THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 5, Issue of February 2, pp. 3188–3194, 2001

Printed in U.S.A.

Isolation of a 40-kDa Huntingtin-associated Protein*

Received for publication, September 5, 2000, and in revised form, October 12, 2000
Published, JBC Papers in Press, October 16, 2000, DOI 10.1074/jbc.M008099200

Matthew F. Peters‡§¶ and Christopher A. Ross‡¶‖**

*From the ‡Laboratory of Molecular Neurobiology, Division of Neurobiology, Department of Psychiatry, ‖Department of Neuroscience, and **The Program in Cellular and Molecular Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2196*

Huntington's disease is caused by an expanded CAG trinucleotide repeat coding for a polyglutamine stretch within the huntingtin protein. Currently, the function of normal huntingtin and the mechanism by which expanded huntingtin causes selective neurotoxicity remain unknown. Clues may come from the identification of huntingtin-associated proteins (HAPs). Here, we show that huntingtin copurifies with a single novel 40-kDa protein termed HAP40. HAP40 is encoded by the open reading frame factor VIII-associated gene A (F8A) located within intron 22 of the factor VIII gene. In transfected cell extracts, HAP40 coimmunoprecipitates with full-length huntingtin but not with an N-terminal huntingtin fragment. Recombinant HAP40 is cytoplasmic in the presence of huntingtin but is actively targeted to the nucleus in the absence of huntingtin. These data indicate that HAP40 is likely to contribute to the function of normal huntingtin and is a candidate for involvement in the aberrant nuclear localization of mutant huntingtin found in degenerating neurons in Huntington's disease.

Huntington's disease (HD)[1] is an autosomal dominant neurodegenerative disorder characterized by motor and cognitive impairments that begin in mid-life and progress over 15–20 years to death. The genetic defect underlying HD is an expanded CAG trinucleotide repeat encoding a polyglutamine stretch within the huntingtin protein (1). Normally, huntingtin's CAG repeat codes for 6–35 consecutive glutamines, whereas in HD, it is expanded beyond 35. At least seven other progressive neurodegenerative diseases result from polyglutamine repeat expansions within otherwise unrelated host proteins (2). In each disorder, polyglutamine expansion causes a toxic gain of function, with longer repeats resulting in increased severity. Accumulating data are consistent with the hypothesis that expanded polyglutamine repeats adopt a conformation that facilitates toxic protein-protein interactions and aggregation (3, 4).

Despite sharing similar pathogenic mechanisms, polyglutamine repeat disorders display distinctive neuropathologies. In each disease, a different subset of central nervous system neurons is targeted that cannot be attributed to the expression pattern of the expanded protein. For example, HD most severely affects medium spiny neurons in the striatum (5, 6), whereas spinocerebellar ataxia 1 targets cerebellar Purkinje cells, although both cell types express the huntingtin and ataxin-1 proteins (for further review, see Ref. 7). Thus, the selective vulnerability to polyglutamine toxicity is influenced by properties of the host protein. Protein characteristics known to alter polyglutamine toxicity in animal and cellular models include nuclear localization, high local protein concentration, protein-protein interactions, and protein size or truncation. Each of these features is likely to be determined by the protein's normal function. Thus, it may be necessary to identify huntingtin's normal cellular roles to understand how polyglutamine expansion in huntingtin results in the selective toxicity characteristic of HD.

Huntingtin is a 348-kDa protein that lacks close similarity to known proteins and has no known function (1). It is expressed in all tissues, with the highest levels found in brain and testis (6). Within the brain, huntingtin is expressed in the cytoplasm of neurons (8). Targeted mutagenesis that significantly reduces huntingtin expression in mice causes aberrant neurogenesis and perinatal death (9). Complete disruption of the huntingtin gene in mice results in embryonic lethality at gastrulation with increased neuronal apoptosis (10–12). Recent data from cultured neurons support an antiapoptotic role for normal huntingtin, although the mechanism remains unclear (13).

Yeast two-hybrid screens have identified a number of proteins that interact with huntingtin, known as huntingtin-interacting proteins (HIPs), huntingtin-yeast partners (HYPs), and huntingtin-associated proteins (HAPs). These associated proteins implicate huntingtin in a number of cellular processes (reviewed in Refs. 7 and 14). Among the interactors, HAP1, HYP-J (α-adaptin-C), SH3GL3 (endophilin III), and HIP3 are believed to be involved in vesicular recycling or trafficking (15–21). Several interactors, including nuclear corepressor, HYP-A, and HYP-I (symplekin), have been linked to nuclear functions (20, 22). Other interactors include: (*a*) HIP1, which is related to yeast Sla2p, a transmembrane protein with a talin-like domain (23, 24); (*b*) HIP2, which is an E2 ubiquitin-conjugating enzyme (25); (*c*) HYP-F, which is part of the 26S proteosome (20); and (*d*) cystathionine β-synthase, which is a key enzyme in the generation of cysteine (26). Six additional HYPs have been identified but have no known cellular roles (20). Other data suggest that cAMP-response element-binding pro-

* This work was supported by Huntington's Disease Society of America "Coalition for the Cure" and National Institutes of Health Grants NS16375, NS34172, and NS38144 (to C. A. R.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF299331.

§ Supported by a Huntington's Disease Society of America postdoctoral fellowship.

¶ To whom correspondence may be addressed: Laboratory of Molecular Neurobiology, Departments of Psychiatry and Neuroscience, Johns Hopkins University School of Medicine, Ross 618, 720 Rutland Ave., Baltimore, MD 21205-2196. Tel.: 410-614-0011; Fax: 410-614-0013; E-mail: mfpeters@jhmi.edu (M. F. P.) or caross@jhu.edu (C. A. R.).

[1] The abbreviations used are: HD, Huntington's disease; HAP, huntingtin-associated protein; NLS, nuclear localization signal; rAb, rabbit polyclonal antibody; mAb, monoclonal antibody; HIP, huntingtin-interacting protein; HYP, huntingtin-yeast partner; PBS, phosphate-buffered saline; MAP, mitogen-activated protein; PYK, proline-rich tyrosine kinase; EST, expressed sequence tag; GFP, green fluorescent protein; F8A, factor VIII-associated gene A.

tein-binding protein, TATA-binding protein, epidermal growth factor receptor, mixed-lineage kinase 2, and calmodulin can associate with huntingtin under some conditions (27–32).

