

# Amyloid-Like Fibril Formation by PolyQ Proteins: A Critical Balance between the PolyQ Length and the Constraints Imposed by the Host Protein

Natacha Scarafone<sup>1</sup>, Coralie Pain<sup>1</sup>, Anthony Fratamico<sup>1#a</sup>, Gilles Gaspard<sup>2#b</sup>, Nursel Yilmaz<sup>2#b</sup>, Patrice Filée<sup>2#b</sup>, Moreno Galleni<sup>2</sup>, André Matagne<sup>1</sup>, Mireille Dumoulin<sup>1\*</sup>

**1** Laboratory of Enzymology and Protein Folding, Centre for Protein Engineering, Institute of Chemistry, University of Liège, Liège, Belgium, **2** Biological Macromolecules, Centre for Protein Engineering, Institute of Chemistry, University of Liège, Liège, Belgium

## Abstract

Nine neurodegenerative disorders, called polyglutamine (polyQ) diseases, are characterized by the formation of intranuclear amyloid-like aggregates by nine proteins containing a polyQ tract above a threshold length. These insoluble aggregates and/or some of their soluble precursors are thought to play a role in the pathogenesis. The mechanism by which polyQ expansions trigger the aggregation of the relevant proteins remains, however, unclear. In this work, polyQ tracts of different lengths were inserted into a solvent-exposed loop of the  $\beta$ -lactamase BlaP and the effects of these insertions on the properties of BlaP were investigated by a range of biophysical techniques. The insertion of up to 79 glutamines does not modify the structure of BlaP; it does, however, significantly destabilize the enzyme. The extent of destabilization is largely independent of the polyQ length, allowing us to study independently the effects intrinsic to the polyQ length and those related to the structural integrity of BlaP on the aggregating properties of the chimeras. Only chimeras with 55Q and 79Q readily form amyloid-like fibrils; therefore, similarly to the proteins associated with diseases, there is a threshold number of glutamines above which the chimeras aggregate into amyloid-like fibrils. Most importantly, the chimera containing 79Q forms amyloid-like fibrils at the same rate whether BlaP is folded or not, whereas the 55Q chimera aggregates into amyloid-like fibrils only if BlaP is unfolded. The threshold value for amyloid-like fibril formation depends, therefore, on the structural integrity of the  $\beta$ -lactamase moiety and thus on the steric and/or conformational constraints applied to the polyQ tract. These constraints have, however, no significant effect on the propensity of the 79Q tract to trigger fibril formation. These results suggest that the influence of the protein context on the aggregating properties of polyQ disease-associated proteins could be negligible when the latter contain particularly long polyQ tracts.

**Citation:** Scarafone N, Pain C, Fratamico A, Gaspard G, Yilmaz N, et al. (2012) Amyloid-Like Fibril Formation by PolyQ Proteins: A Critical Balance between the PolyQ Length and the Constraints Imposed by the Host Protein. *PLoS ONE* 7(3): e31253. doi:10.1371/journal.pone.0031253

**Editor:** Annalisa Pastore, National Institute for Medical Research, Medical Research Council, London, United Kingdom

**Received:** August 9, 2011; **Accepted:** January 5, 2012; **Published:** March 9, 2012

**Copyright:** © 2012 Scarafone et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by grants from the Fonds de la Recherche Fondamentale et Collective (FRFC 2.4530.09 to AM), the Fonds de la Recherche Scientifique (1.C039.09 and MIS-F.4505.11 to MD), and the Belgian program of Interuniversity Attraction Poles administered by the Federal Office for Scientific Technical and Cultural Affairs (PAI number P6/19). N.S. and C.P. are recipients of a FRIA fellowship. M.D. is a Research Associate of the Fonds de la Recherche Scientifique (FRS-F.N.R.S., Belgium). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

**Competing Interests:** The authors have read the journal's policy and have the following conflicts: Current address of GG, NY and PF: ProGenesis S.A., Boulevard du Rectorat, 27b, Sart-Tilman, 4000 Liège, Belgium. These three persons are employed by the company. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials.

\* E-mail: mdumoulin@ulg.ac.be

#a Current address: Laboratory of Plant Biochemistry and Photobiology, University of Liège, Liège, Belgium

#b Current address: ProGenesis S.A., Sart-Tilman, Liège, Belgium

## Introduction

Polyglutamine (polyQ) diseases are neurodegenerative disorders caused by the expansion of unstable CAG trinucleotide repeats in the translated region of unrelated genes. These CAG repeats encode a polyglutamine stretch in the corresponding proteins [1]. At least nine polyQ-related disorders are known including Huntington's disease and several spinocerebellar ataxias [1]. The nine disease-associated proteins show no sequence or structural similarity apart from the expanded polyQ tract which is located at a different position in each protein. The polyQ tract appears, therefore, to be a critical determinant of polyQ diseases and several lines of evidence suggest that it confers a toxic function to the mutant proteins [2,3]. Although the nine diseases present

distinct pathological and molecular phenotypes, they share a number of common features, suggesting a common physiopathological mechanism [4]. Firstly, there is a threshold in the number of repeats above which polyQ proteins become pathogenic. The value of this threshold varies from one disease to another but generally resides between 35 and 45 glutamines [5]. Secondly, above the threshold value, the longer the repeat, the earlier the onset and more severe the disease; this phenomenon is known as the "anticipation phenomenon" [4]. Thirdly, polyQ expansion mediates the deposition of nuclear inclusion bodies that contain amyloid-like fibrils [6]. Although the mechanism of toxicity associated with pathological expansion of polyglutamine tracts remains unclear, a large body of evidence indicates that it is associated with protein misfolding and aggregation [7,8]. Thus,

the cytotoxicity of proteins containing an expanded polyQ tract has been attributed to: (i) the formation of inclusion bodies [9,10], (ii) the presence of misfolded protein monomers [11] and (iii) the transient formation of oligomers during the aggregation process [12,13,14].

Two major scenarios, which are not necessarily mutually exclusive, have been put forward to explain how expanded polyQ tracts promote the aggregation of proteins [5,15]. Firstly, it has been suggested that long polyQ repeats (>36Q) have a high intrinsic propensity to form “polar zippers” which consist of  $\beta$ -sheets stabilized by hydrogen bonds between both main-chain and side-chain amides [16]. The formation of such structures is thought to trigger aggregation into amyloid fibrils. Alternatively, expanded polyQ tracts have been suggested to destabilize the proteins and thereby facilitate the formation of partially unfolded species [17,18,19]. Such species generally expose at least part of their main chain and hydrophobic residues to solvent and are therefore prone to intermolecular interactions leading to fibril formation. Such a mechanism has been described for other proteins associated with amyloidosis, including transthyretin [20] and human lysozyme [21,22].

Although the aggregation of polyQ proteins critically depends on the expansion of the polyQ stretch above the pathological threshold, it is however becoming increasingly evident that regions outside the polyQ tract can modulate both the kinetics and the pathway of aggregation [15,23]. The observation that the aggregation threshold of polyQ peptides is significantly lower ( $\geq 15Q$ ) than that observed for the proteins associated with diseases is a clear indication of the effects of the surrounding sequences [16,24]. Another evidence was provided by Bhattacharyya and co-workers [23], who showed that the addition of ten prolines at the C-terminal end of polyQ peptides decreased both their rate of *in vitro* aggregation and the stability of the resulting aggregates. The influence of the flanking sequences was also observed in the context of different proteins [25,26,27].

Various mechanisms have been proposed to explain the influence of the polyQ flanking regions on the aggregation properties of polyQ proteins [5]. Host protein domains could protect against aggregation by: (i) enhancing total protein solubility [27,28,29], (ii) sterically hindering polyQ intermolecular interactions and/or (iii) restricting the polyQ conformational changes required for fibril formation [23,30]. On the other hand, host protein domains adjacent to the polyQ tract can assist aggregation by providing additional aggregation-prone regions. There is indeed experimental evidence that at least ataxin-3 and the first exon of huntingtin (Htt exon 1) form fibrillar aggregates via a complex multidomain mechanism initiated by intermolecular interactions within non-polyQ regions of the protein. Although the polyQ region is not directly involved in this step, it does indirectly modulate the aggregation propensity of the non-polyQ region through a mechanism that is yet to be fully understood. It possibly involves the destabilization of the non-polyQ region and/or changes in its structure or dynamics [19,31]. At later stages, the expanded polyQ tract is then directly involved in the formation of the core of mature fibrils [17,19,32].

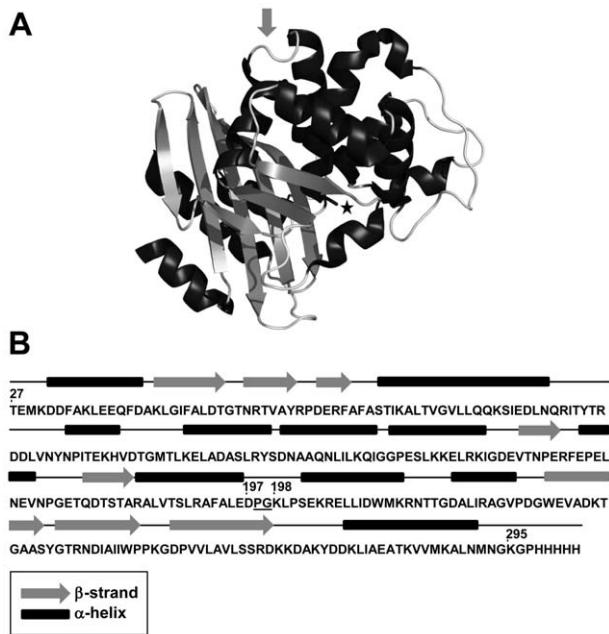
Taken together, these results highlight the existence of a complex interplay between the intrinsic properties of the polyQ tract to trigger aggregation and the modulating effects of the host protein. To gain a better insight into the general principles governing this complex phenomenon, it is necessary to investigate in detail which properties of the host protein influence the ability of polyQ stretches to mediate aggregation; this is therefore the subject of intensive research [17,19,23,30,32–37]. The difficulty in handling proteins involved in diseases, essentially due to their large

size and insoluble character, has prompted the design and use of model proteins [11,18,25,31,34,35,37–40]. The characterization of these proteins unambiguously points to the length and the location of the polyQ tract as important factors [18,34,35,37]. A number of questions are, however, not yet fully addressed; these include the role of the sequence, size, structure, stability, dynamics and topology of the host protein. There is therefore a clear need for the generation of new model polyQ proteins of various sizes, topologies and structures and the extensive characterization of their structural, thermodynamic, dynamic and aggregating properties.

In proteins associated with polyQ diseases, the polyQ tract is separated from the C- or N-terminal ends by at least 50 residues, except for the proteins associated with Huntington’s disease (HD, 17 residues) and spinocerebellar ataxia type 6 (SCA6, 30 residues) [34]. Many model proteins characterized so far, however, display a polyQ tract fused to their N- or C-terminus [35,37,39,41], probably because polypeptide chains with an inserted polyQ sequence are difficult to express [35,37]. It is therefore essential to investigate more model proteins containing an inserted polyQ stretch since the presence of flanking sequences at both extremities of the polyQ tract is likely to impose constraints that are more physiologically relevant than when the polyQ tract has a free extremity. With this aim in mind, we inserted polyQ sequences of various lengths into a solvent-exposed loop of a well-characterized globular protein, and we investigated the effects of these insertions on its structure, stability and aggregating properties.

The  $\beta$ -lactamase from *Bacillus licheniformis* 749/C (BlaP) was used as the protein scaffold. This 264-residue enzyme is organized into two structural domains, with the active site located at the interface of the two domains (Figure 1A) [42]. We chose this scaffold for several reasons: (i) detailed information concerning its thermodynamic and catalytic properties are available [43,44], providing a strong basis to investigate the effects of polyQ insertions, (ii) the solvent-exposed loop located between helices 8 and 9 (the position 197–198, Figure 1A) has been clearly shown to tolerate amino acid insertion [44,45], and (iii) chimeras with inserts of various lengths and structures can be readily produced in *Escherichia coli* ([44,45] and unpublished results).

