Neuron, Vol. 15, 493–496, September, 1995, Copyright © 1995 by Cell Press

When More Is Less: Pathogenesis of Glutamine Repeat Neurodegenerative Diseases

Minireview

Christopher A. Ross
Laboratory of Molecular Neurobiology
Departments of Psychiatry and Neuroscience
Program in Cellular and Molecular Medicine
Johns Hopkins Medical School
Baltimore, Maryland 21205

Five neurodegenerative diseases are caused by expanding CAG triplet repeats coding for polyglutamine. They all involve loss of selected populations of neurons, usually in adulthood. How the genetic lesions cause progressive neuronal degeneration is unknown. Unraveling the pathogenesis of these disorders will help elucidate pathways of neuronal cell death.

Triplet repeat disorders involve the recently identified phenomenon of unstable DNA, in which repeating units of 3 nucleotides such as CAG expand during vertical transmission and cause disease (La Spada et al., 1994). They can be divided into two types. The first type (Figure 1) involves expanding CAG repeats coding for polyglutamine. The normal length of the repeat is from 6 or 7 to about 30, and the disease is present when the repeat expands to 35–40 or more. Except for the androgen receptor of spinal and bulbar muscular atrophy (SBMA), none of the causative genes has a known function or substantial homology with any known gene in the database.

In the second type of disease (Table 1), triplet repeats such as CTG or CCG are present in mRNA outside the open reading frame (La Spada et al., 1994; Ross et al., 1993).

In all of the glutamine repeat diseases, there is loss of neurons with gliosis, but without deposition of extracellular material or intracellular inclusion bodies. Though the expression patterns of the causative genes are widespread, the disorders involve loss of neurons within specific regions of the brain and spinal cord (Figure 1). It is possible that the glutamine repeat diseases share similar mechanisms of pathogenesis.

Glutamine Repeat Diseases and Causative Genes

Huntington's disease (HD) is the most common glutamine repeat disease. It is an autosomal dominant neurodegenerative disorder characterized by abnormal movements including chorea, dementia, and emotional disorders. The pathology is notable for selective neuronal vulnerability of medium spiny projection neurons in the caudate and putamen. After a long search, the HD gene was identified and found to have an expanding CAG repeat (Huntington's Disease Collaborative Research Group, 1993). The mRNA contains a long open reading frame coding for a predicted protein of 348 kDa. Immediately adjacent to the glutamine repeat, there are two proline repeats interrupted by a proline-rich region. The protein product of the HD gene (huntingtin), like the mRNA, has a widespread expression in both the brain and the periphery (Sharp et al., 1995; Trottier et al., 1995). Within the brain, it is predominantly neuronal with no enrichment in the basal ganglia. It is present in cell bodies and dendrites and appears enriched in neuronal terminals, possibly associated with vessicles or microtubules (DiFiglia et al., 1995; Gutekunst et al., 1995).

Dentato-Rubral and Pallido-Luysian Atrophy (DRPLA or Smith's Disease) is a rare neurodegenerative disorder whose clinical features for adult onset cases are similar to those of HD. DRPLA was the first triplet repeat disease–causing gene (*CTG-B37* or *atrophin-1*) to be cloned by screening for triplet repeats directly from cDNA libraries (Li et al., 1993). The predicted 120 kDa protein has a glutamine repeat, a proline repeat, and a region of alternating acidic and basic amino acids (Nagafuchi et al., 1994; Margolis et al., 1995). The protein product of the DRPLA gene has recently been identified and migrates at 190 kDa, slower than expected (Yazawa et al., 1995). Immunocytochemical data suggest a cytoplasmic distribution densest in the perinuclear region but also present in proximal dendrites of neurons.

Spinal cerebellar ataxia type 3 (SCA3 or Machado–Joseph Disease [MJD]) is a member of a class of diseases with related phenotypes and overlapping regions of neuronal degeneration. MJD presents with cerebellar ataxia, dystonia or rigidity, ophthalmoplegia, and sometimes bulging eyes. The causative gene for SCA3 was also identified by direct screening of human brain cDNA libraries for triplet repeats (Kawaguchi et al., 1994). The gene is widely expressed; its protein product has not yet been identified.

Spinal cerebellar ataxia type 1 (SCA1) is a progressive neurodegenerative disorder characterized by ataxia, dysarthria, ophthalmoparesis, and bulbar motor weakness. The gene for SCA1, termed *ataxin-1*, was identified by positional cloning and screening of YACs for triplet repeats (Banfi et al., 1994). The ataxin-1 protein has also recently been identified using antibodies (Servadio et al., 1995). Immunocytochemical data suggest both cytoplasmic and nuclear localization with a widespread distribution.

