

MBPDB: Help

Contents

1. MBPDB Search.....	1
1.1 Single Peptide Sequence.....	1
1.1.1 Sequence search options.....	2
1.1.2 Select similarity threshold.....	2
1.1.3 Select scoring matrix.....	2
1.2 Protein ID.....	3
1.3 Function.....	3
1.4 Download results.....	3
2. MBPDB add single entry.....	3
2.1 Peptide Sequence (single amino acid).....	3
2.2. Category (optional).....	3
2.3 Protein ID.....	3
2.4 Function (choose from dropdown or type in your own).....	4
2.5 Secondary Function (optional).....	4
2.6 Title.....	4
2.7 Authors.....	4
2.8 Abstract (optional).....	4
2.9 DOI / Patent No.....	4
3. MBPDB add multiple entries.....	4
4. MBPDB add proteins.....	4

1. MBPDB Search

This function is for searching a single peptide or batch of peptides towards the bioactive peptides in the MBPDB database. Several search options are available and they can be used individually or in combination.

1.1 Single Peptide Sequence

Input a single amino acid sequence of you peptide such as ENLLRFFVAPFPEVFG.

Tip: You can also input part of a protein sequence to identify bioactive peptides in that region of the protein. To do this you would have to select precursor in the .

1.1.1 Sequence search options

Select a search option of your desire. The search functions are based on the protein blast () search algorithm.

1.1.1.1 Sequence

Matches the entire input to all database entries and returns matches that has a equal or higher than the threshold value set.

1.1.1.2 Truncated

Searches for any of the bioactive database entries containing the with an equal or higher similarity than the set . The input amino acid sequence is indicated with **bold** text in the result sequences.

1.1.1.3 Precursor

Searches for bioactive peptides within the database which is contained within the with equal or higher similarity than the set .

1.1.2 Select similarity threshold

This option allows for searching database entries with similarities to the input single peptide sequence. The search is based on a scoring matrix which is either or . A restriction of the similarity search is that the peptides is at least four amino acids long. Sequences below four amino acids will automatically only search sequences that has a 100% similarity even though the indicated value is set lower than this.

1.1.3 Select scoring matrix

The scoring matrix assigns an alignment score for all pair of amino acid residue matches to calculate the between the and database entries. Currently the database contains the and scoring matrix.

1.1.3.1 Identity

This scoring matrix is based on exact pairwise amino acid residues matches.

1.1.3.2 BLOSUM62

The BLOSUM62 scoring matrix uses an amino acid substitution matrix for the alignment score of all pairs of amino acid residue matches to calculate the (Figure 1).

C	9
S	-1 4
T	-1 1 5
P	-3 -1 -1 7
A	0 1 0 -1 4
G	-3 0 -2 -2 0 6
N	-3 1 0 -2 -2 0 6
D	-3 0 -1 -1 -2 -1 1 6
E	-4 0 -1 -1 -1 -2 0 2 5
Q	-3 0 -1 -1 -1 -2 0 0 2 5
H	-3 -1 -2 -2 -2 -2 1 -1 0 0 8
R	-3 -1 -1 -2 -1 -2 0 -2 0 1 0 5
K	-3 0 -1 -1 -1 -2 0 -1 1 1 -1 -2 5
M	-1 -1 -1 -1 -2 -1 -3 -2 -3 0 -2 -1 -1 5
I	-1 -2 -1 -3 -1 -4 -3 -3 -3 -3 -3 -3 1 4
L	-1 -2 -1 -3 -1 -4 -3 -4 -3 -2 -3 -2 -2 2 2 4
V	-1 -2 0 -2 0 -3 -3 -3 -2 -2 -3 -3 -2 1 3 1 4
F	-2 -2 -2 -4 -2 -3 -3 -3 -3 -1 -3 -3 0 0 0 -1 6
Y	-2 -2 -2 -3 -2 -3 -2 -1 2 -2 -2 -1 -1 -1 3 7
W	-2 -3 -2 -4 -3 -2 -4 -4 -3 -2 -2 -3 -3 -1 -3 -2 -3 1 2 11
C S T P A G N D E Q H R K M I L V F Y W	

Figure 1: BLOSUM62 substitution matrix (&, 1992).

1.1.4 Get extra output?

This function provides additional information on the subject and query match obtained by protein blast and does not work for peptides below four amino acids long. The additional information obtained for each match are: , query and subject start and stop position, e-value, alignment length, mismatches and gaps.

1.2 Protein ID

Searches for bioactive peptides derived from a specific protein. Input the protein ID also known as protein entry on uniprot.org. For example, bovine beta casein has protein ID 'P02666'. If the protein ID is not present in the database, it can be added in fasta file format through the function. Through this function it is also possible to add genetic variants of a protein and assigning it a new protein ID.

