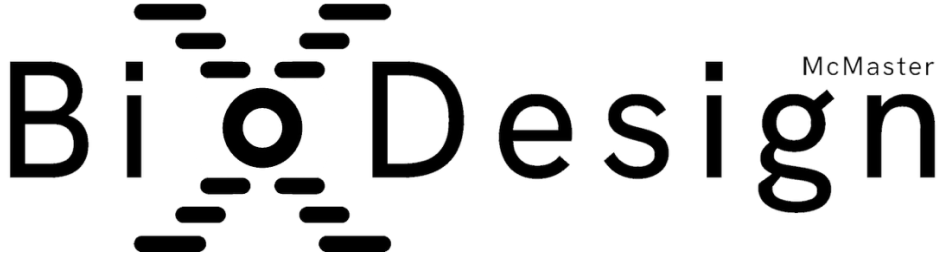
Investigation of the Cold-Shock Response Element of *cspA* as a Target for Directed Evolution

McMaster BioDesign – iDEC 2025

Olamide Asa[a], Ainoor Arora[b], Jordan Classen[e], Mischa Esmail[b], Saejin (Grace) Hur[a], Andrew Lian[b], Irma Lozica[c], Varnigha Mayooran[c], Zachary McKay[d], Sarah Moran[b], Eshana Pararajasegaram[c], Tvara Parikh[b], William Pihlainen-Bleecker[f], Oviya Sathiyanarayanan[c], Madelyn Uhm[a], Bob-Shen Yan[d], Brandon Yoo[b], Jiayi (Christina) Zheng[e], Suky Zheng[e], Timothy Zheng[ba]

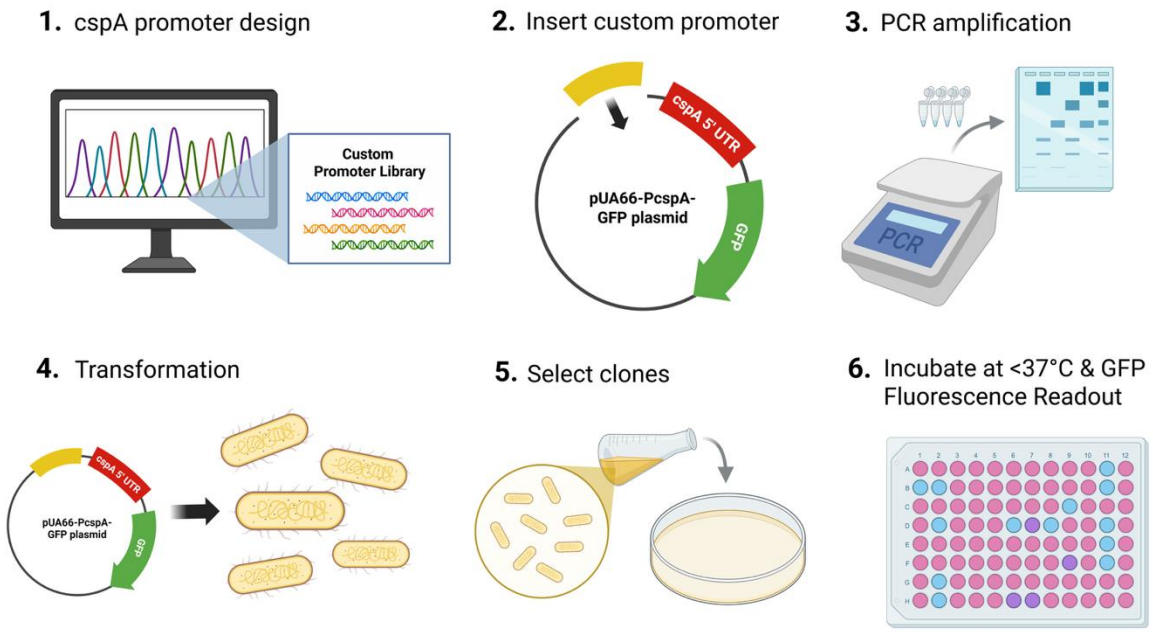
(Contributors have been alphabetized by last name and team roles indicated by superscript annotations.)

[a]*Wet Lab General Members*
[b]*Dry Lab General Members*
[c]*Wet Lab Leaders*
[d]*Dry Lab Leaders*
[e]*Team Leaders*
[f]*Mentor*



Abstract

Recombinant proteins have become one of the most impactful products of the biotechnology revolution. These proteins are typically produced in cell cultures grown at the physiological temperature of the cell line used, with the vast majority at 37ºC. However, decreasing this temperature can be exploited to increase production and improve the quality of these proteins. Present approaches to temperature downshifts focus on the benefits resulting from physiological changes to the cells themselves as opposed to the genetic elements regulating the recombinant protein's production. In this study, we report on the *cspA* cold-shock inducible expression system of *Escherichia coli* as a target for directed evolution to increase protein production at reduced temperatures. Using a reporter plasmid containing GFP under the control of the native *cspA* regulatory element, we identified and investigated 3 regions as targets. We used computational prediction of ribosome binding sites to identify potentially beneficial mutations in the upstream box of the 5' untranslated region and a predicted downstream box for error-prone PCR. Additionally, we investigated replacing the native promoter region to determine its viability as a region for directed evolution. Through this work, we provide an outline for using directed evolution as a method for modifying genetic regulation for increasing recombinant protein production at sub-physiological temperatures.



Introduction

Production of recombinant proteins has become one of the main driving forces behind modern biotechnology, with a global market worth over $271 billion.[1] These proteins have a vast array of applications, with biopharmaceuticals like insulin used for treating diabetes, to industrial enzymes like proteases used in laundry detergent.[2,3] As the demand for these proteins continues to grow, challenges in their production process will put strain on the industry's ability to match.

One of the most significant challenges is in production efficiency. For example, production efficiency for Chinese Hamster Ovary cells (the most common and industry standard cell line used for biopharmaceutical production) has only reached up to ~60% of the maximum predicted productivity through significant cell engineering efforts.[4] Furthermore, recombinant proteins are often quite sensitive, being susceptible to denaturation, misfolding, aggregation, and other problems that result in an unusable protein product and overall reduction in product yield and purity.[5] Therefore, improving the biomanufacturing processes for recombinant proteins has become a major area of research and investment.

One method commonly used in industrial bioprocesses for increasing protein production and quality has been through the use of hypothermic temperature shifts. These processes involve lowering cultures from typical growth temperatures of 37ºC to temperatures as low as 15ºC, resulting in increased production and quality of proteins in mammalian, yeast-based and bacterial production systems.[6,7] However, these interventions act globally on cell physiology, driving cells towards a production-focused phenotype as opposed to targeted effects on the genetic system containing the protein of interest, and can have undesirable effects such as reduced growth.