We have created and characterized a series of chimeras with 23, 30, 55 and 79 glutamines inserted at position 197–198 (Figure 1A and B) using a range of biophysical techniques including fluorescence, circular dichroism (CD), transmission electron microscopy (TEM) and X-ray diffraction. We have found that the insertion of a polyQ tract consisting of up to 79 residues does not modify the structure of the enzyme, although it significantly reduces its conformational stability. The results of this study also show that the aggregating properties of the chimeras are similar to those of proteins associated with polyQ diseases. Thus, we observed a threshold number of inserted glutamines above which the chimeric  $\beta$ -lactamase aggregates into amyloid-like fibrils and that the kinetics of aggregation are faster with longer glutamine repeats. Most importantly, this work clearly demonstrates that the threshold number of glutamines above which the chimeric proteins aggregate into amyloid-like fibrils critically depends on the structural integrity of the  $\beta$ -lactamase moiety and thus on the constraints applied to the polyQ tract. It also suggests that the modulating effects of the protein context on the aggregating properties of proteins associated with polyQ diseases could be negligible when the latter contain particularly long polyQ tracts. Finally, our results indicate that BlaP is a promising scaffold for further investigation of the delicate balance between the propensity of polyQ tracts to trigger aggregation and the modulating effects of the host proteins.



**Figure 1. X-ray structure and topology of the host protein BlaP.**

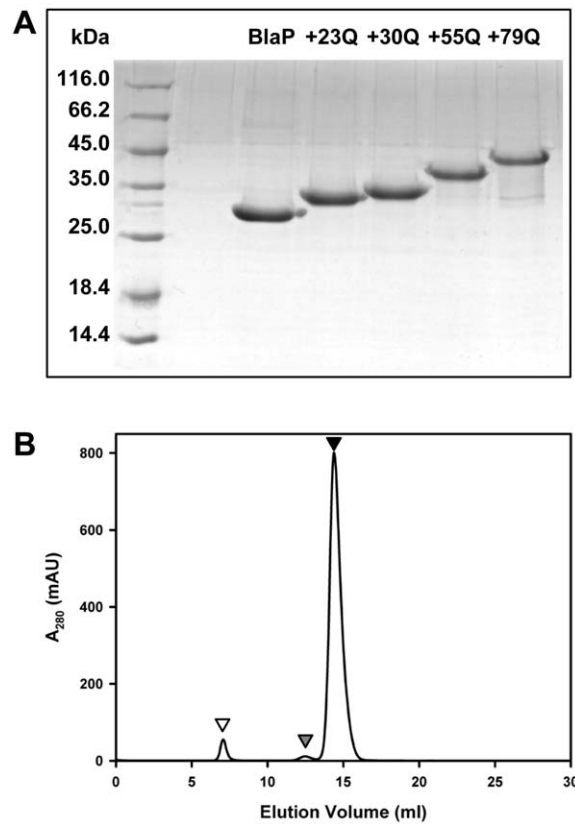
(A) The structure of BlaP was produced using PyMOL (DeLano Scientific LLC, South San Francisco, CA, USA) and the PDB ID is 4BLM [42]. The active site serine (Ser70) is indicated by a star and the insertion site is shown by an arrow. (B) Topology of BlaP. The position 197–198 of the insertion site refers to the numbering scheme of class A  $\beta$ -lactamases [65]. The underlined PG dipeptide between residues 197 and 198 corresponds to the SmaI restriction site inserted into the gene of BlaP for the cloning of CAG repeats.

doi:10.1371/journal.pone.0031253.g001

## Results and Discussion

### Production of the chimeras

Four expression vectors encoding for chimeric  $\beta$ -lactamases containing 23, 30, 55 and 79 glutamines [BlaP(Gln)<sub>23</sub>, BlaP(Gln)<sub>30</sub>, BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub>, respectively] were constructed as described in the Materials and Methods, by inserting CAG repeats into a SmaI restriction site incorporated into the BlaP gene between codons for residues 197 and 198. The introduction of this SmaI restriction site adds a dipeptide proline-glycine to the sequence of BlaP; the polyglutamine tract was inserted between these two amino acids (Figure 1B). The proteins were produced in *E. coli* and purified to homogeneity in one step using a Ni-PDC affinity column. This procedure led to about the same amount (i.e. 10–20 mg per liter of culture) of the different proteins, irrespective of the presence or absence of a polyQ tract. The purity of all samples was >95% as assessed by SDS-PAGE (Figure 2A). Although BlaP, BlaP(Gln)<sub>23</sub> and BlaP(Gln)<sub>30</sub> migrated at positions expected, according to the molecular mass markers, chimeras with either 55 or 79 glutamines showed lower mobility and thus, greater apparent molecular masses than expected. Such anomalous behaviour, which was also observed for chimeras composed of myoglobin and polyQ tracts [40], suggests that long polyQ sequences interfere with SDS binding. The integrity of the proteins was checked by mass spectrometry (ESI-QTOF-MS). This analysis revealed the presence of three distinct peaks in the spectrum of each protein. Although in all cases, the main peak (estimated at >70% of the population) corresponds to the theoretical mass deduced from the complete amino acid sequence, the two other peaks correspond to higher molecular mass species, i.e. +182 Da



**Figure 2. Purification of BlaP and the chimeras.** (A) Purified BlaP and polyQ-containing chimeras separated on a 15% (w/v) SDS-polyacrylamide gel and stained with coomassie blue. First lane on the left shows protein markers with molecular masses as indicated, whereas the other lanes show BlaP and the various chimeric proteins as indicated. Expected molecular masses are 30 368, 33 315, 34 212, 37 416 and 40 491 Da for BlaP and chimeras with 23, 30, 55 and 79 glutamines, respectively. (B) SEC analysis of BlaP(Gln)<sub>79</sub> at 120  $\mu$ M in PBS, pH 7.5; mAU, milli-absorbance units. Monomeric, dimeric and high molecular weight species are indicated by black, grey and white arrows, respectively. The high molecular weight oligomeric species are eluted in the void volume of the column.

doi:10.1371/journal.pone.0031253.g002

(10–20% of the population) and +200 Da (<10% of the population). N-terminal sequencing indicated that the main peak corresponds to the enzyme with a well-processed N-terminal sequence (T-E-M-K) (Figure 1B), whereas the peak at +200 Da corresponds to the full-length protein plus the last two residues of the signal peptide (Q-A-T-E-M-K). The same N-terminal misprocessing was also observed for another BlaP chimera [44]. The nature of the +182 Da adduct is not clear. Since the proportion of each molecular species is similar between various protein preparations, it is assumed that they do not interfere with the main conclusions of this work.

Size-exclusion chromatography (SEC) analysis of protein preparations at high concentrations (50  $\mu$ M and 120  $\mu$ M) revealed that all proteins, particularly BlaP, form dimers although in small amount (at 120  $\mu$ M, 5.6% for BlaP and <1.5% for the four chimeras, Table 1). Minute quantities of high molecular weight species were also observed in BlaP(Gln)<sub>79</sub> solution (3.5% at 120  $\mu$ M and 2.5% at 50  $\mu$ M) (Table 1 and Figure 2B). Based on their volume of elution and assuming that these oligomeric species are spherical, their apparent molecular mass can be estimated to be higher than 600 kDa, and thus formed by 15 or more

**Table 1.** Percentages of the different species observed by SEC.

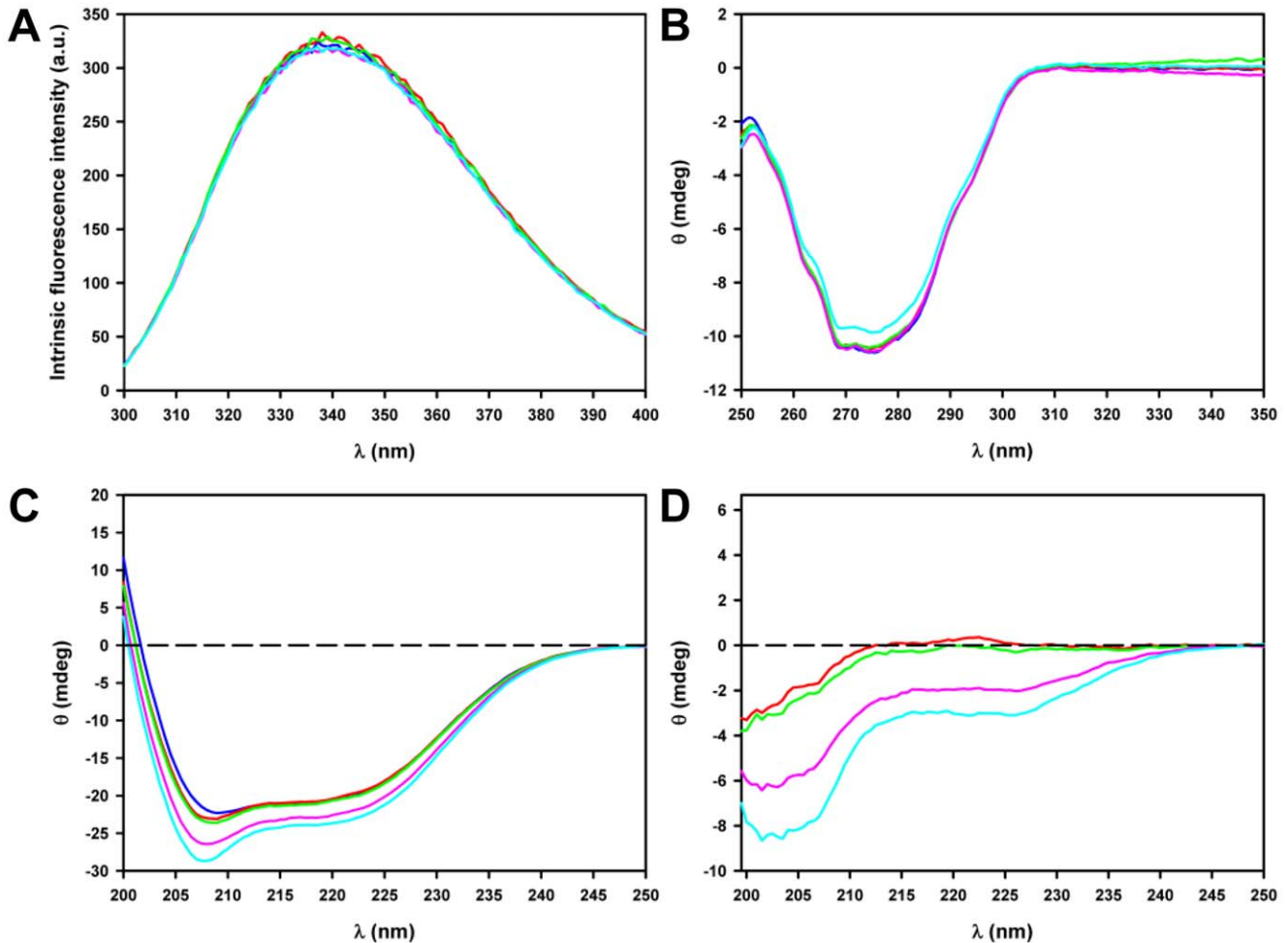
	[protein] ( $\mu\text{M}$ )	M (%)	D (%)	O (%)
BlaP	50	95.1	4.9	0
	120	94.4	5.6	0
BlaP(Gln) <sub>23</sub>	50	99.4	0.6	0
	120	99.3	0.7	0
BlaP(Gln) <sub>30</sub>	50	99.6	0.4	0
	120	99.5	0.5	0
BlaP(Gln) <sub>55</sub>	50	98.9	1.1	0
	120	98.6	1.4	0
BlaP(Gln) <sub>79</sub>	10	97.3	0.9	1.8
	50	96.4	1.1	2.5
	120	95.3	1.2	3.5

This analysis was carried out using a Superdex 200 GL 10/300 column using PBS buffer, pH 7.5. **M**, **D**, **O** stand for monomeric, dimeric, and high molecular weight oligomeric species, respectively.  
doi:10.1371/journal.pone.0031253.t001

monomers. Decreasing the concentration of BlaP(Gln)<sub>79</sub> samples from 50  $\mu\text{M}$  to 10  $\mu\text{M}$  reduces the percentage of these oligomeric species to less than 2% (Table 1). Following the observation that in all cases more than 95% of the protein population is monomeric at concentrations as high as 50  $\mu\text{M}$ , we assumed that the presence of dimers and oligomers did not interfere with the determination of the catalytic and thermodynamic parameters which were carried out with protein concentrations  $\leq 4.6 \mu\text{M}$ .

