Design of Experiments for the Development of Biotechnology Products

10

Suryakanta Swain, Bikash Ranjan Jena, and Sarwar Beg

Abstract

The development of biotechnology products containing large molecules such as nucleic acids, monoclonal antibodies, hormones, etc., requires critical understanding of the product-specific quality attributes and close monitoring of the manufacturing process for attaining quality consistency without batch-to-batch variability. The conventional product development strategy using one factor at one time approach only provides limited solution without any significant influence on the product quality for attaining the robust performance. However, the systematic approach of using design of experiments (DoE) is very useful for minimizing the variability and improving the product performance. A score of literature reports are available which vouch the applicability of experimental designs for efficient development of the biotechnology products. The present chapter, therefore, provides a holistic account on the implementation of DoE approach for manufacturing of the biotechnology, and also highlights the current challenges and opportunity associated with them.

Keywords

Quality · Variability · Product development · Experimental designs · Robustness

S. Swain (✉)
Department of Pharmaceutical Sciences, School of Health Sciences, The Assam Kaziranga University, Jorhat, Assam, India

B. R. Jena
School of Pharmacy and Life Sciences, Centurion University of Technology and Management (CUTM), Bhubaneswar, Odisha, India

S. Beg
Department of Pharmaceutics, School of Pharmaceutical Education and Research, Jamia Hamdard, New Delhi, India

S. Beg (ed.), *Design of Experiments for Pharmaceutical Product Development*,
https://doi.org/10.1007/978-981-33-4351-1_10

10.1 Introduction

Statistical experimental planning, factorial design, and design of experiments (DoE) are synonymous concepts for investigating the mathematical relationships between input and output variables of a system. Although the fundamentals of the methodology have been known since the early 1900s, 1–3, it was not until recent years that it was widely applied in biotechnology. Methods such as factorial design, response surface methodology, and (DoE) provide robust and efficient ways to optimize cultivations and other unit operations and procedures using a reduced number of experiments. The multitude of interdependent parameters involved within a unit operation or between units in a bioprocess sequence may be substantially refined and improved by using such methods [1]. In an experiment, one or more process variables (or factors) are deliberately changed to observe the effect on one or more response variables. Statistical design of experiments is an efficient procedure for planning operations and analyzing the results such that objective conclusions can be drawn. Determining the objectives of analysis is the first step of experimental design. Selecting the process factors and responses to be analyzed is the second important step in optimizing the essential variables in a process [2].

In general, experimental design is the design of a system to obtain information using various innovative techniques. Experimental design can be used to evaluate physical objects, chemical formulations, structures, components, and materials. Bioengineering studies different production areas, including enzyme production, biological water treatment, tissue culture, nanoparticles, biofuel production, the isolation of microorganisms, and the industrial biological production of compounds such as proteins, lipids, and aromatic compounds. Every process is affected by different factors. Optimizing these parameters is an essential step in obtaining maximum production with minimum costs. Developing an experimental design is the most effective means of process optimization. This review summarizes the experimental design methods that are used to investigate the effects of various factors on various bioengineering processes, including full factorial design, fractional factorial design, Plackett–Burman (PB) design, Taguchi design, Box–Behnken design, and central composite design. Each design method is briefly introduced, the advantages and disadvantages of each are briefly discussed, and various processes are analyzed [3].

Other bioprocess-related applications include strain screening evaluation and cultivation media balancing. Because of the emerging regulatory demands on pharmaceutical manufacturing processes, exemplified by the process analytical technology (PAT) initiative of the United States Food and Drug Administration, the use of experimental design approaches improves process development for safer and more reproducible production is becoming increasingly important. Here, these options are highlighted and discussed with a few selected examples from antibiotic fermentation, expanded bed optimization, virus vector transfection of insect cell cultivation, feed profile adaptation, embryonic stem cell expansion protocols, and mammalian cell harvesting [4].

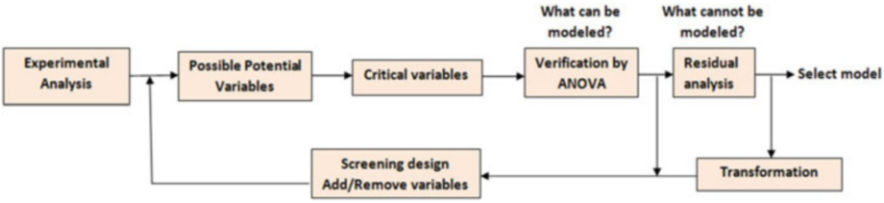

Fig. 10.1 Different critical steps for experimental design (Adapted from [5])

Biotechnology processes can be considered a transformation of nutrients and other medium components to biomass and molecular ones. The utilization of carbon and nitrogen sources is relatively quickly overviewed via total stoichiometric balances. To this, a degree of complexity is added by shifts in metabolism caused by activating or deactivating medium components such as growth, vitamins, and microbial stressors. Figure 10.1 illustrates the situation where all inputs exert their effects as an initial load of multiplicity effects onto the transforming bioprocess in a steady-state condition. However, in reality, the steady-state is unstable since the nutrient factors are degraded and depleted [6].

In most bioprocesses, products are recovered downstream. Additional factors, e.g., eluents, are sequentially supplied to purify the product. Here, too, input factors are distributed over time. The influence of upstream procedures follows the same pattern. Thus, any attempt to encompass all elements in a bioprocess ends up in an incomprehensible number of interacting and non-interacting factors that may, or may not, affect the specific outputs [7].