The known huntingtin-associated proteins are likely to represent only a partial list of proteins that complex with huntingtin *in vivo*. All 18 of the huntingtin-interacting proteins identified in yeast bind to the N-terminal 18% of huntingtin. The absence of proteins binding the central and C-terminal regions may reflect technical limitations of expressing large proteins in yeast. The central 1500 amino acids of huntingtin comprise 10 HEAT repeats that are likely to form a single extended structure and mediate protein-protein interactions. No domain information is available to guide the design of small yeast baits from within the C-terminal 1500 amino acids of huntingtin. Thus, to identify proteins that associate with full-length huntingtin *in vivo*, we purified huntingtin complexes from rat brain extracts and identified the protein components.

MATERIALS AND METHODS

Fractionation—Rat brain extracts were prepared in a glass/Teflon homogenizer in 10 volumes of PBS containing Complete protease inhibitors (Roche Molecular Biochemicals) and 1 mM phenylmethylsulfonyl fluoride. Homogenates were successively centrifuged at 1,000 × *g* for 10 min, 10,000 × *g* for 20 min, and 100,000 × *g* for 1 h. Pellets were washed once and resuspended in PBS. Protein concentration was determined by BCA assay (Pierce).

Antibody Preparation—Antibodies were prepared according to standard procedures (Cocalico Biological). A goat polyclonal antibody was generated against a fusion, HntΔex1, corresponding to amino acids 1–23 and 69–92 of human huntingtin. For immunization, a HntΔex1-glutathione *S*-transferase fusion protein was expressed, purified on glutathione-agarose, resolved by SDS-polyacrylamide gel electrophoresis, and injected as a Coomassie Blue-stained gel slice. A purified HntΔex1-His-tagged fusion protein was used for affinity purification of huntingtin antibodies from serum and anti-glutathione *S*-transferase antibodies. Rabbit polyclonal antibodies (rAbs) were prepared against peptides coupled to keyhole limpet hemocyanin through a terminal cysteine. Huntingtin rAbs were generated against the peptides MATLEKLMKAFESLKSFQC and CITEQPRSQHTLQ corresponding to amino acids 1–18 and 497–508, respectively. HAP40 rAbs were prepared against peptides CRY- RQVSNKLKKRFLRKPN, CPQPPSG-PQPPLSGPQPRP, and CDGHGQDTSGQLPEE corresponding to amino acids 25–42, 226–243, and 314–326, respectively. With the exception of the crude antisera used for the immunoprecipitation in Fig. 5*A*, all antibodies were affinity-purified using antigen coupled to Affi-Gel (Bio-Rad) as described previously (33). MAP2 and PYK2 antibodies were purchased from Transduction Laboratories. The huntingtin monoclonal antibody (mAb) 2166 was purchased from Chemicon. The M2 anti-FLAG antibody was obtained from Sigma. The HAP1 antibody was obtained from Alan Sharp, the Johns Hopkins University, Baltimore, MD. The HIP2 antibody was a gift of Cecile Pickart, the Johns Hopkins University, Baltimore, MD.

Immunoaffinity Purification—Protein complexes were purified from rat brain (Pel-Freez, Rogers, AZ). Brains were homogenized in 10 volumes of PBS, 1% Triton X-100. Complete protease inhibitors (Roche Molecular Biochemicals), and 1 mM phenylmethylsulfonyl fluoride. This homogenization buffer was modified with either 1 mM EDTA or 2 μM Ca^{2+}, with or without phosphatase inhibitors, and with or without detergent added after 100,000 × *g* centrifugation. In all experiments, extracts were cleared by centrifugation at 100,000 × *g* for 1 h. For immunoaffinity purification, 20 ml of cleared brain extract was combined with antibody resins prepared by coupling 5 mg of affinity-purified antibody to 2 ml of Affi-Gel-10. Beads and brain extract were mixed in batch for 2 h and collected in a column. Beads were successively washed with 50 ml of PBS/1% Triton X-100, 10 ml of PBS/0.1% Triton X-100, and 5 ml of PBS. Proteins were eluted with 5 ml of 0.2 M glycine, pH 2.5. Samples were concentrated by lyophilization or trichloroacetic acid precipitation. For immunoprecipitation, 2 μl of antibody was combined with 1 ml of extract according to the standard procedures (34).

Protein Identification—Samples were resolved by SDS-polyacrylamide gel electrophoresis on 4–15% gradient gels (Bio-Rad) and stained with Coomassie Blue. Proteins were excised, and sequence analysis was performed at the Harvard Microchemistry Facility, Cambridge, MA by microcapillary reverse-phase high pressure liquid chromatography nano-electrospray tandem mass spectrometry. For two-dimensional SDS-polyacrylamide gel electrophoresis**,** isoelectric focusing was performed in tube gels (Kenderick Laboratories, Madison, WI) or on pH 3–10 linear IPGphor strips. Duplicate two-dimensional gels were run for all samples. Immunoblotting was performed as described previously (33).

HAP40 Cloning and Expression—A mouse EST known to contain at least the 5′ 763 nucleotides of the F8A/HAP40 open reading frame was obtained from the Japanese National Institute of Infectious Disease (Tokyo, Japan) (EST MNCb-6402, GenBank accession number AU035947). The 1589-base insert was subcloned into the *Eco*RI/*Xba*I sites in pcDNA3.1 and sequenced on both strands. An epitope-tagged construct was prepared by cloning oligonucleotides encoding the C-terminal 21 amino acids of HAP40 and a FLAG tag into the *Blp*I and *Xba*I sites. The amino acids KKR in the putative nuclear localization signal were changed to STS. HEK293 cells were transfected with LipofectAMINE Plus (Life Technologies, Inc.) and immunolabeled as described previously (35). Fluorescence images are 0.8-μm optical sections captured on a Zeiss LSM-410 confocal microscope.

HAP40 in HD Mouse Models and HD Patient Brain—Brain sections from mice expressing an N-terminal 171–82Q (36) or littermate non-transgenic controls were immunostained using the Vectastain ABC labeling system (Vector Laboratories). Biochemical enrichment of huntingtin aggregates from post-mortem human brains was performed by homogenizing aggregates in PBS and centrifuging them at 1,000 × *g* for 5 min. Pellets were collected, boiled in 2% SDS, and resolved by SDS-polyacrylamide gel electrophoresis. Immunoblotting with anti-peptide antibodies against huntingtin but not HAP40 clearly detected proteins in the stacking gel from HD patients (data not shown).

RESULTS

To prepare quantities of antibody suitable for immunoaffinity purification, we immunized goats with a fusion protein corresponding to huntingtin exon 1 with an in-frame deletion of part of the polyglutamine and polyproline regions. We chose this immunogen because N-terminal fusion proteins with deletions have been shown to generate rabbit antibodies that strongly label native huntingtin, especially with an expanded polyglutamine (37).[2]

Consistent with previous observations in rodents (38), we noted that rat huntingtin was found predominately in the soluble fraction (Fig. 1*A*). Lower levels of huntingtin were consistently detected in particulate fractions, suggesting specific association with membranes, the nucleus, or other cellular structures. To extract the maximum amount of soluble huntingtin for purification, we used rat brain homogenates treated with Triton X-100.

