deMix

User manual

Table of Contents

1. Introduction

1.1. About deMix

One of the most popular methods for examining protein conformational changes and dynamics is hydrogen/deuterium exchange (HDX) with mass spectrometry (MS). We previously developed a fully automated algorithm to analyze deuterated isotopic distributions in-depth called deMix (Na et al. 2019 [1]). Here, we introduce the graphical user interface(GUI) version of deMix. The software automatically analyzes HDX data and facilitates the interrogation of data and results. deMix offers visualization of isotopic cluster distributions and sequence coverage maps in heat map form to compare deuteration rates over time.

What deMix provides:

Peptide centric view (Deuterated Distributions)

- Allows users to compare a natural isotope distribution and the corresponding deuterated isotope distribution in one of HDX experiments across D2O labeling with theoretical, aggregated (over elution time spans), and manually annotated distributions for the chosen peptide.
- Offers a deuteration rate (or deuterium uptake) plot across D2O labeling time for the selected peptide.

Protein centric view (Dynamics)

- Allows users to view the dynamics of the protein.
- Offers sequence coverage maps for HDX-MS data. Visually represents the HDX rate of each D2O labeling times within the HDX-MS dataset using colors (= Heat Map).

User-friendly environment

- Includes intuitive and interactive GUI or features.
- Allows users to focus more on analysis rather than spend time adapting to the software.

2. Installation

2.1 Requirements

java version > 17.05

To check (type this in the command prompt) = **java - version**

Download java here: https://www.oracle.com/java/technologies/downloads/

deMix Downlaods:

Windows

- Download Windows version. & Extract the compressed file.
- Double Click the bat file → Run anyway/Run

MacOs

- Download Windows version. & Extract the compressed file.
- Right-click the deMix → Open with Terminal → Open.

Linux

- Download the linux version. & Extract the compressed file.
- Double Click the .sh

* Remove visited_directories file if exists (when first downloaded).

.

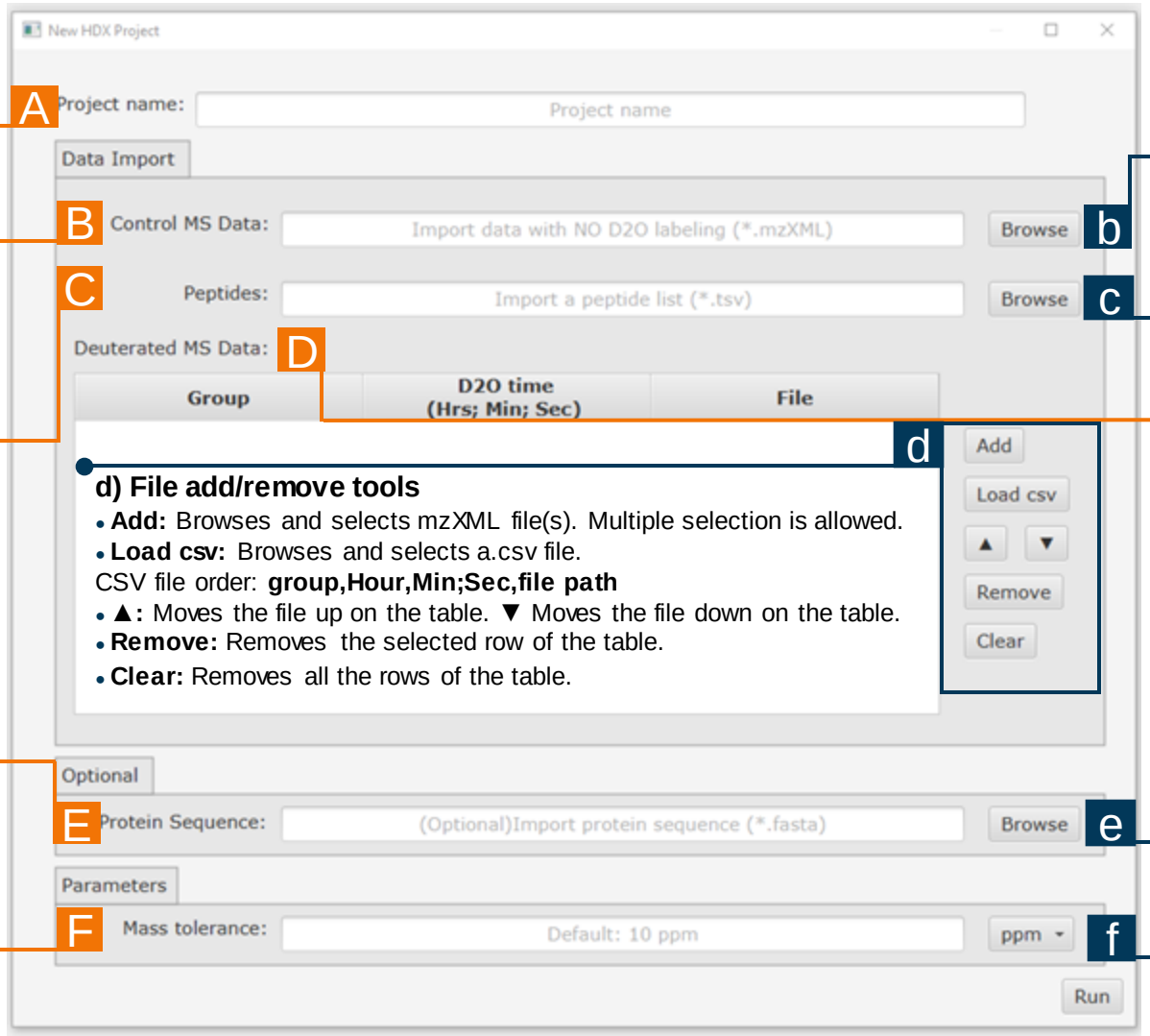
3. Peptide Centric View

3. Protein Centric View

3.1. File Control

3.1.1. Creating a New deMix Project

Click Path: File > New Project



* required

A) Project name *
- Specify Project Name

B) Control MS Data *
- Import control data which does not have D2O labeling.
- File Format: mzXML

C) Peptides *
- Import a file containing peptides to be examined.
- Make sure that the file contains two columns named 'peptide' and 'charge' for peptides and charge states, respectively.
- File Format: tsv

E) Protein Sequence
- Import a protein sequence data
- File Format: fasta

F) Mass Tolerance
- Specify the mass tolerance (ppm or da).
- default = 100ppm

b) Browse button
- Browses and selects .mzXML files. Multiple selection is allowed.

c) Browse button
- Browses and selects a .tsv file.

D) Deuterated MS Data *
- Import D2O labeled data set(s) (With Group and D2O Time). User must provide D2O labeling.
- File Format: MSXML

d) File add/remove tools
- **Add:** Browses and selects mzXML file(s). Multiple selection is allowed.
- **Load csv:** Browses and selects a.csv file.

CSV file order: **group,Hour,Min;Sec,file path**
- **▲:** Moves the file up on the table. ▼ Moves the file down on the table.
- **Remove:** Removes the selected row of the table.
- **Clear:** Removes all the rows of the table.

e) Browse button
- Browses and selects a .fasta file.

