UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dottorato di Ricerca in Biologia XXIV ciclo

EXTRINSIC FACTORS AFFECTING AMYLOID AGGREGATION

Tutor: Dr. Maria Elena Regonesi
Cotutor: Prof. Paolo Tortora

Francesco Antonio Aprile
Matr: 055035

Table of contents

Abstract

Protein conformational diseases (PCD) include a range of degenerative disorders in which specific peptides or proteins misfold and aberrantly self-assemble, eventually forming amyloid-like fibrils, which constitute the hallmark of many neurodegenerative diseases. Plenty of works demonstrated that amyloid aggregation is strongly influenced by several extrinsic factors, such as high concentrations of macromolecules or the presence of proteases and chaperons.

Molecular chaperones have been recognised as key players in the avoidance of amyloid fibril formation and, in particular, recent evidences demonstrate that low levels of chaperone heat-shock protein 70 kDa (Hsp70) are strictly related to the formation of intra-neuronal inclusions associated with Parkinson's (PD) and polyglutamine (polyQ) diseases.

Human Hsp70 is composed of two major functional domains connected by a conserved interdomain linker: the 44-kDa N-terminal nucleotide-binding domain (NBD), with ATPase activity, followed by the 30-kDa substrate-binding domain (SBD) that contains a C-terminal lid subdomain (LS). Using a battery of Hsp70 variants, including full-length Hsp70 and SBD truncated variants, we have been able to discover an interaction between the LS of SBD and the interdomain linker, which we propose could play an important role in the allosteric communication between NBD and SBD. Therefore, we analyzed the anti-amyloidogenic activity of Hsp70, using two model proteins: alpha-synuclein (AS), whose deposition in the brain is associated with PD and the polyQ protein ataxin-3 (AT3), the causative agent of the Machado-Joseph disease (MCD). We demonstrated Hsp70 is able to interact and stabilize pre-fibrillar species formed during amyloid aggregation and that the binding mechanism of these species is different from that of the monomeric protein.

Plenty of evidence supports the idea that protein aggregation observed in *in vitro* experimental conditions is different from that naturally occurring in *in vivo* systems. This is also related with the fact that the high concentration of macromolecules present in the intra- and extra-cellular compartments, a condition known as molecular crowding (MC), strongly affects protein folding and aggregation. Here, we successfully employed *Escherichia coli* as *in vivo* model for studying the aggregation mechanism of the polyQ protein ataxin-3

(AT3) in the presence of MC. In particular we investigated the relationship between the aggregation pathway and cytotoxicity and we were able to characterize the kinetic of formation of aggregated toxic and non-toxic species of AT3.

Our future efforts will be aimed to investigate *in vivo* Hsp70 action by analyzing structural and physiological features of AT3 aggregated species formed in the intracellular environment of *E. coli* that co-express Hsp70 under different conditions.

1. Hsp70 chaperone: molecular mechanism of action

1.1 Protein folding and misfolding

One of the peculiarities of living systems is the capability of even the most intricate of its component molecular structures to self-assembly with precision and fidelity. The folding of proteins into their compact and native three-dimensional structures is the most fundamental and universal example of biological self-assembly mechanism (1).

The wide variety of highly specific structures that result from protein folding has enabled living system to develop astonishing diversity and selectivity in their underlying chemical process. Moreover, protein folding is coupled to many other biological processes, including the trafficking of molecules to specific cellular locations and the regulation of cellular growth and differentiation (2). It is therefore not surprising that the failure of proteins to fold correctly or to remain correctly folded is the origin of a wide variety of pathological conditions.

Although Anfinsen demonstrated in his studies on bovine pancreatic ribonuclease already in 1961 that all the necessary information for native structure is contained in the amino acid sequence, the mechanism by which a polypeptide chain folds to a specific three-dimensional protein structure has until recently been a mystery. Native states of proteins almost always correspond to the structures that are the most thermodynamically stable under physiological conditions. Nevertheless, the total number of possible conformations for any polypeptide chain is so large that a systematic search for native structure is not possible (3). It is now clear that protein folding process does not involve a sequence of mandatory steps between specific partially folded states, but rather a stochastic search for the many conformations accessible to a polypeptide chain. Since, on average, native-like interactions between residues are more stable than non-native ones, they are more persistent and, as such contacts form preferentially, the number of available conformations is reduced, driving the polypeptide chain towards the native conformation by a process of "trial and error" (Fig. 1.1) (4).

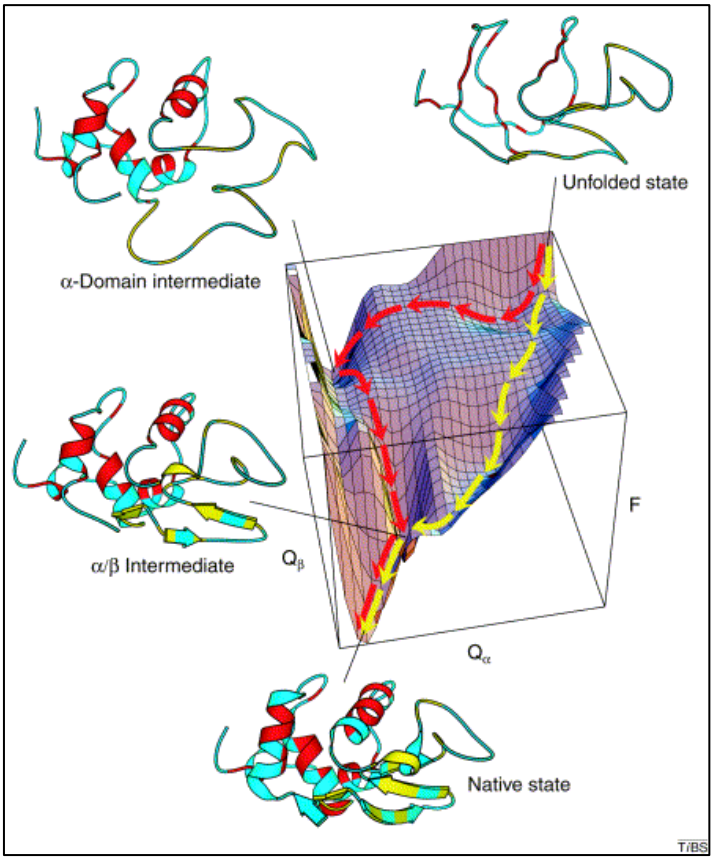

Fig. 1.1 Example of free-energy (F) landscape representing the folding of the hen lysozyme, a protein of 129 amino acids constituted by an α- and a β- domain. Q_α and Q_β indicate the number of native contacts in each domain. The yellow trajectory represents a 'fast track' in which domains fold concurrently. The red trajectory represents a 'slow track' in which the chain becomes trapped in a long-lived intermediate with persistent structure only in the α-domain; further folding from this intermediate involves either a transition over a higher barrier, or partial unfolding to enable the remainder of the folding process to occur along the fast track (4).

A key question is how the native structure is able to emerge from such a huge amount of possible conformations and so which are the determinants and the driving forces of the folding process. The results of many studies suggest that the fundamental mechanism of protein folding involves the interaction of a relatively small number of residues to form a folding nucleus, characterized by a rudimentary native-like architecture, about which the remainder of the structure rapidly condenses (5). Although it is not yet clear exactly how the sequence encodes such characteristics, the essential elements of the

fold are likely to be determined primarily by the pattern of hydrophobic and polar residues that favours preferential interactions of specific residues as the structure becomes increasingly compact. Once the correct topology has been achieved, the native structure will then almost invariably be generated during the final stages of folding (6).

Small single domain proteins (< 100 amino acids in length), in general, fold to the native state on a sub-second timescale and have been the focus of many experimental and theoretical studies of folding. The folding landscape of these proteins is usually relatively smooth, resulting in only two species being stably populated during the folding reaction. For proteins with more than about 100 residues, experiments generally reveal that one (or more) intermediate is significantly populated during the folding process. There has, however, been considerable discussion about the significance of such species: whether they assist the protein to find its correct structure or whether they are traps that inhibit the folding process (7-9). Regardless of the outcome of this debate, the structural properties of intermediates provide important hints about the folding of these larger proteins. In particular, they suggest that these proteins generally fold in modules; in other words, folding can take place largely independently in different segments or domains of the protein (4). In such cases, interactions involving key residues are likely to establish the native-like fold within local regions or domains and also to ensure that the latter then interact appropriately to form the correct overall structure (9,10). The fully native structure is only acquired when all the native-like interactions have been formed both within and between the domains; this happens in a final cooperative folding step when all the side chains become locked in their unique close-packed arrangement and water is excluded from the protein core (11). This modular mechanism is appealing because it suggests that highly complex structures might be assembled in manageable pieces. Moreover, such a principle can readily be extended to describe the assembly of other macromolecules, particularly nucleic acids, and even large ‘molecular machines’ such as the ribosome.

Partially folded or misfolded states often tend to aggregate, particularly when they represent major kinetic traps in the folding pathway. This is due to the fact that these forms typically expose hydrophobic amino acid residues and regions of unstructured polypeptide backbone, features that are mostly buried in the native state. Like intra-molecular folding, aggregation—the association of

two or more non-native protein molecules is largely driven by hydrophobic forces and primarily results in the formation of amorphous structures. Alternatively, aggregation can lead to the formation of highly ordered, fibrillar aggregates called amyloid, in which β-strands run perpendicular to the long fibril axis (cross-β structure). Although apparently restricted to a subset of proteins under physiological conditions, these thermodynamically highly stable structures are accessible to many proteins under denaturing conditions, largely independent of sequence, suggesting that their formation is an inherent property of the polypeptide chain (1). The formation of amyloid fibrils is usually toxic to cells and may give rise to some of the most debilitating neurodegenerative diseases.

In a cell, proteins are synthesized on ribosomes from the genetic information encoded in the cellular DNA. Folding *in vivo* is in some cases co-translational; that is, it is initiated before the completion of protein synthesis, whereas the nascent chain is still attached to the ribosome. Other proteins, however, undergo most of their folding process in the cytoplasm after release from the ribosome, whereas yet others fold in specific compartments, such as mitochondria or the endoplasmic reticulum, after trafficking and translocation through membranes (12,13). Many details of the folding process depend on the particular environment in which folding takes place, although the fundamental principles of folding, discussed above, are undoubtedly universal. But since incompletely folded proteins must inevitably expose to the solvent at least some regions of structure that are buried in the native state, they are prone to inappropriate interactions with other molecules within the crowded environment of a cell.

1.2 *In vivo* folding and the "Chaperone paradigm"

Recently, new fluorescence-based techniques have been developed in order to characterize protein folding and aggregation *in vivo* in real time (14). These new experimental approaches have allowed to demonstrate that partially folded proteins show an high aggregation propensity in the highly crowded context of the cytoplasm (15).

In cell, actually, folding occurs in the presence of 300–400 mg/ml of macromolecule concentration. This fact substantially enhances the affinity between interacting protein molecules, including intermediates of the folding process, increasing the risk of harmful

interactions. Even translation can potentially increase the risk of misfolding and aggregation, as incomplete polypeptide chains cannot fold into stable native conformation. Moreover, the exit channel of the large ribosomal subunit, which is ~100 Å long but, at most, 20 Å wide, largely precludes folding beyond the formation of α-helical elements (16,17) and thus prevents the C-terminal 40–60 residues of the chain from participating in long-range interactions. As a consequence, productive folding can occur only after a complete protein or at least a domain (~50–300 amino acids) has emerged from the ribosome, consistent with the general rules of folding established *in vitro* and supported by recent simulations of nascent chain folding (18). Because translation is relatively slow (~15–75 s for a 300-amino-acid protein), nascent chains are exposed in partially folded, aggregation-sensitive states for prolonged periods of time. Moreover, non-native intra-chain contacts formed during translation could block folding upon completion of synthesis.

For all these reasons, evolution has developed several protein quality control systems, including molecular chaperones, in order to prevent co-translational protein misfolding and avoid aggregation. We define a molecular chaperone as any protein which interacts, stabilizes or helps a non-native protein acquire its native conformation but it is not present in the final functional structure (19). The molecular mechanism of action of chaperones in assisting protein folding consists in the specific interaction with hydrophobic exposed regions in non-native proteins in order to promote their folding through ATP-regulated cycles of binding and release. This interaction reduces the concentration of folding intermediates, consequently blocking misfolding and dangerous interactions, whereas transient release of bound hydrophobic regions is necessary for folding to proceed (19). It is important to consider that chaperones act not by contributing steric information to the folding process but rather by optimizing the efficiency of folding. Notably, a number of essential proteins have extremely low intrinsic folding efficiencies and essentially do not fold in the absence of chaperones. For example, actin and tubulin seem to have highly energetically frustrated folding pathways and can overcome kinetic folding barriers only through assistance by chaperones.

Apart from promoting protein *de novo* folding, chaperones are involved in other fundamental cellular functions: they refold stress-denatured proteins and disaggregate oligomeric assemblies; they participate in intra-cellular protein transport and they assist proteolytic

degradation pathways. Moreover, as mutations often disrupt a protein's ability to fold, it follows that the chaperone system is also important in buffering such deleterious mutations. This buffering function is thought to be crucial in the evolution of new protein functions and phenotypic traits.

Numerous classes of structurally unrelated chaperones have been described both in *Eukaryotes* and in *Prokaryotes* (20,21). Many of these are known as stress proteins or heat-shock proteins, as they are up-regulated by cells under conditions of conformational stress in which the concentration of aggregation-prone folding intermediates increases. Chaperones are usually classified according to their molecular weight (Hsp40, Hsp60, Hsp70, Hsp90, Hsp100 and the so-called small Hsp proteins). Although all these components have the capacity to prevent aggregation, only certain members of the Hsp100 family in bacteria and fungi can actively dissociate aggregates for subsequent protein refolding or degradation. The cellular chaperone machinery forms complex networks that are indispensable for protein quality control and maintenance of protein homeostasis.

1.3 Hsp70 chaperones: functions and molecular mechanism of action

Heat shock protein (Hsp) 70 family members constitute the central part of a ubiquitous chaperone machinery present in *Eukaryotes*, *Eubacteria* and many *Archaea*. They assist a wide range of folding processes, including the folding and assembly of newly synthesized proteins, refolding of misfolded and aggregated proteins, cell sorting and targeting and control of the activity of regulatory proteins. They are, by far, the most conserved proteins in evolution (22-24): the prokaryotic Hsp70 protein DnaK from *Escherichia coli* shares approximately 50% amino acid identity with eukaryotic Hsp70s (25), which has allowed the vast amount of results obtained from this protein to be extrapolated to the whole Hsp70 family.

Hsp70 chaperones consist of two functional major domains, connected by an hydrophobic linker with a well conserved leucine-rich motif: the N-terminal nucleotide binding domain (NBD) of 45 kDa and the C-terminal substrate binding domain (SBD) of 25 kDa which is further divided into a β-sandwich sub-domain (SBSD) of 15 kDa and a C-terminal α-helical lid sub-domain (LSD) (Fig. 1.2A), which plays a key role in regulating the kinetics of substrate binding (26,27).

A major feature for Hsp70 chaperones is the ability to interact with a large spectrum of protein substrates, characterized by very different size and conformation, ranging from extended nascent polypeptides to fully assembled protein oligomers and aggregates. This peculiarity of Hsp70 mechanism of action is possible thank to the low specificity of the interaction surface as well as to the high intrinsic dynamicity of the lid sub-domain that is able to assume different conformations, adapting the binding pocket to substrates of different size (28).

The role of Hsp70 in the folding process of non-native proteins can be divided into three related activities: promotion of folding to the native state (i.e. *de novo* folding), prevention of aggregation and refolding of aggregated proteins.

The mechanism by which Hsp70-chaperones contribute to the folding of non-native protein substrates is not yet completely understood. Hsp70-mediated protein folding occurs *in vitro* typically on the time scale of minutes or longer. Hsp70 chaperone binds substrates in an ATP-dependent manner until the ensemble of substrate molecules has reached the native conformation. There are at

least two alternative proposed mechanisms of action. In the first Hsp70 plays a rather passive role; through repetitive substrate binding and release cycles, it keeps the free concentration of the substrate sufficiently low to prevent aggregation, allowing free molecules to fold to the native state ('kinetic partitioning'). In the second mechanism, the binding and release cycles induce local unfolding in the substrate, e.g. the untangling of a misfolded β-sheet, which helps to overcome kinetic barriers for folding to the native state ('local unfolding') (26,29-31).

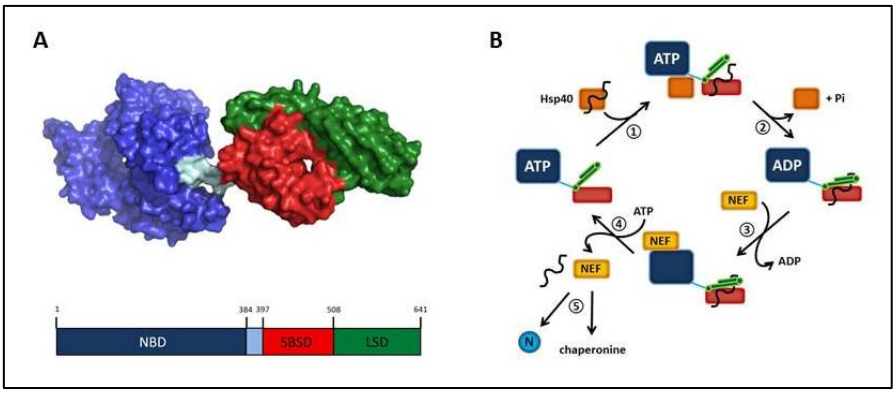

Fig. 1.2 Structures of *E. coli* DnaK (PBI 2KHO). NBD is represented in blue, the inter-domain linker in cyan, the β sub-domain in red and the helical lid in green. A schematic representation of the domains is reported below the structure (the residue numbers refer to human Hsp70 1A). (b) Hsp70 reaction cycle. (1) Hsp40-mediated delivery of substrate to ATP-bound Hsp70. (2) Hydrolysis of ATP to ADP, accelerated by Hsp40, results in closing of the helical lid and tight binding of substrate by Hsp70. Hsp40 dissociates from Hsp70. (3) Dissociation of ADP catalyzed by NEF. (4) Opening of the a-helical lid, induced by ATP binding, results in substrate release. (5) Released substrate either folds to native state (N) or it is transferred to downstream chaperones or rebinds to Hsp70 (19).

Most accepted theories about Hsp70 mechanism of action sustain that substrate binding occurs through the transient ATP-mediated association of Hsp70 with short hydrophobic tracts exposed in non-native protein substrates (Fig. 1.2B). This interaction is regulated by an allosteric mechanism of coupling between the two functional domains SBD and NBD. Actually, when ATP is bound to the NBD, the SBD acquires an *open* conformation that allows the accommodation of the substrates into SBD binding pocket. In this state, the SBD has a low affinity for substrates with fast kinetics of both binding and release (32). Upon ATP hydrolysis, SBD turns into a

so called *close* conformation, characterized by a much higher binding affinity (about 2 orders of magnitude decrease in the dissociation constant K_d) with lower on/off rates. The SBD is therefore able to sense the nucleotide status of the NBD, consequently adapting its conformation. Thus, the energy from ATP hydrolysis can be used to efficiently regulate peptide substrate binding and release, whereby the conformation of the SBD changes between at least two states in a nucleotide-dependent manner. In particular, the hydrolysis of ATP induces an increase in the affinity of Hsp70 for the substrate in a process that is synergistically stimulated by substrate binding. Moreover, it has recently been observed that the low intrinsic ATPase activity of Hsp70 increases ca. ten times just upon substrate binding (33,34), indicating that the two domains communicate in a bidirectional manner.

The chaperone activity of Hsp70 is dramatically enhanced by two classes of co-factors: the J-domain co-chaperones (JDPs) and the prokaryotic nucleotide exchange factors (NEFs) (35). JDPs represent a heterogeneous class of multi-domain proteins that share a conserved stretch of approx. 70 residues (the J-domain), essential for the interaction with Hsp70. This class of proteins includes the *E. coli* DnaJ co-chaperone and eukaryotic heat shock protein 40 kDa family members (Hsp40s). It has been clearly demonstrated that JDPs are able to stimulate the low intrinsic NBD ATPase activity of around 1000 fold, even if the molecular mechanism of this process has not yet been clearly characterized. Hsp40s also interact directly with unfolded polypeptides and they can recruit Hsp70 to protein substrates (36). Following ATP-hydrolysis, various NEFs (GrpE in bacteria; Bag, HspBP1 or Hsp110 in eukaryotes) bind to the Hsp70 NBD and catalyse ADP-ATP exchange, which results in substrate release, thereby completing the reaction cycle (Fig. 1.2B).

Apart from assisting non-native protein folding, Hsp70 also plays an important role in protein aggregation. Actually, Hsp70 has been demonstrated to prevent the aggregation process by directly interacting with hydrophobic patches of substrate molecules that potentially could form harmful inter-molecular interactions (‘holder’ activity). Moreover, Hsp70 has been reported to be able to solubilize and refold protein aggregates in cooperation with Hsp100 family chaperones. This system has been extensively characterized for the chaperones DnaK and ClpB (*E. coli* Hsp100 homologs); in particular ClpB recruitment has been shown to be regulated by the size of the aggregates: small aggregates (2- to 3-mers) of glucose-6- phosphate

dehydrogenase were efficiently solubilized and refolded just by DnaK alone. Larger aggregates (up to 10-mers) could be only partially solubilized by DnaK, which then was necessary in a high molar excess. In the presence of ClpB these larger aggregates were disaggregated more efficiently even at sub-stoichiometric concentrations of DnaK (37). Starting from these data, Goloubinoff and co-workers proposed a model, in which the relatively small amount of Hsp70 molecules retains, once bound to the large protein aggregates, a higher thermal mobility as compared to the aggregate. This difference in thermal mobility contributes to generate a force that leads to the solubilisation of the aggregated proteins (38).

1.4 Hsp70 oligomerization is mediated by an interaction between the interdomain linker and the helical lid of the substrate-binding domain: implications for chaperone allosteric mechanism

Francesco A. Aprile[1,2], Anne Dhulesia[1], Florian Stengel[3], Cintia Roodveldt[4], Justin L Benesch[3], Carol V. Robinson[3], Xavier Salvatella[5], Christopher M. Dobson[1,*] and Nunilo Cremades[1,*]

[1] Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW Cambridge, United Kingdom.
[2] Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza dellaScienza 2, 20126 Milan, Italy.
[3] Department of Chemistry, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK
[4] CABIMER-Andalusian Center for Molecular Biology and Regenerative Medicine, Consejo Superior de Investigaciones Científicas, University of Seville-UPO-Junta de Andalucia, Seville, Spain
[5]Institute for Research in Biomedicine, BaldiriReixac 10, 08028 Barcelona, Spain.

*corresponding author

ABSTRACT

Oligomerization in the heat shock protein (Hsp) 70 family has been extensively documented and even associated with possible functional roles related to the transient character of substrate binding in the chaperone. However, the oligomerization mechanism and the specific protein regions involved are still unclear, being subjects of intense debate in the field. To address these issues, we have studied the oligomeric properties of a various human Hsp70 variants by means of nanoelectrospray mass spectrometry and quantitative size exclusion chromatography and found that, for some of the truncated substrate binding domain (SBD) species, oligomeric forms are highly abundant in solution even at low micromolar concentrations. Hsp70 oligomerization takes place by a specific interaction between helices B to E of the helical lid subdomain in the SBD of one molecule and the leucine-rich tract of the interdomain linker of a different molecule,

generating dimers and higher order oligomers by a domain swapping-like mechanism. Interestingly, this interaction presents an affinity in the same order of magnitude as typical chaperone substrate binding, with which it competes for the optimal spatial orientation of the helical lid subdomain. Based on our findings, we propose that such an interaction, in addition to allowing for the formation of chaperone oligomers, could also take place intramolecularly, playing a possible role in Hsp70 interdomain communication.

INTRODUCTION

The 70-kDa heat shock proteins (Hsp70s) are essential members of the cellular chaperone machinery, assisting protein folding, disaggregation, and trafficking, and their malfunction has been associated with several diseases, including cancer, neuro-degeneration, allograft rejection, and infection (39,40). They are ubiquitously expressed in organisms ranging from *Archaea* to *Homo sapiens* and they are by far the most conserved proteins in evolution (22-24): for example, the prokaryotic Hsp70 protein DnaK from *Escherichia coli* shares approximately 50% amino acid identity with eukaryotic Hsp70s (25), which has allowed the vast amount of results obtained from DnaK to be extrapolated to the entire Hsp70 family.

Hsp70 proteins consist of two structurally independent domains: a conserved nucleotide-binding domain (NBD) with ATPase activity, which corresponds to residues 1-384 in *Homo sapiens* Hsp70 (Hs-Hsp70), and a more variable C-terminal substrate-binding domain (SBD), residues 397-641. Two subdomains can be differentiated within the SBD: a β-sandwich subdomain that contains the hydrophobic substrate-binding cleft (residues 397-508), and a 10-kDa subdomain of α-helical structure (residues 509-641), which has been suggested to play a key role in regulating the kinetics of substrate binding (27,41). The helical subdomain comprises five helices (named helix A-E) that act as a lid over the β-sandwich subdomain, with direct contacts between α-helices A and B and the β-subdomain (42-45). This region has been shown to be highly dynamic and this dynamic behaviour has been found to be in fact a requisite for the chaperone to be able to accommodate a broad spectrum of client polypeptides into the binding pocket (28). Nevertheless, the details of the substrate binding mechanism remain unclear and the question about possible interactions between the helical lid and substrates is still not yet fully resolved. Indeed, although there is no evidence for

direct contacts in the reported crystal structures (42), the helical subdomain has been shown to be essential in stabilizing the Hsp70-substrate complexes (46-48), and it might induce conformational changes in the bound substrates, which could be important for chaperone function (28). In agreement with this idea, a lidless variant of DnaK has been reported to be unable to refold chemically denatured firefly luciferase, suggesting that the lid may actually play an important part in the refolding of some substrates (41). Finally, the NBD and the SBD are connected by a hydrophobic linker of 11 residues (385-396) that carries a highly conserved leucine-rich motif. Recent evidence suggests that this linker might play an active role in the allosteric communication between domains and in the recruitment of the co-chaperone Hsp40 (33,34,49-53).

Oligomerization has been reported for an important number of Hsp70 family members, including Hsc70 (54-61), mouse inducible Hsp72 (62,63), BiP (64-66), Hsp70 from *C. Elegans* (67), and DnaK from *E. Coli* (68), which has led to suggest that the oligomerization process in Hsp70 homologues could be involved in the chaperone functionality. Specifically, it has been observed that the ADP-bound form of the chaperone leads to self-oligomerization *in vitro* in a concentration- and temperature-dependent manner, process that is reversed with the addition of ATP, substrate binding, and some co-chaperones (54,55,68). In addition, dimers and higher order oligomers of DnaK (69), hamster BiP (Grp78) (65) and human Hsp70 (70) have been found in cellular extracts, which become more prevalent upon heat shock and lower levels of ATP. However, the relevance of these oligomeric species in the chaperone function is still unclear, with a few studies suggesting an important role of the oligomerization process in the chaperone cycle, mainly as a regulatory mechanism for the availability of the active form of the chaperone (being monomeric chaperone the only functionally active form) (55,70) or for the selection of different functional forms of the chaperone, depending on their aggregation state (69).

In spite of the considerable amount of studies reporting the formation of Hsp70 oligomers both *in vitro* and *in vivo*, the mechanism by which these oligomeric species are formed still remains unclear. Different locations for the interactions leading to oligomerization have been proposed, including regions within the β-subdomain of the SBD, such as the region composed of residues 385-646 (55) and 385-540 (56) in bovine Hsc70, as well as regions in the helical lid subdomain of SBD, such as the tract between 554-646

residues in rat Hsc70 (71) or even regions in both subdomains (residues 382-561 of human Hsp70 (72)). Interestingly, a crystal structure of a homodimer generated using a truncated variant of rat Hsc70 consisting of the last 96 residues of the protein (residues 542-646) has been reported (71). With the aim of shedding more light on the specific protein regions involved in Hsp70 oligomerization, as well as on the possible role that this process could play in Hsp70 chaperone function, we have characterized the oligomerization propensity and its impact on substrate binding of a series of *Homo sapiens* Hsp70 1A constructs using nanoelectrospray mass spectrometry (nESI-MS), analytical size exclusion chromatography (SEC), fluorescence and circular dichroism (CD).

MATERIALS AND METHODS

Protein expression and purification. Recombinant N-hexa-His-tagged Hsp70 (Hsp70 1A, gi:194248072) and all the truncated SBD variants were overexpressed from pET-28b vector (Novagen) in *E. coli* BL21 (DE3) Gold Strain (Stratagene) and purified by affinity chromatography and size exclusion, as previously described (73). The monomeric fraction was then isolated by size exclusion chromatography (SEC) using a Superdex 26/60 G75 column (GE Healthcare). Thrombin cleavage efficiency, estimated by mass spec analysis, was greater than 99% and the protein purity exceeded 95% as determined by SDS–PAGE gel. The protein was then aliquoted, nitrogen flash-frozen and stored at -80 ºC, and each protein aliquot was thawed only once before using it.

Protein concentrations were determined by absorbance measurements at 280 nm using theoretical extinction coefficients calculated with Expasy ProtParam (74).

Circular dichroism. The far-UV CD spectra, as well as the thermal unfolding scans, of all protein variants were performed using a *Jasco810* spectropolarimeter equipped with a Peltier holder, using a 0.1 cm path length cuvette. Typically, samples contained 20 µM protein in 7 mM Tris pH 7.4, 170 mM KCl, 5 mM $MgCl_2$. The far-UV CD spectra of all our variants in their native, thermally denaturated and refolded states were recorded from 198 to 250 nm, and the spectrum of the buffer was systematically subtracted from the spectra recorded from protein samples. The estimation of the secondary

structure content derived from the far-UV CD spectrum for each protein variant was obtained using K2D software (75,76).

Protein structural stability was analyzed by monitoring the CD signal at 222 nm upon thermal denaturation. Thermal runs were performed from 5 to 90 °C at a rate of 1 °C.min^{-1}. Data points were acquired every 0.1 °C with a response time of 2s and a bandwidth of 1mm. After thermal denaturation, the same protein sample was cooled downback to the starting temperature, and a second denaturation curve was recorded in order to analyse the degree of reversibility of the denaturation process. Analysis of the thermal unfolding curves was performed assuming a two-state unfolding model for the C-term variant, using the following equation:

$$S = \frac{S_N + m_N T + (S_U + m_U T) e^{-\left(\frac{\Delta G}{RT}\right)}}{1 + e^{-\left(\frac{\Delta G}{RT}\right)}} \qquad \text{Eq.1}$$

where the spectroscopic signal of each state (S_N and S_U at T=0 for the native and unfolded state, respectively) is assumed to vary linearly with temperature with slopes m_N and m_U, and ΔG, the free energy difference between native and unfolded state, is calculated using the following expression:.

$$\Delta G = \Delta H_m \left(1 - \left(\frac{T}{T_m}\right)\right) - \Delta C_p \left((T_m - T) + T ln \left(\frac{T}{T_m}\right)\right) \qquad \text{Eq.2}$$

Where ΔH_m refers to the enthalpy of unfolding at T_m, which is the melting temperature (temperature at which ΔG=0), and ΔC_P is the heat capacity change that characterizes the unfolding transition.

For SBD556 and SBD641, two apparent transitions were clearly observed and, therefore, a three-state model was used to fit the experimental data, using the following equation:

$$S = \frac{S_N + m_N T + (S_I + m_I T) e^{-\left(\frac{\Delta G_1}{RT}\right)} + (S_U + m_U T) e^{-\left(\frac{\Delta G_1 + \Delta G_2}{RT}\right)}}{1 + e^{-\left(\frac{\Delta G_1}{RT}\right)} + e^{-\left(\frac{\Delta G_1 + \Delta G_2}{RT}\right)}} \qquad \text{Eq.3}$$

where S_I and m_I are the spectroscopic signal of the intermediate state at T=0 and its linear temperature dependency constant, and ΔG_1 *and* ΔG_2 correspond to the free energy between the native/intermediate and

intermediate/unfolded equilibria, respectively, and are defined by expressions analogous to that of Eq.2.

Fluorescence spectroscopy. Steady-state tryptophan fluorescence emission spectra of native, thermally denaturated and refolded SBD641 and ΔLSBD641were recorded from 300 to 400 nm, selectively exciting the only Trp residue in SBD (W580) at 295 nm, using a Cary Eclipse spectrofluorimeter (Varian). Excitation and emission band-passes were set to 5 nm each. We also analysed the variant SBD556, which lacks the C-terminal tract of SBD in which W580 is located, as a negative control.

Mass Spectrometry. Prior to MS analysis, samples were buffer exchanged into 150 mM ammonium acetate, pH 7.4, using Micro Bio-spin 6 columns (Bio-Rad). For all the experiments a protein concentration ranging from 6 to 18 μM was used. In the case of the substrate binding experiment, SBD641 was analyzed in the presence of a sub-stoichiometric concentration of NR peptide (1:0.05 molar ratio). This substrate concentration, lower than in the SEC experiments, was chosen to guarantee a high enough resolution of the various charge state series, which is critical for unambiguous assignment of substrate-bound states. NR peptide also outcompeted the substrate for charge during the ionization process, thus limiting the ratios that could be investigated by MS.

Nanoflow electrospray mass spectra were obtained according to a previously published protocol (77) on an orthogonal time of flight mass spectrometer with a Z-spray source (Waters) in positive ion mode using the following parameters: capillary voltage, 1500 V; sample cone, 120 V; extraction cone,1V; and ion transfer stage pressure, 8.8 mbar. Complementary denatured spectra were obtained by adding AG501-X8 resins (BioRad) to the samples.

The relative abundance of the different oligomeric states was estimated from the intensities of the two most abundant charge states for each construct and is expressed as a percentage of the total intensity of the peaks assigned to each oligomer divided by the sum over all oligomeric states for each construct.

Spectra were calibrated externally using 33 $mg.ml^{-1}$ cesium iodide. Data was acquired and processed with MassLynx software (Waters), and are shown with minimal smoothing.

Analytical size exclusion chromatography. Analytical SEC was performed using a Superdex 75 10/300 GL column (GF Healthcare), which was previously calibrated using a mixture of standard proteins (Gel Filtration LMW Calibration Kit; GE healthcare). Typically, the samples contained 70 µM of protein in 50 mM Tris buffer pH 7.4, 150 mM KCl, 5mM $MgCl_2$ (the same buffer was used for all the variants) and were incubated for 15 min at room temperature before being loaded into the column for their analysis. To evaluate the reversibility of the oligomerization process, the different elution peaks obtained for each protein variant were subsequently analysed after concentrating the different fractions using AmiconUltra-0.5, Ultracel-3 Membrane, 3 kDa (Millipore). A similar analysis was also performed at 6 µM concentration for all the protein variants to allow for a direct comparison between the results obtained by SEC and MS experiments.

The influence of chaperone substrate binding on the oligomerization properties of all the variants was performed using the well-characterized Hsp70 substrate NR peptide (NRLLLTG, Genemed Synthesis Inc.). The peptide was added to a fresh sample of protein in a concentration 15 times higher than the protein, and the sample was loaded into the Superdex 75 10/300 GL column for its analysis after having incubated the mixture for 1 hour at room temperature.

The relative contribution of the different protein species was estimated by analysing the chromatographic profile using a Gaussian function for each elution peak (OriginPro 8, Origin lab. Corporation) and quantifying the relative area under each peak.

NR peptide binding to Hsp70 variants. The interaction between the different Hsp70 variants (except the C-term) and NR peptide was studied in 50 mM Tris buffer pH 7.4, 150 mM KCl, 5mM $MgCl_2$by fluorescence spectroscopy by titrating different concentrations of chaperone into a solution containing 2 µM DANSYL-NR peptide (NR peptide was N-terminally labeled; Genemed Synthesis Inc.) after incubating the mixtures for 30 min at 25°C. The concentration of DANSYL-NR peptide was calculated using an extinction coefficient at 327.25 nm of 4300 $M^{-1}.cm^{-1}$ (78). The fluorescence emission spectrum of each sample was recorded from 400 to 630 nm after excitation at 330 nm, and the reported data corresponds to the average intensity at the maximum emission wavelength of 2 different samples for which 10 spectra were recorded. The increase in fluorescence

signal was plotted as a function of chaperone concentration, and analysed assuming a one site binding model using the following equation:

$$F = \frac{Fmax \times [chaperone]}{Kd + [chaperone]} \quad \text{Eq.4}$$

where F is the fluorescence intensity observed at a given concentration of Hsp70, F_{max} is the fluorescence at saturation and K_d is the apparent dissociation constant of the complex (73).

RESULTS AND DISCUSSION

Design of truncated variants.

Different protein regions within the SBD of Hsp70 family members have been proposed to be involved in chaperone oligomerization, the majority located in either the substrate binding site (55,56) or in the C-terminal segment of the helical lid subdomain (67,71,72,79-81). Based on this information, we decided to generate a set of truncated SBD variants of human Hsp70 (Fig. 1A-B): SBD641 (388-641), corresponding to the full-length SBD, SBD556 (388-556), which lacks helices C, D and E and the C-terminal disordered tail, ΔLSBD641 (396-641), corresponding to SBD641 without the first 10 residues of the interdomain linker, and the C-terminal variant (540-641), consisting of the helices C-E and the disordered C-terminal tail of the lid subdomain. The SBD556 variant contains the L543Y mutation to prevent self-binding of L543 to the binding pocket, as previously reported (82). The truncation points were designed on the basis of published structures of the homologous proteins (1dkx, 7hsc, 1ud0) (Fig. 1B).

The structural integrity of all the variants, including full-length Hsp70, was studied at 25 °C using far-UV CD (Fig. 1C-D). Secondary structure content for all the variants at 25 °C, derived from the analysis of their far-UV CD spectrum, resulted to be in excellent agreement with the secondary structure content predicted from the available x-ray structures of NBD (from *Homo sapiens*; 1hjo) and SBD (from *E.coli*; 1dkx), suggesting that the different truncated

variants are able to fold maintaining the native-like structure of the two domains (see Table 1).

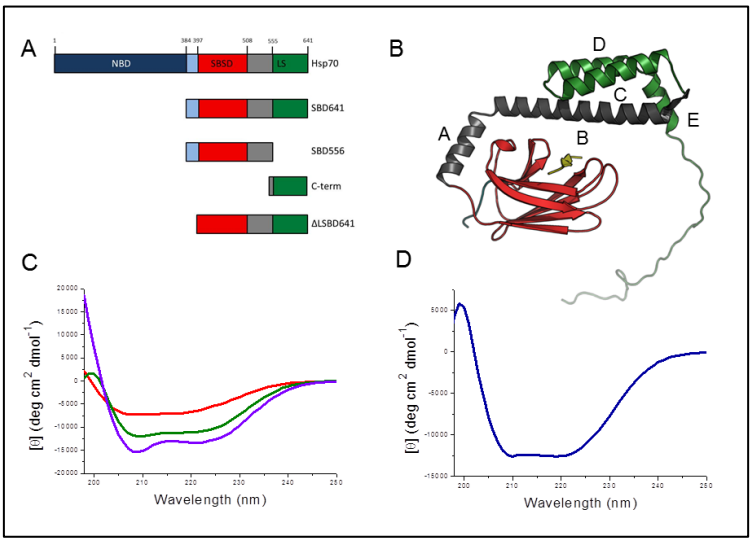

Fig. 1. (A) Schematic representation of the engineered human Hsp70 variants used in this study, where the different functional domains and truncated regions of the protein are shown (see text for a detailed description of the truncated variants). (B) Representation of the different truncated constructs on the crystal structure of the DnaK SBD in complex with the NR peptide substrate, represented in yellow (1dkx; (42)). (C) Far-UV CD spectra of native SBD641 (green), SBD556 (red) and C-term (violet) and (D) of FL-Hsp70 were recorded in order to verify that all the variants were well folded.

