4.3. Reversibilitat de la formació dels cossos d'inclusió

El fet d'haver trobat que els cossos d'inclusió eren estructures dinàmiques, en les que hi havia una transició entre la forma soluble i la insoluble, ens va portar a analitzar amb més profunditat la reversibilitat de l'agregació. Per això vam estudiar l'evolució dels cossos d'inclusió després d'haver aturat la síntesi proteica. La seva desintegració i l'aparició de proteïna soluble demostra que l'agregació en forma de cossos d'inclusió és reversible. Aquesta troballa ens va portar a formular la hipòtesi de que els cossos d'inclusió eren el resultat d'un equilibri dinàmic entre l'agregació i la solubilització, desplaçat cap a l'agregació, durant la sobreproducció de proteïna recombinant.

FEBS 24826 FEBS Letters 489 (2001) 29–33

Protein aggregation as bacterial inclusion bodies is reversible

M. Mar Carrió, Antonio Villaverde*

Institut de Biologia Fonamental and Departament de Genètica and Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

Received 14 November 2000; accepted 4 December 2000

First published online 8 January 2001

Edited by Felix Wieland

Abstract **Inclusion bodies are refractile, intracellular protein aggregates usually observed in bacteria upon targeted gene overexpression. Since their occurrence has a major economical impact in protein production bio-processes, in vitro refolding strategies are under continuous exploration. In this work, we prove spontaneous in vivo release of both β-galactosidase and P22 tailspike polypeptides from inclusion bodies resulting in their almost complete disintegration and in the concomitant appearance of soluble, properly folded native proteins with full biological activity. Since, in particular, the tailspike protein exhibits an unusually slow and complex folding pathway involving deep interdigitation of β-sheet structures, its in vivo refolding indicates that bacterial inclusion body proteins are not collapsed into an irreversible unfolded state. Then, inclusion bodies can be observed as transient deposits of folding-prone polypeptides, resulting from an unbalanced equilibrium between in vivo protein precipitation and refolding that can be actively displaced by arresting protein synthesis. The observation that the formation of big inclusion bodies is reversible in vivo can be also relevant in the context of amyloid diseases, in which deposition of important amounts of aggregated protein initiates the pathogenic process.** © **2001 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.**

Key words: Inclusion body; Protein aggregation; Refolding; Recombinant protein; β Galactosidase; TSP

1. Introduction

Inclusion bodies (IBs) are major protein aggregates commonly occurring in recombinant bacteria when the expression of plasmid-encoded genes is directed at high rates [1–3]. Despite some physiological factors influencing IB formation have been identified [2], attempts to prevent protein aggregation are in general unsuccessful [3]. Therefore, and since polypeptides embedded in IBs are devoid of any biological activity and therefore usefulness in a biotechnological context, refolding procedures for in vitro protein recovery from purified IBs are under continuous development [4–7]. On the other hand, bacterial IBs are interesting and convenient models for dynamic and structural analysis of protein aggregation, that could serve to better understand biological situations which are difficult to approach experimentally, such as protein deposition in amyloid diseases. Despite this potential interest, bacterial IBs, often viewed as an obstacle in bioproduction processes, have been in general poorly investigated.

In this work, we prove that bacterial IBs are not inert protein aggregates that undergo a parsimonious volumetric grow by product accumulation, but on the contrary, the result of an unbalanced equilibrium between in vivo protein aggregation and solubilisation. Interestingly, this equilibrium can be spontaneously displaced towards protein refolding when protein synthesis is arrested, a situation that conducts to an almost complete IB disintegration and to the appearing of fully active protein forms. This fact can have a significant impact on protein recovery when using bacteria as cell factory. In addition, it offers new data about the general mechanics of protein aggregation and an interesting model to monitor protein aggregation and solubilisation in real time, a possibility that could be relevant also in the context of prionic and other amyloid diseases.

2. Materials and methods

2.1. Bacterial strains, plasmids, proteins and culture conditions

The Lon$^-$ *Escherichia coli* strain BL21 [8] and its LacZ$^-$ derivative BL26 were used for protein production. Plasmid pJVP1LAC is a pJLA602 derivative that encodes a β-galactosidase fusion harbouring the VP1 capsid protein of foot and mouth disease virus (FMDV) [9], whose production is controlled by the lambda p_R and p_L lytic promoters and the temperature-sensitive CI857 repressor. Plasmid pITSPA is a pITe99 derivative that encodes a pseudo-wild-type TSP protein (TSPA) with the full biological activities of TSP [10], whose production is under the control of p_{tac} promoter. Production of both VP1LAC and TSPA in *E. coli* results in IB formation. Luria Bertani (LB) medium [8] plus 100 μg/ml ampicillin was used for culture of recombinant cells, and induction of gene expression was achieved by temperature upshift from 28 to 42°C for VP1LAC and by IPTG addition (up to 1 mM) at 37°C for TSPA. At different times after induction of gene expression, protein synthesis was arrested by chloramphenicol addition (up to 200 μg/ml). In TSPA-producing cultures, media were also washed by centrifugation to remove IPTG, immediately prior to chloramphenicol addition. After chloramphenicol addition, pJVP1LAC-carrying cells were incubated at 28°C to allow a proper folding and activity of CI857^{ts} repressor. For some in vivo experiments, the strain JGT17, a *dihp* derivative of MC4100 [11] was also used.

2.2. IB analysis and purification and determination of protein activity

Procedures for numeric and volumetric IB analysis inside the cells were already described [12], as well as the IB purification protocol by repeated detergent treatment [13]. Since TSP IBs are rather smaller than VP1LAC's, in vivo analyses were done on 24 h- and 3 h-aged IBs respectively. Soluble and insoluble cell fractions were analysed by PAGE and proteins were detected in Western blot by using appropriate sera. β-Galactosidase activity was determined according to Miller's method [14]. TSP activity was determined by plaque counting on confluent cultures of *Salmonella typhimurium* LT2 after incubating cell

*Corresponding author. Fax: (34)-93-5812011.
E-mail: antoni.villaverde@uab.es

PII: S0014-5793(01)02073-7

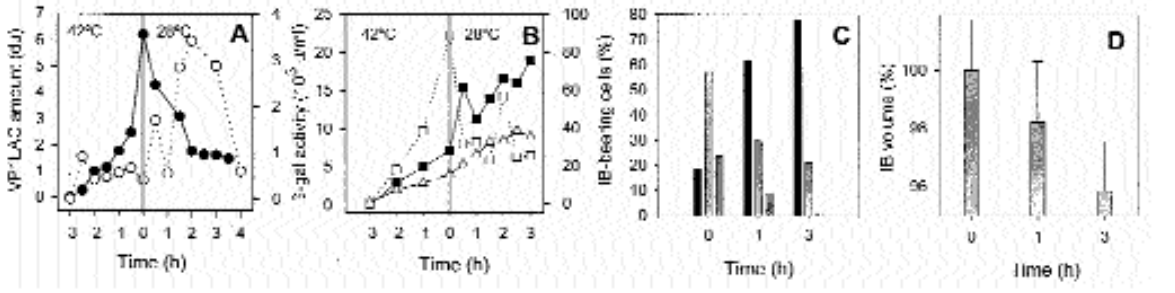

Fig. 1. A: Presence of VP1LAC in the insoluble (black circles, left scale) and soluble (white circles, right scale) cell fractions from BL26 bacterial cultures. The vertical line indicates chloramphenicol addition and temperature downshift. DU are densitometric units. B: β-Galactosidase activity in bacterial cell cultures producing either VP1LAC (black squares, left scale) or CO46 (white squares, right scale). CO46 is a pseudo-wild type β galactosidase not forming IBs. The activity in a VP1LAC producing, IbpAB⁻ strain is also indicated (white triangles, left scale). U are Miller β-galactosidase units [14]. C: Percentage of cells carrying either none (black bars), 1 (light grey bars), or 2 IBs (dark grey bars), in BL21/pJVP1LAC from the same culture than in A and B. D: Variation in the average volume of remaining IBs. The induction time is indicated in all the panels.

extracts and tailless P22 particles for assembling [10]. When required, cells were disrupted by sonication as described [15]. For in vitro analysis, purified TSP and VP1LAC IBs (10× concentrated) from 3 h induced cultures, were incubated at about 1.3 μg of protein per ml at 37°C with 20× concentrated, sonic cell extracts from plasmid-free BL26 cells at late exponential growth phase. These cells, were also grown in LB medium at 37°C as indicated above and shifted to 42°C at the middle exponential phase. Protease inhibitors PMSF (to 1 mM) and benzamidine (to 2.5 μg/ml) were also added. As a control, IBs were incubated under the same conditions with the cell-free buffers used for further activity analysis, using Z buffer for VP1LAC IBs [14] and TMBS for TSP IBs [10]. Samples were taken every 30 min for activity analysis as described above, and soluble and insoluble protein fractionation. Proteins in the obtained fractions were analysed by Western blot. In the figures, representative experiments are shown from sets of four replicas.

3. Results

3.1. In vivo refolding of IB protein

In a previous work, we have shown a dynamic transition between soluble and insoluble forms of β-galactosidase fusion proteins in IB-forming cells, that might involve a significant fraction of IB polypeptides and in which proteolysis is connected [16]. Since it had not been still solved whether proteolysis could be a mechanism of IB protein solubilisation or on the contrary, it occurs over already solubilised polypeptides, we have studied β-galactosidase enzymatic activity and IB volumetric evolution after arresting the synthesis of the β galactosidase fusion protein VP1LAC. Under these conditions, both soluble protein (Fig. 1A) and β galactosidase activity (Fig. 1B) increase, concomitantly with a reduction in the insoluble VP1LAC fraction (Fig. 1A). Note that in a recombinant β-galactosidase, that does not aggregate as IBs (protein CO46), enzymatic activity is rapidly lost after the arrest of protein synthesis (Fig. 1B) by degradation of the soluble CO46. Moreover, in BL26/pJVP1LAC cells, a simultaneous decrease of IB number and average volume is also observed (Figs. 1C, D, 2, top). Altogether, these results strongly suggest that IB embedded protein is solubilised in vivo resulting in disintegration of IB particles trough their progressive volumetric reduction. In addition, at least a fraction of IB protein can reach the native conformation with fully biological activity, proving that proteolysis occurs on polypeptides previously released by a proteolytic-independent process. The fact that in an Ibp⁻ strain, in which IBs are also formed (not shown), β galactosidase activity is recovered as in the wild-type (Fig. 1B), suggests that IbpA,B proteins are not critical in this refolding process.

In an additional approach to evaluate in vivo IB evolution in a different expression system we monitored IB evolution in cells producing P22 TSP protein IBs. TSP is a homotrimeric protein that undergoes a complex folding pathway from which unfolded intermediates collapse as IBs under overproduction conditions [17,18]. In addition, TSP has been extensively used as a model for the analysis of molecular interactions between aggregated polypeptides [19] and of protein folding and unfolding pathways [20–24]. The arrest of recombinant protein synthesis in IB carrying cells results, as in the case of β-galactosidase, in IB disintegration (Fig. 2, bottom; Fig. 3D). Loss of TSP in the insoluble cell fraction can be also monitored, as in the case of VP1LAC, by a numeric IB reduction and a concomitant rise of TSP with full biological activity (Fig. 3B). This indicates that the aggregated protein can be released from IBs in absence of protein synthesis and transferred to the soluble cell fraction as TSP monomers, that, also in absence of protein synthesis, can refold as fully active TSP trimers. Note that a rapid increase in the trimeric native form (Fig. 3B) is accounted by an initial lost of the monomeric form from the soluble cell fraction (Fig. 3A) and that a rate transition in the folding pathway between 1 and 2 h is coincident with the appearing of monomeric forms from IB particles, that results again in a further folding rate transition at about 4 h. Some folding intermediates and degradation fragments have also been observed in blots (not shown).

Although it cannot be excluded that IBs formed by other proteins could undergo alternative or less efficient refolding processes, the fact that two structurally different proteins are released with a comparable efficiency from IBs reveal an unexpected context for protein aggregation in bacterial cells, in which IB formation and growth must be regarded as the result of an unbalanced equilibrium between protein synthesis and aggregation. During production of misfolding prone proteins, bottlenecks in the protein folding assistant network would conduct to deposition of unfolded and misfolded polypeptides, that would be progressively refolded upon chaperone availability, a situation that is strongly favoured when the de novo protein production is arrested.

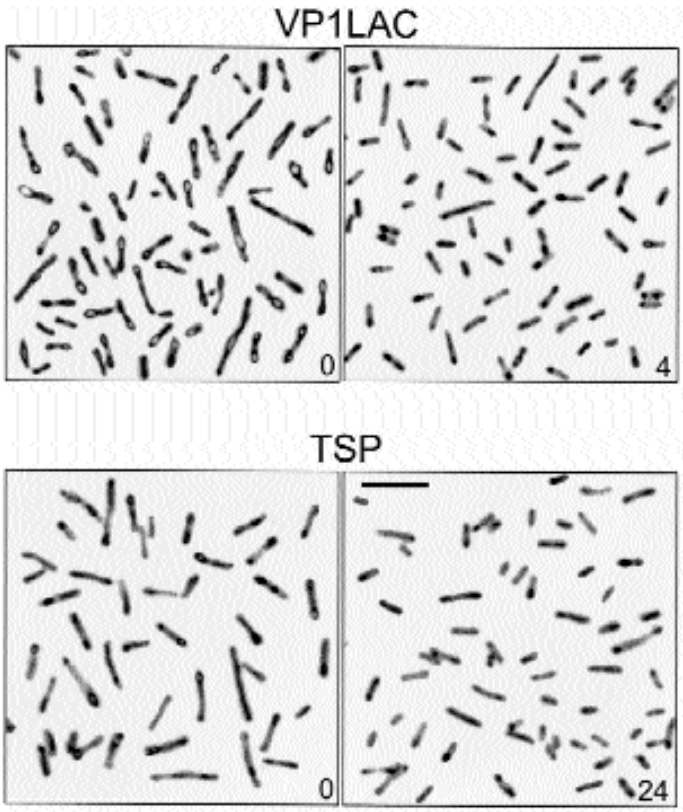

Fig. 2. Micrographies of VP1LAC (top) and TSP (bottom) BL26 producing cells. Numbers indicate time in h after stopping the synthesis of recombinant proteins. The black bar represents 5 μm.

3.2. In vitro refolding of IB protein

Some components of the bacterial folding assistant network have been identified, although their sequential implication in the cascade folding process is still under exploration [25–27]. Irrespective of the specific factors acting on bacterial cells on IB protein refolding, the possibility to induce in vitro protein solubilisation from protein aggregates could be of interest in a variety of purposes. In Fig. 4A, we show a rising of a major VP1LAC degradation product (fragment B) [9] and a concomitant increase of β-galactosidase enzymatic activity by incubation of purified IBs in sonic extracts of *E. coli* cells. A simultaneous rising of intact VP1LAC is also observed by high sensitive developing of blots (not shown), but the amount of this species is too low for an accurate determination. β-Galactosidase activity is not detected when IBs are incubated in saline buffer, although in some of these samples, low protein

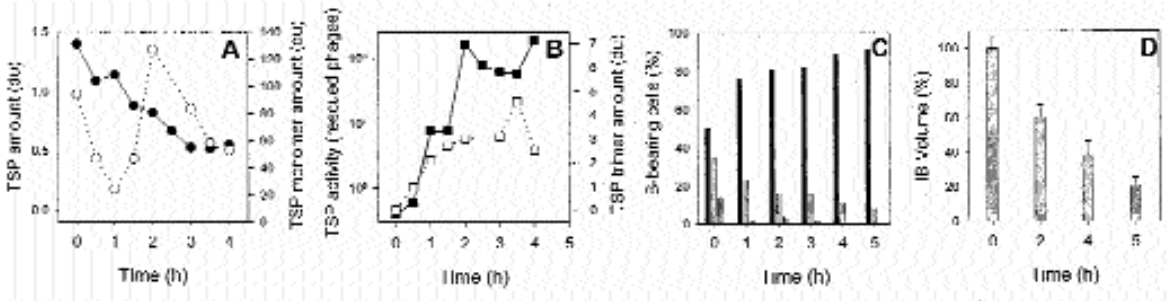

Fig. 3. A: Monitoring of TSP monomer amounts after chloramphenicol addition (time 0) in insoluble (black circles, left scale) and soluble (white circles, right scale) cell fractions. B: TSP activity monitored by phage rescue in the total cell fraction (black squares, left scale) and TSP found in active, trimeric forms in the soluble cell fraction (white squares, right scale). C: Percentage of cells carrying either none (black bars), 1 (light grey bars), or 2 IBs (dark grey bars), in BL26/pTTSPA from the same culture as in A and B. Note that in A, the detection of monomers in the insoluble cell fraction does not necessarily imply that TSP is actually in the monomeric form as embedded into IBs, since for electrophoretic analysis of insoluble protein samples must be denatured before loading. D: Variation in the average volume of remaining IBs. The induction time is indicated in all the panels.

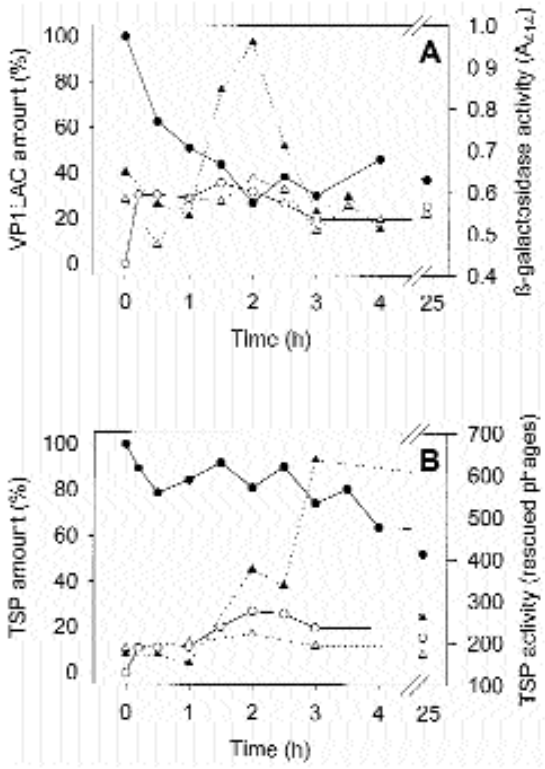

Fig. 4. A: β-Galactosidase activity (right scale) from VP1LAC IBs incubated with crude cell extracts (black triangles) or with saline buffer (white triangles). Protein amounts (left scale) in the insoluble (black circles) and soluble (white circles) cell fractions are also indicated. Whereas the insoluble protein corresponds to the intact VP1LAC, the soluble fraction is mainly composed by a major degradation fragment of 90 kDa [9]. B: TSP activity as monitored by infectious phages rescued from P22 tailless particles (right scale), in TSP IBs incubated with crude cell extracts (black triangles) or with saline buffer (white triangles). Protein amounts (left scale) in the insoluble (black circles) and soluble (white circles) cell fractions are also indicated. Intact TSP monomers are the main component of the soluble cell fraction.

amounts were eventually observed in the soluble cell fraction of some samples (not shown). The intact form of the soluble enzyme in cell extract-treated IB samples is hardly detectable in blots, and since the main degradation fragment is not present in pure IBs [9], this indicates a rapid proteolysis after the release of polypeptides from aggregates, in agreement with the late fading of the enzymatic activity. On the other hand, about 60% of IB protein is lost in about 2 h. In Fig. 4B, an analogous increase of TSP phage rescuing activity is observed in TSP IBs cell extract mixtures, simultaneously to a lost up to 40% of IB protein. Again, the late decrease of TSP activity and concentration of soluble TSP indicates further protein degradation.

4. Discussion

Since early studies on bacterial physiology, IBs were believed to be compact protein aggregates of unfolded polypeptide chains, that being unreachable by proteases and chaperones, remain inert once produced in the cell [1,28]. However, it has been observed enzymatic activity associated to some enzyme-based IBs [29,30] and also the presence of native-like secondary structure in IB protein [31]. In addition, the finding of different conformational states within IB particles [32] has been very recently confirmed by combining scanning electron microscopy and kinetic modelling of in vitro IB tryptic digestion [13]. These observations indicate that as in the case of amyloid fibrils, IBs might be formed by a range of folding intermediates rather that by completely unfolded chains.

