LINEAR AND EXTENDED: A COMMON POLYGLUTAMINE CONFORMATION RECOGNIZED BY THE THREE ANTIBODIES MW1, 1C2 AND 3B5H10

Fabrice A. C. Klein¹,*, Gabrielle Zeder-Lutz⁴, Alexandra Cousido-Siah², André Mitschler², Aline Katz²,⁵, Pascal Eberling³, Jean-Louis Mandel¹, Alberto Podjarny² and Yvon Trottier¹,*

¹Translational Medicine and Neurogenetics
²Integrative Biology Departments
³Peptides Synthesis Facility, Institute of Genetics and Molecular and Cellular Biology (IGBMC), UMR 7104-CNRS/INSERM/University of Strasbourg, BP10142, 67404 Illkirch Cédex, France
⁴UMR 7242 CNRS Biotechnologie et Signalisation Cellulaire, Ecole Supérieure de Biotechnologie, Bd Sébastien Brant, BP 10413 - 67412 ILLKIRCH Cédex
*Address correspondence to: fabrice.klein@gmail.com or yvon.trottier@igbmc.fr - IGBMC. 1, rue L. FRIES, 67404 Illkirch Cédex, France - Tel. +33 3 88653413 ; Fax. +33 3 88653246
⁵present address: Computational Chemistry and Biology Group - DETEMA, Facultad de Química, UdelaR, Isidoro de María 1620 piso 3, CC1157, Montevideo, Uruguay
ABSTRACT

A long-standing pathomechanistic model proposes that the polyglutamine (polyQ) length-dependent toxicity threshold observed in all polyQ diseases is triggered by a conformational change within the monomer that occurs only above a certain polyQ length. If true, this yet undefined and elusive mutant-specific toxic conformation would constitute a direct therapeutic target.

Three anti-polyQ antibodies - MW1, 1C2 and 3B5H10 - have been extensively used to probe the conformation of polyQ. The crystal structure of the MW1 epitope reveals a linear, non-pathogenic polyQ. In contrast, although the detailed structure of its epitope is unknown, the 3B5H10 antibody is widely advertised and used as a conformational antibody that recognizes the toxic conformation of expanded polyQ.

We solved the crystal structure of the 1C2 antigen binding domain (1C2-Fab) and performed a direct comparison between the 1C2, MW1 and 3B5H10 structures. The MW1 and 1C2 antibodies have similar sequences and structures, consistent with their binding to short polyQ and their polyQ length-discrimination properties. Unexpectedly the 3B5H10 antibody also shares striking features with MW1 and 1C2, which prompted us to revisit its binding properties. We show that the 3B5H10 epitope is actually a short, non-pathogenic polyQ. All three antibodies MW1, 1C2 and 3B5H10, interact similarly with polyQ of various lengths, and bind small polyQ epitopes in similar linear and extended conformations.

Together with studies published during the recent years, our work argues against the hypothesis that a mutant-specific conformation in monomeric polyQ molecules is the toxic entity responsible for polyQ diseases.
INTRODUCTION

Polyglutamine (polyQ) disorders are a group of 9 inherited neurodegenerative disorders caused by the aberrant expansion of a trinucleotide repetition (CAG)n within the coding region of a protein. These pathologies become penetrant above a certain size threshold that varies among polyQ diseases and reflects the gain of toxic functions of mutant polyQ proteins (1). It has been proposed that the polyQ length-dependent toxicity threshold is triggered by a conformational change within the monomer that occurs only above a certain polyQ length. This stable and aggregation-prone conformation would convey toxic properties to the soluble monomeric mutant protein and provoke aberrant interactions with its partners (2-9). This "structural toxic-threshold" model has been challenged by another model, in which polyQ tracts are inherently toxic sequences whose deleterious effects gradually increase with polyQ length and aggregation kinetics (10-19). In the latter model, polyQ toxicity manifests once the protein homeostasis machinery, that normally prevents proteotoxicity, is overwhelmed by the accumulation of toxic aggregated species of mutant polyQ proteins.

Addressing whether monomeric expanded polyQ of pathogenic length can adopt a unique and toxic conformation remains an important question as such a "mutant conformation" would constitute a direct therapeutic target. Anti-polyQ antibodies have been commonly used as tools to discriminate different protein conformations. In particular, three antibodies share the ability to interact specifically and strongly with monomeric expanded polyQ and are extensively used in the polyQ research field: the 1C2, MW1 and 3B5H10 antibodies (3, 9, 20). We initially proposed that the strong interaction of 1C2 with mutant polyQ could reflect the existence of a mutant polyQ conformational epitope (9). However, we and others later demonstrated that the strong interaction of 1C2 and MW1 is not due to the recognition of a conformational epitope, but due to avidity - a phenomenon in which the valence of both the antibody and the antigen can result in enhanced binding capacity (10, 14). The crystal structure of MW1 bound to a non-pathogenic Q10 peptide further revealed that the polyQ epitope adopts a linear and extended conformation (21). More recently, the X-ray crystal structure of unbound 3B5H10 was reported (22), and although the X-ray structure...
of 3B5H10:polyQ complex has not been determined, it is proposed - based on cell-imaging, biochemistry, small-angle X-ray scattering and modelling - that the 3B5H10 antibody specifically recognizes a monomeric and toxic mutant-polyQ conformation that triggers neuronal death (20, 22). If true, this antibody, which is widely advertised and used in the polyQ field as a conformational antibody, would be a unique tool that could help in understanding the structure and toxicity relationship of mutant polyQ.

