Limited in Vivo Proteolysis of Aggregated Proteins

J. L. Corchero,* R. Cubarsí,† S.-O. Enfors,‡ and A. Villaverde*¹

*Institut de Biologia Fonamental and Departament de Geneètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain; †Departament de Matemàtica Aplicada i Telemàtica, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain; and ‡Department of Biochemistry and Biotechnology, Royal Institute of Technology, S-100 44, Stockholm, Sweden

Received July 2, 1997

Degradation pathways of insoluble proteins have been analyzed in Escherichia coli by using a N-terminal β-galactosidase fusion protein (VP1LAC) that aggregates immediately after its synthesis. In recombinant E. coli cells, lower molecular mass products, antigenically related to the entire fusion, accumulate together with the entire fusion. In absence of protein synthesis, the insoluble intact protein declines, suggesting that degradation of the recombinant protein also affects aggregated protein. Time course analysis of both soluble and insoluble cell fractions has revealed a limited proteolysis of the insoluble protein that removes the heterologous domain and permits the resulting β-galactosidase fragments to refold and solubilize. Further extensive degradation occurs exclusively on soluble protein. The restricted proteolysis of misfolded, insoluble protein is the initiating event of a subsequent degradative pathway in which rate-limiting steps permit the accumulation of stable degradative intermediates.

In vivo proteolysis is an essential cell mechanism that regulates the intracellular pool of functional proteins and eliminates abnormal polypeptides (1-3). Many recombinant proteins produced in bacteria fail to adopt their native conformation, especially if deriving from mammalian organisms or viruses. The synthesis of such proteins can induce the expression of heat shock genes encoding, among others, proteases and molecular chaperones (4-6). The folding assistant activities of chaperones and related proteins exhibited only a limited success on recombinant protein refolding (7), specially when the target gene is overexpressed at unusually high rates. These misfolded proteins can either aggregate or be degraded, in both cases reducing the yield of active biomolecules. Insoluble proteins are more resistant to proteases than their counterpart soluble forms, and are usually found entrapped into inclusion bodies (8). In fact, the adoption of growth or expression conditions favouring aggregation, such as high temperatures (9), has been proposed as an useful strategy to protect proteins from proteolytic degradation and consequently to increase the yield (10). However, the analysis of inclusion bodies has often revealed the presence of protein fragments resulting from proteolysis (11, 12) that could either aggregate after their release, or be produced by a limited proteolysis activity within the same aggregates. Despite the increasing knowledge of protein folding and aggregation mechanisms (13, 14, 15), there is still an insufficient basis to completely understand the degradation dynamics of misfolding-prone proteins. Therefore, we have here explored the in vivo degradative pathways of a protease-sensitive β-galactosidase fusion protein (VP1LAC) mainly present in the cell cytoplasm in insoluble form (12). Our results prompt us to propose a model for degradation of aggregated protein based on two different activities of the proteolytic cell machinery. This model predicts a limited, two-step proteolysis on specific sites of the insoluble protein, a concomitant solubilization of the resulting fragments and an extensive degradation of these soluble, derived products in a cascade process in which rate-limiting steps can be identified.

MATERIAL AND METHODS

Plasmid and bacterial strain. Plasmid pJ VP1LAC (12), a pJ CO46 derivative (16), expresses the VP1 gene of foot-and-mouth disease
virus (FMDV) fused to the 5' end of lacZ gene, in a p18-C1857 expression system. The resulting VP1LAC protein has an estimated molecular mass of 140 kDa and is recognized by sera and monodonal antibodies against both the homologous and the heterologus segments (12, 17). The E. coli strain MC1061, K12 hsdR2 mcrB araD139 Δ(araABC-leu)7697 Δ(lacX74 galU galK rpsL thi strA, (18) was used in all the experiments.

Western blot analysis of cell extracts. To study the half live of VP1LAC, the following procedures were applied. Cells were grown in shaker flasks at 28°C in LB rich medium (19) in presence of 25 μg/ml streptomycin and 100 μg/ml ampicillin. When the cultures reached an OD550 of about 1 unit, induction of recombinant gene expression was triggered by transferring the cultures to a prewarmed bath at 42°C. After two hours of thermal induction, 100 μg/ml chloramphenicol was added to the culminal tumol at 28°C in synthesis and the flasks were immediately transferred to 37°C. At different times, samples were taken, centrifugated at 4°C for 5 min at 12,000, and pellets stored at −20°C until further analysis. To discriminate between stabilities of soluble and insoluble protein, these cell pellets were resuspended in PBS, sonicated for 10 min in 0.5 sec pulses at 50 W, and centrifugated at 12,000 rpm for 5 min to obtain soluble and insoluble fractions. Protein separation was performed by SDS-PAGE in 3.5% stacking and 7% separation gel according to Laemmli (20). After electophoresis, gels were electrobotted onto nitrocellulose membranes. VP1LAC protein and its degradation products were immunodetected with a rabbit serum elicited against the native enzyme. A monoclonal antibody against a VP1 B-cell epitope, 3E5, reacts with a B-cell epitope within the antigenic site A of VP1 (amino acids 138 to 150) (21), and it was also employed for some experiments. Finally, membranes were submitted to densitometric analysis.

