

## Primer

# A concise guide to choosing suitable gene expression systems for recombinant protein production

Anja Schütz,<sup>1</sup> Frank Bernhard,<sup>2</sup> Nick Berrow,<sup>3</sup> Johannes F. Buyel,<sup>4</sup> Frederico Ferreira-da-Silva,<sup>5</sup> Jurgen Hastraete,<sup>6</sup> Joop van den Heuvel,<sup>7</sup> Jan-Erik Hoffmann,<sup>8,16</sup> Ario de Marco,<sup>9</sup> Yoav Peleg,<sup>10</sup> Sabine Suppmann,<sup>11,17</sup> Tamar Unger,<sup>12</sup> Martine Vanhoucke,<sup>13</sup> Susanne Witt,<sup>14</sup> and Kim Remans<sup>15,\*</sup>

<sup>1</sup>Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Technology Platform for Protein Production & Characterization, Robert-Rössle-Str. 10, 13125 Berlin, Germany

<sup>2</sup>Institute of Biophysical Chemistry, Centre of Biomolecular Magnetic Resonance, Goethe-University of Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany

<sup>3</sup>Protein Expression Core Facility, Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology, Baldiri Reixac 10, 08028 Barcelona, Spain

<sup>4</sup>University of Natural Resources and Life Sciences, Vienna (BOKU), Department of Biotechnology (DBT), Institute of Bioprocess Science and Engineering (IBSE), Muthgasse 18, 1190 Vienna, Austria

<sup>5</sup>Instituto de Biologia Molecular e Celular (IBMC) and Instituto de Investigação e Inovação em Saúde (i3S), Universidade do Porto, Porto, Portugal

<sup>6</sup>VIB, Center for Inflammation Research & Ugent, Department of Biomedical Molecular Biology, Technologiepark-Zwijnaarde 71, 9052 Ghent, Belgium

<sup>7</sup>Helmholtz Centre for Infection Research (HZI), Department of Structure and Function of Proteins, Inhoffenstrasse 7, 38124 Braunschweig, Germany

<sup>8</sup>Protein Chemistry Facility, Max-Planck-Institute of Molecular Physiology, Otto-Hahn-Str. 11, 44227 Dortmund, Germany

<sup>9</sup>Laboratory of Environmental and Life Sciences, University of Nova Gorica, Vipavska Cesta 13, 5000 Nova Gorica, Slovenia

<sup>10</sup>Structural Proteomics Unit (SPU), Department of Life Sciences Core Facilities (LSCF), Weizmann Institute of Science, Rehovot 7610001, Israel

<sup>11</sup>Protein Expression and Purification Core Facility, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

<sup>12</sup>Structural Proteomics Unit (SPU), Department of Life Sciences Core Facilities (LSCF), Weizmann Institute of Science, Rehovot 7610001, Israel

<sup>13</sup>BCCM/GeneCorner Plasmid Collection, Department of Biomedical Molecular Biology, Ghent University, Technologiepark-Zwijnaarde 71, 9052 Ghent, Belgium

<sup>14</sup>Centre for Structural Systems Biology (CSSB), University Medical Center Hamburg-Eppendorf (UKE), Notkestr. 85, 22607 Hamburg, Germany

<sup>15</sup>European Molecular Biology Laboratory (EMBL), Protein Expression and Purification Core Facility, Meyerhofstrasse 1, 69117 Heidelberg, Germany

<sup>16</sup>Present address: ATTO-TEC GmbH, Martinshardt 7, 57074 Siegen, Germany

<sup>17</sup>Present address: Immunology, Infection and Pandemic Research IIP, Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Nonnenwald 2, 82377 Penzberg, Germany

\*Correspondence: [kim.remans@embl.de](mailto:kim.remans@embl.de)  
<https://doi.org/10.1016/j.xpro.2023.102572>

## SUMMARY

This overview guides both novices and experienced researchers facing challenging targets to select the most appropriate gene expression system for producing a particular protein. By answering four key questions, readers can determine the most suitable gene expression system following a decision scheme. This guide addresses the most commonly used and accessible systems and provides brief descriptions of the main gene expression systems' key characteristics to assist decision making. Additionally, information has been included for selected less frequently used "exotic" gene expression systems.



## INTRODUCTION

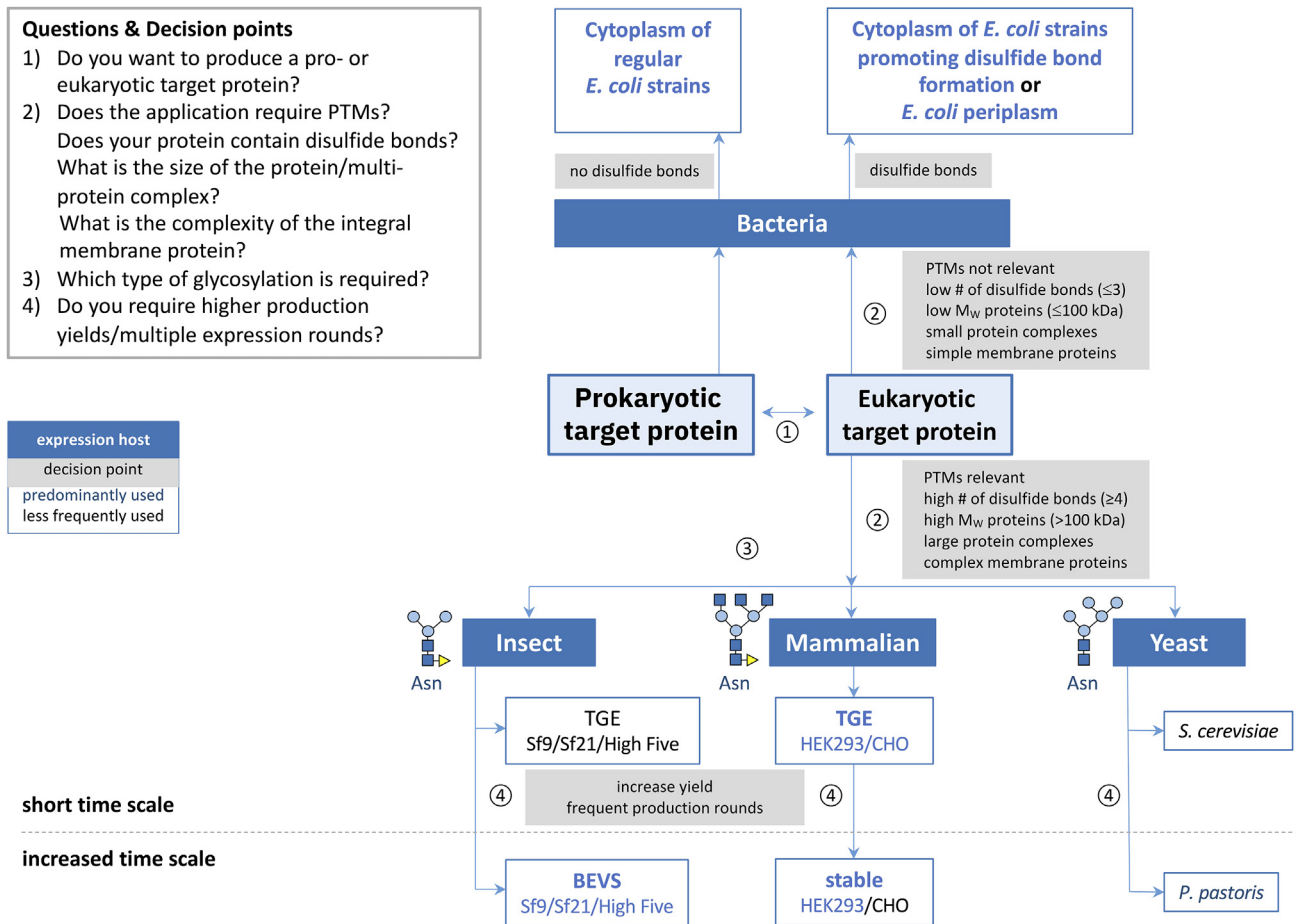
The ready availability of biological resources and related genetic sequence data combined with advances in protein production systems have enabled many laboratories to begin production of their own proteins for use as biological reagents. This allows researchers to control the costs and the availability and quality of the proteins used in their experiments.<sup>1</sup> However, many researchers that are tasked with producing recombinant proteins in their respective laboratories have little or no previous experience with the gene expression systems available. This guide evaluates the key characteristics of the most commonly used gene expression systems in order to direct researchers wishing to begin protein production to the most appropriate system for their needs and resources. The evaluation of the main features of the systems are based on a survey (see [supplemental information](#), “P4EU survey results”) conducted among the members of the Protein Production and Purification Partnership in Europe (P4EU, <https://p4eu.org>), which is a network of professionals active in various protein production laboratories and platforms. We gathered and evaluated information from (mainly European) protein production centers represented by 60 experienced scientists. Their overall experience corresponds to the production of thousands of proteins belonging to many different classes.

The information on the different gene expression systems is presented in two ways:

1. A decision scheme that uses four key questions to help determine the most optimal gene expression system for a certain target protein. These questions are based on the biological characteristics of the protein of interest and direct the reader through key decision points, from which the different branches of the scheme can be followed to decide on the most appropriate gene expression system ([Figure 1](#)).
2. At-a-glance comparison of the key characteristics of the most commonly used gene expression systems, which includes features such as the ease of use, the speed, the capacity of each system for protein production, folding, (complex) assembly and secretion, and the estimated running costs. The results of these evaluations are summarized graphically in [Figure 2](#).

The biological characteristics of the target protein and, to a lesser extent, the planned downstream applications will dictate the most appropriate gene expression system.<sup>2</sup> Therefore, it is important to collect information about the native localization of the protein of interest (intracellular, secreted, or membrane protein), the size/molecular weight, whether it is a single- or multi-domain protein, the number of disulfide bonds that are present, and post-translational modifications (e.g., glycosylation) and/or cofactors that might be required for correct folding and structural integrity. Some proteins that form part of multi-subunit complexes might not be stable on their own and hence require co-expression with their interaction partners. This type of information can be gathered by searching the scientific literature, consulting the Uniprot database (<https://www.uniprot.org>) and using bio-informatic tools such as ProtParam (<https://web.expasy.org/protparam/>) and AlphaFold structural predictions (<https://alphafold.ebi.ac.uk>).

