A Guide to Optimizing Protein Synthesis in the S30 T7 High-Yield Protein Expression System

ABSTRACT Optimal protein yield with the S30 T7 High-Yield Protein Expression System relies on a number of variables including: having a suitable expression vector, using a high-quality DNA template, adding the right amount of template to the reaction, and finding the best reaction time and temperature for your protein. Genetic manipulations of the protein-coding sequence, such as codon usage, truncation/domain and fusion tags can also be explored.

Kate Qin Zhao and Don Creswell, Promega Corporation

INTRODUCTION

The S30 T7 High-Yield Protein Expression System[a] (Cat.# L1110) is an *E. coli* extract-based cell-free protein synthesis system (1). This system simplifies the transcription and translation of DNA sequences cloned into vectors containing a T7 promoter by providing an extract that contains T7 RNA polymerase for transcription and all the necessary components for translation. The extract is prepared from an *E. coli* strain B that is deficient in OmpT endoproteinase and lon protease activity. This results in greater stability of expressed proteins that otherwise would be degraded by proteases. An optimized premix provides all other required components, including amino acids, rNTPs, tRNAs, an ATP-regenerating system, IPTG and appropriate salts. The system expresses high levels of recombinant proteins (up to hundreds of micrograms of recombinant protein per milliliter of reaction) within an hour at 37 °C. You supply only the cloned DNA containing a protein-coding region downstream of a T7 promoter and a ribosome binding site (RBS).

Use positive and negative controls to test optimal protein synthesis conditions.

HOW TO OPTIMIZE AN S30 REACTION

Proper Controls and Nuclease-Free Reagents

Cell-free reactions are very sensitive to nuclease contamination. Always wear gloves while setting up cell-free reactions and use RNase- and DNase-free reagents such as high-quality template DNA, nuclease-free water, reaction tubes and filtered pipette tips. Proper positive (S30 T7 Control DNA) and negative (no DNA) controls need to be performed for each set of experiments.

Sequences within and surrounding the protein-coding region influence protein expression.

DNA Coding Sequence and Vector Context

There can be large differences in gene expression as a result of the size of the protein, gene of interest and the context in which the gene resides. Changes in the position of the gene relative to the RBS will also affect expression levels (2). The RBS is generally located approximately 7 bases upstream of the AUG start codon.

Many eukaryotic genes contain sequences within the coding region that can function as ribosomal binding sites when they precede a methionine codon. The presence of such internal sequences can result in internal translation initiation and synthesis of undesired truncated proteins in the prokaryotic systems. An example of this can be seen in the expression of the firefly luciferase (*luc*) gene with a C-terminal HaloTag® fusion. We cloned *luc* into the pFC20K and pFC20A (HaloTag® 7) T7 SP6 Flexi® Vectors to generate a firefly luciferase-HaloTag® 7 (*luc*-HT) fusion vector. The construct was expressed in the S30 T7 High-Yield System, and the reactions were labeled with HaloTag® TMR Ligand (Cat.# G8252) and resolved by SDS-PAGE analysis followed by fluorescent imaging. Distinct bands visible below the full-length *luc*-HT fusion indicate internal starts that produce truncated translational products (Figure 1).

If the full-length protein cannot be expressed well, expression of domains or constructs with N- or C-terminal truncations or both (3) might lead to higher protein yield or solubility. Other cell-free

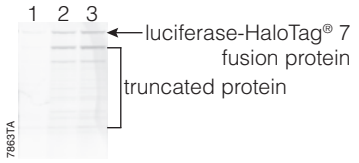

Figure 1. Truncated protein expression from internal starts of firefly luciferase (*luc*) as a C-terminal HaloTag® fusion. The *luc* gene was cloned into the pFC20A or pFC20K (HaloTag® 7) T7 SP6 Flexi® Vector. One microgram of plasmid DNA was used in each 50 µl reaction. The reactions were performed at 37 °C for 1 hour with vigorous shaking in a floor shaker. A 5 µl aliquot of each reaction was mixed with 1 µl of 5 µM HaloTag® TMR Ligand (Cat.# G8251) and incubated at room temperature for 10 minutes. After adding SDS-loading dye and heating at 95 °C for 2 minutes, the equivalent of 1 µl of the original reaction was loaded per lane, resolved by SDS-polyacrylamide gel electrophoresis (PAGE; 4–20% Tris-glycine), and visualized on a Typhoon® scanner to detect the fluorescence. Lane 1, no DNA; lane 2, pFC20A (HaloTag® 7) T7 SP6 Flexi® Vector (Cat.# G1681) expressing a luciferase-HaloTag® 7 fusion protein; lane 3, pFC20K (HaloTag® 7) T7 SP6 Flexi® Vector (Cat.# G1691) expressing a luciferase-HaloTag® 7 fusion protein.

protein expression systems such as wheat germ extract (4) or rabbit reticulocyte lysate (5) and insect extract (6) can be used in parallel with the S30 system to ensure expression of the full-length protein.

