nuclera

eProtein Discovery™ construct design and expression optimization guide

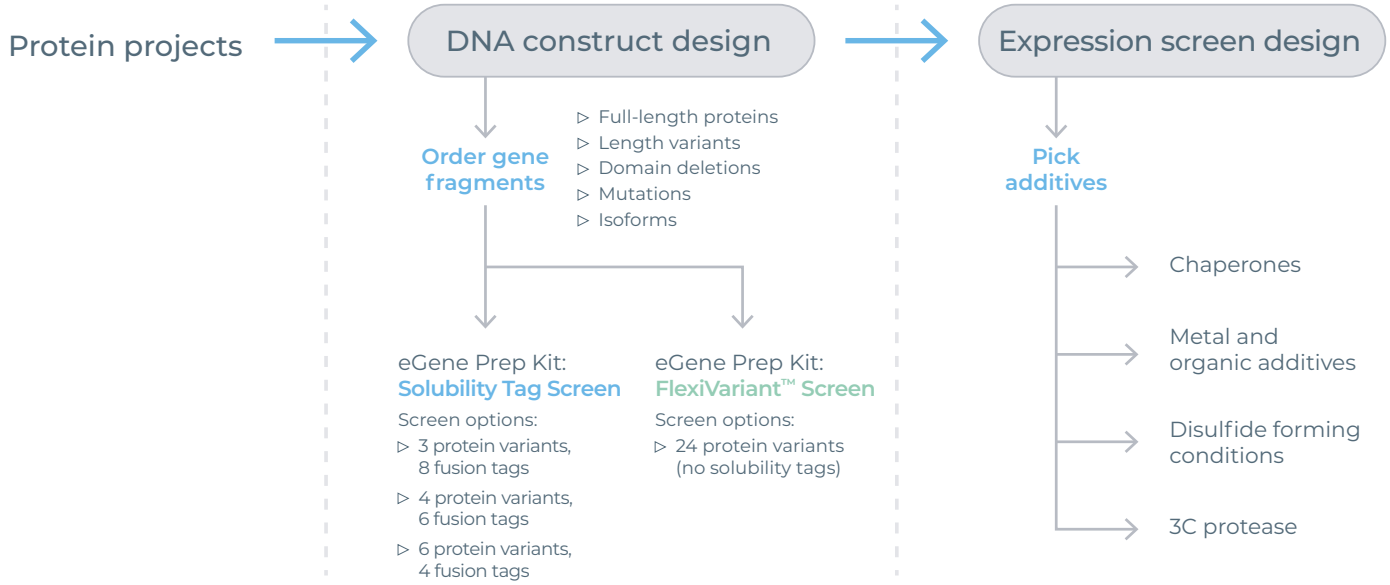
Producing proteins is often challenging, time-consuming, and prone to failure. eProtein Discovery enhances both the speed and success rate of protein production with its robust screening capabilities. Our technology allows you to simultaneously evaluate 24 DNA constructs in 8 customizable Cell-free Blends, identifying the highest purifiable yield and scaling up to have protein in hand within 48 hours.

This guide will help you determine the optimal construct and experimental design to maximize your chances of successful protein production. The document covers:

- Designing DNA constructs to create eGene™ constructs
- Optimizing expression screens by selecting the best additives for optimal protein expression

www.nuclera.com

Overview of DNA construct design and expression screen options



Which eGene kit – Solubility tag or FlexiVariant™?

You can choose from two different types of eGene™ Prep Kit formats allowing you to design the correct DNA constructs for your experiment needs.

Choose eGene Prep kit

eGene Prep Kit: Solubility Tag Screen

When focused on obtaining certain targets, explore a combination of POI variations & solubility tags to increase your chances of obtaining a soluble, active protein.

- Provides robust screen of constructs with different solubility tag options to increase chances of obtaining soluble proteins
- 3, 4 or 6 proteins
- 4, 6 or 8 fusion tag options

eGene Prep Kit: FlexiVariant™ Screen

Allows you to maximize construct screening on the same cartridge by minimizing the use of fusion tags.