Generally, the first choice for the production of prokaryotic target proteins is *E. coli* ([Figure 1](#)), although there are also other bacterial gene expression systems available ([Table S1](#)). For the production of eukaryotic target proteins, multiple factors play a role in the decision-making process. For simple eukaryotic target proteins that do not require post-translational modifications and that possess a limited amount of disulfide bonds, *E. coli* can be considered as an expression host as well ([Figure 1](#)). However, in many cases, eukaryotic gene expression systems such as yeast, insect cells, or mammalian cells might be more suitable. One of the main differences between these eukaryotic expression hosts lies in the type of glycosylation (N- and O-glycosylation) they can provide. Mammalian cells produce mainly complex type N-glycans, in which the glycan branches are modified with N-acetylglucosamine, galactose, fucose, and sialic acid.<sup>3,4</sup> In contrast, N-glycans from insect cells are generally not processed into terminally sialylated complex type structures and are instead modified into paucimannose or oligomannose structures.<sup>5,6</sup> Furthermore, the presence of

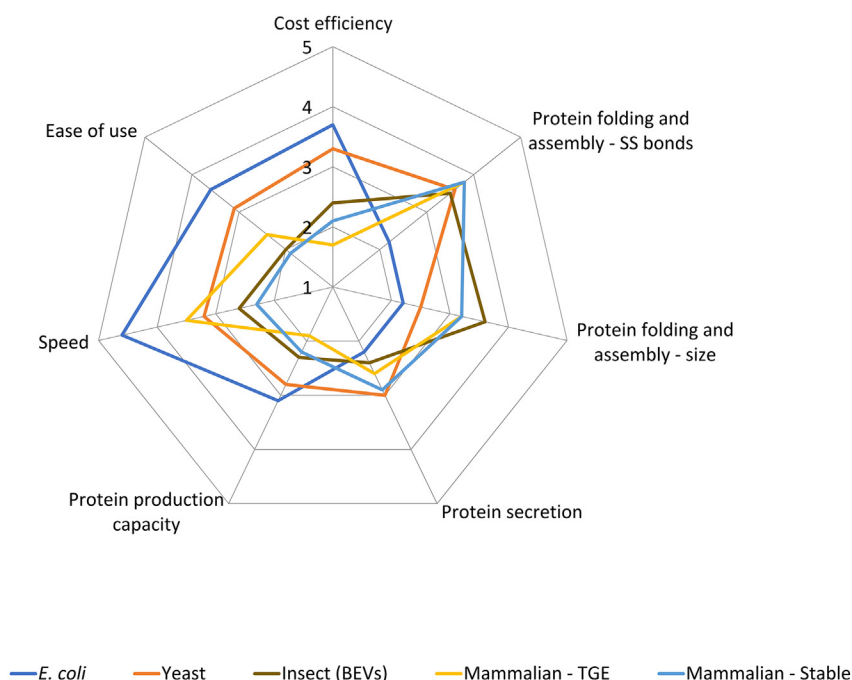


**Figure 1. Decision scheme for gene expression system selection**

There are four key decision points (circled numbers 1–4), each referring to specific questions shown in the inset text box. The blue arrows indicate the reading direction. The associated gray boxes describe the parameters to be considered at the various decision points. Expression hosts are presented in blue boxes, and predominantly and less frequently used systems are colored in blue and black, respectively. Decision points: (1) the initial decision point relates to the origin of the target protein to be produced, either being prokaryotic or eukaryotic in nature. Generally, prokaryotic proteins are produced in bacteria using different strains of *E. coli*. (2) For eukaryotic target proteins, however, multiple parameters have to be considered in the decision process. The production of such proteins in bacteria is only recommended for proteins that do not require post-translational modifications (PTMs; primarily glycosylation) for functional activity and/or stability, for proteins with up to 3 disulfide bonds, for proteins and protein complexes with a molecular weight ( $M_w$ ) of up to 100 kDa, and for small integral membrane proteins (IMPs). Generally, for disulfide-containing proteins produced in bacteria, *E. coli* strains promoting cytoplasmic disulfide bond formation are used or proteins are secreted to the periplasm. On the contrary, the production of eukaryotic target proteins in eukaryotic systems is recommended for proteins requiring functional PTMs, for proteins with multiple ( $\geq 4$ ) disulfide bonds, and large ( $>100$  kDa) proteins/complexes and larger IMPs. (3) The decision as to which eukaryotic system (insect, mammalian, yeast) to use depends on the glycan type required for obtaining functional protein (see cartoon models for the different asparagine [Asn]-linked glycans). (4) If an increased protein yield and/or frequent production rounds are needed, the additional time investment (indicated by dashed line) for the generation of stable cell lines (mammalian systems) or baculovirus expression vector system (BEVS) compared to TGE (transient gene expression) is warranted.

core  $\alpha(1,3)$ -linked fucose modifications, which are common in invertebrates but totally absent in mammals, can be immunogenic. Unicellular yeasts are capable of both N- and O-glycosylation,<sup>7</sup> but the pattern is quite different from mammalian cells. Yeast N-glycosylation is of the high/hyper-mannose type, which can cause antigenicity.<sup>8</sup> If the glycosylation type is important for the protein of interest and/or the intended downstream applications, this might be a critical factor to consider when choosing the optimal expression system.

When the target protein is a membrane-associated or integral membrane protein (IMP), the selection of a suitable gene expression system is essential. While it might be possible to produce small



#### Scoring:

**Ease of use:** (1) Possible with SOP + user training + 1 year hands-on experience; (2) possible with SOP + user training + short-term experience; (3) possible with SOP + user training; (4) possible with SOP + short-term experience; (5) possible with SOP only.

**Speed:** (1) >8 weeks; (2) 4–8 weeks; (3) 1–4 weeks; (4) 3–7 days; (5) 1–3 days.

**Protein production capacity:** (1) <1 mg/L; (2) 1–5 mg/L; (3) 5–20 mg/L; (4) 20–100 mg/L; (5) >100 mg/L.

**Protein secretion:** (1) <1 mg/L; (2) 1–5 mg/L; (3) 5–20 mg/L; (4) 20–100 mg/L; (5) >100 mg/L.

**Protein folding and assembly—size:** (1) <50 kDa; (2) 50–100 kDa; (3) 100–250 kDa; (4) 250–500 kDa; (5) >500 kDa.

**Protein folding and assembly—SS bonds:** (1) 1 SS bond; (2) 2 SS bonds; (3) 3–4 SS bonds; (4) 5–10 SS bonds; (5) >10 SS bonds.

**Cost efficiency:** (5) <50 €/L; (4) 50–100 €/L; (3) 100–500 €/L; (2) 500–1,000 €/L; (1) >1,000 €/L.

**Figure 2. Comparative overview of the characteristics associated with the major gene expression systems**

Currently, the most commonly used protein production systems are *E. coli*, yeast, mammalian cells, and insect cells. In mammalian cells, both TGE and stable cell lines are frequently used, whereas in insect cells baculovirus-mediated expression is the predominant method of choice. The main characteristics associated with protein production in these systems are ranked on a scale of 1–5, which allows for an easy comparison of the individual characteristics between the different gene expression systems. The results presented here are based on a survey, which was organized among the members of the Protein Production and Purification Partnership in Europe (P4EU) network. The scores are weighted averages calculated from the survey responses. The survey first queried about the different gene expression systems used in the participants' home laboratories. These data then formed the basis for deciding on the most commonly used gene expression systems in the community. Next, the participants were asked to score the individual characteristics associated with the gene expression systems they were familiar with based only on their own personal experiences (not on textbook knowledge). Sixty complete responses to the survey were received, which might seem like a small number of participants, but which in reality corresponds to a cumulative experience with thousands of different expression constructs. The characteristics that were assessed in the survey were: (i) ease of use, indicating how much experience/training is necessary to use a particular gene expression system; (ii) speed, which is the time required from plasmid DNA/expression construct to biomass (expressed protein) for processing; (iii) protein production capacity, which represents the average intracellular protein production capacity in mg/L of culture; (iv) protein secretion, which is the average range of secreted protein production capacity in mg/L of culture (secretion to the periplasm for *E. coli*, secretion to the extracellular milieu for yeast, mammalian, and insect cells); (v) protein folding and assembly related to the size of the protein(s) of interest, representing the ability to produce functional and correctly folded single-chain multi-domain proteins or multi-subunit protein complexes depending on their

### Figure 2. Continued

respective maximum size; (vi) protein folding and assembly related to the number of disulfide (SS) bonds, indicating the ability to produce functional and correctly folded (secreted) proteins depending on their respective number of disulfide bonds; and (vii) cost efficiency, estimating the consumable costs (e.g., media, transfection reagents, disposable flasks, plasmid preparation, cell maintenance, virus production, cell counting, etc.) for a 1-L production. All criteria are scored in a positive way, meaning higher scores correspond to more beneficial outputs.

membrane proteins in *E. coli*, eukaryotic host organisms and cell lines are generally preferred for this more challenging class of target proteins.<sup>9,10</sup> Currently, the most commonly used gene expression systems for larger IMPs—such as, for example, GPCRs, ion channels, and transporters—are insect and mammalian cells.<sup>11</sup> Even though many complex membrane proteins can be produced successfully in insect cells,<sup>12</sup> it is useful to keep in mind that the lipidic membrane environments are not identical to those in mammalian cells. As insect cells are generally cultured at 27°C, the types of lipids required to maintain membrane fluidity are different from those in mammalian cells, which are mostly cultured at 37°C.<sup>13</sup>

In order to obtain milligram quantities of recombinant proteins, *in vivo* cell-based gene expression systems are the preferred way to go. However, if either a few micrograms of protein suffice for the downstream application or *in vivo* production is impossible due to toxicity, or if specific ligands or additives are required, then *in vitro* cell-free expression (CFE) might be a suitable alternative. As the proper set-up of CFE with homemade reagents generally requires specialist training and might not be so easily accessible, CFE is neither included in the decision scheme for gene expression system selection (Figure 1) nor in the key-characteristics comparison of gene expression systems (Figure 2). However, as it might be applicable for some specific projects, detailed information and appropriate references about CFE are provided in the section “[cell-free expression](#).”

Figures 1 and 2 focus on *E. coli*, yeast, insect cells, and mammalian cells, as these are commonly used, well-characterized, and easily accessible gene expression systems. Nevertheless, there are many other alternative gene expression systems available, which possess different features and might be suitable choices for specific target proteins as well. However, as these more “exotic” host organisms are generally less frequently used, we recommend seeking experts in these systems before attempting to set up such a system in-house. For example, plants and plant cells are able to fold and secrete more complex proteins and also possess the ability to direct the recombinantly produced proteins to different cellular compartments, which can be useful for, for example, toxic proteins ([supplemental information](#), “protein production in plants”). Even though *E. coli* is by far the best-known prokaryotic gene expression system, other bacterial gene expression systems such as *Vibrio natriegens*, *Pseudomonas putida*, *Mycobacterium smegmatis*, and the Gram-positive bacteria *Lactococcus lactis* and *Bacillus subtilis* can be relevant options as well (Table S1). Furthermore, the eukaryotic expression hosts *Drosophila* S2 and the unicellular green algae *Chlamydomonas reinhardtii* represent other interesting alternative gene expression systems (Table S1).

