Kód předmětu: Bi8980

MASARYKOVA UNIVERZITA

Protein expression and purification

VII. Fusion proteins and affinity purification

Lubomír Janda, Blanka Pekárová and Radka Dopitová

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

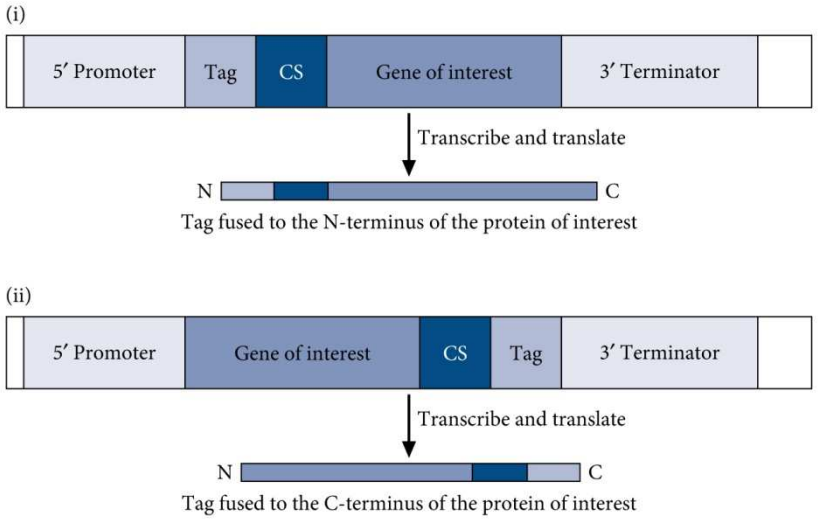
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Fusion proteins (tagged proteins)

Translation fusion of sequences coding a recombinant protein and

a) short peptide [ex. $(His)_n$, $(Asp)_n$, $(Arg)_n$ ...]

b) oligopeptide [ex. MBP, GST, thioredoxin …]



Engineering a tagged protein requires adding DNA encoding the tag to either the **5' or 3' end** of the gene encoding the protein of interest to generate recombinant protein with a tag at the **N- or C-terminus**. The stretch of amino acids containing a **target cleavage sequence** (CS) is included to allow selective removal of the tag.

Uses of fusion tags

➢**Increasing the yield of recombinant proteins –** Fusion of the N-terminus of the target protein to the C-terminus of a highly expressed fusion partner results in high level expression of the target protein.

➢ **Enhancing the solubility of recombinant proteins –** Fusion of the N-terminus of the target protein to the C-terminus of a soluble fusion partner often improves the solubility of the target protein.

➢**Improving detection –** Fusion of the target protein to either terminus of a short peptide (epitope tag) or protein which is recognized by an antibody (Western blot analysis) or by biophysical methods (e.g. GFP by fluorescence) facilitates the detection of the resulting protein during expression or purification.

➢**Localization –** A tag, usually located on the N-terminus of the target protein, which acts as an address for sending a protein to a specific cellular compartment.

➢**Facilitating the purification of recombinant proteins –** Simple purification schemes have been developed for proteins used at either terminus which bind specifically to affinity resins.

Fusion partner (tag)	Size	Tag placement	Uses
His-tag	6, 8, or 10 aa	N- or C-terminus	Purification, detection
Thioredoxin	109 aa (11.7 kDa)	N- or C-terminus	Purification, solubility enhancement
Calmodulin-binding domain (CBD)	26 aa	N- or C-terminus	Purification
Avidin/streptavidin *Strep*-tag	8 aa	N- or C-terminus	Purification, secretion
Glutathione *S*-transferase (GST)	26 kDa	N-terminus	Purification, solubility enhancement
Maltose binding protein (MBP)	396 aa (40 kDa)	N- or C-terminus	Purification, solubility enhancement
Green fluorescent protein (GFP)	220 aa (27 kDa)	N- or C-terminus	Localization, detection, purification
Poly-Arg	5-16 aa	N- or C-terminus	Purification, solubility enhancement
N-utilization substance A (NusA)	495 aa (54.8 kDa)	N-terminus	Solubility enhancement

Advantages and disadvantages of used fusion partners

Proteins do not naturally lend themselves to high-throughput analysis because of their diverse physiological properties. Affinity tags have become indispensable tools for structural and functional proteomics.

No single tag is ideally suited for all purposes. Therefore, combinatorial tagging might be the only way to harness the full potential of affinity tags in a high-throughput setting.

Because affinity tags have potential to interfere with structural and functional studies, provisions must also be made for removing them.

Tag[a]	Advantages	Disadvantages
GST	Efficient translation initiation	High metabolic burden
	Inexpensive affinity resin	Homodimeric protein
	Mild elution conditions	Does not enhance solubility
MBP	Efficient translation initiation	High metabolic burden
	Inexpensive affinity resin	
	Enhances solubility	
	Mild elution conditions	
NusA	Efficient translation initiation	High metabolic burden
	Enhances solubility	
	Not an affinity tag	
Thioredoxin	Efficient translation initiation	Not an affinity tag[b]
	Enhances solubility	
Ubiquitin	Efficient translation initiation	Not an affinity tag
	Might enhance solubility	
FLAG	Low metabolic burden	Expensive affinity resin
	High specificity	Harsh elution conditions
BAP	Low metabolic burden	Expensive affinity resin
	Mild elution conditions	Variable efficiency of enzymatic biotinylation
	Provides convenient means of immobilizing proteins in a directed orientation	Co-purification of *E. coli* biotin carboxyl carrier protein on affinity resin
		Does not enhance solubility
His_6	Low metabolic burden	Specificity of IMAC is not as high as other affinity methods
	Inexpensive affinity resin	
	Mild elution conditions	
	Tag works under both native and denaturing conditions	Does not enhance solubility
STREP	Low metabolic burden	Expensive affinity resin
	High specificity	Does not enhance solubility
	Mild elution conditions	
SET	Enhances solubility	Not an affinity tag
CBP	Low metabolic burden	Expensive affinity resin
	High specificity	Does not enhance solubility
	Mild elution conditions	
S-tag	Low metabolic burden	Expensive affinity resin
	High specificity	Harsh elution conditions (or on-column cleavage)
		Does not enhance solubility

[a]GST, glutathione S-transferase; MBP, maltose-binding protein; NusA, N-utilization substance A; FLAG, FLAG-tag peptide; BAP, biotin acceptor peptide; His_6, hexahistidine tag; STREP, streptavidin-binding peptide; SET, solubility-enhancing tag; CBP, calmodulin-binding peptide.
[b]Derivatives of thioredoxin have been engineered to have affinity for immobilized metal ions (His-patch thioredoxin) or avidin/streptavidin [38].

Combinatorial tagging

The aim is to get the maximum possible benefit from affinity tags.

Combinations:

Solubility-enhancing tag + purification tag: MBP + His6 tag

2x purification tag: IgG-binding domain + streptavidin-binding domain

Localization tag + purification tag: GFP + His6 tag

Localization tag + 2x purification tag + immunodetection: GFP + SBP domain + His_8 tag + c-Myc

Increasing the yield of recombinant proteins using fusion protein technology

Yield enhancing tags are proteins and peptides which can be involved in:

- increasing the efficiency of translation initiation,
- protection against proteolytic degradation, and
- helping to properly fold their partners leading to increased solubility of the target protein (*in vivo* and *in vitro*).

Increasing the yield of recombinant proteins using fusion technology

- **Increasing the efficiency of translation initiation (e.g. GST, MBP, NusA…)**

- Advantage of N-terminal tags

- Providing a reliable context for efficient translation initiation

- Ribosome efficiently initiate translation at the N-terminal methionine of the tag

- Deleterious secondary structures are more likely to occur in conjunction with short N-terminal tags because short RNA-RNA interactions tend to be more stable than long-range interactions.

- **Protection against proteolytic degradation**

- Several studies have shown that the nature of terminal residues in protein can play a role in recognition and subsequent action by proteases and in some cases affinity tags might improve the yield of recombinant proteins by rendering them more resistant to intracellular proteolysis.

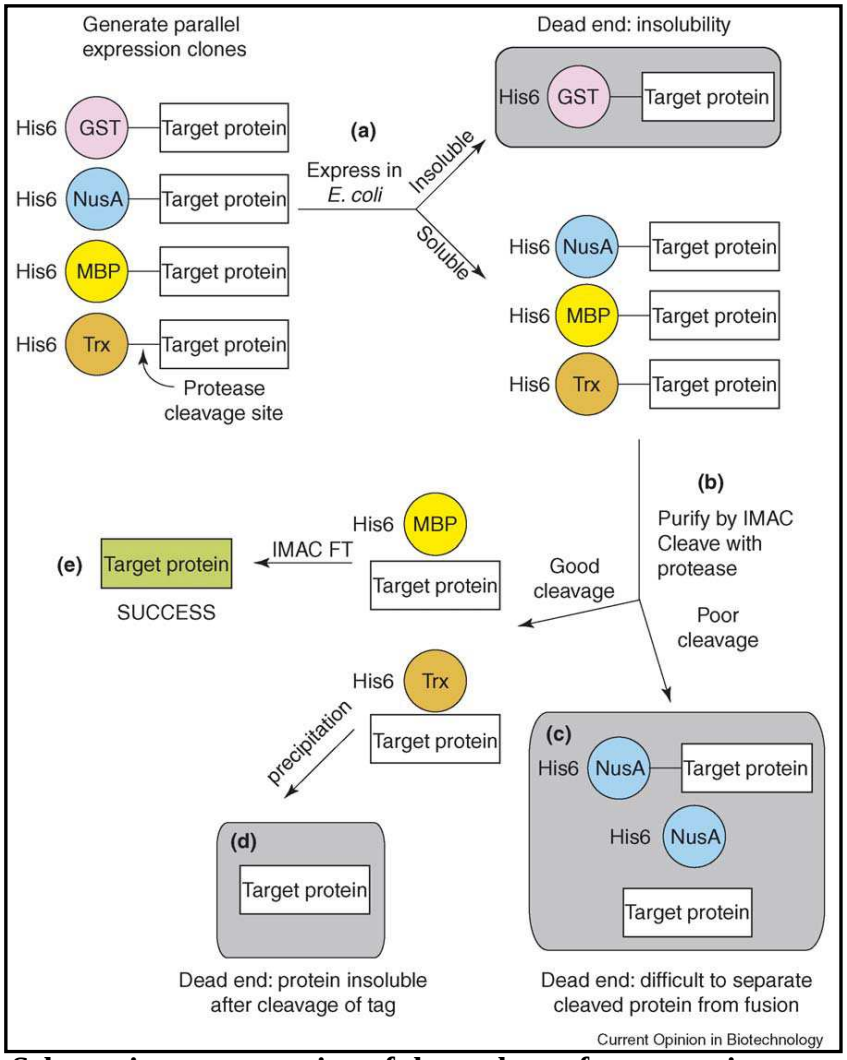
Solubility-enhancing tags

- Are generally proteins or peptides that enhance solubility and even promote the proper folding of the target proteins.

