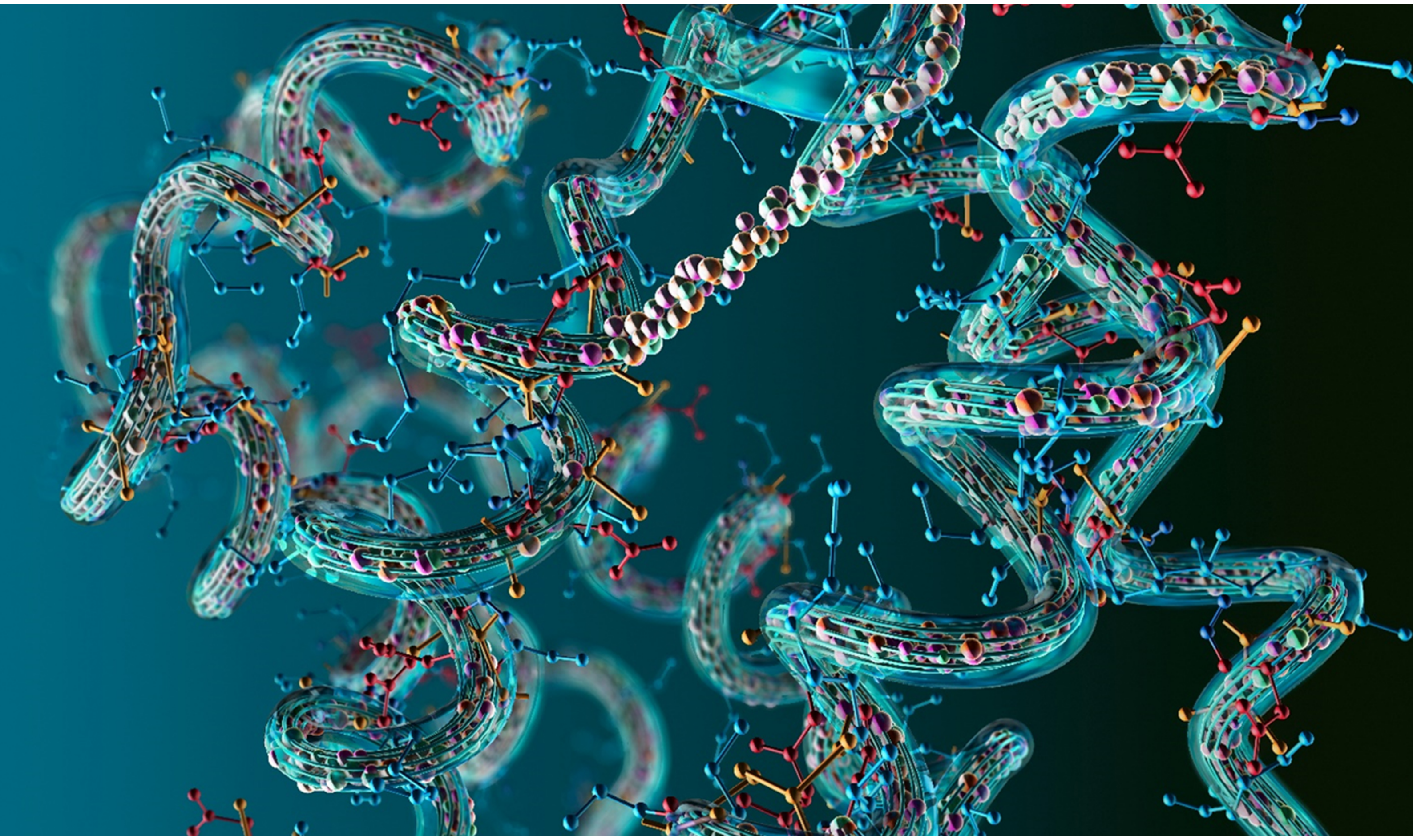
How to Use Fusion Tags



Creative Diagnostics

Fusion Tag Guide

Creative Diagnostics provides a Fusion Tag Guide to support your research.

Creative Diagnostics®

01 Introduction
To Fusion Tags

There are numerous reasons to highlight a protein of interest, including the need to purify it for crystallization or to examine its cellular location. Although many proteins have commercially available antibodies, it can occasionally be challenging to develop a particular antibody against a target. As a solution to these problems, scientists have created an extensive molecular toolbox of fusion tags.

