Journal of Experimental Microbiology and Immunology (JEMI) Vol. 16: 79 – 84

Production of a Recombinant Vector to Enable the Study of Thioredoxin Function as a Bound or Detached Solubilizer of Proteinase Inhibitor 2 in a Bacterial Protein Overexpression System

Chris Duronio
Department of Microbiology & Immunology, University of British Columbia

A useful method of preventing formation of insoluble aggregates during bacterial protein overexpression is to fuse a reductase enzyme to the expressed protein. One example of this is the pET system from Novagen. The pET32a plasmid is an expression vector that fuses insert gene sequences to thioredoxin. This construct has been shown to reduce protein aggregates even when not bound to the protein. In this study, a vector with the fusion tag removed and a fully separate thioredoxin reinserted was created to enable expression of both proteins separately from the same plasmid. This vector is based on a preexisting plasmid with the thioredoxin tag removed and a proteinase inhibitor 2 gene inserted. The thioredoxin sequence along with the appropriate promoter and ribosome-binding site from the commercial pET32a plasmid were isolated. This difficult procedure was followed by the insertion of the sequence into the previously created vector. This study will require further work to get a conclusive result, but potential candidates were isolated.

A common method of obtaining large quantities of protein efficiently is the bacterial overexpression system. This system allows the generation of clones that contain a desired protein sequence that can be induced to unnaturally high levels by addition of a compound specific to the system. This is inexpensive and quick, but can occasionally lead to complications with the proteins that have been produced. One problem is the formation of insoluble aggregates of the expressed protein that inhibit proper isolation of the functional units. These aggregates are called inclusion bodies, and their cause is related to a stress response similar to that of heat shock: heat shock proteins are produced. This results in the cell processing the recently created and incomplete new protein as a temporary measure to allow later processing and prevent misfolding (7). This is seen even more frequently when using bacterial overexpression systems for production of eukaryotic proteins. Several methods have been used to combat the formation of the inclusion bodies, most involve a spin step to isolate the aggregates, a denaturing or unfolding step, then a final renaturing or folding step (4). These are all done post isolation from the bacteria, and are quite time-consuming to carry out. More recent methods of reducing inclusion body formation involve tagging a small sequence onto the end of the recombinant to express a protein that will in some way change the solubility of the protein or has some kind of activity that interferes with the aforementioned aggregate accumulation. One such system is the pET32 expression vector line from Novagen (5). In this system, the expressed protein is bound to the common redox enzyme thiol reductase, also known as thioredoxin. This enzyme has many functions in the cell, including participating as a cofactor for phage T7 DNA polymerase and metabolism of sulphate, but most importantly for this scenario there is the ability to reduce disulphide linkages in proteins (5).

Thioredoxin function has been well studied: when fused to the expression protein, it increases solubility and therefore reduces inclusion body formation (3). The exact method by which thioredoxin has the ability to change solubility of the proteins is not perfectly understood, but it has been shown that thioredoxin, when overproduced on a separate vector, allows for some increase in solubility of the protein (9). This suggests that at least part of the activity of thioredoxin as a fusion partner is due to the redox function of the enzyme, and not merely the presence of the fused tag somehow preventing inclusion body formation as a physical restraint. What has not been well studied is the degree to which a thioredoxin enzyme, when produced at equal rates to its partner protein, is able to prevent inclusion body formation relative to the production of a

thioredoxin and target protein fusion construct. Such a system, if able to prevent inclusion body formation easily while not modifying the target protein with any type of fusion, would allow for an easier system of protein overexpression while still maintaining the use of one vector, not two as has been used in prior research. This system would be particularly useful for production of large amounts of artificially expressed protein in both commercial and research settings. Therefore the purpose of this study was to create a vector that contains a thioredoxin that can be induced by the same system as another protein on the same plasmid.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions: *E. coli* DH5α strains containing JP061, JP062, JP063, JP064 vectors were obtained from stocks frozen by Jei Eung Park in 2006 (6). *E. coli* DH5α strain containing pET32a was obtained from the MICB 421 culture collection. Figure 1 shows the restriction map of wildtype pET32a. *E. coli* DH5α strains were cultured in Luria-Bertani (LB) broth (10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride) or on LB agar plates (LB broth with 15g/L agar added). Each had 100 ug/ml Ampicillin added. The liquid cultures were grown up at 37°C, and shaken at approximately 150-200 rpm. Purified plasmids were prepared using the Invitrogen PureLink HQ Mini Plasmid Purification Kit (Cat# K2100-01) or PureLink HiPure Midiprep Plasmid Kit (Cat# K2100-04).

