

HaDeX: an R package and web-server for analysis of data from hydrogen-deuterium exchange mass spectrometry experiments

Weronika Puchała^{1,*}, Michał Burdukiewicz², Michał Kistowski¹, Katarzyna Dabrowa-Dabrowska¹, Aleksandra E. Badaczewska-Dawid³, Dominik Cysewski¹, Michał Dadlez¹
*puchala.weronika@gmail.com

¹Institute of Biochemistry and Biophysics Polish Academy of Sciences, Poland, ²Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland, ³Faculty of Chemistry, Biological and Chemical Research Center, University of Warsaw, Poland.

Introduction

Hydrogen-deuterium mass spectrometry (HDX-MS) is a staple tool for monitoring dynamics and interactions of proteins. Due to the sheer size of the HDX-MS results, the data analysis require a dedicated software suite. However, the majority of existing tools provides only point-and-click interfaces to black-box models or does not offer a complete analytic workflow. We propose **HaDeX**, a novel tool for processing, analysis and visualisation of HDX-MS experiments. **HaDeX** covers the whole analytic process, including quality control, ISO-based uncertainty and publication-quality figures. The reproducibility of the whole procedure is ensured with advanced reporting functions. It is important to us that our analytic methodology is transparent and understandable for the users so it is consulted with international experts and discussed in-depth in the package vignette.

Coverage of the sequence

The **HaDeX** web server relies only on a single data format: DynamiX datafile. A user can customize each step of the analysis with his own input parameters. We start our analysis with summary of protein coverage (as shown below). It is provided in two forms, histogram and graphic comparison of positions. Also the reconstruction of the sequence is available - based on experimental data.

