FEMS Yeast Research, 2025, **25**, foaf010
DOI: 10.1093/femsyr/foaf010
Advance access publication date: 12 March 2025
Minireview – Industrial and environmental biotechnology

Advancing recombinant protein expression in *Komagataella phaffii*: opportunities and challenges

Wen Lv[1] and Menghao Cai[1,2,*]

[1]State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China
[2]Shanghai Collaborative Innovation Center for Biomanufacturing, 130 Meilong Road, Shanghai 200237, China
*Corresponding author. State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China. E-mail: cmh022199@ecust.edu.cn
Editor: [Hyun Ah Kang]

Abstract

Komagataella phaffii has gained recognition as a versatile platform for recombinant protein production, with applications covering biopharmaceuticals, industrial enzymes, food additives, etc. Its advantages include high-level protein expression, moderate post-translational modifications, high-density cultivation, and cost-effective methanol utilization. Nevertheless, it still faces challenges for the improvement of production efficiency and extension of applicability. This review highlights the key strategies used to facilitate productivity in *K. phaffii*, including systematic advances in genetic manipulation tools, transcriptional and translational regulation, protein folding and secretion optimization. Glycosylation engineering is also concerned as it enables humanized glycosylation profiles for the use in therapeutic proteins and functional food additivities. Omics technologies and genome-scale metabolic models provide new insights into cellular metabolism, enhancing recombinant protein expression. High-throughput screening technologies are also emphasized as crucial for constructing high-expression strains and accelerating strain optimization. With advancements in gene-editing, synthetic and systems biology tools, the *K. phaffii* expression platform has been significantly improved for fundamental research and industrial use. Future innovations aim to fully harness *K. phaffii* as a next-generation cell factory, providing efficient, scalable, and cost-effective solutions for diverse applications. It continues to hold promise as a key driver in the field of biotechnology.

Keywords: *Komagataella phaffii*; non-conventional yeast; protein expression regulation; advanced methods; biotechnology applications

Introduction

Recombinant proteins, as core products in the field of biotechnology, have found extensive applications in production of biopharmaceuticals (Duarte and Monteiro. 2021), food additives (Knychala et al. 2024), industrial enzymes (Chen et al. 2024), etc. The increasing demand for recombinant protein production, driven by rapid biotechnological advancements (Lu et al. 2024), has underscored the importance of selecting the optimal expression system. The primary biological expression systems can be categorized into prokaryotic microorganisms, eukaryotic microorganisms, and mammalian cell systems. The *Escherichia coli* expression system is widely utilized due to its rapid growth rate, low culture cost, and straightforward operation. However, it faces many challenges, including improper protein folding, inclusion body formation and the absence of post-translational modifications (Jiang et al. 2024a). Mammalian expression systems, such as Chinese hamster ovary cells, excel at performing complex post-translational modifications and ensuring proper protein folding and glycosylation (Shi and McHugh. 2023). Despite these advantages, they are hampered by high production costs, slow cell growth, and expensive nutritional requirements (Elegheert et al. 2018). In comparison, yeast expression systems, particularly the non-conventional yeast *Komagataella phaffii* (syn. *Pichia pastoris*), offer a balance between high-efficiency expression and cost-effectiveness, making them uniquely suited for recombinant protein production.

Komagataella phaffii is a generally recognized as safe strain approved by the the American Food and Drug Administration (FDA) (Vuree 2020). It is an ideal platform for recombinant protein production due to its clear genetic background, fully sequenced genome and accessible molecular manipulation tools (Sun et al. 2022). As an eukaryotic host, *K. phaffii* supports proper protein folding and performs essential post-translational modifications, such as glycosylation (Deng et al. 2023) and phosphorylation (Papala et al. 2017), which are critical for the functionality of many biopharmaceutical products. Its robust secretory capacity allows recombinant proteins to be efficiently secreted into the culture medium, minimizing intracellular accumulation and streamlining downstream purification processes (Gao et al. 2023a). Moreover, *K. phaffii* produces less endogenous secreted proteins, reducing interference from non-target proteins and improving the purity of the desired product (Tian et al. 2024). *Komagataella phaffii* also demonstrates an efficient ability to utilize inexpensive raw materials such as methanol, making it advantageous for large-scale industrial production (Karbalaei et al. 2020b). It can achieve high-density fermentation (approximately 150 g/l dry cell weight) within a short period (Barone et al. 2023), which is beneficial for improving recombinant protein yield and overall production efficiency. Till now, *K. phaffii* has been extensively applied for the pro-

Received 18 December 2024; **revised** 19 February 2025; **accepted** 11 March 2025

duction of vaccines, enzymes and antibodies, which represents a preferred eukaryotic expression system in industry (Nand et al. 2024).

This review summarizes the advantages and challenges of *K. phaffii* in recombinant protein expression, focusing on the processes of transcription, translation, protein folding, secretion, and omics technologies and genome-scale metabolic models (GEMs). Additionally, high-throughput screening of high-expression strains is discussed as a crucial strategy for accelerating strain optimization. By addressing these aspects, it is expected to further unlock the potential of *K. phaffii* as an expression host, offering more efficient and cost-effective solutions for thepharmaceutical and industrial sectors.

Host characteristics of *K. phaffii* for recombinant protein production

Strain types and characteristics

Komagataella phaffii strains were isolated from oak trees in California and designated as *P. pastoris* in 1950s. These isolated *P. pastoris* strains were re-classified into the two species of *K. pastoris* and *K. phaffii* in 1990s. As *K. phaffii* has been widely used before re-classification, its original name *P. pastoris* is still in use especially in industry nowadays. In the early phase, the wild-type *K. phaffii* Y-11430 was widely used due to its rapid growth and efficient methanol utilization (Ogata et al. 1969). Then, the histidine auxotrophic strain GS115, obtained through nitrosoguanidine mutagenesis, simplifies the process of gene integration and screening, enabling its commercial use as a host (De Schutter et al. 2009). To enhance the extracellular expression of recombinant proteins, the protease-deficient SMD series of strains were developed to minimize degradation by endogenous proteases (Ahmad et al. 2014). The X-33 strain, a revertant of GS115 generated through *HIS4* complementation, is sometimes referred to as the wild type. It offers improved transformation efficiency and membrane protein secretion, making it a popular commercial expression strain. (Brady et al. 2020). Each strain shows distinct characteristics that contributing to the overall efficiency of the *K. phaffii* system, from facilitating gene integration to increasing the stability of protein secretion.

Flexible culture conditions and industrial application potentiality

Among various yeast species, *K. phaffii* stands out for its highly developed fermentation process, offering remarkable production advantages for recombinant protein production. Its cell dry weight can even reach up to 150 g/l and the fermentation scale can be expanded to 80 000 l, making it applicable for large-scale industrial production (Mattanovich et al. 2012). According to data from Research Corporation Technologies (http://pichia.com/), over 5 000 proteins have been successfully expressed in this host, among which over 70 commercial products have been commercialized, highlighting its widespread application and strong production performance (Zahrl et al. 2017).

During fermentation process, the composition of medium greatly impacts on protein expression by affecting cell growth and viability. *Komagataella phaffii* is capable of utilizing methanol as the sole carbon source. Methanol undergoes assimilation and dissimilation in peroxisomes, which supports cell growth and generates NADH (Antoniewicz 2019). Compared to traditional carbon sources like glucose and glycerol, methanol is more effective in promoting biosynthesis. For example, Cai et al. (2022) successfully enhanced the production of fatty acids and fatty alcohols by optimizing the methanol metabolic pathway and increasing the supply of precursors and cofactors. However, the flammability and explosiveness of methanol pose safety risks. To address this, Wang et al. (2018a) reconstructed the P_{AOX1} transcriptional regulation system, enabling efficient protein synthesis with decreased methanol consumption. Additionally, methanol-free expression systems have been developed in recent years. Liu et al. (2022a) replaced methanol with formate as the carbon source and inducer, improving safety and sustainability while increasing xylanase production by 103.5%. Similarly, Lee et al. (2017) developed a buffered extra-YNB glycerol methanol auto-induction media, which shortened the screening and cultivation time while maintaining membrane protein expression levels similar to traditional media. Furthermore, modification or reconstruction of methanol metabolic network may facilitate methanol metabolism and improve production processes of target products, advancing yeast platform upgrades.

In addition to methanol, *K. phaffii* has a broad substrate spectra including glucose, glycerol, sorbitol, ethanol, etc. Four genes encoding H^+/glycerol transporters have been identified in *K. phaffii*, enhancing its glycerol utilization efficiency compared to other yeasts (Mattanovich et al. 2009). It reached a maximum glycerol utilization rate (qSmax) of 0.37 h^{-1}. In contrast, the baker's yeast (*Saccharomyces cerevisiae*) showed a higher glucose utilization rate (Hagman et al. 2014), with a maximum glucose utilization rate of 2.88 h^{-1}, whereas *K. phaffii* achieved only 0.35 h^{-1}. However, as a typical Crabtree-negative yeast, *K. phaffii* avoids high-sugar fermentation and by-product formation under aerobic conditions, giving it a distinct advantage in industrial production. When methanol is used for inducible expression, the presence of carbon sources such as glycerol and glucose can inhibit the expression of foreign proteins. Therefore, sorbitol is commonly used as a substitution for glycerol. It does not affect activity of the P_{AOX1} promoter, but reduces cell growth rate and increase protein expression. Azadi et al. (2017) found that when sorbitol concentration in the medium reached 50 g/l, recombinant protein expression was highly enhanced. Ethanol is converted into acetyl-CoA through a three-step metabolic pathway. The efficient ethanol assimilation ability of *K. phaffii* provides ample energy and synthetic precursors for the cells. Liu et al. (2019b) developed an ethanol-inducible regulatory system ESAD in *K. phaffii* and established a novel biosynthesis system where ethanol serves as a carbon source, precursor, and inducer simultaneously. This system could be used for the expression of recombinant proteins, offering a versatile approach for metabolic engineering on this host strain.

Genetic manipulation tools and applications

Design and selection of expression systems

Most protein expression in *K. phaffii* relies on the integration of target genes into the genome through single- or double-crossover homologous recombination (HR). Commercial expression vectors including pPIC 9 K, pPIC 3.5 K, pAO815 are commonly used especially for the multi-copy gene integration. The P_{AOX1} promoter region and auxotrophic gene *HIS4* are widely used as single-crossover insertion sites, with integration efficiencies ranging from 50%∼80% (Cereghino and Cregg. 2000). However, this integration method may lead to a second-round single-crossover especially for multi-copy constructions, resulting in the loss of the target gene. Therefore,

researchers have developed double-crossover methods to replace specific sequences in the genome with an expression cassette, resulting in more stable and precise gene integration.

Although integrated expression vectors provide stable protein production, their complex operation and low transformation efficiency limit the optimization of industrial strains. To overcome these challenges, episomal expression vectors have emerged as valuable tools for high-throughput screening due to their easier manipulation and higher transformation efficiency. The autonomous replication sequence (ARS), a key part of episomal plasmids, controls DNA replication initiation, stability and copy number of plasmids (Gu et al. 2019). Although ARS such as PPARS1 and mitochondrial DNA fragments have been successfully used to construct episomal plasmids with high transformation efficiency, these plasmids often exhibit poor stability during mitosis and thus limiting their applications (Schwarzhans et al. 2017). Researchers have identified centromere (CEN) sequences on four chromosomes in this host to improve plasmid stability (Piva et al. 2020). However, the long length of the CEN sequence complicates plasmid transformation and may lead to genome integration, increasing the difficulties in screening procedure (Nakamura et al. 2018).

