Journal of Inclusion Phenomena and Macrocyclic Chemistry **50**: 37–41, 2004.

Mini Review

Purification and Refolding of Cyclodextrin Glycosyltransferase Expressed from Recombinant *Escherichia coli*

DAE-HYUK KWEON[1,2], SUNG-GUN KIM[1,3] and JIN-HO SEO[1,3]*
[1]*Department of Food Science and Technology and School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, Korea;* [2]*School of Bioresource Sciences, Andong National University, Kyungbuk 760-749, Korea;* [3]*Interdisciplinary Program for Biochemical Engineering and Biotechnology, Seoul National University, Seoul 151-742, Korea*

(Received: 15 September 2003; in final form: 9 October 2003)

Key words: cyclodextrin glycosyltransferase, polycationic tag, purification, solid-state refolding

Abstract

Recombinant DNA technology and protein engineering are currently utilized in the cost-effective production of pharmaceutical and industrial proteins with native conformation. *Escherichia coli* retains its dominant position as the first choice of host for speed, simplicity and well-established production protocols. However, protein production using recombinant *E. coli* occasionally encounters complex purification and refolding steps. This paper introduces an efficient scheme for purification and *in vitro* refolding of industrially important proteins including cyclodextrin glycosyltransferase (CGTase) expressed in recombinant *E. coli* employing a polycationic amino acid fusion system. Fusion of polycationic amino acids to CGTase allowed purification and refolding of CGTase to be simple and efficient. A novel CGTase production strategy will be discussed by considering recent progress in protein purification and refolding techniques.

Introduction

Cyclodextrin glycosyltransferase (CGTase) gains more interests because of its sugar-transferring activity as well as the original cyclodextrin (CD)-producing activity. CGTase are produced by a variety of bacteria including: (i) aerobic mesophilic bacteria such as *Bacillus macerans*, *B. megaterium*, *B. cereus*, *B. ohbensis*, *Klebsiella pneumoniae*, *K. oxytoca*, *Micrococcus luteus*; (ii) aerobic thermophilic *B. stearothermophilus*; (iii) anaerobic thermophilic *Thermoanaerobacterium thermosulfurigenes*; (iv) aerobic alkalphilic bacteria such as *B. circulans*, *Bacillus* sp. AL-6; (v) aerobic halophilic *B. halophilus* [1]. Many studies on the production of CGTase from original bacterial sources can be found [2–6].

Production of heterologous polypeptides in *Escherichia coli* has become a routine matter thanks to the well-established recombinant DNA methodologies, the availability of various efficient expression vectors, an ability to grow rapidly on inexpensive substrates and well-established high cell density culture techniques [7, 8]. The production of CGTase using recombinant *E. coli* has been studied extensively in our laboratory [9–11].

Successful commercialization of proteins, especially for industrial enzymes like CGTase, usually depends upon the ability to produce the target proteins inexpensively in a large scale. However, high level expression of foreign proteins in recombinant *E. coli* tends to result in aggregation and accumulation as inclusion bodies. Refolding of inclusion bodies into soluble active enzymes needs high cost and tedious jobs. Further, complicated purification steps are a cost-determining stage even when the enzyme is expressed soluble. Thus, an efficient protein production strategy has been sought by taking into account methodologies both at a gene level and at a process scale [12, 13].

The present review summarizes recent progress in purification and refolding of heterologous proteins that may shed light on the efficient production of CGTase. It is also suggested that expression of CGTase fused to polycationic amino acids is suited for CGTase production since the polycation provides versatile tools for the efficient purification and *in vitro* refolding of inclusion bodies.