Even more intriguing is the observation of a dynamic protein transition between soluble and insoluble cell fractions during volumetric IB growth in vivo [16], that appeared to be linked to the proteolytic digestion. By exploring this event in both a β-galactosidase fusion (VP1LAC) and P22 TSP, we prove here that IB growth is the result of an unbalanced equilibrium between protein precipitation and cell-mediated refolding and solubilisation, that can be spontaneously redirected towards body disintegration by arresting the synthesis of the recombinant polypeptide. Under this situation and when IBs had been already formed, their number and volume rapidly decrease in living cells (Figs 1, 2 and 3), concomitant with a transition of the target protein from the insoluble to the soluble cell fraction and a protein refolding process to gain full functionality. Note that TSP folding is a specially long and complex molecular process [17] to reach deep β-sheet interdigitation between subunits [33]. In the view of these results, it seems reasonable to speculate that components of the bacterial folding assistant network are able to act over aggregates of misfolded protein prompting protein release and correct protein folding from IBs. This solubilisation process is eclipsed when protein synthesis is directed at high rates, but becomes evident when de novo synthesis is arrested. However, note that in both in vivo and in vitro situations, a small protein amount remains insoluble after prolonged experiments, a fact that could be in agreement with the observation of a heterogeneity in the molecular organisation of IB polypeptides as recently described [13]. It is worthy to note, that whereas TSP IBs are homogeneously affected by protein solubilisation (Fig. 3D), recalcitrant species of aggregated VP1LAC seem to be heterogeneously distributed among IBs in the bacterial culture, since a small fraction of these particles displays essentially the same volume in absence of protein synthesis (Fig. 1D). This observation could indicate that different polypeptides can be packaged in different organisation patterns within IB particles.

Independent approaches have proved that chaperones can prevent protein aggregation in different cellular contexts and also favour in vivo solubilisation from aggregates from different origin [34–40]. However, as far as we know, bacterial inclusion bodies have not been explored in this context and its transient nature never reported, despite being extremely common and inconvenient in recombinant bacteria. The data presented here indicate that IB formation is not a dead end process in the protein quality network [41] of bacterial cells, but a transient situation in which IBs act as a reservoir of misfolded polypeptide chains, over which cell components can still apply its folding assistant potential. The fact that IB solubilisation can also be done in vitro by using cell extracts (Fig. 4), offers intriguing possibilities to solve or diminish one of the main bottlenecks in the use of bacteria as a cell factory for foreign proteins, namely IB for-

mation. More importantly, the fact that protein aggregation can be reverted even from poorly structured elements such as IBs prompts to deeply scrutinise possibilities for therapeutic approaches to amyloid diseases based on the reversion of the aggregation process. This possibility is specially appealing since the recent report that chaperones can act as neurodegenerative repressors [35,36] and modulators of fibril formation [37].

Acknowledgements: We are indebted to J. Closa, V. Serrano for technical assistance and Dr. F. Baneyx for generously providing strain JGT17. This work has been supported by grant BIO98-0527 (CICYT) and by Mª Francesca de Rovira lta Foundation. M.M.C. is a recipient of a predoctoral fellowship from UAB, Barcelona, Spain.

References

[1] Marston, F.A.O. (1986) Biochem. J. 240, 1–12.
[2] Strandberg, L. and Enfors, S.O. (1991) Appl. Environ. Microbiol. 57, 1669–1674.
[3] Georgiou, G. and Valax, P. (1996) Curr. Opin. Biotechnol. 7, 190–197.
[4] Clark, E.D.B. (1998) Curr. Opin. Biotechnol. 9, 157–163.
[5] Misawa, S. and Kumagai, I. (1999) Biopolymers 51, 297–307.
[6] Guise, A.D., West, S.M. and Chaudhuri, J.B. (1996) Mol. Biotechnol. 6, 53–64.
[7] Shi, P.Y., Maizels, N. and Weiner, A.M. (1997) Biotechniques 23, 1036–1038.
[8] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[9] Corchero, J.L., Viaplana, E., Benito, A. and Villaverde, A. (1996) J. Biotechnol. 48, 191–200.
[10] Carbonell, X. and Villaverde, A. (1998) Biochem. Biophys. Res. Commun. 244, 428–433.
[11] Thomas, J.G. and Baneyx, F. (1998) J. Bacteriol. 180, 5165–5172.
[12] Carrió, M.M., Corchero, J.L. and Villaverde, A. (1998) FEMS Microbiol. Lett. 169, 9–15.
[13] Carrió, M.M., Cubarsi, R. and Villaverde, A. (2000) FEBS Lett. 471, 7–11.
[14] Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[15] Telin, J.X., Cubarsi, R. and Villaverde, A. (1998) Biotechnol. Bioeng. 58, 536–540.
[16] Carrió, M.M., Corchero, J.L. and Villaverde, A. (1999) Biochim. Biophys. Acta 1434, 170–176.
[17] King, J., Haase-Pettingell, C., Robinson, A.S., Speed, M. and Mitraki, A. (1996) FASEB J. 10, 57–66.
[18] Haase-Pettingell, C.A. and King, J. (1988) J. Biol. Chem. 263, 4977–4983.
[19] Speed, M.A., Wang, D.I. and King, J. (1996) Nat. Biotechnol. 14, 1283–1287.
[20] Schuler, B., Rachel, R. and Seckler, R. (1999) J. Biol. Chem. 274, 18589–18596.
[21] Speed, M.A., Morshead, T., Wang, D.I. and King, J. (1997) Protein Sci. 6, 99–108.
[22] Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J. and King, J. (1991) Science 253, 54–58.
[23] Speed, M.A., Wang, D.T. and King, J. (1995) Protein Sci. 4, 900–908.
[24] Carbonell, X. and Villaverde, A. (1998) FEBS Lett. 432, 228–230.
[25] Buchberger, A., Schroder, H., Hesterkamp, T., Schonfeld, H.J. and Bukau, B. (1996) J. Mol. Biol. 261, 328–333.
[26] Rudiger, K. (1998) Curr. Opin. Struct. Biol. 8, 159–165.
[27] Veinger, L., Diamant, S., Buchner, J. and Goloubinoff, P. (1998) J. Biol. Chem. 273, 11032–11037.
[28] Hellebust, H., Murby, M., Abrahmsen, L., Uhlén, M. and Enfors, S.O. (1989) Biotechnology 7, 165–168.
[29] Wormill, D.M. and Goss, N.H. (1989) Aust. J. Biotechnol. 3, 28–32.
[30] Tokatlidis, K., Dhurjati, P., Millet, J., Béguin, P. and Albert, J.P. (1991) FEBS Lett. 282, 205–208.
[31] Oberg, K., Chrunk, B.A., Wetzel, R. and Fink, A.L. (1994) Biochemistry 33, 2628–2634.
[32] Bowden, G.A., Paredes, A.M. and Georgiou, G. (1991) Biotechnology 9, 725–730.
[33] Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. and Reinemer, P. (1994) Science 265, 383–386.
[34] Krobitsch, S. and Lindquist, S. (2000) Proc. Natl. Acad. Sci. USA 97, 1589–1594.
[35] Warrick, J., Chan, H.Y.E., Gray-Board, G.L., Chai, Y.H., Paulson, H.L. and Bonini, N.M. (1999) Nat. Genet. 19, 148–154.
[36] Kazemi-Esfarjani, P. and Benzer, S. (2000) Science 287, 1837–1840.
[37] Muchowski, P.J., Schaffar, G., Sittler, A., Wanker, E.E., Hayer-Hartl, M.K. and Hartl, F.U. (2000) Proc. Natl. Acad. Sci. USA 97, 7841–7846.
[38] Charmichel, J., Chatellier, J., Woolfson, A., Milstein, C., Fersht, A.R. and Rubinsztein, D.C. (2000) Proc. Natl. Acad. Sci. USA 97, 9701–9705.
[39] Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H. and Bukau, B. (1999) EMBO J. 18, 6934–6949.
[40] Diamant, S., Ben-Zvi, A.P., Bukau, B. and Goloubinoff, P. (2000) J. Biol. Chem. 275, 21107–21113.
[41] Wickner, S., Maurizi, M.R. and Gottesman, S. (1999) Science 286, 1888–1893.

DESINTEGRACIÓ DELS COSSOS D'INCLUSIÓ QUAN S'ATURA LA SÍNTESI PROTEICA

Com acabem de veure, l'agregació proteica en els cossos d'inclusió de VP1LAC i TSPA és reversible. Quan s'atura la síntesi de proteïna recombinant, es produeix una ràpida desintegració dels cossos d'inclusió. Aquí es mostra més detalladament com es perd la proteïna agregada dels cossos d'inclusió de VP1LAC, LACVP1 i TSPA. Els cossos de VP1LAC i LACVP1 es desintegren més ràpidament que els de TSPA, perden un 80% del seu contingut en 2 hores i el 20 % restant de proteïna roman agregada 24 hores més tard. Encanvi, els cossos d'inclusió de TSPA es desintegren més lentament, però completament. La fracció de proteïna recalcitrant dels cossos de VP1LAC i LACVP1 que no és desagregada, podria representar una part dels polipèptids agregats entre els quals s'han creat interaccions més fortes, com podrien ésser ponts disofre.

Comparant la desintegració dels cossos d'inclusió de VP1LAC en les soques BL21 (Lon-) i MC4100, veiem que tenen un comportament similar, la qual cosa ens indica que el sistema encarregat de dissociar dels agregats deu ser independent de la proteasa Lon. Tanmateix, la lleugera diferència entre les dues soques, d'un 6%, podria estar associada a l'absència de la proteasea Lon en BL21, o a la variabilitat dels experiments. Per altra banda, el fet de trobar la mateixa fracció de proteïna més resistent a ésser dissociada en diferents soques, suggereix que depèn de la proteïna recombinant.

La ràpida desintegració dels cossos d'inclusió a l'aturar la incorporació de proteïna als agregats suggereix que hi ha un mecanisme cel.lular dirigit a processar la proteïna agregada. Una gran part de la proteïna desagregada és proteolitzada, i una altra replegada (Fig. 4, article 3) Així doncs, en aquest cas, l'agregació proteica podria representar un estat transitori de polipèptids mal plegats que seran proteolitzats o replegats.

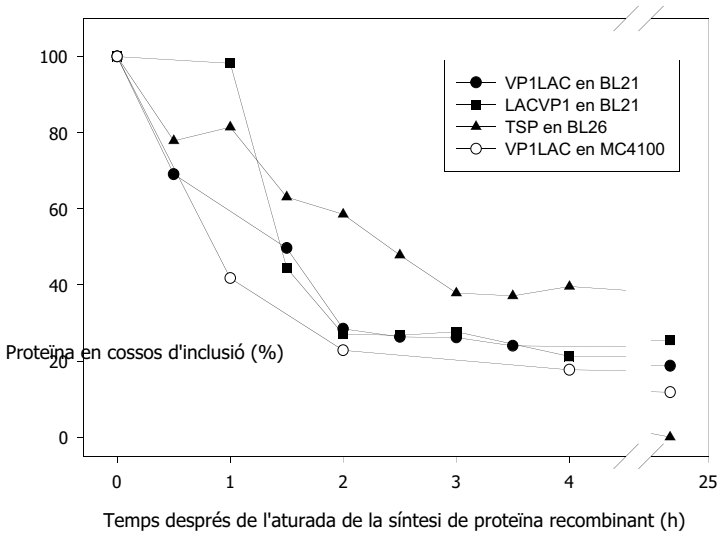

Figura 1. Desintegració dels cossos d'inclusió de VP1LAC, LACVP1 i TSP de 3 hores produïts en les soques BL26, BL21 i MC4100. Quan s'atura la síntesi de proteïna recombinant, s'observa com va disminuïnt la quantitat de proteïna agregada que forma els cossos d'inclusió.

4.4. Organització interna dels cossos d'inclusió bacterians

Es van realitzar una sèrie de treballs per estudiar com s'estructura la proteïna agregada dins dels cossos d'inclusió. Per una banda es va analitzar la sensibilitat *in vitro* de les proteïnes dels cossos d'inclusió de VP1LAC i LACVP1 a la tripsina. Combinant la visualització del procés de fragmentació dels cossos d'inclusió per microscòpia electrònica i el modelatge matemàtic del procés de digestió dels agregats, ens va donar una idea de la seva organització interna. Per altra banda, es van estudiar les propietats estructurals de la proteïna agregada en els cossos d'inclusió de VP1LAC i LACVP1, mitjancant tècniques d'espectroscòpia, com fluorimetria i dicroïsme circular.

FEBS 23495 FEBS Letters 471 (2000) 7–11

Fine architecture of bacterial inclusion bodies

M. Mar Carrió[a], Rafael Cubarsi[b], Antonio Villaverde[a,*]

[a] *Institut de Biologia Fonamental and Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain*
[b] *Departament de Matemàtica Aplicada i Telemàtica, Universitat Politècnica de Catalunya, Barcelona, Spain*

Received 14 February 2000

Edited by Pierre Jolles

Abstract The molecular organisation of protein aggregates, formed under physiological conditions, has been explored by in vitro trypsin treatment and electron microscopy analysis of bacterially produced inclusion bodies (IBs). The kinetic modelling of protein digestion has revealed variable proteolysis rates during protease exposure that are not compatible with a surface-restricted erosion of body particles but with a hyper-surfaced disintegration by selective enzymatic attack. In addition, differently resistant species of the IB proteins coexist within the particles, with half-lives that differ among them up to 50-fold. During in vivo protein incorporation throughout IB growth, a progressive increase of proteolytic resistance in all these species is observed, indicative of folding transitions and dynamic reorganisations of the body structure. Both the heterogeneity of the folding state and the time-dependent folding transitions undergone by the aggregated polypeptides indicate that IBs are not mere deposits of collapsed, inert molecules but plastic reservoirs of misfolded proteins that would allow, at least up to a certain extent, their in vivo recovery and transference to the soluble cell fraction.

Key words: β-Galactosidase; Protein folding; Aggregation; Proteolysis; Recombinant protein

1. Introduction

Chaperones and proteases are major components of the cellular protein quality control system devoted to prevent the occurrence of misfolded polypeptides. In mammalian cells, the proteolytic escape of incorrectly folded proteins promotes their irreversible accumulation as insoluble amyloid plaques and the consequent development of severe degenerative diseases of growing impact [1]. Different protein species are involved in such diseases. However, a common, defined fibril structure is observed upon aggregation [2] irrespective of the primary sequence and the native folding pattern [2–5].

On the other hand, inclusion bodies (IBs) are refractile aggregates of misfolded, insoluble protein commonly observed upon targeted gene overexpression in bacterial cells [6]. IBs are not homogeneous in protein composition. Many cell folding-assistant proteins and proteolytic fragments of the main polypeptide component are found embedded as variable fractions of the total protein content [7–9]. Despite this heterogeneity, transmission electron microscopy has often pictured IBs as amorphous aggregates lacking defined structural organisation [10]. Here, by combining scanning electron microscopy and a kinetic modelling of IB protein digestion during trypsin treatment, we have observed unexpected architectural features of IBs and the coexistence of distinct populations of aggregated protein with different conformational states. These intriguing results prove a particular organisation of IB proteins that, at difference of the mammalian counterpart aggregates, can explain the in vivo dynamic transition between aggregated and soluble polypeptides recently observed in IB-forming bacteria [11].

2. Materials and methods

2.1. Bacterial strain, plasmids and culture conditions

The *Escherichia coli* strain BL21 [12] was used as IB factory, since the absence of protease La stimulates IB formation [13]. Plasmids pILACVP1 and pIVP1LAC [13,14] encode two closely related hybrid proteins in which the foot and mouth disease virus VP1 capsid protein is fused to either the carboxy- or amino-termini of *E. coli* β-galactosidase, respectively. In these vectors, transcription of the encoding genes is directed by the lambda p_R and p_L lytic promoters and controlled by the temperature-sensitive cI857 repressor. Luria–Bertani medium [12] plus 100 μg/ml ampicillin was used for the culture of recombinant bacteria. Cultures were first grown at 28°C and 250 rpm. At the beginning of the exponential phase (around 0.4 OD_{550} units), flasks were transferred to a pre warmed bath at 42°C and further incubated at 250 rpm at this temperature up to 24 h.

2.2. IB sampling and purification

Culture samples of 1.5 ml were periodically taken during the production phase, at 1, 3, 5 and 24 h after temperature shift, and IBs were purified by repeated detergent treatment as follows. Cells were harvested by centrifugation at 12000×*g* (at 4°C) for 15 min and resuspended in 200 μl of lysis buffer (50 mM Tris Cl pH 8, 1 mM EDTA, 100 mM NaCl), plus 30 μl of 100 mM protease inhibitor PMSF and 6 μl of 10 mg/ml lysozyme. After 30 min of incubation at 37°C under gentle agitation, NP-40 was added at 1% (v/v) and the mixture incubated at 4°C for 30 min. Then, 3 μl of DNase I (from a 1 mg/ml stock) and 3 μl of 1 M $MgSO_4$ were added and the resulting mixture was further incubated at 37°C for 30 min. Protein aggregates were separated by centrifugation at 12000×*g* for 15 min at 4°C. Finally, IBs were washed once with the same buffer containing 0.5% Triton X-100 and once with sterile PBS. After a final centrifugation at 12000×*g* for 15 min, pellets were stored at −80°C until analysis.

2.3. Trypsin treatment and electrophoretic analysis of IBs

Purified IBs containing around 15 ng/μl of recombinant protein were resuspended in 300 μl of Z buffer (0.06 M Na_2HPO_4, 0.04 M NaH_2PO_4-H_2O, 0.01 M KCl, $MgSO_4$-$7H_2O$, pH 7.0) and incubated with trypsin (to 100 μg/ml) at 37°C. During 24 h incubation, samples were periodically withdrawn and soybean trypsin inhibitor (to 200 μg/ml) was added to stop the digestion. Afterwards, samples were stored at −20°C until Western blot analysis. Simultaneously, negative controls were performed with the same samples without adding trypsin

*Corresponding author. Fax: (34)-93-5812011.
E-mail: antoni.villaverde@uab.es

PII: S0014-5793(00)01357-0

Trypsin-treated IBs were boiled for 10 min in denaturing buffer (125 mM Tris Cl pH 6.8, 6% SDS, 20% glycerol, 10% β-mercaptoethanol) plus 8 M urea and loaded on 3.5% stacking and 7% polyacrylamide gel electrophoresis separation gels. Protein bands were detected by immunoblot, using anti β galactosidase polyclonal serum. Protein amounts were estimated by measuring band areas with one-dimension Intelligent Quantifier® software after high resolution scanning.

2.4. Electronic microscopy

Purified IBs were fixed with 2.5% glutaraldehyde in phosphate buffer (Na_2HPO_4 0.9 M, NaH_2PO_4 0.06 M, pH 8.0) for 1 h at 4°C. To remove the cross-linker agent, samples were retained on a nucleopore membrane and washed twice with the same buffer. Afterwards, they were dehydrated by consecutive, 5 min washing steps with increasing concentrations of ethanol in water (30, 50, 70, 90 and 100%). Ethanol was finally evaporated by reaching the critical point. Finally, membranes were covered with a 100 nm layer gold for a further observation under a scanning microscope Hitachi S-570.

