مجله دنیای میکروب‌ها
سال هجدهم، شماره دوم (پیاپی ۶۳) تابستان ۱۴۰۴
صفحات ۱۲۵ –۱۱۴

Original Article

# طراحی محاسباتی ادغام ریبوسوئیچ تتراسایکلین در بالادست ژن ERG11 در مخمر *کومگاتائلا فافی* با استفاده از سیستم کریسپر-کس۹

**سجاد یزدان‌پناه[۱]، حسن محبت‌کار[۲*]، محمد برشان‌تشنیزی[۳]، ساره ارجمند[۴]**

[۱] دانشجوی دکترا، گروه زیست فناوری، دانشکده علوم و فناوری‌های زیستی، دانشگاه اصفهان، اصفهان، ایران.

[۲] استاد، گروه زیست فناوری، دانشکده علوم و فناوری‌های زیستی، دانشگاه اصفهان، اصفهان، ایران.

[۳] استادیار، گروه مهندسی زیستی، دانشکده مهندسی علوم زیستی، دانشکدگان علوم و فناوری‌های میان رشته ای، دانشگاه تهران، تهران، ایران.

[۴] استادیار، مرکز تحقیقات پروتئین، دانشگاه شهید بهشتی، تهران، ایران.

## چکیده

*سابقه و هدف:* *کومگاتائلا فافی* (*Komagataella phaffii*) به‌طور گسترده به‌عنوان گزینه‌ای برتر برای تولید پروتئین‌های نوترکیب شناخته شده است. بااین‌حال، کارایی آن در ترشح این پروتئین‌ها هنوز ایده‌آل نیست. یکی از عوامل محدودکننده، سختی غشای سلولی در این مخمر است که تا حد زیادی تحت‌تأثیر میزان ارگوسترول قرار دارد. ژن ERG11 مسئول کدگذاری آنزیم مهمی در مسیر بیوسنتز ارگوسترول است. کاهش بیان ERG11 می‌تواند باعث افزایش سیالیت غشاء و به طور بالقوه بهبود ترشح پروتئین شود. در این مطالعه، هدف طراحی روشی برای کنترل پویای بیان ERG11 با الحاق ریبوسوئیچ تتراسایکلین (TcRs) در ناحیه غیرکدشونده ۵' (5'UTR) آن بوده است.

*مواد و روش‌ها:* از تکنیک‌های محاسباتی شامل پیش‌بینی ساختار ثانویه RNA و داکینگ مولکولی استفاده شد تا محل‌های بهینه برای الحاق ریبوسوئیچ شناسایی شوند. علاوه بر این، راهبردی مبتنی بر کریسپر-کس۹ (CRISPR-Cas9) طراحی شد تا الحاق دقیق TcRs بالادست ERG11 از طریق نوترکیبی هم‌ساخت ممکن شود.

*یافته‌ها:* با تحلیل‌های محاسباتی، چهار محل الحاق مناسب برای TcRs شناسایی و یک سازه DNA دهنده شامل TcRs با دو بازوی همسان برای ورود دقیق طراحی شد. همچنین، سه sgRNA به‌صورت راهبردی طراحی شدند تا توالی‌های مجاور محل الحاق ریبوسوئیچ را هدف قرار دهند که این امر موجب افزایش کارایی و دقت الحاق از طریق نوترکیبی هم‌ساخت می‌شود.

*نتیجه‌گیری:* این مطالعه محل‌های مؤثر برای الحاق را مشخص و روش معتبری برای وارد کردن ریبوسوئیچ در جایگاه ژنی ERG11 ارائه می‌دهد. این روش می‌تواند به عنوان راهکاری برای تنظیم پویای سیالیت غشاء و افزایش ترشح پروتئین‌های نوترکیب در *کومگاتائلا فافی* مورد استفاده قرار گیرد.

*کلمات کلیدی:* پروتئین‌های نوترکیب، *کومگاتائلا فافی*، ارگوسترول، ریبوسوئیچ، کریسپر-کس۹، سیالیت غشاء.

دریافت مقاله: ۱۴۰۴/۴/۱۲ ویرایش مقاله: ۱۴۰۴/۵/۳۰ پذیرش مقاله: ۱۴۰۴/۶/۴

*) آدرس برای مکاتبه: گروه زیست فناوری، دانشکده علوم و فناوری‌های زیستی، دانشگاه اصفهان، اصفهان، ایران.
تلفن: ۰۹۱۳۴۰۱۹۴۳۶ پست الکترونیک: h.mohabatkar@ast.ui.ac.ir

Journal of Microbial World
Volume 18, No. 2, September 2025

Original Article

# Computational Design of Integrating Tetracycline Riboswitch into the 5'UTR of ERG11 Gene in *Komagataella phaffii* using CRISPR-Cas9 System

**Sajjad Yazdanpanah[1], Hassan Mohabbatkar[2], Mohammad Barshan-Tashnizi[3], Sareh Arjmand[4]**

[1] Ph.D student, Department of Biotechnology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran.
[2] Professor, Department of Biotechnology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran.
[3]Assistant Professor, Department of Bioengineering, School of Life Science Engineering, College of Interdisciplinary Science and Technology, University of Tehran, Tehran, Iran. [4]Assistant Professor, Protein Research Center, Shahid Beheshti University, Tehran, Iran.

