Journal of Experimental Microbiology and Immunology (JEMI) Vol. 10:41-45

Attempts to Amplify EDTA Monooxygenase A for Cloning into the pET-32a Expression Vector

STEPHANIE YOUNG
Department of Microbiology and Immunology, UBC

One significant problem with bacterial overexpression systems is the production of inactive protein in the form of inclusion bodies. The Novagen pET-32a vector was developed as an expression system for increasing the solubility of overexpressed proteins. This system works by fusing *Escherichia coli* thioredoxin with an insoluble protein of choice. The goal of this study was to compare the effectiveness of thioredoxin and another *E. coli* redox protein, NADH:flavin oxidoreductase, in increasing the solubility of the highly insoluble protein EDTA monooxygenase A. A strategy for amplifying the EDTA monooxygenase A sequence from another vector pEmoA for its insertion into pET-32a was developed in which mutagenic primers were designed to introduce EcoRI recognition sites flanking the gene to facilitate the cloning process. Several attempts were made to amplify the fragment by PCR in which different experimental conditions were varied, but due to time constraints, the correct PCR product was not obtained. Further work must be done to optimize the PCR for the amplification of the EDTA monooxygenase A gene and to generate the constructs necessary to compare the effectiveness of thioredoxin and flavin oxidoreductase in increasing protein solubility.

Protein overexpression systems in bacteria such as *Escherichia coli* are widely used as a method of producing large quantities of recombinant protein. A major problem with these bacterial systems is the accumulation of insoluble protein aggregates known as inclusion bodies upon high levels of expression that render the protein biologically inactive (10). The recovery of functionally active product from these inclusion bodies involves a complex process of solubilizing, refolding, and purifying the protein (10). To avoid this cumbersome procedure, a novel approach was developed in which *E. coli* thioredoxin (TrxA) is used as a gene fusion partner. This Trx gene fusion expression system has been shown to increase the solubility and biological activity of several mammalian cytokines and growth factors expressed as thioredoxin fusion proteins (6). Other studies have also shown the effectiveness of this system. Yasukawa *et al.* compared the solubility of eight vertebrate proteins, including mouse c-Myb, cAMP response element-binding protein 1 (CRE-BP1), the p53 tumor suppressor gene product, and the Ser/Thr kinases Mos and Lck when expressed with and without Trx. While the proteins formed insoluble aggregates in the absence of TrxA, the solubility of all eight proteins was increased significantly when coproduced with TrxA (13). The expression of thioredoxin as a fusion protein with procathepsin D, the precursor of the lysosomal enzyme cathepsin D, was also found to increase the yield of soluble protein (9). It is not clear why coexpression of thioredoxin as a fusion protein prevents the formation of insoluble aggregates. Since thioredoxins facilitate the reduction of other proteins by disulfide bond formation, it may be that solubility is affected by the redox state of the protein (13).

The focus of this study was to determine whether another redox protein from *E. coli*, NADH:flavin oxidoreductase (Fre), could be used instead of TrxA in this gene fusion expression system, and to compare the effectiveness of TrxA and Fre in enhancing the solubility of EDTA monooxygenase A (EmoA), a highly insoluble protein when overexpressed. Three constructs needed to be generated: one containing *emoA* alone, one containing *emoA* and *trxA* expressed as a fusion protein, and one containing *emoA* and *fre* expressed as a fusion protein. The Novagen vector pET-32a (Fig. 1) designed for the cloning and overexpression of proteins fused with TrxA was used in this study (4). Previous studies (2, 5, 12) described a strategy for excising *emoA* from pEmoA by restriction enzyme digest before insertion into pET-32a. Upon careful analysis of this approach, it was determined that in order for *emoA* to be inserted in the correct reading frame in the pET-32a vector, a different strategy was needed. An attempt was made to amplify *emoA* from pEmoA by designing sense and antisense primers and optimizing the PCR experimental conditions. To replace *trxA* with *fre*, the intention was to excise the *trxA* fragment from the pET-32a vector by restriction enzyme digest and replace it with *fre* amplified from the pES1 plasmid (7), as previously described (5, 12).