### The polyQ tract adopts a disordered structure and does not perturb the overall structure of BlaP

The effects of the insertion of polyQ tracts of 23, 30, 55 and 79 residues on the structure of BlaP were investigated by intrinsic fluorescence and CD measurements. Fluorescence and near-UV CD data (Figure 3A and B) suggest that none of the polyQ insertions into the loop between the  $\alpha$ -helices 8 and 9 of BlaP significantly affects the tertiary structure of the protein. This is further supported by the observation that insertion of polyQ sequences of various lengths does not significantly modify the catalytic properties of the  $\beta$ -lactamase (Table 2). In contrast, the far-UV CD spectra of the chimeras show marked differences with that of the wild-type protein (Figure 3C). Subtraction of the BlaP



**Figure 3. PolyQ insertions have no effect on the structure of BlaP.** (A) Intrinsic fluorescence, (B) near-UV CD and (C) far-UV CD spectra of BlaP (blue) and the chimeras with 23 (red), 30 (green), 55 (pink) and 79 (cyan) glutamines. a.u., arbitrary units. (D) Difference spectra obtained by subtraction of the far-UV CD spectrum of BlaP from that of each chimeric protein. Spectra were recorded at 25°C in PBS, pH 7.5, using protein concentrations of 4.6  $\mu\text{M}$  (fluorescence and far-UV CD) and 20  $\mu\text{M}$  (near-UV CD).  
doi:10.1371/journal.pone.0031253.g003

**Table 2.** Kinetic parameter values for the wild-type and chimeric  $\beta$ -lactamases.

	$k_{cat}$ ( $s^{-1}$ )	$K_m$ ( $\mu M$ )	$10^{-6} \times k_{cat}/K_m$ ( $M^{-1} \cdot s^{-1}$ )
<b>BlaP</b>	90 $\pm$ 10	25 $\pm$ 3	3.6 $\pm$ 0.4
<b>BlaP(Gln)<sub>23</sub></b>	120 $\pm$ 10	32 $\pm$ 4	3.7 $\pm$ 0.3
<b>BlaP(Gln)<sub>30</sub></b>	87 $\pm$ 7	27 $\pm$ 3	3.1 $\pm$ 0.2
<b>BlaP(Gln)<sub>55</sub></b>	107 $\pm$ 10	31 $\pm$ 3	3.5 $\pm$ 0.5
<b>BlaP(Gln)<sub>79</sub></b>	77 $\pm$ 2	27 $\pm$ 2	2.9 $\pm$ 0.1

Cephalothin was used as the substrate. Errors are given as standard deviations. doi:10.1371/journal.pone.0031253.t002

spectrum from the spectra of the chimeras gives an indication of the structure adopted by the polyQ insert. Interestingly, the difference spectra display a negative peak with a minimum at *ca.* 202–203 nm, the amplitude of which increases with the number of glutamines (Figure 3D). This result suggests that the polyQ tract: (i) does not significantly modify the secondary structure of BlaP and (ii) adopts a disordered structure at the surface of the protein irrespective of the number of glutamines.

A similar absence of structural modification of the non-polyQ regions was observed when a polyQ tract is fused to the B domain of the protein A from *Staphylococcus aureus* (SpA) or apomyoglobin [35,37]; however, radically different results were reported for inserted polyQ tracts. The most characterized chimeric proteins with inserted polyQ sequences were created from three all- $\alpha$  proteins: myoglobin [18], SpA [35] and apomyoglobin [37]. In these model proteins, the polyQ tract was inserted into a loop between  $\alpha$ -helices. The insertion of a polyQ tract into either myoglobin [18] or SpA [35] induces changes in the tertiary structure of the host proteins, with no significant effect on their secondary structure. The insertion of at least 16Q in apomyoglobin leads, however, to significant changes in secondary structure, besides a polyQ-length-independent loss of tertiary structure [37]. This partial loss of structure was also observed when serine-glycine repeats were inserted at the same location, suggesting that it was not specific to polyQ insertion. As a likely consequence of these structural changes, the production yield for at least two of these model proteins was significantly lower than that of their respective wild-type counterparts, and this was shown to be directly related to the length of the polyQ tract inserted [35,37]. To date, the model protein with the longest inserted polyQ sequence reported in the literature consists of a chimeric myoglobin containing 50 glutamines only [18]. In contrast, the production yield of our BlaP chimeras remains unchanged (*i.e.* 10–20 mg per liter of culture), even when a 79-glutamine sequence is inserted at position 197–198. This makes BlaP an ideal scaffold to investigate the effects of long inserted polyQ sequences.

Our CD data suggest that the polyQ tracts of 23–79Q, inserted into the solvent-exposed loop of BlaP, adopt a disordered structure. This observation is in very good agreement with the results obtained with other model proteins containing variable-length polyQ tracts, *e.g.* (i) polyQ sequences fused either to SpA or GST [35,39,41] and (ii) the exon 1 of huntingtin fused to thioredoxin [38]. In contrast, data obtained with myoglobin and apomyoglobin indicated that inserted polyQ sequences of sufficient length (*ca.* 16–28Q) could form  $\beta$ -structures [18,37]. These findings suggest that the conformation adopted by a polyQ stretch is strongly context-dependent. The observation that a 79-residue long polyQ sequence is disordered when inserted into BlaP, suggests that it does not adopt any stable, aggregation-prone,

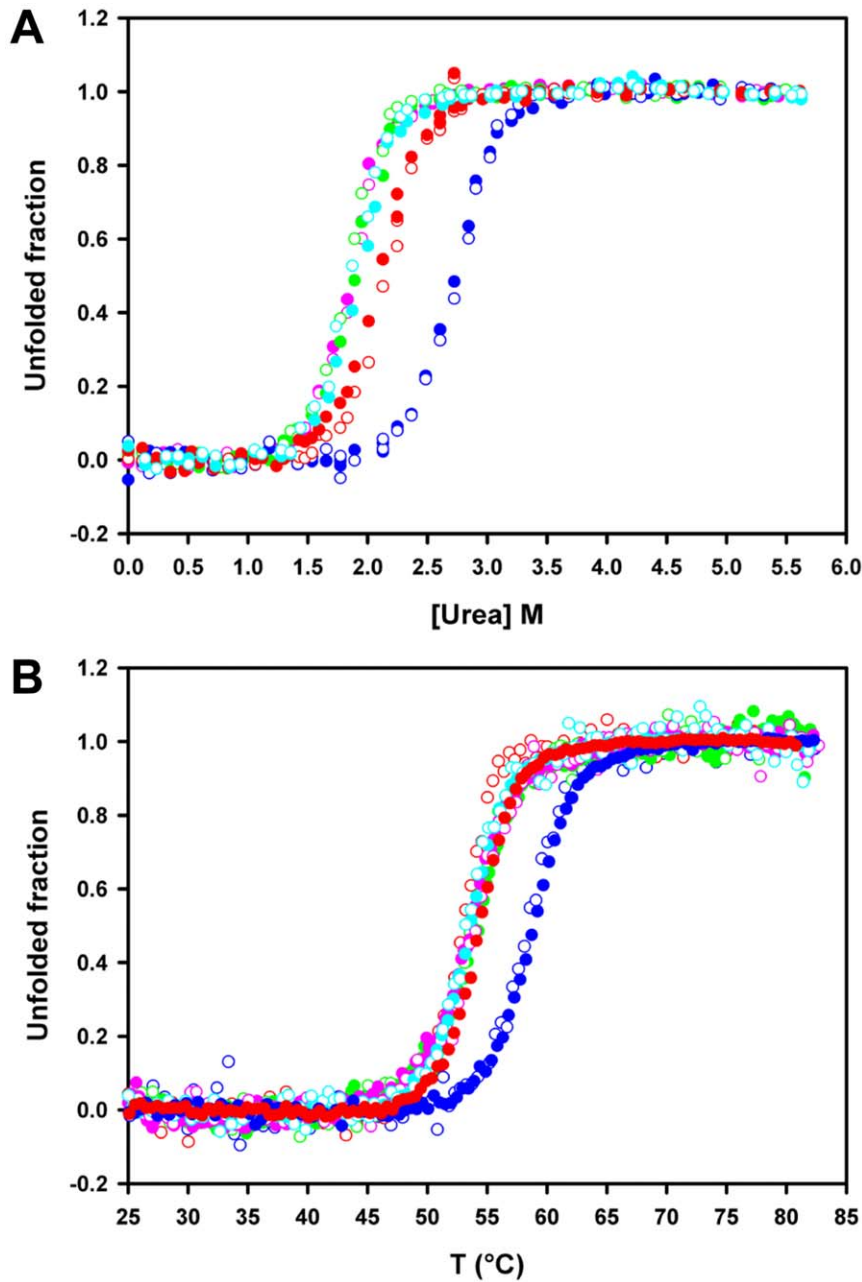
$\beta$ -sheet structure in solution prior to aggregation into amyloid-like fibrils. Because of the low sensitivity of the structural techniques used in this work, we cannot exclude, however, the possibility that the polyQ tract might rarely and/or transiently access more organized structures.

### The polyQ tract destabilizes BlaP and the extent of destabilization is largely independent of the polyQ length

The effects of polyQ insertions of different lengths (23, 30, 55 and 79Q) on the stability of BlaP were determined by urea-induced unfolding experiments. For each protein, unfolding was found to be fully thermodynamically reversible (data not shown) and single transitions were observed (Figure 4A). The data obtained by intrinsic fluorescence and far-UV CD coincide, indicating that the five proteins unfold according to a simple two-state mechanism, with no intermediate species significantly populated between the native and unfolded states. Based on this model, the characteristic thermodynamic parameters were determined and are given in Table 3. These data show that all four chimeras are destabilized with respect to the wild-type protein. Remarkably, the extent of destabilization is similar, *i.e.* 7.6–8.8  $\text{kJ} \cdot \text{mol}^{-1}$ , for all four chimeric proteins and thus is largely independent of the length of the polyQ stretch. Thermal unfolding of BlaP and its chimeras (Figure 4B) was also monitored by both intrinsic fluorescence and far-UV CD measurements. Although this process was not reversible (data not shown), cooperative transitions were observed and apparent temperatures of mid-transition ( $T_m^{\text{app}}$ ) were determined (Table 3). In good agreement with urea-unfolding data, all chimeras are less stable than the wild-type protein ( $\Delta T_m^{\text{app}} = 4.5\text{--}5.3^\circ\text{C}$ ) and the extent of destabilization is independent of the number of glutamines.

The fact that the extent of BlaP destabilization is independent of the number of glutamines contrasts with observations on other model systems reporting: (i) no destabilization at all, *e.g.* the protein moiety SpA when fused to a polyQ tract up to 52 residues in length [35], (ii) selective destabilization, *e.g.* the cellular retinoic acid binding protein I (CRABP I) moiety when fused to Htt exon 1 containing polyQ stretches longer than a threshold size [31], (iii) cumulative destabilization, *e.g.* chymotrypsin inhibitor 2 (CI2) [46], myoglobin [18] and SpA [35] containing inserted polyQ tracts, where the extent of destabilization increases with increasing polyQ length, independently of a threshold. Finally, a polyQ-length independent destabilization was also described for apomyoglobin containing an inserted polyQ tract [37] but this effect does not seem specific to the glutamines, since a similar destabilization was observed when serine-glycine repeats were inserted at the same location. In the case of BlaP, insertion of the 73-residue long chitin-binding domain of human macrophage chitotriosidase [44] destabilizes the host protein significantly less than the polyQ tracts (*i.e.* 3.2  $\text{kJ} \cdot \text{mol}^{-1}$  *vs.* 7.6–8.8  $\text{kJ} \cdot \text{mol}^{-1}$ ). This suggests that the extent of destabilization of BlaP depends more on the nature of the polypeptide inserted than its length, and that the insertion of the disordered polyQ stretch is more destabilizing than that of the folded chitin-binding domain of human macrophage chitotriosidase. Taken together, these data show that the effects of polyQ insertion on the stability of the host protein vary greatly depending on the properties of the latter.

The observation that the chimeric  $\beta$ -lactamases with 30, 55 and 79 glutamines are destabilized to the same extent relative to wild-type BlaP (Table 3) gave us the unique opportunity to investigate independently the influence of (i) the length of the polyQ sequence and (ii) the conformational state of the  $\beta$ -lactamase moiety on the aggregating properties of the chimeras. Accordingly, the aggregating



**Figure 4. PolyQ insertions destabilize BlaP and the extent of destabilization is largely independent of the polyQ length.** (A) Normalized urea-induced unfolding transitions at pH 7.5 and 25°C and (B) normalized heat-induced unfolding transitions at pH 7.5, monitored by the changes in fluorescence intensity at 323 nm (filled circles) and in ellipticity at 222 nm (open circles), using a protein concentration of 4.6  $\mu$ M. BlaP (blue), BlaP(Gln)<sub>23</sub> (red), BlaP(Gln)<sub>30</sub> (green), BlaP(Gln)<sub>55</sub> (pink) and BlaP(Gln)<sub>79</sub> (cyan). Non-normalized data were analysed on the basis of a two-state model and the values of the thermodynamic parameters obtained are reported in Table 3. doi:10.1371/journal.pone.0031253.g004

gating properties of BlaP and the different chimeras were investigated under both native and denaturing conditions.