## Abstract

*Background & Objectives:* *Komagataella phaffii* (*K. phaffii*) is widely recognized as a top choice for producing recombinant proteins. However, its efficiency in secreting these proteins remains less than ideal. One limiting factor is the rigidity of the yeast cell membrane, which is largely influenced by ergosterol levels. The ERG11 gene is responsible for coding a crucial enzyme in the ergosterol biosynthesis pathway. Reducing the expression of ERG11 could increase membrane fluidity, potentially improving protein secretion. In this study, we aim to design a method to dynamically control the expression of ERG11 by integrating a tetracycline riboswitch (TcRs) into its 5′ untranslated region (5′UTR).

*Materials & methods:* Computational techniques, including RNA secondary structure prediction and molecular docking, were utilized to identify optimal sites for riboswitch insertion. Additionally, a CRISPR/Cas9-based knock-in strategy was designed to enable precise integration of the TcRs upstream of ERG11 through homologous recombination.

*Results:* Based on computational analyses, four suitable insertion sites for TcRs were identified, and a DNA donor construct comprising TcRs with two homologous arms was designed for precise integration. Additionally, three single-guide RNAs (sgRNAs) were strategically designed to target sequences adjacent to the riboswitch insertion sites, thereby enhancing the efficiency and accuracy of knock-in via homologous recombination.

*Conclusion:* This work highlights effective insertion sites and provides a validated approach for introducing a regulatory riboswitch into the ERG11 locus. This method offers a potential means to modulate membrane fluidity dynamically, enhancing the secretion of recombinant proteins in *K. phaffii.*

*Keywords:* Recombinant Proteins, *Komagataella phaffi*, Ergosterol, Riboswitch, CRISPR-Cas9, Membrane Fluidity.

Received: 3 July 2025 Revised: 21 Augest 2025 Accepted: 26 Augest 2025

**Correspondence to: Hassan Mohabbatkar**
Tel: +98 9134019436
E-mail: h.mohabatkar@ast.ui.ac.ir
Journal of Microbial World 2025, 18 (2): 114 - 125
**DOI:** https://doi.org/10.82363/jmw.2025.1219004

## Introduction

*Komagataella phaffii (K. phaffii)*, formerly known as *Pichia pastoris*, is widely used as a host organism for recombinant protein production due to its rapid growth, scalability, and ability to perform post-translational modifications. However, protein secretion efficiency often remains suboptimal, mainly because of limitations in protein folding, intracellular trafficking, and membrane translocation (1–3). Recent studies have highlighted the important role of membrane composition in regulating protein secretion in *K. phaffii*. Ergosterol, the main sterol found in fungal cell membranes, is crucial for maintaining membrane integrity and fluidity. The ERG11 gene encodes lanosterol 14α-demethylase, a key enzyme in the ergosterol biosynthesis pathway responsible for removing the 14α-methyl group from lanosterol. Prior research has shown that suppressing ergosterol synthesis can improve recombinant protein secretion in *K. phaffii* by modifying membrane composition and increasing membrane fluidity (4-8).

Riboswitches are regulatory RNA elements that directly bind small molecules and control gene expression without requiring protein cofactors (9-13). The tetracycline riboswitch system has been successfully implemented in yeast for dose-dependent translational control, with insertion into the untranslated region 5'UTR resulting in up to ninefold suppression of gene expression upon tetracycline addition (14). The positioning of the riboswitch within the 5'UTR significantly affects regulatory efficiency, with insertions near the start codon generally providing more effective repression than those near the cap (15-17).

This study aims to employ in silico approaches to identify optimal insertion sites for the tetracycline riboswitch within the ERG11 5’UTR of *K. phaffii*. A knock-in experiment was designed using the CRISPR-Cas9 system to insert the TcRs into the 5’UTR of ERG11. This strategy has the potential to enhance recombinant protein secretion by modulating membrane fluidity through controlled expression of ERG11.

## Materials and methods

*A) Sequence Retrieval and Analysis of the ERG11 Upstream Regulatory Region:* The genomic sequence encompassing the ERG11 coding region (NC-012965.1) and a non-coding 504-nucleotide sequence immediately upstream of its annotated start codon (ATG) was retrieved from the NCBI Nucleotide database (Fig 1). *In silico* analysis of the upstream region (positions -504 to -1 relative to the ERG11 ATG) was performed using the NCBI ORF FINDER tool (default parameters: minimum ORF length = 30 nt, genetic code = standard). A 40-nucleotide segment, immediately before the start codon, was identified as the likely 5′ untranslated region (5′UTR)(18).

*B) Riboswitch Insertion Library Design:* A 15-nucleotide segment spanning positions -30 to −15 relative to the start codon was selected as the target region for TcRs integration (Fig 1). Within this segment, fourteen unique insertion sites were identified at single-nucleotide intervals. The entire sequence, consisting of the 40-nucleotide 5′ untranslated region (5′UTR) with the TcRs element inserted at each of the 14 positions, was transcribed into RNA, generating a library of 14 distinct mRNA variants (19).

*C) Stability Assessment of 5'UTR Variants containing TcRs*: RNAfold (ViennaRNA v2.6.0) applied to predict the secondary structures and folding energies (ΔG) of all 14 variants under default settings (temperature = 28°C, no constraints) (14). From these, a single

structure was selected for further analysis based on its minimal deviation in ΔG from the native 5′UTR (ΔG = 0 kcal/mol), thereby ensuring minimal structural perturbation (21).

*D) Molecular Docking*: The three-dimensional structures of these variants were modeled using RNAComposer, which provided accurate predictions of their tertiary conformations (22). Molecular docking simulations were then performed between the tetracycline ligand and each of the 14 variants using HDOCK, a rigid-body docking tool that exhaustively samples ligand orientations to estimate binding affinities. The resulting docking scores allowed for a comparative assessment of tetracycline's binding potential across all insertion sites, identifying those with the most favorable interactions. Among them, three structures were selected for further analysis based on their higher docking scores (23).

