# Fibril branching dominates self-assembly of mutant huntingtin exon-1 aggregates *in vitro*

Anne S. Wagner[1,§], Antonio Z. Politi[2,3,§], Anne Steinhof[1], Kenny Bravo-Rodriguez[4,5], Alexander Buntru[1], Nadine U. Strempel[1], Lydia Brusendorf[1], Christian Hänig[1], Annett Boeddrich[1], Stephanie Plassmann[1], Juan M. Ramirez-Anguita[4], Katharina Baum[2], Elsa Sanchez-Garcia[4,5], Jana Wolf[2*] and Erich E. Wanker[1*]

1 Neuroproteomics, Max Delbrueck Center for Molecular Medicine, Robert-Rössle-Straße 10, 13125 Berlin, Germany

2 Mathematical Modelling of Cellular Processes, Max Delbrueck Center for Molecular Medicine, Robert-Rössle-Straße 10, 13125 Berlin, Germany

3 Present address: European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany

4 Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 2, 45470 Mülheim an der Ruhr, Germany

5 Computational Chemistry and Biology, University Duisburg-Essen, Universitätsstr. 5, 45141 Essen, Germany

§ contributed equally

***Corresponding authors:**

*ewanker@mdc-berlin.de*, tel. +49 30 9406 2157, fax +49 30 9406 2552 and

*jana.wolf@mdc-berlin.de*, tel. +49 30 9406 2641, fax +49 30 9406 2394

## Abstract

Huntingtin (HTT) fragments with extended polyglutamine (polyQ) tracts self-assemble into amyloid-like fibrillar aggregates. Elucidating the fibril formation mechanism is critical for understanding Huntington's disease pathology and for developing novel therapeutic strategies. Here, we performed systematic experimental and theoretical studies to examine the self-assembly of an aggregation-prone N-terminal HTT exon-1 fragment with 49 glutamines (Ex1Q49). We demonstrate that two nucleation mechanisms control spontaneous Ex1Q49 fibrillogenesis: (1) a relatively slow primary fibril-independent nucleation process, which involves the spontaneous formation of aggregation-competent monomers, and (2) a fast secondary fibril-dependent nucleation process, which involves branching and promotes the rapid assembly of highly complex fibril bundles with multiple ends. The proposed aggregation mechanism is supported by studies with the small molecule O4, which perturbs primary nucleation and delays Ex1Q49 fibril assembly, comprehensive mathematical and computational modelling studies, and seeding experiments with small, preformed fibrillar Ex1Q49 aggregates. All results indicate that *in vitro*, HTT exon-1 fibrillar aggregates are formed by a branching mechanism.

## Keywords

Huntingtin fibrillogenesis, aggregation mechanism, primary and secondary nucleation, amyloidogenesis, fibril branching

## Introduction

Proteins with diverse sequences, structures and functions can spontaneously self-assemble into morphologically similar insoluble protein aggregates, often referred to as amyloid [1-3]. Amyloid aggregates are highly stable fibrillar structures with a characteristic cross-β-sheet conformation [4]. They are associated with a wide range of neurodegenerative and localized non-neuropathic diseases [5-7]. However, they also play important roles in normal cellular processes like the biogenesis of melanosomes or the regulation of translation in yeast [8, 9].

The process of fibril self-assembly can be divided into two phases: (1) a lag phase where little or no fibrils are formed and (2) a fibril growth phase where an increase of fibril mass is observed [10]. Theoretical models describing these two phases of fibrillogenesis typically suggest a polymerization process that depends on a primary, possibly also a secondary, nucleation mechanism [10-14]. The formation of an initial “nucleus” is regarded as the rate-limiting step in amyloid formation. The nucleus is currently thought to be a kinetically unstable oligomer of low abundance spontaneously forming from interacting monomers during the lag phase [12]. Once assembled, it is assumed to grow into larger β-sheet-rich fibrils by monomer addition. The characteristic fibril growth profile observed for many amyloidogenic polypetides reflects the greater ease of addition of monomers onto existing aggregates compared to the *de novo* formation of new amyloidogenic oligomers (nuclei) from monomers through primary nucleation [14]. Several recent studies have demonstrated the involvement of secondary nucleation in the aggregation process of polypetides [15-18] implying that theoretical models with a single nucleation mechanism cannot fully explain aggregation, as they do not account for the rapid increase in fibril growth [19]. This suggests that kinetic models considering both primary and secondary nucleation events might better describe the complex polymerization processes that lead to the rapid formation of mature amyloid fibrils [10].

Evidence has been provided that N-terminal huntingtin exon-1 (HTTex1) fragments with pathogenic polyglutamine (polyQ) tracts spontaneously self-assemble into insoluble aggregates with fibrillar morphology [20-26]. This indicates that polymerization of amyloid-like HTTex1 aggregates is mechanistically similar to α-synuclein or amyloid-β aggregation observed in various model systems [22]. Spontaneous HTTex1 fibrillogenesis critically depends on the length of the polyQ tract as well as on protein concentration and time. It can be stimulated by preformed fibrils, hinting that it is dominated by a nucleation-dependent aggregation mechanism [21]. Based on studies with simple polyQ peptides [27] it is currently assumed that a single, fibril-independent nucleation step is critical for HTTex1 polymerization *in vitro* and *in vivo*. However, more recent studies indicate that short HTT fragments consisting of a flanking sequence of 17 N-terminal amino acids and a polyQ tract ($HTT^{NT}$) show a more complex aggregation mechanism [28], which includes the formation of transient oligomers that over time convert into larger fibrillar aggregates. Interestingly, it was observed that the first 17 amino acids stimulate the initial oligomerization of $HTT^{NT}$ molecules, while the polyQ sequence promotes a later step in the aggregation cascade, namely the self-assembly of ordered fibrils [28]. Thus, published studies suggest that β-sheet-rich $HTT^{NT}$ fibrils are spontaneously formed by a complex multistep process.

However, the molecular events that control the efficient conversion of polyQ-containing HTTex1 fragments from a soluble into an aggregated state are still poorly understood. It is important to note that HTTex1 fragments, which are produced by abnormal splicing in HD patients and transgenic animals [29], are longer and more complex than the previously studied simple polyQ and $HTT^{NT}$ peptides [28, 30, 31]. In addition to the N-terminal 17 amino acids and the polyQ tract, they contain a C-terminal proline-rich domain of ~50 amino acids, which was shown to influence the kinetics of spontaneous amyloid polymerization *in vitro* [32].

In the present study, we have systematically analysed the aggregation mechanism of a disease-relevant HTTex1 fragment with 49 glutamines (Ex1Q49), using biochemical, biophysical and computational methods. We show that self-assembly of Ex1Q49 fragments released by proteolytic cleavage from soluble GST fusions is controlled by a double nucleation mechanism. Initially, a primary fibril-independent nucleation process leading to the formation of aggregation-competent monomers dominates Ex1Q49 aggregation. It is followed by the rapid formation of fibrillar structures involving secondary nucleation, in particular branching, resulting in the formation of large, highly complex bundles of fibrils. Perturbation studies with the small molecule O4 that directly targets Ex1Q49 polypeptides and mathematical modelling of time-resolved Ex1Q49 aggregation data support our hypothesis that complex Ex1Q49 fibrils grow by a branching mechanism. Furthermore, seeding experiments with small preformed Ex1Q49 seeds substantiate the view that branching dominates fibril growth in cell-free assays. The implications of these mechanistic investigations for the pathogenesis of HD are discussed.

## Results

### Ex1Q49 spontaneously self-assembles into large, SDS-stable protein aggregates in cell-free assays

To study the kinetics of polyQ-mediated HTT exon-1 (HTTex1) polymerization and the nature of the nucleating aggregate species, a cell-free aggregation assay was established which allows the systematic investigation of aggregation reactions in a time- and concentration-dependent manner under non-denaturing conditions [33]. A purified, soluble glutathione S-transferase (GST) HTT exon-1 fusion protein with 49 glutamines (GST-Ex1Q49) was produced and purified from *E. coli* protein extracts as the starting material for the investigation of spontaneous Ex1Q49 polymerization *in vitro* (**Fig. 1a**). It migrates at ~60 kDa in denaturating SDS gels (**Fig. S1a**). Analysis by gel filtration under non-denaturating conditions, however, revealed that the fusion protein is a stable oligomer with a size of ~500 kDa (**Fig. S1b**).

Next, the purified GST-Ex1Q49 fusion protein (2 μM) was treated with the sequence-specific PreScission protease (PP) in order to release aggregation-prone Ex1Q49 fragments (**Fig. 1a and S1c**). Samples were incubated at 20°C in microtiter plates and the formation of stable Ex1Q49 protein aggregates was monitored over time using a denaturating membrane filter retardation assay (FRA). With this method the abundance of stable fibrillar HTTex1 aggregates with lengths >0.2 μm can be specifically quantified in complex protein solutions [21]. We found that PP treatment disintegrates soluble GST-Ex1Q49 oligomers and results in the rapid formation of SDS- and heat-stable Ex1Q49 protein aggregates, which are detectable after a short lag phase of ~2 h (**Fig. 1b and c**). This confirms previous observations that HTTex1 fragments with pathogenic polyQ tracts rapidly self-assemble into very stable protein aggregates, once they are released from GST fusion proteins by proteolytic cleavage [21, 34]. As expected, no SDS-stable Ex1Q49 aggregates were detected by FRAs in the absence of PP

(**Fig. 1b**). The cleavage of the GST-Ex1Q49 fusion protein with PP was confirmed by SDS-PAGE and immunoblotting (**Fig. S1d**).

We also addressed the question of whether the kinetics of spontaneous Ex1Q49 polymerization in cell-free assays is influenced by agitation. Previous investigations have demonstrated that shaking of samples with amyloidogenic polypeptides stimulates amyloid growth due to fibril breakage and polymerization at the ends of newly formed fibrils [16, 35], suggesting that also the aggregation of Ex1Q49 might be influenced by mechanical forces. Aggregation reactions with GST-Ex1Q49 fusion protein and PP were systematically analysed with (300 and 600 rpm) and without agitation using FRAs. We found that shaking of samples does not promote spontaneous Ex1Q49 polymerization *in vitro* (**Fig. S1e**). In particular, we did not observe an impairment of Ex1Q49 polymerization without agitation. The kinetics of spontaneous Ex1Q49 aggregation remained unaltered when increasing concentrations of additional proteins such as bovine serum albumin (BSA) were added to the reactions (**Fig. S1f),** indicating that this process is very robust and cannot be easily influenced by shaking or the addition of other proteins.

Finally, we examined the kinetics of spontaneous Ex1Q49 aggregation by dynamic light scattering (DLS), which provides size distribution profiles of aggregating protein molecules [36, 37]. For DLS analysis, a higher concentration of GST-Ex1Q49 fusion protein (20.9 μM) was needed than for FRA analysis (2 μM, **Fig. 1b and c**). From the autocorrelation curves, the size distribution was computed (**Fig. 1d**). Proteolytic cleavage of GST-Ex1Q49 oligomers with a diameter of 19.27±1.28 nm ($n$ = 4) resulted in the release of smaller particles (GST and Ex1Q49 molecules) with a size distribution that peaked at 10.89±0.92 nm ($n$ = 4). This is slightly larger than the expected hydrodynamic diameter for monomeric GST (7.4 nm, for a globular protein) or Ex1Q49 (>6.4 nm), estimated for an exon-1-Q37 peptide [28]. GST has been shown to dimerize [38], the observed diameter could therefore be explained from a mixed population of GST-

dimers and Ex1Q49 monomers. After a lag phase of ~1 h, large particles (diameter > 700 nm) rapidly started to appear, indicating the spontaneous formation of Ex1Q49 aggregates. After 2 h the size distribution reached a steady state and peaked at 1,204±93 nm ($n$ = 4). Thus, our DLS studies confirm the results obtained by FRAs (**Fig. 1c**) and demonstrate that large Ex1Q49 aggregates are rapidly formed from monomers or very small oligomers that are released from larger GST-Ex1Q49 oligomers by proteolytic cleavage (**Fig. 1d**).

