UAB
Universitat Autònoma
de Barcelona

Regulation of recombinant protein solubility and conformational quality in *Escherichia coli*

PhD Thesis

Elena Garcia i Fruitós

Facultat de Biociències

Departament de Genètica i de Microbiologia

Director: Dr. Antonio Villaverde i Corrales

2008

V.Discussion

Inclusion bodies are formed during overproduction of proteins, especially of those with mammalian or viral origin[114,221]. These aggregates are visible by phase contrast microscopy few hours post-induction of gene expression[143], showing variable size and morphology[200,202]. From an industrial point of view, we should consider that many proteins forming inclusion bodies have been excluded from the biotechnological and pharmaceutical market. In consequence, in the last years, different approaches have been developed to reduce inclusion body formation and, hence, increase the recovery of soluble and active protein. To achieve this goal successfully, it must be considered that the extent of protein aggregation is determined, at least partially, by a wide number of parameters such as media composition, growth temperature, gene dosage, promoter strength, mRNA stability, codon usage, fusion tags, protein sequence and availability of heat shock proteins[100,153,174]. However, the control of these parameters does not usually permit to obtain high amounts of native-like, soluble protein and, therefore, other approaches are required. To date, one of the main strategies utilised to recover functional, properly folded, soluble protein is the use of *in vitro* refolding procedures with pure inclusion bodies as starting material[221]. Another commonly used approach for this purpose has been the coproduction of chaperones, during protein production, a group of proteins with folding, holding and aggregation prevention properties. Even though many efforts have been addressed to achieve desirable yields of native, soluble protein, the strategies mentioned above often lead to inconsistent and discouraging results[3,160,181,313,314]. Thus, any of these approaches cannot be taken as universal and the general strategy has to be adapted for each protein in each new situation, becoming a trial-and-error process.

Although there is a considerable number of publications which have characterized inclusion body formation in bacteria[315,316], just a few of them mentioned the presence of biological activity in proteins embedded in these aggregates[317,318]. Therefore, we considered interesting to study in detail the biological activity associated to these protein deposits. For this, VP1LAC and VP1GFP, consisting of VP1 protein of foot and mouth disease virus fused to either the amino-termini of β-galactosidase or green fluorescent protein respectively, have been used as model proteins. Moreover, these proteins form inclusion bodies in all the selected strains.

V.1 DnaK is essential for *Escherichia coli* β-galactosidase folding

DnaK and its cochaperones DnaJ and GrpE are cell elements critical for the proper folding of misfolding-prone proteins in *Escherichia coli*[104]. Additionally, although inclusion bodies formed in cells devoid of functional DnaK are larger than in a DnaK[+] background[319,320], there is still a low but significant yield of proteins remaining in the soluble fraction[319]. In this context, our results show that, indeed, the absence of DnaK promotes VP1LAC aggregation and prevents proper folding (paper 1 and annex II.B), the remaining soluble protein being poorly active (paper 1 table 1 and annex II.B table 2) and less stable (paper 1 figure 5) than their counterparts produced in wild type cells. Specifically, wild type cells display an enzymatic activity between 9 and 15 times higher than mutant cells deficient in DnaK (paper 1 table 1). Moreover, protein stability shown by cells expressing DnaK is almost two-fold higher than that found in a DnaK[-] background (paper 1 figure 5). Therefore, we can conclude that DnaK folding activities are limiting in the conformational surveillance of a misfolding-prone β-galactosidase fusion protein, being such activities not complemented by other elements of the heat-shock response. In fact, in agreement with these findings, it has been previously described that DnaK is necessary for the interactions between complementing fragments that take place in the α-complementation of *Escherichia coli* β-galactosidase[321]. Therefore, we propose DnaK as a key element involved in promoting the tetrameric disposition needed for the complete activation of *Escherichia coli* β-galactosidase enzyme[264].

Summarizing, in the *dnaK* mutant, the folding of VP1LAC is only partially achieved, indicating that folding-assistant proteins alternative to DnaK are only moderately efficient (paper 1 and annex II.B). Thus, considering that DnaK is a critical element for proper folding and which activity cannot be completely complemented by other chaperones, we could partially explain the variable success under the coexpression of chaperones as a strategy to improve solubility[174,178,314].

V.2 Improving solubility in late exponentially phase protein production

The use of bacteria for the production of foreign recombinant proteins usually results in a poor yield of soluble, functional proteins. This phenomenon is especially evident when protein production is driven at high rates. The study in which we compared the soluble VP1LAC activity as well as its stability in DnaK⁻ and DnaK⁺ backgrounds (paper 1), shows that there are clear differences when gene expression is induced at the beginning or at the end of the stationary phase. Even though the thermal stability of the soluble protein produced in the *dnaK* mutant is two-fold lower than that found in their counterparts in wild type cells (paper 1 figure 5), we have also observed that, surprisingly, under late induction conditions, the half-life of the soluble VP1LAC is enhanced about two times, irrespective of the chaperone DnaK (paper 1 figure 5). Additionally, in DnaK⁻ cells not only is the soluble protein stability improved, but also the β-galactosidase enzymatic activity is doubled (paper 1 figure 3b). Moreover, interestingly, in both strains, the specific activity measured in old cultures is improved between 2 and 3-fold more that when produced in young cultures (paper 1 table 1).

Taking all these data and also considering that the fraction of soluble VP1LAC in the *dnaK* mutant increases from 13.9 ± 0.5 % in young cultures to 29.3 ± 0.9 % in old ones (data not published) and in wild type cells from 11.8 ± 1.6 % to 80.7 ± 7.5 % (data not published), this clearly means that protein production at late exponential phase is favoring protein folding as well as solubility. These results could be explained by the reduced biosynthesis of recombinant proteins in old culture, as it has been previously described in our group[322], which might favour proper folding[181,323-325] in a context of a low substrate load for the cell chaperones.

As mentioned in the introduction, there is a wide number of strategies used to improve solubility and functionality in protein production processes (section II.4). These strategies include gene expression at low temperatures[14,156-158,326], the use of weak promoters[166], *Escherichia coli* genetically modified strains[7,17,172], and the modification of media composition and cultivation strategies[162,171], among others. In

this context, we put forward gene expression in recombinant *Escherichia coli* at late exponential phase, to improve DnaK-independent folding processes.

V.3 Inclusion bodies: a new biological concept

V.3.1 Biological and structural composition of inclusion bodies

Even though it has been widespreadly believed that the aggregation process is mainly promoted by a high substrate load of the quality control system, in fact, the biological meaning of aggregation is still controversial.

In the last decade, inclusion body definition has dramatically changed. Until the last years, it was generally accepted that inclusion bodies were deposits of misfolded polypeptides that have escaped from the quality control system[106,191,218], becoming biologically inactive particles from which individual proteins cannot be recovered[191]. On the other hand, the literature[191,215] also described inclusion bodies as protease-resistant aggregates formed through a nonspecific process. However, an increasing number of studies has proved that, contrary to what has been accepted, proteins embedded in these protein aggregates precipitate through a sequence specific pathway[211] and are accessible to proteolysis[143,144,166,204,327], this proteolysis being not surface restricted[204] and occurring as a cascade process[327,328]. Additionally, although some authors still consider the biological activity associated to the proteins embedded in inclusion bodies as a mere contamination[207], it is becoming more evident that inclusion bodies formed by recombinant enzymes are capable to retain enzymatic activity[317-319]. Complementarily to this fact, pioneering studies in the early 1990s[260,329,330] as well as more recent investigations[211,331] suggest that proteins packaged as bacterial inclusion bodies adopt different conformational states ranging from native or native-like to enriched β-sheet structures (stabilized by a network of hydrogen bonds). In this context, our results not only support the presence of a well-defined molecular architecture and a noted biological activity, which contributes to the total activity (paper 1) associated to inclusion bodies (papers 2, 3 and 4), but also clearly prove that this phenomenon is general and not a mere contamination of soluble, functional protein that have unspecifically coaggregated. Analyzing four

structurally different proteins (VP1LAC, hDHFR, VP1GFP and Aβ42(F19D)-BFP), we found that inclusion bodies formed in all these cases display a significantly important specific activity (paper 2 table 1). Although the percentage of activity present in the aggregated form relative to that of the soluble fraction varies depending on the protein (6 % in hDHFR, 20 % in VP1GFP, 31 % in Aβ42(F19D)-BFP and 166 % in VP1LAC) (paper 2 table 1), it is considerably high in all the cases. Complementarily to this observation, the micrographs of the strains overexpressing the fluorescent proteins (VP1GFP and Aβ42(F19D)-BFP) clearly represent the noticeable fluorescence present in these inclusion bodies (paper 2 figure 2, paper 3 figure 1 and paper 4 figure 3). Moreover, when analyzing in detail the IB-fluorescence distribution, we found a heterogeneous fluorescence distribution, being the core of such aggregates, but not the surface layer, particularly rich in active protein forms (paper 3 figure 1). This particular fluorescence distribution pattern is found in VP1GFP inclusion bodies growth at 37, 30, 25, 20 and 16°C (paper 3 figure 1 and annex I -Authors correction and figure 1-). As shown by immunodetection, the distribution of VP1LAC[208] and VP1GFP (annex I figure 2) proteins in inclusion bodies is rather homogeneous. Therefore, the fluorescence distribution cannot be accounted by a variable protein density in inclusion bodies. To explain this fact, we propose the occurrence of an unbalanced equilibrium between protein deposition and removal[114,126]. Whereas protein aggregation involves both functional and misfolded polypeptides[260,329,331-333], we suggest a more selective process removing specially misfolded polypeptides at the inclusion bodies surface by the disaggregating chaperones (DnaK, ClpB and small heat shock proteins)[92,94,334,335]. However, a recent work of our group, in which this phenomenon has been studied in detail, supports an alternative hyptothesis based on a spontaneous *in situ* DnaK-dependent folding or refolding of inclusion body proteins (annex II.C). The structural reorganization of proteins within inclusion bodies suggested in this recent publication (annex II.C) is in agreement with other works in which a structural reorganization of other type of aggregates has been described[336-339].

Our group and others have recently published other examples of enzymatically active inclusion bodies occurring not only in the cytoplasm[340-347] but also in the periplasm[348] (table 6). On the other hand, it has also been proposed that the final amount of active

protein in inclusion bodies depends on how fast the aggregation takes place[349]. Hence, the longer the time that proteins remain in the soluble fraction before aggregation, the higher the activity found associated to these aggregates is. In this context, proteins that remain longer in the soluble fractions can fold better and become functional prior to aggregation.

Besides, we have also demonstrated that the existence of native-like structures in inclusion bodies[260,331-333] is not anecdotic (paper 2 figure 1 and paper 4 figure 1) and that in this kind of aggregates both properly folded polypeptides and enriched β-sheet structures coexist in a natural way (paper 2 figure 1, paper 4 figure 1, annex II.B and II.C). Analyzing the results obtained with the Fourier transform infrared spectroscopy (FTIR) (paper 2 figure 1), the presence of a peak at around 1620 cm^{-1} in the amide I region makes evident the presence of a tightly packed, extended intermolecular β-sheet architecture, an observation which is in agreement with what has already been described in other works in which inclusion bodies[211,331] (annex II.B and II.C) or other aggregates with a predominant β-sheet architecture such as amyloid fibrils[350] have been analyzed from a structural point of view. When analyzing the secondary structure of different mutants deficient in the main proteases and chaperones of the quality control system, we observed again the existence of a consistent pattern of secondary structure characteristic of amyloid depositions (paper 4 figure 1). Interestingly, in this case, a progressive downshift in the β-sheet peak (from 1627cm^{-1} to around 1623 cm^{-1}) shown by all the strains devoid of functional chaperones and proteases, when comparing with the wild type strain (paper 4 figure 1), indicates a tendency of all the mutants to form more compact β-sheet structures. Therefore, our data, together with other observations, indicate that, despite the intermolecular β-sheet-rich structure, aggregation as inclusion bodies does not necessarily disturb the conformation and, thus, the functionality, of all protein domains embedded in these aggregates (table 6). To explain this phenomenon we propose the following hypothesis: considering that enzyme active sites or fluorophores involved in the intermolecular β-sheet organization cannot be functional themselves, these functional sites must be located in properly folded molecule segments, distant from the aggregation-prone regions. This

hypothesis clearly supports the possible coexistence of active and inactive (β-sheet structure) molecules in a single aggregate unit. In fact, the variable extent of protein activity in our model proteins (paper 2 table 1) depends on the ratio of polypeptides forming native-like structures versus those organized as intermolecular β-sheet. One of the parameters favoring the loss of β-sheet pattern at expenses of native-like structure is a decrease in the growth temperature[329,351].

Interestingly, it has also been described that, within inclusion bodies, there are sub-classes of aggregates with distinctive proteolytic sensitivity[204,327]. Therefore, we suggest that, whereas proteins or protein segments in native conformation would be the protease-sensitive nuclei of inclusion bodies, enriched β-sheet polypeptides or polypeptide regions could constitute the protease-resistance sub-classes described by Carrió and coworkers[204,327] (figure 11).
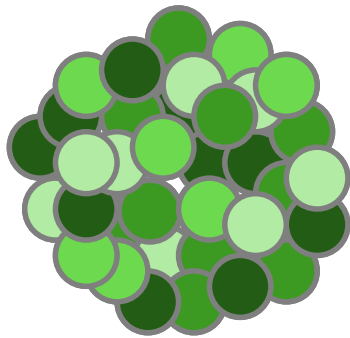

Figure 11. Schematic model of subclasses of aggregates. The set of colours shows the range of heterogenicity (e.g. in dark grey protease-resistance, enriched β-sheet polypeptides and in light green protease-sensitive, native proteins).

To sum up, and against the previous view, inclusion bodies are protein reservoirs formed in a specific manner with a surprisingly high conformational flexibility and biological activity, rather than being "molecular dust-balls" in cells overexpressing recombinant proteins.

Table 6. Inclusion body protein activity and structure.

Inclusion body protein	Biological activity	References
Green- and blue-fluorescent protein fusions	High inclusion body fluorescence emission *in vivo* and *in vitro*	Paper 2 / paper 3 / paper 4 / Mónica Martínez-Alonso *et al.* (2007) *FEMS Microbiology Letters* vol. 273(2): 187-95 / Andrea Vera *et al.* (2007) *Biotechnology and Bioengineering* vol. 96(6): 1101-6 / Kouhei Tsumoto *et al.* (2003) *Biochemical and Biophysical Research Communications* vol. 312(4): 1383-6.
β-galactosidase and β-galactosidase fusion proteins	High specific activity in purified inclusion bodies	Paper 1 / paper 2 / annex II.B and II.C/ D Margaret Worrall and Neil H. Goss (1989) *Australian Journal of Biotechnology* vol. 3(1): 28-32 / Kyung-Hwan Jung *et al.* (2008) *Journal of Industrial Microbiology and Biotechnology* in press.
Endoglucanase D	High activity in purified inclusion bodies	Kostas Tokatlidis *et al.* (1991) *FEBS Letters* vol. 282(1): 205-8.
Dihydrofolate reductase	Low activity in purified inclusion bodies	Paper 3
β-lactamase	Detectable activity in purified inclusion bodies	Dorota Kuczynska-Wisnik *et al.* (2004) *Acta Biochimica Polonica* vol. 51(4):925-31.
rHtrA1 serine protease	Detectable activity in purified inclusion bodies	Dorota Kuczynska-Wisnik *et al.* (2004) *Acta Biochimica Polonica* vol. 51(4):925-31.
Polyphosphate kinase	High activity in purified inclusion bodies	Jozef Nahálka *et al.* (2006) *Artificial Cells Blood Substitutes and Biotechnology* vol. 34(5): 515-21.
D-amino acid oxidase fusion protein	High activity in purified inclusion bodies	Jozef Nahálka and Bernd Nidetzky (2007) *Biotechnology and Bioengineering* vol. 97(3): 454-61.
Maltodextrin phosphorylase fusion protein	High activity in purified inclusion bodies	Jozef Nahálka (2007) *Journal of Industrial Microbiology and Biotechnology* in press.
Sialic acid aldolase fusion protein	High activity in purified inclusion bodies	Jozef Nahálka *et al.* (2008) *Journal of Biotechnology* vol. 134(1-2): 146-153.
Inclusion body protein	Structure	Reference
Interleukin-1 β	Inclusion bodies with native-like secondary structure (FTIR[1])	Kjell Oberg *et al.* (1994) *Biochemistry* vol. 33(9): 2628-34.
α-helix-rich hyperthermophilic proteins	Inclusion bodies with native-like secondary structure (FTIR, NMR[2], CD[3])	Mitsuo Umetsu *et al.* (2004) *FEBS Letters* vol. 557(1-3): 49-56.
TEM β-lactamase	Inclusion bodies with native-like secondary structure (FTIR)	Todd M. Przybycien *et al.* (1994) *Protein Engineering* vol. 7(1): 131-6.
Lipase	Inclusion bodies with native-like secondary structure (FTIR)	Diletta Ami *et al.* (2005) *FEBS Letters* vol. 579(16): 3433-6.
Human granulocyte-colony stimulating factor	Inclusion bodies with native-like secondary structure (FTIR)	Simona Jevsevar *et al.* (2005) *Biotechnology Progress* vol. 21(2):632-9.
Human growth hormone	Inclusion bodies with native-like secondary structure (FTIR)	Diletta Ami *et al.* (2006) *Biochimica et Biophysica Acta* vol. 1764(4): 793-9.
Human interferon α 2β	Inclusion bodies with native-like secondary structure (FTIR)	Diletta Ami *et al.* (2006) *Biochimica et Biophysica Acta* vol. 1764(4): 793-9.

[1]FTIR: Fourier transformed infrared spectroscopy.

[2]NMR: Nuclear magnetic resonance.

[3]CD: Circular dichroism.

V.3.2 Inclusion bodies as biocatalysers

Inclusion body formation has always been seen as one of the main drawbacks in the biotechnological and pharmaceutical industry, with many aggregation-prone proteins being discarded for commercialization. However, contrarily to this view, our data (paper 2, 3 and 4), together with other recent insights, draw inclusion bodies as active particles with an important extent of native-like form[352] and significant amounts of functional polypeptides[317,318,340-343,345-347]. Consequently, we propose the potential use of inclusion bodies as catalysers in bioprocesses as an economical alternative to the time-consuming, expensive protein refolding procedures[3,221]. To test this inclusion bodies application, we have analyzed the behavior of enzymatic bioprocesses driven by VP1LAC and hDHFR proteins (paper 2 and 3). Our results show that β-galactosidase substrates (ONPG and CPRG) are efficiently hydrolyzed by a suspension of purified VP1LAC inclusion bodies (paper 2 figure 3 and paper 3 figure 2); in fact, inclusion body-embedded VP1LAC enzymes perform the substrate hydrolysis faster than that mediated by the same amount of soluble VP1LAC protein (paper 2 figure 3a). Additionally, the conversion of NADPH into NADP$^+$ by hDHFR enzyme was also monitored and, although the substrate processing is slower when performed by inclusion bodies than by the soluble counterpart, it is high enough to consider inclusion body-embedded enzymes as efficient catalysts for enzymatic reactions (paper 2 figure 3b). In addition, inclusion bodies are a source of relatively pure polypeptides[126] (which can even reach 90 % of the total embedded polypeptides[114,315]), with a porous and highly hydrated architecture[202,204] that would facilitate the substrate diffusion. Hence, the biological activity associated to the enzyme-based inclusion bodies, together with its homogeneous, porous and hydrated structure, makes these aggregates a very attractive candidate to perform a wide number of bioprocesses.

To further investigate the process by which inclusion bodies perform bioreactions, we have deeply analyzed the hydrolysis reaction mediated by VP1LAC fusion protein (paper 3 figure 2). Firstly and interestingly, upon inclusion body resuspension, we observed an immediate release of functional protein to the solvent. Therefore, we decided to quantify not only the activity but also the amount of protein in both

inclusion bodies and soluble fraction. Specifically, even though between 7 and 8 % of the enzymatic activity is localized in the soluble fraction immediately upon inclusion body resuspension, only a very low amount of protein (lower than 0.0002 %) was found in this fraction, leading to a really high specific activity associated to the solubilised protein (paper 3). This phenomenon could be easily explained considering that, despite the nuclear localization (paper 3 and annex I figure 1), active forms might be exposed to the solvent due to the highly porous architecture and hydrated nature of these particular protein aggregates[202,204]. On the other hand, when following kinetically the substrate hydrolysis in an inclusion body suspension, we observed a time-dependent increase of the enzymatic activity linked to solubilised protein (paper 3 figure 3a and 3b). However, when comparing the inclusion body behavior in the presence (paper 3 figure 3a) or absence (paper 3 figure 3b) of substrate (ONPG), we can point up that there is an almost significative ($p=0.057$) substrate-mediated modulation of the activity-fractioning. In consequence, the release of active, aggregated polypeptides could be improved due to a protein conformational modification that seems to be substrate-dependent. In agreement with this observation, during the kinetics, the total enzymatic activity decreases more than sixfold in absence of substrate (paper 3 figure 3b), while the decrease is only moderate when the substrate is present (paper 3 figure 3a).

Though it is widely accepted that chaperones are tightly associated to inclusion bodies[208,353-355], more research is needed to clarify if the observed protein release is modulated by them or is rather a mechanical process.

In conclusion, the catalytic properties of porous, highly hydrated inclusion bodies open intriguing possibilities for a new industrial market of enzymatically active inclusion bodies. While the protein *in vitro* refolding procedures using these aggregates as a starting material[356-359] are in general expensive, complex, unsuccessful[3,221] and need to be adapted for each specific protein, the direct use of inclusion bodies as biocatalyers is a particularly appealing alternative. Although hormones and other drugs to be used *in vivo* would still require *in vitro* solubilisation[360], all the enzymes produced to be used in biotechnological processes could be immediately employed skipping any refolding step, because of the porous nature of these aggregates allowing substrate and product

diffusion. Therefore, after a rapid purification from disrupted cells, inclusion bodies can be resuspended in the desired reaction buffer and, once the reaction is accomplished, they can be easily removed from the reaction mixture by low speed centrifugation. In fact, several studies where inclusion bodies are used as biocatalysers have just been published[343,345-347]. In one of them, D-amino acid oxidase from *Trigonopis variabilis*[361-363], an enzyme of industrial relevance, is produced as insoluble inclusion body particles[343], as an efficient alternative to the immobilized whole cells or enzymes commonly used in bioprocesses[364]. In another one, the authors overproduce polyphosphate kinase obtaining high amounts of surprisingly active inclusion bodies which they immobilize and efficiently use in ATP/NTP synthesis[345]. Moreover, these authors describe an effective way to produce activated sugar monomers by using *Pyrococcus furiosus* maltodextrin phosphorylase entrapped as inclusion bodies[347]. Finally, this group also synthesizes sialic acid aldolase[365,366], an industrial enzyme used for neuraminic acid production, as active insoluble inclusion bodies trapped in alginate beads[346]. The results mentioned above, in agreement with our hypothesis, show that the use of inclusion bodies as biocatalysts is a powerful alternative to the technologies used up to now. Intriguingly, the results obtained show that all the used enzymes exhibit a high specific activity, and a long operational stability, being possible to use these aggregates repetively in different conversion cycles [343,345-347].

Our studies also show that activity and size of inclusion bodies can be modulated by choosing the appropriate strain (paper 4 and annex II.B) and the specific environmental conditions[341,342,351] (paper 1). In that way, highly active inclusion bodies with the desired size can be obtained for immediate use.

V.3.3 Biological activity: inclusion bodies and soluble protein versions

To better understand the biological meaning of the biological activity and native-secondary structure found in the inclusion body-embedded proteins, we have also compared what occurs in these aggregates and in the corresponding soluble counterpart concerning activity and molecular organization (paper 2). An interesting result is the occurrence of VP1LAC inclusion bodies with a higher specific activity (1162.5 ± 256 enzymatic units/mg) than that found in the soluble fraction (698.3 ±

153.0 enzymatic units/mg) (paper 2 table 1). This finding can be partially explained taking into account the occurrence of soluble but inactive proteins, a phenomenon that has already been described by other groups[225]. The so-called "soluble aggregates" are clusters of soluble but biologically inactive protein which might be inclusion body precursors[225] and which can be responsible for reducing the average specific activity in the soluble cell fraction. This could explain our results as well as the variability in the specific activities observed in different soluble enzymes when produced under different environmental conditions (paper 1 and annex II.B)[165,367,368].

Therefore, whereas in the biologically active inclusion bodies both properly folded and β-enriched polypeptides coexist, the structural and functional composition of the soluble fraction might also be diverse.

V.4 Solubility and conformational quality are not coincident events

The occurrence of active polypeptides and native-like structures in inclusion bodies indicates that the conventional model regarding recombinant protein production, in which active proteins are localized in the soluble cell fraction and inactive, misfolded polypeptide chains are aggregated as inclusion bodies (figure 12), is not appropriate. This model (figure 12) assumes that protein misfolding prevents both solubility and functionality, becoming solubility an indicator of conformational quality. Nevertheless, as mentioned above, we observed that protein solubility and conformational quality are not matching properties (paper 1, 2, 3 and 4). In this regard, Chiti and co-workers also described, through an exhaustive mutational analysis, that determinants of protein

Figure 12. Conventional model. Soluble, functional proteins (green spheres) and insoluble, inactive proteins (red spheres) (annex II.A figure 1a).

misfolding aggregation are not coincident[369]. Therefore, to develop a new, suitable model, in first term, we have considered that physiological aggregation as inclusion bodies does not split protein population into active and inactive fractions (paper 1, 2, 3

and 4). Additionally, the coexistence of aggregation patches[370] with native-like regions in inclusion bodies (papers 1, 2, 3 and 4)[352] must be taken into account. Besides, this model should also support the existence of "soluble aggregates"[172,225] in the soluble cell fraction, a protein population with a lower activity than the equivalent "true" soluble species observed in the production of proteins such as glutathione transferase[371], β-galactosidase[372] and maltose-binding protein[373,374]. Considering all these observations, we have developed a new model that fits in this new concept (figure 13), in which both solubility and functionality are not necessarily associated events and in which protein quality of inclusion body protein is representative of that found in the whole cell.
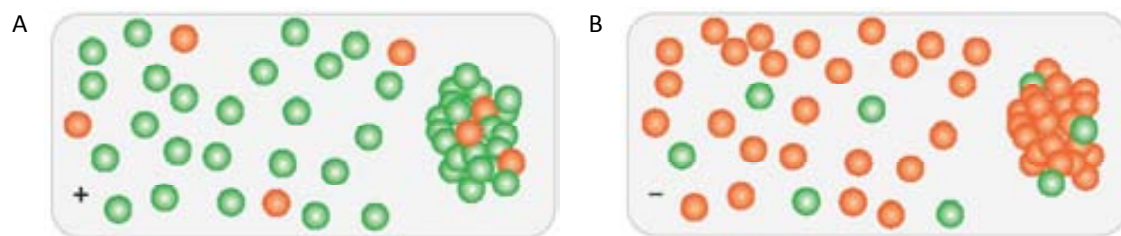

Figure 13. New model: A (favorable conditions), B (unfavorable conditions). Soluble, functional proteins (green spheres) and insoluble, inactive proteins (red spheres) (annex II.A figure 1b and 1c).

In this context, the divergence between solubility and conformational quality shown in figure 13 is also illustrated by the negative relationship between the occurrence of soluble species and their specific emission (paper 4 figure 6a). The variability in the occurrence of functional protein (compare figure 13 A and B) can be partially due to environmental parameters. For example, conditions such as reduced culture growth temperature[351,375] and the coexpression of critical chaperones[341] favors proper folding, therefore, increasing the yield of active protein present in both soluble and insoluble cell fractions (figure 13A). On the other hand, disfavoring conditions like growth at high temperatures and low levels of chaperones, among others, reduce in parallel the conformational quality of proteins localized either in the soluble fraction or in inclusion bodies (figure 13B). In consequence, contrary to what has been generally believed, solubility does not appear to be an all-or-nothing attribute and polypeptides might exhibit a continuum of folding states in both soluble and insoluble cell fractions (figure 13), between which they can be transferred with the assistance of cellular folding

modulators such as chaperones and cochaperones[126]. Moreover, the specific activity of soluble and inclusion body-embedded proteins can be, at least in some cases, quite similar (paper 1 and figure 13)[342], especially in absence of functional DnaK (paper 1 table 1 and annex II.B). DnaK, through the selective removal of active species from inclusion body superficial layers (paper 4), might be involved in the portioning of functional proteins between soluble and insoluble cell fractions. This phenomenon is in accordance with the DnaK inclusion body surface localization[208] as well as with the absence of functional proteins in the external layers of these aggregates (paper 3).

To sum up, we can conclude that solubility (resulting from the combination of the protein folding process and an overcommitted, highly selective proteolysis) should be seriously reconsidered as a protein quality universal indicator, being biological activity a most convenient reporter of the conformational quality (section V.4). Consequently, the discrimination between conformational quality and solubility would help in the optimization of production processes to recover higher yields of functional polypeptides.
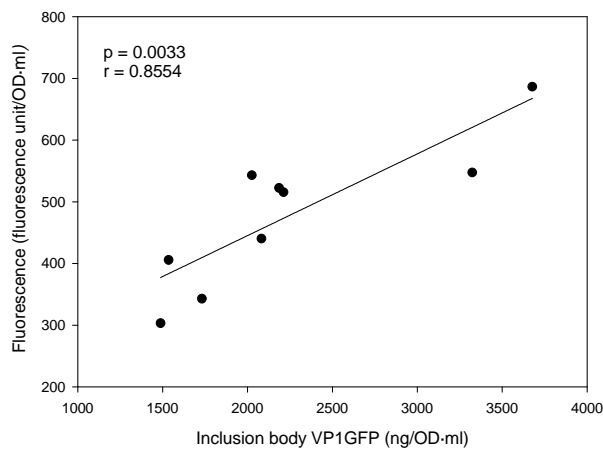
V.4.1 GFP: a protein quality indicator

As mentioned above biological activity is a convenient protein quality indicator. Because of the fact that GFP is only fluorescent if the chromophore has been properly formed[376] and the whole polypeptide has reached the mature structure[301,376,377], it is an excellent candidate to determine protein conformational quality. In accordance with this fact, our data (paper 2 and paper 4) as well as other works[342,378,379] show that GFP emission is really a good reporter of protein folding-misfolding.