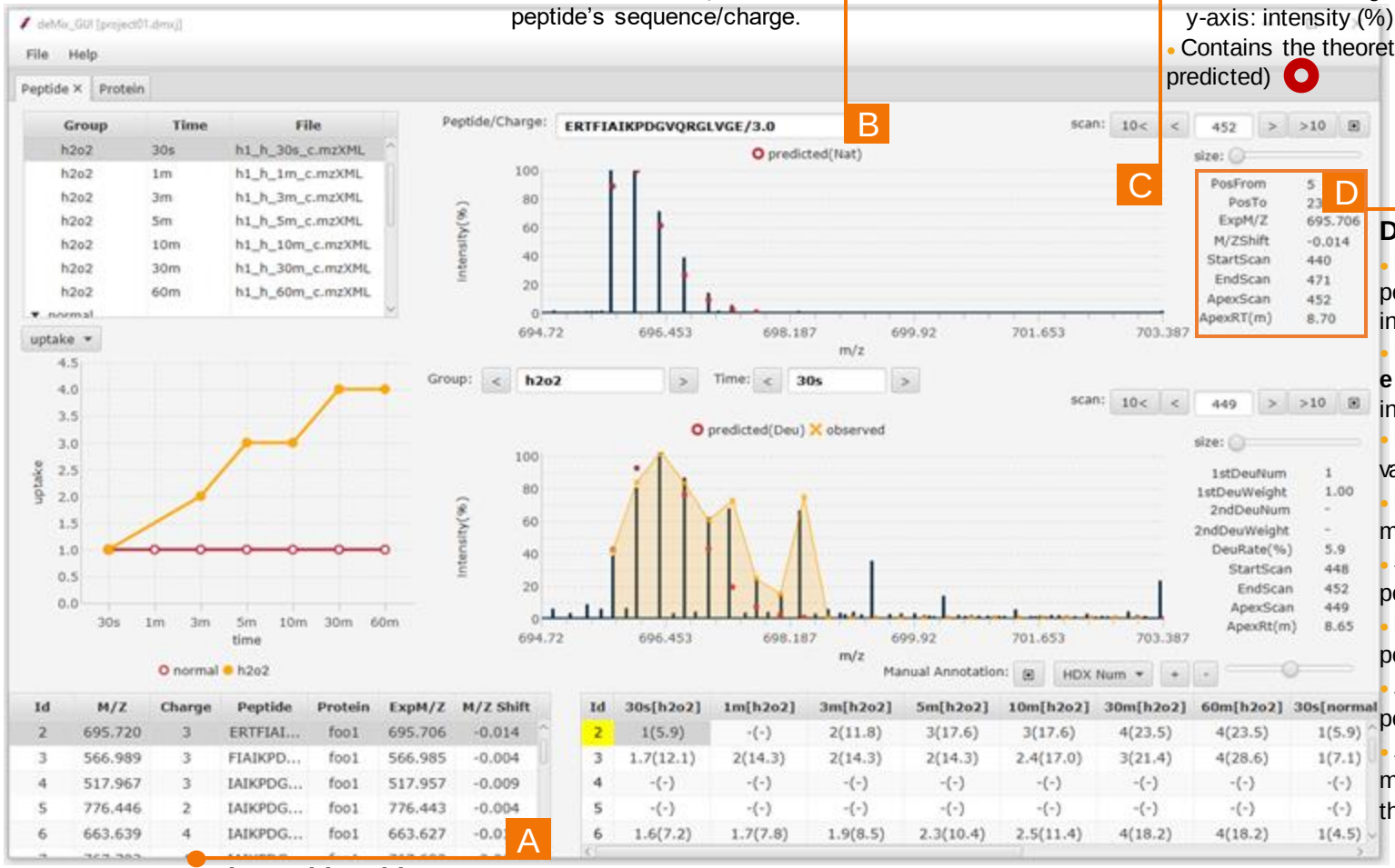
f) ppm/da menu
- Sets the unit of the mass tolerance. (Default: ppm)

3.1.2. Opening an deMix project

Click Path: File > New Project > Open (selected file)

3.2. Natural Isotope Distribution

3.2.1. Distribution Explanations



B) Current Peptide

Shows the currently selected peptide's sequence/charge.

C) Natural Isotope Distribution (Nat.)

- x-axis: mass/charge(m/z)
  y-axis: intensity (%)
- Contains the theoretical distribution of the Nat. (= predicted)

D) Nat. Information

- **PosFrom:** Where the selected peptide **starts** in the given protein.
- **PosTo:** Where the selected peptide **ends** in the given protein.
- **ExpM/Z:** The expected starting m/z value of the distribution.
- **M/Zshift:** Observed m/z - Theoretical m/z
- **StartScan:** Scan number where the peptide elution begins.
- **EndScan:** Scan number where the peptide elution ends.
- **ApexScan:** Scan number where the peptide elution is at the peak.
- **ApexRT(m):** Retention time in minutes when the peptide elution is at the peak.

A) PeptideTable

- Contains peptide's information.

