Biochemical and Biophysical Characterization of Huntingtin

Thesis by

Gwen Ellen Owens

In Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2016

(Defended October 29, 2015)

Acknowledgements

I would like to thank the members of the Bjorkman lab for making my time at Caltech a productive and enjoyable experience. I would like to thank Pamela for providing me with encouragement and advice, as well as giving me great freedom to pursue independent work. You set an incredible example as a researcher, thoughtful mentor, and role model.

Thank you to the other members of my thesis committee: Rob, Doug, and Kai. Your guidance, discussion, ideas and feedback have been invaluable to me.

Thank you to Marta, Lyn, and original Beth. Marta, your cookies at 4pm and help with paper figures got me through many long days in the lab. Lyn, you are amazingly funny, and seamlessly made things in the lab run. Beth, I love your enthusiasm about science, and how you have always been willing to help or lend a sympathetic ear. I was able to send part of my thesis project to the International Space Station, and I wouldn't have submitted my application without your late-night encouragement in the lab! Thank you to members of the Bjorkman Lab who specifically worked on the Huntington's Disease project with me, including Danielle New, my star summer student, Alejandra Olvera, Allen Ninh, and Anthony West. Thank you to Jost for helping with SPR experiments, Pavle for assistance setting up crystal trays and screens, and Jens for teaching me how to collect X-ray diffraction data.

I am deeply indebted to the UCLA-Caltech MSTP, especially Drs. Steven Smale, Kelsey Martin, Siavash Kurdistani, and Carlos Portera, for their guidance through my first two years of medical school at UCLA and during my PhD at Caltech. Thank you also to Dr. Yvette Bordelon, a physician-scientist at UCLA who allowed me to shadow her on many

occasions in the Motor Neuron Disease Clinic when she met with patients with Huntington's Disease.

I would like to thank my amazing new husband Ethan for supporting me throughout graduate school. Ethan, you have been a constant source of encouragement, and you bring out the very best in me.

My biggest thank you goes my wonderful parents as I wouldn't be where I am today if it wasn't for them. You have always been an inspiration to me. Mom and Dad, thank you for all the emotional, practical and financial support leading up to today. You do so much for me– and you've given me the confidence to think I can achieve anything I want in my life– even completing a PhD at Caltech! You've shown me so much support and I want you to know that I thank you from the bottom of my heart.

Abstract

Huntington's disease (HD) is a fatal autosomal dominant neurodegenerative disease. HD has no cure, and patients pass away 10-20 years after the onset of symptoms. The causal mutation for HD is a trinucleotide repeat expansion in exon 1 of the huntingtin gene that leads to a polyglutamine (polyQ) repeat expansion in the N-terminal region of the huntingtin protein. Interestingly, there is a threshold of 37 polyQ repeats under which little or no disease exists, and above which, patients invariably show symptoms of HD. The huntingtin protein is a 350 kDa protein with unclear function. As the polyQ stretch expands, its propensity to aggregate increases with polyQ length. Models for polyQ toxicity include formation of aggregates that recruit and sequester essential cellular proteins, or altered function, producing improper interactions between mutant huntingtin and other proteins. In both models, soluble expanded polyQ may be an intermediate state that can be targeted by potential therapeutics. In the first study described herein, the conformation of soluble, expanded polyQ was determined to be linear and extended using equilibrium gel filtration and small-angle X-ray scattering. While attempts to purify and crystallize domains of the huntingtin protein were unsuccessful, the aggregation of huntingtin exon 1 was investigated using other biochemical techniques including dynamic light scattering, turbidity analysis, Congo red staining, and thioflavin T fluorescence. Chapter 4 describes crystallization experiments sent to the International Space Station and determination of the X-ray crystal structure of the anti-polyQ Fab MW1. In the final study, multimeric fibronectin type III (FN3) domain proteins were engineered to bind with high avidity to expanded polyQ tracts in mutant huntingtin exon 1. Surface plasmon resonance was used to observe binding of monomeric and multimeric FN3 proteins with huntingtin.

Published Content

G.E.O. and P.J.B. conceived the study; G.E.O. and D.M.N. performed protein expression, purification, Western and dot blots, and SEC-MALS; G.E.O. performed equilibrium gel-filtration assays; G.E.O., A.P.W., and P.J.B. analyzed the data; and G.E.O. and P.J.B. wrote the paper with all co-authors contributing to scientific planning and discussions.

Table of Contents

Abbreviations

AUC, analytical ultracentrifugation

CDR, complementarity determining region

DLS, dynamic light scattering

EDT, Eastern Daylight Time

EK, enterokinase

Fab, antigen-binding fragment

FN3, fibronectin type III

Fv, variable fragment

HC, heavy chain

HD, Huntington's disease

HDPCG, High Density Protein Crystal Growth

ISS, International Space Station

KSC, Kennedy Space Center

LC, light chain

MALS, multiangle light scattering

MELFI, Minus Eighty Laboratory Freezer for ISS

N17, 17 N-terminal amino acids of huntingtin

polyQ, polyglutamine

R_g, radius of gyration

R_h, hydrodynamic radius

SAXS, small angle X-ray scattering

SEC, size-exclusion chromatography

SPR, surface plasmon resonance

SSRL, Stanford Synchrotron Radiation Lightsource

TRX, thioredoxin

V_H, heavy chain variable domain

V_L, light chain variable domain

Chapter 1: Huntingtin

Chapter 1: Huntingtin

1. Huntington's Disease

Huntington's disease (HD) is a fatal neurodegenerative disease in humans that is caused by an expanded trinucleotide repeat $(CAG)_n$ in the huntingtin gene on chromosome 4p16.3 [1], in the first exon of the huntingtin gene. The encoded huntingtin protein has a series of repeated glutamine amino acid residues known as a polyglutamine (polyQ) repeat. The pathogenic mechanism of HD is thought to involve a conformational transition that occurs in expanded polyQ tracts (>37 glutamines) that leads to formation of toxic aggregates, although the nature of this conformational transition is unclear. Despite knowledge of the causal mutation, there is currently no effective treatment to delay or prevent HD onset or progression.

1.1 Genetics

The first exon of the huntingtin gene contains a repetitive DNA stretch of three repeated nucleotides: C (cytosine), A (adenine), and G (guanine). In normal individuals, this stretch of trinucleotides consists of 18 to 22 CAG repeats. When 36-39 CAG repeats are present, HD occurs in some, but not all, people [2], known as 'variable penetrance.' When the CAG repeat length expands to more than 39 CAG repeats, HD is observed in 100% of patients, known as 'complete penetrance.' While the normal CAG repeat is stably inherited, longer CAG repeats change in length when transmitted to subsequent generations. Large increases in CAG repeat length are sometimes seen with paternal transmission of mutant huntingtin, resulting in an earlier age of onset of the disease,

known as 'anticipation' [3].

The severity of the disease and the younger age at which the onset of Huntington's disease occurs are correlated positively with the length of the CAG repeat [4]; the longer the CAG repeat, the earlier in life patients begin to exhibit HD symptoms. Although age of onset of HD can be predicted by CAG length, CAG length only accounts for half of the variation in age of onset, with the remaining variance attributable to other genes and environmental factors [4]. Several genes involved in DNA repair mechanisms, including *MLH1*, have recently been shown to modify the age of onset of HD by several years [5].

Most patients with HD have repeat lengths of 40 to 50 CAG, and onset of symptoms typically occurs in the fourth or fifth decade of life [6]. When the CAG tract expands to >62 repeats, onset of HD occurs in the first or second decade of life with symptoms distinct from adult-onset HD, and is known as 'juvenile HD' [7]. Because onset of juvenile HD occurs prior to reproductive age, the longest mutations are not passed down to the next generation. On average, patients die 17 years after symptoms begin [8], regardless of polyQ length or age of onset (Marcy MacDonald, personal communication, May 10, 2015).

HD is an autosomal dominant disorder, so there is a 50% chance of passing the HD mutation from parent to offspring, regardless of gender. Intriguingly, individuals with *two* copies of the HD disease gene have the same age of HD onset and severity as individuals with only one mutated copy, which is unique among autosomal dominant disorders.

HD has an estimated prevalence of 2.7 per 100,000 people worldwide, and 5.7 per 100,000 people in North America, with a higher prevalence in populations with European

heritage [9, 10]. At least 30% of *HD* chromosomes appear to be ancestrally related [1]. The Maracaibo region of Venezuela has one of the highest prevalences of HD in the world at 700 in 100,000 people, which is thought to be due to Spanish colonization of Venezuela in the 16th century [10]. Using polymorphic DNA markers of families in Venezuela with HD, the genetic defect was mapped to chromosome 4 [11]. Several years later, the $(CAG)_n$ repeat expansion was mapped to the IT15 gene [1], later renamed huntingtin.

1.2 Clinical Presentation

Patients with HD suffer from a triad of clinical findings: 1) movement disorders including chorea, dystonia, and lack of coordination; 2) cognitive deficiencies such as impaired executive function and memory; and 3) psychiatric and behavioral symptoms. *Chorea* is derived from the Greek word for "dance" (with the same root as *choreography*) [12], and is characterized by short, involuntary movements that increase in frequency and severity during the progression of HD in many patients. The symptoms of HD are correlated with progressive neuron loss in the brain, particularly in medium spiny neurons in the basal ganglia, caudate nucleus, and putamen [13]. These areas of the brain are important in control of motor movements and behavior, and damage to these areas is thought to lead to the characteristic symptoms of HD.

HD is diagnosed by family history and genetic testing. Testing of individuals suspected of inheriting the CAG mutation in the huntingtin gene is rarely performed in individuals less than 18 years old due to ethical concerns [7].

The course of HD is divided into several stages. In the first stage, patients have minor symptoms, and they are able to live and function with minor accommodations. In the middle stages of HD, patients often are unable to drive and may begin to have significant motor difficulties, such as falls and trouble swallowing, but can perform some functions with assistance. In late-stage HD, patients require 24-hour care, and have significant communication difficulties.

While the onset of HD symptoms typically occurs during adulthood, 6% of HD patients have symptoms before the age of 20. In contrast with adult HD patients, patients with juvenile HD exhibit tremor and sustained involuntary muscle contractions instead of chorea [7].

Monitoring symptoms of patients in a clinical setting is difficult, and few tests exist to judge neural health. The Unified Huntington's Disease Rating Scale (UHDRS) [14-16] has been used for the last 20 years to evaluate patients with HD. Tests include an assessment of motor skills, including ocular pursuit, finger taps, and gait appraisal; cognitive assessment, including verbal fluency and word reading; behavioral assessment, with questions to judge the frequency and severity of the patients' behavior, suicidal thoughts, depression, hallucinations, and other behaviors associated with HD. Also included in the UHDRS is an assessment of the independence and functional capacity of the patient throughout disease progression.

While many clinical trials for potential HD therapeutics have relied on the UHDRS for patient evaluation, several new, less subjective tests have been developed in the past five years that directly monitor CSF concentration of huntingtin. The first of these tests is a Forster resonance energy transfer (FRET) assay [17, 18], in which two labeled

monoclonal antibodies directed against huntingtin N-terminal epitopes are bound to huntingtin. Aggregates of mutant huntingtin are recognized preferentially by this antibody pair, and the donor and acceptor fluorophores are able to generate a FRET signal. The second assay uses a micro-bead based immunoprecipitation and flow cytometry (IP-FCM) [19], using similar antibodies as the FRET assay. While these tests are still under clinical evaluation, it is hoped that they will improve assessment of HD progression as well as more clearly delineate treatment efficacy.

1.3 Treatment

Several medications are prescribed for patients with HD, including antidepressants, sedatives, and botulinum toxin injection for treating dystonia [20]. Because HD is a progressive disease, medications must be evaluated on a regular basis. Education, exercise, and diet are also important parts of management of HD.

Only one drug, tetrabenazine (Xenazine®), is approved for treatment of Huntington's Disease [21]. Tetrabenazine was approved in 2008 for treatment of chorea, and while it is effective at reducing chorea in patients with HD by decreasing dopamine release, it also can cause sedation and depression, problems for which patients with HD are already at elevated risk. All other drugs have failed in clinical trials.

It is possible that previous clinical trials have been unsuccessful due to a focus on single pathways that mutant huntingtin affects, while mutant huntingtin affects a large number of integrated pathways. Therefore, changing the level of mutant huntingtin protein expression may be a more comprehensive method of treating HD. Methods to decrease the amount of huntingtin protein include intrabodies [22-24] and gene silencing techniques such as single-stranded RNAs (RNAi) [25] and antisense oligonucleotides (ASO) [26, 27] that

target either the polyQ stretch or disease-linked polymorphisms. Several observational trials of HD, including PREDICT-HD, Registry, ENROLL-HD, and COHORT [20], are helping to prepare for upcoming gene silencing clinical trials.

2. Huntingtin function

Huntingtin is a 3,144-amino-acid protein with a molecular mass of 348 kDa. However, the role of wild-type huntingtin is poorly understood. Huntingtin is expressed in the cytoplasm of all human and mammalian cells, with the highest concentrations found in the brain and the testes. In neurons and fibroblasts, huntingtin has also been found in the nucleus [28, 29]. The first exon of huntingtin, which contains the disease-causing CAG repeat in humans, is highly conserved across all mammals. In mice, huntingtin has an 86% sequence identity with human huntingtin DNA and a 91% sequence identity with human huntingtin protein [30]. However, the murine gene has a CAG repeat that encodes only 7 consecutive glutamines, which explains the lack of a natural mouse model of HD. In addition to mammals, zebrafish, drosophila, and slime molds also have orthologs of huntingtin, suggesting that huntingtin is required for some basic cellular function. A series of mouse studies has shown that huntingtin is essential early in embryonic development, even before the emergence of the nervous system, and that it may also play an important role in basal ganglia function during adulthood. When the murine homolog of huntingtin was inactivated, Hd -/- homozygous embryos died soon after gastrulation, while heterozygotes had increased levels of motor activity and "cognitive deficits" [31]. Interestingly, physical disruption or deletion of portions of the huntingtin gene in humans does not cause HD, suggesting that the mutant huntingtin

gene may give the huntingtin protein a toxic gain of function. One study reported that the polyQ tract in huntingtin was an important regulator of mitochondrial ADP-phosphorylation in both non-HD and HD CAG repeat ranges [32]; however, many other functions for huntingtin have been proposed, including cytoskeletal anchoring and facilitation of signal transduction.

While HD is the most common inherited neurodegenerative disease, nine other human diseases in addition to HD are caused by expressed, expanded polyQ repeats. These diseases include spinocerebellar ataxia types 1, 2, 3, 6, 7, 12, and 17, detatorubropallidoluysian atrophy, and spinobulbar muscular atrophy. In each of these diseases, there is selective neurodegeneration, and neurons can be found with inclusions containing aggregated polyglutamine proteins. Additionally, for each of these disorders*, irrespective of the protein affected*, the length of the polyQ repeat inversely correlates with the age of onset and directly correlates with disease severity, with the longest polyQ repeats associated with early-onset (juvenile) forms of the disease. Interestingly, several of these disorders have similar polyQ repeat ranges to HD, with disease onset occurring after expansion to ≥ 35-40 polyQ repeats. This range may be affected by the location of the polyQ repeat within the expressed protein or other properties of the affected protein.

3. Huntington Aggregation

Although the precise role of the polyQ stretch within huntingtin in the pathology of Huntington's disease is uncertain, this abnormal region of huntingtin is presumed to cause excessive protein binding interactions with other cellular proteins or with other

polyQ repeats. These excessive interactions result in the formation of large, insoluble, inactive protein aggregates in affected brain regions called inclusions. *In vitro*, polyQ repeats containing least 37 glutamines aggregate, forming dimers, trimers, and higher-order oligomers. The rate of aggregation is faster with increasing numbers of glutamine residues, which seems to explain the correlation between disease severity and polyQ repeat length.

The exact mechanism whereby polyQ aggregation leads to neurotoxicity in HD has not yet been determined. However, several important processes leading to neurodegeneration have been elucidated. Mutant huntingtin has a higher proteolysis rate than wild-type huntingtin. Truncated huntingtin causes increased aggregates and may translocate to the nucleus [8]. Mutant huntingtin may harm neurons through cell-cell interactions by disrupting the function of nearby neurons or glia that support that neuron. Aggregates also may interfere with normal proteins by recruiting normal interaction partners of wild-type huntingtin into their aggregate matrix. At some point, the aggregate formation overcomes the cells' ability to degrade them, leading to inclusion bodies and neuron death. However, the toxicity of aggregated huntingtin has been intensely debated due to conflicting evidence that the aggregates are harmful, neutral, or neuroprotective. Infrared microspectroscopy was used to demonstrate that huntingtin inclusions in brains of patients with HD are, in fact, polymorphic; beta-sheet conformations appeared to be highly neurotoxic, while other inclusions that lacked structural rearrangements appeared to be non-toxic [33]. One recent study found that proteins with disordered domains are recruited to polyQ aggregates, including many RNA-binding proteins [34], which could account for toxicity independent of the function of the polyQ stretch. Other studies have suggested that mutant huntingtin spreads through the CSF with prion-like propagation,

where the small nuclei of mutant huntingtin are able to seed aggregation of monomeric

mutant huntingtin [35], cascading into a larger effect both intra- and extra-cellularly.

4. Huntington Structure

4.1. Huntingtin Exon 1

The protein encoded by exon 1 of the huntingtin gene (hereafter referred to as huntingtin

exon 1 protein) includes 90 amino acids (for a Q23 reference sequence) and four distinct

regions. An N-terminal 17-amino-acid region (N17) is located immediately before the

polyQ region. The polyQ region is followed by a polyproline (polyP) region. Interestingly,

polyQ stretches >6 glutamines and the polyP region are each found only in mammals

[36], leading to the hypothesis that the polyP region may help to stabilize the polyQ

region [37]. C-terminal to the polyP region is a proline-rich region (PRR).

```
        10          20          30          40          50          60
MATLEKLMKA FESLKSFQQQ QQQQQQQQQQ QQQQQQQQQQ PPPPPPPPPP PQLPQPPPQA

        70          80          90
QPLLPQPQPP PPPPPPPPGP AVAEEPLHRP
```

Fig. 1-1. Amino acid sequence of huntingtin exon 1 protein with 23 glutamines. The N17

region is shown in blue, polyP region in red, and PRR in green. NCBI Reference

Sequence: NP_002102.4, huntingtin [Homo sapiens].

The N17 region has several post-translational modifications, including phosphorylation,

acetylation, and sumoylation, and N17 is proposed to mediate nuclear transport of huntingtin and modify the association of huntingtin to with the cytoplasmic membrane. Transgenic mice with mutant huntingtin without the N17 domain show increased cellular pathology and more overt disease compared with mice with huntingtin containing the N17 domain [38].

Several partial structures have been solved and deposited into the Protein Data Bank repository (www.rcsb.org) (Table 1-1); however, none of the deposited structures contain the entire polyQ region thought to be critical for pathogenesis of mutant huntingtin. Two NMR solution structures of the first 17 amino acids in the N-terminal domain show that the N17 domain can transition between an alpha-helical conformation in micelles to a random coil in aqueous solutions [39].

Six medium- to high-resolution X-ray crystal structures including the N17 domain and several glutamines in non-pathologic polyQ stretches of huntingtin were crystallized as part of a maltose-binding protein (MPB) fusion protein at 3.5 Å resolution [40] (PDB:3IO4, 3IO6, 3IOR, 3IOT, 3IOU, 3IOV, 3IOW). In this series of structures, the modeled polyQ region was conformationally flexible and was affected by the conformation of nearby residues.

Four additional X-ray crystal structures of huntingtin with a Q36 stretch with 3 histidine residues introduced within the polyQ stretch, also as part of a MBP fusion protein [41]. While the polyQ region adopts several conformations, the histidine insertions and the non-pathogenic polyQ repeat length add uncertainty to any conclusions drawn from this structure.

STRUCTURE	PDB ID	RESOLUTION
Solution structure of the N-terminal domain of huntingtin (htt17) in 50% TFE [39]	2LD0	Solution NMR
Solution structure of the N-terminal domain of huntingtin (htt17) in presence of DPC micelles [39]	2LD2	Solution NMR
Huntingtin amino-terminal region with 17 Gln residues-crystal C95 [40]	3IOR	2.68 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C92-b [40]	3IOT	2.71 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C94 [40]	3IOU	2.7 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C99 [40]	3IOV	3.5 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C99-Hg [40]	3IOW	3 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C90 [40]	3IO4	3 Å
Huntingtin amino-terminal region with 17 Gln residues-crystal C90 [40]	3IO6	3.7 Å
Anti-huntingtin VL domain in complex with huntingtin peptide (EKLMKAFESLKSFQ)	3LRH	2.6 Å
Htt36Q3H-EX1-X1-C1 (Alpha) [41]	4FE8	2.8 Å
Htt36Q3H-EX1-X1-C2 (Beta) [41]	4FEB	2.8 Å
Htt36Q3H [41]	4FEC	2.8 Å
Htt36Q3H [41]	4FED	2.8 Å

Table 1-1. Molecular structures of components of huntingtin exon 1.

4.2. Full-Length Huntingtin

Full-length huntingtin has proven challenging to study, due to the large size of the huntingtin protein and extensive post-translational modifications of the protein. Three research groups have described production of recombinant full-length huntingtin protein: production in insect cells with dimerization and yields of less than 1 mg/L [42]; production in insect cells with yields of 1-3 mg/L [43]; and production in mammalian cells with significant oligomerization [44]. Huntingtin is predicted to contain 10-36 HEAT repeats, which help proteins act as scaffolds for other proteins (MacDonald 2003).

5. Anti-Huntingtin Antibodies

Dozens of antibodies have been generated that bind to huntingtin. These antibodies include several monoclonal anti-polyQ antibodies, including MW1 [45], 3B5H10 [46],1C2 [47, 48], and 1F8 [49, 50], which is reported to be similar to 1C2 [49]. The X-ray crystal structures of 3B5H10 antigen-binding fragment (PDB: 3S96, 4DCQ) [51, 52], 1C2 Fab (PDB:4ISV, 4JJ5) [53], and MW1 Fv alone (PDB: 2GSG) and MW1 Fv in complex with a GQ$_{10}$G peptide (PDB: 2OTU, 2OTW) [54] are similar [53]. While each of these antibodies has a different apparent affinity for polyQ, 1C2, 3B5H10 and MW1 all have lambda light chains, homologous sequences, and strong structural similarity [55]. Several additional anti-huntingtin exon 1 antibodies were developed at Caltech by the Patterson Lab [56, 57].

mAb	Isotype	Epitope	Antigen	PDB ID (Fragment)
3B5H10	IgG	PolyQ	Huntingtin exon 1-66Q	4DCQ (Fab) 3S96 (Fab)
1C2	IgG	PolyQ	TATA-binding protein-38Q	4JJ5 (Fab) 4ISV (Fab)
MW1	IgG2b	PolyQ	DRPLA-19Q	2GSG (Fv) 2OTU (Fv + $GQ_{10}G$) 2OTU (Fv + $GQ_{10}G$)
MW2	IgM	PolyQ	DRPLA-35Q and TRX-35Q	
MW6	IgM	PolyQ	Soluble huntingtin exon 1-67Q	
MW7	IgM	PolyP	Aggregated and soluble huntingtin exon 1-67Q	
MW8	IgG2a	AEEPLHRPK	Aggregated and soluble huntingtin exon 1-67Q	

Table 1-2. Monoclonal anti-huntingtin exon 1 antibodies.
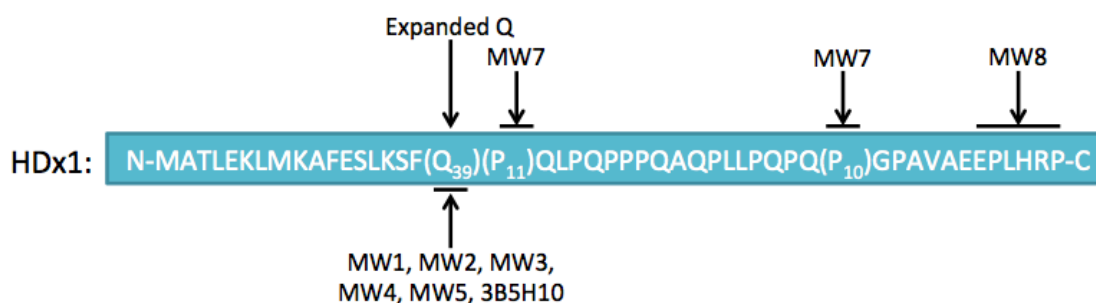

Figure 1-1. Schematic of anti-huntingtin exon 1 epitope specificity.

6. Linear Lattice Model for PolyQ

The "linear lattice" hypothesis for the polyQ repeat of huntingtin proposes that polyQ retains a random-coil structure for both normal and expanded polyQ in the soluble,

monomeric form. However, the increase in number of binding epitopes in expanded polyQ compared with normal polyQ results in avidity effects that cause higher apparent affinities for bivalent proteins such as antibodies [58]. This could induce altered binding interactions with other proteins or other huntingtin polyQ repeats, leading to the neuronal toxicity seen in HD.