2.5. Modelling of IB protein digestion

Proteolysis of IB intact recombinant proteins was modelled as follows. Being $N(t)$ the non-fragmented protein (in absolute densitometric units) at time t, digestion of IB protein occurs experimentally according to the following non linear differential equation:

$$N'(t)/N(t) = -qN(t)^{p-1};\ p>1,\ q>0 \quad (1)$$

For heterogeneous IBs containing more than one protein species, namely n, composition is described as follows:

$$N(t) = \sum_{i=1}^{n} N_i e^{-q_i t} \quad (2)$$

The specific parameters for each protein species are q_i, while the fractions $N_i/N(0)$ represent the mixing proportions. The function described in Eq. 2 satisfies an n-order linear homogeneous differential equation with constant coefficients, such as

$$f(N, N', \ldots, N^{(n)}; q_1, q_2, \ldots, q_n) = 0 \quad (3)$$

From numerical evaluations of $N, N', \ldots, N^{(n)}$ in intermediate degradation times, a linear over-determinate system of equations was built in order to estimate the parameters q_i in Eq. 3. Details of this procedure will be described elsewhere (Cubarsi et al., in preparation). The numerical estimation of the derivatives directly from experimental data led to non-consistent results. For this reason, the derivatives in Eq. 3 were explicitly evaluated from Eq. 1. This clearly produced a smoother and better determination of the derivatives. Thus, Eq. 3 was converted into the following form

$$g(N, p, q; q_1, \ldots, q_n) = 0 \quad (4)$$

Finally, the mixing proportions were obtained by solving another linear over-determinate system of equations, associated with Eq. 2, for the complete set of experimental data.

3. Results

3.1. Proteolytic stability of IB proteins

Isolated IBs, built up by either VP1LAC or LACVP1 β galactosidase fusion proteins, were submitted to in vitro proteolytic digestion by incubation with trypsin. For this analysis, we used bodies obtained from cells in which recombinant gene expression had been induced for different periods forming differently aged IBs, spanning from 1 to 24 h. During trypsin treatment, loss of the intact fusion species was observed in all the cases (see one example in Fig. 1). This was concomitant with a transient increase of anti-β-galactosidase immunoreactive degradation fragments, derived from a proteolytic cascade, that is comparable to that observed in vivo [15]. By plotting an intact protein amount versus time, a non-constant degradation velocity was revealed, being higher during the first minutes but rapidly declining with time (Fig. 2). Previous proteolysis assays performed under the same conditions [16] did not suggest any enzyme limitation responsible for variation in the proteolysis rate. In addition, proteolysis of the degradation intermediates was completed after 6 h (Fig. 1) under nearly constant rates (not shown), indicative of a peculiar protease susceptibility of the full-length forms of the proteins.

Half-lives of these intact length forms were roughly approached as $\ln 2/b$ by adjusting a simple exponential law

$$N(t) = N(0)e^{-bt} \quad (5)$$

and the obtained values are depicted in Fig. 2. For both proteins, protease resistance increased with the age of the containing bodies. Interestingly, while this observed variation is rather constant for LACVP1, it is rapidly stabilised in VP1LAC. Surface digestion has reasonably been suggested for IB protein proteolysis under in vitro protease exposure [10], and consequently, IB topology could eventually influence digestion rate. However, for a surface-restricted IB erosion, the dramatic difference in IB volume found when comparing LACVP1 and VP1LAC IBs [17] would generate greater discrepancies in the observed half-lives. In addition, the volumetric growth of VP1LAC IBs continues up to 24 h, as it also occurs for LACVP1 IBs [11], proving that IB volume is not a key factor in determining the degradation kinetics.

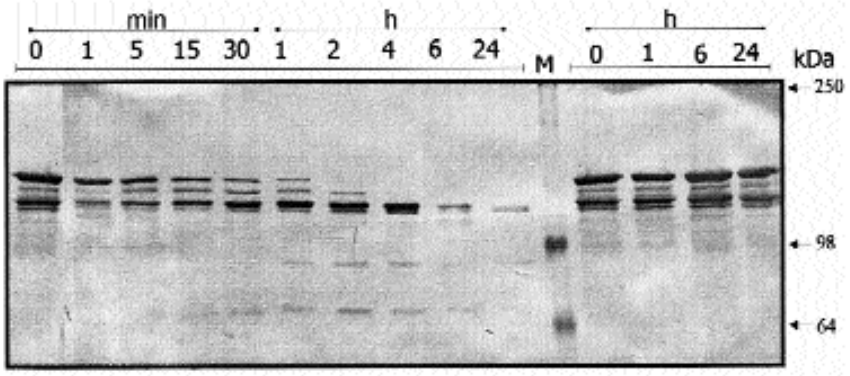

Fig. 1. Western blot of trypsin-treated LACVP1 IBs (aged 5 h) at different incubation times (expressed in min and h). The upper band corresponds to the full-length LACVP1. At the right side, IBs incubated under the same conditions without protease. Lane M indicates molecular weight markers whose apparent masses are indicated by arrows.

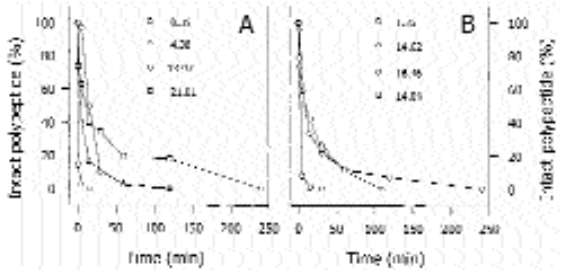

Fig. 2. Evolution of intact IB fusion proteins (LACVP1 in A, VP1LAC in B) during trypsin exposure. Differently aged bodies were processed and are indicated as follows: 1 h, circles; 3 h, triangles; 5 h, diamonds; 24 h, squares. Half-lives estimated from Eq. 1 (in min) are also indicated for each experiment.

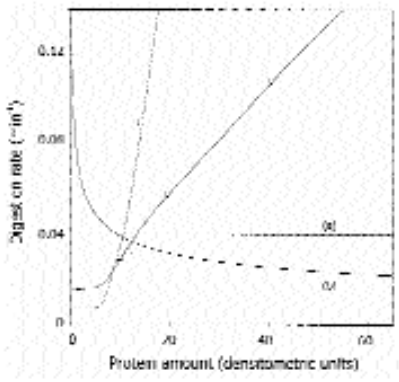

Fig. 3. Digestion rates of IB protein versus protein amount obtained from a three point estimation of 5 h-aged bodies for VP1LAC (circles) and LACVP1 (squares). In addition, expected rates are plotted according to different models of IB digestion, namely constant rate (a, dashed line), surface-restricted erosion (b, continuous line) and mixture of protein species with constant but distinguishable digestion rates, as modelled from the experimental data for both VP1LAC (continuous plot) and LACVP1 (dashed plot). Note that data at zero protein value indicate the digestion rate for the most protease-resistant protein form found in each type of IB, and it is the inverse of the half lives of fraction 3 as indicated in Table 1. This resistant form is the unique remaining in extensively digested IB particles.

3.2. Heterogeneity of IB protein stability

The results presented above suggest that factors other than IB topology, like intrinsic protein stability or molecular organisation within IBs, would be main determinants of IB digestion dynamics. Therefore, to investigate in more detail the mechanics of protein digestion and the putative influence of body volume, digestion experiments were repeated for 5 h-aged VP1LAC IBs and 3, 5 and 24 h-aged LACVP1 bodies. Additional sampling of LACVP1 IBs was of particular interest because of the time-dependent variation of half-life observed in LACVP1-built bodies as discussed above (Fig. 2). With this set of duplicated data, evolution of intact protein was modelled according to Eq. 1, and the relevant parameter p was obtained for each individual experiment (Table 1). In principle, a Poisson stochastic process [18] would be expected to account for the digestion kinetics. Then, the probability density function for the time interval from trypsin addition up to the first molecular proteolytic event would be given by Eq. 1, rendering $p = 1$.

However, such behaviour would imply a constant digestion rate (Fig. 3, plot a), which is obviously inconsistent with the experimental data (Figs. 2 and 3). An alternative possibility of a concentric layer erosion (a surface-restricted proteolytic attack of IB protein) also predicts a digestion profile that does not fit with the experimentally obtained kinetics (Fig. 3, plot b). Also, note that in this case, p would be 2/3 and that the approximated p values for all the performed experiments are indeed higher than one (Table 1).

Therefore, a more plausible hypothesis can be given by assuming a heterogeneous nature of the IB protein in which more than one different protein species (1, 2, 3 and so on, with increasing proteolytic susceptibility) would coexist. According to this model, the composition of IBs is accurately described by Eq. 2. Under protease exposure, each of these species follows an individual Poisson fragmentation process resulting in distinguishable half-lives, whose combination would render the average half lives indicated in Fig. 2. While a two component system explains qualitatively the degradation kinetics, for most of the experiments, the minimal number of protein species required to actually account for the observed data is three (Table 1).

As predicted, these protein species exhibit distinguishable proteolytic sensitivity and they occur at different molar proportions (Table 1), being data from duplicated experiments highly consistent. Except experiment 1 of 5 h-aged LACVP1 IBs, in which only two protein forms could explain numerically the degradation kinetics, a third fraction of a highly sensitive form (half life between 1 and 6.5 min) is observed

Table 1
Protein composition and stability in differently aged IBs

Protein	IB age (h)	Experiment number	p	Protein species					
				1		2		3	
				Half-life (min)	Content (%)	Half-life (min)	Content (%)	Half-life (min)	Content (%)
VP1LAC	5	1	1.91 ± 0.11	1	1.5	3.5	65.5	26.5	33
VP1LAC	5	2	1.62 ± 0.06	3.5	8	11	68	77	24
LACVP1	3	1	2.46 ± 0.37	1	27	4	27	50.5	46
LACVP1	3	2	1.72 ± 0.22	1	24	4.5	23	22.5	53
LACVP1	5	1	3.74 ± 1.08	nd	nd	2.7	58	64	42
LACVP1	5	2	1.88 ± 0.32	1.6	34	6	15	43	51
LACVP1	24	1	1.81 ± 0.63	4	19	10.5	42	75.5	39
LACVP1	24	2	1.55 ± 0.31	6.5	18	19.5	56	109.5	26

nd, not detected

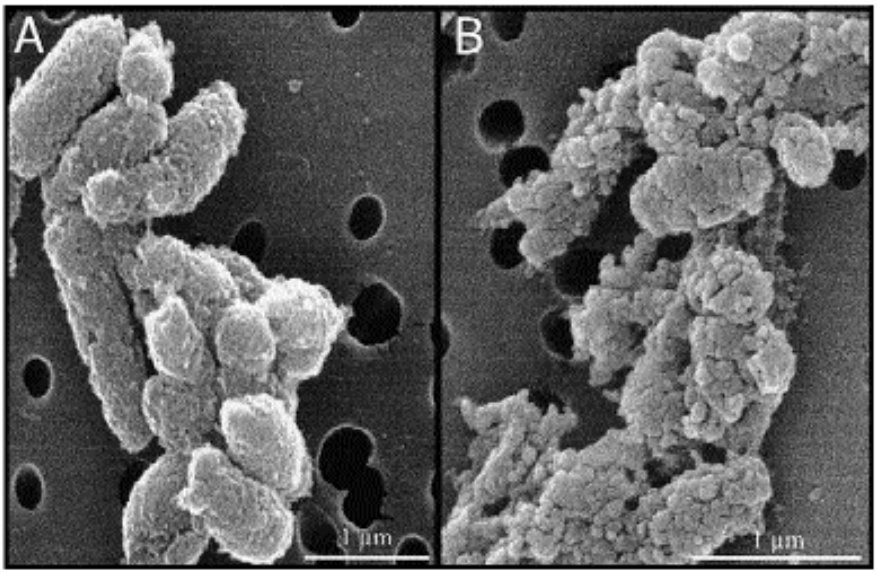

Fig. 4. Scanning electron micrographs of purified VP1LAC (aged 5 h) before (A) and after 5 min of trypsin treatment (B). Similar images were obtained with the smaller LACVP1 IBs (not shown). These pictures are representative of a large number of examinations.

in all cases, abounding more in LACVP1 (from 18 to 34%) than in VP1LAC (less than 8%). The main component of VP1LAC bodies is an intermediate stable protein form (about 68% of total IB protein), while a highly stable species accounts for between 24 and 33% of protein. Interestingly, the half life of the more stable fraction increases dramatically, although its proportion in the IB protein pool remains nearly constant or even declines. This age-dependent enhancement of fraction 3 stability (between 0.5 and 5 fold) could be a main contributor of the progressively increasing trypsin resistance observed concomitantly with the age of LACVP1 IBs (Fig. 2). Lesser but still detectable stabilisation of fractions 1 and 2 could also be relevant.

3.3. Architectural analysis of digested IB

The model presented above explains successfully the dynamics of IB protein degradation by the presence of at least three protein species that are digested at different rates (Table 1). However, it does not provide information about if these fractions co locate in heterogeneous IBs or, alternatively, they segregate into different body particles. To approach experimentally this question, purified IBs were examined under SEM during trypsin treatment. While non-treated bodies exhibited a rood-shaped morphology with a smooth surface (Fig. 4A), protease exposure generates profound fragmentation lines that sub-particulate IBs and released sub-body spherical and rood-shaped pieces (Fig. 4B) that further disappear at longer incubation times (not shown). This observation confirmed that the proteolytic attack is not surface restricted and indicated that protease activity has a selective impact on IB solvent exposed surface. In addition, the obtained images are compatible with a heterogeneous protein composition, regarding proteolytic stability, that is not dependent on surface accessibility to proteases. Interestingly, this fragmentation pattern is seen in the vast majority of the observed particles, suggesting that IBs built of homogeneously resistant polypeptides, if they occur, are rare among IB-producing bacteria.

4. Discussion

Although poorly studied at a molecular level, bacterial IBs are believed to be inert, unorganised aggregates of completely protein that protect the embodied polypeptides from in vivo proteolytic attack [6]. However, a number of independent observations seriously challenge this assumption. Firstly, under certain conditions, enzymatic activity is found associated to enzyme-based IB particles even after stringent washing [19,20]. Since the specific activities ranged from one-third [19] to almost the same values [20] than that of soluble enzyme, the presence of correctly folded polypeptides in IBs, rather than a mere contamination, should be considered. This is in agreement with the finding of native-like secondary structure in IB protein [21] and the suggestion of different protein conformations within bodies [10], that are greatly influenced by chemical and physical parameters during IB building [22].

On the other hand, unexpectedly short in vivo half-lives have been described for both loose aggregated [15] and true IB protein [23], being even lower than those of the soluble counterparts. In this line, protein release and degradation from body-like proteinaceous structures had already been reported in early studies [24,25]. Very recently, we have observed extremely fast protein release from cytoplasmic IBs of β-galactosidase fusion proteins [11] and other recombinant proteins (Carrió et al., manuscript in preparation), during volumetric IB growth in living cells [17]. This finding proves an important dynamic transition between soluble and insoluble polypeptides that would be in agreement with a more flexible model of the IB structure.

In this context, results presented here strongly suggest that more than one single protein species with distinct protease susceptibility coexist (Table 1), thus accounting for a declining

degradation rate of IB protein (Figs. 2 and 3). Since in β-galactosidase, like in β-lactamase [10] and likely in most of the proteins, protease sensitivity depends on protein conformation [26,27], different folding states of IB protein must be present in LACVP1 and VP1LAC IBs. In addition, a time dependent increase of half-lives in all the protein fractions during body building (Table 1) is indicative of structural transitions occurring within LACVP1 IBs during protein accumulation. Since degradation does not occur through surface erosion (Figs. 3 and 4), differential protease sensitivity cannot only be explained by substrate exposure, and determinants more complex than the simple surface topology would be actually modulating the digestion dynamics. The relevant percentage of highly sensitive protein (Table 1) and the appearing of deep fragmentation lines (Fig. 4) suggest an invasive trypsin activity selectively attacking defined protein sectors. Apart from the heterogeneous IB protein composition itself, protease invasiveness could be favoured by the high level of IB voidage previously revealed by density analysis [28]. Accordingly, parameter p from Eq. 1 can be observed as a numerical indicator of the IB erosion topology. For values close to 2/3, it would indicate surface erosion, close to 1, homogeneous protease diffusion within single species bodies, and for p values around 2, what is in fact the actual case (Table 1), a polarised protease resistance generating complex, hyper-surfaced IB structures during body disintegration that permit efficient enzyme penetration. The release of protease-resistant pseudo-spherical particles (Fig. 4) and the mutually exclusive pattern of IB growth in the bacterial cytoplasm [17] could also be in agreement with an IB construction process in which pre-aggregated protein particles with protease resistant cores combine to form one unique higher order body element per cell.

The data presented here, together with other published side observations, show IBs as highly dynamic, flexible and self-organised proteinaceous structures formed by heterogeneously misfolded polypeptides, that support both inner protein folding transitions but also an active protein exchange with the soluble cell fraction.

Acknowledgements: We are indebted to J. Clotet and V. Ferreres for excellent technical assistance and M. Martí for SEM training. This work has been supported by Grant BIO98-0937 (CICYT) and by Mª Francesca de Roviralta Foundation (Spain). M.M.C. has been a recipient of a doctoral fellowship from UAB, Barcelona, Spain.

References

[1] Dobson, C.M. (1999) Trends Biochem. Sci. 24, 329–332.
[2] Serpell, L.C., Sunde, M. and Blake, C.C. (1997) Cell. Mol. Life Sci. 53, 871–887.
[3] Perutz, M.F. (1999) Trends Biochem. Sci. 24, 58–63.
[4] Guijarro, J.I., Sunde, M., Jones, J.A., Campbell, I.D. and Dobson, C.M. (1998) Proc. Natl. Acad. Sci. USA 95, 4224–4228.
[5] Litvinovich, S.V., Brew, S.A., Aota, S., Akiyama, S.K., Haudenschild, C. and Ingham, K.C. (1998) J. Mol. Biol. 280, 245–258.
[6] Marston, F.A.O. (1986) Biochem. J. 240, 1–12.
[7] Rinas, U. and Baily, J.E. (1992) Appl. Microbiol. Biotechnol. 37, 609–614.
[8] Rinas, U. (1996) Biotechnol. Prog. 12, 196–200.
[9] Georgiou, G. and Valax, P. (1999) Methods Enzymol. 309, 48–58.
[10] Bowden, G.A., Paredes, A.M. and Georgiou, G. (1991) Biotechnology 9, 725–730.
[11] Carrió, M.M., Corchero, J.L. and Villaverde, A. (1999) Biophys. Biochim. Acta 1434, 170–176.
[12] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] Corchero, J.L., Viaplana, E., Benito, A. and Villaverde, A. (1996) J. Biotechnol. 48, 191–200.
[14] Corchero, J.L. and Villaverde, A. (1998) Biotechnol. Bioeng. 58, 625–632.
[15] Corchero, E., Cubarsi, R., Enfors, S.-O. and Villaverde, A. (1997) Biophys. Biochem. Res. Commun. 237, 335–340.
[16] Peñin, J.X. and Villaverde, A. (1998) FEBS Lett. 434, 25–27.
[17] Carrió, M.M., Corchero, J.L. and Villaverde, A. (1998) FEMS Microbiol. Lett. 169, 9–15.
[18] Papoulis, A. (1989) Probability, Random Variables, and Stochastic Processes, MacGraww Hill, Singapore.
[19] Worrall, D.M. and Goss, N.H. (1989) Aust. J. Biotechnol. 3, 28–32.
[20] Tokarlidis, K., Dhurjati, P., Millet, J., Béguin, P. and Albert, J.P. (1991) FEBS Lett. 283, 205–208.
[21] Oberg, K., Chrunk, B.A., Wetzel, R. and Fink, A.L. (1994) Biochemistry 33, 2628–2634.
[22] Przybycin, T.M., Duran, J.P., Valax, P. and Georgiou, G. (1994) Protein Eng. 7, 131–136.
[23] Weickert, M.J. and Curry, S.R. (1997) Arch. Biochem. Biophys. 348, 337–346.
[24] Prouty, W.F. and Goldberg, A.L. (1972) Nat. New Biol. 240, 147–150.
[25] Prouty, W.F., Karnovsky, M.J. and Goldberg, A.L. (1975) J. Biol. Chem. 250, 1112–1122.
[26] Viaplana, E., Feliu, J.X., Corchero, J.L. and Villaverde, A. (1997) Biophys. Biochem. Acta 1343, 221–226.
[27] Peñin, J.X. and Villaverde, A. (1998) FEBS Lett. 434, 25–27.
[28] Taylor, G., Hoare, M., Gray, D.R. and Marston, F.A.O. (1986) Biotechnology 4, 553–557.

