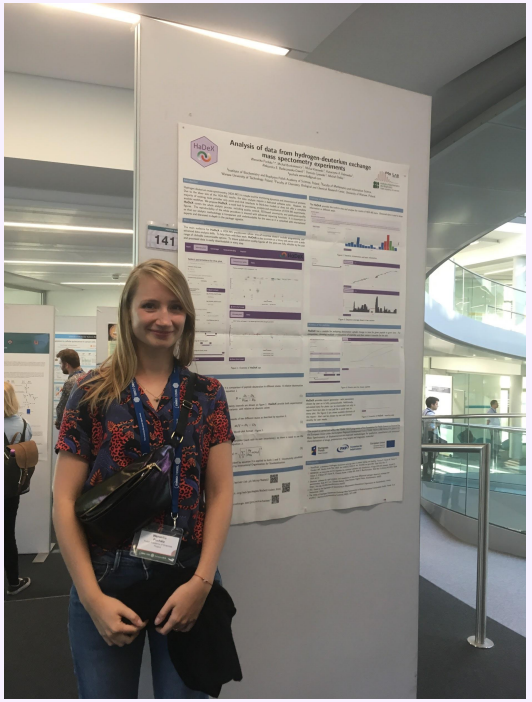
HaDeX - journey in the HDX-MS data

Weronika Puchała

About

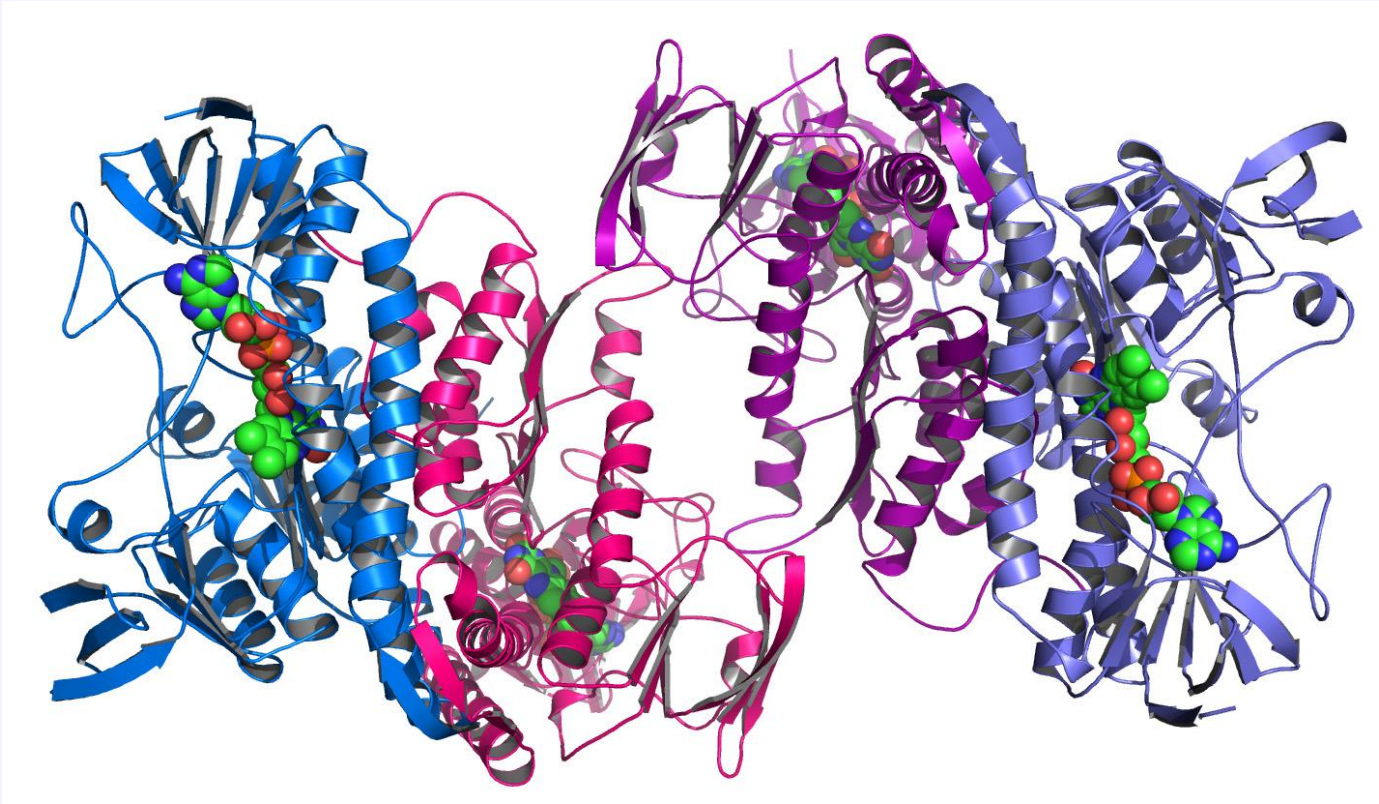
- PhD student in Biophysics
- commercial background
- HDX-MS data analyst
- in the project from the beginning