This led our team to investigate the *cspA* gene of *Escherichia coli* – a known cold-shock inducible protein involved with the cold-shock response.[8] The regulatory element of this gene contains a native promoter PcspA alongside a long 5' untranslated region (UTR) that acts as an RNA thermometer that while unstable at 37ºC, rapidly stabilized mRNA at lower temperatures. While *cspA* itself has been well-characterized, it has not been explored for use in the context of recombinant protein biomanufacturing. We hypothesized that the directed evolution of this system can be tuned to further increase the natural improvements in protein production resulting from growth at lower temperatures.

To test this hypothesis, we identified 3 regions of the *cspA* regulatory element for further investigation – the native PcspA promoter, the upstream box of the 5' UTR, and a suspected downstream box in the *cspA* protein coding sequence fragment. Using 2 promoters from the well-characterized Anderson promoter collection, we investigated the potential for replacing the PcspA promoter with a stronger one to increase production after induction at lower temperatures. Additionally, we developed an error-prone PCR (epPCR) workflow to induce mutations in the upstream and downstream boxes to investigate their potential in increasing production after induction. Using the reporter plasmid pUA66-PcspA-GFP in the study, we report the results of our initial investigation of this system and lay out a foundation for future study.

Results

1. Dry Lab

1.1 - Rationale for Regional Selection and Mutagenesis Strategy

The research team's directed evolution approach targeted the upstream and downstream boxes of the PcspA 5' UTR rather than the cold box region based on distinct regulatory mechanisms. The cold box functions primarily as a transient regulatory element, abolishing derepression when modified.[9] Given its distance from the Shine-Dalgarno sequence and its role in maintaining temperature-responsive regulation, mutations to the cold box risked creating constitutive expression that would eliminate the cold-shock response entirely—contrary to the objective of preserving cold-inducibility while enhancing expression efficiency.

In contrast, the upstream and downstream boxes have a direct influence on mRNA stability and translation efficiency. Literature demonstrates that upstream box modifications reduce *cspA* expression to less than 10% of wild-type during cold shock, while downstream box alterations abolish cold-shock induction entirely, indicating their essential roles in low-temperature translation.[10,11] These regions presented optimal targets for enhancing translational output while preserving temperature-responsive characteristics.

1.2 - Sequence Conservation Analysis

A comparative analysis of cold-shock proteins revealed distinct conservation patterns that informed the mutagenesis strategy. The upstream and downstream box sequences exhibited high conservation across *cspA, cspB, cspG,* and *cspI* homologs, particularly at consensus positions that maintain critical secondary structures.[10,11] Specific positions within the upstream box displayed specific unconserved variations between genes, suggesting these sites could tolerate mutations without completely disrupting function. As a result, mutagenesis was targeted specifically to unconserved positions in the UB and DB. This design preserved conserved core motifs while allowing sequence diversity at positions most likely to modulate ribosome accessibility. A 4^2 mini-library was generated and analyzed using the Salis Lab RBS Calculator, which models initiation rates based on nucleotides of UTR sequence upstream of the GFP start codon and the first nucleotides of the GFP coding sequence. Key outputs included the predicted translation initiation rate (TIR) and associated free energy terms (ΔG_total, ΔG_mRNA–rRNA, and ΔG_mRNA structural penalties).

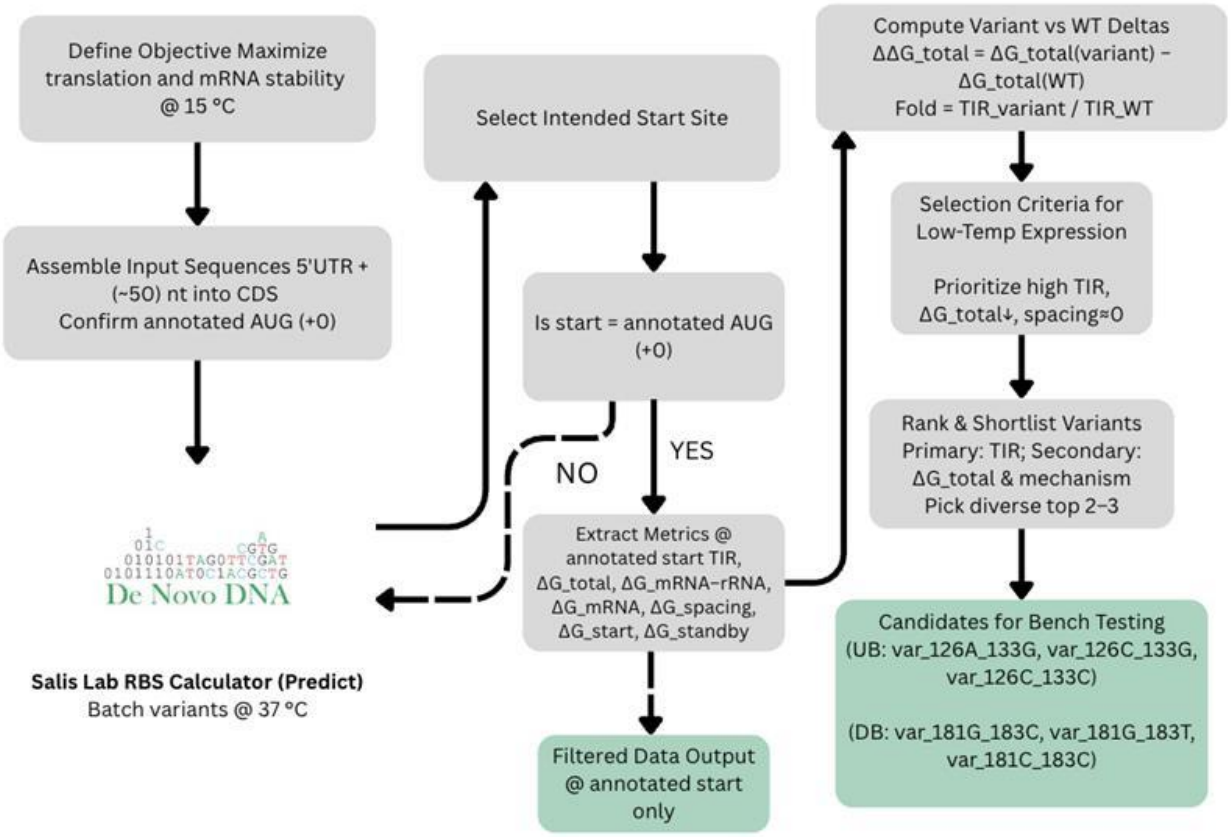


Figure 1: Workflow for selecting PcspA 5′UTR variants using the Salis Lab RBS Calculator (Predict). Input sequences (5′UTR + early CDS) were analyzed to identify the annotated start, extract TIR and ΔG components, compute variant-vs-WT differences, and rank candidates for predicted high translation at 15 °C.

