Chapter

Strategies for Enhancing Product Yield: Design of Experiments (DOE) for *Escherichia coli* Cultivation

Puneet Kumar Gupta and Jyotheeswara Reddy Edula

Abstract

E. coli is considered one of the best model organism for biopharmaceutical production by fermentation. Its utility in process development is employed to develop various vaccines, metabolites, biofuels, antibiotics and synthetic molecules in large amounts based on the amount of yield in shake flasks, bioreactors utilised by batch, fed-batch and continuous mode. Production of the desired molecule is facilitated in the bioreactor by employing strategies to increase biomass and optimised yield. The fermentation is a controlled process utilising media buffers, micronutrients and macronutrients, which is not available in a shake flask. To maximise the production temperature, dissolved oxygen (aerobic), dissolved nitrogen (anaerobic), inducer concentration, feed or supplementation of nutrients is the key to achieving exponential growth rate and biomass. Design of experiments (DOE) is critical for attaining maximum gain, in cost-effective manner. DOE comprises of several strategies likewise Plakett-Burman., Box–Behnken, Artificial Neural Network, combination of these strategies leads to reduction of cost of production by 2–8 times depending on molecules to be produced. Further minimising downstream process for quickly isolation, purification and enrichment of the final product.

Keywords: *E. coli*, Design of experiments (DOE), Response surface methodology, Bacterial Fermentation, Batch, Fed-batch, Process optimisation, glycosylation, disulfide bridges, codon optimisation, host, post-translational modifications, inclusion bodies

1. Introduction

E. coli is most studied bacteria learned of symbiotic relationships with human for years derived after culturing F Plasmid of a 1922 isolate from a diptheria patient. Production of biopharmaceuticals from *E. coli* is in practice since 1965. Also, the it served for production of biopharmaceuticals such as recombinant proteins, metabolites by several companies namely BPB Bioscience, Agilent technologies, Promega, Takara, Tonbo Biosciences, New England Biolabs, Novagen and Lucigen. They are optimised for expression of challenging proteins difficult to express, purify and folding in native conformation. Moreover, formation of proper disulphide bonds and refolding of membrane proteins is also achieved by using newly commercially

available strains. These strains are suitable for production of metabolites and enzymes in easy to scale-up process development. The cultivation of *E. coli* is facilitated by optimisation of concentrations of tryptone (plant and animal derived), yeast extract, casamino acids. These are divided in three categories carbon, nitrogen and micronutrients, as per sustainable development goals set by united nations for 2030. Innovation for increasing productivity using technology is one of the 17 subject areas, it is need to follow typical strategy, which is usually unique for production of an active compound or protein. The **Table 1**, below describes different strategies to discovered for optimisation of product by utilising model organism *E. coli* (**Table 1**) [57].

Production of Biopharmaceuticals and Biomolecules during fermentation require media formulated using cheap ingredients for supplementation of carbon and nitrogen. The carbon sources are typically glucose, glycerol, molasses (**Table 1**). Nitrogen sources are typically yeast extract and digested proteins by enzymes. The utilisation of carbon by *E. coli* during batch can be inhibited due to accumulation of acetate, formate and succinate. Historically several attempts have been made to design defined or semi-defined media recipes that do not require bio nutrient additives. Traditionally, media developed generally utilise one-factor-at-a-time (OFAT) where all but one ingredient (factor) is fixed [58, 59]. This approach guides towards determining individual factors that influence culture growth. Effect of utilisation of amino acids, vitamins and minerals interaction are not possible using OFAT. These interactions can be studied using DoE and multifactorial designs.

DoE is a multifactorial methodology utilising statistical approaches to design and analyse an experimental process by which better outcome and results are obtained. Bacterial growth is the system response impact of "DoE planning" by varying critical parameters and simultaneous factors, such as phosphate, sulphate, vitamin and amino acid concentrations. The full factorial design, DoE is attributed by important principals to allow for robust, valid experimentation. These are known as treatments, replicated, randomised to protect against hidden factors [59]. All treatments must be performed in blocks to control sources of variation. In an experimental process design, all four principles, the statistical approach, conform to regulations and complete exploration of an extensive experimental design process, helps in reducing the number of experiments to get a good outcome [60]. For cultivation using *E. coli*, statistically relevant experimental designs for industrially important proteins (rB-glucosidase, human superoxide dismutase) is studied. A similar approach is applied enzymes, recombinant antibodies and therapeutic proteins(Ranibizumab, Somatropin Certolizumab pegol, PEG interferon alfa-2b, Romiplostim, Interferon-beta 1b, Pegloticase, etc.) [61]. The optimisation of media and critical components is needed for *E. coli* fermentation in batch and fed-batch fermentation is required to achieve a better growth rate, target product per litre of fermentation and reduced cost of production. The cost of various fermentation ingredients for one batch size 10 L of fermentation ranges from 20 to 155 USD (**Table 2**). The cost of media ingredients reduces after scaling up of process for production from 15 to 5000 L.

The DoE with a correct statistical model to understand the effect is needed to be adopted by industrial producers to achieve a robust and high yielding process. The DOE is statistically planned trials for optimising factors associated with increased product output, yield and cost reduction. Typically, DOE experiments are done to screen up to 2–15 elements. Traditionally, using OFAT, the number of experiments needed to determine the effect on yield, the output is numerous. Using the DOE approach, several components, additives and sources of nutrients are compared

Name of Product/ Molecule	Type of Fermentation	Design optimization	Media	Carbon Source	Reference
1,3-propanediol	Batch	Glycerol and glucose Co-fermentation	Reinforced Clostridium Medium (RCM)	Glucose	[1]
D-amino acid Oxidase	Batch	Combination fermentation	Complex medium	Glycerin	[2]
O-acetylhomoserine	Batch	Glycerol-Oxidative pathway	MS medium	Glycerol	[3]
Heparosan	High cell density	High cell density fermentation	Chemically defined fermentation media	Glucose	[4]
O-succinyl-l-homoserine	Batch	Multilevel fermentation optimization	Defined media	Glucose	[5]
(2S)-Naringenin	Batch	Fermentation and metabolic pathway optimization	Optimised MOPS minimal medium	Glucose	[6]
L-tryptophan	Batch	Batch Fermentation	Complex medium with additives	Glucose	[7]
Beta-farnesene	Batch	Batch Fermentation	2 × YT medium	Glycerol	[8]
L-tryptophan	Batch	Enzymatic hydrolyzate	Complex medium with yeast extract	Glucose	[9]
VP2 VLP vaccine	High cell density	Fed batch	Luria Bertani Media	Glycerol	[10]
Vitamin B12	Batch Fermentation	Metabolic engineering	LB media	Glucose	[11]
L-Valine	Batch fermentation	Two stage fermentation	Complex Medium	Glucose	[12]
PhoA	Batch Fermentation	Design of experiments for *Escherichia coli* fermentation	Complex medium Phosphate optimization	Glycerol	[13]
L-methionine	Batch fermentation	Design of experiments with Genetic engineered Escherichia coli	Complex medium with yeast extract	Glucose	[14]
Reteplase	Batch fermentation	Optimization using response surface methodology	LB medium (Response surface methodology for pH, Temp, RPM)	—	[15]
Tauroursodeoxycholic acid	Batch fermentation	Optimization of engineered Escherichia coli cell factory	Complex medium	Glucose and Glycerol	[16]

Name of Product/ Molecule	Type of Fermentation	Design optimization	Media	Carbon Source	Reference
Tyrosine phenol lyase	Batch fermentation	Statistical medium optimization	Complex medium	Sucrose	[17]
Phytochromes Production	High-cell density *E. coli* fermentation	High cell density fermentation	Minimal Medium	Glucose	[18]
Keratinolytic protease	Batch fermentation	Protein engineering to enhance keratinolytic protease activity	Complex Medium	Glucose	[19]
L-Threonine	Batch fermentation	Biofilm-Based fermentation strategies	Complex Medium	Glucose	[20]
Therapeutic DNA vaccine	Batch fermentation	Response surface methodology	2 x YT medium	Mannitol	[21]
Pyruvate Oxidase	High cell density	Temperature gradient based methodology	Yeast, Tryptone complex medium	Glycerol	[22]
rPDT fusion protein	Piolet scale batch fermentation	Statistical experimental design	LB medium	—	[23]
Platelet aggregation inhibitor	High cell density fermentation	Optimization of production	Yeast, Typtone	Glycerol	[24]
Vitreoscilla haemoglobin	Batch fermentation	Immobilised *E. coli*	Cheese Whey, Yeast extract	—	[25]
D-Lactose	Fed-batch fermentation	Optimization of D-lactose production	Modified mineral salt medium	Glucose	[26]
Curcumin	Batch Fermentation	Optimization of fermentation	TB medium (Tryptone, Yeast extract)	Glycerol	[27]
Alkaline beta-mannanase	Fed-batch fermentation	Induction conditions optimization	TB medium (Tryptone, Yeast extract) with additives	Glycerol	[28]
HIV-1 P17	Fed batch fermentation	Multiparameter optimization for secreted protein	TB medium (Tryptone, Yeast extract)	Glucose	[29]
Heptatitis Delta Virus (HDV) antigen	Fed batch fermentation	Optimization of Fed batch	LB medium	Glycerol	[30]
Chitosanase	Fed batch fermentation	Optimization of secretion efficiency PelB-CSN-N	Terrific Broth with trace elements	Glycerol	[31]
Nitrilase	High cell density	Optimization of temperature, substrate and IPTG	Super Optimal	Glycerol	[32]

Name of Product/ Molecule	Type of Fermentation	Design optimization	Media	Carbon Source	Reference
			broth (SOB) medium		
Human Interferon Beta	Batch fermentation	Reduction of acetate by Response surface methodology	Terrific Broth (TB)	Glucose	[33]
Beta-chitin	Batch fermentation	Selection of strain and IPTG optimization	Luria Bertani (LB)	—	[34]
Beta-glucosidase	Batch fermentation	Ag43-mediated thermostable for high temperature fermentation	Mineral Media	Glucose	[35]
46 KDa Antibody Fab' fragment	Batch fermentation	ultra-scale-down (USD) approach based on focused acoustics	Modified defined medium	Glycerol	[36]
Therapeutic protein production (10 KDa)	Fed-Batch fermentation	Optimization of inclusion bodies by autoinduction without adding IPTG due to galactose in the complex media	Yeast and Soy Peptone	Glycerol, Glucose	[37]
Catechol-O-methyltransferase	Batch fermentation	Optimization of production by artificial neural network	Tryptone, Yeast extract (SOB Medium)	Glycerol	[38]
Platelet aggregation inhibitor, salmosin	Fed-batch fermentation	Optimization of carbon and nitrogen sources	Yeast and Tryptone	Glycerol	[39]
VB4–845, an immunotoxin	High cell density fermentation	Optimization of VB4–845 titre in the supernatant	Glycerol Minimal medium	Glycerol, Arabinose	[40]
GDP-mannose	Coupling fermentation	Optimization by Coupling fermentation	Defined media	Fructose, Mannose	[41]
Heparosan	Fed-Batch fermentation	Defined medium using exponential fed-batch glucose addition with oxygen enrichment	Defined media	Glucose	[42]
Pig liver esterase	Batch fermentation	Optimization of co-expression in combination with chaperones GroEL/ES allowed the production of soluble and active enzyme	Complex medium	Glucose	[43]
Penta-N-acetyl-chitopentaose	Two step fermentation	Optimization of oligosaccharides production	Minimal Medium	Glycerol	[44]
Resilin	Batch fermentation	Primary induction step before lactose induction	Lactose, Casein, Yeast	Glycerol	[45]