Ms Lab IBB PAN

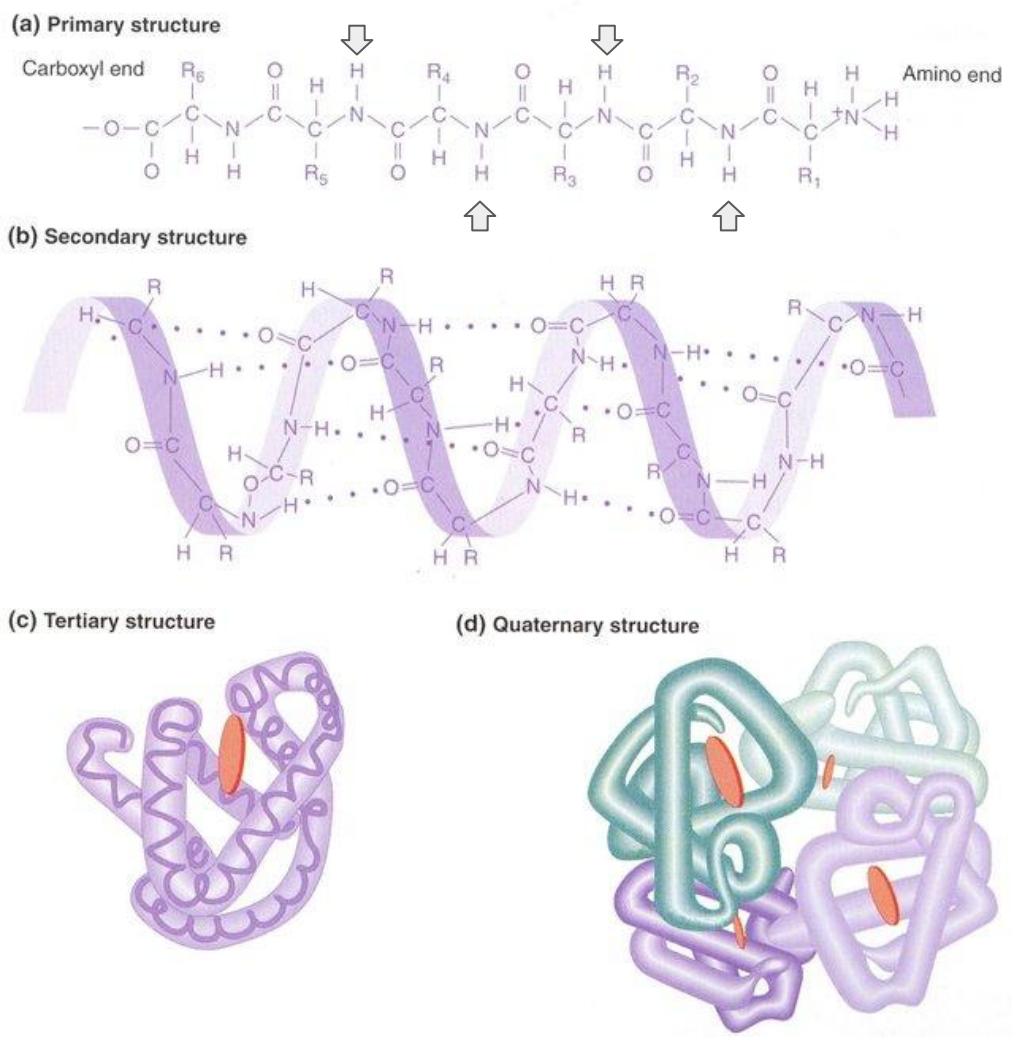
Institute of Biochemistry and Biophysics
Polish Academy of Sciences

Proteins!



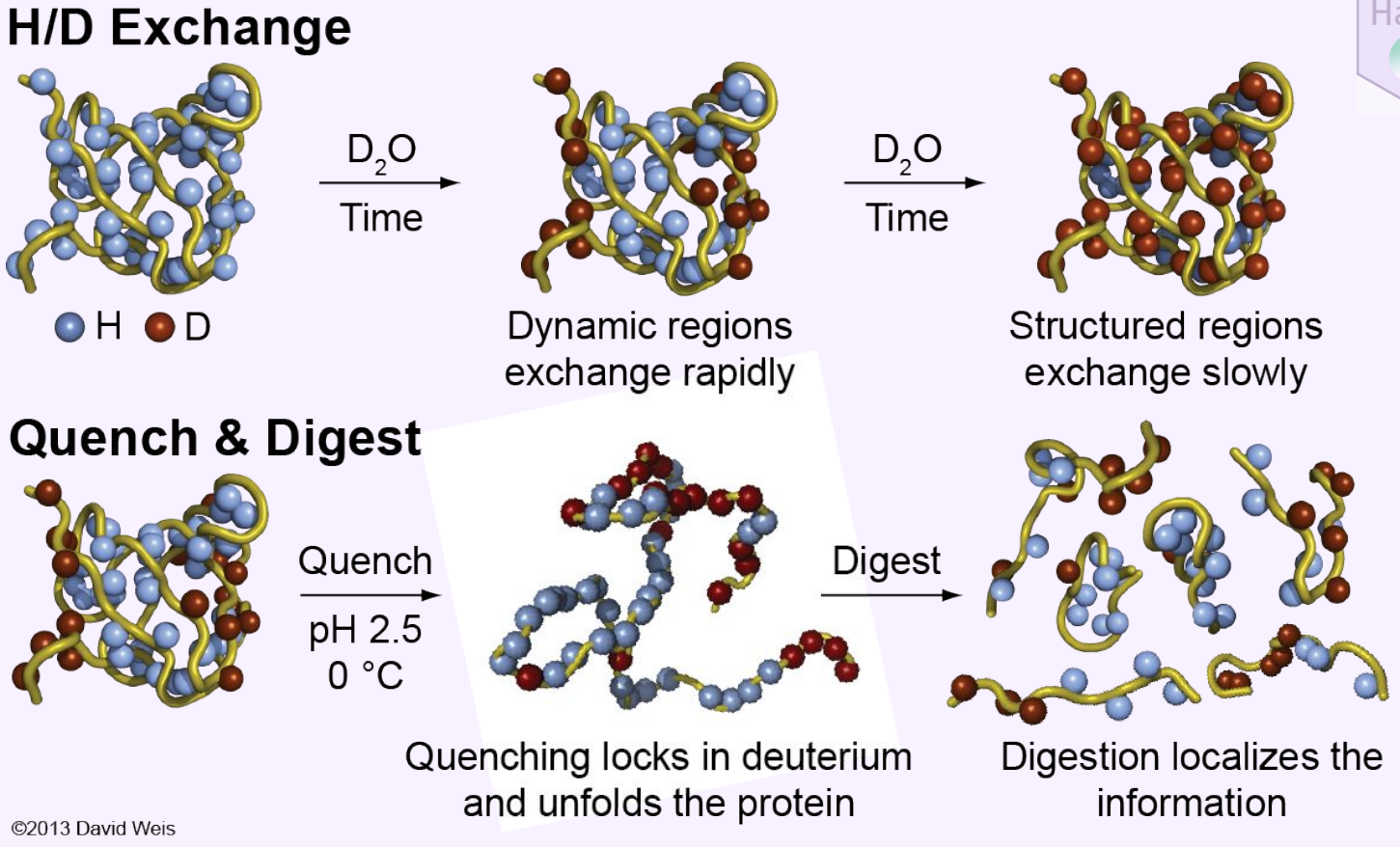
By Matt Howard - Argonne's Midwest Center for Structural Genomics deposits 1,000th protein structure, CC BY-SA 2.0, https://commons.wikimedia.org/w/index.php?curid=9752165