Previous results from the Bjorkman lab have indicated that the polyQ repeats in huntingtin exon 1 are unstructured, even in huntingtin with expanded polyQ repeats. The affinity of the anti-polyQ antibody MW1 to huntingtin exon 1 protein was shown to increase in a polyQ-length dependent manner, and binding of multiple antigen-binding fragments (Fabs) of MW1 to expanded polyQ tracts was observed. In addition, circular dichroism studies demonstrated that huntingtin exon 1-TRX fusion proteins with 16–46 glutamines exhibited a random coil conformation in solution, and no evidence was found for a global conformation change above 37 glutamines [54, 58]. Surface plasmon resonance (SPR) and analytical ultracentrifugation (AUC) studies also demonstrated that multiple MW1 Fabs bound to expanded polyQ tracts [58].

From these binding and structural data, we formulated a linear lattice model to explain the apparent targeting of expanded polyQ tracts by anti-polyQ antibodies [58]. This model is based on models for binding of DNA to DNA binding proteins [59], which suggest that multivalent polyQ binding proteins, such as antibodies, will be able to discriminate between normal and expanded polyQ. This model for polyQ repeats of huntingtin plays a central role in each project described in this thesis.

Chapter 2: Anti-polyQ Antibodies Recognize a Short PolyQ Stretch in Both Normal and Mutant Huntingtin Exon 1

1. HIGHLIGHTS

- We assessed binding of MW1 and 3B5H10 anti-polyQ monoclonal antibodies and Fabs to normal and expanded huntingtin exon 1 polyQ repeats.

- Western and dot blots revealed binding of MW1 and 3B5H10 IgGs to both short and expanded polyQ tracts in huntingtin exon 1 fusion proteins.

- Equilibrium gel filtration studies showed that multiple MW1 or 3B5H10 Fabs bound to a huntingtin exon 1 fusion protein with 39 glutamines.

- Selectivity of antibodies for specific conformations of polyQ to distinguish species of huntingtin exon 1 fusion protein was not observed, and no evidence for a conformational transition between soluble wild type and mutant huntingtin exon 1 was found.

2. ABSTRACT

Huntington's disease (HD) is caused by expansion of a polyglutamine (polyQ) repeat in the huntingtin protein. A structural basis for the apparent transition between normal and disease-causing expanded polyQ repeats of huntingtin is unknown. The 'linear lattice' model proposed random-coil structures for both normal and expanded polyQ in the preaggregation state. Consistent with this model, the affinity and stoichiometry of the anti-polyQ antibody MW1 increased with the number of glutamines. An opposing 'structural toxic threshold' model proposed a conformational change above the pathogenic polyQ threshold resulting in a specific toxic conformation for expanded

polyQ. Support for this model was provided by the anti-polyQ antibody 3B5H10, which was reported to specifically recognize a distinct pathologic conformation of soluble expanded polyQ. To distinguish between these models, we directly compared binding of MW1 and 3B5H10 to normal and expanded polyQ repeats within huntingtin exon 1 fusion proteins. We found similar binding characteristics for both antibodies. First, both antibodies bound to normal, as well as expanded, polyQ in huntingtin exon 1 fusion proteins. Second, an expanded polyQ tract contained multiple epitopes for antigen-binding fragments (Fabs) of both antibodies, demonstrating that 3B5H10 does not recognize a single epitope specific to expanded polyQ. Finally, small angle X-ray scattering and dynamic light scattering revealed similar binding modes for MW1 and 3B5H10 Fab-huntingtin exon 1 complexes. Together, these results support the linear lattice model for polyQ binding proteins, suggesting that the hypothesized pathologic conformation of soluble expanded polyQ is not a valid target for drug design.

3. INTRODUCTION

Huntington's disease (HD) is a fatal neurodegenerative disorder characterized clinically by psychiatric symptoms, cognitive decline, and uncontrolled movements [36]. HD is caused by expansion of a CAG repeat within exon 1 of *HTT* (previously *HD*) that encodes an expanded polyglutamine (polyQ) tract in the N-terminal portion of the huntingtin protein. A pathologic threshold exists for HD, in which HD is fully penetrant in patients with 42 or more glutamines in the huntingtin protein, but no disease is found in individuals with 36 or fewer glutamines, while huntingtin with 37 to 41 glutamines exhibits reduced HD penetrance [60]. Although a structural basis for an apparent

normal-disease threshold is unknown, several hypotheses exist for the conformation of monomeric, soluble polyQ in normal and expanded huntingtin protein.

The 'linear lattice' hypothesis proposed that polyQ retains a random-coil structure for both normal and expanded polyQ in the preaggregation state. In this model, the increase in number of binding epitopes in expanded polyQ compared with normal polyQ results in avidity effects that cause higher apparent affinities for bivalent proteins such as antibodies [58]. This could induce altered binding interactions with other cellular proteins or other polyQ repeats, leading to neuronal toxicity. Consistent with this model, the affinity of the anti-polyQ antibody MW1 to huntingtin amino terminal protein encoded by exon 1 (hereafter called huntingtin exon 1 protein) increased in a polyQ-length dependent manner, and binding of multiple antigen-binding fragments (Fabs) of MW1 to expanded polyQ tracts was observed. In addition, huntingtin exon 1 protein with 16 – 46 glutamines exhibited a random coil conformation in solution, and no evidence was found for a global conformation change above 37 glutamines [54, 58]. Surface plasmon resonance (SPR) and analytical ultracentrifugation (AUC) studies also demonstrated that multiple MW1 Fabs bound to expanded polyQ tracts [58]. The X-ray crystal structure of a $GQ_{10}G$ peptide bound to the variable regions of MW1 revealed that a short polyQ tract adopted an extended structure in a diagonal binding groove across the antigen-binding site of MW1 [54]. Additional binding studies using the anti-polyQ antibody 1C2 [48] showed that 1C2 also exhibited preferential binding to expanded polyQ due to avidity effects, and this preferential binding was not due to a mutant huntingtin-specific toxic structure recognized by 1C2 [53].

In contrast, the 'structural toxic threshold' model proposed that a conformational transition occurs in polyQ repeats that are longer than the pathological threshold, which

results in a specific toxic conformation for monomeric expanded polyQ that could potentially be recognized by antibodies [61]. According to this model, the postulated pathologic conformation could be directly toxic or it could alter interactions between mutant huntingtin and its binding partners; in either case, the pathologic conformation could be targeted for drug design. Support for this model was provided by studies of the anti-polyQ antibody 3B5H10, which was reported to recognize a single epitope representing a distinct pathologic conformation of soluble expanded polyQ [52, 62]. In these studies, 3B5H10 IgG preferentially bound to expanded polyQ, and a two-stranded β-hairpin conformation of polyQ was modeled into the predicted polyQ-binding groove of the 3B5H10 Fab structure [52]. Support for this model was provided by a gel filtration assay of 3B5H10 Fab binding to a Q_{39}-containing huntingtin exon 1 fusion (HD-39Q) protein, which was interpreted to demonstrate a binding stoichiometry of 1:1 3B5H10 Fab:HD-39Q [62]. These results were suggested to indicate that 3B5H10 binds to a single structured polyQ epitope only present in expanded polyQ, as per the structural toxic threshold hypothesis. Modeling of small-angle X-ray scattering (SAXS) data was interpreted as showing that 3B5H10 Fab bound to HD-39Q in a 2:2 3B5H10 Fab:HD-39Q complex in which each 3B5H10 Fab recognized one subunit of an HD-39Q dimer through binding to a two-stranded β-hairpin conformation of polyQ [52]. Contradictory evidence was provided by a recent report demonstrating that pull-down assays and surface plasmon resonance (SPR) studies showed that 3B5H10 IgG, like MW1 and 1C2 IgGs, could bind to short polyQ tracts, as expected given the high sequence and structural similarities between the three antibodies [53].

Here we compared the recognition properties of the anti-polyQ monoclonal antibodies MW1 and 3B5H10 by studying their interactions with a polyQ-containing fragment of

huntingtin. Using expressed and purified huntingtin exon 1-thioredoxin (TRX) fusion

proteins containing 16 to 46 glutamines (HD-16Q, HD-25Q, HD-39Q, HD-46Q) (Fig. 1a),

we directly compared the interactions between soluble huntingtin and these anti-polyQ

antibodies using biochemical and biophysical analysis techniques. We found that both

MW1 and 3B5H10 antibodies exhibited similar binding properties, with neither providing

evidence for a toxic conformation of expanded polyQ. These results argue against

strategies designed to target a novel toxic conformation of soluble mutant huntingtin

exon 1 protein in the preaggregation state.

4. RESULTS

Both MW1 and 3B5H10 antibodies bind to normal and expanded polyQ within huntingtin exon 1 proteins

Western blots were used to evaluate the binding of 3B5H10 and MW1 IgGs to equimolar

amounts of huntingtin exon 1-TRX fusion proteins and to the TRX tag alone (Fig. 1b). If

3B5H10 recognizes a toxic conformation present only in expanded polyQ, then unlike

MW1, it should not bind to short polyQ repeats. In contrast with some previous results

[52], but consistent with other results [17, 63], we found that both MW1 and 3B5H10

IgGs bound in a similar manner to huntingtin exon 1 fusion proteins, each capable of

binding to huntingtin exon 1 proteins containing both normal and expanded polyQ

repeats. Both IgGs bound to huntingtin exon 1 proteins in a polyQ-dependent manner,

with a progressively more intense signal with increased polyQ length. Based on these

results, and previous western blots demonstrating the ability of 3B5H10 to bind to GST-

polyQ with both short and long polyQ repeats [53], we conclude that both antibodies

recognize a similar polyQ epitope that is present in both normal and expanded huntingtin exon 1 proteins.

In order to determine how it is possible to obtain results appearing to indicate that 3B5H10 binds only to expanded polyQ, we examined binding of 3B5H10 and MW1 as a function of concentration to huntingtin exon 1 proteins with different polyQ repeat lengths. Dot blots of serial dilutions of huntingtin exon 1 proteins demonstrated that both 3B5H10 and MW1 IgGs bound to huntingtin exon 1 proteins in a length- and concentration-dependent manner (Fig. 1c). At higher concentrations of huntingtin exon 1 protein, both IgGs recognized huntingtin exon 1 constructs with polyQ tracts ranging from Q_{16} to Q_{46}. However, at lower concentrations of huntingtin exon 1 protein, a more intense signal was observed for the huntingtin exon 1 construct with the longest polyQ repeat (HD-46Q) compared to the construct with the shortest polyQ repeat (HD-16Q) (Fig. S1b,c). Due to the length- and concentration-dependent binding, conditions can be found in which MW1 or 3B5H10 IgG appeared to only bind expanded polyQ, thus explaining previously-reported results that 3B5H10 only recognizes huntingtin with expanded polyQ [52]. However, for both antibodies, the binding dependence on polyQ length was progressive, without a distinct threshold at polyQ lengths >37Q. Thus, avidity effects resulted in bivalent IgG versions of 3B5H10 and MW1 showing preferential binding to expanded polyQ, as predicted by the linear lattice model for antibody interactions with polyQ repeats [55, 58].

Huntingtin exon 1 proteins are monomeric in solution

During purification of huntingtin exon 1 fusion proteins, we observed anomalous

migration by gel filtration chromatography such that huntingtin exon 1 fusion proteins appeared to migrate as higher molecular weight proteins (e.g., dimers) when compared with molecular weight standards of globular proteins. To determine the oligomeric state of the huntingtin exon 1 fusion proteins, we used a combination of size-exclusion chromatography (SEC) with in-line multiangle light scattering (SEC-MALS), a technique that can be used to determine the absolute molecular mass of a protein or complex independent of shape and model [64]. To evaluate the methodology, we first analyzed 3B5H10 Fab alone, which migrated as a single monodisperse peak whose derived molecular mass closely matched the mass calculated from the amino acid sequence (Fig. 2; Table 1). HD-16Q and HD-39Q fusion proteins also migrated as monodisperse peaks, and their calculated molecular masses corresponded to monomers in each case (Table 1). Thus the anomalous migration of each huntingtin exon 1 protein in positions expected for a dimeric version of a globular protein of the same molecular mass results from slower migration due to an elongated structure rather than from dimerization. In particular, no evidence was found for dimer formation for HD-39Q as predicted in a previous study involving the modeling of small-angle X-ray scattering (SAXS) data [52].

Non-equilibrium gel filtration chromatography analyses yield inconsistent apparent binding stoichiometries

We next replicated published non-equilibrium gel filtration chromatography experiments that were conducted to determine the stoichiometry of binding between 3B5H10 Fab and huntingtin exon 1 protein with an expanded polyQ repeat using the same proteins: 3B5H10 Fab and HD-39Q [62]. By varying the ratio of 3B5H10 Fab to HD-39Q, we

determined the ratio where the least amount of excess Fab or excess HD-39Q was detected, the method previously used to report a 1:1 3B5H10:HD-39Q binding stoichiometry [62]. Similar to the published results, we found that unbound 3B5H10 Fab was present at ratios greater than 1:1 3B5H10 Fab:HD-39Q, unbound HD-39Q was present at ratios less than 1:1, and that the ratio where the least amount excess of Fab or excess HD-39Q could be detected was 1:1 (Fig. 3a). As a control, we repeated the non-equilibrium gel filtration stoichiometry experiment to evaluate the binding behavior of MW1 Fab and HD-39Q (Fig. 3b), which was previously shown to form a complex with a greater than 1:1 stoichiometry [58]. Under non-equilibrium conditions, the stoichiometry of MW1 Fab:HD-39Q appeared to be less than 1:1. Thus it appeared that non-equilibrium gel filtration could not be reliably used to derive an accurate binding stoichiometry for an anti-polyQ Fab binding to huntingtin exon 1 fusion proteins with expanded polyQ.

However, we noted that with increasing molar ratios of 3B5H10 Fab:HD-39Q, the complex peak eluted earlier on the gel filtration column, suggesting that a complex larger than 1:1 was forming at the same time as 3B5H10 Fab was dissociating from HD-39Q. Due to the anomalous migration of HD-39Q compared with globular proteins using gel filtration chromatography (see above), the molecular mass of the 3B5H10 Fab:HD-39Q complex peak could not accurately be estimated based on gel filtration migration. Using SEC-MALS, we found that the complex of 3B5H10 Fab and HD-39Q was polydisperse, and the molecular mass of the peak fraction corresponded to a complex composed of greater than a 1:1, but less than a 2:1, ratio of 3B5H10 Fab:HD-39Q (Table 1), demonstrating that dissociation of the complex occurred during the experiment. In contrast, the complex of 3B5H10 Fab and HD-16Q migrated as a monodisperse peak, and the calculated

molecular mass for a 1:1 3B5H10 Fab:HD-16Q stoichiometric ratio was in close agreement with the molecular mass obtained by SEC-MALS (Table 1).

Taken together, the results for these experiments suggested that binding stoichiometries for MW1 and 3B5H10 Fabs binding to huntingtin exon 1 proteins as determined by non-equilibrium gel filtration are incorrect. This is likely because protein complexes that dissociate during this procedure are unable to rebind due to separation by the gel filtration column. Therefore depending on the binding kinetics and the amount of separation between anti-polyQ Fabs and huntingtin exon 1 proteins on the gel filtration column, the binding stoichiometries determined by non-equilibrium gel filtration techniques may be artificially low, as has been found in other protein-protein interaction systems evaluated by this technique [65, 66].

Expanded polyQ tracts within huntingtin exon 1 proteins contain multiple epitopes for the antigen-binding fragments (Fabs) of MW1 and 3B5H10

To determine an accurate stoichiometry of binding between HD-39Q and the Fabs of 3B5H10 and MW1, we used equilibrium gel filtration [67]. In this technique, a chromatography column is run with an equilibration buffer containing one of the binding partners (e.g., protein A). Different ratios of the binding partners (e.g., protein A and protein B) are then injected onto the column. When the amount of additional protein A injected is less than that required for formation of a complex, a trough will form at the position that protein A migrates. If the amount of additional protein A injected is in excess for complex formation, a peak is observed at the position that protein A migrates. When the amount of additional protein A injected is at the amount required for complex

formation, a flat baseline is observed at the position that protein A migrates. However, unless the protein concentration in the equilibration buffer is much greater than the affinity of the protein-protein complex, the ratio of protein A to protein B at which a flat baseline is observed will not be an integer, in which case the correct stoichiometry can be obtained by rounding up to the next integer [65] or by Scatchard analysis [66]. Previous measurements demonstrated a higher polyQ binding affinity for 3B5H10 Fab than for MW1 Fab: K_D = 1.0 µM for 3B5H10 Fab binding to a Q_{22} or Q_{41} peptide [53] versus K_D = 2.2 µM for MW1 Fab binding to HD-39Q [58]. Therefore we used higher concentrations of MW1 Fab (1-10 µM) than 3B5H10 Fab (0.5-5 µM) in these experiments.

Complexes containing different ratios of Fab and HD-39Q were incubated together and injected onto a gel filtration column equilibrated with the appropriate Fab. A series of experiments with different concentrations of 3B5H10 Fab in the equilibration buffer were performed. For example, for 5 µM of 3B5H10 Fab in the equilibration buffer, the stoichiometry at which no peak or trough was observed was ~2.6:1 (Fig. 4a). Analysis with MW1 Fab yielded similar results, with the MW1 Fab:HD-39Q stoichiometry approaching 3:1 (Fig. 4b). Therefore for both antibodies, the stoichiometry of binding was determined as 3:1 Fab:HD-39Q.

Dynamic light scattering (DLS) revealed similar sizes for complexes of MW1 and 3B5H10 Fabs with huntingtin exon 1 proteins

To further investigate complexes of anti-polyQ Fabs bound to huntingtin exon 1 proteins, we used dynamic light scattering (DLS) to compare hydrodynamic radii (R_h) of 3B5H10

and MW1 Fabs alone, HD-16Q and HD-39Q alone, and complexes of Fab and

huntingtin exon 1 protein. Fab complexes with HD-16Q were prepared at a 1:1 Fab:HD-16Q molar ratio, while complexes with HD-39Q were prepared at a 3:1 molar ratio.

Concentrations of proteins and complexes varied from 1-7 mg/mL, higher than the

concentrations used for equilibrium gel filtration experiments. As expected, the R_h values

derived for the 3B5H10 and MW1 Fabs, which are globular proteins of similar

dimensions, were roughly the same, and both R_h values were smaller than the R_h values

for HD-16Q and HD-39Q (Table 2), consistent with the proposed elongated structures of

HD-16Q and HD-39Q [58] (Table 2). Notably, the R_h value for HD-16Q was smaller than

for HD-39Q, inconsistent with the compact structure proposed for expanded polyQ [52].

When complexed with huntingtin exon 1 proteins, both Fabs exhibited qualitatively

similar behavior: the R_h values were lower for the Fab complexes with HD-16Q than with

HD-39Q. Based on our gel filtration and SEC-MALS data (Fig. 2,4), the complexes being

examined by DLS were likely to be 1:1 Fab:HD-16Q complexes and a mixture of 2:1 and

3:1 Fab:HD-39Q complexes for both Fabs. Consistent with our other results, the 3B5H10

and MW1 Fabs did not exhibit different properties when binding to huntingtin exon 1

proteins as would have been expected if 3B5H10, but not MW1, recognized a pathologic

conformation of expanded polyQ.

Small-angle X-ray scattering revealed similar predicted characteristics for complexes of MW1 and 3B5H10 Fabs with huntingtin exon 1 proteins

We next repeated published small-angle X-ray scattering (SAXS) studies for HD-39Q

alone and complexed with 3B5H10 Fab [52], comparing results with SAXS data for

analogous complexes with MW1 Fab and HD-16Q. We collected SAXS data for samples

of 3B5H10 and MW1 Fabs alone, HD-16Q, HD-25Q, HD-39Q, and HD-46Q alone, and

complexes of each Fab with HD-16Q and HD-39Q. With the exception of the HD-46Q

alone sample, the scattering profiles for the Fabs, huntingtin exon 1 fusion proteins

alone, and Fab:HD-16Q and Fab:HD-39Q complexes showed ideal sample quality

characteristics (Fig. S2). Guinier analysis [68] indicated minimal aggregation for both

Fabs alone and Fabs complexed with HD-16Q and HD-39Q and for all huntingtin exon 1

fusion proteins alone (Fig. S3). Radii of gyration (R_g) determined by SAXS were

consistent with R_h values determined by DLS (Table 2). In particular, for measurements

of huntingtin exon 1 fusion proteins alone, we did not see substantially increased R_h or

R_g values for HD-39Q compared with HD-16Q, consistent with computational modeling

of polyQ in aqueous solution suggesting that radii of polyQ tracts of increasing lengths

are similar [69], but inconsistent with the structural toxic threshold model predicting a

conformational transition for expanded polyQ tracts [52]. The predicted molecular

weights of the complexes of 3B5H10 or MW1 Fab bound to HD-16Q in a 1:1 complex

were similar to the molecular weights calculated based on the extrapolated scattering

intensity at zero angle [70]. However, 3B5H10 Fab or MW1 Fab bound to HD-39Q

formed complexes with observed molecular weights larger than a 1:1 complex. Based on

molecular weight alone, these complexes could be 2:1, 3:1, or 2:2 Fab:HD-39Q

complexes. A 2:2 3B5H10 Fab:HD-39Q stoichiometry was postulated to account for

previous SAXS data [52]. However, our SEC-MALS and equilibrium gel filtration data

demonstrated that 3B5H10 Fab does not bind to HD-39Q in a 2:2 ratio (Fig. 2).

Therefore, we interpret our SAXS data for both 3B5H10 and MW1 Fab binding to HD-

39Q as evidence for mixtures of 2:1 and 3:1 Fab:HD-39Q complexes.

Kratky analysis was used to evaluate the relative degree of folding of each sample [71].

The Kratky plots for MW1 and 3B5H10 Fabs alone yielded bell-shaped peaks consistent

with globular proteins [72, 73] (Fig. 5a). By contrast, the Kratky plots for HD-16Q and

HD-39Q were broader, with less degrease at higher scattering angles, consistent with

flexible or unfolded proteins [72, 73] (Fig. 5b). Thus we found no evidence for a

conformational change occurring for HD-39Q relative to HD-16Q. Similarly, we found no

systematic differences for 3B5H10 versus MW1 Fab complexes with either HD-16Q or

HD-39Q (Fig. 5c,d). These results are consistent with both Fabs exhibiting similar

recognition properties for polyQ tracts.

Three-dimensional structures can be modeled into SAXS profiles; however, modeling is

limited by the one-dimensional nature of SAXS data, and more than one 3-D shape can

produce the same one-dimensional scattering profile [74]. We did not attempt to fit

atomistic models into the SAXS data as done in a previous study [52] because (*i*) the

complete 3-D structure of huntingtin exon 1 protein is unknown, (*ii*) the polyQ tract within

huntingtin exon 1 fusion proteins adopts flexible random coil structures in solution [58],

and (*iii*) the arrangements between polyQ tracts, the remainder of huntingtin exon 1, the

TRX fusion partner, and the His purification tag cannot be predicted. Nor did we assume

that the huntingtin exon 1 fusion proteins were dimeric, as also done for interpretation of

SAXS data involving HD-39Q bound to 3B5H10 Fab [52], because our SEC-MALS data

demonstrated that HD-39Q is monomeric in solution (Fig. 2). Instead, we used minimal

assumptions to generate *ab initio* models that predicted molecular envelopes from the

SAXS data for each of the well-behaved samples. We did not find notable differences

between 3B5H10:HD-39Q and MW1:HD-39Q complexes based on R_g, D_{max}, or shape or

volume of calculated envelopes (Fig. 6, Table 2), as would be predicted by the toxic

conformation model suggesting that 3B5H10, but not MW1, recognizes a compact conformation of expanded polyQ [52]. Instead, in agreement with experiments described above and in previous reports [53-55, 58], the SAXS results were consistent with recognition of multiple epitopes within a linear lattice of expanded polyQ by both Fabs.

5. DISCUSSION

The structure of huntingtin exon 1 protein in the preaggregation state, particularly the conformation of the expanded polyQ repeat, is hypothesized to be critical in understanding the pathogenesis of HD. However, the structure of the basic components of huntingtin exon 1 remains controversial. An X-ray crystal structure of a Q_{17} huntingtin N-terminal region fused to MBP showed that a short polyQ region could adopt either α-helical, loop, or random coil conformations [40]. The structure of a Q_{10} peptide bound to the anti-polyQ antibody MW1 revealed an extended structure [54]. Other recent work suggested that the polyQ repeat acts as a flexible hinge that exhibits reduced flexibility at extended polyQ lengths [75]. In the present study, we show that the binding properties of the anti-polyQ antibodies MW1 and 3B5H10 support the 'linear lattice' model for the structure of soluble polyQ in the context of a huntingtin exon 1 fusion protein. This model postulates that both normal and expanded polyQ tracts in the preaggregation state are random-coil structures, with expanded polyQ repeats containing more epitopes recognized by antibodies or other binding proteins than normal polyQ tracts [58]. Several lines of evidence, reported here and in previous publications [17, 53], have shown that 3B5H10 and MW1 IgGs can bind to a normal polyQ repeat, demonstrating that neither antibody preferentially recognizes a novel structure formed by expanded polyQ, but

instead both recognize a short stretch of polyQ. This conclusion is in contrast to other studies suggesting that 3B5H10 bound preferentially to expanded polyQ repeats of mutant huntingtin according to a 'structural toxic threshold' model, in which a conformational transition occurs in the polyQ repeat of huntingtin exon 1 protein at the pathologic threshold (>37Q) [52]. Instead our results agree with the conclusions of a recent study comparing the binding of anti-polyQ antibodies 1C2 and 3B5H10 to polyQ repeats [53].