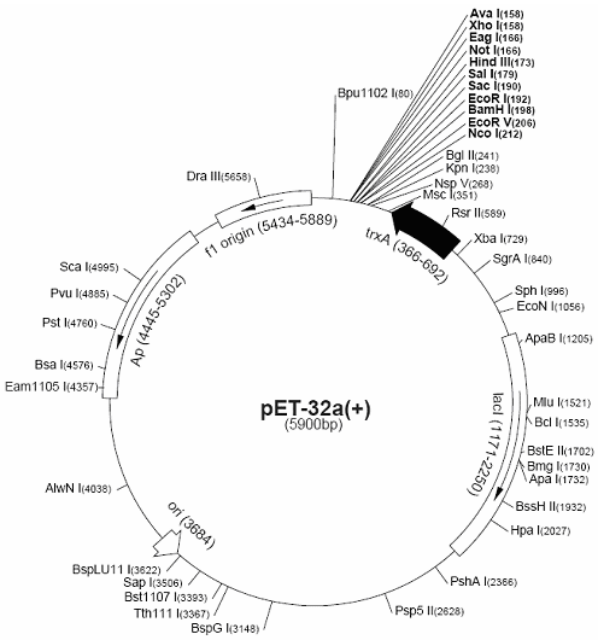


FIG. 1 pET-32a vector from Novagen (4). The EcoRI site used for inserting *emoA* is located in the multiple cloning site at position 192.

pEmoA (2.4 kb EcoRI fragment containing *emoA* in cloning vector pTrc99A)

CGATCACGCCGGTCGAACGGACCGGAATGTGGCGG**C**G**T**C**T**TTTTGGCGGCAGAACCCAGGAGGCCTC
TTA*ATGCGCAAACGCAGGATGTATCTTGTCAGCTGGCTGAACTCCTCGGGGGTTCTCCCCAATTCCTGGAAT*
GAGGGCCGGGGAAACCGGGCGAGGATCTTTGACCTGGAGAATTATATCCGCAGCGCCGAGATCGCCCGGC
GCGGCCGGATCGATGCCTTCTTTCTAGCCGACCAGCCGCAACTAACCCCCAACCCGAAGGTTCGTCCTGAGT
ATCCGTTCGACCCTATCGTGCTTGCGGCAGCGATCACAGGCCGCGTGCCAGACATCGGCGGGATTGTGACT
GCCTCCACCAGCTTCAGCCTGCCCTACACGCTCGCCAGGCAAATCGCCTCGGTGAACCTGCTTTCGGGCGG
CCGCATCGGCTGGAACGCGGTCACAACCGCCAACCCCGCCGTCGCGGCCAACTACGGCGCCGCGATTGCC
ACACATGACAATCGCTATGAGCGGGCGGAAGAGTTCCTGGAGGTTGTCCATGGCCTTTGGAACAGCTGGAAA
TTCCCGTGGGATGAGGCGATCGGCCCCAATCCAAATCCGTTCGGCGAAGTGATGCCGATCAACCATGAGGG
GAAATATTTCAAGGTTGCGGGTCCCCTCAACGTGCCTTTGCCGCCTTACGGCCCGCCGGTGGTCGTGCAGGC
AGGCGGCTCGGACCAGGGCAAAAGGCTCGCCAGCCGTTTCGGCGAGATCATCTATGCGTTCCTGGGGAGCA
AGCCGGCCGGCCGGCGCTTCGTCGCCGAGGCGCGAGCGGCCGCTCGCGCGCAGGGCCGCCCCGAGGGC
TCCACGCTGGTCCTGCCGAGCTTTGTGCCGCTGATCGGGTCGACCGAGGCAGAGGTCAAGCGGCTCGTCGC
CGAGTACGAAGCCGGGCTCGATCCGGCCGAGCAGCGGATCGAGGCGCTGTCGAAGCAATTGGGCATTGATC
TTGAAAGGATCAACGTTGACCAGGTGCTGCAGGAGAAGGACTTCAATCTTCCGAAGGAGTCCGCCACGCCGA
TCGGCATCCTGAAATCCATGGTCGACGTCGCTCTCGACGAGAAGCTTTCACTGCGGCAACTGGCTTTGCGCA
TGCGGCTGATTGCGGGCACGCCGGATCAGGTCGCCGACCGTCTCATCGATTGGTGGCAAGACGAGGCGGC
GGATGGATTTGTGATCAATGCCCCGCTGCTACCCGACGCTCTGGAGATTTTTGTCGATCAGGTCGTACCGATC
CTGCAGTCGCGAGGCGTTTTTCCCCGCAGCTACACCGAGTCGACCTTGCGTGAGAGGTTGGGTTTGCCTCGA
*AATCCGCTA*GGTTGACCAAGCGAA**AA**TCATGGTTTGTTCCACATCTGCATACATTTACAAAT