PROTEINS
GST (glutathione S-transferase), **MBP** (maltose binding protein)
- Also act as affinity tags for protein purification

NusA (N-utilization substance A), **TRX** (thioredoxin)
- Require additional affinity tags for use in protein purification

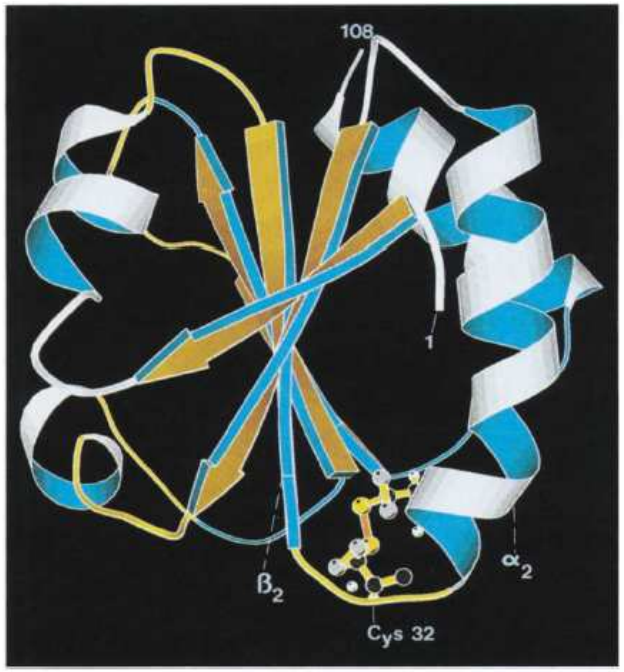
PEPTIDES
Poly-Arg (also acts as affinity tag for protein purification)
Poly-Lys



Schematic representation of the pathway from protein expression to purification using solubility tags.

The mechanism by which partners exert their solubilising function is not clear (they might act through a chaperone-like mechanism) and possibly differs between fusion proteins.

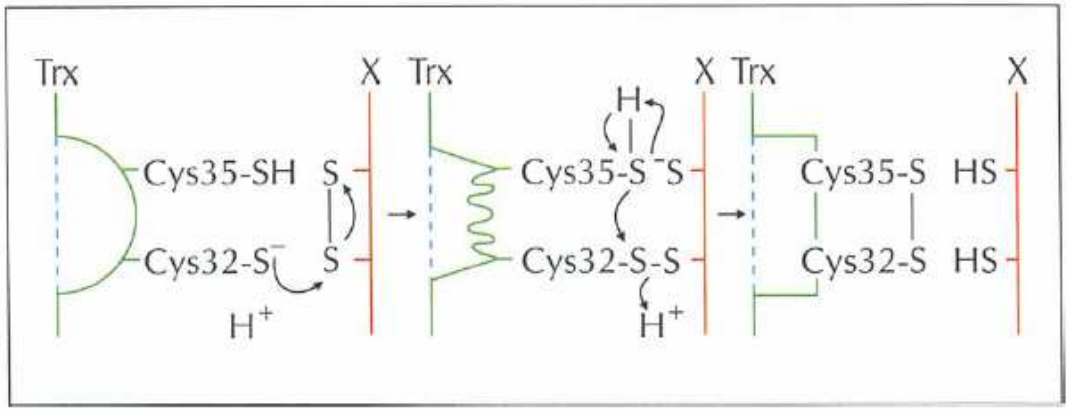
Solubility-enhancing tag - Thioredoxin



- Serves as a general protein disulfide oxidoreductase.

- Is present in all species from Archaebacteria to humans.

Folding of thioredoxin. The redox-active disulfide in the active site (Cys32-Cys35) is located on a protrusion between the strand β2 and the helix α2 . Only the sulphur of Cys32 is exposed to the solvent.



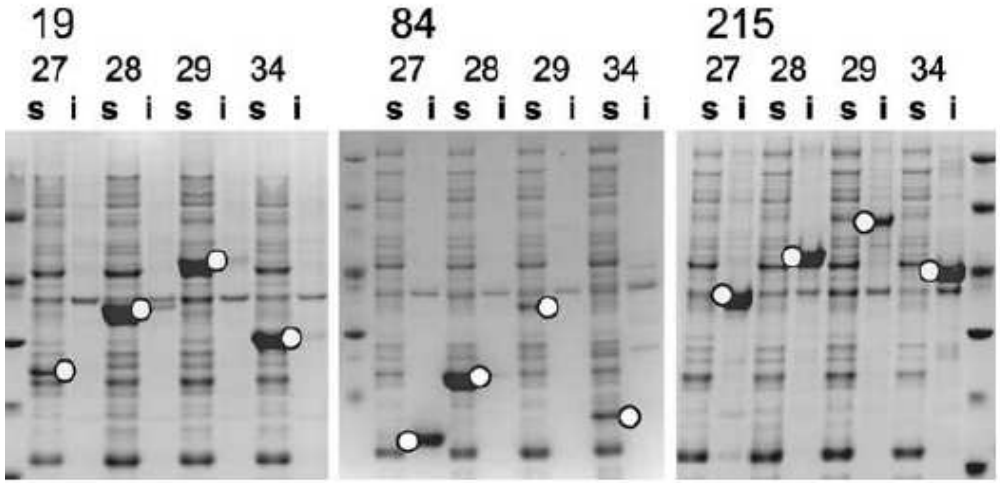
Proposed mechanism of thioredoxin-catalyzed protein disulfide reduction. Reduced thioredoxin [Trx-(SH)2] binds to a target protein via its hydrophobic surface area. Nucleophilic attack by the thiolate of Cys32 results in formation of a transient mixed disulfide, which is followed by nucleophilic attack of the deprotonated Cys35 generating Trx-S2 and the reduced protein. Conformation changes in thioredoxin and the target protein occur during the reaction.

Trx-$(SH)_2$ reduces insulin disulfides at pH 7 with a rate constant of 10^5 M^{-1} s^{-1}, which is about five orders of magnitude faster than insulin reduction by dithiothreitol (DTT), a well-known dithiol reductant.

Solubility-enhancing tag – Thioredoxin

SUMMARY

➢ The active-site surface in thioredoxin is designed to fit many proteins. Thioredoxin thus uses a chaperone-like mechanism of conformational changes to bind a diverse group of proteins and fast thiol-disulfide exchange chemistry in a hydrophobic environment to promote high rates of disulfide reduction.



19, 84, 215 – human proteins involved in cancer

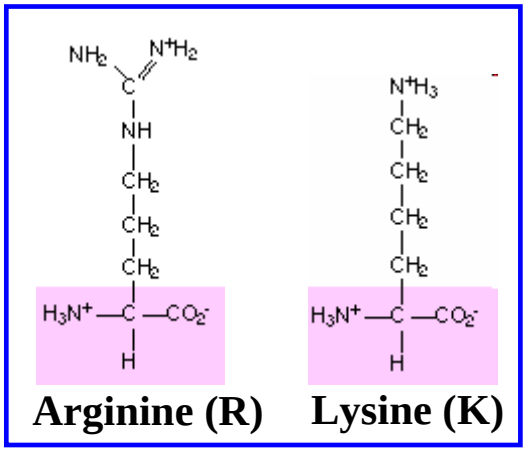
Example of SDS PAGE gels with **soluble (s) and insoluble (i) fraction** following lysis. The results when produced from the four different expression vectors **(27: His tag only; 28: thioredoxin + His tag; 29: GST + His tag; 34: GB1 + His tag)** are shown for three different target proteins (*Hammarstrom et al., 2005*).

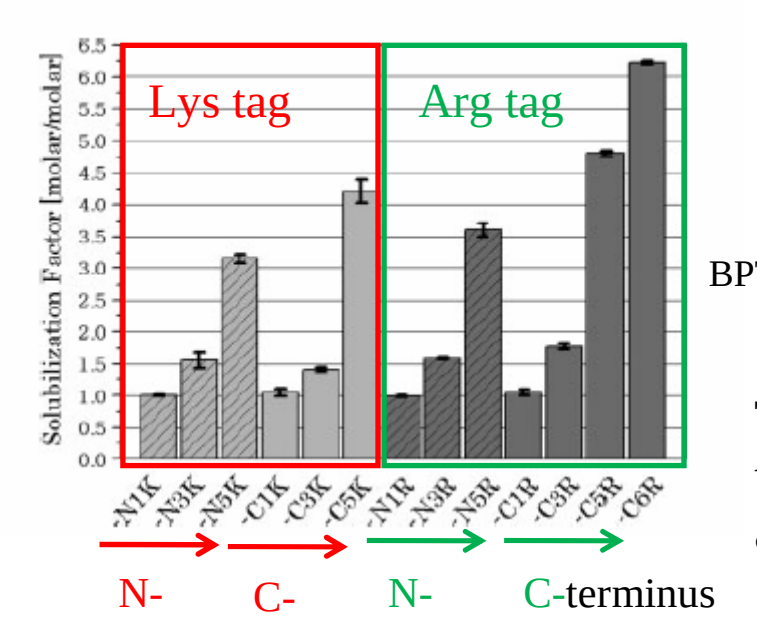
➢Increasing the solubility of the target proteins by overproduction of thioredoxin strongly suggests that the redox state affects the solubility of target proteins.

Solubility-enhancing tags – Short peptide tags

Poly-Lys tag, poly-Arg tag = one, three and five lysine or arginine residues fused to the C- or N-terminus of the target protein

Solubility as defined here is the maximum protein concentration of the supernatant after centrifugation of the supersaturated protein sample.





The solubilization factor is defined as the molar ratio between the solubility of tagged BPTI-22 variants and that of the reference BPT-22 molecule.

BPTI-22 = bovine pancreatic trypsin inhibitor variant containing 22 alanines

The solubilization effect of poly-Lys tags is lower than that of poly-Arg tags (lysines are less hydrophilic than arginines).

Charged residues seem to act through repulsive electrostatic interaction and thus hamper intermolecular interaction arising from the hydrophobic cluster.

Kato et al., 2006

Biochemical properties of poly-Arg and poly- Lys tagged BPTI-22 protein

Protein	Protein Solubility: Conc. [m*M*] (Conc. [mg/ml])[a]	Protein Solubility: Solubilization Factor[b]	T_m (°C)	Rel. Trypsin Inhibitory Activity (%)[c]
BPTI-22	1.70 (10.00)	—	38.4	—
-N1K	1.70 (10.40)	1.00 (1.04)	35.2	1.05
-N3K	2.66 (19.97)	1.56 (2.00)	34.4	1.04
-N5K	5.37 (35.60)	3.16 (3.56)	34.3	1.05
-C1K	1.79 (10.95)	1.05 (1.10)	34.6	1.05
-C3K	2.41 (15.28)	1.42 (1.53)	36.2	1.05
-C5K	7.16 (47.47)	4.21 (4.75)	35.0	1.02
-N1R	1.69 (10.34)	0.99 (1.03)	35.5	1.02
-N3R	2.70 (17.23)	1.59 (1.72)	35.6	0.99
-N5R	6.20 (41.11)	3.65 (4.11)	35.5	0.99
-C1R	1.81 (11.07)	1.06 (1.11)	35.0	1.05
-C3R	3.02 (19.26)	1.78 (1.93)	34.4	1.05
-C5R	8.23 (54.56)	4.84 (5.46)	34.8	1.08
-C6R	10.59 (73.41)	6.22 (7.34)	32.7	1.1
BPTI-22[d]	5.63 (33.11)	3.31 (3.31)	ND[e]	1.09
BPTI-22[f]	2.01 (11.82)	1.18 (1.18)	ND[e]	NA[g]

Protein solubility was determined as the maximum supernatant concentration of a supersaturated protein solution at 4°C in 100 m*M* acetate buffer pH 4.7.