Here, we solved the crystal structure of the antigen binding domain (Fab) of the 1C2 antibody and compared the primary sequences and 3D structures of 1C2, MW1 (apo and holo) and 3B5H10, in order to better understand the structural determinants that would allow different antibodies to tell apart various polyQ conformations. Surprisingly, we found that both the sequences and 3D structures of 3B5H10, 1C2 and MW1 are strikingly similar. Furthermore, we demonstrate that 3B5H10, like 1C2 and MW1, binds polyQ as short as Q_{11} with micromolar affinities. We conclude that all three antibodies recognize non-pathogenic polyQ in a linear and extended conformation.

RESULTS

Crystal structure determination of 1C2-Fab.

We solved the X-ray structure of 1C2-Fab apo form in two different space groups (Supplementary table 1 and figure S1). In addition, to determine the conformation of the polyQ epitope recognized by the antibody, we aimed at crystallizing 1C2-Fab holo form in complex with a polyQ peptide. The length of the polyQ epitope of 1C2-Fab was previously determined to be around Q_{11} (14). Accordingly the purification of 1C2-Fab involves an affinity step with Q_{11} (figure S1a). We thus attempted to co-crystallize 1C2-Fab with polyQ peptides ≤ Q_{11} using several strategies (figure S2a). Although stable complexes between 1C2-Fab and Q_{10} peptides form in solution (figure S2a), we failed to obtain co-crystals of 1C2-Fab:polyQ complex. This is likely due to the Fab crystal packing blocking its antigen-binding surface (figure S2b), which results in the crystallization of the 1C2-Fab apo form only. To circumvent this limitation, we attempted to co-crystallize the minimal domain of interaction of
1C2 (1C2-Fv) with polyQ peptides (figure S2c), a strategy that was successful to obtain the MW1:Q_{10} structure (21). Although 1C2-Fv actively bound polyQ in solution, as determined by biochemistry and surface plasmon resonance (SPR) (figure S2d, e), we did not obtain 1C2-Fv crystals, neither in apo nor in holo forms.

Taking advantage of the apo 1C2-Fab and the recently published apo 3B5H10 crystal structures, we thus performed a thorough comparison of the three antibodies MW1, 1C2 and 3B5H10 in search of variations in their antigen-binding surfaces, which would explain their binding different polyQ conformations.

**Comparison of the Light chain Variable domains (Lv) of 1C2, MW1 and 3B5H10.**

We compared the primary sequences and 3D structures of the Lv domains of 1C2, MW1 and 3B5H10. All three antibodies contain a Lv domain encoded by a rare variable V_{\lambda x} light chain gene present in less than 1% of normal mice antibodies (23, 24) and known to undergo very little somatic mutations (25). Consequently, the V_{\lambda x} aminoacid sequences of 1C2, MW1 and 3B5H10 are almost identical (figure 1a).

We searched in the Protein Data Bank (PDB) for other antibodies carrying a Lv domain encoded by this rare V_{\lambda x} gene, which could help to us understand how 1C2, MW1 and 3B5H10 interact with their epitopes. We found only 2 other V_{\lambda x}-encoded antibodies of known 3D structure: 13F6-1-2, an antibody that recognizes an Ebola virus glycoprotein ((26), PDB code 2QHR) and a monoclonal antibody directed against the parathyroid hormone-related protein: anti-PTHRP ((27), PDB code 3FFD). We also included in our comparisons the F28C4 antibody directed against an encephalitogenic peptide because, although its 3D structure remains unsolved, detailed binding properties and epitope conformation are available and rendered them relevant to our analysis (28, 29).

As expected, the V_{\lambda x} sequences of these antibodies are also highly homologous to the V_{\lambda x} chains of 1C2, MW1 and 3B5H10 (figure 1a). Consequently, the 3D structures of the V_{\lambda x} domains of 1C2, MW1, 3B5H10, 13F6-1-2 and anti-PTHRP are extremely similar (figure 1b). The three complementarity determining region (CDR) loops L1, L2 and L3 of V_{\lambda x} are
longer than in other Lvλ subtypes and their conformations may represent new canonical
classes for Vλx-encoded light chains, as proposed by Lee et al. (26). Indeed, comparison of
apo (1C2, MW1 and 3B5H10) and holo (MW1, 13F6-1-2 and anti-PTHRP) structures reveal
only minute variations in the conformations of loops L1 and L2, which are rigid and
unaffected by ligand binding (figures 1b-d). The loop L3 is more flexible, as illustrated by its
important conformational changes observed between apo and holo structures of MW1 ((21)
and figure 1b), or between two neighbouring apo MW1-Fv molecules in the same crystal
(figure S3a). Nevertheless, we found that loop L3 adopts a very similar conformation in the
apo structures of 1C2 and 3B5H10, and in one apo MW1 molecule (figure 1e). The only
variations observed mostly concern solvent-exposed and less structured amino acids with
elevated B-factors (I98-Q101, figure 1e and figure S1c-d).

Given the high homology of sequence and 3D structure of different Lvλx antibody
domains, common structural determinants may underly their reported binding properties to
their respective ligands (26). Indeed, the interaction of 13F6-1-2, F28C4 and MW1 with their
ligand peptide is atypical and differs from most antibody-peptide structures. Usually, the
ligand peptide lies at the light chain/heavy chain interface and is mainly in contact with the
heavy chain (21, 26). In contrast, the ligand peptides of 13F6-1-2, F28C4 and MW1 adopt a
linear and extended conformation (21, 26, 29). Moreover, the ligand peptides of 13F6-1-2
and MW1 lie diagonally in a shallow groove formed at the surface of the light and heavy
variable domains, and mediate important contacts with the Lvλx domain: 27% of the buried
surface area for 13F6-1-2 (26), and 55% of all contacts for MW1 (21). Importantly, all Lvλx
structures display an almost identical groove at their surface, which strongly suggests that
the Lvλx domains of 1C2, 3B5H10 and MW1 could bind a polyQ stretch in a similar manner.