Italicized: emoA sequence
Bold: Induced mutation sites
Underlined: Primer binding region

PCR Primers

	Sequence	Length	GC%	Tm
Forward	5' – GGAATTCTTTTGGCGGCAGA – 3'	20	50	75 °C
Reverse	5' – TGGAATTCGCTTGGTCAACC – 3'	20	50	74 °C

FIG. 2 EDTA monooxygenase A (*emoA*) and PCR primer sequences. pEmoA was constructed by inserting the *emoA* gene into the cloning vector pTrc99A. The italicized region shows the *emoA* gene and the underlined regions are the binding regions for the forward and reverse primers. The forward primer contains four mutation sites and the reverse primer contains three mutation sites to introduce EcoRI recognition sites into the amplified *emoA* fragment. The expected size of the PCR product is approximately 1.3 kb.

Due to the time constraints, we were only able to work on optimizing the PCR reaction for amplifying *emoA*.

MATERIALS AND METHODS

Bacterial growth conditions and plasmid isolation. Plasmid-carrying *E. coli* strains, including pET-32a (Novagen) and pEmoA (UBC), were grown in Luria-Bertani (LB) broth (10g/L tryptone, 5g/L yeast extract, 10g/L sodium chloride) and filter-sterilized ampicillin at a concentration of 100µg/mL. Cultures were incubated at 37°C and grown overnight with shaking at 250 rpm. Plasmid DNA was isolated from these bacterial cultures using the standard protocol from the PureLink HiPure Plasmid Maxiprep Kit (Invitrogen, Cat. No. K21007).

Amplification of the *emoA* fragment from pEmoA. To amplify *emoA* from pEmoA, forward and reverse primers were designed using the software Primer Designer v. 2.01 (Scientific and Educational Software, Cary NC) and synthesized by Integrated DNA Technologies (IDT, Coralville IA). Mutagenic primers were designed to introduce EcoRI recognition sites in close proximity to the start and stop codons of the gene. This would allow for the amplified fragment to be digested with EcoRI, after which it could be more easily inserted into the pET-32a multiple cloning site. The binding sites of the forward and reverse primers in pEmoA as well as their sequences are shown in Figure 2. The Basic Local Alignment Search Tool (BLAST) network service provided on the National Center for Biotechnology Information (NCBI) website was used to ensure that the proposed primer sequences did not bind nonspecifically to any other regions in pEmoA. The primers were reconstituted in sterile water at a concentration of 100µM. Two microliters of uncut pEmoA (at a concentration of 58.95 ng/µL) was added to the PCR Master Mix (Table 1). All PCR reagents used were obtained from Fermentas. Each PCR reaction was carried out over 40 cycles in the Biometra T-gradient thermocycler. During each cycle, the DNA was denatured at 94°C, annealed to the primers at varying temperatures, and extended at 72°C (Table 1). Four separate trials were conducted in which the composition of the PCR Master Mix and the annealing temperature was varied (Table 2). Five microliters of each PCR reaction was then electrophoresed on a 1% agarose gel to determine the presence or absence of the correct PCR product.

