Experimental Optimization for Investigating Theoretical Protein-RNA Interactions of the Huntingtin Protein in the Post-Translation Regulatory System

William Brown
New York University
Graduate Biology Department, GSAS
April, 2010

INTRODUCTION

Huntington's disease (HD) is a late onset neurodegenerative disease that is a genetically dominant disorder. The pathogenesis of HD originates with the mutant Huntingtin protein (Htt) that is produced when a trinucleotide CAG repeat is expanded in the *IT15* gene, which contains the Htt coding sequence (1). This repetitive repeat codes for a polyglutamine region (*PolyQ region*) that normally contains less than 27 glutamines in the wild type Htt protein but when exceeding 36 residues results in pathogenesis. Recent investigation has shown that Htt associates with a large cadre of proteins through protein-protein interactions (2). Recent studies have also intimated the potentiality of Htt interacting with RNA molecules through these protein complexes in the post-transcriptional regulatory system (3). Developing a methodology that allows for direct identification of RNA targets by Htt would greatly facilitate our understanding of post-transcriptional gene regulatory networks (4).

The protein Htt can be found trafficking in neuronal vesicles and has been implicated in many cellular functions. However the most pertinent to this study is the localization of Htt in RNA complexes involved in RNA processing. In these complexes Htt has been found to associate with numerous proteins through co-immunoprecipitation (co-IP) experiments (3). One such class of proteins that was shown to associate with Htt are the Argonaute (Ago) proteins that function in microRNA (miRNA) mediated translational regulation of mRNA molecules. Another class of proteins pertinent to the research performed herein is the TNRC proteins, which were similarly identified in the co-IP experiments performed by Savas to associate in the processing complexes. In these associations it is hypothesized that Htt is involved in the repression of mRNA transcripts. However the precise interaction between Ago and Htt has yet to be clearly elucidated. It is not known whether the proteins associate through direct protein to protein interaction or if they are associating indirectly via binding of the mRNA transcripts.

Therefore if an experiment were devised in which Htt and Ago could be coexpressed *in vivo* and retrieved via isolation techniques then the precise nature of the molecular association could be investigated. However, coexpression of the Htt and Ago proteins as well as the protein used for a positive control (TNRC6C co-transfected with the Ago construct) requires a great deal of troubleshooting. Therefore before it can be investigated whether the Htt and Ago association is RNAse sensitive or not, the hypothesis must be tested whether the proteins identified in the co-IP studies can be successfully coexpressed via cellular transfection with plasmid constructs. To this end HEK 293 cells were used in transfection optimization experiments with TNRC6C, Htt59025Q, and Ago2.

Another intriguing question is whether or not there is a specific microRNA element (miRE) within the 3' untranslated region (UTR) of mRNA transcripts that the Htt protein recognizes. Within the 3' UTR there are miREs that are complementary to seed sequences 7-8 nucleotides long in the 5' UTR of miRNA. A bionformatic analysis was performed on data generated by Zhang *et al* from a microarray profiling assay of the transcriptome of embryonic mouse fibrolast cells (EMF) that showed differential expression levels of mRNA transcripts when the Huntingtin *Hdh* gene was knocked out (*Hdh*-KO) (5). When the *Hdh*-KO transcriptomes were compared to that of the wild type it was found that 106 transcripts were down-regulated and 173 were up-regulated at least two-fold with $p < 0.05$.

In the bioinformatic analysis described in this report miREs found by Zhand in the UTRs of the differentially expressed genes were compared to identify common ones that are candidates for specific recognition by the Htt protein. The miREs were examined under the hypothesis that one or more miREs common to a subset of the disregulated genes could be identified that would be specific to interaction with Htt protein. In this model the miREs would provide specificity to the intermediary interaction of Htt protein with mRNA transcripts and miRNAs. This intermediary action performed by Htt would modulate translation of a select subset of mRNAs, targeted by Htt because of the specific miREs. Furthermore, due to the highly efficacious modulation of gene expression resultant from differential polyadenylation these sites were investigated within the differentially regulated genes hitherto discussed (6).

Another experimental method utilized in this thesis was to test for mRNAs and miRNAs that Htt may be interacting with. The method was adapted from an experiment

performed by Wook Chi (7) and in this adaptation crosslinking immunoprecipitation (CLIP) of Flag-Htt59025Q was performed to try and covalently link any miRNA or mRNA that Htt-59025Q is directly interacting with. It was hypothesized that upon comparison of the molecular weights of CLIP Htt and non-crosslinked Htt in a SDS-PAGE electrophoresis would show Htt protein bound RNA molecules because of the difference in size produced from UV-induced crosslinking.

MATERIALS AND METHODS

Much troubleshooting and optimization of coexpression was required before the hypothesis of whether or not the association between Ago2 and Htt is mediated via an mRNA transcript and as such is RNase sensitive. The methods herein describe the optimization necessary to perform such an experiment. HEK293 cells were transfected for co-expression with plasmids containing myc epitope-tagged wild type Htt (myc-Htt59025Q) and flag epitope-tagged Ago2 (flag-Ago2). The Htt construct Htt59097Q was used as a positive control. Other controls include myc-TNRC6C and myc-Htt59025Q transfected alone in order to show that there is no non-specific pull down of these proteins in the Co-IP. After a two day incubation period cells were lysed and a protein extraction was performed (see experimental descriptions below labeled in bold print). Proteins of specific interest were then detected using Western Blotting and/or isolated using Immunoprecipitation.

The methodology for the CLIP co-experiment is similar to the Ago-Htt RNase experiment in regards to transfection, protein extraction, RNase treatment, IP and Western Blotting except for the addition of a UV irradiation step. If Htt-59025Q is

interacting with RNA molecules *in vivo* and is crosslinking to them upon UV exposure then when the RNase treated CLIP Htt-59025Q is compared to the untreated CLIP Htt-59025Q there will be a difference in size. In order to show that any potential shift was due to bound RNA, select samples of the CLIP Htt-59025Q were treated with RNase A (1:100). The incubation was such that the reaction was performed to completion and the RNA was fully digested. The RNase treated sample will have a normal Htt-59025Q MW because the RNA bound molecule will have been digested. While the untreated sample will be ~8 kD or ~16 kD larger if miRNA or mRNA is bound, respectively. The differential sizes of the CLIP Htt-59025Q treated with RNase A and the untreated CLIP Htt-59025Q will support the hypothesis that it is indeed RNA that Htt-59025Q is binding.

To identify a microRNA binding element (miRE) that is specifically recognized by the Htt protein the data produced from a differential gene expression analysis project was used to identify genes that had differential expression under the wild type Htt or the mutant Htt protein. In the study by Zhang *et al* genome-wide profiling using microarray was performed on mouse embryonic fibroblast cell lines: four containing a single Htt gene (*Hdh*-HET) and four with the Htt gene deleted (*Hdh*-KO). The 3' UTRs of these differentially expressed genes were then queried in the TargetScan mouse database to identify conserved miREs.