*E) Single-Guide RNA (sgRNA) Selection*: The sgRNA sequences targeting ERG11 from *K. phaffii* were designed using the online tool CRISPOR (http://crispor.org/). CRISPOR was employed to select sgRNAs with high on-target efficiency and minimal off-target effects, based on the *K. phaffii* genome sequence (24).

CRISPi plasmids developed by Brigitte Gasser's lab have been suggested for CRISPR/Cas9-mediated genome editing in *K. phaffii*. These plasmid vectors, which express the Cas9 nuclease, are available at Addgene as (#1000000136). They contain a cloning site for the sgRNA. The protocol for cloning the sgRNA into these vectors should follow the associated publications (25).

*F) Design of the Donor DNA or Homologous DNA:* For the knock-in experiment, a donor DNA template was designed to contain the TcRs flanked by homology arms, each ranging from 400 to 700 base pairs in length. These homology arms are homologous to the genomic sequences upstream and downstream of the Cas9 cleavage site in the ERG11 gene.

## Results

*A) Analysis of the ERG11 Upstream Regulatory Region to recognize the possible insertion sites for TcRs:* The analysis of the ERG11 upstream sequence using NCBI ORF FINDER identified two upstream Open Reading Frames (uORFs) within the 5' leader sequence. As shown in Fig. 2, uORF1 (highlighted in grey in the annotated sequence) initiates at position -346 and terminates at position -242, encoding a putative peptide of 34 amino acids. uORF2 (highlighted in blue) is located closer to the main ERG11 ORF, initiating at position -220 and terminating at position -41, encoding a putative 59 amino acid peptide.

The uORFs in the 5′ leader sequence, such as those identified upstream of ERG11, are important regulatory elements that modulate ERG11 protein expression by controlling the translation of its mRNA (26,27). Therefore, to avoid disrupting natural gene regulation mediated by these uORFs, riboswitch insertion sites should be selected outside of these regions. Ideally, insertion points should be located close to the start codon to allow better control of ERG11 expression without interfering with the native regulatory mechanisms.

Immediately downstream of uORF2 and directly upstream of the start codon, a 40-nucleotide segment (positions -40 to -1 relative to the ERG11 ATG) was defined as the core putative 5'UTR. The region showed no secondary structure (ΔG = 0 kcal/mol), suggesting potential inclusion of the native 5′UTR, which indicates a predominantly linear, unstructured conformation under the modeled conditions. Within the 5′UTR, a 15-nucleotide region was targeted as the potential insertion site for TcRs.

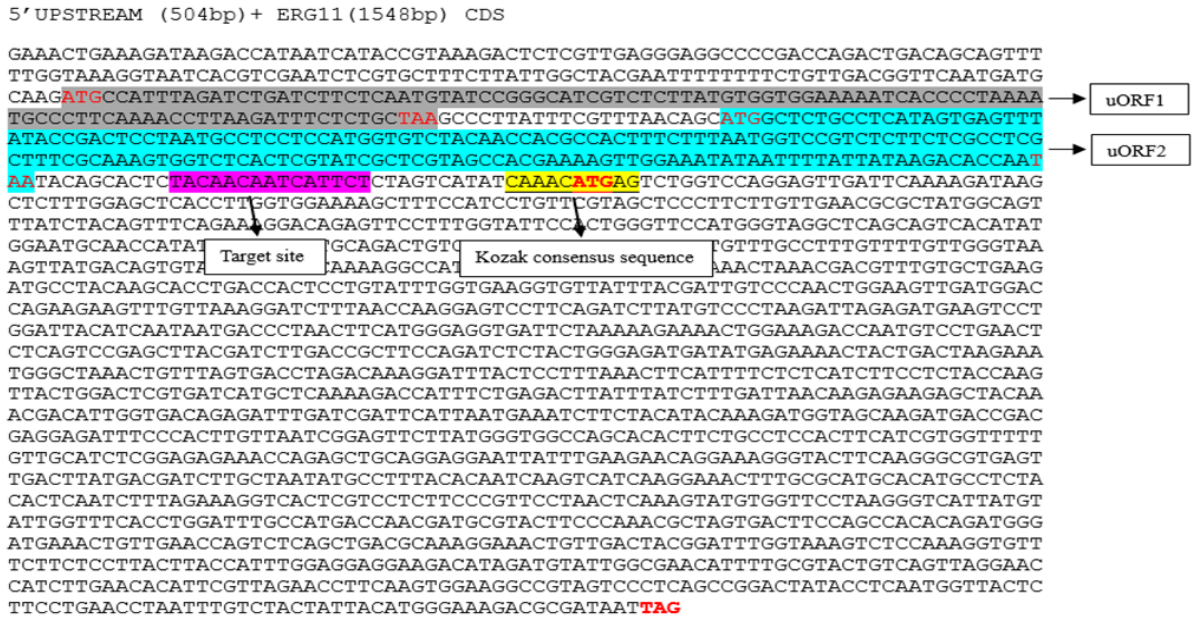


**Fig 1.** The annotated sequence highlights two upstream open reading frames (uORFs) probably regulating ERG11 translation. uORF1 (grey) initiates at position –346 and terminates at –242 relative to ERG11 ATG. uORF2 (blue), positioned closer to the main ERG11 ORF, starts at –220 and ends at -41. The canonical Kozak consensus sequence (CAAACATGAG) surrounding the ERG11 start codon (ATG, +1) is underlined, highlighting its role in translation initiation efficiency. The 40-nucleotide region from –40 to –1 relative to the ATG, presumed to be the 5' untranslated region (5' UTR) of ERG11, contains a 15-nucleotide segment highlighted in pink. This segment was selected as the target site for TcRs insertion.

