# Response Surface Methodology-Based High-Throughput BioLector Fermentation Screening of Rosetta-2 for Enhanced Production of Insulin Variant: Towards Improved Production Efficiency

Esra Ayan[1], Ali Özhan Aytekin[2], Ahmet Katı[3], and Hasan DeMirci[1]

[1]Koc Universitesi
[2]Yeditepe Universitesi
[3]University of Health Sciences

August 27, 2023

**Abstract**

Background: To obtain high yields of recombinant insulin and advancing therapeutic avenues for diabetes patients, the development of innovative designer insulin analogs have critical importance. The modified insulin analog presents a cost-effective remedy by being produced as inclusion bodies (IBs) within Escherichia coli BL21 (DE3) Rosetta-2 strain. This approach to production not only offers reduced production time but also yields high recovery rates. The prime aim of this investigation was to optimize the composition of the cultivation media, thereby accomplishing higher cell density fermentation of the proinsulin. Result: Various factors, including carbon and nitrogen sources, salts, metal ions, and pH, were systematically investigated through experimental screening using the BioLector multiwell bright plate. Additionally, computational analysis employing the Plackett-Burman Design within the Design Expert software was utilized to assess their effectiveness in terms of insulin concentration as a surrogate measure of insulin yield. Among the tested variables, glucose, glycerol, MgSO4, and lower Luria-Bertani mix concentration have a significant influence on insulin production, as determined by the screening experiment. Subsequently, the Central Composite Design approach was operated to further evaluate and optimize the precise levels of these influential variables. This systematic methodology achieved an optimized cultivation media formulation, resulting in a remarkable enhancement of insulin production, with levels reaching up to 13 mg/ml when applied in BioLector fermentation. Conclusion: The formulated cultivation media exhibited suitability for promoting high cell density fermentation of the modified insulin, thereby facilitating the attainment of optimal yields for the proinsulin expression.

## 1. Introduction:

The prevalence of diabetes in the United States has reached epidemic levels, with more than one-quarter of those aged 65 years and older suffer with the disease (*Stailey et al., 2017* ;*IDF Diabetes Atlas*). Aging is a primary risk factor for type 2 diabetes mellitus (T2DM) (*Ahima 2009* ), attributed to the concomitant increase in insulin resistance (*Ahima 2009* ;*Ayan and DeMirci 2023*) and decline in pancreatic beta-cell function (*Ahima 2009* ;*Ayan and DeMirci 2023*). T2DM in older adults may manifest with distinct clinical features, such as postprandial hyperglycemia (*Taylor et al., 2021*). However, the diagnosis of T2DM in older individuals may be complicated by the limitations of commonly used diagnostic tests, including hemoglobin A1C and fasting plasma glucose, which may fail to detect postprandial hyperglycemia (*Korkiakangas et al., 2009* ). Moreover, older adults with T2DM are predisposed to a heightened risk of developing microvascular and cardiovascular complications, further underscoring the need for effective management and monitoring of the disease in this population (*Vaidya et al., 2015*). The therapeutic management of T2DM in older

adults can pose limitations due to comorbidities, polypharmacy that is characterized by simultaneous use of multiple drugs, and renal or hepatic impairment (*Dardano et al., 2014* ). While oral hypoglycemic agents are usually the preferred choice for the initial treatment of T2DM, the progressive deterioration of pancreatic beta-cell function may eventually require insulin therapy (*Snyder et al., 2004*). Nevertheless, administering insulin in older adults necessitates a judicious approach as this population is particularly susceptible to hypoglycemia and unawareness of its occurrence (*Herman et al., 2005*).

Basal-bolus therapy requires the administration of basal insulin referred as intermediate and long-acting insulins, prandial insulin, and correction doses as needed to replace endogenous insulin (*Garg et al., 2010*). Prandial insulin is administered to simulate the physiological response of endogenous insulin to food intake, which involves a rapid and vigorous first phase followed by a more prolonged second phase of insulin secretion into the portal circulation (*Owens et al., 2001*). Recent advances in insulin therapy, such as the development of insulin analogs, including rapid-acting insulin analogs (RAIAs), have transformed the management of type 1 diabetes (T1DM). RAIAs, such as insulin Aspart, Glulisine, and Lispro, offer superior pharmacokinetic and pharmacodynamic profiles compared to regular human insulin (RHI) (*Ashwell et al., 2006*;*Hermansen et al., 2004*;*Becker et al., 2008*). RAIAs may confer notable clinical benefits over RHI, including a lower risk of hypoglycemia and improved glycemic control (*Garg et al., 2010*).

Various strategies have been employed to express recombinant insulin in*E. coli* , including using different expression methods, tags, and host strains (*Khalilvand et al., 2022* ). One efficient approach involves designing insulin constructs with a protease cleavage site using the home-made proteases to facilitate the cost-efficient production of mature insulin (*Akbarian and Yousefi, 2018* ). In addition, host strains can be engineered to enhance insulin expression, such as the use of Rosetta 2 host strains, which are BL21 derivatives optimized for the expression of eukaryotic proteins with rare codons (*Tegel et al., 2010* ). These strains carry a compatible chloramphenicol-resistant plasmid encoding tRNA genes for seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) under their native promoters (*Tegel et al., 2010* ). With IPTG induction, such strains are ideal for high-yield insulin production from target genes cloned into pET vectors (*Pan and Malcolm, 2018*).

During recombinant protein production, the chemical and nutritional composition of the culture medium can significantly impact host cell growth (*Shokri et al., 2003*). Factors such as carbon and nitrogen sources, metal ions, and medium pH play critical roles in determining cell growth, making it crucial to optimize the culture composition for high yields of the target protein (*Kusuma et al., 2019*;*Nikerel et al., 2006*). However, the sheer number of factors that can affect cell growth is vast, and evaluating their effects using the One-Factor-At-A-Time (OFAT) approach is time-consuming and labor-intensive (*Abu et al., 2017*). Furthermore, the OFAT approach must account for the potential dependent or independent effects and interactions between these factors (*Abu et al., 2017*). To address these limitations, the factorial approach offers a more comprehensive and efficient means of examining the impact of multiple factors on cell growth (*Abu et al., 2017*). By simultaneously evaluating all levels of all factors, this approach enables the determination of their independent effects and interactions (*Abu et al., 2017*). The Design of Experiment (DoE) is a statistical tool that employs fractional factorial models, such as Response Surface Methodology (RSM), to evaluate relevant interactions among variables via fewer experiments (*Papaneophytou and Kontopidis, 2014* ;*Sopyan et al., 2022*;*Dentener 2002*). Furthermore, the BioLector XT high-throughput microbioreactor offers real-time monitoring of crucial cultivation parameters for aerobes and anaerobes, such as biomass, pH, dissolved oxygen in the liquid phase (DO), and fluorescence (Drummen;*Osthege et al., 2022*). This tool provides rapid and detailed insights into bioprocess development experiments, allowing for more efficient optimization of the culture medium and, ultimately, higher target protein yields (*Osthege et al., 2022* ).

The objective of our study was to optimize the composition of the culture media to enhance the biomass of Rosetta 2 (DE3), which expresses a novel designer fast acting proinsulin. We employed the*Design-Expert* (DoE) method, utilizing the BioLector XT system to support our investigations. By employing these sophisticated approaches we sought to identify a more comprehensive and efficient analysis of the various factors affecting the growth and yield of the modified insulin from Rosetta 2 (DE3); thus, we have been

gain real-time insights into crucial cultivation parameters, enabling us to identify the optimal culture media composition to achieve maximum biomass production of the modified insulin. Accordingly, a screening experiment was conducted to assess the impact of various cultural components on the growth of Rosetta 2 (DE3), using the *Plackett-Burman Design* (PBD) and BioLector XT Microbioreactor. The optimal culture media composition was thereafter determined through further experimental design utilizing the BioLector XT Microbioreactor. The consequential significant factors were subsequently optimized, employing the Response Surface Methodology (RSM)*Central Composite Design* (CCD) of Design Expert suite. Following the completion of the optimization calculation, the obtained statistically relevant results were subsequently verified through experimental validation employing the BioLector XT Microbioreactor.

