May 23, 2019

Purification of the HTT C-HEAT domain (2088-3144)

SGC code	Construct ID	Cloned Sequence	C-terminal Tag	Mol. Weight (Da)[1]
TOC019 C01	HTT_2088-3144_pBac Mam2-Diex-LIC	LETVSPDKDWYVHLVKSQCWTRSDSALLEGAELVNRIPAEDMNAFMMNSEFNL SLLAPCLSLGMSEISGGQKSALFEAAREVTLARVSGTVQQLPAVHHVFQPELP AEPAAYWSKLNDLFGDAALYQSLPTLARALAQYLVVVSKLPSHLHLPPEKEKD IVKFVVATLEALSWHLIHEQIPLSLDLQAGLDCCCLALQLPGLWSVVSSTEFV THACSLIHCVHFILEAVAVQPGEQLLSPERRTNTPKAISEEEEEVDPNTQNPK YITAACEMVAEMVESLQSVLALGHKRNSGVPAFLTPLLRNIIISLARLPLVNS YTRVPPLVWKLGWSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYRINTLGWTS RTQFEETWATLLGVLVTQPLVMEQEESPPEEDTERTQINVLAVQAITSLVLSA MTVPVAGNPAVSCLEQQPRNKPLKALDTRFGRKLSIIRGIVEQEIQAMVSKRE NIATHHLYQAWDPVPSLSPATTGALISHEKLLLQINPERELGSMSYKLGQVSI HSVWLGNSITPLREEEWDEEEEEEADAPAPSSPPTSPVNSRKHRAGVDIHSCS QFLLELYSRWILPSSSARRTPAILISEVVRSLLVVSDLFTERNQFELMYVTLT ELRRVHPSEDEILAQYLVPATCKAAAVLGMDKAVAEPVSRLLESTLRSSHLPS RVGALHGILYVLECDLLDDTAKQLIPVISDYLLSNLKGIAHCVNIHSQQHVLV MCATAFYLIENYPLDVGPEFSASIIQMCGVMLSGSEESTPSIIYHCALRGLER LLLSEQLSRLDAESLVKLSVDRVNVHSPHRAMAALGLMLTCMYTGKEKVSPGR TSDPNPAAPDSESVIVAMERVSVLFDRIRKGFPCEARVVARILPQFLDDFFPP QDIMNKVIGEFLSNQQPYPQFMATVVYKVFQTLHSTGQSSMVRDWVMLSLSNF TQRAPVAMATWSLSCFFVSASTSPWVAAILPHVISRMGKLEQVDVNLFCLVAT DFYRHQIEEELDRRAFQSVLEVVAAPGSPYHRLLTCLRNVHKVTTC	DYKDDDDK	118117

Rationale

The purification of HTT fragments is a useful approach to learn more about the function of huntingtin in the cell. By obtaining soluble and monomeric samples of HTT domains namely the C-HEAT, N-HEAT and bridge domains, specific protein-protein interactions can be studied. Furthermore, domains of HTT in soluble monomeric form could enable crystallization studies.

Expression and purification of these fragments can be found on these posts https://zenodo.org/record/2600051#.XKU89aeZPOQ and https://zenodo.org/record/2628060#.XULMtnspDb0 (performed by Dr. Rachel Harding). The latest post shows the successful purification of a monomeric and mono disperse sample of the HTT C-HEAT domain. The results here presented are a follow up of those experiments and aim to further characterize the HTT C-HEAT domain as well as explore strategies to improve buffer conditions for crystallization purposes.

Purpose

To perform crystallography studies and biophysical characterization of the HTT C-HEAT domain of huntingtin (HTT C-HEAT domain).

To determine if limited proteolysis of the HTT C-HEAT domain enables crystal formation.

To determine the best buffer conditions for the HTT C-HEAT domain.

Methods

Expression of the HTT C-HEAT domain

Expression of the HTT C-HEAT domain was done following similar expression procedures as the full length HTT[2]. 4 L of Sf9 cell culture at a density of ~4.5 million cells per mL were infected with 8 mL (for the 4 L cell culture) of P3 recombinant baculovirus and grown at 130 rpm and 27 °C. HyQ SFX insect serum medium containing 10 µg/mL gentamicin was used as the culture medium. Infected cells were

harvested when viability dropped to 80%–85%, normally after ~72 h post-infection. Cells were harvested by centrifugation and stored at -80 ˚C prior to purification.