To evaluate whether an expanded polyQ tract contains one epitope for anti-polyQ Fabs as predicted by the structural toxic threshold model, or multiple epitopes for the Fabs as predicted by the linear lattice model, we evaluated complexes using equilibrium gel filtration chromatography. Our results demonstrated that the stoichiometry of both 3B5H10 Fab:HD-39Q and MW1 Fab:HD-39Q complexes was ~3:1 Fab:HD-39Q for both Fabs, thus neither Fab preferentially recognizes a novel structure formed by expanded polyQ. Consistent with this result, we confirmed that both 3B5H10 and MW1 IgGs recognized unexpanded polyQ, in direct contradiction to the structural toxic threshold model. We also used SAXS and DLS to further study the conformation of normal and expanded forms of huntingtin exon 1 protein, alone and in complex with MW1 or 3B5H10 Fab. This allowed us to determine the globularity of the protein complexes and their approximate oligomeric states, which revealed striking similarities between MW1 and 3B5H10 Fabs, both when unliganded and when bound to HD-16Q or HD-39Q. Thus a combination of equilibrium gel filtration chromatography, DLS, and SAXS data are consistent with a linear lattice mode of recognition of unstructured polyQ for both 3B5H10 and MW1 antibodies. Sharing the same general ligand-binding properties is consistent with the high degree of sequence and structural similarity relating 3B5H10

and MW1 [53]: the variable regions are related by 53% (V_H domain) and 99% (V_L domain) sequence identity and a root mean square deviation of 0.59 Å for superposition of V_H-V_L regions of the 3B5H10 (PDB code 3S96) and MW1 (PDB code 2GSG) crystal structures (calculated for all Cα atoms). In addition, the antigen-binding sites of both antibodies include an unusual diagonal groove [53] shown to accommodate a single extended stretch of polyQ in an MW1-polyQ co-crystal structure [54], and thus unlikely to bind to a two-stranded β-hairpin structure of polyQ, as modeled for the HD-39Q interaction with 3B5H10 Fab [52].

Reduced penetrance is seen in patients with between 37 and 41 glutamine repeats in the huntingtin protein, which may be best explained by a quantitative change in a rate-limiting process in which the effects can be countered in some patients and not in others due to environmental or genetic modifiers. This reduced penetrance is consistent with a continuous linear lattice effect that is weak at lower polyQ repeat lengths and progressively stronger at larger repeat lengths. These results are relevant to potential therapeutic approaches to target soluble expanded polyQ in a lag period preceding aggregation. As we find no evidence for recognition of a specific conformation of expanded polyQ within huntingtin exon 1 proteins in either this study or previous studies [53-55, 58], we suggest that efforts to target expanded polyQ using monomeric binding partners are unlikely to be successful in discriminating polyQ stretches found in non-disease proteins such as transcription factors [76, 77] from expanded polyQ within mutant huntingtin exon 1. Instead, we suggest strategies in which reagents that recognize short stretches of polyQ are covalently linked to allow avidity effects to discriminate between short and expanded polyQ tracts.

6. MATERIALS AND METHODS

Protein expression and purification

Human huntingtin exon 1 encoded protein (comprising 91 amino acids when containing 16 glutamine residues) including different sized polyQ segments (Q16, Q25, Q39, and Q46) coded for by CAG or CAA/CAG repeats was expressed as a fusion protein with thioredoxin (TRX). Exon 1 fusion proteins were purified as previously described [58] with the following modifications: autoinduction was used to culture cells to high densities [78], and sonication was used for cell lysis. Purified proteins were flash frozen and stored at -80°C in 50 mM Tris (pH 8.0) and 150 mM NaCl.

MW1 IgG2b was purified from ascites fluid by protein A affinity chromatography (GE Healthcare). MW1 Fab was prepared by papain cleavage of MW1 IgG using a ratio of 1:25 (papain:MW1 by weight) for 30 minutes at 37°C. Fabs were separated from the Fc fragment using protein A affinity chromatography (GE Healthcare) and then further purified by gel filtration chromatography (Superdex 200 10/300 GL).

3B5H10 Fab was expressed and purified as previously described for other IgGs [79]. Briefly, the 3B5H10 light chain (LC) gene and a C-terminally 6x-His tagged heavy chain (HC) gene were subcloned separately into the pTT5 mammalian expression vector (NRC Biotechnology Research Institute), and 3B5H10 Fab was expressed by transient co-transfection of HEK293-6E (NRC Biotechnology Research Institute) cells and purified from supernatants using Ni^{2+}-NTA affinity chromatography and gel filtration chromatography (Superdex 200 10/300 or 16/60).

Protein concentrations were determined using 280 nm extinction coefficients of 80,830

M^{-1} cm^{-1} (3B5H10 Fab), 78,310 M^{-1} cm^{-1} (MW1 Fab), 14,180 M^{-1} cm^{-1} (TRX), 14,180 M^{-1}

cm^{-1} (HD-16Q), and 14,180 M^{-1} cm^{-1} (HD-39Q). Extinction coefficients were calculated

based on amino acid sequence using ProtParam

(http://web.expasy.org/protparam/protpar-ref.html).

Western and dot blot analyses

Equimolar amounts of purified huntingtin exon 1 protein and TRX were loaded and

separated on an Any kD Mini-PROTEAN gel (Bio-Rad Laboratories, Hercules, CA) at

175 V for 40 minutes, followed by transfer to a nitrocellulose membrane at 100 V for 1

hour. After blocking for 1 hour in TBST with 3% BSA, membranes were incubated

overnight at 4°C with 3B5H10 IgG (Sigma, St. Louis, MO), MW1 IgG (purified from

ascites), or rabbit polyclonal N17 huntingtin IgG [80] at 1:75,000 or, 1:10,000, or 1:7,500

respectively. Membranes were washed with TBST and incubated with HRP-conjugated

anti-mouse IgG (Jackson Immunoresearch Laboratories, West Grove, PA) or HRP-

conjugated anti-rabbit IgG (Jackson Immunoresearch Laboratories) for 1 hour at room

temperature. Membranes were washed again, and antibody binding was detected using

Amersham Enhanced Chemiluminescence (ECL) Prime Western Blotting Detection

Reagent (GE Life Sciences, Uppsala, Sweden). Western blots were imaged using a

ChemiDoc MP Imaging System (Bio-Rad Laboratories). Equimolar amounts of purified

huntingtin exon 1 protein and TRX were also analyzed by SDS-PAGE on an Any kD

Mini-PROTEAN gel (Bio-Rad Laboratories) and stained with Coomassie.

Equimolar amounts of huntingtin exon 1 fusion protein and TRX were serially diluted in

50 mM Tris (pH 8.0), 150 mM NaCl and spotted onto nitrocellulose membranes. Membranes were washed with TBST and probed with 3B5H10, MW1, or mouse monoclonal anti-TRX (Genscript, Piscataway, NJ) IgGs overnight at 4°C. HRP-conjugated goat anti-mouse secondary antibody (Jackson Immunoresearch Laboratories) was used to quantitate antibody binding to proteins on the membrane, and antibody binding was detected using Amersham ECL Prime Western Blotting Detection Reagent (GE Life Sciences). Densitometry of blots was performed using Image Lab 5.2.1 Software (Bio-Rad Laboratories). Densities were expressed as a ratio relative to the density observed for HD-16Q (western blots) or 20 pmol HD-16Q (dot blots).

SEC-MALS

Purified proteins or protein complexes were characterized by SEC-MALS to determine absolute molecular masses [64]. Proteins were concentrated to 1 mg/mL, passed through a 0.2 μm filter (Millipore), and injected onto a Superdex 200 10/300 GL gel-filtration chromatography column equilibrated in a buffer containing 20 mM Tris (pH 8.0) and 100 mM NaCl. The chromatography system was connected with an 18-angle light-scattering detector (DAWN HELEOS II; Wyatt Technology), a dynamic light-scattering detector (DynaPro Nanostar; Wyatt Technology), and a refractive index detector (Optilab t-rEX; Wyatt Technology). Data were collected every second at a flow rate of 0.5 mL/min at 25°C. Data analysis was carried out using the program ASTRA 6, yielding the molar mass and distribution of mass (polydispersity) of the sample.

Non-equilibrium gel filtration chromatography

Non-equilibrium protein interaction experiments were carried out on a Superdex 200 PC 3.2/30 gel-filtration column (GE Healthcare) equilibrated in a buffer containing 50 mM Tris (pH 8.0) and 150 mM NaCl. A final concentration of 7 µM HD-39Q was used for all experiments, with 3B5H10 and MW1 Fab concentrations varied to create Fab:HD-39Q complexes with final molar ratios of 0.25:1, 0.5:1, 1:1, 1.5:1, 2:1, and 3:1. 50 µL of each complex was injected and flowed through the column at 50 µL/min at room temperature. The absorbance of the eluent was monitored at 280 nm.

Equilibrium gel filtration chromatography

A Superdex 200 PC 3.2/30 (GE Healthcare) gel filtration column was equilibrated and run with equilibration buffer: 50 mM Tris (pH 8.0), 150 mM NaCl and a specific concentration of 3B5H10 Fab (0.5 µM, 0.75 µM, 1 µM, or 5 µM) or MW1 Fab (1 µM, 3 µM, 5 µM, or 10 µM). Complexes containing 0:1, 1:1, 2:1, or 3:1 molar ratios of a variable concentration of Fab:HD-39Q, where the concentration of HD-39Q was equal to the concentration of Fab in the equilibration buffer, were incubated for 30 minutes at room temperature in equilibration buffer and then injected onto the column. Chromatography was performed at a flow rate of 50 µL/min using a SMART micropurification system (Pharmacia), and the absorbance of the eluent was monitored at 280 nm.

Dynamic light scattering (DLS)

DLS measurements were conducted on a DynaPro® NanoStar[TM] (Wyatt Technology, Goleta, CA) at 25°C. All samples were purified using a Superdex 200 10/300 column in 50 mM Tris (pH 8.0), 150 mM NaCl prior to DLS and SAXS measurements, and the same sample preparations were used for both experiments. Fractions were pooled and concentrated to at least 2 mg/mL and filtered through 0.2 μm membranes (Millipore). Samples were equilibrated to 25°C prior to DLS measurements. Data were analyzed using Dynamics V7.1.2 software (Wyatt Technology) to calculate hydrodynamic radii (R_h).

Small angle X-ray scattering (SAXS)

Small angle X-ray scattering experiments were conducted at beamline 4-2 at SSRL using a Rayonix MX225-HE detector at a distance of 2500 mm, using 1.13 Å wavelength X-rays. Protein preparations for DLS measurements were used for SAXS data collection. For each protein or complex, scattering intensity was measured at four protein concentrations (0.5-7 mg/mL), collecting 10 exposures of 1 second each, covering a momentum transfer (q) range of 0.0047-0.375 1/Å. The scattering profile for the buffer was obtained in the same manner. Scattering curves collected from protein samples were corrected for background scattering using the intensity data collected from the buffer alone using SasTool [81].

SAXS scattering curves were scaled, high and low q regions of scattering curves were merged to extrapolate to infinite dilution, and Guinier analysis was performed using PRIMUS [82]. R_g values were calculated from Guinier plots. Scattering curves were overlaid to check for concentration-dependent effects on the scattering profile. GNOM

[68] was used to calculate pairwise distribution functions. Porod volumes were calculated from DATPOROD [83]. Molecular weight was calculated by using the formula $MM_p = I(0)_p/c_p*(MM_{st})/(I(0)_{st}/c_{st})$, where MM_p and MM_{st} are molecular masses of the protein and lysozyme standard, respectively, $I(0)_p$ and $I(0)_{st}$ are the scattering angles at zero intensity, and c_p and c_{st} are the concentrations [70]. The protein concentration, c_p, was calculated using the equation $A = \varepsilon L c_p$, where A is absorbance at 280 nm, ε is the molar extinction coefficient, and L is the path length. The extinction coefficient and theoretical molecular weight for a 3:1 and a 1:1 Fab:huntingtin exon 1 protein complex are different. Because we most likely observed a mixture of 1:1, 2:1, and 3:1 Fab:huntingtin complexes, we listed a lower limit (100% 1:1) and upper limit (100% 3:1) for potential complex sizes.

For each protein or complex, at least 10 *ab initio* models were generated using DAMMIF [84]. Models were superimposed and averaged using DAMMIN [85] and DAMAVER [86] in the ATSAS package [83], and resulting models were filled with dummy atoms.

Author contributions

G.E.O. and P.J.B. conceived the study; G.E.O. and D.M.N. performed protein expression, purification, western and dot blots, and SEC-MALS; G.E.O. performed equilibrium gel filtration assays; G.E.O., A.P.W., and P.J.B. analyzed the data; and G.E.O. and P.J.B. wrote the paper with all co-authors contributing to scientific planning and discussions.

ACKNOWLEDGEMENTS

7. FIGURE LEGENDS

Fig. 1. Biochemical analyses of huntingtin exon 1:3B5H10 IgG and huntingtin exon
1:MW1 IgG complexes. (a) Schematic of organization of human huntingtin exon 1-
thioredoxin (TRX) fusion proteins used in this study. The bar above the domain structure
represents the huntingtin exon 1 fragment with the polyQ tract indicated by a bracket.
N17, N-terminal 17 amino acid domain. PRR, proline rich region. (b) Western blot
analysis of 3B5H10 and MW1 IgG binding to huntingtin exon 1 fusion proteins with
variable numbers of glutamines. Both MW1 and 3B5H10 IgGs bound to huntingtin exon
1 proteins with normal and expanded polyQ repeats, but did not bind the TRX-tag control
(top panels). Equimolar loading of huntingtin exon 1 fusion proteins was verified by
blotting with the N17 antibody that recognizes the first 17 residues of huntingtin [80]

(bottom panel). Densitometry of bands in blots is shown in Fig. S1b. (c) Dot blot analysis of 3B5H10 and MW1 IgG binding to huntingtin exon 1 fusion proteins (top panels). Equimolar loading of huntingtin proteins was verified by blotting with anti-TRX (bottom panel). Densitometry results are shown in Fig. S1c. As huntingtin concentrations decreased, binding to short polyQ repeats of huntingtin exon 1 protein was reduced more than binding to long polyQ repeats for both anti-polyQ antibodies MW1 and 3B5H10.

Fig. 2. SEC-MALS profiles of huntingtin exon 1 protein and Fab:huntingtin exon 1 complexes. Complexes were prepared with ~3-fold molar excess of Fab for experiments with HD-39Q and ~equal molar ratios for experiments with HD-16Q and then injected onto a Superdex 200 10/300 GL gel filtration column. The differential refractive index (left axis) is plotted against elution time from a gel filtration column and overlaid with the molar mass determined for each peak (right axis).

Fig. 3. Non-equilibrium gel filtration chromatography analyses of Fab:HD-39Q complexes. Fabs of MW1 or 3B5H10 and HD-39Q were incubated at Fab:HD-39Q molar ratios of 0.25:1, 0.5:1, 1:1, 1.5:1, 2:1, and 3:1, and passed over a gel filtration column run under non-equilibrium conditions. An HD-39Q concentration of 7 μM was used for all experiments. (a) 3B5H10 Fab in complex with HD-39Q. (b) MW1 Fab in complex with HD-39Q.

Fig. 4. Equilibrium gel filtration chromatography analyses of Fab:HD-39Q complexes. MW1 or 3B5H10 Fabs and HD-39Q were incubated at the indicated molar ratios and passed over a gel filtration column run in an equilibration buffer containing the indicated concentrations of 3B5H10 or MW1 Fab. For each experiment, HD-39Q at the same concentration as the Fab in the equilibration buffer was incubated with Fab at Fab:HD-39Q molar ratios of 0:1, 1:1, 2:1, and 3:1 in the equilibration buffer and injected onto a column that had been equilibrated in the equilibration buffer. The peak eluting first corresponds to a Fab:HD-39Q complex. The second peak or trough occurs at the volume where free Fab elutes. (a) 3B5H10 Fab in complex with HD-39Q using 0.5 µM, 0.75 µM, 1.0 µM, and 5 µM 3B5H10 Fab in the equilibration buffer. (b) MW1 Fab in complex with HD-39Q using 1 µM, 3 µM, 5 µM, and 10 µM MW1 Fab in the equilibration buffer. The stoichiometry approached 3:1 Fab:HD-39Q for both Fabs.

Fig. 5. Kratky plot analyses of SAXS data for huntingtin exon 1 fusion proteins, MW1 and 3B5H10 Fabs, and complexes of Fabs and huntingtin exon 1 proteins. Each plot shows the intensity of scattering plotted as Iq^2 versus q, where q is scattering angle (Å^{-1}) and I is the scattering intensity. (a) For MW1 and 3B5H10 Fabs, each plot exhibited one maximum, indicating that these are globular proteins. (b) HD-16Q and HD-39Q showed a plateau at higher q values, suggesting that these proteins include disordered regions, with decreasing globular character as the polyQ repeat length increased. (c) Curves for 3B5H10:HD-16Q and MW1:HD-16Q complexes were similar, each exhibiting broad single peaks. (d) Curves for 3B5H10:HD-39Q and MW1:HD-39Q complexes were similar, each exhibiting a similar low plateau.
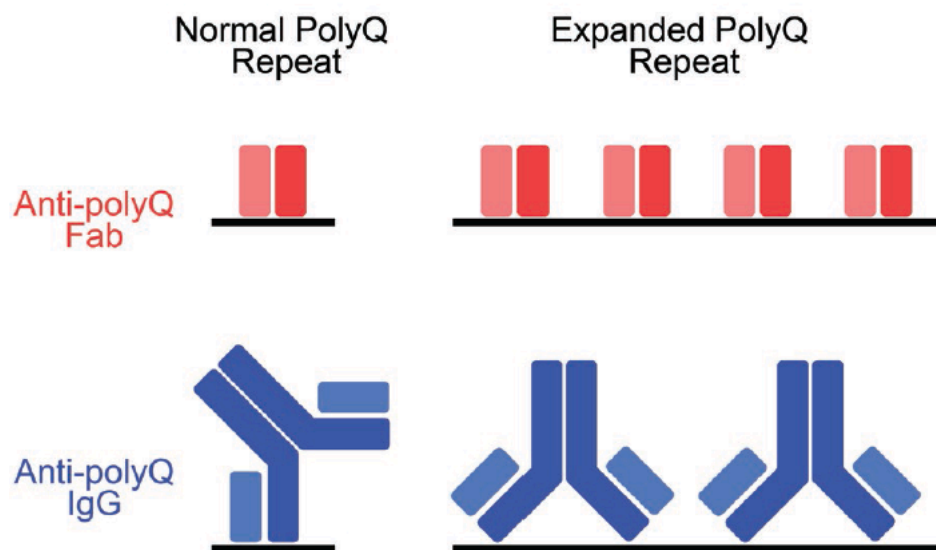
Fig. 6. *Ab initio* models derived from SAXS data for huntingtin exon 1 fusion proteins, MW1 and 3B5H10 Fabs, and Fabs:huntingtin exon 1 complexes. Calculated molecular envelopes filled with dummy atoms reveal the shapes of (a) 3B5H10 Fab, MW1 Fab, HD-16Q, HD-39Q, (b) 3B5H10 Fab:HD-16Q, MW1 Fab:HD-16Q, 3B5H10 Fab:HD-39Q, MW1 Fab:HD-39Q.

Fig. S1. Biochemical analyses of huntingtin exon 1:3B5H10 IgG and huntingtin exon 1:MW1 IgG complexes. (a) SDS-PAGE analysis of purified huntingtin exon 1 fusion proteins and thioredoxin under reducing conditions. (b) Densitometry analysis of western blots in Fig. 1b. (c) Densitometry analysis of dot blots in Fig. 1c.

Fig. S2. SAXS scattering profiles. Log(I) vs. q scattering profiles were superimposable when scaled for concentration, indicating high sample quality. HD-46Q showed variation at different concentrations, suggesting aggregation or other non-ideal behavior. The three lines indicate a serial dilution series of protein. Green is the scattering profile from the highest and light blue is the scattering profile of the lowest protein concentration.

Fig. S3. Guinier plots (left) and pair distribution functions (right). Guinier plots showed linear behavior in low q ranges, except for HD-46Q, which showed non-linear behavior indicative of aggregation. Pair distribution functions, P(r), were obtained from scattering

curves using AUTOGNOM [87], and fell smoothly to zero at D_{max}, the maximum linear

dimension of scattered particles.
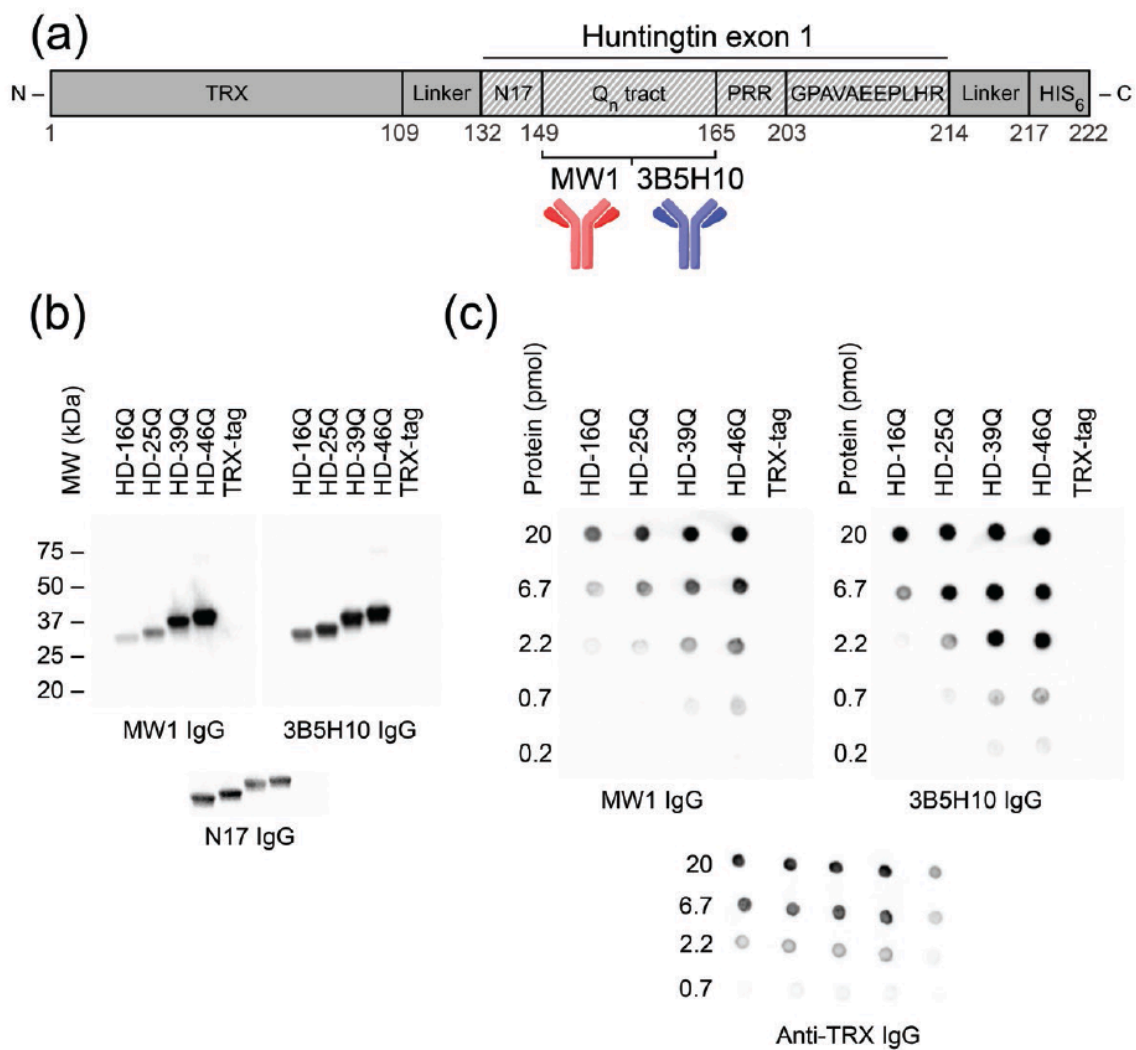
Normal PolyQ Repeat