Immunoaffinity purification with the anti-HntΔex1 antibody dramatically enriched for huntingtin (Fig. 1*B*). Purification was specific because immunoblotting did not detect huntingtin in samples eluted from goat IgG columns (Fig. 1*B*). As expected from a one-step preparation, Coomassie Blue-stained gels revealed that huntingtin samples were highly enriched in a ≈350-kDa protein but were not pure (Fig. 1*C*). Compared with the proteins present in control samples, four major proteins of ≈40, 70, 115, and 350 kDa were consistently enriched in the anti-huntingtin samples (Fig. 1*C*; data not shown). Several variations in the extraction conditions (see "Materials and Methods") did not significantly alter the proteins that specifically coenriched in huntingtin samples (data not shown). The four protein bands were identified by mass spectroscopy/mass spectroscopy sequencing of tryptic peptides. The 350-kDa band yielded 71 nonoverlapping peptide sequences corresponding to huntingtin and 5 peptides corresponding to MAP2(a/b). The 115-kDa band was comprised of PYK2. The 70-kDa band contained MAP2(c). The sequence from the 40-kDa protein band corresponded to open reading frame F8A located on chromosome X within intron 22 of the factor VIII gene (39). We refer to this 40-kDa huntingtin-associated protein as HAP40.

[2] J. D. Wood and C. A. Ross, unpublished data.

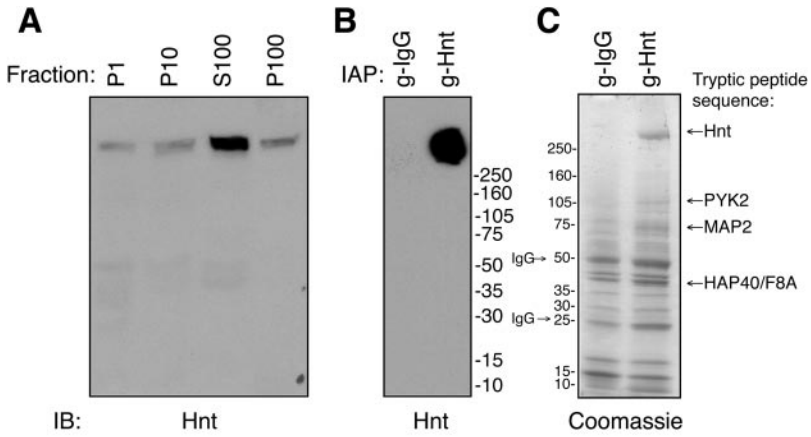

FIG. 1. **Immunoaffinity purification of huntingtin.** *A*, rat brain extracts were successively centrifuged at 1,000 × *g*, 10,000 × *g*, and 100,000 × *g*. 40 μg of protein from the resulting pellets (*P1, P10*, and *P100*, respectively) and the final supernatant (*S100*) were immunoblotted for huntingtin (mAb 2166). *B*, control and huntingtin samples were immunoaffinity-purified from Triton X-100 solubilized rat brain extracts using goat IgG and goat polyclonal antibody HntΔex1 columns. Samples (0.5 μl) were blotted for huntingtin (mAb 2166). *C*, goat IgG and goat polyclonal antibody HntΔex1 immunoaffinity-enriched samples (5 μl) were resolved on 4–15% gels and stained with Coomassie Blue. Four protein bands of ~350, 115, 70, and 40 kDa were coenriched in huntingtin samples compared with IgG samples. The 40-kDa protein migrates below a doublet of more intensely stained nonspecific proteins at ~41 and 45 kDa. Protein sequencing identified these four proteins as huntingtin, PYK2, MAP2, and a novel protein termed HAP40, respectively. PYK2, MAP2(b), and MAP2(c) have predicted molecular masses of 118, 199, and 70 kDa, respectively.

Immunoblot analysis with commercial antibodies to PYK2 and MAP2 and newly prepared HAP40 antibodies (described below) confirmed that each protein was highly enriched in samples eluted from anti-HntΔex1 columns compared with those eluted from goat IgG columns (data not shown). However, sequence comparisons revealed a common proline-rich sequence, PPP*X*P, in the HntΔex1 fusion protein, PYK2, and MAP2(a/b/c). This raises the possibility that both PYK2 and MAP2 were enriched by binding to antibodies against huntingtin's proline-rich region rather than by associating with the huntingtin protein. To distinguish between these possibilities, we prepared new antibodies against huntingtin peptides 1–18 and 497–508, which lack consecutive prolines. Immunoaffinity purification with each of these antibodies greatly enriched for huntingtin and HAP40 but not for PYK2 or MAP2. With long exposures, PYK2 and MAP2 could be detected nonspecifically in both anti-huntingtin and rabbit IgG samples (Fig. 2*A*). In contrast, HAP40 was highly enriched in both rAb $Hnt_{1–18}$ and rAb $Hnt_{497–508}$ samples in comparison with the IgG control (Fig. 2*A*). Thus, we conclude that HAP40 is tightly associated with huntingtin after Triton X-100 solubilization.

To identify additional proteins copurifying with huntingtin in our assay, we immunoblotted for the previously identified huntingtin-interacting proteins. Immunoblots for HAP1 and HIP2 clearly detected proteins migrating at the predicted size in crude soluble fractions. However, neither protein was found to copurify with huntingtin despite increasing the amount of sample loaded by 20-fold compared with that needed to saturate the immunoblot signal for huntingtin (Fig. 2*B*). Similar immunoblots probed for other huntingtin-binding proteins including HIP1, SH3GL3, symplekin (HYP-I), α-adaptin (HYP-J), calmodulin, epidermal growth factor receptor, and the nuclear corepressor failed to reveal any enrichment in huntingtin samples (data not shown). We conclude that none of these proteins form Triton X-100-resistant complexes with soluble huntingtin (see "Discussion").

