Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Peptide Design Guideline

Version 3.1, Revision 2011-05-17

Facts in Designing and Analyzing Your Desired Peptides

The sequence, amino acid composition and length of a peptide will influence whether correct assembly and purification are feasible. These factors also determine the solubility of the final product. The following points are what should be taken into consideration in the design of a peptide.

Overview

1. Design of Peptide Sequence

Peptides can be designed de novo, but most peptides of biological interest are derived from N-terminal, C-terminal, or internal sequences of native proteins. Unfortunately, there are various reasons why certain native sequences sometimes need to be altered. Even for relatively short sequences, there are essential and non-essential (or less important) amino acid residues, although the relative importance of the individual amino acid residues is not always easy to determine. The "not-so-straightforward" rule of thumb is to make the changes in the "non-essential residues". These changes may include amino acid substitution (e.g., for solubility, stability, etc.), chemical modification (e.g., for stability, structure-function studies), attachment of ligands (e.g., chromophores, affinity ligands), conjugation (antisera production), etc. For internal sequences, it may be necessary to cap either or both the N- and C-termini to avoid introducing a charge where there is none in the native sequence. For bulky ligands, it is also common practice to attach a spacer between the peptide and the ligands to minimize the influence of the ligand on the folding of the peptide.

2. Length of Sequence

The purity of a crude peptide typically decreases as the length increases. The yield of peptide for sequences less than 15 residues is usually satisfactory, and such peptides can typically be made without difficulty. In addition, peptides of 10-15 residues in length are satisfactory for raising antisera to linear epitopes of intact proteins.

3. Hydrophobic Stretches

The overall amino acid composition of a peptide is an important design variable that is frequently overlooked. Peptide solubility is strongly influenced by amino acid composition. Peptides with a high content of hydrophobic residues, such as Leu, Val, Ile, Met, Phe and Trp, will either have limited solubility in aqueous solution or be completely insoluble. Under these conditions, it will be difficult to use the peptide in experiments, and it may be difficult to purify the peptide if necessary. It is advisable to keep the hydrophobic amino acid content below 50% and to make sure that there is at least one charged residue for every five amino acids. At physiological pH Asp, Glu, Lys, and Arg all have charged side chains. A single conservative replacement, such as replacing Ala with Gly, or adding a set of polar residues to the N- or C-terminus, may also improve solubility.

4. Difficult Amino Acids

Peptides containing multiple Cys, Met, or Trp residues are also difficult to obtain in high purity, partly because these residues are susceptible to oxidation and/or side reactions. If possible, one should choose sequences to minimize these residues. Alternatively, conservative replacements can be made for some residues. For instance, Norleucine can be used as a replacement for Met, and Ser is sometimes used as a less reactive replacement for Cys. If a number of sequential or overlapping peptides from a protein sequence are to be made, making a change in the starting point of each peptide may create a better balance between hydrophilic and hydrophobic residues. A change in the number of Cys, Met, and Trp residues contained in individual peptides may produce a similar effect. For example, if the stop or start point for choosing peptides separates two Cys residues into two peptides, this may allow better synthesis and purer final product.

5. Secondary Structure

β-sheet formation is a final consideration in peptide design. During synthesis, β-sheet formation causes incomplete solvation of the growing peptide and results in a high degree of deletion sequences in the final product. This problem can be avoided by choosing sequences which do not contain multiple and adjacent residues comprised of Val, Ile, Tyr, Phe, Trp, Leu, Gln, and Thr. If sequences cannot be chosen to avoid stretches of these residues, it often helps to break the pattern by making conservative replacements, for example, inserting a Gly or Pro at every third residue, replacing Gln with Asn, or replacing Thr with Ser.

Analyzing Your Desired Peptide Sequence

Positively charged residues: K, R, H, N-terminus
Negatively charged residues: D, E, C-terminus
Hydrophobic uncharged residues: F, I, L, M, V, W, and Y
Uncharged residues: G, A, S, T, C, N, Q, P, acetyl, amide

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Peptide Design Guideline

Version 3.1, Revision 2011-05-17

1. N-terminus
N-terminal Glutamine (Q) will cyclize to pyroglutamate when exposed to the acidic conditions of cleavage. Recommendation: synthesize with pyroglutamate instead of Q, remove the Q, substitute Q with another amino acid, or acetylate the N-terminus. Any of these suggestions will result in a peptide of higher quality.
N-terminal Asparagine (N) has a protecting group, which can be difficult to remove, when placed at the N-terminus. Recommendation: remove the N or substitute N with another amino acid to the N-terminus.

2. C-terminus
If there is a nonstandard amino acid, including D-amino acids, at the C-terminus, the peptide should be amidated. If there is a modification at the C-terminus (fluorescein, biotin, etc.), the modification must be attached via the side chain of a lysine. These peptides must also be amidated.

3. Sequence
Length: as the sequence length increases, the purity of the peptide decreases. This may result a peptide containing several deletion products. Coupling efficiencies are compromised after ~30 residues. Sequences less than 5 amino acids can be somewhat problematic during the cleavage and purification steps of production. Our baseline for accepting 3-5mer sequences is that the sequence should have at least one hydrophobic residue (L, I, W, V, F, Y, M). It will help in purification process if sequence does not have a hyrdophobic residue, but contains a modification that contributes to the hydrophobic nature (such as Flc, Dansyl, Dabsyl, Btn, Lissamine, etc).

Multiple Prolines (P) in a sequence may undergo a cis/trans isomerization, resulting in a lower purity product.

Adjacent Serines (S) in a sequence frequently result in product that is low in purity and/or contain many deletions.

Multiple Aspartic Acids (D) in a sequence frequently result in the formation of aspartimide adducts, resulting in a product of lower purity.

Multiple modifications within a sequence often result in a product with a low yield and/or purity.

Multiple consecutive Glycines (G) (4 or more) tend to undergo hydrogen bonding (gel formation) in the peptide backbone. The hydrogen bonding may cause difficulty in dissolving and purifying the peptide.

Coupling efficiencies are greatly reduced after a phospho amino acid. Therefore, sequences containing phospho amino acids should have no more than 10 amino acid couplings after the phospho amino acid. Synthesis is performed from the C-term to the N-term. This means that there should be no more than 10 residues after (towards the N-terminus) the phospho amino acid.

4. Solubility
Count the number of charged residues in the peptide, including the uncapped N and C termini.
Typically you want at least 1 charge for every 5 residues. Fewer charges may result in an insoluble product.
Even if a peptide has enough charges, make sure there are not long stretches (more than 5 amino acids) of uncharged residues.

Sequences containing long stretches of charged amino acids or peptides that are short and hydrophilic may not be retained well on the HPLC column, resulting in purification problem.