Expanded PolyQ Repeat

Anti-polyQ Fab

Anti-polyQ IgG

Graphical Abstract
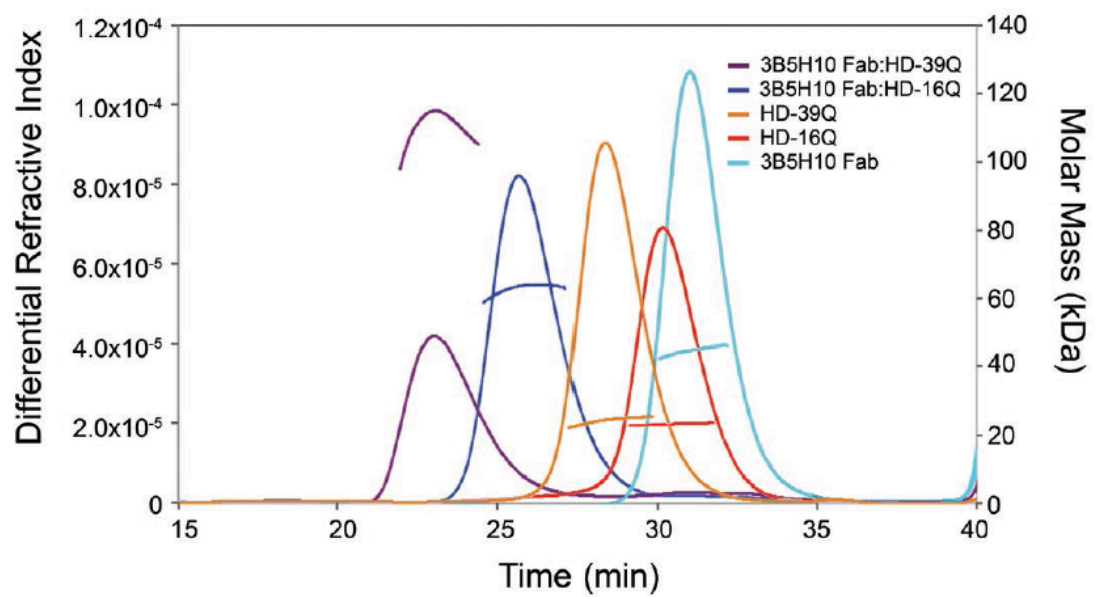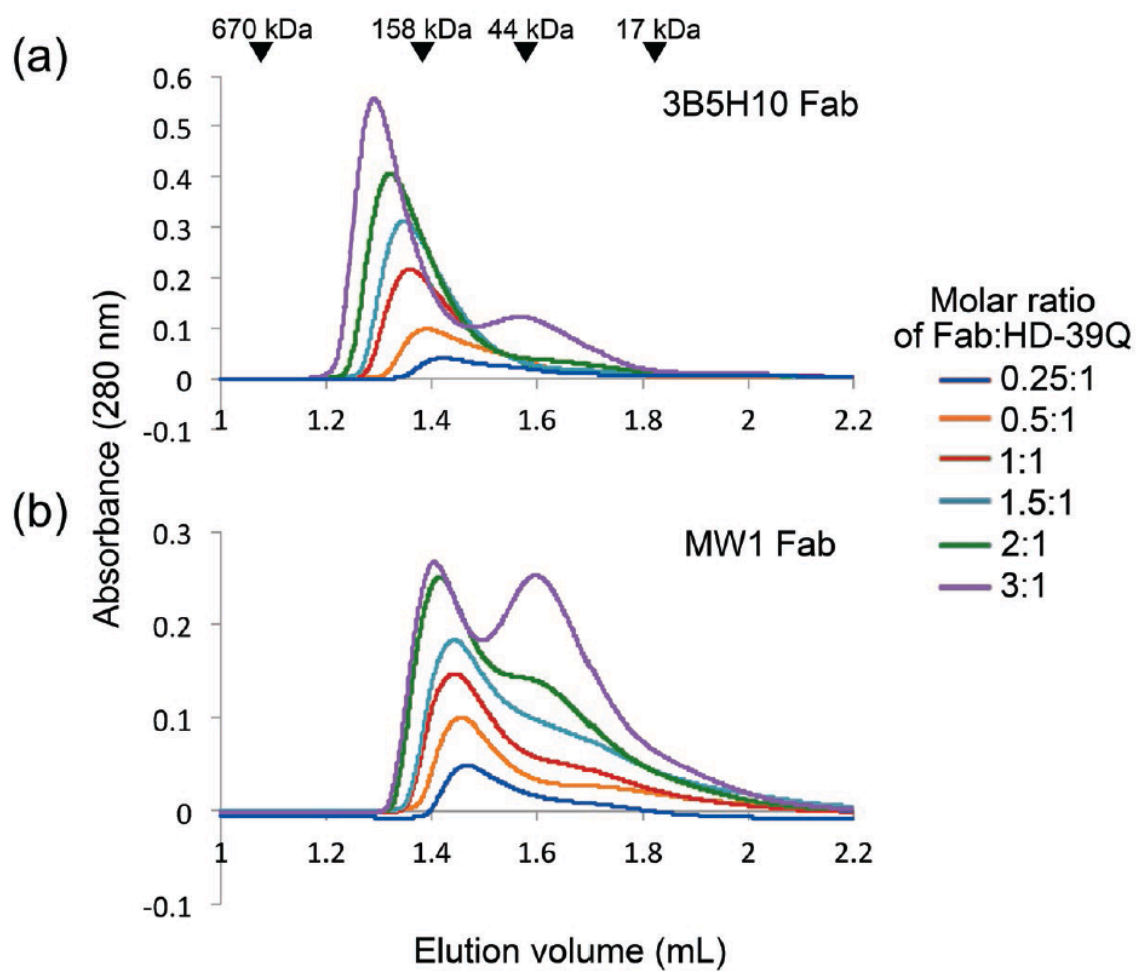
Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

(a)

100 Å

3B5H10 Fab   MW1 Fab   HD-16Q   HD-39Q

(b)

3B5H10
Fab:HD-16Q

MW1
Fab:HD-16Q

3B5H10
Fab:HD-39Q

MW1
Fab:HD-39Q

Figure 6

(a)

(c)

MW1 IgG Dot Blot

Protein (pmol)	HD-16Q	HD-25Q	HD-39Q	HD-46Q	TRX
20	1.0	1.4	1.8	2.4	0
6.7	0.3	0.8	1.2	1.5	0
2.2	0.1	0.2	0.6	1.0	0
0.7	0	0	0.1	0.4	0
0.2	0	0	0	0.04	0

3B5H10 IgG Dot Blot

Protein (pmol)	HD-16Q	HD-25Q	HD-39Q	HD-46Q	TRX
20	1.0	1.7	2.8	3.2	0
6.7	0.3	1.2	2.0	2.2	0
2.2	0.02	0.3	1.1	1.5	0
0.7	0	0.03	0.2	0.3	0
0.2	0	0	0.03	0.06	0

Anti-TRX IgG Dot Blot

Protein (pmol)	HD-16Q	HD-25Q	HD-39Q	HD-46Q	TRX
20	1.0	1.0	1.0	1.1	0.6
6.7	1.0	0.9	1.0	1.0	0.4
2.2	0.5	0.6	0.6	0.6	0.2
0.7	0.1	0.2	0.2	0.1	0.04
0.2	0	0	0	0	0

(b)

Western blots

	HD-16Q	HD-25Q	HD-39Q	HD-46Q	TRX
MW1 IgG	1.0	3.4	12.3	19.4	0
3B5H10 IgG	1.0	1.7	2.4	2.6	0
N17 IgG	1.0	1.0	0.8	0.9	0

Figure S1

3B5H10 Fab   MW1 Fab   HD-16Q   HD-39Q

3B5H10 Fab   3B5H10 Fab   MW1 Fab   MW1 Fab
+HD-16Q   +HD-39Q   +HD-16Q   +HD-39Q

HD-25Q   HD-46Q (less ideal)

$q \ (\mathring{A}^{-1})$

Figure S2

Figure S3

TABLES

Molecule or complex	SEC-MALS Molecular Mass (kDa)	
	Observed	Calculated
HD-16Q	23.1	24.3
HD-39Q	24.3	27.2
3B5H10 Fab	44.7	48.1
HD-16Q + 3B5H10 Fab	63.0	72.4 (1:1)
HD-39Q + 3B5H10 Fab	110.3	75.3 (1:1) 123.5 (2:1) 171.5 (3:1)

Table 1. SEC-MALS analysis of huntingtin exon 1 proteins and 3B5H10 Fab:huntingtin exon 1 complexes. The observed and calculated molecular masses of the proteins are listed.

Molecule or complex	MW (kDa), calculated	MW (kDa), observed	R_h (Å)	R_g (Å)	D_{max} (Å)	Porod volume (10^5) (Å^3)
3B5H10 Fab	48,118	54,900	33	28	98	0.6
MW1 Fab	47,330	49,300	33	26	87	0.6
HD-16Q	24,279	40,000	40	49	168	0.8
HD-39Q	27,226	21,700	47	52	164	0.9
3B5H10 Fab:HD-16Q	72,379	62,600	37	35	128	0.8
MW1 Fab:HD-16Q	71,609	71,200	41	40	141	0.9
3B5H10 Fab:HD-39Q	75,344 (1:1) 171,579 (3:1)	137,000 - 178,000	59	54	194	3
MW1 Fab:HD-39Q	74,556 (1:1) 169,216 (3:1)	125,000 - 162,000	52	50	173	2

Table 2. Structural parameters of huntingtin exon 1 proteins, MW1 and 3B5H10 Fabs, and Fab:huntingtin exon 1 protein complexes from SAXS and DLS. MW observed, R_g, D_{max}, and Porod volume were obtained from SAXS. R_h was obtained from DLS. The extinction coefficient and theoretical molecular weight for a 3:1 and a 1:1 Fab:huntingtin exon 1 protein complex are different; therefore two molecular weights are listed as observed for Fab:HD-39Q complexes. DLS and SAXS experiments were performed on the same samples. MW, molecular weight. R_h, hydrodynamic radius. R_g, radius of gyration. D_{max}, the maximum linear dimension of scattered particles.

Chapter 3: Biophysical Characterization of Huntingtin

Chapter 3: Biophysical Characterization of Huntingtin

1. Huntingtin Expression and Purification

Because huntingtin rapidly aggregates in solution (Chen et al., 2002; Colby et al., 2006), huntingtin exon 1 was expressed as a fusion protein with *E. coli* thioredoxin (TRX) to increase the expression of soluble huntingtin [88] in a pET-32a vector (Novagen) with a single N-terminal His$_6$ tag. Constructs contained the TRX gene, a linker segment (GSGSGERQHMDSPDLGTDDDDK), the HD exon 1 insert, and a His$_6$ tag. Proteins were overexpressed in BL21(DE3) cells (Novagen) by one of two methods: IPTG induction or autoinduction. Proteins were released by osmotic shock by using a modification of a published method [88], or by sonication. For osmotic shock, the cells were spun at 6,000 x g for 10 min at 4°C, and pellets were resuspended in 15 mM Tris-HCl, pH 8.0, and incubated 45 min with shaking at 4°C. Osmotically shocked cells were spun at 15,000 x g for 10 min at 4°C, and the supernatant containing the protein of interest was incubated with Ni nitriloacetate beads (Qiagen). Proteins were eluted in 500 mM imidazole, purified by gel filtration FPLC (Superdex-200; Amersham Pharmacia Biotech), and concentrated with an Amicon concentrator (Millipore). Proteins were stable for up to 2 weeks at 4°C in 50 mM Tris-HCl, pH 8.0 with 150 mM NaCl, 1 mM PMSF, and 1 mM EDTA.

Five different mutant forms of huntingtin exon 1 were successfully purified, ranging from 16 glutamines (normal) to 72 glutamines (juvenile Huntington's disease). Using a new lysis protocol involving sonication rather than osmotic shock, yields were increased from 1-2 mg/L to 10-15 mg/L for huntingtin with 16 glutamines. Additionally, this new protocol allowed for purification of huntingtin with 72 glutamines, which had not been successfully

purified previously.

2. Assay Development to Measure Huntingtin Aggregation

2.1 Thioflavin T (ThT) Assay for Amyloid Formation

Thioflavin T (ThT) is a dye that is widely used to visualize and quantify amyloid in Alzheimer's disease patients and in experiments on amyloid beta [89]. When ThT binds to structures that are rich in β sheets, such as in amyloid aggregates, the dye has an enhanced fluorescence signal that is thought to be caused by changes in charge distribution upon binding to highly ordered amyloid structure [90]. In initial experiments on ThT interaction with huntingtin exon 1, constructs were used with only 16 glutamine repeats, and no increase in ThT signal was seen. However, when constructs ranging from 16 to 72 glutamines were used, a striking correlation of increased ThT signal with increased polyQ length was seen (Figure 3-1). This had not been published in the literature at the time of the experiment, but was reported shortly after [91].
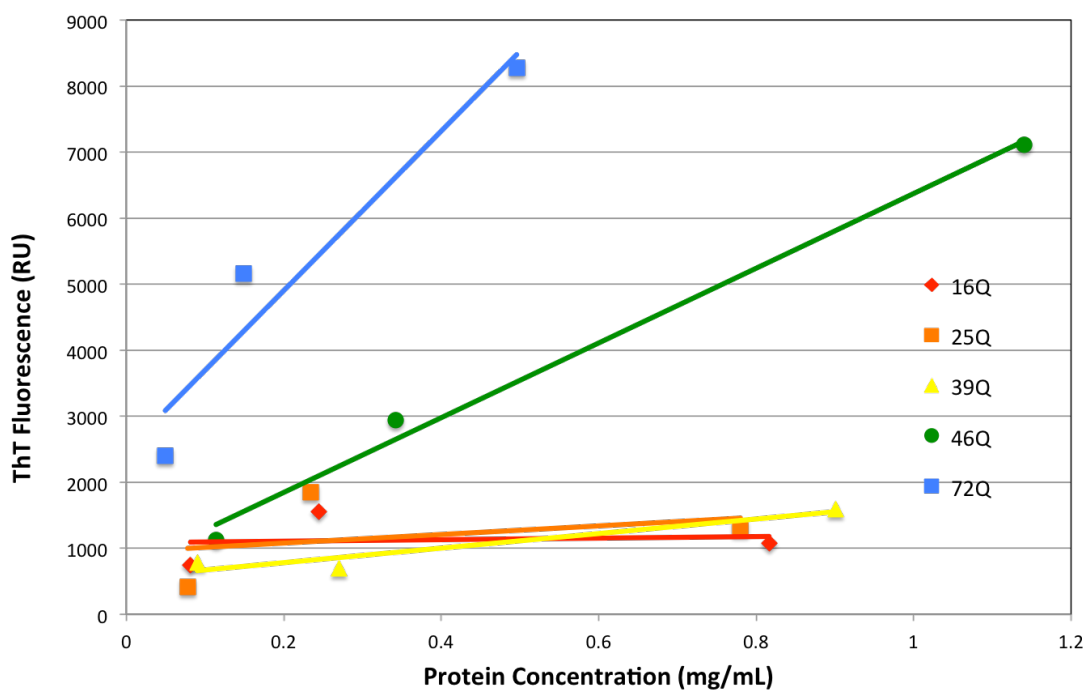
Figure 3-1. ThT assay on huntingtin exon 1 constructions with a range of polyQ repeats.

2.2 Turbidity Assay

Turbidity assays were performed by measuring the absorbance of huntingtin exon 1 at 405 nm on a Nanodrop UV-Vis spectrophotometer before and after enterokinase (EK) cleavage of the stabilizing TRX tag. After EK cleavage, the turbidity of the solution increased in a polyQ-repeat-length-dependent manner (Figure 3-2). The initial turbidity of the HD-46Q protein may be due to aggregation of the protein despite the presence of the TRX tag, as well as a small amount of spontaneous cleavage of the tag. The apparent aggregation of HD-72Q appears to be lower than for HD-46Q likely because the concentration of the HD-72Q protein was approximately half that of the other proteins due to difficulty purifying large amounts of this protein.
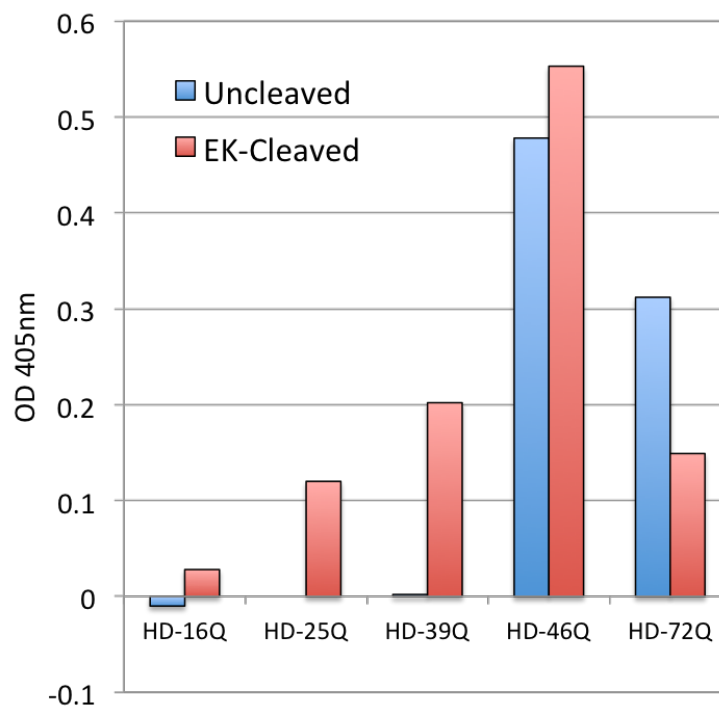
Figure 3-2. Turbidity assay results revealed significant aggregation after cleavage of the

TRX tag.

2.3 Dynamic Light Scattering

Dynamic light scattering (DLS) is a technique used to determine the size distribution

profile of small particles in suspension and is another method that is sensitive to

aggregate formation. DLS measurements were carried out on a Wyatt DynaPro light

scattering instrument to obtain size distributions and hydrodynamic radii (R_h). DLS was

used to observe the aggregation of soluble HD-16Q upon EK cleavage of the stabilizing

thioredoxin tag (Figure 3-3).
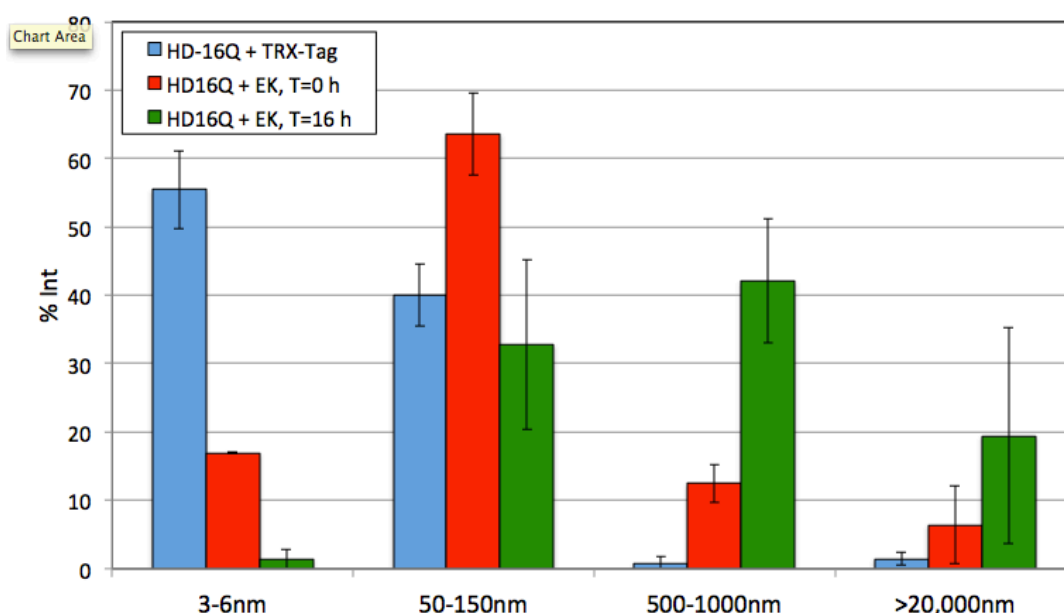
Figure 3-3. Cleavage of TRX tag by EK leads to aggregation over 16 hours as measured by DLS.

DLS confirmed the relative absence of aggregates in starting stocks of monomeric huntingtin exon 1 peptide, and also tested for the presence of aggregates upon EK cleavage. Unfortunately, DLS appears to break down at later stages of aggregation due to large particle size and complexity of the aggregate mixture; therefore, meaningful data could only be obtained for the first 16 hours post-EK cleavage.

2.4 Congo Red Staining

Congo red is a dye that has classically been used to indicate the presence of amyloid fibrils in histology slides [92]. When congo red-stained preparations are imaged under polarized light, apple-green birefringence is indicative of the presence of amyloid fibrils [93]. In this experiment, samples of huntingtin exon 1 with 46 glutamines (HD-46Q) were

stained after 30 day incubation at 4°C, and also after 3 days of shaking at 37°C. While

no staining was visible for the huntingtin sample kept at 4°C, some possible congo red

staining was seen after shaking at 37°C.

HD-46Q after 30 d at 4°C   HD-46Q after 3 d shaking at 37°C
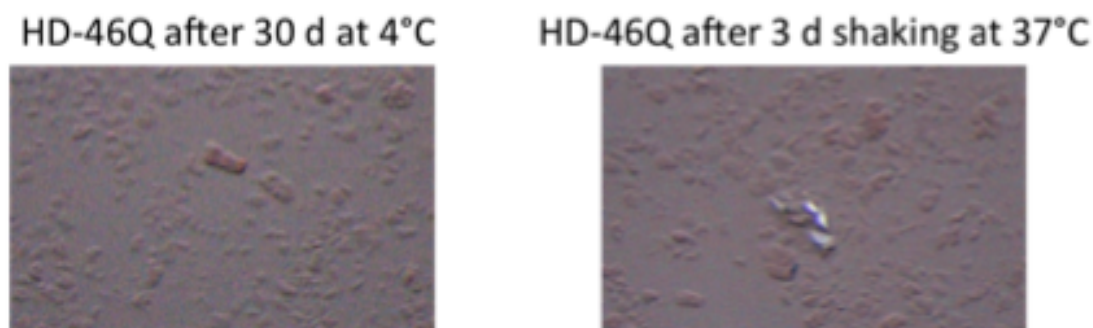
Figure 3-4. Congo red staining for presence of amyloid.

3. Huntingtin Crystallization Trials

Many previous experiments in the Bjorkman lab have been conducted in an attempt to

crystallize huntingtin, and while several researchers have documented the presence of

microcrystals, diffraction-quality crystals have never been obtained.

96-well crystallization trials were set up in the Caltech macromolecular crystallization

laboratory using the Art Robbins Gryphon Nano, a robot used to set up high-throughput

crystallization screens. The Rigaku Minstrel was used to automatically take images of

crystallization plates at given time points; these images were uploaded online for rapid

access to crystallization hits. Several conditions yielded spherolites or imperfect crystals,

which were looped using 30% glycerol as a cryoprotectant, frozen in LN_2, and sent to

the Stanford Synchrotron Radiation Laboratory (SSRL) for crystallographic data

collection on beamline 12-2. However, no diffraction was obtained.

Difficulty growing diffraction-quality crystals may be due to lack of strong secondary structure in this region or slow aggregation of huntingtin during crystallization. In 2005, attempts were made to co-crystallize short polyQ peptides with antibodies known to bind huntingtin. The Fv portion of one of these antibodies, MW1, was successfully crystallized from a solution containing both a polyQ peptide and MW1 Fv [54]. We proposed that the MW1 Fab might assist in crystallization of huntingtin in a similar fashion as the MW1 Fv by stabilizing crystal-packing surfaces.

Additionally, we attempted to crystallize huntingtin exon 1 using the unique environment provided by microgravity to increase the crystal quality and size compared to protein crystals grown on Earth (further discussed in Chapter 4). During protein crystal growth, convection in the solution surrounding growing protein crystals occurs as protein molecules from the surrounding solution move and assemble to become part of the growing crystal lattice. The movement of proteins toward a crystal and incorporation into a growing crystal lattice leads to a lower protein concentration, and a lower relative density, in the solution bordering the crystal than in the remainder of the solution. In an environment with gravity, the less dense solution has a tendency to rise, while the denser solution sinks, which causes convective fluid flow next to the growing crystal. These convective currents may have a negative effect on the quality of the crystal being formed because they may alter the position of the protein molecules as they become a part of the crystal lattice, producing disorder in the crystal lattice. These lattice imperfections adversely affect X-ray diffraction and subsequent analysis. However, in a microgravity environment, convective fluid flow is greatly reduced, allowing for slow, diffusion-driven protein crystal growth [94]. This suggests that microgravity may be an ideal environment for growing macromolecular crystals. In fact, previous experiments on

the International Space Station (ISS) have shown that reduction in buoyancy-driven convection and sedimentation leads to stabilization of growing crystals and more organized, larger protein crystals. Since the inception of NASA's protein crystal growth program in the late 1980s, more than 200 different proteins have been flown on Space Shuttle missions or on the Russian space station *Mir* [95].