Plasmids. The plasmid constructs were obtained from stock supply. The plasmids were constructed using standard molecular techniques involving PCR amplification and cloning with vectors. Two plasmids were obtained from other labs: TNRC6C and SYMPLK.

Cell Culture. Cells were grown in DMEM media (obtained commercially) with Fetal Bovine Serum (FBS) at a 10% concentration and 1% PSQ (antibiotic and glutamine).

Transfection. All transfections were performed using Lipofectamine 2000 following the manufacturers protocol.

Protein Extraction. Cells were washed two times with PBS and scraped from the plates in 1mL of PBS. Cells were pelleted at low centrifugation: 5 minutes, 2700 RPMs, at 4°C. The PBS was removed and 100µL of TNEN buffer (10mM Tris-HCl pH7.8, 150 mM NaCl, 1 mM EDTA, 1% Nonidet P-40) with protease inhibitors was added. This was incubated on ice for 20 minutes. The Lysate was cleared by centrifugation: 20 minutes, 12000 RPMs, at 4°C. The clarified lysate was transferred to a new Eppendorf tube.

Protein Quantification: Quantitation of the protein extracts was performed using a Bradford assay with a spectrometer as per the Bradford protocol.

Immunoprecipitation. HEMG buffer (25 mM Hepes-KOH pH7.9, 0.1 mM EDTA, 12.5 mM MgCl2, 20% glycerol) containing 300 mM KCl was added to the cellular lysates and 5% input was taken for SDS/PAGE electrophoresis. Immunoprecipitations were performed using α -Flag M2 (Sigma) beads. The α-Flag beads were washed with 300 mM HEMG 3 times and an equal volume of 300 mM HEMG was added to create a 50/50 solution. The 50/50 slurry of 40 µL washed beads was added to the protein extractions. The IP was then washed in the same manner as the beads except for a final wash in HEMG with 150mM KCl. All excess HEMG buffer was removed and 20 µL of 6X SDS loading buffer was added. Proteins were eluted by boiling in 6X SDS loading buffer for 3 min before separation by SDS/PAGE.

SDS/PAGE and WB Analysis. All gels were 10% poly-acrylamide except for those used in the CLIP experiments which were 7% poly-acrylamide. Proteins transferred to nitrocellulose membranes by Western blotting were imaged on an Odyssey infrared imager and Odyssey software was used for all imaging and analysis.

Bioinformatic Analysis. For the bioinformatic analysis 168 annotated genes that showed at least a 2-fold difference of expression were selected from 279 genes found to be disregulated in the Zhang study (5). These sequences were then queried in the Targetscan database (8) and the major conserved miRs were recorded. These were cross-referenced with the Miranda algorithm (9) to increase confidence and narrow down potential targets. Then the 3'UTR sequence of the mouse mRNA transcripts had to be found, which was accomplished using the dbUTR (it is now recommended to use The PolyA Cleavage Site and 3'-UTR Database (PACdb, 10). The 3'UTR sequence could then be entered into the estParser algorithm which would query the sequence for potential poly(A) signals (6). The most common miR was identified for the Targetscan query as well as the miRanda query. A statistical analysis was performed in which the numbers of genes with known miRs were compared to the number that did not contain a miR within the Targetscan or miRanda database. The number of genes with known Poly(A) signals were similarly compared to all the genes to generate a percentage of representation.

Clip Procedural Outline. For the CLIP experiment the basic procedural outline is as follows:

1. Obtain HeLa cells that express Flag-Htt59025Q.
2. Include HeLa cells that do not express Flag-Htt59025Q as a control.

3. Expose cells to UV irradiation to try and crosslink any potential Htt complexes: Grow cells in a 100 or 150mm plates, rinse with PBS, and place in a Stratalinker with the cover removed. Irradiate for 400mJ/cm^2 and a second time for 200 mJ/cm^2 (7).
4. Immunoprecipitate the Flag-Htt from crosslinked cells (+XL), non-crosslinked cells (-XL), and regular HeLa cells using Flag ab beads.
5. Split the immunoprecipitated Flag-Htt59025Q +XL into two Eppendorf tubes and incubate one with RNase A (1:100) for 30 min at room temperature while the other is left untreated. The RNase A treated CLIP (+XL) Htt-59025Q serves as a control to show RNA interaction with Htt-59025Q is a potential source of any observable shift in MW.
6. Repeat for the immunoprecipitated Flag-Htt59025Q –XL. There should be no difference between RNase A treated Htt59025Q –XL and the non-treated –XL sample.
7. Perform SDS-PAGE and nitrocellulose transfer with anti-flag probing: The gel will be a 7% gel and will be run until the 75 kD marker is at the bottom so as to increase the likelihood of detecting the shift.
8. Compare CLIP Htt-59025Q to unbound Htt-59025Q using the Odyssey Infrared imager. Compare CLIP Htt treated with RNase A to untreated CLIP Htt59025Q.

RESULTS

Expression via Transfection -

Initially verification had to be obtained that the proteins of interests could be successfully expressed via transfection. To this end lysates from HEK293 transfections were analyzed with WB immuno-blotting to detect the expression of the specified proteins. It was observed that the proteins were being expressed in the transfected HEK293 cells (Figure 1). Myc-Dicer was weakly detected but was not selected for further experimentation. TNRC6C had a faint band which we hoped to strengthen through subsequent optimization. Furthermore we were interested in seeing if Myc-Htt59097Q would express, however in subsequent experimentation Myc-Htt59025Q was used so as to assess the activity of the functional truncation from wild type Htt.

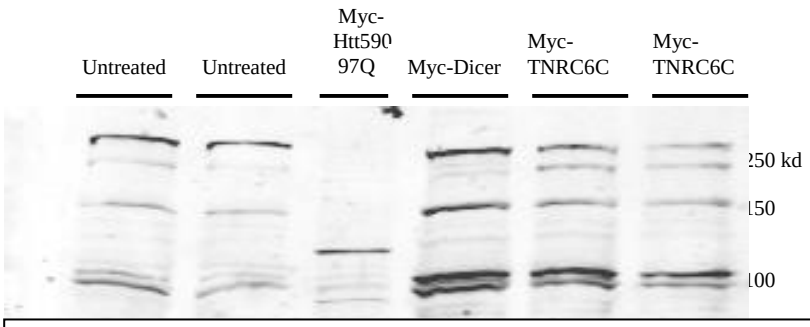

Figure 1. Detection of protein expression after transfection.

Western blot analysis of a transfection performed on HEK293 cells with Myc epitope-tagged Htt59097Q, Dicer and TNRC6C. Lysates were produced from cultures prepared at 1 million cells and for all samples 30µg was loaded. After probing with α-Myc ab all Myc-tagged proteins were detected. The expanded poly(Q) region protein Htt59097Q is 125 kD, Dicer is 218 kD, and TNRC6C is 176 kD. All protein bands were observed at their respective size. TNRC6C was loaded in duplicate. Dicer is very weakly detected. Untreated samples were included and can be seen at the far left.