**Ex1Q49 spontaneously forms bundles of highly connected amyloid fibrils**

Our studies indicate that Ex1Q49 fragments rapidly self-assemble into large, highly stable protein aggregates after a short lag phase (**Fig. 1b-d**). To investigate the morphology of these structures, we analysed the time-dependent formation of Ex1Q49 aggregates by atomic force microscopy (AFM). We observed that soluble, uncleaved GST-Ex1Q49 oligomers are very small spherical particles (**Fig. 1e**), confirming previous reports [34, 39]. However, large bundles of Ex1Q49 fibrils (>1 µm length) were detectable by AFM when soluble GST-Ex1Q49 oligomers were incubated over time with PP **(Fig. 1e, S2a and b)**. This suggests that the Ex1Q49 molecules released from the GST fusion proteins very rapidly self-assemble into complex fibrillar structures with multiple ends.

To substantiate these observations, aggregation reactions with GST-Ex1Q49 and PP were systematically analysed by transmission electron microscopy (TEM), confirming the formation of complex bundles of Ex1Q49 fibrils (**Fig. 1f**). We observed typical branching points, which are potential sites for lateral nucleation from the surface of already existing fibrils. This suggests that fibrillar Ex1Q49 aggregates *in vitro* are rapidly formed by branching and that both primary and secondary nucleation events are involved in the aggregation process.

**Epitope-specific antibodies reveal conformational changes in soluble aggregation-prone Ex1Q49 molecules**

To study the polymerization of Ex1Q49 molecules over time under non-denaturating conditions, dot blot assays (DBAs) with epitope-specific anti-HTT antibodies were performed [40]. We initially examined the time-dependent aggregation of Ex1Q49 using the anti-polyQ antibodies MW1 and 3B5H10 (**Fig. 1a**), both of which recognize short polyQ epitopes in soluble HTT proteins [41-43]. We found that both antibodies readily recognize uncleaved GST-Ex1Q49 fusion protein (**Fig. 2a and b**), indicating that linear polyQ epitopes are exposed on the surface of GST-Ex1Q49 oligomers (**Fig. S1b**). Proteolytic cleavage of GST-Ex1Q49 with PP, however, caused a time-dependent decrease of antibody binding to the polyQ tract in Ex1Q49 fragments (**Fig. 2a and b**), indicating that the polyQ epitopes are less accessible for antibody binding in released Ex1Q49 fragments. We observed a more rapid loss of immunoreactivity with the MW1 than with the 3B5H10 antibody, indicating that the MW1-reactive polyQ conformation is accessible for antibody binding for a shorter time. We found that the rate of GST-Ex1Q49 fusion protein cleavage detected by SDS-PAGE and immunoblotting was similar to the rate of MW1-immunoreactivity loss observed by DBAs (**Fig. S3a**), indicating that proteolytic cleavage of the GST fusion protein is responsible for the time-dependent decrease in MW1 immunoreactivity. Thus, our studies with DBAs suggest that Ex1Q49 fragments, when released from GST fusions by proteolytic cleavage, undergo a rapid conformational rearrangement, which prevents MW1 antibody binding. This conformational change is likely to occur in soluble Ex1Q49 molecules and precedes spontaneous amyloid polymerization.

The decrease of 3B5H10 immunoreactivity was inversely correlated to the time-dependent increase of SDS-stable Ex1Q49 protein aggregates (compare **Fig. 1c and 2a and b**), suggesting that the 3B5H10 antibody exclusively binds to soluble Ex1Q49 molecules. To address this question further, preformed fibrillar Ex1Q49 aggregates were systematically

analysed with DBAs in independent experiments (**Fig. S3b**). We found that the 3B5H10 antibody did not recognize these structures, confirming our hypothesis that it preferentially binds to soluble Ex1Q49 molecules.

Using DBAs, we finally examined the binding of the epitope-specific antibody MW8 to aggregating Ex1Q49 molecules. This antibody recognizes a C-terminal proline-rich region in the Ex1Q49 fragment (**Fig. 1a**). We found that MW8 weakly binds to uncleaved GST-Ex1Q49 fusion protein in DBAs. Treatment of the fusion protein with PP further decreased MW8 immunoreactivity (**Fig. 2a and b**), indicating that the proline epitope is less accessible for antibody binding in released, aggregation-competent Ex1Q49 molecules than in uncleaved GST-Ex1Q49 oligomers. After a lag phase of ~2 h, however, a steep increase of MW8 immunoreactivity was observed (**Fig. 2b**), indicating that the antibody efficiently binds to spontaneously formed fibrillar Ex1Q49 aggregates (**Fig. 1e**). A similar result was observed when Ex1Q49 aggregation reactions were analysed in a time-dependent manner with denaturating FRAs that specifically detect large SDS-stable aggregates (**Fig. S3c**).

**Establishment of a theoretical branching model to describe spontaneous Ex1Q49 aggregation**

Our experimental results suggest that fibril branching could be responsible for the rapid formation of Ex1Q49 fibril bundles in cell-free assays (**Fig. 1f**). To better understand the molecular events underlying spontaneous Ex1Q49 polymerization, we next developed a theoretical aggregation model that, besides primary nucleation and templated polymerisation [11, 12, 16], also includes fibril branching as a secondary nucleation process (see Supplemental Data, Kinetic modelling of Ex1Q49 aggregation). Previously, various theoretical models have been reported that include primary nucleation, templated fibril polymerization, secondary

nucleation, and fragmentation [16, 17, 35]. However, none of these models explicitly account for the fate of newly nucleated branches that remain attached to mother fibrils as observed for instance for glucagon fibrils [44]. Since we did not observe an effect of agitation in our experimental studies (**Fig. S1e**), fragmentation was not included in our model. Our theoretical branching model describes the release of aggregation-prone Ex1Q49 molecules by proteolytic cleavage of GST-Ex1Q49 through the action of PP and the initial formation of a growth nucleus (primary nucleation; characterized by the size parameter $n_c$ and the rate constants $k_{n1}$ and $k_{-n1}$, **Fig. 3a**). Then, the nucleus grows into larger fibrillar structures by templated polymerization (red arrows) through the attachment of Ex1Q49 monomers. This process includes in one rate constant $k_1$ the binding to the template, i.e. to all aggregates equal or larger than the growth nucleus, and a conformational change of the monomers. Moreover, the generated Ex1Q49 fibrils can promote further nucleation (order $n_b$ and rate constant $k_b$), resulting in the formation of multiple growing fibril branches (**Fig. 3a**). Spontaneously formed fibrillar Ex1Q49 aggregates are very stable structures and have a very low propensity to dissociate (**Fig. 1b and c**), our theoretical model is therefore based on the assumption that polymerization and branching reactions are irreversible, allowing us to formulate a mathematically tractable model (see Materials and Methods and Supplemental Data).

**Theoretical models with or without branching can both describe the concentration-dependent formation of SDS-stable Ex1Q49 fibrils**

To compare experimental data and computational simulations, we first developed a normalization procedure and characterized the initial PP cleavage reaction. Using serial dilutions of reaction mixtures, we estimated the relationship between relative protein amounts and relative immunoreactivities for the antibodies applied for aggregate detection in FRAs (HD1 and CAG53b, **Fig. S4a**, Supplemental Data, Eqs. S30, S31). Finally, a log2 dependency was used to

convert simulated protein amounts to relative immunoreactivities, which could then be compared to the measured values. We also performed DBAs with the CAG53b antibody at different time points after addition of PP (**Fig. S4c**), which showed that CAG53b detection efficiency was independent of the Ex1Q49 aggregation state. We finally measured GST-Ex1Q49 cleavage by PP over time and found that the cleavage followed a Michaelis-Menten kinetics for which we estimated the kinetic parameters (**Fig. S4b, Supplemental Data, Eq. S5**).

Previous investigations have shown that through the analysis of kinetic aggregation data, which were collected with different protein concentrations, specific microscopic steps can be defined in complex polymerization reactions [16, 18], suggesting that such an approach may also provide insights into the mechanism of spontaneous HTTex1 aggregation. We therefore first used different concentrations of the GST-Ex1Q49 fusion protein in order to produce multiple time-resolved Ex1Q49 aggregation profiles. The experimental data were obtained with FRAs using the CAG53b antibody (**Fig. 3b**, dots) and then systematically compared with computational simulations. We found that the maximal aggregation rate was not affected by a change in the initial GST-Ex1Q49 concentration whereas the lag phase decreased with increasing protein concentrations (**Fig. 3b)**. Then, two theoretical HTTex1 aggregation models, one with ($k_b > 0$) and the other without branching ($k_b = 0$) were applied to fit the experimental data (see Supplemental Data**,** Fitting of HTTex1 kinetic models to data). By systematically varying the growth nucleus size $n_c$ and the order of the branching reaction $n_b$, we found that the experimental data is best reproduced for a growth nucleus of size 1-2 (**Fig. 3b-c**). Surprisingly, we obtained qualitatively very similar fits (according to the corrected Akaike information criterion) using both theoretical models, indicating that variations of protein concentrations alone are not sufficient to distinguish between a theoretical branching and a non-branching model. We suggest that this might be due to the fact that FRAs only detect Ex1Q49 fibrils when they are SDS-stable and have a length of >0.2 μm [33]. Therefore, to simulate the generated Ex1Q49 aggregation curves, we allowed for a size threshold for aggregate detection ($C_{min}$, see

Supplemental Data, **Eqs. S32-S33**), which likely prevents the distinction of the theoretical branching and non-branching aggregation models.

**The small molecule O4 delays spontaneous Ex1Q49 polymerization in cell-free assays**

We hypothesized that perturbation of spontaneous Ex1Q49 aggregation through the action of small molecules might yield data sets that could be helpful for theoretical modelling to elucidate specific aggregation mechanisms. We selected four previously reported amyloid-binding compounds, namely curcumin (Curc), methylene blue (MB), PGL-34 and O4 (**Table S1**) and systematically tested their activity in the established Ex1Q49 aggregation assay. We found that only the compound O4 affected Ex1Q49 aggregation in a concentration-dependent manner (**Figs 4a and b and S5a-c**), indicating that it directly targets aggregation-prone Ex1Q49 molecules. SDS-PAGE and immunoblotting confirmed that the effect of O4 on Ex1Q49 aggregation is not due to inhibition of PP cleavage (**Fig. S5d**).

More detailed investigations revealed that O4 significantly extends the lag phase of Ex1Q49 aggregation in a concentration-dependent manner (**Fig. S5e**), but does not significantly affect the maximal fibril growth rate. A similar effect of the compound was observed with DLS (**Fig. S5f**), indicating that O4 perturbs an early rather than a late event in the Ex1Q49 aggregation cascade.

Finally, we examined the effect of O4 on spontaneous Ex1Q49 aggregation using DBAs (**Fig. 4c-f**). We observed that the time-dependent decrease of 3B5H10 immunoreactivity is significantly delayed upon compound treatment, while no compound effect was observed on the decrease of MW1 immunoreactivity. This suggests that O4 preferentially targets soluble, 3B5H10-reactive Ex1Q49 molecules that are released from GST fusions and delays their incorporation into fibrillar aggregates. Strikingly, a delay in the increase of MW8

immunoreactivity was also observed (**Fig. 4f**), suggesting a causal relationship between disappearance of 3B5H10- and appearance of MW8-reactive Ex1Q49 molecules in spontaneous aggregation reactions. Together, these results support the data obtained by FRAs and DLS (**Fig. 4b and S5f**).

**Molecular dynamic simulations indicate that O4 directly targets aggregation-prone polyQ-containing HTTex1 molecules and alters their conformation**

We employed computational approaches for an assessment of how O4 might interact with aggregation-prone HTTex1 molecules, using a computationally generated β-sheet-rich structure of an HTT exon-1 protein with 47 glutamines (Ex1Q47), available in the literature [45]. This protein is similar in size to the Ex1Q49 fragment that was used in our experimental studies. Systematic docking studies suggested that O4 preferentially associates with the pathogenic polyQ tract and with the N-terminal 17 amino acids in Ex1Q47 (**Fig. 5a and Table S2**). Additional Quantum Mechanics/Molecular Mechanics Molecular Dynamics (QM/MM MD) simulations (see Materials and Methods for details) supported these results. The N-terminal residues in HTT were previously shown to promote protein aggregation in cell-free assays [28], substantiating our hypothesis that direct binding of O4 to soluble, aggregation-prone Ex1Q49 molecules might slow down the formation of fibrillar aggregates (**Fig. 4c-f**).

