PART 2

CHAPTER 3

Inclusion Bodies: specificity in their aggregation process and amyloid-like structure

3.1 Introduction

3.1.1 Protein folding

Living organisms contain thousands of different proteins that stimulate or control virtually every chemical process on which their lives depend. Despite their molecular diversity, proteins possess the common property that, in general, they can rapidly fold into elaborate three-dimensional structures that are required for their specific functions. It is now known that within the cells there are large numbers of auxiliary factors that assist in the folding process, including folding catalysts and molecular chaperones. These factors serve to enable polypeptide chains to fold efficiently in the complex and crowded milieu of the cell but they do not determine their native structures; the latter are fully encoded by their amino acid sequences. The question of how proteins attain their unique native states is still a mostly unanswered question.

3.1.1.1 The mechanism of protein folding

The mechanism by which a protein could, even in principle, fold to a specific structure was until very recently shrouded in mystery[1]. There is considerable evidence that the native state of a protein corresponds to the structure that is most stable under physiological conditions. Nevertheless, the total number of possible conformations of a polypeptide chain is so large that it would take an astronomical length of time to find this particular structure by means of a systematic search of all conformational space (the Levinthal Paradox).

Recent experimental and theoretical studies have, however, provided a resolution of this apparent paradox. It is now evident that the folding process does not involve a series of mandatory steps between specific partially folded states, but rather a stochastic search of the many conformations accessible to a polypeptide chain[1-4]. Because the correct (native-like) interactions between different residues are on average more stable than the incorrect (non-native) ones, such a search mechanism is in principle able to find the lowest energy structure[5]. Nevertheless, there is a lot of controversy on the possible role played by non-native interactions during the folding of proteins.

This stochastic description of protein folding is often referred to as the “new view” of this complex process[6]. It implies the existence of a particular “energy landscape” for any specific protein, describing the free energy of the polypeptide chain as a function of its

conformational properties (Figure 3.1). For small proteins, this landscape appears to be funnel-like and reflects the evolutionary selection of polypeptide sequences able to fold rapidly and reliably towards a unique native state. The very rapid and efficient search to the native state is encoded by a network of interactions between key residues that establishes the native topology in the transition state of the folding reaction.

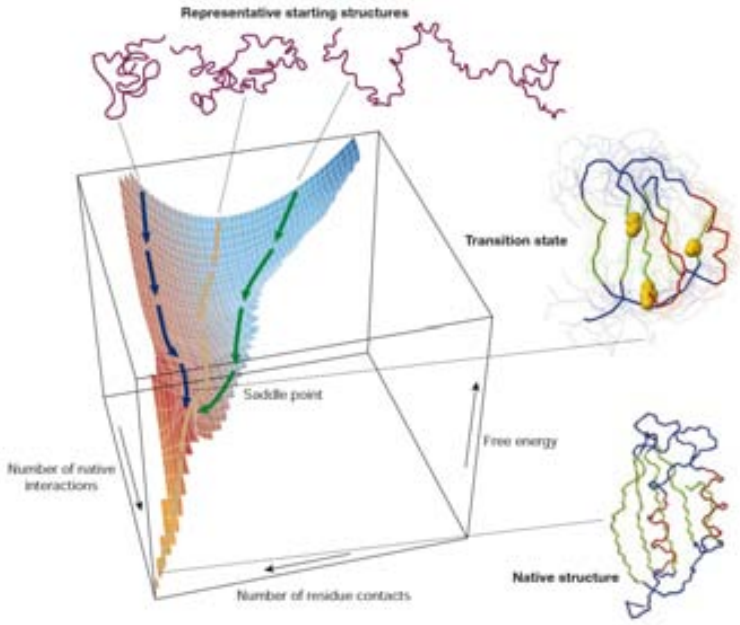

Figure 3.1. A schematic energy landscape for protein folding. The surface is derived from a computer simulation of the folding of a highly simplified model of a small protein. The surface 'funnels' the multitude of denatured conformations to the unique native structure. The critical region on a simple surface such as this one is the saddle point corresponding to the transition state, the barrier that all molecules must cross if they are to fold to the native state. The yellow spheres in this ensemble represent the three 'key residues' in the structure; when these residues have formed their native-like contacts the overall topology of the native fold is established.

For larger proteins, the energy landscapes become rougher, allowing the population of partially folded species that may be on- or off- pathway to the native fold.

3.1.1.2 Protein folding and misfolding in the cell

The crowding and collision of proteins in the living cell renders their *in vivo* energy landscapes even more complex, because the competition between intramolecular (folding) and intermolecular (aggregation) interactions results in a dramatic increase in landscape ruggedness. In the Figure 3.2, it is depicted the complexity of the protein folding and aggregation energy landscapes under these circumstances.

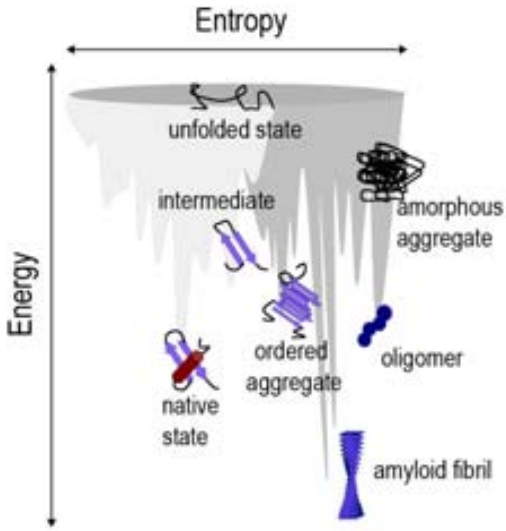

Figure 3.2. Illustration of a combined energy landscape for protein folding and aggregation. The surface illustrates the roughness of the protein energy landscape, showing the multitude of conformational states available to a polypeptide chain. While rather simple folding funnels (light grey) can describe the conformational search of a single polypeptide chain to a functional monomer, intermolecular protein association dramatically increases ruggedness (dark grey).

3.1.2 Protein aggregation and amyloid fibril formation

In order to prevent the inappropriate interaction of incompletely folded proteins with other molecules in the cellular environment, living systems have evolved several strategies that control protein folding such as molecular chaperones or the quality control system at the endoplasmatic reticulum[7].

Specific types of cellular activity are generated and abolished by folding and unfolding events. Moreover, folding and unfolding are associated to many essential processes as translocation across membranes, trafficking secretion, the immune response and regulation of the cell cycle[8]. Thus, failure to achieve a correct folding or to remain correctly folded will cause a malfunctioning of living systems and consequently to disease.

In the last few years, protein aggregation has emerged from a neglected area of protein chemistry as a transcendental issue in biological and medical sciences[9]. In this regard, an increasing body of evidence points out at the anomalous misassembly of proteins into insoluble amyloid deposits as the fundamental cause behind some debilitating human disorders of growing incidence such as Alzheimer's disease (AD), Parkinson's disease (PD), type II diabetes, the transmissible spongiform encephalopathies and many others[10]. A common trait of these disorders is that the aggregated protein deposits in internal organs and interferes with normal cellular function, sometimes lethally[11].

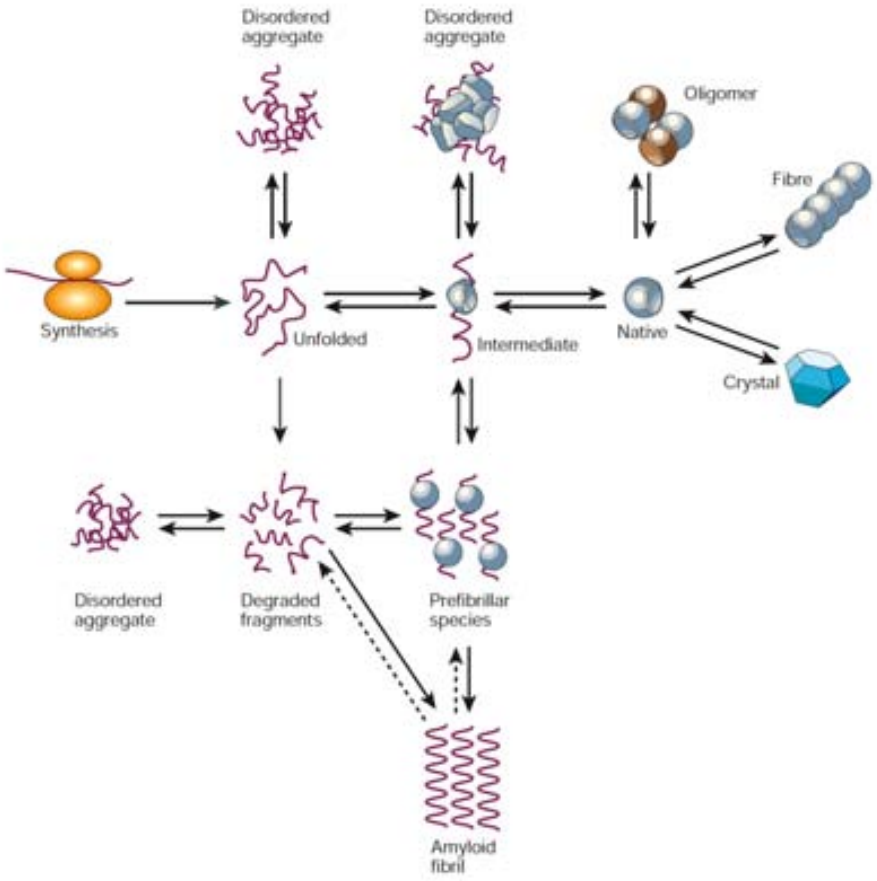

Figure 3.3 A unified view of some of the types of structure that can be formed by polypeptide chains. An unstructured chain, for example synthesized on a ribosome, can fold to a native structure. It can, however, experience other fates such as degradation or aggregation. An amyloid fibril is just one form of aggregate. Other assemblies, including functional oligomers, macromolecular complexes and natural protein fibers, contain natively folded molecules. The populations and interconversions of the various states are determined by their relative thermodynamic and kinetic stabilities under any given conditions[7].

Each amyloid disease is related predominantly with the aggregation of a specific protein, although some other components, including proteins and carbohydrates are also incorporated into the *in vivo* deposits. The quantity of aggregates can range from being nearly undetectable to weight some kilos in only one organ[12]. The soluble forms of the proteins involved in the well-defined amyloidosis are highly dissimilar without any relationship between their sequence, structure or size and comprising from globular proteins to unstructured peptides. However, their aggregated forms have many characteristics in common[13].

The amyloid deposits all show specific optical behavior (such as birefringence) when their are bind to certain dyes like Congo Red or Thioflavine-T. The morphology of the amyloid fibrils can be easily observed in the electronic microscope: the fibrils have usually a 7-12 nm of diameter and are formed by protofilaments (0.2-0.35 nm) that are

intimately associated. By X-ray diffraction, all the amyloid fibrils show a typical 'cross β-sheet' diffraction pattern. It reveals that the organized core structure is composed of β-sheets whose strands run perpendicularly to the fibril axis.

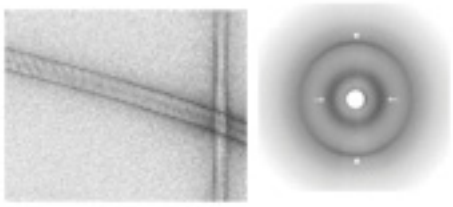

Figure 3.4 Electron micrograph of amyloid fibers (left). X-ray diffraction pattern of fibers. Arrowheads indicate the 4.63-Å reflection; arrows indicate the 10.11-Å reflection (right).

The ability of polypeptide chains to form amyloid structures is not restricted to the relatively small number of proteins associated with recognized clinical disorders, and it now seems to be a generic feature of polypeptide chains[14]. The most compelling evidence for the latter statement is that fibrils can be formed *in vitro* by many other peptides and proteins, including such well-known molecules as myoglobin (a globular protein rich in α-helix structure), and also by homopolymers such as polythreonine or polylysine under the appropriate conditions[14, 15].

Although the ability to from fibrils seems generic, the propensity to do so under certain circumstances can vary markedly between different sequences. Mutational studies of the kinetics of aggregation of full-length proteins revealed simple correlations between aggregation and physico-chemical properties such as β-sheet propensity, hydrophobicity and charge[16, 17].

3.1.2.1 The mechanism of amyloid fibril formation

Independently of the native structure of the protein, all amyloid fibrils show a cross β-structure. This fact implies that fibril formation supposes a strong conformational reorganization of the polypeptides as well as the establishment of new non-covalent interactions. These changes are not possible in the native conformation of many proteins: the polypeptide main chain and the hydrophobic side chains are largely buried within the folded structure. Only when the native conformation is destabilized (e.g. by the addition of denaturants, low pH, high temperature, truncations or mutations), the protein aggregation will be possible. Specific protein destabilization results in an increased population of partially folded conformations, whose exposed aggregation-prone regions enhance the

probability of intermolecular interactions. The factors that cause destabilization of native state are now becoming clear for at least some proteins: mutations, an increase in the concentration of a key amyloid precursor or a change in the local environment or alterations in the protein concentration.

Several models for amyloid fibril formation have been suggested based on monitoring fibrillogenesis using microscopy, spectroscopic techniques or the binding of dyes. Recent *in vitro* studies using electron microscopy and atomic force microscopy have identified and characterized several intermediate structures populated during fibril formation, including small oligomers, membrane embedded pores, and protofibrils, the latter having a characteristic 'beaded' appearance[18, 19].

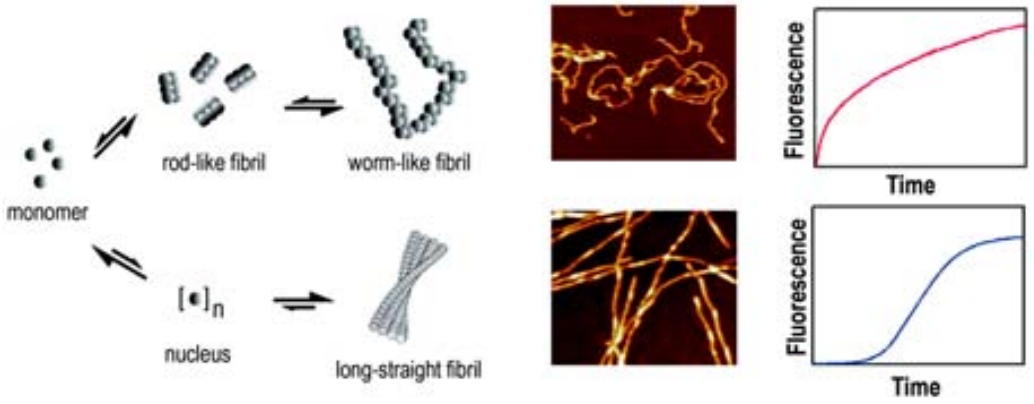

Figure 3.5 Pathway complexity of β_{2m} amyloid fibril formation. Proposed model for competing pathways that lead to the formation of worm-like fibrils or long-straight amyloid fibrils. AFM images of worm-like and long-straight fibrils are also shown. Worm-like fibrils are formed with nucleation-independent kinetics, whilst the formation of long-straight fibrils is nucleation-dependent and shows a clear lag-phase[20].

Classically, amyloid fibrils are formed in a nucleation-dependent manner: the protein monomer is converted into fibrillar structure via a transiently populated aggregation nucleus. After the rate-limiting step of nucleus formation, the aggregate growth proceeds rapidly by further addition of monomers or other assembly-competent species. The formation of the nucleus is thermodynamically unfavorable, whereas its elongation is highly favorable and proceeds quickly to the ultimate fibril structure (Figure 3.5).

By contrast, the assembly of spherical oligomers and other prefibrillar forms occurs with nucleation-independent kinetics, and results in the formation of spherical particles or worm-like fibrils[21-23]. Here, polymerization proceeds in the absence of a lag-phase. In some cases, oligomeric species have been suggested to be direct precursors of long-straight amyloid fibrils, whilst in other cases, an off-pathway role has been proposed[21, 24-26]. More

experiments have to be performed in order to clarify the exact mechanism of the fibril formation and acquire a deep knowledge of how environmental factors (such as pH) affect the fibril formation pathway.

3.1.3 Aggregation in bacterial cells: Inclusion Bodies (IBs)

Misfolded polypeptides accumulate in the cells due to malfunctioning or overloading of the protein quality control pathways or of the components of the degradative pathway[27]. The recombinant expression of proteins tends to saturate the cell folding machinery, resulting frequently in the mistargeting of the heterologous polypeptide into aggregates that accumulate as cytoplasmic aggresomes in eukaryotes[28] or inclusion bodies (IBs) in prokaryotes[29].

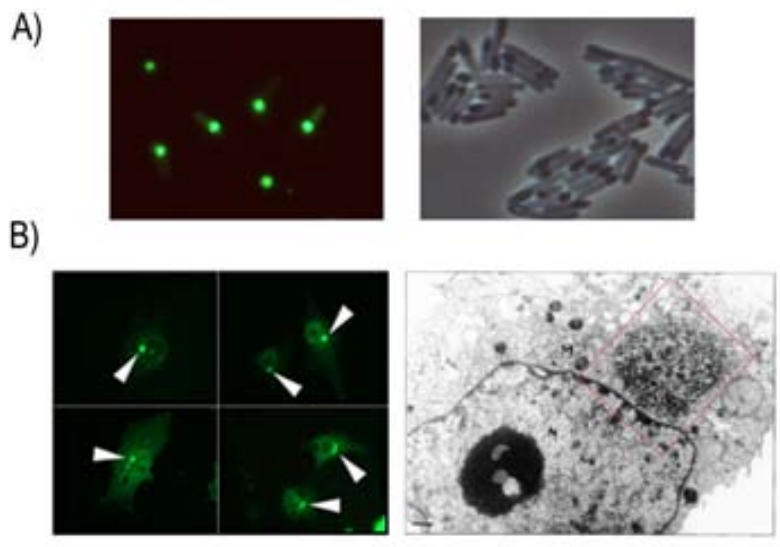

Figure 3.6 Inclusion Bodies and aggresomes. A) Fluorescence image of *E.coli* cells expressing an aggregation–prone peptide (Alzheimer's peptide Aβ42) fused to GFP (left). Inclusion Bodies can be detected inside the bacteria in phase contrast image (right) B) Aggresomes (indicated by white arrowheads) formed due to the expression of an aggregating-prone protein in COS cells. They are detected by immunofluorescence (left)[30]. By electron microscopy, one aggresome is detected close to the nucleus[31] (M mitochondria, N, nucleus) Bar, 1μm (right).

Bacteria (and specifically *Escherichia coli*) are widely used as factories for the production of recombinant polypeptides that do not require post-translational modifications, such as glycosilation, to attain their native and bioactive conformation. *E.coli* cells grow rapidly to high cell density in inexpensive substrates and offer inducible protein expression to extremely high levels.

However, its intrinsic propensity to accumulate heterologous products in IBs represents a major challenge in downstream bioprocessing. For long time, biotechnology efforts in protein production have focused on strategies aimed to increase protein solubility and reduce IBs formation. Nevertheless, the formation of IBs is often unavoidable. In some large structural genomics projects, up to half of the tested targets failed to fold properly but instead accumulate as insoluble protein.

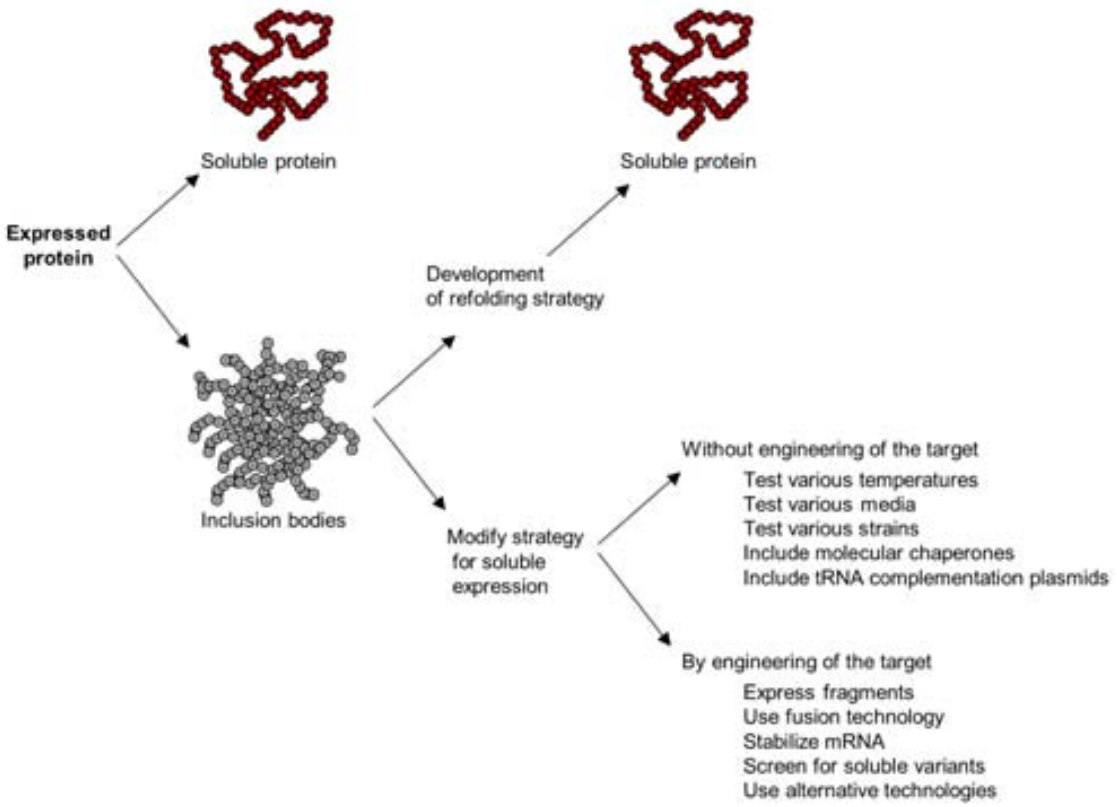

Figure 3.7 Strategies in order to overcome the aggregation of a recombinant protein in *E.coli*[32].

3.1.3.1 Internal composition and structure of IBs

Using phase contrast microscopy, IBs can be seen as refractile particles with a diameter of 0.5-1.3 μm inside the cytoplasm of bacteria[33] or in the periplasmic space, if the protein is secreted[33]. In general, mature IBs contain very little host protein; on many occasions, the overexpressed protein accounts for more than 90% of the polypeptides embedded in the aggregates[34]. The rest of material in the IBs is likely to be ribosomal components, proteolytic fragments of the recombinant protein, traces of hydrophobic membrane proteins[35] or phospholipids and DNA/RNA fragments[36]. These elements might have been trapped in the IBs by way of non-specific intermolecular interactions during the aggregation of the target protein or may have co-purified the aggregates under low-stringency conditions. In contrast, components of the cellular protein quality machinery,

including the small heat-shock proteins IbpA and IbpB and the main chaperones DnaK and GroEL, have been shown to specifically associate with IBs[37].

3.1.3.2 New perspectives of the IBs study

Traditionally, little attention has been paid to the structural and functional characteristics of these intracellular aggregates. They have been considered as non-structured and inactive mass of proteins compared to dust balls within the cells. Nevertheless, in the last years, this classical view has changed thanks to experiments that have provided new data about their structure, activity and mechanism of formation.

3.1.3.2.1 Specificity during IBs formation

Due to the lack of any noticeable pattern (sequence, structure, size or origin) between the numerous proteins able to form IBs inside prokaryotic cells, their formation has been long considered to be driven simply by the establishment of non-specific intermolecular contacts between nascent, partially-folded species. However, nowadays there are some experiments that demonstrate *in vitro* the protein-specific assembly. For example, it has been shown that the aggregation of tryptophanase depends on its own concentration and it is not affected by the presence of foreign proteins[38]. This specificity has also been observed in the *in vitro* aggregation of a mixture of folding intermediates from P22 coat and tailspike proteins, two polypeptides that preferentially self-associate in the test tube[39]. Besides, it has been demonstrated that purified IBs are able to capture homologous soluble proteins in a dose-dependent manner[40]. Moreover, this incorporation is highly specific: a particular IB does not recognize soluble heterologous polypeptide chains, suggesting that this process is directed by selective interactions between the soluble protein chain and its IBs. In addition to the sequence, the conformation of the folding intermediates must be central to the establishment of such intermolecular contacts because IBs are not able to capture properly folded molecules from the solution[40].

Less information is available about the determinants of protein aggregation in the highly complex intracellular environment, where many cellular factors might alter dramatically the aggregation behavior of polypeptides[41]. These include the high protein concentration, molecular crowding effects, the action of molecular chaperones and proteases and the continuous supply of unfolded or partially folded polypeptides by the ongoing translational machinery. This difference between *in vitro* and *in vivo* experimental

conditions constitutes an important barrier for the molecular understanding of protein aggregation in living organisms[42]. Kopito and coworkers have demonstrated that the formation of cytoplasmic protein aggregates in animal cells exhibits a remarkable specificity[43]. We sought to analyze if the same selectivity applies in the *in vivo* formation of bacterial IBs where protein accumulation is likely to be a diffusion-limited process not mediated by cellular substructures in contrast to eukaryotic backgrounds[44].

3.1.3.2.2 Amyloid like properties of IBs

Analyzing in depth IBs structure and function, there are some characteristics shared with amyloid aggregates (summarized in Table 3.1).

Table 3.1 Main functional and structural traits of bacterial IBs resembling those of amyloids.

Feature
High purity of the aggregate[33]
Aggregation mainly from folding intermediates[40, 45]
Sequence-specific aggregation[39, 40]
Chaperon-modulated aggregation[46, 47]
Seeding-driven aggregation[48]
Aggregation propensities strongly affected by point mutations[49-53]
Reduced aggregation by stabilization of the native structure[54, 55]
Enrichment of intermolecular β structure[40, 56]
Fibril-like organization (of soluble protein aggregates)[57]
Amyloid dye binding[40]
Enhanced proteolytic resistance (of a fraction of IB protein species[58, 59]

The observed selectivity in IB formation is reminiscent of the behavior of amyloid aggregates. During a long time, intracellular aggregates had been thought to be amorphous precipitates devoid of any molecular structure, quite the opposite to the almost crystalline packing of polypeptides in extra cellular amyloid fibrils[60]. However, recent studies run against this view by demonstrating the recurrent presence of intermolecular β-sheet secondary structure in IBs[61, 62]. This feature is independent of the protein's native structure and it can be measured by Fourier-transform infrared spectroscopy (FT-IR)[61, 63, 64]. The supramolecular organization of the extended β-strands inside the aggregates is still unknown, but the available data suggest that they are tightly packed with short hydrogen bonds, a hallmark of amyloid fibrils. The recurring presence of intermolecular β-structure

inside IBs explains their ability to bind amyloid diagnostic dyes such us Congo Red and Thioflavin-T[57, 62]. The resemblance in the inner organization of both types of aggregates opens an intriguing possibility: the existence of amyolid-like polypeptidic conformations inside bacterial cytoplasmatic aggregates.

3.1.3.2.3 Protein activity inside IBs

FT-IR structural analysis of α-helical proteins, such as interleukin-1β, demonstrated that their IBs contain, in addition to the characteristic intermolecular β-structure, a significant amount of α-helix secondary conformation[65]. How these two kinds of structures are organized inside the aggregate is still not known. In any case, the existence of partial native conformation inside IBs could explain the detection of protein activity inside IBs. In this sense, IBs of different proteins as DHFR, fluorescent proteins or β-galactosidase have been shown to present activity[66].This fact implies that aggregation of recombinant proteins in bacterial IBs does not necessarily completely inactivate them and opens the possibility to consider the application of IBs as biocatalysts.