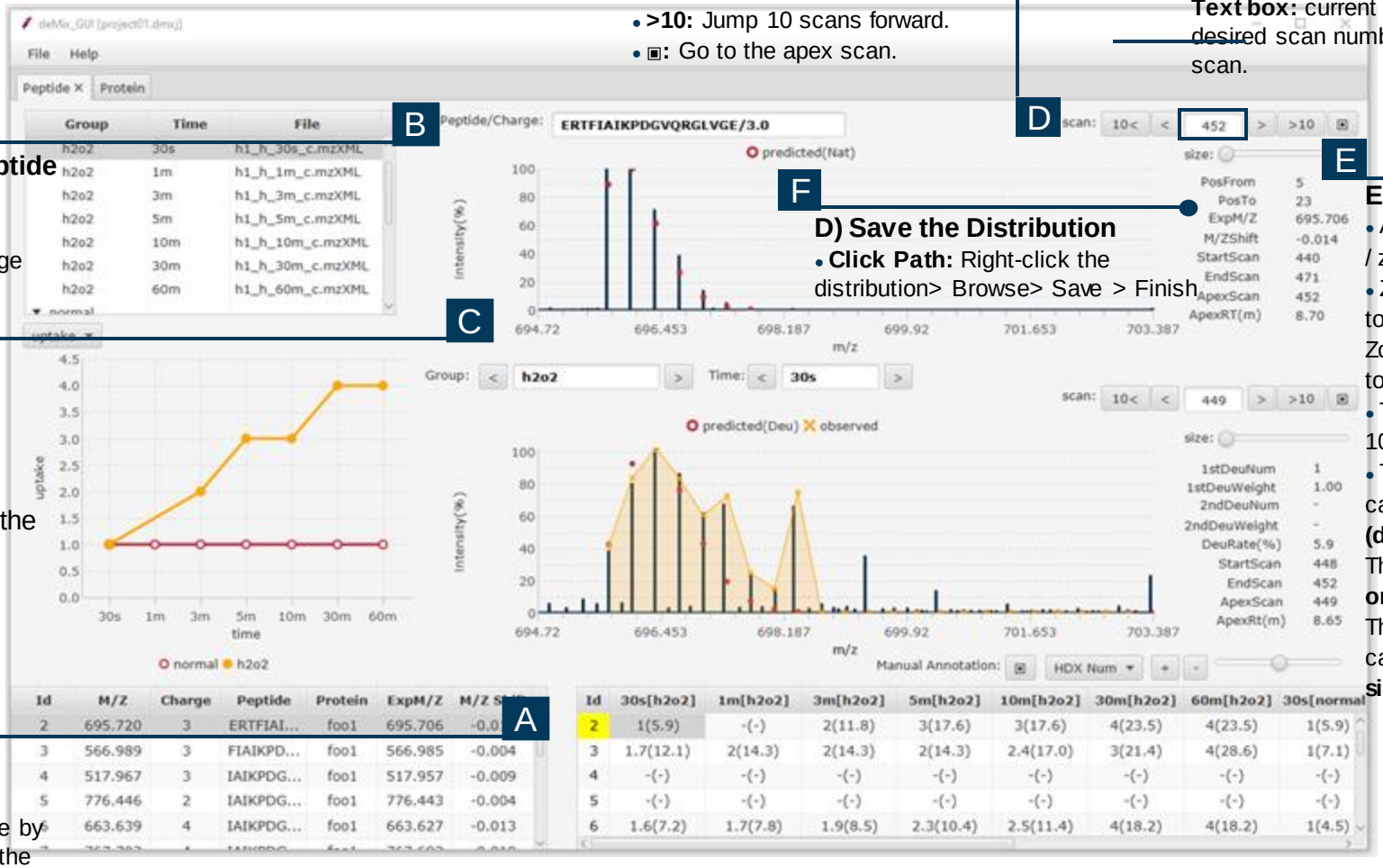
3.2. Natural Isotope Distribution

3.2.2. Distribution Functions

Distribution Control: A → B → C

D) Scan Mover
- **10<:** Jump 10 scans backward.
- **<:** Go to previous scan.
- **>:** Go to the next scan.
- **>10:** Jump 10 scans forward.
- ▣: Go to the apex scan.

Text box: current scan & type the desired scan number to go to that scan.



B) Current Peptide
- The selected peptide's sequence & charge appears.

C) Distribution
- The selected peptide's natural isotope distribution & Information (on the right) appears.

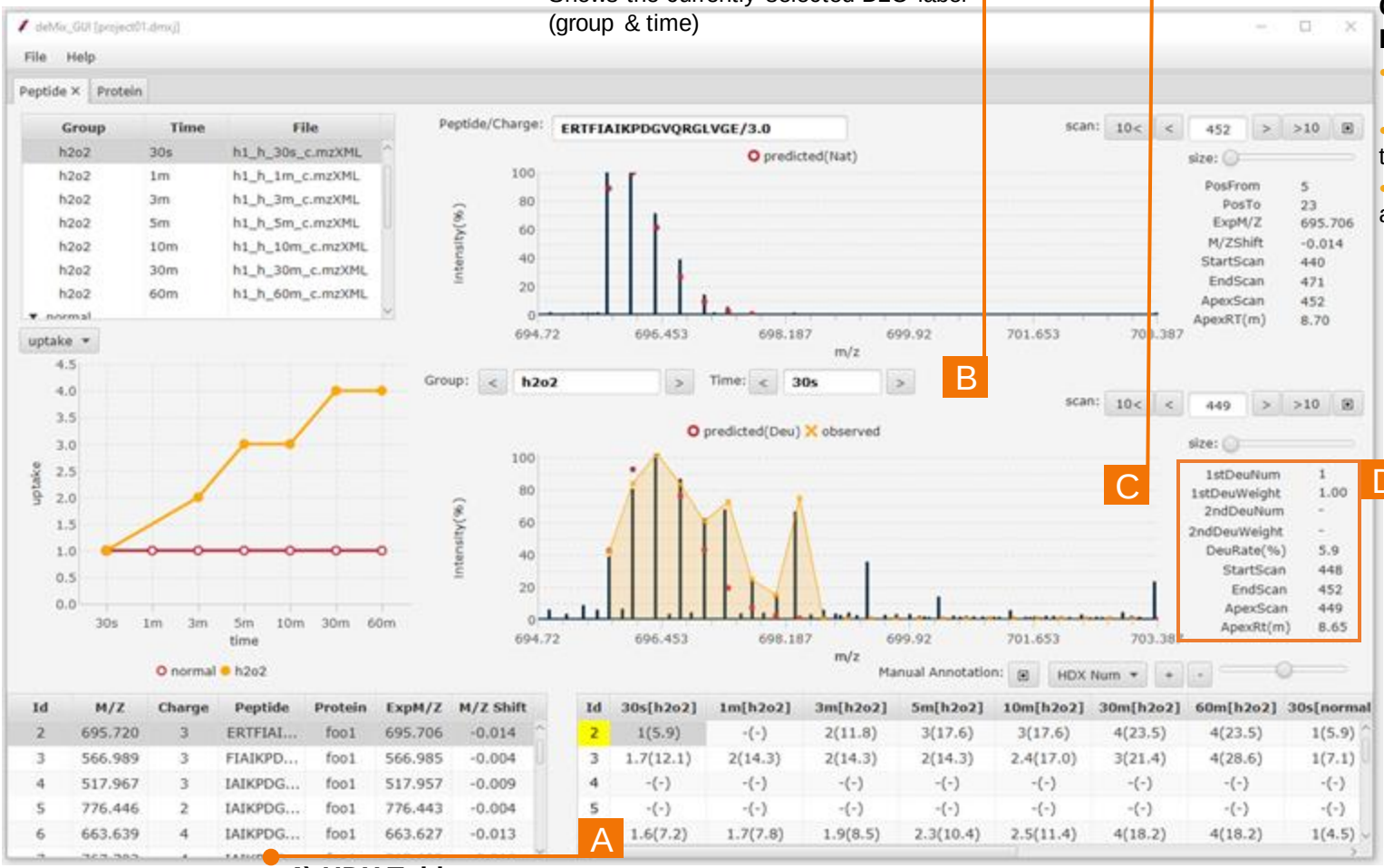
A) Peptide Selection
- Select a peptide by clicking a row in the table.
- Keyboard(up & down arrow keys) support available

D) Save the Distribution
- **Click Path:** Right-click the distribution> Browse> Save > Finish

E) Size Controller
- Adjusts the size (zoom out / zoom in) of the distribution.
- Zoom out: drag the slider to the right.
  Zoom in: drag the slider to the left.
- The slider ranges from 0 to 10 (positive real numbers).
- The zoomed in/out amount calculation**: 5 * size (dragged position).**
  The minimal m/z calculation: **original m/z - size * 5.**
  The maximum m/z calculation: **original m/z + size * 5.**

3.3. Deuterated Isotope Distribution

3.3.1. Distribution Explanations



B) Current D2O Label

Shows the currently selected D2O label (group & time)

C) Deuterated Isotope Distribution (Deu.)

- x-axis: mass/charge(m/z)
  y-axis: intensity (%)
- Contains the theoretical distribution of the Deu. (=predicted)
- Contains the distribution of the aggregated scans. (=observed)

D) Deu. Information

- **1stDeuNum:** Deuterated number (Deuterium Uptake) assuming a unimodal distribution (a single deuterium value).
- **1stDeuWeight:** Proportion of the 1stDeuNum form in the distribution. The weight factor for more abundant num is over 90%, only the abundant one is reported (not accepted as a bimodal distribution).
- **2ndDeuNum:** Second deuterated number assuming the bimodal distribution (i.e., two deuterated forms are simultaneously observed).
- **2ndDeuWeight :** Proportion of the 2ndDeuNum form in the distribution.
- **StartScan:** Scan number where the peptide elution begins.
- **EndScan:** Scan number where the peptide elution ends.
- **ApexScan:** Scan number where the peptide elution is at the peak.
- **ApexRT(m):** Retention time in minutes when the peptide elution is at the peak.