The aim of this manuscript is to guide the reader to the most appropriate gene expression system by posing key questions regarding the characteristics of their proteins and matching them to the characteristics of the different available systems. Once an initial choice has been made regarding the most appropriate gene expression system(s), the reader can find more detailed descriptions in the specific sections of this primer. The different sections offer details about the individual systems, including key reviews and relevant references that can be consulted. Basic information, including the pros and cons of each system, is provided, as are ample references to relevant reading materials. As the availability of equipment might be an important factor as well, a more detailed overview of the instrumentation required for protein production in the respective gene expression systems can be found in Table S3. Additional information about the features of various commonly used expression strains/cell lines and vectors and how biological resources such as vectors, plasmids, and related host

strains can be acquired is provided in the [supplemental information](#), “expression vectors and strains” and “biological resources.”

## E. COLI, ONE OF THE MOST COMMONLY USED GENE EXPRESSION SYSTEMS

*E. coli* is one of the most commonly used host organisms for protein production thanks to its ease of use, cost efficiency, speed, and minimal requirement in terms of equipment. *E. coli* is generally the first organism of choice for production of prokaryotic proteins, but many eukaryotic proteins can be produced successfully in *E. coli* as well. However, compared to eukaryotic systems, *E. coli* cannot provide most of the post-translational modifications (notably glycosylation) and often fails in folding complex proteins, such as those containing multiple disulfide bonds, eukaryotic membrane proteins, or large multi-domain assemblies and multi-subunit complexes.<sup>14</sup>

In *E. coli*, proteins can be produced intracellularly in the cytoplasm, directed into the periplasm, or secreted to the extracellular milieu. The cytoplasm is a reducing environment, whereas the periplasm is an oxidizing environment that allows the formation of disulfide bonds and also has lower proteolytic activity. However, directing produced proteins into the periplasm often results in a lower yield than cytosolic production and usually not all expressed protein will be secreted into the periplasm. To direct a recombinant protein to the periplasm, one needs to add a periplasmic signal sequence (such as *phoA*, *pelB*, *ompA*, *ompT*, *dsbA*, *torA*) to the N-terminus of the protein, which will be removed after crossing the inner membrane. Proteins can be secreted either post-translationally (Sec mechanism) or co-translationally (SPR mechanism).<sup>15</sup>

A large collection of *E. coli* expression vectors is widely available, either commercially or via institutional or non-profit plasmid repositories (see [supplemental information](#), “biological resources”). Such expression vectors contain a set of genetic elements (e.g., promoter, terminator, origin of replication, antibiotic resistance cassette, etc.) that allow a regulated expression of the coding sequence of the protein(s) of interest (see [supplemental information](#), “expression vectors and strains”). One of the most frequently used bacterial gene expression systems makes use of vectors in which the gene(s) of interest are placed under control of the strong T7 promoter, which requires the T7 RNA polymerase for transcription.

Although many different *E. coli* expression strains have been developed in the past decades, the most commonly used strains are based on *E. coli* BL21. The popular *E. coli* BL21(DE3)<sup>16</sup> strain and its derivatives contain a lambda prophage encoding the T7 RNA polymerase under control of the *lacUV5* promoter, allowing IPTG-regulated expression of gene(s) under control of the T7 promoter. Various *E. coli* expression strains also have specific characteristics (see [Table S2](#)), making them more suitable for specific subtypes of proteins. For example, some strains can be engineered to produce extra copies of rare tRNAs, which is very useful if the codon usage of the gene of interest is non-optimized for expression in *E. coli*.<sup>17</sup> Other strains are better equipped to deal with the expression of toxic proteins or are more suitable for the expression of disulfide bond-rich proteins in the cytoplasm.<sup>18</sup> The required plasmid-related host strains are also accessible on a non-profit (see [supplemental information](#), “biological resources”) or profit basis. When starting with the production of a new protein in *E. coli*, it is generally recommended to assess different strains and different expression conditions (e.g., different media,<sup>19–21</sup> growth and induction temperatures, time of induction, concentration of inducer, etc.). This type of approach is also amenable to automation and hence to high-throughput screening.<sup>22</sup>

Over the years, various approaches have been developed to alleviate some *E. coli* shortcomings regarding the production of more complex proteins. For example, a commonly used method is the addition of solubility-enhancing fusion tags to the protein of interest.<sup>23</sup> Often, slowing down the rate of gene expression by using low-copy plasmids and/or low induction temperatures improves solubility as well. Alternatively, co-expression of molecular chaperones can result in proper

folding in *E. coli*.<sup>24</sup> Auto-induction media<sup>20</sup> may also improve yields of soluble protein in *E. coli*. Another option is the engineering of protein sequences to increase their solubility in *E. coli*, for which easy-to-use and validated open-access algorithms are available.<sup>25</sup> In some cases, aggregation of the recombinantly produced proteins into insoluble inclusion bodies can also be exploited to purify relatively homogeneous target proteins and refold them.<sup>26</sup> However, it must be stressed that the re-folding of proteins from inclusion bodies<sup>27</sup> requires time-consuming protocol optimization, and the yields are often low and the recovery of the native structure must be carefully verified.

Many useful general papers<sup>28–32</sup> and protocols to start approaching protein production in *E. coli* are available.

### THE USE OF YEAST AS A PROTEIN PRODUCTION SYSTEM

Yeasts are single-cell eukaryotic host organisms which combine some of the advantages of prokaryotic and eukaryotic-based gene expression systems. They are amenable to high-density fermentation and possess the necessary cellular machinery to carry out certain post-translational modifications such as glycosylation, disulfide bond formation, and proteolytic processing.<sup>33</sup> Several yeasts are being used for protein production, including *Pichia pastoris* (syn. *Komagataella phaffii*), *Saccharomyces cerevisiae*, *Yarrowia lipolytica*, and *Kluyveromyces lactis*.<sup>33,34</sup> Among these, the methylotrophic yeast *P. pastoris* has emerged in the past 20 years as one of the most popular yeast-based gene expression systems,<sup>35,36</sup> whereas *S. cerevisiae* is used as a major genetic tool.

In yeast, proteins can be produced intracellularly, or they can be secreted to the extracellular milieu, which requires the presence of an N-terminal signal peptide (e.g.,  $\alpha$ -mating factor or Ost1). *P. pastoris* is capable of both N- and O-linked glycosylation.<sup>37</sup> Glycosylation in yeast is rich in non-homogeneous hypermannosyl structures, which is different from the more complex mammalian glycan structures and can lead to antigenicity. Therefore, much effort has been put into developing *P. pastoris* strains capable of performing humanized N-glycosylation.<sup>38</sup>

*P. pastoris* is an easy-to-handle and relatively cheap gene expression system. Generating expression strains is more time consuming than for *E. coli*, but it can deliver very high recombinant protein yields and properly folded complex proteins without lipopolysaccharide contamination, which is highly beneficial for pharmaceutical and therapeutical proteins. *P. pastoris* expression vectors are generally integrated into the genome to create stable, high-expressing strains. Small-scale expression tests can be performed to screen for the highest-yielding clones. Commonly used strong promoters are the methanol-inducible AOX1 promoter or the constitutively active GAP promoter. There's also a wide selection of *P. pastoris* expression vectors available (see Table S2 and supplemental information, "biological resources"), which can be wild-type strains used in combination with antibiotic selection or auxotrophic strains that allow complementation with specific marker genes present in the expression vectors.<sup>36,39</sup>

Due to the broad applicability of *P. pastoris* both in academic research labs and in industrial protein production setups, extensive efforts have been made to further improve protein yields and to optimize growth. New elements are being added to the *P. pastoris* expression toolkit continuously, such as the OPENPichia strains,<sup>40</sup> different promoters (e.g., AOX1, UPP, PDF),<sup>41</sup> signal peptides ( $\alpha$ -mating factor, Ost1<sup>39</sup>), and optimized media with reduced protease activity and oxidation levels. Furthermore, novel high-cell-density fermentation methods are being developed.<sup>36</sup> Thanks to all these efforts in the field, there are currently already more than 70 licensed commercial products derived from *P. pastoris* available on the market ([www.pichia.com](http://www.pichia.com)).

*P. pastoris* can be used for the production of various types of (complex) proteins, but it's especially popular for the production of cytokines (IL3<sup>42</sup>), certain growth factors (GM-CSF<sup>43</sup>), and antibody derivatives without Fc fusion such as nanobodies,<sup>44</sup> bibodies, and tribodies.<sup>45</sup> Even though

*S. cerevisiae* is less popular for protein production purposes than *P. pastoris*, it is being used for the large-scale manufacturing of, for example, insulin, certain vaccines, and enzymes for industrial applications.<sup>34,46</sup>

For readers that are interested in using yeast for protein production, we recommend the following papers: Matsuzaki et al.,<sup>44</sup> De et al.,<sup>47</sup> Mastropietro et al.,<sup>48</sup> Rinnofer et al.,<sup>49</sup> and Higgins et al.<sup>50</sup> These are good starting papers to learn more about the technology in general and to obtain some initial protocols.

## BACULOVIRUS-MEDIATED GENE EXPRESSION IN INSECT CELLS

Baculovirus-mediated gene expression in insect cells is one of the most widely used systems for heterologous protein production in academia and industry and has become a major technology for the manufacturing of membrane proteins, especially GPCRs and ion channels, multi-subunit protein complexes, secreted growth factors, virus-like particles (VLPs), and gene delivery vectors for mammalian cells (reviewed in Errey et al.,<sup>11</sup> Gupta et al.,<sup>51</sup> and Mahajan et al.<sup>52</sup>). A multitude of tools developed in the past four decades—extensively engineered and improved variants of the baculovirus *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV), commercially available insect cell lines (*Spodoptera frugiperda* cell lines Sf9 and Sf21, *Trichoplusia ni* cell lines High Five and Tnao38), and the manufacturing of serum-free media—have contributed to its success.

In insect cells, proteins can be produced intracellularly, or they can be secreted to the extracellular milieu, which requires the presence of an N-terminal signal peptide. In many cases, the insect cell peptidase can recognize mammalian signal sequences,<sup>53,54</sup> but it's possible to use native insect cell signal sequences (e.g., gp67, HBM, SP1, SP2) as well.<sup>55–58</sup> Even though insect cells are capable of N- and O-glycosylation, they lack complex type N-glycans, which is a limitation for the production of therapeutic proteins. However, different types of approaches are possible to obtain proteins with a more mammalian-like N-glycosylation from *lepidopteran* insect cell lines, which are generally based on the co-expression of various glycozymes.<sup>59,60</sup>

For protein production in insect cells, the gene(s) of interest are integrated into the baculovirus genome either by Tn7-mediated transposition within *E. coli* cells (DH10Bac, Thermo Fisher Scientific; DH10MultiBac and DH10EMBacY, Geneva Biotech) or by co-transfection of insect cells with a transfer vector and baculovirus DNA (*flashBAC* and derivatives, Oxford Expression Technologies; FlexiBac<sup>61</sup>). Although more time-consuming, transposition-based integration is easier to adopt for first-time users, as it allows users to control and monitor target gene insertion by antibiotic selection, blue-white screening, and bacmid PCR/sequencing. More advanced users may instead prefer the shorter co-transfection/homologous recombination protocol within insect cells. Due to the strong baculovirus-derived *polH* and *p10* promoters driving expression of the target protein(s), high yields can be achieved in infected insect cells, no matter which of the two integration tools has been applied, as shown in a benchmarking study conducted by 14 different expert laboratories.<sup>62</sup> Most standard procedures for baculovirus generation use amplification of the first transfection-based baculovirus passage P0 to generate P1, P2, or P3. However, due to the limited stability of baculovirus particles, shortened protocols using P0<sup>63</sup> or even working virus free (titerless infected-cells preservation and scale-up [TIPS]<sup>64</sup>) has been introduced as well.