V.5 *Escherichia coli* quality control apparatus

It is becoming more obvious that inclusion bodies, apart from sharing a wide number of characteristics with amyloid depositions[211], are partially formed by biologically active polypeptides[317,318,340-347] (paper 2, 3, 4, annex II.B and II.C) with important extents of native-like secondary structure[352] (paper 2, annex II.B and II.C). However, nowadays, it is still not known how this fact is regulated by the cell. The complex protein quality control system of all living organisms, essentially composed by chaperones and proteases, controls protein folding and the degradation of folding-reluctant species, therefore, preventing protein aggregation[75,135,342,353,380-383]. The action of the quality control machinery is especially relevant under conformational stress situations such as high temperature growth or high rates of protein synthesis[146]. Even though protein conformational quality and solubility have been traditionally assumed to be intimately connected, considering that solubility is indicative of conformational quality, this view has dramatically changed in the last few years (section V.4). In order to explore the eventual divergent control of these distinguishable protein properties within the cell, we studied in detail both the distribution of biological activity between the soluble and insoluble fractions and the aggregation propensity of a fluorescent recombinant protein in different mutants devoid of the main cytosolic chaperones and proteases (paper 4). Analyzing the results obtained we observed that, as expected, the total or partial disruption of different arms of *Escherichia coli* quality control apparatus results in more protein aggregation (paper 4 figure 2a and 2b). Nevertheless, all the tested strains, except IbpAB⁻, produce, surprisingly, much more functional proteins than wild type strain (paper 4 figure 2c, 2d and table 1), being the gain in the total fluorescence due to, in most of the cases, an increase in the fluorescence of both soluble and insoluble fractions (paper 4 figure 4a and 4b). Interestingly, fluorescence was considerably intense in genetic backgrounds with high aggregation (figure 14 and paper 4 figure 2 and 4). This phenomenon was especially appealing in cells lacking functional DnaK, where inclusion bodies are extremely big[319,320] and fluorescent (paper 4 figure 3).

On the other hand, we also observed high amounts of recombinant protein in all the mutant strains, excluding IbpAB⁻ (paper 4 figure 5a). This enhancement in VP1GFP yields perfectly correlates with an increase in proteolytic stability in the absence of relevant chaperones or proteases (paper 4 figure 5b), indicating that the diminished fluorescence emission in wild type strain is not due to a saturation of the folding machinery but a consequence of a low protein stability (paper 4 figure 5).

Figure 14. Pair-wise comparaison of inclusion body-embedded VP1GFP and the total cell fluorescence in different producing strains.

V.5.1 Proteolysis mediated by DnaK

While in *Escherichia coli* protease deficient cells (Lon⁻, ClpP⁻ and ClpA⁻ -ClpP ATPase subunit-) proteolysis is widely minimized for most recombinant proteins[146,384-386], leading to an important yield of stable protein, proteolysis inhibition is not evident in mutant strains deficient in chaperones such as GroES, GroEL, ClpB and DnaK (paper 4 figure 5b), although it has been previously reported that in DnaK⁻ cells recombinant proteins are proteolytically more stable[319]. Particularly, VP1GFP half-life in DnaK⁻ or ClpB⁻ is not only noteworthy but close to that shown in absence of Lon, ClpP or ClpA (paper 4 figure 5b), the main cytosolic proteases degrading recombinant proteins[149,386]. Therefore, all these data indicate that both DnaK and ClpB positively mediate proteolysis, indicating that chaperones and proteases play interconnected roles. Specifically, our results suggest that in cells with a fully functional quality control system, ClpP and Lon proteases, directed by DnaK in association with ClpB, degrade aggregation-prone but functional (or suitable to be activated) proteins by an overcommitted activity that, while minimizing aggregation, dramatically reduces the cellular amounts of functional protein species. There are some reports indicating that

DnaK could be involved in degradation by delivering folding intermediates to proteases[96,146,387,388] and, moreover, a recent publication of our group describes a physical interaction between DnaK and VP1GFP through which the recombinant protein is inactivated[341]. However, in spite of these evidences, further investigation is required to elucidate the precise mechanics of this DnaK-dependent proteolysis.

V.5.2 IbpAB: antagonist in the proteolysis

On the other hand, while most of the tested mutants seem to be involved, at least partially, in protein degradation, recombinant proteins produced in cells devoid of the small heat shock chaperones IbpA and IbpB are not more stable than those produced in wild type cells (paper 4 figure 5b). This behavior is in agreement with the fact that IbpAB, which acts in cooperation with DnaK and ClpB chaperones, seems to display a protective role against proteolysis[389], therefore acting as an antagonist in the proteolysis mediated by DnaK-ClpB.

V.5.3 Divergent role of the quality control system

Analysing the positive, significative correlation between the specific fluorescence of soluble and insoluble cell fractions of the strains used in paper 4, we can conclude that conformational quality control system acts irrespectively of protein solubility (paper 4 figure 6b). Even though both slopes are similar (0.0906 -upper line- and 0.0767 -lower line-), DnaK⁻, ClpA⁻, ClpB⁻ and ClpP⁻ inclusion bodies show higher specific fluorescence than GroEL, GroES, IbpAB⁻ and Lon⁻ (paper 4 figure 6b); this interestingly means that DnaK, ClpA, ClpB and ClpP are elements of the protein quality control that actively participate in the discrimination of functional species between soluble and insoluble fractions.

Besides, our results also illustrate that the bacterial quality control promotes solubility instead of conformational quality, through an overcommitted proteolysis of aggregation-prone polypeptides (paper 4). In this context, it has also been described that solubility, yield and conformational quality of soluble proteins cannot be favoured simultaneously (annex II.D). Therefore, to date, the mechanics of this quality control has been largely misunderstood, because, contrary to what has been widely believed,

recombinant protein conformational quality and solubility show a divergent control. In this context, the proposed divergence is supported by the negative relationship between solubility and the specific fluorescence of the soluble cell fraction (paper 4 figure 6a). Additionally to this fact, our group has also described that, under low growth temperature conditions, VP1GFP conformational status of soluble and insoluble proteins is improved in parallel[342], supporting the idea that virtual cell compartments do not have very much sense regarding protein quality. In agreement with this fact, we have also observed, specially in cells devoid of DnaK, that VP1LAC specific activity is quite similar in both soluble fraction and inclusion bodies (paper 1 table 1 and annex II.B table 2). Moreover, the specific fluorescence emission of soluble and inclusion body VP1GFP behave exactly in the same way along protein yield and DnaK availability ranges[341]. Thus, considering the role of the chaperones and proteases concerning solubility and conformational quality (paper 4), the coexpression of these elements of the conformational quality control apparatus to optimize the production and recovery of soluble, functional polypeptides, should be reconsidered. In fact, this new concept might partially explain the inconsistent results obtained upon coexpression of chaperones to gain solubility[160,178,390].

V.5.4 Inclusion bodies: protein reservoirs integrated in the protein quality system
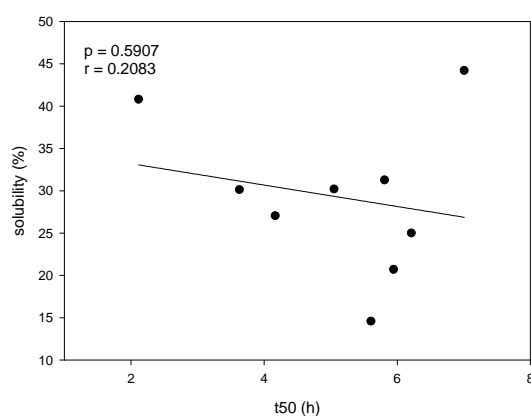

Figure 15. Pair-wise comparaison of VP1GFP half life (t50) and solubility in different producing strains.

It is generally believed that proteins localized in inclusion bodies are protected against proteolysis[160]. However, our results, in agreement with other reports[327,386,391], point out that inclusion body size is highly regulated by proteolysis (paper 4) and, consequently, these aggregates are completely integrated in the cell protein quality control. When analyzing the possible

existence of a correlation between protein solubility and half-life, we do not observe a positive correlation or a negative one either (figure 15). Therefore, the fact that less protein solubility does not correlate with higher half-life indicates that proteolytic resistance drives to aggregation and not the opposite. Additionally, inclusion body formation is not only proteolytically regulated, but also completely integrated in the conformational quality control (paper 4 figure 6b).

VI.Conclusions

1. The enzymatic activity of an aggregation-prone β-galactosidase (VP1LAC) produced in wild type *Escherichia coli* cells is between 9 and 15 times higher than in mutant cells deficient in DnaK. Moreover, the stability of soluble VP1LAC in a DnaK⁻ background is almost two-fold lower than that found in their counterparts in DnaK⁺ cells. Therefore, since the soluble protein produced in absence of DnaK is less active and less stable than that produced in wild type cells, we can affirm that DnaK is critical in the conformational surveillance of aggregation-prone proteins, which folding cannot be totally complemented by other elements of the heat-shock response.

2. The specific activity and also the half-life of the soluble VP1LAC produced in old batch cultures is about two fold higher than that produced in young cultures, irrespective of the chaperone DnaK. Additionally, VP1LAC solubility in both strains when produced at late exponential phase is between 2 and 7 fold higher than that found in young cultures. In consequence, recombinant gene expression at late exponential phase favors protein folding, stability and solubility, probably due to the reduced biosynthesis of recombinant proteins that improve DnaK-independent folding.

3. The biological activity of hDHFR, VP1GFP, Aβ42(F19D)-BFP and VP1LAC present in inclusion bodies relative to that of the soluble fraction is variable but significant (6 %, 20 %, 31 % and 166 %, respectively). Consequently, the presence of biological activity associated to inclusion bodies is not anecdotic but a general event. Therefore, enzymatically active inclusion bodies, due to its easy obtention, purity, porous nature and highly hydrated architecture, might be used as catalysers in bioprocesses, as an economical alternative to the time-consuming and expensive protein refolding procedures

4. Fourier transform infrared spectroscopy (FTIR) of inclusion bodies formed in a wide number of *Escherichia coli* strains showed a peak at around 1620 cm⁻¹ in the amide I region, characteristic of the intermolecular β-sheet architecture. Moreover, in strains devoid of functional chaperones and proteases, there is a downshift from 1627cm⁻¹ to around 1623 cm⁻¹ in the β-sheet peak compared with the wild type strain, indicating a tendency to form more compact β-sheet structures. Therefore,

native-like polypeptides coexist with enriched β-sheet structures in inclusion bodies, without compromising the functionality of properly folded polypeptides. Consequently, the ratio of properly folded polypeptides with native-like structure versus those organized as intermolecular β-sheet structures would define the specific biological activity associated to inclusion bodies.

5. VP1GFP and Aβ42(F19D)-BFP inclusion bodies are highly fluorescent. In this context, the core of VP1GFP aggregates at 37, 30, 25, 20 and 16°C is fluorescent, while the surface layer is poor in active proteins forms. Besides, VP1LAC and VP1GFP immunodetection analyzed by transmission electron microscopy reveals a homogeneous inclusion body protein distribution in both cases. Therefore, functional polypeptides are not surface-limited contaminants of inclusion bodies but true structural components.

6. VP1LAC inclusion bodies, once resuspended in PBS buffer, release functional protein to the solvent in a time-dependent way and such protein release, that has not been observed in VP1GFP inclusion bodies, is substrate-mediated.

7. The specific activity of VP1LAC protein aggregated as inclusion bodies is almost two-fold higher than in the corresponding soluble counterpart. Hence, this fact supports the occurrence of biologically inactive proteins in the soluble fraction.

8. *Escherichia coli* cells devoid of GroEL, GroES, ClpA, ClpB, Lon, ClpP or DnaK, overproducing VP1GFP, are more fluorescent than wild type cells. Besides, most of these strains, when compared to wild type, show an enhanced fluorescence in both soluble and insoluble fractions, an increased protein aggregation and enhanced amounts of recombinant protein. Interestingly, the most fluorescent strains were those in which aggregation was higher and backgrounds with improved VP1GFP yields show an increase in proteolytic stability. Consequently, the total or partial disruption of different arms of *Escherichia coli* quality control apparatus results in proteins with a high conformational quality and more stable than those obtained in wild type cells, irrespective of their solubility state.

9. VP1GFP half-life in DnaK$^-$ or ClpB$^-$ strains is similar to that observed in absence of the main cytosolic proteolytic apparatus formed by Lon, ClpP and ClpA. Therefore, it seems that ClpP and Lon proteases, directed by DnaK in association with ClpB,

degrade aggregation-prone but functional (or suitable to be activated) proteins by an overcommitted activity that, while minimizing aggregation, dramatically reduces the cellular amounts of functional protein species.

10. As an exception, VP1GFP stability in cells devoid of IbpA and IbpB is close to that found in the wild type strain. Consequently, while the other tested chaperones and proteases seem to be involved in proteolyisis, the small heat shock chaperones are not involved in this phenomenon.

11. There is a positive correlation between the specific fluorescence in soluble and inclusion body fractions, being DnaK$^-$, ClpA$^-$, ClpB$^-$ and ClpP$^-$ the strains with the highest specific fluorescence in the aggregated fraction. Therefore, the conformational quality control system acts irrespectively of protein solubility. Moreover, DnaK, ClpA, ClpB and ClpP are those elements of the quality control machinery specifically involved in the fractioning of functional species between soluble and insoluble fractions.

12. Solubility and conformational quality of recombinant proteins are not matching properties and the physiological aggregation as inclusion bodies does not split protein population into active and inactive fractions.

13. The quality control machinery acts irrespectively of protein solubility, improving or deteriorating conformational status of soluble and insoluble proteins in parallel. Moreover, this system promotes solubility instead of conformational quality, through an overcommitted proteolysis that results into a divergent control of conformational quality and solubility.

14. Solubility should be seriously reconsidered as a universal indicator of protein quality, being biological activity a most convenient reporter of the conformational quality.

VII.Annex I

AUTHOR'S CORRECTION

Localization of Functional Polypeptides in Bacterial Inclusion Bodies

Elena García-Fruitós, Anna Arís, and Antonio Villaverde

Institut de Biotecnologia i de Biomedicina and Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

Volume 73, no. 1, p. 289–294, 2007. Page 290: The confocal microscopy pictures of inclusion body-producing cells shown in Fig. 1 correspond to 1.8-μm sections instead of 0.04-μm sections as indicated in the text. VP1GFP inclusion bodies analyzed through 0.08-μm sections (the minimal thickness technically feasible for bacterial cells with the used equipment) offered a layered and concentric distribution of fluorescence similar to that shown in the published material. Therefore, all the conclusions of our work are perfectly valid.
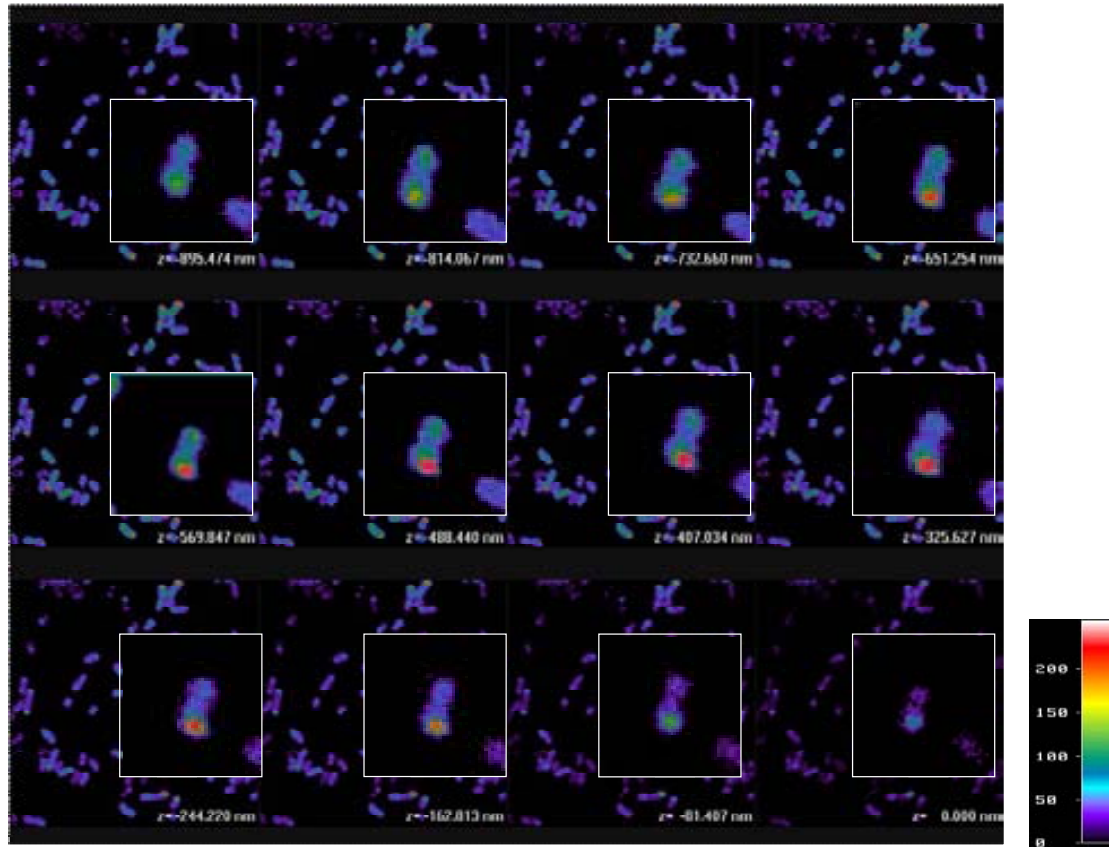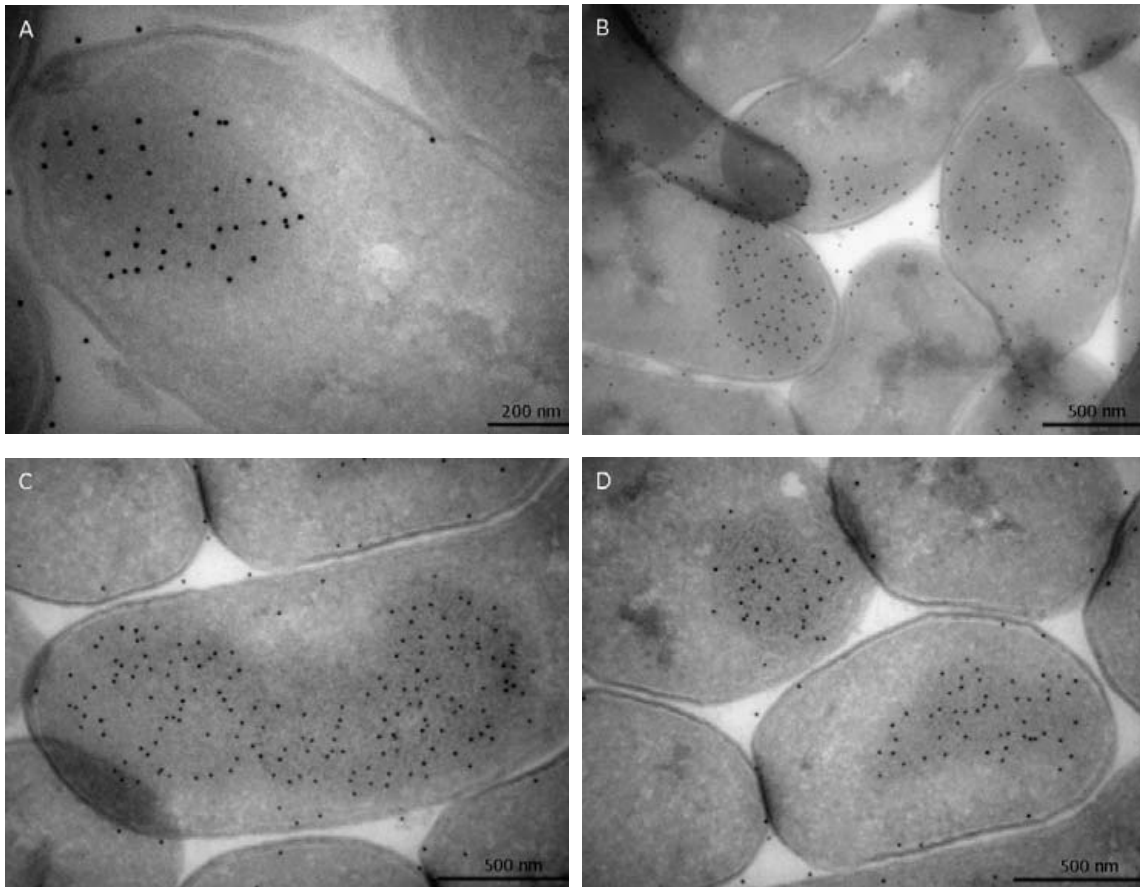
Figure 1. Metamorph image analysis of VP1GFP inclusion bodies formed at 37°C analysed through 0.08 μm serial sections. The color scale is despicted at the right.

Figure 2. *In situ* immunolocalization of GFP in VP1GFP-producing *Escherichia coli* MC4100 cells (A), and in the derivatives DnaK⁻ (B) and ClpP⁻ (C, D). Samples were obtained upon the induction of VP1GFP gene expression as described previously in paper 3. At 3h after induction of gene expression, cells were harvested by centrifugation and fixed with 4% (v/v) paraformaldehyde (EM grade, Merck) in 0.1 M phosphate buffer (PB) pH=7.4. Later, samples were rinsed with PB, incubated in 20 mM glycine solution to quench the free aldehyde groups and embedded in 12% (w/v) gelatine, cryoprotected in 2.1 M sucrose solution and cryofixed in liquid nitrogen. All sections were cut at -120ºC with a cryoultramicrotome (Leica Ultracut UCT, Vienna) and deposited onto formvar coated Cu/Pd grids. Ultrathin sections were initially blocked in PBS containing 1 % (v/v) BSA, incubated with a primary polyclonal antibody anti-GFP (sc-8334 Santa Cruz Biotechnology, Inc., 1/50 dilution) for 40 minutes. After washing in PBS, sections were incubated with gold-labelled protein A (Utrech, 10 nM). Following the final washes with PBS, samples were fixed with 1% glutaraldehyde and washed with distilled water. Samples were finally mounted in metilcelulose-uranyl acetate and visualized in a transmission electron microscope Hitachi H-7000.

VIII.Annex II

VIII.1 Annex II.A

Recombinant protein solubility-does more mean better?

Núria González-Montalbán*, Elena García-Fruitós* and Antonio Villaverde

*Contributed equally to this work

CORRESPONDENCE

Nevertheless, some areas merit further attention. A first problem is that the majority (>70%) of studies on GM rice have been conducted under laboratory conditions. These studies have limited predictive ability regarding large-scale, long-term effects in the field. For this reason, the impact of human activities and agricultural practices on GM rice performance has not been adequately incorporated into the Chinese risk assessment process. For example, although there is a considerable amount of research on natural pollen dispersal in China, under certain conditions human-mediated seed dispersal may have a stronger influence on the risks associated with gene flow[5]. To address such issues, China should consider undertaking something similar to the 'farm-scale evaluations' undertaken in the United Kingdom, the results of which were published in 2003 (ref. 6).

Furthermore, relatively few data have been gathered on the development of resistance by insects and pathogens to Bt rice or Xa21 rice in China. Assessment of these issues should be undertaken before commercialization of GM rice in the country.

Third, socioeconomic considerations need to be considered concerning the actual and potential consequences of adoption of GM rice, such as the potential impact on farmers' incomes and welfare, cultural practices, community well-being, traditional crops and varieties, rural employment, trade and competition, ethics and religion, consumer benefits and ideas about agriculture, technology and society. Taking such considerations into account during the risk assessment process is not legally required in China. But experience has shown they are important factors in China; for example, several empirical economic studies on cotton have revealed income gains for small farmers who plant Bt cotton seeds[7]. More attention needs to be paid to this issue if GM rice is to be commercialized in a sustainable manner.

Fourth, China's sheer size poses a headache for compliance and risk management. Any coordinated management effort must oversee millions of farms. One report assessing the productivity and health effects of two insect-resistant transgenic rice varieties[8] highlighted the fact that farmers were cultivating GM rice without the assistance of knowledgeable technicians. This assessment method was contrary to that set down in the 2001 regulations and raises questions about implementation of regulations and whether the safeguards to prevent GM contamination are effective[9,10].

Questions have also been raised about the

extent that farmers are using GM rice illegally, and the significant international consequences of possible 'contamination' of the Chinese rice supply. The Ministry of Agriculture has conducted a series of investigations and clarifications[11,12], issuing The Guidelines of Biosafety Investigation on Field-Testing of GM Crops in May 2006 to clarify legal requirements for field testing of GM crops. Even so, as the experience with illegal use in India, Argentina and Brazil and admixture in the United States have demonstrated, illegal use and mixing GM and non-GM staples in the food supply are likely to be ongoing problems. Clearly, greater awareness about the requirements of the regulations among farmers using GM crops would be helpful. Greater vigilance by the Ministry of Agriculture is also needed. If compliance problems persist, then stronger sanctions will need to be considered. But as a first step, a more independent investigation of compliance and illegal use is warranted in China.

We conclude that Chinese scientists are using procedures to assess the risks associated with GM rice that meet international standards. Even so, more attention needs to be paid to some critical areas, such as field testing, scientific uncertainty and socioeconomic considerations. Comparative and collaborative studies with other countries would be helpful in developing better procedures. Perhaps the most pressing issue though is addressing compliance issues and

the illegal use of GM crops. The need for a more independent evaluation of these issues is clear.

Yanqing Wang & Sam Johnston

United Nations University-Institute of Advanced Studies, 6F, International Organizations Center, Pacifico-Yokohama, 1-1-1 Minato Mirai, Nishiku, Yokohama, 220-8502, Japan.
e-mail: wang@ias.unu.edu

1. Marris, E. Escaped Chinese GM rice reaches Europe. news@nature.com 5 Sept, (2006). <http://www.nature.com/news/2006/060904/full/060904-5.html>
2. European Commission (2007). Summary Record of the Standing Committee on the Food Chain and Animal Health Held in Brussels on January 16, 2007 (EC, Brussels, 2007). <http://ec.europa.eu/food/committees/regulatory/scfcah/modif_genet/summary12_en.pdf>
3. Wang, Y. & Johnston, S. UNU-IAS Working Paper No. 152 (United Nations University-Institute of Advanced Studies, Yokohama, Japan, 2007). <http://www.ias.unu.edu/resource_centre/152%20Yanqing%20Wang.pdf>
4. Chen, M. et al. Insect Sci. 13, 409–420 (2006).
5. Snow, A.A. et al. Ecol. Appl. 15, 377–404 (2005).
6. Andow, D. Nat. Biotechnol. 21, 1453–1454 (2003).
7. Huang, J. & Wang, Q. AgBioForum. 5, 122–135 (2002).
8. Huang, J., Hu, R., Rozelle, S. & Pray, C. Science 308, 688–690 (2005).
9. <http://english.people.com.cn/200506/14/print20050614_190130.html>
10. <http://english.people.com.cn/200603/15/eng20060315_250850.html>
11. <http://news.xinhuanet.com/newscenter/2005-08/10/content_3335353.htm>; <http://www.china.org.cn/english/2005/Apr/125745.htm>
12. <http://www.stee.agri.gov.cn/biosafety/zhbd/t20060331_583313.htm>; <http://news3.xinhuanet.com/english/2006-03/31/content_4367924.htm>

Recombinant protein solubility—does more mean better?

To the editor:

Since the inception of recombinant DNA technology, maximizing the solubility of a heterologous protein has been the goal of large-scale biomanufacture[1]. In bacterial cell factories, recombinant proteins usually fail to fold properly and accumulate as refractive, insoluble particles called inclusion bodies. Higher yields of soluble proteins have been pursued either by reducing the culture temperature, engineering the protein sequence, adding fusion partners or coproducing selected chaperones[2]. In general, these approaches have been rather hit-or-miss and have proven to be particularly inefficient with respect to both membrane proteins and other difficult-to-express proteins. In the haste to improve solubility, conformational quality has often been disregarded or assumed to be intimately linked to solubility (**Fig. 1a**). Here,

we summarize evidence that challenges the broadly accepted notion that low recombinant protein solubility is indicative of poor conformational quality and compromised biological activity. On this basis, we suggest recombinant proteins expressed as inclusion bodies in bacteria could retain biological activity and thus be of greater potential biotechnological utility than previously appreciated.

Inclusion bodies have long been regarded as clusters of polypeptides unable to reach native conformation that are deposited as inert reservoirs sequestered from protease activity. An increasing number of reports indicate that bacterial inclusion bodies comprise proteins in different conformations, a significant proportion of which may be biologically active. For example, a pioneering study in 1989 by Worrall and Goss[3] showed
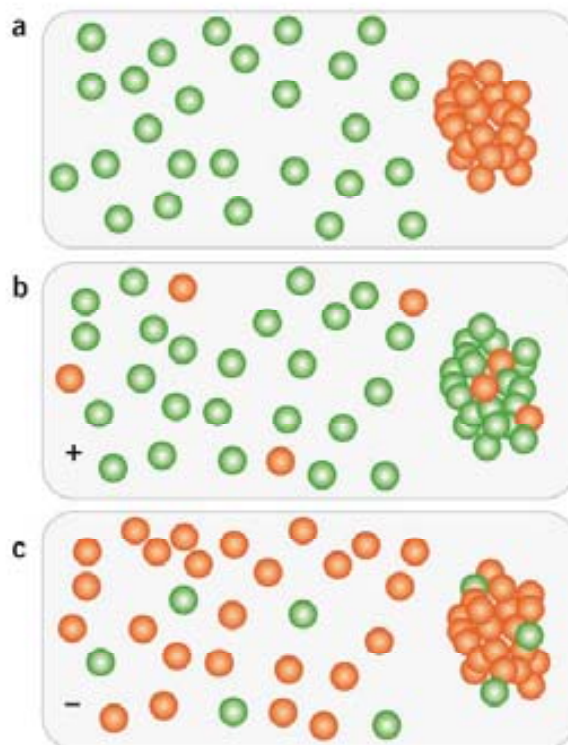
Figure 1 The functionality of recombinant proteins aggregated as inclusion bodies. (a) Conventional model of recombinant protein fractioning in inclusion body–forming cells, where functional species (green spheres) occur in the soluble cell fraction, whereas inactive, misfolded polypeptides (red spheres) are deposited in inclusion bodies. According to this model, protein misfolding prevents both solubility and biological activity. Therefore, solubility would be a straightforward indicator of conformational quality (and insolubility an indicator of poor protein quality). In this model, solubility and biological activity are to all intents and purposes all-or-nothing properties. (b) The new model, which represents inclusion bodies as misfolding-prone recombinant polypeptides clustered in specific cell locations, in this case under conditions (+) that favor proper folding. (c) The new model, but this time under conditions (–) that disfavor proper folding. In b and c, the biological activity of recombinant proteins ranges from no detectable activity to 100% specific activity, and misfolded proteins might exist along a continuum of forms, from fully soluble versions to true refractile inclusion bodies. Because solubility and biological activity are therefore not matching events, mixtures of active protein species (exhibiting a considerable amount of native-like secondary structure) and non-functional polypeptides coexist in both soluble and insoluble cell fractions. The average conformational quality of inclusion body protein is representative of that found in the whole cell.

that inclusion bodies formed by *Escherichia coli* β-galactosidase exhibit specific activities around one-third of those shown by soluble enzyme. Other enzymes, produced alone or as fusion proteins, form inclusion bodies that are also biologically active (for a review, see ref. 4). Our group[5] and others[6] have demonstrated this occurs not only in the cytoplasm, but also in the periplasm if secreted. In some cases, the specific activities of the aggregated and soluble enzyme versions are very close[5]. Similarly, work by our group[5,7] and others[8] at the Institute for Biotechnology and Biomedicine has shown that fluorescent proteins (e.g., green fluorescent protein (GFP), blue fluorescent protein and derived fusions) form inclusion bodies that remain highly fluorescent.

