Volume 282, number 1, 205–208 FEBS 09657 April 1991

ADONIS 001457939100364L

High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D

Kostas Tokatlidis, Prasad Dhurjati*, Jacqueline Millet, Pierre Béguin and Jean-Paul Aubert

Unité de Physiologie Cellulaire, Département des Biotechnologies and URA CNRS 1300, Institut Pasteur, 28, rue du Dr. Roux, 75724 Paris Cedex 15, France

Received 26 January 1991; revised version received 27 February 1991

The formation of cytoplasmic inclusion bodies by *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D (EGD) was investigated. EGD was found in inclusion bodies as a 68 kDa form, whereas the size of the cytoplasmic form was 65 kDa. Upon solubilization with urea followed by dialysis, the 68 kDa form was converted to the 65 kDa species. Proteolysis occurred within the COOH-terminal, reiterated region of the 68 kDa form, which is conserved among most *C. thermocellum* endoglucanases, but is not required for catalytic activity. The specific activity of the enzyme embedded in inclusion bodies was close to that of the purified protein. Thus, inclusion body formation does not involve denaturation of the catalytic domain of EGD, but, more likely, the participation of the reiterated, conserved region in intermolecular interactions.

Inclusion body; *Clostridium thermocellum*, Endoglucanase D; Protein folding

1. INTRODUCTION

A number of recombinant gene products synthesized in *Escherichia coli* precipitate and form insoluble cytoplasmic inclusion bodies [1–5], which appear as highly refractile, amorphous granules, not surrounded by a membrane [6–8]. Inclusion bodies are thought to be formed from in vivo protein folding intermediates leading to improperly folded, insoluble forms of the product [9]. Solubilization of the precipitated protein usually requires treatment of inclusion bodies with a chaotropic agent, such as urea [3,5], and recovery of the desired product in a native form is often difficult.

In this respect, *Clostridium thermocellum* endoglucanase D (EGD) stands out as an exception. Fusion of the *celD* gene encoding EGD with the *lacZ'* gene of plasmid pUC8 results in levels of EGD expression reaching up to 15% of total protein, and about 20–30% of the enzyme is found in inclusion bodies. In contrast to the majority of inclusion body proteins, EGD extracted with urea from the insoluble fraction is highly active; furthermore, the protein crystallizes readily, providing additional evidence that it has a highly homogeneous and presumably native conformation [8]. To investigate the basis of such a behaviour, the properties of EGD present in inclusion bodies were characterized prior to urea solubilization and dialysis.

Correspondence address: P. Béguin, Unité de Physiologie Cellulaire, Département des Biotechnologies and URA CNRS 1300, Institut Pasteur, 28, rue du Dr. Roux, 75724 Paris Cedex 15, France. Fax: (33) (1) 45 68 87 90

**Permanent address:* Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, USA

Abbreviations: EG, endoglucanase (E.C. 3.2.1.4); PAGE, polyacrylamide gel electrophoresis; *p*-NPC, *p*-nitrophenyl-β-D-cellobioside; *p*-NPCase, *p*-nitrophenyl-β-D-cellobiosidase; SDS, sodium dodecyl sulfate

2. MATERIALS AND METHODS

2.1. *Preparation of EGD fractions*

E. coli JM101(pCT603) [8] was grown at 37°C in Luria broth [10] containing 100 μg/ml carbenicillin. Cells from 1 liter overnight cultures were harvested in stationary phase, resuspended in 100 ml 50 mM potassium phosphate-12.5 mM citrate buffer, pH 6.3 and disrupted in a French pressure cell at 90 MPa (13 000 psi). The broken cell suspension was fractionated by centrifugation at 27 000 × g_{max} for 20 min into a soluble, cytoplasmic fraction and an insoluble pellet containing inclusion bodies. The insoluble fraction was washed twice by resuspending in 0.15 M NaCl followed by centrifugation for 20 min at 27 000 × g_{max}. The washed pellet was treated with 0.1 M Tris-HCl pH 8.5 containing 5 M urea [8]. The solubilized protein was dialyzed for at least 30 h (4 buffer changes) against 20 mM Tris-HCl, pH 7.7, and purified as described previously [8].

