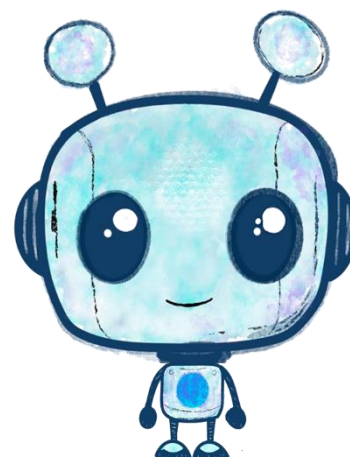# HydroBot - User Manual

Software for interactive hydrogen/ deuterium-exchange mass spectrometry data analysis

Authors: Monika Kiš

🔽 Download HydroBot

You can download the latest HydroBot executable here:

👉 [HydroBot v1.0.0 (Windows .exe)]

(https://github.com/monikakish/HydroBot/releases/tag/v1.0.0)

---

1. Introduction

HydroBot is an interactive desktop application for comprehensive **HDX-MS data analysis and visualization**. It provides:

- **Uptake analysis** at peptide level.

- **Statistical validation** of differences between protein states (Welch's t-test, global thresholds).

- **Visualization tools**: uptake plots, Woods plots, volcano plots, error distributions, and heatmaps.

- **Clustering** (k-means, hierarchical) and peptide trajectory analysis.

- **Direct export to PyMOL** for structural mapping.

HydroBot streamlines HDX-MS workflows from raw .csv datasets to publication-ready figures and 3D structural insights.

---

2. Installation

HydroBot is distributed as a stand-alone application. No Python installation is required.

Simply download the latest version for your operating system from the HydroBot GitHub releases page.

Windows

1. Download the latest HydroBot.exe from the GitHub releases page [monikakish/HydroBot: Software for interactive hydrogen/ deuterium-exchange mass spectrometry data analysis].

2. Place it in a folder of your choice or desktop.

3. Double-click HydroBot.exe to launch the program.

⚠️ **Note***: On first launch, Windows Defender or macOS Gatekeeper may warn that the file is from an unknown developer. Select **Run anyway** to proceed.*

---
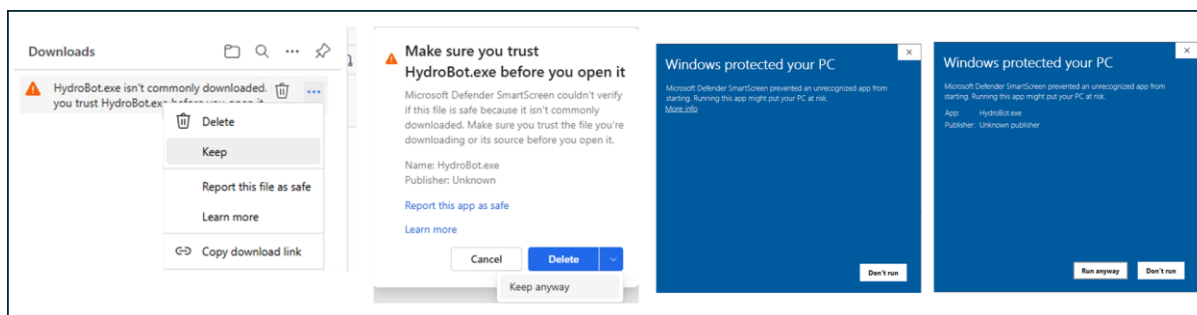
3. Input Data

- **Supported**: CSV output files from **DynamX** or converted files from other HDX software.

- If your data is not from DynamX, use the **File Conversion tab** to map required columns. **Refer to the File conversion section of this document!**

- Required fields typically include:

  o Protein

  o Start / end residues

  o Sequence

  o MaxUptake

  o Charge

  o MHP

  o State

  o Exposure

  o File

o   Inten

o   Center

---

4. Using HydroBot

HydroBot has **four main tabs**:

4.1 HDX and Difference Analysis

- Import dataset and select states for comparison.