Table 1. Comparison of secondary structure content derived from far-UV CD spectra deconvolution using K2D software (75,76) and that according to the crystal structure of homologous proteins: Hu-NBD (pdb 1hjo; (83)) and DnaK-SBD (pdb 1dkx; (42)).

	Predicted			**K2D**		
	Alpha %	**Beta %**	**Random %**	**Alpha %**	**Beta %**	**Random %**
SBD556	22.6	26.9	50.5	21	25	54
SBD641	37.4	21.8	40.8	33	18	49
ΔLSBD641	37.4	22.5	40.1	35	18	46
C-term	64	0	36	60	0.07	33
Hsp70	39.8	22.1	38.1	46	23	32

Oligomerization analysis of Hsp70 variants using nanoelectrospray mass spectrometry and analytical size exclusion chromatography.

In order to study the oligomerization properties of the different Hsp70 variants by MS, we first checked the structural integrity of all the variants in the gas phase, by comparing the nanoelectrospray (nESI) mass spectra of each variant under native (Fig. 2) and denaturing conditions (Fig. S1) (77). In general, the charge states of the proteins shift towards high charge ranges (lower *m/z*) when examined under denaturing conditions (see Materials and Methods), indicative of a higher degree of unfolding (84). Intriguingly, the MS experiments on the C-term variant under these conditions suggest that, under the gas phase, it behaves similarly under both native and denaturing conditions. In Tris buffer at neutral pH (7.4) and 25 °C, however, the C-term variant shows a degree of helical structure according its far-UV CD spectrum (Fig. 1C) that is consistent with that expected from the crystal structures of homologous proteins (see Table 1). It is therefore likely to maintain a high degree of three-dimensional structure under acidified denaturing conditions relative to the native state.

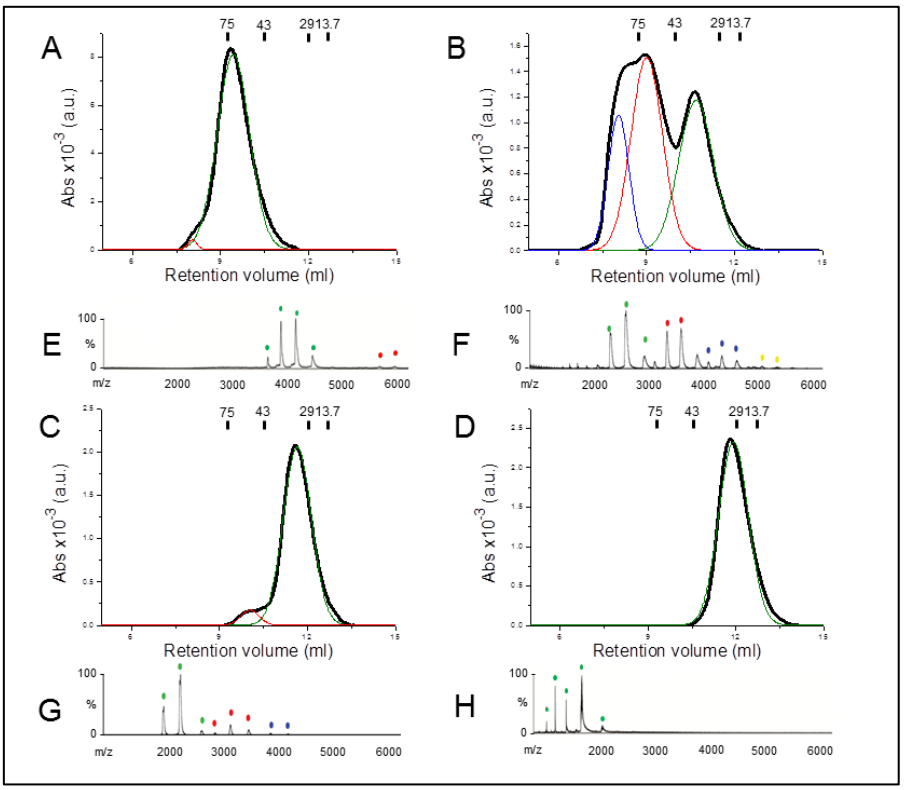

Fig. 2. Analytical SEC and nanoflow electrospray mass spectra of the different proteins constructs analyzed at the same protein concentration (6 μM):FL-Hsp70 (A, E), SBD641 (B, F), SBD556 (C, G) and C-term (D, H). The SEC elution profiles were fitted to multi-peak Gaussian functions to evaluate the relative fractions of monomeric and oligomeric protein species. On the top of each chromatogram the molecular weight of the following standard proteins used to calibrate the column is reported: conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa) andribonuclease A (13.7 kDa). The elution peaks in SEC and the peaks in MS corresponding to monomeric protein are coloured in green for all the variants, while those for the dimeric protein are in red, with trimers in blue and tetramers in yellow.

nESI MS has proven to be a powerful means for quantifying the relative abundances for polydisperse proteins (85). We have exploited this ability to assess the relative proportions of the monomers and oligomers in the different protein variants (Table 2) (Fig. 2E-H). The data indicate that some of the truncated protein variants show a clear tendency to oligomerize, even at the relatively low micromolar concentrations used for the MS experiments, and reveal important differences between them. SBD641 variant presents the highest degree of oligomerization, with the majority of the protein

in an oligomeric state (40% dimers, 12% trimers and 2% tetramers). SBD556, by contrast, exhibits a considerably smaller proportion of oligomers under the same conditions (ca. 18%), which strongly supports the hypothesis that the C-terminal part of the helical lid subdomain is directly involved in chaperone oligomerization. Interestingly, according to the MS analysis, both FL-Hsp70 and the C-terminal variant are essentially monomeric, and only FL-Hsp70 presents a very small percentage of dimers (around 2%).

Table 2. Relative amount of oligomeric forms of Hsp70 variants at 6 μM of protein concentration.

	MS					**SEC**			
	Mon %	**Dim %**	**Tri %**	**Tetr %**	**Tot olig%**	**Mon %**	**Dim %**	**> Dim %**	**Tot olig%**
Hsp70	98	2	0	0	2	99	1	0	1
SBD641	46	40	10	4	54	37	44	19	63
ΔLSBD641	92	8	0	0	8	91	9	0	9
SBD556	82	16	2	0	18	97	3	0	3
C-term	100	0	0	0	0	100	0	0	0

Estimated experimental error is about 5% for MS and 3.5% for SEC

The data obtained by MS in the gas phase was complemented with analytical SEC performed in parallel using the same protein concentrations. Very similar results were obtained (Fig. 2, and table 2). SBD641 was found to be mainly oligomeric (ca. 60 %), in agreement with the MS data, which confirmed a much higher oligomerisation propensity for SBD641 than for the other constructs. The rest of the protein variants remained mainly monomeric, with only SBD556 presenting a significant quantity of oligomers (around 5-20 % according to SEC and MS analysis, see table 2).

The C-terminal variant was found to elute in a single peak, in agreement with the MS results where only monomers were detected. The Stoke radius derived from the retention volume (21.4 Å) is consistent with that calculated from the crystal structure of the same truncated variant from *C. elegans* (22.5 Å), with which it shares 96% of sequence similarity (67). This result, however, contrasts with other

study where dimers of a very similar truncated variant of rat Hsc70 were reported and crystallized (71). To prove further that, in our experimental conditions, this protein variant is monomeric, samples at concentrations ranging from 1 to 125 μM, were analyzed by analytical SEC. In agreement with our predictions, a unique elution peak centred at exactly the same retention volume (11.8 ml) was obtained for all concentrations (Fig. S2). For the rest of the protein variants, on the contrary, a significant increase in the percentage of oligomeric species was observed when the protein concentration was increased from 6 to 70 μM (Table 3 and Fig. S3), which also indicates that the interactions that lead to oligomerization in Hsp70 present an affinity in the micromolar range.

Table 3. Analysis of the effect of protein concentration on chaperone oligomerization.

	6 μM			**70 μM**		
	Mon%	**Dim%**	**> Dim%**	**Mon%**	**Dim%**	**> Dim%**
Hsp70	99	1	0	93	7	0
SBD641	37	44	19	23	55	22
ΔLSBD641	91	9	0	86	14	0
SBD556	97	3	0	93	7	0
C-term	100	0	0	99	1	0

Comparing the initial SEC chromatograms with those obtained by individually re-loading the monomeric and oligomeric fractions in the same chromatography column (see Materials and Methods), we were able to assess the reversibility of the oligomerization process. Both isolated monomeric and oligomeric fractions of FL-Hsp70 and all truncated variants (except the C-terminal variant) led to chromatograms with two elution peaks, corresponding to the same monomeric and oligomeric peaks originally obtained (Fig. S4), indicating that chaperone self-assembly process is a rapid, reversible equilibrium between monomeric and oligomeric species.

Influence of the interdomain linker on oligomerization.

The large difference between the fraction of oligomeric species formed with the SBD641 and SBD556 variants (see Table 2), which only differ in the presence or absence of the C-terminal part of the lid subdomain (Fig. 1A-B), indicates that this C-terminal region of the protein is directly involved in protein oligomerization, as previously suggested (62,71,72). In addition, we also found that the C-terminal region is unable to dimerize in isolation, at least in our experimental conditions. These results, together with the fact that the C-terminal/C-terminal interaction would not allow the formation of higher order oligomers such as trimers and tetramers (as observed in this and other previous studies on homologous Hsp70 proteins (55,56,69,70), led us to search for a different region the helical C-terminal lid of the protein interacts with forming different oligomeric species.

NBD and SBD are connected by a short hydrophobic linker (11 residues), which was included originally in all the truncated SBD protein variants containing the β-sheet subdomain. To explore the possibility of an interaction between the C-terminal tract of SBD and the interdomain linker, we generated a modified SBD641 variant, denoted as ΔLSBD641, where the leucine-rich tract of the interdomain linker was removed. This variant is structurally identical to the original SBD641 as assessed by far-UV CD (Fig. 3A), although its propensity to oligomerize was drastically reduced. Only a small fraction of dimers, representing around 10% of all protein species, was observed by both analytical SEC and MS (Fig. 3B-C), unlike the SBD641 original variant, where the oligomeric species represent more than 50%. A similar result was recently observed in rat Hsc70 where the deletion of the interdomain linker in the full-length SBD truncated variant resulted in a drastic reduction of protein oligomerization (86).

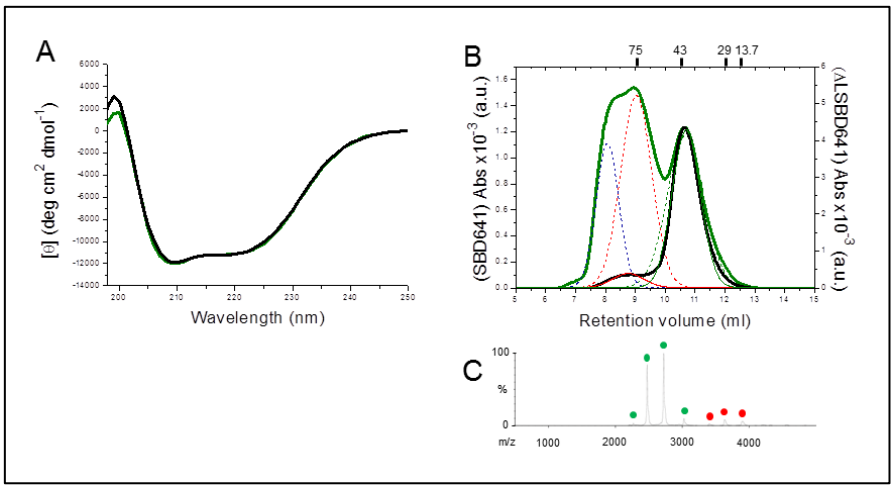

Fig. 3. (A) Far-UV CD spectrum of native ΔLSBD641 (black) compared to that obtained for SBD641variant (green). (B) Comparison of the SEC chromatograms obtained for SBD641 (in green) and ΔLSBD641 (in black) at 6 μM protein concentration. The molecular weight of standard proteins is reported on the top of the chromatogram (Fig. 2). (C) MS analysis on ΔLSBD641 (monomer in green, dimer in red).

This result provides some insight on oligomerisation process of Hsp70, where two distinct regions of the protein, located in different folding and functional domains, are involved: the C-terminal tract of the lid subdomain of the SBD, and the leucine-rich motif of the interdomain linker between the SBD and the NBD. The differences in the oligomerization propensities between the different constructs can indeed be explained by the presence or absence of one of these two oligomerization domains: the C-term construct does not oligomerize because it is lacking the interdomain linker, whereas the SBD556 construct is mainly monomeric as it is lacking the C-terminal region. Both the SBD641 and FL-Hsp70 are shown to oligomerise to some extent, due to the presence of both regions in the construct, although to a different extent: the oligomerisation is indeed much less pronounced in FL-Hsp70 than in SBD641. This suggests that the ability of the SBD lid subdomain of one protein molecule to have access to the interdomain linker of another molecule is very different in SBD641 or FL-Hsp70. In SBD641, the interdomain linker would be fully accessible to the lid subdomain of another molecule, whereas in the full-length protein, it is very likely that there are steric repulsion between the NBD of one molecule and the SBD of another molecule

hampering the intermolecular interaction between the lid subdomain and the interdomain linker, and therefore chaperone oligomerization.

The interaction between the interdomain linker and the helical lid subdomain of SBD is lost following irreversible thermal misfolding of helices C-D.

With the aim to understand the nature of the interaction between the interdomain linker and the C-terminal part of SBD, we analysed the structural stability of each protein variant in relation with its oligomerization propensity. We first performed thermal denaturation experiments, by monitoring the secondary structure content of each protein variant as a function of temperature (Fig. 4). In the case of FL-Hsp70, however, this analysis was not possible because of the aggregation of the protein triggered at temperatures above 37 °C, as revealed when the thermal unfolding of the protein was followed by light scattering measurements at 400 nm: this aggregation was not observed in any of the SBD truncated variants (light scattering experiments for these variants not shown).

The far-UV CD thermal unfolding of SBD556 and SBD641 variants was found to be very similar, with two well-resolved transitions that we have analysed using a classical three-state model (Fig. 4A). The first transition is essentially identical for both variants, with the temperature of the mid-point of the transition being 45.8±1 °C, whereas the second transition shows marked differences between variants, with a T_m of 64.5± 0.2 °C for SBD556 and 73.1± 0.2 °C for SBD641 (see Table 4). This result indicated that both variants share a structural motif in common that denatures at relative low temperatures, but have another structural motif that behaves differently and is more stable in SBD641. Since both variants have the intact β-sheet subdomain in common, it is reasonable to assume that the lower temperature transition corresponds to the denaturation of this β subdomain and that the higher temperature transition is related to the unfolding of the helical lid subdomain, which corresponds to helices A and B in the case of SBD556, and to the complete subdomain in SBD641. This assumption is in agreement with the differences in signal at 222 nm for each transition between proteins: similar signal at 222 nm is lost during the first transition for both proteins, while the difference in signal in the second transition in SBD641 is much higher than in SBD556, as expected for the additional 3 α-helices (helices C-E) in the helical subdomain of

SBD641 (Fig. 4A). These results reveal that helices A and B are able to fold independently, but the presence of helices C-E confers an additional stabilization to the full helical subdomain, resulting in a significant increase in the T_m for the second transition in SBD641 with respect to SBD556.

Table 4. Thermodynamic parameters of protein thermal stability (ΔHm (kcal*mol^{-1}) and Tm (°C)) obtained for the different Hsp70 truncated variants.

	1st transition				**2nd transition**			
	ΔHm1	**Tm1**	**ΔHm2**	**Tm2**	**ΔHm1**	**Tm1**	**ΔHm2**	**Tm2**
SBD556	63.9 ± 1.9	46.8 ± 0.1	85.4 ± 2.6	64.5 ± 0.1	71 ± 3.7	45.5 ± 0.2	80.3 ± 3.7	64.9 ± 0.2
SBD641	49.4 ± 3.3	45.6 ± 0.4	53.3 ± 1.2	73.1 ± 0.1	46.6 ± 4.8	43.9 ± 0.5	50 ± 1	72.9 ± 0.1
ΔLSBD641	45.6 ± 2.4	44.9 ± 0.3	61 ± 1.1	72.3 ± 0.1	38.5 ± 2.7	46.3 ± 0.5	60.5 ± 1.4	72.1 ± 0.1
C-term	51.5 ± 1.5	49.9 ± 0.2	-	-	59.6 ± 2	52.2 ±0.2	-	-

Comparing the degree of reversibility of the thermal unfolding process in SBD641 and SBD556 reveals differences between the two constructs. While the unfolding of SBD556 is fully reversible, the one of SBD641 is not, indicating that the unfolding/folding process of the C-terminal lid subdomain is not fully reversible. However, when we studied the thermal denaturation of the C-terminal variant, we found that this truncated variant, in addition to adopting a thermodynamically stable structure (with a T_m of 50.4 ± 0.2 °C; see Table 4), denatures both cooperatively and reversibly (Fig. S6A).

To characterize further the origin of the irreversibility of the thermal unfolding of the C-terminal part of the lid subdomain in the SBD641 construct, we assessed the solvent accessibility of the only tryptophan residue in SBD (W580), located at the end of helix C, in both the native, thermally unfolded and refolded states of SBD641. We observe that, in the refolded conformation, W580 presents fluorescence properties that are intermediate between the native state and the thermally denatured conformations (Fig. 4B). The maximum fluorescence emission of the native state of the protein is located at

340 nm and it is red-shifted to 350 nm when the protein is thermally denatured, which is typical for a transition from a highly hydrophobic environment in a native protein form to a hydrophilic (solvent exposed) environment in the denatured state of the protein. The spectrum of the refolded protein presents its maximum of intensity at 345nm, an intermediate wavelength between that found for the native and denatured states (Fig. 4B). The same results were obtained with ΔLSBD641 variant (Fig. 4D).

In order to check whether the observed degree of irreversibility (ca. 20%) found in the SBD641 far-UV CD thermal unfolding curves is able to account for this red-shift of the fluorescence emission spectra in the refolded state, we calculated the fluorescence spectrum of the protein corresponding to 80% of the molecules in the native state and 20% of the molecules in the thermally denatured state, assuming that the total fluorescence signal at one given wavelength is a linear combination of the fluorescence signals for each populated state. The spectrum calculated in this way, however, does not even qualitatively agree with the experimental spectrum recorded for the refolded form of the protein. In addition, the intensity signal at the maximum wavelength is higher for the refolded state than either the native or the unfolded state. These results indicate that the region where W580 is located has a much higher degree of irreversibility during thermal unfolding than the value of 20% observed for the unfolding of the helical subdomain by far-UV CD, and that this region does not recover either the native or the fully-denatured conformation following thermal denaturation. Instead, it adopts an alternative conformation where the region around helix C is partially solvent exposed. When this type of analysis was performed on the C-terminal variant, however, the red-shift in the maximum wavelength of the fluorescence emission spectrum from the native to the refolded form of the protein was not observed (Fig. S6B), suggesting that the thermal irreversibility of the C-terminal region of the helical subdomain much likely arises from a non-native interaction of this region with the rest of the protein upon refolding.

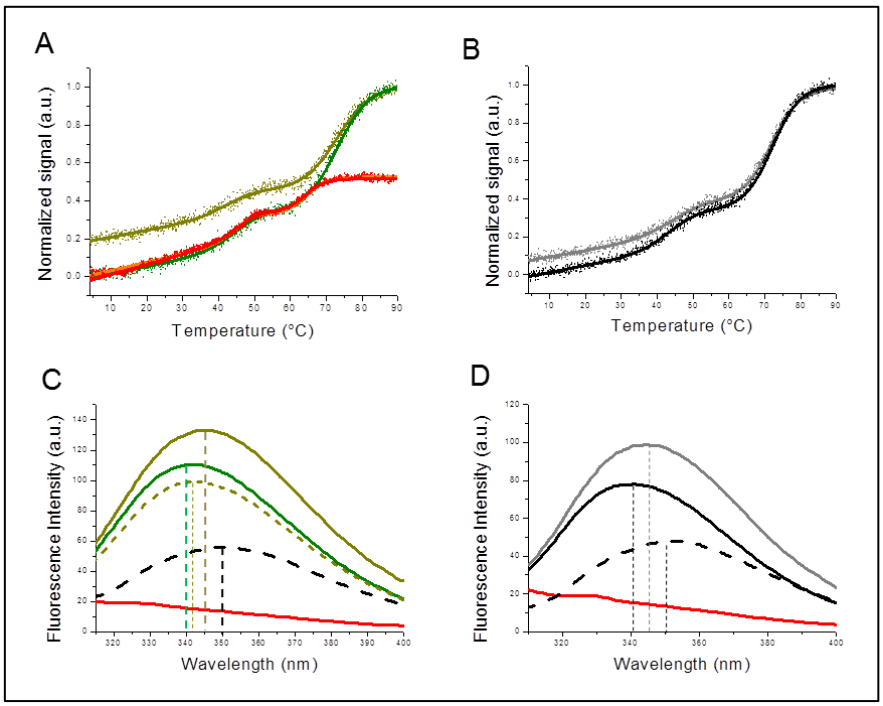

Fig. 4. Far-UV CD thermally-induced denaturation of the native and refolded proteinSBD556 (red and orange points, respectively) and SBD641 (dark green and brown, respectively) in A) and ΔLSBD641 (black and grey) in B). Triptophan (W580) fluorescence spectra of SBD641 (C) and ΔLSBD641 (D) in their native states (dark green and black continuous line for SBD641 and ΔLSBD641, respectively), thermally denatured state (black dashed line in both C) and D)) and refolded states (brown and grey continuous lines for SBD641 and ΔLSBD641, respectively); the spectrum of SBD556 in its native state is shown as a negative control (red continuous line). The calculated spectrum of SBD641corresponding to 80% of the protein molecules in the native structure and 20% of the molecules in the thermally denatured state (brown dashed lines; see text and Materials and Methods section for a detailed explanation) is also shown for comparison with the spectrum of the protein in its refolded state.

The degree of reversibility of the thermal unfolding of ΔLSBD641 variant observed is higher than that of SBD641 (92% and 80 %, respectively, Fig. 4A, C). The only difference between these two variants is the absence or presence of the interdomain linker, which also causes a different oligomerization profile in both variants. To test if the irreversibility of the unfolding of the C-terminal segment of the protein affects the oligomerization of the protein, we performed analytical SEC experiments on the refolded form of SBD641. The chromatogram obtained showed a dramatic increase in the monomeric

fraction of the refolded protein compared to the native protein (from 37% to 71%) (Fig. 5). That the refolded protein is less prone to oligomerize than the native protein rules out the involvement of the last C-terminal intrinsically disordered region of the protein (last 35 residues) in the oligomerization process, as this region is originally unfolded and no differences should be expected between the native and refolded protein. Instead, this finding locates the oligomerization domain of SBD in the helical region, around helices B-E (note that helix B is likely involved to some extent in chaperone oligomerization as SBD556 is still able to form oligomers, although to a much lower extent than SBD641), with helix C playing a major role, according to our thermal unfolding analysis and as previously reported (62,71). In addition, our findings demonstrate that the helical part of the lid subdomain needs to be in a native-like conformation to be able to establish interactions with the leucine-rich motif of the interdomain linker and then be able to oligomerize, since the misfolding of this region prevents oligomerization. This requirement, in turn, highlights the high degree of specificity of the interaction between the SBD and the interdomain linker. The helical lid subdomain contains seven leucine residues that are spread around helices B-E in hHsp70, which could be able to establish leucine/leucine interactions with the highly conserved leucine-rich motif of the interdomain linker in a similar way as in proteins with leucine-rich repeats such as ribonuclease inhibitor or toll-like receptor (87-89).

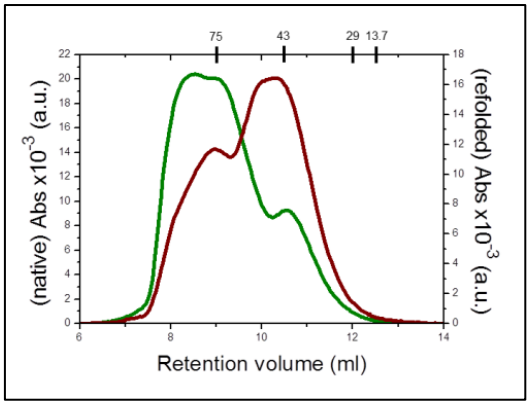

Fig. 5. Analytical SEC performed on native and refolded (after thermal denaturation of the protein) conformations of SBD641 (green and brown continuous lines, respectively). The molecular weight of the standard proteins used to calibrate the column is reported on the top of the chromatogram.

Across prokaryotes and eukaryotes, the NBD, the beta-sandwich subdomain in SBD and the interdomain linker are well conserved in both sequence and structure. The C-terminal subdomain in SBD is, however, more evolutionary variable both in sequence and structure (67). This variability in the helical SBD subdomain could explain the differences in oligomerization propensity found for various members of the Hsp70 family. We notice here that there is a significant variability in the number and possibly in the structural location of some of the leucine residues present in helices B-E between Hsp70 homologous (not shown).

Interaction with substrates.

In order to evaluate the influence of chaperone oligomerization on substrate binding, we characterized the ability of some of the protein variants we utilized, with different oligomerization propensities, to bind NR peptide (NRLLLTG) as a model susbstrate, (42,45,90). NR was N-terminally labelled with a DANSYL fluorophore and complex formation was studied by titration experiments, adding increasing concentrations of chaperone variants into a solution containing DANSYL-NR and following the fluorescence properties of the DANSYL moiety, whose quantum yield increases from a hydrophilic environment (i.e. free in solution) to a hydrophobic environment (i.e. in complex with Hsp70-SBD). All protein variants analyzed were able to bind to the NR peptide (Fig. 7, Table 5), with the exemption of C-terminal variant (not shown).

Table 5. Apparent K_d (μM) of the interaction of the different Hsp70 protein variants with NR peptide.

	App. K_d
SBD556	53.0 ± 5
SBD641	12.4 ± 1
Hsp70	9.8± 0.8
Hsp70/ADP	8.7± 1.3
ΔLSBD641	9.1± 1.2

Full length SBD was found to bind the NR peptide with an identical affinity (within error) as FL-Hsp70 (K_d between 9 and 12 μM) and independently of the presence or absence of the interdomain linker (Fig. 7, Table 5), results that are in concordance with previous

studies where SBD was found to behave independently from the NBD in the nucleotide-free state of both DnaK and human Hsp70 (34,52,91,92). In contrast, SBD556 showed a reduced ability to bind the NR peptide (K_d = 53 ± 5 μM), a feature that was also found for a similar truncated variant in its bacterial homologous DnaK (46,47,93). This result indicates that helices C-E are also important, although not essential, for substrate binding, although this region does not make direct contacts with the NR peptide according to the crystal structure of the complex in DnaK (pdb code 1dkx; (42)). It is likely to play an important role in stabilizing the substrate binding site in a manner similar to that previously proposed for the role of helix B in substrate binding in DnaK (42,46,47,94). The increase in Kd of a DnaK mutant protein lacking the C-terminal 100 residues (DnaK 2-538) with respect to wild type DnaK for two different peptide substrates was found to result mainly from a considerable accelerated dissociation constant rate, while the association constant rate was unaffected (41). All these results indicate that the entire helical lid sub-domain is needed to stabilize the substrate:chaperone complex once formed, preventing it from dissociation. The helical lid subdomain then seems to regulate the kinetics of substrate release, which should be further regulated by allosteric communication between NBD and SBD depending on the nucleotide status of NBD and where the interdomain linker has been found to be a major player (33,34,49,95).

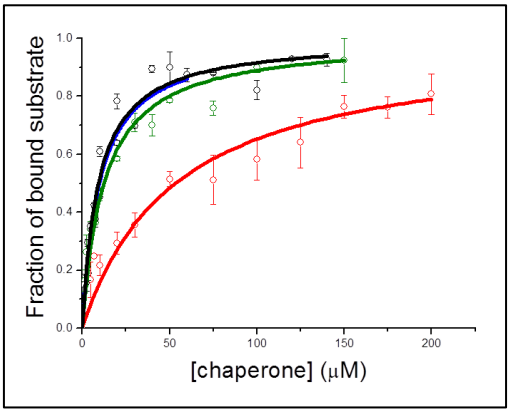

Fig. 6. Fluorescence titration of the binding of the NR peptide to the different Hsp70 constructs obtained by fluorescence titrations: SBD556 (red points), SBD641 (green points), ΔLSBD641 (black points), and FL-Hsp70 (blue points). The continuous lines represent the best fit of the data to a single binding site model.

Influence of substrate binding on oligomerization.

The fact that SBD641 and ΔLSBD641 variants have the same apparent affinity for NR peptide in spite of the fact that these two constructs have drastically different fractions of oligomeric species, lead us to consider that not only monomeric SBD, but also oligomeric species are able to bind NR peptide. To analyse this possibility, we performed MS analysis of SBD641 under native conditions in the presence of NR peptide. NR peptide was found to outcompete for charge during the ionization process, which limited the protein:ligand ratios that could be investigated by MS. A 1:0.05 molar ratio of protein versus peptide was chosen to guarantee the various charge state series to be unambiguously assigned. As it can be seen in Fig. 8, both SBD641 monomers and dimers were found to be in complex with one or two molecules of NR peptide. The fact that dimeric SBD can bind two NR peptide molecules agrees with our previous results where a very similar apparent affinity was found for SBD641 and ΔLSBD641, and demonstrates that the substrate binding pocket is not involved in chaperone oligomerization, in contrast with previous propositions (64).

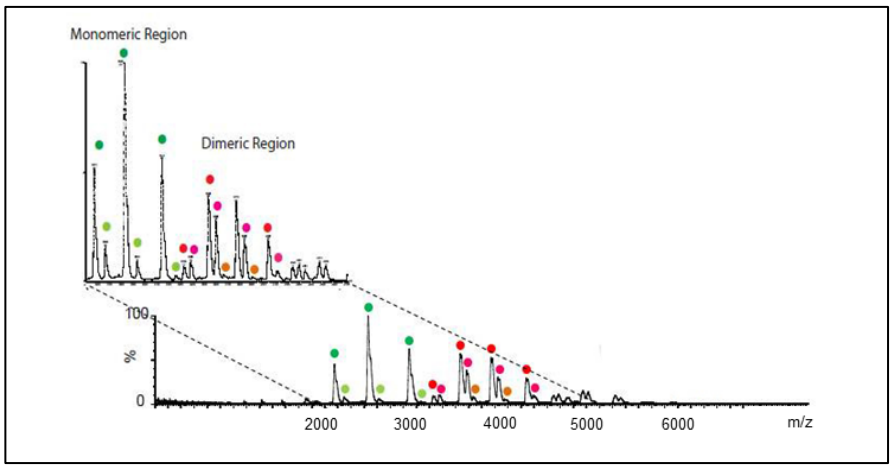

Fig. 7. MS analysis of SBD641 in the presence of the NR peptide in a protein:peptide ratio of 1:0.05. The coloured points represent charge states that correspond to the same protein species: uncomplexed monomeric protein in dark green, monomeric SBD641 complexed with the NR peptide in light green, the uncomplexed dimeric protein in red, dimeric SBD641 complexed with 1 NR peptide molecule in pink, and dimeric SBD641 complexed with 2 NR peptide molecules in orange.

From the MS analysis of the SBD641-NR interaction, we observed a preferential binding to the monomeric form of the protein with respect to the oligomeric form (Fig. 7 and 2F). This indicates the affinity for NR peptide to be slightly higher in the monomeric state of the chaperone, in comparison with the oligomeric state. In the monomeric protein, the complex is stabilized by the helical lid subdomain in the same way as in ΔLSBD641 or FL-Hsp70 variants, with an apparent dissociation constant around 9-10 μM. Whereas in the oligomeric species, the helical lid interacts with the interdomain linker of another protein molecule, and therefore it cannot stabilize the complex, which results in a decreased apparent affinity for NR peptide similar to that found for the SBD556 (dissociation constant of ca. 50 μM; see table 5). These slight differences in affinities for NR peptide between protein species would result in a ligand concentration dependent re-distribution of species upon substrate binding.

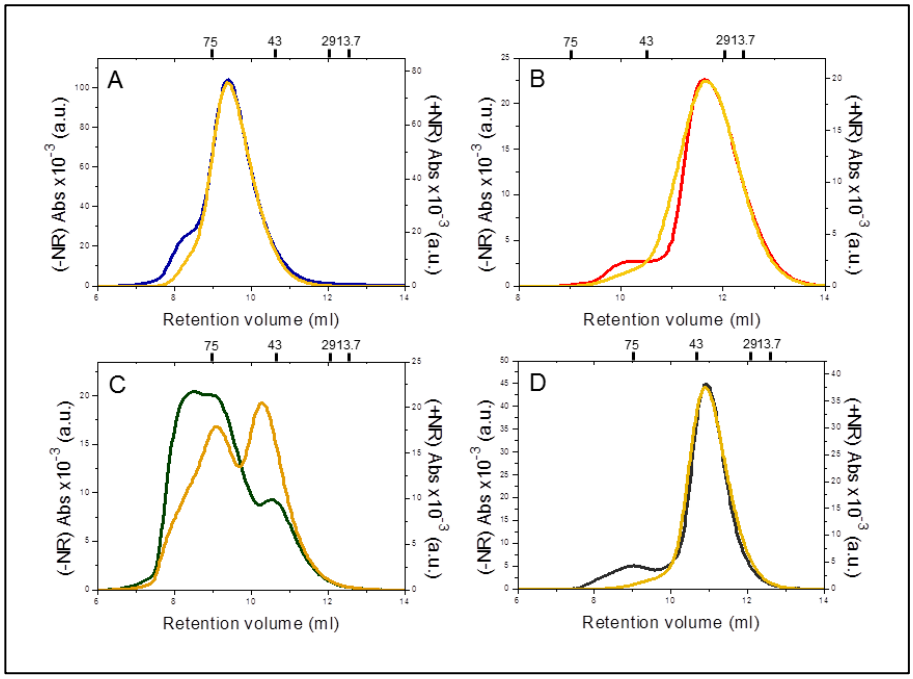

Fig. 8. Analytical SEC each for FL-Hsp70 (A), SBD556 (B), SBD641 (C) and ΔLSBD641 (D) in the presence of a 15fold excess of the NR peptide (orange continuous lines) in comparison with the profiles obtained for the free protein. On the top of each chromatogram the molecular weight of the standard proteins used to calibrate the column are reported.

Table 6. Relative amount of oligomeric forms of Hsp70 variants in presence or absence of an excess of NR peptide obtained by analytical SEC analysis.

	- NR			**+ NR**		
	Mon%	**Dim%**	**> Dim%**	**Mon%**	**Dim%**	**> Dim%**
Hsp70	83	7	0	99.3	0.7	0
SBD641	23	55	22	58	24	18
ΔLSBD641	86	14	0	98.2	1.8	0
SBD556	93	7	0	98	2	0

The substrate:protein ratio used in tese experiments was 15:1. The samples were previously incubated at room temperature for 1 h.

To investigate further the effect of substrate binding on chaperone oligomerization, we performed SEC analysis on all our variants upon incubation of the proteins with a 15-fold excess of the NR peptide. A relatively high excess of peptide with respect to the protein was used to ensure a significant change in the distribution between species according to the estimated affinities. In agreement with our previous results, we observed a dramatic accumulation of the monomeric fraction in the protein samples incubated with excess of NR peptide with respect to the protein samples without peptide (Fig. 9). Upon NR peptide addition, the monomeric fraction of SBD641 increased from 37% to 58% with respect to the free protein, whereas dimers and higher order oligomers were significantly reduced (to 24% and 18% respectively, in comparison with 55 and 22% in absence of peptide). A similar result has previously been observed for the BiP chaperone purified from bovine liver upon incubation of the chaperone with the NR peptide (64), where the observed competition between substrate binding and chaperone oligomerization was interpreted as a competition for the substrate binding site. However, we here demonstrate that this effect is not due to a direct competition between binding sites, but, instead, for the optimal spatial orientation of the SBD helical lid subdomain either for stabilization of substrate:chaperone complex or for chaperone oligomerization through its interaction with the interdomain linker of another protein molecule. Indeed, the entire helical lid subdomain of SBD has been found to occupy different relative positions with respect to the □-sheet subdomain, which has been proposed to be one of the mechanism by

which the chaperone regulates substrate binding and release (43,44,96).

Possible role of the interaction between the helical lid of SBD and the interdomain linker in Hsp70 allosteric interdomain communication.

Our analysis of Hsp70 oligomerization has lead us to discover that the interaction between the helices B-E of the lid subdomain of SBD and the well-conserved leucine-rich motif of the linker connecting the NBD and the SBD results in the formation of oligomeric species. This interaction is highly specific, in the micromolar range, and competes with substrate binding for the optimal spatial orientation of the SBD helical lid subdomain, which makes us hypothesize that it could likely take place intramolecularly as well. In this way, chaperone oligomerization would be a consequence of an intramolecular interaction between the C-terminal lid of the SBD and the interdomain linker becoming intermolecular by a domain swapping-like mechanism, as another example of oligomer assembly by three-dimensional domain swapping (97,98). Three-dimensional domain swapping has been defined as the event by which oligomers are formed from stable monomers by exchanging domains, and it has been proposed to serve as a mechanism for functional interconversion between monomers and oligomers (97), as well as a potentially harmful process leading misfolding and aggregation (99). In the case of Hsp70 oligomerization, there is not a domain exchange between monomers, but an interaction exchange, where the C-terminal lid of a monomer instead of interacting with the interdomain linker of the same molecule, it establishes the interaction with the interdomain linker of another monomer. In this way, dimers and higher order oligomers can be formed.

Given the degree of specificity of this interaction, and the degree of conservation of the oligomerization process in the whole Hsp70 family, together with the relevance of the interdomain linker in the Hsp70 allosteric regulation between NBD and SBD recently proposed (33,34,49,95), we hypothesize that the interaction between the lid subdomain of SBD and the interdomain linker, either intramolecularly or intermolecularly through the formation of oligomers, could likely have an important role in the allosteric regulation of Hsp70.

The allosteric cycle of Hsp70 involves an alternation between the ATP-bound state, with low affinity and fast exchange rates for substrates (the open conformation), and the ADP-bound state, with

high affinity and low exchange rates for substrates (the close conformation). The substrate-bound structure of SBD from *E. coli* DnaK (42) revealed the helical C-terminal part is acting as a lid, encapsulating the substrate and stabilizing the close conformation (48). A more recent model proposes that SBD opens and closes periodically in both ADP and ATP states, and that the differences in substrate association and dissociation constant rates between the two states are a consequence of the different probabilities of opening the lid and forming the substrate binding site (41). The mechanism by which the lid subdomain of SBD changes from the close to the open conformation to allow substrate binding and dissociation remain unclear, and models involving a small hinge movement (42), and a pivot of the entire helical subdomain (43,44) have been proposed, although its regulatory mechanism remains unknown.