We were encouraged that the X-ray crystal structure of insulin, a protein that significantly aggregates over time, and PPG10, a collagen-like protein with a polyproline segment similar in sequence to that in the unsolved region of huntingtin exon 1, were both successfully crystallized at high resolution aboard the ISS despite relatively poor crystal quality on Earth [96-98]. However, only 20-30% of protein crystal growth experiments carried out under microgravity have been successful. Differences in mass transport in microgravity and experimental setup limitations require that crystal growth conditions be optimized. To reduce potential problems with crystal growth, we coordinated with the Center for Advancement of Science in Space (CASIS) to evaluate appropriate conditions for crystallization of huntingtin.

We selected a team of engineers at the University of Alabama at Birmingham Center for Biophysical Science and Engineering (CBSE) as our service providers due to their extensive experience as implementation partners for previous space flight missions. The CBSE team provided the engineering, integration, operations, safety, and post-flight recovery support for our microgravity research. Additionally, they provided necessary documentation and expertise for safety review, toxicology assessments, and Certification of Flight Readiness Reviews (CoFR). We used the Hand-Held High Density Protein Crystal Growth (HDPCG) crystallization growth cell assembly for our huntingtin crystallization studies on the ISS. The HDCPG grows crystals by vapor diffusion and had

been used successfully in previous ISS crystallography experiments as part of the

CGCG-H (Commercial Protein Crystal Growth- High Density) and CMPCG (Commercial

Macromolecular Protein Crystal Growth) studies. The HDPCG requires few crew

resources, and it is compatible with existing NASA-provided cold stowage facilities. Each

Hand-Held HDPCG container includes 6 individual HDPCG blocks for a total of 30

sample wells per hand held device. The CBSE research team provided us with two

hand-held HDPCG assemblies for space flight, for a total of 60 sample chambers for

crystallization experiments. Two additional Hand Held HDPCG assembles were provided

as back up hardware or ground control hardware. An individual well of the HDPCG

apparatus is pictured below (Fig. 3-5).
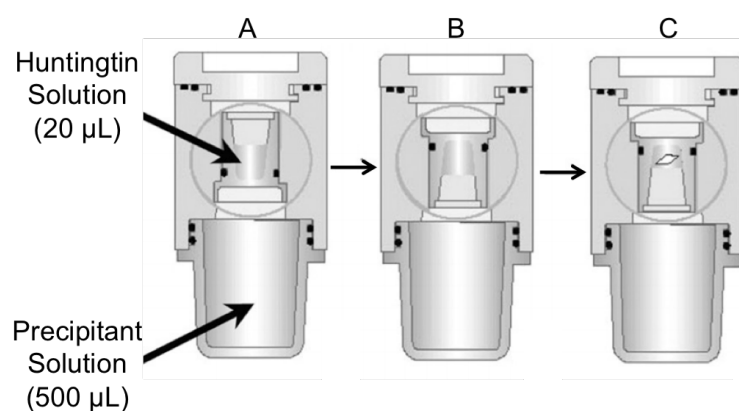
Figure 3-5. HDPCG hardware schematic. Adapted from [99].

The HDPCG apparatus was oriented in position A (Figure 3A) for launch. In this position,

the protein solution and precipitant solution were kept separate. On the ISS,

crewmembers turned the center barrel to position B to allow vapor diffusion between the

huntingtin protein solution and the precipitant solution, leading to crystal growth (Figure

3C). Prior to return to Earth, the barrel was returned to position A. Upon return to Earth,

crystals were harvested from several growth wells.

Chapter 4: Comparative Analysis of Anti-Polyglutamine Fab Crystals Grown on Earth and in Microgravity

Chapter 4: Comparative Analysis of Anti-Polyglutamine Fab Crystals Grown on Earth and in Microgravity

1. Abstract

Huntington's disease is one of nine neurodegenerative diseases caused by a polyglutamine (polyQ) repeat expansion. An anti-polyQ antigen-binding fragment, MW1 Fab, was crystallized both on Earth and on the International Space Station, a microgravity environment where convection is limited. Once the crystals returned to Earth, the number, size, and morphology of all crystals were recorded, and X-ray data were collected from representative crystals. Results generally agreed with previous microgravity crystallization studies. On average, microgravity-grown crystals were 20% larger than control crystals grown on Earth (hereafter called ground or control crystals), and microgravity-grown crystals had slightly improved mosaicity (decreased by 0.03°) and diffraction resolution (decreased by 0.2 Å) compared to ground crystals. However, the highest resolution and lowest mosaicity crystals were formed on Earth, and the highest-quality crystal overall was formed on Earth after return from microgravity. We determined the 1.8 Å crystal structure of MW1 Fab from an Earth-grown crystal, revealing an antigen-binding site that can accommodate extended stretches of glutamines.

2. Introduction

Huntington's disease (HD) is a progressive, late-onset neurodegenerative disease characterized by neuronal death resulting in choreaform movements, cognitive decline, and behavioral abnormalities [100]. There currently is no disease-modifying treatment or

cure for HD [26]. HD is found in individuals with an abnormally expanded N-terminal polyglutamine (polyQ) repeat in huntingtin, a 350 kDa protein of unknown function. The length of this polyQ tract influences the mechanisms of aggregation and associated binding kinetics, with increasing aggregation propensity with increasing polyQ tract length [101]. HD is completely penetrant when the polyQ repeat expands beyond a 37-glutamine threshold. However, the mechanism underlying this transition is unclear, and the relationship between polyQ-mediated aggregation, cellular toxicity, and HD symptoms has not been well characterized. The structure of polyQ with >37Q has been suggested to adopt a new β-sheet conformation [61] or a random coil conformation [69]. In the absence of an X-ray crystal structure, the nature of the expanded polyQ region and the interactions between mutant polyQ and other proteins are unclear.

The polyQ repeat in huntingtin is recognized by several monoclonal anti-polyQ antibodies, including MW1 [45]. Two recent papers measured huntingtin protein in cerebrospinal fluid of patients with HD using the MW1 antibody as one of a pair of antibodies in immunoprecipitation-flow cytometry [19] or FRET [100]. Information about the three-dimensional (3-D) structure of MW1 and its interactions with mutant huntingtin could assist in development of polyQ-length-based methods for quantitation of huntingtin in patients with HD in clinical trials [36]. X-ray crystal structures of the MW1 fragment variable (Fv) (the variable heavy and variable light domains; VH and VL) alone (PDB: 2GSG) and in complex with a $GQ_{10}G$ peptide (PDB: 2OTU, 2OTW) [54] demonstrated that the polyQ epitope could adopt a linear and extended conformation within a shallow groove of the MW1 Fv, and also demonstrated that the binding epitope for MW1 encompasses ~10 glutamines. Major structural changes occurred in MW1 Fv upon polyQ binding, including movement of amino acids in the third complementarity determining regions (CDRs) of the

heavy and light chain variable domains (CDRH3 and CDRL3) to allow hydrogen bond formation between the antigen-binding site and polyQ.

We previously attempted to crystallize huntingtin alone and complexed with Fabs (antigen binding fragments) of anti-polyQ antibodies, including MW1 Fab. While stable complexes of Fabs with the polyQ-containing N-terminal domain of huntingtin with 16, 25, 39, and 46 glutamines (HD-16Q, HD-25Q, HD-39Q, HD-46Q) formed in solution [102], no crystals of huntingtin or of a Fab:huntingtin complex could be obtained. Crystallization of the polyQ stretch of huntingtin is particularly challenging due to polyQ-length-dependent aggregation [103].

Reduced gravity environments may improve crystal formation for proteins that have a propensity to form a disordered aggregate at high concentrations such as huntingtin due to reduction in buoyancy-driven convection. In low convection environments, mass transport is primarily driven by diffusion. Aggregates and large impurities diffuse more slowly than monomers; therefore, monomers may have greater access to the surface of the growing crystal than aggregates in microgravity [104]. Microgravity has led to improved crystal size and quality for several proteins including insulin [105, 106], a protein that has been used as a model of amyloid formation [107], and PPG10, a collagen-like protein with a polyproline region similar to that flanking the polyQ region of huntingtin [96]. To gain further insight into the interaction between anti-polyQ Fabs and the polyQ repeat of huntingtin, we conducted protein crystallization experiments on the International Space Station (ISS); this provided an environment where protein crystals could grow undisturbed for several months in a microgravity environment. Crystallization studies in microgravity and parallel ground control tests were designed to examine whether we could generate high quality crystals of polyQ proteins, or crystals of a polyQ-containing protein in complex

with anti-polyQ Fabs. We were unable to generate crystals of poly-Q containing proteins on the ISS or in ground controls, but present here a comparative analysis of MW1 Fab crystals grown in microgravity and on Earth.

3. Methods

3.1. Protein expression and purification

MW1 Fab was purified as previously described [102]. Briefly, MW1 Fab was prepared by papain cleavage of MW1 IgG and protein A affinity chromatography (GE Healthcare, Little Chalfont, United Kingdom), with further purification by size-exclusion chromatography (SEC) (Superdex 200 10/300 GL) (Fig. S1). Purified protein was stored at 4°C in 50 mM Tris (pH 8.0) and 150 mM NaCl for up to three months. Other proteins used in our microgravity crystallization trials included human huntingtin exon 1-thioredoxin (TRX) fusion proteins (HD-16Q, HD-25Q, HD-39Q, HD-46Q) [58, 102], GFP-huntingtin [108], and the Fab from an MW1-related antibody called 3B5H10 [62]. The GFP-huntingtin construct was the gift of Dr. Robert Hughes (Buck Institute, Novato, CA). Each of these proteins was purified using Ni^{2+}-NTA affinity chromatography (GE Healthcare) and SEC (Superdex 200 10/300 or 16/60), flash frozen, and stored at -80 °C in 50 mM Tris (pH 8.0) and 150 mM NaCl, with the exception of GFP-huntingtin which was stored at -80 °C in 10 mM HEPES (pH 7.5), 50 mM NaCl, 10% glycerol, and 0.5 mM TCEP (tris(2-carboxyethyl)phosphine), a reducing agent. Purified full-length huntingtin protein [43] was the gift of Dr. IhnSik Seong (Massachusetts General Hospital) and was stored at -80 °C in 50 mM Tris (pH 8.0) and 100 mM NaCl. Protein concentrations were determined using 280 nm extinction coefficients of 78,310 M^{-1} cm^{-1} (MW1 Fab), 80,830

M^{-1} cm^{-1} (3B5H10 Fab), 14,180 M^{-1} cm^{-1} (HD-16Q, HD-25Q, HD-39Q, and HD-46Q), and 22,015 M^{-1} cm^{-1} (GFP-huntingtin). Extinction coefficients were calculated based on amino acid sequence using ProtParam [109]. A bicinchoninic acid (BCA) assay (Thermo Scientific, Rockford, IL) was used to determine total protein concentration of full-length huntingtin.

For crystallization of MW1 Fab alone, the protein was concentrated to 7 mg/mL in a centrifugal filter (EMD Millipore, Darmstadt, Germany). For MW1 Fab:HD-39Q co-crystallization trials, purified MW1 Fab and HD-39Q were incubated at a 3:1 molar ratio for at least 1 hr at 4°C. Crystallization conditions were optimized on Earth prior to microgravity experiments. Initial concentrations for all proteins are listed in Table S1.

3.2. Vapor diffusion crystallization

Vapor diffusion protein crystal growth experiments in microgravity were performed using the Hand-Held High Density Protein Crystal Growth (HDPCG) hardware (Fig. 1a) developed by the Center for Biophysical Sciences and Engineering at the University of Alabama, Birmingham. The Handheld HDPCG hardware was designed to reproduce a sitting or hanging drop crystallization experiment in a microgravity environment [110], and it has been used to crystallize dozens of proteins on the Space Shuttle or ISS [111-113]. Each aluminum Handheld HDPCG hardware unit held five HDPCG sample blocks (Fig. 1b), which were molded from Zeonor plastic. Each sample block consisted of six individual growth cells that each contained a single vapor diffusion crystal growth experiment (Fig. 1c). Each growth cell was isolated by triple O-ring containment.

For preparation of experiments prior to launch, 2.5 µL of protein solution was mixed with an equal volume of precipitant solution and placed in a 5 µL well in a growth cell. A separate reservoir in the same growth cell was loaded with ~500 µL precipitant solution at the desired final concentration. The precipitant solution was immobilized in the reservoir using an insert made of chromex, a porous absorbent material composed of ultra high molecular weight polyethylene (Porex, Fairburn, GA). A total of 120 crystallization experiments were set up, 60 for microgravity experiments and 60 replicates as ground controls. Initial conditions for all experiments are listed in Table S1. After all solutions were loaded, each well was checked to ensure there were no bubbles, and the sample blocks were then sealed. Ten HDPCG sample blocks total were loaded into two Handheld HDPCG hardware units. To prevent movement or mixing of solutions prior to or during launch, the sample block barrels were rotated 90° clockwise to "launch configuration" (Fig. 1d).

For experiment activation in-orbit, an astronaut rotated the sample block barrels another 90° clockwise using an Activation Tool (silver object in Fig. 1a) to establish an air path between the protein solution in the well and the precipitant solution in the reservoir (Fig. 1e). After activation, the experiments were stored undisturbed on the ISS. Initially, the protein solution contained an insufficient concentration of precipitant for crystallization, but as water vaporized from the droplet and transferred to the reservoir, the precipitant concentration in the protein well increased to an optimal level for crystallization in some experiments. Before return to Earth, an astronaut resealed the experiment by using the Activation Tool to rotate the sample block barrel counterclockwise 90° to "launch configuration" to turn the protein inserts away from the precipitant reservoir.

3.3. Timeline

All protein and precipitant stock solutions were prepared 7-52 days before launch. Proteins, precipitant solutions, and other equipment were transported to the Space Station Processing Facility, Kennedy Space Center (KSC), Florida between March 9 and April 10, 2014. Proteins were maintained at -20 °C (HD-16Q, HD-25Q, HD-39Q, HD-46Q, GFP-huntingtin, full-length huntingtin) or 4 °C (MW1 Fab, 3B5H10 Fab) during transport and storage. Before launch, flight HDPCG growth cells were prepared at 4 °C in a cold room previously used for astronaut food storage. HDPCG sample filling and hardware integration was completed on April 16, 2014 (Fig. 2). The HDPCG hardware units were turned over to Cold Stowage for integration on April 17, 2014 and were installed at 4 °C on the same day in a Double Cold Bag, a Nomex bag with vacuum insulation panels for passive thermal insulation [114]. Phase-change material Ice Bricks were added to the Double Cold Bag to maintain a 4 °C environment for the HDPCG hardware units prior to launch and during ascent to the ISS. Due to specific problems immediately before a launch, all experimental materials were prepared repeatedly for launch prior to the successful launch on April 18, 2014.

A total of 60 crystallization experimental trials in two Handheld HDPCG hardware units were launched from Cape Canaveral, Florida in an unmanned SpaceX-3 supply vehicle on April 18, 2014 at 15:25 Eastern Daylight Time (EDT). The Dragon capsule berthed with the ISS on April 21, 2014, the units were transferred to the ISS, and activation was completed at 07:50 EDT. After activation, the experiments were placed in a Minus Eighty Laboratory Freezer for ISS (MELFI) at 2 °C. Crystals were allowed to grow undisturbed in microgravity for exactly 6 months (183 days). The CASIS PCG hardware remained in the

MELFI until just prior to unberth of SpaceX-4. HDPCG deactivation and transfer to a Double Cold Bag at 4 °C for return was completed at 15:00 EDT on October 20, 2014.

The deactivated experiments descended to Earth on October 25, 2014 in SpaceX-4. The Dragon capsule landed in the Pacific Ocean on October 25, 2014 at 15:38 EDT, and the experiments were handed over for transfer to Caltech in Pasadena, California at 18:30 EDT on October 26, 2014. A temperature of 4±2 °C was maintained during all transport operations. The HDPCG hardware units were stored and imaged, and crystals were harvested, in a 4 °C room. There were no pre-flight, in-flight, or post-flight anomalies.

3.4. Ground control and comparison studies

The results of the microgravity crystallization experiments were evaluated using the best crystals that could be grown in identical conditions to the space flight experiments, termed "ground controls." Ground controls to replicate the conditions in the space flight experiments were set up at 4 °C at Caltech in HDPCG sample blocks with identical purified proteins, buffers, and precipitant solutions as used for the microgravity payload, with a seven day delay compared to flight experiments.

3.5. Crystal number, size and morphology analysis

Immediately upon return from the ISS, bright field images of all microgravity crystallization wells were taken at 40X magnification on an Olympus SZX16 microscope (Olympus Corporation, Tokyo, Japan) using a Canon DS126311 EOS Rebel camera (Canon Inc.,

Tokyo, Japan). The ground control wells also were imaged, with an approximate seven day delay. Fiji [115] was used to measure the long axis and short axis of each crystal using the line tool. The area of each crystal was calculated both by multiplying the long and short axis, and by either the polygon selection tool or freehand selection tool based on crystal shape. A MicroRuler™ (MiTeGen, Ithaca, NY) was used to scale images in microns. Crystal number, morphology, and visible area were recorded. Morphology was judged by the sharpness of crystal edges and shape of the crystal. Crystals were photographed again five months after return to Earth to evaluate changes in crystal size and morphology.

After crystals were harvested for X-ray diffraction data collection, the remaining crystals were imaged with a Korima PRS-1000 UV microscope (Korima Inc., Carson, CA) at 25 °C to distinguish protein crystals from salt crystals based on tryptophan fluorescence under UV light. Representative bright field and UV images of wells containing protein crystals are shown in Figs. S1 and S2, respectively. Some small crystals seen with UV microscopy were not visible using bright field microscopy; size and X-ray diffraction data were not collected for these crystals.

3.6. Crystallographic data collection, structure determination and refinement

Protein crystals were removed from HDPCG sample wells and briefly soaked in mother liquor solution supplemented with 7.5%, 15%, and then 30% glycerol before flash cooling in liquid nitrogen. X-ray diffraction data were collected from 155 total representative microgravity-grown and Earth-grown crystals at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2 using a Pilatus 6M pixel detector (Dectris). X-ray

diffraction data sets were collected for 32 crystals. X-ray diffraction data were unobtainable for some small crystals grown in microgravity due to technical limitations. The collected data sets were indexed, integrated, and scaled using XDS, a crystallographic data processing program [116, 117].

Data quality was analyzed using XDS and the PHENIX crystallography package [118]. Overall resolution limits of each dataset were estimated using $I/\sigma I > 1.50$ as well as $CC_{1/2}$ (correlation coefficient between two random halves of the data set) [119], where $CC_{1/2} > 0.3$. Average mosaicity was determined using the scaling program AIMLESS [120].

4. Results

4.1. MW1 Fab crystals formed in microgravity had increased size and decreased abundance compared to crystals grown on Earth, while morphology remained similar.

Crystals formed in several wells both in microgravity and on Earth. All crystals formed in wells containing MW1 Fab alone or MW1 Fab with HD-39Q. No crystals were observed in any wells containing full-length huntingtin or GFP-huntingtin; instead, the presence of UV-fluorescent aggregate was noted in these wells. Crystals were observed in 10 of 60 wells in the microgravity HDPCG wells, and 9 of 60 wells in the ground control HDPCG wells (Fig. 3b, Table 1). Of the wells containing crystals, one was found only in the flown samples, i.e., there were no crystals in the corresponding ground control well. This well contained needle crystals that did not diffract beyond 5 Å. No crystals were observed in the ground control wells without also being observed in the sample wells in the flown HDPCG growth cells.

The morphologies of crystals tended to be similar in microgravity and corresponding ground control wells. Most crystals had sharp edges, although several wells contained crystals with plate or needle morphologies (Fig. 3a). In wells 1 and 2, crystals grew larger and were thicker than in parallel ground control wells (Fig. S2). Also, in well 29, ground control samples grew only microcrystals from which no diffraction data could be collected, while crystals were larger in the ISS samples, and diffraction data could be recorded. Conversely, crystals in well 35 grew larger on Earth than in microgravity, demonstrating that the size effect was not consistent between microgravity and Earth conditions. Microgravity-induced changes in crystal morphology have been previously reported [121-124]; however, the morphologies seen in our experiments have all been seen previously for the analogous crystals on Earth and did not represent new crystal forms.

Quantitative analyses of crystal number and visible crystal area from microscopy images demonstrated that significantly fewer crystals of a size suitable for diffraction (> 20 μm in each dimension) were grown per well in microgravity compared with ground controls (Fig. 3b, Table 1). To our knowledge, this is the first microgravity study to statistically analyze differences in crystal size. Two wells in each condition formed microcrystalline precipitate, microcrystals, or stacks of needle crystals; these crystals were not included in crystal number and size analysis, which could have changed the data dramatically. Microgravity well 10 contained a large stack of needle crystals (~250 crystals) that could not be accurately counted, and microgravity well 35 had ~3,000 microcrystals that were below the threshold of 20 μm in each visible dimension In ground controls, well 29 contained ~46 microcrystals, and well 30 contained ~230 microcrystals; these also were not part of the analysis.

Analyses of the visible crystal area showed that crystal size was larger in microgravity (Fig. 3b, Table 1). The largest microgravity-grown crystal was larger than the largest ground control crystal. This agrees with previous reports of increased crystal size and decreased crystal number in microgravity [111]. Convection on Earth may increase the rate of nucleation in solutions containing crystals that are growing, termed secondary crystal nucleation, due to flow of partially nucleated proteins from growing crystal surfaces [125]. Increased secondary nucleation would theoretically yield more and smaller crystals in Earth compared to microgravity.

We recorded images of crystals immediately upon receipt of experiments from the ISS, but no photographs of crystals could be taken on-orbit during microgravity crystal growth due to incompatibility of current ISS microscope hardware with the HDPGC growth cells, particularly due to the variable opacity of the chromex insert. To keep ground controls matched to microgravity experiments, ground controls also were not imaged during the six-month duration of the experiment. Thus, the optimal time for crystal nucleation and growth in microgravity is unclear. A different time frame may have produced more crystals. Additionally, we found that crystals grew several months after return to Earth in two wells sent to the ISS that did not contain visible crystals upon initial return to Earth, and 3-D crystal growth also occurred in one well that had only irregularly-shaped crystals upon initial return to Earth. This was confirmed by analysis of the second set of images taken of all crystallization wells five months after the experiment returned to Earth. Although crystal nucleation may have taken place on the ISS, we have categorized these as "ground" crystals because most crystal growth occurred in a $1g$ environment. These ground crystals were not included in morphology, size, or number analyses, but X-ray diffraction data were collected for several of these crystals. To our knowledge, this is the first microgravity

experiment that has reported crystal growth post-flight.

3.2. Microgravity-grown crystal showed improved X-ray diffraction resolution on average, but the highest resolution and lowest mosaicity crystals grew on Earth

DDiffraction from microgravity and ground crystals was evaluated for resolution limit and mosaicity. High resolution is desirable to allow interpretation of the chemical details of a protein structure. Mosaicity is defined as the full width at half maximum of diffraction peaks. High average mosaicity values are a sign of a poorly ordered crystal and are generally undesirable because wider diffraction maxima can result in overlapping reflections.

MW1 Fab crystals were obtained in three space groups (Table 2). Crystals of MW1 Fab (space group $P2_1$, a = 42 Å, b = 72 Å and c = 89 Å; β = 91°; one molecule per asymmetric unit) were obtained upon mixing MW1 Fab at 7 mg/mL with 0.1 M sodium citrate tribasic dihydrate (pH 5.0) and 18% w/v PEG 20,000. This condition yielded crystals that diffracted to 1.6-2.3 Å, the highest resolution of any MW1 Fab crystals. These crystals formed after return from the ISS. Crystals that diffracted to 3.0 Å were obtained in this space group and unit cell by mixing MW1 Fab at 7 mg/mL plus HD-16Q at 7 mg/mL with 0.2 M magnesium chloride hexahydrate, 0.1 M sodium citrate tribasic dihydrate (pH 5.0) and 10% w/v PEG 20,000, at 4 °C, also in a well where crystal formation occurred after return from the ISS. Crystals did not form in microgravity or in ground controls in space group $P2_1$, so the effects of microgravity on MW1 Fab crystals in this space group is unclear. For all crystals grown in the presence of an HD protein, only MW1 Fab crystallized.

Crystals of MW1 Fab in a second space group ($P2_12_12_1$, a = 44 Å, b = 55 Å and c = 299 Å; one molecule per asymmetric unit) were obtained upon mixing MW1 Fab at 7 mg/mL plus

HD-39Q with one of four precipitant solutions at 4 °C: 0.1 M sodium acetate trihydrate (pH 4.5), and 30% w/v PEG 300; 1.8 M ammonium sulfate, 0.1 M BIS-TRIS (pH 6.5) and 2% v/v PEG MME 550; 0.2 M magnesium chloride hexahydrate, 0.1 M sodium citrate tribasic dihydrate (pH 5.0), and 14% w/v PEG 20,000; or 0.2 M magnesium chloride hexahydrate, 0.1 M sodium citrate tribasic dihydrate (pH 5.0), and 18% w/v PEG 20,000. Over 70% of crystals that formed in microgravity or ground control wells that diffracted to beyond 5.0 Å were in this space group and unit cell. The highest resolution crystals in this space group were formed in ground controls; however, the average resolution improved by 0.2 Å and the average mosaicity of the diffraction data decreased by 0.03° in microgravity wells compared to ground control wells. If the analysis is limited to only crystals that were looped immediately upon return from the ISS, the average resolution improved by 0.4 Å and mosaicity decreased by 0.07° in microgravity wells compared to ground control wells. This is consistent with previous findings of improvements of 0.2-0.4 Å resolution for crystals grown in microgravity compared with ground control crystals [126].

