nuclera

Technical Note

Customize DNA constructs (eGene™) with pre-set fusion tags to optimize protein expression

Namita Khanna, James Allum, Raman Pandey,
Maria Tsanakopoulou and Toby Ost.
Nuclera Ltd, Cambridge UK

Introduction

The eProtein Discovery™ platform provides rapid access to high-quality, soluble, active proteins at the benchtop allowing researchers to accelerate their protein projects.

Using Nuclera's eProtein Discovery™ system scientists can automate construct expression and purification screen to select best constructs for an off-platform scale-up, delivering purified protein in-hand (up to mg amounts) in less than 48 hours. Nuclera's technology integrates cell-free protein synthesis (CFPS) and digital microfluidics on cartridges, allowing rapid progress on protein projects through a benchtop, automated, high throughput protein access system.

Nuclera has developed eGene™ Prep Kits to rapidly convert your Gene of Interest into expression cassettes that are immediately compatible with eProtein Discovery™ reagents for Cell-Free Protein Synthesis (CFPS). CFPS utilizes linear DNAs (eGene™ constructs) as one of its ingredients to power the cell-free machinery to produce protein. eGene constructs are DNA constructs that encode for your Gene-of-Interest (GOI) with individually added varieties of solubility tag, purification tag and detection tag to aid in soluble expression and downstream identification of expressed proteins. Production of eGene constructs for CFPS is a simple and easy to follow workflow, requiring minimal hands-on time. You can consistently achieve high DNA yields of eGene constructs in just one PCR reaction. The protocol is compatible with all types of GOI with diverse sequence variations. Selected GOI should be between 125-2955 base pairs with no TEV or 3C sequences within the GOI sequence.

The eGene PCR reaction set-up can be completed in less than 30 minutes. Once purified and their concentration normalized, your eGene constructs are ready to be used in CFPS reactions on the eProtein Discovery™ Cartridge.

Creating your eGene constructs

eGene preparation begins with the design of DNA constructs. Users can access the eProtein Discovery Software; a cloud-based platform which assists users in the design of DNA constructs. Depending on the downstream target application and the potential challenges with expressing native DNA constructs, users are guided to design DNA constructs to either include the standard solubility tags, or have no solubility tag and generate 24 different DNA variants (see Figure 1). Once construct design is completed, users will be provided with a codon optimized DNA sequence for the protein target of interest. One would then place an order for the gene fragment through Nuclera. Depending on the intention to expand genes with solubility tags or not, the user will pick one from the eGene Prep Kits described in the following sections. The eGene Prep kits are designed to rapidly convert your Gene of Interest into expression cassettes that are immediately compatible with eProtein Discovery reagents for Cell-Free Protein Synthesis (CFPS).

