Increase of Solubility of Foreign Proteins in Escherichia coli by Coproduction of the Bacterial Thioredoxin

(Received for publication, July 12, 1995, and in revised form, September 1, 1995)

Takashi Yasukawa, Chie Kanei-Ishii, Toshio Maekawa, Jiro Fujimoto, Tadashi Yamamoto, and Shunsuke Ishii

From the Laboratory of Molecular Genetics, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan; the Institute of Medical Sciences, University of Tsukuba, 1-1 Tennoudai, Tsukuba, Ibaraki 305, Japan; and the Department of Oncology, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokane, Minato-ku, Tokyo 108, Japan.

Eukaryotic proteins are frequently produced in Escherichia coli as insoluble aggregates. This is one of the barriers to studies of macromolecular structure. We have examined the effect of coproduction of the E. coli thioredoxin (Trx) or E. coli chaperones GroESL on the solubility of various foreign proteins. The solubilities of all eight vertebrate proteins examined including transcription factors and kinases were increased dramatically by coproduction of Trx. Overproduction of E. coli chaperones GroESL increased the solubilities of four out of eight proteins examined. Although the tyrosine kinase Lck that was produced as an insoluble form and solubilized by urea treatment had a very low autophosphorylating activity, Lck produced in soluble form by coproduction of Trx had an efficient activity. These results suggest that the proteins produced in soluble form by coproduction of Trx have the native protein conformation. The mechanism by which coproduction of Trx increases the solubility of the foreign proteins is discussed.

Production of large amounts of proteins in Escherichia coli is a first step for structural studies of macromolecules. However, many eukaryotic proteins, especially full-length proteins, are produced in E. coli as insoluble aggregates (inclusion bodies). Our laboratory is interested in the three-dimensional structure of transcription factors, especially nuclear oncogene products such as the myb proto-oncogene product (c-Myb). We identified the solution structure of the Myb-DNA complex by using the bacterially produced Myb protein containing the DNA-binding domain alone (1). However, the structural study of the full-length Myb has been unsuccessful, because it was produced in E. coli as insoluble aggregates. Although production of a protein as insoluble aggregates can offer the advantage of easy purification, a protein solubilized by some appropriate process such as urea treatment does not have a guarantee that it has the native protein conformation. Thus, large-scale production of eukaryotic proteins in E. coli in soluble form is the first step for structural studies.

The mechanism by which proteins become soluble is not understood. The formation of inclusion bodies might be thought as "inappropriate" protein-protein interactions due to the lack of proper polypeptide folding (2, 3). Why are many eukaryotic proteins produced as insoluble aggregates in E. coli? Two factors appear to affect the solubility of eukaryotic proteins in E. coli. The first parameter is the E. coli heat shock chaperone GroESL (encoded by the groE operon, Fig. 1). The role of the GroESL complex in catalyzing the correct folding of a newly synthesized polypeptide has become firmly established recently (4–7). To produce eukaryotic proteins in E. coli, a strong promoter like the T7 promoter is often used. In this case, a high level of production of the E. coli chaperones GroESL may be needed. For example, when λ phage infects E. coli, the production of GroESL is induced. If the level of functional GroESL does not increase, λ phage cannot form the phage particles, because the folding of λ coat proteins does not occur correctly (8). Thus, the coordinate induction and high level production of E. coli chaperones may be required for proper folding of the foreign proteins. The second parameter that affects the solubility of eukaryotic proteins in E. coli could be the difference of redox state between E. coli and eukaryotic cells. Most of the fusion proteins with GST (glutathione S-transferase) containing various mammalian proteins produced in E. coli bind to glutathione-Sepharose beads very efficiently. In contrast, the GST-fusion proteins produced in mammalian cells bind to glutathione beads only with low efficiency. This observation suggests that mammalian cells have a different redox environment from E. coli. Consistent with this observation, it was reported that quite high concentrations of glutathione are maintained in mammalian cells (9).

We report here that the solubility of various eukaryotic proteins in E. coli is increased by coproducing E. coli thioredoxin (Trx). By coproducing Trx, eight foreign proteins including transcription factors and oncogene products were successfully produced in soluble form. The effect of Trx coproduction on the solubility of foreign proteins is compared with that of GroE coproduction.