Name of Product/ Molecule	Type of Fermentation	Design optimization	Media	Carbon Source	Reference
Malaria antigen PvRII	High cell density	High cell density for production of recombinant proteins	Defined media	Glucose	[46]
Bacterial phytase	Fed-batch fermentation	Optimization of two different feeding strategies	Terrific Broth	Glucose	[47]
Human interferon-gamma (hIFN-gamma)	High cell density fermentation	Exponential growth by using glucose as carbon source	M9 Modified medium	Glucose	[48]
B subunit of Escherichia coli heat-labile enterotoxin (LTB)	High cell density	Step wise addition of IPTG	Complex media	Glucose	[49]
Plasmid pAQN carrying the aqualysin I (AQI) gene	Batch fermentation	Optimization of yeast extract supplementation	LB and M9 minimal media	Glucose	[50]
Prochymosin	Batch fermentation	Use of lactose as energy source and inducer	Complex medium	Glycerol, Glucose	[51]
HIV-1 protease	Batch fermentation	Separation of growth and production phases in a two-step process	LB and M9CA medium	Glucose	[52]
Porcine somatotropin protein	High cell density	Oxygen enrichment, yeast extract (YE) effect, optimal specific growth in switching on gene expression, and feeding strategies.	Defined media with additives	Glucose	[53]
Human interferon-alpha 1	High cell density	Optimization of nutrient source	Defined media	Glucose	[54]
Human interleukin-1 beta	High cell density	Maintenance of low glucose and acetate for production	Defined media	Glucose	[55]
EcoRI	Batch fermentation	Control of pH and Oxygen	LB media	—	[56]

Table 1.
Strategies and optimization parameters and type of fermentation in production of various proteins and small molecules.

simultaneously, and their impact is evaluated. Based on the response critical factors are ranked. Soon after the response effect by variables are generated, the analysis is successively recorded. Various DoE experiments are studied to understand influencing factors determined by a statistically relevant experimental plan. The difference in factors is measured by Analysis of variance (ANOVA). The results obtained after analysing responses are plotted using contour plots and response surface methodology (RSM). There are several software packages for optimisation of critical factors and parameters Design Expert (Stat-Ease Inc.), GT-SUITE (DOE), DOE ++ (RecCom), MODDE (Umetrics), DoE Fusion PRO (S-Matrix Corp.), STAVEX (Aicos), Minitab (Minitab Inc.), and JMP (SAS) [62].

Component	Batch	Basal and Initial medium	Components added in feed solutions	Cost Per 10 Litre of Fermentation media (USD)
	Batch	Fed batch (Basal)	Fed Batch (Feed Media)	
NH_4Cl	7	7	28	71
KH_2PO_4	1.5	7.5	0	22
Na_2HPO_4	1.5	7.5	0	18
K_2SO_4	0.35	0.85	0	1
$MgSO_4.7\ H_2O$	0.17	0.17	0.86	0
Trace Elements[a]	0.8	0.8	3.2	1
Yeast extract	5	10	15	12
Glucose				0
Batch	1.0–5.0	20		27
Fed-Batch		1	99	1
Glycerol				0
Batch	1.0–3.0			1
Fed-Batch		2	30–60	1
			Total Cost	155[b]

Amount of component added to the medium listed for purpose mentioned (g/L).
[a]Millilitres of trace elements solution, prepared as described added as 1X to medium or feed solution.
[b]Cost of media components calculated with consumption in fermentation using high grade reagents.

Table 2.
Components of media used to grow E. coli in batch and fed-batch fermentations.

2. Critical fermentation ingredients

Critical fermentation ingredients are media components which cannot be replaced. These can only be standardised for maximised yield. Typically, buffers and nitrogen sources such as yeast extract, tryptone are not changed. Since *E. coli* is the most studied and highly utilised systems for producing various enzymes, antibodies, and biological products. Bacteria require specific conditions for growth attributed to factors such as oxygen, pH, temperature, and light. Bacterial growth is divided into lag, exponential (log) and stationary phase. During the initial stage, cellular activity in a rich nutrient medium allows cells to synthesise proteins, cells increase in size, but no cell division occurs in the phase. During the exponential phase, metabolic activity is high as DNA, RNA, cell wall components, or machinery needed for division are generated. The stationary phase is triggered due to the accumulation of waste products and depletion of nutrients. During the late log phase, proteins are induced by the addition of allolactose analogue, Isopropyl β-D-1-thiogalactopyranoside (IPTG) [63, 64]. The expression of recombinant products is controlled by promoter systems like T5 and T7 RNA Polymerases. Alternative promoter systems, such as auto-inducible phoA promoter system [13], the salt-inducible promoter (proU), arabinose-inducible promoter (pBAD) [65], the heat-inducible phage Lamba promoters (pL and pR), the cumate-inducible T5 promoter-based system [66], and the cold-inducible cspA promoter-based system [67] are also valuable for the biologics production. The cost of biologics production is due to the high cost of raw material and fermentation media. In the biologics industry, the more straightforward, cheaper, and reproducible process is highly appreciated.

Fermentation media is a critical component, and a balance of nutrients is needed for increasing productivity. Standardisation of *E. coli* fermentation requires identifying a combination of various media components available, e.g., Yeast extract, Soyabean meal, Bactotryptone, Meat extract and Enzymatic digest of plant and animal protein (Trypic or casein enzymes). There are various carbon sources (glucose, glycerol, sucrose, lactose etc). Additives for fermentation are vitamins, amino acids, and trace elements. Designing a media needs to evaluate the requirement of each of the individual component along with the additives. The design of model using statistical approach having multiple parameters in consideration, followed by validation of defined parameters using fermentation. This is achieved using DoE experimentation (**Figure 1**). Experiments are carried out at Shake flash level with selected nutrients such as carbon and nitrogen sources. Small scale studies are carried out to define as batch or fed-batch fermentation. Next stage is to screen the components available for fermentation of batch/fed batch. Once the components are finalised the possibility of Scale-up is evaluated based on the availability from the source. Finalisation of media components is carried out using shake flask with DOE of media buffer additives and inducers. Evaluation of various product outcome biomass ratio and validation of protein quality is also study with 3 to 5 selected designs. Once the nutrient and components are finalised pilot scale batches setup to study biomass to product ratio. Further optimisation of dissolved oxygen and temperature in fermentation is carried out by the DOE approach. If results are not reproducible with the selected condition, other near possible designs are studied to finalise the medium and process for fermentation (**Figure 1**). The process of selection of components is based on outcome in an experiment calculated by Biomass (OD), product output g L^{-1} and cost of ingredients. The process is clearly defined in (**Figure 1**).

2.1 Batch fermentation

In microbial batch, fermentation cultivation is done in a fixed volume of medium in a fermenter. The standard inoculum in the fermenter is 50–200 mL of shake flask volume in 2–5 L of fermentation media. The batch fermentation

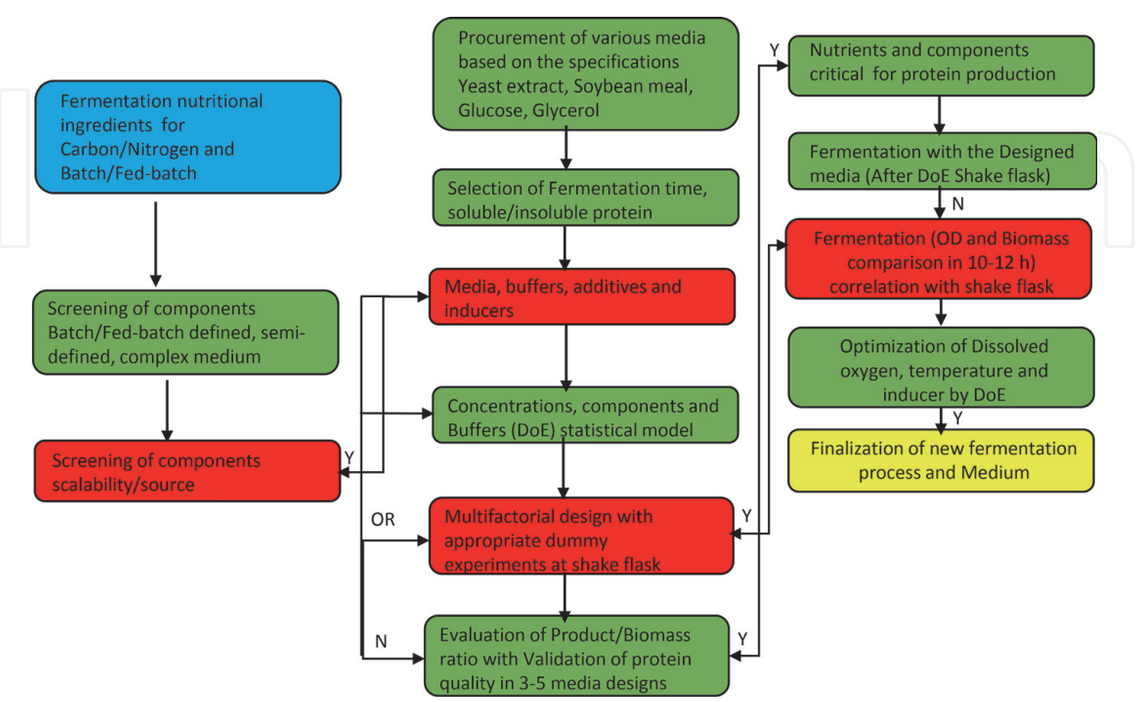

Figure 1.
Flow diagram for logical designing of upstream process for fermentation for E. coli. *Legend: Y = yes, N = No.*

typically OD_{600} is 20–40 in 8–12 h time. The microbial growth depletes the nutrients resulting in the accumulation of by-products; there is a continuous change in the culture environment. After completion of the batch, media and cells are harvested. The advantages are batch fermentation, ease of operation, low risk of contamination, high yield of protein to biomass in less time of fermentation, and majorly for soluble or excreted proteins. Typical disadvantages are relatively long downtime between batches due to vessel setup and sterilisation, low cell/biomass densities, due to cleaning. DOE is needed to optimise the required nutrients and minimise product accumulation during fermentation. Typical batch fermentation media constitutes Yeast extract, Bactotyptone (or Soybean meal) 10 g^{-l} to 24 g^{-l}, respectively. Buffers of Sodium and potassium phosphate in combination to reach pH 6.8. to 7.0, 100 X amino acid solution, Trace elements (400x) is $CoCl_2$. $6H_2O$ (1 mg/ml^{-1}), $MnCl_2.4H_2O$ (6 mg ml^{-1}), $CuCl_2.2H_2O$ (0.6 mg ml^{-1}), H_3BO_3 (1.2 mg ml^{-1}), $Na_2MoO_4.2H_2O$(1 mg ml^{-1}), Zn(II)acetate.$2H_2O$ (5.2 mg ml^{-1}) and Fe(III)citrate (40 mg ml^{-1}) [54]. The typical medium components are listed in the (**Table 2**) for batch fermentation as a base design to start optimisation.

