INVESTIGATION OF THE HUNTINGTIN-HAP40 INTERACTION IN HUNTINGTON'S DISEASE

INVESTIGATION OF THE HUNTINGTIN-HAP40 INTERACTION IN HUNTINGTON'S DISEASE

By JENNIFER WILLIAMSON, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE, Biochemistry (2013) Hamilton, Ontario

TITLE: Investigation of the Huntingtin-HAP40 interaction in Huntington's disease

AUTHOR: Jennifer Williamson, B.Sc.

SUPERVISOR: Dr. Ray Truant

NUMBER OF PAGES: xi, 90

ABSTRACT

Huntington's disease (HD) is a neurodegenerative disorder caused by the expansion of a polyglutamine tract in the huntingtin protein. The huntingtin protein has many roles in vesicular and endocytic trafficking, which can be modified in HD cells. When mutant huntingtin is expressed in HD, protein levels of the huntingtin interacting protein, Huntingtin-associated protein of 40kDa (HAP40) are increased. This increase in HAP40 protein levels causes the formation of a complex between carboxyl terminal huntingtin, HAP40 and active Rab5 on the early endosome. This complex induces a switch from early endosomal movements on microtubules to movements on actin filaments, greatly reducing both the speed and distance of movement. The main objective of this research is to determine where the interaction occurs between huntingtin and HAP40 on the huntingtin protein. Here we show that HAP40 interacts with the amino terminus of huntingtin, specifically the first 17 amino acids (N17). In co-localization studies, we found that HAP40 co-localizes with the carboxyl terminal fragment of huntingtin corresponding to amino acids 2570-2634; however, GFP trap immunoprecipitation analysis revealed no interaction between the carboxyl terminal fragments of huntingtin and HAP40. An interaction was discovered between HAP40 and N17, which was then confirmed by far western blot. These results demonstrate that HAP40 interacts with N17. By identifying the interaction site between HAP40 and huntingtin, modifications can be explored to prevent the interaction of HAP40 with huntingtin. This would restore motility on microtubules reinstating fast, long-range movements of early endosomes.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Ray Truant for providing me with the opportunity to work in his lab. I am grateful for the words of encouragement when experiments were not going as planned. I would also like to thank my committee members Dr. Alba Guarné and Dr. Jon Draper for valuable conversations on how to approach problems a little differently.

I would like to thank both former and current members of the Truant lab for their continued support, especially when things were not working. Specifically, I would like to thank Nick Caron for his help at the beginning of my studies in this lab and Tanya Woloshansky for making the transition to working in a new environment fun. I would also like to thank Claudia Hung who has an amazing ability to get experiments to work simply by touching them and Mina Falcone who finally was able to get HAP40 cloned into the pTandem-1 vector.

I would like to express my gratitude to my family and friends for their continued emotional support and encouragement. I would like to thank my parents for the numerous things they have done to help me succeed at what I want to do. Finally I would like to thank my boyfriend, Jamie, I am fortunate to have met someone who cares about me as much as you do. Thank you for being there through good times and bad, I am lucky to have you in my life.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

αMEM	Alpha minimal essential medium
2xLB	2x Luria broth
μg	Microgram
μL	Microliter
AD	Alzheimer's disease
AP2	Adapter protein 2
BIR	Baculovirus inhibitor of apoptosis protein repeat
CHAPs	3-[(3-Cholamidopropyl)dimethylammonio]-1- propanesulfonate
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
F8A	Factor VIII-associated
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EEA1	Early endosome antigen 1
ELM	Eukaryotic linear motif resource
FBS	Fetal bovine serum
GAP	Rab GTPase-activating protein
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDF	GDI displacement factor
GDI	GDP dissociation inhibitor
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
HAP40	Huntingtin-associated protein of 40kDa
HAP1	Huntingtin-associated protein 1
HD	Huntington's disease
HEAT	Huntingtin, Elongation factor 3, subunit of protein phosphatase 2A (PP2A) and TOR1 – P13 kinase target of rapamycin
HEK293	Human embryonic kidney 293
HeLa cells	Human cervical cancer cell line
HPV	Human papillomavirus
HRP	Horseradish peroxidase
HSC70	Heat shock cognate 70
IAP	Inhibitor of apoptosis protein
IBM	Inhibitor of apoptosis protein binding motif
IgG	Immunoglobulin type G
ILV	Intraluminal vesicles
INR	Intranuclear rodlet

Int22h-1	Intron 22 homology region 1
Int22h-2	Intron 22 homology region 2
Int22h-3	Intron 22 homology region 3
IP	Immunoprecipitation
IPTG	Isopropylthio-β-galactoside
IRES	Internal ribosomal entry site
*IT*15	Interesting transcript 15
kDa	Kilodalton
LDH	Lactate dehydrogenase
MCS	Multiple cloning site
mCh	mCherry fluorophore
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
mL	Milliliters
N17	The first 17 amino acids of huntingtin
N17M8P	The first 17 amino acids of huntingtin with a mutation where methionine 8 is substituted for a proline
N17S13PS16P	The first 17 amino acids of huntingtin with serine 13 and 16 phosphorylated
ND2/ND4	Neutral density filters
ng	Nanogram
NIH3T3	Mouse fibroblast cell line
nm	Nanometers
OD600	Optical density at 600nm
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
Phospho-huntingtin	Phosphorylated huntingtin at amino acids 13 and 16
PI	Propidium iodide
PI(3)P	Phosphatidylinositol 3-phosphate
$PI(3,5)P_2$	Phosphatidylinositol 3,5-bisphosphate
PP2A	Protein phosphatase 2A
PVDF	Polyvinylidene fluoride
Rab	Ras-related in brain
REP	Rab escort protein
RNAi	Ribonucleic acid interference
rpm	Rotations per minute
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
siRNA	Small interfering ribonucleic acid
ST*Hdh*$^{Q7/Q7}$	Wildtype immortalized striatal mouse cell line that produces huntingtin from both alleles that contains 7 consecutive glutamine residues in the polyglutamine tract.
ST*Hdh*$^{Q7/Q111}$	Heterozygous immortalized striatal mouse cell line that produces 2 types of huntingtin. One allele produces wildtype huntingtin with 7

	consecutive glutamine residues in the polyglutamine tract. The other allele produces mutant huntingtin with 111 consecutive glutamine residues in the polyglutamine tract.
ST*Hdh*$^{Q111/Q111}$	Homozygous mutant immortalized striatal mouse cell line that produces mutant huntingtin from both alleles that contains 111 consecutive glutamine residues in the polyglutamine tract.
TBS	Tris buffered saline
TBST	Tris buffered saline with tween 20
tGFP	Turbo green fluorescent protein
WT	Wildtype
YFP	Yellow fluorescent protein

TABLE OF FIGURES

DECLARATION OF ACADEMIC ACHIEVEMENT

Jennifer Williamson performed all the experiments and data analysis except for those stated below:

- Nicholas Caron created huntingtin fragments corresponding to amino acids 2570-2700 and 3110-3144.
- Randy Atwal created the huntingtin fragments corresponding to amino acids 1-17 and the construct of amino acids 1-17 with the methionine at position 8 mutated to a proline.
- Claudia Hung performed the LumaScope experiment of the YFP control construct.

Chapter 1 Introduction

1.1 Neurodegenerative Disease

Neurodegenerative diseases are a group of diseases caused by the degeneration of neurons in the body. Depending on the disease, this degeneration can occur in the central nervous system and/or the peripheral nervous system. Neurodegenerative diseases are one of the major causes of disability in Canada, costing the healthcare system billions of dollars each year[1]. The population in Canada is aging, and as individuals age, the chances of diagnosis with a neurodegenerative disease increases[1]. With our aging population it is important that we find effective long-term treatments for these neurodegenerative diseases.

Huntington's disease (HD) is a neurodegenerative disease. It shares features with other neurodegenerative diseases such as Alzheimer's disease (AD) at both the clinical and molecular level; however, unlike other neurodegenerative diseases, HD is caused by a single genetic mutation. The availability of a single genetic marker for the disease allows for identification of patients prior to symptom onset, which permits observation of patients as the disease progresses. Programs such as PREDICT-HD[2] and TRACK-HD[3] were developed to monitor patients that are positive for the genetic mutation beginning when they are healthy until they reach a certain stage of the disease. With programs such as these in place, researchers and clinical professionals can gain insight into how the neurodegenerative disease starts, when it starts, and the progression of disease to the final stages.

1.2 Huntington's Disease

In 1993, a major breakthrough in Huntington's disease (HD) research occurred wherein both the gene and the mutation involved in HD were identified. HD is caused by a mutation in the *IT15* (interesting transcript 15) gene on chromosome 4, which codes for a 350kDa protein called huntingtin[4]. In patients with HD, a polymorphic CAG trinucleotide repeat sequence close to the amino terminus of the protein is expanded, and

this expansion is translated to an abnormally expanded polyglutamine tract in the huntingtin protein[4]. Individuals who produce normal huntingtin have a polyglutamine tract containing 36 or less consecutive glutamine residues. In HD patients, this tract has 37 or more glutamine residues[4, 5]. The mutation is inherited in a dominant fashion where individuals who have one mutant huntingtin allele and one normal allele display the neurodegenerative phenotype[4, 6].

Symptoms of HD are progressive, with both a psychological and a motor aspect. Psychiatric symptoms appear years before the onset of motor symptoms. Patients experience depression, anxiety and irritability and the majority of patients endure disruptions in their sleeping patterns[7]. Motor impairments begin as fragments of voluntary movements that usually occur in the face and limbs[8, 9]. Over time, motor impairments become more severe producing unintentional body and limb twisting movements, abnormal posture, and, occasionally paralysis[8, 10]. Patients are unable to control these movements and are fully conscious when they occur[8]. Symptom onset typically occurs in midlife; however, individuals with a longer polyglutamine tract have an earlier onset of symptoms, though additional factors can influence the age of onset[11]. Following symptom onset and diagnosis, patients live for an additional 15-25 years, eventually succumbing to secondary illnesses[7].

The primary site of HD pathology is in the brain. Many areas of the brain experience degeneration, including the cerebral cortex and other nuclei such as the thalamus. The majority of degeneration involves the medium spiny neurons of the striatum[4, 7, 10, 12, 13]. The striatum is part of the basal ganglia, an important group of nuclei that receive inputs from all areas of the cerebral cortex and sends information back to the cortex via thalamic projections[9, 12, 14]. The motor and cognitive information gathered from the cerebral cortex is used by the basal ganglia to create and initiate a motor program to produce a desired movement or series of movements[12]. The striatum has two important nuclei: the putamen and the caudate nucleus[9]. The putamen is involved with motor movements. Inputs to this nucleus include those from the motor and somatosensory areas of the cerebral cortex and the projections from the putamen are relayed by the globus

pallidus, another region of the basal ganglia, and the thalamus to the motor, premotor and supplementary motor cortical areas[9]. The caudate nucleus is less involved with motor movements and more involved with the cognitive functions associated with the movements. Input to this nucleus comes from the association areas of the cortex and the projections from this nucleus are sent through the globus pallidus and thalamus to the prefrontal cortex[9]. Since 90-95% of all striatal neurons are medium spiny neurons[10], degeneration of this neuronal population, as seen in HD, can lead to severe disruption of this motor and cognitive circuitry.

Currently there are no long-term treatments available for HD; however, there are treatments to reduce the severity of symptoms[15, 16]. Unfortunately, at time of symptom onset, patients have already lost a significant amount of neurons in the striatum and treatment at this point may be too late. Investigation of the normal roles that the huntingtin protein has in the cell could lead to answers about how this disease can be treated long-term and when the treatment should begin.

1.3 The Huntingtin Protein

1.3.1 Structure and Domains

Huntingtin is a large protein that is expressed at various levels in all tissues of the body, with higher levels in the brain and testis[4, 5]. In the cell, huntingtin can be found in both the nucleus and the cytoplasm[17-23]. The mutation that causes HD is at the amino terminus of the protein, and most studies are targeted to that area. Three important domains are located at the amino terminus: the first 17 amino acids (N17), the polyglutamine tract, and a proline rich region. The N17 region of huntingtin is a highly conserved sequence that forms an amphipathic alpha helix[21, 24, 25]. This sequence can be phosphorylated at serine residues 13 and 16 causing a change in huntingtin localization from the cytoplasm, specifically from the endoplasmic reticulum, to the nucleus[21, 22]. At the carboxyl terminus of N17 is the polyglutamine tract, the location of the expansion in HD. In crystal structures, this area has conformational variability. At the beginning of the tract, the alpha helicity of the N17 region is maintained; however, the structure gradually

degenerates to a random coil[24, 25]. Immediately downstream of the polyglutamine tract, there is a proline rich region composed of a sequence of 11 consecutive proline residues followed by a 17 residue proline and glutamine linker. This linker is followed by another sequence of 10 consecutive proline residues[4] forming a polyproline helix that is involved in protein-protein interactions[24, 25 26].

Due to the size of the protein, huntingtin requires nuclear import and export signals to allow facilitated transport into and out of the nucleus. A nuclear localization signal has been identified between amino acids 174 and 207 allowing nuclear entry using the karyopherin β1/β2 pathway[27]. Two nuclear export signals have been discovered: one within N17 and the other close to the carboxyl terminus at amino acids 2364 to 2414 of human huntingtin[28, 29]. Both of these export signals are recognized by the transporter exportin/CRM1 enabling the proteins to exit from the nucleus[28-30].

The majority of huntingtin is composed of helical structures called HEAT repeats[31]. HEAT repeats were named after the first 4 proteins that they were discovered in: Huntingtin, Elongation factor 3, subunit of protein phosphatase 2A (PP2A) and TOR1 – P13 kinase target of rapamycin[32, 33]. HEAT repeat proteins are found in many cellular compartments and function as a scaffold to mediate protein-protein interactions by positioning proteins in contact with other proteins or cellular molecules[31, 32]. Known HEAT repeat proteins have roles in transcriptional control, nuclear/cytoplasmic transport, chromosomal segregation, DNA repair, mitosis, microtubule dynamics and many other processes[33-35]. HEAT repeat motifs are repeated sequences of approximately 40 to 50 amino acids, which are assembled into 2 α-helices, an A helix and a B helix. A 2-3 amino acid linker allows the helices to fold together creating a hairpin-like structure. These repeats are stacked together to form a superhelix with a continuous hydrophobic core which allows the protein to change conformation without rupturing when forces below a certain threshold are exerted on the protein[31-33, 35].

Molecular modeling of yeast importin β and the PR65 scaffold of PP2A have demonstrated that pushing or pulling forces that are exerted at one end of the protein change the arrangement of intermolecular forces between the residues in the hydrophobic

core. This evenly distributes the force along the length of the protein so the protein does not rupture[32, 33]. Torsional effects were also studied, and the forces exerted by torsion were also evenly distributed along the length of the protein[32]. These findings show that mechanical forces exerted on a HEAT repeat protein, like huntingtin, can have an effect on the conformation of the protein, transiently revealing or concealing interaction sites or other signaling motifs. It also demonstrates that transduction can occur where an interaction or force at one end of the protein can alter events that may occur at the other end of the protein suggesting that the polyglutamine expansion that occurs in HD could effect interactions or processes at the carboxyl terminus.

1.3.2 Function

Huntingtin is an essential protein for central nervous system development[36-38]. The mouse huntingtin gene homologue has been mutated or deleted to explore the effect of huntingtin in both cellular and animal systems. When studying cells individually, huntingtin is not required for cell viability; however, the huntingtin protein is required for proper embryonic development[36-38]. Homozygous mice with both huntingtin alleles inactivated, die at embryonic day 7.5 and are resorbed[36]. Additionally, homozygous mice with a mutation that expanded the polyglutamine tract length and reduced the expression level of huntingtin to a third of wildtype levels developed an abnormal central nervous system. These mice were either stillborn or died shortly after birth[38]. Homozygous mice with an expanded glutamine tract and wildtype levels of huntingtin developed normally[38]. This demonstrates that normal levels of huntingtin are required from at least one allele for correct central nervous system development, and this is independent of the polyglutamine tract length.

Huntingtin has a vast range of cellular functions. In the cytoplasm, huntingtin has an important role at the endoplasmic reticulum, where it is required for proper maintenance and cell stress response[21, 39]. It is also required at the mitochondria for proper mitochondrial localization in the perinuclear area[39]. Huntingtin-associated protein 1 (HAP1) interacts with huntingtin to transport brain derived neurotrophic factor, a protein that is required for normal neuronal activity, to required locations in the neuron[40]. In the

nucleus, huntingtin regulates transcription of many genes and changes in gene expression are seen when huntingtin is mutated in HD[41-43]. Huntingtin interacts with many interacting and associated proteins to carry out cellular functions[44]. One major cellular pathway that huntingtin and huntingtin interacting proteins are involved in is the endocytic pathway.

1.4 The Endocytic Pathway

The endocytic pathway is important for cellular regulation, communication, growth and overall cell survival[45]. It is the process used by cells to sort, recycle, degrade and deliver internalized cargo to the appropriate compartment[46]. Endocytosis has a range of biological functions including the regulation of plasma membrane components, the internalization of nutrients and the control of intracellular signaling by modulating receptors on the cell surface[45, 47-49]. The ability to modulate receptor numbers can lead to long-term signal reduction or even signal extinction and the constant turnover of plasma membrane components allows the cell to adapt to the extracellular environment[45, 46].

Endocytosis is a process that is regulated by small Ras-related in brain (Rab) proteins. Rab proteins are small GTPase proteins that are ubiquitously expressed in all tissues[50]. Rab proteins use guanine nucleotide exchange as a switch to regulate almost all of the membrane trafficking that occurs in cells[50, 51]. Figure 1 demonstrates the Rab cycle in the cytoplasm of cells. When Rab proteins are bound to GTP, they recruit effector molecules. Effector molecules are molecules, proteins, or a complex of proteins that interact with a GTP-bound Rab protein and carry out specific tasks that the Rab protein is responsible for regulating[50, 52]. Through the use of these effector molecules, Rab proteins regulate budding of the vesicle into the cell, delivery of the vesicle to specific cytoskeletal elements, attachment of the vesicle to the target membrane prior to fusion and, finally, fusion of the vesicle with the target membrane[50, 53]. Each endocytic compartment has a specific set of Rab proteins[54]. The different Rab proteins and their conversion between endocytic compartments provide a direction for progression through the endocytic pathway[50, 51]. Each step of endocytosis is regulated by a Rab protein and associated effector molecules.

1.4.1 Endocytosis

The endocytic pathway begins with the internalization of a vesicle by one of two pathways: clathrin-dependent endocytosis or clathrin-independent endocytosis. Clathrin-dependent endocytosis is a well-characterized process that occurs in all known eukaryotic cells[49, 55]. This pathway uses clathrin to encapsulate and stabilize a budding vesicle[56]. A clathrin triskelion is composed of 3 clathrin heavy chains and 3 clathrin light chains that interact at the carboxyl terminus[57, 58]. Specialized adaptor proteins are required to recruit clathrin to the site of vesicle formation[49] and to select cargo for internalization[59]. The plasma membrane specific guanine nucleotide exchange factor (GEF) hRME-6 recruits Rab5 to the plasma membrane[60]. Without Rab5, the clathrin-coated vesicle will not form[52, 56]. The clathrin triskelia polymerize into a lattice of pentagons and hexagons to form a clathrin-coated vesicle attached to the plasma membrane[49, 61]. Dynamin, a mechanochemical enzyme, is recruited to the vesicle neck where it undergoes a conformational change, gaining energy from GTP hydrolysis to constrict the neck of the vesicle, detaching it from the plasma membrane[62-65]. Once detached, the clathrin-coat and the adaptor proteins are shed allowing the vesicle to fuse with the early endosome.

Clathrin-independent mechanisms of endocytosis are classified as such because a clathrin coat is not required for internalization of the vesicle. Examples include: RhoA-regulated endocytosis, caveolae-mediated endocytosis, ARF6-regulated endocytosis and CDC42-regulated endocytosis[66]. These pathways are difficult to study because the processes change depending on the cell type, the environment that they are in, and the treatment that they receive. Many of the proteins involved participate in multiple pathways, at different times and under various conditions[67]. The cargo itself can also affect how the endosome is internalized and what pathway is followed. Additionally, very little is known about how the cargo is selected for internalization by these pathways, unlike clathrin-dependent endocytosis where cargo specific adaptors have been discovered[66].

1.4.2 Early endosome

Almost all vesicles, regardless of how they are internalized, fuse with the early endosome[48, 66, 68]. The major function of the early endosome is to sort the incoming cargo via sequence motifs and receptor interactions[45, 48, 68]. The GTPase Rab5 is a key component of the early endosomal machinery. When Rab5 levels are reduced below 15% of normal using small interfering RNA (siRNA), the entire endosomal and lysosomal system is shut down[69]. Rab5 recruits many effector molecules that regulate early endosomal functions such as homotypic fusion, endosomal maturation, transportation along cytoskeletal elements and the conversion from early to late endosome[68, 70, 71]. Rab5 mediates the transport of endocytic vesicles and early endosomes along two cytoskeletal networks: the microtubule network and the actin filament network[71, 72]. The microtubule network supports fast, long distance transport using both kinesin family and dynein motor proteins[68, 71, 72]. The actin filament network supports slower, more local, short-range movements using motor proteins of the myosin family[72]. GTP-bound Rab5 regulates recruitment of effector molecules to promote the interaction of the microtubule motor protein with the endosome[71].

The early endosome is a very dynamic compartment due to the high rate of fusion that occurs between early endosomes to deposit the cargo for sorting[48]. The pH of the early endosome lumen is mildly acidic, which favors the dissociation of ligands from their receptors[48]. The receptors are sent back to the plasma membrane through the recycling pathway, and the ligand is sorted for degradation[48]. The membrane of the early endosome contains many domains, each with different functions and contents[68]. Domains that occur can be enriched in Rab5, Rab4 or Rab11. The domains are present to provide an area with a high concentration of Rab protein for Rab-associated tasks and to aid with sorting so that the membrane is ready to produce a vesicle and send it to the appropriate compartment[68]. The contents of the early endosomal compartment can be sorted to one of three compartments: the recycling compartment, the late endosomal/lysosomal compartment or the Golgi network[48]. The majority of the cargo is recycled back to the plasma membrane[68].

1.4.3 Recycling endosome

The majority of cargo taken into the cell is returned to the plasma membrane through the recycling pathway. Transport in this pathway requires the actin cytoskeleton and myosin motors recruited by the effector molecules of the associated Rab protein[48]. There are two recycling pathways, each regulated by a different Rab protein[46]. The fast route of recycling is mediated by Rab4 and is used by the cell to sustain signaling by rapidly returning receptors to the plasma membrane[46]. The cargo recycled by this route contains a sequence that identifies it as cargo to be recycled[73]. Many receptors are recycled via this route. When receptors are sorted into recycling endosomes, they are re-sensitized by Rab4 effector molecules, which allow the receptors to bind their ligand when they return to the plasma membrane[45].

The other recycling pathway is the slow recycling pathway, mediated by Rab11. The cargo in this pathway is first sent to the pericentriolar recycling endosome before returning to the plasma membrane[46, 74, 75]. Rab11 acts as a regulator of the processes necessary for sorting and exiting the pericentriolar recycling endosome, including vesicle formation from the recycling endosome.

1.4.4 Late endosome

The final opportunity to be sorted out of the degradative pathway occurs at the late endosome, and only a small proportion of cargo reach this point[68]. Late endosomes contain a mixture of cargo to be degraded and cargo that the late endosome and lysosome require to function such as new membrane proteins and hydrolases[68, 76]. Many changes occur at the late endosome prior to fusion with the lysosome[68]. There is a switch in Rab proteins from Rab5 to Rab7[70]. Intraluminal vesicles (ILVs) are formed by the inward budding of the endosomal membrane to help sort cargo for degradation[45, 68]. Cargo that will be degraded is incorporated into ILVs so it is accessible to the hydrolases in the lumen of the lysosome. ILVs also restrict receptor access to the cytoplasm, which deactivates them[68]. Proton pumps are added to the membrane of the endosome to lower the pH to between 6.0 - 4.9[68, 77]. The point at which a late endosome becomes a lysosome is difficult to determine because both compartments are so similar in composition and pH.

1.4.5 Huntingtin in the Endocytic Pathway

Endocytosis is a very important process in the cell for regulating many signaling pathways and overall homeostasis of the cell. It is tightly regulated by Rab GTPases and their associated effector molecules and without these components, endocytosis would not occur. Huntingtin is involved in many steps of the endocytic pathway both in normal conditions and in disease conditions.