A known protein or peptide is fused to the protein of interest to create a fusion tag. A large selection of high performance antibodies are readily available due to the high level of characterization of these tags, making it easy to detect a specific protein for a variety of applications. Recombinant DNA, in which the DNA of your target protein is inserted into a plasmid that also contains the fusion tag sequence, is the most popular method of attaching the known sequence to your protein. The fusion tag is attached to the protein after the plasmid is translated.

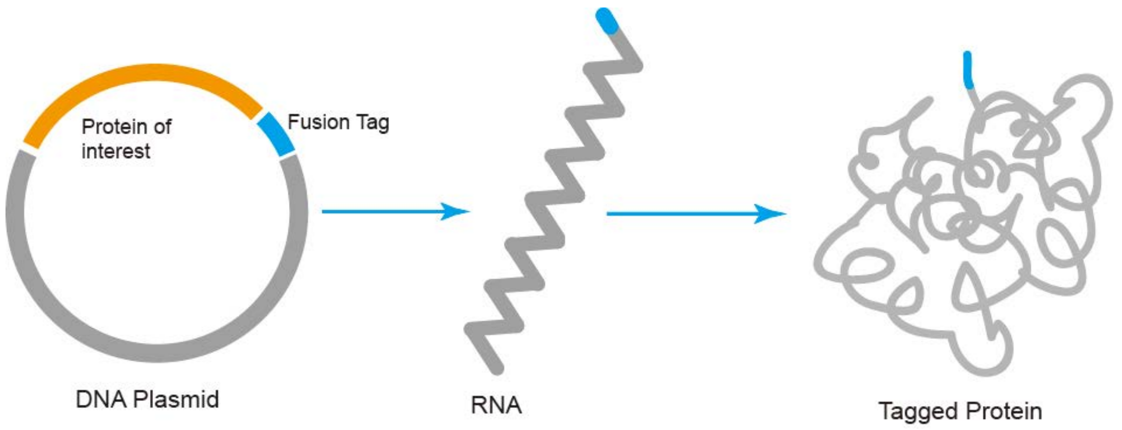


Figure 1. First, the fusion tag and DNA for your target protein are inserted into a plasmid. Translation and transcription then occur to produce the target protein with the fusion tag.

In addition to tag selection, the linker sequence should be carefully considered. The linker sequence that connects your protein to the tag is key to ensuring proper protein folding and functionality. Linker sequences can be flexible, rigid, or cleavable, similar to fusion tags, and each has its own unique properties.

Why Should You to Consider Fusion Tags?

Advantages

- It can isolate a protein of interest without the need for a specific antibody.
- It is sometimes possible to cleave the tag after purification.
- Multiple tags can be attached to the same protein, increasing versatility.
- It avoids antibody interference during immunoprecipitation.
- It can be used to visualize proteins in live cells.

Disadvantages

- Optimal tag insertion may require several attempts, increasing the expense of the experiment.
- Some tags may interfere with protein functionality.

We try to walk you through the potential benefits and drawbacks of the most popular tags, but regrettably we can't say for sure which tag or fusion site will be the most effective for your experiment. In accordance with your aim and application, we recommend following this guidance to select a starting place. Keep an eye out for our [Tips], which will give you information on common roadblocks.

02 Tags Types:
Fluorescence, Affinity, and Epitope Tags

Epitope tags, affinity tags, and fluorescence tags are the three primary types of fusion tags utilized for various applications.

- Affinity tags are typically longer and are used for protein purification or increasing protein solubility.
- Fluorescent tags are typically utilized for imaging studies, such as cellular localization and co-expression tests, and can be applied to both live and dead cells.
- Short peptide sequences, known as epitope tags, are commonly applied in immunological procedures like co-immunoprecipitation and western blotting.

Which Terminal, C or N?

Your decision to attach the tag to the C- or N- terminal of your target protein primarily relies on how the protein folds and whether the terminus you select is necessary for the protein's function. For instance, if the protein's C-terminal is folded inside the protein, you probably won't see any fusion protein signals. Similarly, if your protein is post-translationally cleaved at the terminal to which your tag is fused, your tag will be removed from your protein of interest. It might be desirable to clone both C- and N-terminally tagged constructs to find the optimal choice if you have the resources or if your experiment is innovative. One study discovered that C-terminal tagged fusion proteins are more likely to relocate to the desired subcellular compartment than N-terminal tagged fusion proteins. Although C-terminal tagged proteins typically localize and function as expected, it is crucial to emphasize that this is not always easy to predict. The proper localization of the fusion protein can be verified by immunofluorescence, its size and expression levels can be determined by immunoblot, and its interactions with known substrates can be evaluated by co-immunoprecipitation.