Table 1. PCR conditions for amplification of *emoA* from pEmoA.

Master Mix		Thermocycler Parameters	
Reagents	*Volume/Reaction*	*Temperature*	*Duration*
25 µM forward primer	1 µL	1) 95°C	3 min.
25 µM reverse primer	1 µL	2) 95°C	30 sec. or 1 min.
10X Taq Buffer + $(NH_4)_2SO_4$	5 µL	3) Annealing temp.	1 min.
25 mM $MgCl_2$	1, 2, or 3 µL	4) 72°C	1 min.
25 mM dNTPs	0.5 µL	5) Return to step 2	39 cycles
Taq polymerase (5 U/µL)	0.25 µL	6) 72°C	5 min. or 10 min.
dH_2O	X µL	7) 10°C	Until end
Total volume	48 µL		

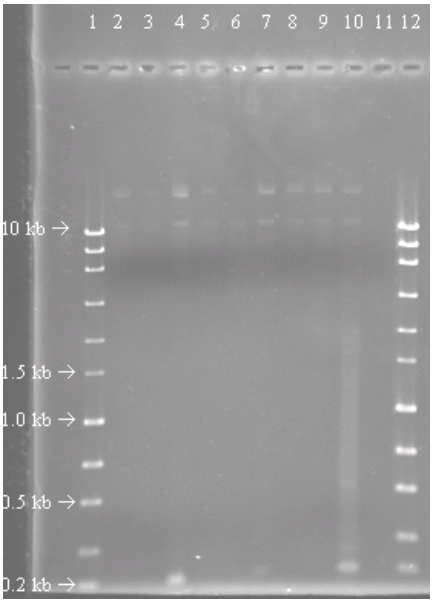
Table 2. Variations in the PCR conditions for each trial in the amplification of *emoA* from pEmoA.

Trial	Conditions		
	Amount of $MgCl_2$	*Run Cycle Parameters*	*Annealing Temperatures*
1	2 µL	Step 2 = 30 sec. Step 6 = 5 min.	51.9°C, 53.5°C, 55.3°C, 57.1°C, 58.9°C
2	2 µL	Step 2 = 1 min. Step 6 = 10 min.	51.9°C
3	1, 2, and 3 µL	Step 2 = 1 min. Step 6 = 10 min.	45.5°C, 50.9°C, 56.1°C
4	3 µL	Step 2 = 1 min. Step 6 = 10 min.	39.8°C, 59.4°C, 60.1°C, 61.9°C, 62.5°C

RESULTS

In the first two trials, no band corresponding to the expected 1.3 kb *emoA* fragment was observed in the PCR reactions (data not shown). Two large bands greater than the 10 kb marker were observed in all reactions, except for the negative control that was lacking the template plasmid, and were likely due to the existence of the plasmid in varying topologies. When the amount of $MgCl_2$ added to each reaction was varied in the third trial, a small band approximately 200 bp in size was observed in the reactions containing 3 µL of $MgCl_2$ (Fig. 3). Another faint band approximately 500 bp in size was also observed in the PCR reaction containing 3 µL of $MgCl_2$ that was subjected to the annealing temperature of 56.1°C. However, none of these bands corresponded to the expected size of the 1.3 kb *emoA* fragment. Since the higher concentration of $MgCl_2$ yielded a product, albeit a nonspecific product, the fourth trial was designed with the PCR reactions containing 3 µL of $MgCl_2$ (data not shown). This, however, did not yield any observable product. A possible reason for this observed difference between Trials 3 and 4 is that the PCR reactions from Trial 4 were electrophoresed on a gel with wider lanes, which may have caused the DNA to diffuse into a less obvious band.