1.3 - Computational Prediction of Variant Performance

The team analyzed variants in both the upstream and downstream boxes using the Salis Lab RBS Calculator to predict translation initiation rates (TIR). The wild-type sequence established the baseline for improvement. Analysis of sixteen upstream box variants revealed substantial improvements over wild-type. Two variants with C and A substitutions achieved maximum TIR of 2057.15, representing a 2.6-fold increase, while a double C variant reached maximum TIR of 1957.79, showing a 2.5-fold increase. A U/G combination variant showed maximum TIR of 1805.44, demonstrating a 2.3-fold increase over baseline.

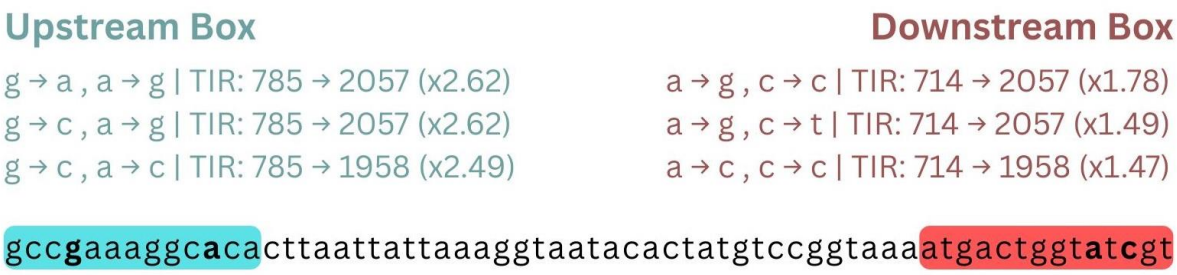


Figure 2: Upstream and Downstream Box Mutations in PcspA 5′UTR Enhance Translation Initiation. Computational predictions using the Salis Lab RBS Calculator (Predict) analyzing unconserved variations between genes sequence variants within the upstream and downstream boxes that significantly alter translation initiation rate (TIR). Variants were evaluated relative to the wild-type sequence.

Parallel analysis of downstream box modifications yielded more modest but still significant improvements. The top G/C variant achieved a maximum TIR of 1271.07, representing a 1.6-fold increase, while G/T and C/C variants reached TIR values above 1050, showing 1.3-fold increases. Several variants maintained TIR values between 700 and 900, remaining comparable to wild-type performance. These variants shared a common feature: reduced mRNA secondary structure stability, as indicated by ΔG_mRNA values, which suggested enhanced accessibility compared to the wild-type. The ΔG_total values, which inversely correlate with TIR, showed more negative values for high-performing variants, indicating stronger overall binding affinity.

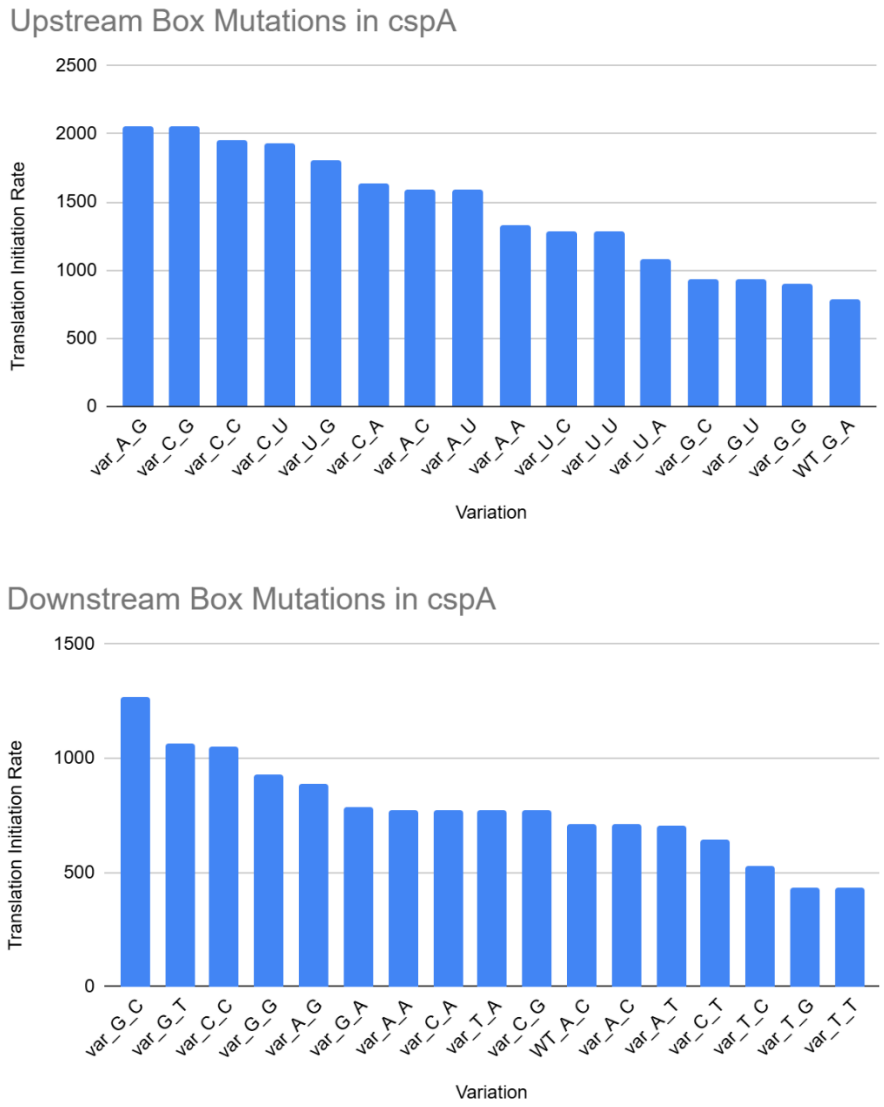


Figure 3: Upstream and downstream box mutations in the PcspA 5′UTR modulate translation initiation rates.