Polycationic fusion tag for efficient purification

Development of sophisticated protein-fusion systems has facilitated cost-effective production and purification of foreign proteins in *E. coli*. Fusion partners offer several advantages such as prevention of inclusion body

* Author for correspondence. E-mail: jhseo94@snu.ac.kr

formation, improved folding characteristics, limited proteolysis and generic protein purification schemes. Fusion proteins widely used include staphylococcal protein A, streptococcal protein G, *Schistosoma japonicum* glutathione-*S*-transferase (GST), maltose-binding protein, thioredoxin, DsbA and ubiquitin. The detection and purification of product proteins may be also facilitated by the use of affinity tags such as the FLAG, His_6, and c-Myc peptides. There are several extensive reviews on this topic [14, 15]. Thus, only fusion tags suitable for the purification of commercial enzymes are within the interest of this review.

Protein purification is achieved by exploiting physicochemical differences in charge, hydrophobicity and size or by using bioaffinity methods based on specific structural or functional properties. Rather than straying around to find optimal purification conditions, protein fusion offers a straightforward strategy to recover the target protein. By fusing a protein of interest to a partner with well-known physicochemical characteristics, the target protein can be easily purified by capitalizing on the characteristics of the fusion tag. Acceptable fusion partners should possess properties that can be employed for an easy purification scheme.

The most commonly used tags for purification purpose are the polyhistidine (His6), GST and FLAG (AspTyrLysAsp4Lys) tags. Polyhistidine and FLAG tags can permit efficient purification and detection of target proteins by using metal-chelate chromatography resins and a FLAG specific antibody, respectively. GST fusions bind to the glutathione-agarose followed by elution with free glutathione. Even though the GST tag has been widely used in laboratory because of improved soluble expression as well as provision of simple purification schemes, very high cost of the glutathione-agarose matrix limits its application to commercial enzymes. Only when small quantities are required for experimental evaluation or when the protein is of high value, these expensive affinity tags may be acceptable.

If a protein is isolated in a large scale, as usual for commercial enzymes like CGTase, the cost of the purification matrix may be critical. The most widely used and the most inexpensive media might be an ion-exchanger. A fusion tag that can be purified based on ionic binding force may provide a cost-effective scheme over the affinity fusion tags.

A charged fusion tail was first pioneered by Sassenfeld and his colleagues to recover proteins from *E. coli* cell extracts [16]. They genetically engineered the fusion of poly-arginine (six Arg residues) tails to several proteins in *E. coli* and recovered the fusion proteins by ion-exchange chromatography. β-Urogastrone and bacterial aspartate aminotransferase fused with six arginines at their C-terminal were simply isolated from the cell lyzates by ion-chromatography. Dalbøge *et al.* have used glutamic acid residues in an amino-terminal tail to distinguish the tailed from the untailed human growth hormone by ion-exchange high pressure liquid chromatography [17]. Zhao *et al.* have fused target proteins with aspartic acid tails and recovered by precipitation using the polycation polyethyleneimine (PEI) [18]. They used the intracellular *E. coli* enzyme β-galactosidase (β-gal) as a model protein. Purificaiton methods for the recovery of negatively charged fusion proteins using precipitation by PEI, hollow ion-exchange membrane, and aqueous two-phase systems have been extensively studied by Glatz and co-workers [18–24].

Recently, fusion proteins with charged polycationic amino acid tails were constructed for the purpose of simple ion-exchange purification of CGTase with high purity [25]. A number of positively charged lysine and arginine tails were fused to the C-terminus of CGTase derived from *B. macerans* and expressed in *E. coli*. The ionic binding forces provided by the tails allowed the selective recovery of CGTase from recombinant *E. coli* cell extracts, while CGTase by itself could not bind to the cation exchanger at neutral pH. The type of amino acids used and the length of the tail directly affected the purification factors. Most intracellular proteins of *E. coli* adsorbed on the cation exchanger could be removed by washing with 400 mM NaCl solution at pH 7.4, suggesting that a fusion partner suitable for purification purpose should be provided with high binding strength and be maintained adsorbed against washing with NaCl solution. Among the fusion CGTases constructed, the CGTK10ase containing 10 lysine residues at the C-terminal provided sufficiently high binding strength to allow purification to its homogeneity through simple ion-exchange chromatography.