Crystals of MW1 in a third space group (P622, a = 189 Å, b = 189 Å and c = 64 Å; γ = 120°; one molecule per asymmetric unit) were obtained upon mixing MW1 Fab at 7 mg/mL with 0.1 M sodium citrate tribasic dihydrate (pH 5.5), 16% w/v PEG 8,000 or 0.1 M sodium citrate tribasic dihydrate (pH 5.0) and 18% w/v PEG 20,000 at 4 °C. Crystals formed both in microgravity and in ground controls in this space group; however, the resolution was poor for all diffracting crystals, ranging from 3.2 Å to >8 Å. The highest resolution crystals were formed in microgravity, with an improvement of 0.8 Å in resolution and a decrease of 0.07° in mosaicity for the highest resolution microgravity crystal in this space group compared to the highest resolution ground control crystal in the same space group.

Previous studies have addressed whether particular space groups are more amenable to

crystallization in a convection-free environment, and found that no space group appeared to be more amenable to improvement in microgravity [95]. Our data support this conclusion.

5. Discussion

Microgravity affects crystal growth by preventing sedimentation and decreasing buoyancy-driven forces on the crystal, thereby creating a more stable depletion zone around a growing crystal [127]. Through these mechanisms, microgravity-grown crystals have been reported to have increased resolution, decreased mosaicity, and increased crystal volume compared to Earth-grown controls [128, 129]; however, some negative studies have been published, and the benefit of microgravity crystallization has been fiercely debated [130] since the first microgravity crystallization experiment [131]. In our studies, we found microgravity-grown crystals of MW1 Fab showed a slight increase in size and an improvement in resolution and mosaicity when compared to Earth-grown crystals in one space group in agreement with some published results [132, 133]; however, the highest overall resolution X-ray data were obtained from a crystal grown on Earth.

Improvements in crystal volume in microgravity are driven by two factors: 1) decreased sedimentation, leading to fewer fused aggregates and increased uniformity of crystals [95], and 2) Maragoni convection and transient accelerations, which also promote increases in crystal volume under microgravity conditions, despite possible deleterious effects on crystal packing [124, 134, 135]. Marangoni convection arises in vapor diffusion experiments, and occurs at the phase boundary between the concentrated solution of protein and the air. Concentration gradients that form during crystallization or precipitation

result in differences in surface tension, which lead to different rates of transfer of vapor at the surface of the protein drop. While Marangoni convection is not the predominant method of mass transfer in crystallization experiments on Earth, it becomes an important factor when buoyancy-induced convection is substantially reduced in microgravity [134]. An analysis of microgravity experiments found greater improvements in crystal quality in liquid-liquid diffusion experiments compared to vapor diffusion experiments [95], and may have been a factor in our microgravity vapor diffusion experiments. Additionally, transient accelerations on the ISS, such as residual accelerations from crew movement and exercising, vibrations imposed by equipment operating near crystallization experiments, and vehicle accelerations from reboost or collision avoidance maneuvers (CAMs), could have also led to deviation from a true microgravity environment. In a perfect microgravity environment, crystal nucleation occurs but growth is slowed because nutrients are depleted in the area of the crystal-solution interface. Brief accelerations could stir the solution to replenish nutrients around the crystal to help it grow larger. On the Space Shuttle, continuous visual feedback on the crystal growth has been used to correlate increased crystal growth with increases in transient accelerations [135]; no similar studies have been published for microgravity crystallization experiments on the ISS. We found that microgravity increased mean MW1 Fab crystal size, although this effect was small.

In contrast with crystal size, resolution improvements in microgravity are due to improved short-range order within a crystal, while mosaicity improvements are driven by improved long-range order; both of these are caused by reduction in buoyancy-driven convection. Marangoni convection is a limiting factor in these improvements. The combination of these effects is complex.

In parallel with our microgravity experiments, we analyzed possible alternatives to microgravity crystallization. The Grashof number (Gr) describes the relationship between inertial and viscous forces, and gives insight into the parameters such as gravity that determine the movement of solutions around a growing crystal:

$$Gr = \frac{\rho U_b L_0}{\eta}$$

where ρ is density, U_b is a measure of buoyant velocity dependent on gravity and concentration gradient, L_0 is length, and η is shear viscosity [136]. When Gr is small, the inertial forces are small compared to viscous forces, and fluid interfaces spread by diffusion. Gr is decreased in microgravity, but other environments also decrease convection, e.g., increased shear viscosity through use of agarose gels [137] or high molecular weight polyethylene glycol (PEG) [138]; decreased length through use of microfluidics, or even magnets [139]. Several proteins have been crystallized in a centrifuge under hypergravity [140] to nucleate crystals at low concentrations of precipitants. However, our experiments to crystallize huntingtin within gels and microfluidic crystallization trays did not improve crystallizability (data not shown).

While our experiments show that ISS is a potential platform for crystal growth, crystallization of proteins in space remains a challenge. Given the expense and time involved in crystallization trials in microgravity, future experiments should consider the potentially deleterious effects of Marangoni convection on vapor diffusion crystallization experiments. Additionally, a comparison of microgravity-grown crystals with the best crystals obtainable through ground-based methods is necessary to realistically determine the relative value of microgravity protein crystallization.

Funding

Author contributions

G.E.O. and P.J.B. conceived the study; G.E.O., A.O., and D.M.N. performed protein expression and purification; G.E.O. and J.T.K. collected X-ray diffraction data and solved the MW1 Fab structure; G.E.O., A.M., B.M., R.R., and D.C. set up experiments at KSC prior to SpaceX-3 launch; G.E.O. and D.C. retrieved samples after SpaceX-4 return to Earth; D.M.N. analyzed crystal size data; and G.E.O. and P.J.B. wrote the paper with all co-authors contributing to scientific planning and discussions.

Acknowledgements

figures; the Protein Expression Center at Caltech for purification of 3B5H10 Fab; the

scientific staff of SSRL beamline 12-2 for assistance with X-ray diffraction experiments;

and members of the Bjorkman Lab for critical reading of the manuscript. We thank the

team at CASIS for support, particularly April Spinale and Ken Shields. We would also

like to thank Scott Slack and his film team at High Impact and Dr. Edward Wild for

helping to make our research accessible to a broad audience. We also thank the two

astronauts who activated and deactivated our experiments on the ISS.

6. Figures
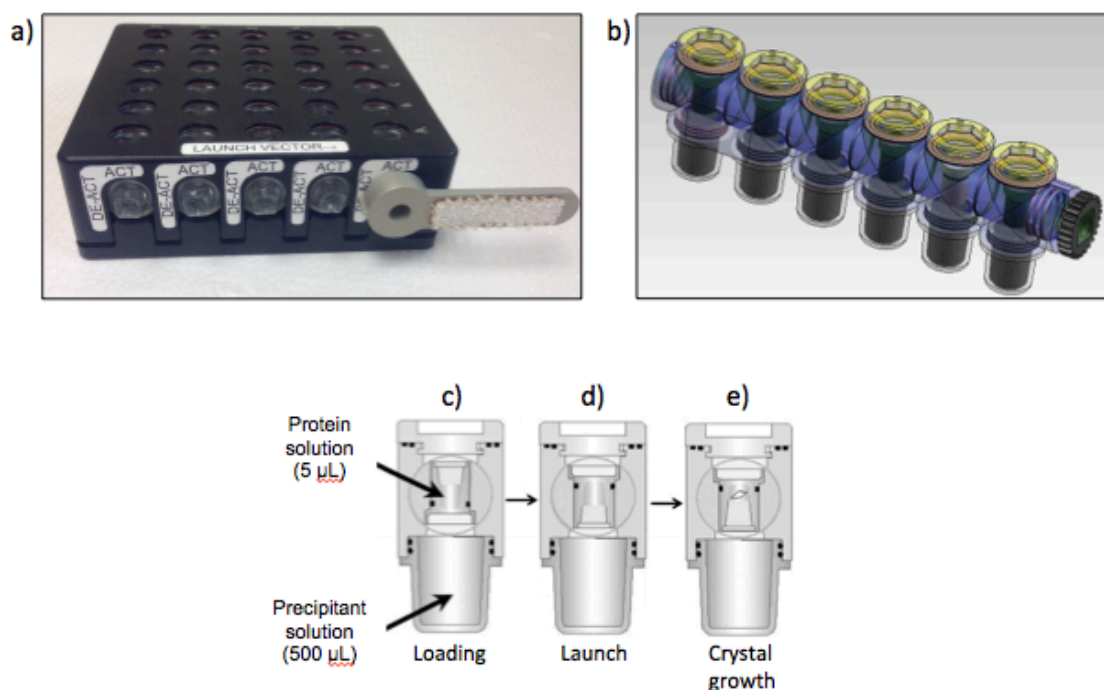
Fig. 1. Schematics of the HDPCG device used for microgravity crystallization experiments:

a) Handheld HDPCG assembly, b) HDPCG sample block, c) HDPCG growth cell in

loading configuration, d) HDPCG growth cell in launch configuration, e) HDPCG growth

cell in microgravity crystal growth configuration.
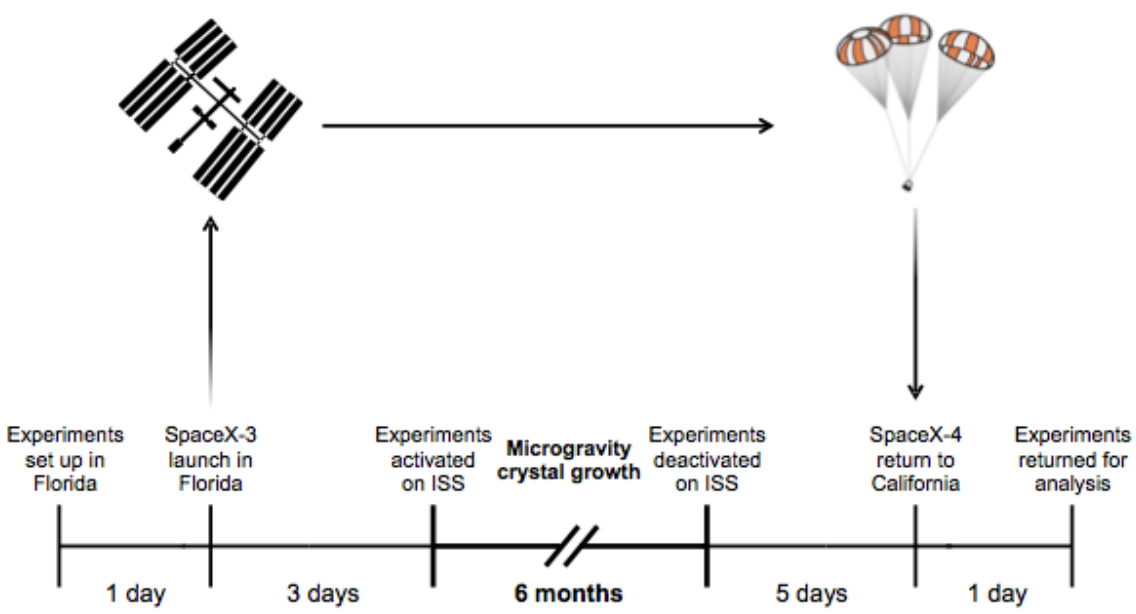
Fig. 2. Timeline of microgravity experiments.
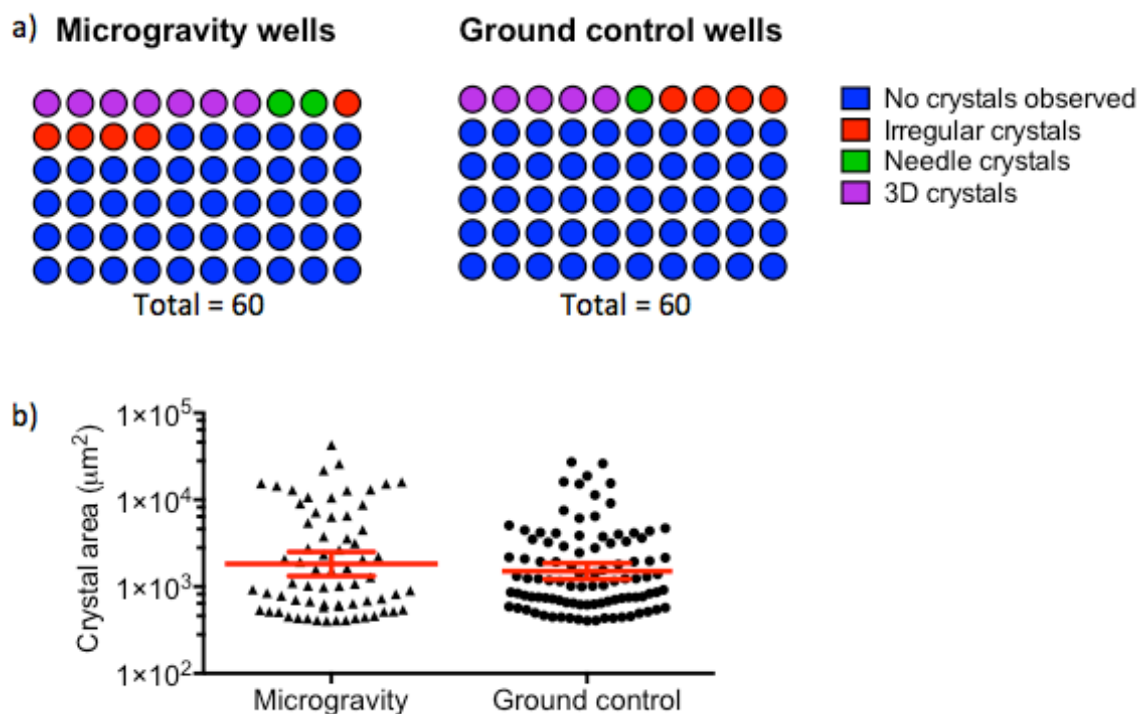
Fig. 3. Comparison of crystal morphology and size. Protein crystals grew in both microgravity and ground control conditions. a) Morphology of crystals and number of wells containing crystals. b) Area of crystals greater than 400 μm^2 grown in microgravity (n=67) and Earth (n=97). Data shown are geometric means with 95% confidence intervals. The geometric mean is suitable for data that range over several orders of magnitude [141].

Well	Microgravity	Ground control	Well	Microgravity	Ground control
1			28		
2			29		
10			30		
17			32		
18			34		
27			35		
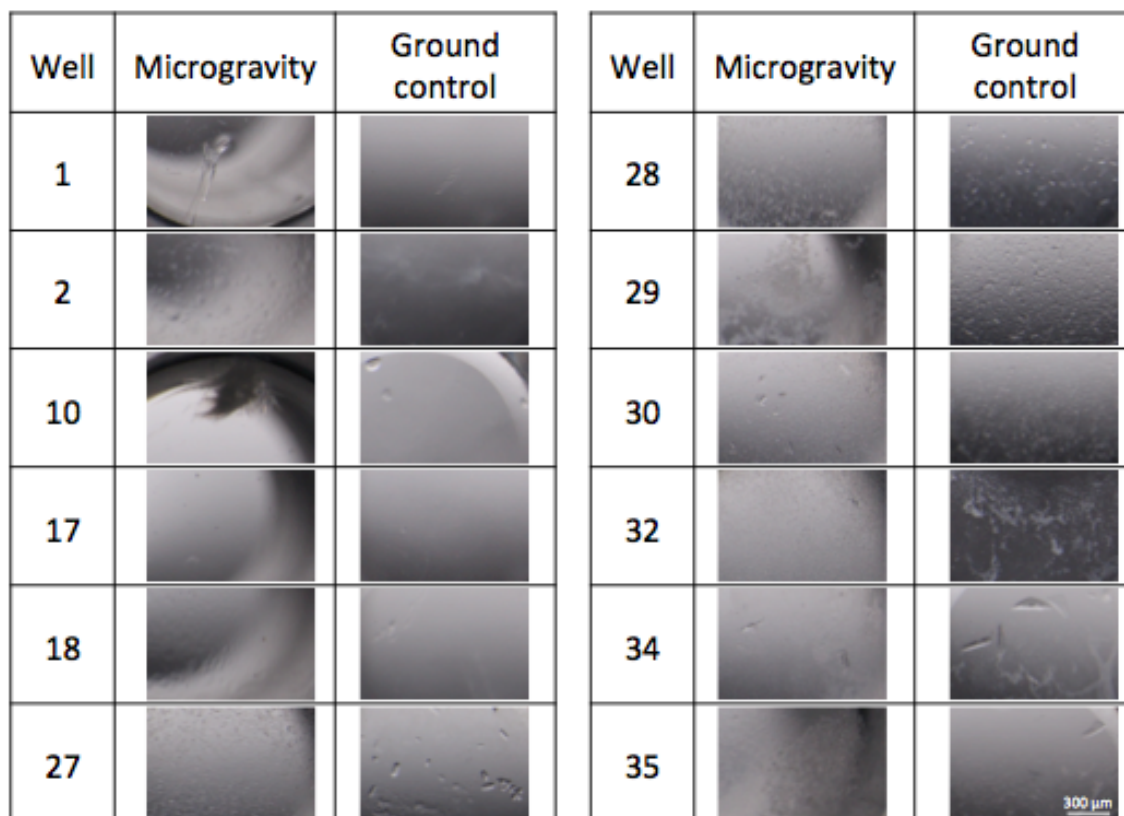
Fig S1. Representative crystal images taken using bright field microscopy immediately upon completion of experiment.

Well	Microgravity		Ground control	
1				
2				
10			*	
17	*			
18			*	
27				300 μm
	UV	Vis	UV	Vis

*Crystals formed post-flight

Well	Microgravity		Ground control	
28				
29				
30				
32	*		*	
34				
35				300 μm
	UV	Vis	UV	Vis
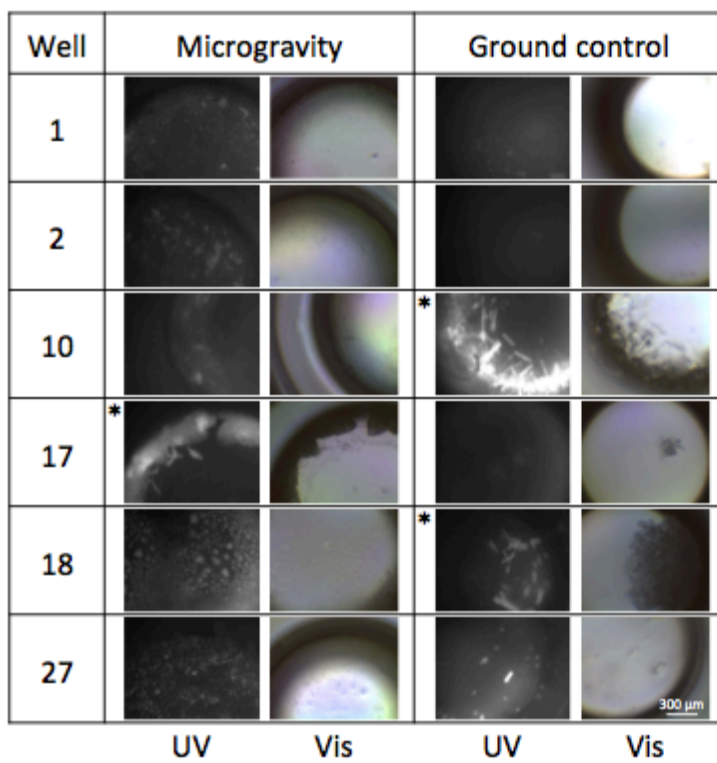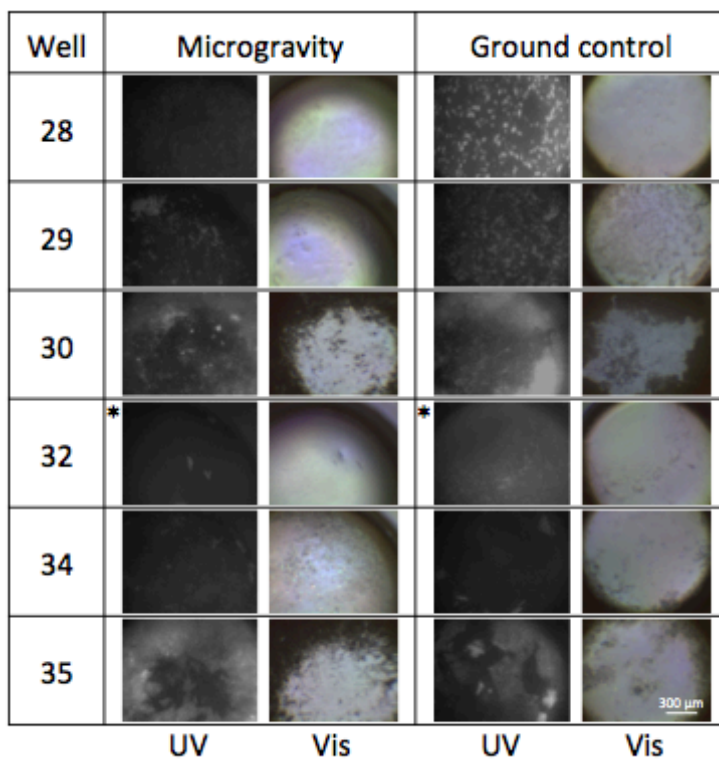
*Crystals formed post-flight

Fig S2. Representative crystal images taken using UV microscopy 10 months after completion of experiment.

7. Tables

Table 1.
Comparison of crystal number and size.

Environment of crystals	Number of wells with crystals	Average number of crystals per well* (range)	Mean crystal area** (μm^2)	Largest crystal (μm^2)
Microgravity	10	7 (5-13)	1,840	42,700
Ground control	9	14 (1-49)	1,500	27,200

*Average number of crystals > 400 μm^2 per well containing crystals
**Geometric mean area

Table 1. Comparison of crystal number and size.

Environment of Crystals	Well	Overall Resolution Limit (Å)		Average Mosaicity	Cell Dimensions: *a, b, c* (Å), α, β, γ (°)	Space group
		$CC_{1/2} > 0.3$	$<\bar{I}/\sigma(\bar{I})> > 1.50$			
Microgravity	29	2.47	2.68	0.08	42.10, 71.20, 207.72, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	34	2.55	2.67	0.13	42.35, 71.31, 207.31, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	34	2.57	2.71	0.07	42.36, 71.38, 207.31, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	34	2.58	2.68	0.13	42.59, 71.60, 208.37, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	29	2.63	2.83	0.07	41.98, 71.63, 207.64, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	34	2.63	2.81	0.18	42.60, 71.55, 208.62, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	29	2.67	2.68	0.11	42.05, 71.39, 207.72, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Microgravity	1	3.20	3.35	0.12	189.07, 189.07, 64.37, 90, 90, 120	$P\,6\,2\,2$
Ground*	2	1.59	1.71	0.06	42.28, 71.62, 89.19, 90, 91.51, 90	$P\,2_1$
Ground*	2	1.65	1.80	0.26	42.21, 72.19, 89.92, 90, 91.95, 90	$P\,2_1$
Ground*	2	1.72	1.80	0.19	42.19, 71.53, 89.08, 90, 90.96, 90	$P\,2_1$
Ground*	2	1.87	2.03	0.09	42.23, 71.69, 89.05, 90, 91.39, 90	$P\,2_1$
Ground*	2	2.19	2.32	0.06	42.19, 71.61, 88.88, 90, 91.34, 90	$P\,2_1$
Ground*	2	2.29	2.61	0.08	42.26, 71.70, 89.04, 90, 91.59, 90	$P\,2_1$
Ground*	17	3.00	3.00	0.15	42.48, 72.37, 89.78, 90, 91.43, 90	$P\,2_1$
Ground control	27	1.98	2.25	0.10	41.84, 70.28, 206.78, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	27	2.13	2.45	0.06	41.94, 70.43, 207.36, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	27	2.40	2.69	0.05	42.03, 70.50, 207.82, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	35	2.72	2.88	0.07	42.35, 71.30, 207.37, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	35	2.76	2.90	0.15	42.31, 71.20, 207.30, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	35	2.76	2.95	0.34	42.37, 71.10, 208.32, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	35	2.84	2.96	0.17	42.63, 71.78, 208.79, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	27	2.89	2.98	0.16	42.22, 71.01, 208.95, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	35	2.91	2.96	0.28	42.38, 71.29, 207.61, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	27	3.96	3.70	0.06	41.90, 70.43, 207.75, 90, 90, 90	$P\,2_1\,2_1\,2_1$
Ground control	2	4.05	4.54	0.19	190.64, 190.64, 64.91, 90, 90, 120	$P\,6_2\,2\,2$
Ground control	1	4.33	4.07	0.29	323.00, 63.74, 186.24, 90, 90, 90	$P\,2_1\,2_1\,2_1$

*Visible crystals grew in microgravity wells after return from ISS

Table 2. X-ray data processing statistics for MW1 Fab crystals < 5.0 Å. $<I/\sigma(I)>$ is the empirical signal to noise ratio [119].