SBMA is a rare late-onset form of motor neuron degeneration. Patients often have signs of androgen insensitivity, including gynecomastia and testicular atrophy. The affected gene product, the androgen receptor (La Spada et al., 1991), is a member of a class of ligand-binding intracellular steroid receptors that can translocate to the nucleus to activate gene transcription via a zinc finger DNA-binding domain. While the expression of the androgen receptor mRNA is rather widespread, high levels of functional androgen-binding activity appear to be more restricted, including in spinal motor neurons.

Glutamine Repeats

The functions of glutamine repeats in proteins are unknown. CAG repeats were originally identified in Drosophila, where they are found in many genes involved in regulation of development and neurogenesis. Glutamine repeats are far more common than repeats of any other amino acid (Green and Wang, 1994). However, the length of the repeats is poorly conserved in homologous genes from different species. For example, the mouse HD gene encodes only seven repeats. The glutamine repeat in the rat

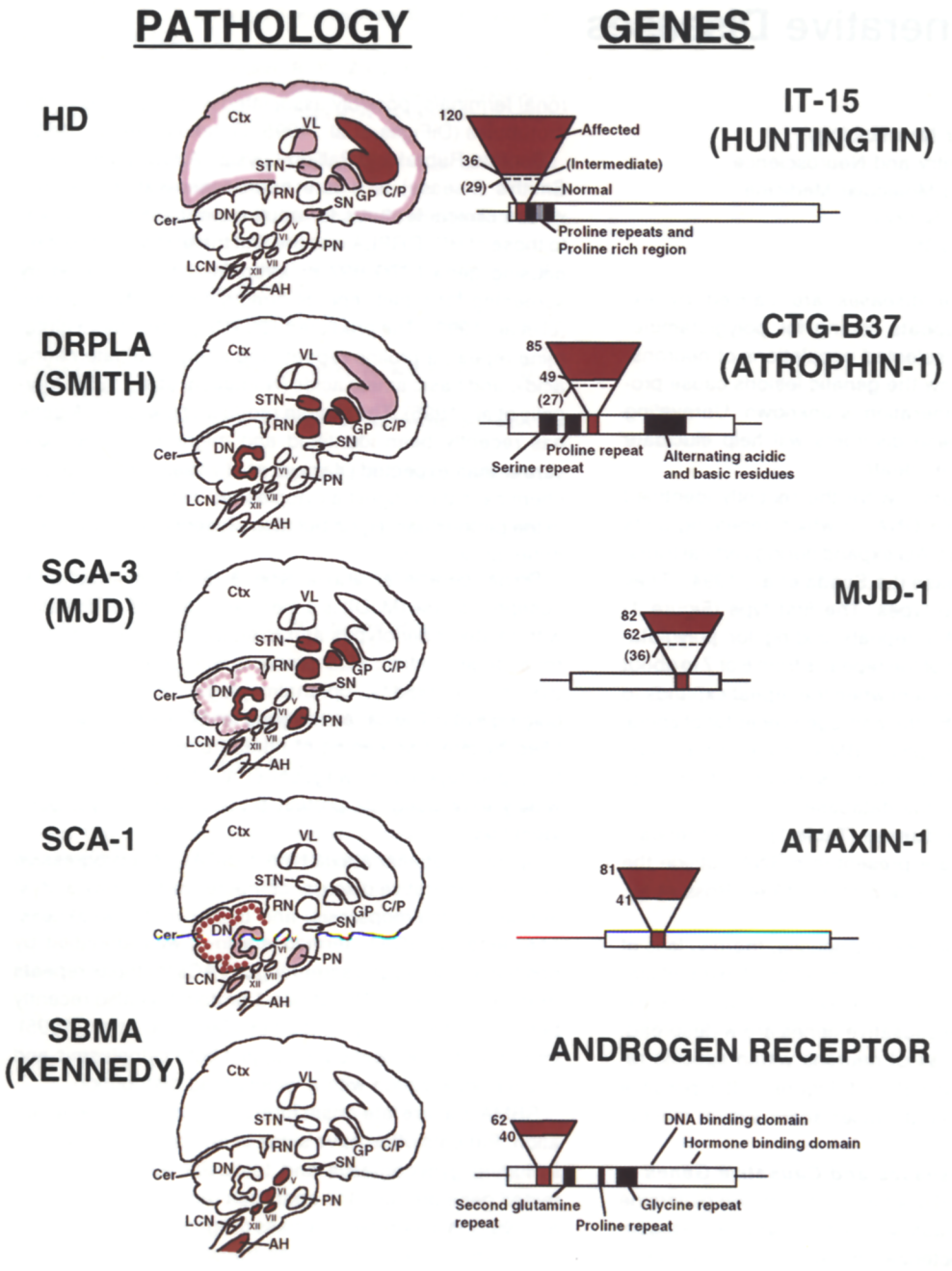

Figure 1. Glutamine Repeat Diseases

The left side shows a schematic diagram of the major sites of neuronal loss in each disease. Dark red indicates severe or selective neuronal loss; half-tone red indicates moderate or variable cell loss. The circles in the cerebellar cortex represent Purkinje cells. The right side shows a schematic of the open reading frame of each causative gene. The glutamine repeat expansion is shown in red, with the approximate number of repeats in controls and in patients. AH, anterior horn; Cer, cerebellar cortex; C/P, caudate/putamen; Ctx, cerebral cortex; DN, dentate nucleus; GP, globus pallidus; LCN, lateral cuneate nucleus; PN, pontine nucleus; RN, red nucleus; SN, substantia nigra; STN, subthalamic nucleus; VL, ventrolateral thalamic nucleus; V, VI, VII, and XII, cranial motor nuclei.

homolog of *atrophin-1* has only 5 glutamines followed by glutamine alternating with proline 4 times.

Glutamine-rich regions have been described in the factor interaction domain of several transcription factors. Although it is unclear whether glutamine repeats generally function like glutamine-rich regions, the expanding glutamine repeat in the androgen receptor is in a region thought to be important for interaction with other cell type–specific transcription factors (Adler et al., 1992). Lengthening the repeat may alter transactivation. Thus, glutamine repeats may function as protein–protein interaction motifs.

Disease Pathogenesis

How do these proteins cause disease? Is it due to loss of a protein's normal function or to a gain of function? The

Table 1. Type 2 Triplet Repeat Diseases

Disease	Gene	Expanded Triplet	Translation
