Copyright © 2006 by Humana Press Inc.
All rights of any nature whatsoever reserved.
0273-2289/1559-0291 (Online)/06/133/0113/$30.00

Modeling and Optimization of Biopolymer (Polyhydroxyalkanoates) Production From Ice Cream Residue by Novel Statistical Experimental Design

KWANG-MIN LEE*,[1] AND DAVID F. GILMORE[2]

[1]Celltrion, Inc., 1001-5, Dongchun-dong, Yeonsu-gu Inchen, 406-130, South Korea, E-mail: Klee0922@yahoo.co.kr; and [2]Environmental Sciences Program, Arkansas State University, Jonesboro, AR 72467;

Received July 31, 2005; Revised October 12, 2005; Accepted October 30, 2005

Abstract

Polyhydroxyalkanoates (PHAs) are thermoplastic polyesters synthesized by *Ralstonia eutropha* and other bacteria as a form of intracellular carbon and energy storage and are accumulated as lipid inclusions in the cytoplasm of these bacteria. The modeling and optimization of PHA production by fermentation from industrial waste (ice cream residue) was studied by employing statistical experimental design methods. A series of iterative experimental designs was used to find optimal factor conditions (medium components and fermentation process time) in the order of fractional factorial design, path of steepest ascent, and full factorial augmented with axial design (rotational central composite design). An optimal range characterized by lipid (15 mg/mL) and % lipid (88%) values was found and further investigated to verify the optimal conditions for PHA production from ice cream (56.68 mL of ice cream or 56.68% ice cream in water [v/v], 5.03 mL of buffer, 1 mL of mineral salts solution, 100 µL of trace element solution, 100 mL of seed culture, and 213.76 h of fermentation time).

Index Entries: Polyhydroxyalkanoate; optimization; design of experiment; modeling; repeated measures.

*Author to whom all correspondence and reprint requests should be addressed.

Introduction

Bioplastics

Biodegradable polymers are reliable options for solid waste management. Whenever plastic is discarded, it remains at the disposal site and its property of resisting biologic or chemical breakdown results in its accumulation in the environment *(1)*. Increasing environmental pollution and potential exhaustion of nonrenewable fossil resources have spurred investigations to find biodegradable and biosynthetic materials *(2)*.

Polyhydroxyalkanoates (PHAs) are naturally occurring polyesters of various hydroxyalkanoates that are synthesized by a broad range of microorganisms as intracellular inclusions of carbon and energy reserve compounds under unbalanced growth conditions. Poly-3-hydroxybutyrate (PHB) is the most well-known member of the family of PHAs. PHB is similar to polypropylene in its physical properties, but it is completely biodegradable by bacteria and fungi into water and carbon dioxide under aerobic conditions, and water, carbon dioxide, and methane are the final end products when it is degraded under anaerobic conditions *(3–7)*.

In the United States, Monsanto (St. Louis, MO) produced PHB and PHB-co-PHV on a scale of metric tons, and these PHAs are marketed under the trade name BIOPOL at a cost of US$16/kg and are used in the manufacture of cosmetic bottles, packing materials, and biochemical devices *(8)*. In May 2001, Metabolix (Cambridge, MA) took over the BIOPOL from Monsanto and started to commercialize biopolyesters. However, the price of biopolyesters is not competitive with that of petrochemical-based polypropylene, which is less than US$1/kg *(1)*. Thus, one of the major problems preventing the commercial application of PHAs is their high price. Much effort has been devoted to reduce the price of PHAs by developing better bacterial strains (recombinant bacteria such as *Escherichia coli*); more efficient fermentation; more economical recovery processes; and, most important, cheaper carbon sources *(9)*.

The carbon source contributes most significantly to the overall cost in PHA production. A number of carbon sources, including carbohydrates, oils, alcohols, acids, and hydrocarbons, can be used by various bacteria. Recently, owing to their low price and potential availability, crude carbon substrates (food wastes or byproducts), such as cane, cheese whey, alpectin, plant oils, tallow, cellulose, and beet molasses, have attracted much attention *(8–19)*.

Statistical Approach to Solve a Bioprocess Problem

In most bioprocesses such as fermentation and other cell culture methods (mammalian or insect), there are no true theoretical or mathematical models that can describe the whole process with 100% certainty. Because of this limitation arising from the incredible complexity of cellular metabolism, efficient empirical approaches to explain these processes are neces-

sary to solve research problems. These statistical or empirical methods must provide lots of data to enable a researcher to reach meaningful conclusions. However, any problem-solving approach is limited by time, money, and resources for research. Because there are limited opportunities to generate and collect data, it is critical that the data be rich in information. A statistically designed experiment is one solution for obtaining the information-rich data from the process being studied, given these limitations *(20,21)*.

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 so that the data collected can be effectively analyzed to derive valid and objective conclusions. The combination of settings for the process variables (also called factors; predictors; regressors; or independent, explanatory variables) is called a run or treatment. A measure of the treatment is referred to as a response (dependent or performance variable), and each output of the response variable is called an observation. Multifactors can have multiresponses *(21,22)*.

DOE begins with identifying the experimental objectives and choosing the process variables (factors) for the study. An experimental design is the laying out of a well-prepared and detailed plan prior to conducting the actual experiment. Well-selected experimental designs economically maximize the amount of information. DOE is widely used in scientific research as well as in industrial settings to draw statistical significance of an effect out of particular factors or combinations of the treatments and to achieve the maximum amount of information about factors affecting a production process with the least amount of time and cost possible.

The statistical theory behind DOE comes from the concept of process models. A process model consists of several discrete and continuous input factors (process variables) capable of being controlled, i.e., varied by the experimenter, and a single or multimeasurement of output (responses). Experimental data collected from each run are employed to derive an empirical model linking process variables (inputs) and responses (outputs). These empirical models are usually simple polynomials *(21–26)*.

DOE is most often applied in planning experiments for variable screening model building, and optimization. If there are many factors whose importance cannot be ruled out at the beginning of a study, a screening experiment should be conducted to eliminate the unimportant ones. Typical screening experiments are Plackett-Burman and fractional factorial designs. Once a small number of important factors are identified, the investigation can continue with subsequent experiments to explore their effects on the response. The relationship between the response and these variables is sometimes called response surface *(21,22)*.

When the purpose of an investigation is to maximize or minimize the response, first-order designs are used to move the experimental region closer to the optimum conditions of the input factors. Once a first-order model is fitted, a search for higher values of the response can be conducted

in the steepest ascent direction. When the experimental region is near or within the optimum region of response surface, a second-order model (one that allows curvature) is used to approximate the relationship between the response and input factors. A second-order design such as central composite designs (CCDs) and Box-Behnken designs allows the model parameters to be estimated. Canonical analysis, contour plots, or desirability functions can be used to identify optimal factor settings *(21,22,27–31)*.

In the present research, the factors affecting PHA production were identified by employing fractional factorial design and path of steepest ascent (PSA), and the PHA production process was optimized using a CCD. The objectives of this investigation were to search for the optimal fermentation culture medium and process factors and to evaluate the effects of the significant factors on PHA production.

Materials and Methods

Bacterial Strain and Stock Culture

The aerobic Gram-negative bacterium *Ralstonia eutropha* H16, obtained from the University of Massachusetts, was used. This culture was maintained on Trypticase Soy Agar (Difco, Detroit, MI) slants at 30°C and subcultured every 2 wk.

Preparation of Seed Culture

The seed culture was prepared by inoculating a loopful of stock culture into a 250-mL shake flask containing 80 mL of Luria-Bertani (LB) medium (10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl/L ; pH 6.6–7.2) and incubating the flask in a rotary shaker for 24 h at 30°C and 175 rpm. Then, the culture was transferred into a 2.5-L shake flask containing 1 L of LB medium. This culture was incubated for another 24 h under the same conditions.

Carbon Substrate for PHA Production

Ice cream left over from flavor switching was obtained from the Yarnell Ice Cream Company, Searcy, AR. Melted ice cream was used as the sole carbon source for the assay medium following pretreatment by centrifugation to remove particulates. Table 1 gives its dietary composition.

Preparation of Assay Media

Assay media for the production of PHA were prepared by adding pretreated ice cream to basal medium containing mineral salts, trace elements, and buffer, depending on the experimental design of a given experiment. Table 2 provides compositions of the concentrated stock solutions used in the basal medium. All media were adjusted to pH 6.5, sterilized at 121°C for 15 min, and cooled to room temperature prior to use.