The construction of efficient protein expression chassis strains requires extensive metabolic reprogramming and expression of multiple genes. Compared to traditional plasmid-based expression systems, genome integration is considered as a viable alternative for the stable expression of genes. For genome integration, identification of neutral sites, i.e. chromosomal locations that can be disrupted with minimal impact on cell physiology and metabolism, is essential for stable multi-gene expression. Cai et al. (2021) identified 46 neutral sites in *K. phaffii*. Gao et al. (2022) further screened and identified 23 potential integration sites. Most recently, Ruan et al. (2024) utilized CRISPR-Cpf1 technology to discover 15 new and efficient integration sites in *K. phaffii*, offering more options for gene integration.

For genetic manipulation, marker genes are commonly used to select recombinant strains. In *K. phaffii*, antibiotic resistance genes, i.e. *Sh ble* (Zeocin^R) (Drocourt et al. 1990), *kan* (G418^R) (Lin-Cereghino et al. 2008), *hph* (hygromycin^R) (Yang et al. 2014), *nat* (nourseothricin^R) (Baghban et al. 2018), and *bsd* (blasticidin^R) (Masuda et al. 2005), and auxotrophic markers like *URA5*, *URA3*, *MET2*, and *HIS4* (Nett et al. 2005), are commonly used. Nevertheless, efficient marker genes are still limited when performing multiple gene manipulations and exploration of appropriate screening markers still keeps challenging. Introduction of reverse selection markers offers a solution by enabling marker-free gene manipulation, which provides a new tool for microbial genetic engineering. Researchers have incorporated the Cre/loxP recombinase system and reverse screening markers, such as *URA3* (Wang et al. 2023a) and *mazF* (Yang et al. 2009), into *K. phaffii* to enable marker removal through site-specific recombination. This reverse screening strategy improves the efficiency of marker recovery, while challenges such as long cycle times and complex procedures still needs to be solved.

Progress and potential applications of gene editing systems

Gene-editing technologies have revolutionized the ability to manipulate *K. phaffii* strains for recombinant protein production. In addition to traditional methods like golden-gate assembly and Cre/loxP system, the CRISPR/Cas system (Weninger et al. 2016) has highly improved the efficiency and accuracy of gene manipulation (Fig. 1A and B). By screening ARS and adjusting the promoter of gRNA expression, Gu et al. (2019) improved the editing efficiency of CRISPR/Cas9 by 10 times. Liu et al. (2019a) established an efficient multi-site integration technology based on CRISPR/Cas9 in *K. phaffii*. The gene editing achieved a double-site integration efficiency of 57.7%∼70% and a three-site integration efficiency of 12.5%∼32.1%. Moreover, Zhang et al. (2021) established a CRISPR-Cpf1-mediated genome editing method in *K. phaffii*, which efficiently enabled multi-gene knockout with deletions of up to 20 kb. It also facilitated one-step integration of multiple genes, with a three-gene editing efficiency of 30%.

Despite these advancements, the primary mechanism for repairing chromosome double-strand breaks (DSBs) in *K. phaffii* is non-homologous end joining (NHEJ) (McCarty et al. 2020), which leads to the randomness and uncontrollability of gene integration, making screening a time-consuming and labor-intensive process (Fig. 1C). Knockout of NHEJ-related gene *Ku70* and overexpression of HR-related proteins RAD52 and RAD59, increasing the efficiency of seamless knockout and gene integration to over 90% (Wang et al. 2023a). The targeted recognition function of Cas9 relies on the PAM sequence. A series of Cas9 proteins, including StCas9, spCas9, NmCas9, and CjCas9, which recognize different PAM sequences, have been successfully applied in gene-editing. These proteins are expected to become effective tools for gene-editing in *K. phaffii* in the future. Besides, dCas9 is widely used in gene inhibition and activation (CRISPRi and CRISPRa) (Fig. 1D). Unlike Cas9, dCas9 does not cause DNA double-strand breaks but instead regulates gene expression by fusing transcriptional repression (TRD) or activation (TAD) domains (Qiao et al. 2023). CRISPRi and CRISPRa have been well applied in *K. phaffii*, generating a strong and precise regulatory SynPic-X expression platform of user-defined signal inducible behavior (Liu et al. 2022b).

Transcriptional regulation strategies

Design and application of promoters and terminators

High titer, high yield, and efficient productivity are always desired for recombinant protein expression. To achieve these goals, optimization strategies typically focus on enhancing the efficiency of heterologous gene expression, modifying chassis cells (Wang et al. 2024a), and refining fermentation processes (Shemesh and Fishman 2024). Therein, modifying expression regulatory elements remains one of the most fundamental strategies for expression optimization.

Komagataella phaffii possesses a highly effective transcriptional regulation system. Promoter plays a critical role as a core element to determine transcription level and regulatory mode of a certain gene. Currently, several high-intensity inducible and constitutive promoters have been identified and applied for protein expression in *K. phaffii* (Table 1).

Inducible promoters decouple cell growth from product synthesis, allowing for precise regulation of recombinant protein expression (Wang et al. 2024d). One of the earliest and most widely used promoters is the alcohol oxidase promoter P_{AOX1}, which is methanol-inducible and exhibits strict regulation mode (Wang et al. 2017). There are several other methanol-inducible promoters with varying strengths, such as P_{DAS1}, P_{DAS2}, and P_{FLDH1}, which provide valuable tools for methanol-induced expression systems and metabolic engineering. However, the use of methanol poses

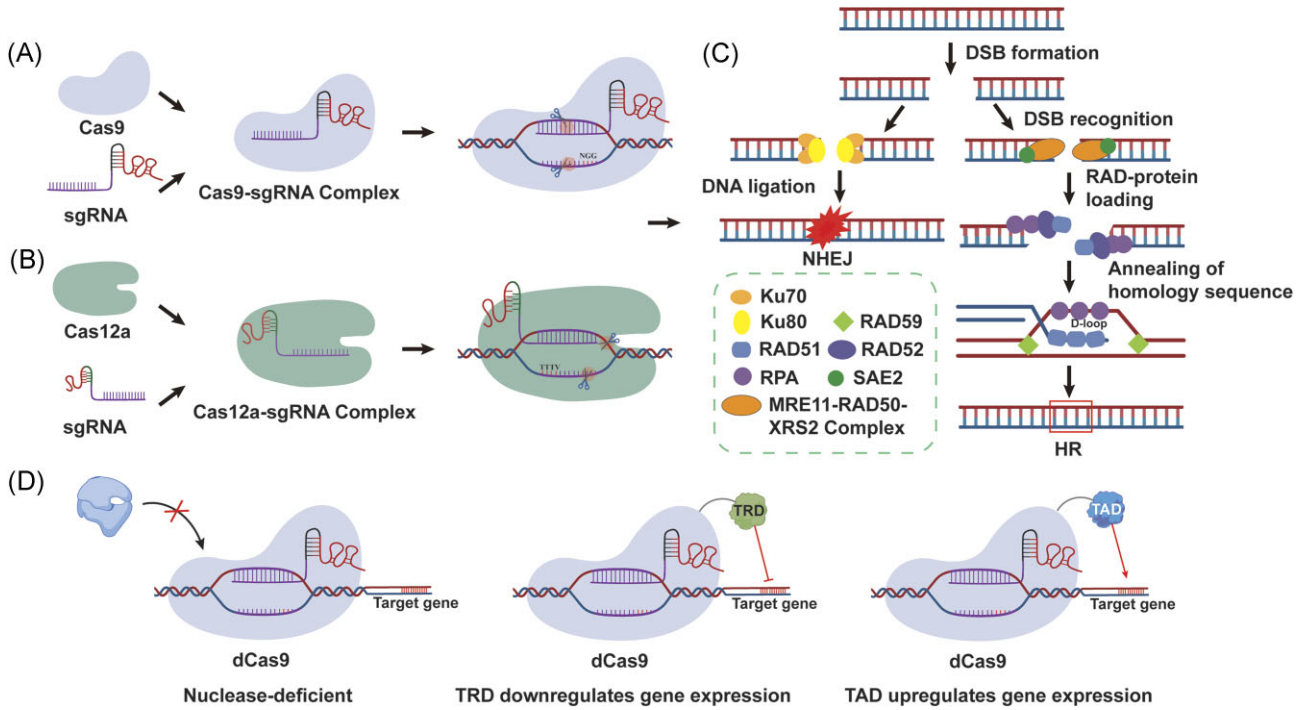

Figure 1. Genome editing in *K. phaffii*. (A) CRISPR/Cas9-mediated gene editing. (B) CRISPR/Cas12a-mediated gene editing. (C) DNA double-strand break repair mechanisms produced by CRISPR systems. (D) CRISPRi (CRISPRa) inhibits (activates) target gene expression and translation.

Table 1. Promoters used to express recombinant protein in *K. phaffii*.

Promoter	Production	Yield	Reference
P_{AOX1}	human insulin precursor	1.05 g/l	(Nurdiani et al. 2021)
P_{AOX1}	lacto-protein	5.4 g/l	(Lv et al. 2023)
P_{AOX1}	human interleukin-3	2.23 g/l	(Dagar and Khasa. 2018)
P_{AOX1}	glucose oxidase	30.7 g/l	(Zhou et al. 2024)
P_{AOX1}	phytase	9.58 g/l	(Li et al. 2015)
P_{AOX1}	human collagen II	3.04 g/l	(Wang et al. 2024b)
P_{AOX1}	soy hemoglobin	3.5 g/l	(Shao et al. 2022)
P_{AOX1}	patatin	1.4 g/l	(Dai et al. 2022)
P_{AOX1}	rHSA	17.47 g/l	(Zhu et al. 2021)
P_{AOX1}	A33scFv	8 g/l	(Das et al. 2024)
P_{GAP}	SAM synthetase 2	2.49 g/l	(Zhang et al. 2009)
P_{GTH1}	i-body	4.2 g/l	(Prielhofer et al. 2018)
P_{GCW14}	xylanase	8.7 g/l	(Xia et al. 2021)
P_{FLD1}	lipase LIP2	35 000 U/ml	(Wang et al. 2012)
P_{DAS1}	lipases	22 342 U/ml	(Wang et al. 2019b)
P_{ADH3}	xylanase	3725 U/ml	(Karaoglan et al. 2016)
P_{TEF1}	human serum albumin	63 556 U/ml	(Huang et al. 2020)
P_{MOX}	endoglucanase 3	711 U/ml	(Mombeni et al. 2020)

safety risks, limiting its application in pharmaceutical and food area. As a result, researchers continue to explore alternative inducible promoters, such as rhamnose-induced P_{LRA3} and P_{LRA4}, ethanol-responsive $P_{ICL1}$and P_{ADH2}, thiamine starvation-induced promoter P_{THI11}, and so on. These alternative promoters provide effective solutions for regulating protein and metabolite production, thereby expanding the versatility of *K. phaffii*. Jiao et al. (2020) developed rhamnose-induced strains and revealed their survival mechanism without carbon sources and the pathway to enhance protein expression. Flores-Villegas et al. (2023) screened a new P_{GTH1} promoter with stronger expression intensity and similar regulatory characteristics, increasing scFv yield by 55%.

