Structure and Dynamics of Huntingtin. A Segmental Labelling Approach

PhD Project

Xamuel Loft Lund

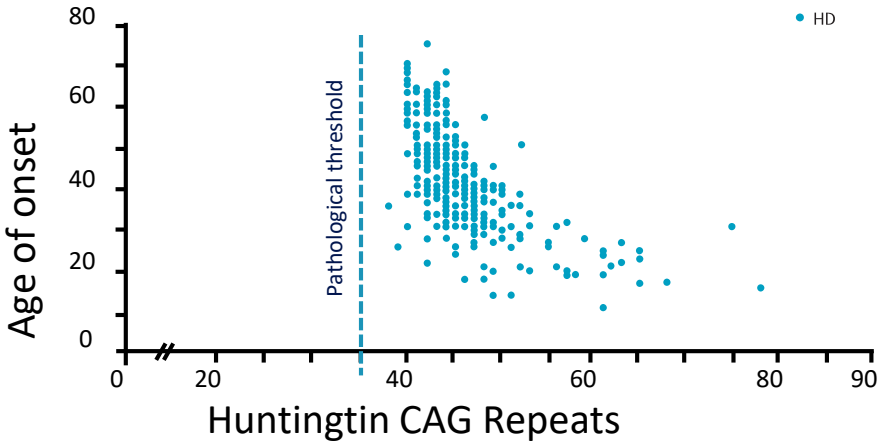
Supervisors:
Pau Bernadó, CBS Montpellier
Anne Martel, ILL Grenoble
Frank Gabel, IBS Grenoble

NEUTRONS FOR SOCIETY

Huntington's Disease and Huntingtin

- Huntington's Disease is a genetically inheritable neurodegenerative disease
- Age of onset is typically 35 to 50 years.
- The disease is caused by abnormal CAG repeats in the gene encoding Huntingtin (Htt)
- CAG repeats beyond the pathological threshold of 35, causes Htt to both loose endogenous functions in neurons and gain functions due to the expanded protein.

Pathological Threshold

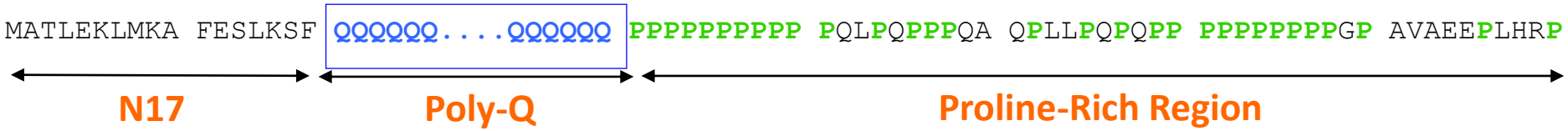


Margolis *et al. Arch. Gen. Psychiatry* 1999, 56, 1019.

NEUTRONS FOR SOCIETY

CBS

The pathology is linked to the Htt Exon1



- The CAG repeats encode the Poly-Q region, located in the exon-1 of Huntingtin.

- This fragment of the protein consists of three regions:
  - N17, which is the N-terminal 17 residues
  - Poly-Q, a low-complexity domain of repeated glutamines
  - Proline rich region(PRR), comprising two poly-P tracts.

NEUTRONS FOR SOCIETY

Aims of the project

- To elucidate the structure of the intrinsically disordered Htt exon1, with a special focus on the poly-Q region in sub-pathological and pathological constructs.

- For this, as a main technique, we will use Small Angle Neutron Scattering (SANS) in segmentally labeled Htt exon1 constructs. These data will be combined with Small Angle X-ray Scattering (SAXS) and computational methods.

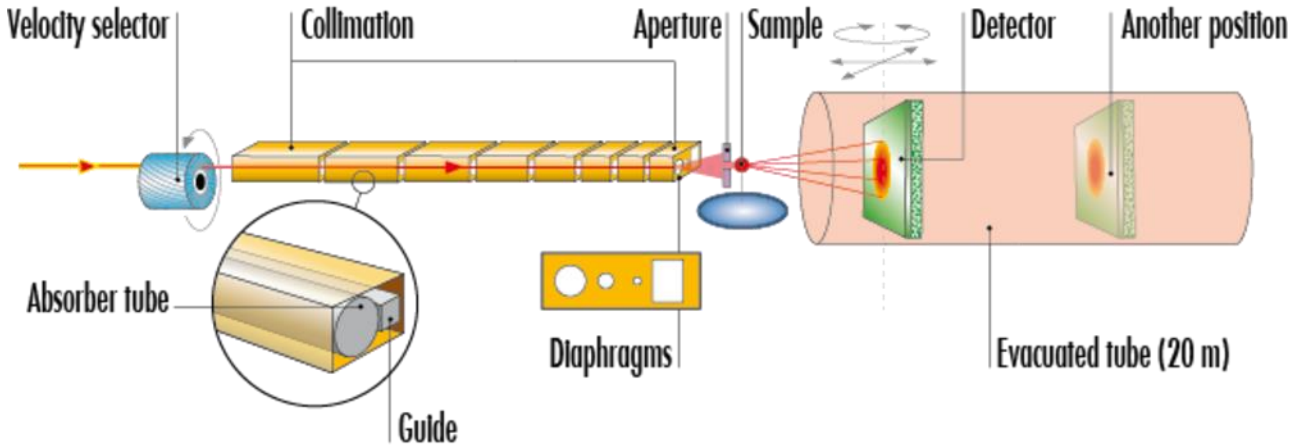
NEUTRONS FOR SOCIETY

Overview

- Methods
  - Small Angle Neutron Scattering
  - Cell-free expression
  - Computational simulations
- Initial Results
  - First beamtime measurements
  - Ensemble simulations
- Perspectives

NEUTRONS FOR SOCIETY

Small Angle Neutron Scattering



Specific setup of the D22 beamline at ILL Grenoble.
https://www.ill.eu/users/instruments/instruments-list/d22/description/instrument-layout

- Small Angle Scattering allows us to gather structural information about biomolecules in solution.
- Small Angle Neutron Scattering (SANS) measures the resulting scattering of a passing beam of neutrons through a sample and it combines elastic, coherent and incoherent scattering.
- Samples can be measured either in batch mode, allowing several samples to be run sequentially, or in SEC-SANS mode. (*Johansen et al. 2018*)

NEUTRONS FOR SOCIETY

SLD varies depending on object and deuteration state

- Using Small Angle Neutron Scattering (SANS) we can use the difference in signal between protonated and deuterated residues to examine the poly-Q domain of Htt exon-1.
- Difference in signal (scattering length density, SLD) is caused when the neutrons interact with the nuclei of the sample during exposure.
- The H/D in biomolecules is modified depending on the % D_2O of the measuring buffer.
- The H/D is equally modified in regards to specific residues.

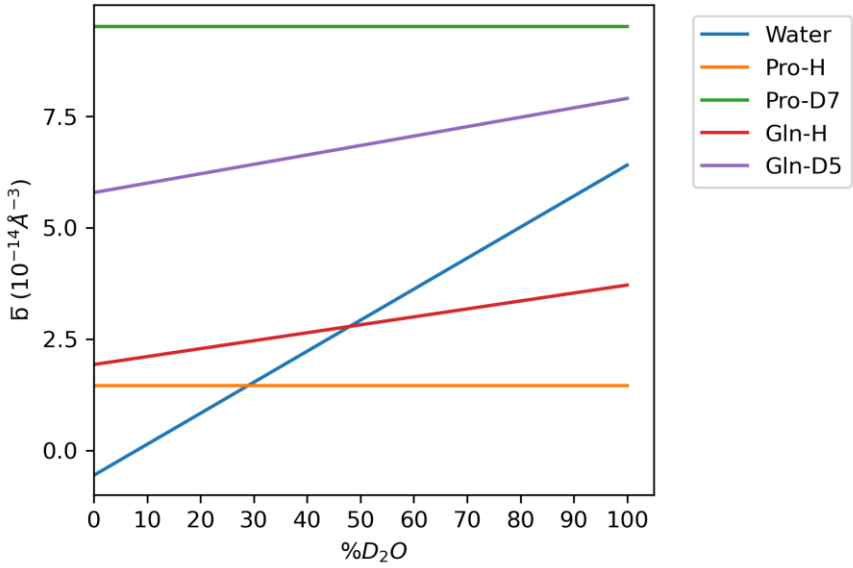
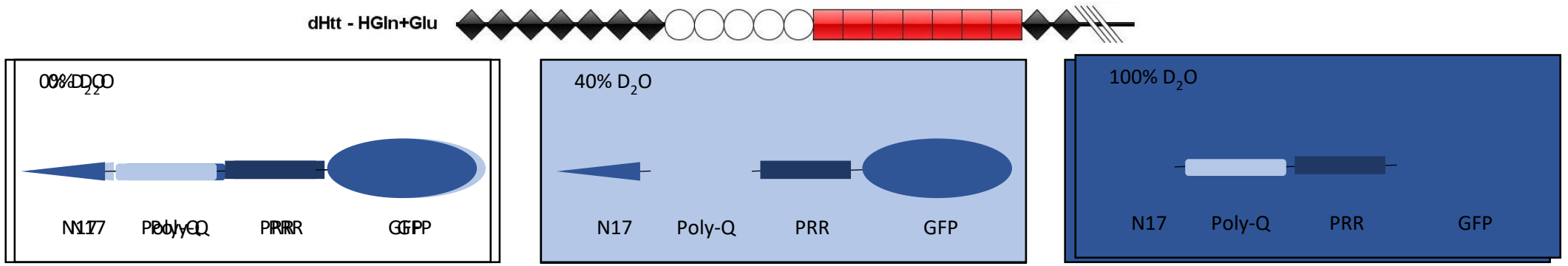