Table 1
Dietary Composition of Vanilla Ice Cream Manufactured by Yarnell[a]

Total fat: 7 g; saturated fat: 4.5 g
Total carbohydrate: 18 g; sugar: 13 g
Protein: 2 g
Ingredients: milk, cream, sucrose, skim milk, fructose, egg yolks, buttermilk, cellulose gel, cellulose gum, mono- and diglycerides, carrageenan, vanilla, polysorbate 80

[a]Nutrition facts are for a single serving size (71 g) (data from manufacturer's label).

Table 2
Basal Medium Components *(32)*

Mineral (100X)[a]		Trace element (1000X)[b]			KPO_4 buffer (10X)[c]	
$NaSO_4$	0.74 g	Microelement A:	$CuSO_4 \cdot 5H_2O$	0.4 g	KH_2PO_4	2.8 g
$MgCl_2$	0.12 g		$ZnSO_4 \cdot 7H_2O$	0.4 g	$K_2HPO_4A3H_2O$	2.16 g
KH_2PO_4	0.09 g		$Na_2MoO_4 \cdot 2H_2O$	0.4 g		
$CaCl_2$	0.64 g		$MnSO_4 \cdot H_2O$	0.2 g		
$NaHCO_3$	2.75 g		$CoCl_2 \cdot H_2O$	0.2 g		
KCl	0.91 g		H_3BO_3	0.2 g		
NaCl	4.22 g		KI	0.1 g		
			$Na_2WO_4 \cdot 2H_2O$	0.1 g		
			$AlK(SO_4)_2 \cdot 12H_2O$	0.1 g		
		Microelement B:	$FeSO_4 \cdot 7H_2O$	0.5 g		

[a]Per liter of H_2O.
[b]Per 100 mL of H_2O.
[c]Per 100 mL of H_2O.

Shake-Flask Fermentation and Culture Conditions

Batch fermentation for PHA production was carried out in 250-mL shake flasks containing 100 mL of assay medium. For each experiment, portions of a 1-L seed culture (volume depending on each experimental design) were centrifuged (10,000*g* for 5 min) to provide cell pellets for inoculation. Fermentation was initiated by resuspending each cell pellet with about 50 mL of the assay medium, then combining the suspension with the remaining culture medium in the fermentation flask (assay flask). The assay flasks were incubated at 30°C in a rotary shaker for 168–240 h, depending on the experimental design. Samples were collected at regular 24-h intervals h and analyzed immediately after being drawn from the cultures.

Sampling and PHA Recovery Analysis

In every collection, 5 mL of assay culture was withdrawn from each flask and transferred into a 20-mL vial. The pH was measured and 1 mL of

culture was aliquoted into a microtube. The cells were harvested by centrifuging for 2 min at 12,000*g* in a microcentrifuge and washed with distilled water twice to remove residual culture medium. The initial supernatant from the culture samples was frozen for further analysis. The biomass of each cell pellet was determined by gravimetry after drying to constant weight in a vacuum oven at 65°C overnight.

The same gravimetric method was employed to measure PHA after a recovery treatment as follows. One milliliter of commercial bleach (Clorox, 6% sodium hypochlorite) was added to the wet biomass pellet, and this was vortexed thoroughly and incubated in a glass bead bath at 100°C until non-PHA biomass dissolved and white flocculant lipid material precipitated (about 20–30 min). The lipid material was recovered by centrifugation and washed with distilled water twice to remove the bleach solution, which contained dissolved cell debris. To remove contaminating lipid cell debris, the wet lipid pellets were resuspended with 1 mL of acetone, centrifuged, and the supernatant was discarded. The pellet was resuspended with 1 mL of ethanol and centrifuged. The resulting white pellet was dried in the heating bath and weighed. All microtubes were preweighed and all measurements were duplicated.

The data presented are defined as follows: dry cell weight (DCW) is the total biomass dried to a constant weight. Lipid refers to extracted and dried PHA. Percent lipid (% lipid) is the percentage of DCW that is PHA (% yield). Net lipid (d lipid) is the amount of PHA accumulated in a 24-h sampling interval. Specific lipid rate (q lipid) is the specific PHA production rate, which is net lipid divided by DCW. This is based on the fact that the rate of PHA synthesis is proportional to biomass present. Lipid productivity is the volumetric PHA productivity, which is defined as the final PHA concentration divided by the fermentation time from inoculation. From nuclear magnetic resonance (NMR) analysis, it was found that bacterial lipid obtained by the lipid extraction protocol was almost pure PHB (>99%) (C. Scholz, personal communication, Sept. 2002). Thus, gravimetry was a reliable way to quantify PHA.

Experimental Designs for Screening and Optimization

The factor treatments and experimental settings were performed by experimental design. Design Expert (version 6.06, Stat-Easy, Minneapolis, MN) was employed for building and analyzing experimental designs. To optimize the composition of the assay medium and the fermentation process for PHA production, a series of statistical designs was created prior to performing experiments. On the basis of the results of pilot experiments in which glucose was used for the major carbon source instead of ice cream (data not shown), several potential significant factors were selected and varied according to the design layout obtained from Design Expert.

Table 3
Actual Values and Coded Levels of Factors for 2^{4-1} Fractional Factorial Design

Factor	Coded symbol	Unit	Coded level/actual values[a]		
			–1	0	+1
Ice cream	x_1	mL	10	30	50
Seed culture	x_2	mL	100	150	200
Trace element[b]	x_3	:L	50	100	150
Buffer[b]	x_4	mL	1	3	5

[a]Code for high = (high – average)/step size.
Code for low = (low – average)/step size, in which step size = factor range/2.
[b]Volume of stock solutions (*see* Table 2).

Table 4
PSA Experimental Design and Treatments

Treatment (run)	Ice cream (mL)		Seed culture (mL)	Trace element (μL)	Buffer (mL)
	Coded	Actual			
1	–1	30	100	100	5
2	–0.5	35	100	100	5
3	0	40	100	100	5
4	0.5	45	100	100	5
5	1	50	100	100	5

2^{4-1} Fractional Factorial Design for Factor Screening

Factors initially considered significant were the amount of ice cream, trace elements, buffer, and volume of seed culture as media formulation factors, and fermentation time as a process factor. Table 3 gives the experimental factor range of actual values and coded levels of variables (factors) used for initial screening in the 2^{4-1} fractional factorial design.

PSA From Screening Design

On the basis of the values of the regression coefficients from the first-order model equation calculated by the 2^{4-1} fractional factorial design, a series of new treatments was conducted in the direction of the steepest ascent (Table 4). The PSA was to increase the amount of ice cream because only the ice cream factor showed statistical significance (main effect) in the previous 2^{4-1} fractional factorial design model.

2^2 Full Factorial Design From PSA

A new 2^2 full factorial design was created based on the result of PSA. A smaller scaling factor range than the one used for screening fractional factorial design was used to narrow down the potential optimal region. Table 5 gives the new experimental range of actual values and coded level of factors.

Table 5
Actual Values and Coded Level of Factors for 2^2 Full Factorial Design

			Coded level/actual values		
Factor	Coded symbol	Unit	–1	0	+1
Ice cream	X_1	mL	40	50	60
Buffer	X_2	mL	5	7.5	10

Rotatable CCD Augmented With Axial Design

For the optimization step, the linear model should be augmented with axial design points to build a CCD when significant curvature or lack of fit is detected. The selection of the proper value of α is required to create a rotatable CCD (RCCD). Since significant lack of fit and curvature was detected in the 2^2 full factorial design, CCD is the next step to approach and identify the optimal point. With α = 1.414 according to the rule of $\alpha = (k)^{1/2}$ in which k = factor number, the 2^2 full factorial design augmented with four axial points (2FRCCD, which means 2 factor RCCD) in Table 6 was employed to investigate the final optimal conditions for PHA production under the factors selected.

Verification by Mixture Design

For verification of the optimal conditions, a new experimental design (mixture design combined with process factor, also called crossed design) was used for medium formulation, and robust design. New design points were created based on the optimal points of the RCCD. The purpose of this design is to verify the model and the optimization from RCCD. Medium formulation and robust design will be described in another article *(15)*.