2.2. *Antisera*

Anti-EGA antiserum was obtained from rabbits immunized against endoglucanase A (EGA) purified from *C. thermocellum* culture supernatant [11]. Anti-EGD antiserum was raised in rabbits immunized against EGD purified from inclusion bodies produced by *E. coli* JM101(pCT603) [8].

2.3. *Western blotting*

Samples were heated at 100°C for 5 min in 2% SDS/5% β-mercaptoethanol sample buffer, followed by SDS-PAGE [12]. Western blotting was performed according to Towbin et al. [13]. The anti-EGD antiserum was saturated with a crude extract of *E. coli* TG1(pUC19) [14] to prevent adsorption to non-specific bands.

2.4. *Enzyme assay*

The release of *p*-nitrophenol from *p*-nitrophenyl-β-D-cellobioside

(*p*-NPC) was followed according to Chauvaux et al. [15]. One unit of *p*-NPCase activity hydrolyzes one μmol *p*-NPC per min. The specific activity of EGD purified from inclusion bodies was calculated after measuring the protein concentration using the Coomassie blue binding assay [16], with bovine serum albumin as a standard.

To obtain the specific activity of EGD present in inclusion bodies and in crude cytoplasmic supernatant, the *p*-NPCase activity present in each fraction was divided by the amount of EGD antigen, estimated by cutting out and counting radioactive bands from a Western blot. Activity and antigen assays were run in quadruplicate and compared with similar assays, run in duplicate with purified, 65 kDa EGD. Errors in the determination of specific activity were calculated by adding the relative errors of the activity assay and of the Western blot quantification of EGD, as determined from the standard deviation of each set of values.

3. RESULTS

3.1. *EGD is present in inclusion bodies as a 68 kDa form.*

EGD present in inclusion bodies was analyzed after dissolving in SDS-mercaptoethanol sample buffer [12] at 100°C (Fig. 1). A major species with an M_r of 68000 kDa was detected by Western blotting using anti-EGD antiserum (lane B). The same species was present after treatment of inclusion bodies with 5 M urea (lanes C-D). Upon dialysis, however, the 68 kDa form was truncated to 65 kDa (lane E), presumably owing to the reactivation of a protease present in the extract. A similar observation was made by Babbit et al. [17], who found that, upon dissolving of aggregates formed in *E. coli* followed by refolding, *Torpedo californica* creatine kinase and bovine pancreatic trypsin inhibitor are degraded by a protease associated with the inclusion body fraction. The 65 kDa form, corresponding to the species of EGD described previously [8], was purified by three precipitation steps using streptomycin sulfate, heat treatment and ammonium sulfate [8] (lanes E-H, and data not shown). Only the 65 kDa species of EGD was detected in the soluble cytoplasmic fraction (lane A).

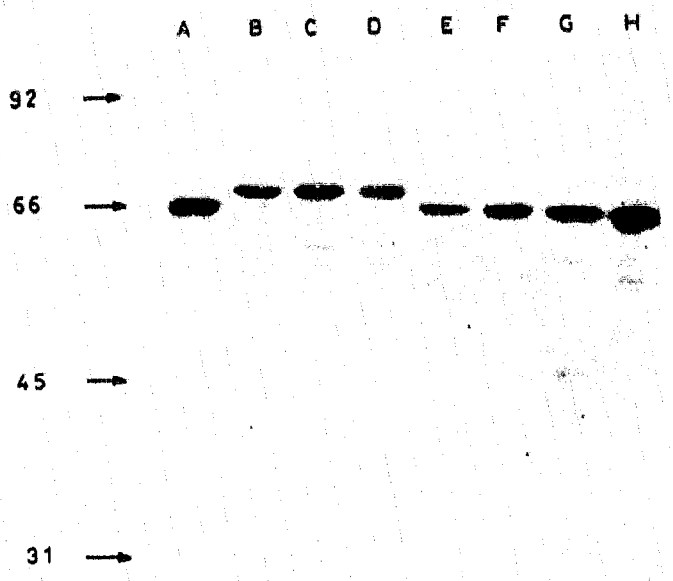

Fig. 1. Western blotting analysis of various fractions of EGD isolated from *E. coli* JM101(pCT603). The blot was probed with anti-EGD antiserum. A, cytoplasmic supernatant; B, inclusion bodies washed twice with 0.15 M NaCl; C, inclusion bodies resuspended in 5 M urea before centrifugation; D-H, fractions from the purification of EGD from inclusion bodies [8]; D, supernatant of urea-treated inclusion bodies; E, dialyzed urea supernatant; F, supernatant after streptomycin sulfate treatment; G, supernatant after heat treatment; H, ammonium sulfate precipitate. 166 nU of enzyme were loaded in each lane.