10.2 Efficient Strategy, Concept, and Parameters of Design of Experiments

Several simulated research problems have been programmed for the IBM 704 computer. Six different empirical research strategies have also been programmed for the computer. The "research problems" are "solved" by the computer using one of the six strategies to learn which is best. The computer uses each research strategy on each problem several times to get a statistical evaluation of its effectiveness [8].

Application of the design of experiments allows investigators to understand multiple method parameters and variables that tend to impact critical responses while unraveling the occurrence of (any) interactions and diminishing complexities ([9]). For the successful implementation of the experimental study design, the knowledge of response variables, critical method variables, their ranges, and suitable mathematical model(s) is mandatory. Response surface methodology (RSM) based various design of experiments like factorial design (FD), central composite design (CCD), Box–Behnken design (BBD), optimal design (OD), etc., are useful in the systematic development of analytical methods comprising significant variable-response relationship(s) [8]. The experimental designs assist in mapping the

responses based on the studied objective(s) and exploring the critical responses at high (coded as +1), medium (coded as 0), or low (coded as −1) levels of the variables. It tends to reveal the mechanistic understanding of the variables-responses relationship and their associated interactions via various pictorial/graphical tools. 3D and 2D-plots like response surface plots, contour plots, perturbation plots, linear correlation plots, outlier plots, and Box–Cox plot are beneficial in this regard ([9]). Once data acquired by the chosen design has been collected, the results can be analyzed using statistical methods like multiple linear regression (MLR) analysis so that objective conclusions can be drawn [10]. The flow layout of a typical design of the experiments-based regression model employed for method development and data analysis is illustrated below (Fig. 10.1).

10.3 Potential Benefits and Types of Experimental Design Used for the Development of Biotechnology Products

The design of experiments (DOE) based approach offers a solution to this problem and allows for efficient estimation of the main effects and the interactions with minimal operations. A systematic approach for evaluating the different DOE designs and choosing the optimal model (I-optimal and D-optimal), central composite, and Box–Behnken designs is mainly used for the development of biotechnology products. Other potential benefits and types of DoE-based software's for the development of biotechnology products are as follows [8, 10];

- Better understanding and control over critical variables
- Beyond traditional approach of method validation
- Flexibility in the analysis of the sample in various matrices
- Improved method robustness thereby reduction of variability
- Analytical attributes within the pharmacopeia restrictions, and away from Out-of-Specification (OOS) limits
- Smooth method transfers to the production level
- No obligation of revalidation within the design space.

10.4 Role of Central Composite Designs (I- and D-Optimal) and Box–Behnken Designs for Bioprocess Optimization

The schematic classifications of experimental design are shown in Fig. 10.2 and briefly described the central composite designs (I- and D-optimal) and Box–Behnken designs used for bioprocess optimization techniques as below.

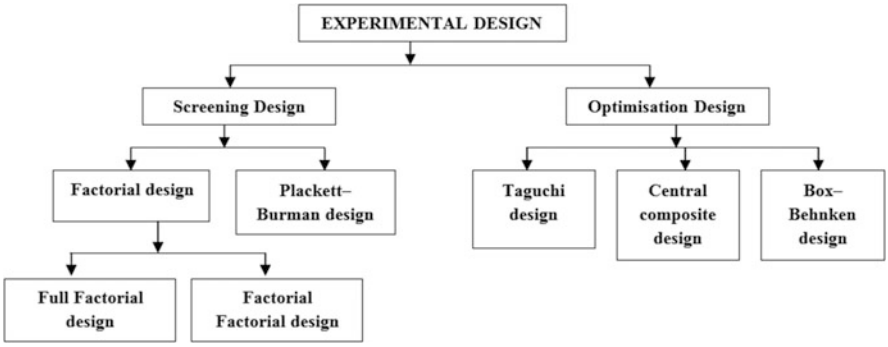

Fig. 10.2 Classification of experimental designs (Adapted from [5])

10.4.1 Central Composite Design

Box and Wilson developed a central composite design (CCD). Both linear and quadric models are allowed to be determined by this design. CCD seems to be a good alternative of a three-level full factorial design as it provides comparable results with a smaller number of experiments. CCD usually consists of a full factorial design or factorial design with two levels, additional axial or star points, and at least one central point of the experimental design. The axial design and the central design are almost the same for the two-level full factorial design except for one factor that may take on levels either above the high level or below the low level. CCD requires experiment numbers according to $N = k2 + 2k + \text{cp}$, here cp is replicate numbers of the center point, and k is factor numbers [1, 11–15].

10.4.2 Box–Behnken Design

Box and Behnken developed Box–Behnken design (BBD). This design consists of a factorial design with three levels and an incomplete block design in such a way to present as a rotatable or nearly rotatable design and to avoid the extreme vertices. BBD requires experiment numbers based on $N = 2k\ (k - 1) + \text{C0}$, here k is factor numbers, and C0 is central point numbers. BBD is useful in avoiding experiments that are in extreme conditions because the highest level and lowest level combinations for every factor cannot be included in BBD. Unsatisfactory results might be avoided in BBD [16–18]. ANOVA can examine the significant effects of BBD on the response, and the optimal response can be determined by the regression model with calculating the derivatives of the model. CCD has more factor levels than BBD. Thus BBD can be an economical alternative of CCD [19, 20].

