Recombinant Protein Expression, Challenges and Solutions

...

Introduction

Proteins are bio-macromolecules with complex higher-order structures and a variety of functionalities. As the end products of the "central dogma" in biology, proteins are the key orchestrators of life. A deep understanding of the correlation between protein structure and function is the key to unlock the "code of life" and reveal the mechanisms of any given biological processes. While the extraction of proteins from natural sources provides key starting materials for their characterizations, the scarce in quantity and inconsistency of protein quality due to sample variations impose challenges to the field of protein biochemistry.

Hence enters the concept of recombinant protein expression, a process that introduces a foreign gene encoding the protein-of-interest (POI) into a host cell via a vector and produces the POI by hijacking the host's protein manufacturing machineries (Figure 1). Discovered more than four decades ago, plasmids (usually presented as circular double-stranded DNA molecules) were identified as the main courier of foreign genes among bacteria and thus opened the gateway to a new era of genetic engineering and manipulation to enable recombinant protein expression. During the course of four decades, various vector systems as well as host cells, including prokaryotes such as *E.coli* and eukaryotes (e.g. yeast, insect cells, and mammalian cells, Figure 1) have been developed to satisfy the generation of a broad spectrum of recombinant proteins close to their natural statues to promote the advances in biochemical research, industrial catalyzation, food processing, vaccine and therapeutics development...



Figure 1: Core concept of recombinant protein expression and common host cells. PTM= post-translational modification.

Challenges in Recombinant Protein Expression

High-quality recombinant proteins are important starting materials for successful research efforts and drug development campaigns. Key attributes of the quality of any recombinant protein include purity, oligomeric status, thermo and chemical stability, folding, post-translational modifications (PTMs), activity.... Due to the intrinsic complexity of proteins in general, protein expression and purification are often challenging. Certain sequence or structural features within a protein can impact the yield and stability of the recombinant protein product. For instance, the presence of transmembrane domains or GPI-anchor sequences often causes the association of target proteins with plasma membrane, resulting in a compromised protein yield. In the meantime, the hydrophobic nature of transmembrane domains also impacts protein stability. A detergent or lipid-based stabilization reagent is often required during the purification and formulation of such proteins. On the other hand, the over-expression and accumulation of the target protein might cause unexpected physiological changes to the host cells, resulting in the production of mis-folded proteins and protein degradation by the host proteases.

www.sinobiological.com

Tel: +1-215-583-7898 Email: order_us@sinobiologicalus.com Address: 1400 Liberty Ridge Drive, Suite 101, Wayne, PA 19087



Most proteins are delicate molecules susceptible to environmental stresses during the expression and purification processes. Thus, recombinant protein expression often requires careful planning and process optimization to identify the suitable host, culture condition and duration, as well as the optimum purification strategy to ensure the high yield of POI with minimally sacrificed quality attributes. Figure 2 summarizes the common challenges often encountered in recombinant protein expression.



Figure 2: Key factors impacting recombinant protein expression (left) and challenges in obtaining high-quality protein-of-interest (right). TM= transmembrane domain.

Last but not least, with the development of high-throughput antibody/protein screening platforms, especially these various "display" techniques, large libraries of antibody/therapeutic protein candidates are generated routinely. However, a recombinant form of the lead candidates is required to verify their functionality and efficacy. Thus, it is essential to establish a robust yet versatile high-throughput protein expression platform in parallel to facilitate lead discovery. Such a notion also brings forward new challenges to the field of recombinant protein expression.

Solutions and Perspectives

1. Approaches to Optimization

As mentioned earlier, it usually requires a systematic optimization effort to formulate a workable procedure to express high-quality target proteins. Unfortunately, there is no generalized "one-size-fits-all" approach and the expression and purification of each protein should be carefully tailored. Here, we will use a few case studies to demonstrate the optimization of key components in the recombinant protein expression workflow.

Suitable Hosts

Host cells determine the potential folding and pattern of PTM of the recombinant protein. Depending on the desired attributes, the selection of host cells should be considered carefully. In the case demonstrated in Figure 3, the attribute-of-interest for the target protein was its ability to form unified oligomers. The protein was initially expressed in insect cells due to its intracellular location but during purification, higher molecular weight polymers were observed. Optimization of the formulation buffer failed to remove the polymers. Thus, the expression host was switched to *E.coli*. The protein expressed by the 1st *E.coli* strain was prone to degradation. A second *E.coli* strain with extended endogenous protease knock-out was then used. This strain successfully produced the stable target protein with desired oligomer formation.



