Local Protein Unfolding and Pathogenesis of Polyglutamine-Expansion Diseases

Yu Wai Chen*
Centre for Protein Engineering and Cambridge University Chemical Laboratory, Cambridge, United Kingdom

ABSTRACT **Polyglutamine expansion diseases are caused by the abnormal lengthening of a glutamine repeat in the respective protein and are believed to progress via a common mechanism. Here a hypothesis based on protein unfolding is formulated to enhance understanding of the pathogenesis of these diseases. Nine disease proteins of this family were investigated with an unfolding prediction protocol based on calculating the net charge and the overall hydrophobicity from the primary sequence. These protein sequences were analysed with and without their respective nonpathogenic glutamine repeats. When full-length protein sequences were studied, the calculations suggest that unexpanded polyglutamine repeats have a minimal effect on the global folding state of the parent proteins. When confined to a local context approximating the size of a single-protein domain, the calculations suggest that all nine sequences including a normal-length polyglutamine repeat correspond to the unfolded state. The unfolding predictions were applied to two model systems and were found to be consistent with published experimental data. Despite its simplicity, this analysis can be extended to simulate repeat length-dependent protein unfolding, an event speculated to precede aggregation and disease progression. This work highlights the correlation between glutamine repeats and protein unfolding and provides insight into the as yet unknown polyglutamine expansion pathogenic mechanism. Proteins 2003;51:68–73.** © 2003 Wiley-Liss, Inc.

Key words: CAG trinucleotide repeat; spinocerebellar ataxia; ataxin; Huntington's disease; huntingtin; neurodegeneration

INTRODUCTION

In the last decade, a growing number of inherited neurodegenerative diseases, including Huntington's disease and several forms of spinocerebellar ataxia have been shown to be caused by abnormal lengthening of an otherwise harmless polyglutamine (polyQ) repeat sequence inside the respective proteins.[1–3] These diseases share several common features and appear to progress via a similar pathogenic mechanism. Expansion of the glutamine repeat beyond a threshold of 35–40 residues confers on the disease protein a toxic gain of function that causes tissue-specific neuronal death. Apart from the polyQ stretch, these proteins and their genes have nothing in common, suggesting that the pathogenesis may be related to some global intrinsic properties of the respective proteins that promote misfolding and aggregation.

It has been known for some time that some proteins exhibit a large extent of disorder in physiological conditions and that intrinsic disorder confers functional advantages to these proteins.[4–6] It is interesting that some of these intrinsically unstructured proteins (e.g., α-synuclein) are implicated in neurodegenerative diseases. In this work, the author investigated the hypothesis that polyQ disease proteins may share a common feature of being "natively unfolded" and are thus susceptible to misfolding and aggregation. Uversky and colleagues[7] established that proteins that are "natively unfolded" are characterized by high net charge and low hydrophobicity. They found that for a given protein with a mean net charge of $\langle R \rangle$, a boundary value exists for its mean hydrophobicity, $\langle H \rangle$, below which the protein is likely to be natively unfolded. This empirical relationship is expressed as:

$$\langle H \rangle = \frac{\langle R \rangle + 1.151}{2.785}$$

From this equation, a critical value of $\langle H \rangle = 0.413$ can be derived, which is the minimal hydrophobicity required for a protein to remain folded, when it carries no net charge. Because both $\langle H \rangle$ and $\langle R \rangle$ can be calculated from a protein's amino acid sequence, this condition represents a quick assessment for the folding state of a protein. Here this analysis was applied to studying the pathogenic mechanism of polyQ expansion disease proteins.

MATERIALS AND METHODS

Protein Sequences

The amino acid sequences of the proteins were obtained from the SWISS-PROT database,[8] namely, Huntington's disease protein (or huntingtin, Htt, accession number P42858), androgen receptor (AR, P10275), atrophin-1 (Atf1, also known as dentatorubral-pallidoluysian atrophy protein, P54259), spinocerebellar ataxin-1 (Atx1, P54253), ataxin-2 (Atx2, Q99700), ataxin-3 (Atx3, also known as Machado-Joseph disease protein, MJD, P54252), ataxin-6

Grant sponsor: Wellcome Trust; Grant number: 061836.