The threshold length of the polyQ tract above which chimeras aggregate into amyloid-like fibrils depends on the structural integrity of BlaP

**Aggregation under denaturing conditions.** The aggregating properties of the proteins were first investigated in the presence of 1.85 M urea (in PBS, pH 7.5), at 25°C. Data in Figure 4A and Table 3 indicate that under these conditions, all BlaP molecules are native, whereas *ca.* 18% and 50% of the molecules of, respectively,

BlaP(Gln)<sub>23</sub> and the three other chimeric enzymes, are unfolded. The kinetics of aggregation were monitored by measuring the amount of protein that remained soluble after different incubation times. BlaP, BlaP(Gln)<sub>23</sub> and BlaP(Gln)<sub>30</sub> display little, if any, tendency to aggregate, even after *ca.* 720 hours of incubation, whereas both BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> readily aggregate (Figure 5A). Aggregation of BlaP(Gln)<sub>55</sub> is characterized by a lag phase (*ca.* 50 hours) while very fast aggregation with no discernable lag phase is observed with BlaP(Gln)<sub>79</sub>. The lag phase is consistent with the nucleation-polymerization mechanism that has been proposed for amyloid fibril formation [47] and corresponds to the

**Table 3.** Thermodynamic parameter values of unfolding of wild-type and chimeric BlaP.

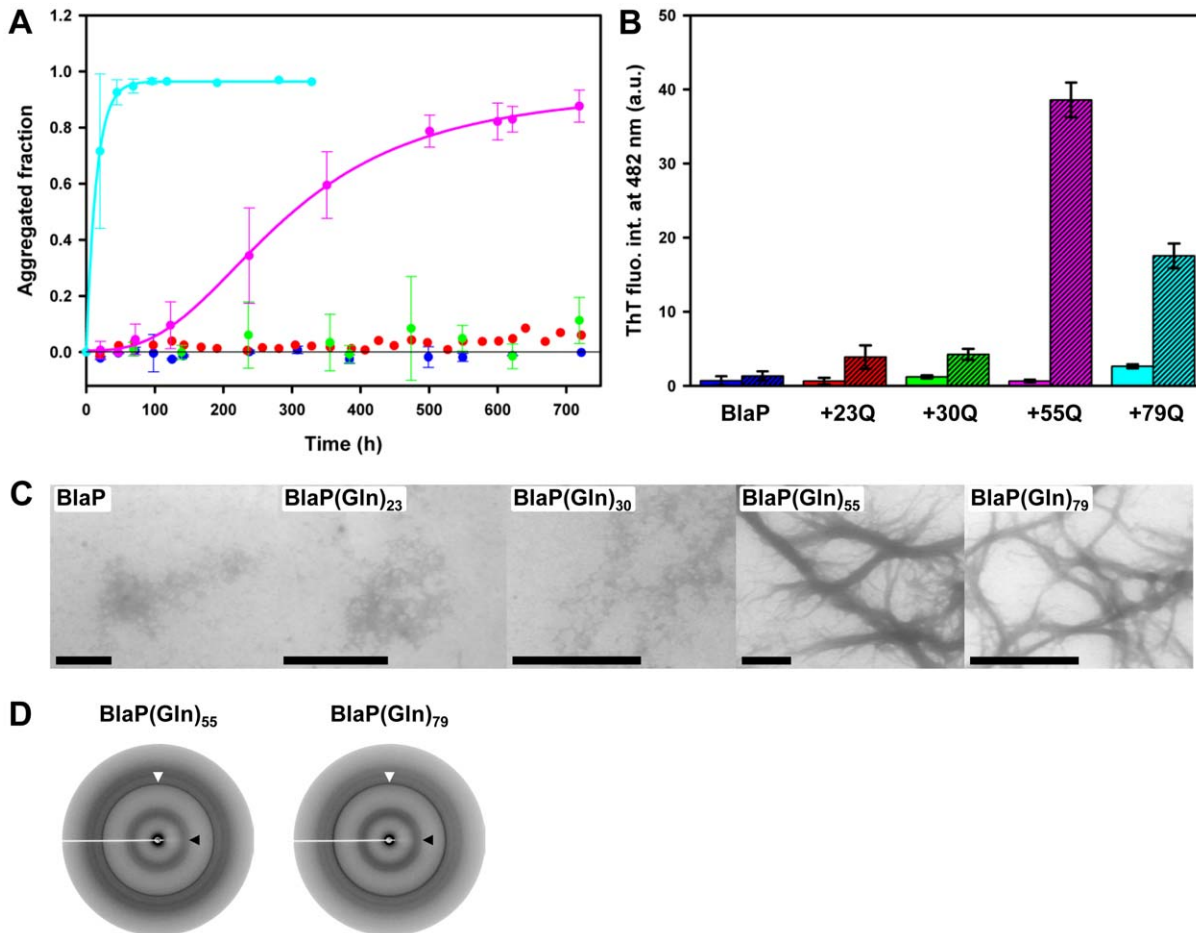
	$\Delta G^{\circ}_{\text{NU}}(\text{H}_2\text{O})$ ( $\text{kJ}\cdot\text{mol}^{-1}$ )	$-m_{\text{NU}}$ ( $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$ )	$C_m$ (M)	$T_m^{\text{app}}$ ( $^{\circ}\text{C}$ )
<b>BlaP</b>	$37.0\pm 1.8$	$13.5\pm 0.6$	$2.74\pm 0.23$	$58.5\pm 0.4$
<b>BlaP(Gln)<sub>23</sub></b>	$29.3\pm 2.4$	$13.7\pm 0.9$	$2.13\pm 0.19$	$53.9\pm 0.7$
<b>BlaP(Gln)<sub>30</sub></b>	$29.4\pm 1.2$	$15.9\pm 0.7$	$1.85\pm 0.15$	$53.7\pm 0.3$
<b>BlaP(Gln)<sub>55</sub></b>	$29.0\pm 0.6$	$15.6\pm 0.6$	$1.86\pm 0.15$	$53.5\pm 0.6$
<b>BlaP(Gln)<sub>79</sub></b>	$28.2\pm 1.3$	$14.9\pm 0.6$	$1.89\pm 0.17$	$53.2\pm 0.3$

The values were deduced from the analysis of unfolding transitions monitored by intrinsic fluorescence and far-UV CD measurements. The urea unfolding experiments were carried out at 25°C.  $\Delta G^{\circ}_{\text{NU}}(\text{H}_2\text{O})$ ,  $m_{\text{NU}}$ ,  $C_m$  and  $T_m^{\text{app}}$  values result from the mean of the values obtained by far-UV CD and intrinsic fluorescence; note that for thermal unfolding, three measurements were carried out both by fluorescence and CD. Thermal unfolding was not fully reversible and only the apparent mid-unfolding temperature ( $T_m^{\text{app}}$ ) was determined.

doi:10.1371/journal.pone.0031253.t003

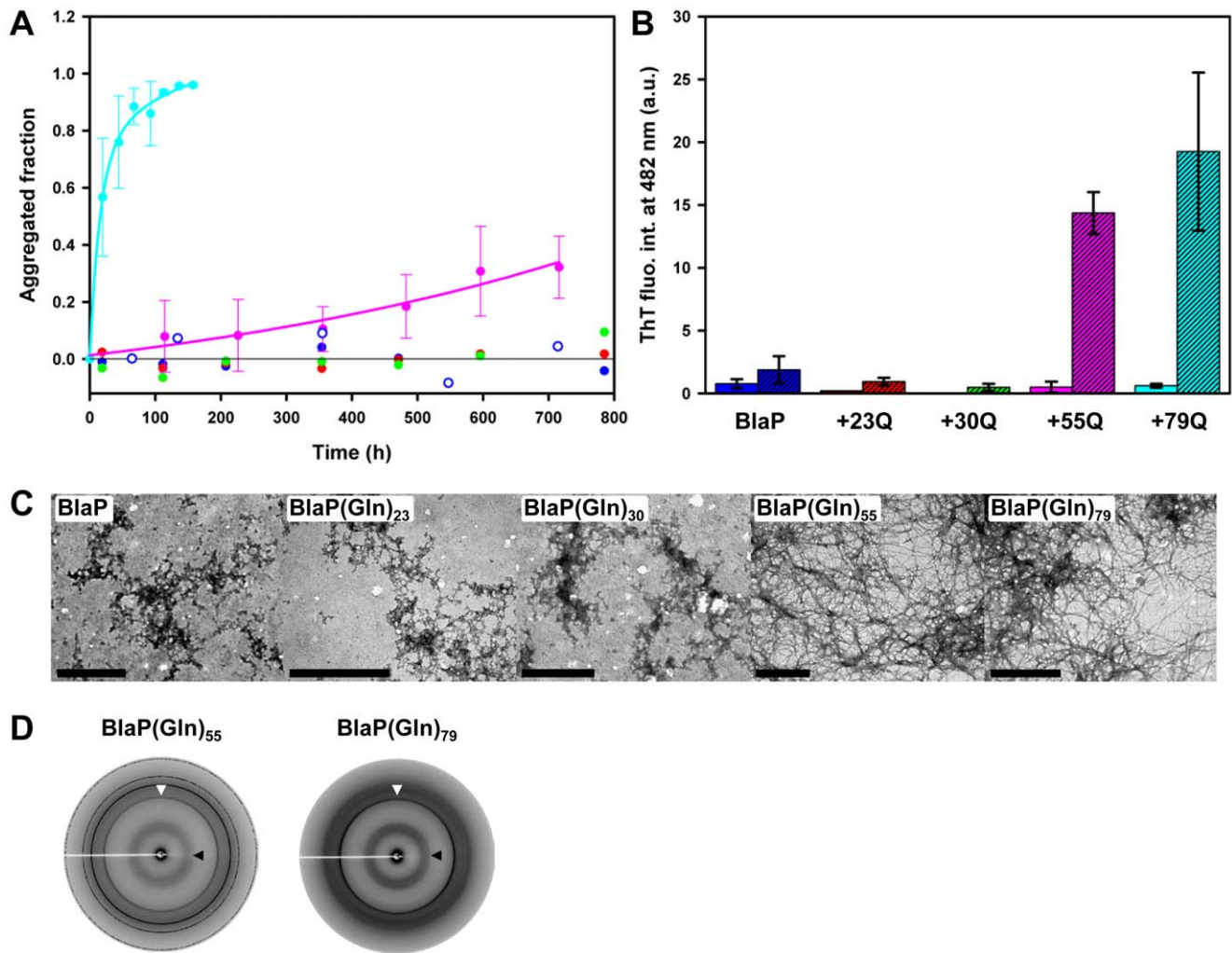
nucleation phase. After *ca.* 720 and *ca.* 330 hours of incubation, fibrillar aggregates (Figure 5C) are clearly visible by TEM for BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub>, respectively. In contrast, even after

720 hours of incubation, only minute amounts of amorphous aggregates, but no fibrils, were observed with BlaP, BlaP(Gln)<sub>23</sub> and BlaP(Gln)<sub>30</sub> (Figure 5C). The aggregates formed by BlaP(Gln)<sub>55</sub> and



**Figure 5. Only BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> form amyloid-like fibrils when incubated in the presence of 1.85 M urea.** (A) Aggregation kinetics of 110  $\mu\text{M}$  BlaP (blue), BlaP(Gln)<sub>23</sub> (red), BlaP(Gln)<sub>30</sub> (green), BlaP(Gln)<sub>55</sub> (pink) and BlaP(Gln)<sub>79</sub> (cyan) at 25°C in the presence of 1.85 M urea in PBS, pH 7.5, followed by measuring the concentration of protein remaining soluble. Time-points shown with an error bar are the average of three independent time-courses for BlaP(Gln)<sub>30</sub>, BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> and two independent time-courses for BlaP. Error bars show the standard deviations. Only one time-course was carried out with BlaP(Gln)<sub>23</sub>. (B) ThT fluorescence intensities at 482 nm in the presence of BlaP and chimeras samples at T<sub>0</sub> (solid bars) and T<sub>f</sub> (dashed bars). T<sub>0</sub> and T<sub>f</sub> correspond to the initial and final points of one time-course for each protein. Data are the average of three measurements and error bars represent the standard deviations. a.u., arbitrary units. (C) TEM images of the protein samples at T<sub>f</sub>. The scale bar is 1  $\mu\text{m}$ . (D) X-ray fibre diffraction patterns from BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> fibrils. White and black arrows indicate meridional and equatorial reflections at 4.7 Å and *ca.* 9.5 Å, respectively.

