Improved recovery of active GST-fusion proteins from insoluble aggregates: solubilization and purification conditions using PKM2 and HtrA2 as model proteins

Dae-Wook Park1, Sang-Soo Kim1, Min-Kyung Nam1, Goo-Young Kim1, Jungho Kim2 & Hyangshuk Rhim1,*
1Department of Biomedical Sciences, College of Medicine, The Catholic University of Korea, Seoul 137-701, 2Department of Life Science, Sogang University, Seoul 121-742, Korea

The glutathione S-transferase (GST) system is useful for increasing protein solubility and purifying soluble GST fusion proteins. However, purifying half of the GST fusion proteins is still difficult, because they are virtually insoluble under non-denaturing conditions. To optimize a simple and rapid purification condition for GST-pyruvate kinase muscle 2 (GST-PKM2) protein, we used 1% sarkosyl for lysis and a 1 : 200 ratio of sarkosyl to Triton X-100 (S-T) for purification. We purified the GST-PKM2 protein with a high yield, approximately 5 mg/L culture, which was 33 times higher than that prepared using a conventional method. Notably, the GST-high-temperature requirement A2 (GST-HtrA2) protein, used as a model protein for functional activity, fully maintained its proteolytic activity, even when purified under our S-T condition. This method may be useful to apply to other biologically important proteins that become highly insoluble in the prokaryotic expression system.

INTRODUCTION

Proper systems are necessary for rapid and high-yield purification of functionally active proteins to investigate the structures and functions of proteins using biochemical and biophysical methods. A method frequently used for this purpose is the glutathione S-transferase (GST) fusion system introduced by Smith and Johnson (1). The GST-fusion system is an affinity protein purification system, which can inducibly express GST-fusion proteins in E. coli and purify the fusion proteins in their native forms by applying GST binding characteristics to glutathione (2, 3). Because this system can only be used to purify soluble proteins, the insolubility of many GST-fusion proteins still remains a major limitation with this approach (4-7).

To overcome this limitation, methods for improving protein solubility have been designed using various reagents that remove the interactions involved in the formation of protein aggregates during purification steps of proteins expressed in a prokaryotic system (8-10). An example of such a method is purification of the chicken pyruvate kinase muscle 2 (cPKM2) recombinant protein with amino acids 17-476 (7). This study has shown that the insoluble GST-fusion protein can be solubilized using the alkyl anionic detergent N-lauroylsarcosine sodium salt (sarkosyl), and the binding affinity of the sarkosyl-solubilized GST-fusion protein to glutathione beads was improved by subsequent treatment with nonionic detergent Triton X-100. Nevertheless, in practice, technical hurdles still exist during the purification of primarily insoluble proteins under non-denaturing conditions.

In the present study, the full-length mouse PKM2 and the serine protease high-temperature requirement A2 (HtrA2) were selected as model proteins for protein purification and functional activity tests, respectively. The purification conditions of these GST-fusion proteins were optimized by various combinations of sarkosyl and Triton X-100 at different ratios, and the proteolytic activity of the serine protease HtrA2 was analyzed to determine the effects of the ionic detergent sarkosyl on its activity. Our study shows that the S-T purification method not only solubilizes insoluble proteins expressed in E. coli, thereby facilitating the purification of a large amount of proteins with high yields, but also is an optimal method with no adverse effects on protein function.

RESULTS AND DISCUSSION

Selection of a challenging model protein for purifying proteins insoluble during prokaryotic expression

Protein insolubility observed during protein expression and purification is a limitation in the purification of a large quantity of soluble proteins (11-15). To overcome this limitation, the protein solubilization and purification conditions were optimized to increase the solubility and binding affinity of GST-fusion proteins to glutathione beads in their native forms. Based
on a previous study for purifying the cPKM2 (aa 17-476) protein purification study, we selected the mouse full-length PKM2 (aa 1-531) as a model protein, which is supposed to be present mostly in the insoluble fraction during conventional purification (7, 16). First, we compared the amino acid sequence homology of PKM2, cPKM2, and human PKM2 (hPKM2) (Supplementary Table 1, Supplementary Fig. 1A). The sequence identity of cPKM2 was almost 85% to those of both hPKM2 and PKM2, whereas PKM2 revealed 96% sequence identity to hPKM2. This alignment result suggests that the structure and function of hPKM2 may be much closer to those of PKM2 than to those of cPKM2; thus, a purification system for PKM2 must be developed to provide a useful reagent for characterizing the common biochemical properties of the PKM2 proteins derived from different species. Subsequently, to examine the correlation between amino acid composition of the proteins and protein solubility, we analyzed the distribution of polar and nonpolar amino acid residues on both the PKM2 and cPKM2 proteins (Supplementary Fig. 1B). The protein hydrophobicity was presented as the Kyte-Doolittle scale by the CLC bio program, which is a widely applied scale for delineating hydrophobic characteristics of proteins. Positive and negative values represent the hydrophobic and hydrophilic properties of proteins, respectively (17). The N-termini of both the PKM2 and cPKM2 proteins (from the N to A2 domains) showed almost similar hydrophobicity, whereas amino acid clusters (blue box, aa 388-414, aa 468-480) with higher hydrophobicity existed in the C domain of PKM2. This finding as well as that of the previous study suggests that PKM2 can be used as a model protein to solve protein insolubility problems in the prokaryotic expression system and to establish optimal conditions for protein purification.