Huntingtin and dynein are both responsible for the positioning of endosomes and lysosomes along cytoskeletal elements in the cell[78]. Normally, late endosomes and lysosomes are located in the perinuclear area, but when levels of huntingtin are decreased, they are redistributed throughout the cytoplasm[78]. Reduction of dynein levels using RNA interference (RNAi) displays a more extreme phenotype where late endosomes and lysosomes are redistributed to the cell periphery. In huntingtin and dynein depleted cells, the motility of the late endosomes and lysosomes were decreased but their function remained normal[78]. Early and recycling endosome positioning is also redistributed when huntingtin and dynein heavy chain or the $p150^{Glued}$ subunit of dynactin are decreased using RNAi[78]. Full-length huntingtin transfected into huntingtin-deficient cells was able to fully redistribute all types of endosomes and lysosomes to their correct location, implicating huntingtin in endosomal positioning along microtubules[78]. Currently, the effect of mutant huntingtin on this function is unknown.

In HD, there is an upregulation of huntingtin-associated protein of 40kDa (HAP40), which alters the movements of early endosomes. The upregulation of HAP40 induces the formation of a complex wherein HAP40 links huntingtin to GTP-bound Rab5 on the early endosome[79]. This interaction causes a switch in early endosomal transport in which movements that normally occur on microtubules switch to movements on actin filaments. This switch reduces long-range movements of the early endosome on microtubules[79]. This will be discussed further in section 1.5.

Huntingtin is required for the activation and membrane localization of Rab11 at the recycling endosome[80]. When huntingtin is not expressed in the cell, the amount of Rab11 bound to membranes, and the activation of Rab11 by nucleotide exchange from GDP to

GTP, is reduced[80]. Huntingtin was found to interact preferentially with the GDP-bound Rab11 at the recycling endosome membrane suggesting that huntingtin is required to stabilize and activate Rab11 at the recycling endosome[80]. Analysis of fibroblasts from control and HD patients revealed impairments in receptor recycling, which would occur if Rab11 was not activated and could not recruit the appropriate effector molecules[81]. Mutant huntingtin and Rab11 co-localize to abnormal tubular structures at the pericentriolar recycling endosome. These abnormal structures were identified as recycling transport intermediates. Membrane association of mutant huntingtin and Rab11 to recycling endosome structures was greater in the HD fibroblast cells than controls[81]. Active Rab11 is required for vesicle formation from the recycling endosome, and this process is impaired in HD fibroblasts and can be reversed by expressing a dominant active Rab11[81]. The presence of transport intermediates, the increased localization of mutant huntingtin and Rab11 on recycling endosome membranes and the impaired vesicle formation indicate a defect in recycling from the pericentriolar recycling endosome to the plasma membrane. This defect is caused by the inhibitory effect that mutant huntingtin has on Rab11 nucleotide exchange[81].

At the late endosome, huntingtin co-localizes with Rab7 and with autophagic vesicles[21, 82]. Wildtype huntingtin appears to function in the late endosome/lysosome compartment and is responsible for the association of the vesicle with microtubules and associated motors[78]. Very little information is available about the effect that mutant huntingtin has on this portion of the endocytic pathway.

Mutant huntingtin causes disruptions in the endocytic pathway through abnormal interactions with both Rab11 and Rab5. Although these disruptions in the pathway are subtle and do not cause immediate cell death, they can still contribute to the pathogenesis of HD. The slow effect of altered endocytic events and vesicular transport parallels the long-term progression of the symptoms and pathology of HD, suggesting that there could be endocytic involvement in the progression of disease.

1.5 Huntingtin-associated Protein of 40kDa

The huntingtin-associated protein of 40kDa (HAP40) gene is associated with the Coagulation Factor VIII locus on the X chromosome. Due to this association, another name used to identify HAP40 is Factor VIII-associated gene A (F8A). The full HAP40 gene is located in intron 22 of the Coagulation Factor VIII gene[83-85] in a region identified as the intron 22 homologous region 1 (*int22h*-1). *Int22h*-1 is a repetitive sequence that contains a CpG island where HAP40 transcription is initiated at a bidirectional promoter in the opposite direction of Factor VIII transcription[86-88]. In humans, there are at least 2 additional copies of *int22h*-1 found outside of the Factor VIII gene closer to the Xq telomere called intron 22 homologous region 2 (*int22h*-2) and region 3 (*int22h*-3)[84, 85, 87, 88]. *Int22h*-2 is oriented in the same direction as *int22h*-1, and *int22h*-3 is oriented in the opposite direction of *int22h*-1[89]. The orientation of these regions is demonstrated in figure 2. These 3 homologous regions are 99.9% similar and all contain a copy of the HAP40 gene[87]. Studies have suggested that all 3 regions contribute to the total amount of HAP40 produced, though the contribution of each of these regions to the total HAP40 population is still unknown[84]. Mouse HAP40 is 85% similar to human HAP40 and is expressed in a variety of different cell types; however, there is only one copy of the gene in the mouse genome and it is encoded outside of the Factor VIII locus[84, 85]. Since HAP40 is expressed in many different cell types[83-85] and is conserved in mice[85], HAP40 was hypothesized to have a housekeeping role in the cell, although no specific roles were identified[84].

In 2001, HAP40 was discovered to interact with huntingtin. Using affinity chromatography with antibodies that recognize multiple epitopes on huntingtin, Peters and Ross (2001) were able to isolate huntingtin complexes from rat brain homogenate (soluble fraction)[86]. One of the proteins isolated with huntingtin in a complex was a 40kDa protein that was named HAP40[86]. Coimmunoprecipitation using HAP40 antibodies and transfected human embryonic kidney (HEK) 293 cell lysates revealed that HAP40 interacts with full length huntingtin, but not with a construct corresponding to the first 513 amino acids of huntingtin[86]. This was the first time that a huntingtin interacting protein was shown to associate with a central or carboxyl terminal portion of huntingtin.

The expression level of HAP40 was found to be low in brain samples through analysis of immunolabeled wildtype mouse brain slices[86]. The same antibodies used to probe brain slices were able to detect endogenous levels of HAP40 in cell lines where localization of HAP40 varied. HAP40 contains a nuclear localization signal close to the amino terminus of the protein, which enables entry to the nucleus. Nuclear entry is reported to occur in the absence of detectable huntingtin[86]. In the presence of huntingtin, HAP40 is found in the cytoplasm diffuse and at puncta[79, 86]. Recently, in hypothalamic neurons, HAP40 was shown to form an intranuclear rodlet (INR), which also contains ubiquitin[90]. Monoubiquitination is a post-translational modification that allows endocytic proteins to bind to ubiquitinated endocytosed proteins like membrane receptors. It can also regulate the nuclear or cytoplasmic localization of a protein[90], which could explain the ability of HAP40 to be targeted to the nucleus or to the cytoplasm. The function of HAP40 in the nucleus is currently unknown; however, investigation of the role of HAP40 in the cytoplasm has been investigated.

HAP40 is a Rab5 effector molecule that is involved with vesicular trafficking of early endosomes[79]. Using affinity chromatography, Pal *et al* (2006) identified HAP40 and huntingtin in a complex with active GTP-bound Rab5 but not inactive GDP-bound Rab5. Further analysis revealed that HAP40 is required for huntingtin to interact with GTP-bound Rab5 and that the carboxyl terminal fragment of huntingtin, corresponding to amino acids 2226-3144 is the area of huntingtin involved in this complex[79]. This suggests that the carboxyl terminus of huntingtin interacts with active Rab5 using HAP40 as a connector.

In HD cell lines, patient brain samples and patient fibroblasts HAP40 protein levels are increased[79]. When HAP40 protein levels are elevated, HAP40 interacts with huntingtin and active Rab5 at the early endosome, but when endogenous levels of HAP40 are expressed, this interaction does not occur. HAP40 overexpression was specifically required for this interaction, as over-expression of Rab5 did not change the amount of huntingtin or HAP40 localized to the early endosome[79]. Reduction of HAP40 levels using siRNA prevented the localization of huntingtin to early endosomes demonstrating that

HAP40 overexpression was required for the localization of huntingtin to the early endosome[79].

Both huntingtin and Rab5 regulate positioning of endosomes on the cytoskeleton; therefore, the effect of this complex on motility along microtubules and actin filaments was examined[79]. Pal *et al* (2006) discovered that the addition of glutathione S-transferase (GST)-HAP40 inhibited endosomal movement along microtubules and reduced the amount of early endosomes that were associated with microtubules[79]. The effect that this complex had on early endosomal motility along actin filaments was tested and it was discovered that the addition of HAP40 stimulated binding of early endosomes to actin filaments[79]. Additionally, the concentration of HAP40 required to inhibit endosomal binding to microtubules was enough to stimulate endosomal binding to actin[79]. This suggests that the huntingtin-HAP40-Rab5 complex stimulates early endosomal interaction with actin filaments and inhibits interaction with microtubules[79]. This was confirmed in fibroblasts from HD patients where the levels of HAP40 are increased[79]. Immunofluorescence was used to examine the association between early endosomes and the actin and microtubule networks in HD and control fibroblasts. In control fibroblasts, early endosomes were aligned along microtubules. In the HD fibroblasts, the alignment of early endosomes occurred on the actin cytoskeleton and only a small proportion of endosomes were aligned along microtubules. Proper endosomal positioning and motility was restored in these cells by reducing HAP40 levels using RNAi, suggesting that the abnormal positioning of endosomes was a result of elevated HAP40 protein levels.

The results presented by Pal *et al* (2006) demonstrate that the complex formed between HAP40 and huntingtin acts as a Rab5 effector that can regulate early endosomal movement by switching to movements on the actin cytoskeletal network. Endosomal movements on actin filaments are short and slow, compared to movements on microtubules, which are fast and occur over long distances. The interaction of HAP40 with huntingtin occurs at the carboxyl terminus and is therefore not likely to be directly affected by the polyglutamine expansion, however it is affected indirectly. Transcriptional abnormalities occur in HD, and Pal *et al* (2006) demonstrated that HAP40 is upregulated

in HD patient brain tissue, HD patient fibroblasts, and HD cell lines[79]. This upregulation of HAP40 influences the interaction of huntingtin with Rab5, altering early endocytic transport

1.6 Project Rationale

Endocytosis is an important process in the cell and impairment of this pathway can affect many cellular processes. HD patient cells and HD cell lines are known to have a early endosome motility defect due to the overexpression of HAP40. The complex that forms between huntingtin, HAP40 and Rab5 at the early endosome shifts endosomal movements to microtubules, limiting both the range and speed of movement. When HAP40 levels are decreased, motility is restored. If we can find where the HAP40 interaction occurs on huntingtin then methods to prevent this interaction can be explored and early endosomal motility can be restored. Additionally, if an increase of HAP40 levels triggers a standstill of the endocytic pathway through the formation of this complex than the increased levels of HAP40 may also be toxic to the cells and disrupting the interaction could prevent cell death.

Chapter 2 Materials and Methods

2.1 Fusion Constructs

Huntingtin Fusion Proteins

N17-YFP

N17M8P-YFP

Huntingtin2570-3144-YFP (H2570-3144-YFP)

Huntingtin2570-2700-YFP (H2570-2700-YFP)

Huntingtin2570-2634-YFP (H2570-2634-YFP)

Huntingtin2635-2700-YFP (H2635-2700-YFP)

Huntingtin3110-3144-YFP (H3110-3144-YFP)

Huntingtin2570-2700-mCh (H2570-2700-mCh)

Huntingtin2570-2634-mCh (H2570-2634-mCh)

Huntingtin2635-2700-mCh (H2635-2700-mCh)

Huntingtin3110-3144-mCh (H3110-3144-mCh)

HAP40 Fusion Proteins

YFP-6xGlycine-HAP40 (YFP-HAP40)

HAP40-6xGlycine-YFP (HAP40-YFP)

F8A-tGFP (Origene Technologies Inc.)

> Note: This is the HAP40 sequence but will be referred to as F8A to distinguish between constructs.

GST-HAP40

Venus-6xGlycine -HAP40-6xglycine-mCherry (venus-HAP40-mCherry)

Vectors and Controls

All YFP fusion constructs were cloned into the YFP-N1 vector (Clontech).

All mCherry fusion constructs were cloned into the mCherry vector (Clontech).

pGEX-5X-1 (GE Life Sciences)

2.2 Antibodies

2.2.1 Primary Antibodies

Rabbit anti-HAP40 (Abcam)

Western blot 1:1500-1:500

Immunofluorescence 1:100-1:50

Rabbit anti-GST (Sigma)

Western blot 1:2000

Mouse anti-GAPDH (Abcam)

Western blot 1:10000

Rabbit anti-N17 huntingtin – epitope: amino acids 9-17 of N17

Western blot 1:2500

Rabbit anti-N17S13PS16P huntingtin – epitope: amino acids 9-17 of N17 with serines 13 and 16 phosphorylated

Western blot 1:2500

Rabbit anti-N17pan huntingtin – epitope: the first 8 amino acids of N17

Western blot 1:2500

Mouse anti-YFP (Clontech)

Western blot 1:2500

Mouse anti-tGFP (Origene Technologies Inc.)

Western blot 1:2000

Sheep anti-cytochrome c (kind gift from Dr. David Andrews lab)

Immunofluorescence 1:2000

2.2.2 Secondary Antibodies

Goat anti-rabbit HRP (Abcam) 1:50000

Rabbit anti-mouse HRP (Abcam) 1:10000

Goat anti-mouse HRP (Biorad) 1:10000

Alexa488 donkey anti-rabbit (Molecular Probes) 1:500

Alexa594 chicken anti-rabbit (Molecular Probes) 1:250

Alexa594 donkey anti-sheep (Molecular Probes) 1:500

2.3 Synthetic Peptides

Peptide corresponding to the first 17 amino acids of huntingtin (N17WT):

MATLEKLMKAFESLKSFC (GenScript)

Peptide corresponding to the first 17 amino acids of huntingtin where the methionine in position 8 is mutated to a proline (N17M8P):

MATLEKLPKAFESLKSFC (GenScript)

Peptide corresponding to the first 17 amino acids of huntingtin where serines at positions 13 and 16 are phosphorylated (N17(pS13)(pS16)):

MATLEKLMKAFE(pS)LK(pS)FC (New England Peptide)

Peptide corresponding to the HAP40 antibody epitope at amino acids 314-325:

DGHGQDTSGQLP (GenScript)

2.4 General Solutions

1% CHAPS lysis buffer: 30mM Tris-HCl pH 7.5, 150mM NaCl, 1% CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate)

Phosphate buffered Saline (PBS): 10mMNaH_2PO_4, 0.14M NaCl, 1mM EDTA, pH7.4

Tris buffered saline (TBS): 10mM Tris, 150mM NaCl, pH7.4

TBST: TBS with 0.05% Tween 20 (Sigma)

2.5 ST*Hdh* Cell Lines

ST*Hdh* cell lines (a kind gift from M. E. MacDonald, MGH) are embryonic, day 14 mouse striatal cells. There are 3 different cell lines. ST$Hdh^{Q7/Q7}$ cells are wildtype where the polyglutamine tract in the huntingtin protein contains 7 consecutive glutamine repeats. ST$Hdh^{Q7/Q111}$ cells are heterozygous where one huntingtin allele contains a knock in mutation to extend the polyglutamine tract to 111 consecutive glutamine repeats in the huntingtin protein. The other huntingtin allele contains 7 consecutive glutamine repeats in the polyglutamine tract. Finally, ST$Hdh^{Q111/Q111}$ cells contain a knock in mutation to both huntingtin alleles that extends the polyglutamine tract length to 111 consecutive glutamine repeats in the huntingtin protein.

2.6 Cell Culture, Transfection and Treatment

All cells were grown in 5% carbon dioxide and 95% air containing atmospheric oxygen levels. ST*Hdh* cell lines were grown at 33°C in Dulbecco's modified eagle's medium (DMEM) (Invitrogen) with 10% fetal bovine serum (FBS) (Sigma) under G418 drug selection at a concentration of 300μg/mL. HeLa cells and NIH3T3 cells were grown in DMEM with 10% FBS at 37°C. HEK293 cells were grown in Alpha minimum essential medium (αMEM) with 10% FBS at 37°C. Human HD (Coriell Institute #GM21756) and control (Coriell Institute #GM2149) fibroblasts were grown at 33°C in minimum essential medium (Sigma) with 15% FBS (Sigma) and 2mM L-Glutamine (BioShop).

Cells were seeded 24 hours prior to transfection. They were seeded into 35mm glass bottom dishes for imaging or into 10cm dishes for western blot. Transfections of 35mm dishes required 3μg of DNA and transfections of 10cm dishes used 12μg DNA. Cells were transfected with either Turbofect (Fermentas), Lipofectamine2000 (Invitrogen) or using the 4D nucleofector system (Lonza). Cells that were transfected using Turbofect or Lipofectamine2000 were transfected according to the manufacturer's directions and without antibiotic. The cells that were electroporated using the Lonza 4D nucleofection system were transfected according to Lonza's NIH3T3 protocol using the "NeuraHE" program. Ten picomoles of siRNA (HAP40 siRNA (Santa Cruz Biotechnology) or scrambled medium G-C content siRNA (Sigma)) was transfected into 35mm dishes using Lipofectamine2000 (Invitrogen) according to the manufacturer's directions.

Cells treated with staurosporine were seeded 24 hour prior to treatment. A 0.5μM solution of staurosporine in serum free media was applied to cells following a brief PBS wash. The cells were returned to the incubator for 4 hours. Following incubation the cells were removed from the incubator and washed once with PBS. The cells were then fixed and immunofluorescence, as described in section 2.13 was completed.

2.7 Recombinant Protein Preparation

Buffers

Elution Solution: 50mM TRIS-HCl pH 8.0 and 10mM reduced glutathione

SDS Sample buffer: 3x Blue Loading Buffer and 30x Dithiothreitol (DTT) (sold by New England Biolabs as a Blue loading buffer pack) diluted to 1x in the sample

Procedure

HAP40 was cloned into the pGEX-5X-1 expression vector (GE Healthcare) and sequenced to ensure that it was fully inserted. Both the pGEX-5X-1 empty vector and the vector containing GST-HAP40 were transformed into *Escherichia coli* BL-21 cells and incubated for 16-18 hours in a 37°C incubator. The cultures were inoculated into 2x luria broth (2x LB) with ampicillin to select for cells that had the plasmid. This culture was incubated at 37°C overnight in a shaker incubator. The following day, a 1:10 dilution of one part overnight culture to ten parts fresh 2x LB containing ampicillin was made. This diluted culture was returned to the 37°C shaker incubator and monitored to obtain an optical density at 600nm (OD600) of 0.6-0.8. When the OD600 is met, 0.5mM isopropyl-beta-D-thiogalactopyranosid (IPTG) is added to the appropriate cultures. IPTG was not added to all of the cultures, if IPTG was not added, the cultures remained in the incubator. The induced cultures were returned to the 37°C shaker incubator for an additional 4-6 hours. The cultures were centrifuged at 4000rpm for 15 minutes at 4°C. The supernatant was removed and the pellet was washed with cold 150mM NaCl and was spun again with the same settings. The supernatant was removed and the pellet was stored at -20°C overnight.

The pellet was suspended in cold PBS. The amount of PBS used was 10% of the amount of culture used during the induction with IPTG. The cells were pulse sonicated on ice for 5 minutes (output 4, duty cycle 30-40%). The sonicated solution was centrifuged at 10000rpm for 15 minutes at 4°C and the supernatant was transferred to a 50mL falcon tube. Glutathione Sepharose 4B beads (GE Healthcare) were washed three times with cold PBS and 100µL of beads (or 200µL of a 50% beads and 50% PBS mixture) for every 10mL of supernatant was added. Beads were incubated with the supernatant by end over

end rotation at 4°C for 1 hour. The samples were then spun down at 500rpm for 5 minutes at 4°C. The supernatant was removed and the beads were washed three times with cold PBS. The bound proteins were eluted; the volume of elution solution used was four times amount of beads used (ex: if 100µL of beads or 200µL of the 50% beads and 50% PBS solution was used, 400µL of elution solution was required). The elution solution was incubated with the beads by end over end rotation for 10 minutes at 4°C. The solution was then spun down at 500rpm for 5 minutes at 4°C and the supernatant was transferred to a new tube and stored at -20°C.

To determine the purity or concentration of the sample, the samples were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and the proteins in the gel were visualized with Coomassie blue stain (Life Technologies Inc.). If the sample was pure, a Bradford assay (Biorad) was used to determine the concentration of protein.

2.8 Western Blot

Buffers

SDS Running Buffer: 25mM Tris, 0.192M Glycine, 0.1% SDS (w/v)

SDS Sample buffer: 3x Blue Loading Buffer and 30x DTT (sold by New England Biolabs as a Blue loading buffer pack) diluted to 1x in the sample

TBST: Tris buffered saline containing 0.05% Tween 20

Sample Preparation

Samples for western blot were grown on 10cm dishes and harvested when the dish was approximately 80-90% confluent. Cells were rinsed once with 1xPBS and then scraped off the dish and centrifuged at 3000rpm for 5 minutes. The supernatant was removed and 100µl of CHAPs lysis buffer with 1x protease inhibitor cocktail (Roche) was added to the cell pellet. The cells and lysis buffer were incubated on ice for 15-20 minutes with gentle mixing every 5 minutes. The cells were then centrifuged at 13000rpm for 15 minutes at 4°C. A Bradford assay was used to determine the concentration of protein in the sample. A total of 30µg of protein was loaded into each well with SDS

sample buffer. The samples were boiled at 100°C for 10 minutes and stored at -20°C until analyzed by SDS-PAGE.

Procedure

A gel was created according to the proteins that were to be resolved. Most samples were run on a 12% gel containing 29 parts acrylamide to 1 part bisacrylamide. To resolve large proteins, samples were run on a 7% gel containing 80 parts acrylamide to 1 part bisacrylamide. To resolve small proteins and peptides, samples were run on a 12% gel containing 16 parts acrylamide to 1 part bisacrylamide. Equal amounts of sample were loaded into each well and run at 120V for 2-5 hours. Large and medium protein containing gels were transferred to a poly-vinyl difluoride (PVDF) membrane (Pall Corporation) with a pore size of 0.45µm and small protein containing gels were transferred to a nitrocellulose membrane (Pall Corporation) with a 0.2µm pore size. Membranes were blocked for 1-3 hours at room temperature in a 5% (w/v) solution of powdered skim milk in TBST. The blot was then incubated overnight at 4°C with the required primary antibody (see section 2.2.1). The next day, the blot was washed 3 times for 15 to 30 minutes with 2.5% (w/v) powdered skim milk in TBST. The appropriate HRP conjugated secondary antibody (see section 2.2.2) was applied at room temperature for 45 minutes. This was followed by 4 additional washes: three 15-30 minute washes in 2.5% milk and one 15 minute wash in TBST. Blots were visualized using diluted Immobilon western chemiluminescent HRP substrate (Millipore) or Amersham enhanced chemiluminescence (ECL) western blotting detection reagent (GE Healthcare) and pictured using the MicroChemi System (DNR Bio-Imaging Systems Ltd.). Exposure times varied between 30 seconds and 2 minutes. Further analysis and quantification of western blots was performed using GelQuant (DNR Bio-Imaging Systems Ltd.).

2.9 Dot Blot

Protein samples were prepared as outlined in section 2.7. Varying concentrations of the proteins were blotted onto either a 0.2µm or a 0.45µm pore size membrane. The membranes were left to dry for 30 minutes or for 1 hour. The dot blot was then blocked in 5% (w/v) skim milk in TBST for 1 hour. A primary antibody (see section 2.2.1) was

applied for 2 hours at room temperature or overnight at 4°C. The blot was washed 3 times in TBST, 5 minutes each wash. The blot was then incubated with the appropriate secondary antibody (see section 2.2.2) for 45 minutes at room temperature. The blot was washed 4 more times in TBST, 5 minutes each wash, then imaged using diluted Immobilon western chemiluminescent HRP substrate (Millipore) or Amersham ECL western blotting detection reagent (GE Healthcare) The blot was pictured using the MicroChemi System (DNR Bio-Imaging Systems Ltd.). Exposure times ranged between 3 and 5 minutes. Dot blots were confirmed to contain protein using a Ponceau S stain (0.5% (w/v) Ponceau S, 1% (v/v) glacial acetic acid).

2.10 Far Western Blot

Buffers

Protein binding buffer: 500mM NaCl, 20mM TRIS-HCl pH 7.5, 0.5M EDTA, 10% glycerol, 0.1% tween 20, 2% skim milk powder, 1mM DTT.

Note: It is important that the protein binding buffer is made fresh each time it is needed.