2.2 Fed-batch fermentation

Fed-batch fermentation is a standard mode of fermentation in the bioprocess industry. Typically, fed-batch fermentation starts at the end of batch fermentation. *E. coli* is adapted and cultivated in defined media. In Fed-batch fermentation, cells are inoculated and grown in batch mode for 10–15 h. Once all the nutrients are depleted, evident by analysing the amount of glucose in the medium, dissolved oxygen levels are increased to 60–80%. The Fed-batch fermentation is initiated by starting to feed of Glucose, Vitamins, amino acids, and trace elements. The feed is added to the medium to allow the volumetric cell to increase the mass concentration exponentially. The growth rate of is changed to 0.12–0.22 μ h^{-1} during fed batch stage. These equations determine the growth rate in the medium.

$$ms(i) = (set + mYX/S)VSo, \tag{1}$$

$$ms(t) = ms(i)eset * (t - ti) \tag{2}$$

The first equation, ms (i), is the value of feed rate at the initiation of the fed-batch phase at time ti. μ_{set} is the specific growth rate, m is specific maintenance coefficient, $Y_{X/S}$ is yield coefficient, V is the bioreactor volume, and So is the initial glucose concentration. In the second equation, ms(t) is the rate of addition of substrate (g^{h-1}). After induction of protein expression, the specific growth rate of *E. coli* is typically reduced to 0.1 μ^{h-1}. The cells are harvested after completion of the run. Suppose the growth rate is not specified during fermentation, constant accumulation of several toxic metabolites produced during the fed-batch process acetate, formate, succinate, and lactate, resulting in oxygen limitation fed-batch. Therefore, it is recommended to wash cells with Tris-EDTA buffers after washing *E. coli* cells are stored or lysed for downstream processing. The distinct advantage of resuspending *E. coli* after completion of batch reduces protein degradation due to metalloproteases [46].

Start of feed is determined by measuring the concentration of substrate in the fermentation broth typically after 10–12 hrs of batch. The feeding strategy should be designed so that the growth rate is maintained to limit the production of toxic formate, acetate and other metabolic compounds, enhancing bacterial growth. The growth of bacteria and conversion of feed to biomass is maximum when the exponential growth phase is maintained. The utility of fed-batch and importance is

obtaining high cell density and biomass, leading to increased production of the high amount of product yield. The fed-batch is applicable to increase product yield by limiting growth rate and controlled substrate utilisation (**Table 2**). The media for batch and fed-batch fermentation is listed (**Table 2**).

3. Culture nutrient optimization strategies

3.1 One factor at a time- classical media optimisation methods

Selection of one-factor-at-a-time (OFAT) is a traditional method for optimisation of media. In this strategy only one factor is varied keeping all other parameters constant. The usual choice is ease and convenience; it makes OFAT the most preferred choice for formulating, designing, optimising, and scaling up the fermentation medium [68]. This method is still popular among many research groups for developing the medium for fermentation. Physical parameters, supplementation, removal, replacement and feedback experiments are primary considerations during OFAT. They comprise of growth temperature, operating pressure, size of nutrients or extracts. The bioprocess in fermentation controls constant supply of nutrients, removing metabolic products and toxic compounds, and constantly disseminating the nutrient solids, buffers and salts in liquid and gaseous phases. There is a constant evolution of design improvement for agitation and aeration; these allow better control over flow dynamics, the minimal effect of viscosity, and even circulation of components and nutrients. The healthy growth of culture in the batch is maintained by supplementing nutrients such as M9 medium with $FeSO_4 \bullet 7H_2O$; M63 medium with KOH; A medium $MgSO_4 \bullet 7H_2O$, 20% Glucose or sugar, vitamins, casamino acids or L-amino acids [69]. Removal experiments are required for the identification of critical components needed for the media. Certain media affect the reduction of formate, acetate, and reduction of pH. There are few examples associated with removing glucose from complex media to prevent inhibition of bacterial growth. Replacement experiments identifies correct nutrients complexes for nitrogen source yeast, soy peptone, bactotryptone, meat extract and protein powder. The carbon sources utilised are glucose, glycerol, sucrose, lactose and others. The use of OFAT for designing of media for fermentation limits number of experiments, the approach is suitable for production of metabolite. In one study, precursor carbohydrate phosphotransferase system (PTS) encoding genes a vital DXP pathway were deleted. This resulted in the enrichment of Isoprenoid phosphoenolpyruvate. Growth medium and production of lycopene (a C40 isoprenoid) resulted in maximisation by these culture conditions [70].

Defined media recipes used in the fermentation of *E. coli* include nine mineral salts in usually salts of ammonium, potassium and sodium cations; and carbonate, chloride, nitrate and sulphate anions. Glucose, Glycerol and ammonia were identified as potential additional sources of carbon and nitrogen, respectively. EDTA is a chelating agent, and seven trace elements are Iron chloride, Zinc chloride, Cobalt chloride, Sodium molybdate, calcium chloride, cuprous chloride, and Boric acid. The vitamins included in the experimental design solutions were Riboflavin, Panthothenic acid, nicotinic acid, Piridoxin, Biotin, and folic acid. In complex media, yeast extract ranges from 1 to 1.6 g $L^{-1,}$ is varied during preliminary concentrations. Screening designs often involve many factors and allow for initial differentiation of significant and non-significant factors and an estimation of the magnitude of the critical factors. A Full Factorial design, including 24 factors, would require almost 20 million experimental treatments. Fractional Factorial platform of the Design expert or JMP software can be used to generate 32 experimental

treatments, randomly distributed into eight blocks. Each block comprised eight treatments and provided information on a technical error, a positive control, and negative control. These results are calculated using standard algorithms in the software.

E. coli growth was studied with nine continuous factors. The media ingredients in the first iteration were found to be influencing optical density in relation to time. For this study custom design platform of the Design expert software was used to construct a design that balanced the need to maximise the information that could be gathered from the experiment whilst minimising resources and time. Total number of experiments and concentration of yeast extract tested in 50–60 experiments is determined 10 g L^{-1} to be the optimal concentration [15].

3.2 Statistical designs for *E. coli* Media optimization

Typically, statistical medium optimisation is beneficial in improving overall product output, reduces time needed for process development and cost. The microbial processes have complex reactions. Evaluation of results statistically increases the reliability of results, further reduction in the number of experiments. In one of the study, the GDP mannose pyrophosphate yield was improved upto 100% after conducting 33 experiments [61]. Improvement of media by DOE is for understanding various test variants, multiple investigations, and uniform pattern. The results obtained after the various experiments are used to predict media improvement using mathematical models. The current advancement in statistical techniques provides rapid analysis of experimental findings. Meticulously planned experiments can enhance the desired outcomes using DOE strategies. For designing a full factorial, possible combinations of relevant factors, e.g. temperature, pH, buffers, carbon, nitrogen sources, strain, are considered. Similarly, partial factorial design is considered if knowledge about few components is not available. These experiments if planned and output is studied properly results in quick and definitive reproducible processes.

3.3 Identification of critical components: Plakett Burman design

Cultivation at a large scale requires a medium that will produce maximum yield of product per gram of substrate, maximum biomass and minimum undesirable by-products. Also, consisten with minimal problems during media preparation and sterilisation. While considering the biomass in isolation, it must be recognised that efficiently grown biomass produced by an optimised high productivity growth rate is not necessarily best suited for its ultimate purpose, such as synthesising the desired product. Different combination, and sequences of process condition need to be investigated to determine phases, specific sets of conditions during optimisation. OFAT for media optimisation using traditional replacement experiments with keeping one factor at a time for nutrient, antifoam, pH, temperature are highly time-consuming and expensive. Minimum number of experiments and development of process in short duration of time is prerequisite to media optimisation. Therefore, other alternative strategies must be considered, which allows more than one variable to be changed at a time, and these methods have been described in earlier studies by Placket and Burman 1946, and Hendrix 1980 (**Table 3**).

The Plakett Burman algorithm is a rapid statistical approach enables us to obtain the physicochemical parameters and factors influencing the fermentation process with the limited number of planned experiments [71]. For the given number of observation, the linear effect of all aspects are screened with maximum accuracy. The design is practical when investigating a large number of factors to produce an

		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Std	**Run**	**A:Select Soytone**	**B:Yeast Extract**	**C:Glucose**	**D:Buffer**	**E:Trace Metals**	**F:MgSO4**
		g/L	g/L	g/L			g/L
6	1	0	0	0	Present	Absent	0.4
5	2	0	0	20	Absent	Present	0.4
9	3	16	24	20	Absent	Absent	0
11	4	16	0	20	Present	Present	0
1	5	16	24	0	Present	Present	0.4
8	6	16	24	0	Absent	Absent	0.4
10	7	0	24	20	Present	Absent	0
3	8	16	0	20	Present	Absent	0.4
7	9	16	0	0	Absent	Present	0
4	10	0	24	0	Present	Present	0
2	11	0	24	20	Absent	Present	0.4
12	12	0	0	0	Absent	Absent	0

Table 3.
Plackett-Burman design of experiments for media optimization (complex media example).

optimal or near-optimal response. Statistically optimised media design along with kinetic models characterises the fermentation behaviour more rapidly to achieve maximum productivity. Also, when complex carbon-nitrogen substrates, such as yeast extract or peptone, are used together with carbohydrate substrates, the Dissolved Oxygen (DO) change is not as significant when the carbon source is depleted, as cells continue to utilise the complex substrates [72]. The use of a good reliable model is essential to develop better strategies for optimising the fermentation process [73]. In one the study, during production of succinic acid [71] increasing output was achieved by combining Plackett-Burman design (PBD), steepest ascent method (SA), and Box–Behnken design (BBD) for fermentation medium. PBD identified Glucose, yeast extract, and $MgCO_3$ as critical components with optimal concentration was located to be 84.6 g L^{-1} of glucose, 14.5 g L^{-1} of yeast extract, and 64.7 g L^{-1} of $MgCO_3$ [2]. Also, the productivity was enhanced by 67.3% and 111.1%, respectively. Microbial fermentation for L-methionine (L-Met) production was enhanced by Plackett-Burman (PB) design, and Box–Behnken design (BBD) estimated glucose 37.43 g/L, yeast extract 0.95 g/L, KH_2PO_4 1.82 g/L, and $MgSO_4.7H_2O$ 4.51 g/L), L-Met titre was increased to 3.04 g/L from less than 2.0 g/L. an increase of 38.53% and 30.0% compared with those of the basal medium, respectively. Furthermore, higher L-Met productivity of 0.261 g/L/h was obtained, representing 2.13-fold higher in comparison to the original medium [14].

In another study, yield of O-succinyl-l-homoserine (OSH) was improved through multilevel fermentation optimisation; Plackett-Burman design was used to screen out three factors (glucose, yeast and threonine) from the original 11 factors that improved the titre of OSH.