A) HDX Table

- Contains each D2O label's HDX Number.

3.3. Deuterated Isotope Distribution

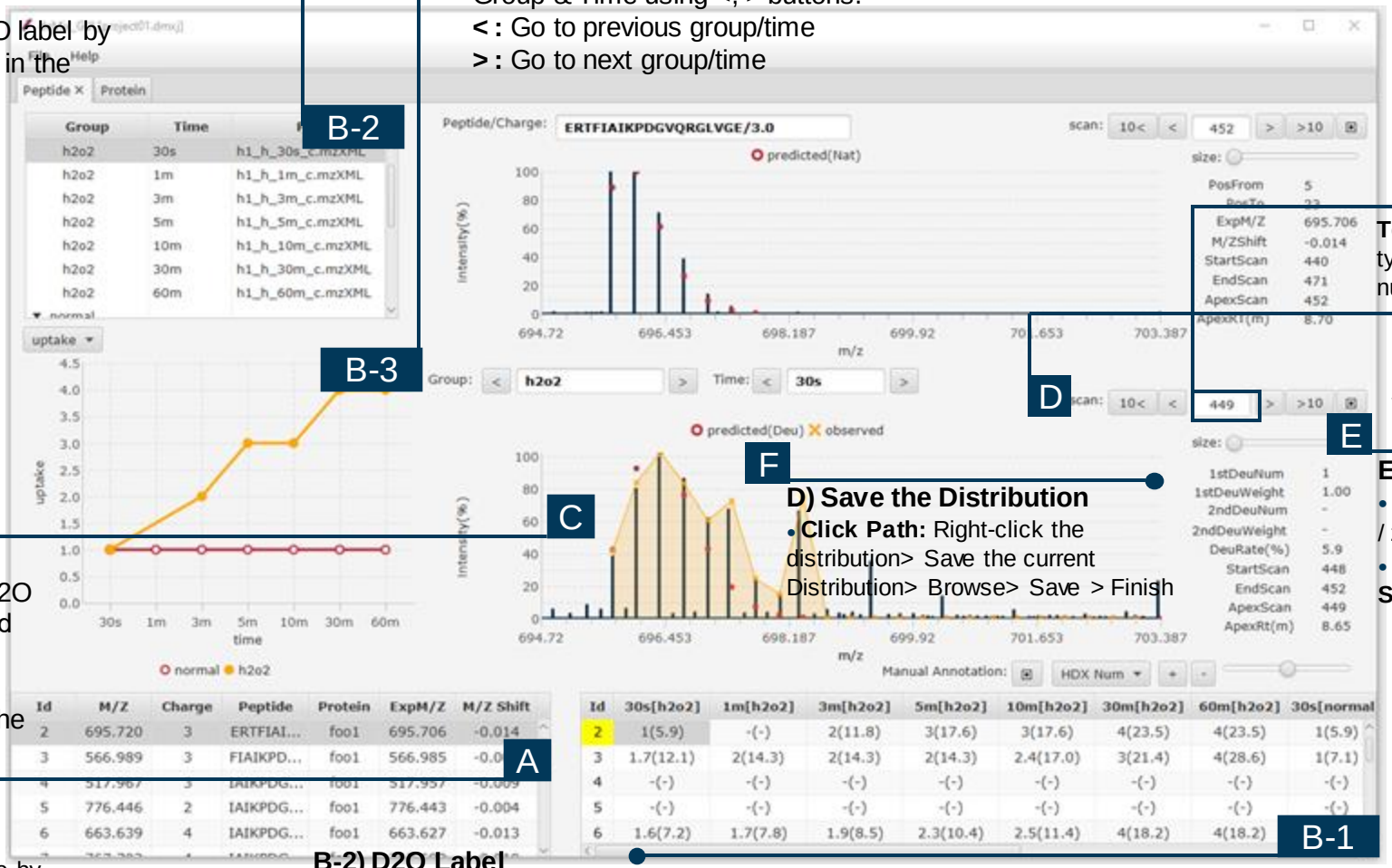
3.3.2. Distribution Functions

Distribution Control: A → B-1/B-2/B-3 → C

B-2) D2O Label Selection
- Select a D2O label by clicking a row in the sample view.

B-3) D2O Label Selection
- Select a D2O label by selecting Group & Time using <, > buttons.

< : Go to previous group/time
> : Go to next group/time



Text box: current scan & type the desired scan number to go to that scan.

D) Scan Mover
- Functions same **as Section 3.3.2, D.**

E) Size Controller
- Adjusts the size (zoom out / zoom in) of the distribution.
- Function same as **Section 3.3.2, E.**

C) Distribution
- The selected D2O label's deuterated isotope distribution & Information (on the right) appears.

D) Save the Distribution
- **Click Path:** Right-click the distribution> Save the current Distribution> Browse> Save > Finish

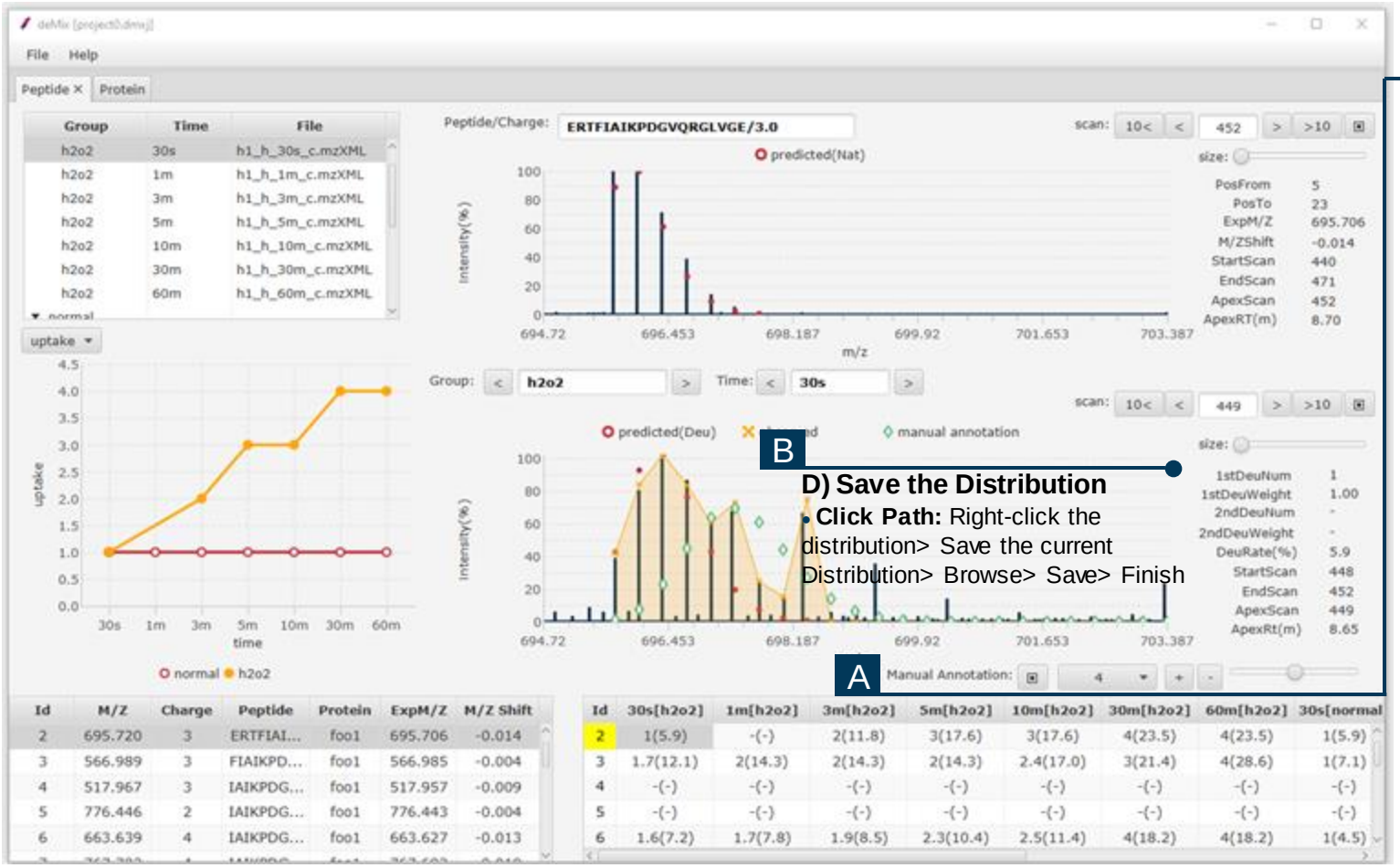
A) Peptide Selection
- Select a peptide by clicking a row in the table.
- Keyboard(up & down arrow keys) support available