To examine whether O4 binding influences the β-sheet conformation in Ex1Q47 molecules, we performed additional molecular dynamics (MD) simulations, focussing, in this case, on putative compound bindings sites (e.g., A7, D1, A5 and A8) in the Ex1Q47 protein (**Table S2**). These simulations suggested that O4 binding promotes a time-dependent disassembly of the initial β-sheet-rich Ex1Q47 structure (**Fig. S6 and Movies S1 and S2**). For instance, cluster analysis of the MD trajectories of Ex1Q47 without and with O4 in the binding

site A7 indicated that 93% of the Ex1Q47 monomers feature a four-strand β-sheet structure in the absence of the compound (**Fig. 5b and Movie S1**). In the presence of O4, however, the population of molecules containing four-strand β-sheet is dramatically reduced to only 1% and structures without β-sheet prevail (74%, **Table S2**, **Fig. 5c and Movie S2**).

**The theoretical branching model correctly predicts the effect of O4 on spontaneous Ex1Q49 aggregation**

Our experimental studies indicate that O4 treatment delays the lag phase of spontaneous Ex1Q49 aggregation in a concentration-dependent manner, while it does not significantly influence the more rapid amyloid growth phase (**Fig. 4b and S5e**). These results together with the MD simulations suggest that the compound interferes with early rather than late events in the Ex1Q49 amyloid polymerization cascade. We thus asked, whether theoretical models (**Fig. 3a**) with or without branching could predict this behaviour when O4 inhibits the primary nucleation reaction (**Fig. 6a**, green line). We modelled the effect of O4 according to a general inhibition function for a non-competitive inhibitor (**Supplemental Data, Eq. S27**), where an increase in O4 decreases the nucleation rate constant ($k_{n1}$). For all simulations, we assumed that $n_c = n_b$. Models where $n_c \neq n_b$ gave less or equally good fits. We fitted the models to aggregation reactions with or without addition of O4 measured by FRAs (see Supplemental Data, Fitting of HTTex1 kinetic models to data; **Tables S3 and S4**). We found that the branching model with a growth nucleus of $n_c = 1$ and with O4 inhibiting primary nucleation accounts best for the observed kinetics (**Fig. 6b**, green squares; **Fig. 6c**), whereas the model without branching ($k_b = 0$) fails to reproduce the data (**Fig. 6b**, green circles; **Fig. S7a**).

We also tested whether the effect of O4 could be accounted for by an inhibition of the polymerization (**Fig. 6a**, red line) or the branching process (**Fig. 6a**, blue line). For a model

without branching, inhibition of polymerization only achieved a slight improvement in the fits (**Fig. 6b**, red circles compared to green circles). Interestingly, for models including branching and $n_c$ = 1, O4 inhibition of the branching or the polymerization process could also reproduce the data (**Fig. 6b**, red and blue squares).

To further test if O4 affects primary Ex1Q49 nucleation, we studied *in silico* whether the time point of O4 addition to aggregation reactions influences the onset of polymerization. If O4 inhibits primary nucleation ($k_{n1}$), the model predicted that addition of O4 at later time points (e.g., 1-2 h after aggregation initiation with PP) will be considerably less efficient in delaying Ex1Q49 aggregation onset than compound addition at t = 0 h together with PP (**Fig. 6d**). We confirmed experimentally that adding O4 at later time points to aggregation reactions was indeed less efficient in delaying Ex1Q49 aggregation (**Fig. 6e**). In contrast, simulations with theoretical models where O4 inhibits polymerization or branching, predicted that O4 addition at different time points can only partially relieve its inhibitory effect. Here, we observed a significant delay in Ex1Q49 aggregation even when O4 addition was set at 2 h after PP (**Fig. S7b**, shown for inhibition of branching). Taken together, our experimental and theoretical studies using the small molecule O4 support the hypothesis that branching shapes the aggregation kinetics of Ex1Q49 and likely is critical for the formation of bundles observed by AFM and EM (**Fig. 1e and f**).

**Predicting the impact of seeding on spontaneous Ex1Q49 aggregation**

Preformed aggregated β-sheet-rich polypeptide structures of varying sizes can act as nuclei or seeds that potently stimulate the formation of fibrillar structures [46, 47]. When they are added to aggregation reactions, they promote fibrillogenesis and shorten the lag phase prior to the fibril growth phase [48]. To characterize seeding effects, we studied the addition of preformed Ex1Q49 seeds on spontaneous HTTex1 aggregate formation using our theoretical

branching model. The model predicts that addition of increasing amounts of seeds to polymerization reactions causes a decrease in the lag phase without affecting the rate of fibril formation (slope, **Fig. 6f**, showing the percentage of protein bound in complexes larger than the simulated seed size). Interestingly, we found that the presence of seeds also changes the final aggregate size distribution in our model. In **Fig. 6g**, we show the average size at steady state, normalized to the average aggregate size without seed addition (seed fraction = 0), as a function of the seed fraction. For higher seed concentrations, fibril size is strongly decreased. The explanation for this is that the only pathways for aggregate growth in the model are templated polymerization and branching upon monomer addition, without considering higher order secondary aggregation processes of already formed fibrils. Therefore, high numbers of seeds compete for a limited pool of monomers leading to an overall smaller aggregate size when the monomers are eventually consumed. For low numbers of seeds, we observed a small increase in aggregate size. The reason for this increase is that the added seeds are outcompeting spontaneous nucleation and distributing the limited pool of monomers among fewer growing structures.

**Preformed Ex1Q49 seeds shorten the onset of spontaneous aggregate polymerization in a concentration-dependent manner**

To study the impact of preformed Ex1Q49 seeds on aggregate formation, we designed a FRET assay. We first produced and purified recombinant GST-HTTex1 fusion proteins that are C-terminally fused to the fluorescent proteins CyPet or YPet (GST-Ex1Q48-CyPet or -YPet; **Fig. S8a and b**). We hypothesized that these fusion proteins, when co-incubated with PP, release the aggregation-prone fragments Ex1Q48-CyPet and -YPet, which over time should co-assemble into ordered fibrillar aggregates (**Fig. S8b**), inducing a time-dependent increase of FRET.

First, we incubated different concentrations of the fusion proteins GST-Ex1Q48-CyPet and -YPet (1:1 molar ratio) in 384-well plates with PP and quantified the appearance of FRET over time. We observed a rapid increase of FRET efficiency after ~4-10 h, indicating that the proteolytically released HTTex1 fusion proteins, after a lag phase, rapidly self-assemble into ordered aggregates (**Fig. 7a**). Analysis of protein samples by AFM confirmed that the spontaneously formed Ex1Q48-CyPet/-YPet co-aggregates have a typical fibrillar morphology (**Fig. 7d**).

Next, we produced well-defined, fibrillar Ex1Q49 aggregates to investigate them for their seeding activity in the FRET-based HTTex1 aggregation assay. We incubated GST-Ex1Q49 fusion protein for 24 h with PP in order to obtain large bundles of fibrillar aggregates. Then, the generated fibrils were sonicated for 1 min on ice to break large fibrils and to produce small fibrillar seeds. Analysis by AFM confirmed that small, fibrillar Ex1Q49 seeds were obtained, while large bundles of HTTex1 fibrils (>1 μm) were detectable in non-sonicated samples (**Fig. 7b**).

Finally, we assessed whether the sonicated Ex1Q49 aggregates influence the co-polymerization of the reporter proteins Ex1Q48-CyPet and –YPet. We treated a mixture (1:1 molar ratio) of the fusion proteins GST-Ex1Q48-CyPet/-YPet with PP and various concentrations (~0.1-100 nM) of preformed sonicated Ex1Q49 seeds and analysed the co-polymerization of the released fusion proteins Ex1Q48-CyPet/-YPet by quantification of FRET. As shown in **Fig. 7c**, we observed a robust shortening of the lag phase when preformed Ex1Q49 seeds were added to reactions, indicating that these structures stimulate Ex1Q48-CyPet/-YPet co-polymerization in a concentration-dependent manner. With the highest seed concentrations, the maximum FRET efficiency was reached after ~4-5 h, while it was reached after ~15 h with the lowest seed concentrations. Thus, our experimental studies support our computational prediction that

addition of preformed HTTex1 seeds to reactions shortens the phase before onset of aggregate polymerization (**Fig. 6f**).

**High concentrations of preformed Ex1Q49 seeds promote the formation of small fibrillar HTTex1 aggregates**

Our theoretical investigations suggest that large numbers of very small HTTex1 fibrils should be obtained when spontaneous polymerization reactions are treated with high concentrations of preformed seeds (**Fig. 6g**). To validate this prediction experimentally, we collected samples from both seeded and non-seeded reactions after 50 h (**Fig. 7c,** black arrow) and systematically analysed them by AFM. Similar to untreated reactions, we observed few, very large bundles of HTTex1 fibrils in reactions treated with low seed concentrations (**Fig. 7d**). In contrast, many small fibril bundles and individual fibrils were detected in reactions treated with high seed concentrations. This indicates that complex bundles of HTTex1 fibrils are formed from small seeds/nuclei and grow by addition of monomers or very small oligomers. Once they are formed, they remain stable and do not further aggregate with one another in aqueous solutions. Thus, polyQ-containing HTTex1 fibrils are stable structures that do not continue to grow unless they are incubated with aggregation-competent HTTex1 molecules. These findings support our hypothesis that polyQ-mediated HTTex1 fibrillogenesis *in vitro* is dominated by a branching mechanism.

Finally, we analysed all aggregation reactions after 50 h with FRAs, which exclusively detect large (>0.2 μm), SDS-stable fibrillar HTTex1 aggregates retained on cellulose acetate membranes [33]. In reactions treated with high concentrations of Ex1Q49 seeds, no Ex1Q48-CyPet/-YPet aggregates were detectable on filter membranes (**Fig. 7e**), while they were readily observed in untreated reactions and reactions treated with low seed concentrations. This

confirms the initial hypothesis that addition of high seed concentrations to polymerization reactions leads to the formation of a multitude of small HTTex1 aggregates (**Fig. 6g)**. In reactions treated with high seed concentrations, however, anti-HTT immunoreactivity was detectable with DBAs (**Fig. 7e)**, which detect both soluble and aggregated HTTex1 proteins independent of their size. Together, these investigations support the observations made by AFM (**Fig. 7d**) and our theoretical predictions, indicating that Ex1Q49 seeds *in vitro* grow rapidly through branching and fibril elongation into large, complex fibrillar bundles.

## Discussion

In this study, we systematically investigated the aggregation mechanism of a pathogenic N-terminal HTT fragment released from a soluble recombinant GST-Ex1Q49 fusion protein by site-specific proteolytic cleavage using a combination of experimental and computational methods. We were able to show that two nucleation processes (primary and secondary nucleation) and fibril polymerization characterize Ex1Q49 aggregation. An earlier study already showed the importance of primary nucleation for the aggregation of polyQ peptides [27]. We found that all three processes are concomitant with conformational changes in Ex1Q49 molecules. In particular, primary nucleation can be accounted for by a slow conformational change in Ex1Q49, leading to the formation of aggregation-competent molecules. We propose that secondary nucleation is mediated by branching events at an existing fibril where the newly formed branches remain attached to the parent fibril. To study the importance of branching in the Ex1Q49 aggregation cascade, we developed a computational model that includes primary nucleation, branching and templated polymerization. Using independent experimental perturbations with the drug O4 and pre-formed seeds at different concentrations, we demonstrate that a theoretical branching model can well describe the experimental results, supporting a prominent role of branching in the Ex1Q49 aggregation process. Branching is essential to explain the fast generation of large, highly connected Ex1Q49 structures observed by AFM and EM (**Fig. 1e and f**). At the date of publication, active branching along existing fibrils has only been described for glucagon [44], but to our knowledge has not been theoretically modelled in the context of HD.

Our investigations of the early steps in the aggregate formation cascade indicate that uncleaved GST-Ex1Q49 fusion protein under non-denaturating conditions is a stable oligomer, which has a very low tendency to self-assemble. Upon proteolytic cleavage with PP, GST-Ex1Q49 oligomers disintegrate and release aggregation-prone Ex1Q49 molecules, which are

most probably monomers, according to our DLS experiments (**Fig. 1d)**. The released molecules then rapidly polymerize into SDS-stable aggregates with a typical fibrillar morphology (**Fig. 1d and e**). Our observations lead us to conclude that large, fibrillar Ex1Q49 aggregates are directly formed from interacting monomers. This is in contrast to previous reports that the assembly of transient, relatively stable spherical oligomers is a critical early event in the spontaneous HTTex1 aggregation cascade [49].

