

PERCEPTRON *v 1.0.0.0*

A Next Generation Top-Down Proteoform
Identification and Characterization Platform

USER MANUAL

Biomedical Informatics Research Laboratory, Department of Biology
Lahore University of Management Sciences
<http://birl.lums.edu.pk/>

Last Modified: 20-12-2020

1. Table of Contents

1. Table of Contents	2
2. Table of Figures	4
3. Introduction to PERCEPTRON	6
3.1. About PERCEPTRON	6
3.2. Features	6
4. Hardware and Software.....	8
4.1. Hardware	8
4.2. Software	8
4.3. Testing.....	8
5. Video Tutorials	9
6. Getting Started with PERCEPTRON.....	9
7. GUI Description.....	11
7.1 Window 1: PERCEPTRON Tool for Top-down Proteomics.....	11
7.2 Window 2: Protein Search Query.....	12
7.2.1 Basic Parameters.....	12
7.2.2 Set Experimental Parameters	13
7.2.3 Set <i>De Novo</i> Sequencing Parameters	14
7.2.4 Set Protein Modifications Parameters.....	15
7.2.5 Set Scoring Components Weight	16
7.3 Window 3: User Search Results.....	17
7.4 Window 4: Summary and Detailed Results View.....	18
8. Interpreting Search Results	21
9. Proteoform Search Using Formatted Data Files	25

9.1	File Formats.....	25
9.1.1	Raw to mzXML File Format Conversion	25
9.1.2	Raw to mzML File Format Conversion	25
9.1.3	MzXML to MGF File Format Conversion	26
9.1.4	MGF to Flat Text File Format Conversion	26
9.2	Parameters	27
9.3	Databases.....	27
9.4	Modes	27
10.	Deploying PERCEPTRON	28
10.1	Installing Visual Studio Code (VS Code).....	28
10.2	Installing SQL Server Management Studio (SSMS).....	30
10.3	Installing Node.js.8.11.1	31
10.4	Installing SQL 2017	33
10.5	Installing Angular 1.7.4.....	34
10.6	Importing files into SQL.....	36
11.	Building PERCEPTRON	37
11.1	PERCEPTRON Application Programming Interface (API)	37
11.2	PERCEPTRON Front-End	41
11.3	Connecting PERCEPTRON API to PERCEPTRON Front-End.....	42
11.4	PERCEPTRON Local Service.....	45
12.	References.....	48

2. Table of Figures

Figure 1. PERCEPTRON Homepage and Log in button	9
Figure 2. PERCEPTRON Login options	10
Figure 3. PERCEPTRON - Overview of User Interface	11
Figure 4. PERCEPTRON - Overview of Basic Parameters	12
Figure 5. PERCEPTRON - Overview of Experimental Parameters	13
Figure 6. PERCEPTRON - Overview of De Novo Sequencing Parameters	14
Figure 7. PERCEPTRON - Overview of Protein Modifications Parameters	15
Figure 8. PERCEPTRON - Overview of Scoring Component Weight	16
Figure 9. User Search History window showing search results	17
Figure 10. Protein spectral matches window showing results for the protein search query	17
Figure 11. Summary Results window showing candidate proteins	18
Figure 12. Detailed Results window showing search parameters and general results (top). Detailed Protein Hit View window showing Protein Search Time and Protein Sequence (bottom)	19
Figure 13. Detailed Results Visualization window with mass spectral visualization	19
Figure 14. Detailed Results Visualization window with Detailed Protein Hit View (left) and Legends (right)	20
Figure 15. Detailed Results Visualization window with Matched Fragments	20
Figure 16. Conversion of raw to mzXML	25
Figure 17. Conversion of raw to mzML	26
Figure 18. Conversion of mzXML to MGF	26
Figure 19. Load Default Parameters	27
Figure 20. Selecting Protein Database	27
Figure 21. Downloading Visual Studio Code (VS Code) setup	28
Figure 22. Choosing destination folder to download VS Code setup	28
Figure 23. Steps to install Visual Studio Code setup	29
Figure 24. Downloading SQL Server Management Studio (SSMS) setup	30
Figure 25. Steps to install SQL Server Management Studio (SSMS)	30
Figure 26. Downloading Node.js (version 8.11.1) setup	31
Figure 27. Steps to install Node.js setup	32

Figure 28. Downloading SQL 2017 setup	33
Figure 29. Steps to install SQL 2017 setup	33
Figure 30. Downloading Angular 1.7.4 setup	34
Figure 31. Steps to install Angular 1.7.4	35
Figure 32. SQL Server Management Studio Pop Up	36
Figure 33. Building PERCEPTRON API. (a) Opening solution PerceptronAPI.sln; (b) Opening WebApiConfig.cs from Solution Explorer of Visual Studio	37
Figure 34. Searching for a statement in PERCEPTRON API solution	37
Figure 35. Copying Server Name from SQL Server Management Studio	37
Figure 36. (a) Opening Web.config from Solution Explorer of Visual Studio (b) Changing data source (Server Name) in Web.config	38
Figure 37. Changing server name in Sqldatabase.cs	38
Figure 38. Building PERCEPTRON API	38
Figure 39. Steps to build PERCEPTRON API	39
Figure 40. Publishing Perceptron API	39
Figure 41. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine”	40
Figure 42. Opening PerceptronFrontEnd folder in Visual studio code	41
Figure 43. Installing extension in VS Code	42
Figure 44. Building Perceptron Front-End	42
Figure 45. Hosting Perceptron Front-End on system with default URL	43
Figure 46. Steps to build PERCEPTRON Front-End	44
Figure 47. Copying Server Name from SQL Server Management Studio	45
Figure 48. Building Perceptron Local Service. (a) Opening PerceptronLocalService.sln; (b) Opening App.config from Solution Explorer of Visual Studio	45
Figure 49. Changing data source value (Server Name) in App.config	45
Figure 50. Steps to execute Protein Search Query	46
Figure 51. Testing Perceptron Local Service	47

3. Introduction to PERCEPTRON

This chapter introduces the user to the PERCEPTRON platform along with a description of its features.

3.1. About PERCEPTRON

PERCEPTRON is a freely available web-based proteoform identification pipeline for Top-Down Proteomics (TDP). Top-down proteomics is an emerging experimental protocol for analysis of intact proteoforms. PERCEPTRON search pipeline brings together algorithms for: (i) intact mass tuning, (ii) *de novo* peptide sequence tag extraction, (iii) *in silico* spectral comparison, (iv) identification of post-translational modifications as well as truncated proteins, and (v) a novel composite scoring scheme for candidate protein scoring. PERCEPTRON achieves high performance by leveraging NVIDIA GPU technology coupled with Microsoft ASP.NET and ANGULAR frameworks. The search results obtained include a list of proteins, their scores and details on the matching information. This information can be visualized as well as downloaded. Overall, PERCEPTRON is aimed at filling the crucial void of open-source and open-architecture protein identification software for TDP data, employing state-of-the-art algorithms.

3.2. Features

The salient features of the pipeline are summarized below:

- **Graphical User Interface (GUI)** - A set of rich and intuitive graphical user interface has been developed for setting up the search parameters as well as for integrating the main components of the engine.
- **Whole Protein Molecular Weight Estimation** - The protein identification begins with the tuning of precursor protein's monoisotopic MW (MS1) as guided by its fragmentation spectra (MS2). Relative abundances and mass/charge (m/z) ratios are used to calculate the consensus MW which is then employed in the search and scoring process.
- **Peptide Sequence Tag Extractor** - Peptide sequence tag ladders (PST) are extracted from the spectra by enumerating successive peaks having MW differences equal to an amino acid and within the user specified mass tolerance. Protein database is then filtered for proteins reporting these PSTs. The length of PST ladders, cumulative mass off-sets and relative abundances are used in calculating the PST scores.
- ***In silico* fragmentation** – *In silico* fragments of candidate proteins are generated by the user selected fragmentation techniques. *In vitro* and *in silico* spectral comparisons are performed and scored.
- **Post-translational Modification (PTM) Search** - Support for predicting typical PTMs has been provided in the tool. Users can select and search variable and fixed PTMs of their choice along with blind-PTMs by simply selecting them from the GUI.
- **Multifactorial Composite Scoring System** - A multifactorial candidate protein scoring scheme incorporating the aforementioned algorithms has been developed. User customization of the parameters and weights in the scoring function is admitted via a GUI.
- **Single and Batch Search** – PERCEPTRON provides support for search in single as well as batch modes. Towards an automated processing of multiple spectral data files, a batch processing mode allows for the selection of multiple files from the folder by clicking the attach file button. The

PERCEPTRON: User Manual

experimental spectra, search parameters and results are automatically stored in the project directory for further processing and visualization.

4. Hardware and Software

4.1. Hardware

For in-house deployment of PERCEPTRON, the hardware requirement includes a graphical processing unit (GPU) that supports CUDA TOOLKIT 7.0.

4.2. Software

Compilation and software build of PERCEPTRON requires the following software:

- Windows Server 2012 R2
- Visual Studio 2013
- Angular 1.7.4
- Node.js 8.11.1
- SQL Server Management Studio (17.6 or 17.9.1)
- CUDA TOOLKIT 7.0
- CUDAFy.NET.1.29.5576.13786
- Microsoft Windows Server v6.2
- .NET Framework 4.5

4.3. Testing

The freely available public deployment of PERCEPTRON is hosted on a Dell Power Edge R730, 2 x Intel Xeon E5-2620, 160 GB RAM (16GBx10) and an NVIDIA Tesla K40C (2880 Cores). Following Windows versions are compatible with PERCEPTRON:

- Windows 8.1
- Windows 7
- Win Server 2012 R2
- Win Server 2008 R2

5. Video Tutorials

Several video tutorials have been developed for employment of PERCEPTRON in protein search. These tutorials are available as a playlist at: https://www.youtube.com/playlist?list=PLaNVq-kFOn0Z_7b-iL59M_CeV06JxEXmA

6. Getting Started with PERCEPTRON

PERCEPTRON manual, samples and issues database is freely available (under the MIT open license) at (<https://perceptron.lums.edu.pk/>)

To initiate use of PERCEPTRON, the user has to register and log in by clicking on the link provided above.

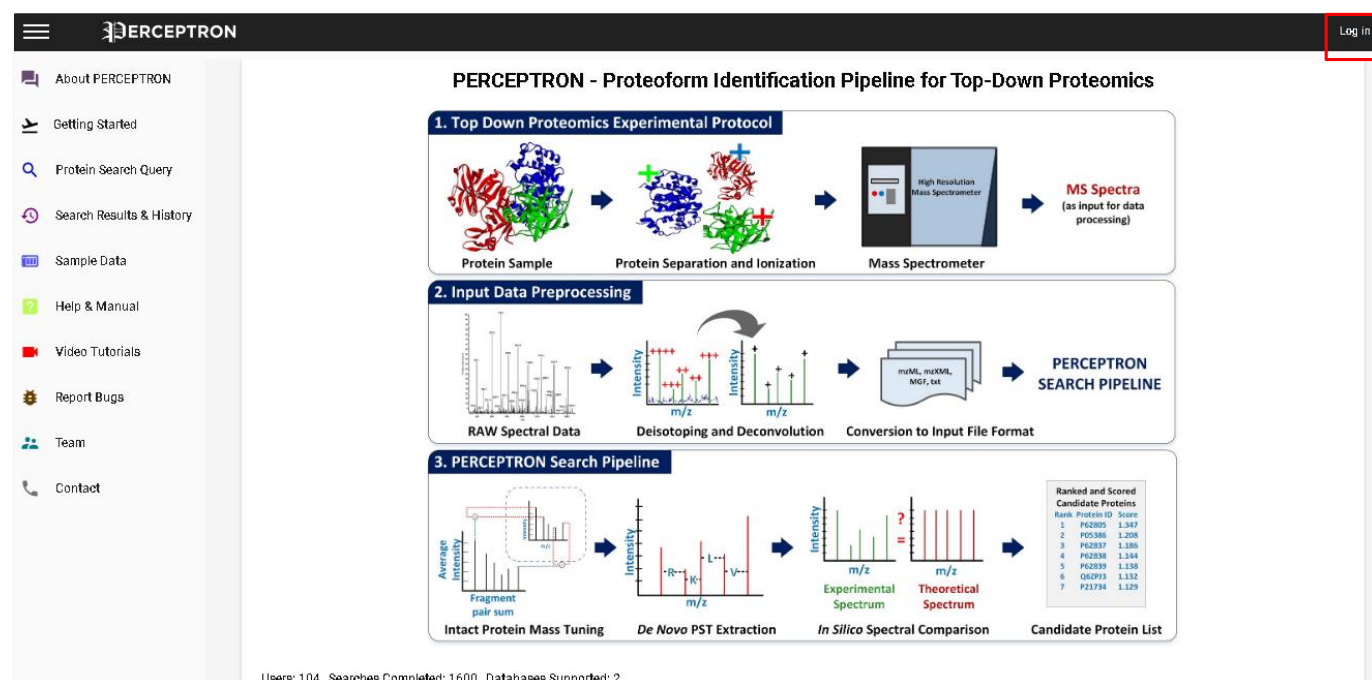


Figure 1. PERCEPTRON Homepage and Log in button

Click the **'Log in'** button on the top right. A window will appear. Enter user credentials to proceed.

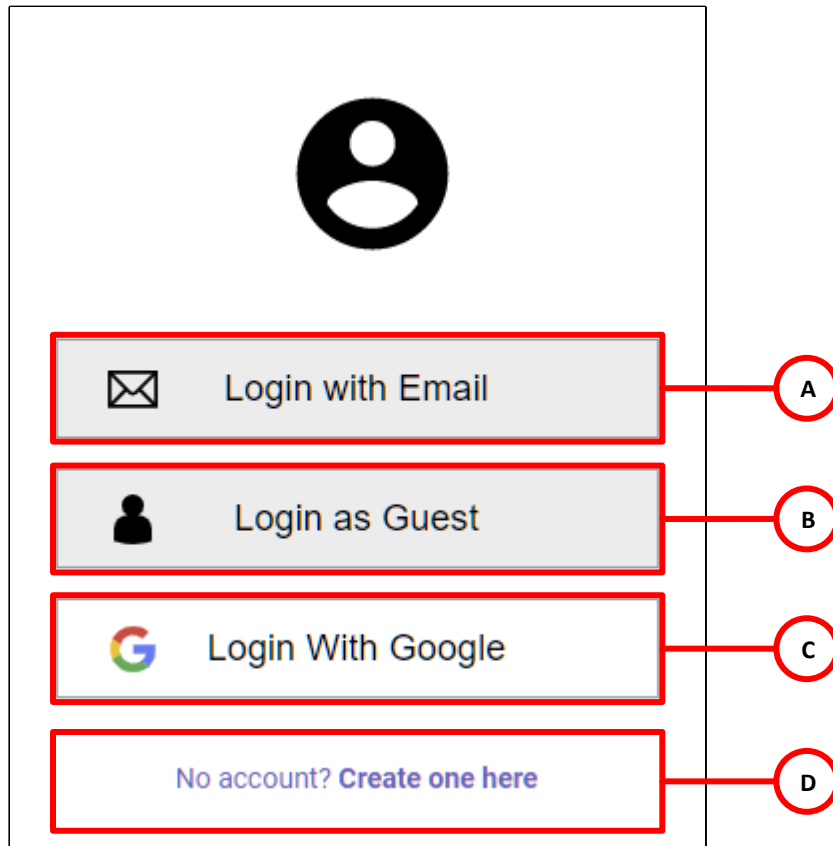


Figure 2. PERCEPTRON Login options

- A. User can create PERCPETRON account to login using an existing email account
- B. Enables user to login as a guest* without the need to create a PERCEPTRON user account to perform search in PERCEPTRON
- C. User can login using an existing google account to use PERCEPTRON
- D. Create a user account to perform search in PERCEPTRON

*Note that results will not be saved in this case.

7. GUI Description

This chapter presents the interface overview for user facilitation.