10.5 DoE for Early Bioprocess Development or Modeling, Optimization, and Characterization

The universal implementation of DoE methodology to bioprocesses development is uncomplicated and unique outline of development. It investigates defined input factors to a converting bio system from which mostly common and well-defined output factors or responses are generated, such as product yield and productivity. The statistical design of experiments (DOE) is a collection of predetermined settings of the process variables of interest, which provides an efficient procedure for planning experiments. Experiments on biological processes typically produce long sequences of successive observations on each experimental unit (plant, animal, bioreactor, fermenter, or flask) in response to several treatments (combination of factors). Cell culture and other biotech-related experiments used to be performed by repeated-measures method of experimental design coupled with different levels of several process factors to investigate dynamic biological process. Data collected from this design can be analyzed by several kinds of general linear model (GLM) statistical methods such as multivariate analysis of variance (MANOVA), univariate ANOVA (time split-plot analysis with randomization restriction), and analysis of orthogonal polynomial contrasts of repeated factor (linear coefficient analysis). Statistical analysis of bioprocess with repeated measurements can help investigate environmental factors and effects affecting physiological and biochemical processes in analyzing and optimizing biotechnology based production. However, the strength of DoE is that it also reveals how interactions between the input factors influence the output responses. These interactions are often difficult to discover and interpret with other methods. The basic outline or layout containing different input factors, output responses, and stages involved in bioprocess and its upstream stages of experimental design model are depicted in Fig. 10.3 [21–23].

10.6 Statistical Analysis of Variance of Bioprocess for Analyzing and Optimizing Biotechnology Production

Cell culture and other biotech-related experiments used to be performed by repeated-measures technique of DoE coupled with diverse levels of several process factors to investigate dynamic biological process. Statistical analysis of bioprocess with repeated measurements can help investigate environmental factors and effects affecting physiological and bioprocesses in analyzing and optimizing biotechnology production. Data collected from this design can be analyzed by several kinds of general linear model (GLM) statistical methods such as multivariate analysis of variance (MANOVA), univariate ANOVA (time split-plot analysis with randomization restriction), and analysis of orthogonal polynomial contrasts of repeated factor (linear coefficient analysis). The statistical design of experiments (DOE) is a collection of predetermined settings of the process variables of interest, which provides an efficient procedure for planning experiments. Experiments on biological processes typically produce long sequences of successive observations on each experimental

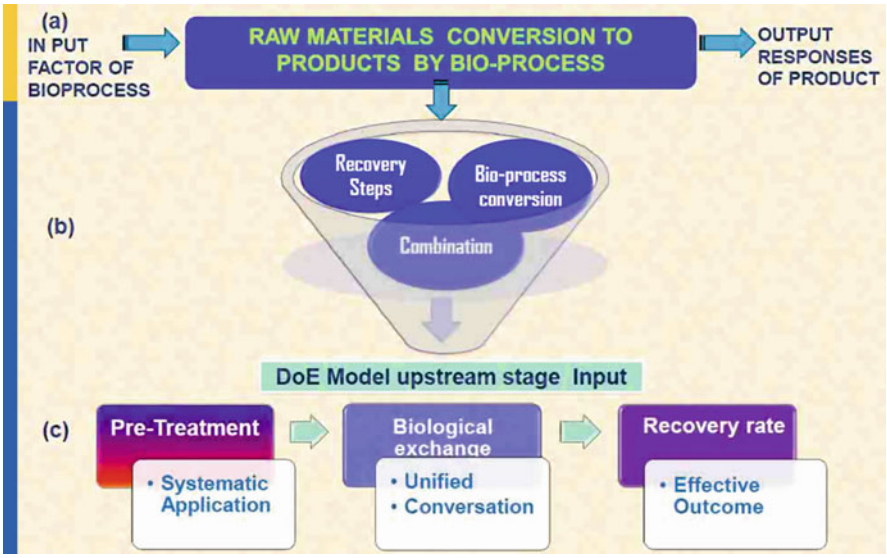

Fig. 10.3 Schematic diagram elucidating (**a**) Mechanism of input coded factors and output responses of a systematic bioprocess system as a basis for DoE, (**b**) Different steps involved in a typical bioprocess methodology: bioreactor and recovery (route of downstream processing), (**c**) Bioprocesses system upstream stage input factors be measured that add on further complication to the DoE models

unit (plant, animal, bioreactor, fermenter, or flask) in response to several treatments (combination of factors) [21, 24–26].

10.7 Application of DoE Methods Through the Use of Different Statistical Software's in Biotechnology

10.7.1 In Expanded Bed Optimization

In view of the emerging regulatory demands on pharmaceutical manufacturing processes, exemplified by the process analytical technology (PAT) initiative of the United States Food and Drug Administration, the use of experimental design approaches to improve process development for safer and more reproducible production is becoming increasingly important. Methods such as factorial design, response surface methodology, and (DoE) provide powerful and efficient ways to optimize cultivations and other unit operations and procedures using a reduced number of experiments. Design of experiments (DOE) based approach offers a solution to this conundrum and allows for an efficient estimation of the main effects and the interactions with minimal number of experiments. Statistical experimental planning, factorial design, and design of experiments (DoE) are more or less

synonymous concepts for investigating the mathematical relationships between input and output variables of a system [7].

10.7.2 In Antibiotic or Microbial Fermentation

Most notable applications of DoE have concerned optimization of the composition of growth and production culture media or microbial fermentation techniques. An elucidating example of this is from the production of the antibiotic clavulanic acid using the fungus *Streptomyces clavuligerus*. Wang et al. [27] optimized the medium composition by first screening a variety of media ingredients by a two-level fractional factorial design approach which subsequently was followed by optimizing their levels by response surface methodology [27]. Other successful examples of applying DoE to optimization of media composition for antibiotics production come from neomycin production by *S. marinensis* with solid-state fermentation [28] nisin by *Lactococcus lactis*, and meilingmycin by *S. nanchangensis*, [29–33] where fractional design methodology in combination with response surface was applied. Overviews of DoE methodology applicable for the optimization of microbial fermentation process and optimization of culture media of some vital antibiotics are enlisted in Table 10.1.