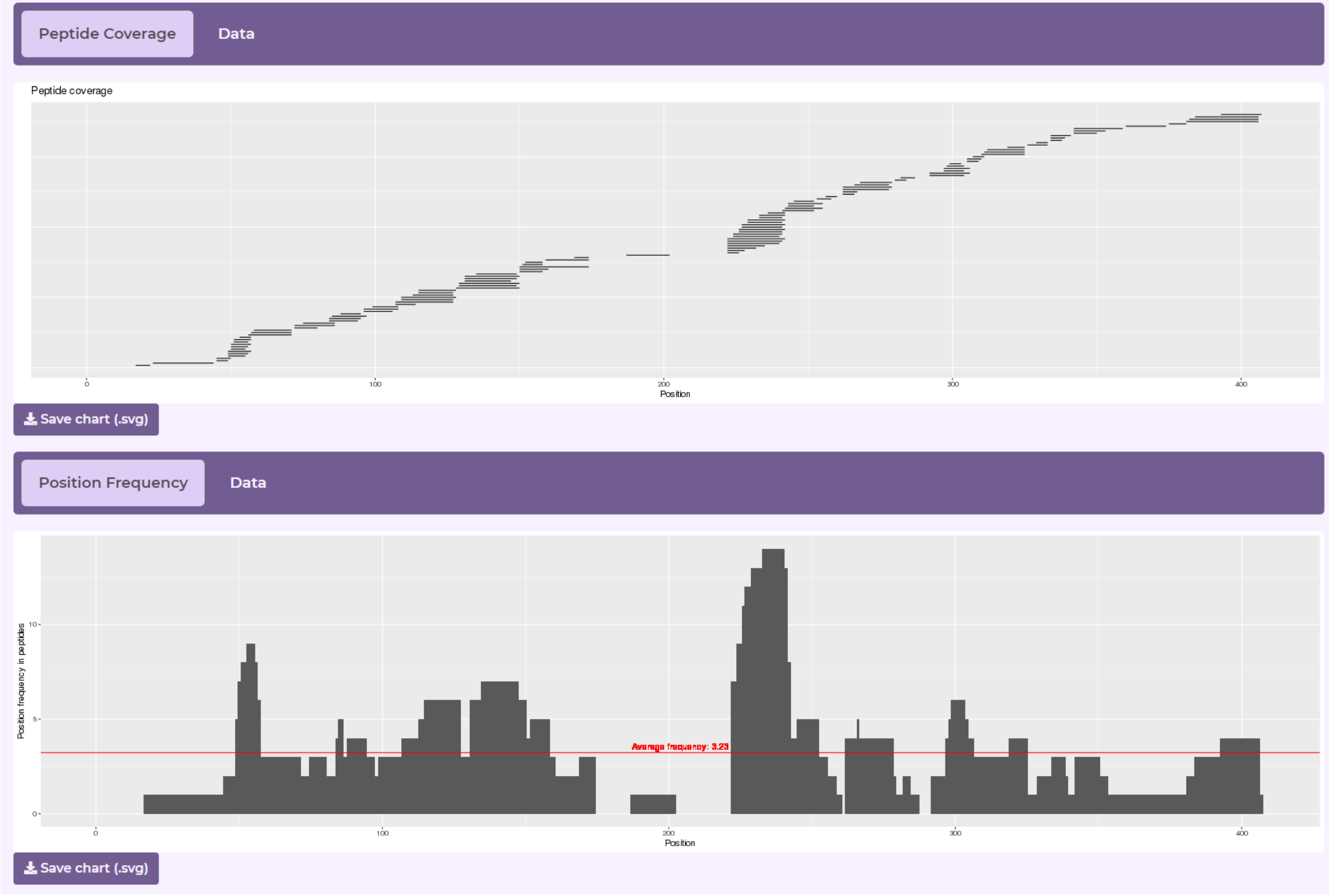


Figure 1: Sequence coverage analysis (**HaDeX** web server).

Differential plot

An important aspect of analysing data is a comparison of peptide deuteration in different states. A relative deuteration level used by **HaDeX** is defined by the equation 1.

$$D = \frac{D_t - D_0}{D_{100} - D_0} \quad (1)$$

The equation 1 is a function of three variables (each with its own uncertainty), so there is need to use the Law of Propagation of Uncertainty, defined in equation 2:

$$u_c(y) = \sqrt{\sum_k \left[\frac{\partial y}{\partial x_k} u(x_k) \right]^2} \quad (2)$$

Applying equation 2 to equation 1, we get deuteration level uncertainty described by equation 3:

$$u_c(D) = \sqrt{\left[\frac{1}{D_{100} - D_0} u(D_t) \right]^2 + \left[\frac{D_t - D_{100}}{(D_{100} - D_0)^2} u(D_0) \right]^2 + \left[\frac{D_0 - D_t}{(D_{100} - D_0)^2} u(D_{100}) \right]^2} \quad (3)$$

The levels of deuteration along with uncertainty intervals are shown on comparison plot 2. **HaDeX** provide both experimental and theoretical levels of deuteration in two variants: with relative or absolute values.