Biochemical and Biophysical Research Communications **282**, 436–441 (2001)
doi:10.1006/bbrc.2001.4583, available online at http://www.idealibrary.com on IDEAL®

In Situ Proteolytic Digestion of Inclusion Body Polypeptides Occurs as a Cascade Process

Rafael Cubarsí,* M. Mar Carrió,† and Antonio Villaverde†,[1]

*Institut de Biologia Fonamental and Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain; and *Departament de Matemàtica Aplicada i Telemàtica, Universitat Politècnica de Catalunya, Barcelona, Spain*

Received February 12, 2001

Misfolded proteins undergo a preferent degradation ruled by the housekeeping bacterial proteolytic system, but upon precipitation as inclusion bodies their stability dramatically increases. The susceptibility of aggregated polypeptides to proteolytic attack remains essentially unexplored in bacteria and also in eukaryotic cells. We have studied here the *in vitro* proteolysis of β-galactosidase fusion proteins by trypsin treatment of purified inclusion bodies. A cascade digestion process similar to that occurring *in vivo* has been observed in the insoluble fraction of the digestion reaction. This suggests that major protease target sites are not either lost or newly generated by protein precipitation and that the digestion occurs *in situ* probably on solvent-exposed surfaces of inclusion bodies. In addition, the sequence of the proteolytic attack is influenced by protein determinants other than amino acid sequence, the early digestion steps having a dramatic influence on the further cleavage susceptibility of the intermediate degradation fragments. These observations indicate unexpected conformational changes of inclusion body proteins during their site-limited digestion, that could promote protein release from aggregates, thus partially accounting for the plasticity of *in vivo* protein precipitation and solubilization in bacteria. © 2001 Academic Press

Key Words: **β-galactosidase; protein misfolding; inclusion bodies; proteolysis; aggregated protein.**

The housekeeping proteolytic cell systems remove misfolded polypeptides from the cell cytoplasm, probably connected to the incapability of chaperones to properly interact for the assisted folding. The signals targeting such proteins for either chaperone refolding attempts or proteolytic attack remain unknown (1). However, misfolded proteins can also aggregate and consequently accumulate in the cell, resulting from a kinetic unbalance between digestion and precipitation pathways. In turn, protein aggregation might reduce the availability of the insoluble protein to proteases thus enhancing its stability (2). While in bacteria, this fact often results in the generation of refractile proteinaceous particles named inclusion bodies (IBs) (3), in mammalian cells, protein precipitation can derive into a set of related degenerative diseases of growing concern (4) that are presumably linked to the slow or null proteolytic turnover of misfolded protein. Recent studies on protein refolding reveal a deep implication of chaperone activity in the separation of aggregated polypeptide chains (5–8), proving a significant susceptibility of insoluble polypeptides to solubilization. In recombinant bacteria, the arrest of protein synthesis allows the almost complete disintegration of bacterial IBs and the recovery of functional, properly folded protein (9), revealing that even dense aggregates are not a dead end of an irreversible precipitation process.

Apart from the anticipated therapeutic potential of chaperone overexpression to prevent or dissolve *in vivo* proteinaceous aggregates, the proteolytic susceptibility of insoluble protein is also of interest to understand the molecular features connecting misfolding and proteolysis. Additionally, controlled proteolysis of insoluble polypeptides might eventually represent an alternative approach to remove pathogenic deposits of protein. Despite in bacteria it has been proven that loosely aggregated proteins retain protease susceptibility (10), very little is know about the digestion mechanism of insoluble protein upon precipitation as IBs, where tighter hydrophobic interactions occur between unfolded chains. In this work we have kinetically explored the attack of bacterial IB proteins by trypsin and the subsequent generation of degradation fragments. This approach, that has revealed an *in situ* cascade digestion process very similar to that occurring on soluble protein *in vivo*, indicates the maintenance of protease target sites in inclusion body-embedded pro-

[1] To whom correspondence should be addressed. Fax: 34 935812011. E-mail: Antoni.Villaverde@uab.es.

0006-291X/01 $35.00

tein and their availability for a sequential cleavage, that appears to be also responsible for a chaperone independent process that releases soluble polypeptides from aggregates.

MATERIALS AND METHODS

Protein production and IB purification. Plasmids pJLACVP1 and pJVP1LAC derive from the temperature inducible expression vector pJLA602 (11) and encode closely related β galactosidase fusion proteins (12, 13) plus a constitutively produced lambda CI thermosensitive repressor. The high expression level achieved by the use of p_R and p_L lytic promoters and the low protein digestion rates in the protease I.a deficient *Escherichia coli* strain BL21 (14) promote accumulation of significant amounts of both proteins as IBs (11). Luria-Bertani (LB) medium (14) plus 100 μg/ml ampicillin was used to culture plasmid containing BL21 cells. Detailed procedures for bacterial culture and temperature up-shift can be found elsewhere (9). IBs were purified 5 h after induction of recombinant gene expression by detergent treatment, DNase digestion and repeated washing (15) from two independent batch production processes.

Trypsin digestion, protein fractionating and digestion process analysis. Purified VP1LAC and LACVP1 IBs obtained after 5 h of recombinant gene expression were resuspended in 300 μl of Z buffer (0.06 M Na_2HPO_4, 0.04 M $NaH_2PO_4H_2O$, 0.02 M KCl, $MgSO_4 \cdot 7H_2O$, pH 7.0) and incubated with 100 μg/ml trypsin at 37°C. At different times, samples were withdrawn and the digestion stopped by adding 200 μg/ml soybean trypsin and immediately freezing. When necessary, soluble and insoluble protein fractions were separated by centrifugation of 60 μl aliquots at 9500*g* and 4°C for 15 min. After boiling the samples for 10 min in denaturing buffer (125 mM Tris–Cl pH 6.8, 6% SDS, 20% glycerol, 10% β-mercaptoethanol) plus 8 M urea, digestion products were analysed in Western blot by detection of β galactosidase immunoreactive bands, after PAGE in 3.5% stacking and 7% separation gels. Blot images were obtained in a Hewlett Packard resolutive scanner and relevant bands were quantified by using Quantify One software from Bio Rad. Data from two independent experiments were used for the modeling.

Numerical analysis and modeling the cascade digestion process. According to the molecular masses of relevant bands, the population of IB protein species was organized in the following groups of components. First, the higher level group, of 139 kDa, corresponds to the intact non treated protein. As previously described (9), a heterogeneous nature of full-length IB protein was detected, and more than one different protein classes with absolute densitometric units $N_1(t)$, $N_2(t), \ldots N_n(t)$, at time t, would coexist. Under protease exposure, the fragmentation of these species follows an independent Poisson process (16) resulting in distinguishable expected lives. According to this model, the composition of non fragmented IBs is described as a mixture of exponential density functions for the specific time interval up to the digestion event:

$$\sum_{i=1}^{n} N_i(0) e^{-q_i t};\ 0 < q_1 < q_2 < \cdots < q_n. \qquad [1]$$

While a two component mixture explains qualitatively the IB degradation kinetics, for most of data analysis from independent experiments, the minimal number of protein species required to actually account for the observed data was three (Fig. 3).

A second group of components corresponds to partially degraded IB protein, in decreasing order of molecular masses, $N_{n+1}(t)$, $N_{n+2}(t), \ldots N_{n+m}(t)$. At this level, the presence of more than one species of full-length protein in IBs before protease addition was in general not distinguishable. In fact, the initial protein amount of any of the protein fragments was much lower than that of intact protein in the first level, where the superposition model was applied. The progressive digestion process occurs in one flow direction, from higher to lower molecular mass levels, and the protein amount of each component increased by receiving protein cleaved from upper levels, and decreased by loss of fragmented material toward lower levels also according to a Poisson process. Thus the kinetics of this cascade process can be modeled according to the following differential equation

$$\frac{dN_k(t)}{dt} = \sum_{i=1}^{k-1} p_{ki} N_i(t) - q_k N_k(t);\ k = n+1, \ldots, n+m, \qquad [2]$$

where the constant p_{ki} (the fraction of protein coming from level i to k per unit of time) and q_k are nonnegative. In any level, the characteristic parameter for each protein component is the partial digestion rate q_k (inverse of the expected life (although sometimes the half life $T_i = \ln 2/q_i$ was used as component describing parameter). The constants involved in Eq. [2], were obtained from protein amounts $N_i(t)$, $i = 1, \ldots, n+m$ at different times, by solving a linear overdeterminate system of equations. The derivative was computed from a three point estimation of the derivative (17), and all overdeterminate systems of equations were solved from an iterative weighted least-squares method, according to the algorithm previously described (18). To obtain minimum variance parameters, weights were evaluated from the inverse covariance matrix of experimental errors.

RESULTS

Cascade proteolysis of IB proteins. LACVP1 and VP1LAC are β-galactosidase fusion proteins of the same molecular mass (139 kDa), in which the FMDV VP1 capsid protein had been joined either to the enzyme carboxy or amino termini respectively (12). Under overproduction conditions in an *E. coli* Lon⁻ background, both proteins produce inclusion bodies whose kinetic formation and final structure are distinguishable by analytical methods (15, 19). During trypsin exposure of purified IBs, a progressive digestion of the aggregated protein took place, resulting in a time dependent evolution of immunoreactive, low molecular weight protein fragments concomitant with a decrease in the amounts of the full length forms of the fusions (Fig. 1). Although an important background of immunoreactive material also appeared in some of the lanes, mainly at early times of VP1LAC IB digestion (Fig. 1A, bottom), discrete bands were evident, their intensity in the blots fading during exposure to trypsin. Among these bands, those corresponding to protein fragments of apparent molecular masses of 128, 116, 98, 90, and 69 kDa for VP1LAC, and 128, 116, 114, 100, 90, and 69 kDa for LACVP1, were repeatedly detected in several samples. Therefore, they must represent bottle necks within a cascade digestion process in which systematic, site limited cleavage events on defined protease target sites generate some stable protein fragments. This sequential protein cleavage is confirmed by the time dependent evolution of protein fragments in which transient increases in the amount of some degradation intermediates were observed (Fig. 1).

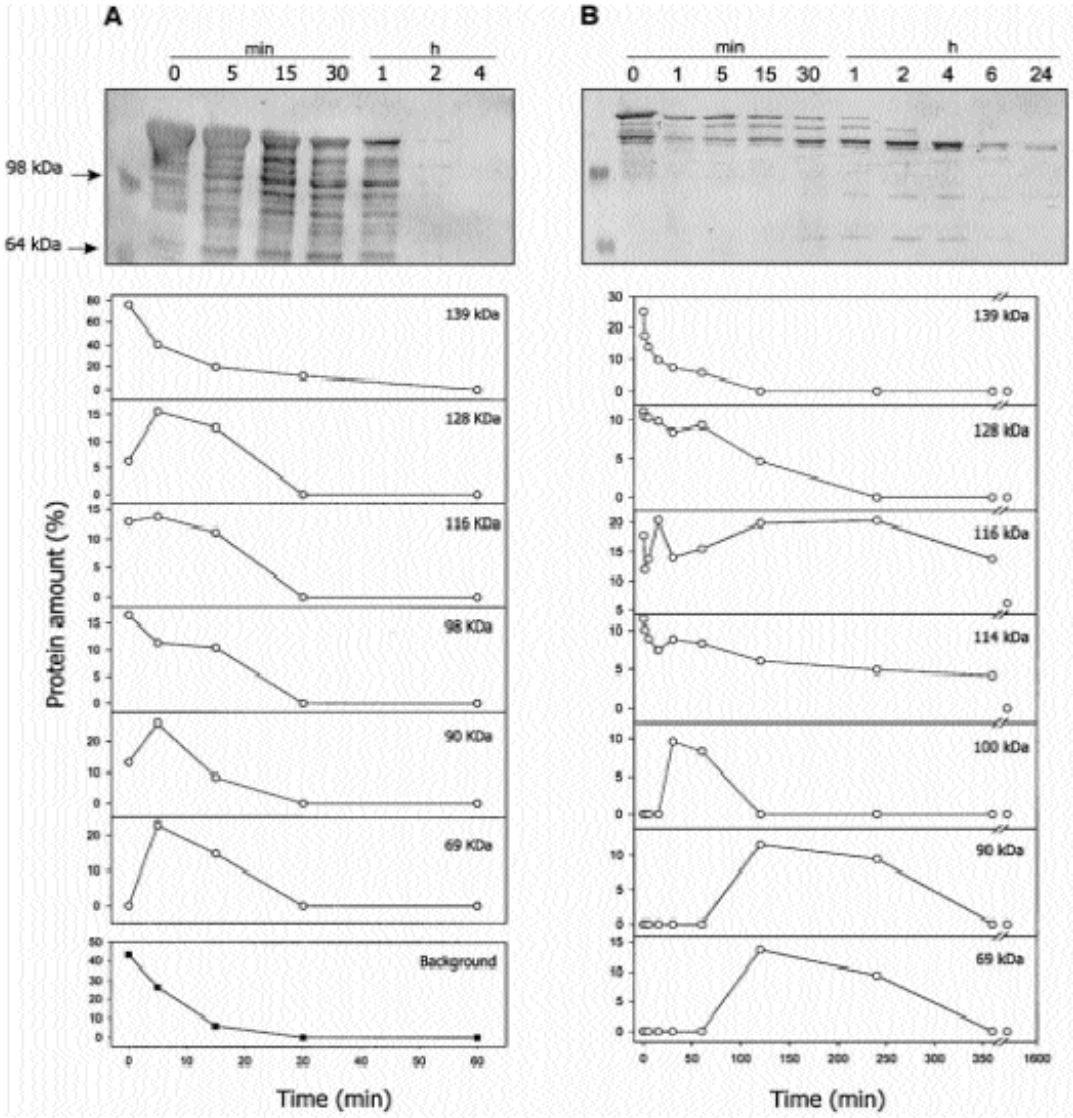

FIG. 1. Western blots revealing β galactosidase immunoreactive proteins during trypsin treatment of purified IBs (VP1LAC in A and LACVP1 in B). Molecular markers were loaded in the left lanes. The digestion time is indicated on the top of each lane. At the bottom, the evolution of different protein species is indicated as amounts relative to those of total protein bands. In A, bottom, the background found between discrete bands along the lane is also represented for VP1LAC IBs, relative to the total densitometric units of bands plus the background itself (in black symbols).

At the exception of 98, 90 and 69 kDa-fragments of protein LACVP1, all the digestion fragments of LACVP1 and VP1LAC were already present in the IB particles. Since under *in vivo* overproduction conditions, protein fragments embedded in IBs are at least mainly generated by proteolysis of soluble protein followed by further precipitation of the digestion products (20), the similar repertoire of protein bands indicates that *in vitro* trypsin degradation of aggregated protein follows a pattern similar to that occurring *in vivo*, at the exception of few specific cleavages. However, it cannot be completely excluded that LACVP1 98, 90, and 69 kDa peptides could also be produced *in vivo* but at lower amounts, or alternatively that they could be poorly incorporated into IBs. It is noteworthy that the 116 and 114 kDa LACVP1 fragments, that are not

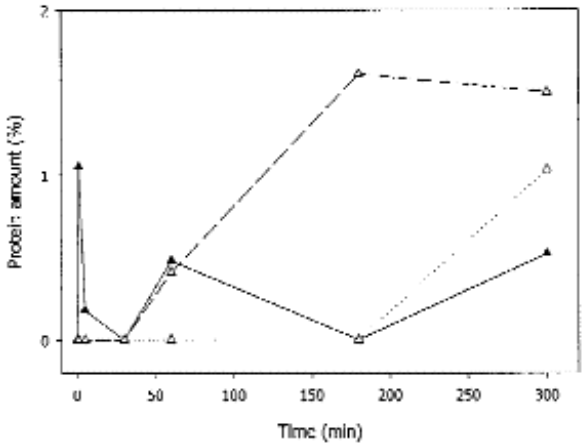

FIG. 2. Protein amounts, relative to total protein, in the soluble cell fraction of trypsin treated VP1LAC IBs. Only these small amounts of intact VP1LAC (black symbols) as well as the 116 kDa (gray symbols) and 90 kDa (white symbols) fragments have been seen in Western blots. No soluble protein has been detected in VP1LAC IBs incubated in buffer in absence of trypsin (not shown).

produced *in vitro*, were extremely resistant to the trypsin attack, being still detected after prolonged incubation (Fig. 1B).

Solubility of IB digestion products. Soluble and insoluble protein was fractionated during trypsin attack of purified VP1LAC IB particles. As shown in Fig. 2, no soluble protein was detected before the addition of the protease in IBs in suspension. However, immediately after starting the digestion process, small amounts of intact protein were transiently found in the soluble cell fraction. The full length protein was lately replaced by major degradation products (Fig. 2), being these few species the only components of soluble protein fraction detected during the digestion process. These proteins are found at very low amounts, representing a maximum of 1.05% with respect to the total protein (at time 1 min). Their dynamic evolution in the soluble fraction suggests that the soluble protein released during IB disintegration could be more protease sensitive than the aggregated polypeptide chains. In addition, these observations also indicate that the time-dependent evolution of protein fragments presented in Fig. 1 involves insoluble protein, and that the cascade protein digestion occurs *in situ* on IB particles. If a parallel cascade digestion process of soluble protein does also occur, it would not be detectable under these digestion conditions.

Modeling the in situ cascade digestion process of IB-embedded protein. Partitioning results presented in Fig. 2, proving that the cascade process observed during *in vitro* IB digestion (Fig. 1) corresponds to insoluble IB protein, offer an unexpected model to investigate kinetically the proteolytic attack of insoluble protein. To approach this unexplored aspect of protein digestion, the sequential appearing of lower molecular mass degradation fragments was modeled for both VP1LAC and LACVP1. Figure 3 represents the flow diagrams obtained for both proteins. Three differently stable intact polypeptide species were detected in the starting protein material, in agreement with previous analysis of bacterial IB architecture (15). While the digestion process of the unstable fractions (and the more stable fraction of VP1LAC) cannot be determined, probably because of the important amount of heterogeneous protein that is distributed along the lanes in the blot (Fig. 1A), a cascade, multi-stepped process is evidenced for the most recalcitrant species. The *in vitro* digestion of VP1LAC involves the appearing of 128, 116, 90, and 69 kDa intermediates, while an additional 100 kDa fragment is evidenced during LACVP1 digestion. Interestingly, a 98 kDa VP1LAC fragment and three LACVP1 degradation intermediates generated *in vivo* and precipitated within IB particles (of 128, 116 and 114 kDa respectively) are not produced during trypsin treatment (Fig. 3). Also, some of these LACVP1 fragments (of 116 and 114 kDa) show unusually high half lives, being even detectable after 24 h of protease treatment (Fig. 1B). This fact, that is not observed in VP1LAC, can only be attributed to a different folding state of both fusion proteins, that having the same amino acid sequence but alternative disposition of the partner VP1 and β galactosidase proteins (12, 13), exhibit distinguishable structural properties that affect, among other features, the digestion pattern and the stability of derived fragments.