10.7.3 In Virus Vector Transfection of Insect Cell Cultivation

The cultivation of recombinant insect cell lines has been the subject of intense research since the 1980s and also allows the industrial production of recombinant proteins, vaccines, and insecticides. Gene therapy is a promising technology for the treatment of several acquired and inherited diseases. However, for gene therapy to be a commercial and clinical success, scalable cell culture processes must be in place to produce the required amount of viral vectors to meet market demand [34]. However, optimization of transformation conditions is carried out using OFAT approaches. Even though the optimization of transformation conditions for cloning experiments using DoE has not been reported, fractional factorial approaches have been successfully employed to identify the factors that significantly affect the transformation efficiency of bacteria for other purposes (e.g., drug development) [35]. evaluated the effect of five factors (Cell density, voltage, resistance, plasmid DNA concentration, and Mg^{2+} concentration) on the transformation efficiency of *Acinetobacter baumannii* using a three-level fractional factorial approach and the transformation efficiency was increased by four times. Thus, DoE approaches could be probably used as a tool to maximize the transformation efficiency of bacteria during cloning experiments. Each type of vector has its own distinct characteristics and consequently its own challenges for production [36].

Table 10.1 Applications DoE methodology for production of antibiotics fermentation and their responses outcome

Name of antibiotics	Rationale of methodology	Optimization process	Outcome and improvement	Reference
Neomycin (*S. marinensis*)	Medium composition maximizing neomycin yield by optimizing	Full FD and RSM	Yield of neomycin, optimum determined	[28]
Nisin (*Lactococcus lactis*)	Optimizing medium composition Maximizing nisin specific productivity by	statistical analysis Fractional factorial design and	productivity Nisin yield; specific. Increased activity yield	[31]
Clavulanic acid (*Streptomyces clavuligerus*)	Medium composition Maximizing clavulanic acid yield by optimizing	Optimizing by RSM Screening by fractional FD and	Yield of clavulanic acid with 50%	[27]
Meilingmycin	Medium composition maximizing meilingmycin yield by optimizing	Optimization by steepest ascent and RSM screening by fractional FD and	4.5-fold increase of Melingomycin yield	[33]
6-Aminopenicillanic acid (penicillin acylase)	Process parameters Maximizing 6-APA yield by optimizing	Full FD and modeling	Optimization has made by complete enzyme yield	[29], [32]
Antifungal antibiotic (*S. chattanoogensis*)	Process parameters Maximizing antibiotic yield by optimizing	Full FD and statistical analysis	Optimization has made by complete enzyme yield	[30]

RSM Response surface methodology, *FD* Factorial design

10.7.4 In Feed Profile Adaptation

Design of Experiment especially concerning the biological sciences, in particular for food microbiology, we can define DoE as a structured, systematic, and rigorous approach to problem solving that applies principles and techniques at the data collection stage, so as to ensure the generation of valid, defensible, and supportable conclusions. DoE can be successfully being implemented in feed profile adaptation. As the PAT initiative emphasizes, the use of DoE to exploit new optimization possibilities in biotechnology can be a very useful resource for bioprocess development. Especially the PAT is used to investigate in a more systematic manner the

factors that affect intrinsic responses in the cells/product protein of the bioprocess, such as glycosylation pattern and other modifications. Here, data mining techniques may provide a useful resource of analytical evaluation methods [37].

10.7.5 In Embryonic Stem Cell Expansion Protocols, and Mammalian Cell Harvesting

The development of new culture production systems for embryonic stem (ES) cells requires substantial work to find suitable production conditions. This is a particular concern in the pre-processing of time-consuming and labor-demanding in vitro differentiation procedures, where the complexities of factors that impact upon ES cell differentiation are profound. Chang and Zandstra [38] have developed and validated such a technology for ES cell differentiation analysis. They used a quantitative screening platform based on automated fluorescence microscopy, which enumerated the ES cells that had entered endodermal differentiation through expression of two biomarkers (Cytokeratin-8 and hepatocyte nuclear factor 3β). Chang and Zandstra [38] developed two-level fractional factorial design model based on 32 triplicate experiments, they screened important medium components for the differentiation process to endodermal cells (glucose, insulin, basic fibroblast growth factor, retinoic acid, and epidermal growth factor) with the biomarkers and cell numbers as responses. The model was further refined using a subsequent three-level factorial experiment for two of the factors. A statistical regression model was used to identify major and interactive effects on the endoderm formation. Retinoic acid was found to have an inhibitory effect on endoderm formation, while low glucose levels were beneficial. DoE proved to be a powerful tool for studying the factors impacting endoderm-specific ES cell differentiation; but it does require a relevant and sufficiently sensitive technique for the analysis of responses.

10.7.6 In Cancer Targeting and Gene Delivery

Implementation of a manufacturing process that assures a predefined quality of product is a critical requirement for the licensing and marketing of every cell and gene therapy (CGT) product. However, inadequate process knowledge and understanding constricts implementation of process changes as the impacts on product safety and efficacy are unknown. This often leads to the adoption of processes that, although compliant with established regulations, are not optimal for assuring broad availability to patients who depend on those therapies [39]. To improve the manufacturing of CGTs, Quality by Design (QbD) principles widely recognized as integration of scientific knowledge and risk assessment into process and build products can be adopted. Mostly the retro virus, Adeno virus, herpes viruses are used in gene therapy applications. QbD approach will involve development of an effective control strategy for it would ensure safe encapsulation of anticancer drugs for successful product development using polymeric nanoparticles. Chien et al. [40]

have developed a comparison technique where he has discussed about paclitaxel (PTX), a common chemotherapeutic agent, was loaded into poly-lactic-co-glycolic acid (PLGA) nanoparticles (NPs), and the coating process of chitosan (CS) onto PLGA NPs was focused for optimization by d-optimal designs and Artificial Neural Networks for targeting lungs and cervical cancer cells.