The baculovirus is unbeatable with regard to the size of the DNA cargo it can carry and transduce into insect or mammalian cells. As many as 17 subunits of a multiprotein complex have been successfully co-expressed in insect cells<sup>65</sup> and as many as 9 subunits have been transduced into HEK293 cells with BacMam baculovirus.<sup>66</sup> Different molecular cloning technologies—Golden Gate (GoldenBac),<sup>67</sup> Gibson assembly of PCR fragments (biGBac),<sup>65</sup> Cre-lox recombination (MultiBac)<sup>68</sup>—allow efficient multi-gene assembly in the baculovirus genome (see Table S2).

The main drawbacks regarding baculovirus-mediated expression in insect cells are the time required to go from DNA to target protein and the decay of the baculovirus over time. Therefore, transient plasmid-based gene expression methods have been developed as an alternative as well (see “[transient gene expression in insect cells](#)”).

For readers that are interested in using baculovirus-mediated gene expression in insect cells, we recommend papers<sup>68–72</sup> as a good start to learn more about the technology in general and to obtain some initial protocols.

## TRANSIENT GENE EXPRESSION IN INSECT CELLS

The use of baculoviral expression vectors (BEVs) to drive heterologous protein production in lepidopteran-derived insect cells is very well established. BEVs are a transient expression system lasting 3–4 days due to the disassembly of the insect cell secretion machinery, loss of cellular structures, and, finally, cell lysis. Plasmid-based transient gene expression (TGE), using chemical transfection of insect cells with expression plasmids, allows protein expression that is free of virus. The transfected cells remain viable and continue growing unhindered by a baculoviral infection process. However, plasmid-based TGE in insect cells is dependent on using strong endogenous insect cell promoters or immediate-early baculoviral promoters.

Since 2014, virus-free TGE in *Sf9/Sf21* insect cells using plasmid-based vectors has been developed<sup>73,74</sup> to avoid the time-consuming generation of baculoviruses (the generation of high-titer baculoviral stocks may require more than 3 weeks<sup>75</sup>). The initial attempts to establish insect TGE resulted in low yields of produced protein in cells of *Sf* origin until the method was hugely improved by Beckmann et al.,<sup>74,76,77</sup> Shen et al.,<sup>73</sup> Mori et al.,<sup>78</sup> and Puente-Massaguer et al.<sup>79,80</sup> Replacement of *Sf*-derived cells with *Trichoplusia ni* (High Five) cells as the expression host and introduction of the strongest available RNA polymerase II-dependent immediate-early promoter (the pOpIE2 promoter from the *Orygia pseudotsugata* multicapsid nucleopolyhedrosis virus, OpMNPV) allowed the development of a rapid and simple virus-free gene expression system in High Five insect cells. Many other experimental parameters have since been optimized and TGE in High Five insect cells was established as a robust and efficient method to produce intra-cellular and secreted protein within one week.<sup>80</sup> Briefly, transient transfection in High Five cells is performed by the addition of, first, ultra-pure expression plasmids, harboring the gene of interest cloned between the insect-specific pOpIE2 promoter and an adequate terminator and, second, polyethylenimine (PEI40) as transfection agent to logarithmically growing High Five insect cells at high density. After a short 3–4 h incubation, the cells are diluted, and growth is continued for several days. The efficiency of each transfection may be followed by co-transfection of a GFP control vector (as 5% of the total plasmid DNA transfected). The transfected cells can be harvested and adequate amounts of correctly folded protein may be isolated from either cell biomass (for cytoplasmic proteins or IMPs) or the cell culture supernatant after removal of the cells (for secreted proteins) by standard affinity chromatographic techniques.

The main advantage of the TGE insect cell system is a simple scale up to several liters in affordable insect media, while the cells are cultivated in a 27°C incubator with shaker platform without the use of CO<sub>2</sub> (in contrast with mammalian cell growth requirements). The expression timeline is fast and requires only one week once the expression plasmid is available. The insect TGE also benefits from the homogeneous paucimannose type of glycosylation, which is ideal for structural analysis of secreted proteins.<sup>81</sup> Recently, its application for producing membrane proteins has been shown as well.<sup>82</sup>

Interested readers are recommended to check Shen et al.,<sup>73</sup> Bleckmann et al.,<sup>74,77</sup> Puente-Massaguer et al.,<sup>79</sup> and Shen et al.<sup>83</sup> as excellent papers to learn about the development of the technology and how to establish TGE in High Five insect cells.

## PROTEIN PRODUCTION IN MAMMALIAN CELLS

Protein production in mammalian cells is particularly suited to larger or more complex eukaryotic proteins, as it can offer a cellular environment closely resembling the native one. Mammalian cells are a popular choice for the production of IMPs<sup>84–86</sup> and other (secreted) eukaryotic proteins requiring functional native-like post-translational modifications. Mammalian cell lines for protein production are generally derived from human embryonic kidney 293 (HEK293) or Chinese hamster ovary (CHO) cells (see Table S2). HEK293 cell lines are frequently used for research applications due to their ease of transfection, whereas CHO cells are often the system of choice for the production of bio-pharmaceutical proteins.

Mammalian cells can be grown as adherent cells or in suspension cultures. Adherent HEK293 have been used for almost 5 decades for transient transfections, as they are easy to culture and to maintain with high reproducibility, and high transfection efficiencies can be obtained with cheap reagents. The growth medium is inexpensive as well and can be prepared in house. However, for large-scale protein production, roller bottles may be necessary to avoid the need for manipulating a large number of culture plates. In contrast, HEK293-based suspension cultures, with simple passaging by dilution, present a more attractive alternative for obtaining production-level quantities of biomass. Popular suspension culture cell lines are, for example, HEK293-6E (293-EBNA1),<sup>87</sup> HEK293F, and Expi293F (see Table S2). The HEK293-6E cell line (transformed with Epstein-Barr virus nuclear antigen 1) combined with plasmids containing an oriP origin of replication allow the transfected expression plasmids to be replicated episomally, in turn leading to increased protein yields. Other suspension-adapted HEK293 derivatives include HEK293F and Expi293F, which are generally cultivated in a commercially available serum-free medium. The medium required for suspension cultures is much more expensive than for adherent cells though, and the composition is often proprietary. The high-density Expi293F commercial system combines both proprietary media and proprietary transfection reagents and may not be suitable for many academic research lab budgets.

Recombinant proteins can be produced transiently in mammalian cells by transfection with plasmid DNA or by transduction with baculoviruses (BacMam). The most widely used method for TGE is transfection with plasmid DNA, as it is fast and easy to adopt and affordable transfection reagents such as polyethylenimine (PEI) are readily available<sup>88–91</sup>. BacMam<sup>92,93</sup> is more time consuming, as it requires the generation of recombinant baculoviruses, but it can be efficient for difficult-to-transfect cell lines or when large DNA fragments need to be introduced—for example, for the expression of multi-component protein complexes (MultiBacMam).<sup>94,95</sup>

Stable mammalian cell pools can be generated by either non-targeted gene integration, using lentiviruses,<sup>96,97</sup> or transposase enzymes such as Sleeping Beauty, Frog Prince, Minos, or piggyBac.<sup>98</sup> PiggyBac transposase, isolated from cabbage looper moth *Trichoplusia ni*, and its hyperactive mutants can efficiently integrate up to 15 gene copies with a cargo capacity of 9–14 kb.<sup>99</sup> Stable pools of HEK293 and CHO cells generated with piggyBac transposase have been increasingly applied in protein production in the past 10 years for several reasons. Small amounts of plasmid DNA are needed for transfection, selection times are short (typically 11 days), the process is adaptable to many cell lines, the pools can produce high levels of protein, and the stable pools can be easily cryo-preserved. Transposase-based systems also allow the integration/expression of multiple genes, and it is possible to express cytotoxic proteins by using an inducible tetracycline promoter. Stable pools offer a lower-cost alternative to multiple rounds of TGE.<sup>100–105</sup>

For readers that are interested in using mammalian cells for protein production, Pieprzyk et al.,<sup>85</sup> Goehring et al.,<sup>86</sup> Baldi et al.,<sup>91</sup> Fornwald et al.,<sup>92</sup> Behiels and Elegheert,<sup>97</sup> and Suppmann<sup>105</sup> are recommended as good starting papers to learn more about the technology in general and to obtain some initial protocols.

### CELL-FREE EXPRESSION

CFE is defined as the production of proteins using the components required for transcription and translation in a cell-free environment. CFE systems are based on lysates of *E. coli* or eukaryotic cells such as wheat germs or insect or tobacco cells.<sup>106–108</sup> Most CFE systems work with relatively crude cell lysates, although defined systems reconstituted from purified protein and RNA components are available as well.<sup>109</sup> The cell lysates are devoid of low-molecular substances and are complemented in CFE reactions by addition of amino acids, nucleotides, energy regeneration systems, and expression templates in the form of plasmid DNA, linear DNA, or mRNA.

The protein production efficiency of CFE strongly depends on the origin of the cell lysate as well as the reaction configuration. CFE systems based on *E. coli* or wheat germ lysates can reach protein synthesis levels of mg/mL reaction in two-compartment configurations, separating reaction mixtures from feeding mixtures that provide fresh low-molecular-weight precursors. Simpler one-pot batch configurations and CFE systems based on insect or mammalian cell lysates operate in the µg/mL production levels.