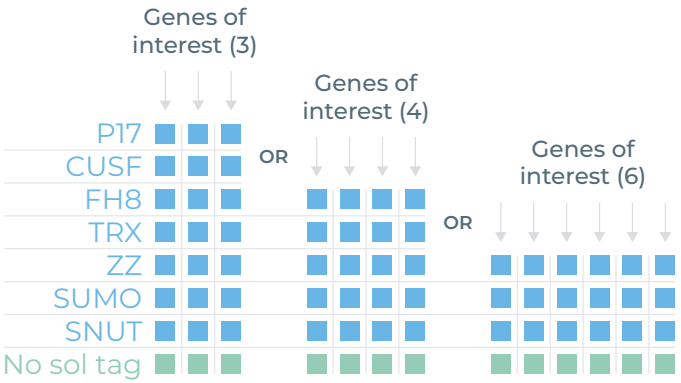
- Screen 24 different proteins to get a quick idea of obtainability and narrow focus
- Maximize POI variation screening of a particular protein
- No solubility tag

The Solubility Tag Screen Kit

The Solubility Tag Screen Kit can expand 3, 4 or 6 Genes of Interest (GOI) into 24 variants with various solubility tags added to the N-terminus and Strep purification and detection tags added to the C-terminus.

eGene™ Prep Solubility Tag Screen

Solubility Tag Screen formats:
- 3 GOI x 8 fusion tag options
- 4 GOI x 6 fusion tag options
- 6 GOI x 4 fusion tag options



Solubility tag options are: P17 tag, FH8 tag, SUMO tag, TRX tag, SNUT tag, CUSF tag, ZZ tag. The following table details the properties of each of these solubility tags.

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
CUSF	9.9 kDa copper-binding periplasmic protein; part of the CusCBFA efflux complex, and forms a beta-barrel structure when used as a solubility tag
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 but can also act as a robust purification handle
TRX	Linkage to thioredoxin from *E. coli* dramatically increases the solubility of heterologous proteins, molecular weight 11.7 kDa
ZZ	13.2 kDa IgG repeat domain ZZ of Protein A from *S. aureus*
SUMO	Molecular weight 11.5 kDa. Human Small Ubiquitin-like Modifier, exerts a detergent-like effect on otherwise insoluble proteins
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*
No solubility tag	For the creation of protein that is untagged at the N-terminus

Which screen format to pick?

Screen format: 3 protein targets x 8 fusion tags

Here are a few examples of scenarios where **3 protein variants against 8 fusion tag** expansions were chosen to give optimal protein expression:

Use case 1: Insoluble protein. An initial assessment indicated that full-length protein expression resulted in inclusion body formation.

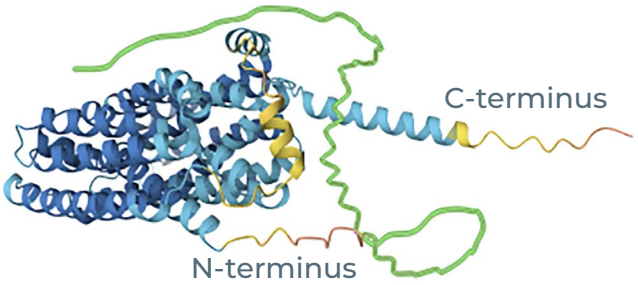
Solution: Explore all solubility tag options with 3 different protein targets:

- Screen 3 different full-length protein targets, each with 7 solubility tags (plus one construct without a solubility tag) for a total of 24 constructs on one cartridge

Use case 2: Target protein has flexible or disordered N-terminus (structure region highlighted in green) and low structure confidence C-terminus as determined by AlphaFold2 structural prediction.