But...



Schmid, Silvan. (2020). Immobilization of DNA and protein to polymerized SU-8 photoresist investigated with microarray assays.

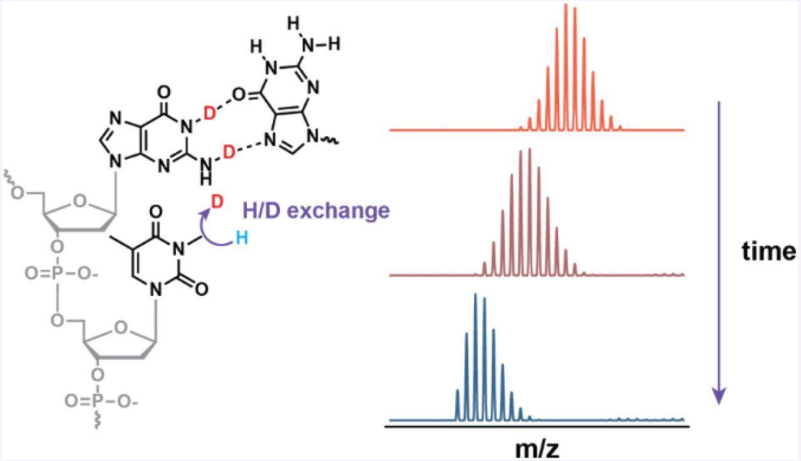
HDX



+ back-exchange!

H/D Exchange. *Weis Research Group* https://weisgroup.ku.edu/hd-exchange (2013).

MS



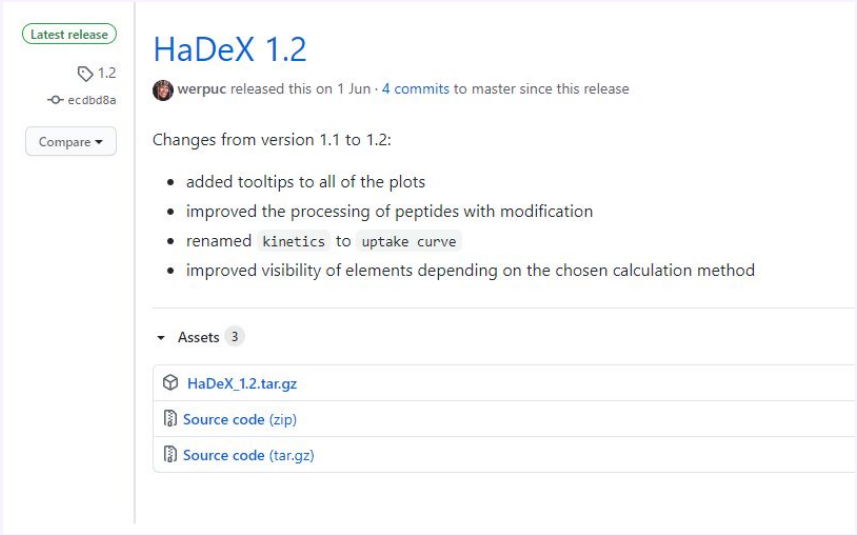
Largy, E. & Gabelica, V. Native Hydrogen/Deuterium Exchange Mass Spectrometry of Structured DNA Oligonucleotides. *bioRxiv* 848598 (2019) doi:10.1101/848598.

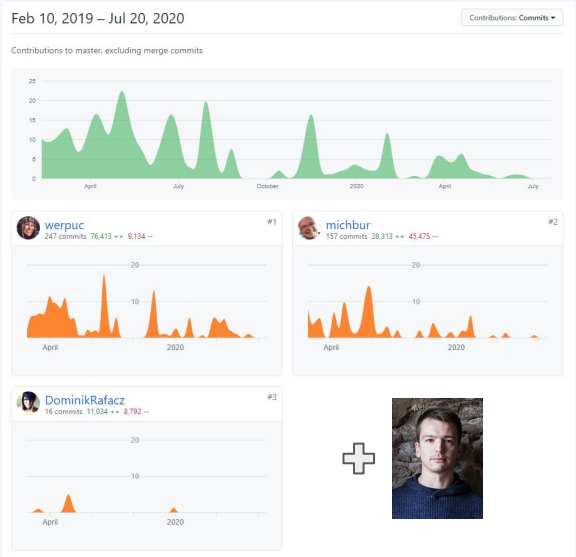
The project and its difficulties

- new knowledge and context changes the workflow
- different fields of expertise
- terms changing the meaning
- local protocoles
- advanced but narrow specialization

This Is Fine. *Know Your Meme*
https://knowyourmeme.com/memes/this-is-fine.