[a] Maximum concentrations calculated in milligrams per milliliter are indicated in parenthesis. The Mw of BPTI-22, -N1K and -C1K, -N3K and -C3K, -N5K and -C5K, -N1R and -C1R, -N3R and -C3R, -N5R and -C5R, and -C6R are, respectively: 5880, 6123, 6379, 6636, 6151, 6463, 6776, and 6932 Da.

[b] Calculated as the ratio between the molar protein solubility of BPTI-22 and that of tagged BPTI-22. Values in parenthesis indicate the ratio calculated in milligrams per milliliters.

[c] Relative trypsin inhibitory activity calculated as the ratio between the activity of BPTI-22 and that of tagged BPTI-22. BPTI-22, which lacks R39, an arginine residue involved in two hydrogen bonding interactions with the trypsin residue backbone,[34] has a reduced trypsin inhibitor activity corresponding to ~60% of the wt-BPTI and BPTI-[5,55] at stoichiometry and a protein concentration of 280 n*M*.[19]

[d] Solubility in the same buffer as above but with the addition of 50 m*M* L-Arg + L-Glu.

[e] The CD thermal melting curve could not be determined due to the strong absorption of arginine and glutamic acid.

[f] Protein solubility with 500 m*M* Arg-HCl added to the above buffer.

[g] The trypsin activity could not be determined because the high arginine concentration inhibited trypsin activity.

The addition of 0.5 M Arg barely increased its solubility, and trypsin activity was inhibited by the high arginine concentration. On the other hand, addition of 50 mM Arg+Glu was more effective and increased protein solubility over threefold.

The addition of a poly-Arg or poly-Lys tag to the N- or C-terminus of BPTI-22 can increase its solubility without significantly affecting its structure, stability or activity.

Kato et al., 2006

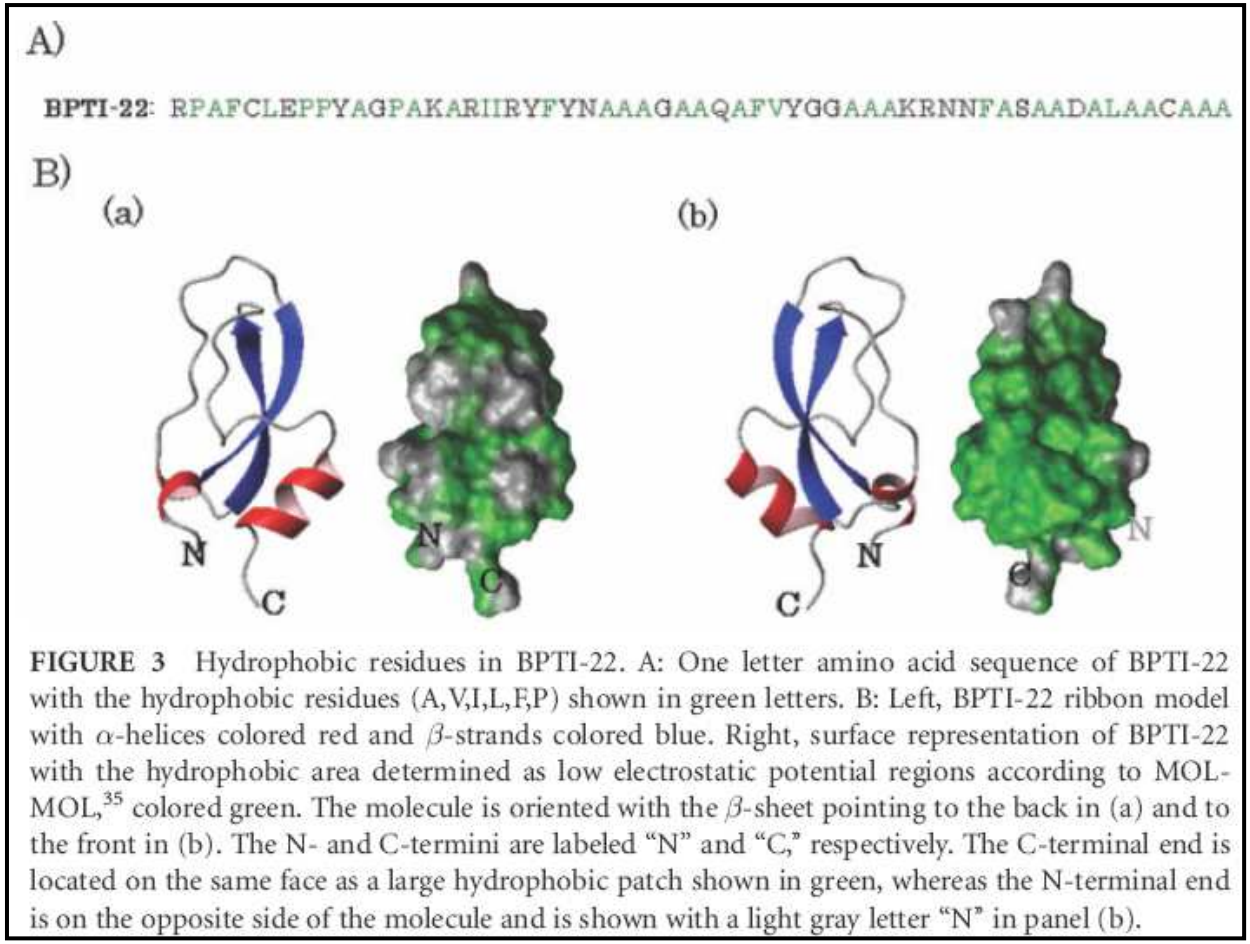


FIGURE 3 Hydrophobic residues in BPTI-22. A: One letter amino acid sequence of BPTI-22 with the hydrophobic residues (A,V,I,L,F,P) shown in green letters. B: Left, BPTI-22 ribbon model with α-helices colored red and β-strands colored blue. Right, surface representation of BPTI-22 with the hydrophobic area determined as low electrostatic potential regions according to MOLMOL,[35] colored green. The molecule is oriented with the β-sheet pointing to the back in (a) and to the front in (b). The N- and C-termini are labeled "N" and "C," respectively. The C-terminal end is located on the same face as a large hydrophobic patch shown in green, whereas the N-terminal end is on the opposite side of the molecule and is shown with a light gray letter "N" in panel (b).

The solubilization factor of all C-terminal tags was slightly higher than that of the respective N-terminal tags.

The C-terminus of BPTI-22 is close to a large hydrophobic patch, whereas the N-terminus is located on the opposite side of the molecule, away from the hydrophobic patch.

Kato et al, 2006

Solubility-enhancing tags – Comparison of peptide and protein tags, conclusions

- Protein tags are inherently large and need to be correctly folded in order to enhance solubility.

- Proteins tag are often natural affinity tags

- Peptide tags are small, and, importantly, they do not need to be folded, which provides a significant advantage over protein tags.

- The use of small tags (< 30 amino acids long) does not increase protein size substantially and diminishes steric hindrance, which simplifies downstream structural and functional applications without the need to remove the tag.

- The solubilization enhancement effect depends on the size of the target protein. Solubility enhancement of fusion partners is less pronounced for larger target proteins (above 25 kDa).

MANY TAGS SUFFER FROM THE SAME PROBLEM – THEY DO NOT FUNCTION EQUALLY WELL WITH ALL TARGET PROTEINS.

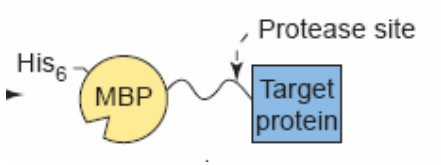
Removal of fusion tags

All tags, whether small or large, have the potential to interfere with biological activity of a protein, impede its crystallization, or otherwise influence its behavior.

1. Site-specific proteolytic cleavage

Enterokinase **Asp-Asp-Asp-Asp-Lys/X**

Amino acid in position X_1	Cleavage of enterokinase (%)
Alanine	88
Methionine	86
Lysine	85
Leucine	85
Asparagine	85
Phenylalanine	85
Isoleucine	84
Aspartic acid	84
Glutamic acid	80
Glutamine	79
Valine	79
Arginine	78
Threonine	78
Tyrosine	78
Histidine	76
Serine	76
Cysteine	74
Glycine	74
Tryptophan	67
Proline	61



TEV protease optimal cleavage sequence: **Glu-Asn-Leu-Tyr-Phe-Gln/Ser**

Thrombin **X4-X3-Pro-Arg[Lys]/X1-X2**

(X4, X3-hydrophobic residues; X1, X2-non-acidic residues)

Some frequently used recognition sites: **Leu-Val-Pro-Arg/Gly-Ser**; **Leu-Val-Pro-Arg/Gly-Phe**; **Met-Tyr-Pro-Arg/Gly-Asn.**

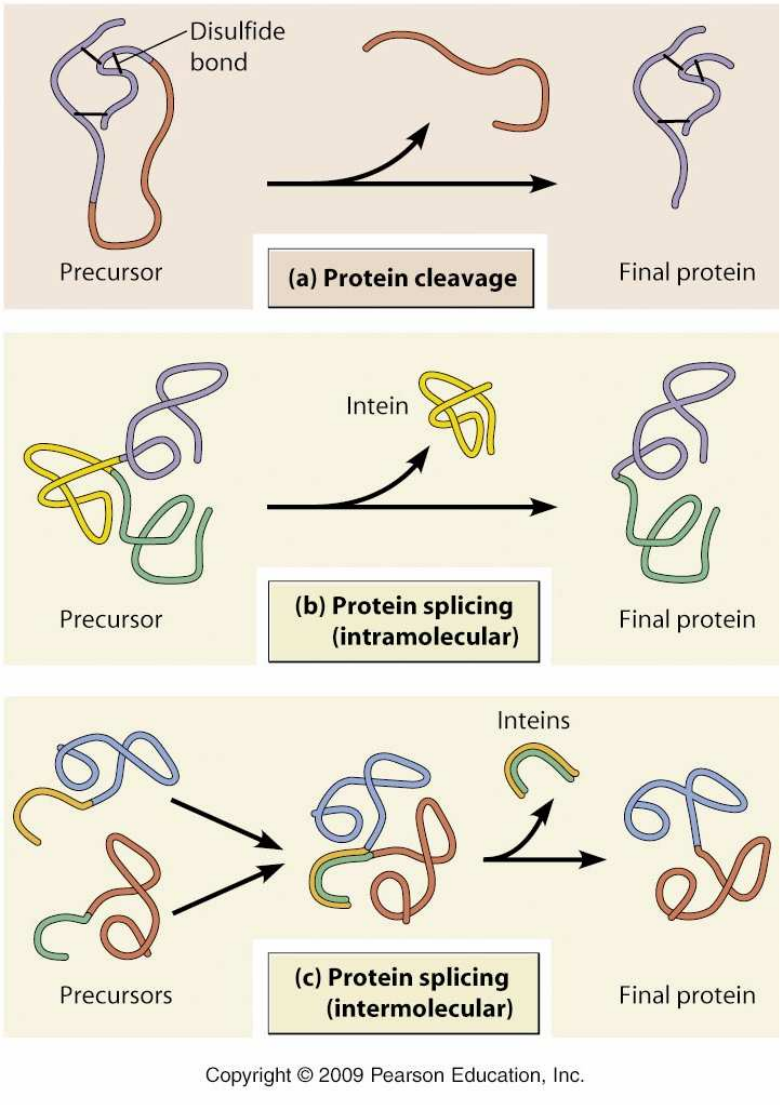
Factor Xa **Ile-Glu[Asp]-Gly-Arg/X1**

X1 can be any amino acid except arginine and proline.