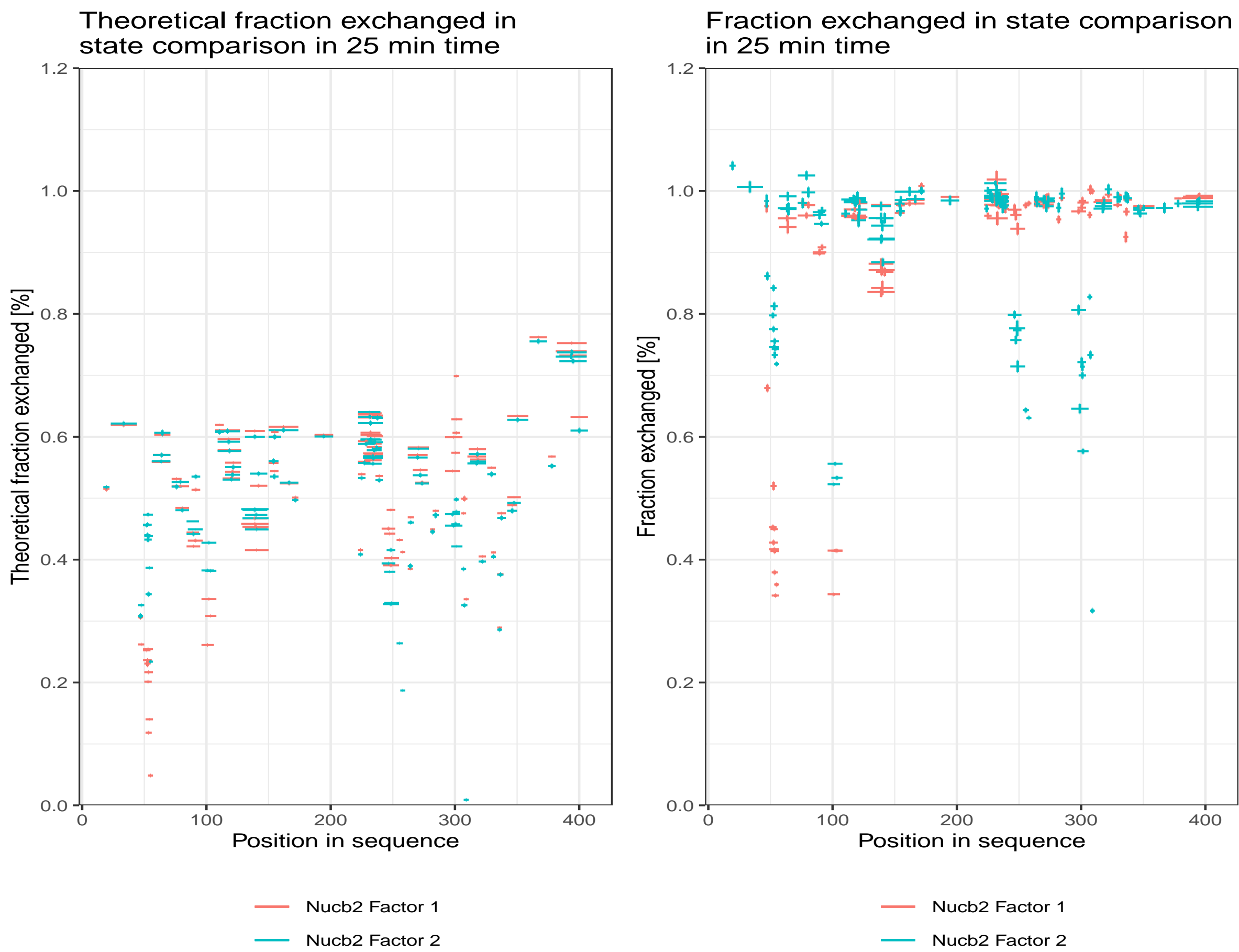


Figure 2: Comparison plots - comparing experimental and theoretical calculations, based on example data.

Availability

HaDeX is available as a web server: <http://mslab-ibb.pl/shiny/HaDeX/>



as the R package: <https://cran.r-project.org/web/packages/HaDeX/index.html>

and as a standalone software: <https://sourceforge.net/projects/hadex/>

Woods plot format

Woods plot format shows difference between results of two different states as described by equation 4.

$$diff = D_1 - D_2 \quad (4)$$

Equation 4 is a function of two variables, so applying 2 we get uncertainty described by equation 5.

$$u_c(diff) = \sqrt{u(D_1)^2 + u(D_2)^2} \quad (5)$$

Described calculations lead to results shown on Woods plot format - figure 3 below.