Interestingly, the affinity of F28C4 and 13F6-1-2 for their corresponding antigen was
shown to be strongly dependent on the presence of a glutamine in the epitope (26, 29). Lee
et al. proposed that the anchoring of this glutamine on germline Lvλx residues might
constitute a general binding rule on the shallow groove of a Lvλx domain (26). As shown in
figure 1f, the MW1:polyQ structural interaction reported by Li et al., (21) supports this
hypothesis. Similarly, we show that, although the peptide recognized by anti-PTHRP is helical (27), it also contains a glutamine in a conformation that allows contacts with the same conserved residues of the Lvλx domain (figure 1f). Hence, it is very likely that the interactions of 1C2 and 3B5H10 Lvλx domains with polyQ ligand are also mediated by a similar binding rule involving a key glutamine and specific Lvλx residues.

All together, the high primary sequence and structural homology of 1C2, 3B5H10 and MW1 Lvλx domains, their highly similar shallow groove at the ligand binding site, and the specific structural determinants of Lvλx:antigen binding revealed by these comparisons, suggest that the 1C2 and 3B5H10 Lvλx domains should bind a polyQ peptide in a conformation very similar to that observed in the MW1:polyQ crystal structure.

Comparison of the Heavy chain Variable domains (Hv) of 1C2, MW1 and 3B5H10.

The Hv sequences of 1C2, MW1 and 3B5H10 are more divergent than the Lv chains, with their CDR H3 being the most variable part, both in sequence and in length (figure 1g). Accordingly, the 3D structures of the Hv domains of 1C2, MW1 and 3B5H10 are more disparate than the Lv ones, especially in the H3 region (figure 1h).

Although the MW1 sequence is globally more distant to that of 1C2 and 3B5H10, some similarities exist between the three antibodies, such as an enrichment of Hv-CDRs in aromatic residues (22), and the fact that H1, responsible alongside H3 for interactions between Hv-MW1 and polyQ (21), adopts a similar conformation in the three antibody structures (figure 1h). Moreover, if CDR-H3 is omitted, the Hv sequences of 1C2 and 3B5H10 display a remarkably high homology (94 identical amino acids out of 108 = 87% identity), and their CDR H1 and H2 only vary by a single amino acid respectively (figure 1g). The overall 3D structures of Hv-1C2 and Hv-3B5H10 are extremely similar, as underlined by their loops H1 and H2, with the exception of their H3 (figure 1h-j). Furthermore, homology searches with the program BLAST (30) identifies 3B5H10 as the 2nd closest homolog of the 1C2-Fab heavy chain out of 2911 hits in the PDB with an E-Value<0.01 (E-value = 2.48237E-100). The closest homolog identified was the heavy chain of the anti-cholera toxin TE33-Fab.
(E-value = 6.75231E-102). The light chains of TE33-Fab and 3B5H10 however lack remarkable homology, and the TE33-Fab recognizes non-polyQ related antigens (31).

In summary, when taking into account the sequences of both their light and heavy chains, 1C2 and 3B5H10 are the closest homologs found in the PDB. Furthermore, the 3D structures of L1, L2, L3, H1 and H2 CDR loops of 1C2 and 3B5H10 exhibit striking similarities, and only the H3 loop significantly differs in the two antibodies.

3B5H10, like MW1 and 1C2, interacts with micromolar affinity with small, non pathologic polyQ.

The strong sequence homology and structural similarities between 3B5H10, 1C2, and MW1 led us to consider the possibility that 3B5H10 could recognize non-pathogenic polyQ, as do 1C2 and MW1. To clarify this point, we analyzed the interaction properties of 3B5H10 with polyQ spanning from non-pathogenic to pathogenic lengths, or flanked with various sequences.

We performed GST pull-downs using well characterized GST-polyQ proteins (14, 15) to study the ability of 3B5H10 to interact with polyQ of various non-pathogenic (Q11, Q22) and pathogenic (Q41) lengths. We found that 3B5H10 can specifically bind to polyQ as small as Q11 (figure 2a).

The surface plasmon resonance (SPR) technique was previously used to show that the Fabs of MW1 and 1C2 have a micromolar affinity for polyQ stretches of non-pathogenic and pathogenic lengths, and to conclude that they do not specifically recognize a mutant specific polyQ conformation (10, 14). We thus studied by SPR the binding properties of 3B5H10-Fab with polyQ of various lengths, using 1C2-Fab as a control. As shown in figure 2b the intensity of the SPR response at the end of the injection phase (R120s) increased as a function of the polyQ length for all concentrations of 3B5H10-Fab or 1C2-Fab injected on immobilized GST-polyQ proteins. This phenomenon has already been documented for MW1 and 1C2, and can be explained by the progressive valence increase of the polyQ sequence, which allows longer polyQ tracts to interact simultaneously with several Fabs (10, 14). A similar effect was
observed on western blot (WB) when using the Fabs of 3B5H10 or 1C2 as primary antibodies, as the intensity of the signal increased progressively as a function of polyQ length (Figure 2c). Interestingly, soluble polyQ – of non-pathogenic and pathogenic lengths, and surrounded by different flanking sequences - is a disordered homopolymer (10, 11, 14, 15, 32, 33), bearing a number of epitopes that increases in parallel with polyQ length (e.g., in the case of MW1, it increases from 13 epitopes in Q22 to 32 epitopes in Q41, figure S4a). The multiplicity of epitopes, combined with the possibility of several Fabs interacting simultaneously on longer polyQ, constitute thus an unusually complex interaction system (figure 2b-c, figure S4a and (10)). Therefore, conditions can be found to detect only the longer polyQ, even when similar quantities of short and long polyQ are present: for example, the Q41 stretch, but not the Q11 one, is detected with the lower concentrations of 1C2 and 3B5H10 Fabs (figure 2b and c). This was already illustrated in similar WB experiments previously performed by Bennett et al. with MW1 and 1C2 (10), by us with 1C2 (14) and by Miller et al. (20) with 3B5H10, where a Q68 stretch was strongly revealed by the Fab or IgG form, but a smaller Q17 sequence was only weakly detected by the full antibody. Importantly, the polyQ length dependent increase of R120s and WB responses (figure 2b-c) is progressive and is not specific to non-pathogenic/pathogenic length transitions (Q22 to Q41), since it also occurs between two non-pathogenic lengths (Q11 to Q22). In light of all this information, the increased SPR and WB responses on expanded polyQ compared to short polyQ, are not due to the recognition of a mutant-specific conformation by the Fabs.