Confirmation of Identity of *E. coli* DH5α: JP061, JP062, JP063, JP064, possible pET32aΔTrx/PI2$^+$ and pET32a/PI2$^+$ Clones: In a previous study by Park (6), 4 potential clones were frozen for further analysis and screening. Of the 4 strains, JP061 and JP063 were expected to be pET32a/PI2$^+$, while JP062 and JP064 strains were expected to be pET32aΔTrx/PI2$^+$. To confirm the identity of the clones, the 4 strains were grown up and their plasmids isolated. From these, a double digest of approximately 300 ng of plasmid DNA was cleaved with 20 units of New England BioLabs EcoRI (Cat# R0101S) and 10 units of New England BioLabs NcoI (Cat# R0193S) with 5 ul of 10x NEBuffer EcoRI (Cat# B0101S). Reactions were performed in 50 ul final volume for 90 minutes at 37°C. At the end of the digestion, samples were added with 6x Gel Loading Buffer and resolved on a 1% agarose gel (Agarose 3:1, Amresco, Cat# E776-250) in 0.5x Tris-Borate-EDTA (TBE) at 110 V.

Isolation of Trx Sequence from *E. coli* DH5α:pET32a: Upon confirmation of the identity of the JP062 and JP064 strains as pET32aΔTrx/PI2$^+$, the aforementioned *E. coli* DH5α:pET32a clones were grown up and the plasmids were isolated as mentioned previously. From these, a double digest of approximately 1000 ng of plasmid DNA was cleaved with 30 units of Fermentas KpnI (Cat# ER0521) and 30 units of New England BioLabs SgrAI (Cat# R0603S) with 2.5 ul of 10x NEBuffer I (Cat# B7001S) and 2.5 ul of 10x Invitrogen REact 4 Buffer (Cat# 16304-016). Reactions were performed in 40 ul final volume for 150 minutes at 37°C. At the end of the digestion, samples were added with 6x Gel Loading Buffer and resolved on a 1% agarose gel (Agarose 3:1, Amresco, Cat# E776-250) in 0.5x Tris-Borate-EDTA (TBE) at 110 V. Under UV, an appropriate fragment was cleaved out and isolated from gel via Sigma GenElute Gel Extraction system (Cat# NA1111-1KT) and eluted in 40 ul final volume.

Construction of pET32aΔTrx/PI2$^+$Trx$^+$ vector by cloning of Trx into pET32aΔTrx/PI2$^+$: First step to build this vector was to cleave the appropriate vector (from the previously isolated JP062 strain) at a single cleavage site. For this, a digest was performed on 300 ng of DNA, cleaved with 10 units of New England BioLabs SgrAI (Cat# R0603S), with 5 ul of 10x NEBuffer I (Cat# B7001S) for 90 minutes at 37°C in a final volume of 50 ul. Next, previously isolated Trx fragment was ligated into this newly cut vector. This was done by ligation of 9 ul of insert to 3 ul of vector with 1 unit of T4 DNA Ligase (Invitrogen, Cat# 15224-017) with 4 ul of 5x DNA Ligase Buffer (Invitrogen, Cat# 46300-018). Final reaction volume was 20 ul and incubation was 120 minutes at 22°C.

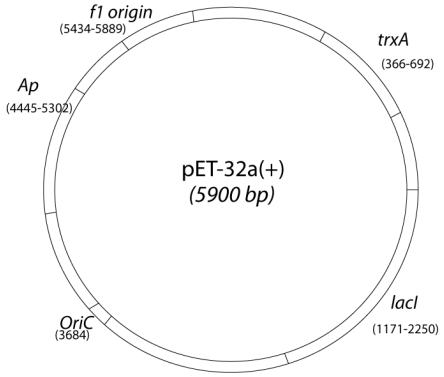

FIG. 1. The restriction map of wildtype pET32a. Multiple cleavage site discussed throughout Materials and Methods section are shown above, including SgrAI at 840 bp, KpnI at 238 bp, RsrII at 589 bp, and EcoRI at 192 bp.