Other vector or sequence-related elements that have been reported to affect protein synthesis in S30 systems include:

- 5´ and 3´ untranslated regions (UTRs; 7,8)
- N- or C-terminal fusion tags (9,10)
- codon usage (11)
- mRNA secondary structure (12,13)
- mRNA stability (12,14)

Therefore, target protein expression might be optimized by considering the genetic elements surrounding the gene as well as those within its protein-coding sequence.

Template DNA Quality and Quantity

To achieve optimal protein yield, use only highly purified plasmid DNA templates, preferably those with concentrations of 500 ng/µl or higher. In addition, avoid adding high concentrations of salts or glycerol to the DNA template. To isolate sufficient quantity of high-quality DNA in 10 minutes, we recommend the PureYield™ Plasmid Miniprep System (Cat.# A1221; 15). For greater amounts of high-quality DNA, we recommend using the PureYield™ Plasmid Midiprep System (Cat.# A2492) with the Eluator™ Vacuum Elution Device (Cat.# A1071).

The amount of DNA required for optimal protein expression might be template-dependent. We recommend 0.5–1 µg of <5 kb plasmid DNA containing a T7 promoter per 50 µl reaction. Higher DNA concentrations (such as 2 µg/50 µl reaction) can be used for larger plasmids. If needed, the amount of DNA added can be optimized. To demonstrate the effect of template concentration, we used different amounts of S30 T7 Control DNA, which contains the coding sequence for *Renilla* luciferase (*hRluc*), in the S30 reaction. We measured protein expression by assaying *Renilla* luciferase activity. We discovered that using 0.5–2 µg DNA/50 µl reaction gave the optimal *Renilla* protein expression (Figure 2).

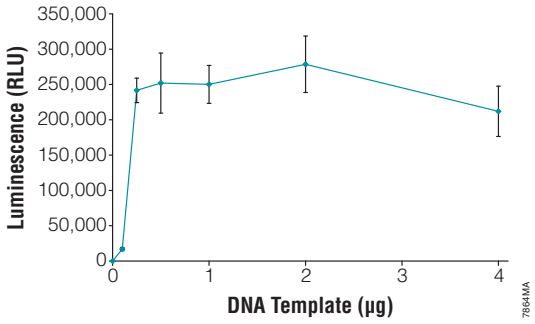

Figure 2. DNA titration of the S30 T7 Control DNA. Template DNA was purified using the PureYield™ Plasmid Midiprep System (Cat.# A2492), and 1 µg was added to triplicate 50 µl reactions. The translation reactions were incubated at 37 °C for 1 hour with vigorous shaking. *Renilla* luciferase activity was assayed using the *Renilla* Luciferase Assay System (Cat.# E2810).

In general, we recommend starting with 1 µg DNA/50 µl reaction. Optimize your DNA concentration if needed; however, an increase in the amount of DNA might increase the number of internal translational starts or produce prematurely arrested translation products.

Reaction Conditions

For the S30 T7 High-Yield Protein Expression System, optimal reaction time and temperature might be different for each protein. We analyzed the expression level of *Renilla* luciferase and Monster Green® Fluorescent Protein (hMGFP) cloned into the pFN6A (HQ) Flexi® Vector (Cat.# C8511) to investigate the effect of time and temperature. We monitored *Renilla* luciferase activity (Figure 3, Panel A) and the fluorescent signal of MGFP (Figure 3, Panel B).

High-quality DNA template benefits the coupled transcription/translation reaction.

Start with 0.5–1 µg of <5 kb plasmid containing a T7 promoter in a 50 µl reaction.

Test a range of temperatures and incubation times to maximize protein synthesis.