Co-expression via Transfection -

After verification of successful expression from transfection it was necessary to assess if the proteins could be co-expressed. Immunological probing was performed with Flag and Myc antibodies to detect co-expression of Myc-Htt59025Q with Flag-Ago2 and Myc-TNRC6C with Flag-Ago2. In the Htt-59025Q + Ago2 both proteins expressed at readily detectable levels (Figure 2). For TNRC6C + Ago2 both proteins expressed however TNRC6C was weakly detected (compare the discontinuous band of the TNRC6C in the co-expression with the adjacent band of TNRC6C transfected alone, which has a much stronger signal). The expression level of Myc-Htt59025Q is relatively suppressed in the co-expression with Ago2 as well which can be observed by comparing the strength of the bands between the co-expressed Htt59025Q and Htt59025Q alone.

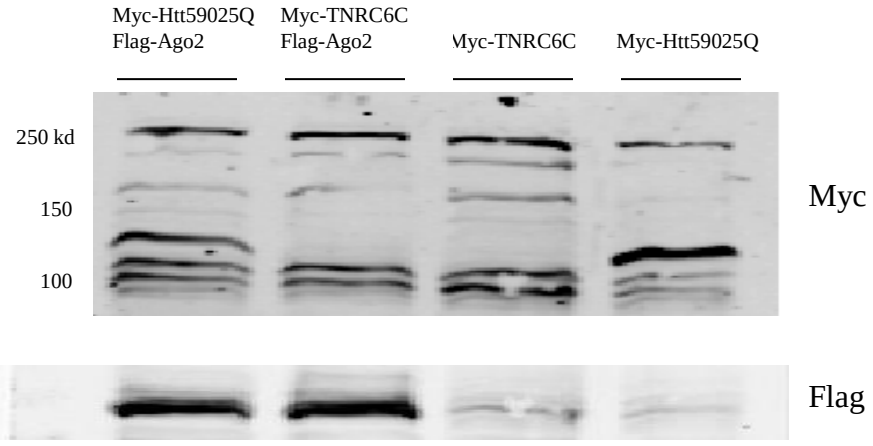

Figure 2. Detection of co-expression after transfection.

The proteins Htt59025Q and Ago2 were co-expressed in HEK293 as were TNRC6C and Ago2. Lysates were produced from 1 million cells and for all samples 30µg was loaded. The wild type Htt protein Htt59025Q was used in the co-expression. The top image is probed for Myc and the bottom image is probed for Flag. In the Top image TNRC6C is observed in the single transfection and in the co-transfection with Ago2. Htt59025Q obtained the same results; being detected in single expression and co-expression. The bottom image shows detection of Flag-tagged Ago2 protein. It is detected in the lanes for co-expression with TNRC6C and Htt59025Q.

Transfection Optimization -

With the verification of co-expression of transfected DNA an attempt was made to increase the expression levels of the transfected proteins, specifically TNRC6C. Protein expression was examined after more than doubling the amount of cells used in the transfection and increasing the amount of Lipofectamine. Unfortunately at this stage the amount of DNA used for the transfection was held constant at 1.5µg. There was strong expression of Flag-Ago2 and Myc-Htt59025Q (Figure 3). However TNRC6C was not detected.

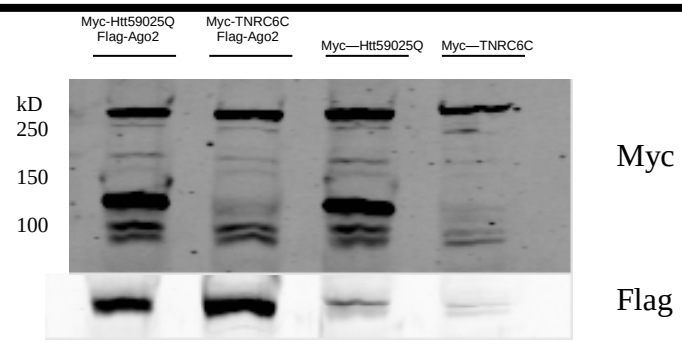

Figure 3. Experiment to increase the yield of TNRC6C.

Transfections performed with 2.5 million cells per plate and 75µg of lysate loaded for each sample. There is no expression of TNRC6C. The other proteins Htt59025Q and Ago2 expressed.

From the results of the initial adjustments to the experimental methodology a second optimization attempt was made in which the number of cells used in the transfection was decreased in order to mitigate the high stress of transfection. The same

analysis procedure was performed however the expression level of TNRC6C became the primary focus. Under the new conditions TNRC6C expression increased to a useable level (Figure 4). However in a rather retrogressive fashion expression of TNRC6C with Ago2 ceased being detected.

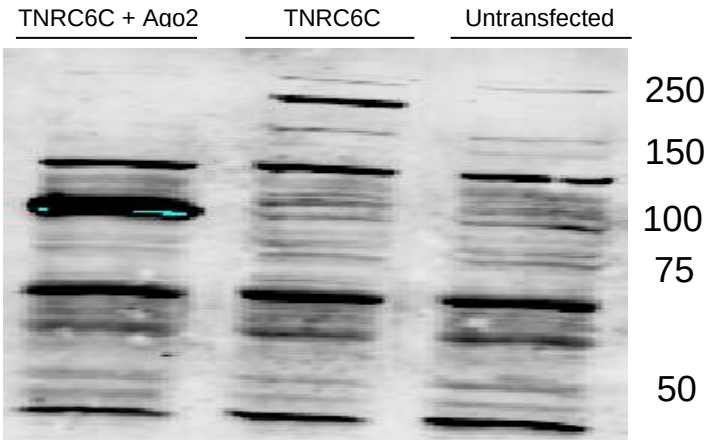

Figure 4. Analysis showing TNRC6C expression from transfection.

TNRC6C expression is observed for the single TNRC6C transfection but not in the co-expression with TNRC6C and Ago2 transfection. The transfection was performed at 1.5 million cells and 50µg of lysate was used for the SDS-PAGE. Untransfected cells are included in the right lane as a control.

Co-immunoprecipitation -

With an increased level of TNRC6C expression, but no detectable TNRC6C co-expression with Ago2, an experiment was devised to see if TNRC6C would bind to Ago2 *in vitro* and co-immunoprecipitate. As a control assessment Htt-59025Q was similarly mixed with Ago2 so that this could be compared to a sample in which Htt-59025Q and Flag-Ago2 were co-expressed. The IP samples were run in a WB with nitrocellulose transfer and immunologically probed. The Flag-Ago2 protein was recovered in all

samples however no other proteins were detected in the IPs (Figure 5). While Htt-59025Q was weakly detected in the input samples no TNRC6C was detected at all.

