Essential strategies to optimize protein expression

Many scientific research projects are based on the investigation of either transiently or stably expressed proteins. Achieving detectable and reliable amounts of recombinant protein may be challenging, especially in heterologous expression systems or while expressing highly regulated proteins. In this white paper we summarize several different technologies that can be applied to maximize the success of your protein expression experiments.

Gene optimization → **Vector optimization** → **Cell culture optimization** → **Transfection optimization**

Gene optimization

After you have chosen your favorite expression system (e.g., mammalian, bacterial, or many other systems) there are several ways to obtain the DNA starting material for your expression experiments. Gene synthesis offers the utmost flexibility in realizing individual sequence requirements like functional motifs, cloning sites, and detection tags. Since codon usage is diverse in different organisms, a good starting point is to optimize the DNA sequence for expression in your host and obtain your gene by *de novo* synthesis. By using our GeneOptimizer™ algorithm, you not only adapt the gene to the codon usage of your host system, but you also remove elements that potentially inhibit expression (e.g., killer motifs, splice sites, and RNA secondary structures). Overall, the GeneOptimizer algorithm takes more than 50 parameters into account in order to determine the optimal gene sequence for more reliable and higher-level protein expression without altering the protein sequence. Figure 1 shows an example of increased expression by optimized gene sequences in different host cells [1].

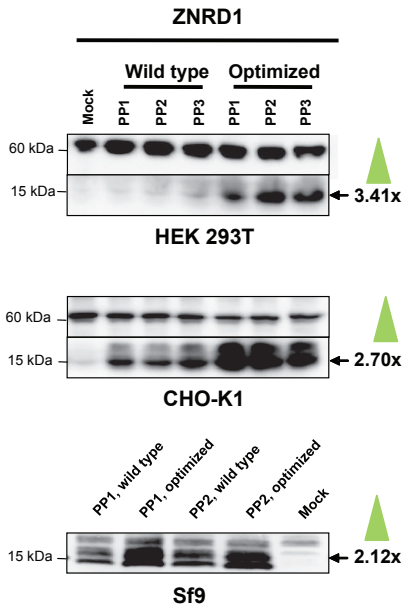

Figure 1. Gene optimization effects are not restricted to a distinct mammalian cell system. Western blot analyses of ZNRD1 protein transfected using three independent plasmid preparations (PP) into HEK 293T and CHO-K1 cells, or two independent plasmid preparations into Sf9 cells. Right: the fold increase in expression of the optimized gene [1].

Vector optimization

The coding sequence of a gene is not the only factor that needs to be considered when optimizing a construct for expression. It is also recommended to optimize the surrounding noncoding DNA elements. Tuning the expression level by choosing the optimal promoter and terminator combination could also be an essential part of an expression project, as high expression levels of foreign protein driven by a strong promoter or insufficient termination by a weak terminator can lead to growth inhibition of the host. The copy number of the vector, which is determined

by the origin of replication, also has a significant influence on the expression level of a foreign protein [2]. There is a broad range of commercially available, predesigned vectors optimized for various expression systems. You can use our Vector Selection Tool available at **thermofisher.com/vectors** to see if one fits your research purposes.

In some cases you might need a vector that is not commercially available. The GeneArt™ Elements™ Vector Construction service provides you with individually designed vectors, serving your personal experimental needs. Example applications that might require tailored vectors are gene therapy where the on/off regulation of a gene could be a major goal [3] or DNA vaccines where the presence of CpG motifs in plasmid DNA [4] plays an important role.

Cell culture and transfection optimization

The choice of expression system is of further importance for getting optimal and reliable protein expression. Options include stable cell line expression systems and transient expression systems. Factors that can be optimized in transient expression systems include cell density, the expression host, and transfection efficiency. For expression in mammalian cells, we have developed the Expi293™ Expression System that optimizes all three of these factors. In collaboration with 22 labs, expression levels of 98 different proteins were tested using the Expi293 Expression System (Figure 2). The following results were obtained:

- 87% of proteins demonstrated increased expression in the Expi293 system compared with the user's current system
- 4.6x average increase for all proteins (n = 98)
- 4.0x average increase for mAbs (n = 54); highest level was 826 mg/L
- 5.3x average increase for non-mAbs (n = 44); highest level was 790 mg/L

Conclusions

We have summarized various factors influencing protein expression. These factors, while not comprehensive, are of considerable importance for achieving reliable and high-level protein expression for your research. Please see the following resources for additional information on the technologies presented here.

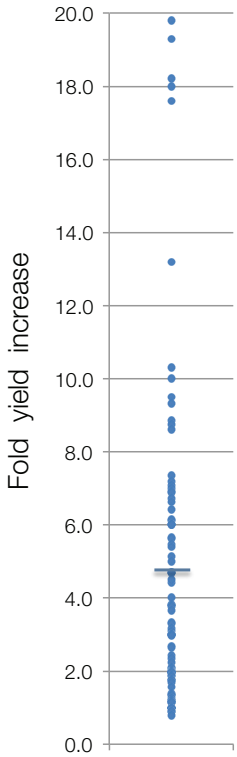

Figure 2. External collaborator results. The expression levels of 98 proteins was examined by 22 labs using the Expi293 Expression System. An average increase of 4.6-fold was observed.

Resources

thermofisher.com/genesynthesis

thermofisher.com/elementsvc

thermofisher.com/expi293

References

1. Fath S, Bauer AP, Liss M et al. (2011) Multiparameter RNA and codon optimization: a standardized tool to assess and enhance autologous mammalian gene expression. *PLoS One* 6:e17596.
2. Da Silva NA, Baily JE (1991) Influence of plasmid origin and promoter strength in fermentations of recombinant yeast. *Biotechnol Bioeng* 37:318–324.
3. Goverdhana S, Puntel M, Xiong W et al. (2005) Regulatable gene expression systems for gene therapy applications: progress and future challenges. *Mol Ther* 12:189–211.
4. Chen YS, Hsiao YS, Lin HH et al. (2006) CpG-modified plasmid DNA encoding flagellin improves immunogenicity and provides protection against *Burkholderia pseudomallei* infection in BALB/c mice. *Infect Immun* 74:1699–1705.

Find out more at **thermofisher.com/geneartproducts**

ThermoFisher SCIENTIFIC