3.1.4 Protein aggregates inside the cell

3.1.4.1 Protein quality

IBs have been considered during a long time as reservoirs of inactive protein. However, the new consideration of IBs as particles for industrial catalysis might lead to rethinking many biotechnological strategies. While most protein production pipelines had aimed to increase the amount of soluble protein in the cell, little was known about the protein quality in this fraction: the proteins were usually assumed to be fully functional in this fraction.

Recently, it has been shown that, in fact, there are a huge variety of protein conformations in the soluble fraction (including "soluble aggregates")[67] with different level of functionality[57, 68]. Therefore, the soluble fraction may contain some inactive or partially inactive protein forms in addition to the active well-folded molecules[69]. The results indicate that incorrect folding is not always paired with aggregation and correct folding cannot always be associated with solubility (Figure 3.8). Thus, solubility *per se* is not the best reporter of protein quality during protein production, because the presence of active polypeptides in IBs and inactive protein forms in the soluble cell fraction result in very similar specific activities. Due to the fact that the distribution of protein conformers in

the soluble and insoluble fractions is considerably influenced by the production conditions and/or the host cell genetic backgrounds[70], the distinction between conformational quality and solubility can be exploited in new strategies aimed to optimize quality rather than quantity during protein production.

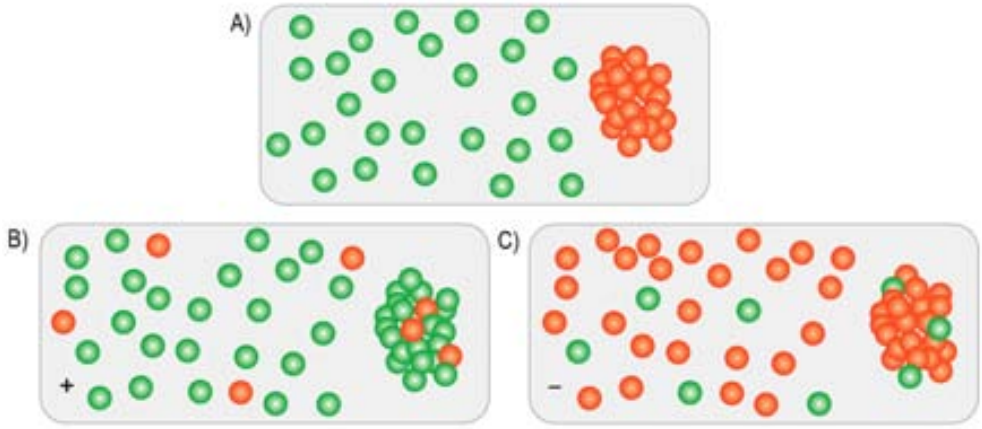

Figure 3.8 The functionality of recombinant proteins aggregated as IBs. A) The classical model of recombinant protein fractioning, where functional species (green spheres) occur in the soluble cell fraction, whereas inactive, misfolded polypeptides (red spheres) are deposited in IBs. With the new perspective, the average conformational quality of IB protein is representative of that found in the whole cell either under favorable (B), or unfavorable (C) conditions for the protein folding. In these cases, the biological activity of recombinant protein ranges from no detectable activity to 100% specific activity, and misfolded proteins might exist along a continuum of forms, from fully soluble versions to aggregated ones[70].

3.1.4.2 Dynamic equilibrium with chaperones

It has been shown that there is an unbalanced, highly dynamic equilibrium between protein deposition and removal involving a continuous exchange of polypeptides between the soluble and insoluble forms of recombinant proteins[71]. In this way, if protein synthesis becomes interrupted, cytoplasmic IBs are almost totally dispersed within a few hours, because the arrest of new protein translation makes larger ratios of chaperones and foldases available to refold the precipitated protein in IBs[72]. This observation suggests that IBs may act as cellular protein reservoir from where recombinant protein can be extracted.

The sophisticated cell quality control system is involved in this protein flux through disaggregation, unfolding and polypeptide reactivation[68]. According to their mode of action, the different elements of the cell quality machinery can be classified as unfolders, folders and holders[73]. Folder chaperones, mainly the DnaK/Hsp70 family and GroEL have the ability to refold misfolded or aggregated proteins. Sometimes folder chaperones require a previously unfolded substrate to act, in which case they work together with unfolder chaperones, principally ATPase-associated chaperones. Finally, the holder activity protects

polypeptides against aggregation. This function is executed by small Hsps, which are able to bind aggregates and folding intermediates and isolate them from the crowded environment.

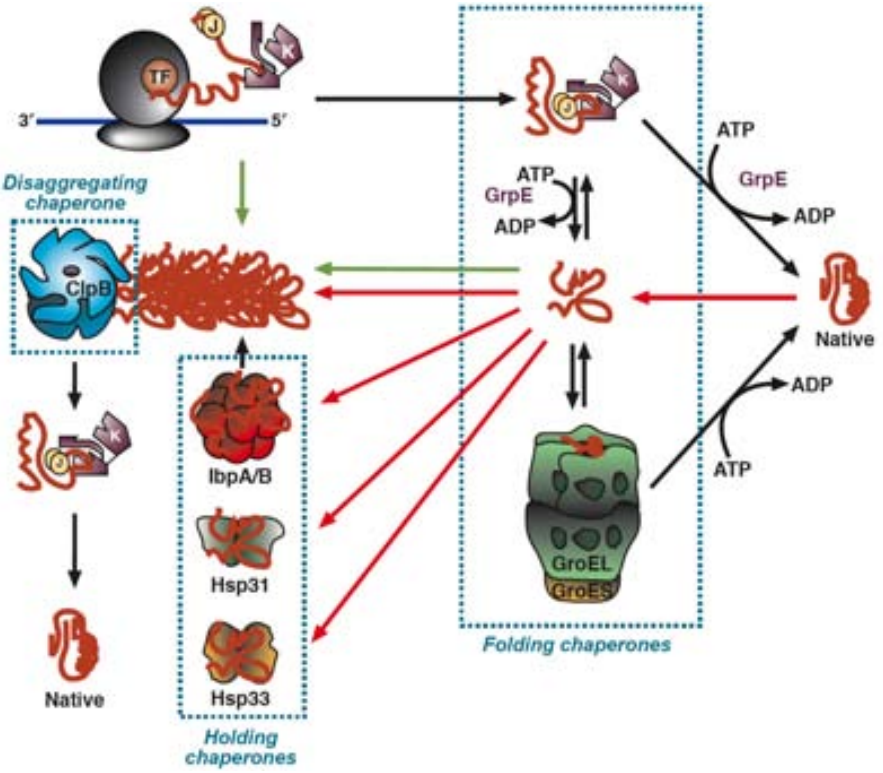

Figure 3.9 Chaperone-assisted protein folding in the cytoplasm of *E.coli*. Nascent polypeptides first encounter TF or DnaK-DnaJ. Both chaperones engage hydrophobic solvent-exposed stretches. After undocking from TF- or GrpE-mediated release from DnaK, folding intermediate may reach a native conformation, cycle back to DnaK-DnaJ or be transferred to the central chamber of GroEL for folding upon GroES capping. In times of stress (red arrows), thermolabile proteins unfold and aggregate. IbpB binds partially folded proteins on its surface until folding chaperones become available and intercalates within large aggregates. The holding chaperones become important under thermal stress. ClpB promotes the shearing and disaggregation of thermally unfolded host proteins and cooperates with DnaK-DnaJ-GrpE to reactivate them once stress has abated. Recombinant proteins that miss an early interaction with TF or DnaK/DnaJ, that undergo multiple cycles of abortive interactions with folding chaperones or titrate them out, accumulate in IBs (green arrows)[74].

All these activities have a remodeling effect on the structure and composition of IBs. In this sense, it has been suggested that the removal of misfolded protein from the surface of IBs results in a progressive enrichment of the native-like and active protein in the IB core[75]. Besides, when the expression of chaperones is knockout, in general, a decrease in the amount of soluble protein and an increase in the β-sheet compactness inside IBs are detected[69].

3.2 Objectives

The main objective is the study of the common properties between amyloid aggregates and inclusion bodies. Specifically two aspects will be investigated:

→ Specificity in inclusion bodies formation

→ Inner Structure of inclusion bodies

3.3 Experimental procedures

3.3.1 Strain, plasmids, culture conditions

The plasmids encoding for Aβ42-GFP and Vp1-GFP were as previously described[66, 76]. DNA encoding Aβ42-BFP was obtained by site-directed mutagenesis of Aβ42-GFP using the QuickChange kit from Stratagene: forward and reverse primers were designed to introduce the Y66H and Y144F mutations in GFP. Aβ42-GFP was also cloned in pBAT4 (EMBL). All constructs were verified by DNA sequencing. The plasmid encoding for Vp1-GFP was a kind gift from Dr. Villaverde.

Table 3.2 Plasmids used in this study

Fusion Protein	Plasmid	Restriction sites	Antibiotic resistance
Vp1-GFP	pTRC99A	*Eco*RI, *Bam*HI	Ampicillin
Aβ42-BFP	pET28a(+)	*Nde*I, *Bam*HI	Kanamycin
	pQBI67	*Nhe*I	Ampicillin
Aβ42-GFP	pET28a(+)	*Nde*I, *Bam*HI	Kanamycin

3.3.2 IBs purification and denaturation

IBs were purified from cell extracts sixteen hours after the induction of gene expression, by detergent-based procedures as described[77]. Briefly, cells were harvested by centrifugation at 12000×*g* (at 4ºC) for 15 min and resuspended in 200 μl of lysis buffer (50 mM Tris-Cl pH=8, 1 mM EDTA, 100 mM NaCl), plus 30 μl of 100 mM protease inhibitor PMSF and 6 μl of 10 mg/ml lysozyme. After 30 min of incubation at 37ºC under gentle agitation, NP-40 was added at 1% (v/v) and the mixture incubated at 4ºC for 30 min. Then, 3 μl of DNase I (from a 1 mg/ml stock) and 3 μl of 1 M $MgSO_4$ were added and the resulting mixture was further incubated at 37ºC for 30 min. Protein aggregates were separated by centrifugation at 12000×*g* for 15 min at 4ºC. Finally, IBs were washed once with the same buffer containing 0.5% Triton X-100 and once with sterile PBS. After a final centrifugation at 12000x*g* for 15 min, pellets were stored at −80ºC until analysis.

To discard the presence of nucleic acids in IBs, they were digested with Dnase and

Rnase A (25 μg/ml) for 1 hour at 37ºC in the presence of $MgSO_4$ (10 mM). Afterwards, IBs were centrifuged and dissolved in PBS. No significant spectral, morphological or compositional differences between untreated and treated IBs could be detected.

Protein in IBs was quantified after overnight denaturation in 8 M urea by the Bradford assay. For kinetic experiments, 50 μl of a solution of purified and homogenized IBs was added to 950 μl of 10 mM Tris–HCl buffer (pH=7.5) containing 2.3 M GuHCl. The reaction was monitored by measuring the changes in A_{350nm} in a Cary-100 Varian spectrophotometer. Double-exponential decay curves were fitted to the data using Sigmaplot non-linear regression software (Jandel Scientific, San Rafael, CA, USA), and apparent rate constants were derived from these regressions[78].

3.3.3 Confocal microscopy analysis

Cells expressing the different constructs were fixed with 0.1% formaldehyde and stored at 4°C until observed. Photographs were taken using a Leica TCS SP2 confocal microscope (excitation wavelength 488nm (in the case of GFP) and excitation wavelength 409-468nm (in the case of BFP) and emission wavelength at 500 to 600 nm in both cases; optical lens magnification 63X; 1.024 by 1.024 pixels; zoom 4).

3.3.4 FRET analysis

3.3.4.1 By confocal microscopy

Cells were fixed and imaged as described in the general Experimental procedures section. The analysis of digitized microscope images allow selection of regions of interest (ROI) in the cells. To record BFP emission spectra, the BFP donor in the selected ROI was excited with both 351 nm and 364 nm laser lines and emission spectra were generated by scanning the ROI and recorded using Leica Confocal Software (LCS). To record GFP emission spectra, the GFP-fluorophore was excited by the 488 nm laser line. For GFP acceptor photobleaching experiments, cells were photobleached exiting at 488 nm. Fluorescence emission spectra were recorded before and after photobleaching following BFP excitation at 409-468 nm to assess changes in donor GFP fluorescence. FRET acceptor photobleaching software was used to document that FRET occurred by showing that the intensity of donor BFP fluorescence increased after the GFP acceptor was photobleached. FRET efficiency (E_{FRET}) was calculated using the formula:

$$E_{FRET} = \frac{I_{post} - I_{pre}}{I_{post}}$$

where I_{pre} and I_{post} were the fluorescence intensities of BFP before and after photobleaching, respectively. The percentage of FRET was measured in one hundred samples in each case (Aβ42-BFP+Aβ42-GFP or Aβ42-BFP+Vp1-GFP) and afterwards it was averaged and plotted.

3.3.4.2 By fluorescence spectrometry

After 18h from the protein expression induction, cells expressing different protein fusions were washed with PBS and diluted until A_{600nm}= 0.01. Their fluorescence emission was measured from 420-560 nm at 25°C using an excitation wavelength of 389 nm.

3.3.5 Transmission Electronic Microscopy

IBs (50 μg/ml) were digested using PK (20 μg/ml) at 37°C 50 mM Tris-HCl and 150 mM NaCl, pH=8.0. At different times of digestion, the samples were centrifuged, the insoluble part was resuspended in water, placed on carbon-coated copper grids, and left for five minutes. The grids were washed with distilled water and stained with 2% (w/v) uranyl acetate for another two minutes before analysis using a HitachiH-7000 transmission electron microscope operating at accelerating voltage of 75 kV.

3.3.6 Thioflavin-T Binding

Thioflavin-T binding assays were carried out using aliquots of 100 μl drawn from 100 μg/ml soluble or IBs protein samples. These aliquots were diluted into buffer (10 mM sodium phosphate (pH=7.0, 150 mM NaCl) containing 65 μM Th-T, and adjusted to a final volume of 1 ml. PK was added in a final concentration of 1 μg/ml. Fluorescence emission spectra were recorded using an excitation wavelength fixed at 440 nm. Apertures of 10 nm were fixed in both excitation and emission slits.

3.3.7 FT-IR spectroscopy

The Aβ42-GFP fibrillar core was analyzed using a Bruker Tensor 27 FT-IR Spectrometer (Bruker Optics Inc) with a Golden Gate MKII ATR accessory. Each spectrum consists of 20 independent scans, measured at a spectral resolution of 2 cm^{-1} within the 1800-1500 cm^{-1} range. All spectral data were acquired and normalized using the

OPUS MIR Tensor 27 software. Second derivatives of the spectra were used to determine the frequencies at which the different spectral components were located.

3.3.8 Atomic Force Microscopy

3.3.8.1 AFM imaging

Imaging was performed with a commercial MultiMode atomic force microscope controlled by Nanoscope IV electronics (Digital Instruments, Santa Barbara, CA), equipped with either a 12-µm scanner (E-scanner) or a 120-µm scanner (J-scanner). All images were taken in liquid using a tapping mode liquid cell without the O-ring seal. Oxide-sharpened pyramidal Si_3N_4 tips mounted on triangular 100-µm-long cantilevers (k = 0.08 N/m) were purchased from Olympus (Tokyo, Japan). The liquid cell and the tip were cleaned with ethanol and thoroughly rinsed with deionized water before use. For imaging, the microscope head was placed on a vibration-isolated plate. 20 µl of the sample were allowed to adsorb for about 15 min at room temperature on highly ordered pyrolytic graphite (Nt-MDT Co., Zelenograd, Moscow, Russia) and finally overlaid with 100 µl of incubation buffer.

3.3.8.2 Force spectroscopy

The resistance of IBs to indentation was measured by acquiring force curves with the AFM in force spectroscopy experiments. A commercial AFM (MFP-3D, Asylum Research, Santa Barbara, CA) integrated with a Nikon Eclipse TE 2000-S inverted fluorescence microscope was used to study the rupture forces and the fluorescence visualization of IBs. Tapping mode imaging and force measurements were performed in PK digestion buffer solution using the tips described above. Prior to each experiment, 20 µl of the sample were allowed to adsorb for about 15 minutes at room temperature on poly (methyl methacrylate) (PMMA). All indentations were made using a displacement rate of 2 µm/s.

3.3.9 Mass spectrometry analysis of the IBs proteolysis

IBs of Aβ42-GFP were digested as described above for 20 hours. The reaction mixture was centrifuged for 10 minutes at 10000x*g*, the pellet was resuspended in GuHCl 6 M and left for 12 hours. Then it is diluted ten times and analyzed by MALDI-TOF. The sample was mixed with an equal volume of the matrix: 2',6'-Dihydroxyacetophenone

(DHAP) (Fluka). This matrix was prepared in 30% acetonitrile in 20 mM of ammonium citrate. The samples were loaded on MALDI plates by the dried droplet method. The calibration standards were from Bruker and the MALDI-TOF mass spectrometer was a Bruker Ultraflex.

3.3.10 Analysis of *in vivo* Aβ42-GFP oligomerization

For the analysis of Aβ42-GFP oligomers formation, cells were collected at different time points after induction of protein expression and analyzed by Western blot. The protein content of cellular fractions was resolved on 10% SDS-PAGE gels, transferred on to PVDF membranes, and recombinant proteins detected with a polyclonal anti-GFP antibody. The membranes were developed with the ECL method. The existence of strong non covalent intermolecular contacts stabilizing the oligomers, was demonstrated by boiling the cellular fractions for 10 min in the presence of a reducing agent (15% β-mercaptoetanol) and SDS (9%) before electrophoretic analysis.

3.3.11 Seeding of Aβ42 amyloid fibrils

A stock solution of 100 μM of peptide Aβ42 was prepared in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) and the solution was incubated at room temperature for 10 min. The solution was divided into aliquots of 100 μl and the HFIP was removed by evaporation under a gentle stream of nitrogen, leaving a slightly yellow film. Afterwards, they were stored at -20°C for later use. Before the experiment, each aliquot was resuspended in 50 μl of dimethyl sulfoxide (DMSO) and sonicated for 15 min. Sonication was critical to remove any traces of undissolved seeds that may resist solubilization. These aliquots of Aβ42 stock were dissolved in 750 μl of 10mM Tris HCl pH 7.4. Then, Th-T was dissolved until a final concentration of 25 μM. In the control (aggregation of Aβ42 mononers), Tris-HCl was added until a final concentration of 1ml. In the seeding assay, a solution Aβ42 only IBs (to a final OD_{350}=0.125) was also added. As a positive control, one reaction was performed under the presence of preformed Aβ42 fibers (10% w/w). The fluorescence emission at 480 nm (exciting at 445nm) was measured at 1 min intervals in a fluorescence spectrophotometer using quartz microcuvettes placed in a thermostated cell holder at 37°C. The solutions were stirred continuously. Absorbance values were transformed into fraction of fibrillar form in the system as previously described[79].

3.4 Results

3.4.1 *In vivo* protein aggregation as IBs displays remarkable specificity

In order to evaluate if the specificity in protein aggregation was also applied in the far more complex cellular environment, two aggregation-prone polypeptides were tagged with different fluorescent proteins and were co-expressed in *E. coli*. In particular, the Alzheimer-related peptide Aβ42 was fused to the Blue and Green Fluorescent Proteins (Aβ42-BFP and Aβ42-GFP) and the foot-and-mouth disease virus VP1 capsid protein (Vp1), to GFP (Vp1-GFP). All these protein fusions accumulated mainly as cytoplasmic IBs when they were expressed in *E. coli* and thus, they could be tracked due to their fluorescence[80].

First, we used fluorescence resonance energy transfer (FRET). FRET consists in the energy transfer from one fluorescent molecule to another via dipole–dipole coupling[81] as it has been explained in the first chapter. This phenomenon only occurs when the two fluorophores are within 10-100 Å of each other, turning it into a high-resolution approach to interrogate protein vicinity. In our experiment, equal amount of cells expressing different protein fusions were exposed to the BFP excitation wavelength (389 nm) and their fluorescence emission was recorded above 420 nm to detect simultaneously BFP and GFP signals.

Control cells expressing Aβ42-GFP or VP1-GFP exhibited little GFP emission (510 nm) when excited at 389 nm, as a result of the reduced excitation of the GFP fluorophore at this wavelength[82]. In contrast, cells expressing Aβ42-BFP fusion displayed a strong, single emission maximum at 445 nm corresponding to BFP signal (Figure 3.10). When GFP and BFP were co-expressed in the same cell, the resulting spectra depended on the identity of the aggregation-prone peptides fused to them. Sensitized emission of GFP at 510 nm was detected in cells co-expressing Aβ42-GFP+Aβ42-BFP and in cells co-expressing Vp1-GFP+Aβ42-BFP when they were excited at 389 nm; but BFP emission decreased only when Aβ42-GFP and Aβ42-BFP were expressed in the same cell. The ratio between emission at the donor and acceptor wavelengths reflected both quenching of donor fluorescence and sensitized emission by the acceptor. Thus, we used the ratio of GFP fluorescence (at 510 nm) to BFP fluorescence (at 445 nm) to quantify the degree of FRET: 1.5 and 0.9 ratios were obtained for cells expressing the fluorescent proteins fused

to homologous and heterologous aggregation-prone peptides, respectively. This indicated a higher degree of energy transfer and suggested a higher proximity between the two fluorophores in cells co-expressing Aβ42-GFP+Aβ42-BFP than in cells co-expressing Vp1-GFP+Aβ42-BFP. As a negative control, we mixed equal amounts of cells expressing Aβ42-GFP and cells expressing Aβ42-BFP and the mixture was excited at 389 nm. The resulting spectrum was very similar to that of cells containing only the blue fluorophore, indicating the absence of significant FRET because of the large distance between cells expressing the different fluorescent proteins.

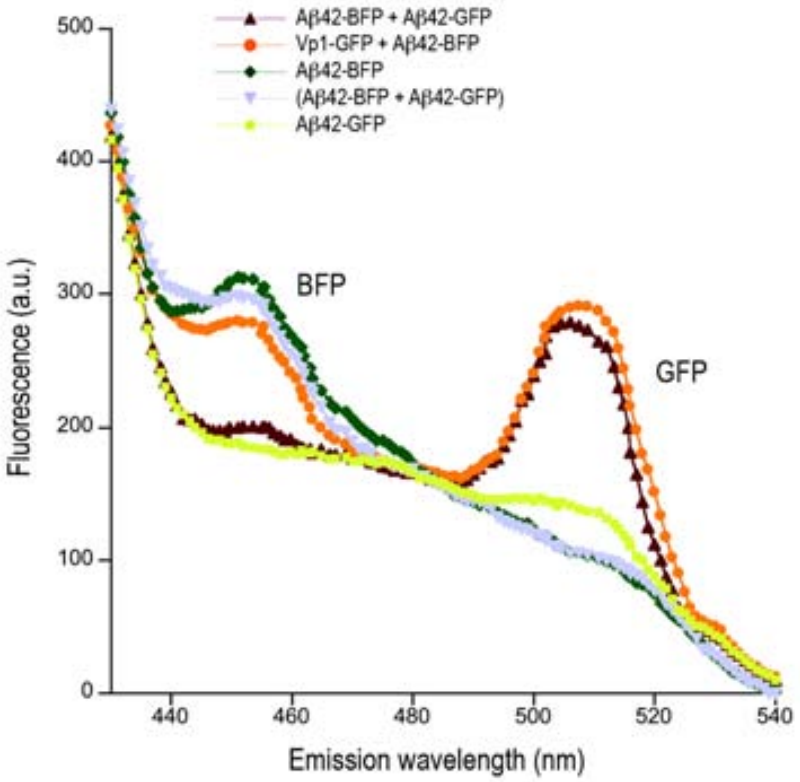

Figure 3.10 FRET efficiency between BFP and GFP measured in live *E.coli* cells by recording the fluorescence spectra between 400 and 540 nm after excitation of BFP. The fluorescence emission of BFP and GFP are indicated. The expressed protein fusions are specified in the legend. Brackets signify an equal mixture of cells expressing the different fusions.

We wondered if the observed differences in FRET signals reflected distinct biophysical states within the structural organization of the aggregates formed by the different proteins. To address this issue, cells co-expressing homologous and heterologous aggregation-prone peptides fused to BFP and GFP were visualized using confocal microscopy (Figure 3.11). Phase contrast visualization showed that the two types of aggregates displayed different morphology. In cells co-expressing Aβ42-GFP and Aβ42-BFP, IBs appeared as the usual round shaped refractile structures at the poles of the cells.

In contrast, IBs in cells co-expressing Vp1-GFP and Aβ42-BFP displayed an anomalous segmented morphology with the poles being denser than the inner part. Fluorescence imaging indicated that this different morphology was related to a non-homogenous distribution of the tagged aggregation-prone proteins in the IBs. In cells co-expressing Vp1-GFP and Aβ42-BFP fusions, the green fluorescence signal was concentrated at the poles, coincident with the dense regions observed by phase contrast, whereas the blue fluorescence was distributed homogenously in the IB. This means that the external part of IBs was selectively enriched with aggregated Vp1-GFP protein. No segregation of fluorescent tags could be observed in the IBs of cells co-expressing Aβ42-GFP and Aβ42-BFP fusions: they exhibited homogeneous and overlapping fluorescence signals.

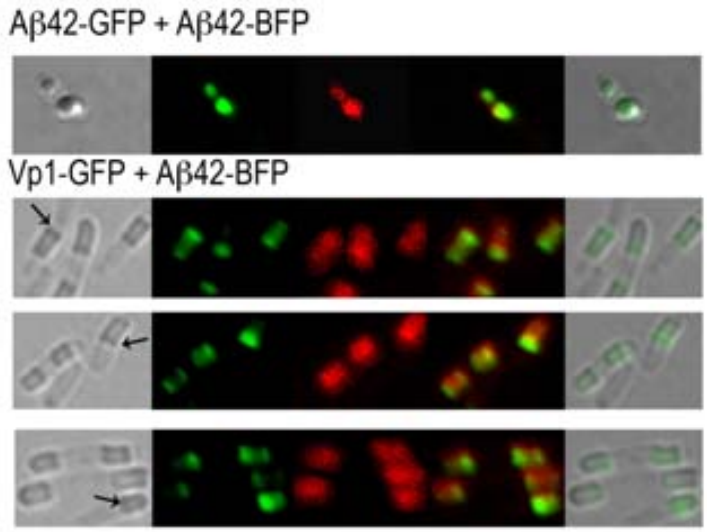

Figure 3.11 Confocal fluorescence imaging of cells expressing Aβ42-BFP+Aβ42-GFP or Aβ42-BFP+Vp1-GFP. BFP signal is displayed in red for better visualization. The images from left to right correspond to phase, GFP channel, BFP channel, merged images of BFP and GFP channels and merged images of GFP channel and phase, respectively. IBs could be observed inside the cells as refractile structures. When Aβ42-BFP and Vp1-GFP were co-expressed, the green fluorescence was more intense in the poles of the IBs coinciding with a higher density region pointed by arrows in the phase contrast image.

To corroborate these results, FRET was directly quantified over the confocal microscopy samples. A widely used approach for FRET quantification is measuring the increase in the fluorescence intensity of the donor after acceptor photobleaching. Because the excitation light used to bleach the acceptor usually does not bleach the donor efficiently, the FRET quantification by this approach is less affected by cross-talk. Energy transfer efficiency depends not only on the distance between the donor and acceptor but also on the relative orientation of their transitional dipole moments and thus aggregation

might, in principle, affect the FRET signal. Nevertheless, recently Petersen and co-workers have demonstrated that intermolecular energy transfer efficiency depends only slightly on the degree of aggregation[83], validating our approach.