In the ADP-bound or nucleotide free state, the NBD and the interdomain linker are structurally and dynamically independent (82,90,100-103), while in the ATP-bound state, the interdomain linker adopts a β-strand conformation able to fit in a solvent-accessible hydrophobic cleft on the interface between subdomain IA and IIA in NBD (34). On the other hand, it has been shown that SBD undergoes a conformational rearrangement that is concerted with the conformational exchange in NBD (34,104), and that such allosteric communication from the NBD nucleotide-binding site to the SBD substrate-binding site is via the interdomain linker (34,105). Significant advances have been made in understanding the propagation of the allosteric signal from the nucleotide-binding site in NBD to the interdomain linker (34,53,95); however, much less it is known about the transmission of the nucleotide status of NBD from the interdomain linker to SBD or the exact mechanisms by which NBD senses the substrate status of SBD.

It has been observed that the presence of the conserved interdomain linker sequence motif VLLL in the isolated NBD of DnaK results in an about 10-fold stimulation of the ATPase activity, a value that is similar to that found in FL-Hsp70 upon substrate binding to SBD (95). This observation suggests that the accessibility of the interdomain linker to be able to interact with NBD and stimulate ATP hydrolysis depends on whether or not there is a substrate bound to SBD (i.e. the interdomain linker becomes accessible to interact with NBD upon substrate binding), and that the protein region that interacts with the linker to prevent it from docking into NBD is located in SBD.

It has been previously reported that upon ATP binding to NBD, there are major conformational changes in both NBD and SBD. NBD becomes stabilized in its closed conformation, and is able to interact with the interdomain linker (95). The region around the helical lid subdomain of SBD, on the contrary, becomes destabilized and several well-dispersed peaks in its NMR spectrum disappear, indicating that there is an exchange broadening in this region upon ATP binding (34). However, in the nucleotide-free and ADP-bound states, or upon substrate binding, the NBD and SBD are more independent and the linker, flexible, solvent-exposed and highly labile to proteolysis (82,106).

All these evidences, together with the recent proposed role of the interdomain linker as the protein switch motif involved in Hsp70 allosterism (34,53,95), and the interaction we have identified here between the helical lid subdomain of SBD and the interdomain linker, which is competing with substrate binding, make us propose that this interaction could have a regulatory role for the specific activation of ATP hydrolysis and substrate binding/release. In this way, by restricting the accessibility of the interdomain linker for its interaction with NBD, depending on whether or not there is a substrate bound to SBD, the chaperone could regulate the specific ATP hydrolysis only when it is needed. In the substrate-free and ATP-bound state of the chaperone, the lid subdomain of SBD might be interacting with the leucine-rich motif of the interdomain linker, in a similar way as we have observed in chaperone oligomerization, preventing it from interacting with NBD and therefore preventing unnecessary ATP hydrolysis. However, upon substrate binding, the lid subdomain of SBD is required to stabilize the complex and the interdomain linker becomes solvent-exposed and free to interact with the ATP-bound state of NBD (34,95), interaction that in turn stimulates ATP hydrolysis (34), which can be further accelerated by Hsp40/DnaJ (107). In the substrate-bound and ADP-bound state of the chaperone, the lid is in its closed conformation, with the polypeptide substrate tightly bound to the SBD, at the same time that NBD loses its interaction with the interdomain linker (34,95), which becomes accessible again to interact with the lid. To prevent inappropriate substrate release, dissociation of ADP catalyzed by one of several nucleotide-exchange factors (NEFs) is required in order to recycle the chaperone in its ATP-bound state, and then to maintain the substrate tightly bound. Once the substrate starts to fold, the hydrophobic residues would become buried in the folding core, and its affinity for

the substrate-binding pocket of SBD would be drastically reduced, resulting in the release of the substrate, now as a non-damaging species, and allowing the lid subdomain to be able to restore its interaction with the linker if needed to prevent hydrolysis of additional ATP molecules. Alternatively, the chaperone could regulate the process of substrate binding through the regulation of the accessibility of the interdomain linker to the SBD helical lid subdomain depending on the nucleotide state of NBD. In this way, the interaction between the lid subdomain of SBD and the interdomain linker could play a regulatory role for both substrate binding and NBD-SBD communication.

SUPPLEMENTARY IMFORMATION

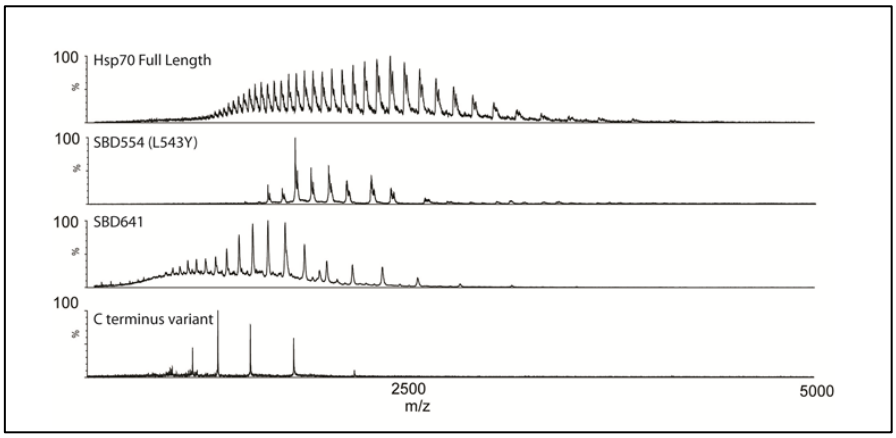

Fig. S1. Mass spectra of FL-Hsp70, SBD556, SBD641 and C-terminal variants under acidified denaturing conditions. The denaturing conditions were achieved using AG 501 –X8 (BioRad) beads (see Materials and Methods). All protein variants show a mass spectrum typical of a denatured protein, and, therefore, very different from that show at native conditions (see Fig. 2 in main text), except the C-terminal variant.

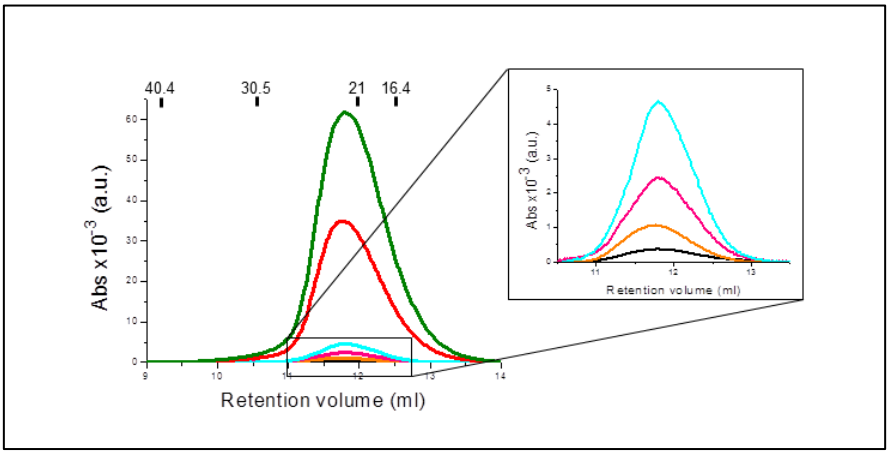

Fig. S2. Concentration dependency of the analytical SEC chromatograms of the C-term variant. The protein concentrations used were 1µM (black line), 2µM (orange line), 5 µM (pink line), 10 µM (cyan line), 70 µM (red line) and 125 µM (green line). Essentially the same elution profile was found for all the protein concentrations used. At the top the chromatogram the previously reported Stokes radii of the previously reported standard proteins used to calibrate the column is reported: conalbumin (40.4 Å), ovalbumin (30.5 Å), carbonic anhydrase (20.1 Å), ribonuclease A (16.4 Å).

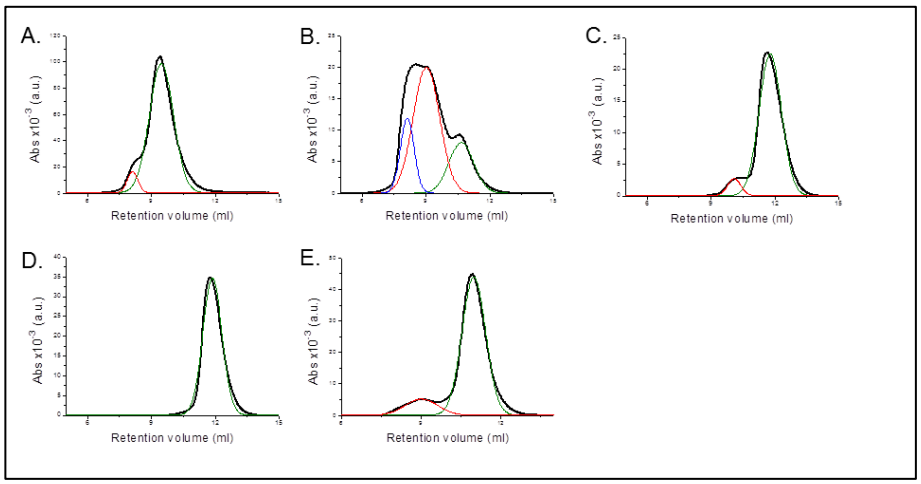

Fig. S3. Analytical SEC analysis of the different protein constructs at 70 µM of protein concentration. FL-Hsp70 (A), SBD641 (B), SBD556 (C) C-term (D) and ΔLSBD641 (E).

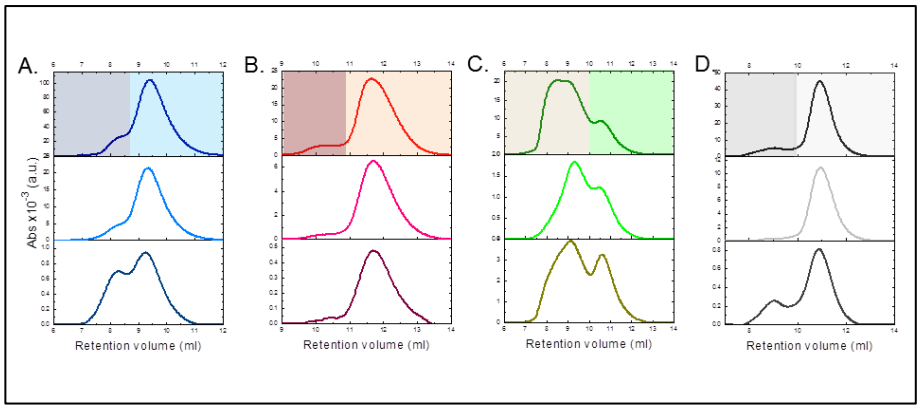

Fig. S4. Analytical SEC analysis of the different oligomeric fractions of FL-Hsp70 (A), SBD556 (B), SBD641 (C) and ΔLSBD641 (D). The monomeric and oligomeric peaks isolated by SEC (upper panel) were re-loaded into the same column for analysis (middle and bottom panels, respectively).

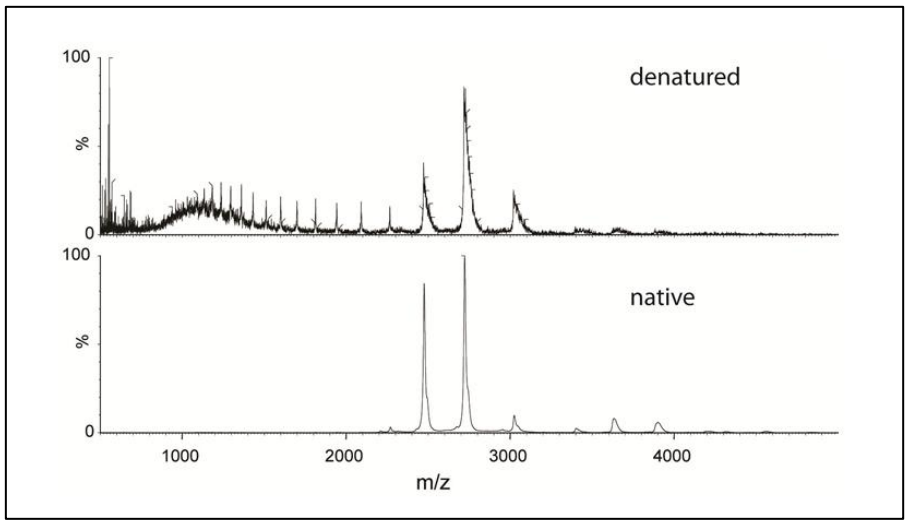

Fig. S5. Mass spectra of ΔLSBD641 under acidified denaturing (upper panel) and native conditions (bottom panel). This protein variant behaves similarly as SBD641 under denaturing conditions, while it shows a much less oligomerization propensity under native conditions, indicating that the leucine-rich sequence of the interdomain linker plays a key role in chaperone oligomerization.

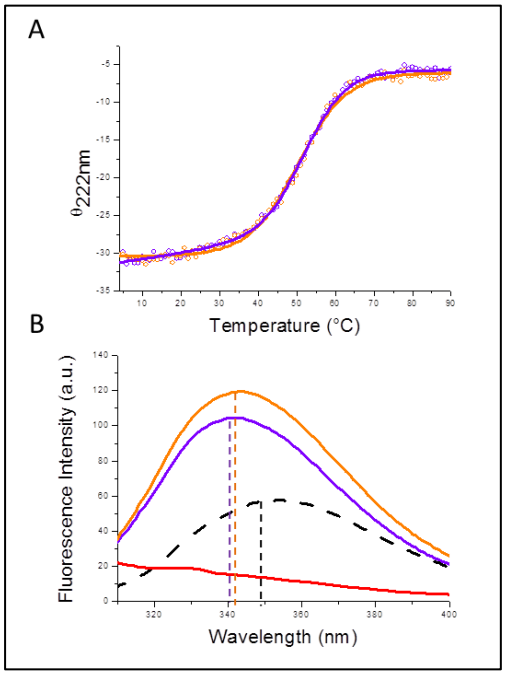

Fig. S6. (A) Thermal denaturation of C-term variant followed by CD signal at 222nm (violet points). The thermal denaturation curve of the refolded protein is shown as orange points. Both experimental curves were fitted to a two-state model; the best fit is shown as continuous lines. B) Tryptophan fluorescence emission spectra of the C-term variant in its native (violet continuous line), thermally denatured (black dashed line) and refolded conformations (orange continuous line). The spectrum of SBD556 is shown in red continuous line as a negative control.

1.5 References

1. **Dobson, C. M. (2003) *Nature* 426, 884-890**
2. **Radford, S. E., and Dobson, C. M. (1999) *Cell* 97, 291-298**
3. **Dobson, C. M., and Ellis, R. J. (1998) *EMBO J* 17, 5251-5254**
4. **Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000) *Trends Biochem Sci* 25, 331-339**
5. **Fersht, A. R. (2000) *Proc Natl Acad Sci U S A* 97, 14121-14126**
6. **Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2001) *Nature* 409, 641-645**
7. **Roder, H., and Colón, W. (1997) *Curr Opin Struct Biol* 7, 15-28**
8. **Sánchez, I. E., and Kiefhaber, T. (2003) *J Mol Biol* 325, 367-376**
9. **Khan, F., Chuang, J. I., Gianni, S., and Fersht, A. R. (2003) *J Mol Biol* 333, 169-186**
10. **Vendruscolo, M., Zurdo, J., MacPhee, C. E., and Dobson, C. M. (2003) *Philos Transact A Math Phys Eng Sci* 361, 1205-1222**
11. **Cheung, M. S., García, A. E., and Onuchic, J. N. (2002) *Proc Natl Acad Sci U S A* 99, 685-690**
12. **Bukau, B., and Horwich, A. L. (1998) *Cell* 92, 351-366**
13. **Hartl, F. U., and Hayer-Hartl, M. (2002) *Science* 295, 1852-1858**
14. **Ignatova, Z., and Gierasch, L. M. (2004) *Proc Natl Acad Sci U S A* 101, 523-528**
15. **Ellis, R. J., and Minton, A. P. (2006) *Biol Chem* 387, 485-497**
16. **Woolhead, C. A., McCormick, P. J., and Johnson, A. E. (2004) *Cell* 116, 725-736**
17. **Lu, J., and Deutsch, C. (2005) *Nat Struct Mol Biol* 12, 1123-1129**
18. **Elcock, A. H. (2006) *PLoS Comput Biol* 2, e98**
19. **Hartl, F. U., and Hayer-Hartl, M. (2009) *Nat Struct Mol Biol* 16, 574-581**
20. **Chang, H. C., Tang, Y. C., Hayer-Hartl, M., and Hartl, F. U. (2007) *Cell* 128, 212**
21. **Tang, Y. C., Chang, H. C., Hayer-Hartl, M., and Hartl, F. U. (2007) *Cell* 128, 412**
22. **Hunt, C., and Morimoto, R. I. (1985) *Proc Natl Acad Sci U S A* 82, 6455-6459**
23. **Lindquist, S., and Craig, E. A. (1988) *Annu Rev Genet* 22, 631-677**
24. **Gupta, R. S., and Singh, B. (1994) *Curr Biol* 4, 1104-1114**
25. **Daugaard, M., Rohde, M., and Jäättelä, M. (2007) *FEBS Lett* 581, 3702-3710**
26. **Mayer, M. P., Rüdiger, S., and Bukau, B. (2000) *Biol Chem* 381, 877-885**

27. **Pellecchia, M., Montgomery, D. L., Stevens, S. Y., Vander Kooi, C. W., Feng, H. P., Gierasch, L. M., and Zuiderweg, E. R. (2000) *Nat Struct Biol* 7, 298-303**

28. **Schlecht, R., Erbse, A. H., Bukau, B., and Mayer, M. P. (2011) *Nat Struct Mol Biol* 18, 345-351**

29. **Pierpaoli, E. V., Sandmeier, E., Baici, A., Schönfeld, H. J., Gisler, S., and Christen, P. (1997) *J Mol Biol* 269, 757-768**

30. **Ben-Zvi, A. P., and Goloubinoff, P. (2001) *J Struct Biol* 135, 84-93**

31. **Slepenkov, S. V., and Witt, S. N. (2002) *Mol Microbiol* 45, 1197-1206**

32. **Schmid, D., Baici, A., Gehring, H., and Christen, P. (1994) *Science* 263, 971-973**

33. **Vogel, M., Mayer, M. P., and Bukau, B. (2006) *J Biol Chem* 281, 38705-38711**

34. **Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M., and Gierasch, L. M. (2007) *Mol Cell* 26, 27-39**

35. **Lowe, J., Hand, N., and Mayer, R. J. (2005) *Methods Enzymol* 399, 86-119**

36. **Young, J. C., Barral, J. M., and Ulrich Hartl, F. (2003) *Trends Biochem Sci* 28, 541-547**

37. **Diamant, S., Ben-Zvi, A. P., Bukau, B., and Goloubinoff, P. (2000) *J Biol Chem* 275, 21107-21113**

38. **Ben-Zvi, A., De Los Rios, P., Dietler, G., and Goloubinoff, P. (2004) *J Biol Chem* 279, 37298-37303**

39. **Mayer, M. P., and Bukau, B. (2005) *Cell Mol Life Sci* 62, 670-684**

40. **Evans, C. G., Chang, L., and Gestwicki, J. E. (2010) *J Med Chem* 53, 4585-4602**

41. **Mayer, M. P., Schröder, H., Rüdiger, S., Paal, K., Laufen, T., and Bukau, B. (2000) *Nat Struct Biol* 7, 586-593**

42. **Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) *Science* 272, 1606-1614**

43. **Wang, H., Kurochkin, A. V., Pang, Y., Hu, W., Flynn, G. C., and Zuiderweg, E. R. (1998) *Biochemistry* 37, 7929-7940**

44. **Morshauser, R. C., Hu, W., Wang, H., Pang, Y., Flynn, G. C., and Zuiderweg, E. R. (1999) *J Mol Biol* 289, 1387-1403**

45. **Stevens, S. Y., Cai, S., Pellecchia, M., and Zuiderweg, E. R. (2003) *Protein Sci* 12, 2588-2596**

46. **Buczynski, G., Slepenkov, S. V., Sehorn, M. G., and Witt, S. N. (2001) *J Biol Chem* 276, 27231-27236**

47. **Slepenkov, S. V., and Witt, S. N. (2002) *Biochemistry* 41, 12224-12235**

48. Moro, F., Fernández-Sáiz, V., and Muga, A. (2004) *J Biol Chem* 279, 19600-19606
49. Han, W., and Christen, P. (2001) *FEBS Lett* 497, 55-58
50. Laufen, T., Mayer, M. P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J., and Bukau, B. (1999) *Proc Natl Acad Sci U S A* 96, 5452-5457
51. Mayer, M. P., Laufen, T., Paal, K., McCarty, J. S., and Bukau, B. (1999) *J Mol Biol* 289, 1131-1144
52. Liu, Q., and Hendrickson, W. A. (2007) *Cell* 131, 106-120
53. Kumar, D. P., Vorvis, C., Sarbeng, E. B., Cabra Ledesma, V. C., Willis, J. E., and Liu, Q. (2011) *J Mol Biol* 411, 1099-1113
54. Benaroudj, N., Triniolles, F., and Ladjimi, M. M. (1996) *J Biol Chem* 271, 18471-18476
55. Benaroudj, N., Fouchaq, B., and Ladjimi, M. M. (1997) *J Biol Chem* 272, 8744-8751
56. Fouchaq, B., Benaroudj, N., Ebel, C., and Ladjimi, M. M. (1999) *Eur J Biochem* 259, 379-384
57. Guidon, P. T., and Hightower, L. E. (1986) *Biochemistry* 25, 3231-3239
58. Heuser, J., and Steer, C. J. (1989) *J Cell Biol* 109, 1457-1466
59. Kim, D., Lee, Y. J., and Corry, P. M. (1992) *J Cell Physiol* 153, 353-361
60. King, C., Eisenberg, E., and Greene, L. (1995) *J Biol Chem* 270, 22535-22540
61. Palleros, D. R., Welch, W. J., and Fink, A. L. (1991) *Proc Natl Acad Sci U S A* 88, 5719-5723
62. Ohno, M., Kitabatake, N., and Tani, F. (2004) *FEBS Lett* 576, 381-386
63. Brown, C. R., Martin, R. L., Hansen, W. J., Beckmann, R. P., and Welch, W. J. (1993) *J Cell Biol* 120, 1101-1112
64. Blond-Elguindi, S., Fourie, A. M., Sambrook, J. F., and Gething, M. J. (1993) *J Biol Chem* 268, 12730-12735
65. Carlino, A., Toledo, H., Skaleris, D., DeLisio, R., Weissbach, H., and Brot, N. (1992) *Proc Natl Acad Sci U S A* 89, 2081-2085
66. Freiden, P. J., Gaut, J. R., and Hendershot, L. M. (1992) *EMBO J* 11, 63-70
67. Worrall, L. J., and Walkinshaw, M. D. (2007) *Biochem Biophys Res Commun* 357, 105-110
68. Schönfeld, H. J., Schmidt, D., Schröder, H., and Bukau, B. (1995) *J Biol Chem* 270, 2183-2189
69. Thompson, A. D., Bernard, S. M., Skiniotis, G., and Gestwicki, J. E. (2011) *Cell Stress Chaperones*

70. **Angelidis, C. E., Lazaridis, I., and Pagoulatos, G. N. (1999) *Eur J Biochem* 259, 505-512**

71. **Chou, C. C., Forouhar, F., Yeh, Y. H., Shr, H. L., Wang, C., and Hsiao, C. D. (2003) *J Biol Chem* 278, 30311-30316**

72. **Nemoto, T. K., Fukuma, Y., Itoh, H., Takagi, T., and Ono, T. (2006) *J Biochem* 139, 677-687**

73. **Roodveldt, C., Bertoncini, C. W., Andersson, A., van der Goot, A. T., Hsu, S. T., Fernández-Montesinos, R., de Jong, J., van Ham, T. J., Nollen, E. A., Pozo, D., Christodoulou, J., and Dobson, C. M. (2009) *EMBO J* 28, 3758-3770**

74. **Gill, S. C., and von Hippel, P. H. (1989) *Anal Biochem* 182, 319-326**

75. **Andrade, M. A., Chacón, P., Merelo, J. J., and Morán, F. (1993) *Protein Eng* 6, 383-390**

76. **Merelo, J. J., Andrade, M. A., Prieto, A., and Morán, F. (1994) *Neurocomputing* 6, 443-454**

77. **Hernández, H., and Robinson, C. V. (2007) *Nat Protoc* 2, 715-726**

78. **Chen, R. F. (1968) *Anal Biochem* 25, 412-416**

79. **Farr, C. D., and Witt, S. N. (1997) *Biochemistry* 36, 10793-10800**

80. **Amor-Mahjoub, M., Gomez-Vrielyunck, N., Suppini, J. P., Fouchaq, B., Benaroudj, N., and Ladjimi, M. (2007) *Protein Pept Lett* 14, 761-765**

81. **Amor-Mahjoub, M., Gomez-Vrielyunck, N., Suppini, J. P., Fouchaq, B., Benarouj, N., and Ladjimi, M. (2006) *Arch Inst Pasteur Tunis* 83, 53-62**

82. **Swain, J. F., Schulz, E. G., and Gierasch, L. M. (2006) *J Biol Chem* 281, 1605-1611**

83. **Osipiuk, J., Walsh, M. A., Freeman, B. C., Morimoto, R. I., and Joachimiak, A. (1999) *Acta Crystallogr D Biol Crystallogr* 55, 1105-1107**

84. **Konermann, L. (2007) *J Phys Chem B* 111, 6534-6543**

85. **Benesch, J. L., and Ruotolo, B. T. (2011) *Curr Opin Struct Biol* 21, 641-649**

86. **Amor-Mahjoub, M., Suppini, J. P., Gomez-Vrielyunck, N., and Ladjimi, M. (2006) *J Chromatogr B Analyt Technol Biomed Life Sci* 844, 328-334**

87. **Gay, N. J., Packman, L. C., Weldon, M. A., and Barna, J. C. (1991) *FEBS Lett* 291, 87-91**

88. **Kobe, B., and Deisenhofer, J. (1994) *Trends Biochem Sci* 19, 415-421**

89. **Kobe, B., and Kajava, A. V. (2001) *Curr Opin Struct Biol* 11, 725-732**

90. **Bertelsen, E. B., Chang, L., Gestwicki, J. E., and Zuiderweg, E. R. (2009) *Proc Natl Acad Sci U S A* 106, 8471-8476**

91. **Revington, M., Zhang, Y., Yip, G. N., Kurochkin, A. V., and Zuiderweg, E. R. (2005) *J Mol Biol* 349, 163-183**
92. **Rist, W., Graf, C., Bukau, B., and Mayer, M. P. (2006) *J Biol Chem* 281, 16493-16501**
93. **Slepenkov, S. V., Patchen, B., Peterson, K. M., and Witt, S. N. (2003) *Biochemistry* 42, 5867-5876**
94. **Moro, F., Fernández, V., and Muga, A. (2003) *FEBS Lett* 533, 119-123**
95. **Zhuravleva, A., and Gierasch, L. M. (2011) *Proc Natl Acad Sci U S A* 108, 6987-6992**
96. **Cupp-Vickery, J. R., Peterson, J. C., Ta, D. T., and Vickery, L. E. (2004) *J Mol Biol* 342, 1265-1278**
97. **Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) *Protein Sci* 4, 2455-2468**
98. **Rousseau, F., Schymkowitz, J. W., and Itzhaki, L. S. (2003) *Structure* 11, 243-251**
99. **Bennett, M. J., Sawaya, M. R., and Eisenberg, D. (2006) *Structure* 14, 811-824**
100. **Jiang, J., Prasad, K., Lafer, E. M., and Sousa, R. (2005) *Mol Cell* 20, 513-524**
101. **Jiang, J., Lafer, E. M., and Sousa, R. (2006) *Acta Crystallogr Sect F Struct Biol Cryst Commun* 62, 39-43**
102. **Jiang, J., Maes, E. G., Taylor, A. B., Wang, L., Hinck, A. P., Lafer, E. M., and Sousa, R. (2007) *Mol Cell* 28, 422-433**
103. **Chang, Y. W., Sun, Y. J., Wang, C., and Hsiao, C. D. (2008) *J Biol Chem* 283, 15502-15511**
104. **Moro, F., Fernández-Sáiz, V., and Muga, A. (2006) *Protein Sci* 15, 223-233**
105. **Taneva, S. G., Moro, F., Velázquez-Campoy, A., and Muga, A. (2010) *Biochemistry* 49, 1338-1345**
106. **Buchberger, A., Theyssen, H., Schröder, H., McCarty, J. S., Virgallita, G., Milkereit, P., Reinstein, J., and Bukau, B. (1995) *J Biol Chem* 270, 16903-16910**
107. **Fan, C. Y., Lee, S., and Cyr, D. M. (2003) *Cell Stress Chaperones* 8, 309-316**

2. Molecular basis of Hsp70 action in Parkinson's disease

2.1 Neurodegenerative diseases, a particular class of protein conformational diseases

Protein conformational diseases (PCD), also known as 'proteinopathies', are clinically and pathologically different disorders associated with the misfolding of specific proteins or peptides. PCD can be divided into two groups. In the group of diseases known as amyloidoses, misfolded proteins undergo a structural rearrangement that leads to the formation of toxic protein species affecting neuronal cells and other tissues (gain of toxic function). In the other group, a small genetic error, usually a single point mutation, leads to a misfolded conformation that either affects its function or makes it extremely susceptible to cellular proteases (loss of function).

Table 2.1 Principal neurogenerative diseases with the main identified protein involved in the pathogenesis (1).

Disease	Protein involved
Alzheimer's disease (AD)	Amyloid β peptide (Aβ); Tau
Amyotrophic Lateral Sclerosis (ALS)	Superoxide dismutase, TDP-43, FUS
Parkinson's disease	α-synuclein (AS)
Dementia with Lewy Bodies (DLB)	α-synuclein
Prion disease	Prion protein (PrP)
Tauopathies	Microtubule-associated protein tau (Tau protein)
Huntington's disease (HD)	Huntingtin (HTT)
Spinocerebellar ataxia (SCA)	Proteins with tandem glutamine expansions (i.e. ataxin-1, ataxin-3, etc.)

Neurodegenerative diseases (Table 2.1) are a particular class of amyloidoses, including Alzheimer disease (AD), Parkinson disease (PD), and the inheritable polyglutamine (PolyQ) diseases, that leads to the degeneration and death of specific neuronal types.

Although the proteins involved share few or no functional and structural similarities, the molecular mechanisms of the pathogenesis are essentially the same (Fig. 2.1). In the pathology, a particular protein misfolds, assuming an atypical, three-dimensional

conformation. Under these conditions condition, the protein aberrantly self-assembles and starts an aggregation process encompassing soluble oligomeric species, and eventually leads to the formation of amyloid aggregates.

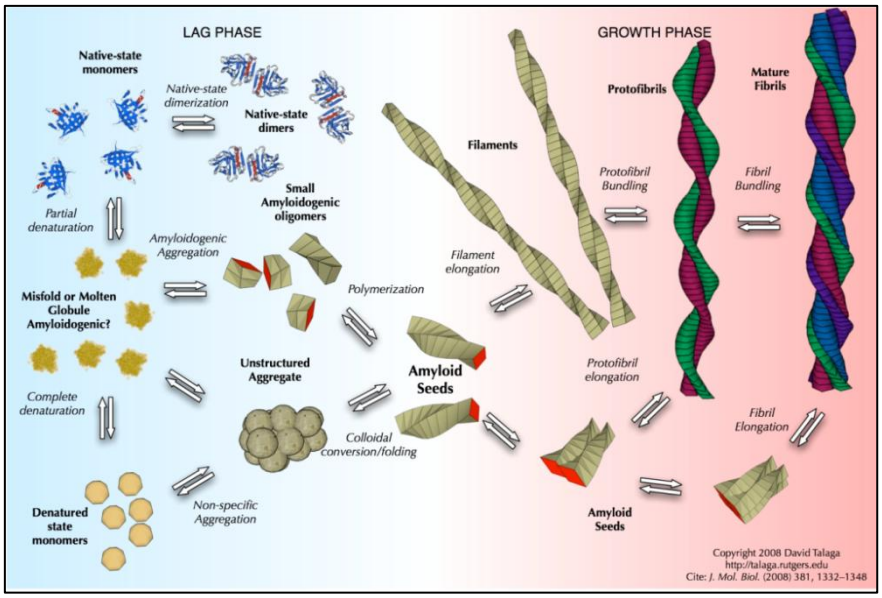

Fig. 2.1 Possible pathways of amyloid formation, starting from denatured monomeric protein. Normally a protein, co-translationally or just after its synthesis, acquires its correct native folding. If, for some reasons, the protein is not able to reach the native conformation, it can go through an aggregation process, leading to the formation of amyloid fibrils (2).

Amyloids are insoluble fibrous protein aggregates exhibiting cross-beta sheet quaternary structure. Similar to most cross-beta type structures, they are generally identified by apple-green birefringence when stained with Congo red and seen under polarized light.

The reason for amyloid association with disease is unclear. Recent theories point out pre-fibrillar intermediates, rather than mature amyloid fibrils, as the element triggering the diseases (3,4), leading to the consequent conclusion that amyloid fibrils could even be exploited as defence mechanism by the organism (5).

2.2 Parkinson's disease: clinical features and molecular mechanisms

Parkinson's disease (PD) currently represents one of the most serious and widespread neurodegenerative diseases, as about 1-2% of over 65 of the world population is affected. Symptomatology becomes evident in the second half-life with rising incidence with the age. PD is related to progressive loss in the ability of starting movements and it is clinically characterized by bradikinesia, resting tremor and muscle rigidity. In some cases, patients can even show non-motoric symptoms such as autonomic, cognitive and psychiatric dysfunctions. PD symptoms are caused by the progressive loss of dopaminergic neurons of the *substantia nigra* (SN), which is located in the area of the brain involved in movement planning and activation (Fig. 2.2A).

Post-mortem histological analysis of PD patient brain sections shows in the majority of the cases intra-neuronal proteinaceous aggregates, called Lewy bodies (LBs), from the name of Frederick Lewy, who first described them in 1912 (6). Proteomic analysis of LBs revealed the presence of amyloid fibrils composed of the protein α-synuclein (AS) (7,8), which normally localizes at the pre-synaptic level in all the central nervous system.

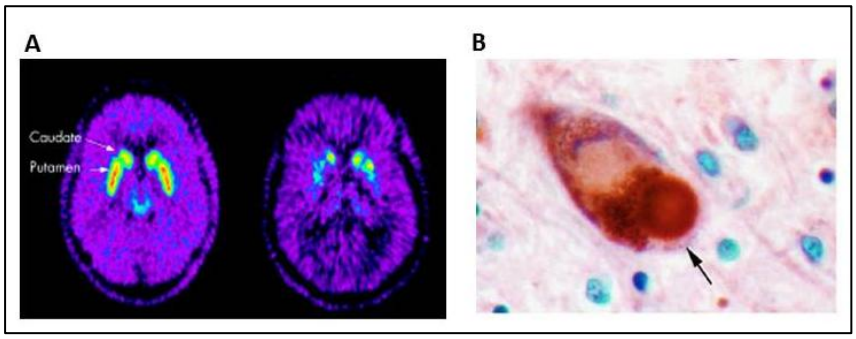

Fig. 2.2 (A) Comparison between the tomography of a control and PD patient shows a massive denegeneration of the substantia nigra dopaminergic neurons. (B) Histological sections of patient brains demonstrate the presence of intracellular LB inclusion.

So far, PD diagnosis is still an intricate issue: the symptomatology can be associated with a wealth of different causes, not just with AS amyloid deposition; moreover, only 10% of PD cases are characterized by genetic inheritance. The resting 90% is sporadic and often idiopathic. For this reason, PD main symptoms are grouped

by the term Parkinsonism, or "Parkinsonian syndrome". In contrast, all the pathologies related to AS accumulation, such as PD, dementia with Lewy bodies and multiple system atrophy are typically classified as Synucleinopathies. All PD patients show Parkinsonism, but not all of them present AS accumulation. In general, these latter cases consist of autosomal recessive forms of PD (Table 2.2). The most important protein involved in recessive forms of PD is the mammalian parkin (9,10). This protein is an E3 ubiquitin protein ligase that under normal conditions targets misfolded proteins to the ubiquitin proteasome pathway for degradation. Mutated parkin seems unable to tag proteins that have to be removed, which accumulate in neurons.

Table 2.2 Genes involved in inherited forms of PD (11).

Locus name	Chromosome	Gene name	Gene symbol	Inheritance	Protein
PARK1	4q21.3-q-22	SNCA, NACP, PD1	SNCA	AD	AS
PARK2	6q25.2-q27	Parkin; PDJ; AR-PD	PARK2	AR††	Parkin
PARK3	2p13 PD3	PD3 (AD with Lewy bodies)	PARK3	AD	?
PARK4	4q21	PD4 (AD with Lewy bodies)	PARK4**	AD	?**
PARK5	4q14	UCHL1	UCHL1	AD	UCHL1
PARK6	1p36-p35	PD6 (AR)	PARK6	AR††	Pink1
PARK7	1p36	PD7 (AR)	PARK7 or DJ-1 or DJ1	AR	DJ1
PARK8	1 2p11.2-q13.1	PD8 (AR)	LRRK2†	AD	LRRK2
PARK9	1p36	PD9 (AR) Rufor-Raked syndrome	PARK9	AR	?
PARK10	1p	PD10	PARK10	?	?
PARK11	2q36-q37 2q22-q23	PD11 (AD with early onset) NR4A2, TINUR, NOT, RNR1, HZF-3	PARK11 NR4A2	? AD?	? NR4A2

* Both the name and the symbol of the gene follow the nomenclature recommended by the Human Genome Organization (HUGO).
** The genome region associated with the disease presents a large deletion which includes the gene SNCA and several others; so far, it is still not clear if this gene would be the only one responsible for the phenotype.
† Symbol still not approved by HUGO.
†† Autosomal recessive inheritance in most cases; however, there are heterozygous and compound heterozygous patients.
PD: Parkinson's disease; AR: autosomal recessive; AD: autosomal dominant.

Another gene whose loss-of-function is associated with rare forms of autosomal recessive early-onset Parkinsonism is DJ-1 (12). DJ-1 is a homodimeric highly conserved protein of 19 kDa expressed in several tissues including brain (13). It may work as a scavenger of reactive oxygen species, as it has been proven that oxidative stress leads to Dj-1 auto-oxidation in order to remove H_2O_2 excess (14,15). Recently, some research groups suggested that oxidized DJ-1 acts as a chaperone that prevents early steps in the formation of AS aggregates (16,17). Some other rare forms of familial PD have been found to be correlated with mutations in PINK1 gene, a mitochondrial serine/threonine kinase. Overexpression of wild type PINK1 in some cell lines prevents mitochondrial cytochrome c release and subsequent apoptosis; this function is abolished in familial PD-linked PINK1 mutants (18).

Autosomal dominant PD is mainly linked to point mutations in AS gene, as A53T (19), A30P (20) and E46K (21), that correlate with the aggregation propensity of the protein. So far, the only known autosomal dominant PD form not related to AS mutations is linked to mutated variants of the leucin-rich repeat kinase 2 (LRRK2) (22,23), a modulator in the mitochondrial dependent cell death pathway. Recessive forms of PD are instead associated with several genes involved in the oxidative-stress response or in the ububiquitin proteasomal system.