Fragile X syndrome, type A (FraX-A)	*FMR1*	CCG × >200	5′-Untranslated
Fragile X syndrome, type E (FraX-E)	?	CCG × >200	? Untranslated
Myotonic dystrophy or dystrophia myotonica (DM)	Myotonin protein kinase (DMK)	CTG × >200	3′-Untranslated
Jacobsen syndrome with Fra11B	*CBL2* proto-oncogene	CCG	5′-Untranslated

recent development of an HD gene knockout (Duyao et al., 1995; Nasir et al., 1995; Zeitlin et al., 1995) provides an approach to determine this gene's normal function. Mice homozygous for the disruption die before embryonic day 8.5, indicating that the gene product has some vital cellular function. In one model, animals heterozygous for the disruption may express a 20 kDa truncated N-terminal protein from the allele with the knockout as well as one copy of the normal allele (Nasir et al., 1995) and appear normal with no overt nervous system dysfunction. However, on specific behavioral tests they show abnormalities of cognition and motor activity similar to those shown by animals with lesions of the striatum. In addition, quantitative analysis of two of the animals suggests decreased size of the subthalamic nucleus and fewer neurons in the globus pallidus. If not due to strain differences, these results would suggest that the HD gene is important for development or maintenance of the structure of the basal ganglia. Previous magnetic resonance imaging studies of the basal ganglia in presymptomatic humans with the HD expansion have indicated decreased volumes (Aylward et al., 1994). The other knockout model, which does not express the 20 kDa N-terminal truncated protein, appears to show no abnormalities in the heterozygous animals (Duyao et al., 1995).

Several considerations suggest that the pathogenesis of these diseases involves a gain of function. Most of the disorders have dominant inheritance, which is usually associated with gain of function. Patients with loss of a single allele of the HD gene do not appear to have progressive loss of neurons (Ambrose et al., 1994). Patients with SBMA show some signs of loss of androgen receptor activity; androgen receptors with expanded glutamine repeats have decreased activity (Kazemi-Esfarjani et al., 1995). However, SBMA cannot be due simply to loss of functional receptor. Patients with deletions of the androgen receptor, or point mutations that inactivate it, have a different phenotype—that of pure androgen insensitivity. The dissimilarity between effects of loss of function of the relevant genes in HD and SBMA and the neuronal degeneration of the diseases argues against a dominant negative model. Finally, in all the disorders studied so far (HD, SCA1, and DRPLA), the allele with the expanded repeat is present in mRNA and protein in brains of patients with the disease (Persichetti et al., 1995; Servadio et al., 1995; Sharp et al., 1995; Trottier et al., 1995; Yazawa et al., 1995).