Purification of the HTT C-HEAT domain

The purification methods used followed those by Harding *et al*. with minor modifications[2]. Briefly, the cell pellets were thawed and then diluted in FLAG buffer consisting of 20 mM HEPES pH 7.4, 300 mM NaCl, 5 % glycerol. Benzonase and $MgCl_2$ were added at concentration of 0.02µg/mL and 2 mM, respectively. Cell debris was separated by centrifugation at 15,000 rpm, 1 h, 4 °C (Beckman JLA16.2500).

The supernatant was incubated with anti-FLAG slurry (Sigma M2) pre-equilibrated with FLAG buffer at 4 ˚C with rocking for 2 hours. Then, the resin was washed 3 times. The first wash was done with 80 CV of FLAG buffer. The second was as done with 80 CV of FLAG-buffer supplemented with 5 mM MgCl2 and 5 mM ATP. The third was 15 CV of FLAG buffer. Then, HTT C-HEAT domain was eluted with 8 CV of FLAG buffer supplemented with 250 µg/mL 3xFLAG peptide. The anti-FLAG resin was cleaned immediately after use with 100 mM glycine at pH 3.5 and then stored in buffer containing 20 mM Hepes, 300 mM NaCl and 50 % glycerol.

The HTT C-HEAT domain was eluted from the anti-FLAG resin was concentrated to 1 mL at 4,500 rpm for 15 min (with resuspension of the sample every 5 min) in a pre-equilibrated 50 KDa concentrator (Millipore) with Superose 6 buffer containing 20 mM Hepes pH 7.4, 300 mM NaCl, 1 mM TCEP, 5 % (v/v) glycerol. Then, the concentrated sample was loaded onto a pre-equilibrated Superose 6 10/300 GL column with Superose6 buffer. Samples at every step of the purification were run with 4-20 % tris-glycine SDS-PAGE. The molecular weight of the HTT C-HEAT domain was confirmed by LCMS.

Crystallography

The eluted sample from the gel filtration column was collected and concentrated in a 50 KDa concentrator (Millipore) at 4,500 rpm to a concentration of 8 mg/mL. The concentration was measured using the extinction coefficient (122270 M^{-1} cm^{-1}) at 280 nm. Then, the sample was divided in three parts. One part was left untouched while the other two parts were exposed to trypsin in 1:1000 and 1:10000 w/w ratio, respectively. Crystal trays were set using the SGC[3], Redwing[4] and Peg-Ion[5] (Hampton Research) conditions.

LC-MS

1 pmol of HTT C-HEAT domain was used for the LC-MS analysis in an Agilent 1260 capillary HPLC system coupled to an Agilent Q-TOF 6545 mass spectrometer via the DUAL Agilent Jetstream.

Differential scanning fluorimetry (DSF)

DSF experiments with TOC019C01 were performed in a real-time PCR device (Mx3005p from Stratagene) at different protein concentrations (0.2, 0.1, 0.05, 0.025 mg/mL) using a 96-well PCR microplates. The protein unfolding was monitored by the increase in the fluorescence of the fluorophore SYPRO Orange (Invitrogen). The protein in the experiment was in either size exclusion buffer (buffer 1) or buffer containing 20 mM HEPES pH 7.4 and 150 mM NaCl (buffer 2,). Optical foil was used to cover the plates to prevent evaporation. The samples were heated at 0.5 °C per minute from 20-95 °C. All experiments were done in duplicate. The fluorescence intensity was plotted as a function of temperature and then fitted to the Boltzmann equation by nonlinear regression.

Differential Static Light Scattering (DSLS)

Temperature-dependent aggregation was measured by using static light scattering (StarGazer)[6, 7]. Three different protein concentration were tested: 0.4, 0.2 and 0.1 mg/mL as well as two buffers: buffer 1 is the

size exclusion buffer and buffer 2 was the same used in the DSF experiments. Fifty microliters of protein were heated from 25 to 85 °C at a rate of 1 °C per min in each well of a clear bottom 384-well plate (Nunc). Light was shone on the protein from underneath at an angle of 30 °. Aggregation of the sample was measured by the intensity of scattered light every 30 s with a CCD camera. The pixel intensity in a preselected region of each well were integrated to generate values of the total amount of scattered light. The intensities were plotted against temperature and fitted to the Boltzmann equation by nonlinear regression. The temperature at the inflection point of the curve is represents the temperature of aggregation (T_{agg}).

