Designing peptides for peptide antibodies

Rule of thumb for peptide design

1. Length between 10-20 amino acids.
    a. Lack of sequence complexity if too short. The generated antibody is more likely to bind to the exact sequence in off-target proteins.
    b. Decreased synthesized peptide purity if too long. The percentage of peptides carrying the wrong sequence increased exponentially during each round of peptide synthesis. Less likely to generate the antibody bind to the correct sequence in the target protein.

2. Possibly need to include 1 charged amino acid for every 4-5 uncharged amino acids.

3. Contain both hydrophobic and hydrophilic regions, including 1 or 2 hydrophilic amino acids to the N or C terminus to increase hydrophilicity.
    a. Aim to have <25% hydrophobic amino acids as they will likely be soluble in an aqueous solution.
    b. Avoid having 25%-50% hydrophobic amino acids in a sequence that may be **only partly soluble or insoluble**
    c. Not to have >50% hydrophobic amino acids, as it may **not be soluble at all.** If >50% hydrophobic, dissolve sequence in an organic solvent first, to help with solubility.

4. Avoid following amino acids
    a. Internal cysteines, as it could form undesired disulfide bonds. If this cannot be avoided, cysteine can be replaced with serine.
    b. Methionine, as it likely becomes oxidized and can disrupt the cleavage of protecting groups.
    c. Pairing aspartic acid with glycine, proline, or serine, as it causes cleavage of the peptide, making the antigen unstable.
    d. Multiple serine or proline together, as this causes deletions during peptide synthesis.

Tools for peptide design

AbDesginer, hosted and supported by NIH. It identifies optimal immunizing peptides for the production of antibodies.

1. Select the desired type of input and put corresponding information regarding the gene or peptide in the input box. There are 4 types of input allowed:
    a. Gene symbol and organism: abbreviation of a particular gene of interest (ie. TP53).
    b. Swiss-Prot accession number: a sequence of alphanumeric characters (ie. A-Z,0-9)
    c. Swiss-Prot entry name: a protein identification code and a species identification code (ie. P53_Human).
    d. Fasta amino acid sequence: peptide sequence with one letter abbreviations for amino acids (ie. MEDSALD…).
2. Set output antigen peptide length to 10-15 amino acids long, and epitope length to >5 amino acids.
3. Once submitting, multiple antigen sequences will be generated in order of strongest to weakest antigenicity. There are 3 lists to choose from.
    a. Ig-score rank list: Antigen sequences shown from strongest to weakest antigenicity. This is a list may be used if there are no specific limitations (ie. paralogs with similar sequences) for the antigen.
    b. Uniqueness-optimized rank: Shows the Ig-score rank list but with more specificity for the desired protein. May be used when there are other proteins similar in sequence within the same species.
    c. Conservation-optimized rank: Shows the Ig-score rank for sequences more likely to be recognized in proteins of multiple species. May be used if the focus is on proteins that are common to different species.
4. Focus on the antigen sequences that are higher ranked, as they are generally **more immunogenic**

NovoFocus, hosted by NovoPro, based on Neural networks and Deep Learning.

1. Paste protein sequence into NovoFocus and set antigen length to 10-15 amino acids.
2. Antigen sequences produced are scored by NovoFocus; Scores >3 indicate acceptable antigen sequences with larger numbers for better antigenicity.
3. Cross check sequences produced from AbDesign with ones from NovoFocus. Make sure if a sequence is ranked high with one tool, that it is ranked high using the second tool. Avoid sequences that are not ranked high on both tools.

Antigen Profiler Peptide Tool, hosted by ThermoFisher, takes peptide sequence and compares the antigenicity with *in vivo* data.

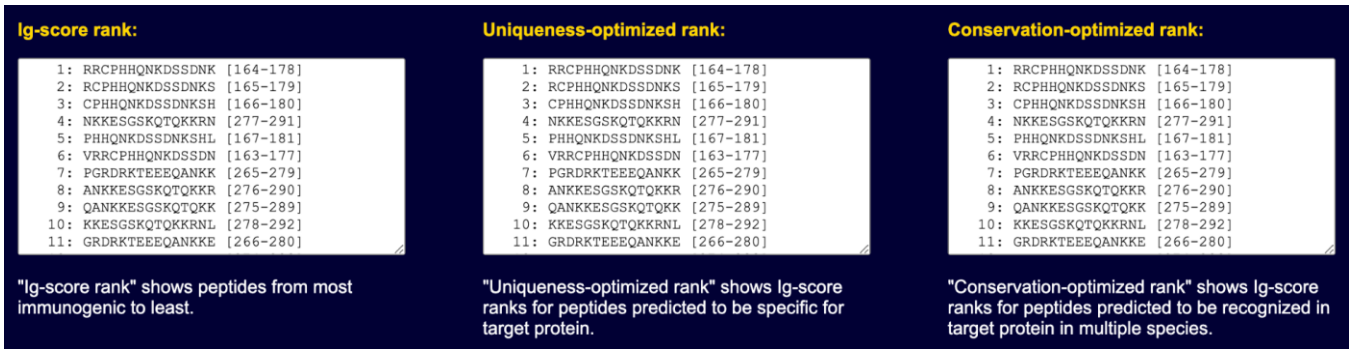
1. Paste the desired antigen sequences from AbDesigner and NovoFocus one by one into the Antigen Profiler Peptide Tool to receive a "score." A score >2.6 is deemed a 'good antigen.' Avoid using sequences that do not have a score >2.6.