Figure from Jeffries et al. 2020

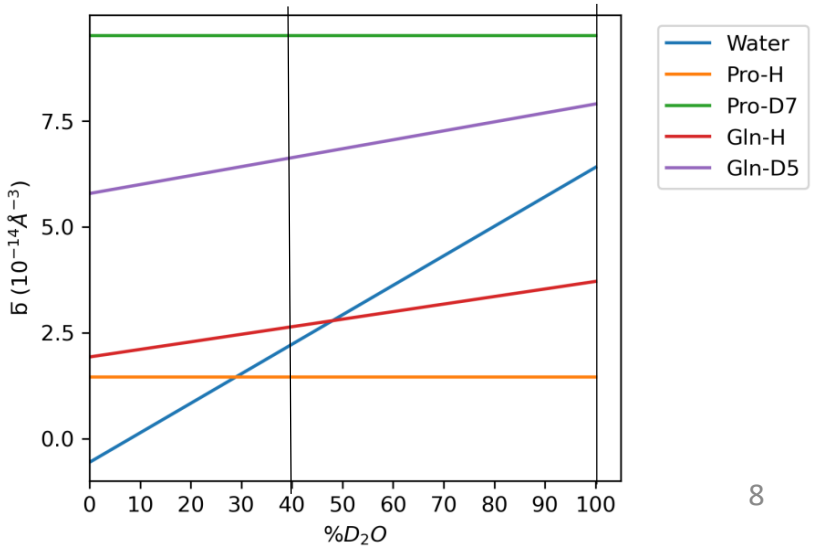
NEUTRONS FOR SOCIETY

Using the SLD of specific residues in combination with low complexity regions



- Htt contains 23 Glutamines (24%) and 32 Prolines (33%).

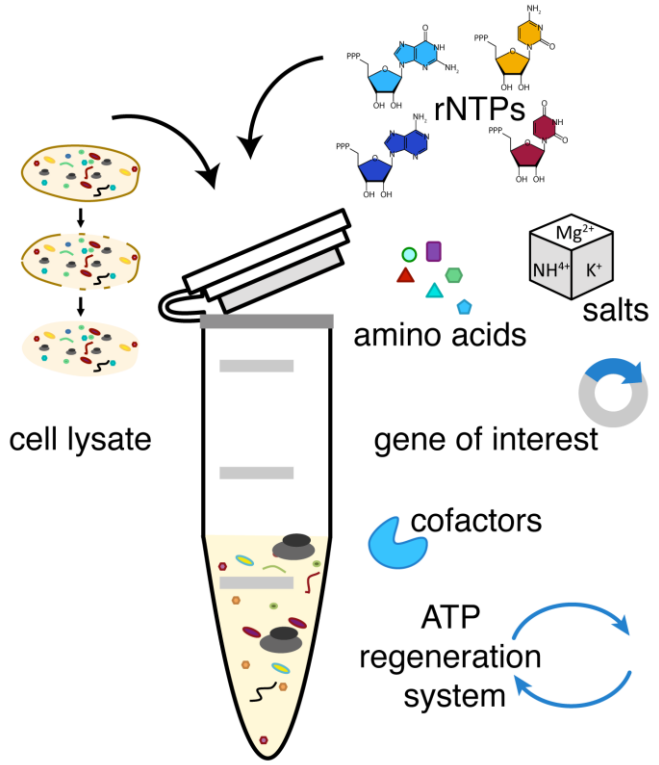
- Different deuteration patterns can produce situations where the majority of the signal will relate to either the GFP and N17, Poly-Q region or the PRR region



Specific deuteration patterns by Cell-free expression

- In Cell-free expression we use the translational system of E. coli to express protein directly in a tube.

- Cell-free expression have several attributes suitable for our project.
  - Control of amino acid mixture, allowing specific deuteration
  - Reduces aggregation of the proteins.
  - Fast expression

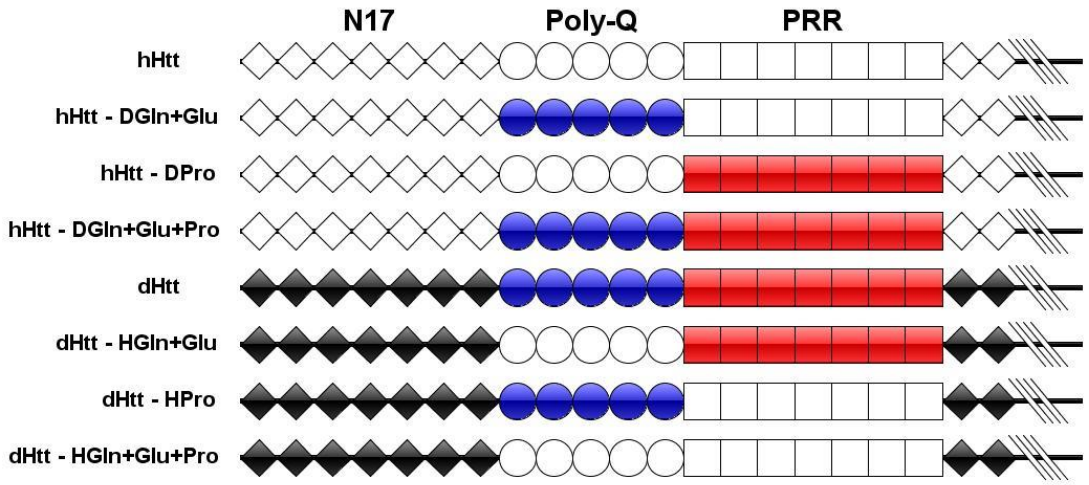


Based on Hong, Kwon, Jewett 2014.

Deuteration scheme

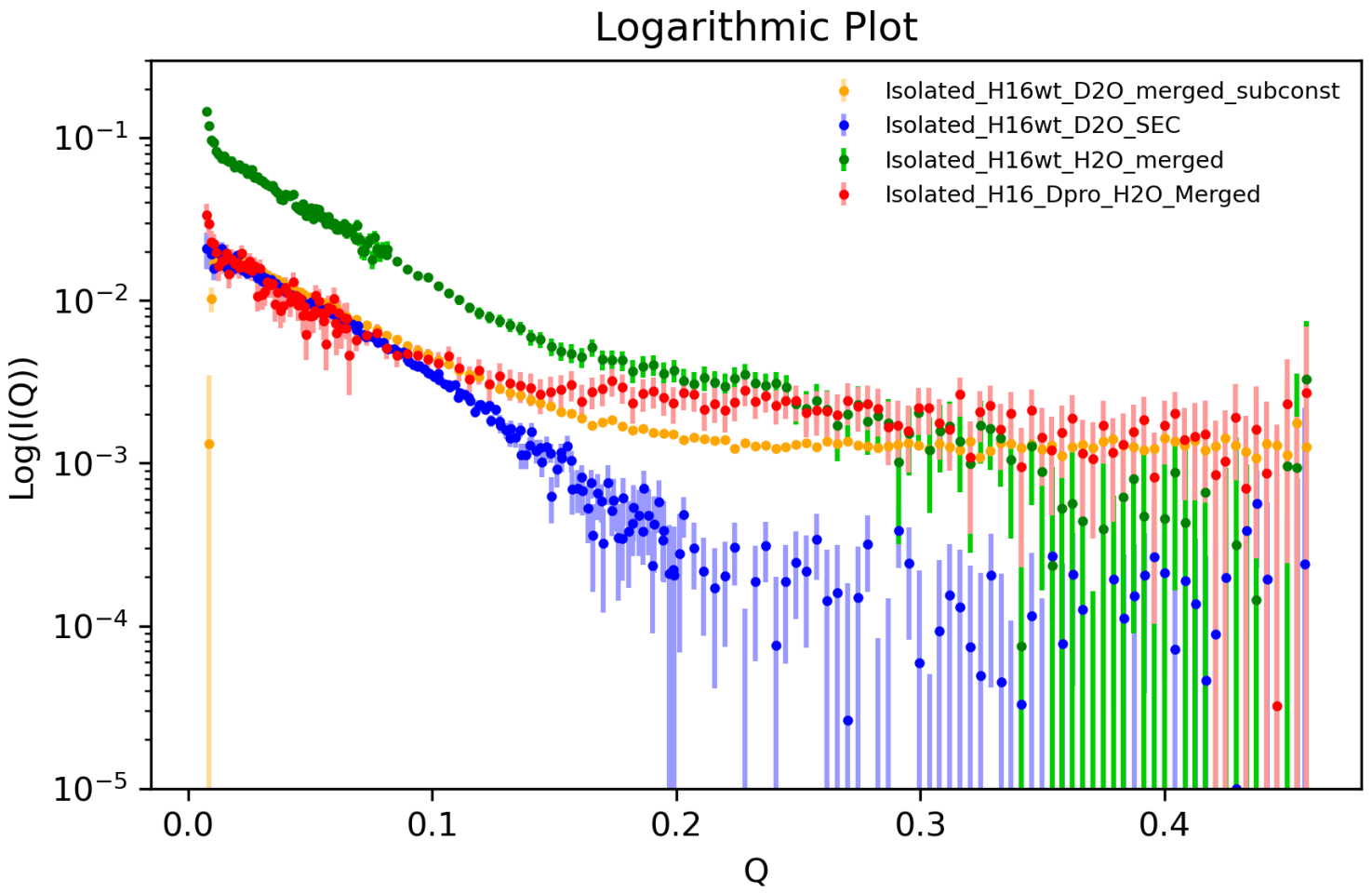
- Deuteration scheme based on the three domains of the Htt exon-1.

