Technical Brief:

ACES™ Promoter Selection Vector

Sheldon E. Broedel, Jr., Ph. D. and Sharon M. Papciak, Ph. D.
Athena Environmental Sciences, Inc., Baltimore , MD

Athena Enzyme Systems™
1450 South Rolling Road
Baltimore, MD 21227
www.athenaes.com

ACES™ Promoter Selection Vector

December 2007

In the early days of producing proteins in *E. coli*, limitations to transcription initiation were believed to lead to lower expression levels.[1] This was often true and the historical result was the almost exclusive construction and deployment of expression vectors which carried strong promoters. These include the phage promoters T7 and T5, the synthetic promoters *tac* and *trc*, and the arabinose inducible *ara*.[2] The T7 IPTG-inducible system has become the dominant expression system currently employed. However, the use of a strong promoter, which leads to hyper-expression levels, can adversely effect recombinant protein expression.[3,4]

The most frequent and challenging problem encountered with hyper-expressed proteins is that the proteins can accumulate as inclusion bodies. These insoluble particles require that the target protein be denatured and then refolded in order to recover soluble protein.[5] The process of refolding proteins can be a daunting task. The optimum conditions for recovering active protein requires the evaluation of several variables, is not always possible and, when successful, often results in poor yields. The reasons for proteins accumulating as inclusion bodies are varied, involving both the intrinsic properties of the protein as well as interactions with host proteins. One approach used to limit the accumulation of inclusion bodies involves lowering the level of expression. This is done by decreasing the culture temperature during expression, decreasing the amount of inducer (if a titratable promoter is used), employing alternative strains (with higher levels of chaperone proteins or other accessory proteins), or using weaker promoters. None of the approaches by themselves will give the best solution and maximum production of soluble protein still requires the use of matrix analyses to identify the optimum expression conditions. The first three alternatives are relatively easy to control experimentally. The latter, however, currently requires the subcloning of the target gene sequences into a small number of alternative expression vectors which carry different inducible promoters. Thus, the ability to easily examine a range of promoter strengths to identify the best for the expression of a particular protein has been currently limited.

To facilitate the construction of promoter variants that yield the maximum level of soluble active target protein we designed a simplified method for creating a library of promoter mutants. The coding sequence for the desired protein is first inserted into the expression vector pAES25 (Fig. 1). Promoters with different strengths that direct the expression of the target protein are made from a set of specifically designed PCR primers. These are made using an inverse PCR technique that introduces random sequences within the promoter region. The promoter variant that yields maximum levels of soluble, active protein is identified by screening the library. Since the promoter selection is performed in the expression vector pAES25, no further subcloning is needed to express and purify the target protein.

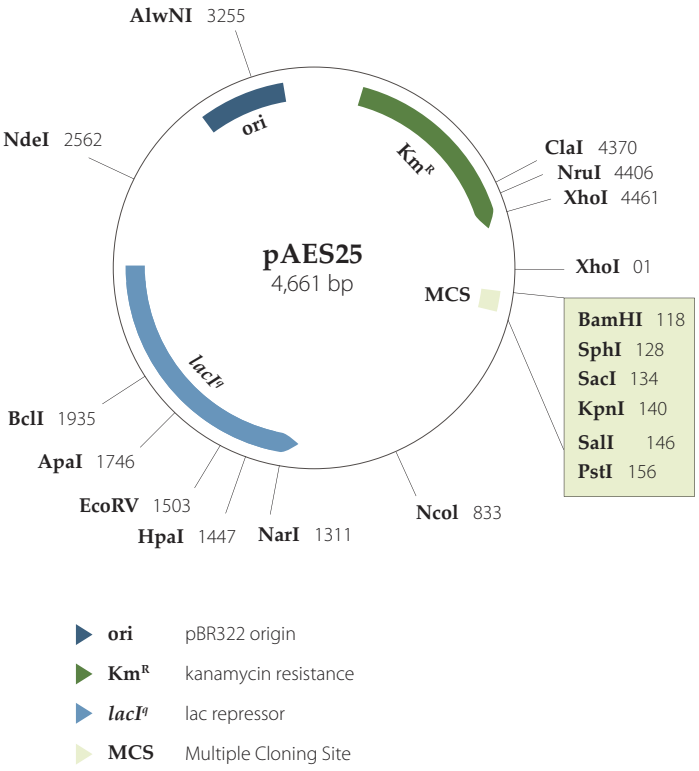

Figure 1. Plasmid map of pAES25.