3.2. *Conversion of the 68 kDa form into the 65 kDa form affects the conserved, reiterated domain located at the COOH terminus of EGD.*

The COOH-terminal region encoded by *celD* contains two highly homologous segments of 23 amino acids each [18] (Fig. 2). It is strongly conserved in a xylanase and in most endoglucanases of *C. thermocellum* whose sequence is known to date, but it is not directly involved in catalytic activity [14,15,19].

Western blotting using antibodies directed against *C. thermocellum* endoglucanase A (EGA) indicated that conversion of the 68 kDa polypeptide into the 65 kDa form removed at least part of the conserved COOH-terminal sequence. EGA has no detectable homology with EGD, except for the reiterated, conserved domain present at the COOH terminus of both proteins. Any immunological cross-reactivity of EGD with EGA antibodies would therefore be expected to be due to the presence of the conserved domain within EGD protein. As shown in Fig. 3, the 68 kDa form of EGD (lane b) reacted with anti-EGA antibodies. In contrast, no cross-reactivity was observed with the 65 kDa form (lane c) obtained after dialysis of the urea extract, nor with the 63 kDa form synthesized from a truncated gene not encoding the COOH-terminal reiterated region (lane a) [15]. The result confirms previous COOH-

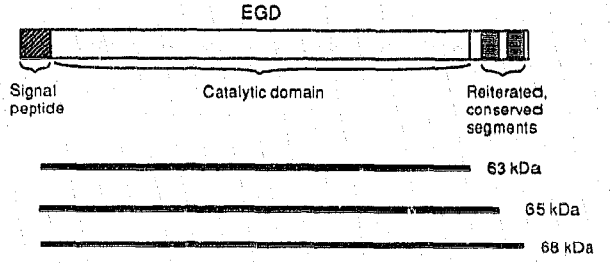

Fig. 2. Domain structure of EGD. The extent of the catalytic domain was defined by deletion analysis of *celD* (data not shown). The portion of the EGD sequence present in the 63 kDa, in 65 kDa and in the 68 kDa forms discussed in the text are shown by horizontal bars. Each species resulted from the fusion of *celD* in frame with the start of *lacZ'*, using the *Hin*dIII site present in the polylinker of pUC8 (65 and 68 kDa forms) [8] or pUC19 (63 kDa form) [15].

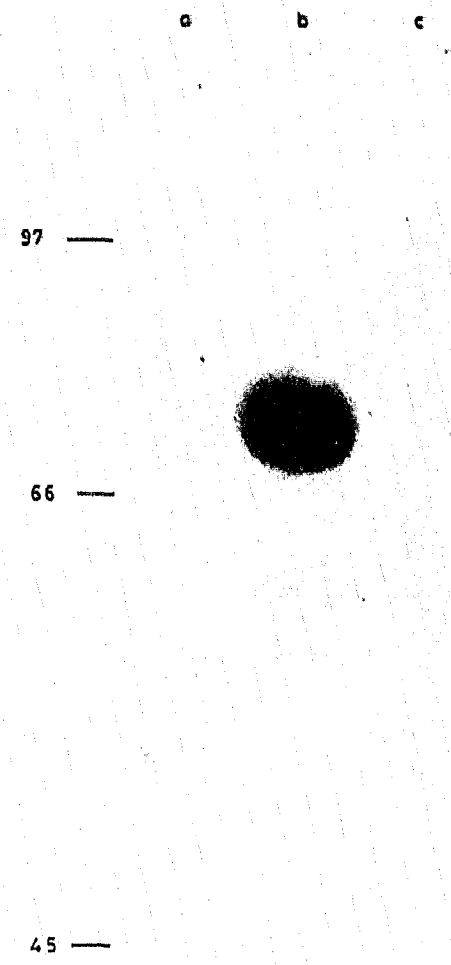

Fig. 3. Recognition of the 68 kDa form of EGD by anti-EGA antiserum. Samples were submitted to SDS-PAGE followed by Western blotting using anti-EGA antiserum. a, 63 kDa EGD (10 μU); b, 68 kDa EGD (20 μU); c, 65 kDa EGD (23 μU).

terminal analysis showing that the 65 kDa species did not end at Ile-649, the last residue predicted from the nucleotide sequence of *celD* [18], but at Leu-605 [15], which is located at the end of the first repeated segment (Fig. 2).