Solution: You would then create the following eGene constructs:

- Protein variant #1: full-length construct
- Protein variant #2: N-terminus truncation
- Protein variant #3: C-terminus truncation
- All three protein variants would be expanded with 8 fusion tags each to determine optimal construct for protein expression

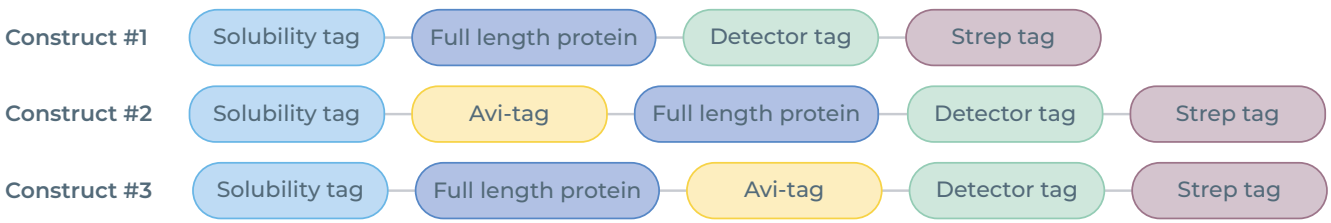


Feline leukemia virus subgroup C receptor 2 protein (FLVCR2)

Use case 3: You want to evaluate the best position of an Avi-tag for downstream biotin labeling.

Solution: You would then create the following eGene constructs:

- Protein variant #1: Full-length protein
- Protein variant #2: Protein with N-term Avi tag
- Protein variant #3: Protein with C-term Avi tag
- All three protein variants would be expanded with 8 fusion tags each to determine the optimal construct(s) for protein expression



Screen format: 4 protein targets x 6 fusion tags

Here are a few example of scenarios where **4 protein variants against 6 fusion tag** expansions were chosen to give optimal protein expression:

Use case 1: You want to maximize the eProtein Discovery Cartridge capacity – therefore you would screen 4 new target proteins (full-length) per cartridge.

Solution: You could screen 4 different target proteins against 6 selected solubility tag options or pick one without.

Use case 2: You want to explore a target protein which has flexible N- and C-terminus as determined by AlphaFold2 structural prediction.

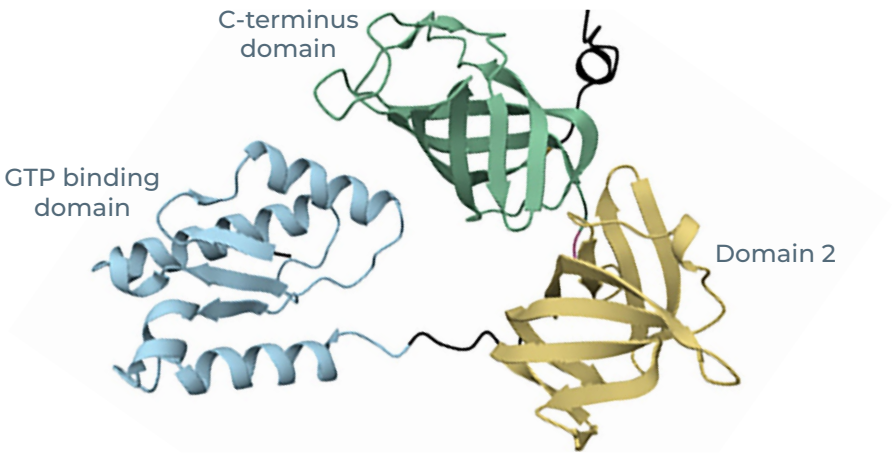
Solution: You would then create eGene constructs for:

- Protein variant #1: full-length protein
- Protein variant #2: protein with N-terminus truncated
- Protein variant #3: protein with C-terminus truncated
- Protein variant #4: protein with both N- and C-terminus truncated
- All four variants would be expanded with 6 solubility tags each

Use case 3: You have a protein that contains 3 main structured domains.