03 Tandem Tag Cleavage and Affinity Purification

Tandem Affinity Purification

Tags have different properties. Sometimes you need a tag to increase the solubility of your protein and a tag for purification at the same time. Tandem Affinity Purification (TAP) is a technique used to meet similar needs. Calmodulin-binding peptide (CBP) and Protein A from Staphylococcus aureus (ProtA), split by a tobacco etch virus (TEV) cleavage site, are examples of proteins that can be joined by TAP. A number of 2-3 tags can currently be fused to your protein of interest using the approach, even though they often contain one of the original TAP components. These tags can either be fused to any of the terminals, or inserted simultaneously into one end of your protein. The tags should typically be fused together to facilitate cleavage if the tag needs to be cut after use.

Due to their diminutive size and efficiency as purifying agents, his-tags are a common choice for TAP. In many cases, they are coupled to the maltose-binding protein (MBP), which increases their solubility and lessens the possibility of aggregates in E. coli systems. GFP and His is another popular combination because it allows your protein to be detected by fluorescence techniques.

TAP tags have many advantages, but because they are so large, they can interfere with the way the protein works. Additionally, calcium signaling may be hampered if you use a tag that contains CBP. It may be tempting to add multiple tags to your protein because each tag offers a unique benefit. As a general rule, you should avoid using more tags than the protein of interest.

Cleavage of Your Tag from Your Protein of Interest

While certain tags, particularly those of large size, pose little risk to protein functionality, others may have consequences further down the line. In these circumstances, it is often beneficial to cleave the tag from your protein of interest after initial detection. Inteins or site-specific proteases can be used for this purpose.

The functionality of your protein shouldn't be affected by site-specific proteases, although it's frequently preferable to get rid of the protease after cleavage. One option is to use a protease fused to the affinity tag itself. Affinity chromatography can then be used to quickly remove the protease. Depending on your individual requirements, there are a variety of protease recognition sites, each with unique advantages and limitations. Fusion of the tag to the N- or C-terminal is one thing to consider. The protein of interest will have very few additional residues when a tag is cleaved from the N-terminal. Cleavage of C-terminal tags, on the other hand, will leave 4-6 additional residues on your target protein. Carboxypeptidases, however, can be utilized in some specific situations to get rid of these brief C-terminal sequences. The most typical protease sites are shown below.

04 Composite Fluorescent Dyes for Flow Cytometry

Enterokinase

Cleavage site:	DDDDK^X
Advantages:	The DDDK (FLAG) tag has an internal recognition site.
Limitations:	The varying degrees of efficacy depend on the amino acid following the lysine residue, from 61% (X = proline) to 88% (X = alanine). Any location can be cleaved.

Factor X

Cleavage site:	I(E/D)GR^X
Advantages:	Widely available
Limitations:	Proline or arginine cannot be an amino acid named "X". Factor X should not be combined with chelating drugs like EDTA since it can bind calcium ions. Cleavage may occur at non-specific sites.

SUMO protease (S. cerevisiae Ulp1)

Cleavage site:	Identification of the tertiary structure of SUMO. The fusion protein is cleaved at its N-terminus.
Advantages:	Applicable in a wide range of buffer temperatures (4–37°C) and pH conditions (5.5–10.5). It can be utilized with 2 M urea to aid in the removal of SUMO-tagged proteins from inclusion bodies.
Limitations:	The effectiveness of your protein of interest depends on its sequence; for instance, efficacy is decreased if a proline follows the cleavage site.

Tobacco etch virus (TEV) protease

Cleavage site:	ENLYFQ^S
Advantages:	Widely produced using recombinant technology, ensuring precise cleavage at a reasonable price. It can be utilized in a variety of buffers and at low temperatures. The natural N-terminal residue of your protein of interest can be used without loss of potency due to the flexibility of the post-cleavage residue.
Limitations:	The efficiency of the wild type TEV protease can be decreased by cleaving it.

Thrombin

Cleavage site:	LVPR^GS
Advantages:	Suitable for use in the presence of detergents; relatively specific. Benzamidine sepharose removes the substance effectively.