The advantages of CFE systems are their open, accessible nature and operation in low volumes. A wide range of ligands, stabilizers, and other additives, even those that are toxic or difficult to implement into cell-based expression systems, are tolerated. Tailored environments for the production of individual proteins can thus be created by co-expression of targets in the presence of cofactors, interaction partners, or ligands. CFE is of particular value for the production of membrane proteins as well. Insect and tobacco cell lysates retain microsomal fragments able to translocate and glycosylate synthesized membrane proteins. However, these modifications only work efficiently at low expression levels of a few µg/mL and may become readily overloaded.<sup>110,111</sup> Alternatively, membrane mimetics in the form of liposomes, nanodiscs, or even detergents can be supplied into CFE reactions to facilitate the instant co-translational solubilization of synthesized membrane proteins.<sup>112</sup> These strategies allow high-throughput applications<sup>113</sup> and are suitable to determine the functionality and even structures of membrane proteins by crystallization, NMR, or electron microscopy.<sup>114–117</sup>

Either commercial or individual in-house CFE systems may be used. Commercial systems are usually operated in one-pot batch configurations and the costs per milligram of product can become excessive. These systems may rather be considered if synthesis of a few micrograms of protein is sufficient. Protein synthesis is completed within a few hours, and no equipment other than pipets and a thermostat is required. For more frequent use and in order to profit from the full potential of CFE, in-house systems, ideally based on easy-to-prepare *E. coli* lysates and operated in two-compartment configurations, might be preferred. Necessary infrastructure would just be an adequately equipped biochemistry lab, whereas CFE protocol development might require some training and experience. The power and perspectives of this workflow were recently reviewed.<sup>118</sup>

Obtaining high-quality samples usually results from systematic screening to identify supporting additives, suitable template designs as well as optimal concentrations of additives, and critical basic reaction components. CFE is therefore not competitive for the production of standard protein samples that can be obtained in reasonable amounts from conventional cell-based systems. However, it could become a perfect choice for difficult targets such as membrane proteins, toxins, or the production of labeled protein samples for, for example, NMR studies.<sup>119</sup>

In summary, CFE can become a system of choice if either the entire platform, including cell lysate production, is available or if intended applications would require only low amounts of sample.

### CONCLUSION

This gene expression system selection guide is based on the results of the consultation of more than 60 specialists in protein production and reflects the extensive practical experience of the authors. The decision scheme and the key characteristics comparison cover the currently most broadly

used, most widely available, and best understood gene expression systems. Unfortunately, there is no gene expression system which “fits” all, and, generally, the specific characteristics of the required protein and planned downstream application will determine which will be the most adequate gene expression system. The availability of local expertise and equipment should also be considered, as this may render a less commonly used gene expression system both accessible and viable/economical. Readers are encouraged to investigate potential gene expression systems more fully using the provided references before embarking on protein production in their own laboratories. Finally, this review is based on the authors’ experience at the time of writing. As these gene expression systems continue to evolve, it is vital that readers regularly review their options for protein production systems. Today’s “exotic” gene expression systems may become tomorrow’s widely used gene expression systems for even more challenging protein targets.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xpro.2023.102572>.

## ACKNOWLEDGMENTS

The authors thank all members from the P4EU community who responded to the survey and contributed their expertise in the preparation of this manuscript. We acknowledge Elena Gubanov (MDC, Berlin, Germany) for administrative support in conducting the survey. We are very grateful to Dr. Christian Loew (EMBL Hamburg, Germany) and Dr. Eric Geertsma (MPI-CBG, Germany) for the discussions about membrane protein production. We thank Dr. Anna Arís and Dr. Elena Garcia-Fruitós (IRTA, Spain) for their insights about the use of *Lactococcus lactis* as a recombinant protein production system.

## AUTHOR CONTRIBUTIONS

Conceptualization, A.S., N.B., F.F.d.S., J.v.d.H., Y.P., K.R.; survey, A.S., J.v.d.H.; writing & original draft, A.S., N.B., F.F.d.S., J.v.d.H., T.U.; figures, A.S., N.B., J.F.B., F.F.d.S., J.v.d.H., Y.P., S.S., T.U., S.W., K.R.; “*E. coli*, one of the most commonly used gene expression systems,” A.S., F.F.d.S., J.E.H., A.d.M., S.W., K.R.; “the use of yeast as a protein production system,” J.H., Y.P., K.R.; “baculovirus-mediated gene expression in insect cells,” J.v.d.H., S.S., K.R.; “transient gene expression in insect cells,” J.v.d.H., K.R.; “protein production in mammalian cells,” S.S., T.U., K.R.; “cell-free expression,” F.B., S.W., K.R.; supplemental information, “plants as an alternative protein production system,” J.F.B.; supplemental information, “‘exotic’ gene expression systems” and Table S1, A.d.M., Y.P., S.W., K.R.; supplemental information, “expression vectors and strains/cell lines: how to choose them?” and Table S2, A.d.M.; supplemental information, “access to biological resources,” M.V.; supplemental information, “equipment list” and Table S3, J.v.d.H.; review & editing, A.S., N.B., F.F.d.S., J.v.d.H., K.R.

## DECLARATION OF INTERESTS

Kim Remans is a member of the STAR Protocols advisory board.

## REFERENCES

- De Marco, A., Berrow, N., Lebendiker, M., Garcia-Alai, M., Knauer, S.H., Lopez-Mendez, B., Matagne, A., Parret, A., Remans, K., Uebel, S., and Raynal, B. (2021). Quality control of protein reagents for the improvement of research data reproducibility. *Nat. Commun.* 12, 2795. <https://doi.org/10.1038/s41467-021-23167-z>.
- Remans, K., Lebendiker, M., Abreu, C., Maffei, M., Sellathurai, S., May, M.M., Vaněk, O., and De Marco, A. (2022). Protein purification strategies must consider downstream applications and individual biological characteristics. *Microb. Cell Factories* 21, 52. <https://doi.org/10.1186/s12934-022-01778-5>.
- Schjoldager, K.T., Narimatsu, Y., Joshi, H.J., and Clausen, H. (2020). Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* 21, 729–749. <https://doi.org/10.1038/s41580-020-00294-x>.
- Heffner, K.M., Wang, Q., Hizal, D.B., Can, Ö., and Betenbaugh, M.J. (2018). Glycoengineering of Mammalian Expression Systems on a Cellular Level. In *Advances in Glycobiotechnology Advances in Biochemical Engineering/Biotechnology*, E. Rapp and U. Reichl, eds. (Springer International Publishing), pp. 37–69. [https://doi.org/10.1007/10\\_2017\\_57](https://doi.org/10.1007/10_2017_57).
- Tomiya, N., Narang, S., Lee, Y.C., and Betenbaugh, M.J. (2004). Comparing N-glycan processing in mammalian cell lines to native and engineered lepidopteran insect cell lines. *Glycoconj. J.* 21, 343–360. <https://doi.org/10.1023/B:GLYC.0000046275.28315.87>.
- Palomares, L.A., Srivastava, I.K., Ramírez, O.T., and Cox, M.M.J. (2018).