To search for additional interacting proteins that may be present at levels too low to be detected by Coomassie Blue staining or obscured by comigrating contaminate proteins, we analyzed immunoaffinity-purified huntingtin samples by two-dimensional gel electrophoresis and silver staining. Samples eluted from rAb $Hnt_{1–18}$, rAb $Hnt_{497–508}$, goat polyclonal antibody HntΔex1, rabbit IgG, and goat IgG columns were com-

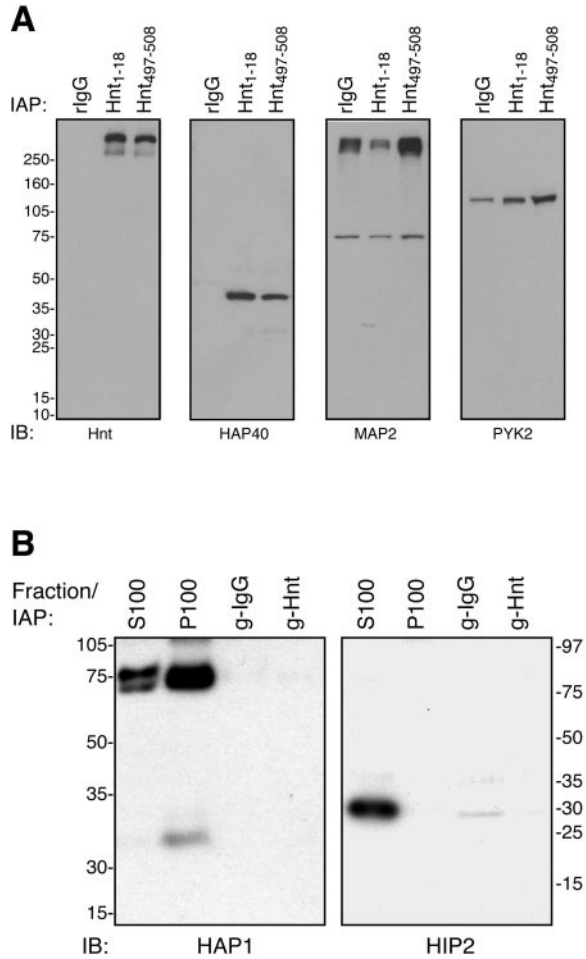

FIG. 2. **Immunoblot analysis of partially purified huntingtin-associated protein complexes.** *A*, samples immunoaffinity-purified with rAb $Hnt_{1–18}$ and $Hnt_{497–508}$ were compared with rabbit IgG samples for enrichment of HAP40, PYK2, and MAP2. Only HAP40 was specifically enriched in each huntingtin sample. PYK2 and MAP2 blots were deliberately overexposed to demonstrate that similar levels were present in huntingtin and control samples. *B*, crude brain fractions (40 μg) and immunoaffinity-purified samples (10 μl) were immunoblotted with for HAP1 and HIP2. HAP1 and HIP2 have predicted molecular masses of 68 and 28 kDa, respectively.

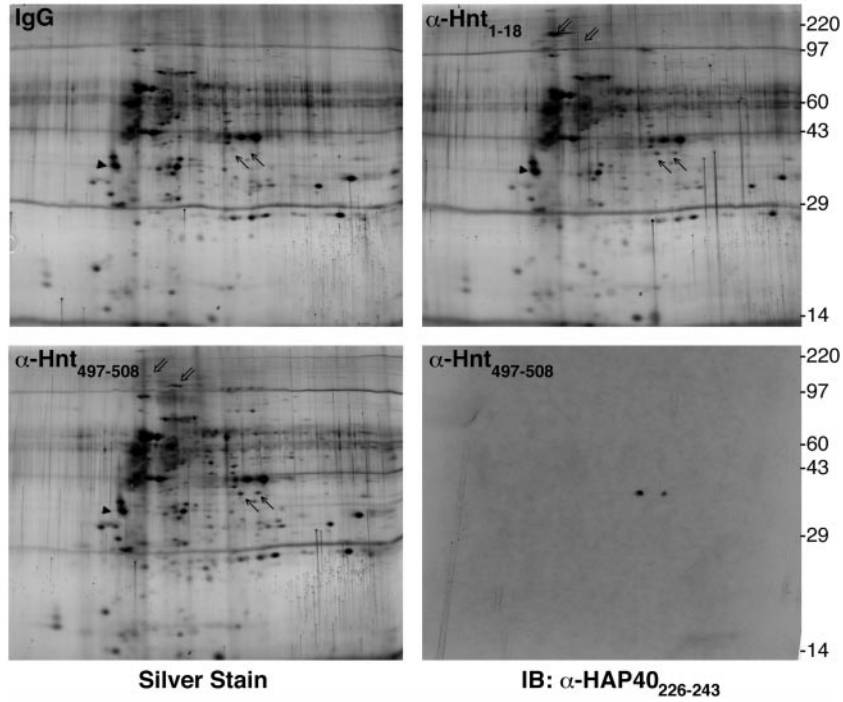

FIG. 3. **Huntingtin-associated protein complexes resolved by two-dimensional gels and detected by silver-staining.** Two 40-kDa protein spots were present in samples purified with rAbs Hnt_{1-18} and $Hnt_{497-508}$ but not in samples purified with rabbit IgG (*single arrows*). All additional proteins in anti-huntingtin samples were either present in IgG controls or were unique to Hnt_{1-18}, $Hnt_{497-508}$, or HntΔex1 preparations (compare *double arrows* for Hnt_{1-18} and $Hnt_{497-508}$). Immunoblotting the $Hnt_{497-508}$ preparation with $HAP40_{226-243}$ revealed two 40-kDa spots in a pattern indistinguishable from the unique silver-stained spots. 25 μg of protein was loaded on each gel. Duplicate gels for each sample showed similar results. Molecular mass markers are in kDa; acidic and basic pH are on the *left* and *right* respectively.

pared (Fig. 3; data not shown). The pattern of protein spots was highly similar in all samples and reproducible in several experiments. Each of the five samples contained some proteins that were unique to that antibody column (Fig. 3, compare *double arrows*). Two 40-kDa protein spots were detected in all three huntingtin samples but were not detected in either IgG control (Fig. 3, *single arrows*). No other potential HAPs were identified in more than one huntingtin sample and absent in IgG samples. Immunoblotting similar gels with HAP40 antibodies revealed a pair of 40-kDa spots in a pattern indistinguishable from the two silver-stained spots (Fig. 3*D*). The reason for two HAP40 spots is unknown. We conclude that HAP40 is a major component of the huntingtin protein complex.

For expression analysis, we obtained mouse EST MNCb-6402, which was previously reported to code for at least the N-terminal 254 amino acids of HAP40/F8A. Complete sequencing of this EST revealed a Kozak start site followed by an open reading frame encoding 381 amino acids with a calculated molecular mass of 40,472.78 Da. This EST sequence differed from the previously reported mouse F8A sequence (39) by one amino acid insertion and two substitutions (Fig. 4). These changes are unlikely to result from errors in the mouse EST sequence because each of these three residues in the mouse EST was identical to human ESTs. None of the differing residues were confirmed by mass spectroscopy/mass spectroscopy peptide sequences (Fig. 4).