Limitations:	Although rare and typically caused by impurities in commercial preparations, cleavage may occur at non-specific sites.

3C and PreScission™ (human rhinovirus)

Cleavage site:	ETLFQ^GP
Advantages:	Optimal activity at 4°C
Limitations:	Efficiency loss at higher temperatures.

Another strategy is the use of inteins. These are protein fragments that catalytically remove themselves from their host proteins. This approach has drawbacks even though it eliminates the requirement for proteases. First, the large pieces of catalytic machinery involved in inteins increase the metabolic burden on the cells and do not improve solubility or ease of purification. The second issue is that it can be a slow procedure and hasn't been thoroughly tested in a high-throughput setting. The sequence context at the fusion junction affects the effectiveness of cleavage.

A modified intein sequence and a chitin-binding domain (CBD) work together to provide an efficient affinity tag. The connection between the tag and resin is quite robust, although the elution phase requires proteolytic cleavage.

05 Affinity Tags

Affinity tags get their name because they are often used in affinity purification, a method of isolating proteins from cell lysates. Your protein of interest can be removed from solution by chemical or physical interaction of the affinity tag with the immobilized substrate. The precise affinity chromatography technique depends on the tag chosen. You can discover some of the most popular affinity tags in this section.

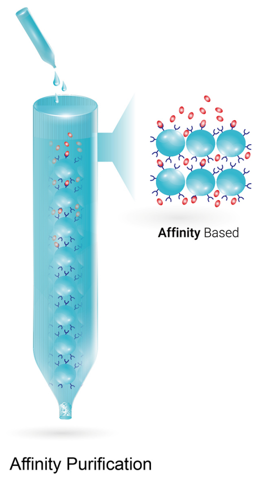


Figure 2. Schematic diagram of affinity chromatography.

Glutathione-S-transferase

Molecular Weight:	~26 kDa
Size:	211 amino acids
Tag location:	C- or N- terminals
Affinity Resin:	Glutathione
Applications:	Protein purification, protein-DNA interactions, protein-protein interactions
Strengths:	They can promote solubility and stabilize fusion protein structure
Limitations:	Tag is relatively large, which may affect the function of the fusion protein. Heat-shock proteins may be present in the eluent.

Overview of Glutathione-S-transferase

The protein glutathione-S-transferase (GST) has 211 amino acids and a molecular weight of around 26 kDa. Native GST is in charge of defending the cell against harmful substances and oxidative stress. One of GST's key characteristics is its affinity for glutathione, a tripeptide that can be used in affinity chromatography. GST binds to glutathione and separates it from the solution when a solution containing GST is passed through a column with immobilized glutathione. In order to obtain the best binding of GST to the immobilized glutathione, it is advised to employ buffers that are close to neutral. GST can be eluted after binding using 10 mM glutathione, a gentle eluent that helps retain the function of the fusion protein.

What to be Wary of:

- Contamination: GST can occasionally elute heat-shock proteins. With SDS-PAGE, this contamination can be found out: HSP60 and HSP10 will appear as a band at about 68 kDa, while HSP70 will manifest as a doublet at 70 kDa. Prior to purification, cell lysates should be treated with 5 mM MgCl2 and 5 mM ATP to eliminate heat-shock protein impurities.
- A reduction in protein functioning: Due of GST's size and tendency to spontaneously dimerize in solution, your protein's activity or functionality can be affected. By using proteases like thrombin, factor Xa, or PreScission to cleave the GST from your protein of interest, this problem can be resolved. While attached to the glutathione column, this action can be carried out.
- Denaturing GST: GST must be appropriately folded to bind to the glutathione column. Therefore, it is recommended to apply reducing or denaturing agents in low concentrations: 2-3 M urea, less than 10 mM 2-mercaptoethanol, or 2% by volume Tween 20.

Tips

Before purification, cell lysates should be treated with 5 mM MgCl2 and 5 mM ATP to eliminate heat-shock protein impurities.

His-tags

Molecular Weight:	0.2–1.6 kDa. 6x-His tag is 0.8 kDa
Size:	2–10 histidine residues
Tag location:1	C- or N- terminals, or internal.
Affinity Resin:	Transition metal ions, usually Ni^{2+}
Applications:	Protein purification
Strengths:	Small size means less chance of influencing the functionality of the fusion protein, causing no inclusion bodies.
Limitations:	In insect and human cells, background binding can be rather strong.