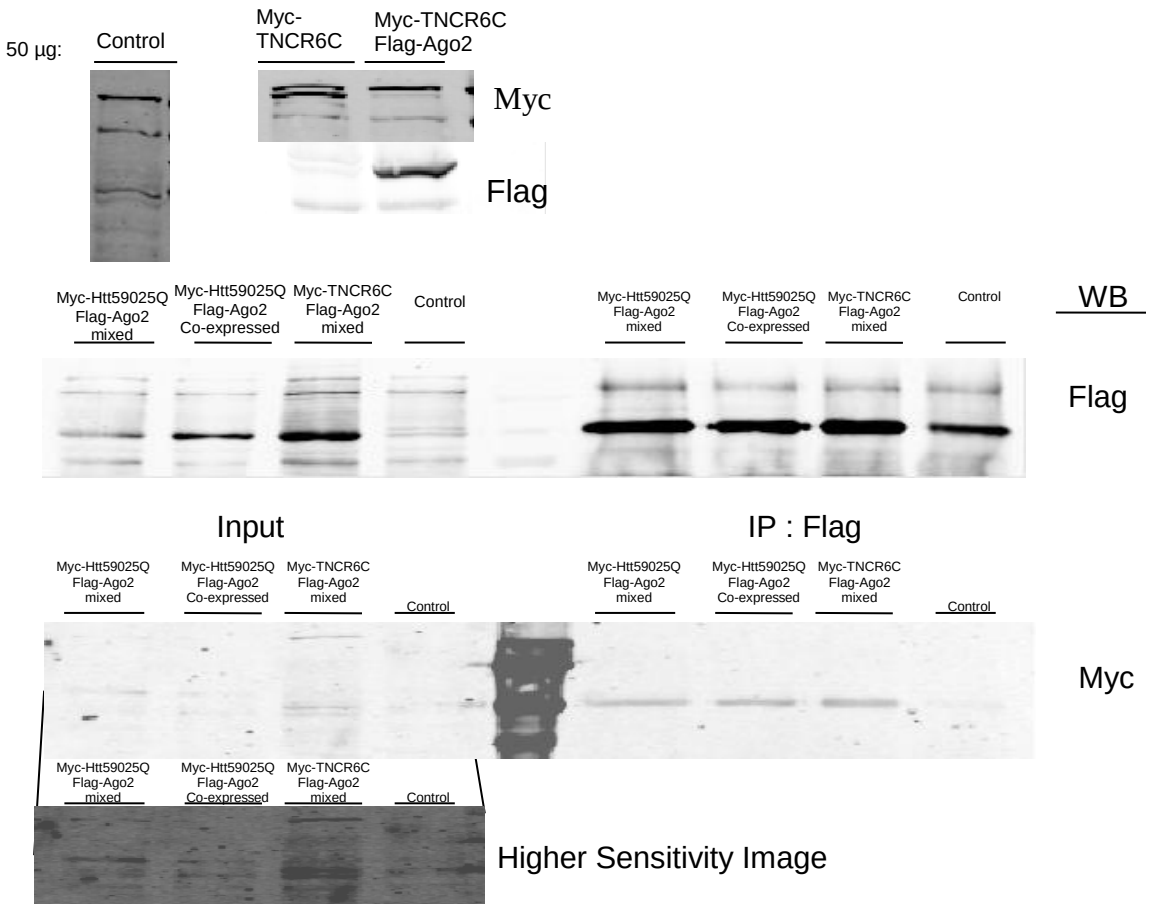

Figure 5. Analysis of IP of TNRC6C mixed with Ago.

Experiment to detect TNRC6C in IP

The top image shows the expression of TNRC6C from transfection, verifying that TNCR6C is being expressed in the HEK293 cells. The middle image is detection of Flag, which detects Flag-Ago from inputs and IPs. The bottom image shows detection of Myc tagged proteins, and primarily in the higher sensitivity image the band for Myc-Htt59025Q is observed in the input. No TNRC6C is detected in the input or the IP samples. Htt59025Q is seen in the input but not in the IP. For the IP 370 micrograms of protein was used and 5% of that (17 µg) for the inputs.

Increasing the Amount of TNRC6C for the IP –

Since TNRC6C was not detected at all in the input an attempt was made to increase the yield of TNRC6C such that sufficient protein levels could be used within the IP. This time differential amounts of DNA for transfection were used. Doubling the amount of DNA used for the TNRC6C transfection increased the expression level of the protein (Figure 6). The lysate from the 1.5µg TNRC6C DNA transfection had approximately the same concentration of the lysate from the 3.0µg TNCR6C DNA transfection: 3.9µg/µL and 4.4µg/µL respectively. However as can be seen the strength of the TNRC6C band from the 3.0µg transfection is much greater in magnitude than the TNRC6C band from the 1.5µg transfection, which is barely detectable at all.

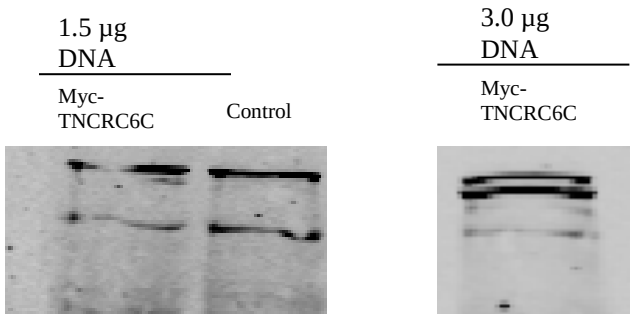

Figure 6. Expression of TNRC6C from transfection using different quantities of DNA.

The left image shows the results of transfection using 1.5µg of TNRC6C DNA. The image at right shows the results of transfection using 3.0µg of DNA.

Experimentation for Htt protein-RNA Interaction

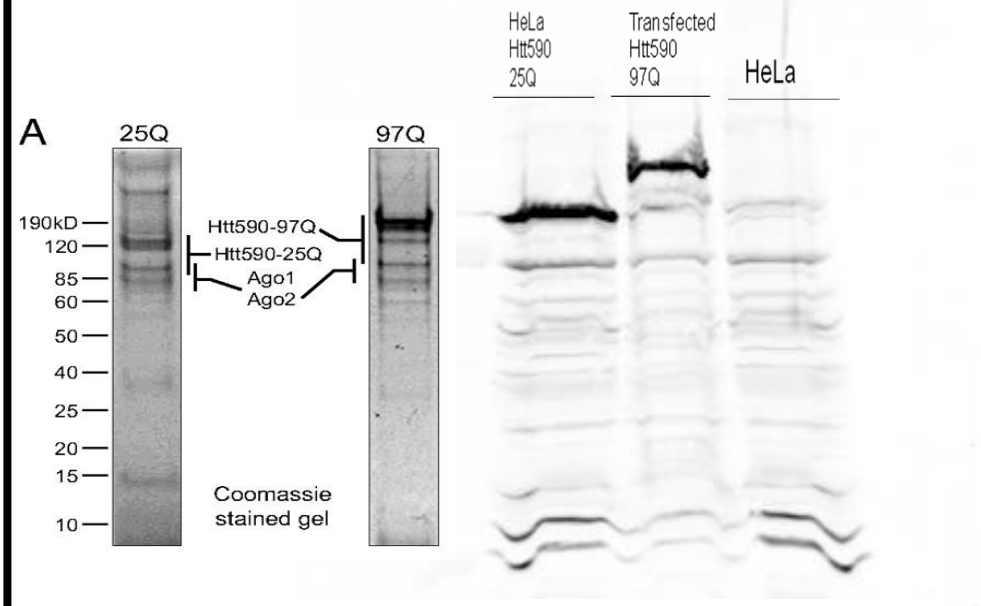

Figure 7. Verifying constitutive expression of the exogenous plasmid construct HeLa Htt59025Q.