10.7.7 In Recombinant Protein Expression

Quality by Design (QbD) is a new approach to the development of recombinant therapeutic protein products that promotes a better understanding of the product and its manufacturing process [41]. Recombinant proteins are widely used in diverse fields in laboratory and industry, while many applications require sufficient amounts of high-quality proteins in terms of purity and activity. Viral transduction of eukaryotic cell lines is one possibility to efficiently generate recombinant proteins, and among all viral-based expression systems the baculovirus/insect cell expression system (BEVS) is certainly the most well-known and applied. In order to produce recombinant proteins, cost effectively, a satisfactory expression level has to be achieved in one of several species available for recombinant protein expression. Suitable hosts include bacteria (*Escherichia coli*), yeasts (such as *Pichia pastoris*), and cell lines of mammalian or insect origin [34]. These expression systems differ in terms of complexity, space-time-yield, and the ability to support protein folding and posttranslational modification. Several designs like full factorial, fractional factorial, Plackett–Burman, Taguchi orthogonal, RSM, and incomplete fractional are implemented in recombinant protein biotechnology. The design has so much advantages like both categorical and continuous variables can be simultaneously tested by full factorial design, results from the whole set of experiments are utilized. The Plackett–Burman method screens the design space to detect a large main effect. Taguchi orthogonal is a highly fractional orthogonal design allowing examining a selected subset of combinations of multiple factors at multiple levels with the fewest number of experiments. RSM can be useful for the only "real" optimization process and can be used to fine-tune the optimum conditions. Incomplete fractional is highly considered and can easily be set up using a freeware package (SAmBA)—Advance knowledge in statistics is not necessary—Factors can be examined in more than two levels. The other crucial steps include PRC amplification, Ligation, Purification, Functional assay, and Protein Crystallography, etc. [42].

10.8 DoE-Based Reported Bioprocess Case Studies

Mendes et al. [43] investigated that the daily relative growth rates of the red macroalgae *Gracilaria domingensis* in synthetic seawater. The effect of the combined influence of five factors [light (L), temperature (T), nitrate (N), phosphate (P), and molybdate (M) concentrations] was determined using a full factorial design. The ranges of the experimental cultivation conditions used were as follows: T, 18–26

C; L, 74–162 micromol photons m-2 s-1; N, 40–80 mmol/L; P, 8–16 mmol/L; and M, 1–5 nmol/L. The results were analyzed using the analysis of variance (ANOVA) test method. According to the ANOVA test, the factors N (p¼ 0.01327) and T (p¼ 0.00025) are highly significant at a¼ 0.05 level [43].

The production of poly-3-hydroxybutyrate (PHB) by *Methylocystis hirsuta* from natural gas was first studied in two different media by Rahnama et al. [44]. After selecting the medium, the effects of methane-to-air ratio (20/80–50/50–80/20) and nitrogen content (0–50–100 mg/L) on PHB production were determined in a bubble column using a 3^2 full factorial design. Both of these factors were significant, and a maximum accumulation of 42.5% w/w of dry cell weight was achieved [44].

Nalakath et al. [45] investigated the effect of four factors on the bioconversion of carbon monoxide to ethanol and acetic acid by *Clostridium autoethanogenum* using a 2^4 full factorial design. The four studied factors were initial pH (4.75–5.75), initial total pressure (0.8–1.6 bar), cysteine-HCl and H_2O concentration (0.5–1.2 g/L), and yeast extract (YE) concentration (0.6–1.6 g/L). The maximum ethanol production was enhanced by up to 200% at conditions pH 4.75 and a YE concentration of 1.6 g/L. All the main effects and the interaction effects were found to be statistically significant (p50.05). Main effect plots indicated that increasing the initial pH and using higher YE concentrations negatively affected ethanol production, whereas increasing the initial pressure and the cysteine-HCl and H_2O concentration had a positive effect. Ethanol production of 0.065 g/L was achieved using the following values of the factors: pH (4.75), pressure (1.6 bar), cysteine-HCl.H2O (1.2 g/L), and YE concentration (1.6 g/L; [45]).

The optimal conditions for microwave-assisted enzymatic biodiesel synthesis were investigated using a 22 full factorial design. The critical factors affecting biodiesel production were the ethanol-to-beef tallow ratio ($X1$, 6–12) and the reaction temperature ($X2$, 40–50° C). The transesterification yield (Y) was selected as the response. The reactions were catalyzed by a lipase from *Burkholderia cepacia*, which was immobilized on silica, and were performed using 8–15 W of microwave radiation. High conversions were achieved using lower molar ratios of ethanol-to-beef tallow, and the effect of temperature was observed to be not significant based on statistical analysis. Under optimized conditions (a 1:6 molar ratio of beef tallow to ethanol at 5° C), the fatty acids in the original beef tallow were almost completely converted into ethyl esters in 8 h of reaction time, and productivity of 92 mg ethyl esters/g.h was achieved (a six-fold increase; [46]).