Results

The HTT C-HEAT domain was purified using a two-step purification protocol that involved first binding to anti-FLAG resin and second separation by gel filtration chromatography using a Superose 6 column (Figure 1 and Figure 2). By using this purification protocol, the HTT C-HEAT domain was obtained in an estimated 90 % purity and a yield of 0.45 mg/ L of Sf9 cell culture. Notably, the A_{260}/A_{280} ratio of this sample was 0.7 indicating that there is considerable amount of nucleic acid material in sample. This is similar to the full length HTT sample [2]. Furthermore, due to overloading of the Superose 6 column, there is poor separation of other proteins at low molecular weights ~ 63 - 48 KDa (Figure 2).

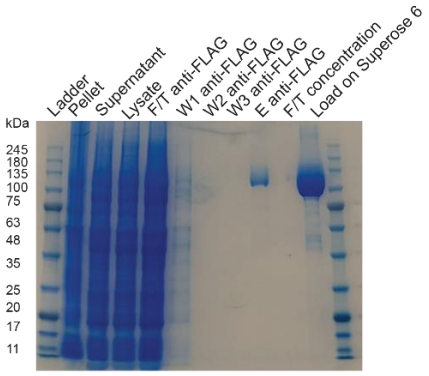

Figure 1. Purification of the HTT C-HEAT domain. SDS-PAGE gel containing the samples post purification of the HTT C-HEAT domain. The Sf9 cells were thawed and resuspended in FLAG buffer from which the soluble (supernatant) and non-soluble (pellet) are separated by centrifugation. A sample of the anti-FLAG load is taken as well as the flow through, wash 1, wash 2, wash 3 and elution. Additionally, a sample prior loading the HTT C-HEAT domain on the Superose 6 column is also taken as well as the flow through in the concentration step. For each sample 5 µL were loaded on the gel.

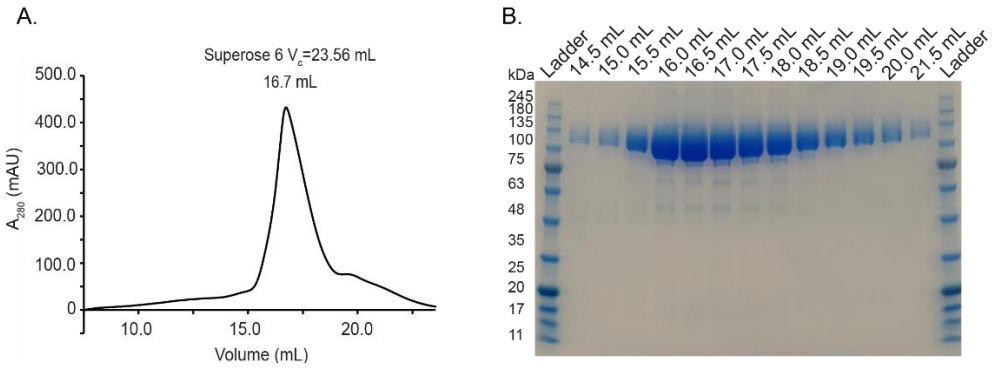

Figure 2. Gel filtration chromatography A. The concentrated sample of the HTT C-HEAT domain was loaded on a Superose 6 10/300 GL column. The absorbance of the eluted sample is monitored at 280 nm. The elution volume of the sample is 16.7 mL. The total volume of the column is Vc=23.56 mL. B. SDS-PAGE gel of the eluted fractions of the gel filtration column. For each sample 15 µL were loaded.

The freshly purified sample of the HTT C-HEAT domain was used to prepare crystal trays. In total twelve different crystal trays were made that screened for a variety of conditions (Table 1).