*Correspondence to: Yu Wai Chen, Centre for Protein Engineering and Cambridge University Chemical Laboratory, MRC Centre, Hills Road, Cambridge CB2 2QH, UK. E-mail: ywc@cantab.net

Received 12 September 2002; Accepted 17 October 2002

(Atx6, O00555), ataxin-7 (Atx7, O15265), and sperm whale myoglobin (Mb, P02185). The TATA-box-binding protein (TBP or TF2D, referred to as Atx17 here, P20226) is recently implicated in ataxia 17[9] and is also included.

Folding/Unfolding Prediction

Folding/unfolding predictions were based on a previous study[6,7] and were performed with the ExPASy server.[10] ProtParam was used for charge calculations. The mean net charge, $\langle R \rangle$, is the absolute value of the net charge divided by the number of residues. Hydrophobicity was calculated with the Kyte and Doolittle protocol[11] implemented in ProtScale, using a segment size of five residues and normalized to a value between 0 and 1. The mean hydrophobicity, $\langle H \rangle$, is the arithmetic average over the range of residues being studied.

For each protein, the global values of mean net charge $\langle R \rangle$ and mean hydrophobicity $\langle H \rangle$ were calculated for the intact sequence. To study the local environment embedding the polyQ repeat, local $\langle R \rangle$ and $\langle H \rangle$ values were calculated for fragments of the respective proteins. For Htt, AR, Atx3, and Atx17, there are weak evidences suggesting the natural boundaries of the local domain that harbors the polyQ repeat (see later), and these domains were used in the calculations. For the remaining proteins, there is no information concerning what constitutes a local environment. Therefore, calculations were performed for two or three fragments of each protein, strategically chosen so that the polyQ resides near the ends or in the middle of the fragment. The protein fragments contain 100 neighboring residues in addition to the polyQ. A fragment size of 100 residues was used for several reasons. First, it is a good approximation of protein domain or subunit sizes.[12] Second, the original study of Uversky et al.[7] used monomeric globular proteins of 50–200 residues as reference for the natively folded set.[7] By using a size in this range, the results can be compared directly with theirs. Third, most "natively unfolded" proteins are 50–100 residues in length.[7]

For prediction validation using Atx3, data for the maltose-binding protein fusions (MBP-Atx3-Q_{27} and MBP-Atx3-Q_{78}) and hexahistidine-tagged ataxin-3 (H6-Atx3-Q_{27}) were calculated by using the respective translated amino acid sequence corresponding to these fusion constructs.[13,14] The polyQ insertion mutant sequences of myoglobin are constructed as described.[15] For simulation of polyQ expansion, the calculations were performed with models having the glutamine repeat replaced by a stepwise increasing number (0, 1, 2, … up to 80) of glutamine residues in the respective sequences. For the Atx3 model system, the homogenous 22Q-repeat was replaced and for AR, the sequence Q_{21}ETSPRQ_6 was replaced with Q_0 up to Q_{80}.

Approximation of Local Domain for Htt, AR, Atx3, and Atx17

Htt has a protease-sensitive region near the N-terminus producing fragments of 20–50 kDa in size.[16] With this knowledge, the first 180 residues (corresponds to ca. 20 kDa) are approximated to be the local domain harboring the polyQ. The polyQ repeat of Sca17 (or TBP) is found in the variable N-terminal domain of 158 amino acids.[9] In AR, the polyQ is found in the N-terminal domain of ~500 residues. A recent study identified a protease-sensitive region at residues 170–180 and may indicate a natural domain boundary.[17] In Atx3, comparison of homologues among different species revealed a conserved N-terminal half (Josephin domain, residues 1–198) and a more variable C-terminal half (residues 199–360) that contains the polyQ (Pfam server[18]). For this study, the C-terminal 162 residues (Atx3:199–360) is considered to be a natural domain. In these three cases, the fragments represent reasonable domain sizes, and there is no evidence that these can be further divided into subdomains.