- Glycobiotechnology of the Insect Cell-Baculovirus Expression System Technology. In *Advances in Glycobiotechnology Advances in Biochemical Engineering/Biotechnology*, E. Rapp and U. Reichl, eds. (Springer International Publishing), pp. 71–92. [https://doi.org/10.1007/10\\_2018\\_61](https://doi.org/10.1007/10_2018_61).
7. Delic, M., Valli, M., Graf, A.B., Pfeffer, M., Mattanovich, D., and Gasser, B. (2013). The secretory pathway: exploring yeast diversity. *FEMS Microbiol. Rev.* 37, 872–914. <https://doi.org/10.1111/1574-6976.12020>.
  8. De Wachter, C., Van Landuyt, L., and Callewaert, N. (2018). Engineering of Yeast Glycoprotein Expression. In *Advances in Glycobiotechnology Advances in Biochemical Engineering/Biotechnology*, E. Rapp and U. Reichl, eds. (Springer International Publishing), pp. 93–135. [https://doi.org/10.1007/10\\_2018\\_69](https://doi.org/10.1007/10_2018_69).
  9. He, Y., Wang, K., and Yan, N. (2014). The recombinant expression systems for structure determination of eukaryotic membrane proteins. *Protein Cell* 5, 658–672. <https://doi.org/10.1007/s13238-014-0086-4>.
  10. I. Mus-Veteau, ed. (2022). *Heterologous Expression of Membrane Proteins: Methods and Protocols* (Springer US). <https://doi.org/10.1007/978-1-0716-2368-8>.
  11. Errey, J.C., and Fiez-Vandal, C. (2020). Production of membrane proteins in industry: The example of GPCRs. *Protein Expr. Purif.* 169, 105569. <https://doi.org/10.1016/j.pep.2020.105569>.
  12. Hu, N.-J., Rada, H., Rahman, N., Nettleship, J.E., Bird, L., Iwata, S., Drew, D., Cameron, A.D., and Owens, R.J. (2015). GFP-Based Expression Screening of Membrane Proteins in Insect Cells Using the Baculovirus System. In *Structural Proteomics Methods in Molecular Biology*, R.J. Owens, ed. (Springer New York), pp. 197–209. [https://doi.org/10.1007/978-1-4939-2230-7\\_11](https://doi.org/10.1007/978-1-4939-2230-7_11).
  13. Marheineke, K., Grünewald, S., Christie, W., and Reiländer, H. (1998). Lipid composition of *Spodoptera frugiperda* (Sf9) and *Trichoplusia ni* (Tn) insect cells used for baculovirus infection. *FEBS Lett.* 441, 49–52. [https://doi.org/10.1016/S0014-5793\(98\)01523-3](https://doi.org/10.1016/S0014-5793(98)01523-3).
  14. Structural Genomics Consortium; Architecture et Fonction des Macromolécules Biologiques; Berkeley Structural Genomics Center; China Structural Genomics Consortium; Integrated Center for Structure and Function Innovation; Israel Structural Proteomics Center; Joint Center for Structural Genomics; Midwest Center for Structural Genomics; New York Structural Genomics Research Center for Structural Genomics, and Northeast Structural Genomics Consortium, et al. (2008). Protein production and purification. *Nat. Methods* 5, 135–146. <https://doi.org/10.1038/nmeth.f.202>.
  15. De Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Factories* 8, 26. <https://doi.org/10.1186/1475-2859-8-26>.
  16. Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130. [https://doi.org/10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2).
  17. Eschenfeldt, W.H., Makowska-Grzyska, M., Stols, L., Donnelly, M.I., Jedrzejczak, R., and Joachimiak, A. (2013). New LIC vectors for production of proteins from genes containing rare codons. *J. Struct. Funct. Genom.* 14, 135–144. <https://doi.org/10.1007/s10969-013-9163-9>.
  18. Bertelsen, A.B., Hackney, C.M., Bayer, C.N., Kjølgaard, L.D., Rennig, M., Christensen, B., Sørensen, E.S., Safavi-Hemami, H., Wulff, T., Ellgaard, L., and Nørholm, M.H.H. (2021). DisCoTune: versatile auxiliary plasmids for the production of disulphide-containing proteins and peptides in the *E. coli* T7 system. *Microb. Biotechnol.* 14, 2566–2580. <https://doi.org/10.1111/1751-7915.13895>.
  19. Atlas, R.M., and Atlas, R.M. (2004). *Handbook of Microbiological Media* 0, ed. (CRC Press). <https://doi.org/10.1201/9781420039726>.
  20. Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234. <https://doi.org/10.1016/j.pep.2005.01.016>.
  21. Taylor, T., Denson, J.-P., and Esposito, D. (2017). Optimizing Expression and Solubility of Proteins in *E. coli* Using Modified Media and Induction Parameters. In *Heterologous Gene Expression in E.coli Methods in Molecular Biology*, N.A. Burgess-Brown, ed. (Springer New York), pp. 65–82. [https://doi.org/10.1007/978-1-4939-6887-9\\_5](https://doi.org/10.1007/978-1-4939-6887-9_5).
  22. Saez, N.J., and Vincentelli, R. (2014). High-Throughput Expression Screening and Purification of Recombinant Proteins in *E. coli*. In *Structural Genomics Methods in Molecular Biology*, Y.W. Chen, ed. (Humana Press), pp. 33–53. [https://doi.org/10.1007/978-1-62703-691-7\\_3](https://doi.org/10.1007/978-1-62703-691-7_3).
  23. Dümmler, A., Lawrence, A.-M., and De Marco, A. (2005). Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microb. Cell Factories* 4, 34. <https://doi.org/10.1186/1475-2859-4-34>.
  24. De Marco, A. (2007). Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. *Nat. Protoc.* 2, 2632–2639. <https://doi.org/10.1038/nprot.2007.400>.
  25. Peleg, Y., Vincentelli, R., Collins, B.M., Chen, K.-E., Livingstone, E.K., Weeratunga, S., Leneva, N., Guo, Q., Remans, K., Perez, K., et al. (2021). Community-Wide Experimental Evaluation of the PROSS Stability-Design Method. *J. Mol. Biol.* 433, 166964. <https://doi.org/10.1016/j.jmb.2021.166964>.
  26. Humer, D., and Spadiut, O. (2018). Wanted: more monitoring and control during inclusion body processing. *World J. Microbiol. Biotechnol.* 34, 158. <https://doi.org/10.1007/s11274-018-2541-5>.
  27. Singh, S.M., and Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 99, 303–310. <https://doi.org/10.1263/jbb.99.303>.
  28. De Marco, A. (2017). Acting on Folding Effectors to Improve Recombinant Protein Yields and Functional Quality. In *Heterologous Gene Expression in E.coli Methods in Molecular Biology*, N.A. Burgess-Brown, ed. (Springer New York), pp. 197–210. [https://doi.org/10.1007/978-1-4939-6887-9\\_12](https://doi.org/10.1007/978-1-4939-6887-9_12).
  29. Rosano, G.L., Morales, E.S., and Ceccarelli, E.A. (2019). New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Sci.* 28, 1412–1422. <https://doi.org/10.1002/pro.3668>.
  30. Karyolaimos, A., and De Gier, J.-W. (2021). Strategies to Enhance Periplasmic Recombinant Protein Production Yields in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 9, 797334. <https://doi.org/10.3389/fbioe.2021.797334>.
  31. Royes, J., Talbot, P., Le Bon, C., Moncoq, K., Uzan, M., Zito, F., and Miroux, B. (2022). Membrane Protein Production in *Escherichia coli*: Protocols and Rules. In *Heterologous Expression of Membrane Proteins Methods in Molecular Biology*, I. Mus-Veteau, ed. (Springer US), pp. 19–39. [https://doi.org/10.1007/978-1-0716-2368-8\\_2](https://doi.org/10.1007/978-1-0716-2368-8_2).
  32. Rong, Y., Jensen, S.I., Lindorff-Larsen, K., and Nielsen, A.T. (2023). Folding of heterologous proteins in bacterial cell factories: Cellular mechanisms and engineering strategies. *Biotechnol. Adv.* 63, 108079. <https://doi.org/10.1016/j.biotechadv.2022.108079>.
  33. Patra, P., Das, M., Kundu, P., and Ghosh, A. (2021). Recent advances in systems and synthetic biology approaches for developing novel cell-factories in non-conventional yeasts. *Biotechnol. Adv.* 47, 107695. <https://doi.org/10.1016/j.biotechadv.2021.107695>.
  34. Baghban, R., Farajnia, S., Rajabibazli, M., Ghasemi, Y., Mafi, A., Hoseinpour, R., Rahbarnia, L., and Aria, M. (2019). Yeast Expression Systems: Overview and Recent Advances. *Mol. Biotechnol.* 61, 365–384. <https://doi.org/10.1007/s12033-019-00164-8>.
  35. de Sá Magalhães, S., and Keshavarz-Moore, E. (2021). *Pichia pastoris* (Komagataella phaffii) as a Cost-Effective Tool for Vaccine Production for Low- and Middle-Income Countries (LMICs). *Bioengineering* 8, 119. <https://doi.org/10.3390/bioengineering8090119>.
  36. García-Ortega, X., Cámara, E., Ferrer, P., Albiol, J., Montesinos-Seguí, J.L., and Valero, F. (2019). Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (Komagataella phaffii) using the methanol-free GAP promoter. Where do we stand? *N. Biotech.* 53, 24–34. <https://doi.org/10.1016/j.nbt.2019.06.002>.
  37. Liu, C., Gong, J.-S., Su, C., Li, H., Li, H., Rao, Z.-M., Xu, Z.-H., and Shi, J.-S. (2022). Pathway engineering facilitates efficient protein expression in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 106, 5893–5912. <https://doi.org/10.1007/s00253-022-12139-y>.
  38. Laukens, B., Jacobs, P.P., Geysens, K., Martins, J., De Wachter, C., Ameloot, P., Morelle, W., Hastraete, J., Renauld, J.C., Samyn, B., et al. (2020). Off-target glycans encountered along the synthetic biology route toward humanized N-glycans in *Pichia*