In our samples, the efficiency of FRET was measured in IBs regions in which green and blue fluorescence apparently co-localized (100 independent measurements for each case). As it is plotted in Figure 3.12, the FRET percentage was nearly three fold higher in the aggregates of cells co-expressing Aβ42 fused to both fluorescent proteins than the one measured in cells co-expressing GFP and BFP fused to different peptides. Hence, although no absolute compartmentalization of the two fluorescent proteins was detected, the observed decrease in FRET efficiency would reflect a larger mean distance between the two proteins (Aβ42–BFP and Vp1-GFP) inside IBs.

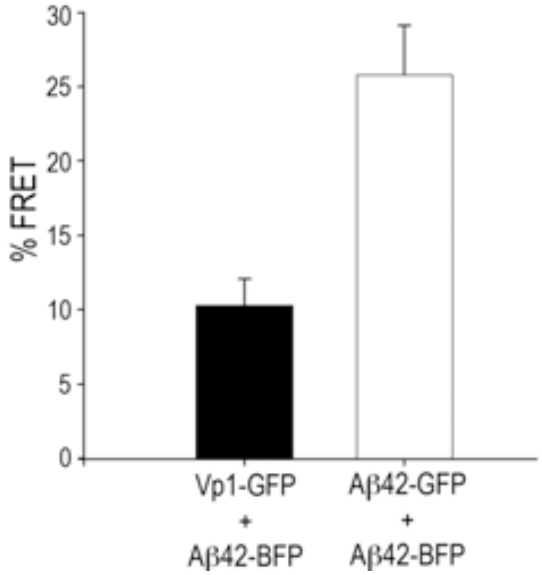

Figure 3.12 FRET efficiency measured over confocal microscopy samples of cells expressing Aβ42-BFP/Aβ42-GFP or Aβ42-BFP/Vp1-GFP.

3.4.2 Intracellular aggregates display differential stability

If, as deduced from the data presented above, the molecular interactions in Vp1-GFP+Aβ42-BFP IBs were different from those in Aβ42-GFP+Aβ42-BFP IBs, the two types of aggregates should exhibit different stability. One method to determine IBs stability is the kinetic study of their solubilization by a fixed concentration of a chaotropic agent[78] like guanidinium hydrochloride (GuHCl). Specifically, IBs were incubated in 2.3 M GuHCl and the change in the turbidity signal (A_{350}) was monitored in a time-course manner (Figure 3.13). The data could be fitted to a double-exponential decay equation with good accuracy (R>0.99) allowing the calculation of the apparent rate constants of the fast

phase. Differences in the velocity of solubilization could be observed between both samples with the following fast rate constants: 0.113±0.001min^{-1} (Aβ42-GFP+Aβ42-BFP) and 0.127±0.001min^{-1} (Aβ42-BFP+Vp1-GFP). In addition, the turbidity values at which the reaction stopped were significantly different: 0.11 and 0.23, respectively.

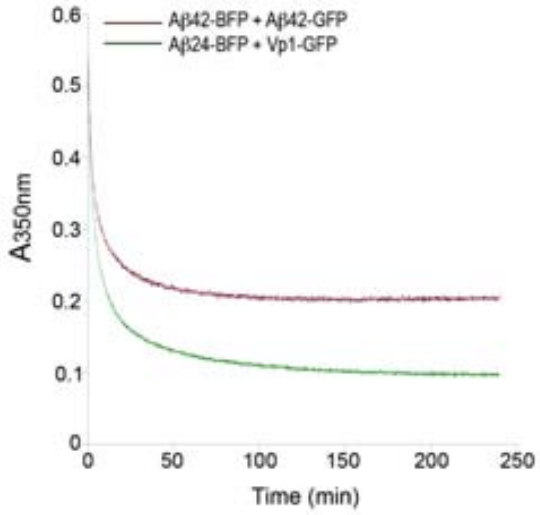

Figure 3.13 Kinetics of solubilization using GuHCl of IBs formed by the proteins indicated in the legend and monitored by a time-dependent decrease of turbidity at 350 nm.

Therefore, different types of IBs show different behavior in front of denaturation, being the IBs containing different aggregation-prone peptides less stable than those containing only Aβ42. The stability of IBs depends on the conformational properties of the polypeptide chains embedded in the aggregates[78]. The present data suggest that homogeneity promotes stable aggregates because IBs structure is glued through a stable network of intermolecular contacts. On the other hand, in heterogeneous IBs, segregation and distance between the different aggregated species would difficult the establishment of solid inter-molecular bonds resulting in less stable structures. Therefore, it could be concluded that aggregation specificity influences IBs structure and stability.

3.4.3 Kinetic control of intracellular protein aggregation

The formation of IBs is considered to be a diffusion-limited process. Therefore, even if two different proteins finally accumulate in the same IB, if one of them aggregates faster, it would effectively exclude the other one from being incorporated in the core of the aggregate. Several bioinformatic tools as AGGRESCAN[84] or TANGO[85] have been developed to predict the relative aggregation propensities of polypeptides. Both programs predict the aggregation propensity of Aβ42-GFP to be higher than the one of Vp1-GFP.

This might result in different aggregation kinetics and could explain the differential location of both proteins in the aggregates. To confirm this extent, the relative amount of both peptides in the insoluble cellular fraction was determined at selected time points.

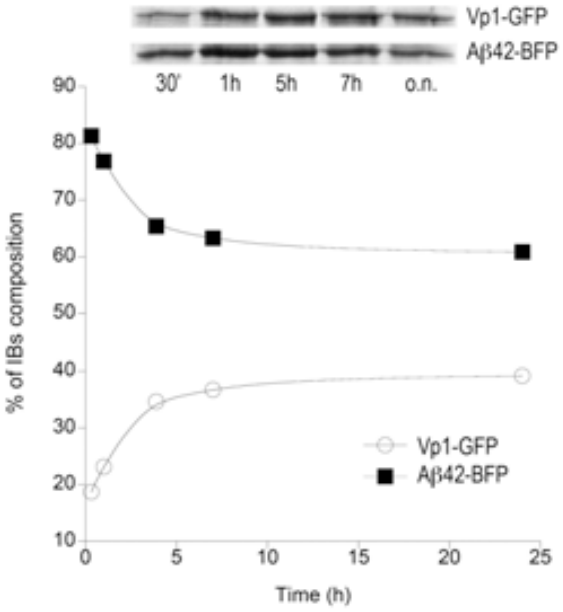

Figure 3.14 Composition of IBs along the time. The insoluble fraction of cells expressing Vp1-GFP+Aβ42-BFP was isolated and analyzed by SDS-PAGE. The percentage of each fusion was plotted versus the time after induction of protein expression.

Even if the two proteins were expressed under the same promoter, the initial aggregates contained over 80% of Aβ42-BFP, indicating that this protein fusion aggregated *in vivo* faster than Vp1-GFP forming the core of mixed IBs. The percentage of insoluble Vp1-GFP increased progressively probably by deposition on the outer face of the aggregates. The IBs composition evolved along the time to reach equilibrium around ten hours after protein expression induction in which Aβ42-BFP and VP1-GFP constituted around 60% and 40% of the protein in the aggregates, respectively.

It seems that *in vivo* kinetic partitioning could exert a tight control on the distribution of proteins in intracellular aggregates being responsible of the preferential localization of Vp1-GFP at the poles of the mixed aggregates. The observed kinetic segregation indicates that proteins tend to aggregate independently *in vivo*, and suggests that specific contacts are responsible for this behavior. The observed segregation is even more striking if one takes into account that Aβ42-BFP and Vp1-GFP share the fluorescent protein domain and that the establishment of contacts between unfolded or partially folded BFP and GFP moieties during aggregation cannot be discarded. This fact supports selectivity during intracellular deposition in physiological conditions.

3.4.4 Proteolytic digestion of IBs

For a long time, the inner structure of IBs was considered amorphous. However, our *in vivo* and *in vitro* data suggest that aggregation is the result of specific interactions between aggregation-prone species that should render, at least theoretically, some degree of internal order in the final aggregate, as it occurs in amyloid fibrils. Recent data confirm certain resemblance in the inner organization of both types of aggregates. However, IBs also contain significant amounts of functional polypeptides[80]. It is still controversial the way in which active and inactive polypeptides can coexist in cytoplasmic aggregates. The possibility that single polypeptide chains might simultaneously display misfolded protein regions engaged in the β-sheet architecture of IBs and properly folded functional domains is still under discussion. Interestingly enough, enzymes and fluorescent proteins have been successfully attached in their active forms to aggregated amyloid fibrils[86-88], suggesting that this organization of domains might exist in other type of macromolecular aggregates. In the last times, several groups are using solid-state NMR as a high-resolution approach to study *in situ* the conformational states of polypeptides within IBs. Unfortunately, no high-quality 3D information on the structure of such intracellular aggregates is available yet.

To gain insights into the molecular organization of IBs, we took advantage of the broad specificity of proteinase K (PK). This protease is applied in mapping the core of amyloid structures, because it is highly active against globular domains or disordered regions but cannot attack the densely packed cross-β-conformations residing inside amyloids[89].

We added PK to purified Aβ42-GFP IBs and measured its accessibility to properly folded domains in the aggregates by monitoring the IBs GFP fluorescence emission along time. As it is shown in Figure 3.15, PK promoted a gradual loss of green fluorescent signal that was complete after ten minutes of reaction. An SDS-PAGE analysis of the digestion shows a concomitant disappearance of the coomassie-stained band corresponding to Aβ42-GFP. The PK effect over Aβ42-GFP IBs compactness was monitored by light scattering. PK promoted a gradual decrease in the turbidity of the solution indicating a disintegration of the initial aggregates. The reaction was completed after 10 minutes (Figure 3.15).

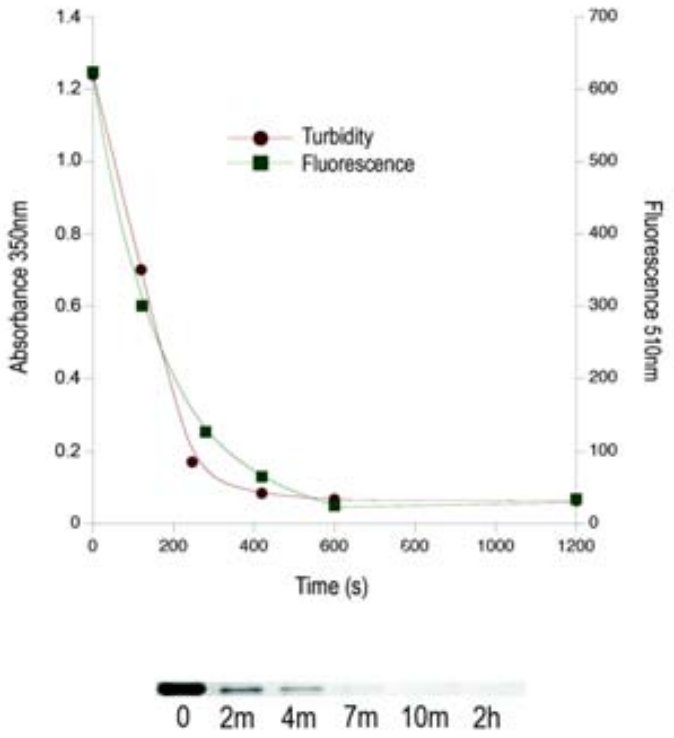

Figure 3.15 Kinetics of Aβ42-GFP IBs fluorescence and turbidity decrease in the presence of PK. A progressive disappearance of the SDS-PAGE Aβ42-GFP band could be noticed in agreement with fluorescence and turbidity data.

The parallel kinetics of light and fluorescence suggest two possibilities concerning the organization of polypeptides in IBs. First, active (fluorescent) and inactive protein regions can be mixed and organized in a structure that is homogenously sensible to proteolysis. Therefore, the inactivation of functional domains would coincide with the global disintegration of the aggregate yielding similar kinetics for fluorescence loss and light scattering. A second possibility is that only the digestion of a fraction of the protein embedded in the aggregate is monitored, whereas another part remains resistant. In this case, the fluorescence data would imply that the protease sensitive regions include polypeptides displaying a conformation compatible with fluorescence. In this scenario, it is tempting to speculate that the PK-resistant structures could correspond to the intermolecular β-sheets recurrently observed in IBs.

3.4.5 Inner structure of IBs

3.4.5.1 Transmission electron microscopy (TEM)

To distinguish between the two above-mentioned possibilities, TEM was used to explore the morphology of the aggregates before and after adding PK. In TEM images,

intact Aβ42-GFP IBs displayed the typical electro-dense ellipsoidal shape (Figure 3.16). In contrast, in samples incubated with PK, the presence of abundant fibrillar structures connecting apparently amorphous regions was observed. These fibers were not organized as a tight assembly, but they formed a loose network. The size and morphology of these structures were compatible with an amyloid-like nature.

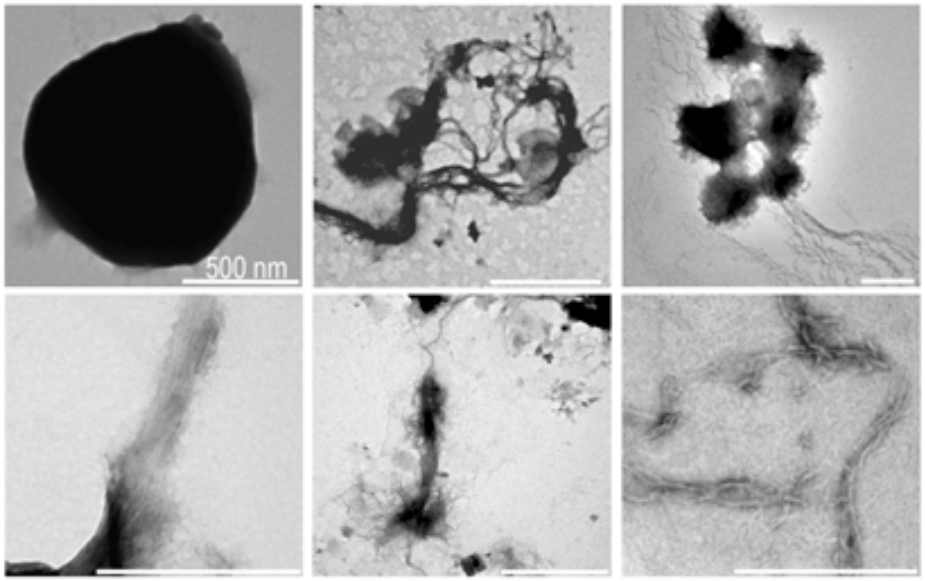

Figure 3.16 TEM images of PK digestion of Aβ42-GFP IBs. In the upper left corner, the micrograph shows a purified IB. The rest of micrographs correspond to samples of IBs digested by PK. The scale bar represents 500nm.

If the observed fibrillar structures included Aβ42 polypeptide regions in the IBs, the same kind of structures should be observed in IBs formed by the Aβ42 peptide without any fusion (Aβ42-only). To test this hypothesis, we expressed Aβ42 peptide in *E. coli*. The formation of typical spheric and refractile aggregates at the poles of the cells could be observed. The IBs formed by Aβ42-only were smaller (0.3 μm) than those formed by the Aβ42-GFP fusion (1.0 μm).

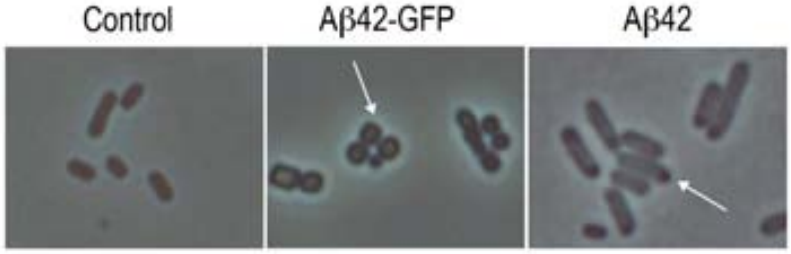

Figure 3.17 Phase contrast images of bacteria expressing Aβ42-GFP or Aβ42. In the control experiment, the bacteria were transformed with a plasmid that does not have any insert.

Incubation of Aβ42-only IBs with PK also resulted in formation of fibrillar structures, as observed by TEM (Figure 3.18).

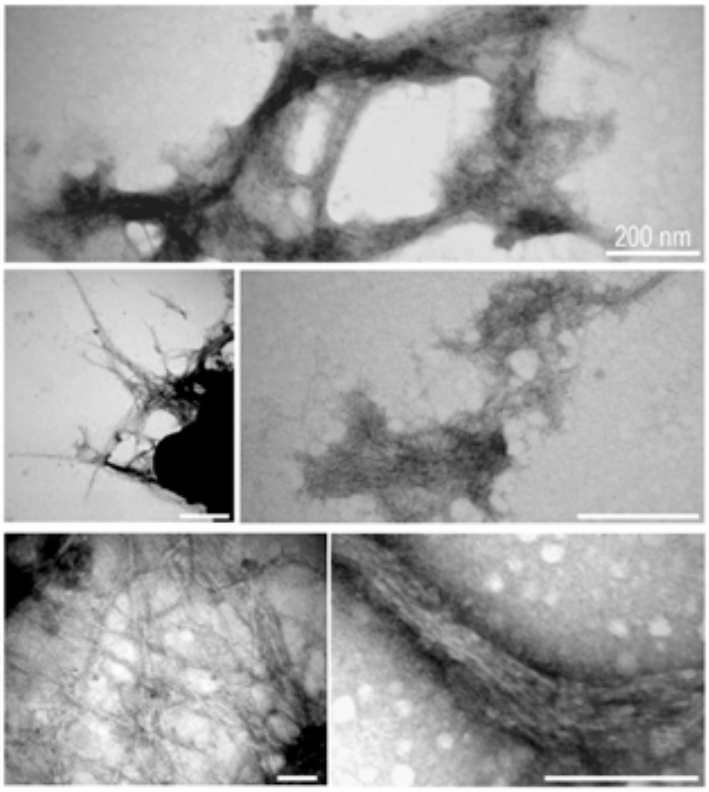

Figure 3.18 TEM images of PK digestion of Aβ42-only IBs. The scale bar represents 200nm.

Images of early digestion times show that this fibrillar material protrudes from the surface of the aggregates, suggesting the existence of ordered PK resistant cores with fibrillar structure inside the aggregates formed in the prokaryotic cytoplasm. These data provide one of the first direct evidences of the existence of fibrillar structure inside intracellular aggregates.

3.4.5.2 Atomic force microscopy (AFM)

AFM has been used for the study of several protein systems that form amyloids[90]. Recently it has been reported its use for the analysis of mammalian intracellular aggregates[91], but no data referring to its application in prokaryotic aggregates are still available. In contrast to TEM, AFM allows direct observation of protein aggregates in solution and permits following the morphology of individual molecules and aggregates along time. When AFM was applied to investigate the time course IBs PK digestion, a first surprising observation was that aggregates in close vicinity reacted very differently in front

of PK treatment. Whereas some aggregates became largely disassembled, others remained essentially intact.

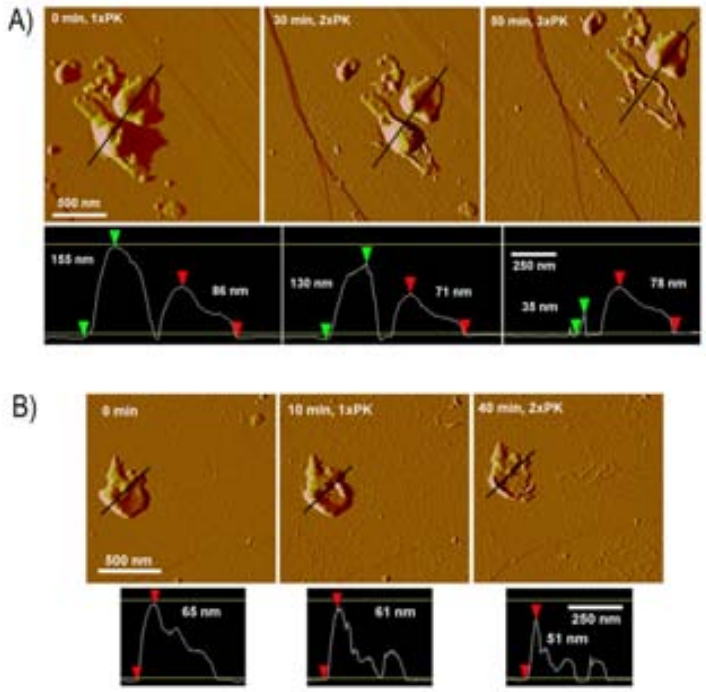

Figure 3.19 A) A group of IBs was imaged at the beginning (0 min), 30 min after addition of a first PK aliquot (1× PK), and after the addition of a second PK aliquot (2× PK). Note the fibril-like structures that settled on the surface as digestion progresses. Lines indicate the path of the topographic profiles shown below that illustrate the dramatic differences in susceptibility to PK digestion between the two largest IBs. Arrowheads indicate the points that define the height measurements in nm. B) In a parallel experiment to that described in panel A, images of a PK-digested IB were taken at 0, 10, and 40 min.

This different behavior in front of PK digestion suggests heterogeneity in the internal structure, which may arise, among other factors, from variations in the environment and aggregate growth conditions. We can speculate that each aggregate comes from an individual cell and differences in expression levels or differing metabolic states between cells might result in aggregates with different properties. Accordingly, cell-to-cell heterogeneity in the molecular interactions of intracellular aggregates formed by polyglutamine fusions inside human cells has been reported[92].

In agreement with TEM data, the action of the protease in PK sensitive aggregates promoted the progressive appearance of elongated structures with heights ranging from 2 to 5 nm, consistent with the dimensions of amyloid protofilaments and fibrils[93].

The above-discussed data indicate the coexistence in prokaryotic IBs of PK-resistant fibrillar structure and less packed regions sensitive to proteolytic attack. Using the AFM tip as force sensor, force-distance curves can be obtained providing information on single-molecule or –particle rupture forces. When applying such force spectroscopy

approach, indentations of IBs before and after PK treatment revealed that freshly prepared IBs had regions opposing relatively strong resistance to indentation, together with other softer areas (Figure 3.20). However, PK removed completely the hardest zones from the IBs simultaneously to the release of the amyloid-like fibrils.

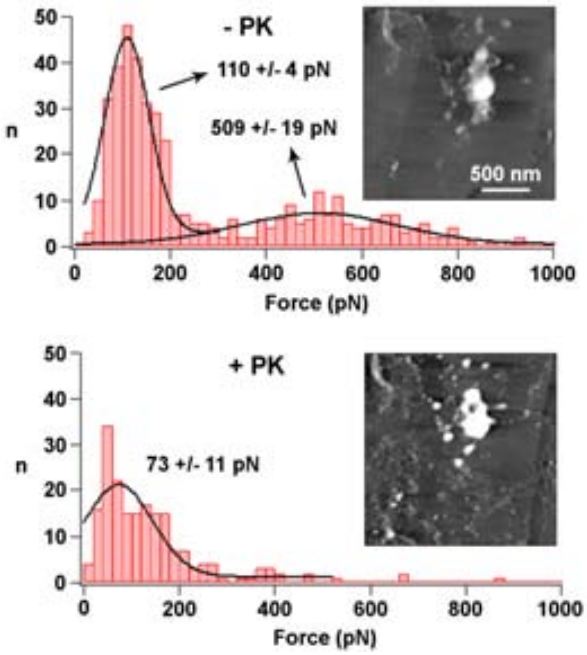

Figure 3.20 Indentation of Aβ42-GFP IBs. A single IB was indented with an AFM tip before (upper panel) and after (lower panel) PK treatment. The inset height AFM images showed the appearance on the graphite surface of abundant fibrillar material after PK digestion. The two histograms represented the forces required to penetrate IBs in both samples. The group of indentations centered at ca. 500 pN in pre-PK samples was totally absent from post-PK assays, indicating that PK treatment had eliminated from the IBs preexisting harder inner areas.

Analysis of individual approaching curves for the indentations performed before and after PK treatment revealed astonishing differences. Indentations on PK-treated IBs yielded a typical profile of progressive penetration with increasing load. Before PK digestion, on the other hand, the AFM tip penetrated the IBs once the applied load overcame the resistance of the structure, but surprisingly, the AFM tip was apparently pushed off the IB until much greater loads of ca. 6 nN finally penetrated the IB again (Figure 3.21).

This result strongly suggests internal rearrangements of the IB that might rebuild strong structures temporarily disrupted by the indenting AFM tip. This indicates that the fibrillar material confers resistance to IBs in front of indentation and confirms the existence of structural diversity inside the aggregates.

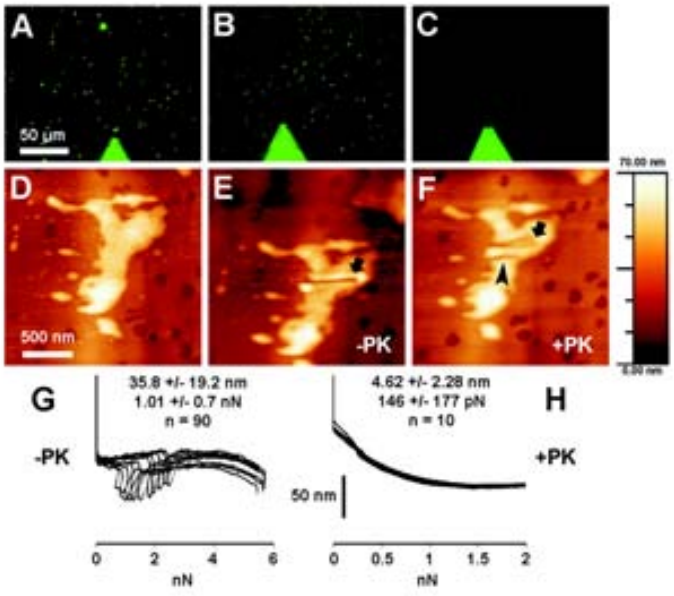

Figure 3.21 Combined force spectroscopy-fluorescence analysis study of the effect of PK treatment on IBs. A-C) GFP fluorescence of IBs immediately after deposition on the PMMA surface (A), and immediately before indentation without PK treatment (B), and after PK digestion (C). D-F) An individual IB without PK treatment before (D) and after indentation (E), and the same IB indented again after PK treatment (F). The indentation paths before and after PK treatment are indicated by an arrow and an arrowhead, respectively. (G,H) Analysis of 10 individual approaching curves for the indentations performed on the IB from panels D-F before (G) and after PK treatment (H).

3.4.6 Amyloid-like properties of PK-resistant fibrillar IBs core

3.4.6.1 FT-IR spectroscopy

Fourier-transform infrared (FT-IR) spectroscopy has been proved to be a powerful tool for investigation of secondary structure in aggregated proteins[61, 94]. In this technique, the amide I band essentially corresponds to the absorption of the carbonyl peptide bond group of the protein main chain and therefore, it is a sensitive marker of the protein secondary structure.

Using this technique, we decided to study the PK-resistant material of Aβ42-GFP IBs. The second derivative of FT-IR spectrum was dominated by two bands at 1625 cm^{-1} and 1692 cm^{-1} in the amide I region. Both bands are typically indicative of extended intermolecular β-sheet structures in the fibrillar aggregates (Figure 3.22). Specifically, the component at 1625 cm^{-1} indicates that the polypeptide backbones involved in the β-sheet architecture are tightly packed, sharing short hydrogen bonds. The presence and relative position of these two bands are generally used as a diagnostic tool for the presence of cross β-structure in most amyloid fibrils[95]. Therefore, it appears that the formation of an inter-backbone hydrogen-bonded network leading to the formation of similar fibrillar structures

enriched in β-sheet conformation constitutes a common driving force for protein deposition in different aggregates.