Virological testing of bottled water is another bioengineering application. A study was conducted to choose the best tool for detecting viruses in bottled water. Different approaches were examined; for example, the recovery of viral RNA was measured following the in-situ lysis of virus particles in the aqueous phase. A second method detected the healing of viral RNA following the lysis of virus particles. The third detection method generated the lowest genome recovery, regardless of water and virus type. Two ways were compared because they were considered viruses. Using a 3^4 4^1 2^1 full factorial design, viral RNA recovery was determined. The independent variables used were three types of water with three different mineral compositions; four viruses (poliovirus, hepatitis A virus, Norovirus, and MS2

phage); three incubation times (1, 10, and 20 days), and two methods (A and B). Every factor except incubation time was found to be significant. Method A provided the best results, suggesting that this method should be used to detect hepatitis A [47].

A study of the metabolic responses of a new neuronal human cell line, AGE1HN, to various substrate values, showed that reduced substrate and pyruvate load improves metabolic efficiency leading to improved growth and a1 antitrypsin (A1AT) production. A 3^2 full factorial design was used to analyze metabolism's adaptation to various pyruvate and glutamate concentrations. The concentrations of pyruvate tested were 2, 5, and 9 mm, and the strengths of glutamate tested were 5, 7.5, and 10 mM. The most important result was that higher pyruvate concentrations in the medium decreased cell proliferation and reduced the efficiency of substrate use. However, the highest viable cell density and A1AT concentration (167% of the batch) could be achieved without adding pyruvate [48].

Most biotechnology unit operations are complex with numerous process variables, feed material attributes, and raw material attributes that can have significant impact on the performance of the process. Design of experiments (DOE) based approach offers a solution to the bioprocessing unit operations. I-optimal and D-optimal designs to the commonly used central composite and Box–Behnken designs for bioprocess applications [49]. Chang and Zandstra [38] have developed and validated such a technology for ES cell differentiation analysis [38]. They used a quantitative screening platform based on automated fluorescence microscopy, which enumerated the ES cells that had entered endodermal differentiation through expression of two biomarkers (cytokeratin-8 and hepatocyte nuclear factor 3β). Using a two-level fractional factorial design model based on 32 triplicate experiments, they screened important medium components for the differentiation process to endodermal cells (glucose, insulin, basic fibroblast growth factor, retinoic acid, and epidermal growth factor) with the biomarkers and cell numbers as responses.

The model was further refined using a subsequent three-level factorial experiment for two of the factors. A statistical regression model was used to identify major and interactive effects on the endoderm formation. Similarly, Yu Ji [50] discussed on Model based process design for bioprocess optimization: case studies on precipitation with its applications in antibody purification. The objective of this study was to design and optimize a precipitation based mAb purification process [50].

The process was selected from two precipitation systems with ammonium sulfate and PEG 6000 as precipitant, respectively. Then it was further evaluated as an alternative mAb capturing step in the general purification platform. The conditions of initial DoE and validation DoE used in model based algorithm for mAb precipitation by ammonium sulfate have carried out with that of the conditions of initial DoE and validation DoE used in model based algorithm for mAb precipitation by PEG. Final optimization was performed of ammonium sulfate using d-optimal design by taking total of 21 points.

Similarly, it has carried out for PEG precipitation also and evaluated. The model based process design approach includes bioprocess modeling, model based experimental design, and high-throughput microwell experimentation. The bioprocess design is based on experimental data and a computational framework with

optimization algorithm. Innovative model based experimental design is a core part in this approach. The method also employs Random design and Simplex to identify extra experiments to increase the accuracy, and will iteratively improve the process design solutions [50].

10.9 DoE-Based Media Formulation Optimization Technique

Design of Experiment (DoE) approaches allow the interaction effects of components to be explored. The response surface methodology (RSM) developed by Box and Wilson has been used to formulate media for bacteria, fungi, and mammalian cells. Several researchers have used the Central Composite design (CCC or CCF) with the RSM optimization technique for media optimization. Media Blending is a common technique used to optimize many media ingredients simultaneously. In this approach, mixtures are ranked based on the highest-performing blends by mixing with already available media. When blending media, their composition should be known and significantly different so that the resulting combinations are unique. However, since all components are changed simultaneously, concluding causal relationships is impossible. Ultimately, media blending reduces the time and cost of optimization compared to the traditional and DoE approaches. For example, Roullier et al. used media blending to increase the process titer by 40% using a high-throughput methodology that took 6 weeks. Sixteen base CDMs containing the same 47 components were combined to create 376 unique media formulations. The collected data was used to build statistical models that could identify critical elements in the media. It was found that ferric ammonium citrate, pantothenic acid, valine, methionine, arginine, biotin, and serine were the most significant components of titer. Effective media formulation requires accurate analytical methods to measure metabolites over the culture duration. Generally, metabolic analyzers (using combinations of biosensors and ion-selective electrodes) weigh around ten primary metabolites and ions. To measure other components, several different techniques are applied, for example, capillary electrophoresis (CE), HPLC, liquid chromatography-coupled with mass spectrometry (LCMS), gas chromatography (GC), and gas chromatography-coupled with mass spectrometry (GCMS). Genomic and transcriptomic studies have also been performed to increase production by altering the temperature of the cultivation, supplying butyric acid, and inducing osmotic shock. Griffin et al. used different levels of gene expression to identify metabolism changes. Millipore Sigma has developed a targeted approach to find the media components to which cells are expected to respond to, eliminating the need for random testing. Specifically, a high-throughput microarray analysis was employed to screen the cytokine receptors responsible for growth in a particular cell line and process. Media components that could then activate those receptors could be tested and used for a more targeted media optimization study. They found four ligands that could enable the receptors in the media. This information could be used to design a much smaller group of experiments around those four components that could optimize cell growth [51].