1.3 Function

Searches for bioactive peptides with a specific function. Write your own or select from the dropdown menu (ACE-inhibitory, antimicrobial, DPP-IV-inhibitory, Antioxidant, Immunodmodulatory or Antinflammatory).

1.4 Download results

The search result can be downloaded as a Tab-separated values file by checking this box and opened using e.g. Microsoft excel.

2. MBPDB add single entry

This function is for adding a single entry into the database. Before the entry is added, it needs to be accepted by a database administrator.

2.1 Peptide Sequence (single amino acid)

Input the single letter amino acid sequence of your peptide of the bioactive peptide.

2.2. Category (optional)

Add the category that the bioactive peptide belongs to. This could be Dairy, Cheese, casein etc. Note that this is optional.

2.3 Protein ID

Input the protein ID of the protein from where the bioactive peptide derives. This is similar to the entry identifier for protein in uniprot.org, e.g. P02666 for bovine beta-casein and is a unique ID for each known protein. The input peptides sequence has to match a sequence in the protein ID, otherwise the database will not accept the entry. The database will automatically identify the species based on the protein ID, and stored information about the peptides sequence start and stop position in the protein. In case of a genetic variant of a protein, add the genetic variant as a new protein in fasta format through the function. Use the newly inputted protein ID for the genetic variant in this field afterward.

2.4 Function (choose from dropdown or type in your own)

Input the function of the bioactive peptide you wish to add to the database. Write your own or select from the dropdown menu (ACE-inhibitory, antimicrobial, DPP-IV-inhibitory, Antioxidant, Immunodmodulatory or Antinflammatory).

2.5 Secondary Function (optional)

Add the additional function information in this field in connection to the added. This field is often used to add information on the activity level and assay conditions of the function given in the previous field. If you peptide is ACE-inhibitory, then added the activity. Note that this field is optional but highly recommended.

Example:

If your peptide is ACE-inhibitory and used the assay described by Cushman and Cheung (1971), then your secondary function field should state the activity and assay information: 182 µM, Cushman and Cheung (1971).

2.6 Title

Input the title of the paper published describing the bioactivity.

2.7 Authors

Input the authors of the paper published describing the bioactivity.

2.8 Abstract (optional)

Input the abstract of the paper published describing the bioactivity. Note this is optional.

2.9 DOI / Patent No.

Input the DOI of the paper published describing the bioactivity or Patent no. of the patented peptide. If this information is not available, input N/A.

3. MBPDB add multiple entries

This function is for adding multiple entries into the database by uploading the information as a .tsv file. An example of the file format is available on the online database and an example is shown below. The data required are similar as those for the . Before the entry is added, it needs to be accepted by a database administrator.

	A	B	C	D	E	F	G	H	I
1	proteinID	peptide	category	function	secondary_function	title	authors	abstract	doi
2	P02668	AIPPKKNQD	dairy	ACE-inhibitory	Cushman and Cheung (1971) 19.9 µM	Angiotensin I-converting enzyme inhibitory pep	Quan, Shuang		10.1002/sfa.3394
3	P02663	ALNEINQFY	dairy	ACE-inhibitory	Cushman and Cheung (1971) 219 µM	Angiotensin-I-converting enzyme inhibitory pep	Tauzin, Jérôme		10.1016/S0014-5793(02)03576-7
4	P02663	ALNEINQFYQK	dairy	ACE-inhibitory	Cushman and Cheung (1971) 264 µM	Angiotensin-I-converting enzyme inhibitory pep	Tauzin, Jérôme		10.1016/S0014-5793(02)03576-7
5									

4. MBPDB add proteins

This function is made for adding protein information into the database in fasta format (example shown below). Select one or multiple fasta files to be uploaded and press run. This function should be used when adding entries to the database which gives the error that the is not in the database. This function could also be used to add genetic variant of proteins by adding their sequence information in fasta format. Several proteins are milk proteins are already added to the database and can be viewed through the “View current protein fasta headers (IDs)” link.

```
>sp|P02666|CASN_BOVIN Beta-casein OS=Bos taurus GN=CSN2 PE=1 SV=2  
MKVLILACLVALALARELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDEL  
QDKIHPFAQTQSLVYPFPGPPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK  
HKEMPFPKYPVEPTESQSLTLTDVENLHLPLPLLQSWMHQPHQPLPPTVMFPPQSVLSL  
SQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFFIIIV
```

References:

Henikoff S, Henikoff J.G. 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992 Nov 15; 89(22): 10915-10919.