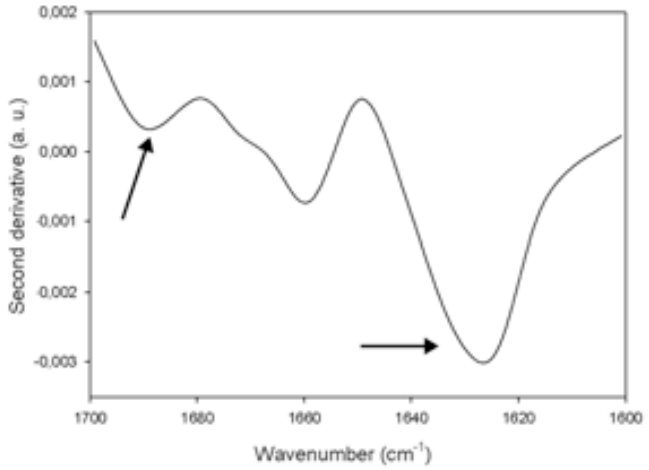

Figure 3.22 FT-IR spectrum of the PK resistant fraction of Aβ42-GFP IBs. The position of the spectral components is estimated from the second derivative. The β-sheet related bands (at 1625cm^{-1} and 1692cm^{-1}) are indicated by arrows.

3.4.6.2 Thioflavine T binding

The presence of intermolecular crossed β-pleated sheet structure in amyloid fibrils can be detected using Th-T, a specific marker for this conformation[96]. In agreement with previous results obtained with several IBs[48, 97], purified Aβ42-GFP aggregates bound to Th-T. Treatment of the aggregates with PK triggered a two-fold increase in the typical emission maximum at 485 nm (Figure 3.23).

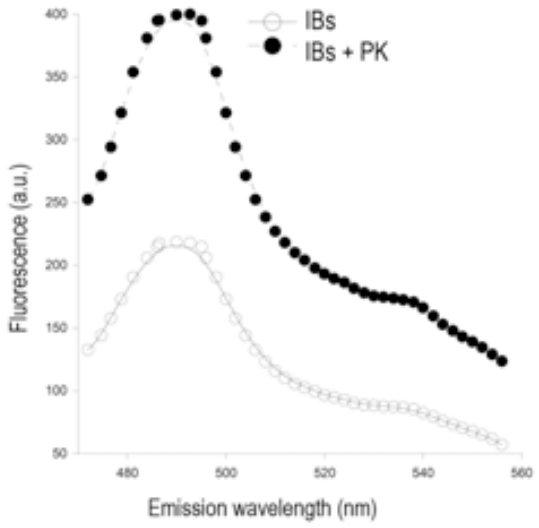

Figure 3.23 Th-T binding to Aβ42-GFP IBs before (empty circles) and after (solid circles) PK digestion.

In agreement with TEM, AFM and FT-IR data, this proteolysis promoted effect was likely the result of exposure to the solvent and increased accessibility to Th-T of previously hidden fibrillar structures displaying β-sheet structure.

3.4.6.3 Protein regions involved in the formation of IBs PK-resistant core

Recently, it has been proved that very short specific amino acid stretches can act as facilitators for the incorporation of proteins into amyloid fibrils. Usually these sequences also form part of the inner β-sheet core of the fibrils.

In order to identify the specific protein regions involved in the formation of the fibrillar Aβ42-GFP IBs PK-resistant core, MALDI-TOF MS was used. A collection of different low molecular weight peptides was identified coming from both Aβ42 peptide and GFP (Figure 3.24). Interestingly enough, all the detected peptides corresponded to aggregation-prone regions according to TANGO and AGGRESCAN calculations: 15-21, 30-45, 60-64, 112-124, 202-207 and 276-282. In particular, the region comprising the Aβ42 residues Q15-A21 was one of the most recurrently protected regions. This protein stretch corresponds to the central hydrophobic cluster (CHC) of Aβ42, a fragment involved in the formation of the Aβ42 fibrils core, as reported by solid state-NMR[98] and site-directed spin labelling[99]. Accordingly, we have also shown that point mutations in this region are able to abolish completely the formation of amyloid fibrils[17] and tightly modulate the kinetics of aggregation of Aβ42-GFP into IBs in *E. coli*[100].

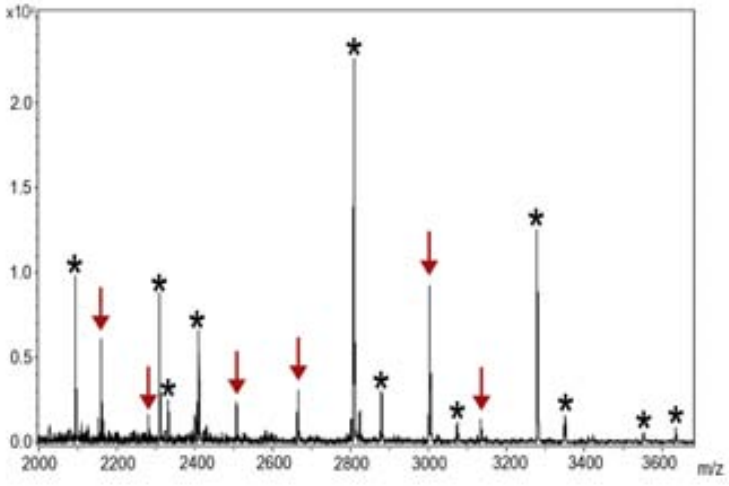

Figure 3.24 Identification of the IBs fibrillar core by mass spectroscopy. In the mass spectrum of the PK resistant fraction of Aβ42-GFP IBs, red arrows and asterisks mark peptides coming from Aβ42 peptide and GFP, respectively.

Theoretical mass	Dev (Da)	Range	Sequence
2091,099745	-0,014	[54-72]	FSKGEE**LFTGV**VPILVELD
2156,947143	-0,045	[6-23]	HDSGYEVHH**QKLVFFA**ED
2278,618544	0,086	[18-40]	**VFFA**EDVGSNKG**AIIGLMVGGVV**
2306,593353	0,088	[50-71]	GSGEFSKGEE**LFTGV**VPILVEL
2329,533299	0,078	[55-76]	SKGEE**LFTGV**VPILVELDGDVN
2356,500228	0,085	[261-281]	SKDPNEKRDHMVLLE**FVTAAG**
2406,340937	0,075	[104-125]	GKLPVPWP**TLVTTLTYGVQCF**S
2505,125915	-0,066	[7-28]	DSGYEVHH**QKLVFFA**EDVGSNK
2661,892561	-0,090	[0-21]	MDAEFRHDSGYEVHH**QKLVFFA**
2798,819907	0,071	[54-79]	FSKGEE**LFTGV**VPILVELDGDVNGHK
2805,718163	0,08	[95-119]	LTLKFICTTGKLPVPWP**TLVTTLTY**
2856,64939	0,094	[57-83]	GEE**LFTGV**VPILVELDGDVNGHKFSVS
	0,094	[58-84]	EE**LFTGV**VPILVELDGDVNGHKFSVSG
	0,094	[59-85]	E**LFTGV**VPILVELDGDVNGHKFSVSGE
2876,64972	0,093	[192-215]	HKLEYNYNSH**NVYITA**DKQKNGIK
3000,527426	-0,026	[10-36]	YEVHH**QKLVFFA**EDVGSNKGAIIGLMV
3071,479557	-0,029	[49-78]	AGSGEFSKGEE**LFTGV**VPILVELDGDVNGH
3131,47018	-0,015	[3-29]	EFRHDSGYEVHH**QKLVFFA**EDVGSNKG
3275,463188	-0,08	[113-139]	**LVTTLTYGVQCF**SRYPDHMKRHDFFKS
3346,433746	-0,06	[117-144]	**LTYGVQCF**SRYPDHMKRHDFFKSAMPEG
3548,51503	-0,07	[115-144]	**TTLTYGVQCF**SRYPDHMKRHDFFKSAMPEG
3630,629074	-0,08	[186-217]	DGNILGHKLEYNYNSH**NVYITA**DKQKNGIKAN

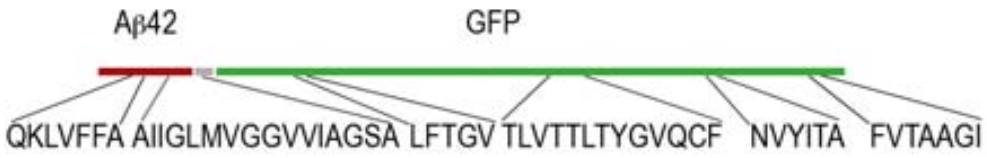

Figure 3.25 In the table, the theoretical mass, the standard deviation of the identification, the amino acid range in the protein fusion and the primary sequence of all the detected peptides are shown. The scheme below represents the protein fusion with the Aβ42 in red, the linker in grey and GFP in green. The regions depicted correspond to aggregation-prone sequences, as predicted by AGGRESCAN and TANGO. These segments are highlighted in bold in the table.

The presence of both Aβ42 and GFP fragments in the PK-resistant core is consistent with the observation that for chimerical polyglutamine-proteins: regions from both the aggregation-prone polyQ ladder and the attached globular protein are found in the core of bacterial aggregates[101].

3.4.6.4 Detection of SDS-resistant oligomers

Wealth of data converges to indicate that the early steps in the fibrillization pathway of Aβ42 involve the formation of non-covalent, SDS-resistant oligomers[102]. These species have also been detected in mammalian cell culture, and in the human brain[103]. And interestingly, the expression in yeast of Aβ42 fused to the globular domain of the Sup35 prion renders SDS-stable oligomers[104].

In order to know if these oligomers were also present in bacteria, western blot analysis of GFP was performed in *E. coli* cells expressing Aβ42-GFP. With this analysis, the formation of high molecular weight, SDS-stable, oligomers shortly after induction of protein expression could be demonstrated (Figure 3.26).

This observation adds evidence to the hypothesis that the molecular pathways that drive proteins to associate *in vitro* as fibrils are the same that also favor aggregation as intracellular aggregates in both eukaryotic and prokaryotic environments.

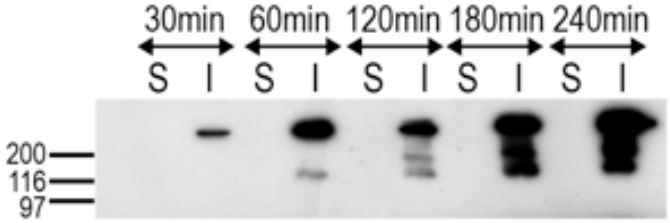

Figure 3.26 Western Blot analysis of high molecular weight species in the insoluble (I) and soluble (S) cell fractions of cells expressing Aβ42-GFP at different times after IPTG addition.

3.4.6.5 IBs seeds Aβ42 fibril formation

It is well established that the rate of fibrils formation by an amyloidogenic protein is enhanced by the addition of preformed fibrils, a phenomenon known as seeding. This effect has been associated with the transmission of the TSEs and with the rapid development of Alzheimer's disease once clinical symptoms are detected[105, 106]. Seeding displays a high sequence specificity and protein aggregation can be nucleated by homologous fibrils, but not by fibrils from closely related sequences[107]. This behavior is thought to be the cause of the high purity of protein depositions in amyloid diseases. Moreover, a recent study claims that the amino acid sequence is important only insofar as it confers a preference for a particular fibril conformation: it is in fact fibril conformation that correlates with seeding capability[108]. This fact provides us with a unique tool to demonstrate that intracellular prokaryotic aggregates display amyloid conformations.

Monitoring the changes in Th-T fluorescence emission as a function of time, we assayed whether Aβ42-only IBs were able to specifically seed the formation of amyloid fibrils by a synthetic Aβ42 peptide (Figure 3.27). The formation of amyloid fibrils followed a characteristic sigmoid curve displaying a nucleation phase of 251 min in which no increase in fluorescence was detected, followed by a fast fibril growth phase that reached a plateau after 380 min of incubation. Seeding the reaction with 10% of *in vitro*

preformed Aβ42 fibrils completely abrogated the lag phase and the fibril formation reaction was completed after only 12 min. Importantly, the addition of purified Aβ42-only IBs also seeded Aβ42 peptide fibril formation, significantly reducing the nucleation phase to 105 min and allowing the reaction to be completed in 285 min.

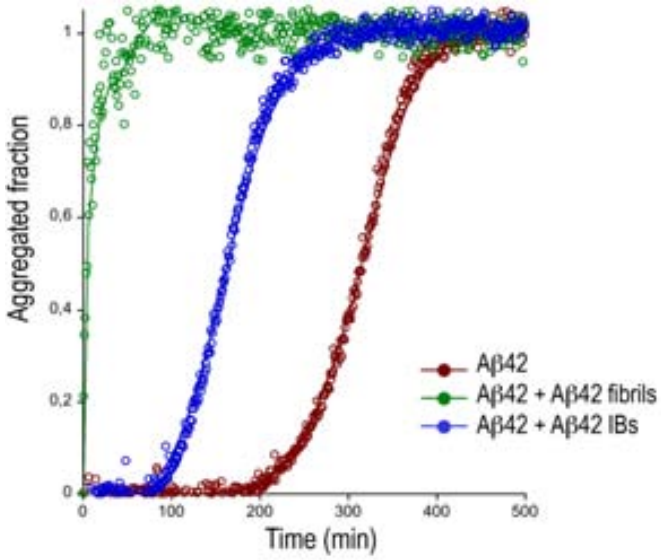

Figure 3.27 Seeding dependent maturation of synthetic Aβ42 peptide amyloid fibril growth monitored through Th-T fluorescence emission. The fibrillar fraction of Aβ42 is represented as a function of time. The formation of Aβ42 amyloid fibrils was accelerated by the addition of Aβ42 preformed fibrils or Aβ42-only IBs.

Thus, the polypeptides embedded in the IBs can specifically recognize soluble Aβ42 peptide molecules and promote their fibrillization into β-sheet rich structures. Two recent independent works have demonstrated the existence of fibrillar structure in the intracellular aggregates formed by polyQ-GFP fusions inside mammalian cells[109, 110].

3.5 Discussion

Our data indicate that the existence of fibrillar structure in intracellular aggregates is not restricted to disease-linked eukaryotic aggregates. It is rather a general property of the intracellular inclusions formed by aggregation-prone polypeptides, independently of the cellular background. In this sense, it is demonstrated that the prokaryotic intracellular aggregates and amyloid fibrils of a given protein share conformational and mechanistic features. In addition, it is shown that, as in eukaryotes, the formation of prokaryotic aggregates exhibits an important *in vivo* specificity even among extremely aggregating substrates expressed at very high levels.

Our data regarding the amyloid-like structure of IBs have been recently confirmed by another parallel work performed by Riek and coworkers[111]. Taking three proteins with dissimilar native folds, they demonstrate that their corresponding bacterial IBs are amyloid-like and are comprised of amino acid sequence-specific cross-β structure by CR birefrigerence and X-ray diffraction, respectively. Furthermore, they prove that IBs contain sequence-specific positions of regular secondary structure by analyzing the H/D exchange by NMR. Taking into account the structural data, it is suggested that the primary structure of each polypeptide will determine the number, length and form of the cross β-sheet segments. And, they hypothesize that IBs may contain a core of ordered structured β-sheet adjacent to amorphously disordered segments or also folded domains. It is worth to notice that our results coincide with this model because together with the fibrillar part of the IBs, the fluorescence displayed by IBs points to a partially native folded GFP embedded in them.

In many conformational disorders, aggregation-prone, causative proteins are incorporated into intracellular microscopic IBs. Therefore, during a long time, it was supposed that these intracellular aggregates were the toxic structures associated to the disease. However, the onset of cell dysfunction has been found to be independent of the formation of IBs[112]. In addition, recent data suggest that pre-aggregated, diffusible assemblies are the most harmful species; on the contrary, high order aggregates might display a protective role[113]. Accordingly, a comparison of the cytotoxicity of poliQ oligomers and their homologous eukaryotic IBs in neuronally differentiated cells reveals that cells containing IBs have longer survival periods[114]. In a similar way, immature bacterial IBs are more toxic in both bacterial and neuronal cell lines than mature forms of

the same aggregates[115, 116]. These results are also obtained working with *in vivo* models. For example, the injection of Aβ dimers in rats impair synapse structure and function whereas the injection of amyloid plaques does not have the same effect[117].

In this later scenario, because all organisms face challenges of protein misfolding and aggregation, it is not surprising the existence of an evolutionary conserved strategy to avoid the harmful effects of undesired protein aggregation by sequestering sticky folding intermediates into conformationally related stable aggregated structures through selective interactions. Accordingly, the genes encoding for the protein quality control machinery genes are highly conserved across species. Furthermore, an increasing number of reports indicates that even the proteins themselves have acquired structural and sequence-based strategies to escape undesired general protein aggregation, but also to avoid the specific formation of amyloid-like fibrils[118]. In this sense, in a recent study, using an algorithm that predicts protein aggregation, twenty-eight complete proteomes were surveyed. The results demonstrate that the evolution selectively capes strong aggregating protein sequences with amino acids with low aggregation propensity such as arginine or proline. Furthermore, the substrate specificity of various chaperones is geared by these positive amino acids flanking hydrophobic sequences. Elucidating such adaptations is important for identifying those strategies that nature has designed to protect proteins from the uncontrolled formation of amyloid-like fibrils as well as their precursors. This nature-derived knowledge is critical because it could be extremely useful to design strategies to rationally control aggregation in pathology or biotechnology.

The tight link between aggregation in eukaryotic and prokaryotic systems suggests that prokaryotes should be seriously considered when exploring the *in vivo* determinants and cellular effects of protein aggregation. One recent example illustrating this statement is the demonstration that the segregation of protein aggregates in the cell is associated with cellular aging and rejuvenation, conclusion resulting from the study of aggregation in bacterial backgrounds[119].

In addition, the formation in *E. coli* of SDS-resistant oligomers and fibrillar structures able to seed amyloid fibril formation raise important safety concerns relative to the biotechnological and biomedical applications of recombinant polypeptides, especially for aggregation-prone molecules. In this sense, the exposure to exogenous substances with amyloid enhancing potential may be an important epigenetic or environmental factor in the development of amyloidosis in a susceptible population[120]. For example, in a recent work, birefringent congophilic material composed of serum amyloid A-related protein was

detected in commercial foie gras. The oral or intravenous administration of this material in transgenic murine model of secondary (amyloid A protein) amyloidosis acted as a potent amyloid enhancing factor: the animals developed extensive systemic pathological deposits. On this basis, it is hypothesized that this and perhaps other forms of amyloidosis may be transmissible, akin to the infectious nature of prion-related illnesses.

3.6 References

1. Karplus, M. The Levinthal paradox: yesterday and today. *Fold Des* **2**, S69-75 (1997).
2. Dill, K.A. & Chan, H.S. From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-19 (1997).
3. Wolynes, P.G., Onuchic, J.N. & Thirumalai, D. Navigating the folding routes. *Science* **267**, 1619-1620 (1995).
4. Dobson, C., A., S. & M., K. Protein folding: a perspective from theory and experiment. *Angew. Chem., Int. ed. Engl.* **37**, 868-893 (1998).
5. Dinner, A.R., Sali, A., Smith, L.J., Dobson, C.M. & Karplus, M. Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem Sci* **25**, 331-339 (2000).
6. Jackson, S.E. How do small single-domain proteins fold? *Fold Des* **3**, R81-91 (1998).
7. Dobson, C.M. Protein folding and misfolding. *Nature* **426**, 884-890 (2003).
8. Radford, S.E. & Dobson, C.M. From computer simulations to human disease: emerging themes in protein folding. *Cell* **97**, 291-298 (1999).
9. Thomas, P.J., Qu, B.H. & Pedersen, P.L. Defective protein folding as a basis of human disease. *Trends Biochem Sci* **20**, 456-459 (1995).
10. Dobson, C.M. Protein-misfolding diseases: Getting out of shape. *Nature* **418**, 729 - 730 (2002).
11. Rochet, J.C. & Lansbury, P.T. Amyloid fibrillogenesis: themes and variations. *Curr Opin Struct Biol* **10**, 60 - 68 (2000).
12. Tan, S.Y. & Pepys, M.B. Amyloidosis. *Histopathology*, 403-414 (1994).
13. Sunde, M. & Blake, C.C.F. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. . *Adv. Protein Chem.* **50**, 123-159 (1997).
14. Dobson, C.M. The structural basis of protein folding and its links with human disease. *Phil. Trans. R. Soc. Lond. B* **356**, 133-145 (2001).
15. Fändrich, M.D., C. M. The behaviour of polyamino acids reveals an inverse side-chain effect in amyloid structure formation. . *EMBO J.* **21**, 5682-5690 (2002).
16. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & C.M., D. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424**, 805-808 (2003).
17. de Groot, N.S., Aviles, F.X., Vendrell, J. & Ventura, S. Mutagenesis of the central hydrophobic cluster in Abeta42 Alzheimer's peptide. Side-chain properties correlate with aggregation propensities. *Febs J* **273**, 658-668 (2006).
18. Lashuel, H.A., Hartley, D., Petre, B.M., Walz, T. & Lansbury, P.T., Jr. Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* **418**, 291 (2002).
19. Walsh, D.M., Lomakin, A., Benedek, G.B., Condron, M.M. & Teplow, D.B. Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate. *J Biol Chem* **272**, 22364-22372 (1997).
20. Jahn, T.R. & Radford, S.E. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys* **469**, 100-117 (2008).
21. Gosal, W.S. et al. Competing pathways determine fibril morphology in the self-assembly of beta2-microglobulin into amyloid. *J Mol Biol* **351**, 850-864 (2005).

22. Modler, A.J., Gast, K., Lutsch, G. & Damaschun, G. Assembly of amyloid protofibrils via critical oligomers--a novel pathway of amyloid formation. *J Mol Biol* **325**, 135-148 (2003).
23. Smith, A.M., Jahn, T.R., Ashcroft, A.E. & Radford, S.E. Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionisation mass spectrometry. *J Mol Biol* **364**, 9-19 (2006).
24. Baskakov, I.V., Legname, G., Baldwin, M.A., Prusiner, S.B. & Cohen, F.E. Pathway complexity of prion protein assembly into amyloid. *J Biol Chem* **277**, 21140-21148 (2002).
25. Bitan, G. et al. Amyloid beta -protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci U S A* **100**, 330-335 (2003).
26. Necula, M., Kayed, R., Milton, S. & Glabe, C.G. Small molecule inhibitors of aggregation indicate that amyloid beta oligomerization and fibrillization pathways are independent and distinct. *J Biol Chem* **282**, 10311-10324 (2007).
27. Wickner, S., Maurizi, M.R. & Gottesman, S. Posttranslational quality control: folding, refolding, and degrading proteins. *Science* **286**, 1888-1893 (1999).
28. Kopito, R.R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol* **10**, 524-530 (2000).
29. Ventura, S. & Villaverde, A. Protein quality in bacterial inclusion bodies. *Trends Biotechnol* **24**, 179-185 (2006).
30. Hishiya, A. et al. A novel ubiquitin-binding protein ZNF216 functioning in muscle atrophy. *Embo J* **25**, 554-564 (2006).
31. Garcia-Mata, R., Bebok, Z., Sorscher, E.J. & Sztul, E.S. Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J Cell Biol* **146**, 1239-1254 (1999).
32. Sorensen, H.P. & Mortensen, K.K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact* **4**, 1 (2005).
33. Carrio, M.M., Corchero, J.L. & Villaverde, A. Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. *FEMS Microbiol Lett* **169**, 9-15 (1998).
34. Villaverde, A. & Carrio, M.M. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* **25**, 1385-1395 (2003).
35. Rinas, U. & Bailey, J.E. Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli. *Appl Microbiol Biotechnol* **37**, 609-614 (1992).
36. Valax, P. & Georgiou, G. Molecular characterization of beta-lactamase inclusion bodies produced in Escherichia coli. 1. Composition. *Biotechnol Prog* **9**, 539-547 (1993).
37. Carrio, M.M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. *J Biotechnol* **96**, 3-12 (2002).
38. London, J., Skrzynia, C. & Goldberg, M.E. Renaturation of Escherichia coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. *Eur J Biochem* **47**, 409-415 (1974).
39. Speed, M.A., Wang, D.I. & King, J. Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat Biotechnol* **14**, 1283-1287 (1996).
40. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* **347**, 1025-1037 (2005).

41. Bukau, B., Weissman, J. & Horwich, A. Molecular chaperones and protein quality control. *Cell* **125**, 443-451 (2006).
42. Luheshi, L.M. et al. Systematic In Vivo Analysis of the Intrinsic Determinants of Amyloid beta Pathogenicity. *PLoS Biol* **5**, e290 (2007).
43. Rajan, R.S., Illing, M.E., Bence, N.F. & Kopito, R.R. Specificity in intracellular protein aggregation and inclusion body formation. *Proc Natl Acad Sci U S A* **98**, 13060-13065 (2001).
44. Johnston, J.A., Ward, C.L. & Kopito, R.R. Aggresomes: a cellular response to misfolded proteins. *J Cell Biol* **143**, 1883-1898 (1998).
45. Chrunyk, B.A., Evans, J., Lillquist, J., Young, P. & Wetzel, R. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. *J Biol Chem* **268**, 18053-18061 (1993).
46. Thomas, J.G., Ayling, A. & Baneyx, F. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold. *Appl Biochem Biotechnol* **66**, 197-238 (1997).
47. Carrio, M.M. & Villaverde, A. Role of molecular chaperones in inclusion body formation. *FEBS Lett* **537**, 215-221 (2003).
48. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. *J. Mol. Biol.* **347**, 1025-1037 (2005).
49. Izard, J., Parker, M.W., Chartier, M., Duche, D. & Baty, D. A single amino acid substitution can restore the solubility of aggregated colicin A mutants in Escherichia coli. *Protein Eng* **7**, 1495-1500 (1994).
50. Krueger, J.K., Stock, J. & Schutt, C.E. Evidence that the methylesterase of bacterial chemotaxis may be a serine hydrolase. *Biochim Biophys Acta* **1119**, 322-326 (1992).
51. Wetzel, R. & Chrunyk, B.A. Inclusion body formation by interleukin-1 beta depends on the thermal sensitivity of a folding intermediate. *FEBS Lett* **350**, 245-248 (1994).
52. Wetzel, R., Perry, L.J. & Veilleux, C. Mutations in human interferon gamma affecting inclusion body formation identified by a general immunochemical screen. *Biotechnology (N Y)* **9**, 731-737 (1991).
53. Nieba, L., Honegger, A., Krebber, C. & Pluckthun, A. Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng* **10**, 435-444 (1997).
54. Chan, W. et al. Mutational effects on inclusion body formation in the periplasmic expression of the immunoglobulin VL domain REI. *Fold Des* **1**, 77-89 (1996).
55. Yan, G. et al. A single residual replacement improves the folding and stability of recombinant cassava hydroxynitrile lyase in E. coil. *Biotechnol Lett* **25**, 1041-1047 (2003).
56. Przybycien, T.M., Dunn, J.P., Valax, P. & Georgiou, G. Secondary structure characterization of beta-lactamase inclusion bodies. *Protein Eng* **7**, 131-136 (1994).
57. Schrodel, A. & de Marco, A. Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem* **6**, 10 (2005).
58. Cubarsi, R., Carrio, M.M. & Villaverde, A. In situ proteolytic digestion of inclusion body polypeptides occurs as a cascade process. *Biochem Biophys Res Commun* **282**, 436-441 (2001).