RESULTS AND DISCUSSION

Some PolyQ Disease Proteins Are Globally "Natively Unfolded"

First, the hypothesis that all polyQ disease proteins have an intrinsically unstructured context that makes them susceptible to misfolding were investigated. With Atx3 as a model, when the homogenous 22Q-repeat were excluded from the sequence, the protein context has values of $\langle H \rangle = 0.438$ and $\langle R \rangle = 0.059$ and is classified marginally as a properly folded protein [Fig. 1(a)]. If the calculation is performed on the sequence of Atx3 with 22Q (this corresponds to a real nonpathogenic Atx3 protein sequence from a normal individual), it has values of $\langle H \rangle$ and $\langle R \rangle$ that characterize it to be natively unfolded [Table I and Fig. 1(a)]. Apparently, the inclusion of a polyQ can render the parent protein unfolded.

The calculations were extended to other eight members of the polyQ disease family to see if this repeat-related unfolding is universally observed. The results showed that inclusion of the polyQ has a minimal consequence on the global folding state of these proteins [Table I and Fig. 1(a)]. Three proteins, Atf1, Atx2, and Atx7, have a protein context that is characteristic of the "natively unfolded" state, whereas the rest are not. Therefore, the global protein context being intrinsically unfolded is not generally observed for polyQ expansion disease proteins.

All PolyQ Disease Proteins Are Locally Unfolded

With the exception of two smallest members, Atx3 and Atx17, proteins of this family have sizes ranging from 800 to >3000 residues and are likely to consist of multiple domains. It is necessary to dissect out the local context in the vicinity of the polyQ repeat for studying. The results of local context calculations are shown in Table I and Figure 1(b). In about half of the cases, the local protein contexts are predicted to be unfolded. However, in all cases, when a polyQ repeat of normal (nonpathogenic) length is included, the local environment of the respective protein is predicted to be unfolded. Although the effect of polyQ on the global protein context appears to be minimal, in the cases of Htt, AR, Atx1, Atx2, and Atx17, the calculations suggest that the presence of a polyQ correlates to local unfolding.

These predictions are not surprising because glutamine has a very low hydrophobicity of 0.111 according to the Kyte and Doolittle hydropathy scale.[11] Inclusion of a