Table 1. Crystal trays set up with pure HTT C-HEAT domain

# of trays	[protein] mg/mL	Trypsin ratio	Conditions
1	8	0	SGC
1	8	1 to 1000	SGC
1	8	1 to 10000	SGC
1	8	0	Redwing
1	8	1 to 1000	Redwing
1	8	1 to 10000	Redwing
1	8	1 to 1000	PegIon
1	8	1 to 10000	PegIon
1	5	1 to 1000	Redwing
1	5	1 to 10000	Redwing
1	3	1 to 1000	Redwing
1	3	1 to 10000	Redwing

Note: SGC[3], Redwing[4] and Peg-Ion[5]

LC-MS analysis shows that molecular weight of the HTT C-HEAT domain is 182282 Da which differs by 165 Da from the theoretical MW (Figure 3). This difference is likely due to post translational modifications such as phosphorylation as it is seen with full length HTT [2]. Thus, the difference of 165 Da seen is likely due to the HTT C-HEAT domain becoming phosphorylated at two sites (phosphate group has molecular weight of 80 Da).

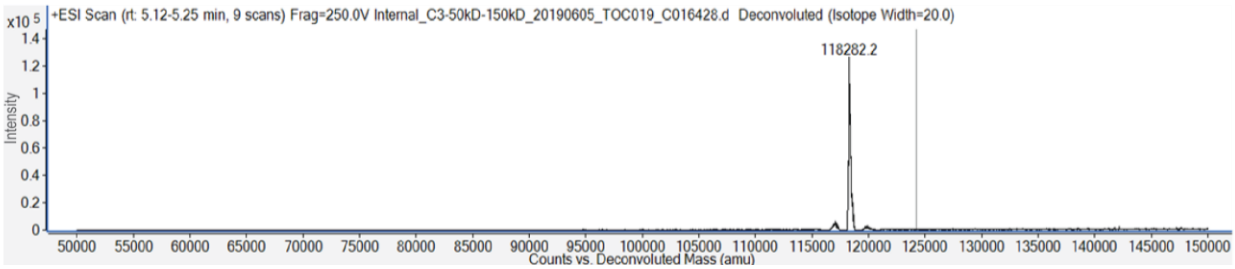

Figure 3 LC-MS spectra of the HTT C-HEAT domain.

We used the remaining sample from the purification to check for sample stability as well as to determine if limited proteolysis of the C-HEAT domain produced stable domains that could be used to determine the boundaries of new constructs. The sample C-HEAT domain preserves well at 4 °C (Figure 4) even after three weeks in storage. Limited proteolysis of C-HEAT domain 1 shows the formation of some products at around 25 kDa which implies the presence of stable domains that resist digestion. In turn, using a stable domain for crystallography would be more favourable for crystal formation [8, 9]. Preliminary analysis by LC-MS of the fragments show that the there is a more stable construct for C-HEAT domain which starts from residue 2284 to residue 3144. If our current attempts at crystallizing C-HEAT domain fail, in the future we could use this new domain boundaries to determine if this construct is more amenable for crystallography. Because the results of these experiments only represent preliminary experiments, future experiments would test the reproducibility of these results.

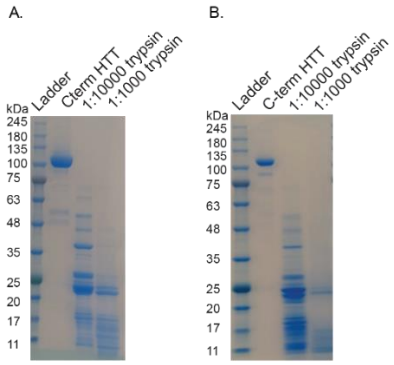

Figure 4 Sample stability of the HTT C-HEAT domain construct. Samples of HTT C-HEAT domain as well as trypsin treated samples of HTT C-HEAT domain were stored at 4 °C for one week (A.) and three weeks (B). The SDS-PAGE of the samples is shown

Differential scanning fluorimetry (DSF)