We estimated by SPR that the affinities of 3B5H10-Fab or 1C2-Fab for Q11, Q22 and Q41 sequences are all in the micromolar range, with very little variation from Q11 to Q41 (figure S4b), consistently with the previous characterization of 1C2 and MW1 antibodies by SPR (10, 14, 21). Moreover, to assess the potential influence of sequences flanking the polyQ stretch on the interaction with 3B5H10 and 1C2, we also studied their binding with non-pathogenic polyQ stretches surrounded by different flanking sequences: Q11 bound to GST (GST-Q11) and Q16 in the context of the Huntingtin exon-1 fused to an MBP carrier (MBP-Htt-exon1-Q16, (12)). We found that the sequences flanking the polyQ stretch do not influence
the antibody:polyQ interaction and that both antibodies have similar micromolar affinities for Q\textsubscript{11} and for Htt-exon1-Q\textsubscript{16} (figure 2d and figure S4c).

It is noteworthy that the interaction properties of 1C2 and 3B5H10 with polyQ are not exactly the same (figure 2b-d). This is not surprising, since 1C2 and 3B5H10 differ significantly for the H3 CDR domain (figure 1), and are thus not identical. However, our results show that the 3B5H10, MW1 and 1C2 interact with polyQ of various lengths in an extremely similar fashion and that the epitope recognized by 3B5H10 is not a mutant-polyQ specific conformation, but can be a non-pathologic polyQ stretch as small as Q\textsubscript{11}.

**Discussion**

*MW1, 1C2 and 3B5H10 recognize a non-pathogenic polyQ in a linear and extended conformation.*

The three anti-polyQ antibodies MW1, 1C2 and 3B5H10 were generated independently, in three different laboratories and with three different antigen contexts containing polyQ stretches of different lengths: 1C2 was generated against the TATA-Binding Protein carrying Q\textsubscript{38}, MW1 against the Atrophin-1 carrying Q\textsubscript{19}, and 3B5H10 against an amino-terminal Huntingtin fragment carrying Q\textsubscript{66} (3, 9, 20). Despite the dissimilarity of antigens, all three antibodies share very striking similarities: their interaction properties with polyQ stretches, as studied with different techniques, are remarkably alike ((10, 14, 21) and see Figure 2 and S4); they all carry a rare variable light chain Lv\textsubscript{λ} domain, which differs from other Lv\textsubscript{λ} domains through its important role in binding antigenic peptides in an extended conformation (21, 25, 26, 29); the Lv\textsubscript{λ}x domains of the 3 antibodies have almost identical sequence and 3D structure shaping a similar groove at the ligand binding surface (figure 1). In addition, the Hv moieties of 1C2 and 3B5H10 are also very similar (see figure 1g-j). The remarkable likeness of MW1, 1C2 and 3B5H10 suggests that during the immune response process in mouse, elicited by different antigenic context, a similar polyQ epitope was used for the generation of the three antibodies.
The unique atomic structure of a polyQ peptide bound to the MW1 antibody shows that the MW1 epitope consists of a small Q_{10} peptide in an extended conformation, sitting non-canonically in a shallow groove on the surface of the Lv and Hv domains (21, 26), see figure 3a). The Lvλx domain of MW1 mediates 55% of all contacts with the Q_{10} peptide, and 6 out of 10 glutamines are sitting on the Lvλx surface. To bind the Q_{10} peptide, the antigen-binding surface of MW1 undergoes some structural rearrangements, including important conformational changes of the L3 and H3 CDR loops (figure 3b). Since the Lvλx domains of MW1, 1C2 and 3B5H10 display more than 97% identity (figure 1a) and their 3D structures reveal an almost identical shallow groove at their ligand-binding interfaces (figure 1c-e), these Lvλx domains should all interact with a polyQ peptide in an almost identical manner.

To visualize the binding of polyQ to the Lvλx moieties of 1C2 and 3B5H10, we positioned a 6-glutamines peptide in an extended conformation on their surface (figure 3c-d), as observed for the MW1:Q_{10} structure (figure 3a). A structural change similar to that seen for MW1 Lvλx (see Figure 3b) is likely to occur on the surface of 1C2 and 3B5H10, to fully accommodate the linear and extended Q_{6} stretch. Since 1C2 and 3B5H10 recognize polyQ as short as Q_{10} or Q_{11} (see figure 2, S2a and (14)), the 4 to 5 remaining glutamines that are not in contact with the Lvλx domain are likely to be positioned in the groove formed by the hyper-variable loops of the Hv domain of 1C2 and 3B5H10, in an unknown conformation (see figure 3c-d).

