Fusion Protein Nomenclature

A fusion protein is defined as a multifunctional protein derived from a single nucleotide sequence that may contain 2 or more genes or portions of genes with or without amino acid linker sequences. The genes should originally code for separate proteins, both with pharmacological action (e.g., action and targeting).

The suffix "*-fusp*" for "fusion protein" will be used with a syllable formed from 1 consonant and 1 vowel added before the suffix to indicate: (a) the pharmaceutical action (consonant); and (b) the target (vowel). The meanings of these infix letters are given below.

The "*-fusp*" naming scheme is not designed to give comprehensive information about the substance, but to indicate it is a fusion protein and to indicate its action and target. The description on the USAN adoption statement will provide more extensive information about the precise content and action of the fusion protein.

Bifunctional fusion proteins with 1 component being a peptide will only be included in the "*–fusp*" naming scheme if it has a clear pharmacological action or targeting role. If the role of the peptide is not clear or is vague, other naming conventions should be used.

All components must be endowed with a pharmacological activity.

In a bifunctional fusion protein, if 1 component has a purely stabilizing function (e.g., to increase half-life), no "*-fusp*" will be assigned. For instance, if the component is a stabilizing Fc fragment, the "*ef-*" prefix should be used, not "*-fusp*".

In a multifunctional fusion protein that has more than 1 pharmacological action, but also contains a stabilizing Fc fragment, both "*ef-*" and "*-fusp*" should be used.

If both components of the fusion protein have a targeting action, and 1 of them is derived from a monoclonal antibody (mAb) or from a mAb fragment, when assigning the identifying infix letters, the "*-a-*" for *antibody* takes priority. For instance, the fusion of a receptor with an antibody will be "*-ra-*" (where *r* stands for *receptor* and *a* for *antibody*) not "*-be-*" (where *b* stands for *binding protein* and *e* for *receptor*).

The infix letters will not distinguish between mAb or mAb fragments; in all these cases the letter "*a*" will be selected.

Multiple mAb or mAb fragments will be named using the "*-mab*" nomenclature scheme, not the "*-fusp*" scheme.

If more than 2 components are present, the 2 infix letters will still be used to represent the different action/targeting by class: e.g., if a fusion protein is composed by 2 mAbs and 1 receptor, the USAN will end in "*-rafusp*."

Infix Letters and Their Meaning

These letters have the following "action" meanings associated with them:

- b = Binding protein
- c = Encapsulation protein
- d =
- f = Hormone
- g = Antigen
- k = Cytokine
- l =
- m = Membrane protein
- n = Enzyme
- p = Apoptosis
- r = Receptor
- s =
- t = T-cell receptor
- v = Multiple actions/proteins
- x = Toxin
- z =

These letters have the following "targeting" meanings associated with them:

- a = Antibody
- e = Receptor
- i = Antigen
- o = Other
- u = Untargeted

Please be aware of the additional notes:

- *v* = will be used when a multifunctional fusion protein has multiple and not related actions;
- *o* = will be used when some other targeting mechanism (i.e., not antibody, receptor or antigen) is used in a bifunctional fusion protein or in a multifunctional fusion protein with multiple unrelated targeting;
- *u* = will be used when a fusion protein has multiple actions and no targeting;
- h, j, q, w, y = these consonants were excluded to facilitate translation.