

## Section 1C

# Getting Started with Epitope Tagging

Typical epitope tagging experiments require a combination of techniques from different disciplines, including molecular biology, cell biology, protein chemistry, and immunology. For those unfamiliar with some or all those techniques, establishing a successful epitope tagging system could prove a time-consuming endeavor.

As with most things, however, learning epitope tagging is easier if you can learn from those who have already done it. They can not only describe the procedure, but they know the critical steps that must be done “just so” if the experiment is to succeed. They can help you recover when something goes wrong. They know the answers to the questions you are most likely to ask.

So, think of this section as a dialogue between someone learning epitope tagging and someone with epitope tagging experience. The questions are those someone new to “epitope tagging” might ask. The answers are a distillation of the epitope tagging expertise and advice of Boehringer Mannheim Research and Development scientists, Technical Services representatives, and customers.

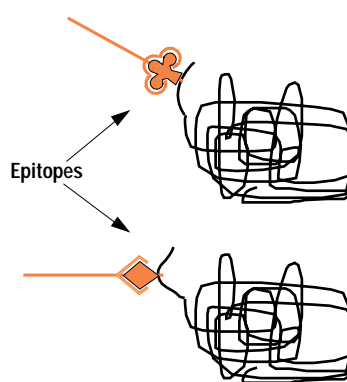
### How should I prepare the tagged protein?

The first step in epitope tagging is the addition of an epitope peptide to the target protein. Here are some basic things to consider.

### What’s the difference between a fusion protein and an epitope tagged protein?

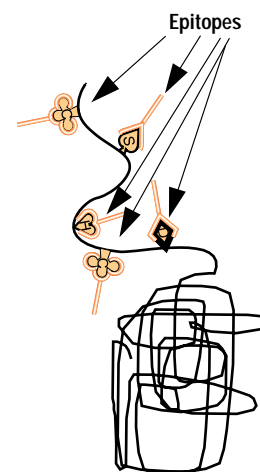
Strictly speaking, there is no difference. Each is a single protein constructed, by recombinant DNA techniques, from a target protein and another unique peptide sequence. However, each added sequence has unique properties that make it a good “tag” for detection and purification.

In an epitope tagged protein (Figure 1C.1), the added sequence is a short peptide (3–14 amino acids), usually with no function of its own. Table 1C.1 (on page 1.6) lists several epitope tags. The important property of the epitope tag (Figure 1C.1) is its ability to be recognized and bound by a single, tag-specific antibody.



**Figure 1C.1: Epitope tagged proteins.** Two proteins, each tagged with a unique small peptide. Each peptide serves as an epitope (sequence recognized by an antibody) of a single, tag-specific antibody.

In a fusion protein (Figure 1C.2), the added sequence is a sizable polypeptide, usually with a function of its own. Examples of polypeptides used to construct fusion proteins include  $\beta$ -galactosidase ( $\beta$ -Gal, approx. 120 kD amino acids), glutathione-S-transferase (GST, approx. 26 kD), ketosteroid isomerase (KIS, approx. 15 kD) and maltose binding protein (MBP, approx. 44 kD). A fusion protein may have many epitopes (Figure 1C.2), but it is usually detected by its function (for instance, enzymatic activity or ability to interact with another protein) rather than by its antigenic properties.



**Figure 1C.2: Fusion protein.** The tagging unit is a large polypeptide, which contains many epitopes. Fusion proteins are usually detected by function, rather than antigenic properties.

### Which tag should I use?

A wide range of epitopes have been used to tag proteins (Table 1C.1) and many are available commercially. However, the “best” epitope tag for a particular experimental system is the one that does not interfere with the function or cellular processing of the tagged protein, yet gives a strong detection signal on Western blots, in immunofluorescence microscopy, or in other analytical techniques.

There is no single “best” epitope for every tagging application, and no one can reliably predict how a particular tag will behave in a particular protein. However, as a starting point for determining the right tag sequence for your particular experimental system, you can do the following:

- ▶ Do a review of the literature to see if proteins similar to yours have been successfully tagged with a particular epitope. If it worked for them, it may work well for you.
- ▶ Consider the location of the tag. Most epitope tags can be placed at the amino-terminus (N-terminus), at the carboxy-terminus (C-terminus), or within the reading frame of your target protein. However, if you do not know much about the function of your protein, attaching the tag to either end of the protein may be less likely to interfere with protein function.

**Note:** It may be useful to place the tag at the C-terminus, rather than the N-terminus, since detection of a C-terminal tag *in vivo* demonstrates complete translation of the tagged protein.

- ▶ Consider the copy number of the epitope tag. Some researchers have problems with the sensitivity of their detection assay. Thus, it might be better to add several copies of the same tag sequence in tandem to your protein to increase the detection signal.

**Caution:** A longer multicopy tag sequence may give a stronger detection signal, but it may also be more likely to interfere with the function of your protein.