7.1 Window 1: PERCEPTRON Tool for Top-down Proteomics

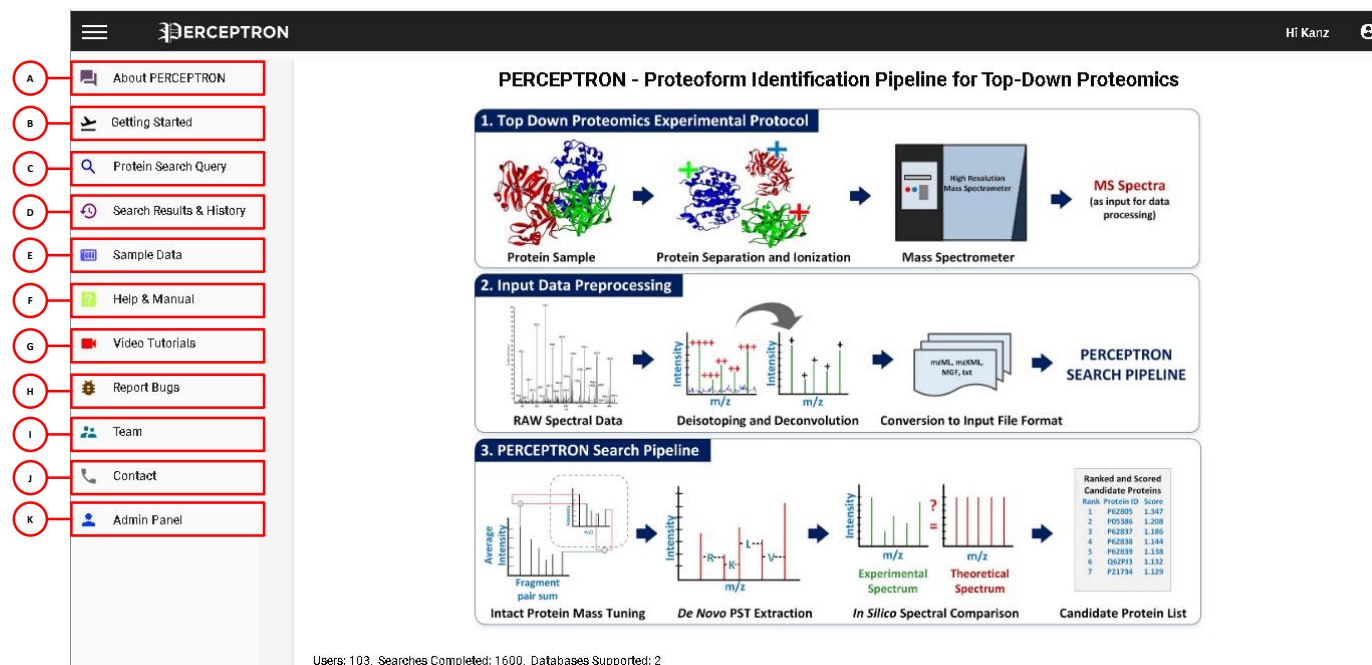


Figure 3. PERCEPTRON - Overview of User Interface

- A. About PERCEPTRON: A next-generation top-down proteoform search and identification platform
- B. Getting Started: Quick guide to proteoform search and identification using PERCEPTRON
- C. Protein Search Query: Job submission - Search top-down proteomics data files to identify and characterize proteoforms
- D. Search Results & History: View search results and history
- E. Sample Data: Sample top-down proteomics data for use with PERCEPTRON
- F. Help & Manual: Get assistance with using PERCEPTRON and download manual
- G. Video Tutorials: View step-by-step video tutorials demonstrating usage of PERCEPTRON
- H. Report Bugs: Report problems and issues here
- I. Team: PERCEPTRON project team members
- J. Contact: Contact us for further information
- K. Admin Panel: For updating or downloading protein database

In order to start protein search, click on the 'Protein Search Query' tab and the following window will appear:

7.2 Window 2: Protein Search Query

7.2.1 Basic Parameters

The screenshot shows the 'Submit Protein Identification Query' form in the PERCEPTRON application. The form is divided into several sections. On the left, a sidebar contains navigation links: About PERCEPTRON, Getting Started, Protein Search Query (highlighted), Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Team, Contact, and Admin Panel. The main content area is titled 'Submit Protein Identification Query' and includes a 'Required Fields *' warning. The 'Basic Parameters' section contains the following fields: 'Protein Search Title *' (annotated with A), 'Select Database *' (annotated with B), 'Enter your E-mail Address' (annotated with D), 'FDR Cut-off' (set to 1, annotated with E), and 'Relevé Top' (set to 10, annotated with F). There is also an 'Upload Mass-Spectrometry Data *' button with a file icon (annotated with C) and a note 'Please select a data file, zip file...'. Below the basic parameters are four expandable sections: 'Set Experimental Parameters', 'Set De Novo Sequencing Parameters', 'Set Protein Modifications Parameters', and 'Set Scoring Components Weight *'. At the bottom right are 'Reset' and 'Submit' buttons.

Figure 4. PERCEPTRON - Overview of Basic Parameters

- A. In order to start protein search, user must enter the 'Protein Search Title' (for example: "MyProject")
- B. Select protein database
- C. Browse and upload experimental data (.mzXML/ .MGF/ .txt) for Single mode; Peak-list files for Batch mode
- D. Enter the email address at which the link for downloading results will be sent. Note that users are notified about the availability of search results upon completion of the job
- E. Enter an FDR percentage to shortlist identified proteoforms at desired FDR cut off
- F. Provide the number of candidate protein hits to be received in results.

7.2.2 Set Experimental Parameters

The screenshot displays the PERCEPTRON web application interface. The sidebar on the left contains navigation links: About PERCEPTRON, Getting Started, Protein Search Query, Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Team, Contact, and Admin Panel. The main content area is titled 'Set Experimental Parameters' and includes several sections: 'Set Experimental Parameters' (with sub-sections for Mass Mode, Peptide Tolerance, Slider Value, Select Fragmentation Type, Filter Database using MS1, Tune Intact Protein Mass, Intact Mass Tolerance, Neutral Mass Loss, and Select Special Ions), 'Set De Novo Sequencing Parameters', 'Set Protein Modifications Parameters', and 'Set Scoring Components Weight *'. The interface also features a top header with a database selector, an upload button for mass spectrometry data, and a user profile section. Red boxes and letters A-I highlight specific features: A (Mass Mode), B (Peptide Tolerance), C (Slider Value), D (Select Fragmentation Type), E (Filter Database using MS1), F (Intact Mass Tolerance), G (Neutral Mass Loss), H (Tune Intact Protein Mass), and I (Select Special Ions).

Figure 5. PERCEPTRON - Overview of Experimental Parameters

- Select Mass Mode. MS data can only be provided in either m/z form with $z = 1$ or neutral masses
- Select the tolerance value for Peptide and select its unit
- Select the 'slider value' to provide a range for tuning MS1 using 'Tune Intact Protein Mass' option
- Select the 'Fragmentation type' from drop down menu
- User can filter database by checking the option 'Filter Database using MS1'
- Set the tolerance value for Protein Mass and select its unit
- Provide the value of Neutral loss, if any
- Check the option 'Tune Intact Protein Mass' to allow for tuning of MS1 using MS2 data
- Choose the corresponding special ions for the type of fragmentation selected (i.e. a', b', y', z'', a*, b*, y*, z' ions)

7.2.3 Set *De Novo* Sequencing Parameters

The screenshot shows the PERCEPTRON web interface. The sidebar on the left contains links: About PERCEPTRON, Getting Started, Protein Search Query, Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Team, Contact, and Admin Panel. The main content area has a top section with 'Peptide Tolerance *', 'Tune Intact Protein Mass' checkbox, 'Neutral Mass Loss', 'Slider Value', 'Select Fragmentation Type *', and 'Select Special Ions'. Below this is the 'Set De Novo Sequencing Parameters' section, which includes:

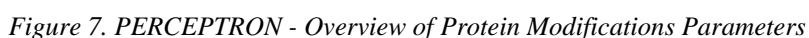
- A**: ☐ Enable PST Filtering
- B**: Tolerance For Each Hop (Da)
- C**: Minimum Tag Length (dropdown menu)
- D**: Maximum Tag Length (dropdown menu)
- E**: Overall tolerance for PST (slider)

 Below the 'Set De Novo Sequencing Parameters' section are 'Set Protein Modifications Parameters' and 'Set Scoring Components Weight *' sections. At the bottom right are 'Reset' and 'Submit' buttons.

Figure 6. PERCEPTRON - Overview of *De Novo* Sequencing Parameters

- A. Check 'Enable PST Filtering' to filter PSTs
- B. Set the 'Tolerance for each Hop'
- C. Tags will be filtered above the minimum length of PST selected from the drop-down menu by the user
- D. Tags will be filtered below the maximum length of PST selected from the drop-down menu by the user
- E. Overall mass error tolerance shows error margin for the whole PST

Note: Input value for Minimum and Maximum Tag length should be different. Minimum Tag Length should be smaller than Maximum Tag Length.



- 15

7.2.5 Set Scoring Components Weight

The screenshot displays the PERCEPTRON web application interface. The top navigation bar includes the PERCEPTRON logo and a user profile 'Hi Kanz'. The left sidebar lists various menu items: About PERCEPTRON, Getting Started, Protein Search Query, Search Results & History, Sample Data, Help & Manual, Video Tutorials, Report Bugs, Team, Contact, and Admin Panel. The main content area features a search form with fields for 'Protein Search Title *', 'Enter your Email Address', and a 'Select Database *' dropdown. Below the search form are four expandable sections: 'Set Experimental Parameters', 'Set De Novo Sequencing Parameters', 'Set Protein Modifications Parameters', and 'Set Scoring Components Weight *'. The 'Set Scoring Components Weight *' section is expanded and highlighted with a red box. It contains three horizontal sliders for adjusting weights: 'Intact Protein Mass Score Weightage (%)', 'Peptide Sequence Tags Score Weightage (%)', and 'Spectral Comparisons Score Weightage (%)'. A red circle with the letter 'A' points to the 'Spectral Comparisons' slider. A note below the sliders states: '* Set weight for at least one of the components.' At the bottom right of the section are 'Reset' and 'Submit' buttons.

Figure 8. PERCEPTRON - Overview of Scoring Component Weight

- A. Set the desired weights of Scoring Components towards computing scores by shifting the slider left or right accordingly

7.3 Window 3: User Search Results

User Search History			
Sr. #	Search Title	Creation Time	Status
1	Default Run	2020/12/19 14:05:46	Completed
2	MzXML_18122020	2020/12/18 15:02:56	Completed
3	MGF_18122020	2020/12/18 14:39:55	Completed
4	MzML_18122020	2020/12/18 14:39:24	Completed
5	MzXML_18122020	2020/12/18 14:38:49	Completed
6	CS1_Hela_batch	2020/12/18 14:19:33	Completed
7	CS1_Hela_140s_single	2020/12/18 14:19:09	Completed
8	CS2_ETD_17122020	2020/12/17 23:48:37	Completed
9	Ecoli_ETD_CS2_batch	2020/12/16 18:18:17	Completed
10	Hela_CS1_batchmode	2020/12/16 13:00:30	Completed
11	Hela_CS1_single_140s	2020/12/16 01:08:14	Completed
12	Hela_CS1_batch	2020/12/16 01:06:34	Completed

Figure 9. User Search History window showing search results

User can view search history in ‘Search Results and History’ tab with respective search titles. Latest search results are provided on the top. Click on the ‘Search Title’ to go to ‘Protein Spectral Matches’ which enlists the proteoforms identified from the search.

Protein Spectral Matches				
Sr. #	File Name	Protein ID	Protein Score	Molecular Weight
1	HELA_pk19_sw1_210sc_mono.txt	P62805	0.0696	11360.3815

Items per page: 10 0 of 0

* Click on the table row for corresponding candidate proteins list.

Contact Us (c) 2021 Biomedical Informatics Research Laboratory (BIRL)

Figure 10. Protein spectral matches window showing results for the protein search query

Click on the results under ‘File Name’ to go to ‘Summary Results’ window. User can download results as a zip folder by clicking ‘Results Download.’

Note: Results are downloadable and kept on the server for 2 days. There is no way to retrieve the data older than 48 hours.

7.4 Window 4: Summary and Detailed Results View

Summary Results View					
Filter					
Protein Rank	Protein ID	Molecular Weight	Terminal Modification	No. of PTM(s)	Protein Score
1	P62805	11360.3815	None	0	0.06962
2	Q9NPA8	11390.2726	NME	0	0.05063
3	Q9NPA8	11432.2832	NME_Acetylation	0	0.05063
4	A0A0B4J280	11417.6263	NME	0	0.03797
5	L0R6Q1	11125.8844	None	0	0.03797
6	Q75964	11290.2023	NME	0	0.03797
7	P0CE67	11240.7981	NME_Acetylation	0	0.03797
8	P22362	10984.5913	None	0	0.03797
9	Q5T7P2	10974.5096	None	0	0.03797
10	Q5T7P2	10843.4691	NME	0	0.03797

Figure 11. Summary Results window showing candidate proteins

Proteins found after the search are reported along with their Protein ID, Molecular Weight, Terminal Modifications, Number of Post-Translational Modifications (PTMs) and Protein Scores. User can click on any 'Protein ID' to see the details of the selected proteoform in 'Detailed Protein Hit View.'

Detailed Protein Hit View					
User Search Parameters					
Search Title:	Default Run	Protein Database:	Human	Mass Mode:	M(Neutral)
Protein Mass Tolerance:	500	Tune Mass:	False	Peptide Tolerance:	15
Fragmentation Type:	HCD	Special Ions:	bo,bstaryo,ystar	PST Filter:	True
PST Length:	3,6	Blind PTM:	False	Terminal Modifications:	None
Intact Protein Mass %:	0	PST %:	0	Spectral Comparison %:	100
FDR Cut-off %:	0.0				
General Results					
Protein Rank:	1	Protein ID:	P62805	Protein Name:	P62805
Protein Score:	0.0696202531645569	Molecular Weight:	11360.381505	# Matched Fragments:	5

PERCEPTRON: User Manual

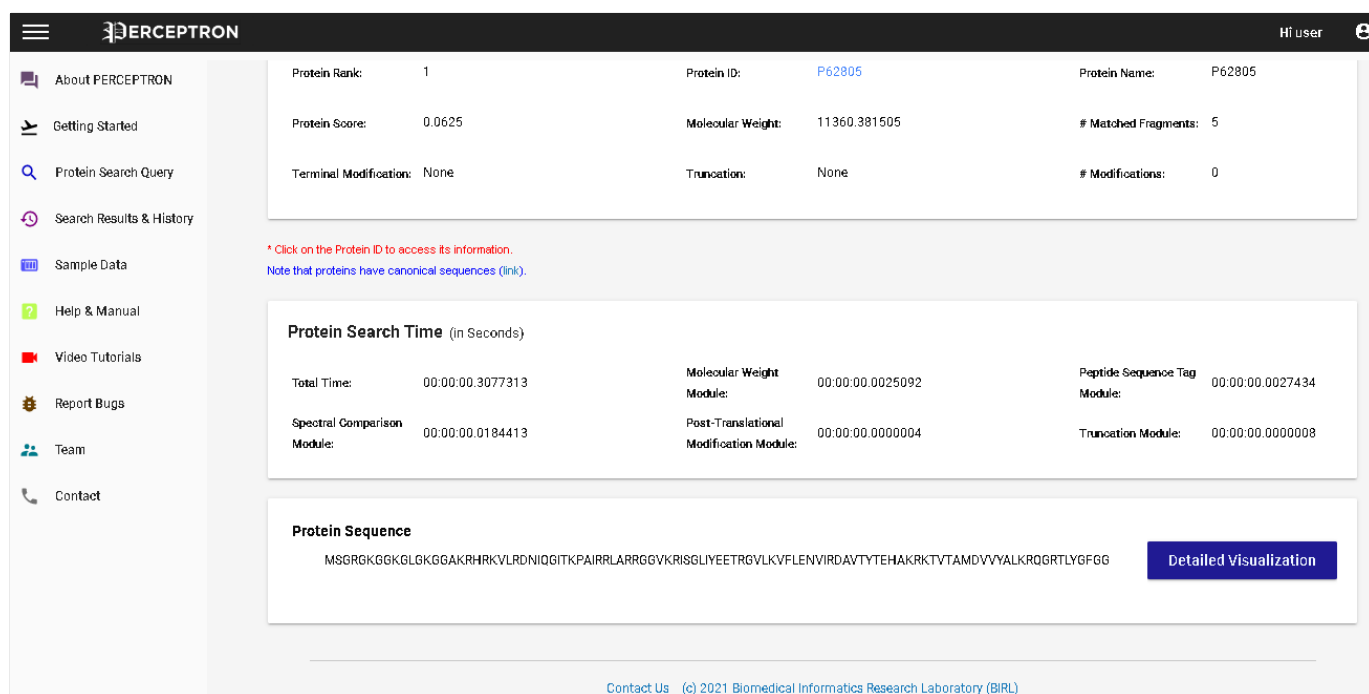


Figure 12. Detailed Results window showing search parameters and general results (top). Detailed Protein Hit View window showing Protein Search Time and Protein Sequence (bottom)

‘Detailed Protein Hit View’ window shows Search Parameters added while submitting protein search query. ‘General Results’ tab indicates an overview of search results. User can scroll down to view ‘Protein Search Time’ indicating runtimes for each algorithmic component. ‘Protein sequence’ shows the sequence of selected proteoform. PERCEPTRON offers visualization of search results. User can click ‘Detailed Visualization’ to visualize search results in detail.

