
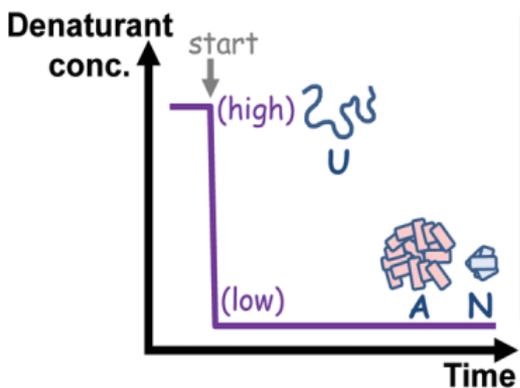
Impiego di agenti redox

Reducing/oxidizing reagents: BME, BMC, TCEP, DTT, DTE,
cysteine/cystine, glutathione, cysteamine/cystamine, and EDTA

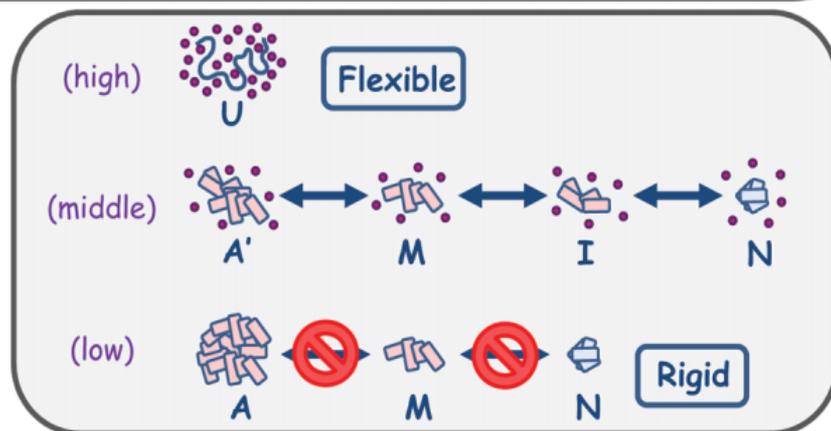
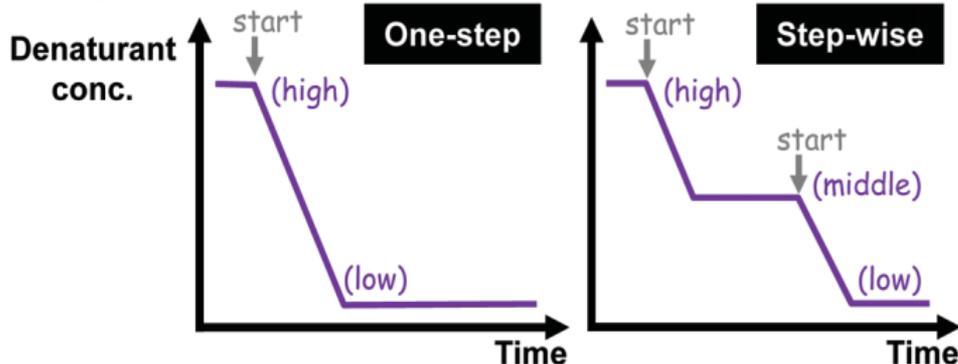
Espressione di proteine ricombinanti per studi strutturali

Metodi di refolding in soluzione:

A. Dilution method



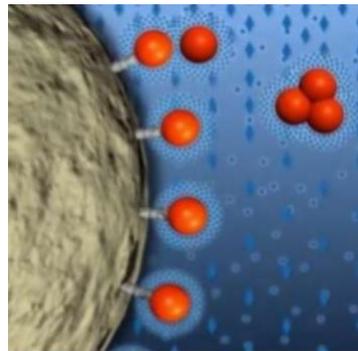
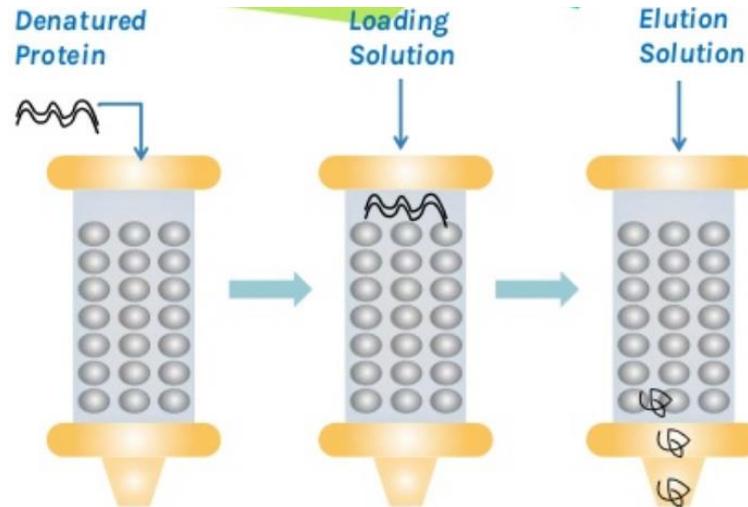
B. Dialysis method



Espressione di proteine ricombinanti per studi strutturali

Metodi di refolding:

C) REFOLDING SU COLONNA

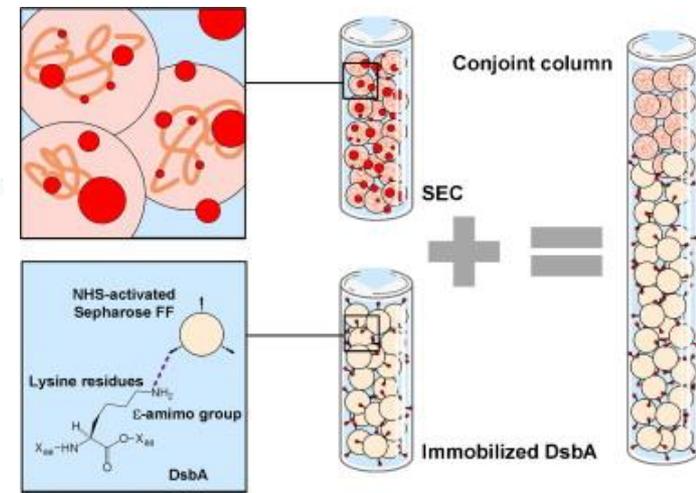
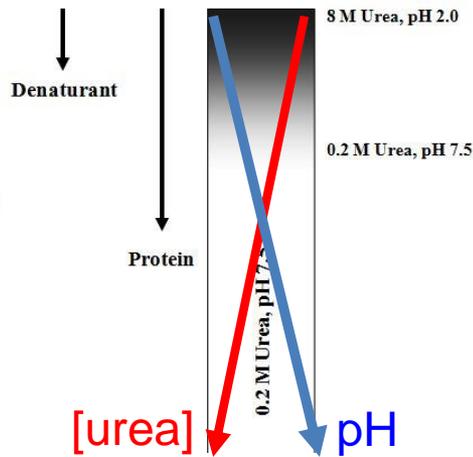
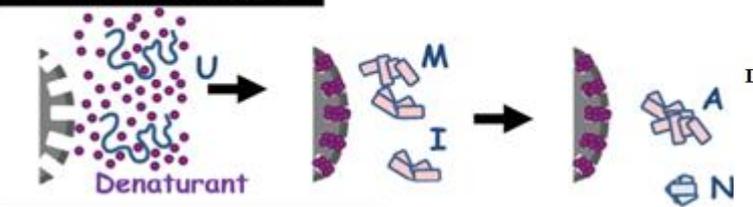


Adsorbimento/Immobilizzazione sulla matrice inibiscono interazioni aspecifiche → riducono l'aggregazione

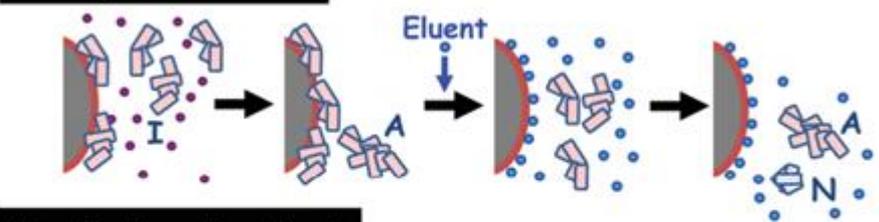
Espressione di proteine ricombinanti per studi strutturali

Refolding su colonna:

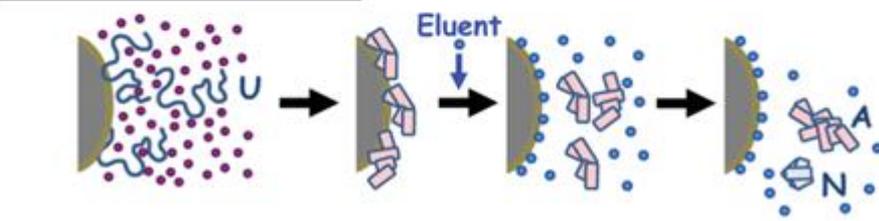
SEC-based method



Adsorption method

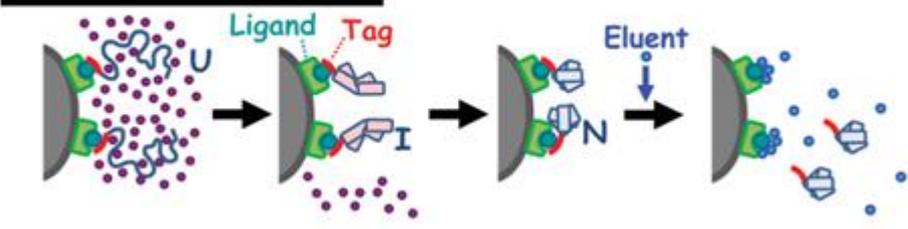


Zeolite-based method



porous, high adsorption capacity, ion exchange properties

Immobilization method



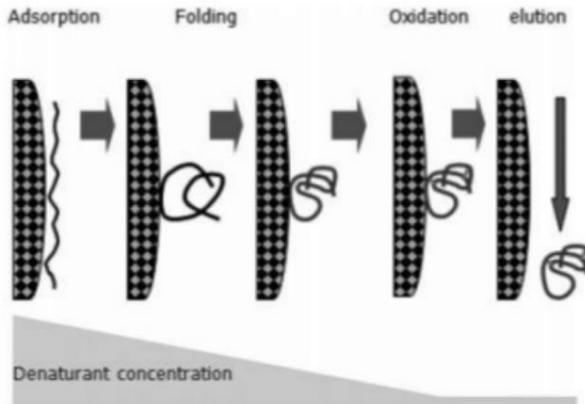
IMAC

Espressione di proteine ricombinanti per studi strutturali

Metodi di refolding:

ADSORBIMENTO

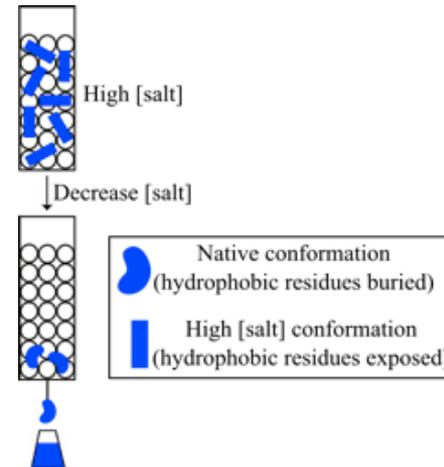
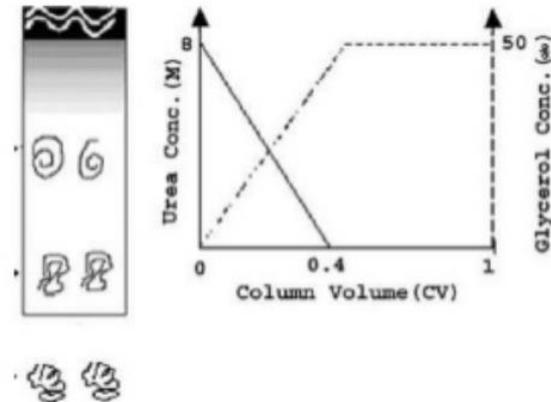
A) IEX



Gradiente duplice:

- 1) pH più elevato del pI (riduce l'aggregazione)
- 2) pH vicino al pI (compatibile con il folding nativo)

B) HIC



Purificazione di proteine ricombinanti per studi strutturali

purificazione

Mass spectrometry

Antigen for immunization

Functional studies

Structural studies

Therapeutic proteins

Purity

Moderate > 80%

High > 95-99%

Very high > 99%

Mass spectrometry

Functional studies

Antigen for immunization

Structural studies

Therapeutic proteins

Amount

pg

ng

μ g

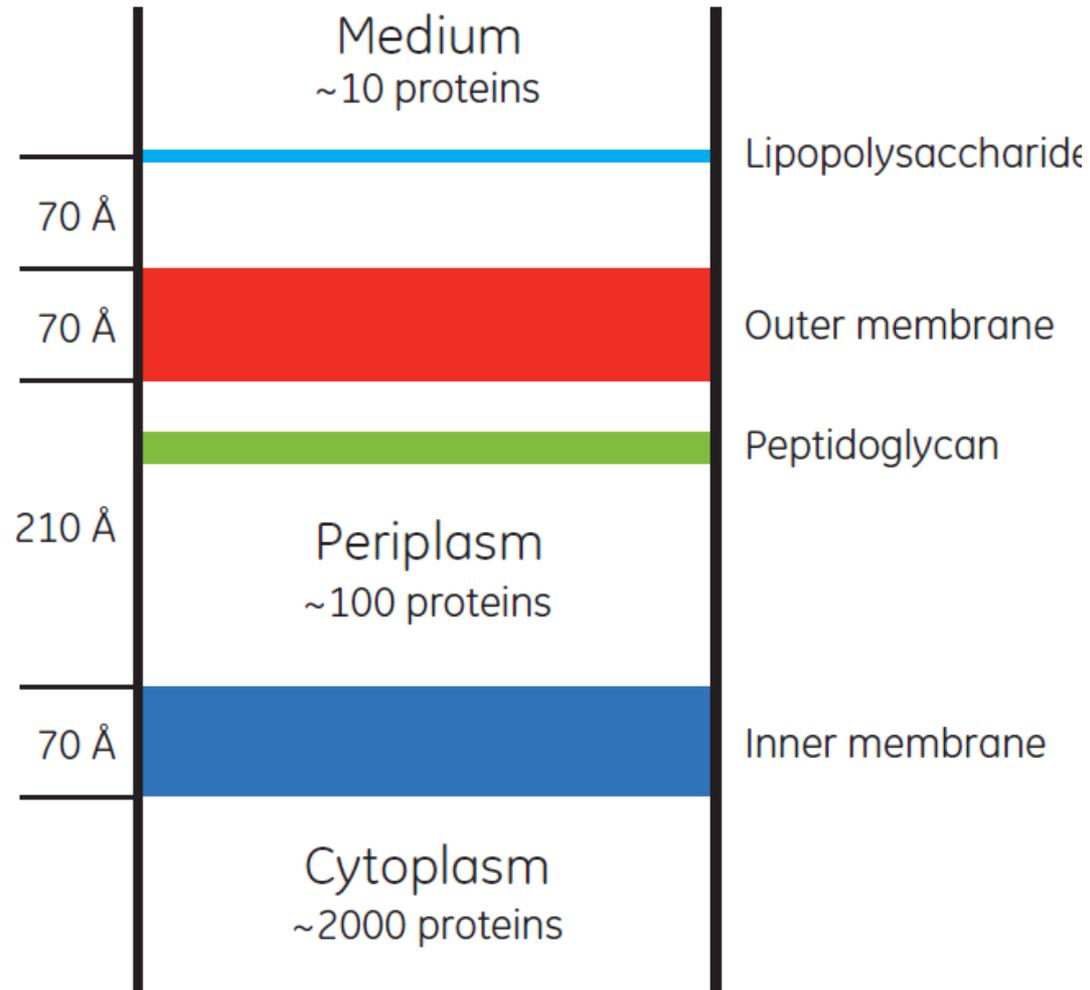
mg

g

kg tonnes

Purificazione di proteine ricombinanti per studi strutturali

purificazione



Purificazione di proteine ricombinanti per studi strutturali

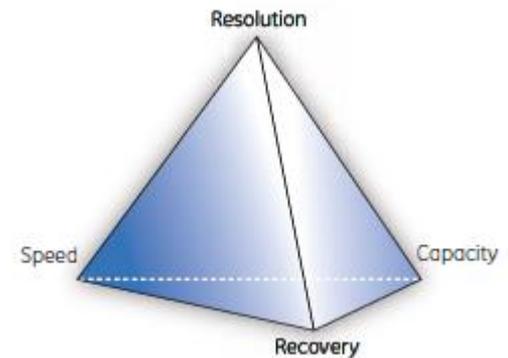
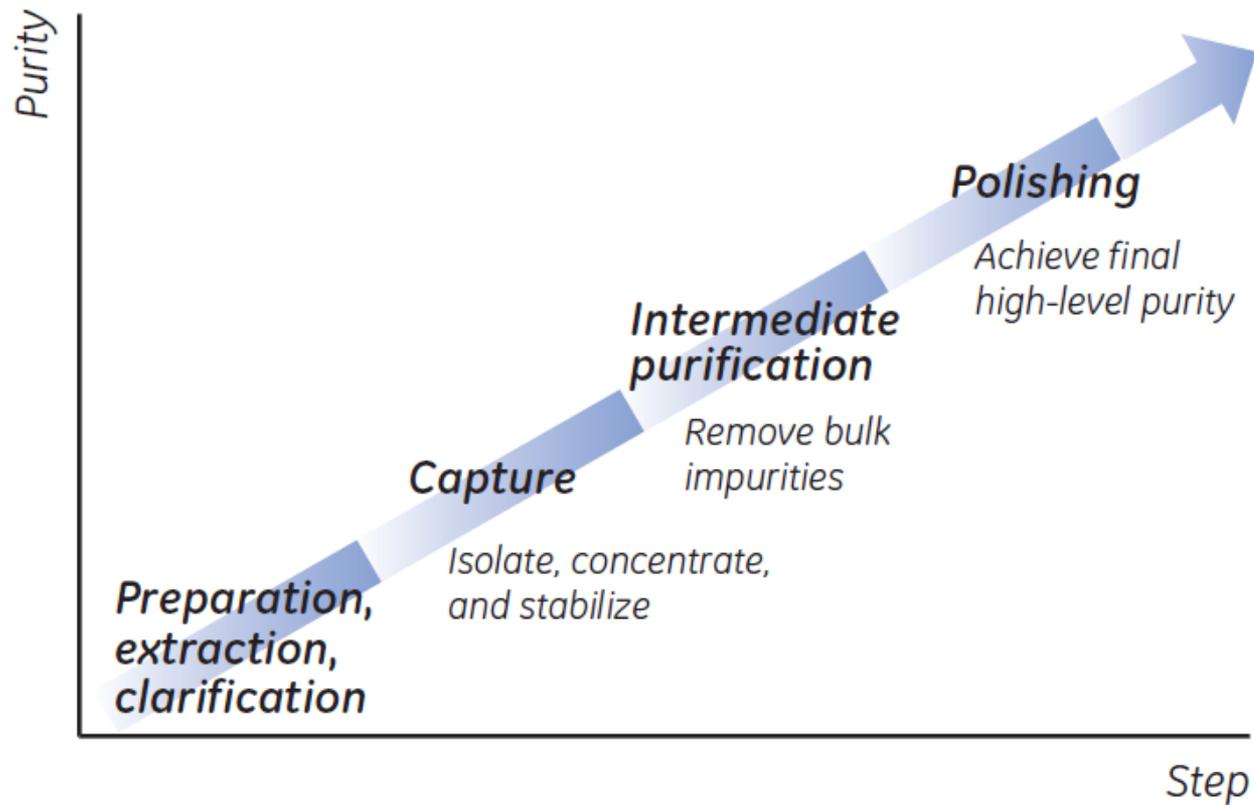
evoluzione dei metodi di purificazione

Table 1. Some milestones in the history of protein purification

Precipitation	1789	Fourcroy
Crystalline ovalbumin	1889	Hofmeister
Chromatography	1903	Tswett*
Ultracentrifugation	1924	Svedberg
Moving boundary electrophoresis	1937	Tiselius
Ion exchange chromatography	1940s	The Manhattan Project
Two-phase partitioning	1955	Albertsson (1)
Size-exclusion chromatography (gel filtration) [†]	1955, 1959	Lindqvist and Storgårds (2), Porath and Flodin (3), Ingelman
Hydroxyapatite chromatography	1956	Tiselius, et al. (4)
Sephadex™ (gel filtration medium)	1959	Pharmacia (now GE Healthcare)
Polyacrylamide gel electrophoresis	1959	Raymond and Weintraub (5)
Isoelectric focusing [‡]	1959	Kolin (6)
High-performance liquid chromatography (HPLC) [§]	1941; 1966	Martin and Synge (7)
Sepharose™	1967	Pharmacia (now GE Healthcare)
SDS-PAGE	1967	Shapiro, et al. (8)
Affinity-ligand coupling chemistry for affinity chromatography [¶]	1967	Axén, et al. (9)
Affinity chromatography	1968	Cuatrecasas, et al. (10)
Reversed-phase chromatography	1970s	Kirkland (11); Molnar and Horvath (12)
Protein A adsorbent for IgG purification	1972	Hjelm, et al. (13); Kronwall (14)
Hydrophobic interaction chromatography	1973	Porath (15); Hjertén (16)
Two-dimensional chromatography	1975	O'Farrell (17)
Immobilized metal ion affinity chromatography	1975	Porath, et al. (18)
Chromatofocusing	1977**	Sluyterman and Wijdenes (19)
Capillary electrophoresis	1981 ^{††}	Jorgenson and Lukacs (20)
Fast Protein Liquid Chromatography	1982	Pharmacia (now GE Healthcare)
Histidine affinity tagging	1988 ^{‡‡}	Smith et al. (21), Hochuli et al. (22)
ÅKTATM design	1996	Pharmacia Biotech (now GE Healthcare)

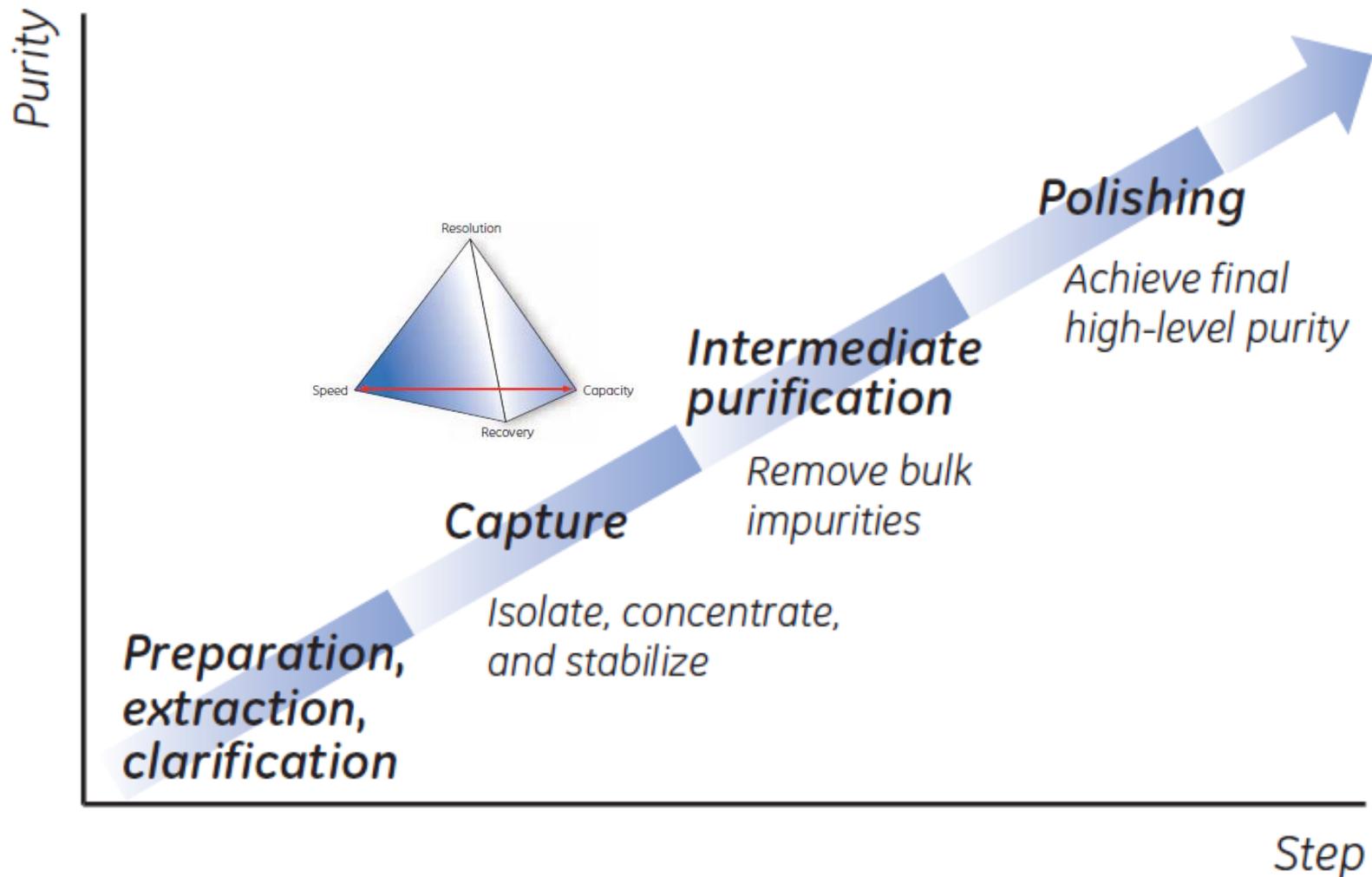
Purificazione di proteine ricombinanti per studi strutturali

strategie di purificazione



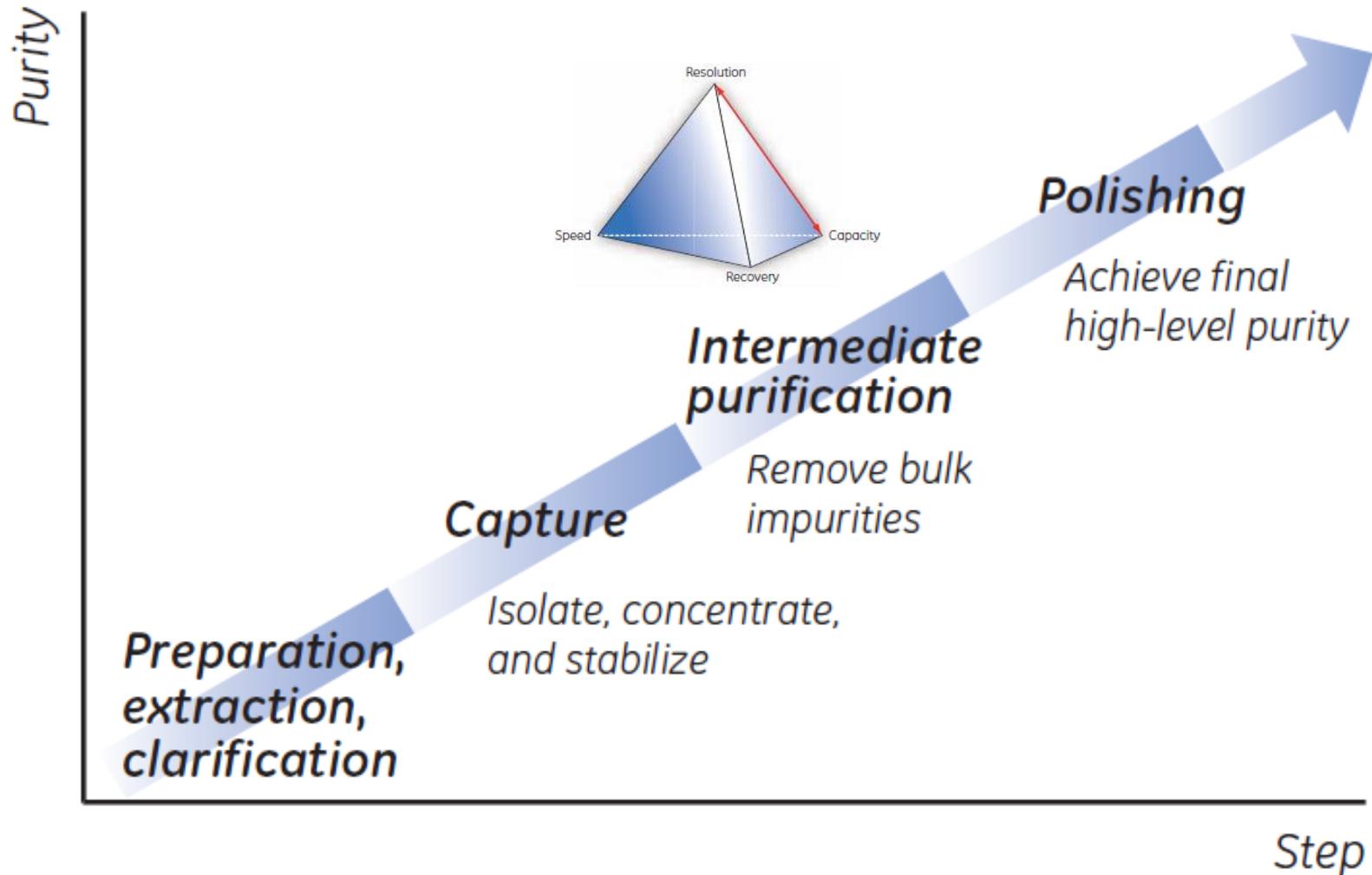
Purificazione di proteine ricombinanti per studi strutturali

strategie di purificazione: cattura



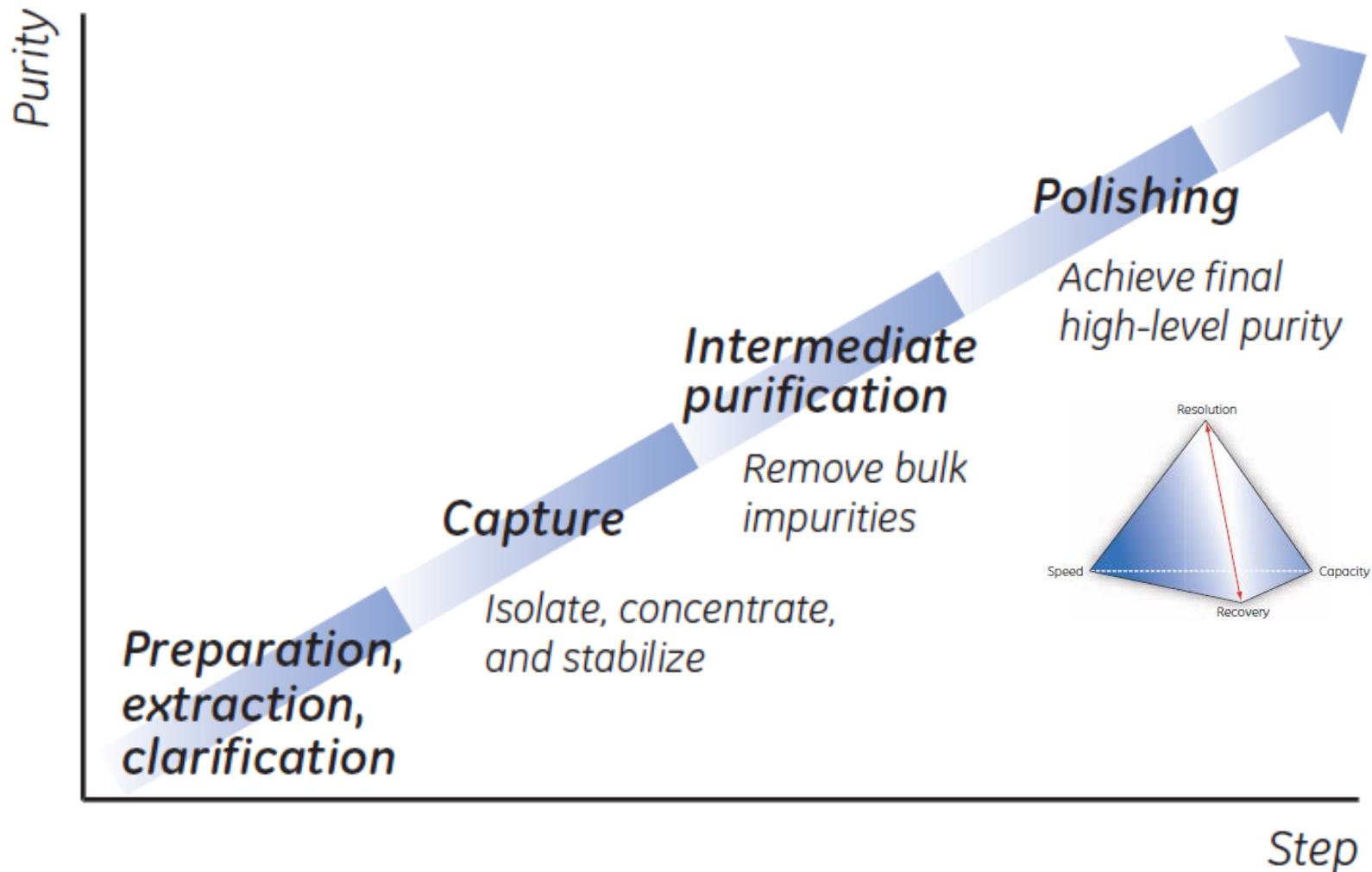
Purificazione di proteine ricombinanti per studi strutturali