Computational predictions from the Salis Lab RBS Calculator (Predict) were used to assess how sequence variations within the upstream and downstream boxes of the PcspA 5′UTR alter the translation initiation rate (TIR). Both regions contain unconserved positions where substitutions can significantly enhance or reduce ribosome recruitment efficiency

1.4 - Structure-Function Relationships

Analysis of thermodynamic parameters revealed that improved variants primarily achieved higher TIR through optimization of mRNA-rRNA interactions while maintaining similar spacing and standby site accessibility. The strongest variants exhibited enhanced ΔG_mRNA-rRNA values, suggesting improved ribosome binding compared to wild-type. Notably, all high-

performing variants maintained consistent ΔG_start values, preserving start codon recognition fidelity.

The downstream box analysis yielded more modest improvements, with the best variant achieving a TIR of 1271.07, a 1.6-fold increase over wild-type. Compared to the best UB variant with a maximum TIR of 2057.15, the upstream box represents a more significant bottleneck for translation efficiency and a more promising target for optimization.

These computational predictions established a framework for experimental validation, identifying specific mutations predicted to enhance cold-inducible expression while maintaining regulatory control. The correlation between reduced secondary structure stability and increased TIR values supports the hypothesis that disrupting inhibitory RNA structures at sub-37ºC temperatures can substantially improve recombinant protein yields.

2. Wet Lab

2.1 - pUA66-PcspA-GFP Extraction

Upon receipt of the pUA66-PcspA-GFP plasmid from AddGene, individual colonies were generated by streaking on LB-agar plates supplemented with kanamycin. A negative control plate with no cells showed no colony growth, confirming sterility of media and reagents. 3 distinct colonies were selected and grown overnight in 3 mL of LB supplemented with kanamycin before miniprep using the GeneJET Plasmid Miniprep Kit following the manufacturers protocol. Duplicates of each miniprep were then analyzed using a NanoQuant plate in a Tecan M200 PRO microplate reader, with yields between 17.7 and 25.6 ng/µL, and OD260/280 ratios between 1.57 and 1.83.

Replicate Number	**OD260/280 Ratio**	**Yield (ng/µL)**
1a	1.78	18.2
1b	1.78	17.8
Sample 1 Average:	**1.78**	**18.0**
2a	1.83	25.6
2b	1.81	25.1
Sample 2 Average:	**1.82**	**25.4**
3a	1.57	17.7
3b	1.64	17.9
Sample 3 Average:	**1.61**	**17.8**

Table 1: Initial pUA66-PcspA-GFP DNA concentration and purity after miniprep. Three replicates of pUA66-PcspA-GFP plasmid cultures were miniprepped with the GeneJET Plasmid Miniprep Kit following overnight growth in LB supplemented with overnight. Samples were analyzed in duplicate using a NanoQuant plate in a Tecan M200 PRO microplate reader. Yields ranged from 17.7 to 25.6 ng/µL, with OD260/280 ratios between 1.57 and 1.83.

As recovered DNA mass and quality were on the lower end of the acceptable range, slight modifications were made to the manufacturers protocol to decrease any risk of wash carryover and increase elution efficiency. An additional colony from a freshly streaked plate was grown overnight, and miniprepped using the modified protocol. Triplicates were then analyzed using a NanoQuant plate in a Tecan M200 PRO microplate reader, with yield increasing to 100.4 ng/µL, and an OD260/280 ratio of 2.01. This sample was then used for all further experiments.

Replicate Number	**OD260/280 Ratio**	**Yield (ng/µL)**
1	2.04	98.5
2	2.00	103.7
3	2.00	99
Average:	2.01	100.4 ng/µL

Table 2: pUA66-PcspA-GFP DNA yield and purity after protocol modification. A second iteration of pUA66-PcspA-GFP plasmid extraction was done using a modified GeneJET Plasmid Miniprep protocol. Triplicate measurements showed higher DNA yield and purity than the previous protocol, with an average concentration of 100.4 ng/µL and OD260/280 of 2.01.

2.2 - PCR and Gel Electrophoresis

To identify our optimal PCR conditions, we then performed a screen of annealing temperatures and template DNA masses. PCR reactions were prepared with 250 pg, 500 pg, and 1 ng of miniprepped plasmid DNA, and annealed at 66.5, 65.9, 65.4, and 65ºC based on the predicted annealing temperature of 66ºC and the temperature gradient of our thermocycler. The resulting gel can be found below, indicating optimal amplification at an annealing temp of 65.9ºC and vector mass of 500 pg.