B-2) D2O Label Selection
- Select a D2O label by clicking a row in the sample view.

B-1) D2O Label Selection
- Select a D2O label by clicking a cell in the table.
- The cell must be in the same row as the highlighted id.

3.3. Deuterated Isotope Distribution

3.3.3. Manual Annotation



A

B

A) Manual Annotation ◊

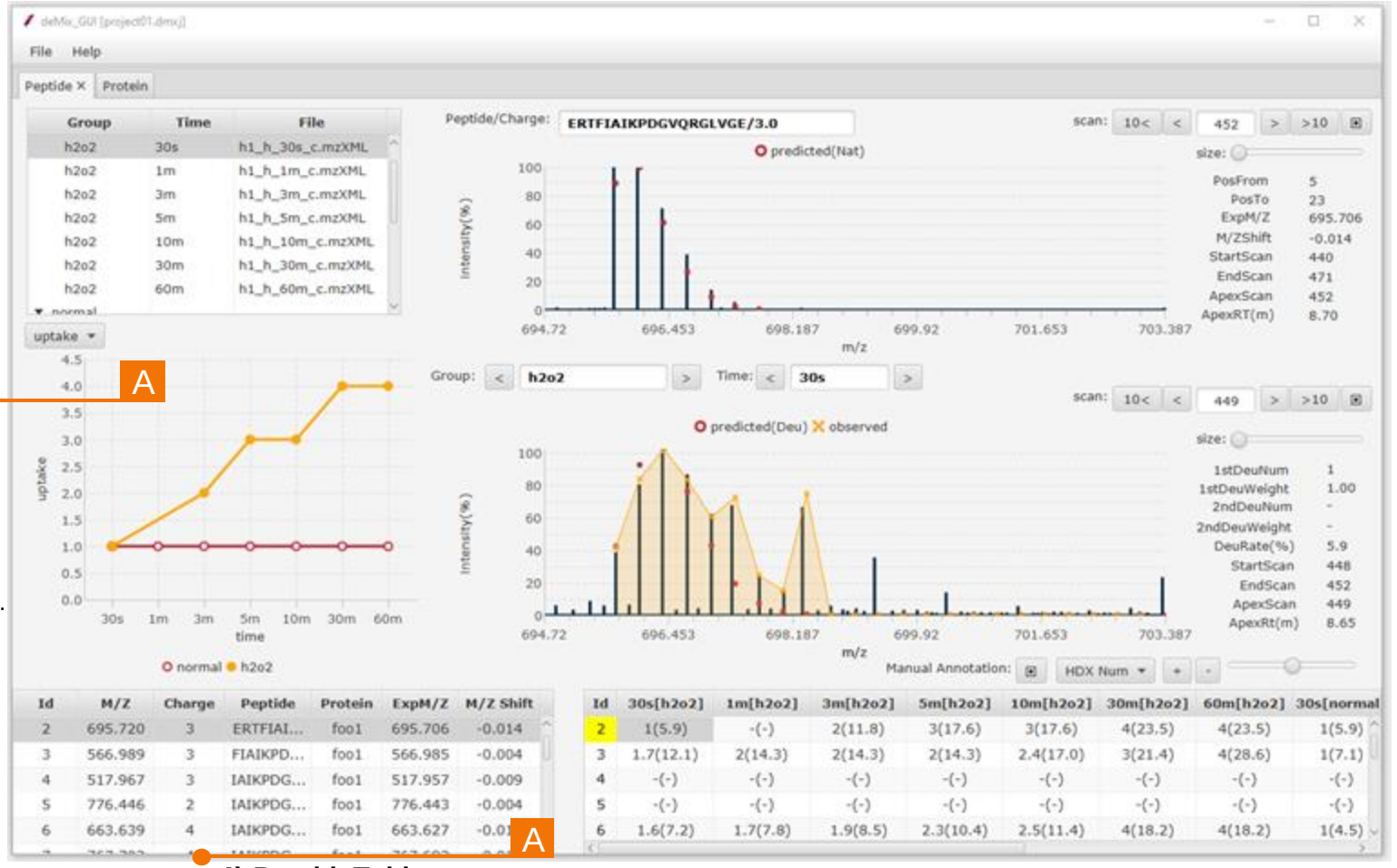
- ▣ : Set manual annotation to system's theoretical distribution (predicted)
- **HDX Num Menu:** A menu with all the possible HDX numbers, which are from 1 to length of the peptide -1. The menu box shows the currently selected HDX Number.
- **+:** Increases the y-axis value of the distribution for the set amount.
  Calculation (each data point): Current y-axis value * 1.05
- **-:** Decrease the y-axis value of the distribution for the set amount.
  Calculation (each data point): Current y-axis value / 1.05
- **:** Increases/decreases the y-axis value of the distribution
- The slider ranges from 0 to 10 (positive real numbers). The default value is 5 (original).
- Increase: drag the slider to the right (dragged value > 5).
  Decrease: drag the slider to the left(dragged value < 5)
- Dragged Amount: |5- dragged value|
  Decrease Calculation (each data point): y-axis value / (1 + Dragged Amount)
  Increase Calculation (each data point): y-axis value * (1 + Dragged Amount)

D) Save the Distribution

- **Click Path:** Right-click the distribution> Save the current Distribution> Browse> Save> Finish

3.4. HDX Plot

3.4.1. Plot Explanations



B) HDX Plot

- Shows the change in the amount of deuterium uptake/HDX rate (can be over time (all D2O labeling time).
- The range of the y-axis varies for each peptide selection.
- The maximum number of groups is 8.
- It disregards the D2O labeling time point where the HDX number unavailable.