Transformation of *E. coli* DH5α cells with vector and screening for potential clones of new vector CD11n: Newly created vector solution was transformed into an electrocompetent *E. coli* DH5α strain from the MICB 421 culture collection. This procedure was done using a BioRad electroporator, in a 0.2 cm electroporation vial. 6 ul of vector was added to 50 ul electrocompetent *E. coli* DH5α cells in each vial (4 replicates). Pulse on setting Ec2 was performed, and the average readout was 2.51 kV, for an average of 2.15 ms, with no extreme variability. 250 ul of SOC medium was added immediately and the samples were incubated for 90 minutes at 37°C with gentle shaking. Screening of colonies was done by random picking from suitable plates, grown up, then cleaved via a double digest of approximately 1500 ng of plasmid DNA was cleaved with 15 units of New England BioLabs EcoRI (Cat# R0101S) and 10 units of New England BioLabs RsrII (Cat# R0501S) with 5 ul of 10x NEBuffer 4 (Cat# B7004S). Reactions were performed in 50 ul final volume for 90 minutes at 37°C. At the end of the digestion, samples were added with 6x Gel Loading Buffer and resolved on a 1% agarose gel (Agarose 3:1, Amresco, Cat# E776-250) in 0.5x Tris-Borate-EDTA (TBE) at 110 V.

RESULTS

Confirmation of Identity of *E. coli* DH5α: JP061, JP062, JP063, JP064, possible pET32aΔTrx/PI2$^+$ and pET32a/PI2$^+$ Clones: The restriction map of JP061, JP062, JP063, and JP064 clones following NcoI and EcoRI digestion is shown below in Figure 2. The possible results were a 5900 bp and 590 bp fragment from each of the pET32a/PI2$^+$ clones, and a 5600 bp and 590 bp fragment from the

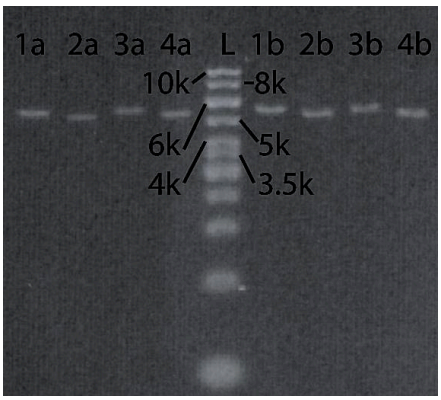

FIG. 2. EcoRI and NcoI restriction digest of potential clones. In this gel, 2 replicates were performed for each sample, and are distinguished from each other via the "a" or "b" designation. L refers to the ladder, and the numbers simply refer to the ending number of the code of the potential clones (JP061 is referred to by 1, and so on). The ladder contains 3 ul of GeneRuler 1 kb Ladder from Fermentas. Important molecular weight sizes of standards are referred to in short form as 10k meaning 10,000 bp molecular weight, and so on.

pET32aΔTrx/PI2^{+}clones. It was expected that the JP061 and JP063 clones would be pET32a/PI2^{+}, and therefore had a 5600 bp fragment, while the JP062 and JP064 clones would be pET32aΔTrx/PI2^{+}, and therefore had a 5900 bp fragment. The difference in size was due to the cleavage of the *trxA* sequence (the Trx gene).

As seen in Figure 2, there were two different sizes of cleaved vectors: one at approximately 6000 or just under, and the other at somewhere between 5000 and 6000. Therefore it is inferred that lanes 1 and 3 contained the 5900 bp fragments, while lanes 2 and 4 contained the 5600 bp fragments. This confirmed the prior expectations of the previous study. The 590 bp bands that were expected did not show as they were too faint on the gel after running and too close to the edge of the gel to photograph clearly. More importantly, the bands that were shown were enough to differentiate between the two different clones and allow for further research with them, without requiring further selection. From this gel, the decision was made to choose the JP062 strain as the clone to use and follow up on. It appeared to be identical to the JP064 strain but was chosen because it appeared to be a cleaner sample.