Results and Discussion

2^{4-1} Fractional Factorial Design for Factor Screening

A 2^{4-1} fractional factorial design was employed for the experimental plan for screening the most significant factors from the potential factors. One can judge the significance of the effects of factors on the response by making normal plots of the effects, as shown in Fig. 1. The squares plotted on the graphs correspond to estimates of the model effects, including all possible interactions between factors. Those effects that can be ignored are required to have estimates that follow normally distributed noise with a mean of zero and a constant variance along with pure errors and are represented by the triangles. They are on the straight line. However, significant effects can be picked out as the isolated squares that do not line up along the straight line, with their distance from the line depending on the magnitude of their effects.

Using the data from the this experiment, it was discovered that the concentration of ice cream is a significant factor for lipid production, and that there is an interaction between ice cream and buffer in % lipid. Fermentation time was used as a blocking factor to control the extraneous source of variability in the response. However, it was found that there is a large value for the sum of squares for fermentation time, which means that fermentation time is another significant factor that influences responses. Thus, it is advised not to use fermentation time as a blocking factor, owing to its significant effects on response variables. Based on these discoveries, the steepest ascent method was used to move quickly to the potential vicinity of the optimum point.

Path of Steepest Ascent

The PSA was taken as a sequential movement along the direction of maximum increase in the response of interest, starting from the central design point of the previous experiment. The step sizes along the path are proportional to the regression coefficients of significant factors from the previous 2^{4-1} fractional factorial design. The actual (uncoded) step size was determined based on experience and knowledge about the process, considering the practical aspect of the treatment and outcome (Table 4).

According to part a of Table 7, lipid shows an R^2 value of 0.98, meaning that 95% of the variability in the response can be explained by the model, leaving only 5% of variability owing to other factors. There are time and ice cream interactions in lipid, which means that the effect of time on the response depends on the level of ice cream. More important, these interactions indicate that when the amount of ice cream in the medium is large, an increase in time will cause a greater increase in lipid than when the amount of ice cream in the medium is small. Part b of Table 7 also shows that there is a significant time effect on % lipid. This is a critical discovery that must be considered when optimizing any bioprocess economically and practically.

RCCD for Process Optimization

New 2^2 Full Factorial Design for Phase I Optimization

Based on the results of the former PSA experiment, a new 2^2 full factorial experiment was conducted with two center points, as shown in Table 8, block 1, which are potential optimal points in terms of ice cream and buffer. Considering the interaction between ice cream and buffer on the response of % lipid in the 2^{4-1} fractional factorial design and the overall increased lipid production in PSA with 5 mL of buffer for the whole process, an increased range of buffer was used in the new factorial combination of ice cream, as shown in Table 5. In addition, as a result of the discovery of interactions between time and ice cream concentration, a longer fermentation time (240 h) was employed to display more fermentation time courses, as shown in Figs. 2 and 3.

Table 6
Design Matrix for Rotatable Central Composite Design (RCCD)

Experiment	Treatment	Assay order	Time (h) actual	Ice cream (mL) actual	Buffer (mL) actual	Time coded	Ice cream coded	Buffer coded	Point type
1	1	1	24	40.00	5.00	–1.00	–1.00	–1.00	Fact
2	2	5	24	60.00	5.00	–1.00	1.00	–1.00	Fact
3	3	3	24	40.00	10.00	–1.00	–1.00	1.00	Fact
4	4	6	24	60.00	10.00	–1.00	1.00	1.00	Fact
5	0	4	24	50.00	7.50	–1.00	0.00	0.00	Center
6	0	2	24	50.00	7.50	–1.00	0.00	0.00	Center
7	5	7	24	35.86	7.50	–1.00	–1.41	0.00	Axial
8	6	11	24	64.14	7.50	–1.00	1.41	0.00	Axial
9	7	10	24	50.00	3.96	–1.00	0.00	–1.41	Axial
10	8	8	24	50.00	11.04	–1.00	0.00	1.41	Axial
11	0	9	24	50.00	7.50	–1.00	0.00	0.00	Center
12	0	12	24	50.00	7.50	–1.00	0.00	0.00	Center
13	1	17	48	40.00	5.00	–0.78	–1.00	–1.00	Fact
14	2	16	48	60.00	5.00	–0.78	1.00	–1.00	Fact
15	3	14	48	40.00	10.00	–0.78	–1.00	1.00	Fact
16	4	15	48	60.00	10.00	–0.78	1.00	1.00	Fact
17	0	18	48	50.00	7.50	–0.78	0.00	0.00	Center
18	0	13	48	50.00	7.50	–0.78	0.00	0.00	Center
19	5	19	48	35.86	7.50	–0.78	–1.41	0.00	Axial
20	6	23	48	64.14	7.50	–0.78	1.41	0.00	Axial
21	7	21	48	50.00	3.96	–0.78	0.00	–1.41	Axial
22	8	22	48	50.00	11.04	–0.78	0.00	1.41	Axial
23	0	20	48	50.00	7.50	–0.78	0.00	0.00	Center
24	0	24	48	50.00	7.50	–0.78	0.00	0.00	Center
25	1	30	72	40.00	5.00	–0.56	–1.00	–1.00	Fact

26	2	28	72	60.00	5.00	–0.56	1.00	–1.00	Fact
27	3	29	72	40.00	10.00	–0.56	–1.00	1.00	Fact
28	4	26	72	60.00	10.00	–0.56	1.00	1.00	Fact
29	0	27	72	50.00	7.50	–0.56	0.00	0.00	Center
30	0	25	72	50.00	7.50	–0.56	0.00	0.00	Center
31	5	33	72	35.86	7.50	–0.56	–1.41	0.00	Axial
32	6	32	72	64.14	7.50	–0.56	1.41	0.00	Axial
33	7	36	72	50.00	3.96	–0.56	0.00	–1.41	Axial
34	8	31	72	50.00	11.04	–0.56	0.00	1.41	Axial
35	0	34	72	50.00	7.50	–0.56	0.00	0.00	Center
36	0	35	72	50.00	7.50	–0.56	0.00	0.00	Center
37	1	38	96	40.00	5.00	–0.33	–1.00	–1.00	Fact
38	2	37	96	60.00	5.00	–0.33	1.00	–1.00	Fact
39	3	42	96	40.00	10.00	–0.33	–1.00	1.00	Fact
40	4	41	96	60.00	10.00	–0.33	1.00	1.00	Fact
41	0	39	96	50.00	7.50	–0.33	0.00	0.00	Center
42	0	40	96	50.00	7.50	–0.33	0.00	0.00	Center
43	5	46	96	35.86	7.50	–0.33	–1.41	0.00	Axial
44	6	47	96	64.14	7.50	–0.33	1.41	0.00	Axial
45	7	45	96	50.00	3.96	–0.33	0.00	–1.41	Axial
46	8	43	96	50.00	11.04	–0.33	0.00	1.41	Axial
47	0	48	96	50.00	7.50	–0.33	0.00	0.00	Center
48	0	44	96	50.00	7.50	–0.33	0.00	0.00	Center
49	1	50	120	40.00	5.00	–0.11	–1.00	–1.00	Fact
50	2	49	120	60.00	5.00	–0.11	1.00	–1.00	Fact
51	3	53	120	40.00	10.00	–0.11	–1.00	1.00	Fact
52	4	54	120	60.00	10.00	–0.11	1.00	1.00	Fact
53	0	51	120	50.00	7.50	–0.11	0.00	0.00	Center
54	0	52	120	50.00	7.50	–0.11	0.00	0.00	Center
55	5	57	120	35.86	7.50	–0.11	–1.41	0.00	Axial

(continued)