This study has demonstrated experimentally that ion-exchange chromatography can be employed for the recovery of commercial enzymes with high purity in a cost-effective manner. When CGTase is fused to a polycation, inexpensive ion-exchangers may substitute for the expensive affinity chromatography media without loss of yield and purity.

Refolding harnessing polycationic fusion tag

Inclusion bodies are very dense particles of aggregated proteins, exhibiting an amorphous or paracrystalline structure independent of their subcellular location. The nature of the expressed proteins, the rate of their expression, and the level of expression are known to exert a profound influence on the formation of inclusion bodies [26]. High expression rate causes insufficient time for the nascent polypeptide chain to fold into the native conformation. Subsequent highly localized concentration of misfolded or folding intermediate polypeptide in the cytoplasm leads to non-specific precipitation.

Deposition of heterologous proteins in inclusion bodies can be either advantageous or disadvantageous. Where renaturation is problematical, the preferred way is to avoid or at least to reduce inclusion body formation. In contrast, when a simple and efficient

renaturation procedure exists, deposition of the protein in inclusion bodies and subsequent isolation and renaturation of the aggregated protein often means the most straightforward strategy to get large amounts of the product protein [12, 13, 27].

The formation of inclusion bodies offers several advantages for the production of heterologous proteins. Heterologous proteins may be unstable in the cytoplasm of a bacterium due to proteolysis (e.g., insulin chains A and B) and may be toxic to the host cell in the native conformation (e.g., immunotoxins). However, inclusion bodies are generally not very sensitive to proteolytic breakdown. Moreover, under appropriate conditions the recombinant protein deposited in inclusion bodies amounts to about 50% or more of the total cell protein. This obviously enables cost-efficient downstream processing. However, refolding of inclusion bodies is not a straightforward process. The main challenges lie in the tedious and inefficient steps of unfolding and refolding for the recovery of biological activity.

The efficiency of a refolding process depends on the competition between correct folding and aggregation [12, 13]. In order to minimize aggregation, refolding is usually performed at low protein concentrations in a range of 10–100 mg/L. Refolding conditions must be carefully selected based on temperature, pH, additives, ionic strength, and concentrations of the denaturant and the protein itself [12]. Even under an optimized condition, the yield of renaturation is often relatively low, necessitating large process volumes for the preparation of the native protein. Therefore, research activities have been often focused on the protection of intermolecular aggregation, either by changing the processes for the removal of the denaturants or by adding aggregation-protecting molecules.

Dialysis and diafiltration are the most frequently employed techniques for the exchange of buffer. In dialysis and diafiltration, membranes of defined molecular weight cut-off are used. Since the membrane molecular weight cut-off is much lower than the target protein, the protein is retained by the membrane, whereas buffer exchange results in refolding of the protein. In gel filtration, the denaturant enters the pores of the matrices, whereas the protein is exposed to the refolding buffer. The buffer-exchange method is much faster than dialysis, but the relative solubilities of the folding intermediates determine the success of the technique [28, 29].

A less time-consuming method than the previous ones is dilution. The solubilized inclusion body is diluted 10–100 times in the refolding buffer, thereby the concentration of the denaturant is reduced to restore the native secondary structure. At the same time, in order to minimize the formation of aggregates, the dilution is carried out by slow adding the denatured protein to the refolding buffer, or by quick pouring the refolding buffer into the denatured protein solution. This method is the most aggregation-provocative one among refolding methods.

The minimization of protein aggregation is the key factor of a refolding process. This goal have been achieved by incorporating various additives in the refolding buffer. It is believed that the additives either bind to the folding intermediates and inhibit their aggregation, or stabilize the refolded native structure of the protein and thereby improve the yield of the refolding process. Amino acids, sugars, neutral surfactants, and polymers (polyethylene glycol) have been used as additives for the refolding of various proteins such as Fab-fragments, relaxin, β-lactamase, lysozyme, RNaseA, tPA, and so forth [12, 13, 27–29].