*B) Stability Assessment of 5′UTR Variants containing TcRs:* To evaluate the potential impact of TcRs insertions on the secondary structure of the ERG11 5′ untranslated region (5′UTR), we used RNAfold (ViennaRNA v2.6.0) to predict the minimum free energy (MFE) structures for all 14 insertions (Table 1). The native 40-nucleotide 5′UTR had an MFE of 0 kcal/mol, which served as the baseline reference, while the isolated TcRs sequence alone had an MFE of -23.71 kcal/mol. Among the variants, MFE values ranged from -24.56 to -26.48 kcal/mol, indicating varying degrees of structural perturbation relative to the native sequence. The differences in MFE values were not significant. However, variant 13 showed the smallest change in free energy (-24.56 kcal/mol) compared to the other variants, suggesting that insertion at this position causes the least disruption to the native RNA secondary structure. This minimal perturbation is critical, as preserving the native 5′UTR folding is likely essential for the natural translational regulation of ERG11. Therefore, variant 13 represents a promising candidate for further functional analysis, as it balances effective riboswitch integration with preservation of the native mRNA's structural integrity.

*C) Molecular Docking:* The HDOCK scores revealed notable variation in binding affinity depending on the insertion position, indicating that the structural context within the 5′UTR significantly modulates tetracycline accessibility (Table 2). Several sites, particularly those at positions 6, 7, and 8, exhibited the most favorable (highest) docking scores, suggesting these regions may provide optimal environments for ligand interaction and are promising candidates for engineering functional riboswitches. In contrast, other sites, such as 1 and 4, showed

lower binding affinities. These findings highlight the importance of insertion site selection for effective riboswitch insertion and offer a rational basis for prioritizing candidates for experimental validation. It is important to note that while HDOCK scores provide comparative binding estimates, they do not fully capture the complexities of RNA-ligand dynamics or cellular context, underscoring the need for molecular dynamics and the subsequent experimental confirmation. Molecular docking analysis illustrating the interaction between tetracycline and TcRs is shown in Fig 2.

*D) sgRNA Selection:* A sequence of 75 nt surrounding position 13 was analyzed using CRISPOR. Three sgRNAs were initially selected and are listed in Table 3. As shown in Fig. 3, they are located in the green region, indicating a GC% of 40-49. Studies have found that GC content in the sgRNA sequence between 40% and 60% increases on-target activity because moderate GC content stabilizes the DNA-RNA duplex, thereby enhancing binding strength without causing excessive rigidity or misfolding (20). Among these, 54forw, with zero off-targets, was chosen as the primary candidate for the experiment. The selection was based on criteria to ensure high editing efficiency and minimal off-target effects.

*E) Constructing donor DNA for homologous recombination:* Among four insertion sites that were screened based on the minimal deviation in ΔG from the native 5′UTR (ΔG = 0 kcal/mol) and the docking scores, position 13 was selected as the first candidate for integrating the riboswitch DNA sequence, due to its close distance to the start codon, its high docking score, and offering the minimal deviation in ΔG from the native 5′UTR (-24.56 kcal/mol). A sketch of the donor DNA, or homologous DNA, is shown in Figure 4. The DNA

**Table 1.** Predicted minimum free energy (MFE) values for TcRs and the various TcRs-inserted 5'UTR variants, as determined by RNAfold. More negative MFE values indicate increased RNA secondary structure stability, indicating more degrees of structural perturbation relative to the native sequence of TcRs.

| Positions | Variants | RNAfold predicted MFE (kcal/mol) |
|---|---|---|
| | TcRs | -23.71 |
| | -40 to -1 | 0 |
| 1 | -40…-30/TcRs/-29…-1 | -25.96 |
| 2 | -40…-29/TcRs/-28...-1 | -24.84 |
| 3 | -40…-28/TcRs/-27...-1 | -25.32 |
| 4 | -40…-27/TcRs/-26...-1 | -25.43 |
| 5 | -40…-26/TcRs/-25…-1 | -24.84 |
| 6 | -40…-25/TcRs/-24…-1 | -25.32 |
| 7 | -40…-24/TcRs/-23…-1 | -25.43 |
| 8 | -40…-23/TcRs/-22…-1 | -26.48 |
| 9 | -40…-22/TcRs/-21…-1 | -24.8 |
| 10 | -40…-21/TcRs/-20…-1 | -25.32 |
| 11 | -40…-20/TcRs/-19…-1 | -26.08 |
| 12 | -40…-19/TcRs/-18…-1 | -25.12 |
| 13 | -40…-18/TcRs/-17…-1 | -24.56 |
| 14 | -40…-17/TcRs/-16…-1 | -25.22 |

**Table 2.** HDOCK-predicted binding scores (in kcal/mol) for TcRs inserted at various positions within the 5'UTR of the ERG11 gene. More negative HDOCK scores indicate stronger predicted binding affinities between tetracycline and the riboswitch.