Compared with inducible promoters, constitutive promoters typically do not require inducers, which simplifies the culture process and offers broad compatibility under various conditions. This makes constitutive promoters highly suitable for studying physiological traits, component assembly, and large-scale fermentation (Nieto-Taype et al. 2020). For non-toxic proteins, the robust

nature of the constitutive system helps maximize yield per unit time, greatly enhancing the optimization potential of *K. phaffii* cell factories. The P_{GAP} promoter is a commonly used constitutive promoter in *K. phaffii* in academic research and industrial production due to its ease of use and stable expression under various conditions (García-Ortega et al. 2019). Lin et al. (2024) enhanced the activity of P_{GAP}, resulting in a 2.43-fold increase in cellulase expression. Besides, several other constitutive promoters with varying expression characteristics and strengths have been reported in *K. phaffii*, such as P_{TEF1}, P_{GCW14}, and P_{PGK1}.

Although natural inducible and constitutive promoters have been widely used, they may not always meet the evolving requirements of *K. phaffii* for more complex production. Therefore, promoter engineering has emerged as an effective approach to precisely regulate metabolism at the transcriptional level, offering various possibilities to meet different production demands. Eukaryotic promoters consist of a core promoter and an upstream regulatory region. The core promoter binds to RNA polymerase to determine the transcription initiation site and strength, while the upstream regulatory region contains transcription factor binding sites that regulating promoter strength (Gibbons et al. 2022). Promoter engineering typically focuses on modifying these elements. Vogl and Glieder (2013) firstly synthesized and modified the core promoter of P_{AOX1}, creating variants with distinct expression levels. Wang et al. (2018b) created a hybrid promoter by modifying the upstream activation sequence and core element, achieving transcription activity twice that of a common yeast constitutive promoter P_{TEF1}. Lai et al. (2024) constructed a gradient strength promoter library in *K. phaffii* through random hybridization and high-throughput screening. With GFP as a reporter, the expression level of P_{hy47} was 2.93 times higher than that of the P_{GAP} promoter.

Additionally, the terminator also plays a crucial role in regulating gene expression levels by influencing mRNA stability and affecting the subsequent translation process. Protein expression levels can be finely regulated by adjusting appropriate terminator sequences (Zhang et al. 2024a). To date, only 20 terminators have been identified in *K. phaffii*, with activity ranging from 57% (*CAT1*t) to 100% (*AOX1*t), indicating a limited ability to regulate pathway flux. To explore terminator sequences that can regulate protein expression levels in *K. phaffii*, a terminator library of 72 sequences was constructed, exhibiting adjustable expression intensity range of 17-fold (Ito et al. 2020). These advances in promoter and terminator engineering enable the precise regulation and expression improvement of heterologous proteins in *K. phaffii*.

Gene dosage regulation

Increasing gene copies is an effective strategy to boost recombinant protein production (Zheng et al. 2014). Antibiotic-based multicopy screening methods are widely used for constructing high-copy strains (Wang et al. 2023c), e.g. geneticin, zeocin, and hygromycin. However, high concentrations of antibiotics may reduce the number of positive transformants, thereby increasing the difficulties of the screening process (Aw and Polizzi 2016). To solve this problem, the post-transformation vector amplification (PTVA) approach was developed, screening low-copy strains by gradually increasing the concentration of antibiotics, allowing the strains to undergo multiple recombination events that ultimately resulting in high-copy transformants (Sunga et al. 2008). Although effective, this method faces challenges such as long duration, high costs and the potential for gene mutations. To optimize this process, the liquid PTVA method was proposed to avoid the time and expense (Aw and Polizzi 2016). Another significant advancement involves CRISPR-mediated rDNA integration technology. Jiang et al. (2024b) designed a glycerol-induced and glucose-inhibited promoter system (CRISPOi) and also used a strong promoter (CRISPOc) to achieve high-copy expression of the target gene through one-step integration, resulting in a 19.5-fold increase in the yield of geraniol.

For all that, increasing gene dosage does not always result in a linear increase in expression levels. For example, L-asparaginase yield reached the highest level with only three-copy genes in *K. phaffii*, and more copies could not improve production further (Erden-Karaoğlan and Karaoğlan 2023). It might be ascribed to that increased gene dosages cause oxidative stress related to protein folding, as well as a shortage of carbon and energy supply, which in turn affects the transcription and translation processes (Mao et al. 2024). Therefore, balancing gene dosage with the physiological burden on the cells remains a topic that requires further exploration.

Translation regulation strategies

The transcription level of a gene does not always align with the final expression level of the protein. The translation from mRNA to protein involves multiple variables. Translation factors and their regulation play a crucial role in the quality of protein synthesis, folding, modification, and secretion (Vogel and Marcotte 2012). As a result, optimizing translation regulation has proved to be an effective strategy to enhance protein expression. In *K. phaffii*, some translation-related factors have been studied. Lin et al. (2013) found that the upregulation of certain translation factors was positively correlated with the increased expression of xylanase by the Isobaric Tag for Relative and Absolute Quantitation- Liquid Chromatography-Tandem Mass Spectrometry (iTRAQ–LC–MS/MS method). The translation process can be divided into three stages: initiation, elongation, and termination. The initiation of translation, which is regulated by multiple subunits, is the rate-limiting step (Zhang et al. 2022). By strengthening the key factors involved in translation initiation, Staudacher et al. (2022) highly increased the yield of recombinant nanobody protein and enhanced overall protein translation intensity. Ribosomes are the core of mRNA translation, Liao et al. (2020) demonstrated that overexpressing these factors not only promoted cell growth but also increased the yield of reporter protein. Recent studies have also focused on the overexpression of factors involved in the closed-loop structure of translation, which remarkably enhanced the stability and efficiency of translation initiation. By overexpressing these factors, either individually or in combination, the fermentation yield of exogenous proteins can be increased by up to three times (Staudacher et al. 2022). Codon optimization is also a commonly used strategy to improve translation efficiency. The effect of codon preference on gene expression is largely due to its influence on translation, as codons impact translation efficiency through elongation and accuracy (Wu et al. 2024). Li et al. (2021) achieved a 62.2-fold increase in acetylcholinesterase enzyme activity by employing codon substitution in *K. phaffii*. However, achieving a balance between translation rate and mRNA stability remains a challenge. Excessively rapid translation or overly stable mRNA can increase the metabolic burden on cells, thereby affecting protein expression efficiency and cell growth (Dave et al. 2023). Therefore, it is essential to establish a precise relationship between mRNA sequence and protein synthesis yield, including both protein synthesis rate and mRNA half-life (Bicknell et al. 2024).

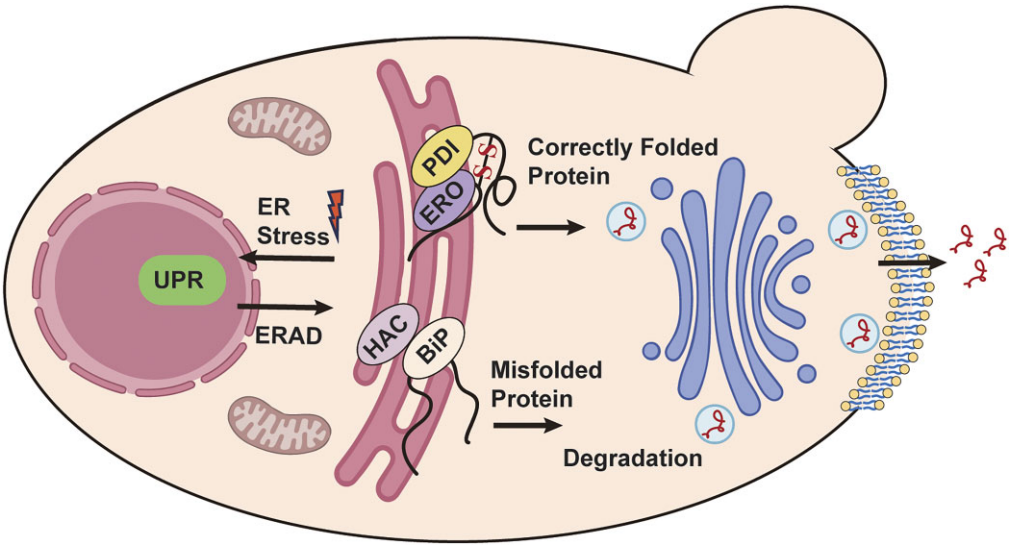

Figure 2. Schematic diagram of improving recombinant protein expression in *K. phaffii* at the folding level.

Optimization of protein folding

Folding bottlenecks and regulatory strategies

In *K. phaffii*, the secretion of recombinant proteins relies on a complex folding process within the endoplasmic reticulum (ER), which serves as the central hub for protein folding and post-translational modification (Fig. 2). Secretory proteins enter the ER in an unfolded state and undergo folding with the assistance of molecular chaperones, ensuring their proper function and stability (Idiris et al. 2010). However, when recombinant protein expression levels are excessively high, unfolded or misfolded proteins can accumulate in the ER, overwhelming its folding capacity and triggering the unfolded protein response (UPR). The UPR mitigates ER stress by reducing translation rates, upregulating molecular chaperone expression, and expanding ER capacity to restore homeostasis. If the misfolded protein burden exceeds the ER's processing capabilities, it may lead to apoptosis that compromising protein expression efficiency (Hetz et al. 2020).

In addition to UPR, ER relies on the ER-associated degradation (ERAD) pathway to remove misfolded proteins. Through ERAD, misfolded proteins are transported from the ER to the cytoplasm, where they undergo ubiquitination and subsequent degradation by the proteasome (Krshnan et al. 2022). The synergy of UPR and ERAD can reduce ER stress, but excessive activation may hinder protein synthesis and lower recombinant protein yield (Lin et al. 2023). Balancing the activities of UPR and ERAD has therefore emerged as a critical challenge in optimizing protein folding and enhancing production efficiency.

To enhance the folding and secretion efficiency of recombinant proteins in *K. phaffii*, researchers have focused on overexpression of endogenous or exogenous molecular chaperones to promote proper protein folding. Key molecular chaperones, such as heat shock proteins (Qu et al. 2024) and protein disulfide isomerase (Pdi) (Fu et al. 2020), play critical roles in this process. Kar2 facilitates correct folding by binding to unfolded polypeptide chains, thereby preventing aggregation between molecules. For instance, Sun et al. (2017) demonstrated that overexpression of Kar2 in recombinant *K. phaffii* increased β-galactosidase secretion by 57.5%. Similarly, Pdi improved the secretion of disulfide bond-containing proteins by forming and rearranging disulfide bonds to stabilize their structure. Inan et al. (2006) reported that overexpression of Pdi significantly improved the secretion levels of the recombinant protein NA-ASPI. Additionally, a variety of molecular chaperones have been identified and characterized in *K. phaffii* (Table 2) to support recombinant protein expression.

While co-expression of molecular chaperones is effective in enhancing protein secretion efficiency, the varying folding requirements of different proteins and associated metabolic burden on host cells pose challenges. Developing standardized and universally applicable optimization strategies remains for further exploration to achieve high-efficiency protein expression and secretion but minimizing cellular stress.