Since the concentration of non-native polypeptides is the most critical factor of aggregation in the refolding process, a strategy employing the immobilization of target proteins on a solid phase has been proposed to circumvent aggregation of unfolded proteins or folding intermediates (so called solid-state refolding). Solid-state refolding is based on that proteins attached to an insoluble carrier may avoid intermolecular aggregation. In conventional refolding techniques, the process volume is usually very large in order to maintain the fully or partially unfolded proteins at low concentrations. The dilute solution in large volumes makes it very difficult to recover the correctly folded protein. Furthermore, side reactions such as aggregation cause exceedingly low yield of renaturation. In solid-state refolding, proteins are prevented from intermolecular aggregation. Further, it becomes easy to alter the solvent readily and rapidly, expediting manipulation of the folding conditions. Such effects allow the refolding process to obtain a good yield in a minimized process volume, labor and time.

Proteins covalently attached to the matrix have been unfolded and refolded with improved yields when intermolecular interactions are particularly prevented, when the degradation by proteolysis was avoided [30–32]. However, preparative protein refolding in the solid phase requires that the proteins be reversibly attached to the solid support such as Ni-resin or ion-exchange resin in the presence of a denaturant.

Refolding of the proteins reversibly adsorbed to an ion-exchanger has been performed for several proteins [33]. These proteins were absorbed on cation- or anion-exchangers through their own charges. Ni-NTA agarose has been also employed using the His_6 fusion tag [34]. However, solid-phase refolding using the own charge of the protein is often encountered with the problem associated with the promiscuous electrostatic interactions between the arbitrarily distributed local charge of the denatured protein and the counter-charge on solid ion-support (hereafter we term non-specific electrostatic interaction). This apparently hinders the application of refolding on ionic solid support. Although Ni-NTA agarose and His_6 tag can be employed for a good solid-phase scheme as well as for excellent purification, the

weak nature of Ni-attachment to a solid support might be sometimes problematic. For example, a strong reducing agent like dithiothreitol (DTT) or chelating agent like EDTA cannot be used because they might reduce the nickel ions and thereby prevent them from binding His_6-tagged proteins. High concentrations of buffer components containing strong electron-donating groups (e.g., NH_4^+), or amino acids such as arginine, glutamine, glycine, or histidine should be also avoided. This might have inhibited a broad application of Ni-agarose especially to the solid-phase refolding, because reduction or oxidation of cysteine residues and the addition of above components are sometimes indispensable.

A good alternative for the solid support must be a charged matrix, an ion-exchanger. In fact, the denatured α-glucosidase fused to a 6-arginine tag, which is bound to heparin-sepharose, has been renatured [35]. Renaturation, under conditions at which the protein was still bound to the matrix electrostatically, was shown to be attractive because it yielded high refolding efficiency even a g/L range of protein concentration. As discussed above, the polycationic fusion tag provides an efficient purification scheme, too. It is noted that the 10-lysine tag provided much better feasibility for the purification than the 6-arginine tag [25]. The 10-lysine tag had stronger binding force than the 6-arginine tag that enabled much better purification yield and purity.

The same polycation tail attached to CGTase has been shown to hold versatility for a solid-phase refolding method that utilizes a charged adsorbent as a supporting material (Biotechnology Progress, in press). CGTase fused with 10 lysine residues at the C-terminal (CGTK10ase) retained an ability to bind to a cation exchanger even at urea-denatured state. When the denatured and adsorbed CGTK10ase was induced to refold, the bound CGTK10ase aggregated little even at a g/L range. The renatured CGTK10ase could be simply recovered from the solid support by adding high concentration of NaCl. The CGTK10ase refolded on a solid support retained specific enzyme activity virtually identical to the native CGTK10ase. Several factors that were thought to be important in improving the refolding efficiency were explored. Experimental results indicated that non-specific electrostatic interactions between the charge of ion-exchanger and the local charge of CGTase other than the polycationic tag should be reduced to obtain higher refolding yield. The solid-phase refolding method involving a strong polycationic tag resulted in a remarkable increase in the refolding performance.