Preliminary DSF experiments were done with the HTT C-HEAT domain at different concentrations and in two different buffer systems in order to screen for optimal protein buffer conditions. Buffer 1 contained 300 mM and 5 % v/v glycerol, 20 mM HEPES pH 7.4 and TCEP 1 mM while buffer 2 was contained 20 mM HEPES pH 7.4 and 150 mM NaCl. No significant differences were observed in the T_m for TOC019C01 (≤ 2 °C), implying that the protein is equally stable in both buffers (Table 2). However, this preliminary experiment demonstrated that the ideal concentration at which the experiments should be conducted is 0.1 mg/mL. Experiments ran with concentrations higher or lower than that had significantly lower signal to noise ratio (Figure 5). Future studies with DSF will involve other buffer conditions as well as testing of the other HEAT domain constructs. Similarly, DSF experiments would also be useful to test the full length HTT as preliminary experiments with full length HTT also show that we can get curves with good signal to noise ratio suitable for fitting.

Table 2 Melting temperature of the HTT C-HEAT. The melting temperature T_m is given in °C. The average and standard deviation for each T_m listed was calculated from two trials.

	HTT C-HEAT domain concentration			
Buffer conditions	**0.2 mg/ mL**	**0.1 mg/mL**	**0.05 mg/mL**	**0.025 mg/mL**
1	47.0 ± 1.2	45.3 ± 0.1	45.0 ± 1.0	44.2 ± 0.03
2	45.9 ± 1.5	45.0 ± 0.1	N/A	N/A

Note: N/A= not available

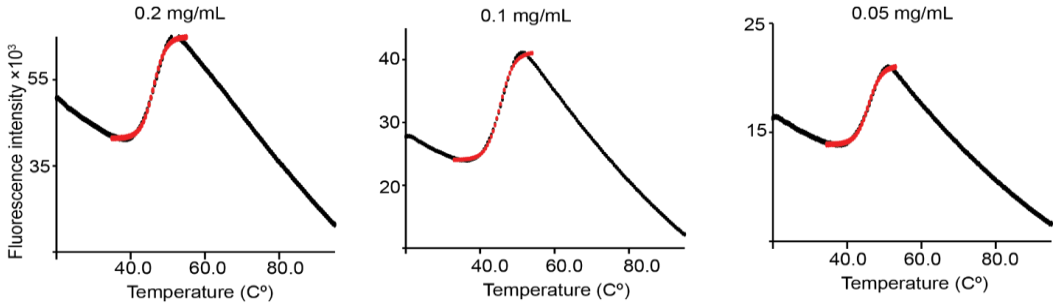

Figure 5 Differential scanning calorimetry curves of the HTT C-HEAT domain. Three different concentrations of the HTT C-HEAT domain are shown. The figure on the right has an unusually high starting fluorescence making the non-linear regression difficult. The figure on the left has overall low fluorescence intensity and like the figure on the right has high starting fluorescence. The figure in the middle where the experiment was run with 0.1 mg/mL of the HTT C-HEAT domain provides the best signal to noise ratio. The black dots represent the signal obtained from the experiment while the red dots are the non-linear regression of the curve.

Differential Static Light Scattering (DSLS)

We performed DSLS experiments with the construct the HTT C-HEAT domain. DSLS experiments in combination with DSF experiments can help us determine the most optimal protein stability which should aid for protein crystallization [10]. Our preliminary results show that the construct the HTT C-HEAT domain at a concentration of 0.2 mg/mL has an increase in the temperature of aggregation (T_{agg}) of 5.4 °C when diluted in buffer 2 which contains no glycerol and and lower salt concentration than buffer 1 (300 mM NaCl in buffer 1 *vs*. 150 mM NaCl in buffer 2) (Figure 6, Table 3). Based on this information, future crystal trays set with this construct will contain the buffer components of buffer 2. Conditions in the absence of glycerol and low salt are the most ideal for crystallography a s high glycerol and high salt concentration may prevent protein crystal formation. Furthermore, our results show that the best concentration of the HTT C-HEAT domain to conduct DSLS experiments at is 0.2 mg/mL. The concentration of 0.4 mg/mL precipitates of solution forming large aggregates which makes data collection difficult. On the other hand, 0.1 mg/mL signal is too low and for this reason makes both data acquisition and fitting the data challenging.