Procedure

Peptide N17WT was suspended in a solution of 12% glacial acetic acid and 9% dimethyl sulfoxide (DMSO) in deionized water to a concentration of 5mg/mL. Peptides N17M8P and N17(pS13)(pS16) were dissolved in PBS to a concentration of 5mg/mL. These peptides were then diluted to a concentration of 2.5mg/mL in a solution of 24% glacial acetic acid and 18% DMSO to lower the pH of the solution. The peptides were dot blotted onto a 0.2µm pore sized nitrocellulose membrane, each spot containing 10µg of peptide, and left to dry for 1.5-2 hours. The blot was incubated overnight at 4°C on a rocker in 5% skim milk powder in TBST. The following morning, fresh protein binding buffer was prepared and 10µg of the bait protein was added to 7mL of protein binding buffer. Bait proteins were GST-HAP40, GST or water. GST-HAP40 and GST were produced using the procedure outlined in section 2.7. Bait proteins were incubated with the dot blots for 2 hours at room temperature on a rocker. The blot was washed 3 times in TBST, each wash lasting 5 minutes. The primary antibody that recognizes GST (see section 2.2.1) was incubated on a rocker with the dot blots for 2 hours at room

temperature. This was followed by 3 additional 5 minute washes with TBST. The secondary antibody (see section 2.2.2) was applied and incubated with the blot for 45 minutes at room temperature on a rocker. Prior to visualization, an additional four 5 minute washes using TBST were applied. The dot blots were visualized using Amersham ECL western blotting detection reagent (GE Healthcare) and pictured using the MicroChemi System (DNR Bio-Imaging Systems Ltd.). Exposure time for these dot blots were between 3 and 5 minutes.

2.11 GFP Trap

Buffers

Lysis buffer: 10mM Tris-HCl pH 7.5, 150mM NaCl, 0.5mM EDTA, 0.5% NP-40
Protease inhibitor (Roche) was added to the lysis buffer prior to cell lysis.
Dilution/Wash buffer: 10mM Tris-HCl pH 7.5, 150mM NaCl, 0.5mM EDTA
2X SDS Sample buffer: 3x Blue Loading Buffer and 30x DTT (sold by New England Biolabs as a Blue loading buffer pack) diluted to 2x in deionized water.

Procedure

Cells were grown in 10cm plates and transfected with the appropriate YFP construct. Twenty-four hours post transfection, the cells were harvested and lysed and the cell lysate was frozen at -80°C. The following day, dilution buffer was added to the thawed lysate to bring the volume of the sample to 500µL. Magnetic GFP trap beads were incubated with the cell lysate for 2-2.5 hours at 4°C by end over end rotation. Following incubation, the beads were magnetically separated and an unbound fraction was removed from the supernatant and set aside for analysis. The beads were washed 3 times with wash buffer. To dissociate the protein from the beads and prepare for western blot, 2X SDS sample buffer was applied to the beads and the unbound fraction sample. The samples were then boiled at 100°C for 10 minutes. The beads were magnetically separated and the supernatant was collected. The supernatant and unbound fractions were stored at -20°C until run on a western blot (see section 2.8).

2.12 Peptide Competition Assay

Procedure

Samples containing the protein of interest were harvested, run on a gel, transferred and blocked using the western blot procedure in section 2.8. The lowest possible amount of HAP40 antibody to use was determined by running cell lysates of ST*Hdh*$^{Q7/Q7}$ cells on a western blot (see section 2.8) and conducting a serial dilution of the antibody. A 1:5000 dilution (1µg/5mL) of HAP40 antibody was the lowest concentration of antibody that could still recognize the protein in the sample.

This procedure is the same as a western blot (see section 2.8) except that the primary antibody was pre-incubated with peptide prior to application to the blot. The required volume of each peptide was calculated (see calculations below). To begin, our assay was conducted using a 200-fold molar excess of peptide to antibody, which was incubated in TBST prior to the application to the blot. When the 200-fold molar excess was unsuccessful, a 500-fold molar excess was tested. The peptide was incubated with the antibody in TBST for 1.5-2 hours at room temperature or for 1.5 hours at 4°C. Following incubation, 2.5mL of 10% (w/v) milk in TBST was added to create a solution of 5% (w/v) milk in TBST plus the antibody and peptide. This antibody/peptide/milk solution was added to the western blot instead of a primary antibody. The remainder of the western blot procedure was then completed.

Peptide Calculations

1. Find how many moles of antibody are being used in the solution.

$$mol\ of\ antibody = \frac{\mu g\ of\ antibody\ used\ in\ solution}{molecular\ weight\ of\ antibody}$$

2. Find out how many moles of peptide/protein are required for a 200-fold or 500-fold increase above the moles of antibody used.

For a 200-fold increase:

$$mol\ of\ peptide = mol\ of\ antibody \times 200$$

For a 500-fold increase:

$$mol\ of\ peptide = mol\ of\ antibody \times 500$$

3. Find how many grams of peptide/protein are needed.

$$grams\ of\ peptide = mol\ of\ peptide \times molecular\ mass\ of\ peptide$$

4. Find the volume of peptide/protein required for the amount of moles needed.

$$volume\ of\ peptide = mol\ of\ peptide \times concentration\ of\ peptide$$

In these experiments, the antibody is an IgG antibody with a molecular mass of 150kDa. The molecular mass of GST-HAP40 was 68kDa and the molecular mass of GST was 28kDa. The molecular mass of the HAP40 epitope peptide (GenScript) was 1.2kDa. The concentration of peptides varied. Deionized water was used in place of a peptide as a control.

2.13 Immunofluorescence

Buffers

Fixation Solution: 4% paraformaldehyde (PFA) in PBS or 100% methanol (ice cold)
Permeabilization Solution: 2% FBS, 0.5% Triton X-100 in PBS
Blocking Solution (Immunofluorescence): 2% FBS in PBS
Antibody Dilution Solution: 2% FBS, 0.5% Tween 20 in PBS

Procedure

Cells were seeded in 35mm glass bottom plates 24 hours prior to fixation. They were fixed one of two ways: using methanol or paraformaldehyde. Methanol was pre-chilled in a -20°C freezer. The cold methanol was applied to cells and the cells were either put on ice or were put into a -20°C freezer for a 12 minute incubation. Methanol was removed and the cells were washed 3 times with PBS. A four percent PFA solution was warmed to room temperature. It was then applied to cells for 20 minutes. Following incubation the cells were washed 3 times with PBS. The 4% PFA fixed cells were then permeabilized with either cold methanol (same procedure outlined above) or with permeabilization solution for 15 minutes at -20°C.

Following fixation, the cells were incubated 3 times in blocking solution for 15 minutes. The primary antibody was applied at the appropriate concentration in antibody dilution solution (see section 2.2.1) and left to incubate for 2 hours at room temperature or overnight at 4°C. The primary antibody was then removed and 3 additional washes

with blocking solution were applied as described above. The appropriate secondary antibody (see section 2.2.2) was applied and incubated at room temperature for 45 minutes. Following this incubation, the antibody was removed and three PBS washes were applied, each 5-10 minutes long. If a second primary antibody was to be used, the same procedure (without the fixation/permeabilization) was conducted with the second antibody where application of the second primary antibody immediately followed the final PBS washes. Cells were then stained with Hoechst 33342 (Bisbenzimide H) at a final concentration of 1μg/mL and incubated on cells for 10 minutes at room temperature. The cells were then washed 3 more times with PBS prior to imaging.

2.14 Widefield Fluorescence Imaging

Widefield fluorescent imaging was conducted on a Nikon Eclipse Ti inverted widefield epifluorescent microscope. The light source was a xenon arc lamp (Sutter Instruments). A 10x eyepiece and a plan apochromatic 60x objective lens (Nikon) with a numerical aperture of 1.4 were used with a corn-based oil (Cargille) (refractive index of 1.51). The images were captured with a Hamamatsu ORCA Flash 4.0 camera (Hamamatsu Photonics) using Nikon NIS Elements 4.0 (Nikon) software. Videos of live cells were taken over the course of approximately 8 hours with an image taken every minute. The cells were maintained at 33°C in 5% carbon dioxide. A neutral density filter, ND2 or ND4 (Nikon) was used to reduce photobleaching.

2.15 LumaScope Imaging

Cells were seeded and transfected by the methods described in section 2.6. A 40x air objective was used in the LumaScope inverted widefield epifluorescent microscope model 500 (Etaluma Inc., California). Cells were maintained at 33°C with 5% carbon dioxide and atmospheric oxygen levels. Imaging began 16-18 hours post-transfection and images were taken every 5 minutes over a 24-48 hour time period.

2.16 Chromatin dependence assay

Cells were seeded into a 35mm glass bottom dish and were transfected 24 hours later with 3μg of DNA (see section 2.6 for transfection methods). Twenty-four hours after

transfection, the cells were imaged. To avoid bias during imaging, the focus was shifted so that the cells were blurry and nuclear puncta were not visible. This method was chosen because transfection efficiency was not high enough to select fields using DIC. When a field of view was selected, it was focused and captured. The numbers of cells with and without nuclear spots were counted and an image was taken. Once imaged, each dish was returned to the incubator for approximately 1 hour. Hoechst 33342 (Bisbenzimide H) stain was mixed with media to a final concentration of 10μg/mL and applied to the cells. The cells were incubated for 3 hours then were fixed using 4% (w/v) paraformaldehyde (4% PFA) (20 minute incubation followed by three 5 minute 1xPBS rinses). The cells were imaged in the same way described above after fixation, and the numbers of cells with or without nuclear puncta were recorded. A total of 4 trials were completed for this experiment.

2.17 Lactate dehydrogenase activity assay

Cells were seeded into a 35mm dish and were transfected 24 hours later using Lipofectamine2000 (Invitrogen) according to the manufacturer's instructions. YFP-HAP40, HAP40-YFP and YFP were transfected into separate dishes. Two dishes were left not transfected, one as a non-transfected, low release control and the other for a high release control. Additionally, a sample of media was obtained to measure any background. One hour after transfection, samples of 800μL of media was collected from each well and was stored at -20°C until all the samples were collected. This was repeated every 6 hours until 6 samples were obtained. At the final time point, 800μL of the high release control was collected by incubating the designated dish for 10 minutes with a 2% solution of Triton-X in media. The Roche cytotoxicity kit, using lactate dehydrogenase (LDH) activity as an indicator of cell death was used according to manufacturer's instructions. The samples were loaded into a 96 well dish in triplicate. Absorbance readings were taken at 492nm to detect LDH activity and 692nm as a reference.

To calculate the percent cytotoxicity, the average absorbance was calculated from the 3 corresponding samples (including the controls). The average background

absorbance from the media sample was then subtracted from the samples. The percent cytotoxicity was then calculated using this equation:

$$\%\, cytotoxicity = \frac{transfected\ sample\ absorbance}{high\ control - low\ control} \times 100$$

2.18 Statistical Analysis

All statistical analysis was conducted using SigmaPlot 11.0 (Systat Software Inc.). When two groups were compared a student t-test was preformed if normality assumption was passed. If the normality assumption was not passed, the Mann-Whitney method was used. Mander's overlap and Pearson's correlation coefficient were measured using the ImageJ co-localization plugin.

For the western blot band quantification of HAP40 protein levels, band volume was measured using the GelQuant program (DNR Bio-Imaging Systems Ltd.). The band volume of the loading controls in each sample was normalized to the loading control in the first lane of each blot. The volumes of the HAP40 band were then adjusted up or down relative to the normalized value of protein present in the corresponding loading control. The statistical analysis using SigmaPlot 11.0 was completed.

Chapter 3 Results and Discussion

3.1 Fusion protein and Antibody Validation

3.1.1 HAP40 fusion protein

The mouse HAP40 fusion construct, F8A-tGFP, was obtained commercially from OriGene. YFP-HAP40 and HAP40-YFP fusion proteins were created using the F8A-tGFP fusion construct as a template. A flexible linker consisting of 6 glycines was added between the HAP40 protein sequence and the YFP fluorophore sequence to reduce any effect that the addition of the fluorophore may have on protein folding. When transfected into ST*Hdh*$^{Q7/Q7}$ cells, the fusion constructs all displayed similar localizations. The localizations were similar to the staining observed following immunofluorescence using a HAP40 antibody. HAP40 could be seen in the nucleus at punctate structures and diffuse throughout the cytoplasm, as seen in the transfection of ST*Hdh*$^{Q7/Q7}$ cells with F8A-tGFP in figure 3a. The opposite was also seen, where HAP40 had a punctate distribution throughout the cytoplasm and was diffuse through the nucleus. The puncta could also be seen in both compartments as demonstrated in figure3b where ST*Hdh*$^{Q7/Q7}$ cells were transfected with YFP-HAP40. Finally, HAP40 could be diffuse throughout the whole cell as seen in the transfection of F8A-tGFP in live ST*Hdh*$^{Q7/Q7}$ cells in figure 3c. In the literature, HAP40 fusion proteins were found in both the nucleus and the cytoplasm, in both punctate and diffuse patterns[79, 86]. These HAP40 fusion proteins contain the full mouse HAP40 sequence, confirmed by sequencing, and they display the correct localization, therefore, we can conclude that the constructs are producing YFP or tGFP tagged HAP40.

A HAP40 recombinant GST fusion protein was created and grown in *Escherichia coli* BL21 bacterial cells. This vector, prior to transformation into bacterial cells for expression, was also sequenced and determined to be mutation free, complete and in frame. The protein sample, isolated from the soluble fraction of the BL21 cell lysates, was run by SDS-PAGE. Figure 3d shows a Coomassie stained gel with a band present at 66kDa, the calculated molecular weight for GST-HAP40 (40kDa for HAP40 and 26kDa

for GST). An antibody that recognizes HAP40 when a western blot of the same protein sample is analyzed also labels this band as seen in figure 3e. This, along with the complete sequencing of the vector, suggests that the HAP40 fusion protein was created correctly. Since the yield of GST-HAP40 protein was low, future experiments should include analysis of the portion of cell lysate that was not bound to the beads to ensure that all the available GST-HAP40 is binding to the beads. In addition, analysis of the insoluble phase should also be included to determine if GST-HAP40 is sequestered into inclusion bodies in the bacterial cell.

3.1.2 HAP40 antibody

A commercially available antibody that recognizes mouse HAP40 at the carboxyl terminus was used in these experiments. The antibody was acquired from Abcam where a western blot was used to validate the antibody (Abcam, personal communication). Since a western blot cannot determine the specificity of the antibody binding to the target protein, we conducted a variety of experiments in an attempt to validate the antibody.

A western blot was run using ST*Hdh*$^{Q7/Q7}$ and HEK293 cell lysates to determine if the antibody detects a signal band at the appropriate molecular weight in only mouse cells. Figure 4a demonstrates that this antibody does recognize protein at approximately 45kDa in ST*Hdh*$^{Q7/Q7}$ cell lysates, but does not in human HEK293 cell lysates. Although this molecular weight is larger than the calculated mass of 40kDa[86], this signal could be from HAP40 since the protein bands that are produced when prestained ladders are run by SDS-PAGE are an approximation of the assigned molecular weight. The observed size of HAP40 on the western blot is approximately 5kDa greater then the 40kDa estimate of the molecular weight of HAP40, and it is therefore possible that this band contains HAP40.

Immunofluorescence conducted on fixed ST*Hdh*$^{Q7/Q7}$ cells using this antibody displayed weak nuclear staining and strong cytoplasmic staining as seen in figure 4b and 4c. When primary antibody concentrations were increased, punctate nuclear staining was observed, which is demonstrated in figure 4b. The control in figure 4d shows that when the primary antibody was excluded, no signal occurred; therefore, the staining observed was specific to the HAP40 antibody in the test plate. Both the cytoplasm and/or the

nucleus contained either diffuse or punctate staining of HAP40, which is consistent with the staining obtained in previous publications[79, 86]. It is also consistent with the localization of the HAP40 fusion protein constructs expressed ST*Hdh*$^{Q7/Q7}$ cells and described in section 3.1.1.

Small interfering RNA (siRNA) was used to reduce the HAP40 protein levels in ST*Hdh*$^{Q7/Q7}$ cells. A scrambled siRNA was used as a control to ensure the result was not due to the transfection of the siRNA. Figure 4e demonstrates that the HAP40 levels in lane 3, as detected by the HAP40 antibody, remained the same as the non-transfected control in lane 1 and the scrambled siRNA control in lane 2. This indicates that the HAP40 antibody could be recognizing a different protein at approximately 40kDa on the western blot, or the siRNA is not working correctly. Analysis of the HAP40 mRNA levels following siRNA transfection is required to ensure that the siRNA is working correctly.

The HAP40 antibody was raised against a peptide corresponding to amino acids 314 to 325 of mouse HAP40. The mouse peptide sequence is different than the corresponding human sequence as seen in figure 5a, where the different amino acids are highlighted in red. Of the 12 residues in the peptide, there are only 5 consecutive residues that are the same. This antibody should not recognize human HAP40. A panel of human and mouse cell lysates were run on a western blot and probed for HAP40 using the HAP40 antibody. The western blot in figure 5b shows that most of the human cell lysates were negative for any signal from the antibody, and all of the mouse cell lysates contained a protein at 45kDa that was identified by the antibody. HeLa cells were the only human cell line that contained a protein that was recognized by the antibody.

HeLa cells are immortalized human cervical cancer cells and he most common cause of cervical cancer in females is infection with the human papillomavirus (HPV)[91, 92]. When a BLAST search was run using the HAP40 epitope peptide sequence and specific search restrictions to HPV and associated proteins, many similarities were identified. Figure 5c shows that the major capsid protein of HPV type 147 contains 6 consecutive amino acids at the carboxyl terminus of the peptide (boxed in green), and upstream of this sequence are two groups of two consecutive residues (also boxed in

green), which match the sequence of the epitope peptide. Additionally, the predicted molecular weight of the HPV type 147 major capsid protein is approximately 59kDa, which is very close to the molecular weight at which the band appears on the blot in figure 5b. Human HAP40 is 10 amino acids smaller than mouse HAP40; therefore the HAP40 band in human cells should be smaller then the band seen in mouse cells, the opposite of what is seen on the western blot. Given the epitope similarities and the size of the protein, the antibody may be detecting the capsid of HPV type 147 and future experiments are needed to determine if this is HPV, HAP40 or a different protein. The HeLa cell genome and transcriptome is full of errors when compared to normal human cells, including incorrect chromosome numbers and many single nucleotide polymorphisms[91]. As a result, the signal from the HeLa cell population should be disregarded in this experiment. Human endogenous HAP40 was not recognized in any of the other human cell lines; therefore, the HAP40 antibody recognizes mouse HAP40 and not human HAP40.

Experiments to determine if the antibody could recognize overexpressed HAP40 and HAP40 recombinant peptides were completed. The three ST*Hdh* cell lines, ST*Hdh*$^{Q7/Q7}$, ST*Hdh*$^{Q7/Q111}$, and ST*Hdh*$^{Q111/Q111}$ were transfected with YFP-HAP40 and the cell lysates were run on a western blot. The blot was first probed using the HAP40 antibody and the results are seen in figure 6. The YFP-HAP40 construct was expected to weigh approximately 67kDa (40kDa for HAP40 and 27kDa for YFP), and, as such, a band was expected on the western blot probed for HAP40 at approximately this 67kDa. Figure 6 shows that no band is present at 67kDa on the anti-HAP40 blot; however, the band previously identified to be endogenous HAP40 at approximately 45kDa is still present. When the same blot was probed with an antibody recognizing YFP, the fusion proteins were detected at 67kDa (see Anti-YFP blot in figure 6). The same results were obtained when HAP40-YFP and F8A-tGFP were expressed and analyzed by western blot.

Folding of a protein can greatly affect the ability of an antibody to recognize an antigen. The folding of the protein when the YFP or tGFP fluorophore is attached could be masking the epitope from the antibody. Although there is a flexible linker consisting of

6 glycines between YFP and HAP40 in the YFP-HAP40 and HAP40-YFP fusion constructs, the fusion protein could be folding in such a way that the epitope is blocked by the fluorophore. Though it is possible that this epitope masking occurred, it is unlikely since the cell lysates are thoroughly denatured prior to analysis by SDS-PAGE. Another possibility is that the epitope on HAP40 is required to have a particular structure to be recognized by the antibody, and since it is denatured for western blot analysis, the structure is not present and the epitope not recognized. This possibility also seems unlikely, as the endogenous HAP40 can be seen when denatured on a western blot; however, if the endogenous HAP40 on the blot is not actually HAP40, this could be true.

A GST-HAP40 recombinant protein was produced in an *Escherichia coli* BL21 bacterial system and was identified using the HAP40 antibody. The samples were analyzed by western blot where the protein was denatured and by dot blot where it was not denatured and maintained the folded structure that was produced in the bacterial cells. Figure 7a and 7b demonstrates that the HAP40 antibody in both western blot and dot blot applications recognized the GST-HAP40 recombinant protein at 66kDa. Figure 7a shows a western blot where purified recombinant protein from the soluble phase of cultures induced with IPTG (i) and not induced with IPTG (ni) were run. The HAP40 antibody recognized the GST-HAP40 fusion protein at 66kDa. It also identified a HAP40 recombinant protein present at approximately 45kDa that was cleaved from the GST tag. The blot probed with a GST antibody identifies GST-HAP40 at 66kDa as well as the cleaved GST tag present at 26kDa. The band present at approximately 66kDa in the induced GST lane of the blot probed with the HAP40 antibody in figure 7a is likely due to the large amount of protein loaded into that lane. This blot was not equally loaded. The dot blot in figure 7b shows that there is no detectable HAP40 in the GST alone samples when equal amounts of protein are loaded onto a dot blot. Figure 7c is a blot probed with only secondary antibody and exposed for 5 minutes, demonstrating that the signal from the blot probed with HAP40 is due to identification of HAP40 by the HAP40 antibody.

Figure 7a shows the isolated protein from the pGEX gene fusion system. Levels of induced GST-HAP40 isolated in this system are very low compared to the level of

induced GST isolated. This may be due to the system in which the proteins are expressed. Bacteria lack the machinery required for translation of rare codons, post-translational modifications, and complex protein folding that is seen in eukaryotic cells. When overexpressing protein in bacteria using recombinant systems, it is not unusual for the overexpressed fusion protein to fold incorrectly. This improper folding can lead to degradation or incorporation into inclusion bodies[93]. The drastic difference in isolated protein levels in figure 7a suggests that GST-HAP40 is not able to fold correctly and is either degraded or sequestered into inclusion bodies. The soluble fraction of cell lysates are used to isolate the fusion protein and if HAP40 is sequestered in the insoluble phase it would not be isolated and not seen on the gel. This may be why the purified recombinant GST-HAP40 protein levels were so low.

The conflicting results obtained between recognition of recombinant GST-HAP40 and the other mammalian derived HAP40 fusion proteins could be due to post-translational modifications. The antibody was raised against a synthetic peptide that has no post-translational modifications. Bacterial cells also lack machinery required for post-translational modifications; therefore the GST-HAP40 fusion protein produced in the BL21 cells will not be modified. The YFP-HAP40, HAP40-YFP and F8A-tGFP constructs were expressed in a system where post-translation modifications are possible. If the HAP40 fusion proteins were modified at the epitope then the antibody would not be able to recognize the fusion protein as we saw with our western blots; however, it is unlikely that all of the overexpressed HAP40 fusion constructs would be modified. Some signal would be expected on the western blots identifying the fusion protein since we know that the epitope is intact in these fusion proteins and since we can identify the endogenous HAP40.

A peptide competition assay was the final test conducted to establish if this antibody recognized HAP40. A concentration of antibody was found in which the HAP40 signal was detectable but not very strong on a western blot. Immunoglobulin G (IgG) antibodies have 2 binding sites which can each bind one peptide. The first experiments used a 200-fold molar excess of bait peptide, a 100-fold molar excess of peptide per

binding site on the antibody. This was later increased to a 500-fold molar excess of bait peptide to antibody. The bait peptide was pre-incubated with no blocker in an attempt bind the bait peptide to the antibody. If the bait peptide was bound to the antibody, the antibody could not bind to the protein on the blot and therefore, no signal would be seen on the western blot. The bait peptides used were: GST-HAP40, the HAP40 epitope peptide (the mouse sequence in figure 5a), GST, and water. Figure 8 shows the typical results obtained from these experiments. Incubation with the GST-HAP40 protein or the HAP40 epitope peptide was unsuccessful at reducing or eliminating the signal seen on the western blot where ST*Hdh* cell lysates were run. The volume of bait peptide used for each trial was very high and should have been sufficient to bind to the antibody and prevent interaction on the western blot, but it was not. We know from previous dot blot and western blot data (see figure 7) that the antibody does recognize the GST-HAP40 fusion construct, so recognition of the peptide itself was not an issue.