## 2. Result

### 2.1. Plackett–Burman design and BioLector growing conditions

The study assessed the effectiveness of eleven factors on bacterial growth by designing a series of experiments employing *Design of Expert* (Stat Ease Inc.) 7.0.0 software, resulting in 16 trial configurations (Supplementary Table 1). This study aimed to scrutinize the Rosetta 2 (DE3) the modified insulin strain's growth kinetics utilizing the BioLector (mp2-labs, Baesweiler, Germany) system. Two distinct concentrations of LB broth media, precisely 30 g/L or 50 g/L, were utilized, and a comprehensive DoE-PBD with 16 parameters was executed in triplicate on a single plate (Fig. 1). The cells were then inoculated for 4 h in each well, and the experiments were conducted under constant agitation (800 rpm) at 37°C in 48-well FlowerPlates (mp2-labs, Baesweiler, Germany) with a working volume of 1000 µL. This study continuously monitored scattered light intensities to observe growth kinetics in real-time (Fig. 1). Furthermore, the induction with Isopropyl β-D-1-thiogalactopyranoside (IPTG) was performed in each well at distinct final concentrations (0.2 mM, 0.3 mM, 0.4 mM) compliance with DoE-PBD parameters, and was conducted overnight at 37°C. The final insulin density, measured in grams per liter by UV/Vis spectrophotometer (280 nm) and analyzed by SDS-PAGE, has been considered the response variable and recorded in the PBD table (Supplementary Table 1; Fig. 1; Supplementary Fig. 1). The statistical inference conducted on the data revealed that the model under consideration attained statistical significance, as deduced from a $p$ -value of 0.05 and an R2 coefficient of determination of 85.37% (Table 1). Notably, the model terms, encompassing $MgSO_4$, glycerol, glucose, and LB broth concentration were established to be influential factors, with p-values that fell below the conventional threshold of 0.05 for statistical significance. This observation suggests that these terms exerted a substantial effect on the response variable. Moreover, the F-statistic was employed to measure the magnitude of association between each term and the response variable, where the outcomes showed that higher F-values corresponded to stronger associations.

The half-normal plot is a graphical representation of the distribution of the residuals in a regression model, which allows for the examination of the normality assumption of the residuals*(Zahn et al., 1975)* . In Fig. 2, a linear pattern in the half-normal plot indicates that the residuals conform to a normal distribution. Conversely, any deviation from linearity in the plot suggests that the residuals deviate from normality, implying that the normality assumption has been violated. Following the standardized effect, the half-normal plot depicts the relative significance of the independent variables in a regression model, where the percentage probability of the standardized effect for the $MgSO_4$, glycerol, glucose, and LB broth concentration values are more substantial than those of the other variables (*Fig. 2*).

A *Pareto* chart also serves as a graphical representation of the standardized effect exerted by each independent variable on the response variable (*Kenett 1991* ). The effectiveness of the factors is gauged by employing two statistical thresholds, namely the*Benforrini* limit and t-value limit, both of which are ascertained at 3,898 and 2.306, respectively, for a significance level (α) of 0.05. The *Pareto* chart elucidates the magnitude of the standardized effects via bars, with the four bars exhibiting the most prominent values corresponding to $MgSO_4$, glycerol, glucose, and LB broth concentration, thereby attesting to their paramount significance in influencing the response. Consequently, these four variables were shortlisted for the optimization of the experimental design by applying the CCD approach, a popular RSM technique (*Fig.3*).

The perturbation plot serves as a valuable instrument for discerning and juxtaposing the effects of multiple factors within a given point in the design space (*Bonnans and Shapiro, 2013* ). It facilitates the visualization of the response by systematically varying a singular factor across its entire range while upholding the constancy of all other factors (*Bonnans and Shapiro, 2013* ). A comprehensive elucidation of the influences exerted by the immediate and interactive effects of independent variables on insulin yield was attained through the utilization of perturbation plots. The perturbation plot (Fig. 4) visually represents the direct effects of variables C, D, E, and H on insulin yield.

### 2.2. Optimizing BioLector fermentation screening through central composite design of response surface methodology

The DoE packagefacilitated the generation of 30 experiments using the RSM to optimize carefully selected model terms, encompassing $MgSO_4$, glycerol, glucose, and certain LB broth concentration (referred as TB). The experimental runs were conducted utilizing the BioLector microfermenter in a 1 mL screening culture that maintained a constant concentration of 10 mM KCI, 15 mM $MgCI_2$, 15 mM $KH_2PO_4$, 10 mM thiamine, and 50 g/L 2X LB Miller media (Supplementary Table. 2 Fig. 5), complemented by 0.25 mM IPTG tailored to each specific design point. Once conducting analyses using various models, the quadratic model emerged as a suitable candidate for predicting and validating optimal conditions. Notably, the model exhibited a significant p-value of 0.0037, thereby providing a reliable prediction framework. Additionally, the model's lack of fit displayed insignificance at 0.4479 relative to the pure error, suggesting the error's immateriality in influencing the accuracy of the proposed model (Table 2). The model's R2 value of 0.6866 indicated that it reasonably fits the experimental data. Furthermore, the adequate precision value of 7.861 indicates the sufficiency of the signal for the biological system under scrutiny. The perturbation analysis, factorial contour plots and 3D response surface of the central composite design was operated to gain a deeper understanding of the impact of independent variables' major effects and interactions on insulin yield (Fig. 6, Fig. 7, and Fig. 8). Perturbation of CCD analysis specifically clarified the primary marks of variables A, B, C, and D, on insulin yield, as depicted in Fig. 6. The two-dimensional contour and three-dimensional response surface plot effectively portrayed the intricate relationship between the independent variables and the ensuing insulin yield (Fig. 7, Fig.8). These independent variables were deliberately manipulated across a predefined range to systematically investigate their combined influence on insulin production. The contour plots (Fig.7), in particular, offer invaluable insights into the inherent characteristics and magnitude of the interactions among various contributing factors. Notably, when glycerol levels were elevated, there was a substantial reduction in the resultant insulin yield, plummeting from 12.6259 mg/ml to 7.25534 mg/ml. Likewise, the simultaneous increase of glycerol and media content to elevated levels (exceeding 30 g/l) resulted in a decrease in insulin yield from 11.124 mg/ml to 6.84772 mg/ml. Conversely, the presence or absence of $MgSO_4$exhibited no discernible impact on the insulin yield of significance. The investigation of the interplay between the dependent and independent variables was further elucidated by implementing 3D response surface plots (Fig. 8). These plots effectively affirmed the intricate and interactive influence of specific glucose, glycerol, $MgSO_4$ concentrations, and media content on insulin yield. Notably, when considering the actual factors of LB (30 g/l) and $MgSO_4$ (5mM), it was regarded that a lower concentration of glycerol and the highest concentration of glucose displayed a positive correlation with the production of insulin at higher levels compared to the other variables. Consequently, based on these findings, four distinct scenarios were statistically derived and subsequently subjected to experimental validation. The experimentation was achieved using the BioLector micro-fermenter, with each scenario being replicated three times to ensure the reliability and accuracy of the results (Table 3). Contempt the absence of substantial discrepancies among the scenarios, it is noteworthy that scenario III, characterized by a glucose concentration of 8.78 mM, glycerol concentration of 10 g/L, LB of 30 g/l, and $MgSO_4$ concentration of 15 mM, exhibits the slightest deviation and most heightened insulin product in comparison to the additional scenarios (Table 3).

## 3. Discussion

*E. coli* is a prominent host organism for the production of recombinant proteins, primarily owing to its advantageous characteristics encompassing expeditious growth, facile genetic manipulations, and commendable

rates of recombinant protein synthesis (*Wang and Li, 2014*). Notably, *E. coli* emerged as the pioneering expression host for the manufacture of human insulin in 1982 (*Baeshen et al., 2015* ). However, the inherent propensity of insulin expression in *E. coli* to aggregate culminates in the formation of inclusion bodies (IBs) (*Baeshen et al., 2016* ). Furthermore, proinsulin conversion to mature insulin necessitates miscellaneous enzymatic cleavage at a precisely timed interval, manifesting a concomitant reduction in insulin yield (*Kemmler et al., 1971*). In this study, we present a novel designer insulin variant purposefully engineered to exhibit enhanced cleavability via trypsin. The focal objective of our investigation revolves around optimizing the compositional makeup of the culture media, thereby potentiating the augmentation of the modified insulin's growth rate and biomass yield.