- Generate **uptake plots** (normalized to max uptake = 100%).

- Perform **Welch's t-test** with a **global significance threshold**.

- Visualization options:

    o   **Bar plots** - per peptide/amino acid difference.

    o   **Woods plots** - uptake differences across sequence coverage.

    o   **Heatmaps** - per-residue differences.

    o   **Volcano plots** - ΔD vs. significance.

    o   **Error distribution plots** - replicate variability.

👉 **Tip**: *Use volcano plots with error distributions to evaluate both magnitude and confidence of changes.*

To generate Uptake plots:

Annotations on screenshot:

1. Input the Cluster and State data, Fasta file and select an output folder!

2. Select the time units from the dropdown

3. Select the states to be plotted

4. Choose colors for each state selected

5. To view the plots and interact with them click on the peptides – only 3 can be viewed (hold ctrl + click on the peptides to display them)

- Upload your files → Add Cluster + State data, a FASTA file, and choose an output folder.
- Choose time units → Select from the dropdown.
- Pick your states → Click on the states you want to plot.
- Set colors → Assign a color to each selected state in the order they were selected.
- View plots → Click a peptide to see its plot (max 3 at once). Hold Ctrl + click to select multiple peptides.
- Saved results → Plots are saved automatically in two subfolders inside your chosen output folder: Matplotlib and Plotly.

To perform difference analysis:
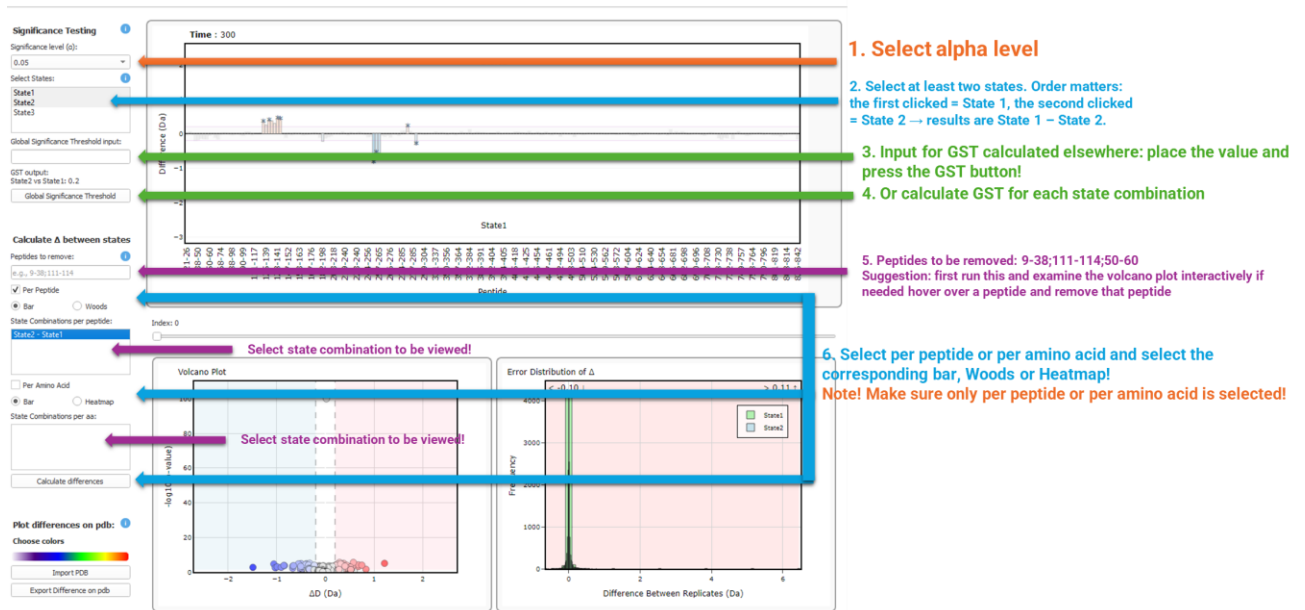
- Select alpha level → Choose your significance level.
- Pick states → Select at least two states. Order matters: the first clicked = State 1, the second clicked = State 2 → results are State 1 – State 2.
- Input GST → If you already have a value, enter it and press Global Significance Threshold.
- Or calculate GST → Let the app calculate GST for each state combination.
- Remove peptides → Example: 9-38; 111-114; 50-60.
- 💡 Tip: *First run the volcano plot, explore interactively, hover over peptides, and then remove sequences if needed*.
- Choose analysis type → Select per peptide or per amino acid, then pick the corresponding Bar, Woods, or Heatmap plot.
- ⚠️ Note! *Only one mode (per peptide OR per amino acid) should be selected at a time.*
- View state combinations → Choose which state comparison you want to display.
- Drag the slider to view the differences per time point!
- Saved results → Plots are stored in your chosen output folder inside two subfolders: per peptide and per amino acid! Plots are also saved as plotly and matplotlib. Also a csv file with the difference data is saved!