| Positions | Variants | HDOCK scores (kcal/mol) |
|---|---|---|
| 1 | -40…-30/TcRs/-29...-1 | -212.94 |
| 2 | -40…-29/TcRs/-28...-1 | -235.04 |
| 3 | -40…-28/TcRs/-27...-1 | -231.34 |
| 4 | -40…-27/TcRs/-26...-1 | -229.11 |
| 5 | -40…-26/TcRs/-25…-1 | -232.43 |
| 6 | -40…-25/TcRs/-24…-1 | -254.36 |
| 7 | -40…-24/TcRs/-23…-1 | -245.76 |
| 8 | -40…-23/TcRs/-22…-1 | -239.91 |
| 9 | -40…-22/TcRs/-21…-1 | -226.9 |
| 10 | -40…-21/TcRs/-20…-1 | -225.94 |
| 11 | -40…-20/TcRs/-19…-1 | -221.27 |
| 12 | -40…-19/TcRs/-18…-1 | -210.1 |
| 13 | -40…-18/TcRs/-17…-1 | -227.91 |
| 14 | -40…-17/TcRs/-16…-1 | -220.51 |

**Table 3.** Selected sgRNAs targeting the 75 nt region surrounding position 13.

| guideId | targetSeq | mitSpecScore | cfdSpecScore | PAM | offtargetCount |
|---|---|---|---|---|---|
| 48forw | AGTCATATCAAACATGAGTCTGG | 100 | 100 | TGG | 1 |
| 52rev | TTATCTTTTGAATCAACTCCTGG | 100 | 100 | TGG | 1 |
| 54forw | ATCAAACATGAGTCTGGTCCAGG | 100 | 100 | AGG | 0 |

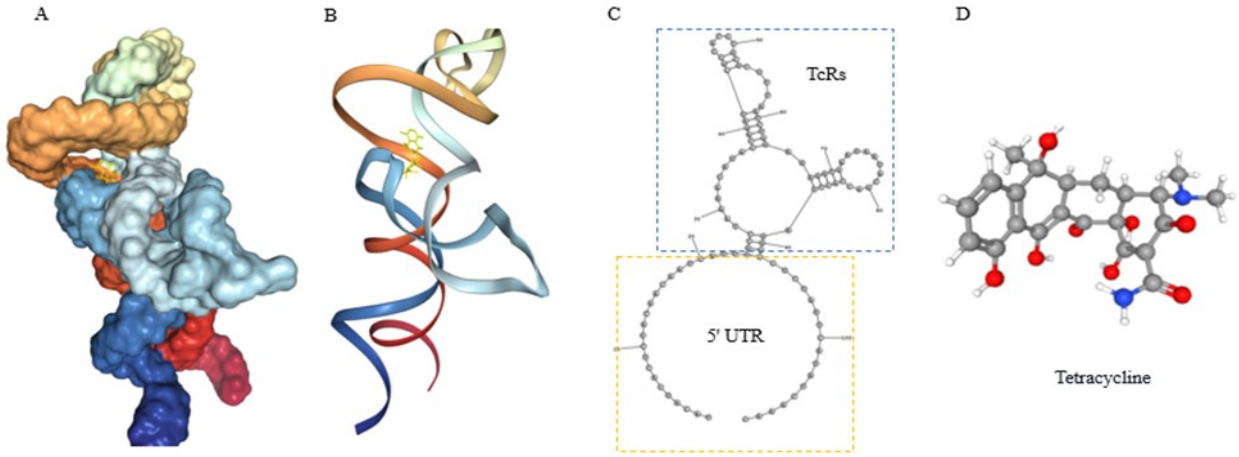


**Fig 2.** (A, B) Molecular docking analysis illustrating the interaction between tetracycline and a TcRs-inserted 5'UTR variant. (C) Predicted secondary structure of the tetracycline with TcRs, positioned within the 5' untranslated region (5'UTR) of the ERG11 gene, created by the AllSub web server. (D) Chemical structure of tetracycline obtained from PubChem.

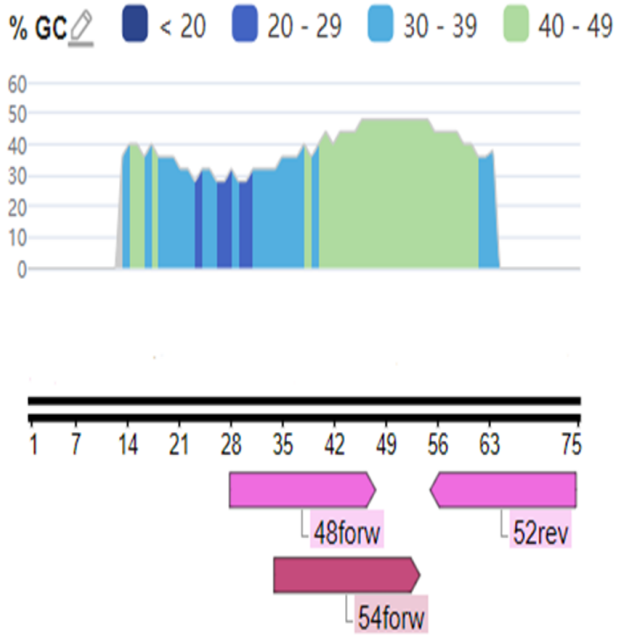


**Fig 3.** sgRNA locations on the 75-nt sequence, which includes 40 nt upstream and 35 nt downstream of the start codon. The image is the output from the CRISPOR server, visualized using Vectorbee.

sequence of TcRs (69 bp) is located 17 bp upstream of the start codon (ATG). The Cas9 cut site is also located a few nucleotides after it (recognized by CRISPOR.org).