Much like other natural processes, the growth rate of bacteria is subject to the influence of numerous contributing parameters. The identification and optimization of these factors present substantial challenges from a financial and economic standpoint (*Packiam et al., 2020* ). Considering these challenges, the combination of experimental BioLector micro-bioreactors and computational DoE methodology offers a robust approach for bioprocesses' experimental and statistical optimization. This combinatorial approach facilitates attaining superior outcomes while minimizing the expenditure of time and resources (*Elibol, 2004* ). Several studies have effectively employed DoE methods to augment the yield of recombinant protein expression in *E. coli* by optimizing culture media (*Sunitha et al., 1999*;*Shahbazmohammadi and Omidinia, 2017*;*Zare et al., 2019*;*Duan et al., 2020*). The existing body of literature confines similar investigations encompassing diverse protein types, design methodologies, factors under evaluation, and the consequent optimization outcomes (*Sunitha et al., 1999*;*Shahbazmohammadi and Omidinia, 2017*;*Zare et al., 2019*;*Duan et al., 2020*). Notable instances include optimizing culture media for producing L-Asparaginase, Phytase, Streptokinase, and Reteplase in*E. coli* , employing DoE-based strategies that have enhanced production yields (*Sunitha et al., 1999*;*Kenari et al., 2011*;*Ghoshoon et al., 2011*). A comprehensive review of the literature indicates the presence of numerous variables capable of influencing bacterial growth rate and biomass production (*Sunitha et al., 1999*;*Shahbazmohammadi and Omidinia, 2017*;*Zare et al., 2019*;*Duan et al., 2020*;*Kenari et al., 2011*;*Ghoshoon et al., 2011*). Parameters such as the nature and concentration of carbon and nitrogen sources, pH conditions, and the inclusion of trace elements have received considerable attention as key factors.

Furthermore, several studies have centered around utilizing the 48-well BioLector microbioreactor system (Osthege et al., 2022;*Sparviero et al., 2023*;*Flitsch et al., 2016*;*Lennen et al., 2016* ). For instance, this system has proven effective in the real-time monitoring of lipid production and the tracking of growth in *Y. lipolytica*(*Back et al., 2016*). The literature also showcases the result transferability and comparability assessment between the BioLector system and fully controlled bioreactor systems operating in fed-batch mode, specifically at moderate to high cell densities (*Toeroek et al., 2015* ). Additionally, various investigations illustrate the efficacy of the BioLector system in screening optimal growth conditions and engineered strains prior to scaling up (*Kensy et al., 2009* ). This robust cultivation protocol conducted on microtiter plates enables the screening of *E. coli* systems under conditions closely resembling lab-scale bioreactor cultivations (*Kensy et al., 2009* ).

In this study, the PBD was operated to investigate the effects of eleven factors on cellular growth. These factors encompass diverse nitrogen (N) and carbon (C) sources, salts, metal ions, pH, and the buffering system. Among these factors, the lower concentration of LB (Luria-Bertani) medium (30 g/L) and, in contrast to prior investigations, the $MgSO_4$, glucose, and glycerol levels exerted the most pronounced effect on cellular growth. Consequently, these factors were singled out for subsequent optimization by implementing the RSM and the CCD. Moreover, 0.25 mM IPTG exhibited advantageous effects on cellular growth. Model terms displaying non-significant p-values were deliberately excluded from the culture media during the CCD experiments. According to the outcomes derived from RSM, glycerol concentration emerged as the most influential factor, exhibiting a direct relationship with heightened cellular proliferation. These findings concur with prior investigations that have likewise underscored the favorable impact of glycerol on *E. coli* growth, notably by fostering anaerobic fermentation (*Dharmadi et al.,2006* ). When juxtaposing the coexistence of glycerol as a carbon source and IPTG as an inducer (*Malakar and Venkatesh, 2012*), our findings establish

that the optimal glycerol concentration, in the presence of IPTG is a mere 1% (v/v). This observation aligns with studies evidencing a decline in cellular yield as the glycerol concentration escalates (*Malakar and Venkatesh, 2012*). Moreover, the fitting of the Hill equation has been validated as yielding a prototypical Monod-type expression for growth on glycerol, both in the presence and absence of IPTG, as corroborated by previous scholarly works (*Malakar and Venkatesh, 2012*). On the other hand, glucose and $MgSO_4$ have garnered recognition as pivotal factors exerting a profound influence on *E. coli* growth in extant literature (*Izaki and Arima, 1965* ). However, the constructive ramifications of limited glycerol presence on growth may also be ascribed to the deleterious consequences of glucose, acting as a carbon source, under conditions of meager nitrogen availability (*Bren et al., 2016*;*Shiloach et al., 1996*;*Michaels et al.,1983*).

The gram-negative bacterium *E. coli* is widely regarded as the preferred host organism for the heterologous expression of various recombinant proteins. This selection is primarily attributed to its cultivation's facile nature, the culture media's cost-effectiveness, and the potential for achieving high product titers (*Kopp et al., 2017*). However, using harsh induction strategies involving IPTG as an inducer often triggers stress responses, giving rise to the phenomenon known as "metabolic" or "product burden." These stress reactions are characterized by diminished growth rates and cellular lysis during prolonged induction. Alternative approaches have been explored to mitigate these challenges, emphasizing "gentle" or "modifiable" induction techniques employing lactose as an alternative inducer (*Kopp et al., 2017*). This approach aims to alleviate the strain exerted on the production host. In contrast, conventional induction methods using glucose as the primary carbon source and lactose can lead to catabolite repression effects on lactose uptake kinetics, ultimately compromising the overall product yield. Conversely, glycerol, an alternate carbon source, has demonstrated promising outcomes when combined with glucose and lactose in auto-induction systems (*Kopp et al., 2017*). Glycerol has been shown to exhibit favorable effects on recombinant protein production, exhibiting reduced signs of catabolite repression during co-cultivation with lactose. Furthermore, an investigation into the mechanistic aspects of glycerol uptake in the presence of lactose as an inducer highlighted significantly enhanced inducer uptake rates in product-producing strains compared to non-producing strains. These findings underscore glycerol's practical and less disruptive nature on cellular viability and recombinant protein productivity compared to glucose (*Kopp et al., 2017*).

Moreover, the metabolic breakdown of the carbon source results in the accumulation of acidic by-products, mainly acetate, within the culture medium (*Kusuma et al., 2019*;*Ukkonen et al., 2011*). These acidic conditions can significantly impede cell growth and compromise recombinant protein production (*Kusuma et al., 2019*;*Ukkonen et al., 2011*). To counteract this issue, the supplementation of yeast extract and tryptone in the culture medium serves to mitigate medium acidification, primarily by counterbalancing the elevated ammonia levels generated during their metabolic utilization (*Kusuma et al., 2019*;*Ukkonen et al., 2011*). Among the various scenarios examined, Scenario-3 represents an optimal formulation of the culture medium, characterized by an optimal combination of yeast (0.5 g/L), tryptone (1 g/L), and salt (1 g/L) concentrations, in addition to a comparatively lower glucose concentration (8.8 mM). This particular configuration may elicit enhanced cell growth and delay entry into the death phase, thereby signifying its suitability for promoting favorable outcomes regarding cellular dynamics and extended cell viability.

## 4. Conclusion

This study proposes a novel formulation for the high-throughput cultivation medium of modified insulin, combining experimental methodologies using the BioLector microbioreactor and computational approaches employing DoE software. The efficacy of each step was validated by utilizing high-throughput PBD and CCD within the DoE software framework, complemented by the screening capabilities of the BioLector micro-bioreactor, enabling comprehensive assessments in a complex batch environment. The investigation primarily focused on Scenario-3, which exhibited superior attributes and demonstrated its potential as a more streamlined and cost-effective medium for achieving high cell density fermentation of the modified insulin in the *E. coli* Rosetta-2 strain.

## 5. Methods

5.1 Microorganism, culture media, chemicals, BioLector and Software

Utilized the *E. coli* BL21 Rosetta 2 strain that had been transformed with the pET28a(+) vector containing the the modified insulin construct, featured a cleavable N-terminal 6XHis-tag. To assess the growth and expression of the the modified insulin clone, we employed Luria-Bertani (LB) Miller media as the foundational culture medium. The necessary chemical components were procured from reputable suppliers such as Merck and Sigma. Our internal resources provided a protein weight marker known as KUYBIIGM-Ladder. We utilized *Design-Expert 7.0.0* software package (Stat-Ease, Inc., based in Minneapolis, MN, USA) to design and analyze optimization experiments. We employed the BioLector XT Microbioreactor (Brandd City, Country) to conduct screening experiments, considering statistically determined parameters.