MATERIALS AND METHODS

Construction of Plasmids—To make the pACYC plasmid containing the T7 promoter (pACYC-T7), the HindIII-SphI 0.5-kb DNA fragment of pACYC184 (10) was replaced by the 0.6-kb HindIII-SphI fragment containing the T7 promoter from the pAR2156 vector (11, 12). The NdeI-BglII 2.1-kb DNA fragment containing the GroESL-coding region was made by the polymerase chain reaction using the groE plasmid (pK1561) (13) and inserted into the NdeI-BamHI site of pACYC-T7 to generate the plasmid to produce E. coli GroESL under the control of T7 promoter (pT-GroE). To make the plasmid in which the T7 promoter was linked to the E. coli Trx coding region (pT-Trx) (14), the NdeI-HindII fragment of pT-GroE was replaced by the NdeI-HindII fragment containing the Trx-coding region and the aspA transcription ter-

1 The abbreviations used are: GST, glutathione S-transferase; CREBP1, CAMP response element-binding protein 1; E3A, adenosine E3A gene product; GroESL, groES and groEL gene products; Trx, E. coli thioredoxin; kb, kilobase(s); IPTG, isopropyl-β-D-thiogalactopyranoside; PAGE, polyacrylamide gel electrophoresis.

2 T. Yasukawa, C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii, unpublished data.
minator. The plasmids to produce various vertebrate proteins were constructed using the appropriate pET expression vector containing the T7 promoter (11, 12) and CDNA clones.

Bacterial Strains and Production of Proteins—The E. coli strain BL21(DE3) (11, 12) harboring pT-GroE or pT-Trx, or pET vector to produce various vertebrate proteins was made by transformation with each plasmid. To generate the bacteria to produce both foreign protein and GroE or Trx, the E. coli strain BL21(DE3) harboring pT-GroE or pT-Trx was transformed by pET vector encoding various vertebrate proteins. The bacteria was cultivated in 2.5 ml of superbroth to A550 of 0.7 (IPTG(−)), or was then cultivated for 4 h more in the presence of 1 mM IPTG (IPTG(+)). Total lysates were prepared and centrifuged to separate the soluble and insoluble fractions. Samples prepared from 0.4 and 0.7 ml of culture were analyzed by 10% (for GroE) (left) and 15% (for GroES and Trx) (right) SDS-PAGE followed by Coomassie staining.

RESULTS

We tried to examine the effect of coproduction of E. coli Trx or E. coli chaperones GroESL or E. coli Trx on the solubility of vertebrate proteins in E. coli. The pET vector, which contains the T7 promoter and a pBR322-based replicon (11, 12), was used to produce foreign proteins. To coproduce GroESL or Trx with a level similar to that of foreign proteins of interest, the GroESL- or Trx-coding region was also linked to the T7 promoter and inserted into the pACYC vector, which contains a p15A replicon and a chloramphenicol resistance marker (10) (Fig. 1A). The generated plasmids, pT-GroE or pT-Trx, allowed for cotransformation with the pET plasmids to produce various vertebrate proteins based upon plasmid compatibility. To confirm the production of GroESL or Trx from these plasmids, the BL21(DE3) bacteria transformed with pT-GroE or pT-Trx plasmids was cultured in the presence or absence of IPTG. Fig. 1B shows the Coomassie staining of the total proteins of each culture following SDS-PAGE. The overproduction of GroEL and GroESL from pT-GroE or Trx from pT-Trx is striking, with over 30% of the total cellular protein being GroESL or Trx.

We first examined the effects of coproduction of GroESL on the solubility of various vertebrate proteins (Fig. 2, see lanes marked by +GroE). To produce various foreign proteins, the BL21(DE3) transformants harboring the pET plasmid alone or both the pET plasmid and pT-GroE were made. To assess the production and solubility of each of the proteins, cultures were harvested 4 h postinduction, the cell pellets were lysed by sonication, and the resulting lysates were separated into soluble and insoluble fractions by centrifugation. The proteins in both fractions were separated by SDS-PAGE followed by Coomassie staining. Without coproduction of GroESL, mouse c-Myb (16) was produced completely as insoluble aggregates. However, coproduction of GroESL significantly increased the solubility of c-Myb, and approximately 10% of c-Myb was produced in soluble form, resulting in the production level of about 30 mg of soluble form per liter of culture. Similar increases in solubility were observed with two other human transcription factors, cAMP response element-binding protein 1 (CRE-BP1) (also called ATF-2) (17) and the p53 tumor suppressor gene product (18). However, coproduction of GroESL did not increase the solubility of three other nuclear factors, the skil-related gene product SnoN (19), myc proto-oncogene product (Myc) (20), and adenovirus oncogene product E1A (21). We have also examined the effects of coproduction of GroESL on the solubility of vertebrate kinases. The solubility of one of the Ser/Thr kinases, the Xenopus mos proto-oncogene product