A transgenic animal model related to SCA1 has recently been developed (Burright et al., 1995). The human *SCA1* cDNA with an expanded repeat was inserted into the mouse genome under the control of a Purkinje cell–specific promoter, yielding very high expression specifically in Purkinje cells of the cerebellum. Transgenic animals from five of six lines developed ataxia and Purkinje cell degeneration similar to the human disorder. Curiously, while the mRNA for the transgene could be detected, the protein expressed by the transgene could not. It is possible that the abnormal protein is unusually labile. It is unlikely that expression of the mRNA alone is responsible for the pathology since, in a similar experiment, animals transgenic for the HD gene with an expanded repeat were created in which the transgene inadvertently contained a frameshift mutation. These animals expressed mRNA but not protein and did not have neuronal degeneration. Thus, these data together appear most consistent with the hypothesis of a gain of function mechanism.

Protein–Protein Interactions

What might this gain of function be? As noted above, glutamine repeats may be involved in protein–protein interactions. There might be a stronger association with a protein with which the gene product normally associates. Alternatively, the gain of function might involve some novel interaction.

Several hypotheses regarding protein–protein interactions have recently been proposed. Green (1993) has suggested that proteins with expanding glutamine repeats might become substrates for transglutaminase activity, which could result in cross-linked products involving an ε–γ glutamyl lysine isopeptide. The cross-linked proteins would presumably be degraded by proteolysis, but the residual isodipeptide could not be degraded and might conceivably have toxic effects within the cells.

Another suggestion involves noncovalent protein interactions. Perutz (1994; Stott et al., 1995) has suggested that glutamine repeats within proteins can form a "polar zipper" involving a β sheet held together by hydrogen bonds between the main chain and the side chain amides. These could form between two proteins with glutamine repeats or by self-bonding within one glutamine repeat–containing protein. Presumably, lengthening the glutamine repeat would increase the stability of the association.

Since both of these models involve direct interactions of the glutamine repeats themselves, these mechanisms could be common for all the diseases. The differences in distribution of pathology might be explained by differences in the localization of each of the proteins to various neuronal populations.

Other mechanisms might be specific to the gene products for individual diseases and to specific cell types. For instance, the HD gene product with the expanded glutamine repeat present in brain appears to have altered migration in a gel, compared with protein from nonneuronal cells (Schilling et al., 1995; Trottier et al., 1995). The proteins with long expanded repeats appear to migrate more slowly than expected, in a broad, diffuse band of lower

intensity than the band representing the normal allele. Regions of the brain affected in the illness, such as striatum and cortex, show this change more prominently than other regions of the brain. This could be due to an alteration of the protein; however, an effect due to somatic mosaicism (Telenius et al., 1994) also cannot be ruled out.

Finally, it is possible that the cell type specificity of the pathology of each disease relates to cell type–specific protein–protein interactions. For instance, a putative HAP1 protein associates with the huntingtin protein in a glutamine repeat length–dependent manner and has brain-selective expression.

Relationship to Cell Death Pathways

Excitotoxicity has long been hypothesized to be a pathogenic mechanism in HD, based on the similarity of the pattern of cell loss in HD and in excitotoxic lesions of the striatum (Coyle and Puttfarcken, 1993). An elaboration on this model suggests that metabolically compromised neurons might be more sensitive to excitotoxicity or other forms of toxicity (Albin and Greenamyre, 1992; Beal, 1992). Oxidative stress may be involved in several forms of cell death (Coyle and Puttfarcken, 1993), either as a pathogenic mechanism or as a signaling pathway to activate other cell death pathways. Finally, neurons may die via apoptosis or programmed cell death, which might activate a genetic program involved in the normal death of certain neurons during development.

Different neurons in the brain appear to have different thresholds for various cell death mechanisms. Thus, part of the selective vulnerability of neurons in the brain might reflect intrinsic vulnerability. The results of the SCA1 transgenic model experiments suggest that cell-autonomous mechanisms can be sufficient, though pathways involving cell–cell interactions have also been hypothesized (Burright et al., 1995; Sharp et al., 1995; Trottier et al., 1995). Unraveling the pathogenesis of these diseases will provide opportunities to explore pathways leading to neuronal cell death as the issues move from genetics to biochemistry and neuronal cell biology.

References

Adler, A. J., Danielsen, M., and Robins, D. M. (1992). Proc. Natl. Acad. Sci. USA *89*, 11660–11663.

Albin, R. L., and Greenamyre, J. T. (1992). Neurology *42*, 733–738.

Ambrose, C. M., Duyao, M. P., Barnes, G., Bates, G. P., Lin, C., Srinidhi, J., Baxendale, S., Hummerich, H., Lehrach, H., Altherr, M., et al. (1994). Somat. Cell Mol. Genet. *20*, 27–38.