Solution: You would then create eGene constructs for:

- Protein variant #1: Full-length construct
- Protein variant #2: Domain I
- Protein variant #3: Domain II
- Protein variant #4: Domain III
- Expand all four constructs against 6 solubility tags each



An example of protein with 3 distinct domains: Elongation factor Tu protein (EF-Tu)

Screen format: 6 protein targets x 4 fusion tags

Here are a few examples of uses where **6 protein variants against 4 fusion tag** expansion were chosen to give optimal protein expression:

Use case 1: You want to maximize the eProtein Discovery cartridge Capacity - therefore you would screen 6 new target proteins (full-length) per cartridge.

Solution: You could screen 6 different target proteins against 4 selected solubility tag options or pick one without.

Use case 2: You want to explore variants of a target protein to maximize protein expression success.

Solution: You would then create the following protein variants:

- Protein variant #1: Full-length protein
- Protein variant #2: Length variant I
- Protein variant #3: Length variant II
- Protein variant #4, 5 and 6: Different isoforms
- Each protein variant would be expanded against 4 solubility tag options

Investigators would use this strategy when working on a very important target protein where both protein boundaries and closely related isoforms need to be explored.

The FlexiVariant™ Screen Kit

If you want to explore either 24 different proteins, or conversely 24 different variations of the one protein, then the eGene prep FlexiVariant Screen kit can facilitate this. Maximize gene-of-interest variation (e.g. truncations, domain deletions, mutations, homologs, isoforms) while avoiding solubility tags. Conversely, you can add your own selection of solubility tags. This kit adds detector tag and Strep purification tag to the C-terminus of your proteins.

eGene™ Prep FlexiVariant™ Screen

24 unique constructs

Example: Full-length proteins, truncations, isoforms, homologs, mutants

Use case 1: You want to assess expression levels of novel proteins

Solution: You would then create the following:

- 24 high throughput expression screening constructs of different target POIs derived from genomic sequencing data
- Evaluate expressability and solubility of novel full-length proteins

Use case 2: You want to map binding sites on your target protein

Solution: Map binding sites by mutating key residues in and around potential binding sites by creating 24 constructs with mutation sites, and then evaluate the constructs expression and purifiability.

Use case 3: You want to evaluate as many sequence variations as possible for an important target to maximize chances of obtaining soluble protein.

Solution: Explore a combination of full-length, multiple length variants, remove flexible loops, mutations and isoforms.

The best expression condition for each GOI is facilitated by the addition of 8 different Cell-free Blends using differing expression components.