59. Cubarsi, R., Carrio, M.M. & Villaverde, A. A mathematical approach to molecular organization and proteolytic disintegration of bacterial inclusion bodies. *Math Med Biol* **22**, 209-226 (2005).
60. Nelson, R. & Eisenberg, D. Structural models of amyloid-like fibrils. *Adv Protein Chem* **73**, 235-282 (2006).
61. Ami, D., Natalello, A., Taylor, G., Tonon, G. & Maria Doglia, S. Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. *Biochim Biophys Acta* **1764**, 793-799 (2006).
62. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* **347**, 1025 - 1037 (2005).
63. Ami, D. et al. FT-IR study of heterologous protein expression in recombinant Escherichia coli strains. *Biochim. Biophys. Acta* **1624**, 6-10 (2003).
64. Ami, D., Natalello, A., Gatti-Lafranconi, P., Lotti, M. & Doglia, S.M. Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Lett* **579**, 3433-3436 (2005).
65. Oberg, K., Chrunyk, B.A., Wetzel, R. & Fink, A.L. Nativelike secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. *Biochemistry* **33**, 2628-2634 (1994).
66. Garcia-Fruitos, E. et al. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact* **4**, 27 (2005).
67. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol Bioeng* **96**, 1101-1106 (2007).
68. Gonzalez-Montalban, N., Garcia-Fruitos, E., Ventura, S., Aris, A. & Villaverde, A. The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body-forming cells. *Microb Cell Fact* **5**, 26 (2006).
69. Garcia-Fruitos, E. et al. Divergent genetic control of protein solubility and conformational quality in Escherichia coli. *J Mol Biol* **374**, 195-205 (2007).
70. Gonzalez-Montalban, N., Garcia-Fruitos, E. & Villaverde, A. Recombinant protein solubility - does more mean better? *Nat Biotechnol* **25**, 718-720 (2007).
71. Carrio, M.M., Corchero, J.L. & Villaverde, A. Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. *Biochim Biophys Acta* **1434**, 170-176 (1999).
72. Carrio, M.M. & Villaverde, A. Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett* **489**, 29-33 (2001).
73. Dougan, D.A., Mogk, A. & Bukau, B. Protein folding and degradation in bacteria: to degrade or not to degrade? That is the question. *Cell Mol Life Sci* **59**, 1607-1616 (2002).
74. Baneyx, F. & Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nat Biotechnol* **22**, 1399-1408 (2004).
75. Garcia-Fruitos, E., Aris, A. & Villaverde, A. Localization of functional polypeptides in bacterial inclusion bodies. *Appl Environ Microbiol* **73**, 289-294 (2007).
76. Wurth, C., Guimard, N.K. & Hecht, M.H. Mutations that reduce aggregation of the Alzheimer's Abeta42 peptide: an unbiased search for the sequence determinants of Abeta amyloidogenesis. *J Mol Biol* **319**, 1279-1290 (2002).
77. Carrio, M.M., Cubarsi, R. & Villaverde, A. Fine architecture of bacterial inclusion bodies. *FEBS Lett* **471**, 7 - 11 (2000).

78. de Groot, N.S. & Ventura, S. Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett* (2006).
79. Sabate, R., Gallardo, M. & Estelrich, J. An autocatalytic reaction as a model for the kinetics of the aggregation of beta-amyloid. *Biopolymers* **71**, 190-195 (2003).
80. Garcia-Fruitos, E. et al. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microbial Cell Factories* **4**, 27 (2005).
81. Pfleger, K.D. & Eidne, K.A. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat Methods* **3**, 165-174 (2006).
82. Heim, R., Cubitt, A.B. & Tsien, R.Y. Improved green fluorescence. *Nature* **373**, 663-664 (1995).
83. Anikovsky, M., Dale, L., Ferguson, S. & Petersen, N. Resonance Energy Transfer in Cells: a New Look at Fixation Effect and Receptor Aggregation on Cell Membrane. *Biophys J* (2008).
84. Conchillo-Sole, O. et al. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65 (2007).
85. Fernandez-Escamilla, A.M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* **22**, 1302 - 1306 (2004).
86. Baxa, U., Speransky, V., Steven, A.C. & Wickner, R.B. Mechanism of inactivation on prion conversion of the Saccharomyces cerevisiae Ure2 protein. *Proc Natl Acad Sci U S A* **99**, 5253-5260 (2002).
87. Hamada, D. et al. Effect of an amyloidogenic sequence attached to yellow fluorescent protein. *Proteins* (2008).
88. Sondheimer, N. & Lindquist, S. Rnq1: an epigenetic modifier of protein function in yeast. *Mol Cell* **5**, 163-172 (2000).
89. Wilson, L.M. et al. A structural core within apolipoprotein C-II amyloid fibrils identified using hydrogen exchange and proteolysis. *J Mol Biol* **366**, 1639-1651 (2007).
90. Ban, T., Yamaguchi, K. & Goto, Y. Direct observation of amyloid fibril growth, propagation, and adaptation. *Acc Chem Res* **39**, 663-670 (2006).
91. Mukai, H. et al. Formation of morphologically similar globular aggregates from diverse aggregation-prone proteins in mammalian cells. *Proc Natl Acad Sci U S A* **102**, 10887-10892 (2005).
92. Kim, S., Nollen, E.A., Kitagawa, K., Bindokas, V.P. & Morimoto, R.I. Polyglutamine protein aggregates are dynamic. *Nat Cell Biol* **4**, 826-831 (2002).
93. Arimon, M. et al. Fine structure study of Abeta1-42 fibrillogenesis with atomic force microscopy. *Faseb J* **19**, 1344-1346 (2005).
94. Fink, A.L. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold Des* **3**, R9-23 (1998).
95. Hiramatsu, H. & Kitagawa, T. FT-IR approaches on amyloid fibril structure. *Biochim Biophys Acta* **1753**, 100-107 (2005).
96. Sabate, R. & Saupe, S.J. Thioflavin T fluorescence anisotropy: an alternative technique for the study of amyloid aggregation. *Biochem Biophys Res Commun* **360**, 135-138 (2007).
97. Espargaro, A., Castillo, V., de Groot, N.S. & Ventura, S. The in vivo and in vitro aggregation properties of globular proteins correlate with their conformational stability: the SH3 case. *J Mol Biol* **378**, 1116-1131 (2008).

98. Petkova, A.T. et al. A structural model for Alzheimer's beta-amyloid fibrils based on experimental constraints from solid state NMR. *Proc Natl Acad Sci U S A* **99**, 16742 - 16747 (2002).
99. Torok, M. et al. Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling. *J Biol Chem* **277**, 40810 - 40815 (2002).
100. de Groot, N.S. & Ventura, S. Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. *J Biotechnol* **125**, 110-113 (2006).
101. Bulone, D., Masino, L., Thomas, D.J., San Biagio, P.L. & Pastore, A. The interplay between PolyQ and protein context delays aggregation by forming a reservoir of protofibrils. *PLoS ONE* **1**, e111 (2006).
102. Kryndushkin, D.S., Alexandrov, I.M., Ter-Avanesyan, M.D. & Kushnirov, V.V. Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *J Biol Chem* **278**, 49636-49643 (2003).
103. Takahashi, R.H. et al. Oligomerization of Alzheimer's beta-amyloid within processes and synapses of cultured neurons and brain. *J Neurosci* **24**, 3592-3599 (2004).
104. Bagriantsev, S. & Liebman, S. Modulation of Abeta42 low-n oligomerization using a novel yeast reporter system. *BMC Biol* **4**, 32 (2006).
105. Perutz, M.F. & Windle, A.H. Cause of neural death in neurodegenerative diseases attributable to expansion of glutamine repeats. *Nature* **412**, 143-144 (2001).
106. Selkoe, D.J. Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. *Nat Cell Biol* **6**, 1054-1061 (2004).
107. Krebs, M.R., Morozova-Roche, L.A., Daniel, K., Robinson, C.V. & Dobson, C.M. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci* **13**, 1933-1938 (2004).
108. Jones, E.M. & Surewicz, W.K. Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell* **121**, 63-72 (2005).
109. Hazeki, N. et al. Ultrastructure of nuclear aggregates formed by expressing an expanded polyglutamine. *Biochem Biophys Res Commun* **294**, 429-440 (2002).
110. Mitsui, K. et al. Purification of polyglutamine aggregates and identification of elongation factor-1alpha and heat shock protein 84 as aggregate-interacting proteins. *J Neurosci* **22**, 9267-9277 (2002).
111. Wang, L., Maji, S.K., Sawaya, M.R., Eisenberg, D. & Riek, R. Bacterial Inclusion Bodies Contain Amyloid-Like Structure. *PLoS Biol* **6**, e195 (2008).
112. Saudou, F., Finkbeiner, S., Devys, D. & Greenberg, M.E. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* **95**, 55-66 (1998).
113. Walsh, D.M. & Selkoe, D.J. A beta oligomers - a decade of discovery. *J Neurochem* **101**, 1172-1184 (2007).
114. Takahashi, T. et al. Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. *Hum Mol Genet* **17**, 345-356 (2008).
115. Gonzalez-Montalban, N., Carrio, M.M., Cuatrecasas, S., Aris, A. & Villaverde, A. Bacterial inclusion bodies are cytotoxic in vivo in absence of functional chaperones DnaK or GroEL. *J Biotechnol* **118**, 406-412 (2005).
116. Gonzalez-Montalban, N., Villaverde, A. & Aris, A. Amyloid-linked cellular toxicity triggered by bacterial inclusion bodies. *Biochem Biophys Res Commun* **355**, 637-642 (2007).

117. Shankar, G.M. et al. Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat Med* **14**, 837-842 (2008).
118. Monsellier, E. & Chiti, F. Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep* **8**, 737-742 (2007).
119. Lindner, A.B., Madden, R., Demarez, A., Stewart, E.J. & Taddei, F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A* **105**, 3076-3081 (2008).
120. Solomon, A. et al. Amyloidogenic potential of foie gras. *Proc Natl Acad Sci U S A* **104**, 10998-11001 (2007).

CHAPTER 4

Linking protein aggregation and yeast survival

4.1 Introduction

In the last few years, protein aggregation has emerged from a neglected area of protein chemistry as a transcendental issue in biological and medical sciences[1]. In this regard, an increasing body of evidence points out at the anomalous misassembly of proteins into insoluble amyloid deposits as the fundamental cause behind some debilitating human disorders of growing incidence such as Alzheimer's disease (AD), Parkinson's disease (PD), type II diabetes, the transmissible spongiform encephalopathies and many others[2]. A common trait of these disorders is that the aggregated protein deposits in internal organs and interferes with normal cellular function, sometimes lethally[3]. On the other hand, protein aggregation in cell factories represents a major bottleneck in recombinant protein production, narrowing the spectrum of polypeptides obtained by recombinant techniques and hampering the development of priority research areas such as structural genomics and proteomics[4]. Therefore, there is an increasing interest in the development of protein solubility screening methods that allow foreseeing genes, chemical compounds or culture conditions that would modulate protein aggregation.

4.1.1 Methods to evaluate *in vivo* folding and aggregation of polypeptides

4.1.1.1 Methods to screen protein solubility and aggregation in prokaryotic cells

To date, several protein solubility assays have been developed in order to study and detect aggregation in prokaryotic cells, specifically in *Escherichia coli* [5-7]. The majority of these systems rely on the ability of aggregating proteins to co-translationally induce improper folding of a fused marker protein. Thus, different reporter proteins have been used allowing different readouts as fluorescence, chloramphenicol resistance or β-galactosidase activity.

4.1.1.1.1 Fluorescence based methods

One of the first developed methods used the GFP as a reporter of concomitant target protein aggregation[6]. It is based on the finding that the fluorescence in GFP fusions correlates with the tendency of the N-terminal fusion partner to form insoluble aggregates. Fusions with non-aggregating partners allow proper folding of the GFP reporter. Therefore, colonies expressing such fusions exhibit green fluorescence (Figure 4.1). On the contrary, GFP fused to an aggregating-prone protein do not fold properly and the

fluorescence emission decreases. As a test case, amyloid peptide Aβ42 was fused to GFP and expressed in bacteria. After performing random mutagenesis over this peptide, mutations that reduce Aβ42 tendency to aggregate resulted in increased fluorescence of the bacteria expressing such fusions[8].

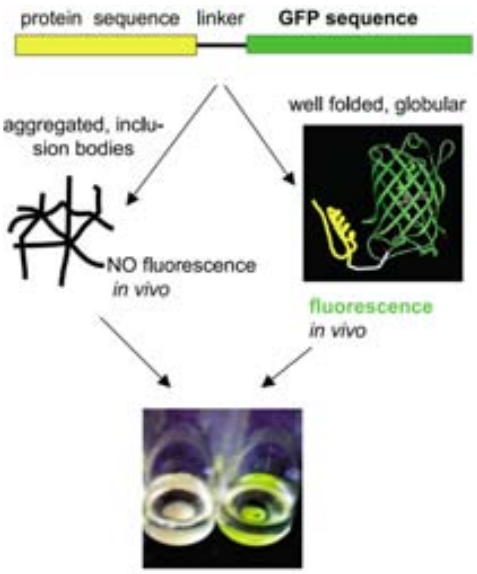

Figure 4.1. Schematic depiction of the properties of Aβ42–GFP fusion proteins. Wild-type Aβ42 forms insoluble amyloid (left) and prevents the GFP portion of the fusion protein from forming its native fluorescent structure. However, mutants in the Aβ42 sequence that prevent aggregation enable GFP to form its native green fluorescent chromophore (right).

Fluorescence resonance energy transfer (FRET) has also been used to establish a novel *in vivo* screening system that allows rapid detection of protein folding and protein variants with increased thermodynamic stability in the *E.coli* cytoplasm[9] (Figure 4.2).

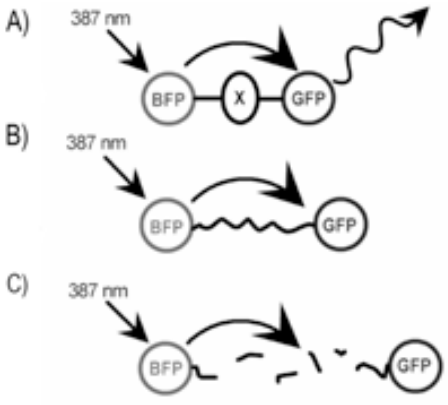

Figure 4.2 Concept of *in vivo* screening for protein folding and stability: the system is based on a ternary fusion in which a protein of interest (X) is fused to the green fluorescent protein (GFP) at its C terminus and BFP, at its N terminus. FRET from BFP to GFP is observed only when both proteins are in close proximity due to tertiary structure formation of protein X (A). In contrast, FRET from BFP to GFP is lost when protein X is either unfolded (B) or degraded by cellular proteases (C)[9].

The system is based on the simultaneous fusion of the green fluorescent protein (GFP) to the C terminus of a protein X of interest, and of blue fluorescent protein (BFP) to the N terminus. Efficient FRET from BFP to GFP in the ternary fusion protein is observed *in vivo* only when protein X is folded and brings BFP and GFP into close proximity; whereas FRET is lost when BFP and GFP are far apart due to unfolding or intracellular degradation of protein X.

4.1.1.1.2 β-galactosidase based methods

Other approaches have exploited the intracellular activity of β-galactosidase to report on protein solubility or aggregation.

One method is an adaptation of the classical β-galactosidase protein complementation assays[7] (Figure 4.3). Each monomer of the homotetrameric enzyme can be divided into two fragments, the small α-fragment and the larger ω-fragment. In the presence of α-fragment, dimers of ω-fragments achieve a dynamic equilibrium to form a tetramer with enzymatic activity. If the target protein is fused to the α-fragment, redistribution of the α-fragment into the insoluble cellular fraction will lead to a reduction in the level of β-gal activity, reporting on the solubility of the target.

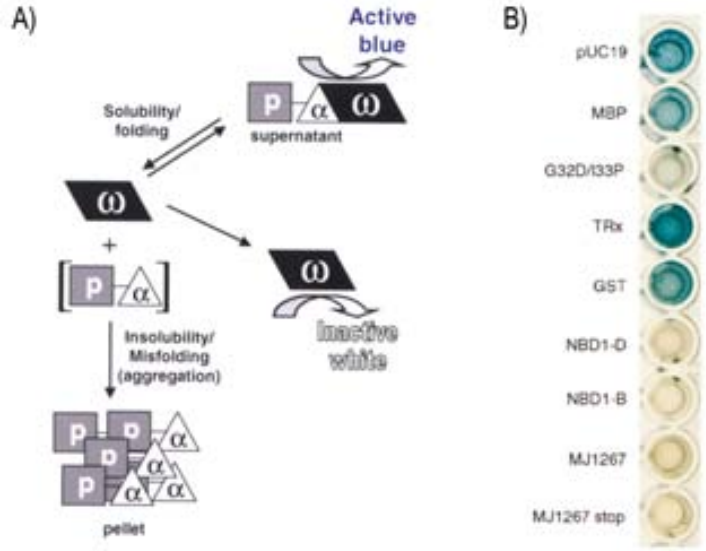

Figure 4.3 A) Schematic depiction of the complementation solubility assay: p (gray squares) represents the target protein, and α (white triangles) and ω (black trapezoids) represent each of the complementing fragments of the tetrameric β-galactosidase (β-gal). Brackets indicate the concentration dependence of complementation reflecting the availability of soluble (folded) target/α fusion. B) The complementation assay is adaptable to a 96-well plate assay format, a configuration well suited to rapid-throughput screening for compounds that affect target solubility[7].

Another interesting assay has been developed as a result of the identification of specific genes responding to protein misfolding. In this method, one of these genes promoters (*Ibp*AB) is fused to β-galactosidase in order to quantify the response of the promoter to intracellular misfolding and aggregation[10] (Figure 4.4). In this way, β-gal expression (and activity) becomes linked to protein aggregation inside the cell. Recently, it has been demonstrated the utility of this approach to evaluate different factors that modulate solubility during recombinant expression in *E.coli* [11].

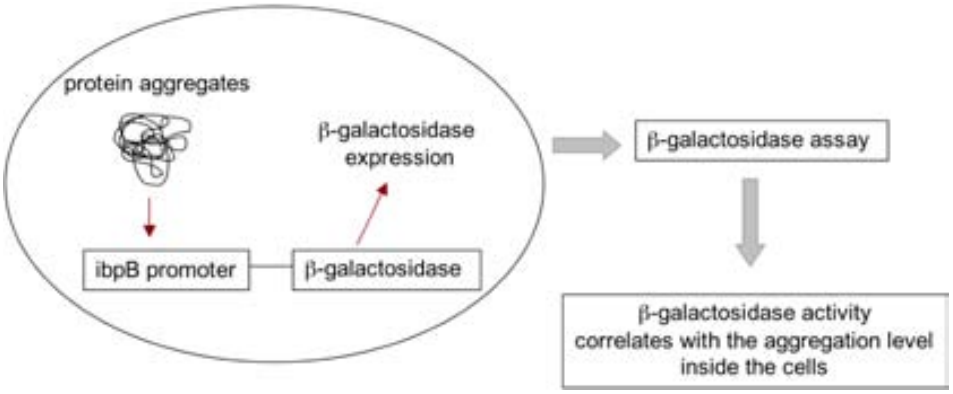

Figure 4.4 The native promoter of the β-galactosidase enzyme has been substituted by the aggregate-inducible *IbpB* promoter. As the consequence of protein aggregate crowding, the β-galactosidase accumulates and its activity serves as an indirect measurement of the aggregate amount in the cell.

4.1.1.1.3 Chloramphenicol resistance based methods

Chloramphenicol resistance has also been used as a readout to detect soluble mutants of an aggregating-prone protein *in vivo* in *E. coli*[5]. In this case, the reporter protein is chloramphenicol acetyltransferase (CAT). Resistance to high levels of chloramphenicol will be equivalent to the expression of soluble mutant fusions of the target protein. The selection can be carried out growing the cells in plates with high concentration of the antibiotic.

4.1.1.1.4 Methods based on the twin-arginine translocation pathway

The basis for this assay is the protein dependence on correct folding in order to be transported through the bacterial twin-arginine translocation (Tat) pathway[12]. In this system, a target protein is expressed as a tripartite fusion between a N-terminal Tat signal peptide and a C-terminal TEM1 β-lactamase reporter protein (Bla). If the protein folds correctly, it will be translocated through the TAT pathway to the periplasm. Due to the fact that the target protein is also fused to β-lactamase, it will confer ampicillin resistance to the

bacteria. Then, survival of *E. coli* cells expressing a Tat-targeted test protein/β-lactamase fusion on selective medium correlates with the solubility of the protein of interest. Using this assay, variants of the Alzheimer's Aβ42 peptide with an enhanced solubility could be detected and isolated from a large combinatorial library.

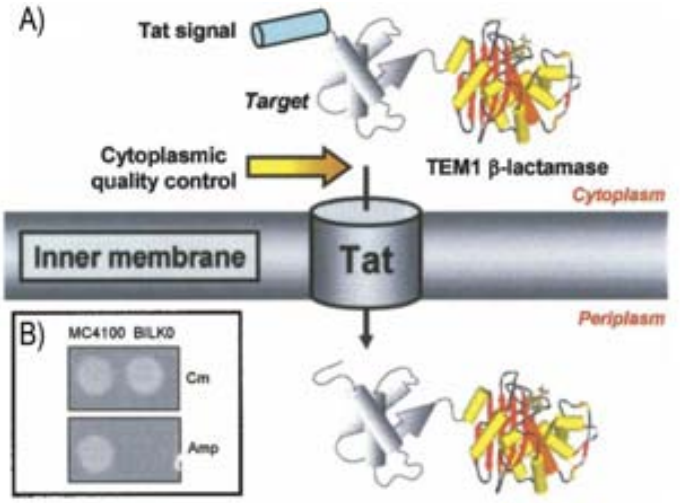

Figure 4.5 Exploiting the Tat pathway's folding quality control feature for monitoring protein solubility. (A) A tripartite fusion protein is created between a Tat signal peptide, a target protein, and the TEM1 β-lactamase protein (Bla). Discrimination between folded and misfolded target sequences is accomplished by the Tat machinery such that only correctly folded, soluble proteins are localized to the periplasm. Concomitant delivery of Bla to the *E. coli* periplasm confers an Amp resistant phenotype to cells. (B) The growth of cells on LB agar plates supplemented with either Cm (control) or 100 Amp, indicating unambiguously that Bla can be used as a Tat-specific reporter[12].

4.1.1.2 Methods to screen protein solubility and aggregation in eukaryotic cells

The majority of the developed methods have been applied in bacterial cells and surprisingly, only one approach has been available to screen solubility in yeast[13]. This assay is based on the capacity of the translational termination factor Sup35p to form self-propagating infectious amyloid aggregates. This factor manifests a prion phenotype referred to as [PSI+] and it is composed of three domains. The N-terminal domain (N) is dispensable for viability, and it is required and sufficient for the prion properties of Sup35p. While the function of the highly charged middle (M) domain remains unclear, the C-terminal RF (release factor) domain performs termination of protein translation and is essential for viability

In this assay, the activity of the termination factor Sup35p (NMRF) is conveniently assayed *in vivo* by examining the efficiency with which protein synthesis terminates at a premature stop codon (a nonsense-suppression assay). The assay uses the *ade1-14* nonsense allele. Strains carrying this mutation and bearing fully active NMRF produce

only a truncated (inactive) version of Ade1p, and as a result cannot grow on synthetic medium lacking adenine (-Ade), while they grow normally on synthetic medium supplemented with adenine (+Ade). In addition, these cells accumulate a red intermediate of the adenine synthesis pathway when grown on complex medium. However, if the efficiency of translational termination at the premature stop codon of the *ade1-14* allele is compromised, the cells gain the ability to grow on -Ade (i.e. they become Ade+) and do not accumulate red pigment. For example, cells expressing the complete Sup35p containing all three domains are white and Ade+ when NMRF is in the aggregated [*PSI*+] prion form (Fig 4.6). Cells expressing an aggregation-deficient and therefore fully functional form of Sup35p lacking the non-essential N-terminal domain (MRF) are red and Ade-. Thus, this well established system reliably distinguishes between fully active monomer, and malfunctioning aggregated forms of NMRF. It can be applied to study the aggregation of a target protein like $A\beta_{42}$. If this aggregation-prone peptide is fused to MRF (AβMRF), the yeast cells become white and grow on -Ade (they will be Ade+), whereas mutations that promote increased solubility of $A\beta_{42}$ render dark pink, Ade- cells.

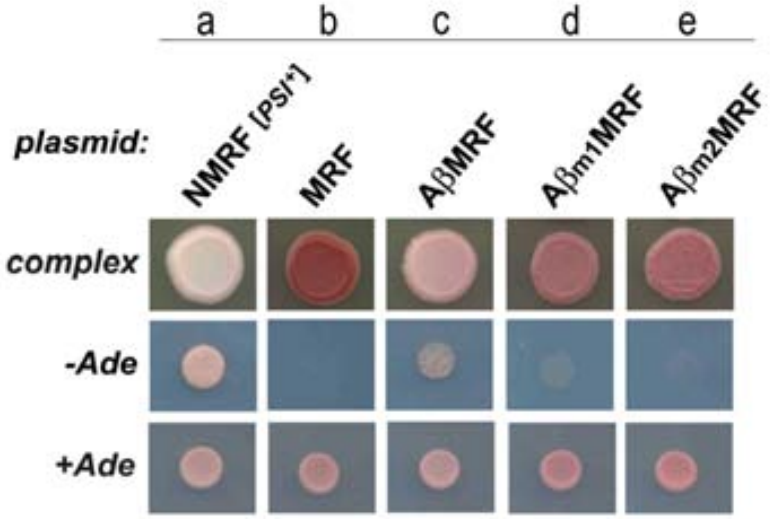

Figure 4.6 Equal numbers of ade1-14 cells containing a genomic deletion of SUP35 (*sup35*Δ), and carrying the indicated constructs (a-e) on a plasmid were grown on complex medium, or synthetic medium supplemented (+Ade) or not (-Ade) with adenine. (a) Cells with inactivated NMRF ([PSI+]) had an impaired translational termination activity, were white and grew on -Ade. (b) Cells with fully active MRF (lacking the aggregation-prone prion, N, domain), were red and failed to grow on -Ade. (c) Cells expressing AβMRF have an impaired translational termination activity, as they were white and grew on -Ade. (d, e) The translational termination activity was restored by F19, 20T (Aβm1MRF) and F19,20T/I31P (Aβm2MRF) mutations in the Aβ42 region of the fusion protein, making the cells dark pink and preventing their growth on -Ade[13].

4.1.2 Linking DHFR activity to protein aggregation.

Protein aggregation in eukaryotic environments is receiving increasing attention due to its relevance in biomedicine and biotechnology. Therefore, we decided to develop a method that could couple an easily measurable phenotype like cell survival to protein aggregation using yeast as a model of eukaryotic organism. It is based in the fusion of the target protein to the human dihydrofolate reductase (h-DHFR)[14]. DHFR is a key enzyme in thymidine synthesis that catalyses the reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate with NADPH as a coenzyme. This enzyme has been used previously in protein fragment complementation assays[15] and in the three-hybrid method[16].

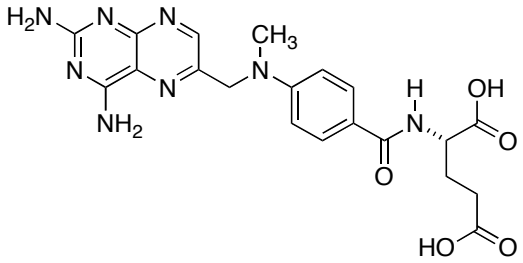

Figure 4.7 Formula of Methotrexate (MTX).

Prokaryotic and eukaryotic DHFRs are central to cellular one-carbon metabolism and are absolutely required for cell survival. And its activity can be specifically inhibited by the drug methotrexate (MTX). In the present work, we sought to exploit our observation that yeast cells can become insensitive to MTX if they express high levels of h-DHFR.