To further test the huntingtin-HAP40 association, we prepared anti-HAP40 peptide antibodies. Two HAP40 antisera coimmunoprecipitated huntingtin from mouse brain extracts (Fig. 5*A*). This was judged to be specific because preimmune serum failed to coimmunoprecipitate huntingtin. When considered together, the coimmunoprecipitation of huntingtin with two anti-HAP40 antibodies and the copurification of HAP40 with three anti-huntingtin antibodies strongly indicate that huntingtin and HAP40 form a stable complex.

To determine the region of huntingtin required for association with HAP40, HEK293 cells were cotransfected with HAP40 and huntingtin constructs of various lengths. Full-length huntingtin specifically coimmunoprecipitated with HAP40 (Fig. 5*B*). Despite being expressed at levels higher than

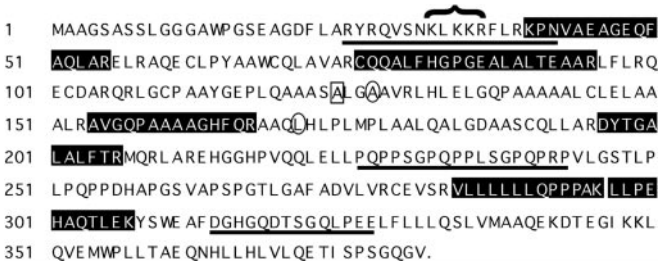

FIG. 4. **Protein sequence of HAP40/F8A.** The translated sequence for mouse EST MNCb-6402 is shown. The EST sequence differs from the previously published mouse sequence, termed F8A, by one insertion and two substitutions (denoted by an *open box* and *circles*, respectively). Six mass spectroscopy/mass spectroscopy peptide sequences obtained from rat HAP40 are indicated by *shaded boxes*. Leucine 203 in the mouse EST was a valine in rat the mass spectroscopy/mass spectroscopy peptide and in human ESTs. Synthetic peptides used for antibody production are *underlined*. The basic amino acids mutated to inactive the putative nuclear localization signal are *bracketed*.

full-length huntingtin, a construct corresponding to the N-terminal 513 amino acids failed to coimmunoprecipitate with HAP40 (Fig. 5*B*). HAP40 is the first protein known to require either the central or C-terminal regions of huntingtin for association. Although additional experiments will be needed to prove that the HAP40-huntingtin interaction is direct, it seems unlikely that other proteins mediate this interaction because none were detected in the purified huntingtin complexes (Fig. 3).

To compare the subcellular localization of HAP40 and huntingtin, we immunolabeled sections of mouse brain and transfected cells. None of the anti-peptide antibodies against HAP40 labeled normal brain sections. Because each antibody recognized native protein in transfected cells, these data are consistent with low expression levels of HAP40 proteins in normal brain. On sections of HD-N171–82Q transgenic mouse brain, the anti-huntingtin antibodies but not the anti-HAP40 antibodies labeled inclusions (data not shown). In HEK293 cells transiently transfected with full-length huntingtin and HAP40, immunostaining for both proteins showed a similar cytoplasmic distribution (Fig. 6, *top panels*). Cotransfection of either

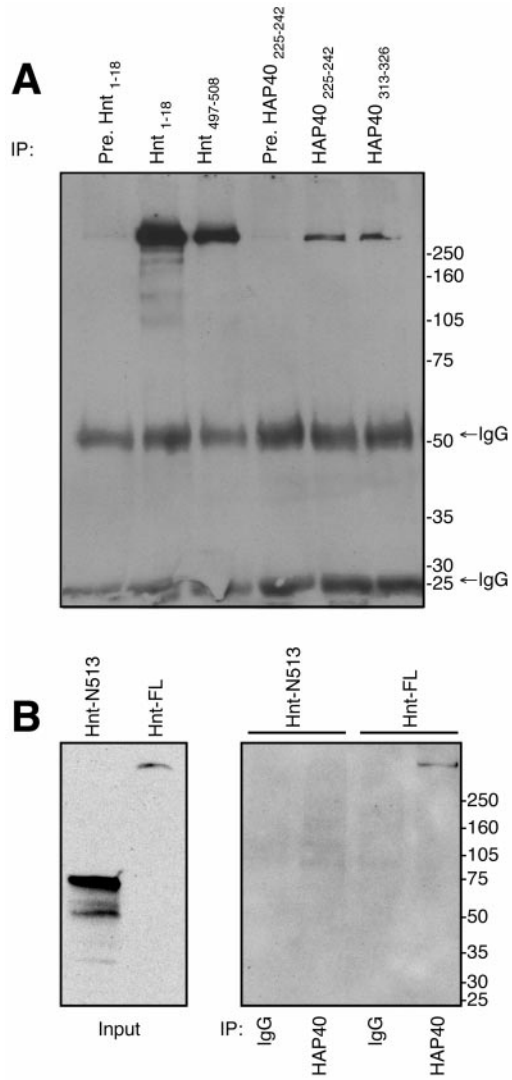

FIG. 5. **HAP40 associates with full-length huntingtin but not with an N-terminal huntingtin fragment.** *A*, HAP40 was immunoprecipitated from mouse brain extracts. Huntingtin was coimmunoprecipitated with rAb $HAP40_{25-42}$ and rAb $HAP40_{226-243}$ but not with preimmune sera. *B*, HEK293 cells were cotransfected with HAP40 and either full-length huntingtin (FL-23Q) or a construct corresponding to the N-terminal 513 amino acids (N513–18Q). Cleared lysates used as starting material were blotted to show expression levels (*left panel*). Immunoprecipitation with rAb $HAP40_{226-243}$ specifically coenriched for full-length huntingtin (FL-23Q) but not for N513–18Q (*right panel*). Samples were blotted for huntingtin with mAb 2166.

normal (23Q) or expanded (82Q) full-length huntingtin with HAP40 resulted in similar colocalization in the cytoplasm (data not shown). Surprisingly, in cells that lack detectable endogenous huntingtin, HAP40 transfected alone was concentrated in the nucleus, both diffusely and in punctate nuclear structures (Fig. 6, *bottom panels*). The HAP40 nuclear dots did not overlap with staining for a splicing factor (mAb SC35) or promyelocytic leukemia protein oncogenic domains detected with anti-PML antibodies (data not shown). Although the nature of HAP40 nuclear structures remains unclear, they are unlikely to result from nonspecific aggregation of overexpressed protein because they were observed in individual cells expressing lower levels of protein (Fig. 6).

HAP40 is smaller than the ≈ 45-kDa threshold for passive diffusion through nuclear pores. However, passive diffusion alone would not be expected to result in the relative concentration of HAP40 in the nucleus. To test the possibility that HAP40 is actively targeted to the nucleus, we examined the localization of a 68-kDa HAP40-GFP fusion protein. The HAP40-GFP fusion was concentrated in the nucleus (Fig. 7, *upper panels*). Mutation of a lysine/arginine-rich region resembling a nuclear localization signal (NLS) resulted in a significant redistribution of HAP40-GFP to the cytoplasm (Fig. 7, *lower panels*). Low levels of HAP40-GFP mutant NLS were detectable in the nucleus, suggesting that HAP40 may be actively targeted to the nucleus by multiple mechanisms (Fig. 7).