strategie di purificazione: purificazione intermedia



Purificazione di proteine ricombinanti per studi strutturali

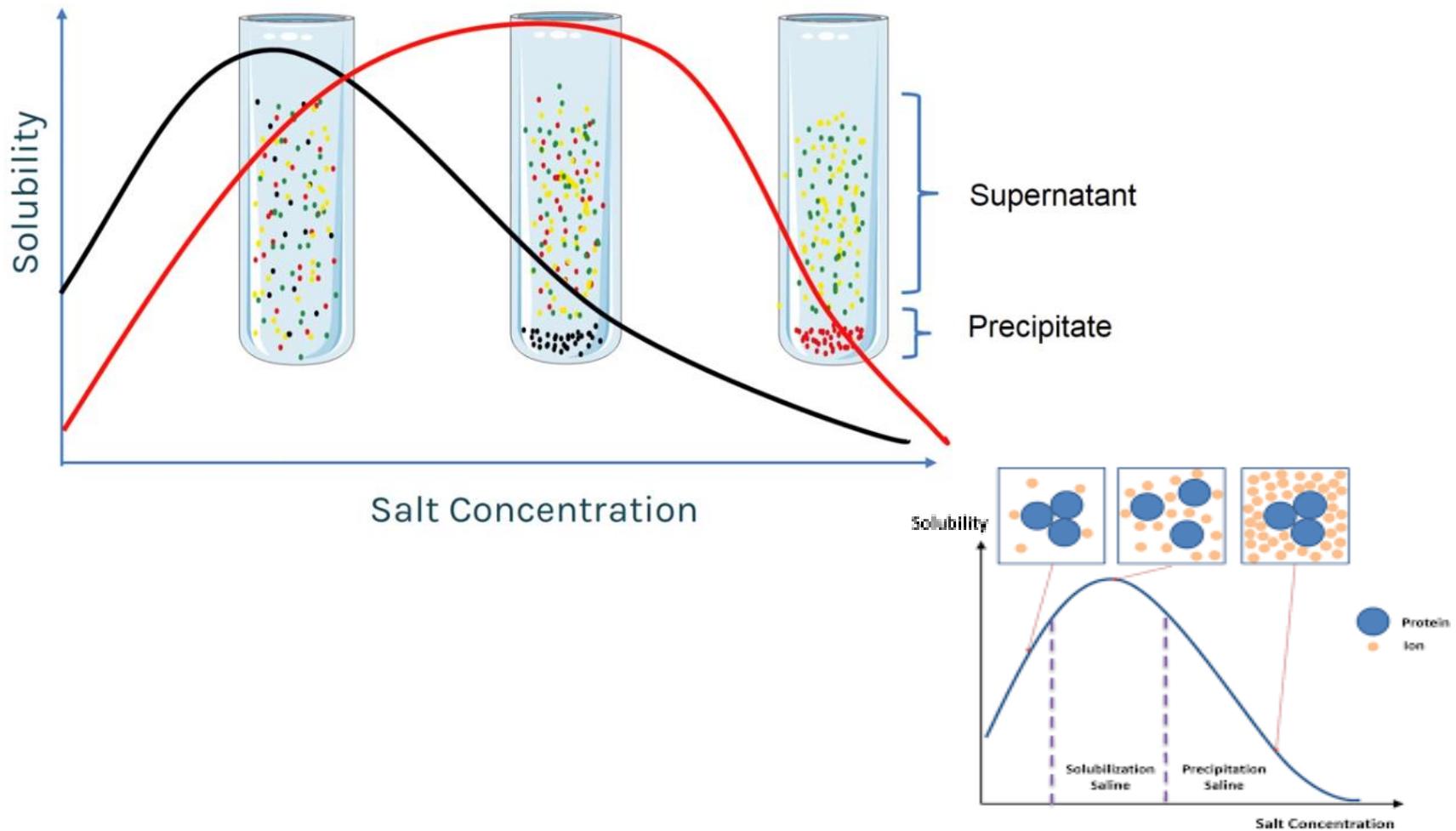
strategie di purificazione: polishing



Purificazione di proteine ricombinanti per studi strutturali

Precipitazione frazionata:

separazione sulla base della solubilità (effetto del salting out)



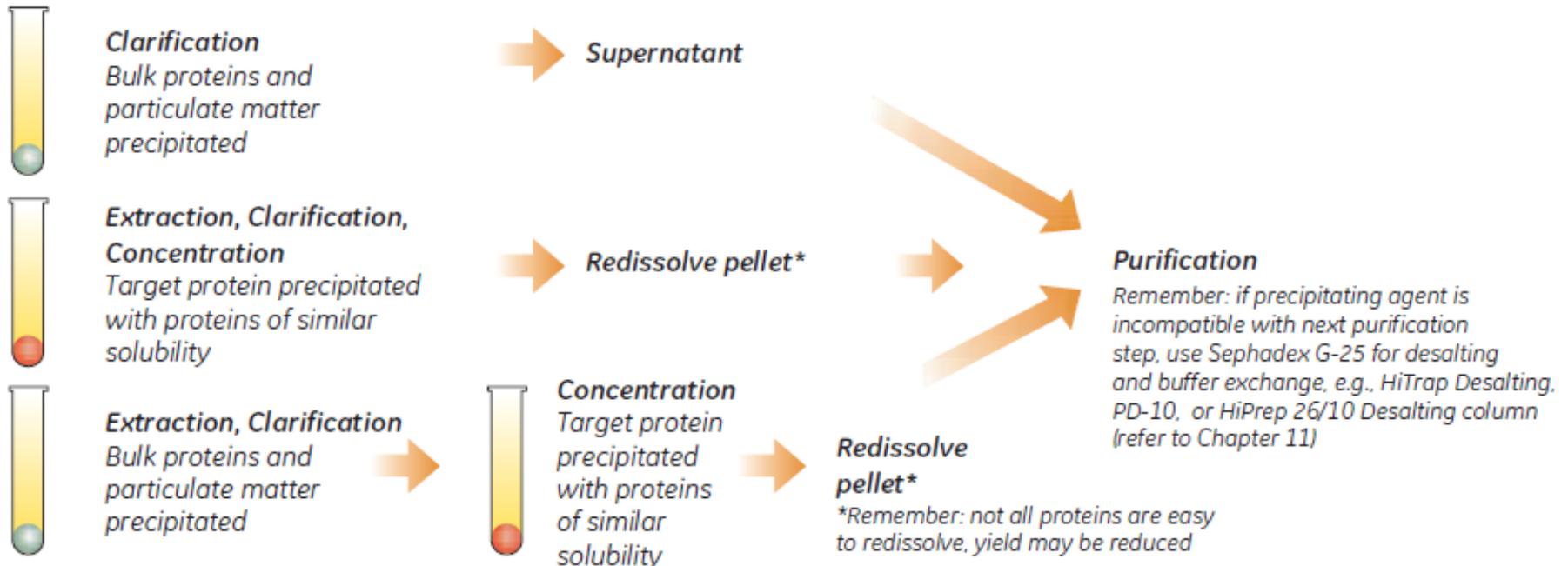
Purificazione di proteine ricombinanti per studi strutturali

Precipitazione frazionata

Precipitation agent	Typical conditions	Sample type	Comment
Ammonium sulfate	As described later in this chapter.	> 1 mg/ml proteins, especially immunoglobulins.	Stabilizes proteins, no denaturation. Supernatant can go directly to HIC. Helps to reduce lipid content.
Dextran sulfate	Add 0.04 ml of 10% dextran sulfate and 1 ml of 1 M CaCl ₂ per ml of sample, mix 15 min, centrifuge at 10 000 × g, discard pellet.	Samples with high levels of lipoproteins, e.g., ascites.	Precipitates lipoproteins.
Polyvinylpyrrolidone	Add 3% (w/v), stir 4 h, centrifuge at 17 000 × g, discard pellet.	Samples with high levels of lipoprotein, e.g., ascites.	Alternative to dextran sulfate.
Polyethylene glycol (PEG, M _r > 4000)	Up to 20% (w/v).	Plasma proteins.	No denaturation, supernatant goes directly to IEX or AC. Complete removal of PEG may be difficult. Stabilizes proteins.
Acetone (cold)	Up to 80% (v/v) at 0°C. Collect pellet after centrifugation at full speed in chilled microcentrifuge.	Useful for peptide precipitation or concentration of sample for electrophoresis.	May denature protein irreversibly.
Polyethyleneimine	0.1% (w/v).		Precipitates aggregated nucleoproteins.
Protamine sulfate	1% (w/v).		Precipitates aggregated nucleoproteins.
Streptomycin sulfate	1% (w/v).		Precipitates nucleic acids.
Caprylic acid	1:15 (w/w).	Antibody concentration should be > 1 mg/ml.	Precipitates bulk of proteins from sera or ascites, leaving immunoglobulins in solution.

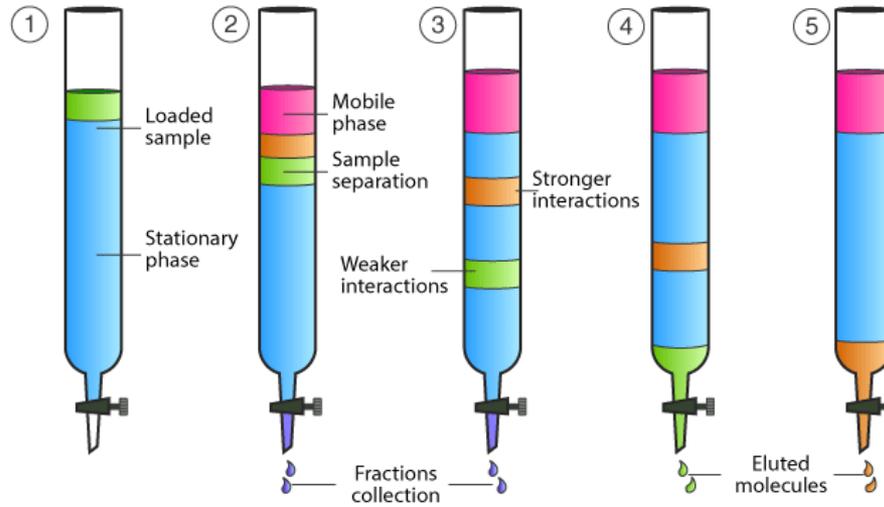
Purificazione di proteine ricombinanti per studi strutturali

Precipitazione frazionata



Purificazione di proteine ricombinanti per studi strutturali

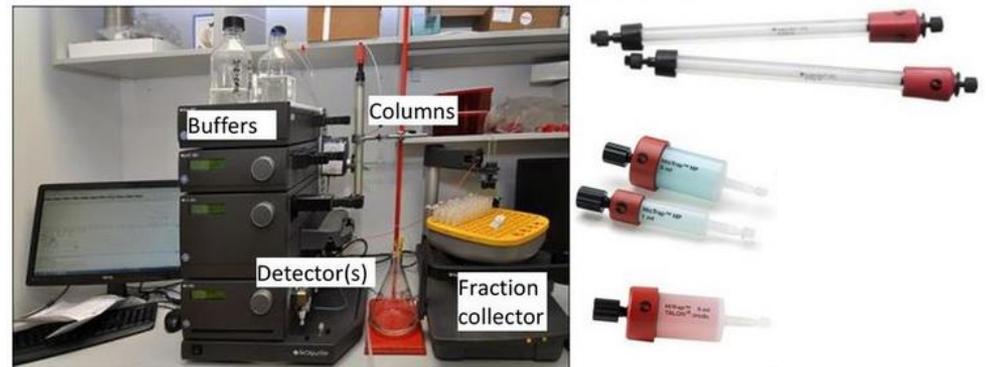
Metodi cromatografici:



IN BATCH



CON FPLC (Fast Protein Liquid Chromatography)

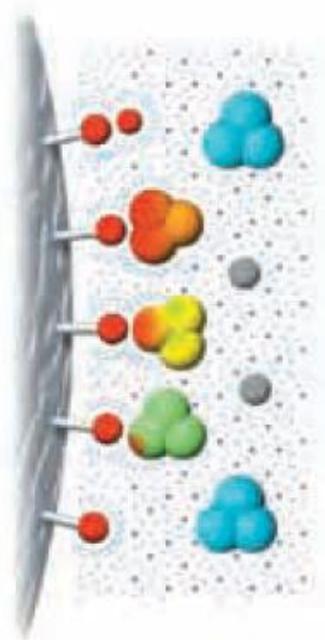


Purificazione di proteine ricombinanti per studi strutturali

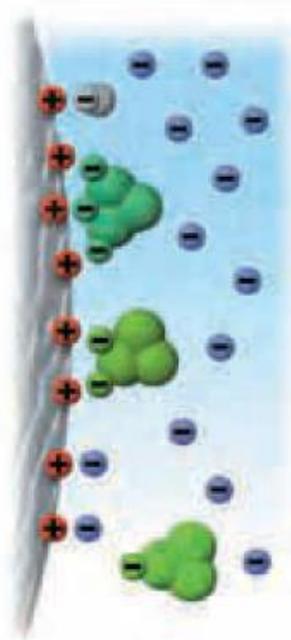
Metodi cromatografici:



Gel filtration



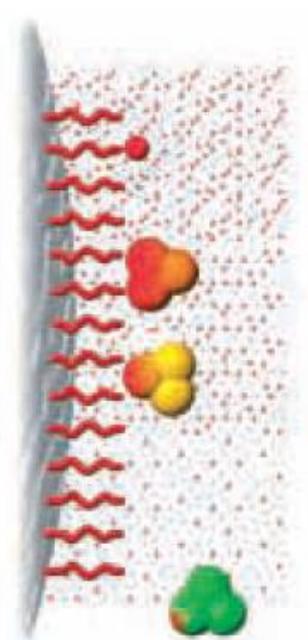
Hydrophobic interaction



Ion exchange



Affinity

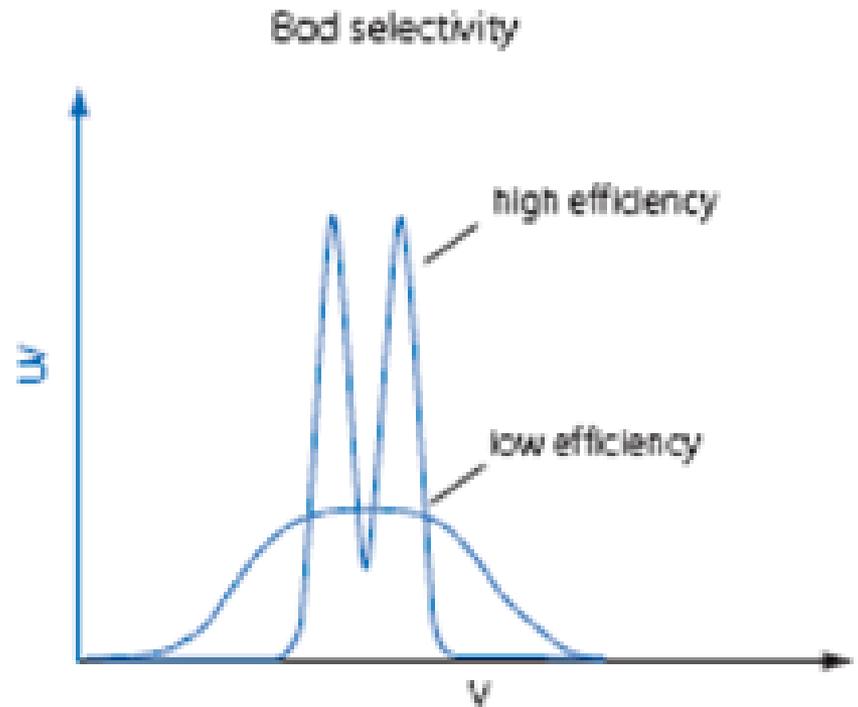
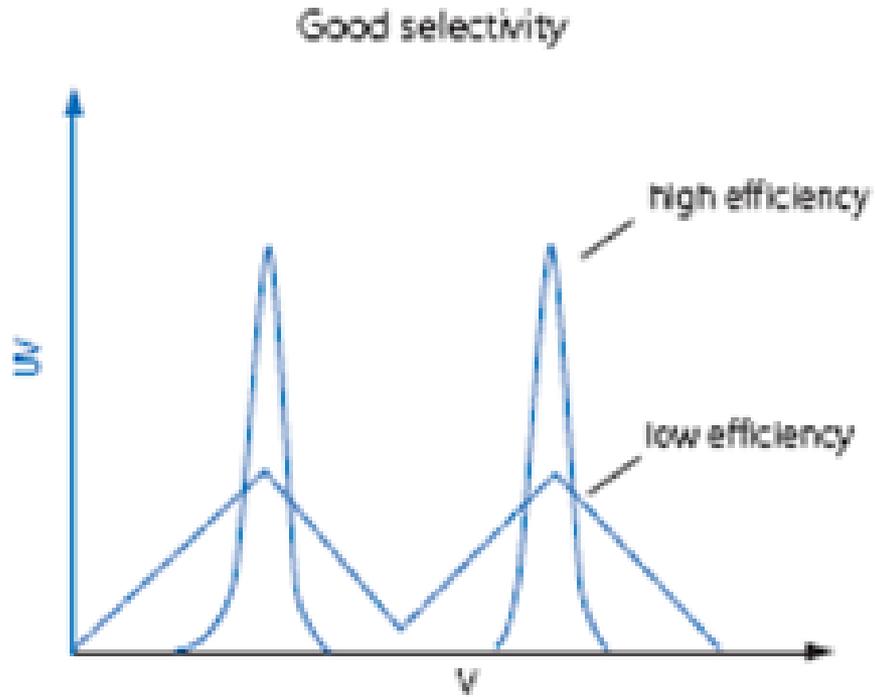


Reversed phase

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

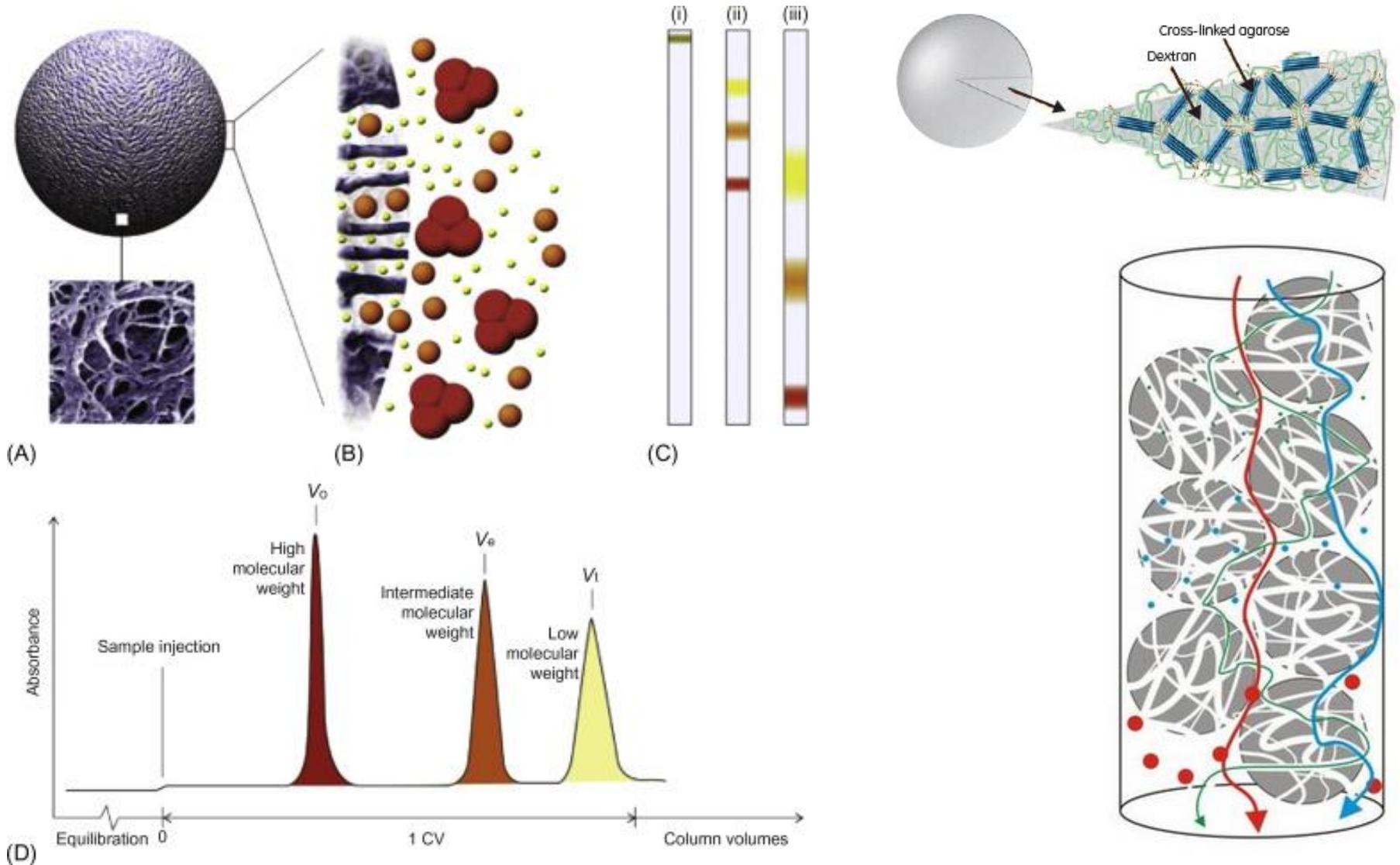
Metodi cromatografici : selettività vs efficienza



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

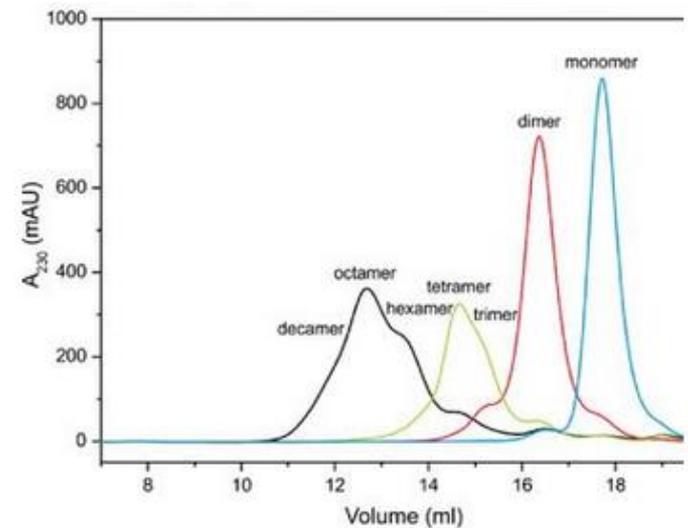
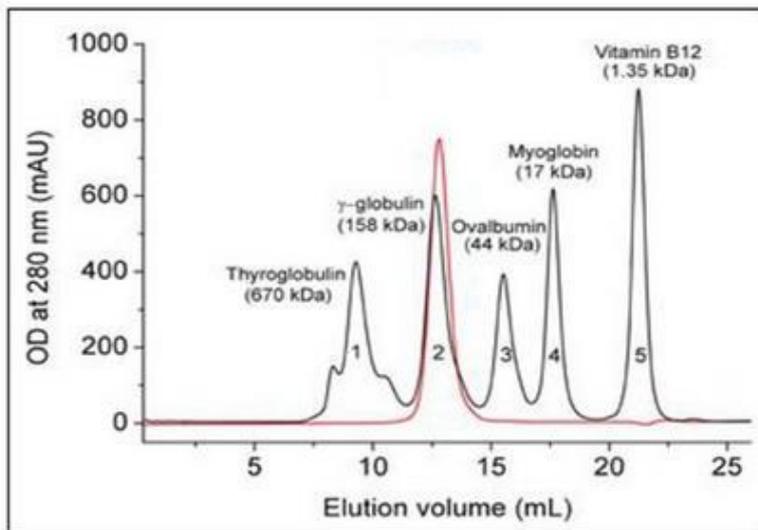
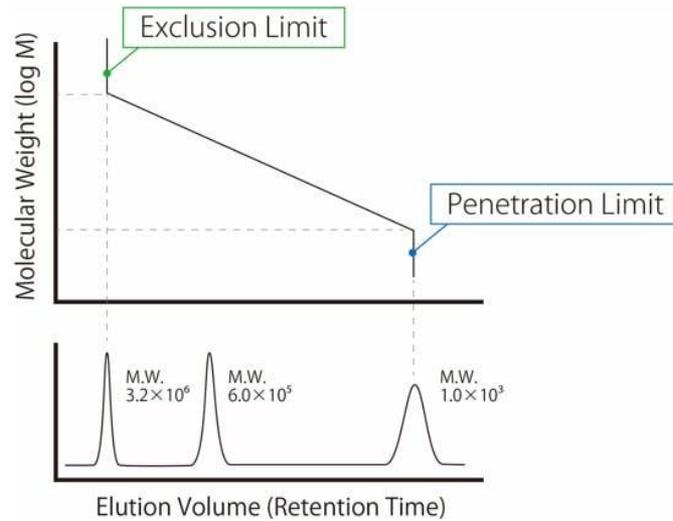
SEC



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

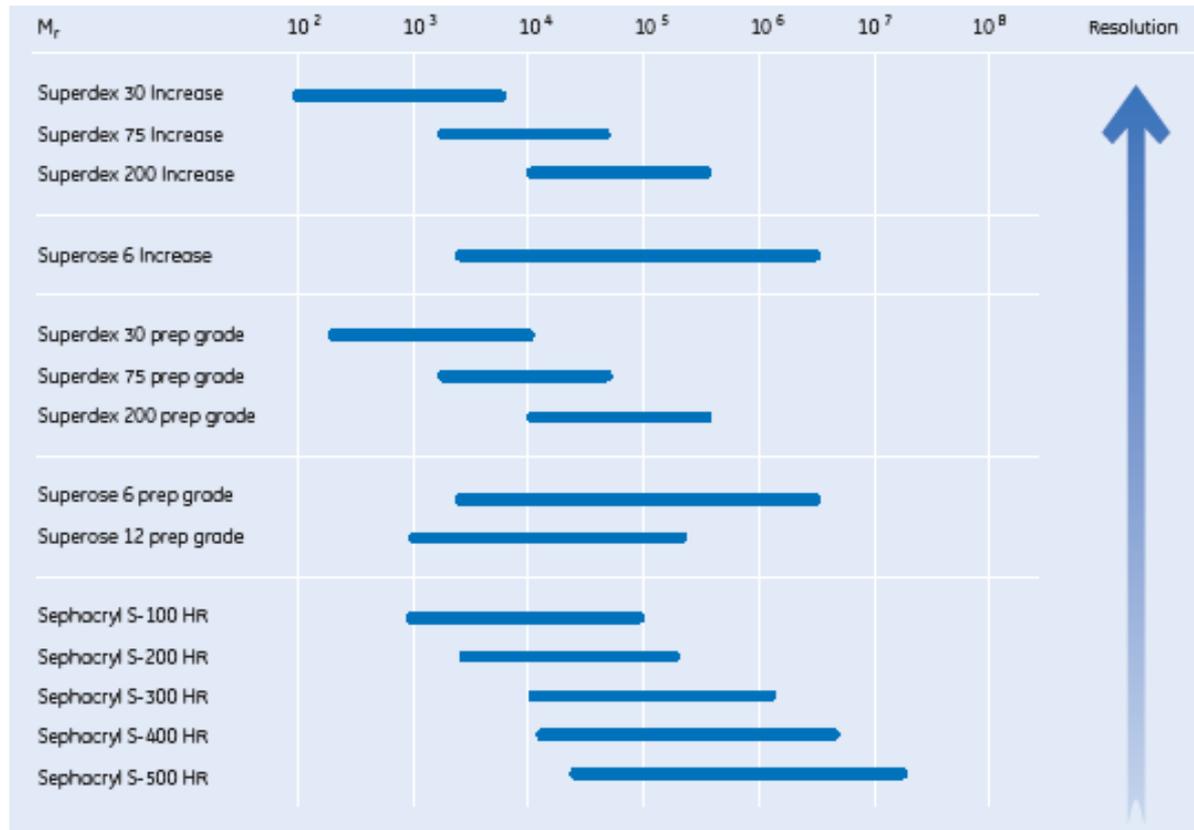
SEC



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

SEC

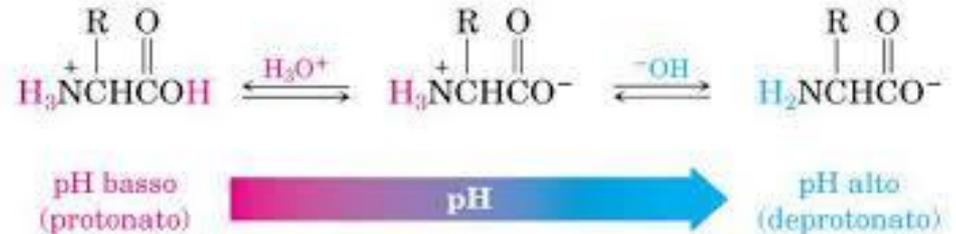
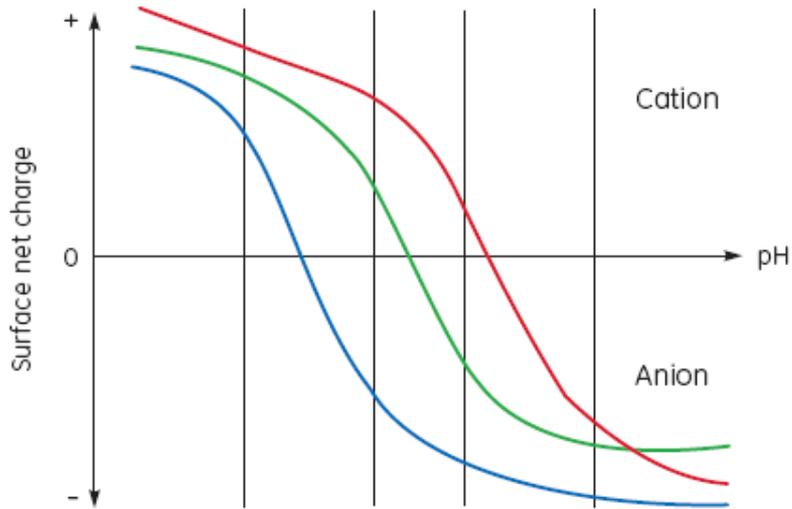


Product	Efficiency ¹	Operational ²	Cleaning-in-place (CIP) ³	Particle size, d_{50} (μm) ⁴
Superdex 30 Increase	> 43 000	3 to 12	1 to 14	~ 9
Superdex 75 Increase	> 43 000	3 to 12	1 to 14	~ 9
Superdex 200 Increase	> 48 000	3 to 12	1 to 14	~ 8.6
Superdex 30 prep grade	\geq 13 000	3 to 12	1 to 14	~ 34
Superdex 75 prep grade	\geq 13 000	3 to 12	1 to 14	~ 34
Superdex 200 prep grade	\geq 13 000	3 to 12	1 to 14	~ 34