This evidence of biological activity in inclusion bodies indicates that aggregation of recombinant proteins in bacteria can be driven by specific interactions between solvent-exposed hydrophobic stretches in partially structured species that are not necessarily involved in, or critical to, the formation of the active site or chromophore. Our group[9] has suggested that the interactions between the aggregation motifs, through cross β-sheet-based intermolecular interactions, sustain the particular amyloid-like molecular architecture of inclusion bodies. The evidence from functional analysis is corroborated by

structural analyses that show inclusion bodies can be highly enriched in native-like secondary structure (reviewed by A. Ventura and A.V. in ref. 4).

The occurrence of active polypeptides as structural components of inclusion bodies indicates that solubility and functionality are not necessarily associated. In a 1999 study, Waldo and coworkers[10] analyzed the fluorescence of 20 different GFP fusion proteins. For nine of them, the total fluorescence of producing cells correlated well with the solubility of the nonfused heterologous partner, indicating that the aggregation tendency of the aggregation-prone domain modulated the conformational status of the chromophore in the whole fusion. Similarly, collaborators at our institute[8] have recently described how 20 amyloid beta 42 peptide (Aβ42)–GFP fusions containing single point mutations in Aβ42 show variable specific fluorescent emission when deposited as inclusion bodies. Interestingly, fluorescence of these inclusion bodies is inversely dependent on the aggregation rates predicted for the protein variants, indicating that the conformational quality of inclusion bodies depends on how fast the aggregation occurs after protein synthesis. Unfortunately, the study by Waldo and colleages[10] did not investigate the solubility of the GFP

fusions themselves nor the contribution of their inclusion body versions to total cell fluorescence.

Looking at the problem from another perspective, our work at the institute[7,11,12] has shown that the specific activities of enzymes or fluorescence emission properties of fluorescent reporter proteins present as so-called soluble aggregates (oligomeric versions of recombinant proteins) are highly variable, depending on the production conditions and/or the host cell genetic background. A recent structural analysis of soluble GFP fusions has shown that that they assume a wide spectrum of forms, ranging from amorphous particles to true fibers[13]. These observations indicate that the soluble cell fraction contains inactive or partially active protein forms that might be related to the soluble aggregates.

The combined occurrence of functional protein in inclusion bodies and inactive protein versions in the soluble cell fraction (Fig. 1b,c) results in very similar specific activities when comparing soluble and inclusion body protein forms[7,11], especially in the absence of functional DnaK[12]. This chaperone controls the partitioning of functional proteins between soluble and insoluble cell fractions, probably through the selective removal of active, better-folded species from inclusion bodies. Although

CORRESPONDENCE

other explanations are also plausible, we have posited that this hypothesis is compatible with the strong affinity shown by DnaK for the inclusion body surface (A.V. and M.M. Carrio[14]) and the absence of functional protein in the external layers of these aggregates (A.V., E.G.-F. & A. Aris[16]). Although in wild-type cells, the specific activity (or fluorescence emission) of soluble proteins is up to eight times higher than that of proteins in inclusion bodies (still surprisingly close), such values tend to be equal in DnaK-deficient cells[11,12].

As our recent results have revealed that the conformational quality and functionality of protein—whether soluble or insoluble—increase in parallel to reduced culture growth temperature or to elevated intracellular concentrations of critical chaperones[7,16], recombinant proteins expressed in inclusion bodies are representative of the average protein quality in heterologous cells (Fig. 1b,c). In this context, as protein quality correlates only poorly with occurrence in the soluble fraction, solubility should

be reconsidered as a universal indicator of protein quality. What's more, greater judicious discrimination between conformational quality and solubility would likely benefit the optimization of production processes for higher yields of functional polypeptides.

Nuria González-Montalbán,
Elena García-Fruitós & Antonio Villaverde

Institute for Biotechnology and Biomedicine,
Department of Genetics and Microbiology and
CIBER in Bioengineering, Biomaterials and
Nanomedicine, Autonomous University
of Barcelona, Bellaterra, 08193 Barcelona,
Spain.
e-mail: avillaverde@servet.uab.es

1. Baneyx, F. & Mujacic, M. Nat. Biotechnol. 22, 1399–1408 (2004).
2. Sorensen, H.P. & Mortensen, K.K. Microb. Cell Fact. 4, 1 (2005).
3. Worrall, D.M. & Goss, N.H. Aust. J. Biotechnol 3, 28–32 (1989).
4. Ventura, S. & Villaverde, A. Trends Biotechnol. 24, 179–185 (2006).
5. Garcia-Fruitos, E. et al. Microb. Cell Fact. 4, 27 (2005).
6. Arie, J.P., Miot, M., Sassoon, N. & Betton, J.M. Mol. Microbiol. 62, 427–437 (2006).
7. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. Biotechnol. Bioeng. 96, 1101–1106 (2007).
8. de Groot, N.S. & Ventura, S. J. Biotechnol. 125, 110–113 (2006).
9. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. J. Mol. Biol. 347, 1025–1037 (2005).
10. Waldo, G.S., Standish, B.M., Berendzen, J. & Terwilliger, T.C. Nat. Biotechnol. 17, 691–695 (1999).
11. Garcia-Fruitos, E., Carrio, M.M., Aris, A. & Villaverde, A. Biotechnol. Bioeng. 90, 869–875 (2005).
12. Gonzalez-Montalban, N., Garcia-Fruitos, E., Ventura, S., Aris, A. & Villaverde, A. Microb. Cell Fact. 5, 26 (2006).
13. de Marco, A. & Schroedel, A. BMC Biochem. 6, 10 (2005).
14. Carrio, M.M. & Villaverde, A. J. Bacteriol. 187, 3599–3601 (2005).
15. Garcia-Fruitos, E., Aris, A., & Villaverde, A. Appl. Environ. Microbiol. 73, 289–294 (2007).
16. Martinez-Alonso, M., Vera, A. & Villaverde, A. FEMS Microbiol. Lett., published online 11 June 2007, doi: 10.111/j.1574-6968.2007.00788.x

VIII.2 Annex II.B

The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body forming cells

Núria González-Montalbán, Elena García-Fruitós, Salvador Ventura, Anna Arís and Antonio Villaverde

Microbial Cell Factories, Vol. 5, August 7, 2006

Microbial Cell Factories

Research

The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body-forming cells

Nuria González-Montalbán[1,2], Elena García-Fruitós[1,2], Salvador Ventura[1,3], Anna Arís[1,2] and Antonio Villaverde*[1,2]

Address: [1]Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain, [2]Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain and [3]Departament de Bioloquímica i de Biologia Molecular, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

Email: Nuria González-Montalbán - nuria.gonzalez.montalban@uab.es; Elena García-Fruitós - Elena.Garcia.Fruitos@uab.es; Salvador Ventura - salvador.ventura@uab.es; Anna Arís - anna.aris@uab.es; Antonio Villaverde* - avillaverde@servet.uab.es

* Corresponding author

Abstract

Background: The molecular mechanics of inclusion body formation is still far from being completely understood, specially regarding the occurrence of properly folded, protein species that exhibit natural biological activities. We have here comparatively explored thermally promoted, *in vivo* protein aggregation and the formation of bacterial inclusion bodies, from both structural and functional sides. Also, the status of the soluble and insoluble protein versions in both aggregation systems have been examined as well as the role of the main molecular chaperones GroEL and DnaK in the conformational quality of the target polypeptide.

Results: While thermal denaturation results in the formation of heterogeneous aggregates that are rather stable in composition, protein deposition as inclusion bodies renders homogenous but strongly evolving structures, which are progressively enriched in the main protein species while gaining native-like structure. Although both type of aggregates display common features, inclusion body formation but not thermal-induced aggregation involves deposition of functional polypeptides that confer biological activity to such particles, at expenses of the average conformational quality of the protein population remaining in the soluble cell fraction. In absence of DnaK, however, the activity and conformational nativeness of inclusion body proteins are dramatically impaired while the soluble protein version gains specific activity.

Conclusion: The chaperone DnaK controls the fractioning of active protein between soluble and insoluble cell fractions in inclusion body-forming cells but not during thermally-driven protein aggregation. This cell protein, probably through diverse activities, is responsible for the occurrence and enrichment in inclusion bodies of native-like, functional polypeptides, that are much less represented in other kind of protein aggregates.

Background

In bacteria, formation of inclusion bodies is common during overexpression of plasmid-encoded recombinant genes, and this fact represents an important matter of concern in biotechnology [1]. Like in mammalian aggresomes, inclusion body formation is stimulated when proteolysis is impaired in protease-deficient mutants [2,3], and these protein deposits act as reservoirs of misfolded polypeptide chains [4] for their further refolding or proteolysis [3,5,6]. Bacterial inclusion bodies are dynamic structures, they grow resulting from an unbalanced equilibrium between constant protein deposition and removal that is lost in absence of protein synthesis [4,7]. Intriguingly, they contain significant amounts of protein in a native-like form [8-12], a fact that is reflected by the important extent of biological activity exhibited by inclusion bodies formed by very different target proteins [13-15]. Why active protein is found in inclusion bodies is still controversial, and the mechanics of the aggregation process that involves properly folded polypeptides (or polypeptides with properly folded domains critical for activity) remains obscure. In this context, it has been recently proposed that protein aggregation in bacteria is not an all-or-nothing process [16], since the quality of recombinant proteins extends over a continuum of conformational forms [17], that include soluble aggregates [18,19] and active protein entrapped in true, refractile inclusion bodies [9,13]. The conformational status of the inclusion body protein is influenced, among others, by environmental factors such as the growth temperature [20] and the gene expression strategy [21], but little is known about the role of cellular factors on the quality of protein species in both soluble and insoluble cell fractions.

In this work, we have explored the occurrence of active, properly folded polypeptides in inclusion bodies and in thermally driven aggregates formed by the same protein species, and the influence of the main chaperones DnaK and GroEL in the quality of the deposited polypeptides but also of those remaining in the soluble fraction. Intriguingly, while both type of aggregates display a few common physiological traits, the occurrence of active protein species is much higher in inclusion bodies, at expenses of a poorer quality (when compared to thermal aggregates) of the protein population remaining in the soluble fraction. Also, the chaperone DnaK has a main role in the distribution of active polypeptides between soluble and insoluble cell fractions in inclusion body forming cells but not during thermally driven protein aggregation.

Results

Composition of β-galactosidase-based thermal aggregates and inclusion bodies

E. coli β-galactosidase is a huge, homotetrameric enzyme formed by the *lacZ* gene product. When overproduced in bacteria, the enzyme remains soluble in the cell cytoplasm and is clearly functional. In an engineered version of the enzyme, the VP1LAC fusion, the presence of a small viral capsid protein at the amino terminus promotes aggregation as cytoplasmic inclusion bodies, and VP1LAC is distributed in the soluble and insoluble cell fractions at comparable proportions [22]. Interestingly, VP1LAC inclusion bodies are enzymatically active [13] at an extend not much different than that found in the soluble protein version [21]. To compare the performance of the enzyme in either thermal aggregates and inclusion bodies, we have used a particular thermo-inducible expression system that enables a comparative study. Expression of both *lacZ* and *VP1LAC* genes was triggered from a temperature-inducible plasmid vector encoding a temperature sensitive lambda repressor, essentially inactive at 42° [23]. The temperature shift from 28 to 42° induced efficient recombinant protein production (without signs of cell toxicity) (Figure 1A). The lower amounts of β-galactosidase compared to that of VP1LAC (Figure 1B) were probably caused by a slightly higher proteolytic sensitivity of the parental protein as previously reported [24]. Under this conditions, cells undergo a mild heat shock that results in thermal denaturation and aggregation of cellular proteins. In particular, the production of the misfolding prone VP1LAC resulted in its accumulation as inclusion bodies [4]. Also, a small part of the recombinant β-galactosidase present in the cells (up to around 5%) was found in the insoluble cell fraction as part of thermal aggregates, and this figure remained nearly constant throughout the heat shock (Figure 2A). In contrast, a progressively higher fraction of VP1LAC (up to 45% at 3 h) occurred as inclusion bodies (Figure 2A). Despite at 42°C the recombinant β-galactosidase is the most abundantly produced protein in the cell, the enzyme only represented around 3% of the protein species found in the insoluble cell fraction, while VP1LAC accounted for 90% of the inclusion body material (Figure 2B). During the experiment time, inclusion bodies were steadily enriched with VP1LAC species and therefore their homogeneity dramatically increased, while the β-galactosidase fraction in thermal aggregates randomly fluctuated between 1.5 and 3%. These results are compatible with the seeding process recently shown to drive inclusion body formation [9] and indicate that, in contrast, thermal aggregation does not involve interaction between homologous protein patches and it is not, at least strictly, sequence-specific. On the other hand, polypeptides embedded in both kinds of aggregates undergo important changes in their global secondary structure (Figure 3; Table 1), through the continuous formation of
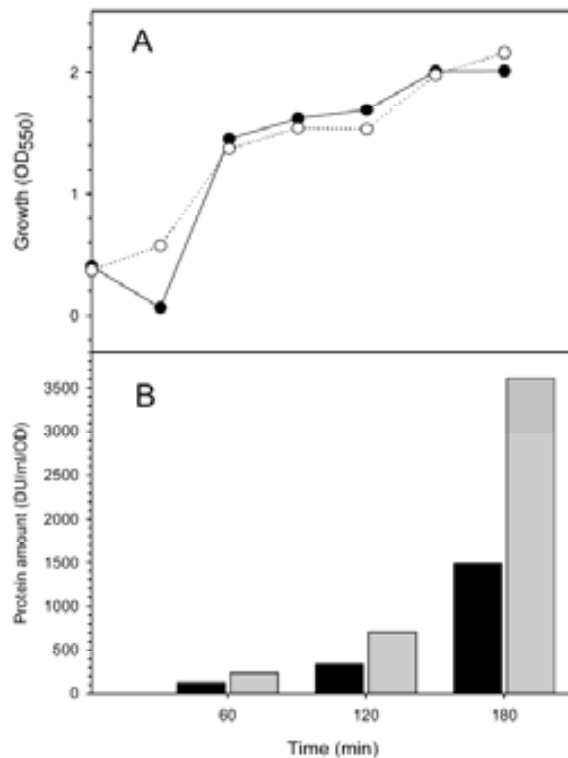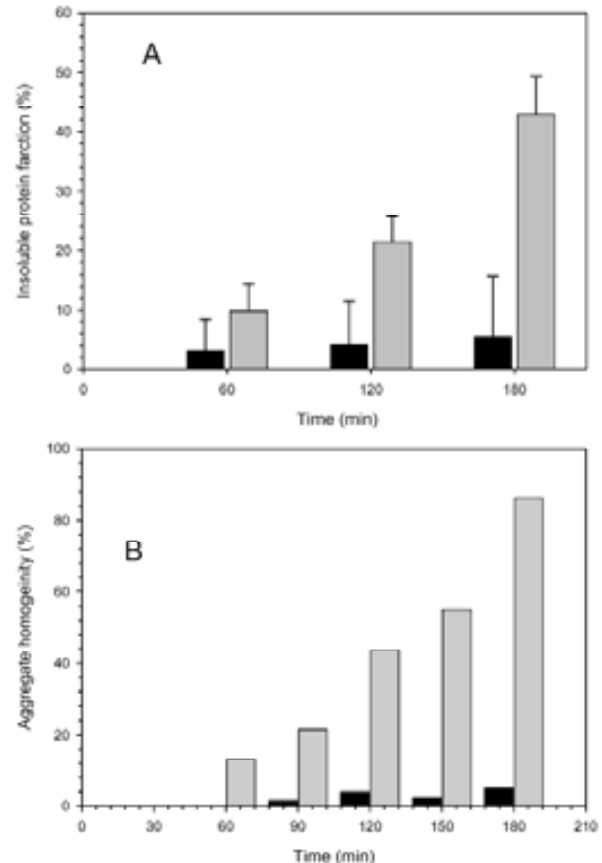
Figure 1
A. Cell growth measured through optical density for MC4100 cultures producing either β-galactosidase (black symbols) or VP1LAC (white symbols). Time 0 represents the temperature up shift. B. Total yield of β-galactosidase (black bars) and VP1LAC (grey bars), as measured by Western blot densitometric units.

Figure 2
A. Fraction of the produced recombinant polypedptides found in protein deposits, either thermal aggregates of β-galactosidase (black bars) or VP1LAC inclusion bodies (grey bars). B. Percentage of β-galactosidase (black bars) and VP1LAC (grey bars) found in thermal aggregates and inclusion bodies respectively.

extended, intermolecular β-sheet structure, being more pronounced in inclusion bodies than in thermal aggregates. This was deduced from the evolution of the bands at 1627 cm⁻¹ and 1692 cm⁻¹ (β-sheet) relative to that at 1652 cm⁻¹ (disordered and/or α-helix) (Table 1). The presence of a band at 1638–1640 cm⁻¹, even if not well resolved, can be attributed to the occurrence of some intramolecular β-sheet. This band appeared only in aged inclusion bodies and it was absent in thermal aggregates. According to previous analysis [9] this peak corresponds to native-like species, that could be accounted by β-galactosidase moieties.

Impact of DnaK and GroEL in β-galactosidase aggregation and activity

The formation of β-galactosidase thermal aggregates and VP1LAC inclusion bodies was explored in absence of the main cytoplasmic chaperones, either DnaK or GroEL. It has been previously reported that when DnaK is not avail-

able, inclusion bodies are larger than in the wild type strain and the amounts of soluble VP1LAC much lower [25]. Such alteration in inclusion body formation can be accounted for by two described DnaK activities, namely preventing aggregation [26] or actively disaggregating proteins [27-29], both done in combination with other chaperones and small heat shock proteins. As observed in the DnaK⁻ background (Figure 4), the deposition of the recombinant enzyme was enhanced in both types of aggregates, although the negative impact on solubility was dramatically higher in those formed by the parental form of the enzyme. The parallel stimulation of aggregation would indicate that DnaK is managing both thermal aggregates and inclusion bodies, although the chaperone could be more active in controlling deposits of denatured
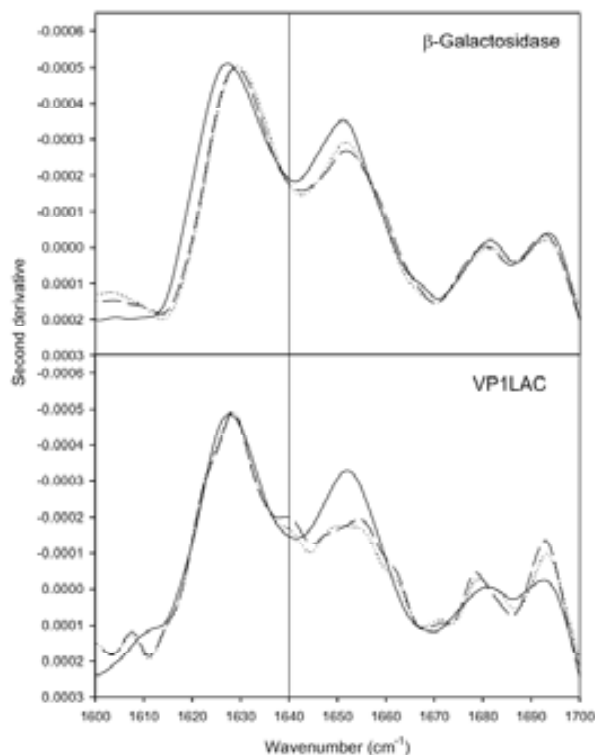
Figure 3
FTIR of β-galactosidase aggregates (top) and VP1LAC inclusion bodies (bottom) formed during either 1 (continuous), 3 (dotted) or 5 (dashed) hours. The vertical line at 1640 cm⁻¹ indicates the position of the band that can be attributed to intramolecular β-sheet.

Table 1: Time evolution of the secondary structure in both β-galactosidase thermal aggregates and VP1LAC inclusion bodies as measured by FTIR peak ratios.

Protein[a]	Time (h)	Ratio 1627/1652[b]	Ratio 1692/1652[b]
β-galactosidase	1	1.31	0.42
	3	1.53	0.40
	5	1.63	0.52
VP1LAC	1	1.36	0.54
	3	2.14	0.83
	5	2.21	0.96

[a]Data are from Figure 3
[b]Peaks at 1627 cm⁻¹ and 1692 cm⁻¹ can be attributed to extended intermolecular β-sheet while that at 1652 to disordered structure and/or α-helix.

As it has recently been proven that deposition as bacterial inclusion bodies does not necessarily represent functional protein inactivation [13], the specific activity of both model proteins was investigated in wild type cells and in absence of either DnaK or functional GroEL. As expected (Table 2), the soluble β-galactosidase was more active (from 2 to 8 fold) than the soluble VP1LAC. Despite this fact, protein aggregated as inclusion bodies was much more active than that occurring in thermal aggregates (up to 10 fold in wild type cells), indicating a higher occurrence of properly folded protein. While GroEL seems to be poorly relevant, this fact is clearly depending on DnaK, since in JGT20, insoluble VP1LAC is around 10 fold less active than insoluble β-galactosidase.

polypeptides. This is suggested by the fact that the amount of insoluble VP1LAC is not even doubled in its absence, while the increase of aggregated β-galactosidase is nine fold higher than that of the wild type enzyme. The presence of a non-functional form of the chaperone GroEL (GroEL44), only had a minor, non-significant impact on protein solubility in both aggregation conditions (Figure 4). The comparative ATR-FTIR of both types of aggregates formed in the mutant strains indicated a different structural pattern compared to the wild type (Figure 5). For VP1LAC in inclusion bodies, the absence of GroEL results in a significant enrichment of native like intramolecular β-sheet structures (peaking at 1638–1640 cm⁻¹). In the case of thermal aggregates the absence of either DnaK or GroEL results in more complex FTIR spectra relative to that recorded for the aggregates formed in the wild-type strain, reflecting a higher degree of conformational heterogeneity.
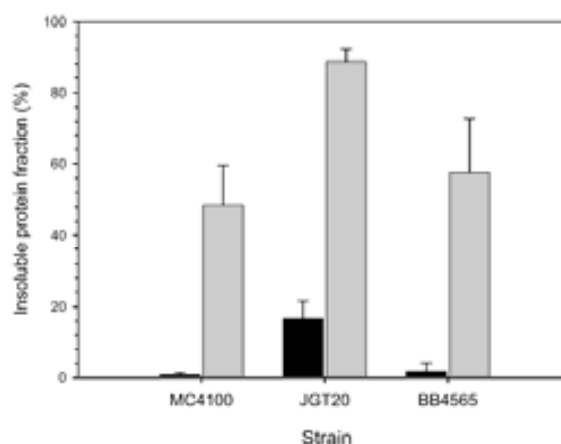

Figure 4
Percentage of the recombinant protein found in protein deposits, either thermal aggregates of β-galactosidase (black bars) or VP1LAC inclusion bodies (grey bars), in MC4100 (wild type). JGT20 (DnaK⁻) and BB4565 (GroEL44) strains.

http://www.microbialcellfactories.com/content/5/1/26
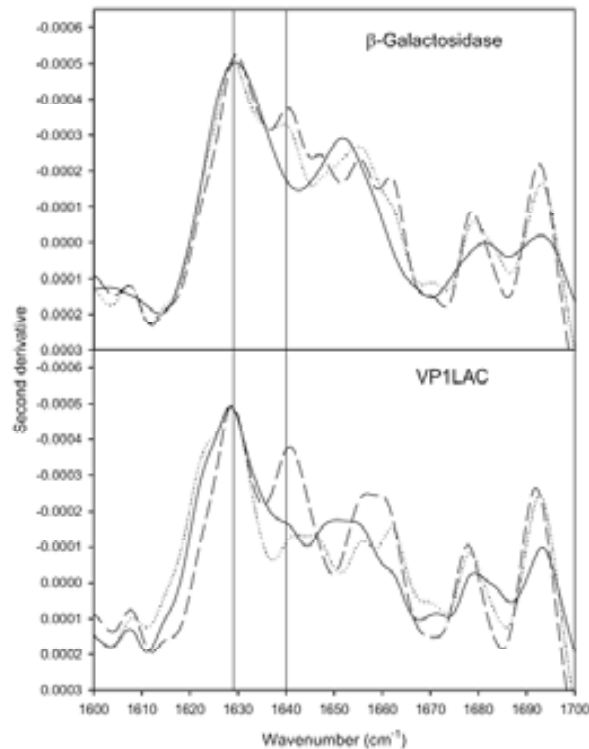
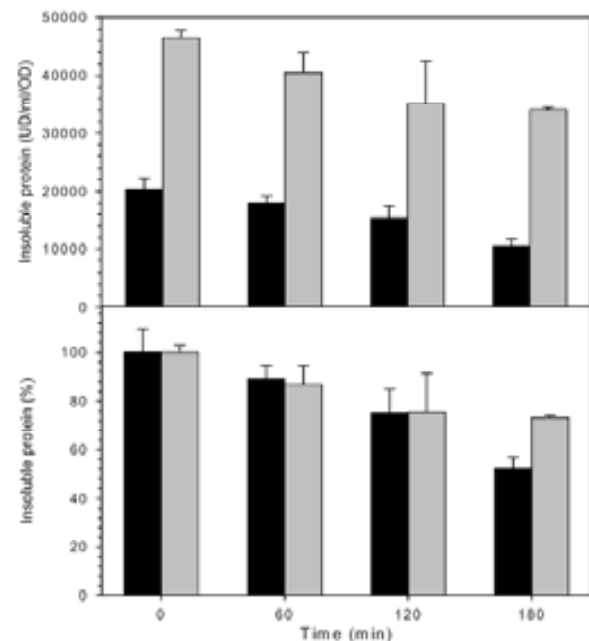

Figure 5
FTIR of β-galactosidase aggregates (top) and VP1LAC inclusion bodies (bottom) formed in MC4100 (continuous), JGT20 (dotted) and BB4565 (dashed) strains. Vertical lines at 1628 and 1640 cm⁻¹ indicate intermolecular β-sheet and intramolecular β-sheet, respectively.

Figure 6
A. Amount of β-galactosidase (black bars) or VP1LAC (grey bars) retained in the insoluble cell fraction after arrest of protein synthesis, as determined by Western blot densitometric units. B. Representation of the above values referred to the starting insoluble material amount.

Physiological disintegration of thermal aggregates and inclusion bodies

The kinetics of physiological disintegration of inclusion bodies and thermal aggregates were compared upon arrest of protein synthesis to investigate the cell ability to process both kinds of structures when chaperones and proteases become available. As shown in Figure 6, the protein removal process is similarly efficient on both aggregate

types, although inclusion body disintegration might be slightly delayed from 3 hours on, with respect to the disintegration of denatured protein clusters.

Discussion
Under mild heat-shock conditions, most of a recombinant β-galactosidase produced in *E. coli* remains in the soluble cell fraction, while an engineered derivative containing an aggregation-prone viral peptide (VP1LAC), forms cytoplasmic inclusion bodies. Up to around 45% of the produced VP1LAC is found trapped in such structures (Figure 2). When comparing with thermal aggregation, the formation of bacterial inclusion bodies appears as a highly specific event, that renders homogenous particles species regarding composition (90% purity in inclusion bodies versus 5% in thermal aggregates, Figure 2). The heterogeneous nature of *in vivo* formed thermal aggregates was not unexpected as many termolabile cellular proteins are deposited as misfolded versions at high temperatures [30]. The high purity of inclusion bodies, however, is reached only 3 hours after inducing gene expression and before that, these particles are progressively gaining homogeneity (Figure 2). In agreement to previous observations [4,5,7], this fact reflects the dynamic nature of

Table 2: Specific activity (in U/ng) of β-galactosidase and its derivative VP1LAC produced in different strains, in the soluble and insoluble cell fractions.

Strain	Soluble fraction	Insoluble fraction
MC4100/pJCO46	628.2 ± 40.5	6.3 ± 0.3
MC4100/pJVP1LAC	234.1 ± 52.9	65.2 ± 19.4
BB4565/pJCO46	689.7 ± 164.9	63.6 ± 2.2
BB4565/pJVP1LAC	230.2 ± 25.7	129.6 ± 45.9
JGT20/pJCO46	888.9 ± 179.3	175.2 ± 34.9
JGT20/pJVP1LAC	12.5 ± 3.8	10.3 ± 6.3

inclusion bodies versus the poor evolution of β-galactosidase present in thermal aggregates, despite this protein is much more abundant in the cell than any of the other deposited species. The seeding mechanics of inclusion body formation [9] and the sequence-dependent aggregation determinants acting there [9,31] have not been described in thermal aggregation, and their absence could account for the different time-dependent composition patterns.

However, ATR-FTIR analysis shows that polypeptides embedded in both kinds of aggregates undergo a structural evolution during formation (Figure 3, Table 1) that can be seen as a continuous formation of new, non-native, extended intermolecular β-sheet structure, more pronounced in inclusion bodies than in thermal aggregates.

The presence of native-like intramolecular β-sheet structure in inclusion bodies aged 3 and 5 h, and absent in the thermal aggregates (peaking at 1638–1640, Figure 3), would be indicative of the presence of a fraction of properly folded proteins or protein domains, in agreement with previous structural analysis [8,10-12,32]. Also, although aggregation reduces the β-galactosidase activity in both β-galactosidase and VP1LAC (Table 2), many descriptions of biological activity in structurally different inclusion body proteins [13-15,21,33,34] indicate that the presence of active protein could be a general trait of such protein deposits. In fact, we prove here that inclusion body protein is 10-fold more active than its thermally denatured counterpart (Table 2). On the other hand, the disintegration of inclusion bodies and thermal aggregates upon arrest of protein synthesis shows comparable rates (Figure 6). This fact indicates that both aggregate types are under the surveillance of disaggregating chaperones [27-29,35,36] and/or proteases [3,4]. Protein removal in both kind of aggregates also suggests that physiological disaggregation is not specifically involving residual native-like structure, as it occurs also on heat denatured protein in which the presence of properly folded polypeptide backbones cannot be detected (Figure 3). Contrarily, the possibility of refolding (or digestion) specifically targeted towards misfolded polypeptides needs to be explored.

Interestingly, the lack of either GroEL or DnaK major cytosolic chaperones globally enhances the activity of the aggregated proteins in both thermal deposits and inclusion bodies (Table 2). The comparative FTIR analysis of both type of aggregates formed in the mutant strains indicates a different general structural pattern compared to the wild type (Figure 5). Aggregates formed in the absence of chaperones are more heterogeneous than those in the wild type strain. The presence of native-like intramolecular β-sheet structure (peaking at 1638–1640 cm⁻¹), corresponding to native-like VP1LAC in inclusion bodies is

enriched specially in the absence of functional GroEL. This coincides with an increased activity of this aggregates, suggesting that this signal corresponds to the accumulation of native and functional β-galactosidase [9]. For thermal aggregates, the presence of a band in the region assignable to intramolecular β-sheet conformations is also detected in the absence of both chaperones. Although, due to the heterogeneous composition of this aggregates, the band cannot attributed to a unique protein species, the significant increased enzymatic activity exhibited by thermal aggregates produced in the absence of chaperones suggests that native functional β-galactosidase contributes, at least partially, to this band in the FTIR spectra.