Overview of His-Tags

Since PolyHis tags are compact and have stable binding, they are frequently utilized for protein purification. The most common His-tag is the 6x-His tag, often known as hexatag, which contains six histidine residues. For the purpose of purifying proteins, histidine forms coordination connections with immobilized transition metal ions. For immobilized metal affinity chromatography (IMAC), cobalt and zinc columns are available, but nickel columns are most frequently utilized. His-tags are short peptide sequences so that rarely alter the characteristics of the fused protein.

What to be Wary of:

- Endogenous His: Since his residues are common in insect and mammalian systems, use of these systems may result in a significant background. Washing in 5–10 mM imidazole will prevent background binding, but may prematurely elute your target protein. Background binding must also be considered when using anti-His antibodies.

- Using proteins with metal centres: Because the metal could be absorbed by the affinity resin, it is not recommended to fuse a protein with a metal centre to a His-tag.

- Prevent reducing circumstances: Affinity resins should not be subjected to reducing conditions such as dithiothreitol (DTT) or -mercaptoethanol.

Tips

To improve IMAC, try changing the location of your His-tag.

Biotin

Molecular Weight:	244 Da
Affinity Resin:	Avidin/Streptavidin
Applications:	Protein purification, surface plasmon resonance, protein expression.

Strengths:	Utilises avidin's and streptavidin's high affinity.
Limitations:	In mammalian systems, co-elution with other biotinylated proteins is possible.

Overview of Biotin

Biotin, sometimes referred to as vitamin H, has a molecular weight of 244 Da. It is a frequently employed molecular tool and is often used in ELISA and western blot experiments as a visualisation approach when coupled to secondary antibodies. Affinity purification can take use of the fact that avidin and streptavidin molecules form very strong bonds with biotin. Biotin is a tiny molecule, hence it is not likely to have an impact on the function of proteins. The Strep-tag is another useful tag. The fact that it is 8 amino acids long (WRHPQFGG) and binds reversibly to the same location as biotin makes it unlikely to affect protein function. This indicates that under mild buffer conditions, proteins fused to the Strep-tag can be easily purified with streptavidin resin and eluted with biotin.

What to be Wary of:

- Tetrameric vs. monomeric affinity resins: Avidin-coated columns, either tetrameric or monomeric, can be used to purify biotin-tagged proteins. While monomeric resins allow a softer elution using a 10 mM biotin buffer, tetrameric avidin columns require strong denaturing agents, such as urea or guanidine hydrochloride, for the elution process.

- Application in mammals: At least four different biotinylated protein species can co-elute with your target protein in mammalian systems.

06 Epitope Tags

The portion of an antigen known as an epitope is recognized by antibodies. Therefore, tags that are often utilized in antibody-based assays are called epitope tags. Since epitope tags are often shorter than affinity tags, they are less likely to interfere with protein function. Immobilized antibody-based columns can be used for affinity purification, but are typically more expensive or less effective than columns for affinity tags. Epitope tags, however, are a helpful tool for the identification of fusion proteins because of their affinity for the corresponding primary antibodies. Immunoprecipitation (IP), including protein complex immunoprecipitation (co-IP), and cell culture both make extensive use of epitope tags.