## What technique should I use to insert the tag?

Epitope tags are incorporated at the DNA level. To fuse the coding sequences of the epitope and the target protein, most researchers use either (1) simple recombinant DNA techniques or (2) oligonucleotide-mediated mutagenesis via the Polymerase Chain Reaction (PCR).

Which technique you choose depends upon what expression vectors you have available to carry your tagged protein into the target cell (Figure 1C.3).

Name	Number of amino acids	Remark/Sequence*
$\alpha$ -tubulin	3	C-terminal sequence of yeast protein. Sequence: EEF.
B-tag	6	VP7 protein of blue-tongue virus. Sequence: QYPALT
E tag	13	Sequence: GAPVPYPDPLEPR.
c-myc <sup>§</sup>	10	Human c-myc gene protein. Sequence: EQKLISEEDL.
FLAG <sup>®</sup> epitope	8	Synthetic peptide. Sequence: DYKDDDDK.
HA <sup>§</sup>	9	Peptide from human influenza hemagglutinin protein. Sequence: YPYDVPDYA.
His <sub>6</sub> or His <sub>10</sub>	6 or 10	Polyhistidine; binds metal ligand (affinity chromatography); antibody used for detection.
HSV	11	Peptide from herpes simplex virus glycoprotein D. Sequence: QPELAPEDPED.
Pk-tag	9 or 14	P/V proteins of paramyxovirus SV5. Sequence: GKPIPPLLGLDST.
Protein C	12	Sequence: EDQVDPRLIDGK.
T7	11	Major capsid protein of the T7 phage. Sequence: MASMTGGQQMG.
VSV-G <sup>§</sup>	11	Vesicular Stomatitis Virus Glycoprotein. Sequence: YTDIEMNRLGK.

\*Abbreviations for amino acids: A, Alanine; C, Cysteine; D, Aspartic Acid; E, Glutamic Acid; F, Phenylalanine; G, Glycine; H, Histidine; I, Isoleucine; K, Lysine; L, Leucine; M, Methionine; N, Asparagine; P, Proline; Q, Glutamine; R, Arginine; S, Serine; T, Threonine; V, Valine; W, Tryptophan; Y, Tyrosine.

<sup>§</sup> Antibodies to these epitopes are available from Boehringer Mannheim

Table 1C.1: Some sequences that have been used as epitope tags

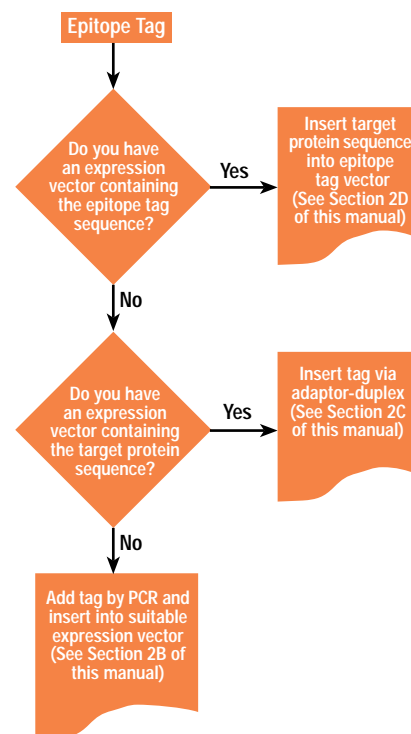


Figure 1C.3: Flow chart for deciding which epitope tagging technique to use.

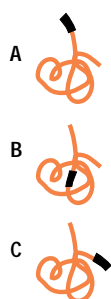
Of the available tagging techniques, PCR is the most flexible, since PCR:

- ▶ Can insert the tag sequence at any point in the target protein sequence, rather than depending upon the existence of a certain restriction site
- ▶ Is fast, resulting in a usable product in just a few hours
- ▶ Produces clonable amounts of the hybrid DNA without using a bacterial expression system

For more information on adding an epitope by PCR and other tagging techniques, see Chapter 2 of this manual.

### At what location in my protein do I insert the tag?

In theory, most epitope tags can be placed at the amino-terminus (N-terminus), at the carboxy-terminus (C-terminus), or anywhere within the coding sequence of your target protein (Figure 1C.4). If it does not interfere with biological function, you can place the tag sequence near a particular site on the protein to address questions about the function of that site.



**Figure 1C.4: Location of epitope tag on protein.** Theoretically, an epitope tag (colored block) may be placed at the N-terminus (A), the C-terminus (C) or within the coding sequence (B) of a protein. Note that a tag at either end of the protein (A, C) may be more accessible to antibodies than a tag within the body of the protein (B).

In the majority of cases, however, the tag should be placed at either the N-terminus or the C-terminus of the protein coding sequence. Locating the tag at one end of the protein minimizes the possibility that the tag will interfere with protein function. It also maximizes the probability that the tagged protein will be recognized by the tag-specific antibody, since it is less likely to be buried by the native folding of the protein (Figure 1C.4).