Purificazione di proteine ricombinanti per studi strutturali

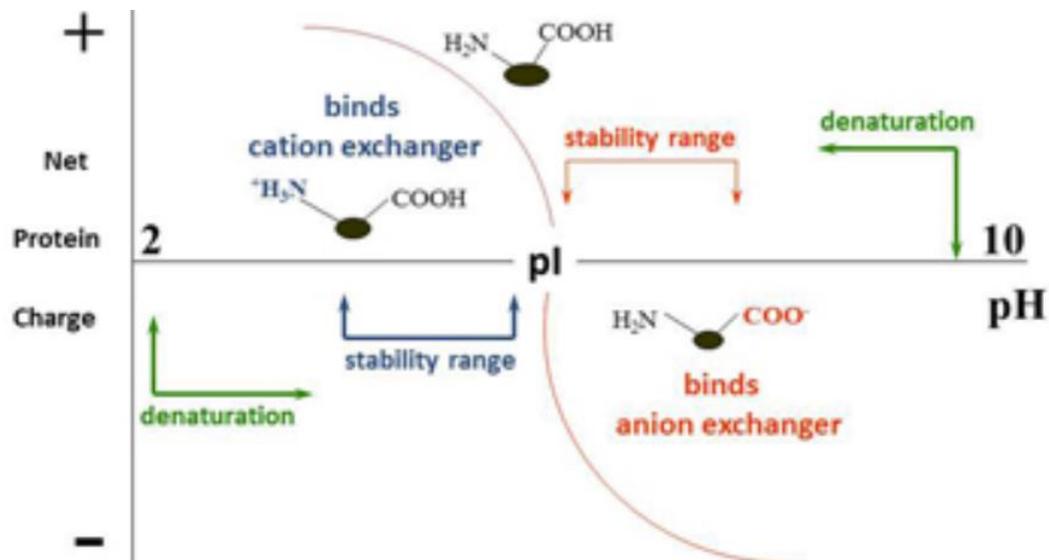
Metodi cromatografici

IEX



Punto isoelettrico (zwitterione neutro)

(pI 5.5-7.5)



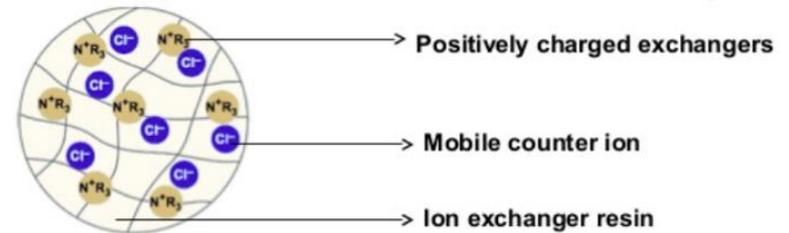
Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

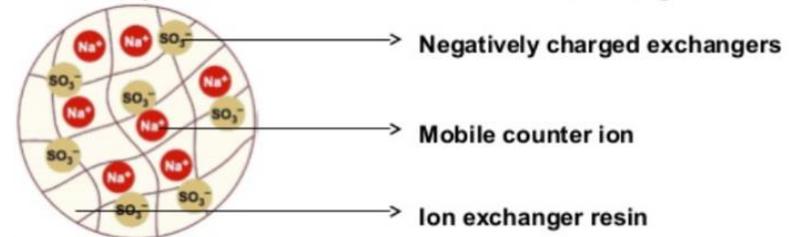
IEX

Resin	Functional group	Weak or Strong	Functional pH Range
DEAE	diethylaminoethyl [-N ⁺ (C ₂ H ₅) ₂ H ⁺]	weak anion	pH 2 - 9
ANX	diethylaminopropyl [-N ⁺ (C ₂ H ₅) ₂ H ⁺]	weak anion	pH 2 - 9
Q	quarternary amine [-N ⁺ (CH ₃) ₃]	strong anion	pH 1 - 14
CM	carboxymethyl [-O-CH ₂ COO ⁻]	weak cation	pH 5 - 10
S	methyl sulfonate [O-CH ₂ -CHOH-CH ₂ -O- CH ₂ -CHOH-CH ₂ -SO ₃ ⁻]	strong cation	pH 2 - 12
SP	sulphonyl group [-CH ₂ -CH ₂ -CH ₂ SO ₃ ⁻]	strong cation	pH 2 - 14

Tertiary amines → Strong anion exchangers
 Secondary amines → Weak anion exchangers



Sulphonic acid → Strong cation exchangers
 Carboxylic acid → Weak cation exchangers

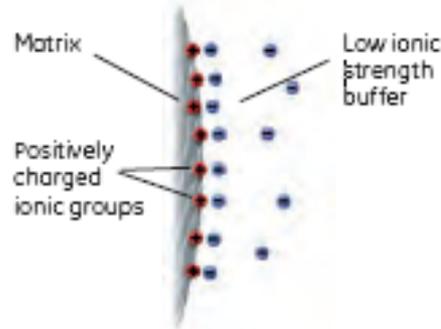


- selettività diversa
- capacità varia con il pH per i deboli

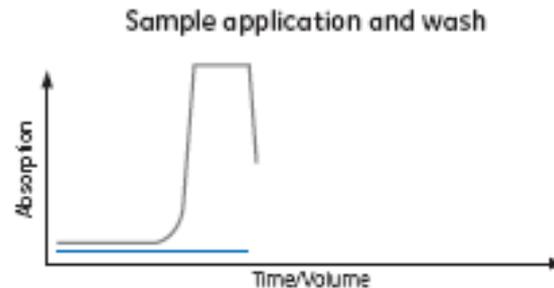
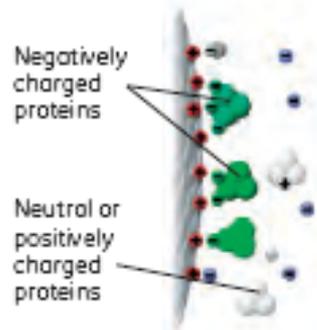
Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

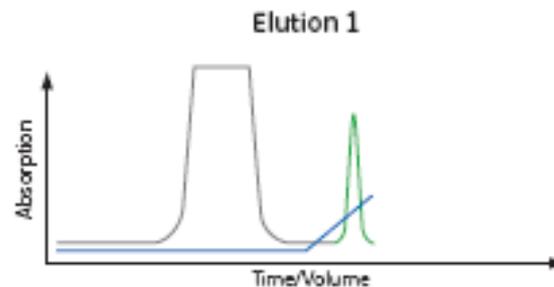
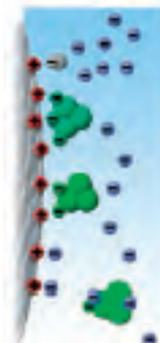
IEX: fasi della purificazione



IEX medium equilibrated with start buffer.



Oppositely charged proteins bind to ionic groups of the IEX medium, becoming concentrated on the column. Uncharged proteins, or those with the same charge as the ionic groups, elute during sample application or just after, during the wash with start buffer.

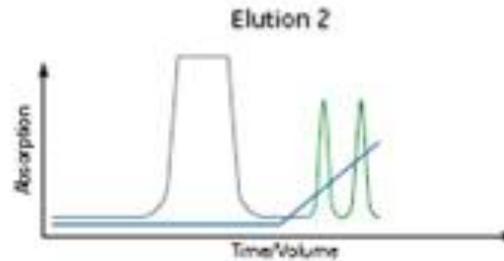
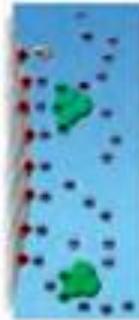


Increasing ionic strength [using a gradient] displaces bound proteins as ions in the buffer compete for binding sites.

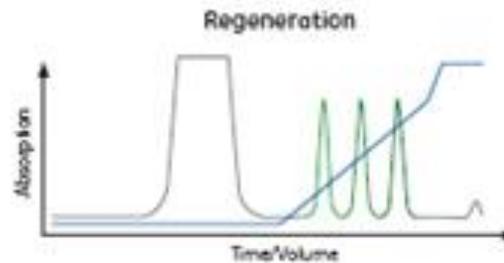
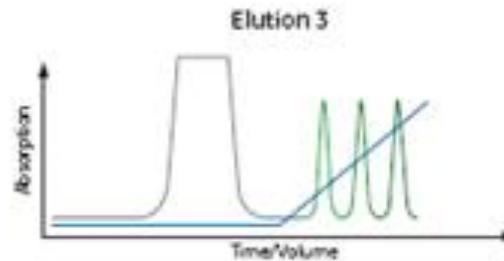
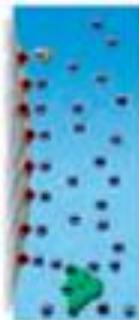
Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

IEX: fasi della purificazione



Further increases in ionic strength displace proteins that are more highly charged (more tightly bound)



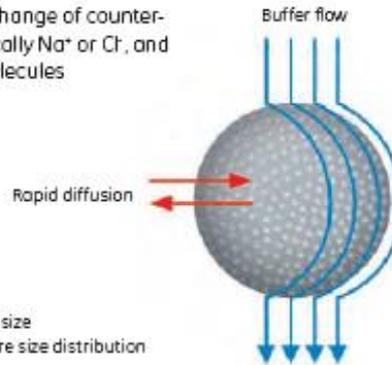
Final high ionic strength wash removes any ionically bound proteins before re-equilibration

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

IEX: dimensioni delle componenti di matrice

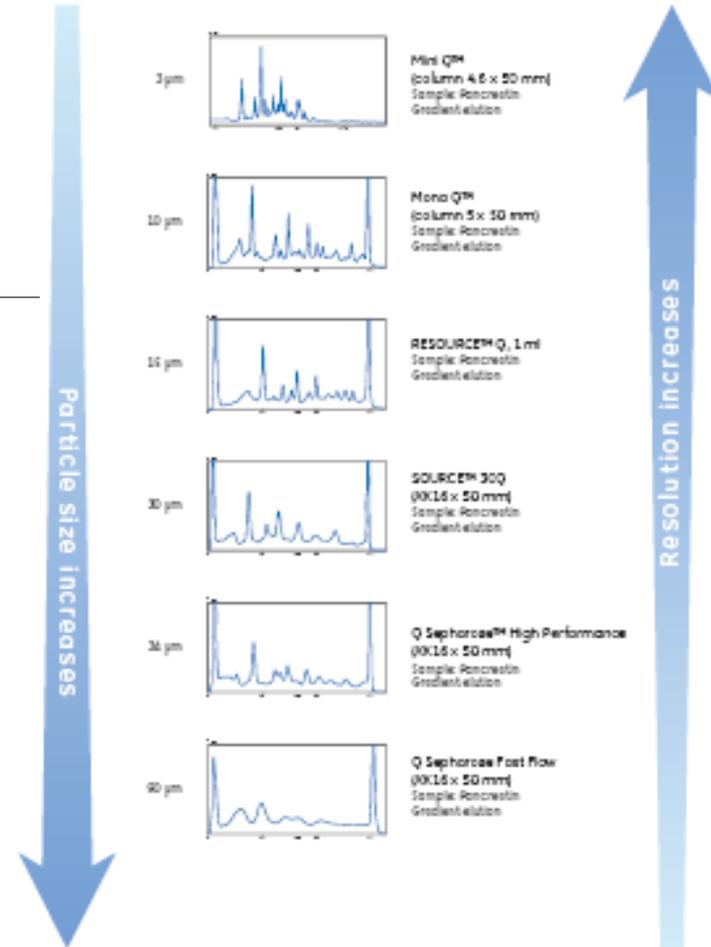
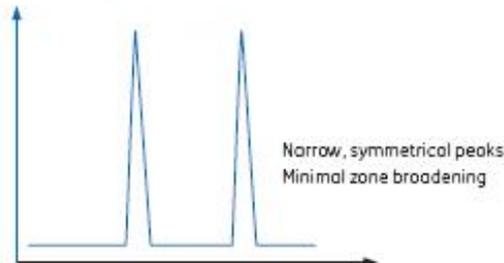
Rapid exchange of counter-ions, typically Na⁺ or Cl⁻, and solute molecules



	Form	Mean particle size (µm)
MiniBeads™	Polystyrene/divinyl benzene	3
MonoBeads™	Polystyrene/divinyl benzene	10
SOURCE 15	Polystyrene/divinyl benzene	15
SOURCE 30	Polystyrene/divinyl benzene	30
Sepharose High Performance	Agarose 6%	34
Sepharose Fast Flow	Agarose 6%	90
Sepharose 4 Fast Flow	Agarose 4%	90
Sepharose XL	Agarose 6%, dextran chains coupled to agarose	90
Sepharose Big Beads	Agarose 6%	200
Capto™ ImpRes	High-flow agarose	40
Capto ImpAct	High-flow agarose	50
Capto	High-flow agarose	90

Small bead size
Uniform pore size distribution

Even buffer flow distribution
Uniform packing
Narrow particle size distribution

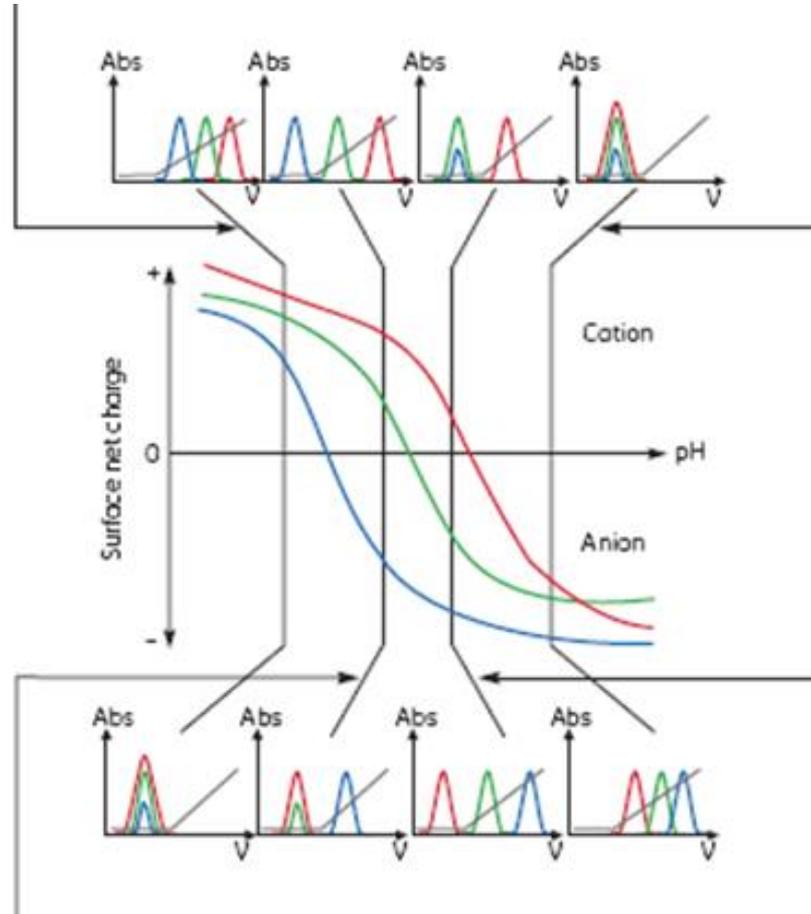


Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

IEX

Most acidic pH: all three proteins are below their pI , positively charged, and bind only to a cation exchanger. Proteins are eluted in the order of their net charge.



Most alkaline pH: all three proteins are above their pI , negatively charged, and bind only to the anion exchanger. Proteins are eluted in the order of their net charge.

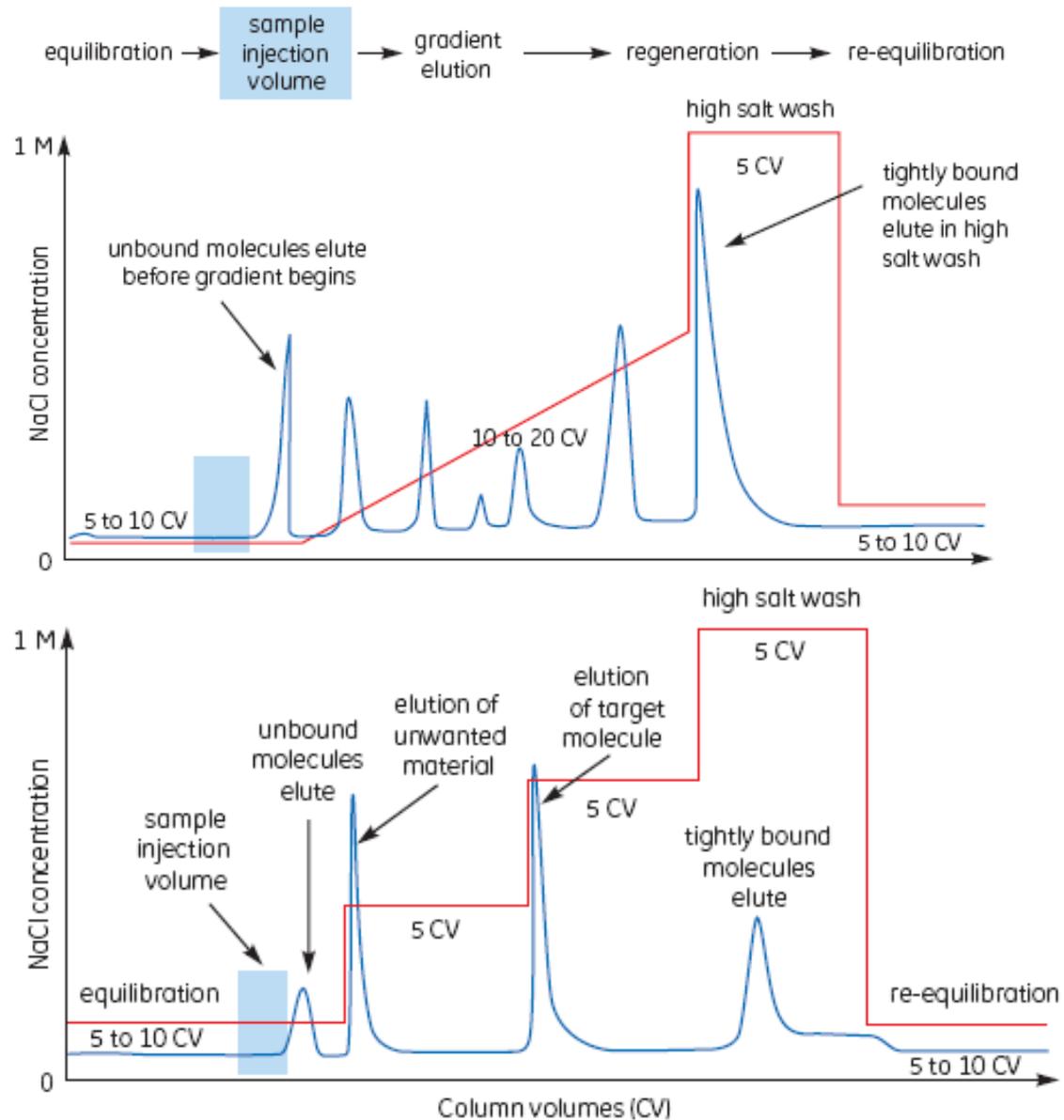
Less acidic pH: blue protein is above its pI , negatively charged, other proteins are still positively charged. Blue protein binds to an anion exchanger and can be separated from the other proteins which wash through. Alternatively, red and green proteins can be separated on a cation exchanger and the blue protein washes through.

Less alkali pH: red protein below its pI , positively charged. Red protein binds to cation exchanger and can be separated from the other proteins which wash through. Alternatively, blue and green proteins can be separated on an anion exchanger and the red protein washes through.

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici

IEX: tipi di eluizione



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

IEX: CAPTURE

Capture

Isolate, concentrate, and stabilize target protein(s)
Sample conditions:
clarified or nonclarified

Easy scale-up, Broad choice of selectivity, including alternatives to Q or S IEX media
Sepharose Fast Flow (Q, SP, DEAE, CM, ANX)

Try weak ion exchangers such as DEAE, CM, or ANX if the selectivity of Q or S is unsatisfactory.

High volume throughput and high capacity, Easy scale-up
Capto (Q, S, DEAE, adhere, MMC)

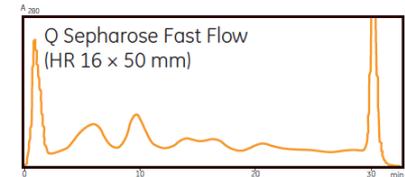
Use high bed heights for increased productivity.
Use Capto MMC for high salt feed.

High binding capacity for selected proteins, Easy scale-up
Sepharose XL (Q or SP)

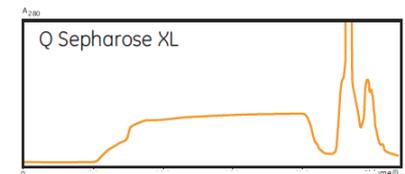
Use for capture with high binding capacity/rapid separation of proteins from clarified samples

Large scale, viscous samples
Sepharose Big Beads (Q or SP)

Use with step elution.



Sample:
Pancreatin
Gradient elution



Sample:
Recombinant α -amylase
Pilot scale:
Gradient elution begins after 20 l

R
I
S
O
L
U
Z
I
O
N
E

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

IEX: INTERMEDIATE PURIFICATION

Intermediate purification

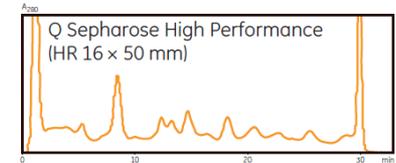
Remove bulk impurities
Sample condition:
partially purified

High resolution
Easy scale-up
Sephacrose High Performance (Q or SP)

Use HiTrap columns
prepacked with
Sephacrose High
Performance, Sepharose
XL and Sepharose Fast
Flow for media selection
and pH scouting.

High resolution and throughput
Flexibility of process design
Capto ImpRes (Q or SP)
Capto S ImpAct

Use Capto ImpRes and
Capto S ImpAct for
increased productivity in
large-scale production



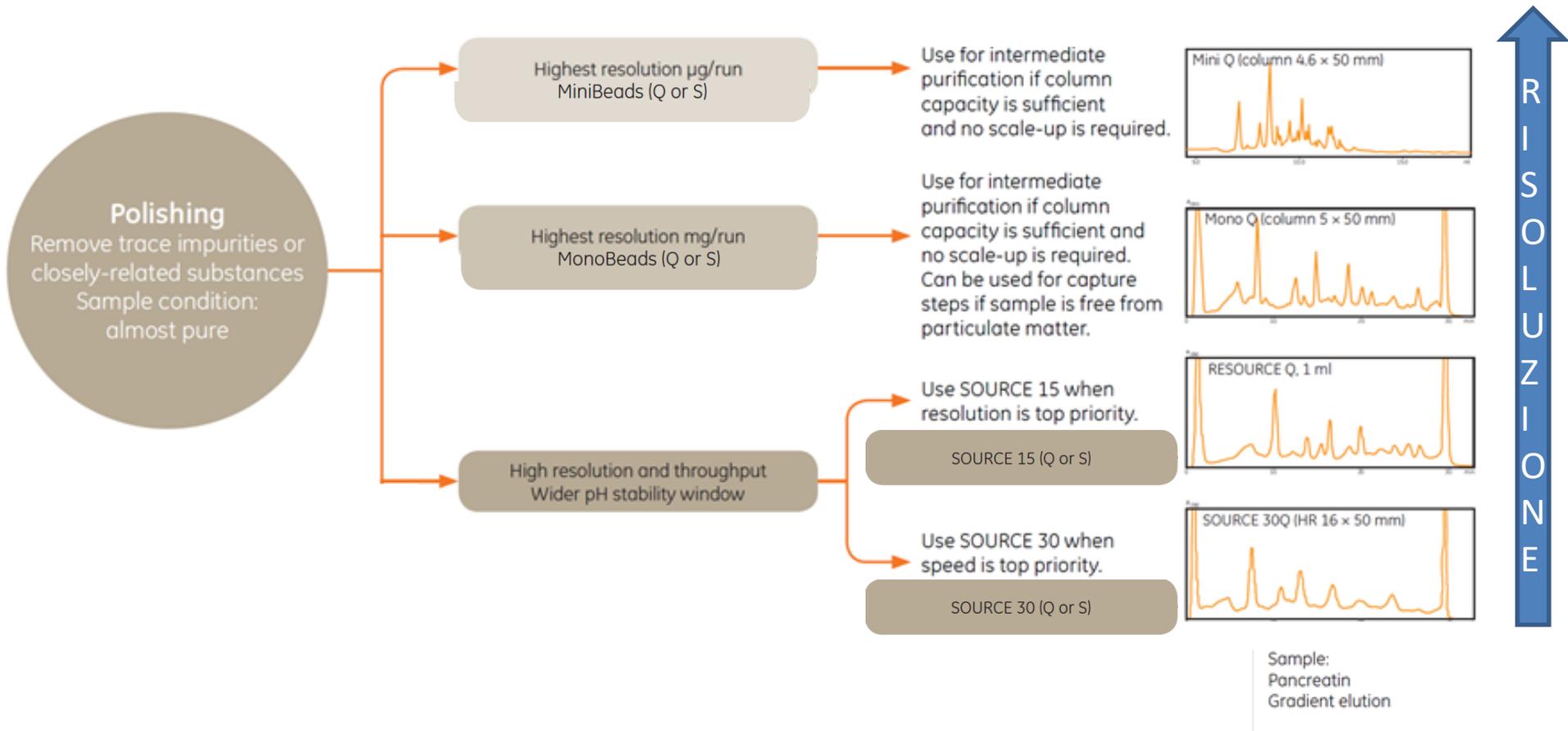
Sample:
Pancreatin
Gradient elution

R
I
S
O
L
U
Z
I
O
N
E

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

IEX: POLISHING

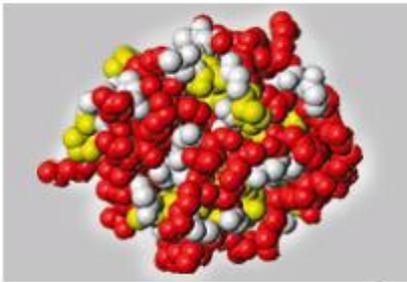


Purificazione di proteine ricombinanti per studi strutturali

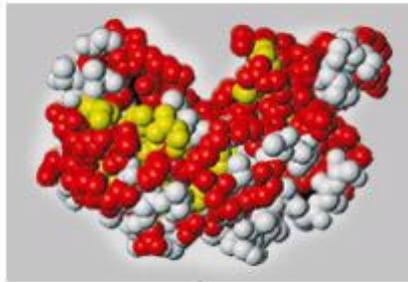
Metodi cromatografici:

HIC

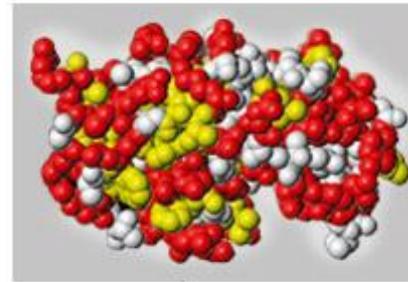
Cytochrome c



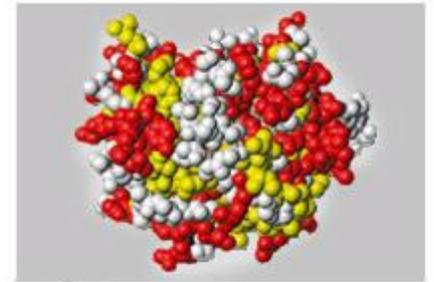
RNase A



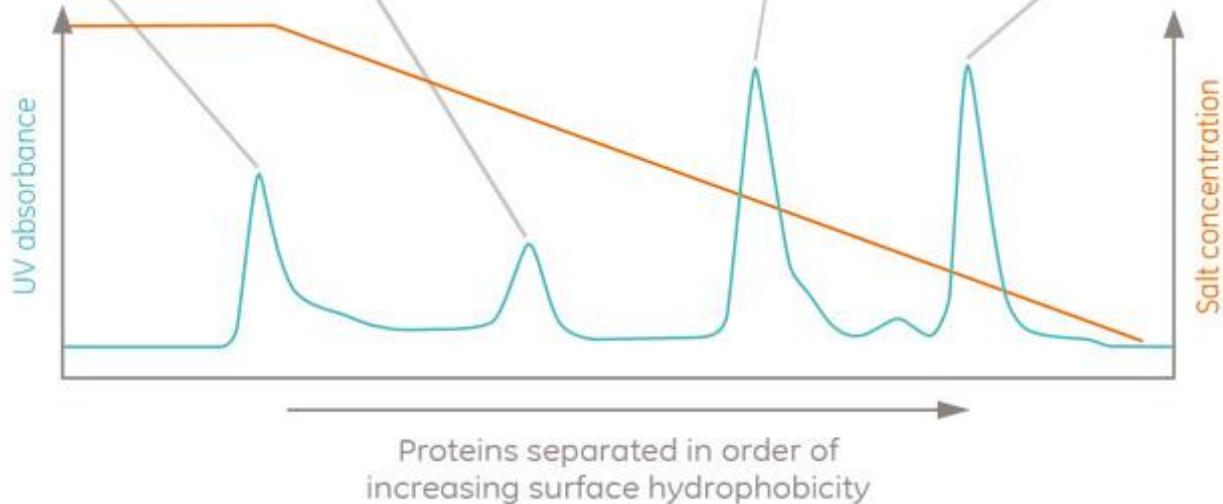
Lysozyme



α -chymotrypsin



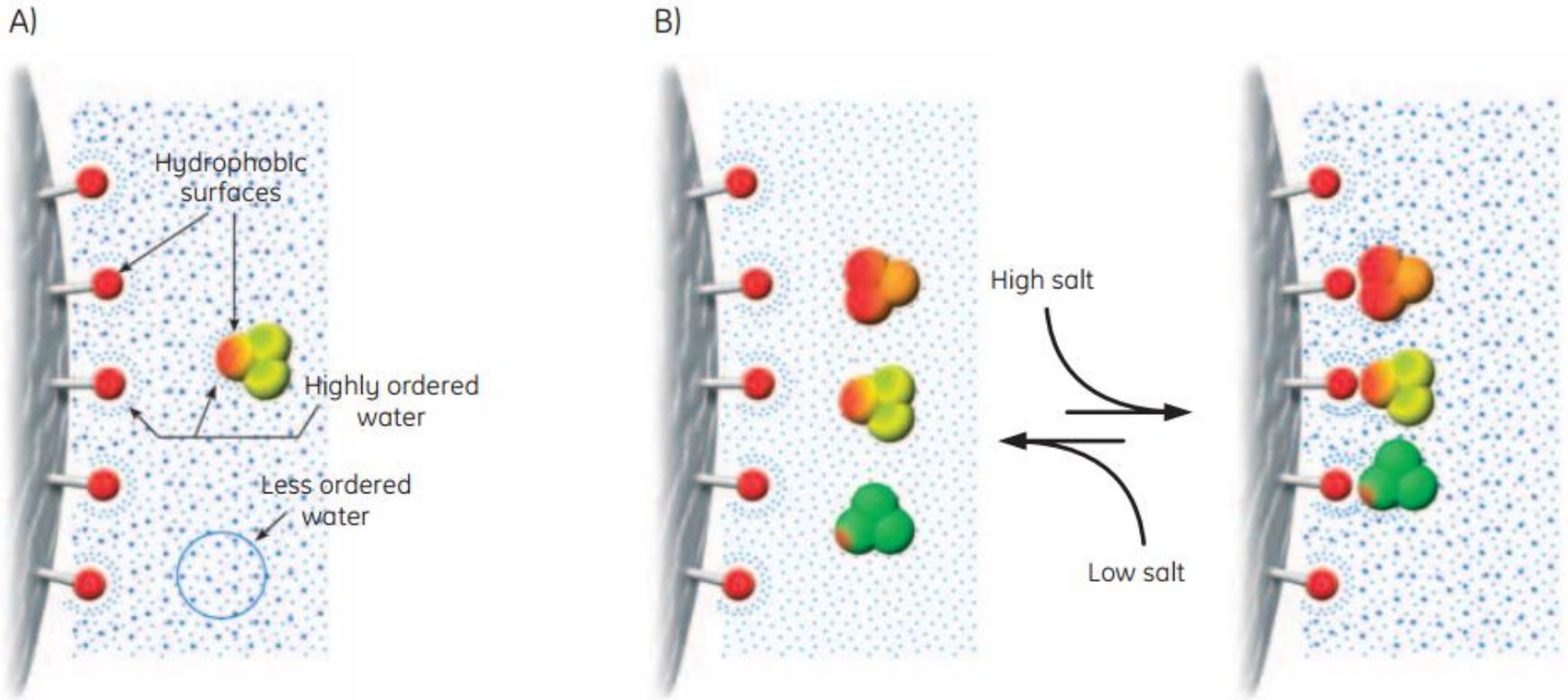
Yellow :
hydrophobic residues



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

HIC

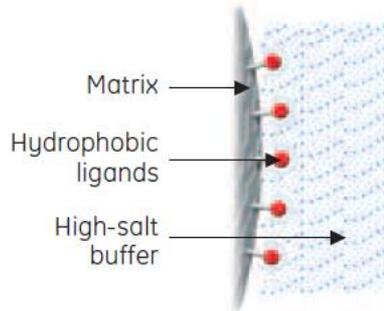


A) Highly ordered water shells surround the hydrophobic surfaces of ligands and proteins. Hydrophobic substances are forced to merge to minimize the total area of such shells (maximize entropy). Salts enhance the hydrophobic interaction. B) The equilibrium of the hydrophobic interaction is controlled predominantly by the salt concentration.