Plackett Burman randomisation is an excellent tool for the determination of the effect of variables for optimisation. Once such approach for preparation of Bacterial Ghosts (BGs) preparation is established using these methods. The twelve experiments containing either the +1 or − 1 value for each variable in each experiment in random arrangement have been conducted simultaneously to get the best results and enable the best possible comparison. The BGQ has been given 100% quality as

10, while ten cells have been evaluated as either bad or good. This will decrease the range of the differences if we use %. Unexpectedly, *E. coli*, which is more sensitive to the SDS than *E. coli* BL21 (DE3), gives better results with most of the experiments. Nine experiments provide the number 10 out of the twelve experiments. Two give the number eight, and the only one shows the number 0, which means inferior preparation. The experimental Design is based mainly on the determination of Minimum Inhibition Concentration (MIC) and the Minimum Growth Concentration (MGC) of critical concentrations from chemical compounds able to convert viable cells to BGs. The mean of +1 experiments has been calculated using the following formula:$(\sum +\mathbf{1})/\boldsymbol{n}(+\mathbf{1})$. While the standard of −1 experiments has been calculated using the following formula: $(\sum -\mathbf{1})/\boldsymbol{n}(-\mathbf{1})$.The main effect of both +1 and − 1 for each variable has been calculated from the following formula: Main product = $\sum(+\mathbf{1})/\boldsymbol{n}(+\mathbf{1}) - \sum(-\mathbf{1})/\boldsymbol{n}(-\mathbf{1})$. Multiple linear regression analysis with ANOVA test of Plackett–Burman design has been performed on the BGQ as responses. A multiple linear regression analysis for the data of the BGQ has been committed to study the relationship between different variables and their level of significance regarding BGQ as a response. From the analysis of the Coefficient, Standard error, T Statistic, P-value and Confidence level % for each has been calculated. The confidence level has been calculated from the formula The confidence level% = $100*(1 - \text{P-value})$. The P-value from the ANOVA analysis for the BGQ response was determined to analyse the relationship between the variables at the 90% or higher confidence level. The model created from the analysis of Plackett–Burman experimental design using multiple regression analysis is based on the 1st order-model $\boldsymbol{Y} = \boldsymbol{ß0} + \sum \boldsymbol{ßiXi}$. Where Y is the predicted response, ß0 model intercept, ßi variables linear coefficient. ANOVA test was generated for each response to determine the relationship between the variables at the 90% or higher confidence level [74]. This improvement was applied for optimisation of production of chimeric protein PfMSP3-MSP1$_{19}$ resulted in critical concentrations are calculated are listed in **Table 4**.

3.4 Optimisation of fermentation conditions: Box-Behnken response surface methodology

George E. P. Box and Donald Behnken in 1960 developed the Box–Behnken response surface method. This algorithm establishes a comparison between composite central, three-level full factorial and Doehlert designs to optimise the fermentation conditions. In one of the example result of optimisation, the titre of O-succinyl-l-homoserine (OSH) reached 102.5 g l^{-1}, which is 5.6 times higher than before (15.6 g l^{-1}) [5]. Similarly, by Box–Behnken combination and Plackett-Burman design and were optimised further by employing the Response Surface Methodology, O-acetylhomoserine OAH production was up to 9.42 and 7.01 g/L. The effect of glycerol, ammonium chloride and yeast extract were screened for fermentation conditions [3].

scFv anti-HIV-1 P17 protein was optimised by the sequential simplex method. Plackett-Burman design (PBD) and sequential simplex were combined with the aim of improving feed medium for enhanced cell biomass, relative protein to biomass ratio. The scFv anti-p17 activity was enhanced by 4.43, 1.48, and 6.5 times more than batch cultivation, respectively [29].

DNA vaccine pcDNA-CCOL2A1 production was increased using the response surface method (RSM) in *E. coli* DH5alpha in fermentation, therapeutic DNA vaccine pcDNA-CCOL2A1 markedly increased from 223.37 mg/L to 339.32 mg/L under optimal conditions, and a 51.9% increase was observed compared with the original medium [21]. Statistical experimental design methodology for fermentation

Source of Variation	Degrees of Freedom	Sum of Squares [Partial]	Mean Squares [Partial]
Model	11	32.0001	2.9091
Main Effects	11	32.0001	2.9091
Residual	0	—	
Total	11	32.0001	
R-sq =	100.00%	—	
Regression Information		—	
Term	Effect	Coefficient	
Intercept		2.0933	
A:Select Soytone	1.5	0.75	
B:Yeast Extract	1.93	0.965	
C:Glucose	0.32	0.16	
D:Buffer	1.6467	0.8233	
E:Trace Metals	−0.4667	−0.2333	
F:MgSO4	−0.5733	−0.2867	
G:Dummy 1	−0.5667	−0.2833	
H:Dummy 2	0.6133	0.3067	
J:Dummy 3	0.3733	0.1867	
K:Dummy 4	−0.4467	−0.2233	
L:Dummy 5	−0.5433	−0.2717	

Table 4.
Plackett Burman design of experiments for media optimization after ANOVA analysis.

conditions (dissolved oxygen, IPTG, and temperature) improved rPDT production by *E. coli*. 15 Box–Behnken design augmented with centre points revealed that IPTG and DO at the centre point and low temperature would result in high yield. The optimal condition for rPDT production was found to be 100 mM IPTG, DO 30%, and temperature 20°C [23]. In another application, *E. coli* drug susceptibility testing was done by on-chip bacterial culture conditions using the Box–Behnken design response surface methodology for faster drug susceptibility, optimal growth parameters were determined within 6–8 h, MICs determination in 2–6 h of individual drugs (antibiotics and TCMs) to improve the clinical management of bacterial infection [75].

3.5 Functional characteristics with minimum experiments—Taguchi design

There are several challenges associated with the PBD and Box–Behnken design. To overcome these challenges of Box–Behnken new array based on "Orthogonal Array" was developed. Using this method, less number of experiments, instead of full factorial, is implemented. The system and technique provide control over three stages, likewise system strategy, tolerance design and parameters designing. The strategy design helps in determining tolerance, affecting factors in product output. Taguchi design is using a number of OAs to initiate the experimental setup, these arrays are utilised to suit the number of experimental iterations. Second step is conducting total tests with orthogonal arrays. These experiments are decided as per number of trial experiments as per Taguchi design, followed by randomisation of

experiments for determining the output. This design analyses main effect and two-factor interactions. Noise for uncontrolled experimental variables is considered, focal point for two-point analysis. Taguchi methodology removes effect of noise due to uncontrolled variables; this is better as compared to PBD [76]. The Taguchi method provides help in functional characteristic for capturing acceptable deviations. Human insulin-like growth factor I (hIGF-I) was produced in one study in *E. coli*, 32y media, 32°C and 0.05 mM IPTG. The unimproved hIGF-I was 0.694 g L^{-1} which improved to 1.26 g L^{-1} using optimum conditions [77].

3.6 Deciphering outcome with- central composite design

This design is widely used in building a second-order (quadratic) model in response surface methodology (RSM). It consists of factorial Design with two levels +1 and − 1; centre points, factorial Design in experiments with median values; and star points identical runs for centre points except for one factor considering values above median and below the median. The number of star points is double the number of factors used. CCD is defined on the level of factors: as Face centred CCD (CCF), Inscribed CCD (CCI) and Circumcentered CCD (CCC). *E. coli* BL21(DE3) is utilised for optimum Design for extracellular production of recombinant human epidermal growth factor (rhEGF) by CCD. This resulted in 122.40 μg mL^{-1} rhEGF concentration in medium 20 h after induction. In 2 L fermentation, medium optimised yield to 1.5 fold and induction time to 3 h [78].

3.7 Predicting effects of responses—partial least squares modelling

The effect of media ingredients, interactions with the system (X) and co-relation in response to culture ΔOD_{600} (Y) is defined by the Partial least squares (PLS) model. PLS covariance of matrix design and outcome are inferred accurately by virtue of small underlying events not measured directly. The ideal or latent variable (LVs) to study outcome needs to be carefully evaluated to avoid overfitting training data. The prediction of the accuracy of models for DoE iterations with multiple values, the Root Mean Predicted Residual Error Sum of Squares, is with the lowest value of LVs, having the slightest error. The significance of LVs is calculated by the Voet T2 test. The score of media design and component determined by OD600, the threshold of 0.8 is accepted. The lower score of model threshold defines to remove from subsequent designs; these threshold values are considered to study the positive and negative effect on the contribution of various factors associated with the increasing growth of the culture. In one study, 2D spectrofluorometry was utilised for fermentation processes to monitor the fermentation process online to produce extracellular 5-aminolevulinic acid (ALA). Various chemometric methods used for analysis of the spectral data are principal component analysis (PCA), partial least square regression (PLS) and principal component regression (PCR). PCA results visualised and considered for online fermentation monitoring. PCR and PLS compared for correlation between the 2D fluorescence spectra, PLS had slightly better calibration and prediction performance than PCR [79].

3.8 Minimal product trial experiments—definitive screening designs

Traditional definitive screening designs (DSD) require a low number of experiments and trials to determine the positive outcome. Jones and Nachtscheim developed Jones DSD methodology. These designs are popular in biopharma due to the relatively small number of experiments. These designs use three levels for each factor, allows estimation of nonlinear effects. Evaluate the number of runs to

determine X is 2 X + 1 or 2X + 3 for even and odd values. Typically, the Design of X = 6 is used if X < 6. There are few dummies runs with additional factorial or centre points are added to precisely determine the experimental error. In DSD, few different factorial trials or centre point trials are added to the initial design to define and evaluate the experimental error. One such example of DSD is for upstream process development for cell growth and increased product output in fed-batch high-cell-density fermentation. The expression of the desired gene cloned in the plasmid was under the control of the phoA promoter [13]. Simultaneous evaluation of phosphate concentration from 2.79 mM to 86.4 mM was designed using DoE. Several parameters, phosphate content, temperature, pH, and DO evaluated using a Definite screening design (DSD), resulting in determining each parameter's impact on product formation. Similarly, a 24-bioreactor ambr250™ system for fermentation utilised 10-factor DSD to characterise the process of demonstrating 16 batches reproducible workflow for recombinant protein production. This strategy was further evaluated by QbD approaches to assess techniques for late-stage depiction in small experiments and subsequently leading to large scale fermentation parameters improvement [80].

3.9 Stepwise regression and artificial neural network modelling

Artificial Neural Networks (ANN) are known for parallel, and continuous learning capabilities are known to interpret nonlinear functions. These are utilised to predict steady-state and dynamic processes. One iteration multi-layer perceptron (MLP) is famous for estimating hidden layers between output and input layers. Using this method, simulation of dissolve Oxygen (DO) parameters, Feeding (F), Biomass, Glucose, Acetate, and output production of γ-interferon is modelled. Several DoE iteration modelled using stepwise regression; these models are fitted with linear regression, six terms per model are allowed with Heredity Restriction. The goodness-of-fit of the resulting model is evaluated using Akaike Information Criterion (AICc). An artificial neural network (ANN) is used to create weighted ensemble of regression models. There are three nodes in single hidden layer of ANN. Sigmoid activation functions were used, cross-validation of 19 of the media formulations defined in the second DOE iteration were randomly selected and withheld from ANN training set to do validation studies [59].

4. DOE optimisation techniques for *E. coli*

4.1 Evaluation of production and process—response surface methodology (RSM)

RSM is simple, robust and efficient, in predicting processes of metabolite or product production. Also, this method helps in the determination of factors for specifications, changes in levels of the elements, response with specified levels, quantitative understanding of system behaviour, predict product properties, factor combinations not run and stability of the designed process.

RSM methodology consists of different phases. Typically, performed in three steps, First is the screening factors by steepest ascent/descent, secondly by quadratic regression model fitting, third optimization using canonical regression analysis.

For a cost-effective and robust process, improvising parameters related to medium, productivity, safety and usefulness are desired outputs. The

interdependency of factors associated with productivity is difficult to understand, and this slows downs the enhancement process and yield evaluation.