Given the major contribution of the Lvλx moiety to the interaction, the polyQ epitopes of 1C2 or 3B5H10 are thus mostly linear and extended.

*The existence of a mutant-specific toxic conformation in monomeric polyQ is controversial.*

A long-standing pathomechanistic model, which we previously referred to as "structural toxic-threshold" model (14), postulates that a mutant-specific and monomeric polyQ conformation is the toxic entity responsible for polyQ diseases. This model was initially proposed based on differences observed between wild-type and mutant polyQ proteins: expanded polyQ was shown to be specifically recognized by the 1C2 antibody (9) and to
perturb the interaction of a carrier protein with a partner (5); moreover, only expanded polyQ proteins were initially found to aggregate \textit{in vitro} and \textit{in vivo} and to be toxic (8, 34).

Therefore, it was proposed early on that monomeric polyQ above a certain length-threshold can adopt a unique, stable, aggregation prone and toxic conformation, which can be specifically recognized by antibodies, and which can perturb protein interactions (5, 6, 8, 9).

However, the properties of polyQs were later studied under new experimental conditions, and cumulative evidence now contradict the "structural toxic-threshold" model: in contrast with earlier findings, non-expanded polyQ can form aggregates \textit{in vitro}, although at much lower kinetics and/or at higher concentrations, and these aggregates resemble those of expanded polyQ (11, 14); non-expanded polyQ can also be toxic when administered as aggregates in cellular cultures (19), or when overexpressed in \textit{C. elegans} (16), showing that toxicity is not an exclusive property of elongated polyQ; using carefully purified recombinant proteins and quantitative methods to study protein interactions, we recently demonstrated that the presence of an expanded polyQ does not \textit{per se} alter the interactions of a monomeric mutant protein with its partners (12); finally, anti-polyQ antibodies such as MW1 and 1C2 were shown to recognize short polyQ stretches, and not a mutant-specific conformation (10, 14, 21).

Although the detailed characterization of the 3B5H10 antibody has only been published quite recently (20, 22), it has been used as a conformational anti-polyQ antibody for almost a decade, and remained thus one of the key arguments in favour of the "structural toxic-threshold" model. In our study, we now demonstrate that the 3B5H10 antibody does not recognize a mutant-specific and toxic polyQ conformation, as previously suggested (20, 22), but a short non-pathogenic polyQ peptide, as do the 1C2 and MW1 antibodies (10, 14, 21). As a consequence, none of the known anti-polyQ antibodies are able to recognize a mutant-specific polyQ conformation, challenging even further the relevance of the "structural toxic-threshold" model.
A number of findings indicate that non-pathogenic and pathogenic polyQ share many biophysical and structural properties: they are intrinsically disordered (10, 11, 14, 15, 33) but rather compact (35, 36) as soluble monomers, and they undergo similar conformational changes, leading to the formation of aggregates with overall similar amyloid-like properties and structural features (11, 17, 37-39). However, some properties of polyQ vary with polyQ length: for instance, progressive polyQ length-dependent variations were reported in the mature aggregates, such as differences in the size and shape of the fibres, in their exhibition of Congo Red birefringence (37), or in their ability to form intramolecular contacts (39); along the same line, the nucleation process and the aggregation kinetics are influenced by polyQ length variations (11, 40). It is now crucial to decipher whether and how such polyQ length-dependent variations contribute to trigger nine neurodegenerative diseases, intrinsically caused by polyQ elongation.

MATERIALS AND METHODS
Protein expression and purification: the 1C2 antibody (IgG1) from mice ascites fluids was purified using a single affinity purification step against a GST-Q_{11} fusion protein pre-bound on a Glutathione sepharose column. Treatment with thrombin protease, on the column, in phosphate 50 mM pH7 buffer, allowed for release of the Q_{11}:1C2 complex off the column. The Q_{11} peptide was removed by dialysis (cut-off 50kDa), the pure 1C2 was then cleaved by papain treatment as reported earlier (Klein et al., 2007). Fc fragments were retained by protein-A sepharose beads and 1C2-Fab retrieved pure in the supernatant as reported (Klein et al., 2007). 1C2-Fab was either dialyzed in ammonium acetate (50 mM) for ESI-TOF analysis, or in various buffers for crystallization (10 mM TRIS, 10 mM NaCl, pH 7.3) or SPR (10 mM HEPES, pH 7.4, 150 mM NaCl,3.4 mM EDTA, 0.005% (v/v) Tween20; (P6585 Sigma-Aldrich)). 1C2-Fab was stored at 4°C at concentration ~10 to 20 µM and concentrated on Amicon Ultra systems (Millipore) at up to 200 µM (~10 mg/ml) for crystallization purposes. 1C2 and 1C2-Fab production is summarized in figure S1a. 3B5H10 was purchased from...
SIGMA-ALDRICH (P-1874). 3B5H10-Fab was prepared following methods identical to that reported for 1C2-Fab (14). GST-polyQ proteins were purified as reported (14).

Crystallization and structural resolution methods are detailed in SI Materials and Methods and figures S1b-d and S2.

GST Pull-down assays: Phosphate buffer 50 mM pH 7 (4°C) was used at all steps. 300 µl of pre-washed and equilibrated Glutathione-sepharose beads (GE 17-0756-01) were loaded with equal quantities of GST (negative control), GST-Q_{11}, GST-Q_{22} or GST-Q_{41} proteins (270 µl à 80 µM), washed (10 column volumes) and kept at 4°C as sources of GST and GST-polyQ beads for 4 days (time needed to perform three GST pull-down experiments). To realize the pull-down while using as little 1C2 or 3B5H10 IgGs as possible, only 10 µl of either GST or GST-polyQ beads, were loaded on home-made micro-columns (pipette tips plugged with cotton). Equal quantities of either 1C2 or 3B5H10 (15 µl à 5 µM) were loaded on GST, GST-Q_{11}, GST-Q_{22} or GST-Q_{41} columns. After the IgGs had passed through the column and after washing (8 column volumes), the beads were resuspended in 15 µl of SDS/PAGE running buffer containing Laemmli loading dye, boiled for 2 min and loaded on SDS/PAGE (16%) for electrophoresis.