Purificazione di proteine ricombinanti per studi strutturali

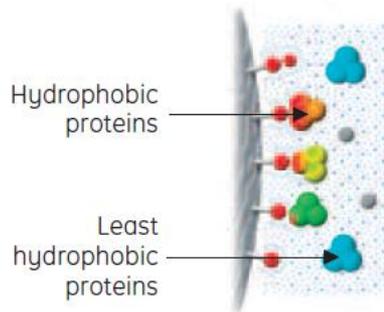
Metodi cromatografici:

HIC: fasi della purificazione



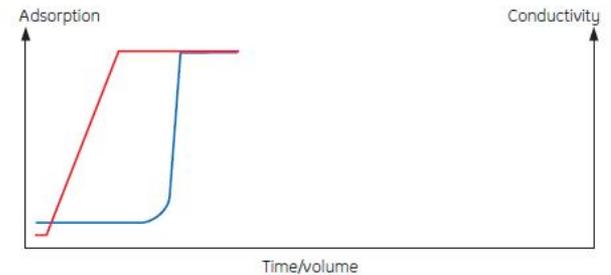
Equilibration

HIC medium equilibrated with high-salt start buffer.



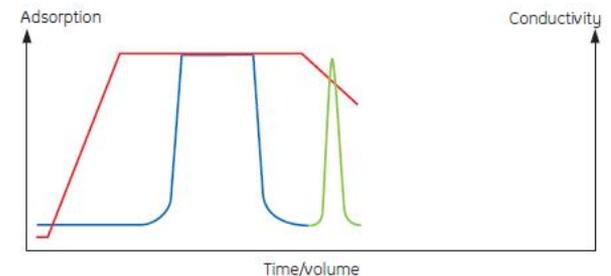
Sample application

Start buffer causes hydrophobic proteins bind to hydrophobic ligands on the medium, becoming concentrated on the column. Proteins with insufficient hydrophobicity elute during or just after sample application.



Elution 1

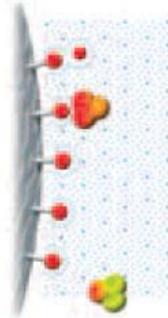
Decreasing salt content (using a linear gradient) causes hydrophobic proteins to elute: the least hydrophobic proteins elute first.



Purificazione di proteine ricombinanti per studi strutturali

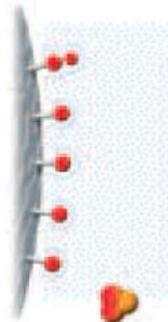
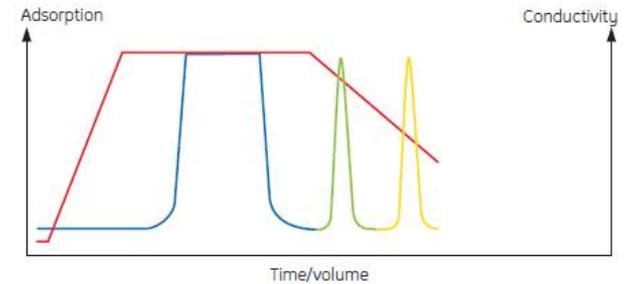
Metodi cromatografici:

HIC: fasi della purificazione

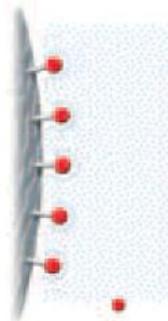
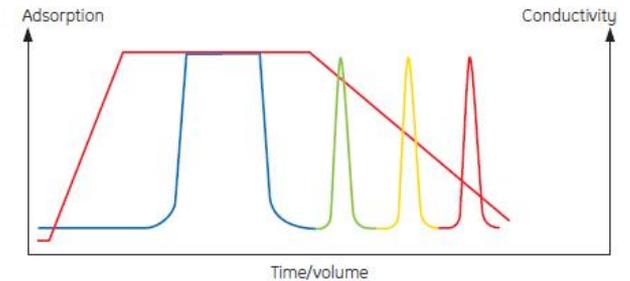


Elution 2

Further decreases in salt displace the more hydrophobic proteins (more tightly bound).

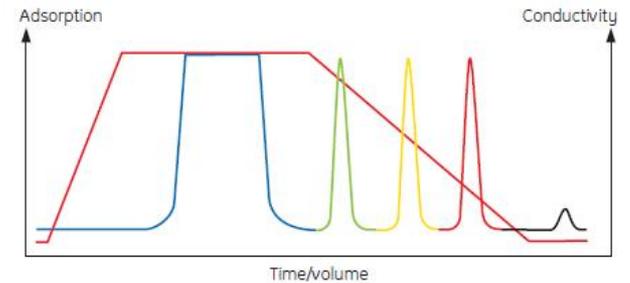


Elution 3



Wash

Final "salt-free" wash removes any hydrophobically bound proteins before re-equilibration.

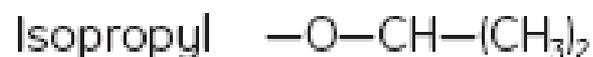
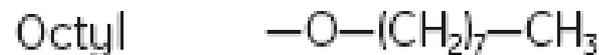
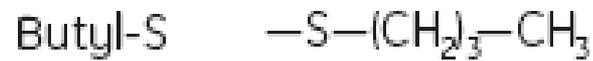
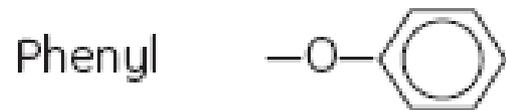


Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

HIC: tipi di matrici

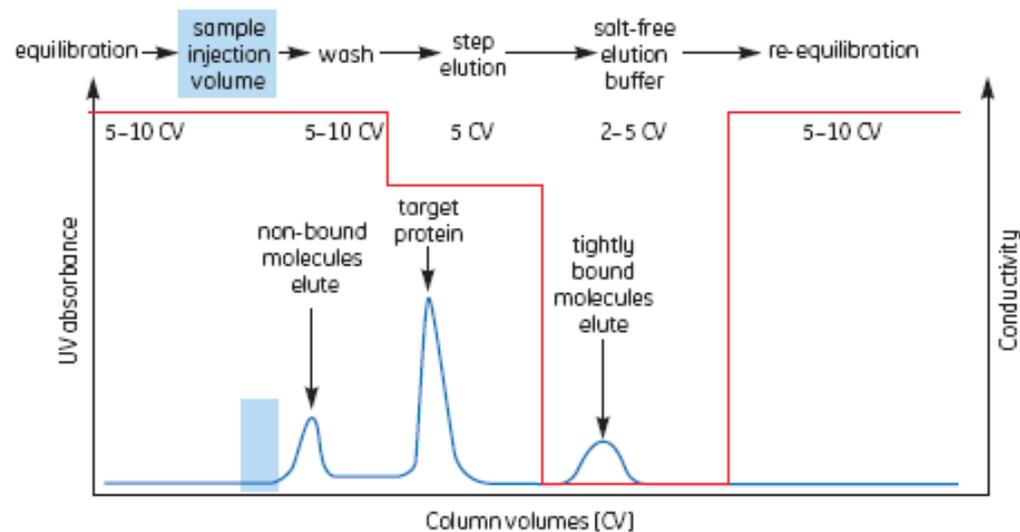
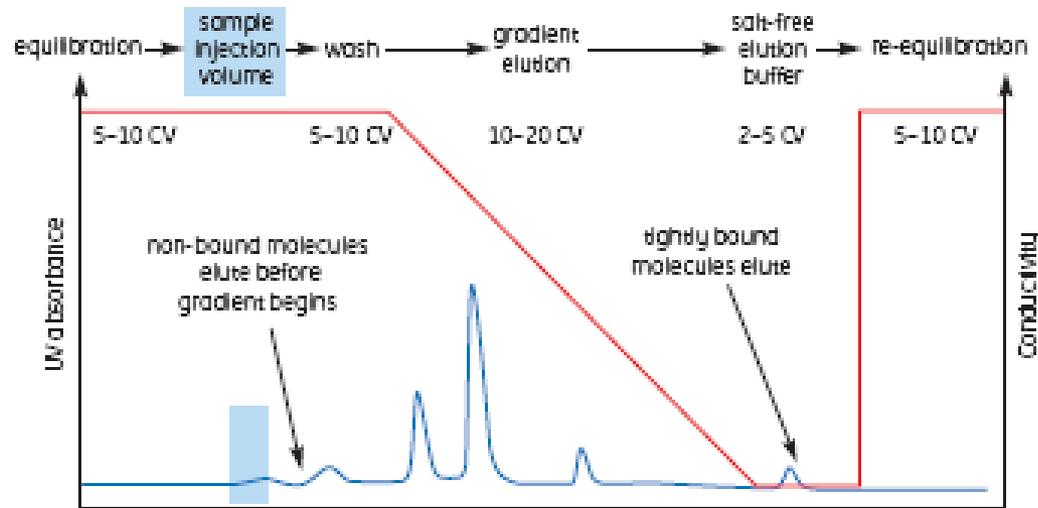
Matrix	Form	Mean particle size
SOURCE 15	Polystyrene/divinyl benzene	15 μm
Sepharose High Performance	Agarose 6%	34 μm
Sepharose 6 Fast Flow	Agarose 6%	90 μm
Sepharose 4 Fast Flow	Agarose 4%	90 μm



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

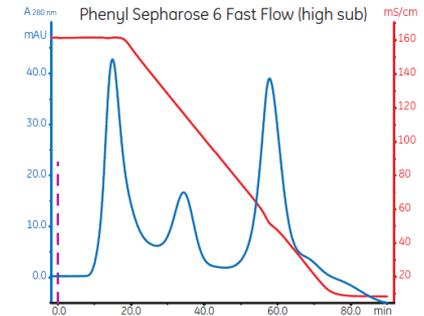
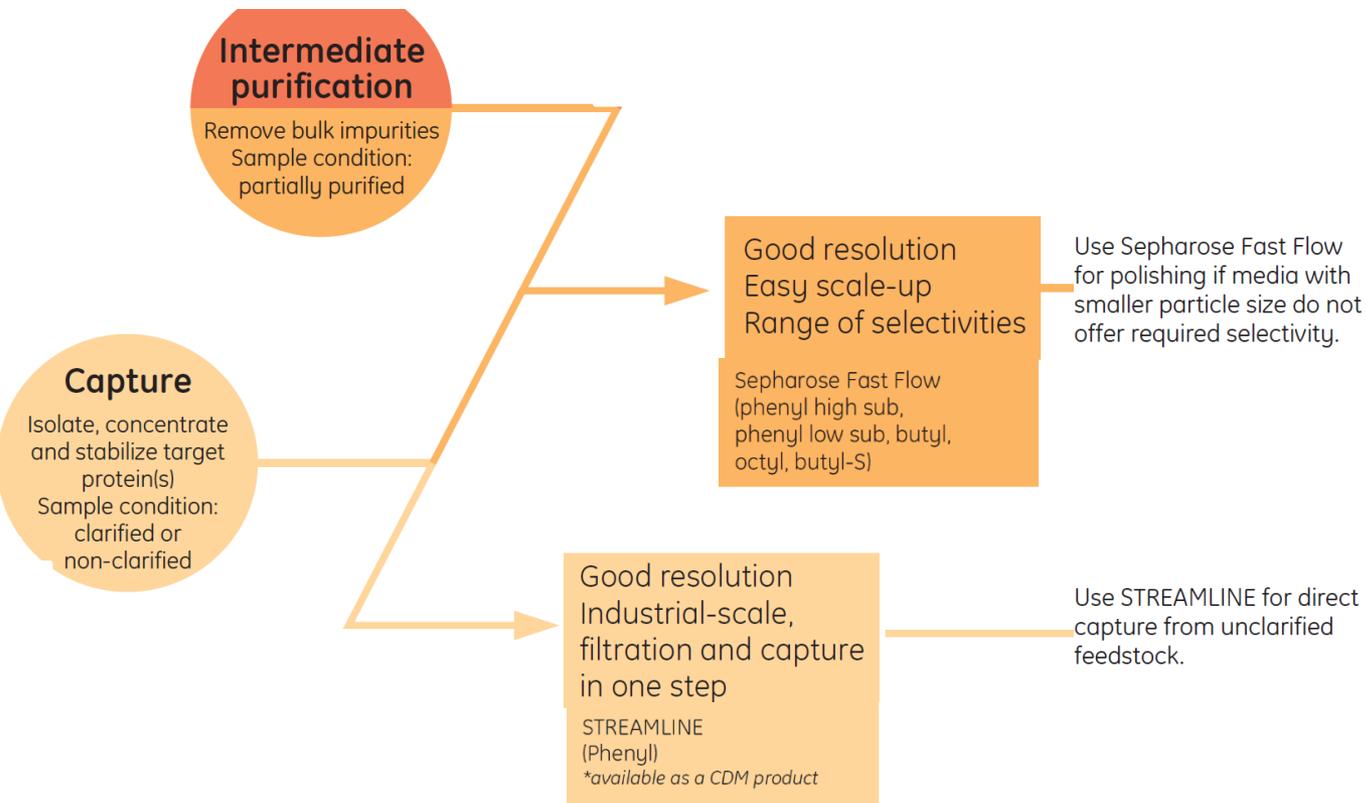
HIC: tipi di eluizione



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

HIC: CAPTURE INTERMEDIATE PURIFICATION

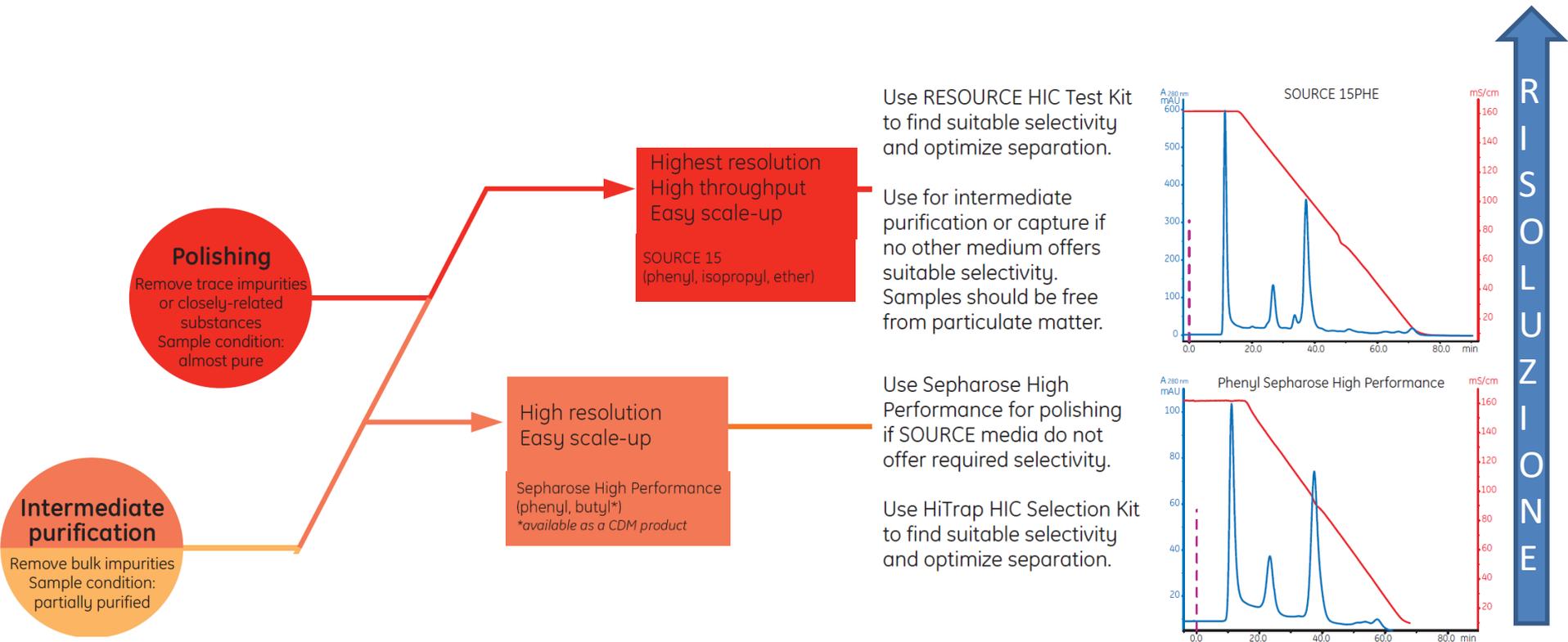


R
I
S
O
L
U
T
I
O
N
E

Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

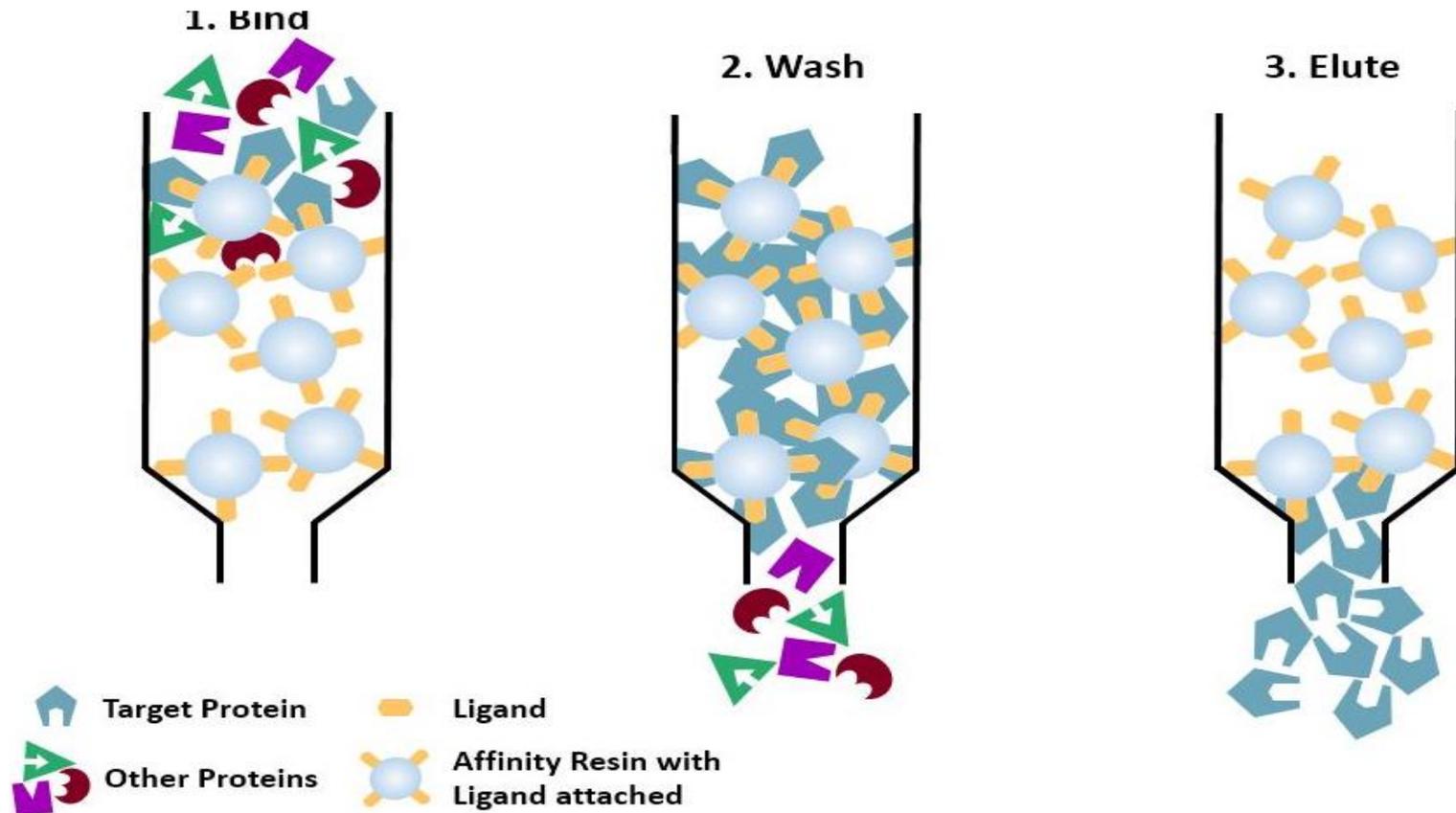
HIC: INTERMEDIATE PURIFICATION & POLISHING



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

AFFINITY

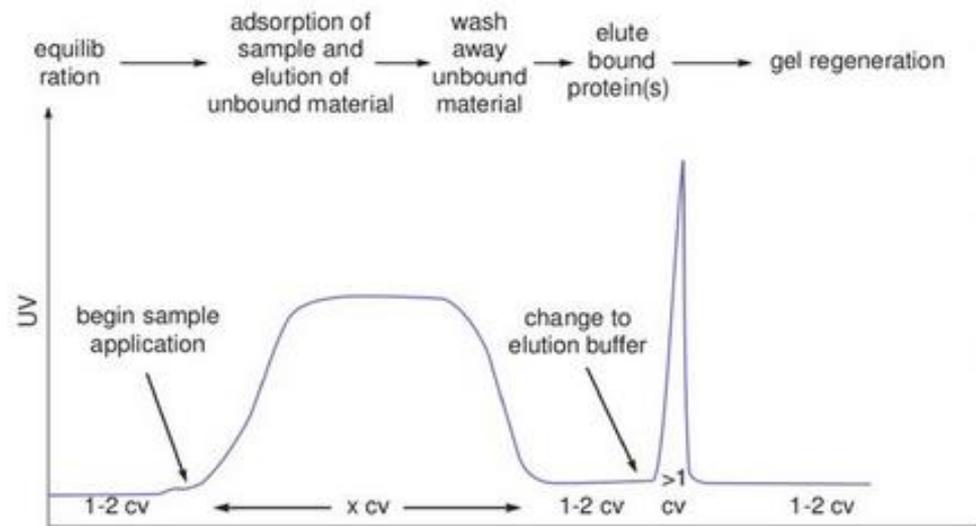


Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

AFFINITY

<i>Types of ligand</i>	<i>Target molecules or molecules of interest</i>
Enzyme	Substrate analogue, inhibitor, cofactor
Antibody	Antigen
Lectin	Polysaccharide, glycoprotein, cell surface receptor, cell
Nucleic acid	Complementary base sequence, nucleic acid binding protein
Hormone	Receptor
Avidin	Biotin
Calmodulin	Calmodulin-binding molecule
Poly(A)	RNA containing poly(U) sequences
Glutathione	Glutathione-S-transferase or GST fusion proteins
Proteins A and G	Immunoglobulins
Metal ions	Poly (His) fusion proteins, native proteins with histidine

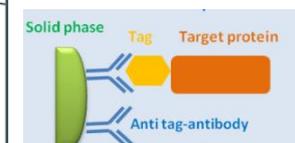
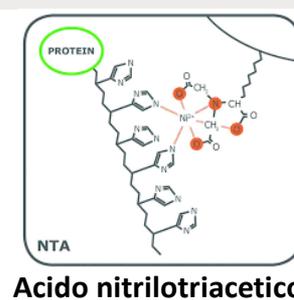
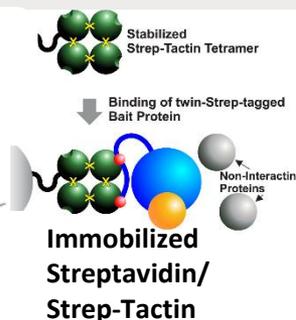
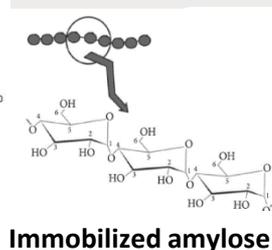
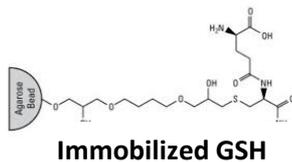


Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

AFFINITY

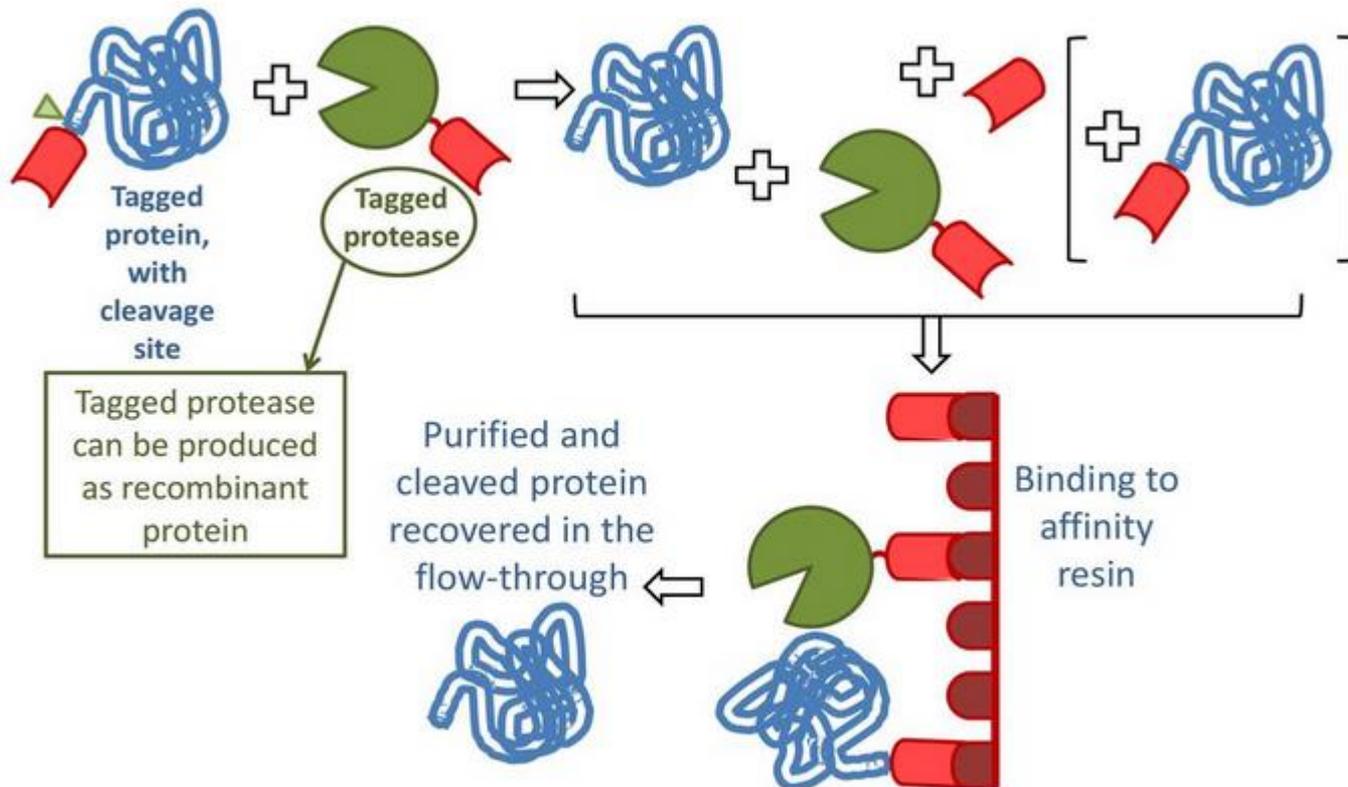
	GST	MBP	<i>Strep</i> TM -tag II	His	FLAG TM
Size	26 kDa	40 kDa	8 aa	6 aa	8 aa
Protein binding capacity	30 mg/ml	10 mg/ml	6 mg/ml	40 mg/ml	0.6 mg/ml
Purity	++	++	+++	+	+++
Increased solubility	Yes	Yes	No	No	No
Risk for interference with function	Yes	Yes	No	No	No



Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

REVERSE AFFINITY



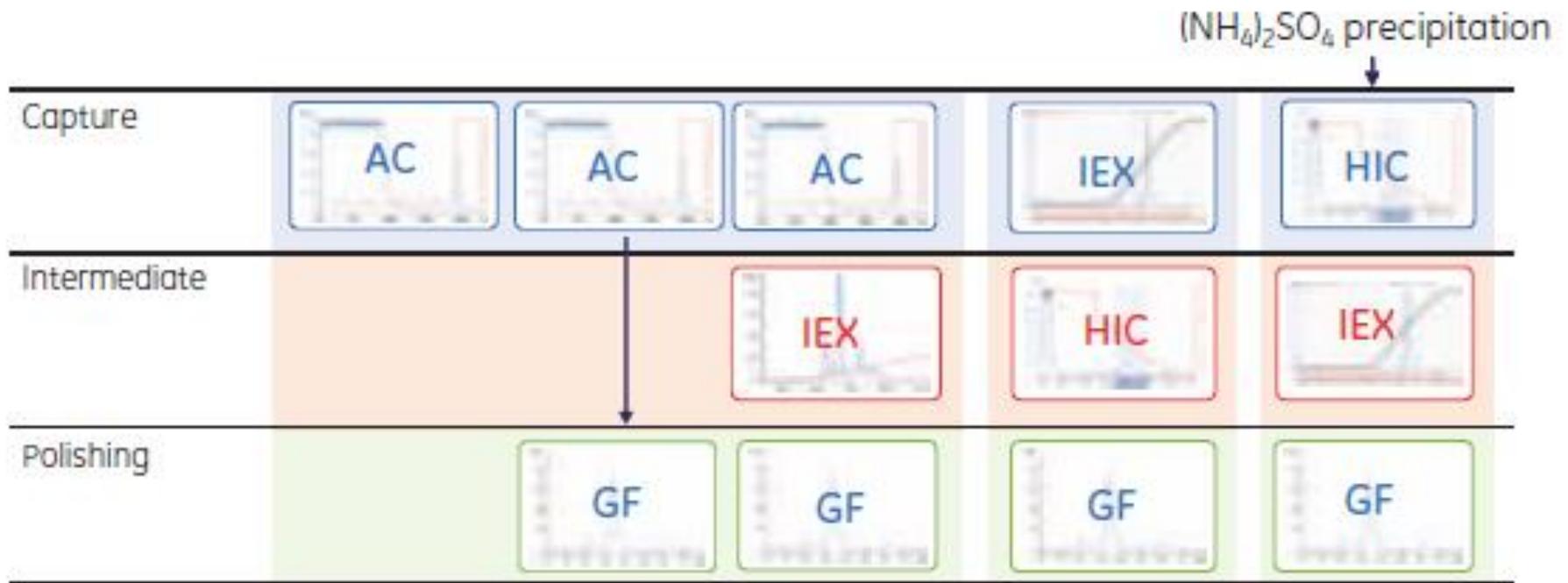
Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:

Method	Typical characteristics		Purification phase			Sample start conditions	Sample end conditions
	Resolution	Capacity	Capture	Intermediate	Polishing		
AC	+++ or ++	+++ or ++	+++	++	+	Various binding conditions	Specific elution conditions
IMAC	+++	++	+++	++	+	For purifying histidine-tagged proteins using Ni Sepharose columns: 20-40 mM imidazole; pH > 7; 500 mM NaCl; no chelators Other proteins: low concentration of imidazole	High concentration of imidazole, pH > 7, 500 mM NaCl
GF	++	+	+		+++	Most conditions acceptable, limited sample volume	Buffer exchange possible, diluted sample
IEX	+++	+++	+++	+++	+++	Low ionic strength. pH depends on protein and IEX type	High ionic strength or pH changed
HIC	+++	++	++	+++	+++	High ionic strength, addition of salt required	Low ionic strength