Engineering of protein glycosylation

Compared to prokaryotic systems, *K. phaffii* showed obvious advantages in protein post-translational modifications. Glycosylation stands out as a crucial modification that profoundly influences protein secretion efficiency and overall quality of proteins. The *K. phaffii* host possesses the capability for both *N*- and *O*-glycosylation (Thak et al. 2018). Furthermore, compared to *S. cerevisiae*, the glycosylation pattern of *K. phaffii* was more like mammalian cells, with a lower degree (only 9~14 mannose residue) of high-mannose glycosylation (Sun et al. 2022). This characteristic provides a crucial foundation for the application of this host in the production of therapeutic glycoproteins. The Nb11-59 nanobody was successfully produced at a large scale by *K. phaffii*, achieving a yield of 20 g/l (Liu et al. 2022c).

However, the glycosylation pathway of *K. phaffii* differs from that of human, potentially leading to the introduction of non-target glycosylation structures that can affect the activity, function, and immunogenicity of recombinant glycoproteins (Swiech et al. 2015). Glycosylation engineering has been performed to modify the glycosylation pathway in *K. phaffii*, aiming to achieve a humanized glycosylation profile. By knocking out endogenous high-mannose-related gene *OCH1* and introducing mammalian-derived glycosylation enzymes like MnsI and GnTI, Choi et al. (2003) successfully developed strains capable of producing hybrid glycan structures. Complex glycoproteins, such as recombinant erythropoietin (rEPO) (Hamilton et al. 2006) and monoclonal antibodies (Liu et al. 2018), were also successfully produced in *K. phaffii*.

Table 2. Molecular chaperones used for recombinant protein expression in *K. phaffii*.

Chaperones	Target protein	Final-yield	Improvement	Reference
BiP	A33scFv	4 g/l	3-fold	(Damasceno et al. 2007)
Kar2	HFBI	242 mg/l	22-fold	(Sallada et al. 2019)
Ero1	HFBI	330 mg/l	30-fold	(Sallada et al. 2019)
Pdi	thermophilic lipase	57 521 U/ml	1.46-fold	(Wang et al. 2019a)
Ero1-Pdi	human lysozyme	81 600 U/ml	1.38-fold	(Wang et al. 2023c)
Hac1	nanobodies	2.13 g/l	1.5-fold	(Zheng et al. 2024)
Cpr5	carboxylesterase	161.34 U/ml	4.57-fold	(Jiang et al. 2023)
Sec1	IL2-HSA	1.25 g/l	2.5-fold	(Guan et al. 2016)
Sly1	IL2-HSA	0.95 g/l	1.9-fold	(Guan et al. 2016)
Aha1	β-glucuronidase	0.97 g/l	1.59-fold	(Huangfu et al. 2016)
Ypt6	β-glucuronidase	0.95 g/l	1.56-fold	(Huangfu et al. 2016)
Ssa4	*Rhizopus oryzae* lipase	4470 U/ml	1.45-fold	(Jiao et al. 2018)
Ssa1-Sis1	porcine growth hormone	340 mg/l	/	(Deng et al. 2020)
Ssa1-Ydj1-BiP-Lhs1	E2-Spy antigen protein	168.3 mg/l	6.18-fold	(Li et al. 2024)
Ssa1	*Candida antarctica* lipase B	260 U/ml	1.4-fold	(Samuel et al. 2013)
GroES-GroEL	phytase	495.5 U/mg	2.3-fold	(Summpunn et al. 2018)

Additionally, advancements in glycosylation engineering from other yeast platforms offer valuable insights for the future development of glycosylation engineering in *K. phaffii*. For example, Makrydaki et al. (2024) developed the SUGAR-TARGET platform, which enables sequential and customizable N-glycosylation in vitro using purified recombinant glycosyltransferases. While this system is not directly related to *K. phaffii*, it provides innovative perspectives for *in vitro* glycosylation that could inspire future developments in *K. phaffii* glycoengineering.

Optimization of protein secretion pathways

In *K. phaffii*, the secretion of recombinant proteins primarily relies on the classical secretion pathway. Its initial and critical step is the translocation of proteins across the ER membrane into the ER. The efficiency of this translocation highly influences the overall secretion levels of recombinant proteins, with the signal peptide playing a pivotal role in guiding this process. The sequence of the signal peptide directly determines both the efficiency of protein translocation into the ER and the subsequent secretion levels (Yang et al. 2023). Till now, numerous signal peptides (Table 3) have been developed in *K. phaffii* to meet the high-efficiency secretion demands of recombinant protein production.

Currently, the α-mating factor signal peptide (α-MF) from *S. cerevisiae* is one of the most commonly used signal peptides. It has been successfully integrated into commercial expression vectors such as pPIC 9 K and pPIC ZαA (Wang et al. 2019b). The α-MF signal peptide demonstrates high efficiency in the secretion of small peptides and proteins; however, its effectiveness is limited when applied to the secretion of larger macromolecular proteins (Aw et al. 2017). To enhance the secretion efficiency of α-MF, Chahal et al. (2017) optimized its hydrophobic region and codon usage, significantly improving its secretion ability of specific exogenous proteins.

Advancements in proteomics have revealed that endogenous signal peptides in *K. phaffii* possessed a strong capacity to guide protein secretion (Shen et al. 2022). These endogenous peptides are often more readily recognized by the host cells, thereby enabling more efficient secretion of foreign proteins (Liang et al. 2013). However, the choice of an optimal signal peptide largely depends on the specific properties of the target protein. As such, experimental validation remains an essential step in optimizing signal peptides for enhanced secretion efficiency.

Beyond optimizing the signal peptide itself, enhancing the protein translocation capability of host cells represents an alternative strategy to improve secretion efficiency (Dong et al. 2024). For instance, Marsalek et al. (2019) demonstrated that overexpression of ER translocation pore subunit Sbh1 in *K. phaffii* led to a more than 4-fold increase in the secretion of an antibody fragment HyHEL-Fab. Zhang et al. (2024b) improved the co-translational translocation capacity of *K. phaffii* by overexpressing components of the signal recognition particle, resulting in a 48% increase in hyaluronidase secretion. Once the protein is processed in the ER, downstream transport processes may become bottlenecks that limiting secretion efficiency (Kaneyoshi et al. 2019). This limitation can be effectively mitigated by overexpressing transport-related factors. For example, Wang et al. (2024c) enhanced the secretion of human lactoferrin by 38.5% by maintaining a balanced bidirectional vesicle transport system. Liu et al. (2024) improved xylanase secretion by 9.6% with co-expression of vesicle transport-related factors.

Omics approaches and metabolic modeling in *K. phaffii* engineering

Omics technologies provide powerful tools for the comprehensive elucidation of cellular functions. Techniques such as genomics, transcriptomics, proteomics, and metabolomics facilitate the intricate mapping of interactions between genes, proteins, and metabolites. These technologies are particularly transformative in biotechnology and metabolic engineering, where they enable precise optimization of metabolic pathways. For microbial factories like *K. phaffii*, omics technologies are indispensable for deciphering metabolic profiles, enhancing production efficiency, and optimizing fermentation processes.

Genomics, as one of the fundamental omics technologies, provides crucial support for strain improvement of *K. phaffii*. Through comparative genomic analysis, researchers can systematically identify genetic differences among various strains, thereby providing a basis for establishing a scientific foundation for selecting a dominant strain for specific applications. Brady et al. (2020) found that cell wall integrity is a key factor that influencing strain performance. By combining genomics and transcriptomics, optimized *K. phaffii* mutants were developed, which significantly improved transformation efficiency and protein expression levels.

Table 3. Signal peptides used for recombinant protein secretion in *K. phaffii*.

Signal peptide	Origin	Target protein	Yield	Reference
α-MF	*S. cerevisiae*	insulin precursor	1.5 g/l	(Gurramkonda et al. 2010)
MF4I	*S. cerevisiae*	Phytase	6.1 g/l	(Xiong et al. 2005)
αpre	*S. cerevisiae*	endo-1,4-β-mannosidase	/	(Wang et al. 2015)
SUC2	*S. cerevisiae*	dextranase	1706 U/l	(Martínez et al. 2021)
Killer protein	*S. cerevisiae*	mouse α-amylase	240 mg/l	(Kato et al. 2001)
OST1	*S. cerevisiae*	type III collagen	/	(Wang et al. 2023b)
SCW	*K. phaffii*	*Candida antarctica* lipase B	/	(Liang et al. 2013)
DSE	*K. phaffii*	*Candida antarctica* lipase B	/	(Liang et al. 2013)
EXG	*K. phaffii*	*Candida antarctica* lipase B	/	(Liang et al. 2013)
PHO1	*K. phaffii*	*Aspergillus niger* xylanase	139 U/ml	(Karaoglan et al. 2014)
PIR1	*K. phaffii*	*Aspergillus niger* xylanase	250 U/ml	(Karaoglan et al. 2014)
GAS1	*K. phaffii*	β-galactosidase	191.8 U/g	(Duan et al. 2019)
SP23	*K. phaffii*	hyaluronidase	138 000 U/ml	(Huang et al. 2020)
nsB-AP	*Candida antarctica*	hyaluronidase	1 680 000 U/ml	(Kang et al. 2016)
INU	*Kluvveromyces marxious*	phytase	/	(Helian et al. 2020)
HFBI	*Trichoderma reesei*	β-galactosidase	/	(Cao et al. 2017)
α-amylase	*Aspergillus niger*	brazzcin	283 mg/l	(Neiers et al. 2021)
Serum albumin	*Homo sapiens*	human lysozyme	/	(Xiong et al. 2008)
IgGl	Murine	anti-HIV VRC0l antibody	6.5 mg/l	(Aw et al. 2018)

Transcriptomics plays a crucial role in studying gene expression and regulatory mechanisms in *K. phaffii*. Through transcriptomic analysis, the expression changes of key genes under specific conditions can be revealed, providing insights into their regulatory roles in metabolic pathways. Gao et al. (2023b) identified the key role of the *PAS_0305* gene in cell wall thickness and stress responses by transcriptomic analysis to. Through activating cell wall sensors and increasing cell wall permeability, they significantly improved methanol conversion efficiency and enhanced the production of single-cell protein, effectively addressing the issues of carbon loss and metabolic efficiency in methanol fed-batch fermentation of *K. phaffii*. Gupta and Rangarajan (2022) used identified a series of interactions among transcription factors that functioning in the transcriptional regulation of genes in the methanol utilization pathway of *K. phaffii* by transcriptomic analysis, providing new insights for optimizing this pathway. Boojari et al. (2023) integrated dynamic flux balance analysis (DFBA) with transcriptomic data to simulate the growth process of recombinant protein expression in *K. phaffii* under different feeding strategies. They further developed a novel feeding strategy that increased protein concentration by 16% and improved production yield by 85% compared to the previously optimized culture conditions.

The application of proteomics complements the findings from genomics and transcriptomics. By systematically analyzing the protein expression profile within cells, researchers can achieve deeper insights into intracellular metabolic and secretion processes, thereby optimizing recombinant protein expression. Hou et al. (2022) revealed the upregulation of the transaldolase isoenzyme (*Tal2*) in methanol medium through proteomic analysis, highlighting its important role in methanol metabolism. This study provided valuable information for understanding methanol metabolism in *K. phaffii* and facilitated optimization of metabolic pathways for the increase of protein production efficiency.