Conclusions

We have demonstrated experimentally that expression of a target protein fused with polycationic amino acid residues provides an efficient protein preparation scheme. CGTase fused with strong polycation could be purified to its near homogeneity from *E. coli* cell extract by using simple ion-exchange chromatography (Figure 1, left panel). Further, it has been shown that inclusion bodies can be easily refolded by using a solid-state refolding scheme, which may be applicable to other industrial enzymes including CGTase (Figure 1, right panel). Thus, it is likely that the polycationic tag is a versatile tool for the purification and refolding of

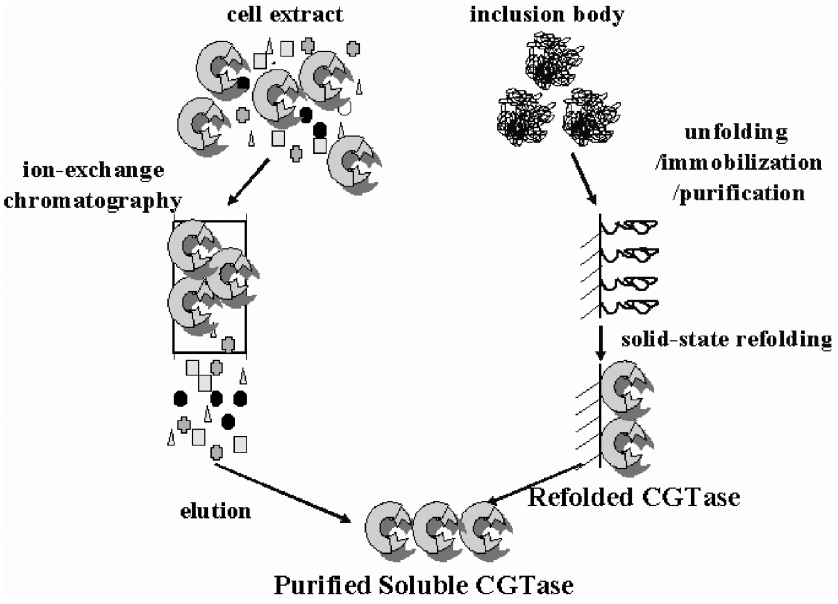

Figure 1. Application of polycation tag for production of CG Tase. Left panel: Purification of CGTK10ase (CGTase fused with 10 lysine residues) using ion-exchange chromatography. Right panel: Solid-state refolding scheme involving a polycationic tag.

recombinant proteins, leading to the cost-effective production of CGTase.

Acknowledgements

This work was supported by Center for Advanced Bioseparation Technology at Inha University and Ministry of Commerce, Industry and Energy.