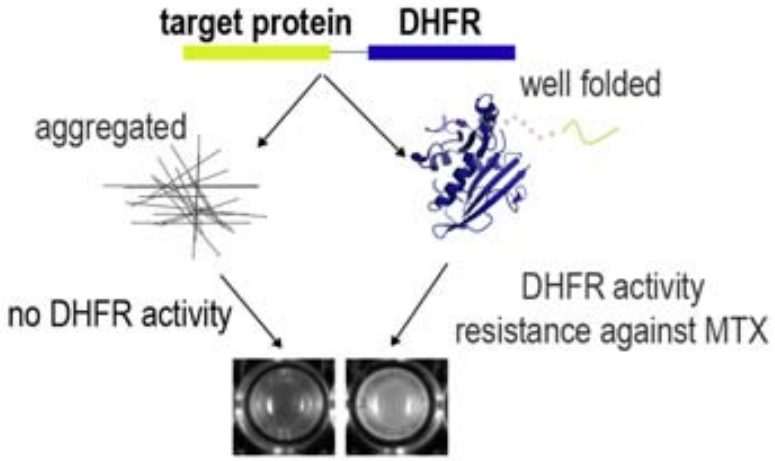

Figure 4.8. Schematic depiction of the properties of DHFR fusion proteins. Aggregating proteins form insoluble deposits (left) and prevents DHFR from being functional. However, mutants in the target protein that prevent aggregation enable DHFR to fold correctly and perform its activity conferring MTX resistance to the cells (right).

This enzyme is a very soluble protein and, in our approach, it is expressed at concentrations that allow cell survival in MTX concentrations that otherwise would be lethal. Therefore, we hypothesize that the fusion of h-DHFR to aggregation-prone polypeptides might inactivate the enzyme and render the cells expressing these kinds of fusions MTX susceptible.

In principle, this might allow for the design of an easy high-throughput method to monitor protein folding inside yeast based on the reversal of MTX growth inhibition. And it might permit to monitor the effect of different factors (mutations, genetic backgrounds, chemical compounds or growth conditions) over protein aggregation. To demonstrate the applicability of the assay, different aggregation prone proteins (Alzheimer's amyloid β (Aβ) peptide, polyglutamine expansions in the huntingtin protein (poliQ) and alpha-synuclein (α-Syn) will be used as models.

4.2 Objectives

→ Development of a method to detect protein aggregation using simultaneously yeast survival and fluorescence emission as reporter signals. As a test models different aggregation-prone proteins involved in neurodegenerative diseases will be used: Alzheimer's amyloid β (Aβ) peptide, polyglutamine expansions in the huntingtin protein (poliQ) and alpha-synuclein (α-Syn).

→Study the ability of the method to detect the effect of different factors that modulate protein aggregation *in vivo*:

- Chemical compounds
- Overexpression of chaperones
- Deletion of chaperones
- Growth conditions (i.e. temperature)

4.3 Experimental procedures

4.3.1 Reagents and strains

In this study we used different yeast strains depending on the assay to be performed. The strain FY834 was used in the preliminary assays and in the study related to overexpression of chaperones whereas drug testing was done in the drug permeable strain *erg6*Δ in the BY4741 parental background. Strains with a deletion of specific chaperone were provided by Euroscarf and were also based in the BY4741 strain. Methotrexate (MTX), sulfanilamide, quercetin and Congo Red were purchased from Sigma Aldrich.

4.3.2 Plasmid construction

All the polypeptides (Aβ42, polyQ expansions and α-synuclein as well as their variants) were fused to DHFR by a linker with the sequence GSAGSAAGSG. The table 4.1 summarizes the protein fusions designed for the experiments. In order to obtain the different protein fusions, it was used the strategy described in the general Experimental Procedures section. DHFR was also cloned in the same plasmids (Figure 4.9). Plasmids encoding the polyQ expansions and α-synuclein variants were a kind gift of Dr. Lindquist[17].

Table 4.1 Design of the fusion proteins

Target proteins	Fusion proteins	Plasmid	Restriction sites	Selectable markers
Aβ42	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	pESC	*Cla*I, *Bgl*II	Trp
Aβ42 F19D	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	pESC	*Cla*I, *Bgl*II	Trp
Q25	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p416	*Sal*I, *BamH*I	Ura
Q72	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p416	*Sal*I, *BamH*I	Ura
Q103	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p416	*Sal*I, *BamH*I	Ura
α-syn wt	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p426	*Spe*I, *Xho*I	Ura
α-syn A30P	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p426	*Spe*I, *Xho*I	Ura
α-syn A53T	DHFR	pESC	*Cla*I, *Bgl*II	Ura/Trp
	GFP	p426	*Spe*I, *Xho*I	Ura

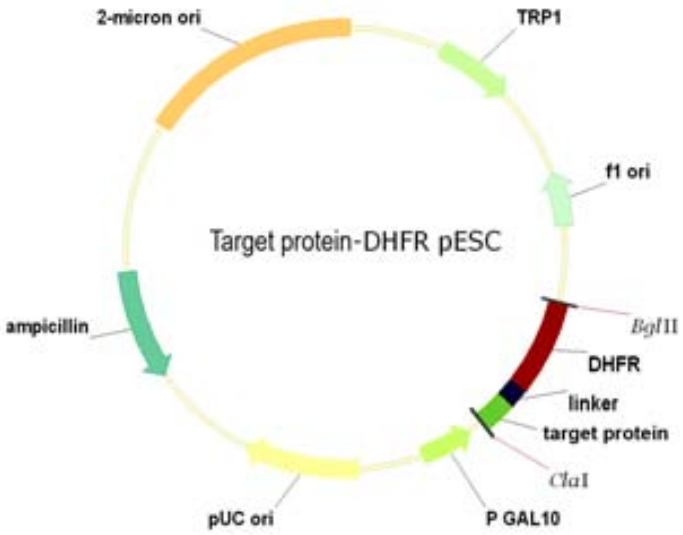

Figure 4.9 Scheme of the plasmid used in this assay. The target protein fused to DHFR was cloned between the restriction sites *Cla*I i *Bgl*II in the plasmid pESC

Plasmids encoding for the different yeast chaperones[18] are listed in the following table and were a kind gift of Dr. Lindquist.

Table 4.2 Plasmids encoding for the chaperones used in this study

Chaperone	Plasmid	Restriction sites	Selectable markers
Ssa1	pRS425	*BamH*I, *Hind*III	Leu
Hsp104	p2HG	*BamH*I	His
Hsp82	pTGpd	*BamH*I	Trp
Sis1	pTV3	*BamH*I	Trp

4.3.3 Growth curves

Cells were grown overnight at 30°C in a selective synthetic complete (SC) media containing raffinose. They were inoculated at OD_{600} of 0.02 in minimal galactose medium containing the appropriate MTX concentration (20-100 μM) and 1 mM of sulfanilamide (compound that enhances sensibility against MTX[19]). Growth was followed measuring OD_{600} using Cary 400Bio spectrophotometer.

In drug screenings, the protocol was the same being the cells inoculated in a minimal galactose medium containing 20 μM MTX, 1mM sulfanilamide and the tested compound: quercetin (30 μM) or CR (10 μM). In this case, a previous incubation before the addition of galactose was performed to assure the presence of the compound in the cell

when the fusion protein started to be expressed. This way, before the induction with galactose, quercetin or Congo Red were added at the same concentration as the final assays during 90 minutes. After this period of time, the cells were inoculated in a media containing galactose, MTX and sulfanilamide.

It has to be taken into account that some compounds were dissolved in DMSO. To control the effects of DMSO on the cells, equal amounts of this solvent were added in the negative control. After 30 hours from the induction, the cultures were diluted 1/100 and its OD_{600} was measured. In all the experiments OD_{600} measurement was the average of triplicate measurements from several independent transformants.

4.3.4 Spotting assays

Yeast cells were grown overnight in minimal media at 30°C containing raffinose. Cell density was determined by measuring the OD_{600} and cells were diluted to a final OD_{600} of 0.18 using PBS. Afterwards, 10 μl of each dilution (1/10, 1/10 and 1/100) was spotted in plates of selective synthetic complete media with galactose, MTX (20 μM) and sulfanilamide (1 mM). The Petri dishes were incubated for 48 hours at 30°C. Images of the plates were taken using the molecular imager Gel Doc XR system from Bio-Rad.

4.3.5 Fluorescent detection of h-DHFR with fMTX

Yeast cells transformed with plasmids encoding the target proteins fused to DHFR were grown overnight in minimal media at 30°C containing raffinose. They were inoculated at OD_{600} of 0.02 in minimal galactose medium. Yeast cells were grown during 24 hours and then, they were incubated with sulfanilamide (1 mM) and 10 μM of MTX labeled with the fluorescent molecule Alexa (Invitrogen) for another period of 24 hours. Afterwards, the medium was removed and the cells were washed with PBS and reincubated for 30 min in the selective medium to allow for efflux of unbound fMTX. Then, the cells were centrifuged (1300x*g* for 5 min) and washed with 50 mM Tris–HCl pH=7 five times. Cells were visualized by microscopy as it has been explained in the general Experimental Procedures section.

4.4 Results

4.4.1 Development of the method using Alzheimer's Aß42 peptide as a proof of principle

4.4.1.1 Alzheimer's peptide Aβ42 forms inclusions in yeast

The main component of AD lesions is the hydrophobic polypeptide Aβ(1–42)[20]. In previous works, our group has shown that mutation of Phenylalanine 19 to Aspartate (F19D) abolishes the amyloidogenicity of Aβ42 *in vitro*[21] and also reduces Aβ42 aggregation propensity inside *E.coli*[22]. These two extreme behaviours make of peptide Aβ42 and its F19D point mutant a promising pair of targets to explore factors influencing protein aggregation in eukaryotic backgrounds, particularly in yeast.

In order to study the aggregation behaviour of the proteins in an eukaryotic background, both variants were expressed in *S. cerevisiae* as fusion proteins with GFP. While the GFP-fusion with mutant Aβ42(F19D) did not aggregate intracellularly and its fluorescence was distributed diffusely throughout the cell, the Aβ42-GFP protein fluorescence was concentrated in a single large aggregate in a juxtanuclear position, as shown by co-staining with Hoechst (Figure 4.10). Immunoblotting of total cellular protein indicated that both proteins were expressed at similar levels (data not shown), demonstrating that the degree of coalescence exhibited by Aβ42 forms is more dependent on their sequence than on the level of protein expressed.

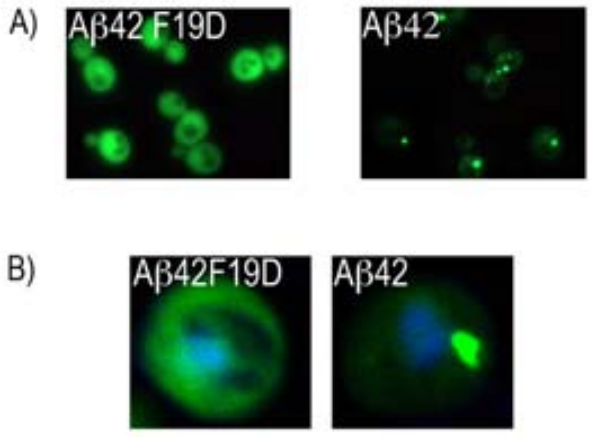

Figure 4.10 A) Visualization of intracellular wild-type and Aβ42(F19D) distribution with GFP expressed in *S. cerevisiae* B) Co-staining of the cell nucleus with Hoechst (blue). The aggregated Aβ42-GFP has a juxtanuclear position.

4.4.1.2 Intracellular DHFR activity allows linking protein aggregation and cell growth

The lethality of MTX on *S.cerevisiae* FY384 cells above certain concentrations (such as 25 μM with 1 mM of sulfanilamide)[23] can be overcome by transformation with a plasmid encoding human DHFR (h-DHFR) under the control of Gal10 promoter. Different degrees of sensitivity to MTX may thus be correlated with the intracellular activity of the heterologously expressed enzyme, which is likely to vary depending on its expression alone, as a fusion with soluble molecules or as a fusion with an aggregation-prone polypeptide. In the latter case, the fused protein would promote, at least in part, its deposition lowering the intracellular DHFR activity and causing a higher sensitivity to MTX. Since the intrinsic aggregation propensities of Aβ42 and Aβ42(F19D) peptides determined the fate of the fused GFP within yeast, both peptides were fused to h-DHFR. The differential growth abilities of cells expressing these fusions were compared with that of cells expressing h-DHFR alone.

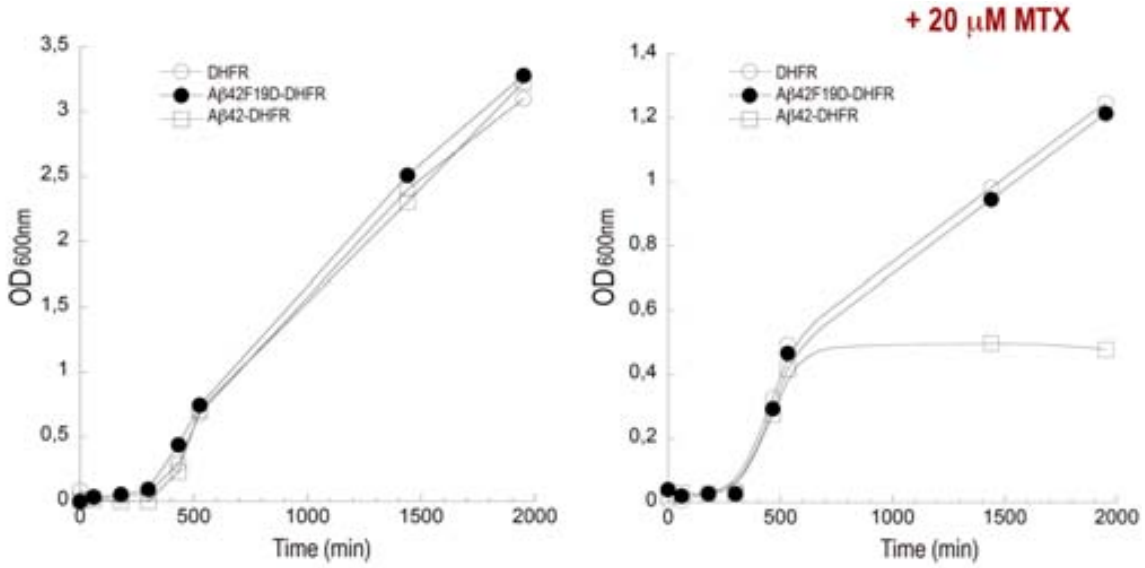

Figure 4.11 Growth assays of FY834 yeast cells expressing DHFR (empty circles), peptide Aβ42-DHFR (empty squares) or peptide Aβ42(F19D)-DHFR (solid circles) in the presence of 0 μM (left) and 20 μM of MTX (right).

In the presence of MTX, the growth of yeast expressing Aβ42-DHFR became stationary after a short period of time; whereas cells expressing h-DHFR or peptide Aβ42(F19D)-DHFR displayed clearly higher and similar growth rates (Figure 4.11). In the absence of MTX, no significant differences in growth rates could be observed indicating that the growth divergence could not be attributed to a differential toxicity of the expressed proteins.

In order to assess if the detected differences in viability were caused by the dissimilar solubility of h-DHFR due to its fusion to the Aβ42 variants, a filter trap assay was performed. As it could be inferred from the target protein fusions with GFP, protein aggregates were only detected in the case of yeast cells expressing the wild-type form fused to DHFR, whereas Aβ42(F19D)-DHFR or DHFR alone did not present any aggregate within the cell.

Figure 4.12 Filter trap assay of cells expressing DHFR, Aβ42(F19D)-DHFR or Aβ42-DHFR. Protein aggregates were detected by immunoblot analysis using specific antibody against DHFR.

The system should allow also monitoring the influence of culture conditions on polypeptide aggregation. It is assumed that high temperatures promote *in vivo* and *in vitro* protein aggregation by reinforcing hydrophobic intermolecular interactions among polypeptides[24]. Increasing the growth temperature from 30°C to 37°C resulted in a decrease in viability of cells expressing Aβ42-DHFR at all the MTX concentrations assayed, but also of cells expressing Aβ42(F19D)-DHFR at high MTX concentrations (Figure 4.13). No such phenotypic effect was observed in cells expressing h-DHFR alone. This result suggested that temperature was specifically increasing, directly or indirectly, the aggregation propensity of the Aβ moiety and that the system was sensitive enough to detect such effect.

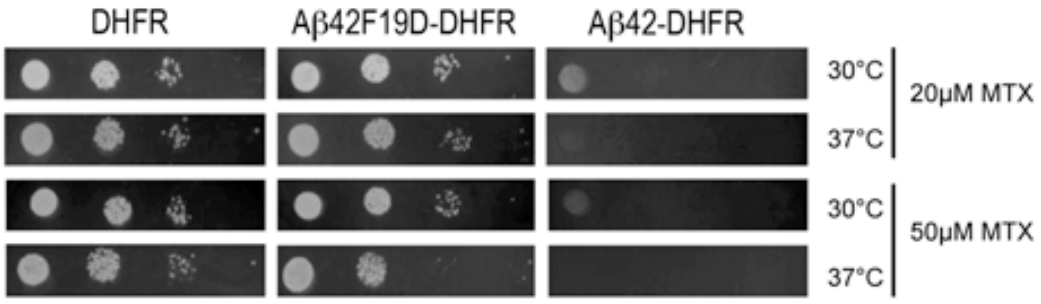

Figure 4.13 Cell viability (spotting) assays for yeast expressing DHFR, peptide Aβ42-DHFR or peptide Aβ42 F19D-DHFR at different temperatures and MTX concentrations. Four-fold serial dilutions starting with equal number of cells are shown.

Overall, these results indicated that h-DHFR activity, as reflected in cell growth, could be used as a reporter to monitor the influence of both intrinsic and extrinsic factors on the aggregation of a given polypeptide.

4.4.2 Detection of the cellular localization of h-DHFR fusions

The reliability of the indirect methods to study *in vivo* aggregation has to be always confirmed by measuring the solubility of the target protein using another approach. In our case, one strategy was the fusion of the same target to GFP. Then, it was expressed in the same eukaryotic background and the aggregation state of the target protein was imaged by monitoring the localization of the fluorescence. Alternatively, the solubility of the target protein fused to DHFR was also detected by filter trap assay.

In this context, we sought to exploit the ability of h-DHFR to bind MTX with high affinity in a 1:1 complex in order to visualize the presence of the active enzyme within the cells. With this purpose, we used a version of the inhibitor labeled with a fluorescent compound (fMTX). It has been proved before that fMTX is retained into the cells through this binding to DHFR, whereas the unbound fMTX is actively and rapidly transported outside[25]. In addition, the interaction of fMTX and DHFR results in a 4.5-fold increase in quantum yield. Thus, bound fMTX, and by inference the properly folded DHFR, could be monitored by fluorescence microscopy.

The data obtained by imaging the location of fMTX in cells expressing h-DHFR, Aβ42-DHFR and Aβ42(F19D)-DHFR were absolutely coincident with those obtained using GFP as a tag.

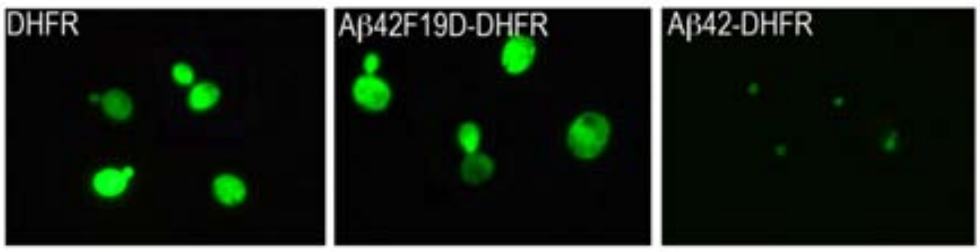

Figure 4.14 The addition of a fluorescent inhibitor of DHFR (fMTX) enables the visualization of the intracellular distribution of wild-type and F19D mutant Aβ42 fused to DHFR.

While cells expressing h-DHFR and Aβ42(F19D)-DHFR had no aggregates and their fluorescence was high and distributed diffusely throughout the cell, the fluorescence of the ones expressing Aβ42-DHFR was lower, indicating less amount of folded enzyme, and it was concentrated in a single large aggregate per cell. Therefore, from the results in the previous sections, it could be deduced that the sequence of the target protein determined the aggregated state of the enzyme. Also, the solubility of the target protein could be easily and simultaneously assayed by measuring cellular viability in the presence of MTX as well as by imaging the location of the fluorescent inhibitor inside the cells.

4.4.3 Application of the method to another aggregating proteins

To test the general applicability of the method, other aggregation-prone proteins should be fused to DHFR verifying the link between protein aggregation and yeast survival. With this purpose, polyglutamine expansions[26, 27] and α-synuclein[28, 29] protein were chosen due to their aggregation propensity and their implication in human pathologies: Huntington's and Parkinson's diseases, respectively.

4.4.3.1 Polyglutamine expansions

In the case of polyglutamine expansions, it has been previously demonstrated that their aggregation propensity depends on their length[18]. As it was performed in the previous case, the distribution of the different variants in living yeast cells was studied through their fusion to GFP (Figure 4.15). Whereas cells expressing extensions of 25 glutamines (Q25) displayed fluorescence diffusely distributed trough the cell, yeasts expressing 72 (Q72) or 103 glutamines (Q103) exhibited various fluorescent foci. Also, it was noticeable that the number and size of foci increased with the length of polyQ expansions.

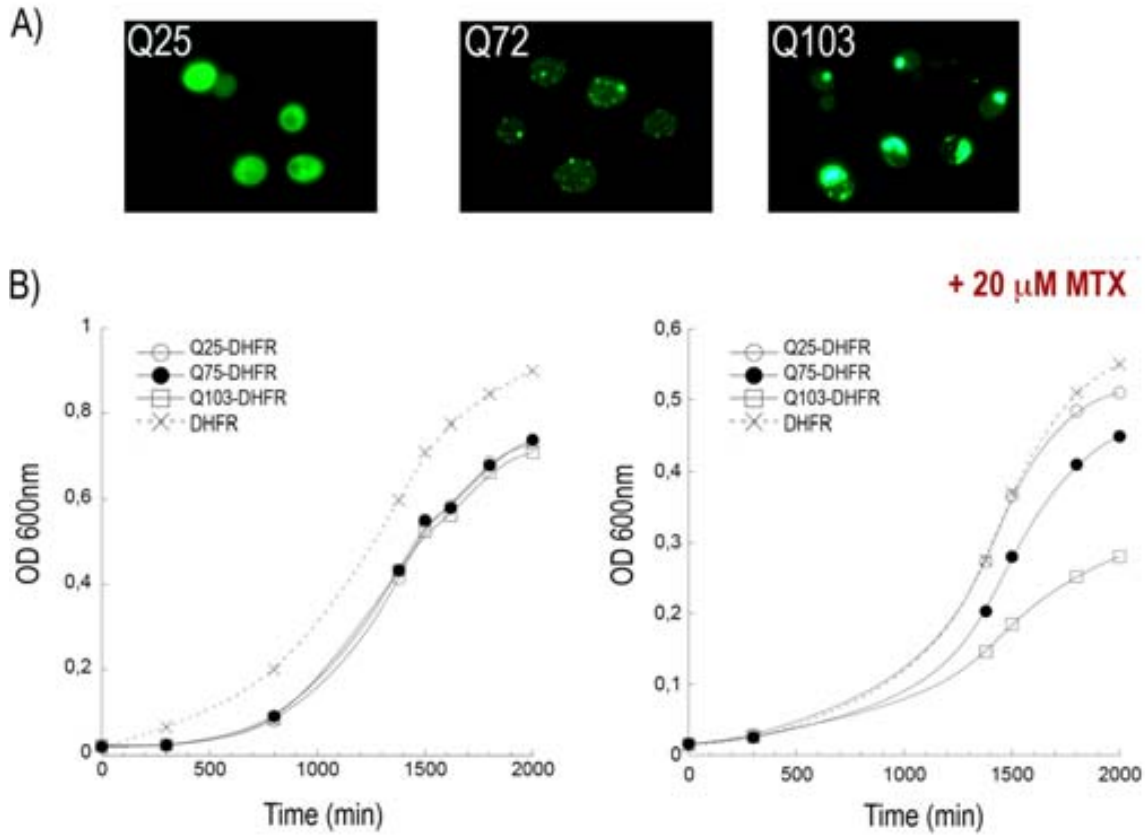

Figure 4.15 A) Fluorescence microscopy of yeast cells expressing different polyQ expansions (Q25, Q72 or Q103) fused to GFP B) Growth assays of yeast cells expressing the different polyQ expansions fused to DHFR in the presence of 20 μM MTX.

In the absence of MTX, the fusion of poly-Q fragments to DHFR reduced cell viability. Nevertheless, all polyQ repeats displayed similar effect on viability independently of their length.

In the presence of MTX, the yeast growth was inversely proportional to polyQ expansions length. In other words, yeasts expressing Q25-DHFR exhibited the highest growth rate (Figure 4.15). Therefore, the survival of yeast cells expressing different polyQ fused to DHFR correlated with the observed solubility of the GFP fusions. This link indicated that, under the conditions of the assay, the presence of properly folded DHFR moiety determined the cell growth.

The use of fMTX provided results in excellent agreement with these data. In cells expressing Q25-DHFR, the enzyme was soluble because the fluorescence was high and homogenously distributed in the cell. On the contrary, the presence of Q72 or Q103 expansions strongly reduced the fluorescence emission and the active enzyme became mainly located into aggregates.

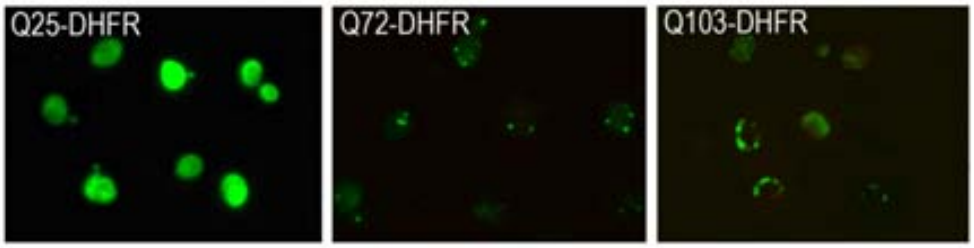

Figure 4.16 The addition of a fluorescent inhibitor of DHFR (fMTX) enables the visualization of the intracellular distribution of the different poliQ expansions fused to DHFR.

4.4.3.2 α-Synuclein

The protein α-synuclein (α-Syn) forms the fibrous portion of Lewy Bodies, cytoplasmatic inclusions present in Parkinson's disease (PD). Two rare early-onset forms of PD are linked with mutations in the α-Syn gene: A53T and A30P[30]. Both variants have distinct physical properties: A53T is accumulated at the plasma membrane or in cytoplasmatic foci like wild-type α-Syn; whereas A30P is dispersed through the cell[31, 32].

When all the α-Syn variants were fused to GFP, the above described phenotypes were reproduced (Figure 4.17). When the variants were fused to DHFR and expressed in yeast in the absence of MTX, cell viability was reduced in a similar extent indicating certain toxicity of the α-Syn gene when it was expressed in yeast.

In the presence of MTX, the expression of α-SynA30P-DHFR allowed a cell growth rate very close to the one of cells expressing h-DHFR alone, whereas the α-SynA53T-DHFR or α-Syn-DHFR variants displayed a significantly reduced viability.

Overall, the survival of yeast expressing the α-Syn variants fused to DHFR correlated well with the observed solubility of the GFP-fusions. This fact demonstrated again that the presence of properly folded DHFR controlled cell growth under the conditions of the assay.

