HYDRA: a flexible software package for 'one-click' HDX-MS data analysis

Gordon W. Slysz, CJ Baker, Benjamin M. Bozsa, Anthony Dang, David C. Schriemer

Dept. Biochem. and Molecular Biology, University of Calgary, Calgary, AB T2N 1N4, Canada

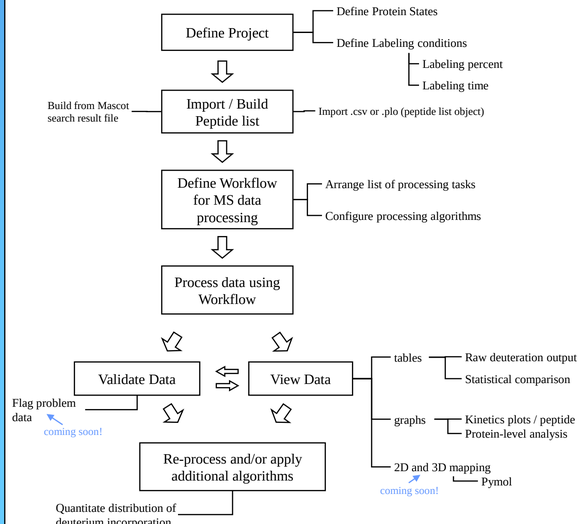
INTRODUCTION

- Hydrogen-deuterium exchange mass spectrometry (HDX-MS) is a powerful method for relaying information on protein dynamics, protein-ligand interaction and protein structure interpretation.
- The methods for collecting HDX-MS are relatively straightforward and can be achieved on a relatively short timescale. In contrast, HDX-MS data analysis and visualization, if done manually, can require many times more labor than that involved in data collection.
- Here we present a new software application that automates all major steps in HDX-MS data analysis and visualization.
- Important aspects of the software include: forward-looking flexibility to allow alternative data inputs and data analysis workflows, auto-building peptide lists using MASCOT search result output, and ability to automate MS/MS data analysis.

METHODS

- HYDRA was developed in .NET 3.0 using C# and the Windows Presentation Foundation (WPF). The software architecture allows different data inputs and data analysis scenarios to be plugged-in with minor changes to the software's code. Currently, the software supports the use of Analyst QS 1.1 for data input. Future iterations of HYDRA will allow the analysis of mzXML data.
- HYDRA workflow and performance were demonstrated using data on bottom-up HDX-MS experiments on calmodulin, performed on a QSTAR Pulsar i QqTOF instrument operated under turboionspray mode. A total of 18 runs were collected – triplicate analysis of two calmodulin states, under three labeling conditions (10%, 50%, 75% D2O labeling). See our poster ThP497 for additional details.
- Theoretical isotopic distributions were calculated using the .NET library provided by PNNL's Molecular Weight Calculator.[1]
- Linear Least Squares calculations for deuterium distribution calculations were made using the .NET library from Lutz Roeder's Mapack for .NET.[2]

WORKFLOW:



SAMPLE PROJECT: HDX-MS analysis of apo-CaM vs holo-CaM

Goal: Use HYDRA to fully automate HDX-MS analysis of apo-calmodulin (apo-CaM) vs. holo-calmodulin (holo-CaM)

Automated Analysis includes:

1) 'One-click' data processing
   - *XIC generation*
   - *XIC smoothing*
   - *XIC peak selection*
   - *MS generation*
   - *MS smoothing*
   - *Isotopic profile extraction*
   - *Theoretical isotopic distribution calculation*
   - *Centroid mass calculation*
   - *Label incorporation calculation*

2) Aggregate statistics
   - *Average SD*
   - *Student's T-test and p-value reporting for indicating differences between two protein states*

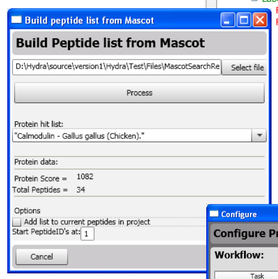
3) Tabular & graphical representation of the data (see below)

Project Stats

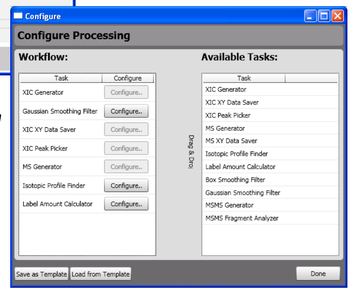
Protein States	2	apo-CaM vs holo-CaM
Labeling conditions	3	Labelled in 10%, 50%, and 75% D_2O
Replicates	3	
Peptides analyzed	21	96% Sequence coverage; see poster ThP497 for details
Total MS spectra evaluated	**378**	