Table 6 *(continued)*

Experiment	Treatment	Assay order	Time (h) actual	Ice cream (mL) actual	Buffer (mL) actual	Time coded	Ice cream coded	Buffer coded	Point type
56	6	60	120	64.14	7.50	–0.11	1.41	0.00	Axial
57	7	58	120	50.00	3.96	–0.11	0.00	–1.41	Axial
58	8	55	120	50.00	11.04	–0.11	0.00	1.41	Axial
59	0	56	120	50.00	7.50	–0.11	0.00	0.00	Center
60	0	59	120	50.00	7.50	–0.11	0.00	0.00	Center
61	1	62	144	40.00	5.00	0.11	–1.00	–1.00	Fact
62	2	64	144	60.00	5.00	0.11	1.00	–1.00	Fact
63	3	65	144	40.00	10.00	0.11	–1.00	1.00	Fact
64	4	63	144	60.00	10.00	0.11	1.00	1.00	Fact
65	0	66	144	50.00	7.50	0.11	0.00	0.00	Center
66	0	61	144	50.00	7.50	0.11	0.00	0.00	Center
67	5	67	144	35.86	7.50	0.11	–1.41	0.00	Axial
68	6	69	144	64.14	7.50	0.11	1.41	0.00	Axial
69	7	71	144	50.00	3.96	0.11	0.00	–1.41	Axial
70	8	70	144	50.00	11.04	0.11	0.00	1.41	Axial
71	0	68	144	50.00	7.50	0.11	0.00	0.00	Center
72	0	72	144	50.00	7.50	0.11	0.00	0.00	Center
73	1	75	168	40.00	5.00	0.33	–1.00	–1.00	Fact
74	2	76	168	60.00	5.00	0.33	1.00	–1.00	Fact
75	3	73	168	40.00	10.00	0.33	–1.00	1.00	Fact
76	4	77	168	60.00	10.00	0.33	1.00	1.00	Fact
77	0	78	168	50.00	7.50	0.33	0.00	0.00	Center
78	0	74	168	50.00	7.50	0.33	0.00	0.00	Center
79	5	81	168	35.86	7.50	0.33	–1.41	0.00	Axial
80	6	83	168	64.14	7.50	0.33	1.41	0.00	Axial

81	7	80	168	50.00	3.96	0.33	0.00	–1.41	Axial
82	8	82	168	50.00	11.04	0.33	0.00	1.41	Axial
83	0	79	168	50.00	7.50	0.33	0.00	0.00	Center
84	0	84	168	50.00	7.50	0.33	0.00	0.00	Center
85	1	86	192	40.00	5.00	0.56	–1.00	–1.00	Fact
86	2	87	192	60.00	5.00	0.56	1.00	–1.00	Fact
87	3	89	192	40.00	10.00	0.56	–1.00	1.00	Fact
88	4	90	192	60.00	10.00	0.56	1.00	1.00	Fact
89	0	88	192	50.00	7.50	0.56	0.00	0.00	Center
90	0	85	192	50.00	7.50	0.56	0.00	0.00	Center
91	5	96	192	35.86	7.50	0.56	–1.41	0.00	Axial
92	6	93	192	64.14	7.50	0.56	1.41	0.00	Axial
93	7	94	192	50.00	3.96	0.56	0.00	–1.41	Axial
94	8	92	192	50.00	11.04	0.56	0.00	1.41	Axial
95	0	95	192	50.00	7.50	0.56	0.00	0.00	Center
96	0	91	192	50.00	7.50	0.56	0.00	0.00	Center
97	1	97	216	40.00	5.00	0.78	–1.00	–1.00	Fact
98	2	100	216	60.00	5.00	0.78	1.00	–1.00	Fact
99	3	99	216	40.00	10.00	0.78	–1.00	1.00	Fact
100	4	98	216	60.00	10.00	0.78	1.00	1.00	Fact
101	0	101	216	50.00	7.50	0.78	0.00	0.00	Center
102	0	102	216	50.00	7.50	0.78	0.00	0.00	Center
103	5	103	216	35.86	7.50	0.78	–1.41	0.00	Axial
104	6	104	216	64.14	7.50	0.78	1.41	0.00	Axial
105	7	107	216	50.00	3.96	0.78	0.00	–1.41	Axial
106	8	108	216	50.00	11.04	0.78	0.00	1.41	Axial
107	0	106	216	50.00	7.50	0.78	0.00	0.00	Center
108	0	105	216	50.00	7.50	0.78	0.00	0.00	Center
109	1	112	240	40.00	5.00	1.00	–1.00	–1.00	Fact
110	2	110	240	60.00	5.00	1.00	1.00	–1.00	Fact

(continued)

Table 6 *(continued)*

Experiment	Treatment	Assay order	Time (h) actual	Ice cream (mL) actual	Buffer (mL) actual	Time coded	Ice cream coded	Buffer coded	Point type
111	3	109	240	40.00	10.00	1.00	–1.00	1.00	Fact
112	4	114	240	60.00	10.00	1.00	1.00	1.00	Fact
113	0	113	240	50.00	7.50	1.00	0.00	0.00	Center
114	0	111	240	50.00	7.50	1.00	0.00	0.00	Center
115	5	116	240	35.86	7.50	1.00	–1.41	0.00	Axial
116	6	115	240	64.14	7.50	1.00	1.41	0.00	Axial
117	7	117	240	50.00	3.96	1.00	0.00	–1.41	Axial
118	8	120	240	50.00	11.04	1.00	0.00	1.41	Axial
119	0	118	240	50.00	7.50	1.00	0.00	0.00	Center
120	0	119	240	50.00	7.50	1.00	0.00	0.00	Center

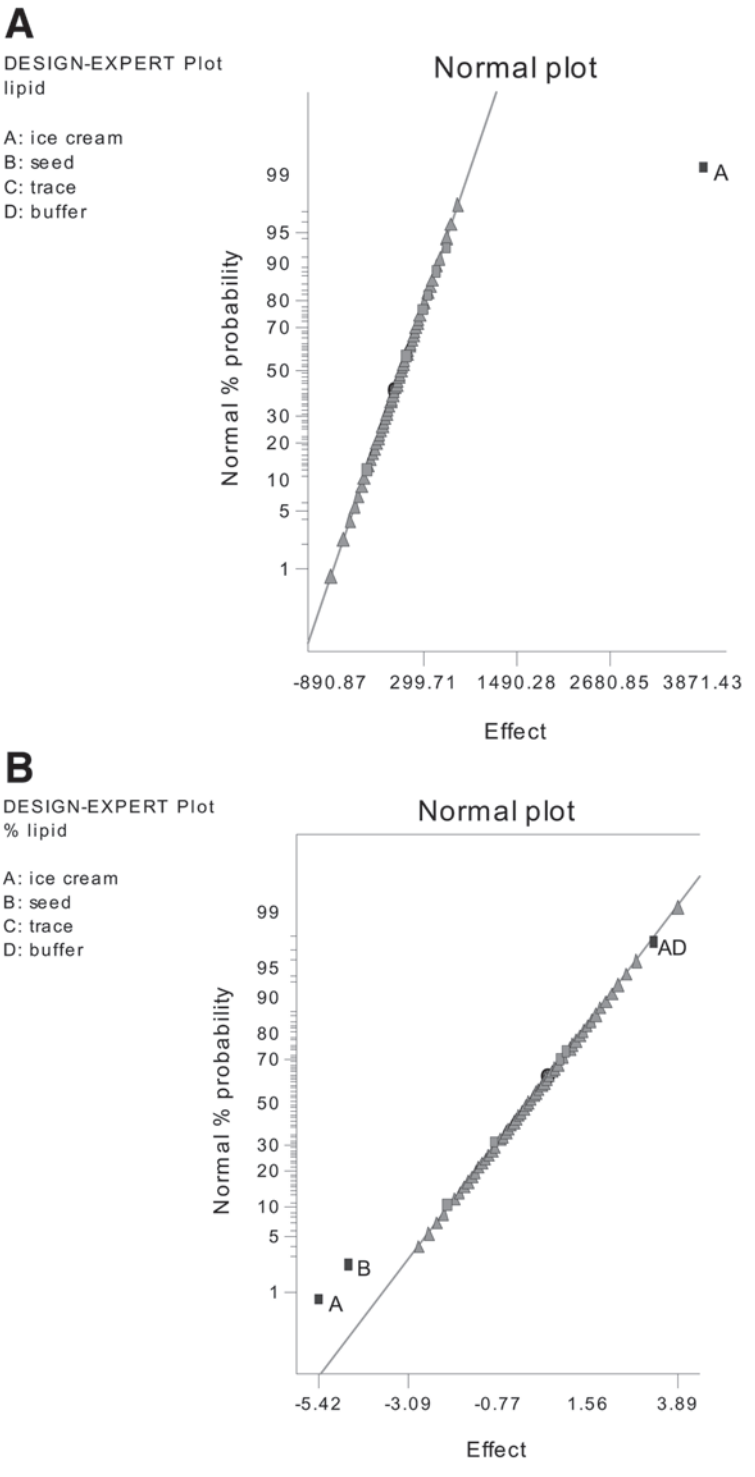

Fig. 1. Normal plots for 2^{4-1} fractional factorial design: **(A)** lipid; **(B)** % lipid.