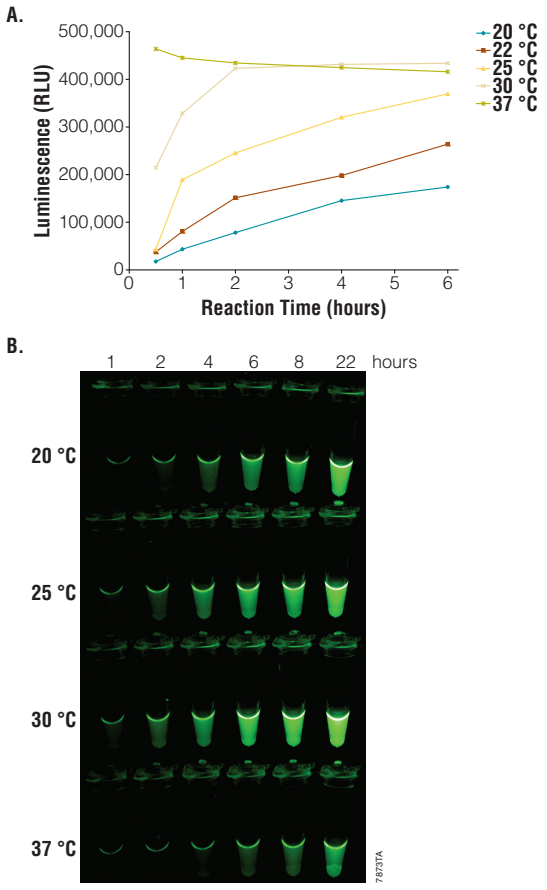

Figure 3. The effect of time and temperature on protein expression in the S30 T7 High-Yield System. Plasmids were purified using the PureYield™ Plasmid Midiprep System (Cat.# A2492). **Panel A.** Control DNA expressing *Renilla* luciferase (0.25 µg) was added to 25 µl reactions for each time point at the indicated temperature. The reactions were shaken at 1,400 rpm in an Eppendorf Thermomixer at a constant temperature. Samples were taken at various time points and frozen at –20 °C. All samples taken at each time point were analyzed simultaneously. *Renilla* expression was monitored by using the *Renilla* Luciferase Assay System (Cat.# E2810). **Panel B.** Monster Green® Fluorescent Protein (hMGFP) was synthesized in 1 ml reactions with 20 µg of pFN6A-MGFP Flexi® Vector. Parallel reactions were set up at each temperature, a 50 µl aliquot was taken at each indicated time point, and each sample was stored at –20 °C. The samples taken at each time point were analyzed together by monitoring GFP fluorescence on a UV box.

For *Renilla* luciferase, optimal functional protein production can be achieved between 25–37 °C. To reach similar amounts of *hRluc* activity, a longer reaction time was required with decreased temperatures. For hMGFP, lower temperatures (20, 25 and 30 °C) yielded more functional protein, while incubating at 37 °C resulted in less functional protein than that measured at lower temperature reactions. Therefore, for each protein of interest, reaction temperature and time should be tested for optimal functional protein expression. We recommend testing the following combinations: 37 °C for 1 hour, 30 °C for 2–6 hours, 25 °C for 4–8 hours and 20 °C for 6 hours to overnight.

CONCLUSIONS

The level of protein expression is dependent on several factors. For optimal protein yield with the S30 T7 High-Yield Protein Expression System, consider the following parameters:

1. Choose a suitable expression vector (see reference 1 for recommendations). A good starting point is using vectors suited for high-level *E. coli*-based expression under the control of a T7 promoter.
2. Always perform proper positive and negative controls.
3. Use high-quality template DNA. We recommend using plasmid DNA template purified by the PureYield™ Plasmid Systems.
4. Optimize the amount of DNA. We recommend starting with 1 µg/50 µl reaction volume.
5. Optimize reaction time and temperature.
6. Optimize the protein-coding sequence by matching codon usage to that of *E. coli* using fusion tags, testing constructs with N- and C-terminal truncation, and expressing protein domains instead of full-length protein.

REFERENCES

1. Zhao, K., Creswell, D. and Slater, M. (2008) *Promega Notes* **100**, 9–10.
2. Baranov, V.I. and Spirin, A.S. (1993) *Methods Enzymol.* **217**, 123–42.
3. Gourdon, P. *et al.* (2008) *Protein Expr. Purif.* **58**, 103–13.
4. Slater, M.R. *et al.* (2005) *Promega Notes* **91**, 21–5.
5. Hillebrecht, J.R. and Chong, S. (2008) *BMC. Biotechnol.* **8**, 58.
6. Leippe, D. *et al.* (2008) *Promega Notes* **100**, 11–2.
7. Ahn, J.H. *et al.* (2005) *Biochem. Biophys. Res. Commun.* **338**, 1346–52.
8. Son, J.M. *et al.* (2006) *Anal. Biochem.* **351**, 187–92.
9. Torizawa, T. *et al.* (2004) *J. Biomol. NMR* **30**, 311–25.
10. Palmer, E. *et al.* (2006) *Protein Sci.* **15**, 2842–6.
11. Chumpolkulwong, N. *et al.* (2006) *J. Struct. Funct. Genomics* **7**, 31–6.
12. Paulus, M., Haslbeck, M. and Watzele, M. (2004) *Nucleic Acids Res.* **32**, e78.
13. Voges, D. *et al.* (2004) *Biochem. Biophys. Res. Commun.* **318**, 601–14.
14. Ahn, J.H., Kang, T.J. and Kim, D.M. (2008) *Biotechnol. Bioeng.* **101**, 422–7.
15. Smith, D. and Vincent, E. (2008) *Promega Notes* **100**, 6–8.

ORDERING INFORMATION

Product	Size	Cat.#
S30 T7 High-Yield Protein Expression System	8 reactions	L1115
	24 reactions	L1110
pFC20A (HaloTag® 7) T7 SP6 Flexi® Vector	20 µg	G1681
pFC20K (HaloTag® 7) T7 SP6 Flexi® Vector	20 µg	G1691
HaloTag® TMR Ligand	15 µl	G8252
	30 µl	G8251
pFN6A (HQ) Flexi® Vector	20 µg	C8511
Renilla Luciferase Assay System	100 assays	E2810
	1,000 assays	E2820
Monster Green® Fluorescent Protein phMGFP Vector	20 µg	E6421
PureYield™ Plasmid Miniprep System	50 preps	A1221
	250 preps	A1222
PureYield™ Plasmid Midiprep System	25 preps	A2492
	100 preps	A2495
Eluator™ Vacuum Elution Device*	4 each	A1071

*For Laboratory Use.