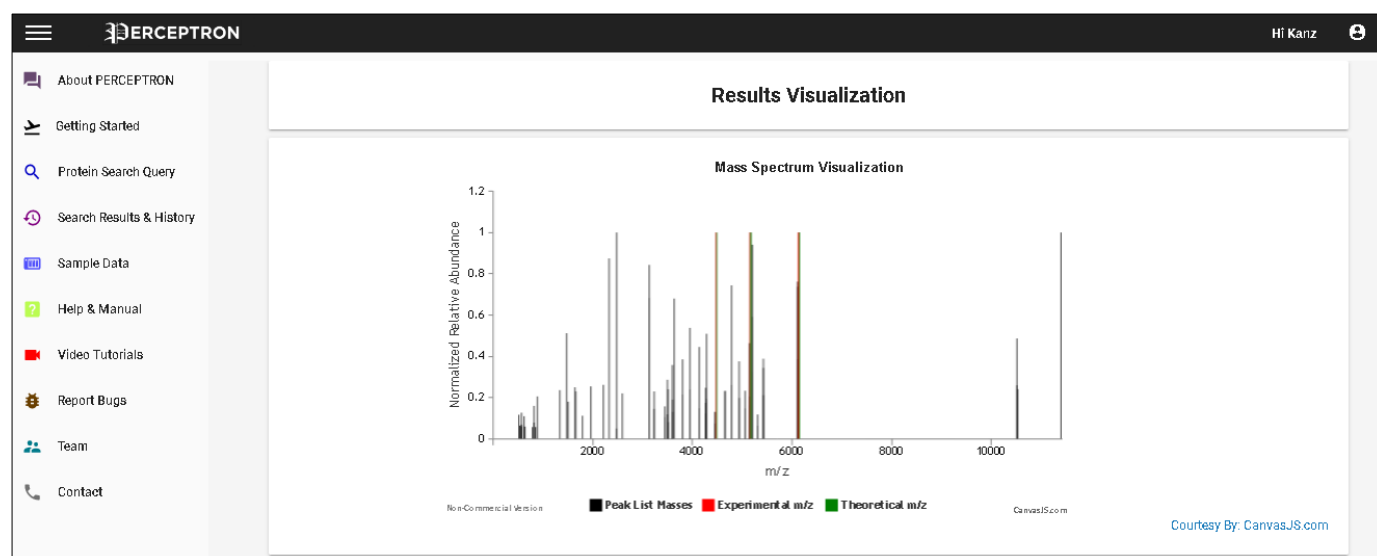


Figure 13. Detailed Results Visualization window with mass spectral visualization

PERCEPTRON: User Manual

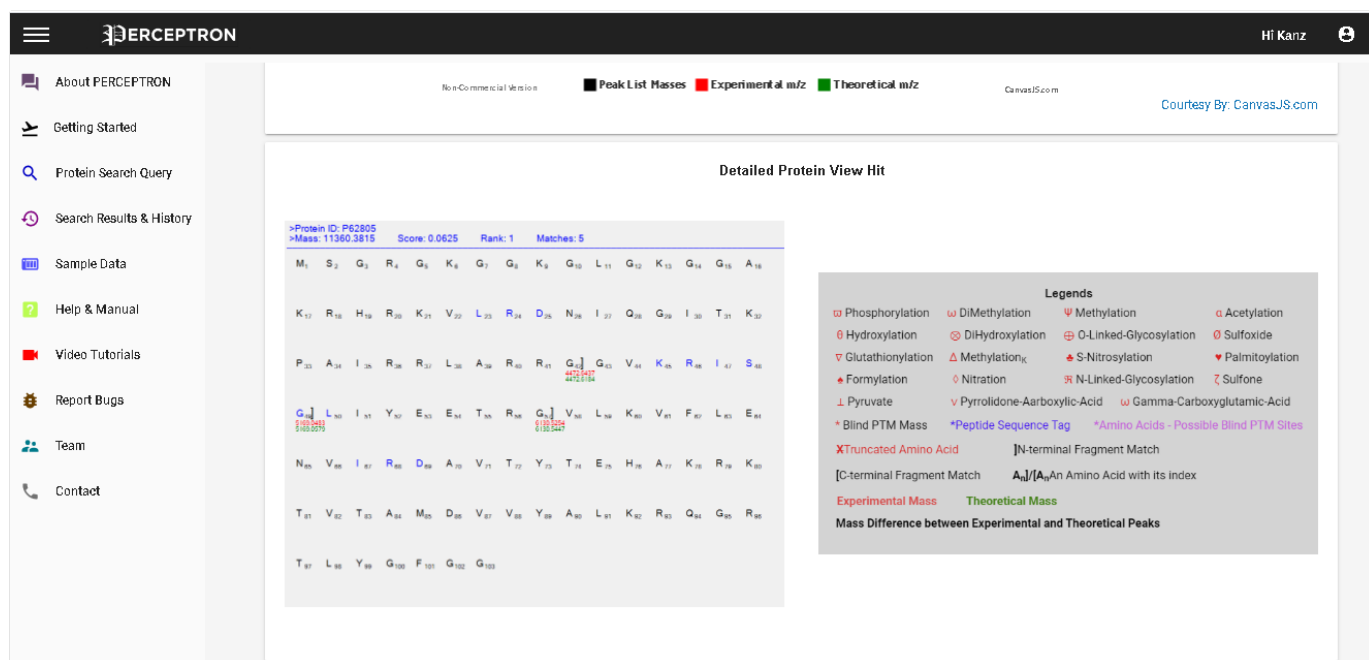


Figure 14. Detailed Results Visualization window with Detailed Protein Hit View (left) and Legends (right)

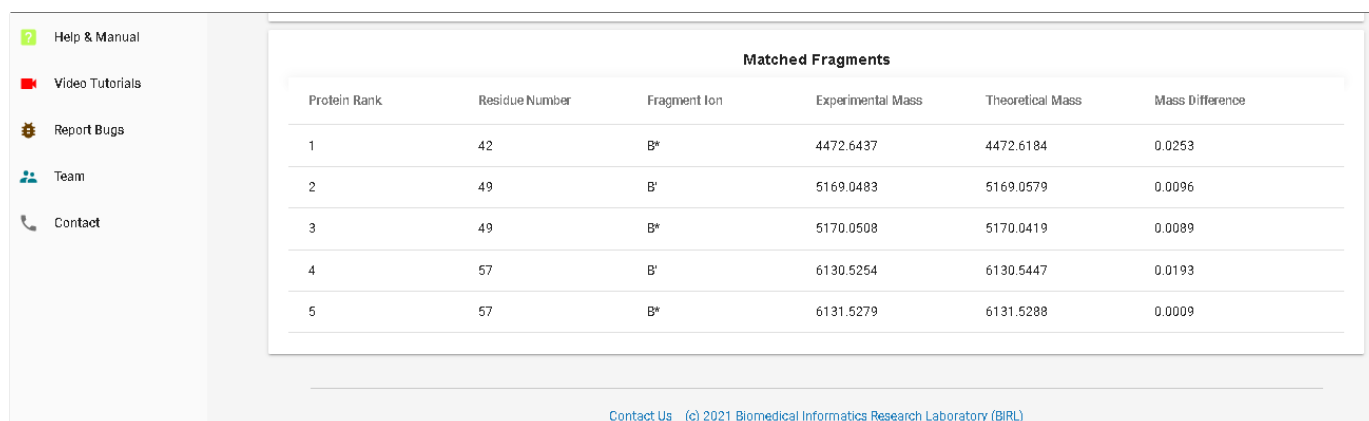


Figure 15. Detailed Results Visualization window with Matched Fragments

8. Interpreting Search Results

This chapter presents an overview of interpreting results obtained using PERCEPTRON.

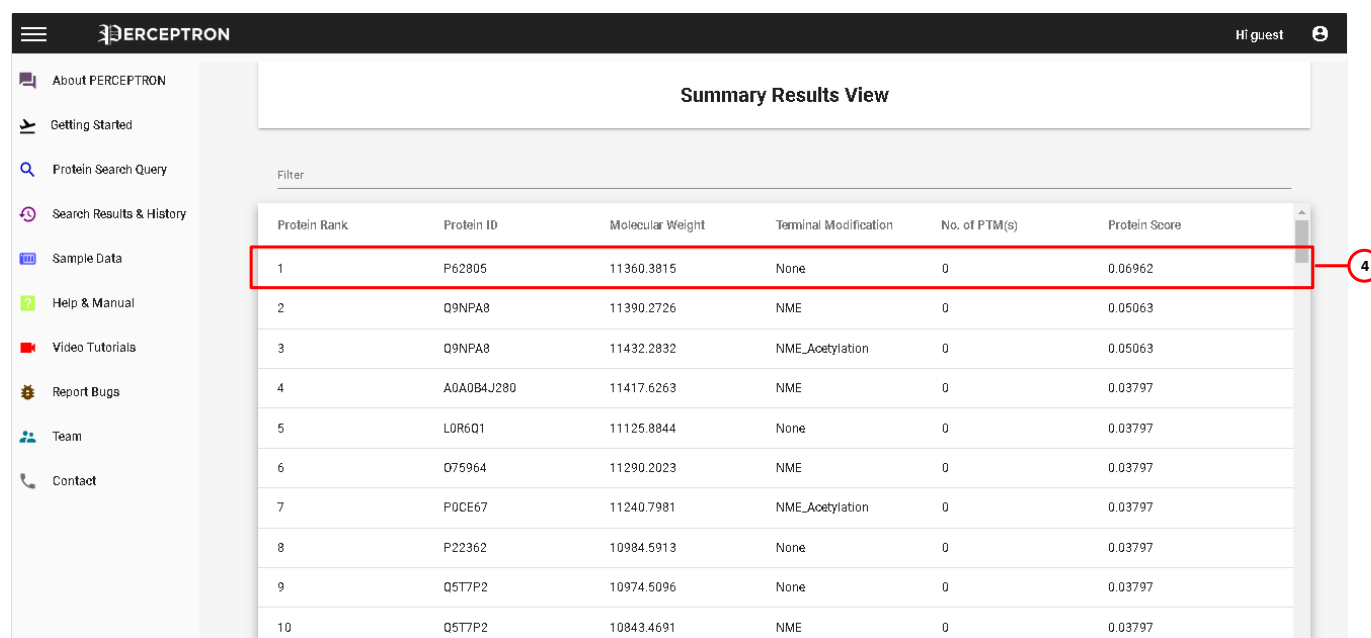
Sr. #	Search Title	Creation Time	Status
1	Default Run	2020/12/19 14:05:46	Completed
2	MzXML_18122020	2020/12/18 15:02:56	Completed
3	MGF_18122020	2020/12/18 14:39:55	Completed
4	MzXML_18122020	2020/12/18 14:39:24	Completed
5	MzXML_18122020	2020/12/18 14:38:49	Completed
6	CS1_Hela_batch	2020/12/18 14:19:33	Completed
7	CS1_Hela_140s_single	2020/12/18 14:19:09	Completed
8	CS2_ETD_17122020	2020/12/17 23:48:37	Completed
9	Ecoli_ETD_CS2_batch	2020/12/16 18:18:17	Completed
10	Hela_CS1_batchmode	2020/12/16 13:00:30	Completed
11	Hela_CS1_single_140so	2020/12/16 01:08:14	Completed
12	Hela_CS1_batch	2020/12/16 01:06:34	Completed

1. ‘User Search History’ in ‘Search Results and History’ tab shows search results with ‘Search Title.’ User can go to ‘Protein Spectral Matches’ by clicking the ‘Search Title.’

Sr. #	File Name	Protein ID	Protein Score	Molecular Weight
1	HELA_pk19_sw1_210sc_mono.txt	P62805	0.0696	11360.3815

2. In ‘Protein Spectral Matches,’ user can download results by clicking ‘Results Download’ option provided on top right. The results along with search parameters will be downloaded as a .zip folder.
3. File name indicates the data file used for performing protein search. User can view the top-ranked proteoform identified (Protein ID) along with its protein score and molecular weight. By clicking the ‘Protein ID,’ user can go to ‘Summary Results View.’

Note: ‘Detailed Visualization’ feature is not provided for proteoform search in batch mode. For batch mode, users can only download results as a .zip folder using the ‘Results Download’ option.

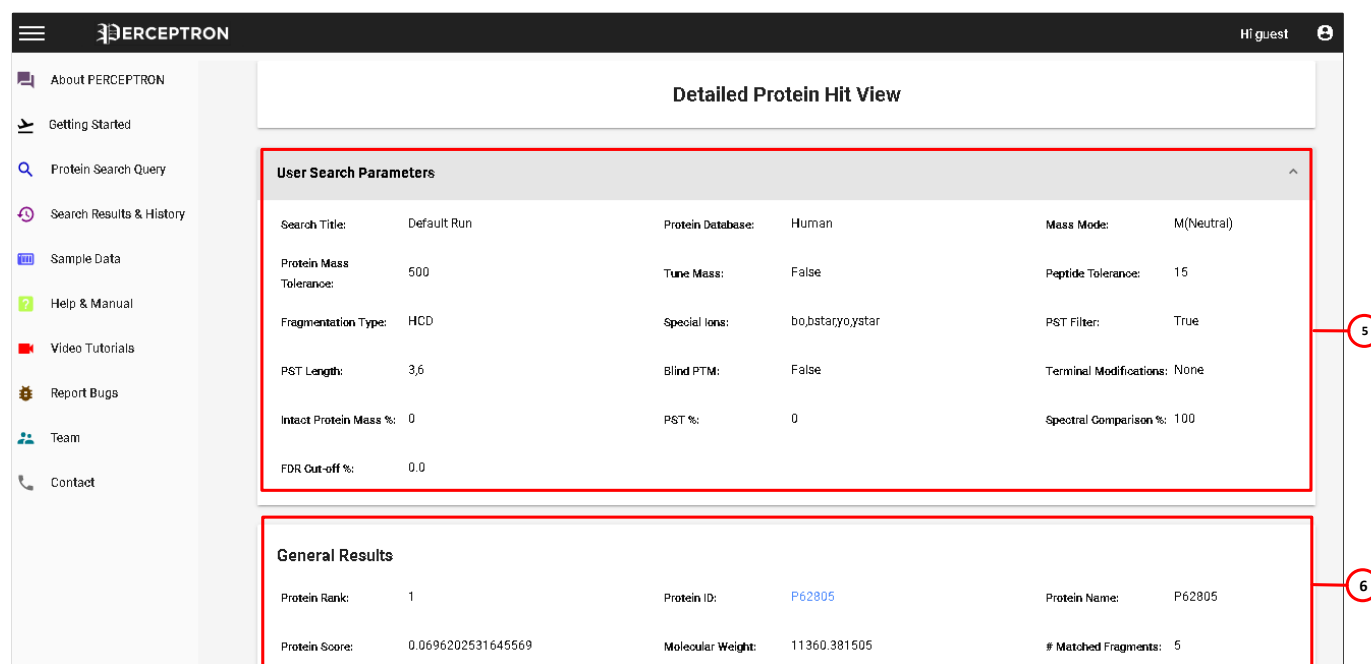


Summary Results View

Filter

Protein Rank	Protein ID	Molecular Weight	Terminal Modification	No. of PTM(s)	Protein Score
1	P62805	11360.3815	None	0	0.06962
2	Q9NFA8	11390.2726	NME	0	0.05063
3	Q9NFA8	11432.2832	NME_Acetylation	0	0.05063
4	A0A0B4J280	11417.6263	NME	0	0.03797
5	L0R6Q1	11125.8844	None	0	0.03797
6	Q75964	11290.2023	NME	0	0.03797
7	P0CE67	11240.7981	NME_Acetylation	0	0.03797
8	P22362	10984.5913	None	0	0.03797
9	Q5T7P2	10974.5096	None	0	0.03797
10	Q5T7P2	10843.4691	NME	0	0.03797

- ‘Summary Results View’ shows the list of proteoforms identified from the data file provided by the user. Clicking the ‘Protein ID’ leads user to ‘Detailed Protein Hit View.’



Detailed Protein Hit View

User Search Parameters

Search Title:	Default Run	Protein Database:	Human	Mass Mode:	M(Neutral)
Protein Mass Tolerance:	500	Tune Mass:	False	Peptide Tolerance:	15
Fragmentation Type:	HCD	Special Ions:	bo,bstaryo,ystar	PST Filter:	True
PST Length:	3,6	Blind PTM:	False	Terminal Modifications:	None
Intact Protein Mass %:	0	PST %:	0	Spectral Comparison %:	100
FDR Cut-off %:	0.0				

General Results

Protein Rank:	1	Protein ID:	P62805	Protein Name:	P62805
Protein Score:	0.0696202531645569	Molecular Weight:	11360.381505	# Matched Fragments:	5

- ‘Detailed Protein Hit View’ has ‘User Search Parameters’ on top.
- Below search parameters panel is ‘General Results’ which gives user details of the selected proteoform.

PERCEPTRON: User Manual

PERCEPTRON Hi user

Protein Rank: 1 **Protein ID:** [P62805](#) **Protein Name:** P62805

Protein Score: 0.0625 **Molecular Weight:** 11360.381505 **# Matched Fragments:** 5

Terminal Modification: None **Truncation:** None **# Modifications:** 0

* Click on the Protein ID to access its information.
Note that proteins have canonical sequences ([link](#)).