Table 7
Analysis of Variance for Selected Model for PSA

a. Lipid
ANOVA table [partial sum of squares]

Variation source	Sum of squares	Degree of freedom	Mean square	F value	Prob > *F*
Model	158,445,118.00	4	39,611279.43	372.54	<0.0001
A (ice cream)	11,700.36	1	11,700.36	0.11	0.7424
B (time)	126,987,302.00	1	126,987,301.60	1194.30	<0.0001
B^2	31,076,000.10	1	31,076,000.06	292.27	<0.0001
AB	370,115.71	1	370,115.71	3.48	0.0719
Residual	3,189,839.40	30	106,327.98		
Cor total	161,634,957.00	34			
SD	326.08		R^2	0.98	
Mean	7129.29		Adjusted R^2	0.98	
CV	4.57		Pred R^2	0.97	
PRESS	4,471,971.32		Adeq precision	52.62	

b. % lipid
ANOVA table [partial sum of squares]

Variation source	Sum of squares	Degrees of freedom	Mean square	F value	Prob > *F*
Model	1516.22	3	505.41	31.40	<0.0001
B (time)	519.11	1	519.11	32.25	<0.0001
B^2	204.16	1	204.16	12.68	0.0012
B^3	116.47	1	116.47	7.24	0.0114
Residual	498.94	31	16.09		
Cor total	2015.16	34			
SD	4.01		R^2	0.75	
Mean	74.05		Adjusted R^2	0.73	
CV	5.42		Pred R^2	0.69	
PRESS	628.46		Adeq precision	13.34	

SD, standard deviation; CV, coefficient of variation; Cor, correction; Pred, predicted; Adeq, adequate; PRESS, predicted residue sum of squares.

Augmented Design for Phase II Optimization

There was a significant curvature effect and strong evidence of lack of fit in phase I, which implied that the optimization procedure needed to proceed using augmented second-order designs to estimate the fit and optimization of a higher-order effect (usually quadratic). This process was the second step (phase II) of the optimization procedure.

Since significant curvature was detected, the first-order design used in the phase I optimization (2^2 full factorial design) was augmented with axial design points (also called star points) on each factor to create a CCD. This design can estimate a quadratic effect from additional axial runs. To develop

an RCCD for a circular isocontour standard error of design, the 2^2 full factorial design was augmented with axial points at $\alpha = \sqrt{2}$ with additional center points and performed as shown in Table 8, block 2 and Fig. 3.

Analysis of Variance, Regression, and Prediction Equation

Table 8 indicates the design layout for RCCD and results collected over the fermentation period. The lack-of-fit tests and other diagnostic statistics are followed by the analysis of variance (ANOVA) table. Although the quadratic model has a low *p* value for the lack-of-fit test, the higher R^2 and adjusted R^2 indicate a good explanation of the variability by the selected model for lipid and % lipid, as shown in parts a and b of Table 9 and parts a and b of Table 10, respectively. Therefore, the quadratic model appears to be a reliable model for lipid and % lipid from RCCD. The significance of the lack-of-fit test seems to be owing to covering the wide range of fermentation time (24–240 h), which produces inappropriate mean value for the lack-of-fit test.

Part c of Table 9 and part c of Table 10 introduce regression analysis for the selected model for lipid and % lipid, respectively, by the stepwise model reduction method, in which the subset models are identified sequentially by addition or deletion. Stepwise regression is the combination of forward addition and backward elimination methods. The estimated regression coefficients equal one half the factorial effect in orthogonal designs. The standard error of the regression is the estimated standard deviation associated with the regression coefficient estimate. The 95% confidence interval (CI) gives the estimated range in which the true coefficient can be found.

The variance inflation factor (VIF) represents how much the variance of that model coefficient increases from the lack of orthogonality in the design. If a coefficient is orthogonal to the rest of the model term, then its VIF is 1, as shown in these model terms. The larger the VIF, the more the multicollinearity. In general, a term with a VIF greater than 10 should be avoided, owing to excessive multicollinearity.

The purpose of the prediction equation is to fit the data to the model for prediction or optimization. The final regression functions for lipid and % lipid in terms of coded factors, shown next, were used for making various graphic models:

Final equation in terms of coded factors:

Lipid (μg/mL)	% lipid (%)
12,154.22	86.03069
113.845 × *A* (ice cream)	–0.57597 × *A*
–70.8463 × *B* (buffer)	8.457571 × *C*
5742.705 × *C* (time)	–8.12918 × C^2
–472.063 × A^2	0.732886 × *A* * *C*
–142.063 × B^2	
–3379.22 × C^2	
696.4817 × *A* * *C*	
–279.168 × *B* * *C*	

Table 8
Design Layout and Results for RCCD[a]

Id	Run	Block	Factor 1 *A*:ice cream (mL)	Factor 2 *B*:buffer (mL)	Factor 3 *C*:time (h)	Resp. 1 DCW (µg/mL)	Resp. 2 **lipid** (µg/mL)	Resp. 3 *d* lipid (µg/mL·d)	**Resp. 4 % lipid (%)**	Resp. 5 *q* lipid (d^{-1})	Resp. 6 lipid productivity (µg/mL·h)
1	1	Block 1	40.00	5.00	24	4265	**2900**	2450	**68.00**	0.57	102.08
2	5	Block 1	60.00	5.00	24	2695	**1855**	1405	**68.82**	0.52	58.54
3	3	Block 1	40.00	10.00	24	4305	**2905**	2455	**67.47**	0.57	102.29
4	6	Block 1	60.00	10.00	24	3050	**2145**	1695	**70.33**	0.56	70.63
0	4	Block 1	50.00	7.50	24	3755	**2615**	2165	**69.64**	0.58	90.21
0	2	Block 1	50.00	7.50	24	3760	**2610**	2160	**69.51**	0.58	90.00
5	7	Block 2	35.86	7.50	24	4130	**2830**	2380	**68.53**	0.58	99.17
6	11	Block 2	64.14	7.50	24	3330	**2065**	1615	**62.08**	0.49	67.29
7	10	Block 2	50.00	3.96	24	3865	**2470**	2020	**63.94**	0.52	84.17
8	8	Block 2	50.00	11.04	24	3775	**2580**	2130	**68.35**	0.56	88.75
0	9	Block 2	50.00	7.50	24	3955	**2575**	2125	**65.10**	0.54	88.54
0	12	Block 2	50.00	7.50	24	3845	**2550**	2100	**66.32**	0.55	87.50
1	17	Block 1	40.00	5.00	48	7285	**5870**	2970	**80.58**	0.41	112.92
2	16	Block 1	60.00	5.00	48	5675	**4295**	2440	**75.71**	0.43	80.10
3	14	Block 1	40.00	10.00	48	7000	**5800**	2895	**82.87**	0.41	111.46
4	15	Block 1	60.00	10.00	48	6125	**4710**	2565	**76.97**	0.42	88.75
0	18	Block 1	50.00	7.50	48	6880	**5465**	2850	**79.43**	0.41	104.48
0	13	Block 1	50.00	7.50	48	6795	**5250**	2640	**77.33**	0.39	100.00
5	19	Block 2	35.86	7.50	48	7505	**5695**	2865	**76.00**	0.38	109.27
6	23	Block 2	64.14	7.50	48	6395	**4535**	2470	**70.96**	0.39	85.10
7	21	Block 2	50.00	3.96	48	6905	**5075**	2605	**73.51**	0.38	96.35
8	22	Block 2	50.00	11.04	48	6980	**5270**	2690	**75.52**	0.39	100.42
0	20	Block 2	50.00	7.50	48	7245	**5150**	2575	**71.18**	0.36	97.92
0	24	Block 2	50.00	7.50	48	6960	**5185**	2635	**74.51**	0.38	98.65
1	30	Block 1	40.00	5.00	72	10,650	**8620**	2750	**80.96**	0.26	113.47