DISCUSSION

Bacterial inclusion bodies formed upon recombinant gene overexpression are quite homogeneous in protein composition (19, 21, 22), apparently due to a high specificity during intermolecular hydrophobic interactions (23). However, some chaperones like Ibp are also present (24) being probably trapped during the precipitation events. Apart from the full-length recombinant protein, it is also common to observe degradation fragments of different molecular masses (25, and Fig. 1). These peptide fragments seem to be the result of a dynamic equilibrium between protein aggregation and solubilization during the volumetric growth of IB particles, that allows site limited proteolysis trough transient protein accessibility to cell proteases (20). The truncated polypeptides include species of discrete molecular masses, corresponding to stable degradation products, but also peptides with extremely heterogeneous molecular masses that generate an important background. In VP1LAC IBs, this represents about 40% of the immunoreactive material (Fig. 1A). This background becomes spectrophotometrically undetectable at about 30 min of trypsin treatment of IB

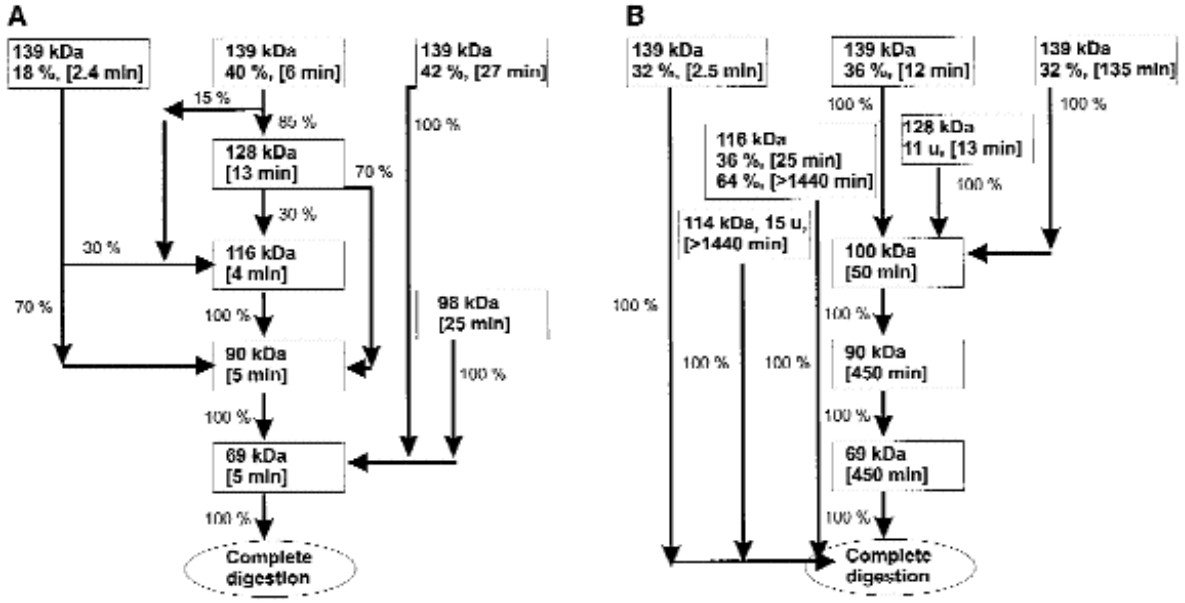

FIG. 3. Modeling of IB digestion for both VP1LAC (A) and LACVP1 (B) IBs. Percentages indicate the fractions of different species in the initial IB samples as well as the distribution between different digestion pathways. Numbers in brackets represent the expected life time of each protein species.

particles, indicating that protein fragments are then either completely digested or degraded until stable intermediates concentrating as bands of increasing intensity. The presence of this inter band material makes the numerical modeling of the cleavage cascade more difficult, resulting in some unsolved steps (Fig. 3). However, both the transient increase on some of these intermediates (Fig. 1) and the global process modeling (Fig. 3) reveal a subsequent appearing of increasingly shortened species and therefore a sequential exposure of trypsin target sites during IB incubation with the protease. Interestingly, the described cascade process occurs in the insoluble fraction of the digestion mixture, while the soluble fraction contains hardly detectable amounts of protein (Fig. 2). In absence of protein synthesis, the presence of replacing protein fragments in the IB fraction cannot be accounted by a dynamic equilibrium between soluble and insoluble forms and preferential digestion of the soluble protein, as probably occurs in biosynthetically active cells (20). On the contrary, the site-limited cleavage is directed on IB polypeptides, generating progressively shorter fragments that remain insoluble and probably retained in the particles during their progressive disintegration (9).

A high protease sensitivity of some IB proteins has been suggested by experimental data (26). Also, looser aggregates of VP1LAC undergo an *in vivo* site limited digestion generating stable fragments within the insoluble cell fraction (10). In this report, we have observed a sequential, site-limited digestion process that implies 3 (for LACVP1) or 4 (for VP1LAC) detectable steps before the complete protein digestion (Fig. 3), only involving insoluble IB protein. Since the activation of target sites for site limited proteolysis requires conformational modifications in the protein (27, 28), these conformational changes must occur *in vitro* during trypsin digestion of IB particles. On the other hand, it is already known that a restricted cleavage event can modify the conformation of soluble protein, thus dramatically affecting susceptibility to further attack by proteases (28). The data presented here indicates that a such conformational events can also take place within protein aggregates. In this context, the release of small amounts of soluble protein observed during the digestion process (Fig. 2) could be linked to the degradation of more sensitive protein species, affecting the molecular environment of stable peptide chains allowing their solubilization. Alternatively, the discriminating erosion of the IB particles mediated by the heterogeneity of IB protein conformation (15), could promote a mechanical release of IB material to the soluble fraction during body fragmentation.

Despite bacterial IBs have been often considered as homogeneous, inert depositions of insoluble protein, several independent observations reporting enzymatic activity associated to IB particles (29, 30), high voidage accessible to the solvent (31), native-like secondary structure of aggregated IB protein (32), dynamic het-

erogeneity in IB protein organization (15), refolding potential of IB proteins (9, 20) and both *in vivo* and *in vitro* reversibility of IB formation (9) indicate an unexpected molecular plasticity within these aggregates. The results presented in this work indicate that at least a significant fraction of IB protein is susceptible to sequential, site-limited proteolytic attack, with a cascade activation of protease target sites, as it occurs in soluble protein by a conformation dependent way. In addition, minor amounts of stable protein are released from IBs incubated with protease but not in its absence (not shown), being this a possibility for an alternative, chaperone-independent mechanism of *in vivo* IB disintegration.

ACKNOWLEDGMENTS

We are indebted to J. Checa and V. Ferreres for technical assistance. This work has been supported by Grant BIO98-0527 (CICYT) and by Mª Francesca de Roviralta Foundation. M.M.C. has been recipient of a predoctoral fellowship from UAB, Spain.

REFERENCES

1. Herman, C., and DAri, R. (1998) Proteolysis and chaperones: The destruction/reconstruction dilemma. *Curr. Opin. Microbiol.* **1,** 204–209.
2. Hellebust, H., Murby, M., Abrahmsen, L., Uhlén, M., and Enfors, S-O. (1989) Different approaches to stabilize a recombinant fusion protein. *BioTechnology* **7,** 165–168.
3. Marston, F. A. O. (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem. J.* **240,** 1–12.
4. Hope, J. (2000) Prions and neurodegenerative diseases. *Curr. Opin. Genet. Dev.* **10,** 568–574.
5. Warrick, J., Chan, H. Y. E., Gray Board, G. L., Chai, Y. H., Paulson, H. L., and Bonini, N. M. (1999) Suppression of polyglutamine mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. *Nat. Genet.* **23,** 425–428.
6. Kazemi Esfarjani, P., and Benzer, S. (2000) Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science* **287,** 1837–1840.
7. Muchowski, P. J., Schaffar, G., Sittler, A., Wanker, E. E., Hayer Hartl, M. K., and Hartl, F. U. (2000) Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid like fibrils. *Proc. Natl. Acad. Sci. USA* **97,** 7841–7846.
8. Diamant, S., Ben Zvi, A. P., Bukau, B., and Goloubinoff, P. (2000) Size dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. *J. Biol. Chem.* **275,** 21107–21113.
9. Carrió, M. M., and Villaverde, A. (2001) Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett.* **489,** 29–33.
10. Corchero, J. L., Cubarsi, R., Enfors, S. O., and Villaverde, A. (1997) Limited *in vivo* proteolysis of aggregated proteins. *Biochem. Biophys. Res. Commun.* **237,** 325–330.
11. Schauder, B., Blöcker, H., Frank, R., and McCarthy, J. E. G. (1987) Inducible vectors incorporating the *Escherichia coli atpE* translational initiation region. *Gene* **52,** 279–283.
12. Corchero, J. L., Viaplana, E., Benito, A., and Villaverde, A. (1996) The position of the heterologous domain can influence the solubility and proteolysis of β galactosidase fusion proteins. *J. Biotechnol.* **48,** 191–200.
13. Corchero, J. L., and 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.
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. Carrió, M. M., Cubarsi, R., and Villaverde, A. (2000) Fine architecture of bacterial inclusion bodies. *FEBS Lett.* **471,** 7–11.
16. Papoulis, A. (1989). Probability, Random Variables, and Stochastic Processes, McGraw Hill, Singapore.
17. Kincaid, D., and Cheney, W. (1991) Numerical Analysis, Brooks-Cole Publishing Co., Belmont, CA.
18. Cubarsi, R., Corchero, J. L., Vila, P., and Villaverde, A. (1998) Numerical techniques and mathematical modeling for C1857-controlled gene expression and cell growth in recombinant *E. coli*. *IMA J. Math. Appl. Biol. Med.* **15,** 257–278.
19. Carrió, M. M., Corchero, J. L., and Villaverde, A. (1998) Dynamics of *in vivo* protein aggregation: Building inclusion bodies in recombinant bacteria. *FEMS Microbiol. Lett.* **169,** 9–15.
20. Carrió, M. M., Corchero, J. L., and Villaverde, A. (1999) Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. *Biochem. Biophys. Acta* **1434,** 170–176.
21. Bowden, G. A., Paredes, A. M., and Georgiou, G. (1991) Structure and morphology of protein inclusion bodies in *Escherichia coli*. *BioTechnology* **9,** 725–730.
22. Valax, P., and Georgiou, G. (1993) Molecular characterization of beta-lactamase inclusion bodies produced in *Escherichia coli*. 1. Composition. *Biotechnol. Prog.* **9,** 539–547.
23. Speed, M. A., Wang, D. I., and King, J. (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat. Biotechnol.* **14,** 1283–1287.
24. Allen, S. P., Polazzi, J. O., Gierse, J. K., and Easton, A. M. (1992) Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *J. Bacteriol.* **174,** 6938–6947.
25. Rinas, U., and Bailey, J. E. (1992) Protein compositional analysis of inclusion bodies produced in recombinant *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **37,** 609–614.
26. Weickert, M. J., and Curry, S. R. (1997) Turnover of recombinant human hemoglobin in *Escherichia coli* occurs rapidly for insoluble protein and slowly for soluble protein. *Arch. Biochem. Biophys.* **348,** 337–346.
27. Hubbard, S. J., Eisenmenger, F., and Thornton, J. M. (1994) Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci.* **3,** 757–768.
28. Viaplana, E., Feliu, J. X., Corchero, J. L., and Villaverde, A. (1997) Reversible activation of a cryptic cleavage site within *E. coli* β galactosidase in β galactosidase fusion proteins. *Biophys. Biochem. Acta* **1343,** 221–226.
29. Tokatlidis, K., Dhurjati, P., Millet, J., Béguin, P., and Albert, J. P. (1991) High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D. *FEBS Lett.* **282,** 205–208.
30. Worrall, D. M., and Goss, N. H. (1989) The formation of biologically active β galactosidase inclusion bodies in *Escherichia coli*. *Aust. J. Biotechnol.* **3,** 28–32.
31. Taylor, G., Hoare, M., Gray, D. R., and Marston, F. A. O. (1986) Size and density of protein inclusion bodies. *BioTechnology* **4,** 553–557.
32. Oberg, K., Chrunik, B. A., Wetzel, R., and Fink, A. L. (1994) Native like secondary structure in interleukin 1β inclusion bodies by attenuated total reflectance FTIR. *Biochemistry* **33,** 2628–2634.

APROXIMACIÓ A LES PROPIETATS ESTRUCTURALS DE LA PROTEÏNA AGREGADA EN COSSOS D'INCLUSIÓ

INTRODUCCIÓ

Conèixer l'estructura que presenta la proteïna agregada en els cossos d'inclusió pot ser important per a entendre com es formen els cossos d'inclusió i per avaluar la potencialitat de recuperar la proteïna nativa. Estudis preliminars proposaven que els cossos d'inclusió estaven compostos per intermediaris de plegament en estat desnaturalitzat compacte o en estat de "molten globule". Aquests dos estats presenten un elevat contingut en estructura secundària i algunes interaccions terciaries, formant estats intermediaris entre la forma nativa i desnaturalizada, amb residus hidrofòbics exposats al solvent que els fan propensos a l'agregació. Aquests dos estats es diferencien en que l'estat de molten globule presenta una estructura quasi nativa, mentre que el desnaturalitzat compacte no (Mitraki *i col*., 1991).

Hi ha alguns treballs posteriors que han analitzat l'estructura secundària dels cossos d'inclusió, com és el cas dels agregats d'interleukina-1β, analitzats mitjançant espectroscòpia d'infraroig (FTIR). Aquests mostren una estructura secundaria molt propera a la nativa i l'única diferència observada és un increment en el nombre de ponts d'hidrogen de les fulles beta, el que indica que les fulles beta estan empaquetades per unions més fortes en els cossos d'inclusió que en l'estructura nativa. Això suggereix que hi ha una topologia diferent en la seva estructura supersecundària, és a dir, en la manera en com interaccionen els sudbomins proteics, entre els quals s'afavoreix la formació de ponts d'hidrogen entre les fulles beta inter i intramoleculars (Oberg *i col*., 1994). Els cossos d'inclusió de β-lactamasa, produïts en diferents condicions i diferents compartiments cel·lulars, han estat estudiats per espectroscopia de Raman i s'observa que presenten una estructura secundària pertorbada, en la que incrementa el contingut de fulles beta a expenses d'hèlix alfa. Aquesta pertorbació és més acusada quan els agregats es produeixen en temperatures elevades, però no es veu afectada per la compartimentació subcel·lular (Pryzbyecien *i col*., 1994).

Així doncs, la formació d'interaccions intermoleculars entre fulles beta sembla que sigui un denominador comú en els processos d'agregació *in vivo* i *in vitro* (Ellis i Pinheiro, 2000), però les dades que es tenen sobre els polipèptids que intervenen en la formació de cossos d'inclusió són escasses i variables, i es desconeix si la proteïna agregada en forma de cossos d'inclusió presenta propietats estructurals comunes.

Per això vam creure interessant fer un estudi estructural de la proteïna recombinant que forma els cossos d'inclusió de VP1LAC i LACVP1. Com ja s'ha descrit prèviament, aquests presenten diferències morfològiques i de composició destacades, essent els agregats de VP1LAC més grans i amb una composició proteica més homogènia i els de LACVP1, més petits i heterogenis (treball 2).

Per altra banda, en estudis previs havíem observat que els cossos d'inclusió són estructures dinàmiques, formades per diferents espècies polipeptídiques amb diferent sensibilitat a proteases, que varia al llarg del seu creixement, el que indica un dinamisme conformacional (treball 5). Aquest treball analitza cinèticament els canvis estructurals que pateixen els polipèptids que formen els cossos d'inclusió de VP1LAC i LACVP1 durant el seu creixement en el citoplasma d' *E.coli*, utilitzant tècniques espectroscòpiques de fluorimetria i dicroïsme circular i analitzant la susceptibilitat dels diferents agregats envers agents caotròpics. Aquests anàlisis es van realitzar al laboratori de bioquímica de la Dra. Debora Foguel, on vaig realitzar dues estades curtes durant la tesi.

MATERIALS I MÈTODES

1. Producció i purificació de cossos d'inclusió

Els cossos d'inclusió es produeixen a la soca d'*Escherichia coli* Lon- BL21. Els plasmidis pJVP1LAC i pJLACVP1 codifiquen dues proteïnes de fusió molt relacionades entre si, portant la proteïna VP1 de la càpside del virus de la febre aftosa a l'extrem carboxi o amino- terminal de la β-galactosidasa, respectivament. Aquests vectors d'expressió termoinduïbles estan regulats pels promotors p_R i p_L del bacteriòfag lambda. Els cultius creixen en medi ric (LB) a 28°C i 250 rpm i quan entren en fase exponencial s'indueix la producció de les proteïnes recombinants per temperatura, transferint-los a 42°C. A partir d'aquest moment es recullen mostres de 100ml cada hora, durant 5 hores i a les 24 hores, per a purificar els cossos d'inclusió.

La purificació dels cossos d'inclusió es fa a través de repetits tractaments amb detergents després de la lisi cel·lular. Les cèl·lules es recullen per centrifugació a 12396g durant 15 minuts, es resuspenen amb tampó de llisi (50 mM Tris-HCl, 1mM EDTA, 100mM NaCl, pH 8) i s'hi afegeix lisozima i l'inhibidor de proteases PMSF (Phenylmethylsulfonyl fluoride). Després d'haver incubat la barreja durant 30 minuts, a 37°C amb agitació, s'afegeix NP40 (1%) i s'incuba a 4°C amb agitació suau, durant 30 minuts. Seguidament s'incuben amb DNAasa i MgSO4 a 37°C durant 30 minuts més i es recullen els agregats per centrifugació a 12396g durant 30 minuts. Finalment, els cossos d'inclusió es renten per centrifugació dues vegades amb tampó de lisi més 0,5% triton X-100 i es resuspenen amb tampó fosfat fins que s'utilitzen. La proteïna total dels cossos d'inclusió es quantifica amb el reactiu bradford (Bradford, 1976), havent estat prèviament desnaturalitzades per urea.

2. Anàlisis de fluorimetria

2.1. Seguiment de la fluorescència del triptòfan

L'emissió de fluorescència té lloc quan un electró excitat retorna a l'estat basal emetent energia en forma de llum. Les proteïnes tenen aquesta propietat gràcies als fluoròfors naturals que tenen, que són els aminoàcids fenilalanina, tirosina i triptòfan. De tots ells, el triptòfan és el que contribueix més en l'emissió de fluorescència, ja que tant la seva absorció d'energia a la

longitud d'ona d'excitació com la intensitat d'emissió són considerablement més altes que els valors que presenten la tirosina i la fenilalanina (Lakowicz, 1986). El seguiment de la fluorescència que emeten els triptòfans ens dóna una idea de l'estat estructural en què es troba la proteïna, ja que ens informa sobre la seva exposició al solvent, i comparant-ho amb la proteïna nativa, podem detectar canvis estructurals.

Per analitzar la fluorescència del triptòfan dels diferents cossos d'inclusió purificats, es preparen les mostres a una mateixa D.O. a 280, de 0,4, i s'obté l'espectre de fluorescència de cada mostra, excitant a 280nm i recullint l'espectre d'emissió de 300 a 400 nm. D'aquest espectre s'obté l'àrea total de fluorescència i la màxima emissió de triptòfan, que prenem com a valor representatiu de l'estat conformacional de la proteïna. Els espectres de triptòfan s'han fet en un espectrofluoròmetre ISS K2 (Foguel *i col*., 1995).

2.2. Seguiment de l'espatllament de llum

L'espatllament de llum és una mesura òptica que ens dóna un valor d'agregació donat que mesura la quantitat de llum que és rebotada al xocar amb una partícula agregada. Quanta més agregació hi hagi, més quantitat de llum espatllada. Les mostres que s'utilitzen per a analitzar l'espatllament de llum són les mateixes utlitzades per obtenir els espectres del triptòfan i s'obtenen excitant les mostres a 320 i recullint l'espectre d'emissió de llum incident a 90ºC, a través de la intensitat de llum integrada entre 315 i 325 nm.

2.3. Unió de la proteïna agregada al ANS

L' ANS (bis 8-anilinonaphthahalene-1-sulfonate) és una molècula que exhibeix poca fluorescència en solució aquosa, però que es torna altament fluorescent quan s'uneix a regions hidrofòbiques estructurades (Horowitz i Butler, 1993) i per això s'utilitza per a prevenir l'agregació o per a fer-ne el seguiment. La unió a l'ANS es detecta excitant les mostres a 360nm i recollint l'emissió entre el rang de 400-600nm (Silva *i col.*, 1992). Per a mesurar la intensitat de la unió s'utilitza l'àrea de l'espectre de fluorescència. La concentració d' ANS que s'ha usat en totes les mostres és de 5 mM.

2.4. Anàlisis de dicroïsme circular

El dicroïsme circular (DC) és una tècnica d'espectroscopia que mesura les diferències d'absorció de la llum polaritzada, circularment a la dreta i a l'esquerra, per a cada substància de la mostra. S'ha vist que els espectres de dicroïsme circular de la zona entre 260 i 180nm, també anomenada regió amida, ens donen informació sobre l'estat dels enllaços peptídics i és molt útil per a identificar l'estructura secundària i detectar-ne variacions, ja que es coneixen els espectres de DC característics per les estructures secundàries de tipus hèlix alfa, fulles beta paral·leles i antiparal·leles,"turn" i altres, que es mostren a la figura 1. L'elipticitat és la unitat del dicroïsme circular (es defineix com la tangent del radi des del menor al major axis elíptic) i s'utilitza per a donar un valor quantitatiu de cada tipus d'estructura.