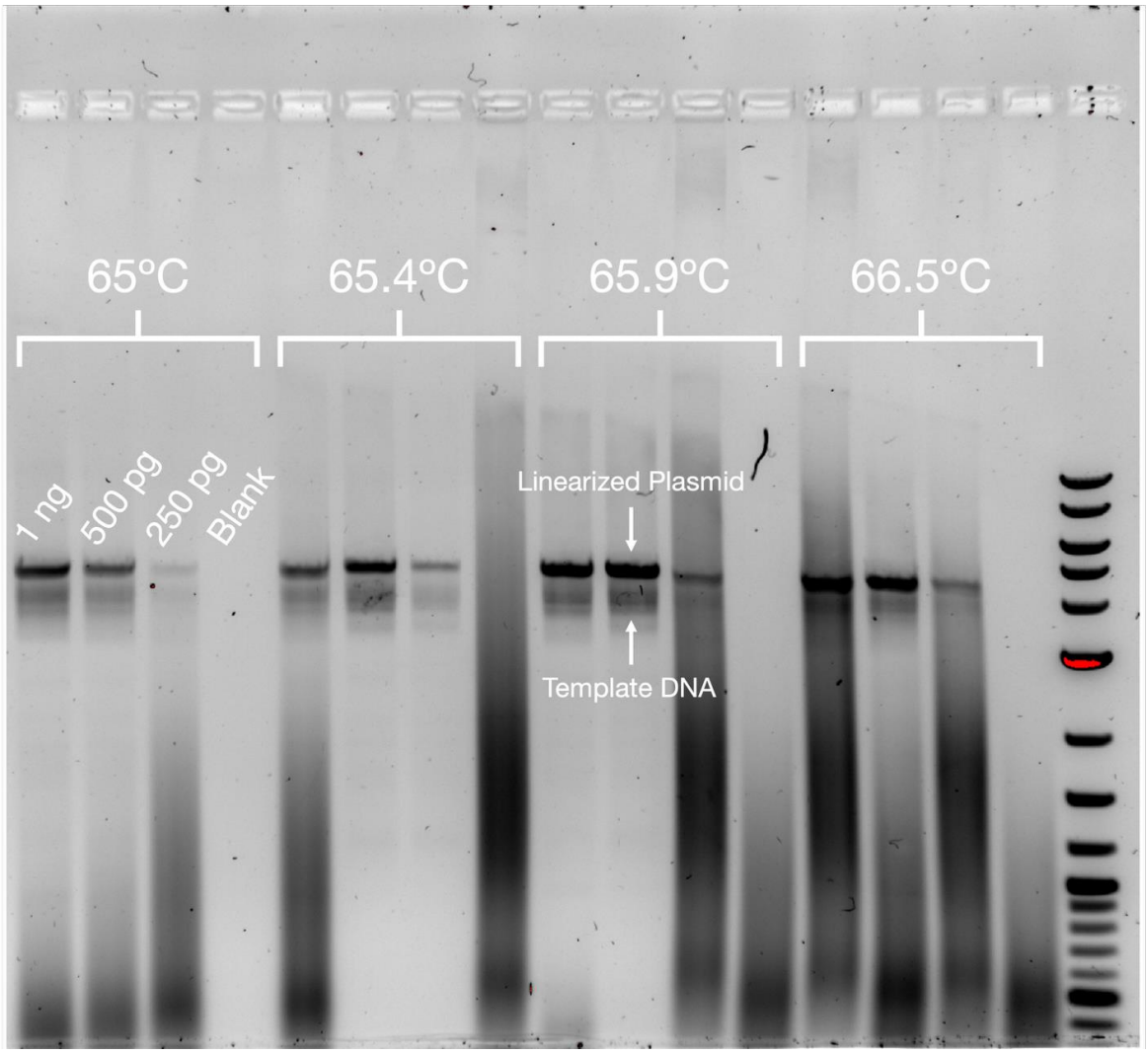


Figure 4: Optimization of PCR conditions for pUA66-PcspA-GFP linearization. A screening PCR was performed using annealing temperatures of 65, 65.4, 65.9, and 66.5ºC, and template masses of 1 ng, 500 pg, 250 pg, and no template controls. 65.9ºC and 250 pg were selected as the optimal conditions based on high product mass and low secondary product formation.

PCR was then performed using these conditions and treated with DpnI before PCR cleanup to remove any template DNA that could act as a false-positive in our subsequent transformations. A single replicate of this sample was then measuring NanoQuant plate to preserve sample volume, yielding 20.1 ng/µL of linearized DNA with an OD260/280 of 1.95.

2.3 – Transformation of pUA66-PcspA-GFP

Before attempting to transform our assembled constructs, we first validated our transformation protocols using purified pUA66-PcspA-GFP. 1 ng of purified plasmid was transformed with 50 µL of DH10B, and serial diluted 10x and 100x before plating and overnight growth. A control plate of untransformed DH10B was also prepared to confirm viability of the competent cells.

After incubation, colonies were only observed on the undiluted LB agar plate without antibiotics, confirming growth of the cells (Fig. xA). Unexpectedly, no colonies were observed on any selective LB plates and the positive control with untransformed DH10β cells (Fig. xH). Given the lack of growth in the control plate and individual colonies on the transformed plate without antibiotics, procedural errors were suspected to be the cause of failure, and the experiment repeated. The following experiment was successful, with colonies observed on all plates containing kanamycin, and bacterial lawns on all plates without antibiotics.

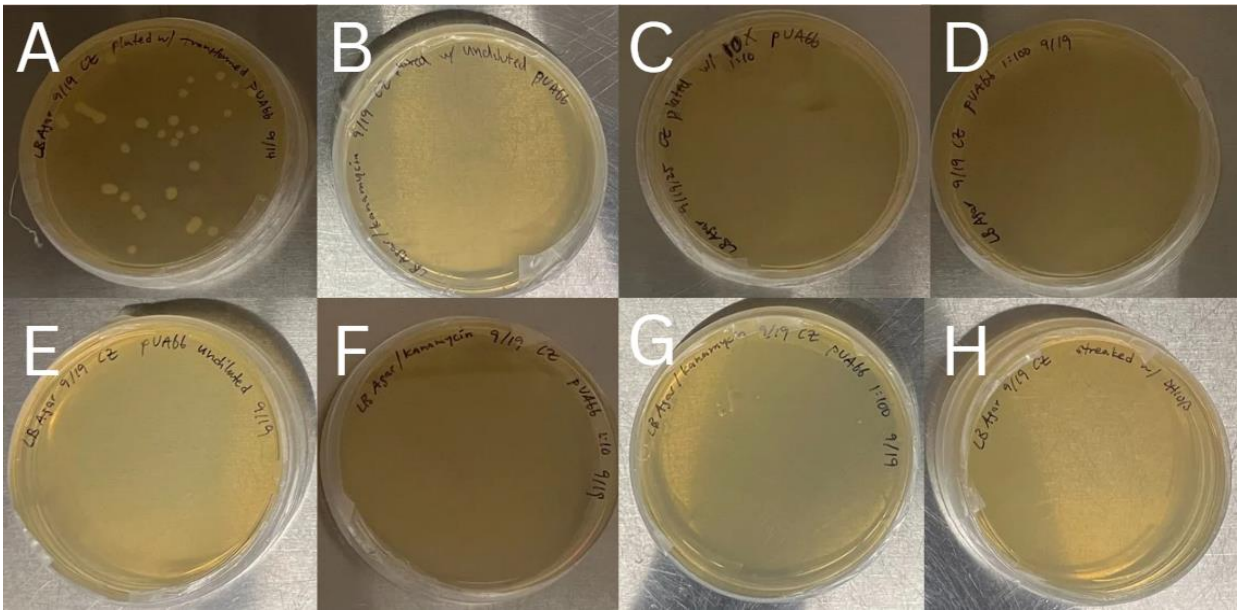


Figure 5: Colony growth following transformation of pUA66-PcspA-GFP into DH10B. Purified pUA66-PcspA-GFP was transformed into DH10B MAX competent cells and plated on LB agar plates without antibiotics (A, B, C, D) or with kanamycin (E, F, G). Serial dilutions of each culture were performed for each condition. Colonies were exclusively observed on the undiluted transformation plate without antibiotics (A). A control (H) was plated with untransformed cells produced no colonies.

2.4 - Assembly and Transformation

With the transformation protocol validated, we then progressed to insertion of our 2 promoters. As a proof-of-concept for replacing the PcspA promoter, we chose the promoters J23100 and J23119 from the Anderson promoter collection as our test promoters. These promoters were selected for two main reasons – their short length (35 base pairs) made it feasible to order them as dsDNA oligos as opposed to gene fragments, and their activity has been well-characterized.

To insert these promoters, we assembled the linearized plasmid with the promoters using both CPEC and NEBuilder to increase our chances for a successful reaction. 200 fmol of J23100 and J23119 were assembled with 50 ng of linearized plasmid, and 2 µL of each transformed into 50 µL of DH10B MAX E. coli, along with 1 ng of pUA66-PcspA-GFP to act as a positive control.