Construction of pET32aΔTrx/PI2^{+}Trx^{+} by cloning of *trxA* into pET32aΔTrx/PI2^{+}: In this section, there are several figures to discuss, the first of which is Figure 3, shown above. This result was from a screen of potential clones after the transformation of the newly constructed vector into electrocompetent *E. coli* DH5α cells. First item to note from this figure was the possible contaminating DNA in the JP062 control well. However the band that was most strongly there is the 5600 bp band, which was the same sample as the cleaved by double digest vector that was discussed earlier and used to identify potential clones from the frozen cell stocks. There appeared to be 4 or 5 potential clones that were larger than the control. This was a necessity as there should have been a 490 bp fragment added into the JP062 from the *trxA* sequence. The ones that definitely were larger were in lanes 4, 5, 7, and 10 and the one that appeared to be approximately the same size was in lane 1. The clones in lanes 4, 5, 7, and 10 will henceforth be referred to as clones CD111, CD112, CD113, and CD114. Each of these will be considered a potential clone with the identity pET32aΔTrx/PI2^{+}Trx^{+} and should be identified as the CD11n series. The next stage of this set of results is shown below in Figure 4.

The above figure shows the screen of the 4 potential clones pET32aΔTrx/PI2^{+}Trx^{+} after double digest with RsrII and EcoRI. The size of the complete plasmid should be approximately 6600 bp, but the removal of approximately 1000 bp (in a separate

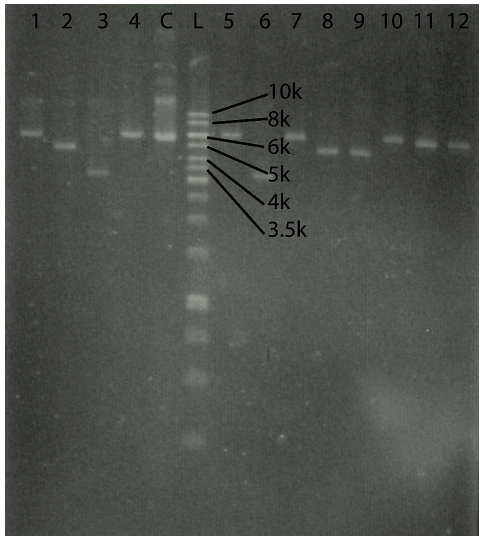

FIG. 3. 1.0% Agarose gel showing screen of 12 potential clones. In this gel, there are 12 single replicates of 12 separate potential clones in the numbered lanes. They have been run uncut. L refers to the ladder, and C refers to control, which in this case is the cleaved JP062 E. coli from previous experiments. The ladder contains 3 ul of GeneRuler 1 kb Ladder from Fermentas. Important molecular weight sizes of standards are referred to in short form as 10k meaning 10,000 bp molecular weight, and so on.

fragment, too faint for photographing) should yield around 5600-5700 bp. The fragments appear to be slightly more than the double-digest cleaved JP062 plasmid, which has been shown to be around 5600 bp. See Figure 5, below, showing the potential completed vector CD11n.

DISCUSSION

The identification of potential clones from the frozen stocks of JP061, JP062, JP063 and JP064 was straightforward. It was based on the work of Jei Eung Park and his predecessors from the student lab over a number of years that are not cited here. The larger bands in this gel photo (Fig. 2) are very clearly larger or smaller, based on the expected size of the cleaved out *trxA* gene. Potential problems seen at this stage were the error in allowing the smaller bands (approximately 590 bp) to run to the end of the gel or run so far that they became so blurred as to be nearly invisible in the photograph. This makes the results not as clear or as certain as they would be. Proper controls were also not included at this stage, such as an uncut sample, which would help to alleviate concerns about activity of the restriction enzymes in the double digest. Given that the sizes are different, it shows absolutely that the *trxA* gene was removed, however. It would have needed to have controls and additionally, to be run at a higher concentration in order to show the smaller bands as a final confirmation. The isolation of the *trxA* fragment was extremely difficult (data not shown). The difficulty of successfully isolating the sought-after fragment of *trxA* with its own T7 promoter site and ribosome binding site (rbs) was exacerbated by the choice of using a double digest to cleave out the fragment from the plasmid, then running an agarose gel and purifying the fragment from there. This yielded very little DNA at the end, and was therefore very difficult to isolate. The results from this section were not photographed because of the necessity of the removal of the fragment from the gel immediately upon viewing on the transilluminator. Additionally, the figure would have appeared inconclusive with the resulting image, as the band was very faint and would not photograph well. This part could have been improved by the use of PCR with appropriate primers to isolate the fragment required in much higher concentration. This was the most time-consuming step for the research, as it took many attempts to get a visible band of *trxA* from the gel after double digest. This required double digests of very high concentration DNA in order to achieve this, as the smaller fragment was only seen at approximately 10% of the intensity of the larger band on the gel due to the much smaller size.