For example, as shown in Figure 4, there is a region of approximately 450 bp between the EcoRI site and the start codon of *emoA* that could potentially interfere with the expression of EmoA as a fusion protein with either TrxA or Fre. In designing the mutagenic primers, several conditions were met to ensure their effectiveness. Most importantly, the mutagenic region was positioned towards the 5' end rather than the 3' end of the primer to ensure sufficient binding of the DNA polymerase when initiating elongation. The primers were also designed so that their GC content was approximately 50%, the Tm values of the forward and reverse primers were within 1°C of each other, and they did not form any hairpin loops within themselves. These conditions were all taken into consideration when designing the primers to improve the chances of successfully amplifying *emoA* from pEmoA. One final consideration that was made was to ensure that the EcoRI recognition sites were not positioned too closely to the end of the expected PCR product, as the amplified product would have to be digested with EcoRI prior to insertion into the pET-32a. If the EcoRI sites were too close to the end of the fragment, the enzyme would not be able to bind and thus would not be able to catalyze the reaction (8).



Lane	Volume of 25 mM $MgCl_2$/Reaction	Annealing Temperature
2	1 µL	45.5°C
3	2 µL	45.5°C
4	3 µL	45.5°C
5	1 µL	50.9°C
6	2 µL	50.9°C
7	3 µL	50.9°C
8	1 µL	56.1°C
9	2 µL	56.1°C
10	3 µL	56.1°C
11	Negative control (no plasmid template)	

FIG. 3 Amplification of *emoA* from pEmoA, third trial. Two conditions were varied in this trial: the amount of $MgCl_2$ added per reaction and the annealing temperature. Lanes 2 through 11 contained the different PCR reactions. Lanes 1 and 12 contained 5 µL of the Mass Ruler Express DNA Ladder Mix Forward (Fermentas, Cat. No. SM1283).

This concern was addressed by ensuring that there were at least two nucleotides from the end of the expected PCR product to the EcoRI site.

None of the four PCR trials, however, yielded the correct 1.3 kb *emoA* fragment from pEmoA. A potential explanation for this is that the template DNA contained inhibitory molecules that prevented its amplification by PCR (1). It is also possible that by adding the template plasmid in its uncut form, the primer binding regions were somehow obstructed, preventing the primers from annealing properly to the plasmid to initiate amplification. Another possible reason is that the primers themselves were poorly designed, with too many mutations to allow for successful binding to the plasmid DNA, in which case they would have to be redesigned. Finally, the system may merely require some further optimization, and that by adjusting some of the experimental conditions (e.g. annealing temperature, amount of $MgCl_2$ in the reaction), it will allow for the correct fragment to be amplified. It is unlikely that the reagents themselves were defective as they were used successfully by other students to amplify different gene products.

FUTURE EXPERIMENTS

The amplification of *emoA* from pEmoA by PCR needs to be further optimized to generate the correct product. As only a few DNA molecules are needed to initiate the reaction, the template could first be diluted ten-fold to remove any inhibitory molecules (1). Linearized pEmoA could also be used as the template plasmid to determine whether the uncut form of the plasmid was hindering the reaction. To address the possible issue of poor primer design, other software programs could be used to design different primers, such as *Primer3* which is available free on Internet (1). As well, different reaction parameters could be further varied. From Trial 3, it appeared that an annealing temperature below 56°C was the most favorable for generating a product, albeit a nonspecific one, so further annealing temperatures in this range could be tested. The concentrations of the reaction components could also be further adjusted by increasing the amount of primers and $MgCl_2$ used, as it has been reported that the optimum range for primer concentration is 0.1 to 1 µM while an $MgCl_2$ concentration between 1 to 4 mM has been recommended (1, 11). Finally, high-fidelity DNA polymerases such as *Pfu* and *Vent*, which might be better suited to this particular PCR reaction, could be used to reduce the introduction of errors in amplification (3).

Upon successful amplification of *emoA* from pEmoA, the next step in this project would be to generate the three constructs described in the Introduction and transform them into the *E. coli* strain BL21 (DE3) (7). Assays would then need to be developed and performed to determine the solubility of EmoA produced by each culture to compare the amount of soluble EmoA produced from each of the three constructs.

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. William Ramey and Jennifer Sibley for their invaluable guidance and support throughout this study. I would also like to thank my classmates for their advice and suggestions. Finally, I would like to acknowledge the funding provided by the UBC Department of Microbiology and Immunology to support this study.