10.10 Current Challenges and Future Prospective

Modeling has always been an attractive prospect due to its potential to replace laboratory experiments or significantly reduce the amount of experimentation required. Models simply calculate an output as a function of given inputs. If the model inputs are media components, then the formulation can be optimized to produce the desired output. While no formal model based approach has been developed (although it could be argued the DoE approach is a model based approach), it is worth briefly introducing the types of models that have been developed for bio-manufacturing processes and how they might be adapted to media formulation applications. Customizing cell culture media according to individual cell line needs for each method is of paramount importance. Different cell lines and even cell clones require different formulations to optimize performance, and the process type, batch, fed-batch, and continuous require vastly different media. The intensification and concentration of fed-batch feed media have led to problems with precipitation and feed media without significant performance drops needed for perfusion culture to be more cost-competitive. Emerging high-throughput methodologies will accelerate the development and validation of such models. This then leads to a scenario in which on-line, in-line, or at-line PAT could provide inputs to models and controllers on the floor that can optimize feed media on the manufacturing floor, removing the need for months of process development and creating an incredibly robust process from a regulatory standpoint.

10.11 Concluding Remarks

The experimental design is a form of process analysis in which specific factors are selected to obtain the desired responses of interest. It may also be used to determine the effects of various independent factors on a dependent factor. The bioengineering discipline includes many different areas of scientific interest, and each study area is affected and governed by many different factors. Briefly analyzing the essential factors and selecting an experimental design for optimization are handy tools for the design of any bioprocess under question. This chapter summarizes innovative design methods that can be used to investigate various factors relating to bioengineering processes or biotechnology. The basic concepts and applications of the design of experiments (DoE) in recombinant protein biotechnology are also briefly discussed.

References

1. Montgomery DC (2005) Design and analysis of experiments, 6th edn. Wiley, New York
2. Karen LJ (2012) Experimental design in biotechnology. Technometrics 33(1):106–108
3. Perry DH (1991) Book Review. Dry Technol 9(3):817
4. Mandenius CF, Brundin A (2008) Beioprocess optimization using design-of-experiments methodology. Biotechnol Prog 24(6):1191–1203. https://doi.org/10.1002/btpr.67

5. Sahu PK, Rao RN, Cecchi T, Swain S, Patro CS, Panda J (2018) An overview of experimental designs in HPLC method development and validation. J Pharm Biomed Anal 147(5):590–611
6. Yates F (1937) Design and analysis of factorial experiments. Technical communication no. 35. Imperial Bureau of Soil Science, London
7. Pilat P, Votruba J, Dobersky P, Prokop A (1976) Application of mathematical optimization methods in microbiology. Folia Microbiol 21:391–405
8. Singh B, Beg S (2013) Quality by design in product development life cycle. Chron Pharmabiz 22:72–79
9. Singh B, Kumar R, Ahuja N (2004) Optimizing drug delivery systems using systematic "design of experiments." part I: fundamental aspects. Crit Rev Ther Drug Carrier Syst 22(1):27–105
10. Singh B, Raza K, Beg S (2013) Developing "optimized" drug products employing "designed" experiments. Chem Ind Dig 12:1–7
11. Box GE, Wilson K (1951) On the experimental attainment of optimum conditions. J R Stat Soc Ser B (Methodological) 13(1):1–45
12. Bruns RE, Scarminio IS, Barros, N.B.d. (2006) Statistical design-chemometrics. Elsevier, Amsterdam
13. Ferreira SLC, Bruns RE, da Silva EGP, dos Santos WNL, Quintella CM, David JM, de Andrade JB, Breitkreitz MC, Jardim ICSF, Neto BB (2007a) Statistical designs and response surface techniques for the optimization of chromatographic systems. J Chromatogr A 1158(1):2–14
14. Hibbert DB (2012) Experimental design in chromatography: a tutorial review. J Chromatogr B 910:2–13
15. Tarley CRT, Silveira G, dos Santos WNL, Matos GD, da Silva EGP, Bezerra MA, Miro M, Ferreira SLC (2009) Chemometric tools in electroanalytical chemistry: methods for optimization based on factorial design and response surface methodology. Microchem J 92(1):58–67
16. Box GE, Behnken D (1960) Some new three level designs for the study of quantitative variables. Technometrics 2(4):455–475
17. Kuehl RO, Kuehl RO (2000) Design of Experiments: statistical principles of research design and analysis. Duxbury/Thomson Learning, Pacific Grove, CA
18. Otto M (1999) Chemometrics: statistics and computer application in analytical chemistry. Wiley-VCH, Weinheim, New York
19. Ferreira SLC, Bruns RE, Ferreira HS, Matos GD, David JM, Brandao GC, da Silva EGP, Portugal LA, dos Reis PS, Souza AS, dos Santos WNL (2007b) Box–Behnken design: an alternative for the optimization of analytical methods. Analytica Chim Acta 597(2):179–186
20. Hanrahan GZJ, Gibani S, Patil DG (2005) Chemometrics: experimental design. In: Worsfold PJTA, Poole CF (eds) Encyclopedia of analytical science. Elsevier, Oxford, pp 8–13
21. Haaland PD (1989) Experimental design in biotechnology. Marcel Dekker, New York
22. Kwang-Min L (2006) David Gilmore. Statistical experimental design for bioprocess modeling and optimization analysis: repeated-measures method for dynamic biotechnology process. Appl Biochem Biotechnol 135(2):101–116
23. Vijesh K, Akriti B, Anurag SR (2008) Design of experiments applications in bioprocessing: concepts and approach. Biotechnol Prog 24(6):1191–1203
24. Kwang-Min L, David FG (2006) Statistical experimental design for bioprocess modeling and optimization analysis: repeated-measures method for dynamic biotechnology process. Appl Biochem Biotechnol 135(2):101–116
25. Myers RH, Montgomery DC (2002) Response surface methodology: process and product optimization using designed experiments, 2nd edn. Wiley, New York
26. Wu CFJ, Hamada M (2000) Experiments: planning, analysis, and parameter design optimization. Wiley, New York
27. Wang YH, Yang B, Ren J, Dong ML, Liang D, Xu AL (2004) Optimization of medium composition for the production of clavulanic acid by Streptomyces clavuligerus. Process Biochem 40:1161–1166