**To export difference on a pdb structure:**
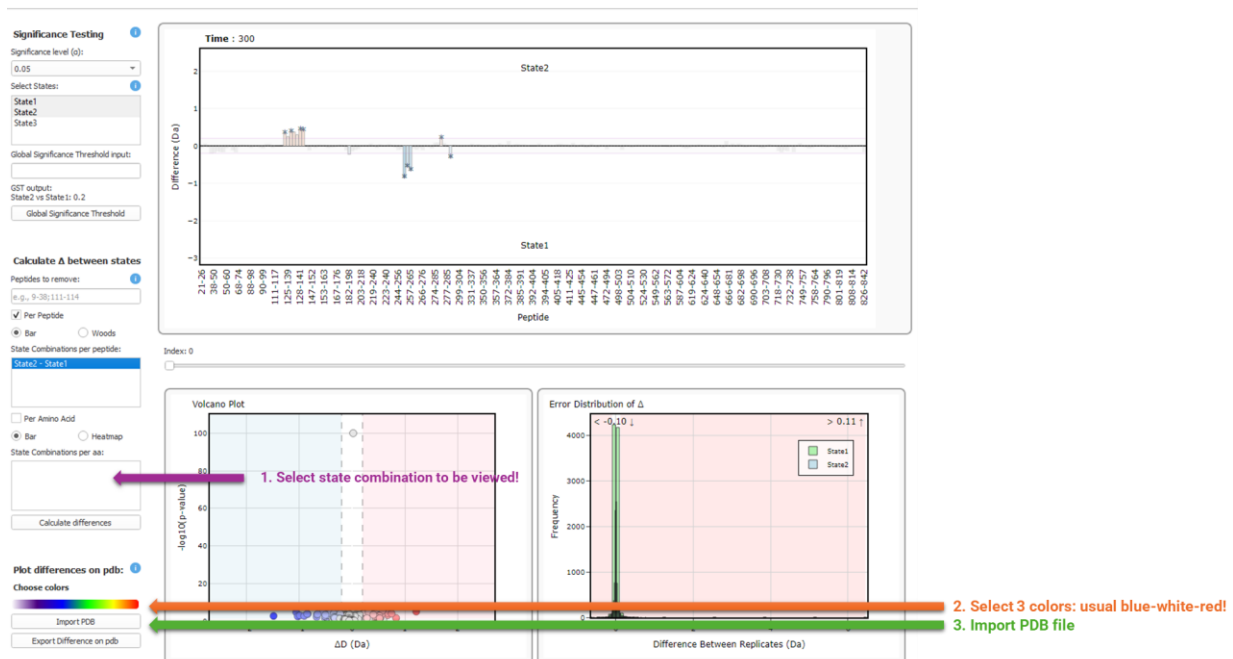
- Select state combination → Pick one from the per amino acid list.
- Choose color palette → Usually blue–white–red!
- Import PDB file → Load your PDB.