Response surface methodology (RSM), is based on factorial designs to elevate the process and final product yield. RSM is considered a sturdy, robust, and efficient mathematical approach. It includes experimental statistical methods, multiple regression design, and analysis, resulting in developing the best strategies guided by constrained equations. RSM is typically applied to study the response of different media ingredients [21]. The production of Examples for *E. coli* fermentation. One such example is standardisation of production of human interferon-gamma (hIFN-γ).

$$Y = \beta_0 + \sum \beta_i x_i + \sum \beta_i i x^2 i_i + \sum \beta_i j x_i x_j + \varepsilon. \quad (3)$$

Where β_0 is defined as the constant, B_i the linear coefficient, $B_i i$ the quadratic coefficient and $B_i j$ the cross-product coefficient. X_i and X_j are levels of the independent variables, while ε is the residual error. This variable and RSM predicted 7.81 g L^{-1} glucose, 30°C for fermentation and induced at OD600 1.66, Combined with BBD to get the 95.50% acetate and 97.96% productivity of rhIFN-β [81].

Plackett Burman design and Response surface methodology are utilised together to increase the production of the desired product multiple-fold. Combining these techniques is usually employed to enhance the product outcome of several microbial processes, batch, fed-batch fermentation. RSM is widely used, with Plackett Burman, CCD, Box Benken. Even after much success, several limitations are associated with RSM, likewise predicting responses based on second-order polynomial equation [82], results in poor estimation of optimal designs, leading to low levels of yield or outcome. One limitation is developing a model for many variables on physical and chemical inputs due to nonlinear biochemical network interactions, with partial knowledge of these systems [83]. Another limitation is the study of multiple interactions and significant variations, resulting in error, bias and or no reproducibility. These challenges are dealt with better using Artificial neural networks (ANN) [84].

4.2 Study of interaction of pathways and multiple parameters—artificial neural network

An artificial neural network (ANN) is designed for a computing system to simulate the information and process the data similar to the human brain, guided by artificial intelligence (AI). It solves problems impossible or difficult by human or statistical standards. Handling units consist of inputs and outputs; using these inputs, ANNs produce desired or defined work. ANNs are built as neurons, are interconnected like a web in the human brain. There are hundreds or thousands of artificial neurons or processing units interconnected by nodes [85].

Similarly, as human functions, ANNs have a set of rules for learning backpropagation, an abbreviation for backward propagation of error, to perfect their output results. Typical processes in ANN are the training phase to recognise patterns in data, whether visually, aurally, or textually; in this supervised learning, the actual output is compared with the desired outcome. The differences are adjusted using backpropagation. The program runs backwards as we advance, and adjustments are made until the actual and expected output difference results in a minimum possible error. Designing of medium or metabolic process ANN is highly suitable, and it generates large amounts of data. The architecture of ANN consists of three layers: a layer of "input" connected to "hidden" units, ultimately connected with "output". The conditions for ANNs three types are Supervised, Unsupervised

and Reinforced learning. The objective of supervised learning for the neural network is to provide input training data and possible experimental output. An unsupervised output unit is trained to respond to clusters or patterns present in the input data. Reinforced learning is an intermediate system; learning systems' actions are considered good or bad based on environmental responsibility. These parameters are adjusted till the time equilibrium state is attained. These systems are applied for system designing, modelling, optimisation. It leads to control the noisy signals and generalise through system training procedure. The ANNs are employed in various fermentation processes to optimise nutrient and prediction biomass outcome in different culture conditions. ANN has several limitations; likewise, it needs proper training, also based on input data to get the quality output [86]. To overcome some of these challenges, ANN combined with a genetic algorithm (GA) is applied to improve the concentration and shelf life of aspartate-β-semialdehyde dehydrogenase protein [87].

4.3 Study of biological process of evolution- genetic algorithm (GA)

Genetic algorithm (GA), developed in 1975 by Holland and Long, is based on Charles Darwin's theory of natural selection. GA is a model for the study of biological evolution by testing crossover and recombination mutation in adaptive and artificial systems. A genetic algorithm works as a problem-solving strategy using essential genetic operators. There are several GA designed to deal with complex problems and parallelism for stationary or non-stationary functions, linear or nonlinear, continuous or discontinuous, random or noise. Improvement in yield of recombinant G-CSF was obtained in auto-induction medium. The backpropagation (BP) algorithm and radial basis function (RBF) algorithm combined with the Genetic Algorithm improve G-CSF yield. The yield of models was 72.24 and 76.09%, respectively, and are higher than those obtained using non-optimised auto-induction mediums.

There are some disadvantages or genetic algorithms as well. The formulation of the fitness function, population size, choice of factors for mutation and cross over, selection of criteria for these factors needed to be carried out carefully. Despite drawbacks, GA is one of the widely used algorithms in modern nonlinear optimisation [88].

4.4 Geometric function evaluation- Nelder-Mead simplex algorithm

Nelder–Mead published this algorithm in 1965. The objective is to solve the classical unconstrained optimisation problem of minimising a given nonlinear function; without derivative only numerical evaluation of the objective function is needed. This algorithm is based on geometry, and in three-dimensional space, simplex is a tetrahedron determined by four points (vertices) and their interconnecting line segments. For two dimensions, simplex is an equilateral triangle, and three dimensions should be tetrahedron. The objective function is evaluated every point with the highest numerical value of all four points is perpendicularly mirrored against the opposite plain segment, generating reflection [89]. An expansion can accompany the reflection to take more significant steps or a contraction to shrink the simplex where an optimisation valley floor is reached. The optimisation procedure continues until the termination criteria are met. The termination criterion is usually the maximum number of reflections with contractions or tolerance for optimisation variables. The algorithm can be implemented in N dimensions, where simplex is a hypercube with N + 1 vertex points. The NM

method provides significant improvements in primary iteration and improves outcome. NM is combined with ANN to optimise the production of metabolites [90].

5. Problems and bottlenecks in *E. coli* media optimisation

Medium optimisation involves many experiments irrespective of the media chosen, which accounts for labour cost and is an open-ended experiment. Many experiments are carried out at shake flask, even after generation of large amount of data using single experiments. The results obtained at piolet scale batch fermentation are not reproducible. During shake flask experiments, the precise control over pH, oxygen transfer and evaporation is not controlled. The experiments carried out at shake flask may or may replicate during fermentation. Also, soluble proteins expression may lead to inclusion bodies formation. Optimisation of media is time consuming due to the requirement of rigorous experimental planning. Moreover, the media utilised in the production of recombinant products faces challenges due to variability in different batches, media availability, cost of media, bulk storage, transport time. For Biotherapeutics, Enzymes and Probiotics, the cost of media needs to lower in Probiotics compared to Enzymes and Biotherapeutics, respectively. The choice of fermentation depends on the solubility of protein from batch to fed-batch. *E. coli* cells are dynamic, and every product requires different media compared to the earlier optimised process. Optimisation of media depends on considering dynamic internal control mechanisms. For the production of metabolites after engineering of bacterial strains, metabolic pathways needed to be optimised to regulate the desired product by choice of media. The influence of using different strains for the production is dependent on toxicity, complexity (Disulphide bonds in the sequence), AT-rich sequences. In our previous study, *E. coli* cell Shuffle 3030H for production of *Plasmodium falciparum* MSP-3 and the MSP-1_{19} fusion protein was successfully optimised to generate protective antibodies [91].

Improvement of production of recombinant products is also guided by downstream processing of protein. Therefore, series of experiments designed for correct folding and confirmation are most important. Significant protein amounts can be achieved using pH, time for fermentation, oxygen transfer and temperature for fermentation. Also, inducer and harvest time are critical for increasing output. The critical factors for fermentation in batch and fed-batch are different. Therefore, the choice of media defined, semi-defined or complex media with vitamins, minerals and trace elements needed to be considered for evaluation in DOE experiments. To evaluate the considerable amount of output and variables combination and application of various algorithms is done to achieve desired output. In all the optimisation process and advanced algorithms such as Artificial neural networks and Genetic Algorithm are applied to achieve the desired output efficiently. The need for innovation as per sustainable development goals (SDGs) for United Nations 2030 plan is needed to increase the reach of technologies to low income countries. The application of DOE can improve the yield and cost leading to improved access to Biomolecules, Biopharmaceuticals, enzymes and metabolites.

6. Future strategies

The selection of host cells for industrial application has some technical difficulties despite the availability of many gene manipulations theoretically in various organisms. The availability of a genetic map, gene exchange system, useful vector and transformation procedures, and metabolic pathways leading from raw material

to the desired product are essential criteria for selecting a suitable host strain. The most popular organisms used to date for the expression of the recombinant proteins are *E. coli*, *Bacillus subtilis, Bacillus stearothermophillus, Streptomyces spp, Corynebacterium, Saccharomyces cerevisiae, Pichia pastoris, Hansenula polymorpha* and various animal/plant cells. *E. coli* remains an important host system for the industrial protein production from cloned genes as one of the main applications of genetic engineering in biotechnology. Various efficient expression vector systems have been developed, and a variety of mutants are available as host strains for different purposes [92, 93]. Overexpression of a heterologous protein is possible in *E. coli,* making it suitable for industrial production. Fermentation DOE is an essential tool for basic research that greatly facilitates efficient purification and analysis of such proteins [94].

For the successful production of the recombinant protein-based vaccine, producing biologically active protein is an essential requirement that can be further scaled up. Production of a biologically active recombinant protein depends on the host cell's microenvironment for expression and compatibility of codon usage. *E. coli* has been a widely used expression host for the high-level production of heterologous protein. Differences in usage of codons in prokaryotes (*E. coli*) and eukaryotes Chinese hamster ovary cells (CHO) can substantially impact heterologous protein production. The compatibility of codon usage can significantly increase protein expression [95, 96]. Moreover, the presence of rare codons in cloned genes affects protein expression level and mRNA & plasmid stability. The excessive presence of rare codons may result in ribosome stalling, slow translation errors [96, 97]. In some cases, rare codons inhibit protein synthesis and cell growth [98]. Earlier studies of codon usage patterns in *E. coli* have established that a clear codon bias exists in the mRNA. The level of each cognate tRNA seems to be directly proportional to the codon frequency [99, 100]. The strategy widely used is to change rare codons in the target gene to the favoured codons of *E. coli* without affecting the encoded amino acid sequence [101, 102]. The second approach is to expand the intracellular tRNA pool by introducing a plasmid encoding additional copies of tRNAs for codons rarely used in *E. coli* [103]. The co-presence of the RIG plasmid encoding three tRNAsAG(A/G), ATA, GGA in the host cells significantly increases the expression level of Dihydropteroate synthase, Aldolase, Phosphatase, and Orotidine-5′-monophosphate decarboxylase of *P. falciparum* [104–106]. Codon optimisation for maximum expression of foreign proteins by changing host cell favourable codons is beneficial and crucial for large-scale proteins [107].

The recombinant plasmid carrying cloned gene would behave differently compared to the original vector plasmid. It can be easily understood, as it is preserved under a delicate quasi-equilibrium state in the host cell. There are several reasons for the instability of recombinants. The higher the plasmid gene expression, the more segregants (plasmid free cells) tend to appear. The recombinant plasmid is relatively unstable when the cloned gene products are inhibitory to the host cells. Phenotypic instability of plasmid is due to the disappearance of the entire plasmid or the deletion of a specific region [108]. Both plasmid copy number and plasmid loss rate are features affected by factors such as media composition growth rate and culture strategy [109] and other factors such as temperature, agitation rate, and pH [110].