---

**Image 199x457 to 551x742**

**Fig. 1.** Expression vector for E. coli chaperones GroESL or E. coli Trx. A, structure of the GroESL or Trx expression vector. The GroESL- or Trx-coding region was inserted into the T7 promoter and inserted into the pACYC vector containing the chloramphenicol resistance marker. B, induction of GroESL or Trx production. E. coli BL21(DE3) harboring GroESL or Trx expression vector was cultivated to A550 of 0.7 (IPTG(−)), or was then cultured for 4 h more in the presence of 1 mM IPTG (IPTG(+)). Total lysates were prepared and centrifuged to separate the soluble and insoluble fractions. Samples prepared from 0.4 and 0.7 ml of culture were analyzed by 10% (for GroE) (left) and 15% (for GroES and Trx) (right) SDS-PAGE followed by Coomassie staining.
Increased Solubility of Foreign Proteins by Coproduction of Trx

FIG. 2. Increase of solubility of mammalian proteins by coproduction of GroESL or Trx. E. coli BL21(DE3) harboring the pET expression vector to produce various proteins indicated on the right with (+) or without (−) GroESL expression vector (GroE) or Trx expression vector (Trx) was cultivated to an OD of 0.7 (IPTG(−)), and was then cultivated for 4 h more in the presence (IPTG(+)) or absence (IPTG(−)) of 1 mM IPTG. The soluble (marked by S) and insoluble (marked by I) fractions were prepared and analyzed on 10% or 8% SDS-PAGE followed by Coomassie staining.

DISCUSSION

We have demonstrated that coproduction of Trx dramatically increases the solubility of all eight foreign proteins examined in the protein sample solubilized by urea. In contrast, the proteins produced in soluble form by coproduction of Trx appear to have the native protein conformation.

E. coli. However, an increase of the solubility by coproduction of GroESL was observed with only four of the eight proteins examined. These results indicate that the Trx coproduction system is more useful than the GroE system to produce foreign proteins as soluble forms. By making the plasmid to coproduce both Trx and GroE, we have also examined whether overproduction of both Trx and GroE can more effectively increase the solubility of vertebrate proteins than that of either protein. However, we have observed no additional effects of overproduction of either Trx or GroE. Improvement of the solubility of foreign proteins by overproduction of GroESL was reported recently by other groups with human procollaganase (25) or human Csk (26). In spite of the obvious effect on the solubility of procollaganase, Lee and Ollins (25) reported that overproduction of GroESL had no effect on the solubility of three other proteins. Our better results with GroE (four successful cases out of eight) than in the other report could be due to the high level of production of GroESL under the control of the T7 promoter.

It was reported that the use of Trx as a gene fusion partner increases the solubility of foreign proteins like cytokines (27). However, our finding that the overproduction of free Trx dramatically increases the solubility of foreign proteins is clearly different from this. This finding gives us a big advantage in preparing the large amount of proteins for structural study. In the case of fusion proteins, Trx has to be cleaved off by a specific peptidase, but the efficiency of cleavage is often very low. Increase of the solubility of foreign proteins by overproduction of Trx strongly suggests that the redox state affects the solubility of foreign proteins. Our observation that many GST-fusion proteins produced in mammalian cells bound to glutathione-Sepharose beads much less efficiently than those produced in E. coli cells suggest that E. coli cells have the relatively oxidative environment compared to mammalian cells. This could induce formation of abnormal intramolecular disulfide bonds that aggregate the proteins. Another observation that addition of a reducing reagent in the process of solubilization of insoluble aggregates of c-Myc by urea or guanidine hydrochloride increases the efficiency of obtaining the functional proteins2 may support this speculation.

