

PFP: Protein Function Prediction server

For any questions regarding PFP contact the administration at info@kiharalab.org

This method allows user to submit protein sequence(s) and obtain annotation prediction for the sequence(s) in terms of [Gene Ontology](#) terms. The underlying PFP algorithm can be found [here](#) . To learn more about each input you can click on the question mark next to it.

Part1: Steps for submitting job as anonymous user

1. Go to <http://www.kiharalab.org/web/software.php> . To start a new job click on the PFP logo highlighted in red below. This will open PFP job submission page in your browser.

Our Software



ESG is our new sequence similarity-based protein function prediction server. In essence, it further applies PFP iteratively and obtains superior performance in terms of prediction accuracy. [Visit the server](#) to submit a sequence or [read the paper in Bioinformatics](#). ESG annotates query sequences with Gene Ontology terms by assigning probability to each annotation. Statistical framework of ESG improves the prediction accuracy by iteratively taking into account the neighborhood of query protein in the similarity based sequence space.



PFP is our sequence similarity-based protein function prediction server designed to predict GO annotations for a query protein sequence beyond what can be found by searching conventional databases. [Visit the server](#) to submit a sequence or [read the paper on Protein Science](#). PFP has achieved the highest total score among participating servers in a function prediction contest held at AFP-SIG'05, ISMB 2005, and also was the best method in the head-to-head performance comparison in the function prediction category at CASP7.

The logo for 3D-SURFER features the text '3D-SURFER' in a bold, black, sans-serif font, enclosed in a white rectangular box with a black border. The background is dark with some colorful, abstract shapes.

3D-SURFER

3D-Surfer is web-based software for protein surface comparison and analysis. The server integrates repertoire of methods to assist in high throughput screening and visualization of protein surface comparisons. The surface representation enables a very fast structure search; It takes a couple of seconds to perform an exhaustive comparison between a single protein surface to all protein structures in

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [?](#) [Clear](#) [Load Sample](#)

or

Upload your FASTA File: [?](#)

No file chosen

Email Notifications

To receive email notifications you must first login or create a new account

Submit

This website is free and open to all users and there is no login requirement.

2. You can enter one or more query protein sequence(s) in [fasta](#) format in the text box titled “Enter Query Sequence(s)”.

Consider the following sequence that you can enter.

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMFIIYISWYKIEHICTSDNWMDRFRNAYVWVQNPLKVLKCHQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
```

You can also click on “Load Sample” link to load this sequence in the text box and follow the next steps. Clicking on “Clear” link will clear the text box for sequence and load default ESG parameters in the boxes below.

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [\[?\] Clear](#) [Load Sample](#)

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMF IYISWYKIEHICTSDNWMDFRNAYVWVQNPLKVLKCHQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
```

or

Upload your FASTA File: [\[?\] Choose File](#) No file chosen

Email Notifications

To receive email notifications you must first login or create a new account

Submit

This website is free and open to all users and there is no login requirement.

[Submit](#)

Or you can click on browse button and upload a fasta file containing query sequences. Currently we support maximum 1000 query sequences in one job.

To clear sequence, click on clear link next to the label for protein sequence text box.

3. When you are submitting the job without logging in, you can book mark the results url and refer back to the job results. If you want to get email notifications about your jobs, view their progress/status and refer to all your jobs at once place, you should register with the website. This will allow you to log in and submit jobs and enjoy additional features free of cost.

4. After loading the query sequence(s) click on submit button. This will take you to the results page and show your job waiting for processing. Bookmark this url and refer to it for the results. Or you can save/remember the job id number to come back to the results.

Job Status

Job ID	Type	Status	Date Submitted
000426	PFP	Job Processing	2009-12-31

You may bookmark this page to view your results at a later date or use the Job Retrieval form with your Job ID#.

When you job status shows complete, click on the jobID to view predicted Gene Ontology terms for the query protein. Or you can refresh the page to go to results. The

results page will show PFP input sequence and prediction terms for each Gene Ontology category (Molecular function, Biological Process and Cellular Component) that have confidence greater than 5%. The results page will also provide link to the results in xml format which you can download for further processing.