Therefore, future strategies for optimisation of cultivations needed to be shift to conclusions evaluated during experimental phases before actual fermentation to identify role of batch, fed-batch or different media components. The utilisation of carbon, nitrogen and other minimally required nutrients during batch and fed-batch is critical for delivering output and achieve sustainable development goals (SDGs) for technological innovation. The method design and modelling approaches are future strategies for increasing output during a process development. Utilisation

of one factor and carrying out experiments by statistical media optimisation can be improved by combining several algorithms such as Plakett Burman, Box–Behnken, Taguchi design, Central composite design, partial least squares modelling in determining optimal factors. Response surface methodology with Artificial neural network (ANNs) can be applied to difficult model kind of fermentation processes. A free artificial neural network is applicable for carrying out nonlinear regression models to optimise metabolic processes. These algorithms are combined and applied to increase productivity and optimise the product output by reducing the cost of fermentation and product development.

7. Conclusions

Optimisation of critical factors and nutrient sources is an essential step for metabolite, recombinant proteins before pilot fermentation. In this chapter, strategies, conventional, advanced process design are reviewed and detailed. DOE approaches with statistical evaluation are critical for process development are essential for saving experimentation time. The strategies and examples shared in this review have been analysed for ease of implementation, time consumption. The conditions and media designed needs to be further tested under realistic conditions, full scale process with replication to production setup.

Overall, this chapter detailed need of critical factors identification, their significant contribution in enhancing process of metabolite production. Also, recently, co-fermentation of glycerol and glucose in engineered *E. coli* increased production of 1, 3 propanediol [1]. Similarly, O-acetylhomoserine production is increased by suitable designs for fermentation and modification of glycerol-Oxidative pathway [3]. Production of recombinant protein in one study response surface methodology was utilised for production of repletase and improved yield to 188 mg $L^{-1\ of}$ fermentation [15]. The approaches discussed in this chapter have several advantages for improving the yield and reduction of resource utilisation. These approaches are efficient for achieving the access for biotechnologically produced products to reach the larger population across the globe.

Acknowledgements

Puneet Kumar Gupta acknowledges financial support from Grace Lifetech Pvt. Ltd. for the study. Jyotheeswara Reddy Edula acknowledges support from the Tata Institute of Genomics and Society (TIGS).

Conflict of interest

The authors declare no conflict of interest.

Acronyms and Abbreviations

DOE	Design of Experiments
E. coli	*Escherichia coli*
OFAT	one-factor-at-a-time
ANOVA	Analysis of variance
RSM	Response surface methodology

DNA	Deoxyribonucleic acid
RNA	Ribonucleic Acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
EDTA	Ethylenediamine tetra acetic acid
DO	Dissolved Oxygen
PBD	Plackett-Burman design
SA	Steepest ascent method
BBD	Box–Behnken design
OSH	O-succinyl-l-homoserine
BGs	Bacterial Ghosts
MIC	Minimum Inhibition Concentration
MGC	Minimum Growth Concentration
BGQ	Bacterial Ghosts Quantification
OA	Orthogonal Array
CCF	Face centred CCD
CCI	Inscribed CCD
CCC	Circumcentered CCD
DSD	Definitive screening designs
QbD	Quality by design
ANN	Artificial Neural Network
GA	Genetic Algorithm
NM	Nelder–Mead
SDGs	Sustainable Development Goals
CHO	Chinese Hamster Ovary

Author details

Puneet Kumar Gupta[1*] and Jyotheeswara Reddy Edula[2]

1 Grace Lifetech Pvt. Ltd., New Delhi, India

2 Tata Institute of Genetics and Society, Bangalore, Karnataka, India

*Address all correspondence to: guptapuneetin@gmail.com

References

[1] Yun J, Zabed HM, Zhang Y, Parvez A, Zhang G, Qi X. Co-fermentation of glycerol and glucose by a co-culture system of engineered *Escherichia coli* strains for 1,3-propanediol production without vitamin B12 supplementation. Bioresour Technol. 2020/10/14 ed2021. p. 124218.

[2] Xu JM, Cao HT, Wang M, Ma BJ, Wang LY, Zhang K, et al. 2021. Development of a Combination Fermentation Strategy to Simultaneously Increase Biomass and Enzyme Activity of D-amino Acid Oxidase Expressed in Escherichia coli. Appl Biochem Biotechnol. DOI: 10.1007/s12010-021-03519-7.

[3] Liu P, Liu JS, Zhang B, Liu ZQ, Zheng YG. 2021. Increasement of O-acetylhomoserine production in Escherichia coli by modification of glycerol-oxidative pathway coupled with optimisation of fermentation. Biotechnol Lett 43 (1):105-117. DOI: 10.1007/s10529-020-03031-8.

[4] Datta P, Fu L, Brodfuerer P, Dordick JS, Linhardt RJ. 2021. High density fermentation of probiotic E. coli Nissle 1917 towards heparosan production, characterisation, and modification. Appl Microbiol Biotechnol 105 (3):1051-1062. DOI: 10.1007/s00253-020-11079-9.

[5] Zhu WY, Niu K, Liu P, Fan YH, Liu ZQ, Zheng YG. 2020. Enhanced O-succinyl-l-homoserine production by recombinant Escherichia coli DeltaIJBB*TrcmetL/pTrc-metA(fbr) -Trc-thrA(fbr) -yjeH via multilevel fermentation optimization. J Appl Microbiol. DOI: 10.1111/jam.14884.

[6] Zhou S, Hao T, Zhou J. 2020. Fermentation and Metabolic Pathway Optimization to De Novo Synthesize (2S)-Naringenin in Escherichia coli. J Microbiol Biotechnol 30 (10):1574-1582. DOI: 10.4014/jmb.2008.08005.

[7] Zhao C, Fang H, Wang J, Zhang S, Zhao X, Li Z, et al. 2020. Application of fermentation process control to increase l-tryptophan production in Escherichia coli. Biotechnol Prog 36 (2):e2944. DOI: 10.1002/btpr.2944.

[8] Yao P, You S, Qi W, Su R, He Z. 2020. Investigation of fermentation conditions of biodiesel by-products for high production of beta-farnesene by an engineered Escherichia coli. Environ Sci Pollut Res Int 27 (18):22758-22769. DOI: 10.1007/s11356-020-08893-z.

[9] Xu D, Zhang Z, Liu Z, Xu Q. 2020. Using enzymatic hydrolyzate as new nitrogen source for L-tryptophan fermentation by E.coli. Bioengineered 11 (1):1-10. DOI: 10.1080/21655979.2019.1700092.

[10] Wang J, Liu Y, Chen Y, Wang A, Wei Q, Liu D, et al. 2020. Large-scale manufacture of VP2 VLP vaccine against porcine parvovirus in Escherichia coli with high-density fermentation. Appl Microbiol Biotechnol 104 (9):3847-3857. DOI: 10.1007/s00253-020-10483-5.

[11] Li D, Fang H, Gai Y, Zhao J, Jiang P, Wang L, et al. 2020. Metabolic engineering and optimisation of the fermentation medium for vitamin B12 production in Escherichia coli. Bioprocess Biosyst Eng 43 (10):1735-1745. DOI: 10.1007/s00449-020-02355-z.

[12] Hao Y, Ma Q, Liu X, Fan X, Men J, Wu H, et al. 2020. High-yield production of L-valine in engineered Escherichia coli by a novel two-stage fermentation. Metab Eng 62:198-206. DOI: 10.1016/j.ymben.2020.09.007.

[13] Agbogbo FK, Ramsey P, George R, Joy J, Srivastava S, Huang M, et al. 2020.

Upstream development of Escherichia coli fermentation process with PhoA promotdesignng design of experiments (DoE). J Ind Microbiol Biotechnol 47 (9–10):789-799. DOI: 10.1007/s10295-020-02302-7.

[14] Zhou HY, Wu WJ, Niu K, Xu YY, Liu ZQ, Zheng YG. 2019. Enhanced L-methionine production by genetically engineered Escherichia coli through fermentation optimisation. 3 Biotech 9 (3):96. DOI: 10.1007/s13205-019-1609-8.

[15] Zare H, Mir Mohammad Sadeghi H, Akbari V. 2019. Optimisation of Fermentation Conditions for Reteplase Expression by Escherichia coli Using Response Surface Methodology. Avicenna J Med Biotechnol 11 (2):162-168.

[16] Xu Y, Yang L, Zhao S, Wang Z. 2019. Large-scale production of tauroursodeoxycholic acid products through fermentation optimisation of engineered Escherichia coli cell factory. Microb Cell Fact 18 (1):34. DOI: 10.1186/s12934-019-1076-2.

[17] Tang XL, Wang ZC, Yang J, Zheng RC, Zheng YG. 2019. Statistical medium optimisation and DO-STAT fed-batch fermentation for enhanced production of tyrosine phenol lyase in recombinant Escherichia coli. Prep Biochem Biotechnol 49 (2):117-126. DOI: 10.1080/10826068.2018.1541808.

[18] Horner M, Gerhardt K, Salavei P, Hoess P, Harrer D, Kaiser J, et al. 2019. Production of Phytochromes by High-Cell-Density E. coli Fermentation. ACS Synth Biol 8 (10):2442-2450. DOI: 10.1021/acssynbio.9b00267.

[19] Fang Z, Sha C, Peng Z, Zhang J, Du G. 2019. Protein engineering to enhance keratinolytic protease activity and excretion in Escherichia coli and its scale-up fermentation for high extracellular yield. Enzyme Microb Technol 121:37-44. DOI: 10.1016/j.enzmictec.2018.11.003.

[20] Chen T, Liu N, Ren P, Xi X, Yang L, Sun W, et al. 2019. Efficient Biofilm-Based Fermentation Strategies for L-Threonine Production by Escherichia coli. Front Microbiol 10:1773. DOI: 10.3389/fmicb.2019.01773.

[21] Long J, Zhao X, Liang F, Liu N, Sun Y, Xi Y. 2018. Optimisation of fermentation conditions for an Escherichia coli strain engineered using the response surface method to produce a novel therapeutic DNA vaccine for rheumatoid arthritis. J Biol Eng 12:22. DOI: 10.1186/s13036-018-0110-y.

[22] Liang J, Zhao J, Wang Z, Wang Y. 2018. Temperature gradient-based high-cell density fed-batch fermentation for the production of pyruvate oxidase by recombinant E. coli. Prep Biochem Biotechnol 48 (2):188-193. DOI: 10.1080/10826068.2018.1425709.

[23] Koopaei NN, Khadiv-Parsi P, Khoshayand MR, Mazlomi MA, Kebriaeezadeh A, Moloudian H, et al. 2018. Optimisation of rPDT fusion protein expression by Escherichia coli in pilot scale fermentation: a statistical experimental design approach. AMB Express 8 (1):135. DOI: 10.1186/s13568-018-0667-3.

[24] Jing J, Chen Y, Sheng L, Wu M. 2018. Optimised production of insulin variant, a recombinant platelet aggregation inhibitor, by high cell-density fermentation of recombinant Escherichia coli. Protein Expr Purif 152: 7-12. DOI: 10.1016/j.pep.2018.07.001.

[25] Sar T, Seker G, Erman AG, Stark BC, Yesilcimen Akbas M. 2017. Repeated batch fermentation of immobilised E. coli expressing Vitreoscilla hemoglobin for long-term use. Bioengineered 8 (5):651-660. DOI: 10.1080/21655979.2017.1303024.

[26] Fu X, Wang Y, Wang J, Garza E, Manow R, Zhou S. 2017. Semi-industrial scale (30 m(3)) fed-batch fermentation for the production of D-lactate by

Escherichia coli strain HBUT-D15. J Ind Microbiol Biotechnol 44 (2):221-228. DOI: 10.1007/s10295-016-1877-9.