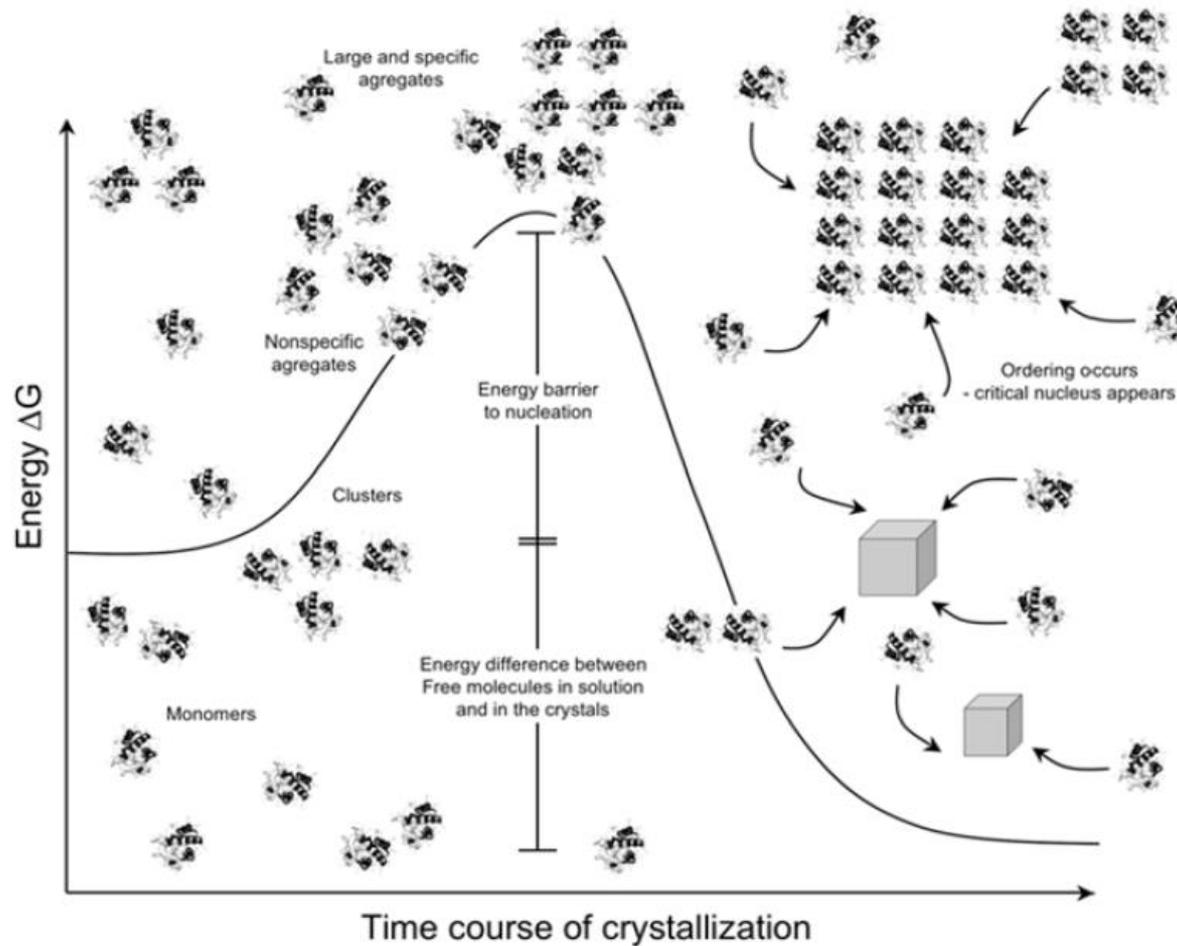
Purificazione di proteine ricombinanti per studi strutturali

Metodi cromatografici:



Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: considerazioni energetiche



$$\Delta G_{\text{cryst}} = \Delta H_{\text{prot}} - T\Delta S_{\text{prot}} - T\Delta S_{\text{solv}}$$

Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: considerazioni energetiche

autoassemblaggio in cristalli: $\Delta G_c = \Delta H_c - T(\Delta S_{\text{protein}} + \Delta S_{\text{solvent}})$

- numero di **contatti relativamente piccolo/interazioni deboli**:
 ΔH_c negativi/poco significativi (**-17/0 kcal/mol**)
- perdita significativa dei gradi di libertà traslazionale/rotazionale

conformazionale dei residui coinvolti

$-T\Delta S_{\text{protein}} = +7,5/25 \text{ kcal / mol (a } T_{\text{amb}})$

contributo destabilizzante positivo alla ΔG_c

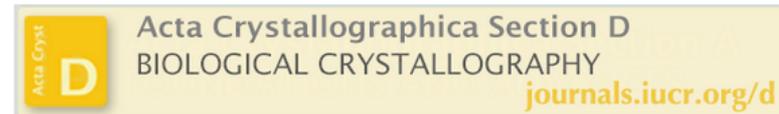
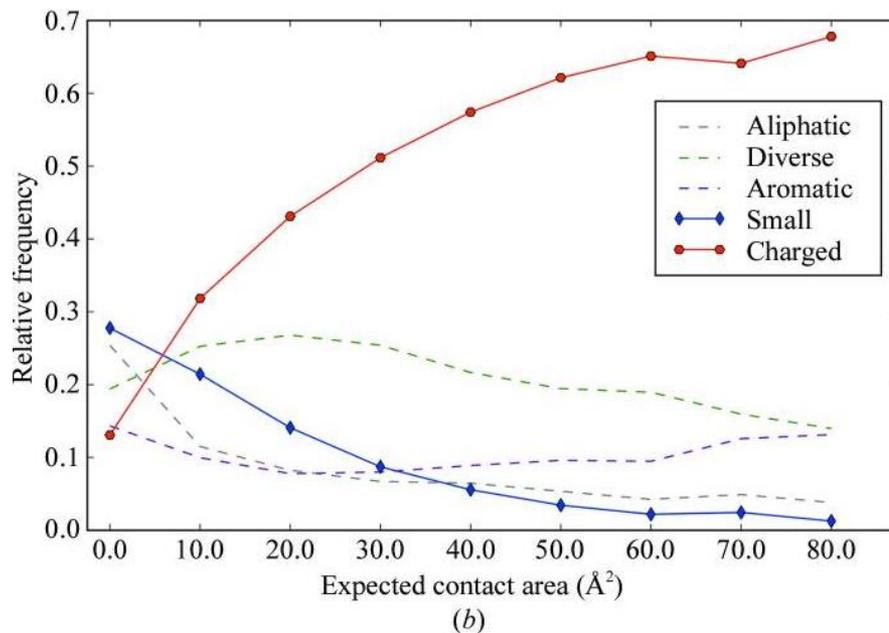
- compensato dal rilascio di solvente attraverso residui di superficie sia idrofobici che polari durante l'impacchettamento (guadagno entropico del solvente)

→ **$-T\Delta S_{\text{sov}} = -25/75 \text{ kcal / mol (a } T_{\text{amb}})$** corrisponde al rilascio di **5-30** H₂O ordinate

Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: parametri chimico-fisici critici

Database = 821 proteine funzionalmente monomeriche



The role of entropy and polarity in intermolecular contacts in protein crystals

Marcin Cieřlik and Zygmunt S. Derewenda



superficie accessibile al solvente dei residui

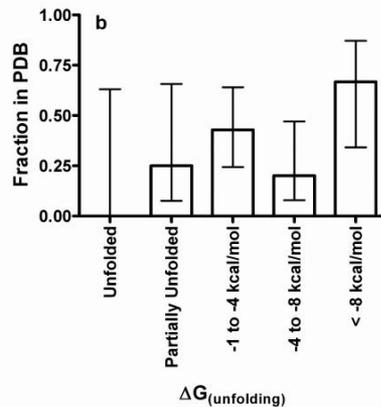
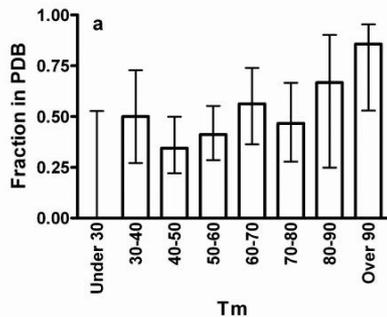


entropia delle catene laterali

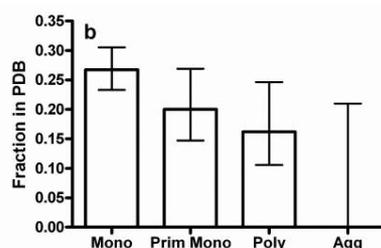
Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: parametri chimico-fisici critici

Database = 679 proteine espresse (179 cristallizzano)



stabilità



proprietà idrodinamiche

epitopi superficiali ordinati (G) → interazioni stereospecifiche nell' impacchettamento

nature
biotechnology

Analysis | Published: 14 December 2008

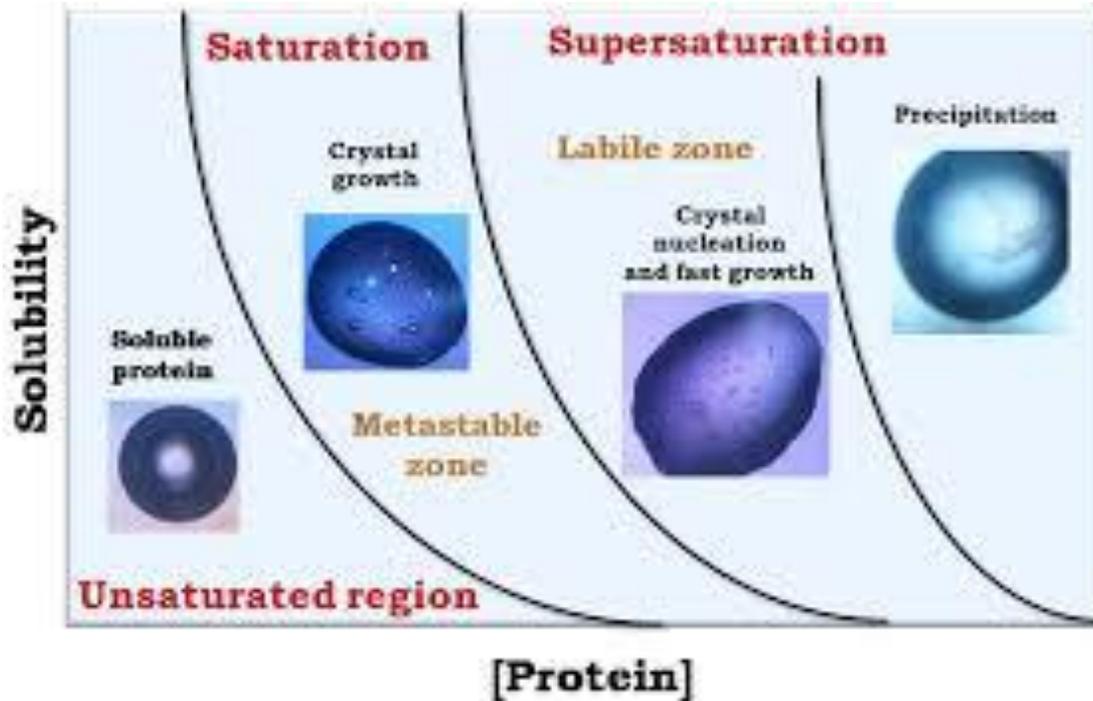
Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data

W Nicholson Price II, Yang Chen, Samuel K Handelman, Helen Neely, Philip Manor, Richard Karlin, Rajesh Nair, Jinfeng Liu, Michael Baran, John Everett, Saichiu N Tong, Farhad Forouhar, Swarup S Swaminathan, Thomas Acton, Rong Xiao, Joseph R Luft, Angela Lauricella, George T DeTitta, Burkhard Rost, Gaetano T Montelione & John F Hunt

Nature Biotechnology 27, 51–57 (2009) | [Download Citation](#)

Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: parametri critici



solubilità

regioni con solvente strutturato

Gravy (Grand average hydrophobicity) index



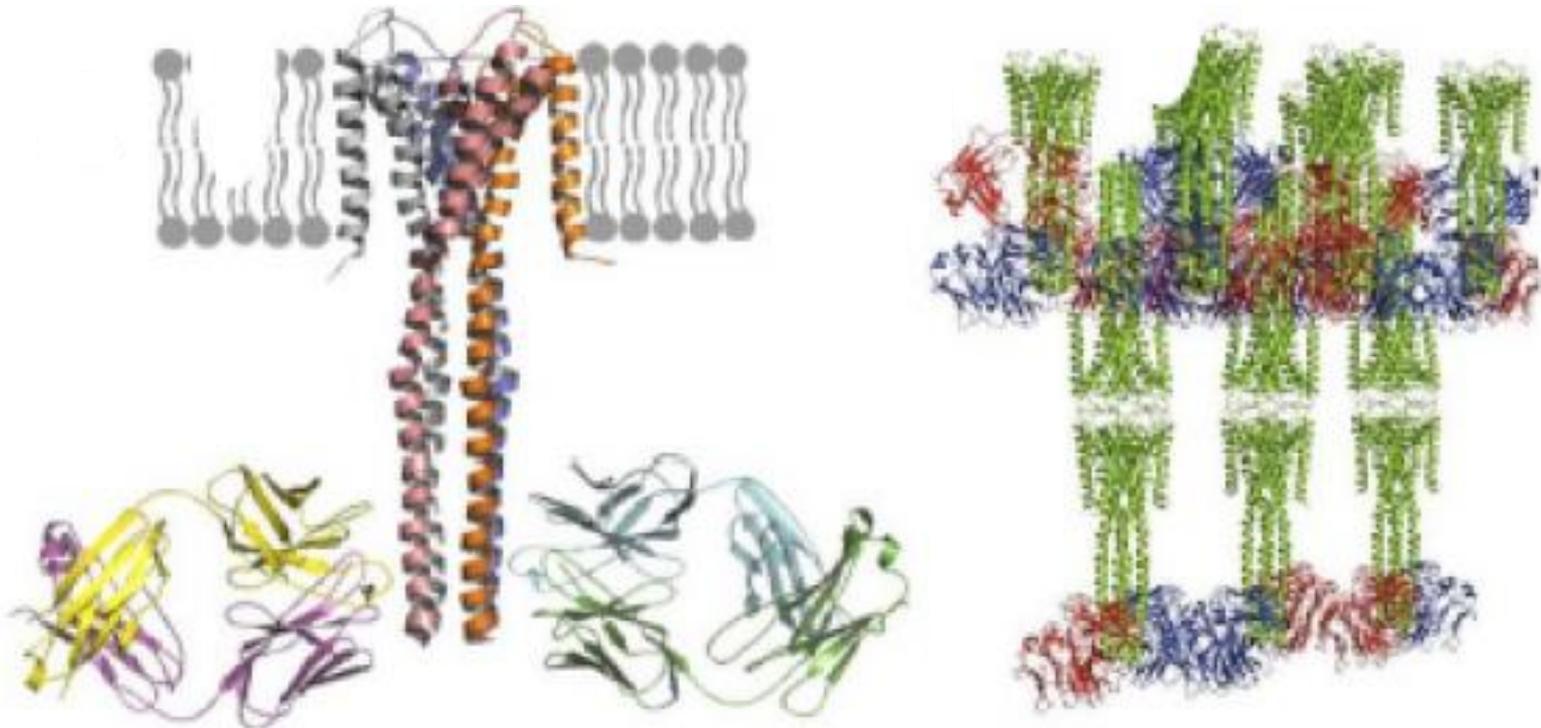
Flessibilità/regioni disordinate

pl

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Solubilità

1) FAb



Crystal structure of full-length KcsA in its closed conformation

Serdar Uysal^{a,b}, Valeria Vázquez^{a,b}, Valentina Tereshko^{a,b}, Kaori Esaki^a, Frederic A. Fellouse^c, Sachdev S. Sidhu^c, Shohei Koide^a, Eduardo Perozo^{a,b,1}, and Anthony Kossiakoff^{a,b,1}

^aDepartment of Biochemistry and Molecular Biology, and ^bInstitute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637; and ^cDepartment of Protein Engineering, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080

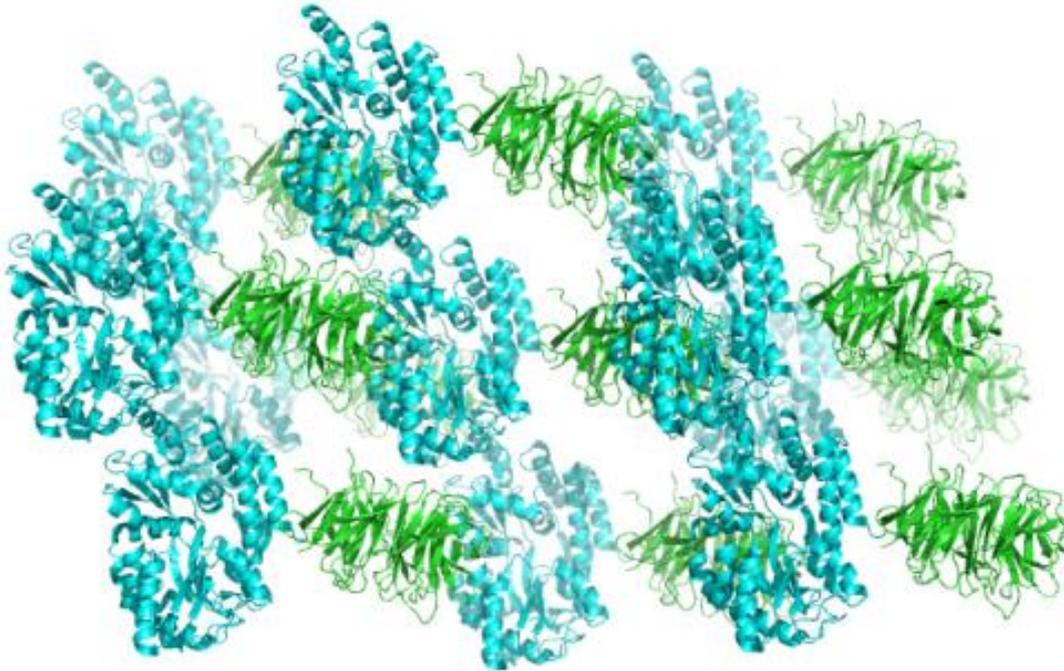
Edited by John Kuriyan, University of California, Berkeley, CA, and approved February 18, 2009 (received for review October 29, 2008)

KcsA is a cytosol-activated, voltage-modulated K⁺ channel that has been shown to be a member of the Shaker family of voltage-gated K⁺ channels. The crystal structure of full-length KcsA in its closed conformation is presented here.

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Solubilità

2) fusion partners (ex: MBP)



Structure of a signal transduction regulator, RACK1,
from *Arabidopsis thaliana*

HEMAYET ULLAH,¹ ERICA LOUISE SCAPPINI,² ANDREA FLORENCE MOON,³
LATANYA VERONICA WILLIAMS,^{1,4} DAVID LEE ARMSTRONG,²
AND LARS CHRISTIAN PEDERSEN³

¹Department of Biology, Howard University, Washington, DC 20059, USA

²Laboratory of Neurobiology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709, USA

³Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709, USA

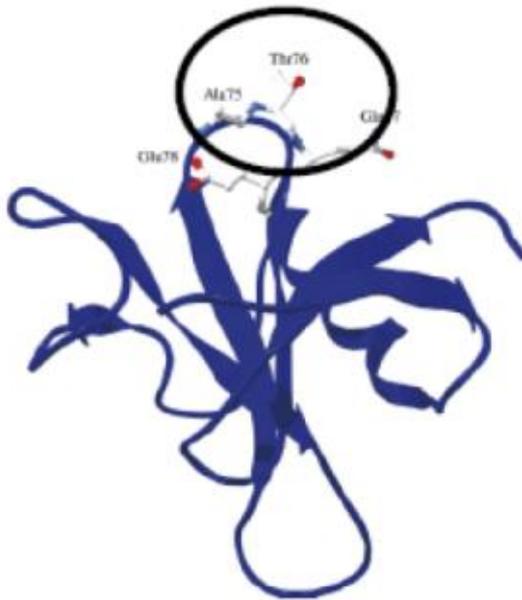
(RECEIVED February 26, 2008; FINAL REVISION June 20, 2008; ACCEPTED June 25, 2008)

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Solubilità

3) direct surface engineering

Amino acid at position 76	Solubility (mg/ml) ^a
Asp	43
Arg	42
Glu	42
Ser	39
Lys	31
Gly	27
Ala	27
His	24
Asn	21
Thr	20
Gln	20
Pro	15
Cys	12
Met	11
Val	10
Leu	9.3
Ile	8.2
Tyr	5.6
Phe	4.4
Trp	3.6



^a The error in these measurements is $\pm 10\%$.

jmb

Journal of Molecular Biology

Volume 366, Issue 2, 16 February 2007, Pages 449-460



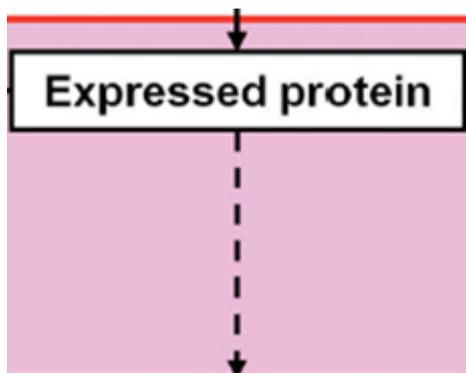
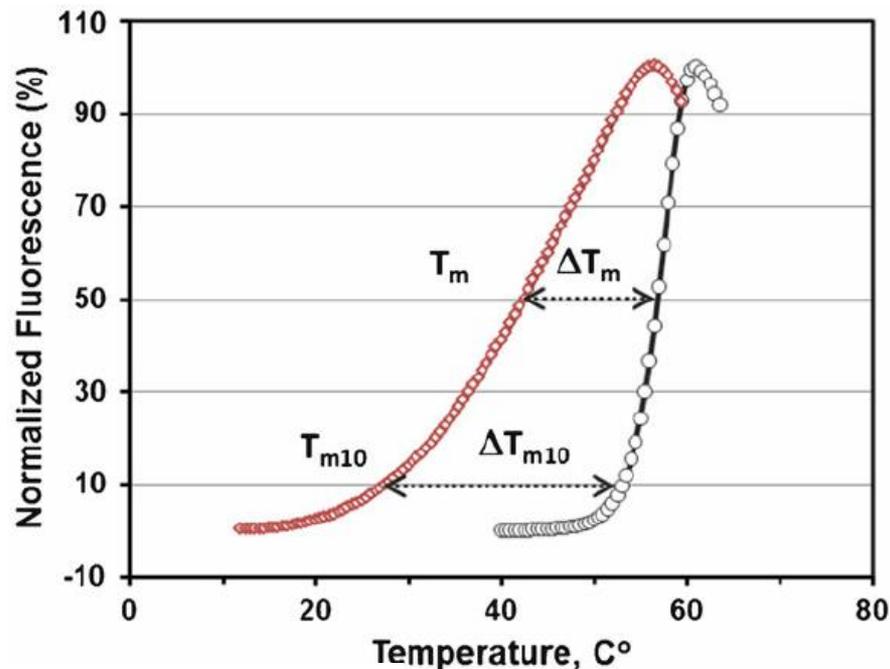
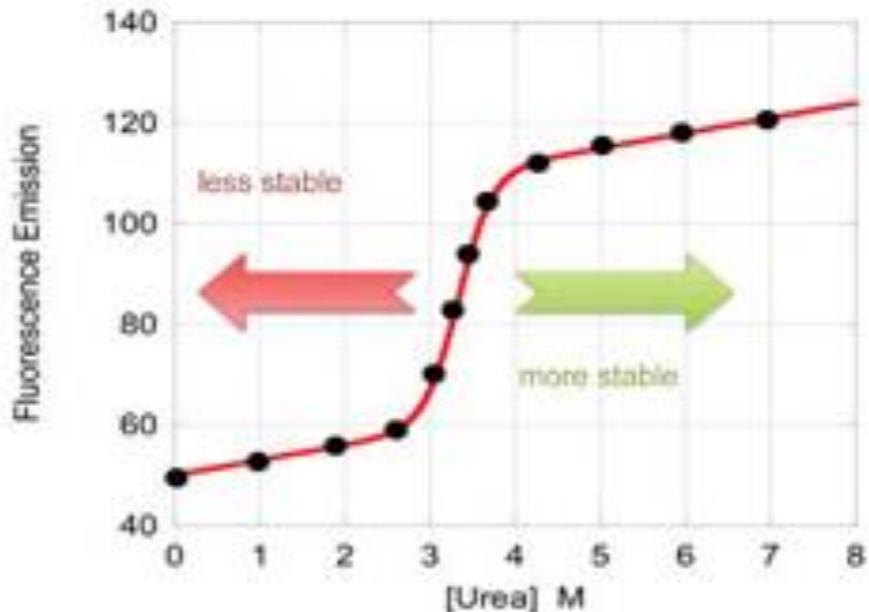
Amino Acid Contribution to Protein Solubility:
Asp, Glu, and Ser Contribute more Favorably
than the other Hydrophilic Amino Acids in
RNase Sa

Saul R. Trevino¹, J. Martin Scholtz^{1, 2} , C. Nick Pace^{1, 2} 

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Stabilità/Solubilità

Ottimizzazione delle curve di denaturazione



screening di formulazioni

Buffer, pH

- None (Deionized water)
- Sodium acetate, 4.5
- Sodium citrate, 5.0
- Succinic acid, 5.5
- MES, 6.0
- BIS-TRIS, 6.5
- Imidazole, 7.0
- HEPES, 7.5
- Tris, 8.0
- BIS-TRIS propane, 8.5
- AMPD, 9.0
- Glycine, 9.5

12
unique
buffers

Ionic Strength (NaCl) (Suggested Assay)

- 0
- 0.05 M
- 0.1 M
- 0.15 M
- 0.2 M
- 0.25 M
- 0.5 M
- 1.0 M

8 levels
of ionic
strength

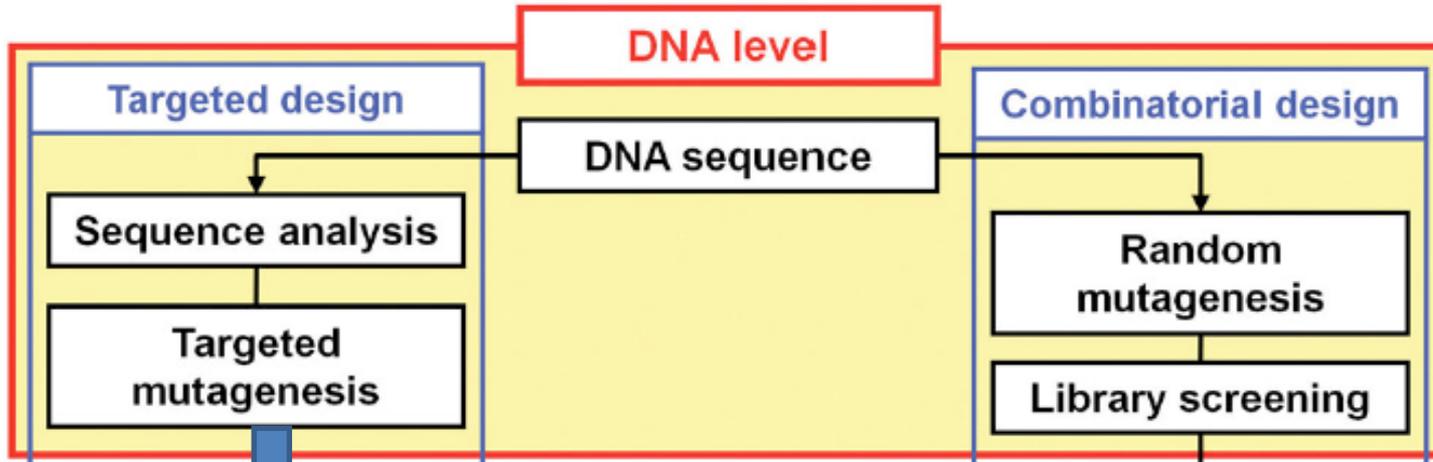
Creates
96
reagents



Produzione di proteine ricombinanti per studi strutturali

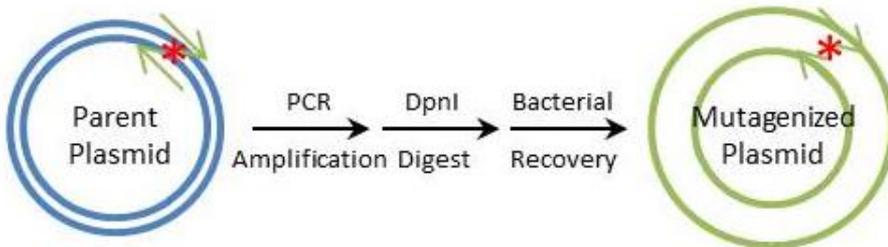
Strategie di improvement: Stabilità/Solubilità

•Biologia molecolare:

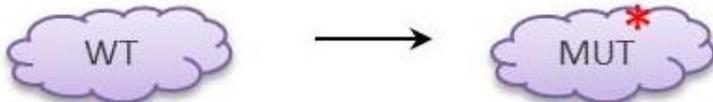


Mutagenesi sito specifica:

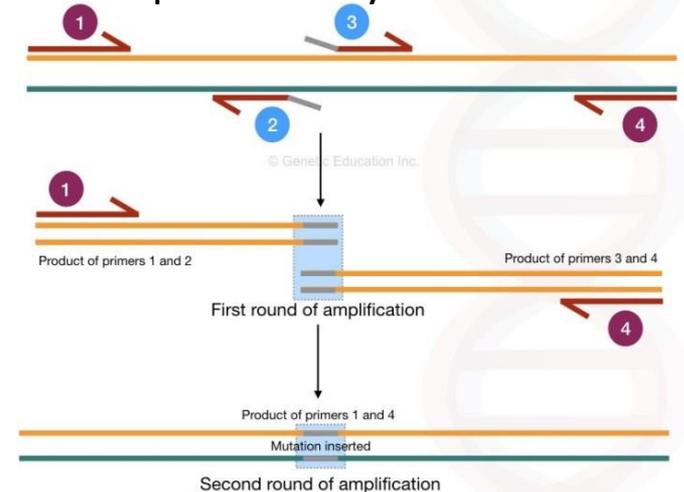
1. QuikChange



Protein



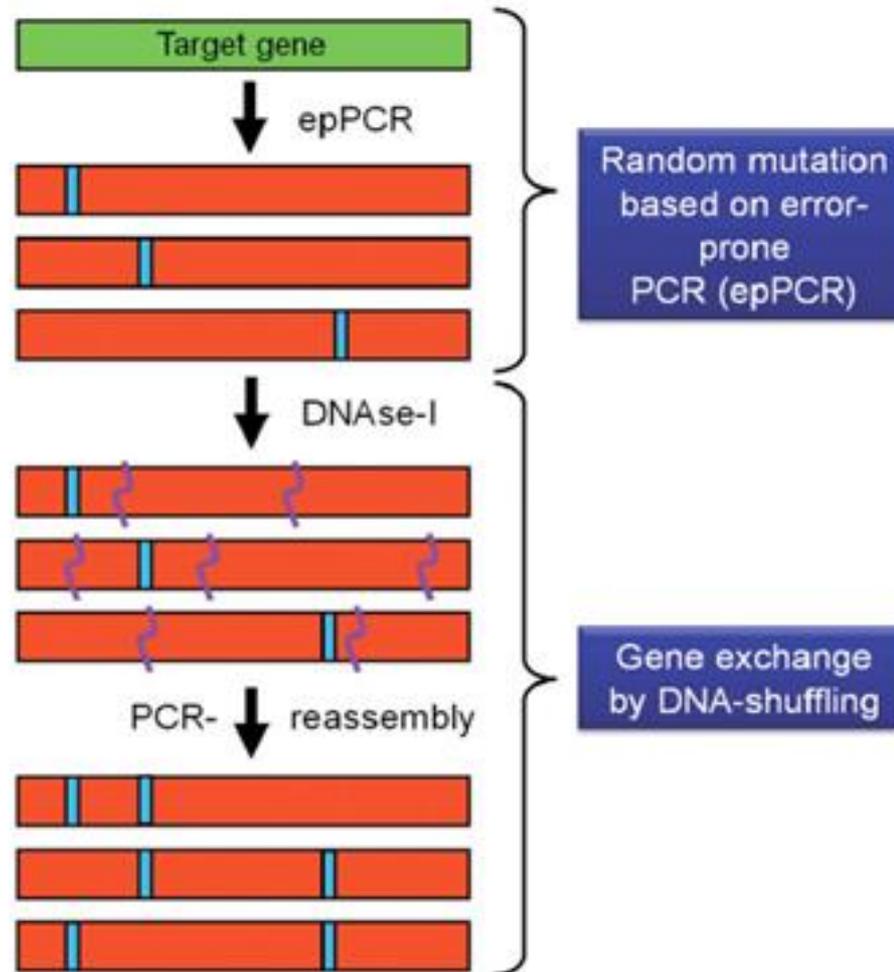
2. overlap- assembly PCR



Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Stabilità/Solubilità

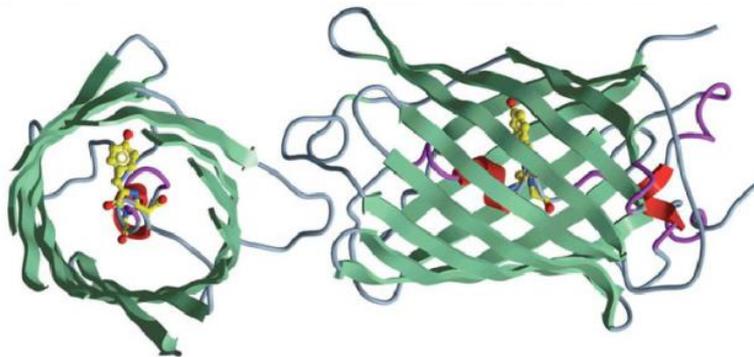
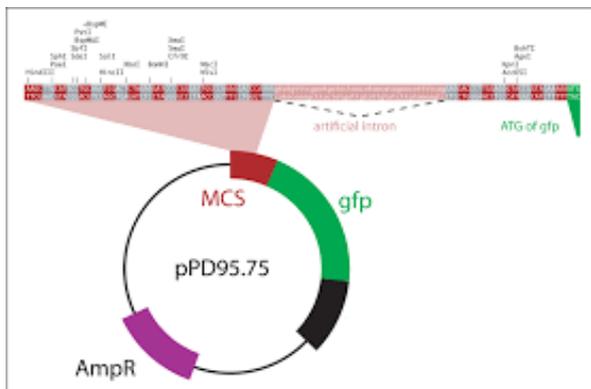
Directed evolution