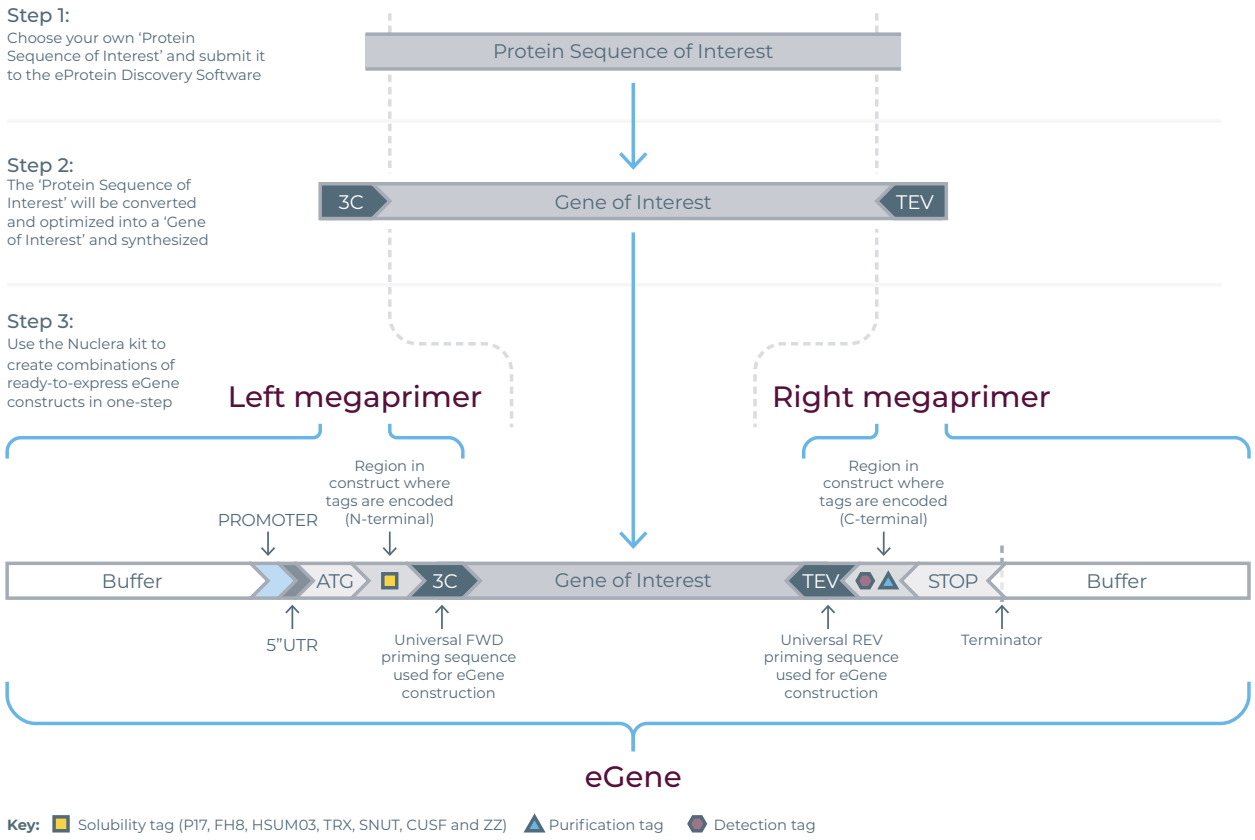


Figure 1. Overview of the eGene workflow, from inputting the Protein Sequence of Interest in the eProtein Discovery Software to generation of an eGene construct.

eGene Prep Kit Options

Nuclera offers two versions of the eGene Prep Kit based on the type of experiment and construct variation that a scientist is performing: Solubility Tag Screen and FlexiVariant Screen.

Solubility Tag Screen

This kit format provides a robust screen of constructs with different solubility tag options. This is the most popular kit format and allows you to screen 3, 4 or 6 proteins against 4, 6 or 8 solubility tag options. This kit format is best used when you are focused on obtaining certain targets as this kit allows exploration of a combination of POI variations & solubility tags to increase your chance of obtaining soluble protein.

FlexiVariant Screen

This kit format maximizes POI variation screening of either a particular protein or a screen of up to 24 different proteins. This kit format does not include solubility fusion tags. This kit format is best used when you want to screen 24 different proteins to get a quick idea of obtainability and narrow focus.

eGene Prep Kit: Solubility Tag Screen

7 solubility tags were carefully curated to be included in the eGene Prep Kit: Solubility Tag Screen. These 7 solubility tags were selected based on the unique properties of each and their proven ability to solubilize target proteins when added to the N-terminus of a protein. The unique feature of each solubility tag is described in Table 1. Also included in the standard screen is a construct that does not contain any solubility tag; with only purification and detection tags added. That way, users can identify if solubility tags confer added benefits and decide to pick a construct with or without solubility tags to work on, Figure 2. Simultaneously testing multiple versions of the same synthsized DNA will save on financial costs, as opposed to sequential testing of the same synthesized DNA.

Figure 2. eGene Prep Kit: Solubility Tag Screen - an overview of the possible solubility tag options and screening strategies.

The eProtein Discovery Cartridge can enable a 24 construct screen per cartridge and hence with the standard solubility tag screen, users can perform an expression screen with one of the following configurations:

1) 3 GOI expressed with 7 different solubility tagged constructs, and solubility tag-free construct
2) 4 GOI expressed with any 6 solubility tagged constructs, or 5 solubility tagged constructs and one untagged construct picked from Table 1, or
3) 6 GOI expressed with any 4 solubility tagged constructs, or 3 solubility tagged constructs and one untagged construct picked from the list in Table 1