doi:10.1371/journal.pone.0031253.g005



**Figure 6. The unfolding of the  $\beta$ -lactamase moiety is not the critical factor that triggers the aggregation process.** (A) Aggregation kinetics of 110  $\mu$ M BlaP (blue), BlaP(Gln)<sub>23</sub> (red), BlaP(Gln)<sub>30</sub> (green), BlaP(Gln)<sub>55</sub> (pink) and BlaP(Gln)<sub>79</sub> (cyan) at 25°C in the presence of 3.5 M urea in PBS, pH 7.5, followed by measuring the concentration of protein remaining soluble. Time-points shown with an error bar are the average of three independent time-courses for BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub>. Error bars show the standard deviations. For BlaP, two independent experiments were carried out (filled and open blue circles); however, since the times at which samples were taken differ from one time-course to the other, the data could not be averaged. For BlaP(Gln)<sub>23</sub> and BlaP(Gln)<sub>30</sub>, only one time-course was carried out. (B) ThT fluorescence intensities at 482 nm in the presence of BlaP and chimeras samples at T<sub>0</sub> (solid bars) and T<sub>F</sub> (dashed bars). T<sub>0</sub> and T<sub>F</sub> correspond to the initial and final points of one time-course for each protein. Data are the average of three measurements and error bars represent the standard deviations. a.u., arbitrary units. (C) TEM images of the protein samples at T<sub>F</sub>. The scale bar is 1  $\mu$ m. (D) X-ray fibre diffraction patterns from BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> fibrils. White and black arrows indicate meridional and equatorial reflections at 4.7 Å and *ca.* 9.5 Å, respectively. doi:10.1371/journal.pone.0031253.g006

BlaP(Gln)<sub>79</sub> after, respectively, 720 and 330 hours of incubation significantly bind ThT (Figure 5B), suggesting that the fibrils observed are amyloid-like. The degree of ThT binding observed for BlaP(Gln)<sub>79</sub> is, however, significantly lower than that observed for BlaP(Gln)<sub>55</sub>, a surprising observation considering that BlaP(Gln)<sub>79</sub> aggregates to a higher extent than its 55-glutamine homolog (Figure 5A). The significantly lower ThT binding to BlaP(Gln)<sub>79</sub> fibrils may be due to an alternative overall packing of the aggregates. This is consistent with the observation that aggregates formed by BlaP(Gln)<sub>79</sub> were clearly larger and more difficult to resuspend in the ThT solution. For BlaP(Gln)<sub>79</sub>, the ThT fluorescence measurements were performed with samples incubated for *ca.* 330 hours and since the protein is fully aggregated after 100 hours, the aggregates could mature (e.g. laterally associate) during the last 200 hours of incubation. In the case of BlaP(Gln)<sub>55</sub>, the aggregation is much slower and there is therefore less time for fibril maturation until the

final point of the time-course (i.e. when samples were taken for ThT fluorescence measurements at *ca.* 720 hours) (Figure 5A). Aggregates formed by BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> were analyzed by X-ray fibre diffraction. Both diffraction patterns obtained show two reflections: a sharp meridional one at 4.7 Å and a broad and more diffuse equatorial one at *ca.* 9.5 Å (Figure 5D). Despite difficulties encountered in aligning the fibrils, these reflections are consistent with a cross- $\beta$  structure characteristic of amyloid fibrils and reflect the distance between the  $\beta$ -strands in each sheet of the amyloid protofilament (4.7 Å) and the spacing between the  $\beta$ -sheets (*ca.* 9.5 Å) [48].

These results show that, although the three chimeras BlaP(Gln)<sub>30</sub>, BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> are equally destabilized under the conditions used to monitor aggregation (i.e. 50% of the molecules of each protein are unfolded at 25°C in the presence of 1.85 M urea), only those with 55 and 79 glutamines significantly



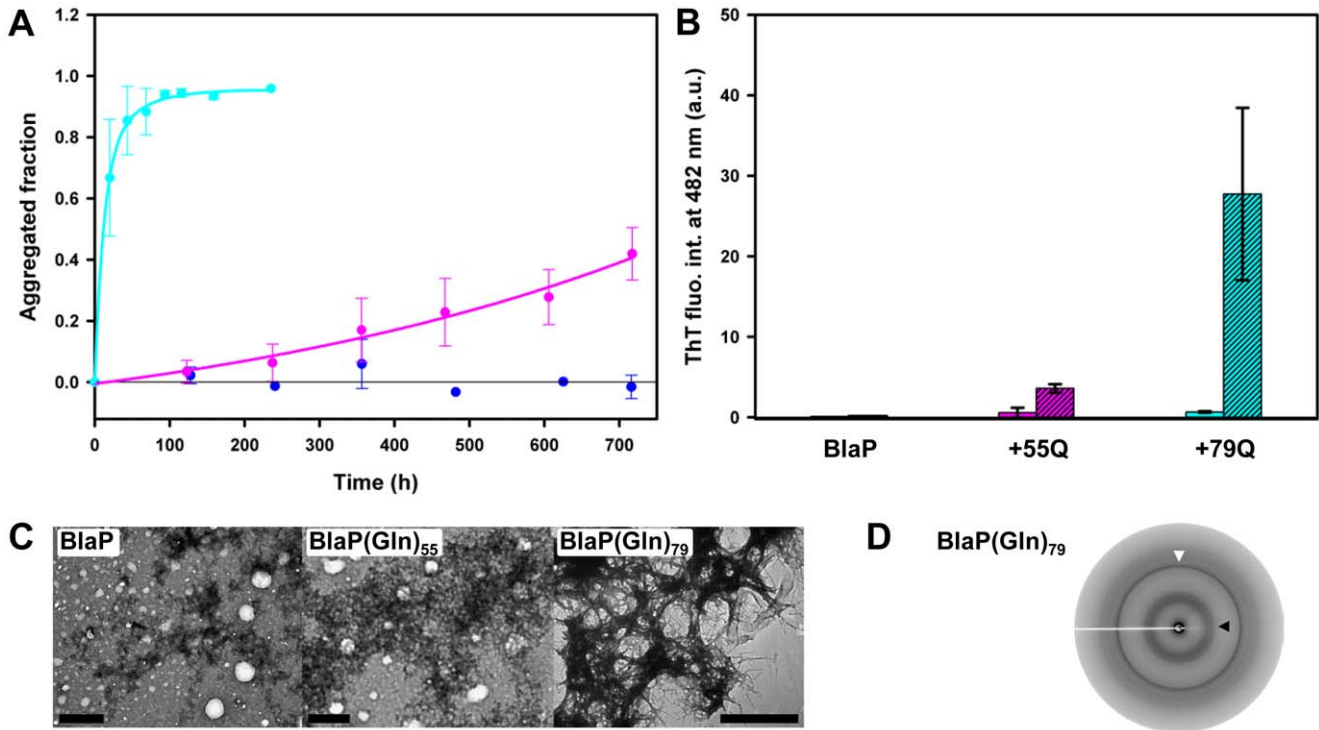
aggregate into amyloid-like fibrils within the time scale of the experiment. Hence, independently of the extent of destabilization, there seems to be a minimal length (i.e. a threshold) of the polyQ tract beyond which chimeric  $\beta$ -lactamases readily aggregate. This value is comprised between 30 and 55 residues, which is reminiscent of the threshold number of glutamines above which other model proteins fused to a polyQ tract aggregate both *in vivo* and *in vitro* [31,49] and, more importantly, of the pathological threshold observed in polyQ diseases [5].

BlaP and all the chimeras (with 23 to 79Q) were also incubated at 25°C in the presence of 3.5 M urea, a concentration at which all proteins are unfolded (Figure 4A). Under these conditions, again, only BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> noticeably aggregate into fibrils that possess characteristics of amyloid such as ThT binding and cross- $\beta$  structure (Figure 6A–D). Although BlaP, BlaP(Gln)<sub>23</sub> and BlaP(Gln)<sub>30</sub> remain soluble, small amounts of amorphous aggregates are observed by TEM for these proteins; these aggregates do not bind ThT (Figure 6A–C).

Taken together, the results of these two experiments give a clear indication that the unfolding of the  $\beta$ -lactamase moiety is not the driving force underlying the fibrillar aggregation. Rather, it is purely the expansion of the polyQ tract above a threshold length that promotes the formation of amyloid-like fibrils. Moreover, in the presence of 1.85 M or 3.5 M urea, the chimera with 79 glutamines aggregates faster than that containing 55 glutamines. Such a polyQ length-dependent rate of aggregation has been observed in all *in vitro* studies of polyQ peptides [23,50] and

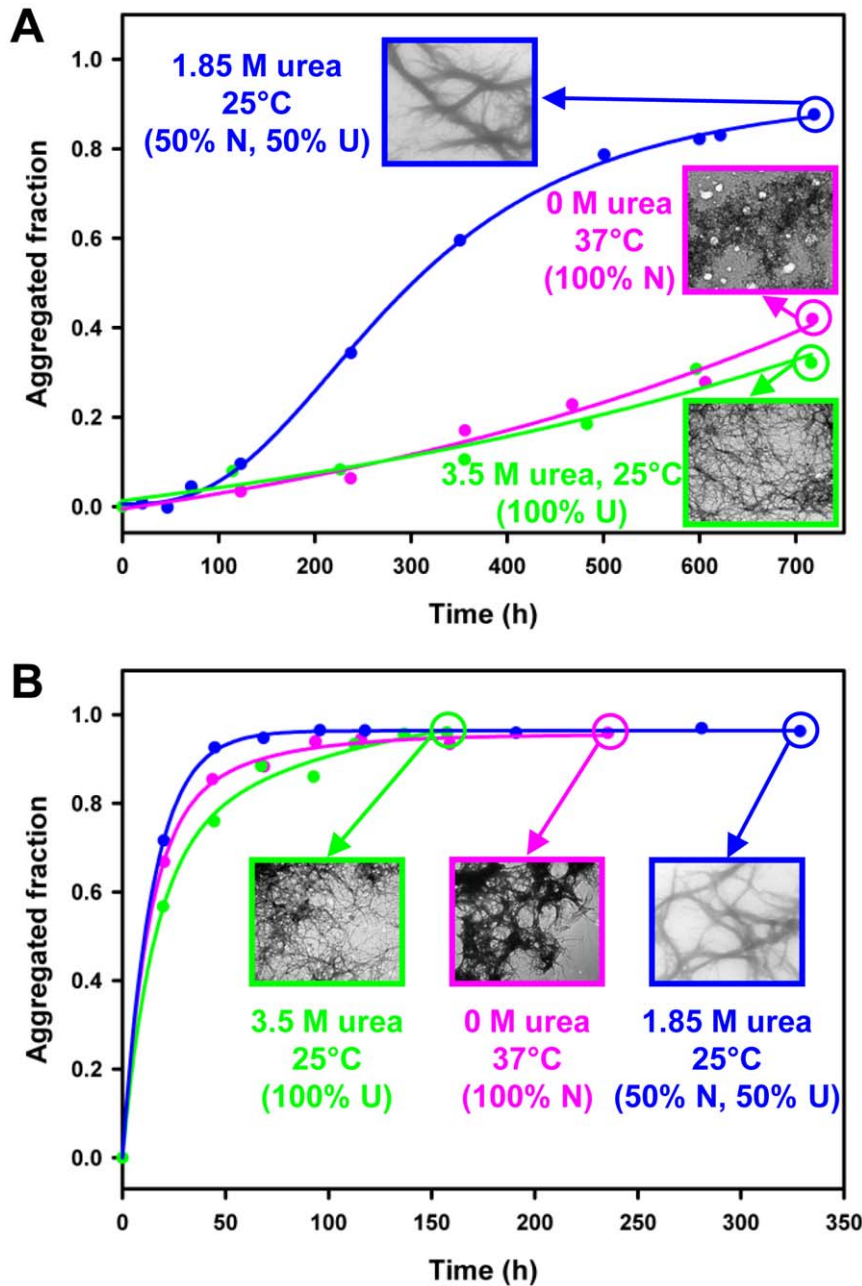
different protein systems [31,35], and can be related to the so called “anticipation phenomenon” characteristic of polyQ diseases [4].