As shown in Figure 5, to synthesize the construct of the donor DNA (1103 bp), it was divided into three fragments: the upstream fragment (approximately equal to the right homology arm and 370 bp in length), a fragment containing the riboswitch sequence in the middle (201 bp), and the downstream fragment (approximately equal to the left homology arm and 597 bp in length). As illustrated in Figure 6, the riboswitch-containing fragment, which overlaps with both the upstream and downstream fragments, is synthesized separately. The upstream and downstream fragments are amplified by PCR from *K. phaffii* genomic DNA. Finally, all three fragments are joined together through SOEing PCR reactions.


**Fig 4.** The precise positioning of the right homology arm (432 bp), the tetracycline riboswitch insertion site, the start codon, the Cas9 cut site, the PAM site, and the left homology arm (602 bp) in donor DNA. The original PAM site was AGG modified to AAG to prevent recutting by CAS9.

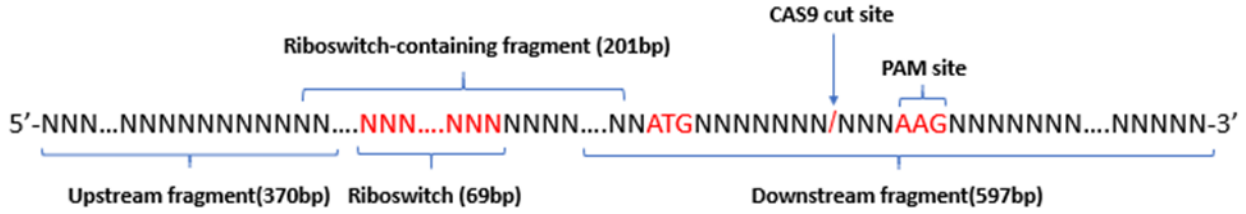


**Fig 5.** Donor DNA containing three fragments for SOEing PCR: the upstream fragment (370bp), the riboswitch-containing fragment (201), and the downstream fragment (597).

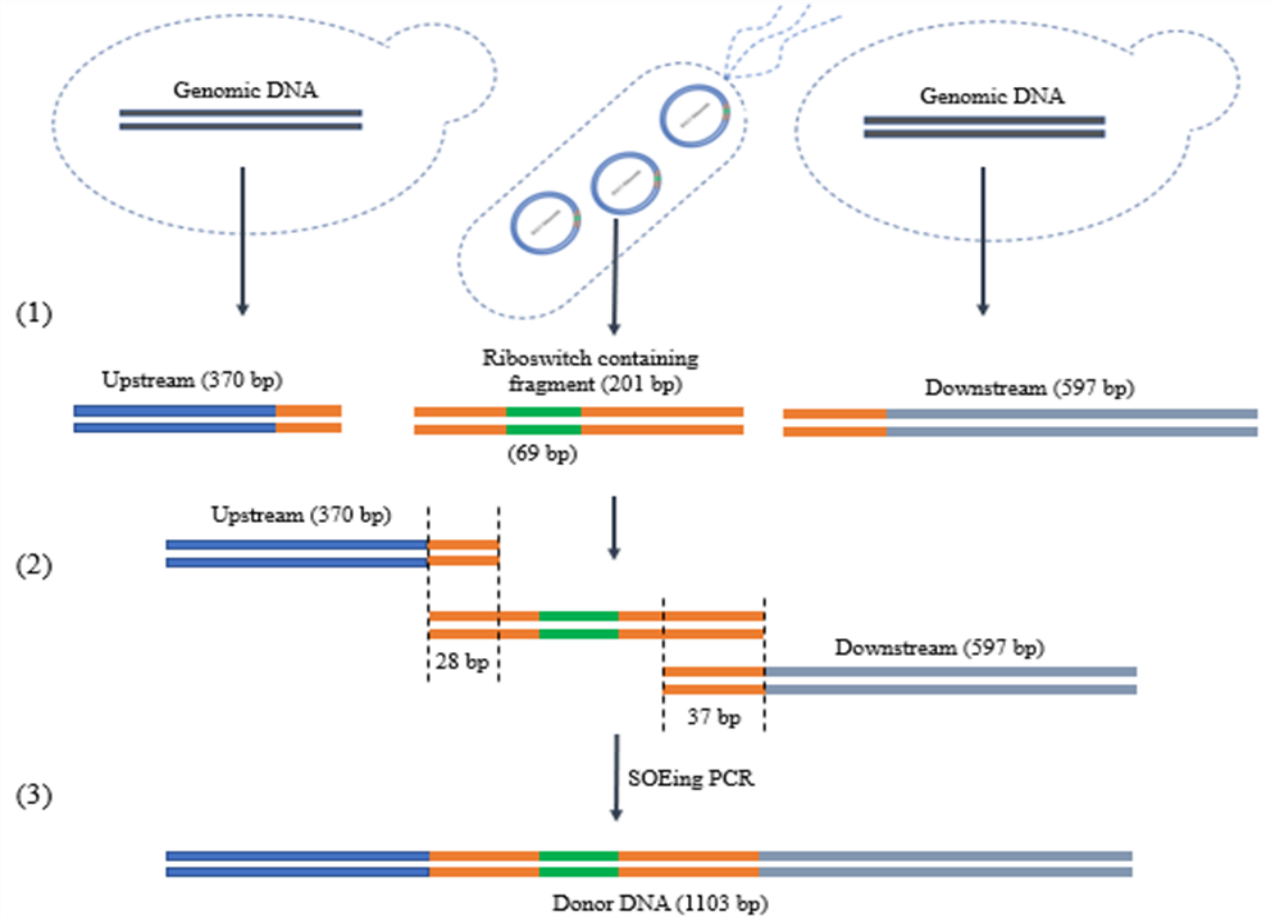


**Fig 6.** (1) The upstream and downstream fragments were obtained from the genomic DNA of K. phaffii via PCR. The riboswitch-containing fragment was also obtained by PCR from the PUC57-TcRs plasmid. (2) The riboswitch-containing fragment has a 28 bp overlap with the upstream fragment and a 37 bp overlap with the downstream fragment. (3) The three fragments were joined using SOEing PCR.