Protein Search Time (in Seconds)

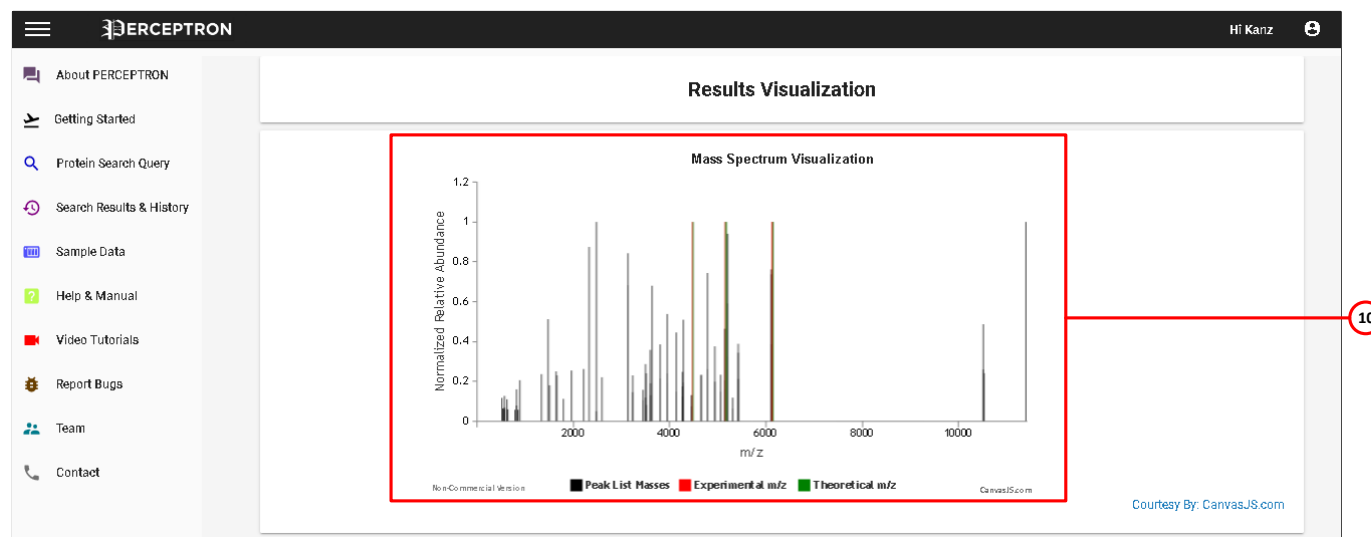
Total Time:	00:00:00.3077313	Molecular Weight Module:	00:00:00.0025092	Peptide Sequence Tag Module:	00:00:00.0027434
Spectral Comparison Module:	00:00:00.0184413	Post-Translational Modification Module:	00:00:00.0000004	Truncation Module:	00:00:00.0000008

Protein Sequence
MSGRGKGGKGLGKGGAKRHRKVLRLDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKVFLENVIRDVITYTEHAKRKTVTAMDVVYALKRQGRITLYGFGG

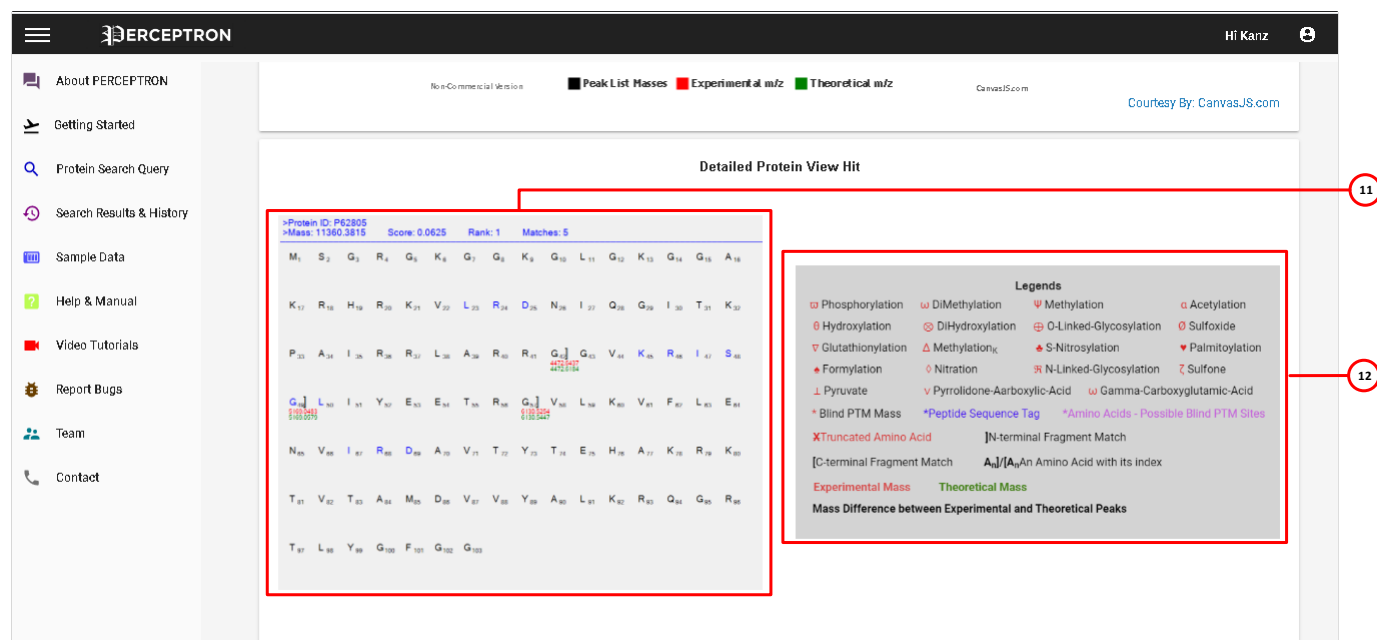
[Detailed Visualization](#)

Contact Us (c) 2021 Biomedical Informatics Research Laboratory (BIRL)

7. User can scroll down to see ‘Protein Search Time’ panel below ‘General Results’ indicating total time taken for processing protein query. Additionally, these runtimes have been split for highlighting time taken by each algorithmic component.
8. ‘Protein Sequence’ panel shows the sequence of the selected proteoform.
9. User can visualize search results in detail by clicking ‘Detailed Visualization’ button provided on bottom right.



10. ‘Mass Spectrum Visualization’ on ‘Results Visualization’ page shows Experimental Spectrum superimposed with Theoretical Spectrum highlighting peaks matched during spectral comparison. User can hover the cursor over to find peak list masses.



Detailed Protein View Hit

Protein ID: P62805
Mass: 11360.3815
Score: 0.0625
Rank: 1
Matches: 5

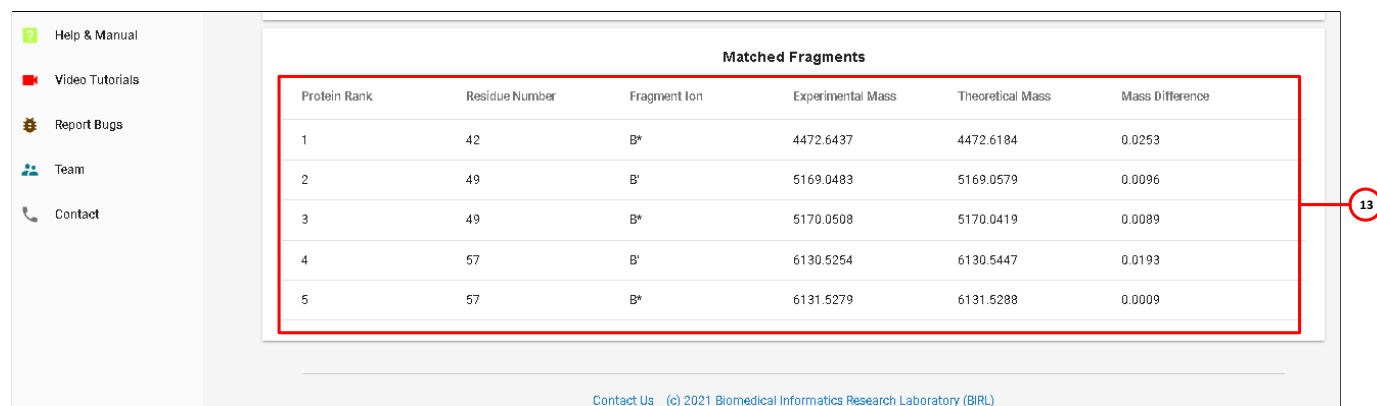
Sequence: M₁ S₂ G₃ R₄ G₅ K₆ G₇ G₈ K₉ G₁₀ L₁₁ G₁₂ K₁₃ G₁₄ G₁₅ A₁₆
K₁₇ R₁₈ H₁₉ R₂₀ K₂₁ V₂₂ L₂₃ R₂₄ D₂₅ N₂₆ I₂₇ Q₂₈ G₂₉ I₃₀ T₃₁ K₃₂
P₃₃ A₃₄ I₃₅ R₃₆ R₃₇ L₃₈ A₃₉ R₄₀ R₄₁ G₄₂ G₄₃ V₄₄ K₄₅ R₄₆ I₄₇ S₄₈
G₄₉ L₅₀ I₅₁ Y₅₂ E₅₃ E₅₄ T₅₅ R₅₆ G₅₇ V₅₈ L₅₉ K₆₀ V₆₁ F₆₂ L₆₃ E₆₄
N₆₅ V₆₆ I₆₇ R₆₈ D₆₉ A₇₀ V₇₁ T₇₂ Y₇₃ T₇₄ E₇₅ H₇₆ A₇₇ K₇₈ R₇₉ K₈₀
T₈₁ V₈₂ T₈₃ A₈₄ M₈₅ D₈₆ V₈₇ V₈₈ Y₈₉ A₉₀ L₉₁ K₉₂ R₉₃ Q₉₄ G₉₅ R₉₆
T₉₇ L₉₈ Y₉₉ G₁₀₀ F₁₀₁ G₁₀₂ G₁₀₃

Legends

- Phosphorylation
- Hydroxylation
- Glutathionylation
- Formylation
- Pyruvate
- Blind PTM Mass
- Truncated Amino Acid
- DiMethylation
- DiHydroxylation
- Methylation_{ox}
- Nitration
- Pyrrolidone-Aarboxylic-Acid
- Peptide Sequence Tag
- N-terminal Fragment Match
- C-terminal Fragment Match
- Methylation
- O-Linked-Glycosylation
- S-Nitrosylation
- Gamma-Carboxyglutamic-Acid
- Amino Acids - Possible Blind PTM Sites
- Acetylation
- Sulfoxide
- Palmitoylation
- Sulfone

Experimental Mass Theoretical Mass
Mass Difference between Experimental and Theoretical Peaks

11. User can scroll down to find detailed results. Left panel in 'Detailed Protein Hit View' shows complete information of the selected proteoform. These details include:
 - i. Sequence of the proteoform containing one-letter symbols for amino acid residues along with the number at bottom right indication the position of the respective residue within the proteoform sequence
 - ii. Molecular weight of the proteoform selected from the list of identified proteoforms
 - iii. Rank of the respective proteoform in the list of proteoforms identified
 - iv. Proteoform score computed by PERCEPTRON
 - v. Matched fragments along with experimental (in red) and theoretical (in green) fragment masses
 - vi. Post-translational modifications as well as other modifications (e. g. truncation)
12. 'Legends' is a key to help user understand 'Detailed Protein Hit View.'



Matched Fragments

Protein Rank	Residue Number	Fragment Ion	Experimental Mass	Theoretical Mass	Mass Difference
1	42	B*	4472.6437	4472.6184	0.0253
2	49	B'	5169.0483	5169.0579	0.0096
3	49	B*	5170.0508	5170.0419	0.0089
4	57	B'	6130.5254	6130.5447	0.0193
5	57	B*	6131.5279	6131.5288	0.0009

Contact Us (c) 2021 Biomedical Informatics Research Laboratory (BIRL)

13. 'Matched Fragments' shows a list of theoretical fragments generated by PERCEPTRON's algorithmic pipeline matched with experimental fragments which are originating from data file provided as an input during protein search.
 - (i) 'Protein Rank' indicates the rank of the protein within the list of identified proteoforms
 - (ii) 'Residue Number' is the location of residue within proteoform sequence at which the match for 'fragment ion' was obtained
 - (iii) 'Fragment Ion' indicates the type of fragment ion matched
 - (iv) 'Experimental Mass' is the mass of experimental fragment from data file
 - (v) 'Theoretical Mass' is the mass of *in silico* fragment generated by PERCEPTRON
 - (vi) 'Mass Difference' is the difference between the masses of experimental and theoretical fragments

9. Proteoform Search Using Formatted Data Files

9.1 File Formats

PERCEPTRON provides support for plain text files (data in columns containing mass to charge ratios (m/z) and relative intensities), eXtensible Markup Language (XML) files with m/z and relative abundances (mzXML)¹, Mass Spectrometry Markup Language (mzML)^{2,3} and Mascot Generic Format (MGF)⁴ data formats in both single and batch file processing modes.

9.1.1 Raw to mzXML File Format Conversion

User can convert raw data files to mzXML file format by using MS-Convert⁵.

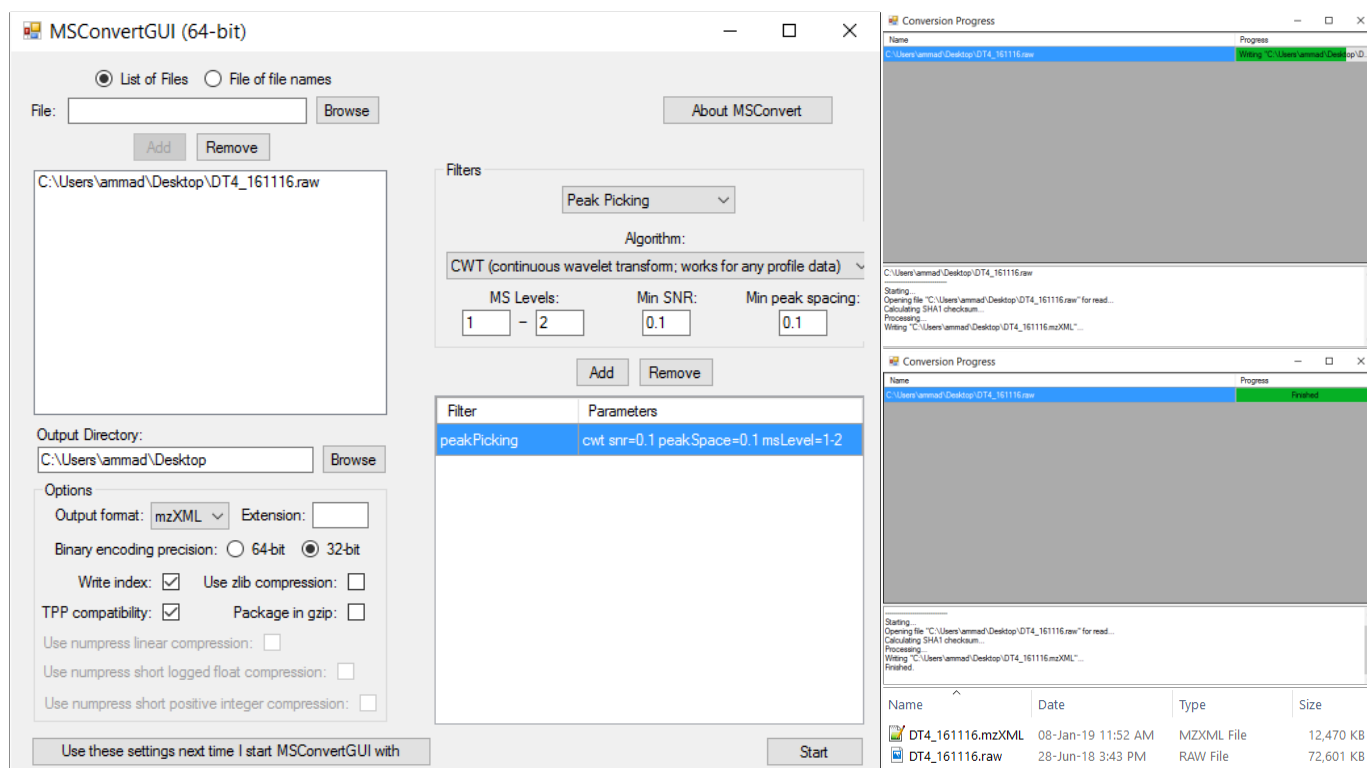


Figure 16. Conversion of raw to mzXML

9.1.2 Raw to mzML File Format Conversion

Raw data files can be converted to mzML file format by using MS-Convert⁵.

PERCEPTRON: User Manual

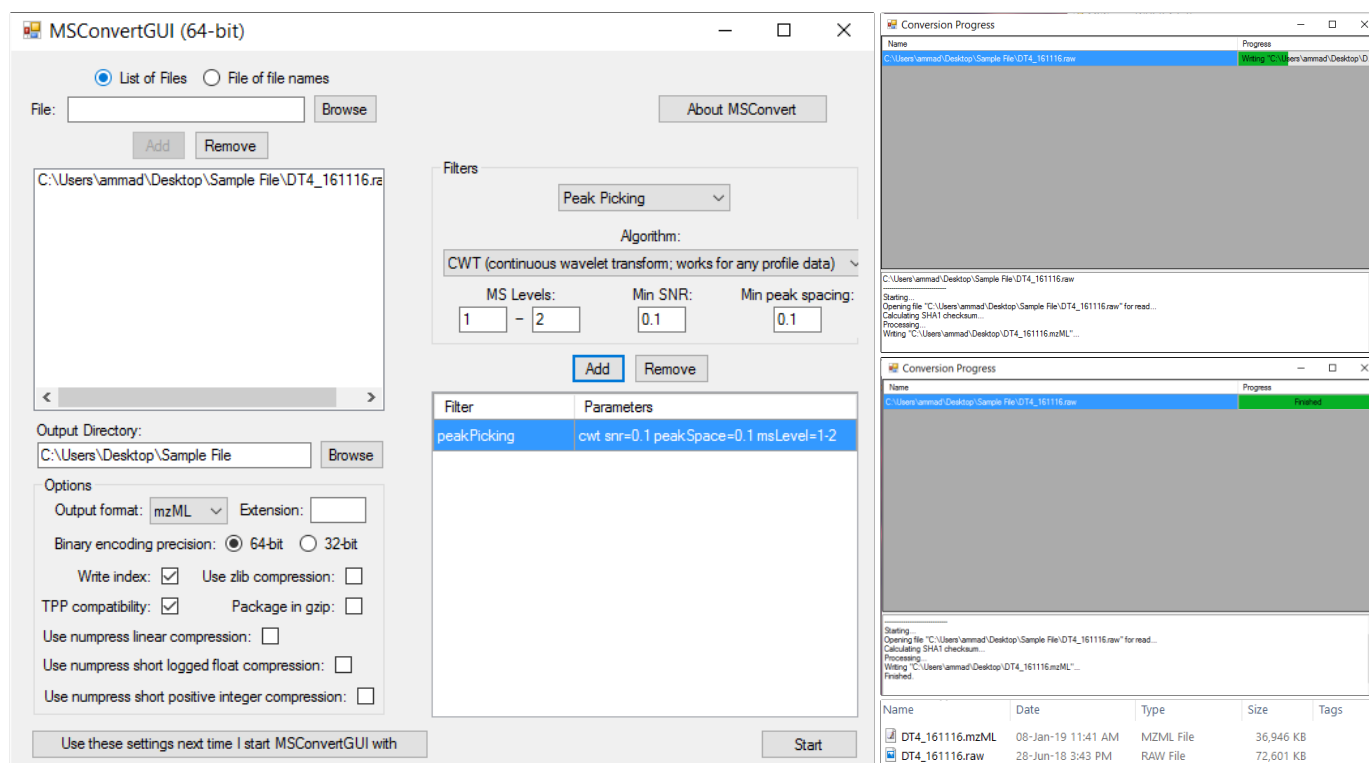


Figure 17. Conversion of raw to mzML

9.1.3 MzXML to MGF File Format Conversion

User can convert mzXML files to MGF using MS-Decov⁶.