On the other hand, the specific activity of soluble VP1LAC is between 2 and 3 fold lower than that of the parental enzyme (for wild type and GroEL44 strains), as it would be expected for a fusion protein. However, in absence of DnaK, soluble VP1LAC (but not β-galactosidase) is much more inactive, indicating that this chaperone importantly participates in the VP1LAC (but not β-galactosidase) folding process as previously suggested [21]. Also, the specific activity of inclusion body VP1LAC is surprisingly higher than that of denatured β-galactosidase, only when DnaK is present (Table 2). This intriguing observation indicates an enrichment of inclusion body active species in which DnaK might have a positive role. It cannot be discarded that DnaK, acting as a disaggregase at inclusion body's surface [37], could selectively remove inactive (misfolded) protein. Alternatively, DnaK could preferentially prevent the incorporation of inactive protein into inclusion bodies. In the case of β-galactosidase, the presence of DnaK modulates the deposition of the enzyme under heat stress, as shown by the nine fold increase of β-galactosidase in the aggregated fraction in the absence of this chaperone. The low activity and amount of β-galactosidase in thermal aggregates suggest that they are formed by highly aggregation-prone protein conformations which escape DnaK control. In a DnaK⁻ background, this control does not longer exist and a more heterogeneous set of polypeptide conformations, including some functional or partially functional ones, can aggregate as thermal deposits. This is in accordance both with the higher conformational heterogeneity, as seen by FTIR, and the higher activity of thermal aggregates in the absence of DnaK.

Altogether, these observations point out significant differences between inclusion body formation and *in vivo* thermal aggregation, as revealed by a convenient comparative expression system. While both types of aggregates are controlled by the quality cell system, inclusion bodies are homogeneous and highly organized structures progressively enriched in properly folded versions of the main protein component.

Conclusion

The formation of both protein deposits induced in bacteria by heat shock and inclusion bodies is negatively controlled by DnaK, and both type of aggregates efficiently disintegrate when the conformational stress is over. Despite such similarities, inclusion bodies are more homogeneous in composition and result progressively enriched in native-like forms of the target protein during their construction, what results in a detectable evolution of the global secondary structure of the embedded proteins. In this regard, precipitation as inclusion bodies keeps the target protein in a more functional form than in thermal aggregates, but only when DnaK is present. Interestingly, the biological activity of the soluble counterparts is especially poor when inclusion bodies are more active, suggesting that active polypeptides from the soluble cell fraction are used for inclusion body construction. Therefore, this particular chaperone is important to ensure the biological activities of inclusion body polypeptides that are not conserved in other aggregation conditions, by controlling the distribution of functional protein species between soluble and insoluble cell fractions. Protein packaging as bacterial inclusion bodies is then a cell driven deposition process.

Methods

Bacterial strains, plasmids, proteins and gene expression conditions

Recombinant proteins were produced in *Escherichia coli* MC4100 *araD139* Δ(*argF-lac*) *U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*, and their derivatives GroEL44 *groEL44 zdj::Tn10 zje::kan* (BB4565) and DnaK⁻ *dnaK thr::Tn10* (JTG20). Plasmid pJCO46 encodes a soluble, pseudo-wild type *E. coli* β-galactosidase, and the closely related pJVP1LAC, a derivative β-galactosidase fusion protein containing the aggregation-prone VP1 capsid protein of foot-and-mouth disease virus joined at the amino terminus [22]. The presence of the viral protein segment promotes aggregation of the whole fusion and under our gene expression conditions, approximately 50% of VP1LAC is found as cytoplasmic inclusion bodies. Both *lacZ* and *VP1LAC* genes are under the control of tandem lambda $p_L p_R$ lytic promoters and repressed by a plasmid-encoded and constitutively expressed temperature-sensitive CI[857] repressor. Bacterial cells were cultured in shake flasks up an OD_{550} of 0.3, in Luria-Bertani (LB) rich medium [38] with 100 μg/ml ampicillin. Then, the expression of both *lacZ* and *VP1LAC* genes was triggered by temperature up-shift from 28 to 42°C. When required, protein synthesis was arrested by adding chloramphenicol at 200 μg/mL and the cultures were further incubated at 28°C. Usually, data were obtained from three or more independent experiments.

Quantitative protein analysis

Samples of bacterial cultures (10 ml) were low-speed centrifuged (15 min at 12000 g) and cell pellets resuspended in denaturing buffer. For the analysis of soluble and insoluble cell fractions, samples were resuspended in 500 μl of Z buffer without β-mercaptoethanol [39] with one tablet of protease inhibitor cocktail (Roche, ref. 1 836 170) per 10 ml buffer. Such mixtures, once jacketed in ice, were sonicated for a minimum of 5 min at 50 W under 0.5 s cycles, and centrifuged for 15 min at 12000 g. Soluble and insoluble fractions were separately resuspended in denaturing buffer [40] for Western Blot and Coomassie blue staining. After boiling for 20 min, small sample volumes were loaded onto gels. For Western blot, a rabbit anti β-galactosidase sera was used to immunodetect both β-galactosidase and VP1LAC proteins. Full-length forms of VP1LAC and its major proteolysis fragments (both know to be functional) were considered in the analysis. Dried gels and blots were scanned at high resolution and bands quantified by using the Quantity One software of Bio Rad. All determinations were done at least in quadruplicate.

Conformational analysis by ATR-FTIR spectroscopy

For ATR-FTIR spectroscopy analysis, inclusion bodies and thermal aggregates were purified from cell extracts by repeated detergent washing as described [41]. Then, both kinds of aggregates were dried for two hours in a Seed-Vac system before analysis to reduce water interference in the infrared spectra. A Bruker Tensor 27 FT-IR Spectrometer (Broker Optics Inc.) with a Golden Gate MKII ATR accessory (Specac) was employed for ATR FT-IR experiments. Each spectrum comprises 16 scans measured at a spectral resolution of 4 cm⁻¹ in the 4000–600 cm⁻¹ range. Spectral data were acquired with OPUS MIR Tensor 27 software version 4.0 (Broker Optics Inc.). All the absorbance spectra were normalized to correct for concentration dependent effects and the second derivatives of the amide I band spectra were used to determine the frequencies at which the different spectral components were located.

Determination of the specific activity

To determine the specific activity of both soluble and aggregated β-galactosidase and VP1LAC proteins, 2.5 ml culture samples were disrupted by sonication as described [42] and centrifuged for 15 min at 15000 g. The soluble fraction was directly used for the analysis, and inclusion bodies and thermal aggregates were purified from cell extracts by repeated detergent washing [41]. Substrate hydrolysis was quantified espectrophotometrically as described [21] and the amounts of recombinant protein either soluble or within the aggregates, was specifically determined by Western blot as indicated above, by using serial dilutions of a commercial β-galactosidase of known concentration as pattern. All determinations were done in triplicate.

http://www.microbialcellfactories.com/content/5/1/26

Abbreviations
ATR Attenuated total reflection

FTIR Fourier transformed infrared

LB Luria-Bertani

Competing interests
The author(s) declare that they have no competing interests.

Authors' contributions
NGM performed most of the experimental and Figures, EGF analysed the biological (and specific) activities of recombinant proteins, SV and AA designed the experimental and analysed structural data and AV directed the work and prepared the manuscript.

Acknowledgements

References
1. Baneyx F, Mujacic M: **Recombinant protein folding and misfolding in Escherichia coli.** *Nat Biotechnol* 2004, **22**:1399-1408.
2. Rosen R, Biran D, Gur E, Becher D, Hecker M, Ron EZ: **Protein aggregation in Escherichia coli: role of proteases.** *FEMS Microbiol Lett* 2002, **207**:9-12.
3. Vera A, Aris A, Carrio M, Gonzalez-Montalban N, Villaverde A: **Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies.** *J Biotechnol* 2005, **119**:163-171.
4. Carrio MM, Villaverde A: **Protein aggregation as bacterial inclusion bodies is reversible.** *FEBS Lett* 2001, **489**:29-33.
5. Carrio MM, Corchero JL, Villaverde A: **Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms.** *Biochim Biophys Acta* 1999, **1434**:170-176.
6. Corchero JL, Cubarsi R, Enfors S, Villaverde A: **Limited in vivo proteolysis of aggregated proteins.** *Biochem Biophys Res Commun* 1997, **237**:325-330.
7. Carrio MM, Villaverde A: **Construction and deconstruction of bacterial inclusion bodies.** *J Biotechnol* 2002, **96**:3-12.
8. Ami D, Bonecchi L, Cali S, Orsini G, Tonon G, Doglia SM: **FT-IR study of heterologous protein expression in recombinant Escherichia coli strains.** *Biochim Biophys Acta* 2003, **1624**:6-10.
9. Carrio M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S: **Amyloid-like properties of bacterial inclusion bodies.** *J Mol Biol* 2005, **347**:1025-1037.
10. Oberg K, Chrunyk BA, Wetzel R, Fink AL: **Nativelike secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR.** *Biochemistry* 1994, **33**:2628-2634.
11. Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM: **Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy.** *FEBS Lett* 2005, **579**:3433-3436.
12. Ami D, Natalello A, Taylor G, Tonon G, Maria DS: **Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy.** *Biochim Biophys Acta* 2006, **1764**:793-799.
13. Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A, et al.: **Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins.** *Microb Cell Fact* 2005, **4**:27.
14. Worrall DM, Goss NH: **The formation of biologically active beta-galactosidase inclusion bodies in Escherichia coli.** *Aust J Biotechnol* 1989, **3**:28-32.
15. Tokatlidis K, Dhurjati P, Millet J, Beguin P, Aubert JP: **High activity of inclusion bodies formed in Escherichia coli overproducing Clostridium thermocellum endoglucanase D.** *FEBS Lett* 1991, **282**:205-208.
16. Ventura S, Villaverde A: **Protein quality in bacterial inclusion bodies.** *Trends Biotechnol* 2006, **24**:179-185.
17. de Marco A, Schroedel A: **Characterization of the aggregates formed during recombinant protein expression in bacteria.** *BMC Biochem* 2005, **6**:10.
18. Sorensen HP, Mortensen KK: **Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.** *Microb Cell Fact* 2005, **4**:1.
19. Sorensen HP, Mortensen KK: **Advanced genetic strategies for recombinant protein expression in Escherichia coli.** *J Biotechnol* 2005, **115**:113-128.
20. Jevsevar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V: **Production of nonclassical inclusion bodies from which correctly folded protein can be extracted.** *Biotechnol Prog* 2005, **21**:632-639.
21. Garcia-Fruitos E, Carrio MM, Aris A, Villaverde A: **Folding of a misfolding-prone beta-galactosidase in absence of DnaK.** *Biotechnol Bioeng* 2005, **90**:869-875.
22. Corchero JL, Viaplana E, Benito A, Villaverde A: **The position of the heterologous domain can influence the solubility and proteolysis of beta-galactosidase fusion proteins in E. coli.** *J Biotechnol* 1996, **48**:191-200.
23. Villaverde A, Benito A, Viaplana E, Cubarsi R: **Fine regulation of cI857-controlled gene expression in continuous culture of recombinant Escherichia coli by temperature.** *Appl Environ Microbiol* 1993, **59**:3485-3487.
24. Corchero JL, Villaverde A: **Tolerance of Escherichia coli beta-galactosidase C-terminus to different-sized fusions.** *Biotechnol Bioeng* 1999, **64**:644-649.
25. Carrio MM, Villaverde A: **Role of molecular chaperones in inclusion body formation.** *FEBS Lett* 2003, **537**:215-221.
26. Mujacic M, Bader MW, Baneyx F: **Escherichia coli Hsp31 functions as a holding chaperone that cooperates with the DnaK-DnaJ-GrpE system in the management of protein misfolding under severe stress conditions.** *Mol Microbiol* 2004, **51**:849-859.
27. Mogk A, Deuerling E, Vorderwulbecke S, Vierling E, Bukau B: **Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation.** *Mol Microbiol* 2003, **50**:585-595.
28. Schlieker C, Tews I, Bukau B, Mogk A: **Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides.** *FEBS Lett* 2004, **578**:351-356.
29. Mogk A, Schlieker C, Friedrich KL, Schonfeld HJ, Vierling E, Bukau B: **Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK.** *J Biol Chem* 2003, **278**:31033-31042.
30. Mogk A, Tomoyasu T, Goloubinoff P, Rudiger S, Roder D, Langen H, et al.: **Identification of thermolabile Escherichia coli proteins: prevention and reversion of aggregation by DnaK and ClpB.** *EMBO J* 1999, **18**:6934-6949.
31. Speed MA, Wang DI, King J: **Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition.** *Nat Biotechnol* 1996, **14**:1283-1287.
32. Przybycien TM, Dunn JP, Valax P, Georgiou G: **Secondary structure characterization of beta-lactamase inclusion bodies.** *Protein Eng* 1994, **7**:131-136.
33. Kuczynska-Wisnik D, Zurawa-Janicka D, Narkiewicz J, Kwiatkowska J, Lipinska B, Laskowska E: **Escherichia coli small heat shock proteins IbpA/B enhance activity of enzymes sequestered in inclusion bodies.** *Acta Biochim Pol* 2004, **51**:925-931.
34. Lethanh H, Neubauer P, Hoffmann F: **The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant Escherichia coli and delay degradation of inclusion bodies.** *Microb Cell Fact* 2005, **4**:6.
35. Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T, Bukau B: **Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network.** *Proc Natl Acad Sci USA* 1999, **96**:13732-13737.

36. Weibezahn J, Bukau B, Mogk A: **Unscrambling an egg: protein disaggregation by AAA+ proteins.** *Microb Cell Fact* 2004, **3:**1.
37. Carrio MM, Villaverde A: **Localization of chaperones DnaK and GroEL in bacterial inclusion bodies.** *J Bacteriol* 2005, **187:**3599-3601.
38. Sambrook J, Fritsch E, Maniatis T: *Molecular Cloning, A Laboratory Manual* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1989.
39. Miller JH: *Experiments in Molecular Genetics* Cold SpringHarbor Laboratory Press, Cold Spring Harbor, NY; 1972.
40. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, **227:**680-685.
41. Carrio MM, Cubarsi R, Villaverde A: **Fine architecture of bacterial inclusion bodies.** *FEBS Lett* 2000, **471:**7-11.
42. Feliu JX, Cubarsi R, Villaverde A: **Optimized release of recombinant proteins by ultrasonication of E. coli cells.** *Biotechnol Bioeng* 1998, **58:**536-540.

VIII.3 Annex II.C

In situ protein folding and activation in bacterial inclusion bodies

Núria González-Montalbán, Antonino Natalello, Elena García-Fruitós, Antonio Villaverde and Silvia Maria Doglia

Biotechnology and Bioengineering (in press), January 30, 2008

COMMUNICATION TO THE EDITOR

BIOTECHNOLOGY and BIOENGINEERING

In Situ Protein Folding and Activation in Bacterial Inclusion Bodies

Nuria Gonzalez-Montalban,[1] Antonino Natalello,[2] Elena García-Fruitós,[1] Antonio Villaverde,[2] Silvia Maria Doglia[2]

[1]Institute for Biotechnology and Biomedicine, Department of Genetics and Microbiology, Autonomous University of Barcelona and CIBER-BBN Network in Bioengineering, Biomaterials and Nanomedicine, Bellaterra, 08193 Barcelona, Spain; telephone: 34-935812148; fax: 34-935812011; e-mail: avillaverde@servet.uab.es

[2]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy

ABSTRACT: Recent observations indicate that bacterial inclusion bodies formed in absence of the main chaperone DnaK result largely enriched in functional, properly folded recombinant proteins. Unfortunately, the molecular basis of this intriguing fact, with obvious biotechnological interest, remains unsolved. We have explored here two non-excluding physiological mechanisms that could account for this observation, namely selective removal of inactive polypeptides from inclusion bodies or in situ functional activation of the embedded proteins. By combining structural and functional analysis, we have not observed any preferential selection of inactive and misfolded protein species by the dissagregating machinery during inclusion body disintegration. Instead, our data strongly support that folding intermediates aggregated as inclusion bodies could complete their natural folding process once deposited in protein clusters, which conduces to significant functional activation. In addition, in situ folding and protein activation in inclusion bodies is negatively regulated by the chaperone DnaK.
Biotechnol. Bioeng. 2008;xxx: xxx–xxx.

© 2008 Wiley Periodicals, Inc.

KEYWORDS: inclusion bodies; protein folding; DnaK; enzymatic activity

Introduction

Bacterial inclusion bodies (IBs) are amyloid-like protein aggregates usually formed during the overexpression of foreign genes (Villaverde and Carrio, 2003). Since IBs are built up as a result of seeding-driven, sequence-specific protein deposition processes (Carrio et al., 2005) they are highly pure and composed almost exclusively by the recombinant protein itself (Carrio et al., 1998). While for long time these aggregates were believed to be uniquely formed by misfolded versions of recombinant proteins, recent structural data has revealed, for many model proteins, the presence of significant proportions of native-like secondary structure (Ami et al., 2003, 2005, 2006; Gonzalez-Montalban et al., 2006; Oberg et al., 1994; Przybycien et al., 1994). In parallel, the analysis of the biological properties of IBs formed by enzymes and fluorescent proteins reveal enzymatic activities or strong fluorescence respectively (Arie et al., 2006; Garcia-Fruitos et al., 2005b; Kuczynska-Wisnik et al., 2004; Tokatlidis et al., 1991; Worrall and Goss, 1989). Therefore, IBs are composed, at least partially, by functional polypeptides, whose deposition is necessarily driven by discrete aggregation determinants, that act irrespective of the global folding state of the protein (Gonzalez-Montalban et al., 2007). The prevalence and extent of biological activity of IB proteins is variable depending on the protein itself, the genetic background of the producing cells and the protein production conditions (Garcia-Fruitos et al., 2005a,b; Gonzalez-Montalban et al., 2006). Intriguingly, functional polypeptides seem to be excluded from the IBs surface (Garcia-Fruitos et al., 2007a).

In some cases, the specific activity or specific fluorescence emission observed in IBs is very similar to that found in the soluble protein versions (Garcia-Fruitos et al., 2005b; Gonzalez-Montalban et al., 2006). This fact results from the combination of highly active IB proteins and the presence, in the soluble cell fraction, of inactive protein forms, that might be related to the occurrence of the so-called 'soluble aggregates' (de Marco and Schroedel, 2005). In particular, the differences in the biological activity (either enzymatic

activity or fluorescence) of soluble and IB proteins are extremely minimized in cells devoid of the main chaperone DnaK (Garcia-Fruitos et al., 2005a; Gonzalez-Montalban et al., 2006), and in this genetic background, green fluorescent protein aggregated as IBs exhibit specific fluorescence values even higher than those shown by the soluble counterpart (Garcia-Fruitos et al., 2007b). DnaK, an homologous of the eukaryotic Hsp70, has several recognized activities in IB-forming cells, namely preventing aggregation, folding and refolding of misfolded species and protein dissaggregation (Gragerov et al., 1992; Langer et al., 1992; Mogk et al., 2003a,b; Schlieker et al., 2004; Thomas and Baneyx, 1996). The possibility of obtaining highly active IBs in DnaK$^-$ backgrounds is very attractive since it might be a straightforward source of pure, naturally immobilized enzymes ready for catalytic processes (Garcia-Fruitos et al., 2005b, 2007a). However, the molecular basis of the high relative biological activity of IBs in absence of DnaK remains unexplored. Since during active protein production IBs are under continuous reconstruction (Carrio and Villaverde, 2001, 2002; Hoffmann et al., 2001; Rinas et al., 2007), one possibility would be a selective release of highly functional, better folded species by DnaK and their partners in the disaggregase complex, ClpB and IbpAB (Mogk et al., 2003a). This would account for the similar activity of IB and soluble protein versions in absence of DnaK and for the external layer of non-functional species occurring on the IBs surface (Garcia-Fruitos et al., 2007a), were DnaK abounds (Carrio and Villaverde, 2005). However, the present analysis, designed to address this question, does instead support an alternative hypothesis based on the in situ folding or refolding of inclusion body proteins, a process in which DnaK seems to act as a negative modulator.

Materials and Methods

Bacterial Strains, Plasmids, and Purification of Protein Aggregates

Recombinant proteins were produced in *Escherichia coli* MC4100 *araD139* Δ(*argF-lac*) *U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR* (Sambrook et al., 1989), used as wild-type (wt), and in its derivative DnaK$^-$ (*dnaK thr::Tn10*, JTG20) (Thomas and Baneyx, 1998). Plasmid pJCO46 encodes a soluble, pseudo-wt *E. coli* β-galactosidase, and the closely related vector pJVP1LAC, a β-galactosidase fusion protein containing the aggregation-prone VP1 capsid protein of foot-and-mouth disease virus (Corchero and Villaverde, 1998). The presence of the viral protein promotes aggregation of the whole fusion and under our gene expression conditions, approximately 50% of VP1LAC is found as cytoplasmic inclusion bodies. Both *lacZ* and *VP1LAC* genes are under the control of tandem lambda pLpR lytic promoters and repressed by a plasmid-encoded and constitutively expressed temperature-sensitive CI857

repressor (Corchero and Villaverde, 1998). Bacterial cells were cultured in shake flasks up an OD$_{550}$ of 0.3, in Luria-Bertani (LB) rich medium (Sambrook et al., 1989) with 100 μg/mL ampicillin. The expression of both *lacZ* and *VP1LAC* genes was then triggered by temperature up-shift from 28 to 42°C. When required, protein synthesis was arrested by adding chloramphenicol at 200 μg/mL and the cultures were further incubated at 28°C. Data obtained from three or more independent cultures was used for further analysis. IBs were purified by repeated detergent washing as described (Carrio et al., 2000) and properly stored at −20°C until use.

Quantitative Protein Analysis

Samples of IBs deriving from 1 mL of culture were resuspended in denaturing buffer (Sambrook et al., 1989) for Western Blot. After boiling for 20 min, appropriate sample volumes were loaded onto SDS–polyacrilamide gels. A rabbit anti-β-galactosidase serum was used to immuno-detect and quantify both β-galactosidase and VP1LAC proteins, by comparing to serial dilutions of a commercial β-galactosidase of known concentration. Full-length forms of VP1LAC and its major proteolysis fragments (both know to be functional) were considered in the analysis. All determinations were done at least in quadruplicate.

Determination of β-Galactosidase Activity

The analysis of total β-galactosidase activity in inclusion bodies was performed through a variant of Miller's protocol (Miller, 1972) in absence of toluene. Briefly, inclusion bodies were resuspended in 1 mL Z buffer (0.06 M Na$_2$HPO$_4$, 0.04 M NaH$_2$PO$_4$, 0.01 M KCl, 1 mM MgSO$_4$) and 100 μL of the samples were taken and resuspended again in 1 mL of Z buffer. Then, 200 μL of 4 mg/mL *o*-nitrophenyl-β-D-galactopyranoside (ONPG) were added in each sample and the mixture incubated at 28°C until yellow color was apparent (usually a few min). The reaction was stopped with 500 μL 1 M NaCO$_3$ and the amount of colored product was determined spectophotometrically at 420 nm. Enzymatic units were finally calculated according to Miller's (1972) equation. The specific β-galactosidase activity was obtained by dividing enzymatic units by the amounts either β-galactosidase or VP1LAC found in each sample and the final, shown value is the average of those obtained from three or more independent experiments.

FT-IR Measurements

Cells or IBs were extracted from 25 mL-culture samples (for 1-h sample) or from 10 mL-samples (for other samples) and were resuspended in 50 or 200 μL distilled water respectively. Between 5 and 15 μL of these suspensions were deposited on a BaF$_2$ infrared support. After 20 min at room

temperature, the water evaporated resulting in hydrated films of cells or IBs. The FT-IR absorption of these films was measured from 4,000 to 800 cm^{-1} in the transmission mode by an UMA 500 infrared microscope (Orsini et al., 2000) coupled to a FTS 40A spectrometer (both from Bio-Rad, Digilab Division, Cambridge, MA), under conditions previously described (Ami et al., 2005). Concerning the reproducibility of the data (see also Supplementary Figs. 1 and 2), we performed at least three independent cultures (five in many cases) and each cell and extracted IB samples were measured at least three times, for a total of more than 200 produced spectra.

FT-IR Spectral Analysis

The second derivative of the absorption spectrum was obtained to resolve the broad Amide I band, which consists of the overlapping of several components, each due to the C=O absorption in the different secondary structures and aggregates. The negative bands in the derivative spectrum allow identifying the Amide I components (Susi and Byler, 1986). Their relative intensity enables also to evaluate the extent of each component (Dong et al., 1992). For the kinetics studies of IB aggregation and disruption in intact cells, we followed the FT-IR method recently reported (Ami et al., 2003, 2005). Briefly, FT-IR absorption spectra of intact cells taken from the culture broth were examined at different times, during 5 h after induction. The kinetics of aggregation was then monitored within intact cells by evaluating in the second derivative spectra the time dependence of the aggregation band intensity at 1,630 cm^{-1} (Ami et al., 2005). The spectra of intact cells from wt and DnaK$^-$ strains were examined after normalization at the tyrosine band (1,517 cm^{-1}), in order to compensate possible differences in the optical paths and in the total protein content of the samples.

The aggregation percentage was calculated in each strain from the intensity value of the negative band shoulder at 1,630 cm^{-1} in the intact cell second derivative spectra (Ami et al., 2005). The highest aggregation percentage, namely 100%, was attributed to the highest intensity, so that the lowest aggregation percentage was the lowest intensity value at 1,630 cm^{-1} during the culture induction time. Every strain percentages were calculated from the intensity absolute values, after normalization at the tyrosine band (1,517 cm^{-1}).

Results and Discussion

To check any biased removal of inclusion body polypeptides by the disaggregases of the quality control system, we monitored the secondary structure of IBs formed by the misfolding-prone β-galactosidase protein VP1LAC (Carrio et al., 2005), during their disintegration in presence and absence of DnaK, after arrest of protein synthesis (Carrio
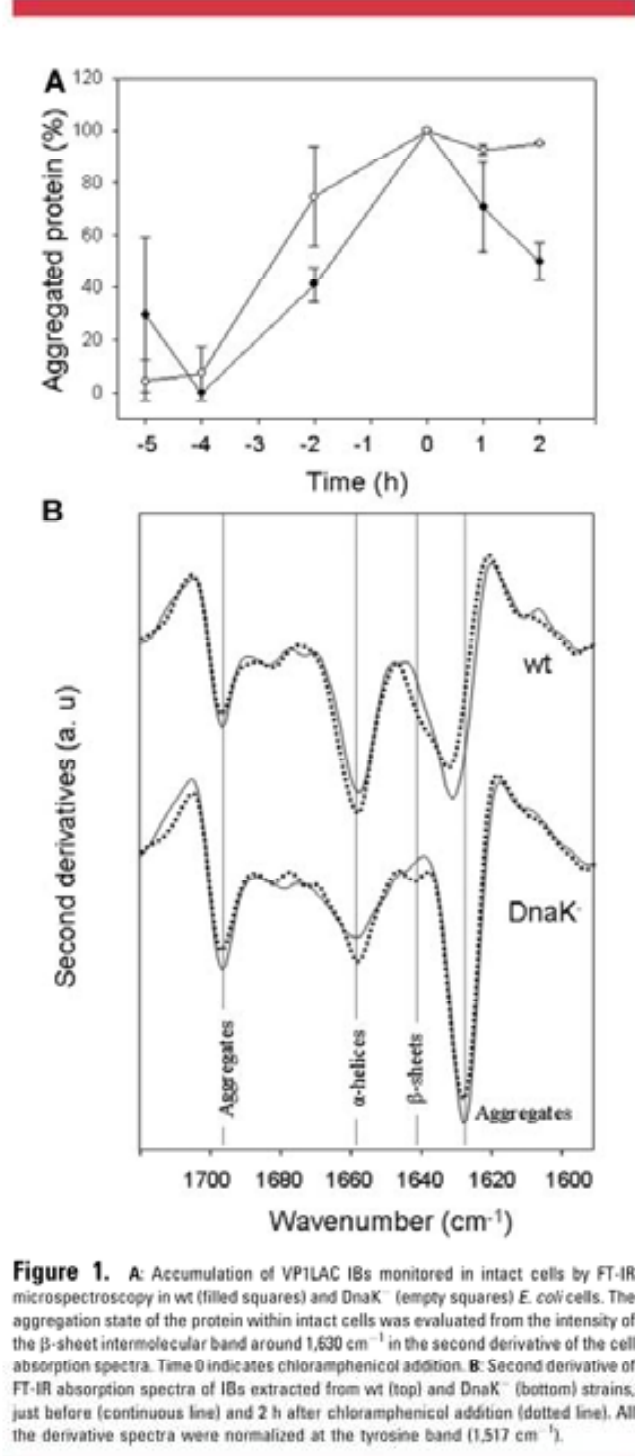
Figure 1. A: Accumulation of VP1LAC IBs monitored in intact cells by FT-IR microspectroscopy in wt (filled squares) and DnaK$^-$ (empty squares) E. coli cells. The aggregation state of the protein within intact cells was evaluated from the intensity of the β-sheet intermolecular band around 1,630 cm^{-1} in the second derivative of the cell absorption spectra. Time 0 indicates chloramphenicol addition. B: Second derivative of FT-IR absorption spectra of IBs extracted from wt (top) and DnaK$^-$ (bottom) strains, just before (continuous line) and 2 h after chloramphenicol addition (dotted line). All the derivative spectra were normalized at the tyrosine band (1,517 cm^{-1}).

and Villaverde, 2001). Under these circumstances, cell chaperones promote the physiological dissolution of IBs in a few hours (Carrio and Villaverde, 2001, 2003). If DnaK-mediated protein removal were preferentially targeted to better folded species we should observe a reduction of native-like secondary structure in IBs from wt but not DnaK$^-$ cells. By Fourier transform infrared spectroscopy methodologies adapted to in vivo IB analysis (Ami et al.,
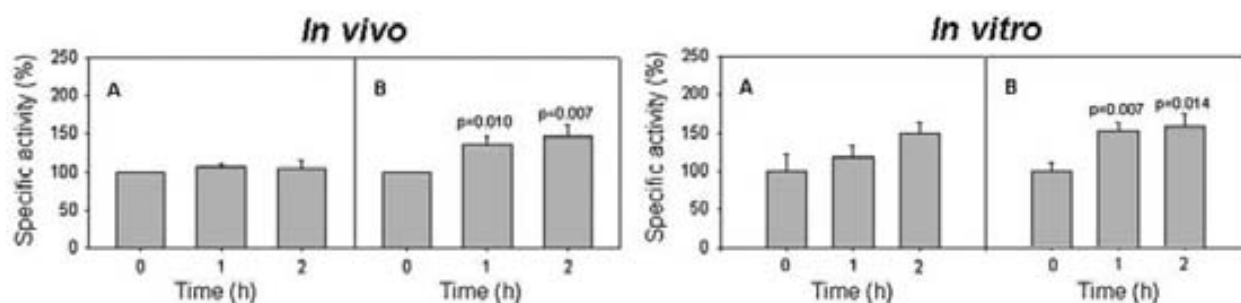
Figure 2. In vivo: β-Galactosidase specific activity of IBs in living cells after chloramphenicol addition formed either in wt (A) or DnaK⁻ cells (B) In vitro: β-Galactosidase specific activity of IBs purified just after chloramphenicol addition, from either wt (A) or DnaK⁻ cells (B). Values derive from six independent experiments, and significant differences data from time 0 are indicated through Pvalues of comparative ANOVA tests. The range of absolute values of β-galactosidase specific activities of IBs formed in wt and DnaK⁻ cells can be found elsewhere (Gonzalez-Montalban et al., 2006).