Currently, in the project...

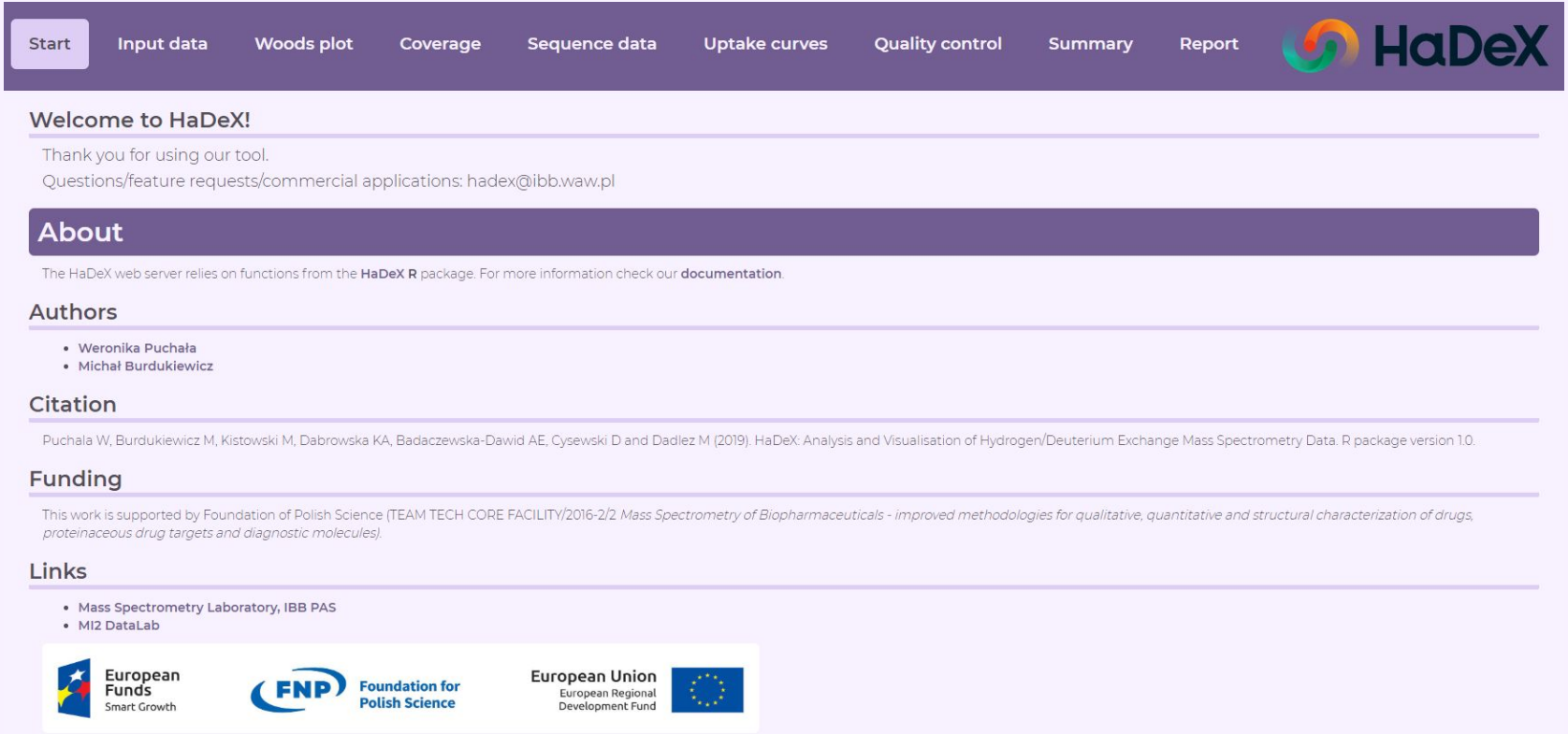




Our data

	Protein	Start	End	Sequence	Modification	Fragment	MaxUptake	MHP	State	Exposure	File	z	RT	Inten	Center
1:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.000	KD_160527_CD160_sekw_05	1	3.232479	6592	1591.2584
2:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.000	KD_160527_CD160_sekw_05	2	3.238079	394066	796.3552
3:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.000	KD_160527_CD160_sekw_05	3	3.238759	173526	531.2633
4:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.001	KD_160527_CD160_IN_01	2	3.258598	232221	796.3634
5:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.001	KD_160527_CD160_IN_01	3	3.256844	110675	531.2849
6:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.167	KD_160530_CD160_10s_01	2	3.262774	99894	800.3610
7:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.167	KD_160530_CD160_10s_02	2	3.264558	117541	800.3852
8:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.167	KD_160530_CD160_10s_03	2	3.263795	90562	800.3682
9:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	0.167	KD_160530_CD160_10s_04	2	3.273656	66131	800.4242
10:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1.000	KD_160530_CD160_1min_01	2	3.262300	109301	800.7878
11:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1.000	KD_160530_CD160_1min_02	2	3.255346	112491	800.6889
12:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1.000	KD_160530_CD160_1min_03	2	3.254521	119940	800.8194
13:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1.000	KD_160530_CD160_1min_04	2	3.261150	105349	800.7712
14:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	5.000	KD_160530_CD160_5min_02	2	3.252980	139481	800.9808
15:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	5.000	KD_160530_CD160_5min_03	2	3.258085	108544	800.9552
16:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	5.000	KD_160530_CD160_5min_01	2	3.223357	329153	800.9635
17:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	5.000	KD_160530_CD160_5min_04	2	3.262770	68601	800.8628
18:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	25.000	KD_160531_CD160_25min_01	2	3.262472	116561	801.2475
19:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	25.000	KD_160531_CD160_25min_02	2	3.269511	62188	801.1659
20:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	25.000	KD_160531_CD160_25min_03	2	3.249441	95892	801.2836
21:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	25.000	KD_160531_CD160_25min_04	2	3.256427	92053	801.2446
22:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	120.000	KD_160531_CD160_2h_01	2	3.247047	140394	801.4678
23:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	120.000	KD_160531_CD160_2h_02	2	3.249840	115181	801.3489
24:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	120.000	KD_160531_CD160_2h_03	2	3.244796	147080	801.3217
25:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	120.000	KD_160531_CD160_2h_04	2	3.254010	143965	801.2839
26:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1440.000	KD_160601_CD160_24h_01	2	3.275502	96743	801.6348
27:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1440.000	KD_160601_CD160_24h_02	2	3.270558	100373	801.5983
28:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1440.000	KD_160601_CD160_24h_03	2	3.262421	82944	801.5474
29:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160	1440.000	KD_160601_CD160_24h_04	2	3.268210	86853	801.5782
30:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.000	KD_160527_CD160_sekw_05	1	3.232479	6592	1591.2584
31:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.000	KD_160527_CD160_sekw_05	2	3.238079	394066	796.3552
32:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.000	KD_160527_CD160_sekw_05	3	3.238759	173526	531.2633
33:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.001	KD_160527_CD160_IN_01	2	3.258598	232221	796.3634
34:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.001	KD_160527_CD160_IN_01	3	3.256844	110675	531.2849
35:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.167	KD_160531_HVEM_CD160_10s_01	2	3.251068	147694	800.2299
36:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.167	KD_160531_HVEM_CD160_10s_02	2	3.242326	178169	800.2436
37:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.167	KD_160531_HVEM_CD160_10s_03	2	3.254681	141814	800.2545
38:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	0.167	KD_160531_HVEM_CD160_10s_04	2	3.241787	194042	800.1728
39:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	1.000	KD_160531_HVEM_CD160_1min_01	2	3.250153	150273	800.6576
40:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	1.000	KD_160531_HVEM_CD160_1min_02	2	3.241649	190512	800.6585
41:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	1.000	KD_160531_HVEM_CD160_1min_03	2	3.249339	183445	800.4920
42:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	1.000	KD_160531_HVEM_CD160_1min_04	2	3.255107	100346	800.5765
43:	db_CD160	1	15	INITSSASQEGTRLN	NA	NA	14	1590.8082	CD160_HVEM	5.000	KD_160531_HVEM_CD160_5min_01	2	3.254450	121446	800.6951

Welcome to HaDeXversum!



Upload your file. Otherwise, you will see the example data.
Currently HaDeX is limited to `cluster` files from DynamX 3.0 or 2.0 and `tables` file from HDeXaminer
Accepted file extensions: .csv, .xls, .xlsx.
The supplied file should contain at least two repetitions of the measurement for the uncertainty to be calculated.
If the supplied file contains modified peptides, maximal exchange control cannot be applied.

Choose file:

Browse... CK2_all results table.cs

Upload complete

File status:

Supplied file is valid. Detected data source: HDeXaminer. User action needed below!

Please be aware that loading data (including example file) may take a while. Be patient.
For the program to behave correctly, please make sure supplied file fulfills following requirements:

Show requirements

Settings

Values chosen here are propagated into all of the tables for coherent results.

Choose protein:

db_CD160

Choose maximal exchange control for chosen protein:

db_CD160 | CD160 | 1440

Choose D20 concentration [%]:

100

Move sequence start:

1

Correct sequence length:

132

Apply changes in `Input Data` tab.

File from HDeXaminer detected!

Some of the information from the data file requires your confirmation.

For the additional information on how the data from HDeXaminer is processed, check the requirements above. Keep in mind that the MHP value is generated based on the peptide sequence and therefore, may differ from actual value in case of the modifications.

FD timepoint [min]:

1440

Protein name:

Protein State 1

States names:

Protein State 1

Accepted confidence values:

- [x] High
- [x] Medium
- [ ] Low

Apply changes to continue

The calculated values of MPH might slightly differ based on data used and its precision.

Used amino mass data

Protein name

db_CD160

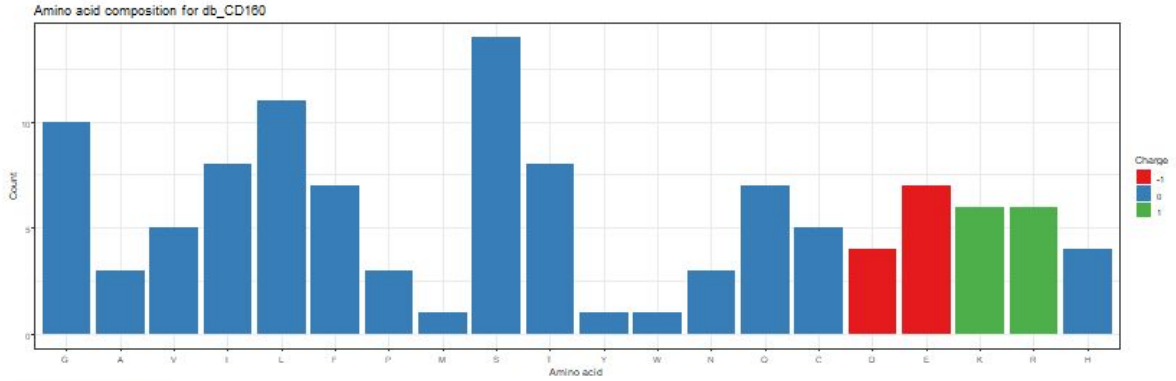
Reconstructed sequence

INITSSASQEGTRLNLICTVWHKKEEAEGFVVFLCKDRSGDCSPETSLKQLRLKRDPGIDGVGEISSQLMFTISQVTPLHSGTYQCCARSQKSGIRLQGHFFSILFxxxxxxxxxxxxxxxxxxFSHNEGTL

Name	Value
Length	132
Coverage	86.36%
Cys	5

Hydro-

- [x] Hydrophilic
- [x] Hydrophobic



Save chart (.svg)

Source: Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol 157, 105–132.