Tag name	Description
P17	P17 protein from the tail of T7 phage. Molecular weight of 3.8 kDa. The P17 tag's hydrophilic sequence contributes to the solubility enhancement and also elevates the thermostability[1]
FH8	FH8 is a highly soluble and unusual thermal stable small antigen (7.5 kDa) secreted by the parasite *F. hepatica*. It not only helps solubilize protein[2] but can also act as a robust purification handle[3]
HSUMO3	Molecular weight 11.5 kDa. Human Small Ubiquitin-like Modifier, exerts a detergent-like effect on otherwise insoluble proteins[4]
TRX	Linkage to thioredoxin from *E. coli* dramatically increases the solubility of heterologous proteins, molecular weight 11.7 kDa[5]
SNUT	Solubility enhancing Ubiquitous Tag, 16.7 kDa protein tag derived from a portion of the bacterial trans-peptidase sortase (SrtA)1 found in *S. aureus*[6]
CUSF	9.9 kDa and forms a beta-barrel structure. A periplasmic protein that is part of the CusCBFA efflux complex[7]
ZZ	13.2 kDa IgG repeat domain ZZ of Protein A from *S. aureus*[8,9]
No solubility tag	For the creation of protein that is untagged at the N-terminus

Table 1. Properties of eGene construct solubility tags.

eGene Prep Kit: FlexiVariant™ Screen

The FlexiVariant Screen allows you to maximize target screening (24 different protein constructs) on the same cartridge without the use of solubility tags. This is good for situations when a solubility tag is not tolerated or when the removal of solubility tag is not favored. Each of the variants is expressed with Strep purification tags. The experimental formats supported by eGene Prep Kit: FlexiVariant Screen are:

- Screening 24 different proteins to get a quick idea of obtainability and narrow down focus
- Exploring full length and 23 combinations of truncations/isoforms/homologs/site directed mutations to compare expression yield
  All of the above are assessed in parallel to determine the most favorable DNA/protein construct that would enable expression of soluble proteins.
- Screen 24 different mutation sites

eGene Prep Kit Primer Construct and PCR Workflow

The eGene constructs are assembled based on a one pot, one step overlap extension PCR that primarily requires four key components: left megaprimer, right megaprimer, universal terminal primer pair and the gene sequence of interest corresponding to the target protein of interest (POI) as shown in Figure 3.