PROJECT CONFIGURATION:

- *Step 1: Define project's Protein States and Labeling conditions & associate individual experiments*



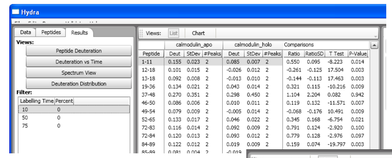
- *Step 2: Either import / merge a previous peptide list (.csv) or build a list from a Mascot search result (.csv format) as shown to the left.*



- *Step 3: Build the data processing workflow using available algorithms/tasks.*
- *Configure each task as desired*
- *New or alternate tasks are easily added to workflows (i.e. alternate smooths, baseline subtraction, etc.).*

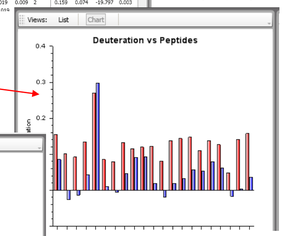
DATA VISUALIZATION

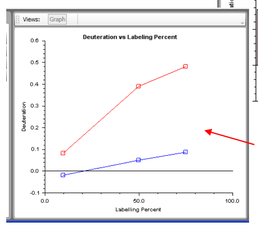
Table showing aggregate statistics for each protein state and labeling condition



- *All data (including MS data) are saved as part of the .HDX project*
- *Alternatively, results can be exported as a .csv file (i.e. for Excel)*

Bird's eye view of all peptides for two Calmodulin states (apo=red, holo=blue) and one labeling condition (10%)

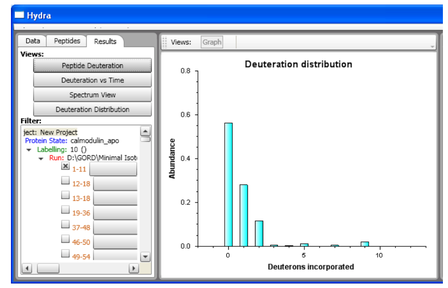




Graphical analysis at a per peptide level for two calmodulin states (apo=red; holo=blue).

- *Kinetic plots are also available (deuteration vs labeling time).*

DATA ANALYSIS: Deuterium Distribution



- *HYDRA includes deuterium distribution analysis based on the linear least squares method described by Chik et al, 2006.[3] This allows greater insight into deuterium incorporation and may help resolve HDX-MS data for peptides that exist in multiple states*

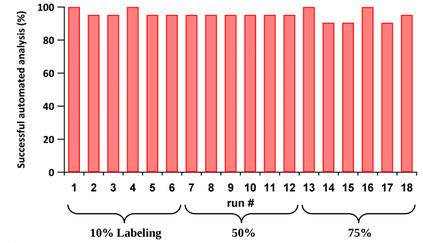
REFERENCES:

1. Molecular Weight Calculator. *http://ncrr.pnl.gov/software/*
2. Lutz Roeder's Mapack for .NET: *http://www.aisto.com/roeder/dotnet*
3. Chik JK, Graaf JLV, Schriemer DC (2006). Quantitating the statistical distribution of deuterium incorporation to extend the utility of H/D exchange MS data. *Analytical Chemistry.* 78(1): 207-214.

SOFTWARE PERFORMANCE

Success Rates

- *HYDRA was used to visually validate XIC peak selection and MS spectra from each run. The figure below shows the run-by-run success rate for HYDRA's automated data analysis. The overall success rate was 96%.*



Speed

Task	Time Required
Automated deuterium level extraction for 378 mass spectra	~ 11 minutes*
Software-assisted visualization of each spectrum for validation purposes	~ 1 hour
Statistical analysis & Deuterium distribution measurement	< 1 second

* Computer: 2.0 GHz AMD cpu with 1G of RAM

Flexibility

- Alternate data analysis workflows can be easily added to the software architecture.
- MS/MS data analysis workflows have been added and will enable the large-scale evaluation of MS/MS data for HDX-MS studies.

MS/MS Capability:

SUMMARY & KEY FEATURES

- HYDRA provides an automated means of extracting deuterium incorporation data including distribution information.
- Visual tools provide a quick way of validating results.
- HYDRA uses step-by-step wizards to walk users through project configuration, and thus requires little training.
- Flexibility shown by the ability to adapt or add new data processing workflows.

SOFTWARE AVAILABILITY

- Requests for HYDRA should be directed to Dr. David Schriemer (dschriem@ucalgary.ca)

FUTURE WORK:

- Automatic interference detection and correction.
- 2D and 3D visualization
  - Interactive displays
- mzData/mzXML as data inputs

ACKNOWLEDGEMENTS

Research was funded by NSERC & CRC