**Aggregation under native conditions.** Finally, the kinetics of aggregation of BlaP, BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> were monitored under conditions favouring the native state (PBS, pH 7.5, 37°C) (Figure 4B). As observed under denaturing conditions (Figures 5 and 6), only chimeras with 55 and 79 glutamines aggregate, whereas BlaP remains soluble throughout the duration of the experiment (Figure 7A). The kinetics of aggregation of BlaP(Gln)<sub>79</sub> was similar to that obtained in the presence of urea (Figures 5A and 6A) and the aggregates formed significantly bind ThT (Figure 7B), display a fibrillar morphology (Figure 7C), and exhibit a X-ray fibre diffraction pattern consistent with a cross- $\beta$  structure (Figure 7D). These observations are all indicative of amyloid-like fibril formation by BlaP(Gln)<sub>79</sub> under native conditions. In contrast, BlaP(Gln)<sub>55</sub> aggregates to a lesser extent than in the presence of 1.85 M urea, and the resulting species bind ThT only weakly (Figure 7B) and appear amorphous when viewed by TEM (Figure 7C). These aggregates did not mature into amyloid-like fibrils upon further incubation of up to 1500 hours (data not shown). Similarly, only small, disperse and amorphous aggregates that do not bind ThT are visible by TEM in BlaP samples at the end of the experiment (Figure 7B and C). These results show that under native conditions, only the chimeric  $\beta$ -lactamase with 79 glutamines forms amyloid-like fibrils. Our data also show that under native conditions, the polyQ threshold



**Figure 7. Only BlaP(Gln)<sub>79</sub> form amyloid-like fibrils when incubated under native conditions.** (A) Aggregation kinetics of 110  $\mu$ M BlaP (blue), BlaP(Gln)<sub>55</sub> (pink) and BlaP(Gln)<sub>79</sub> (cyan) at 37°C in PBS, pH 7.5, followed by measuring the concentration of protein remaining soluble. Time-points shown with an error bar are the average of three independent time-courses for BlaP(Gln)<sub>55</sub> and BlaP(Gln)<sub>79</sub> and two independent time-courses for BlaP. Error bars show the standard deviations. (B) ThT fluorescence intensities at 482 nm in the presence of BlaP and chimeras samples at  $T_0$  (solid bars) and  $T_f$  (dashed bars).  $T_0$  and  $T_f$  correspond to the initial and final points of one time-course for each protein. Data are the average of three measurements and error bars represent the standard deviations. ThT fluorescence intensities in the presence of BlaP samples are weak and for more visibility, error bars (which are equal or less than to 0.3) are not shown. a.u., arbitrary units. (C) TEM images of the protein samples at  $T_f$ . The scale bar is 1  $\mu$ m. (D) X-ray fibre diffraction pattern from BlaP(Gln)<sub>79</sub> fibrils at  $T_f$ . White and black arrows indicate meridional and equatorial reflections at 4.7 Å and ca. 9.5 Å, respectively.

doi:10.1371/journal.pone.0031253.g007



**Figure 8. The ability of BlaP(Gln)<sub>79</sub> to form amyloid-like fibrils does not depend on the structural integrity of BlaP.** Comparison between the aggregation kinetics and the morphology of aggregates obtained with BlaP(Gln)<sub>55</sub> (A) and BlaP(Gln)<sub>79</sub> (B) under the following conditions of incubation: (i) PBS, pH 7.5 and 0 M urea at 37°C (pink), (ii) PBS, pH 7.5 and 1.85 M urea at 25°C (blue) and (iii) PBS, pH 7.5 and 3.5 M urea at 25°C (green). N is the native state and U is the unfolded state. The data correspond to those shown in Figures 5, 6, 7 without error bars. doi:10.1371/journal.pone.0031253.g008

length for aggregation into fibrillar aggregates is higher than that observed under denaturing conditions, suggesting that its value depends on the structural integrity of the  $\beta$ -lactamase moiety. Indeed, BlaP(Gln)<sub>55</sub> is only able to form amyloid-like aggregates when BlaP is unfolded (Figure 8A) whereas, remarkably, BlaP(Gln)<sub>79</sub> aggregates into amyloid-like fibrils at the same rate irrespective of the incubation conditions and therefore, of the conformational state (i.e. native or unfolded) of BlaP (Figure 8B). This observation shows that the presence of folded BlaP suppresses the intrinsic propensity of the 55-glutamine tract to trigger the self-association of the chimera into amyloid-like fibrils. In other words, in the presence of 55Q, the protective role of BlaP in polyQ-driven

aggregation into amyloid-like fibrils depends on the integrity of the BlaP structure. In contrast, the conformational state of BlaP has no discernable effect on the ability of the 79-glutamine tract to promote the formation of amyloid-like fibrils.

The dependence of BlaP(Gln)<sub>55</sub> aggregation on the structural integrity of BlaP can result from both steric and conformational constraints. The folded  $\beta$ -lactamase moiety may limit the accessibility of the 55-glutamine tract via steric hindrances and thus abrogate its propensity to form the highly ordered intermolecular  $\beta$ -sheets characteristic of amyloid fibrils. The 55-glutamine tract in the presence of folded BlaP leads, nevertheless, to less organized intermolecular interactions and thus to the

formation of amorphous aggregates. The unfolding of BlaP can however render the 55Q tract more accessible and thus permits the formation of amyloid-like fibrils. In contrast, when expanded to 79 glutamines, the polyQ tract is similarly accessible to interact with other monomers to form amyloid-like fibrils whether BlaP is denatured or not. The need to unfold BlaP in order to generate amyloid-like fibrils from the 55Q chimera is reminiscent of that of proteolysis observed with some proteins associated with polyQ diseases. For example, nuclear inclusions from HD patients have been shown to contain Htt fragments including the expanded polyQ tract rather than the full-length protein [51]. It has been proposed that these fragments have a higher tendency to self-associate and thus may be crucial to initiate the aggregation phenomenon [28,29]. Their greater propensity to aggregate is probably due to the removal by proteolysis of some of the steric hindrances otherwise exerted by the non-polyQ domains.

The protective effects mediated by the folded state of BlaP could also be due to the imposition of some conformational constraints on the polyQ tract. In a similar vein of what has been suggested by a number of experimental and computational studies [23,52,53,54], we propose that the polyQ tract which extends from the solvent-exposed loop in BlaP exists as an ensemble of heterogeneous disordered conformations, in dynamic equilibrium with rare more structured conformational species, some of which are competent for amyloid fibril formation. The longer the polyQ repeat, the more dynamic it is and the wider the conformational space it can sample. As a consequence, more aggregation-prone conformations can be transiently sampled by longer polyQ tracts, thus explaining their higher propensity to aggregate. The length-dependent frequency at which the amyloid-competent conformations are visited, and thus the propensity of the polyQ tract to aggregate into amyloid-like fibrils, is further modulated by the conformational state of the  $\beta$ -lactamase moiety. Short polyQ stretches ( $\leq 30$  residues) rarely or never access amyloid-competent conformations, even if the  $\beta$ -lactamase is in its unfolded state. In contrast, such conformations are equally sampled by the long polyQ tracts ( $\geq 79$  residues) regardless of whether BlaP is native or unfolded. Finally, for polyQ stretches of intermediate lengths (i.e. 30–79 residues), the frequency at which amyloid-competent conformations are adopted critically depends on the structural integrity of BlaP. For instance, the presence of folded BlaP prevents the 55-glutamine tract from accessing conformational precursors to amyloid-like fibril formation. Conversely, the unfolding of the  $\beta$ -lactamase moiety would allow the 55-glutamine tract to interconvert more freely between various conformations and hence access, more frequently, conformations that are competent for amyloid-like fibril formation. The probability of accessing these conformations is however lower for the 55- than the 79-glutamine tract, and thus a lag phase is observed for BlaP(Gln)<sub>55</sub> fibril formation. These results are consistent with previous reports showing the importance of the constraints imposed by adjacent sequences on the polyQ tract [23,30]. For example, it was reported that the C-terminal addition of 10 or 11 prolines to polyQ peptides tends to decrease both aggregation rate and aggregate stability, and increases the threshold for fibril formation by disfavoring aggregation-competent conformations [23,30]. The study of Darnell *et al.* strongly suggests that it is the formation of constraining polyproline type II helices by the C-terminal prolines that tips the balance in favor of aggregation-incompetent conformations [30]. Moreover, a recent study has shown that when a polyQ tract of pathological length is positioned between two SpA domains, it triggers slower aggregation than when it is located at the N-terminus of one SpA domain [34]. The authors proposed that the reduction of the

aggregation rate is due to lower conformational freedom and higher steric constraints.

Results obtained for the other model proteins with an inserted polyQ tract can tentatively be analyzed in terms of the balance between the intrinsic propensity of polyQ tracts to trigger aggregation into amyloid-like fibrils and the constraints imposed by the host protein on the polyQ tract. It is interesting to note that the chimera of myoglobin containing 50Q aggregates into amyloid-like fibrils when incubated under conditions similar to those where BlaP(Gln)<sub>55</sub> does not [18]. This observation suggests that steric/conformational constraints imposed on the polyQ tract in myoglobin are lower than those in BlaP. These lower constraints could, at least in part, originate from the fact that (i) the structure of myoglobin is perturbed by the insertion of the polyQ tract while the structure of BlaP is not and (ii) the loop in which the polyQ tract is inserted is significantly longer in myoglobin than in BlaP, essentially due to the addition of several amino acids from ataxin-3 at both sides of the polyQ tract [37]. In the case of apomyoglobin, the insertion of 38Q, the longest tract used in the aggregation studies, does not induce amyloid-like fibril formation. This observation suggests that a 38Q tract inserted into apomyoglobin is not long enough to access amyloid-competent conformations and/or to overcome the steric hindrances exerted by apomyoglobin.

Under native conditions, traces of dimeric species are observed for both BlaP and all chimeras with the highest amount being observed for wild-type BlaP (Table 1) which does not form amyloid-like fibrils. It is therefore very unlikely that the observed dimers act as seeds to facilitate fibril formation by BlaP(Gln)<sub>79</sub> but not, for example, BlaP(Gln)<sub>55</sub>. The high molecular weight oligomeric species are however observed only for BlaP(Gln)<sub>79</sub>; since this protein is the only one to aggregate into amyloid-like fibrils under native conditions, the observed oligomers could indeed act as seeds of fibril formation and thus accelerate the process. A deeper characterization of the structural properties of the oligomeric species and of their role in the process of aggregation is under investigation. The potential of these species to act as seeds does not, however, invalidate the conclusions of the work. Indeed, this would purely imply that the constraints applied by the BlaP moiety act to prevent the formation of amyloid-like fibrils, at least in part, by preventing the formation of oligomeric species that are formed early on the pathway of aggregation.

## Conclusions

We have created and characterized a series of chimeras with 23, 30, 55 and 79 glutamines inserted into a solvent-exposed loop of a globular protein, the  $\beta$ -lactamase BlaP. The threshold number of glutamines above which the BlaP chimeras aggregate into amyloid-like fibrils critically depends on the structural integrity of the  $\beta$ -lactamase moiety. This result suggests that this threshold value results, at least in part, from a delicate balance between the intrinsic propensity of polyQ tracts to aggregate and the extrinsic protective conformational/steric constraints originating from BlaP. While it has been suggested that, in proteins associated with diseases, polyQ tracts are located in regions that are essentially unstructured and generally N- or C-terminal to a structured domain [35], several studies have shown that sequences flanking the polyQ tract could, however, adopt some elements of secondary structure [30,55]. The latter could therefore exert constraints on the polyQ tract similar to those exerted by the BlaP moiety. The threshold for amyloid-like fibril formation by the BlaP chimeras under native conditions ( $>55Q$ ) is higher than the highest threshold length observed for proteins associated with diseases (49 residues in atrophin-1, [5]). This observation suggests

that the constraints applied to the polyQ tract by BlaP in its folded state are higher than those imposed by the proteins associated with polyQ diseases, probably because BlaP is a more structured and rigid scaffold. The use of this globular protein has however allowed us to produce and characterize, for the first time, a model protein with an inserted 79Q tract. Interestingly, our results clearly show that the structural integrity of BlaP, and thus the constraints imposed on such a long tract, has negligible impact on the propensity of the latter to mediate amyloid-like fibril formation. Based on these observations, we speculate that the modulating effects of the protein context on the aggregating properties of proteins associated with polyQ diseases could also be negligible when a particularly long polyQ tract is present. PolyQ tracts of 88–306 residues were reported for some proteins associated with polyQ diseases [1]; the aggregating properties of proteins with such long tracts could be, therefore, dictated essentially by the intrinsic propensity of the polyQ expansion to form intermolecular  $\beta$ -sheets.

Finally, this study demonstrates that BlaP is an appropriate scaffold to further investigate the delicate balance between the propensity of polyQ tracts to trigger aggregation and the modulating effects of the host proteins.