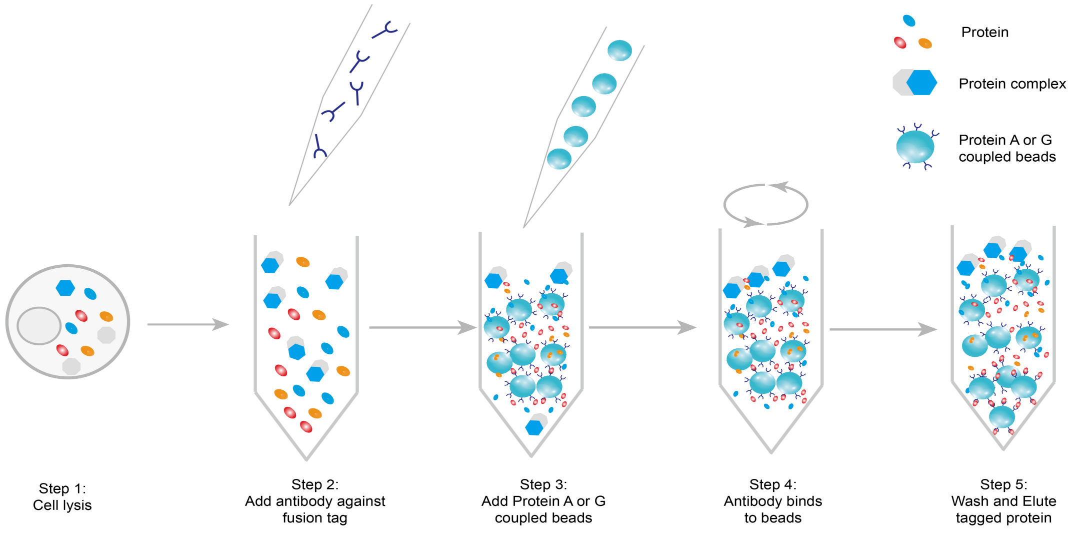


Figure 3. The process for co-immunoprecipitation is depicted in this diagram.

c-myc

Molecular Weight:	1.2 kDa
Size:	10 amino acids (EQKLISEEDL)
Tag location:	C- or N- terminals
Affinity Resin:	Divinyl sulphone-activated agarose
Applications:	Western blot, immunoprecipitation, flow cytometry. It can be used in affinity purification.
Strengths:	Either the N- or C-terminal can be used for it. Proteins that have been tagged with c-myc can crystallize effectively.
Limitations:	Low pH elution for affinity chromatography may have an impact on the functionality of fusion proteins.

Overview of c-myc

Human c-myc is expressed at low levels in growing cells and is essential for the development of human cancer. The c-terminal of the c-myc gene is the origin of the c-myc tag, which is well recognized by antibodies. As a result, it is a widely used tag for the detection of proteins in procedures including flow cytometry, immunoprecipitation, and western blot. Successful crystallisation of c-myc-tagged purified proteins has been achieved.

What to be Wary of:

Fusing to a secretory signal: The c-myc tag can be attached to either the C or N terminals, but it is not advised to fuse the tag to secretory signals as this may affect translocation to the secretory pathway.
Affinity purification: Although the c-myc tag can be used to purify proteins, this is a rather uncommon

application. This is because elution requires a low pH, which can be detrimental to the functionality of proteins.

Tips

Fusing c-myc to a secretory signal is not recommended as it may affect translocation to the secretory pathway.

Human influenza hemagglutinin (HA)

Molecular Weight:	1.1 kDa
Size:	9 amino acids (YPYDVPDYA)
Tag location:	C- or N- terminals
Affinity Resin:	Anti-HA antibody immobilized onto agarose beads
Applications:	ELISA, western blot, immunoprecipitation and immunofluorescence, protein purification
Strengths:	Unlikely to affect protein functionality.
Limitations:	Not recommended for use in apoptotic cells.

Overview of HA

The HA glycoprotein, located on the surface of influenza viruses and responsible for the virus's contagiousness, is the source of the human influenza hemagglutinin (HA) tag. As a useful tool for protein identification by ELISA, western blot, immunoprecipitation, and immunofluorescence, the HA-tag is a short peptide tag that rarely interferes with protein activity. For protein purification, anti-HA antibodies can also be immobilized on agarose beads.

What to be Wary of:

It is not advised to utilize HA-tags for proteins derived from apoptotic cells when using affinity purification. Caspases 3 and 7 have the ability to cleave the HA-tag, resulting in loss of immunoreactivity.

Tips

The HA-tag is a tiny peptide tag and is therefore unlikely to interfere with protein function.

DDDK (FLAG®, Sigma)

Molecular Weight:	1.01 kDa
Size:	8 amino acids (DYKDDDDK)
Tag location:	C- or N- terminals, or internal
Affinity Resin:	Immobilized DDDK antibodies

Applications:	ELISA, western blot, protein purification, protein crystallization
Strengths:	Not expected to have an impact on the fusion protein's functionality. There is a cleavage point inside.
Limitations:	Affinity Resin is less durable than other materials and can be expensive.

Overview of DDDK

The DDDK, or FLAG®, tag (Sigma), is the only patented tag. It is more hydrophilic than other epitope tags, so it is less likely to affect the function of the protein to which it is fused. If necessary, the DDDK tag can easily and swiftly be removed from the target protein thanks to its enterokinase cleavage sequence.

What to be Wary of:

Affinity purification: Anti-DDDK antibody-immobilized columns can be efficient, but they are also more expensive than other kinds of affinity columns.