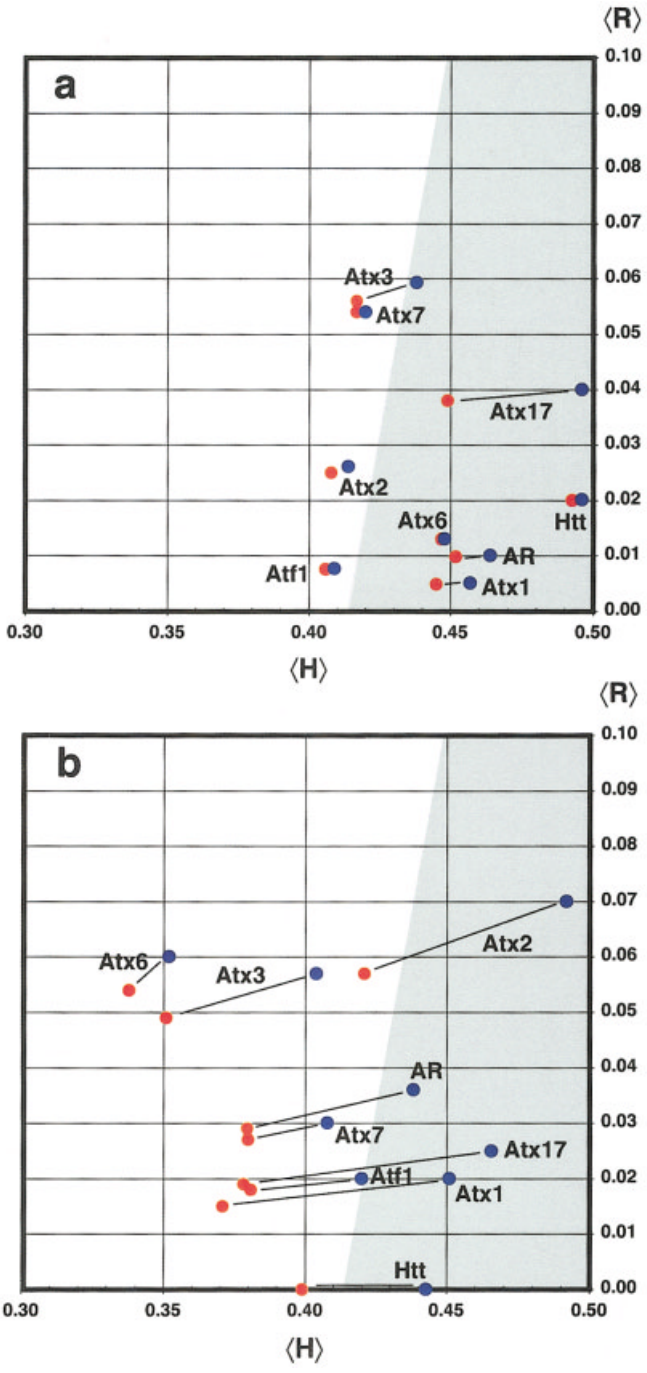

Fig. 1. Prediction of folding/unfolding for polyQ expansion disease protein contexts. The results are shown as blue dots (excluding polyQ) and red dots (including the polyQ) for (**a**) global protein contexts and for (**b**) local protein contexts. Multiple calculations of local domains of individual proteins mostly agree (Table I), and only one representative set is shown for each protein. A data point falling in the shaded area predicts that the protein is folded; otherwise, the protein is unfolded.

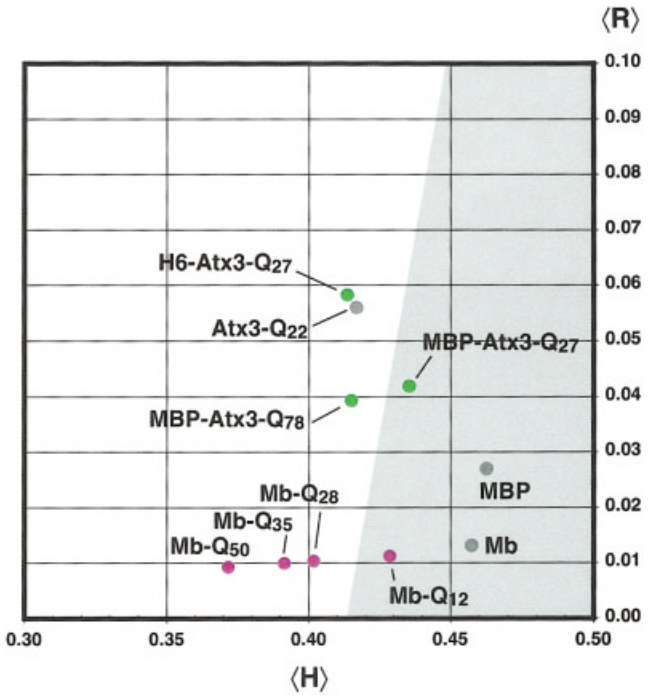

Fig. 2. Folding/unfolding prediction of ataxin-3 and myoglobin constructs. The sequences of H6-tagged and MBP-tagged fusion constructs of Atx3 were used in folding predictions and shown as green dots. The predictions of a real sequence of Atx3 with 22Q and of MBP are shown as gray dots for references. Note that the prediction of MBP—Atx3—Q_{27} is approximately located at the midpoint between those of MBP and H6—Atx3—Q_{27}; this finding illustrates graphically how protein production with a soluble partner can render the whole fusion more soluble. The predictions corresponding to polyQ insertion mutants of myoglobin are shown as purple dots. The prediction of native myoglobin sequence is shown as a gray dot.