The ethiopathogenesis of PD is not completely understood. However, in the cases of sporadic PD patients, cells undergo a series of toxic insults that lead to proteasomal impairment and increase of oxidative stress damage (24). Several genes and cellular pathways are implicated and single characters of neuronal death are becoming evident step by-step. Neuronal death cannot be prevented yet. Until the causes that trigger PD will not be understood, the development of drugs able to block nigral impairment and possibly restore motor functionality in patients will be very hard.

2.3 Alpha-synuclein

Alpha-synuclein (AS) is a 14 kDa protein consisting of 140 amino acids. It is expressed in all neurons of the central nervous system, even if predominantly in the neocortex, hippocampus, substantia nigra, thalamus, and cerebellum (25). AS mainly localises at the presynaptic level, where it can reach concentrations 70 to 140 μM. Although recent works supports the idea that AS might exist in cell as a folded α-helical tetramer (26-28), the most accepted view about its structure is that AS is an intrinsically disordered protein that acquires a helical conformation upon binding to membranes.

Although the function of AS is still unknown, several evidences pointed out a possible involvement in synaptic vesciculation. It has been actually observed that roughly 15% of AS is bound to the membrane in any moment and it has also been demonstrated that AS directly interacts with phospholipidic surfaces (29). In addition, recent evidence suggests that AS might act as a molecular chaperone in the formation of SNARE complexes (30). In fact, AS overexpression rescues lethality induced by CSPα knockout, which is a co-chaperone associated to synaptic vesicles and implicated in folding of SNARE proteins, leading to the conclusion that AS may act as an auxiliary chaperone preserving the function and integrity of the synapse (31). Finally, it has been suggested AS might play an important role in synaptic plasticity and learning: AS is specifically up-regulated in a discrete population of pre-synaptic terminals of the brain during the period of acquisition-related synaptic rearrangement (32,33). Nevertheless, AS knockout mice present only minor deficits in dopamine system development (34) as compared to wild type ones.

AS structure consists of three regions: an amphipathic N-terminal part, including the first sixty residues, an intermediate region called *non-amyloid component* (NAC) and an acidic and Pro-rich C-terminal region (Fig. 2.3A). The N-terminal region contains six imperfect repeats of the motif KTKEGV that is involved in the binding of detergents, micelles and liposomes. The NAC region, which corresponds to amino acids 61-95, is highly hydrophobic and it can acquire β-sheet structure as it is the leader of amyloid fibril formation (Fig. 2.3B). In particular, aa 71-82 was described to be the responsible of AS aggregation, as deletion mutants do not form β-sheet structured fibrils (35). The C-terminal tail of AS contains a large number of acidic amino acids and several prolines, so this region does not acquire structure under any condition. The region has been found

to bind metal ions such as Ca^{2+}, $Fe^{2+/3+}$, Al^{3+}, Co^{2+} and Mn^{2+} (36) and divalent metal binding has been shown to increase fibril formation rate. Moreover Ser129 is site of phosphorylation (37), whereby phosphorylated AS fibrilization rate is slower than wild type protein (38).

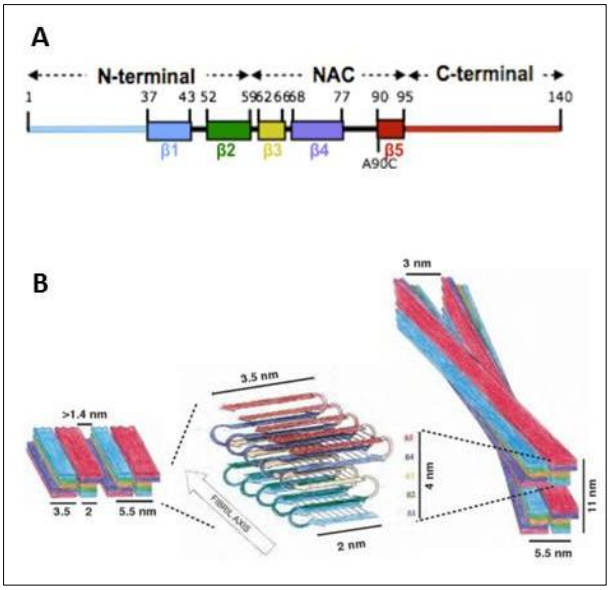

Fig. 2.3 (A) Linear organization of AS, highlighting the three main regions: N-terminal (residues 1-60), NAC (61-95) and C-terminal (95-140). (B) the five regions proposed to form the strands of the beta-sheet sandwich core of the fibrilar structure according to solid-state NMR (*19*): strand β1 comprising residues 37-43; strand β2, 52-59; strand β3, 62-66; strand β4, 68-77 and strand β5, 90-95 (39).

It is well reported in literature that AS is able to form fibrils *in vitro*, explaining the LB composition in Parkinson's patients. Nevertheless, the molecular details about the aggregation process are still unclear. Moreover, so far the protein species triggering the disease have not yet been clearly identified, although several evidences pointed to soluble oligomers as the toxic ones (40,41). The transition rate from monomer to oligomers and from oligomers to fibril is one of the most critical topics in the comprehension of the aggregation process of AS. Actually, AS fibril formation is a nucleation dependent process so the rate limiting step is the formation of a nucleus composed of a critical numbers of monomers, as it is energetically unstable (42,43). However, once this passage has been achieved, elongation or enlargement is a thermodynamically favoured process (Fig. 2.4).

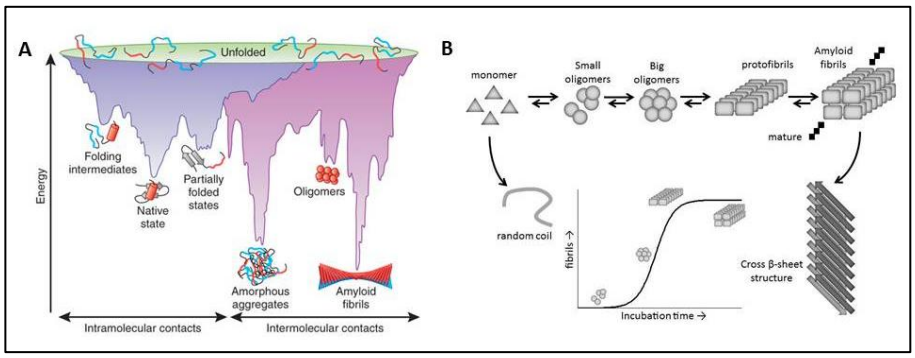

Fig. 2.4 (A) Energy landscape scheme of protein folding and aggregation. The purple surface shows the multitude of conformations 'funneling' to the native state via intra-molecular contacts and the pink area shows the conformations moving toward amorphous aggregates or amyloid fibrils via intermolecular contacts. Both parts of the energy surface overlap. Aggregate formation can occur from intermediates populated during de novo folding or by destabilization of the native state into partially folded states and is normally prevented by molecular chaperones. Cell-toxic oligomers may occur as off-pathway intermediates of amyloid fibril formation. (B) Scheme for amyloid fibril formation: monomeric protein undergoes a structural rearrangement that, toward soluble oligomeric species, leads to final mature fibril formation (44).

2.4 Hsp70 role in Parkinson's disease

In neurodegenerative diseases, amyloidogenic protein species are considered a potential therapeutic target and many studies have been focused on the neuroprotective role of Hsp70 chaperones. Actually, it has been demonstrated that amyloid aggregates contain, apart from the pathological protein, several important components of the protein *turn-over* machinery, such as molecular chaperones.

In particular, Hsp70 activity has been associated with several neurodegenerative pathologies (Table 2.3), including PD and polyQ diseases. First of all, strong indications for Hsp70 involvement in neurodegeneration came directly from medical studies on patients: for instance, post-mortem histological analyses on brain sections have shown in many cases the presence of Hsp70 chaperones in amyloid deposits (45). Moreover, protein expression profile of patient tissue often revealed highly perturbed levels of Hsp70 family members. Further demonstrations have been obtained from *in vivo* model systems and in particular it has been observed that over-expression of Hsp70 is related to a significant reduction of the pathological phenotype. Therefore, several studies have demonstrated that activation of the Hsp70 response might have a protective role against neurodegenerative conditions and this effect has been postulated to be principally related to its chaperone role. Hsp70 has been linked to multiple steps of the protein misfolding and aggregation pathway, including in preventing misfolding, blocking early stages of aggregation, and mediating the degradation of misfolded intermediates through coupling to the ubiquitin-proteasome system (Fig. 2.5).

The first hints about the role of Hsp70 chaperones in neurodegeneration have been obtained on polyQ diseases. Indeed, it has been widely demonstrated in several cellular models that the over-expression of polyQ proteins produces an increase in endogenous expression of Hsp70 family protein (46,47). Moreover, Hsp70 levels have been observed to decrease along with disease progression in HD mouse brain (48,49) and in the SCA3 *Drosophila* model (50).

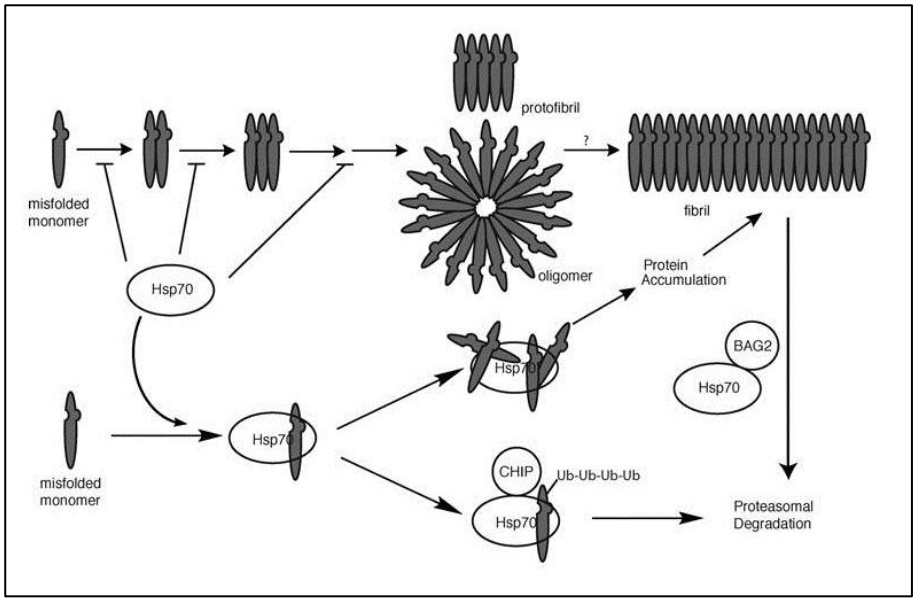

Fig. 2.5 Potential roles for Hsp70 in protein misfolding and aggregation. Hsp70 action has been associated to several steps in the protein aggregation pathway, such as monomer misfolding, early stages of aggregation and degradation of misfolded intermediates through the ubiquitin-proteasome system. The Hsp70 cochaperones BAG2 and CHIP have both been linked to clearance of misfolded proteins. In some systems (including yeast prions), Hsp70 activity is required for fibril formation. For simplicity, this schematic encompasses broad aspects of the misfolding pathway of amyloid β, polyQ, and tau, although important differences likely exist.

Following the discovery that Hsp70 in able to abrogate the neurotoxicity exerted by polyQ aggregates (51), it has been shown that Hsp70 can also prevent dopaminergic neuronal loss associated with AS in mouse and *Drosophila* PD models (52-54). Numerous studies that followed have reported that over-expression of Hsp70 is able to reduce AS aggregation and/or toxicity in various cellular models (54-59).

Unlike the many research works performed with PD cellular and animal models that have contributed to our understanding on Hsp70 function under physiological or pathological conditions, only a handful of studies have focused on the molecular mechanism and interactions that underlie the modulation of AS aggregation exerted by Hsp70. Similarly to what had been found *in vitro* for Hsp70 and the HD-related huntingtin protein (60,61), several studies with AS have shown that Hsp70 is able to suppress AS fibril assembly. In this case,

a variety of *in vitro* studies have shown that this efficient inhibition of amyloid assembly can occur in the absence of collaborating co-chaperones and in an ATP-independent manner (62-66), while promoting formation of small oligomeric species of moderate toxicity (66). However, even though Hsp70 is indeed able to control AS fibrillation and its associated toxicity in an ATP-independent manner, it does so more efficiently in the presence of ATP (54,66). Nevertheless, this ATP-dependent activity was found to require Hip co-chaperone activity. Indeed, it has recently been proved that Hsp70 in the presence of ATP is prone to coaggregate with aggregating AS, presumably consequently to the formation of a highly insoluble (ADP)Hsp70/AS complex (66). This co-aggregation can be prevented by the addition of the co-chaperone Hip (66), which had been found to be down-regulated in PD patients since the early stages of the disease (67). This finding indicates that Hip co-chaperone could be important not only in modulating the chaperone's molecular mechanism, but also in guarding the functionality and availability of Hsp70 under certain conditions.

One relevant question relates to the nature of the species along the aggregation pathway of AS that are specifically recognized and targeted by Hsp70. It is currently understood that Hsp70 inhibits fibril formation via interactions with soluble pre-fibrillar forms of AS rather than with mature amyloid fibrils (63,66), since so far Hsp70 has not been reported Hsp70 to disaggregate or alter the structural properties of mature AS fibrils. Although previous works did not report any interaction between Hsp70 and the monomeric form of AS, it has recently been shown by FRET and NMR analyses that, contrary to what was previously thought, Hsp70 can also interact with AS monomeric species (66). Nevertheless the molecular mechanism of binding and the role of this interaction in the inhibition of AS amyloid fibrils have not been clarified.

Although AS fibrillar assembly has been demonstrated to be inhibited by Hsp70, an active refolding process mediated by Hsp70 is unlikely (65). A hypothesis that summarizes many of the results of PD studies obtained in recent years predicts that Hsp70 solubilizes AS and promotes the degradation of its insoluble forms, via both chaperone-mediated autophagy and the proteasome (68).

Table 2.3 Role of Hsp70 chaperone in the main neurodegenerative diseases (69).

Disease	Hsp70 role
HD	Hsp70 colocalizes with polyQ aggregates. Hsp70 and Hsp40 prevent aggregation of purified HD exon. Overexpression of Hsp70 and Hsp40 reduces polyQ aggregation and cytotoxicity. Hsp70 inhibits polyQ-induced caspases 3 and 9 activation and M3/6 JNK phosphatase aggregation. Overexpression of Hsp70 does not significantly ameliorate the disease symptoms of R6/2 HD model mice. HSF-1 activating compounds reduce polyQ aggregation and rescue neurodegeneration in cultured cells and HD model mice. Hsp70 overexpression suppresses neurodegeneration and improves motor function in SCA1 mice.
SCA1	Hsp70 overexpression suppresses neurodegeneration and improves motor function in SCA1 mice.
SCA3	Hsp70 colocalizes with nuclear inclusions of ataxin-3. Overexpression of Hsp70 suppresses polyQ-mediated neuropathy in a Drosophila model of SCA3.
SCA7	Hdj2 and Hsp70 prevent mutant ataxin-7 aggregation in cultured cells but not in a mouse model.
SBMA	Hsp70 colocalizes with nuclear inclusions of polyQ expanded androgen receptor (AR). Hsp70 and Hsp40 increase the SDS solubility and proteasomal degradation of mutated AR in cultured cells. Overexpression Hsp70 ameliorates disease phenotypes in a SBMA model. Oral administration of GGA, an HSF1 inducer, ameliorates the SBMA phenotype in mouse models.
AD-Aβ	Hsp70 and Hsp40 block *in vitro* Aβ self-assembly. Hsp70 reduces steady state Aβ levels and Aβ-induced cytoxicity in cultured cells.
AD-tau	Hsp70 interacts with sites in tau important for aggregation. Hsp70-Bag2 captures and delivers insoluble and phosphorylated tau to the proteasome for ubiquitin-independent degradation. Hsp70-Bag1 associates with tau and inhibits proteasomal degradation.
PD	Lewy bodies contain Hsp40 and Hsp70. Overexpression of Hsp70 reduces R-synuclein aggregation and cytotoxicity. Hsp70 overexpression prevents dopaminergic neuronal loss in PD models.

2.5 Insights into the molecular mechanisms of Hsp70-mediated inhibition of alpha-synuclein amyloid formation

INTRODUCTION

Molecular chaperones have been recognised as key players in the avoidance of misfolding and amyloid fibril formation. In particular, recent evidences demonstrate that loss of chaperone heat-shock protein 70 kDa (Hsp70) activity is strictly related to the formation of intra-neuronal inclusions of alpha-synuclein (AS), one of the main hallmarks of Parkinson's disease (68).

Human Hsp70 is composed of two functional domains; the N-terminal nucleotide binding domain (NBD, 1-384), with ATPase activity, connected by a conserved intra-domain linker (385-397) to the C-terminal substrate-binding domain (SBD, 398-641). SBD is organized in other two sub-domains: the beta (398-507) and the lid sub-domain (LS, 508-641). With the aim of gaining insight into Hsp70 substrate recognition process of AS and the resulting capability of the chaperone to inhibit AS aggregation, we have characterised both protein interaction as well as its effect on amyloidogenesis using full-length Hsp70 and several truncated variants and by means of a wide range of biophysical techniques.

We have demonstrated that an helical lid sub-domain (LS) in the SBD is essential for monomeric AS binding, but not for the anti-aggregation activity of the chaperone, suggesting that Hsp70 is able to interact with pre-fibrillar oligomeric species formed during AS aggregation and that, then, the mechanism of binding for these species is different from that of the monomeric protein.

MATERIAL AND METODS

Protein expression and purification. Recombinant N-hexa-His-tagged Hsp70 (Hsp70 1A, gi:194248072) variants were expressed in and purified from *E.coli* BL21 (DE) gold strain (Stratagene) as previously described (66). Thrombin cleavage efficiency, estimated by mass spec analysis, was more than 99% and protein purity exceeded 95% as determined by SDS–PAGE and MS. Protein concentration was determined by absorbance measurement at 280 nm using theoretical extinction coefficients calculated with Expasy ProtParam (70).

Human *wild-type* AS (gi:80475099) was purified as early reported (71) as well as A90C variant (with addition of 1 mM DTT

to avoid cross-link reaction between mutagenized Cys). Protein purity exceeded 99% as determined by SDS–PAGE, and MS, and the A90C AS concentration was determined measuring absorbance at 275 nm using an extinction coefficient of 5600 M^{-1} cm^{-1}.

Labelling reaction. A90C AS was labelled with DANSYL-MTT (Toronto research) via the cysteine thiol moiety. The protein was incubated in the presence of 5 molar equivalent excess of the dye in PBS for 3 h at room temperature in the dark. Then the labeled protein was purified from the excess of free dye by a P10 desalting column with Sephadex G25 matrix (GE Healthcare), divided into aliquots, flash frozen and stored at -80 °C. Each aliquot was thawed immediately prior to use. The labelling efficiency was more than 70 % as estimated by mass spectrometry. The labeled protein concentration was estimated by absorbance measurement at 335 nm using the extinction coefficient of the free dye (4100 M^{-1} cm^{-1}) (72).

Determination of the affinity for monomeric AS by fluorescence titration. Fluorescence titration experiments were carried out by incubating 2 μM DANSYL-AS in 50 mM Tris pH 7.4, 5 mM $MgCl_2$ and 150 mM KCl for 30 min at 25°C in the presence of different concentrations of chaperone. Fluorescence emissions spectra from 400 to 630 nm were recorded as the average of 10 spectra, following an excitation at 330 nm (66); the increase in fluorescence intensity at the maximum emission was plotted as a function of chaperone concentration, and analysed assuming a one site binding model using the following equation:

$$F = \frac{Fmax \times [chaperone]}{Kd + [chaperone]} \qquad \text{Eq. 1}$$

where F is the fluorescence intensity observed at a given concentration of free Hsp70 in equilibrium (for practical reasons the total Hsp70 concentration is employed instead), F_{max} is the fluorescence at saturation and K_d is the apparent dissociation constant of the complex.

Aggregation conditions. 70 μM AS alone or together with 7 μM chaperones (i.e. 1:10 molar ratio) was incubated in Tris 30 mM pH 7.4, 110 mM KCl, 25 mM NaCl, 4 mM $MgCl_2$ (with 0.01% NaN_3 to prevent bacterial growth) under shaking at 200 rpm at 37 °C. Where

specified, samples contained 2 mM ATP and an ATP-buffer system, composed by 0.2 units/ml pyruvate kinase and 5 mM phosphoenol pyruvate. In order to evaluate the intrinsic stability of chaperone variants in the experimental conditions, aggregation experiments were also performed on 7 μM chaperones alone.

At specific incubation times, 10 μl aliquots were analyzed for soluble protein amount or Thioflavin T (ThT) binding. For ThT analysis, samples were incubated for 30 min in the presence of 20 μM ThT. Fluorescence emissions were recorded as the average of 3 spectra from 460 to 600 nm, following an excitation at 446 nm using a Cary Eclipse fluorescence spectrophotometer (Varian). Soluble protein fraction was isolated by centrifuging the samples at 16000 xg and analysing the supernatant by SDS-PAGE, using 4–12% Bis-Tris NuPAGE gels (Invitrogen) in MES buffer under reducing conditions. Protein amount was estimated by densitometry analysis using Image-J software (NIH).

Fluorescence and densitometry values were analyzed with the following equation generally used to fit fibrillization kinetics (36):

$$y = (A0 + mA0 * t) + \frac{(A+mA*t)}{1+\exp(-(t-t50\%)*\mathrm{Kagg})} \qquad \text{Eq. 2}$$

Where A_0 and A are the values at beginning and the end of the aggregation, m_{A0} and m_A are the slopes of the lag phase and the plateau, assuming for them a linear dependency of normalized fluorescence values with the incubation time, $t_{50\%}$ is the midpoint of aggregation and k_{agg} is the apparent aggregation rate constant. The lag time of aggregation, t_{lag}, was determined from k_{agg} and $t_{50\%}$ as follows:

$$\mathrm{tlag} = t50\% - \frac{2}{\mathrm{Kagg}} \qquad \text{Eq. 3}$$

ThT results are reported as relative fluorescence, normalizing absolute values with the value estimated for the plateau from the fitting; densitometry data are represented as insoluble protein fraction, normalizing them with the value corresponding to initial soluble protein amount and plotting reciprocal values.

Guanidinium thiocyanate (GuTCN) fiber dissolution assay. 600 μl fibril samples of AS alone or in the presence of Hsp70 variants were centrifuged for 15 min at 16,000 x g and resuspended in 70 ul of fresh

aggregation buffer (Tris 30 mM pH 7.4, 110 mM KCl, 25 mM NaCl, 4 mM $MgCl_2$). 10 ul aliquots were then incubated at different concentrations of guanidinium thiocyanate (GuTCN) for 1 h at 25°C and analyzed by SDS-PAGE.

Seeding experiments. First generation AS fibrils were obtained as previously described (see Material and methods *Aggregation conditions*). Fibrils were then centrifuged for 15 min at 16,000 xg and resuspended in the same volume of aggregation buffer in order to remove monomer in solution if any. 20 μM monomeric AS alone or with 2 μM chaperone variants (i.e. 1:0.1 AS:chaperone molar ratio) was then incubated in the presence of 1% (v/v) of pre-formed first generation fibrils and 20 μM ThT at 37 °C without shaking in a VICTOR TM X3 Multilabel Plate Reader (Perkin Elmer, MA, USA). In the case of FL-Hsp70, the nucleotide-free state has been also tested, adding 5 mM ATP to the mix as previously described (66). Emissions at 535 nm were recorded every 30 min, upon excitation at 445 nm. Plates were sealed to prevent evaporation.

Fibril interaction assay. 20 μM chaperone variants were incubated in the presence of different molar ratios of second generation AS fibrils at room temperature for 30 min, if not differently specified. For FL-Hsp70, both the nucleotide-free and -bound states have been tested, adding to the sample 2 mM ATP or ADP. At the end of the incubation, samples were centrifuged at 16,000 x g for 15 min and the supernatant was analyzed by SDS-PAGE for the amount of soluble chaperone.
Second generation fibrils were obtained incubating 70 μM monomeric AS in the presence of 2% (v/v) of preformed fibrils for 50 h at 37 °C in shaking condition at 200 rpm.

Transmission Electron Microscopy (TEM). TEM images of AS fibrils (obtained in absence of presence of chaperone variants) were obtained using a Philips CEM100 transmission electron microscope. The samples were applied on Formvar-carbon coated nickel grids and negatively stained with 2% (w/v) uranyl acetate. For the aliquots of early incubation times, the sample was applied directly onto the grids. For late incubation times, the sample was previously centrifuged for 10-15 min at 16,000 x g, and then the supernatant was applied to the grid and stained, while fibrilllar material was applied to the grid, washed with ddH_2O, and stained.

Fourier Transform Infra-Red Spectroscopy (FT-IR) on AS fibrils. AS fibrils formed as described below were centrifuged at 16000 xg, washed in 1 volume of 50 mM Tris pH 7.4,50 mM KCl, 5 mM $MgCl_2$ and finally resuspended in the same buffer at the concentration of around 1 mg/ml. 5 µl-samples were deposited on the crystal of a bio-ATR cell in a Bruker Equinox 55 system and let dry for 15 minutes. When the buffer had completely evaporated and the fibrils had formed, a dry film on the crystal, 256 scans (IRIS Aperture 7000 µm, resolution 4 cm^{-1}) were accumulated to generate a spectrum, which was then processed through atmospheric compensation and baseline correction. A spectrum recorded with an empty crystal was used for the atmospheric compensation each time. The processed spectrum between 1580 and 1720 cm^{-1} was used for the fitting and interpretation of the data. The fitting approach was based on the analysis of the second derivative and on a Fourier self deconvolution of the spectrum. By these means the peak maxima were identified. Once the significant frequencies were defined, a set of Gaussian or Lorentzian curves was fitted below the whole absorption spectrum using the OPUS software (Bruker). The curves generated were divided into secondary structure categories based on the frequency of their maxima absorption. By calculating the area below each curve, the percentage of structure was calculated for each contribution. The area of curves with centre between 1620 and 1640 cm^{-1}, and between 1680 and 1700 cm^{-1}, was assigned to β-sheet. The area of curves with centre between 1640 and 1660 cm^{-1} was assigned to random coil and alpha helix, and finally, curves with centre between 1660 and 1680 cm^{-1} were attributed to turns and loops. The side chain contributions, mainly relevant below 1610 cm^{-1} were not taken into account in the spectra analysis.

RESULT AND DISCUSSION

Design of Hsp70 constructs. Plenty of works show that Hsp70 chaperones are related to the onset of several neurodegenerative diseases, by affecting the aggregation pathway of the amyloidogenic proteins involved. In particular, it has been demonstrated that human Hsp70 is able to inhibit AS fibril formation *in vitro* (63,64) and that Hsp70 allosteric regulation, fundamental for many activities of the chaperone, also strongly influences its effect on AS aggregation (66). Despite all these evidences, a detailed description of the nature of AS aggregated species recognized by Hsp70 is still missing, as well as for the mechanisms by which the allosteric regulation influences the anti-

amyloidogenic effect. Therefore, in order to get insight into the features of Hsp70 target species and into the nucleotide-dependence of the anti-amyloidogenic activity of the chaperone, we studied the effect on AS aggregation of a battery of Hsp70 1A variants, including the full-length protein (FL-Hsp70) and two forms of SBD (Fig. 1A): the full length SBD (residues 385-641), called SBD641, and a truncated SBD variant that we named SBD556 (residues 385-556) and that lacks the C-terminal tract carrying the helices C, D and E, which have been demonstrated to be important for the stabilization of monomeric substrates inside the binding pocket of the β-subdomain (73). The SBD556 variant contains the L543Y mutation to prevent self-binding of L543 to the binding pocket, as previously reported (74). The truncation points were designed based on the published structure of homolog proteins (1dkx, 7hsc, 1ud0) so as not to disrupt the folding motifs (Fig. 1B). The structural integrity of all the variants, including full-length Hsp70 was studied by far-UV circular dichroism (CD) as we reported in the previous chapter (see § 1.4).

Role of the C-terminal tract and nucleotide regulation in Hsp70 binding to monomeric AS. First we decided to analyse the binding capability of Hsp70 variants for the monomeric form of AS, quantifying the constants of binding by fluorescence titrations. We employed a cysteine mutant of AS (A90C AS) labelled with the fluorophore dansyl, which increases its quantum yield in the hydrophobic environment of the binding pocket of the SBD, and we monitored dansyl fluorescence emissions as a function of the chaperone concentration, as we previously reported (66). The experimental data were fitted and the apparent constants of binding (K_d) were estimated assuming a one binding site model (see Material and Methods) (Fig. 2, Table 1).

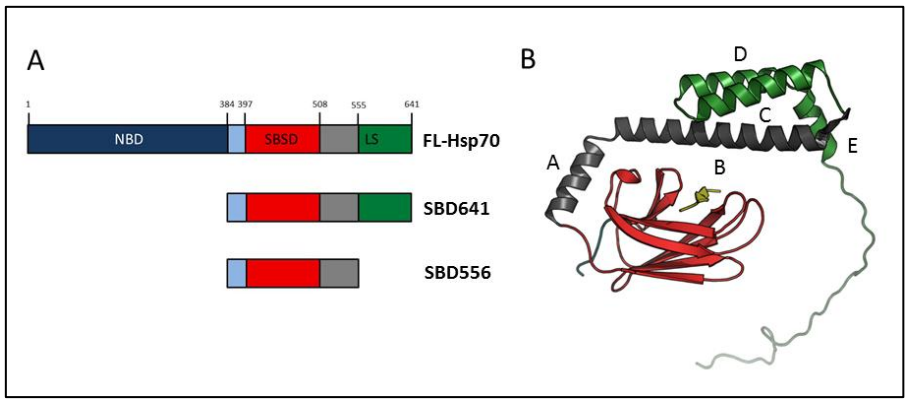

Fig. 1. (A) Scheme representing the engineered Hsp70 variants: the full length Hsp70 (FL-Hsp70), the full length SBD (SBD641) and a shorther SBD missing the last C-terminal 75 amino acids (SBD556). (B) Representation of the different truncations on the crystal structure of DnaK SBD (73).

FL-Hsp70 in the ADP-bound state showed twice the affinity than in the nucleotide-free state (11.5 ± 3 μM and 23.9 ± 3 μM respectively). The low values of these affinities and the higher K_d for the nucleotide-free complex agree with previous reports for the binding of Hsp70 chaperones with model substrates (75-83) and with another AS cysteine mutant (66). Interestingly, we observed that the K_d of SBD641 (19.3 ± 2.8 μM) was almost the same of nucleotide-free FL-Hsp70 (Table 1), supporting the idea that NBD and SBD act independently in the Hsp70 nucleotide-free state (84-87). On the contrary, SBD556 didn't show any binding for monomeric AS in the experimental conditions (Fig. 2). These results are strongly in agreement with the evidences that we obtained for the binding of the same Hsp70 variants with the model substrate NR (see § 1.4) and suggest that helices C-E are important for the binding of Hsp70 with misfolded proteins. Indeed, although it has been reported that this region is unable to make direct contacts with the substrate (according to DnaK:NR crystal structure (73)), it might play an important role in stabilizing the chaperone:substrate complex, similar to what has been already proposed for the helix B of DnaK SBD (73,88-90).

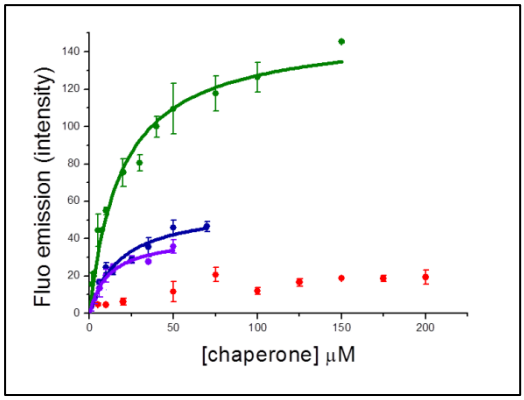

Fig. 2. Fluorescence titrations of dansyl-AS with SBD556 (red), SBD641 (green), nucleotide-free (blue) and ADP-bound (purple) FL-Hsp70.

Table 1. Apparent K_d estimation

	Kd (μM)
SBD556	n.d.*
SBD641	19.3 ± 2.8
FL-Hsp70	23.9 ± 3
FL-Hsp70/ADP	11.5± 2.8

*not detectable

Effect of Hsp70 variants on AS aggregation. In order to analyse how the effect of Hsp70 on the aggregation of AS is related to the affinity for the monomeric protein as well as the role of ATP in the process, we monitored AS amyloid formation by ThT assay and we estimated the soluble amount of AS during the aggregation, in the presence or the absence of 1:10 molar ratio of our Hsp70 variants (Fig. 3A, B).

ThT analysis showed that all our Hsp70 variants were able to affect AS aggregation (Fig. 3A, Table 2). In particular, nucleotide-free FL-Hsp70 demonstrated the stronger inhibitory effect: the apparent constant (K_{app}) of AS aggregation in the presence of nucleotide-free FL-Hsp70 is 20 fold lower (0.025 ± 0.006 h^{-1}) with respect to the aggregation of AS alone (0.58 ± 0.09 h^{-1}) (Table 2). Interestingly, the K_{app} obtained in the presence of nucleotide-bound FL-Hsp70 is just five times lower (0.12 ± 0.01 h^{-1}), suggesting that the nucleotide regulation might interfere with the anti-amyloidogenic action of Hsp70. Further, a clear effect was also detected for SBD641 and surprisingly for SBD556, which does not show any capability to bind monomeric AS (K_{app} of 0.08 ± 0.006 h^{-1} and 0.05 ± 0.004 h^{-1} respectively). Similar results have been obtained by monitoring

soluble AS quantity during aggregation experiments (Fig. 3B, Table 2).

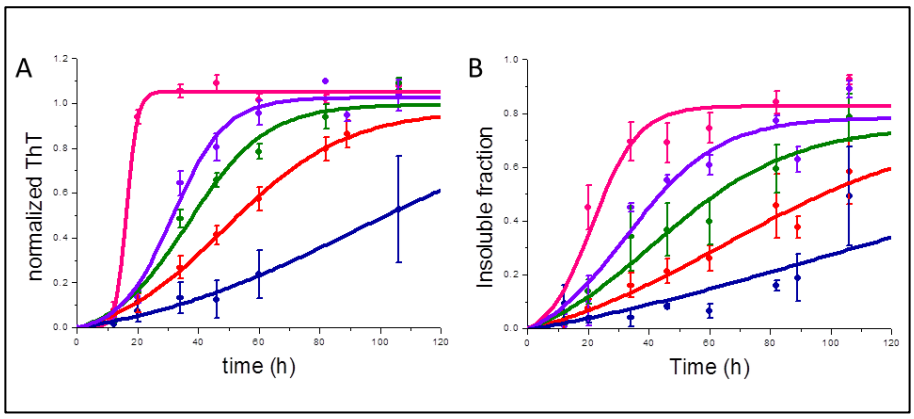

Fig. 3. Normalized ThT fluorescence intensities (A) and normalized levels of insoluble AS (B) of the aggregation of AS alone (pink), AS in the presence of SBD556 (red), SBD641 (green), nucleotide-free (blue) and ADP-bound (purple) FL-Hsp70.

Table 2. K_{app} (h^{-1}) estimation from ThT and soluble fraction analysis

	ThT	**Sol. frac.**
AS	0.58 ± 0.09	0.13 ± 0.02
AS/SBD556	0.05 ± 0.004	0.028 ± 0.006
AS/SBD641	0.08 ± 0.006	0.044± 0.006
AS/FL-Hsp70	0.025 ± 0.006	0.015± 0.006
AS/FL-Hsp70/ATP	0.12 ± 0.01	0.065 ± 0.008

Chaperone variant stability under the aggregation conditions. In order to explain the different effects exerted by Hsp70 constructs on AS aggregation, we decided to first characterize their solubility, as indication of the structural stability, in the experimental conditions used for AS aggregation. Therefore, we analyzed the variation of soluble protein fraction of each of them in the presence or absence of AS (Fig. S1).

In the absence of AS, SBD556 was the variant with the highest instability in the experimental conditions, as it was almost all precipitated after 50 h of incubation (Fig. S1A). On the contrary, SBD641 showed a similar stability with respect to the nucleotide-free full length chaperone (Fig. S1B, C). The truncation of the C-terminal

tract in SBD556 does not compromise protein conformation as assessed by CD spectra. Nevertheless this deletion produces a shorter and therefore less thermodynamically stable SBD, as we extensively discussed in the previous chapter (see § 1.4). This fact can explain the instability of SBD556 in the aggregation conditions with respect to SBD641 and FL-Hsp70. Further, for the full length Hsp70, we observed that the addition of ATP clearly increased the stability of the protein, which remains 50% soluble even after 110 h of incubation (Fig. S1D). This fact strongly supports the idea that the NBD of human Hsp70 is prone to aggregate and that the interaction with nucleotide leads to its stabilization.

We noticed that the addition of AS exerts two different effects on the stability of our chaperone variants, depending on the presence of the nucleotide regulation. Indeed, we observed that, for the truncated SBDs and for nucleotide-free FL-Hsp70, the presence of AS correlates with the increase of the soluble amount of chaperones along the incubation time (Fig. 1), indicating that the binding of AS increases their structural stability. Moreover, it was evident a good positive correlation between the soluble fraction of AS and of the chaperones (Fig. 4). In fact, the levels of AS in solution have been observed to decay with respect to those of the chaperone variants with a kinetic that can be well described with an exponential function (Fig. 4). Particularly, we observed that the amount of AS in solution decreases by just the 20%, as long as the soluble chaperone remains above 40%. Below this percentage, the relative amount of soluble AS drastically drops. Further, the soluble amounts of all Hsp70 variants show in these conditions almost the same exponential correlation with respect to soluble AS. These evidences suggest that all of them can potentially exert the same inhibitory effect on AS aggregation, if corrected for their structural stability in the experimental conditions used for the aggregation. This leads to the conclusion that the two SBD variants conserve almost the same capability to inhibit AS fibrillation as the full length chaperone, suggesting that the NBD and so the nucleotide regulation of Hsp70 is not fundamental for its anti-aggregation activity (in the case of AS aggregation). Moreover, the fact that even SBD556 maintains the inhibitory proprieties of FL-Hsp70 and that the correlation between soluble fractions is exponential suggests that Hsp70 inhibitory action is not fulfilled by binding monomeric AS, but via the interaction with pre-fibrillar oligomeric species formed along AS fibrillation process. Interestingly, we observed, that in the presence of ATP, FL-Hsp70 acted in a

different manner (Fig. 4). In this condition, the correlation between AS and chaperone soluble fractions is well approximated by a linear regression. Further, AS in this case turns 50% insoluble although Hsp70 is still 80% in solution. It has recently been suggested that AS, when Hsp70 is in the presence of ATP and so it is under the nucleotide- regulatory cycle, is able to trigger the co-aggregation of the chaperone into fibrils (66).