2003, 2005, 2006), we observed a rapid decrease of the amount of IB protein immediately after chloramphenicol addition in *E. coli* wt cells but not in the DnaK⁻ strain (Fig. 1A). A finest comparative analysis of IB protein secondary structure was done on IBs purified before and after arresting protein synthesis. In both wt and DnaK⁻ cells, the second derivative absorption spectra of such IBs (Fig. 1B) displayed two bands at approximately 1,631 and 1,697 cm⁻¹, due to the antiparallel β-sheet intermolecular interaction that characterizes amyloid-like protein aggregation in IBs (Carrio et al., 2005). In DnaK⁻ cells, the 1,631 cm⁻¹ aggregation band was more pronounced and the wavenumber downshifted of about 3 cm⁻¹, indicative of a more compact IB architecture. This was accompanied by less evident native-like structure (Fig. 1B, bottom), represented by a band at around 1,658 cm⁻¹ (α-helical and unordered structures) and a shoulder at around 1,640 cm⁻¹ (intramolecular β-sheet of the native protein, previously found to correlate with the enzymatic activity of IBs (Gonzalez-Montalban et al., 2006)).

In absence of protein synthesis and contrarily to what we initially expected, the extent of residual native-like structures at around 1,658 and 1,640 cm⁻¹ increased in both wt and DnaK⁻ cells, while the aggregation peak at 1,631 cm⁻¹ decreased moderately and shifted towards higher wavenumbers (of about 1–2 cm⁻¹). These facts indicated a tendency to form native-like secondary structure, what was contrary to the hypothesis of a preferential removal of functional species.

On the other hand, since no protein release from IBs was detected in DnaK⁻ cells (Fig. 1A), any structural modification observed in IBs should be attributed to internal molecular reorganizations. Any structural contribution to IBs architecture of soluble polypeptides newly deposited in absence of protein synthesis should be in principle discarded, since in absence of DnaK, the amounts of soluble VP1LAC are negligible (Carrio and Villaverde, 2003). Therefore, since these results strongly suggested a spontaneous in situ folding of IB polypeptides in absence of DnaK, we explored potential changes in the specific activity of IB

polypeptides, at different times after protein synthesis arrest. This was done by determining both β-galactosidase activity and VP1LAC protein amounts by standard enzymatic assays and quantitative Western blot analysis as previously described (Garcia-Fruitos et al., 2005b). After the arrest of protein synthesis, the specific activity of VP1LAC IBs formed in wt cells remained nearly constant while interestingly, they undergone a significant activation up to more than 50% in absence of DnaK (Fig. 2A). This observation was indeed fully compatible with in situ protein folding of the aggregated β-galactosidase in DnaK⁻ cells. To additionally confirm these data, we monitored the specific β-galactosidase activity of IBs once purified, during incubation for 2 h in a physiological buffer (Garcia-Fruitos et al., 2007a) at 37°C. Again, in vitro, the specific β-galactosidase activity of IBs from DnaK⁻ cells significantly increased around 50% (Fig. 2B), in a spontaneous process that did not require any soluble cellular factor.

Altogether, results presented in Figure 1 prove an in situ folding of IB proteins in absence of DnaK, and strongly suggested that this process also occurs in wt cells. However, only in absence of DnaK it conduces to enzyme activation (Fig. 2). In this context, we have recently shown that a molar excess of DnaK inactivates both soluble and insoluble recombinant protein forms in IB-forming *E. coli* cells (Martinez-Alonso et al., 2007). This is not caused by a cascade effect of DnaK activities but by a direct interaction between DnaK and partially folded protein species. Since the release of DnaK from folding polypeptides is required for a complete refolding to native state (Thomas and Baneyx, 1996), the high concentration of DnaK molecules at the IBs surface might inhibit protein folding. The slight enzymatic activation of wt IBs once purified, not seen in vivo, could be due to a lost of surface-attached DnaK molecules (Carrio and Villaverde, 2005) during purification.

Very recently, we have shown a dramatic inhibition of recombinant protein degradation in DnaK⁻ cells that permits a massive deposition of misfolded protein species targeted for proteases (Garcia-Fruitos et al., 2007b). If folding intermediates are among those targets, their

deposition in IB, that are highly hydrated and porous aggregates might allow them to continue their natural folding process. In combination with the extended half-life and enhanced aggregation of functional proteins, the absence of DnaK acting also as a molecular inhibitor of IB protein activation explains both the similar specific activity of soluble and IB enzyme versions (Garcia-Fruitos et al., 2005a) and the high specific fluorescence (of GFP and BFP) found in DnaK$^-$ when comparing to wt cells (Garcia-Fruitos et al., 2007b).

Finally and very interestingly, the structural reorganization of proteins within bacterial IBs, which we report here for the first time, seems to be a feature common to protein aggregates, even of different types and from different origins. Indeed, it has been described that the amyloid structures of *Sulfolobus solfataricus* acylphosphatase arise from the molecular reorganization of previously formed aggregates, without disaggregation and re-nucleation of the initial protein assemblies (Plakoutsi et al., 2005). Also thermal aggregates of bovine carbonic anhydrase were found to undergo structural reorganization upon cooling, as probed by a fluorescent dye binding investigation (Kundu and Guptasarma, 2002). Similarly, when thermal aggregates are cooled, a downshift of the infrared aggregate band (around 1,628–1,620 cm^{-1}) are often observed, as for instance in the case of recombinant human factor XIII (Dong et al., 1997) and of *Candida rugosa* lipase 1 ((Natalello et al., 2005) and data not shown). This downshift indicates the occurrence of protein reorganization into strong protein–protein interactions. Overall, these results indicate an unexpected conformational plasticity of aggregated proteins, with important biotechnological and medical implications.

References

Ami D, Bonecchi L, Cali S, Orsini G, Tonon G, Doglia SM. 2003. FT-IR study of heterologous protein expression in recombinant *Escherichia coli* strains. Biochim Biophys Acta 1624:6–10.

Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM. 2005. Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. FEBS Lett 579:3433–3436.

Ami D, Natalello A, Taylor G, Tonon G, Doglia SM. 2006. Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. Biochim Biophys Acta 1764:793–799.

Arie JP, Miot M, Sassoon N, Betton JM. 2006. Formation of active inclusion bodies in the periplasm of *Escherichia coli*. Mol Microbiol 62:427–437.

Carrio MM, Villaverde A. 2001. Protein aggregation as bacterial inclusion bodies is reversible. FEBS Lett 489:29–33.

Carrio MM, Villaverde A. 2002. Construction and deconstruction of bacterial inclusion bodies. J Biotechnol 96:3–12.

Carrio MM, Villaverde A. 2003. Role of molecular chaperones in inclusion body formation. FEBS Lett 537:215–221.

Carrio MM, Villaverde A. 2005. Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. J Bacteriol 187:3599–3601.

Carrio MM, Corchero JL, Villaverde A. 1998. Dynamics of in vivo protein aggregation: Building inclusion bodies in recombinant bacteria. FEMS Microbiol Lett 169:9–15.

Carrio MM, Cubarsi R, Villaverde A. 2000. Fine architecture of bacterial inclusion bodies. FEBS Lett 471:7–11.

Carrio M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S. 2005. Amyloid-like properties of bacterial inclusion bodies. J Mol Biol 347: 1025–1037.

Corchero JL, Villaverde A. 1998. Plasmid maintenance in *Escherichia coli* recombinant cultures is dramatically, steadily, and specifically influenced by features of the encoded proteins. Biotechnol Bioeng 58:625–632.

de Marco A, Schroedel A. 2005. Characterization of the aggregates formed during recombinant protein expression in bacteria. BMC Biochem 6: 10.

Dong A, Caughey B, Caughey WS, Bhat KS, Coe JE. 1992. Secondary structure of the pentraxin female protein in water determined by infrared spectroscopy: Effects of calcium and phosphorylcholine. Biochemistry 31:9364–9370.

Dong A, Kendrick B, Kreilgard L, Matsuura J, Manning MC, Carpenter JF. 1997. Spectroscopic study of secondary structure and thermal denaturation of recombinant human factor XIII in aqueous solution. Arch Biochem Biophys 347:213–220.

Garcia-Fruitos E, Carrio MM, Aris A, Villaverde A. 2005a. Folding of a misfolding-prone beta-galactosidase in absence of DnaK. Biotechnol Bioeng 90:869–875.

Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A, Ventura S, Villaverde A. 2005b. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. Microb Cell Fact 4:27.

Garcia-Fruitos E, Aris A, Villaverde A. 2007a. Localization of functional polypeptides in bacterial inclusion bodies. Appl Environ Microbiol 73:289–294.

Garcia-Fruitos E, Martinez-Alonso M, Gonzalez-Montalban N, Valli M, Mattanovich D, Villaverde A. 2007b. Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. J Mol Biol 374:195–205.

Gonzalez-Montalban N, Garcia-Fruitos E, Ventura S, Aris A, Villaverde A. 2006. The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body-forming cells. Microb Cell Fact 5:26.

Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A. 2007. Recombinant protein solubility-does more mean better? Nat Biotechnol 25:718–720.

Gragerov A, Nudler E, Komissarova N, Gaitanaris GA, Gottesman ME, Nikiforov V. 1992. Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. Proc Natl Acad Sci USA 89:10341–10344.

Hoffmann F, Posten C, Rinas U. 2001. Kinetic model of in vivo folding and inclusion body formation in recombinant *Escherichia coli*. Biotechnol Bioeng 72:315–322.

Kuczynska-Wisnik D, Zurawa-Janicka D, Narkiewicz J, Kwiatkowska J, Lipinska B, Laskowska E. 2004. *Escherichia coli* small heat shock proteins IbpA/B enhance activity of enzymes sequestered in inclusion bodies. Acta Biochim Pol 51:925–931.

Kundu B, Guptasarma P. 2002. Use of a hydrophobic dye to indirectly probe the structural organization and conformational plasticity of molecules in amorphous aggregates of carbonic anhydrase. Biochem Biophys Res Commun 293:572–577.

Langer T, Lu C, Echols H, Flanagan J, Hayer MK, Hartl FU. 1992. Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356:683–689.

Martinez-Alonso M, Vera A, Villaverde A. 2007. Role of the chaperone DnaK in protein solubility and conformational quality in inclusion body-forming *Escherichia coli* cells. FEMS Microbiol Lett 273:187–195.

Miller JH. Experiments in molecular genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. 1972.

Mogk A, Deuerling E, Vorderwulbecke S, Vierling E, Bukau B. 2003a. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol Microbiol 50:585–595.

Mogk A, Schlieker C, Friedrich KL, Schonfeld HJ, Vierling E, Bukau B. 2003b. Refolding of substrates bound to small Hsps relies on a dis-aggregation reaction mediated most efficiently by ClpB/DnaK. J Biol Chem 278:31033–31042.

Natalello A, Ami D, Brocca S, Lotti M, Doglia SM. 2005. Secondary structure, conformational stability and glycosylation of a recombinant Candida rugosa lipase studied by Fourier-transform infrared spectro-scopy. Biochem J 385:511–517.

Oberg K, Chrunyk BA, Wetzel R, Fink AL. 1994. Nativelike secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. Biochemistry 33:2628–2634.

Orsini F, Ami D, Villa AM, Sala G, Bellotti MG, Doglia SM. 2000. FT-IR microspectroscopy for microbiological studies. J Microbiol Methods 42:17–27.

Plakoutsi G, Bemporad F, Calamai M, Taddei N, Dobson CM, Chiti F. 2005. Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. J Mol Biol 351:910–922.

Przybycien TM, Dunn JP, Valax P, Georgiou G. 1994. Secondary structure characterization of beta-lactamase inclusion bodies. Protein Eng 7:131–136.

Rinas U, Hoffmann F, Betiku E, Estape D, Marten S. 2007. Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in Escherichia coli. J Biotechnol 127:244–257.

Sambrook J, Fritsch E, Maniatis T. Molecular cloning, a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. 1989.

Schlieker C, Tews I, Bukau B, Mogk A. 2004. Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides. FEBS Lett 578:351–356.

Susi H, Byler DM. 1986. Resolution-enhanced Fourier transform infrared spectroscopy of enzymes. Methods Enzymol 130:290–311.

Thomas JG, Baneyx F. 1996. Protein folding in the cytoplasm of Escherichia coli: Requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines. Mol Microbiol 21:1185–1196.

Thomas JG, Baneyx F. 1998. Roles of the Escherichia coli small heat shock proteins IbpA and IbpB in thermal stress management: Comparison with ClpA, ClpB, and HtpG In vivo. J Bacteriol 180:5165–5172.

Tokatlidis K, Dhurjati P, Millet J, Beguin P, Aubert JP. 1991. High activity of inclusion bodies formed in Escherichia coli overproducing Clostridium thermocellum endoglucanase D. FEBS Lett 282:205–208.

Villaverde A, Carrio MM. 2003. Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. Biotechnol Lett 25:1385–1395.

Worrall DM, Goss NH. 1989. The formation of biologically active beta-galactosidase inclusion bodies in Escherichia coli. Aust J Biotechnol 3:28–32.

VIII.4 Annex II.D

Yield, solubility and conformational quality of soluble proteins cannot be simultaneously favoured in recombinant *Escherichia coli*

Mónica Martínez-Alonso, Elena García-Fruitós and Antonio Villaverde

Biotechnology and Bioengineering

Yield, solubility and conformational quality of soluble proteins cannot be simultaneously favoured in recombinant Escherichia coli

Journal:	Biotechnology and Bioengineering
Manuscript ID:	draft
Wiley - Manuscript type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Martínez-Alonso, Mónica; Universitat Autonoma de Barcelona, Institut de Biotecnologia i de Biomedicina García-Fuitós, Elena; Universitat Autonoma de Barcelona, Institut de Biotecnologia i de Biomedicina Villaverde, Antonio; Institut de Biotecnologia i de Biomedicina, Universitat Autonoma de Barcelona
Key Words:	protein folding, protein quality, solubility, inclusion bodies

scholarONE™
Manuscript Central

1

Yield, solubility and conformational quality of soluble proteins cannot be simultaneously favoured in recombinant *Escherichia coli*.

Mónica Martínez-Alonso, Elena García-Fruitós and Antonio Villaverde*

Institute for Biotechnology and Biomedicine, Department of Genetics and Microbiology, Autonomous University of Barcelona, and CIBER-BBN Network in Bioengineering, Biomaterials and Nanomedicine, Bellaterra, 08193 Barcelona, Spain

*Corresponding author: Institut de Biotecnologia i de Biomedicina. Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain. Phone: 34 935812148; Fax: 34 935812011; e mail: avillaverde@servet.uab.es

Running title: Fine tuning of protein solubility, yield and quality

Keywords: protein folding, protein quality, solubility, inclusion bodies

22 **Abstract**

23 Many enzymes or fluorescent proteins produced in *Escherichia coli* are

24 enzymatically active or fluorescent respectively when deposited as inclusion

25 bodies. The occurrence of insoluble but functional protein species with native-

26 like secondary structure indicates that solubility and conformational quality of

27 recombinant proteins are not coincident parameters, and suggests that both

28 properties can be engineered independently. We have here proven this

29 principle by producing elevated yields of a highly fluorescent but insoluble

30 Green Fluorescent Protein (GFP) protein in a DnaK⁻ background, and further

31 enhancing its solubility through adjusting the growth temperature and *GFP* gene

32 expression rate. The success of such a two-step approach confirms the

33 independent control of solubility and conformational quality, advocates for new

34 routes towards high quality protein production and intriguingly, proves that high

35 protein yields dramatically compromise the conformational quality of soluble

36 versions.

37

Introduction

Very often, the bacterial production of recombinant proteins results in the formation of insoluble protein aggregates known as inclusion bodies (Villaverde and Carrio 2003). Improving solubility has been a main goal in protein production, and a spectrum of genetic, process engineering and physicochemical approaches have been explored with relative degree of success (Sorensen and Mortensen 2005b). In particular, the co-production of appropriate sets of chaperones along with a misfolding-prone protein results in enhanced solubility ratios (de *et al.* 2007) although, at least in some cases, in clearly lower protein yield and stability (Garcia-Fruitos *et al.* 2007).

Recently, by using FourierTransform-InfraRed (FTIR) spectroscopy procedures (Ami *et al.* 2005; Oberg *et al.* 1994; Ami *et al.* 2006), it is being recognized that inclusion bodies contain important extents of properly folded, functional polypeptides (Ventura and Villaverde 2006), and that protein aggregation in recombinant *Escherichia coli* does not necessarily imply loss of biological activity, neither in the cytoplasm (Garcia-Fruitos *et al.* 2005b) nor in the periplasm (Arie *et al.* 2006). The occurrence of functional protein versions in inclusion bodies seems to be inversely dependent on the aggregation rate (de Groot and Ventura 2006). Interestingly, specific activity or fluorescence emission of recombinant enzymes and fluorescent proteins respectively is similar when comparing soluble and inclusion body versions (Gonzalez-Montalban *et al.* 2006; Martinez-Alonso *et al.* 2007; Garcia-Fruitos *et al.* 2005a). This might result from a combination of functional protein species forming inclusion bodies as mentioned above and the occurrence of soluble aggregates (de Marco and Schroedel 2005) that might contain, at different extents, misfolded and non functional proteins. Therefore, enhancing the solubility of a recombinant protein, irrespective of the used procedure, does not necessarily enhance the yield of functional versions (Gonzalez-Montalban *et al.* 2007).

Since conformational quality and solubility are not completely matching protein properties, we wondered if both parameters might be modulated by selectable conditions to enhance the yield of soluble but also biologically efficient protein. By using a misfolding-prone Green Fluorescent Protein (GFP) variant we show here that the total cellular amount of functional protein can be dramatically enhanced by producing it in a DnaK⁻ background, although it occurs as large

72 inclusion bodies. In absence of DnaK, solubility of such functional polypeptides

73 can be estimulated by appropriately adjusting growth temperature and gene

74 expression rate. The success of such a combined, two-step (genetic and

75 process) approach proves that solubility and conformational quality can be

76 independently engineered, offering new strategies to optimize protein

77 recombinant production processes. The results presented here indicate,

78 however, that high protein yield dramatically compromises the conformational

79 quality of the soluble product versions, being both parameters mutually

80 exclusive. Therefore, recombinant production processes should be designed on

81 the basis of the preferential outcome regarding yield and functionality.

82 **Materials and methods**

83

84 *Strains and plasmids*

85 *Escherichia coli* pseudo wild type strain MC4100 (*araD139 Δ(argF-lac) U169*

86 *rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*)(Sambrook *et al.* 1989; Gonzalez-

87 Montalban *et al.* 2007) and its derivatives JGT3 (*ΔclpB::kan*), JGT4 (*clpA::kan*),

88 JGT6 (*zjd::Tn10 groES30*), JGT17 (*Δibp::kan*), JGT19 (*clpP::cat*), JGT20

89 (*dnak756 thr::Tn10*)(Thomas and Baneyx 1996), BB4564 (*groEL140 zjd::Tn10*

90 *zje::ΩSpc^r/Str^r*) (Ziemienowicz *et al.* 1993) and BB2395 (*Δlon146::miniTn10*)

91 (Tomoyasu *et al.* 2001) were used in this work. All these strains were

92 transformed with plasmid pTVP1GFP (Garcia-Fruitos *et al.* 2007), which was

93 used to drive the expression of a GFP fusion protein (mGFP) containing the

94 aggregation-prone VP1 capsid protein of the foot-and-mouth disease virus. The

95 chimerical *VP1GFP* gene is under the control of the ITPG-inducible *trc*

96 promoter.

97

98 *Culture and gene expression conditions*

99 Bacterial strains were cultured at 37ºC and 250 rpm in shake flasks, in Luria-

100 Bertani (LB) rich medium with 100 µg/ml ampicillin, up to an OD_{550} of 0.4. Then,

101 the expression of the recombinant gene was triggered by addition of IPTG at

102 different final concentrations (0.01mM, 0.1 mM or 1 mM) and aliquots of the

103 culture were submitted then at different growth temperatures (16ºC, 22ºC, 27ºC,

104 32ºC, 37ºC or 42ºC). Samples for analysis were taken when the culture reached

105 an OD_{550} around 3. All experiments were performed in triplicate.

106

107 *Protein analysis*

108 Samples of bacterial cultures (15 ml) were centrifuged (for 15 min at 15000 g) to

109 harvest cells, and pellets were resuspended in 2 ml of Phosphate-Buffered

110 Saline (PBS) with one tablet of Protease Inhibitor Cocktail (Roche, ref. 1 836

111 170) per 10 ml of buffer. For analysis of the soluble fraction, 1 ml-aliquots of the

112 resuspended cells were ice-jacketed and sonicated for a minimum of 5 min at

113 50 W under 0.5 s cycles, or longer when required for total disruption of the cells

114 (Feliu *et al.* 1998). After centrifugation for 15 min at 15000 g, the supernatant,

115 corresponding to the soluble fraction, was mixed with denaturing buffer
116 (Laemmli 1970) at appropriate ratios for further Western Blot analysis.

117 The remaining 1 ml-aliquots were used to purify inclusion bodies by repeated
118 washing with detergent as described (Carrio *et al.* 2000) and resuspended in
119 denaturing buffer. Samples were boiled for 20 minutes, and appropriate
120 volumes were loaded onto denaturing gels for Western Blot analysis. mGFP
121 was immunodetected using a rabbit polyclonal antibody against GFP (Santa
122 Cruz Biotechnology, Inc.). Blots were scanned at high resolution and bands
123 quantified using Quantity One software from Bio Rad, using different amounts of
124 commercial GFP as standards. Determinations were always done in triplicate
125 and within the linear range, and they were used to calculate the specific activity
126 values.

127

128 *Fluorescence determination*

129 Soluble cell fraction samples were appropriately diluted in PBS and their
130 fluorescence measured without any further treatment. Inclusion bodies were
131 purified as described above, and resuspended in PBS for fluorescence analysis.
132 Determinations were carried out using a Cary Eclipse Fluorescence
133 Spectrophotometer (Variant) and under continuous stirring. Excitation
134 wavelength was 450 nm, and measures were taken at 510 nm. All experiments
135 were performed in triplicate. The obtained data, combined with mGFP protein
136 amounts determined by immunoanalysis, were used to calculate the specific
137 fluorescence emission of both soluble and mGFP inclusion bodies.

138

Results

140 *E. coli genetic background and yield of active protein*

141 GFP fusions are excellent models to monitor conformational quality, as the

142 proper conformation for fluorescence emission is reached during the last folding

143 steps (maturation) of GFP (Herberhold *et al.* 2003; Zhang *et al.* 2006;

144 Scheyhing *et al.* 2002). In a recent study, we have observed that *E. coli* mutant

145 cells deficient in different chaperones or proteases were more fluorescent than

146 wild type cells when producing an aggregation-prone GFP (mGFP) (Garcia-

147 Fruitos *et al.* 2007). We explored here the fluorescence distribution between

148 soluble and insoluble cell fraction in these mutants to select one with higher

149 total fluorescence per cell. For further engineering attempts, and to explore up

150 to what extent solubility and functionality can be modulated, we were interested

151 in strains with the fluorescent protein population being mainly insoluble. As

152 observed (Table 1), DnaK⁻, ClpB⁻, Lon⁻ and ClpP⁻ mutants produced significantly

153 higher fluorescence emission than wild type cells. Interestingly, in all these

154 cases, most of the fluorescent GFP accumulated as inclusion bodies, a fact that

155 has been associated to a strong inhibition of DnaK-surveyed proteolysis of

156 functional protein species (Garcia-Fruitos *et al.* 2007). Interestingly, among

157 these highly fluorescent mutants, DnaK⁻ cells showed the lowest ratio between

158 soluble and insoluble fluorescence (0.8 versus 8.0 in the wild type). Hence, we

159 decided to use this mutant to explore if conventional methods to enhance

160 solubility could promote a more favourable distribution of functional protein

161 between inclusion bodies and the soluble cell fraction, thus enhancing the

162 occurrence of both soluble and fluorescent GFP.

163

164 *Impact of temperature and gene expression rates on protein solubility and*

165 *conformational quality*

166 Therefore, we analyzed the fluorescence emission in DnaK⁻ cells producing

167 mGFP at different temperatures, from 16 to 42 °C. As observed (Figure 1a),

168 both the total fluorescence per biomass and the particular fraction of emission

169 associated with inclusion bodies increased with temperature, showing a sudden

170 up-shift between 27 and 32 °C. However, the fluorescence associated with

171 soluble protein only slightly decreased at the same temperature range, proving

172 a positive effect of temperature on the absolute yield of functional and insoluble
173 (but not soluble) protein. In this context, the ratio between soluble and insoluble
174 fluorescence significantly increased at low temperatures, reaching 4.2 at 16ºC
175 (and dropping to 0.6 at 42ºC). At 27 ºC or below, the prevalence of soluble
176 fluorescent protein was then more than 4 times higher than at 32 ºC or higher
177 temperatures.
178 With regard to protein production at each growth condition, we observed that
179 while the amount of soluble mGFP showed a slight peak at 27 ºC, amounts of
180 both total and insoluble mGFP increased with temperature (Figure 1b). This
181 resulted in a strong dependence of solubility (from 19.5 to 54 %) on
182 temperature. Altogether, these data suggested important differences in the
183 temperature-mediated evolution of protein quality, depending on the soluble-
184 insoluble protein status. In agreement (Figure 1c), the specific emission of
185 insoluble mGFP was poorly affected by temperature, although a minimum was
186 observed at 27 ºC. However, the conformational quality of total mGFP
187 increased with decreasing temperatures in an exponential pattern what was
188 essentially accounted by the soluble fraction, since the specific fluorescence of
189 aggregated mGFP was unaffected by temperature. At 27 ºC then, the soluble
190 fluorescence was slightly higher than at other temperatures (Figure 1a),
191 probably because quality and solubility were both favoured and yield was still
192 high when compared to that obtained at lower temperatures (Figure 1b).
193 We used then this intermediate growth temperature to analyze the effects of the
194 IPTG concentration on solubility and protein quality in the range of doses
195 commonly used for recombinant gene expression. As observed (Figure 2a),
196 total fluorescence per biomass was significantly lower at 0.01 mM than at the
197 other tested concentrations (namely 0.1 and 1 mM), that produced very similar
198 values. However, protein yield was strongly dependent on IPTG concentration
199 (Figure 2b). When combined with fluorescence data, these results suggest
200 dramatic effects of IPTG on protein quality. This was indeed confirmed when
201 determining the specific fluorescence of produced mGFP as distributed among
202 different fractions (Figure 2c). Medium IPTG values (0.1 mM) resulted in higher
203 quality protein than that obtained at 1mM. This fact accounts for the similar
204 fluorescence per cell observed at these two IPTG doses (Figure 1a) even when
205 the higher protein yield was obtained at 1 mM IPTG (Figure 2b).

206

207

Discussion

209 Solubility has been universally considered as the best indicator of recombinant

210 protein quality. Therefore gaining solubility is a main goal in protein production

211 processes, and numerous strategies have been tested in this regard (Sorensen

212 and Mortensen 2005b). Many of them are based on the production of

213 chaperones along with the target protein, since they are believed to be limiting

214 for recombinant protein folding. The selection of appropriate combinations of

215 chaperones has resulted in higher solubility values (Nishihara *et al.* 1998; de *et*

216 *al.* 2007), in general expressed as the percentage of soluble over total protein.

217 However, a detailed analysis of published data suggests that at least in some

218 cases, increasing solubility through chaperone co-production would reduce the

219 final protein yield. This concept has been clearly shown by the co-production of

220 the DnaK-DnaJ pair, which dramatically reduces the proteolytic stability and

221 yield of an IB-forming GFP (Garcia-Fruitos *et al.* 2007). In fact, a

222 comprehensive genetic analysis of protein production in *E. coli* has recently

223 indicated that cell mutations increasing solubility minimize the conformational

224 quality of the soluble protein (Garcia-Fruitos *et al.* 2007). This fact, and other

225 findings relevant to functionality of soluble and insoluble polypeptides in

226 recombinant bacteria clearly prove that solubility and conformational quality are

227 non matching (and potentially divergent) protein properties (Gonzalez-

228 Montalban *et al.* 2007).

229

230 For industrial processes requiring functional products, the production of highly

231 active polypeptides (irrespective of their solubility) would be more appealing

232 than high percentages (but poor yields) of soluble and moderately active

233 polypeptides. However, solubility is obviously required for applications such as

234 crystallographic determination or in vivo protein delivery for therapeutic

235 purposes among others. In this work, we have explored how solubility of highly

236 functional proteins (produced in a convenient DnaK$^-$ background) can be

237 successfully manipulated through process engineering by manipulating growth

238 temperature and gene expression rate. In this regard, temperature, in the

239 physiological ranges between 16 and 42ºC has a positive and progressive

240 impact on the total yield of mGFP. This is exclusively accounted for by an
241 increase in the amount of aggregated protein since the yield of the soluble
242 version is only slightly affected (Figure 1b). The total fluorescence per cell
243 undergoes an up-shift above 27°C, but again it is accounted for exclusively by
244 the insoluble cell fraction (Figure 1a). Finally, the conformational quality of
245 soluble mGFP is dramatically and progressively impaired by temperature while
246 specific fluorescence of IB mGFP remains nearly constant (Figure 1c). The
247 influence of IPTG concentration is more modest regarding the variation range of
248 the studied parameters, which follow a less progressive pattern than the one
249 defined by temperature. However, the divergent evolution of yield (and total
250 fluorescence) and the functional quality of the soluble protein version is also
251 evident (Figure 2). Importantly, by combining the appropriate temperature
252 (27°C) and IPTG dose (0.1 mM), the distribution of fluorescence between
253 soluble and insoluble shifted from 0.8 (Table 1 and Figure 1a) to 2.1 (Figure
254 2a). Of course, better distribution values can be reached at 1 mM IPTG (3.5),
255 but at the expense of protein quality measured by specific fluorescence (Figure
256 2c).

257 More intriguingly, the data presented here indicate that yield, solubility and
258 conformational quality of soluble proteins cannot be favoured simultaneously in
259 recombinant *E. coli*. This fact must be seriously considered in protein production
260 processes, since the production strategy should be clearly targeted to either
261 protein yield, solubility or product quality. In this regard, many of the non-
262 coincident reports regarding the success of given strategies for improved
263 protein production (Sorensen and Mortensen 2005a; de Marco *et al.* 2000; de
264 and De, V 2004; de *et al.* 2007; Schultz *et al.* 2006; Baneyx and Palumbo 2003;
265 Baneyx and Mujacic 2004) and the unpredictability and product-dependence of
266 the chaperone co-production approach (de 2007; de *et al.* 2007) are probably
267 accounted for (at least in many cases) by the different parameters through
268 which process success are measured, namely solubility, yield or functionality.
269 While evidences that enhancing solubility does not imply better protein quality
270 are now stronger (Gonzalez-Montalban *et al.* 2007), the results presented here
271 furthermore indicate that conditions promoting high protein yield and high
272 soluble yield are clearly adverse for conformational quality (Figure 3). In this
273 context, the distribution of fluorescence between soluble and insoluble cell

274 fractions (Figure 3a) and the specific fluorescence of soluble mGFP (Figure 3b)
275 are negatively affected by the total production of mGFP. Likewise, the lower the
276 yield of soluble mGFP, the higher its conformational quality is (Figure 3c),
277 strongly supporting the concept that gaining yield and quality cannot be reached
278 simultaneously.
279

280 **Acknowledgements**

284

285

286 **Legends for Figures**

287

288 **Figure 1.** Effect of growth temperature on the fluorescence per cell biomass
289 (A), yield (B) and specific fluorescence of mGFP (C). The ratio of soluble and
290 insoluble fluorescence (panel A) and mGFP solubility (panel B) are also
291 indicated for each temperature.