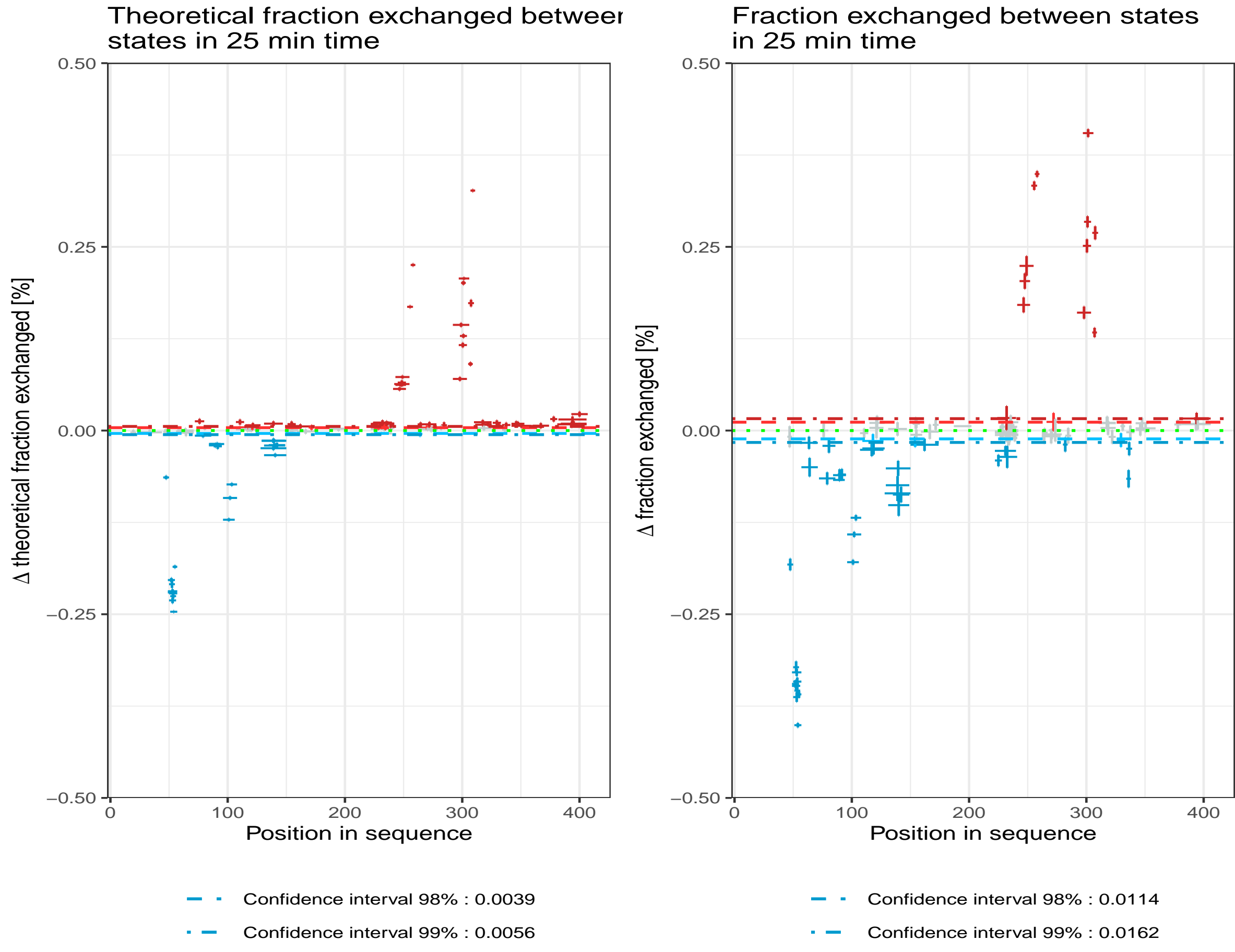


Figure 3: Plots in Woods format - comparing experimental and theoretical calculations, based on example data.

Quality control

HaDeX provides additional tools for assesment of experiments. For example, the quality control function shows how the mean uncertainty per peptide changes with time points of an experiment.