- By adding a variation of protonated and deuterated amino acids to the Cell-free expression, a series of constructs with different deuteration schemes will be produced.



NEUTRONS FOR SOCIETY

SANS data from test experiment



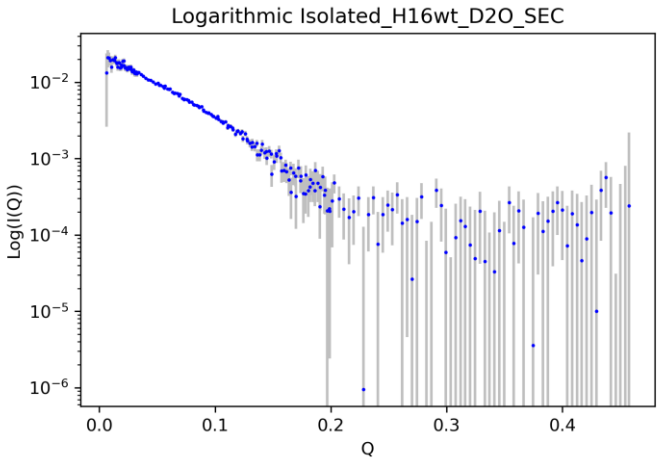
Test experiment at ILL:
TEST-3129
Performed the 21/09-2020

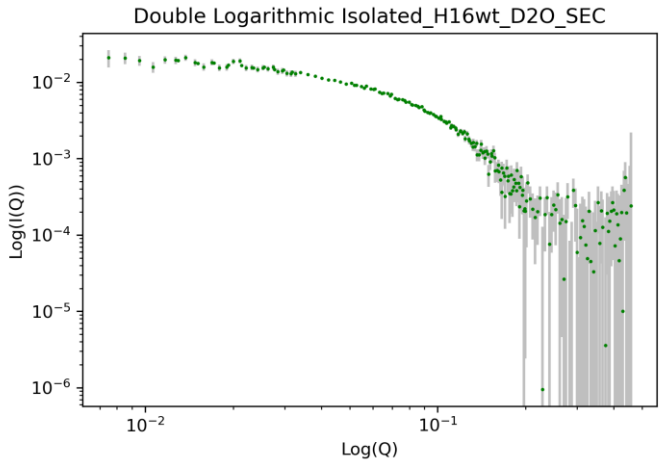
Protonated Htt16 was tested using both SEC-SANS and batch SANS. Htt16-D-Pro was tested only using batch SANS

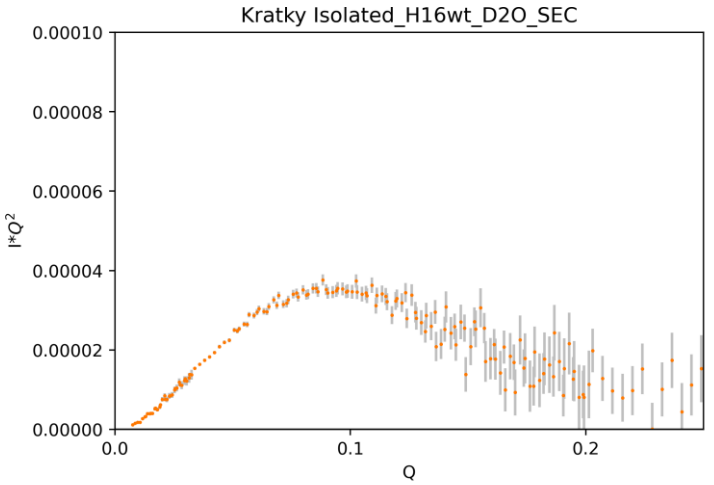
NEUTRONS FOR SOCIETY

CBS

Initial SANS data of Htt16 suggest monomeric protein







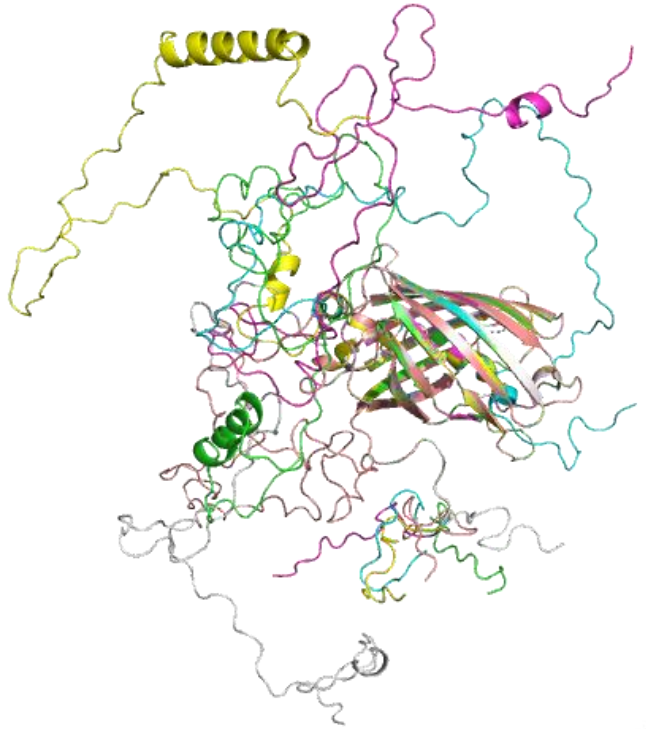
- The Kratky plot shows a predominantly structured protein, which could be caused by sfGFP dominating the signal.

- R_g was 31.5 ± 0.6 Å

Computational study of Htt16

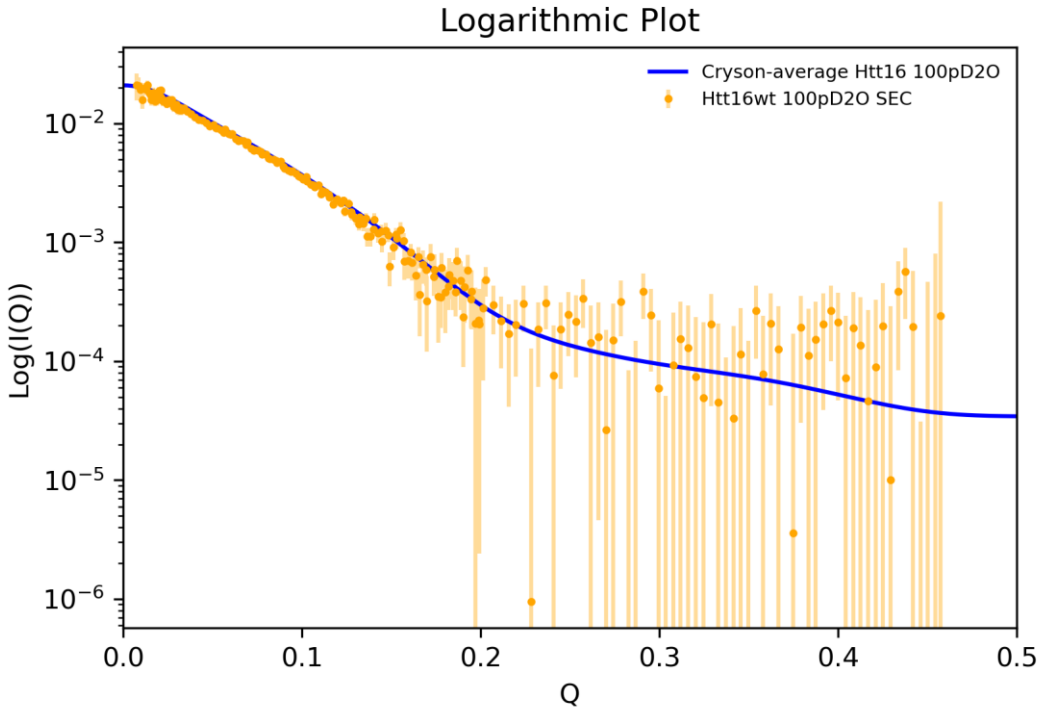
- The group previously produced an ensemble of 11,061 structures based on NMR data (*Urbanek et al. 2020*).

- Crysol and Cryson can be used to generate simulated SAXS and SANS profiles respectively, from these structures.