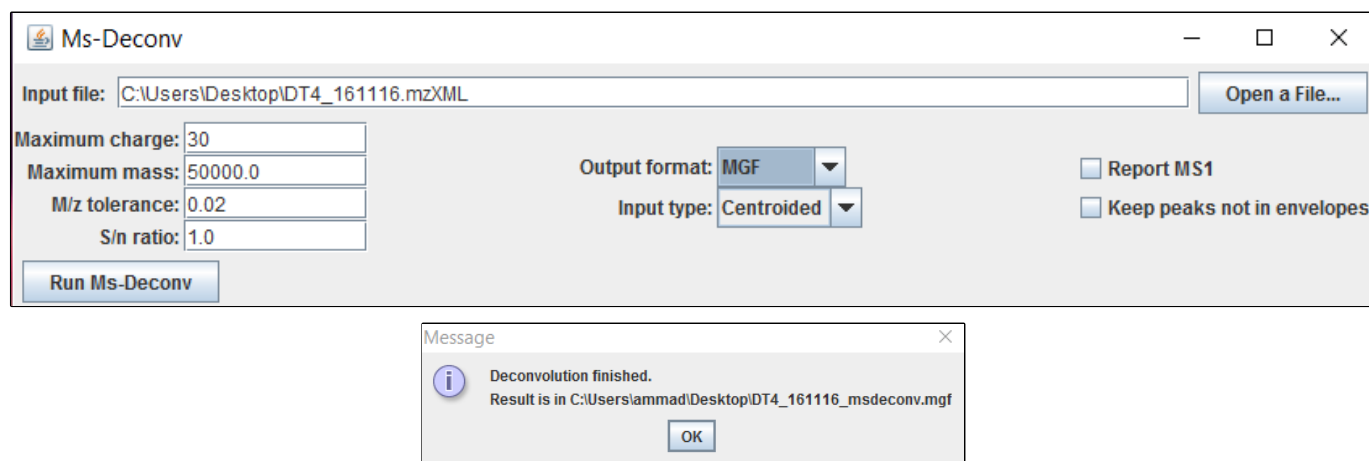


Figure 18. Conversion of mzXML to MGF

9.1.4 MGF to Flat Text File Format Conversion

PERCEPTRON converts MGF files to flat text (peak list) using built-in custom file reader.

9.2 Parameters

PERCEPTRON can employ from the following set of search parameters including: (i) Default Parameters, and (ii) Selected parameters.

How to load default Parameters?

To submit the job using default parameters, select 'Load Default Parameters' option in front of Basic Parameters.

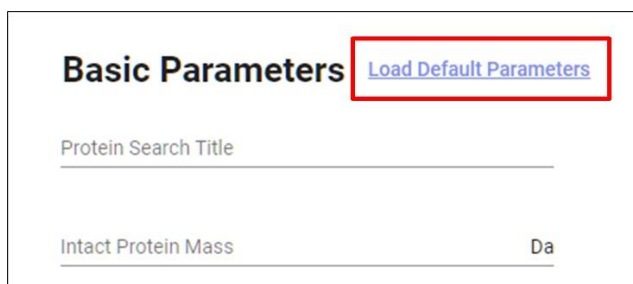
A screenshot of a web form titled "Basic Parameters". To the right of the title is a blue hyperlink "Load Default Parameters" which is enclosed in a red rectangular box. Below the title, there is a text input field labeled "Protein Search Title". At the bottom, there is a text input field labeled "Intact Protein Mass" with a unit label "Da" to its right.

Figure 19. Load Default Parameters

9.3 Databases

SwissProt database is included in PERCEPTRON by default. User can take any protein sequence from other databases (such as Uniprot) in .fasta format.

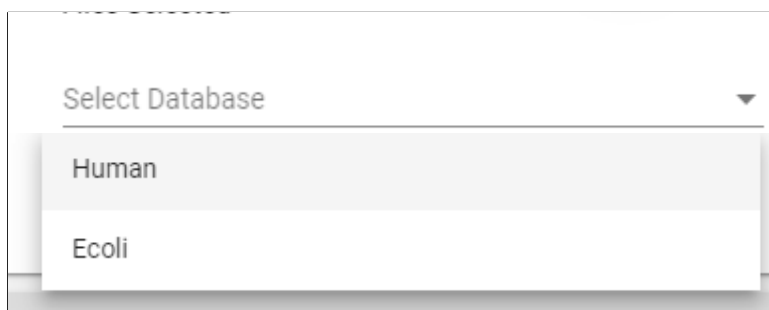
A screenshot of a dropdown menu titled "Select Database". The menu is open, showing two options: "Human" and "Ecoli". The "Human" option is highlighted with a light gray background.

Figure 20. Selecting Protein Database

9.4 Modes

The search modes are auto-selected based on the number of files that the user inputs. If one file is given as input, PERCEPTRON runs single search mode whereas if multiple files are given as input, the mode switches to batch mode.

- (i) Single Search Mode
- (ii) Batch Mode

Batch mode takes more processing time as it deals with larger data. The experimental spectra, search parameters and results are automatically stored in the project directory for further processing and visualization.

10. Deploying PERCEPTRON

10.1 Installing Visual Studio Code (VS Code)

Users can download VS Code setup using the following link:

<https://code.visualstudio.com/download>

Click **'64 bit'** in front of **User Installer** to download the setup. Run the downloaded .exe file.

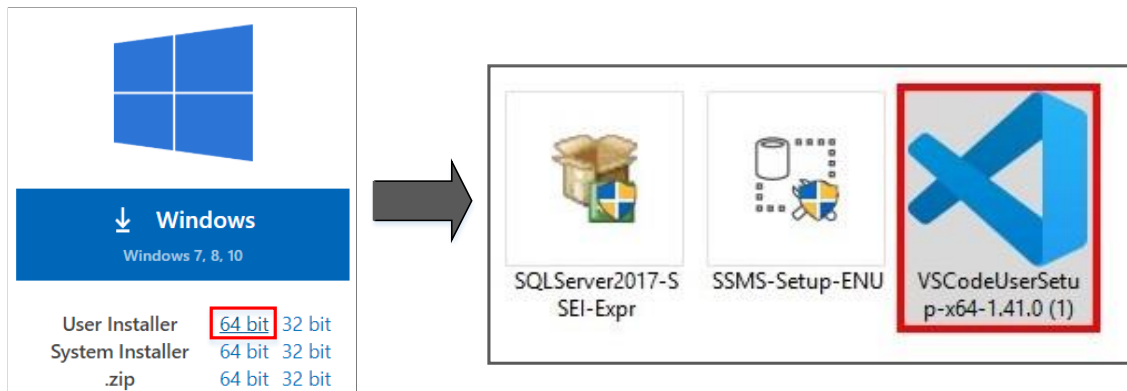


Figure 21. Downloading Visual Studio Code (VS Code) setup

Accept license terms and choose a destination folder to download Visual Studio Code setup.

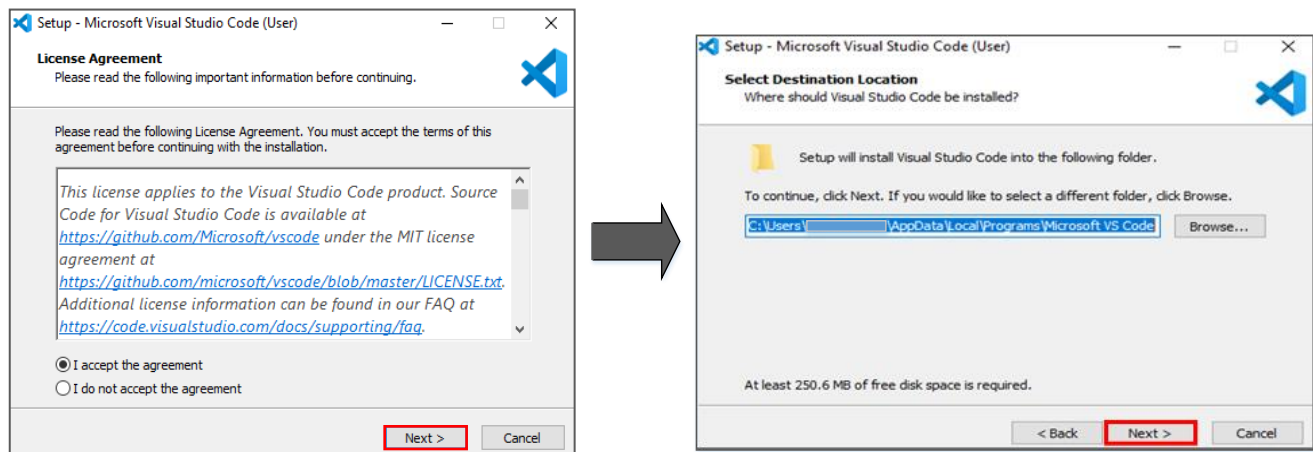
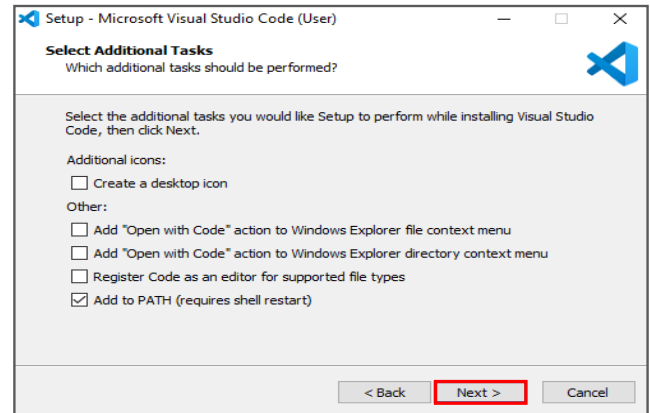
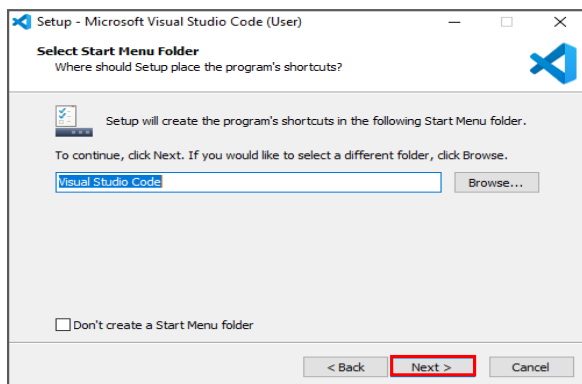


Figure 22. Choosing destination folder to download VS Code setup

PERCEPTRON: User Manual

Click **'Next'** then select **'Add to PATH.'** Click **'Next'** to proceed.



Click **'Install'** to begin installation. Click **'Finish'** to complete installation.

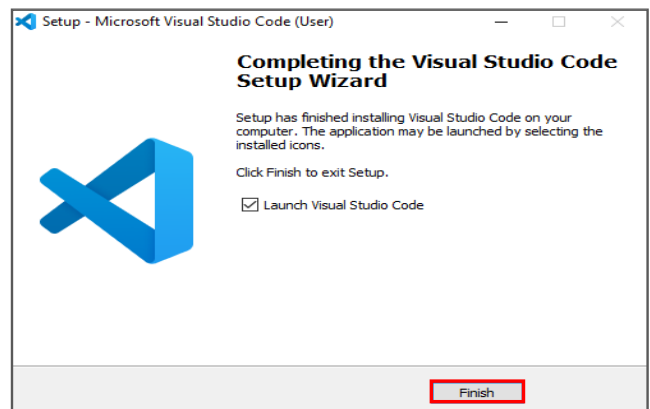
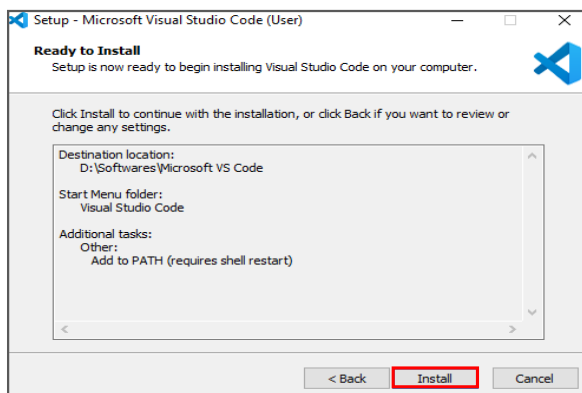


Figure 23. Steps to install Visual Studio Code setup

10.2 Installing SQL Server Management Studio (SSMS)

Users can download SSMS setup (version 17.6 or version 17.9.1) using the following link:

<https://docs.microsoft.com/en-us/sql/ssms/release-notes-ssms?view=sql-server-ver15#additional-downloads>

Click **‘Download’** to download the setup. Run the downloaded .exe file.

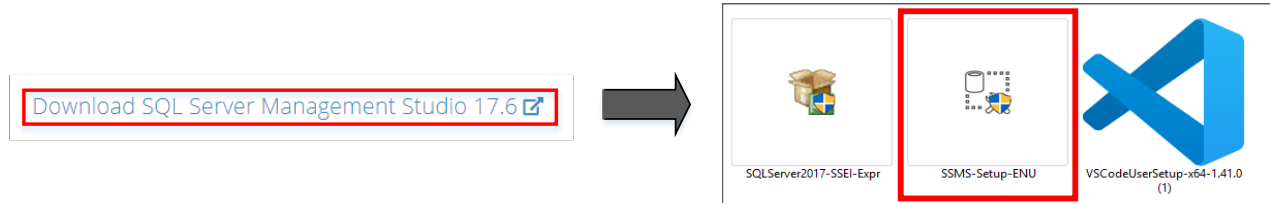


Figure 24. Downloading SQL Server Management Studio (SSMS) setup

Click **‘Install’** to begin installation. Click **‘Close’** to complete installation.

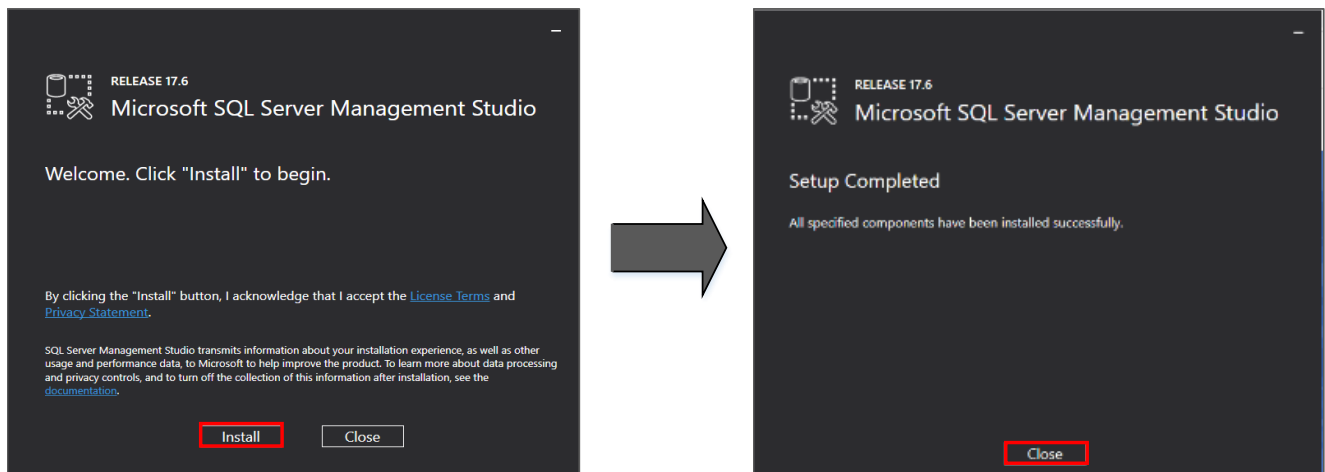


Figure 25. Steps to install SQL Server Management Studio (SSMS)

10.3 Installing Node.js.8.11.1

Users can download the setup using the following link:

<http://www.npackd.org/p/org.nodejs.NodeJS64/8.11.1>

Click on **'Download Node.js 64 bit 8.11.1'** to download node.js setup. Click on **'Run'** to proceed with downloading the .exe file.