DISCUSSION

We immunoaffinity-purified huntingtin and characterized the associated proteins. We found that most rat huntingtin is soluble and that soluble huntingtin associates with a novel protein, HAP40. None of the previously identified huntingtin-associated proteins were found to copurify with huntingtin in our assay. This result may be due to several reasons, the most notable of which is that our purification system requires soluble huntingtin and the presence of detergent to reduce nonspecific protein associations. Thus, we would not expect to purify huntingtin-associated protein complexes that were either detergent-sensitive or detergent-insoluble. Interestingly, HAP1 and SH3GL3 have been shown to form detergent-resistant complexes with huntingtin from transfected cells (15, 19), yet neither was found to copurify with huntingtin from brain, raising the possibility that a number of proteins may associate with huntingtin in a regulated fashion. The stable association of HAP40 with huntingtin suggests that it is likely to hold important clues to the cellular role of huntingtin.

Peptide sequencing of HAP40 revealed that it is encoded by an open reading frame previously termed F8A. F8A was identified as a putative gene located within intron 22 of the factor VIII gene but transcribed in the direction opposite factor VIII. In humans, a portion of factor VIII intron 22 including F8A is duplicated with two additional copies located nearer to the Xq telomere. Intrachromosomal recombination between intron 22 and either of these two copies interrupts the factor VIII gene and underlies almost half of all severe hemophilia A cases (40). Our data prove that the single F8A sequence in mouse is expressed. However, because the three intron 22 sequences in human are 99% identical, it remains unclear which of the three human F8A sequences are expressed.

F8A/HAP40 mRNA is expressed in a wide range of mouse tissues, similar to huntingtin (39). When recombinant HAP40 and huntingtin are coexpressed in cells, they colocalize in the cytoplasm. In the absence of huntingtin, HAP40 is concentrated in the nucleus, where it is both diffuse and in punctate nuclear structures. We found that a HAP40-GFP fusion protein too large to diffuse through nuclear pores was concentrated in the nucleus. Mutation of a NLS largely redistributed HAP40-GFP to the cytoplasm, indicating that HAP40 is actively targeted to the nucleus. Future experiments will be needed to determine the function of nuclear HAP40.

The mechanism by which huntingtin redistributes HAP40 to the cytoplasm in transfected cells is unclear. Huntingtin is normally concentrated in the cytoplasm and may simply anchor HAP40. Alternatively, several lines of evidence are consistent with regulated active nuclear transport of huntingtin. Huntingtin contains several consensus NLSs (41) and >10 sequences matching leucine-rich nuclear export signals.[3] Huntingtin is comprised of 10 HEAT repeats (42) plus an additional flanking sequence highly similar to β-importin, a key component of the nuclear transport machinery.[3] Low levels of full-length huntingtin have been detected in purified nuclei from brain (43) and in the the nucleus of cultured cells (44). Thus,

[3] M. F. Peters and C. A. Ross, unpublished observations.

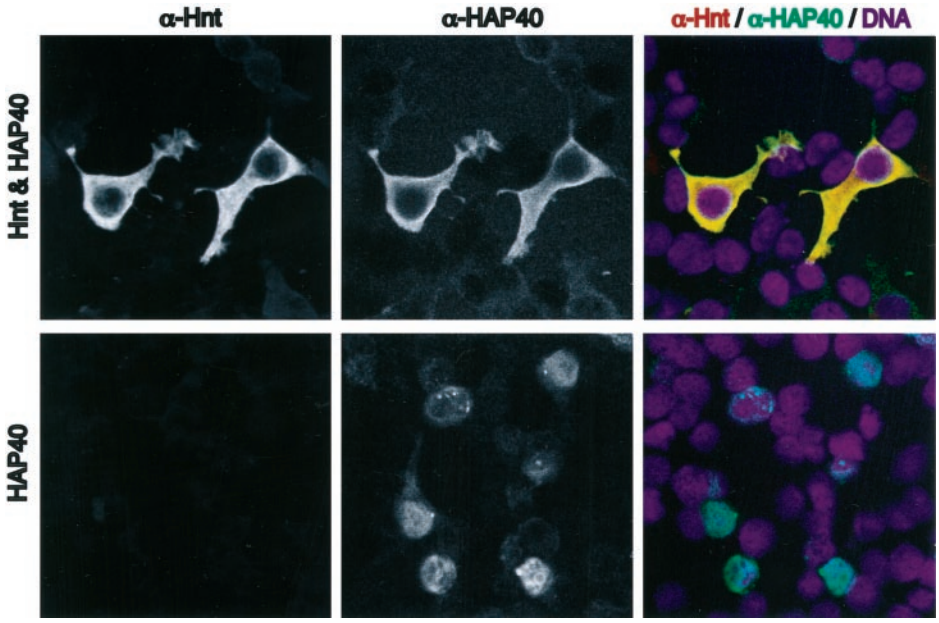

FIG. 6. **Huntingtin redistributes HAP40 from the nucleus to the cytoplasm.** In cotransfected HEK293 cells, full-length huntingtin and HAP40 have indistinguishable cytoplasmic distributions (*top panels*). In cells transfected with HAP40 alone, HAP40 immunostaining was both diffuse nuclear and in nuclear dots (*bottom panels*). Endogenous huntingtin was undetectable. Cells were labeled with rAb Hnt_{1-18} and mAb anti-FLAG. DNA was stained with 4′,6-diamidino-2-phenylindole. Similar results were observed on labeling with mAb Hnt, mAb 2166, and rAb $HAP40_{25-42}$ or $HAP40_{226-243}$ and in transfected HeLa cells.