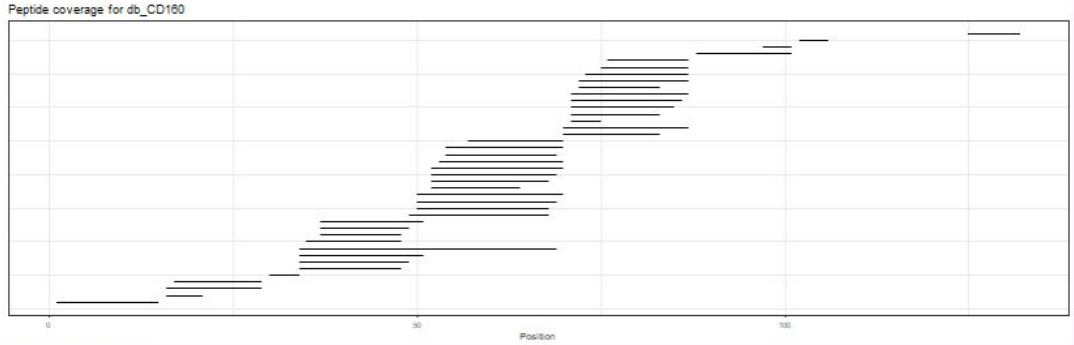
Choose state:

- CD160
- CD160_HVEM

Choose range:

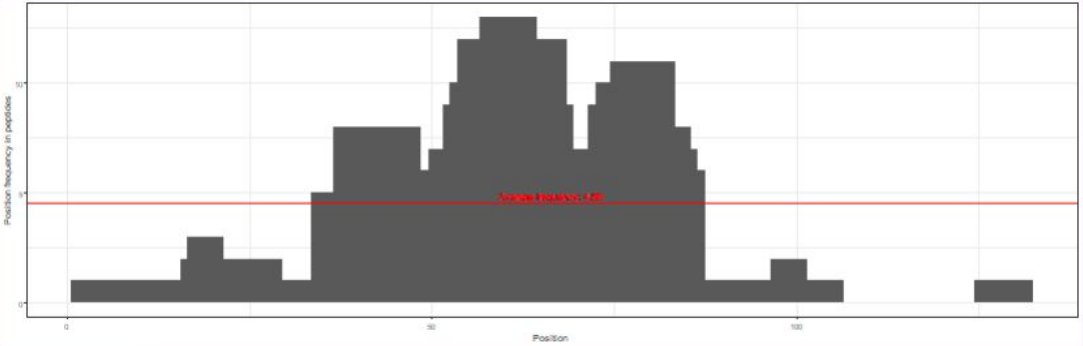
1 132

Peptide Coverage | Data

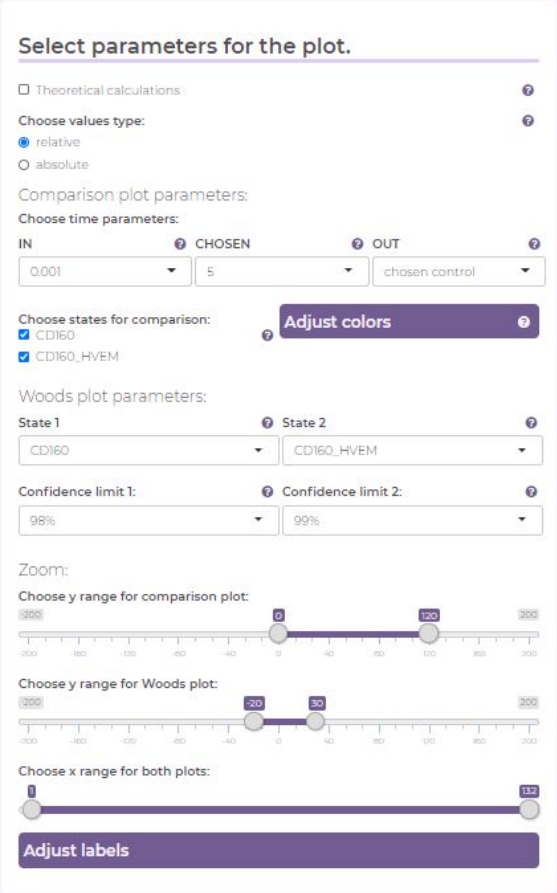


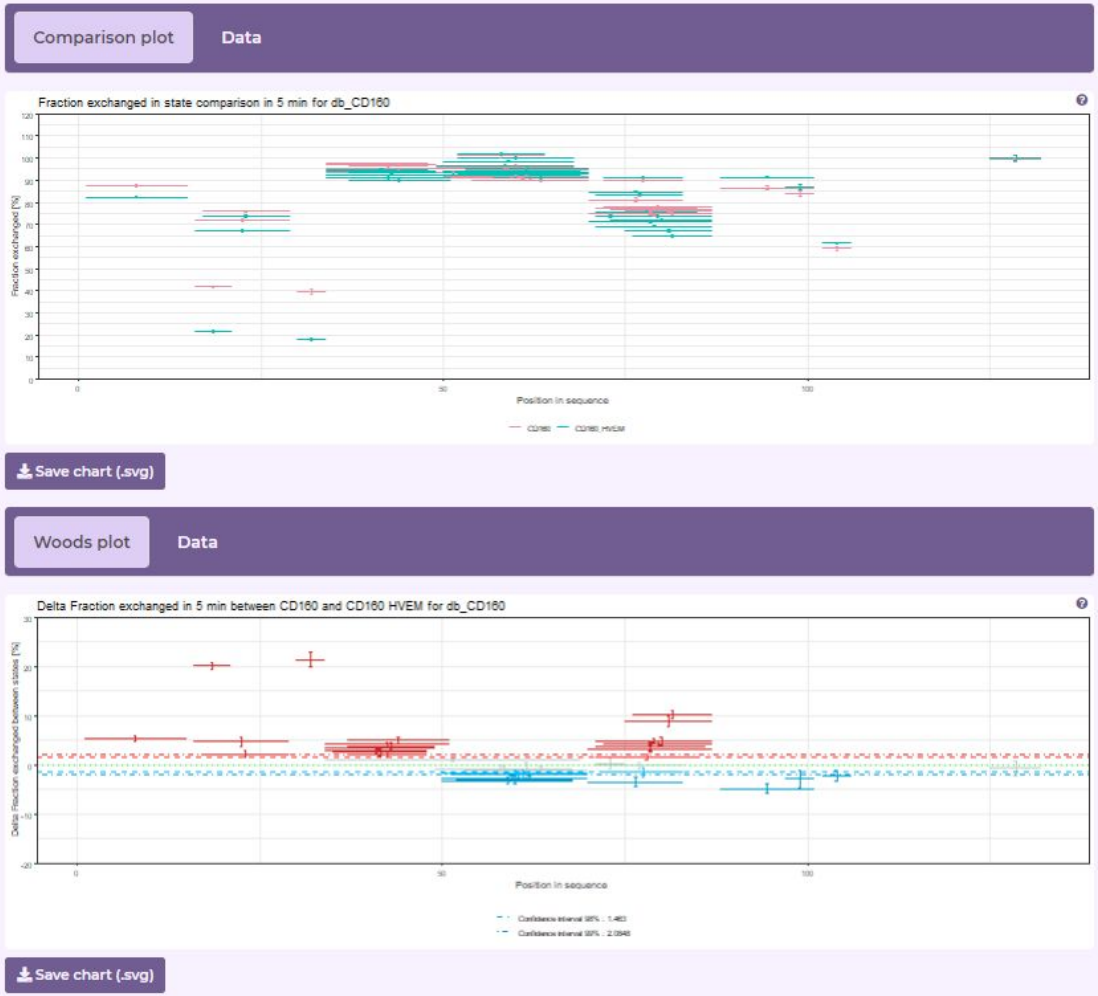
Save chart (.svg)

Position Frequency | Data



Save chart (.svg)


Select parameters for the plot.
Theoretical calculations
Choose values type:
relative
absolute
Comparison plot parameters:
Choose time parameters:
IN
CHOSEN
OUT
0.001
5
chosen control
Choose states for comparison:
CD160
CD160_HVEM
Adjust colors
Woods plot parameters:
State 1
State 2
CD160
CD160_HVEM
Confidence limit 1:
Confidence limit 2:
98%
99%
Zoom:
Choose y range for comparison plot:
Choose y range for Woods plot:
Choose x range for both plots:
Adjust labels


Comparison plot
Data
Fraction exchanged in state comparison in 5 min for db_CD160
Fraction exchanged [%]
Position in sequence
Save chart (.svg)
Woods plot
Data
Delta Fraction exchanged in 5 min between CD160 and CD160 HVEM for db_CD160
Delta Fraction exchanged between states [%]
Position in sequence
Save chart (.svg)

Select parameters for the plot.

☐ Theoretical calculations

Choose values type:

◉ relative

○ absolute

Choose time parameters:

IN: 0.001

OUT: chosen control

Choose peptide:

Sequence	State	Start	End
INITSSASQEGTRLN	CD160	1	15
INITSSASQEGTRLN	CD160_HVEM	1	15
LICTVW	CD160	16	21
LICTVW	CD160_HVEM	16	21
LICTVWHKKEEAEG	CD160	16	29
LICTVWHKKEEAEG	CD160_HVEM	16	29
ICTVWHKKEEAEG	CD160	17	29
ICTVWHKKEEAEG	CD160_HVEM	17	29
FVVFL	CD160	30	34
FVVFL	CD160_HVEM	30	34
All	All	All	All

Showing 1 to 10 of 82 entries

Previous 1 2 3 4 5 … 9 Next

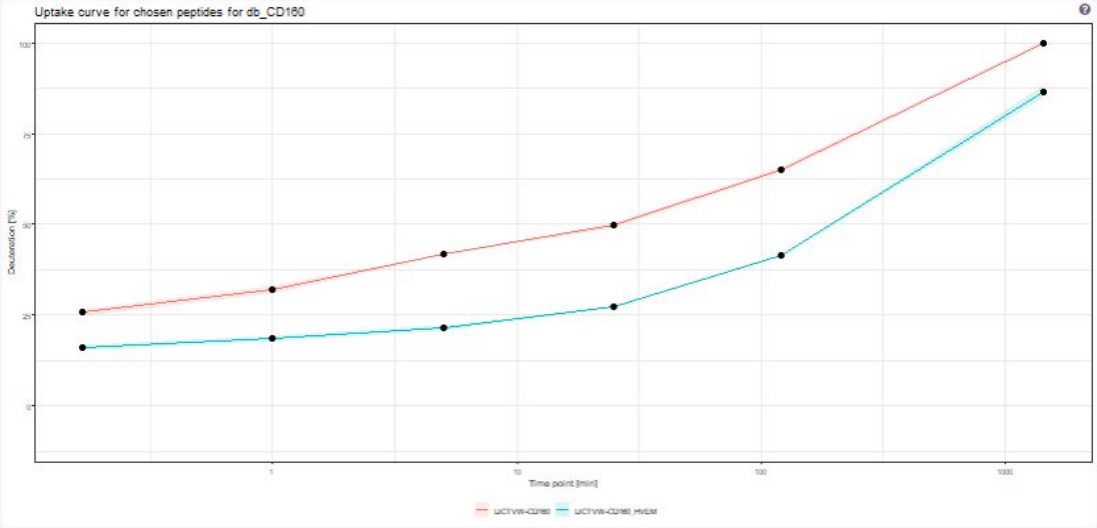
Reset chosen peptides

Choose y range for the plot:

-50 -10 100 200

Adjust labels

Uptake curve | Data



Save chart (.svg)

Select parameters for the plot.

The plot can be rendered only for relative experimental values.

Choose in time:

0.001

Choose time:

0.167

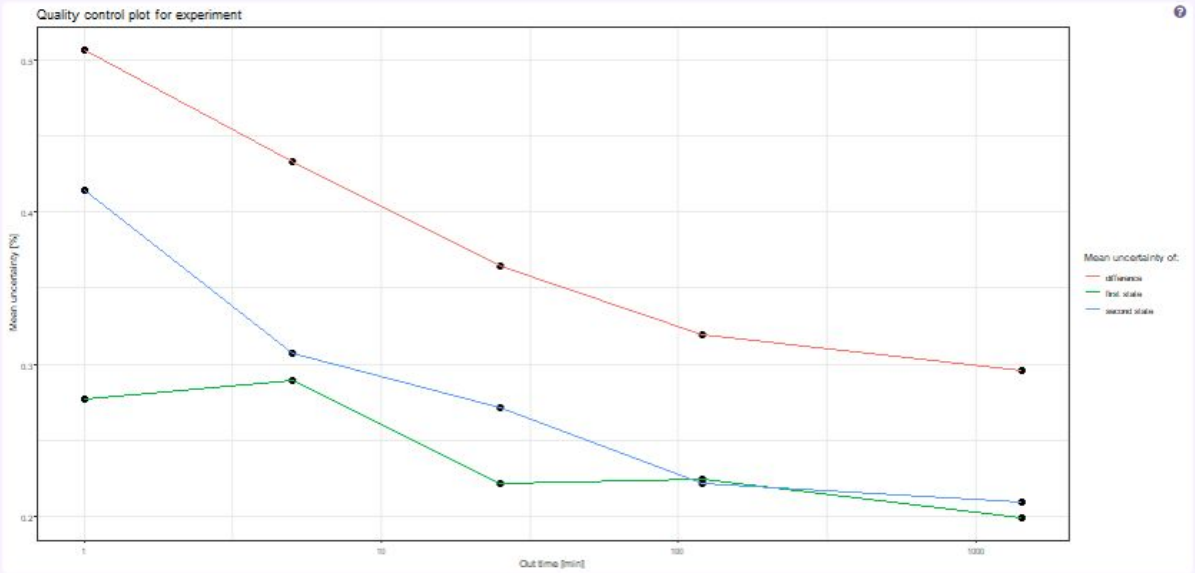
State 1

CD160

State 2

CD160_HVEM

Quality control plot | Data



Save chart (.svg)

This function plots the change in the uncertainty of deuteration levels as a function of incubation time. The uncertainty is averaged over all peptides available at a given time point in a selected state. This chart has a double function: firstly, it allows checking if the measurement uncertainty is decreasing over time (which is the expected behavior). Secondly, it helps to plan the appropriate incubation length for the tested protein (whether we obtain the desired data reliability values).

Name	Value
HDX time course	8
Number of peptides	41
Sequence coverage	86.36%
Average peptide length	14.6829
Redundancy	4.5606
Replicates	4
Significant differences in HDX	0.98 | 0.99
All	All

Excel PDF

Showing 1 to 7 of 7 entries

This table was adapted from the HDX Data Summary (Masson, G.R., Burke, J.E., Ahn, N.G., Anand, G.S., Borchers, C., Brier, S., Bou-Assaf, G.M., Engen, J.R., Englander, S.W., Faber, J., et al. (2019). Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. Nat Methods 16, 595–602).

Choose items for report:

- [x] Position Frequency
- [x] Peptide Coverage
- [x] Comparison Plot
- [x] Theoretical Comparison Plot
- [x] Woods Plot
- [x] Theoretical Woods Plot
- [ ] Uptake Curve
- [ ] Theoretical Uptake Curve
- [ ] Quality Control Plot
- [ ] Position Frequency Data
- [ ] Peptide Coverage Data
- [ ] Comparison Plot Data
- [ ] Theoretical Comparison Plot Data
- [ ] Woods Plot Data
- [ ] Theoretical Woods Plot Data
- [ ] Uptake Curve Data
- [ ] Theoretical Uptake Curve Data
- [ ] Quality Control Plot Data

Elements chosen for report have the same parameters as chosen in panel e.g. axis range and title. Adjust parameters for plots as needed in the report.

Create report!

Keeping up with the competitors

MEMHDX

Finally!

ACCEPTED MANUSCRIPT

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

Weronika Puchała, Michał Burdukiewicz, Michał Kistowski, Katarzyna A Dkşbrowska, Aleksandra E Badaczewska-Dawid, Dominik Cysewski ✉, Michał Dadlez ✉ Author Notes

Bioinformatics, btaa587, https://doi.org/10.1093/bioinformatics/btaa587

Published: 24 June 2020 **Article history** ▾

Thanks!