Les mostres dels cossos d'inclusió purificats resuspeses en tampó fosfat, a una mateixa absorvància a 280nm (0,4), es van col·locar en una cubeta de quars de 1mm i es van obtenir els espectres a una velocitat d'escaneig de 50 nm/min., a una resolució de 1 nm i una sensibilitat de 50 mdeg.

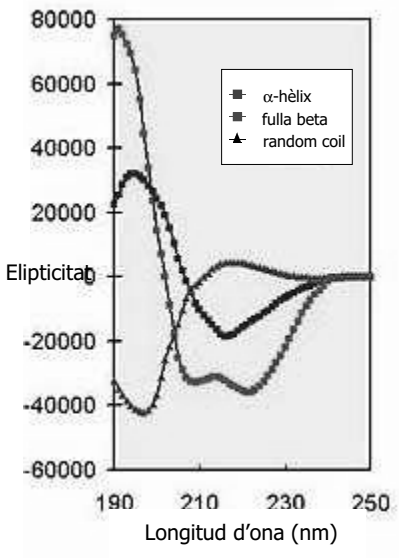

Figura 1. Espectres de dicroïsme circular típics d'hèlix α (amb un mínim a 208 nm i un a 222 nm),de fulla beta (amb un mínim a 218 nm) i de "random coil".

3. Pressió hidrostàtica

Les mostres dels cossos d'inclusió purificats i resuspesos en tampó fosfat es sotmeten a una pressió hidrostàtica de 40Kpsi, que s'aplica de cop, mitjançant una bomba de pressió que va afegint etanol a la mostra, tancada al buit dins d'una ampolleta de vidre amb un tap de plàstic moldejable (veure treball 11). El seguiment de l'agregació i l'estat de la mostra es fa mitjançant el fluorímetre connectat a la bomba i així es recullen espectres de triptòfan i d'espatllament de llum durant la pressurització.

RESULTATS I DISCUSSIÓ

1. Seguiment dels espectres de fluorescència del triptòfan

Les dues proteïnes estudiades, VP1LAC i LACVP1, tenen un espectre de fluorescència similar quan estan agregades en els cossos d'inclusió, la qual cosa podria indicar que presenten una estructura tridimensional similar (Fig. 2).

La β-galactosidasa salvatge, en estat natiu, presenta un màxim d'emissió de fluorescència de 340,75 ± 1nm (Feliu *i col*., 1998), i no sembla variar quan aquesta és fusionada amb la proteïna VP1, formant la proteïna VP1LAC (340,5), com es veu a la Fig. 2.

L'espectre del triptòfan de la proteïna dels cossos d'inclusió és força diferent de quan es troba en el seu estat soluble. La proteïna agregada presenta un màxim d'emissió del triptòfan bastant menor, amb 8nm de diferència, i aquest canvi suggereix que els triptòfans es deuen trobar més amagats en els polipèptids agregats. Tot i que els dos tipus de cossos d'inclusió estudiats, VP1LAC i LACVP1, són morfològicament diferents, els polipèptids que els formen no semblen presentar grans diferències estructurals, ja que els espectres de triptòfan respectius són molt similars.

Com es veu a la figura 2, la proteïna present als agregats no sembla patir canvis radicals en l'espectre d'emissió del triptòfan durant les primeres fases de creixement dels cossos d'inclusió, però sí després de 24 hores, quan augmenta considerablement el màxim d'emissió de triptòfan. Aquest augment

indica que la proteïna que resta en els cossos d'inclusió al cap de 24 hores de la seva formació presenta un estat conformacional diferent al de les primeres hores.

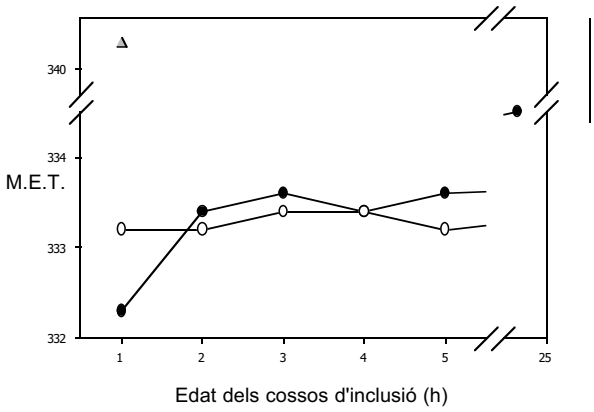

Figura 2. Màxima emissió del triptòfan dels cossos d'inclusió de VP1LAC (cercle negre) i LACVP1 (cercle blanc) a les primeres 5 hores i 24 hores. La màxima emissió del triptòfan de la proteïna VP1LAC en estat soluble es mostra com a referència (triangle gris)

2. Anàlisi de l'estructura secundària dels cossos d'inclusió durant la seva formació, per dicroïsme circular

Els cossos d'inclusió de VP1LAC estan compostos bàsicament per la proteïna recombinant intacta (90%), mentre que els de LACVP1 presenten un 70-80% de proteïna recombinant, repartida entre la forma intacta i el fragment de degradació estable "β-galactosidasa-like". Per analitzar els espectres de DC d'aquests agregats s'ha de tenir en compte que encara que les proteïnes coagregades representin un percentatge baix en l'agregat, poden intervenir d'una manera important en el

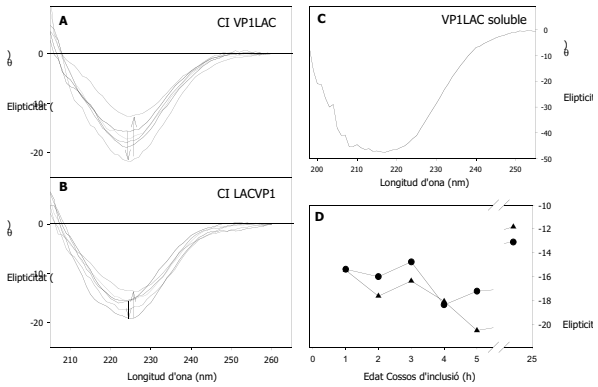

Figura 3. Espectres de DC de la proteïna VP1LAC soluble (B) i dels cossos d'inclusió purificats de VP1LAC (A) i LACVP1 (C) de les primeres 5 hores i 24 hores de creixement (negre, 1h., verd, 2h., vermell, 3h., blau, 4 h., rosa, 5h., gris, 24 h.). En el panel D es mostra l'evolució de l'elipticitat a 222 nm dels cossos d'inclusió de VP1LAC (triangles) i LACVP1 (cercles) per veure l'increment del senyal durant les primeres hores i la seva pèrdua a les 24 hores.

senyal de l'espectre.

Els espectres dels cossos d'inclusió de VP1LAC i LACVP1 purificats indiquen que una part de la proteïna de la mostra conté estructura secundària. L'espectre mostra un mínim atípic, als 225nm, que es podria explicar com un artefacte produït pels agregats de la mostra o per la presència d'anells aromàtics, que desplaçarien el mínim típic d'hèlix α (222nm). El senyal de DC de les mostres dels cossos d'inclusió de les primeres 5 hores de la seva formació incrementa amb el creixement dels agregats (Fig. 3). Com que el DC mostra estructura secundària soluble (els agregats no dónen senyal), el senyal detectat en els cossos d'inclusió podria representar la fracció de proteïna més susceptible a la solubilitazació. El fet que la forma dels espectres siguin iguals (Fig. 4) i augmenti el senyal durant aquest període de temps, indica que la forma plegada existent a les

mostres es podria estar estabilitzant, això significa que l'equilibri que hi ha entre la forma desplegada i la plegada està desplaçat cap a la forma plegada. Tanmateix, en els cossos d'inclusió de 24 hores la senyal de DC disminueix significativament, la qual cosa suggereix que la proteïna està més agregada (Fig. 3). Per altra banda, l'espectre presenta un lleuger desplaçament cap a la dreta, que també podria estar relacionat amb el fenòmen de l'agregació.

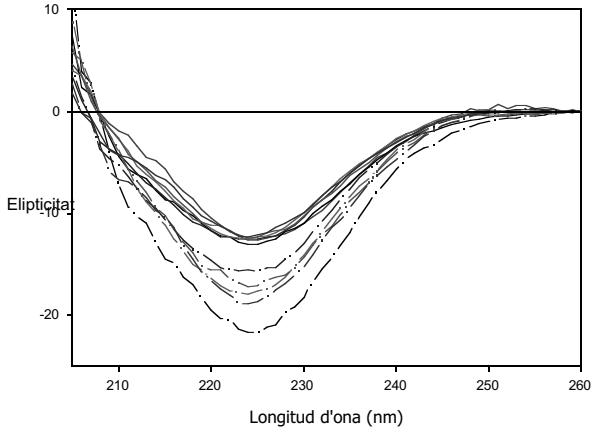

Figura 4. Els espectres de DC dels cossos d'inclusió de VP1LAC (línies discontinues) s'han multiplicat per un factor (línies continues), per a sobreposar-los i veure si presenten la mateixa forma. Es veu que els espectres de les primeres 5 hores tenen la mateixa forma, però el de les 24 hores (línia vermella) presenta un petit desplacament.

3. La unió dels agregats a la molècula ANS indica canvis estructurals durant el creixement dels cossos d'inclusió:

El fet que la proteïna agregada s'uneixi a la molècula ANS indica que presenta regions hidrofòbiques ben exposades. La molècula d' ANS emet fluorescència quan s'uneix a regions hidrofòbiques. Per tant, la intensitat de la fluorescència de l' ANS es pot utilitzar per a detectar la quantitat de parts hidrofòbiques que presenta la proteïna estudiada, les quals poden intervenir en les interaccions intermoleculars que porten a l'agregació.

La gràfica mostra com augmenta la unió de l' ANS amb els cossos d'inclusió durant el seu procés de creixement, indicant una major quantitat de regions hidrofòbiques exposades, que seria compatible amb l'estabilització de la forma plegada, que suggereixen els espectres de DC. Els cossos d'inclusió de LACVP1 presenten una unió més dèbil a ANS que els de VP1LAC i l'augment no és tan acusat.

Una vegada més s'observa que els cossos d'inclusió de 24 hores presenten diferències estructurals respecte els més joves, el que aquí es reflexa amb una menor unió l' ANS. Això suggereix que la proteïna en els cossos de 24 hores presenta menys regions hidrofòbiques exposades, el que podria ser degut a que la proteïna estigui més

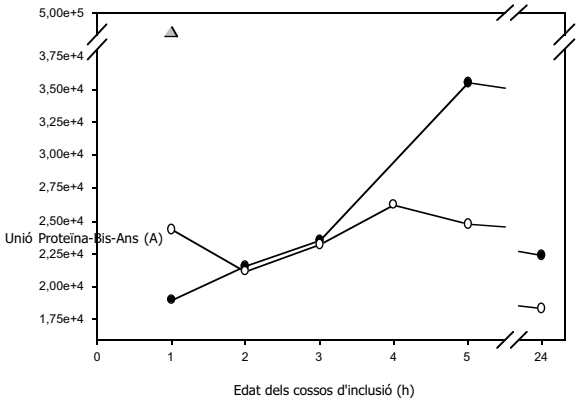

Figura 5. Fluorescència de l'ANS quan s'uneix als cossos d'inclusió de VP1LAC (cercles negres) i LACVP1 (cercles blancs), de diferents edats de creixement.

agregada.

4. Dissociació dels cossos d'inclusió per urea

4.1. Els cossos d'inclusió vells mostren més sensibilitat a la desnaturalització per urea que els joves

S'ha analitzat el patró de dissociació dels cossos d'inclusió de LACVP1 i VP1LAC, de 2 i de 24 hores amb urea, incubant les mostres dels agregats purificats amb concentracions creixents entre 0 i 8 M d'urea. La dissociació dels agregats es detecta amb la caiguda de l'absorvància a 320nm i l'espatllament de llum (Fig. 6), que ens dóna una estima d'agregació, i l'activitat i l'espectre de triptòfan (Fig. 6), que ens informa sobre el plegament de la proteïna.

Els cossos de LACVP1 i de VP1LAC de 24 hores mostren una major sensibilitat a la urea, això es reflecteix amb la ràpida caiguda de l'absorvància a 320nm i l'espatllament de llum, que mostra la dissociació dels agregats. Els espectres de triptòfan mostren la desestructuració de la proteïna, que també es reflecteix amb la pèrdua d'activitat enzimàtica.

4.2. La proteïna soluble és més sensible a la urea que la proteïna agregada

Per a comparar l'efecte de la urea sobre la proteïna nativa i l'agregada, s'han fet corbes de dissociació del cossos d'inclusió de VP1LAC de 2 i 24 hores i de la proteïna soluble incubats a concentracions creixents

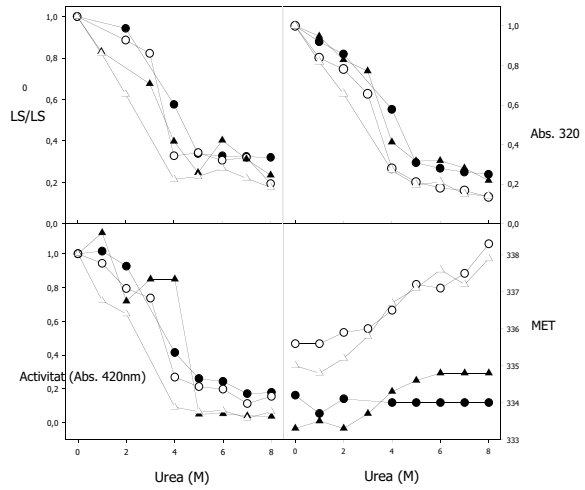

Figura 6. Dissociació dels cossos d'inclusió de VP1LAC (triangles) i LACVP1 (cercles) de 2 hores (negres) i de 24 hores (blancs) al ser incubats amb concentracions creixents d'urea. El panel A mostra la caiguda de l'espatllament de llum normalitzat amb el valor inicial. La baixada de l'absorvància a 320 nm dóna una estima de la desagregació dels cossos (B), acompanyada amb la pèrdua d'activitat enzimàtica (C). En el panel D es mostra com canvia la fluorescència del triptòfan durant la desnaturalització per urea.

d'urea (entre 0 i 8M).

Seguint la unió de la proteïna amb l'ANS durant la desintegració dels cossos, es detecten els canvis estructurals que pateix la proteïna durant la dissociació dels agregats, diferents dels que pateix la proteïna soluble. Curiosament, s'observa que la proteïna dels cossos d'inclusió passa per diferents estats, abans d'ésser desnaturalitzada completament. Amb petites concentracions d'urea, s'observa un augment de la unió amb l'ANS, la qual cosa suggereix que la proteïna adquireix un estat on grups hidrofòbics que estan interaccionant formant grups (clusters), estan exposats al solvent, aquest estat és característic de molten globule. A partir d'aquest estat, la proteïna es desnaturalitza ràpidament, amb concentracions d'urea més elevades. Aquests resultats indiquen que la proteïna dels cossos d'inclusió pot estar en forma de molten globule (o "globule fos"). La urea fa que les

proteïnes s'hidratin, que pugui entrar solvent a la mostra, la qual cosa deu fer que augmentin els grups hidrofòbics típics de molten globule i per això hi ha una major unió amb l'ANS.

Aquí s'observa que els cossos de 24 hores tenen més sensibilitat a la urea i es requereix menys concentració de l'agent per a formar el globule fos previ a la desnaturalització.

A diferència de la proteïna dels cossos, la proteïna soluble es desnaturalitza directament, ja que no presenta interaccions entre residus hidrofòbics exposats, que hagi de trencar primer.

Comparant com varien els espectres de triptòfan dels cossos d'inclusió amb la proteïna soluble després d'ésser incubats amb urea, s'observa que la proteïna soluble pateix un canvi molt més destacat, variant 10 nm el màxim d'emissió del triptofan d'una manera cooperativa, a diferència dels cossos d'inclusió, en els que sols canvia 2 nm i de manera no cooperativa. Això pot indicar que la diferència conformacional de la proteïna nativa respecte la desnaturalitzada és molt més acusada que respecte la proteïna agregada, així doncs la proteïna agregada deu presentar una conformació intermitja entre l'estructura nativa i la desnaturalitzada.

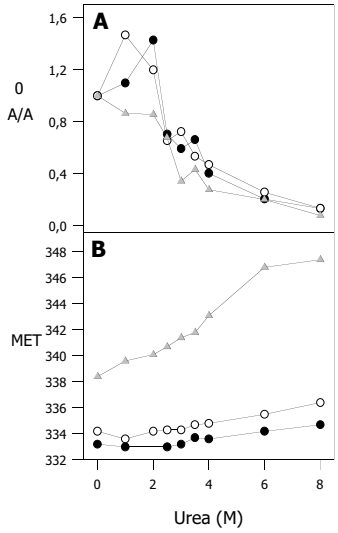

Figura 7. Seguiment de la unió de la proteïna dels cossos d'inclusió de VP1LAC (cercle) de 2h (negre) i 24h (blanc) amb la molècula ANS, durant la seva dissociació per urea (panel A). Els triangles grisos mostren la VP1LAC soluble. En el panel B es compara la fluorescència del triptòfan dels cossos d'inclusió amb la de la proteïna soluble.

CONCLUSIONS

Els cossos d'inclusió són el resultat de l'acumulació de polipèptids sobreproduits, que es troben en un equilibri dinàmic entre l'agregació i la solubilització, desplaçat cap a l'agregació quan les condicions de sobreproducció són òptimes (treball 2). Aquest treball mostra que la proteïna dels cossos d'inclusió presenta un dinamisme conformacional durant el seu creixement, que pot estar relacionat amb aquest equilibri. Els resultats obtinguts semblen indicar que durant les primeres hores de formació dels cossos d'inclusió, la proteïna que es troba en un balanç entre la forma plegada i desplegada, es desplaça cap a la forma plegada. La fracció de proteïna que dóna senyal en els espectres de DC podria correspondre a la proteïna dels cossos d'inclusió que es troba en un estat transitori entre l'agregació i la solubilització. El desplaçament d'aquesta fracció cap a la forma plegada, podria significar que durant aquest període de temps, augmenta la tendència cap a la solubilització. Els cossos d'inclusió de 24 hores no segueixen aquesta tendència, sinó que sembla que estiguin formats per una proteïna més agregada (Fig. 3 i 5), això ens suggereix que podrien representar el resultat de l'equilibri entre l'agregació i la solubilització, que amb el temps s'ha estabilizat, quedant la proteïna més compactada en els cossos d'inclusió. Per altra banda, en estudis anteriors havíem vist que els cossos d'inclusió de 24 hores són més resistents a la proteolisi *in vitro* que els cossos més joves (treball 5), la qual cosa és compatible

amb aquests agregats més compactes.

Aquests resultats també són interessants perquè es veu que la proteïna dels cossos d'inclusió de VP1LAC i LACVP1 es troba en estat de molten globule, el que s'observa clarament amb petites concentracions d'urea (Fig. 6). El fet que resultats anteriors ja trobin estructura de tipus molten globule en els cossos d'inclusió, i que aquest estat sigui molt proper al natiu, dóna esperances positives per a la recuperació d'aquesta proteïna agregada.

4.5. Paper dels cossos d'inclusió en la biologia de l'estrés cel·lular

Al veure que l'agregació en forma de cossos d'inclusió era un procés que podia ésser revertit en el citoplasma cel·lular, ens va semblar interessant identificar quins factors cel·lulars n'eren responsables. Les proteïnes de xoc tèrmic, xaperones i proteases porten a terme el sistema de control de qualitat proteic a la cèl·lula, i per tant estan implicades en l'agregació i la solubilització de proteines. Per això es va estudiar la formació i la desintegració dels cossos d'inclusió en una sèrie de soques mutants deficients en les xaperones DnaK, GroEL, GroES i IbpAB; les proteases ClpA i ClpB i el doble mutant DnaKClpB. El fet d'observar que els components de la resposta a l'estrés cel·lular estan implicats en la formació dels cossos d'inclusió, ens mostra que es tracta d'un mecanisme quiat per la cèl·lula per afrontar les condicions adverses provocades per la sobreproducció d'una proteïna recombinant.