After overnight incubation, only the control plates without antibiotics and the plates transformed with pUA66-PcspA-GFP contained colonies. While this initially suggested an assembly failure, the low number of positive control colonies suggested a potential issue with transformation efficiency. We then repeated the assemblies and transformed using Stbl3, where we obtained significantly larger number of positive control colonies. However, there were still no colonies present on the plates containing the assembled plasmids, confirming a failure in DNA assembly. At this point, we were unable to continue experiments due to lab space availability and time constraints, leaving us unable to complete the experimental workflow.

Discussion

In this study, we investigated potential targets for directed evolution of the *cspA* regulatory element to increase protein production at sub-37ºC temperatures. We first identified and investigated 3 regions of the element - the native PcspA promoter, 5' UTR upstream box, and *cspA* downstream box. We selected 2 well-characterized promoters from the Anderson promoter collection to investigate replacing the PcspA promoter and used an RBS calculator to identify potential mutations in the upstream and downstream boxes to increase transcription. In the lab, we successfully purified and transformed our reporter plasmid pUA66-PcspA-GFP and excised the native PcspA promoter from the plasmid through PCR linearization. However, we failed to successfully insert either of the Anderson promoters into the plasmid using CPEC or NEBuilder and were unable to perform the epPCR workflow.

Our primary challenge in the lab occurred during assembly of our linearized plasmid with the 2 Anderson promoters to test their effect on the *cspA* element's function. One consideration that led to the selection of the 2 promoters was due to their short length, allowing them to be ordered as oligos relatively inexpensively compared to longer gene fragments. With the overlap sequences required for CPEC and NEBuilder, these oligos had a total length of 77 base pairs. While oligos this short have been successfully used with these techniques by other groups, we suspect that the extra challenge posed by inserting such a small fragment resulted in the failed assembly.

Another challenge that arose during the sequence design phase came about during primer generation for epPCR. The region between the upstream and downstream boxes is very AT rich (>70%), preventing us from designing primers to specifically target those regions for epPCR. As such, a compromise was developed to encompass both regions for epPCR, still allowing for mutations to be generated in both sections, though with less specificity. Though we were not able to perform the epPCR workflow, this would have significantly increased the screening required to generate the mutations of interest generated during dry lab work.

Future Directions and Applications

While we were unable to successfully execute it ourselves, we believe our main contribution from this work stems from the proposed outline and applications of directed evolution in cold-shock inducible expression systems. If PcspA is validated as a potential target for modification – indicated by differing transcriptional activity of GFP – libraries of natural and engineered promoters could be screened to identify promoters with the strongest transcriptional activity at reduced temperatures. Given the predicted beneficial effects of modifications to the upstream and downstream boxes, we believe directed evolution of these regions has the potential to generate beneficial mutants of PcspA. With the challenges of performing mutations throughout these elements, alternative methods such as CRISPR-based mutagenesis could instead be implemented for more flexibility than epPCR.[12] This approach could be also expanded to other regions of the *cspA* regulatory element, such as the RNA thermometer itself that are more challenging to modify due to their secondary structures. Furthermore, while we focused on *cspA*, several other cold-shock inducible elements exist in *E. coli* and other relevant protein production platforms such as CHO, and this same outline could be applied to them.

Screening for beneficial promoters or mutations has the potential to uncover several different functional benefits for protein production. Increased transcriptional activity (indicated by GFP expression) is the main metric for comparing these, as it shows the potential for increasing the overall yield of protein. Another additional metric we identified was the difference between activity at 37ºC and lowered temperatures, with promoters or mutations having a greater difference potentially allowing for tighter control of expression with different temperatures. We also hypothesized that the rate of transcriptional activation/inhibition when changing temperatures could be a source of improvement, though gains from this would likely be less significant given the time frame this occurs is very short relative to overall culturing time during manufacturing.

Another promising direction is the integration of machine learning (ML) models for predictive promoter engineering. By utilizing computational tools, experimental burden can be reduced through in silico prediction models of promoter strength, identification of promising mutations, and guided design prior to wet-lab testing. Sequence-to-activity data generated from the expanded promoter libraries, including both high-performing and negative samples, can train predictive models. Current approaches, such as Extended Vision Mutant Priority (EVMP) show improved prediction metrics by measuring mutation structures (k-mers) relative to base promoters.[13] Deep generative models (DGMs) and deep learning (DL) frameworks offer enhanced promoter identification and de novo design.[14] These models can consider features at the transcriptional, post-transcriptional, and translational level to deduce favourable traits (i.e., a high cold induction, low warm basal level, or favourable dynamic range). Iterative Design-Build-Test-Learn (DBTL) cycles, where each cycle of experimental data refines the model, improves accuracy and efficiency over time.

Long-term application of this work focuses on the industrial production of proteins, where it could be used to improve biomanufacturing processes for a variety of proteins, especially those susceptible to problems post-transcription at 37ºC. A notable example of this is monoclonal antibodies, which require several post-translational modifications for their activity, and are sensitive to aggregation in culture.[15] While temperature shifts alone are currently being explored to improve their production, use of a genetic element designed specifically for low temperature expression could be valuable to making it a viable process.

Materials and Methods

1. Bacterial Strains

The following chemically competent bacterial strains were used: DH10B MAX (Invitrogen) and Stbl3 (Invitrogen).[16,17] Transformations were performed following the manufacturers protocol.

2. Plasmid and Antibiotic

The plasmid pUA66 PcspA GFP No Linker was acquired from AddGene in DH5alpha as an agar stab and referred to in this study as pUA66-PcspA-GFP. As the plasmid contains the NeoR neomycin resistance gene, all cultures containing the plasmid or its assemblies were cultured with 50 µg/mL of kanamycin. pUA66 PcspA GFP No Linker was a gift from Pamela Silver (Addgene plasmid # 105605[18]; RRID:Addgene_105605[19]).

3. Bacterial Culture

3.1 – Plasmid Streaking

Upon receipt of the plasmid from AddGene, single colonies were obtained by streaking the punctured area of the stab on an LB agar plate supplemented with kanamycin. The plate was then incubated inverted overnight in a 37°C incubator.

3.2 – Liquid Culture

Single colonies of bacteria containing pUA66-PcspA-GFP were obtained from overnight plates, and individual colonies inoculated into 15 mL Falcon tubes containing LB supplemented with kanamycin. Cultures were grown loosely capped and sealed with parafilm overnight in a 37°C incubator and shaken at 250 rpm.

3.3 – Miniprep

All minipreps were performed using a GeneJET Plasmid Miniprep Kit (Thermo Scientific) according to the manufacturers protocol with the following modifications: The elution buffer was preheated to 60°C before use.[20] The elution buffer was incubated in the spin column for 5 minutes before centrifugation.