5.2 Media optimization

Firstly, a preliminary screening of various factors was performed using the Plackett-Burman Design (PBD) method to identify the most influential variables. Subsequently, the Central Composite Design (CCD) method of Response Surface Methodology (RSM) was employed to optimize the levels of these significant variables. All experimental procedures were conducted in 2 mL volume wells within the Biolector XT Microbioreactor, each containing 1 mL of the respective culture media. Prior to the experiments, the media were prepared according to the specific design points and then inoculated with an overnight seed culture. The inoculated cells grew for 4 hours in each well, with constant agitation at 800 rpm, and maintained at 37°C. To facilitate these experiments, 48-well FlowerPlates (mp2-labs, Baesweiler, Germany) was used, with a working volume of 1000 µL. Throughout the study, real-time monitoring of growth kinetics was conducted by continuously measuring scattered light intensities. Additionally, the final insulin density, expressed in grams per liter, was determined using a UV/Vis spectrophotometer at a wavelength of 280 nm. To further analyze the results, SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) was performed.

*5.2.1. Screening of factors using the Plackett-Burman factorial design approach*

A total of eleven factors were investigated, including the concentrations of different nitrogen (N) and carbon (C) sources, pH, the concentration of LB media, as well as the concentrations of salts and metal ions (Supplementary Table 1). To design the experimental layout, a Plackett-Burman factorial design consisting of sixteen experiments was employed using DoE software (Supplementary Table. 1). Following the execution of the experiments via BioLector, statistical analyses were conducted on the obtained responses. The results, including model validation parameters and the significance values of the variables, were presented in ANOVA (Analysis of Variance) and fit statistic tables. Variables with significant effects (p-values $< 0.05$) were identified based on the ANOVA table, as well as through the examination of the Half-normal plot, Pareto chart, and Perturbation chart of standardized effects. These analytical techniques were employed to assess the significance and influence of each variable, facilitating the identification of key factors that significantly impacted the observed responses in the experiment.

*5.2.2. Optimizing factors using response surface methodology*

Based on PBD results, four of the most influential factors were selected for further optimization by CCD in DoE software, resulting in thirty experimental runs (Supplementary table 2). The specified composition of model terms (chosen factors) was employed to prepare culture media for each run according to the designed points via BioLector. Besides, constant values of other media components not in the model were supplemented in the media according to the insulin results (mg/ml) in PBD design and SDS-PAGE results. The concentration corresponding to the highest response for less significant variables and -1 level of insignificant factors were supplied. Following the execution of experiments via BioLector, responses were scrutinized via different models. The best model was picked based on model validation parameters documented in the ANOVA table, fit statistic tables, and diagnostic analysis. DoE software generated the perturbation chart, and the effect of each significant independent and dependent variable on response was registered graphically via contour and 3D plots. Finally, DoE software generated four scenarios about optimal points based on the acquired regression equation. Predicted design scenarios with the highest desirability were dissected and compared to each other using Biolector XT microbioreactor in triplicates again. The suggested optimal

media was defined as Scenario-3 (13 mg/ml) with a 3.3 deviation and is ready for scale-up optimization.

**Acknowledgments**

The authors dedicate this manuscript to the memory of Albert E. Dahlberg and Nizar Turker. The authors gratefully acknowledge the use of the services and facilities of the University of Health Science-Validebag DETAUM (Experimental Medicine Research and Application Center).

**REFERENCES**

Stailey, Matthew; Conway, SusanE. (2017). Review of the Next Generation of Long-Acting Basal Insulins: Insulin Degludec and Insulin Glargine. The Consultant Pharmacist, 32(1), 42–46. doi:10.4140/TCP.n.2017.42

Ahima, R. S. (2009). Connecting obesity, aging and diabetes. Nature medicine, 15(9), 996-997.

Ayan, E., & DeMirci, H. (2023). A brief atlas of insulin. Current Diabetes Reviews, 19(6), 18-79.

Taylor, S. I., Yazdi, Z. S., & Beitelshees, A. L. (2021). Pharmacological treatment of hyperglycemia in type 2 diabetes. The Journal of Clinical Investigation, 131(2).

Korkiakangas, E. E., Alahuhta, M. A., & Laitinen, J. H. (2009). Barriers to regular exercise among adults at high risk or diagnosed with type 2 diabetes: a systematic review. Health promotion international, 24(4), 416-427.

Vaidya, V., Gangan, N., & Sheehan, J. (2015). Impact of cardiovascular complications among patients with Type 2 diabetes mellitus: a systematic review. Expert review of pharmacoeconomics & outcomes research, 15(3), 487-497.

Dardano, A., Penno, G., Del Prato, S., & Miccoli, R. (2014). Optimal therapy of type 2 diabetes: a controversial challenge. Aging (Albany NY), 6(3), 187.

Snyder, R. W., & Berns, J. S. (2004, September). Reviews: use of insulin and oral hypoglycemic medications in patients with diabetes mellitus and advanced kidney disease. In Seminars in dialysis (Vol. 17, No. 5, pp. 365-370). Oxford, UK: Blackwell Science Inc.

Herman, W. H., Ilag, L. L., Johnson, S. L., Martin, C. L., Sinding, J., Al Harthi, A., ... & Raskin, P. (2005). A clinical trial of continuous subcutaneous insulin infusion versus multiple daily injections in older adults with type 2 diabetes. Diabetes care, 28(7), 1568-1573.

Garg, Satish; Ampudia-Blasco, Francisco; Pfohl, Martin (2010). Rapid-Acting Insulin Analogues in Basal-Bolus Regimens in Type 1 Diabetes Mellitus. Endocrine Practice, 16(3), 486–505. doi:10.4158/EP09294.RA

Owens, D. R., Zinman, B., & Bolli, G. B. (2001). Insulins today and beyond. The Lancet, 358(9283), 739-746.

Ashwell, S. G., Amiel, S. A., Bilous, R. W., Dashora, U., Heller, S. R., Hepburn, D. A., ... & Home, P. D. (2006). Improved glycaemic control with insulin glargine plus insulin lispro: a multicentre, randomized, cross-over trial in people with Type 1 diabetes. Diabetic Medicine, 23(3), 285-292.

Hermansen, K., Fontaine, P., Kukolja, K. K., Peterkova, V., Leth, G., & Gall, M. A. (2004). Insulin analogues (insulin detemir and insulin aspart) versus traditional human insulins (NPH insulin and regular human insulin) in basal-bolus therapy for patients with type 1 diabetes. Diabetologia, 47, 622-629.

Becker, R. H., & Frick, A. D. (2008). Clinical pharmacokinetics and pharmacodynamics of insulin glulisine. Clinical pharmacokinetics, 47, 7-20.

Khalilvand, A. B., Aminzadeh, S., Sanati, M. H., & Mahboudi, F. (2022). Media optimization for SHuffle T7 Escherichia coli expressing SUMO-Lispro proinsulin by response surface methodology. BMC biotechnology, 22(1), 1-13.

Akbarian, M., & Yousefi, R. (2018). Human αB-crystallin as fusion protein and molecular chaperone increases the expression and folding efficiency of recombinant insulin. PLoS One, 13(10), e0206169.

Tegel, H., Tourle, S., Ottosson, J., & Persson, A. (2010). Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta (DE3). Protein expression and purification, 69(2), 159-167.

Pan, S. H., & Malcolm, B. A. (2000). Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). Biotechniques, 29(6), 1234-1238.

Shokri, A., Sandén, A., & Larsson, G. (2003). Cell and process design for targeting of recombinant protein into the culture medium of Escherichia coli. Applied microbiology and biotechnology, 60, 654-664.

Kusuma, S. A. F., Parwati, I., Rostinawati, T., Yusuf, M., Fadhlillah, M., Ahyudanari, R. R., ... & Subroto, T. (2019). Optimization of culture conditions for Mpt64 synthetic gene expression in Escherichia coli BL21 (DE3) using surface response methodology. Heliyon, 5(11).

Nikerel, İ. E., Öner, E., Kirdar, B., & Yildirim, R. (2006). Optimization of medium composition for biomass production of recombinant Escherichia coli cells using response surface methodology. Biochemical Engineering Journal, 32(1), 1-6.

Abu, M. L., Nooh, H. M., Oslan, S. N., & Salleh, A. B. (2017). Optimization of physical conditions for the production of thermostable T1 lipase in Pichia guilliermondii strain SO using response surface methodology. BMC biotechnology, 17(1), 1-10.