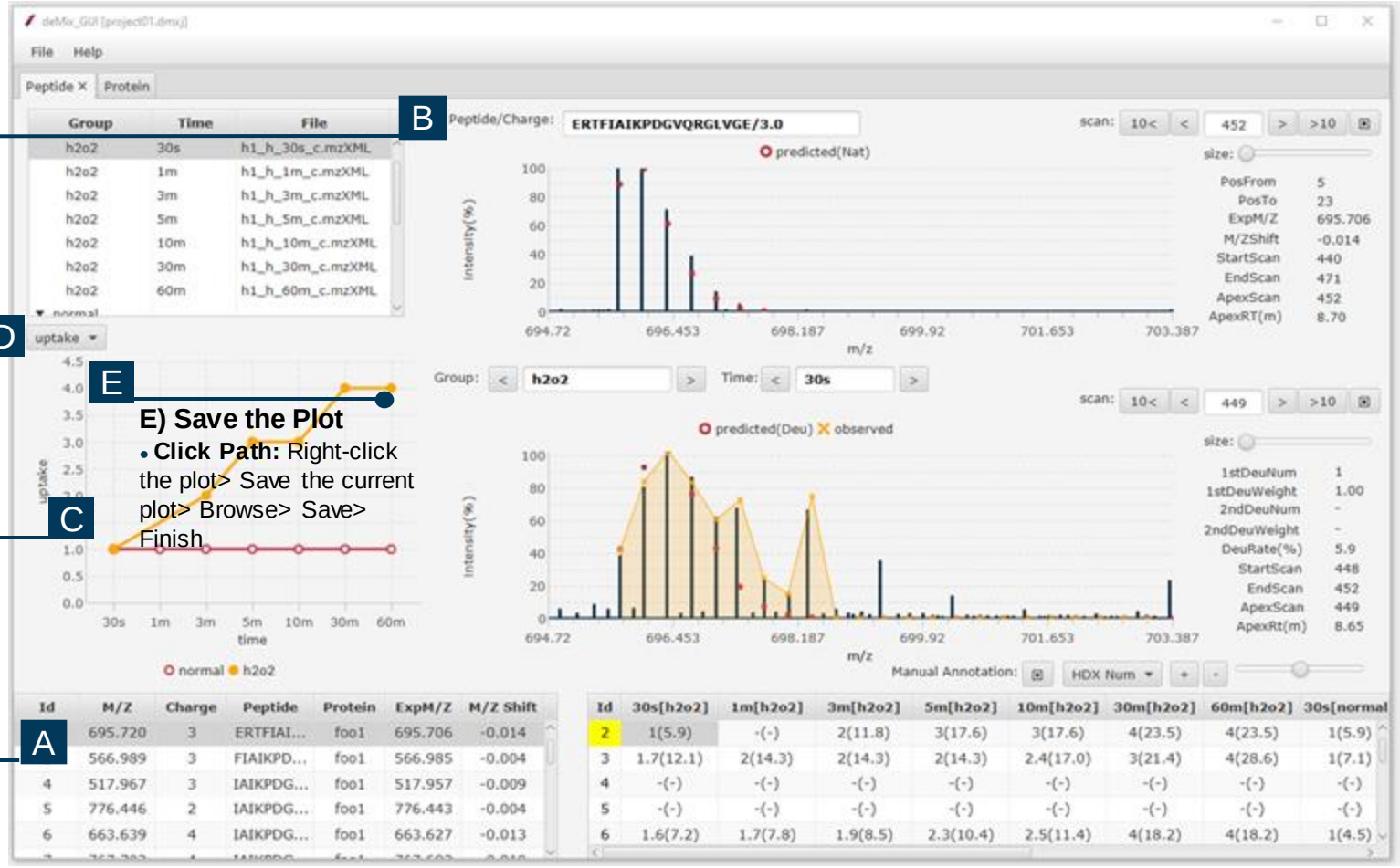
A) PeptideTable

- Contains peptide's information.

3.4. HDX Plot

3.4.2. Plot Functions

Plot Control: A → B → C



B) Current Peptide

- The selected peptide's sequence & charge appears.

D) Rate/Uptake Button

- Changes the metric of the HDX measurement.
- The metric for the y-axis is set to "uptake" by default.
- For rate(%), the range of the y-axis is set to 0 to 100 (does not vary).

E) Save the Plot

- **Click Path:** Right-click the plot> Save the current plot> Browse> Save> Finish

C) HDX Plot

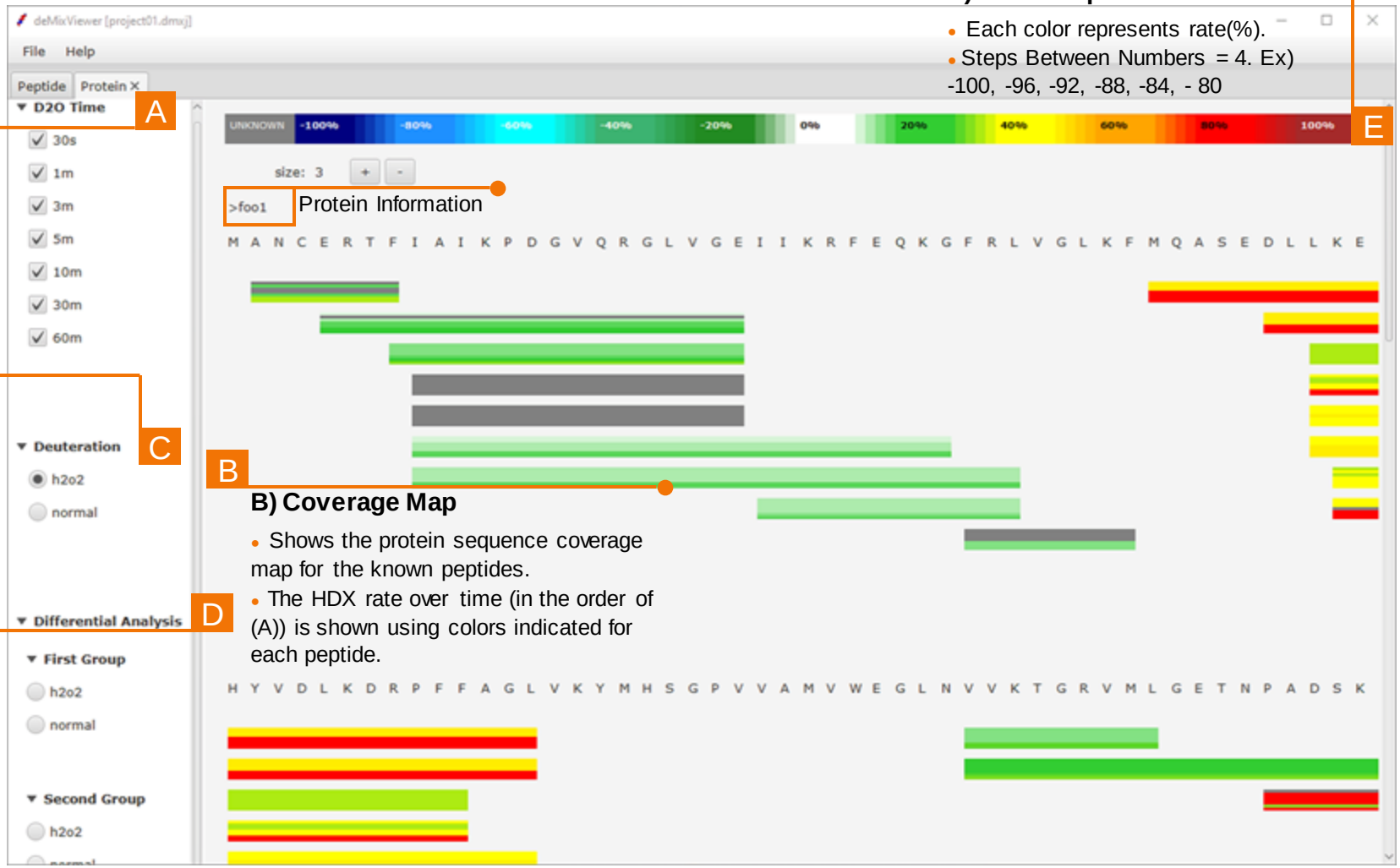
- The HDX plot appears across all the D2O labeling time.
- The range of the y-axis varies for each peptide selection.
- **Order Control**: Right-click the plot> Ascend/Descend

A) Peptide Selection

- Select a peptide by clicking a row in the table.
- Keyboard(up & down arrow keys) support available

3.5. Sequence Coverage Map

3.5.1. Coverage Map Explanations



A) D2O time

- Shows all the D2O times across the data.
- The deuteration rate at the only checked times are shown in the heat map.