3.3. *Insoluble EGD present in inclusion bodies has close to full activity.*

EGD present in inclusion bodies could be assayed prior to urea solubilization using the low molecular weight substrate *p*-NPC. Control experiments showed that the activity was not released in soluble form, but remained in the insoluble fraction under the conditions of the assay (data not shown).

As shown in Table I, about 25% of the total activity present in the broken cell suspension prior to centrifugation is accounted for by the insoluble fraction. Association of the activity with inclusion bodies and not with some other insoluble fraction (e.g. cell debris) was suggested by the concomitant appearance of insoluble EGD activity and of refractile bodies in recombinant cells (unpublished data). Furthermore, more than 90% of the activity remained in insoluble form after treatment with 2 M NaCl, 20 mM EDTA, or 1% Triton X-100, or in buffers with a pH range of 4.8 to 8.8. In addition, only 12% of the activity was released in soluble form in the presence of 2 M urea (data not shown). Neither the urea treatment, nor the following dialysis resulted in loss or increase of activity. This is consistent with data showing that purified EGD is fully active in the presence of 5 M urea, and that, at 20°C, 8 M urea does not induce any change detectable by circular dichroism (A. Chaffotte and M. Goldberg, personal communication).

The specific activity of the enzyme present in each fraction was calculated after estimating the amount of EGD by Western blotting. Since antibody probing occurred after the samples had been denatured in SDS at 100°C, the immunoreactivity of the EGD species detected was independent of their initial conformation. The lowest specific activity was found for the soluble cytoplasmic enzyme (0.9 ± 0.2 U/mg). EGD present in inclusion bodies prior to solubilization had a specific activity of 1.4 ± 0.25 U/mg, close to the activity of the purified enzyme (1.6 ± 0.4 U/mg). The two minor bands detected in Fig. 1, lanes B to H, were not taken into account for the calculation of specific activity, since, from previous deletion analysis of *celD* (data not shown), they corresponded to polypeptides too short to contain a functional active domain. It is possible that other, still smaller EGD fragments were present in solubilized inclusion bodies and escaped antibody detection. However, the data show that most or all of intact size EGD protein present in inclusion bodies (and not only a subfraction that would be preferentially solubilized in the presence of urea) had nearly normal activity. Furthermore, the presence of active EGD in the insoluble fraction does not seem to be due to passive, non-specific trapping of soluble proteins, since the insoluble and the cytoplasmic fraction of JM101(pCT603) show quite different electrophoretic patterns [8]. As pointed out before, the enzyme purified

Table I

Distribution of the activity between various EGD fractions isolated from *E. coli* JM101(pCT603)

Sample	Activity (U)	Activity yield (%)
Broken cell suspension	130	100
Cytoplasmic supernatant	100	76
Washed inclusion bodies	32	25
Solubilized inclusion bodies	32	25
Urea dialysate	32	25

Cells from a 1 liter overnight culture (OD_{600} = 4.5) were processed and the total activity present in each fraction was assayed as described in section 2.

from the granule fraction does not consist of a heterogeneous mixture of species with different conformations since it can be readily crystallized.

4. DISCUSSION

How do the results presented above fit with the view that inclusion bodies result from improper protein folding leading to the formation of aggregates? Activity measurements show that aggregated EGD displays little change of conformation of the catalytic core, which comprises 90% of the total sequence of the protein. Similarly, urea treatment does not affect the activity of the enzyme significantly. These features are consistent with the thermostable character of the enzyme, which implies a high stability of its native structure. However, changes in the conformation of the COOH-terminal region, which is distinct from the catalytic domain and not required for activity [15], would not be detected by activity assays. It is therefore tempting to speculate that, in the case of EGD, inclusion body formation is mediated, at least in part, by intermolecular interactions involving the COOH-terminal domain, which would be sensitive to urea. Two observations are consistent with the proposed model: (i) no insoluble, active EGD could be pelleted from broken cells carrying a truncated version of *celD* encoding a form of EGD devoid of the COOH-terminal domain [15]; (ii) inclusion bodies contained almost exclusively the 68 kDa species, which carries most or all of the COOH-terminal domain, whereas only the truncated 65 kDa species was found in the soluble cytoplasmic fraction.