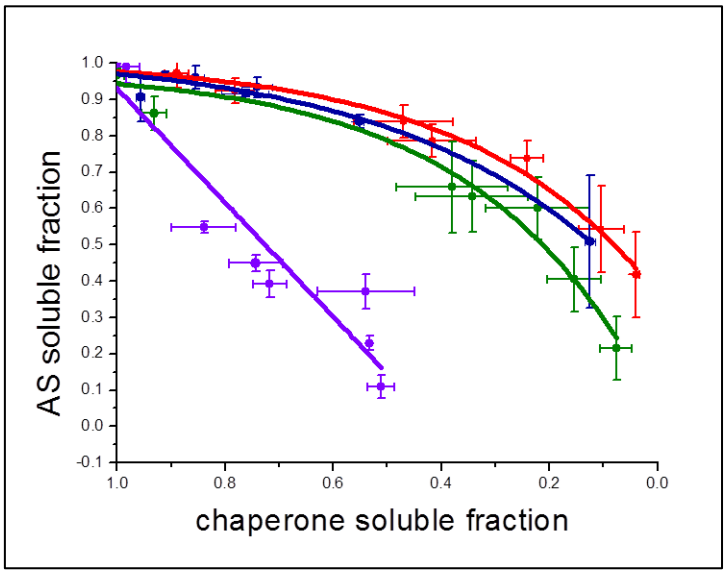

Fig. 4. Correlation between the levels of soluble AS and soluble SBD556 (red), SBD641 (green), nucleotide-free- (blue) and bound- (violet) FL-Hsp70.

In order to confirm this hypothesis, the amount of chaperone variants and AS in the fibrils at the end of the aggregation (i.e. after 110 h of incubation at 37°C) has been analyzed comparing guanidinium thiocyanate (GuTCN) dissolution profiles of the final aggregates (Fig. 5A). Interestingly, we were able to detect the presence of chaperone variants in the insoluble material at the end of the incubation for all the variants. Further, our results obtained by GuTCN dissolution experiment show that SBD556, SBD641, FL-Hsp70 in the nucleotide-free state are totally dissolved just by the SDS present in the gel used for the electrophoresis, suggesting that during aggregation experiments these chaperones were aggregating by themselves without any synergistic action by the aggregation process of AS. In contrast, FL-Hsp70 aggregates formed in the presence of ATP follow exactly the same dissolution profile as AS fibrils, strongly

indicating that under these conditions the chaperones co-aggregate with AS. Therefore, to assess if this was due to a specific and stable interaction of chaperones with mature fibrils, we incubated 20 μM chaperone in the presence of increasing amount of AS fibrils and monitored the amount of soluble AS by SDS-PAGE analysis (Fig. 5B). Surprisingly, we did not observe any changes in soluble fraction of AS even after 24 h of incubation, suggesting that the presence of the chaperones in the final aggregates was mainly a consequence of their aggregation. These evidences well explain the apparent contradictory behavior between the nucleotide-free and bound FL-Hsp70. Actually, the lower inhibitory activity shown by the chaperone in the presence of ATP could be explained by a progressive loss of function due to the sequestering action of AS during its aggregation.

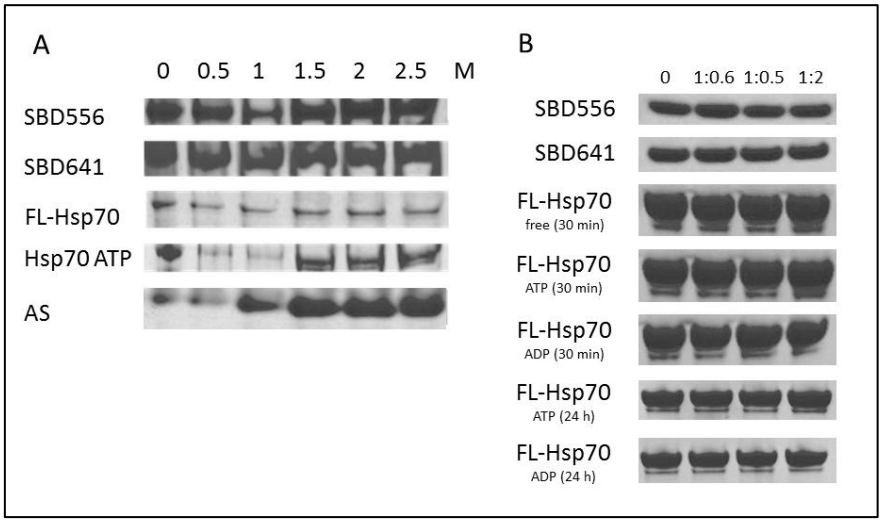

Fig. 5. (A) GuTCN dissolution profiles of final AS aggregates formed in the absence or in the presence of the different chaperone variants (B) Fibril interaction assay for SBD556, SBD641 and FL-Hsp70 in nucleotide-free, ADP-, ATP-bound state.

Hsp70 inhibits AS aggregation by interfering on the nucleation with a nucleotide-independent mechanism. We have so far demonstrated that the deletion of the C-terminal tract of the SBD abolishes the binding of Hsp70 to the monomeric form of AS, although it does not affect the anti-aggregation activity of the chaperone. Indeed, we observed that SBD556 exerts almost the same inhibition on AS aggregation as the full length SDB and the full length chaperone, based on the structural stability of the variants under the experimental conditions. This fact

strongly supports the idea that Hsp70 inhibitory effect might be independent from the binding to the monomeric form of AS, in agreement with previous evidences (63,64). In order to elucidate the mechanism by which Hsp70 inhibits AS aggregation without interacting with its monomeric form, we decided to investigate the effect of our chaperone variants on the elongation step of AS aggregation. Therefore, we performed seeding experiments with or without Hsp70 variants and, for FL-Hsp70, we tested both nucleotide free- and bound- state (Fig. 6). When AS was seeded in the presence of any of our chaperone variants, there was a total inhibition of the fibril growth in comparison with the control sample (i.e. AS alone). This fact leads to the consideration that our chaperones may act on or prior to nucleation, similarly to what observed for the flavonoid baicalein (91). Interestingly nucleotide-bound FL-Hsp70 did not lose the inhibiting effect during the seeding experiment, in contrast to what observed for normal aggregation. This can be explained by the fact that the co-aggregation of Hsp70 with AS may be promoted by the binding of the chaperone to different (and earlier) AS aggregated species with respect to those formed during seeding aggregation. Moreover the fact that all the variants, including nucleotide-bound Hsp70, had the same effect on fibril growth strongly indicates that this inhibition on fibril elongation must be nucleotide-independent.

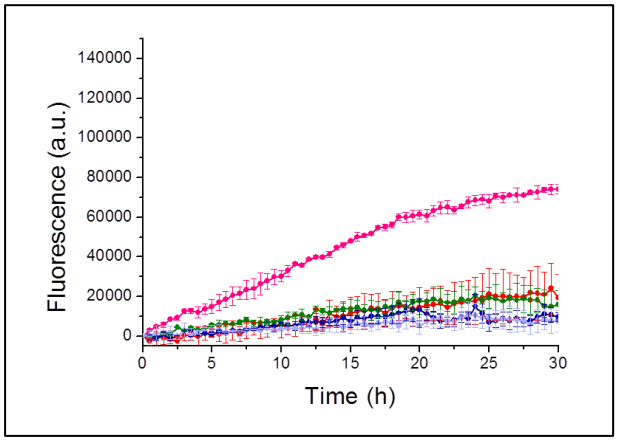

Fig. 6. Analysis on AS fibril elongation in the presence of chaperones variants. AS alone (pink), SBD556 (red), SBD641 (green) and FL-Hsp70 in the nucleotide-free state (blue) and in the presence of ATP (violet).

Effect of Hsp70 variants on fibril maturation. To obtain a detailed description of the effect of Hsp70 on the aggregation of AS, we performed a morphological and biophysical characterization of the amyloid fibrils formed in the absence or the presence of chaperone variants. First, we performed Infra-red (IR) spectroscopy analysis, coupled with TEM morphological characterization, on insoluble protein material obtained after 20, 45 and 110 h of incubation of AS alone (Fig. 7). Secondary structure quantification from FT-IR spectra of the insoluble aggregates (Fig. 7A) clearly shows that between 20 and 45 h AS insoluble aggregates undergo an important structural rearrangement that brings β-aggregate relative fraction (main peak at 1628 cm^{-1}) to be almost identical to the one of mature fibrils. Furthermore, the same holds true for random coil component (main peak at 1660 cm^{-1}), which reaches the same relative percentage with respect to mature fibrils already after 45 h. Interestingly, AS mature fibril (i.e. after 110 h of incubation), compared to the insoluble fractions at 20 and 45 h, show a β-aggregate peak with a pronounced shoulder at 1614 cm^{-1} (Fig. S2). This peak increases during aggregation of AS, reaching the 40% of the total secondary structures at the end of the incubation (Fig. 7B). Such further reorganization of fibrils can be interpreted as a quaternary structural change that leads to a more packed conformation of the fibrils in the final stage of their maturation. Indeed, the appearance of this shoulder is coupled with a right shift of the β-aggregate peak of 4 cm^{-1} (Fig. S2) between 45 and 110 h that is a strong indication of more packed aggregated structures.

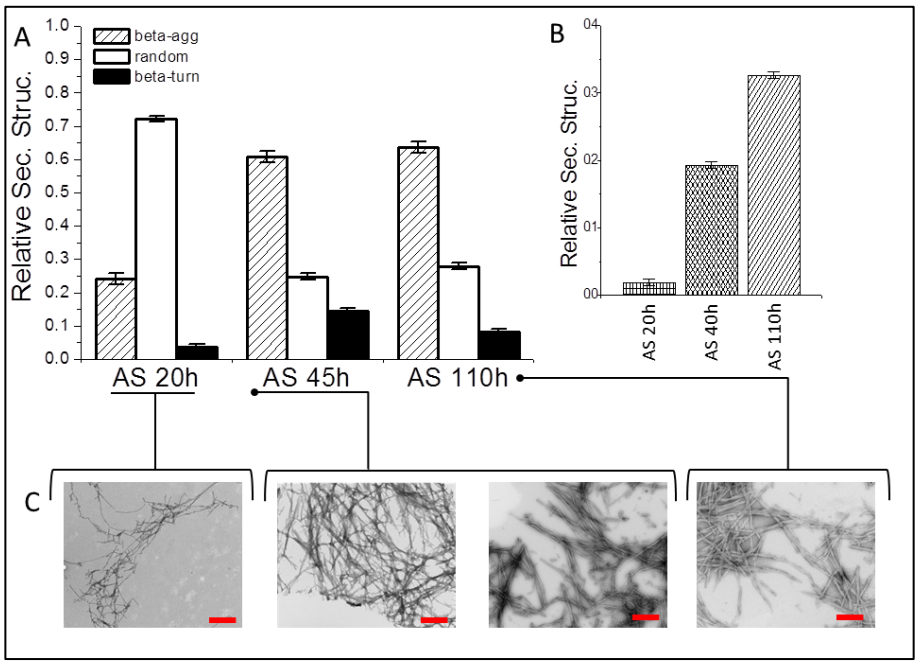

Fig. 7. (A) Secondary structure estimation from FT-IR spectra deconvolution of AS insoluble aggregates at 20, 45 and 110 h of incubation. The histogram shows the relative contribution of β-aggregate (dash), random coil (white) and β-turn (black). (B) Relative contribution of the β-aggregate component associated to 1614 cm^{-1} peak. (C) TEM analysis on the same AS aggregates.

FT-IR data are strongly supported by TEM morphological analysis we performed in parallel on the same aggregates (Fig. 7C). Indeed, just few short pre-fibrillar species with an average diameter of 14 nm have been detected after 20 h of incubation, in agreement with the low relative percentage of β inter-molecular structures determined by FT-IR; the aggregated material at 45 h is characterized by a high heterogeneity, but it is possible to detect longer fibrillar-like aggregates with an average diameter of 30 -40 nm. In some cases these species are almost indistinguishable from the mature fibrils obtained at the end of the aggregation, in keeping with IR spectra deconvolution. At 110 h, fibrils have a diameter of almost 40 nm and they often assembly into twisted (and more compact) structures that potentially may represent the final stage of maturation discussed above. Indeed, it is interesting to notice that the shoulder at 1614 cm^{-1} and the shift of the β-aggregate peak at 110 h observed in IR correlate with the appearance of these twisted fibrils.

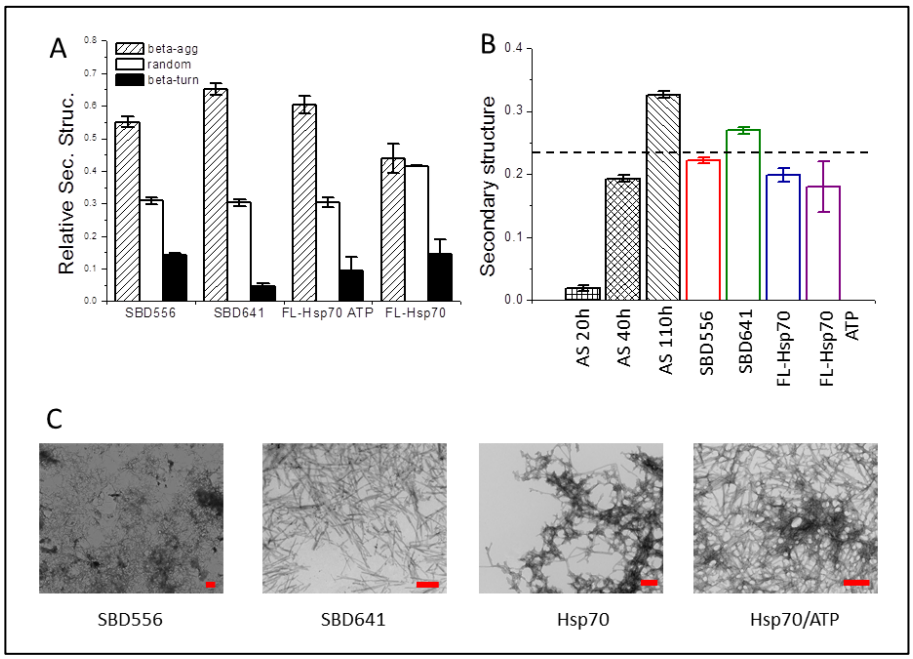

Fig. 8. (A) Secondary structure estimation from FT-IR spectra deconvolution of AS insoluble aggregates formed after 110 h of incubation in the absence or in the presence of the different chaperone variants. The histogram shows the relative contribution of β-aggregate (dash), random coil (white) and β-turn (black). (B) Relative contribution of 1614 cm^{-1} peak of AS aggregates formed in the presence of SBD556 (red), SBD641 (green), FL-Hsp70 (blue) and FL-Hsp70/ATP (purple). (C) TEM analysis on the same AS aggregates analysed by FT-IR.

To get insight into the effect of Hsp70 on the formation and maturation of AS fibrils, we performed the same investigations described above in the presence of our chaperone variants. The FT-IR analysis of the final aggregates (i.e. after 110 h) formed under the action of SBD556, SBD641 and nucleotide-bound FL-Hsp70 showed almost the same percentages of secondary structures as for AS alone, indicating that they have almost the same main structural features of normal fibrils (Fig. 8B). On the contrary, nucleotide-free FL-Hsp70 has been observed to strongly affect the structure of AS aggregates, as in this case the relative fraction of the random component increases up to 43% with respect to the 28% for the aggregates in normal conditions. Interestingly, in the presence of all the chaperone variants, the β-aggregate peak does not show any evident shoulder at 1614 cm^{-1} as observed for normal AS mature fibrils (Fig. S4). Further, in the case of the nucleotide-free FL-Hsp70, the position of the main β-

aggregate peak is the same observed for AS alone at 20 and 45 h (1628 cm^{-1}) (Fig. S3C). Therefore, AS aggregates in the presence of the full length chaperone appear to be strongly modified acquiring structural features resembling pre-fibrillar-like species observed after 20-45 h of AS normal aggregation. Interestingly, although the TEM analysis does not reveal any major morphological difference from AS aggregates formed under normal conditions (i.e. in the absence of chaperones), in the presence of the chaperones, no twisted fibril structure has been observed (Fig. 8A). This strongly confirms the idea that all our chaperone variants (even if deprived of the NBD or of the C-terminal tract of the lid subdomain) are able to affect the maturation grade of AS fibrils, particularly the final stages, which consist in a further packaging of the fibrils in twisted structures. Moreover, in the absence of nucleotides FL-Hsp70 produces even a stronger effect on fibril morphology that is likely to stop their maturation at the pre-fibrillar stage shown between 20 and 45 h.

Molecular mechanisms of Hsp70-mediated inhibition of AS aggregation and role of the nucleotide regulation. An increasing amount of studies indicate that molecular chaperones, in particular the Hsp70 family members, have an important influence on the aggregation of amyloidogenic proteins (92). In our work, we intended to analyse the effect of human Hsp70 on the aggregation process of AS, which is the main responsible of Parkinson's disease.

Currently, many theories about the mechanisms of Hsp70 inhibition of AS aggregation have been proposed. In particular, there is a heated debate regarding the nature of the misfolded species recognized by Hsp70 and about the role that the nucleotide regulation may play for this activity. Indeed, recent studies have shown that Hsp70 is able to inhibit AS aggregation and the associated toxicity by interacting preferentially with oligomeric species and even in the absence of ATP (63,64). Nevertheless, it has been also observed that physiological concentrations of ATP can potentially influence this action: under these conditions, Hsp70 acts more efficiently than in the nucleotide-free state, but only when the co-chaperone Hip is present. In fact, it has been observed that just the presence of ATP produces a specific depletion of Hsp70 as a consequence of the tendency of the chaperone in the ADP-bound state to co-aggregate with AS. Hip is required for Hsp70 activity as it is able to stabilize the chaperone and to prevent this co-aggregation (66).

Our *in vitro* data lead to important considerations, which may constitute the basis for the effect of Hsp70 against AS toxicity *in vivo*. First, we observed that the deletion of the NBD of Hsp70 does not compromise the binding for the monomeric form or the activity on the aggregation process of AS (Fig. 2, Fig. 3). In fact, the SBD641 variant, consisting of the full length SBD, showed almost the same affinity for the monomeric AS and the same anti-aggregation activity as the full length chaperone. Although the capability of Hsp70 to inhibit AS aggregation in a nucleotide-independent manner has been already documented (63-66,93), here we demonstrate for the first time that this activity is exactly the same shown by the chaperone under nucleotide regulation. Indeed, the different effects on AS aggregation rate (Fig. 3, Table 2) and on the structure of the insoluble aggregates (Fig. 8) can be easily explained by the different structural stability of the chaperone variants in the experimental conditions (Fig. 4). These evidences confirm the fact that in general the two Hsp70 domains act independently when the chaperone is in the nucleotide-free conformation, as already demonstrated for the binding of the chaperone with monomeric model peptides (87).

Further indications on the mechanism of Hsp70 inhibition on AS amyloid aggregation can be extrapolated from experiments using the SBD556 variant. This construct, which is deleted of the C-terminal tract of the lid subdomain, lacks of the capability to bind monomeric AS (Fig. 2). This strongly agrees with the evidences that we have obtained for the binding of the same Hsp70 variant to the model substrate NR (see § 1.4) and that suggest that helices C-E are important for the binding of Hsp70 with misfolded proteins. Indeed, although this region has been reported to not direct contact with the substrate (according to DnaK:NR crystal structure (73)), it might play an important role in stabilizing the chaperone:substrate complex, similar to what already proposed for the helix B of DnaK SBD (73,88-90). Although SBD556 lacks of the capability to bind monomeric AS, it still maintains the same inhibitory effect on AS aggregation as the full length chaperone. This demonstrates that the inhibition exerted by Hsp70 may be achieved by the interaction of the chaperone with pre-fibrillar species, in agreement with previous works (63,64,66,94,95). Further, the fact that the SBD556 shows the same effect with respect to FL-Hsp70 suggests that the mechanism of binding of Hsp70 for aggregated species is similar to those for monomeric misfolded substrates.

Finally, our data show a totally different behaviour for nucleotide-bound FL-Hsp70: during a normal aggregation (i.e. without seeding), the effect of the chaperone is attenuated by the presence of the nucleotides. Indeed we demonstrated that AS is able to trigger the co-aggregation of the chaperone and therefore to promote its depletion from solution (Fig. 3, Fig. 5B), confirming previous suggestions (66). During a seeded aggregation, Hsp70 is as active as the other chaperone variants, which suggests that FL-Hsp70, under this condition, does not aggregate with AS. This indicates that Hsp70 may be able to interact with a heterogeneous class of pre-fibrillar species and that those promoting its co-aggregation arise in the early oligomerization step of the aggregation process.

SUPPLEMENTARY INFORMATION

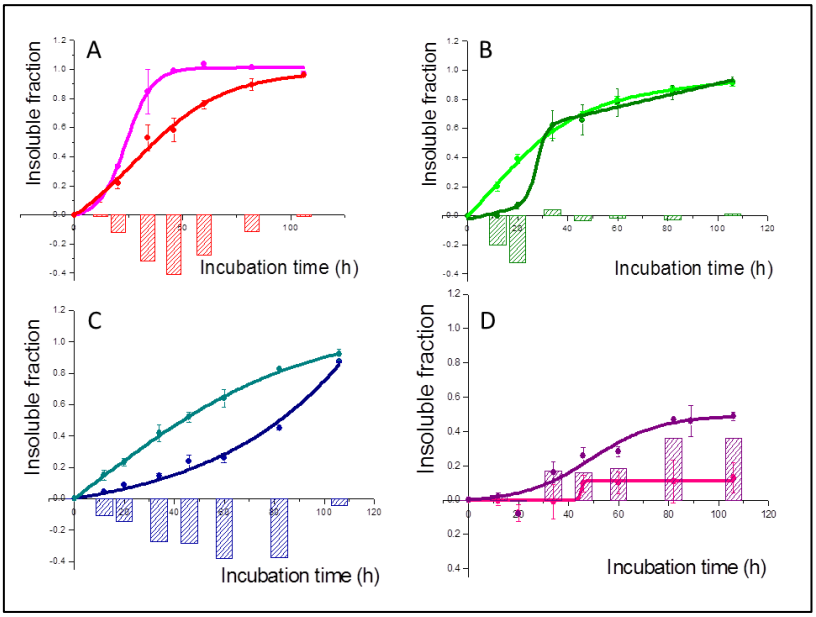

Fig. S1. Characterization of the stability of the chaperone variants in the aggregation conditions by insoluble fraction estimation. Insoluble protein fractions derived for SBD556 (A), SBD641 (B), FL-Hsp70 in the presence of ATP (C) and in the nucleotide-free state (D) during aggregation experiments performed in absence (light colours) or in presence of AS (dark colours) are shown. Each bar of the histograms is calculated subtracting chaperone quantity in absence of AS with the quantity in presence of AS.

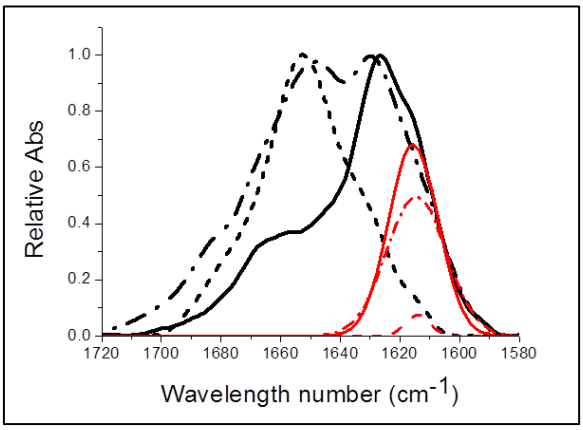

Fig. S2. Normalized FT-IR spectra of AS insoluble material at 20 (spots), 40 (dashes) and 110 h (line) of incubation. The 1614 cm^{-1} peak relative contribution for each spectrum is indicated in red.

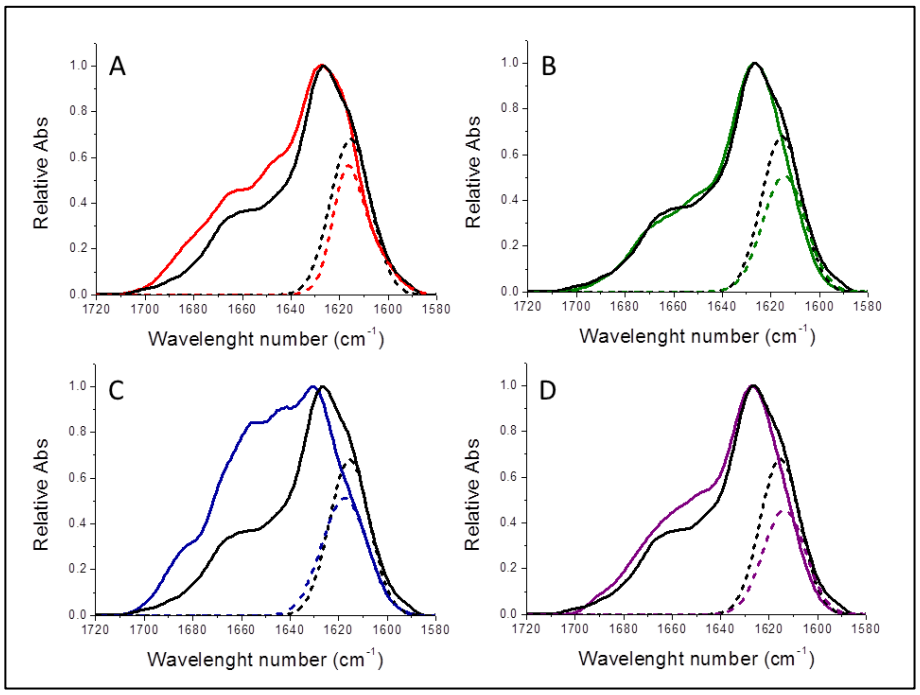

Fig. S4. Normalized FT-IR spectra of AS insoluble aggregates formed in the presence of SBD556 (A), SBD641 (B), nucleotide-free (C) and nucleotide-bound FL-Hsp70 (D). In all the panels the spectrum of the aggregates of AS alone is represented with a black line and the spectrum of AS in the presence of the chaperones with a coloured line (red for SBD556, green for SBD641, blue for nucleotide-free FL-Hsp70 and purple for nucleotide-bound FL-Hsp70). For each spectrum the relative contribution of 1614 cm^{-1} peak is indicated with a dashed line of the same colour.

2.6 References

1. **Soto, C. (2003) *Nat Rev Neurosci* 4, 49-60**
2. **Giurleo, J. T., He, X., and Talaga, D. S. (2008) *J Mol Biol* 381, 1332-1348**
3. **Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) *Nature* 416, 507-511**
4. **Campioni, S., Mannini, B., Zampagni, M., Pensalfini, A., Parrini, C., Evangelisti, E., Relini, A., Stefani, M., Dobson, C. M., Cecchi, C., and Chiti, F. (2010) *Nat Chem Biol* 6, 140-147**
5. **Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) *Nature* 431, 805-810**

6. **Lewy, F. H. (1912) Paralysis agitans. Pathologische Anatomie., In: Handbuch der Neurologie Band III (Lewandowsky M, Abelsdorff G, eds), pp. 920-933. Berlin: Springer Verlag.**
7. **Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) *Nature* 388, 839-840**
8. **Spillantini, M. G., Crowther, R. A., Jakes, R., Hasegawa, M., and Goedert, M. (1998) *Proc Natl Acad Sci U S A* 95, 6469-6473**
9. **Zhang, Y., Gao, J., Chung, K. K., Huang, H., Dawson, V. L., and Dawson, T. M. (2000) *Proc Natl Acad Sci U S A* 97, 13354-13359**
10. **Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K., and Suzuki, T. (2000) *Nat Genet* 25, 302-305**
11. **Lesage, S., and Brice, A. (2009) *Hum Mol Genet* 18, R48-59**
12. **Bonifati, V., Oostra, B. A., and Heutink, P. (2004) *J Mol Med (Berl)* 82, 163-174**
13. **Bandopadhyay, R., Kingsbury, A. E., Cookson, M. R., Reid, A. R., Evans, I. M., Hope, A. D., Pittman, A. M., Lashley, T., Canet-Aviles, R., Miller, D. W., McLendon, C., Strand, C., Leonard, A. J., Abou-Sleiman, P. M., Healy, D. G., Ariga, H., Wood, N. W., de Silva, R., Revesz, T., Hardy, J. A., and Lees, A. J. (2004) *Brain* 127, 420-430**
14. **Canet-Avilés, R. M., Wilson, M. A., Miller, D. W., Ahmad, R., McLendon, C., Bandyopadhyay, S., Baptista, M. J., Ringe, D., Petsko, G. A., and Cookson, M. R. (2004) *Proc Natl Acad Sci U S A* 101, 9103-9108**
15. **Taira, T., Saito, Y., Niki, T., Iguchi-Ariga, S. M., Takahashi, K., and Ariga, H. (2004) *EMBO Rep* 5, 213-218**
16. **Shendelman, S., Jonason, A., Martinat, C., Leete, T., and Abeliovich, A. (2004) *PLoS Biol* 2, e362**
17. **Zhou, W., Zhu, M., Wilson, M. A., Petsko, G. A., and Fink, A. L. (2006) *J Mol Biol* 356, 1036-1048**
18. **Petit, A., Kawarai, T., Paitel, E., Sanjo, N., Maj, M., Scheid, M., Chen, F., Gu, Y., Hasegawa, H., Salehi-Rad, S., Wang, L., Rogaeva, E., Fraser, P., Robinson, B., St George-Hyslop, P., and Tandon, A. (2005) *J Biol Chem* 280, 34025-34032**
19. **Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E. S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W. G., Lazzarini, A. M., Duvoisin, R. C., Di Iorio, G., Golbe, L. I., and Nussbaum, R. L. (1997) *Science* 276, 2045-2047**
20. **Krüger, R., Kuhn, W., Müller, T., Woitalla, D., Graeber, M., Kösel, S., Przuntek, H., Epplen, J. T., Schöls, L., and Riess, O. (1998) *Nat Genet* 18, 106-108**

21. **Zarranz, J. J., Alegre, J., Gómez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atarés, B., Llorens, V., Gomez Tortosa, E., del Ser, T., Muñoz, D. G., and de Yebenes, J. G. (2004) *Ann Neurol* 55, 164-173**

22. **Paisán-Ruíz, C., Jain, S., Evans, E. W., Gilks, W. P., Simón, J., van der Brug, M., López de Munain, A., Aparicio, S., Gil, A. M., Khan, N., Johnson, J., Martinez, J. R., Nicholl, D., Carrera, I. M., Pena, A. S., de Silva, R., Lees, A., Martí-Massó, J. F., Pérez-Tur, J., Wood, N. W., and Singleton, A. B. (2004) *Neuron* 44, 595-600**

23. **Zimprich, A., Müller-Myhsok, B., Farrer, M., Leitner, P., Sharma, M., Hulihan, M., Lockhart, P., Strongosky, A., Kachergus, J., Calne, D. B., Stoessl, J., Uitti, R. J., Pfeiffer, R. F., Trenkwalder, C., Homann, N., Ott, E., Wenzel, K., Asmus, F., Hardy, J., Wszolek, Z., and Gasser, T. (2004) *Am J Hum Genet* 74, 11-19**

24. **Jenner, P. (2003) *Ann Neurol* 53 Suppl 3, S26-36; discussion S36-28**

25. **Kahle, P. J., and Haass, C. (2004) *EMBO Rep* 5, 681-685**

26. **Bartels, T., Choi, J. G., and Selkoe, D. J. (2011) *Nature* 477, 107-110**

27. **Ullman, O., Fisher, C. K., and Stultz, C. M. (2011) *J Am Chem Soc***

28. **Wang, W., Perovic, I., Chittuluru, J., Kaganovich, A., Nguyen, L. T., Liao, J., Auclair, J. R., Johnson, D., Landeru, A., Simorellis, A. K., Ju, S., Cookson, M. R., Asturias, F. J., Agar, J. N., Webb, B. N., Kang, C., Ringe, D., Petsko, G. A., Pochapsky, T. C., and Hoang, Q. Q. (2011) *Proc Natl Acad Sci U S A* 108, 17797-17802**

29. **Uversky, V. N. (2007) *J Neurochem* 103, 17-37**

30. **Burré, J., Sharma, M., Tsetsenis, T., Buchman, V., Etherton, M. R., and Südhof, T. C. (2010) *Science* 329, 1663-1667**

31. **Chandra, S., Gallardo, G., Fernández-Chacón, R., Schlüter, O. M., and Südhof, T. C. (2005) *Cell* 123, 383-396**

32. **Clayton, D. F., and George, J. M. (1998) *Trends Neurosci* 21, 249-254**

33. **Clayton, D. F., and George, J. M. (1999) *J Neurosci Res* 58, 120-129**

34. **Abeliovich, A., Schmitz, Y., Fariñas, I., Choi-Lundberg, D., Ho, W. H., Castillo, P. E., Shinsky, N., Verdugo, J. M., Armanini, M., Ryan, A., Hynes, M., Phillips, H., Sulzer, D., and Rosenthal, A. (2000) *Neuron* 25, 239-252**

35. **Giasson, B. I., Murray, I. V., Trojanowski, J. Q., and Lee, V. M. (2001) *J Biol Chem* 276, 2380-2386**

36. **Uversky, V. N., Li, J., and Fink, A. L. (2001) *J Biol Chem* 276, 44284-44296**

37. **Takahashi, M., Kanuka, H., Fujiwara, H., Koyama, A., Hasegawa, M., Miura, M., and Iwatsubo, T. (2003) *Neurosci Lett* 336, 155-158**

38. **Paleologou, K. E., Schmid, A. W., Rospigliosi, C. C., Kim, H. Y., Lamberto, G. R., Fredenburg, R. A., Lansbury, P. T., Fernandez, C. O., Eliezer, D., Zweckstetter, M., and Lashuel, H. A. (2008)** ***J Biol Chem*** **283, 16895-16905**
39. **Vilar, M., Chou, H. T., Lührs, T., Maji, S. K., Riek-Loher, D., Verel, R., Manning, G., Stahlberg, H., and Riek, R. (2008)** ***Proc Natl Acad Sci U S A*** **105, 8637-8642**
40. **Conway, K. A., Lee, S. J., Rochet, J. C., Ding, T. T., Williamson, R. E., and Lansbury, P. T. (2000)** ***Proc Natl Acad Sci U S A*** **97, 571-576**
41. **Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T. (2002)** ***Nature*** **418, 291**
42. **Murphy, R. M., and Kendrick, B. S. (2007)** ***Biotechnol Prog*** **23, 548-552**
43. **Wood, S. J., Wypych, J., Steavenson, S., Louis, J. C., Citron, M., and Biere, A. L. (1999)** ***J Biol Chem*** **274, 19509-19512**
44. **Takahashi, T., and Mihara, H. (2008)** ***Acc Chem Res*** **41, 1309-1318**
45. **Muchowski, P. J., and Wacker, J. L. (2005)** ***Nat Rev Neurosci*** **6, 11-22**
46. **Jana, N. R., Tanaka, M., Wang, G., and Nukina, N. (2000)** ***Hum Mol Genet*** **9, 2009-2018**
47. **Chai, Y., Koppenhafer, S. L., Bonini, N. M., and Paulson, H. L. (1999)** ***J Neurosci*** **19, 10338-10347**
48. **Hay, D. G., Sathasivam, K., Tobaben, S., Stahl, B., Marber, M., Mestril, R., Mahal, A., Smith, D. L., Woodman, B., and Bates, G. P. (2004)** ***Hum Mol Genet*** **13, 1389-1405**
49. **Yamanaka, T., Miyazaki, H., Oyama, F., Kurosawa, M., Washizu, C., Doi, H., and Nukina, N. (2008)** ***EMBO J*** **27, 827-839**
50. **Huen, N. Y., and Chan, H. Y. (2005)** ***Biochem Biophys Res Commun*** **334, 1074-1084**
51. **Warrick, J. M., Chan, H. Y., Gray-Board, G. L., Chai, Y., Paulson, H. L., and Bonini, N. M. (1999)** ***Nat Genet*** **23, 425-428**
52. **Auluck, P. K., Chan, H. Y., Trojanowski, J. Q., Lee, V. M., and Bonini, N. M. (2002)** ***Science*** **295, 865-868**
53. **Auluck, P. K., Meulener, M. C., and Bonini, N. M. (2005)** ***J Biol Chem*** **280, 2873-2878**
54. **Klucken, J., Shin, Y., Masliah, E., Hyman, B. T., and McLean, P. J. (2004)** ***J Biol Chem*** **279, 25497-25502**
55. **Danzer, K. M., Ruf, W. P., Putcha, P., Joyner, D., Hashimoto, T., Glabe, C., Hyman, B. T., and McLean, P. J. (2011)** ***FASEB J*** **25, 326-336**
56. **McLean, P. J., Klucken, J., Shin, Y., and Hyman, B. T. (2004)** ***Biochem Biophys Res Commun*** **321, 665-669**

57. **Opazo, F., Krenz, A., Heermann, S., Schulz, J. B., and Falkenburger, B. H. (2008) *J Neurochem* 106, 529-540**
58. **Outeiro, T. F., Putcha, P., Tetzlaff, J. E., Spoelgen, R., Koker, M., Carvalho, F., Hyman, B. T., and McLean, P. J. (2008) *PLoS One* 3, e1867**
59. **Zhou, W., and Freed, C. R. (2004) *J Biol Chem* 279, 10128-10135**
60. **Muchowski, P. J., Schaffar, G., Sittler, A., Wanker, E. E., Hayer-Hartl, M. K., and Hartl, F. U. (2000) *Proc Natl Acad Sci U S A* 97, 7841-7846**
61. **Wacker, J. L., Zareie, M. H., Fong, H., Sarikaya, M., and Muchowski, P. J. (2004) *Nat Struct Mol Biol* 11, 1215-1222**
62. **Ahmad, B., Winkelmann, J., Tiribilli, B., and Chiti, F. (2010) *Biochim Biophys Acta* 1804, 223-234**
63. **Dedmon, M. M., Christodoulou, J., Wilson, M. R., and Dobson, C. M. (2005) *J Biol Chem* 280, 14733-14740**
64. **Huang, C., Cheng, H., Hao, S., Zhou, H., Zhang, X., Gao, J., Sun, Q. H., Hu, H., and Wang, C. C. (2006) *J Mol Biol* 364, 323-336**
65. **Luk, K. C., Mills, I. P., Trojanowski, J. Q., and Lee, V. M. (2008) *Biochemistry* 47, 12614-12625**
66. **Roodveldt, C., Bertoncini, C. W., Andersson, A., van der Goot, A. T., Hsu, S. T., Fernández-Montesinos, R., de Jong, J., van Ham, T. J., Nollen, E. A., Pozo, D., Christodoulou, J., and Dobson, C. M. (2009) *EMBO J* 28, 3758-3770**
67. **Scherzer, C. R., Eklund, A. C., Morse, L. J., Liao, Z., Locascio, J. J., Fefer, D., Schwarzschild, M. A., Schlossmacher, M. G., Hauser, M. A., Vance, J. M., Sudarsky, L. R., Standaert, D. G., Growdon, J. H., Jensen, R. V., and Gullans, S. R. (2007) *Proc Natl Acad Sci U S A* 104, 955-960**
68. **Witt, S. N. (2010) *Biopolymers* 93, 218-228**
69. **Evans, C. G., Chang, L., and Gestwicki, J. E. (2010) *J Med Chem* 53, 4585-4602**
70. **Gill, S. C., and von Hippel, P. H. (1989) *Anal Biochem* 182, 319-326**
71. **Rivers, R. C., Kumita, J. R., Tartaglia, G. G., Dedmon, M. M., Pawar, A., Vendruscolo, M., Dobson, C. M., and Christodoulou, J. (2008) *Protein Sci* 17, 887-898**
72. **Gasymov, O. K., Abduragimov, A. R., and Glasgow, B. J. (2010) *Biophys Chem* 149, 47-57**
73. **Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) *Science* 272, 1606-1614**
74. **Swain, J. F., Schulz, E. G., and Gierasch, L. M. (2006) *J Biol Chem* 281, 1605-1611**

75. **Burkholder, W. F., Zhao, X., Zhu, X., Hendrickson, W. A., Gragerov, A., and Gottesman, M. E. (1996)** ***Proc Natl Acad Sci U S A*** **93, 10632-10637**

76. **Gao, B., Eisenberg, E., and Greene, L. (1995)** ***Biochemistry*** **34, 11882-11888**

77. **Kasper, P., Christen, P., and Gehring, H. (2000)** ***Proteins*** **40, 185-192**

78. **Maeda, H., Sahara, H., Mori, Y., Torigo, T., Kamiguchi, K., Tamura, Y., Hirata, K., and Sato, N. (2007)** ***J Biol Chem*** **282, 26956-26962**

79. **Palleros, D. R., Shi, L., Reid, K. L., and Fink, A. L. (1994)** ***J Biol Chem*** **269, 13107-13114**

80. **Rajapandi, T., Wu, C., Eisenberg, E., and Greene, L. (1998)** ***Biochemistry*** **37, 7244-7250**

81. **Tapley, T. L., Cupp-Vickery, J. R., and Vickery, L. E. (2005)** ***Biochemistry*** **44, 12307-12315**

82. **Theyssen, H., Schuster, H. P., Packschies, L., Bukau, B., and Reinstein, J. (1996)** ***J Mol Biol*** **263, 657-670**

83. **Wittung-Stafshede, P., Guidry, J., Horne, B. E., and Landry, S. J. (2003)** ***Biochemistry*** **42, 4937-4944**

84. **Revington, M., Zhang, Y., Yip, G. N., Kurochkin, A. V., and Zuiderweg, E. R. (2005)** ***J Mol Biol*** **349, 163-183**

85. **Rist, W., Graf, C., Bukau, B., and Mayer, M. P. (2006)** ***J Biol Chem*** **281, 16493-16501**

86. **Liu, Q., and Hendrickson, W. A. (2007)** ***Cell*** **131, 106-120**

87. **Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M., and Gierasch, L. M. (2007)** ***Mol Cell*** **26, 27-39**

88. **Buczynski, G., Slepenkov, S. V., Sehorn, M. G., and Witt, S. N. (2001)** ***J Biol Chem*** **276, 27231-27236**

89. **Slepenkov, S. V., and Witt, S. N. (2002)** ***Biochemistry*** **41, 12224-12235**

90. **Moro, F., Fernández-Sáiz, V., and Muga, A. (2004)** ***J Biol Chem*** **279, 19600-19606**

91. **Zhu, M., Rajamani, S., Kaylor, J., Han, S., Zhou, F., and Fink, A. L. (2004)** ***J Biol Chem*** **279, 26846-26857**

92. **Mayer, M. P., and Bukau, B. (2005)** ***Cell Mol Life Sci*** **62, 670-684**

93. **Ahmad, A. (2010)** ***Int J Biol Macromol*** **46, 275-279**

94. **Zhou, Y., Gu, G., Goodlett, D. R., Zhang, T., Pan, C., Montine, T. J., Montine, K. S., Aebersold, R. H., and Zhang, J. (2004)** ***J Biol Chem*** **279, 39155-39164**

95. **Klucken, J., Outeiro, T. F., Nguyen, P., McLean, P. J., and Hyman, B. T. (2006)** ***FASEB J*** **20, 2050-2057**

3. The effect of molecular crowding on polyglutamine-containing protein aggregation

3.1 PolyQ diseases: inheritance and molecular mechanisms of the pathology

The human genome is characterized by the presence of several tracts of repeated trinucleotides. When these exceed a critical length, they may result in pathological conditions that have been classified as *Trinucleotide Repeat Expansion Diseases* (TREDs) (1). The mechanism of expansion is based on the formation of loops and hairpins and consequently to the insertion of additional repeats during DNA replication.