Aylward, E. H., Brandt, J., Codori, A. M., Mangus, R. S., Barta, P. E., and Harris, G. J. (1994). Neurology *44*, 823–828.

Banfi, S., Servadio, A., Chung, M.-Y., Kwiatkowski, T. J., Jr., McCall, A. E., Duvick, L. A., Shen, Y., Roth, E. J., Orr, H. T., and Zoghbi, H. Y. (1994). Nature Genet. *7*, 513–520.

Beal, M. F. (1992). Ann. Neurol. *31*, 119–130.

Burright, E. N., Clark, H. B., Servadio, A., Matilla, T., Feddersen, R. M., Yunis, W. S., Duvick, L. A., Zoghbi, H. Y., and Orr, H. T. (1995). Cell *82*, in press.

Coyle, J. T., and Puttfarcken, P. (1993). Science *262*, 689–695.

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.

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.

Green, H. (1993). Cell *74*, 955–956.

Green, H., and Wang, N. (1994). Proc. Natl. Acad. Sci. USA *91*, 4298–4302.

Gutekunst, C.-A., Levey, A. I., Heilman, C. J., Whaley, W. L., Hong, Y., Nash, N. R., Rees, H. D., Madden, J. J., and Hersch, S. M. (1995). Proc. Natl. Acad. Sci. USA, in press.

Huntington's Disease Collaborative Research Group (1993). Cell *72*, 971–983.

Kazemi-Esfarjani, P., Trifiro, M. A., and Pinsky, L. (1995). Hum. Mol. Genet. *4*, 523–527.

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I., et al. (1994). Nature Genet. *8*, 221–227.

La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E., and Fishbeck, K. (1991). Nature *352*, 77–79.

La Spada, A. R., Paulson, H. L., and Fischbeck, K. H. (1994). Ann. Neurol. *36*, 814–822.

Li, S.-H., McInnis, M. G., Margolis, R. L., Antonarakis, S. E., and Ross, C. A. (1993). Genomics *16*, 572–579.

Margolis, R. L., Li, S. H., Young, W. S., Wagster, M. V., Stine, O. C., Kidwai, A. S., Ashworth, R. G., and Ross, C. A. (1995). Mol. Brain Res., in press.

Nagafuchi, S., Yanagisawa, H., Ohsaki, E., Shirayama, T., Tadokoro, K., Inoue, T., and Yamada, M. (1994). Nature Genet. *8*, 177–182.

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.

Persichetti, F., Ambrose, C. M., Ge, P., McNeil, S. M., Srinidhi, J., Anderson, M. A., Jenkins, B., Barnes, G. T., Duyao, M. P., Kanaley, L., et al. (1995). Mol. Med. *1*, 374–383.

Perutz, M. (1994). Protein Sci. *3*, 1629–1637.

Ross, C. A., McInnis, M. G., Margolis, R. L., and Li, S. H. (1993). Trends Neurosci. *16*, 254–260.

Schilling, G., Sharp, A. H., Loev, S. J., Wagster, M. V., Li, S.-H., Stine, O. C., and Ross, C. A. (1995). Hum. Mol. Genet. *4*, 1365–1371.

Servadio, A., Koshy, B., Armstrong, D., Antalffy, B., Orr, H. T., and Zoghbi, H. Y. (1995). Nature Genet. *10*, 94–98.

Sharp, A. H., Loev, S. J., Schilling, G., Steiner, J. P., Li, S.-H., Li, X.-J., Bao, J., Wagster, M. V., Kotzuk, J. A., Steiner, J. P., et al. (1995). Neuron *14*, 1065–1074.

Stott, K., Blackburn, J. M., Butler, P. J. G., and Perutz, M. (1995). Proc. Natl. Acad. Sci. USA *92*, 6509–6513.

Telenius, H., Kremer, B., Goldberg, Y. P., Theilmann, J., Andrew, S. E., Zeisler, J., Adam, S., Greenberg, C., Ives, E. J., Clarke, L. A., and Hayden, M. R. (1994). Nature Genet. *6*, 409–414.

Trottier, Y., Devys, D., Imbert, G., Saudou, F., An, I., Lutz, Y., Weber, C., Agid, Y., Hirsch, E. C., and Mandel, J.-L. (1995). Nature Genet. *10*, 104–110.

Yazawa, I., Nukina, N., Hashida, H., Goto, J., Yamada, M., and Kanazawa, I. (1995). Nature Genet. *10*, 99–103.

Zeitlin, S., Liu, J.-P., Chapman, D. L., Papaioannou, V. E., and Efstratiatis, A. (1995). Nature Genet. *10*, in press.