2. Use of self-processing fusion partners derived from self-splicing inteins

(analogy to RNA splicing)

Inteins are selfish DNA elements inserted in-frame and translated together with their host proteins. This precursor protein undergoes an autocatalytic protein splicing reaction. **The process of protein splicing** removes inteins and splices the exteins together to make a mature protein.



The principle disadvantages of the intein approach:

- the large size of the catalytic machinery that must be incorporated into the fusion protein, which increases the metabolic burden on the cells
- the dependence of processing efficiency on the sequence context at the fusion junction
- the slow rate of auto-processing
- the fact that inteins neither enhance the solubility nor facilitate the purification of their fusion partners

3. Chemical cleavage

- Rarely used

Cyanogen bromide Met/X

Hydroxylamine Asn-Gly

Chemical cleavage is a harsh method leading to non-specific cleavage, whereas enzymatic cleavage can be specific but inefficient.

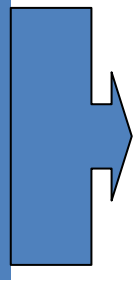
Removal of fusion tags – Achilles' heel of the fusion approach

1. **Unspecific cleavage** (SOLUTION: optimization of protein cleavage conditions or using re-engineered proteases with increased specificity such as ProTEV and AcTEV)
2. **Optimization of protein cleavage conditions** (mainly enzyme-to-substrate ratio, temperature, pH, salt concentration, length of exposure)
3. **Precipitation of target protein when the fusion partner is removed** (so-called soluble aggregates; SOLUTION: other approach for protein solubilization has to be found)
4. **Cleavage efficiency** (varies with each fusion protein in an unpredictable manner, probably due to aggregation or steric issues; the problem can be solved by introducing short linkers between the protease site and the fusion tag)
5. **High cost of proteases**
6. **Re-purification step**
7. **Failure to recover active or structurally intact protein**

The alternative is to leave the affinity tag in place for structural analysis:

These multi-domain proteins are usually

- less conducive to forming well-ordered diffraction crystals, presumably due to the conformational heterogeneity allowed by the flexible linker region
- too large for NMR analysis



Small tags are better choice in structural biology

Unspecific His-tag cleavage

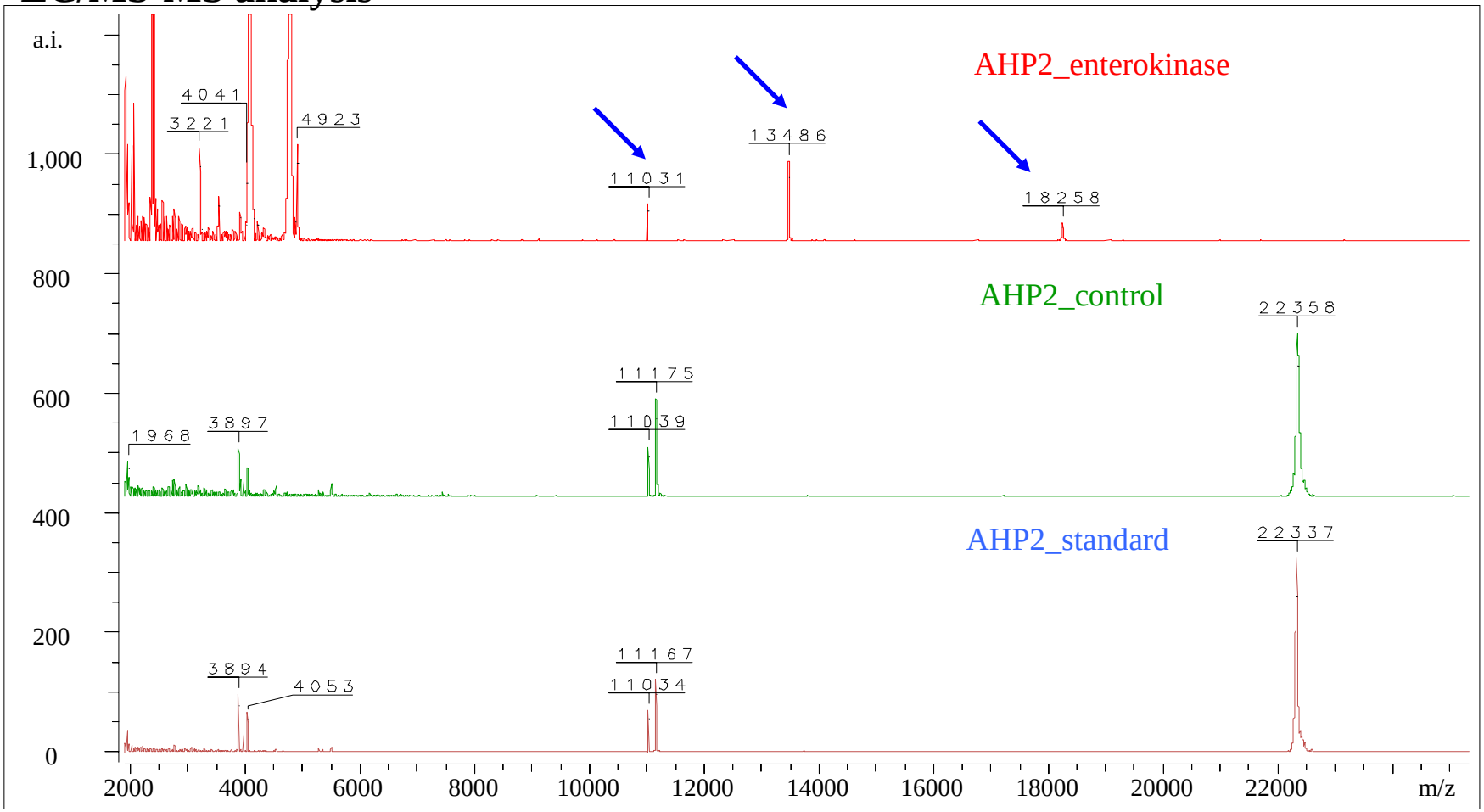
pRSETB::AHP2

Enterokinase cleavage site

N'MRGSHHHHHHGMASMTGGQQMGRDLY**DDDDK**↓**D**PSSRSAAGTMEFMDALIA….……GIVPQVDIN **C'**

Theoretically: 3.4 kDa 18.9 kDa

LC/MS-MS analysis



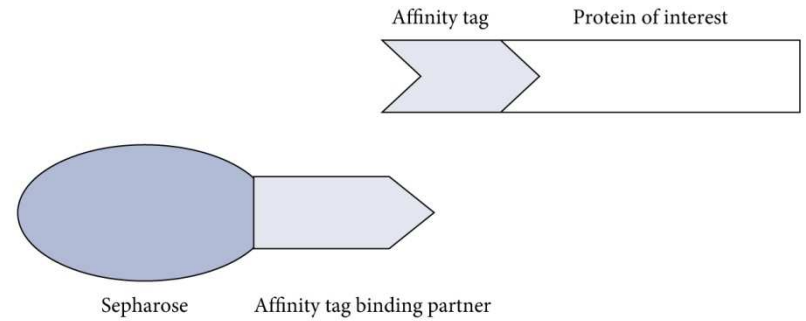
Affinity chromatography

- Affinity chromatography exploits the natural specific recognition between biological molecules.

Ligand	**Affinity to**
Enzyme	substrate analogue, inhibitor, cofactor
Antibody	antigen, virus, cell
Lectin	polysaccharide, glycoprotein, cell surface receptor, cell
Nucleic acid	complementary base sequence, histones, nucleic acid polymerase, nucleic acid binding protein.
Hormone, vitamin	receptor, carrier protein
Glutathione	glutathione-S-transferase or GST fusion proteins
Metal ions	poly (His) fusion proteins, native proteins with histidine, cysteine and/or tryptophan residues on their surfaces.

- Essential tool for protein purification (protein preparation for structural genomics, antibody generation, and biochemical analysis) and protein complex isolation.
- The essential element of affinity chromatography is the affinity ligand immobilized onto an inert, hydrophilic solid support (or matrix), which is used in a column to purify the desired target molecule.

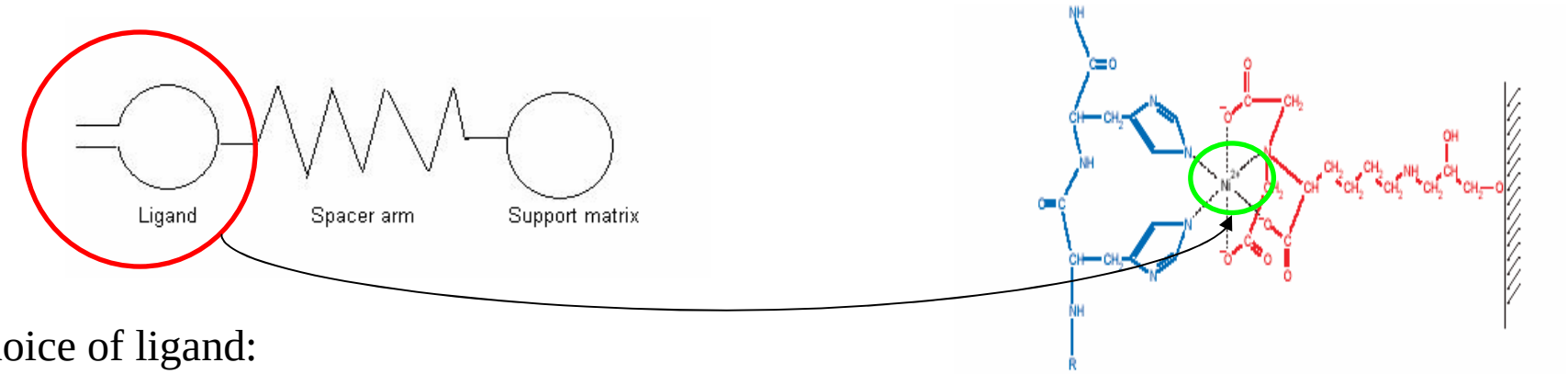
Affinity purification of tagged proteins



Immoblized binding partner of affinity tag	Affinity tag fused to N- or C-terminus of protein
TPEG (substrate analogue of β-galactosidase)	β-galactosidase
Glutathione	Glutathione-S-Transferase
Immunoglobulin G	Protein A
Cu II, Co II or Ni II	poly His or poly Cys

A tag is fused to the N- or C-terminus of the protein of interest to facilitate purification, which relies on a specific interaction between the affinity tag and its immobilized binding partner. **Genetically engineered fusion tags allow the purification of virtually any protein without any prior knowledge of its biochemical properties.**

Components of a matrix for affinity chromatography



Choice of ligand:

➢ The ligand must bind strongly (e.g., Kd < 1 nM) to the target molecule to facilitate its capture from a complex protein mixture.