Normally, antibody-epitope/antigen interactions are quite strong; however, some antibodies can have lower affinities for their antigen than others. Peptide specific antibodies in particular are known for having a moderate binding affinity for their corresponding peptide antigens[94]. This antibody is a polyclonal antibody made using a peptide; therefore, it will have multiple antibodies that recognize multiple epitopes with varying affinities. It is possible that the mixture of antibodies have an average low affinity for HAP40 allowing for a more reversible interaction. If the antibody-peptide interaction breaks apart while the complex is exposed to the blot, the antibody would then be free to recognize the protein on the blot. In an attempt to reduce the amount of time available for this to happen, incubation times were decreased in our experiments, but this did not lead to any change in the results.

The final test to validate an antibody is reproducibility. The reproducibility between lots of this antibody has not yet been tested since there currently is only one lot available. A final step in proper validation of an antibody would be to repeat the same tests with a new lot and determine if the results are the same.

When all this information is taken together, it is difficult to trust that this antibody is able to recognize HAP40. The apparent molecular weight on a western blot is at the correct weight, and the localization of the HAP40 signal in immunofluorescence experiments is consistent with the literature and the localization of validated fusion proteins. As expected, the antibody does not recognize human HAP40, but does detect recombinant mouse HAP40. Attempts at reducing HAP40 levels using siRNA were not successful, however it is not clear whether the siRNA is not working, or if the antibody is not detecting the reduction because it is not detecting the correct protein. The antibody does not recognize the YFP/tGFP fusion proteins on western blot, regardless of where the fluorophore is located; however, the antibody does recognize the GST-HAP40 recombinant protein on both dot and western blots. In the peptide competition assay, pre-incubation with bait peptides was unsuccessful at eliminating or reducing signal on the western blot; however, the binding affinity of the antibody for the antigen may be low. When taken together, the HAP40 antibody does not identify human HAP40 but it likely does detect mouse HAP40.

3.2 Endogenous HAP40 levels are highest in ST*Hdh*$^{Q7/Q111}$ cell lines.

The interaction between huntingtin, HAP40 and Rab5 occurs when the levels of HAP40 are increased; therefore the levels of HAP40 in our cell lines were quantified to determine the cell line with the highest amounts of HAP40. In a previous publication, Pal *et al* (2006) showed that HAP40 levels were highest in ST*Hdh*$^{Q111/Q111}$ cells and lowest in ST*Hdh*$^{Q7/Q7}$ cells by western blot. To ensure that the HAP40 levels in our cell lines are consistent with these results, we conducted a western blot using the cell extracts of ST*Hdh*$^{Q7/Q7}$, ST*Hdh*$^{Q7/Q111}$ and ST*Hdh*$^{Q111/Q111}$ cells. Figure 9a demonstrates that in our cell lines, ST*Hdh*$^{Q7/Q111}$ cells had the highest level of HAP40, and ST*Hdh*$^{Q111/Q111}$ cells had the least amount of HAP40. The HAP40 band volume of the western blot was quantified and adjusted to reflect the differences in loading controls. The results are shown graphically in figure 9b. A student t-test analysis was conducted to compare the levels of HAP40 in ST*Hdh*$^{Q7/Q7}$ cells to the HAP40 levels in ST*Hdh*$^{Q7/Q111}$ cells or ST*Hdh*$^{Q111/Q111}$ cells. The difference in HAP40 levels between ST*Hdh*$^{Q7/Q7}$ cells and

ST*Hdh*$^{Q7/Q111}$ cells was statistically significant (p=0.042). Additionally, the difference in HAP40 levels between ST*Hdh*$^{Q7/Q7}$ cells and ST*Hdh*$^{Q111/Q111}$ cells was also statistically significant (p=0.004).

The changes in HAP40 levels are likely a result of altered transcription by mutant huntingtin. Many transcription factors contain domains that are activated by glutamine-rich sequences, such as the polyglutamine tract found in huntingtin[43]. Both wildtype and mutant huntingtin can associate with DNA; however, mutant huntingtin has a greater binding affinity and activates transcription more than wildtype[43]. If HAP40 transcription is activated by mutant huntingtin, a cell expressing two copies of mutant huntingtin would be expected to have higher levels of HAP40 then a cell expressing one copy of mutant huntingtin. Pal *et al* (2006), using the same HAP40 antibody, found that ST*Hdh*$^{Q111/Q111}$ cells had the highest amount of HAP40 and ST*Hdh*$^{Q7/Q7}$ cell lines had the lowest levels of HAP40 protein. This fits the hypothesis. When our cell lines were analyzed, HAP40 levels were decreased in ST*Hdh*$^{Q111/Q111}$ cell lines, far below the HAP40 levels detected in ST*Hdh*$^{Q7/Q7}$ cell lines and ST*Hdh*$^{Q7/Q111}$ cells had the most HAP40. This contradictory data could be due to the cell lines. We have hypothesized that HAP40 overexpression is toxic. If this is the case, when the cells were selected for immortalization, the stronger cells with less HAP40 will survive and the weaker ones, with more HAP40, will not. Pal *et al* (2006) may have had a different clonal cell population.

3.3 Overexpression of HAP40 may be toxic to the cell

We have demonstrated in section 3.2 that HAP40 protein levels are increased in our ST*Hdh*$^{Q7/Q111}$ cell lines. Previously, HAP40 levels in human HD patient samples (brain tissue and fibroblasts) and in ST*Hdh*$^{Q7/Q111}$ and ST*Hdh*$^{Q111/Q111}$ cell lines were also shown to be elevated compared to the wildtype counterparts[79]. Additionally, the pathogenic complex between huntingtin, Rab5 and HAP40 only occurs when HAP40 protein levels are increased[79]. This suggests that an increase of HAP40 protein levels may be toxic.

LumaScope movies were set up for 24-48 hour periods to watch the expression of YFP-HAP40 in ST*Hdh*$^{Q7/Q7}$ cell lines for an extended period of time. At the beginning of the experiment, 16 to 18 hours post transfection, many cells were expressing YFP-

HAP40; however, as time passed, these cells began to die. As seen in the movie stills in figure 10a, by the end of the 24 hour acquisition time, many of the transfected cells were condensed, floating and dead. The YFP control in figure 10b still contained many live cells. These movies were used to determine if cell death occurred and the extent to which cell death occurred, further experiments are required to quantify the amount of cell death.

The location of the fluorophore tag on the HAP40 fusion protein affects the expression of HAP40. Figure 11a shows a western blot of lysates from the three ST*Hdh* cell lines, each transfected with YFP-HAP40 or HAP40-YFP. The expression level of HAP40-YFP protein was much lower than the expression of YFP-HAP40 in each of the cell types. Additionally, when the ST*Hdh*$^{Q7/Q7}$ cells are imaged following transfection with either YFP-HAP40 or HAP40-YFP, fewer cells appear transfected in the HAP40-YFP dishes as demonstrated in figure11b. It is unclear why there is such a difference in transfection efficiency. One possibility is that the transfections are not equal, though the same pattern is seen over multiple trials. Transfection with a second empty vector would rule out questions about equal transfections. Another possibility is that the fusion protein is not folded correctly due to the addition of the YFP fluorophore and is degraded by the proteasome or clumped into insoluble aggregates.

The higher expression level of the YFP-HAP40 construct compared to the HAP40-YFP construct may be because of a potential inhibitor of apoptosis protein binding motif (IBM) located at the amino terminus of the protein. This motif is predicted by ELM (the Eukaryotic Linear Motif resource for functional sites in proteins) to occur within the first five amino acids of HAP40[95]. It is present in proapoptotic proteins that bind to a family of antiapoptotic proteins called inhibitors of apoptosis proteins (IAP). IAPs are a group of proteins that inhibit caspases in non-apoptotic cells. The important sequence in IAPs is the baculovirus IAP repeat (BIR) domain, which binds to the caspase[96-98]. There are 3 types of BIR domains: BIRI, BIRII and BIRIII. IBM containing proteins that bind the BIR domain of IAPs prevent the BIR domains of the IAPs from binding to caspases, which promotes apoptosis[97, 98]. The IBM consensus sequence requires an alanine residue in the first position. The consensus motif in HAP40 is located at the amino terminus, after

the initiator methionine residue. When the initiator methionine is removed by methionine aminopeptidase, the amino terminal alanine is revealed[99]. This sequence specifically binds to BIRII domains on IAPs. When the initiator methionine is removed from HAP40-YFP, this sequence is revealed, however, when the initiator methionine is removed from the YFP-HAP40, this sequence is not revealed. Apoptosis is a highly regulated cellular process[99] and therefore, a protein that promotes apoptosis should be tightly controlled in the cell. If HAP40-YFP is a proapoptotic protein, its expression may be repressed. Another possibility is that the cells expressing HAP40-YFP died quickly and were therefore never seen on the western blot since dead cells in the media were not collected with the protein samples. Further investigation is required to determine if this IBM is functional. Amino terminal alanine residues can be acetylated in the cytoplasm, rendering them unable to bind to the BIR domain[99] and, if HAP40-YFP is acetylated, the IBM consensus sequence activity would be lost. Additionally, it is unclear if overexpression of a construct containing an IBM consensus sequence would be sufficient to inhibit IAPs enough to induce apoptosis, especially since IAPs work with other proapoptotic factors to carry out apoptosis.

Many methods were attempted to determine cell viability upon HAP40 overexpression. An apoptosis assay using immunofluorescence to stain cytochrome c was used to determine if cells transfected with YFP-HAP40 were undergoing apoptosis. During apoptosis, cytochrome c is released from the mitochondria into the cytoplasm and forms an apoptosome with other proapoptotic proteins. The apoptosome then activates caspase 9, which activates other proteases for a controlled cell death[100]. In a healthy cell, cytochrome c is found at the mitochondria, but in an apoptotic cell, it is found more diffuse in the cytosol. Cells were transfected with YFP-HAP40, HAP40-YFP and YFP and 2 control dishes were left not transfected. The one control dish was treated with staurosporine to induce apoptosis; the other dish was left untouched. All 4 dishes were fixed and immunofluorescence using a cytochrome c antibody was conducted. The analysis was to be done by counting the number of transfected cells and the number of apoptotic transfected cells. When analyzed, it was not easy to tell which cells were

undergoing apoptosis and which ones were not. The cells undergoing apoptosis were extremely bright, similar to the cells treated with staurosporine, whereas the healthy cells were quite dim; however, many of the transfected cells were in between and it was difficult to class them individually. As a result, a better method for quantifying apoptotic cells was considered.

Propidium iodide (PI) stain was used to look for transfected cells that were dying. PI is a dye that cannot enter the cell unless the membrane is permeabilized, therefore healthy cells exclude the dye and unhealthy cells take up the dye. This method was also difficult to quantify on the microscope, so flow cytometry was used to determine how many cells were transfected and healthy and how many were transfected and unhealthy. The results from the flow cytometry experiments were inconclusive as the controls showed similar uptake of PI to the tests. During these experiments our cell lines were stressed due to a change of FBS used in the media and, shortly afterward, detection of mycoplasmosis. These issues have recently been addressed through treatment of mycoplasmosis with plasmocin and an upgrade to a higher quality FBS and these experiments can now be continued.

Lactate dehydrogenase (LDH) assays were conducted on YFP-HAP40, HAP40-YFP and YFP transfected cells. LDH is an enzyme present in the cytoplasm of cells. When the cell dies and the membrane is permeable, the LDH is released from the cell. The assay measures the amount of LDH released from dying cells into the media. If cells die quickly, LDH is released faster, and if the cells die slowly, the LDH is released slowly. Two trials were performed with ST*Hdh*$^{Q7/Q7}$ cells, one dish was not transfected and three dishes were transfected each with YFP, HAP40-YFP or YFP-HAP40. During the first trial, samples were collected every 6 hours for a total of 31 hours. When analyzed, the results revealed no difference between the time points. When the cells were lysed and run on a western blot, transfection efficiency for HAP40-YFP was very low. This may be the reason why there was no change in the LDH levels. Additionally, samples were taken only up to 31 hours post-transfection. Longer times between samples and a longer expression time were required. During the next trial, cells were transfected in

the same way and samples were harvested every 12 hours for 60 hours. When analyzed, the percent cytotoxicity did increase for the tests, but it also increased for the controls. This was likely due to the confluency of the dish, which, by the end of the trial, was very confluent. Additionally, there was still very little expression of HAP40-YFP. Although the LDH assay is an excellent indicator of cell death, it may not be appropriate for the toxicity of HAP40. Since the expression levels of HAP40-YFP are so low, it would be best to look at the population of transfected cells on a cell-by-cell basis as opposed to a total cell population.

Since many of the experiments that test the toxicity of HAP40 thus far have been inconclusive, further experiments are required to determine if HAP40 overexpression is toxic.

3.4 The Huntingtin-HAP40 interaction: Amino or carboxyl terminus?

Pal *et al* (2006) narrowed the interaction region between HAP40 and huntingtin to huntingtin amino acids 2226-3144. Areas between the HEAT repeats in this region that are conserved between species were chosen for our investigation. We created two constructs corresponding to huntingtin amino acids 2570-2700 and 3110-3144. These constructs were tagged with a red mCherry fluorophore creating the constructs H2570-2700-mCh and H3110-3144-mCh. They were then co-transfected with F8A-tGFP into ST*Hdh*$^{Q7/Q7}$ cells and analyzed for co-localization at endosomes with HAP40. Figure 12a shows that there is co-localization between H2570-2700-mCh, identified by the location of the arrows. Figure 12b demonstrates that there is no co-localization between H3110-3144-mCh at endosomes. To identify a smaller region where HAP40 interacts with huntingtin, the H2570-2700-mCh construct was split in half to form H2570-2634-mCh and H2635-2700-mCh, and each construct was co-transfected with F8A-tGFP into ST*Hdh*$^{Q7/Q7}$ cells. Figure 13a demonstrates the co-localization between H2570-2635-mCh and F8A-tGFP at the arrows. Figure 13b shows that H2635-2700-mCh does not co-localize with F8A-tGFP at any puncta. Since co-localization only puts these two proteins in close proximity to each other and does not necessarily mean that there is an interaction,

other methods were required to search for an interaction between HAP40 and the huntingtin fragments.

GFP-trap magnetic beads were used to determine if there was an interaction between the YFP fusion constructs and HAP40. These magnetic beads use a proprietary GFP binding protein that binds to the GFP or YFP fluorophore. ST*Hdh*$^{Q7/Q111}$ cells were transfected with H2570-2700-YFP, H2570-2634-YFP, H2635-2700-YFP and H3110-3144-YFP. An empty YFP vector was transfected as a negative control. In addition, a construct corresponding to the first 17 amino acids of huntingtin fused to YFP (N17-YFP) was also used as a negative control as previous papers have stated that there is no interaction between HAP40 and the amino terminus of huntingtin[79, 86]. When the unbound fraction (the sample taken after incubation with the beads) and the bound fraction (the sample attached to the beads) were run on western blot, an unexpected interaction was discovered. Figure 14a shows that in the ST*Hdh*$^{Q7/Q111}$ cell lines where HAP40 levels are increased, N17-YFP pulled down HAP40, not the carboxyl terminal fragments. The YFP portion of the blot demonstrates that all of the constructs were successfully transfected, although, on the YFP blot in figure 14a there are ghost bands identified by *'s. These bands contain high levels of protein where the ECL substrate was used quickly. Although there is no signal from the ECL in this area, there are still high levels if protein present. Next, ST*Hdh*$^{Q7/Q7}$ and ST*Hdh*$^{Q111/Q111}$ cell lines were transfected with YFP and N17-YFP to determine if this interaction only occurred in ST*Hdh*$^{Q7/Q111}$ cell lines. Figures 14b and 14c show that there is no interaction between N17 and HAP40 in ST*Hdh*$^{Q7/Q7}$ or ST*Hdh*$^{Q111/Q111}$ cell lines. The interaction between N17-YFP and HAP40 is exclusively in ST*Hdh*$^{Q7/Q111}$ cell lines, where HAP40 levels are elevated. Additionally, this interaction is very sensitive to cell stress. When heat shocked for 10 minutes at 42°C, the interaction disappears (figure 14d, compare lanes 6 and 8 – note that this was only done one time, additional repeats are necessary). When the cells in our lab became stressed due to FBS and mycoplasma infection, this interaction disappeared and we were unable to analyze it further using this method.

A far western blot was completed to confirm the interaction between HAP40 and N17-YFP. The far western blot required optimization. First, a new construct was created which included both the carboxyl terminal fragments of huntingtin into one construct. This new fragment corresponded to huntingtin amino acids 2570-3144 and was tagged with YFP. ST*Hdh*$^{Q7/Q111}$ cell lysates transfected with the huntingtin constructs N17-YFP, H2570-3144-YFP and YFP were analyzed on a western blot and incubated with the bait protein GST-HAP40 and associated controls GST, and water. The resulting blots were blank and no interactions with any fragments were seen.

Next, we focused on the interaction between N17 and HAP40. GST, GST-HAP40 and ST*Hdh*$^{Q7/Q111}$ cell lysates were run on the western blot and incubated with the bait peptides: N17, N17M8P (a mutant that disrupts the alpha-helicity of the N17 peptide), N17(pS13)(pS16) (N17 where serines 13 and 16 are phosphorylated), and water. This method also did not work because the signal on the control blots was just as strong as the signal on the test blot. When the N17 antibody that was used to detect the N17 peptides was tested for recognition of HAP40, it was discovered that the N17 antibodies also see HAP40. Since the N17 peptides are small, it would be difficult to run them on a western blot. A first attempt was made using a 16:1 acrylamide to bis-acrylamide gel. The peptides did not transfer well to the membrane so the dot blot method of blotting the peptides onto the membrane was chosen. The isoelectric point of these peptides is approximately 8.9. When the peptide pH was lowered to approximately pH 3, the peptides became more positively charged and were able to bind to the membrane. The blots were incubated with the bait proteins: GST-HAP40, GST, and water. The blot was then probed using the anti-GST antibody. Figure 14e reveals that by this method, the N17 peptide does interact with GST-HAP40.

Here, we show by far western blot that HAP40 interacts with the first 17 amino acids of huntingtin when huntingtin is not phosphorylated. We also show, using the GFP trap method that HAP40 interacts with N17-YFP in ST*Hdh*$^{Q7/Q111}$ cells; however the phosphorylation status of this construct is unknown. The far western blot uses synthetic peptides and recombinant proteins to demonstrate an interaction. This system is artificial

and lacks any potential modifications that occur within a living cell. The GFP trap experiment looks at cell lysates from the cell and isolates the protein of interest in a complex; however, the over expressed N17-YFP fusion protein could produce false interactions that do not occur when levels are normal. This experiment also uses only the cell lysates of cells, killing the cell in order the harvest the protein. It would be beneficial to have experiments conducted in live cells to confirm that the interaction occurs in live cells. Live cell experiments could also tell us where and when the interaction occurs. Since huntingtin is involved in a variety of cellular activities, this information could help to reveal a cellular function of HAP40. It is still possible that HAP40 interacts with the carboxyl terminal portion of huntingtin. The full sequence from amino acids 2226 to 3144 identified by Pal *et al* (2006) to interact with huntingtin was not fully examined.

3.5 HAP40 nuclear puncta are dependent on chromatin structure

HAP40 nuclear puncta are observed with both the HAP40 antibody and the HAP40 YFP fusion proteins. A chromatin dependence assay was used to determine if the presence of the nuclear puncta were dependent on chromatin structure. To test this, Hoechst 33342 (Bisbenzimide H), a fluorescent blue dye that binds to the minor groove of DNA, was used to disrupt chromatin structure. The addition of hoechst to YFP-HAP40 transfected cells significantly reduced the amount of cells with YFP-HAP40 nuclear puncta from 45.0% to 30.0% ($p=0.029$) indicating that YFP-HAP40 nuclear puncta are dependent upon undamaged chromatin structure.

Phosphorylated huntingtin (phospho-huntingtin) also forms puncta in the nucleus. In a 2011 publication from our lab, a similar experiment demonstrated that the phospho-huntingtin spots present in the nucleus were dependent upon undamaged chromatin structure[22]. HAP40 may be interacting with phospho-huntingtin at these nuclear spots, although the interaction would likely not be with N17 since our far western blot results in section 3.4 indicated that HAP40 does not interact with phosphorylated N17(pS13)(pS16) peptides.

3.6 HAP40 Cleavage Products

While testing the levels of HAP40 in ST*Hdh* cell lines, breakdown products of varying sizes were noted. The western blot in figure 15a demonstrates the differences in breakdown products between each cell line. In ST$Hdh^{Q7/Q7}$ cell lines, both 30kDa and 35kDa breakdown products occur (see figure 15a red box). In ST$Hdh^{Q7/Q111}$ cell lines, only the 35kDa breakdown band is observed (see figure 15a blue box) and, in ST$Hdh^{Q111/Q111}$ cell lines, only the 30kDa band is seen (see figure 15a purple box). These results suggest a change in the breakdown of HAP40 in each of these cell lines. When the HAP40 sequence is entered into the protein analysis software $(PS)^2$, a large loop is predicted to occur between the helices at amino acid positions 75 and 95[101]. If HAP40 is cleaved at this loop, the fragments that could be produced fall into the size range of the breakdown products seen on the western blot, suggesting that the difference in breakdown size in each cell type is a result of different cleavage patterns.

Since the cleavage of HAP40 could be a result of how the cell lysates were handled during the lysis process, we sought to determine if the cleavage of HAP40 could be seen in the cell. A HAP40 construct was created with a fluorophore tagged at each end of the HAP40 protein sequence. A venus fluorophore was tagged to the amino terminus of HAP40 and an mCherry fluorophore was tagged to the carboxyl terminus, creating venus-HAP40-mCherry. The fluorophores were separated from the HAP40 sequence with a flexible linker of 6 glycine residues. This construct was then transfected into ST$Hdh^{Q7/Q7}$ cell lines. The venus fluorophore and the mCherry fluorophore will co-localize perfectly if there is no cleavage. In figure 15b there are areas (arrows) where the venus fluorophore and the mCherry fluorophore do not co-localize suggesting that there is cleavage of HAP40. This is seen in a small subset of transfected cells in the dish. The co-localization between channels can be measured using Mander's overlap and Pearson's correlation coefficient. A Mander's overlap value of 1 indicates perfect co-localization, and a value of 0 indicates no co-localization. A Pearson's correlation coefficient value of -1 indicates that there is negative correlation between the channels, and a value of +1 indicates that there is positive correlation between the channels. The image was first cropped to reduce

the amount of background and to include only one of the cells in an image. The resulting image was then deconvolved to remove unfocussed light. Using image J and an image for the cell on the left, Mander's overlap was calculated to be 0.931 and the Pearson's correlation value was 0.840. Both of these values indicate that there is not perfect co-localization between the channels as would be expected if the construct remained intact. No conclusions can be drawn from this data since the sample population is only a small amount of cells. Cleavage of the HAP40 construct did occur in multiple trials. There was also no empty vector control to ensure that the cleavage is occurring due to cleavage of HAP40 and not cleavage of the vector. A larger sample population and controls using a venus-MCS-mCherry construct are required. The localization of venus-HAP40-mCherry in the majority of transfected cells is diffuse in both the nucleus and the cytoplasm.

Since the cleavage of HAP40 did occur in the cell, cleavage prediction software was used to identify potential cleavage sites in HAP40. Amino acid 81, a phenylalanine located within the proposed loop which could produce breakdown products at the correct size, was predicted to be cleaved by the protease, calpain[102]. Calpains are a family of cysteine proteases that are activated in the presence of intracellular calcium[103, 104]. Activation of calpains can occur under normal physiological conditions where activation is controlled and only a small subset of calpains are activated; or, under pathological conditions where a high intracellular concentration of calcium causes activation of all calpains in the cell[103]. In HD, active and inactive calpain levels are increased[105]. Furthermore, deregulation of calcium levels occurs in HD cells. Mutant huntingtin associates with mitochondria and with HAP1 and a calcium channel on the endoplasmic reticulum, both of these associations contribute to the deregulation of calcium levels in HD cells[106]. This change in calcium signaling in the presence of mutant huntingtin could cause an increase in calpain activation and an increase in cleavage of calpain substrates. Chemical treatments were investigated to determine the effect of an increase or decrease of calpain activity with little success. The treatments may not have been strong or long enough to see an effect on HAP40 cleavage.

At this time it is early to draw conclusions about the HAP40 cleavage products. There is little information about the proteases that may be cleaving the protein and where the cleavage would occur. Although a potential loop has been identified, this is only a predicted structure and the potential cleavage sites are also only predictions. Prior to further investigation of these cleavage products, the bands on the blot should be identified by mass spectrometry to determine what sequences of HAP40 are present. This would provide a strong starting point to continue analysis of these breakdown products.