In the first lane is the plasmid construct of interest Htt59025Q, which has migrated to a position below the larger construct Htt59097Q in the middle lane. To the left of the gel image is a key to the molecular weight of Htt59025Q and Htt59097Q, allowin for a comparison of the relative migration positions of the two proteins as compared between the key and the gel. Untransfected HeLa extracts were ran in parallel (third lane) as a negative control.

Considering that Htt59025Q is an artificial plasmid construct and was introduced into the HeLa cell line through transfection it was initially necessary to verify that the HeLa Htt59025Q cells indeed expressed the exogenous plasmid in a constitutive manner. Analysis of the Western blot image showed a thick band within the Htt59025Q extract at the correct molecular weight, indicating successful expression of the exogenous

Htt59025Q within the transfected HeLa cells (Figure 7). The Htt59097Q plasmid construct was included as a positive control; it was not subsequently worked with after this analysis.

Assessing the Functionality of the Flag-epitope Fused Htt59025Q Construct

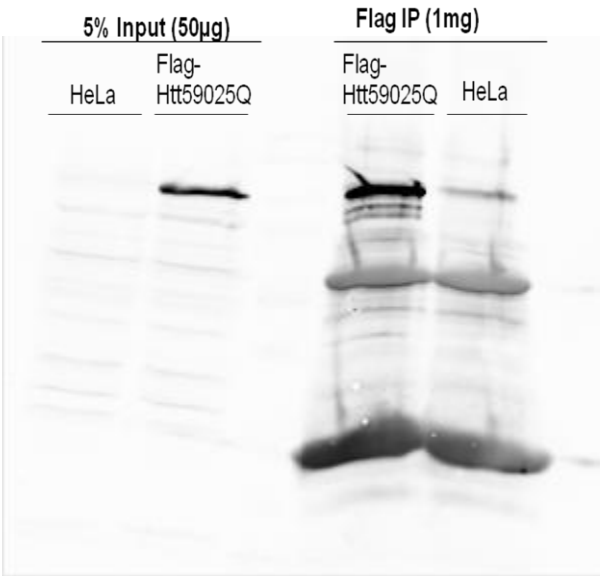

Figure 8. Western blot results for the trial immunoprecipitation of Flag-Htt59025Q.

On the right side of the image the IP for Flag-Htt59025Q including a HeLa negative control can be viewed. There is a slight background band in the HeLa IP, but as can be seen in the input (left side of image) this is not a specific protein. The IP was performed using 1mg of lysate, of which 5% was taken for input.

With the verification of Htt59025Q expression it was then necessary to insure that the Flag-fusion epitope was functioning correctly and as such that the Flag-Htt59025Q protein could be immunoprecipitated. In the lanes labeled Flag-Htt59025Q there was a thick black band at the uppermost portion of the extract in the input and the recovery product from the IP, which corresponded to the molecular weight of Htt59025Q (Figure

8). These results implied that the Flag epitope on the exogenously expressed Htt59025Q plasmid construct was indeed functional and that the protein was being successfully immunoprecipitated.

Commencing UV-induced Cross-linking of Htt59025Q and Immunoprecipitation

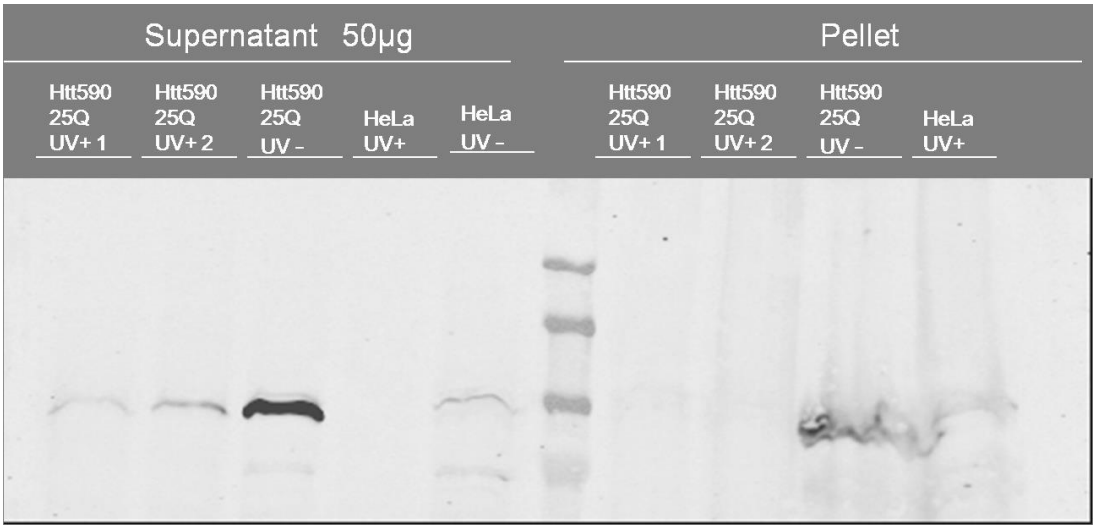

Figure 9. Analysis of Htt59025Q UV-induced cross-linking without immunoprecipitation; searching for aggregates.

On the left side of the image are the samples of the supernatant from the cellular extract. The first two lanes sub-labeled Htt59025Q UV +1 and Htt59025Q UV +2 contain the UV-treated samples. The lane sub-labeled Htt59025Q UV- that contains the thick black band is the untreated Htt59025Q sample. Untransfected HeLa samples that are UV-treated and untreated are included as negative controls. In the center of the image is the molecular ladder, the uppermost band is 250 Kd, below it is 150 Kd, then 100 Kd, and the faint band at the bottom is 75 Kd (value of Kd markers not indicated on image). On t he right side of the image is solubilized material of the pellet from the same cellular extract as the supernatant on the left side of the image. The amorphous black material is RNA.