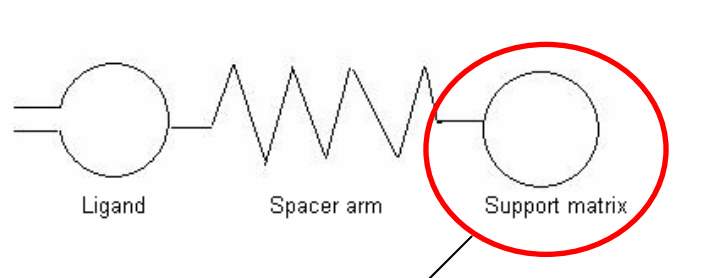
➢ Higher affinity produces better specificity, and thus better purification.

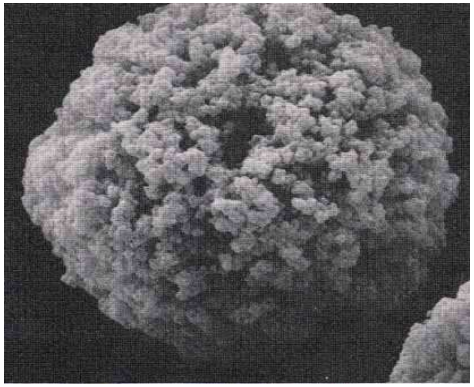
➢ When the affinity is high enough (e.g., Kd < 1 nM) , small amounts of affinity gel matrices can be used to purify proteins from large volumes of crude extracts.

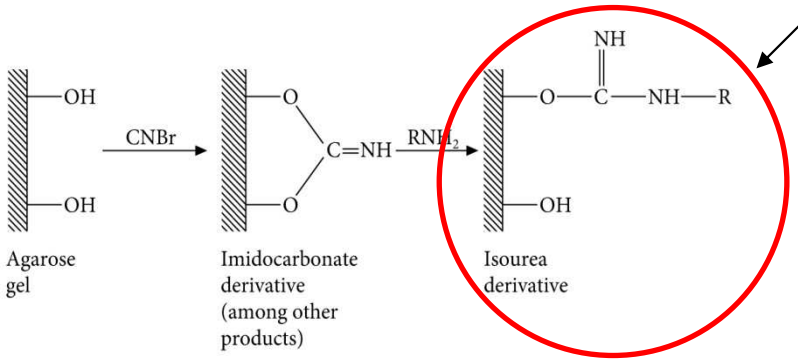
➢ When the affinity is not high enough and the protein is of low abundance, partial purification using other methods to enrich the protein of interest may be required.

➢ Affinity tag procedures are particularly useful when target proteins must be isolated from complex protein mixtures.

Components of a matrix for affinity chromatography



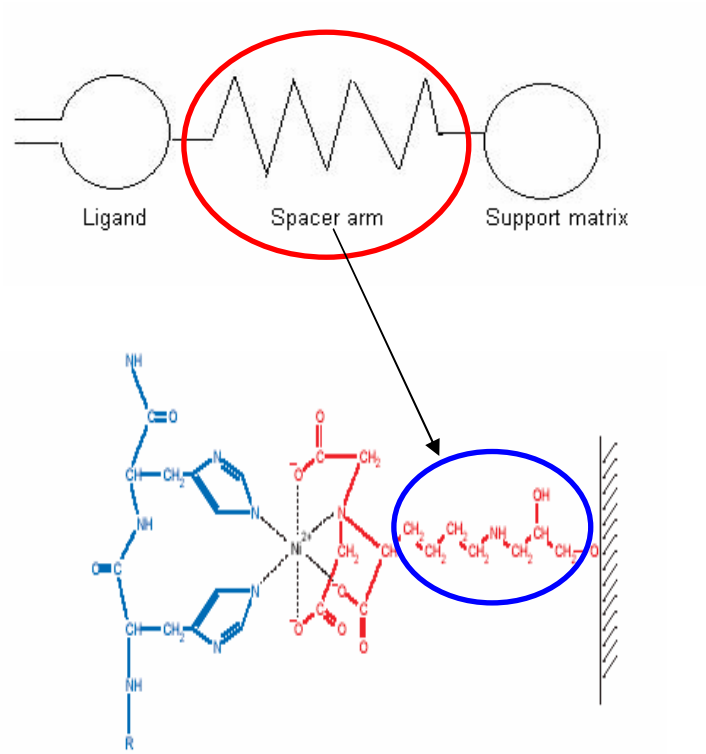




One of the most common methods for immobilizing ligands involves cyanogen bromide activation of agarose to produce imidocarbonate derivatives, which react with amino groups to generate isourea linkages.

A **matrix supports** (typically, a macroporous polysaccharide bead such as agarose) tether the active ligands and provide a porous structure so that there is an increased surface area to which the target molecule can bind. A ligand can be covalently affixed to substituent groups within the matrix (e.g., amino, hydroxyl, carbonyl, and thio groups) that are easily activated using conventional chemical methods.

Components of a matrix for affinity chromatography



Ni^{2+} NTA sepharose

A **spacer arm** will be required in cases where direct coupling of the ligand to the matrix results in steric hindrance and subsequently the target protein will fail to bind to the immobilized ligand efficiently. The introduction of a spacer arm between the ligand and the matrix minimizes this steric effect and promotes optimal adsorption of the target protein to the immobilized ligand.

Overview of tags using in affinity chromatography

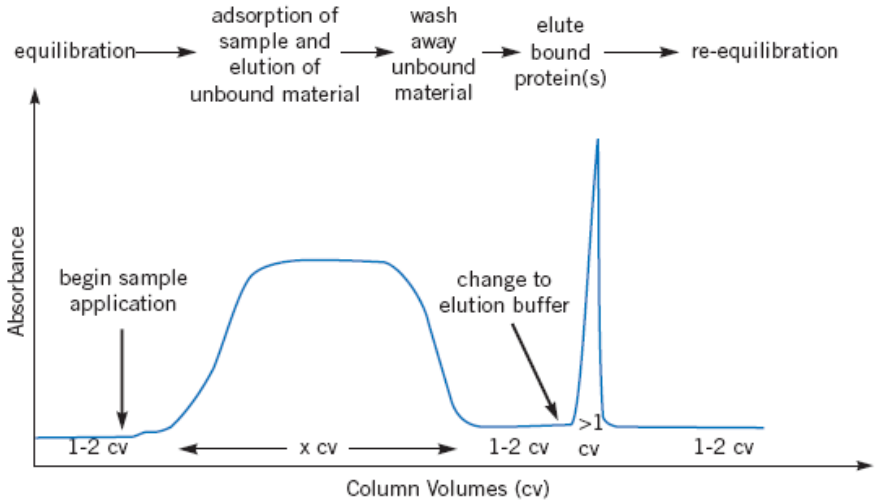
Table 1 Matrices and elution conditions of affinity tags

Affinity tag	Matrix	Elution condition
Poly-Arg	Cation-exchange resin	NaCl linear gradient from 0 to 400 mM at alkaline pH>8.0
Poly-His	Ni^{2+}-NTA, Co^{2+}-CMA (Talon)	Imidazole 20–250 mM or low pH
FLAG	Anti-FLAG monoclonal antibody	pH 3.0 or 2–5 mM EDTA
Strep-tag II	Strep-Tactin (modified streptavidin)	2.5 mM desthiobiotin
c-myc	Monoclonal antibody	Low pH
S	S-fragment of RNaseA	3 M guanidine thiocyanate, 0.2 M citrate pH 2, 3 M magnesium chloride
HAT (natural histidine affinity tag)	Co^{2+}-CMA (Talon)	150 mM imidazole or low pH
Calmodulin-binding peptide	Calmodulin	EGTA or EGTA with 1 M NaCl
Cellulose-binding domain	Cellulose	Family I: guanidine HCl or urea>4 M Family II/III: ethylene glycol
SBP	Streptavidin	2 mM Biotin
Chitin-binding domain	Chitin	Fused with intein: 30–50 mM dithiothreitol, β-mercaptoethanol or cysteine
Glutathione S-transferase	Glutathione	5–10 mM reduced glutathione
Maltose-binding protein	Cross-linked amylose	10 mM maltose

Table 2 Sequence and size of affinity tags

Tag	Residues	Sequence	Size (kDa)
Poly-Arg	5–6 (usually 5)	RRRRR	0.80
Poly-His	2–10 (usually 6)	HHHHHH	0.84
FLAG	8	DYKDDDDK	1.01
Strep-tag II	8	WSHPQFEK	1.06
c-myc	11	EQKLISEEDL	1.20
S-	15	KETAAAKFERQHMDS	1.75
HAT-	19	KDHLIHNVHKEFHAHAHNK	2.31
3x FLAG	22	DYKDHDGDYKDHDIDYKDDDDK	2.73
Calmodulin-binding peptide	26	KRRWKKNFIAVSAANRFKKISSSGAL	2.96
Cellulose-binding domains	27–189	Domains	3.00–20.00
SBP	38	MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP	4.03
Chitin-binding domain	51	TNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ	5.59
Glutathione S-transferase	211	Protein	26.00
Maltose-binding protein	396	Protein	40.00

Typical affinity purification steps



➢ In the equilibration phase, buffer conditions are optimized to ensure that the target molecules interact effectively with the ligand and are retained by the affinity medium as all other molecules wash through the column.

➢ During the washing step, buffer conditions are created that wash unbound substances from the column without eluting the target molecules or that re-equilibrate the column back to the starting conditions (in most cases the binding buffer is used as a wash buffer).

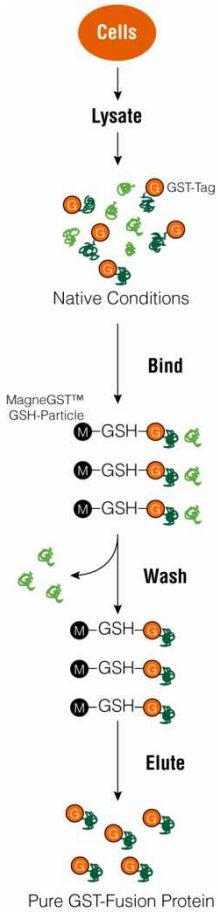
➢ In the elution step, buffer conditions are changed to reverse (weaken) the interaction between the target molecules and the ligand so that the target molecules can be eluted from the column.

Affinity chromatography

Glutathione S-transferase

GSTs perform a protective role in the cell by detoxifying endogenous compounds during oxidative stress, chemical carcinogens, environmental pollutants and a range of pharmaceutical compounds, leading to drug resistance.

- Enzymatically active fusion partner.
- GST ensures a high concentration of GST in the cell extract – this acts effectively as a purification step, i.e. GST may form 10% of total cell protein.
- One-step affinity chromatography is used.
- This chromatographic approach relies on the specificity of the interaction between GST and its substrate – glutathione.
- GST affinity column, which contains immobilized glutathione, binds GST whereas most contaminating proteins fail to bind to the column.
- Glutathione interacts with GST and promotes the specific elution of the enzyme.