2	28	Block 1	60.00	5.00	72	9260	**7380**	3085	**79.72**	0.33	96.25
3	29	Block 1	40.00	10.00	72	11,130	**9115**	3315	**81.90**	0.30	120.35
4	26	Block 1	60.00	10.00	72	9325	**7430**	2720	**79.69**	0.29	96.94
0	27	Block 1	50.00	7.50	72	10,530	**8400**	2935	**79.77**	0.28	110.42
0	25	Block 1	50.00	7.50	72	10,230	**8280**	3030	**80.94**	0.30	108.75
5	33	Block 2	35.86	7.50	72	9085	**7760**	2065	**85.59**	0.23	101.53
6	32	Block 2	64.14	7.50	72	8560	**6810**	2275	**79.59**	0.27	88.33
7	36	Block 2	50.00	3.96	72	9230	**7410**	2335	**80.29**	0.25	96.67
8	31	Block 2	50.00	11.04	72	9115	**7560**	2290	**82.96**	0.25	98.75
0	34	Block 2	50.00	7.50	72	9165	**7385**	2235	**80.59**	0.24	96.32
0	35	Block 2	50.00	7.50	72	9115	**7340**	2155	**80.56**	0.24	95.69
1	38	Block 1	40.00	5.00	96	11,660	**9850**	1230	**84.48**	0.11	97.92
2	37	Block 1	60.00	5.00	96	11,535	**9315**	1935	**80.78**	0.17	92.34
3	42	Block 1	40.00	10.00	96	12,875	**11,020**	1905	**85.59**	0.15	110.10
4	41	Block 1	60.00	10.00	96	11,900	**9615**	2185	**80.80**	0.18	95.47
0	39	Block 1	50.00	7.50	96	12,505	**10,185**	1785	**81.45**	0.14	101.41
0	40	Block 1	50.00	7.50	96	12,615	**10,295**	2015	**81.61**	0.16	102.55
5	46	Block 2	35.86	7.50	96	11,210	**8470**	710	**75.65**	0.06	83.54
6	47	Block 2	64.14	7.50	96	10,545	**8340**	1530	**79.09**	0.15	82.19
7	45	Block 2	50.00	3.96	96	11,205	**8790**	1380	**78.45**	0.12	86.88
8	43	Block 2	50.00	11.04	96	11,195	**9050**	1490	**80.87**	0.13	89.58
0	48	Block 2	50.00	7.50	96	11,310	**9245**	1860	**81.74**	0.16	91.61
0	44	Block 2	50.00	7.50	96	11,350	**9055**	1715	**79.79**	0.15	89.64
1	50	Block 1	40.00	5.00	120	12,460	**10,565**	715	**84.78**	0.06	84.29
2	49	Block 1	60.00	5.00	120	13,150	**11,190**	1875	**85.09**	0.14	89.50
3	53	Block 1	40.00	10.00	120	13,515	**11,785**	765	**87.19**	0.06	94.46
4	54	Block 1	60.00	10.00	120	13,620	**11,585**	1970	**85.06**	0.14	92.79
0	51	Block 1	50.00	7.50	120	13,875	**11,740**	1555	**84.61**	0.11	94.08
0	52	Block 1	50.00	7.50	120	14,305	**12,200**	1905	**85.29**	0.13	97.92
5	57	Block 2	35.86	7.50	120	11,875	**10,195**	1725	**85.85**	0.15	81.21
6	60	Block 2	64.14	7.50	120	11,925	**10,200**	1860	**85.55**	0.16	81.25

(continued)

Table 8 *(continued)*[a]

Id	Run	Block	Factor 1 *A*:ice cream (mL)	Factor 2 *B*:buffer (mL)	Factor 3 *C*:time (h)	Resp. 1 DCW (µg/mL)	Resp. 2 **lipid** (µg/mL)	Resp. 3 *d* lipid (µg/mL·d)	**Resp. 4 % lipid (%)**	Resp. 5 *q* lipid (d^{-1})	Resp. 6 lipid productivity (µg/mL·h)
7	58	Block 2	50.00	3.96	120	12,790	**11,010**	2220	**86.08**	0.17	88.00
8	55	Block 2	50.00	11.04	120	11,985	**10,415**	1365	**86.92**	0.11	83.04
0	56	Block 2	50.00	7.50	120	12,550	**10,835**	1590	**86.41**	0.13	86.54
0	59	Block 2	50.00	7.50	120	13,155	**10,875**	1820	**82.67**	0.14	86.88
1	62	Block 1	40.00	5.00	144	13,765	**11,975**	1410	**86.99**	0.10	80.03
2	64	Block 1	60.00	5.00	144	15,225	**13,045**	1855	**85.69**	0.12	87.47
3	65	Block 1	40.00	10.00	144	14,070	**12,685**	900	**90.16**	0.06	84.97
4	63	Block 1	60.00	10.00	144	15,365	**13,375**	1790	**87.05**	0.12	89.76
0	66	Block 1	50.00	7.50	144	15,525	**13,900**	2160	**89.54**	0.14	93.40
0	61	Block 1	50.00	7.50	144	15,950	**13,810**	1610	**86.58**	0.10	92.78
5	67	Block 2	35.86	7.50	144	12,495	**10,700**	505	**85.63**	0.04	71.18
6	69	Block 2	64.14	7.50	144	12,985	**11,340**	1140	**87.35**	0.09	75.63
7	71	Block 2	50.00	3.96	144	14,250	**12,265**	1255	**86.07**	0.09	82.05
8	70	Block 2	50.00	11.04	144	13,070	**11,295**	880	**86.45**	0.07	75.31
0	68	Block 2	50.00	7.50	144	13,710	**12,070**	1235	**88.04**	0.09	80.69
0	72	Block 2	50.00	7.50	144	13,960	**12,230**	1355	**87.63**	0.10	81.81
1	75	Block 1	40.00	5.00	168	15,100	**13,175**	1200	**87.25**	0.08	75.74
2	76	Block 1	60.00	5.00	168	15,965	**14,190**	1145	**88.91**	0.07	81.79
3	73	Block 1	40.00	10.00	168	14,430	**12,505**	–180	**86.66**	–0.01	71.76
4	77	Block 1	60.00	10.00	168	15,970	**14,075**	700	**88.14**	0.04	81.10
0	78	Block 1	50.00	7.50	168	16,485	**14,515**	615	**88.05**	0.04	83.72
0	74	Block 1	50.00	7.50	168	16,810	**14,790**	980	**87.99**	0.06	85.36
5	81	Block 2	35.86	7.50	168	12,695	**10,930**	230	**86.10**	0.02	62.38
6	83	Block 2	64.14	7.50	168	14,365	**12,165**	825	**84.69**	0.06	69.73
7	80	Block 2	50.00	3.96	168	14,900	**12,995**	730	**87.21**	0.05	74.67
8	82	Block 2	50.00	11.04	168	14,645	**12,520**	1225	**85.50**	0.08	71.85