  ⚠️ *Coloring is applied per amino acid across the whole structure. If your PDB has multiple chains, all chains will be colored.*

- Export difference → Click Export Difference on PDB.
- Open in PyMOL → The export creates a .pml file. In PyMOL, type run, then drag & drop the file.
- View results → Differences are plotted per time point for the selected state combination, normalized to the maximum of the whole dataset.
- Saved results → The .pml file is saved in the Difference per amino acid folder inside your chosen output folder.

4.2 Clustering Analysis

- Compare **three states** at a time (subtract the other two from the reference).

- Supported clustering methods:

  - **k-means** (choose number of clusters, elbow/Silhouette analysis included).

  - **Hierarchical clustering** (Ward's linkage, Euclidean distance, dendrogram view).

To generate elbow and Silhouette plots:
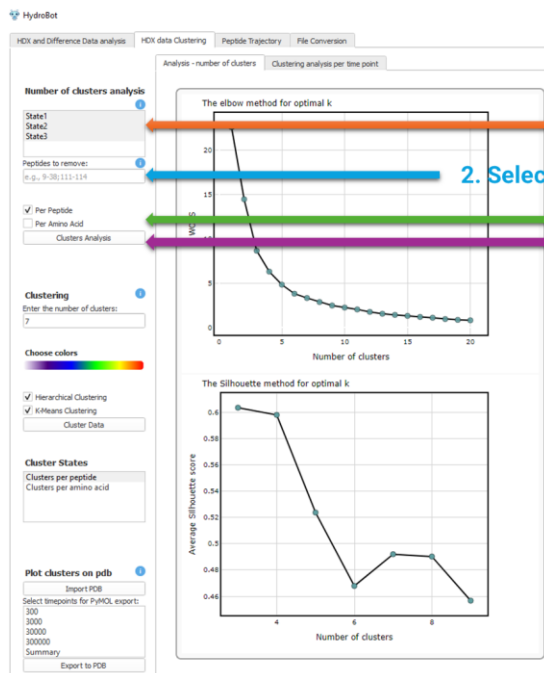
- Select states → Typically in this **_specific_** order: apo, eq, non-eq.
- x-axis: non-eq – apo
- y-axis: non-eq – eq
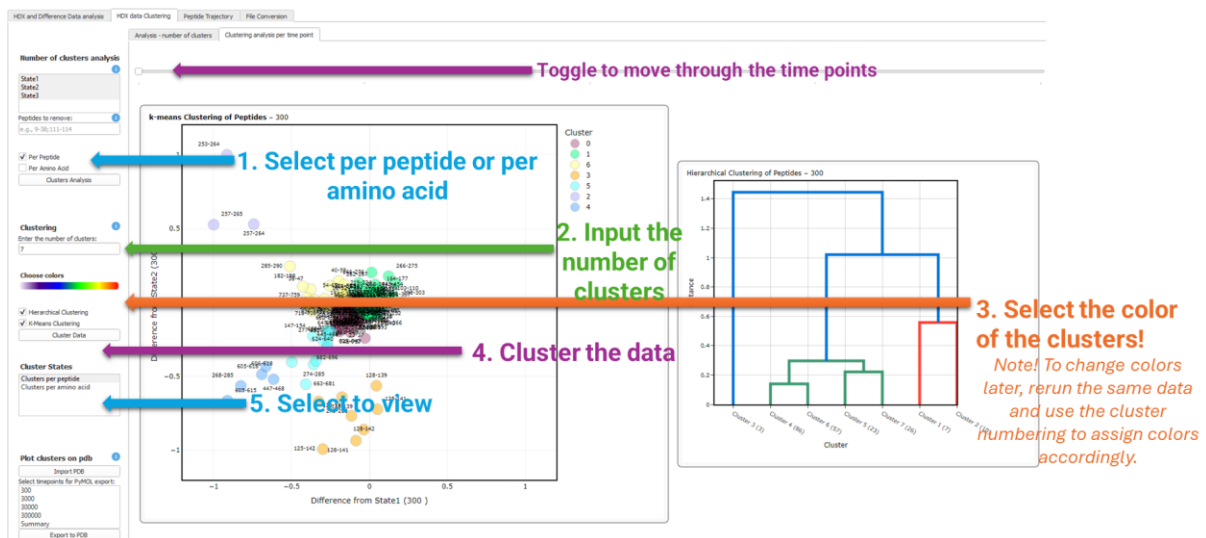- Remove peptides → Select any peptides you want to exclude from the analysis.
- Choose analysis type → Pick per peptide or per amino acid.
- Perform analysis → Generate Elbow plot and Silhouette plot to determine the optimal number of clusters.