Well	Protein Name(s)	Protein Conc (mg/mL)	Crystallization solution
1	MW1 Fab	7	0.1 M Sodium citrate tribasic dihydrate, pH 5.5, 16% w/v PEG 8,000
2	MW1 Fab	7	0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 18% w/v PEG 20,000
3	MW1 Fab	7	0.2 M Ammonium citrate tribasic, pH 7.0, 0.1 M Imidazole, pH 7.0, 20% w/v PEG MME 2,000
4	3B5H10 Fab	7	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 10% w/v PEG 20,000
5	3B5H10 Fab	7	1.8 M Ammonium sulfate, 0.1 M BIS-TRIS, pH 6.5, 2% v/v PEG MME 550
6	3B5H10 Fab + $K_2Q_{10}K_2$	7 + 7	1.8 M Ammonium sulfate, 0.1 M BIS-TRIS, pH 6.5, 2% v/v PEG MME 550
7	MW1 Fab + $K_2Q_{10}K_2$	7 + 7	0.1 M Imidazole, pH 7.0, 30% w/v PEG MME 550
8	MW1 Fab + $K_2Q_{10}K_2$	7 + 7	0.2 M Imidazole, pH 7.0, 28% w/v PEG MME 550
9	MW1 Fab + $K_2Q_{10}K_2$	7 + 7	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 10% w/v PEG 20,000
10	MW1 Fab + $K_2Q_{10}K_2$	7 + 7	0.1 M Imidazole, pH 7.0, 12% w/v PEG 20,000
11	HD-16Q	5	0.1 M Tris-HCl, pH 8.5, 8% w/v PEG 8,000
12	HD-16Q	5	0.1 M Tris-HCl, pH 8.5, 8% w/v PEG 8,000
13	3B5H10 Fab + HD-16Q	7 + 7	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 10% w/v PEG 20,000
14	3B5H10 Fab + HD-16Q	7 + 7	0.1 M Tris-HCl, pH 8.0, 30% v/v Jeffamine M-600, pH 7.0
15	3B5H10 Fab + HD-16Q	7 + 7	0.1 M Sodium acetate trihydrate, pH 4.0, 10% w/v PEG 4,000
16	3B5H10 Fab + HD-16Q	7 + 7	0.25 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 15% w/v PEG 20,000
17	MW1 Fab + HD-16Q	7 + 7	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 10% w/v PEG 20,000
18	MW1 Fab + HD-16Q	7 + 7	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 18% w/v PEG 20,000
19	MW1 Fab + HD-16Q	7 + 7	0.1 M Imidazole, pH 7.0, 25% w/v PEG MME 550
20	MW1 Fab + HD-16Q	7 + 7	0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 30% Jeffamine ED-2001 pH 7.0
21	MW1 Fab + HD-16Q	7 + 7	0.1 M BIS-TRIS, pH 6.5, 20% w/v PEG 1,500
22	MW1 Fab + HD-16Q	7 + 7	0.1 M Tris-HCl, pH 8.0, 28% w/v PEG 4,000
23	3B5H10 Fab + HD-39Q	7 + 2.3	0.1 M Sodium acetate trihydrate, pH 4.0, 15% w/v PEG 400
24	3B5H10 Fab + HD-39Q	7 + 2.3	0.1 M Sodium acetate trihydrate, pH 4.0, 15% w/v PEG 400
25	3B5H10 Fab + HD-39Q	7 + 2.3	0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 18% w/v PEG 20,000
26	3B5H10 Fab + HD-39Q	7 + 2.3	0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 18% w/v PEG 20,000
27	MW1 Fab + HD-39Q	7 + 2.3	0.1 M Sodium acetate trihydrate, pH 4.5, 30% w/v PEG 300
28	MW1 Fab + HD-39Q	7 + 2.3	0.1 M Sodium acetate trihydrate, pH 4.5, 30% w/v PEG 300
29	MW1 Fab + HD-39Q	7 + 2.3	1.8 M Ammonium sulfate, 0.1 M BIS-TRIS, pH 6.5, 2% v/v PEG MME 550
30	MW1 Fab + HD-39Q	7 + 2.3	0.2 M Magnesium formate dihydrate, 0.1 M Sodium acetate trihydrate, pH 4.0, 18% w/v PEG MME 5,000
31	MW1 Fab + HD-39Q	7 + 2.3	0.1 M Sodium citrate tribasic dihydrate, pH 5.5, 18% w/v PEG 3,350
32	MW1 Fab + HD-39Q	7 + 2.3	0.1 M Sodium citrate tribasic dihydrate, pH 5.5, 16% w/v PEG 8,000
33	MW1 Fab + HD-39Q	7 + 2.3	2% w/v 1,4-Dioxane, 0.1 M Tris-HCl, pH 8.0, 15% PEG 3,350
34	MW1 Fab + HD-39Q	7 + 2.3	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 14% w/v PEG 20,000
35	MW1 Fab + HD-39Q	7 + 2.3	0.2 M Magnesium chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.0, 18% w/v PEG 20,000
36	MW1 Fab + HD-39Q	7 + 2.3	0.4 M Sodium malonate, pH 6.0, 0.1 M MES monohydrate, pH 6.0, 0.5% w/v PEG 10,000
37	HD-16Q	10	0.2 M Ammonium acetate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.5, 30% w/v PEG 4,000
38	HD-16Q	10	0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate trihydrate, pH 6.5, 20% w/v PEG 8,000
39	HD-16Q	10	0.5 M Ammonium sulfate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.6, 1.0 M Lithium sulfate monohydrate
40	HD-16Q	10	0.2 M Magnesium chloride hexahydrate, 0.1 M Tris-HCl, pH 8.5, 25% w/v PEG 3,350
41	HD-25Q	10	0.01 M Nickel (II) chloride hexahydrate, 0.1 M Tris-HCl, pH 8.5, 20% w/v PEG MME 2,000
42	HD-25Q	10	0.01 M Nickel (II) chloride hexahydrate, 0.1 M Tris-HCl, pH 8.5, 1.0 M Lithium sulfate monohydrate
43	HD-25Q	10	1.0 M Imidazole, pH 7.0
44	HD-25Q	10	0.1 M BIS-TRIS, pH 6.5, 20% w/v PEG MME 5,000
45	HD-39Q	10	0.01 M Iron (III) chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate, pH 5.6, 10% Jeffamine M-600
46	HD-39Q	10	1.0 M Imidazole, pH 7.0
47	HD-46Q	10	0.2 M Calcium acetate hydrate, 0.1 M Sodium cacodylate trihydrate, pH 6.5, 18% w/v PEG 8,000
48	HD-46Q	10	0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate, 0.01 M Hexadecyltrimethylammonium bromide
49	HTT-GFP	20	0.2 M Zinc acetate dihydrate, 0.1 M Sodium cacodylate trihydrate, pH 6.5, 18% w/v PEG 8,000
50	HTT-GFP	20	0.2 M Zinc acetate dihydrate, 0.1 M Sodium cacodylate trihydrate, pH 6.5, 18% w/v PEG 8,000
51	HTT-GFP	20	0.2 M Zinc acetate dihydrate, 0.1 M Sodium cacodylate trihydrate, pH 6.5, 18% w/v PEG 8,000
52	HTT-GFP	20	0.2 M Lithium sulfate monohydrate, 0.1 M HEPES, pH 6.5, 25% w/v PEG 3,350
53	HTT-GFP	20	0.2 M Lithium sulfate monohydrate, 0.1 M HEPES, pH 6.5, 25% w/v PEG 3,350
54	HTT-GFP	20	0.2 M Lithium sulfate monohydrate, 0.1 M HEPES, pH 6.5, 25% w/v PEG 3,350
55	FL-HTT	0.4	0.1 M Sodium cacodylate trihydrate, pH 6.5, 25% w/v PEG 8000
56	FL-HTT	0.4	0.1 M Sodium cacodylate trihydrate, pH 6.5, 25% w/v PEG 8000
57	FL-HTT	0.4	2.4 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate, pH 4.5
58	FL-HTT	0.4	2.4 M Ammonium sulfate, 0.1 M Sodium acetate trihydrate, pH 4.5
59	FL-HTT	0.4	2.4 M Ammonium sulfate, 0.1 M Sodium cacodylate trihydrate, pH 6.5
60	FL-HTT	0.4	2.4 M Ammonium sulfate, 0.1 M Sodium cacodylate trihydrate, pH 6.5

Table S1. Initial conditions for all microgravity crystallization experiments.

Environment of crystals	Well	Morphology
Microgravity	1	3D
Microgravity	2	Irregular
Microgravity	8	Needle
Microgravity	10	Needle
Microgravity	17	Grew post-flight
Microgravity	27	3D
Microgravity	28	3D
Microgravity	29	3D
Microgravity	30	3D
Microgravity	32	Grew post-flight
Microgravity	34	3D
Microgravity	35	3D
Ground control	1	Irregular
Ground control	2	Needle
Ground control	8	Irregular
Ground control	10	Grew post-flight
Ground control	18	Grew post-flight
Ground control	27	3D
Ground control	28	3D
Ground control	29	3D
Ground control	30	3D microcrystals
Ground control	32	Grew post-flight
Ground control	34	3D
Ground control	35	3D

Table S2. Crystal morphology.

Chapter 5: Development of High-Avidity Anti-PolyQ Therapeutics

Chapter 5: Development of High-Avidity Anti-PolyQ Therapeutics

1. Fibronectin III (FN3) Domains

The soluble, unaggregated, unstructured polyQ stretch currently is a target for design of drugs to combat Huntington's disease [54]. Past drug design studies have focused on targeting the expanded polyQ repeat in huntingtin with either high-affinity small molecules or with antibodies. However, high-affinity polyQ binding proteins may disrupt other proteins that have polyQ repeats that are essential for normal cell function, such as transcription factors.

In order to discriminate between huntingtin with normal versus expanded polyQ repeats, we decided to create multivalent binding agents that exhibit high *avidity* rather than high *affinity* for polyQ stretches. Affinity describes the strength of a single 1:1 interaction between a protein and its binding partner, whereas avidity describes the combined, synergistic strength of multiple interactions. A well-known example of avidity occurs when antibodies bind to antigens that are tethered to a surface– because an antibody has two "arms," it can bind two antigens simultaneously. Thus when a binding interaction breaks, the other binding interaction can keep the antibody bound to the antigens on the surface. Here we constructed binding complexes with multiple polyQ binding sites that bind with high avidity to long polyQ stretches. Our high-avidity binding complex should bind only weakly to proteins with a short polyQ stretch, but the complex would bind much more strongly to proteins with longer stretches of polyQ (Figure 5-1). Through linking multiple individual binding compounds with low affinities, one can create new multicomponent compounds for which the binding avidity is very strong and these compounds are therefore specific for expanded polyQ repeats.

	Binding to <u>short</u> polyQ repeats	Binding to <u>long</u> polyQ repeats
High-Affinity Monovalent Protein	$K_D \sim uM$	$K_{D1} \sim uM$ $K_{D2} > K_{D1}$
High-Avidity Multivalent Protein	$K_D \sim uM$	$K_D < nM$

Figure 5-1. Use of high-avidity multivalent proteins allows for discrimination between short and expanded polyQ repeats. While high-affinity monovalent proteins bind strongly to short or to expanded polyQ repeats, high-avidity multivalent proteins bind more strongly and specifically to expanded polyQ repeats.

To test this linear lattice theory and derive potential novel therapeutic compounds for Huntington's disease by selecting proteins that bind weakly to a short stretch of polyQ, and then assembling together several of these protein building blocks to create a single complex containing two to five copies of the polyQ-binding domain. More than 30 different naturally existing proteins have been engineered for ligand binding, including α-helical, β-barrel, and β-sandwich proteins [142]. These types of protein domains are short peptides that have a simple, thermodynamically stable fold that is able to tolerate substitutions of amino acids within the binding surface region.

We chose to start by using a domain from fibronectin, a large protein that plays an essential role in the interaction between cells and the extracellular matrix. The fibronectin type III (FN3) domain is a small, independently folded domain that exists arranged in tandem repeats within fibronectin and other proteins [143]. The structure of

the FN3 domain resembles structures of binding domains of antibodies. **Since FN3 domains naturally exist in a series of repeats with up to 17 FN3 domains in a row, their structure is ideally suited to engineering complexes of individual domains linked together.** Individual FN3 domains are expressed well in bacterial cell culture, are very stable under a range of conditions, and have been used previously in selection strategies in which mutated FN3 domains that bind a desired protein are selected for from a large pool of mutated FN3 domains [144].
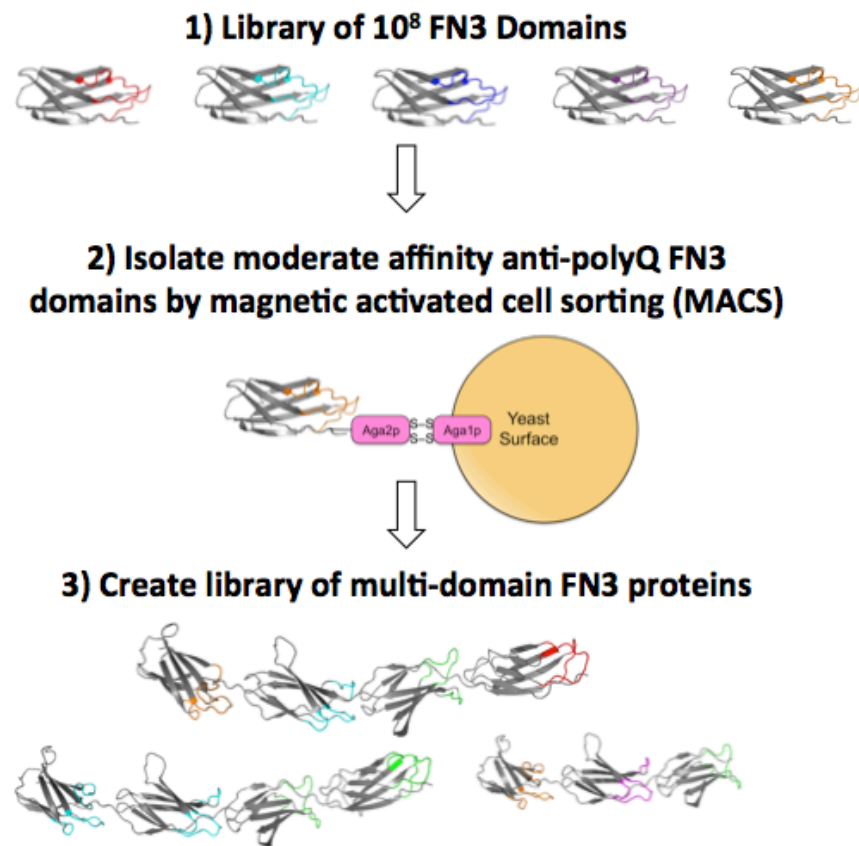

Figure 5-2. Overview of production of monovalent and multivalent FN3 proteins from polyQ-specific FN3 domains.

Using a pool of FN3 proteins containing a combinatorial library of 10^8 different FN3

domains in which multiple positions have been randomized, proteins were generated that bind weakly to soluble polyQ repeats within the first exon of huntingtin. Multiple independent binding proteins were linked together to create binding complexes with avidity and specificity for the pathogenic, expanded polyQ repeats.

2. Design of High Avidity FN3 Binding Proteins

2.1. Yeast Display

The FN3 domain is a protein domain with a beta sandwich structure that can be genetically engineered to bind to antigens. Using a library of 10^8 engineered monomeric FN3 domains, yeast display was used to select for FN3 domains with affinity to polyQ. The G4 FN3 library was selected as the starting material for our studies [144]. Moderate affinity FN3 domains were isolated by magnetic activated cell sorting (MACS) for affinity to a synthesized, biotinylated polyQ peptide, $K_2Q_{10}K_2$ (Genscript). Three sequential rounds of MACS were used to isolate these polyQ-binding domains [145]. FACS was subsequently used to select full-length FN3 domains using the C-terminal c-myc tag on each of these domains. A streptavidin antibody was used to test for binding of the selected yeast to the biotinylated polyQ peptide.

Plasmid DNA was isolated from the selected population of yeast using a Zymoprep yeast plasmid miniprep kit (Zymo Research), with addition of zymolyase to digest the yeast cell wall. *E. coli* (XLI-Blue or XL10-Gold) were transformed and incubated overnight. Plasmids DNA was prepared by miniprep and FN3 domains were sequenced using pCTCON2 sequencing primers. Sequences were aligned with MegAlign, and the

consensus sequences were displayed using WebLogo [146]. Sequencing of 20 clones showed an increase in aromatic amino acids in randomized loops above what would be expected based on the distribution of amino acids in the unsorted FN3 library.

Figure 5-3. Amino acid distribution in anti-polyQ FN3 (*left*). Structure of a single FN3 domain (1FNA) (*right*).

2.2. Cloning Strategy

Using the FN3 domains isolated, we produced monovalent and multivalent FN3 proteins from polyQ-specific FN3 domains. These FN3 domains were cloned in tandem repeats and selected for high avidity to expanded polyQ tracts to create multivalent binding agents. The FN3 genes from sorted G4 libraries were subcloned into the pET-24a(+) vector (Novagen; San Diego, California, United States) by Gibson assembly (New England Biolabs; Ipswich, Massachusetts, United States) (Figure 5-4) [147].

1. PCR FN3 domains & vector
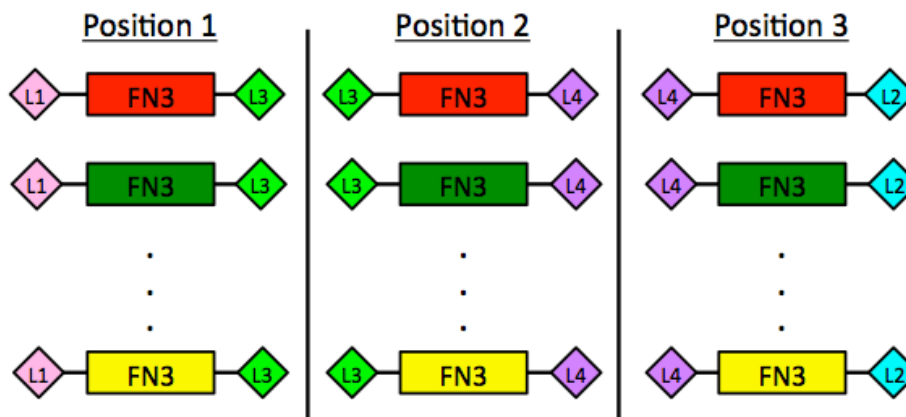
2. Gibson assembly

Figure 5-4. Multivalent FN3 construct construction. Each FN3 domain was flanked with a linker for assembly.

Four to six clones of each were sequenced. The pET-24a(+) vecor encodes a C-terminal His$_6$ tag. The FN3 proteins were expressed in *E. coli* BL21(DE3) cells (Novagen) by induction at OD$_{600}$=0.6-0.8 with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) for 3 to 4 hours at 37°C. Cells were collected by centrifugation at 6000 x *g* for 15 min and resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 10 mM imidazole, and 1 mM PMSF). These were sonicated on ice for 4 min. The cell lysate was centrifuged at 15,000 x *g* for 45 min to isolate the soluble fraction. The supernatant was incubated with Ni nitriloacetate beads, and the histidine-tagged FN3 domain was eluted with 500 mM imidazole. The eluate was concentrated using an Amicon Ultra centrifugal filter with a 5 kDa MWCO membrane. The protein was isolated via gel filtration on a Superdex 200 HR10/300 or 16/60 column, with yields of 0.5-100 mg/L of

bacterial culture.

While 1-domain FN3 constructs were easily cloned and expressed in bacterial culture, cloning of repeated FN3 domains proved to be difficult. Due to similarities between each domain, the reactions were easily mis-primed and truncated clones were made. After making the linker segments between each FN3 domain significantly longer, up to 5 tandem domains could be cloned using the method described above. A total of 24 different constructs were cloned, including thirteen 1-domain, three 2-domain, four 3-domain, and four 5-domain FN3 constructs. Out of these, six 1-domain, three 2-domain, four 3-domain, and two 5-domain FN3 constructs were expressed in cell culture and purified.

The presence of unpaired cysteines in the initial FN3 constructs led to concerns regarding oligomerization of FN3 domains during protein purification and analysis. Cysteine (TGT/TGC) and histidine (CAT/CAC) were included in the original FN3 library at a relatively high frequency due to codon similarity to tyrosine (TAT/TAC), such that only 19% of the naïve G4 library was cysteine-free. Several non-cysteine-containing FN3 constructs have been cloned, expressed, and purified. Additionally, because purification of multi-domain FN3 constructs has produced only small amounts of pure protein, purification techniques must be optimized to increase the production of multi-domain FN3 constructs and decrease the amount of non-specific binding of FN3 domains to other components in the cell lysate. Possible methods to optimize multi-domain FN3 constructs include alternate cloning methods and expression in yeast.

2.3. Surface Plasmon Resonance

Surface plasmon resonance (SPR) is a phenomenon that occurs when polarized light,

under conditions of total internal reflection, hits a layer of gold at the interface between a glass sensor surface, which has high refractive index, and a buffer, which has low refractive index. The change in refractive index in the buffer close to the gold layer can be detected, and thereby yield information about the interaction between proteins immobilized on the sensor surface and molecules in the buffer. SPR can be used to determine the strength of binding of an interaction, K_D. In our experiments, huntingtin with 16 glutamines (HD-16Q) was immobilized on Biacore chips using amino coupling, and FN3 constructs were injected through flow cells over the surface at successive dilutions to obtain the binding constants. A control antibody, MW1, which is known to bind to the polyQ repeat in huntingtin exon 1, was also used in this experiment.

HD-16Q + 1-Domain FN3
$K_D = 18\ \mu M$

HD-16Q + 2-Domain FN3
$K_D = 3.5\ \mu M$

HD-16Q + 3-Domain FN3
$K_D = 2.3\ \mu M$

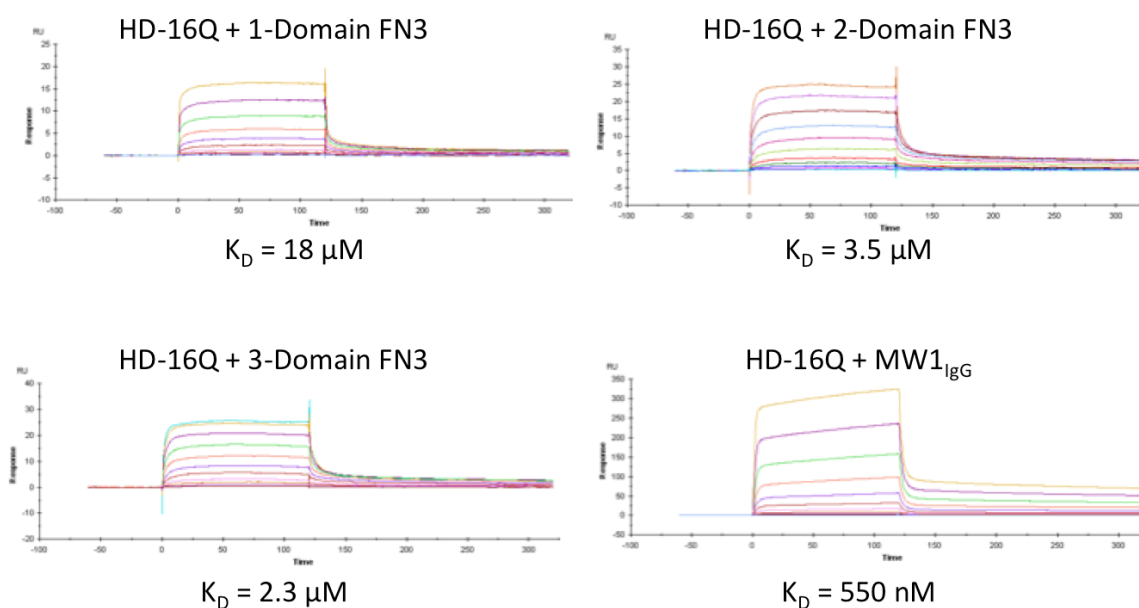HD-16Q + MW1$_{IgG}$
$K_D = 550\ nM$

Figure 5-5. SPR results demonstrate micromolar binding affinity of monomeric FN3 domain (FN3-1) to HD-16Q, but increasing affinity with two (2-Domain FN3) and three (3-Domain FN3) linked FN3 domains. MW1 is an anti-polyQ antibody that acted as a control for this experiment.

Results demonstrated that there was micromolar binding affinity to a 1-domain FN3 construct, as expected. Additionally, a 2-domain FN3 construct bound with a stronger affinity to HD-16Q. Future experiments will be aimed at careful observation of the effect of the newly purified non-cysteine-containing anti-polyQ FN3s on the biochemical features of huntingtin.