Schematic representation of GST purification: Optimal binding of GST to immobilized glutathione requires a low flow rate (≤ 1 mL min^{-1}) due to the relative weak affinity of the enzyme for the substrate. Specific elution of GST is achieved by applying a solution of reduced glutathione.

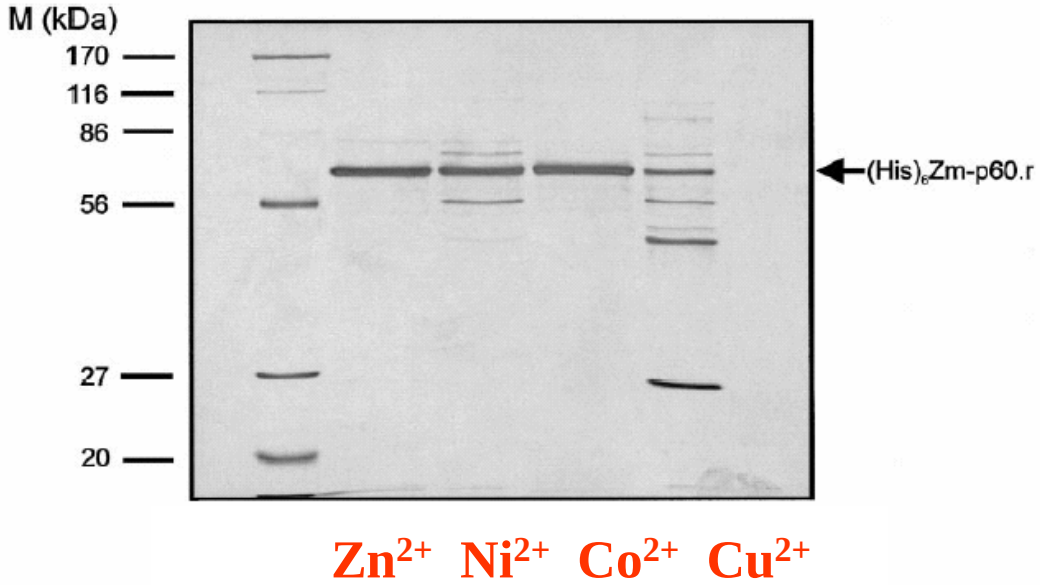
Affinity chromatography - Immobilized metal ion affinity chromatography (IMAC)

- The most common purification tag is typically composed of six consecutive histidine residues.
- Histidine, cysteine, and tryptophan residues are known to interact specifically with divalent transient metal ions such as Ni2+, Cu2+, Co2+, and Zn2+.
- Histidine is the amino acid that exhibits the strongest interaction with immobilized metal ion matrices, as the electron donor groups on the histidine imidazole ring readily form coordination bonds with an immobilized transition metal.

Binding strength of His tag

to metal ions:

$Cu^{2+} > Ni^{2+} > Zn^{2+} \sim Co^{2+}$



- IMAC can be used under native and/or denatured conditions.
- Immobilized Fe3+, Ga3+, Al3 = metal ions that have been used for selective enrichment of phosphopeptides and phosphoproteins.
- A highly purified protein can often be obtained in one or, at most, two purification steps.

His-tagged protein and IMAC under native conditions

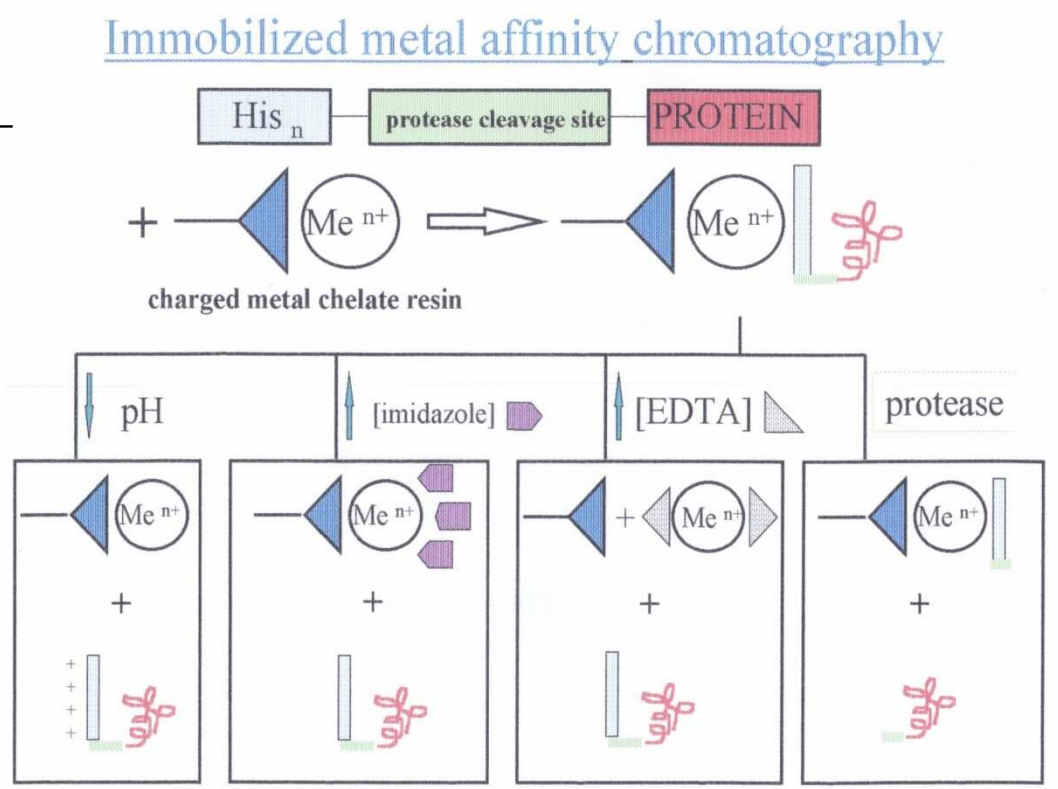
- Optimal binding of recombinant protein with metal ion is achieved at pH 7–8.

- Buffers with high salt concentration (0.5–1 M NaCl-) reduce nonspecific electrostatic interaction.

- Nonionic detergents or glycerol reduce nonspecific hydrophobic interactions.

- Elution of contaminating proteins can be achieved by lowering pH or using low concentrations of imidazole.

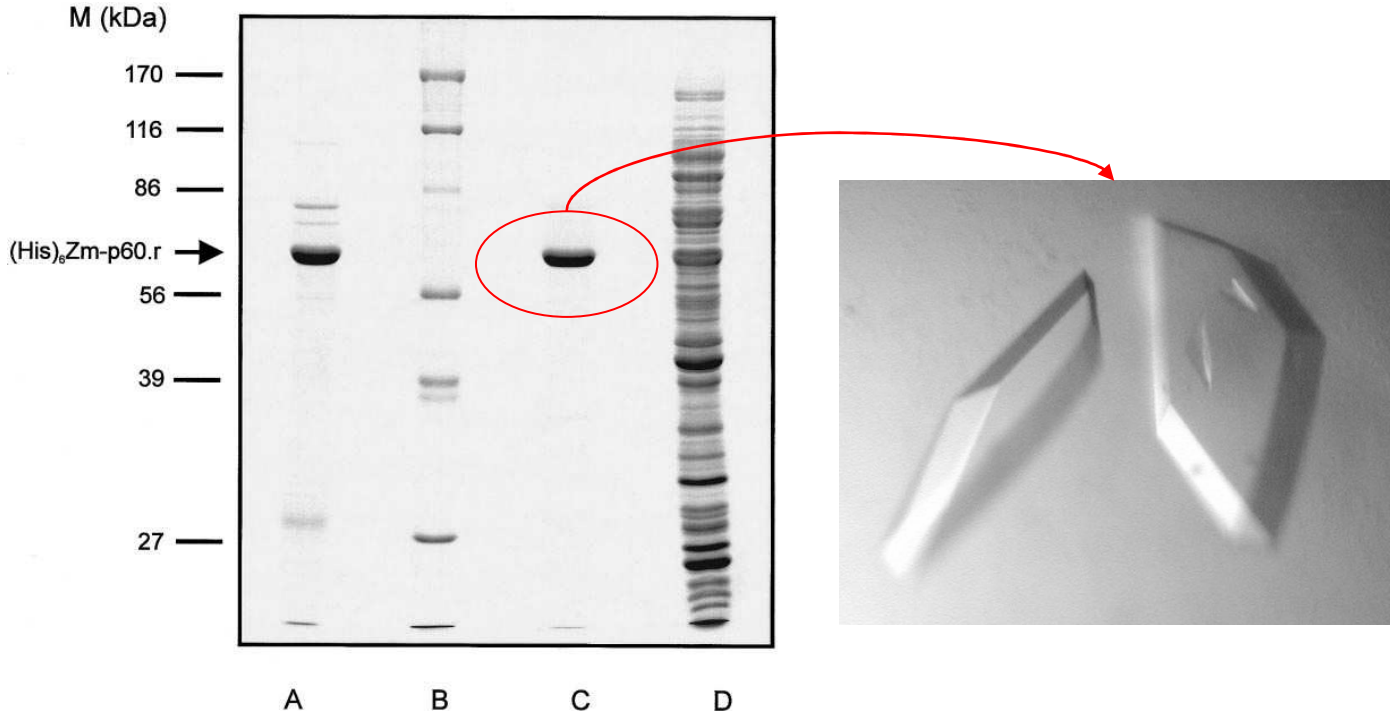
- Elution of tagged protein is achieved at high imidazole concentrations (0–0.5 M), by strongly decreasing pH, or by using EDTA.



His-tagged protein and IMAC under native conditions

One-step purification

- Perfusion matrix: POROS MC/M
- Functional group: iminodiacetate, metal ion Zn^{2+}
- Removing contaminated proteins: linear gradient of imidazole (0–50 mM) and pH (pH 6.1–7)
- Protein elution: 0.1 M EDTA
- 80% recovery, 95 fold purification
- Common production and isolation of wild type and soluble mutant form for enzymatic measurements and crystallization



His-tagged protein purification under denatured conditions

Denaturing IMAC – purification of proteins expressed in inclusion bodies
– purification in high concentration of urea or guanidine chloride
– result is pure protein, but in denatured form (sufficient for immunization)

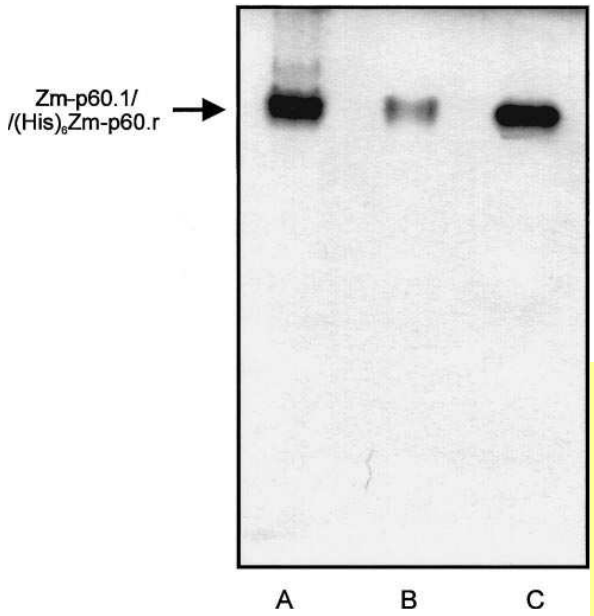
Recovery of native conformers (necessary for functional and structural analysis):

➢ binding to the column under strong denaturing conditions (8 M urea)

➢**Two renaturation possibilities**:

1. Protein is eluted from column and renatured by dialysis or rapid dilution in renaturing buffers (8-0M urea).

2. Renaturation of protein bounded to the column (matrix assisted refolding procedure): gradient from denatured to renatured buffers or pulsion renaturation.