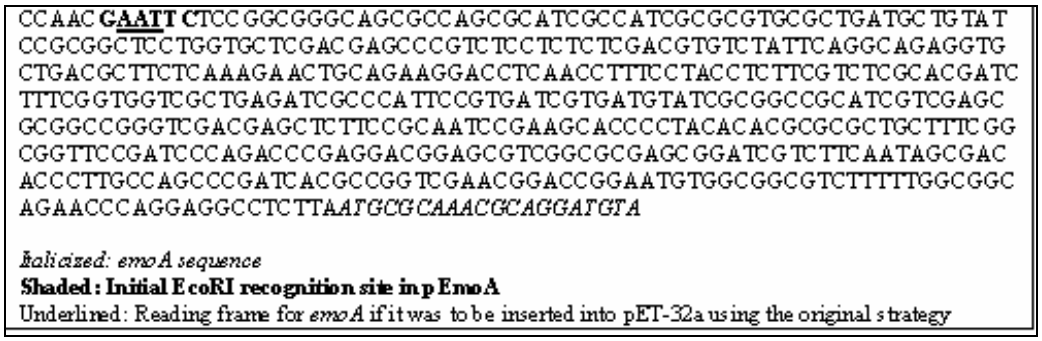


FIG. 4 The reading frame resulting from cloning *emoA* into pET-32a using the original strategy of excising the gene from pEmoA using an EcoRI restriction enzyme digest and inserting this fragment into the vector. The region between the first EcoRI recognition site and the beginning of *emoA* is shown. As observed from the underlined regions, the reading frame resulting from this strategy does not allow for the correct expression of the *emoA* start codon.

REFERENCES

1. **Altshuler, M. L.** 2006. PCR Troubleshooting: The Essential Guide. Caister Academic Press, Norfolk, U.K.
2. **Bohuslavek, J., J. W. Payne, Y. Liu, H. Bolton, and L. Xun.** 2001. Cloning, sequencing, and characterization of a gene cluster involved in EDTA degradation from the bacterium BNC1. Appl. Environ. Microbiol. **67:** 688-695.
3. **Cline, J., J. C. Braman, J. C., and H. H. Hogrefe.** 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res. **24:** 3546-3551.
4. **EMD Biosciences, Novagen.** 1998. pET-32a-c(+) vectors. [Online.] http://www.emdbiosciences.com/docs/NDIS/TB122-000.pdf.
5. **Kazem, M.** 2004. Cloning EDTA monooxygenase as a model protein to characterize the effects of flavin oxidoreductase on solubility of proteins in protein overexpression systems. J. Exp. Microbiol. Immunol. **6:** 26-34.
6. **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 cytoplasm. Bio/Technology. **11:** 187-193.
7. **Louie, T. M., H. Yang, P. Karnchanaphanurach, X. S. Xie, and L. Xun.** 2002. FAD is a preferred substrate and an inhibitor of *Escherichia coli* general NAD(P)H:flavin oxidoreductase. J. Biol. Chem. **277:** 39450-39455.
8. **Moreira, R. F. and C. J. Noren.** 1995. Minimum duplex requirements for restriction enzyme cleavage near the termini of linear DNA fragments. Biotechniques. **19:** 56-59.
9. **Sachdev, D. and J. M. Chirgwin.** 1998. Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. Protein Expr. Purif. **12:** 122-132.
10. **Singh, S. M. and A. K. Panda.** 2005. Solubilization and refolding of bacterial inclusion body proteins. J. Biosci. Bioeng. **99:** 303-310.
11. **Williams, J. F.** 1989. Optimization strategies for the polymerase chain reaction. Biotechniques. **7:** 762-768.
12. **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.
13. **Yasukawa, T., C. Kanei-Ishii, T. Maekawa, J. Fujimoto, T. Yamamoto, and S. Ishii.** 1995. Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. J. Biol. Chem. **270:** 25328-25331.