Our analysis of the early events in the Ex1Q49 polymerization cascade with DBAs and the epitope-specific anti-polyQ antibodies MW1 and 3B5H10 indicate that both antibodies readily detect the polyQ domain in soluble GST-Ex1Q49 oligomers (**Fig. 2**). However, MW1 antibody binding to the polyQ domain was rapidly lost upon cleavage of the GST fusion protein with PP. It has been shown previously that the MW1 antibody recognizes short linear epitopes in unfolded polyQ chains [43]. Our results suggest that linear polyQ chains are exposed to MW1 in GST-Ex1Q49 oligomers but not in released Ex1Q49 molecules. This supports previous reports that pathogenic polyQ tracts collapse and assume compact geometries in aqueous environments [50]. Previously published experiments also indicate that long polyQ chains collapse when they are analysed with FRET or TR-FRET assays [51, 52]. In contrast to MW1, released soluble Ex1Q49 molecules are readily detectable by the anti-polyQ antibody 3B5H10 (**Fig. 2**), indicating that this antibody is less polyQ conformation sensitive and recognizes both linear and collapsed polyQ chains under non-denaturating conditions.

We suggest that the proteolytically released 3B5H10-reactive Ex1Q49 molecules are non-β-sheet structures with a very low propensity to spontaneously self-assemble into fibrillar aggregates. This is also supported by recent time-dependent studies of the Ex1Q49 aggregation process with the dye Thioflavin T (E. Wanker, unpublished results), which preferentially detects amyloidogenic, β-sheet-rich structures [53] and studies with purified polyQ-containing HTTex1 fibrils that are not recognized by the 3B5H10 antibody (**Fig. S3b**). Furthermore, we hypothesize

that the 3B5H10-reactive molecules *per se* are aggregation-incompetent structures and their polyQ domains have to undergo a conformational change from a non-β-sheet into a stable β-sheet-rich structure in order to become aggregation-competent. This conformational conversion is an energetically unfavourable event that occurs only very rarely in a large population of 3B5H10-reactive Ex1Q49 molecules. However, once a stable β-sheet-rich Ex1Q49 molecule is spontaneously formed, it functions as nucleus (or seed) that can convert other non-β-sheet-rich Ex1Q49 molecules into an aggregation-competent state. Thus, we propose a spontaneous fibril formation mechanism in which a highly unfavourable folding event in soluble 3B5H10-reactive Ex1Q49 molecules is the rate limiting step (**Fig. 3a**).

Our theoretical modelling studies support our hypothesis that the spontaneous formation of β-sheet-rich Ex1Q49 monomers is the rate limiting step in the aggregation cascade as simulated aggregation profiles operating with a nucleus size of 1 ($n_c$ = 1) best describe the polymerization of Ex1Q49 molecules into amyloid fibrils (**Fig. 3 and 6**). Fibrillogenesis of simple polyQ peptides similar to Ex1Q49 has been reported to be dominated by an unfavourable protein folding step in monomers [30]. A conformational transition from a non-β-sheet into a β-sheet structure has also been observed in soluble polyQ monomers with biophysical methods [54], further substantiating our results with the aggregation-prone Ex1Q49 protein.

We assume that the polyQ domain in seeding-competent, β-sheet-rich Ex1Q49 molecules has a compact β-hairpin conformation. Such a structure has been previously predicted [55] and was recently identified with high resolution structural methods in fibrillar polyQ-containing HTTex1 aggregates [56]. In these molecules, polyQ β-hairpins are stabilized by steric-zipper-like interactions, supporting our results that fibrillar Ex1Q49 aggregates are highly stable structures. We suggest that 3B5H10-reactive, non-β-sheet-rich Ex1Q49 molecules are converted through templated nucleation into β-hairpin-rich structures that remain attached to each other (**Fig. 3a**). Thus, the observed complex fibril bundles (**Fig. 1e and f**) might be formed

from one Ex1Q49 monomer with a stable β-hairpin structure that grow over time in multiple directions by monomer addition. This fibril assembly mechanism is compatible with the recently proposed stochastic assembly mechanism for HTTex1 fibrils [56].

Previous studies indicate that HTTex1 fibrils once they are formed can further aggregate into even more complex amyloid structures [49]. Our experiments with preformed Ex1Q49 seeds do not support these published results (**Fig. 7b-d**). Rather, we observed that small Ex1Q49 fibrils, once formed in seeding reactions, neither change their size nor their morphology, when incubated in aqueous solutions for long periods of time (**Fig. 7c and d**). Thus, we propose that Ex1Q49 fibrils exclusively grow through monomer assembly; when monomers are used up fibril polymerization reactions cannot proceed.

Our studies provide evidence that Ex1Q49 fibrillogenesis *in vitro* is dominated by branching as the secondary nucleation process (**Fig. 1e and f**). This indicates that the mechanism underlying Ex1Q49 polymerization is distinct from that of other polypeptides such as the 42-residue amyloid-β (Aβ42) peptide, which forms unbranched fibrils [48]. Recent studies have shown that Aβ42 fibrils can catalyse the assembly of neurotoxic oligomers [18, 57], indicating that surface-catalysed secondary nucleation events play a key role in polymerization. In strong contrast to Ex1Q49 branches, however, these oligomers are not stably attached to the nucleating amyloid fibrils.

To clarify the role of branching in Ex1Q49 fibrillogenesis, we performed two perturbation experiments. First, we obtained evidence that the small molecule O4 delays primary nucleation when added to spontaneous Ex1Q49 polymerization reactions. Second, we investigated the impact of preformed seeds at different concentrations on Ex1Q49 aggregation. The resulting aggregation profiles from these studies in combination with our computational models strongly supported branching to be involved in the formation of HTTex1 fibrils (**Fig. 6 and 7**). The proposed perturbations therefore represent an alternative approach to the variation of Ex1Q49

concentrations, in particular when this strategy, as in our experimental system (**Fig. 3**), is not conclusive with regard to the presence or absence of a branching mechanism (**Fig. 3c**). The specific perturbation of aggregation kinetics by inhibitors has been previously shown for other amyloid peptides [57, 58], supporting our experimental strategy.

Our time-resolved investigations of spontaneous Ex1Q49 aggregation *in vitro* indicate that small spherical oligomers, which have been previously described in cell-free assays [59, 60] do not play a major role in the fibril polymerization process. Such structures, however, may form *in vivo* under conditions where polyQ-mediated HTTex1 aggregation occurs in a complex cellular environment and the rate of mutant HTT polymerization is slowed down by interacting proteins or post-translational modifications [61-64]. However, high resolution studies indicate that, besides large perinuclear inclusion bodies with highly complex fibrillar structures [65, 66], also smaller bundles of fibrillar HTTex1 aggregates with pathogenic polyQ tracts are detectable in the cytoplasm of mammalian cells [67, 68], supporting our hypothesis that spontaneous HTTex1 aggregation both *in vitro* and *in vivo* is dominated by the fibril formation pathway. We suggest that complex bundles of fibrillar HTT aggregates are important disease-relevant structures. They are very stable molecules and likely are degraded inefficiently by cellular protein quality systems such as autophagy. This process was shown to be impaired in HD models [69, 70]. Complex bundles of mutant HTT fibrils also might be responsible for the recruitment of multiple, critical factors, which could lead to aberrant cellular functions. Recent studies indicate that multiple cellular proteins aberrantly interact with mutant HTTex1 aggregates [71]. Also, experimental evidence was obtained that proline-rich domains are exposed many times on the surface of HTTex1 fibrils [32, 72], suggesting that such structures might efficiently recruit SH3 and WW domain-containing signalling proteins and thereby perturb their normal cellular functions. Such proteins have previously been demonstrated to bind to mutant HTTex1 aggregates in various disease model systems [73, 74]. Furthermore, recent studies indicate that highly complex fibrillar HTTex1 structures interact with endomembranes in cells and lead to their local impairment [66].

Thus, the formation complex bundles of fibrillar HTT aggregates in neurons through branching may be of high relevance for the development and progression of HD.

Our mechanistic *in vitro* studies that show a critical role of branching for the aggregation of Ex1Q49 have several important implications for the development of disease-modifying therapeutic strategies for HD. For example, we found that the small molecule O4 delays very early nucleation events in cell-free assays (**Fig. 4a-f**), indicating that it slows down the formation of aggregation-competent conformations, which are a prerequisite for fibril assembly. If we assume that seeding-competent, fibrillar HTT aggregates are the toxic species in pathogenesis [24], compounds like O4 should delay fibrillogenesis and improve symptoms in HD patients. In order to be effective, however, treatment with such compounds most probably would need to start before fibrillar, seeding-competent aggregates have formed, because propagation of such structures in patient brains likely cannot be halted. Such specific questions can now be followed up in experiments with transgenic and knock-in mouse models of HD using our set of epitope-specific antibodies and small molecules that target specific events in the Ex1Q49 aggregation cascade.

## Materials and Methods

**Protein purification and *in vitro* analysis of aggregates.** The proteins GST-Ex1Q49, GST-Ex1Q48-CyPet, GST-Ex1Q48-YPet and GST were produced in *E. coli* BL21-RP cells and purified as described previously [21]. Filter retardation assays (FRAs), dot blot assays (DBAs), SDS-PAGE, dynamic light scattering, size exclusion chromatography and atomic force microscopy of aggregates were performed using standard methods. Details are provided in *Supplemental Data*.

**HTTex1 protein aggregation**. Aggregation of the Ex1Q49 fragment was initiated by addition of 0.28 U PreScission (PP) protease to 2 µM GST-Ex1Q49 fusion protein (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA and 1 mM DTT). Reactions were incubated at 20°C (300 rpm) for 1-8 h. Samples were taken at indicated time points, frozen on dry ice and stored at -80°C. Aggregation of the Ex1Q48-CyPet/-YPet fragments was performed under comparable conditions. Details are provided in *Supplemental Data*.

**Computational details for docking studies and MD simulations.** O4-Ex1Q47 docking studies were performed using the program AutodockVina [75] and the previously reported structural model Ex1Q47 [45]. Quantum Mechanics/Molecular Mechanics Molecular Dynamics (QM/MM MD) simulations in explicit solvent were performed using the program CHARMM (v33b13) [76] with the CHARMM22 force field [77] for the MM region and the SCC-DFTB-D [78] method for the QM region (O4 molecule). Additional classical MD simulations were performed with the NAMD program [79]. More details are provided in *Supplemental Data - Computational Details for Docking Studies and MD Simulations*.

**Kinetic modelling.** Models for the various aggregation mechanisms are given by sets of ordinary differential equations (ODEs). Binding and conformational transitions are described by mass-action kinetics. We used moments of the distribution to derive a finite set of equations for

describing the aggregation process. These equations are solved numerically using MATLAB ode15s function. All details are provided in *Supplemental Data - Kinetic Modeling of Ex1Q49 Aggregation*.

## Acknowledgments

E.S.-G. thanks N. V. Dokholyan, University of North Carolina at Chapel Hill, US for kindly providing his calculated Ex1Q47 structure and M. Doerr for useful discussions. E.S.-G. acknowledges a Plus-3 Grant by the Boehringer Ingelheim foundation. E.S.-G. also acknowledges the financial support of the Cluster of Excellence RESOLV (EXC 1069) and the SFB1093, both funded by the DFG. E.E.W thanks K. Rau and G. Grelle for AFM studies, B. Purfürst for performing electron microscopy investigations and S. Schnögl for critical reading of the manuscript. This project was supported by grants from the HDSA, DFG (SFB740 and SFB618), BMBF (NGFN-Plus: NeuroNet), EU (EuroSpin and SynSys) and the Helmholtz Association (MSBN and HeIMA). The relevant grant numbers are SFB740: 740/2-11, SFB618: 618/3-09, NeuroNet: 01GS08169-73, EuroSpin: Health-F2-2009-241498, SynSys: HEALTH-F2-2009-242167, MSBN: NW2, HelMA: HA-215. J.W. was supported by the FORSYS program of the BMBF (grant number 0315289), and the Helmholtz Alliance on Systems Biology.

The authors declare that they have no conflict of interest.