TREDs are grouped into two major classes (i.e. I and II) based on the position of the expansion in the genome: in class I TREDs the expansion is located into non-coding regions (usually in regulation elements) and therefore it can potentially affect the expression of the adjacent genes; in class II, the expansion occurs in translated regions and this leads to a gain of toxic function of the synthetized protein, which eventually acquire the capability to form toxic aggregated species.

A typical class II expansion consists of trinucleotide CAG repeats that are translated as an expanded tract of repeated glutamines (polyQ) in the encoded protein. The associated diseases, called *polyQ diseases*, arise in the case the polyQ length exceeds a certain threshold, typically of around 50Q. This group of disorders includes nine members: Huntington's disease (HD), dentatorubral-pallidoluysian atrophy (DRPLA), spinal and bulbar muscular atrophy (SBMA) and spinocerebellar ataxia types (SCA) 1, 2, 3 6, 7 and 17 (2-4). All these diseases are characterized by selective neuronal loss accompanied by a collection of associated physical and cognitive symptoms, though the particular features vary among the different diseases (2).

Although the different polyQ diseases have several common traits, the proteins associated with each different disorder share few or no homology outside the polyQ tract, being structurally and functionally unrelated (2,4). One important common aspect for all the polyQ diseases is the negative correlation between the age of onset and the number of CAG repeats, which involves that a greater number

of such repeats results in an earlier development of the disease (Fig. 3.1) (4-8).

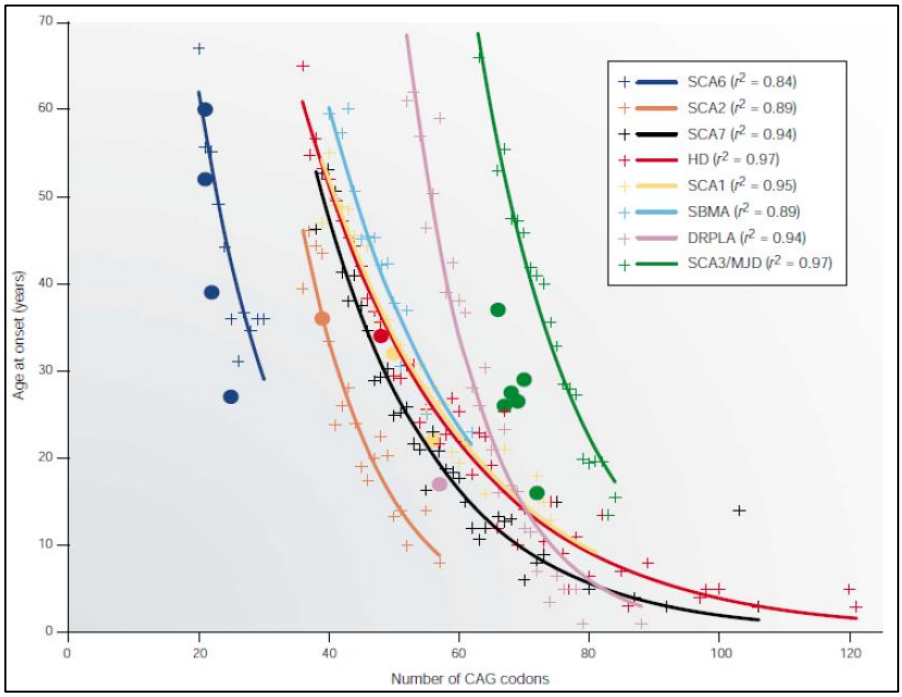

Fig. 3.1 Relationship between age at onset and CAG repeat length for polyQ disorders. Data from literature were used to calculate the mean age at onset (+) associated with various CAG repeat lengths and fitted using a simple exponential decay model. The age of onset of homozygotes is also shown (filled circles), plotted according to the longer of their two expanded CAG repeats (9).

As in Alzheimer's and Parkinson's disease, the hallmark of polyQ diseases is the formation of large amyloid protein aggregates containing the expanded and misfolded disease protein in the nucleus and/or cytoplasm of neurons (4). The mechanisms by which expanded polyQs give rise to the formation of amyloid aggregates and the reason why these are related to the onset of the polyQ diseases remain still unclear.

It has been proposed that polyQ fibrillation occurs by a multi-step process called of *nucleated polymerization* (10,11). The first step of this mechanism consists in the formation of ordered oligomers from polyQ monomers that have undergone a conformational change from random coil to β-sheet. During the second step, these oligomers function as nuclei for the rapid polymerization of further monomers

and oligomers into amyloid-like fibrils. Either the initial conversion or the subsequent oligomerization may be the rate-limiting step in polyQ fibrillogenesis (12).

Amyloid fibrils have been supposed to be responsible for the pathogenesis of polyQ diseases for a long time. Nevertheless recent evidences indicated oligomers as the potentially toxic species and support the idea that, to the contrary, fibrils might represent a protective mechanism exerted by the cell for seizing oligomers from the cytosol (3,13-15).

Although there are some common aspects to the different polyQ diseases that must be taken into account in order to study possible common mechanisms of pathogenesis, nevertheless, each of them also displays peculiar features. The different neurodegenerative and symptomatic profiles of each polyQ expansion disease may be explained as a result of the fact that the polyQ expansion is inserted into a different host protein in each different disease. The diverse properties of each of these proteins, which include their subcellular localization, abundance, size, structure, activity and biological role, along with the way the polyQ expansion affects them, shall constitute the factors responsible for each disease's specific presentation (2).

3.2 Ataxin-3, a model protein for investigating polyQ diseases

One of the most extensively investigated polyQ proteins is ataxin-3 (AT3), which is responsible for Spinocerebellar ataxia type 3 (SCA3) or Machado-Joseph Disease (MJD).

AT3 is widespread among eukaryotes, having been identified in protozoans, plants, fungi and animals, including nematodes, flatworms, arthropods and vertebrates. Despite the localized neuronal degeneration observed in MJD patients, AT3 displays in mice and humans a ubiquitous expression among different body tissues and cell types (16-20). It was found to be widely expressed throughout the brain, though different regions present varying expression levels (20).

AT3 is composed by a globular N-terminal domain followed by a flexible C-terminal tail (21). The N-terminal domain (residues 1-182), called Josephin domain (JD), displays a deubiquitinating (DUB) activity, while the flexible tail presents two Ubiquitine (Ub)-interacting motifs (UIMs) (residues 224-243; 244-263), followed by the polyQ region (Fig. 3.2) (22,23). Other features of the protein are a highly conserved nuclear localization signal (NLS) upstream of the polyQ (residues 282-285) and two nuclear export signal (NES) in the JD (residues 77-99, 141-158) (24,25). Further, five serine residues present in the UIMs (S236, S256, S260/S261, S340, S352) have been identified as potential phosphorylation sites; also, an ubiquitinatable lysine residue was mapped to residue 117, inside the JD.

The NMR structure of the JD reveals that it is mainly composed of two subdomains – a globular catalytic subdomain and a helical hairpin (26,27). The JD surface presents two binding sites for Ub; site 1, localized close to the catalytic cleft separating the two subdomains, and site 2, contiguous to site 1 but placed on the opposite surface (28). The two conserved UIMs located N-terminally of the polyQ region act cooperatively when binding Ub. The JD carries the catalytic triad found in cysteine-proteases (Cys14-His119-Asn134), similar to that found in Ub C-terminal hydrolases (UCH) and Ub-specific processing proteases (USP) (Fig. 3.2 B, C) (26,27).

Different human AT3 isoforms resulting from alternative splicing have been described, the longest having an approximate molecular weight of 42 kDa (17,20,29). Notably, the most common isoform found in the human brain has an extra UIM (residues 331-348) localized in the C-terminal region, downstream of the polyQ sequence (30). A recent study identified otherwhole 56 human

alternative splicing variants, expected to be translated into at least 20 isoforms, with varying predicted domain architecture (31), but the actual biological relevance of such variants remains unknown.

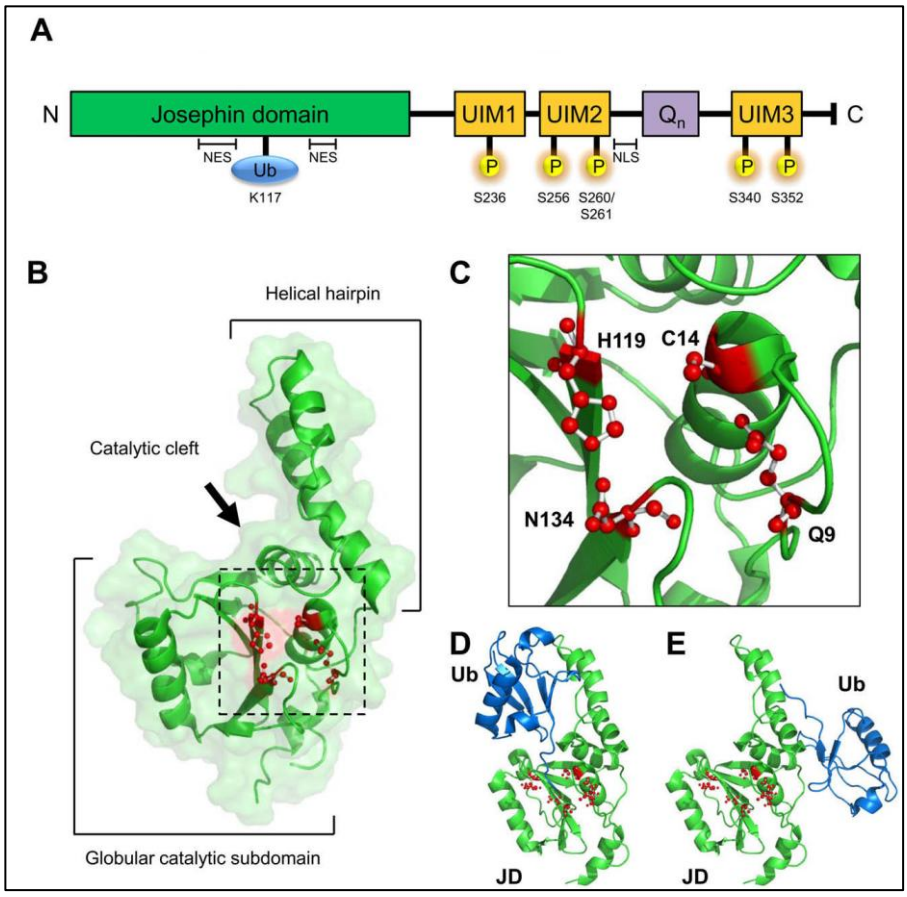

Fig. 3.2 (A) Schematic representation of AT3 isoform 2. (B) Structure of the JD solved by NMR (PBI code 1YZB) where the globular catalytic subdomain, the helical hairpin and the catalytic residues (in red) are shown. (C) Closeup of the catalytic cleft with in red the catalytic triad. (D, E) JD Ub-binding sites: site 1 is located close to the catalytic cleft and site 2 on the opposite surface (PDB code: 2JRI) (32).

Many experimental evidences suggest for AT3 a role in the ubiquitin-proteasome pathway (UPP), one of the main mechanisms of protein turnover. Short-lived or damaged proteins can undergo a covalent modification called ubiquitination (i.e. covalent attachment of Ub molecule or K48- or K63-polyUb chains to lysine residues) that

targets them to the proteasome for the degradation (33). It has been observed that inhibition of the DUB activity of AT3 in mammalian cells leads to an increase in polyubiquitinated proteins to a degree similar to what is observed when the proteasome is inhibited (34). *In vitro* studies have shown that AT3 is able to bind and cut both K48- and K63-polyUb chains in a UIM-dependent manner (23,35-37). However, Winborn and co-workers observed that intracellularly AT3 preferentially cleaves K63-linked chains and chains of mixed K48 and K63 linkage, suggesting that AT3 may function as a regulator of topologically complex polyUb chains (37). Moreover, further results suggest that AT3 functions as a polyUb-editing protease, shortening polyUb chains rather than completely disassembling them, in order to yield free Ub (23,37-40).

AT3 has been shown to interact with the p97/valosin-containing protein (VCP) (26,41-44) and with the Ub-like domains of the human homologs of the yeast DNA repair protein Rad23, HHR23A and HHR23B (27,45). Both p97/VCP and HHR23A and B are implicated in several biological processes, including the UPP; they have both been linked to the shuttling of polyubiquitinated substrates to the proteasome for degradation, particularly in endoplasmic reticulum-associated degradation (ERAD) that mediates the ubiquitination of misfolded proteins or unassembled complex constituents present in the secretory pathway.

A different aspect of AT3 function concerns its possible involvement in transcription regulation. In particular, it has been reported that AT3 is able to repress trascription in different manners: by inhibiting transcription activators as the cAMP response element-binding protein (CREB)-binding protein (CBP), p300 and p300/CBP-associated factor (PCAF) (46); by decreasing histone acetylation (47) through interaction with histone deacetylase 3 (HDAC3), nuclear receptor co-repressor (NCor) (47) and histones (46). Further, it has been proposed that AT3 deubiquitinating activity may interfere with the turnover of transcription regulators with which it interacts, thereby influencing repressor complex formation and activity (47,48).

The importance of AT3 interaction with components of the cytoskeleton such as tubulin, MAP2 and dynein may not be limited to a possible role in *aggresome* formation (38,49). Recent findings indicate that AT3 may play a role in the organization of the cytoskeleton itself, since its absence leads the disorganization of the

several cytoskeleton constituents (microtubules, microfilaments and intermediate filaments) and a loss of cell adhesions (50).

The mechanism by which polyQ-expanded AT3 leads to MJD pathogenesis has not been clarified yet. Although *wild-type* AT3 displays a ubiquitous distribution, in MJD patient, expanded AT3 accumulates as nuclear inclusions (NIs) only in neurons; recently, however, axonal inclusions have also been observed (51).

Further studies suggest that expanded AT3, like any other polyQ expanded protein, tends to form aggregates, as a result of polyQ expansion-induced misfolding and consequent transition to aggregation-prone conformations (52-58). As for most amyloid-forming proteins, several pathways may drive the conversion of the soluble protein to amyloid aggregates, through the formation of different conformationally altered monomeric or self-assembled multimeric species (59), being the small aggregates or oligomers the ones envisioned as the actual toxic species causing cytotoxicity.

Several works have focused on the aggregation mechanism of AT3, highlighting the complexity of this process. To date, it has been shown that the isolated JD has an intrinsic amyloidogenic potential, which results in the capability of the *wild-type* protein to aggregate. This implies that the aggregation pathway consists of two steps. The first gives rise to SDS-soluble oligomers and protofibrils as a consequence of aberrant interactions between the JDs; the second is accessible just to variants carrying expanded polyQs and results in the formation of mature, SDS-insoluble fibrils that are characterized by the formation of hydrogen bonds between polyQ glutamine side-chains (60-62) (Fig. 3.3). Expanded variants display the fastest aggregation kinetics, suggesting that the polyQ tract also affects the mode of JD aggregation (63). JD plays therefore a key role in the early conformational changes modulating the aggregation of both expanded and non-expanded AT3 (63-65) and, interestingly, the surfaces involved in the self-association overlap with the functionally relevant ubiquitin binding sites 1 and 2 (28,66). This observation provides a direct link between protein function and aggregation and a role for intracellular interactors in protecting against AT3 self-assembly, in keeping with the fact that Ub reduces *in vitro* aggregation of the JD (66). The C-terminal region of AT3 may also represent a bridge between physiological interactions and aggregation. In AT3, the regions predicted to form coiled-coils partially include the α-helical UIMs and regions shown to interact with molecular partners such as

VCP (41,67,68). When expanded, the polyQ may provoke aberrant protein interactions leading to AT3 aggregation. The connection between normal molecular interactions and aggregation may help explain the failure of non-expanded protein to self-aggregate in the crowded cell environment (60).

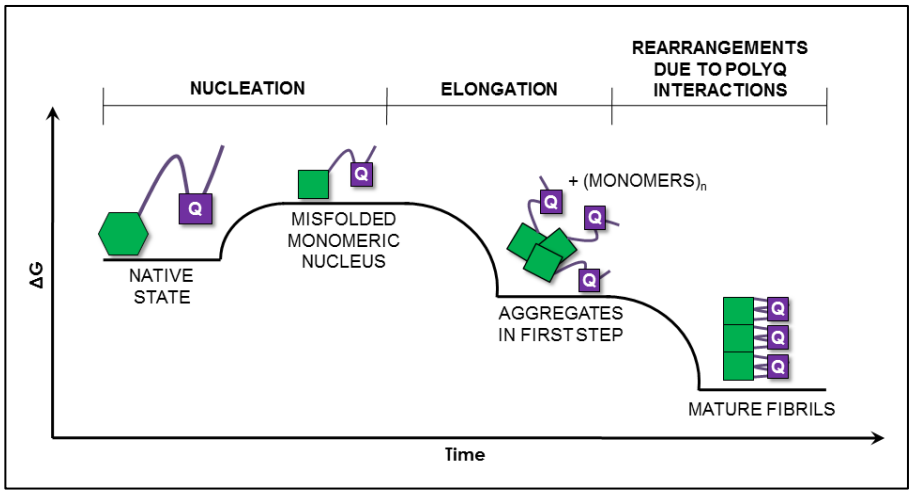

Fig. 3.3 Scheme of AT3 fibrillogenesis**.** In native AT3 the JD is represented as a hexagon and the disordered tract, including the polyQ (square), as a not-structured tail. AT3 fibrillization follows a two-step aggregation process. The first consists in the formation of a misfolded monomeric nucleus that is thermodynamically less stable with respect to the native protein. This conformational change is promoted by a structural rearrangement that does not involve the polyQ and lead to a first elongation step, driven by monomer addition. Only in the presence of an expanded polyQ, AT3 undergoes to a further aggregation step that lead to an increase in size and stability of the fibril (63).

3.3 *In vivo* studies on protein aggregation process

Despite many efforts put in characterizing amyloid formation *in vitro*, it is increasingly evident that the biological context in which aggregation occurs naturally, influences mechanism and rate of the process, as well as the structure and stability of the resulting fibrils (69).

In a typical eukaryotic cell, proteins, nucleic acids and others biopolymers have on the whole a concentration of 300-400 mg/ml and occupy 30% of the volume, a condition known as macromolecular crowding (MC) (Fig. 3.4a). The direct consequence of MC, which also exists in the extracellular space, is that the volume freely available to a molecule is only a fraction of the total volume where the macromolecule is dissolved. Therefore, the resulting excluded volume effects can favour thermodynamically compact states including both natively folded and aggregated states of proteins, though to a different extent depending on the protein (70). As a consequence, when proteins and peptides are unable to fold efficiently into monomeric compact states, aggregation is theoretically favoured, both kinetically and thermodynamically (70). Indeed, MC has been shown to accelerate the aggregation of a wide range of proteins *in vitro* (70,71). Moreover, the presence of chaperones, proteases and the continuous synthesis of the protein of interest make the understanding of the real *in vivo* amyloidogenic process a difficult achievement (72,73).

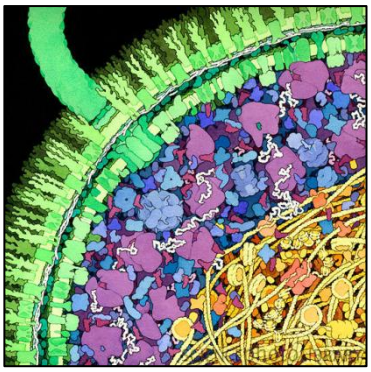

Fig. 3.4a Artwork of a section of *E. coli*. The cell wall (green, lower left to upper right) contains two membrane layers studded with proteins. The cytoplasm contains ribosomes (purple), tRNA molecules (brown), enzymes (blue) and mRNAs (white). The lower right portion of the image shows the genetic material (yellow and orange).

One of the most employed models for studying the mechanisms of amyloid aggregation is based on the capability of the prokaryotic organism *Escherichia coli* to form intracellular macro-structures of aggregated proteins, called Inclusion bodies (IBs). IB formation is strictly related with intra-cellular accumulation of misfolded proteins and it has been recently suggested to be part of a cellular protection

mechanism aimed to seize from the cytoplasm potentially harmful species.

The architecture and mechanisms of IB formation in bacteria have remained unexplored for years. Although IBs were conventionally regarded as disordered aggregates being formed by non-specific interactions of exposed hydrophobic surfaces, an increasing amount of evidence is showing that they are highly ordered protein deposits sharing many structural features with eukaryotic amyloid aggregates. In fact, as for amyloid fibrils, their assembly is driven by selective intermolecular interactions in a nucleation-dependent manner (Fig. 3.4b) (74,75). Proteins embedded in IBs appear to be enriched in β-sheet secondary structure elements (74,76). In addition, amyloid-specific dyes like Congo Red or Thioflavin T and S (ThT and ThS) bind to bacterial IBs with affinity similar for that for amyloid structures (74,76), which confirms a high degree of conformational similarity between the two types of aggregates. IBs display regions with high resistance against proteolytic attack, probably corresponding to a preferentially protected β-sheet core. Further, expression of some amyloidogenic proteins in *E. coli* resulted in the appearance of amyloid fibrils inside IBs and in the ability of isolated IBs to seed protein fibrillogenesis *in vitro* (75-80).

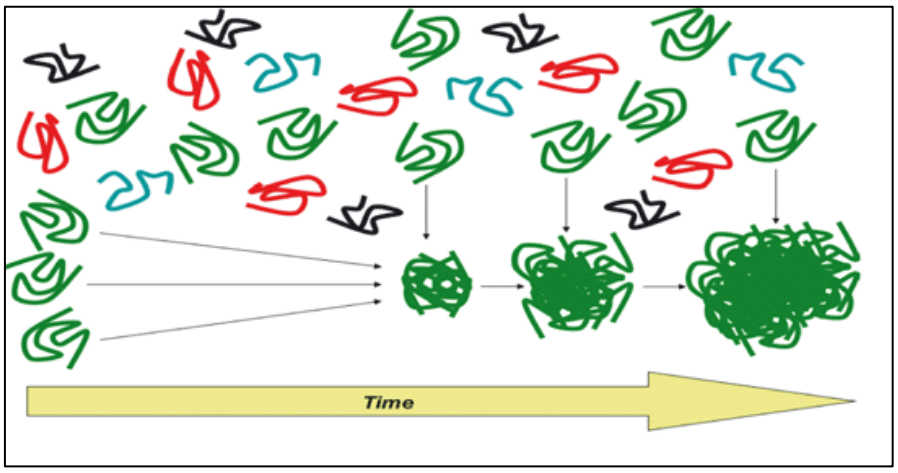

Fig. 3.4b *In vivo* nucleation/polymerization self-assembly process drives the formation of IBs in bacteria. Aggregation-prone versions of the recombinant protein (green) slowly form seeding nuclei by cross-molecular selective interactions. These proto-aggregates recruit further copies of the target protein, in a process compatible with first-order kinetics. This promotes a fast mass growth of IBs. Non-homologous

cellular or recombinant proteins (red, black) are excluded from these seeding events (81).

These similarities between bacterial IBs and amyloid fibrils provide the unique opportunity to dissect the molecular pathways triggering amyloid disorders in a simple, yet physiologically relevant, organism. Accordingly, *E. coli* has been employed for studying the correlation between protein aggregation and ageing (82), the role of the highly conserved protein quality machinery on the conformational properties of aggregated states (83), the effect of the protein sequence on *in vivo* aggregation kinetic (84), the influence of extrinsic factors like temperature on protein aggregation properties (85) or the dependence of polypeptide solubility in biological environments on the thermodynamic (86) and kinetic stability of proteins (87). In addition, the possibility of labelling aggregation-prone proteins with natural or artificial fluorophores (88) allows *in vivo* deposition pathways to be tracked in real time. Finally, bacteria provide a means to trap and study the highly toxic, unstable and transient intermediates in the fibrillation reaction, illuminating one of the more obscure but crucial steps in amyloid fibril formation (80).

3.4 Relationship between aggregation and toxicity of a polyglutamine-containing protein in the intracellular environments of *Escherichia coli*[+]

Gaetano Invernizzi[1]†*, Francesco A. Aprile[1]†, Antonino Natalello[1], Andrea Ghisleni[1¥], Amanda Penco[2], Annalisa Relini[2], Silvia M. Doglia[1], Paolo Tortora[1]*, Maria E. Regonesi[1]

[1] Department of Biotechnologies and Biosciences, University of Milano-Bicocca, Milan, Italy
[2] Department of Physics, University of Genoa, Genoa, Italy

*Running Title: "Two-steps AT3 aggregation in an intracellular environment"

To whom correspondence should be addressed: Gaetano Invernizzi and Paolo Tortora, Department of Biotechnologies and Biosciences, University of Milano-Bicocca, p.zza della Scienza n.2, Milano, Italy, Tel: (+39) 0264483430-3356; Fax (+39) 0264483565; emails: gaetano.invernizzi@unimib.it; paolo.tortora@unimib.it

Keywords: *Aggregation, amyloid, polyQ, ataxin-3, E. coli*

Background: Normal and expanded polyQ ataxin-3 undergo reversible and irreversible aggregation, respectively.
Results: Oligomers from expanded ataxin-3 and a truncated form retard growth of *Escherichia coli*; no significant effect by normal ataxin-3. Final aggregates prevent this effect.
Conclusion: Oligomers from normal protein are non-toxic, unlike those from the two other forms.
Significance: Intrinsic structural differences between oligomers from normal and expanded ataxin-3. Final aggregates are protective.

SUMMARY

Several neurodegenerative diseases are triggered by proteins containing a polyglutamine (polyQ) stretch expanded beyond a critical threshold. Among these, ataxin-3 (AT3) is the causative agent of spinocerebellar ataxia type-3. We expressed in *Escherichia coli* three AT3 variants: normal (AT3-Q24), expanded

(AT3-Q55) and truncated immediately upstream of the polyQ (AT3-291Δ); then, based on growth rate reduction, we quantified protein toxicity. We show that AT3-Q55 and -291Δ, unlike AT3-Q24, strongly reduced growth rate in the early stages (2-4 h). This correlated well with the appearance of soluble cytosolic oligomers, but not with the amount of insoluble protein in inclusion bodies (IBs). The impact of AT3-291Δ on cell growth clearly points to an intrinsic toxicity of the AT3 fragment. The expanded form, besides the typical FT-IR signal for intermolecular β-sheets, displayed an additional infrared signature assigned to glutamine side-chain hydrogen bonding and associated with SDS-insoluble fibrils, elongation was monitored by AFM. This mirrors the well-known *in vitro* two-step aggregation pattern of expanded AT3. Also, we demonstrated that final aggregates of strains expressing expanded or truncated AT3 have a protective role against the soluble toxic oligomers. Noteworthy, although normal AT3 gave rise to soluble oligomers at levels comparable to those of the two other forms, it was essentially non-toxic. Thus, its lack of toxicity *in vivo* should be related to its intrinsic structural features. Further, our findings suggest that the mechanisms of toxicity are evolutionarily conserved.

INTRODUCTION

Protein aggregation into highly stable and structured deposits, the so-called amyloids, is a hallmark of many neurodegenerative diseases (89). Although the proteins involved in the onset of the diseases share no sequence and structure similarities, the aggregation process leading to amyloid fibrils involves a common pathway generally triggered by protein misfolding, followed by soluble oligomeric aggregates formation, which in turn are precursors to fibrils (89). In recent years, plenty of evidence pointed out the soluble oligomers as the causative agents of cell toxicity (90-96), whereas insoluble amyloid deposits raised a controversial debate on their role as a harmful reservoir or sequestering deposits of toxic soluble species (95-103). Interestingly, strategy mechanisms capable of seizing potentially harmful misfolded polypeptides into insoluble deposits seem to be perpetuated throughout the evolution from prokaryotic bacteria to higher complex organisms (104,105). Indeed, in *Escherichia coli*, misfolded proteins are accumulated into inclusion bodies (IBs), insoluble proteins macro-aggregates that share several

distinctive features with amyloid deposits. IBs, like amyloid fibrils, are structurally organized in intermolecular β-strands, can bind classical amyloid dyes (ThT, Congo red) and their assembly is driven by selective intermolecular interactions that make the aggregation process highly protein-specific (106,107). Further, expression of some amyloidogenic proteins in *E. coli* resulted in the appearance of amyloid fibrils inside IBs and in the ability of isolated IBs to seed protein fibrillogenesis *in vitro* (108-110). Coupled with its ease of handling, this makes *E. coli* an effective host for fibrillogenesis studies under intracellular conditions that cannot easily be reproduced *in vitro,* such as high molecular crowding, presence of chaperones and proteases, continuous synthesis of the protein of interest (110-113). In particular, *E. coli* has been successfully employed to monitor *in vivo* the multistep aggregation mechanism of an artificial chimera harboring a polyglutamine (polyQ) tract (113). Proteins containing stretches of repeated glutamines whose length exceeds a critical threshold undergo amyloid aggregation, which results in neurodegeneration (114,115). The aggregation pathway of the most investigated polyQ proteins, huntingtin (Htt) and ataxin-3 (AT3) has been extensively studied *in vitro* and demonstrated to consist of a multistep mechanism involving different domains of the protein (116-118). In particular, regions flanking the polyQ are responsible for the first steps of aggregation, whereas polyQ is involved in a subsequent step leading to the formation of SDS-insoluble fibrillar aggregates. By employing AT3 as a model, we recently demonstrated that the involvement of the polyQ in the second stage of the aggregation pathway results in the formation of hydrogen bonds among glutamine side chains, which leads to the irreversible formation of SDS-insoluble aggregates (62). Nevertheless, the multistep aggregation mechanism of an authentic polyQ protein has never been investigated *in vivo* and, in particular, it has never been verified whether glutamine-glutamine interaction (62,119,120) is a hallmark of polyQ amyloids in the intracellular environment. Further, the possibility to investigate the aggregation pathway *in vivo* would raise the opportunity to study the intriguing relationship between protein aggregation and toxicity. In this view, *E. coli* has proven to be a convenient host to asses cytotoxicity associated with protein expression. In fact, detrimental effects on the cell growth of *E. coli* are reported for the expression of GST harboring an expanded polyQ, as well as for Htt carrying more than 50 glutamine repeats (121,122). Here, we characterize for the

first time in the intracellular environment of *E. coli* the multistep aggregation mechanism of authentic variants of the polyQ protein AT3, investigating in particular the relationship between aggregation pathway and cytotoxicity. Using a protein variant carrying an expanded polyQ, we show a correlation between the appearance of soluble species and cytotoxicity, as well as a protective role for insoluble species appearing at the latest stages of the process. We also show that an AT3 variant deprived of polyQ exerts a detrimental effect on cell growth comparable to that of the polyQ expanded variant, which suggests a possible role of the polyQ context in determining cell toxicity.

EXPERIMENTAL PROCEDURES

AT3 variants cloning and expression

AT3 variants were cloned into pET-21a plasmid (EMD Biosciences) and expressed in *E. coli* Rosetta™ (DE3) pLacI Competent Cells (EMD Bioscience). AT3-Q24 and -Q55 were directly inserted into previously mutagenized NdeI/XhoI restriction sites; AT3-291Δ was obtained by phosphorylated oligonucleotide PCR on Q24 cDNA with the following primers: 291Δ-Rev 5' P-TATTTTTCAAAGTAGGCTTCTCGTCTC 3'; 291Δ-For 5' P-TAGCCCGGGTCGACTCGAG 3'. The integrity of the expression hosts was verified by monitoring the growth rate under non-inductive condition. All *E. coli* employed strains showed identical growth profiles in LB medium without IPTG (Fig.S1).

Growth profile analysis

Growth experiments were performed at 37°C or 26°C, under constant shaking at 150 rpm in LB standard medium. Cells were induced at OD_{600} = 0.1 with 500 µM IPTG (Qiagen), unless otherwise specified, and growth rates were measured by monitoring OD_{600} at each hour after induction. To study the toxic effect of pre-fibrillar species, 2 h-inducted cells were centrifuged for 10 min at 3000 g and 37°C, resuspended in the same volume of fresh media deprived of IPTG and analyzed for the growth profile for another 4 h. The protective role of IBs was tested monitoring the growth profile of 6 h-inducted cells upon re-inoculation in fresh medium at OD_{600} = 0.1.

Propidium iodide staining of E. coli cells

To assess viability, cells were subjected to propidium iodide staining. The dye intercalates DNA with resulting enhancement of fluorescence emission, but cannot penetrate intact membranes. Cells were centrifuged for 15 min at 6000 g and 4°C; pellets were then resuspended in 1 volume of 10 mM $MgSO_4$ and centrifuged again as above. Resulting pellets were resuspended in the same buffer at an OD_{620} of 0.4 and incubated for 30 min in the dark with 5 µg/ml of propidium iodide, in the absence or the presence of 10% isopropanol, the latter treatment resulting in quantitative cell lysis. The samples were centrifuged for 15 min at 6000 g; pellets were resuspended in the same buffer at an OD_{620} of 0.4 and fluorescence emission determined at 616 nm (excitation: 535 nm) using a Cary Eclipse spectrofluorimeter (Agilent Technologies) in a 3 ml quartz cuvette.

Soluble and insoluble protein fraction extraction and inclusion bodies (IBs) purification

Cells were harvested for 15 min at 6000 g and 4°C, frozen at -80 °C in 3 volumes of Lysis Buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA) and stored overnight. Then, the cell suspension was thawed; 0.5 mM PMSF, 0.8 mg/ml lysozyme (final concentrations) was added and the resulting mixture incubated at 37°C for 30-45 min, which caused cell lysis. 1% NP-40 was added to the lysate and further incubated at 4°C for 40-60 min. Finally, 25 µg/ml DNAse, 10 mM $MgSO_4$ was added and the incubation prolonged at 37°C for 30-40 min. IBs were separated from soluble protein by centrifugation at 10000 g for 15 min; the pellet was resuspended in 1 ml of Lysis Buffer, 0.5% Triton X-100 and incubated under shaking at 4°C for 15 min. IBs were then collected by centrifuging for 15 min at 10000 g and 4°C, and washed twice with 1 ml of either milliQ water (for FTIR analysis) or Lysis Buffer, Triton X-100 (for all other analyses), in both cases centrifuging at 3000 g for 15 min. All reagents were from Sigma Aldrich.

SDS-PAGE and densitometry analysis of soluble and insoluble protein fractions

Soluble and insoluble protein fractions from cell cultures collected at the same OD were isolated as reported in the IBs purification protocol and analyzed in 14% SDS-PAGE. Total protein was revealed by Coomassie staining (ImperialTM protein stain, Pierce); AT3 by western blotting using Z46 rabbit polyclonal antibody

(123). Immunoreactive bands were detected using ECL Western blotting reagent (GE Healthcare) and protein band intensity quantified using ImageJ software (NIH).

Formic acid treatment for inclusion body protein quantification

To solubilize protein from isolated IBs, these were resuspended in 1 ml 100% formic acid and incubated at 37°C for 30 min (124). Then, protein was determined by absorbance measurement at 280 nm.

Filter Trap assay for the identification of SDS-insoluble species

Purified IBs from 10 ml growth cultures were resuspended in 500 µl of milliQ water and quantified by formic acid treatment as described above. Samples were centrifuged for 15 min at 10,000 g and 4°C, and pellets resuspended in SDS buffer (20 mM Tris-HCl pH 6.8, 10 mM DTT, 5% SDS) to a final protein concentration of 0.2 mg/ml. Samples were then boiled for 15 min at 99°C and 50 µl vacuum filtered in a 48-well dot-blot apparatus through a 0.2 µm pore-sized cellulose acetate membrane. The membrane was washed twice with SDS buffer and probed with polyclonal Z46 antibody as reported above.