Papaneophytou, C. P., & Kontopidis, G. (2014). Statistical approaches to maximize recombinant protein expression in Escherichia coli: a general review. Protein expression and purification, 94, 22-32.

Sopyan, I. Y. A. N., Gozali, D. O. L. I. H., & Guntina, R. K. (2022). Design-expert software (DOE): An application tool for optimization in pharmaceutical preparations formulation. Int. J. Appl. Pharm., 55-63.

Dentener, A. (2002). Design-Expert DOEs it better. Food Technology in New Zealand Magazine, 12-14.

Drummen, N. Aerobic cultivation of high-oxygen-demanding microorganisms in the BioLector XT microbioreactor.

Osthege, M., Tenhaef, N., Zyla, R., Müller, C., Hemmerich, J., Wiechert, W., ... & Oldiges, M. (2022). bletl-A Python package for integrating BioLector microcultivation devices in the Design-Build-Test-Learn cycle. Engineering in life sciences, 22(3-4), 242-259.

Zahn, D. A. (1975). An empirical study of the half-normal plot. Technometrics, 17(2), 201-211.

Kenett, R. S. (1991). Two methods for comparing Pareto charts. Journal of quality technology, 23(1), 27-31.

Bonnans, J. F., & Shapiro, A. (2013). Perturbation analysis of optimization problems. Springer Science & Business Media.

Wang, Y., & Li, Y. Z. (2014). Cultivation to improve in vivo solubility of overexpressed arginine deiminases in Escherichia coli and the enzyme characteristics. BMC biotechnology, 14, 1-10.

Baeshen, M. N., Al-Hejin, A. M., Bora, R. S., Ahmed, M. M., Ramadan, H. A., Saini, K. S., ... & Redwan, E. M. (2015). Production of biopharmaceuticals in E. coli: current scenario and future perspectives.

Baeshen, M. N., Bouback, T. A., Alzubaidi, M. A., Bora, R. S., Alotaibi, M. A., Alabbas, O. T., ... & Baeshen, N. A. (2016). Expression and purification of C-peptide containing insulin using Pichia pastoris expression system. BioMed research international, 2016.

Kemmler, W., Peterson, J. D., & Steiner, D. F. (1971). Studies on the conversion of proinsulin to insulin: I. Conversion in vitro with trypsin and carboxypeptidase B. Journal of Biological Chemistry, 246(22), 6786-6791.

Packiam, K. A. R., Ramanan, R. N., Ooi, C. W., Krishnaswamy, L., & Tey, B. T. (2020). Stepwise optimization of recombinant protein production in Escherichia coli utilizing computational and experimental approaches. Applied microbiology and biotechnology, 104, 3253-3266.

Elibol, M. (2004). Optimization of medium composition for actinorhodin production by Streptomyces coelicolor A3 (2) with response surface methodology. Process Biochemistry, 39(9), 1057-1062.

Sunitha, K., Lee, J. K., & Oh, T. K. (1999). Optimization of medium components for phytase production by E. coli using response surface methodology. Bioprocess Engineering, 21, 477-481.

Shahbazmohammadi, H., & Omidinia, E. (2017). Medium optimization for improved production of dihydrolipohyl dehydrogenase from Bacillus sphaericus PAD-91 in Escherichia coli. Molecular biotechnology, 59, 260-270.

Zare, H., Sadeghi, H. M. M., & Akbari, V. (2019). Optimization of fermentation conditions for reteplase expression by Escherichia coli using response surface methodology. Avicenna Journal of Medical Biotechnology, 11(2), 162.

Duan, M., Wang, Y., Yang, G., Li, J., Wan, Y., Deng, Y., & Mao, Y. (2020). High-level production of γ-cyclodextrin glycosyltransferase in recombinant Escherichia coli BL21 (DE3): culture medium optimization, enzymatic properties characterization, and product specificity analysis. Annals of Microbiology, 70, 1-13.

Kenari, S. L. D., Alemzadeh, I., & Maghsodi, V. (2011). Production of l-asparaginase from Escherichia coli ATCC 11303: optimization by response surface methodology. Food and Bioproducts Processing, 89(4), 315-321.

Ghoshoon, M. B., Berenjian, A., Hemmati, S., Dabbagh, F., Karimi, Z., Negahdaripour, M., & Ghasemi, Y. (2015). Extracellular production of recombinant L-Asparaginase II in Escherichia coli: Medium optimization using response surface methodology. International Journal of Peptide Research and Therapeutics, 21, 487-495.

Sparviero, S., Barth, L., Keil, T., Dinter, C., Berg, C., Lattermann, C., & Büchs, J. (2023). Black glucose-releasing silicon elastomer rings for fed-batch operation allow measurement of the oxygen transfer rate from the top and optical signals from the bottom for each well of a microtiter plate. BMC biotechnology, 23(1), 5.

Flitsch, D., Krabbe, S., Ladner, T., Beckers, M., Schilling, J., Mahr, S., ... & Büchs, J. (2016). Respiration activity monitoring system for any individual well of a 48-well microtiter plate. Journal of biological engineering, 10(1), 1-14.

Lennen, R. M., Nilsson Wallin, A. I., Pedersen, M., Bonde, M., Luo, H., Herrgård, M. J., & Sommer, M. O. (2016). Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects. Nucleic acids research, 44(4), e36-e36.

Back, A., Rossignol, T., Krier, F., Nicaud, J. M., & Dhulster, P. (2016). High-throughput fermentation screening for the yeast Yarrowia lipolytica with real-time monitoring of biomass and lipid production. Microbial cell factories, 15(1), 1-12.

Toeroek, C., Cserjan-Puschmann, M., Bayer, K., & Striedner, G. (2015). Fed-batch like cultivation in a microbioreactor: screening conditions relevant for Escherichia coli based production processes. SpringerPlus, 4(1), 1-10.

Kensy, F., Engelbrecht, C., & Büchs, J. (2009). Scale-up from microtiter plate to laboratory fermenter: evaluation by online monitoring techniques of growth and protein expression in Escherichia coli and Hansenula polymorpha fermentations. Microbial Cell Factories, 8(1), 1-15.

Dharmadi, Y., Murarka, A., & Gonzalez, R. (2006). Anaerobic fermentation of glycerol by Escherichia coli: a new platform for metabolic engineering. Biotechnology and bioengineering, 94(5), 821-829.

Malakar, P., & Venkatesh, K. V. (2012). Effect of substrate and IPTG concentrations on the burden to growth of Escherichia coli on glycerol due to the expression of Lac proteins. Applied microbiology and biotechnology, 93, 2543-2549.

Izaki, K., & Arima, K. (1965). Effect of various conditions on accumulation of oxytetracycline in Escherichia coli. Journal of Bacteriology, 89(5), 1335-1339.

Bren, A., Park, J. O., Towbin, B. D., Dekel, E., Rabinowitz, J. D., & Alon, U. (2016). Glucose becomes one of the worst carbon sources for E. coli on poor nitrogen sources due to suboptimal levels of cAMP. Scientific reports, 6(1), 24834.

Shiloach, J., Kaufman, J., Guillard, A. S., & Fass, R. (1996). Effect of glucose supply strategy on acetate accumulation, growth, and recombinant protein production by Escherichia coli BL21 (λDE3) and Escherichia coli JM109. Biotechnology and bioengineering, 49(4), 421-428.

Michaels, E. K., Chmiel, J. S., Plotkin, B. J., & Schaeffer, A. J. (1983). Effect of D-mannose and D-glucose on Escherichia coli bacteriuria in rats. Urological research, 11, 97-102.

Kopp, J., Slouka, C., Ulonska, S., Kager, J., Fricke, J., Spadiut, O., & Herwig, C. (2017). Impact of glycerol as carbon source onto specific sugar and inducer uptake rates and inclusion body productivity in E. coli BL21 (DE3). Bioengineering, 5(1), 1.

Ukkonen, K., Vasala, A., Ojamo, H., & Neubauer, P. (2011). High-yield production of biologically active recombinant protein in shake flask culture by combination of enzyme-based glucose delivery and increased oxygen transfer. Microbial Cell Factories, 10, 1-9.