## Author contributions

A.S.W. designed and performed all time-resolved aggregation experiments in this study; N.U.S. and A.B. performed AFM experiments; S.P. produced and characterized recombinant proteins; E.E.W. coordinated experimental and theoretical studies and designed the overall experimental

strategy; A.Z.P. performed the DLS experiments; A.Z.P. designed the kinetic models and performed the computational simulations with the help of J.W. and K.B.; K.B-R and J.M.R-A performed the docking experiments, the MD simulations and processed the corresponding computational data; E.S-G performed the QM/MM MD simulations, planned and directed the computational work; A.S., A.B. and L.B. performed seeding experiments; C.H. analyzed data and generated figures; A.S.W., E.E.W., A.Z.P., J.W., and E.S-G. wrote the paper.

## Supplemental Data

Supplemental Methods

Supplemental References

Legends of Supplemental Figures S1 to S8

Legends of Supplemental Movies S1 and S2

Supplemental Tables S1 to S4 and Legends

Supplemental Figures S1 to S8

Supplemental Movies S1 and S2 (separate files)

## References


[1] Dobson CM. Protein folding and misfolding. Nature. 2003;426:884-90.
[2] Bellotti V, Chiti F. Amyloidogenesis in its biological environment: challenging a fundamental issue in protein misfolding diseases. Curr Opin Struct Biol. 2008;18:771-9.
[3] Rambaran RN, Serpell LC. Amyloid fibrils: abnormal protein assembly. Prion. 2008;2:112-7.
[4] Sunde M, Blake C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. Adv Protein Chem. 1997;50:123-59.
[5] Taylor JP, Hardy J, Fischbeck KH. Toxic proteins in neurodegenerative disease. Science. 2002;296:1991-5.
[6] Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. Annual review of biochemistry. 2006;75:333-66.
[7] Tipping KW, van Oosten-Hawle P, Hewitt EW, Radford SE. Amyloid Fibres: Inert End-Stage Aggregates or Key Players in Disease? Trends in biochemical sciences. 2015;40:719-27.
[8] Fowler DM, Koulov AV, Balch WE, Kelly JW. Functional amyloid--from bacteria to humans. Trends Biochem Sci. 2007;32:217-24.
[9] Shorter J, Lindquist S. Prions as adaptive conduits of memory and inheritance. Nat Rev Genet. 2005;6:435-50.
[10] Cohen SI, Vendruscolo M, Dobson CM, Knowles TP. From macroscopic measurements to microscopic mechanisms of protein aggregation. Journal of molecular biology. 2012;421:160-71.
[11] Morris AM, Watzky MA, Finke RG. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. Biochimica et biophysica acta. 2009;1794:375-97.
[12] Ferrone F. Analysis of protein aggregation kinetics. Methods in enzymology. 1999;309:256-74.
[13] Gillam JE, MacPhee CE. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. Journal of physics Condensed matter : an Institute of Physics journal. 2013;25:373101.
[14] Cohen SI, Vendruscolo M, Welland ME, Dobson CM, Terentjev EM, Knowles TP. Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. The Journal of chemical physics. 2011;135:065105.
[15] Ferrone FA, Hofrichter J, Eaton WA. Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. Journal of molecular biology. 1985;183:611-31.
[16] Xue WF, Homans SW, Radford SE. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:8926-31.
[17] Ruschak AM, Miranker AD. Fiber-dependent amyloid formation as catalysis of an existing reaction pathway. Proceedings of the National Academy of Sciences of the United States of America. 2007;104:12341-6.
[18] Cohen SI, Linse S, Luheshi LM, Hellstrand E, White DA, Rajah L, et al. Proliferation of amyloid-beta42 aggregates occurs through a secondary nucleation mechanism. Proceedings of the National Academy of Sciences of the United States of America. 2013;110:9758-63.
[19] Bernacki JP, Murphy RM. Model discrimination and mechanistic interpretation of kinetic data in protein aggregation studies. Biophysical journal. 2009;96:2871-87.
[20] Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, et al. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell. 1997;90:549-58.
[21] Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, et al. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:4604-9.

[22] Rochet JC, Lansbury PT, Jr. Amyloid fibrillogenesis: themes and variations. Curr Opin Struct Biol. 2000;10:60-8.
[23] Poirier MA, Li H, Macosko J, Cai S, Amzel M, Ross CA. Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. J Biol Chem. 2002;277:41032-7.
[24] Pieri L, Madiona K, Bousset L, Melki R. Fibrillar alpha-synuclein and huntingtin exon 1 assemblies are toxic to the cells. Biophysical journal. 2012;102:2894-905.
[25] Sontag EM, Lotz GP, Yang G, Sontag CJ, Cummings BJ, Glabe CG, et al. Detection of Mutant Huntingtin Aggregation Conformers and Modulation of SDS-Soluble Fibrillar Oligomers by Small Molecules. Journal of Huntington's disease. 2012;1:119-32.
[26] Vieweg S, Ansaloni A, Wang ZM, Warner JB, Lashuel HA. An Intein-based Strategy for the Production of Tag-free Huntingtin Exon 1 Proteins Enables New Insights into the Polyglutamine Dependence of Httex1 Aggregation and Fibril Formation. The Journal of biological chemistry. 2016;291:12074-86.
[27] Chen S, Ferrone FA, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proceedings of the National Academy of Sciences of the United States of America. 2002;99:11884-9.
[28] Thakur AK, Jayaraman M, Mishra R, Thakur M, Chellgren VM, Byeon IJ, et al. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. Nature structural & molecular biology. 2009;16:380-9.
[29] Sathasivam K, Neueder A, Gipson TA, Landles C, Benjamin AC, Bondulich MK, et al. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. Proceedings of the National Academy of Sciences of the United States of America. 2013;110:2366-70.
[30] Bhattacharyya AM, Thakur AK, Wetzel R. polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction. Proceedings of the National Academy of Sciences of the United States of America. 2005;102:15400-5.
[31] Kar K, Jayaraman M, Sahoo B, Kodali R, Wetzel R. Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent. Nature structural & molecular biology. 2011;18:328-36.
[32] Shen K, Calamini B, Fauerbach JA, Ma B, Shahmoradian SH, Serrano Lachapel IL, et al. Control of the structural landscape and neuronal proteotoxicity of mutant Huntingtin by domains flanking the polyQ tract. eLife. 2016;5.
[33] Wanker EE, Scherzinger E, Heiser V, Sittler A, Eickhoff H, Lehrach H. Membrane filter assay for detection of amyloid-like polyglutamine-containing protein aggregates. Methods in enzymology. 1999;309:375-86.
[34] Muchowski PJ, Schaffar G, Sittler A, Wanker EE, Hayer-Hartl MK, Hartl FU. Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:7841-6.
[35] Knowles TP, Waudby CA, Devlin GL, Cohen SI, Aguzzi A, Vendruscolo M, et al. An analytical solution to the kinetics of breakable filament assembly. Science. 2009;326:1533-7.
[36] Liu R, He M, Su R, Yu Y, Qi W, He Z. Insulin amyloid fibrillation studied by terahertz spectroscopy and other biophysical methods. Biochemical and biophysical research communications. 2010;391:862-7.
[37] Nichols MR, Colvin BA, Hood EA, Paranjape GS, Osborn DC, Terrill-Usery SE. Biophysical comparison of soluble amyloid-beta(1-42) protofibrils, oligomers, and protofilaments. Biochemistry. 2015;54:2193-204.
[38] Bulone D, Masino L, Thomas DJ, San Biagio PL, Pastore A. The interplay between PolyQ and protein context delays aggregation by forming a reservoir of protofibrils. PLoS One. 2006;1:e111.
[39] Hollenbach B, Scherzinger E, Schweiger K, Lurz R, Lehrach H, Wanker EE. Aggregation of truncated GST-HD exon 1 fusion proteins containing normal range and expanded glutamine

repeats. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 1999;354:991-4.
[40] Ehrnhoefer DE, Duennwald M, Markovic P, Wacker JL, Engemann S, Roark M, et al. Green tea (-)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models. Hum Mol Genet. 2006;15:2743-51.
[41] Li P, Huey-Tubman KE, Gao T, Li X, West AP, Jr., Bennett MJ, et al. The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model. Nat Struct Mol Biol. 2007;14:381-7.
[42] Peters-Libeu C, Miller J, Rutenber E, Newhouse Y, Krishnan P, Cheung K, et al. Disease-associated polyglutamine stretches in monomeric huntingtin adopt a compact structure. J Mol Biol. 2012;421:587-600.
[43] Klein FA, Zeder-Lutz G, Cousido-Siah A, Mitschler A, Katz A, Eberling P, et al. Linear and extended: a common polyglutamine conformation recognized by the three antibodies MW1, 1C2 and 3B5H10. Human molecular genetics. 2013;22:4215-23.
[44] Andersen CB, Yagi H, Manno M, Martorana V, Ban T, Christiansen G, et al. Branching in amyloid fibril growth. Biophysical journal. 2009;96:1529-36.
[45] Lakhani VV, Ding F, Dokholyan NV. Polyglutamine induced misfolding of huntingtin exon1 is modulated by the flanking sequences. PLoS Comput Biol. 2010;6:e1000772.
[46] Jucker M, Walker LC. Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders. Annals of neurology. 2011;70:532-40.
[47] Jeong JS, Ansaloni A, Mezzenga R, Lashuel HA, Dietler G. Novel mechanistic insight into the molecular basis of amyloid polymorphism and secondary nucleation during amyloid formation. Journal of molecular biology. 2013;425:1765-81.
[48] Harper JD, Lansbury PT, Jr. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. Annual review of biochemistry. 1997;66:385-407.
[49] Duim WC, Jiang Y, Shen K, Frydman J, Moerner WE. Super-resolution fluorescence of huntingtin reveals growth of globular species into short fibers and coexistence of distinct aggregates. ACS chemical biology. 2014;9:2767-78.
[50] Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:16764-9.
[51] Schaffar G, Breuer P, Boteva R, Behrends C, Tzvetkov N, Strippel N, et al. Cellular toxicity of polyglutamine expansion proteins: mechanism of transcription factor deactivation. Molecular cell. 2004;15:95-105.
[52] Fodale V, Kegulian NC, Verani M, Cariulo C, Azzollini L, Petricca L, et al. Polyglutamine- and temperature-dependent conformational rigidity in mutant huntingtin revealed by immunoassays and circular dichroism spectroscopy. PloS one. 2014;9:e112262.
[53] Biancalana M, Koide S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochimica et biophysica acta. 2010;1804:1405-12.
[54] Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, Urade Y, et al. A toxic monomeric conformer of the polyglutamine protein. Nature structural & molecular biology. 2007;14:332-40.
[55] Wanker EE. Protein aggregation and pathogenesis of Huntington's disease: mechanisms and correlations. Biological chemistry. 2000;381:937-42.
[56] Hoop CL, Lin HK, Kar K, Magyarfalvi G, Lamley JM, Boatz JC, et al. Huntingtin exon 1 fibrils feature an interdigitated beta-hairpin-based polyglutamine core. Proceedings of the National Academy of Sciences of the United States of America. 2016;113:1546-51.
[57] Cohen SI, Arosio P, Presto J, Kurudenkandy FR, Biverstal H, Dolfe L, et al. A molecular chaperone breaks the catalytic cycle that generates toxic Abeta oligomers. Nature structural & molecular biology. 2015;22:207-13.