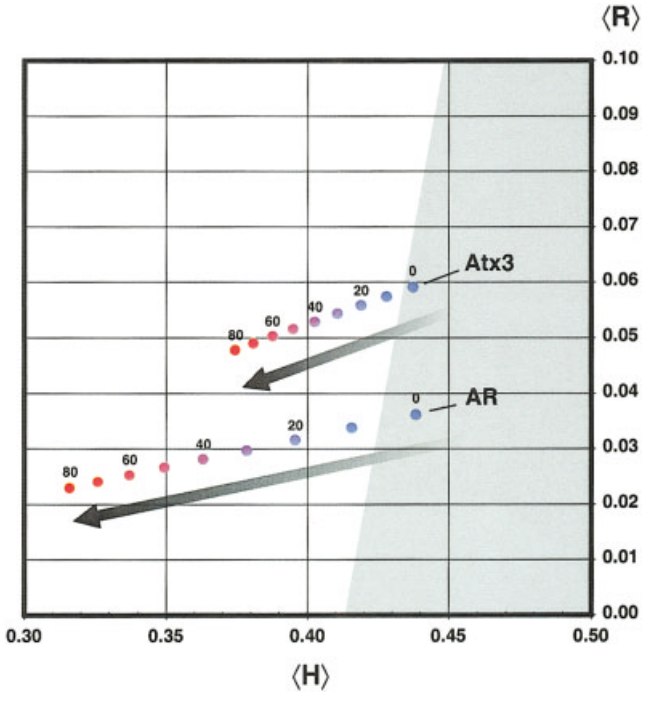

Fig. 3. Simulation of polyQ expansion for the AR (local) and Atx3 (global) proteins. For both proteins, the first blue spot represents prediction excluding the polyQ. A pair of values was calculated with each glutamine residue added, simulating the effect of polyQ expansion up to 80 glutamines. A spot is shown for every 10 glutamines added, and the color of the spot becomes progressively more red as the length of the polyQ increases.

relatively large number (10–40) of glutamine residues into a local context of about 100 residues can bring down the overall hydrophobicity considerably. For example, addition of 35 glutamines to a domain of 100 residues can lower the mean hydrophobicity by as much as 0.1. The value of the critical mean hydrophobicity implies that any domain context having a mean hydrophobicity < 0.51