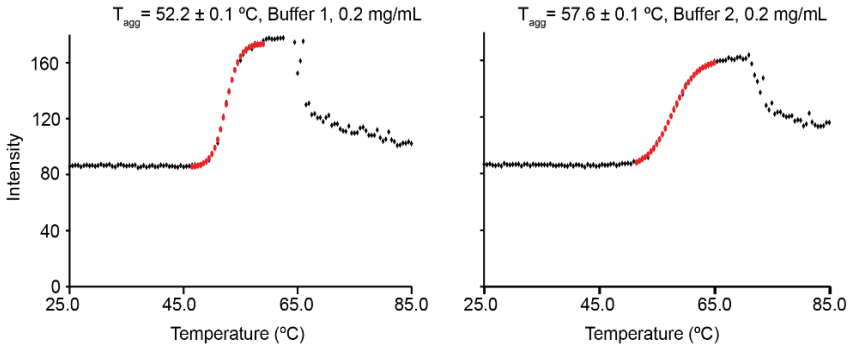

Figure 6 Dynamic static light scattering for the HTT C-HEAT domain 1. The HTT C-HEAT domain at 0.2 mg/mL in buffer 1 (left) and buffer 2 (right) are shown. An increase in the temperature of aggregation of ~ 5 °C is observed at these conditions. The black dots represent the intensity signal while the red dots are the non-linear regression of the curve

Table 3 Temperature of aggregation (T_{agg}) for the HTT C-HEAT domain. The values of T_{agg} for the different experiment conditions are given in °C.

	HTT C-HEAT domain concentration		
Buffer conditions	**0.4 mg/ mL**	**0.2 mg/mL**	**0.1 mg/mL**
1	49.3 ± 0.3	52.2 ± 0.3	57.1 ± 0.3
2	51.4 ± 0.3	57.6 ± 0.1	N/A

Note: N/A not available. Note 2: n = 2

Conclusions

The HTT C-HEAT domain can be purified in large scale and high purity. However, in order to improve the purity of HTT C-HEAT domain, future purifications of this construct will be performed using a Superdex 200 column to determine if we can improve separation and purity.

The increase in T_{agg} indicates that the HTT C-HEAT domain sample is better behaved in a buffer lacking glycerol and less NaCl concentration. Thus, HTT C-HEAT domain will be exchange into this buffer in the future.

Future samples of this construct will be used to determine the most ideal buffer conditions using DSLS and DSF.

References

1. Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A *Protein Identification and Analysis Tools on the ExPASy Server*. https://web.expasy.org/protparam/.
2. Harding, R. J.; Loppnau, P.; Ackloo, S.; Lemak, A.; Hutchinson, A.; Hunt, B.; Holehouse, A. S.; Ho, J. C.; Fan, L. X.; Toledo-Sherman, L.; Seitova, A.; Arrowsmith, C. H., Design and characterization of mutant and wildtype huntingtin proteins produced from a toolkit of scalable eukaryotic expression systems. *Journal of Biological Chemistry* **2019,** *294* (17), 6986-7001.
3. Consortium, S. G. SGC screen. https://www.thesgc.org/sites/default/files/fileuploads/SGC-I_Screen.xls.
4. Structural genomics consortium, T., Canada Redwing screen. https://www.thesgc.org/sites/default/files/fileuploads/Red-Wings-Screen-20050912.xls.
5. Research, H. Peg/Ion screen. https://hamptonresearch.com/documents/product/hr009484_binder1.pdf.
6. Senisterra G, M. E., Yamazaki K, Hui R 2004.
7. Senisterra G, H. R., Vedadi M 2005.
8. Dong, A. P.; Xu, X. H.; Edward, A. M.; Mcsg; Sgc, In situ proteolysis for protein crystallization and structure determination. *Nature Methods* **2007,** *4* (12), 1019-1021.
9. Wernimont, A.; Edwards, A., In Situ Proteolysis to Generate Crystals for Structure Determination: An Update. *Plos One* **2009,** *4* (4).
10. Vedadi, M.; Niesen, F. H.; Allali-Hassani, A.; Fedorov, O. Y.; Finerty, P. J.; Wasney, G. A.; Yeung, R.; Arrowsmith, C.; Ball, L. J.; Berglund, H.; Hui, R.; Marsden, B. D.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A. M., Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** *103* (43), 15835-15840.