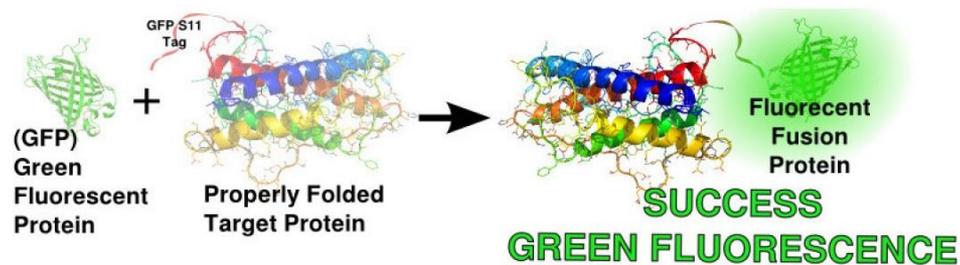
Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Stabilità/Solubilità

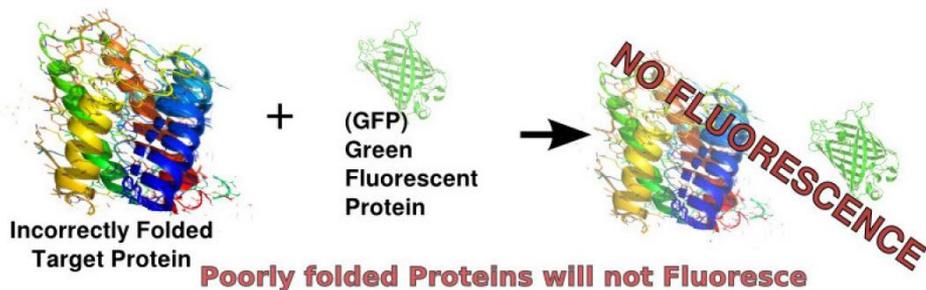
Read out: fusione con GFP



S65-Y66-G67: ciclizzazione, deidratazione e ossidazione → p-idrossi benzil idene-imidazolinone



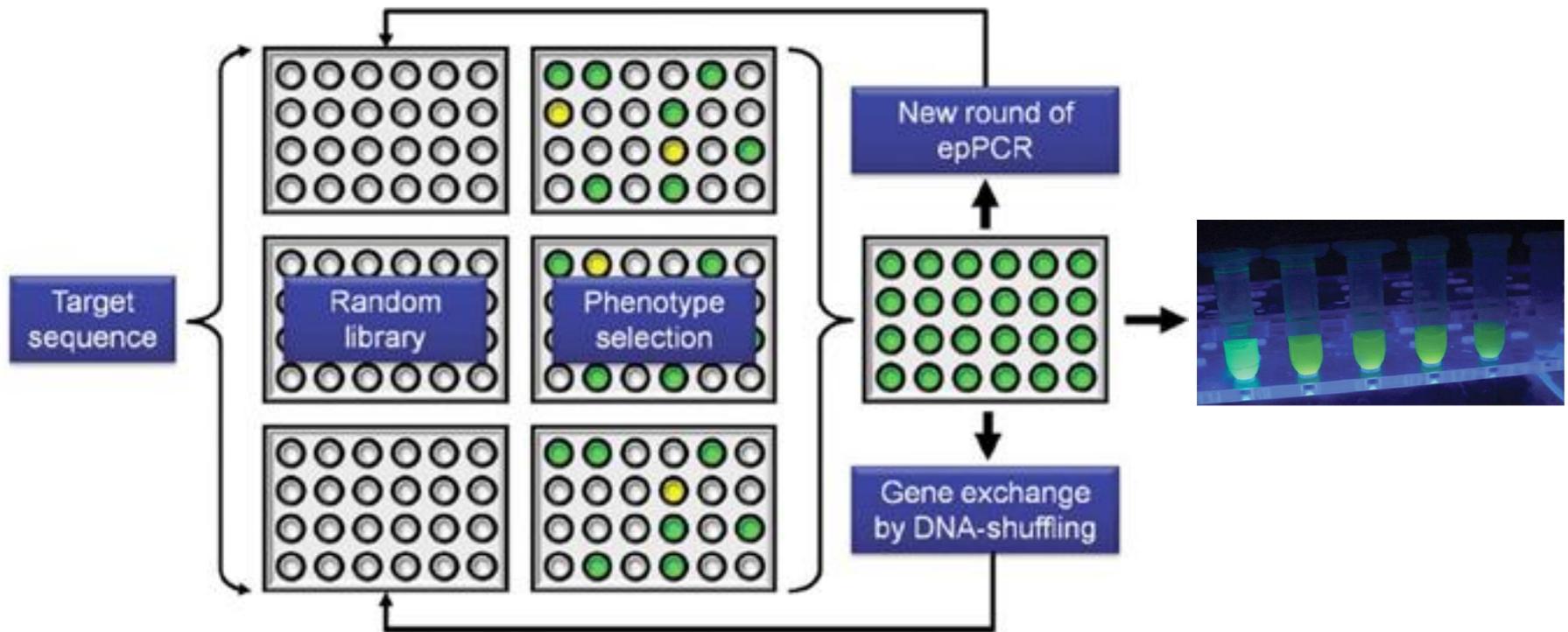
Bacterial Fold-N-Glow™
Split GFP S11 Plasmid



Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Stabilità/Solubilità

Read out: fusione con GFP



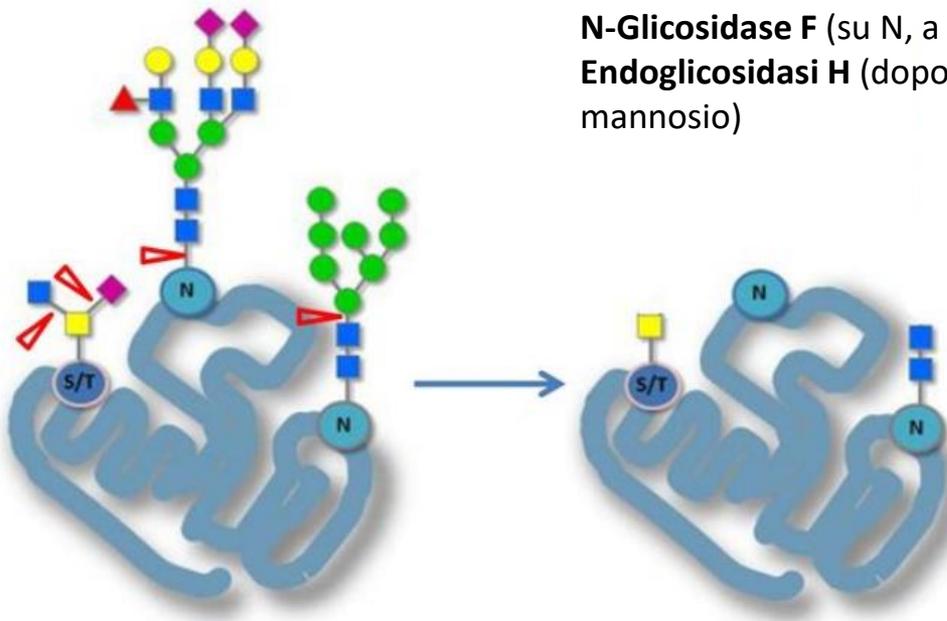
Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

- Riduzione della flessibilità:
- Eliminare modifiche post-traduzionali:

Glicosilazioni:

1. Mutagenesi
2. Inibitori
3. Glicosidasi



N-Glicosidase F (su N, a monte della prima N-acetilglucosamina)
Endoglicosidasi H (dopo la prima N-acetilglucosamina in glicani ricchi di mannosio)

Neuron
Article



Structural and Mechanistic Insights into Nerve Growth Factor Interactions with the TrkA and p75 Receptors

Tom Wehrman,^{3,4} Xiaolin He,^{1,4} Bill Raab,³ Abhiram Dukipatti,¹ Helen Blau,³ and K. Christopher Garcia^{1,2,*}

¹Departments of Molecular and Cellular Physiology, and Structural Biology

²Howard Hughes Medical Institute

³Baxter Laboratory for Genetic Pharmacology, Department of Microbiology and Immunology, The Stem Cell Institute, Stanford University School of Medicine, Stanford, CA 94305, USA

⁴These authors contributed equally to this work.

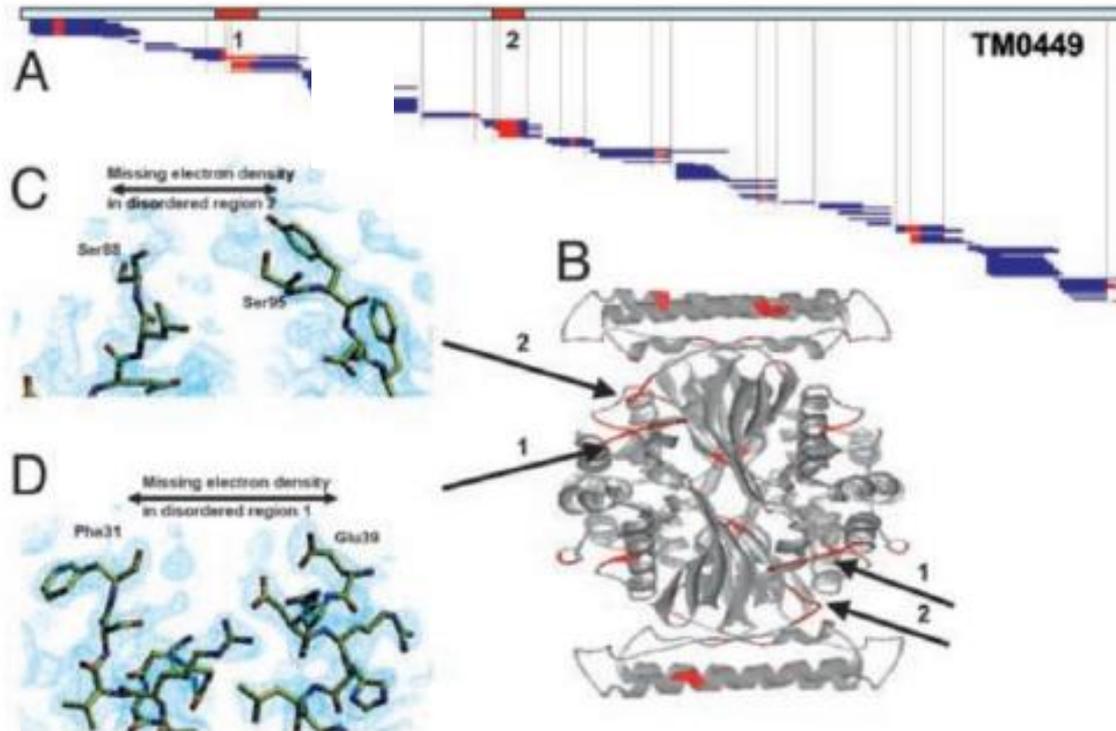
*Correspondence: kcgarcia@stanford.edu

DOI 10.1016/j.neuron.2006.09.034

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

- Riduzione della flessibilità:
- Eliminare linkers/tags (vedi analisi bioinformatica): biologia molecolare/proteolisi



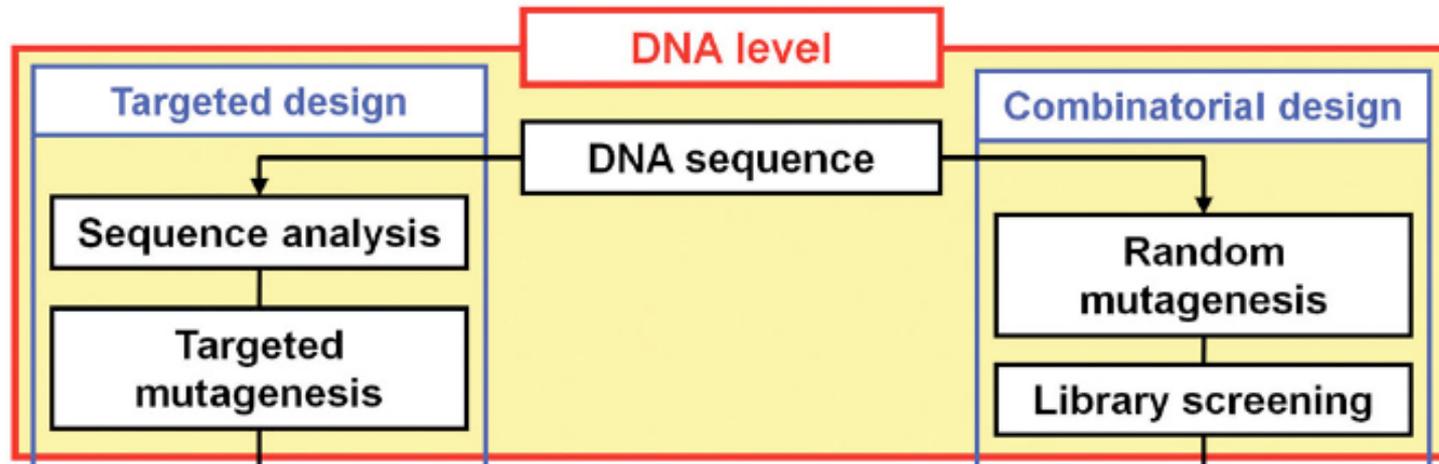
Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS

Dennis Pantazatos*, Jack S. Kim*, Heath E. Klock¹, Raymond C. Stevens², Ian A. Wilson², Scott A. Lesley¹, and Virgil L. Woods, Jr.*^{§1}

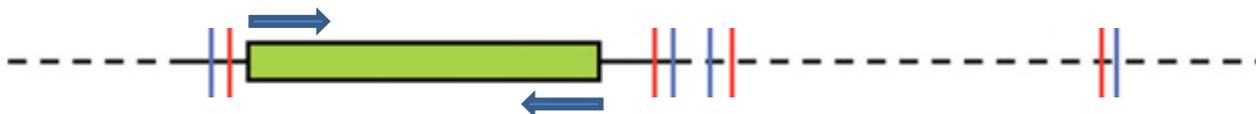
*Department of Medicine, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA, 92093; ¹Joint Center for Structural Genomics, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037; and ²Joint Center for Structural Genomics, Genomics Institute of the Novartis

Produzione di proteine ricombinanti per studi strutturali
Strategie di improvement: Abbassamento della barriera entropica

- **Riduzione della flessibilità:**

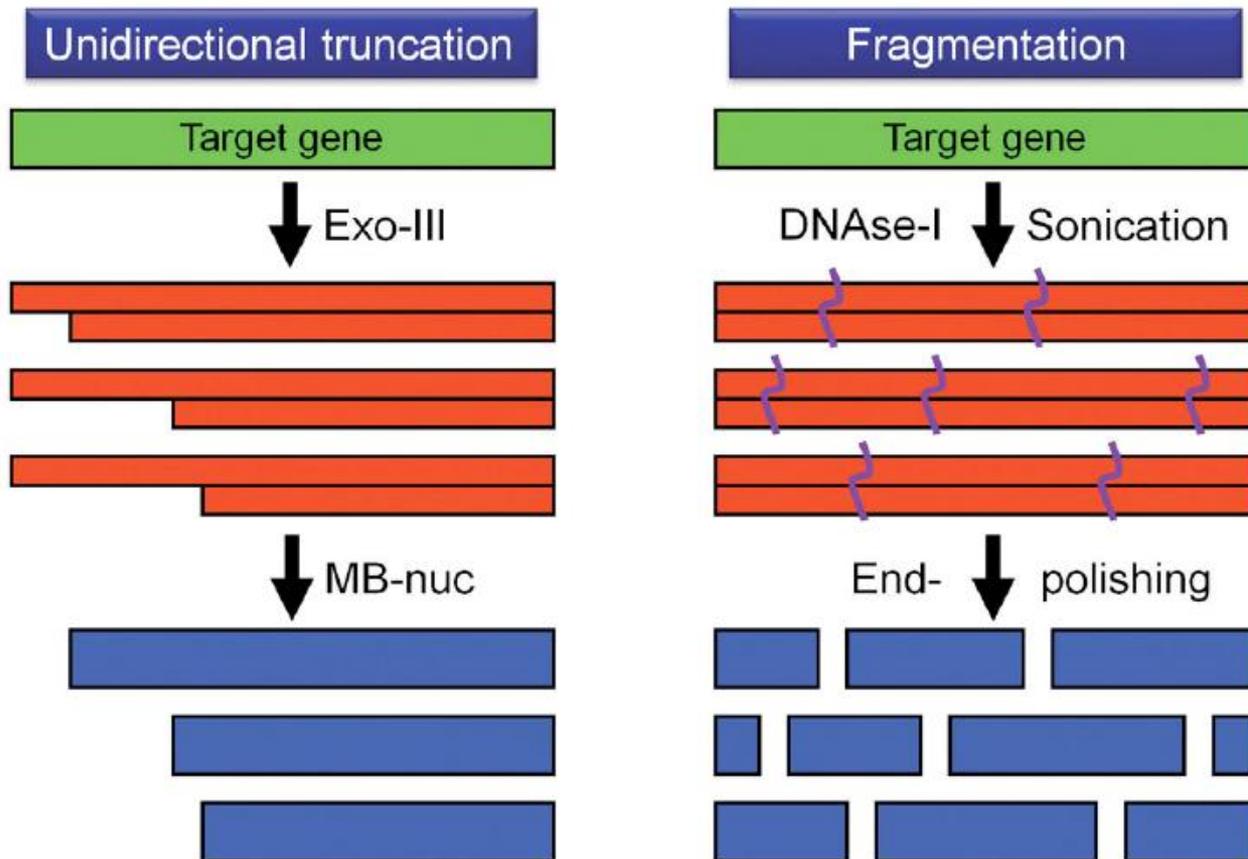


1. Mutanti di delezione



Produzione di proteine ricombinanti per studi strutturali
Strategie di improvement: Abbassamento della barriera entropica

2. **Generazione di libraries di mutanti di delezione mediante:**



Produzione di proteine ricombinanti per studi strutturali

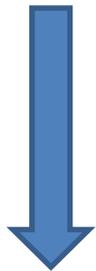
Strategie di improvement: Abbassamento della barriera entropica

• Surface Entropy Reduction:

MODIFICAZIONI CHIMICHE:

1) LYSIN REDUCTIVE METHYLATION

K superficiali



metilate

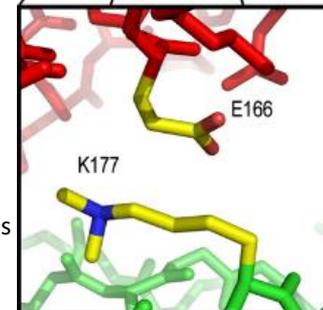
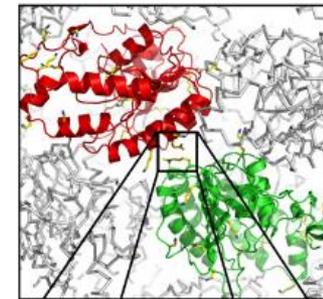
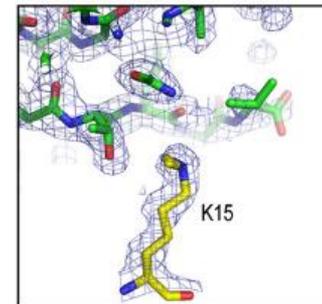
interazioni idrofobiche



Structure 14, 1617–1622, November 2006 ©2006 Elsevier Ltd All rights reserved DOI 10.1016/j.str.2006.09.005

Lysine Methylation as a Routine Ways & Means Rescue Strategy for Protein Crystallization

Thomas S. Walter,^{1,3} Christoph Meler,^{1,2,3,4}
Rene Assenberg,^{1,3} Kin-Fai Au,^{1,2,3,4}
Jingshan Ren,^{1,2,3,4} Anil Verma,^{1,3}
Joanne E. Nettleship,^{1,3} Raymond J. Owens,^{1,3}
David I. Stuart,^{1,2,3,4} and Jonathan M. Grimes^{1,2,3,4,*}



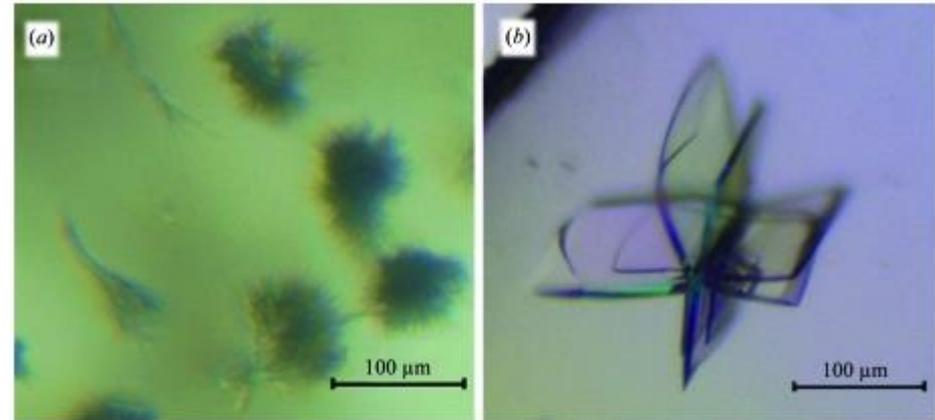
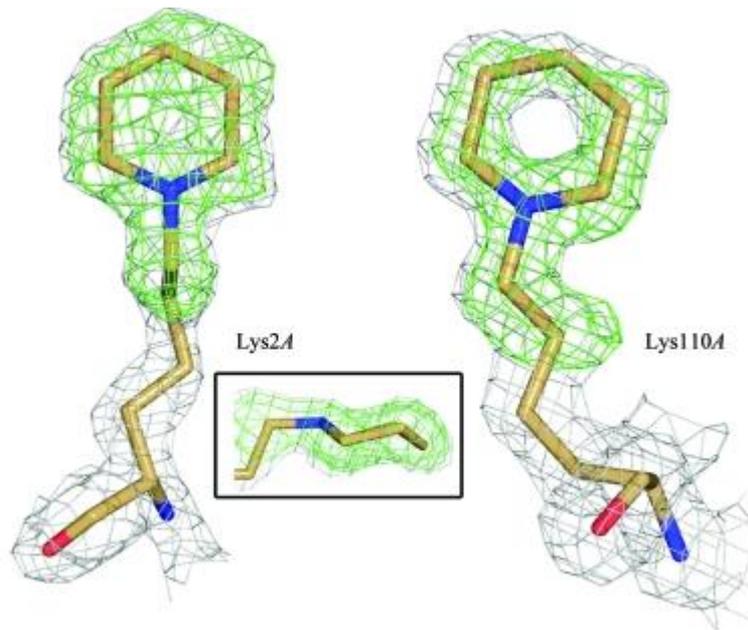
Examples of Methylated Lysines from MVE Methyltransferase

Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

- Surface Entropy Reduction: MODIFICAZIONI CHIMICHE:

2) REDUCTIVE CYCLIC PENTYLATION OF FREE AMINES



Acta Crystallogr D Biol Crystallogr. 2009 May 1; 65(Pt 5): 462–469.
Published online 2009 Apr 18. doi: [10.1107/S0907444909008324](https://doi.org/10.1107/S0907444909008324)

PMCID: PMC2672816
PMID: [19390151](https://pubmed.ncbi.nlm.nih.gov/19390151/)

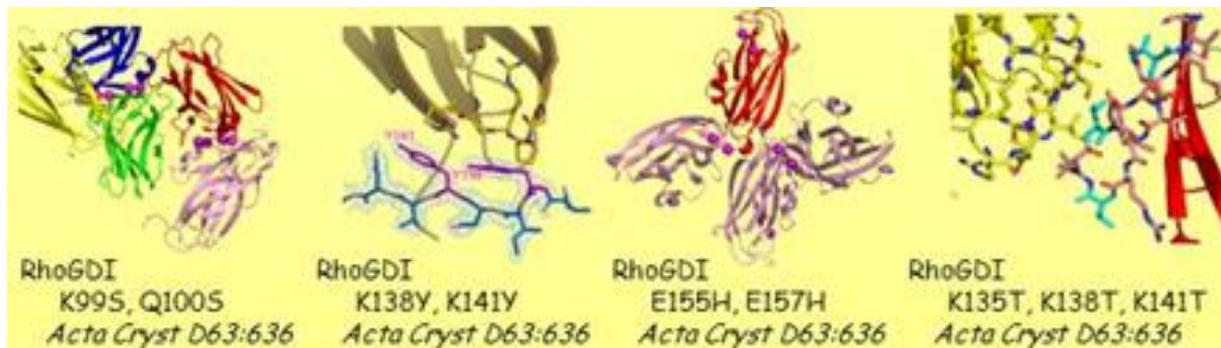
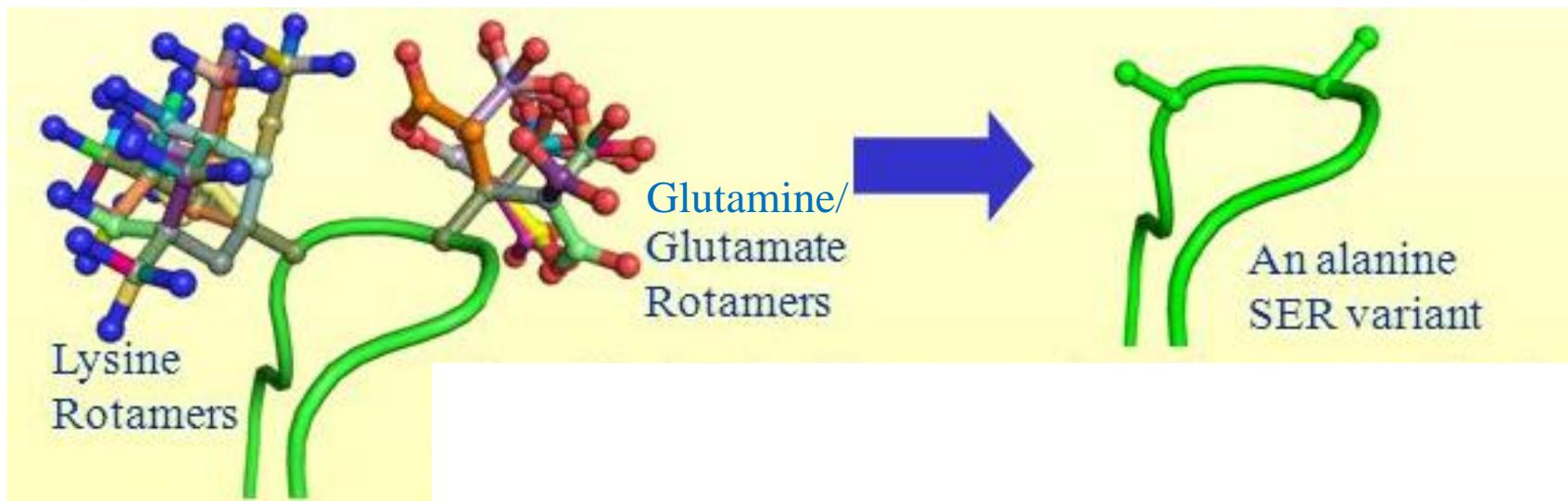
Crystallization of a pentapeptide-repeat protein by reductive cyclic
pentylation of free amines with glutaraldehyde

[Matthew W. Vetting](#),^{a,*} [Subray S. Hegde](#),^a and [John S. Blanchard](#)^a

Produzione di proteine ricombinanti per studi strutturali
Strategie di improvement: Abbassamento della barriera entropica

- Surface Entropy Reduction: biologia molecolare

MUTAGENESI sito specifica



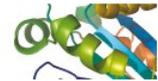
Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

MUTAGENESI sito specifica: selezione dei residui I

<http://services.mbi.ucla.edu/SER/>

PROTEIN|SCIENCE

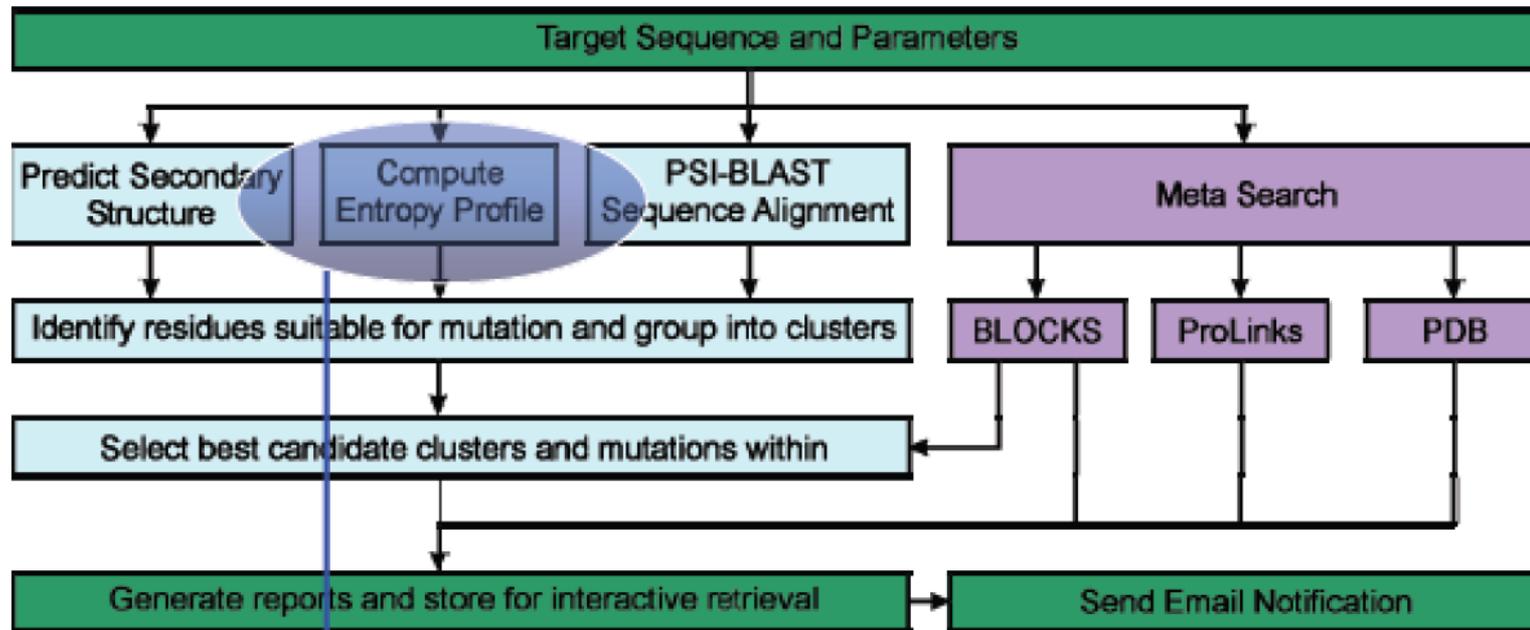


research-article | [Free Access](#)

Toward rational protein crystallization: A Web server for the design of crystallizable protein variants

Lukasz Goldschmidt, David R. Cooper, Zygmunt S. Derewenda, David Eisenberg

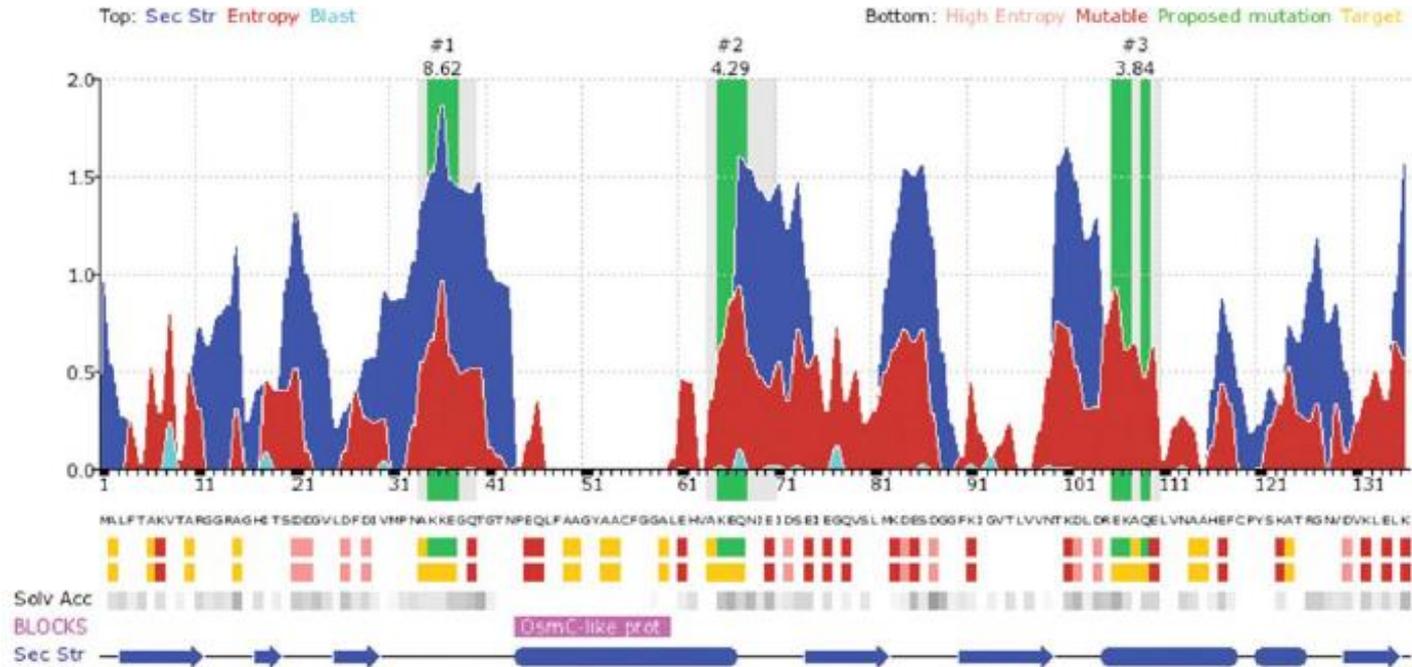
First published: 02 January 2009 | <https://doi.org/10.1110/ps.072914007> | Cited by: 134



Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

MUTAGENESI sito specifica: selezione dei residui II



Cluster #1:

- K 35 => A
- K 36 => A
- E 37 => A

Residues 34 - 37: AKKEGQ [?]