292

293 **Figure 2.** Effect of IPTG concentration on the fluorescence per cell biomass (A),
294 yield (B) and specific fluorescence of mGFP (C). The ratio of soluble and
295 insoluble fluorescence (panel A) and mGFP solubility (panel B) are also
296 indicated for each dose.

297

298 **Figure 3.** Influence of total (A and B) and soluble (C) mGFP yield on
299 soluble/insoluble fluorescence ratio (A) and specific fluorescence of soluble
300 mGFP (B and C). All the conditions shown in Figure 2 and 3 were used in this
301 analysis. In all cases, the data set also fitted to exponential decay, single, 2
302 parameter equations (not shown) with important extents of statistic significance
303 (A, $p=0.1209$; B, $p=0.0444$; C, $p=0.0292$).

304

305

306 **Table 1.** Fluorescence emission observed in the soluble and inclusion body fractions in

307 mGFP-producing cells

308

Phenotype	Total units/OD[a]	fraction %	Soluble units/OD	fraction %	Inclusion units/OD	bodies %	Ratio of soluble/IB fluorescence[b]	Solubility (%)
wt (MC4100)	405.6±13.1	100	331.0±71.6	100	40.9±29.3	100	8.0	40.8±11.1
DnaK⁻	547.3±77.9	134.9±19.2	200.6±33.7	60.6±10.2	235.8±42.5	575.653±103.8	0.8	20.7±3.4
GroEL140	342.9±35.3	84.5±8.7	245.2±23.8	74.1±7.2	47.1±7.5	114.996±18.4	5.2	30.2±5.4
ClpB⁻	515.6±24.4	127.1±6.0	310.7±24.9	93.8±7.5	231.8±25.3	565.879±61.9	1.3	31.2±11.8
ClpA⁻	543.1±45.5	133.9±11.2	46.9±7.1	14.1±2.1	66.2±18.6	161.662±45.5	0.7	44.2±9.9
GroES⁻	440.2±5.3	108.5±1.3	379.2±15.9	114.5±4.8	105.1±13.0	256.726±31.8	3.6	27.1±7.8
IbpAB⁻	303.2±16.1	74.7±4.0	261.3±45.7	78.9±13.8	62.9±9.3	153.622±22.8	4.1	30.1±9.8
ClpP⁻	522.4±31.2	128.7±7.7	431.0±38.0	130.2±11.4	237.8±29.2	580.392±71.4	1.8	25.0±8.7
Lon⁻	686.4±40.9	169.2±10.1	400.7±21.8	121.0±6.5	261.9±12.5	639.428±30.5	1.5	14.6±3.5

309

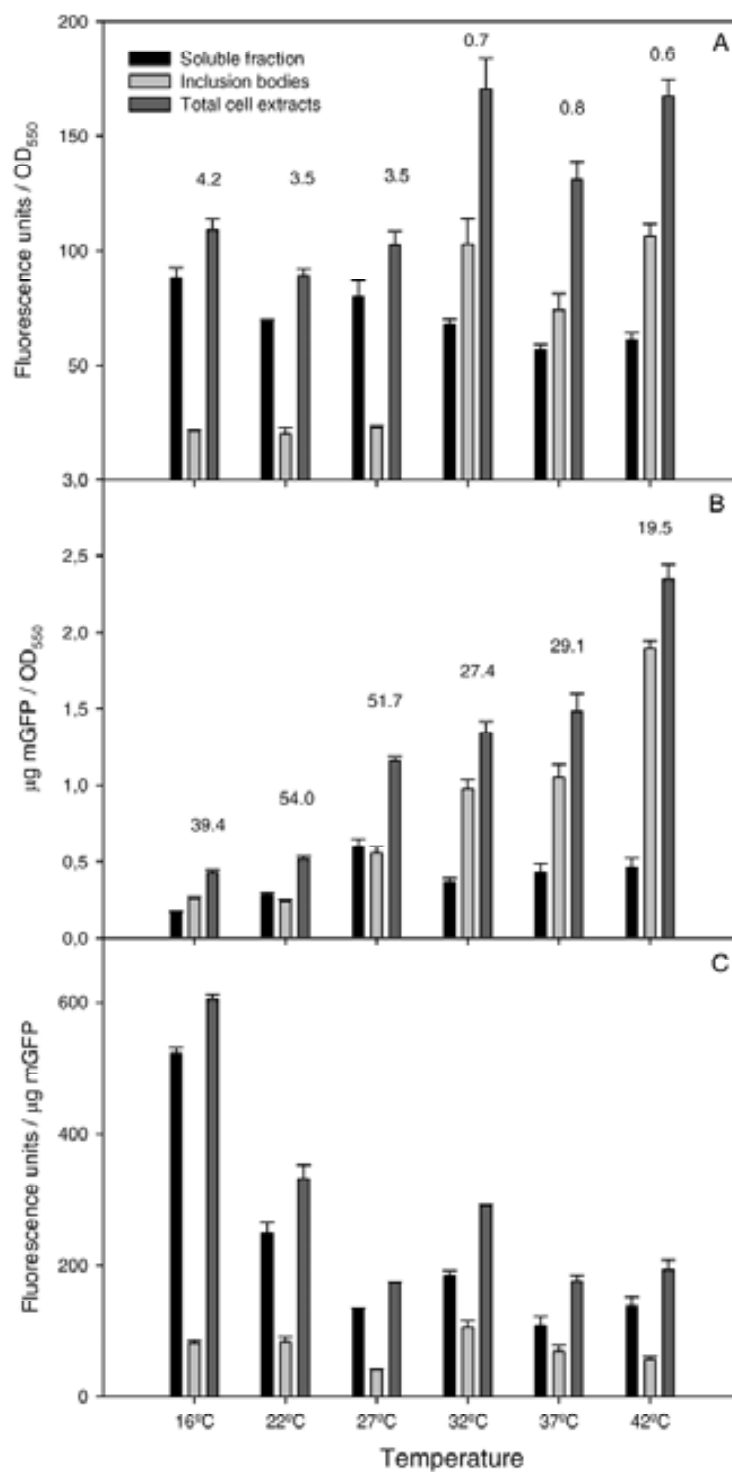310 [a] Total fluorescence data in this strain set have been obtained and shown in a previous

311 study (Garcia-Fruitos et al. 2007).

312 [b] Quotient between soluble and insoluble fluorescence in a given sample.

313 [c] Amount of soluble mGFP relative to its total amount in the cell.

314
315   Figure 1
316

317
318

319
320

Figure 2
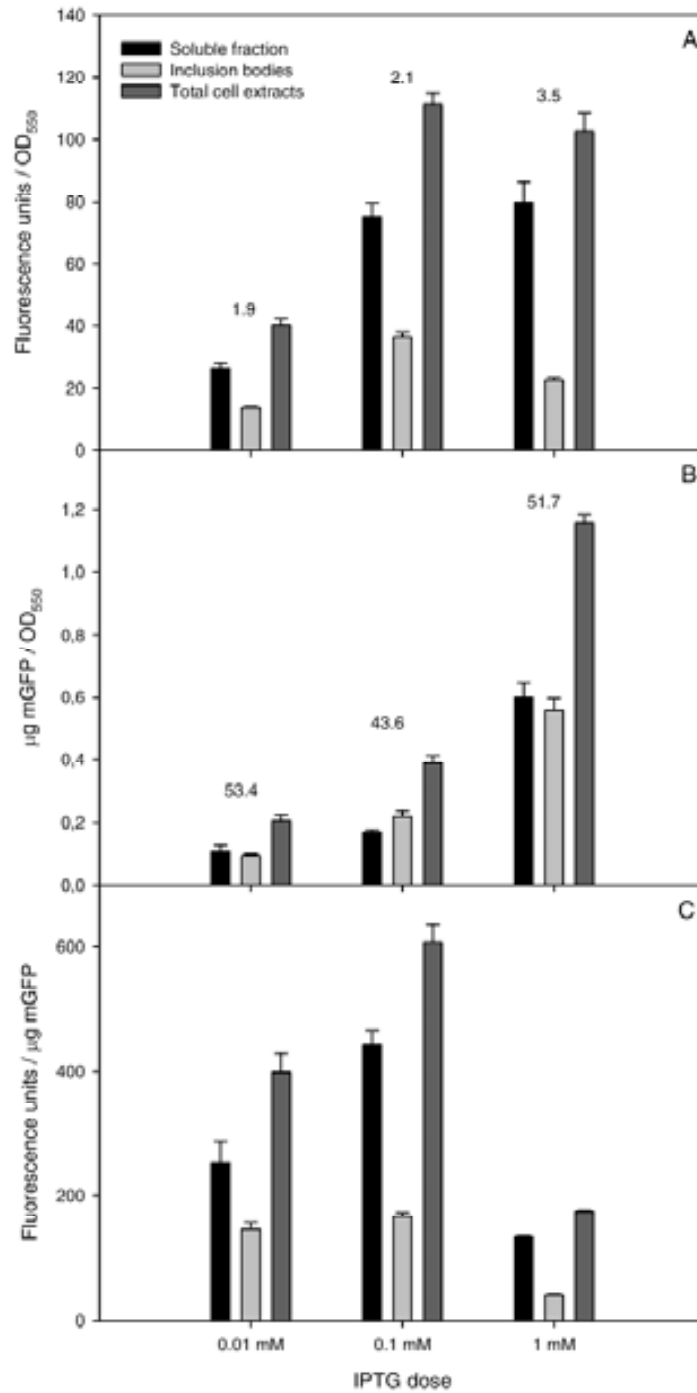
321
322
323
324
325
326
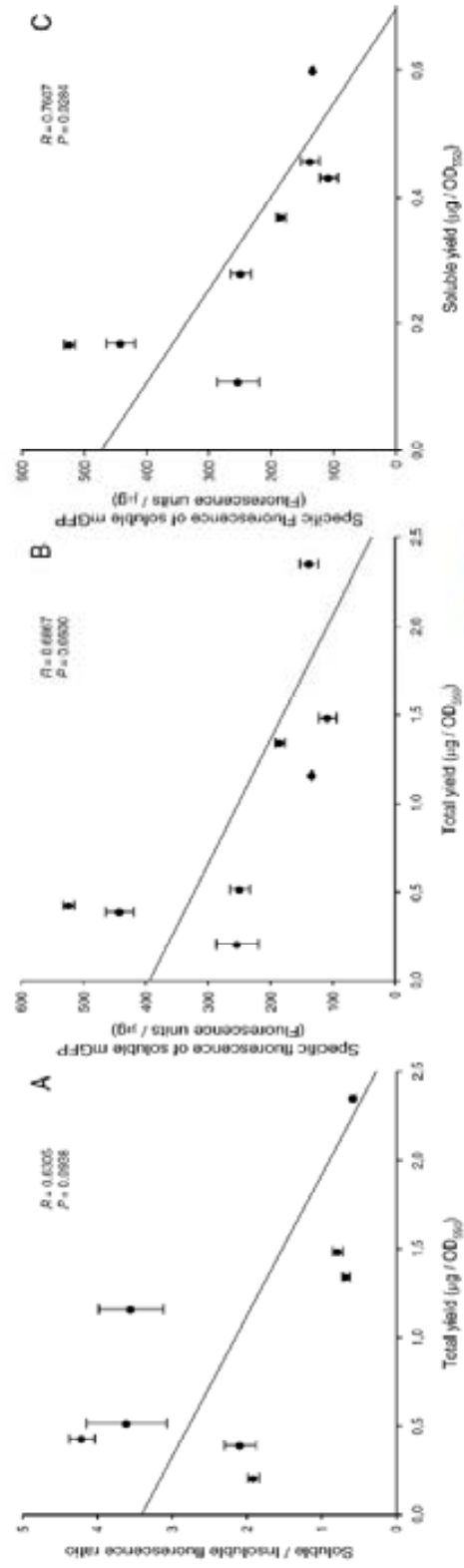
178

Annex II

Figure 3

I5

333
334   Reference List
335

336   Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM (2005) Kinetics of
337   inclusion body formation studied in intact cells by FT-IR spectroscopy. *FEBS Lett* **579,**
338   3433-3436.

339   Ami D, Natalello A, Taylor G, Tonon G, Maria DS (2006) Structural analysis of protein
340   inclusion bodies by Fourier transform infrared microspectroscopy.
341   *Biochim.Biophys.Acta* **1764,** 793-799.

342   Arie JP, Miot M, Sassoon N, Betton JM (2006) Formation of active inclusion bodies in
343   the periplasm of Escherichia coli. *Mol.Microbiol.* **62,** 427-437.

344   Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in
345   Escherichia coli. *Nat.Biotechnol.* **22,** 1399-1408.

346   Baneyx F, Palumbo JL (2003) Improving heterologous protein folding via molecular
347   chaperone and foldase co-expression. *Methods Mol Biol* **205,** 171-197.

348   Carrio MM, Cubarsi R, Villaverde A (2000) Fine architecture of bacterial inclusion
349   bodies. *FEBS Lett* **471,** 7-11.

350   de Groot NS, Ventura S (2006) Protein activity in bacterial inclusion bodies correlates
351   with predicted aggregation rates. *J.Biotechnol.* **125,** 110-113.

352   de Marco A, Schroedel A (2005) Characterization of the aggregates formed during
353   recombinant protein expression in bacteria. *BMC.Biochem* **6,** 10.

354   de Marco A, Volrath S, Bruyere T, Law M, Fonne-Pfister R (2000) Recombinant maize
355   protoporphyrinogen IX oxidase expressed in Escherichia coli forms complexes with
356   GroEL and DnaK chaperones. *Protein Expr Purif* **20,** 81-86.

357   de MA (2007) Protocol for preparing proteins with improved solubility by co-
358   expressing with molecular chaperones in Escherichia coli. *Nat.Protoc.* **2,** 2632-2639.

359   de MA, De M, V (2004) Bacteria co-transformed with recombinant proteins and
360   chaperones cloned in independent plasmids are suitable for expression tuning.
361   *J.Biotechnol.* **109,** 45-52.

362   de MA, Deuerling E, Mogk A, Tomoyasu T, Bukau B (2007) Chaperone-based
363   procedure to increase yields of soluble recombinant proteins produced in E. coli.
364   *BMC.Biotechnol.* **7,** 32.

365   Feliu JX, Cubarsi R, Villaverde A (1998) Optimized release of recombinant proteins by
366   ultrasonication of E. coli cells. *Biotechnol Bioeng* **58,** 536-540.

367   Garcia-Fruitos E, Carrio MM, Aris A, Villaverde A (2005a) Folding of a misfolding-
368   prone beta-galactosidase in absence of DnaK. *Biotechnol.Bioeng.* **90,** 869-875.

369 Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A,
370 Ventura S, Villaverde A (2005b) Aggregation as bacterial inclusion bodies does not
371 imply inactivation of enzymes and fluorescent proteins. *Microb.Cell Fact.* **4**, 27.

372 Garcia-Fruitos E, Martinez-Alonso M, Gonzalez-Montalban N, Valli M, Mattanovich
373 D, Villaverde A (2007) Divergent Genetic Control of Protein Solubility and
374 Conformational Quality in Escherichia coli. *J.Mol.Biol.* **374**, 195-205.

375 Gonzalez-Montalban N, Garcia-Fruitos E, Ventura S, Aris A, Villaverde A (2006) The
376 chaperone DnaK controls the fractioning of functional protein between soluble and
377 insoluble cell fractions in inclusion body-forming cells. *Microb.Cell Fact.* **5**, 26.

378 Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A (2007) Recombinant protein
379 solubility-does more mean better? *Nat.Biotechnol.* **25**, 718-720.

380 Herberhold H, Marchal S, Lange R, Scheyhing CH, Vogel RF, Winter R (2003)
381 Characterization of the pressure-induced intermediate and unfolded state of red-shifted
382 green fluorescent protein--a static and kinetic FTIR, UV/VIS and fluorescence
383 spectroscopy study. *J.Mol Biol.* **330**, 1153-1164.

384 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of
385 bacteriophage T4. *Nature* **227**, 680-685.

386 Martinez-Alonso M, Vera A, Villaverde A (2007) Role of the chaperone DnaK in
387 protein solubility and conformational quality in inclusion body-forming Escherichia coli
388 cells. *FEMS Microbiol.Lett.* **273**, 187-195.

389 Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T (1998) Chaperone
390 coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and
391 GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in
392 Escherichia coli. *Appl Environ Microbiol* **64**, 1694-1699.

393 Oberg K, Chrunyk BA, Wetzel R, Fink AL (1994) Nativelike secondary structure in
394 interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. *Biochemistry*
395 **33**, 2628-2634.

396 Sambrook, J, Fritsch, E., and Maniatis, T. Molecular Cloning, A Laboratory Manual,
397 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989.
398 Ref Type: Generic

399 Scheyhing CH, Meersman F, Ehrmann MA, Heremans K, Vogel RF (2002)
400 Temperature-pressure stability of green fluorescent protein: a Fourier transform infrared
401 spectroscopy study. *Biopolymers* **65**, 244-253.

402 Schultz T, Martinez L, de MA (2006) The evaluation of the factors that cause
403 aggregation during recombinant expression in E. coli is simplified by the employment
404 of an aggregation-sensitive reporter. *Microb.Cell Fact.* **5**, 28.

405 Sorensen HP, Mortensen KK (2005a) Advanced genetic strategies for recombinant
406 protein expression in Escherichia coli. *J.Biotechnol.* **115**, 113-128.

407 Sorensen HP, Mortensen KK (2005b) Soluble expression of recombinant proteins in the
408 cytoplasm of Escherichia coli. *Microb.Cell Fact.* **4,** 1.

409 Thomas JG, Baneyx F (1996) Protein folding in the cytoplasm of Escherichia coli:
410 requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone
411 machines. *Mol Microbiol* **21,** 1185-1196.

412 Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B (2001) Genetic dissection of
413 the roles of chaperones and proteases in protein folding and degradation in the
414 Escherichia coli cytosol. *Mol Microbiol* **40,** 397-413.

415 Ventura S, Villaverde A (2006) Protein quality in bacterial inclusion bodies. *Trends*
416 *Biotechnol.* **24,** 179-185.

417 Villaverde A, Carrio MM (2003) Protein aggregation in recombinant bacteria:
418 biological role of inclusion bodies. *Biotechnol Lett* **25,** 1385-1395.

419 Zhang L, Patel HN, Lappe JW, Wachter RM (2006) Reaction progress of chromophore
420 biogenesis in green fluorescent protein. *J.Am.Chem.Soc.* **128,** 4766-4772.

421 Ziemienowicz A, Skowyra D, Zeilstra-Ryalls J, Fayet O, Georgopoulos C, Zylicz M
422 (1993) Both the Escherichia coli chaperone systems, GroEL/GroES and
423 DnaK/DnaJ/GrpE, can reactivate heat-treated RNA polymerase. Different mechanisms
424 for the same activity. *J Biol Chem* **268,** 25425-25431.
425
426

IX.References

1. Yokoyama, S. Protein expression systems for structural genomics and proteomics. Curr. Opin. Chem. Biol. 7, 39-43 (2003).

2. Enfors, S. O. Control of in vivo proteolysis in the production of recombinant proteins. Trends Biotechnol. 10, 310-315 (1992).

3. Georgiou, G. & Valax, P. Expression of correctly folded proteins in Escherichia coli. Curr. Opin. Biotechnol. 7, 190-197 (1996).

4. del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. & az-Orejas, R. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62, 434-464 (1998).

5. Makrides, S. C. Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol. Rev. 60, 512-538 (1996).

6. Hannig, G. & Makrides, S. C. Strategies for optimizing heterologous protein expression in Escherichia coli. Trends Biotechnol. 16, 54-60 (1998).

7. Terpe, K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. 72, 211-222 (2006).

8. Figge, J., Wright, C., Collins, C. J., Roberts, T. M. & Livingston, D. M. Stringent regulation of stably integrated chloramphenicol acetyl transferase genes by E. coli lac repressor in monkey cells. Cell 52, 713-722 (1988).

9. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189, 113-130 (1986).

10. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. 177, 4121-4130 (1995).

11. Elvin, C. M. et al. Modified bacteriophage lambda promoter vectors for overproduction of proteins in Escherichia coli. Gene 87, 123-126 (1990).

12. Lowman, H. B. & Bina, M. Temperature-mediated regulation and downstream inducible selection for controlling gene expression from the bacteriophage lambda pL promoter. Gene 96, 133-136 (1990).

13. Love, C. A., Lilley, P. E. & Dixon, N. E. Stable high-copy-number bacteriophage lambda promoter vectors for overproduction of proteins in Escherichia coli. Gene 176, 49-53 (1996).

14. Vasina, J. A. & Baneyx, F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. Protein Expr Purif 9, 211-218 (1997).

15. San, K. Y., Bennett, G. N., Chou, C. H. & Aristidou, A. A. An optimization study of a pH-inducible promoter system for high-level recombinant protein production in Escherichia coli. Ann. N. Y. Acad. Sci. 721, 268-276 (1994).

16. Tolentino, G. J., Meng, S. Y., Bennett, G. N. & San, K. Y. A Ph-Regulated Promoter for the Expression of Recombinant Proteins in Escherichia-Coli. Biotechnology Letters 14, 157-162 (1992).

17. Sorensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in Escherichia coli. J. Biotechnol. 115, 113-128 (2005).

18. Gouy, M. & Gautier, C. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. 10, 7055-7074 (1982).

19. Sharp, P. M. et al. Codon usage patterns in Escherichia coli, Bacillus subtilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity. Nucleic Acids Res. 16, 8207-8211 (1988).

20. Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. J. Mol. Biol. 260, 649-663 (1996).

21. McNulty, D. E. et al. Mistranslational errors associated with the rare arginine codon CGG in Escherichia coli. Protein Expr. Purif. 27, 365-374 (2003).

22. Kurland, C. & Gallant, J. Errors of heterologous protein expression. Curr. Opin. Biotechnol. 7, 489-493 (1996).

23. Sorensen, H. P., Sperling-Petersen, H. U. & Mortensen, K. K. Production of recombinant thermostable proteins expressed in Escherichia coli: completion of protein synthesis is the bottleneck. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 786, 207-214 (2003).

24. Kane, J. F. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr. Opin. Biotechnol. 6, 494-500 (1995).

25. Goldman, E., Rosenberg, A. H., Zubay, G. & Studier, F. W. Consecutive low-usage leucine codons block translation only when near the 5' end of a message in Escherichia coli. J. Mol. Biol. 245, 467-473 (1995).

26. Buell, G. et al. Optimizing the expression in E. coli of a synthetic gene encoding somatomedin-C (IGF-I). Nucleic Acids Res. 13, 1923-1938 (1985).

27. Hernan, R. A. et al. Human hemoglobin expression in Escherichia coli: importance of optimal codon usage. Biochemistry 31, 8619-8628 (1992).

28. Kane, J. F. et al. Novel in-frame two codon translational hop during synthesis of bovine placental lactogen in a recombinant strain of Escherichia coli. Nucleic Acids Res. 20, 6707-6712 (1992).

29. Calderone, T. L., Stevens, R. D. & Oas, T. G. High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in Escherichia coli. J. Mol. Biol. 262, 407-412 (1996).

30. Brinkmann, U., Mattes, R. E. & Buckel, P. High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. Gene 85, 109-114 (1989).

31. Del, T. B., Jr. et al. Effects of a minor isoleucyl tRNA on heterologous protein translation in Escherichia coli. J. Bacteriol. 177, 7086-7091 (1995).

32. Dieci, G., Bottarelli, L., Ballabeni, A. & Ottonello, S. tRNA-assisted overproduction of eukaryotic ribosomal proteins. Protein Expr. Purif. 18, 346-354 (2000).

33. Lorimer, G. H. A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo. FASEB J. 10, 5-9 (1996).

34. Anfinsen, C. B. Principles that govern the folding of protein chains. Science 181, 223-230 (1973).

35. Snow, C. D., Nguyen, H., Pande, V. S. & Gruebele, M. Absolute comparison of simulated and experimental protein-folding dynamics. Nature 420, 102-106 (2002).

36. Hardesty, B. & Kramer, G. Folding of a nascent peptide on the ribosome. Prog Nucleic Acid Res Mol Biol 66, 41-66 (2001).

37. Hartl, F. U. & Hayer-Hartl, M. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295, 1852-1858 (2002).

38. Karplus, M. The Levinthal paradox: yesterday and today. Fold. Des 2, S69-S75 (1997).

39. Yon, J. M. Protein folding: a perspective for biology, medicine and biotechnology. Braz. J. Med. Biol. Res. 34, 419-435 (2001).

40. Matagne, A. & Dobson, C. M. The folding process of hen lysozyme: a perspective from the 'new view'. Cell Mol. Life Sci. 54, 363-371 (1998).

41. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. Navigating the folding routes. Science 267, 1619-1620 (1995).

42. Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. Nat. Struct. Biol. 4, 10-19 (1997).

43. Baldwin, R. L. Protein folding. Matching speed and stability. Nature 369, 183-184 (1994).

44. Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M. & Karplus, M. Understanding protein folding via free-energy surfaces from theory and experiment. Trends Biochem. Sci. 25, 331-339 (2000).

45. Otzen, D. E. & Fersht, A. R. Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. Biochemistry 37, 8139-8146 (1998).

46. Jackson, S. E. How do small single-domain proteins fold? Fold. Des 3, R81-R91 (1998).

47. Panchenko, A. R., Luthey-Schulten, Z. & Wolynes, P. G. Foldons, protein structural modules, and exons. Proc. Natl. Acad. Sci. U. S. A 93, 2008-2013 (1996).

48. Vendruscolo, M., Paci, E., Dobson, C. M. & Karplus, M. Rare fluctuations of native proteins sampled by equilibrium hydrogen exchange. J. Am. Chem. Soc. 125, 15686-15687 (2003).

49. Khan, F., Chuang, J. I., Gianni, S. & Fersht, A. R. The kinetic pathway of folding of barnase. J. Mol. Biol. 333, 169-186 (2003).

50. Dobson, C. M. Principles of protein folding, misfolding and aggregation. Semin. Cell Dev. Biol 15, 3-16 (2004).

51. Fink, A. L. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold. Des 3, R9-23 (1998).

52. Gething, M. J. & Sambrook, J. Protein folding in the cell. Nature 355, 33-45 (1992).

53. Selkoe, D. J. Folding proteins in fatal ways. Nature 426, 900-904 (2003).

54. Dobson, C. M. Experimental investigation of protein folding and misfolding. Methods 34, 4-14 (2004).

55. Frokjaer, S. & Otzen, D. E. Protein drug stability: a formulation challenge. Nat. Rev. Drug Discov. 4, 298-306 (2005).

56. Chiti, F. et al. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. Proc. Natl. Acad. Sci. U. S. A 96, 3590-3594 (1999).

57. Babbitt, P. C., West, B. L., Buechter, D. D., Kuntz, I. D. & Kenyon, G. L. Removal of a proteolytic activity associated with aggregates formed from expression of creatine kinase in Escherichia coli leads to improved recovery of active enzyme. Biotechnology (N. Y.) 8, 945-949 (1990).

58. Fandrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle

myoglobin - Even an ordinary globular protein can assume a rogue guise if conditions are right. Nature 410, 165-166 (2001).

59. Fandrich, M. & Dobson, C. M. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. EMBO J. 21, 5682-5690 (2002).

60. Tan, S. Y. & Pepys, M. B. Amyloidosis. Histopathology 25, 403-414 (1994).

61. Sunde, M. & Blake, C. C. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. Q. Rev. Biophys. 31, 1-39 (1998).

62. Pallares, I., Vendrell, J., Aviles, F. X. & Ventura, S. Amyloid fibril formation by a partially structured intermediate state of alpha-chymotrypsin. J. Mol. Biol. 342, 321-331 (2004).

63. Villegas, V. et al. Protein engineering as a strategy to avoid formation of amyloid fibrils. Protein Sci. 9, 1700-1708 (2000).

64. Dobson, C. M. Protein misfolding, evolution and disease. Trends Biochem. Sci. 24, 329-332 (1999).

65. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat. Biotechnol. 22, 1302-1306 (2004).

66. Pawar, A. P. et al. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. J. Mol. Biol. 350, 379-392 (2005).

67. Sunde, M. & Blake, C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. Adv. Protein Chem. 50, 123-159 (1997).

68. Chiti, F. et al. Kinetic partitioning of protein folding and aggregation. Nat Struct Biol 9, 137-143 (2002).

69. Klunk, W. E., Pettegrew, J. W. & Abraham, D. J. Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. J. Histochem. Cytochem. 37, 1273-1281 (1989).

70. Caughey, B. & Lansbury, P. T. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. Annu. Rev. Neurosci. 26, 267-298 (2003).

71. Walsh, D. M. et al. Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. Nature 416, 535-539 (2002).

72. Laskey, R. A., Honda, B. M., Mills, A. D. & Finch, J. T. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature 275, 416-420 (1978).

73. Ellis, J. Proteins as molecular chaperones. Nature 328, 378-379 (1987).

74. Young, J. C., Agashe, V. R., Siegers, K. & Hartl, F. U. Pathways of chaperone-mediated protein folding in the cytosol. Nat. Rev. Mol. Cell Biol. 5, 781-791 (2004).

75. Bukau, B., Weissman, J. & Horwich, A. Molecular chaperones and protein quality control. Cell 125, 443-451 (2006).

76. Kazemi-Esfarjani, P. & Benzer, S. Genetic suppression of polyglutamine toxicity in Drosophila. Science 287, 1837-1840 (2000).

77. Krobitsch, S. & Lindquist, S. Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins.

Proc. Natl. Acad. Sci. U. S. A 97, 1589-1594 (2000).

78. Warrick, J. M. et al. Suppression of polyglutamine-mediated neurodegeneration in Drosophila by the molecular chaperone HSP70. Nat. Genet. 23, 425-428 (1999).

79. Kazemi-Esfarjani, P. & Benzer, S. Genetic suppression of polyglutamine toxicity in Drosophila. Science 287, 1837-1840 (2000).

80. Muchowski, P. J. et al. Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. Proc. Natl. Acad. Sci. U. S. A 97, 7841-7846 (2000).

81. Mogk, A. et al. Identification of thermolabile Escherichia coli proteins: prevention and reversion of aggregation by DnaK and ClpB. EMBO J 18, 6934-6949 (1999).

82. Straus, D. B., Walter, W. A. & Gross, C. A. The heat shock response of E. coli is regulated by changes in the concentration of sigma 32. Nature 329, 348-351 (1987).

83. Lemaux, P. G., Herendeen, S. L., Bloch, P. L. & Neidhardt, F. C. Transient rates of synthesis of individual polypeptides in E. coli following temperature shifts. Cell 13, 427-434 (1978).

84. Ehrnsperger, M., Graber, S., Gaestel, M. & Buchner, J. Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. EMBO J. 16, 221-229 (1997).