PFP Job Results

PFP Parameters

Protein Sequence:

```
MASSLKINGTLLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEALKDKSS
HMFYIYSWYKIEHICTSDNWMDFRFRNAYVWVQNPLKVLKCHQENSKNSYTESSRFMYIEFHCSMD
GYVDSIEDLKMVEPIGM
```

[Download](#)

Molecular Function Terms

Probability	Term	Description
92%	GO:0004522	pancreatic ribonuclease activity
87%	GO:0005488	binding
85%	GO:0016638	oxidoreductase activity, acting on the CH-NH2 group of donors
83%	GO:0003676	nucleic acid binding
62%	GO:0046914	transition metal ion binding
60%	GO:0043169	cation binding
58%	GO:0043167	ion binding
56%	GO:0003705	RNA polymerase II transcription factor activity, enhancer binding
52%	GO:0046872	metal ion binding
50%	GO:0005515	protein binding

Biological Process Terms

Probability	Term	Description
68%	GO:0044270	nitrogen compound catabolism
66%	GO:0009310	amine catabolism
65%	GO:0044237	cellular metabolism
59%	GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolism
58%	GO:0050876	reproductive physiological process
57%	GO:0050875	cellular physiological process
56%	GO:0044260	cellular macromolecule metabolism
55%	GO:0044238	primary metabolism
53%	GO:0043170	macromolecule metabolism
53%	GO:0008152	metabolism

Cellular Component Terms

Probability	Term	Description
60%	GQ:0005623	cell
59%	GQ:0005622	intracellular
59%	GQ:0016020	membrane
57%	GQ:0000151	ubiquitin ligase complex
56%	GQ:0005737	cytoplasm
56%	GQ:0043229	intracellular organelle
54%	GQ:0005634	nucleus
53%	GQ:0005773	vacuole
53%	GQ:0043231	intracellular membrane-bound organelle
51%	GQ:0005739	mitochondrion

5. If you close your browser and need to go back to your job results then you can use the “Job Retrieval” text box on the right side of the web page. It has been highlighted in red in the screenshot below. When you enter a valid job id and click on “Fetch Job”, you will get to view the job status or results depending on if the job is complete or not.

Member Portal Access

Username:

Password:

[Forgot your password?](#)
[Register](#)

Job Retrieval

Job ID#:

Part2: Steps for logging in and submitting ESG job

1. Open PFP job submission page in your browser.

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [\[?\] Clear Load Sample](#)

or

Upload your FASTA File: [\[?\] Choose File](#) No file chosen

Email Notifications

To receive email notifications you must first login or create a new account

Submit

This website is free and open to all users and there is no login requirement.

Kihara Lab Software

Thank you for using Kihara Lab Web Servers. Please refer to the tutorials below for getting started with our servers. You can register with the lab and enjoy additional features free of cost. We would be delighted to receive [feedback](#) from you.

Tutorials

[ESG Tutorial](#)
[PFP Tutorial](#)

Member Portal Access

Username:

Password:

[Register](#)

Job Retrieval

Job ID#:

2. Click on the “Register” link on the right side of the page, it is highlighted in red in the screenshot above. It will take you to the account creation page shown below.

New Account Creation

Account Settings

Username: *

Email Address:

Receive Email Notifications

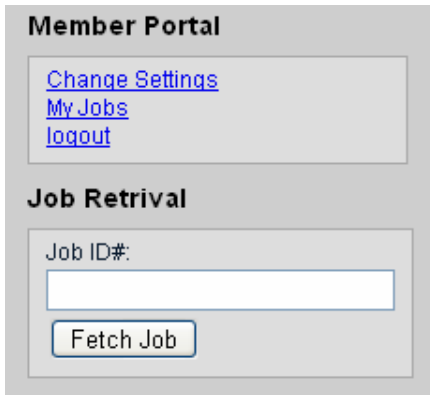
Password: *

Confirm Password: *

* denotes required field

Here you should enter your username and password information. The email address is optional; you should enter it if you wish to receive email notifications about your job status. If you don't check the "Receive Email Notification" check box you will have to log in to the server and check the status of your jobs.

3. Once you have your account set up you can log in and submit jobs for processing. Enter your username and password in the text boxes provided on the right side of the web page. This will redirect you to the page displaying all your jobs with their status and results links.



Member Portal

[Change Settings](#)
[My Jobs](#)
[logout](#)

Job Retrieval

Job ID#:

You can click on "My Jobs" and view your existing jobs and their status. The complete jobs have clickable links where you can view results of those jobs.

Your jobs

Job ID	Type	Status	Date Submitted	Batch Job
000424	ESG	Done - View or Download	2009-12-31	No
000423	PFP	Done - View or Download	2009-12-31	No
2 Jobs				

To start a new job click on "Software" tab at the top and select PFP by clicking on the logo highlighted in red below.

Our Software



ESG is our new sequence similarity-based protein function prediction server. In essence, it further applies PFP iteratively and obtains superior performance in terms of prediction accuracy. [Visit the server](#) to submit a sequence or [read the paper in Bioinformatics](#). ESG annotates query sequences with Gene Ontology terms by assigning probability to each annotation. Statistical framework of ESG improves the prediction accuracy by iteratively taking into account the neighborhood of query protein in the similarity based sequence space.



PFP is our sequence similarity-based protein function prediction server designed to predict GO annotations for a query protein sequence beyond what can be found by searching conventional databases. [Visit the server](#) to submit a sequence or [read the paper on Protein Science](#). PFP has achieved the highest total score among participating servers in a function prediction contest held at AFP-SIG'05, ISMB 2005, and also was the best method in the head-to-head performance comparison in the function prediction category at CASP7.