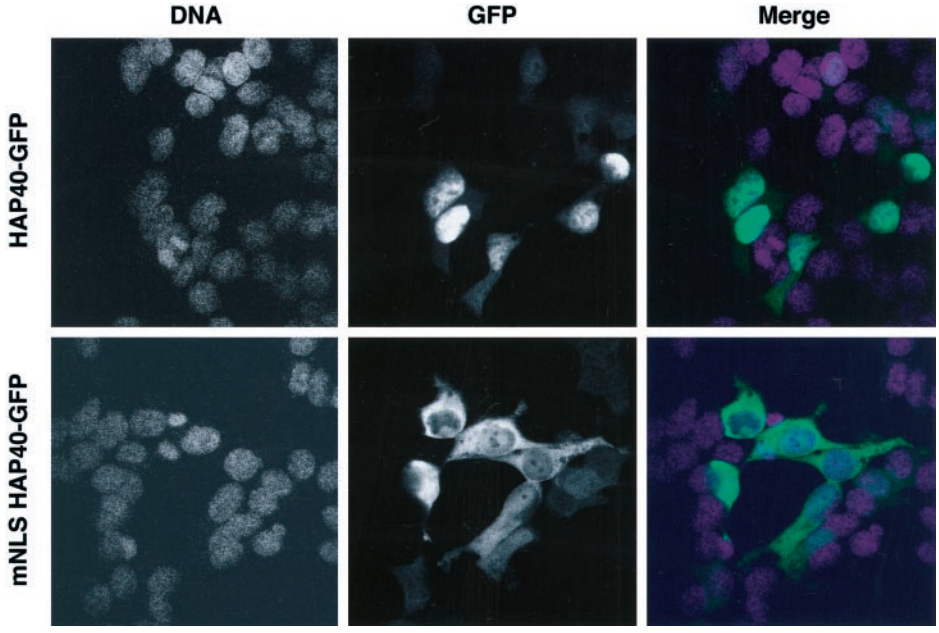

FIG. 7. **Active nuclear targeting of HAP40.** A HAP40-GFP fusion protein too large to passively diffuse through nuclear pores was concentrated in the nucleus (*top panels*). Mutation of a putative NLS in HAP40 clearly reduced but did not eliminate nuclear targeting of HAP40-GFP (*bottom panels*). Thus, HAP40 is actively targeted into the nucleus. DNA was stained with 4′,6-diamidino-2-phenylindole.

under various conditions, huntingtin may import and export HAP40 through nuclear pores.

HAP40 is unlikely to be directly involved in the toxic gain of function caused by huntingtin's expanded polyglutamine. HAP40 does not associate with the N-terminal region of huntingtin that contains the repeat and is not easily detected in huntingtin aggregates of either mice transgenic for an expanded N-terminal fragment or HD post-mortem brain. However, events that regulate nuclear localization of the huntingtin-HAP40 protein complex are candidates for contributing to the selective neuronal vulnerability in HD. In mouse models of HD generated by inserting an expanded polyglutamine stretch into mouse huntingtin, full-length (43) and N-terminal fragments of expanded huntingtin (45) are found in the nucleus. The nuclear redistribution is selective for the striatal neurons most vulnerable in HD. Nuclear targeting exacerbates the toxicity of both expanded huntingtin and ataxin-1 (35, 46, 47). A key role for the nuclear localization is consistent with recent findings that expanded polyglutamine repeats cause toxicity via changes in cAMP-response element-binding protein-binding protein-mediated transcription (30, 48, 49). Investigating the mechanisms that regulate the localization of the normal huntingtin-HAP40 complex may be important to understanding the aberrant nuclear localization and toxicity of expanded huntingtin.

Acknowledgments—We thank Jon Wood for the HntΔex1-glutathione *S*-transferase plasmid, Jon Kushi for affinity-purifying antibodies, Cecile Pickart for providing the HIP2 antibody, Erich Wanker for the SH3GL3 and HIP1 antibodies, Bill Lane for helpful discussions regarding protein sequencing, and Fred Nucifora and Dave Borchelt for comments on the manuscript.

REFERENCES

1. The Huntington's Disease Collaborative Research Group. (1993) *Cell* **72,** 971–983
2. Wells, R. D., and Warren, S. T. (1998) *Genetic Instabilities and Hereditary Neurological Diseases*, Academic Press, San Diego, CA
3. Perutz, M. F. (1999) *Trends Biochem. Sci.* **24,** 58–63
4. Perutz, M. F. (1995) *Mol. Med.* **1,** 718–721
5. Ross, C. A. (1997) *Neuron* **19,** 1147–1150
6. Sharp, A. H., Loev, S. J., Schilling, G., Li, S. H., Li, X. J., Bao, J., Wagster, M. V., Kotzuk, J. A., Steiner, J. P., Lo, A., Hedreen, J., Sisodia, S., Snyder, S. H., Dawson, T. M., Ryugo, D. K., and Ross, C. A. (1995) *Neuron* **14,**