Metabolomics and GEMs represent powerful tools for systematic optimization of metabolic pathways in *K. phaffii*. By integrating genomic data and metabolic flux analysis, these models can predict dynamic changes in intracellular metabolic reactions, thereby guiding the design and implementation of metabolic engineering strategies. Nocon et al. (2014) integrated recombinant protein production into the metabolic model of *K. phaffii*, simulating the impact of overproduction on cellular metabolism and predicting gene targets to enhance productivity. The results showed that overexpression of targets related to the glycolytic pathway or knockout of specific metabolic pathways significantly increased the production efficiency of human superoxide dismutase. Ye et al. (2017) upgraded the existing GEM of *K. phaffii* and the new model iRY1243, resulting in increased number of reactions, metabolites, and unique ORFs, significantly expanding the metabolic coverage. It improved the accuracy of the model in predicting cell growth and metabolic flux distribution.

Screening for high-expression strains

Application of high-throughput characterization technologies

Microbial cell factories for protein production are typically constructed by optimizing expression elements and modifying chassis strains. However, due to the complexity of biological systems and unclarified mechanisms, current modifications still rely on a "trial and error" approach, leading to long development cycles, high costs, and low production efficiency. Therefore, the rapid and efficient mining of candidate genes in the genome is crucial for improving the design efficiency of microbial cell factories. The genotype-to-phenotype correlation map serves as a crucial link between genomic information and efficient protein expression (Costanzo et al. 2019). To construct this map, a high-throughput characterization and screening model for secretory protein production must be firstly established. A common approach is to convert protein yield into a detectable fluorescence signal, integrate a droplet microfluidics (DMFS) system to sort highly fluorescent droplets, then refold the strains and analyze their genotype (Lebrun et al. 2023). For example, Chen et al. (2021) fused eGFP with ER transmembrane protein Sec63, and dynamically monitored the effect of overexpression of this fusion protein on ER morphology (Fig. 3A). Navone et al. (2023) established a split GFP biosensor to detect the expression and folding of recombinant proteins by dividing GFP into two fragments (GFP1-10 and GFP11). The fluorescence intensity is linearly correlated with recombinant protein ex-

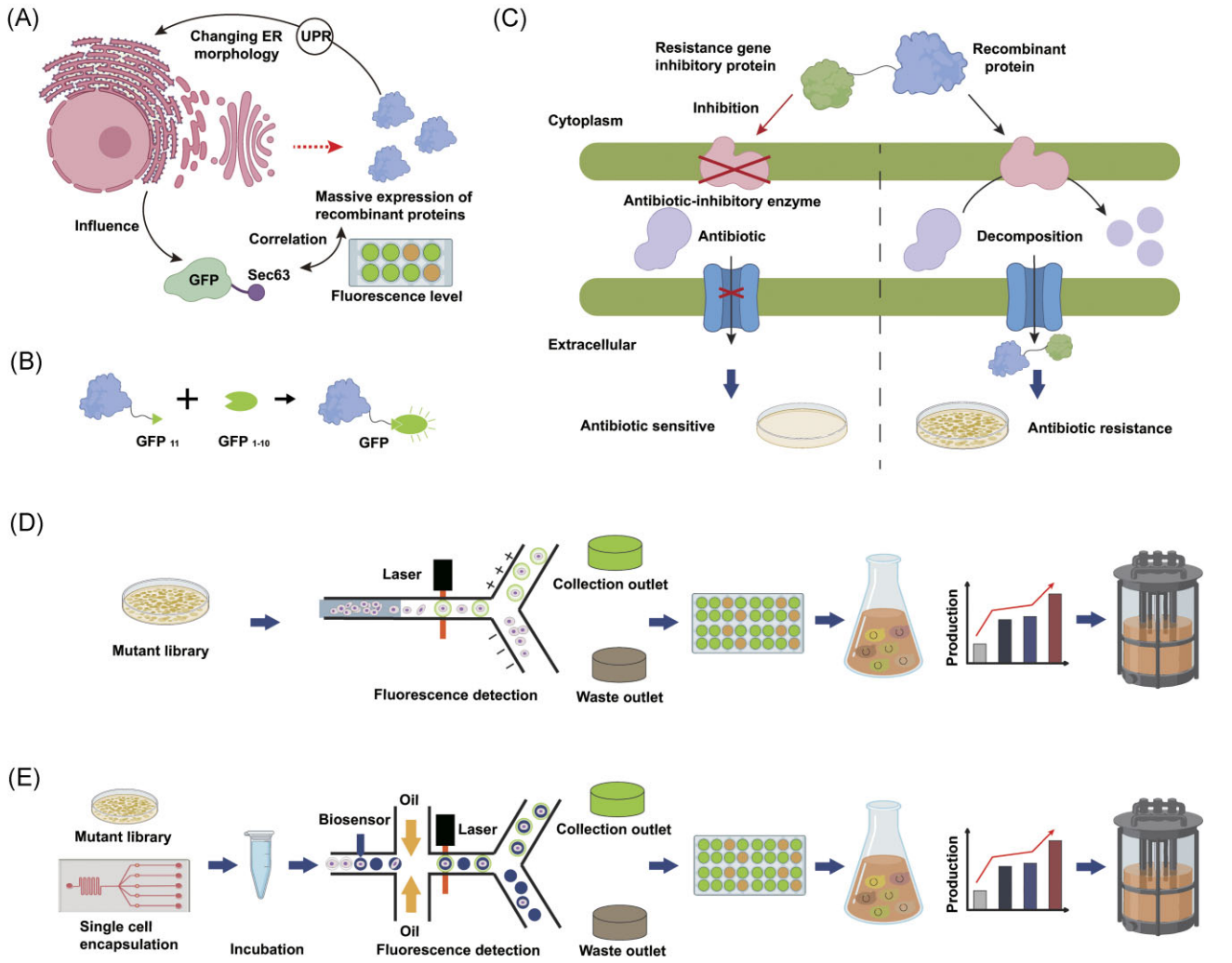

Figure 3. High-throughput characterization and screening strategy for high-yield recombinant protein strains. (A) Correlation between eGFP-Sec63 signal and recombinant protein expression. (B) Split-GFP biosensor. (C) Construction of a growth-lethality model. (D) Fluorescence activated cell sorting technology. (E) Droplet microfluidics screening technology.

pression. If the target protein is insoluble or misfolded, the fragments fail to bind, resulting in the absence of the fluorescence signal, thus enabling dynamic monitoring of protein expression and solubility (Fig. 3B). In addition, recombinant protein production can be coupled with growth to establish a growth difference or lethal screening model (Tuo et al. 2023). This approach allows for the elimination of low-yield strains, amplification of high-yield strains, and tracking changes in genotype abundance during the cell culture process. In *E. coli*, a target protein is fused with β-lactamase (BLa) inhibitor protein, while BLa is expressed in the periplasmic space. When the fusion protein is efficiently secreted, BLa activity is not inhibited, giving the strain a growth advantage under β-lactam antibiotics (Natarajan et al. 2017) (Fig. 3C). Similar strategies still remain to be implemented *K. phaffii*.

Application of high-throughput screening technologies

Recently, high-throughput methods such as microplate screening (Schmitz et al. 2024), fluorescence activated cell sorting (FACS), and DMFS have improved the efficiency of expression elements and strain modification (Fig. 3D and E). The porous plate, combined with an automated liquid handling platform, can process up to 10^4 samples per day (Czerniecki et al. 2018). FACS, utilizing optical analysis and sorting technology, can process up to 10^7 cells per hour, making it a key tool for cell screening (Kjeldsen et al. 2022). DMFS, enable efficient droplet sorting for internal and external cell expression and secretion processes, thereby advancing research in medicine, food, and environmental protection (Yuan et al. 2022). However, as the scale of screening expands and the variety of screening targets increases, improvement of the screening precision and accuracy and reduction of false positives and negatives still remain a major challenge.

Conclusions and future perspectives

As a versatile platform for recombinant protein production, *K. phaffii* has demonstrated success across biopharmaceuticals, industrial enzymes, food additives, biomaterials, etc. In recent years, researchers have introduced various optimization strategies targeting transcription, translation, protein folding, and secretion—key processes in recombinant protein production (Sun et al. 2022). However, the inherent complexity of cellular metabolism and regulatory networks, coupled with an incomplete understanding of their mechanisms, continues to hinder the development of highly efficient cell factories. Moreover, the specific characteristics of tar-

get proteins heavily influence the selection of expression strategies (Schütz et al. 2023), making it difficult for singular optimizations to address diverse production requirements.

With the advancement of omics technologies, researchers have gained deeper insights into the physiological processes and genetic framework of *K. phaffii*. GEMs have proved to be valuable tools for the systematic analysis of *K. phaffii* metabolic networks, successfully guiding the efficient expression of specific recombinant proteins (Ye et al. 2017). However, these models still face challenges in fully capturing the complexity of metabolic networks and exhibit limited predictive accuracy (Fang et al. 2020). The emergence of synthetic biology offers a promising solution to these limitations. By employing genome-scale engineering modifications, it becomes possible to comprehensively intervene in metabolic networks and precisely identify non-intuitive engineering targets, thereby enhancing protein expression levels. Moreover, the introduction and advancement of CRISPR-based genome-scale engineering technologies promise to revolutionize the design of efficient yeast cell factories.

The potential of *K. phaffii* in bio-utilization of C1 and C2 substrates, i.e. methanol, formate, ethanol, and acetate, also warrants further exploration. The C1 and C2 substrates, as inexpensive and readily available raw materials, serve as promising carbon sources for production of recombinant proteins production and bulk chemicals. However, the current utilization efficiency of C1 and C2 substrates in *K. phaffii* is relatively low, with much metabolic flux directed toward the dissimilation pathways and low tolerance of high concentration of substrates (Guo et al. 2022, Meng et al. 2023). Future advancements in metabolic engineering aim to redirect C1 and C2 substrate metabolism toward the assimilation pathway, unlocking their full potential as raw materials for bioproducts manufacturing.

Overall, these innovations, combined with the continued integration of synthetic biology and systems biology tools, are poised to unlock the full potential of *K. phaffii* as a next-generation cell factory, enabling cost-effective and scalable solutions across diverse industries.

Conflict of interest: None declared.

Funding

This work was financially supported by the National Key Research and Development Program of China (2022YFC2805102) and Shanghai Explorer Research Plan (24TS1411600).

Reference

Ahmad M, Hirz M, Pichler H *et al.* Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* 2014;**98**:5301–17.

Antoniewicz MR. Synthetic methylotrophy: strategies to assimilate methanol for growth and chemicals production. *Curr Opin Biotechnol* 2019;**59**:165–74.

Aw R, McKay PF, Shattock RJ *et al.* A systematic analysis of the expression of the anti-HIV VRC01 antibody in *Pichia pastoris* through signal peptide optimization. *Protein Expression Purif* 2018;**149**:43–50.

Aw R, McKay PF, Shattock RJ *et al.* Expressing anti-HIV VRC01 antibody using the murine IgG1 secretion signal in *Pichia pastoris*. *AMB Expr* 2017;**7**:70.

Aw R, Polizzi KM. Liquid PTVA: a faster and cheaper alternative for generating multi-copy clones in *Pichia pastoris*. *Microb Cell Fact* 2016;**15**:29.

Azadi S, Mahboubi A, Naghdi N *et al.* Evaluation of sorbitol-methanol Co-feeding strategy on production of recombinant human growth hormone in *pichia pastoris*. 2017;**16**:1555–64.