The initial UV treatment revealed that a large portion of the Htt59025Q molecules were being affected by the electromagnetic energy exposure however no complexes were

being detected in the Western blot analysis. Thin bands of the Htt59025Q protein in its native migration position were detected in the UV-treated samples, which were comparably less intense than the corresponding band for the untreated Htt59025Q sample (figure 9). A portion of the pellets of cellular debris from the protein extraction process were solubilized and interrogated for potential Htt59025Q aggregates. Furthermore a subsequent Western blot was run to check the wells for any aggregates that may have been too large to navigate through the porous acrylamide (image not shown). However no proteins were detected in the wells.

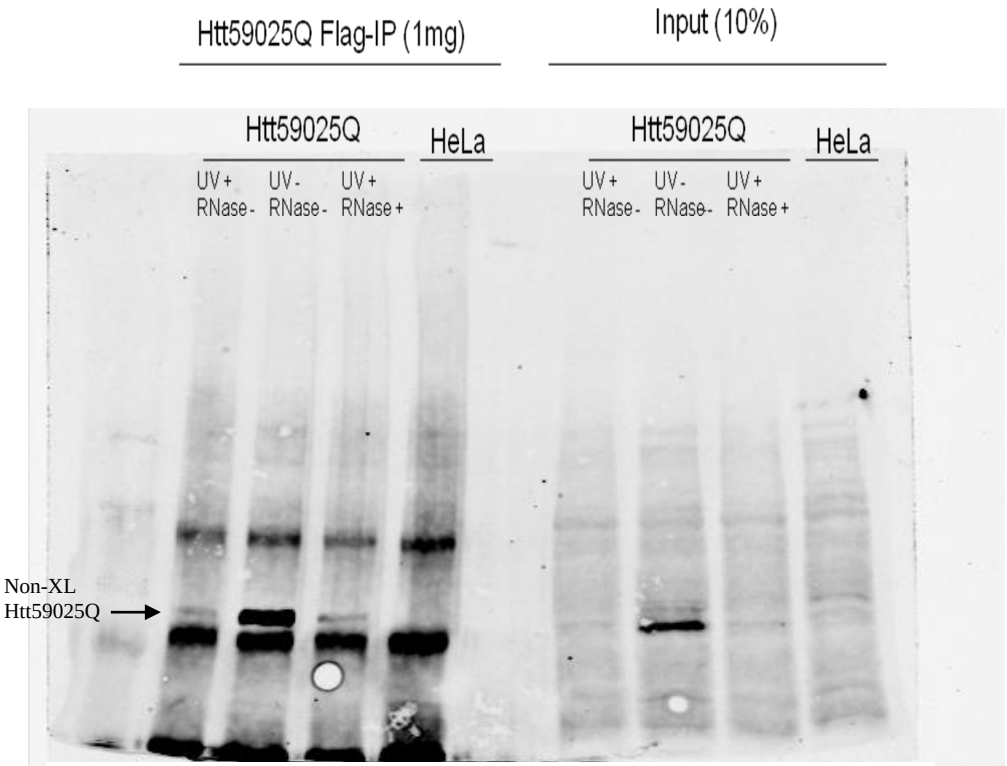

Figure 10. Western blot results from the execution of the CLIP method as applied to the Htt59025Q HeLa cells.

No shift in molecular weight was observed in the Htt59025Q UV+ RNase- band,

On the left under the header Htt59025Q Flag-IP (1mg) is the product of the CLIP-treated cells. The lane labeled UV- RNase- under the sub-header Htt59025Q contains the migration band for Htt59025Q under non-cross-linked conditions. The corresponding bands are observed in the XL samples, both RNase treated and untreated, with both exhibiting mitigated signal intensity. Robust background bands are observed both below and above the specifically IP'd protein. Within the Input only the untreated sample exhibits an Htt59025Q migration pattern, which is 100µg of lysate.

which was the CLIP-treated sample (Figure 10). Similarly no shift was observed in the Htt59025Q UV+ RNase+ band. From a qualitative assessment of the bands from the CLIP-treated samples it appeared to the observer that the Htt59025Q UV+ RNase+ band had stronger signal strength than the non-RNase treated band. Therefore the signal strengths of the bands were quantitatively measured using the imaging software (Figure 11). After normalization the bands for both Htt59025Q UV+ RNase+ and Htt59025Q UV+ RNase- had equal signal intensities of 1.66 units.

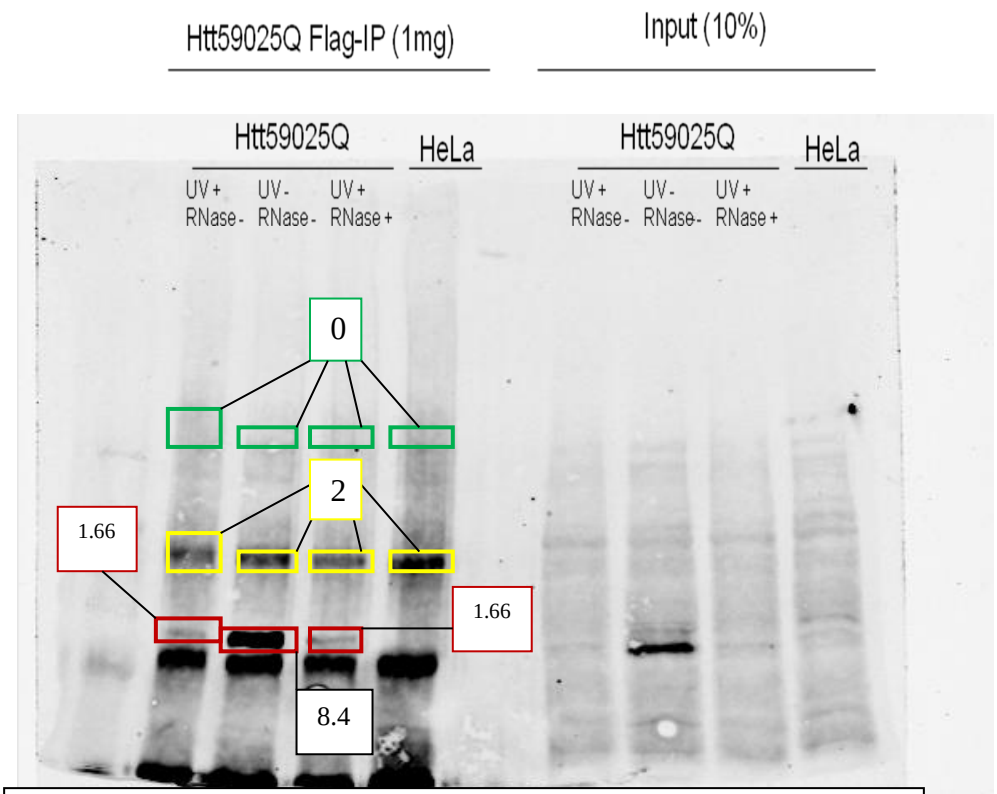

Figure 11. Signal normalization and quantitation of Htt59025Q signal strength.