The logo for 3D-SURFER features the text '3D-SURFER' in a bold, black, sans-serif font. The text is set against a white rectangular background with a slight shadow effect. The background of the entire logo area is dark with some colorful, abstract shapes.

3D-SURFER

3D-Surfer is web-based software for protein surface comparison and analysis. The server integrates repertoire of methods to assist in high throughput screening and visualization of protein surface comparisons. The surface representation enables a very fast structure search; It takes a couple of seconds to perform an exhaustive comparison between a single protein surface to all protein structures in

This will direct you to the PFP job submission page.

4. You can enter one or more query protein sequence(s) in [fasta](#) format in the text box titled “Enter Query Sequence(s)”.

Consider the following sequence that you can enter.

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMF IYISWYKIEHICTSDNWMDFRNAYVWVQNPLKVLKCHQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
```

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [\[?\] Clear Load Sample](#)

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLALLCILCTLLVQSKEVSWREFMRQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMFIIYISWYKIEHICTSDNWMDFRNAYVWVQNPVKVLEKCHQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
```

or

Upload your FASTA File: [\[?\] Choose File](#) No file chosen

Email Notifications

Receive email notifications for this job

Emails will be sent to your account email address, currently (*mgillogly@gmail.com*)

Submit

This website is free and open to all users and there is no login requirement.

Or you can click on browse button and upload a fasta file containing query sequences. Currently we support maximum 1000 query sequences in one job.

5. After loading the query sequence(s) click on submit button. This will take you to the results page and show your job waiting for processing. You can log in again later to check the status of your job and access the results when it is complete. You can also enter job id in “Job Retrieval” text box on right side of the page and click “Fetch Job” to view the status of job and results. If you have chosen to get email notifications in your account settings then you will be notified by an email when your job is complete.

6. When you job status shows complete, click on the jobID to view predicted Gene Ontology terms for the query protein. The results page will also provide link to the results in xml format which you can download for further processing.

Part3: Steps for submitting a batch job using ESG

1. Open PFP job submission page in your browser.

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [\[?\] Clear](#) [Load Sample](#)

or

Upload your FASTA File: [\[?\] Clear](#)

No file chosen

Email Notifications

To receive email notifications you must first login or create a new account

Submit

This website is free and open to all users and there is no login requirement.

2. You can enter more than one query protein sequences in [fasta](#) format in the text box titled “Enter Query Sequence(s)”.

Consider the following sequences that you can enter.

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMFYIYISWYKIEHICTSDNWMDRFRNAYVWVQNPLKVLKCHQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
>sp|P13811|ELBH_ECOLX
MNKVKFYVLF TALLSSLCAHGAPQSITELCSEYHNTQIYTINDKILSYTESMAGKREMI
ITFKSGATFQVEVPGSQHIDSQKKAIERMKDTRLRITYLTETKIDKLCVWNNKTPNSIAAI
SMEN
>sp|P05928|VPR_HV1BR
MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWHLGLGQHIYETYGDTWAGVEAI
IRILQQLLFIFRIGCRHSRIGVTQQRARRNGASRS
```

PFP: Automated Protein Function Prediction Job Submission

Enter Query Sequence(s)

Enter your protein sequence here: [\[?\] Clear Load Sample](#)

```
>sp|P56851|EP3B_HUMAN
MASSLKIWGTLALLCILCTLLVQSKEVSWREFMKQHLYLSPSREFREYKCDVLMRENEAL
KDKSSHMFYISWYKIEHICTSDNWMDRFNAYVWVQNPVKVLCHEQENSKNSYTESRSF
NYIEFHCSMDGYVDSIEDLKMVEPIGN
>sp|P13811|ELBH_ECOLX
MNKVKFYVLFATALLSSLC AHGAPQSITELCSEYHNTQIYTINDKILSYTESMAGKREMI
```

or

Upload your FASTA File: [\[?\] Choose File](#) No file chosen

Email Notifications

To receive email notifications you must first login or create a new account

Submit

This website is free and open to all users and there is no login requirement.

Or you can click on browse button and upload a fasta file containing query sequences. Currently we support maximum 1000 query sequences in one job.

3. Follow steps 3 to 5 from part 1 if you are submitting the job as anonymous or steps 5 to 6 from part 2 if you are logged in as a registered user. On the job status page in both cases the job type will be shown as batch.

Job Status

Job ID	Type	Status	Date Submitted
000403	PFP	Job Processing	2009-12-20

You may bookmark this page to view your results at a later date or use the Job Retrieval form with your Job ID#.

Once the processing is complete the results will be displayed in the table with each sequence, its identifier and a clickable link which will take you to detailed results for that sequence.