

PepHemo Instruction

PepHemo is a user-friendly, easy-to-access web server that can be used to predict the hemolytic nature of peptides with sequences as the input. The hemolytic activity of a peptide indicates its ability to destroy red blood cells, leading to the release of hemoglobin. Numerous therapeutic peptides failed in clinical trials just because of their high hemolytic activity. Since experimental determination of the hemolytic potency of a large number of peptides is a labor-intensive, time-consuming, and costly, we developed **PepHemo** which can serve the scientific community to screen hemolytic peptides in an ultrafast manner. The model was constructed (using experimentally determined 1,140 positive and 713 negatives along with 442 randomly generated negative peptides) based on a supervised autoencoder (SAE) with a new Chaos Game Representation (CGR) scheme to characterize peptide sequences. The web server is implemented based on the Flask framework, with Python 3.8.10, Torch 1.11.0, and Flask 2.4.

1. Input to PepHemo

The input from users is simple and straightforward: just the sequences of peptides to be predicted. There are two ways to input peptide sequences, as shown below:

Input of Peptide FASTA Sequences

The screenshot displays the input section of the PepHemo web server. It features a large text area for direct input and a file upload option. Two callout boxes provide instructions for each method.

1. directly input peptide sequences in the fasta format

```
>peptide_43
RLYRKVYGRLYRKVYGRLYRKVYGRLYRKVYGRLYRKVYGRLYRKV
YGRLYRKVYGRLYRKVYGRLYRKVYGRLYRKVYGRLYRKVYGRLYRKV
>peptide_126
DGVRYSPRLRIVQELNAAAGAHG
```

2. upload a file (tmp_peptide.fasta) of peptide sequences in the fasta format

or Upload a File ([Example](#)):
Choose File tmp_peptide.fasta

SUBMIT **HELP**

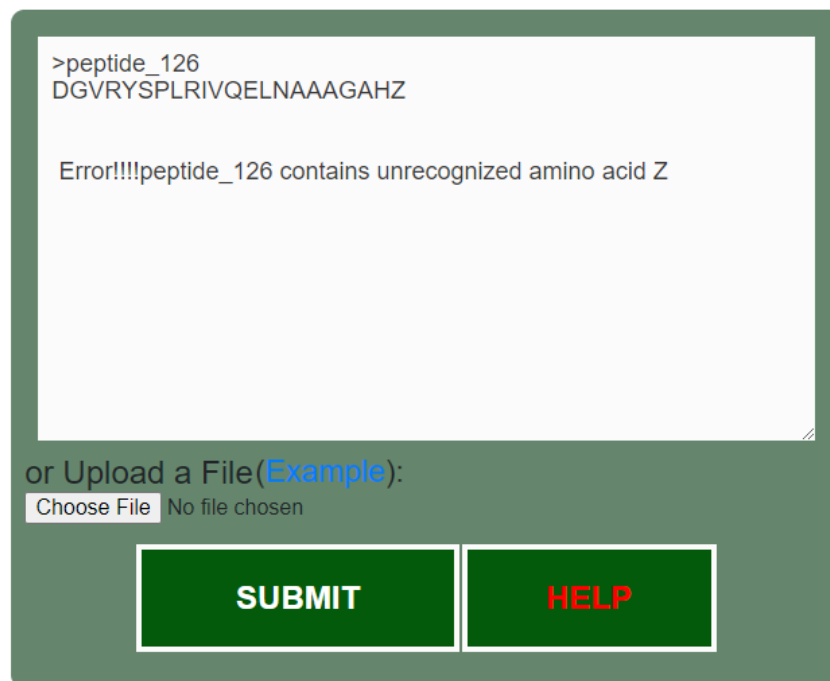
The input peptide sequences must consist of amino acids in single uppercase characters in the fasta format. Otherwise, the input will be considered as “illegal”. If this happens, an error message will show up in the input Form. Here are some examples with incorrect input:

1). Unrecognized amino acid characters

```
>peptide_126  
DGVRYSPLRIVQELNAAAGAHZ
```

An error message will show up to indicate there is an unrecognized amino acid in the sequence (“Z” here).

Input of Peptide FASTA Sequences



The screenshot shows a web form with a text input area containing the following text:

```
>peptide_126  
DGVRYSPLRIVQELNAAAGAHZ
```

Below the input area, an error message is displayed: "Error!!!!peptide_126 contains unrecognized amino acid Z".