References

1. A. Tonkova: *Enzym. Microb. Technol.* **22**, 678 (1998).
2. J.A. DePinto and L.L. Campbell: *Biochemistry* **7**, 114 (1968).
3. S. Kitahata and S. Okada: *Agric. Biol. Chem.* **38**, 2413 (1974).
4. R. Jamuna, N. Saswathi, R. Sheela, and S.V. Ramakrishna: *Appl. Biochem. Biotechnol.* **43**, 163 (1993).
5. S.V. Ramakrishna, N. Saswathi, R. Sheela, and R. Jamuna: *Enzym. Microb. Technol.* **16**, 441 (1994).
6. L.B. Alves, G. Matioli, F.F. De Moraes, G.M. Zanin, and J.E. Olivo: *J. Incl. Phenom.* **44**, 399 (2002).
7. G., Georgiou: In J.L. Cleland and C.S. Craik (eds.), *Principles and Practice of Protein Engineering*, Butterworths, New York (1995), pp. 101–127.
8. R.C. Hockney: *Trends Biotechnol.* **12**, 456 (1994).
9. H.-H. Jin, N.S. Han, D.-H. Kweon, Y.-C. Park, and J.-H. Seo: *J. Microbiol. Biotechnol.* **11**, 91 (2001).
10. C.-I. Kim, M.-D. Kim, Y.-C. Park, N.S. Han, and J.-H. Seo: *J. Microbiol. Biotechnol.* **10**, 632 (2000).
11. Y.C. Park, C.S. Kim, C.-I. Kim, K.-H. Choi, and J.-H. Seo: *J. Microbiol. Biotechnol.* **7**, 323 (1997).
12. A. Mukhopadhyay: *Adv. Biochem. Eng. Biotechnol.* **56**, 61 (1997).
13. A.P. Middelberg: *Trends Biotechnol.* **20**, 437 (2002).
14. K. Terpe: *Appl. Microbiol. Biotechnol.* **60**, 523 (2003).
15. E.R. LaVallie and J.M. McCoy: *Curr. Opin. Biotech.* **6**, 501 (1995).
16. S.J. Brewer and J.M. Sassenfeld: *Trends Biotechnol.* **3**, 119 (1985).
17. H. Dalbøge, H.H.M. Dahl, J. Pedersen, J.W. Hansen, and T. Christensen: *Biotechnology* **5**, 161 (1987).
18. J. Zhao, C.F. Ford, C.E. Glatz, M.A. Rougvie, and S.M. Gendel: *J. Biotechnol.* **147**, 273 (1990).
19. M.Q. Niederauer, I. Suominen, M.A. Rougvie, C.F. Ford, and C.E. Glatz: *Biotechnol. Prog.* **10**, 237 (1994).
20. M.H. Heng and C.E. Glatz: *Biotechnol. Bioeng.* **42**, 333 (1993).
21. M.H. Heng and C.E. Glatz: *Biotechnol. Bioeng.* **44**, 745 (1994).
22. I. Suominen, C. Ford, D. Stachon, H. Heimo, M. Niederauer, H. Nurmela, and C. Glatz: *Enzym. Microb. Technol.* **15**, 593 (1993).
23. J.R. Luther and C.E. Glatz: *Biotechnol. Bioeng.* **44**, 147 (1994).
24. J.R. Luther and C.E. Glatz: *Biotechnol. Bioeng.* **46**, 62 (1995).
25. D.-H. Kweon, D.-H. Lee, N.-S. Han, C.-S. Rha, and J.-H. Seo: *Biotechnol. Prog.* **18**, 303 (2002).
26. A. Mitraki and J. King: *Biotechnology* **7**, 190 (1989)
27. E.D.,Clark: *Curr. Opin. Biotechnol.* **12**, 202 (2001).
28. B. Batas and J.B. Chaudhuri: *Biotechnol. Bioeng.* **50**, 16 (1996).
29. B. Batas, C. Schiraldi, and J.B. Chaudhuri: *J. Biotechnol.* **68**, 149 (1999).
30. A. Patchornik and M.A. Kraus: *Pure Appl. Chem.* **43**, 503 (1975).
31. C.J. Epstein and C.B. Anfinsen: *J. Biol. Chem.* **237**, 2175 (1962).
32. V.G. Janolino, M.X. Sliwkowski, H.E. Swaisgood and H.R. Horton: *Arch. Biochem. Biophys.* **191**, 269 (1978).
33. T.E. Creighton: In D.L. Oxender (eds.), *Protein Structure*, *Folding*, and Design, Alan R. Liss, NewYork. (1986), pp. 249–257.
34. A. Negro, M. Onisto, L. Grassato, C. Caenazzo, and S. Garbisa: *Protein Eng.* **10**, 593 (1997).
35. G. Stempfer, B. Höll-Neugebauer, and R. Rudolph: *Nat. Biotechnol.* **14**, 329 (1996).