85. Veinger, L., Diamant, S., Buchner, J. & Goloubinoff, P. The small heat-shock protein IbpB from Escherichia coli stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. J. Biol. Chem. 273, 11032-11037 (1998).

86. Narberhaus, F. Alpha-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. Microbiol. Mol. Biol. Rev. 66, 64-93 (2002).

87. Malki, A., Kern, R., Abdallah, J. & Richarme, G. Characterization of the Escherichia coli YedU protein as a molecular chaperone. Biochem. Biophys. Res. Commun. 301, 430-436 (2003).

88. Mujacic, M., Bader, M. W. & Baneyx, F. Escherichia coli Hsp31 functions as a holding chaperone that cooperates with the DnaK-DnaJ-GrpE system in the management of protein misfolding under severe stress conditions. Mol Microbiol 51, 849-859 (2004).

89. Graf, P. C. & Jakob, U. Redox-regulated molecular chaperones. Cell Mol. Life Sci. 59, 1624-1631 (2002).

90. Schirmer, E. C., Glover, J. R., Singer, M. A. & Lindquist, S. HSP100/Clp proteins: a common mechanism explains diverse functions. Trends Biochem. Sci. 21, 289-296 (1996).

91. Ben Zvi, A. P. & Goloubinoff, P. Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. J Struct Biol 135, 84-93 (2001).

92. Schlieker, C., Tews, I., Bukau, B. & Mogk, A. Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides. FEBS Lett. 578, 351-356 (2004).

93. Thomas, J. G. & Baneyx, F. ClpB and HtpG facilitate de novo protein folding in stressed Escherichia coli cells. Mol Microbiol 36, 1360-1370 (2000).

94. Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E. & Bukau, B. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol Microbiol 50, 585-595 (2003).

95. Kazemi-Esfarjani, P. & Benzer, S. Genetic suppression of polyglutamine toxicity in Drosophila. Science 287, 1837-1840 (2000).

96. Huang, H. C., Sherman, M. Y., Kandror, O. & Goldberg, A. L. The molecular chaperone DnaJ is required for the degradation of a soluble abnormal protein in Escherichia coli. J. Biol. Chem. 276, 3920-3928 (2001).

97. Bukau, B., Deuerling, E., Pfund, C. & Craig, E. A. Getting newly synthesized proteins into shape. Cell 101, 119-122 (2000).

98. Nishihara, K., Kanemori, M., Yanagi, H. & Yura, T. Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli. Appl Environ Microbiol 66, 884-889 (2000).

99. Huang, G. C., Li, Z. Y., Zhou, J. M. & Fischer, G. Assisted folding of D-glyceraldehyde-3-phosphate dehydrogenase by trigger factor. Protein Sci. 9, 1254-1261 (2000).

100. Hoffmann, F. & Rinas, U. Roles of heat-shock chaperones in the production of recombinant proteins in Escherichia coli. Adv. Biochem. Eng Biotechnol. 89, 143-161 (2004).

101. Lindquist, S. & Craig, E. A. The heat-shock proteins. Annu. Rev. Genet. 22, 631-677 (1988).

102. Gupta, R. S. & Singh, B. Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. Curr. Biol. 4, 1104-1114 (1994).

103. Hunt, C. & Morimoto, R. I. Conserved features of eukaryotic hsp70 genes revealed by comparison with the nucleotide sequence of human hsp70. Proc. Natl. Acad. Sci. U. S. A 82, 6455-6459 (1985).

104. Goloubinoff, P., Mogk, A., Zvi, A. P., Tomoyasu, T. & Bukau, B. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc Natl Acad Sci U S A 96, 13732-13737 (1999).

105. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H. & Yura, T. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. Appl Environ Microbiol 64, 1694-1699 (1998).

106. Thomas, J. G. & Baneyx, F. Protein folding in the cytoplasm of Escherichia coli: requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines. Mol Microbiol 21, 1185-1196 (1996).

107. Yura, T. & Nakahigashi, K. Regulation of the heat-shock response. Curr. Opin. Microbiol. 2, 153-158 (1999).

108. Bukau, B. Regulation of the Escherichia coli heat-shock response. Mol Microbiol 9, 671-680 (1993).

109. Teter, S. A. et al. Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. Cell 97, 755-765 (1999).

110. Mogk, A. & Bukau, B. Molecular chaperones: structure of a protein disaggregase. Curr Biol 14, R78-R80 (2004).

111. Zolkiewski, M. ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. J. Biol. Chem. 274, 28083-28086 (1999).

112. Glover, J. R. & Tkach, J. M. Crowbars and ratchets: hsp100 chaperones as tools in reversing protein aggregation. Biochem. Cell Biol. 79, 557-568 (2001).

113. Ben-Zvi, A. P. & Goloubinoff, P. Review: mechanisms of disaggregation and refolding of stable protein aggregates

by molecular chaperones. J. Struct. Biol. 135, 84-93 (2001).

114. Carrio, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. J Biotechnol 96, 3-12 (2002).

115. Echave, P. et al. DnaK dependence of mutant ethanol oxidoreductases evolved for aerobic function and protective role of the chaperone against protein oxidative damage in Escherichia coli. Proc Natl Acad Sci U S A 99, 4626-4631 (2002).

116. Fredriksson, A., Ballesteros, M., Dukan, S. & Nystrom, T. Defense against protein carbonylation by DnaK/DnaJ and proteases of the heat shock regulon. J. Bacteriol. 187, 4207-4213 (2005).

117. Nagai, H., Yuzawa, H., Kanemori, M. & Yura, T. A distinct segment of the sigma 32 polypeptide is involved in DnaK-mediated negative control of the heat shock response in Escherichia coli. Proc Natl Acad Sci U S A 91, 10280-10284 (1994).

118. Tomoyasu, T., Ogura, T., Tatsuta, T. & Bukau, B. Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in Escherichia coli. Mol Microbiol 30, 567-581 (1998).

119. Morita, M. T., Kanemori, M., Yanagi, H. & Yura, T. Dynamic interplay between antagonistic pathways controlling the sigma 32 level in Escherichia coli. Proc Natl Acad Sci U S A 97, 5860-5865 (2000).

120. Tomoyasu, T., Arsène, F., Ogura, T. & Bukau, B. The C terminus of sigma(32) is not essential for degradation by FtsH. J Bacteriol 183, 5911-5917 (2001).

121. Genevaux, P., Schwager, F., Georgopoulos, C. & Kelley, W. L. The djlA gene acts synergistically with dnaJ in promoting Escherichia coli growth. J Bacteriol 183, 5747-5750 (2001).

122. Genevaux, P., Georgopoulos, C. & Kelley, W. L. The Hsp70 chaperone machines of Escherichia coli: a paradigm for the repartition of chaperone functions. Mol. Microbiol. 66, 840-857 (2007).

123. Hennessy, F., Nicoll, W. S., Zimmermann, R., Cheetham, M. E. & Blatch, G. L. Not all J domains are created equal: implications for the specificity of Hsp40-Hsp70 interactions. Protein Sci. 14, 1697-1709 (2005).

124. Zhu, X. et al. Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272, 1606-1614 (1996).

125. Ewalt, K. L., Hendrick, J. P., Houry, W. A. & Hartl, F. U. In vivo observation of polypeptide flux through the bacterial chaperonin system. Cell 90, 491-500 (1997).

126. Carrio, M. M. & Villaverde, A. Protein aggregation as bacterial inclusion bodies is reversible. FEBS Lett 489, 29-33 (2001).

127. Parsell, D. A., Kowal, A. S., Singer, M. A. & Lindquist, S. Protein disaggregation mediated by heat-shock protein Hsp104. Nature 372, 475-478 (1994).

128. Motohashi, K., Watanabe, Y., Yohda, M. & Yoshida, M. Heat-inactivated proteins are rescued by the DnaK.J-GrpE set and ClpB chaperones. Proc Natl Acad Sci U S A 96, 7184-7189 (1999).

129. Glover, J. R. & Lindquist, S. Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. Cell 94, 73-82 (1998).

130. Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures. J. Bacteriol. 171, 1379-1385 (1989).

131. Kedzierska, S., Staniszewska, M., Wegrzyn, A. & Taylor, A. The role of DnaK/DnaJ and GroEL/GroES systems in the removal of endogenous proteins aggregated by heat-shock from Escherichia coli cells. FEBS Lett 446, 331-337 (1999).

132. Grantcharova, V., Alm, E. J., Baker, D. & Horwich, A. L. Mechanisms of protein folding. Curr. Opin. Struct. Biol. 11, 70-82 (2001).

133. Sakikawa, C., Taguchi, H., Makino, Y. & Yoshida, M. On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. J Biol Chem 274, 21251-21256 (1999).

134. Ranson, N. A. et al. ATP-bound states of GroEL captured by cryo-electron microscopy. Cell 107, 869-879 (2001).

135. Allen, S. P., Polazzi, J. O., Gierse, J. K. & Easton, A. M. Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in Escherichia coli. J. Bacteriol. 174, 6938-6947 (1992).

136. Haslbeck, M. sHsps and their role in the chaperone network. Cell Mol Life Sci. 59, 1649-1657 (2002).

137. Chuang, S. E., Burland, V., Plunkett, G., III, Daniels, D. L. & Blattner, F. R. Sequence analysis of four new heat-shock genes constituting the hslTS/ibpAB and hslVU operons in Escherichia coli. Gene 134, 1-6 (1993).

138. Kuczynska-Wisnik, D. et al. The Escherichia coli small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. Microbiology 148, 1757-1765 (2002).

139. Thomas, J. G. & Baneyx, F. Roles of the Escherichia coli small heat shock proteins IbpA and IbpB in thermal stress management: comparison with ClpA, ClpB, and HtpG In vivo. J Bacteriol 180, 5165-5172 (1998).

140. Shearstone, J. R. & Baneyx, F. Biochemical characterization of the small heat shock protein IbpB from Escherichia coli. J Biol Chem 274, 9937-9945 (1999).

141. Kitagawa, M., Miyakawa, M., Matsumura, Y. & Tsuchido, T. Escherichia coli small heat shock proteins, IbpA and IbpB, protect enzymes from inactivation by heat and oxidants. Eur. J. Biochem. 269, 2907-2917 (2002).

142. Wickner, S., Maurizi, M. R. & Gottesman, S. Posttranslational quality control: folding, refolding, and degrading proteins. Science 286, 1888-1893 (1999).

143. Carrio, M. M., Corchero, J. L. & Villaverde, A. Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. Biochim Biophys Acta 1434, 170-176 (1999).

144. Corchero, J. L., Cubarsi, R., Enfors, S. & Villaverde, A. Limited in vivo proteolysis of aggregated proteins. Biochem Biophys Res Commun 237, 325-330 (1997).

145. Missiakas, D., Schwager, F., Betton, J. M., Georgopoulos, C. & Raina, S. Identification and characterization of HsIV HsIU (ClpQ ClpY) proteins involved in overall proteolysis of misfolded proteins in Escherichia coli. EMBO J. 15, 6899-6909 (1996).

146. Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P. & Bukau, B. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol. Mol Microbiol 40, 397-413 (2001).

147. Wickner, S. et al. A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proc Natl Acad Sci U S A 91, 12218-12222 (1994).

148. Levchenko, I., Luo, L. & Baker, T. A. Disassembly of the Mu transposase tetramer by the ClpX chaperone. Genes Dev. 9, 2399-2408 (1995).

149. Hoskins, J. R., Pak, M., Maurizi, M. R. & Wickner, S. The role of the ClpA chaperone in proteolysis by ClpAP. Proc Natl Acad Sci U S A 95, 12135-12140 (1998).

150. Matouschek, A. Protein unfolding--an important process in vivo? Curr. Opin. Struct. Biol. 13, 98-109 (2003).

151. Wall, J. G. & Pluckthun, A. Effects of overexpressing folding modulators on the in vivo folding of heterologous proteins in Escherichia coli. Curr Opin Biotechnol 6, 507-516 (1995).

152. Chalmers, J. J. et al. Effects of temperature on Escherichia coli overproducing beta-lactamase or human epidermal growth factor. Appl. Environ. Microbiol. 56, 104-111 (1990).

153. Strandberg, L. & Enfors, S. O. Factors influencing inclusion body formation in the production of a fused protein in Escherichia coli. Appl. Environ. Microbiol. 57, 1669-1674 (1991).

154. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol. Bioeng. 96, 1101-1106 (2007).

155. Vasina, J. A. & Baneyx, F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. Protein Expr Purif 9, 211-218 (1997).

156. Niiranen, L. et al. Comparative expression study to increase the solubility of cold adapted Vibrio proteins in Escherichia coli. Protein Expr. Purif. 52, 210-218 (2007).

157. Chesshyre, J. A. & Hipkiss, A. R. Low-Temperatures Stabilize Interferon-Alpha-2 Against Proteolysis in Methylophilus-Methylotrophus and Escherichia-Coli. Applied Microbiology and Biotechnology 31, 158-162 (1989).

158. Schein, C. H. Production of Soluble Recombinant Proteins in Bacteria. Bio-Technology 7, 1141-1147 (1989).

159. Kiefhaber, T., Rudolph, R., Kohler, H. H. & Buchner, J. Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation. Biotechnology (N. Y.) 9, 825-829 (1991).

160. Baneyx, F. & Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. Nat. Biotechnol. 22, 1399-1408 (2004).

161. Hunke, S. & Betton, J. M. Temperature effect on inclusion body formation and stress response in the periplasm of Escherichia coli. Mol Microbiol 50, 1579-1589 (2003).

162. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol. Cell Biochem. 307, 249-264 (2008).

163. Shaw, M. K. & Ingraham, J. L. Synthesis of macromolecules by Escherichia coli near the minimal temperature for growth. J. Bacteriol. 94, 157-164 (1967).

164. Vasina, J. A. & Baneyx, F. Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. Protein Expr Purif 9, 211-218 (1997).

165. Ferrer, M., Chernikova, T. N., Timmis, K. N. & Golyshin, P. N. Expression of a temperature-sensitive esterase in a novel chaperone-based Escherichia coli strain. Appl. Environ. Microbiol. 70, 4499-4504 (2004).

166. Weickert, M. J., Doherty, D. H., Best, E. A. & Olins, P. O. Optimization of heterologous protein production in Escherichia coli. Curr. Opin. Biotechnol. 7, 494-499 (1996).

167. Apiyo, D. & Wittung-Stafshede, P. Presence of the cofactor speeds up folding of Desulfovibrio desulfuricans flavodoxin. Protein Sci. 11, 1129-1135 (2002).

168. Yang, Q., Xu, J., Li, M., Lei, X. & An, L. High-level expression of a soluble snake venom enzyme, gloshedobin, in E. coli in the presence of metal ions. Biotechnol. Lett. 25, 607-610 (2003).

169. Bruser, T., Yano, T., Brune, D. C. & Daldal, F. Membrane targeting of a folded and cofactor-containing protein. Eur. J. Biochem. 270, 1211-1221 (2003).

170. Weickert, M. J., Pagratis, M., Glascock, C. B. & Blackmore, R. A mutation that improves soluble recombinant hemoglobin accumulation in Escherichia coli in heme excess. Appl. Environ. Microbiol. 65, 640-647 (1999).

171. Lewis, G., Taylor, I. W., Nienow, A. W. & Hewitt, C. J. The application of multi-parameter flow cytometry to the study of recombinant Escherichia coli batch fermentation processes. J. Ind. Microbiol. Biotechnol. 31, 311-322 (2004).

172. Sorensen, H. P. & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microb. Cell Fact. 4, 1 (2005).

173. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli. Nature 337, 44-47 (1989).

174. Thomas, J. G. & Baneyx, F. Protein misfolding and inclusion body formation in recombinant Escherichia coli cells overexpressing Heat-shock

proteins. J Biol Chem 271, 11141-11147 (1996).

175. Amrein, K. E. et al. Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an Escherichia coli expression system overproducing the bacterial chaperones GroES and GroEL. Proc. Natl. Acad. Sci. U. S. A 92, 1048-1052 (1995).

176. de Marco, A., Volrath, S., Bruyere, T., Law, M. & Fonne-Pfister, R. Recombinant maize protoporphyrinogen IX oxidase expressed in Escherichia coli forms complexes with GroEL and DnaK chaperones. Protein Expr Purif 20, 81-86 (2000).

177. Ikura, K. et al. Co-overexpression of folding modulators improves the solubility of the recombinant guinea pig liver transglutaminase expressed in Escherichia coli. Prep. Biochem. Biotechnol. 32, 189-205 (2002).

178. de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T. & Bukau, B. Chaperone-based procedure to increase yields of soluble recombinant proteins produced in E. coli. BMC. Biotechnol. 7, 32 (2007).

179. Lee, S. C. & Olins, P. O. Effect of overproduction of heat shock chaperones GroESL and DnaK on human procollagenase production in Escherichia coli. J. Biol. Chem. 267, 2849-2852 (1992).

180. Sato, K., Sato, M. H., Yamaguchi, A. & Yoshida, M. Tetracycline/H+ antiporter was degraded rapidly in Escherichia coli cells when truncated at last transmembrane helix and this degradation was protected by overproduced GroEL/ES. Biochem. Biophys. Res. Commun. 202, 258-264 (1994).

181. Thomas, J. G., Ayling, A. & Baneyx, F. Molecular chaperones, folding catalysts, and the recovery of active

recombinant proteins from E. coli. To fold or to refold. Appl Biochem Biotechnol 66, 197-238 (1997).

182. Rinas, U. & Bailey, J. E. Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli. Appl. Microbiol. Biotechnol. 37, 609-614 (1992).

183. Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J. & King, J. Global suppression of protein folding defects and inclusion body formation. Science 253, 54-58 (1991).

184. Mitraki, A. & King, J. Amino acid substitutions influencing intracellular protein folding pathways. FEBS Lett. 307, 20-25 (1992).

185. Chrunyk, B. A., Evans, J., Lillquist, J., Young, P. & Wetzel, R. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. J. Biol. Chem. 268, 18053-18061 (1993).

186. Forrer, P. & Jaussi, R. High-level expression of soluble heterologous proteins in the cytoplasm of Escherichia coli by fusion to the bacteriophage lambda head protein D. Gene 224, 45-52 (1998).

187. LaVallie, E. R. et al. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. Biotechnology (N. Y.) 11, 187-193 (1993).

188. Esposito, D. & Chatterjee, D. K. Enhancement of soluble protein expression through the use of fusion tags. Curr. Opin. Biotechnol. 17, 353-358 (2006).

189. Zhang, Y. B. et al. Protein aggregation during overexpression limited by peptide extensions with large net negative charge. Protein Expr. Purif. 36, 207-216 (2004).

190. Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. Degradation of abnormal proteins in Escherichia coli. Formation of protein inclusions in cells exposed to amino acid analogs. J. Biol. Chem. 250, 1112-1122 (1975).

191. Marston, F. A. The purification of eukaryotic polypeptides synthesized in Escherichia coli. Biochem. J. 240, 1-12 (1986).

192. Schein, C. H. Solubility as a function of protein structure and solvent components. Biotechnology (N. Y.) 8, 308-317 (1990).

193. Ivanova, M. I., Sawaya, M. R., Gingery, M., Attinger, A. & Eisenberg, D. An amyloid-forming segment of beta2-microglobulin suggests a molecular model for the fibril. Proc. Natl. Acad. Sci. U. S. A 101, 10584-10589 (2004).

194. Ventura, S. et al. Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case. Proc. Natl. Acad. Sci. U. S. A 101, 7258-7263 (2004).

195. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat. Biotechnol. 22, 1302-1306 (2004).

196. Linding, R., Schymkowitz, J., Rousseau, F., Diella, F. & Serrano, L. A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. J. Mol. Biol. 342, 345-353 (2004).

197. Dubay, K. F. et al. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. J. Mol. Biol. 341, 1317-1326 (2004).

198. Conchillo-Sole, O. et al. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. Bmc Bioinformatics 8, (2007).

199. Lopez de la, P. M. & Serrano, L. Sequence determinants of amyloid fibril formation. Proc. Natl. Acad. Sci. U. S. A 101, 87-92 (2004).

200. Taylor, G., Hoare, M., Gray, D. R. & Marston, F. A. O. Size and Density of Protein Inclusion-Bodies. Bio-Technology 4, 553-557 (1986).

201. Carrio, M. M., Corchero, J. L. & Villaverde, A. Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. FEMS Microbiol Lett 169, 9-15 (1998).

202. Bowden, G. A., Paredes, A. M. & Georgiou, G. Structure and morphology of protein inclusion bodies in Escherichia coli. Biotechnology (N. Y.) 9, 725-730 (1991).

203. Arie, J. P., Miot, M., Sassoon, N. & Betton, J. M. Formation of active inclusion bodies in the periplasm of Escherichia coli. Mol. Microbiol. 62, 427-437 (2006).

204. Carrio, M. M., Cubarsi, R. & Villaverde, A. Fine architecture of bacterial inclusion bodies. FEBS Lett 471, 7-11 (2000).

205. Singh, S. M. & Panda, A. K. Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. 99, 303-310 (2005).

206. Rinas, U. & Bailey, J. E. Overexpression of bacterial hemoglobin causes incorporation of pre-beta-lactamase into cytoplasmic inclusion bodies. Appl. Environ. Microbiol. 59, 561-566 (1993).

207. Georgiou, G. & Valax, P. Isolating inclusion bodies from bacteria. Methods Enzymol. 309, 48-58 (1999).

208. Carrio, M. M. & Villaverde, A. Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. J. Bacteriol. 187, 3599-3601 (2005).

209. King, J., Haase-Pettingell, C., Robinson, A. S., Speed, M. & Mitraki, A. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. FASEB J. 10, 57-66 (1996).

210. Mogk, A., Mayer, M. P. & Deuerling, E. Mechanisms of protein folding: molecular chaperones and their application in biotechnology. Chembiochem. 3, 807-814 (2002).

211. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. J. Mol. Biol. 347, 1025-1037 (2005).

212. Speed, M. A., Wang, D. I. & King, J. Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. Nat. Biotechnol. 14, 1283-1287 (1996).

213. Rajan, R. S., Illing, M. E., Bence, N. F. & Kopito, R. R. Specificity in intracellular protein aggregation and inclusion body formation. Proc. Natl. Acad. Sci. U. S. A 98, 13060-13065 (2001).

214. Hart, R. A., Rinas, U. & Bailey, J. E. Protein composition of Vitreoscilla hemoglobin inclusion bodies produced in Escherichia coli. J. Biol. Chem. 265, 12728-12733 (1990).

215. Hellebust H., Murby M., Abrahamsen L. & Enfors SO. Different approaches to stabilize a recombinant fusion protein. Biotechnology 7, 165-168 (1989).

216. Corchero, J. L., Viaplana, E., Benito, A. & Villaverde, A. The position of the heterologous domain can influence the solubility and proteolysis of beta-galactosidase fusion proteins in E. coli. J Biotechnol 48, 191-200 (1996).

217. Meyer, D. E. & Chilkoti, A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. Nat. Biotechnol. 17, 1112-1115 (1999).

218. Rudolph, R. & Lilie, H. In vitro folding of inclusion body proteins. FASEB J. 10, 49-56 (1996).

219. Vallejo, L. F. et al. Renaturation and purification of bone morphogenetic protein-2 produced as inclusion bodies in high-cell-density cultures of recombinant Escherichia coli. J. Biotechnol. 94, 185-194 (2002).

220. Clark, E. D. Protein refolding for industrial processes. Curr. Opin. Biotechnol. 12, 202-207 (2001).

221. Vallejo, L. F. & Rinas, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb. Cell Fact. 3, 11 (2004).

222. Middelberg, A. P. Process-scale disruption of microorganisms. Biotechnol Adv. 13, 491-551 (1995).

223. Valax, P. & Georgiou, G. Molecular characterization of beta-lactamase inclusion bodies produced in Escherichia coli. 1. Composition. Biotechnol. Prog. 9, 539-547 (1993).

224. Schroder, B., Elsasser, H. P., Schmidt, B. & Hasilik, A. Characterisation of lipofuscin-like lysosomal inclusion bodies from human placenta. FEBS Lett. 581, 102-108 (2007).

225. de Marco, A. & Schroedel, A. Characterization of the aggregates formed during recombinant protein expression in bacteria. BMC. Biochem 6, 10 (2005).

226. Singh, S. M. & Panda, A. K. Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. 99, 303-310 (2005).

227. De Bernardez, C. E., Hevehan, D., Szela, S. & Maachupalli-Reddy, J. Oxidative renaturation of hen egg-white lysozyme. Folding vs aggregation. Biotechnol. Prog. 14, 47-54 (1998).

228. Middelberg, A. P. Preparative protein refolding. Trends Biotechnol. 20, 437-443 (2002).

229. Cejka, J., Vodrazka, Z. & Salak, J. Carbamylation of globin in electrophoresis and chromatography in the presence of urea. Biochim. Biophys. Acta 154, 589-591 (1968).

230. Hagel, P., Gerding, J. J., Fieggen, W. & Bloemendal, H. Cyanate formation in solutions of urea. I. Calculation of cyanate concentrations at different temperature and pH. Biochim. Biophys. Acta 243, 366-373 (1971).

231. Puri, N. K. et al. Solubilization of growth hormone and other recombinant proteins from Escherichia coli inclusion bodies by using a cationic surfactant. Biochem. J. 285 (Pt 3), 871-879 (1992).

232. Stempfer, G. & Rudolph, R. Improved refolding of a matrix bound fusion protein. Ann. NY Acad. Sci. 782, 506-512 (1996).

233. Hamaker, K. H., Liu, J., Seely, R. J., Ladisch, C. M. & Ladisch, M. R. Chromatography for rapid buffer exchange and refolding of secretory leukocyte protease inhibitor. Biotechnol. Prog. 12, 184-189 (1996).

234. Tran-Moseman, A., Schauer, N. & De Bernardez, C. E. Renaturation of Escherichia coli-derived recombinant human macrophage colony-stimulating factor. Protein Expr. Purif. 16, 181-189 (1999).

235. Tikhonov, R. V. et al. Recombinant human insulin IX. Investigation of factors, influencing the folding of fusion protein-S-sulfonates, biotechnological precursors of human insulin. Protein Expr. Purif. 26, 187-193 (2002).

236. Maachupalli-Reddy, J., Kelley, B. D. & De Bernardez, C. E. Effect of inclusion body contaminants on the oxidative renaturation of hen egg white

lysozyme. Biotechnol. Prog. 13, 144-150 (1997).

237. Batas, B., Schiraldi, C. & Chaudhuri, J. B. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. J. Biotechnol. 68, 149-158 (1999).

238. Thatcher, D. R. & Hitchcock, A. Protein Folding in Biotechnology. In Mechanisms of Protein Folding. Pain, R. H. Ed. and IRL Press: Oxford. 229-261. 1994.
Ref Type: Serial (Book,Monograph)

239. Dill, K. A. & Shortle, D. Denatured states of proteins. Annu. Rev. Biochem. 60, 795-825 (1991).

240. Vallejo, L. F. & Rinas, U. Optimized procedure for renaturation of recombinant human bone morphogenetic protein-2 at high protein concentration. Biotechnol. Bioeng. 85, 601-609 (2004).

241. Gu, Z., Su, Z. & Janson, J. C. Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding. J. Chromatogr. A 918, 311-318 (2001).

242. Arakawa, T. & Tsumoto, K. The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. Biochem. Biophys. Res. Commun. 304, 148-152 (2003).

243. Rudolph, R. Renaturation of recombinant, disulfide-bonded proteins from inclusion bodies. In: Modern Methods in Protein and Nucleic Acid Research 149-171 (1990).

244. Katoh, S. & Katoh, Y. Continuous refolding of lysozyme with fed-batch addition of denatured protein solution. Process Biochem. 35, 1124 (2000).

245. West, S. M., Chaudhuri, J. B. & Howell, J. A. Improved protein refolding using hollow-fibre membrane dialysis. Biotechnol Bioeng 57, 590-599 (1998).

246. Goldberg, M. E., Rudolph, R. & Jaenicke, R. A Kinetic-Study of the Competition Between Renaturation and Aggregation During the Refolding of Denatured Reduced Egg-White Lysozyme. Biochemistry 30, 2790-2797 (1991).

247. Kim, K., Ramanathan, R. & Frieden, C. Intestinal fatty acid binding protein: a specific residue in one turn appears to stabilize the native structure and be responsible for slow refolding. Protein Sci. 6, 364-372 (1997).

248. Batas, B., Jones, H. R. & Chaudhuri, J. B. Studies of the hydrodynamic volume changes that occur during refolding of lysozyme using size-exclusion chromatography. J. Chromatogr. A 766, 109-119 (1997).

249. Werner, M. H., Clore, G. M., Gronenborn, A. M., Kondoh, A. & Fisher, R. J. Refolding proteins by gel filtration chromatography. FEBS Lett. 345, 125-130 (1994).

250. Li, M., Poliakov, A., Danielson, U. H., Su, Z. & Janson, J. C. Refolding of a recombinant full-length non-structural (NS3) protein from hepatitis C virus by chromatographic procedures. Biotechnol. Lett. 25, 1729-1734 (2003).

251. Suttnar, J., Dyr, J. E., Hamsikova, E., Novak, J. & Vonka, V. Procedure for refolding and purification of recombinant proteins from Escherichia coli inclusion bodies using a strong anion exchanger. J. Chromatogr. B Biomed. Appl. 656, 123-126 (1994).

252. Creighton TE Folding of proteins adsorbed reversibly to ion-exchange resins. UCLA Symp. Mol. Cell Biol. 39, 249-257 (1986).

253. Li M., Zhang G.F. & Su Z. Dual gradient ion-exchange chromatography improved refolding yield of lysozyme. J Chromatogr. A 959, 113-120 (2002).

254. Negro, A., Onisto, M., Grassato, L., Caenazzo, C. & Garbisa, S.

Recombinant human TIMP-3 from Escherichia coli: synthesis, refolding, physico-chemical and functional insights. Protein Eng 10, 593-599 (1997).

255. Rehm, B. H., Qi, Q., Beermann, B. B., Hinz, H. J. & Steinbuchel, A. Matrix-assisted in vitro refolding of Pseudomonas aeruginosa class II polyhydroxyalkanoate synthase from inclusion bodies produced in recombinant Escherichia coli. Biochem. J 358, 263-268 (2001).

256. Cho, T. H., Ahn, S. J. & Lee, E. K. Refolding of protein inclusion bodies directly from E. coli homogenate using expanded bed adsorption chromatography. Bioseparation. 10, 189-196 (2001).

257. Ahmed, A. K., Schaffer, S. W. & Wetlaufer, D. B. Nonenzymic reactivation of reduced bovine pancreatic ribonuclease by air oxidation and by glutathione oxidoreduction buffers. J. Biol. Chem. 250, 8477-8482 (1975).

258. Menzella, H. G., Gramajo, H. C. & Ceccarelli, E. A. High recovery of prochymosin from inclusion bodies using controlled air oxidation. Protein Expr. Purif. 25, 248-255 (2002).

259. Kuhelj, R., Dolinar, M., Pungercar, J. & Turk, V. The preparation of catalytically active human cathepsin B from its precursor expressed in Escherichia coli in the form of inclusion bodies. Eur. J. Biochem. 229, 533-539 (1995).

260. Oberg, K., Chrunyk, B. A., Wetzel, R. & Fink, A. L. Nativelike secondary structure in interleukin-1 beta inclusion bodies by attenuated total reflectance FTIR. Biochemistry 33, 2628-2634 (1994).