Comparison of the activity of Lck between the soluble form and the solubilized form by urea indicates that the soluble form has higher activity than the urea-treated form, indicating that only a portion of the molecules in the urea-treated sample has a native protein conformation. These results show that the artificially solubilized protein sample is not suitable for structural study. Recently, we have succeeded in purifying the large amount of full-length c-Myc protein produced as a soluble form using the Trx system. This c-Myc protein sample will be useful not only for crystal study but also for various experiments to
examine the interaction between c-Myb and other proteins. The Trx coproduction system allows us to prepare many mammalian proteins suitable for structural study and molecular biological study.

Acknowledgments—We thank T. Yura for the groE plasmid, K. Okazaki and N. Sagata for the Mos expression plasmid, and Y. Koga for the Lck-specific antibody.

REFERENCES

1. Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., and Nishimura, Y. (1994) Cell 79, 639–648
2. Schein, C. H. (1989) Bio/Technology 7, 1141–1149
3. Mitraki, A., and King, J. (1989) Bio/Technology 7, 690–702
4. Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Pfeifer, G., J. Viitanen, P. V., Lorimer, G. H., and Buchner, J. (1994) Science 265, 656–659
5. Azen, A., Kasal, M., and Goloubinoff, P. (1994) Science 265, 653–656
6. Martin, J., Mayhew, M., Langer, T., and Hartl, F. U. (1993) Nature 366, 228–233
7. Georgopoulos, C., Ang, D., Liberek, K., and Zylcic, M. (1990) in Stress Proteins in Biology and Medicine (Morimoto, R. I., Tissieres, A., and Georgopoulos, C., eds) pp. 191–221, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
8. Kondo, T., Yoshida, K., Urata, Y., Goto, S., Gasa, S., and Taniguchi, N. (1993) J. Biol. Chem. 268, 20366–20372
9. Chang, A. C. Y., and Cohen, S. N. (1978) J. Bacteriol. 134, 1141–1156
10. Studier F. W., and Moffatt, B. A. (1986) J. Mol. Biol. 185, 113–130
11. Rosenberg, A. H., Lade, B. N., Chui, D.-S., Lin, S.-W., Dunn, J. J., and Studier, F. W. (1987) Gene (Amst.) 56, 125–135
12. Kanemori, M., Mor, H., and Yura, T. (1994) J. Bacteriol. 176, 4235–4242
13. Lunn, C. A., Kathju, S., Wallace, B. J., Kushner, S. R., and Piglet, V. (1984) J. Biol. Chem. 259, 10469–10474
14. Yamanashi, Y., Kakiuchi, T., Mizuguchi, J., Yamamoto, T., and Toyoshima, K. (1991) Science 251, 192–194
15. Sarkar, H., Kanel-Ishii, C., Nagase, T., Nakagoshi, H., Gonda, T. J., and Ishii, S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5758–5762
16. Maekawa, T., Sakuri, H., Kanel-Ishii, C., Sudo, T., Yoshimura, T., Fujisawa, J., Yoshida, M., and Ishii, S. (1989) EMBO J. 8, 2023–2028
17. Vogelstein, B., and Kindler, K. W. (1992) Cell 70, 523–526
18. Nagase, T., Nomura, N., and Ishii, S. (1993) J. Biol. Chem. 268, 13710–13716
19. Luzcher, B., and Eisenman, R. N. (1990) Genes & Dev. 4, 2025–2035
20. Maran, E., and Mathews, M. B. (1987) Cell 48, 177–178
21. Sagata, N., Watanabe, N., Vande Woude, G. F., and Ikawa, Y. (1989) Nature 342, 512–518
22. Martin, J. D., Peet, R., Krebs, E. G., and Perlmutter, R. M. (1985) Cell 43, 393–404
23. Voronova, A. F., and Sefton, B. M. (1986) Nature 321, 682–685
24. Lee, C. S., and Ohins, P. O. (1992) J. Biol. Chem. 267, 2849–2852
25. Amrein, K. E., Tancaks, B., Stieger, M., Mdrno, J., Flint, N. A., and Burn, P. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1048–1052
26. LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., and McCoy, J. M. (1993) Biotech. Technology 11, 187–193