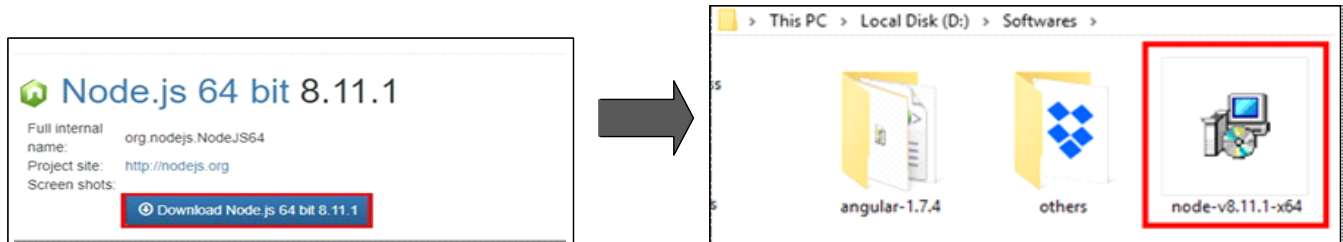
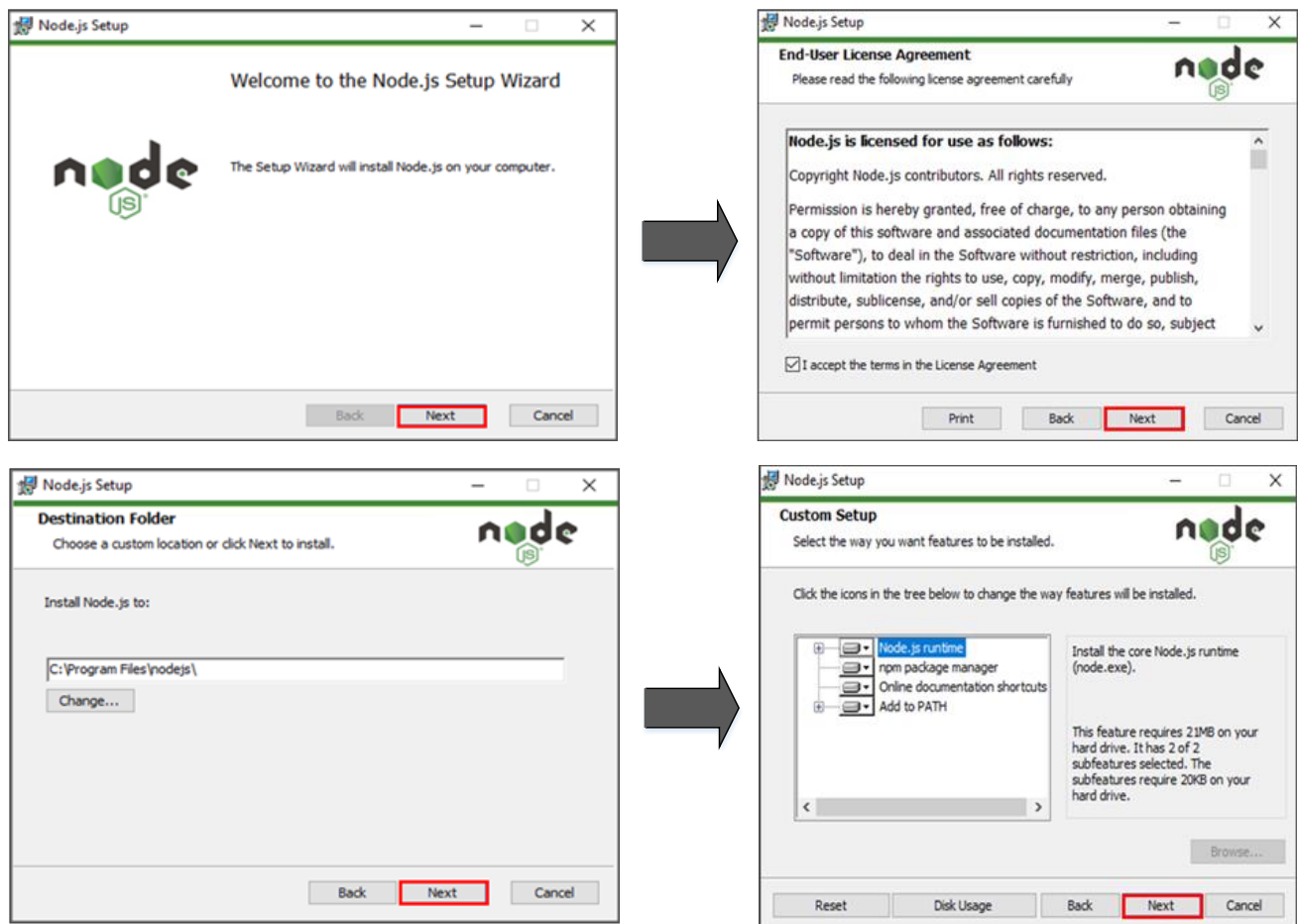
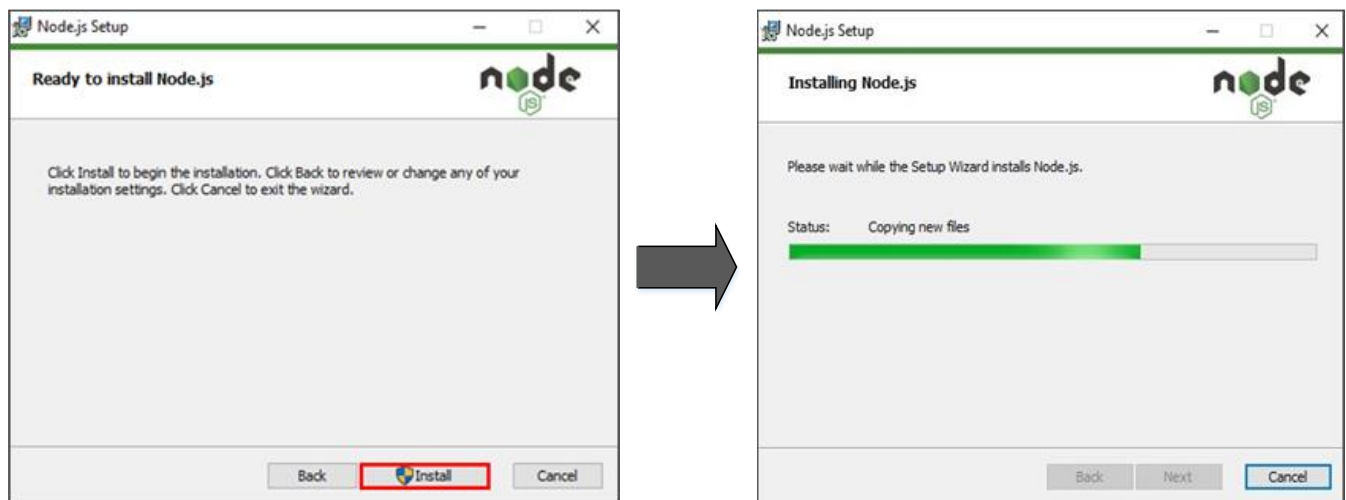


Figure 26. Downloading Node.js (version 8.11.1) setup

Click **'Next'** and accept license terms to proceed.



Click **'Install'** to begin installation.



Click **'Finish'** to complete installation.



Figure 27. Steps to install Node.js setup

10.4 Installing SQL 2017

Users can download Structured Query Language (SQL) 2017 setup using the following link:

<https://www.microsoft.com/en-pk/download/details.aspx?id=55994>

Select a language and click **‘Download.’**

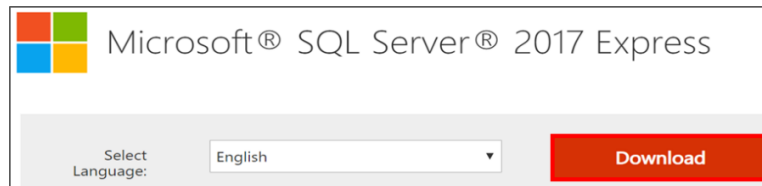
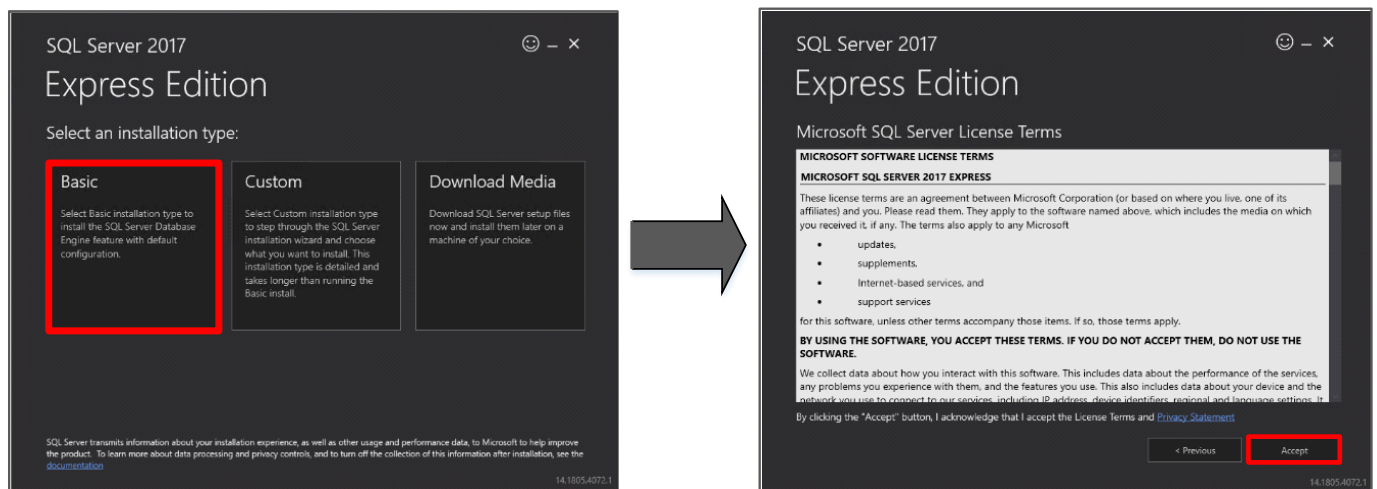


Figure 28. Downloading SQL 2017 setup

Select **‘Basic’** and then click **‘Accept’** to accept license terms.



Allow installation to proceed. Click **‘Close’** to complete installation.

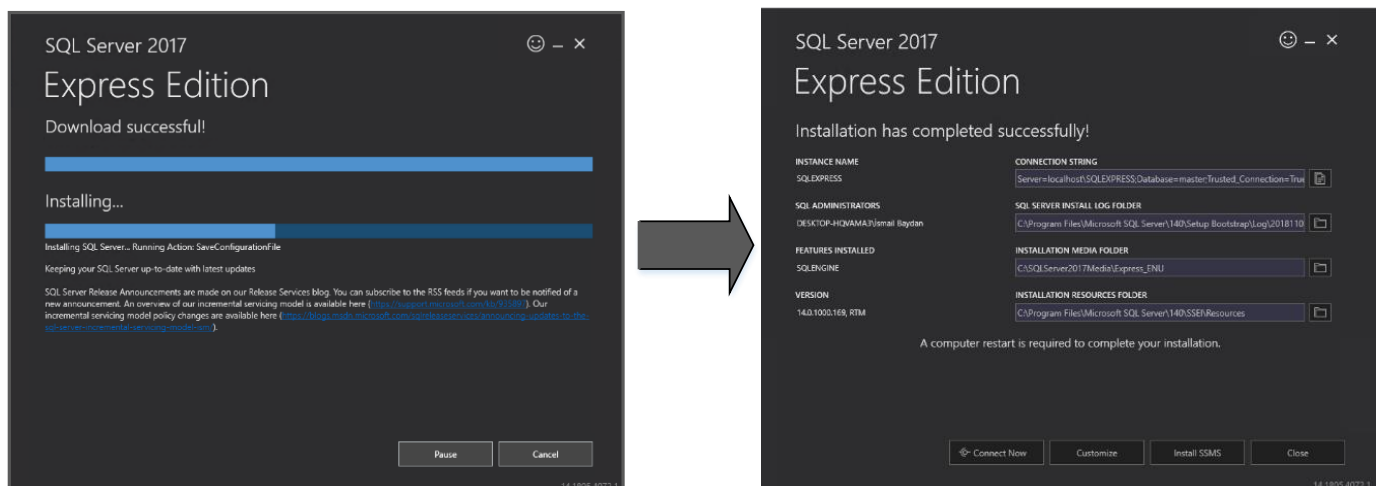


Figure 29. Steps to install SQL 2017 setup

10.5 Installing Angular 1.7.4

Users can download Angular 1.7.4 as a zip file using the following link:

<https://code.angularjs.org/1.7.4/>

Open the link and select the .zip file for installing Angular 1.7.4. Extract the contents in the file.

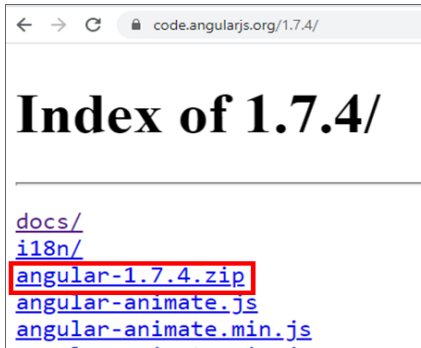
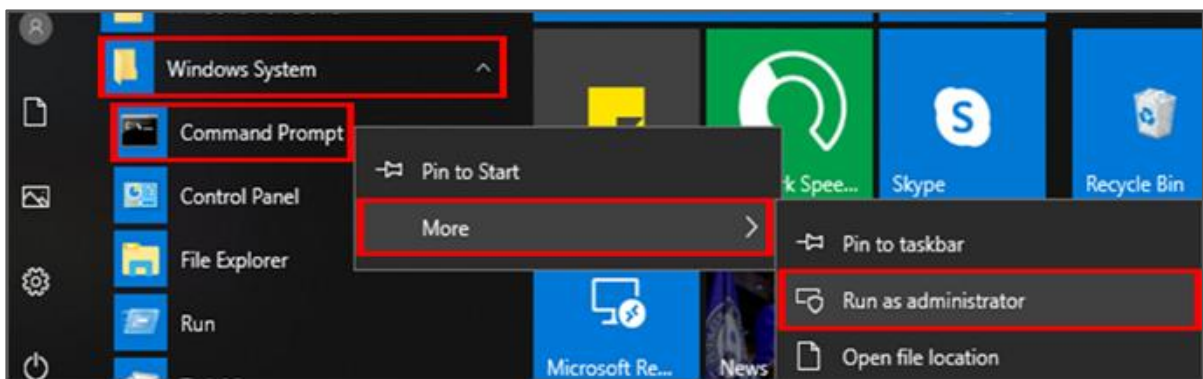


Figure 30. Downloading Angular 1.7.4 setup

Next, open Command Window by selecting **‘Windows System.’** Right click on **‘Command Prompt,’** then select **‘More’** and click **‘Run as administrator.’**



Type **'npm uninstall -g @angular/cli'** in Command Window and press **ENTER**. Next, type **'npm cache verify'** and press **ENTER** again.

```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.17134.407]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>npm uninstall -g @angular/cli
up to date in 0.031s

C:\WINDOWS\system32>
```



```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.17134.407]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>npm uninstall -g @angular/cli
up to date in 0.031s

C:\WINDOWS\system32>npm cache verify
Cache verified and compressed (~\AppData\Roaming\npm-cache\_cacache):
Content verified: 2162 (80052308 bytes)
Index entries: 3387
Finished in 6.055s

C:\WINDOWS\system32>
```

Type **'npm install -g @angular/cli@1.7.4'** and press **ENTER** to install. Type **'ng -v'** and press **ENTER** to ensure successful installation.

PERCEPTRON: User Manual

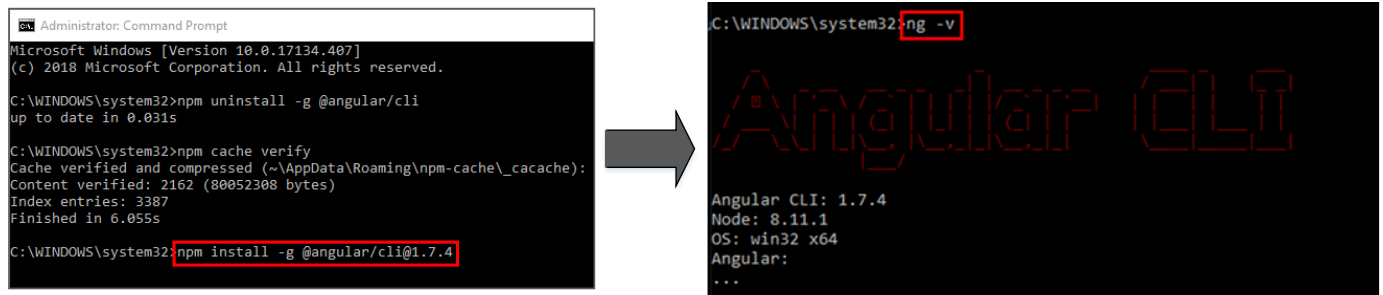


Figure 31. Steps to install Angular 1.7.4

10.6 Importing files into SQL

First, open SQL Server Management Studio.