To demonstrate the potential utility of this technique, the coding sequence for the *Ptilosarcus* green fluorescent protein (Prolume, Ltd.) was subcloned into pAES25 at the *Sal*I site. This resulted in a plasmid, pAES25-PtGFP, where the MCS was left intact and the GFP was expressed beginning at the AUG start codon that is located just upstream of the *Bam*HI site. Expression of the Pt-GFP protein from this plasmid results in the accumulation of the protein into inclusion bodies; however, *in vivo* fluorescence can still be detected. Purified pAES25-PtGFP was then used as a template for amplification with Primer A (N_{17}GATTCAATTGTGAGCGG)

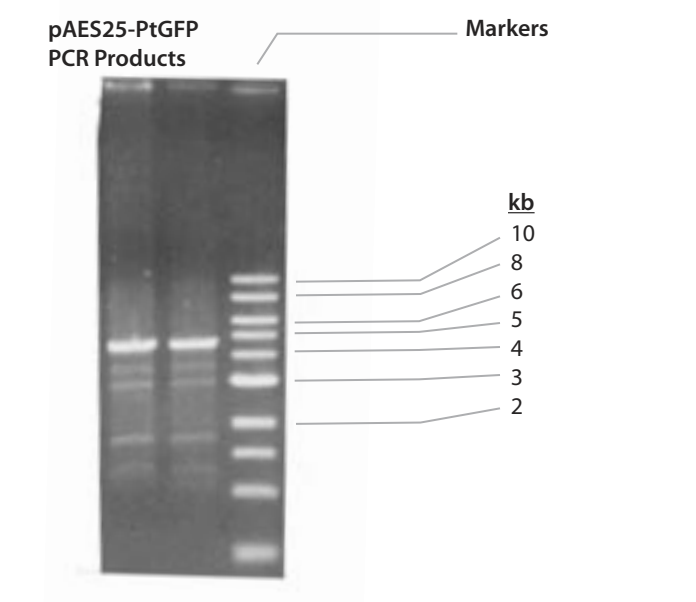

Figure 2. Analysis of the PCR product. A 5 µl sample of each amplification reaction was loaded onto a 0.7% agarose-TBE gel and electrophoresed.

and Primer B (N_{17}ATTTTTTATGATTTCTCGAG). Two 100 µl reactions containing 10 ng of plasmid DNA, 20 pmoles of each primer, 200 µM dNTPs and 0.5 Unit Hot Star DNA Polymerase (Qiagen) was heated to 95°C for 15 min and then subjected to 20 cycles of 95°C, 30s; 60°C, 30s with -0.5°C per cycle; 72°C, 5 min followed by 15 similar cycles except that the annealing temperature was 50°C and the final extention temperature at 72°C was extended by 10 minutes. before chilling to 4°C. The resulting amplification product was analyzed by agarose gel electrophoresis (Fig. 2) and then gel purified using the Qiaex II Gel Purification Kit (Qiagen) according the manufacturer's protocol. The purified DNA was recovered in 40 µl of 10 mM Tris-Cl, pH 8.5. To ligate the vector, a 2 µl aliquot was phosphorylated in a 10 µl reaction (20 Units kinase in 50 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 50 pmoles ATP), heated at 75°C for 2 min., chilled on ice for 5 min. and ligated by adding 1 Weiss unit of T4 DNA ligase. After 1 hour at 22°C, the DNA was introduced into competent cell strain HMS174 and Km^R transformants were selected.

Isolates expressing the GFP protein were identified by inducing expression of 93 independent transformants. Each of 93 colonies were used to inoculate 0.2 ml of Turbo Broth™ (AthenaES™) in a 96-well microplate. Strains harboring pAES25 and pAES25-PtGFP were used as parental controls. The cultures were incubated at 37°C overnight with shaking using a GENIOS Fluorescence Microplate Reader. A 10 µl aliquot of each culture was subcultured into 0.2 ml of fresh medium and the microplate incubated at 37°C for 3 hours. GFP expression was induced by the addition of 1 mM IPTG and the level of fluorescence measured at 30 min intervals for 8 hours. Figure 3 shows the kinetics of fluorescence accumulation after induction. Of the 93 isolates selected 10 showed significant levels of GFP accumulation with several others showing relatively slight increases in fluorescence. The induction experiment was repeated except analysis of the total

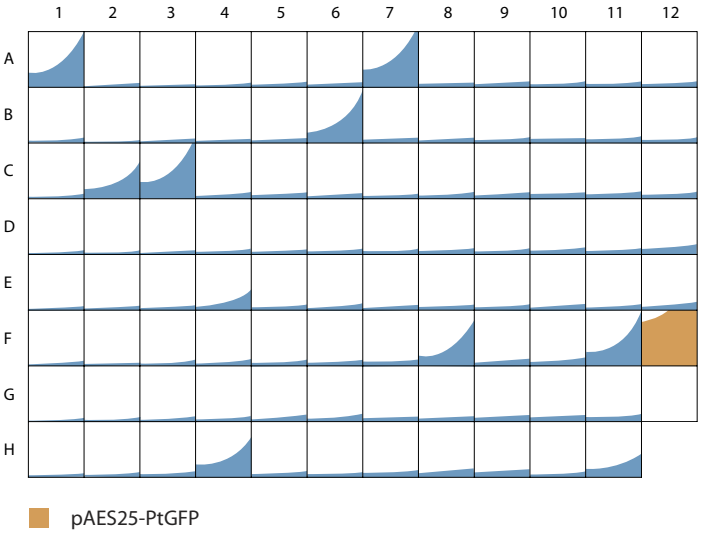

Figure 3. Accumulation of GFP in 93 promoter mutants. The increase in fluorescence was measured in cultures after the addition of IPTG to 1 mM. Fluorescence was measured at 30 min. intervals for 8 hours using a λ_{ex} of 485 nm, λ_{em} of 535 nm, gain 82, 3 flashes, and 40 µs integration. The parent strain harboring pAES25-PtGFP showed high GFP levels before induction which resulted in fluorescence above the detection limit of the plate reader.