3.4 – NanoQuant

DNA quantity and quality was measured using the NanoQuant plate for the Tecan M200 PRO microplate reader according to the manufacturer protocol for dsDNA samples.[21]

4. Molecular Biology

4.1 – PCR Plasmid Linearization

Excision of PcspA and linearization of pUA66-PcspA-GFP was performed using NEB Q5 High-Fidelity 2x Master Mix in 20 µL reactions according to the manufacturers protocol.[22] Template DNA was assessed at varying concentrations from 250 pg to 1 ng, and annealing temperature was assessed from 65 – 66.5ºC, before selecting a final template DNA concentration of 500 pg and annealing temperature of 65.9ºC. The following thermocycling conditions were used for the final protocol:

Step	**Temperature (ºC)**	**Time**
Initial Denaturation	98	30 seconds
20 Cycles	98	10 seconds
	65.9	20 seconds
	72	135 seconds
Final Extension	72	120 seconds
Hold	10	-

4.2 – Gel Electrophoresis

Agarose gel electrophoresis was used to assess optimal PCR conditions. The gel was prepared using 1X TAE buffer (Sartorius), 1% agarose (Fisher), and 1X SYBR Safe (Invitrogen), and ran at 5V/cm for 5 hours to ensure separation of the linearized plasmid from the unlinearized template. A NEB 1kb+ DNA Ladder was run alongside for sizing reference. Imaging was performed post-run using a Bio-Rad ChemiDoc XR.

4.3 – DpnI Restriction Enzyme TreatmentDpnI was used to remove template pUA66-PcspA-GFP after PCR linearization.[23] 2 µL of NEB rCutSmart and 0.5 µL of NEB DpnI were added immediately post-PCR and incubated at 37ºC for 10 minutes before PCR cleanup.

4.4 – PCR Cleanup

PCR cleanup was performed using the Zymo DNA Clean & Concentrator according to the manufacturers protocol with the following modifications: Samples were centrifuged dry for an additional 3 minutes after the second wash step.[24] The elution buffer was preheated to 60°C before use. The elution buffer was incubated in the spin column for 5 minutes before centrifugation. Samples were eluted in 20 µL of elution buffer.

4.5 – Circular Polymerase Extension Cloning (CPEC)

CPEC was used to assemble our promoter oligos with our linearized pUA66-PcspA-GFP with NEB Q5 High-Fidelity 2X Master Mix in 20 µL reactions. 50 ng of linearized plasmid DNA was assembled with 200 fmol of each oligo, resulting in a vector:insert ratio of ~1:10. The following thermocycling conditions were used:

Step	**Temperature (ºC)**	**Time**
Initial Denaturation	98	30 seconds
Annealing	65.9	30 seconds
Extension	72	7 minutes

4.6 – NEBuilder

NEBuilder assembly was used to assemble our promoter oligos with our linearized pUA66-PcspA-GFP with each promoter using NEB HiFi DNA Assembly Master Mix.[25] 50 ng of linearized plasmid DNA was assembled with 200 fmol of each oligo, resulting in a vector:insert ratio of ~1:10. The assemblies were then incubated at 50ºC for 60 minutes in a thermocycler.

5. DNA Sequences

5.1 – Information

An annotated version of pUA66-PcspA-GFP can be found in Genbank format in the supplementary information. All primers were screened for primer-dimers and secondary structures using the IDT OligoAnalyzer tool.[26] Primer Tm's were determined using the NEB Tm Calculator.[27]

5.2 – PcspA Excision Primers

The following primers were designed to directly flank the PcspA promoter sequence in pUA66-PcspA-GFP and linearize the plasmid while excising PcspA. These promoter regions also serve as the overlap regions of the promoter inserts. The primers were ordered from IDT as a RxnReady primer pool and reconstituted to 100 µM in nuclease-free water upon receipt.

PcspA Excision Forward: acggtttgacgtacagaccat

PcspA Excision Reverse: catcgaggtgaagacgaaagg

5.3 – Promoter Inserts

Two promoters from the Anderson promoter collection were selected and their sequences retrieved from the iGEM Registry.[28] The promoter J23100 represents the reference promoter, while J23119 represents the consensus promoter sequence of the collection. The overlap region used by the PcspA excision primers was added to each promoter and acquired from IDT as a double-stranded DNA oligo and reconstituted to 100 µM in nuclease-free water upon receipt.

J23100 with Overlaps:
cctttcgtcttcacctcgatgttgacggctagctcagtcctaggtacagtgctagcacggtttgacgtacagaccat

J23119 with Overlaps:
cctttcgtcttcacctcgatgttgacagctagctcagtcctaggtataatgctagcacggtttgacgtacagaccat

5.4 – epPCR Primers

The following primers were designed using Primer3 to flank the region of *cspA* containing the upstream box in the 5' UTR and the downstream box in the *cspA* CDS scar upstream of the GFP CDS. The first set was designed to excise the target region in pUA66-PcspA-GFP while linearizing the plasmid. The second, reverse complementary set was designed for performing epPCR on the excised target region. The primers were ordered from IDT as 100 μM RxnReady primer pools in pH 8.0 IDTE.

epPCR Excision Forward: aacttttcactggagttgtccc

epPCR Excision Reverse: tggcgtgctttacagattttga

epPCR Forward: tcaaaatctgtaaagcacgcca

epPCR Reverse: gggacaactccagtgaaaagtt

6. RBS Calculator

Translation initiation for each *cspA* variant was estimated with the Salis Lab RBS Calculator (Predict, De Novo DNA).[29] For each sequence comprising the 5′ UTR and the first codons of the CDS, the tool scans candidate start codons and applies a thermodynamic model of ribosome–mRNA binding. Multiple energy terms are combined into a total binding free energy (ΔG_total), including standby site accessibility, Shine–Dalgarno to anti-SD pairing, SD-to-AUG spacing, start codon context, local mRNA structure, and stacking penalties. The translation initiation rate is reported on a proportional scale using a Boltzmann relationship to ΔG_total, with a per-term breakdown indicating the dominant contributions. Comparisons were restricted to the annotated start codon, and variants were ranked by the reported TIR together with their ΔG components.