NEUTRONS FOR SOCIETY

SEC-SANS profile matches the averaged profile of the ensemble



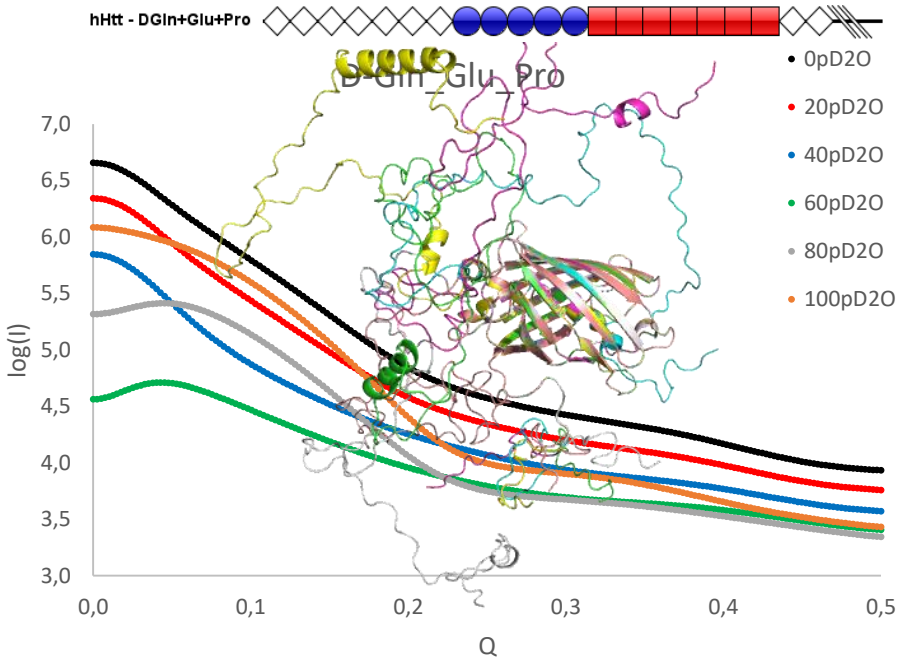
R_g of experimental data: 31.5 $\dot{A}$
R_g of Cryson theoretical curve: 32.7 $\dot{A}$

In the low Q range, the data overlap nicely and shows us that the low concentration sample (SEC-SANS load of 300uL at ~1.5mg/mL) can provide usable data.

NEUTRONS FOR SOCIETY

Cryson

- Using the structures of the ensemble, we have generated Cryson profiles for each of the 11,061 conformations and averaged the profiles per construct.
- This yields us with eight deuteration patterns, at six different solution deuteration levels.
- 48 experimental conditions were monitored.

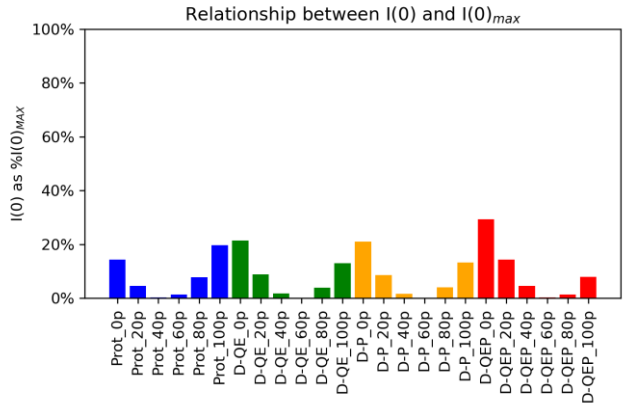


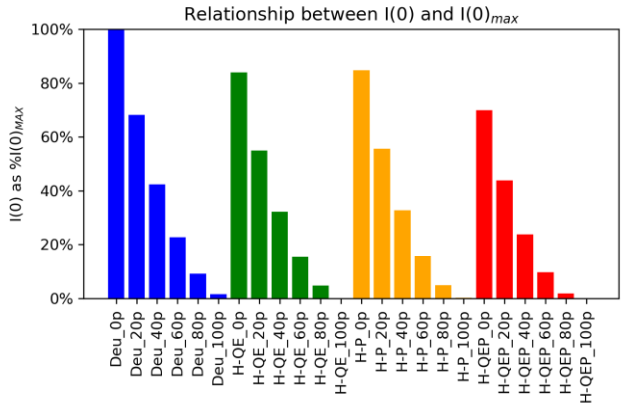
NEUTRONS FOR SOCIETY

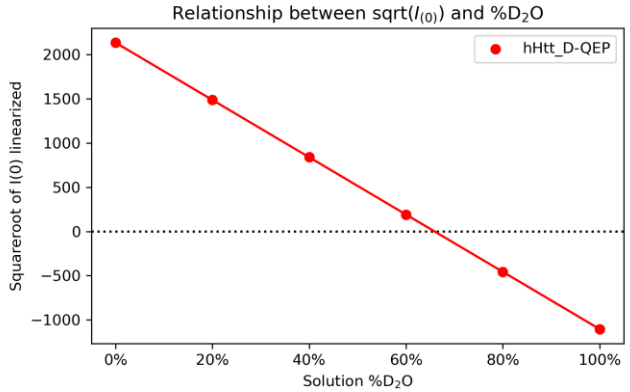
CBS

Using the simulated SANS data, the match point can be estimated

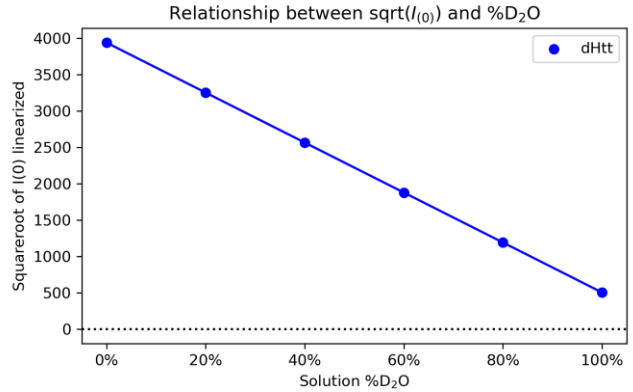
- I(0) values of averaged ensembles plotted in as a fraction of the sample with the highest I(0) ($I(0)_{max}$)
- By plotting the square root of I(0) and inverting datapoints which are after the matchpoint, estimations of each constructs matchpoint can be done.







hHtt_D-QEP's matchpoint is 65.91 %D_2O

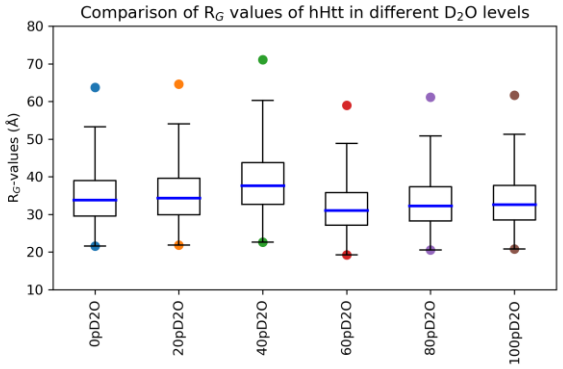


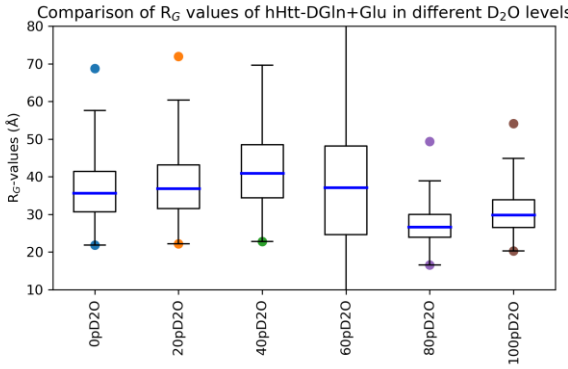
dHtt's matchpoint is 114.73 %D_2O

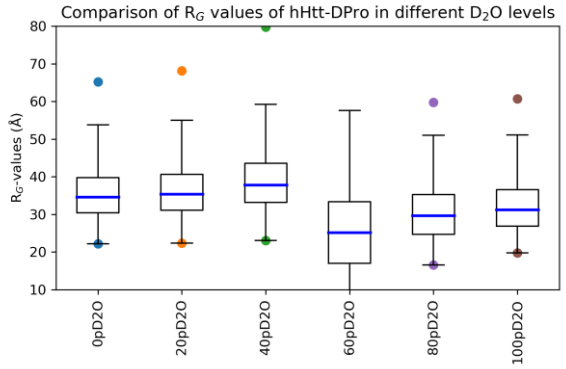
Construct	hHtt	hHtt_D-QE	hHtt _D-P	hHtt _D-QEP	dHtt	dHtt _H-QE	dHtt _H-P	dHtt _H-QEP
Matchpoint	46.0%	56.2%	55.7%	65.9%	114.7%	105.1%	105.6%	97.0%

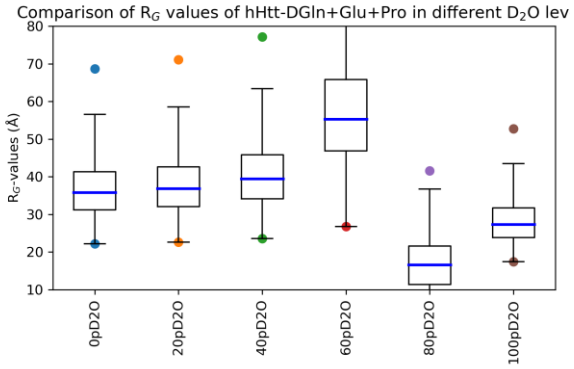
NEUTRONS FOR SOCIETY

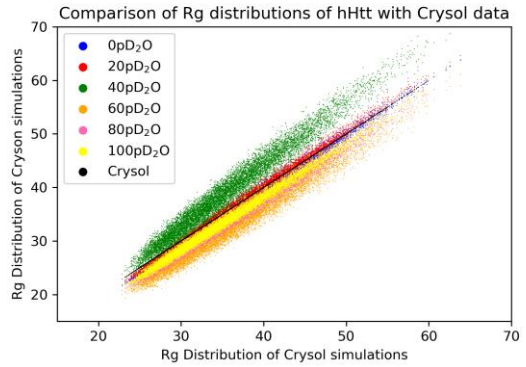
Comparing R_G distributions of protonated samples