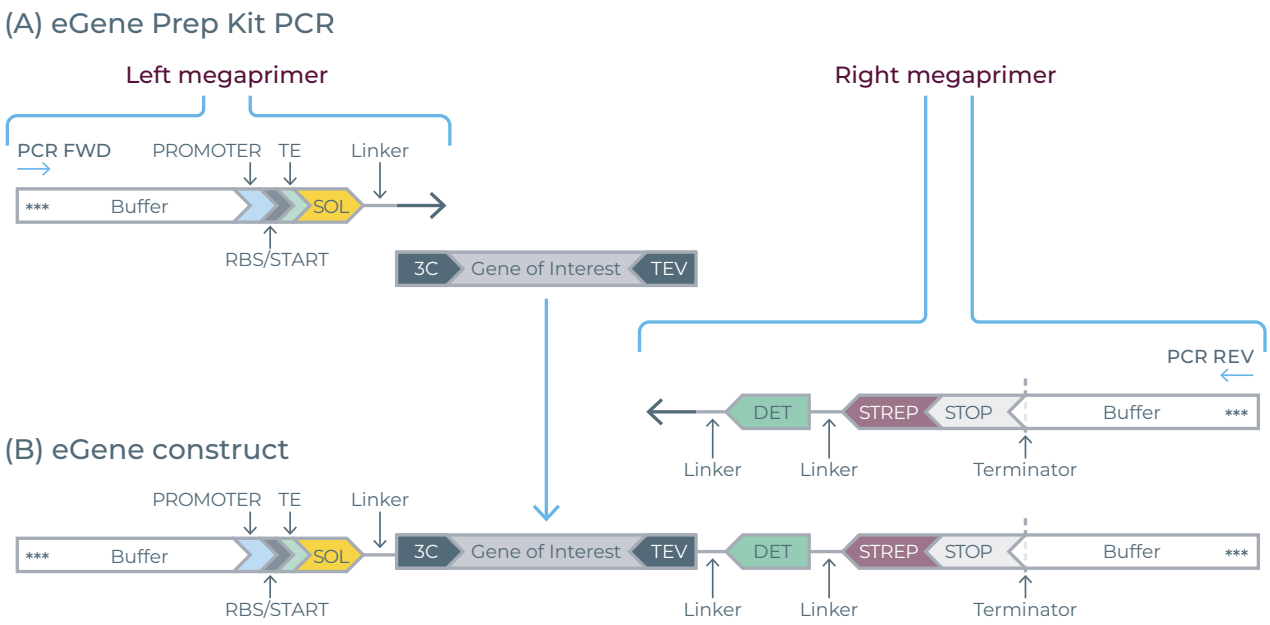


Figure 3. Description of eGene Prep Kit components including megaprimers and resulting eGene construct. Ribosome binding site (RBS), Translation enhancer (TE), Solubility tag (SOL), 3C and TEV (protease binding sites), Detection tag (DET), Streptavidin based purification tag (STREP). The terminal primers used in the assembly process contain three phosphorothioate bonds (***) at their 5' ends. The Flexivariant Screen does not encode the SOL in the Left megaprimer but continues to have the DET and STREP in the Right megaprimer.

The megaprimers are double stranded DNA molecules (599 - 1085 bp) containing all the regulatory elements required for transcription and translation. In addition, the left megaprimer may contain a variety of solubility tags, while the right megaprimer always includes a detection tag (DET, 17 amino acids long, 1.95 kDa) along with a Streptavidin based purification tag (STREP).

Sacrificial buffer regions are added before and after the transcriptional region in the left and right megaprimers, respectively, to prevent the possibility of exonuclease activity. The universal terminal primers are customized to contain three phosphorothioate bonds at the 5' termini to provide additional protection against exonuclease activity.

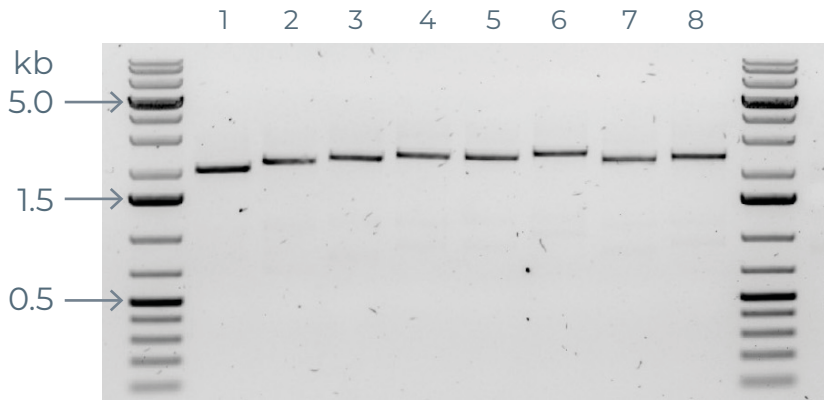
Importantly, the 3' left megaprimer which is at the N-terminus incorporates a 3C protease cleavage site, 8 amino acids long, 0.9 kDa, while the 5' right megaprimer which is at the C-terminus incorporates a TEV protease cleavage site, 7 amino acids long, 0.89 kDa, providing the user flexibility to easily cleave off the additional tags from the purified proteins. The 3C and TEV protease cleavage sites also serve as the universal hybridization site in the overlap extension assembly reaction. Therefore, the GOI must be pre-adapted with 3C and TEV adaptor sequences at the 5' and 3' ends, respectively to be compatible with the eGene Prep workflow and must not have TEV or 3C sequences within the GOI (eGene template).

A summary of the complete workflow is illustrated in Figure 5 below, which takes the user from after the point of eGene template acquisition through to PCR reaction and lastly quality control.

After receiving the eGene template, the PCR reaction is set up which contains the eGene template mixed with PCR master mix and the eGene Primer Mix reagents which comprise of the megaprimers, provided in the eGene Prep Kit. The resulting PCR eGene product can be purified using either a spin column or bead purification technique. The purified DNA product is then quantified and normalized to 5 nM for use on the eProtein Discovery platform. Users are encouraged to run a small quantity of DNA on an agarose gel to ascertain the quality of the DNA constructs generated.

Figure 4. eGene workflow summary - timings applicable for generating eGene constructs using the eGene Prep Kit: Solubility Tag Screen and eGene Prep Kit: FlexiVariant™ Screen.

The quality of DNA generated using the eGene Prep Kit is illustrated in Figure 5, showing a comparison of the base pair sizes of PCR products from incorporating different solubility tags using the eGene Prep Kit: Solubility Tag Screen, while keeping the sequence of the 'Gene of Interest' (GOI), detection tag, Strep tag constant.