Chapter 4 Conclusions and Future Directions

The HAP40 antibody

Prior to the commencement of any new experiments, a new antibody should be produced and properly validated. It would be advantageous to produce an antibody that can recognize both human and mouse forms of HAP40. There is approximately 85% sequence homology between human and mouse HAP40 and, therefore, it should be possible to produce an antibody that can identify both HAP40 species. In addition, if an antibody that could detect human HAP40 was produced, experiments could be conducted on human cells. It would be helpful to examine the HAP40 levels in human brain samples from HD patients and normal controls. Currently, HD and control fibroblast cells are available in the lab and the levels of HAP40 in these cells could be quantified in the same way they were quantified in ST*Hdh* cell lines. This would allow us to determine if HAP40 levels are up regulated in human HD cells.

It is likely that the 45kDa band identified by the antibody is HAP40 since it is the correct molecular weight and we know that the antibody can identify recombinant GST-HAP40; however some uncertainty remains. No experiments could definitively conclude that this band was full length HAP40. Mass spectrometry of the isolated band would provide us with a definitive conclusion. Immunoprecipitation (IP) prior to mass spectrometry would ensure that the protein that is identified by the antibody is the only protein in the sample.

The only human source of HAP40 that this antibody identified was in HeLa cells; however, this may not be HAP40. HeLa cells have been shown previously to contain viral insertions of HPV[91]. The capsid coat of HPV type 147 has a sequence that is similar to the mouse HAP40 epitope. It is more similar to the mouse HAP40 epitope than the human HAP40 epitope is to the mouse epitope. It also has a similar mass to the band shown on the western blot. Using IP, the protein can be isolated from the HeLa cell extract and run on a western blot. The resulting band could then be excised and sent for mass spectrometry. This will definitively identify the protein in the band.

Transfection of siRNA to reduce HAP40 levels was shown to have no effect on the HAP40 protein levels identified by this antibody. It is possible that the siRNA is not effective at reducing HAP40 levels. Reverse transcription PCR can quantify the HAP40 mRNA levels in siRNA transfected cells and compare them to the control cells. This would determine the mRNA levels in the cells, revealing any differences that may be a result of the siRNA.

Endogenous HAP40 levels

Pal *et al* (2006) analyzed the HAP40 levels in ST*Hdh* cells using the same antibody that we used here and found that ST$Hdh^{Q111/Q111}$ cells contain the most HAP40 protein of the three cell lines. Our results identified ST$Hdh^{Q111/Q111}$ as the cell line with the least amount of HAP40 protein. Our lab has recently ordered new ST*Hdh* cell lines. When these arrive it would be interesting to run a western blot using the new cell lysates to determine if HAP40 levels are the same in the new cells as they are in the old cells.

Overexpression of HAP40 and toxicity

The toxicity associated with overexpressed HAP40 still requires further investigation. The cytochrome c immunofluorescence apoptosis assay should be revisited with the addition of a mitochondrial stain that identifies healthy mitochondria. Co-localization of cytochrome c and the mitochondrial marker would indicate that the cell is healthy. This will help to distinguish cells that are healthy from those that are dying where co-localization between cytochrome c and the mitochondrial marker would not occur. The flow cytometry experiments can also be continued since the stress/contamination problems with the cells have been resolved.

Investigation into the function of the IBM consensus sequence at the amino terminus of the HAP40 protein could answer why HAP40-YFP will not express in wildtype cells. If the IBM is functional and can bind to the BIRII domains on IAPs, apoptosis is no longer inhibited. Currently, it is unknown if this motif is functional in HAP40. If it is functional, this would explain why HAP40-YFP expression is low. The cells could be down regulating the expression of the HAP40-YFP construct or the cells expressing a large amount of this protein would die. To determine if this motif is

functional, site-directed mutagenesis can be used to mutate the required amino terminal alanine residue. If there is no alanine at the amino terminus, the consensus will not be able to interact with the BIRII domain binding pocket and no inhibition of the IAP will occur. This allows the IAP to bind and inhibit caspases. Expression of the mutated HAP40-YFP can be quantified to determine if there is a difference between the mutated and non-mutated constructs. Additionally, the mutated construct can be used in the experiments described above to determine if it affects cell death.

The huntingtin-HAP40 interaction

The huntingtin-HAP40 interaction began with analysis of the carboxyl terminus of huntingtin; however, an interaction was found with N17 at the amino terminus. This interaction must be further characterized in live cells. The methods used here to demonstrate this interaction were biochemical methods using cell lysates, recombinant proteins, and synthetic peptides in artificial systems. Additionally, the impact that phosphorylation of serine residues 13 and 16 have on the interaction should be explored in the cell. Phosphomimetic constructs of N17 can be transfected into the cell and, using the GFP trap beads, the interaction can be investigated. The experiments analyzing the effect of cell stress on the interaction between N17 and HAP40 must be repeated.

Investigation of the carboxyl terminus of huntingtin should continue as well. A fusion protein of huntingtin amino acids 2226 to 3144 should be created and tested for interaction with HAP40. This construct covers the full length of huntingtin that was demonstrated to interact with HAP40 in previous studies[79]. Co-localization between the huntingtin fragment and HAP40 should be analyzed followed by GFP trap experiments to determine if there is an interaction.

Huntingtin is a flexible protein that is capable of moving around and changing conformation due to the HEAT repeat motifs. This leaves the potential for many transient interactions to occur when huntingtin is in a specific conformation. Affinity chromatography experiments using recombinant proteins of the carboxyl terminus of huntingtin can be coupled to a column. Cell lysates can then be run through the column and the proteins bound to the carboxyl terminal huntingtin fragments can be eluted from

the column and sent for mass spectrometry. A list will be produced of proteins that interact with the constructs and the list can be investigated. This could potentially identify other interacting proteins at the carboxyl terminus.

HAP40 Cleavage products

Producing a pair of antibodies, one that can identify the amino terminus of HAP40 and one that can detect the carboxyl terminus of HAP40 would help to identify any cleavage products that occur at endogenous HAP40 levels. Alternatively, cells can be transfected with the venus-HAP40-mCh construct. This would analyze cleavage that occurs in the disease state when HAP40 protein levels are elevated. The cell lysates from the transfection can be run on a western blot and the blot probed with antibodies that recognize venus or mCherry to identify the location of each fluorophore on the western blot. Additionally, when venus-HAP40-mCherry is transfected into cells, a video watching what happens to each fragment of HAP40 and the cell after HAP40 is cleaved could unveil a role of HAP40 in the cell.

In conclusion, the exact location of the interaction between carboxyl terminal huntingtin and HAP40 remains elusive; however, we were able to identify an interaction that occurs between HAP40 and the first 17 amino acids at the amino terminus of huntingtin. This interaction only occurs when HAP40 levels are elevated, similar to the carboxyl terminal huntingtin and HAP40 interaction previously published[79]. In the nucleus, HAP40 nuclear spots depend on the structure of chromatin, similar to phosphorylated huntingtin nuclear puncta. We also demonstrated that overexpressed HAP40 can be cleaved in the cell; however, details of when this cleavage occurs or what happens to the cell following HAP40 cleavage are unknown. Investigation of the toxicity associated with increased levels of HAP40 was inconclusive and further experiments are required to determine if increased levels of HAP40 protein causes cell death.

Chapter 5 Figures

Figure 1

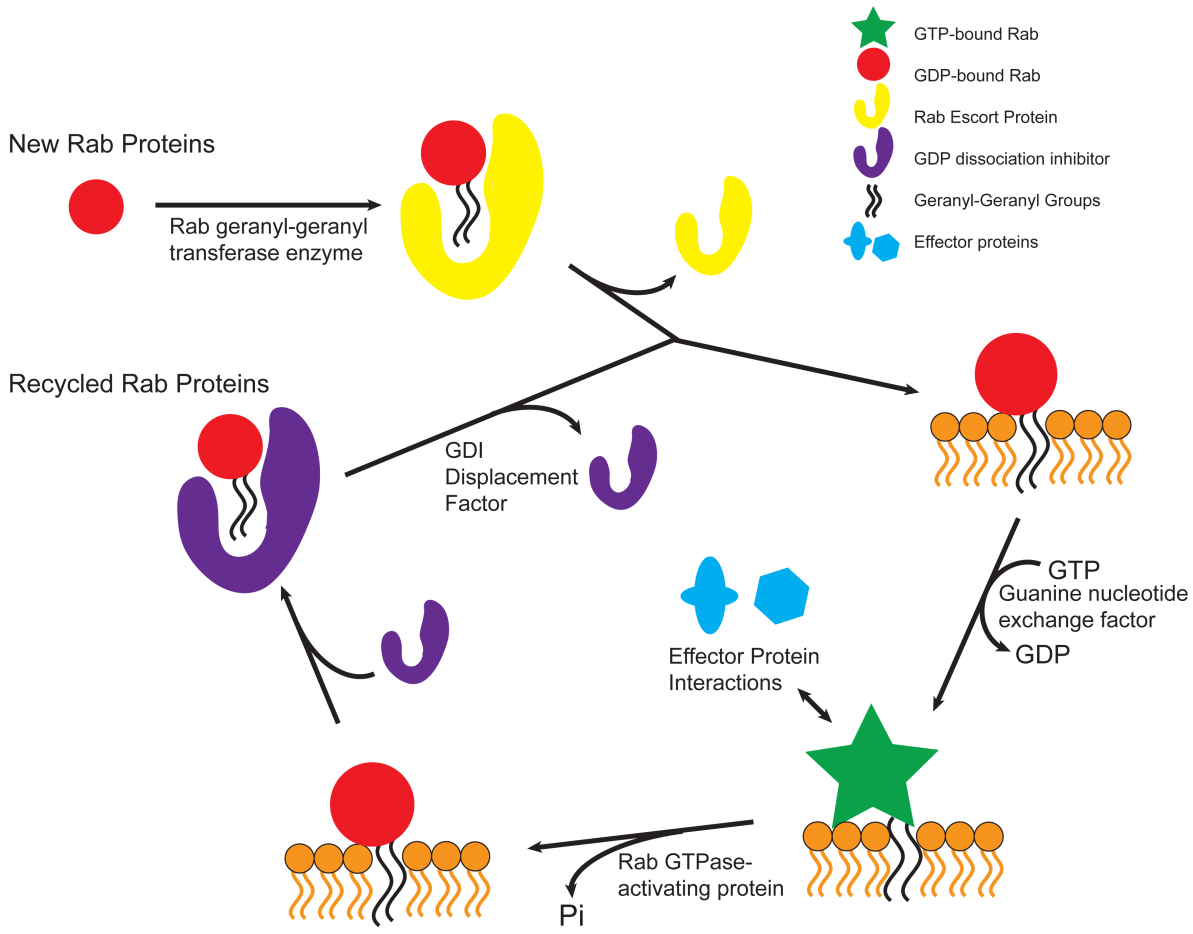

GTP-bound Rab
GDP-bound Rab
Rab Escort Protein
GDP dissociation inhibitor
Geranyl-Geranyl Groups
Effector proteins
New Rab Proteins
Rab geranyl-geranyl transferase enzyme
Recycled Rab Proteins
GDI Displacement Factor
GTP Guanine nucleotide exchange factor
GDP
Effector Protein Interactions
Rab GTPase-activating protein
Pi

FIGURE 1

The Rab cycle. New Rab proteins are synthesized in the cytosol and recognized by a Rab escort protein (REP). The REP transports the new Rab to a Rab geranyl-geranyl transferase enzyme that adds one or two lipid groups to the new Rab protein, which allows for stable insertion into a targeted membrane[107, 108]. The REP then delivers the Rab protein to the target membrane for insertion[51, 107]. Guanine nucleotide exchange factors (GEF) convert the Rab from its inactive GDP-bound state to its active GTP-bound state. This switch from GDP to GTP changes the conformation of the Rab protein, allowing it to interact with the effector molecules that carry out the tasks that each Rab is responsible for such as sorting, transport and fusion of vesicles[51, 109]. When the Rab protein's functions are complete, the Rab protein is deactivated by hydrolysis of GTP to GDP, a reaction that is catalyzed by a Rab GTPase-activating protein (GAP)[50, 51]. The Rab protein is then removed from the membrane by a GDP dissociation inhibitor (GDI), which binds to the geranyl-geranyl moiety allowing the Rab to remain in the cytoplasm to be recycled[51, 109, 110]. When the Rab is to be inserted into the membrane once again, a GDI displacement factor (GDF) displaces the GDI for insertion of the Rab protein back into a membrane[50].

Figure 2

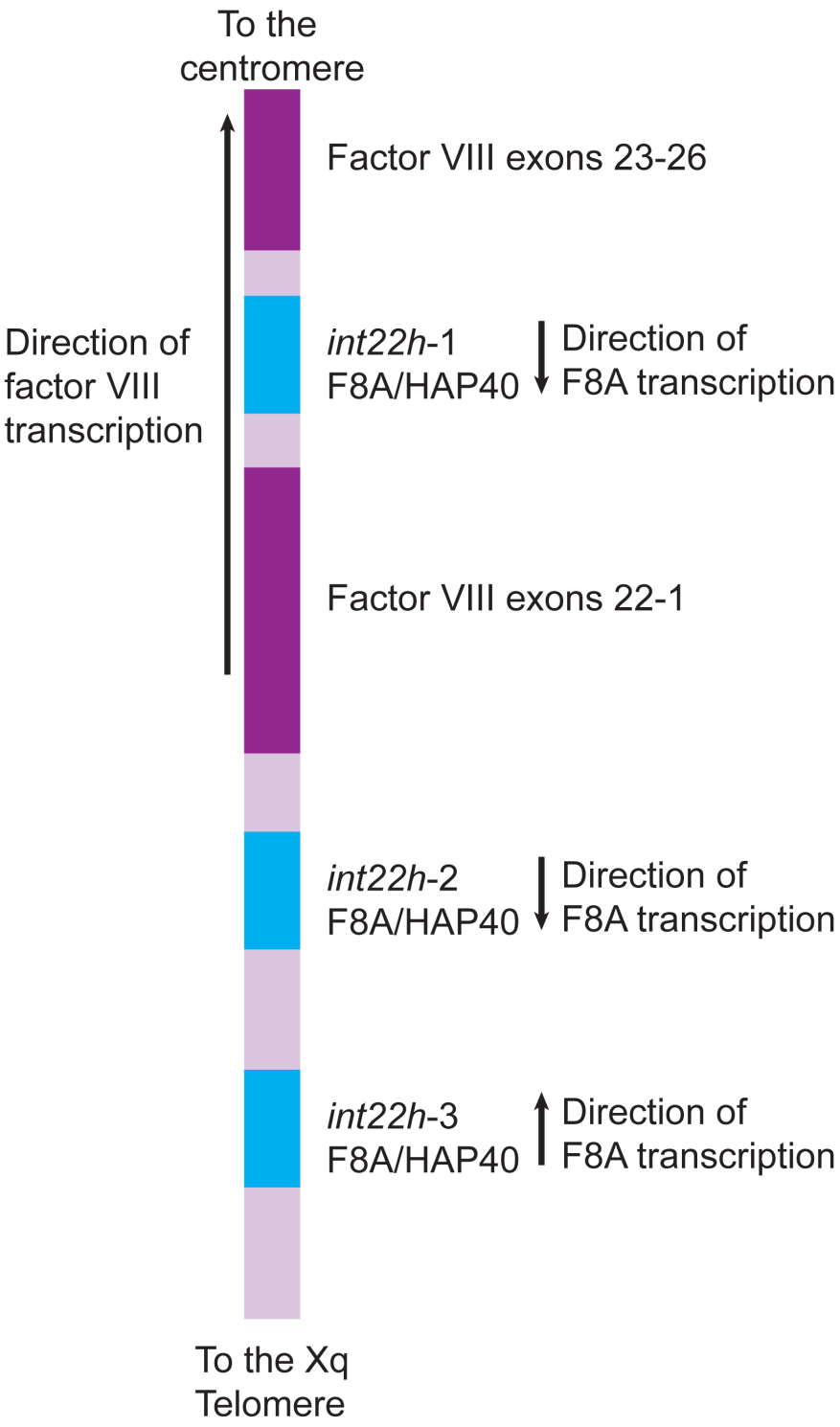

To the centromere
Factor VIII exons 23-26
Direction of factor VIII transcription
int22h-1 F8A/HAP40
Direction of F8A transcription
Factor VIII exons 22-1
int22h-2 F8A/HAP40
Direction of F8A transcription
int22h-3 F8A/HAP40
Direction of F8A transcription
To the Xq Telomere

FIGURE 2

The location of the HAP40 gene on chromosome X. In humans there are 3 copies of the *int22h* region. One is located within intron 22 of the Coagulation Factor VIII gene. The other 2 regions are located closer to the Xq telomere, outside of the Factor VIII gene. All 3 *int22h* homology regions contain a copy of the HAP40 gene. Interchromosomal homologous recombination between *int22h*-1 and *int22h*-2 can result in deletions or duplications of the intervening sequence. Recombination events between *int22h*-1 and *int22h*-3 can result in inversion of the intervening sequence. This figure was modified from figure 1 in DeBrasi *et al* (2008)[88].

Figure 3

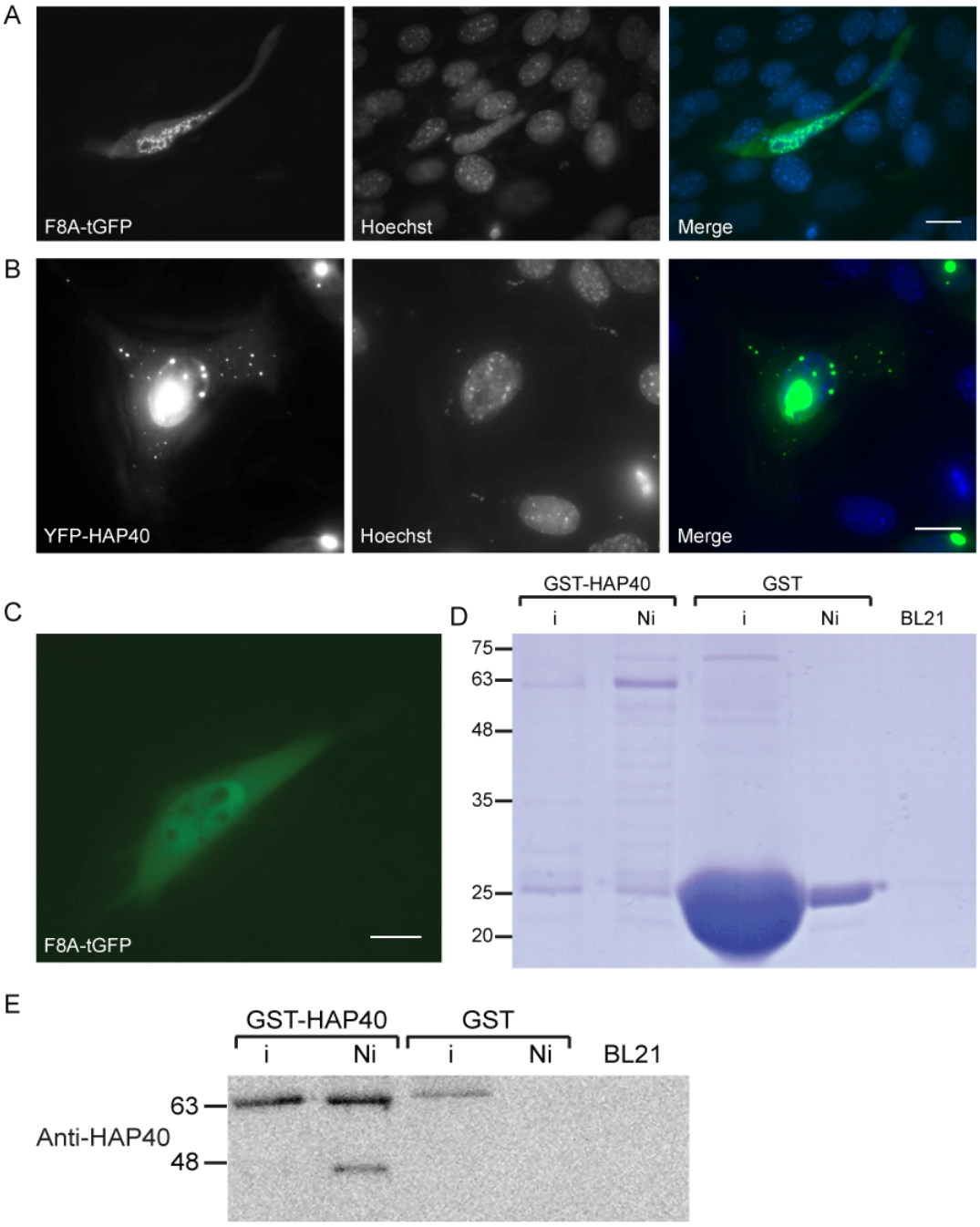

A
F8A-tGFP
Hoechst
Merge
B
YFP-HAP40
Hoechst
Merge
C
F8A-tGFP
D
GST-HAP40
GST
i
Ni
i
Ni
BL21
75
63
48
35
25
20
E
GST-HAP40
GST
i
Ni
i
Ni
BL21
63
Anti-HAP40
48

FIGURE 3

The verification of HAP40 fusion constructs. (A) Transfection of the F8A-tGFP construct demonstrating diffuse localization in the cytoplasm and punctate localization in the nucleus. A hoechst counterstain was used to identify the nucleus. **(B)** Transfection of YFP-HAP40 demonstrating a punctate localization to both the cytoplasm and the nucleus. A hoechst counterstain was used to identify the nucleus. **(C)** A live cell transfection of F8A-tGFP demonstrating a diffuse localization in both the nucleus and the cytoplasm. **(D)** A Coomassie stained SDS-PAGE gel showing the 66kDa GST-HAP40 recombinant protein in the lane not induced (Ni) with IPTG. Levels of GST-HAP40 were low in cells that were induced (i) with IPTG. Induced (i) and not induced (Ni) GST samples were used as a control. The cell lysate of BL21 *Escherichia coli* cultures, which contained no vector, was also used as a control. **(E)** A western blot of the recombinant proteins GST-HAP40, GST and BL21 cell lysate control. The blot was probed with the HAP40 antibody, and the antibody was able to identify GST-HAP40 at 66kDa in the induced (i) and not induced (ni) lanes.

All fusion constructs were transfected into ST*Hdh*$^{Q7/Q7}$ cell lines. Scale bars represent 10μm. Image brightness was adjusted for part C in Adobe Photoshop.