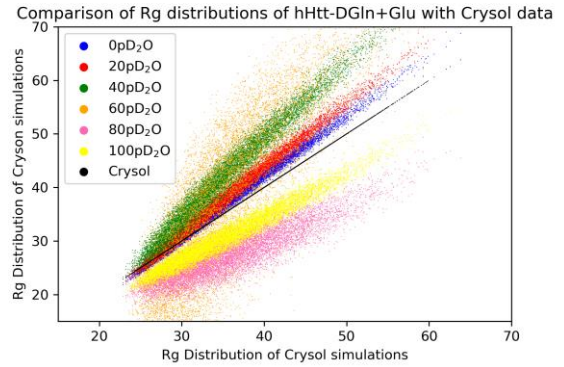


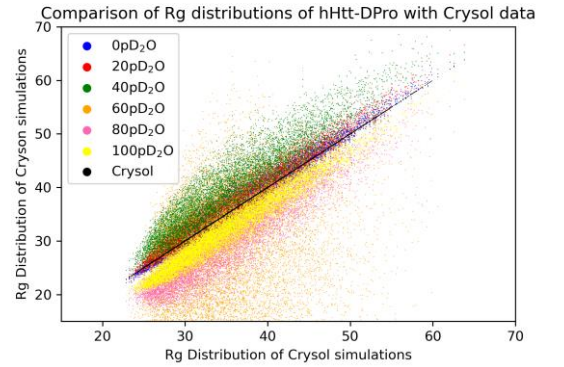


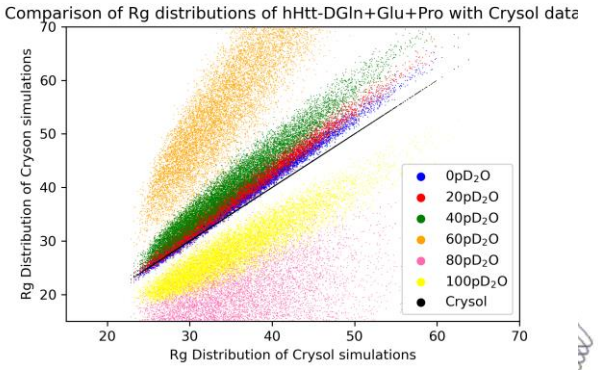










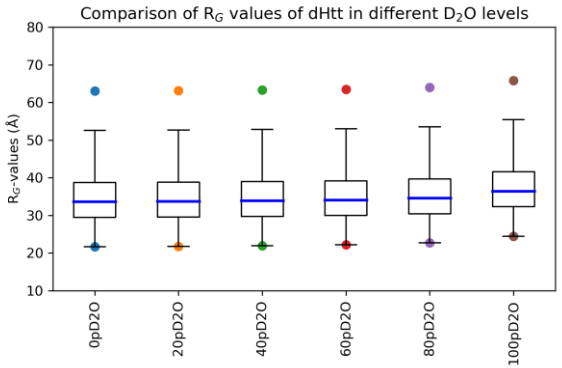


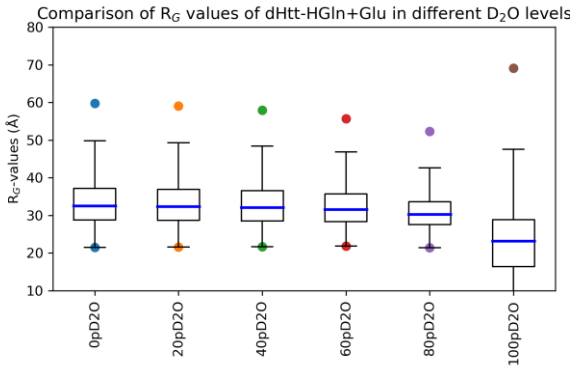
1. The distributions of R_G behave different at different deuteration levels and in different deuteration schemes.

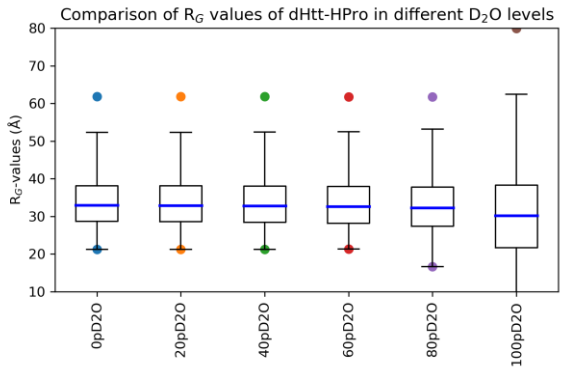
2. R_G distributions becomes incoherent when deuteration level approaches the match point of the construct

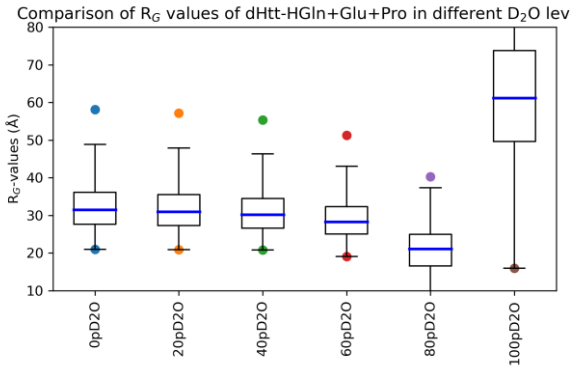
NEUTRONS FOR SOCIETY

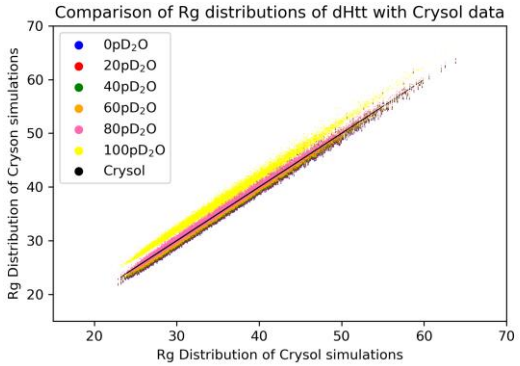
Comparing R_G distributions of deuterated samples