## Discussion

The efficient production of recombinant proteins in microbial hosts like *K. phaffii* is often hampered by bottlenecks in the secretion pathway. This study addresses one such bottleneck, membrane rigidity influenced by ergosterol content, by proposing a novel synthetic biology approach to dynamically control the expression of ERG11. Our computational design successfully identifies optimal sites for integrating a tetracycline-responsive riboswitch into the 5'UTR of ERG11 and outlines a precise CRISPR/Cas9-based strategy for its genomic knock-in.

A critical first step in our design was the in-depth analysis of the upstream regulatory region of the ERG11 gene. Here, the identification of two uORFs (uORF1 and uORF2) is a significant finding, as uORFs are well-established cis-regulatory elements that can profoundly impact translation initiation and mRNA stability in eukaryotes (26,27). By strategically selecting the insertion site for the TcRs within the core 40-nt 5'UTR, downstream of these uORFs, we aimed to avoid disrupting this native layer of translational control. Furthermore, we prioritized sites closer to the start codon. Besides the reasons of minimal structural perturbation and strong docking score, Variant 13 (inserted at position) 17-was chosen as the primary candidate due to its closer proximity to the start codon. Studies have demonstrated that riboswitches positioned nearer to the translation initiation site often exhibit more potent regulatory effects, likely because they exert more direct steric hindrance on the assembling ribosome (14-16). This ensures that the primary regulatory mechanism governing ERG11 expression remains intact, and any observed phenotypic changes can be more confidently attributed to the function of the engineered riboswitch.

Our integrated computational pipeline, combining RNA secondary structure prediction and molecular docking, proved highly effective in prioritizing candidate insertion sites. The use of RNAfold to calculate minimum free energy (MFE) provided a measure of the structural perturbation caused by riboswitch integration. Variant 13 emerged as a top candidate because it induced the smallest change in MFE ($\Delta G$ =-24.56 kcal/mol) relative to the native, unstructured 5'UTR ($\Delta G$ = 0 kcal/mol). This suggests that the insertion at this site preserves the natural structural context of the 5'UTR. Complementing this, molecular docking simulations using HDOCK revealed that several other positions, particularly 6, 7, and 8, also exhibited highly favorable docking scores, indicating they are potentially excellent candidates for insertion as well. This underscores a key principle in riboswitch engineering: the local structural context dictated by the insertion site can dramatically influence ligand accessibility and binding energy(14,15). To further discriminate between these top candidates (e.g., Variants 6, 7, 8, and 13), performing molecular dynamics (MD) simulations would be a valuable next step. MD could provide deeper insights into the stability, flexibility, and conformational changes of the riboswitch-ligand complex over time, helping to select the variant with the most robust and reliable performance *in vivo*.

To translate these computational predictions into a testable biological system, a robust CRISPR /Cas9-mediated homology-directed repair (HDR) strategy was developed. The selection of sgRNAs with high specificity and zero off-target predictions (e.g., 54forw) is paramount for ensuring precise editing and minimizing genotoxic effects (24,28). Furthermore, the design of the donor DNA construct, featuring long homology arms and a modified PAM site

to prevent Cas9 re-cleavage of the edited locus, follows established best practices for achieving high-efficiency knock-in in yeasts like *K. phaffii* (25). The use of SOEing PCR to assemble the final donor template is a practical and reliable method for generating the complex DNA constructs required for HDR.

This study has some limitations. While our computational designs are powerful, they are predictive. The true test of functionality, the dose-dependent repression of ERG11 by tetracycline and its subsequent effect on the membrane fluidity and protein secretion, requires empirical validation *in vivo*. Future work should include the aforementioned molecular dynamics simulations to better understand the binding kinetics. Subsequently, the planned CRISPR/Cas9 knock-in and physiological assays in *K. phaffii* will be essential to confirm that reduced ERG11 expression leads to altered ergosterol levels, increased membrane fluidity, and ultimately, enhanced secretion of target recombinant proteins, as suggested by previous research (4).

## Conclusion

Collectively, this study delivers a comprehensive, *in silico*-informed blueprint for engineering dynamic gene regulation in *K. phaffii*, showcasing the power of riboswitches to optimize recombinant protein yields. The transition to *in vivo* experimentation is now essential to empirically verify the predicted control of ERG11 and its physiological impact on membrane properties and secretory capacity. Proven success here would demonstrate a transferable methodology, paving the way for its application to myriad targets across diverse fungal biotechnological platforms.

## Acknowledgment

This research article is the result of a doctoral dissertation and is supported by the Department of Biotechnology, Faculty of Biological Science and Technology, the University of Isfahan. The authors are grateful to Dr. P. Rabiee and Dr. M. Momeni for their expert advice on RNA structure modeling, which greatly enhanced the quality of this work.

## Conflict of Interest

To the best of our knowledge, no financial or personal relationships exist that could be construed as influencing the design, execution, or interpretation of the findings in this research. This study was conducted in accordance with ethical standards, and all necessary approvals and consents were obtained.