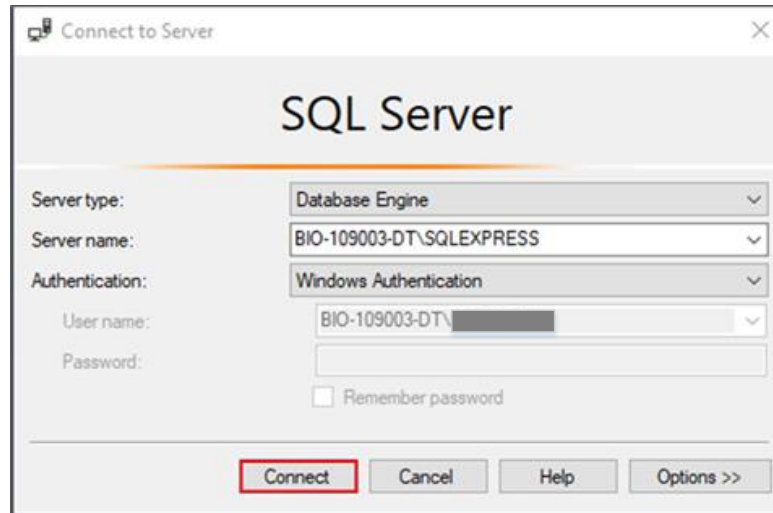



Figure 32. SQL Server Management Studio Pop Up

1. Click **'Connect'**
2. Right click **'Databases'** on left plane (Object Explorer)
3. Click **'Restore Databases'**
4. Choose **'Device'**, click , and add your **.bak** file
5. Click **'OK'**, and then **'OK'** again

11. Building PERCEPTRON

11.1 PERCEPTRON Application Programming Interface (API)

Open 'PerceptronAPI' from the folder **PERCEPTRON_CPU** in visual studio. In the Solution Explorer, click on 'PerceptronAPI'. Go to 'App_Start' and open 'WebApiConfig.cs'.



Figure 33. Building PERCEPTRON API. (a) Opening solution PerceptronAPI.sln; (b) Opening WebApiConfig.cs from Solution Explorer of Visual Studio

Next, search 'config.filters' using **CTRL+F**. Comment out the statement that includes: **config.Filters.Add(new RequireHttpsAttribute())** by typing '//' before the start of the statement.

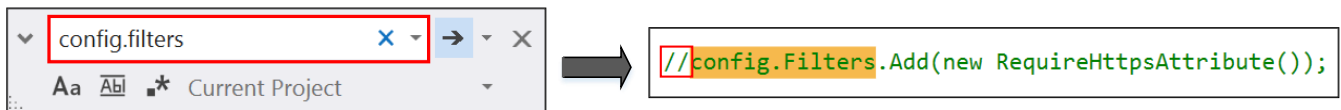


Figure 34. Searching for a statement in PERCEPTRON API solution

Open SQL Server Management Studio. Copy the given server name from SQL Server Pop Up.

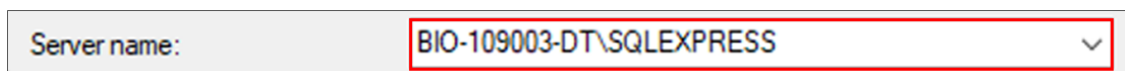


Figure 35. Copying Server Name from SQL Server Management Studio

PERCEPTRON: User Manual

In Visual Studio, open '**Web.config**' from the Solution Explorer. Press **CTRL+F** to search item. Type server name, e.g. '**CHIRAGH-II**' and press **ENTER**. Replace this server name with the name copied in Figure 35. Copying Server Name from SQL Server Management Studio, e.g. **BIO-109003-DT\SQLEXPRESS**.

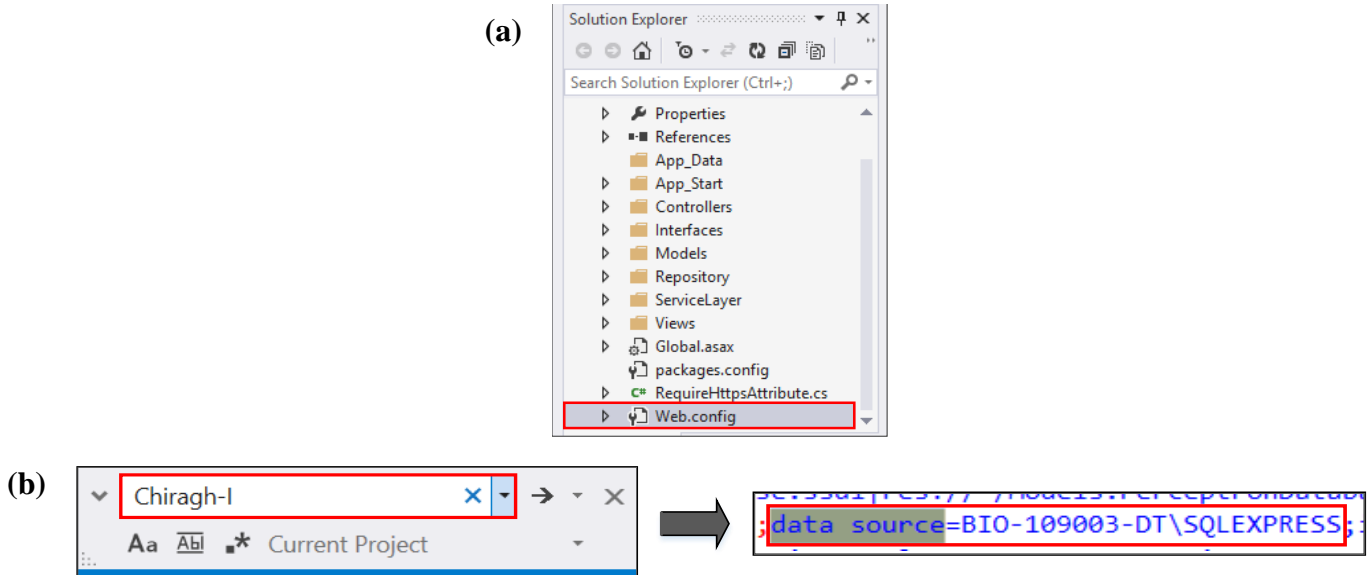


Figure 36. (a) Opening Web.config from Solution Explorer of Visual Studio (b) Changing data source (Server Name) in Web.config

Press **ENTER** again to change the name in *SqlDatabase.cs* (in Repository Folder) as well. Replace the given name for all occurrences of '**Server**' in *SqlDatabase.cs* with the server name from SQL Pop Up. In each case, add '@' right before '**Server**.'

```
using (var db = new PerceptronDatabaseEntities())
{
    var sqlConnection1 =
        new SqlConnection(
            @"Server=BIO-109003-DT\SQLEXPRESS; Database=
            PerceptronDatabase; Integrated Security=SSPI;")
}
```

Figure 37. Changing server name in *SqlDatabase.cs*

In dropdown menu of **BUILD**, select '**Build Solution.**' Next, select '**Publish PerceptronAPI**' under the same menu.

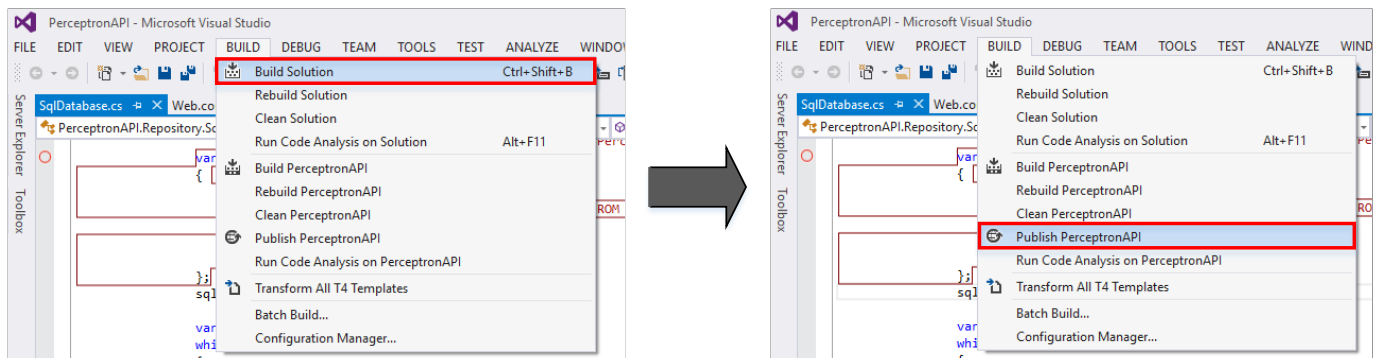


Figure 38. Building PERCEPTRON API

PERCEPTRON: User Manual

In Profile, select '**PERCEPTRON**'. In Connection, go to Target Location and ensure the following pathway: '**C:\inetpub\wwwroot\PerceptronAPI**'. In Settings, set Configuration to '**Debug**'.

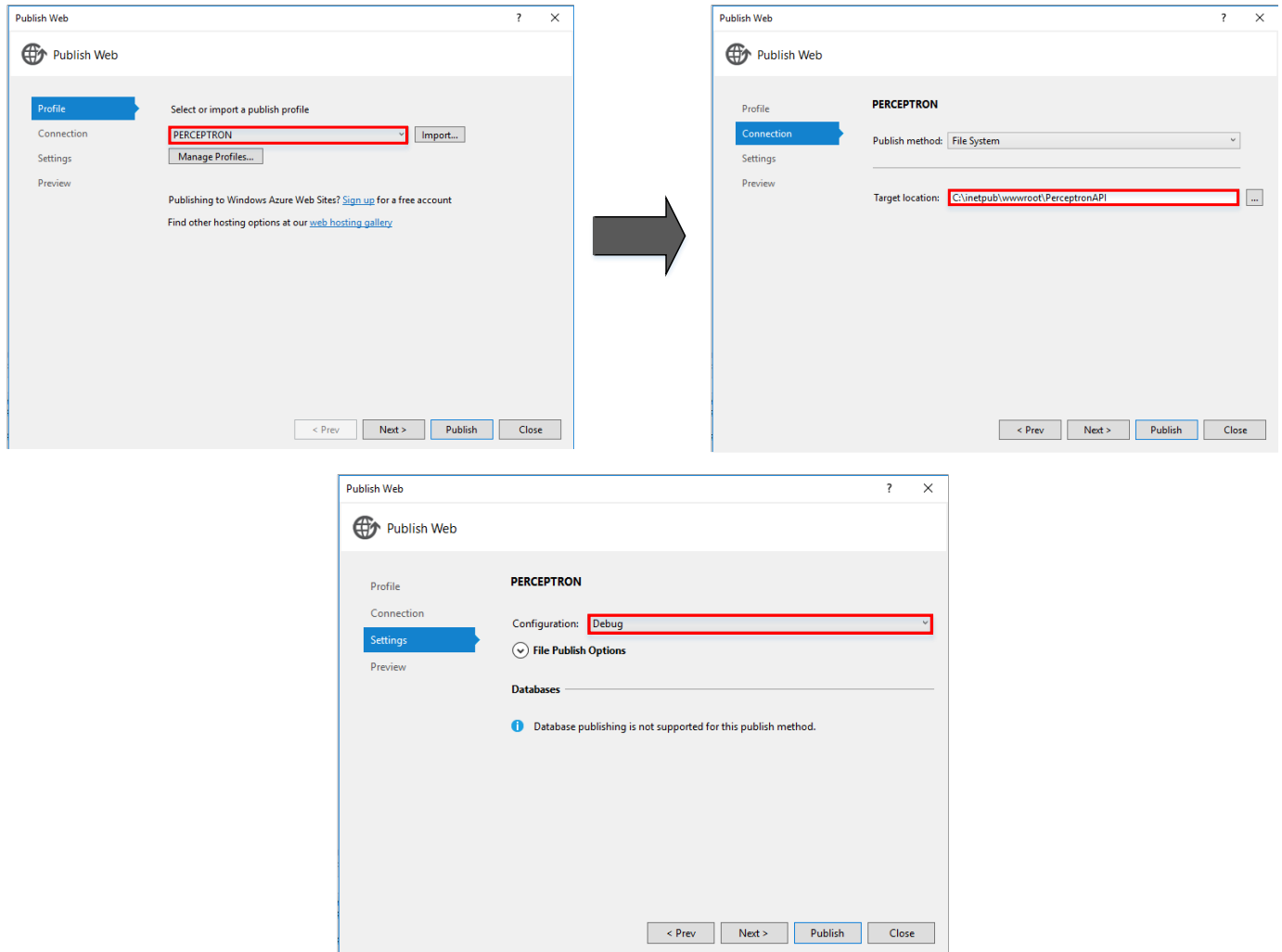


Figure 39. Steps to build PERCEPTRON API

Open the following pathway and make a folder in it by the name '**App_Data**'.

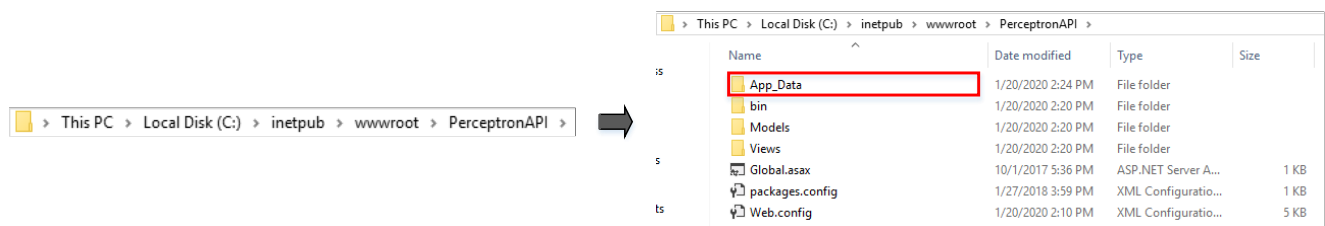


Figure 40. Publishing Perceptron API

PERCEPTRON: User Manual

From the dropdown menu of **BUILD**, select '**Build Solution.**' Run on '**Google Chrome**' and save the URL obtained.

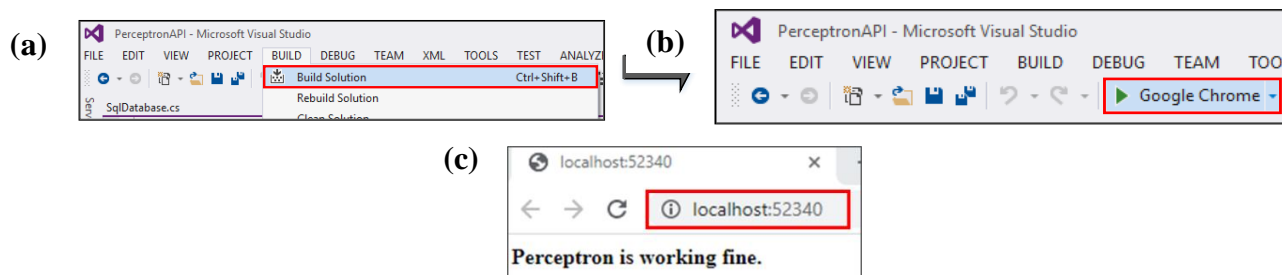


Figure 41. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine”

Message in Figure 41. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine” i.e. “PERCEPTRON API is working fine” ensures that PERCEPTRON API has been successfully built and is running fine.

11.2 PERCEPTRON Front-End

Open Visual Studio Code. From '**File**' dropdown menu, select '**Open Folder**'. Select the folder '**PerceptronFrontEnd**' and click on '**Select Folder**.'

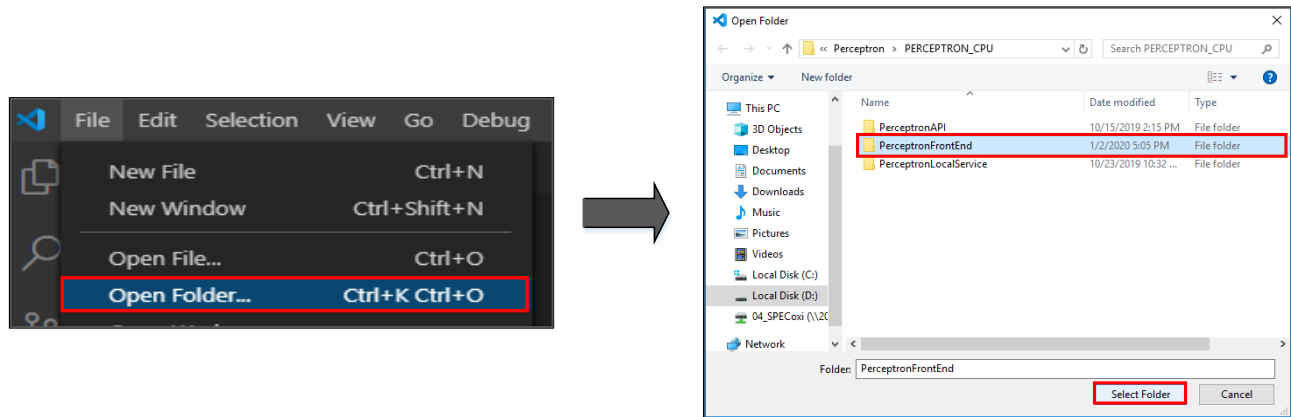
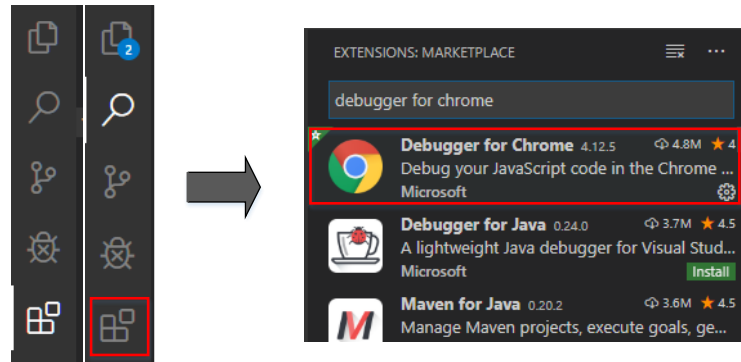


Figure 42. Opening PerceptronFrontEnd folder in Visual studio code

11.3 Connecting PERCEPTRON API to PERCEPTRON Front-End

In Visual Studio Code, click on the **'Extensions'** icon (as indicated below) and type **'Debugger for Chrome'** in the search bar. Select the one shown below.