[58] Abelein A, Graslund A, Danielsson J. Zinc as chaperone-mimicking agent for retardation of amyloid beta peptide fibril formation. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:5407-12.
[59] Legleiter J, Lotz GP, Miller J, Ko J, Ng C, Williams GL, et al. Monoclonal antibodies recognize distinct conformational epitopes formed by polyglutamine in a mutant huntingtin fragment. The Journal of biological chemistry. 2009;284:21647-58.
[60] Nucifora LG, Burke KA, Feng X, Arbez N, Zhu S, Miller J, et al. Identification of novel potentially toxic oligomers formed in vitro from mammalian-derived expanded huntingtin exon-1 protein. The Journal of biological chemistry. 2012;287:16017-28.
[61] Stroedicke M, Bounab Y, Strempel N, Klockmeier K, Yigit S, Friedrich RP, et al. Systematic interaction network filtering identifies CRMP1 as a novel suppressor of huntingtin misfolding and neurotoxicity. Genome research. 2015;25:701-13.
[62] Bustamante MB, Ansaloni A, Pedersen JF, Azzollini L, Cariulo C, Wang ZM, et al. Detection of huntingtin exon 1 phosphorylation by Phos-Tag SDS-PAGE: Predominant phosphorylation on threonine 3 and regulation by IKKbeta. Biochemical and biophysical research communications. 2015;463:1317-22.
[63] Sontag EM, Joachimiak LA, Tan Z, Tomlinson A, Housman DE, Glabe CG, et al. Exogenous delivery of chaperonin subunit fragment ApiCCT1 modulates mutant Huntingtin cellular phenotypes. Proceedings of the National Academy of Sciences of the United States of America. 2013;110:3077-82.
[64] Monsellier E, Bousset L, Melki R. alpha-Synuclein and huntingtin exon 1 amyloid fibrils bind laterally to the cellular membrane. Scientific reports. 2016;6:19180.
[65] Waelter S, Boeddrich A, Lurz R, Scherzinger E, Lueder G, Lehrach H, et al. Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. Molecular biology of the cell. 2001;12:1393-407.
[66] Bauerlein FJB, Saha I, Mishra A, Kalemanov M, Martinez-Sanchez A, Klein R, et al. In Situ Architecture and Cellular Interactions of PolyQ Inclusions. Cell. 2017;171:179-87 e10.
[67] Sahl SJ, Weiss LE, Duim WC, Frydman J, Moerner WE. Cellular inclusion bodies of mutant huntingtin exon 1 obscure small fibrillar aggregate species. Scientific reports. 2012;2:895.
[68] Sahl SJ, Lau L, Vonk WI, Weiss LE, Frydman J, Moerner WE. Delayed emergence of subdiffraction-sized mutant huntingtin fibrils following inclusion body formation. Quarterly reviews of biophysics. 2016;49:e2.
[69] Ashkenazi A, Bento CF, Ricketts T, Vicinanza M, Siddiqi F, Pavel M, et al. Polyglutamine tracts regulate beclin 1-dependent autophagy. Nature. 2017;545:108-11.
[70] Heng MY, Duong DK, Albin RL, Tallaksen-Greene SJ, Hunter JM, Lesort MJ, et al. Early autophagic response in a novel knock-in model of Huntington disease. Human molecular genetics. 2010;19:3702-20.
[71] Kim YE, Hosp F, Frottin F, Ge H, Mann M, Hayer-Hartl M, et al. Soluble Oligomers of PolyQ-Expanded Huntingtin Target a Multiplicity of Key Cellular Factors. Molecular cell. 2016;63:951-64.
[72] Lin HK, Boatz JC, Krabbendam IE, Kodali R, Hou Z, Wetzel R, et al. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core. Nature communications. 2017;8:15462.
[73] Sittler A, Walter S, Wedemeyer N, Hasenbank R, Scherzinger E, Eickhoff H, et al. SH3GL3 associates with the Huntingtin exon 1 protein and promotes the formation of polygln-containing protein aggregates. Molecular cell. 1998;2:427-36.
[74] Davranche A, Aviolat H, Zeder-Lutz G, Busso D, Altschuh D, Trottier Y, et al. Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. Human molecular genetics. 2011;20:2795-806.
[75] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455-61.

[76] Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem. 1983;4:187-217.
[77] MacKerell AD, Jr., Feig M, Brooks CL, 3rd. Improved treatment of the protein backbone in empirical force fields. J Am Chem Soc. 2004;126:698-9.
[78] Elstner M. SCC-DFTB: what is the proper degree of self-consistency? J Phys Chem A. 2007;111:5614-21.
[79] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005;26:1781-802.

## Figure Legends

**Figure 1. Spontaneous formation of large SDS-resistant Ex1Q49 aggregates *in vitro*.**

(a) Schematic representation of GST-Ex1Q49 fusion protein and epitope binding sites of monoclonal antibodies used.

(b) Time-resolved analysis of SDS-resistant Ex1Q49 aggregates by FRA. Ex1Q49 aggregation was initiated by proteolytic cleavage of GST-Ex1Q49 fusion protein (2 µM) with PP. Aggregates retained on filter membranes were detected using the polyclonal anti-HTT antibody CAG53b.

(c) Quantification of spontaneously formed Ex1Q49 aggregates by FRAs (solid line, mean ± SD, n = 4). 2 µM GST-Ex1Q49 fusion protein was incubated with PP at 20°C. The highest signals were set to 100%.

(d) Analysis of Ex1Q49 aggregation (20.9 µM) by dynamic light scattering. Left panel: particle size distribution prior (0 h) and after PP addition (at 0.5 and 3 h). Right panel: color-coded particle diameter distribution over time. Shown are the results for one representative experiment out of four.

(e) Analysis of Ex1Q49 aggregation reactions by AFM at different time points after addition of PP.

(f) Analysis of Ex1Q49 fibril bundles by TEM revealed potential fibril branching points, which are indicated by triangles. Ex1Q49 aggregates were examined 8 h after addition of PP at 20°C.

**Figure 2. Investigation of Ex1Q49 aggregation under non-denaturing conditions using epitope-specific antibodies.**

(a) Analysis of spontaneous Ex1Q49 aggregation by non-denaturing DBAs. Samples were spotted onto a nitrocellulose membrane; retained Ex1Q49 protein was detected using the epitope-specific antibodies MW1, 3B5H10 or MW8.

(b) Quantification of antibody immunoreactivity. The highest signals were set to 100%.

**Figure 3. A branching model to describe Ex1Q49 aggregation.**

(a) Schematic for a branching model. After PP cleavage of GST, Ex1Q49 can spontaneously nucleate and switch to an active aggregation-prone conformation with rate constant $k_{n1}$. Fibrils can then elongate by templated addition of monomers, rate constant $k_1$. In addition, a branching reaction can occur with rate constant $k_b$. This creates a new growing fibril. If the branching rate constant $k_b = 0$, the model reduces to a linear templated aggregation model.

(b) Aggregation of Ex1Q49 at different initial GST-Ex1Q49 concentrations monitored using the FRA assay and the CAG53b antibody (dots). Dashed and solid lines show the results for models without ($k_b = 0$, $n_c = 2$) or with ($k_b > 0$, $n_c = 2$, $n_b = 2$) branching.

(c) Difference in the corrected Akaike information criterion (*AICc*, Eq. S35) with respect to the best model for models without branching (gray circle) or branching (blue symbols) and different values of the nucleus size $n_c$ and the branching order $n_b$. A large value indicates a less good fit compared to the best model. Models were fitted to the experimental data from (b) as described in Supplemental Data.

**Figure 4. The small molecule O4 influences an early step in the Ex1Q49 polymerization cascade.**

(a) Chemical structure of O4.

(b) Analysis of spontaneous Ex1Q49 aggregation in the presence and absence of O4 by FRA. GST-Ex1Q49 fusion protein (2 µM) was incubated with PP and different concentrations of O4 at 20°C. SDS-stable aggregates retained on filter membranes were immunodetected with CAG53b antibody.

(c) Analysis of Ex1Q49 aggregation reactions with DBAs using the monoclonal antibodies MW1, 3B5H10 and MW8. GST-Ex1Q49 (2 µM) fusion protein was incubated with PP and O4 (2 µM).

(d-f) Quantification of dot blot results shown in (c).

**Figure 5. The compound O4 perturbs the structure of an N-terminal, β-sheet-rich Ex1Q47 fragment.**

(a) Computational docking and QM/MM MD studies predict that O4 binds to the N-terminal amino acids and the polyQ regions in Ex1Q47. O4 is colored according to the binding motif: A7 purple, D1 orange, A5 yellow and A8 cyan (Fig. S6). Different colors are used for the Ex1Q47 regions: N-terminus in blue, polyQ in red, polyP in gray and the mixed region including the C-terminus is shown in green. One of the most favored binding motifs, A7, is displayed in more detail. The residues within a 3 Å distance between O4 and Ex1Q47 are shown.

(b-c) O4 alters the β-sheet structure of Ex1Q47 monomers. Representative most populated structures of Ex1Q47 (MD simulations); (b) reference system without O4 and (c) Ex1Q47 in the Ex1Q47-O4 system (binding site A7, O4 omitted). See also Fig. S6.

**Figure 6. The theoretical branching model correctly reproduces the effect of O4 on spontaneous Ex1Q49 aggregation and predicts the impact of seeds on HTTex1 polymerization.**

(a) In the model O4 can either decrease the nucleation rate constant $k_{n1}$ (green), the branching rate constant $k_b$ (blue), or the polymerization rate constant $k_1$ (red) in a concentration dependent way.

(b) Difference in the corrected Akaike information criterion (*AICc*) with respect to the best fitting model for three different O4 mediated inhibition scenarios (shown in colored symbols) and the model without branching ($k_b = 0$, circle) or with branching ($k_b > 0$, squares). Models were fitted to the experimental data from (c) (see Supplemental Data). Parameters for the best fitting model are shown in Tables S3 and S4.

(c) Simulated FRA data from the branching model ($n_c = 1$ and $n_b = 1$) without (solid lines) and with O4 (dashed lines). The experimental data (symbols, mean ± std) show aggregation for 2 different initial GST-Ex1Q49 concentrations and 3 different O4 concentrations as indicated. Experimental data and model were normalized to maximal value of the control with 2 µM initial GST-Ex1Q49.

(d) Simulations of aggregation reactions of the best fitting model from (b) when O4 is added at different time points post PP addition as indicated.

(e) 2 µM GST-Ex1Q49 was incubated in the absence (control) or presence of 2 µM O4 added together with PP (0 h), 30 min, 1 h, or 2 h post PP addition. Aliquots were analysed by FRAs and were immunodetected with CAG53b antibody. Data were normalized to the control value at 7 h.

(f) Simulation of the best fitting model from (b) with addition of different fractions of seeds [Seed]/[GST-Ex1Q49]. Shown is the percentage of protein bound in complexes above

the seed size. In the simulations, we use dimers as seeds, this is the smallest possible seed size.

(g) Average simulated amyloid size at steady-state for different fractions of seeds computed from Eq. S26. The average amyloid size is normalized to the average size without seeds.

**Figure 7. Seed-induced HTTex1 aggregation experiments confirm the predictions of the mathematical branching model.**

(a) Analysis of GST-Ex1Q48-CyPet/-YPet co-aggregation by time-resolved quantification of FRET. Aggregation was initiated by proteolytic cleavage of the GST-tag with PP.

(b) Preformed Ex1Q49 fibrils were visualized by AFM before and after fragmentation through sonication for 60 seconds. Scale bar: 1 µm.

(c) Co-aggregation of the sensor proteins Ex1Q48-CyPet/-YPet was monitored in the presence of different concentrations of sonicated Ex1Q49 aggregates (seeds) by quantification of FRET for 50 h. Indicated seed concentrations are equivalent to initially applied monomer concentrations. The black arrow indicates the time point at which aggregation reactions were analysed by DBAs, FRAs and AFM shown in (d) and (e).

(d) AFM analysis of GST-Ex1Q48-CyPet/-YPet co-aggregates formed in the absence and presence of different seed concentrations; samples were analysed after 50 h. Scale bar: 1 µm.

(e) Ex1Q48-CyPet/-YPet co-aggregates shown in (d) were analysed by DBAs and FRAs. Proteins were detected using an anti-GFP antibody.