Figure 4

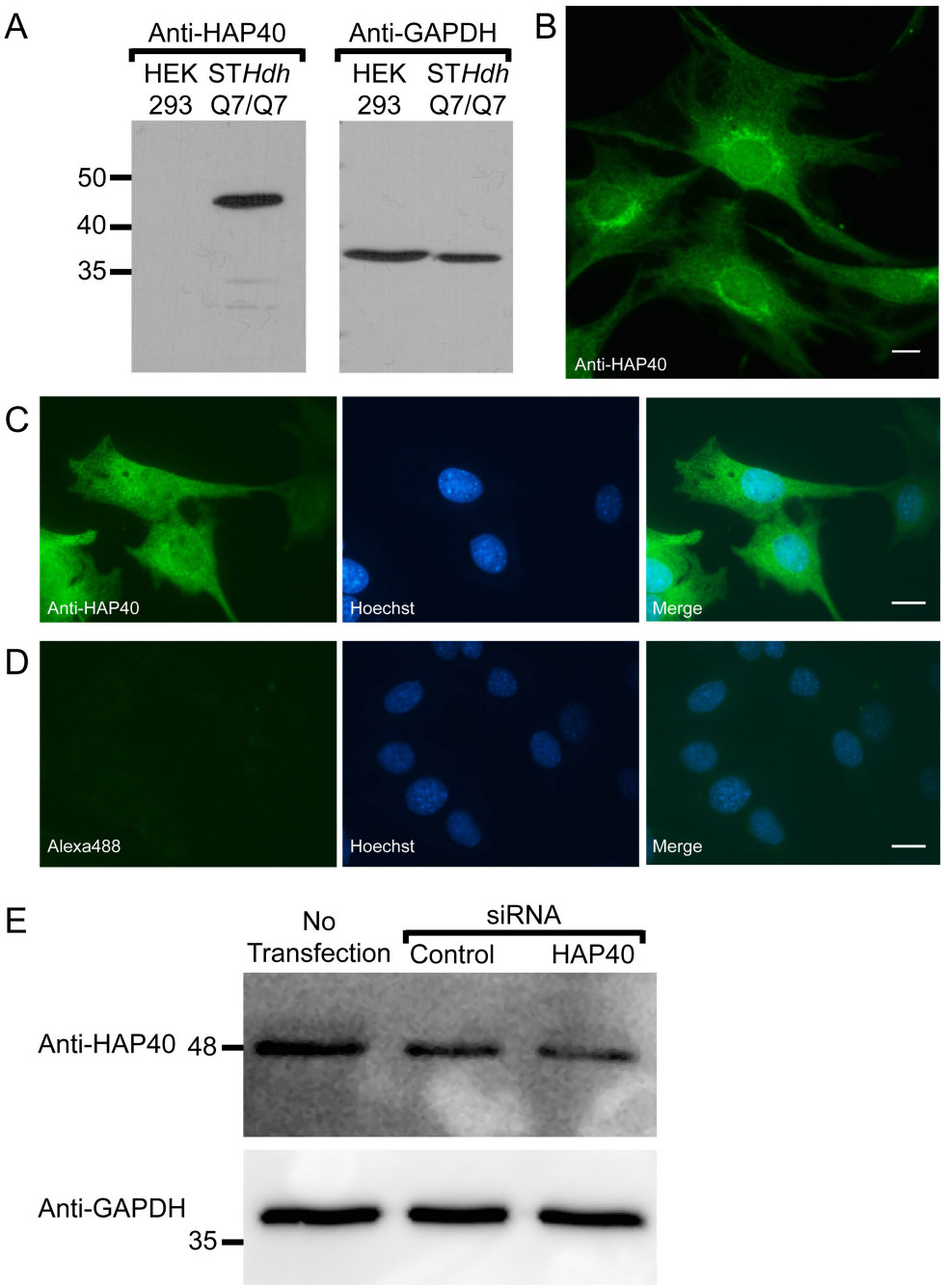

A
Anti-HAP40
Anti-GAPDH
HEK 293
STHdh Q7/Q7
50
40
35
B
Anti-HAP40
C
Anti-HAP40
Hoechst
Merge
D
Alexa488
Hoechst
Merge
E
No Transfection
siRNA
Control
HAP40
Anti-HAP40 48
Anti-GAPDH
35

FIGURE 4

The verification of the HAP40 antibody using endogenous protein methods. (A) A western blot demonstrating that the HAP40 antibody identified mouse HAP40 at approximately 45kDa, but does not identify human HAP40 in HEK293 cells. The GAPDH loading control confirms that an equal amount of sample was loaded into each lane. **(B)** Immunofluorescence using a high concentration of HAP40 antibody reveals dim nuclear puncta. These puncta can also be seen when HAP40 is overexpressed. **(C)** Immunofluorescence demonstrating the typical staining pattern of the HAP40 antibody. Strong cytoplasmic staining and weak nuclear staining occurs. A hoechst counterstain was used to identify the nucleus. **(D)** A control with no primary antibody confirms that the staining pattern seen in parts A and B are as a result of the HAP40 antibody and not non-specific staining from the secondary antibody. A hoechst counterstain was used to identify the nucleus. **(E)** A western blot of ST*Hdh*$^{Q7/Q7}$ cell lysates from cells not transfected (lane 1), transfected with a siRNA scrambled control (lane 2) or transfected with a siRNA to reduce HAP40 protein levels (lane 3). The HAP40 antibody does not demonstrate decreased levels of HAP40 protein in the sample transfected with the siRNA against HAP40 compared to the controls. The GAPDH loading control demonstrates that an equal amount of protein was loaded into each well.

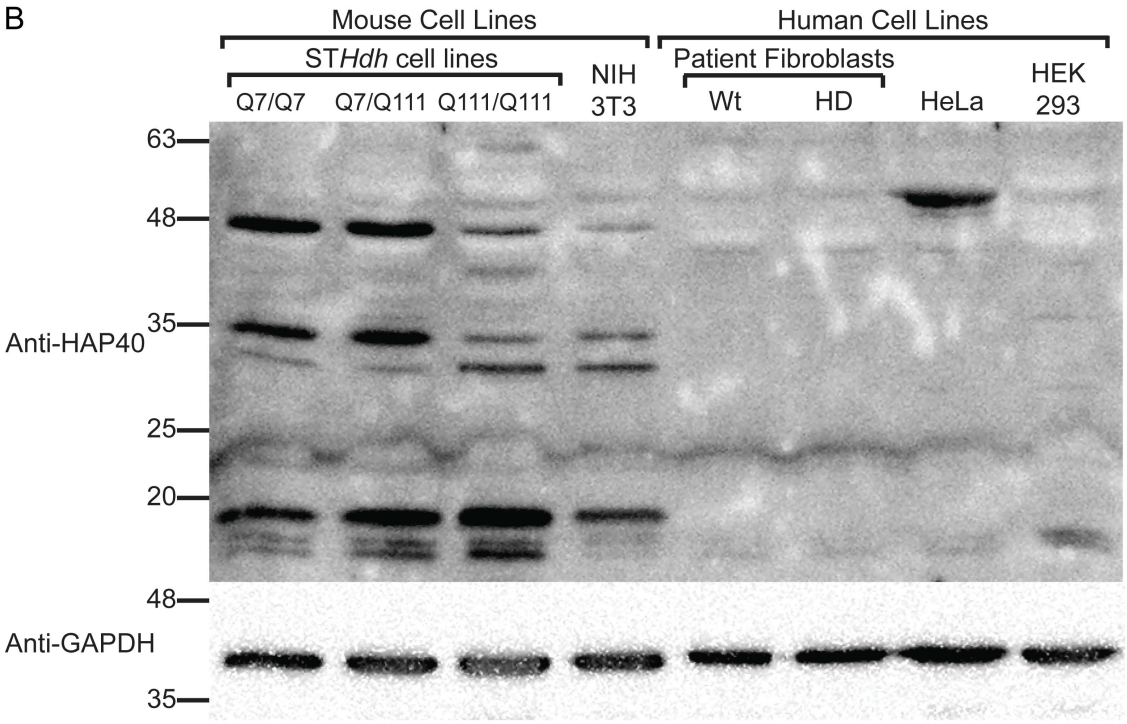
All immunofluorescence experiments were conducted on ST*Hdh*$^{Q7/Q7}$ cell lines. Scale bars represent 10μm. Image brightness was adjusted for part B, C and D in Adobe Photoshop.

Figure 5

A HAP40 Epitope Alignment

Human HAP40:
D S H G G E S S G Q L P
Mouse HAP40:
D G H G Q D T S G Q L P

B



C BLAST Alignment Query - Epitope Sequence DGHGQDTSGQLP

Major Capsid Protein [HPV type 147] - 521 amino acids

Query -	D	G	-	H	G	Q	D	-	-	-	-	-	-	-	T	S	G	Q	L	P
HPV -	D	G	T	N	N	Q	D	L	Y	Y	F	S	P	A	T	S	G	Q	L	P

FIGURE 5

The HAP40 antibody recognizes mouse HAP40 but does not identify human HAP40. **(A)** The human and mouse HAP40 epitope sequences are outlined where the differences between the two epitopes are shown in red. There is only one area of 5 consecutive amino acids where these epitopes align, suggesting that the human epitope is not recognizable by the antibody. **(B)** A panel of human and mouse cell extracts were analyzed by western blot. The HAP40 antibody recognizes a band at approximately 45kDa in all mouse cell lysates (lanes 1-4). No bands are present in the human cell lysates (lanes 5, 6 and 8); however, there is a band present in the lane corresponding to the HeLa cell lysate (lane 7). The HAP40 antibody identified a protein in the HeLa cell lysate at approximately 55kDa, which may or may not be HAP40. The GAPDH control demonstrates that equal amounts of protein were loaded into each lane. **(C)** A BLAST sequence alignment of the HAP40 antibody epitope sequence and the major capsid protein of the human papillomavirus type 147 sequence. This alignment demonstrates that a sequence in the capsid protein of HPV type 147 is more similar to the mouse HAP40 sequence then the human HAP40 sequence is to the mouse sequence as seen in part a. Identical amino acids are outlined in green boxes.

Figure 6

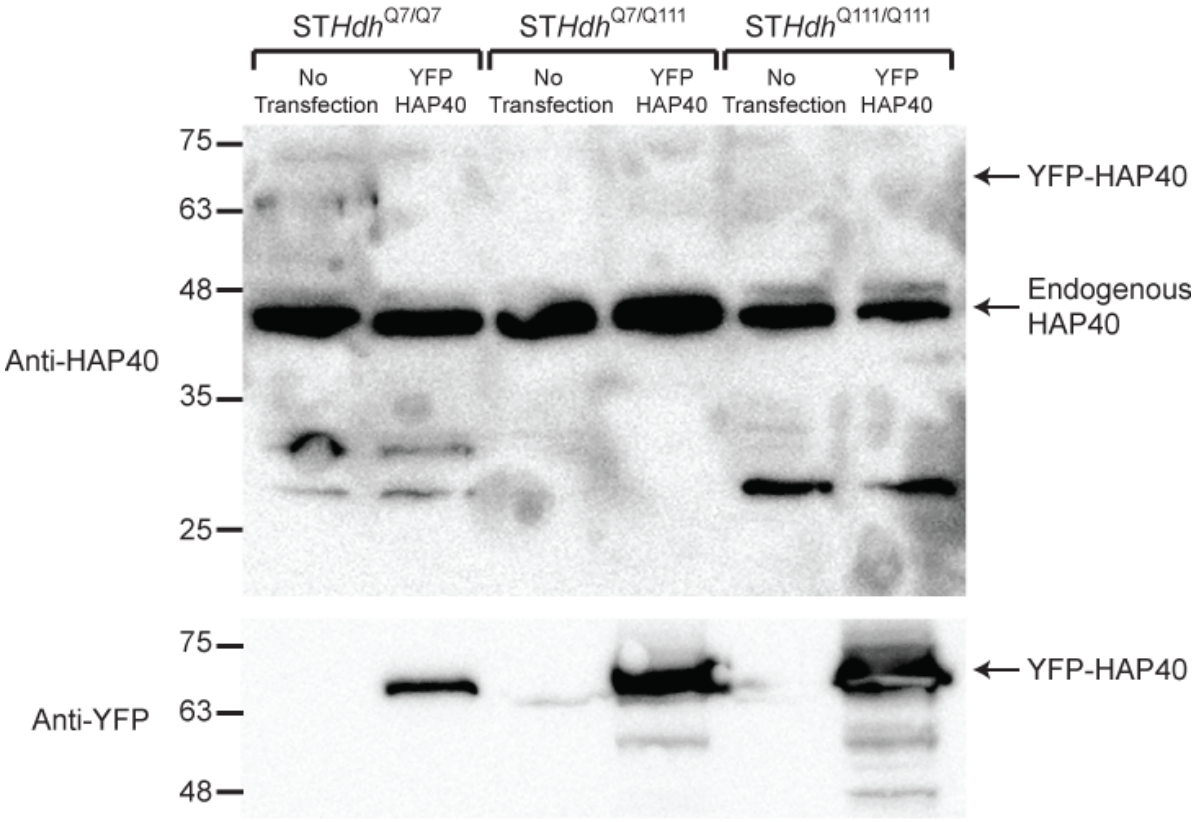

STHdh Q7/Q7
STHdh Q7/Q111
STHdh Q111/Q111
No Transfection
YFP HAP40
No Transfection
YFP HAP40
No Transfection
YFP HAP40
75
63
48
35
25
Anti-HAP40
YFP-HAP40
Endogenous HAP40
75
63
48
Anti-YFP
YFP-HAP40

FIGURE 6

The HAP40 antibody does not recognize YFP-HAP40 on a western blot. A Western blot of cell lysates of each of the ST*Hdh* cell lines transfected with YFP-HAP40 reveals that the HAP40 antibody does not see the HAP40 fusion protein. The YFP-HAP40 fusion protein has a molecular weight of approximately 67kDa and no signal is present on the blot probed with the HAP40 antibody at this molecular weight. On the western blot probed with the YFP antibody, the fusion proteins are identified at approximately 67kDa indicating that the fusion protein is present in the samples.

Figure 7

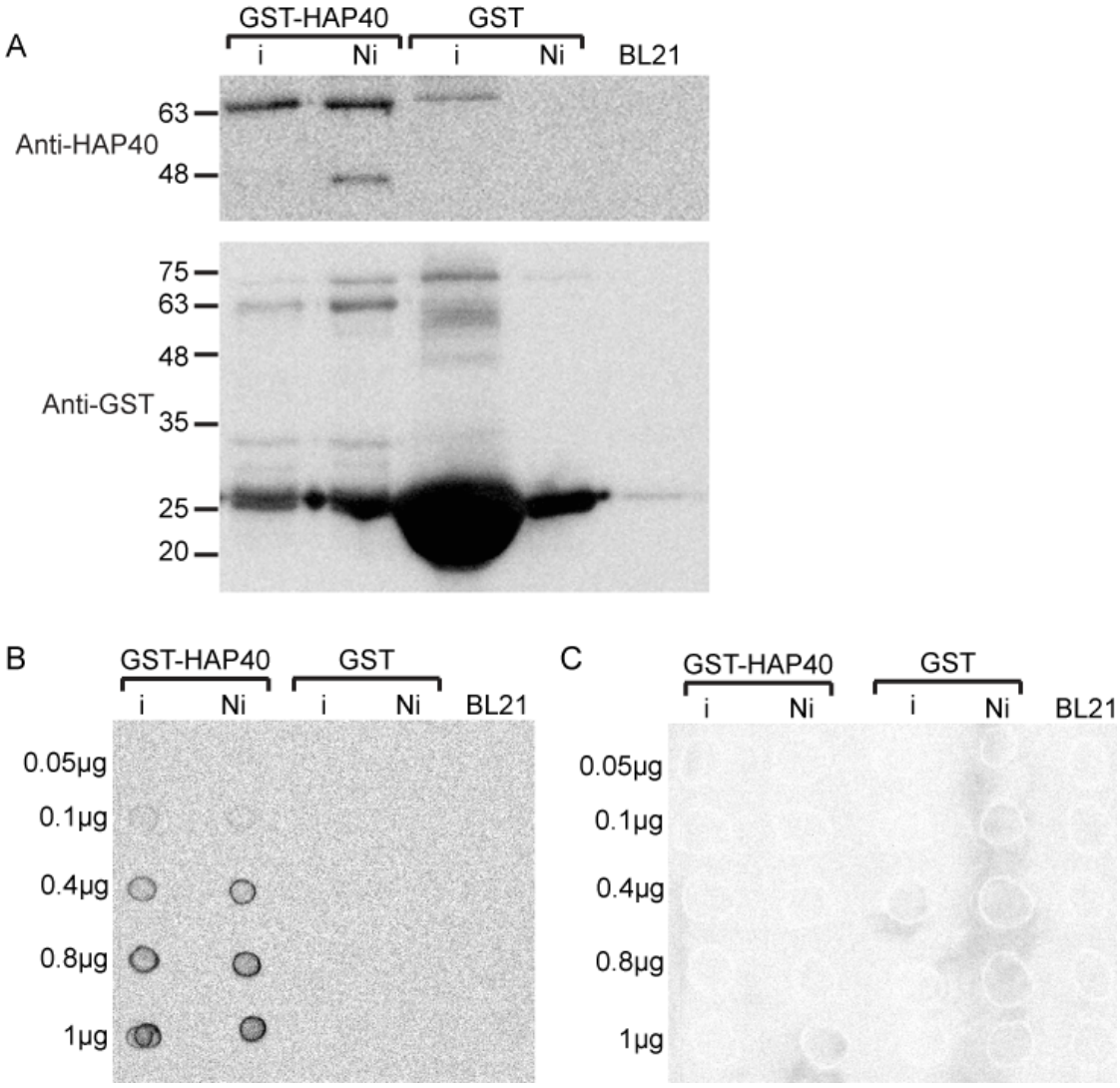

A
GST-HAP40
GST
i
Ni
i
Ni
BL21
63
Anti-HAP40
48
75
63
48
Anti-GST
35
25
20
B
GST-HAP40
GST
i
Ni
i
Ni
BL21
0.05µg
0.1µg
0.4µg
0.8µg
1µg
C
GST-HAP40
GST
i
Ni
i
Ni
BL21
0.05µg
0.1µg
0.4µg
0.8µg
1µg

FIGURE 7

The HAP40 antibody recognizes the GST recombinant protein produced in BL21 *Escherichia coli* cells. (A) A western blot of recombinant protein from the soluble fraction of cultures that were induced (i) or not induced (Ni) with IPTG. The HAP40 antibody could identify the GST-HAP40 fusion protein present in these samples. The samples in this western blot were not equally loaded. **(B)** A Dot blot with an equal loading of protein demonstrated that the HAP40 antibody can see HAP40 and does not see any background staining in the GST sample columns as seen in part a. **(C)** An equally loaded dot blot where primary antibody was not applied to demonstrate that the secondary antibody did not produce non-specific interactions. This blot demonstrates that signal in parts a and b are not a result of non-specific secondary antibody binding.

Figure 8

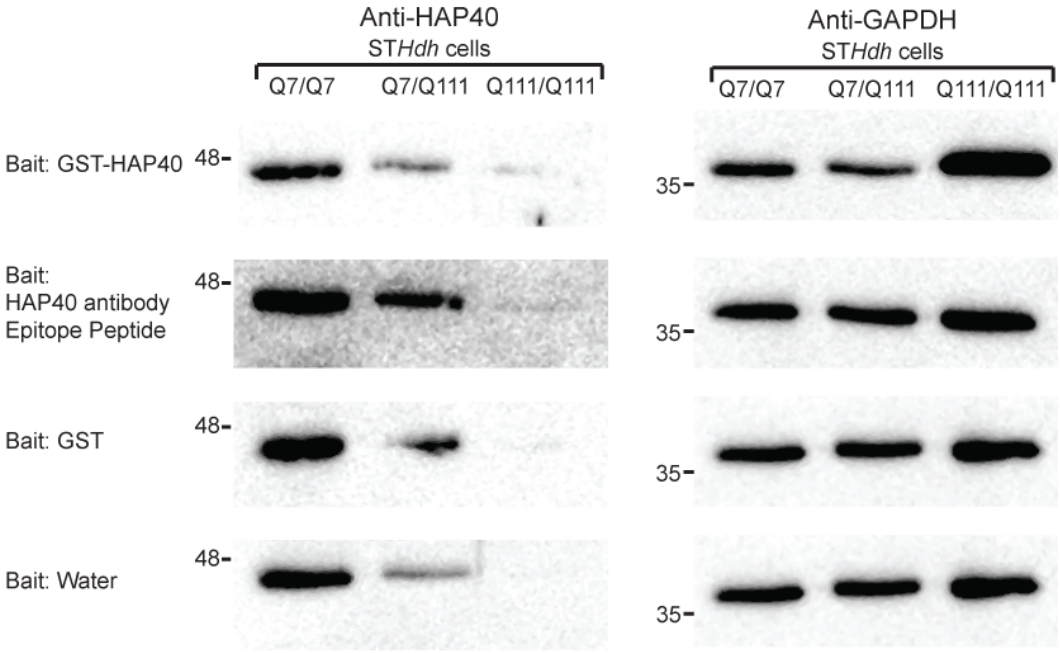

Anti-HAP40
STHdh cells
Q7/Q7
Q7/Q111
Q111/Q111
Anti-GAPDH
STHdh cells
Q7/Q7
Q7/Q111
Q111/Q111
Bait: GST-HAP40
48-
35-
Bait:
HAP40 antibody
Epitope Peptide
48-
35-
Bait: GST
48-
35-
Bait: Water
48-
35-

FIGURE 8

Incubation with HAP40 peptides does not prevent interaction of the antibody with protein on the blot. The cell lysates from each of the ST*Hdh* cell lines were run on a western blot and we probed with an antibody solution that was pre-incubated with a bait peptide. The bait peptides are shown to the left of the western blots. None of the bait proteins were able to reduce binding of the antibody to the protein on the membrane. The GAPDH loading control demonstrates that most samples were equally loaded.

Figure 9

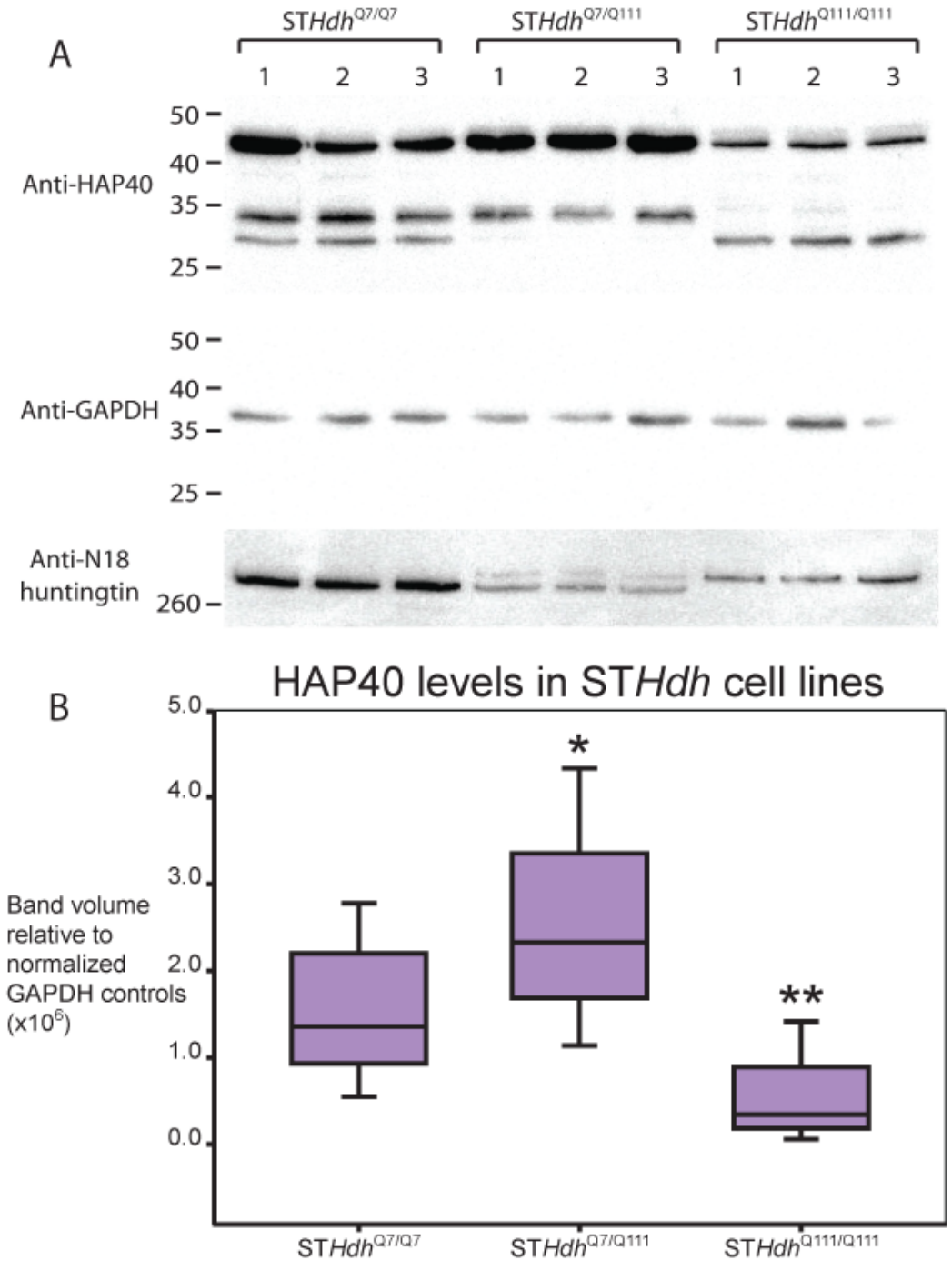

A
STHdhQ7/Q7
STHdhQ7/Q111
STHdhQ111/Q111
1 2 3 1 2 3 1 2 3
Anti-HAP40
50
40
35
25
Anti-GAPDH
50
40
35
25
Anti-N18 huntingtin
260
B
HAP40 levels in STHdh cell lines
Band volume relative to normalized GAPDH controls (x10^6)
5.0
4.0
3.0
2.0
1.0
0.0
*
**
STHdhQ7/Q7
STHdhQ7/Q111
STHdhQ111/Q111

FIGURE 9

HAP40 levels are highest in ST*Hdh*$^{Q7/Q111}$ cells and lowest in ST*Hdh*$^{Q111/Q111}$ cells. (A) A western blot showing the differences in the levels of HAP40 in each of the ST*Hdh* cell lines. HAP40 protein levels are low in ST*Hdh*$^{Q111/Q111}$ cell lysates and high in ST*Hdh*$^{Q7/Q111}$ cell lysates when compared to ST*Hdh*$^{Q7/Q7}$ cells. GAPDH was used to confirm equal loading and an antibody that recognizes the N17 region of huntingtin was used to verify the identity of the cell lines. Huntingtin in ST*Hdh*$^{Q7/Q7}$ cells has a lower molecular weight than the mutant huntingtin in ST*Hdh*$^{Q111/Q111}$ cells. Both mutant and normal huntingtin is expressed in ST*Hdh*$^{Q7/Q111}$ cells and therefore, 2 bands are present, one with a lower molecular weight representing normal huntingtin and one with a higher molecular weight representing mutant huntingtin. **(B)** Band volume analysis of the western blot indicates that the changes in HAP40 levels in ST*Hdh*$^{Q7/Q111}$ cells and ST*Hdh*$^{Q111/Q111}$ cells when compared to ST*Hdh*$^{Q7/Q7}$ cells are statistically significant. *p=0.042, **p=0.004