Cluster #2:

- K 65 => A
- E 66 => A
- Q 67 => A

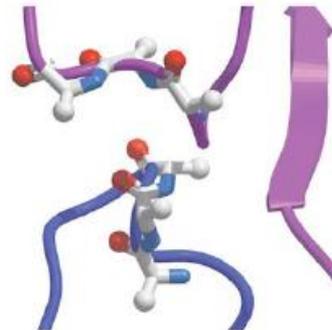
Residues 64 - 67: AKEQNI E [?]



Cluster #3:

- E 106 => A
- K 107 => A
- Q 109 => A

Residues 106 - 109: EKAQE [?]

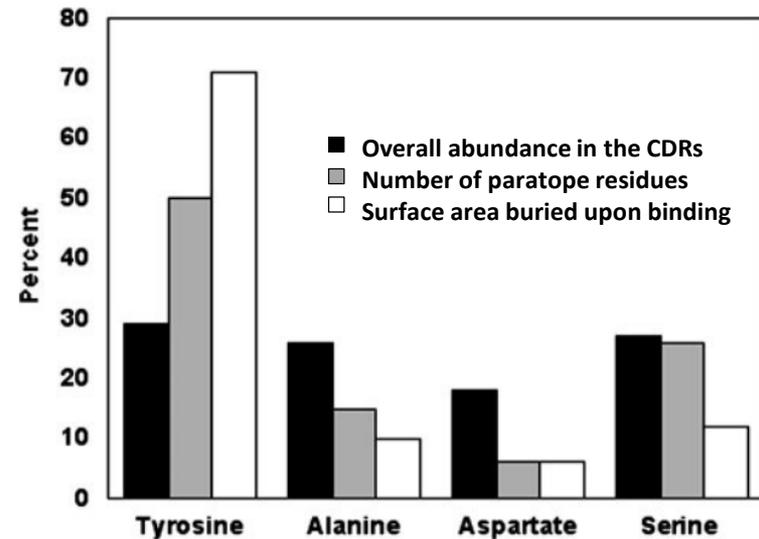
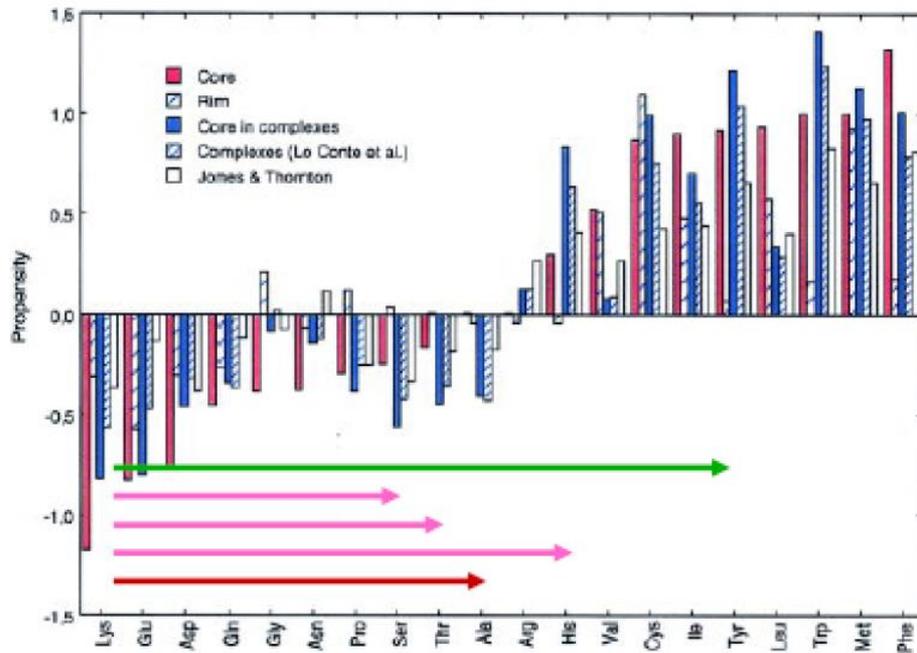


Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: Abbassamento della barriera entropica

•Surface Entropy Reduction: MUTAGENESI sito specifica:

non solo A, ma anche Y, H e S favoriscono interazioni intermolecolari



Proteins: Structure, Function, and Bioinformatics / Volume 53, Issue 3

Dissecting subunit interfaces in homodimeric proteins

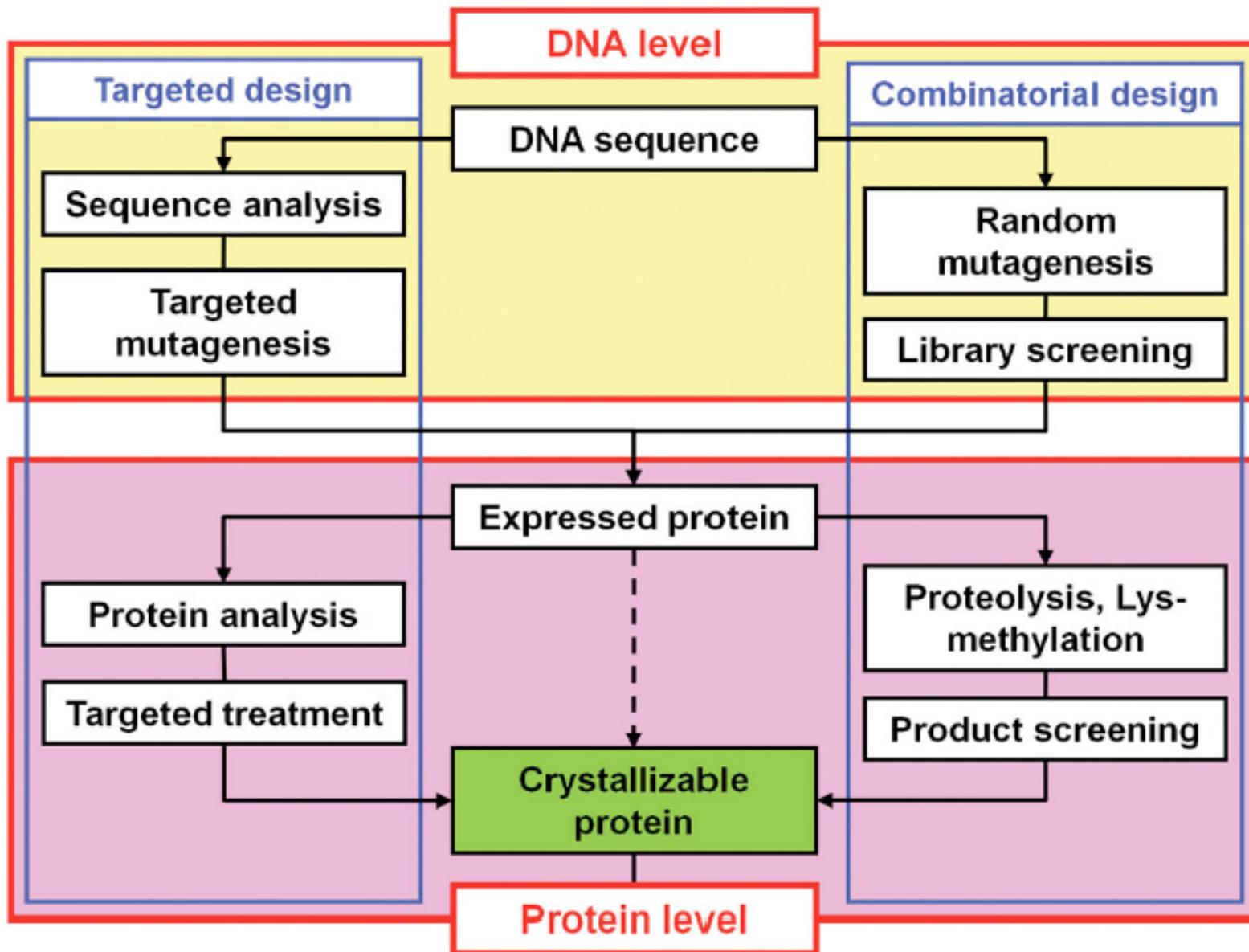
Ranjit Prasad Bahadur, Pinak Chakrabarti, Francis Rodier, Joël Janin

Synthetic antibodies from a four-amino-acid code:
A dominant role for tyrosine in antigen recognition

Frederic A. Fellouse*, Christian Wiesmann*, and Sachdev S. Sidhu**

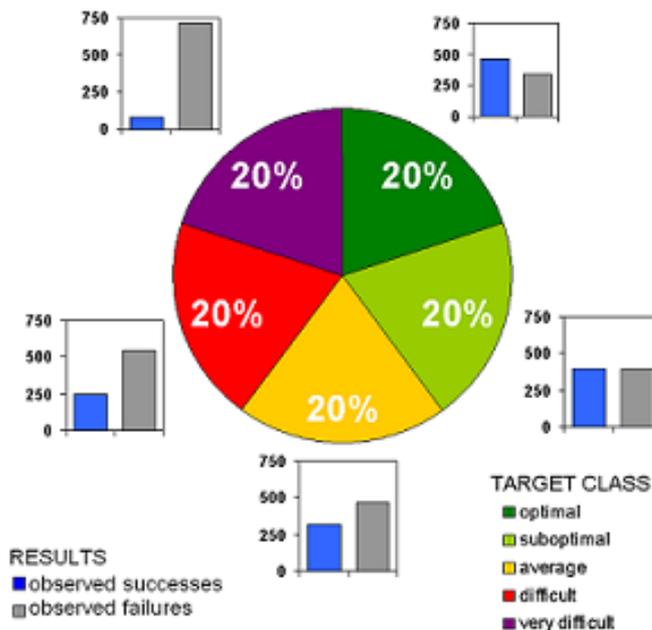
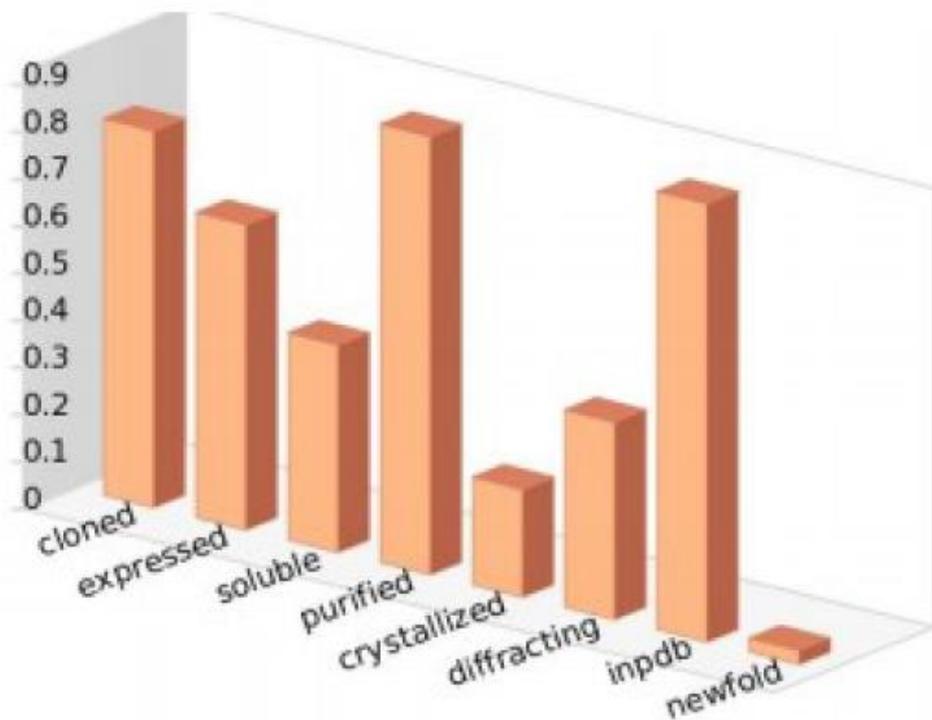
Produzione di proteine ricombinanti per studi strutturali

Strategie di improvement: summary



Produzione di proteine ricombinanti per studi strutturali

Cristallizzazione: probabilità di successo

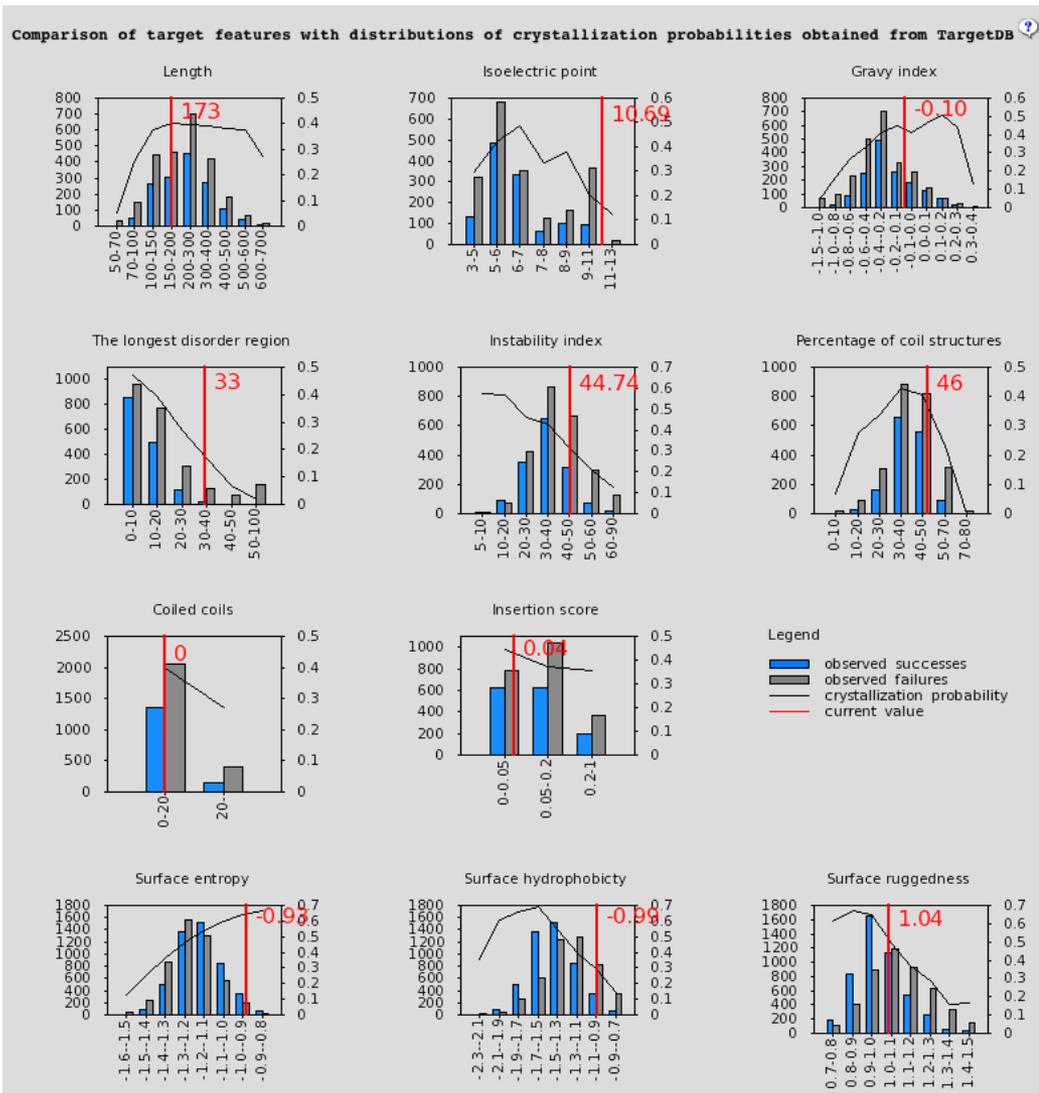


Classi definite per diversa percentuale di successo nella cristallizzazione osservata in TargetDB (database centralizzato dai centri di genomica strutturale)

Espressione di proteine ricombinanti per studi strutturali

Predizione del successo della cristallizzazione

<http://xtalpred.godziklab.org/XtalPred-cgi/xtal.pl>

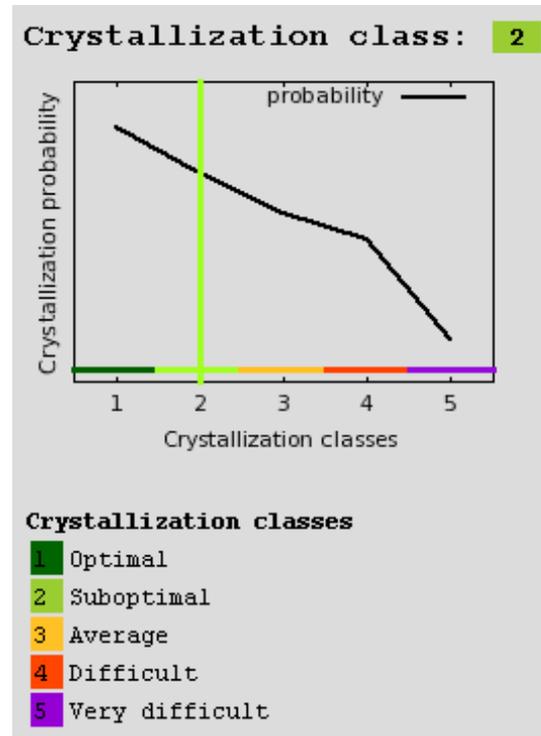


Combinando le probabilità individuali di cristallizzazione calcolate per le classi di parametri,

il target viene assegnato a una delle cinque classi di cristallizzazione.

Espressione di proteine ricombinanti per studi strutturali

Predizione del successo della cristallizzazione



#	Target id (link to target details)	EP-class	RF-class	Length	Gravy index	Instability index (II)	Isoelectric point (pI)	Coiled coils	Longest disorder region	Percentage of coil structure	Transmembrane helices (TM)	Signal peptides (SP)	Insertions score	Homologs in NR (clustered to 60% seq. ident.)	Homologs in PDB
1	sp Q14118	5	11	624	-0.39	45.65	9.31	0	234	69	No	No	0.05	871	227

EP crystallization classes [?](#)

↓ Most promising



↑ Least promising

RF crystallization classes [?](#)

↓ Most promising

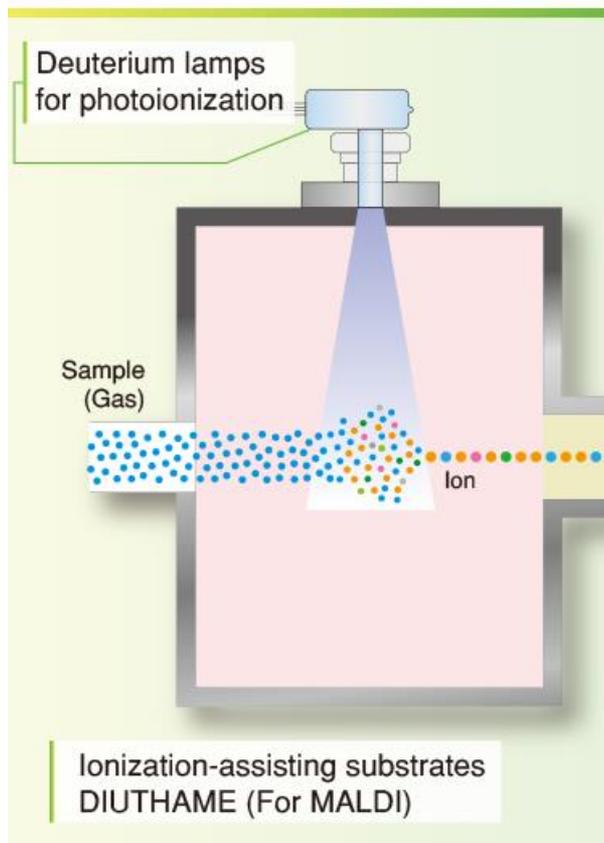


Least promising ↑

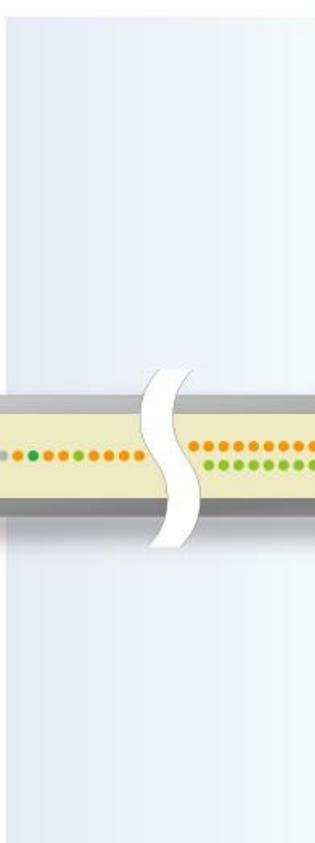
Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa

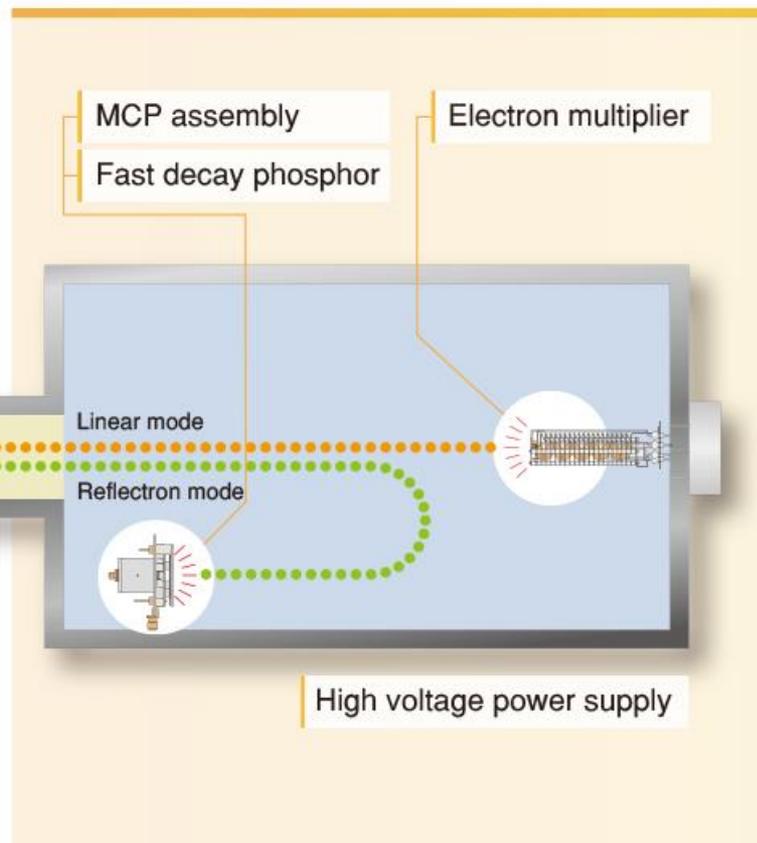
Ionization



Separation



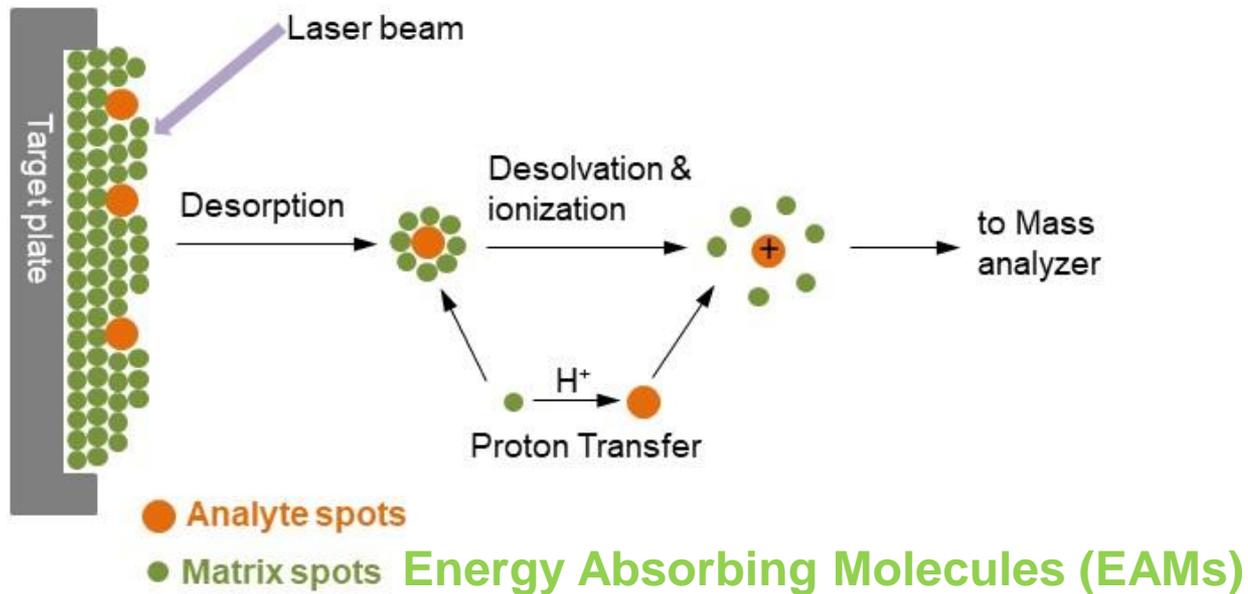
Detection



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

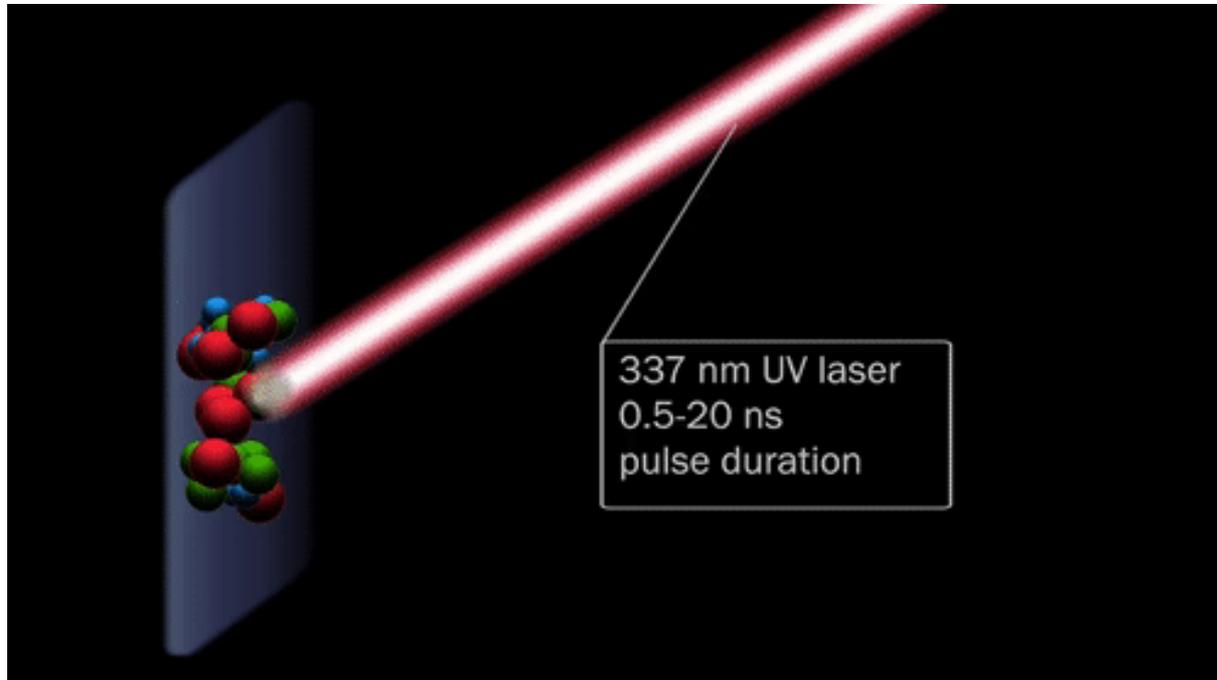
MALDI (Matrix Assisted Laser Desorption Ionization)



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

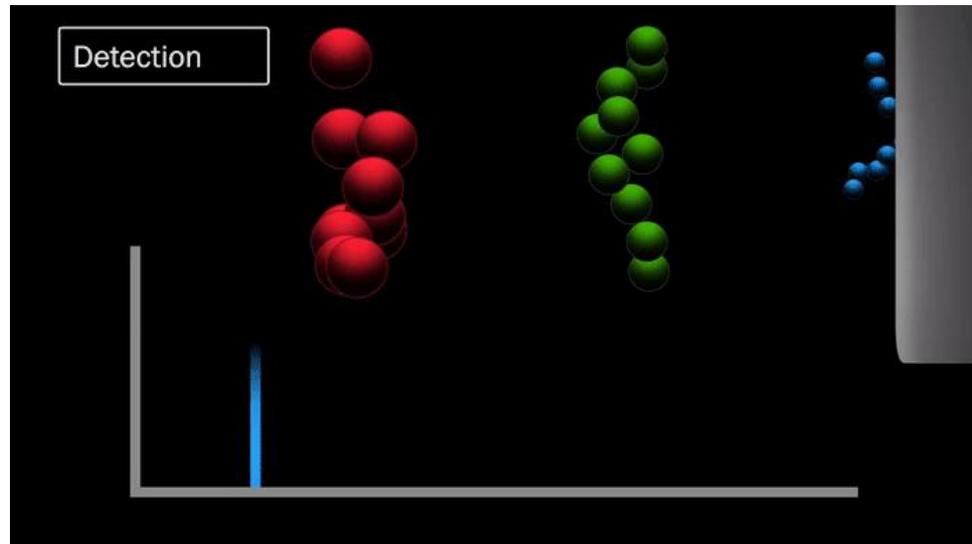
MALDI (Matrix Assisted Laser Desorption Ionization)