- Green boxes –The green boxes are the background normalization: the background was set to a signal strength of 0.
- Yellow boxes – The yellow boxes are signal normalization to a background band, which is a protein non-specifically recovered during the IP: the background band signal strength was set to an arbitrary unit of 2.
- Burgundy boxes – The burgundy boxes are the signal strength of the Htt59025Q bands after normalization: both XL-Htt59025Q bands had signal intensities of 1.66. The untreated Htt59025Q band had a value of 8.4 unit

Attempting a Determination of Optimal Radiation Dosing Parameters

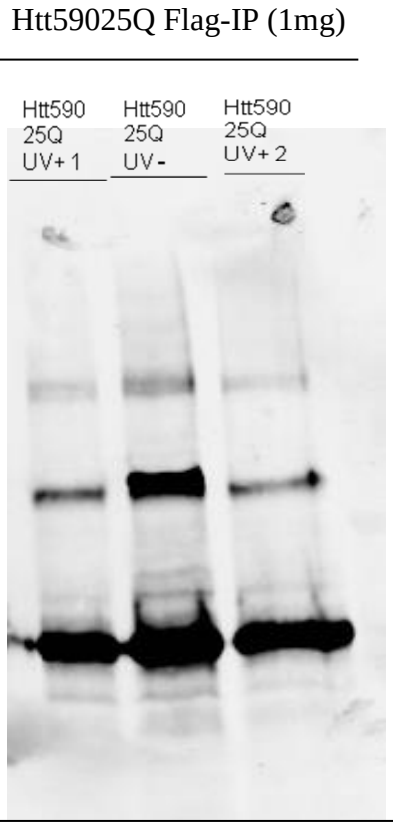

Figure 12. Adjusting the UV cross-linking parameters.

- In the first lane Htt59025Q UV +1 the normal amount of radiation was administered: 500 mJoules/cm^2 then a second dose of 250 mJoules/cm^2.
- Htt59025Q UV +2 received half the normal exposure: 250 mJoules/cm^2 followed by 175 mJoules/cm^2.
- The center lane is untreated HeLa Htt59025Q immunoprecipitated protein.

Two different energy densities were used for the samples Htt59025Q UV+1 and Htt59025Q UV+2 (Figure 12). Each received a different dosage of radiation; with the +1 sample receiving the normal dosage and the +2 sample receiving half of the normal

dosage (see Figure 12 for the specific energy densities administered). There was no difference in the signal intensities between the differentially treated samples. Of importance to note is that the signal intensity of the untreated Htt59025Q samples is several times stronger than that of the treated samples, although uniform cell densities were utilized.

A Bioinformatic Analysis of microRNA-binding Elements

After querying the TargetScan database 41% of the genes were found to contain miRs (Table 1). The average number of elements was 1.43 and the most common recognition element was miR-19. Of the 168 genes miR-19 was found in 13 of the total genes. The percent of up-regulated genes from the initial data set was 59%. After cross-referencing the data set with miRanda the total number of genes was 69 and since only the genes with miRs were preserved 100% of the genes in the new data set had miRs. The average number of elements for the 69 genes was 3 and the most common element was miR-30a/30a-5p/30b/30b-5p/30cde/384-5p being found in 17% of the genes.

TABLE 1. Statistical analysis of TargetScan query results and miRanda query results. The data for % with poly(A) signal is only available for miRanda because only this refined data was used for querying the estParser database.

	TargetScan	**miRanda**
# of genes	168	69
% with miR	41%	100%
average # of elements	1.43	3
most obs miRNA element	miR-19	miR-30a/30a-5p/30b/30b-5p/30cde/384-5p
% with poly(A) signal	N/A	11%
% of genes up-regulated	59%	57%

The percent of up-regulated genes was 57% and the percent with poly(A) signals was 11%. Three of the genes had two poly(A) sites while the rest had only one. When the three common elements miR-19, let-7, and miR-30 were analyzed they were found to be associated with 25% of the genes contained within the data set and of these 65% are up-regulated. Interestingly all three miRs are found within *Cpeb3*, a cytoplasmic polyadenylation element binding protein and this gene is down-regulated. The gene *Cpeb3* has 16 miRs, which is 3 times more than the average number of miRs typically found within the 3'UTR of the transcripts investigated.

DISCUSSION

Understanding the normal functions of the wild type Htt protein is an essential component of understanding what leads to disease in the mutant Htt. Elucidation of this normal functioning will not only lead to a more sophisticated understanding of HD but will allow for novel and more efficacious treatments of HD. Beyond the immediate reward of treating the pathology there is also the benefit of advancing our understanding of the molecular biology of the cell. Benefits derived from such knowledge are more subtle and not as obvious as curing the disease. So to unveil more of the incomplete picture of Htt, its association with Ago2 was researched and the question of whether or not it has an RNase sensitive association with Ago2 was addressed. Furthermore the mechanism by which Htt recognizes mRNA transcripts was examined to see if there is a specific miR within the 3' UTR region that is being targeted by Htt.

The experiment to probe the nature of the association between Ago2 and Htt could not be successfully executed because one of the control proteins TNRC6C could not be co-expressed at sufficient levels for detection. TNRC6C expression levels seemed to decline during attempts of optimization. The apparent regression was due to the nature of the optimization, in which it was thought that TNRC6C yield could be improved by increasing the concentration of cells involved in the extraction process. Logically this would be successful if the same conditions were applied that were used during the initial transfections which yielded adequate TNRC6C levels, so that the only variation would be the scale at which transfection was performed. However by increasing the concentration of cells while holding the volume static the cells were subject to higher stresses and inhospitable conditions which decreased protein yield.

The problem was further exacerbated by increasing the time of expression for the exogenous TNRC6C, which at the non-physiological level involved during expression from the foreign plasmid is toxic to the cell. Furthermore the attempt at co-expression with Ago2 is an ambitious one to begin with, as Ago2 is a protein involved in translational repression. So a condition in which exogenous Ago2 expression is highly elevated will in any case not be very conducive to the parallel increase in expression of any other proteins.

Although it was assumed that varying the DNA amounts used during transfection would have no effect an experiment was finally performed in which this was done. The results showed that TNRC6C levels were increased by administering more DNA during transfection as compared to the normal amount that was used during all previous experimentation. Doubling the amount of TNRC6C plasmid DNA used in the

transfection from 1.5 µg to 3.0 µg effectively increased protein yield as detected from the higher signal intensity in the WB electrophoresis. Upon further investigation it was found that the Lipofectamine protocol actually suggests 8µg of DNA for a 60 mm transfection, roughly a five-fold increase over the standard amount administered during incubation with the liposomal vector.

Obfuscation was fostered in part due to the high performance of the Htt59025Q and Ago2 plasmid constructs. The high rates of successful transfection and efficiency of transcription allows for modest amounts to effectively yield relatively high expression of the proteins. And perhaps it seemed more reasonable to perform all transfections using equiquanta of the three plasmids. However due to the high cost of lipofectamine transfection subsequent investigation was terminated. Future experimentation would most certainly be successfully executed using 8µg of the TNRC6C plasmid when incubating with the lipofectamine vector. Furthermore since Ago2 is a translational repressor, creating a large differential by lowering the amount used in the cotransfection with TNRC6C would greatly increase the yield of TNRC6C expression, facilitating the experiment even further.