## Materials and Methods

### Molecular biology

A library of (CAG)<sub>n</sub> double-strand DNA fragments was constructed by an overlapping PCR strategy [56], using the following oligonucleotides: 5'-(CAG)<sub>13</sub>-3' and 5'-(CTG)<sub>13</sub>-3'. A PCR using Pfu DNA polymerase (Promega, Madison, WI, USA) was performed as follows: 3 min at 95°C, 30 sec at 94°C, 1 min at 55°C and 30 sec at 68°C; the last three steps were repeated 35 times. The library was purified from a 2% agarose gel and the extremities of each double-strand DNA fragment were blunt-ended with Pfu DNA polymerase and dNTPs at 68°C over a period of 30 min. The polyCAG DNA fragments were then dephosphorylated for 30 min at 37°C using calf intestinal alkaline phosphatase (Roche, Mannheim, Germany). Finally, the polyCAG double-strand DNA library was inserted into the SmaI restriction site, which was previously introduced in the gene of BlaP carried by the constitutive expression vector pNYESBlaP [44]. In this vector, the gene of BlaP is followed by the nucleotide sequence coding for an additional C-terminal dipeptide glycine-proline and a (His)<sub>5</sub> tag. The resulting library of expression vectors was used to transform *E. coli* DH5 $\alpha$  cells (Invitrogen, Paisley, UK). Note that the  $\beta$ -lactamase activity which enables *E. coli* to become resistant to antibiotics can be used as a reporter to efficiently select clones producing soluble chimeras in which BlaP is correctly folded. Consequently, the transformed cells were plated on LB (Luria-Bertani) medium containing ampicillin (10  $\mu\text{g}\cdot\text{mL}^{-1}$ , Sigma) in addition to spectinomycin (75  $\mu\text{g}\cdot\text{mL}^{-1}$ , Sigma) for which the plasmid contains a resistance gene. The presence of an insert within the BlaP gene was checked by colony-PCR on more than 50 randomly selected transformants, and plasmids from about 10 colonies carrying an insertion were amplified and extracted. Their sequences were determined by the Sanger method at the GIGA GenoTranscriptomics technology platform (Liège, Belgium).

### Protein expression and purification

The proteins were expressed in *E. coli* JM109 strains (Promega). A 100 mL LB medium preculture, containing 75  $\mu\text{g}\cdot\text{mL}^{-1}$  of spectinomycin and 10  $\mu\text{g}\cdot\text{mL}^{-1}$  of ampicillin, was inoculated with the transformed cells and incubated at 37°C for approximately 7 hours. Two liters of modified TB (Terrific Broth) medium containing no glycerol, but supplemented with 75  $\mu\text{g}\cdot\text{mL}^{-1}$  of

spectinomycin, 10  $\mu\text{g}\cdot\text{mL}^{-1}$  of ampicillin, 4.2  $\mu\text{M}$  of biotin and 2.2  $\mu\text{M}$  of riboflavin, were inoculated with 5 mL of preculture and incubated overnight at 37°C. Periplasmic proteins were then extracted by osmotic shock as described by Vandevenne *et al.* [44]. The proteins of interest, which were expressed with a C-terminal (His)<sub>5</sub> tag, were purified in a single step by metal chelate affinity chromatography, using a 20 mL Ni-PDC column (Affiland, Liège, Belgium). After loading the periplasmic extract, the column was washed successively with 60 mL of: (i) PBS (50 mM sodium phosphate, pH 7.5, containing 150 mM NaCl); (ii) PBS containing 2 M NaCl; (iii) PBS containing 10 mM imidazole. Proteins were eluted with a linear imidazole gradient (0–300 mM) in PBS. Enzymatically active fractions (probed with nitrocefin as the substrate) containing more than 95% of the protein of interest [as assessed by 15% (w/v) SDS-PAGE] were pooled. They were then either dialyzed four times against 15 L of milliQ water, lyophilized and stored at 4°C, or dialyzed two times against 15 L of PBS and stored at –20°C.

### Size-exclusion chromatography

A Superdex 200 GL 10/300 column (G.E. Healthcare, Uppsala, Sweden) was equilibrated with PBS buffer, pH 7.5. Solutions of BlaP and the four chimeras at 50  $\mu\text{M}$  and 120  $\mu\text{M}$ , and a solution of the chimera with 79 glutamines at 10  $\mu\text{M}$ , in PBS, pH 7.5, were injected and eluted at a flow rate of 0.5  $\text{mL}\cdot\text{min}^{-1}$ . Elution of protein was monitored by absorbance measurements at 280 nm. The column was calibrated using 7 protein standards: hen egg white lysozyme (Belovo, Bastogne, Belgium), chicken egg white albumin (Sigma, A2512), phosphor-ylase B from rabbit muscle (Sigma, P6635),  $\beta$ -galactosidase from *E. coli* (Sigma, 48275), aldolase from rabbit muscle (Sigma, A2714), catalase from bovine liver (Sigma, C40) and thyroglobulin from bovine thyroid (Sigma, T9145).

### N-terminal sequencing and mass spectrometry

N-terminal sequencing was performed using the Edman degradation procedure according to Han's protocol [57]. The molecular masses were determined using electrospray ionisation quadrupole time-of-flight mass spectrometry (ESI-QTOF-MS) at the GIGA proteomics platform (Liège, Belgium). The percentage of each species observed within the same sample injection run was estimated from the centroid value of peak.

### Quantification of the protein

The molar extinction coefficient (33000  $\text{M}^{-1}\cdot\text{cm}^{-1}$  at 280 nm) of BlaP was determined experimentally using the BCA assay from Pierce (Rockford, IL, USA). This value was used for the determination of the concentrations of both wild-type and chimeric enzyme solutions.

### Enzymatic activity measurements

Kinetic parameters were determined at 30°C with cephalothin (Sigma C4520, 50–100  $\mu\text{M}$ ) as the substrate, in 50 mM sodium phosphate buffer, pH 7, using a UVIKON 860 spectrophotometer (Kontron Instrument, Zürich, Switzerland), as described by Matagne *et al.* [43].

### Fluorescence and circular dichroism measurements

Fluorescence data were acquired using either a Cary Eclipse spectrofluorimeter equipped with a Peltier-controlled holder (Varian, Mulgrave, Australia) or a LS50B spectrofluorimeter (Perkin-Elmer, Norwalk, CT, USA) and a 1 cm pathlength cell. CD measurements were performed on a Jasco J-810 spectropo-

larimeter (Jasco, Tokyo, Japan) equipped with a Peltier-controlled holder and using 1 mm and 1 cm pathlength cells for far-UV and near-UV measurements, respectively.

### Fluorescence and circular dichroism spectra of native proteins

All spectra were recorded at 25°C in PBS, pH 7.5, using protein concentrations of 4.6  $\mu\text{M}$  for fluorescence and far-UV CD measurements, and 20  $\mu\text{M}$  for near-UV CD measurements. Five fluorescence emission spectra were recorded in the 300–400 nm range ( $\lambda_{\text{exc}} = 295$  nm,  $\text{slit}_{\text{exc}} = 2.5$  nm,  $\text{slit}_{\text{em}} = 5$  nm), at a rate of 600  $\text{nm}\cdot\text{min}^{-1}$  using the Cary Eclipse spectrofluorimeter, and averaged. Twenty CD spectra were acquired at a rate of 50  $\text{nm}\cdot\text{min}^{-1}$ , both in the far-UV (200–250 nm) and near-UV (250–350 nm) regions, and averaged. The bandwidth and the response time were 1 nm and 0.5 sec, respectively. All protein spectra (fluorescence and CD) were corrected for the buffer contribution.

### Urea-induced unfolding experiments

Protein samples (4.6  $\mu\text{M}$ ) in PBS, pH 7.5, at various urea concentrations (VWR BDH Prolabo, 0 to 5.5 M by 0.1 M increments) were unfolded to equilibrium by incubation for *ca.* 16 h at 25°C. Unfolding was monitored by changes in intrinsic fluorescence ( $\lambda_{\text{exc}} = 295$  nm,  $\lambda_{\text{em}} = 323$  nm,  $\text{slits}_{\text{exc/em}} = 2.6$  nm) using the LS50B spectrofluorimeter, and in far-UV CD signal at 222 nm (bandwidth = 1 nm, response = 4 sec), as described previously [58]. The background of the solutions (PBS buffer+denaturant) was subtracted from the fluorescence and CD signals. Urea concentrations were determined from the refractive index measurements [59] using a R5000 refractometer from Atago (Tokyo, Japan). Moreover, other protein samples were denatured for 3 hours at 25°C in 5.5 M urea (under these conditions, all the investigated proteins have been shown to be completely unfolded) and renatured (for *ca.* 16 h) by dilution to different urea concentrations (from 5.5 to 0.55 M). The reversibility of the unfolding transitions was demonstrated by comparing the fluorescence and CD signals recorded with these samples to those obtained for the samples unfolded at similar urea concentrations. Equilibrium unfolding curves were analyzed on the basis of a two-state model ( $\text{N} \rightleftharpoons \text{U}$ ), as previously described [58,60,61,62].

### Heat-induced unfolding

Heat-induced unfolding was monitored by the changes in the intrinsic fluorescence intensity ( $\lambda_{\text{exc}} = 295$  nm,  $\lambda_{\text{em}} = 323$  nm,  $\text{slits}_{\text{exc/em}} = 5$  nm) using the Cary Eclipse spectrofluorimeter, and in far-UV CD signal at 222 nm (bandwidth = 1 nm, response = 4 sec). The protein concentration was 4.6  $\mu\text{M}$  in PBS, pH 7.5, and mineral oil was added on top of the samples to limit solvent evaporation. The temperature was increased from 25 to 85–90°C at a rate of 0.5°C·min<sup>-1</sup>; the fluorescence and the CD data were acquired every 0.5°C. The temperature in the cell was measured with a PT200 thermocouple (IMPO Electronic, Olgod, Denmark). The reversibility of the heat-induced unfolding was assessed by monitoring the changes in the fluorescence and CD signals upon cooling the sample down to 25°C at a rate of 0.5°C·min<sup>-1</sup>. Data were analyzed on the basis of a two-state model ( $\text{N} \rightleftharpoons \text{U}$ ), as described by El Hajjaji *et al.* [63].

### Aggregation kinetics

A series of tubes containing 100  $\mu\text{L}$  of 110  $\mu\text{M}$  protein in PBS, pH 7.5, containing 0.2% sodium azide, were incubated

either at 37°C in the absence of urea or at 25°C in the presence of 1.85 or 3.5 M urea. Airtight tubes (Multiply Safecup, Sarstedt, Nümbrecht, Germany) were used to limit evaporation. At selected times, one tube was removed and centrifuged for 50 min at 12000 rpm. The supernatant was used to determine the quantity of soluble protein by absorbance measurements at 280 nm and the protein integrity was demonstrated by SDS-PAGE analysis. Unless otherwise stated, the aggregation time-courses were repeated three times with proteins originating from different production and purification batches. For one of the aggregation time-courses of each protein, aliquots of samples at the initial ( $T_0$ ) and end ( $T_f$ ) time-points were taken (in triplicate) before centrifugation for thioflavin T (ThT) fluorescence measurements; they were kept frozen until analysis. Samples at  $T_f$  were also analyzed by transmission electron microscopy.

### Thioflavin T (ThT) fluorescence measurements

To 5  $\mu\text{L}$  protein sample was added 1.5 mL of 10 mM sodium phosphate buffer, 150 mM NaCl, 50  $\mu\text{M}$  ThT (Sigma, T3516), pH 7. Ten fluorescence emission spectra were recorded (using the Cary Eclipse spectrofluorimeter) at 25°C with stirring in the 450–600 nm range ( $\lambda_{\text{exc}} = 440$  nm,  $\text{slit}_{\text{exc/em}} = 5$  nm) at a rate of 1200  $\text{nm}\cdot\text{min}^{-1}$ , averaged, and corrected for the background fluorescence of the ThT solution alone.

### Transmission electron microscopy

Samples were left for 4 min on carbon-coated 400-mesh copper grids, before being stained for 1 min with 2% uranyl acetate (w/v). The grids were then washed once with 2% uranyl acetate and finally, three times with milliQ water. Images were recorded on a Philips CEM100 transmission electron microscope operating at 100 kV.

### X-ray fibre diffraction

Fibril suspensions were centrifuged and the pellet subjected to two washing cycles in milliQ water; between each cycle, the suspension was centrifuged and the supernatant discarded in order to remove any trace of soluble protein, buffer, or urea. The fibrils were then aligned using a modification of the stretchframe method as previously described [64]. X-ray diffraction data were collected at room temperature for *ca.* 10 min on a Bruker AXS FR591 diffractometer with Cu K $\alpha$  radiation with a wavelength of 1.5418 Å and equipped with a MARDTB 345 mm image plate detector. The sample-to-detector distance was 300 mm.