- pastoris*. Biotechnol. Bioeng. 117, 2479–2488. <https://doi.org/10.1002/bit.27375>.
39. Gao, J., Jiang, L., and Lian, J. (2021). Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products. Synth. Syst. Biotechnol. 6, 110–119. <https://doi.org/10.1016/j.synbio.2021.04.005>.
40. Van Herpe, D., Vanluchene, R., Vandewalle, K., Vanmarcke, S., Wyseure, E., Van Moer, B., Eeckhaut, H., Fijalkowska, D., Grootaert, H., Lonigro, C., et al. (2022). OPENPichia: Building a Free-To-Operate *Komagataella phaffii* Protein Expression Toolkit. Mol. Biol. <https://doi.org/10.1101/2022.12.13.519130>.
41. Garrigós-Martínez, J., Vuoristo, K., Nieto-Taype, M.A., Tähtiharju, J., Uusitalo, J., Tukiainen, P., Schmid, C., Tolstorukov, I., Madden, K., Penttilä, M., et al. (2021). Bioprocess performance analysis of novel methanol-independent promoters for recombinant protein production with *Pichia pastoris*. Microb. Cell Factories 20, 74. <https://doi.org/10.1186/s12934-021-01564-9>.
42. Dagar, V.K., and Khasa, Y.P. (2018). Combined effect of gene dosage and process optimization strategies on high-level production of recombinant human interleukin-3 (hIL-3) in *Pichia pastoris* fed-batch culture. Int. J. Biol. Macromol. 108, 999–1009. <https://doi.org/10.1016/j.ijbiomac.2017.11.008>.
43. Jacobs, P.P., Inan, M., Festjens, N., Hastraete, J., Van Hecke, A., Contreras, R., Meagher, M.M., and Callewaert, N. (2010). Fed-batch fermentation of GM-CSF-producing glycoengineered *Pichia pastoris* under controlled specific growth rate. Microb. Cell Factories 9, 93. <https://doi.org/10.1186/1475-2859-9-93>.
44. Matsuzaki, Y., Kajiwar, K., Aoki, W., and Ueda, M. (2022). Production of Single-Domain Antibodies in *Pichia pastoris*. In Single-Domain Antibodies Methods in Molecular Biology, G. Hussack and K.A. Henry, eds. (Springer US), pp. 181–203. [https://doi.org/10.1007/978-1-0716-2075-5\\_9](https://doi.org/10.1007/978-1-0716-2075-5_9).
45. Schoonoghe, S., Leoen, J., and Hastraete, J. (2012). Production of Antibody Derivatives in the Methylophilic Yeast *Pichia pastoris*. In Antibody Engineering Methods in Molecular Biology, P. Chames, ed. (Humana Press), pp. 325–340. [https://doi.org/10.1007/978-1-61779-974-7\\_19](https://doi.org/10.1007/978-1-61779-974-7_19).
46. Wang, G., Huang, M., and Nielsen, J. (2017). Exploring the potential of *Saccharomyces cerevisiae* for biopharmaceutical protein production. Curr. Opin. Biotechnol. 48, 77–84. <https://doi.org/10.1016/j.copbio.2017.03.017>.
47. De, S., Mattanovich, D., Ferrer, P., and Gasser, B. (2021). Established tools and emerging trends for the production of recombinant proteins and metabolites in *Pichia pastoris*. Essays Biochem. 65, 293–307. <https://doi.org/10.1042/EBC20200138>.
48. Mastropietro, G., Aw, R., and Polizzi, K.M. (2021). Expression of proteins in *Pichia pastoris*. In Methods in Enzymology (Elsevier), pp. 53–80. <https://doi.org/10.1016/bs.mie.2021.07.004>.
49. Rinnofner, C., Felber, M., and Pichler, H. (2022). Strains and Molecular Tools for Recombinant Protein Production in *Pichia pastoris*. In Yeast Metabolic Engineering Methods in Molecular Biology, V. Mapelli and M. Bettiga, eds. (Springer US), pp. 79–112. [https://doi.org/10.1007/978-1-0716-2399-2\\_6](https://doi.org/10.1007/978-1-0716-2399-2_6).
50. Higgins, D.R., and Cregg, J.M. (1998). Introduction to *Pichia pastoris*. In *Pichia* Protocols Methods in Molecular Biology, D.R. Higgins and J.M. Cregg, eds. (Humana Press), pp. 1–15. <https://doi.org/10.1385/0-89603-421-6:1>.
51. Gupta, K., Tölzer, C., Sari-Ak, D., Fitzgerald, D.J., Schaffitzel, C., and Berger, I. (2019). MultiBac: Baculovirus-Mediated Multigene DNA Cargo Delivery in Insect and Mammalian Cells. Viruses 11, 198. <https://doi.org/10.3390/v11030198>.
52. Mahajan, P., Ellis, K., Mukhopadhyay, S., Fernandez-Cid, A., Chi, G., Man, H., Dürr, K.L., and Burgess-Brown, N.A. (2021). Expression Screening of Human Integral Membrane Proteins Using BacMam. In Structural Genomics Methods in Molecular Biology, Y.W. Chen and C.-P.B. Yiu, eds. (Springer US), pp. 95–115. [https://doi.org/10.1007/978-1-0716-0892-0\\_6](https://doi.org/10.1007/978-1-0716-0892-0_6).
53. Smith, G.E., Ju, G., Ericson, B.L., Moschera, J., Lahm, H.W., Chizzonite, R., and Summers, M.D. (1985). Modification and secretion of human interleukin 2 produced in insect cells by a baculovirus expression vector. Proc. Natl. Acad. Sci. USA 82, 8404–8408. <https://doi.org/10.1073/pnas.82.24.8404>.
54. Le, L.T.M., Nyengaard, J.R., Golas, M.M., and Sander, B. (2018). Vectors for Expression of Signal Peptide-Dependent Proteins in Baculovirus/Insect Cell Systems and Their Application to Expression and Purification of the High-Affinity Immunoglobulin Gamma Fc Receptor I in Complex with Its Gamma Chain. Mol. Biotechnol. 60, 31–40. <https://doi.org/10.1007/s12033-017-0041-8>.
55. Wang, Y., Mao, Y., Xu, X., Tao, S., and Chen, H. (2015). Codon Usage in Signal Sequences Affects Protein Expression and Secretion Using Baculovirus/Insect Cell Expression System. PLoS One 10, e0145887. <https://doi.org/10.1371/journal.pone.0145887>.
56. Soejima, Y., Lee, J., Nagata, Y., Mon, H., Iiyama, K., Kitano, H., Matsuyama, M., and Kusakabe, T. (2013). Comparison of signal peptides for efficient protein secretion in the baculovirus-silkworm system. Open Life Sci. 8, 1–7. <https://doi.org/10.2478/s11535-012-0112-6>.
57. Ailor, E., and Betenbaugh, M.J. (1999). Modifying secretion and post-translational processing in insect cells. Curr. Opin. Biotechnol. 10, 142–145. [https://doi.org/10.1016/S0958-1669\(99\)80024-X](https://doi.org/10.1016/S0958-1669(99)80024-X).
58. Futatsumori-Sugai, M., and Tsumoto, K. (2010). Signal peptide design for improving recombinant protein secretion in the baculovirus expression vector system. Biochem. Biophys. Res. Commun. 391, 931–935. <https://doi.org/10.1016/j.bbrc.2009.11.167>.
59. Palmberger, D., Wilson, I.B.H., Berger, I., Grabherr, R., and Rendic, D. (2012). SweetBac: A New Approach for the Production of Mammalian Glycoproteins in Insect Cells. PLoS One 7, e34226. <https://doi.org/10.1371/journal.pone.0034226>.
60. Maghodia, A.B., Geisler, C., and Jarvis, D.L. (2021). A New Bacmid for Customized Protein Glycosylation Pathway Engineering in the Baculovirus-Insect Cell System. ACS Chem. Biol. 16, 1941–1950. <https://doi.org/10.1021/acscchembio.0c00974>.
61. Lemaître, R.P., Bogdanova, A., Borgonovo, B., Woodruff, J.B., and Drechsel, D.N. (2019). FlexiBAC: a versatile, open-source baculovirus vector system for protein expression, secretion, and proteolytic processing. BMC Biotechnol. 19, 20. <https://doi.org/10.1186/s12896-019-0512-z>.
62. Stolt-Bergner, P., Benda, C., Bergbrede, T., Besir, H., Celie, P.H.N., Chang, C., Drechsel, D., Fischer, A., Geerlof, A., Giabbai, B., et al. (2018). Baculovirus-driven protein expression in insect cells: A benchmarking study. J. Struct. Biol. 203, 71–80. <https://doi.org/10.1016/j.jsb.2018.03.004>.
63. Scholz, J., and Suppmann, S. (2021). A fast-track protocol for protein expression using the BEV system. Methods Enzymol. 660, 171–190. <https://doi.org/10.1016/bs.mie.2021.06.015>.
64. Wasilko, D.J., Lee, S.E., Stutzman-Engwall, K.J., Reitz, B.A., Emmons, T.L., Mathis, K.J., Bienkowski, M.J., Tomasselli, A.G., and Fischer, H.D. (2009). The tierless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with recombinant baculovirus. Protein Expr. Purif. 65, 122–132. <https://doi.org/10.1016/j.pep.2009.01.002>.
65. Weissmann, F., Petzold, G., VanderLinden, R., Huis In 't Veld, P.J., Brown, N.G., Lampert, F., Westermann, S., Stark, H., Schulman, B.A., and Peters, J.-M. (2016). biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. Proc. Natl. Acad. Sci. USA 113. <https://doi.org/10.1073/pnas.1604935113>.
66. Miller-Vedam, L.E., Bräuning, B., Popova, K.D., Schirle Oakdale, N.T., Bonnar, J.L., Prabu, J.R., Boydston, E.A., Seviliano, N., Shurtleff, M.J., Stroud, R.M., et al. (2020). Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. eLife 9, e62611. <https://doi.org/10.7554/eLife.62611>.
67. Neuhold, J., Radakovics, K., Lehner, A., Weissmann, F., García, M.C., Romero, M.C., Berrow, N.S., and Stolt-Bergner, P. (2020). GoldenBac: a simple, highly efficient, and widely applicable system for construction of multi-gene expression vectors for use with the baculovirus expression vector system. BMC Biotechnol. 20, 26. <https://doi.org/10.1186/s12896-020-00616-z>.
68. Berger, I., Fitzgerald, D.J., and Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. Nat. Biotechnol. 22, 1583–1587. <https://doi.org/10.1038/nbt1036>.

69. Berger, I., and Poterszman, A. (2015). Baculovirus expression: old dog, new tricks. *Bioengineered* 6, 316–322. <https://doi.org/10.1080/21655979.2015.1104433>.
70. Clem, R.J., and Passarelli, A.L. (2013). Baculoviruses: Sophisticated Pathogens of Insects. *PLoS Pathog.* 9, e1003729. <https://doi.org/10.1371/journal.ppat.1003729>.
71. Scholz, J., and Suppmann, S. (2021). A fast-track protocol for protein expression using the BEV system. *Methods Enzymol.* 660, 171–190. <https://doi.org/10.1016/bs.mie.2021.06.015>.
72. Weissmann, F., and Peters, J.-M. (2018). Expressing Multi-subunit Complexes Using biGbac. In *Protein Complex Assembly Methods in Molecular Biology*, J.A. Marsh, ed. (Springer New York), pp. 329–343. [https://doi.org/10.1007/978-1-4939-7759-8\\_21](https://doi.org/10.1007/978-1-4939-7759-8_21).
73. Shen, X., Hacker, D.L., Baldi, L., and Wurm, F.M. (2014). Virus-free transient protein production in Sf9 cells. *J. Biotechnol.* 171, 61–70. <https://doi.org/10.1016/j.jbiotec.2013.11.018>.
74. Bleckmann, M., Fritz, M.H.-Y., Bhuj, S., Jarek, M., Schürig, M., Geffers, R., Benes, V., Besir, H., and Van Den Heuvel, J. (2015). Genomic Analysis and Isolation of RNA Polymerase II Dependent Promoters from *Spodoptera frugiperda*. *PLoS One* 10, e0132898. <https://doi.org/10.1371/journal.pone.0132898>.
75. Jarvis, D.L. (2014). Recombinant Protein Expression in Baculovirus-Infected Insect Cells. In *Methods in Enzymology* (Elsevier), pp. 149–163. <https://doi.org/10.1016/B978-0-12-420070-8.00013-1>.
76. Bleckmann, M., Schürig, M., Chen, F.-F., Yen, Z.-Z., Lindemann, N., Meyer, S., Spehr, J., and Van Den Heuvel, J. (2016). Identification of Essential Genetic Baculoviral Elements for Recombinant Protein Expression by Transactivation in Sf21 Insect Cells. *PLoS One* 11, e0149424. <https://doi.org/10.1371/journal.pone.0149424>.
77. Bleckmann, M., Schürig, M., Endres, M., Samuels, A., Gebauer, D., Konisch, N., and Van Den Heuvel, J. (2019). Identifying parameters to improve the reproducibility of transient gene expression in High Five cells. *PLoS One* 14, e0217878. <https://doi.org/10.1371/journal.pone.0217878>.
78. Mori, K., Hamada, H., Ogawa, T., Ohmuro-Matsuyama, Y., Katsuda, T., and Yamaji, H. (2017). Efficient production of antibody Fab fragment by transient gene expression in insect cells. *J. Biosci. Bioeng.* 124, 221–226. <https://doi.org/10.1016/j.jbiosc.2017.03.007>.
79. Puente-Massaguer, E., Lecina, M., and Gòdia, F. (2018). Nanoscale characterization coupled to multi-parametric optimization of Hi5 cell transient gene expression. *Appl. Microbiol. Biotechnol.* 102, 10495–10510. <https://doi.org/10.1007/s00253-018-9423-5>.
80. Puente-Massaguer, E., Strobl, F., Grabherr, R., Striedner, G., Lecina, M., and Gòdia, F. (2020). PEI-Mediated Transient Transfection of High Five Cells at Bioreactor Scale for HIV-1 VLP Production. *Nanomaterials* 10, 1580. <https://doi.org/10.3390/nano10081580>.
81. Korn, J., Schäckermann, D., Kirmann, T., Bertoglio, F., Steinke, S., Heisig, J., Ruschig, M., Rojas, G., Langreder, N., Wenzel, E.V., et al. (2020). Baculovirus-free insect cell expression system for high yield antibody and antigen production. *Sci. Rep.* 10, 21393. <https://doi.org/10.1038/s41598-020-78425-9>.
82. Kaipa, J.M., Krasnoselska, G., Owens, R.J., and van den Heuvel, J. (2023). Screening of Membrane Protein Production by Comparison of Transient Expression in Insect and Mammalian Cells. *Biomolecules* 13, 817. <https://doi.org/10.3390/biom13050817>.
83. Shen, X., Pitol, A.K., Bachmann, V., Hacker, D.L., Baldi, L., and Wurm, F.M. (2015). A simple plasmid-based transient gene expression method using High Five cells. *J. Biotechnol.* 216, 67–75. <https://doi.org/10.1016/j.jbiotec.2015.10.007>.
84. Kesidis, A., Depping, P., Lodé, A., Vaitsoyopoulou, A., Bill, R.M., Goddard, A.D., and Rothnie, A.J. (2020). Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts. *Methods* 180, 3–18. <https://doi.org/10.1016/j.jymeth.2020.06.006>.
85. Pieprzyk, J., Pazicky, S., and Löw, C. (2018). Transient Expression of Recombinant Membrane-eGFP Fusion Proteins in HEK293 Cells. In *Recombinant Protein Expression in Mammalian Cells Methods in Molecular Biology*, D.L. Hacker, ed. (Springer New York), pp. 17–31. [https://doi.org/10.1007/978-1-4939-8730-6\\_2](https://doi.org/10.1007/978-1-4939-8730-6_2).
86. Goehring, A., Lee, C.-H., Wang, K.H., Michel, J.C., Claxton, D.P., Bacongus, I., Althoff, T., Fischer, S., Garcia, K.C., and Gouaux, E. (2014). Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* 9, 2574–2585. <https://doi.org/10.1038/nprot.2014.173>.
87. Durocher, Y., Perret, S., and Kamen, A. (2002). High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res.* 30, 9e–99e. <https://doi.org/10.1093/nar/30.2.e9>.
88. Aricescu, A.R., Lu, W., and Jones, E.Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1243–1250. <https://doi.org/10.1107/S0907444906002979>.
89. Delafosse, L., Xu, P., and Durocher, Y. (2016). Comparative study of polyethylenimines for transient gene expression in mammalian HEK293 and CHO cells. *J. Biotechnol.* 227, 103–111. <https://doi.org/10.1016/j.jbiotec.2016.04.028>.
90. Kadlecova, Z., Nallet, S., Hacker, D.L., Baldi, L., Klok, H.-A., and Wurm, F.M. (2012). Poly(ethyleneimine)-mediated large-scale transient gene expression: Influence of molecular weight, polydispersity and N-propionyl groups. *Macromol. Biosci.* 12, 628–636. <https://doi.org/10.1002/mabi.201100404>.
91. Baldi, L., Hacker, D.L., Meerschman, C., and Wurm, F.M. (2012). Large-Scale Transfection of Mammalian Cells. In *Protein Expression in Mammalian Cells Methods in Molecular Biology*, J.L. Hartley, ed. (Humana Press), pp. 13–26. [https://doi.org/10.1007/978-1-61779-352-3\\_2](https://doi.org/10.1007/978-1-61779-352-3_2).
92. Fornwald, J.A., Lu, Q., Boyce, F.M., and Ames, R.S. (2016). Gene Expression in Mammalian Cells Using BacMam, a Modified Baculovirus System. In *Baculovirus and Insect Cell Expression Protocols Methods in Molecular Biology*, D.W. Murhammer, ed. (Springer New York), pp. 95–116. [https://doi.org/10.1007/978-1-4939-3043-2\\_5](https://doi.org/10.1007/978-1-4939-3043-2_5).
93. Barsoum, J., Brown, R., McKee, M., and Boyce, F.M. (1997). Efficient Transduction of Mammalian Cells by a Recombinant Baculovirus Having the Vesicular Stomatitis Virus G Glycoprotein. *Hum. Gene Ther.* 8, 2011–2018. <https://doi.org/10.1089/hum.1997.8.17-2011>.
94. Mansouri, M., Bellon-Echeverria, I., Rizk, A., Ehsaei, Z., Cianciolo Cosentino, C., Silva, C.S., Xie, Y., Boyce, F.M., Davis, M.W., Neuhauss, S.C.F., et al. (2016). Highly efficient baculovirus-mediated multigene delivery in primary cells. *Nat. Commun.* 7, 11529. <https://doi.org/10.1038/ncomms11529>.
95. Bellón-Echeverría, I., Carralot, J.-P., Del Rosario, A.A., Kueng, S., Mauser, H., Schmid, G., Thoma, R., and Berger, I. (2018). MultiBacMam Bimolecular Fluorescence Complementation (BiFC) tool-kit identifies new small-molecule inhibitors of the CDK5-p25 protein-protein interaction (PPI). *Sci. Rep.* 8, 5083. <https://doi.org/10.1038/s41598-018-23516-x>.
96. Elegheert, J., Behiels, E., Bishop, B., Scott, S., Woolley, R.E., Griffiths, S.C., Byrne, E.F.X., Chang, V.T., Stuart, D.I., Jones, E.Y., et al. (2018). Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat. Protoc.* 13, 2991–3017. <https://doi.org/10.1038/s41596-018-0075-9>.
97. Behiels, E., and Elegheert, J. (2021). High-Level Production of Recombinant Eukaryotic Proteins from Mammalian Cells Using Lentivirus. In *Structural Proteomics Methods in Molecular Biology*, R.J. Owens, ed. (Springer US), pp. 83–104. [https://doi.org/10.1007/978-1-0716-1406-8\\_4](https://doi.org/10.1007/978-1-0716-1406-8_4).
98. Tschorn, N., Berg, K., and Stitz, J. (2020). Transposon vector-mediated stable gene transfer for the accelerated establishment of recombinant mammalian cell pools allowing for high-yield production of biologics. *Biotechnol. Lett.* 42, 1103–1112. <https://doi.org/10.1007/s10529-020-02889-y>.
99. Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. USA* 108, 1531–1536. <https://doi.org/10.1073/pnas.1008322108>.
100. Li, Z., Michael, I.P., Zhou, D., Nagy, A., and Rini, J.M. (2013). Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. *Proc. Natl. Acad. Sci. USA* 110, 5004–5009. <https://doi.org/10.1073/pnas.1218620110>.
101. Balasubramanian, S., Matasci, M., Kadlecova, Z., Baldi, L., Hacker, D.L., and Wurm, F.M. (2015). Rapid recombinant protein production from piggyBac transposon-mediated stable