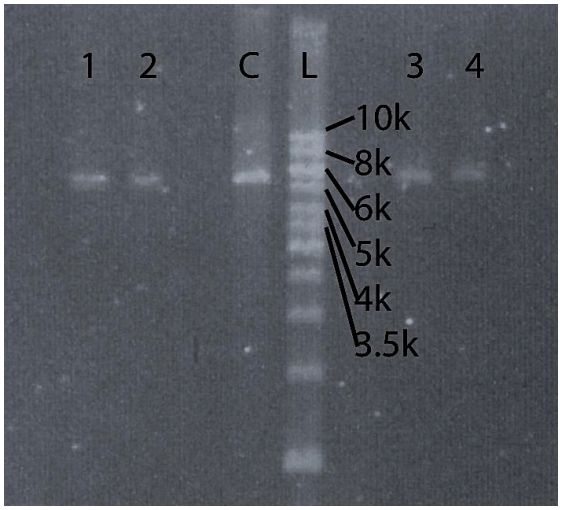

FIG. 4. EcoRI and RsrII restriction digest of potential clones. In this gel, each of the 4 potential clones designated CD111, CD112, CD113, CD114 plasmids were cleaved via double digest. L refers to the ladder, C refers to the control, which in this case is cleaved JP062 plasmid which should show up as a band at approximately 5600 bp. The numbers simply refer to the ending number of the code of the potential clones (CD111 is referred to by 1, and so on). The ladder contains 3 ul of GeneRuler 1 kb Ladder from Fermentas. Important molecular weight sizes of standards are referred to in short form as 10k meaning 10,000 bp molecular weight, and so on.

The screening of the potential clones after the transformation with the ligated plasmid shows the extreme variability in the plasmids found within the potential clones that were screened (Fig. 3). Another area of concern from this gel was the large band at over 10,000 bp (estimation was approximately 12,000 bp) in the control lane. A possible explanation for this is that the cut plasmid that was frozen down several weeks before formed some dimers that would have approximately that size. Additional concerns involve the aforementioned variation in band size, which suggests that there could have been some degradation of genetic material at some stage of the plasmid isolation or that there is perhaps a heterologous mix of the starting plasmids that were found in the bacteria used in the study. There were many similarities between what occurred in this experiment and those that were the basis for this study, some of which was done with the insertion of proteinase inhibitor 2, or PI2, into the pET32a vector, with the eventual goal of replacing the thioredoxin, or *trx* with flavin oxidoreductase, or *fre* (6). This research was based on the previous work of

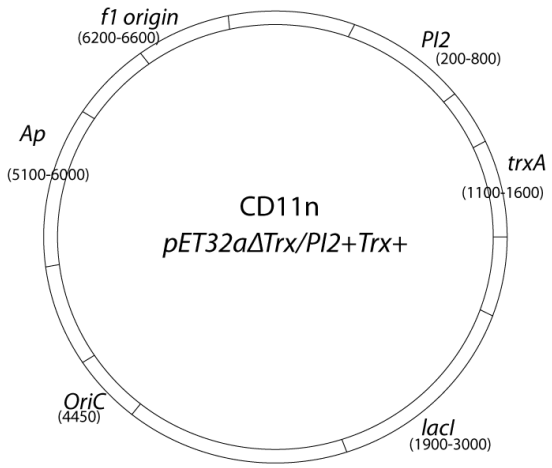

FIG. 5. Map of pET32aΔTrx/PI2^{+}Trx^{+} vector, designated CD11n. This figure shows approximate locations of important genes on the new, potentially created vector. The approximate total size is 6600 bp, and the labels below each gene signify distance in bases from the origin, which in this case is the f1 origin, as in Figure 1. Important to note that the *trxA* regions in figures 1 and 5 are different, as the one in figure 1 is for fusion to the PI2, whereas in figure 5 it is entirely separate.

cleaving of the thioredoxin tag out of the pET32a vector (8). Several other projects attempted to insert EDTA Monooxygenase A (1)(2). There are similarities between this research and those of its predecessors, as it is a continuation of the Park research (6). It was taken in a different direction, however. The goal of their research was to change the tag to something else in order to study its effects instead of thioredoxin. This study was intended to act as the basis for the study of the activity of thioredoxin in order to understand how it works to solubilize the protein aggregates known as inclusion bodies. Therefore it is difficult to compare the projects except as a way of seeing how all projects have had difficulty isolating the small fragment they wish to insert and the act of insertion has shown to be challenging as well.