Western Blot interaction assays of 1C2 and 3B5H10 with GST-polyQ proteins:
Equal amounts of pure GST-Q_{11}, GST-Q_{22} and GST-Q_{41} proteins (200 picomole each) were mixed in 150 µl SDS/PAGE running buffer containing Laemmli loading dye, and boiled for 2 min. They were loaded on a unique and 6.2 cm wide SDS/PAGE (16%) well. After electrophoresis and transfer on nitrocellulose membrane, the unique lane was cut into 15 equivalent lanes. Each ~4mm wide lane was individually incubated for 90 min at room temperature (RT) in 2 ml of different primary antibodies: purified 1C2 or 3B5H10 IgGs, or purified 1C2-Fab or 3B5H10-Fab, or home made monoclonal anti-GST ascites (dilution 1/60,000). The concentrations of primary antibodies and identity of the lanes are shown in figure 2b. After 3 washing steps of 10 min in Tris buffer saline (TBS), the lanes were reassembled (with tape on the extremities) and simultaneously incubated with goat-anti-mouse antibody coupled to peroxidase (GAM-perox, 1/5,000 dilution) in TBS. Immobilon
Western (Millipore) was used for revelation. The chemiluminescent signal was recorded on a Fusion (Fx7) image acquisition system (Vilber Lourmat), by successive acquisition windows, which provided acquisition data between 30 sec and 1,110 sec of exposition.

Surface Plasmon Resonance: The instrumentation and methodologies were adapted from Klein et al., (14) and Davranche et al. (12). The interaction studies were carried out in real time by SPR analysis using a BiacoreT200® instrument (GE Healthcare Biacore, Uppsala, Sweden). Experiments were carried out at 25°C using a flow rate of 30 µl/min and HBS as running and dilution buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.005% (v/v) Tween20; (P6585 Sigma-Aldrich)).

To capture GST-Qn molecules in figure 2b, anti GST antibody (GST capture kit BR100838 GE Healthcare Biacore, Uppsala, Sweden) was directly immobilized to all four flow cells of the chip according to the supplier’s protocol (GE Healthcare). One flow cell was used as a control and was loaded with GST (reference surface) and the other three flow cells were loaded with GST-Q11, GST-Q22 or GST-Q41 (Klein 2007).

For the experiment described in figure 2d, carboxymethylated dextran sensor chips sCM5 (BR-100030, GE Healthcare Biacore) were used to immobilize purified GST (reference surface), GST-Q11 and MBP-Htt-exon1-Q16 proteins, via the standard amine coupling method (amine coupling kit, BR-100050 GE Healthcare Biacore, Uppsala, Sweden).

Serial dilutions of each purified Fab or Mab were injected over the four cells at a flow rate of 30 µl/min. Each cycle consisted of a 80 sec or 120 sec sample injection followed by a post-injection phase of 7,200 sec in running buffer.

Correction of all binding curves was performed by so-called double referencing, i.e. subtraction of the data of the GST control flow cell followed by subtraction of the data from an average of three buffer injections (Myszka, 2000).

The initial reaction rate on the GST-polyQ surfaces was not affected by flow variations between 5 µl/min and 100 µl/min (GST-polyQ immobilization levels less than 120 RU and 20 nM injected 1C2 and 3B5H10 Fabs or Mabs), indicating the absence of mass-transport phenomenon.
To determine Fabs and Mabs binding affinities and/or kinetic parameters for the different Qn antigens, the data from 6 different concentrations were globally fitted to molecular interaction models found in Biacore T200 evaluation software (1 to 1 binding, bivalent analyte (A + B ⇔ AB; AB + B ⇔ AB2); heterogeneous ligand (A + B1 ⇔ AB1; A + B2 ⇔ AB2); and two-state reaction (A + B ⇔ AB ⇔ AB*). Regression analysis of kinetic data was limited to these relatively simple molecular interaction models which provided acceptable fits in order to compare the affinity orders even when experimental data are complex. All tested models provided consistent results, i.e. micromolar affinities. The results for the "1 to 1 binding mode" are shown in supplementary figures S4b-c.

ACKNOWLEDGEMENTS

We thank the Advanced Photon Source (Argonne National Laboratory, USA, operated by UChicago Argonne, LLC, for the U.S. DOE - contract DE-AC02-06CH11357), especially beamlines ID19, SBC (Steve Ginell) and ID23, BioSync (Ruslan Sanishvili), and the Swiss Light Source, PSI, Switzerland, X06SA beamline (Clemens Schulze Briese). We thank Pierre Poussin-Courmontagne and the IGBMC Structural Biology and Genomics Platform. We are grateful to: Dr. Susana Shochat (technical help), Dr. Pingwei Li (sumo-GQ10G and SUMOase constructs (21)), Peran Terrier & Hélène Nierengarten (mass-spectrometry), Mustapha Oulad-Abdelghani (1C2 production), Olga Koutsopoulos and Sebastiaan Werten (critical reading of the manuscript).