ELSEVIER

Journal of Biotechnology 96 (2002) 3–12

Journal of Biotechnology

www.elsevier.com/locate/jbiotec

Review

Construction and deconstruction of bacterial inclusion bodies

M.M. Carrio, A. Villaverde *

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

Received 6 July 2001; received in revised form 17 January 2002; accepted 6 February 2002

Abstract

Bacterial inclusion bodies (IBs) are refractile aggregates of protease-resistant misfolded protein that often occur in recombinant bacteria upon gratuitous overexpression of cloned genes. In biotechnology, the formation of IBs represents a main obstacle for protein production since even favouring high protein yields, the in vitro recovery of functional protein from insoluble deposits depends on technically diverse and often complex re-folding procedures. On the other hand, IBs represent an exciting model to approach the in vivo analysis of protein folding and to explore aggregation dynamics. Recent findings on the molecular organisation of embodied polypeptides and on the kinetics of inclusion body formation have revealed an unexpected dynamism of these protein aggregates, from which polypeptides are steadily released in living cells to be further refolded or degraded. The close connection between in vivo protein folding, aggregation, solubilisation and proteolytic digestion offers an integrated view of the bacterial protein quality control system of which IBs might be an important component especially in recombinant bacteria.

Keywords: Chaperones; *Escherichia coli*; IbpA; IbpB; Inclusion bodies; Protein aggregation; Refolding; Recombinant protein

1. Introduction

The formation of amorphous proteinaceous granules in *Escherichia coli* was first described in cells growing in the presence of the amino acid analogue canavanine (Prouty et al., 1975). These granules were deposits of abnormal cell proteins and they were not surrounded by any defined superficial layer. This observation was initially seen as a rather irrelevant cellular response indeed under non-physiological conditions. However, the implementation of DNA recombinant technologies and the generation of recombinant bacteria for protein production revealed protein precipitation as a common feature during the overexpression of a cloned gene (Marston, 1986). On the other hand, the resulting aggregates, named inclusion bodies (IBs), are a main obstacle

* Corresponding author. Tel.: +34-93581-2148; fax: +34-93581 2011.
E-mail address: antoni.villaverde@uab.es (A. Villaverde).

0168-1656/02/$ - see front matter
PII: S0168-1656(02)00032-9

for the production of recombinant proteins in a soluble, functional form (Buchner and Rudolph, 1991).

Aggregation of recombinant proteins is probably due to a limiting amount of chaperones when recombinant gene expression is directed at high levels (Rinas and Bailey, 1993; Thomas and Baneyx, 1996; Lorimer, 1996). Under these conditions, in which non-physiological amounts of proteins are produced, the intermolecular association of exposed hydrophobic surfaces before the protein folding can be completed can conduce to the deposition of folding intermediates (King et al., 1996), mainly when they are resistant to proteolytic degradation (Corchero et al., 1996). While bacteria are genetically prepared to respond in front of adverse natural conditions such as mild protein denaturation under high temperatures (Schroder et al., 1993; Mogk et al., 1999), no natural mechanisms favouring proper folding have been identified among the abundant stress responses triggered in recombinant cells (Parsell and Sauer, 1989; Kosinski et al., 1992; Wild et al., 1993; Arora and Pedersen, 1995; Anderson et al., 1996; Arns et al., 1998; Hareum and Bentley, 1999; Gill et al., 2000, 2001). However, the expression of some heat-shock genes encoding chaperones and proteases has been identified in response to protein overproduction (Goff and Goldberg, 1985; Allen et al., 1992; Rinas, 1996; Hoffmann and Rinas, 2000; Jurgen et al., 2000), but these stress proteins are clearly not sufficient to prevent IB formation.

Despite the obvious biological interest of bacterial models to monitor in vivo protein aggregation, the formation of IBs has been mainly regarded as a parameter to be controlled in the complex context of the bacterial cell factories (Hockney, 1994). In this regard, IB formation has not been associated to particular protein sequences, being this fact, an additional obstacle to predict the yield of a given protein in a new production process. However, successful protein engineering by directed mutation or gene fusion has resulted in enhanced solubility (Rinas et al., 1992; Mitraki and King, 1992; Chrunyk et al., 1993; LaVallie et al., 1993; Wetzel and Chrunyk, 1994; Forrer and Jaussi, 1998), especially when affecting the extent of hydrophobic regions (Murby et al., 1995). Also, naturally occurring β-lactamase mutants shown higher solubility and therefore enhanced biological activities (Sideraki et al., 2001). Finally, protein engineering can eventually enhance the yield of in vitro refolded protein from purified IBs (Rattenholl et al., 2001). However, the modification of the target protein does not offer universal strategies to prevent IB formation since any engineering approach must be adapted to particular polypeptides.

This general unpredictability of solubility but also of proteolytic stability for a given protein is then defining the trial-and-error landscape of recombinant protein production in bacteria. From a complete different approach to overcome aggregation of recombinant proteins, the co expression of chaperone-encoding genes can minimise IB formation. Although in some cases the soluble yield can be significantly enhanced, this approach renders irregular results when comparing different polypeptides (Gilbert, 1994; Wall and Pluckthun, 1995; Georgiou and Valax, 1996; Hoffmann and Rinas, in press). The important number of co-operating, folding assistant cell proteins (Schwarz et al., 1996; Thomas et al., 1997; Wickner et al., 1999), can account for this restricted success. Only a restricted number of cell proteins can be easily co-produced along with the recombinant protein, while probably, a larger set of chaperones is actually limiting for its folding. The simultaneous co-expression of several chaperone-encoding genes renders in general good results (Caspers et al., 1994; Nishihara et al., 1998, 2000), but again the improvement of solubility and the specific chaperone(s) critical to prevent aggregation vary dramatically upon the target recombinant protein (Nishihara et al., 2000). On the other hand, several environmental conditions such as high temperatures and media acidification clearly stimulate IB formation (Chalmers et al., 1990; Strandberg and Enfors, 1991). Then, the manipulation of media composition (Bowden and Georgiou, 1990; Moore et al., 1993) and the global control of culture and gene expression conditions (Schein, 1991; Weick

ert et al., 1996) can also offer possibilities to minimise IB formation.

Interestingly, IBs have been also observed as a convenient source of relatively pure polypeptide, provided efficient protein refolding protocols can be successfully applied to purified IBs (Clark, 1998). In this context, IB formation can be favoured by culturing at high temperatures (Hellebust et al., 1989), by using protease-deficient strains (Corchero et al., 1996) and through the fusion of aggregation-prone domains (Meyer and Chilkoti, 1999). Then, soluble protein can be obtained in vitro through different strategies that are under continuous methodological development (Lilie et al., 1998), and have to be adapted to particular polypeptides (Parkar et al., 2000; Ward et al., 2001). This implies the consideration of all the key parameters to be adjusted in the sequential steps of the refolding process (Clark, 2001). After global optimisation, the yield of active protein usually ranges between 10 and 50% (Arora and Khanna, 1996; Jin et al., 1994; Batas et al., 1999; Ejima et al., 1999; Babu et al., 2000; Patra et al., 2000) of the initial IB protein, but higher recoveries have also being achieved, reaching in some cases about 85% (Vandenbroeck et al., 1993; Hellman et al., 1995; Cardamone et al., 1995; Khan et al., 1998; Sunitha et al., 2000; Rehm et al., 2001).

More recently, new insights in protein folding, refolding and aggregation in bacteria, transgenic animals and cultured cell models have dramatically shifted the concept of protein aggregation regarding its biological relevance and reversibility (Goloubinoff et al., 1999; Mogk et al., 1999; Warrick et al., 1999; Krobitsch and Lindquist, 2000; Kazemi-Esfarjani and Benzer, 2000; Muchowski et al., 2000; Charmichael et al., 2000; Diamant et al., 2000; Tomoyasu et al., 2001). In the context of the arising conformational diseases (Dobson, 1999) and the presumed adaptive role of some prion-like proteins (True and Lindquist, 2000), a more detailed investigation of bacterial IBs from both structural and kinetic points of view can offer unexpected ways for a better comprehension of the mechanics of protein aggregation and de-aggregation and the role of chaperones in both processes.

2. IB location, morphology and inner organisation

IBs are commonly found in the bacterial cytoplasm, but secreted proteins like β-lactamase can also aggregate in the periplasmic space (Georgiou et al., 1986; Bowden and Georgiou, 1990). The microscopy analysis of IBs in their cellular location but also after purification, have revealed a porous structure and in general cylindrical and ovoid shapes (Bowden et al., 1991; Carrio et al., 1998, 2000). The topology of IB surface varies from rough to smooth (Bowden et al., 1991), but the limited number of morphological studies does not allow to correlate shape and surface properties with features of the recombinant protein. However, the porous structure is in agreement with density analysis data (Taylor et al., 1986) and a high level of hydration found in IB particles (Debora Foguel, personal communication). By combining scanning electron microscopy and the kinetic analysis of IB tryptic digestion, a clustered, sub-unit organisation of the IB packaged protein has been revealed (Carrio et al., 2000). This observation, the low number (one or eventually two) of IB particles inside the cells (Carrio et al., 1998) and the compatibility of a first-order kinetics to account for IB protein aggregation (Hoffmann et al., 2001) could indicate that pre-aggregates might join to a bigger seeding particle that undergo further growth. The formation of such aggregation nuclei might be then an exclusive event, in a process that could be analogous to that observed in eukaryotic cells (Kopito, 2000).

On the other hand, biological activity has been detected associated to enzyme-based IBs (Worrall and Goss, 1989; Tokatlidis et al., 1991; Carrio et al., unpublished results). Although during IB purification, contamination with minor amounts of soluble polypeptides cannot be completely excluded (Georgiou and Valax, 1999), this coincident finding in IBs formed by different types of enzymes indicates the presence of properly folded molecules with full accessibility to the solvent. This possibility is actually supported by the structural analysis of embodied protein proving a high level of secondary structure but also a native-like structure (Oberg et al., 1994). The indications of

different protein conformational states within IBs (Bowden et al., 1991; Carrio et al., 2000; Cubarsi et al., 2001) suggest a conformational flexibility of IB polypeptide chains and in the architecture of the whole IB particle. Moreover, the recent observation of IB protein refolding in actively producing recombinant bacteria, occurring simultaneously to protein aggregation (Carrio et al., 1999; Carrio and Villaverde, 2001), indicated that IBs are under a permanent reconstruction process involving both protein income and outcome. In vivo, in the absence of de novo protein synthesis, IBs are then disintegrated in a few hours resulting in a detectable increase of the soluble protein and the corresponding rise of biological activity (Carrio and Villaverde, 2001). This proves that at least an important fraction of IB protein remains suitable for refolding attempts and eventually for proteolytic attack. Then, IB formation can be seen as the result of an unbalanced equilibrium between aggregation and solubilisation occurring in the recombinant cells, that can also account for the release of protein from IBs observed in stationary-phase *E. coli* cultures (Carrio et al., 1998). IBs would then represent reservoirs of protease resistant protein that steadily accumulates, until chaperones and proteases are again available either for successful protein folding or for proteolysis (Fig. 1).

3. IB protein composition

In recombinant *E. coli* cells, the target protein is the main IB component (Rinas and Bailey, 1992; Rinas et al., 1993; Valax and Georgiou, 1993; Carrio et al., 1998). However, minor amounts of other cell proteins are also found in variable concentrations. Apart from contaminants derived from incomplete purification processes (Georgiou and Valax, 1999) and unspecific co-precipitation of unrelated polypeptides (Rinas and Bailey, 1993), folding assistant proteins could be also entrapped in IBs during aggregation of the target polypeptide. Among them, small heat shock proteins IbpA and IbpB, that are involved in protein folding and cell survival under thermal stress (Laskowska et al., 1996; Thomas and Baneyx, 1998; Kitagawa et al., 2000), are abundant (Allen et al., 1992). These chaperones, could be one among the first cellular binders to misfolded proteins acting as aggregation controllers (Veinger et al., 1998). However, it is not clear if IbpA and IbpB (and also other IB-contained chaperones), are just trapped during the aggregation process or if in addition, they might have a role in the in vivo IB construction and de-construction processes. Interestingly, IbpA and IbpB are found at different extents in IBs depending on the residual solubility of the aggregating recombinant protein (Hoffmann and Rinas, 2000), suggesting that their presence in IBs might be linked to the cell-mediated protein processing.

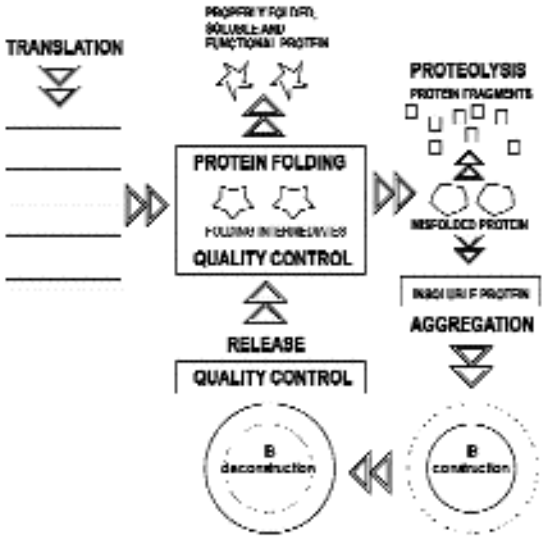

Fig. 1. Integrated model of protein folding, digestion and aggregation in recombinant bacteria. After synthesis, recombinant proteins are assisted in their folding process to reach the native conformation. The cell quality control system delivers folding-recalcitrant intermediates to proteases for destruction, but more stable proteins *aggregate* as IBs before digestion. Signals to discriminate proteins for these two processes are still unknown. Also, elements from the quality control system remove IB polypeptides for further folding attempts in an iterative process. Cell proteins might be actively involved in protein deposition and IB formation.

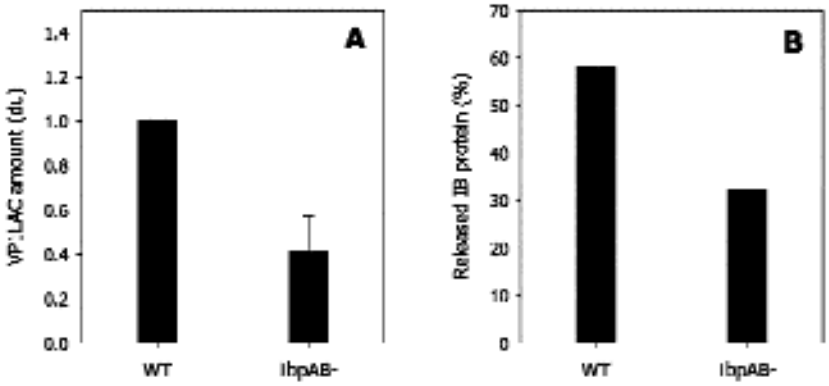

Fig. 2. Relative amounts of the recombinant protein VP1LAC, accumulated in IBs (panel A), as produced in the wild type (wt) *E. coli* MC4100 ([*r*⁻, *m*⁺]) Δ(*argF-lac*)*169 ara*D139 *deoC1 flb*D5301 *fruA*25 *rbs*R22 *rel*A1 *rpsL*150 (str^R)) and in the isogenic IbpAB⁻ mutant JGT17 (MC4100 Δ*ibp1::kan*) (Thomas and Baneyx, 1998). IBs were purified following previously described procedures (Carrio et al., 1998), from 3 h heat-induced cultures. For protein quantification, pure IBs were denatured and submitted to SDS-PAGE and further coomassie blue staining. IbpA and IbpB were determined in combination. Panel B shows the percentage of recombinant protein released from the IBs 1 h after arresting the recombinant protein synthesis with chloramphenicol as described (Carrio et al., 2001). DU are densitometric units. All the experiments were performed at least in duplicate.

4. Role of small heat shock proteins (IbpA and IbpB) in IB formation

The formation of IBs in an *E. coli* IbpAB⁻ strain has been investigated during the production of the misfolded prone β-galactosidase fusion protein VP1LAC (Corchero et al., 1996). Interestingly, IBs were actually formed in this strain, but a smaller amount of recombinant protein was found in the insoluble cell fraction when compared with the wild type cells (Fig. 2A). Since the soluble VP1LAC fraction did not increase concomitantly (not shown) and it is has been reported that proteolysis negatively influences IB formation (Corchero et al., 1996), the absence of small heat shock proteins might favour the proteolytic attack of VP1LAC before its aggregation. On the other hand, IB disintegration also takes place in the IbpAB⁻ strain, although at a lower rate than in the wild type strain (Fig. 2B). These results, while showing that Ibp proteins are positively involved but dispensable for IB formation, also prompt to suggest that they could be involved in the in vivo protein release from IBs. However, the refolding potential of IB proteins is not exclusively dependent on these chaperones.

The presence of Ibp proteins in VP1LAC IBs was then explored in the wild type strain as well as on different mutants deficient in chaperones and proteases belonging to the Clp family (namely ClpA, ClpB and ClpP). These polypeptides are closely involved in the quality control of recombinant proteins (Wickner et al., 1999). In particular ClpB, in co-operation with DnaK, might have a key role in the dissolution of protein aggregates (Ben-Zvi and Goloubinoff, 2001). The presence of both VP1LAC and IbpAB proteins in IBs is shown in Fig. 3. Note that despite the different level of insoluble VP1LAC found in these strains, IB Ibp proteins are detected in all the cases (with the exception of the IbpAB⁻ mutant). Interestingly, the ratio IbpAB/VP1LAC is constant in three of the tested strains (Fig. 3B).

However, the IbpAB/VP1LAC ratio is significantly higher in the ClpA⁻ strain (Fig. 3B). To further explore if this increase in the IB trapping ratio could be extended to other chaperones, the effect of the *clpA* mutation on the co-precipitation of DnaK and GroEL, that in vivo interact with VP1LAC (Boels et al., 1999), was studied. As in the case of Ibp proteins, the *clpA* mutation enhances the molar presence of both DnaK and GroEL in VP1LAC IBs (Fig. 4).

ClpA is an ATPase involved in proteolytic degradation (Hoskins et al., 1998; Singh et al., 2000), and it can act independently as a chaper

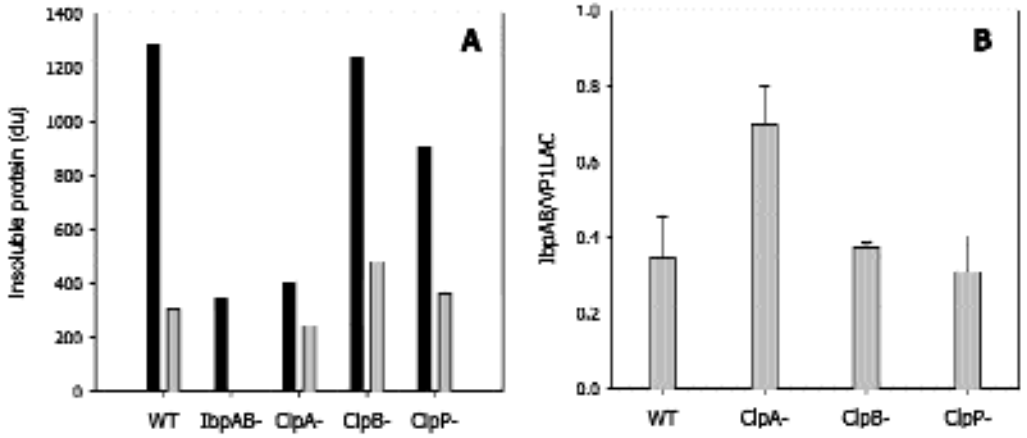

Fig. 3. Amounts of VP1LAC (black bars) and small heat shock proteins IbpA and IbpB (grey bars) present in purified IBs from 3 h induced cultures, as produced in the *E. coli* chaperone mutant strains JGT3 (MC4100 ΔclpB::*kan*), JGT4 (MC4100 ΔclpA::*kan*), JGT17 (MC4100 Δibp1::*kan*) and JGT19 (MC4100 *clpP*::*cat*) (Thomas and Baneyx, 1998) (panel A). Protein amounts were determined by PAGE and coomassie blue staining. Panel B represents the amount of IbpA and IbpB relative to the amount of the recombinant protein VP1LAC present in IBs.