Pulse labelling and protein sequencing. Minimal medium M9 (19) was employed to support cell growth. Streptomycin and ampicillin were added at 25 and 100 μg/ml respectively, glucose at 0.2% (w/v) and thiamin at 1 μg/ml. A mixture containing all the amino acids except methionine and cysteine was also added to 25 μg/ml each. Overnight cultures grown in minimal medium at 28°C and 250 rpm were used to inoculate 10 ml of fresh medium. Fresh cultures were left under the same conditions until the OD550 reached values around 0.3 units. Then, they were transferred to a prewarmed bath at 42°C, and after 45 min, 13 μl of a 35S-methionine-cysteine mixture (14.28 μCi/μl) were added. Incorporation of labelled amino acids was stopped 3 min latter with an excess of non labelled methionine and cysteine followed by the immediate transfer to 37°C. Once at this temperature, samples of ml were taken at different intervals of time. For the obtention of soluble and insoluble fractions, cells were lysed and fractionated by a lysozyme-NP40 standard procedure (19). Finally, samples were submitted to SDS-PAGE, gels dried and further exposed for autoradiography. The resulting bands were analyzed densitometrically. To estimate molar amounts of protein, densitometric units of each of the relevant products were corrected by the number of methionines and cysteines present in them. These arbitrary molar values were used for further mathematical modeling.

Amino terminal microsequencing of specific products was done in a Beckman LFS 3000 protein sequencer after SDS-PAGE and blotting.

RESULTS

Proteolysis of VP1LAC. The disappearing of VP1LAC protein in viable E. coli cells, after arrest of protein synthesis is shown in Fig. 1A. Three major bands, immunoreactive with anti-β-galactosidase serum, are detected in cell extracts. The apparent molecular masses of two of them were 141 and 118 kDa, compatible with those of the VP1LAC and β-galactosidase (140 and 116 kDa, respectively). A faster migrating product of 93 kDa is also evident, and it could be a degradation intermediate of 90 kDa found during degradation of abnormal β-galactosidase mutant proteins, called fragment B (22). This peptide is released during the in vivo endoproteolytic attack on the C-terminal half of β-galactosidase (22), and because of its apparent stability it is a common rate-limiting step in the complete degradation of the enzyme. Only the entire fusion is recognized by anti VP1 3E5 monoclonal antibody (not shown). This suggests that the β-galactosidase-like segment could be the larger product of an endoproteolytic cleavage within or near the linker region. In agreement with these assumptions, N-terminal microsequencing of the β-galactosidase-like products reveals a major sequence (19PGVTQL24) corresponding to the N-terminus of β-galactosidase as encoded by pJ VP1LAC.

In vivo half life of VP1LAC protein. In vivo degradation of produced VP1LACV protein was explored in absence of protein synthesis. Declining of the entire VP1LAC fusion and the concomitant increase of -galactosidase-like product confirm active proteolysis, and reveal a half life for VP1LAC of 138±35 min. Since it has been shown that an important fraction of this protein can accumulate in inclusion bodies (12), we analyzed separately the evolution of the fusion protein in both insoluble and soluble fractions (Fig. 1B and C respectively) to evaluate the protective effect of aggrega-
tion on the proteolysis of VP1LAC. By comparing both Western blots, it is evident an asymmetric distribution of the VP1LAC forms, the entire fusion being predominant in the insoluble fraction, whereas degradation products are mainly observed in the soluble fraction. Signs of heterogeneity in the $\beta$-galactosidase-like species are eventually observed but they were not evident in all the analyzed gels. The time-dependent disappearance of the full-length VP1LAC allowed us to estimate the half-life of this protein in the insoluble and soluble fractions in 107±32 and 180±28 min respectively.