Designing your expression screen

Customizable additives increase your chances of success

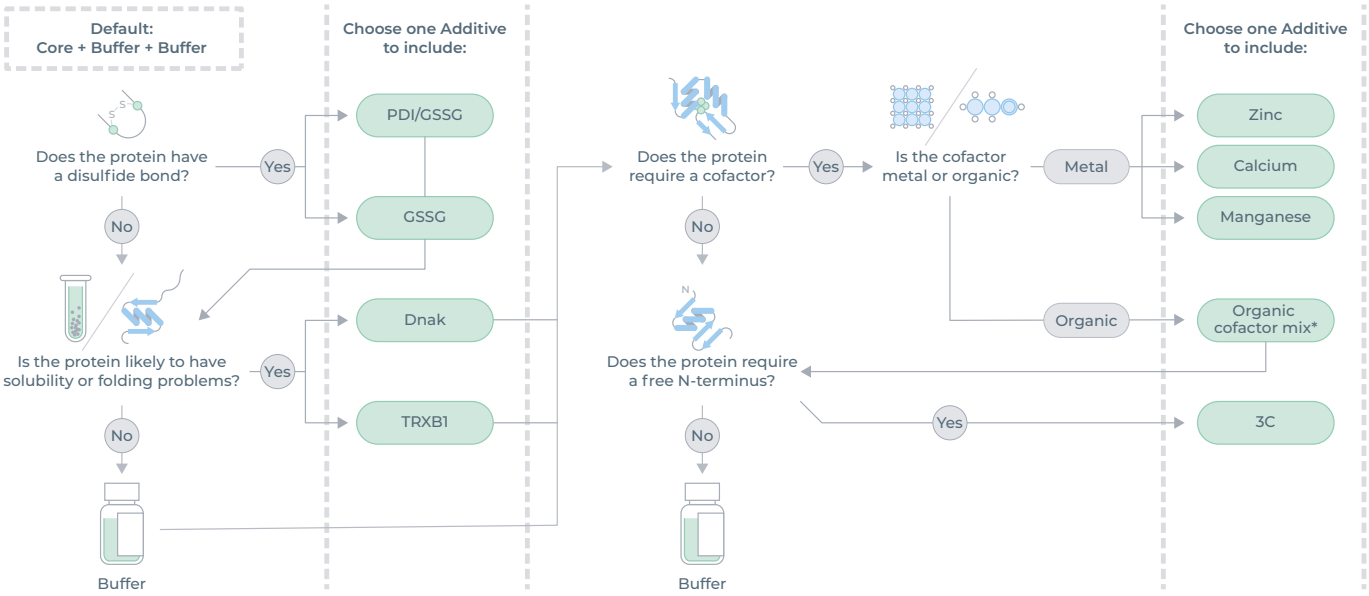
Nuclera's Cartridge Screen Reagents include the basic Cell-free Core Reagents as well as a range of customizable additives to meet the expression needs of a broad range of proteins. Each cartridge can accommodate up to 8 unique combinations of Cell-free Blends which include Cell-free Core Reagents and up to 2 additives. If needed, the expression environment can be enhanced with 2x of the same additives.

Additive options	Benefits
PDI/GSSG	Protein disulfide isomerase to promote correct disulfide bond formation
GSSG	Mimics oxidizing conditions in eukaryotic endoplasmic reticulum to promote disulfide bond formation. Also mimics the periplasmic space of prokaryotic cells
TrxB1	Chaperone to promote correct folding and stabilizes correctly folded proteins
Dnak mix	Chaperone that suppresses and reverses protein aggregation
Metal ions: $ZnCl_2$, $CaCl_2$, $MnCl_2$	Metal ions necessary for folding
Cofactor mix (NAD, Acetyl-CoA, FAD, SAM and PLP)	Promotes correct protein folding, stabilization, and/or activity
3C Protease	*In situ* solubility tag removal
Additive Buffer	Cell-free core diluting buffer; use when no additive is required

55 different additive combinations to choose from

The presence of cofactors, metal ions, chaperones or the right expression environment will have an impact on protein folding and activity.

To help you achieve successful protein formation there are a number of additives to add to the basic Cell-free Core Reagents. The below decision tree will help guide you through which additive should be included based on your protein requirements.



How to choose different combinations of additives

Specific use cases listed below are examples which highlight possible experimental scenarios that you may come across.

Do you have proteins with disulfide bonds?
Proteins that contain functional disulfide bonds will require an oxidizing environment to form properly. Here is a list of combinations of additives that creates an oxidizing environment:

"Redox Reconnaissance"

Use case
Deep dive into how redox conditions and chaperones impact the obtainability of your protein target, which you believe requires disulfide bond formation.