TABLE I. Results of Global and Local Protein Folding Predictions

Protein context	Length	⟨*R*⟩	⟨*H*⟩	State
Huntingtin (polyQ at 18–40)	3144 (3121)	0.020 (0.020)	0.492 (0.495)	F (F)
Htt: 1–180*	180 (157)	0.0 (0.0)	0.399 (0.443)	U (F)
Androgen R (polyQ at 58–89)	919 (887)	<0.01 (0.01)	0.452 (0.464)	F (F)
AR: 1–170*	170 (138)	0.029 (0.036)	0.378 (0.466)	U (F)
Atrophin 1 (polyQ at 484–497)	1185 (1171)	<0.01 (<0.01)	0.406 (0.409)	U (U)
Atf1:384–497*	114 (100)	0.018 (0.02)	0.381 (0.420)	U (U/F)
Atf1:434–547	114 (100)	<0.01 (0.01)	0.365 (0.403)	U (U)
Atf1:484–597	114 (100)	0 (0)	0.367 (0.403)	U (U)
Ataxin-1 (polyQ at 197–226)	816 (786)	<0.01 (<0.01)	0.445 (0.457)	F (F)
Atx1:97–226	130 (100)	0.0 (0.0)	0.404 (0.491)	U (F)
Atx1:147–276*	130 (100)	0.015 (0.02)	0.371 (0.451)	U (F)
Atx1:197–326	130 (100)	0.023 (0.03)	0.365 (0.437)	U (F)
Ataxin-2 (polyQ at 166–187)	1312 (1290)	0.025 (0.026)	0.408 (0.414)	U (U)
Atx2:66–187	122 (100)	0.074 (0.09)	0.422 (0.487)	U (F)
Atx2:116–237*	122 (100)	0.057 (0.07)	0.421 (0.492)	U (F)
Atx2:166–287	122 (100)	0.057 (0.07)	0.425 (0.491)	U (F)
Ataxin-3 (polyQ at 296–317)	360 (338)	0.056 (0.059)	0.417 (0.438)	U (U/F)
Atx3:199–360*	162 (140)	0.049 (0.057)	0.351 (0.404)	U (U)
Ataxin-6 (polyQ at 2314–2324)	2505 (2494)	0.013 (0.013)	0.447 (0.448)	F (F)
Atx6:2214–2324	111 (100)	0.090 (0.1)	0.300 (0.317)	U (U)
Atx6:2264–2374	111 (100)	0.117 (0.13)	0.354 (0.382)	U (U)
Atx6:2314–2424*	111 (100)	0.054 (0.06)	0.338 (0.352)	U (U)
Ataxin-7 (polyQ at 30–39)	892 (882)	0.054 (0.054)	0.417 (0.420)	U (U)
Atx7:1–110*	110 (100)	0.027 (0.03)	0.380 (0.408)	U (U)
Atx7:30–139	110 (100)	<0.01 (0.01)	0.377 (0.401)	U (U)
Ataxin-17 (polyQ at 55–95)	339 (298)	0.038 (0.040)	0.449 (0.496)	F (F)
Atx17:1–158*	158 (120)	0.019 (0.025)	0.380 (0.438)	U (F)

Mean net charge ⟨*R*⟩ and mean normalized hydrophobicity ⟨*H*⟩ calculations are reported. For each protein, the first line represents the global calculations with the full-length sequence, and subsequent lines correspond to various local contexts harboring the polyQ. Values in parenthesis are calculated by excluding the respective polyQ. A sequence is classified as either folded (F) or "natively unfolded" (U) according to its position in the ⟨*H*⟩ and ⟨*R*⟩ space (Fig. 1, 2, and 3). Marginal classifications are marked as "U/F". Data marked with * was used to produce Figure 1b.

would be unfolded, even if the domain carries no net charge. In the cases where the domain context is already unfolded, addition of polyQ renders the domain deeper into the unfolded realm [Fig. 1(b)].

Local Unfolding and Misfolding Mechanism

According to Perutz's polar zipper model, a nonpathogenic polyQ repeat is unstructured and stabilized mainly by its high entropy.[19] When the polyQ length exceeds a certain threshold, formation of an extensive regular structure such as a β-strand leads to liberation of solvent molecules and increase in the overall entropy of the system that compensates the entropy loss due to ordering of the unstructured conformation. This hypothesis can explain why expanded polyQ aggregates and forms fibers. The results reported here support this view: the inclusion of polyQ destabilizes the local context as an intermediate step toward misfolding and aggregation. It is intuitive to recall that natural natively unfolded proteins (or domains) can be induced to become folded on binding of ligands or interaction partners.[5,7] In a similar way, self-self interaction of an expanded polyQ can lead to formation of regular structure (fiber precipitate) from an otherwise unfolded structure.[15] Alternatively, domain swapping (also requires local unfolding) can also lead to polymerization and aggregation.[13,20] Some nonpolyglutamine neurodegeneration conditions also involve deposition of natively unfolded proteins, including α-synuclein, prion protein, tau protein, and amyloid-β, implicated in Alzheimer's disease, Parkinson's disease, and so forth.[6]