Lane	Description
1	**GOI**-DET-STREP
2	P17-**GOI**-DET-STREP
3	FH8-**GOI**-DET-STREP
4	HSUMO3-**GOI**-DET-STREP

Lane	Description
5	TRX-**GOI**-DET-STREP
6	SNUT-**GOI**-DET-STREP
7	CUSF-**GOI**-DET-STREP
8	ZZ-**GOI**-DET-STREP

Figure 5. Example of an agarose gel illustrating the DNA bands generated from PCR products of the GOI CloQ (UniProt ID: Q8GHB2) expanded with different solubility tags as shown in the adjoining table. CloQ PCR reactions produced clean products, as indicated by strong amplification of a single band corresponding to the expected DNA size.

eGene product quality

In a study where a large volume of constructs were tested (n=120), the eGene Prep Kit consistently generated high quality constructs on a diverse set of GOI, up to 40 nM yield from eGene Prep PCR reactions (up to 2.5 kb size of eGene). Results from the study showed that 108/120 (93%) PCR product samples tested gave a single band of the expected size. Where multiple bands are obtained, product can be further purified by performing a band stab PCR.

Summary

The streamlined workflow for the eGene Prep Kit requires minimal hands-on time and achieves high yields with one PCR reaction.

The eGene Prep Kit is designed to streamline DNA preparation workflow to consistently deliver superior quality DNA constructs. Purified eGene constructs are ready for use in cell-free protein synthesis reactions within the eProtein Discovery platform, delivering you on the path towards your desired protein for your downstream applications.

eGene Prep Kit options

eGene Prep Kit: Solubility Tag Screen
- 8 Megaprimer pairs (7 solubility tag options and 1 strep purification tag only)
- Elution buffer
- Re-amplification primers
- Positive control DNA

eGene Prep Kit: FlexiVariant™ Screen
- 1 Megaprimer pair (strep purification tag only)
- Elution buffer
- Re-amplification primers
- Positive control DNA

Other materials required (provided by customer)

Pipette and tips

PCR
- Thermocycler
- PCR tubes/strips
- High Fidelity Polymerase PCR Master Mix
- Nuclease free water
- Template DNA

PCR clean up
- Column or bead based DNA purification kit
- Centrifuge and 1.5mL microcentrifuge tubes

DNA verification
- Agarose gel setup, electrophoresis unit and 1kb ladder, gel visualization tools or fragment analyzer
- DNA concentration determination tools

References

1. Wang, Y., Yuan, W., Guo, S. *et al.* A 33-residue peptide tag increases solubility and stability of *Escherichia coli* produced single-chain antibody fragments. Nat Commun 13, 4614 (2022).

2. Costa, S., Almeida, A., Castro, A. *et al.* The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology. Appl Microbiol Biotechnol 97(15), 6779 (2013).

3. Costa, S.J., Coelho, E., Franco, L., *et al*. The Fh8 tag: a fusion partner for simple and cost-effective protein purification in *Escherichia coli*. Protein Expr. Purif. 91, 163 (2013).

4. Hardeland, U., Steinacher, R., Jiricny, J. *et al.* Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J. 21(6), 1456 (2002).

5. LaVallie, E.R., DiBlasio, E.A., Kovacic, S. *et al.* A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Biotechnology, 11(2), 187 (1993).

6. Caswell, J., Snoddy, P., McMeel, D. *et al.* Production of recombinant proteins in *Escherichia coli* using an N-terminal tag derived from sortase. Protein Expr Purif. 70(2):143 (2010).

7. Cantu-Bustos, J. E., Vargas-Cortez, T., Morones-Ramirez, J/R. *et al.* Expression and purification of recombinant proteins in *Escherichia coli* tagged with the metal-binding protein CusF. Protein Expression and Purification 121, 61 (2016).

8. Rondahl, H., Nilsson, B., Holmgren, E. Fusions to the 5' end of a gene encoding a two-domain analogue of staphylococcal protein A. J Biotechnol. 25(3): 269 (1992).

9. Inouye, S., Sahara-Miura, Y. A fusion protein of the synthetic IgG-binding domain and aequorin: Expression and purification from *E. coli* cells and its application. Protein Expr Purif. 137:58 (2017).

nuclera®

Cambridge, UK
Boston, MA USA
Email: hello@nuclera.com

www.nuclera.com