Relative Fluorescence Units per OD_{600} at 4 hours post-induction						
	1	**2**	**3**	**4**	**5**	**6**
A	23,773	644	296	414	750	662
B	673	451	683	498	626	16,915
C	497	10,696	18,526	834	766	518
D	567	570	581	655	781	596
E	744	562	698	4,894	784	798
F	556	147	582	534	545	527
G	385	549	564	547	608	506
H	461	621	560	12,055	820	420
	7	**8**	**9**	**10**	**11**	**12**
A	21,533	509	589	586	554	579
B	632	454	576	284	775	464
C	495	419	617	513	730	492
D	543	515	476	644	659	1,035
E	528	599	583	481	419	677
F	542	14,045	469	691	16,736	37,832
G	405	479	615	750	491	0
H	448	570	619	377	6,498	-

Wells A1 through E12 are the individual isolates 1 to 93. Well F12 is pAES25-PtGFP, well G12 is pAES25 and well H12 is a medium blank. RFU values of the isogenic parent strain were subtracted.

Table 1.

fluorescence produced after 4 hours induction was determined for each culture and normalized to culture density. Table 1 and Figure 4 show that relative fluorescence units (RFU) per OD ranged from a low of 147 (F2, #14) to a high of 23,773 (A1, #1). Most isolates had RFU/OD values below 850 whereas 10 isolates had RFU levels that were more than 5-fold higher. The 83 isolates exhibiting low fluorescence are likely the result of leaky transcription inherent in plasmids rather than transcription initi-

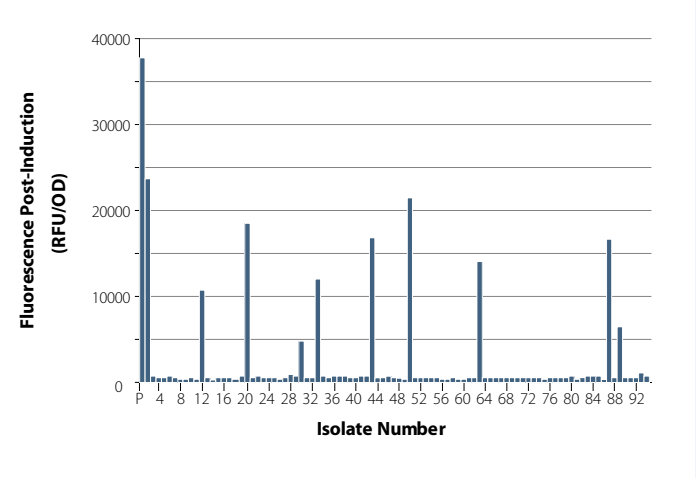

Figure 4. Fluorescence levels at 4 hours post-induction. "P" is the parent strain HMS174/pAES25-PtGFP.

ating at the modified promoter region. The 10 isolates which exhibited inducible expression from the promoter variants showed a 5-fold range in GFP levels. Their fluorescent levels were about 2-fold less than the parent plasmid. Microscopic examination revealed that none of the promoter variants yielded visible inclusion bodies whereas the strain harboring the parent plasmid did. These results suggest that by reducing the transcription rates the amount of soluble protein can be increased. Further quantitative analysis of the amount of Pt-GFP produced should confirm these initial observations.

Decreasing the expression levels of recombinant proteins can alleviate the accumulation of a protein in an insoluble state. Similarly, the adverse effects of over expressing proteins that are toxic to the host, can be mitigated by reducing the expression level of the recombinant protein. Lowering the transcription rates is one approach for decreasing expression levels. The vector described here, provides a simple and rapid method for identifying a promoter that has the transcription initiation rates which are best suited for the production of a soluble or toxic protein when the protein is otherwise difficult to obtain using the strong promoters carried on most expression vectors. This tunable expression system can be used in conjunction with other techniques to maximize the accumulation of the desired protein.

References:

1. Gralla, J., D. 1990. Promoter Recognition and mRNA Initiation by *Escherichia coli* $E\sigma^{70}$. Methods in Enzymology 185:37-54.

2. Trepe, K., 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamental to commercial systems. Appl. Microbiol Biotechnol. 72(2):211-222.

3. Jana, S. and Deb, J. K. 2005. Strategies for efficient production of heterologous proteins in *Escherichia coli*. Appl. Microbiol. Biotechnol. 67(3):289-298.

4. Sorensen, H. P. and Mortensen, K. K. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotechnol. 115(2):113-128.

5. Panda, A. K. 2003. Bioprocessing of therapeutic proteins from inclusion bodies of *Escherichia coli*. Adv. Biochem. Eng. Biotechnol. 85:43-93.