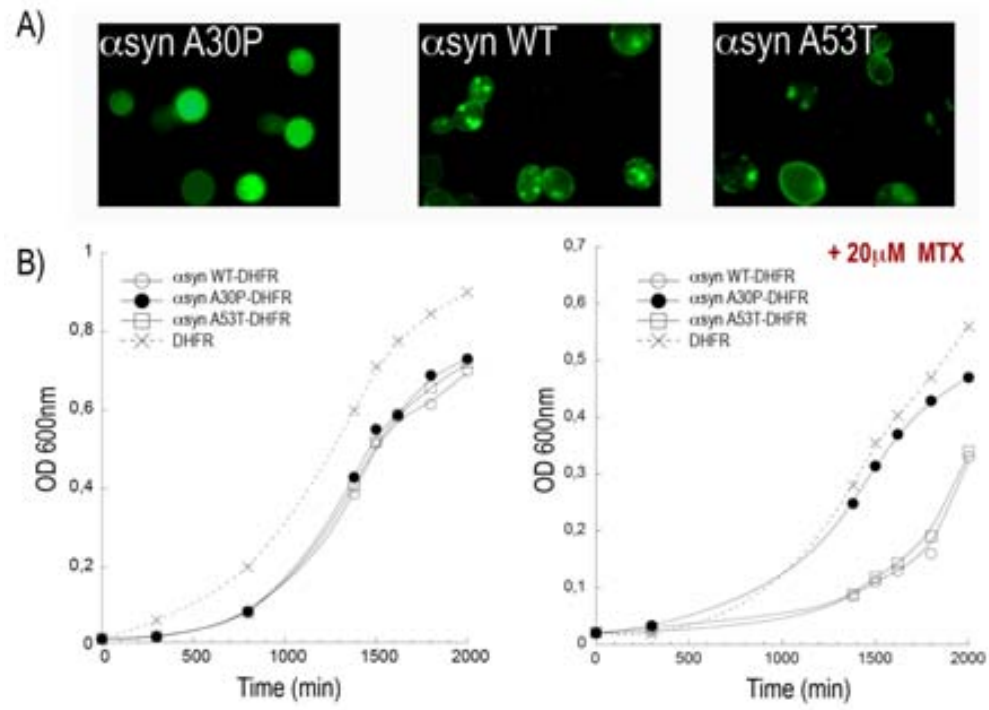

Figure 4.17 A) Fluorescence microscopy of yeast cells expressing different α-synuclein variants fused to GFP.B) Growth assays of yeast cells expressing α-synuclein variants fused to DHFR in the presence of 100 μM MTX.

Accordingly to all these data, the fMTX fluorescence of cells expressing α-SynA30P-DHFR was higher and no aggregates were observed. On the other hand, punctuated nuclei, usually close to the cytoplasmic membrane, were observed for the wild-type and A53T mutant.

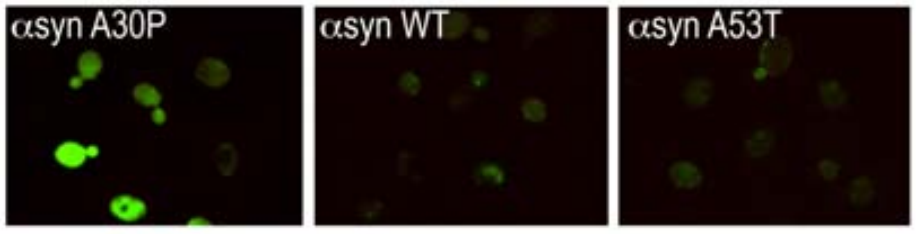

Figure 4.18 The addition of a fluorescent inhibitor of DHFR (fMTX) enables the visualization of the intracellular distribution of the different α-synuclein variants fused to DHFR.

4.4.4 Characterization of small-chemical compounds effect on intracellular Aβ42 peptide aggregation

The identification of small chemical compounds that affect the size or the number of protein inclusions inside cells is one of the approaches towards therapeutic intervention against depositional diseases[33]. We wanted to explore whether the previously outlined method could be useful to screen for such substances. As test case, *erg6*Δ yeast strain cells expressing h-DHFR, and Aβ42-DHFR were grown in the presence of selected concentrations of the compounds quercetin and Congo Red (CR), both previously shown to *in vitro* bind Aβ aggregates.

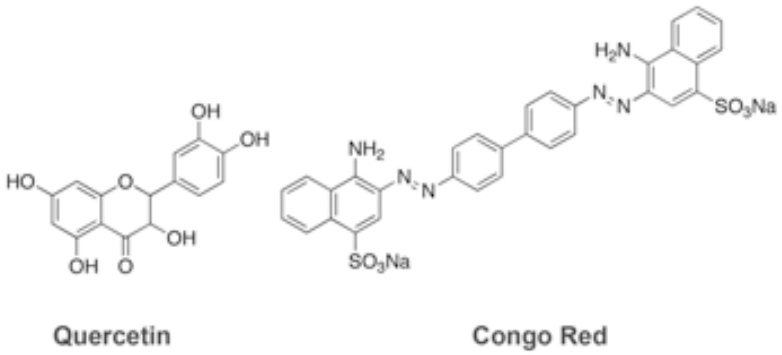

Figure 4.19 Structure of compounds previously described as inhibitors of the protein aggregation.

The erg6 mutation inhibits ergosterol biosynthesis, which enhances membrane fluidity and permeability to various chemical compounds[34]. This results in a four-fold higher sensitivity to MTX when compared to the FY384 strain and in a concomitant decrease in the viability of *erg6*Δ cells expressing Aβ42-DHFR relative to those expressing h-DHFR. Thus, to specifically monitor the effects of compounds on the viability of cells expressing the Aβ42 fusion, growth was always referenced to that of cells expressing h-DHFR under the same conditions. Quercetin is a flavonoid compound shown to inhibit *in vitro* Aβ fibril formation[35] and to reduce the toxicity of Aβ fragments in neuroblastoma cells[36]. In the presence of MTX, cells expressing peptide Aβ42-DHFR displayed a growth restoration dependent on quercetin concentration (Figure 4.20 A). Interestingly enough, this effect was also reported in a yeast-based model of α-synucleinopathy at similar quercetin concentrations[37].

On the other hand, despite the fact that CR function as amyloid ligand was widely reported[38], no significant effect could be observed at 10 μM (Figure 4.20 A). Strikingly,

this CR concentration was shown to reduce significantly the aggregation of huntingtin with expanded polyglutamine (polyQ) inside mammalian cells[33].

The different activity of both substances on Aβ42 aggregation in yeast could be rationalized by analyzing fluorescence microscopy images of cells expressing peptide Aβ42 fused to GFP in media containing selected concentrations of quercetin or CR (Figure 4.20 B). Cells grown in the presence of 30 μM quercetin displayed a unique aggregate per cell with a smaller average diameter (0.6 μm) than those formed in the absence of the compound (1.0 μm), suggesting that it effectively targets *in vivo* intracellular protein insolubility. Cells treated with 10 μM CR presented one big fluorescent focus with an average diameter of 1.1 μm as well as many smaller ones. Thus, although CR might interfere with the aggregation process of Aβ42 in yeast, it did not increase the effective soluble protein concentration and rather promoted the appearance of new aggregation foci.

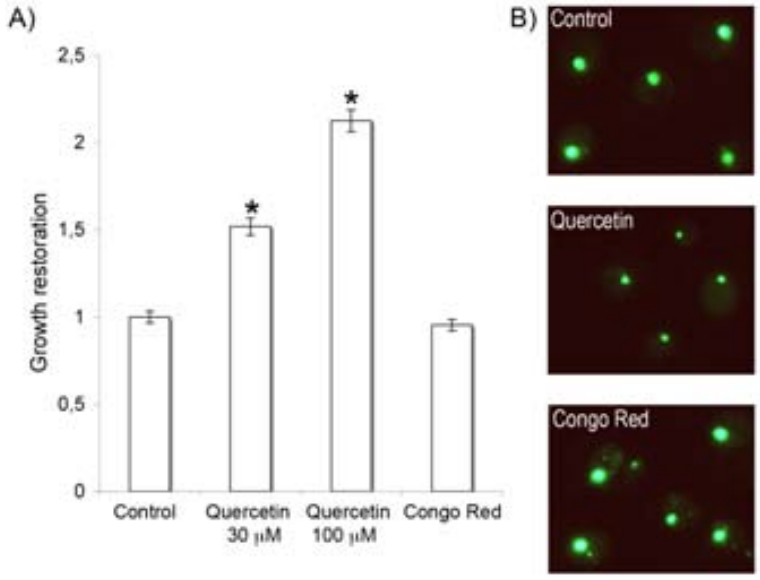

Figure 4.20 A) Growth restoration of *erg6*Δ yeast cells expressing in Aβ42-DHFR in the presence of 20 μM MTX, 1mM sulfanilamide and selected concentrations of quercetin (30 μM and 100 μM) and CR (10 μM). Growth is normalized to 0 μM compound. Significant differences are marked with an asterisk. B) Fluorescence microscopic assessment of Aβ42–GFP aggregation in control or compound treated *erg6*Δ cells.

4.4.5 Molecular chaperones influence intracellular Aβ42 aggregation

It has been shown that changes in the levels of chaperones in the cell modulate protein aggregation in yeast models of both Huntington's and Parkinson's diseases[18, 39]. To evaluate whether our method could be also sensitive to differences in the composition and concentration of proteins involved in the cellular folding machinery, a set of chaperone knockouts and chaperone overexpressing strains producing h-DHFR and Aβ42-DHFR

were engineered. In the table 4.3, there is a classification off the all chaperones used in this section according to their family.

Table 4.3 List of the chaperones studied in this section

Family	Members
Hsp100	Hsp104
Hsp90	Hsc82, Hsp82
Hsp70	Ssa1, Ssa2, Ssa3, Ssa4
Hsp40	Ydj1, Sis1
Small heat shock proteins	Hsp26, Hsp42

4.4.5.1 Overexpression of chaperones

Overexpression of all tested chaperones promoted an increase in the viability of cells expressing Aβ42-DHFR in liquid media containing MTX. The chaperones with greater effect were in decreasing order: Sis1, Hsp104, Hsp82 and Ssa1 with Ydj1 having a rather moderate impact on yeast survival (Figure 4.21 A). However, in parallel spotting assays in the presence of MTX only Sis1 and Hsp82 overexpression promoted clear increase in the cell survival. (Figure 4.21 B).

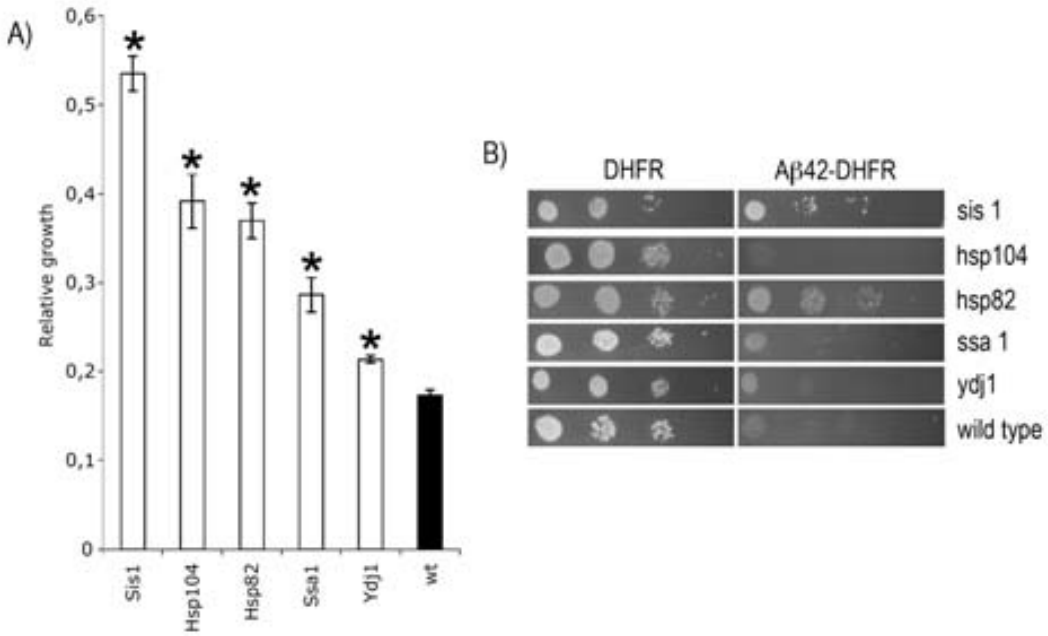

Figure 4.21 A) Growth of yeast FY834 strains overexpressing a chaperone and co-expressing peptide Aβ42-DHFR in the presence of MTX in liquid media. Growth is normalized to the same strain expressing the corresponding chaperone and DHFR. Significant differences are marked with an asterisk. B) Cell viability (spotting) assays for the different strains overexpressing a chaperone and DHFR or Aβ42-DHFR. In each case, four-fold serial dilutions starting with equal number of cells are shown.

Western blot analysis against Aβ42 moiety demonstrated the formation of SDS-stable oligomers[13] (Figure 4.22). These oligomers were reminiscent of the ones formed by the Aβ42 peptide *in vitro*, in mammalian cell culture, and in the human brain. Interestingly, the expression in yeast of Aβ42 fused to the MRF domain of the Sup35 prion resulted in a very similar distribution of SDS-stable oligomers[13]. Overall, the levels and distribution of Aβ42-DHFR were not affected by chaperone overexpression suggesting that they target larger aggregates causing the observed differences in viability.

Accordingly, fluorescence microscopy images of yeast expressing Aβ42-GFP showed that chaperones strongly affected the distribution of fluorescence (Figure 4.22). Cells overexpressing Hsp104 dramatically increased the number and intensity of fluorescent inclusions as well as of background fluorescence. Sis1 overexpressing cells presented generally one big aggregate and several minor aggregation nuclei with high background fluorescence. On the other hand, overexpression of Hsp82 and Ssa1 resulted in a phenotype similar to the one reported in the presence of quercetin, with only a small fluorescent focus per cell. And the overexpression of Ydj caused the appearance of several small foci per cell.

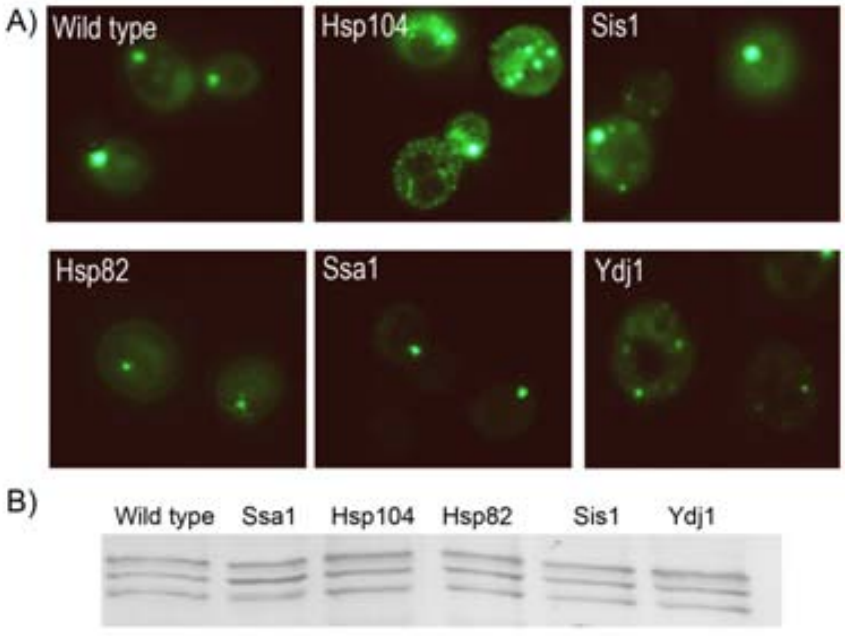

Figure 4.22 A) Fluorescence microscopy of different strains overexpressing a chaperone and peptide Aβ42 fused to GFP. B) Western Blot analysis of cells co-expressing different chaperones and Aβ42-DHFR. The concentration of Aβ42-DHFR did not differ due to the expression of one chaperone. The different bands were correlated to the different oligomerization degrees of Aβ42-DHFR.

4.4.5.2 Knockouts in chaperones

To test if the depletion of chaperones also affected the intracellular aggregation of Aβ42, strains with a knockout in one specific chaperone gene were transformed with plasmids encoding h-DHRF or peptide Aβ42-DHFR and their growth was measured in the presence of MTX. Surprisingly, depletion of some chaperones resulted in increased cell growth relative to the wild-type strain (*hsp35*Δ, *hsp42*Δ and *hsp26*Δ) whereas others resulted in decreased cell density (*ssa1*Δ, *ssa2*Δ, *hsc82*Δ or *hsp104*Δ) and one strain did not grow (*ssa4*Δ).

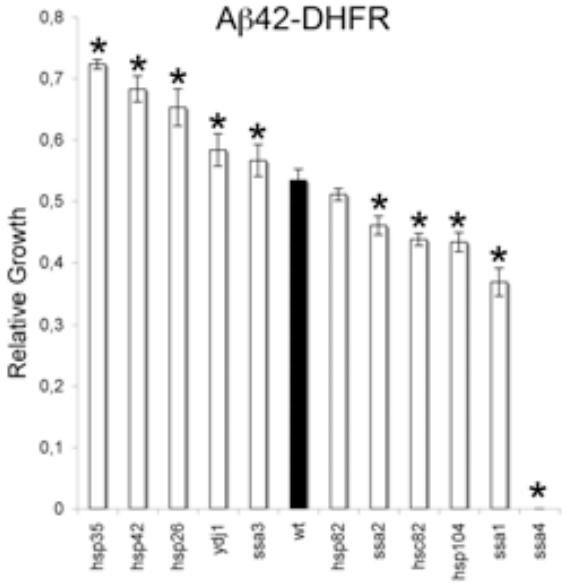

Figure 4.23 Growth of yeast FY834 strains with a deletion in one chaperone and expressing peptide Aβ42-DHFR in the presence of MTX in liquid media. Growth is normalized to the same strain expressing the corresponding chaperone and DHFR. Significant differences are marked by an asterisk.

Western blot analysis indicated that no significant differences in global expression of Aβ42-DHFR existed between knockouts but some chaperones affected oligomer distribution.

To test whether the deletion of chaperones influenced the intracellular distribution of peptide Aβ42 aggregates, some knockout strains were transformed with this peptide fused to GFP (Figure 4.24). *hsp35*Δ cells exhibited a single fluorescent focus with a smaller diameter (0.6 μm) than cells with wild-type background. On the other hand, deletions of Ssa1, Ssa2 or Hsp104 resulted in the appearance of more than one Aβ42-GFP aggregate per cell. It is worth to notice that the *ssa4*Δ strain, which has lost the ability to grow in the presence of MTX albeit overexpressing h-DHFR, displayed a large number of

Aβ42-GFP aggregates indicating an anomalous increased aggregation of polypeptides in the absence of this gene.

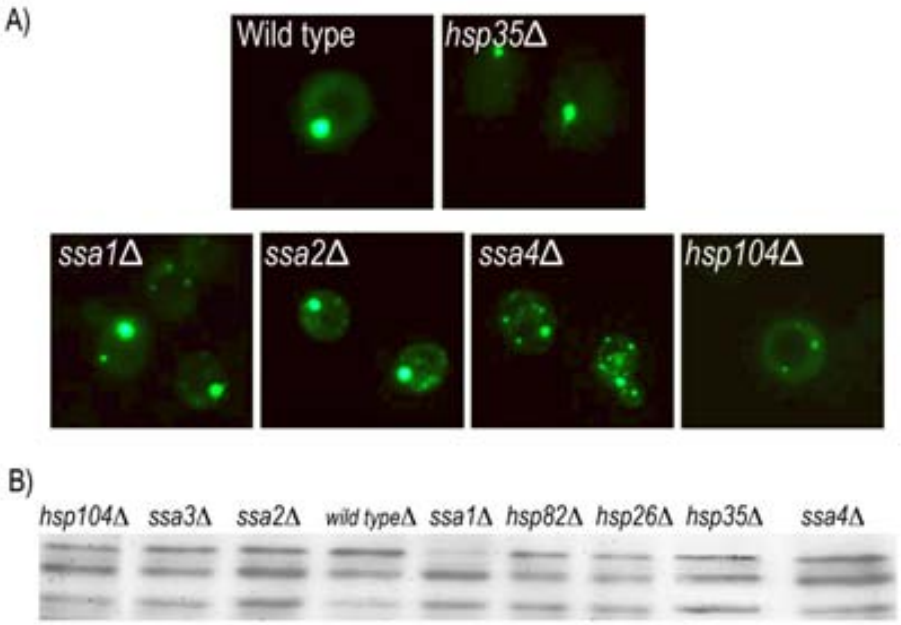

Figure 4.24 A) Fluorescence microscopy of different strains with chaperone knockouts and expressing peptide Aβ42 fused to GFP. B) Western Blot analysis of different yeast strains with a deletion in a chaperone and cells expressing Aβ42-DHFR. The detected bands correlated with the different oligomerization degrees of Aβ42-DHFR.

4.5 Discussion

Yeast is an organism that shares with higher eukaryotes basic cellular mechanisms such replication, protein folding or metabolism[40]. Interestingly enough, several recent studies suggest that the response pathways related to depositional diseases could be present in yeast. Although the modeling of Huntington's and Parkinson's disorders in this organism has provided significant insights into the pathogenesis of these diseases[17, 40-42], the study of Alzheimer's disease (AD) has received less attention, probably because the toxic effect has been traditionally thought to be mediated from the extracellular space. Nevertheless, increasing evidences suggest that intracellular Aβ peptide accumulation is an early step in both AD[43-45] and the neurodegeneration associated with Down syndrome[46]. Accordingly, *Drosophila* and *C.elegans* models to study intracellular Aβ aggregation have been performed[47, 48]. The development of an analogous yeast system could be helpful to dissect the intracellular pathways of Aβ deposition.

Expression of Aβ, extended polyQ and α-Syn in yeast results in their inclusion into aggregates whose formation is modulated by their primary sequence. In fact, the observed single, large juxtanuclear deposit of Aβ42 is an aggregate of aggregates whose cellular position and kinetics of formation exhibit striking similarities with aggresomes in mammalian cells[49]. Aggresomes are thought to be part of a general cellular response to minimize the toxicity of aggregated polypeptides. The recurrent formation of aggregates by conformational disease-linked proteins in yeast might reflect an evolutionary conserved mechanism to modulate the effects of intracellular protein aggregation.

The approach discussed here aims at the easy and reliable evaluation of the effects of intrinsic and extrinsic factors on protein aggregation. And it is based on the correspondence between the intracellular activity and solubility of recombinant h-DHFR and cell growth in the presence of lethal concentrations of MTX. Furthermore, the use of fMTX (a MTX labeled with a fluorescent compound) enables to monitor simultaneously the cell viability and the localization of the aggregates inside the cell.

Overall, the method is able to anticipate the intracellular aggregation propensity of genetic variants of three unrelated polypeptides linked to important human disorders.

The system could become also a convenient platform for chemical screening of agents that interfere with protein aggregation in order to assist in the development of new therapeutic leader compounds targeting protein aggregation and toxicity. The use of *S. cerevisiae* is compatible with these applications due to the availability of drug-permeable

strains (i.e. *erg6*Δ). This way, the assay could detect the previously reported inhibitory aggregation activity of quercetin. The presence of this flavonoid in the medium restored the growth of cells expressing wild-type Aβ42 and reduced the size of intracellular Aβ aggregates, suggesting that it targets *in vivo* Aβ aggregation. In contrast, under the conditions of the assay, CR had not positive effect on growth rate. Consistently, it has been recently shown that, *in vitro,* CR inhibits Aβ42 oligomerization but does not affect its fibrillization[50]. Furthermore, it has been demonstrated that low concentrations of CR can promote fibril formation as shown for the Aβ11–28 fragment[51], the prion protein[52] or immunoglobulin light chains[53], questioning the therapeutic utility of this compound or its analogs to inhibit amyloidosis. The increased number of aggregation foci promoted by CR in our system might reflect a similar effect on Aβ aggregation inside yeast explaining the observed rather negative impact on cell survival.

The analysis of the effect of overexpression or deletion of chaperones demonstrated the applicability of the method in genetic screening. Overall, modification of chaperones levels had huge impact in the survival of cells and in the size and distribution of intracellular Aβ aggregates.

Among the entire chaperone set, overexpression of the Sis1 gene (the yeast homologue of human HDJ1) had the most dramatic effect on viability. This result is not surprising because the knockout of *sis1* gene is not viable, being thus essential for yeast. Its effect is clearly higher than this promoted by Ydj1 (the yeast Hsp40 homologue of human HDJ2). In agreement with our data, both in mammalian and in yeast cells, the Sis1/HDJ1 chaperones had a much stronger effect on modulating the aggregation of polyQ than Ydj1/HDJ2.

Overexpression of Hsp104 also increased cell viability. Hsp104 allows rescuing proteins from aggregated states regaining their function. It is a key protein in the chaperone network[54]. Puzzlingly, Sis1 or specially Hsp104 promoted an increase of fluorescence signal in the cytoplasm and in the aggregates. This coincides with their effect in the yeast model of HD, where it was shown that the protein in those aggregates was more loosely packed[18]. Taking into account the relationship between packing of Aβ-GFP aggregates and the activity of the embedded protein that was established in bacteria[55], these chaperones might promote loosely packed and probably more active Aβ-DHFR aggregates.

In contrast, Hsp82 promoted viability seems to be dependent on direct reduction of Aβ42 aggregation, resulting in smaller intracellular foci. Interestingly, also human

homolog interacts with amyloid precursors in AD[56] and its upregulation protects neurons from Aβ toxicity[57] .

Surprisingly, the knockout of Hsp35, Hsp26 and Hsp42 chaperones promoted cell growth, implying lower Aβ aggregation. Hsp35 is a member of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) family. And it is supposed to be a chaperone because of its heat inducibility and its high abundance in yeast. Intriguingly, in humans, polymorphic variation within GAPDH genes is associated with an elevated risk of developing AD and this effect depends specifically on the interaction with Aβ peptide[58]. Also, the Hsp35 *Caenorhabditis elegans* homolog is down regulated in transgenic animals expressing Aβ42[59].

Hsp26 and Hsp42 are small heat shock proteins (sHsp), which trap misfolded proteins into aggregates that are subsequently reactivated by the Hsp104/Hsp70/Hsp40 chaperone system. Nevertheless, larger substrate/sHsp relationships result in larger (and maybe tighter) complexes, which are poorly reactivated by other chaperones[60]. The knockout of sHsp might prevent the incorporation of overexpressed Aβ42 into sHsp aggregates reducing indirectly its deposition.

The loss of function of Ydj1 has been shown to reduce polyQ aggregation[61], an observation compatible with the increase in yeast survival promoted by its knockout in our model.

On the other hand, deletion of the members of the Hsp70 cytosolic family (Ssa1, Ssa2 and specially Ssa4) reduced cell viability and increased the number of fluorescent foci within the cells. Importantly, in a yeast model of polyQ aggregation, mutation of the SSA1, SSA2 genes also inhibited the expansion of small aggregate foci into a large inclusion body[61]. Finally, the huge impact on both viability and aggregation of SSa4 deletion suggest that it could play an important role in protein folding and/or deposition. Accordingly, under conformational stress the amounts of Ssa4 mRNA increase several fold[62].

Overall, and although the present study was not aimed to characterize in detail the effect of each specific chaperone on the intracellular aggregation of Aβ, the obtained data confirm that the aggregation of disease-related polypeptides in yeast share several hallmarks suggesting also new and differential targets to study Aβ aggregation.