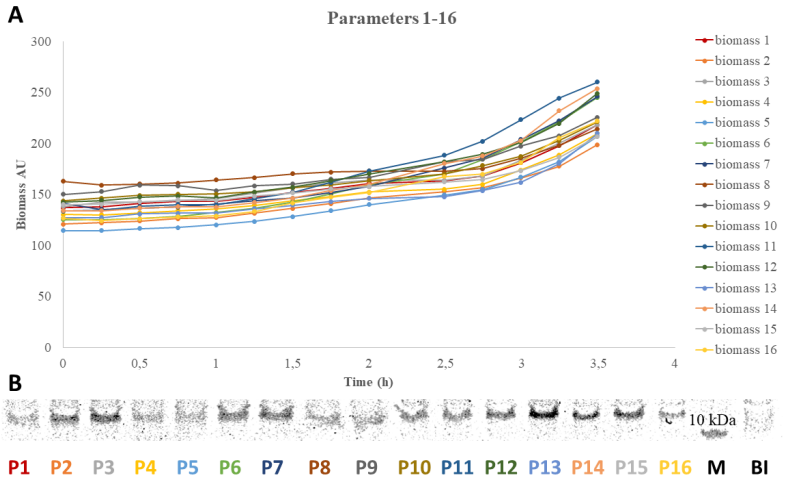



**Figure 1. Production and characterization of modified proinsulin.** Growth Kinetics of 16 parameters using BioLector (A) and analysis of resulting modified proinsulins via 20% SDS-PAGE (B).

**Table 1. ANOVA Table Analysis for Selected Factorial Model of Plackett-Burman Design** . The analysis of variance (ANOVA) table was obtained for the selected factorial model.

| Source | Sum of squares | df | Mean square | F value | p-value Prob > F |
|---|---|---|---|---|---|
| Model | 28155 | 4 | 7038,75 | 11,67 | 0,0020* |
| C-MgSO4 | 11285,33 | 1 | 11285,33 | 18,72 | 0,0025* |
| D-Glycerol | 7803 | 1 | 7803 | 12,94 | 0,0070* |
| E-Glucose | 5461,33 | 1 | 5461,33 | 9,06 | 0,0168* |
| H-LB broth | 3605,33 | 1 | 3605,33 | 5,98 | 0,0402* |

| | | | | | |
|---|---|---|---|---|---|
| Curvature | 2674,71 | 1 | 2674,71 | 4,44 | 0,0683 ns |
| Residual | 4823,5 | 8 | 602,94 | | |
| *Lack of Fit* | 4783 | 7 | 683,29 | 16,87 | 0,1854 ns |
| *Pure Error* | 40,5 | 1 | 40,5 | | |
| Cor Total | 35653,21 | 13 | | | |

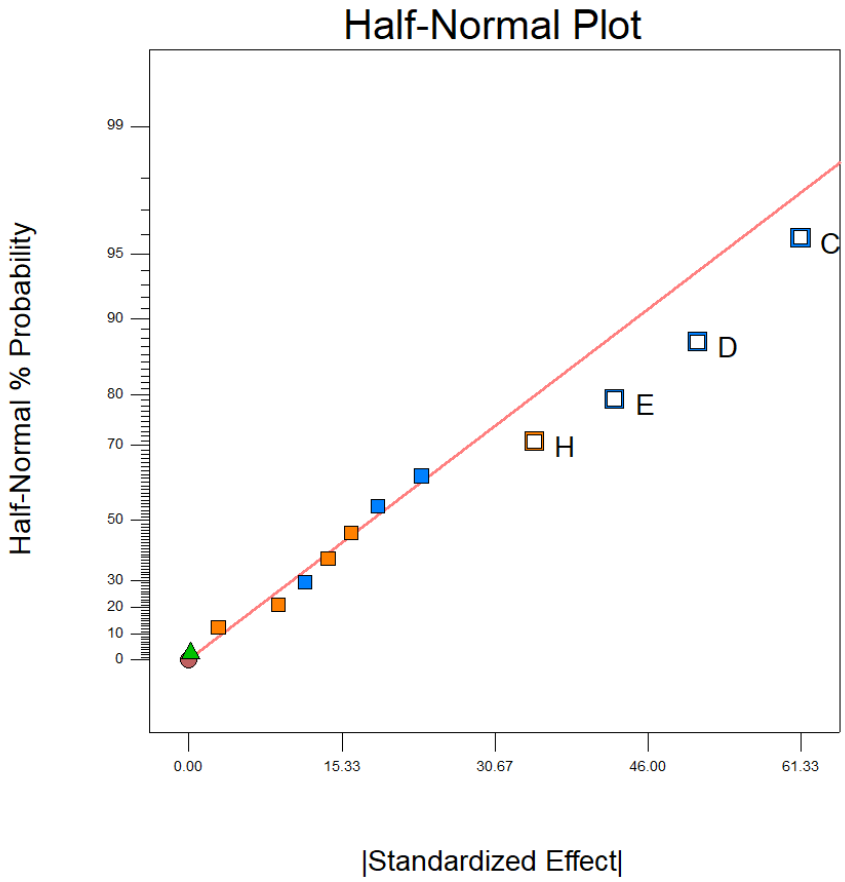



**Figure 2. Half-normal plot of Plackett-Burman Design.** The graph illustrates the relative importance of various independent variables within a regression model, specifically highlighting the significant impact of MgSO4, glycerol, glucose, and LB broth concentration values. These variables exhibit considerably higher percentage probabilities of standardized effects compared to other variables in the model, indicating their greater influence on the outcome.

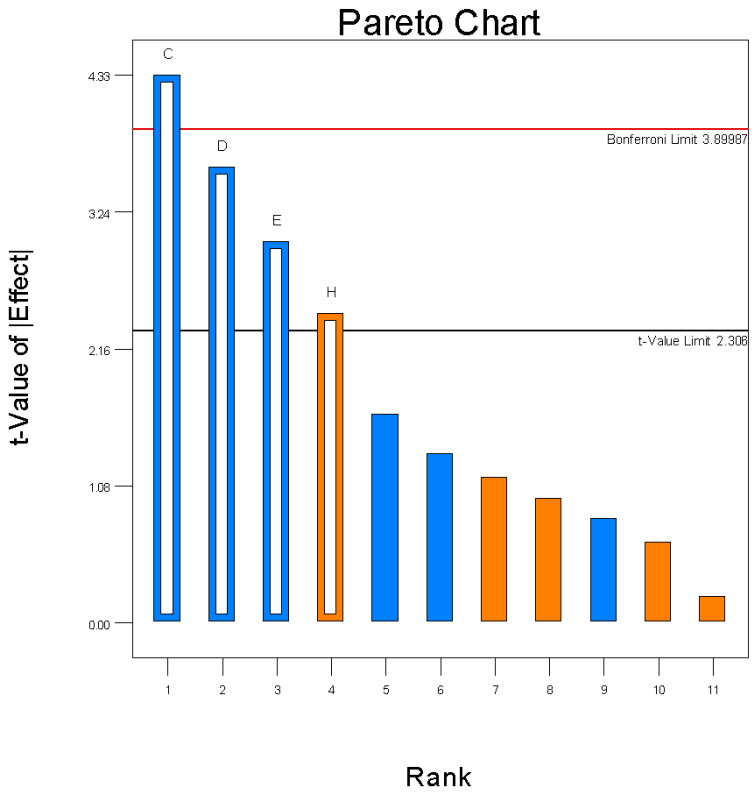



**Figure 3. Pareto chart of Placket-Burman Design.** The effectiveness of factors was evaluated using two statistical thresholds, namely the Benforrini limit and t-value limit. These thresholds were determined to be 3,898 and 2.306, respectively, for a significance level ($\alpha$) of 0.05. The Pareto chart visually represents the magnitude of standardized effects through bars. Notably, four bars stand out with the highest values, corresponding to MgSO4, glycerol, glucose, and LB broth concentration.