## References


1. Juturu V, Wu JC. Heterologous protein expression in *Pichia pastoris*: latest research progress and applications. ChemBioChem. 2018;19.
2. Pan Y, Yang J, Wu J, Yang L, Fang H. Current advances of *Pichia pastoris* as cell factories for production of recombinant proteins. Front Microbiol. 2022;13:1059777.
3. Lv W, Cai M. Advancing recombinant protein expression in *Komagataella phaffii*: opportunities and challenges. FEMS Yeast Res. 2025;25.
4. Baumann K, Adelantado N, Lang C, Mattanovich D, Ferrer P. Protein trafficking, ergosterol biosynthesis and membrane physics impact recombinant protein secretion in *Pichia pastoris*. Microb Cell Fact. 2011;10.
5. Adelantado, N., Tarazona, P., Grillitsch, K., García-Ortega, X., Monforte, S., Valero, F., Feussner, I., Daum, G., & Ferrer, P. (2017). The effect of hypoxia on the lipidome of recombinant *Pichia pastoris*. Microb Cell Fact, 16, 86. https://doi.org/10.1186/s12934-017-0699-4
6. Foroudi, M. R., Khodavandi, A., Alizadeh, F., Mahmoudi, H., & Pishgar, E. Presence pattern of genes related to ergosterol biosynthesis in colonizing *Candida albicans* isolates from patients with multiple sclerosis. J Microbial World, 2019; 17(4) (In Persian).
7. Majdi M,Khazaei Koohpar Z,Nasrollahi Omran A. Investigation of mutations in hotspot regions of ERG11 gene in fluconazole-resistant isolates of *Candida albicans* in the west of Mazandaran. J Microbial World, 2018; 11(3) (In Persian).
8. Solat R, Rahimi P, Bakhshi Khaniki GR, Aghasadeghi MR, Vahabpour R, Shokri M. Expression of full length of HCV-E2 gene in yeast *Pichia pastoris*. J Microbial World. 2016;8(4) (In Persian).
9. Edwards BAL, Batey RT. Riboswitches: a common RNA regulatory element riboregulation in bacteria. Nat Educ. 2010;3(9).
10. Ge H, Marchisio MA. Aptamers, riboswitches and ribozymes in *S. cerevisiae* synthetic biology. Life. 2021;11.
11. Wittmann A, Suess B. Engineered riboswitches: expanding researchers' toolbox with synthetic RNA regulators. FEBS Lett. 2012;586(15):2076–83.
12. Serganov A, Nudler E. A decade of riboswitches. Cell. 2013;152(1-2):17–24.
13. Bauer G, Suess B. Engineered riboswitches as novel tools in molecular biology. J Biotechnol. 2006 Jun 25;124 (1):4–11.
14. Hanson S, Berthelot K, Fink B, McCarthy JEG, Suess B. Tetracycline-aptamer-mediated translational regulation in yeast. Mol Microbiol. 2003;49(6).
15. Suess B, Hanson S, Berens C, Fink B, Schroeder R, Hillen W. Conditional gene expression by controlling translation with tetracycline-binding aptamers. Nucleic Acids Res. 2003;31.
16. Weigand JE, Suess B. Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. Nucleic Acids Res. 2007;35(12).
17. Vieira IPV, Pimentel FSA, Coelho CM, De Marco JL, de Moraes LMP, Torres FAG. Use of an on/off tetracycline riboswitch to control protein production in *Komagataella phaffii*. AMB Express. 2023;13(1).
18. Wieder, N., D'Souza, E. N., Martin-Geary, A. C., Lassen, F. H., Talbot-Martin, J., Fernandes, M., Chothani, S. P., Rackham, O. J. L., Schafer, S., Aspden, J. L., MacArthur, D. G., Davies, R. W., & Whiffin, N.. Differences in 5'untranslated regions highlight the importance of translational regulation of dosage sensitive genes. Genome Biology, 2024; 25(1), 111.
19. Kelvin, D., & Suess, B.. Tapping the potential of synthetic riboswitches: reviewing the versatility of the tetracycline aptamer. *RNA Biology*, 2023; *20*(1), 457.

20. Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. . ViennaRNA Package 2.0. Algorithms for Molecular Biology, 2011; 6(1), 1–14.

21. Yan Y, Zhang D, Zhou P, Li B, Huang SY. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res. 2017;45(W1).

22. Sarzynska J, Popenda M, Antczak M, Szachniuk M. RNA tertiary structure prediction using RNAComposer in CASP15. Proteins. 2023;91(12):1790–9.

23. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125, 1731–1737 (2003).

24. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 2018;46(W1).

25. Gassler T, Heistinger L, Mattanovich D, Gasser B, Prielhofer R. CRISPR/Cas9-mediated homology-directed genome editing in *Pichia pastoris*. Methods Mol Biol. 2019.

26. Gunišová S, Valášek LS. Fail-safe mechanism of GCN4 translational control - uORF2 promotes reinitiation by analogous mechanism to uORF1 and thus secures its key role in GCN4 expression. Nucleic Acids Res. 2014;42 (9):5880–93.

27. Lin Y, May GE, Kready H, Nazzaro L, Mao M, Spealman P, et al. Impacts of uORF codon identity and position on translation regulation. Nucleic Acids Res. 2019;47(17):9358–67.

28. Mengstie MA, Azezew MT, Dejenie TA, Teshome AA, Admasu FT, Teklemariam AB, et al. Recent advancements in reducing the off-target effect of CRISPR-Cas9 genome editing. Biologics. 2024;18:21.