Figure 10

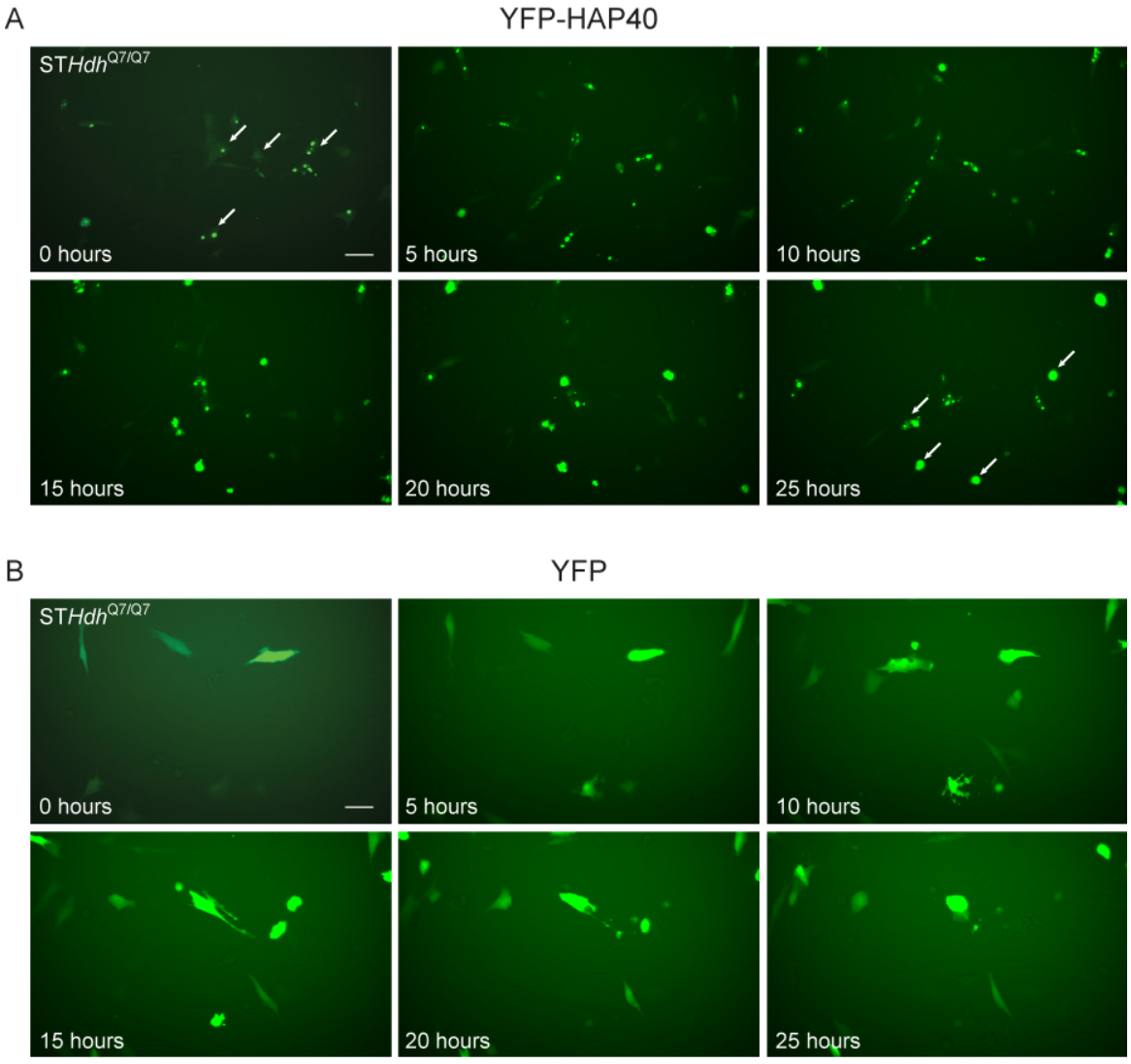

A
YFP-HAP40
STHdhQ7/Q7
0 hours
5 hours
10 hours
15 hours
20 hours
25 hours
B
YFP
STHdhQ7/Q7
0 hours
5 hours
10 hours
15 hours
20 hours
25 hours

FIGURE 10

Cells transfected with YFP-HAP40 die over a 25-hour image acquisition period. (A) Cells transfected with YFP-HAP40. Arrows in the first frame are pointing to healthy cells, which die as time progresses. Arrows in the final frame point to dead cells. **(B)** Cells transfected with an empty YFP vector. Fewer cells died in when transfected with YFP then when transfected with YFP-HAP40.

Both experiments were conducted on ST*Hdh*$^{Q7/Q7}$ cell lines. Cells were transfected 16 hours prior to the 0 hour image. Images were taken every 5 minutes for 25 hours. Scale bars represent 100μm.

Figure 11

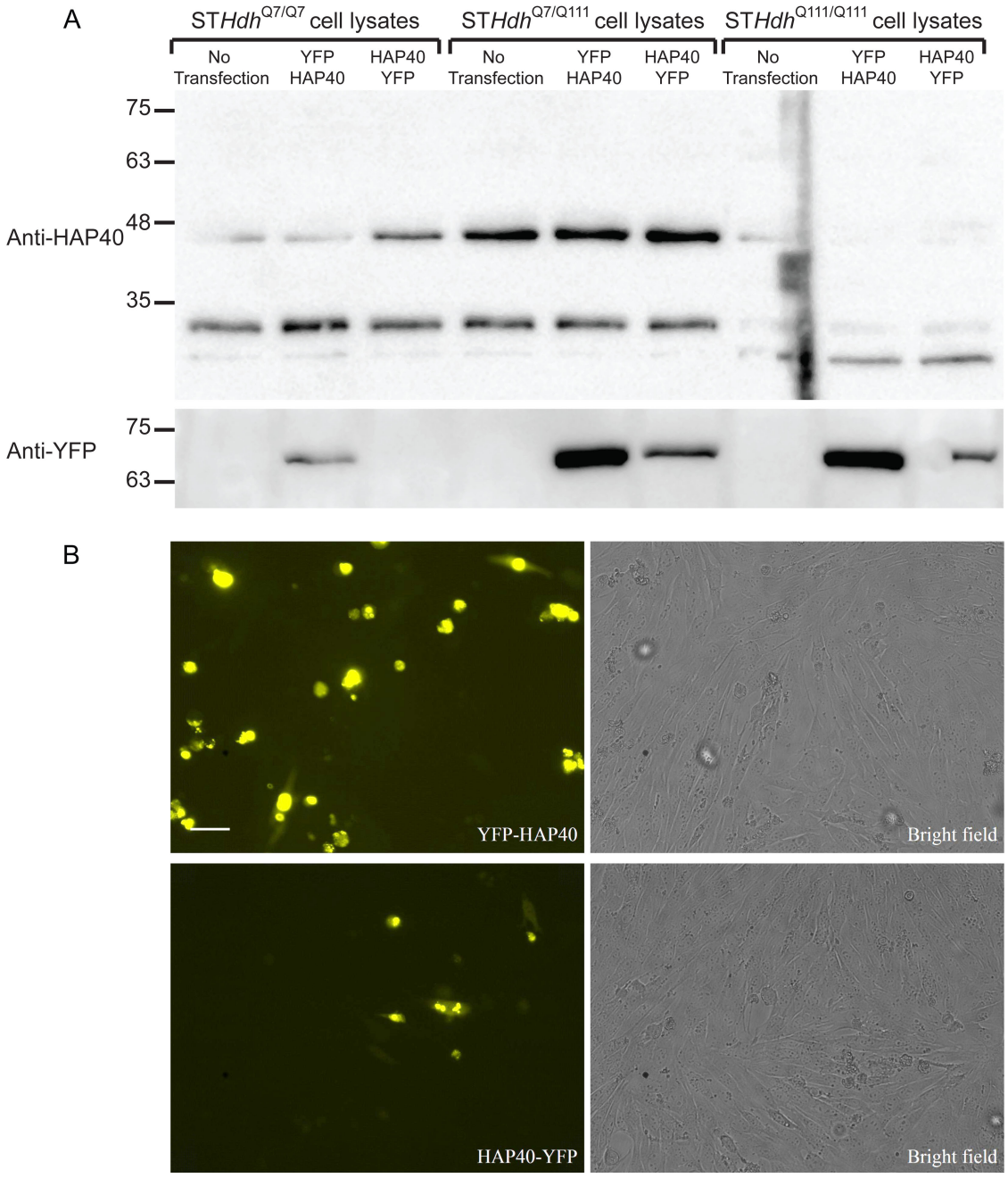

A
STHdhQ7/Q7 cell lysates
STHdhQ7/Q111 cell lysates
STHdhQ111/Q111 cell lysates
No Transfection
YFP HAP40
HAP40 YFP
No Transfection
YFP HAP40
HAP40 YFP
No Transfection
YFP HAP40
HAP40 YFP
75
63
48
35
Anti-HAP40
75
63
Anti-YFP
B
YFP-HAP40
Bright field
HAP40-YFP
Bright field

FIGURE 11

YFP-HAP40 transfects and expresses better than HAP40-YFP. (A) A western blot demonstrating the levels of YFP-HAP40 or HAP40-YFP following transfection into the ST*Hdh* cell lines. Levels of HAP40-YFP seen in the blot probed with YFP are less then the levels of YFP-HAP40. **(B)** When HAP40-YFP is transfected into ST$Hdh^{Q7/Q7}$ cells, a lower expression level of HAP40-YFP is seen. Additionally, the cells that are transfected with HAP40-YFP appear more unhealthy than those transfected with YFP-HAP40. Corresponding bright field images are to the right of the YFP channel images. Scale bar represents 100μm.

Figure 12

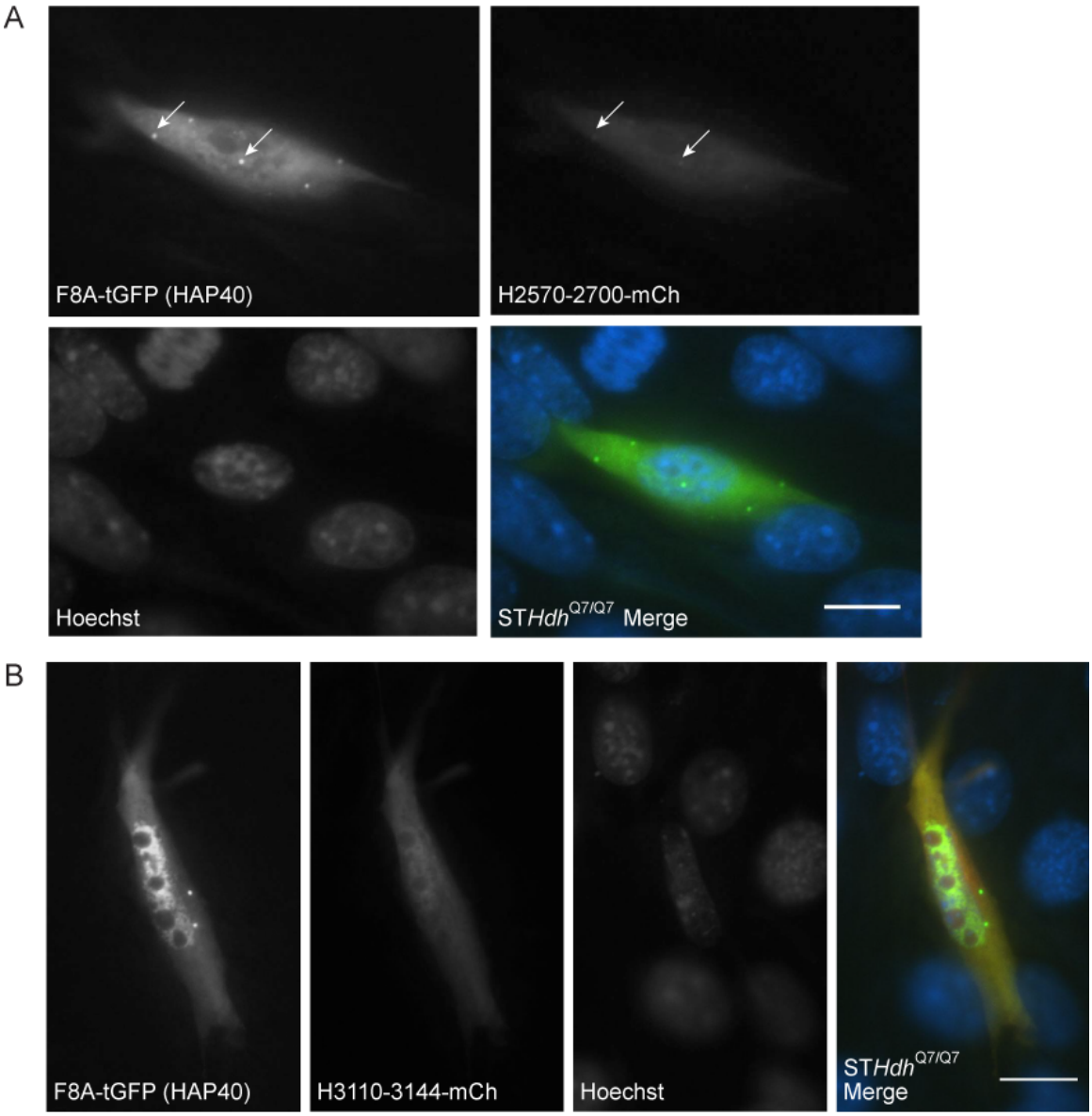

A
F8A-tGFP (HAP40)
H2570-2700-mCh
Hoechst
STHdhQ7/Q7 Merge
B
F8A-tGFP (HAP40)
H3110-3144-mCh
Hoechst
STHdhQ7/Q7
Merge

FIGURE 12

Huntingtin amino acids 2570-2700 co-localize with HAP40, but huntingtin amino acids 3110-3144 do not. (A) Huntingtin fragment H2570-2700-mCherry and F8A-tGFP were transfected into cells. Arrows show areas of co-localization between the fragment and HAP40. **(B)** Co-localization does not occur between the F8A-tGFP and huntingtin fragment H3110-3144-mCherry fusion proteins.

ST*Hdh*$^{Q7/Q7}$ cell lines were used in both experiments. Scale bar represents 10μm. Images were brightened using Adobe Photoshop.

Figure 13

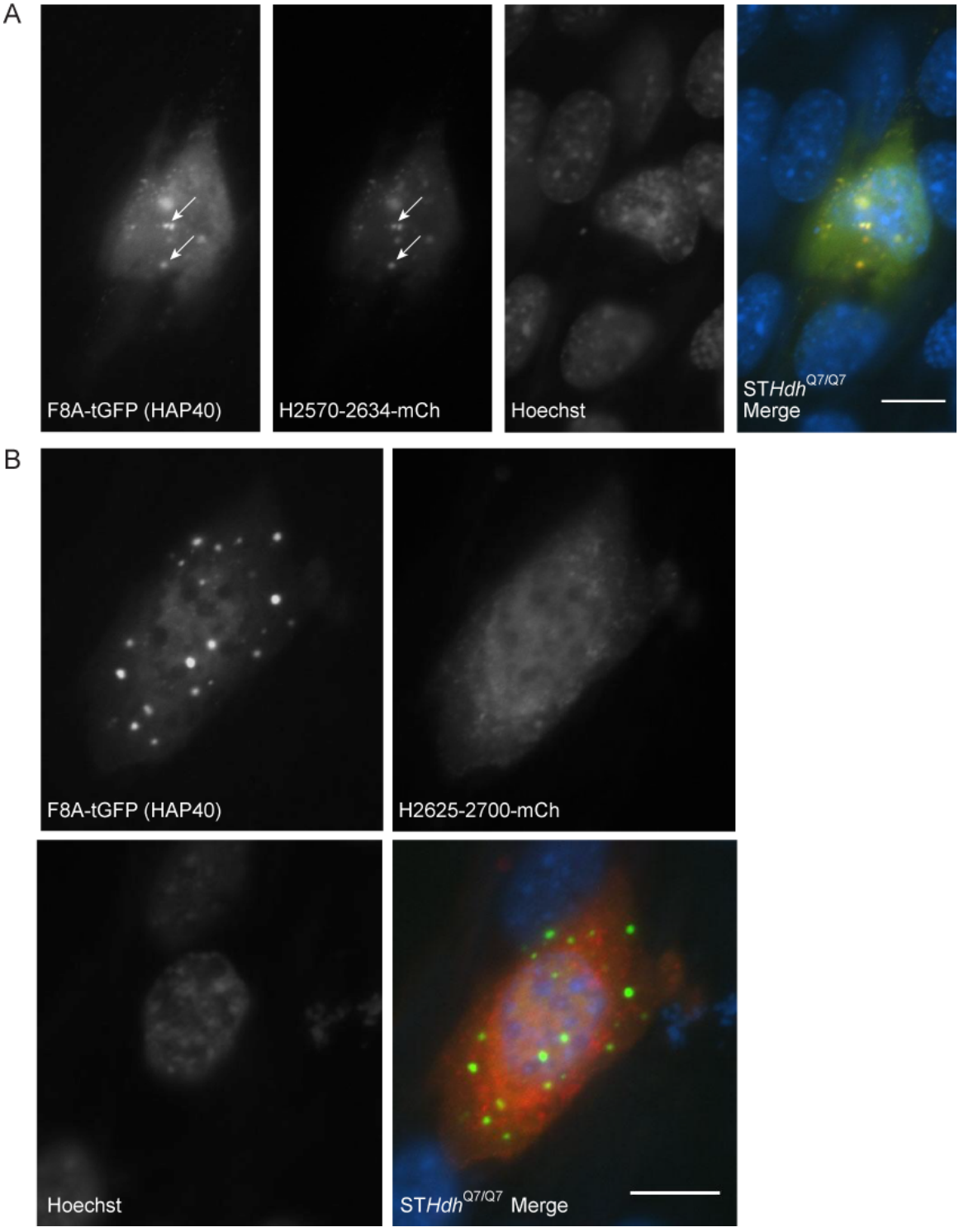

A
F8A-tGFP (HAP40)
H2570-2634-mCh
Hoechst
STHdhQ7/Q7
Merge
B
F8A-tGFP (HAP40)
H2625-2700-mCh
Hoechst
STHdhQ7/Q7 Merge

FIGURE 13

Huntingtin amino acids 2570-2634 co-localize with HAP40, but huntingtin amino acids 2635-2700 do not. (A) Huntingtin fragment H2570-2635-mCherry and F8A-tGFP co-localize in the cells where arrows show the areas of co-localization between the constructs. **(B)** Co-localization does not occur between the F8A-tGFP and huntingtin fragment 2635-2700-mCherry.

ST*Hdh*$^{Q7/Q7}$ cell lines were used in both experiments. Scale bar represents 10μm. Images were brightened using Adobe Photoshop.

Figure 14

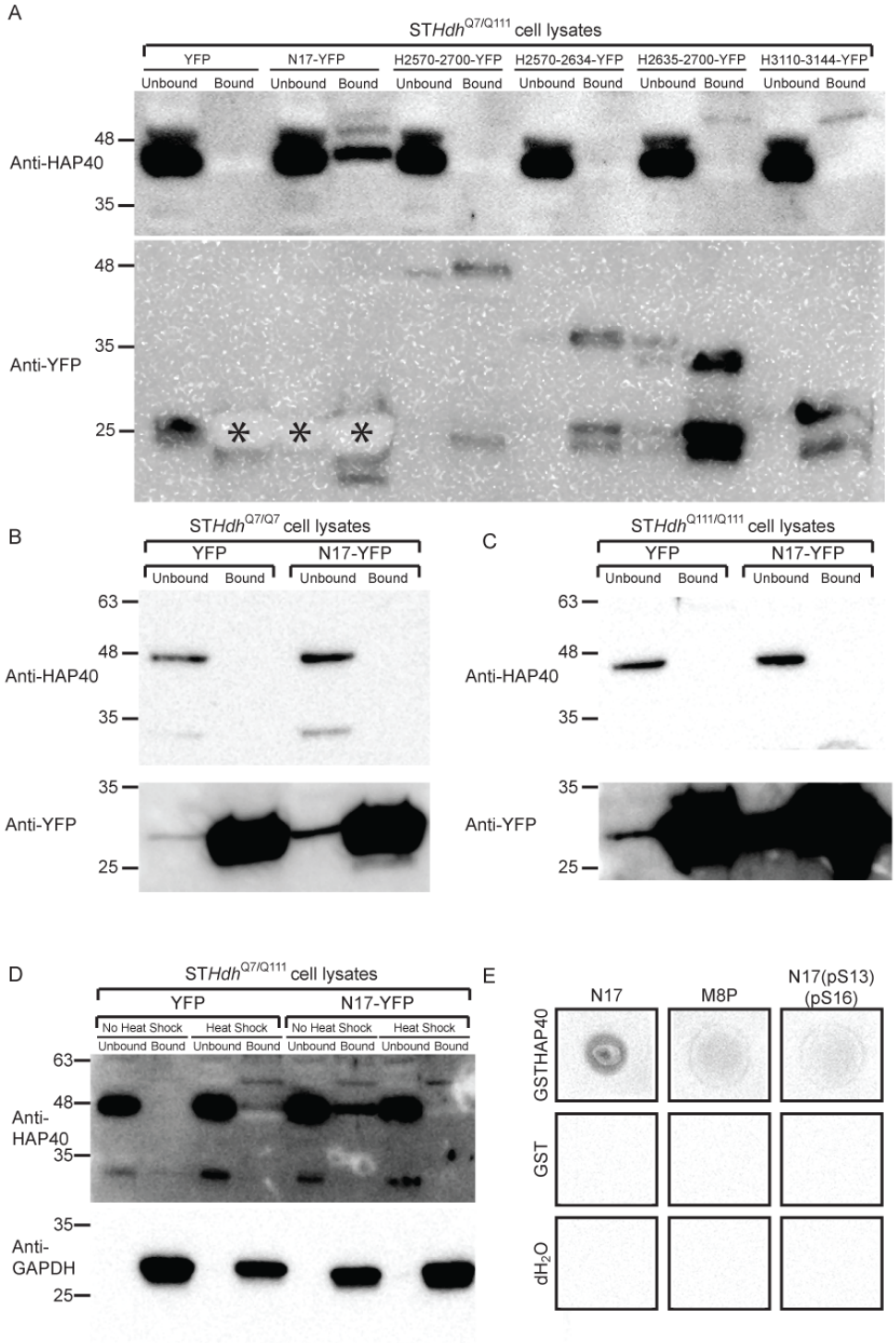

A
STHdhQ7/Q111 cell lysates
YFP
N17-YFP
H2570-2700-YFP
H2570-2634-YFP
H2635-2700-YFP
H3110-3144-YFP
Unbound Bound
Anti-HAP40
Anti-YFP
48
35
25
B
STHdhQ7/Q7 cell lysates
C
STHdhQ111/Q111 cell lysates
63
D
STHdhQ7/Q111 cell lysates
No Heat Shock
Heat Shock
Anti-HAP40
Anti-GAPDH
E
N17
M8P
N17(pS13) (pS16)
GSTHAP40
GST
dH2O

FIGURE 14

HAP40 interacts with N17 when HAP40 levels are upregulated in ST*Hdh*$^{Q7/Q111}$ cell lines. GFP trap experiments run on a western blot show both unbound and bound fractions. Bound fractions are those that interact with the fusion protein that binds to the beads. The unbound fraction is the fraction that does not interact with the beads. **(A)** The western blot of a GFP trap experiment in ST*Hdh*$^{Q7/Q111}$ cell lines indicates that N17-YFP interacts with HAP40 (Compare the bound fraction in lane 4 to the other bound fractions, lanes 2, 6, 8, 10, and 12). All of the YFP fusion proteins were expressed, however some constructs had higher expression then others. Ghost bands are present and are identified by an *. Although there is no signal present there is a lot of protein in this location on the blot. (see anti-YFP blot lanes 1, 2, and 3). **(B and C)** GFP trap experiments looking at the interaction between N17-YFP and HAP40 in ST*Hdh*$^{Q7/Q7}$ cells (B) and in ST*Hdh*$^{Q111/Q111}$ cells (C) where HAP40 levels are low. No interaction was observed. **(D)** Following transfection into ST*Hdh*$^{Q7/Q111}$ cells, a heat shock stress was applied to the cells and the GFP trap experiment immediately followed. This western blot shows that the interaction between HAP40 and N17-YFP does not occur when the cell is stressed. **(E)** A far western blot verifies an interaction between HAP40 and N17. The peptides N17, N17M8P and N17(pS13)(pS16) were dot blotted onto a membrane and were incubated with GST-HAP40 (top row), GST (middle row) or water (bottom row). The blot was then probed using a GST antibody. The GST antibody was only able to detect GST-HAP40 on the N17 dot blot suggesting an interaction between HAP40 and N17.