0	79	Block 2	50.00	7.50	168	14,980	**12,965**	895	**86.55**	0.06	74.49
0	84	Block 2	50.00	7.50	168	14,995	**13,100**	870	**87.36**	0.06	75.30
1	86	Block 1	40.00	5.00	192	16,055	**14,140**	965	**88.07**	0.06	81.49
2	87	Block 1	60.00	5.00	192	16,795	**14,885**	695	**88.63**	0.04	85.92
3	89	Block 1	40.00	10.00	192	14,760	**12,965**	460	**87.83**	0.03	74.49
4	90	Block 1	60.00	10.00	192	16,620	**14,810**	735	**89.11**	0.04	85.48
0	88	Block 1	50.00	7.50	192	17,270	**15,280**	765	**88.48**	0.04	88.27
0	85	Block 1	50.00	7.50	192	17,495	**15,440**	650	**88.25**	0.04	89.23
5	96	Block 2	35.86	7.50	192	13,075	**11,290**	360	**86.34**	0.03	64.52
6	93	Block 2	64.14	7.50	192	15,160	**13,060**	895	**86.15**	0.06	75.06
7	94	Block 2	50.00	3.96	192	15,670	**13,775**	780	**87.91**	0.05	79.32
8	92	Block 2	50.00	11.04	192	15,175	**12,975**	455	**85.50**	0.03	74.55
0	95	Block 2	50.00	7.50	192	15,595	**13,735**	770	**88.09**	0.05	79.08
0	91	Block 2	50.00	7.50	192	15,790	**13,840**	740	**87.64**	0.05	79.70
1	97	Block 1	40.00	5.00	216	16,550	**14,460**	320	**87.37**	0.02	83.39
2	100	Block 1	60.00	5.00	216	17,990	**15,355**	470	**85.35**	0.03	88.72
3	99	Block 1	40.00	10.00	216	14,995	**12,755**	–210	**85.06**	–0.01	73.24
4	98	Block 1	60.00	10.00	216	17,225	**14,735**	–75	**85.55**	0.00	85.03
0	101	Block 1	50.00	7.50	216	17,435	**15,105**	–175	**86.63**	–0.01	87.23
0	102	Block 1	50.00	7.50	216	17,870	**15,375**	–65	**86.04**	0.00	88.84
5	103	Block 2	35.86	7.50	216	13,205	**11,425**	135	**86.53**	0.01	65.33
6	104	Block 2	64.14	7.50	216	16,080	**13,550**	490	**84.27**	0.03	77.98
7	107	Block 2	50.00	3.96	216	16,015	**14,065**	290	**87.82**	0.02	81.04
8	108	Block 2	50.00	11.04	216	15,720	**13,365**	390	**85.02**	0.02	76.88
0	106	Block 2	50.00	7.50	216	16,115	**14,080**	345	**87.37**	0.02	81.13
0	105	Block 2	50.00	7.50	216	16,000	**14,090**	250	**88.06**	0.02	81.19
1	112	Block 1	40.00	5.00	240	17,055	**15,140**	680	**88.77**	0.04	87.44
2	110	Block 1	60.00	5.00	240	17,970	**15,145**	–210	**84.28**	–0.01	87.47
3	109	Block 1	40.00	10.00	240	15,030	**13,170**	415	**87.63**	0.03	75.71
4	114	Block 1	60.00	10.00	240	17,005	**14,945**	210	**87.89**	0.01	86.28
0	113	Block 1	50.00	7.50	240	17,540	**15,325**	220	**87.37**	0.01	88.54

(continued)

Table 8 *(continued)*[a]

Id	Run	Block	Factor 1 *A*:ice cream (mL)	Factor 2 *B*:buffer (mL)	Factor 3 *C*:time (h)	Resp. 1 DCW (µg/mL)	Resp. 2 **lipid** (µg/mL)	Resp. 3 *d* lipid (µg/mL·d)	**Resp. 4 % lipid (%)**	Resp. 5 *q* lipid (d^{-1})	Resp. 6 lipid productivity (µg/mL·h)
0	111	Block 1	50.00	7.50	240	17,855	**15,660**	285	**87.71**	0.02	90.54
5	116	Block 2	35.86	7.50	240	13,825	**12,030**	605	**87.01**	0.04	68.93
6	115	Block 2	64.14	7.50	240	15,455	**13,805**	255	**89.34**	0.02	79.49
7	117	Block 2	50.00	3.96	240	15,915	**13,915**	–150	**87.44**	–0.01	80.15
8	120	Block 2	50.00	11.04	240	15,145	**13,595**	230	**89.77**	0.02	78.24
0	118	Block 2	50.00	7.50	240	15,880	**14,315**	235	**90.15**	0.01	82.53
0	119	Block 2	50.00	7.50	240	15,955	**14,240**	150	**89.25**	0.01	82.08

[a]Bold data used for data analysis (ANOVA and regression) and modeling. Resp., response.

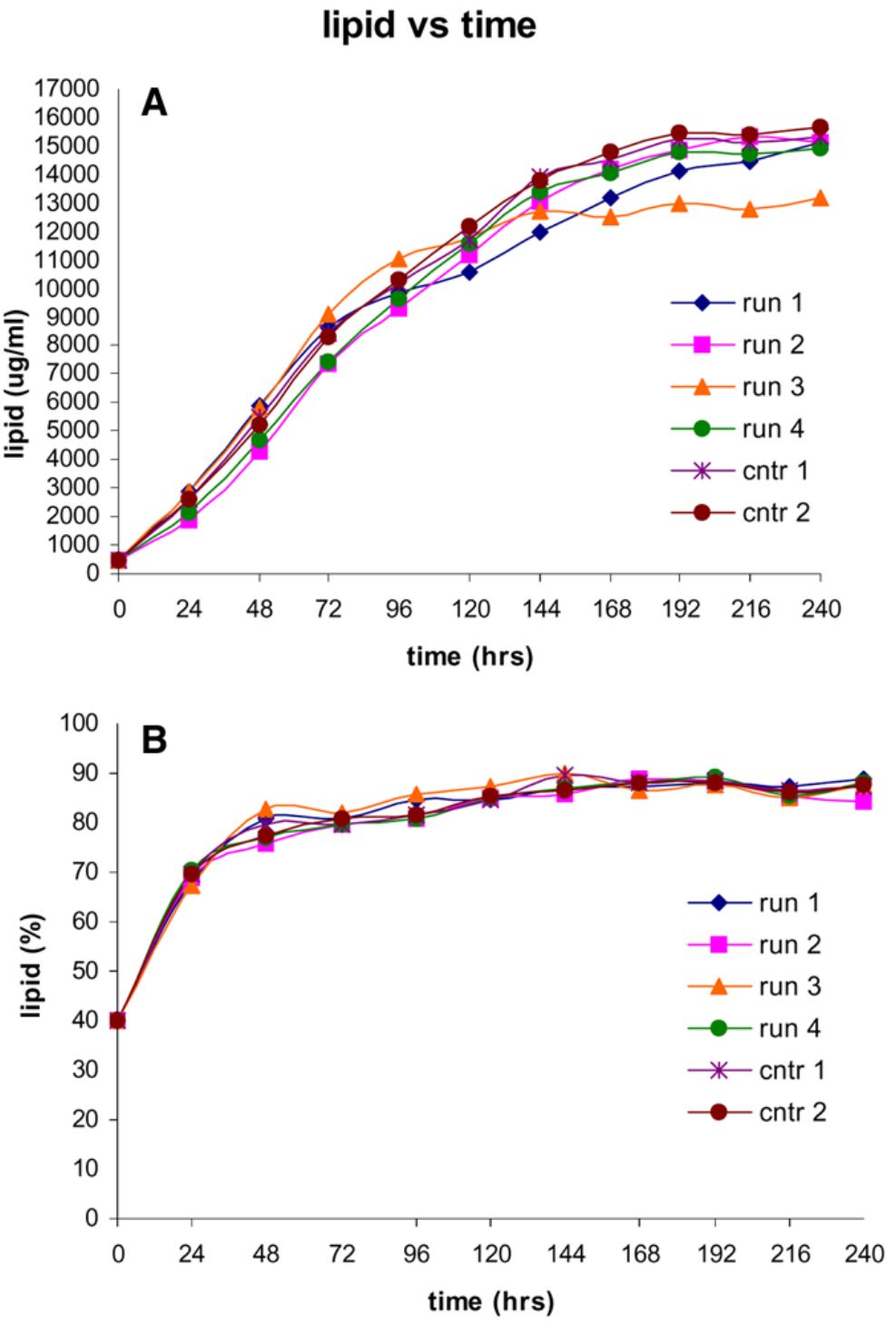

Fig. 2. Time sequences responses for 2^2 full factorial design of RCCD: **(A)** lipid; **(B)** % lipid.