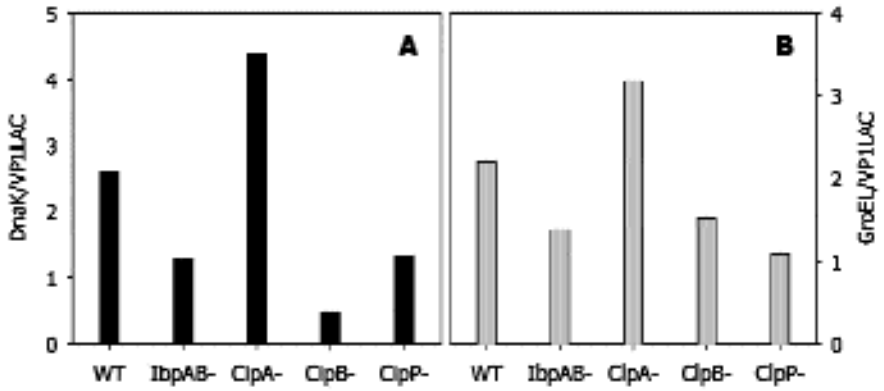

Fig. 4. Amounts of chaperones DnaK (panel A) and GroEL (panel B) related to the total recombinant protein VP1LAC present in IBs. Purified IBs were loaded for SDS-PAGE and further electroblotted onto nitro-cellulose filters to immunodetect DnaK and GroEL chaperones as described (Hoels et al., 1999).

one or co-operating with the protease ClpP (Weber-Ban et al., 1999; Reid et al., 2001). Since, the *clpP*⁻ mutation does not increase the presence of main chaperones in IBs (Figs. 3 and 4), these results suggest that the chaperone activity of ClpA may be involved in the early steps of protein folding and also that its absence could result in an enhanced binding of Ibps, DnaK and GroEL to VP1LAC, followed by their trapping onto IBs. Additional experiments would be required to test this hypothesis.

On the other hand, these results also prompt to propose that the presence of co-precipitating chaperones in bacterial IBs is importantly affected by the checking points in the quality control undergone by the recombinant protein prior to its deposition. Again, this is in the context of the IB formation as a process integrated in the whole mechanics of the cell assisted protein folding and degradation. Data presented here support that IB formation is the result of an extremely complex and dynamic process influenced by both the re

folding success and the extent of proteolysis, in which multiple cell proteins are probably involved. In addition, IBs can now be regarded as extremely plastic structures, being transient protein reservoirs rather than the dead ends of unsuccessful folding pathways.

Acknowledgements

We are indebted to U. Rinas for helpful discussions and to F. Baneyx for generously providing several *E. coli* mutant strains. This work has been supported by grant BIO2001-2443 and by Mª Francesca de Roviralta Foundation. M.M. Carrió is a recipient of a predoctoral fellowship from UAB, Barcelona, Spain.

References

Andersen, L., Yang, S., Neubauer, P., Enfors, S.-O., 1996. Impact of plasmid presence on cellular responses in fed batch cultures of *Escherichia coli*. J. Biotechnol. 46, 255–263.

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

Arie, A., Cordero, J.L., Benito, A., Carbonell, X., Viaplana, E., Villaverde, A., 1998. The expression of recombinant genes from bacteriophage lambda strong promoters triggers the SOS response in *Escherichia coli*. Biotechnol. Bioeng. 60, 551–559.

Arora, K.K., Pedersen, P.L., 1995. Glucokinase of *Escherichia coli*: induction in response to the stress of overexpressing foreign proteins. Arch. Biochem. Biophys. 319, 574–578.

Arora, D., Khanna, N., 1996. Method for increasing the yield of properly folded recombinant human gamma interferon from inclusion bodies. J. Biotechnol. 50, 127–135.

Babu, K.R., Swaminathan, S., Marten, S., Khanna, N., Rinas, U., 2000. Production of interferon-α in high cell density cultures of recombinant *Escherichia coli* and its single step purification from refolded inclusion body proteins. Appl. Microbiol. Biotechnol. 53, 655–660.

Batas, B., Schaalch, C., Chaudhuri, J.B., 1999. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. J. Biotechnol. 68, 149–158.

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

Boels, K., Carrió, M.M., Arís, A., Corchero, J.L., Villaverde, A., 1999. Distinct chaperone affinity to folding variants of homologous recombinant proteins. Biotechnol. Lett. 21, 541–546.

Bowden, G.A., Georgiou, G., 1990. Folding and aggregation of β lactamase in the periplasmic space of *Escherichia coli*. J. Biol. Chem. 265, 16760–16766.

Bowden, G.A., Paredes, A.M., Georgiou, G., 1991. Structure and morphology of protein inclusion bodies in *Escherichia coli*. Biotechnology 9, 725–730.

Buchner, J., Rudolph, R., 1991. Routes to active proteins from transformed microorganisms. Curr. Opin. Biotechnol. 2, 532–538.

Cardamone, M., Puri, N.K., Brandon, M.R., 1995. Comparing the refolding and reoxidation of recombinant porcine growth hormone from a urea denatured state and from *Escherichia coli* inclusion bodies. Biochemistry 34, 5773–5781.

Carrió, M.M., Corchero, J.L., Villaverde, A., 1998. Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. FEMS Microbiol. Lett. 169, 9–15.

Carrió, M.M., Corchero, J.L., Villaverde, A., 1999. Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. Biochim. Biophys. Acta 1434, 170–176.

Carrió, M.M., Cubarsi, R., Villaverde, A., 2000. Fine architecture of bacterial inclusion bodies. FEBS Lett. 471, 7–11.

Carrió, M.M., Villaverde, A., 2001. Protein aggregation as bacterial inclusion bodies is reversible. FEBS Lett. 489, 29–33.

Caspers, P., Stieger, M., Burn, P., 1994. Overproduction of bacterial chaperones improves the solubility of recombinant protein tyrosine kinases in *Escherichia coli*. Cell. Mol. Biol. 40, 635–644.

Chalmers, J.J., Kim, E., Telford, J.N., Wong, P.Y., Tacon, W.C., Shuler, M.L., Watson, D.B., 1990. Effects of temperature on *Escherichia coli* overproducing β lactamase or human epidermal growth factor. Appl. Environ. Microbiol. 56, 101–111.

Charmichael, J., Chatellier, J., Woolfson, A., Milstein, C., Fersht, A.R., Rubinsztein, D.C., 2000. Bacterial and yeast chaperones reduce both aggregate formation and cell death in mammalian cell models of Huntington s disease. Proc. Natl. Acad. Sci. USA 97, 9701–9705.

Chiruvyk, B.A., Evans, J., Lilquist, J., Young, P., Wetzel, R., 1993. Inclusion body formation and protein stability in sequence variants of interleukin-1 β. J. Biol. Chem. 268, 18053–18061.

Clark, E.D.B., 1998. Refolding of recombinant proteins. Curr. Opin. Biotechnol. 9, 157–163.

Clark, E.D.B., 2001. Protein refolding for industrial processes. Curr. Opin. Biotechnol. 12, 202–207.

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

Cubarsi, R., Carrió, M.M., Villaverde, A., 2001. In situ proteolytic digestion of inclusion body polypeptides occurs as a cascade process. Biochem. Biophys. Res. Commun. 282, 436–441.

Diamant, S., Ben-Zvi, A.P., Bukau, B., Goloubinoff, P., 2000. Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J. Biol. Chem. 275, 21107–21113.

Dobson, C.M., 1999. Protein misfolding, evolution and disease. Trends Biochem. Sci. 24, 329–332.

Ejima, D., Watanabe, M., Sato, Y., Date, M., Yamada, N., Takahara, Y., 1999. High yield refolding and purification process for recombinant human interleukin-6 expressed in *Escherichia coli*. Biotechnol. Bioeng. 62, 301–310.

Ferrer, P., Jaussi, R., 1998. High-level expression of soluble heterologous proteins in the cytoplasm of *Escherichia coli* by fusion to the bacteriophage λ head protein D. Gene 224, 45–52.

Georgiou, G., Valax, P., 1996. Expression of correctly folded proteins in *Escherichia coli*. Curr. Opin. Biotechnol. 7, 190–197.

Georgiou, G., Telford, J.N., Shuler, M.L., Wilson, D.B., 1986. Localization of inclusion bodies in *Escherichia coli* overproducing β-lactamase or alkaline phosphatase. Appl. Environ. Microbiol. 52, 1157–1161.

Georgiou, G., Valax, P., 1999. Isolating inclusion bodies from bacteria. Methods Enzymol. 309, 48–58.

Gilbert, H.F., 1994. Protein chaperones and protein folding. Curr. Opin. Biotechnol. 5, 534–539.

Gill, R.T., Valdes, J.J., Bentley, W.E., 2000. A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli*. Metab. Eng. 2, 178–189.

Gill, R.T., DeLisa, M.P., Valdes, J.J., Bentley, W.E., 2001. Genome analysis of high-cell-density recombinant *Escherichia coli* fermentation and cell conditioning for improved recombinant protein yield. Biotechnol. Bioeng. 72, 85–95.

Goff, S.A., Goldberg, A.L., 1985. Production of abnormal proteins in *E. coli* stimulates transcription of lon and other heat shock genes. Cell 41, 587–595.

Goloubinoff, P., Mogk, A., Zvi, A.P., Tomoyasu, T., Bukau, B., 1999. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc. Natl. Acad. Sci. USA 96, 13732–13737.

Harcum, S.W., Bentley, W.E., 1999. Heat-shock and stringent responses have overlapping protease activity in *Escherichia coli*. Implications for heterologous protein yield. Appl. Biochem. Biotechnol. 80, 23–37.

Hellebust, H., Murby, M., Abrahmsen, L., Uhlen, M., Tnfors, S.-O., 1989. Different approaches to stabilize a recombinant fusion protein. Biotechnology 7, 165–168.

Hellman, J., Lassila, P., Mantsala, P., 1995. In vitro refolding of cyclomaltodextrin glucanotransferase from cytoplasmic inclusion bodies formed upon expression in *Escherichia coli*. Protein Expr. Purif. 6, 56–62.

Hockney, R.C., 1994. Recent developments in heterologous protein production in *Escherichia coli*. Trends Biotechnol. 12, 456–463.

Hoffmann, F., Rinas, U., 2000. 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.

Hoffmann, F., Rinas, U. Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli*. Adv. Biochem. Eng. Biotechnol., in press.

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.

Hoskins, J.R., Pak, M., Maurizi, M.R., Wickner, S., 1998. The role of the ClpA chaperone in proteolysis by ClpAP. Proc. Natl. Acad. Sci. USA 95, 12135–12140.

Jin, H.L., Uddin, M.S., Huang, Y.L., Teo, W.K., 1994. Purification and renaturation of recombinant human lymphotoxin (tumour necrosis factor β) expressed in *Escherichia coli* as inclusion bodies. J. Chem. Technol. Biotechnol. 59, 61–72.

Jurgen, B., Lin, H.Y., Riemschneider, S., Scharf, C., Neubauer, P., Schmid, R., Hecker, M., Schweder, T., 2000. 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.

Kazemi-Esfarjani, P., Benzer, S., 2000. Genetic suppression of polyglutamine toxicity in *Drosophila*. Science 287, 1837–1840.

Khan, R.H., Rao, K.B., Eshwari, A.N., Totey, S.M., Panda, A.K., 1998. Solubilization of recombinant ovine growth hormone with retention of native-like secondary structure and its refolding from the inclusion bodies of *Escherichia coli*. Biotechnol. Prog. 14, 722–728.

Kang, J., Hanse-Pattugell, C., Robinson, S., Speed, M., Mitraki, A., 1996. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. FASEB J. 10, 57–66.

Kitagawa, M., Matsumura, Y., Tsuchido, T., 2000. Small heat shock proteins, IbpA and IbpB, are involved in resistances to heat and superoxide stresses in *Escherichia coli*. FEMS Microbiol. Lett. 184, 165–171.

Kopito, R.R., 2000. Aggresomes, inclusion bodies and protein. Trends Cell. Biol. 10, 524–530.

Kosinska, M.J., Rinas, U., Bailey, J.E., 1992. Proteolytic response to the expression of an abnormal β-galactosidase in *Escherichia coli*. Appl. Microbiol. Biotechnol. 37, 335–341.

Krobitsch, S., Lindquist, S., 2000. Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. Proc. Natl. Acad. Sci. USA 97, 1589–1594.

Laskowska, E., Kuczynska-Wisnik, D., Skorko-Glonek, J., Taylor, A., 1996. Degradation by proteases Lon, Clp and HtrA, of *Escherichia coli* proteins aggregated in vivo by heat shock; HtrA protease action in vivo and in vitro. Mol. Microbiol. 22, 555–571.

LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., McCoy, J.M., 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Biotechnology 11, 187–193.

Lilie, H., Schwarz, E., Rudolph, R., 1998. Advances in refolding of proteins produced in *E. coli*. Curr. Opin. Biotechnol. 9, 497–501.

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

Marston, F.A., 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. Biochem. J. 240, 1–12.

Meyer, D.I., Chilkoti, A., 1999. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. Nat. Biotechnol. 17, 1112–1115.

Mitraki, A., King, J., 1992. Amino acid substitutions influencing intracellular protein folding pathways. FEBS Lett. 307, 20–25.

Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H., Bukau, B., 1999. Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. EMBO J. 18, 6934–6949.

Moore, J.T., Uppal, A., Maley, F., Maley, G.F., 1993. Overcoming inclusion body formation in a high level expression system. Protein Expr. Purif. 4, 160–169.

Muchowski, P.J., Schaffar, G., Sittler, A., Wanker, E.E., Hayer-Hartl, M.K., Hartl, F.U., 2000. Hsp70 and hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. Proc. Natl. Acad. Sci. USA 97, 7841–7846.

Murby, M., Samuelsson, E., Nguyen, T.N., Uhlen, M., Stahl, S., 1995. Hydrophobicity engineering to increase solubility and stability of a recombinant protein from respiratory syncytial virus. Eur. J. Biochem. 230, 38–44.

Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., Yura, T., 1998. 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.

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

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

Parkar, A.A., Stow, M.D., Smith, K., Panicker, A.K., Gmilioteau, J.P., Jupp, R., Crowe, S.J., 2000. Large scale expression, refolding, and purification of the catalytic domain of human macrophage metalloelastase (MMP-12) in *Escherichia coli*. Protein Expr. Purif. 20, 152–161.

Parsell, D.A., Sauer, R.T., 1989. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. J. Biol. Chem. 264, 7590–7595.

Patra, A.K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L.C., Panda, A.K., 2000. Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from *Escherichia coli*. Protein Expr. Purif. 18, 182–192.

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

Rattenholl, A., Lilie, H., Grossmann, A., Stern, A., Schwarz, E., Rudolph, R., 2001. The pro sequence facilitates folding of human nerve growth factor from *Escherichia coli* inclusion bodies. Eur. J. Biochem. 268, 3296–3303.

Rehm, B.H., Qi, Q., Beermann, B.R., Hinz, H.J., Steinbuchel, A., 2001. 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.

Reid, B.G., Fenton, W.A., Horwich, A.L., Weber Ban, E.U., 2001. ClpA mediates directional translocation of substrate proteins into the ClpP protease. Proc. Natl. Acad. Sci. USA 98, 3768–3772.

Rinas, U., 1996. 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.

Rinas, U., Bailey, J.E., 1992. Protein compositional analysis of inclusion bodies produced in recombinant *Escherichia coli*. Appl. Microbiol. Biotechnol. 37, 609–614.

Rinas, U., Bailey, J.E., 1993. Overexpression of bacterial hemoglobin causes incorporation of pre-β-lactamase into cytoplasmic inclusion bodies. Appl. Environ. Microbiol. 59, 561–566.

Rinas, U., Tsai, L.B., Lyons, D., Fox, G.M., Stearns, G., Fleischko, J., Fenton, D., Bailey, J.E., 1992. Cysteine to serine substitutions in basic fibroblast growth factor: effect on inclusion body formation and proteolytic susceptibility during in vitro refolding. Biotechnology 10, 435–440.

Rinas, U., Boone, T.C., Bailey, J.E., 1993. Characterization of inclusion bodies in recombinant *Escherichia coli* producing high levels of porcine somatotropin. J. Biotechnol. 28, 313–320.

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

Schroder, H., Langer, T., Hartl, F.U., Bukau, B., 1993. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. 12, 4137–4144.

Sideraki, V., Huang, W., Palzkill, T., Gilbert, H.F., 2001. A secondary drug resistance mutation of TEM-1 β-lactamase that suppresses misfolding and aggregation. Proc. Natl. Acad. Sci. USA 98, 283–288.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., Maurizi, M.R., 2000. Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. USA 97, 8898–8903.

Strandberg, L., Enfors, S.O., 1991. Factors influencing inclusion body formation in the production of a fused protein

in *Escherichia coli*. Appl. Environ. Microbiol. 57, 1669–1674.

Schwarz, E., Lilie, H., Rudolph, R., 1996. The effect of molecular chaperones on in vivo and in vitro folding processes. Biol. Chem. 377, 411–416.

Sunitha, K., Chung, B.H., Jang, K.H., Song, K.B., Kim, C.H., Rhee, S.K., 2000. Refolding and purification of *Zymomonas mobilis* levansucrase produced as inclusion bodies in fed-batch culture of recombinant *Escherichia coli*. Protein Expr. Purif. 18, 388–394.

Taylor, G., Hoare, M., Gray, D.R., Marston, F.A.O., 1986. Size and density of protein inclusion bodies. Biotechnology 4, 553–557.

Thomas, J.G., Baneyx, F., 1996. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing Heat-shock proteins. J. Biol. Chem. 271, 11141–11147.

Thomas, J.G., 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.

Thomas, J.G., Ayling, A., Baneyx, F., 1997. 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.

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

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

True, H.L., Lindquist, S.L., 2000. A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature 407, 477–483.

Valax, P., Georgiou, G., 1993. Molecular characterization of β-lactamase inclusion bodies produced in *Escherichia coli*. 1. Comp. Biotechnol. Prog. 9, 539–547.

Vandenbroeck, K., Martens, E., D'Andrea, S., Billiau, A., 1993. Refolding and single-step purification of porcine interferon-gamma from *Escherichia coli* inclusion bodies. Conditions for reconstitution of dimeric IFN-γ. Eur. J. Biochem. 215, 481–486.

Veinger, L., Diamant, S., Buchner, J., Goloubinoff, P., 1998. 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.

Wall, J.G., Pluckthun, A., 1995. Effects of overexpressing folding modulators on the in vivo folding of heterologous proteins in *Escherichia coli*. Curr. Opin. Biotechnol. 6, 507–516.

Waud, R.J., de Oliveira, A.H., Botteldo, R.K., Rosa, J.C., Faca, V.M., Greene, L.J., 2001. Refolding and purification of Bothropstoxin I, a Lys49 phospholipase A2 homologue, expressed as inclusion bodies in *Escherichia coli*. Protein Expr. Purif. 21, 134–140.

Warrick, J., Chan, H.Y.E., Gray-Board, G.L., Chai, Y.H., Paulson, H.L., Bonini, N.M., 1999. Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. Nat. Genet. 19, 148–154.

Weber-Ban, E.U., Reid, B.G., Miranker, A.D., Horwich, A.L., 1999. Global unfolding of a substrate protein by the hsp100 chaperone ClpA. Nature 401, 90–93.

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

Wetzel, R., Chrunyk, B.A., 1994. Inclusion body formation by interleukin 1 β depends on the thermal sensitivity of a folding intermediate. FEBS Lett. 350, 245–248.

Wickner, S., Maurizi, M.R., Gottesman, S., 1999. Posttranslational quality control: folding, refolding, and degrading proteins. Science 286, 1888–1894.

Wild, J., Walter, W.A., Gross, C.A., Altman, E., 1993. Accumulation of secretory protein precursors in *Escherichia coli* induces the heat shock response. J. Bacteriol. 175, 3992–3997.

Worrall, D.M., Goss, N.H., 1989. The formation of biologically active β-galactosidase inclusion bodies in *Escherichia coli*. Aust. J. Biotechnol. 3, 28–32.