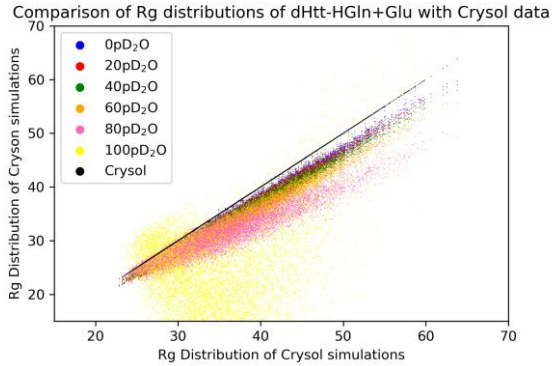


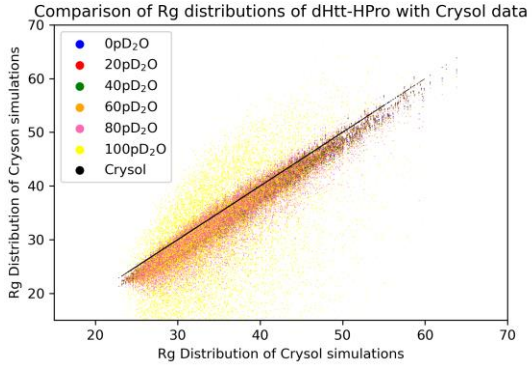


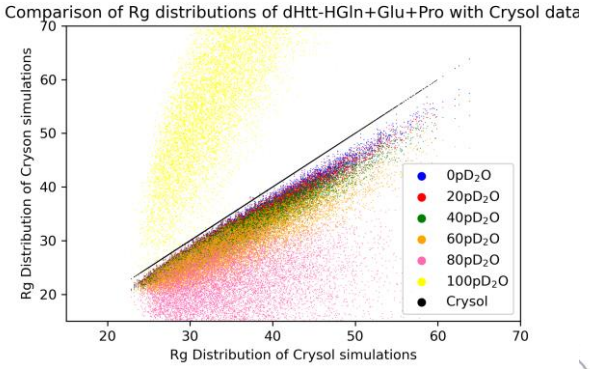












1. The distributions of R_G behave different at different deuteration levels and in different deuteration schemes.

2. R_G distributions becomes incoherent when deuteration level approaches the match point of the construct

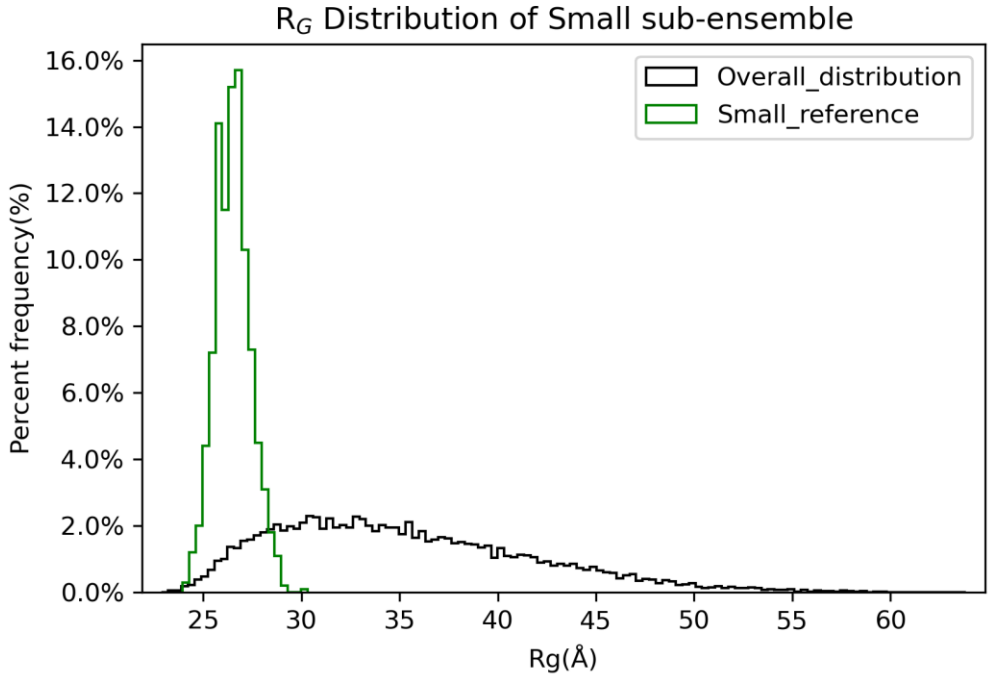
NEUTRONS FOR SOCIETY

Sub-ensembles were created to evaluate the structural information content using EOM

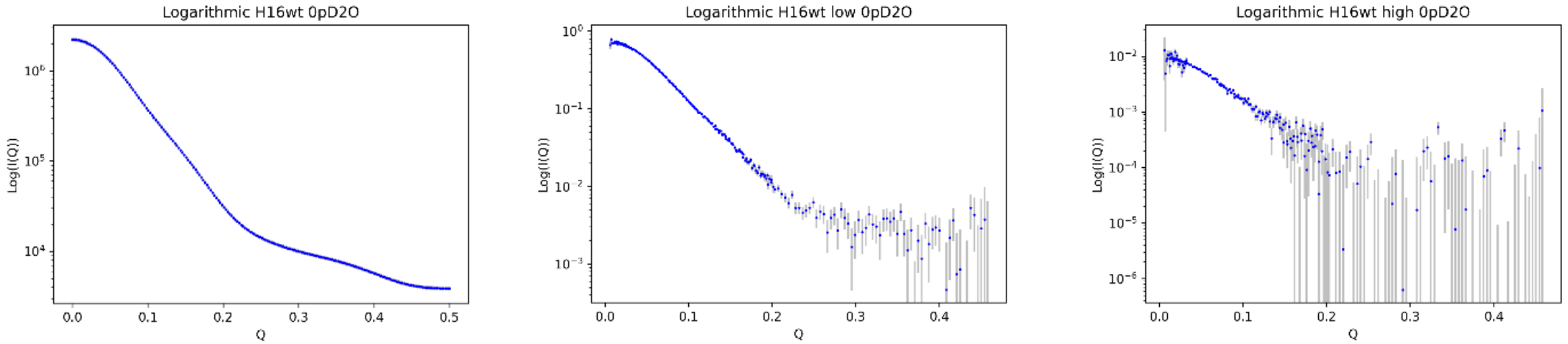
The Ensemble Optimization Method (EOM) is a genetic algorithm used to select an ensemble, that collectively describe a given dataset.

From the simulated ensembles of 11,061 individual Cryson SANS curves, a sub-ensemble of 1,000 structures was selected, distributed randomly around a R_g *of 26* Å

To evaluate the method we used the averaged SANS profile of this biased sub-ensemble as input for EOM.



Synthetic noise was simulated using SEC-SANS noise.



Two levels of synthetic noise was created using the SEC-SANS measurement from the ILL.

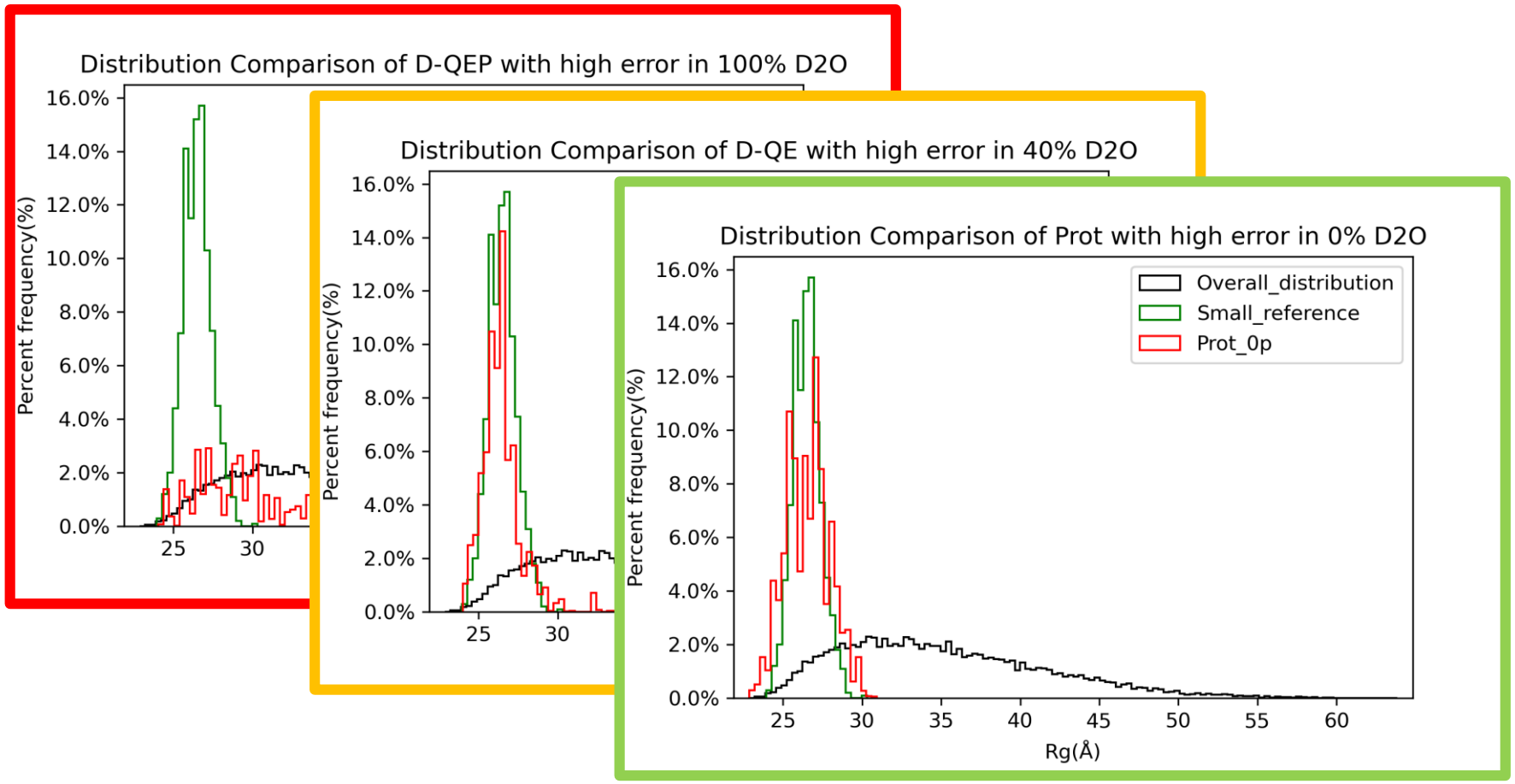
1: High quality SANS data with low noise level

2: Low quality SANS data with high noise level

The relationship between high and low concentrations of D_2O and the amount of incoherent scattering was not taken into account for the noise profiles.

NEUTRONS FOR SOCIETY

EOM results can be divided in three categories



- Red distributions but not recreate a distribution similar to the input distribution.
- Yellow distributions could recreate something similar to the input distribution but all contain some artifacts
- Green distributions could recreate the R_g range of the input

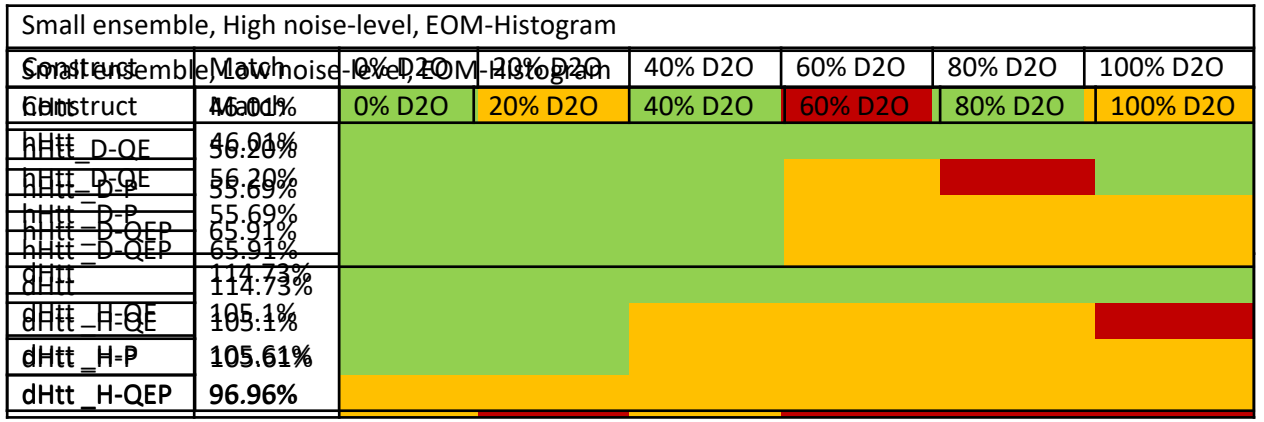
NEUTRONS FOR SOCIETY

Ensemble information from EOM calculations

- Overview of EOM calculations for each deuteration pattern using high noise input.
- Running the same calculations using the low noise profiles yielded more recoverable data.