Thus, the departure of EGD from the usual properties of inclusion body proteins (denatured conformation devoid of activity) could be only apparent. The participation of the COOH-terminal domain in intermolecular interactions leading to aggregation is fully consistent with the model proposed by Mitraki and King [9], and earlier by Goldberg and Zetina [20], based on the interaction of partially misfolded polypeptides. The critical point is that, in the case of EGD, conformational changes do not affect the catalytic domain of the enzyme.

It remains to be determined whether the formation of cytoplasmic aggregates in *E. coli* has any connection with the real function of the reiterated domain. It has been speculated that the conserved region could participate in the anchoring of the various components to the high molecular weight cellulolytic complex, or cellulosome [21], of *C. thermocellum*. Thus, it might act as a 'sticky' tail, which would also enhance inclusion body formation when the proteins are expressed in *E. coli*.

Acknowledgements: We thank A. Chaffotte and M. Goldberg for communicating and discussing their results prior to publication. R. Longin is acknowledged for helpful discussions. K.T. was supported by Presidential Young Investigator Grant ECE-8552492 awarded by the National Science Foundation to P.D. This work was supported by Grant EN3B-0082-F from the Commission of the European Communities and by research funds from the University of Paris 7.

REFERENCES

[1] Williams, D., Van Frank, R.M., Muth, W. and Burnett, J. (1982) Science 215, 687–689.
[2] Krueger, J.K., Kulke, M.H., Schutt, C. and Stock, J. (1989) Biopharm 3, 40–45.
[3] Marston, F.A.O. (1986) Biochem. J. 240, 1–12.
[4] Schein, C.H. (1989) Bio/Technology 7, 1141–1148.
[5] Schein, C.H. (1990) Bio/Technology 8, 308–317.
[6] Schoemaker, J.M., Brasnett, A.H. and Marston, F.A.O. (1985) EMBO J. 4, 775–780.
[7] Schoner, R.G., Ellis, L.F. and Schoner, E.F. (1985) Bio/Technology 3, 151–154.
[8] Joliff, G., Béguin, P., Juy, M., Millet, J., Ryter, A., Poljak, R. and Aubert, J.-P. (1986) Bio/Technology 4, 896–900.
[9] Mitraki, A. and King, J. (1989) Bio/Technology 7, 690–697.
[10] Maniatis, T., Frisch, E.F. and Sambrook, J. (1982) Molecular cloning: A laboratory manual, Cold Spring Harbor Lab., Cold Spring Harbor, New York.
[11] Pètre, J., Longin, R. and Millet, J. (1981) Biochimie 63, 629–639.
[12] Laemmli, U.K. (1970) Nature 227, 680–685.
[13] Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350–4354.
[14] Grépinet, O., Chebrou, M.-C. and Béguin, P. (1988) J. Bacteriol. 170, 4582–4588.
[15] Chauvaux, S., Béguin, P., Aubert, J.-P., Bhat, K.M., Gow, L.A., Wood, T.M. and Bairoch, A. (1990) Biochem. J. 265, 261–265.
[16] Sedmak, J.J. and Grossberg, S.E. (1977) Anal. Biochem. 79, 544–552.
[17] Babbit, P.C., West, B.L., Buechter, D.D. Kuntz, I.D. and Kenyon, G.L. (1990) Bio/Technology 8, 945–949.
[18] Joliff, G., Béguin, P. and Aubert, J.-P. (1986) Nucleic Acids Res. 14, 8605–8613.
[19] Hall, J., Hazlewood, G.P., Barker, P.J. and Gilbert, H.J. (1988) Gene 69, 29–38.
[20] Goldberg, M.E., and Zetina, C.R. (1980) in: Protein Folding (R. Jaenicke, ed) pp. 469–484, Elsevier, Amsterdam.
[21] Lamed, R., Setter, E., Kenig, R. and Bayer, E.A. (1983) Biotechnol. Bioeng. Symp. 13, 163–181.