[27] Couto MR, Rodrigues JL, Rodrigues LR. 2017. Optimisation of fermentation conditions for the production of curcumin by engineered Escherichia coli. J R Soc Interface 14 (133). DOI: 10.1098/rsif.2017.0470.

[28] Zheng H, Yu Z, Fu X, Li S, Xu J, Song H, et al. 2016. High level extracellular production of a truncated alkaline beta-mannanase from alkaliphilic Bacillus sp. N16-5 in Escherichia coli by the optimisation of induction condition and fed-batch fermentation. J Ind Microbiol Biotechnol 43 (7):977-987. DOI: 10.1007/s10295-016-1773-3.

[29] Paopang P, Kasinrerk W, Tayapiwatana C, Seesuriyachan P, Butr-Indr B. 2016. Multiparameter optimisation method and enhanced production of secreted recombinant single-chain variable fragment against the HIV-1 P17 protein from Escherichia coli by fed-batch fermentation. Prep Biochem Biotechnol 46 (3):305-312. DOI: 10.1080/10826068.2015.1031388.

[30] Lu XX, Yi Y, Su QD, Bi SL. 2016. Expression and Purification of Recombinant Hepatitis Delta Virus (HDV) Antigen for Use in a Diagnostic ELISA for HDV Infection Using the High-Density Fermentation Strategy in Escherichia coli. Biomed Environ Sci 29 (6):417-423. DOI: 10.3967/bes2016.054.

[31] Huang L, Wang Q, Jiang S, Zhou Y, Zhang G, Ma Y. 2016. Improved extracellular expression and high-cell-density fed-batch fermentation of chitosanase from Aspergillus Fumigatus in Escherichia coli. Bioprocess Biosyst Eng 39 (11):1679-1687. DOI: 10.1007/s00449-016-1643-4.

[32] Sohoni SV, Nelapati D, Sathe S, Javadekar-Subhedar V, Gaikaiwari RP, Wangikar PP. 2015. Optimisation of high cell density fermentation process for recombinant nitrilase production in E. coli. Bioresour Technol 188: 202-208. DOI: 10.1016/j.biortech.2015.02.038.

[33] Morowvat MH, Babaeipour V, Rajabi Memari H, Vahidi H. 2015. Optimisation of Fermentation Conditions for Recombinant Human Interferon Beta Production by Escherichia coli Using the Response Surface Methodology. Jundishapur J Microbiol 8 (4):e16236. DOI: 10.5812/jjm.8(4)2015.16236.

[34] Minamoto T, Takahashi N, Kitahara S, Shinozaki Y, Hirano T, Hakamata W, et al. 2015. Saccharification of beta-chitin from squid pen by a fermentation method using recombinant chitinase-secreting Escherichia coli. Appl Biochem Biotechnol 175 (8):3788-3799. DOI: 10.1007/s12010-015-1547-9.

[35] Munoz-Gutierrez I, Moss-Acosta C, Trujillo-Martinez B, Gosset G, Martinez A. 2014. Ag43-mediated display of a thermostable beta-glucosidase in Escherichia coli and its use for simultaneous saccharification and fermentation at high temperatures. Microb Cell Fact 13:106. DOI: 10.1186/s12934-014-0106-3.

[36] Li Q, Mannall GJ, Ali S, Hoare M. 2013. An ultra scale-down approach to study the interaction of fermentation, homogenisation, and centrifugation for antibody fragment recovery from rec E. coli. Biotechnol Bioeng 110 (8): 2150-2160. DOI: 10.1002/bit.24891.

[37] Xu J, Qian Y, Skonezny PM, You L, Xing Z, Meyers DS, et al. 2012. Reduction of N-terminal methionylation while increasing titer by lowering metabolic and protein production rates in E. coli auto-induced fed-batch fermentation. J Ind Microbiol Biotechnol 39 (8):1199-1208. DOI: 10.1007/s10295-012-1127-8.

[38] Silva R, Ferreira S, Bonifacio MJ, Dias JM, Queiroz JA, Passarinha LA. 2012. Optimisation of fermentation conditions for the production of human soluble catechol-O-methyltransferase by Escherichia coli using artificial neural network. J Biotechnol 160 (3–4): 161-168. DOI: 10.1016/j.jbiotec.2012.03.025.

[39] Seo MJ, Choi HJ, Chung KH, Pyun YR. 2011. Production of a platelet aggregation inhibitor, salmosin, by high cell density fermentation of recombinant Escherichia coli. J Microbiol Biotechnol 21 (10):1053-1056. DOI: 10.4014/jmb.1105.05014.

[40] Premsukh A, Lavoie JM, Cizeau J, Entwistle J, MacDonald GC. 2011. Development of a GMP Phase III purification process for VB4-845, an immunotoxin expressed in E. coli using high cell density fermentation. Protein Expr Purif 78 (1):27-37. DOI: 10.1016/j.pep.2011.03.009.

[41] Honghong J, Fuping L, Yu L, Xiaoguang L, Yihan L, Hongbin W, et al. 2011. Synthesis of GDP-mannose using coupling fermentation of recombinant Escherichia coli. Biotechnol Lett 33 (6): 1145-1150. DOI: 10.1007/s10529-011-0547-2.

[42] Wang Z, Ly M, Zhang F, Zhong W, Suen A, Hickey AM, et al. 2010. E. coli K5 fermentation and the preparation of heparosan, a bioengineered heparin precursor. Biotechnol Bioeng 107 (6): 964-973. DOI: 10.1002/bit.22898.

[43] Brusehaber E, Schwiebs A, Schmidt M, Bottcher D, Bornscheuer UT. 2010. Production of pig liver esterase in batch fermentation of E. coli Origami. Appl Microbiol Biotechnol 86 (5):1337-1344. DOI: 10.1007/s00253-009-2392-y.

[44] Zhang D, Wang PG, Qi Q. 2007. A two-step fermentation process for efficient production of penta-N-acetyl-chitopentaose in recombinant Escherichia coli. Biotechnol Lett 29 (11): 1729-1733. DOI: 10.1007/s10529-007-9462-y.

[45] Kim M, Elvin C, Brownlee A, Lyons R. 2007. High yield expression of recombinant pro-resilin: lactose-induced fermentation in E. coli and facile purification. Protein Expr Purif 52 (1):230-236. DOI: 10.1016/j.pep.2006.11.003.

[46] Yazdani SS, Shakri AR, Chitnis CE. 2004. A high cell density fermentation strategy to produce recombinant malarial antigen in E. coli. Biotechnol Lett 26 (24):1891-1895. DOI: 10.1007/s10529-004-6040-4.

[47] Kleist S, Miksch G, Hitzmann B, Arndt M, Friehs K, Flaschel E. 2003. Optimisation of the extracellular production of a bacterial phytase with Escherichia coli by using different fed-batch fermentation strategies. Appl Microbiol Biotechnol 61 (5–6):456-462. DOI: 10.1007/s00253-003-1229-3.

[48] Khalilzadeh R, Shojaosadati SA, Bahrami A, Maghsoudi N. 2003. Over-expression of recombinant human interferon-gamma in high cell density fermentation of Escherichia coli. Biotechnol Lett 25 (23):1989-1992. DOI: 10.1023/b:bile.0000004390.98648.25.

[49] Panda AK, Ghorpade A, Mukhopadhyay A, Talwar GP, Garg LC. 1995. High cell density fermentation of recombinant Vibrio cholerae for the production of B subunit of Escherichia coli enterotoxin. Biotechnol Bioeng 45 (3):245-250. DOI: 10.1002/bit.260450309.

[50] Sakamoto S, Terada I, Iijima M, Matsuzawa H, Ohta T. 1994. Fermentation conditions for efficient production of thermophilic protease in Escherichia coli harboring a plasmid. Appl Microbiol Biotechnol 42 (4): 569-574. DOI: 10.1007/BF00173922.

[51] Kapralek F, Jecmen P, Sedlacek J, Fabry M, Zadrazil S. 1991. Fermentation conditions for high-level expression of the tac-promoter-controlled calf prochymosin cDNA in Escherichia coli HB101. Biotechnol Bioeng 37 (1):71-79. DOI: 10.1002/bit.260370111.

[52] Herber WK, Bailey FJ, Carty CE, Heimbach JC, Maigetter RZ. 1991. Production of cytotoxic proteins in Escherichia coli: a fermentation process for producing enzymatically active HIV-1 protease. Appl Microbiol Biotechnol 36 (2):149-152. DOI: 10.1007/BF00164410.

[53] Chang LL, Hwang LY, Hwang CF, Mou DG. 1991. Study of high density Escherichia coli fermentation for production of porcine somatotropin protein. Ann N Y Acad Sci 646: 259-272. DOI: 10.1111/j.1749-6632.1991.tb18589.x.

[54] Riesenberg D, Menzel K, Schulz V, Schumann K, Veith G, Zuber G, et al. 1990. High cell density fermentation of recombinant Escherichia coli expressing human interferon alpha 1. Appl Microbiol Biotechnol 34 (1):77-82. DOI: 10.1007/BF00170927.

[55] Jung G, Denefle P, Becquart J, Mayaux JF. 1988. High-cell density fermentation studies of recombinant Escherichia coli strains expressing human interleukin-1 beta. Ann Inst Pasteur Microbiol 139 (1):129-146. DOI: 10.1016/0769-2609(88)90100-7.

[56] Botterman JH, De Buyser DR, Spriet JA, Zabeau M, Vansteenkiste GC. 1985. Fermentation and recovery of the EcoRI restriction enzyme with a genetically modified Escherichia coli strain. Biotechnol Bioeng 27 (9): 1320-1327. DOI: 10.1002/bit.260270908.

[57] Zhang J, Greasham R. 1999. Chemically defined media for commercial fermentations. Applied microbiology and biotechnology 51 (4): 407-421. DOI: 10.1007/s002530051411.

[58] Nikel PI, Pettinari MJ, Galvagno MA, Méndez BS. 2008. Poly(3-hydroxybutyrate) synthesis from glycerol by a recombinant Escherichia coli arcA mutant in fed-batch microaerobic cultures. Applied microbiology and biotechnology 77 (6):1337-1343. DOI: 10.1007/s00253-007-1255-7.

[59] Singleton C, Gilman J, Rollit J, Zhang K, Parker DA, Love J. 2019. A design of experiments approach for the rapid formulation of a chemically defined medium for metabolic profiling of industrially important microbes. PLoS One 14 (6):e0218208. DOI: 10.1371/journal.pone.0218208.

[60] Singh V, Haque S, Niwas R, Srivastava A, Pasupuleti M, Tripathi CK. 2016. Strategies for Fermentation Medium Optimisation: An In-Depth Review. Front Microbiol 7:2087. DOI: 10.3389/fmicb.2016.02087.

[61] Weuster-Botz D. 2000. Expedesignal design for fermentation media development: Stadesignal design or global random search? Journal of Bioscience and Bioengineering 90 (5): 473–83. DOI: https://doi.org/10.1016/S1389-1723(01)80027-X.

[62] Weissman SA, Anderson NG. 2015. Design of Experiments (DoE) and Process Optimisation. A Review of Recent Publications. Organic Process Research & Development 19 (11): 1605-1633. DOI: 10.1021/op500169m.

[63] Dvorak P, Chrast L, Nikel PI, Fedr R, Soucek K, Sedlackova M, et al. 2015. Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21(DE3) carrying a synthetic metabolic pathway. Microbial Cell Factories 14 (1):201. DOI: 10.1186/s12934-015-0393-3.