Small ensemble, High noise-level, EOM-Histogram

Construct	Match	0% D2O	20% D2O	40% D2O	60% D2O	80% D2O	100% D2O
hHtt	46.01%						
hHtt_D-QE	56.20%						
hHtt_D-P	55.69%						
hHtt_D-QEP	65.91%						
dHtt	114.73%						
dHtt_H-QE	105.1%						
dHtt_H-P	105.61%						
dHtt_H-QEP	96.96%						



NEUTRONS FOR SOCIETY

Conclusions for SANS measurements and computational calculations

- H16 can be produced using Cell-free expression and measured using SEC-SANS and the resulting data obtained resembles monomeric protein of the correct size (R_G).

- Initial results from synthetic data and EOM fittings, show that different information can be gathered from different deuteration schemes.

NEUTRONS FOR SOCIETY

The main focuses going forward

- Finish the EOM optimizations and use the results to evaluate which protein samples might provide the most valuable data. This will be done using sub-ensembles that have compact and extended R_G-distributions as well as sub-ensembles with specific distances between glutamines in the Poly-Q region.
- Measure as many informative samples as possible at D22.
- Analyze data from multiple measurements collectively using a modified version of EOM.
- Express samples of Htt exon1 containing 36 Glutamines in the poly-Q domain and initially perform SAXS measurements.

NEUTRONS FOR SOCIETY

Acknowledgements

- **Supervisors**:

Pau Bernadó, CBS Montpellier
Frank Gabel, IBS Grenoble
Anne Martel, ILL Grenoble

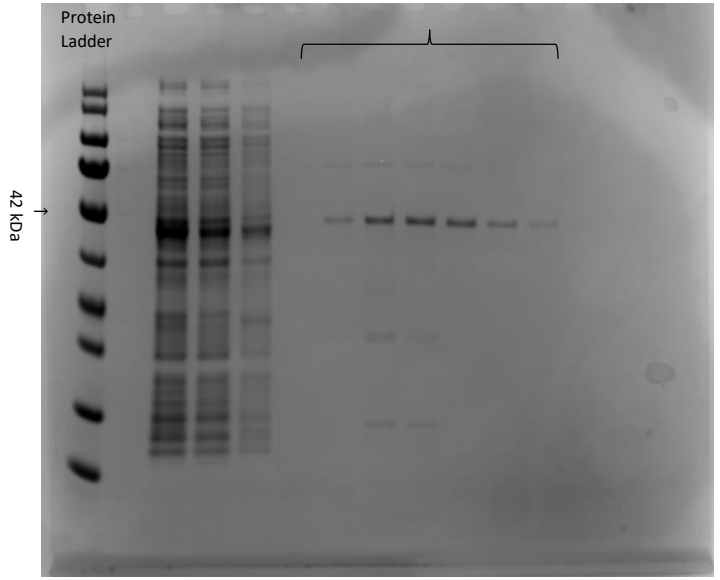
- **Highly flexible proteins, at the CBS**

Amin Sagar, Postdoc
Anna Morató, IR
Annika Urbanek, Postdoc
Carlos Elena-Real, Postdoc
biased

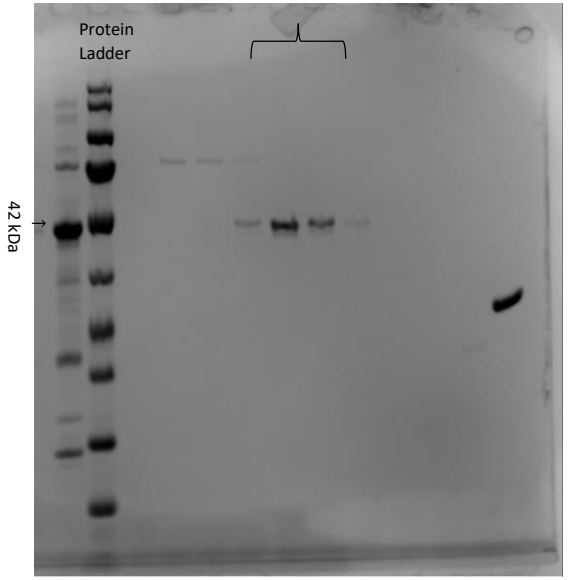
To be deleted

Cell-free expression yields pure protein

IMAC



SEC

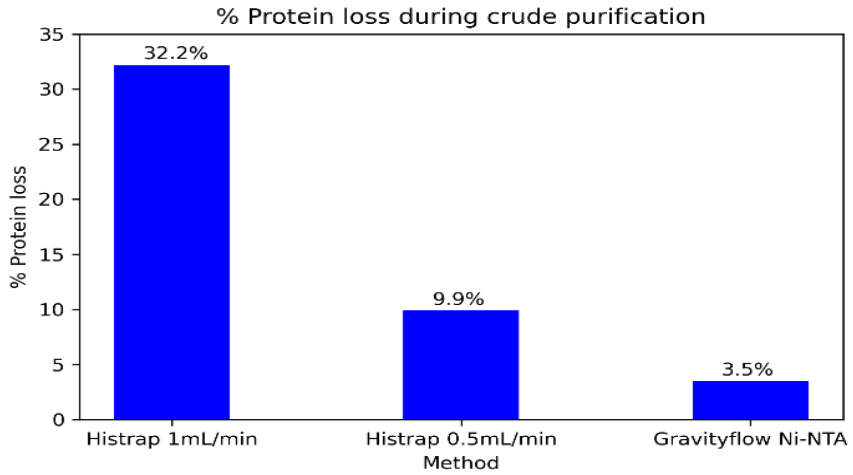


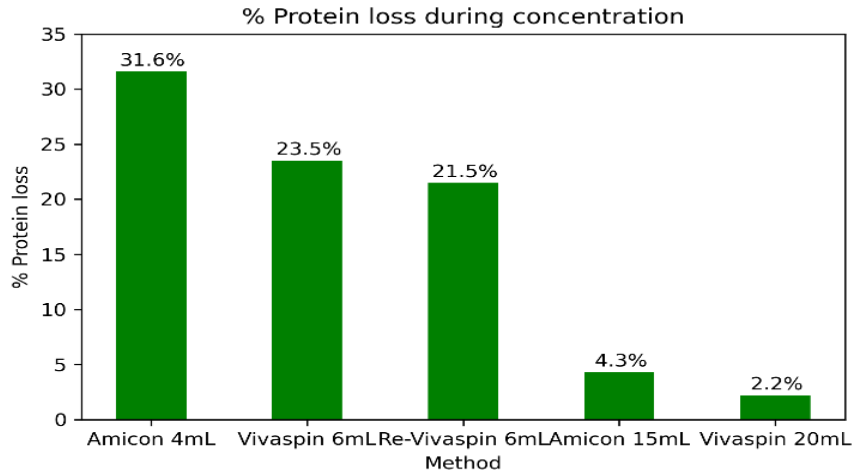
- Protein samples produced of fully protonated Htt16 have been successfully produced using Cell-free expression and purified using Immobilized Ion Affinity Chromatography (IMAC) and Size-Exclusion Chromatography (SEC)
- SDS-PAGE overviews showed clear protein bands around ~40 kDa

NEUTRONS FOR SOCIETY

To be deleted

Optimization of yields





- Crude purification using a FPLC system and 5mL Histrap Excel colums showed a high loss of protein.
- Lowering of flowrate and changing to a gravity flow system both increased yield.

- Initial concentration attempts showed a high loss of protein.
- Using PES membrane units instead of Cellulose membranes increased yield.
- Changing to concentration units with a larger volume increased yield.

NEUTRONS FOR SOCIETY

CBS

Production of protein with deuterated Glutamine and Glutamic acid needs further investigation To be deleted

- Scrambling during expression causes residues of be interchanged due to biological functions of enzymes from the E. coli lysates.

- Replacement of Potassium Glutamate (Kglu) with Potassium Acetate (KAce), shows a decrease in yield of ~70%.