**Note:** Introduction of two or three tag sequences into a gene may increase the avidity of the antibody-antigen reaction and, thus, enhance the ability of the antibody to detect the tagged protein in vivo or in vitro (Berkower et al., 1994; Tyers et al., 1993). However, multiple copies of the tag may be more likely to interfere with protein function.

### How do I get a functional tagged protein expressed in the cell?

The DNA sequence for the tagged protein must be added to an expression vector that contains sequences to allow the tagged protein to be made in the target cell. (See Section 2D of this manual for more information on expression vectors.) Then, the expression vector must be introduced into the target cell by transformation (of bacteria) or transfection (of eukaryotic cells).

There are five basic methods for transfecting the DNA into a eukaryotic cell. The most commonly used are:

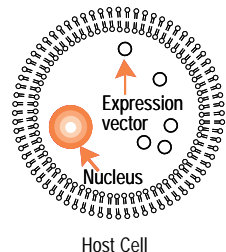
- ▶ **Chemical methods**, in which compounds (such as DEAE dextran or calcium phosphate) make openings in the cell membrane, through which DNA can pass
- ▶ **Physical methods** (such as electroporation, particle delivery [gene gun] or microinjection) which mechanically disrupt the cell membrane, allowing DNA to enter
- ▶ **Lipid-mediated methods**, in which lipid-like molecules (such as DOTAP, DOSPER, DOTMA, or DOPE) enable DNA to enter the cell with minimal disruption of the membrane

**Note:** The molecules used for lipid-mediated transfection have two parts: a positively charged head that can complex with negatively charged DNA and a lipid tail that can pass easily into the cell. For more information on lipid-mediated transfection, call the Boehringer Mannheim representative in your area.

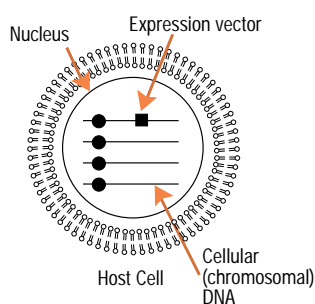
Less commonly used transfection methods include:

- ▶ **Viral methods**, in which a virus carries the transfected DNA into the cell, often as part of its genome
- ▶ **Receptor-mediated methods**, in which a particular ligand complexes the DNA and carries it into the cell via a membrane receptor that recognizes the ligand

## A. Transient transfection



## B. Stable transfection



**Figure 1C.5: Transient vs. stable transfection.** In a transient transfection (panel A), an expression vector makes tagged protein for a brief time in the cell cytoplasm. In a stable transfection (panel B), the expression vector integrates into a chromosome and is replicated, transcribed, and translated indefinitely (like the cell's own DNA).

Depending on the fate of the transfected DNA, a transfection may be:

- ▶ **Transient** (Figure 1C.5, panel A): Transfected DNA remains free within the cell cytoplasm and functions for only a limited time
- ▶ **Stable** (Figure 1C.5, panel B): Transfected DNA recombines with the cell DNA, forming an integrated copy of the vector that the cell will replicate, transcribe, and translate as if it were normal cell DNA

## What immunological techniques can I use with tagged proteins?

As Table 1B.1 (page 1.4) shows, epitope-tagging can be an important tool for addressing many research problems. While the literature describes many applications for epitope-tagged proteins, most laboratories use four techniques to analyze or purify tagged proteins:

- ▶ Western blotting
- ▶ Immunofluorescence microscopy
- ▶ Immunoprecipitation
- ▶ Affinity chromatography

Here are some critical factors to consider as you begin to use each of these techniques.

## How much antibody reagent will I need?

No exact answer can be given, since different tags and different applications will require differing amounts of tag-specific antibody reagent. However, to estimate how much antibody reagent you will need for each experiment you design, determine:

- ▶ The working concentration of antibody reagent that will give optimal results  
**Note:** Based on Boehringer Mannheim laboratory experience, Table 1C.2 gives some guidelines for the use of our tag-specific antibodies in common techniques.

- ▶ The volume of working antibody reagent solution needed for each sample (as detailed in Chapters 4 and 5 of this manual)
- ▶ The number of samples and controls you will need to include in the experiment

## Should I use a direct or indirect detection procedure?

The literature generally refers to two types of detection procedures, direct and indirect. An analytical technique, such as Western blotting, may use either direct or indirect detection. The difference depends upon the type(s) of antibody used and the location of the indicator molecule that allows visualization of the antibody-antigen complexes (Figure 1C.7).

## Note on dilution nomenclature

Often in this manual, a procedure may refer to a 1:X dilution (for example, a 1:10 dilution of secondary antibody conjugate). The nomenclature, 1:10, in this manual means "Take 1 part of the concentrated solution and add it to 9 parts of diluent, then mix thoroughly."