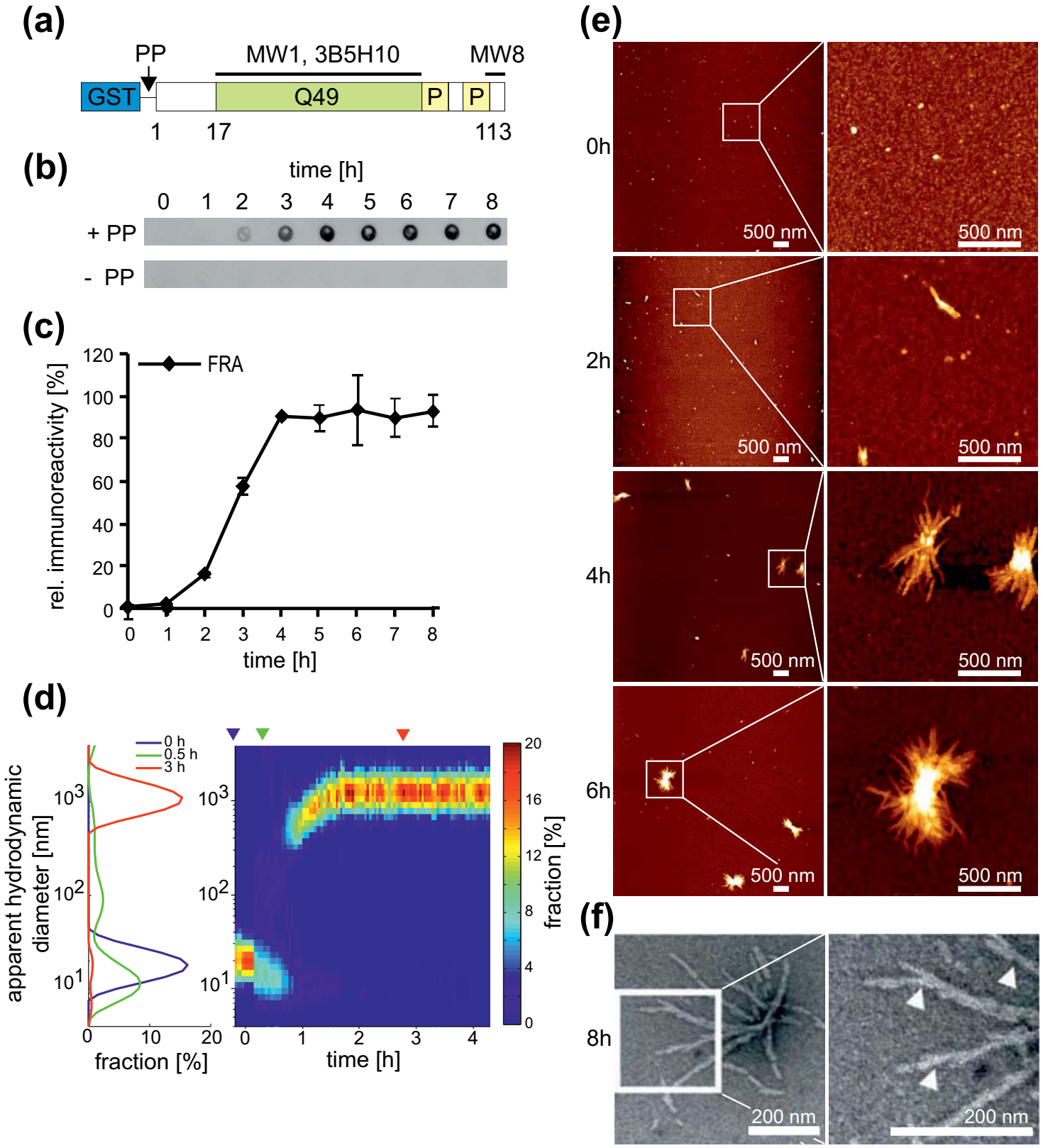
(a)
PP
MW1, 3B5H10
MW8
GST
Q49
P
P
1
17
113
(b)
time [h]
0 1 2 3 4 5 6 7 8
+ PP
- PP
(c)
FRA
rel. immunoreactivity [%]
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
time [h]
(d)
0 h
0.5 h
3 h
apparent hydrodynamic diameter [nm]
$10^3$
$10^2$
$10^1$
0 10 20
fraction [%]
0 1 2 3 4
time [h]
20
16
12
8
4
0
fraction [%]
(e)
0h
2h
4h
6h
500 nm
(f)
8h
200 nm

**Figure 1**

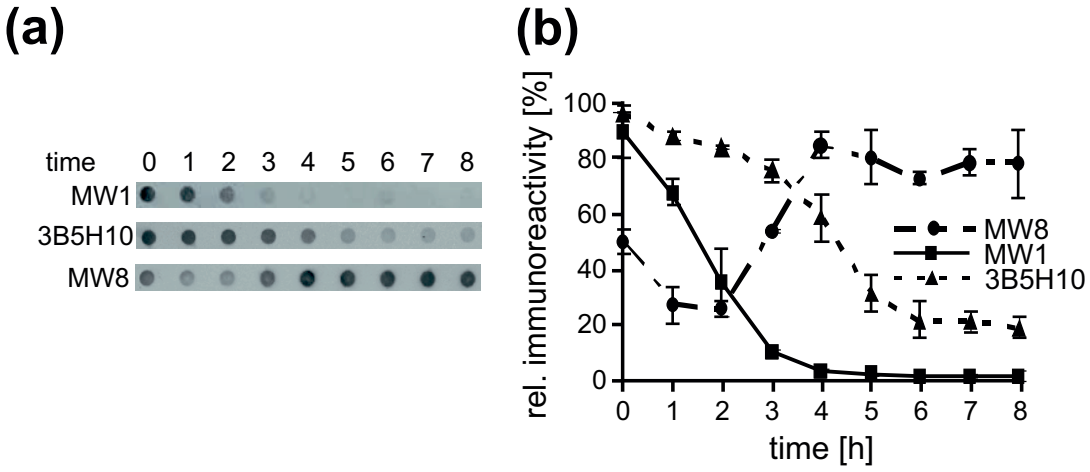
(a)
time 0 1 2 3 4 5 6 7 8
MW1
3B5H10
MW8
(b)
rel. immunoreactivity [%]
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
time [h]
MW8
MW1
3B5H10

**Figure 2**

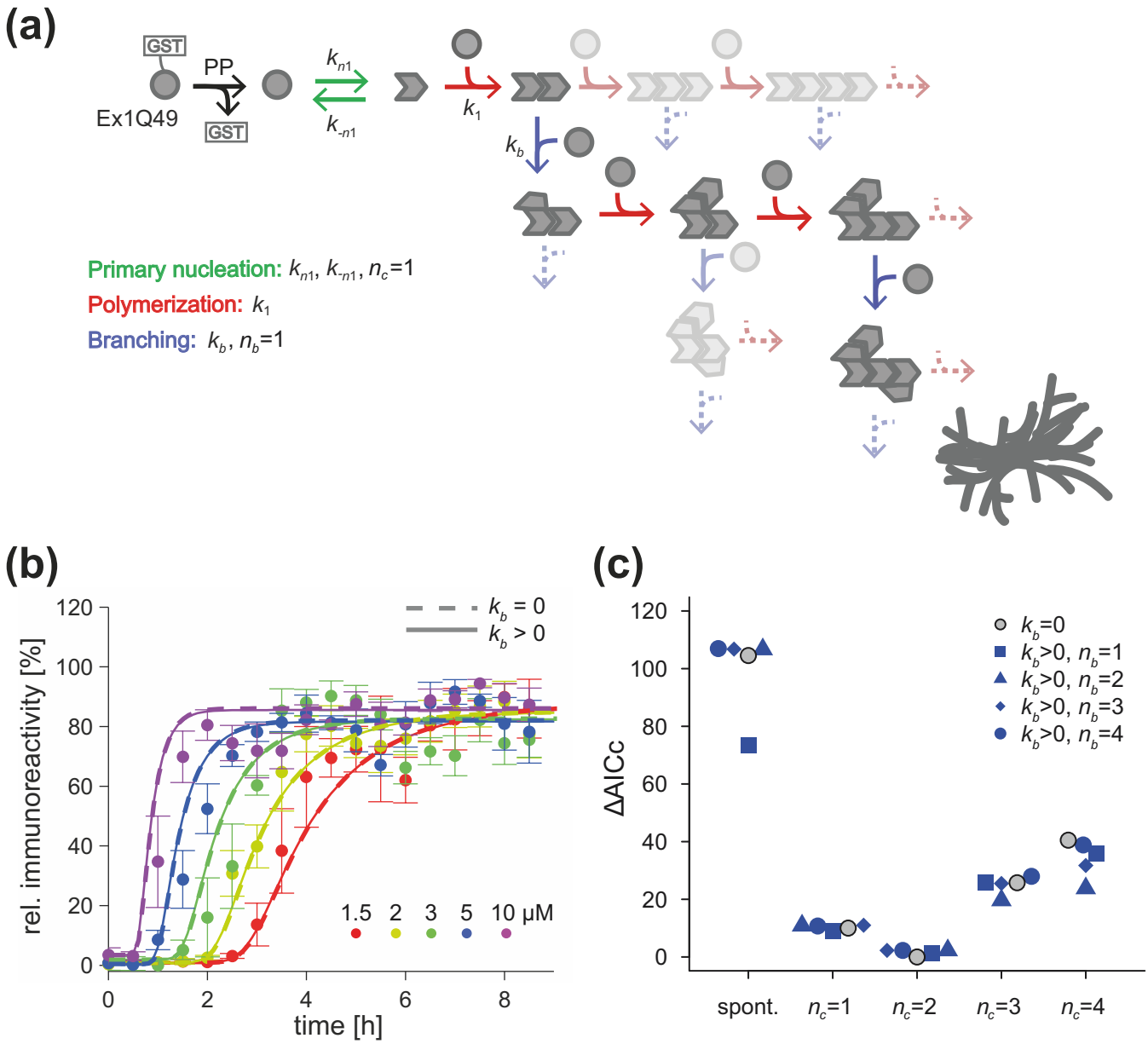
(a)
GST
PP
Ex1Q49
GST
$k_{n1}$
$k_{-n1}$
$k_1$
$k_b$
Primary nucleation: $k_{n1}$, $k_{-n1}$, $n_c$=1
Polymerization: $k_1$
Branching: $k_b$, $n_b$=1
(b)
rel. immunoreactivity [%]
time [h]
$k_b$ = 0
$k_b$ > 0
1.5 2 3 5 10 µM
(c)
ΔAICc
$k_b$=0
$k_b$>0, $n_b$=1
$k_b$>0, $n_b$=2
$k_b$>0, $n_b$=3
$k_b$>0, $n_b$=4
spont.
$n_c$=1
$n_c$=2
$n_c$=3
$n_c$=4

**Figure 3**

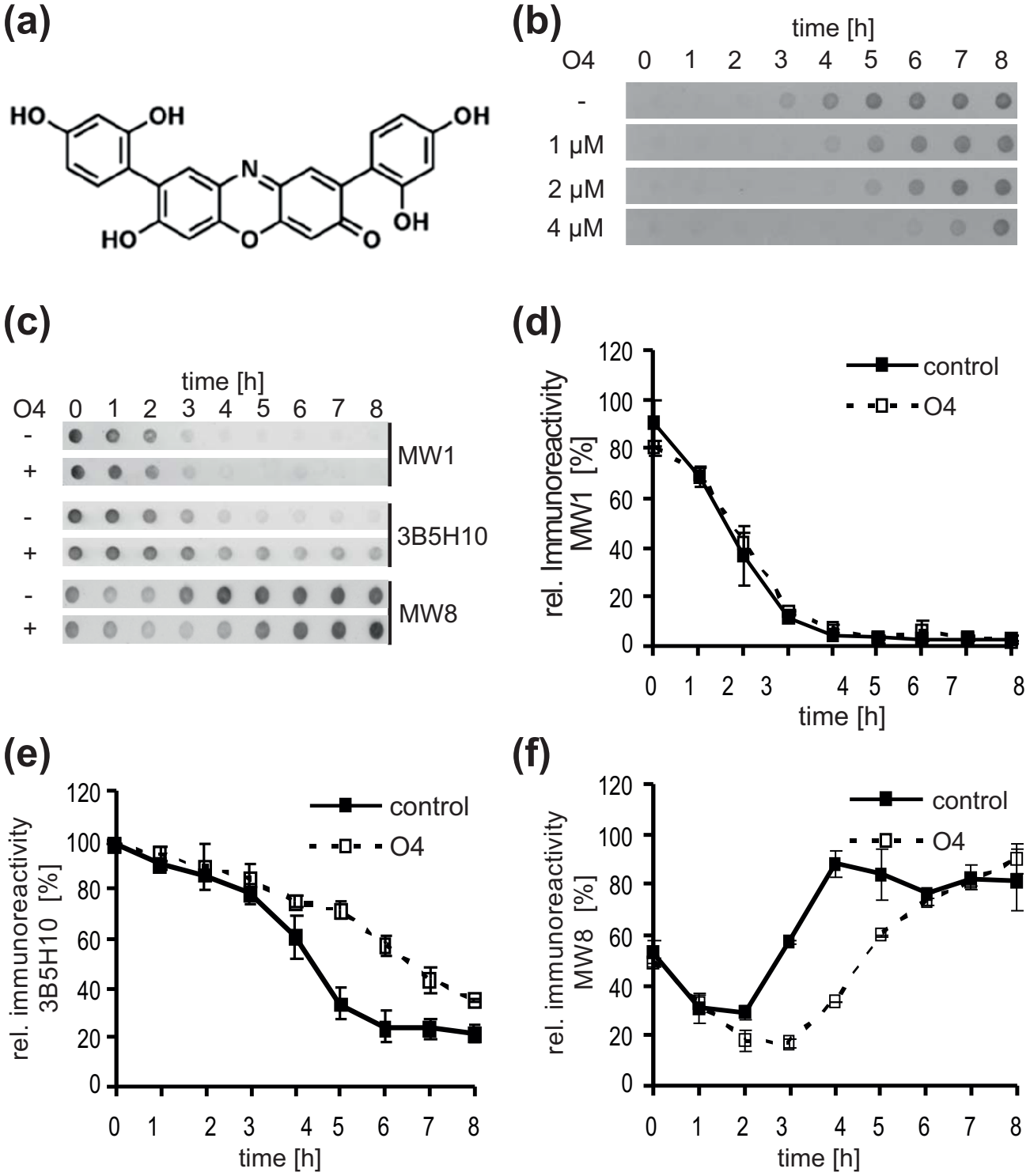
(a)
HO
OH
N
OH
HO
O
O
OH
(b)
time [h]
O4
0 1 2 3 4 5 6 7 8
-
1 µM
2 µM
4 µM
(c)
time [h]
O4
0 1 2 3 4 5 6 7 8
-
+
MW1
-
+
3B5H10
-
+
MW8
(d)
control
O4
rel. Immunoreactivity MW1 [%]
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
time [h]
(e)
control
O4
rel. immunoreactivity 3B5H10 [%]
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
time [h]
(f)
control
O4
rel. immunoreactivity MW8 [%]
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
time [h]