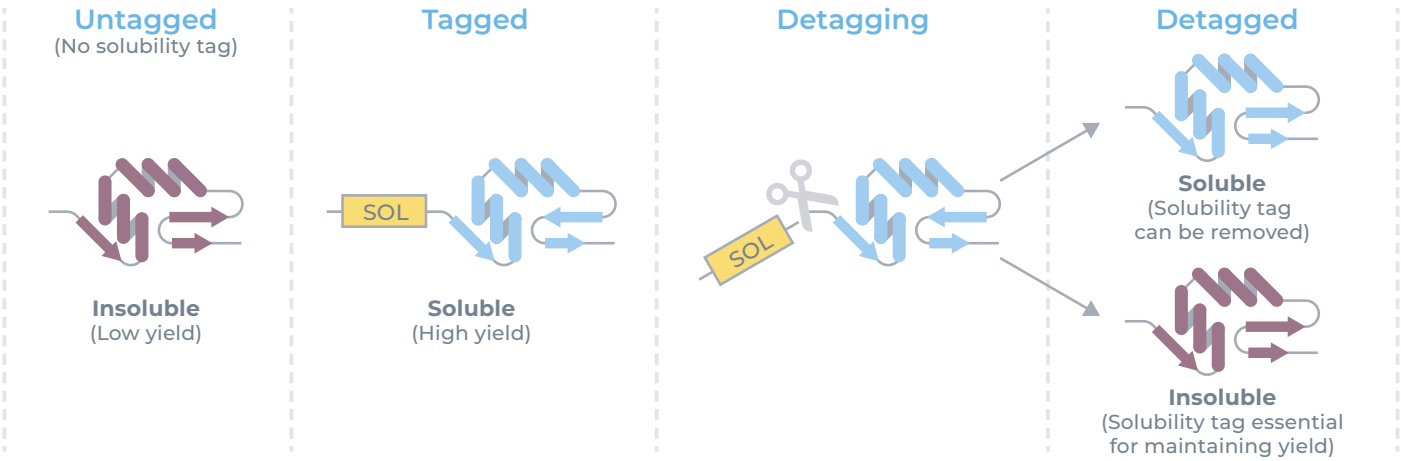
Learning outcomes
- Winning condition
- Optimal redox state (fine grain)
- Chaperone requirements
- Solubility tag requirements

	Additive 1	Additive 2	Screen features
Core	Buffer	Buffer	Reducing conditions
Core	Buffer	GSSG	Oxidizing conditions
Core	Buffer	PDI/GSSG	Oxidizing conditions with chaperone
Core	GSSG	TRXB1 or DnaK	Oxidizing conditions with chaperone
Core	PDI/GSSG	TRXB1 or DnaK	Oxidizing conditions with chaperones
Core	GSSG	GSSG	More oxidizing condition
Core	GSSG	PDI/GSSG	More oxidizing condition with chaperone
Core	PDI/GSSG	PDI/GSSG	More oxidizing condition with chaperone

You can 'detag' your solubility tag with 3C Protease

If you are working on a difficult new target and know that the protein interacts with its substrate at the far N-terminus end, and that a required biochemical assay dictates a free N-terminus, you can add 3C protease to evaluate if expressing a solubility tag helps increase protein yield. You can also check whether stability and solubility of your target protein is retained after tag removal, which we refer to as 'detagging'.

During *in situ* protease detagging, 3C protease is included from the start of cell-free protein synthesis (CFPS) reactions on the eProtein Discovery Cartridge. Solubility tags in some instances can boost expression yields. For example *in situ* detagging enables the generation of native (untagged) protein in higher soluble yields than could be obtained without solubility tags.



"Detagging"

The following gives specific examples of where you can explore detagging in conjunction with creating reducing and oxidizing conditions.

Use case
Simultaneously explore the impact of redox conditions, chaperones and the presence / absence of solubility tags on the obtainability of your protein target.

Learning outcomes
- Winning condition
- Redox state preferences
- Chaperone requirements
- Solubility tag requirements
- Stability of protein once solubility tag removed

	Additive 1	Additive 2	Screen features
Core	Buffer	Buffer	Detagging experiment with reducing expression conditions
Core	Buffer	3C Protease	
Core	TRXB1	Buffer	Detagging experiment with reducing expression conditions with chaperone
Core	TRXB1	3C Protease	
Core	GSSG	Buffer	Detagging experiment with oxidizing expression conditions
Core	GSSG	3C Protease	
Core	PDI/GSSG	Buffer	Detagging experiment with oxidizing expression conditions with chaperone
Core	PDI/GSSG	3C Protease	

Want to explore an unknown protein?
By using a step-wise combination of all the additives that are offered within the system, you can systematically elucidate the optimal conditions for protein expression, purification and ultimately obtainability.