To perform clustering analysis:

- Input number of clusters → Enter how many clusters you want to generate.
- Select cluster colors → Pick a color for each cluster.

  ⚠️ *To change colors later, rerun the same data and use the cluster numbering to assign colors accordingly.*

- Cluster data → Click Cluster Data to perform the clustering.
- View clusters → Choose per peptide or per amino acid to display the plots.
- Navigate time points → Use the slider to move through the different time points.
- Saved results → Plots are stored in your chosen output folder inside two subfolders: Clusters per peptide and per amino acid.

To export clusters on pdb:
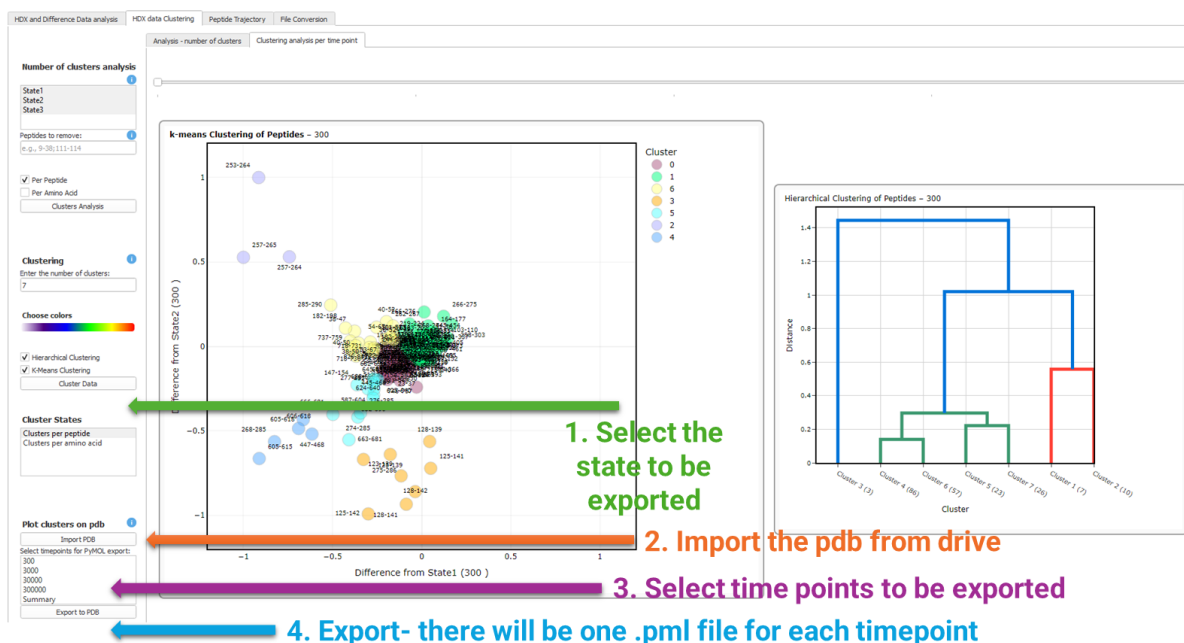
- Select state → Choose the state you want to export.
- Import PDB file → Load the PDB from your drive.
- Select time points → Pick the time points to export.
- Export clusters → Click Export. A separate .pml file will be created for each selected time point.
- Open PyMOL → In PyMOL, type run, then drag & drop the file → press Enter.
- Cluster colors → The colors you selected for clustering will be applied to the PDB structures.
- Saved results → The .pml files are saved in the Clusters per peptide or per amino acid folder inside your chosen output folder.