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

MALDI (Matrix Assisted Laser Desorption Ionization)



Matrici

CHCA (alpha-Cyano-4-hydroxycinnamic)

SA (sinapinic acid)

Super DHS

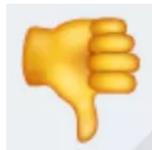
MW < 10kDa

> 10kDa

glicoproteine



Semplice
MW alti

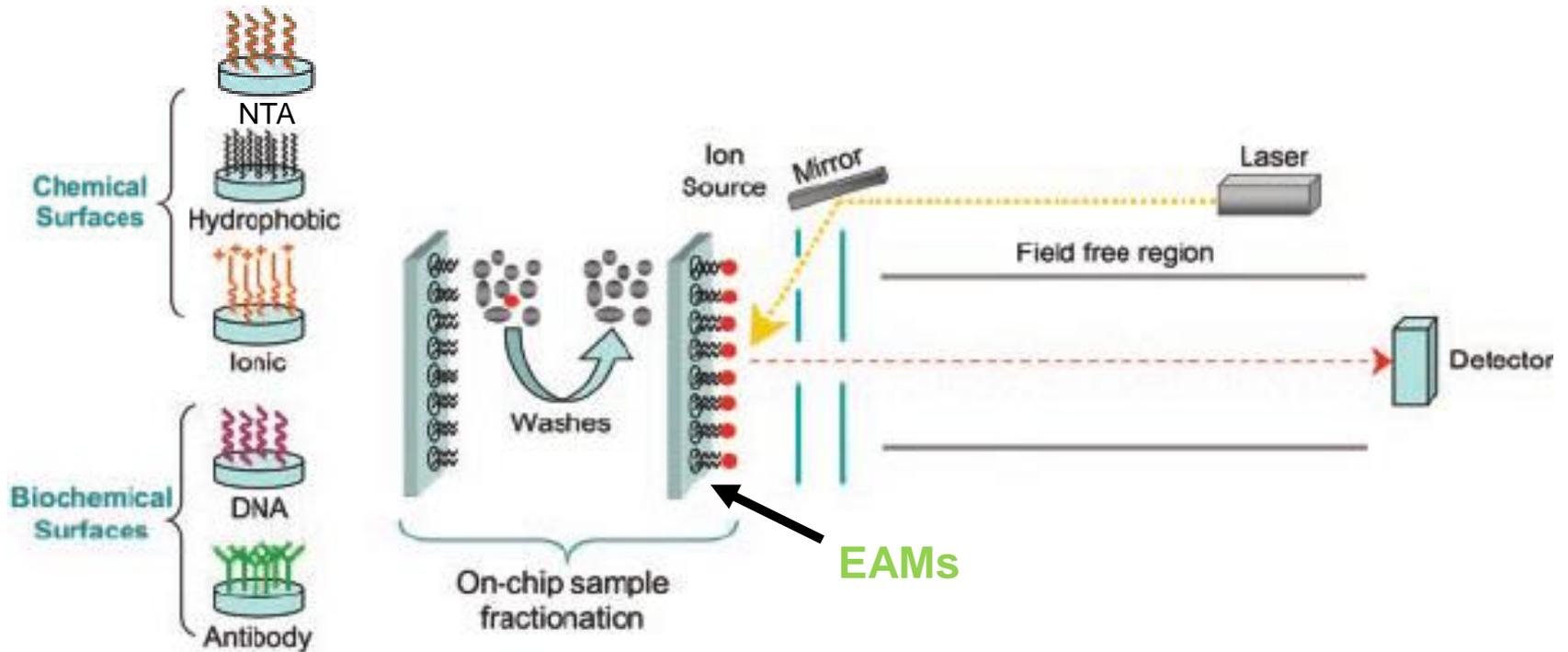


interferenza fra 50-500 m/z (non quantitativa)
distribuzione non omogenea

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

SELDI (Surface Enhanced Laser Desorption Ionization)

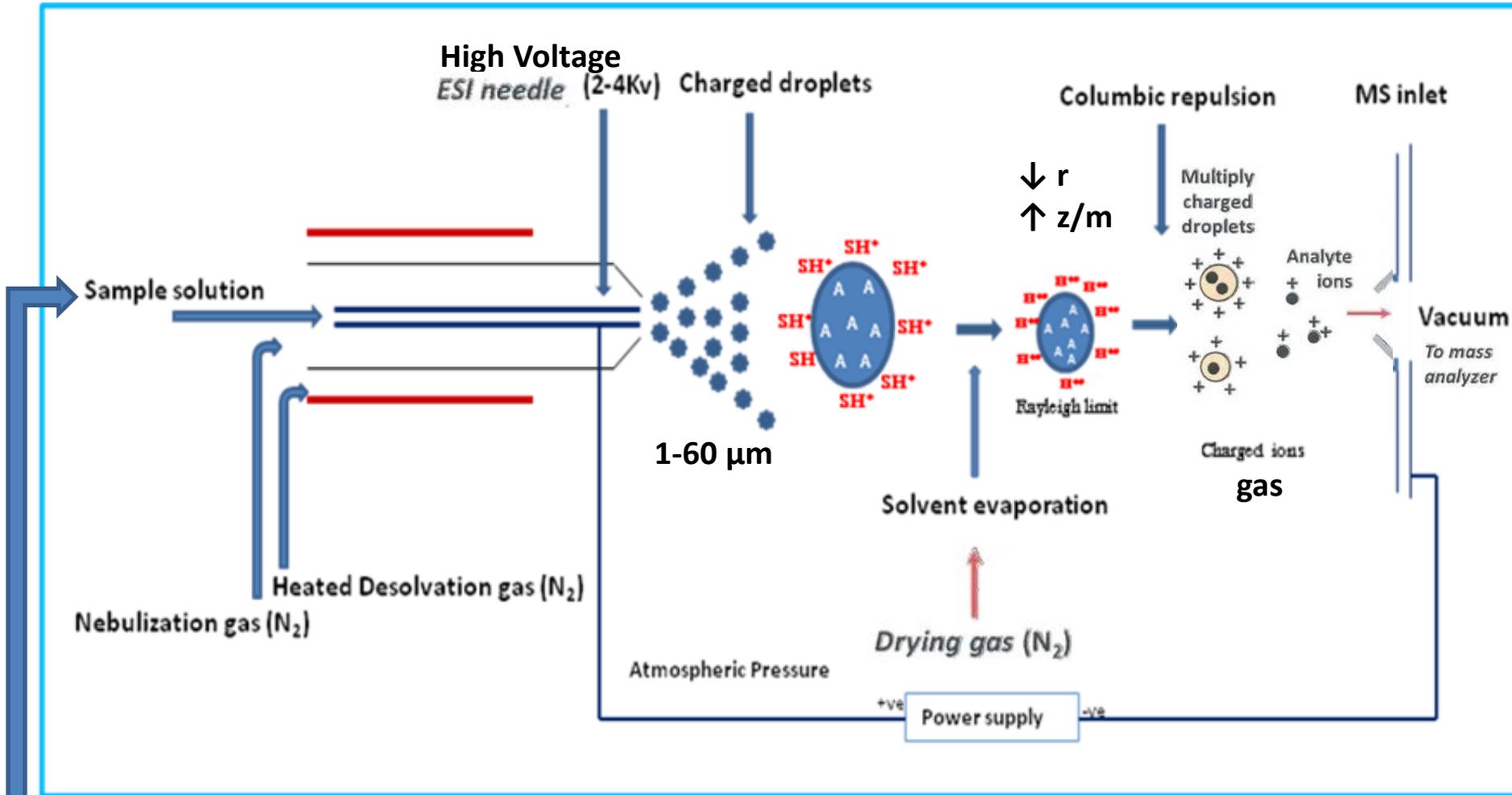


campioni biologici

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

ESI (ElectroSpray Ionization)



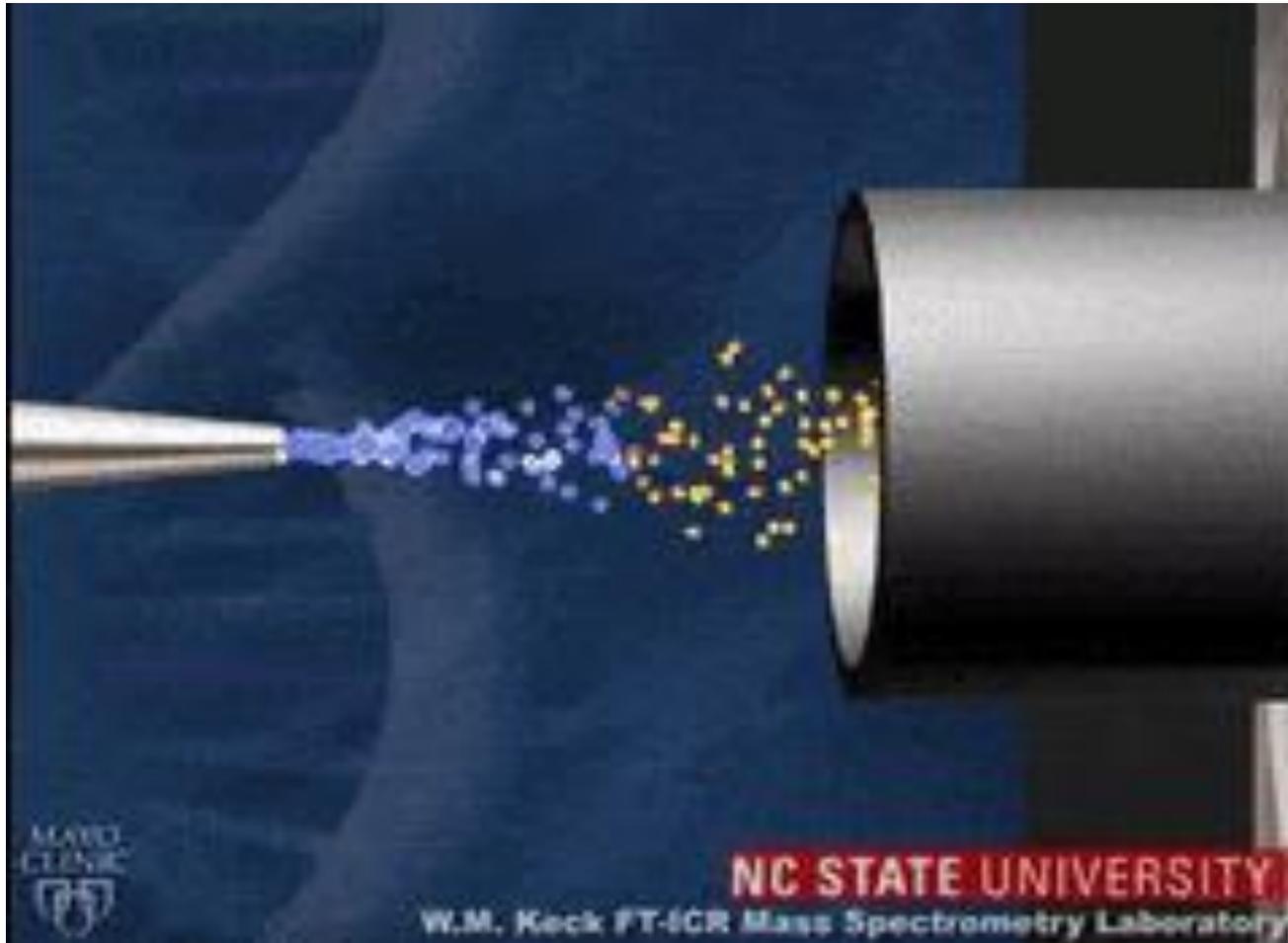
HPLC capillare (5-10 μl/min; r=0.25-0.33 mm)

nano (0.07-0.15 μl/min; r=0.075 mm) → cala r di 100-1000x → 100% ionizzazione

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: ionizzazione

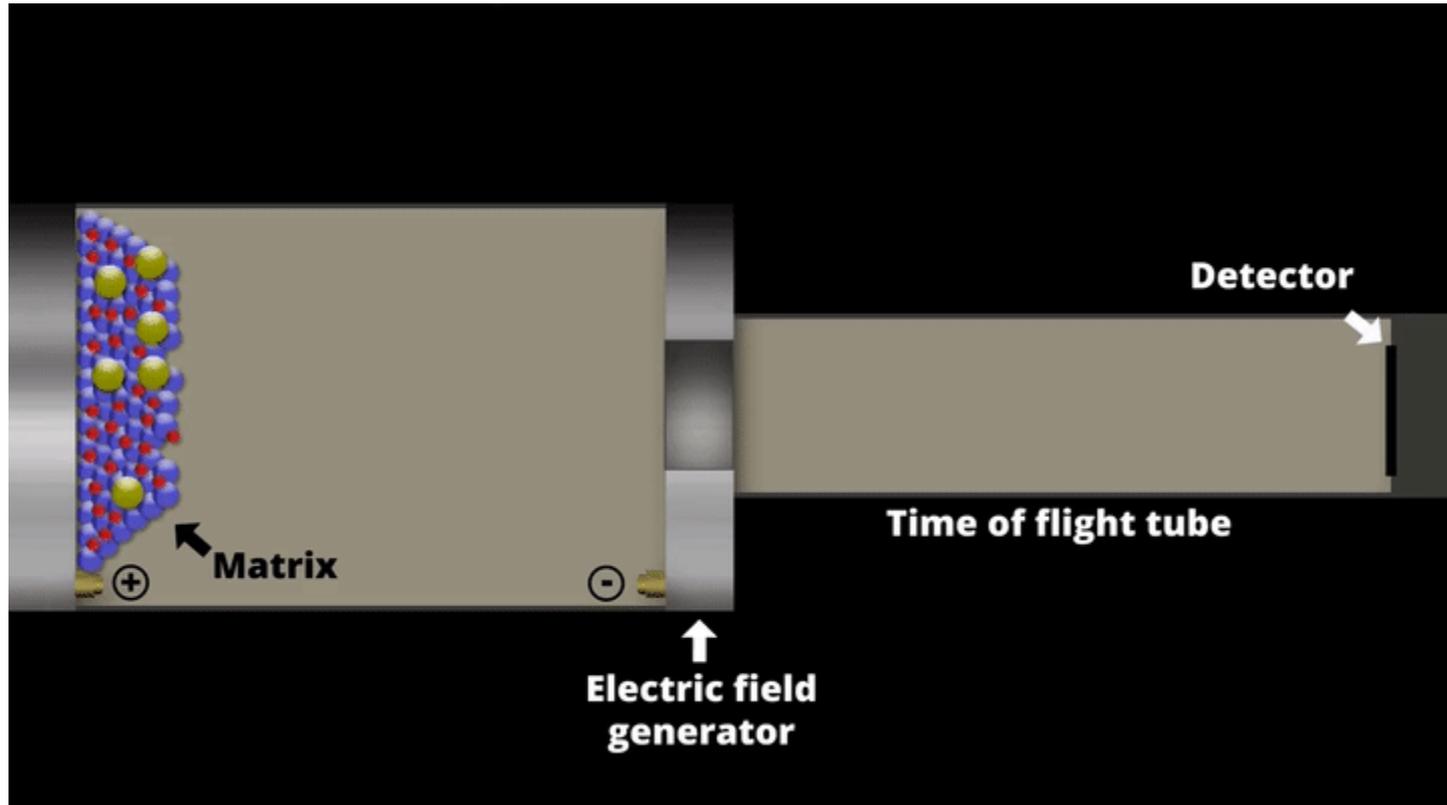
ESI (ElectroSpray Ionization)



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

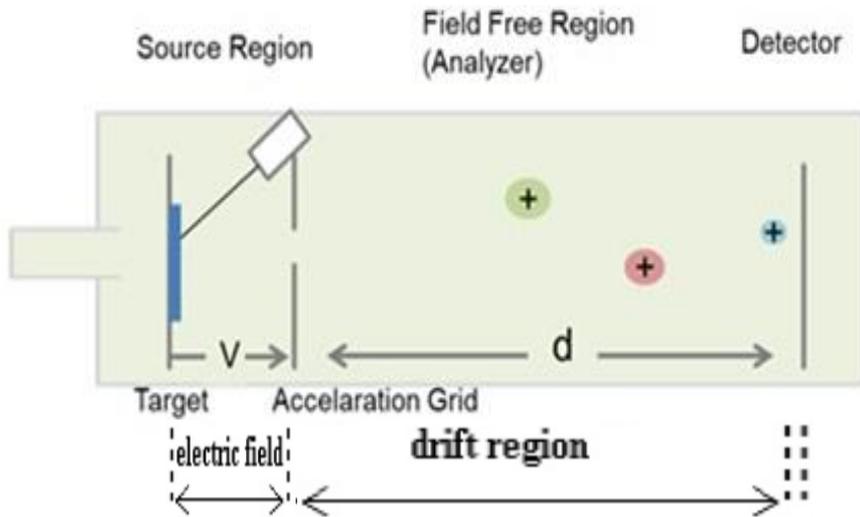
TOF (Time Of Fly)



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

TOF (Time Of Fly)



The potential energy :

$$E = zeV$$

The kinetic energy of ions:

$$E = 1/2mv^2$$

The potential energy is converted into kinetic energy:

$$zeV = 1/2mv^2$$

Thus,

$$v = \sqrt{\frac{2zeV}{m}}$$

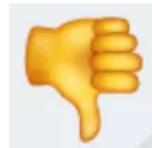
The time taken by ions to reach the detector: $t = \frac{d}{v}$

Finally, we get:

$$t = d \sqrt{\frac{m}{2zeV}}$$



MW 60-80 kDa

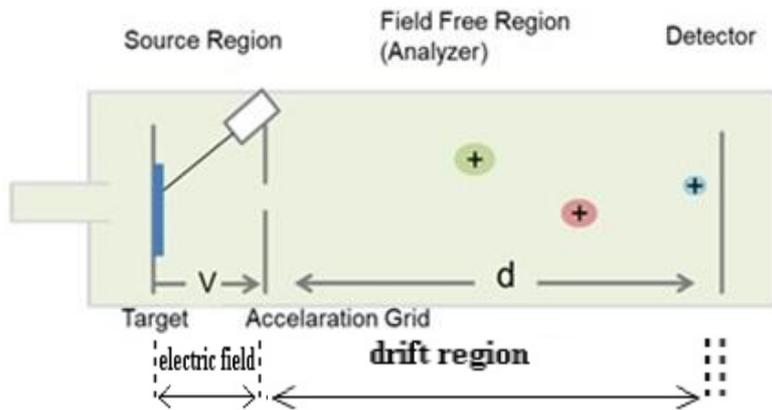


Low resolution

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

TOF (Time Of Fly)



The potential energy :

$$E = zeV$$

The kinetic energy of ions:

$$E = 1/2mv^2$$

The potential energy is converted into kinetic energy:

$$zeV = 1/2mv^2$$

Thus,

$$v = \sqrt{\frac{2zeV}{m}}$$

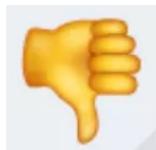
The time taken by ions to reach the detector: $t = \frac{d}{v}$

Finally, we get:

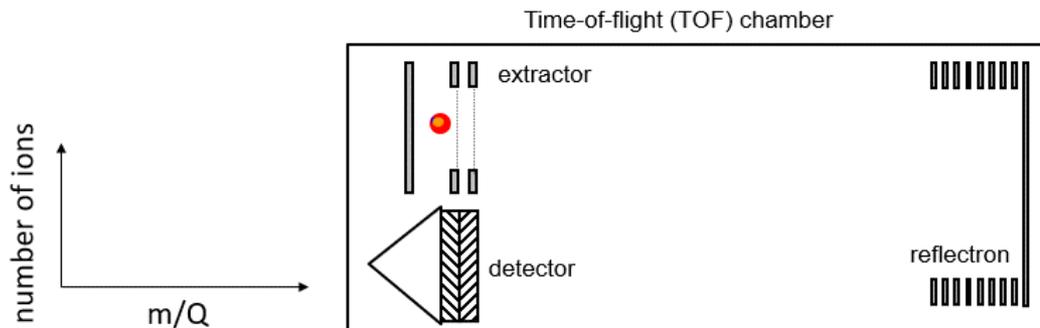
$$t = d \sqrt{\frac{m}{2zeV}}$$



MW 60-80 kDa



Low resolution



MW 5-10 kD

High resolution

Produzione di proteine ricombinanti per studi strutturali

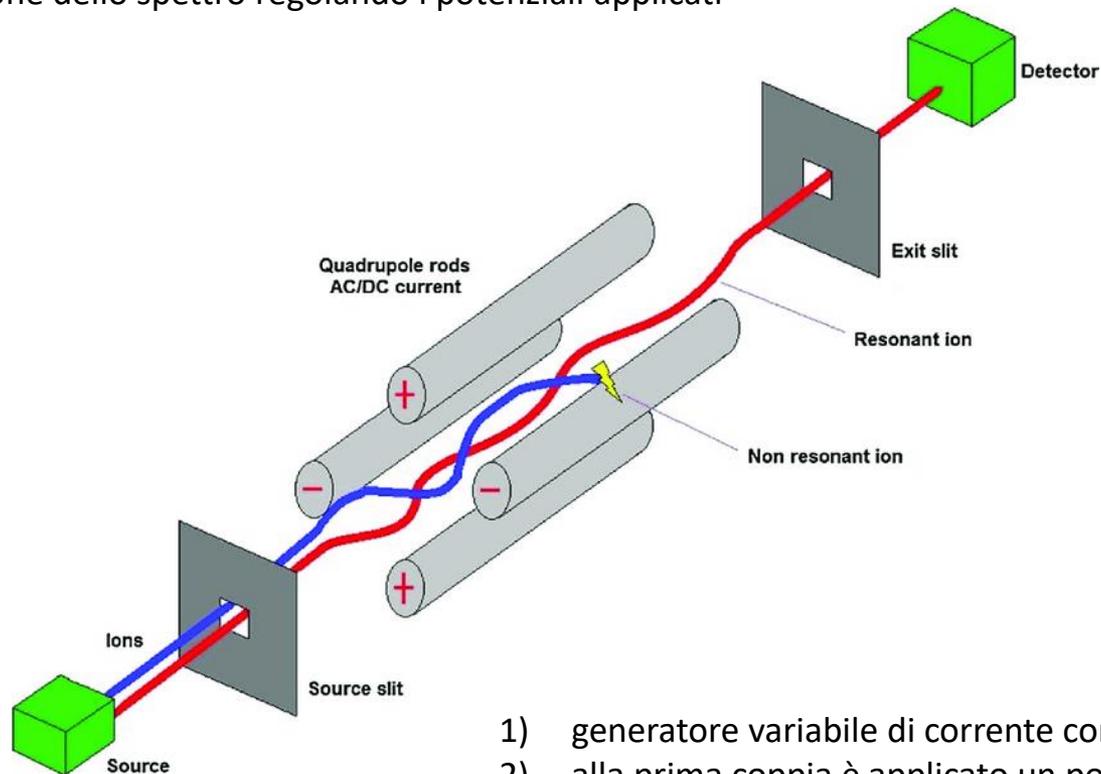
Spettrometria di massa: analizzatori

Quadrupole mass analyser

filtro di massa a banda stretta:

trasmette gli ioni entro un ristretto $[m/z]$

scansione dello spettro regolando i potenziali applicati

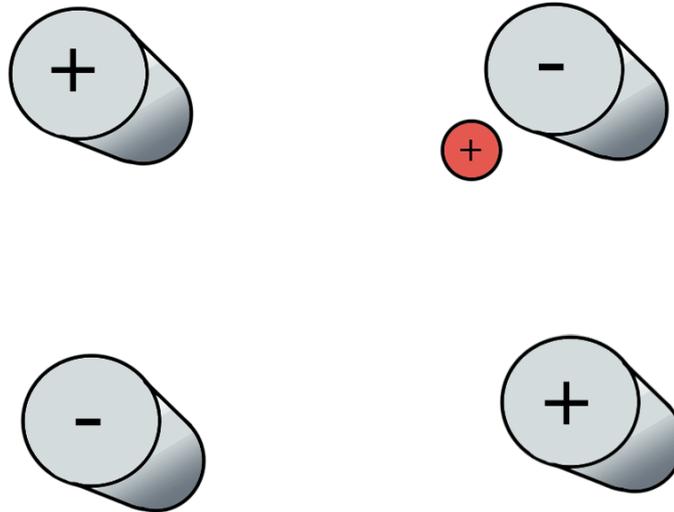


- 1) generatore variabile di corrente continua (DC)
- 2) alla prima coppia è applicato un potenziale variabile in corrente alternata a radiofrequenza (RF) e alla seconda coppia un segnale sfasato di 180°

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

Quadrupole mass analyser

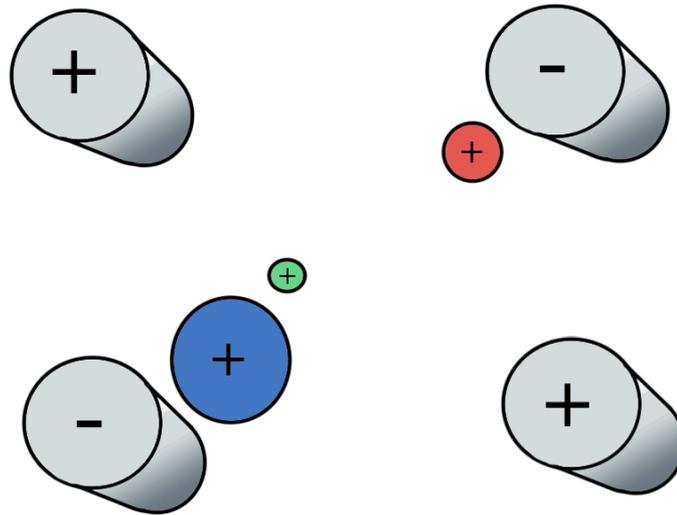


Ogni ione con un dato m/z avrà un percorso stabile attraverso il quadripolo in date condizioni (determinato dalle ampiezze RF e DC).

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

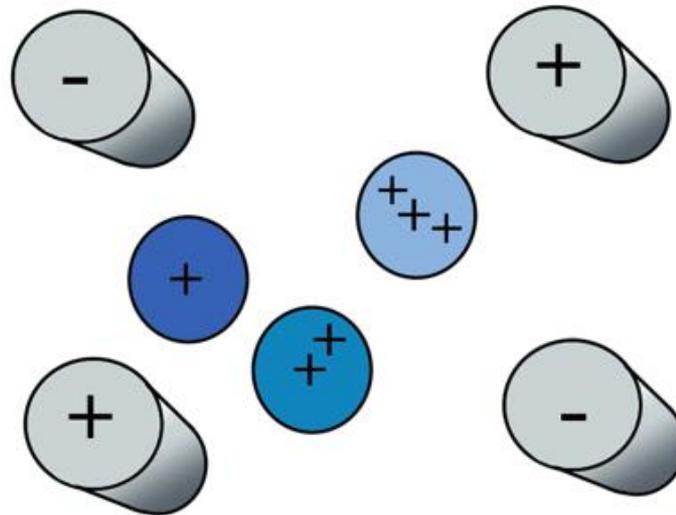
Quadrupole mass analyser



Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

Quadrupole mass analyser



m/z determina il modo in cui uno ione si muove in un campo elettromagnetico.

Stesso ione blu presente in tre diversi stati di carica:

2+ si comporterebbe come se avesse la metà della massa

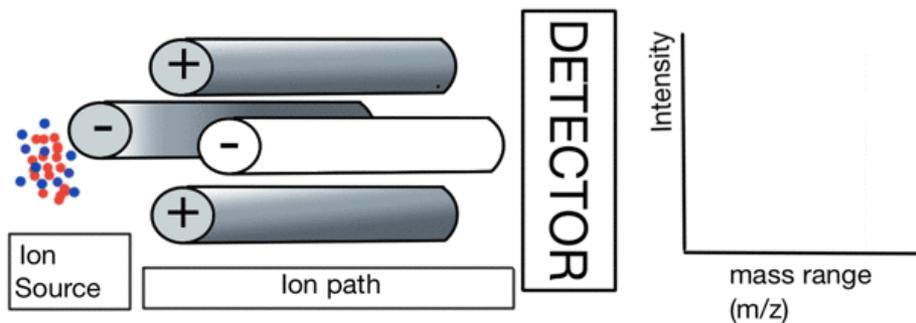
3+ sembrerebbe avere un terzo della massa dello ione 1+.

Produzione di proteine ricombinanti per studi strutturali

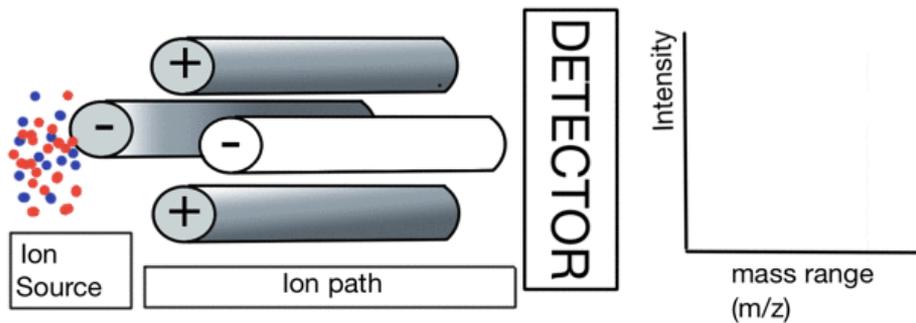
Spettrometria di massa: analizzatori

Quadrupole mass analyser

Under quadrupole conditions "one" only red ions are detected.



Under quadrupole conditions "two" only blue ions are detected.

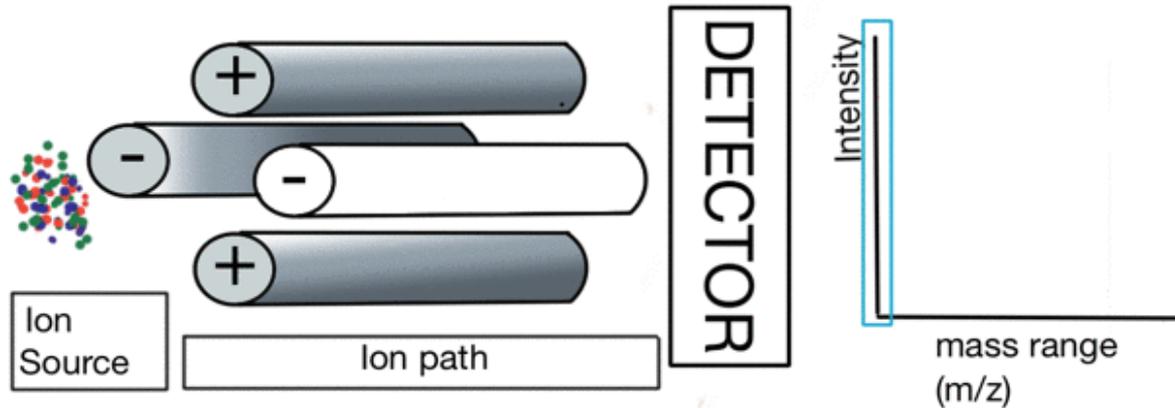


Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

Quadrupole mass analyser

The quadrupole **SCANS** through a range of conditions...



... that lets ions of a specific mass to charge ratio reach the detector.

All the other ions go to waste (over a few milliseconds).

spettro cambiando sequenzialmente dell'ampiezza RF del quadrupolo per consentire a diversi ioni di raggiungere il rivelatore.



Spettri di massa classici
Riproducibilità
Sistemi relativamente piccoli ed economici.



Risoluzione limitata
Altezze dei picchi variabili in funzione di m

Produzione di proteine ricombinanti per studi strutturali

Spettrometria di massa: analizzatori

Single Focusing Magnetic Mass Analyzer