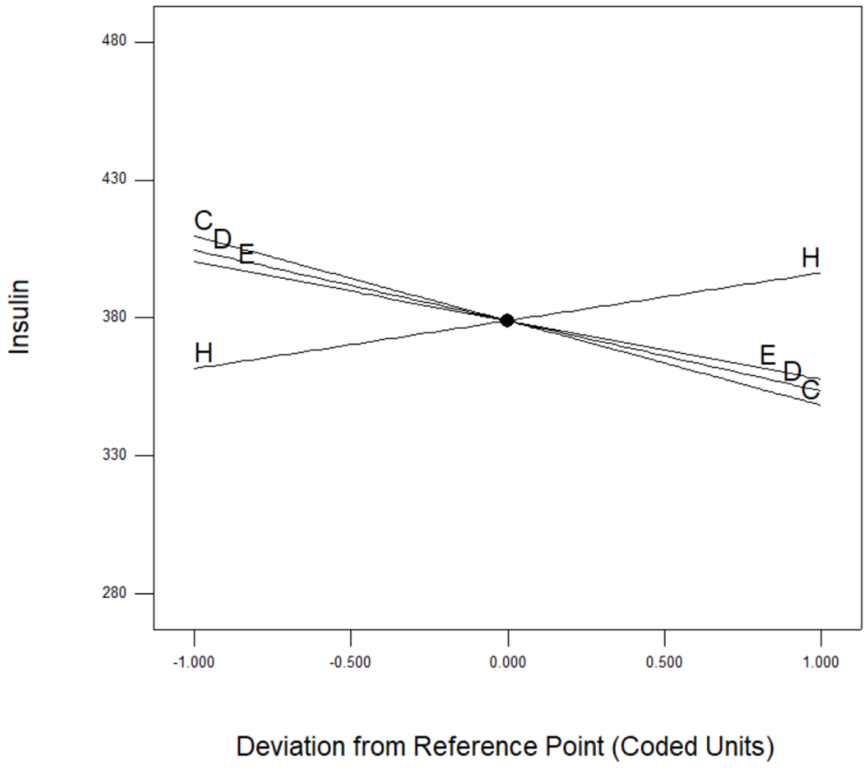



**Figure 4. Perturbation chart of Plackett-Burman Design.** A comprehensive understanding of the impacts caused by the individual and combined effects of independent variables on insulin yield was achieved by employing perturbation plots. These plots visually depict the direct effects of variables C, D, E, and H on insulin yield, providing a clear elucidation of their respective influences.

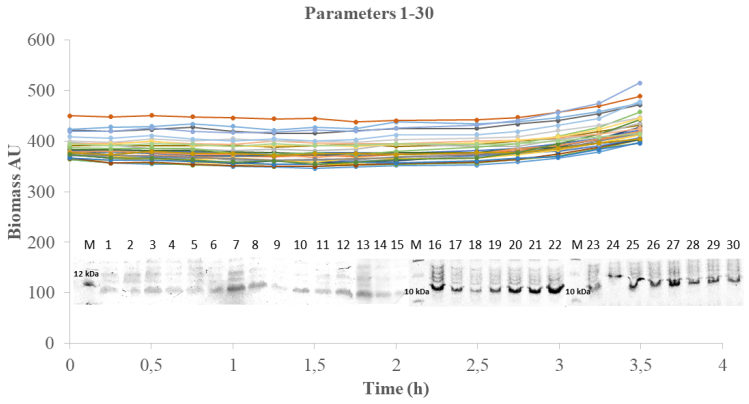



**Figure 5. Production and characterization of modified proinsulin for second round of the experimental validation.** Growth Kinetics of 30 parameters using BioLector to conduct CCD analysis and analysis of resulting modified proinsulins via 20% SDS-PAGE.

**Table 2** . The analysis of variance table is used to perform ANOVA for a Response Surface Reduced Quadratic Model.

| **Source** | **Sum of squares** | **df** | **Mean square** | **F value** | **p-value Prob > F** |
|---|---|---|---|---|---|
| Model | 201.91 | 8 | 25,24 | 5,31 | 0.0010* |
| A-Glucose | 1,25 | 1 | 1,25 | 0,26 | 0.6139 |
| B-Glycerol | 93,10 | 1 | 93,10 | 19,58 | 0.0002* |
| C-TB | 2,82 | 1 | 2,82 | 0,59 | 0.4502 |
| D-MgSo4 | 4,04 | 1 | 4,04 | 0,85 | 0.3671 |
| AB | 7,29 | 1 | 7,29 | 1,53 | 0.2293 |
| BC | 4,20 | 1 | 4,20 | 0,88 | 0.3579 |
| BD | 4,84 | 1 | 4,84 | 1,02 | 0.3245 |
| $B^2$ | 84,38 | 1 | 84,38 | 17,74 | 0.0004 |
| Residual | 99,86 | 21 | 4,76 | | |
| Lack of Fit | 78,33 | 16 | 4,90 | 1,14 | 0.4833 ns |
| Pure Error | 21,53 | 5 | 4,31 | | |
| Cor Total | 301,77 | 29 | | | |

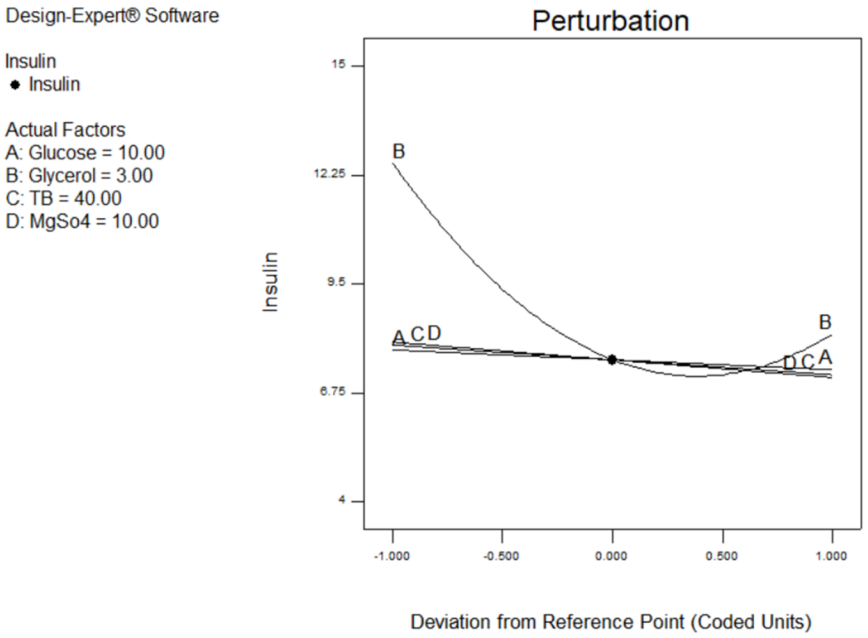



**Figure 6. Perturbation of Cental Composit Design.**Representation the perturbation of CCD analysis, which is clarified the primary marks of variables A, B, C, and D on insulin yield. Actual factors, A: Glucose 10.00, B: Glycerol 3.00, C: TB 40.00, D: MgSO4 10.00

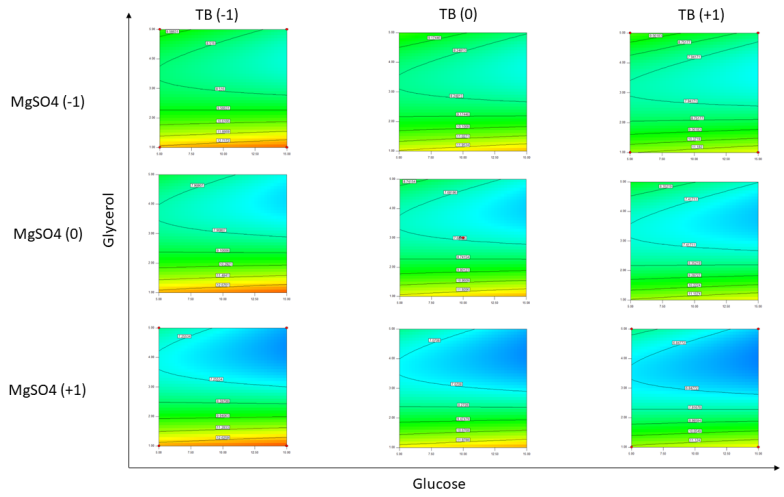



**Figure 7. Factorial contour plots of Central Composite Design.** A representative plot serves as an effective visualization tool, capturing the intricate relationship between the independent variables and the resulting insulin yield. The intentional manipulation of these independent variables within a predetermined range allowed for a systematic exploration of their collective impact on insulin production. X-axis and Y-axis are Glucose and Glycerol, respectively.