Fourier transform infrared spectroscopy

FTIR measurements on IBs were performed following an approach previously validated (125). In particular, the extracted insoluble fractions were resuspended in distilled water and a few microliter1s of the suspension were deposited onto a BaF_2 window and dried at room temperature. FTIR absorption spectra were collected in the transmission mode by coupling the Varian 610-IR infrared microscope – equipped with a nitrogen-cooled Mercury Cadmium Telluride (MCT) detector – to a Varian 670-IR spectrometer (Varian) under the following conditions: 2 cm^{-1} spectral resolution, 25 kHz scan speed, 512 scan co-additions, and triangular apodization. When necessary, spectra were corrected for possible residual water vapor. The second derivatives of the absorption spectra were obtained after a 11-point binomial smoothing followed by the Savitzky–Golay derivative method (3rd polynomial, 5 smoothing points), using the GRAMS/32 software (Galactic Industries).

Atomic force microscopy

For AFM imaging, 10-µl aliquots of the sample obtained from inclusion bodies after washing with SDS were deposited on freshly cleaved mica and then dried under mild vacuum. AFM measurements were performed in air using a Dimension 3100 scanning probe microscope equipped with a G scanning head (maximum scan size 100 µm) and driven by a Nanoscope IIIa controller and a Multimode SPM, equipped with "E" scanning head (maximum scan size 10 mm) and driven by a Nanoscope V controller, (Digital Instruments, Bruker AXS GmbH). Images were acquired in tapping mode in air using single beam uncoated silicon cantilevers (type OMCL-AC160TS, Olympus). The drive frequency was between 290 and 320 kHz and the scan rate was between 0.5 and 1.0 Hz. Aggregate height was measured from the height in cross-section in the topographic AFM images. The length of fibrillar aggregates was measured from the topographic AFM images with a routine implemented using MATLAB (The Mathworks Inc.). Such routine approximates the fibril profile using a sequence of segments.

RESULTS

The model system

To provide insight into the relationships between aggregation and toxicity in an intracellular environment, we expressed in *E. coli* three authentic AT3 variants, i.e., lacking of any tags or fusion partners. This is relevant to the goal of the present investigation, as it has been demonstrated that the particular protein context of polyQ can greatly influence the aggregation pathway and, consequently, aggregate toxicity (126,127). Specifically, we expressed a variant truncated immediately upstream of the polyQ (AT3-291Δ), a *wild-type* and a pathogenic AT3 variant carrying 24 (AT3-Q24) and 55 (AT3-Q55) consecutive glutamines, respectively. As a control, we used β-lactoglobulin (BLG), a non-amyloidogenic protein, which has been shown to aggregate into inclusion bodies (IBs) when overexpressed in *E. coli* cells (128).

AT3-291Δ and AT3-Q55 expression is deleterious to cell growth

It has been reported that in *E. coli* cell growth and viability are affected by polyQ expression. Furthermore, a correlation between polyQ length and toxicity has been clearly demonstrated (121,122). Accordingly, in our model *wild-type*- and expanded AT3-expression

strains displayed substantially different growth profiles (Fig. 1A, B), the latter growing much more slowly than the former in the first 3 h after induction. Surprisingly, the strain expressing AT3-291Δ showed a profile very similar to that of the expanded form. Growth profiles of the strains expressing AT3-Q24 and BLG were almost identical to each other, thus showing that the detrimental effect on cell growth exerted by *wild-type* AT3 is merely due to protein overexpression (129).

To assess cell viability we performed propidium iodide staining, which is based on the capability of the dye to bind DNA with resulting fluorescence enhancement, but not to permeate intact cell membranes. Cells expressing all AT3 variants did not appreciably take up the dye, consistent with the hypothesis that both AT3-291Δ and AT3-Q55 toxicity do not affect cell viability during the experimental time course (Fig. S2).

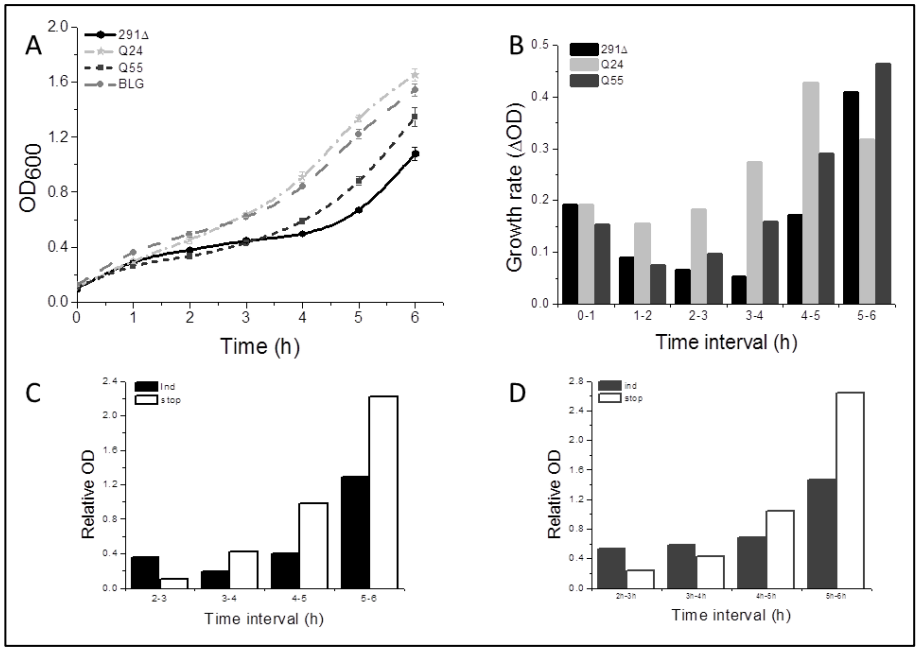

Fig. 1. Expression of AT3-291Δ and AT3-Q55 and cytotoxicity. A) Growth curves of *E. coli* strains expressing BLG (grey, dots), AT3-291Δ (black, dots), AT3-Q24 (light grey, stars) and ATQ55 (dark grey, squares). Cultures were induced with 500 μM IPTG at OD_{600} = 0.1 and growth profiles at 37°C monitored. Error bars represent standard deviations and are derived from at least 5 independent experiments. B) Changes in optical density per hour (ΔOD) derived from growth profiles of AT3-291Δ (291Δ), AT3-Q24 (Q24) and ATQ55 (Q55). C, D) Effect of

removal of the inducer on growth rates of AT3-291Δ (C) and AT3-Q55 (D). After 2 h from the induction, cells were resuspended in an inducer-deprived medium (stop, empty bars); as a control, they were further cultured in the presence of the inducer (ind, full bars). ΔODs were normalized with respect to the growth profile of AT3-Q24.

To confirm that toxic species arising from AT3-Q55 and AT3-291Δ are actually produced in the earliest stages of their intracellular aggregation, we stopped protein induction after 2 h by replacing the growth medium by one deprived of IPTG, and monitored cell growth. Indeed, it has been reported that, following IPTG deprivation, protein expression is strongly inhibited in T7-based inducible strains, like the one which we are employing (130). Thus, after IPTG removal we monitored growth rates and normalized them with respect to that of the *wild-type*, to allow for the effect of overexpression-related cellular stress. By comparing them with those of the same strains under induction (Fig.1C, D), it is apparent that after IPTG removal the deleterious effects of the toxic species still persisted for about 4-5 h.

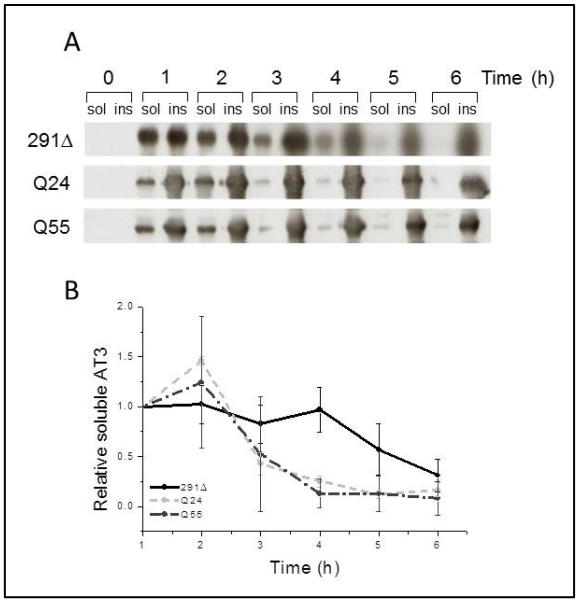

Fig. 2. Soluble protein fraction analysis of *E. coli* strains expressing AT3 variants. A) soluble (sol) and insoluble (ins) fractions of strains expressing AT3 variants were collected at different times after induction, subjected to SDS-PAGE and western blotted using anti AT3-antibodies. B) soluble protein fraction was quantified by densitometry of western blots. Standard deviations are from at least five independent experiments. Signals were normalized at t = 1h.

AT3 toxicity correlates with the amount of soluble protein fraction and declines at the onset of the aggregation

Although the results above reported clearly indicate that the AT3-291Δ and Q55 variants are toxic to the cell, they do not provide,

by themselves, any clue as to the identity of the toxic species. In recent years, plenty of evidence highlighted a critical role for soluble pre-fibrillar amyloid species in triggering cellular toxicity. Actually, detrimental effects of early oligomers have been demonstrated for a number of amyloid proteins including polyQ proteins such as huntingtin and atrophin-1 (94-96). Thus, to settle the issue, at different times after induction we isolated soluble and insoluble protein fractions from our expression strains (see *Experimental Procedures*). AT3 was revealed by western blot analysis (Fig. 2A), and quantified in the soluble fraction by densitometric analysis (Fig. 2B). Already at the first hour after induction soluble AT3 was detected in all three strains and reached a maximum at 2 h. Then, soluble AT3-Q55 and AT3-Q24 declined, reaching almost undetectable levels at 4 h. In contrast, in the bacterial strain expressing AT3-291Δ, soluble protein was abundant and essentially constant over the first 3-4 h, a significant decrease taking place only at later times (Fig. 2B). Remarkably, insoluble protein levels were high and constant in all three strains throughout the growth (Fig. 2A). Our results highlight two important points: i) AT3-291Δ- and Q55-expression strains show a remarkable recovery of growth only concomitant to a significant decrease in AT3 soluble fraction. ii) there is no correlation between growth rates and the amounts of insoluble protein detected

To further substantiate a possible correlation between soluble protein level and toxicity, relative amounts of soluble AT3 at the different sampling times after induction were compared with cell growth rates of the respective strains, normalized with respect to the growth rate of AT3-Q24 (Fig. S3). Then, based on these data, The Pearson Correlation Index (PcI) was determined. A value of -0.57 and -0.60 was assigned to AT3-291Δ and AT3-Q55, respectively, which highlights a good linear correlation. In contrast, no significant correlation was found for AT3-Q24, as supported by a PcI of -0.20. These results are consistent with the hypothesis that the soluble species responsible for the cytotoxic effects are those resulting from protein aggregation in the first 3-4 h from induction.

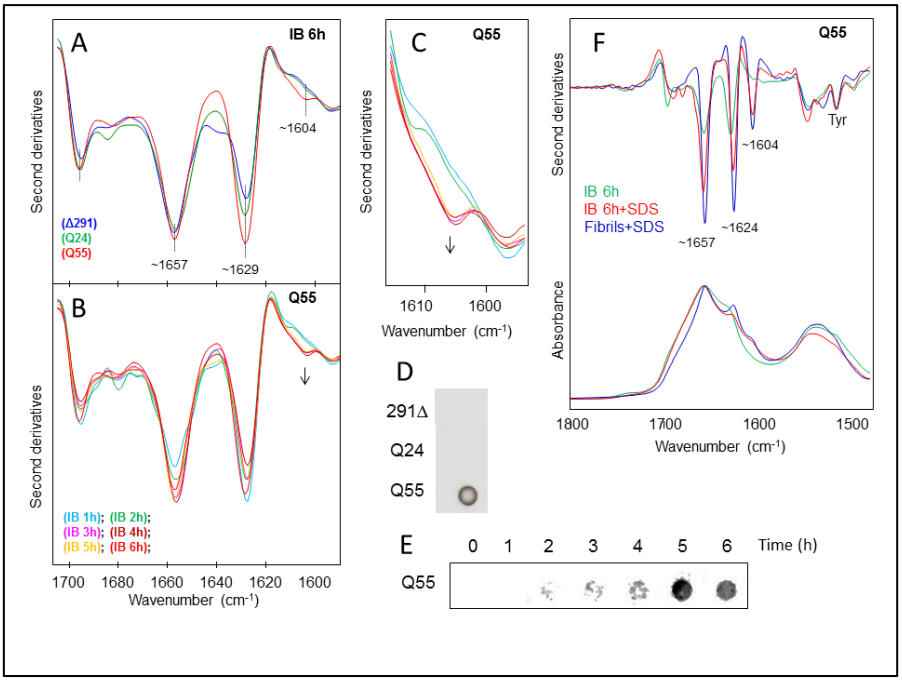

Fig. 3. FTIR and filter trap analyses of insoluble protein fractions of *E. coli* strains expressing AT3 variants. A) Second derivative spectra of the insoluble fractions collected 6 h after induction of the AT3-291Δ, AT3-Q24, and AT3-Q55 expression strains. The peak positions of the main components are indicated. B) Second derivative spectra of the insoluble fractions collected at different times after induction of the AT3-Q55 expression strain. The arrows point to increasing incubation time of the ~1604 cm^{-1} peak, hallmark of glutamine side-chains hydrogen bonding. C) Second derivative FTIR spectra from (B), reported on an enlarged wavenumber-scale in the spectral region of NH_2 deformation modes of glutamines. D) Filter trap assay of protein insoluble fractions of the different AT3-291Δ, AT3-Q24 and AT-Q55 after 6 h of induction. E) Filter trap assay of protein insoluble fractions of AT3-Q55 at different times after induction. F) FTIR absorption spectra and their second derivatives collected from IBs extracted 6 h after induction, before and after SDS treatment. The spectrum of AT3-Q55 SDS-insoluble fibrils, obtained *in vitro* after incubation for 168 h of the purified protein at 37°C (62), is also given for comparison. All second derivative spectra are normalized at the ~1515 cm^{-1} peak of tyrosine to allow for possible differences in protein content.

Structural characterization of the insoluble protein fractions

To provide insight into the structural features of the insoluble aggregated species, we took advantage of FTIR spectroscopy, a powerful technique that provides information on secondary structure content of both protein in solution and insoluble protein species. Also, this technique has been successfully applied to monitor protein

aggregation in bacteria (131). We therefore analyzed insoluble protein fractions from the strains expressing the three AT3 variants. The second derivatives spectra of the insoluble aggregates extracted 6 hours after induction are presented in Fig. 3A. The spectra of the three AT3 variants displayed well resolved components at 1657 cm^{-1}, assigned to random coils and β-helices, and the two bands at 1696 cm^{-1} and 1629 cm^{-1} due to intermolecular β-sheet structures in protein aggregates. Only in the spectrum of the AT3-Q55 variant, we observed an additional component around 1604 cm^{-1}that in a previous work on AT3 fibrillation *in vitro* (62) we assigned to glutamine side chains involved in a network of hydrogen bonds. Interestingly, we also found that this band is a hallmark of AT3-Q55 SDS-insoluble fibrils. Here, during the time course of the expression of AT3-291Δ and Q24, no significant spectral variation of the insoluble aggregates was detected by FTIR spectroscopy (Fig. S4), indicating that proteins did not undergo major structural rearrangement. Instead, for the AT3-Q55 variant we observed the appearance of the ~1604 cm^{-1} component starting from 3 hours after induction (Fig. 3B, C). This shows that structural ordering of the glutamine side chains, involved in a network of hydrogen bonds, also occurs *in vivo* at this stage of aggregation. In keeping with our *in vitro* results (62), we expected that the AT3-Q55 aggregates would consist of SDS-insoluble protein. In fact, filter trap assays performed on insoluble protein fractions (Fig. 3D) revealed the presence of SDS-resistant species only in the strain expressing AT3-Q55 starting from 3 h after induction (Fig. 3E). Noteworthy, SDS-insoluble material extracted from the IBs, displayed a FTIR spectrum almost identical to that collected from the *in vitro* irreversibly aggregated AT3-Q55 (Fig. 3F). In particular, both spectra displayed the two components at 1657 cm^{-1} and 1604 cm^{-1}, respectively due to the C=O and NH vibrations of the glutamine side chains, besides the well resolved band at 1624 cm^{-1} characteristic of intermolecular β-sheet structures (62).

The SDS-insoluble protein fraction of *E. coli* expressing AT3-Q55 was analyzed by tapping mode AFM at different induction times to get insight into aggregates morphology. Short elongated protofibrils with height of about 1 nm were observed after 3 h (Fig. 4A). Longer protofibrils with height between 2 and 20 nm were present after 5 h (Fig. 4B), while after 6 h fibrillar structures with similar height were also found (Fig. 4C). Fig. 4D shows the contour length of elongated aggregates (protofibrils and fibrils) measured from the topographic

AFM images as a function of the induction time. The aggregate population, which after 3 h of induction had a mean length of 63 ± 9 nm, was gradually enriched with longer structures, some of which reached a few microns at 6 h of induction. Even at large induction times the fibrillar structures were found to coexist with globular aggregates. The latter can be identified as oligomeric species.

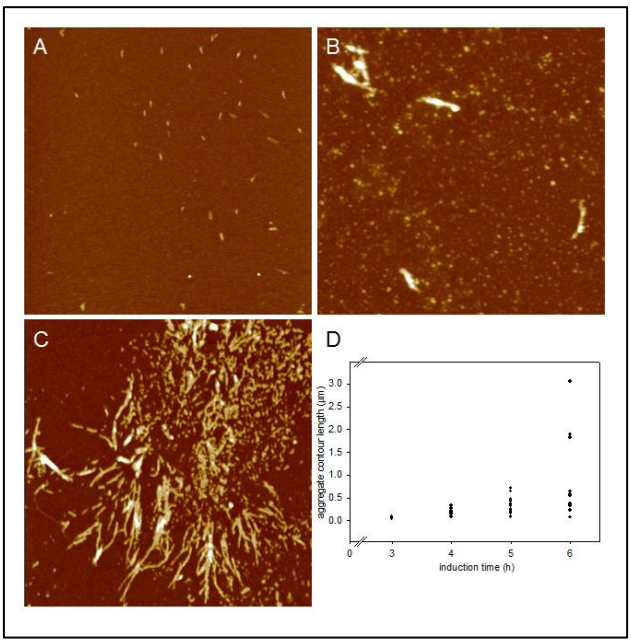

Fig. 4. Atomic force microscopy images of AT3-Q55 insoluble protein fraction. A-C) Tapping mode AFM images (height data) of AT3-Q55 insoluble protein fraction after 3h (A), 5h (B), 6h (C) of induction. Scan size 3.0 μm; Z range 5.0 nm (A), 25 nm(B), 15 nm (C). D) Contour length of elongated aggregates (protofibrils and fibrilis) as a function of the induction time.

AT3-291Δ and Q55 soluble species toxicity is strongly prevented by insoluble aggregates

In recent years, increasing evidence has provided support to the idea that the cytotoxic species are soluble oligomeric aggregates and fibrils or amyloid plaques are not likely to be the causative agent of pathogenesis. It has been also suggested that they might even exert a protective effect against oligomer toxicity (94-98,102). Interestingly, *E. coli* IBs are supposed to be a defensive mechanism against potentially harmful misfolded proteins (105,132). In keeping with this view, we observed that growth rates of AT3-Q55 and AT3-291Δ-expression cells, after 5-6 h from induction increased up to values comparable to those of non-induced strains in exponential growth (Fig. 1A; Fig. S1). Actually, in this lapse of time, the aggregated species detected intracellularly were almost exclusively insoluble (Fig. 2A).

To assess whether insoluble aggregates could also exert a protective effect against soluble species toxicity, strains expressing AT3 variants grown for 6 h under inducing conditions were re-inoculated at the initial optical density (0.1 OD) into a fresh new medium, still in the presence of IPTG. After re-induction, growth rates of AT3-291Δ and AT3-Q55-expression strains were substantially higher than those observed after the primary induction (Fig. 5A; for comparison see Fig. 1A) and comparable to those of re-induced AT3-Q24 (Fig. 5B and C), which points to a protective role of the insoluble species against the soluble toxic oligomers. It should be also noted that growth rates of re-induced AT3-Q24 were higher than those observed during the primary induction, which could be accounted for by the massive recruitment into IBs of soluble misfolded protein, as already reported in the case of another nontoxic protein expressed in the prokaryotic host (132). Cell fractionation of re-induced AT3-291Δ and Q55 showed the massive presence of aggregated species at all times analyzed, whereas soluble species were hardly detectable (Fig. 5D). SDS-insoluble species were present over the entire re-induction time (3 h) in the strain expressing AT3-Q55 (Fig. 5E). Further, it is interesting to note that AFM images of insoluble AT3-Q55 fraction after 6 h of induction revealed the presence not only of fibrillar aggregates but also of oligomeric species that were trapped inside the IBs (Fig. 4C).

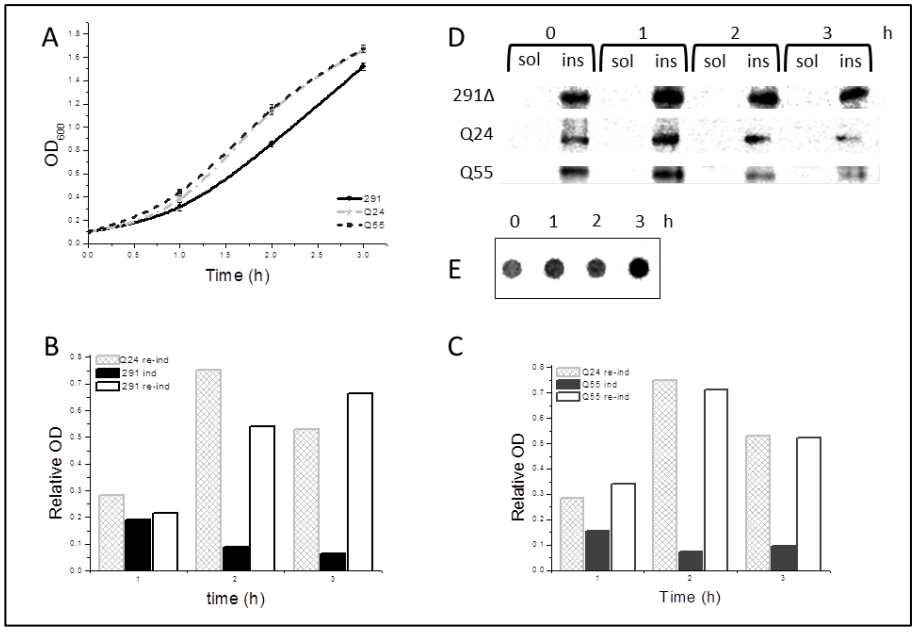

Fig. 5. Effect of insoluble aggregates on cytoxicity induced by the expression of AT3 variants. A) Growth curves at 37°C of *E. coli* strains expressing the three AT3 variants. After 6 h from the primary induction performed as reported in Fig. 1, cells were harvested and re-inoculated in a fresh medium in the presence of 500 μM IPTG. B, C**)** Growth rates of AT3-291Δ (B) and AT3-Q55 (C); bars represent growth rates under the primary induction (ind) or during the re-induction (re-ind). For comparison, the growth rate of AT3Q24 under re-induction is also shown. D) Soluble (sol) and insoluble (ins) fractions of strains expressing AT3 variants were collected at different times after re-induction, subjected to SDS-PAGE and western blotted using anti AT3-antibodies. E) Filter trap assay of protein insoluble fraction of AT3-Q55 at different times after re-induction.

DISCUSSION

Plenty of evidence accumulated in recent years has shown that several neurodegenerative diseases are associated with the accumulation of amyloid aggregates (89). Each disease normally results from the enhanced propensity of one specific protein to aggregate, although the proteins involved are unrelated to each other (89). Initially, this led to the conclusion that such macromolecular assemblies are the causative agents of the pathologies (133). However, compelling evidences achieved more recently points to the oligomeric species arising in the early stages of the fibrillogenesis as the culprits for cell toxicity (2-8). Thus, to provide insight into the mechanisms underlying toxicity, a thorough understanding of aggregation

pathways and structural features of the intermediates is required. Actually, *in vitro* investigations on fibrillogenesis of different amyloid proteins such as α-synuclein, β-amyloid, Htt, helped identify the species formed during the process (103,117,134)**.** However, the intracellular environment might greatly influence the aggregation mechanisms, as a result of high protein concentration and macromolecular crowding effect, presence of chaperones and proteases, constant supply of the polypeptides by the cellular translational complex. This implies that the aggregation intermediates isolated *in vitro* might not be representative of those generated intracellularly. Based on these considerations, not only studies *in vivo* are required to identify the cellular targets of the toxic species, but also to establish whether and how the aggregation pathway is affected by the intracellular environment.

E. coli has been long exploited as a useful cellular model to study aggregation of amyloid-related proteins in a physiological background and also for the screening of anti-amyloidogenic compounds (106-108,110,111,113,135-138). In particular, it was also employed to study the aggregation pattern of a polyQ chimera, which highlighted a multidomain misfolding resembling that demonstrated *in vitro* by the authentic polyQ proteins AT3 and Htt (113). However, investigations on artificial constructs are biased by the effect of polyQ-flanking sequences that in fact can deeply influence the aggregation pathway and, consequently, the toxicity exerted by the protein (126,127). This is why authentic polyQ proteins should be preferably assayed for such studies, although the use of artificial constructs may be justified by the requirement of a physically detectable tag to monitor aggregation.

Based on these premises, we assessed the toxic effects of authentic AT3-Q24 and AT3-Q55, i.e. a normal and expanded polyQ protein, respectively, by expressing them in the prokaryotic host. To the best of our knowledge, this is the first report of authentic full-length polyQ proteins expressed in an intracellular milieu, along with the analysis of their toxicity. To detect possible polyQ-unrelated effects, we also expressed a variant truncated immediately upstream of the polyQ itself (AT3-291Δ). The structural features of the intermediates were characterized in cell lysates by several biochemical and biophysical techniques. In particular, FTIR made it possible to detect the changes in secondary structure content and, most important, the increase in protein-protein interactions. Cellular toxicity was

quantified on the basis of the decrease in growth rates and correlated with the appearance of different aggregated species. Also, our results strongly suggest that the effect of the toxic oligomers mostly resulted in a bacteriostatic rather than bactericidal effect, as substantiated by cell inability to uptake propidium iodide at any time (Fig. S2).

When expressed in *E. coli*, expanded AT3 aggregated essentially in the same way as *in vitro* (62), namely undergoing a two-step fibrillogenesis process, the first giving rise to SDS-soluble oligomers, the second to SDS-insoluble fibrils found in the IBs. In FTIR, the latter displayed the typical pattern of side-chain glutamine hydrogen bonding, a hallmark of irreversible polyQ aggregation (62). AFM highlighted the fibrillar morphology of these aggregates, which elongate during the experimental timeframe, providing a direct representation of the polyQ fibrils formed. These findings are the first evidence of a multidomain misfolding of an authentic amyloid protein in an intracellular environment (139). As expected, AT3-Q24 and AT3-291Δ underwent the first step of aggregation, but did not generate any SDS-insoluble aggregate, as supported by filter trap assays, although aggregation products were found in the IBs as well.

Time courses of soluble protein showed almost identical profiles for AT3-Q24 and AT3-Q55, with a maximum around 2 h, followed by a relatively fast decline (Fig. 2). This result conforms to the "induced misfit" model (140,141) that assigns similar intrinsic aggregation rates to normal and expanded polyQ proteins, but assumes that the former is prevented from aggregating via interactions with its physiological partners, which of course cannot take place in the prokaryotic cell. In contrast, the level of soluble AT3-291Δ declined much more slowly than that of the two other forms until 5-6 h after induction, resulting in comparatively higher levels at the latest times (3-5 h). Interestingly, this matches well with the *in vitro* observation that AT3-291Δ (Santambrogio et al., in press[#]) undergoes slower aggregation kinetics with respect to those of AT3-Q24 and AT3-Q55 (62).

When quantifying the toxic effects of the three proteins, we demonstrated that, as expected, the expanded but not the normal AT3 dramatically reduced growth rate, suggestive of strong toxicity. Strikingly, we observed that the truncated variant was also toxic, which suggests a direct involvement of the polyQ-harboring context in protein toxicity. Recently, a work on an animal model demonstrated that mice, both homozygous and heterozygous for an AT3 variant

truncated at Gly259, developed severe motor coordination dysfunction and altered behavior, followed by premature death (142). The pathology was associated with extranuclear protein inclusions, massive neuronal cell death and altered ERAD response. Thus, despite the obvious differences between prokaryotic and eukaryotic cellular environments, truncated AT3 exhibits toxic effects in both systems, which suggests a conserved gain-of-function mechanism of toxicity triggered by proteolysis.

Growth rates of cells expressing AT3-291Δ or AT3-Q55 displayed a significant inverse correlation with the levels of early soluble oligomeric species. Remarkably, in the case of cells expressing AT3-Q55, not only was the cytotoxic effect apparent prior to the formation of SDS-insoluble, fibrillar aggregates, but also their later accumulation did not affect cell growth. These findings clearly point to the oligomeric species formed in the first step of aggregation process, but not to the final aggregates, as the culprits for cellular toxicity.

As above mentioned, the *wild-type* AT3-Q24 variant should be instead regarded as nontoxic, as its anyway marginal impact on growth rate could be fully accounted for by its massive overexpression. As a consequence, the soluble species arising from its aggregation must be devoid of any detrimental effect. One must therefore infer that the aggregation pathway of AT3-Q24 leads to the formation of soluble oligomeric intermediates structurally different from those generated by AT3-Q55 and AT3-291Δ. As far as AT3-Q55 is concerned, it is likely that the expanded polyQ plays a major role in shaping the soluble oligomeric species into somehow cytotoxic conformations. At the present stage of our investigations we have no obvious explanation to account for these major differences. Actually, we analyzed soluble aggregates in native gradient gel electrophoresis (data not shown), and observed significant discrepancies in the electrophoretic patterns of the oligomers resulting from the three variants. Although this substantiates the idea that each of the three AT3 forms undergoes its own aggregation pathway, an interpretation of the mechanisms of toxicity can be hardly conceived based on the sole molecular mass of the aggregates detected. This will be therefore a subject of our further investigations. In any case, these findings provide an initial clue to the issue of the lack of toxicity of normal AT3, despite its capability to give rise to soluble oligomers (118,143): based on our results, this should be related to their intrinsic structural

properties rather than to different intracellular abundances of the oligomers generated by the normal, compared to the expanded form.

No less interestingly, we observed that the final aggregates could largely prevent cellular toxicity resulting from soluble oligomer accumulation, irrespective of whether they were SDS-insoluble (AT3-Q55) or not (AT3-291Δ). This is clearly supported by an experiment, whereby the two variants were re-induced after the primary induction, a condition that leads to protein new synthesis in the presence of insoluble aggregates (Fig. 5). This result adds direct evidence to the controversial debate on the protective role of intracellular inclusions in polyQ diseases (7-15). In any case, our report makes it possible to rule out that the culprits for cell toxicity in the early growth stages are IBs, as instead recently shown by Ventura and coworkers in investigations on Aβ42. They demonstrated that IBs formed in *E. coli* by this peptide are toxic to mammalian cells in culture, due to a massive presence of oligomeric structures as opposed to prefibrillar species rich in intermolecular β-sheets interactions (108). However, IBs generated by AT3-291Δ and AT3-Q55 (Fig. 3B and S4) display a similar secondary structure content and intermolecular β-sheet interactions throughout the experimental time-course, which suggests that changes in cell growth rates are not caused by conformational changes occurring in IBs.

In conclusion, we show that, also in *E. coli*, the expanded and truncated but not the normal AT3 exert cytotoxic effects, in line with plenty of reports on polyQ protein toxicity in higher organisms, included humans (126,142,144,145). In our system, we could also reproduce the intracellular two-step aggregation pathway of the expanded AT3 and detect the toxic effect of the oligomeric species. Indeed these findings suggest that the mechanisms of toxicity are essentially conserved throughout the evolutionary tree. Also, they show that *E. coli*, thanks to its ease of handling, is a valuable model organism for studying amyloid aggregation in the intracellular environment using authentic polyQ proteins.

Acknowledgements
We thank Alessandra Gliozzi and Elena Papaleo for helpful discussion.

Footnotes

† Authors contributed equally to this work

[+]This work was supported by the following: Fondazione Cariplo (N.O.B.E.L. project); Italian Ministry for Instruction, University and Research, PRIN 2007 project n. 2007XY59ZJ_004; Regione Lombardia ("Network Tecnologico integrato per lo studio proteomico e trascrittomico di malattie neurodegenerative correlate a deposizioni di amiloidi"); Fondo Ateneo per la Ricerca of the University of Milano-Bicocca. A.N. and G.I. acknowledge postdoctoral fellowships of the University of Milano-Bicocca. This work was partially supported by University of Genoa (Progetti di Ateneo).

"The role of the central flexible region on the aggregation and conformational properties of human ataxin-3" Santambrogio, Carlo; Frana, Anna; Natalello, Antonino; Papaleo, Elena; Regonesi, Maria; Doglia, Silvia; Tortora, Paolo; Invernizzi, Gaetano; Grandori, Rita; FEBS J, in press

The abbreviations used are: AT3, ataxin-3; polyQ, polyglutamine; AFM, atomic force microscopy; FTIR, fourier infra red spectroscopy; IPTG, Isopropil β-D-1-tiogalattopiranoside; OD, optical density; IBs, inclusion bodies; Htt, huntigtine;

¥ Present address: King's College London, Muscle Signalling and Development, The Randall Division of Cell and Molecular Biophysics and The Cardiovascular Division, London, UK

SUPPLEMENTARY INFORMATION

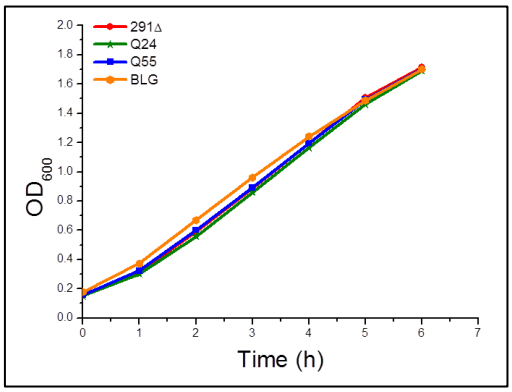

Fig. S1. Comparison of AT3 expression strains under non-inducing conditions. Growth curves at 37 °C of AT3-291Δ (red) -Q24 (green), -Q55 (blue) and BLG (yellow) expression strain.

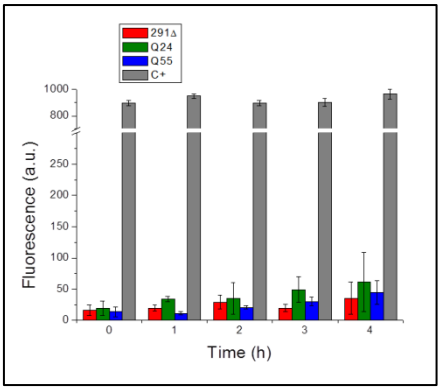

Fig. S2. Propidium iodide staining of AT3-expression strains. Fluorescence emission of propidium iodide in strains expressing the three AT3 variants and in cells treated with 10% isopropanol as a positive control (C+). Error bars are derived from at least 5 independent experiments.

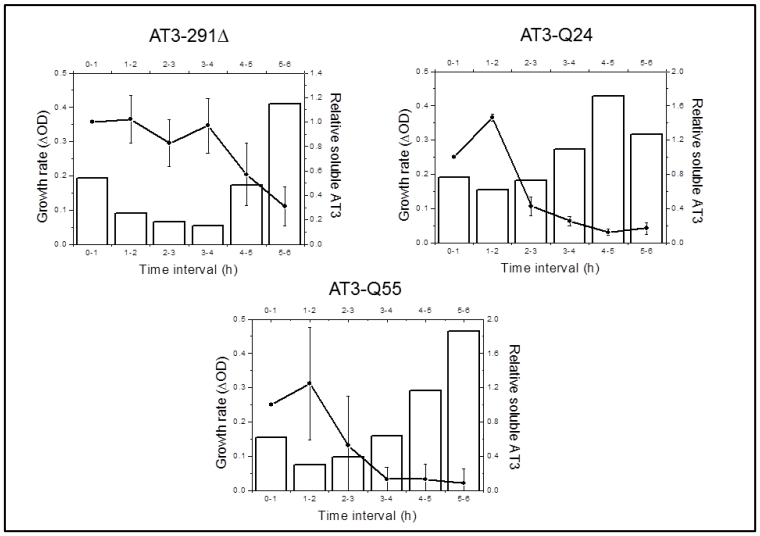

Fig. S3. Comparison between soluble amounts and growth rates of AT3 variants. Growth rates of AT3 expression strains (lines) were plotted with relative amounts of soluble protein fraction produced (columns).

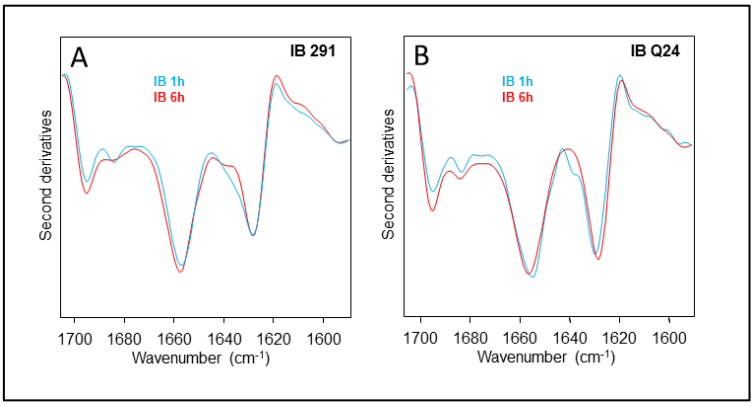

Fig. S4. . FTIR characterization of AT3-291Δ and -Q24 initial and mature aggregates. Second derivative FTIR spectra of the insoluble fractions collected at 1 and 6 h after induction of the AT3-291Δ (A) and -Q24 (B) expression strains. Spectra are normalized at the ~1515 cm-1 peak of tyrosine.