"The Explorer 1"

Use case
For a protein you may not know much about, the Explorer 1 designs help you determine which factors have most impact on expression, purification and ultimately obtainability.

Learning outcomes
- Winning condition
- Redox state preference
- Chaperone preference
- Solubility tag requirements

	Additive 1	Additive 2	Screen features
Core	Buffer	Buffer	Reducing condition
Core	Buffer	DnaK	Reducing condition with chaperone
Core	Buffer	TRXB1	Reducing condition with chaperone
Core	Buffer	PDI/GSSG	Oxidizing condition with chaperone
Core	GSSG	Buffer	Oxidizing condition
Core	GSSG	DnaK	Oxidizing condition with chaperone
Core	GSSG	TRXB1	Oxidizing condition with chaperone
Core	GSSG	PDI/GSSG	More oxidizing condition with chaperone

"The Explorer 2"

Another set of scenarios are laid out in the following table which could be used to decipher the additives to be used for another protein where the properties are unknown.

Use case
For a protein you may not know much about, the Explorer 2 designs help you determine which factors have most impact on expression, purification and ultimately obtainability.

Learning outcomes
- Winning condition
- Redox state preference
- Chaperone preference
- Solubility tag requirements
- Cofactor requirements

	Additive 1	Additive 2	Screen features
Core	Buffer	Buffer	Reducing condition
Core	Buffer	DnaK	Reducing condition with chaperone
Core	Buffer	TRXB1	Reducing condition with chaperone
Core	Buffer	Cofactor mix	Reducing condition with cofactors
Core	PDI/GSSG	Buffer	Oxidizing condition
Core	PDI/GSSG	DnaK	Oxidizing condition with chaperones
Core	PDI/GSSG	TRXB1	Oxidizing condition with chaperone
Core	PDI/GSSG	Cofactor mix	Oxidizing condition with cofactors

Want to express a protein that's previously been difficult to obtain?
The following additive suggestions would be used to express a protein that you've found challenging in the past, or expect to be challenging. You can deploy the full chaperone suite offering to fold the unfoldable!

"The Enabler"

Use case
For difficult protein that you have struggled or expect to struggle with, deploy the full chaperone suite to fold the unfoldable.

Learning outcomes
- Winning condition
- Redox state preference
- Chaperone (or chaperones) requirement

	Additive 1	Additive 2	Screen features
Core	Buffer	Buffer	Reducing condition
Core	Buffer	DnaK	Reducing condition with chaperone
Core	Buffer	TRXB1	Reducing condition with chaperone
Core	DnaK	TRXB1	Reducing condition with chaperones
Core	PDI/GSSG	Buffer	Oxidizing condition with chaperone
Core	PDI/GSSG	DnaK	Oxidizing condition with chaperone
Core	PDI/GSSG	TRXB1	Oxidizing condition with chaperone
Core	PDI/GSSG	Cofactor mix	Oxidizing condition with chaperone & cofactors

Expressing proteins is rarely straightforward, but remember, you're not alone. If you find you are daunted by designing an experiment or feel that you require some reassuring words before starting in the lab, feel free to reach out to us. We have a dedicated technical support team always waiting to hear from you.

Please reach out to Technical Support team – techsupport@nuclera.com if you want to discuss more.

nuclera®

Nuclera Technical Support:
UK Phone / WhatsApp Business: +44 1223 942 761
US Phone: +1 508-306-1297
Email: techsupport@nuclera.com

Offices:
Nuclera UK (HQ): One Vision Park, Station Road, Cambridge, CB24 9NP, UK
Nuclera USA: 1000 Technology Park Drive, Suite B, Billerica MA 01821, USA

www.nuclera.com