PKM2, as the model protein, shows high insolubility in the prokaryotic expression system

To investigate whether the PKM2 protein is highly insoluble under non-denaturing conditions, we expressed PKM2 as a GST-fusion protein in *E. coli* and lysed cells with a typical EBC lysis buffer containing the nonionic detergent 0.5% NP40 (Supplementary Fig. 2). The GST-PKM2 protein with a molecular weight of approximately 85 kDa was expressed at high levels (40 mg/L *E. coli* culture, T lane). Total cell lysates (T) were separated by centrifugation as precipitates (P, NP40-insoluble precipitate plus some cell debris) and supernatants (S, NP40-soluble fraction), and the protein samples were analyzed on a 13% SDS-polyacrylamide gel to assess protein solubility. Over 95% of the total GST-PKM2 (38 mg) was present in the NP40-insoluble fraction (P), indicating that high levels of GST-PKM2 are virtually insoluble under non-denaturing conditions. Additionally, a decrease in protein solubility may be due to the amino acid clusters within the C domain of PKM2. Thus, GST-PKM2 can be used as an insoluble model protein to establish the optimal conditions for purifying proteins that become highly insoluble under non-denaturing conditions.

A dramatic increase in the solubility of the GST-PKM2 protein using the alkyl anionic detergent sarkosyl

The GST portion of the fusion protein improves the solubility of target proteins. Nevertheless, many GST-fusion proteins tend to aggregate rapidly and often become insoluble when overexpressed in *E. coli* and purified under non-denaturing conditions (Supplementary Fig. 2) (10, 18, 19). Generally, it is necessary to break interactions, such as hydrophobic interactions, hydrogen bonds, van der Waals forces, and ionic interactions, to improve protein solubility involved in protein aggregation (20-22). Compounds used for protein solubilization include chaotropic agents, detergents, reducing agents, and salts (20, 23). Notably, detergents are well-established and useful reagents to disrupt the hydrophobic interactions present in protein aggregates and thus to increase protein solubility. NP40 and Triton X-100 are nonionic detergents with a phenyl ring between an uncharged alkyl chain (hydrocarbon straight chain) and a hydrophilic ether group (23, 24) (Supplementary Fig. 3). These nonionic detergents have been used widely to purify proteins into their biologically active forms without denaturation. Several ionic detergents have detrimental properties on proteins, such as denaturation, and substantially destroy protein function. Nevertheless, they could be widely used to solubilize insoluble protein aggregates (7, 10, 25-27).

We lysed *E. coli* cells with various concentrations of alkyl anionic sarkosyl consisting of an anionic head group and a hydrocarbon straight chain to improve protein solubility by disrupting hydrophobic interactions present in the NP40-insoluble GST-PKM2 protein aggregates (Fig. 1, Supplementary Fig. 3). In contrast to the highly insoluble properties of GST-PKM2 under the sarkosyl-free condition, an increase in the sarkosyl concentration dramatically increased the GST-PKM2 solubility. In 0.5% sarkosyl, 75% of all of the GST-PKM2 was present in the insoluble fraction, and 24% was present in the soluble fraction (Fig. 1A). Notably, the levels of 1%- and 2%-sarkosyl soluble GST-PKM2 were 74% (29.6 mg/L culture) and 92% (36.8 mg/L culture) of the total GST-PKM2, respectively (Fig. 2B). The results demonstrate that the solubility of GST-PKM2 increased up to 23 times in the presence of sarkosyl compared to that under sarkosyl-free conditions. This sarkosyl-solubilization method is rapid, simple, and efficient for recovering GST-PKM2 at high levels.

Purification of the GST-PKM2 protein using Triton X-100

The GST-fusion protein can be readily purified by affinity chromatography using glutathione Sepharose 4B under non-denaturing conditions (28-30). After the cell lysates were solubilized with 2% sarkosyl, the protein samples were further diluted in STE buffer to a final concentration of 0.008% sarkosyl to reduce the detrimental effects of ionic detergents on the proteins. Most of the GST-PKM2 was still in the soluble fraction; however, only 2.3% of the total solubilized GST-PKM2 protein was bound to the glutathione beads (Fig. 2, 0% Triton X-100). The final concentration of sarkosyl was further diluted...