

## POINTS TO NOTE BEFORE SUBMITTING A JOB TO THE CASCADE PSI-BLAST WEB SERVER.

1. A proper and complete e-mail address in the format is required for processing the results. Please check for errors in spelling and the use of special characters.

Ex: [name@ncbs.res.in](mailto:name@ncbs.res.in)

2. Protein description must be less than 20 characters. Since result files are generated internally using this name, it would be nice to use a meaningful name such as: tim,1dxy,P15265. Special characters such as `_`, `*`, `.` may be avoided.

3. The protein sequence must include only the 20 amino acids. Special characters such as `>`, `B`, `J`, `O`, `U`, `Z`, `X`, `*`, `0-9` should be excluded. Also, please enter the sequence alone.

Sample sequence:

```
QYSALRSNVSMGLGKVLGETIKDALGEHILERVETIRKLSKSSRAGNDANRQELLTTLQNLNDEL
LPVARAFSQFLNLANTAEQYHSISPKGEAASNPEVIARTLRKLNQPELSEDTIKKAVESLSLELV
LTAHPTEITRRTLIIHKMVEVNACLKQLDNKDIADYEHNQLMRRLRQLIA
```

4. The search is currently enabled against three databases. These include:

- Swiss-Prot. (<ftp://ftp.expasy.org/databases/swiss-prot/release>)
- Scop (release 1.69). (<http://astral.berkeley.edu/>)
- Pfam-A seed alignment (release 19.0).  
([ftp://ftp.sanger.ac.uk/pub/databases/Pfam/current\\_release](ftp://ftp.sanger.ac.uk/pub/databases/Pfam/current_release))

5. E-value and H-value must be selected to run Psi-Blast. Your query sequence and all the sequences of each generation will use these E-value and H-value cut-off to run Psi-Blast. By default E-value=0.001 and H-value=0.0001 will be used.

6. The program employs E-value and length of alignment as filters to select hits at every generation of the search. It is, therefore, recommended that these are selected with some discretion. The alignment length criteria is calculated as:

$$(\text{Alignment length}/\text{Query sequence length}) * 100$$

A choice of 70% as the length criteria will examine only those alignments in the output that are atleast 70% the query length. By default, the alignment criteria is set to 70%. We recommend that if your query sequence has multiple domains and you would like to search either the Pfam or Scop database, to select a low value of alignment length criteria, to allow for detection of hits.

\* E-value criteria differs from E-value to run PSI-BLAST and corresponds to the threshold/ cut-off above which hits will be considered as true hits by our server.

7. A low complexity filter is enabled (by default) in our PSI-BLAST searches. This may be disabled.

8. This is an interactive web server. At the end of each generation, the URL with the link to the results page is sent to the user. The user is given a choice to select the hits with which to propagate further searches.  
By default, the top 30 sequences are selected as queries for propagating the next generation of searches. If the user wishes to employ sequences other than those selected by the server as queries, the default selection may be cleared by clicking on the "Deselect all sequences" feature and ticking against any of the sequences listed in the page.  
An upper limit of 30 is set for the maximum number of sequences used to propagate subsequent generations of the search. If more than 30 queries are to be selected, the user is requested to wait until completion of the first 30 runs before selecting more queries with which to propagate the searches.
  9. Hits are annotated with domain boundaries detected in PSI-BLAST searches. The clickable hit links to its corresponding page from SWISS-PROT and PFAM depending on the database selected for the search.
  10. All the hits from each generation are given in a fasta format into a file which can be downloaded.
  11. Results are also provided in tabular format to help assess which hits in each generation served as intermediates in detecting new hits.
  12. Psi-blast output for all the queries for each generation are provided as a gzip file.
  13. The search time depends on the query, choice of database, the search criteria selected and the job queue. Please bear with us for delays in reporting the results.
  14. To report delays, errors or possible bugs in the program, please write an e-mail to [dinesh@mbu.iisc.ernet.in](mailto:dinesh@mbu.iisc.ernet.in).
- 

#### PLEASE CITE

Sandhya S, Chakrabarti S, Abhinandan KR, Sowdhamini R, Srinivasan N: **Assessment of a rigorous transitive profile based search method to detect remotely similar proteins**. J Biomol Struct Dyn. 2005 Dec;23(3):283-98. Click here for [PDF](#).

R. Bhadra, S. Sandhya, K. R. Abhinandan, S. Chakrabarti, R. Sowdhamini and N. Srinivasan: **Cascade PSI-BLAST web server: a remote homology search tool for relating protein domains**. Nucleic Acids Research, 2006, Vol. 34, Web Server issue W143–W146. Click here for [PDF](#).

---