Agreement With Experimental Observations

The unfolding prediction results are hypothetical, and it is essential to challenge the calculations with known experimental results. The work of Bevivino and Loll[13] included studies of the secondary structures of several fusion constructs of Atx3 using circular dichorism (CD) spectroscopy. The far-UV CD spectrum of H6-Atx3-Q_{27} clearly indicates the presence of unstructured component, as judged by the pronounced peak at 206 nm. Indeed, prediction of H6-Atx3-Q_{27} suggested it to be natively unfolded (Fig. 2). On the other hand, the MBP-Atx3-Q_{27} sequence gives values that are characteristic of a folded protein (Fig. 2). The far-UV CD spectrum of this fusion protein confirms that it is indeed folded and is helical. The MBP-Atx3-Q_{78} sequence leads to an unfolded prediction (Fig. 2). Instead of showing characteristics of an unfolded species, the CD spectrum of this protein indicated the presence of β-sheet. This is consistent with the hypothesis that β-sheet formation (misfolding) is a consequence of polyQ-mediated unfolding.

In a related study, Tanaka et al.[15] used an artificial system to study the structural consequences of introducing repeats of 12, 28, 35, and 50 glutamines into sperm whale

myoglobin (Mb). This work happens to be a good model system for validation because the structure of Mb is known, and the protein context consists of a single globular domain of 153 residues. The unfolding analysis predicted that Mb-Q_{28}, Mb-Q_{35}, and Mb-Q_{50} sequences are unfolded, whereas wild-type Mb and Mb-Q_{12} sequences are not (Fig. 2). Indeed, constructs with 28 and more glutamines aggregate over time. Consistent with the predictions, the extent of destabilization due to unfolding is correlated with the length of the polyQ repeat. Further analysis showed that the Mb-Q_{35} and Mb-Q_{50} constructs behave like molten globule: with a folded core but the surface is unfolded.

Simulation of "Unfolding" Correlated With PolyQ Expansion in Atx3 and in AR

Two models of polyQ expansion are presented here. Global unfolding was demonstrated by using the full-length Atx3 protein sequence, whereas local unfolding was demonstrated with the N-terminal 170 residues of the AR sequence. The modeling results are illustrated in Figure 3.

In the two simulations, the folded protein contexts become unfolded, and the extent correlates with the lengthening of the embedded polyQ. However, the threshold numbers of glutamines obtained from these simulations, both below 10, do not agree with the physiologically observed values; suggesting that these models based on unfolding prediction are oversimplified. Nevertheless, this semiquantitative study does offer physical interpretation to the molecular mechanism of polyQ expansion. Glutamines have low hydrophobicity and no charge. Addition of each glutamine lowers the local hydrophobicity and favors unfolding. On the other hand, addition of uncharged glutamine residues helps to dilute any net charge on the local environment and favors the folded state. The results showed that glutamine residues have a larger effect on promoting unfolding by lowering hydrophobicity than driving folding by spreading the net charge.

The calculations presented here are sensitive to how the local context is defined (see later). For instance, in Atx3, the calculations based on amino acid sequence have ignored the effect of counter ions. Atx3 has a high concentration of negatively charged groups at neutral pH, with a net charge of -20. If these groups are neutralized by bound positive ions so that the overall protein plus solvent cations has no net charge, the number of glutamines that are required to reach the unfolding boundary would become 27 (calculations not shown), much closer to the well-established threshold of pathogenesis. Uversky et al.[7] also noted that bound ligands can affect the local folded or unfolded state.