**Figure 4**

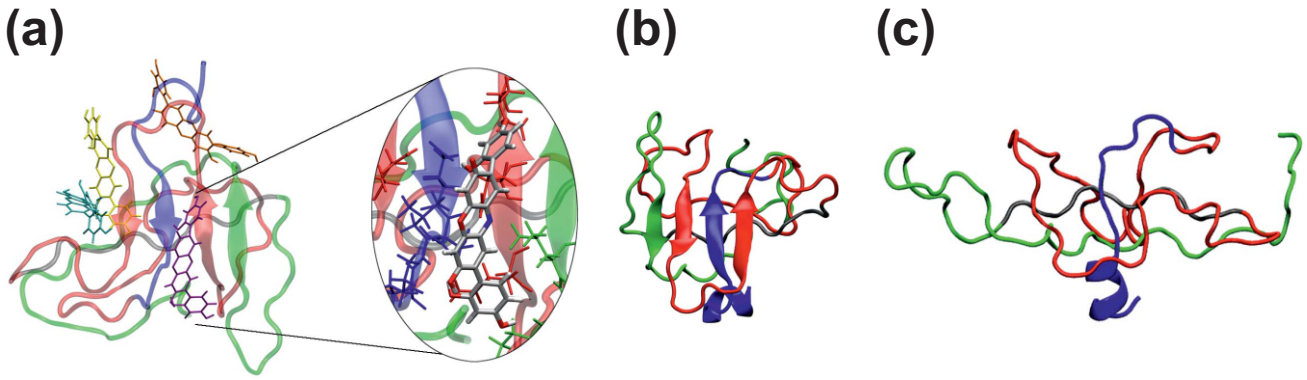


Figure 5

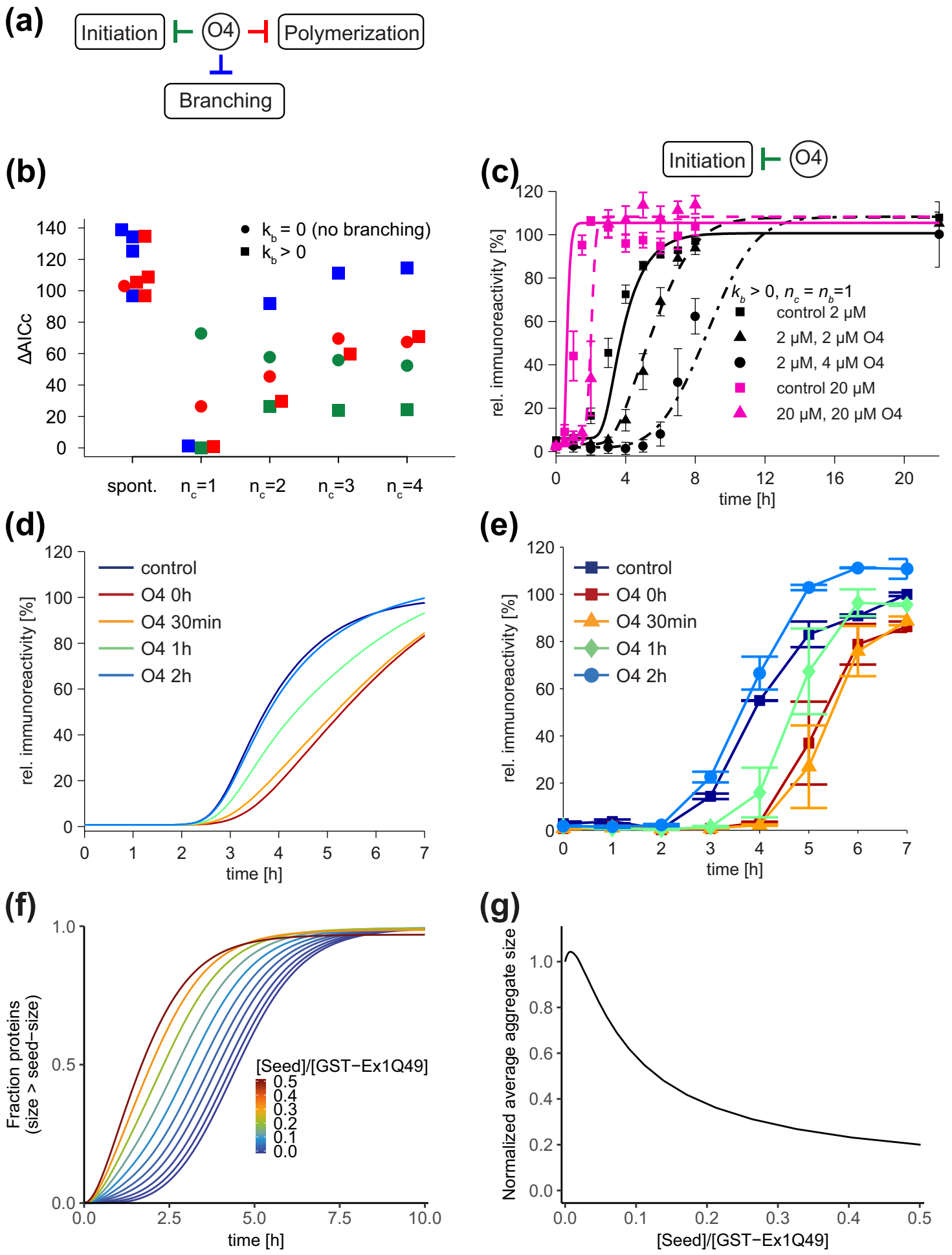
(a)
Initiation
O4
Polymerization
Branching
(b)
ΔAICc
$k_b$ = 0 (no branching)
$k_b$ > 0
spont.
$n_c$=1
$n_c$=2
$n_c$=3
$n_c$=4
(c)
Initiation
O4
rel. immunoreactivity [%]
time [h]
$k_b$ > 0, $n_c$ = $n_b$=1
control 2 µM
2 µM, 2 µM O4
2 µM, 4 µM O4
control 20 µM
20 µM, 20 µM O4
(d)
rel. immunoreactivity [%]
time [h]
control
O4 0h
O4 30min
O4 1h
O4 2h
(e)
rel. immunoreactivity [%]
time [h]
control
O4 0h
O4 30min
O4 1h
O4 2h
(f)
Fraction proteins (size > seed−size)
time [h]
[Seed]/[GST−Ex1Q49]
(g)
Normalized average aggregate size
[Seed]/[GST−Ex1Q49]

**Figure 6**

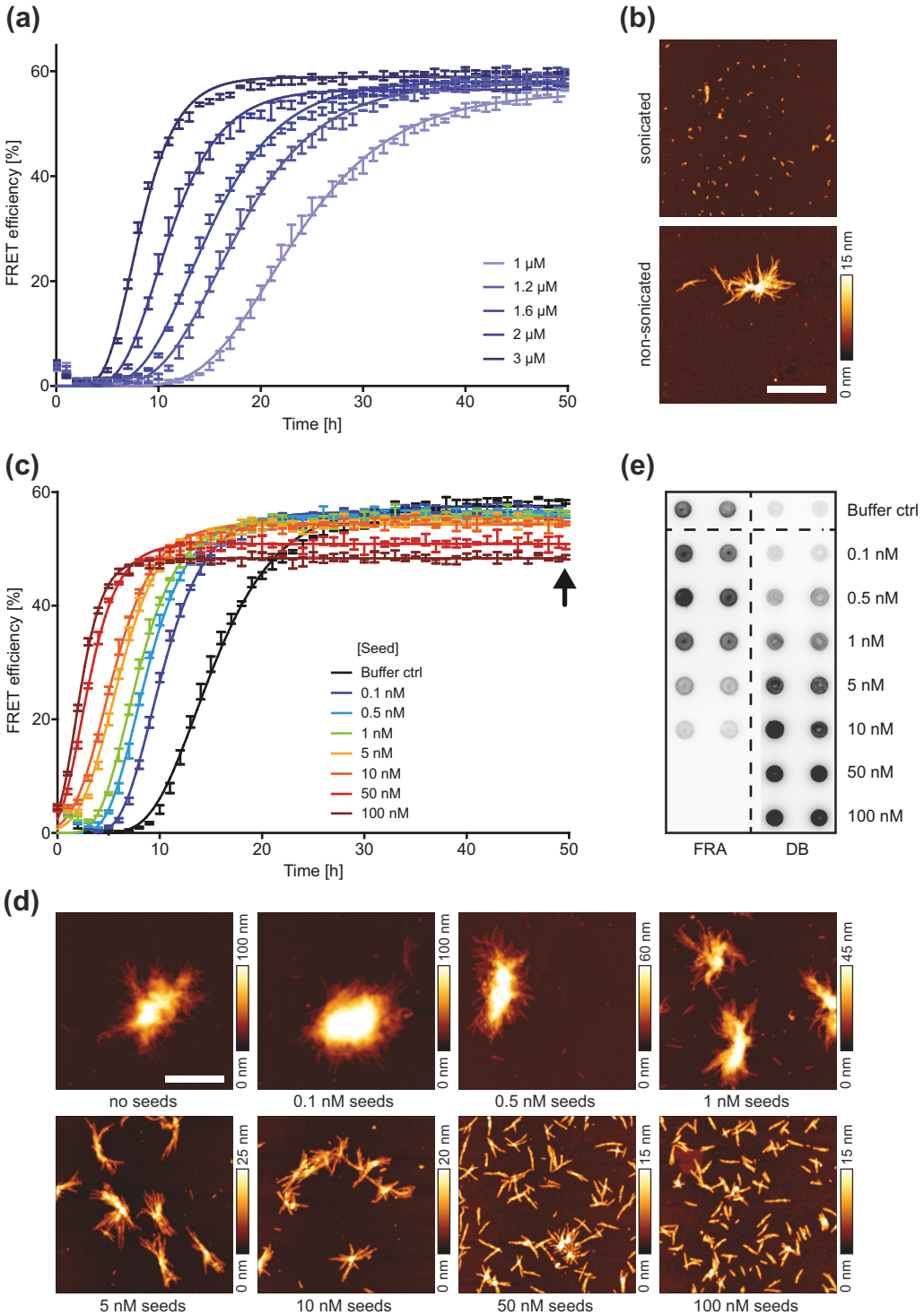
(a)
FRET efficiency [%]
Time [h]
1 µM
1.2 µM
1.6 µM
2 µM
3 µM
(b)
sonicated
non-sonicated
15 nm
0 nm
(c)
FRET efficiency [%]
Time [h]
[Seed]
Buffer ctrl
0.1 nM
0.5 nM
1 nM
5 nM
10 nM
50 nM
100 nM
(e)
Buffer ctrl
0.1 nM
0.5 nM
1 nM
5 nM
10 nM
50 nM
100 nM
FRA
DB
(d)
no seeds
0.1 nM seeds
0.5 nM seeds
1 nM seeds
5 nM seeds
10 nM seeds
50 nM seeds
100 nM seeds

**Figure 7**