28. Adinarayana K, Ellaiah P, Srinivasulu B, Bhavani DR, Adinarayana G (2003) Response surface methodological approach to optimize the nutritional parameters for neomycin production by Streptomyces marinensis under solid-state fermentation. Process Biochem 38:1565–1572
29. Cladera-Olivera F, Caron GR, Brandelli A (2004) Bacteriocin production by Bacillus licheniformis strain P40 in cheese whey using response surface methodology. Biochem Eng J 21:53–58
30. Gupte MD, Kulkarni PR (2002) A study of antifungal antibiotic production by Streptomyces chattanoogensis MTCC 3423 using full factorial design. Lett Appl Microbiol 35:22–26
31. Penna TCV, Moraes DA (2002) Optimization of nisin production by Lactococcus lactis. Appl Biochem Biotechnol 98:775–789
32. Torres BJ, Arroyo M, Torres-Guzman R, De La MI, Acebal C, Castillon MP (2005) Optimization of culture medium and conditions for penicillin acylase production by Streptomyces lavendulae ATCC 13664. Appl Biochem Biotechnol 126:119–131
33. Zhuang YP, Chen B, Chu J, Zhang S (2006) Medium optimization for meilingmycin production by Streptomyces nanchangensis using response surface methodology. Process Biochem 41:405–409
34. Zitzmann J, Sprick G, Weidner T, Schreiber C, Czermak P (2017) Process optimization for recombinant protein expression in insect cells. In: Gowder SJT (ed) New insights into cell culture technology. InTech Open, Rijeka, pp 43–98
35. Yildirim S, Thompson MG, Jacobs AC, Zurawski DV, Kirkup BC (2016) Evaluation of parameters for high efficiency transformation of Acinetobacter baumannii. Sci Rep 6:22110
36. Ikonomou L, Schneider Y-J, Agathos SN (2003) Insect cell culture for industrial production of recombinant proteins. Appl Microbiol Biotechnol 62:1–20
37. Ramirez J, Gutierrez H, Gschaedler A (2001) Optimization of astaxanthin production by Phaffia rhodozyma through factorial design and response surface methodology. J Biotechnol 88:259–268
38. Chang KH, Zandstra PW (2004) Quantitative screening of embryonic stem cell differentiation: endoderm formation as a model. Biotechnol Bioeng 88:287–298
39. Govind S, Ketaki K, Saritha S, Gupta MK, Yadav KS (2020) Quality by design (QbD) approach in processing polymeric nanoparticles loading anticancer drugs by high pressure homogenizer. Heliyon 6(4):e03846
40. Chien NB, Tran BT, Thao D, Thi NN (2018) D-optimal optimization and data-analysis comparison between a DoE software and artificial neural networks of a chitosan coating process onto PLGA nanoparticles for lung and cervical Cancer treatment. J Pharm Innov 14(4):1–15
41. Kozlowski S, Swann P (2009) Considerations for biotechnology product quality by design. In: Rathore AS, Mhatre R (eds) Quality by design for biopharmaceuticals. Wiley, New York, pp 9–30
42. Christos P (2019) Design of experiments as a tool for optimization in recombinant protein biotechnology: from constructs to crystals. Mol Biotechnol 61:873–891. https://doi.org/10.1007/s12033-019-00218-x
43. Mendes LF, Vale LA, Martins AP et al (2012) Influence of temperature, light and nutrients on the growth rates of the macroalga Gracilaria domingensis in synthetic seawater using experimental design. J Appl Phycol 24:1419–1426
44. Rahnama F, Vasheghani-Farahani E, Yazdian F, Shojaosadati SA (2012) PHB production by Methylocystis hirsuta from natural gas in a bubble column and a vertical loop bioreactor. Biochem Eng J 65:51–56
45. Nalakath AH, Veiga MC, Kennes C (2012) Biological conversion of carbon monoxide to ethanol: effect of pH, gas pressure, reducing agent and yeast extract. Bioresour Technol 114:518–522
46. Da Rós PC, de Castro HF, Carvalho AK et al (2012) Microwave assisted enzymatic synthesis of beef tallow biodiesel. J Ind Microbiol Biotechnol 39:529–536

47. Huguet L, Carteret C, Gantzer C (2012) A comparison of different concentration methods for the detection of viruses present in bottled waters and those adsorbed to water bottle surfaces. J Virol Methods 181:18–24
48. Niklas J, Priesnitz C, Rose T et al (2012) Primary metabolism in the new human cell line AGE1. HN at various substrate levels: increased metabolic efficiency and a1-antitrypsin production at reduced pyruvate load. Appl Microbiol Biotechnol 93:1637–1650
49. Kumar V, Bhalla A, Rathore AS (2014) Design of experiments applications in bioprocessing: concepts and approach. Biotechnol Prog 30(1):86–99
50. Ji Y (2012) Model based process design for bioprocess optimisation: case studies on precipitation with its applications in antibody purification. PhD thesis submitted to The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering University College, London
51. Shaun CG, Hemalatha B, Houlong L, Seongkyo Y (2018) Media formulation optimization: current and future opportunities. Curr Opin Chem Eng 22:42–47