Creating new vectors can be very difficult, especially within the time frame available. However each researcher has gotten one step further towards an answer with regards to the mechanism and extent of solubilization of expressed proteins via the pET32a expression system. Assuming the data is correct and the proper vectors have been isolated in the *E. coli* DH5α strain, the next experiment is clear and will involve actually stimulating production of PI2 and Trx in a suitable bacteria. The difficulty with this type of experiment is that the data will not prove something, as it has not yet reached that stage. It is, however, an important foundation for further research.

FUTURE RESEARCH

The first step for follow-up experiments would be to grow up the 4 potential clones CD111, CD112, CD113 and CD114 from frozen stocks. Then there would have to be an analysis to determine if any of them are indeed the pET32aΔTrx/PI2^{+}Trx^{+} clones. The method to do so would be cleavage with RsrII and another restriction enzyme suitable to the purpose, based on the restriction map for pET32a. The reason for this is that the RsrII is only found in the *trxA* region, so if it is not found in the plasmid, it means the *trxA* gene was not correctly inserted. After this confirmation step, the next step would be to transform the vector into *E. coli* BL21(DE) cells. The other vector that should be transformed is the plasmid from the JP061 or JP063 strain into an identical set of *E. coli* BL21(DE) cells. This will provide a pET32a/PI2^{+} clone that will produce a PI2-Trx fusion protein. Following this, appropriate cultures should be grown up and induced with IPTG. This should produce a fusion protein in the JP clone and a set of two individual proteins in the CD clone, allowing for comparison to each other as far as formation of inclusion bodies. It might also be helpful to add other controls for interest and for relevance, such as the JP062/JP064 clone, which should show the PI2 protein on its own, without the tag, which would likely have a much greater incidence of inclusion body formation.

ACKNOWLEDGEMENTS

The research done for this project was done with the support of the UBC Department of Microbiology and Immunology. I would like to thank Dr. Bill Ramey for his support and direction during the difficult stages of the project. I would also like to acknowledge those students who came before me without whom I would have not been able to get to this stage.

REFERENCES

1. **Cameron, A., S. Gray, A. Sribanaia, and T. Withrow.** 2009. Attempted construction of a pET32a(+) vector containing EDTA monooxygenase A. J. Exp. Microbiol. Immunol. **13**:119-124.
2. **Kazem, M.** 2004. Cloning EDTA monooxygenase as a model protein to characterize the effects of flavin oxireductase on solubility of proteins in protein overexpression systems. J. Exp. Microbiol. Immunol. **6**:26-34.
3. **LaVallie, E.R., E.A. DiBlasio, S. Kovacic, K.L. Grant, P.F. Schendel, and J.M. McCoy.** 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E.coli* system. Bio/Technology. **11**:187-193.

4. **Misawa, S. and I. Kumagai.** 2000. Refolding of therapeutic proteins produced in *Escherichia coli* as inclusion bodies. Peptide Science. **51:**297-307.
5. **Novy, B., J. Berg, K. Yaeger, and R. Mierendorf.** 1995. Novagen in Novations. **3:**4 http://www.emdbiosciences.com/docs/NDIS/inno03-003.pdf.
6. **Park, J.E.** 2006. Generation of recombinant plasmid constructs to assess the ability of NADH:flavin oxidoreductase to solubilize proteinase inhibitor 2 in bacterial protein overexpression systems. J. Exp. Microbiol. Immunol. **10**:27-33.
7. **Villaverde, A. and M. Mar Carrio.** 2003. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotech. Letters. **25:**1385-1395.
8. **Wong, S.H.** 2005. Cloning of flavin reductase into pET32a(+) expression vector lacking the thioredoxin A tag to study solubility of EDTA monooxygenase A in overexpression systems. J. Exp. Microbiol. Immunol. **8**:59-66.
9. **Yasukawa, T., C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii.** 1995. Increase in solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. J. Biol. Chem. **270:**25328-25331.