Tips

The target protein can easily be stripped of the DDDK tag due to its enterokinase cleavage sequence.

V5

Molecular Weight:	0.95–1.4 kDa
Size:	9–14 amino acids (IPNPLLGLD or GKPIPNPLLGLDST)
Tag location:	C- or N- terminals
Applications:	Western blot, ELISA, flow cytometry, protein visualization, ChIP, immunoprecipitation.
Strengths:	Available in two peptide lengths
Limitations:	Potential cross-reactivity in mammalian systems

Overview of V5

The P and V proteins of the paramyxovirus simian virus 5 are the source of the V5 tag. The V5 tag comes in two sizes, ranging from 9 to 14 amino acids, although the longer tag is typically employed. Sometimes it is advised to combine a His-tag and a V5-tag.

What to be Wary of:

There is a potential for cross-reactivity when employing a mammalian expression system.

Tips

Sometimes it is recommended to combine a His-tag and a V5-tag.

06 Fluorescent Tags

The number of fluorescent tags has multiplied since the original green fluorescent protein (GFP) gene was cloned in 1992. Fluorescent tags have the significant benefit of being non-toxic, allowing the use of live cells. Despite being large tags, they typically alter most proteins only slightly. Although GFP, which has a molecular weight of 26.9kDa, is one of the most popular fluorescent tags, there are a number of crucial factors to consider.

Table 1. Excitation and emission spectra for different fluorescent proteins

Class	Protein	ʎex (nm)	ʎem (nm)	Extinction coecient	Quantum Yield	Relative brightness (% of EGFP)	pKa	Bleaching t½ (s)
Blue	EBFP2	383	448	32,000	0.56	53	4.5	55
Cyan	Cerulean	433	475	43,000	0.62	79	4.7	36
Green	EGFP	488	507	56,000	0.6	100	6	174
UV-excitable green	T-Sapphire	399	511	44,000	0.6	79	4.9	25
Yellow	mCitrine	516	529	77,000	0.76	174	5.6	49
Orange	mOrange	548	562	71,000	0.69	146	6.5	9
Red	mCherry	587	610	72,000	0.22	47	<4.5	96
Far red	mPlum	590	649	41,000	0.1	12	<4.5	53

References

1. Chen, X., Zaro, J. & Shen, W. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*. 2013, 65, 1357-1369.
2. Costa, S., Almeida, A., Castro, A. & Domingues, L. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in Microbiology*. 2014, 5.
3. Palmer, E. & Freeman, T. Investigation Into the use of C- and N-terminal GFP Fusion Proteins for Subcellular Localization Studies Using Reverse Transfection Microarrays. *Comparative and Functional Genomics*. 2004, 5, 342-353.
4. Snapp, E. Fluorescent proteins: a cell biologist's user guide. *Trends in Cell Biology*. 2009, 19, 649-655.
5. Kimple, M., Brill, A. & Pasker, R. Overview of Affinity Tags for Protein Purification. *Current Protocols in Protein Science*. 2013, 73.
6. Kosobokova, E., Skrypnik, K. & Kosorukov, V. Overview of fusion tags for recombinant proteins. *Biochemistry (Moscow).* 2016, 81, 187-200.
7. Young, C., Britton, Z. & Robinson, A. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal*. 2012, 7, 620-634.
8. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*. 2003, 60, 523-533.
9. Kimple, M., Brill, A. & Pasker, R. Overview of Affinity Tags for Protein Purification. *Current Protocols in Protein Science*. 2013, 73.
10. Brizzard, B. Epitope tagging. *BioTechniques*. 2008, 44, 693-695.
11. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology.* 2003, 60, 523-533.
12. Crivat, G. & Taraska, J. Imaging proteins inside cells with fluorescent tags. *Trends in Biotechnology*. 2012, 30, 8-16.
13. Shaner, N., Steinbach, P. & Tsien, R. A guide to choosing fluorescent proteins. *Nature Methods*. 2005, 2, 905-909.

CONTACT

CREATIVE DIAGNOSTICS

45-1 Ramsey Road, Shirley, NY 11967, USA

Tel: 1-631-624-4882 (USA) 44-161-818-6441 (Europe)

Fax: 1-631-938-8221 | Email: info@creative-diagnostics.com

www.creative-diagnostics.com

CD Creative Diagnostics®