### Acknowledgments

We acknowledge Fabrice Bouillenne and Anne-Marie Matton for their assistance during purification of the proteins, and Dr. David Thorn, Dr. Caroline Montagner and Prof. Jean-Marie Frère for critical reading of the manuscript. We also thank the group of Prof. Kurt Hoffmann (Bioanalytics unit of the Institute of Molecular Biotechnology RWTH-Aachen University, Germany) for their assistance with the X-ray fibre diffraction experiments.

### Author Contributions

Conceived and designed the experiments: NS AM MD. Performed the experiments: NS CP AF. Analyzed the data: NS CP AM MD. Contributed reagents/materials/analysis tools: NY GG PF MG. Wrote the paper: NS AM MD.

## References

- Orr HT, Zoghbi HY (2007) Trinucleotide repeat disorders. *Annu Rev Neurosci* 30: 575–621.
- Ordway JM, Tallaksen-Greene S, Gutekunst CA, Bernstein EM, Cearley JA, et al. (1997) Ectopically expressed CAG repeats cause intranuclear inclusions and a progressive late onset neurological phenotype in the mouse. *Cell* 91: 753–763.
- Robitaille Y, Lopes-Cendes I, Becher M, Rouleau G, Clark AW (1997) The neuropathology of CAG repeat diseases: review and update of genetic and molecular features. *Brain Pathol* 7: 901–926.
- Zoghbi HY, Orr HT (2000) Glutamine repeats and neurodegeneration. *Annu Rev Neurosci* 23: 217–247.
- Robertson AL, Bottomley SP (2010) Towards the treatment of polyglutamine diseases: the modulatory role of protein context. *Curr Med Chem* 17: 3058–3068.
- Diaz-Hernandez M, Moreno-Herrero F, Gomez-Ramos P, Moran MA, Ferrer J, et al. (2004) Biochemical, ultrastructural, and reversibility studies on huntingtin filaments isolated from mouse and human brain. *J Neurosci* 24: 9361–9371.
- Sanchez I, Mahlke C, Yuan J (2003) Pivotal role of oligomerization in expanded polyglutamine neurodegenerative disorders. *Nature* 421: 373–379.
- Williams AJ, Paulson HL (2008) Polyglutamine neurodegeneration: protein misfolding revisited. *Trends Neurosci* 31: 521–528.
- Ross CA (1997) Intranuclear neuronal inclusions: a common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? *Neuron* 19: 1147–1150.
- Yang W, Dunlap JR, Andrews RB, Wetzel R (2002) Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells. *Hum Mol Genet* 11: 2905–2917.
- Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, et al. (2007) A toxic monomeric conformer of the polyglutamine protein. *Nat Struct Mol Biol* 14: 332–340.
- Mukai H, Isagawa T, Goyama E, Tanaka S, Bence NF, et al. (2005) Formation of morphologically similar globular aggregates from diverse aggregation-prone proteins in mammalian cells. *Proc Natl Acad Sci U S A* 102: 10887–10892.
- Olshina MA, Angley LM, Ramdzan YM, Tang J, Bailey MF, et al. (2010) Tracking mutant huntingtin aggregation kinetics in cells reveals three major populations that include an invariant oligomer pool. *J Biol Chem* 285: 21807–21816.
- Takahashi T, Kikuchi S, Katada S, Nagai Y, Nishizawa M, et al. (2008) Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. *Hum Mol Genet* 17: 345–356.
- Wetzel R (2010) Misfolding and aggregation in Huntington disease and other expanded polyglutamine repeat diseases. In: Ramirez-Alvarado M, Kelly JW, Dobson CM, eds. *Protein Misfolding Diseases: Current and Emerging Principles and Therapies* John Wiley and sons, Inc. pp 305–324.
- Perutz MF, Johnson T, Suzuki M, Finch JT (1994) Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci U S A* 91: 5355–5358.
- Ignatova Z, Thakur AK, Wetzel R, Gierasch LM (2007) In-cell aggregation of a polyglutamine-containing chimera is a multistep process initiated by the flanking sequence. *J Biol Chem* 282: 36736–36743.
- Tanaka M, Morishima I, Akagi T, Hashikawa T, Nukina N (2001) Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. *J Biol Chem* 276: 45470–45475.
- Thakur AK, Jayaraman M, Mishra R, Thakur M, Chellgren VM, et al. (2009) Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. *Nat Struct Mol Biol* 16: 380–389.
- Johnson SM, Wiseman RL, Sekijima Y, Green NS, Adamski-Werner SL, et al. (2005) Native state kinetic stabilization as a strategy to ameliorate protein misfolding diseases: a focus on the transthyretin amyloidosis. *Acc Chem Res* 38: 911–921.
- Dumoulin M, Kumita JR, Dobson CM (2006) Normal and aberrant biological self-assembly: Insights from studies of human lysozyme and its amyloidogenic variants. *Acc Chem Res* 39: 603–610.
- Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D, et al. (2003) A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature* 424: 783–788.
- Bhattacharyya A, Thakur AK, Chellgren VM, Thiagarajan G, Williams AD, et al. (2006) Oligoproline effects on polyglutamine conformation and aggregation. *J Mol Biol* 355: 524–535.
- Sharma D, Sharma S, Pasha S, Brahmachari SK (1999) Peptide models for inherited neurodegenerative disorders: conformation and aggregation properties of long polyglutamine peptides with and without interruptions. *FEBS Lett* 456: 181–185.
- Bulone D, Masino L, Thomas DJ, San Biagio PL, Pastore A (2006) The interplay between PolyQ and protein context delays aggregation by forming a reservoir of protofibrils. *PLoS ONE* 1: e111.
- Hollenbach B, Scherzinger E, Schweiger K, Lurz R, Lehrach H, et al. (1999) Aggregation of truncated GST-HD exon 1 fusion proteins containing normal range and expanded glutamine repeats. *Philos Trans R Soc Lond B Biol Sci* 354: 991–994.
- Nozaki K, Onodera O, Takano H, Tsuji S (2001) Amino acid sequences flanking polyglutamine stretches influence their potential for aggregate formation. *Neuroreport* 12: 3357–3364.
- Cooper JK, Schilling G, Peters MF, Herring WJ, Sharp AH, et al. (1998) Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Hum Mol Genet* 7: 783–790.
- Martindale D, Hackam A, Wiczorek A, Ellerby L, Wellington C, et al. (1998) Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nat Genet* 18: 150–154.
- Darnell G, Orgel JP, Pahl R, Meredith SC (2007) Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. *J Mol Biol* 374: 688–704.
- Ignatova Z, Gierasch LM (2006) Extended polyglutamine tracts cause aggregation and structural perturbation of an adjacent beta barrel protein. *J Biol Chem* 281: 12959–12967.
- Ellisdon AM, Thomas B, Bottomley SP (2006) The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J Biol Chem* 281: 16888–16896.
- Ellisdon AM, Pearce MC, Bottomley SP (2007) Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein. *J Mol Biol* 368: 595–605.
- Robertson AL, Bate MA, Buckle AM, Bottomley SP (2011) The rate of polyQ-mediated aggregation is dramatically affected by the number and location of surrounding domains. *J Mol Biol* 413: 879–887.
- Robertson AL, Home J, Ellisdon AM, Thomas B, Scanlon MJ, et al. (2008) The structural impact of a polyglutamine tract is location-dependent. *Biophys J* 95: 5922–5930.
- Saunders HM, Gilis D, Rooman M, Dehouck Y, Robertson AL, et al. (2011) Flanking domain stability modulates the aggregation kinetics of a polyglutamine disease protein. *Protein Sci* 20: 1675–1681.
- Tobelmann MD, Murphy RM (2011) Location trumps length: polyglutamine-mediated changes in folding and aggregation of a host protein. *Biophys J* 100: 2773–2782.
- Bennett MJ, Hucy-Tubman KE, Herr AB, West AP, Jr., Ross SA, et al. (2002) Inaugural Article: A linear lattice model for polyglutamine in CAG-expansion diseases. *Proc Natl Acad Sci U S A* 99: 11634–11639.
- Masino L, Kelly G, Leonard K, Trotter Y, Pastore A (2002) Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. *FEBS Lett* 513: 267–272.
- Tobelmann MD, Kerby RL, Murphy RM (2008) A strategy for generating polyglutamine ‘length libraries’ in model host proteins. *Protein Eng Des Sel* 21: 161–164.
- Klein FA, Pastore A, Masino L, Zeder-Lutz G, Nierengarten H, et al. (2007) Pathogenic and non-pathogenic polyglutamine tracts have similar structural properties: towards a length-dependent toxicity gradient. *J Mol Biol* 371: 235–244.
- Knox JR, Moews PC (1991) Beta-lactamase of *Bacillus licheniformis* 749/C. Refinement at 2 Å resolution and analysis of hydration. *J Mol Biol* 220: 435–455.
- Matagne A, Misselyn-Bauduin AM, Joris B, Erpicum T, Granier B, et al. (1990) The diversity of the catalytic properties of class A beta-lactamases. *Biochem J* 265: 131–146.
- Vandevenne M, Filee P, Scarafone N, Cloes B, Gaspard G, et al. (2007) The *Bacillus licheniformis* BlAP beta-lactamase as a model protein scaffold to study the insertion of protein fragments. *Protein Sci* 16: 2260–2271.
- Vandevenne M, Gaspard G, Yilmaz N, Giannotta F, Frère JM, et al. (2008) Rapid and easy development of versatile tools to study protein/ligand interactions. *Protein Eng Des Sel* 21: 443–451.
- Ladurner AG, Fersht AR (1997) Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J Mol Biol* 273: 330–337.
- Harper JD, Lansbury PT, Jr. (1997) Models of amyloid seeding in Alzheimer’s disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu Rev Biochem* 66: 385–407.
- Serpell LC, Fraser PE, Sunde M (1999) X-ray fiber diffraction of amyloid fibrils. *Methods Enzymol* 309: 526–536.
- Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, et al. (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90: 549–558.
- Chen S, Berthelie V, Yang W, Wetzel R (2001) Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *J Mol Biol* 311: 173–182.
- DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, et al. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277: 1990–1993.
- Crick SL, Jayaraman M, Frieden C, Wetzel R, Pappu RV (2006) Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc Natl Acad Sci U S A* 103: 16764–16769.

53. Lakhani VV, Ding F, Dokholyan NV (2010) Polyglutamine induced misfolding of huntingtin exon1 is modulated by the flanking sequences. *PLoS computational biology* 6: e1000772.
54. Vitalis A, Wang X, Pappu RV (2007) Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. *Biophys J* 93: 1923–1937.
55. Kim MW, Chelliah Y, Kim SW, Otwinowski Z, Bezprozvanny I (2009) Secondary structure of Huntingtin amino-terminal region. *Structure* 17: 1205–1212.
56. Ordway JM, Detloff PJ (1996) In vitro synthesis and cloning of long CAG repeats. *Biotechniques* 21: 609–610, 612.
57. Han KK, Tetaert D, Debuire B, Dautrevaux M, Biserte G (1977) [Sequential Edman degradation]. *Biochimie* 59: 557–576.
58. Dumoulin M, Conrath K, Van Meirhaeghe A, Meersman F, Heremans K, et al. (2002) Single-domain antibody fragments with high conformational stability. *Protein Sci* 11: 500–515.
59. Pace CN, Scholtz M (1997) Measuring the conformational stability of a protein. In: Creighton TE, ed. *Protein structure: a practical approach*. Second Edition ed IRL press. pp 299–321.
60. Pace CN (1990) Measuring and increasing protein stability. *Trends Biotechnol* 8: 93–98.
61. Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27: 8063–8068.
62. Vandenameele J, Lejeune A, Di Paolo A, Brans A, Frère JM, et al. (2010) Folding of class A beta-lactamases is rate-limited by peptide bond isomerization and occurs via parallel pathways. *Biochemistry* 49: 4264–4275.
63. El Hajjaji H, Dumoulin M, Matagne A, Colau D, Roos G, et al. (2009) The zinc center influences the redox and thermodynamic properties of *Escherichia coli* thioredoxin 2. *J Mol Biol* 386: 60–71.
64. Mossuto MF, Dhulesia A, Devlin G, Frère E, Kumita JR, et al. (2010) The non-core regions of human lysozyme amyloid fibrils influence cytotoxicity. *J Mol Biol* 402: 783–796.
65. Ambler RP, Coulson AF, Frère JM, Ghuyens JM, Joris B, et al. (1991) A standard numbering scheme for the class A beta-lactamases. *Biochem J* 276 (Pt 1): 269–270.