C) Deuteration

- Shows all the D2O labeling groups across the data.
- The checked group's deuteration over time across the peptides is shown as a heat map.

D) Differential analysis

- Performs the differential analysis between two groups. (control = First Group, treatment = Second Group)
- The rate differences between the two groups are shown as a heat map

B) Coverage Map

- Shows the protein sequence coverage map for the known peptides.
- The HDX rate over time (in the order of (A)) is shown using colors indicated for each peptide.

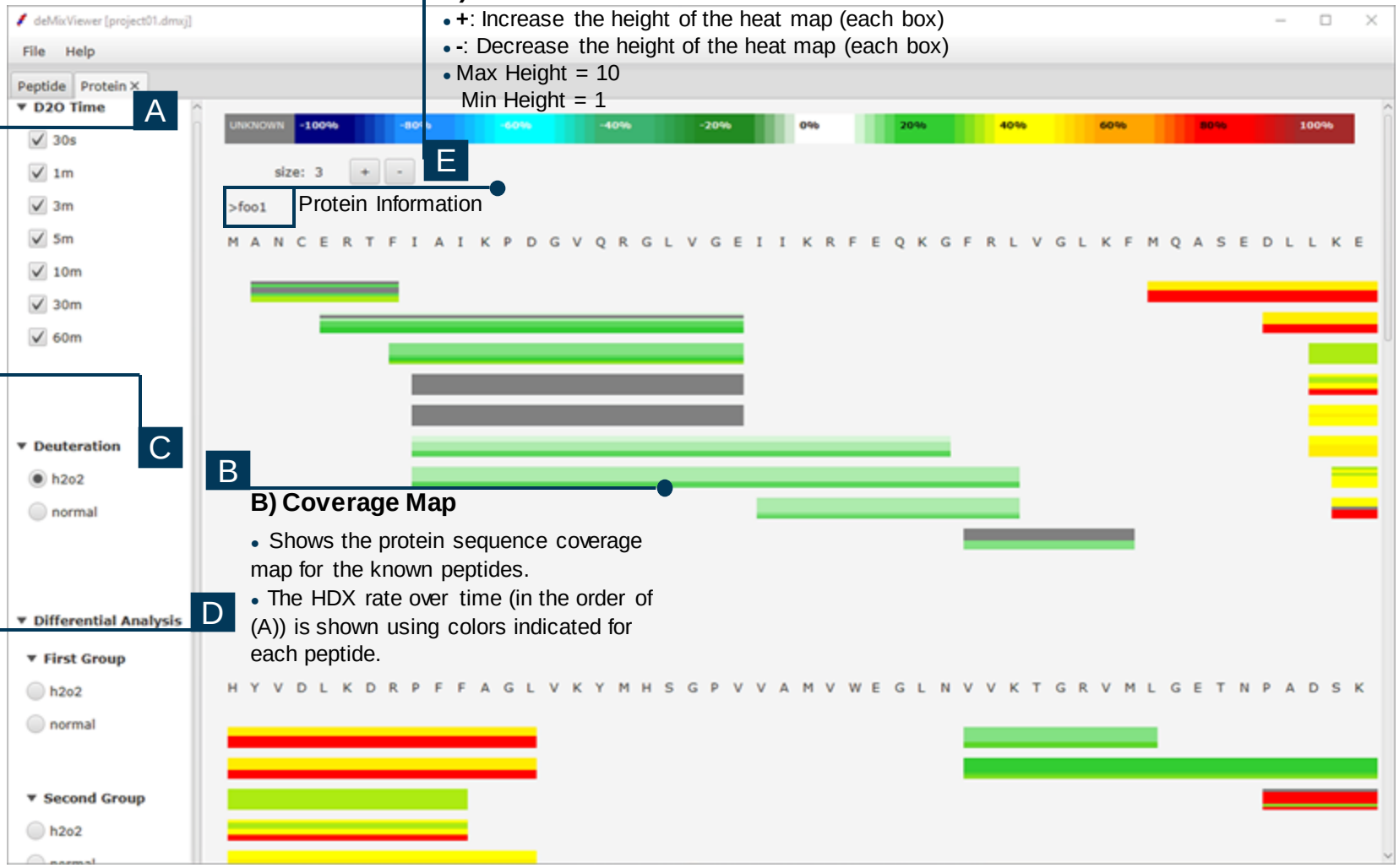
E) Color Map

- Each color represents rate(%).
- Steps Between Numbers = 4. Ex) -100, -96, -92, -88, -84, - 80

3.5. Sequence Coverage Map

3.5.1. Coverage Map Functions

Map Control: A → C/D → B



A

A) D2O time

- Select D2O labeling times to include.

C

C) Deuteration

- Select a group to draw a coverage map (= heat map).

D

D) Differential analysis

- Select control group (First Group) and treatment group (Second Group) to perform differential analysis.
- The difference calculation example: 10s : Second Group – First Group.

B

B) Coverage Map

- Shows the protein sequence coverage map for the known peptides.
- The HDX rate over time (in the order of (A)) is shown using colors indicated for each peptide.