References

1. Walsh G, Walsh E. Biopharmaceutical benchmarks 2022. Nat Biotechnol. 2022 Dec;40(12):1722–60.

2. Singh R, Kumar M, Mittal A, Mehta PK. Microbial enzymes: industrial progress in 21st century. 3 Biotech [Internet]. 2016 Aug 19 [cited 2025 Oct 1];6(2). Available from: https://doi.org/10.1007/s13205-016-0485-8

3. Kesik-Brodacka M. Progress in biopharmaceutical development. 2017 Oct 3 [cited 2025 Oct 1];65(3). Available from: https://iubmb.onlinelibrary.wiley.com/doi/10.1002/bab.1617

4. Hefzi H, Ang KS, Hanscho M, Bordbar A, Ruckerbauer D, Lakshmanan M, et al. A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. Cell Systems [Internet]. 2016 Nov 23 [cited 2025 Oct 1];3(5). Available from: https://www.sciencedirect.com/science/article/pii/S2405471216303635

5. Beygmoradi A, Homaei A, Hemmati R, Fernandes P. Recombinant protein expression: Challenges in production and folding related matters. International Journal of Biological Macromolecules [Internet]. 2023 Apr 1 [cited 2025 Oct 1];233. Available from: https://www.sciencedirect.com/science/article/pii/S0141813023002933

6. Huang CJ, Lin H, Yang X. Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. J Ind Microbiol Biotechnol [Internet]. 2012 Mar 1 [cited 2025 Oct 1];39(3). Available from: https://doi.org/10.1007/s10295-011-1082-9

7. Al-Fageeh MB, Smales CM. Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems. Biochem J [Internet]. 2006 Jun 28 [cited 2025 Oct 1];397(2). Available from: https://doi.org/10.1042/BJ20060166

8. Bae W, Jones PG, Inouye M. CspA, the major cold shock protein of Escherichia coli, negatively regulates its own gene expression. Journal of Bacteriology [Internet]. 1997 Nov [cited 2025 Oct 1]; Available from: https://journals.asm.org/doi/10.1128/jb.179.22.7081-7088.1997

9. Fang L, Hou Y, Inouye M. Role of the Cold-Box Region in the 5′ Untranslated Region of the cspA mRNA in Its Transient Expression at Low Temperature in Escherichia coli. 1998 Jan 1 [cited 2025 Oct 1]; Available from: https://journals.asm.org/doi/10.1128/jb.180.1.90-95.1998

10. Yamanaka K, Mitta M, Inouye M. Mutation Analysis of the 5′ Untranslated Region of the Cold Shock cspA mRNA of Escherichia coli. Journal of Bacteriology [Internet]. 1999 Oct 15 [cited 2025 Oct 1]; Available from: https://journals.asm.org/doi/10.1128/jb.181.20.6284-6291.1999

11. Mitta M, Fang L, Inouye M. Deletion analysis of cspA of Escherichia coli: requirement of the AT-rich UP element for cspA transcription and the downstream box in the coding region for its cold shock induction. [cited 2025 Oct 1]; Available from: https://onlinelibrary.wiley.com/doi/10.1046/j.1365-2958.1997.5771943.x

12. Maes S, Deploey N, Peelman F, Eyckerman S. Deep mutational scanning of proteins in mammalian cells. Cell Reports Methods [Internet]. 2023 Nov 20 [cited 2025 Oct 1];3(11). Available from: https://www.sciencedirect.com/science/article/pii/S2667237523003120

13. Yang W, Li D, Huang R. EVMP: enhancing machine learning models for synthetic promoter strength prediction by Extended Vision Mutant Priority framework. Front Microbiol [Internet]. 2023 Jul 5 [cited 2025 Oct 1];14. Available from: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2023.1215609/full

14. Gu Y, Su J, Xia J, Wu P, Wu H, Su Y, et al. De novo promoter design method based on deep generative and dynamic evolution algorithm. Nucleic Acids Res. 2025 Aug 28;53(16):gkaf833.

15. Sifniotis V, Cruz E, Eroglu B, Kayser V. Current Advancements in Addressing Key Challenges of Therapeutic Antibody Design, Manufacture, and Formulation. Antibodies [Internet]. 2019 Jun [cited 2025 Oct 1];8(2). Available from: https://www.mdpi.com/2073-4468/8/2/36

16. MAX Efficiency™ DH10B Competent Cells 5 x 200 µL | Buy Online | Invitrogen™ [Internet]. [cited 2025 Oct 1]. Available from: https://www.thermofisher.com/order/catalog/product/18297010

17. One Shot™ Stbl3™ Chemically Competent E. coli 21 x 50 µL/tube | Buy Online [Internet]. [cited 2025 Oct 1]. Available from: https://www.thermofisher.com/order/catalog/product/C737303

18. pUA66 PcspA GFP No Linker [Internet]. AddGene. Available from: https://www.addgene.org/105605/

19. Stirling F, Bitzan L, O'Keefe S, Redfield E, Oliver JWK, Way J, et al. Rational Design of Evolutionarily Stable Microbial Kill Switches. Molecular Cell. 2017 Nov 16;68(4):686-697.e3.

20. GeneJET Plasmid Miniprep Kit 50 Preps | Buy Online | Thermo Scientific™ | thermofisher.com [Internet]. [cited 2025 Oct 1]. Available from: https://www.thermofisher.com/order/catalog/product/K0502

21. AG TT. Plate reader Infinite 200 PRO [Internet]. [cited 2025 Oct 1]. Available from: https://lifesciences.tecan.com/plate_readers/infinite_200_pro

22. Q5® High-Fidelity 2X Master Mix | NEB [Internet]. [cited 2025 Oct 1]. Available from: https://www.neb.com/en-ca/products/m0492-q5-high-fidelity-2x-master-mix

23. DpnI | NEB [Internet]. [cited 2025 Oct 1]. Available from: https://www.neb.com/en-ca/products/r0176-dpni

24. DNA Clean & Concentrator-5 – Zymo Research International [Internet]. [cited 2025 Oct 1]. Available from: https://zymoresearch.eu/products/dna-clean-concentrator-5

25. NEBuilder® HiFi DNA Assembly Master Mix | Gene Assembly | NEB [Internet]. [cited 2025 Oct 1]. Available from: https://www.neb.com/en-ca/products/e2621-nebuilder-hifi-dna-assembly-master-mix

26. Primer Analysis Tool and Tm Calculator | IDT [Internet]. Integrated DNA Technologies. [cited 2025 Oct 1]. Available from: https://www.idtdna.com/pages/tools/oligoanalyzer

27. NEB Tm Calculator [Internet]. [cited 2025 Oct 1]. Available from: https://tmcalculator.neb.com/#!/main

28. Promoters/Catalog/Anderson - parts.igem.org [Internet]. [cited 2025 Oct 1]. Available from: https://parts.igem.org/Promoters/Catalog/Anderson

29. GeneticSystemsCalculator [Internet]. SalisLab. [cited 2025 Oct 1]. Available from: https://salislab.net/software/login