Identification of properly refolded $(His)_6$Zm-p60.1 (maize β-glucosidase) using 10% native PAGE, followed by activity in gel staining:

A = crude protein extract prepared from maize seedlings containing native enzyme

B = $(His)_6$Zm-p60.1, renatured product (matrix assisted refolding procedure – 23 renaturing cycles)

C = $(His)_6$Zm-p60.1 purified by native IMAC

K_M $(His)_6$Zm-p60.1 purified by native IMAC: **0.64 ± 0.06 mM**

K_M $(His)_6$Zm-p60.1 renatured product: **0.6 ± 0.08 mM**

Determination of v_{max} and k_{cat} was hampered by the fact that the refolding process yielded a number of improperly folded polypeptides.

Affinity purification for isolation of protein complexes

Many protein-protein associations that exist within the intact cell are conserved during purification. This property can be exploited to facilitate the detection and identification of physiologically relevant protein-protein interactions.

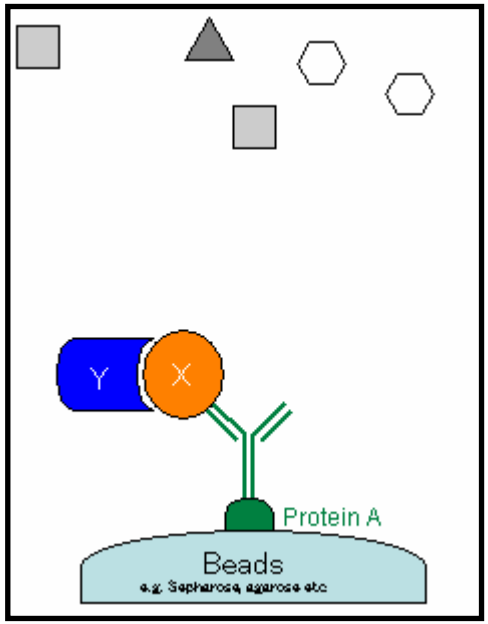
Affinity based method used for detection and identification:

- **Co-immunoprecipitation**
- **Tandem affinity purification**
- **GST pull-down**

- Testing an interaction between two known proteins
- Identification of novel protein-protein interactions

Co-immunoprecipitation

If protein X is immunoprecipitated with an antibody of X, then protein Y, which is stably associated with X in vivo, may also be precipitated. **This precipitation of protein Y, based on a physical interaction with X, is referred to as co-immunoprecipitation**.



1. Cell lysis under mild conditions that do not disrupt protein-protein interactions (using low salt concentrations, nonionic detergents, protease inhibitors, phosphatase inhibitors).
2. The protein of interest (X) is specifically immunoprecipitated from the cell extracts (using an antibody specific to the protein of interest or to its fusion tag).
3. The antibody-protein(s) complex is then pelleted usually using protein-A or G sepharose, which binds most antibodies.
4. Eluted immunoprecipitates are then fractionated by SDS-PAGE.
5. A protein of known identity is most commonly detected by performing a western blot or autoradiography when the interaction partner is labeled with S^{35} methionine. Identification of novel interaction is carried out by mass spectrometry analysis.

Pull-down assay

➢ Pull-down assays are a common variation of immunoprecipitation and are used in the same way, although this approach is more suited to an initial screening for interacting proteins.

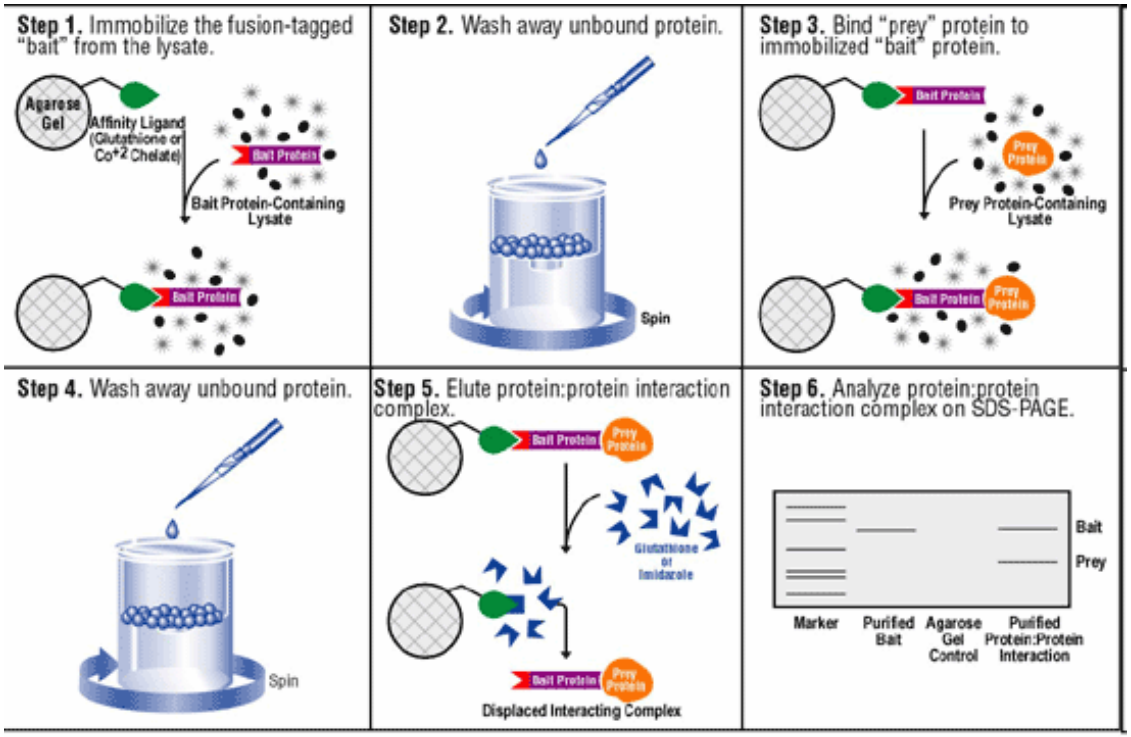
➢ They are used for purification of multiprotein complexes *in vitro*.

➢ The protein of interest is expressed in E. coli as GST (or His) fusion and immobilized on glutathione-sepharose beads (GST alone is often used as a control).

➢ Cellular lysate is applied to the beads or column, and the target protein competes with the endogenous protein for interacting proteins, forming complexes in vitro.

➢ Centrifugation is used to collect the GST fusion probe protein and adhering proteins.

➢ The complexes are washed to remove nonspecifically adhering proteins.



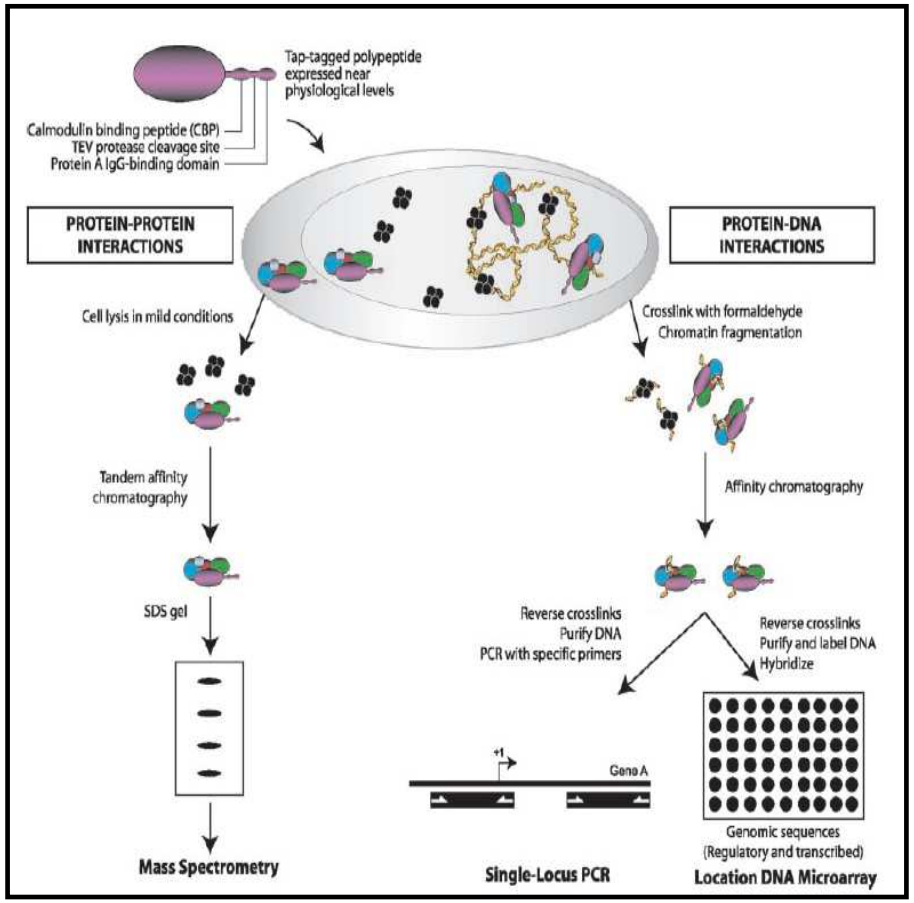
➢ Free glutathione is used to elute the complexes from the beads, or alternatively the beads with attached complexes are boiled directly in an SDS-PAGE sample buffer.

➢ The proteins are resolved on SDS-PAGE and processed for further analysis.

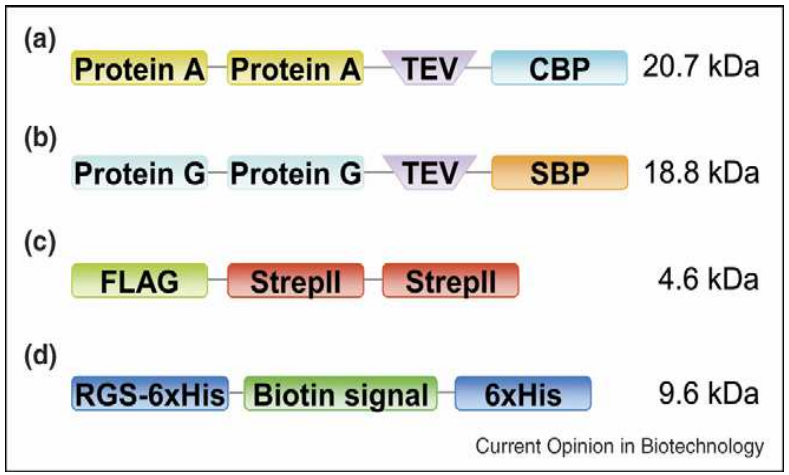
Tandem affinity purification (TAP)

Two-step purification strategy in order to achieve higher purity of isolated multiprotein complexes under near physiological conditions.