261. Henrissat, B. & Bairoch, A. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 293 (Pt 3), 781-788 (1993).

262. Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases. Curr. Opin. Struct. Biol. 7, 637-644 (1997).

263. Fowler, A. V. & Zabin, I. The amino acid sequence of beta galactosidase. I. Isolation and composition of tryptic peptides. J. Biol. Chem. 245, 5032-5041 (1970).

264. Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W. Three-dimensional structure of beta-galactosidase from E. coli. Nature 369, 761-766 (1994).

265. Marchesi, S. L., Steers E Jr & Shifrin, S. Purification and characterization of the multiple forms of beta-galactosidase of Escherichia coli. Biochim. Biophys. Acta 181, 20-34 (1969).

266. Matthews, B. W. The structure of E. coli beta-galactosidase. C. R. Biol. 328, 549-556 (2005).

267. Jacob F & Monod J Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3, 318-356 (1961).

268. Huber, R. E., Kurz, G. & Wallenfels, K. A quantitation of the factors which affect the hydrolase and transgalactosylase activities of beta-galactosidase (E. coli) on lactose. Biochemistry 15, 1994-2001 (1976).

269. Richard, J. P., Huber, R. E., Heo, C., Amyes, T. L. & Lin, S. Structure-reactivity relationships for beta-galactosidase (Escherichia coli, lac Z). 4. Mechanism for reaction of nucleophiles with the galactosyl-enzyme intermediates of E461G and E461Q beta-galactosidases. Biochemistry 35, 12387-12401 (1996).

270. Ring, M., Bader, D. E. & Huber, R. E. Site-directed mutagenesis of beta-galactosidase (E. coli) reveals that tyr-

503 is essential for activity. Biochem. Biophys. Res. Commun. 152, 1050-1055 (1988).

271. Gebler, J. C., Aebersold, R. & Withers, S. G. Glu-537, not Glu-461, is the nucleophile in the active site of (lac Z) beta-galactosidase from Escherichia coli. J. Biol. Chem. 267, 11126-11130 (1992).

272. Roth, N. J. & Huber, R. E. The beta-galactosidase (Escherichia coli) reaction is partly facilitated by interactions of His-540 with the C6 hydroxyl of galactose. J. Biol. Chem. 271, 14296-14301 (1996).

273. Martinez-Bilbao, M. & Huber, R. E. Substitutions for Gly-794 show that binding interactions are important determinants of the catalytic action of beta-galactosidase (Escherichia coli). Biochem. Cell Biol. 72, 313-319 (1994).

274. Wallenfels, K. & Weil R. The enzymes. Boyer P. D. Acad. Press New York Ed. vol. VII cap. 20, 617-663 (1992).

275. Creuger W A textbook of Industrial Microbiology. Biotechnology 1-275 (1984).

276. Ullmann, A. Complementation in beta-galactosidase: from protein structure to genetic engineering. Bioessays 14, 201-205 (1992).

277. Schwartz, D. & Beckwith, J. R. The lactose operon. Cold Spring Harb. Lab. , New York (1970).

278. Miller, J. H. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 466 pp. (1972).

279. Fowler, A. V., Brake, A. J. & Zabin, I. Amino acid sequence of beta-galactosidase. X. Sequence of the COOH-terminal segment, CNBr peptides 18 to 24, residues 654 to 1021. J. Biol. Chem. 253, 5515-5520 (1978).

280. Fowler, A. V. & Zabin, I. Amino acid sequence of beta-galactosidase. VIII. Sequence of the NH2-terminal segment, CNBr peptides 1 to 9, residues 1 to 377. J. Biol. Chem. 253, 5505-5509 (1978).

281. Casadaban, M. J., Chou, J. & Cohen, S. N. In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: Escherichia coli plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. 143, 971-980 (1980).

282. Howell, M. L. & Blumenthal, K. M. Cloning and expression of a synthetic gene for Cerebratulus lacteus neurotoxin B-IV. J. Biol. Chem. 264, 15268-15273 (1989).

283. Khudyakov, Y. E. et al. Artificial mosaic protein containing antigenic epitopes of hepatitis E virus. J. Virol. 68, 7067-7074 (1994).

284. Broekhuijsen, M. P. et al. Fusion proteins with multiple copies of the major antigenic determinant of foot-and-mouth disease virus protect both the natural host and laboratory animals. J. Gen. Virol. 68 (Pt 12), 3137-3143 (1987).

285. Baum, E. Z., Bebernitz, G. A. & Gluzman, Y. beta-Galactosidase containing a human immunodeficiency virus protease cleavage site is cleaved and inactivated by human immunodeficiency virus protease. Proc. Natl. Acad. Sci. U. S. A 87, 10023-10027 (1990).

286. Sassenfeld, H. M. Engineering proteins for purification. Trends Biotechnol 8, 88-93 (1990).

287. Morin, J. G. & Hastings, J. W. Energy transfer in a bioluminescent system. J. Cell Physiol 77, 313-318 (1971).

288. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier,

M. J. Primary structure of the Aequorea victoria green-fluorescent protein. Gene 111, 229-233 (1992).

289. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. Science 263, 802-805 (1994).

290. Inouye, S. & Tsuji, F. I. Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. FEBS Lett. 341, 277-280 (1994).

291. Yang, F., Moss, L. G. & Phillips, G. N., Jr. The molecular structure of green fluorescent protein. Nat. Biotechnol. 14, 1246-1251 (1996).

292. Ormo, M. et al. Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392-1395 (1996).

293. Brejc, K. et al. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. Proc. Natl. Acad. Sci. U. S. A 94, 2306-2311 (1997).

294. Ward, W. W. & Cormier, M. J. An energy transfer protein in coelenterate bioluminescence. Characterization of the Renilla green-fluorescent protein. J. Biol. Chem. 254, 781-788 (1979).

295. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. Gene 111, 229-233 (1992).

296. Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G. & Ward, W. W. Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. Biochemistry 32, 1212-1218 (1993).

297. Rao, B. K. M. a. P. F. Proton nuclear magnetic resonance and fluorescence spectroscopic studies of segmental mobility in aequorin and a green fluorescent protein from Aequorea forskalea. Biophys. J. 32, 630-632 (1980).

298. Li, X. et al. Deletions of the Aequorea victoria green fluorescent protein define the minimal domain required for fluorescence. J. Biol. Chem. 272, 28545-28549 (1997).

299. Cubitt, A. B. et al. Understanding, improving and using green fluorescent proteins. Trends Biochem. Sci. 20, 448-455 (1995).

300. Swaminathan, R., Hoang, C. P. & Verkman, A. S. Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. Biophys. J. 72, 1900-1907 (1997).

301. Scheyhing, C. H., Meersman, F., Ehrmann, M. A., Heremans, K. & Vogel, R. F. Temperature-pressure stability of green fluorescent protein: a Fourier transform infrared spectroscopy study. Biopolymers 65, 244-253 (2002).

302. Heim, R., Cubitt, A. B. & Tsien, R. Y. Improved green fluorescence. Nature 373, 663-664 (1995).

303. Makino, Y., Amada, K., Taguchi, H. & Yoshida, M. Chaperonin-mediated folding of green fluorescent protein. J. Biol. Chem. 272, 12468-12474 (1997).

304. Wachter, R. M. The family of GFP-like proteins: Structure, function, photophysics and biosensor applications. Introduction and perspective. Photochemistry and Photobiology 82, 339-344 (2006).

305. Giavedoni, L. D., Kaplan, G., Marcovecchio, F., Piccone, M. E. & Palma, E. L. Protection conferred by TrpE fusion proteins containing portions of the C-terminal region of capsid protein VP1 of foot-and-mouth

disease virus. J. Gen. Virol. 72 (Pt 4), 967-971 (1991).

306. Bittle, J. L. et al. Protection against foot-and-mouth disease by immunization with a chemically synthesized peptide predicted from the viral nucleotide sequence. Nature 298, 30-33 (1982).

307. Strohmaier, K., Franze, R. & Adam, K. H. Location and characterization of the antigenic portion of the FMDV immunizing protein. J. Gen. Virol. 59, 295-306 (1982).

308. Mateu, M. G. et al. Reactivity with monoclonal antibodies of viruses from an episode of foot-and-mouth disease. Virus Res. 8, 261-274 (1987).

309. Mateu, M. G. et al. Implications of a quasispecies genome structure: effect of frequent, naturally occurring amino acid substitutions on the antigenicity of foot-and-mouth disease virus. Proc. Natl. Acad. Sci. U. S. A 86, 5883-5887 (1989).

310. Mateu, M. G. et al. A single amino acid substitution affects multiple overlapping epitopes in the major antigenic site of foot-and-mouth disease virus of serotype C. J. Gen. Virol. 71 (Pt 3), 629-637 (1990).

311. Fox, G. et al. The cell attachment site on foot-and-mouth disease virus includes the amino acid sequence RGD (arginine-glycine-aspartic acid). J. Gen. Virol. 70 (Pt 3), 625-637 (1989).

312. Tsien, R. Y. The green fluorescent protein. Annu. Rev. Biochem. 67, 509-544 (1998).

313. Baldwin, R. L. Temperature dependence of the hydrophobic interaction in protein folding. Proc. Natl. Acad. Sci. U. S. A 83, 8069-8072 (1986).

314. Thomas, J. G. & Baneyx, F. Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant Escherichia coli. Protein Expr Purif 11, 289-296 (1997).

315. Villaverde, A. & Carrio, M. M. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnol Lett 25, 1385-1395 (2003).

316. Fahnert, B., Lilie, H. & Neubauer, P. Inclusion bodies: formation and utilisation. Adv. Biochem. Eng Biotechnol 89, 93-142 (2004).

317. Worrall, D. M. & Goss, N. H. The formation of biologically active beta-galactosidase inclusion bodies in Escherichia coli. Aust. J Biotechnol 3, 28-32 (1989).

318. Tokatlidis, K., Dhurjati, P., Millet, J., Beguin, P. & Aubert, J. P. High activity of inclusion bodies formed in Escherichia coli overproducing Clostridium thermocellum endoglucanase D. FEBS Lett 282, 205-208 (1991).

319. Carrio, M. M. & Villaverde, A. Role of molecular chaperones in inclusion body formation. FEBS Lett 537, 215-221 (2003).

320. Petersson, L., Carrio, M. M., Vera, A. & Villaverde, A. The impact of dnaKJ overexpression on recombinant protein solubility results from antagonistic effects on the control of protein quality. Biotechnol Lett 26, 595-601 (2004).

321. Lopes Ferreira, N. & Alix, J. H. The DnaK chaperone is necessary for alpha-complementation of beta-galactosidase in Escherichia coli. J Bacteriol 184, 7047-7054 (2002).

322. Vila, P., Corchero, J. L., Cubarsi, R. & Villaverde, A. Enhanced fitness of recombinant protein synthesis within the stationary phase of E. coli batch cultures. Biotechnology Letters 225-228 (1997).

323. Friehs, K. & Reardon, K. F. Parameters influencing the productivity of recombinant E. coli cultivations. Adv Biochem Eng Biotechnol 48, 53-77 (1993).

324. Murby, M., Uhlen, M. & Stahl, S. Upstream strategies to minimize proteolytic degradation upon recombinant production in Escherichia coli. Protein Expr Purif 7, 129-136 (1996).

325. Schein, C. H. Optimizing protein folding to the native state in bacteria. Curr Opin Biotechnol 2, 746-750 (1991).

326. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol. Bioeng. 96, 1101-1106 (2007).

327. Cubarsi, R., Carrio, M. M. & Villaverde, A. In situ proteolytic digestion of inclusion body polypeptides occurs as a cascade process. Biochem Biophys Res Commun 282, 436-441 (2001).

328. Cubarsi, R., Carrio, M. M. & Villaverde, A. A mathematical approach to molecular organization and proteolytic disintegration of bacterial inclusion bodies. Math. Med. Biol. (2005).

329. Przybycien, T. M., Dunn, J. P., Valax, P. & Georgiou, G. Secondary structure characterization of beta-lactamase inclusion bodies. Protein Eng 7, 131-136 (1994).

330. Georgiou, G., Valax, P., Ostermeier, M. & Horowitz, P. M. Folding and aggregation of TEM beta-lactamase: analogies with the formation of inclusion bodies in Escherichia coli. Protein Sci. 3, 1953-1960 (1994).

331. Ami, D. et al. FT-IR study of heterologous protein expression in recombinant Escherichia coli strains. Biochim. Biophys. Acta 1624, 6-10 (2003).

332. Ami, D., Natalello, A., Taylor, G., Tonon, G. & Maria, D. S. Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. Biochim. Biophys. Acta 1764, 793-799 (2006).

333. Ami, D., Natalello, A., Gatti-Lafranconi, P., Lotti, M. & Doglia, S. M. Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. FEBS Lett 579, 3433-3436 (2005).

334. Mogk, A. et al. Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. J Biol Chem 278, 17615-17624 (2003).

335. Schlieker, C. et al. Substrate recognition by the AAA+ chaperone ClpB. Nat. Struct. Mol. Biol. 11, 607-615 (2004).

336. Plakoutsi, G. et al. Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. J Mol Biol 351, 910-922 (2005).

337. Kundu, B. & Guptasarma, P. Use of a hydrophobic dye to indirectly probe the structural organization and conformational plasticity of molecules in amorphous aggregates of carbonic anhydrase. Biochem. Biophys. Res. Commun. 293, 572-577 (2002).

338. Dong, A., Caughey, B., Caughey, W. S., Bhat, K. S. & Coe, J. E. Secondary structure of the pentraxin female protein in water determined by infrared spectroscopy: effects of calcium and phosphorylcholine. Biochemistry 31, 9364-9370 (1992).

339. Natalello, A., Ami, D., Brocca, S., Lotti, M. & Doglia, S. M. Secondary structure, conformational stability and glycosylation of a recombinant Candida rugosa lipase studied by Fourier-transform infrared spectroscopy. Biochem. J 385, 511-517 (2005).

340. Kuczynska-Wisnik, D. et al. Escherichia coli small heat shock proteins IbpA/B enhance activity of enzymes sequestered in inclusion bodies. Acta Biochim. Pol. 51, 925-931 (2004).

341. Martinez-Alonso, M., Vera, A. & Villaverde, A. Role of the chaperone DnaK in protein solubility and conformational quality in inclusion body-forming Escherichia coli cells. FEMS Microbiol. Lett. 273, 187-195 (2007).

342. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol Bioeng 96, 1101-1106 (2007).

343. Nahalka, J. & Nidetzky, B. Fusion to a pull-down domain: a novel approach of producing Trigonopsis variabilisD-amino acid oxidase as insoluble enzyme aggregates. Biotechnol Bioeng 97, 454-461 (2007).

344. Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D. & Arakawa, T. Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine. Biochem. Biophys. Res. Commun. 312, 1383-1386 (2003).

345. Nahalka, J., Gemeiner, P., Bucko, M. & Wang, P. G. Bioenergy beads: a tool for regeneration of ATP/NTP in biocatalytic synthesis. Artif. Cells Blood Substit. Immobil. Biotechnol 34, 515-521 (2006).

346. Nahalka, J., Vikartovska, A. & Hrabarova, E. A crosslinked inclusion body process for sialic acid synthesis. J Biotechnol 134, 146-153 (2008).

347. Nahálka J. Physiological aggregation of maltodextrin phosphorylase from Pyrococcus furiosus and its application in a process of batch starch degradation to alpha-D: -glucose-1-phosphate. J Ind Microbiol Biotechnol. In press, (2007).

348. Arie, J. P., Miot, M., Sassoon, N. & Betton, J. M. Formation of active inclusion bodies in the periplasm of Escherichia coli. Mol. Microbiol. 62, 427-437 (2006).

349. de Groot, N. S. & Ventura, S. Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. J. Biotechnol. 125, 110-113 (2006).

350. Rochet, J. C. & Lansbury, P. T., Jr. Amyloid fibrillogenesis: themes and variations. Curr. Opin. Struct. Biol. 10, 60-68 (2000).

351. Jevsevar, S. et al. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. Biotechnol. Prog. 21, 632-639 (2005).

352. Ventura, S. & Villaverde, A. Protein quality in bacterial inclusion bodies. Trends Biotechnol. 24, 179-185 (2006).

353. Jurgen, B. et al. Monitoring of genes that respond to overproduction of an insoluble recombinant protein in Escherichia coli glucose-limited fed-batch fermentations. Biotechnol. Bioeng. 70, 217-224 (2000).

354. Lethanh, H., Neubauer, P. & Hoffmann, F. The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant Escherichia coli and delay degradation of inclusion bodies. Microb. Cell Fact. 4, 6 (2005).

355. Boels, K., Carrió, M. M., Arís, A., Corchero, J. L. & Villaverde, A. Distinct chaperone affinity to folding variants of homologous recombinant proteins. Biotechnology Letters 21, 531-536 (1999).

356. Clark, E. D. B. Refolding of recombinant proteins. Curr. Opin. Biotechnol 9, 157-163 (1998).

357. Misawa, S. & Kumagai, I. Refolding of therapeutic proteins produced in

Escherichia coli as inclusion bodies. Biopolymers 51, 297-307 (1999).

358. Shi, P. Y., Maizels, N. & Weiner, A. M. Recovery of soluble, active recombinant protein from inclusion bodies. Biotechniques 23, 1036-1038 (1997).

359. Guise, A. D., West, S. M. & Chaudhuri, J. B. Protein folding in vivo and renaturation of recombinant proteins from inclusion bodies. Mol. Biotechnol 6, 53-64 (1996).

360. Panda, A. K. Bioprocessing of therapeutic proteins from the inclusion bodies of Escherichia coli. Adv. Biochem. Eng Biotechnol 85, 43-93 (2003).

361. Tishkov, V. I. & Khoronenkova, S. V. D-Amino acid oxidase: structure, catalytic mechanism, and practical application. Biochemistry (Mosc.) 70, 40-54 (2005).

362. Pilone MS & Pollegioni L D-amino acid oxidase as an industrial biocatalyst. Biocatal. Biotrans. 20, 145-159 (2002).

363. Riethorst W & Reichert A An industrial view on enzymes for the cleavage of cephalosporin C. C. Chimia 53, 600-607 (1999).

364. Buchholz K, Kasche V & Bornscheuer UT Biocatalysts and enzyme technology. Weinheim: Wiley-VCH (2005).

365. Traving, C. & Schauer, R. Structure, function and metabolism of sialic acids. Cell Mol Life Sci. 54, 1330-1349 (1998).

366. Uchida Y, Tsukada Y & Sugimori T Distribution of N-acetylneuraminate lyase in bacteria and its production by Escherichia coli. Agric. Biol. Chem. 49, 181-187 (1985).

367. Vasina, J. A. & Baneyx, F. Recombinant protein expression at low temperatures under the transcriptional control of the major

Escherichia coli cold shock promoter cspA. Appl Environ Microbiol 62, 1444-1447 (1996).

368. Leandro, P., Lechner, M. C., Tavares, d. A., I & Konecki, D. Glycerol increases the yield and activity of human phenylalanine hydroxylase mutant enzymes produced in a prokaryotic expression system. Mol Genet. Metab 73, 173-178 (2001).

369. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature 424, 805-808 (2003).

370. Ventura, S. Sequence determinants of protein aggregation: tools to increase protein solubility. Microb. Cell Fact. 4, 11 (2005).

371. Stegemann, J., Ventzki, R., Schrodel, A. & de, M. A. Comparative analysis of protein aggregates by blue native electrophoresis and subsequent sodium dodecyl sulfate-polyacrylamide gel electrophoresis in a three-dimensional geometry gel. Proteomics. 5, 2002-2009 (2005).

372. Aris, A. & Villaverde, A. Molecular organization of protein-DNA complexes for cell-targeted DNA delivery. Biochem. Biophys. Res. Commun. 278, 455-461 (2000).

373. Sachdev, D. & Chirgwin, J. M. Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. J Protein Chem. 18, 127-136 (1999).

374. Nomine, Y. et al. Formation of soluble inclusion bodies by hpv e6 oncoprotein fused to maltose-binding protein. Protein Expr. Purif. 23, 22-32 (2001).

375. Vera, A., Gonzalez-Montalban, N., Aris, A. & Villaverde, A. The conformational quality of insoluble recombinant proteins is enhanced at low growth

temperatures. Biotechnol. Bioeng. 96, 1101-1106 (2007).

376. Herberhold, H. et al. Characterization of the pressure-induced intermediate and unfolded state of red-shifted green fluorescent protein--a static and kinetic FTIR, UV/VIS and fluorescence spectroscopy study. J. Mol Biol. 330, 1153-1164 (2003).

377. Zhang, L., Patel, H. N., Lappe, J. W. & Wachter, R. M. Reaction progress of chromophore biogenesis in green fluorescent protein. J Am. Chem. Soc. 128, 4766-4772 (2006).

378. de Groot, N. S. & Ventura, S. Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. J. Biotechnol 125, 110-113 (2006).

379. de Groot, N. S. & Ventura, S. Effect of temperature on protein quality in bacterial inclusion bodies. FEBS Lett. 580, 6471-6476 (2006).

380. Goff, S. A. & Goldberg, A. L. Production of abnormal proteins in E. coli stimulates transcription of lon and other heat shock genes. Cell 41, 587-595 (1985).

381. Rinas, U. Synthesis rates of cellular proteins involved in translation and protein folding are strongly altered in response to overproduction of basic fibroblast growth factor by recombinant Escherichia coli. Biotechnol. Prog. 12, 196-200 (1996).

382. Hoffmann, F. & Rinas, U. Kinetics of heat-shock response and inclusion body formation during temperature-induced production of basic fibroblast growth factor in high-cell-density cultures of recombinant Escherichia coli. Biotechnol Prog 16, 1000-1007 (2000).

383. Miot, M. & Betton, J. M. Protein quality control in the bacterial periplasm. Microb. Cell Fact. 3, 4 (2004).

384. Maurizi, M. R. Proteases and protein degradation in Escherichia coli. Experientia 48, 178-201 (1992).

385. Rozkov, A. & Enfors, S. O. Analysis and control of proteolysis of recombinant proteins in Escherichia coli. Adv Biochem Eng Biotechnol 89, 163-195 (2004).

386. Vera, A., Aris, A., Carrio, M., Gonzalez-Montalban, N. & Villaverde, A. Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies. J. Biotechnol. 119, 163-171 (2005).

387. Jubete, Y., Maurizi, M. R. & Gottesman, S. Role of the heat shock protein DnaJ in the lon-dependent degradation of naturally unstable proteins. J Biol Chem 271, 30798-30803 (1996).

388. Sherman, M. Y. & Goldberg, A. L. Involvement of the chaperonin dnaK in the rapid degradation of a mutant protein in Escherichia coli. EMBO J. 11, 71-77 (1992).

389. Han, M. J., Lee, J. W. & Lee, S. Y. Enhanced proteome profiling by inhibiting proteolysis with small heat shock proteins. J Proteome. Res. 4, 2429-2434 (2005).

390. Baneyx, F. & Palumbo, J. L. Improving heterologous protein folding via molecular chaperone and foldase co-expression. Methods Mol Biol 205, 171-197 (2003).

391. Carbonell, X. & Villaverde, A. Protein aggregated into bacterial inclusion bodies does not result in protection from proteolytic digestion. Biotechnology Letters 24, 1939-1944 (2002).

X.Acknowledgements

Ara ha arribat el moment de fer el resum d'una etapa molt important de la meva vida, una etapa de la qual guardaré sempre el record dels bons moments, dels experiments, de les troballes emocionants…però, per sobretot de tot, el record de totes aquelles persones que d'alguna manera amb la seva ajuda, comprensió i dedicació han aportat el seu granet de sorra a aquest projecte i, al cap i a la fi, l'han fet possible.

En primer lloc voldria donar unes gràcies majúscules i molt especials al Toni per obrir-me les portes de l'apassionant món de la ciència. De fet, m'agradaria agrair-li moltíssimes coses, però especialment el fet d'haver-me fet tan planer aquest viatge a través del món de la recerca, d'haver estat tan proper i d'haver diposit en tot moment plena confiança en mi. Toni, de tot cor, moltíssimes gràcies per tot!

Evidentment també voldria dedicar unes ratlles a tots els "aplicats", passats i presents, persones que durant aquests anys han fet que el dia a dia al laboratori tingués un toc especial. A la Mar, la persona que em va apropar al món dels cossos d'inclusió, vull agrair-li tots els coneixements que em va transmetre sobre un món que m'ha aportat algun dels moments més gratificants que mai he viscut. L'ajuda, la comprensió i els consells són només algunes de les infinites coses que tinc per agrair-li a l'Anna, una "aplicada" molt especial en la meva vida i que ha estat un pilar clau durant aquesta etapa, tan a nivell laboral com personal. Les coca-coles fresquetes de l'Andrea, una amiga sempre a punt per donar un cop de mà, han aportat moments molt refrescants al nostre dia a dia. El "salero" de la Núria i la Rosa han convertit el laboratori en "algo" molt més especial que un simple lloc de treball. A totes dues, companyes de travessa, els hi voldria dir que aquests anys al laboratori amb elles han estat una experiència única que mai oblidaré. A la Núria m'agradaria agrair-li l'oportunitat que m'ha ofert de compartir amb ella moments fantàstics durant aquests quatre anys, així com el fet d'haver pogut anar descobrint al seu costat els secrets que guardaven els nostres estimats cossos. La paciència i el gran cor de la "Rouse" m'han ajudat molt a superar les situacions difícils i a gaudir de manera molt especial els bons moments viscuts durant aquest viatge per la ciència. A l'Éscar vull agrair-li molt sincerament el seu recolzament, la seva comprensió, els seus petits detalls i el fet d'haver estat un

exemple de constància i dedicació. A les resta de membres del "súper equipo d'aplicats": moltes gràcies per la vostra ajuda, especialment durant el frenètic període d'escriptura de la tesi. Neus, Mònica, Esther, Rosa, Vero, Pepe, Joan i Fela, amb tots vosaltres ha començat una nova etapa al laboratori, una etapa de noves cares, de nous projectes, però sobretot de grans moments. A més a més de tots vosaltres no puc oblidar a la Pili i els seus eppendorfs aprofitats fins el darrer mil·límetre rotulable, al David, que amb la seu tarannà, va aportar una mica de seny en un galliner de nenes i a l'Edu per la seva ajuda amb els clons.

Més enllà de les parets de Micro Aplicada, voldria agrair-li al Sergi el seu ajut en els moments de crisi davant de l'ordenador i la seva companyia durant el dinars al bar de ciències. Tampoc puc oblidar mencionar l'enorme vitalitat i l'ajuda de l'Almu que fan la nostra feina molt més fàcil cada dia. I dins l'àmbit de l'IBB vull donar les gràcies a tots aquells que heu format part dels records viscuts durant aquests anys, molt especialment als nostres veïns d'Immuno.

Molts dels experiments no haguessin estat el mateix sense el suport i la companyia del Salva durant les múltiples estones que he passat davant del fluorímetre i del microscopi. La Mònica sempre m'ha ajudat amb el confocal, transmetent-me en tot moment interès per la meva feina. I l'Àlex, que durant aquest darrer any ha fet un gran esforç posant a punt la tècnica de TEM, m'ha encomanant el seu entusiasme i m'ha mostrat que amb perseverança i paciència tot s'acaba trobant.

Fora de l'àmbit de la universitat, vull agrair als meus pares la seva comprensió front les meves decisions i el seu suport incondicional al llarg de tots aquests anys. Trobar-me cadascun dels meus articles impresos sobre la taula cada cop que han estat publicats i el vostre interès per un món tan diferent han estat el millor regal que em podíeu haver fet mai. Mare, evidentment no podria deixar d'agrair-te tots els tuppers que m'han fet el dia a dia molt més fàcil! A la meva germaneta li vull donar les gràcies per ensenyar-me a veure les coses des d'un punt de vista tan diferent (i per haver portat amb ella de Sevilla el meu nebodet "el patxu", l'alegria de la casa). Els avis, tot i sentir parlar d'un

món tan estrany, sempre m'han ajudat i recolzat, especialment l'avi Joan que m'ha escoltat sempre que ho he necessitat. El suport dels tiets i cosins, que durant algun moment han sentit a parlar dels meus experiments, ha estat també molt important. Diuen que desconnectar una mica sempre va bé per començar amb més ganes i, la veritat, és que el dinars familiars amb els petits de la casa (el Lluc, la Laia i el Biel) han estat una manera de fugir durant unes horetes d'aquesta última etapa de nervis.

La companyia insubstituïble del Marcel i el simple fet de formar part de la meva vida han fet que tot plegat fos molt més fàcil i, senzillament, millor. Sempre a punt per escoltar com m'ha anat el dia, per entendre la importància dels experiments acabats amb èxit i per ajudar-me a minimitzar la importància dels fracassos. Gràcies per estar durant tots aquests anys al meu costat, formant part activament d'aquest projecte tan important per mi.

Tampoc podria deixar passar l'oportunitat de dedicar unes paraules d'agraïment a tota la família de Can Castellet, particularment per l'interès mostrat per la meva feina durant aquests anys.

No podria acabar aquests agraïments oblidant-me de les meves nenetes (i els seus nenets)! Anneta, Maria, Marteta i Laia amb elles vaig descobrir i compartir la meva passió per la Biologia i allò d'alguna manera va ser el pròleg d'aquest projecte, durant el qual sempre han estat al meu costat. Anneta, la meva Anneta, sempre ha cregut en mi i això estic convençuda que m'ha ajudat a arribar fins aquí! Els cafetons i les xerrades amb la Maria m'han aportat aire nou als moments dolents. Al Marc li voldria donar les gràcies per la seva paciència i per deixar-me "robar-li" la Maria de tant en tant. Ja veus el que es pot arribar a fer amb els papus, una tesi! A la Laia, voldria agrair-li l'enorme paciència escoltant les meves històries sobre la tesi. I Marteta, gràcies per sempre estar pendent de com em van les coses.

Evidentment, voldria dir-li a l'Arantza que ha estat genial poder compartir esmorzars i xerrades amb ella. Amb el Gamito i les seves visites al laboratori he viscut moments

molt divertits. Al Javi voldria dedicar-li un agraïment especial pel seu suport incondicional durant tots aquests anys.

Amb la Nàdia i Montse he compartit moltes experiències i per mi ha estat molt gratificant poder compartir aquesta darrera etapa de la meva vida amb elles. Noies, gràcies!

Finalment, vull dedicar unes gràcies molt sinceres al Lluís i a la Sara simplement per ser al meu costat en tot moment.

Ich möchte die Kollegen von Wien für die herzliche Aufnahme im BOKU Labor danken. Ich habe viel bei Ihnen gelernt und eine sehr schöne Zeit verbracht. Besonders möchte ich Didi dafür danken, dass er mir die Chance gegeben hat, bei ihm im Labor zu arbeiten. Ich bin ihm auch sehr dankbar für seine Hilfe in diesen Jahren. Zum schluβ möchte ich Nicole für ihre Hilfe mit dem cytometer und Michael für seine Unterstützung danken. Vielen Dank allerseits!

Voglio inoltre ringranziare Minoska per il suo grande aiuto durante il mio soggiorno a Vienna. È una buona compagna di lavoro, ma soprattutto una buona amica. Grazie mille per tutto!