- CHO cell pools. *J. Biotechnol.* 200, 61–69. <https://doi.org/10.1016/j.jbiotec.2015.03.001>.
102. Caro, L.N., Li, Z., Balo, A.R., Van Eps, N., Rini, J.M., and Ernst, O.P. (2015). Rapid and Facile Recombinant Expression of Bovine Rhodopsin in HEK293S GnTI<sup>−</sup> Cells Using a PiggyBac Inducible System. In *Methods in Enzymology* (Elsevier), pp. 307–330. <https://doi.org/10.1016/bs.mie.2015.01.005>.
103. Hacker, D.L., and Balasubramanian, S. (2016). Recombinant protein production from stable mammalian cell lines and pools. *Curr. Opin. Struct. Biol.* 38, 129–136. <https://doi.org/10.1016/j.sbi.2016.06.005>.
104. Michael, I.P., and Nagy, A. (2018). Inducible Protein Production in 293 Cells Using the piggyBac Transposon System. In *Recombinant Protein Expression in Mammalian Cells Methods in Molecular Biology*, D.L. Hacker, ed. (Springer New York), pp. 57–68. [https://doi.org/10.1007/978-1-4939-8730-6\\_5](https://doi.org/10.1007/978-1-4939-8730-6_5).
105. Suppmann, S. (2021). Inducible protein expression in piggyBac transposase mediated stable HEK293 cell pools. In *Methods in Enzymology* (Elsevier), pp. 321–339. <https://doi.org/10.1016/bs.mie.2021.06.016>.
106. Silverman, A.D., Karim, A.S., and Jewett, M.C. (2020). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. <https://doi.org/10.1038/s41576-019-0186-3>.
107. Foshag, D., Henrich, E., Hiller, E., Schäfer, M., Kerger, C., Burger-Kentischer, A., Diaz-Moreno, I., García-Mauriño, S.M., Dötsch, V., Rupp, S., and Bernhard, F. (2018). The *E. coli* S30 lysate proteome: A prototype for cell-free protein production. *N. Biotech.* 40, 245–260. <https://doi.org/10.1016/j.nbt.2017.09.005>.
108. Kögler, L.M., Stichel, J., and Beck-Sickinger, A.G. (2019). Structural investigations of cell-free expressed G protein-coupled receptors. *Biol. Chem.* 401, 97–116. <https://doi.org/10.1515/hsz-2019-0292>.
109. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. <https://doi.org/10.1038/90802>.
110. Zemella, A., Grossmann, S., Sachse, R., Son nabend, A., Schaefer, M., and Kubick, S. (2017). Qualifying a eukaryotic cell-free system for fluorescence based GPCR analyses. *Sci. Rep.* 7, 3740. <https://doi.org/10.1038/s41598-017-03955-8>.
111. Merk, H., Rues, R.-B., Gless, C., Beyer, K., Dong, F., Dötsch, V., Gerrits, M., and Bernhard, F. (2015). Biosynthesis of membrane dependent proteins in insect cell lysates: identification of limiting parameters for folding and processing. *Biol. Chem.* 396, 1097–1107. <https://doi.org/10.1515/hsz-2015-0105>.
112. Hein, C., Henrich, E., Orbán, E., Dötsch, V., and Bernhard, F. (2014). Hydrophobic supplements in cell-free systems: Designing artificial environments for membrane proteins. *Eng. Life Sci.* 14, 365–379. <https://doi.org/10.1002/elsc.201300050>.
113. Bruni, R., Laguerre, A., Kaminska, A.M., McSweeney, S., Hendrickson, W.A., and Liu, Q. (2022). High-throughput cell-free screening of eukaryotic membrane protein expression in lipidic mimetics. *Protein Sci.* 31, 639–651. <https://doi.org/10.1002/pro.4259>.
114. Kuruma, Y., and Ueda, T. (2015). The PURE system for the cell-free synthesis of membrane proteins. *Nat. Protoc.* 10, 1328–1344. <https://doi.org/10.1038/nprot.2015.082>.
115. Krug, U., Gloge, A., Schmidt, P., Becker-Baldus, J., Bernhard, F., Kaiser, A., Montag, C., Gauglitz, M., Vishnivetskiy, S.A., Gurevich, V.V., et al. (2020). The Conformational Equilibrium of the Neuropeptide Y2 Receptor in Bilayer Membranes. *Angew. Chem. Int. Ed.* 59, 23854–23861. <https://doi.org/10.1002/anie.202006075>.
116. Reckel, S., Gottstein, D., Stehle, J., Löhr, F., Verhoeven, M.-K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., et al. (2011). Solution NMR Structure of Proteorhodopsin. *Angew. Chem. Int. Ed.* 50, 11942–11946. <https://doi.org/10.1002/anie.201105648>.
117. Boland, C., Li, D., Shah, S.T.A., Habersack, S., Dötsch, V., Bernhard, F., and Caffrey, M. (2014). Cell-free expression and in meso crystallisation of an integral membrane kinase for structure determination. *Cell. Mol. Life Sci.* 71, 4895–4910. <https://doi.org/10.1007/s00018-014-1655-7>.
118. Garenne, D., Haines, M.C., Romantseva, E.F., Freemont, P., Strychalski, E.A., and Noireaux, V. (2021). Cell-free gene expression. *Nat. Rev. Methods Primers* 1, 49. <https://doi.org/10.1038/s43586-021-00046-x>.
119. Takeda, M., and Kainosho, M. (2012). Cell-Free Protein Production for NMR Studies. In *Protein NMR Techniques Methods in Molecular Biology*, A. Shekhtman and D.S. Burz, eds. (Humana Press), pp. 71–84. [https://doi.org/10.1007/978-1-61779-480-3\\_5](https://doi.org/10.1007/978-1-61779-480-3_5).