3. Future of Therapeutics for Huntington's Disease

Although 20 years have passed since the discovery of the Huntington's disease gene, Huntington's disease remains without a cure or even effective treatment. Current medications for treatment of HD include medications for movement disorders, such as antipsychotic drugs, and medications for psychiatric disorders, such as antidepressants and mood-stabilizing drugs. Three HD clinical trials are currently underway in the United States: Enroll-HD, a study to accelerate development of therapies for HD by compiling clinical data and biological samples during the "natural course" of the disease; 2CARE, a test of Coenzyme Q10; and CREST-E, a test of creatine. However, no treatment is available to slow or stop the progressive course of the disease. Creating multivalent binding proteins with high avidity and specificity for the expanded polyQ tracts that cause Huntington's disease may be a key to inhibition of aggregation of polyQ and have implications for future targeted treatments for Huntington's disease.

Bibliography

1.   Group, T.H.s.D.C.R., *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes.* Cell, 1993. **72**: p. 971-783.

2.   Myers, R.H., et al., *De novo expansion of a (CAG)n repeat in sporadic Huntington's disease.* Nat Genet, 1993. **5**(2): p. 168-73.

3.   Wheeler, V.C., et al., *Factors associated with HD CAG repeat instability in Huntington disease.* J Med Genet, 2007. **44**(11): p. 695-701.

4.   Wexler, N.S., et al., *Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset.* Proc Natl Acad Sci U S A, 2004. **101**(10): p. 3498-503.

5.   Genetic Modifiers of Huntington's Disease Consortium. Electronic address, g.h.m.h.e. and M.H.D.C. Genetic Modifiers of Huntington's Disease Ge, *Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease.* Cell, 2015. **162**(3): p. 516-26.

6.   Gusella, J.F., et al., *Molecular genetics of Huntington's disease.* Arch Neurol, 1993. **50**(11): p. 1157-63.

7.   Douglas, I., et al., *Juvenile Huntington's disease: a population-based study using the General Practice Research Database.* BMJ Open, 2013. **3**(4).

8.   Walker, F.O., *Huntington's disease.* The Lancet, 2007. **369**: p. 218-228.

9.   Pringsheim, T., K. Fiest, and N. Jette, *The international incidence and prevalence of neurologic conditions: how common are they?* Neurology, 2014. **83**(18): p. 1661-4.

10.   Warby, S.C., et al., *HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia.* Eur J Hum Genet, 2011. **19**(5): p. 561-6.

11.   Gusella, J.F., et al., *A polymorphic DNA marker genetically linked to Huntington's disease.* Nature, 1983. **306**(5940): p. 234-8.

12.   Wexler, A., *Chorea and Community in a Nineteenth-Century Town.* Bulletin of the History of Medicine, 2002. **76**(3): p. 495-527.

13.   Graveland, G.A., R.S. Williams, and M. DiFiglia, *Evidence for degenerative and regenerative changes in neostriatal spiny neurons in Huntington's disease.* Science, 1985. **227**(4688): p. 770-3.

14.   Youssov, K., et al., *The unified Huntington's Disease Rating Scale for advanced patients: validation and follow-up study.* Mov Disord, 2013. **28**(14): p. 1995-2001.

15.   Siesling, S., et al., *Unified Huntington's disease rating scale: a follow up.* Mov Disord, 1998. **13**(6): p. 915-9.

16.   Group, H.S., *Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group.* Mov Disord, 1996. **11**(2): p. 136-42.

17.   Cui, X., et al., *TR-FRET assays of Huntingtin protein fragments reveal temperature and PolyQ length-dependent conformational changes.* Sci Rep, 2014. **4**: p. 5601.

18.   Baldo, B., et al., *TR-FRET-based duplex immunoassay reveals an inverse correlation of soluble and aggregated mutant huntingtin in huntington's disease.* Chem Biol, 2012. **19**(2): p. 264-75.

19.   Southwell, A.L., et al., *Ultrasensitive measurement of huntingtin protein in cerebrospinal fluid demonstrates increase with Huntington disease stage and decrease following brain huntingtin suppression.* Sci Rep, 2015. **5**: p. 12166.

20. Nance, M., et al., *A Physician's Guide to the Management of Huntington's Disease*. 2011, Huntinton's Disease Society of America.

21. Shen, V., et al., *Safety and Efficacy of Tetrabenazine and Use of Concomitant Medications During Long-Term, Open-Label Treatment of Chorea Associated with Huntington's and Other Diseases.* Tremor and Other Hyperkinetic Movements, 2013. **3**: p. 1-13.

22. Messer, A. and S.N. Joshi, *Intrabodies as neuroprotective therapeutics.* Neurotherapeutics, 2013. **10**(3): p. 447-58.

23. Southwell, A.L., et al., *Intrabodies binding the proline-rich domains of mutant huntingtin increase its turnover and reduce neurotoxicity.* J Neurosci, 2008. **28**(36): p. 9013-20.

24. Colby, D.W., et al., *Potent inhibition of huntingtin aggregation and cytotoxicity by a disulfide bond-free single-domain intracellular antibody.* Proc Natl Acad Sci U S A, 2004. **101**(51): p. 17616-21.

25. Yu, D., et al., *Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression.* Cell, 2012. **150**(5): p. 895-908.

26. Skotte, N.H., et al., *Allele-specific suppression of mutant huntingtin using antisense oligonucleotides: providing a therapeutic option for all Huntington disease patients.* PLoS One, 2014. **9**(9): p. e107434.

27. Fiszer, A. and W.J. Krzyzosiak, *Oligonucleotide-based strategies to combat polyglutamine diseases.* Nucleic Acids Res, 2014. **42**(11): p. 6787-810.

28. Hoogeveen, A.T., et al., *Characterization and localization of the Huntington disease gene product.* Hum Mol Genet, 1993. **2**(12): p. 2069-73.

29. De Rooij, K.E., et al., *Subcellular localization of the Huntington's disease gene product in cell lines by immunofluorescence and biochemical subcellular fractionation.* Hum Mol Genet, 1996. **5**(8): p. 1093-9.

30. Barnes, G.T., et al., *Mouse Huntington's disease gene homolog (Hdh).* Somat Cell Mol Genet, 1994. **20**(2): p. 87-97.

31. Nasir, J., et al., *Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes.* Cell, 1995. **81**(5): p. 811-23.

32. Seong, I.S., et al., *HD CAG repeat implicates a dominant property of huntingtin in mitochondrial energy metabolism.* Hum Mol Genet, 2005. **14**(19): p. 2871-80.

33. Andre, W., et al., *Structure of inclusions of Huntington's disease brain revealed by synchrotron infrared microspectroscopy: polymorphism and relevance to cytotoxicity.* Anal Chem, 2013. **85**(7): p. 3765-73.

34. Wear, M.P., et al., *Proteins with Intrinsically Disordered Domains Are Preferentially Recruited to Polyglutamine Aggregates.* PLoS One, 2015. **10**(8): p. e0136362.

35. Tan, Z., et al., *Huntington's disease cerebrospinal fluid seeds aggregation of mutant huntingtin.* Mol Psychiatry, 2015.

36. Zuccato, C., M. Valenza, and E. Cattaneo, *Molecular mechanisms and potential therapeutical targets in Huntington's disease.* Physiol Rev, 2010. **90**(3): p. 905-81.

37. Steffan, J.S., et al., *SUMO modification of Huntingtin and Huntington's disease pathology.* Science, 2004. **304**(5667): p. 100-4.

38. Gu, X., et al., *N17 Modifies mutant Huntingtin nuclear pathogenesis and severity of disease in HD BAC transgenic mice.* Neuron, 2015. **85**(4): p. 726-41.

39. Michalek, M., E.S. Salnikov, and B. Bechinger, *Structure and topology of the huntingtin 1-17 membrane anchor by a combined solution and solid-state NMR approach.* Biophys J, 2013. **105**(3): p. 699-710.

40. Kim, M.W., et al., *Secondary structure of Huntingtin amino-terminal region.* Structure, 2009. **17**(9): p. 1205-12.

41. Kim, M., *Beta conformation of polyglutamine track revealed by a crystal structure of Huntingtin N-terminal region with insertion of three histidine residues.* Prion, 2013. **7**(3): p. 221-8.

42. Li, W., et al., *Expression and characterization of full-length human huntingtin, an elongated HEAT repeat protein.* J Biol Chem, 2006. **281**(23): p. 15916-22.

43. Seong, I.S., et al., *Huntingtin facilitates polycomb repressive complex 2.* Hum Mol Genet, 2010. **19**(4): p. 573-83.

44. Huang, B., et al., *Scalable production in human cells and biochemical characterization of full-length normal and mutant huntingtin.* PLoS One, 2015. **10**(3): p. e0121055.

45. Ko, J., S. Ou, and P.H. Patterson, *New anti-huntingtin monoclonal antibodies: Implications for huntingtin conformation and its binding proteins.* Brain Research Bulletin, 2001. **56**(3/4): p. 11.

46. Brooks, E., et al., *Using antibodies to analyze polyglutamine stretches.* Methods Mol Biol, 2004. **277**: p. 103-28.

47. Lescure, A., et al., *The N-terminal domain of the human TATA-binding protein plays a role in transcription from TATA-containing RNA polymerase II and III promoters.* EMBO J, 1994. **13**(5): p. 1166-75.

48. Trottier, Y., et al., *Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias.* Nature, 1995. **378**(6555): p. 4.

49. White, J.K., et al., *Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion.* Nat Genet, 1997. **17**(4): p. 404-10.

50. Persichetti, F., et al., *Mutant huntingtin forms in vivo complexes with distinct context-dependent conformations of the polyglutamine segment.* Neurobiol Dis, 1999. **6**(5): p. 364-75.

51. Peters-Libeu, C., et al., *Crystallization and diffraction properties of the Fab fragment of 3B5H10, an antibody specific for disease-causing polyglutamine stretches.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2005. **61**(Pt 12): p. 1065-8.

52. Peters-Libeu, C., et al., *Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure.* J Mol Biol, 2012. **421**(4-5): p. 587-600.

53. Klein, F.A., et al., *Linear and extended: a common polyglutamine conformation recognized by the three antibodies MW1, 1C2 and 3B5H10.* Hum Mol Genet, 2013. **22**(20): p. 4215-23.

54. Li, P., et al., *The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model.* Nat Struct Mol Biol, 2007. **14**(5): p. 381-7.

55. Klein, F.A., et al., *Pathogenic and non-pathogenic polyglutamine tracts have similar structural properties: towards a length-dependent toxicity gradient.* J Mol Biol, 2007. **371**(1): p. 235-44.

56. Khoshnan, A., et al., *Antibodies Against Huntingtin: Production and Screening of Monoclonals and Single-Chain Recombinant Forms*, in *Methods in Molecular Biology*, Y. Kohwi, Editor. 2004. p. 87-102.

57. Ko, J., S. Ou, and P.H. Patterson, *New anti-huntingtin monoclonal antibodies: implications for huntingtin conformation and its binding proteins.* Brain Res Bull, 2001. **56**(3-4): p. 319-29.

58.   Bennett, M.J., et al., *A linear lattice model for polyglutamine in CAG-expansion diseases.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11634-9.

59.   McGhee, J.D. and P.H. von Hippel, *Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice.* J Mol Biol, 1974. **86**(2): p. 469-89.

60.   Hendricks, A.E., et al., *Estimating the probability of de novo HD cases from transmissions of expanded penetrant CAG alleles in the Huntington disease gene from male carriers of high normal alleles (27-35 CAG).* Am J Med Genet A, 2009. **149A**(7): p. 1375-81.

61.   Nagai, Y., et al., *A toxic monomeric conformer of the polyglutamine protein.* Nat Struct Mol Biol, 2007. **14**(4): p. 332-40.

62.   Miller, J., et al., *Identifying polyglutamine protein species in situ that best predict neurodegeneration.* Nat Chem Biol, 2011. **7**(12): p. 925-34.

63.   Fodale, V., et al., *Polyglutamine- and temperature-dependent conformational rigidity in mutant huntingtin revealed by immunoassays and circular dichroism spectroscopy.* PLoS One, 2014. **9**(12): p. e112262.

64.   Wyatt, P.J., *Light scattering and the absolute characterization of macromolecules.* Analytica Chimica Acta, 1993. **272**: p. 40.

65.   Sanchez, L.M., D.M. Penny, and P.J. Bjorkman, *Stoichiometry of the interaction between the major histocompatibility complex-related Fc receptor and its Fc ligand.* Biochemistry, 1999. **38**(29): p. 9471-9476.

66.   West, A.P., Jr., et al., *Mutational analysis of the transferrin receptor reveals overlapping HFE and transferrin binding sites.* J Mol Biol, 2001. **313**(2): p. 385-97.

67.   Hummel, J.P. and W.J. Dreyer, *Measurement of protein-binding phenomena by gel filtration.* Biochim Biophys Acta, 1962. **63**: p. 530-2.

68.   Svergun, D.I., *Determination of the regularization parameter in indirect-transform methods using perceptual critera.* J Appl Crystallogr, 1992. **25**: p. 495-503.

69.   Vitalis, A., X. Wang, and R.V. Pappu, *Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization.* J Mol Biol, 2008. **384**(1): p. 279-97.

70.   Mylonas, E. and D.I. Svergun, *Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering.* Journal of Applied Crystallography, 2007. **40**: p. S245-S249.

71.   Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering.* J Struct Biol, 2010. **172**(1): p. 128-41.

72.   Rambo, R.P. and J.A. Tainer, *Accurate assessment of mass, models and resolution by small-angle scattering.* Nature, 2013. **496**(7446): p. 477-81.

73.   Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.

74.   Trewhella, J., et al., *Report of the wwPDB Small-Angle Scattering Task Force: data requirements for biomolecular modeling and the PDB.* Structure, 2013. **21**(6): p. 875-81.

75.   Caron, N.S., et al., *Polyglutamine domain flexibility mediates the proximity between flanking sequences in huntingtin.* Proc Natl Acad Sci U S A, 2013. **110**(36): p. 14610-5.

76.   Butland, S.L., et al., *CAG-encoded polyglutamine length polymorphism in the human genome.* BMC Genomics, 2007. **8**: p. 126.

77.   Schaefer, M.H., E. Wanker, and M.A. Andrade-Navarro, *Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks.* Nucleic Acids Res, 2012. **40**(10): p. 4273-4287.

78.   Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.

79.   Diskin, R., et al., *Increasing the potency and breadth of an HIV antibody by using structure-based rational design.* Science, 2011. **334**(6060): p. 1289-93.

80.   Atwal, R.S., et al., *Kinase inhibitors modulate huntingtin cell localization and toxicity.* Nat Chem Biol, 2011. **7**(7): p. 453-60.

81.   Smolsky, I.L., et al., *Biological small-angle x-ray scattering facility at the Stanford synchrotron radiation laboratory.* Journal of Applied Crystallography, 2007. **40**: p. S453-S458.

82.   Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**: p. 1277-1282.

83.   Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**: p. 342-350.

84.   Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.* Journal of Applied Crystallography, 2009. **42**: p. 342-346.

85.   Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophysical Journal, 1999: p. 2897-2886.

86.   Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering.* Journal of Applied Crystallography, 2003. **36**: p. 860-864.

87.   Petoukhov, M.V., et al., *ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis.* Journal of Applied Crystallography, 2007. **40**: p. 6.

88.   LaVallie, E.R., et al., *A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm.* Biotechnology, 1996. **11**: p. 187-193.

89.   Naiki, H., et al., *Fluorescent determination of amyloid fibrils in vitro using the fluorescent dye, Thioflavine T.* Anal Biochem, 1989. **177**: p. 244-249.

90.   Wolfe, L.S., et al., *Protein-induced photophysical changes to the amyloid indicator dye thioflavin T.* Proc Natl Acad Sci U S A, 2010. **107**(39): p. 16863-8.

91.   Streets, A.M., et al., *Simultaneous measurement of amyloid fibril formation by dynamic light scattering and fluorescence reveals complex aggregation kinetics.* PLoS One, 2013. **8**(1): p. e54541.

92.   Khurana, R., et al., *Is Congo red an amyloid-specific dye?* J Biol Chem, 2001. **276**(25): p. 22715-21.

93.   Howie, A.J., et al., *Physical basis of colors seen in Congo red-stained amyloid in polarized light.* Lab Invest, 2008. **88**(3): p. 232-42.

94.   NASA, *Educational brief: a NASA recipe for protein crystallography.* 2000. p. 1-18.

95.   Judge, R.A., E.H. Snell, and M.J. van der Woerd, *Extracting trends from two decades of microgravity macromolecular crystallization history.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 6): p. 763-71.

96. Vergara, A., et al., *Crystallization of the collagen-like polypeptide (PPG)10 aboard the International Space Station. 1. Video observation.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 10 Pt 1): p. 1690-4.

97. Vergara, A., et al., *Physical aspects of protein crystal growth investigated with the Advanced Protein Crystallization Facility in reduced-gravity environments.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 1): p. 2-15.

98. Berisio, R., et al., *Crystallization of the collagen-like polypeptide (PPG)10 aboard the International Space Station. 2. Comparison of crystal quality by X-ray diffraction.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 10 Pt 1): p. 1695-9.

99. Yoshizaki, I., et al., *Scientific approach to the optimization of protein crystallization conditions for microgravity experiments.* Ann N Y Acad Sci, 2004. **1027**: p. 28-47.

100. Ross, C.A., et al., *Huntington disease: natural history, biomarkers and prospects for therapeutics.* Nat Rev Neurol, 2014. **10**(4): p. 204-16.

101. Thakur, A.K., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat Struct Mol Biol, 2009. **16**(4): p. 380-9.

102. Owens, G.E., et al., *Anti-PolyQ Antibodies Recognize a Short PolyQ Stretch in Both Normal and Mutant Huntingtin Exon 1.* J Mol Biol, 2015. **427**(15): p. 2507-19.

103. Temussi, P.A., L. Masino, and A. Pastore, *From Alzheimer to Huntington: why is a structural understanding so difficult?* EMBO J, 2003. **22**(3): p. 355-61.

104. McPherson, A. and L.J. DeLucas, *Microgravity protein crystallization.* npj Microgravity, 2015. **1**: p. 15010.

105. Borgstahl, G.E., et al., *A test of macromolecular crystallization in microgravity: large well ordered insulin crystals.* Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 8): p. 1204-7.

106. Dong, J., et al., *Bound-solvent structures for microgravity-, ground control-, gel- and microbatch-grown hen egg-white lysozyme crystals at 1.8 Å resolution.* Acta Crystallogr D Biol Crystallogr, 1999. **55**: p. 745-752.

107. Ivanova, M.I., et al., *Molecular basis for insulin fibril assembly.* Proc Natl Acad Sci U S A, 2009. **106**(45): p. 18990-5.

108. Sabogal, A. and D.C. Rio, *A green fluorescent protein solubility screen in E. coli reveals domain boundaries of the GTP-binding domain in the P element transposase.* Protein Sci, 2010. **19**(11): p. 2210-8.

109. Gasteiger, E., et al., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press. p. 571-607.

110. DeLucas, L., R. Rouleau, and K. Banaslewics, *High Density Protein Crystal Growth*, in *Patent No.: US 6,592,824 B2*, U.S.P. Office, Editor. 2003, UAB Research Foundation, Birmingham, AL: United States.

111. Abd Rahman, R.N.Z.R., et al., *A comparative analysis of microgravity and Earth grown thermostable T1 lipase crystals using HDPCG apparatus.* Protein & Peptide Letters, 2015. **22**: p. 173-179.

112. Krauspenhaar, R., et al., *Crystallization under microgravity of mistletoe lectin I from Viscum album with adenine monophosphate and the crystal structure at 1.9 Å resolution.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(1704-1707): p. 1704.

113. Ponassi, M., et al., *Crystals of the hydrogenase maturation factor HypF N-terminal domain grown in microgravity, display improved internal order.* Journal of Crystal Growth, 2011. **314**(1): p. 246-251.

114. Campana, S.E. and D.T. Melendez, *Cold Stowage Flight Systems*, in *41st International Conference on Environemental Systems*. 2011, American Inst. of Aeronautics and Astronautics: Portland, Oregon. p. 1-5.

115. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.

116. Kabsch, W., *Xds.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.

117. Kabsch, W., *Integration, scaling, space-group assignment and post-refinement.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 133-44.

118. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

119. Karplus, P.A. and K. Diederichs, *Linking crystallographic model and data quality.* Science, 2012. **336**(6084): p. 1030-3.

120. Evans, P.R. and G.N. Murshudov, *How good are my data and what is the resolution?* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 7): p. 1204-14.

121. Takahashi, S., et al., *JAXA protein crystallization in space: ongoing improvements for growing high-quality crystals.* J Synchrotron Radiat, 2013. **20**(Pt 6): p. 968-73.

122. Zorb, C., et al., *Bacteriorhodopsin crystal growth in reduced gravity--results under the conditions, given in CPCF on board of a Space Shuttle, versus the conditions, given in DCAM on board of the Space Station Mir.* Microgravity Sci Technol, 2002. **13**(3): p. 22-9.

123. Snell, E.H., et al., *CCD video observation of microgravity crystallization of lysozyme and correlation with accelerometer data.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 6): p. 747-55.

124. Savino, R. and R. Monti, *Buoyancy and surface-tension-driven convection in hanging-drop protein crystallizer.* Journal of Crystal Growth, 1996. **165**(3): p. 308-318.

125. Snell, E.H. and J.R. Helliwell, *Macromolecular crystallization in microgravity.* Reports on Progress in Physics, 2005. **68**(4): p. 799-853.

126. Strong, R.K., et al., *Long Duration Growth of Protein Crystals in Microgravity Aboard the Mir Space Station.* Journal of Crystal Growth, 1992. **119**(3-4): p. 200-214.

127. Snell, E.H., et al., *Investigating the effect of impurities on macromolecule crystal growth in microgravity.* Crystal Growth & Design, 2001. **1**(2): p. 151-158.

128. Ng, J.D., et al., *Comparative analysis of thaumatin crystals grown on earth and in microgravity.* Acta Crystallographica Section D-Biological Crystallography, 1997. **53**: p. 724-733.

129. Barnes, C.L., E.H. Snell, and C.E. Kundrot, *Thaumatin crystallization aboard the International Space Station using liquid-liquid diffusion in the Enhanced Gaseous Nitrogen Dewar (EGN).* Acta Crystallographica Section D-Biological Crystallography, 2002. **58**: p. 751-760.

130. Stoddard, B.L., et al., *Mir for the crystallographers' money.* Nature, 1992. **360**(6402): p. 293-4.

131. Littke, W. and C. John, *Materials: protein single crystal growth under microgravity.* Science, 1984. **225**(4658): p. 203-4.

132. DeLucas, L. and A. McPherson, *Comprehensive Evaluation of Microgravity Protein Crystallization*. 2012: ISS Utilization Conference.

133. Delucas, L.J., et al., *Preliminary Investigations of Protein Crystal-Growth Using the Space-Shuttle.* Journal of Crystal Growth, 1986. **76**(3): p. 681-693.

134. Kawaji, M., et al., *Investigation of Marangoni and natural convection during protein crystal growth.* Journal of Crystal Growth, 2003. **258**(3-4): p. 420-430.

135. Boggon, T.J., et al., *Protein crystal movements and fluid flows during microgravity growth.* Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 1998. **356**(1739): p. 1045-1061.

136. Squires, T.M. and S.R. Quake, *Microfluidics: Fluid physics at the nanoliter scale.* Reviews of Modern Physics, 2005. **77**(3): p. 977-1026.

137. Sauter, C., B. Lorber, and R. Giege, *Towards atomic resolution with crystals grown in gel: the case of thaumatin seen at room temperature.* Proteins, 2002. **48**(2): p. 146-50.

138. Tanaka, H., et al., *Numerical model of protein crystal growth in a diffusive field such as the microgravity environment.* Journal of Synchrotron Radiation, 2013. **20**: p. 1003-1009.

139. Qi, J.W., N.I. Wakayama, and M. Ataka, *Magnetic suppression of convection in protein crystal growth processes.* Journal of Crystal Growth, 2001. **232**(1-4): p. 132-137.

140. Lorber, B., *Virus and protein crystallization under hypergravity.* Crystal Growth & Design, 2008. **8**(8): p. 2964-2969.

141. West, A.P., Jr., et al., *Evaluation of CD4-CD4i antibody architectures yields potent, broadly cross-reactive anti-human immunodeficiency virus reagents.* J Virol, 2010. **84**(1): p. 261-9.

142. Jeong, K.J., R. Mabry, and G. Georgiou, *Avimers hold their own.* Nat Biotechnol, 2005. **23**(12): p. 1493-1494.

143. Potts, J.R. and I.D. Campbell, *Structure and function of fibronectin modules.* Matrix Biol, 1996. **15**(5): p. 313-20; discussion 321.

144. Hackel, B.J., et al., *Stability and CDR composition biases enrich binder functionality landscapes.* J Mol Biol, 2010. **401**(1): p. 84-96.

145. Chao, G., et al., *Isolating and engineering human antibodies using yeast surface display.* Nat Protoc, 2006. **1**(2): p. 755-68.

146. Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.

147. Gibson, D.G., et al., *Enzymatic assembly of DNA molecules up to several hundred kilobases.* Nat Methods, 2009. **6**(5): p. 343-5.