Baghban R, Farajnia S, Ghasemi Y *et al.* New developments in *pichia pastoris* expression system, review and update. *Curr Pharm Biotechnol* 2018;**19**:451–67.

Barone GD, Emmerstorfer-Augustin A, Biundo A *et al.* Industrial production of Proteins with *Pichia pastoris—Komagataella phaffii*. *Biomolecules* 2023;**13**:441.

Bicknell AA, Reid DW, Licata MC *et al.* Attenuating ribosome load improves protein output from mRNA by limiting translation-dependent mRNA decay. *Cell Rep* 2024;**43**:114098.

Boojari MA, Rajabi Ghaledari F, Motamedian E *et al.* Developing a metabolic model-based fed-batch feeding strategy for *pichia pastoris* fermentation through fine-tuning of the methanol utilization pathway. *Microb Biotechnol* 2023;**16**:1344–59.

Brady JR, Whittaker CA, Tan MC *et al.* Comparative genome-scale analysis of *pichia pastoris* variants informs selection of an optimal base strain. *Biotechnol Bioeng* 2020;**117**:543–55.

Cai P, Duan X, Wu X *et al.* Recombination machinery engineering facilitates metabolic engineering of the industrial yeast *pichia pastoris*. *Nucleic Acids Res* 2021;**49**:7791–805.

Cai P, Wu X, Deng J *et al.* Methanol biotransformation toward high-level production of fatty acid derivatives by engineering the industrial yeast *Pichia pastoris*. *Proc Natl Acad Sci USA* 2022;**119**:e2201711119.

Cao L, Ren G, Qin Z *et al.* Improving the secretion yield of the β-galactosidase Bgal1-3 in *Pichia pastoris* for use as a potential catalyst in the production of prebiotic-enriched milk. *J Agric Food Chem* 2017;**65**:10757–66.

Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* 2000;**24**:45–66.

Chahal S, Wei P, Moua P *et al.* Structural characterization of the α-mating factor prepro-peptide for secretion of recombinant proteins in *Pichia pastoris*. *Gene* 2017;**598**:50–62.

Chen Q, Wang B, Pan L. Efficient expression of γ-glutamyl transpeptidase in *Bacillus subtilis* via CRISPR/Cas9n and its immobilization. *Appl Microbiol Biotechnol* 2024;**108**:149.

Chen Y, Yuan Q, Li C *et al.* Development and evaluation of a novel method for rapid screening of *Pichia pastoris* strains capable of efficiently expressing recombinant proteins. *Chin J Biotechnol* 2021;**37**:939–49.

Choi B-K, Bobrowicz P, Davidson RC *et al.* Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Proc Natl Acad Sci USA* 2003;**100**:5022–7.

Costanzo M, Kuzmin E, Van Leeuwen J *et al.* Global genetic networks and the genotype-to-phenotype relationship. *Cell* 2019;**177**:85–100.

Czerniecki SM, Cruz NM, Harder JL *et al.* High-throughput screening enhances kidney organoid differentiation from Human pluripotent stem cells and enables automated multidimensional phenotyping. *Cell Stem Cell* 2018;**22**:929–40. e4.

Dagar VK, Khasa YP. 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* 2018;**108**:999–1009.

Dai Z, Wu X, Zeng W *et al.* Characterization of highly gelatinous patatin storage protein from *pichia pastoris*. *Food Res Int* 2022;**162**:111925.

Damasceno LM, Anderson KA, Ritter G *et al.* Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in *Pichia pastoris*. *Appl Microbiol Biotechnol* 2007;**74**:381–9.

Das PK, Sahoo A, Veeranki VD. Recombinant monoclonal antibody production in yeasts: challenges and considerations. *Int J Biol Macromol* 2024;**266**:131379.

Dave P, Roth G, Griesbach E *et al.* Single-molecule imaging reveals translation-dependent destabilization of mRNAs. *Mol Cell* 2023;**83**:589–606.

De Schutter K, Lin Y-C, Tiels P *et al.* Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol* 2009;**27**:561–6.

Deng B, Yue Y, Yang J *et al.* Improving the activity and thermostability of PETase from Ideonella sakaiensis through modulating its post-translational glycan modification. *Commun Biol* 2023;**6**:39.

Deng J, Li J, Ma M *et al.* Co-expressing GroEL–GroES, Ssa1–Sis1 and bip–PDI chaperones for enhanced intracellular production and partial-wall breaking improved stability of porcine growth hormone. *Microb Cell Fact* 2020;**19**:35.

Dong X, Lin Y, Zhang J *et al.* Modification of the endoplasmic reticulum to enhance ovalbumin secretion in *saccharomyces cerevisiae*. *J Agric Food Chem* 2024;**72**:19985–93.

Drocourt D, Calmels T, Reynes J-P *et al.* Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucl Acids Res* 1990;**18**:4009–.

Duan G, Ding L, Wei D *et al.* Screening endogenous signal peptides and protein folding factors to promote the secretory expression of heterologous proteins in *Pichia pastoris*. *J Biotechnol* 2019;**306**:193–202.

Duarte SOD, Monteiro GA. Plasmid replicons for the production of pharmaceutical-grade pDNA, proteins and antigens by *lactococcus lactis* cell factories. *Int J Mol Sci* 2021;**22**:1379.

Elegheert J, Behiels E, Bishop B *et al.* Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat Protoc* 2018;**13**:2991–3017.

Erden-Karaoğlan F, Karaoğlan M. Improvement of recombinant L-asparaginase production in *Pichia pastoris*. *3 BioTech* 2023;**13**:164.

Fang X, Lloyd CJ, Palsson BO. Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nat Rev Micro* 2020;**18**:731–43.

Flores-Villegas M, Rebnegger C, Kowarz V *et al.* Systematic sequence engineering enhances the induction strength of the glucose-regulated GTH1 promoter of *komagataella phaffii*. *Nucleic Acids Res* 2023;**51**:11358–74.

Fu J, Gao J, Liang Z *et al.* PDI-regulated disulfide bond formation in protein folding and biomolecular assembly. *Molecules* 2020;**26**:171.

Gao J, Xu J, Zuo Y *et al.* Synthetic biology toolkit for markerless integration of multigene pathways into *pichia pastoris* via CRISPR/Cas9. *ACS Synth Biol* 2022;**11**:623–33.

Gao J, Zuo Y, Xiao F *et al.* Biosynthesis of catharanthine in engineered *Pichia pastoris*. *Nat Synth* 2023a;**2**:231–42.

Gao L, Meng J, Dai W *et al.* Deciphering cell wall sensors enabling the construction of robust *P. pastoris* for single-cell protein production. *Biotechnol Biofuels Bioprod* 2023b;**16**:178.

García-Ortega X, Cámara E, Ferrer P *et al.* 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? *New Biotechnol* 2019;**53**:24–34.

Gibbons MD, Fang Y, Spicola AP *et al.* Enhancer-mediated formation of nuclear transcription initiation domains. *Int J Mol Sci* 2022;**23**:9290.

Gu Y, Gao J, Cao M *et al.* Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in *Pichia pastoris*. *World J Microbiol Biotechnol* 2019;**35**:79.

Guan B, Chen F, Su S *et al.* Effects of co-overexpression of secretion helper factors on the secretion of a HSA fusion protein (IL2-HSA) in *pichia pastoris*. *Yeast (Chichester Engl)* 2016;**33**:587–600. https://doi.org/10.1002/yea.3183.

Guo Y, Liao Y, Wang J *et al.* Methylotrophy of *Pichia pastoris*: current advances, applications, and future perspectives for methanol-based biomanufacturing. *ACS Sustain Chem Eng* 2022;**10**:1741–52.

Gupta A, Rangarajan PN. Coordinate regulation of methanol utilization pathway genes of *komagataella phaffii* by transcription factors and chromatin modifiers. *Front Microbiol* 2022;**13**:991192.

Gurramkonda C, Polez S, Skoko N *et al.* Application of simple fed-batch technique to high-level secretory production of insulin precursor using *pichia pastoris* with subsequent purification and conversion to human insulin. *Microb Cell Fact* 2010;**9**:31.

Hagman A, Säll T, Piškur J. Analysis of the yeast short-term Crabtree effect and its origin. *FEBS J* 2014;**281**:4805–14.

Hamilton SR, Davidson RC, Sethuraman N *et al.* Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* 2006;**313**:1441–3.

Helian Y, Gai Y, Fang H *et al.* A multistrategy approach for improving the expression of *E. coli* phytase in *Pichia pastoris*. *J Ind Microbiol Biotechnol* 2020;**47**:1161–72.

Hetz C, Zhang K, Kaufman RJ. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol* 2020;**21**:421–38.

Hou R, Gao L, Liu J *et al.* Comparative proteomics analysis of *pichia pastoris* cultivating in glucose and methanol. *Synth Syst Biotechnol* 2022;**7**:862–8.

Huang H, Liang Q, Wang Y *et al.* High-level constitutive expression of leech hyaluronidase with combined strategies in recombinant *pichia pastoris*. *Appl Microbiol Biotechnol* 2020;**104**:1621–32.

Huangfu J, Xu Y, Li C *et al.* Overexpressing target helper genes enhances secretion and glycosylation of recombinant proteins in *pichia pastoris* under simulated microgravity. *J Ind Microbiol Biotechnol* 2016;**43**:1429–39.

Idiris A, Tohda H, Kumagai H *et al.* Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl Microbiol Biotechnol* 2010;**86**:403–17.

Inan M, Aryasomayajula D, Sinha J *et al.* Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol Bioeng* 2006;**93**:771–8.

Ito Y, Terai G, Ishigami M *et al.* Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. *Nucleic Acids Res* 2020;**48**:13000–12.

Jiang L, Guan X, Liu H *et al.* Improved production of recombinant carboxylesterase FumDM by co-expressing molecular chaperones in *pichia pastoris*. *Toxins* 2023;**15**:156.

Jiang R, Yuan S, Zhou Y *et al.* Strategies to overcome the challenges of low or no expression of heterologous proteins in *escherichia coli*. *Biotechnol Adv* 2024a;**75**:108417.

Jiang X, Li M, Wang Z *et al.* CRISPR-mediated rDNA integration and fluorescence screening for pathway optimization in *Pichia pastoris*. *Chem Bio Eng* 2024b;**1**:940–51.

Jiao J, Wang S, Tian H *et al.* Decreased expression of LRA4, a key gene involved in rhamnose metabolism, caused up-regulated expression of the genes in this pathway and autophagy in *pichia pastoris*. *AMB Expr* 2020;**10**:37.

Jiao L, Zhou Q, Su Z *et al.* Efficient heterologous production of *rhizopus oryzae* lipase via optimization of multiple expression-related helper proteins. *Int J Mol Sci* 2018;**19**:3372.

Kaneyoshi K, Kuroda K, Uchiyama K *et al.* Secretion analysis of intracellular "difficult-to-express" immunoglobulin G (IgG) in Chinese hamster ovary (CHO) cells. *Cytotechnology* 2019;**71**:305–16.

Kang Z, Zhang N, Zhang Y. Enhanced production of leech hyaluronidase by optimizing secretion and cultivation in *pichia pastoris*. *Appl Microbiol Biotechnol* 2016;**100**:707–17.

Karaoglan M, Karaoglan FE, Inan M. Comparison of *ADH3* promoter with commonly used promoters for recombinant protein production in *Pichia pastoris*. *Protein Expression Purif* 2016;**121**:112–7.