The results of the transcriptomic comparison between Htt-KO cells and non-KO cells demonstrated pellucidly that Htt is involved in the mRNA transcript processing pathway, as a great multitude of transcripts were either up-regulated or down-regulated upon removal of the Htt protein. The bioinformatic analysis identified specific miRs that may contain a canonical binding motif that Htt is recognizing in its regulation of mRNA transcripts. To test this hypothesis empirically a dual-luciferase assay could be designed to show the interaction of Htt with one of the common miRs identified in the

bioinformatic analysis. The Bioinformatic analysis shows that we have at this time an incomplete representation of the elements involved in the post-transcriptional pathway as our databases could only identify miRs for 41% of the genes and poly(A) signals for 11%. The probability that the other genes do not contain these regulatory elements is not high. Therefore a Bioinformatic analysis with the stated objectives such as the one performed herein is at this time unfeasible because more information and data is required of the genes.

Evolving the path of research followed up to this point a further investigation of Htt-RNA interaction was devised using the CLIP method. This is an efficacious method of studying protein-RNA interaction because only molecules within a few angstroms of each other will be covalently bonded by the ultraviolet electromagnetic radiation. The covalent bonds form cross-linking structures between the co-associated molecules making the resulting complex highly robust. Therefore strenuous applications such as immunoprecipitations can be executed without disrupting the molecularly-chimeric complexes. The Htt protein has the potential to form ternary complexes due to its hypothesized function as a chaperone that binds specific miRNA molecules and recognizes their mRNA counterpart, bringing the two into association. However it is not presumed that Htt would be performing such functions independently, as the co-immunoprecipitation studies show that Htt associates with a vast array of proteins.

Although it was hypothesized that the CLIP method would be successful in identifying Htt-RNA complexes they remained elusive throughout this investigation. Western blot analysis revealed that cells exposed to UV irradiation had diminished levels of un-associated Htt protein, so transformations were taking place. However it may

simply be that Htt is actually not amendable to the CLIP method. This could be due to the association of Htt with the dendritic myofibrils, were it transports molecules to the synapses. If the majority of intracellular Htt is interacting with the myofibrils then UV exposure will cross-link Htt to the cytoskeleton, in which case it will most likely not be retrievable via protein extraction and immunoprecipitation.

In the experiment in which different dosages of UV irradiation were administered there was no resulting change in Htt59025Q signal intensity. Since the different dosing parameters did not affect the signal strength then the band must represent the amount of un-associated intracellular Htt59025Q. This further supports the hypothesis that Htt59025Q involved in interactions must be associated with microtubules or large protein complexes such that it cannot be recovered.

During experimentation it was hypothesized that the Htt was forming large aggregates with many proteins. While this could indeed be the case, it is not strongly supported by this analysis because no aggregate complexes were observed in the poly-acrylamide wells or in solublized samples from the cellular extraction pellet. This hypothesis could be further tested by an in depth optimization of the electromagnetic radiation dosages administered during cross-linking. The aggregate formations would occur as UV exposure proceeded for too long, so that initially Htt is covalently bonded to the miRNA/mRNA that it is interacting with, but then from Brownian motion it floats over and bumps into a nearby protein complex and is then cross-linked with it, and so on depending upon the length of time that the exposure is taking place.

Furthermore, the methodology employed in previous CLIP experiments used special pre-fabricated poly-acrylamide apparatuses that maintained a near neutral pH.

Whereas the Laemmeli style, self-poured SDS PAGE produces a high pH that results in alkaline hydrolysis of the RNA. Within the gel images streaks can be seen in the UV-treated samples. These could have possibly been the remnants of Htt-RNA complexes that were hydrolyzed by the alkalinity. Therefore implementing an electrophoresis apparatus that maintains near neutral pH, such as the Nu-Page system, could produce interpretable results.

Conclusion

While CLIP remains an effective means of researching *in vivo* protein-RNA interaction, much optimization is required before insight can be gleamed for Htt using this methodology. This will allow for illumination of a small portion of the process of post-transcriptional regulation. Because post-transcriptional mRNA processing renders another dimension to genetic regulation and gene expression, elucidating the mechanisms by which translational control is implemented within the transcriptome is highly pertinent to understanding development and maintenance of the organism as well as the layers of complexity involved in orchestrating the molecular symphony of the cell. Furthermore understanding the mechanisms by which Htt transports and perhaps modulates translation of transcripts within the dendrite will help in the delineation of neuronal development and synaptic plasticity because mRNA sequestration is involved in functional control and development of the synapse.

CITATIONS

1. Jean Marx. NEURODEGENERATION: Huntingtin's Research Points to Possible New Therapies. Science 7: Vol. 310. no. 5745, pp. 43 – 45. October 2005

2. Shi-Hua Lia and Xiao-Jiang Li. Huntingtin–protein interactions and the pathogenesis of Huntingtin's disease. Trends in Genetics; Volume 20, Issue 3, Pages 146-154, March 2004.

3. Jeffrey N. Savas, Anthony Makusky, Søren Ottosen, David Baillat, Florian Then, Dimitri Krainc, Ramin Shiekhattar, Sanford P. Markey, and Naoko Tanese. Huntingtin's disease protein contributes to RNA-mediated gene silencing through association with Argonaute and P bodies. PNAS, Vol 105, no. 31, 5 August 2008.

4. Jeffrey N. Savas and Naoko Tanese. A combined immunoprecipitation, mass spectrometric and nucleic acid sequencing approach to determine microRNA-mediated post-transcriptional gene regulatory networks. Briefings in Functional Genomics and Proteomics Advance Access published January 6, 2010

5. Hua Zhang, Sudipto Das, Quan-Zhen Li, Ioannis Dragatsis, Joyce Repa1, Scott Zeitlin, György , Hajnóczky and Ilya Bezprozvanny. Elucidating a normal function of huntingtin by functional and microarray analysis of huntingtin-null mouse embryonic fibroblasts. BMC Neuroscience, 15 April 2008.

6. Emmanuel Beaudoing and Daniel Gautheret. Identification of Alternate Polyadenylation Sites and Analysis of their Tissue Distribution Using EST Data. Genome Res. 11: 1520-1526. 2001.

7. Sung Wook Chi, Julie B. Zang, Aldo Mele & Robert B. Darnell. Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. Nature, Vol 460, 23 July 2009.

8. Targetscan Algorithm. http://www.targetscan.org/.

9. miRanda Algorithm. http://cbio.mskcc.org/research/sander/data/miRNA2003/miranda_new.html.

10. PolyA Cleavage Site & 3' UTR Database. http://harlequin.jax.org/pacdb/