The prevalence of the conformational diseases in our society presents a new challenge both for basic and applied research. In the last years the simple, yet powerful,

genetics of *S. cerevisiae* has been exploited for the study of intracellular amyloid protein aggregation. From the present study, it appears that Alzheimer's disease may be yet another disorder whose modelling in yeast could contribute to decipher conserved and/or differential mechanisms of amyloid aggregation and help in the identification of potential therapeutic targets. The ability of the method discussed here to link protein aggregation to cell survival is expected to allow a fast, visual and easily automated screening of target mutations, genes or compounds that modulate protein aggregation of disease-related polypeptides inside eukaryotic cells. Besides, the method should have wide applicability in protein production and design as well as in folding studies.

4.6 References

1. Thomas, P.J., Qu, B.H. & Pedersen, P.L. Defective protein folding as a basis of human disease. *Trends Biochem Sci* **20**, 456-459 (1995).
2. Dobson, C.M. Protein-misfolding diseases: Getting out of shape. *Nature* **418**, 729 - 730 (2002).
3. Rochet, J.C. & Lansbury, P.T. Amyloid fibrillogenesis: themes and variations. *Curr Opin Struct Biol* **10**, 60 - 68 (2000).
4. Ventura, S. & Villaverde, A. Protein quality in bacterial inclusion bodies. *Trends Biotechnol* **24**, 179-185 (2006).
5. Maxwell, K.L., Mittermaier, A.K., Forman-Kay, J.D. & Davidson, A.R. A simple in vivo assay for increased protein solubility. *Protein Sci* **8**, 1908-1911 (1999).
6. Waldo, G.S., Standish, B.M., Berendzen, J. & Terwilliger, T.C. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* **17**, 691-695 (1999).
7. Wigley, W.C., Stidham, R.D., Smith, N.M., Hunt, J.F. & Thomas, P.J. Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. *Nat Biotechnol* **19**, 131-136 (2001).
8. Kim, W. & Hecht, M.H. Sequence determinants of enhanced amyloidogenicity of Alzheimer A{beta}42 peptide relative to A{beta}40. *J Biol Chem* **280**, 35069-35076 (2005).
9. Philipps, B., Hennecke, J. & Glockshuber, R. FRET-based in vivo screening for protein folding and increased protein stability. *J Mol Biol* **327**, 239-249 (2003).
10. Lesley, S.A., Graziano, J., Cho, C.Y., Knuth, M.W. & Klock, H.E. Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng* **15**, 153-160 (2002).
11. Schultz, T., Martinez, L. & de Marco, A. The evaluation of the factors that cause aggregation during recombinant expression in E. coli is simplified by the employment of an aggregation-sensitive reporter. *Microb Cell Fact* **5**, 28 (2006).
12. Fisher, A.C., Kim, W. & DeLisa, M.P. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein Sci* **15**, 449-458 (2006).
13. Bagriantsev, S. & Liebman, S. Modulation of Abeta42 low-n oligomerization using a novel yeast reporter system. *BMC Biol* **4**, 32 (2006).
14. Chen, M.J. et al. The functional human dihydrofolate reductase gene. *J Biol Chem* **259**, 3933-3943 (1984).
15. Michnick, S.W., Remy, I., Campbell-Valois, F.X., Vallee-Belisle, A. & Pelletier, J.N. Detection of protein-protein interactions by protein fragment complementation strategies. *Methods Enzymol* **328**, 208-230 (2000).
16. Lefurgy, S. & Cornish, V. Finding Cinderella after the ball: a three-hybrid approach to drug target identification. *Chem Biol* **11**, 151-153 (2004).
17. Willingham, S., Outeiro, T.F., DeVit, M.J., Lindquist, S.L. & Muchowski, P.J. Yeast genes that enhance the toxicity of a mutant huntingtin fragment or alpha-synuclein. *Science* **302**, 1769-1772 (2003).
18. Krobitsch, S. & Lindquist, S. Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. *Proc Natl Acad Sci U S A* **97**, 1589-1594 (2000).

19. Brophy, V.H. et al. Identification of Cryptosporidium parvum dihydrofolate reductase inhibitors by complementation in Saccharomyces cerevisiae. *Antimicrob Agents Chemother* **44**, 1019-1028 (2000).
20. Barnham, K.J., Cappai, R., Beyreuther, K., Masters, C.L. & Hill, A.F. Delineating common molecular mechanisms in Alzheimer's and prion diseases. *Trends Biochem Sci* **31**, 465-472 (2006).
21. de Groot, N.S., Aviles, F.X., Vendrell, J. & Ventura, S. Mutagenesis of the central hydrophobic cluster in Abeta42 Alzheimer's peptide. Side-chain properties correlate with aggregation propensities. *Febs J* **273**, 658-668 (2006).
22. de Groot, N.S. & Ventura, S. Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. *J Biotechnol* **125**, 110-113 (2006).
23. Miyajima, A., Miyajima, I., Arai, K. & Arai, N. Expression of plasmid R388-encoded type II dihydrofolate reductase as a dominant selective marker in Saccharomyces cerevisiae. *Mol Cell Biol* **4**, 407-414 (1984).
24. Kiefhaber, T., Rudolph, R., Kohler, H.H. & Buchner, J. Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology (N Y)* **9**, 825-829 (1991).
25. Remy, I. & Michnick, S.W. Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays. *Proc Natl Acad Sci U S A* **96**, 5394-5399 (1999).
26. Gusella, J.F., Persichetti, F. & MacDonald, M.E. The genetic defect causing Huntington's disease: repeated in other contexts? *Mol Med* **3**, 238-246 (1997).
27. Martindale, D. et al. Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nat Genet* **18**, 150-154 (1998).
28. Lucking, C.B. & Brice, A. Alpha-synuclein and Parkinson's disease. *Cell Mol Life Sci* **57**, 1894-1908 (2000).
29. Muchowski, P.J. Protein misfolding, amyloid formation, and neurodegeneration: a critical role for molecular chaperones? *Neuron* **35**, 9-12 (2002).
30. Conway, K.A. et al. Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proc Natl Acad Sci U S A* **97**, 571-576 (2000).
31. Kruger, R. et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat Genet* **18**, 106-108 (1998).
32. Polymeropoulos, M.H. et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047 (1997).
33. Zhang, X. et al. A potent small molecule inhibits polyglutamine aggregation in Huntington's disease neurons and suppresses neurodegeneration in vivo. *Proc Natl Acad Sci U S A* **102**, 892-897 (2005).
34. Emter, R., Heese-Peck, A. & Kralli, A. ERG6 and PDR5 regulate small lipophilic drug accumulation in yeast cells via distinct mechanisms. *FEBS Lett* **521**, 57-61 (2002).
35. Matsuzaki, K. et al. Inhibitors of amyloid beta-protein aggregation mediated by GM1-containing raft-like membranes. *Biochim Biophys Acta* (2006).
36. Kim, H. et al. Effects of naturally occurring compounds on fibril formation and oxidative stress of beta-amyloid. *J Agric Food Chem* **53**, 8537-8541 (2005).
37. Griffioen, G. et al. A yeast-based model of alpha-synucleinopathy identifies compounds with therapeutic potential. *Biochim Biophys Acta* **1762**, 312-318 (2006).

38. Roterman, I. et al. Why Congo red binding is specific for amyloid proteins - model studies and a computer analysis approach. *Med Sci Monit* **7**, 771-784 (2001).
39. Flower, T.R., Chesnokova, L.S., Froelich, C.A., Dixon, C. & Witt, S.N. Heat shock prevents alpha-synuclein-induced apoptosis in a yeast model of Parkinson's disease. *J Mol Biol* **351**, 1081-1100 (2005).
40. Botstein, D., Chervitz, S.A. & Cherry, J.M. Yeast as a model organism. *Science* **277**, 1259-1260 (1997).
41. Giorgini, F., Guidetti, P., Nguyen, Q., Bennett, S.C. & Muchowski, P.J. A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease. *Nat Genet* **37**, 526-531 (2005).
42. Outeiro, T.F. & Giorgini, F. Yeast as a drug discovery platform in Huntington's and Parkinson's diseases. *Biotechnol J* **1**, 258-269 (2006).
43. Glabe, C. Intracellular mechanisms of amyloid accumulation and pathogenesis in Alzheimer's disease. *J Mol Neurosci* **17**, 137-145 (2001).
44. Zhang, Y., McLaughlin, R., Goodyer, C. & LeBlanc, A. Selective cytotoxicity of intracellular amyloid beta peptide1-42 through p53 and Bax in cultured primary human neurons. *J Cell Biol* **156**, 519-529 (2002).
45. Casas, C. et al. Massive CA1/2 neuronal loss with intraneuronal and N-terminal truncated Abeta42 accumulation in a novel Alzheimer transgenic model. *Am J Pathol* **165**, 1289-1300 (2004).
46. Mori, C. et al. Intraneuronal Abeta42 accumulation in Down syndrome brain. *Amyloid* **9**, 88-102 (2002).
47. Crowther, D.C. et al. Intraneuronal Abeta, non-amyloid aggregates and neurodegeneration in a Drosophila model of Alzheimer's disease. *Neuroscience* **132**, 123-135 (2005).
48. Cohen, E., Bieschke, J., Perciavalle, R.M., Kelly, J.W. & Dillin, A. Opposing activities protect against age-onset proteotoxicity. *Science* **313**, 1604-1610 (2006).
49. Mukai, H. et al. Formation of morphologically similar globular aggregates from diverse aggregation-prone proteins in mammalian cells. *Proc Natl Acad Sci U S A* **102**, 10887-10892 (2005).
50. Frid, P., Anisimov, S.V. & Popovic, N. Congo red and protein aggregation in neurodegenerative diseases. *Brain Res Brain Res Rev* (2006).
51. Fraser, P.E., Nguyen, J.T., Chin, D.T. & Kirschner, D.A. Effects of sulfate ions on Alzheimer beta/A4 peptide assemblies: implications for amyloid fibril-proteoglycan interactions. *J Neurochem* **59**, 1531-1540 (1992).
52. Rudyk, H. et al. Screening Congo Red and its analogues for their ability to prevent the formation of PrP-res in scrapie-infected cells. *J Gen Virol* **81**, 1155-1164 (2000).
53. Kim, Y.S., Randolph, T.W., Manning, M.C., Stevens, F.J. & Carpenter, J.F. Congo red populates partially unfolded states of an amyloidogenic protein to enhance aggregation and amyloid fibril formation. *J Biol Chem* **278**, 10842-10850 (2003).
54. Weibezahn, J., Bukau, B. & Mogk, A. Unscrambling an egg: protein disaggregation by AAA+ proteins. *Microb Cell Fact* **3**, 1 (2004).
55. de Groot, N.S. & Ventura, S. Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Lett* (2006).
56. Cottrell, B.A. et al. A pilot proteomic study of amyloid precursor interactors in Alzheimer's disease. *Ann Neurol* **58**, 277-289 (2005).
57. Ansar, S. et al. A non-toxic Hsp90 inhibitor protects neurons from Abeta-induced toxicity. *Bioorg Med Chem Lett* **17**, 1984-1990 (2007).

58. Cumming, R.C. & Schubert, D. Amyloid-beta induces disulfide bonding and aggregation of GAPDH in Alzheimer's disease. *Faseb J* **19**, 2060-2062 (2005).
59. Link, C.D. et al. Gene expression analysis in a transgenic Caenorhabditis elegans Alzheimer's disease model. *Neurobiol Aging* **24**, 397-413 (2003).
60. Cashikar, A.G., Duennwald, M. & Lindquist, S.L. A chaperone pathway in protein disaggregation. Hsp26 alters the nature of protein aggregates to facilitate reactivation by Hsp104. *J Biol Chem* **280**, 23869-23875 (2005).
61. Meriin, A.B. et al. Huntington toxicity in yeast model depends on polyglutamine aggregation mediated by a prion-like protein Rnq1. *J Cell Biol* **157**, 997-1004 (2002).
62. Wegele, H., Haslbeck, M., Reinstein, J. & Buchner, J. Sti1 is a novel activator of the Ssa proteins. *J Biol Chem* **278**, 25970-25976 (2003).

DISCUSSION

Discussion

The present work was aimed to develop novel *in vivo* approaches to study and characterize how and where proteins interact with other proteins to control most of the cellular processes and how proteins interact anomaly with themselves to render toxic intracellular aggregates under special circumstances. In the first part of this thesis, we have taken advantage of the properties displayed by fluorescent proteins to study the applications of BIFC as a method to detect weak protein interactions as well as to monitor the specific interference of protein binding. In the second part, some properties of protein aggregation in prokaryotic environment have been deciphered by fusing aggregation prone peptides to fluorescent proteins. Besides, the use of another protein reporter (the enzyme DHFR) has allowed to link cell survival to protein aggregation propensity in more complex, eukaryotic backgrounds.

Part 1

Historically, most of the methods available to study protein interactions were mainly useful to detect strong association between binding partners. Nevertheless, in the cell, the majority of interactions are weak. Therefore, the progress towards assays that enable the *in vivo* study of weak protein contacts is really important in order to have a self-consistent overall view of the cell interactome.

BIFC has been demonstrated to be a useful method to detect and characterize protein interactions as well as to discover specific antagonists of protein-protein binding[1, 2]. Still, it has to take into account that the perfect, universal, method does not exist. Nowadays, there are techniques that provide us with similar analytical answers (e.g. TAP approach, FRET, Y2H). Thus, the choice of one approach will strongly depend on the target interaction and on the specific objectives of the study. In order to find specific and physiologically functional partners of one specific protein, working *in vivo* is very convenient because the cellular environment exerts an enormous influence on the establishment of protein bindings. However, once an interaction has been identified, the molecular and biophysical properties of the complex need to be characterized in detail. And these key aspects can be only addressed using *in vitro* methods.

Many human diseases are related to aberrant protein-protein interactions: the loss of a crucial interaction or the establishment of an abnormal protein complex between

endogenous proteins or with pathogen proteins. For this reason, protein interactions have been recognized as challenging but attractive targets for chemical drugs. In particular, investigations to date suggest that the inhibition of protein interactions by small drug-like chemical compounds could lead to treatments for human diseases[3-6].

The study of the human interactome as well as the inhibition of certain protein associations could open new avenues for drug development. Novelty in drug research is really important: a deep study of the relationship between drugs and their targets has shown that many therapeutic compounds are only "follow-on drugs"[7]. During a long time, novel drugs were not acting against new targets; they were only improvements of the old versions sharing the same objective. In addition, it has been demonstrated that many of the actual drugs are palliative instead of having an etiologic effect. In other words, the common mechanism of their action is to counteract disease symptoms and not to attack directly the proteins causing the illness. Nevertheless, because of the development of methods that enable the high-throughput analysis of protein interactions and their inhibitors, there is more concernment of the multiplicity in protein binders and the high connectivity between proteins (as it is reflected in the interactome networks). Therefore, at the present time drug discovery is moving from the classic protein-centric view to a more holistic pathway-centric analysis[8]. In our opinion, BIFC-related approaches might contribute in this search.

Part 2

Protein aggregation has become one of the most important areas in protein chemistry research, mainly because it has a huge impact in many biotechnological processes as well as in many human degenerative disorders (like Alzheimer or Parkinson diseases)[9]. The aggregated state is now recognized as one of the energy minimums in folding landscape of any polypeptide and therefore, as a stable protein conformation[10]. In the last few years, many studies have addressed this issue *in vitro* by protein engineering using both disease related and non-related polypeptide models. Accordingly, our knowledge on the mechanistics of the aggregation process has increased significantly. Unfortunately, fewer studies have addressed the influence of polypeptide properties on their aggregation within the cell. The *in vivo* aggregation of polypeptides does not necessarily have to correlate with their *in vitro* properties. The protein quality machinery modulates the accumulation of aggregation prone polypeptidic chains by facilitating their

folding, masking hydrophobic regions and targeting improperly folded proteins towards degradation pathways[11].

Our view on how protein aggregation proceeds within the cell will clearly be benefited from the development of simple, yet physiologically relevant models to study protein self-assembly in living organisms. For this reason, the demonstration that prokaryotic and eukaryotic aggregation pathways have common features allow to speculate whether the genetically and biochemically well characterized bacterial organisms could provide important insights on the *in vivo* determinants of protein aggregation

In addition, the availability of assays to detect, localize and characterize *in vivo* protein aggregation open an avenue to dissect the main factors that influence protein aggregation. Besides, they could be useful to test or screen small compounds of therapeutic interest.

References

1. Morell, M. et al. Monitoring the interference of protein-protein interactions in vivo by bimolecular fluorescence complementation: the DnaK case. *Proteomics* **8**, 3433-3442 (2008).
2. Morell, M., Espargaro, A., Aviles, F.X. & Ventura, S. Detection of transient protein-protein interactions by bimolecular fluorescence complementation: the Abl-SH3 case. *Proteomics* **7**, 1023-1036 (2007).
3. Archakov, A.I. et al. Protein-protein interactions as a target for drugs in proteomics. *Proteomics* **3**, 380-391 (2003).
4. Arkin, M.R. & Wells, J.A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* **3**, 301-317 (2004).
5. Pagliaro, L. et al. Emerging classes of protein-protein interaction inhibitors and new tools for their development. *Curr Opin Chem Biol* **8**, 442-449 (2004).
6. Toogood, P.L. Inhibition of protein-protein association by small molecules: approaches and progress. *J Med Chem* **45**, 1543-1558 (2002).
7. Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabasi, A.L. & Vidal, M. Drug-target network. *Nat Biotechnol* **25**, 1119-1126 (2007).
8. Fishman, M.C. & Porter, J.A. Pharmaceuticals: a new grammar for drug discovery. *Nature* **437**, 491-493 (2005).
9. Selkoe, D.J. Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. *Nat Cell Biol* **6**, 1054-1061 (2004).
10. Jahn, T.R. & Radford, S.E. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys* **469**, 100-117 (2008).
11. Muchowski, P.J. Protein misfolding, amyloid formation, and neurodegeneration: a critical role for molecular chaperones? *Neuron* **35**, 9-12 (2002).

CONCLUDING REMARKS

Concluding remarks

Part 1

Chapter 1

- BIFC approach can be used to detect weak protein interactions (in the micromolar range). Specifically, the binding between a SH3 domain and peptides can be visualized in a bacterial background.

- The detected fluorescence is proportional to the interaction strength. One mutation that weakens the interaction causes a decrease in the emitted fluorescence.

- BIFC can be applied to map the interaction surface of one specific protein binding.

- BIFC can be coupled to flow cytometry turning to a good method for the screening of mutations that affect an interaction. The cells displaying lowest or highest fluorescence signal could be isolated afterwards using a fluorescence-activated cell sorting (FACS) even if they are present in a low percentage.

Chapter 2

- BIFC approach can be used to monitor the inhibition of a specific protein interaction *in vivo*. Specifically, it has been applied to detect in *E.coli* the inhibition of the binding between a chaperone and a short peptidic ligand by addition of perturbing antibacterial peptides (pyrrhocoricin).

- The method could be applied to screen inhibitors with different potency because the emitted fluorescence is directly linked with the strength of the modulation effect: higher the inhibition, lower detected fluorescence.

- The reassembly of the fragments of the fluorescent protein is irreversible. Therefore, the inhibitor has to be added before the protein expression has started.

- BIFC coupled to flow cytometry can also be applied to the identification of interaction inhibitors with optimal *in vivo* biological activity (penetrability, stability and affinity).

Part 2

Chapter 3

- It has been demonstrated that the formation of protein aggregates in a prokaryotic cytoplasm exhibits an important *in vivo* specificity even among extremely aggregating polypeptides expressed at very high levels.

- Because all organisms face challenges of protein misfolding and aggregation, it is not surprising the existence of an evolutionary conserved strategy to avoid the harmful effects of undesired protein aggregation by sequestering sticky folding intermediates into conformationally related stable aggregated structures through selective interactions.

- Inclusion bodies that are comprised of different aggregated proteins have a dissimilar stability in front chemical denaturation than the ones with homogeneous composition.

- Digestion of inclusion bodies by a protease (proteinase K) enables the visualization of fibrillar structures inside the aggregates using microscopy techniques (TEM and AFM).

- The detection of SDS resistant oligomers suggests common features in the fibrillar formation pathway between the prokaryotic and eukaryotic environments.

- The formation of fibrillar structures able to seed amyloid fibril formation points to common conformational features between amyloid fibrils and IBs.

Chapter 4

- The enzyme DHFR can be used as a protein reporter (together with its inhibition by methotrexate) to detect *in vivo* protein aggregation.

- The approach enables to detect mutations of the target protein that change its aggregation propensity.

- The use of methotrexate conjugated with a fluorescent dye confirmed the equivalent behavior of the target protein fused to DHFR or GFP.

- The system is sensitive to changes in the growth conditions or to the presence of certain chemical compounds that can reverse the aggregation propensity of the target protein.

- The method is also receptive to changes in the yeast genetic background: the overexpression or deletion of chaperones (proteins belonging to the cellular folding machinery) could be detected.

SUMMARY

Summary in Catalan

Part 1

Chapter 1

Les interaccions entre proteïnes són essencials en molts processos biològics. Molts mètodes proteòmics han reeixit en la identificació d'interaccions fortes però l'estudi d'interaccions dèbils és encara tot un repte. L'objectiu d'aquest estudi va ser analitzar la capacitat de la tècnica anomenada *"Bimolecular fluorescent protein complementation" (BIFC)* per a detectar i discriminar interaccions intracel·lulars proteiques de caràcter dèbil *in vivo*. Amb aquest objectiu, es va prendre com a model la interacció entre el domini SH3 de la proteïna quinasa c-Abl amb lligands naturals o dissenyats. L'assemblatge funcional de la proteïna fluorescent a partir dels seus fragments requereix la prèvia interacció del domini SH3 i dels seus lligands. No obstant, a continuació, el complex queda atrapat degut a l'agrupació de la proteïna fluorescent. D'aquesta manera, les dèbils interaccions del domini SH3 es transformen en interaccions estables i de fàcil detecció. Si s'acobla a la citometria de flux, el mètode BIFC pot ser implementat com una eina d'anàlisis proteòmica d'interaccions proteiques dèbils. La combinació d'ambdós mètodes proporciona una tècnica ràpida i altament sensible per a la validació d'interaccions. D'altra banda, també es demostra que l'emissió de fluorescència depèn de la força de la interacció. D'aquesta manera, el mètode pot ser usat per seleccionar els millors lligands entre un grup de pèptids rics en prolines. A més a més, el mètode BIFC pot ser aplicat per a obtenir informació sobre la superfície d'interacció entre el domini SH3 i una altra proteïna (per exemple, la BRCA1). Així es poden deduir quines regions intervenen en la unió.

Chapter 2

Molts processos cel·lulars depenen de l'establiment d'interaccions entre proteïnes. Moltes malalties estan relacionades amb l'establiment de noves unions proteiques o en disfuncions en interaccions essencials. Per tant, la identificació de molècules capaces de modular aquests contactes proteics és d'especial interès per al descobriment de noves drogues

per pal·liar aquestes malalties. En aquest capítol, es descriu l'ús de BIFC com a mètode per a detectar compostos que bloquegen la interacció de proteïnes diana *in vivo*. Es demostra que la inhibició de una interacció és concomitant a uns disminució en la fluorescència detectada. També es prova que l'acoblament entre el BIFC i la citometria de flux permet detectar la presència d'inhibidors a través de canvis en la senyal de fluorescència. L'aplicació *in vivo* es confirma en la detecció de la activitat inhibitòria d'un pèptid bactericida (anomenat pyrrhocoricin) sobre la interacció entre una xaperona DnaK i un substrat peptídic en *E.coli*. Amb els resultats obtinguts es confirma que l'activitat bactericida està relacionat directament amb la inhibició de l'activitat de la xaperona DnaK.

Part 2

Chapter 3

Tradicionalment, s'ha considerat que els agregats intracel·lulars estan formats a partir d'associacions inespecífiques de proteïnes mal plegades. No obstant, varis estudis demostren que el procés d'agregació proteica *in vitro* sovint involucra interaccions específiques. En canvi, el coneixement que es posseeix de l'especificitat de l'agregació proteica *in vivo* és molt poca. En aquest capítol s'ha investigat el grau de co-agregació *in vivo* entre dues proteïnes amb tendència a agregar (el pèptid amiloide Aβ42 i la proteïna VP1 que constitueix la càpside viral) co-expressades en *E.coli*. D'altra banda, l'estructura dels agregats intracel·lulars ha estat investigada per desxifrar si les fibres amiloides i els cossos d'inclusió comparteixen característiques estructurals. Les dades obtingudes indiquen que l'agregació proteica *in vivo* presenta una remarcable especificitat que depèn de l'establiment d'interaccions selectives i resulta en la formació d'estructures oligomèriques i fibrilars que presenten propietats amiloides. Aquestes característiques permeten als agregats intracel·lulars de Aβ42 formats en *E.coli* d'actuar com a acceleradors de la formació de fibres amiloides de Aβ42 (en un procés anomenat *seeding*). Globalment, els resultats suggereixen que l'agregació proteica en organismes eucariotes i procariotes es basa en principis idèntics. Aquest fet pot tenir importants implicacions en l'ús de proteïnes recombinants en aplicacions biotecnològiques i biomèdiques.

Chapter 4

L'agregació proteica es troba latent en moltes malalties que afecten els humans i també constitueix un coll d'ampolla en la producció recombinant de proteïnes. En aquest capítol, es descriu un nou mètode per a estudiar l'agregació proteica en llevat. Es basa en la fusió de la proteïna d'interès a un enzim, concretament la di-hidrofolat reductasa (DHFR). L'activitat d'aquest enzim és essencial per al llevat. D'aquesta manera, sota la presència d'un inhibidor d'aquest enzim (anomenat metotrexat), la tendència a agregar de la proteïna diana queda directament lligada a la supervivència i creixement del llevat. En altres paraules, si la proteïna agrega, la DHFR no serà funcional i el llevat serà susceptible a la presència del metotrexat en el medi. Al contrari, si la proteïna no agrega, la DHFR es trobarà soluble en el medi intracel·lular conferint resistència al llevat. Degut al fet que l'agregació està relacionada al creixement, no es necessita cap assaig funcional per a detectar-la. A més a més, l'ús d'un derivat fluorescent del metotrexat permet confirmar l'estat d'agregació o solubilitat de la proteïna diana.. Per demostrar una possible aplicació del mètode en el cribratge de gens o compostos que modulen l'agregació de proteïnes en cèl·lules vives, el pèptid amiloide Aβ42, les expansions de poliglutamines, la proteïna α-sinucleïna així com les seves variants no amiloides van ser usades com a models. D'altra banda, les propietats anti-agregacionals de diferents compostos químics que s'uneixen al pèptid Aβ42 també varen ser avaluades emprant el mètode. Finalment, l'efecte de la deleció o la sobre expressió de xaperones en l'agregació de Aβ42 va ser estudiada.

The research presented in this thesis is the basis of the following scientific papers:

1. Morell, M., de Groot N., Vendrell J., Aviles, F.X. & Ventura, S. Linking aggregation and yeast survival. *Submitted.*

2. Morell, M. et al. Inclusion bodies: Specificity in their aggregation process and amyloid-like structure. *Biochim Biophys Acta* (2008).

3. Morell, M. et al. Monitoring the interference of protein-protein interactions in vivo by bimolecular fluorescence complementation: the DnaK case. *Proteomics* **8**, 3433-3442 (2008).

4. Morell, M., Espargaro, A., Aviles, F.X. & Ventura, S. Study and selection of in vivo protein interactions by coupling bimolecular fluorescence complementation and flow cytometry. *Nat Protoc* **3**, 22-33 (2008).

5. Morell, M., Espargaro, A., Aviles, F.X. & Ventura, S. Detection of transient protein-protein interactions by bimolecular fluorescence complementation: the Abl-SH3 case. *Proteomics* **7**, 1023-1036 (2007).