Karaoglan M, Yildiz H, Inan M. Screening of signal sequences for extracellular production of *Aspergillus niger* xylanase in *Pichia pastoris*. *Biochem Eng J* 2014;**92**:16–21.

Karbalaei M, Rezaee SA, Farsiani H. *Pichia pastoris* : a highly successful expression system for optimal synthesis of heterologous proteins. *J Cell Physiol* 2020;**235**:5867–81.

Kato S, Ishibashi M, Tatsuda D *et al.* Efficient expression, purification and characterization of mouse salivary α-amylase secreted from methylotrophic yeast, *Pichia pastoris*. *Yeast* 2001;**18**:643–55.

Kjeldsen A, Kay JE, Baxter S *et al.* The fluorescent protein iLOV as a reporter for screening of high-yield production of antimicrobial peptides in *Pichia pastoris*. *Microb Biotechnol* 2022;**15**:2126–39.

Knychala MM, Boing LA, Ienczak JL *et al.* Precision fermentation as an alternative to animal protein, a review. *Fermentation* 2024;**10**:315.

Krshnan L, van de Weijer ML, Carvalho P. Endoplasmic reticulum-associated protein degradation. *Cold Spring Harb Perspect Biol* 2022;**14**:a041247.

Lai J, Song L, Zhou Y *et al.* Fine-tuned gene expression elements from hybrid promoter libraries in *Pichia pastoris*. *ACS Synth Biol* 2024;**13**:310–8.

Lebrun E, Shenshin V, Plaire C *et al.* Efficient full-length IgG secretion and sorting from single yeast clones in droplet picoreactors. *Lab Chip* 2023;**23**:3487–500.

Lee JY, Chen H, Liu A *et al.* Auto-induction of *pichia pastoris* AOX1 promoter for membrane protein expression. *Protein Expr Purif* 2017;**137**:7–12.

Li B, Zheng Y, Zhao S *et al.* High-level secretory production of recombinant E2-spy antigen protein via combined strategy in *pichia pastoris*. *Fermentation* 2024;**10**:99.

Li C, Lin Y, Zheng X *et al.* Combined strategies for improving expression of citrobacter amalonaticus phytase in *pichia pastoris*. *BMC Biotech* 2015;**15**:88.

Li J, Xie X, Cai J *et al.* Enhanced secretory expression and surface display level of Bombyx mori acetylcholinesterase 2 by *Pichia pastoris* based on Codon optimization Strategy for pesticides setection. *Appl Biochem Biotechnol* 2021;**193**:3321–35.

Liang S, Li C, Ye Y *et al.* Endogenous signal peptides efficiently mediate the secretion of recombinant proteins in *Pichia pastoris*. *Biotechnol Lett* 2013;**35**:97–105.

Liao X, Lin W, Chen N *et al.* Overexpression of the regulatory subunit of protein kinase A increases heterologous protein expression in *Pichia pastoris*. *Biotechnol Lett* 2020;**42**:2685–92.

Lin X, Ding W, Zheng S *et al.* Novel transcriptional regulation of the GAP promoter in *pichia pastoris* towards high expression of heterologous proteins. *Microb Cell Fact* 2024;**23**:206.

Lin X, Liang S, Han S *et al.* Quantitative iTRAQ LC–MS/MS proteomics reveals the cellular response to heterologous protein overexpression and the regulation of *HAC1* in *Pichia pastoris*. *J Proteomics* 2013;**91**:58–72.

Lin Y, Feng Y, Zheng L *et al.* Improved protein production in yeast using cell engineering with genes related to a key factor in the unfolded protein response. *Metab Eng* 2023;**77**:152–61.

Lin-Cereghino J, Hashimoto MD, Moy A *et al.* Direct selection of *pichia pastoris* expression strains using new G418 resistance vectors. *Yeast (Chichester Engl)* 2008;**25**:293–9.

Liu B, Li H, Zhou H *et al.* Enhancing xylanase expression by *komagataella phaffii* by formate as carbon source and inducer. *Appl Microbiol Biotechnol* 2022a;**106**:7819–29.

Liu C-P, Tsai T-I, Cheng T *et al.* Glycoengineering of antibody (Herceptin) through yeast expression and in vitro enzymatic glycosylation. *Proc Natl Acad Sci USA* 2018;**115**:720–5.

Liu C, Zhang Y, Ye C *et al.* Combined strategies for improving the heterologous expression of a novel xylanase from *fusarium oxysporum* Fo47 in *Pichia pastoris*. *Synthetic Systems Biotechnol* 2024;**9**:426–35.

Liu Q, Shi X, Song L *et al.* CRISPR–Cas9-mediated genomic multiloci integration in *pichia pastoris*. *Microb Cell Fact* 2019a;**18**:144.

Liu Q, Song L, Peng Q *et al.* A programmable high-expression yeast platform responsive to user-defined signals. *Sci Adv* 2022b;**8**:eabl5166.

Liu S, Li G, Ding L *et al.* Evaluation of SARS-CoV-2-neutralizing nanobody using virus receptor binding domain-administered model mice. *Research (Wash D C)* 2022c;**2022**:9864089.

Liu Y, Bai C, Liu Q *et al.* Engineered ethanol-driven biosynthetic system for improving production of acetyl-CoA derived drugs in crabtree-negative yeast. *Metab Eng* 2019b;**54**:275–84.

Lu X, Cummings C, Osuala UA *et al.* Characterization of recombinant human lactoferrin expressed in *Komagataella phaffii*. *Analyst* 2024;**149**:3636–50.

Lv X, Cui S, Chen J *et al.* Cascaded *de novo* biosynthesis of lactoproteins from CO_2 by engineered *pichia pastoris*. *Green Chem* 2023;**25**:5460–9.

Makrydaki E, Donini R, Krueger A *et al.* Immobilized enzyme cascade for targeted glycosylation. *Nat Chem Biol* 2024;**20**:732–41.

Mao J, Zhang H, Chen Y *et al.* Relieving metabolic burden to improve robustness and bioproduction by industrial microorganisms. *Biotechnol Adv* 2024;**74**:108401.

Marsalek L, Puxbaum V, Buchetics M *et al.* Disruption of vacuolar protein sorting components of the HOPS complex leads to enhanced secretion of recombinant proteins in *Pichia pastoris*. *Microb Cell Fact* 2019;**18**:119.

Martínez D, Menéndez C, Chacón O *et al.* Removal of bacterial dextran in sugarcane juice by Talaromyces minioluteus dextranase expressed constitutively in *Pichia pastoris*. *J Biotechnol* 2021;**333**:10–20.

Masuda T, Ueno Y, Kitabatake N. High yield secretion of the sweet-tasting protein lysozyme from the yeast *pichia pastoris*. *Protein Expr Purif* 2005;**39**:35–42.

Mattanovich D, Branduardi P, Dato L *et al.* Recombinant protein production in yeasts. In: Lorence A (ed.), *Recombinant Gene Expression*. Vol **824**. Totowa, NJ: Humana Press, 2012,329–58.

Mattanovich D, Graf A, Stadlmann J *et al.* Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb Cell Fact* 2009;**8**:29.

McCarty NS, Graham AE, Studená L *et al.* Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat Commun* 2020;**11**:1281.

Meng J, Liu S, Gao L *et al.* Economical production of *Pichia pastoris* single cell protein from methanol at industrial pilot scale. *Microb Cell Fact* 2023;**22**:198.

Mombeni M, Arjmand S, Siadat SOR *et al.* pMOX: a new powerful promoter for recombinant protein production in yeast *pichia pastoris*. *Enzyme Microb Technol* 2020;**139**:109582.

Nakamura Y, Nishi T, Noguchi R *et al.* A stable, autonomously replicating plasmid vector containing *Pichia pastoris* centromeric DNA. Löffler FE (ed.), *Appl Environ Microb* 2018;**84**:e02882–17.

Nand KN, Gupta JC, Panda AK. Overexpression and characterization of recombinant hCGβ-LTB vaccine in fed-batch culture of yeast *Pichia pastoris*. *Syst Microbiol and Biomanuf* 2024;**5**:171–84.https://doi.org/10.1007/s43393-024-00307-2.

Natarajan A, Haitjema CH, Lee R *et al.* An engineered survival-selection assay for extracellular protein expression uncovers hypersecretory phenotypes in *Escherichia coli*. *ACS Synth Biol* 2017;**6**:875–83.

Navone L, Moffitt K, Behrendorff J *et al.* Biosensor-guided rapid screening for improved recombinant protein secretion in *Pichia pastoris*. *Microb Cell Fact* 2023;**22**:92.

Neiers F, Belloir C, Poirier N *et al.* Comparison of different signal peptides for the efficient secretion of the sweet-tasting plant protein brazzein in *Pichia pastoris*. *Life* 2021;**11**:46.

Nett JH, Hodel N, Rausch S *et al.* Cloning and disruption of the *pichia pastoris ARG1, ARG2, ARG3, HIS1, HIS2, HIS5, HIS6* genes and their use as auxotrophic markers. *Yeast* 2005;**22**:295–304.

Nieto-Taype MA, Garcia-Ortega X, Albiol J *et al.* Continuous cultivation as a tool toward the rational bioprocess development with *Pichia Pastoris* Cell Factory. *Front Bioeng Biotechnol* 2020;**8**:632.

Nocon J, Steiger MG, Pfeffer M *et al.* Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. *Metab Eng* 2014;**24**:129–38.

Nurdiani D, Hariyatun H, Utami N *et al.* Enhancement in Human insulin precursor secretion by *Pichia pastoris* through modification of expression conditions. *HAYATI J Biosci* 2021;**29**:22–30.

Ogata K, Nishikawa H, Ohsugi M. Acetate utilization for yeast cell growth as sole carbon source. *Agric Biol Chem* 1969;**33**:977–8.

Papala A, Sylvester M, Dyballa-Rukes N *et al.* Isolation and characterization of human CapG expressed and post-translationally modified in *Pichia pastoris*. *Protein Expression Purif* 2017;**134**:25–37.

Piva LC, De Marco JL, Moraes LMPD *et al.* Construction and characterization of centromeric plasmids for *Komagataella phaffii* using a color-based plasmid stability assay. Fairhead C (ed.), *PLoS One* 2020;**15**:e0235532.

Prielhofer R, Reichinger M, Wagner N *et al.* Superior protein titers in half the fermentation time: promoter and process engineering for the glucose-regulated GTH1 promoter of *Pichia pastoris*. *Biotechnol Bioeng* 2018;**115**:2479–88.

Qiao S, Bai F, Cai P *et al.* An improved CRISPRi system in *pichia pastoris*. *Synthetic Systems Biotechnol* 2023;**8**:479–85.

Qu X, Zhao S, Wan C *et al.* Structural basis for the dynamic chaperoning of disordered clients by *Hsp90*. *Nat Struct Mol Biol* 2024;**31**:1482–91.

Ruan S, Yang Y, Zhang X *et al.* Screening and characterization of integration sites based on CRISPR-Cpf1 in *pichia pastoris*. *Synth Syst Biotechnol* 2024;**9**:759–65.

Sallada ND, Harkins LE, Berger BW. Effect of gene copy number and chaperone coexpression on recombinant hydrophobin HFBI biosurfactant production in *Pichia pastoris*. *Biotechnol Bioeng* 2019;**116**:2029–40.