Figure 3: Case study-host optimization for recombinant protein expression.

www.sinobiological.com

Tel: +1-215-583-7898 **f in >** Email: order_us@sinobiologicalus.com Address: 1400 Liberty Ridge Drive, Suite 101, Wayne, PA 19087





Recombinant Protein Expression, Challenges and Solutions

• Vectors and Culture Conditions

Vectors are the vehicles that deliver the target genes into a host cell. They contain basic components such as a multiple cloning site (MCS) to harbor the gene-of-interest, a promoter to enhance expression, and antibiotic resistance genes to facilitate screening.... Vectors are usually optimized by commercial vendors to achieve the best transfection and expression efficiency. As shown in Figure 4, Sino Biological has established a vector system that out-performed those of several competitors in protein expression level. Besides vectors, other conditions used in the host cell culture, e.g. duration and temperature, should also be optimized to capture the target protein at its intact form. In the meantime, the use of additives sometimes can help increase the yield of the target protein. The addition of inorganics, such as metal ions and co-factors, has been proven to be helpful for the expression of active enzymes due to their stabilization effects on the protein molecules.



Figure 4: Case study—optimization of vector and culture conditions to enhance protein expression.

Protein Constructs

Certain structural features on a target protein might cause instability in the over-expressed form of the recombinant protein. Regions of high hydrophobicity, high disorders, and repetitive amino acid motifs... are notorious for causing protein instability and the presence of such regions is worth noting. As long as they are not directly involved in protein function, the removal of such regions may help improve protein expression (Figure 5).



Figure 5: Case study-removal of a hydrophobic region in the protein sequence to enhance protein expression.

www.sinobiological.com

Tel: +1-215-583-7898 Email: order_us@sinobiologicalus.com Address: 1400 Liberty Ridge Drive, Suite 101, Wayne, PA 19087



Recombinant Protein Expression, Challenges and Solutions

Purification Procedure

Proteins are sensitive to their surrounding chemical environment. Changes in pH, ionic strength, and oxidative status... would impact protein stability. This rule-of-thumb should be considered in seeking the most suitable buffer formula during purification and in protein storage. During purification, sometimes additives are also needed to stabilize target protein or facilitate tag exposure. As shown in Figure 6, the target protein (a single-pass transmembrane protein) was poorly extracted using detergent formula 1, presumably due to the inadequate exposure of the His-tag. Revision of the detergent formula enhanced protein extraction while during the final polishing step, another detergent (DDM) was used to replace detergent formula 2 and stabilize the final protein product.



2. Going High-throughput

As mentioned earlier, advances in high-throughput screening require an accommodating high-throughput antibody/protein expression platform to help move the process of drug discovery forward. On the other hand, the COVID19 pandemic has demonstrated the power of RNA virus hyper-mutation and suitable tools are also needed to create mutant virus protein libraries to facilitate neutralizing antibody screening and assessment. Sino Biological has established a high-throughput recombinant antibody/protein expression platform to serve therapeutics discovery and infectious disease research. This platform, depicted in Figure 7, is based on HEK293. Briefly, antibody/protein sequence library is synthesized via a PCR-based method and the target genes are then transferred to HEK293 for expression. Purified antibodies/proteins are subjected to quality and activity analysis and the suitable candidates are progressed to scale-up. This system uses flasks as initial culture method and a weekly capacity of 100~200 molecules can be generated depending on the volume of each culture. More than 15 projects have been completed so far by this platform with a maximum library size of ~600 antibodies. Virus proteins, such as influenza HA, NA, and SARS COV2 RBD mutates have also been produced by this platform to demonstrate its versatility.



Figure 7: General work-flow of the high-throughput antibody/protein expression platform.

Conclusive Remarks

Recombinant proteins are fundamental to the current biologics development landscape. Many factors would impact the quality and yield of recombinant proteins, including host cell, vector, culture method, protein construct, as well as purification approaches. Recombinant expression of a protein expression is a highly tailored process and there is no "one-size-fits-all" solution. Lastly, high-throughput recombinant antibody/protein expression has been achieved by Sino Biological using HEK293 cells. This platform is available for contracted research services for the acceleration of novel biologics discovery.

www.sinobiological.com

Tel: +1-215-583-7898 Email: order_us@sinobiologicalus.com Address: 1400 Liberty Ridge Drive, Suite 101, Wayne, PA 19087