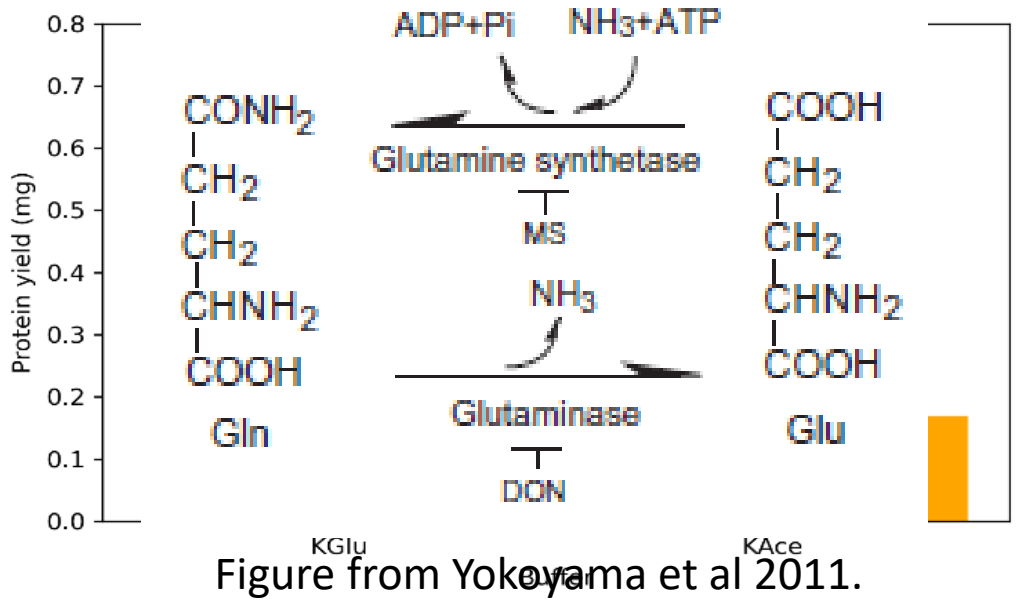


Figure from Yokoyama et al 2011.

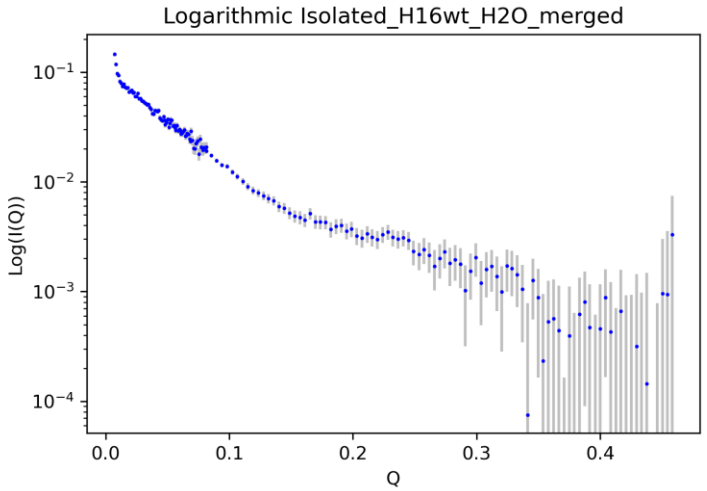
To be deleted

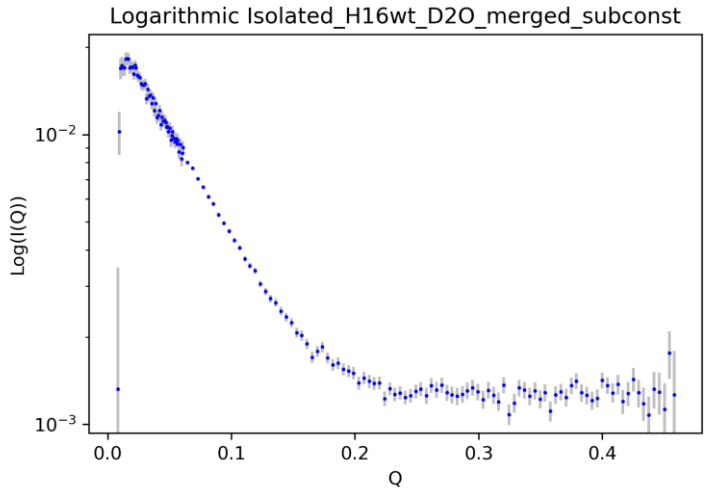
Cellfree conclusions

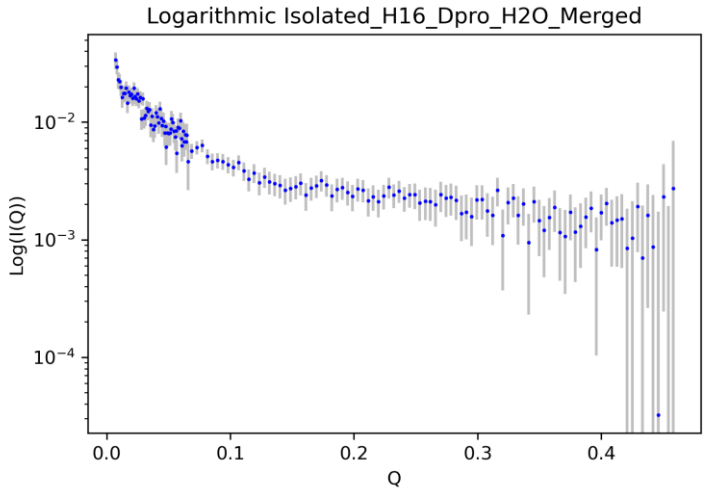
- Htt16 and Htt16-DPro can be produced using Cell-free expression.
- Lower flowrate increased protein yield from the Histrap crude purification.
- Concentration using PES-membrane filters is time consuming, but offers a lower protein loss than Amicon-filters
- Using an acetate-buffer instead of glutamate-buffer during cell-free has to be optimized.
- Overall the loss of each step after initial purification has been improved.

NEUTRONS FOR SOCIETY

Batch samples show signs of either protein aggregation or buffer-mismatch





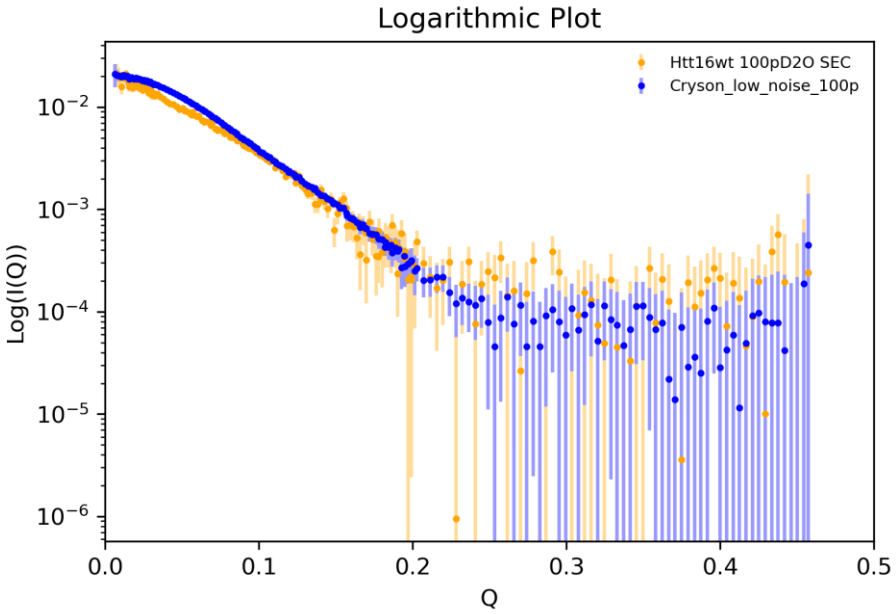


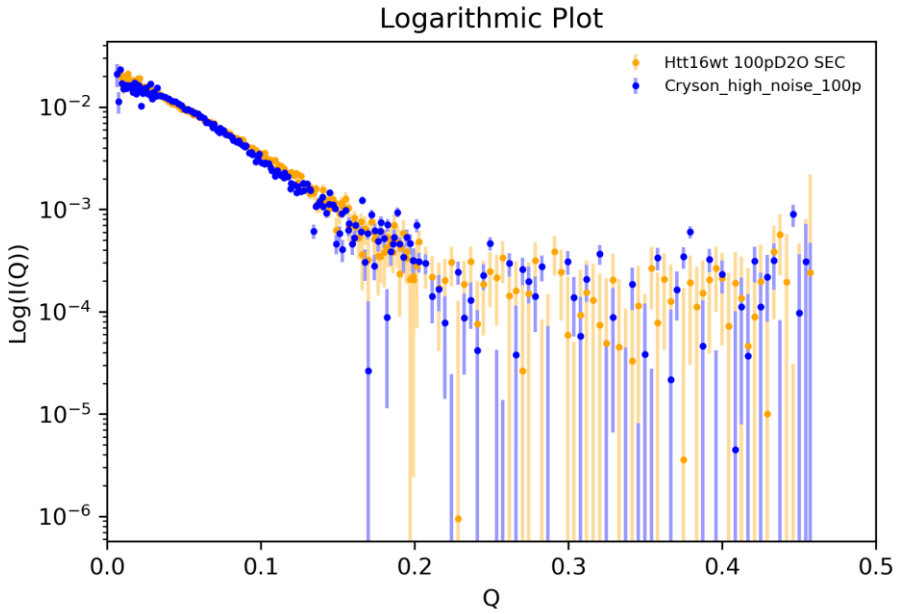
Sharp increase towards the lowest Q-range is a sign of protein aggregation in the sample.

A sharp fall at low Q-range could be a sign of buffer mismatch. Here it could be caused by sample having passed through a S75 5/300 24mL column while buffer was taken before the column.

NEUTRONS FOR SOCIETY

Comparison between simulated and experimental SANS data





The low noise profiles show a lower level of noise than the experimental data while the high noise profiles show an increased error profile.

NEUTRONS FOR SOCIETY