Proteolytic sensitivity of soluble and insoluble forms of VP1LAC. Data presented above show that the insoluble fraction of VP1LAC decreases in absence of protein synthesis. In fact, half life of insoluble VP1LAC is significantly lower than that of soluble protein. This unexpected observation could be explained by two different, non-exclusive events, either refolding and solubilization of aggregated protein permitting its further degradation, or a direct proteolysis of the insoluble polypeptide. To explore these possibilities, a pulse labeling experiment was done to determine the fate of aggregated proteins in absence of further recombinant protein synthesis. In Fig. 2, is shown the evolution of labelled proteins in each fraction. As depicted at the bottom, there is a dramatic decrease in the pool of insoluble VP1LAC which is not reflected by any increase in the amount of its soluble partner. However, the amount of degradation products ($\beta$-galactosidase-like and B fragments) varies significantly in both fractions throughout the experiment, suggesting a flow of protein between the possible states that could involve solubilization, aggregation and proteolytic events. On the other hand, about 30% of the initial labelled molecules are not detected in bands after 120 min of incubation, indicative of complete degradation of this percentage of VP1LAC molecules.

To identify specific pathways of degradation, we have identified the 7 possible states in which we found VP1LAC protein or its degradation products (Fig. 3A). Molecular mass decreasing soluble fractions are numbered as 1, 3, 5 and the insoluble versions with the same electrophoretic mobility as 2, 4 and 6. Completely degraded protein, not detected in the autoradiography, is state number 7. The relative amount of protein that could move from a state to another is represented by the transition coefficients $p_1, q_1, r_1, s_1, \ldots, s_3$, assumed to be constant and positive throughout the process.

FIG. 3. (A) A scheme of biologically possible protein flows between all of the states of VP1LAC protein and its stable degradation products. Transition coefficients are indicated. (B) Degradative pathways of aggregated VP1LAC in Escherichia coli derived from the significant transition coefficient values shown in Table 1. According to the model, two consecutive site-specific cleavages occur only on the misfolded form of the protein and it renders $\beta$-galactosidase-like and B fragments, both lacking the FMDV domain. These truncated forms of VP1LAC can refold and enter the soluble fraction. The soluble, 90 kDa B fragment is the only intermediate for the complete degradation of the protein. Soluble $\beta$-galactosidase-like protein is a stable product that accumulates into the cells. Even assuming a certain degradation of fragment B, it represents a critical rate-limiting step. Soluble VP1LAC is proteolytically stable.
with

\[
A = \begin{bmatrix}
-(q_1 + r_1) & p_1 & 0 & 0 & 0 & 0 \\
q_1 & -(p_1 + s_1) & 0 & 0 & 0 & 0 \\
r_1 & 0 & -(q_2 + r_2) & p_2 & 0 & 0 \\
0 & s_1 & q_2 & -r_2 & 0 & 0 \\
0 & 0 & r_2 & 0 & s_2 & q_3 \\
0 & 0 & 0 & 0 & r_3 & s_3 \\
\end{bmatrix}
\]

where the prima represents the time derivative. To spot the preferential paths for protein flow, the elements of matrix A must be estimated.

The molar amount of protein in each state is represented in Fig. 2 for all the sample times \((t_1, t_2, \ldots, t_6)\). Thus, it is possible to evaluate protein variation \(N'(t)\) for a specific moment \(t_i\), according to

\[
N'(t_i) = \frac{N(t_{i+1}) - N(t_i)}{t_{i+1} - t_i}; \quad i = 1, \ldots, 5
\]

Since protein flow has a slow and smooth time evolution, it is enough to use a first order approximation for the derivative \(N'\). Therefore, according to (Eq 1), the subsequent linear system is built involving all the data from different times:

\[
[N'(t_1), \ldots, N'(t_5)] = A[N(t_1), \ldots, N(t_5)]
\]

This overdetermined system comprises a set of 35 equations involving the 12 unknown constants, and the least squares solution can be found. Solutions are shown in Table 1, where the constants below the 5% of the most significative value have been neglected. Note that \(q_1\) and \(r_1\) remain undetermined since no VP1LAC protein is detected in the state 1 throughout the pulse-labelling experiment (Fig. 2). The results obtained permit to define significant values of transition coefficients and therefore the actual protein flows between all the possible states of VP1LAC and its degradation products (Fig. 3B). Interestingly, the obtained results show a 2-step limited proteolysis that occurs on the insoluble protein but not on the soluble versions, whereas the complete degradation of the resulting 90 kDa protein is exclusive of soluble polypeptide. This model suggests that insoluble protein is not completely resistant to in vivo proteolysis but that it can be cleaved by specific, sequentially activated target sites.