1065–1074
7. Gusella, J. F. (2000) *NeuroScience News* **3,** 15–22
8. DiFiglia, M., Sapp, E., Chase, K., Schwarz, C., Meloni, A., Young, C., Martin, E., Vonsattel, J. P., Carraway, R., Reeves, S. A., Boyce, F. M., and Aronin, N. (1995) *Neuron* **14,** 1075–1081
9. White, J. K., Auerbach, W., Duyao, M. P., Vonsattel, J. P., Gusella, J. F., Joyner, A. L., and MacDonald, M. E. (1997) *Nat. Genet.* **17,** 404–410
10. Duyao, M. P., Auerbach, A. B., Ryan, A., Persichetti, F., Barnes, G. T., McNeil, S. M., Ge, P., Vonsattel, J. P., Gusella, J. F., Joyner, A. L., and MacDonald, M. E. (1995) *Science* **269,** 407–410
11. Nasir, J., Floresco, S. B., O'Kusky, J. R., Diewert, V. M., Richman, J. M., Zeisler, J., Borowski, A., Marth, J. D., Phillips, A. G., and Hayden, M. R. (1995) *Cell* **81,** 811–823
12. Zeitlin, S., Liu, J. P., Chapman, D. L., Papaioannou, V. E., and Efstratiadis, A. (1995) *Nat. Genet.* **11,** 155–163
13. Rigamonti, D., Bauer, J. H., De-Fraja, C., Conti, L., Sipione, S., Sciorati, C., Clementi, E., Hackam, A., Hayden, M. R., Li, Y., Cooper, J. K., Ross, C. A., Govoni, S., Vincenz, C., and Cattaneo, E. (2000) *J. Neurosci.* **20,** 3705–3713
14. Jones, L. (2000) *NeuroScience News* **3,** 55–63
15. Li, X. J., Li, S. H., Sharp, A. H., Nucifora, F. C., Jr., Schilling, G., Lanahan, A., Worley, P., Snyder, S. H., and Ross, C. A. (1995) *Nature* **378,** 398–402
16. Li, S. H., Gutekunst, C. A., Hersch, S. M., and Li, X. J. (1998) *J. Neurosci.* **18,** 1261–1269
17. Colomer, V., Engelender, S., Sharp, A. H., Duan, K., Cooper, J. K., Lanahan, A., Lyford, G., Worley, P., and Ross, C. A. (1997) *Hum. Mol. Genet.* **6,** 1519–1525
18. Engelender, S., Sharp, A. H., Colomer, V., Tokito, M. K., Lanahan, A., Worley, P., Holzbaur, E. L., and Ross, C. A. (1997) *Hum. Mol. Genet.* **6,** 2205–2212
19. Sittler, A., Walter, S., Wedemeyer, N., Hasenbank, R., Scherzinger, E., Eickhoff, H., Bates, G. P., Lehrach, H., and Wanker, E. E. (1998) *Mol. Cell* **2,** 427–436
20. Faber, P. W., Barnes, G. T., Srinidhi, J., Chen, J., Gusella, J. F., and MacDonald, M. E. (1998) *Hum. Mol. Genet.* **7,** 1463–1474
21. Singaraja, R., Hadano, S., Metzler, M., Wellington, C., Leavitt, B., Bannykh, S., Fichter, K., Hackam, A., Gan, L., and Hayden, M. (2000) *Am. J. Hum. Genet.* **65,** A112
22. Boutell, J. M., Thomas, P., Neal, J. W., Weston, V. J., Duce, J., Harper, P. S., and Jones, A. L. (1999) *Hum. Mol. Genet.* **8,** 1647–1655
23. Kalchman, M. A., Koide, H. B., McCutcheon, K., Graham, R. K., Nichol, K., Nishiyama, K., Kazemi-Esfarjani, P., Lynn, F. C., Wellington, C., Metzler, M., Goldberg, Y. P., Kanazawa, I., Gietz, R. D., and Hayden, M. R. (1997) *Nat. Genet.* **16,** 44–53
24. Wanker, E. E., Rovira, C., Scherzinger, E., Hasenbank, R., Walter, S., Tait, D., Colicelli, J., and Lehrach, H. (1997) *Hum. Mol. Genet.* **6,** 487–495
25. Kalchman, M. A., Graham, R. K., Xia, G., Koide, H. B., Hodgson, J. G., Graham, K. C., Goldberg, Y. P., Gietz, R. D., Pickart, C. M., and Hayden, M. R. (1996) *J. Biol. Chem.* **271,** 19385–19394
26. Boutell, J. M., Wood, J. D., Harper, P. S., and Jones, A. L. (1998) *Hum. Mol. Genet.* **7,** 371–378
27. Huang, C. C., Faber, P. W., Persichetti, F., Mittal, V., Vonsattel, J. P., MacDonald, M. E., and Gusella, J. F. (1998) *Somatic Cell Mol. Genet.* **24,** 217–233
28. Bao, J., Sharp, A. H., Wagster, M. V., Becher, M., Schilling, G., Ross, C. A., Dawson, V. L., and Dawson, T. M. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5037–5042
29. Liu, Y. F., Deth, R. C., and Devys, D. (1997) *J. Biol. Chem.* **272,** 8121–8124
30. Steffan, J. S., Kazantsev, A., Spasic-Boskovic, O., Greenwald, M., Zhu, Y. Z., Gohler, H., Wanker, E. E., Bates, G. P., Housman, D. E., and Thompson, L. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6763–6768
31. Kazantsev, A., Preisinger, E., Dranovsky, A., Goldgaber, D., and Housman, D. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 11404–11409
32. Liu, Y. F., Dorow, D., and Marshall, J. (2000) *J. Biol. Chem.* **275,** 19035–19040
33. Peters, M. F., Adams, M. E., and Froehner, S. C. (1997) *J. Cell Biol.* **138,** 81–93
34. Harlow, E., and Lane, D. (1999) *Using Antibodies: a Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY
35. Peters, M. F., Nucifora, F. C., Jr., Kushi, J., Seaman, H. C., Cooper, J. K., Herring, W. J., Dawson, V. L., Dawson, T. M., and Ross, C. A. (1999) *Mol. Cell. Neurosci.* **14,** 121–128
36. Schilling, G., Becher, M. W., Sharp, A. H., Jinnah, H. A., Duan, K., Kotzuk, J. A., Slunt, H. H., Ratovitski, T., Cooper, J. K., Jenkins, N. A., Copeland, N. G., Price, D. L., Ross, C. A., and Borchelt, D. R. (1999) *Hum. Mol. Genet.* **8,** 397–407
37. Li, H., Li, S. H., Cheng, A. L., Mangiarini, L., Bates, G. P., and Li, X. J. (1999) *Hum. Mol. Genet.* **8,** 1227–1236
38. Wood, J. D., MacMillan, J. C., Harper, P. S., Lowenstein, P. R., and Jones, A. L. (1996) *Hum. Mol. Genet.* **5,** 481–487
39. Levinson, B., Bermingham, J. R., Jr., Metzenberg, A., Kenwrick, S., Chapman, V., and Gitschier, J. (1992) *Genomics* **13,** 862–865
40. Naylor, J. A., Buck, D., Green, P., Williamson, H., Bentley, D., and Giannelli, F. (1995) *Hum. Mol. Genet.* **4,** 1217–1224
41. Bessert, D. A., Gutridge, K. L., Dunbar, J. C., and Carlock, L. R. (1995) *Brain Res. Mol. Brain Res.* **33,** 165–173
42. Andrade, M. A., and Bork, P. (1995) *Nat. Genet.* **11,** 115–116
43. Wheeler, V. C., White, J. K., Gutekunst, C. A., Vrbanac, V., Weaver, M., Li, X. J., Li, S. H., Yi, H., Vonsattel, J. P., Gusella, J. F., Hersch, S., Auerbach, W., Joyner, A. L., and MacDonald, M. E. (2000) *Hum. Mol. Genet.* **9,** 503–513
44. De Rooij, K. E., Dorsman, J. C., Smoor, M. A., Den Dunnen, J. T., and Van Ommen, G. J. (1996) *Hum. Mol. Genet.* **5,** 1093–1099
45. Li, H., Li, S. H., Johnston, H., Shelbourne, P. F., and Li, X. J. (2000) *Nat. Genet.* **4,** 385–389
46. Klement, I. A., Skinner, P. J., Kaytor, M. D., Yi, H., Hersch, S. M., Clark, H. B., Zoghbi, H. Y., and Orr, H. T. (1998) *Cell* **95,** 41–53
47. Saudou, F., Finkbeiner, S., Devys, D., and Greenberg, M. E. (1998) *Cell* **95,** 55–66
48. McCampbell, A., Taylor, J. P., Taye, A. A., Robitschek, J., Li, M., Walcott, J., Merry, D., Paulson, H., Chai, Y., Sobue, G., and Fischbeck, K. H. (2000) *Hum. Mol. Genet.* **9,** 2197–2202
49. Nucifora, F. C., Sasaki, M., Huang, H., Troncoso, J., Dawson, V. L., Dawson, T. M., and Ross, C. A. (2000) *Soc. Neurosci. Abstr.* **26,** 1030s