Below the error message, there is a section for file upload: "or Upload a File(Example):". This section includes a "Choose File" button and the text "No file chosen".

At the bottom of the form, there are two buttons: "SUBMIT" and "HELP".

2). Input without “>” or names for peptides

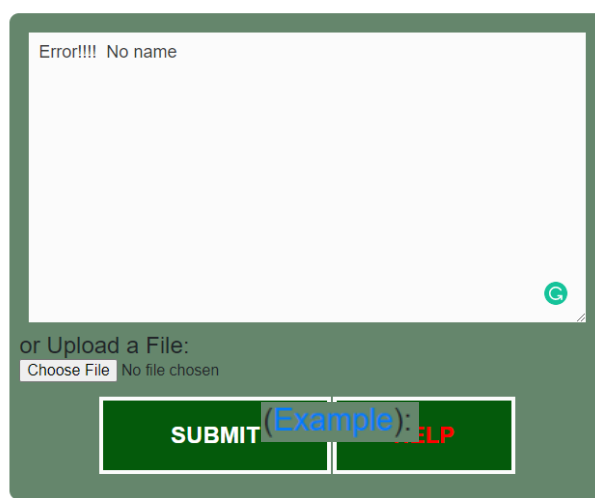
```
peptide_126  
DGVRYSPLRIVQELNAAAGAHG
```

Or

```
>  
DGVRYSPLRIVQELNAAAGAHG
```

An error message will show up to indicate there is no name for the peptide.

Input of Peptide FASTA Sequences



Error!!!! No name

or Upload a File:
Choose File No file chosen

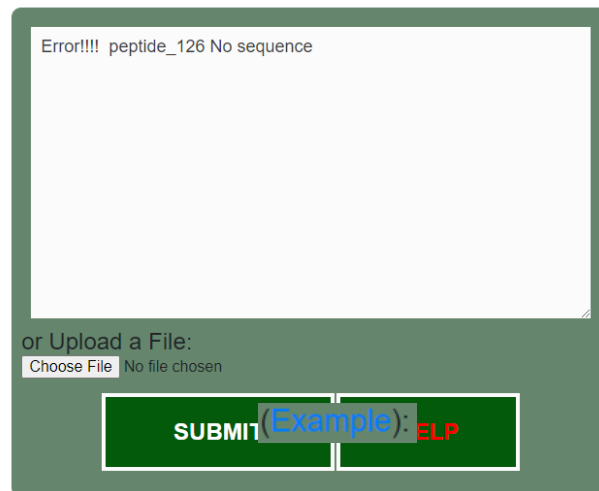
SUBMIT (Example): ELP

3). Peptide name and sequence in the same line

>peptide_126 DGVRYSPLRIVQELNAAAGAHG

An error message will show up to indicate peptide_126 has no sequence.

Input of Peptide FASTA Sequences



Error!!!! peptide_126 No sequence

or Upload a File:
Choose File No file chosen



SUBMIT (Example): ELP

2. Output of PepHemo

The output of **PepHemo** is also easy to understand, with three columns in a table. The first column is the peptide name or ID, as provided by users in the input; The second column is the predicted hemolytic properties, with 1 representing hemolytic and 0 representing non-hemolytic; The third column is the encoded color image of each peptide by our SAE model. The images are meant to illustrate visually how peptides are different from each other and to interpret why and how they have different hemolytic

properties based on the code image colors, patterns, etc. Please refer to our manuscript for more details. The below shows some examples of prediction: peptide_43 is predicted as hemolytic (1) while peptide_126 as non-hemolytic (0), both in agreement with experimental observations.

Output of Predicted Hemolytic Properties ([Download Data](#))

Peptide Name	Hemolytic? Yes (1) No (0)	Encoded Peptide Image
peptide_43	1	
peptide_126	0	



As shown, the images shown in the last column (with zoom-in) illustrate that the two peptides are very different and their predicted hemolytic activities are opposite. For more details about interpretation, please refer to the manuscript.

In addition to visualizing the result in a table, we also dynamically provide users with an option to download the data. This is particularly useful if the input number of sequences is large (e.g., >50). The format of the downloadable table is as follow:

```
Peptide Name      Hemolytic (1) or Non-hemolytic (0)
peptide_43        1
peptide_126       0
```

3. Frequently Asked Questions

This FAQ section is quite self-explanatory. It provides brief answers to the most frequently asked questions, along with some more information about the use of our tools and services.