3.5 References

1. **Richards, R. I., and Sutherland, G. R. (1997)** ***Trends Biochem Sci*** **22, 432-436**
2. **Gatchel, J. R., and Zoghbi, H. Y. (2005)** ***Nat Rev Genet*** **6, 743-755**

3. **Shao, J., and Diamond, M. I. (2007) *Hum Mol Genet* 16 Spec No. 2, R115-123**
4. **Zoghbi, H. Y., and Orr, H. T. (2000) *Annu Rev Neurosci* 23, 217-247**
5. **Dürr, A., Dodé, C., Hahn, V., Pêcheux, C., Pillon, B., Feingold, J., Kaplan, J. C., Agid, Y., and Brice, A. (1995) *J Neurol Sci* 129, 51-55**
6. **Maciel, P., Gaspar, C., DeStefano, A. L., Silveira, I., Coutinho, P., Radvany, J., Dawson, D. M., Sudarsky, L., Guimarães, J., and Loureiro, J. E. (1995) *Am J Hum Genet* 57, 54-61**
7. **Ranum, L. P., Lundgren, J. K., Schut, L. J., Ahrens, M. J., Perlman, S., Aita, J., Bird, T. D., Gomez, C., and Orr, H. T. (1995) *Am J Hum Genet* 57, 603-608**
8. **Riess, O., Rüb, U., Pastore, A., Bauer, P., and Schöls, L. (2008) *Cerebellum* 7, 125-137**
9. **Gusella, J. F., and MacDonald, M. E. (2000) *Nat Rev Neurosci* 1, 109-115**
10. **Rochet, J. C., and Lansbury, P. T. (2000) *Curr Opin Struct Biol* 10, 60-68**
11. **Kelly, J. W. (2000) *Nat Struct Biol* 7, 824-826**
12. **Sakahira, H., Breuer, P., Hayer-Hartl, M. K., and Hartl, F. U. (2002) *Proc Natl Acad Sci U S A* 99 Suppl 4, 16412-16418**
13. **Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) *Nature* 431, 805-810**
14. **Ross, C. A., and Poirier, M. A. (2004) *Nat Med* 10 Suppl, S10-17**
15. **Slow, E. J., Graham, R. K., Osmand, A. P., Devon, R. S., Lu, G., Deng, Y., Pearson, J., Vaid, K., Bissada, N., Wetzel, R., Leavitt, B. R., and Hayden, M. R. (2005) *Proc Natl Acad Sci U S A* 102, 11402-11407**
16. **Costa, M. C., Gomes-da-Silva, J., Miranda, C. J., Sequeiros, J., Santos, M. M., and Maciel, P. (2004) *Genomics* 84, 361-373**
17. **Ichikawa, Y., Goto, J., Hattori, M., Toyoda, A., Ishii, K., Jeong, S. Y., Hashida, H., Masuda, N., Ogata, K., Kasai, F., Hirai, M., Maciel, P., Rouleau, G. A., Sakaki, Y., and Kanazawa, I. (2001) *J Hum Genet* 46, 413-422**
18. **Paulson, H. L., Das, S. S., Crino, P. B., Perez, M. K., Patel, S. C., Gotsdiner, D., Fischbeck, K. H., and Pittman, R. N. (1997) *Ann Neurol* 41, 453-462**
19. **Schmidt, T., Landwehrmeyer, G. B., Schmitt, I., Trottier, Y., Auburger, G., Laccone, F., Klockgether, T., Völpel, M., Epplen, J. T., Schöls, L., and Riess, O. (1998) *Brain Pathol* 8, 669-679**
20. **Trottier, Y., Cancel, G., An-Gourfinkel, I., Lutz, Y., Weber, C., Brice, A., Hirsch, E., and Mandel, J. L. (1998) *Neurobiol Dis* 5, 335-347**

21. **Masino, L., Musi, V., Menon, R. P., Fusi, P., Kelly, G., Frenkiel, T. A., Trottier, Y., and Pastore, A. (2003)** ***FEBS Lett*** **549, 21-25**
22. **Albrecht, M., Golatta, M., Wüllner, U., and Lengauer, T. (2004)** ***Eur J Biochem*** **271, 3155-3170**
23. **Burnett, B., Li, F., and Pittman, R. N. (2003)** ***Hum Mol Genet*** **12, 3195-3205**
24. **Albrecht, M., Hoffmann, D., Evert, B. O., Schmitt, I., Wüllner, U., and Lengauer, T. (2003)** ***Proteins*** **50, 355-370**
25. **Scheel, H., Tomiuk, S., and Hofmann, K. (2003)** ***Hum Mol Genet*** **12, 2845-2852**
26. **Mao, Y., Senic-Matuglia, F., Di Fiore, P. P., Polo, S., Hodsdon, M. E., and De Camilli, P. (2005)** ***Proc Natl Acad Sci U S A*** **102, 12700-12705**
27. **Nicastro, G., Menon, R. P., Masino, L., Knowles, P. P., McDonald, N. Q., and Pastore, A. (2005)** ***Proc Natl Acad Sci U S A*** **102, 10493-10498**
28. **Nicastro, G., Masino, L., Esposito, V., Menon, R. P., De Simone, A., Fraternali, F., and Pastore, A. (2009)** ***Biopolymers*** **91, 1203-1214**
29. **Goto, J., Watanabe, M., Ichikawa, Y., Yee, S. B., Ihara, N., Endo, K., Igarashi, S., Takiyama, Y., Gaspar, C., Maciel, P., Tsuji, S., Rouleau, G. A., and Kanazawa, I. (1997)** ***Neurosci Res*** **28, 373-377**
30. **Harris, G. M., Dodelzon, K., Gong, L., Gonzalez-Alegre, P., and Paulson, H. L. (2010)** ***PLoS One*** **5, e13695**
31. **Bettencourt, C., Santos, C., Montiel, R., Costa, M. o. C., Cruz-Morales, P., Santos, L. R., Simões, N., Kay, T., Vasconcelos, J., Maciel, P., and Lima, M. (2010)** ***Neurogenetics*** **11, 193-202**
32. **Matos, C. A., de Macedo-Ribeiro, S., and Carvalho, A. L. (2011)** ***Prog Neurobiol*** **95, 26-48**
33. **Goldberg, A. L. (2003)** ***Nature*** **426, 895-899**
34. **Berke, S. J., Chai, Y., Marrs, G. L., Wen, H., and Paulson, H. L. (2005)** ***J Biol Chem*** **280, 32026-32034**
35. **Chai, Y., Berke, S. S., Cohen, R. E., and Paulson, H. L. (2004)** ***J Biol Chem*** **279, 3605-3611**
36. **Donaldson, K. M., Li, W., Ching, K. A., Batalov, S., Tsai, C. C., and Joazeiro, C. A. (2003)** ***Proc Natl Acad Sci U S A*** **100, 8892-8897**
37. **Winborn, B. J., Travis, S. M., Todi, S. V., Scaglione, K. M., Xu, P., Williams, A. J., Cohen, R. E., Peng, J., and Paulson, H. L. (2008)** ***J Biol Chem*** **283, 26436-26443**
38. **Burnett, B. G., and Pittman, R. N. (2005)** ***Proc Natl Acad Sci U S A*** **102, 4330-4335**
39. **Kuhlbrodt, K., Janiesch, P. C., Kevei, É., Segref, A., Barikbin, R., and Hoppe, T. (2011)** ***Nat Cell Biol*** **13, 273-281**

40. **Nicastro, G., Todi, S. V., Karaca, E., Bonvin, A. M., Paulson, H. L., and Pastore, A. (2010)** ***PLoS One*** **5, e12430**

41. **Boeddrich, A., Gaumer, S., Haacke, A., Tzvetkov, N., Albrecht, M., Evert, B. O., Müller, E. C., Lurz, R., Breuer, P., Schugardt, N., Plassmann, S., Xu, K., Warrick, J. M., Suopanki, J., Wüllner, U., Frank, R., Hartl, U. F., Bonini, N. M., and Wanker, E. E. (2006)** ***EMBO J*** **25, 1547-1558**

42. **Hirabayashi, M., Inoue, K., Tanaka, K., Nakadate, K., Ohsawa, Y., Kamei, Y., Popiel, A. H., Sinohara, A., Iwamatsu, A., Kimura, Y., Uchiyama, Y., Hori, S., and Kakizuka, A. (2001)** ***Cell Death Differ*** **8, 977-984**

43. **Matsumoto, M., Yada, M., Hatakeyama, S., Ishimoto, H., Tanimura, T., Tsuji, S., Kakizuka, A., Kitagawa, M., and Nakayama, K. I. (2004)** ***EMBO J*** **23, 659-669**

44. **Zhong, X., and Pittman, R. N. (2006)** ***Hum Mol Genet*** **15, 2409-2420**

45. **Wang, G., Sawai, N., Kotliarova, S., Kanazawa, I., and Nukina, N. (2000)** ***Hum Mol Genet*** **9, 1795-1803**

46. **Li, F., Macfarlan, T., Pittman, R. N., and Chakravarti, D. (2002)** ***J Biol Chem*** **277, 45004-45012**

47. **Evert, B. O., Araujo, J., Vieira-Saecker, A. M., de Vos, R. A., Harendza, S., Klockgether, T., and Wüllner, U. (2006)** ***J Neurosci*** **26, 11474-11486**

48. **Rodrigues, A. J., Coppola, G., Santos, C., Costa, M. o. C., Ailion, M., Sequeiros, J., Geschwind, D. H., and Maciel, P. (2007)** ***FASEB J*** **21, 1126-1136**

49. **Mazzucchelli, S., De Palma, A., Riva, M., D'Urzo, A., Pozzi, C., Pastori, V., Comelli, F., Fusi, P., Vanoni, M., Tortora, P., Mauri, P., and Regonesi, M. E. (2009)** ***Int J Biochem Cell Biol*** **41, 2485-2492**

50. **Rodrigues, A. J., do Carmo Costa, M., Silva, T. L., Ferreira, D., Bajanca, F., Logarinho, E., and Maciel, P. (2010)** ***Biochim Biophys Acta*** **1803, 1154-1163**

51. **Seidel, K., den Dunnen, W. F., Schultz, C., Paulson, H., Frank, S., de Vos, R. A., Brunt, E. R., Deller, T., Kampinga, H. H., and Rüb, U. (2010)** ***Acta Neuropathol*** **120, 449-460**

52. **Bevivino, A. E., and Loll, P. J. (2001)** ***Proc Natl Acad Sci U S A*** **98, 11955-11960**

53. **Chow, M. K., Ellisdon, A. M., Cabrita, L. D., and Bottomley, S. P. (2004)** ***J Biol Chem*** **279, 47643-47651**

54. **Jana, N. R., and Nukina, N. (2004)** ***Neurotox Res*** **6, 523-533**

55. **Muchowski, P. J., and Wacker, J. L. (2005)** ***Nat Rev Neurosci*** **6, 11-22**

56. **Nagai, Y., Inui, T., Popiel, H. A., Fujikake, N., Hasegawa, K., Urade, Y., Goto, Y., Naiki, H., and Toda, T. (2007)** ***Nat Struct Mol Biol*** **14, 332-340**

57. **Tanaka, M., Morishima, I., Akagi, T., Hashikawa, T., and Nukina, N. (2001)** ***J Biol Chem*** **276, 45470-45475**

58. **Williams, A. J., and Paulson, H. L. (2008)** ***Trends Neurosci*** **31, 521-528**

59. **Uversky, V. N. (2010)** ***FEBS J*** **277, 2940-2953**

60. **Gales, L., Cortes, L., Almeida, C., Melo, C. V., Costa, M. C., Maciel, P., Clarke, D. T., Damas, A. M., and Macedo-Ribeiro, S. (2005)** ***J Mol Biol*** **353, 642-654**

61. **Ellisdon, A. M., Thomas, B., and Bottomley, S. P. (2006)** ***J Biol Chem*** **281, 16888-16896**

62. **Natalello, A., Frana, A. M., Relini, A., Apicella, A., Invernizzi, G., Casari, C., Gliozzi, A., Doglia, S. M., Tortora, P., and Regonesi, M. E. (2011)** ***PLoS One*** **6, e18789**

63. **Ellisdon, A. M., Pearce, M. C., and Bottomley, S. P. (2007)** ***J Mol Biol*** **368, 595-605**

64. **Masino, L., Nicastro, G., Menon, R. P., Dal Piaz, F., Calder, L., and Pastore, A. (2004)** ***J Mol Biol*** **344, 1021-1035**

65. **Saunders, H. M., and Bottomley, S. P. (2009)** ***Protein Eng Des Sel*** **22, 447-451**

66. **Masino, L., Nicastro, G., Calder, L., Vendruscolo, M., and Pastore, A. (2011)** ***FASEB J*** **25, 45-54**

67. **Fiumara, F., Fioriti, L., Kandel, E. R., and Hendrickson, W. A. (2010)** ***Cell*** **143, 1121-1135**

68. **Song, A. X., Zhou, C. J., Peng, Y., Gao, X. C., Zhou, Z. R., Fu, Q. S., Hong, J., Lin, D. H., and Hu, H. Y. (2010)** ***PLoS One*** **5, e13202**

69. **Bellotti, V., and Chiti, F. (2008)** ***Curr Opin Struct Biol*** **18, 771-779**

70. **Ellis, R. J., and Minton, A. P. (2006)** ***Biol Chem*** **387, 485-497**

71. **Zhou, H. X. (2008)** ***Arch Biochem Biophys*** **469, 76-82**

72. **Tam, S., Geller, R., Spiess, C., and Frydman, J. (2006)** ***Nat Cell Biol*** **8, 1155-1162**

73. **Rousseau, E., Kojima, R., Hoffner, G., Djian, P., and Bertolotti, A. (2009)** ***J Biol Chem*** **284, 1917-1929**

74. **Carrió, M., González-Montalbán, N., Vera, A., Villaverde, A., and Ventura, S. (2005)** ***J Mol Biol*** **347, 1025-1037**

75. **Morell, M., Bravo, R., Espargaró, A., Sisquella, X., Avilés, F. X., Fernàndez-Busquets, X., and Ventura, S. (2008)** ***Biochim Biophys Acta*** **1783, 1815-1825**

76. **Sabaté, R., Espargaró, A., Saupe, S. J., and Ventura, S. (2009)** ***Microb Cell Fact*** **8, 56**

77. **Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D., and Riek, R. (2008) *PLoS Biol* 6, e195**
78. **Wasmer, C., Benkemoun, L., Sabaté, R., Steinmetz, M. O., Coulary-Salin, B., Wang, L., Riek, R., Saupe, S. J., and Meier, B. H. (2009) *Angew Chem Int Ed Engl* 48, 4858-4860**
79. **de Groot, N. S., Sabate, R., and Ventura, S. (2009) *Trends Biochem Sci* 34, 408-416**
80. **Dasari, M., Espargaro, A., Sabate, R., Lopez del Amo, J. M., Fink, U., Grelle, G., Bieschke, J., Ventura, S., and Reif, B. (2011) *Chembiochem* 12, 407-423**
81. **García-Fruitós, E., Vazquez, E., Gonzalez-Montalbán, N., Ferrer-Miralles, N., and Villaverde, A. (2011) *Curr Pharm Biotechnol***
82. **Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J., and Taddei, F. (2008) *Proc Natl Acad Sci U S A* 105, 3076-3081**
83. **Woulfe, J. (2008) *Biochim Biophys Acta* 1783, 2195-2206**
84. **de Groot, N. S., Espargaró, A., Morell, M., and Ventura, S. (2008) *Future Microbiol* 3, 423-435**
85. **Vera, A., González-Montalbán, N., Arís, A., and Villaverde, A. (2007) *Biotechnol Bioeng* 96, 1101-1106**
86. **Espargaró, A., Sabaté, R., and Ventura, S. (2008) *FEBS Lett* 582, 3669-3673**
87. **Castillo, V., Espargaró, A., Gordo, V., Vendrell, J., and Ventura, S. (2010) *Proteomics* 10, 4172-4185**
88. **Ignatova, Z. (2005) *Microb Cell Fact* 4, 23**
89. **Chiti, F., and Dobson, C. M. (2006) *Annu Rev Biochem* 75, 333-366**
90. **Campioni, S., Mannini, B., Zampagni, M., Pensalfini, A., Parrini, C., Evangelisti, E., Relini, A., Stefani, M., Dobson, C. M., Cecchi, C., and Chiti, F. (2010) *Nat Chem Biol* 6, 140-147**
91. **Baglioni, S., Casamenti, F., Bucciantini, M., Luheshi, L. M., Taddei, N., Chiti, F., Dobson, C. M., and Stefani, M. (2006) *J Neurosci* 26, 8160-8167**
92. **Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) *Nature* 416, 507--511**
93. **Tomic, J. L., Pensalfini, A., Head, E., and Glabe, C. G. (2009) *Neurobiol Dis* 35, 352--358**
94. **Takahashi, T., Kikuchi, S., Katada, S., Nagai, Y., Nishizawa, M., and Onodera, O. (2008) *Hum Mol Genet* 17, 345-356**
95. **Olshina, M. A., Angley, L. M., Ramdzan, Y. M., Tang, J., Bailey, M. F., Hill, A. F., and Hatters, D. M. (2010) *J Biol Chem* 285, 21807-21816**
96. **Lajoie, P., and Snapp, E. L. (2010) *PLoS One* 5, e15245**

97. **Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) *Nature* 431, 805-810**

98. **Miller, J., Arrasate, M., Shaby, B. A., Mitra, S., Masliah, E., and Finkbeiner, S. (2010) *J Neurosci* 30, 10541-10550**

99. **Sanchez, L., Madurga, S., Pukala, T., Vilaseca, M., Lopez-Iglesias, C., Robinson, C. V., Giralt, E., and Carulla, N. (2011) *Journal of the American Chemical Society* 133, 6505-6508**

100. **Xue, W. F., Hellewell, A. L., Gosal, W. S., Homans, S. W., Hewitt, E. W., and Radford, S. E. (2009) *Journal of Biological Chemistry* 284, 34272-34282**

101. **Koffie, R. M., Meyer-Luehmann, M., Hashimoto, T., Adams, K. W., Mielke, M. L., Garcia-Alloza, M., Micheva, K. D., Smith, S. J., Kim, M. L., Lee, V. M., Hyman, B. T., and Spires-Jones, T. L. (2009) *Proceedings of the National Academy of Sciences of the United States of America* 106, 4012-4017**

102. **Treusch, S., Cyr, D. M., and Lindquist, S. (2009) *Cell Cycle* 8, 1668-1674**

103. **Haass, C., and Selkoe, D. J. (2007) *Nature Reviews Molecular Cell Biology* 8, 101-112**

104. **Rujano, M. A., Bosveld, F., Salomons, F. A., Dijk, F., van Waarde, M. A. W. H., van der Want, J. J. L., de Vos, R. A. I., Brunt, E. R., Sibon, O. C. M., and Kampinga, H. H. (2006) *PLoS Biol* 4, e417**

105. **Tyedmers, J., Mogk, A., and Bukau, B. (2010) *Nature Reviews Molecular Cell Biology* 11, 777-788**

106. **Carrió, M., González-Montalbán, N., Vera, A., Villaverde, A., and Ventura, S. (2005) *J Mol Biol* 347, 1025-1037**

107. **Morell, M., Bravo, R., Espargaró, A., Sisquella, X., Avilés, F. X., Fernàndez-Busquets, X., and Ventura, S. (2008) *Biochim Biophys Acta* 1783, 1815-1825**

108. **Dasari, M., Espargaro, A., Sabate, R., Amo, J. M. L. D., Fink, U., Grelle, G., Bieschke, J., Ventura, S., and Reif, B. (2011) *Chembiochem***

109. **Wasmer, C., Benkemoun, L., Sabaté, R., Steinmetz, M. O., Coulary-Salin, B., Wang, L., Riek, R., Saupe, S. J., and Meier, B. H. (2009) *Angew Chem Int Ed Engl* 48, 4858-4860**

110. **de Groot, N. S., Sabate, R., and Ventura, S. (2009) *Trends Biochem Sci* 34, 408-416**

111. **García-Fruitós, E., Sabate, R., de Groot, N. S., Villaverde, A., and Ventura, S. (2011) *FEBS J***

112. **de Groot, N. S., Espargaró, A., Morell, M., and Ventura, S. (2008) *Future Microbiol* 3, 423-435**

113. **Ignatova, Z., Thakur, A. K., Wetzel, R., and Gierasch, L. M. (2007) *J Biol Chem* 282, 36736-36743**
114. **Masino, L., and Pastore, A. (2002) *Biochem Soc Trans* 30, 548-551**
115. **Zoghbi, H. Y., and Orr, H. T. (2000) *Annu Rev Neurosci* 23, 217--247**
116. **Ellisdon, A. M., Thomas, B., and Bottomley, S. P. (2006) *J Biol Chem* 281, 16888-16896**
117. **Thakur, A. K., Jayaraman, M., Mishra, R., Thakur, M., Chellgren, V. M., Byeon, I.-J. L., Anjum, D. H., Kodali, R., Creamer, T. P., Conway, J. F., Gronenborn, A. M., and Wetzel, R. (2009) *Nat Struct Mol Biol* 16, 380--389**
118. **Ellisdon, A. M., Pearce, M. C., and Bottomley, S. P. (2007) *J Mol Biol* 368, 595--605**
119. **Sikorski, P., and Atkins, E. (2005) *Biomacromolecules* 6, 425--432**
120. **Schneider, R., Schumacher, M. C., Mueller, H., Nand, D., Klaukien, V., Heise, H., Riedel, D., Wolf, G., Behrmann, E., Raunser, S., Seidel, R., Engelhard, M., and Baldus, M. (2011) *J Mol Biol***
121. **Onodera, O., Roses, A. D., Tsuji, S., Vance, J. M., Strittmatter, W. J., and Burke, J. R. (1996) *FEBS Lett* 399, 135-139**
122. **Nagao, Y., Ishiguro, H., and Nukina, N. (2000) *Biochim Biophys Acta* 1502, 247-256**
123. **Shehi, E., Fusi, P., Secundo, F., Pozzuolo, S., Bairati, A., and Tortora, P. (2003) *Biochemistry* 42, 14626--14632**
124. **Hazeki, N., Tukamoto, T., Goto, J., and Kanazawa, I. (2000) *Biochem Biophys Res Commun* 277, 386-393**
125. **Ami, D., Natalello, A., Taylor, G., Tonon, G., and Doglia, S. M. (2006) *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1764, 793-799**
126. **Duennwald, M. L., Jagadish, S., Muchowski, P. J., and Lindquist, S. (2006) *Proceedings of the National Academy of Sciences of the United States of America* 103, 11045-11050**
127. **Nozaki, K., Onodera, O., Takano, H., and Tsuji, S. (2001) *Neuroreport* 12, 3357--3364**
128. **Invernizzi, G., Annoni, E., Natalello, A., Doglia, S. M., and Lotti, M. (2008) *Protein Expression and Purification* 62, 111-115**
129. **Hoffmann, F., and Rinas, U. (2004) *Advances in biochemical engineering/biotechnology* 89, 73-92**
130. **Hinz, J., Gierasch, L. M., and Ignatova, Z. (2008) *Biochemistry* 47, 4196-4200**
131. **Ami, D., Bonecchi, L., Cali, S., Orsini, G., Tonon, G., and Doglia, S. M. (2003) *Biochimica Et Biophysica Acta-General Subjects* 1624, 6-10**

132. **Ami, D., Natalello, A., Schultz, T., Gatti-Lafranconi, P., Lotti, M., Doglia, S. M., and de Marco, A. (2009)** ***Biochim Biophys Acta*** **1794, 263-269**

133. **Dobson, C. M. (2006)** ***Protein Pept Lett*** **13, 219-227**

134. **Glabe, C. G. (2008)** ***J Biol Chem*** **283, 29639--29643**

135. **Ignatova, Z., and Gierasch, L. M. (2006)** ***J Biol Chem*** **281, 12959--12967**

136. **Wang, X., and Chapman, M. R. (2008)** ***J Mol Biol*** **380, 570-580**

137. **Winkelmann, J., Calloni, G., Campioni, S., Mannini, B., Taddei, N., and Chiti, F. (2010)** ***J Mol Biol*** **398, 600-613**

138. **Dolado, I., Nieto, J., Saraiva, M. J. M., Arsequell, G., Valencia, G., and Planas, A. (2005)** ***Journal of Combinatorial Chemistry*** **7, 246-252**

139. **Saunders, H. M., and Bottomley, S. P. (2009)** ***Protein Eng Des Sel*** **22, 447--451**

140. **Masino, L., Nicastro, G., Menon, R. P., Piaz, F. D., Calder, L., and Pastore, A. (2004)** ***J Mol Biol*** **344, 1021--1035**

141. **Masino, L., Nicastro, G., Calder, L., Vendruscolo, M., and Pastore, A. (2011)** ***FASEB J*** **25, 45-54**

142. **Hübener, J., Vauti, F., Funke, C., Wolburg, H., Ye, Y., Schmidt, T., Wolburg-Buchholz, K., Schmitt, I., Gardyan, A., Drießen, S., Arnold, H.-H., Nguyen, H. P., and Riess, O. (2011)** ***Brain***

143. **Chow, M. K. M., Ellisdon, A. M., Cabrita, L. D., and Bottomley, S. P. (2004)** ***J Biol Chem*** **279, 47643--47651**

144. **Bates, G. (2003)** ***Lancet*** **361, 1642-1644**

145. **Parker, J. A., Connolly, J. B., Wellington, C., Hayden, M., Dausset, J., and Neri, C. (2001)** ***Proceedings of the National Academy of Sciences of the United States of America*** **98, 13318-13323**

4. Future prospects

4.1 Hsp70-mediated inhibition of polyQ aggregation

Plenty of studies demonstrate a strong correlation between the activity of Hsp70 homologues and the severity/progression of polyQ diseases. Indeed, Hsp70 family members have been detected in polyQ amyloid aggregates of diseased brains. Moreover, it has been observed that the deficit of activity of Hsp70 represents a factor of risk for the onset of these pathologies.

In order to obtain the molecular details of the inhibition exerted by Hsp70 on polyQ-induced neurodegeneration, several animal systems have been employed. Many reports on fly and mouse models show that that the overexpression of Hsp70/Hsp40 chaperone system can suppress polyQ neurotoxicity. For example, Warrick and coworkers demonstrated in a fly model of SCA3 that the overexpression of Hsp70 strongly reduces the toxicity associated to AT3 aggregated species (1). Similar results have been also obtained with SCA1 and SBMA mice (2,3). Surprisingly, in all these cases the overexpression of Hsp70/Hsp40 system does not prevent the formation of polyQ aggregates. Further, HD mice (4,5) and a SCA3 fly model (6) showed that the cytosolic level of Hsp70 decreases along with the progression of these diseases. Finally, contradictory but still important results have been obtained in other HD mouse models: despite only modest effects on pathology onset were observed as a consequence of Hsp70 overexpression (7), the absence of even one allele of the hsp70 gene significantly exacerbated the severity of the pathological phenotype (8). The demonstration that Hsp70, either endogenous or overexpressed, is an integral component of the *in vivo* physiological response to polyQ disease highlights the importance of this chaperone in view of its potential use in management of neurodegenerative disorders. Nevertheless it is not still clear if Hsp70s effects always mediated by direct effects on polyQ assembly or through more general buffering of prosurvival signalling (9).

Despite the compelling *in vivo* evidence of the implications of Hsp70 in polyQ diseases, *in vitro* studies of the molecular mechanisms of the Hsp70-mediated inhibition of polyQ amyloid formation are relatively sparse. In particular, several works focused on the effect of Hsp70 on Htt aggregation. Purified mammalian Hsp70 and Hsp40 suppressed, in an ATP-dependent manner, the *in vitro* assembly of polyQ-expanded Htt exon 1 constructs into ordered, SDS-

insoluble amyloid fibrils (10). Importantly, these chaperones are only active when added during the lag phase of the aggregation reaction, which suggests that Hsp70 and Hsp40 preferably act on early, prefibrillar states (10). Further, it is interesting to note that several cochaperones, including J-domain proteins, CHIP, and BAG, have dramatic effects on polyQ aggregate formation either on their own or in concert with Hsp70 (10-12).

In order to evaluate the effect of Hsp70 on the aggregation of polyQ proteins, we decided to employ AT3 as a model, having the unique opportunity to study the effect of this chaperone on a full length polyQ protein. Therefore we planned to characterize the anti-amyloidogenic activity of Hsp70 *in vitro* and *in vivo*, performing co-aggregation experiments in *E. coli*, with AT3 protein variants, as previously discussed.

We recently obtained preliminary data that clearly show that Hsp70 is able to inhibit *in vitro* in a nucleotide-independent manner the aggregation of a pathological expanded variant of AT3 (AT3-Q55). Indeed, analysing the protein soluble fraction of AT3 alone or in the presence of an equimolar concentration of FL-Hsp70, it is possible to observe a decrease in the aggregation rate of the protein (Table 1) and the appearance of a lag phase of approximately 5 h in the presence of the chaperone (Fig. 1A). Interestingly the addition of the chaperone led to a drastic reduction in the formation of SDS-insoluble species (Fig. 1B), as already observed for Htt exon 1 and Hsp70 in the presence of ATP (10). Moreover, it is interesting to notice that the insoluble protein fraction of AT3 in the presence of Hsp70 follows a kinetic specular to that of Hsp70, suggesting that the chaperone is able to interact with AT3 as long as present in solution. These facts demonstrate that Hsp70 is able to affect AT3 aggregation independently from its nucleotide regulation. Further, these data strongly suggest that the mechanism exerted by the chaperone in the nucleotide –free and –bound state on polyQ amyloid formation may be similar.

As extensively discussed above, it has been demonstrated that expanded AT3 fibrillogenesis process follows a two-step mechanism (13): in the first step, monomeric form of AT3 misfolds and then oligomeric species form as a consequence of the aberrant interaction between JDs. During the second step, these species evolve to mature amyloid fibrils, characterized by hydrogen bonds between the side-chains of the glutamines of the polyQ tract. The presence of an expanded polyQ tract is fundamental for the formation of these

interactions and it has been demonstrate to positively influence the first step by accelerating it.

Considering this fact, we considered the possibility that the inhibition exerted by Hsp70 on AT3 fibril formation may be related to an effect on the first step of the aggregation process. Therefore we performed soluble fraction analysis also on the aggregation of the isolated JD (AT3/182Δ) alone or with an equimolar concentration of nucleotide-free FL-Hsp70 (Fig. 1C). Interestingly FL-Hsp70 showed a clear effect also on the aggregation of this protein, suppressing by 60% the formation of insoluble protein species after 70 h of aggregation (Fig. 1C).

Our results demonstrate that Hsp70 is able to affect the amyloid aggregation process of AT3, also including the first JD-dependent step of the aggregation. Future experiments will be aimed to obtain a detailed characterization of the different aggregated species formed during the AT3 fibril formation process in the presence on the chaperone. To this end we will employ ensemble average and single molecule biophysical techniques. Further, we will investigate the role of the nucleotide regulation cycle in the anti-amyloidogenic activity of Hsp70. Since we successfully employed *Escherichia coli* as an *in vivo* model for studying aggregation mechanism of authentic variants of AT3, our future efforts will be also aimed to obtain *in vivo* evidences of Hsp70 action by analyzing structural and physiological features of AT3 aggregated species formed in *E. coli* intracellular environment co-expressing Hsp70 under different conditions.

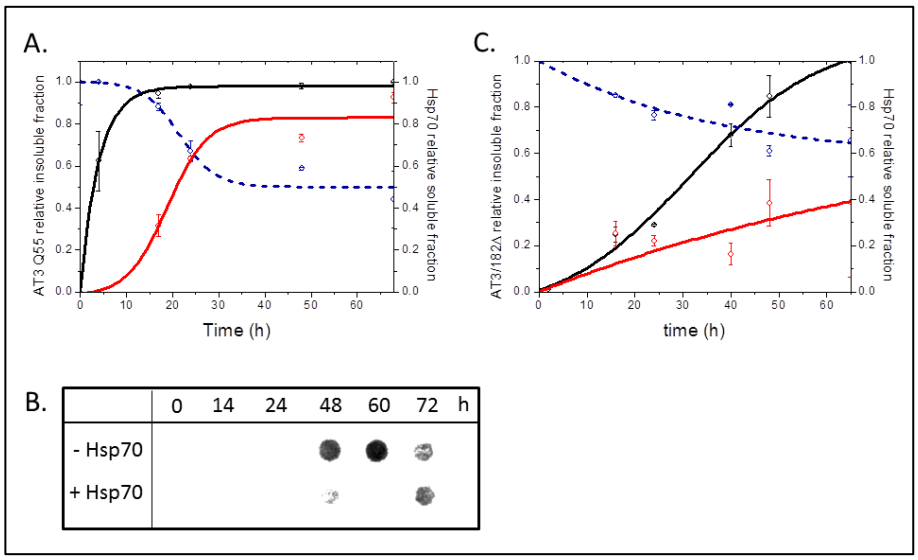

Fig. 1. (A) Changes in the insoluble protein fraction of AT3-Q55 alone (black line) and in the presence of an equimolar concentration of FL-Hsp70 (red line). Blue line represents the variation of soluble Hsp70. (B) Filter-trap assay of AT3-Q55 alone or in the presence of Hsp70. (C) Variation of the insoluble fraction of AT3/182Δ alone (black line) and in the presence of an equimolar concentration of FL-Hsp70 (red line). Blue line represents the variation of soluble Hsp70.

Table 1. K_{app} estimation (h^{-1})

	- Hsp70	**+ Hsp70**
AT3-Q55	0.25 ± 0.01	0.07 ± 0.02
AT3/182Δ	0.22 ± 0.1	0.01 ± 0.02

MATHERIAL AND METODS

Protein expression and purification of AT3 variants. The truncated (AT3/182Δ) and the expanded (AT3Q55) forms of AT3 were expressed as previously reported (14). Both protein variants were subjected to gel filtration on a Superose 12 10/300 qL (GE Healthcare) before each analysis. Recombinant N-hexa-His-tagged Hsp70 (Hsp70 1A, gi:194248072) were expressed and purified as previously described (15). Thrombin cleavage efficiency, estimated by mass spec analysis, was greater than 99%.

Protein purity exceeded 95% as determined by SDS–PAGE gel. Protein concentrations were determined by absorbance

measurements at 280 nm using theoretical extinction coefficients calculated with Expasy ProtParam (16).

Aggregation conditions and soluble fraction analysis. 25 µM AT3/182Δ or 12.5 µM AT3Q55 alone or together with an equimolar concentration of FL-Hsp70 was incubated in Tris 25 mM pH 7.4, 150 mM KCl, 5 mM $MgCl_2$ (with 0.01% NaN_3 to prevent bacterial growth) under shaking at 180 rpm at 37 °C. At specific incubation times, 10 µl aliquots were analyzed for soluble protein amount. Soluble protein fraction was isolated by centrifuging the samples at 16000 xg and analysing the supernatant by SDS-PAGE. Protein amount was estimated by densitometry analysis using Image-J software (NIH). Densitometry values were analyzed with the following equation generally used to fit fibrillization kinetics (17):

$$y = (A0 + mA0 * t) + \frac{(A+mA*t)}{1+\exp(-(t-t50\%)*\mathrm{Kagg})} \quad \text{Eq. 1}$$

Where A_0 and A are the values at beginning and the end of the aggregation, m_{A0} and m_A are the slopes of the lag phase and the plateau, assuming for them a linear dependency of normalized fluorescence values with the incubation time, $t_{50\%}$ is the midpoint of aggregation and k_{agg} is the apparent aggregation rate constant. The lag time of aggregation, t_{lag}, was determined from k_{agg} and $t_{50\%}$ as follows:

$$\mathrm{tlag} = t50\% - \frac{2}{\mathrm{Kagg}} \quad \text{Eq. 2}$$

Densitometry data are represented as insoluble protein fraction, normalizing them with the value corresponding to initial soluble protein amount and plotting reciprocal values.

Filter trap assay. Filter trap assay was performed on AT3Q55 samples incubated alone or in the presence of FL-Hsp70 as described above. At specific incubation times, 10 µl aliquots were mixed with 500 µl of SDS buffer (20 mM Tris-HCl, pH 6.8, 5% SDS, 10 mM DTT). After a 10 min incubation at 100°C, the resulting samples were applied to a cellulose acetate membrane (0.2 µm pore size) mounted on a manifold, vacuum-filtered, washed twice with 100 µl of SDS buffer. SDS-stable aggregates retained on the membrane were detected by

immunoblotting analysis using anti-human AT3Q24 Z46 polyclonal antibody (18).

4.2 References

1. **Warrick, J. M., Chan, H. Y., Gray-Board, G. L., Chai, Y., Paulson, H. L., and Bonini, N. M. (1999) *Nat Genet* 23, 425-428**
2. **Cummings, C. J., Sun, Y., Opal, P., Antalffy, B., Mestril, R., Orr, H. T., Dillmann, W. H., and Zoghbi, H. Y. (2001) *Hum Mol Genet* 10, 1511-1518**
3. **Adachi, H., Katsuno, M., Minamiyama, M., Sang, C., Pagoulatos, G., Angelidis, C., Kusakabe, M., Yoshiki, A., Kobayashi, Y., Doyu, M., and Sobue, G. (2003) *J Neurosci* 23, 2203-2211**
4. **Hay, D. G., Sathasivam, K., Tobaben, S., Stahl, B., Marber, M., Mestril, R., Mahal, A., Smith, D. L., Woodman, B., and Bates, G. P. (2004) *Hum Mol Genet* 13, 1389-1405**
5. **Yamanaka, T., Miyazaki, H., Oyama, F., Kurosawa, M., Washizu, C., Doi, H., and Nukina, N. (2008) *EMBO J* 27, 827-839**
6. **Huen, N. Y., and Chan, H. Y. (2005) *Biochem Biophys Res Commun* 334, 1074-1084**
7. **Hansson, O., Nylandsted, J., Castilho, R. F., Leist, M., Jäättelä, M., and Brundin, P. (2003) *Brain Res* 970, 47-57**
8. **Wacker, J. L., Huang, S. Y., Steele, A. D., Aron, R., Lotz, G. P., Nguyen, Q., Giorgini, F., Roberson, E. D., Lindquist, S., Masliah, E., and Muchowski, P. J. (2009) *J Neurosci* 29, 9104-9114**
9. **Zhou, H., Li, S. H., and Li, X. J. (2001) *J Biol Chem* 276, 48417-48424**
10. **Muchowski, P. J., Schaffar, G., Sittler, A., Wanker, E. E., Hayer-Hartl, M. K., and Hartl, F. U. (2000) *Proc Natl Acad Sci U S A* 97, 7841-7846**
11. **Krobitsch, S., and Lindquist, S. (2000) *Proc Natl Acad Sci U S A* 97, 1589-1594**
12. **Williams, A. J., Knutson, T. M., Colomer Gould, V. F., and Paulson, H. L. (2009) *Neurobiol Dis* 33, 342-353**
13. **Ellisdon, A. M., Thomas, B., and Bottomley, S. P. (2006) *J Biol Chem* 281, 16888-16896**
14. **Natalello, A., Frana, A. M., Relini, A., Apicella, A., Invernizzi, G., Casari, C., Gliozzi, A., Doglia, S. M., Tortora, P., and Regonesi, M. E. (2011) *PLoS One* 6, e18789**
15. **Roodveldt, C., Bertoncini, C. W., Andersson, A., van der Goot, A. T., Hsu, S. T., Fernández-Montesinos, R., de Jong, J., van Ham, T. J.,**

Nollen, E. A., Pozo, D., Christodoulou, J., and Dobson, C. M. (2009) *EMBO J* 28, 3758-3770

16. **Gill, S. C., and von Hippel, P. H. (1989) *Anal Biochem* 182, 319-326**
17. **Uversky, V. N., Li, J., and Fink, A. L. (2001) *J Biol Chem* 276, 44284-44296**
18. **Shehi, E., Fusi, P., Secundo, F., Pozzuolo, S., Bairati, A., and Tortora, P. (2003) *Biochemistry* 42, 14626-14632**