4.3 Peptide Trajectory

- Visualize **temporal changes** in peptide dynamics across labeling times.

- Tracks cluster membership over time.

- Overlay multiple peptide trajectories to detect:

  o Transient unfolding

  o Allosteric transitions

  o Switching cluster dynamics
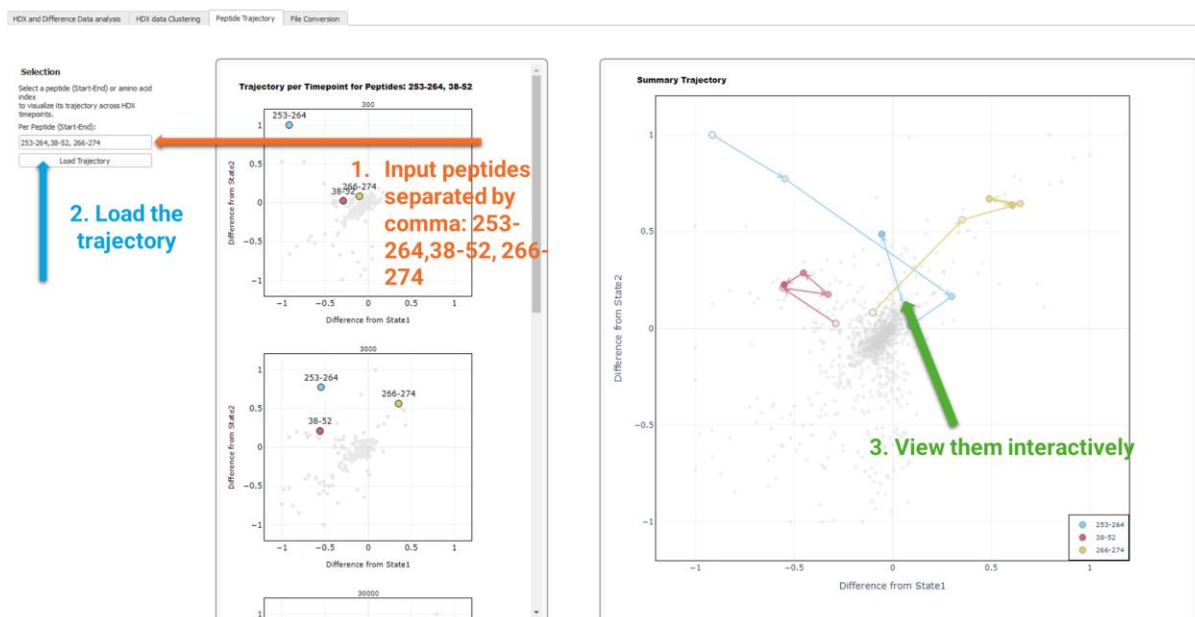
- Input peptides → Enter peptide ranges separated by commas, e.g., 253-264, 38-52, 266-274.

- Load trajectory → Import the trajectory file to visualize.
- View interactively → Explore the selected peptides interactively in the viewer.

4.4 File Conversion

The File Conversion tab is designed to work with output from any HDX-MS processing software. To ensure compatibility, the input CSV file must have column names defined in the first row.

Within the tab, users are required to map each column of their dataset to the corresponding DynamX cluster CSV format by selecting the appropriate column from the dropdown menus. If certain columns (e.g., *Max Uptake* or *MHP*) are not present in the input file, HydroBot can automatically calculate these values from the sequence, which is also available for selection in the dropdown.

5. Citation

If you use HydroBot in your work, please cite:
[placeholder]