Samuel P, Prasanna Vadhana AK, Kamatchi R *et al.* Effect of molecular chaperones on the expression of *Candida antarctica* lipase B in *Pichia pastoris*. *Microbiol Res* 2013;**168**:615–20.

Schmitz F, Röder A, Hoffrogge M *et al.* Agar plate-based activity assay for easy and fast screening of recombinant *Pichia pastoris* expressing unspecific peroxygenases. *Biotechnol J* 2024;**19**:2300421.

Schütz A, Bernhard F, Berrow N *et al.* A concise guide to choosing suitable gene expression systems for recombinant protein production. *STAR Protocols* 2023;**4**:102572.

Schwarzhans J-P, Luttermann T, Wibberg D *et al.* A mitochondrial autonomously replicating sequence from *Pichia pastoris* for uniform high level recombinant protein production. *Front Microbiol* 2017;**8**:780.

Shao Y, Xue C, Liu W *et al.* High-level secretory production of leghemoglobin in *pichia pastoris* through enhanced globin expression and heme biosynthesis. *Bioresour Technol* 2022;**363**:127884.

Shemesh P, Fishman A. Optimal fermentation conditions for growth and recombinant protein production in *pichia pastoris*: strain selection, ploidy level and carbon source. *Current Res Food Sci* 2024;**9**:100840.

Shen Q, Zhou X-T, Guo Q *et al.* Potential of the signal peptide derived from the PAS_chr3_0030 gene product for secretory expression of valuable enzymes in *Pichia pastoris*. *Appl Environ Microb* 2022;**88**:e0029622.

Shi M, McHugh KJ. Strategies for overcoming protein and peptide instability in biodegradable drug delivery systems. *Adv Drug Deliv Rev* 2023;**199**:114904.

Staudacher J, Rebnegger C, Dohnal T *et al.* Going beyond the limit: increasing global translation activity leads to increased productivity of recombinant secreted proteins in *Pichia pastoris*. *Metab Eng* 2022;**70**:181–95.

Summpunn P, Jomrit J, Panbangred W. Improvement of extracellular bacterial protein production in *Pichia pastoris* by co-expression of endoplasmic reticulum residing GroEL–GroES. *J Biosci Bioeng* 2018;**125**:268–74.

Sun H, Bankefa OE, Ijeoma IO *et al.* Systematic assessment of *Pichia pastoris* system for optimized β -galactosidase production. *Synthetic Systems Biotechnol* 2017;**2**:113–20.

Sun W, Zuo Y, Yao Z *et al.* Recent advances in synthetic biology applications of *Pichia* species. In: Darvishi Harzevili F (ed.), *Synthetic Biology of Yeasts*. Cham: Springer International Publishing, 2022, 251–92.

Sunga AJ, Tolstorukov I, Cregg JM. Posttransformational vector amplification in the yeast *pichia pastoris*. *FEMS Yeast Res* 2008;**8**:870–6.

Swiech K, de Freitas MCC, Covas DT *et al.* Recombinant glycoprotein production in human cell lines. *Methods Mol Biol* 2015;**1258**:223–40.

Thak EJ, Kim J, Lee D-J *et al.* Structural analysis of N-/O-glycans assembled on proteins in yeasts. *J Microbiol* 2018;**56**:11–23.

Tian E, Shen X, Xiao M *et al.* An engineered *Pichia pastoris* platform for the biosynthesis of silk-based nanomaterials with therapeutic potential. *Int J Biol Macromol* 2024;**269**:131954.

Tuo J, Nawab S, Ma X *et al.* Recent advances in screening amino acid overproducers. *Engineer Microbiol* 2023;**3**:100066.

Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* 2012;**13**:227–32.

Vogl T, Glieder A. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnol* 2013;**30**:385–404.

Vuree S. *Pichia pastoris* expression system: an impending candidate to express protein in industrial and biopharmaceutical domains. In: *New and Future Developments in Microbial Biotechnology and Bioengineering*. Amsterdam, Netherlands;Elsevier;2020:223–34.

Wang B, Wu Y, Lv X *et al.* Synergistic regulation of chassis cell growth and screening of promoters, signal peptides and fusion protein linkers for enhanced recombinant protein expression in *bacillus subtilis*. *Int J Biol Macromol* 2024a;**280**:136037.

Wang J, Wang X, Shi L *et al.* Methanol-independent protein expression by AOX1 promoter with trans-acting elements engineering and glucose-glycerol-shift induction in *Pichia pastoris*. *Sci Rep* 2017;**7**:41850.

Wang J, Wang X, Shi L *et al.* Reduced methanol input induces increased protein output by *AOX1* promoter in a *trans* -acting elements engineered *Pichia pastoris*. *J Ind Microbiol Biotechnol* 2018a;**45**:25–30.

Wang J, Wu Z, Zhang T *et al.* High-level expression of *thermomyces dupontii* thermophilic lipase in *pichia pastoris* via combined strategies. *BioTech* 2019a;**9**:62.

Wang J, Zhai H, Rexida R *et al.* Developing synthetic hybrid promoters to increase constitutive or diauxic shift-induced expression in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2018b;**18**:foy098.

Wang K, Yu S, Sun R *et al.* Biosynthesis of a functional fragment of human collagen II in *pichia pastoris*. *ACS Synth Biol* 2024b;**13**:2567–76.

Wang L, Li Y, Cui S *et al.* Engineering a transport pathway to boost extracellular production of human lactoferrin in *komagataella phaffii*. *Food Biosci* 2024c;**59**:104057.

Wang X, Li Y, Jin Z *et al.* A novel CRISPR/Cas9 system with high genomic editing efficiency and recyclable auxotrophic selective marker for multiple-step metabolic rewriting in *Pichia pastoris*. *Synthetic Systems Biotechnol* 2023a;**8**:445–51.

Wang X, Shen X, Sun Y *et al.* Production of *yarrowia lipolytica* lipase LIP2 in *pichia pastoris* using the nitrogen source-regulated *FLD1* promoter. *J Chem Technol Biotechnol* 2012;**87**:553–8.

Wang X, Wang P, Li W *et al.* Effect and mechanism of signal peptide and maltose on recombinant type III collagen production in *Pichia pastoris*. *Appl Microbiol Biotechnol* 2023b;**107**:4369–80.

Wang Y, Wang B, Gao Y *et al.* Highly efficient expression and secretion of human lysozyme using multiple strategies in *Pichia pastoris*. *Biotechnol J* 2023c;**18**:2300259.

Wang Y, Zhao N, Ma J *et al.* High-level expression of a novel α-amylase from *Thermomyces dupontii* in *Pichia pastoris* and its application in maltose syrup production. *Int J Biol Macromol* 2019b;**127**:683–92.

Wang Y, Zheng J, Lin F *et al.* Improved extracellular endo-1,4-β-mannosidase activity of recombinant *pichia pastoris* by optimizing signal peptide. *J Cent South Univ* 2015;**22**:2088–95.

Wang Y, Zhou J, Zhang Z *et al.* Efficient carbon flux allocation towards D-pantothenic acid production via growth-decoupled strategy in *Escherichia coli*. *Bioresour Technol* 2024d;**411**:131325.

Weninger A, Hatzl A-M, Schmid C *et al.* Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *pichia pastoris*. *J Biotechnol* 2016;**235**:139–49.

Wu X, Xu M, Yang J-R *et al.* Genome-wide impact of codon usage bias on translation optimization in *Drosophila melanogaster*. *Nat Commun* 2024;**15**:8329.

Xia W, Hu M, Pan Y *et al.* Improved production of streptomyces sp. FA1 Xylanase in a dual-plasmid *Pichia pastoris* system. *CIMB* 2021;**43**:2289–304.

Xiong A-S, Yao Q-H, Peng R-H *et al.* High level expression of a recombinant acid phytase gene in *pichia pastoris*. *J Appl Microbiol* 2005;**98**:418–28.

Xiong R, Chen J, Chen J. Secreted expression of human lysozyme in the yeast *Pichia pastoris* under the direction of the signal peptide from human serum albumin. *Biotechnol Appl Biochem* 2008;**51**:129–34.

Yang J, Jiang W, Yang S. *mazF* as a counter-selectable marker for unmarked genetic modification of *pichia pastoris*. *FEMS Yeast Res* 2009;**9**:600–9.

Yang J, Nie L, Chen B *et al.* Hygromycin-resistance vectors for gene expression in *Pichia pastoris*: *p ichia pastoris* vectors with hygromycin resistance. *Yeast* 2014;**31**:115–25.

Yang Y-X, Li Q, Li W-D *et al.* Factors and mechanisms affecting the secretion of recombinant protein in CHO cells. *Curr Pharm Biotechnol* 2023;**24**:391–400.

Ye R, Huang M, Lu H *et al.* Comprehensive reconstruction and evaluation of *Pichia pastoris* genome-scale metabolic model that accounts for 1243 ORFs. *Bioresour Bioprocess* 2017;**4**:22.

Yuan H, Zhou Y, Lin Y *et al.* Microfluidic screening and genomic mutation identification for enhancing cellulase production in *Pichia pastoris*. *Biotechnol Biofuels* 2022;**15**:50.

Zahrl RJ, Peña DA, Mattanovich D *et al.* Systems biotechnology for protein production in *pichia pastoris*. *FEMS Yeast Res* 2017;**17**:fox068. https://doi.org/10.1093/femsyr/fox068.

Zhang A-L, Luo J-X, Zhang T-Y *et al.* Recent advances on the GAP promoter derived expression system of *pichia pastoris*. *Mol Biol Rep* 2009;**36**:1611–9.

Zhang M, Song J, Xiao J *et al.* Engineered multiple translation initiation sites: a novel tool to enhance protein production in *Bacillus licheniformis* and other industrially relevant bacteria. *Nucleic Acids Res* 2022;**50**:11979–90.

Zhang S, Liu J, Xiao Z *et al.* Systems metabolic engineering of *saccharomyces cerevisiae* for the high-level production of (2S)-eriodictyol. *J Fungi* 2024a;**10**:119.

Zhang X, Gu S, Zheng X *et al.* A novel and efficient genome editing tool assisted by CRISPR-Cas12a/Cpf1 for *Pichia pastoris*. *ACS Synth Biol* 2021;**10**:2927–37.

Zhang Y-S, Gong J-S, Jiang J-Y *et al.* Engineering protein translocation and unfolded protein response enhanced human PH-20 secretion in *Pichia pastoris*. *Appl Microbiol Biotechnol* 2024b;**108**:54.

Zheng J, Guo N, Zhou H-B. A simple strategy for the generation of multi-copy *Pichia pastoris* with the efficient expression of mannanase. *J Basic Microbiol* 2014;**54**:1410–6.

Zheng Y, Li B, Zhao S *et al.* A universal strategy for the efficient expression of nanobodies in *Pichia pastoris*. *Fermentation* 2024;**10**:37.

Zhou H, Zhang W, Qian J. Hypersecretory production of glucose oxidase in *pichia pastoris* through combinatorial engineering of protein properties, synthesis, and secretion. *Biotechnol Bioeng* 2024;**121**:735–48.

Zhu W, Xu R, Gong G *et al.* Medium optimization for high yield production of human serum albumin in *pichia pastoris* and its efficient purification. *Protein Expr Purif* 2021;**181**:105831.

Received 18 December 2024; **revised** 19 February 2025; **accepted** 11 March 2025