**DISCUSSION**

Two major stable degradation products of VP1LAC fusion protein, of about 118 and 93 kDa, are observed in crude cell extracts (Fig. 1). The electrophoretic mobility of these segments is compatible with those of entire \(\beta\)-galactosidase and B fragment respectively, this latter being an intermediate in the cytoplasmatic degradation of abnormal \(\beta\)-galactosidases (22-24). The absence of other immunoreactive bands indicates that the endoproteolytic cleavages that generate these products occur sequentially, firstly removing the viral domain from a site near or within the linker region, and subsequently activating a cryptic cleavage site, probably within the carboxy moiety of the enzyme, that generates the B fragment. The total amount of recombinant protein declines in the cell cytoplasm in absence of further protein synthesis (Fig. 1A), suggesting that apart from limited proteolysis by these two sites, an extensive degradation of the fusion protein takes also place.

It has been previously shown that more that 90% of the overproduced VP1LAC is found in the insoluble cell fraction (12). This fact prompted us to investigate the degradative pathways of this aggregated protein, and to determine if soluble intermediates are required. By exploring the stability of both soluble and insoluble fractions of full-length VP1LAC, half lives of 180 and 107 min respectively have been obtained, suggesting a higher proteolytic stability of the soluble protein. However, a major presence of intact, full-length form of this protein is clearly observed in the insoluble fraction in Western blot, strongly suggesting proteolytic protection of the aggregated protein (Fig. 1). This apparent contradiction seems to indicate solubilization of aggre-
gated protein in absence of protein synthesis, combined with a proteolytic activity that seems to occur mainly on soluble protein. Although it has been reported that protein molecules entrapped into inclusion bodies are hardly disgregatted (7), aggregated VP1LAC do not form clear inclusion bodies in MC1061, although they are manifest in BL21 Lon− cells (12), indicating that the solubility of this protein is delicately controlled and not only a question of protein structure but also depending on the cytoplasmic environment. The fate of recently synthesized VP1LAC protein was traced by pulse-labelling, again in absence of protein synthesis (Fig. 2). Recently synthesised VP1LAC protein is found in the insoluble fraction but rapidly disappears, concomitant to an increase of soluble degradation fragments. The modelling of this dynamics according to a linear differential system suggests that the limited proteolysis occurs exclusively in the insoluble fraction (Fig. 3), releasing soluble β-galactosidase-like and B fragments. Whereas the 118 kDa protein is proteolytically stable, soluble B fragment is degraded, probably by rapid exoproteolysis. An explanation for this degradative pathway can be found if assuming that the initial aggregation steps of VP1LAC imply restricted interactions between molecules, that could allow partial solvent exposure of fusion proteins and their restricted digestion by specific sites. These interactions could occur preferentially between VP1 partners, that contain highly hydrophobic regions necessary for the construction of viral capsids (25). A partial aggregation of VP1 domains could render solvent-exposed regions within β-galactosidase and the linker region allowing protease activities. In this line, it has recently been described that coaggregation is a specific event involving hydrophobic patches of homologous molecules (26). Partially trapped proteins could exhibit local conformational alterations activating solvent-exposed protease target sites (27, 28), that in the case of VP1LAC, would result in the release of either soluble or easily solubilizable VP1-free fragments. After the release of β-galactosidase-like product from the aggregates and its refolding, no further limited cleavage seems to be done over this protein, leading to its accumulation in the soluble fraction (Fig. 1, 3). On the contrary, if a second endo proteolytic step occurs on the still misfolded protein, the resulting B fragment is solubilized and suffers an extensive proteolytic attack that can account for the complete degradation of an important fraction of recombinant protein. This hypothesis is in agreement with the high stability displayed by native β-galactosidase, that could also be shared by the 118 kDa product, and the extreme proteolytic sensitivity of the B fragment (29). The small amount of non aggregated VP1LAC protein remains stable (Fig. 1C).

The results presented in this work strongly suggest that the initial steps of protein coaggregation can activate protease target sites in the misfolded protein, generating a transient susceptibility to limited proteolysis that could be further lost in dense aggregates. The release of partially degraded protein fragments from aggregating intermediates allows an extensive proteolysis from which the completely aggregated protein remains efficiently protected.

ACKNOWLEDGMENTS

We are indebted to E. Brocchi for providing 3E5 mAb. We also thank J. Checa, V. Ferreres, F. Canals, and F. García for helpful assistance. Work in Barcelona has been financed by CICYT (Grant B1095-0801) and partially by CUR of Catalonia (Grant 1995SGHR 00376) and Fundación Mª Francesca Rovira. Work in Stockholm has been supported by the Swedish Research Council for Engineering Sciences (TFR). J. L. Corchero received a fellowship from MEC, Spain.

REFERENCES