Click **'Install'** to install the debugger.

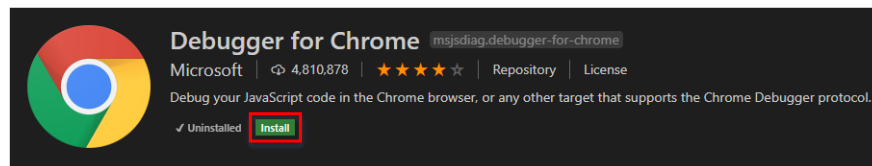
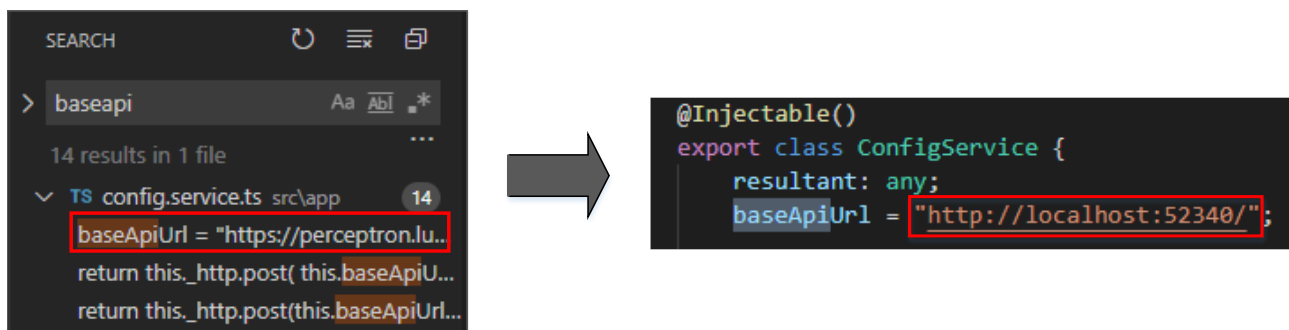


Figure 43. Installing extension in VS Code

Next, type **'baseAPI'** in the search bar and select the one shown below. Replace the given URL with the one saved before in Figure 41. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message "PERCEPTRON is working fine"



Type **'ng build'** in Terminal. A new folder by the name of **'dist'** will form in the same directory.

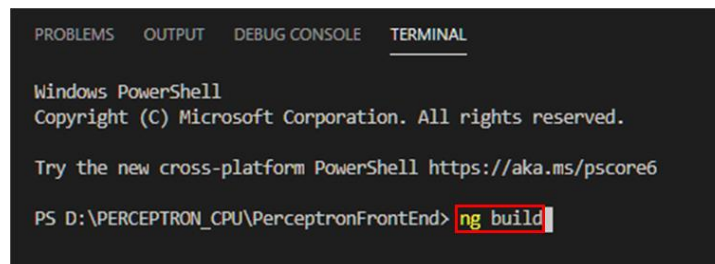


Figure 44. Building Perceptron Front-End

At bottom in Terminal Window, click on the icon  located at right side to clear Terminal and then open **'New Terminal'** from **Toolbox**. Type **'ng serve'** and press **ENTER**. Open the link (<http://localhost:4200/>)

PERCEPTRON: User Manual

provided in Figure 41. Steps to build PERCEPTRON API. (a) Building PERCEPTRON API solution file (PerceptronAPI.sln); (b) Run PerceptronAPI; (c) Google chrome page with URL localhost:52340 displaying the message “PERCEPTRON is working fine” by pressing CTRL button + left mouse button.

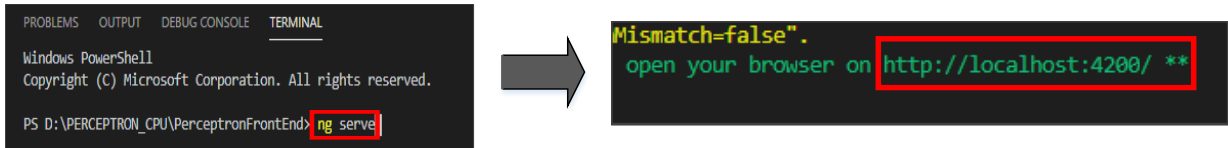
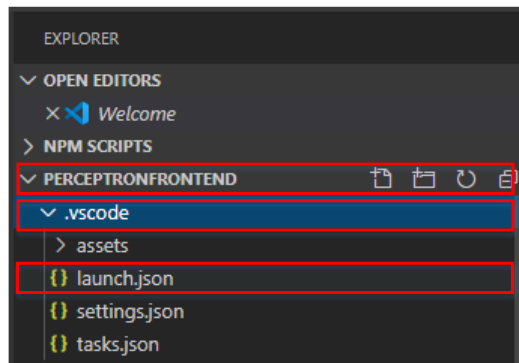
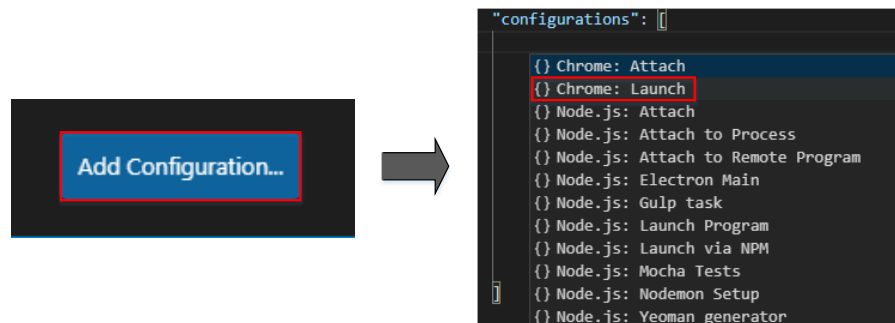


Figure 45. Hosting Perceptron Front-End on system with default URL

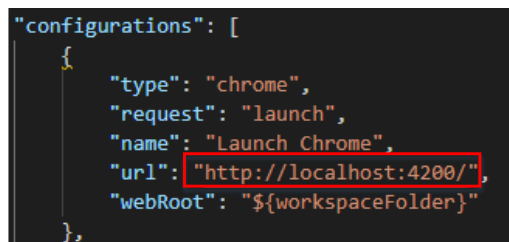
In Explorer, select ‘PERCEPTRONFRONTEND’. Click on ‘.vscode’ and open ‘launch.json’.



Click ‘Add Configuration’ icon on the bottom right corner and select ‘Chrome: Launch’ from the dropdown menu that appears



Replace the URL acquired in Figure 45. Hosting Perceptron Front-End on system with default URL.



In the Toolbox, go to ‘Debug’ and select ‘Start Debugging’. In the Pop Up that appears, click ‘Debug Anyway’ and then ‘Open launch.json’.

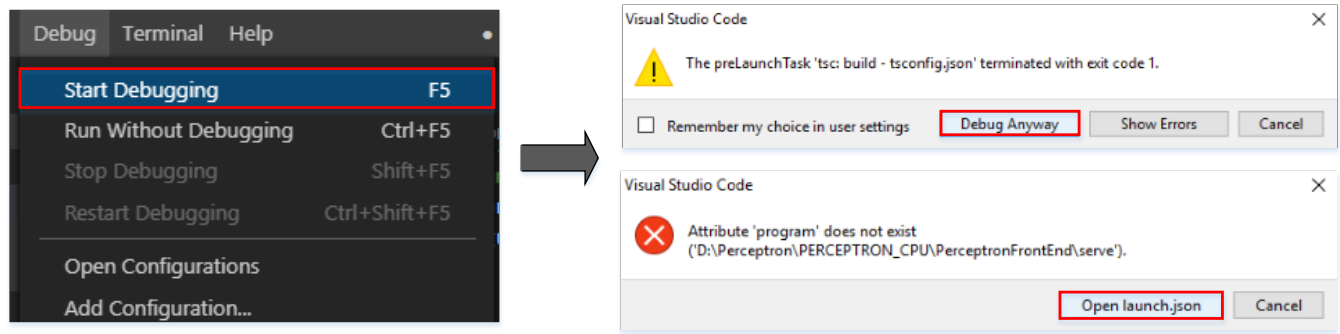


Figure 46. Steps to build PERCEPTRON Front-End

11.4 PERCEPTRON Local Service

Open SQL Server Management Studio. Copy the given server name from SQL Server Pop Up.

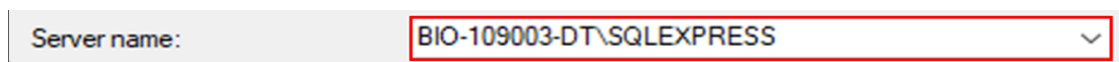


Figure 47. Copying Server Name from SQL Server Management Studio

Open 'PerceptronLocalService.sln' from the folder PERCEPTRON_CPU in visual studio. In the Solution Explorer, select 'App.config'.



Figure 48. Building Perceptron Local Service. (a) Opening PerceptronLocalService.sln; (b) Opening App.config from Solution Explorer of Visual Studio

Press **CTRL+F** to search server name. Type '**CHIRAGH-II**' and press **ENTER**. Replace this server name with the name copied in Figure 47. Copying Server Name from SQL Server Management Studio, e.g. **BIO-109003-DT\SQLEXPRESS**. Replace the given name for all occurrences of '**Data Source**' in App.config with the server name from SQL Pop Up.

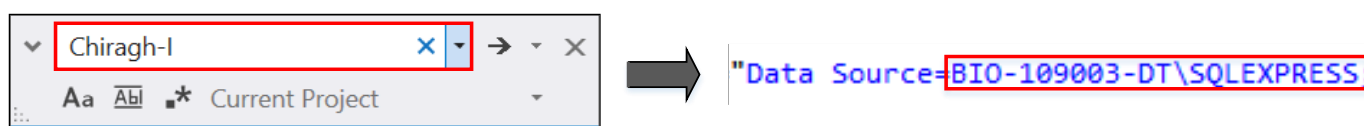
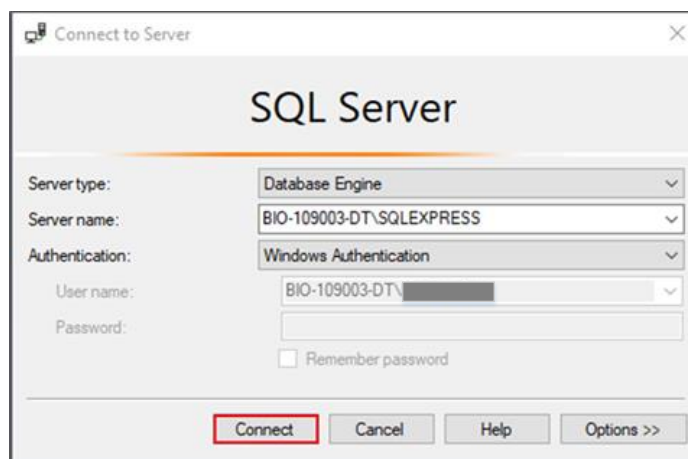


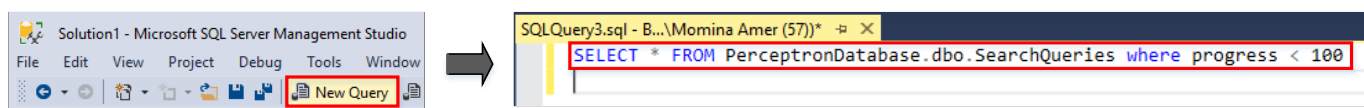
Figure 49. Changing data source value (Server Name) in App.config

Send a Protein Search Query on local host (see Section 7.2)

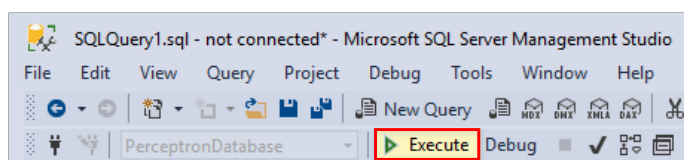
Open SQL Server Management Studio and click on '**Connect**'.



Select 'New Query' and type 'SELECT * FROM PerceptronDatabase.dbo.SearchQueries where progress < 100' in the window that appears.



In SQL, click on 'Execute'.

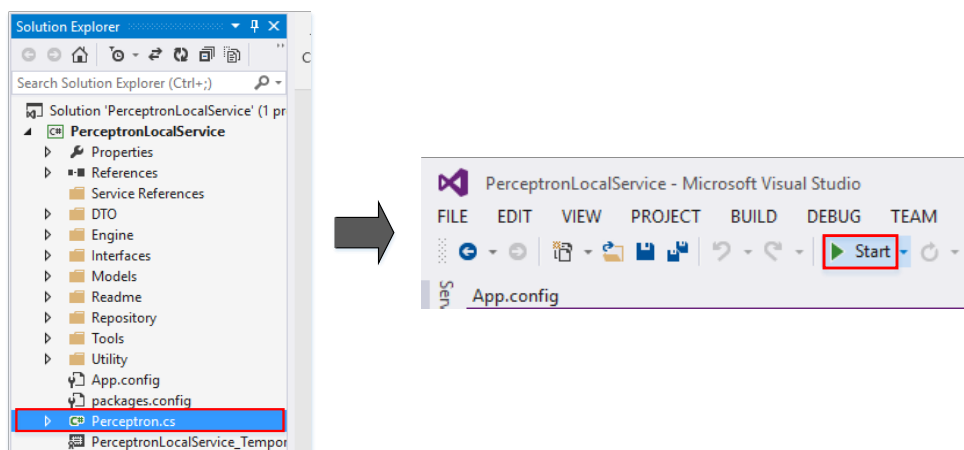


The query will show up in the Results Window.

Results		Messages			
QueryId	UserId	Progress	CreationTime	Email	
1	1ec2fba3-d044-43fc-ba87-15edc62fdd6d	0	01/21/2020 18:55:20	NULL	

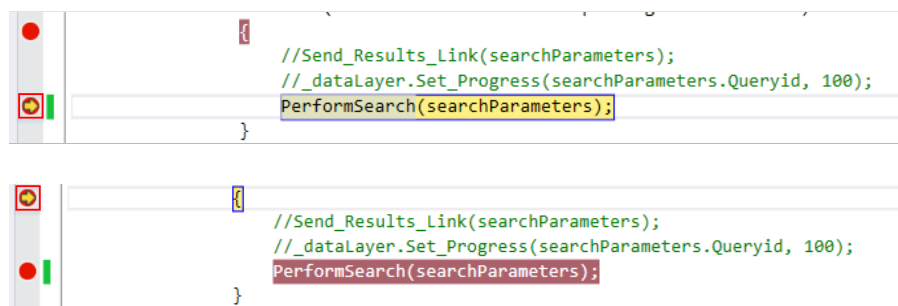
Figure 50. Steps to execute Protein Search Query

In Visual Studio, open 'PerceptronLocalService.sln'. Go to Solution Explorer and select 'Perceptron.cs'. Click 'Start' to run it.



PERCEPTRON: User Manual

Click ‘Continue’ twice to step over the breakpoints.



On the locally hosted PERCEPTRON, click on ‘Search Results & History’ to see the results of your query.

Sr. #	File Name	Protein ID	Protein Score	Molecular Weight	Truncation	Spectral Matches	No. of Modifications	Total Search Time
1	\\App_Data\\HEL A_pk13_sw1_6 6sc_mono.txt	Q4VC39	0.2316	11397.7953	No	1	1	

Figure 51. Testing Perceptron Local Service

12. References

1. Pedrioli PGA, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, Pratt B, Nilsson E, Angeletti RH, Apweiler R. A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* 2004;22(11):1459–1466.
2. Turewicz M, Deutsch EW. Spectra, chromatograms, Metadata: mzML—the standard data format for mass spectrometer output. In: *Data mining in proteomics*. Springer; 2011. p 179–203.
3. Martens L, Chambers M, Sturm M, Kessner D, Levander F, Shofstahl J, Tang WH, Römpp A, Neumann S, Pizarro AD. mzML—a community standard for mass spectrometry data. *Mol Cell Proteomics* 2011;10(1):R110. 000133.
4. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;20(18):3551–3567.
5. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* 2012;30(10):918.
6. Liu X, Inbar Y, Dorrestein PC, Wynne C, Edwards N, Souda P, Whitelegge JP, Bafna V, Pevzner PA. Deconvolution and database search of complex tandem mass spectra of intact proteins a combinatorial approach. *Mol Cell Proteomics* 2010;9(12):2772–2782.