This method was originally developed for use in yeast and quickly adapted to higher eukaryotes such as insect cells, human cells and plant cells.

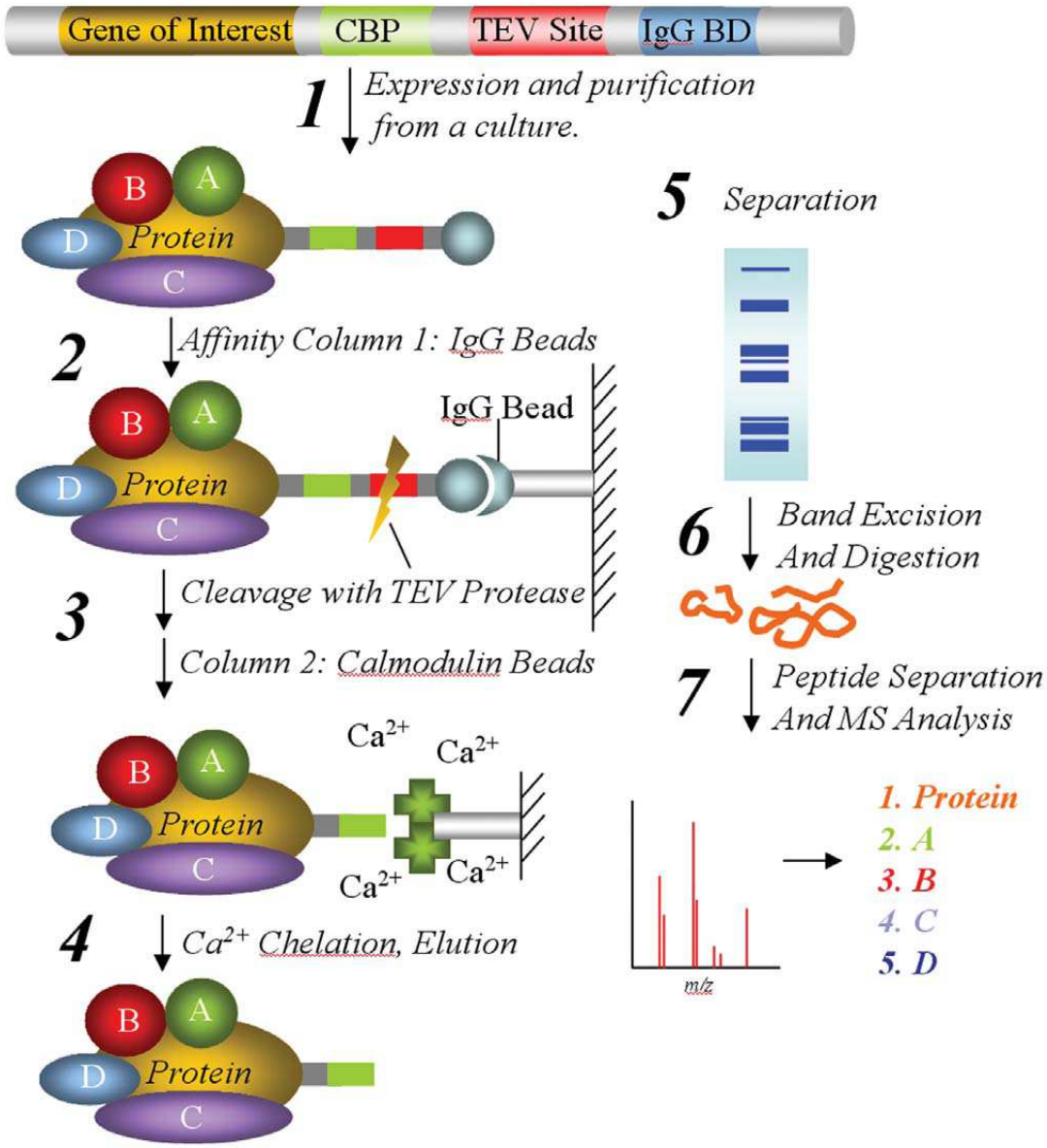


Examples of TAP tags

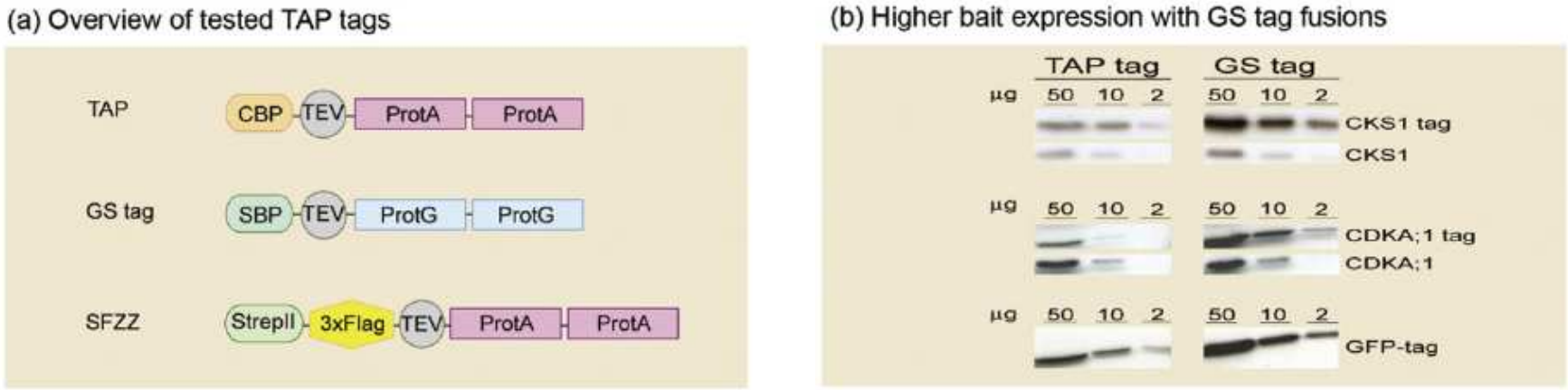


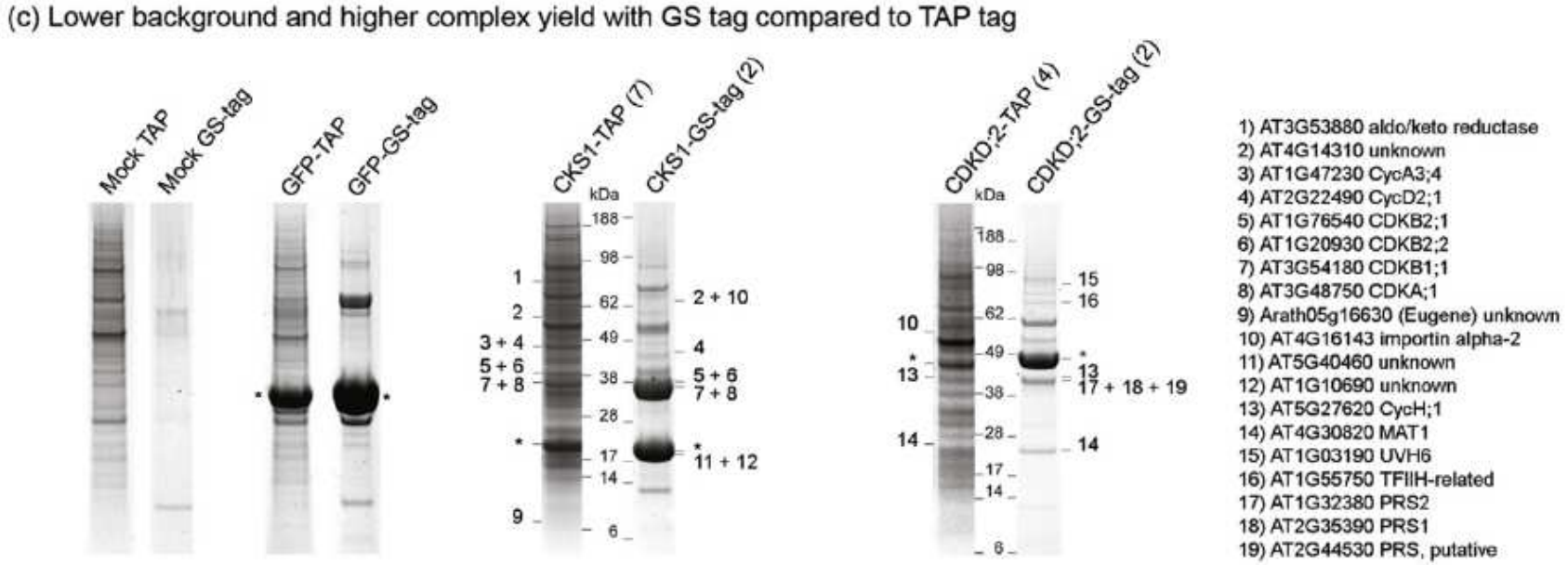
TAP tag: a double affinity tag (highly specific) which is fused to a protein of interest as an efficient tool for purification of native protein complexes.

Tandem affinity purification



A new (and so far the best) TAP tag for plants: The GS tag



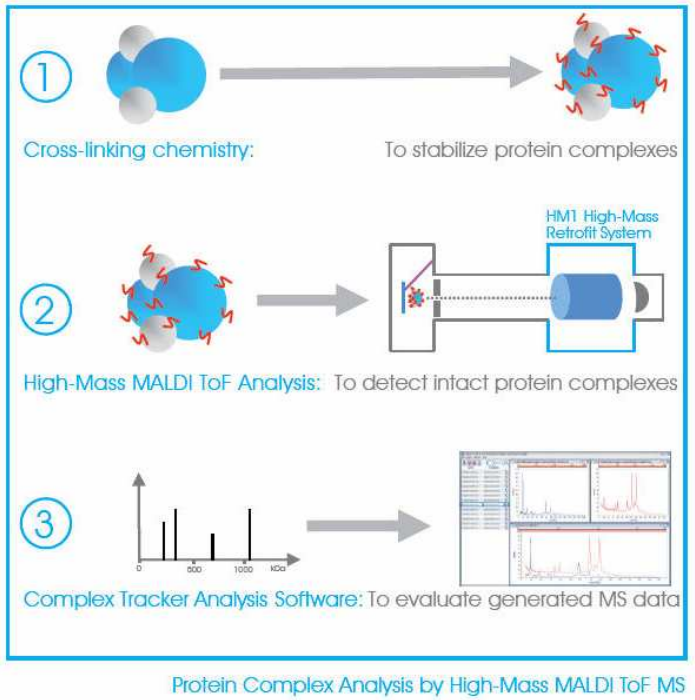


van Leene et al., 2008

Affinity purification for isolation of protein complexes

- Tags can influence protein-protein interactions (testing N- and C-terminal fusion).

- Loss of weak or transient protein-protein interactions (in vivo chemical cross-linking, e.g. using formaldehyde).



Protein Complex Analysis by High-Mass MALDI ToF MS

- Verification of newly identified interactors by other methods and biologically relevant mutants

- Non-specificity: controls, affinity tags with higher specificity