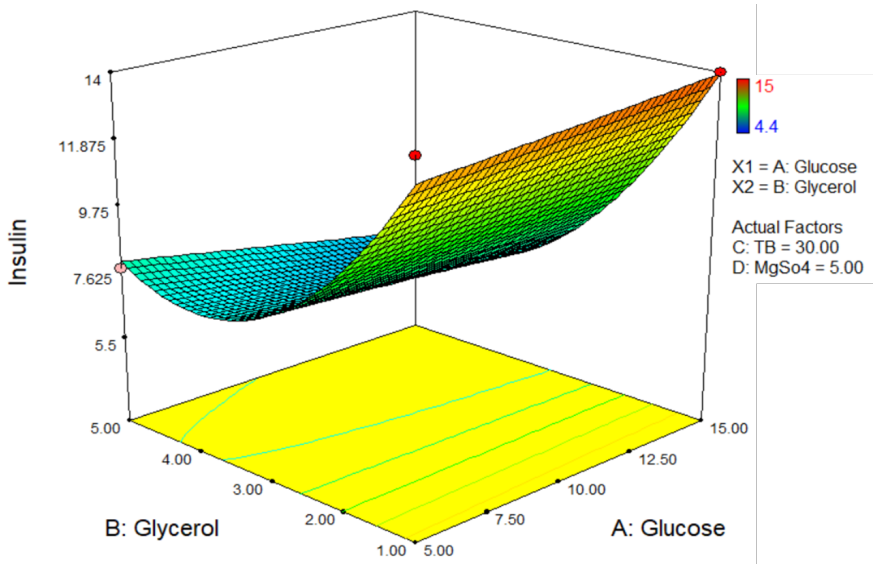



**Figure 8.** (A) Response surface plot illustrating the interactive effects of glucose, glycerol, MgSO4, and media content concentrations on the efficiency of insulin production, while keeping temperature and pH constant.

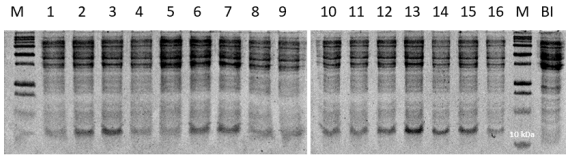
**Table 3.** The forecasted optimum conditions for achieving the maximum production of Insulin.

| Scenario | Glucose (mM) | Glycerol (1%) | TB (g/L) | $MgSO_4$ (mM) | Predicted Insulin (g/L) | Observed Insulin (g/L) | Deviation (%) |
|---|---|---|---|---|---|---|---|
| **1** | 14.58 | 1 | 30.09 | 15 | 13.92 | 11.87 | 14.8 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **2** | 15 | 1 | 30 | 6.73 | 13.79 | 11.54 | 16.3 |
| **3** | 8.78 | 1 | 30 | 15 | 13.44 | 13.00 | 3.3 |
| **4** | 15 | 1 | 37.51 | 10.62 | 13.21 | 12.33 | 6.7 |

**Supplementary Table 1.** Created trial configurations using PBD to assess factors and obtained resultant data

| Std | Run | KCl (5-10mM) | $MgCl_2$ (0-15mM) | $MgSO_4$ (0-15mM) | Glycerol (0-5%) | Glucose (0-15mM) | IPTG (0.2-0.4mM) | pH (6.8-7) | LB (30-50g/L) | $KH_2PO_4$ (0-15mM) | $Na_2HPO_4$ (0-15mM) | Thiamine (0-10mM) | Insulin g/L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 1 | 10 | 0 | 15 | 5 | 15 | 0,2 | 6,8 | 30 | 15 | 0 | 10 | 2,88 |
| 14 | 2 | 5 | 7,5 | 7,5 | 2,5 | 7,5 | 0,3 | 6,9 | 50 | 7,5 | 7,5 | 5 | 3,35 |
| 15 | 3 | 5 | 7,5 | 7,5 | 2,5 | 7,5 | 0,3 | 6,9 | 30 | 7,5 | 7,5 | 5 | 3,38 |
| 13 | 4 | 5 | 7,5 | 7,5 | 2,5 | 7,5 | 0,3 | 6,9 | 30 | 7,5 | 7,5 | 5 | 3,44 |
| 5 | 5 | 0 | 0 | 15 | 0 | 15 | 0,4 | 6,8 | 50 | 15 | 15 | 0 | 3,82 |
| 2 | 6 | 0 | 15 | 15 | 0 | 15 | 0,4 | 7 | 30 | 0 | 0 | 10 | 3,29 |
| 9 | 7 | 10 | 15 | 15 | 0 | 0 | 0,2 | 7 | 30 | 15 | 15 | 0 | 3,49 |
| 1 | 8 | 10 | 15 | 0 | 5 | 15 | 0,4 | 6,8 | 30 | 0 | 15 | 0 | 3,68 |
| 4 | 9 | 0 | 15 | 0 | 5 | 15 | 0,2 | 7 | 50 | 15 | 0 | 0 | 3,32 |
| 8 | 10 | 10 | 15 | 0 | 0 | 0 | 0,4 | 6,8 | 50 | 15 | 0 | 10 | 4,75 |
| 12 | 11 | 0 | 0 | 0 | 0 | 0 | 0,2 | 6,8 | 30 | 0 | 0 | 0 | 4,45 |
| 7 | 12 | 10 | 0 | 0 | 0 | 15 | 0,2 | 7 | 50 | 0 | 15 | 10 | 4,47 |
| 10 | 13 | 0 | 15 | 15 | 5 | 0 | 0,2 | 6,8 | 50 | 0 | 15 | 10 | 3,86 |
| 3 | 14 | 10 | 0 | 15 | 5 | 0 | 0,4 | 7 | 50 | 0 | 0 | 0 | 3,56 |
| 6 | 15 | 0 | 0 | 0 | 5 | 0 | 0,4 | 7 | 30 | 15 | 15 | 10 | 3,91 |
| 16 | 16 | 5 | 7,5 | 7,5 | 2,5 | 7,5 | 0,3 | 6,9 | 50 | 7,5 | 7,5 | 5 | 4,44 |




**Supplementary Figure 1. SDS-PAGE result of 16 Parameters.**

**Supplementary Table 2.** Achieving Factor Optimization through Central Composite Design Experimental Runs and Corresponding Response Assessment. KCl, $MgCl_2$, $KH_2PO_4$, Thiamine, IPTG and pH are constant value for variable parameters.

| | Constant values | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| std | run | KCl mM | MgCl2 mM | KH2PO4 mM | Tiamine mM | IPTG mM | pH | glucose (5-15 mM) | glycerol (1-5%) | LB broth (30-50 g/L) | $MgSO_4$ (5-15 mM) | Insulin g/L |
| 6 | 1 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 1 | 50 | 5 | 6,4 |
| 26 | 2 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 9,5 |
| 19 | 3 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 0,171573 | 40 | 10 | 6,3 |
| 4 | 4 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 5 | 30 | 5 | 11 |
| 14 | 5 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 1 | 50 | 15 | 4,2 |
| 30 | 6 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 3,5 |
| 1 | 7 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 1 | 30 | 5 | 6 |
| 24 | 8 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 17,07107 | 5,7 |
| 17 | 9 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 2,9289322 | 3 | 40 | 10 | 7,7 |
| 12 | 10 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 5 | 30 | 15 | 5 |
| 7 | 11 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 5 | 50 | 5 | 12,5 |
| 15 | 12 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 5 | 50 | 15 | 7,4 |
| 5 | 13 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 1 | 50 | 5 | 9,9 |
| 28 | 14 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 5,8 |
| 29 | 15 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 9 |
| 27 | 16 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 3,9 |
| 11 | 17 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 5 | 30 | 15 | -1,2 |
| 10 | 18 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 1 | 30 | 15 | 12,2 |
| 20 | 19 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 5,828427 | 40 | 10 | 9,2 |
| 23 | 20 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 2,928932 | 6,5 |
| 25 | 21 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 40 | 10 | 6,9 |
| 18 | 22 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 17,071068 | 3 | 40 | 10 | 2,7 |
| 22 | 23 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 54,1421356 | 10 | 5 |
| 9 | 24 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 1 | 30 | 15 | 5,1 |
| 2 | 25 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 1 | 30 | 5 | 11,5 |
| 16 | 26 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 5 | 50 | 15 | 5,1 |
| 3 | 27 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 5 | 30 | 5 | 4,9 |
| 13 | 28 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 5 | 1 | 50 | 15 | 4 |
| 21 | 29 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 10 | 3 | 25,8578644 | 10 | 5,7 |
| 8 | 30 | 10 | 15 | 15 | 10 | 0,25 | 6,8 | 15 | 5 | 50 | 5 | 4,9 |