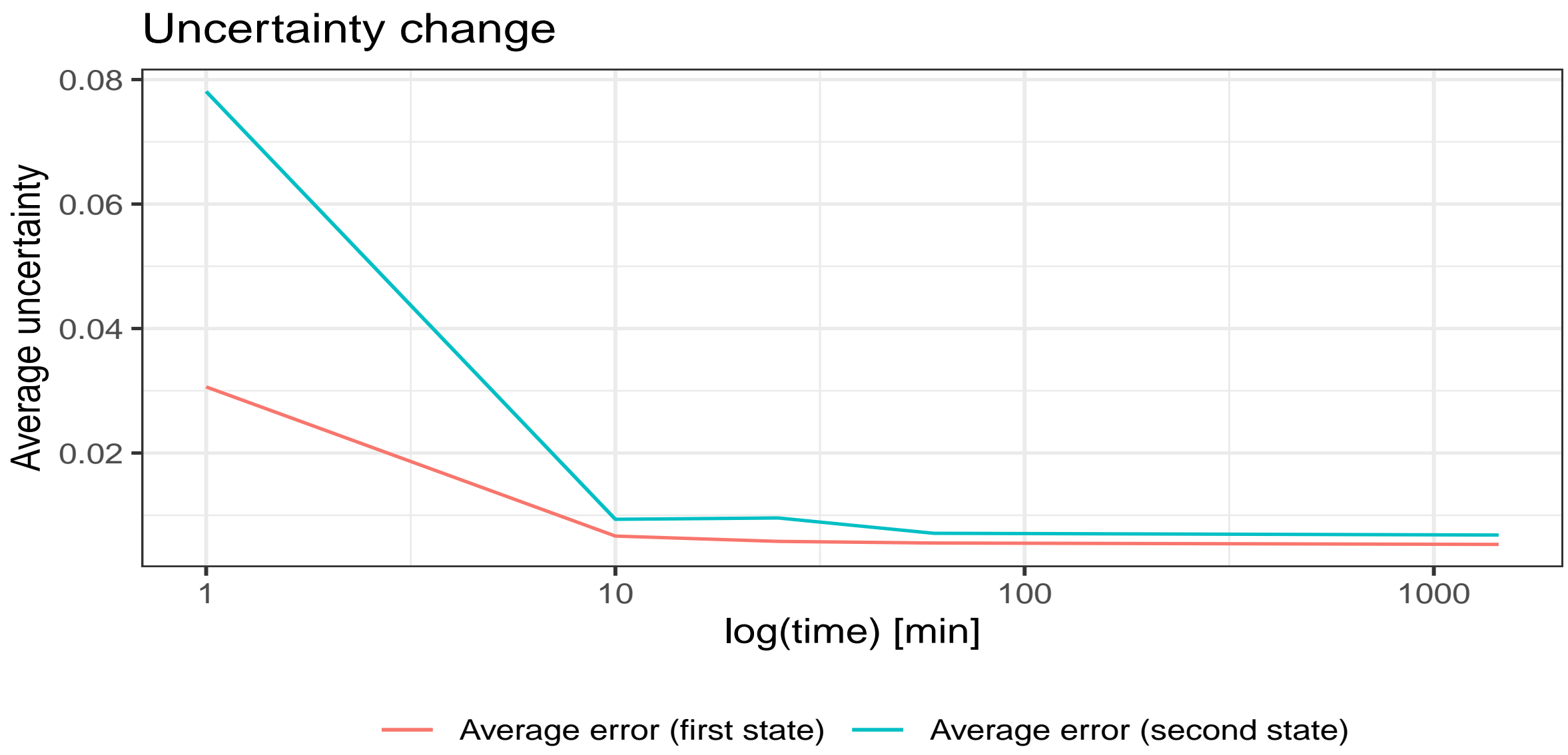


Figure 4: Plot illustrating uncertainty change depending on used out parameter

Report function

HaDeX provides report generation - with parameters chosen by user so is fully personalizable. Additionally, calculated data for plots can be downloaded not only in report form but also in csv/pdf file in panel next to every plot. On figure 5 are shown available elements of the report - that can be chosen by user and limited strickly for user needs.

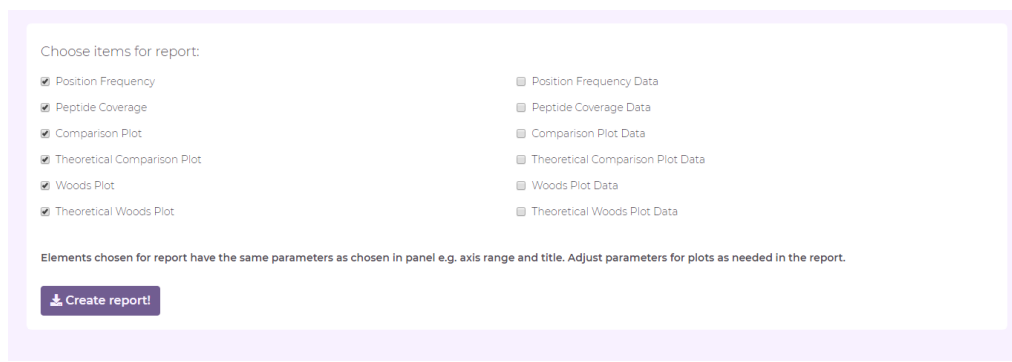


Figure 5: Screenshot of **HaDeX** - reporting page

Acknowledgements

The project is carried out within the TEAM TECH programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.(TEAM TECH CORE FACILITY/2016-2/2) Mass Spectrometry of Biopharmaceuticals-improved methodologies for qualitative, quantitative and structural characterization of drugs, proteinaceous drug targets and diagnostic molecules"



European Funds
Smart Growth



Foundation for
Polish Science

European Union
European Regional
Development Fund



Bibliography

- 1 Start2Fold: a database of hydrogen/deuterium exchange data on protein folding and stability. Pancsa R, Varadi M, Tompa P, Vranken WF. Nucleic Acids Res. 2016 Jan 4;44(D1):D429-34. doi: 10.1093/nar/gkv1185. Epub 2015 Nov 17
- 2 HDX workbench: software for the analysis of H/D exchange MS data. Pascal BD, Willis S, Lauer JL, Landgraf RR, West GM, Marciano D, Novick S, Goswami D, Chalmers MJ, Griffin PR. J Am Soc Mass Spectrom. 2012 Sep;23(9):1512-21. doi: 10.1007/s13361-012-0419-6. Epub 2012 Jun 13.
- 3 MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. Hourdel V, Volant S, O'Brien DP, Chenal A, Chamot-Rooke J, Dillies MA, Brier S. Bioinformatics. 2016 Nov 15;32(22):3413-3419. Epub 2016 Jul 13
- 4 Deuterios: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry, Andy M C Lau, Zainab Ahdash, Chloe Martens, Argyris Politis, , Bioinformatics, btz022, <https://doi.org/10.1093/bioinformatics/btz022>
- 5 ISO, Guide to the Expression of Uncertainty in Measurement (International Organization for Standardization, Geneva, Switzerland, 1993)