This work has progressed at a slow pace because of technical issues inherent to aggregation prone proteins and to crystal structure projects, and also because our preliminary results - which did not support the long-standing structural toxic-threshold model - led to the drop of some initial fundings. In the beginning, this work was fully supported by the Cure Huntington’s Disease Initiative program of High-Q Foundation (2004-2009) (to YT, AP & FK). Partial support was provided by Agence Nationale de Recherche (ANR-05-MRAR-034-01; 2006-2008) (to YT, AP and FK) and the European integrated project EUROSCA (LSHM-CT-2004-503304 for YT; 2005-2009) as well as by the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Centre National de la Recherche Scientifique (CNRS). Apart from institutional fundings, the project was not supported since 2009. However, FK is grateful to the Association Paul and Liba Mandel (APLM) from Strasbourg, for stipends received in 2009-2010 and 2011-2012.

CONFLICT OF INTEREST STATEMENT:

The authors declare no conflict of interest
REFERENCES


FIGURE LEGENDS

Figure 1. Primary sequence and 3D structure comparison of the MW1, 1C2 and 3B5H10 antibodies.

a. primary sequence alignments of the L\(\lambda\)\(x\) domains of the indicated antibodies:
The compared L\(\lambda\)\(x\) sequences are highly similar. The upper numbering corresponds to the 1C2 sequence, the lower numbering to the Kabat scheme. The amino acids that differ from the other sequences are shaded in dark grey. The CDR loops defined as in Lee et al. (26) are shaded in light grey. The amino acids shaded in orange indicate discrepancies between the sequences of 3B5H10 associated with the structures 3S96.pdb and 4DCQ.pdb ((22), see also figure S3b and c). The stars represent the amino acids of MW1 that are in contact with the polyQ peptide in the MW1:Q\(_{10}\) structure (21). The black stars represent two conserved L\(\lambda\)\(x\) amino acids (T\(^{32}\) and G\(^{95}\), 1C2 numbering) that bind a glutamine present in different L\(\lambda\)\(x\) ligands peptides ((21, 26, 27), see also figure 1f).

b. Superimposition of the L\(\lambda\)\(x\) domain structures of MW1 apo (grey = 2GSG.pdb), MW1 holo (green = 2OTU.pdb, cyan = 2OTW.pdb), 1C2 apo (orange = 4ISV.pdb, beige = 4JJ5.pdb), 3B5H10 apo (red = 3S96.pdb, purple = 4DCQ.pdb), 13F6-1-2 holo (dark blue = 2QHR.pdb) and anti-PTHRP holo (light-blue = 3FFD.pdb).

c and d. Superimposition of the L1 (c) and L2 (d) loops of all antibodies presented in b (structures and color codes as in figure 1b).

e. Superimposition of the L3 loop observed in the structures of apo 1C2, 3B5H10, and MW1 (same color codes as in figure 1b).
f. In three different ligand:antibody structures, a glutamine residue from the ligand adopts a similar conformation and mediates similar contacts on the surface of the $\lambda\lambda$ domain. $\text{MW1:Q}_{10}$ (green = 2OTU.pdb), 13F6-1-2: (blue = 2QHR.pdb) and anti-PTHRP (yellow = 3FFD.pdb). The distances in Angstrom ($d_1$; $d_2$; $d_3$) are: MW1:Q$_{10}$ (2.8; 2.9; 3.0), 13F6-1-2: (2.9; 3.1; 2.9) and anti-PTHRP (2.8; 3.4; 2.9). $T^{32}$ and $G^{95}$ (1C2 numbering) are labeled with a black star in figure 1a.

g. primary sequence alignments of the Hv domains of MW1, 1C2 and 3B5H10:
The numbering, shading color and star codes are identical to figure 1a. Aminoacids that are unique to each sequence are shaded in black. The additional box within CDR H3 indicates that the main chain of these amino acids adopt a different configuration in all three antibodies, whereas the surrounding amino acids are in a similar configuration in the three structures.

h. Superimposition of Hv domain structures of MW1 (grey = 2GSG.pdb), MW1:Q$_{10}$ (green = 2OTU.pdb, cyan = 2OTW.pdb), 1C2 (orange = 4ISV.pdb, beige = 4JJ5.pdb) and 3B5H10 (red = 3S96.pdb, purple = 4DCQ.pdb).

i and j. Superimposition of the H1 (i) and H2 (j) loops of 1C2 and 3B5H10 (color codes as in figure 1h).

Figure 2. Interaction of the 3B5H10 and 1C2 antibodies with polyQ of various lengths.

a. GST pull-down assays of 3B5H10 and 1C2 on GST-polyQ bearing different polyQ lengths: 3B5H10 and 1C2 come down in the presence of GST-Q$_{11}$, GST-Q$_{22}$ or GST-Q$_{41}$ but not GST. The Coomassie blue stained gel shown here is representative of three pull-down experiments. In one experiment, the samples were also studied in parallel by western-blot (WB), using a Goat-Anti-Mouse (GAM) antibody coupled to peroxidase that reveals only the
heavy chain of 1C2 and 3B5H10 antibodies (V\(\lambda\)-encoded chains are not recognized by most GAM sera (41)). The WB signal revealing both antibodies on the GST-polyQ lanes is saturated, whereas almost no signal is detectable on the negative-control GST lane, showing that the interactions of 1C2 and 3B5H10 with GST-Q_{11}, GST-Q_{22} or GST-Q_{41} are highly specific.