E

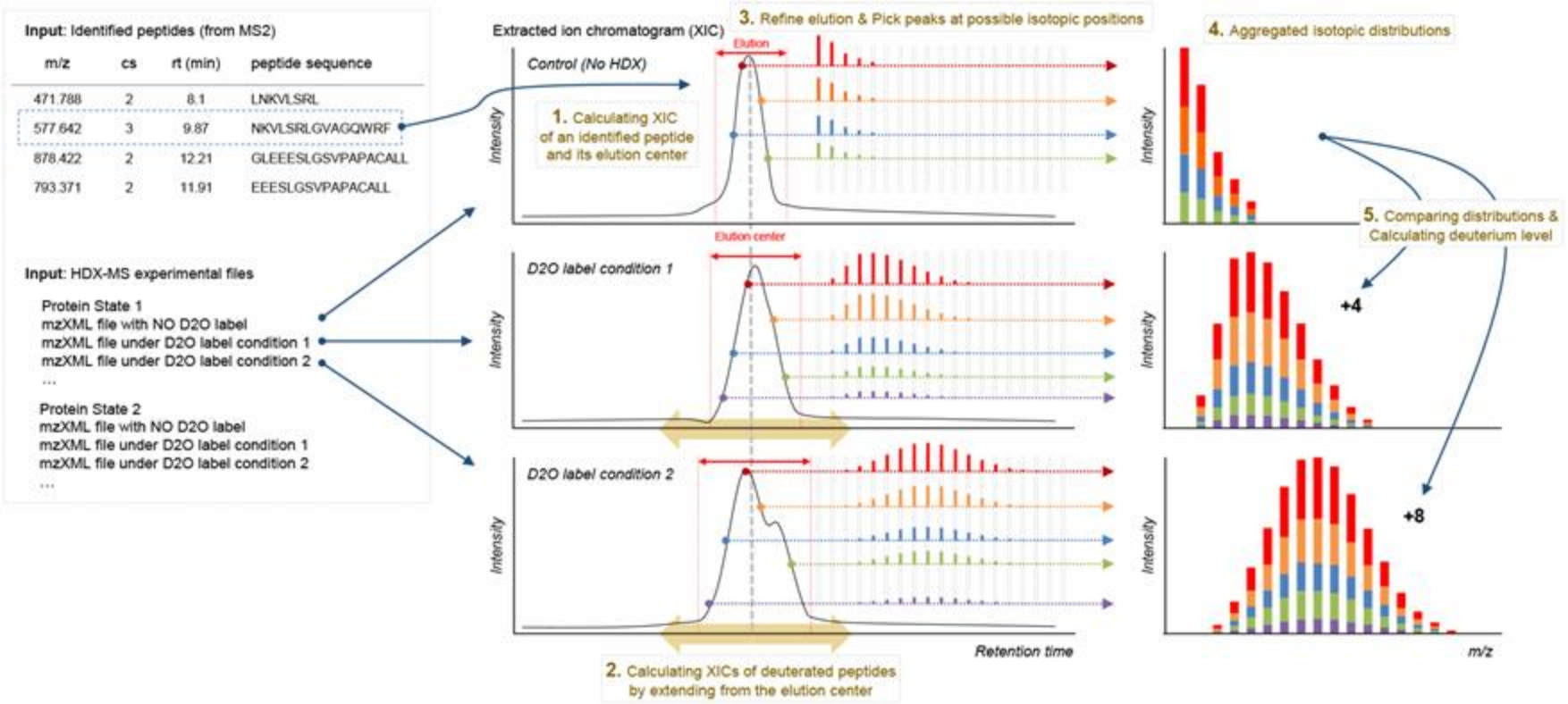
E) Size Controller

- **+**: Increase the height of the heat map (each box)
- **-**: Decrease the height of the heat map (each box)
- Max Height = 10
  Min Height = 1

Protein Information

4. Appendix

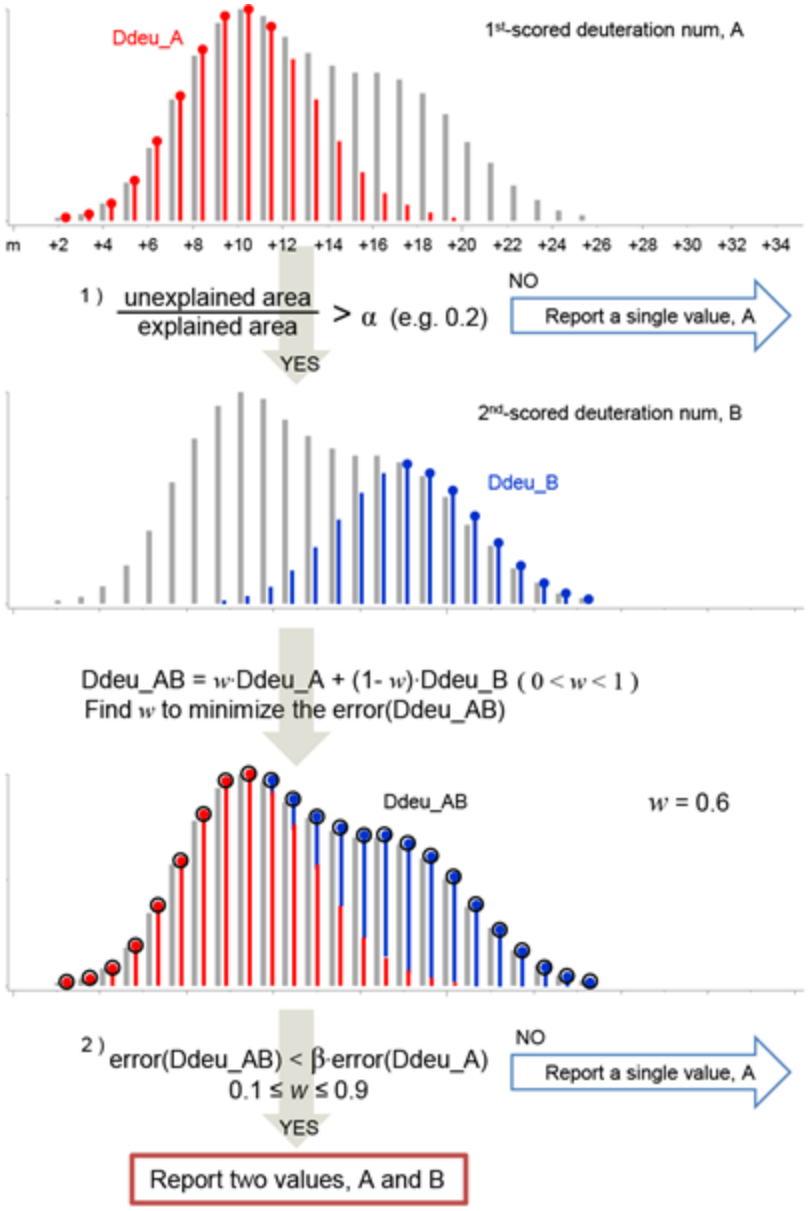
4.1. deMix Workflow



1) given peptides identified from MS2, deMix generates a theoretical isotopic distribution for each peptide and compares observed distributions in MS1 spectra (non-deuterated sample), constructing its extracted ion chromatogram (XIC).
2) based on the XIC in the non-deuterated sample, XICs of peptides in deuterated samples are constructed, where our assumption is that the related XICs across samples partially overlap or are shifted within a certain range (e.g., ±40 scans) although they may not totally overlap.
3) deMix refines each XIC and selects isotopic distribution peaks corresponding to presumably the same peptide ion within a determined elution time span.
4) deMix aggregates all detected isotopic peaks into a single isotopic distribution. The aggregated isotopic distributions are regarded more robust than individual distributions; and 5) based on the aggregated isotopic distributions, deuterium numbers are determined.

4. Appendix

4.2. Bimodal Analysis



Bimodal distribution analysis.

a) After the initial unimodal distribution analysis, it is assessed how well the observed distribution is explained.
b) If a significant portion is not explained, the next bimodal distribution analysis is performed, where weights of the two distributions are optimized.
c) After bimodal distribution analysis, it is assessed how much the error is improved and whether both distributions are all fairly abundant. Only if all criteria are satisfied, two deuterium numbers are reported.

DeuRate(%)= (1stDeuNum*1stDeuWeight+2ndDeuNum*2ndDeuWeight)*100

4. Appendix

4.3. Referencces

[1] Na, S., Lee, JJ., Joo, J.W.J. *et al.* deMix: Decoding Deuterated Distributions from Heterogeneous Protein States via HDX-MS. *Sci Rep* **9**, 3176 (2019). https://doi.org/10.1038/s41598-019-39512-8