Figure 15

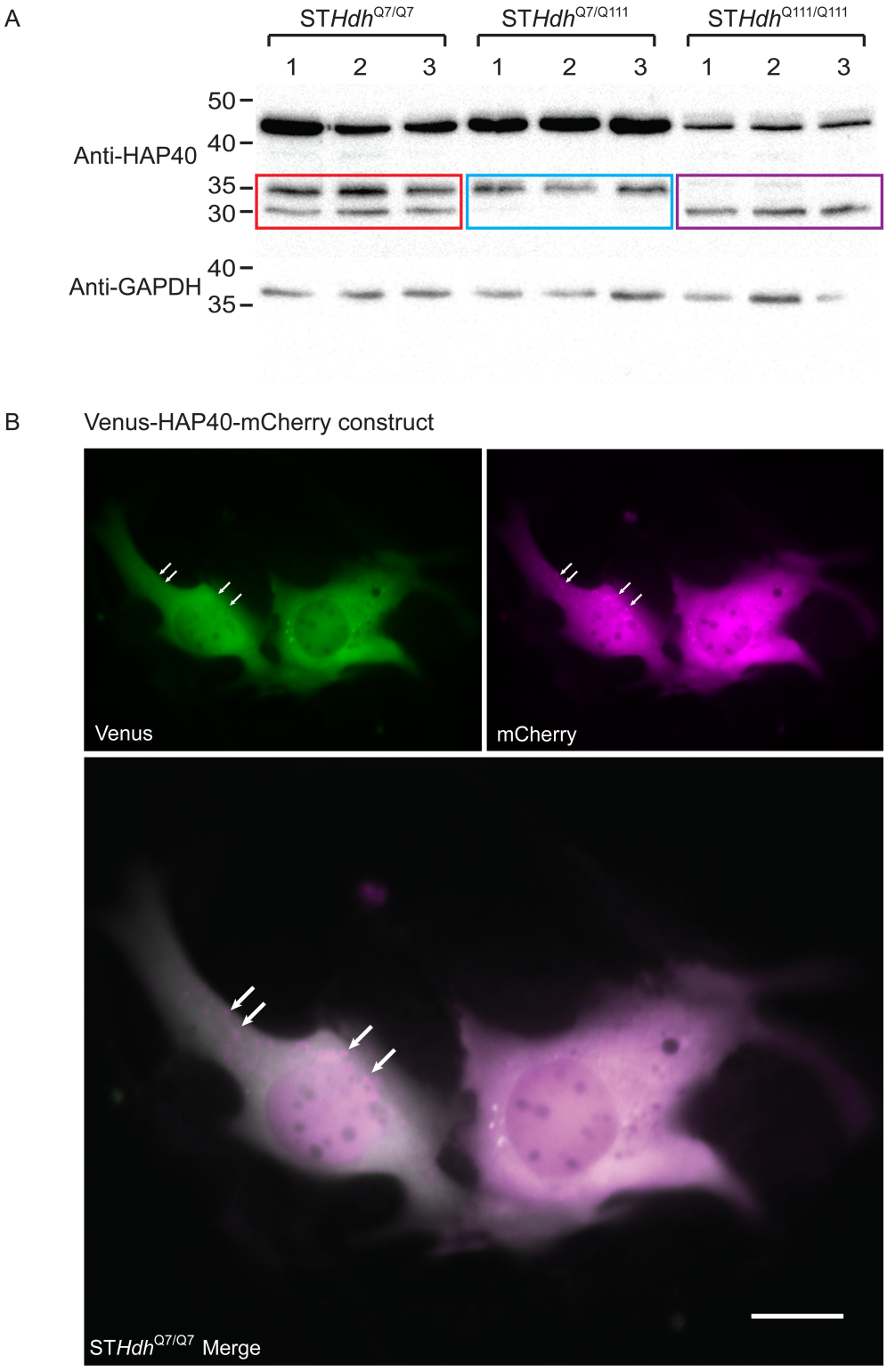

A
STHdhQ7/Q7
STHdhQ7/Q111
STHdhQ111/Q111
1 2 3 1 2 3 1 2 3
50
40
Anti-HAP40
35
30
40
Anti-GAPDH
35
B
Venus-HAP40-mCherry construct
Venus
mCherry
STHdhQ7/Q7 Merge

FIGURE 15

The breakdown of HAP40 occurs both on western blot and in the cell. (A) A western blot demonstrating that breakdown bands of HAP40 occur differently in each ST*Hdh* cell type. The ST*Hdh*$^{Q7/Q7}$ cells have breakdown bands at 35kDa and 30kDa, boxed in red. The ST*Hdh*$^{Q7/Q111}$ cells have a breakdown band at 35kDa boxed in blue, and the ST*Hdh*$^{Q111/Q111}$ cells have a breakdown band at 30kDa boxed in purple. **(B)** A venus-HAP40-mCherry construct, transfected into ST*Hdh*$^{Q7/Q7}$ cells shows that cleavage of HAP40 does occur inside the cell. The arrows point to an area of the cell where puncta of mCherry fluorescence can be seen with no co-localization with venus fluorophore that was attached to the amino terminus of HAP40. The scale bar represents 10μm. Images were brightened using Adobe Photoshop.

References

1. Canadian Insititute for Health Information. The Burden of Neurological Diseases, Disorders and Injuries in Canada. (2007).
2. Paulsen, J.S. et al. Detection of Huntington's disease decades before diagnosis: the Predict-HD study. *J Neurol Neurosurg Psychiatry* **79**, 874-80 (2008).
3. Tabrizi, S.J. et al. Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. *Lancet Neurol* **8**, 791-801 (2009).
4. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington disease chromosomes. *Cell* **72**, 971-983 (1993).
5. Trottier, Y. et al. Cellular localization of the Huntington's disease protein and discrimination of the normal and mutated form. *Nature Genetics* **10**, 104-110 (1995).
6. Persichetti, F. et al. Differential expression of normal and mutant Huntington's disease gene alleles. *Neurobiology of disease* **3**, 183-190 (1996).
7. Petersen, A. & Gabery, S. Hypothalamic and Limbic System Changes in Huntington's Disease. *Journal of Huntington's Disease* **1**, 5-16 (2012).
8. Huntington, G. On chorea. *Med. Surg. Rep.* **26**, 317-321 (1872).
9. Nolte, J. The Human Brain An Introduction to its Functional Anatomy (Mosby, Inc., Missouri, USA, 2002).
10. Galvan, L., Andre, V.M., Wang, E.A., Cepeda, C. & Levine, M.S. Functional Differences Between Direct and Indirect Striatal Output Pathways in Huntington's Disease. *Journal of Huntington's Disease* **1**, 17-25 (2012).
11. Andrew, S.E. et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's Disease. *Nature Genetics* **4**, 398-403 (1993).
12. Joel, D. Open interconnected model of basal ganglia-thalamocortical circuitry and its relevance to the clinical syndrome of Huntington's disease. *Mov Disord* **16**, 407-23 (2001).
13. Vonsattel, J.P. Huntington disease models and human neuropathology: similarities and differences. *Acta Neuropathol* **115**, 55-69 (2008).
14. Vonsattel, J.P. et al. Neuropathological classification of Huntington's Disease. *Journal of Neuropathology and Experimental Neurology* **44**, 559-577 (1985).
15. Aylward, E.H. Change in MRI striatal volumes as a biomarker in preclinical Huntington's disease. *Brain Res Bull* **72**, 152-8 (2007).
16. Zuccato, C., Valenza, M. & Cattaneo, E. Molecular Mechanisms and Potential Therapeutical Targets in Huntington's Disease. *Physiol Rev* **90**, 905-981 (2010).
17. de Roolj, K.E., Dorsman, J.C., Smoor, M.A., Den Dunnen, J.T. & van Ommen, G.B. Subcellular localization of the Huntington's disease gene product in cell lines by immunofluorescence and biochemical subcellular fractionation. *Hum Mol Genet* **5**, 1093-1099 (1996).
18. DiFiglia, M. et al. Huntingtin is a Cytoplasmic Protein Associated with Vesicles in Human and Rat Brain Neurons. *Neuron* **14**, 1075-1081 (1995).

19. Sapp, E. et al. Huntingtin localization in brains of normal and Huntington's disease patients. *Ann Neurol* **42**, 604-12 (1997).
20. Hoogeveen, A.T. et al. Characterization and Localization of the Huntington's Disease Gene Product. *Hum Mol Genet* **2**, 2069-2073 (1993).
21. Atwal, R.S. et al. Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. *Hum Mol Genet* **16**, 2600-15 (2007).
22. Atwal, R.S. et al. Kinase inhibitors modulate huntingtin cell localization and toxicity. *Nat Chem Biol* **7**, 453-60 (2011).
23. Tao, T. & Tartakoff, A.M. Nuclear relocation of normal huntingtin. *Traffic* **2**, 385-94 (2001).
24. Kim, M.W., Chelliah, Y., Kim, S.W., Otwinowski, Z. & Bezprozvanny, I. Secondary structure of Huntingtin amino-terminal region. *Structure* **17**, 1205-12 (2009).
25. Dlugosz, M. & Trylska, J. Secondary structures of native and pathogenic huntingtin N-terminal fragments. *J Phys Chem B* **115**, 11597-608 (2011).
26. Faber, P.W. et al. Huntingtin Interacts with a Family of WW Domain Proteins. *Hum Mol Genet* **7**, 1463-1474 (1998).
27. Desmond, C.R., Atwal, R.S., Xia, J. & Truant, R. Identification of a karyopherin beta1/beta2 proline-tyrosine nuclear localization signal in huntingtin protein. *J Biol Chem* **287**, 39626-33 (2012).
28. Xia, J. Huntingtin contains a highly conserved nuclear export signal. *Hum Mol Genet* **12**, 1393-1403 (2003).
29. Maiuri, T., Woloshansky, T., Xia, J. & Truant, R. The huntingtin N17 domain is a multifunctional CRM1 and Ran-dependent nuclear and cilial export signal. *Hum Mol Genet* **22**, 1383-94 (2013).
30. Zheng, Z., Li, A., Holmes, B.B., Marasa, J.C. & Diamond, M.I. An N-terminal nuclear export signal regulates trafficking and aggregation of Huntingtin (Htt) protein exon 1. *J Biol Chem* **288**, 6063-71 (2013).
31. Li, W., Serpell, L.C., Carter, W.J., Rubinsztein, D.C. & Huntington, J.A. Expression and characterization of full-length human huntingtin, an elongated HEAT repeat protein. *J Biol Chem* **281**, 15916-22 (2006).
32. Grinthal, A., Adamovic, I., Weiner, B., Karplus, M. & Kleckner, N. PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. *Proc Natl Acad Sci U S A* **107**, 2467-72 (2010).
33. Kappel, C., Zachariae, U., Dolker, N. & Grubmuller, H. An unusual hydrophobic core confers extreme flexibility to HEAT repeat proteins. *Biophys J* **99**, 1596-603 (2010).
34. Neuwald, A.F. HEAT Repeats Associated with Condensins, Cohesins, and Other Complexes Involved in Chromosome-Related Functions. *Genome Research* **10**, 1445-1452 (2000).
35. Andrade, M.A., Petosa, C., O'Donoghue, S.I., Muller, C.W. & Bork, P. Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1-18 (2001).

36. Duyao, M.P. et al. Inactivation of the mouse Huntington's disease gene homologue Hdh. *Science* **269**, 407-410 (1995).
37. Woda, J.M. et al. Inactivation of the Huntington's disease gene (Hdh) impairs anterior streak formation and early patterning of the mouse embryo. *BMC Dev Biol* **5**, 17 (2005).
38. White, J.K. et al. Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nature Genetics* **17**, 404-410 (1997).
39. Hilditch-Maguire, P. et al. Huntingtin: an iron regulated protein essential for normal nuclear and perinuclear organelles. *Hum Mol Genet* **9**, 2789-2797 (2000).
40. Gauthier, L.R. et al. Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. *Cell* **118**, 127-38 (2004).
41. Benn, C.L. et al. Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner. *J Neurosci* **28**, 10720-33 (2008).
42. Kegel, K.B. et al. Huntingtin is present in the nucleus, interacts with the transcriptional corepressor C-terminal binding protein, and represses transcription. *J Biol Chem* **277**, 7466-76 (2002).
43. Becanovic, K. et al. Transcriptional changes in Huntington disease identified using genome-wide expression profiling and cross-platform analysis. *Hum Mol Genet* **19**, 1438-52 (2010).
44. Harjes, P. & Wanker, E.E. The hunt for huntingtin function: interaction partners tell many different stories. *Trends in Biochemical Sciences* **28**, 425-433 (2003).
45. Sorkin, A. & von Zastrow, M. Endocytosis and signalling: intertwining molecular networks. *Nat Rev Mol Cell Biol* **10**, 609-22 (2009).
46. Sigismund, S. et al. Endocytosis and signalling: Cell logisticsshape the eukaryotic cell plan. *Physiol Rev* **92**, 273-366 (2012).
47. Jovic, M., Sharma, M., Rahajeng, J. & Caplan, S. The early endosome- a busy sorting station for proteins at the crossroads. *Histol. Histopathol.* **25**, 99-112 (2010).
48. Gruenberg, J. The endocytic pathway: a mosaic of domains. *Nat Rev Mol Cell Biol* **2**, 721-30 (2001).
49. McMahon, H.T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* **12**, 517-33 (2011).
50. Grosshans, B.L., Ortiz, D. & Novick, P. Rabs and their effectors: achieving specificity in membrane traffic. *Proc Natl Acad Sci U S A* **103**, 11821-7 (2006).
51. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. *Nat Rev Mol Cell Biol* **10**, 513-25 (2009).
52. Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nature Reviews Molecular Cel Biology* **2**, 107-119 (2001).
53. Ng, E.L. & Tang, B.L. Rab GTPases and their roles in brain neurons and glia. *Brain Res Rev* **58**, 236-46 (2008).

54. Chavrier, P., Parton, R.G., Hauri, H.P., Simons, K. & Zerial, M. Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. *Cell* **62**, 317-329 (1990).
55. Henne, W.M. et al. FCHo proteins are nucleators of clathrin-mediated endocytosis. *Science* **328**, 1281-4 (2010).
56. McLaughlin, H. et al. A novel role for Rab5-GDI in ligand sequestration into clathrin-coated pits. *Curr Biol* **8**, 34-45 (1997).
57. Pearse, B.M., Smith, C.J. & Owen, D.J. Clathrin coat construction in endocytosis. *Curr Opin Struct Biol* **10**, 220-8 (2000).
58. Mousavi, S.A., Malerod, L., Berg, T. & Kjeken, R. Clathrin-dependent endocytosis. *Biochem J* **377**, 1-16 (2004).
59. Doherty, G.J. & McMahon, H.T. Mechanisms of endocytosis. *Annu Rev Biochem* **78**, 857-902 (2009).
60. Semerdjieva, S. et al. Coordinated regulation of AP2 uncoating from clathrin-coated vesicles by rab5 and hRME-6. *J Cell Biol* **183**, 499-511 (2008).
61. Pearse, B.M. Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. *Proc Natl Acad Sci U S A* **73**, 1255-9 (1976).
62. Mettlen, M., Pucadyil, T., Ramachandran, R. & Schmid, S.L. Dissecting dynamin's role in clathrin-mediated endocytosis. *Biochem Soc Trans* **37**, 1022-6 (2009).
63. Bashkirov, P.V. et al. GTPase cycle of dynamin is coupled to membrane squeeze and release, leading to spontaneous fission. *Cell* **135**, 1276-86 (2008).
64. Roux, A., Uyhazi, K., Frost, A. & De Camilli, P. GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission. *Nature* **441**, 528-31 (2006).
65. Sweitzer, S.M. & Hinshaw, J.E. Dynamin undergoes a GTP-dependent conformational change causing vesiculation. *Cell* **93**, 1021-9 (1998).
66. Mayor, S. & Pagano, R.E. Pathways of clathrin-independent endocytosis. *Nat Rev Mol Cell Biol* **8**, 603-12 (2007).
67. Sandvig, K., Pust, S., Skotland, T. & van Deurs, B. Clathrin-independent endocytosis: mechanisms and function. *Curr Opin Cell Biol* **23**, 413-20 (2011).
68. Huotari, J. & Helenius, A. Endosome maturation. *EMBO J* **30**, 3481-500 (2011).
69. Zeigerer, A. et al. Rab5 is necessary for the biogenesis of the endolysosomal system in vivo. *Nature* **485**, 465-70 (2012).
70. Poteryaev, D., Datta, S., Ackema, K., Zerial, M. & Spang, A. Identification of the switch in early-to-late endosome transition. *Cell* **141**, 497-508 (2010).
71. Nielsen, E., Severin, F., Backer, J.M., Hyman, A.A. & Zerial, M. Rab5 regulates motility of early endosomes on microtubules. *Nat Cell Biol* **1**, 376-382 (1999).
72. Jordens, I., Marsman, M., Kuijl, C. & Neefjes, J. Rab proteins, connecting transport and vesicle fusion. *Traffic* **6**, 1070-7 (2005).
73. Chen, Z.Y., Ieraci, A., Tanowitz, M. & Lee, F.S. A novel endocytic recycling signal distinguishes biological responses of Trk neurotrophin receptors. *Mol Biol Cell* **16**, 5761-72 (2005).

74. Soldati, T. & Schliwa, M. Powering membrane traffic in endocytosis and recycling. *Nat Rev Mol Cell Biol* **7**, 897-908 (2006).
75. Ullrich, O., Reinsch, S., Urbe, S., Zerial, M. & Parton, R.G. Rab11 regulates recycling through the pericentriolar recycling endosome. *Journal of Cell Biology* **135**, 913-924 (1996).
76. Bucci, C., Thomsen, P., Nicoziani, P., McCarthy, J. & van Deurs, B. Rab7: a key to lysosome biogenesis. *Mol Biol Cell* **11**, 467-80 (2000).
77. Maxfield, F.R. & Yamashiro, D.J. Endosome acidification and the pathways of receptor-mediated endocytosis. *Adv Exp Med Biol* **225**, 189-98 (1987).
78. Caviston, J.P., Zajac, A.L., Tokito, M. & Holzbaur, E.L. Huntingtin coordinates the dynein-mediated dynamic positioning of endosomes and lysosomes. *Mol Biol Cell* **22**, 478-92 (2011).
79. Pal, A., Severin, F., Lommer, B., Shevchenko, A. & Zerial, M. Huntingtin-HAP40 complex is a novel Rab5 effector that regulates early endosome motility and is up-regulated in Huntington's disease. *J Cell Biol* **172**, 605-18 (2006).
80. Li, X. et al. A function of huntingtin in guanine nucleotide exchange on Rab11. *Neuroreport* **19**, 1643-7 (2008).
81. Li, X. et al. Mutant huntingtin impairs vesicle formation from recycling endosomes by interfering with Rab11 activity. *Mol Cell Biol* **29**, 6106-16 (2009).
82. Ravikumar, B., Imarisio, S., Sarkar, S., O'Kane, C.J. & Rubinsztein, D.C. Rab5 modulates aggregation and toxicity of mutant huntingtin through macroautophagy in cell and fly models of Huntington disease. *J Cell Sci* **121**, 1649-60 (2008).
83. Lakich, D., Kazazian Jr., H.H., Antonarakis, S.E. & Gitschier, J. Inversions disrupting the factor viii gene are a common cause of hemophilia A. *Nature Genetics* **5**, 236-241 (1993).
84. Levinson, B., Kenwrick, S., Lakich, D., Hammonds, G. & Gitschier, J. A Transcribed gene in an intron of the human factor viii gene. *Genomics* **7**, 1-11 (1990).
85. Levinson, B. et al. Sequence of the human factor VIII-associated gene is conserved in mouse. *Genomics* **13**, 862-865 (1992).
86. Peters, M.F. & Ross, C.A. Isolation of a 40-kDa Huntingtin-associated protein. *J Biol Chem* **276**, 3188-94 (2001).
87. Naylor, J.A. et al. Investigation of the factor viii intron 22 repeated region (int22h) and the associated inversion junctions. *Hum Mol Genet* **4**, 1217-1224 (1995).
88. De Brasi, C.D. & Bowen, D.J. Molecular characteristics of the intron 22 homologs of the coagulation factor VIII gene: an update. *J Thromb Haemost* **6**, 1822-4 (2008).
89. Lannoy, N. et al. Intron 22 homologous regions are implicated in exons 1-22 duplications of the F8 gene. *Eur J Hum Genet* (2013).
90. Milman, P. & Woulfe, J. A novel variant of neuronal intranuclear rodlet immunoreactive for 40 kDa huntingtin associated protein and ubiquitin in the mouse brain. *J Comp Neurol* (2013).
91. Landry, J. et al. The genomic and transcriptomic landscape of a HeLa cell line. *G3: Genes|Genomes|Genetics* **3**, 1213-1224 (2013).

92. Walboomers, J.M.M. et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *Journal of Pathology* **189**, 12-19 (1999).
93. Baneyx, F. & Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nat Biotechnol* **22**, 1399-408 (2004).
94. Olsson, N., Wallin, S., James, P., Borrebaeck, C.A. & Wingren, C. Epitope-specificity of recombinant antibodies reveals promiscuous peptide-binding properties. *Protein Sci* **21**, 1897-910 (2012).
95. Dinkel, H. et al. ELM--the database of eukaryotic linear motifs. *Nucleic Acids Res* **40**, D242-51 (2012).
96. Deveraux, Q.L. & Reed, J.C. IAP family proteins - suppressors of apoptosis. *Genes Dev.* **13**, 239-252 (1999).
97. Salvesen, G.S. & Duckett, C.S. IAP proteins: blocking the road to death's door. *Nat Rev Mol Cell Biol* **3**, 401-10 (2002).
98. Schimmer, A.D. Inhibitor of apoptosis proteins: translating basic knowledge into clinical practice. *Cancer Res* **64**, 7183-90 (2004).
99. Eckelman, B.P., Drag, M., Snipas, S.J. & Salvesen, G.S. The mechanism of peptide-binding specificity of IAP BIR domains. *Cell Death Differ* **15**, 920-8 (2008).
100. Green, John, C., nbsp & Reed, D.R. Mitochondria and Apoptosis. *Science* **281**, 1309-1312 (1998).
101. Chen, C.C., Hwang, J.K. & Yang, J.M. (PS)2: protein structure prediction server. *Nucleic Acids Res* **34**, W152-7 (2006).
102. duVerle, D., Takigawa, I., Ono, Y., Sorimachi, H. & Mamitsuka, H. CaMPDB: a Resource for Calpain and Modulatory Proteolysis. *Genome Informatics* **22**, 202-214 (2009).
103. Liu, J., Liu, M.C. & Wang, K.K. Calpain in the CNS: from synaptic function to neurotoxicity. *Sci Signal* **1**, re1 (2008).
104. Vosler, P.S., Brennan, C.S. & Chen, J. Calpain-mediated signaling mechanisms in neuronal injury and neurodegeneration. *Mol Neurobiol* **38**, 78-100 (2008).
105. Gafni, J. & Ellerby, L.M. Calpain activation in Huntington's disease. *Journal of Neuroscience* **22**, 4842-4849 (2002).
106. Tang, T. Huntingtin and Huntingtin-Associated Protein 1 Influence Neuronal Calcium Signaling Mediated by Inositol-(1,4,5) Triphosphate Receptor Type 1. *Neuron* **39**, 227-239 (2003).
107. Alexandrov, K., Horiuchi, H., Steele-Mortimer, O., Seabra, M.C. & Zerial, M. Rab escort protein-1 is a multifunctional protein that accompanies newly prenylated rab proteins to their target membranes. *The EMBO Journal* **13**, 5262-5273 (1994).
108. Seabra, M.C., Goldstein, J.L., Sudhof, T.C. & Brown, M.S. Rab geranylgeranyl transferase. A multisubunit enzyme that prenylates GTP-binding proteins terminating in Cys-X-Cys or Cys-Cys. *J Biol Chem* **267**, 14497-14503 (1992).
109. Zhu, G. et al. Structural basis of Rab5-Rabaptin5 interaction in endocytosis. *Nat Struct Mol Biol* **11**, 975-83 (2004).

110. Ullrich, O., Horiuchi, H., Bucci, C. & Zerial, M. Membrane association of Rab5 mediated by GDP-dissociation inhibitor and accompanied by GDP/GTP exchange. *Nature* **368** (1994).