Example:

>*Nothobranchius furzeri* TP53

MEDSALDLERHDSFHDMWMDLKDNVYSALESPPIPVTYPDGSDVPDEAWVDQSQMPLLDSQT
YNQLISELPVDMPQKDCILPTSSTVPVTTDHPGDYQLELRFQKSGTTKSVTSTFSEQLNKLFCRL
AKTTPVEVLVSREPPQNAILRATAVYKKSEHVAEAVRRCPHHQNKDSSDNKSHLIRVEGSQLA
QYFEDPFTKRQSVTVPYEPPQLGSEMTTILLSFMCNSSCMGGMNRRPILTILTLETPEGLVLGRR
CFEVRVCACPGRDRKTEEEQANKKESGSKQTQKKRNLAPNTSSLTTPAKKMKSSSSGEDEEKE
MIPLYIQGRKKWNLMKRISDGLDLVEKEKKQLLVQEVLPTSGKRLLKKDRSDSD

Step 1: Generate the three lists. Select the list best fit for the purpose of the experiment. The Ig-score rank was used for this example.



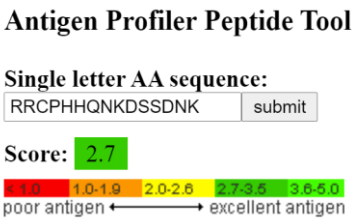
Step 2: Cross check the highly ranked sequences on Abdesigner with NovoFocus. For this example, take sequence 1 and 4 from AbDesign and cross check them with NovoFocus.

LNKLFCRLAKTTPVE	-0.8	120-134
SEQLNKLFCRLAKTT	-0.8	117-131
RRCPHHQNKDSSDNK	-0.8	164-178
QLNKLFCRLAKTTPV	-0.8	119-133
VEKEKKQLLVQEVLP	9	343-357
VLGRRCFEVRVCACP	9	251-265
NKKESGSKQTQKKRN	9	277-291
LVEKEKKQLLVQEVL	8.4	342-356

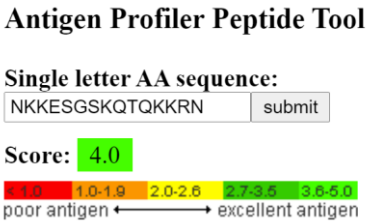
RRCPHHQNKDSSDNK is **below** the recommended score 3 by NovoFocus, so even though it ranked high on AbDesigner, it may not be an ideal sequence.

NKKESGSKQTQKKRN is ranked high on AbDesigner, and has a score **above** 3 on NovoFocus. This should be a good sequence to consider.

Step 3: Enter sequence of choice into Antigen Profiler Peptide Tool to compare antigenicity with *in vivo* data.



Sequence RRCPHHQNKDSSDNK was entered. A score of 2.7 is acceptable. This confirms that this sequence may still be considered.



Sequence NKKESGSKQTQKKRN was entered. A score of 4.0 is very strong. This further confirms that this is an overall good peptide sequence to consider.

Ideal Peptide Sequences for TP53

Antigen Sequence	RRCPHHQNKDSSDNK	NKKESGSKQTQKKRN
Length (residues)	15	15
Hydrophobic Percentage (%)	6.67	6.67
Charged Amino Acid (residues)	6	6
Miscellaneous Problems	Aspartic acid with serine	None
AbDesigner Rank	1	4
NovoFocus Score	-0.8	9
Antigen Profiler Peptide Tool Score	2.7	4.0
Comments	Highly ranked and scored from AbDesigner and Antigen Profiler Peptide Tool. Disagreement with NovoFocus. May still be considered.	All 3 tools agreed that this is an acceptable sequence. This antigen has a high ranking and scores from the 3 tools. This is a good sequence to choose.

v1.0: Nora Singh & Angel Zheng, 2022/01/12