[64] Cardoso VM, Campani G, Santos MP, Silva GG, Pires MC, Gonçalves VM, et al. 2020. Cost analysis based on bioreactor cultivation conditions: Production of a soluble

recombinant protein using Escherichia coli BL21(DE3). Biotechnology Reports 26:e00441. DOI: https://doi.org/10.1016/j.btre.2020.e00441.

[65] Siegele DA, Hu JC. 1997. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. Proceedings of the National Academy of Sciences 94 (15):8168-8172. DOI: 10.1073/pnas.94.15.8168.

[66] Choi YJ, Morel L, Le François T, Bourque D, Bourget L, Groleau D, et al. 2010. Novel, versatile, and tightly regulated expression system for Escherichia coli strains. Applied and environmental microbiology 76 (15): 5058-5066. DOI: 10.1128/AEM.00413-10.

[67] Mujacic M, Cooper KW, Baneyx F. 1999. Cold-inducible cloning vectors for low-temperature protein expression in Escherichia coli: application to the production of a toxic and proteolytically sensitive fusion protein. Gene 238 (2): 325-332. DOI: 10.1016/s0378-1119(99)00328-5.

[68] Nor NM, Mohamed MS, Loh TC, Foo HL, Rahim RA, Tan JS, et al. 2017. Comparative analyses on medium optimisation using one-factor-at-a-time, response surface methodology, and artificial neural network for lysine–methionine biosynthesis by Pediococcus pentosaceus RF-1. Biotechnology & Biotechnological Equipment 31 (5): 935-947. DOI: 10.1080/13102818.2017.1335177.

[69] Elbing KL, Brent R. 2019. Recipes and Tools for Culture of Escherichia coli. Current protocols in molecular biology 125 (1):e83-e. DOI: 10.1002/cpmb.83.

[70] Zhang C, Chen X, Zou R, Zhou K, Stephanopoulos G, Too H-P. 2013. Combining Genotype Improvement and Statistical Media Optimization for Isoprenoid Production in E. coli. PLoS One 8 (10):e75164. DOI: 10.1371/journal.pone.0075164.

[71] Zhu LW, Wang CC, Liu RS, Li HM, Wan DJ, Tang YJ. 2012. Actinobacillus succinogenes ATCC 55618 fermentation medium optimisation for the production of succinic acid by response surface methodology. J Biomed Biotechnol 2012: 626137. DOI: 10.1155/2012/626137.

[72] Riesenberg D, Schulz V, Knorre WA, Pohl HD, Korz D, Sanders EA, et al. 1991. High cell density cultivation of Escherichia coli at controlled specific growth rate. J Biotechnol 20 (1):17-27. DOI: 10.1016/0168-1656(91)90032-Q.

[73] Ghaly AE, Kamal M, Correia LR. 2005. Kinetic modelling of continuous submerged fermentation of cheese whey for single cell protein production. Bioresour Technol 96 (10):1143-1152. DOI: 10.1016/j.biortech.2004.09.027.

[74] Amro AA, Salem-Bekhit MM, Alanazi FK. 2014. Plackett-Burman randomisation method for Bacterial Ghosts preparation form E. coli JM109. Saudi Pharm J 22 (3):273-279. DOI: 10.1016/j.jsps.2013.06.002.

[75] Li H, van den Driesche S, Bunge F, Yang B, Vellekoop MJ. 2019. Optimisation of on-chip bacterial culture conditions using the Box-Behnken design response surface methodology for faster drug susceptibility screening. Talanta 194: 627-33. DOI: https://doi.org/10.1016/j.talanta.2018.10.048.

[76] Aggarwal A, Singh H. 2005. Optimisation of machining techniques — A retrospective and literature review. Sadhana 30 (6):699-711. DOI: 10.1007/BF02716704.

[77] Ranjbari J, Babaeipour V, Vahidi H, Moghimi H, Mofid MR, Namvaran MM, et al. 2015. Enhanced Production of Insulin-like Growth Factor I Protein in Escherichia coli by Optimisation of Five

Key Factors. Iran J Pharm Res 14 (3): 907-917.

[78] Sriwidodo S, Subroto T, Maksum I, Wathoni N, Rostinawati T, Ulya H, et al. 2019. Optimisation of secreted recombinant human epidermal growth factor production using pectate lyase B from *Escherichia coli* BL21(DE3) by central composite design and its production in high cell density culture. Journal of Pharmacy And Bioallied Sciences 11 (8):562-566. DOI: 10.4103/jpbs.JPBS_207_19.

[79] Rhee JI, Kang T-H. 2007. On-line process monitoring and chemometric modeling with 2D fluorescence spectra obtained in recombinant E. coli fermentations. Process Biochemistry 42 (7):1124-34. DOI: https://doi.org/10.1016/j.procbio.2007.05.007.

[80] Tai M, Ly A, Leung I, Nayar G. 2015. Efficient high-throughput biological process characterisation: Definitive screening design with the ambr250 bioreactor system. Biotechnol Prog 31 (5):1388-1395. DOI: 10.1002/btpr.2142.

[81] Feyzdar M, Vali AR, Babaeipour V. 2013. Identification and Optimisation of Recombinant E. coli Fed-Batch Fermentation Producing γ-Interferon Protein. International Journal of Chemical Reactor Engineering 11 (1): 123-134. DOI: doi:10.1515/ijcre-2012-0081.

[82] Zhang C, Fan D, Shang La, Ma X, Luo Ye, Xue W, et al. 2010. Optimisation of Fermentation Process for Human-like Collagen Production of Recombinant Escherichia coli Using Response Surface Methodology. Chinese Journal of Chemical Engineering 18 (1): 137-42. DOI: https://doi.org/10.1016/S1004-9541(08)60334-1.

[83] Aydar AY. 2018. Utilisation of response surface methodology in optimisation of extraction of plant materials. Statistical approaches with emphasis on design of experiments applied to chemical processes:157-169.

[84] Desai KM, Survase SA, Saudagar PS, Lele SS, Singhal RS. 2008. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimisation: Case study of fermentative production of scleroglucan. Biochemical Engineering Journal 41 (3):266-73. DOI: https://doi.org/10.1016/j.bej.2008.05.009.

[85] Kubilius J, Schrimpf M, Kar K, Hong H, Majaj NJ, Rajalingham R, et al. 2019. Brain-like object recognition with high-performing shallow recurrent ANNs. arXiv preprint arXiv:190906161.

[86] Mishra M, Srivastava M, editors. A view of artificial neural network. 2014 International Conference on Advances in Engineering & Technology Research (ICAETR-2014); 2014: IEEE.

[87] Khan S, Bhakuni V, Praveen V, Tewari R, Tripathi CKM, Gupta VD. 2011. Maximising the native concentration and shelf life of protein: a multiobjective optimisation to reduce aggregation. Applied microbiology and biotechnology 89 (1):99-108. DOI: 10.1007/s00253-010-2835-5.

[88] Beg AH, Islam MZ, editors. Advantages and limitations of genetic algorithms for clustering records. 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA); 2016: IEEE.

[89] Lagarias JC, Poonen B, Wright MH. 2012. Convergence of the Restricted Nelder–Mead Algorithm in Two Dimensions. SIAM Journal on Optimization 22 (2):501-532.

[90] Singer S, Nelder J. 2009. Nelder-mead algorithm. Scholarpedia 4 (7):2928.

[91] Gupta PK, Mukherjee P, Dhawan S, Pandey AK, Mazumdar S, Gaur D, et al.

2014. Production and preclinical evaluation of Plasmodium falciparum MSP-119 and MSP-311 chimeric protein, PfMSP-Fu24. Clinical and vaccine immunology : CVI. DOI: 10.1128/CVI.00179-14.

[92] Favre-Bulle O, Schouten T, Kingma J, Witholt B. 1991. Bioconversion of n-octane to octanoic acid by a recombinant Escherichia coli cultured in a two-liquid phase bioreactor. Biotechnology (N Y) 9 (4): 367-371.

[93] Sabatie J, Speck D, Reymund J, Hebert C, Caussin L, Weltin D, et al. 1991. Biotin formation by recombinant strains of Escherichia coli: influence of the host physiology. J Biotechnol 20 (1): 29-49.

[94] Wagner CR, Benkovic SJ. 1990. Site directed mutagenesis: a tool for enzyme mechanism dissection. Trends Biotechnol 8 (9):263-270.

[95] Kane JF. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr Opin Biotechnol 6 (5): 494-500.

[96] Kurland C, Gallant J. 1996. Errors of heterologous protein expression. Curr Opin Biotechnol 7 (5):489-493.

[97] Roche ED, Sauer RT. 1999. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. EMBO J 18 (16):4579-4589. DOI: 10.1093/emboj/18.16.4579.

[98] Zahn K. 1996. Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth. J Bacteriol 178 (10): 2926-2933.

[99] Gouy M, Gautier C. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res 10 (22): 7055-7074.

[100] Zhang SP, Zubay G, Goldman E. 1991. Low-usage codons in Escherichia coli, yeast, fruit fly and primates. Gene 105 (1):61-72.

[101] Prapunwattana P, Sirawaraporn W, Yuthavong Y, Santi DV. 1996. Chemical synthesis of the Plasmodium falciparum dihydrofolate reductase-thymidylate synthase gene. Mol Biochem Parasitol 83 (1):93-106.

[102] Mazumdar S, Sachdeva S, Chauhan VS, Yazdani SS. 2010. Identification of cultivation condition to produce correctly folded form of a malaria vaccine based on Plasmodium falciparum merozoite surface protein-1 in Escherichia coli. Bioprocess and biosystems engineering 33 (6):719-730. DOI: 10.1007/s00449-009-0394-x.

[103] Baca AM, Hol WG. 2000. Overcoming codon bias: a method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. Int J Parasitol 30 (2):113-118.

[104] Sano G, Morimatsu K, Horii T. 1994. Purification and characterisation of dihydrofolate reductase of Plasmodium falciparum expressed by a synthetic gene in Escherichia coli. Mol Biochem Parasitol 63 (2):265-273.

[105] Cinquin O, Christopherson RI, Menz RI. 2001. A hybrid plasmid for expression of toxic malarial proteins in Escherichia coli. Mol Biochem Parasitol 117 (2):245-247.

[106] Lindenthal C, Klinkert MQ. 2002. Identification and biochemical characterisation of a protein phosphatase 5 homologue from Plasmodium falciparum. Mol Biochem Parasitol 120 (2):257-268.

[107] Hoffman SL, Rogers WO, Carucci DJ, Venter JC. 1998. From genomics to vaccines: malaria as a model

system. Nat Med 4 (12):1351-1353. DOI: 10.1038/3934.

[108] Ryan W, Parulekar SJ. 1991. Recombinant protein synthesis and plasmid instability in continuous cultures of Escherichia coli JM103 harboring a high copy number plasmid. Biotechnol Bioeng 37 (5):415-429. DOI: 10.1002/bit.260370504.

[109] Hellmuth K, Korz DJ, Sanders EA, Deckwer WD. 1994. Effect of growth rate on stability and gene expression of recombinant plasmids during continuous and high cell density cultivation of Escherichia coli TG1. J Biotechnol 32 (3):289-298.

[110] Gupta R, Sharma P, Vyas VV. 1995. Effect of growth environment on the stability of a recombinant shuttle plasmid, pCPPS-31, in Escherichia coli. J Biotechnol 41 (1):29-37.