b. SPR analysis of interaction of 3B5H10-Fab and 1C2-Fab with polyQ of various lengths: The two antibodies display similar binding properties with polyQ of various lengths. GST (negative control and reference), GST-Q_{11}, GST-Q_{22} and GST-Q_{41} were linked by an anti-GST antibody (as in Klein et al., (14)) on different channels of the same SPR chip (93, 110, 97 and 90 relative units (RU), respectively). Different concentrations of 3B5H10-Fab and 1C2-Fab (from 62.5nM to 2,000nM) were successively injected on the SPR chip. The SPR sensorgrams show the specific binding signal of increasing concentrations of 1C2-Fab or 3B5H10-Fab on GST-Q_{11}, GST-Q_{22} or GST-Q_{41}. For a given concentration of Fab, the SPR response at the end of the injection phase (R_{120s}) increases progressively in a polyQ length dependent manner. This figure is representative of two independent experiments that included replicates for the 3B5H10-Fab injections. Kinetic and affinity parameters returned from regression analysis of the Fab:Qn curves are reported in figure S4b.

c. Interaction of 3B5H10-Fab and 1C2-Fab with polyQ of various lengths as analyzed by western blot: The efficiency of detection of polyQ stretches by western blot, employing 1C2-Fab or 3B5H10-Fab as primary antibodies, increases in a polyQ-length dependent manner. Equal amounts of GSTQ_{11}, GSTQ_{22} and GST-Q_{41} were mixed and loaded in a unique well. After SDS-PAGE and transfer onto a nitrocellulose membrane, the membrane was cut into 15 equivalent lanes. Each lane was individually incubated with the following primary antibodies, at the indicated concentrations: 1C2 IgG (lane 1); 1C2-Fab (lanes 2 to 7); anti-GST (lane 8); 3B5H10-Fab (lanes 9 to 14); 3B5H10 IgG (lane 15). The 15 lanes were then reassembled,
incubated with GAM-peroxidase, and revealed simultaneously for 30 sec (top panel) or 1,110 sec (lower panel). Note the equal amounts of the three GST-polyQ proteins as revealed with anti-GST after a 30 sec exposure.

d. Determination of the affinity of 3B5H10 and 1C2 for GST-Q₁₁ and MBP-Htt-exon1-Q₁₆ by SPR:

3B5H10 and 1C2 have similar affinity for GST-Q₁₁ and MBP-Htt-exon1-Q₁₆. GST-Q₁₁, MBP-Htt-exon1-Q₁₆ and GST (negative control and reference) were cross-linked on different channels of the same SPR chip (37, 96 and 45RU, respectively). The figure shows the specific binding of 3B5H10 and 1C2 that were injected at different concentrations. The pattern of interaction of each antibody with GST-Q₁₁ and MBP-Htt-exon1-Q₁₆ are similar, and both antibodies have a micromolar affinity for GST-Q₁₁ and MBP-Htt-exon1-Q₁₆ (see figure S4c).

Figure 3. Structural insights into MW1, 1C2 and 3B5H10 interactions with small polyQ in a linear and extended conformation

a. Representation of the MW1:Q₁₀ interaction surface (pdb code 2OTU, (21)):

The Q₁₀ peptide (yellow/orange) adopts an extended conformation and sits diagonally in a shallow groove on the surface of the Lvλ (green) and Hv (blue) MW1 subunits. The aminoacids of the Lv moiety in contact with the polyQ peptide (21) are coloured in dark green. The glutamine side chains of the polyQ are displayed as thin lines, except two of them that are shown as sticks: the yellow one (N°4) that mediates canonical contacts with the Lvλ domain ((26) and figure 1f) and the orange one (N°6) that mediates 6 contacts with the Lvλ domain, and only one with the Hv domain, representing the limit of the polyQ peptide in contact with the Lvλ domain. The approximate positions of the loops L1, L3, H1 and H3 – that contact the polyQ peptide – are indicated. The same codes (color and representation) apply in figures 3b, 3c and 3d.
b. Representation of the Fv domain of apo MW1 (pdb code 2GSG, (21)), on which the polyQ peptide was positioned as observed in the MW1:Q_{10} structure (a):

The structure of the apo FV domain of MW1 was aligned on that of MW1 in complex with Q_{10}, and the Q_{10} peptide was superimposed on the apo MW1 structure. In this position, the polyQ peptide is closer to H3 (steric clash) and further away from L3, when compared to the MW1:Q_{10} structure (a). Indeed, it was reported that upon polyQ binding, major structural changes involve H3, which moves to avoid direct contacts with the glutamine N°9, and L3 which moves towards the polyQ peptide (21). These structural changes observed upon polyQ binding are indicated by two red arrows.

c and d. Representation of the Fv domains of 1C2 (c) and 3B5H10 (d), on which a polyQ peptide Q_6 was positioned on the Lv\lambda x surface, as observed in the MW1:Q_{10} structure (a):

The structures of the Fv domains of 1C2 (c, PDB code 4ISV) and 3B5H10 (d, PDB code 3S96) were aligned with that of MW1 in complex with Q_{10}. A Q_6 peptide was positioned on the binding interface of 1C2-Lv (c) and 3B5H10-Lv (d), in an identical conformation and similar position as that observed in the MW1:Q_{10} structure. The dotted yellow line indicates the putative position of the groove on the Hv domains of 1C2 (c) and 3B5H10 (d), on which 4 to 5 C-terminal residues of a small Q_{10} or Q_{11} peptide should be positioned, with respect to the MW1:Q_{10} structure (a). Structural changes similar to those occurring for MW1 (as shown in (b)) are also likely to occur upon polyQ binding on 1C2 and 3B5H10.

ABBREVIATIONS

The abbreviations used are: polyQ, polyglutamine; Q_n, polyQ of n residues; GST, glutathione S-transferase; SPR, surface plasmon resonance; mab, monoclonal antibody; Fab, antigen binding fragment; Hv, Heav-chain variable domain; Lv, Light chain variable domain; Fv, Hv:Lv heterodimer; RU, resonance unit; RT, room temperature.