Prediction of Unfolding

In this work, a prediction tool that requires the protein sequence alone was used. Despite its simplicity, the folding prediction based on calculations of net charge and hydrophobicities does capture some most important factors contributing to the folded/unfolded state of a protein. Hydrophobicity is related to solvation and entropy changes leading to folding, whereas net charge is related to repulsion leading to unfolding. This prediction method relies on these two properties as determinants of protein folding and unfolding. However, this method is of limited use when applied to large multidomain proteins because individual domains on a large protein can have very different folding properties. In these cases, an overall folded/unfolded classification would be difficult to interpret. The authors of this method have noticed this limitation, and the original folded set consists of small proteins that are globular and monomeric.[7] For large proteins, knowledge of the domain boundaries is required to delineate the region of interest. It is likely that the properties of the polyQ stretch can be propagated to but limited to affecting the host domain only.

In addition to a good estimate of domain boundaries, the real local environment of a polyQ repeat should be considered three-dimensionally and atoms that are close in space rather than in sequence should be included. The approximation of the local context without these spatial details accounts for the discrepancies of the predictions from the real conditions. On the other hand, without the need for a three-dimensional structure, the prediction method used here represents a good compromise and can lead to semiquantitative results.

ACKNOWLEDGMENTS

The author is grateful to Patrick Loll and Randy Pittman for supplying the sequences of the ataxin-3 fusion constructs.

REFERENCES

1. Perutz MF. Glutamine repeats and neurodegenerative diseases: molecular aspects. Trends Biochem Sci 1999;24:58–63.
2. Gusella JF, MacDonald ME. Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. Nat Rev Neurosci 2000;1:109–115.
3. Paulson HL, Bonini NM, Roth KA. Polyglutamine disease and neuronal cell death. Proc Natl Acad Sci USA 2000;97:12957–12958.
4. Wright PE, Dyson HJ. Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. J Mol Biol 1999;293:321–331.
5. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. J Mol Graph Model 2001;19:26–59.
6. Uversky VN. What does it mean to be natively unfolded? Eur J Biochem 2002;269:2–12.
7. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 2000;41:415–427.
8. http://www.ebi.ac.uk/swissprot
9. Nakamura K, Jeong S-Y, Uchihara T, Anno M, Nagashima K, Nagashima T, Ikeda S, Tsuji S, Kanazawa I. SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. Hum Mol Genet 2001;10:1441–1448.
10. http://www.expasy.org
11. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–132.
12. Wheelan SJ, Marchler-Bauer A, Bryant SH. Domain size distributions can predict domain boundaries. Bioinformatics 2000;16:613–618.
13. Bevivino AE, Loll PJ. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel β-fibrils. Proc Natl Acad Sci USA 2001;98:11955–11960.
14. Paulson HL, Das SS, Crino PB, Perez MK, Patel SC, Gotsdiner D, Fischbeck KH, Pittman RN. Machado-Joseph disease gene prod-

uct is a cytoplasmic protein widely expressed in brain. Ann Neurol 1997;41:453–462.
15. Tanaka M, Morishima I, Akagi T, Hashikawa T, Nukina N. Intra- and intermolecular β-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. J Biol Chem 2001;276:45470–45475.
16. Mende-Mueller LM, Toneff T, Hwang S-R, Chesselet M-F, Hook VYH. Tissue-specific proteolysis of Huntingtin (htt) in human brain: evidence of enhanced levels of N- and C-terminal htt fragments in Huntington's disease striatum. J Neurosci 2001;21:1830–1837.
17. Reid J, Kelly SM, Watt K, Price NC, McEwan IJ. Conformational analysis of the androgen receptor amino-terminal domain involved in transactivation influence of structure-stabilishing solutes and protein interactions. J Biol Chem 2002;14:14.
18. http://www.sanger.ac.uk/Software/Pfam
19. Perutz MF. Glutamine repeats as polar zippers: their role in inherited neurodegenerative disease. Mol Med 1995;1:718–721.
20. Liu Y, Gotte G, Libonati M, Eisenberg D. A domain-swapped RNase A dimer with implications for amyloid formation. Nat Struct Biol 2001;8:211–214.