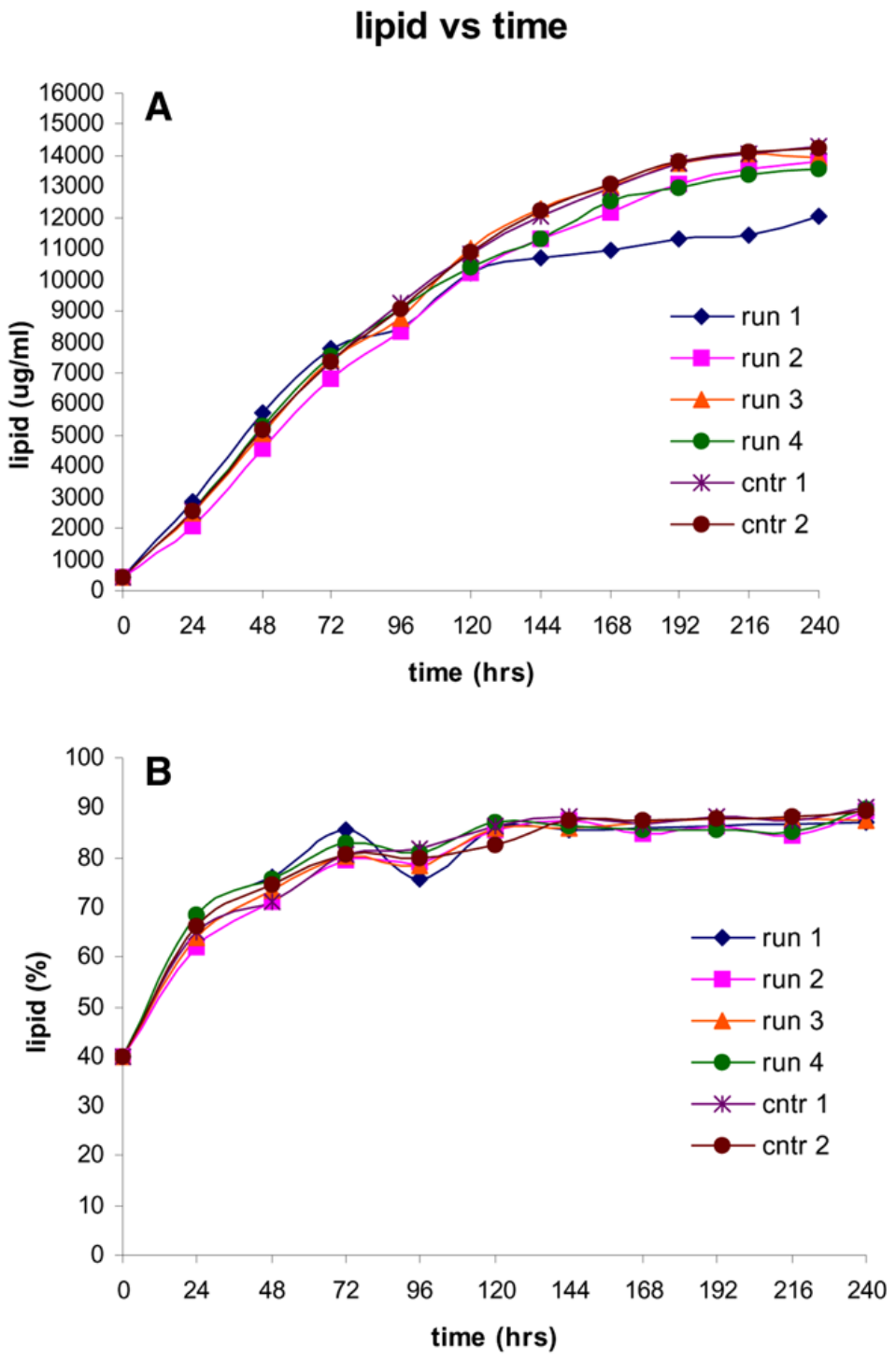

Fig. 3. Time sequences responses for axial (augmented) design of RCCD: **(A)** lipid; **(B)** % lipid.

Table 9
ANOVA and Regression for Selected Model for Lipid From RCCD

a. Stepwise regression with alpha to enter = 0.100, alpha to exit = 0.100
Forced terms: intercept, block 1

Term added	Coefficient estimate	t for H_0 coefficient = 0	Prob > |t|	R^2	MSE
C	5742.71	28.27	<0.0001	0.87	2,017,031.00
C^2	–3379.22	–18.54	<0.0001	0.97	513,382.10
AC	696.48	6.45	<0.0001	0.98	380,365.90
A^2	–443.65	–6.99	<0.0001	0.98	268,600.00
BC	–279.17	–3.20	0.0018	0.98	248,498.20
B^2	–142.06	–2.32	0.0219	0.99	239,184.60
A	113.85	2.11	0.0367	0.99	231,998.30

Hierarchical terms added after stepwise regression: B
b. Analysis of variance table [Partial sum of squares]

Variation source	Sum of squares	Degrees of freedom	Mean square	F value	Prob > F
Block	29,274,441	1	29,274,440.83		
Model	1.82E+09	8	227,866,517.30	988.76	<0.0001
A	1,036,855	1	1,036,854.87	4.50	0.0362
B	401,535.5	1	401,535.53	1.74	0.1896
C	1,610,000,000	1	1,612,289,824.00	6996.05	<0.0001
A^2	14,261,952	1	14,261,952.25	61.89	<0.0001
B^2	1,291,632	1	1,291,632.25	5.60	0.0197
C^2	176,000,000	1	176,440,275.00	765.61	<0.0001
AC	15,810,235	1	15,810,234.91	68.60	<0.0001
BC	2,540,094	1	2,540,093.79	11.02	0.0012
Residual	25,350,280	110	230,457.09		
Lack of fit	25,009,805	90	277,886.72	16.32	<0.0001
Pure error	340,475	20	17,023.75		
Cor total	1,880,000,000	119			
SD	480.0595		R^2	0.99	
Mean	10,368.08		Adjusted R^2	0.99	
CV	4.630166		Pred R^2	0.98	
PRESS	31,183,937		Adeq precision	106.71	

c. Regression analysis

Factor	Coefficient estimate	Degrees of freedom	SE	95% CI low	95% CI high	VIF
Intercept	12,154.22	1	90.76	11,974.36	12,334.08	
Block 1	493.92	1				
Block 2	–493.92					
A–ice cream	113.85	1	53.67	7.48	220.21	1.00
B–buffer	–70.85	1	53.67	–177.21	35.52	1.00
C–time	5742.71	1	68.66	5606.64	5878.77	1.00
A^2	–472.06	1	60.01	–590.98	–353.14	1.04
B^2	–142.06	1	60.01	–260.98	–23.14	1.04
C^2	–3379.22	1	122.13	–3621.25	–3137.19	1.00
AC	696.48	1	84.09	529.84	863.13	1.00
BC	–279.17	1	84.09	–445.81	–112.53	1.00

SD, standard deviation; CV, coefficient of variation; PRESS, predicted residue sum of squares; SE, standard error; MSE, mean square error; CI, confidence interval; VIF, variance inflation factor; Pred, predicted; Adeq, adequate; Cor, correction.

Table 10
ANOVA and Regression for Selected Model for % Lipid From RCCD

a. Stepwise regression with alpha to enter = 0.100, alpha to exit = 0.100
Forced terms: intercept, block 1

Term added	Coefficient estimate	t for H_0 coefficient = 0	Prob > | t |	R^2	MSE
C	8.46	15.73	<0.0001	0.68	14.14
C^2	–8.13	–13.68	<0.0001	0.88	5.46
A	–0.58	–2.24	0.0268	0.88	5.27
AC	0.73	1.84	0.0683	0.89	5.17

b. ANOVA table [partial sum of squares]

Variation source	Sum of squares	Degrees of freedom	Mean square	F value	Prob > F
Block	43.80	1	43.80		
Model	4562.17	4	1140.54	220.75	<0.0001
A	26.54	1	26.54	5.14	0.0253
C	3497.05	1	3497.05	676.84	<0.0001
C^2	1021.08	1	1021.08	197.63	<0.0001
AC	17.51	1	17.51	3.39	0.0683
Residual	589.01	114	5.17		
Lack of fit	564.90	94	6.01	4.99	<0.0001
Pure error	24.11	20	1.21		
Cor total	5194.97	119			
SD	2.27		R^2	0.89	
Mean	82.72		Adjusted R^2	0.88	
CV	2.75		Pred R^2	0.87	
PRESS	670.78		Adeq precision	43.29	

c. Regression analysis

Factor	Coefficient estimate	Degrees of freedom	SE	95% CI low	95% CI high	VIF
Intercept	86.03	1	0.31	85.41	86.65	
Block 1	0.60	1				
Block 2	–0.60					
A–ice cream	–0.58	1	0.25	–1.08	–0.07	1.00
C–time	8.46	1	0.33	7.81	9.10	1.00
C^2	–8.13	1	0.58	–9.27	–6.98	1.00
AC	0.73	1	0.40	–0.06	1.52	1.00

SD, standard deviation; CV, coefficient of variation; PRESS, prediction residue sum of squares; SE, standard error; MSE, mean square error; CI, confidence interval; VIF, variance inflation factor; Pred, predicted; Adeq, adequate; Cor, correction.

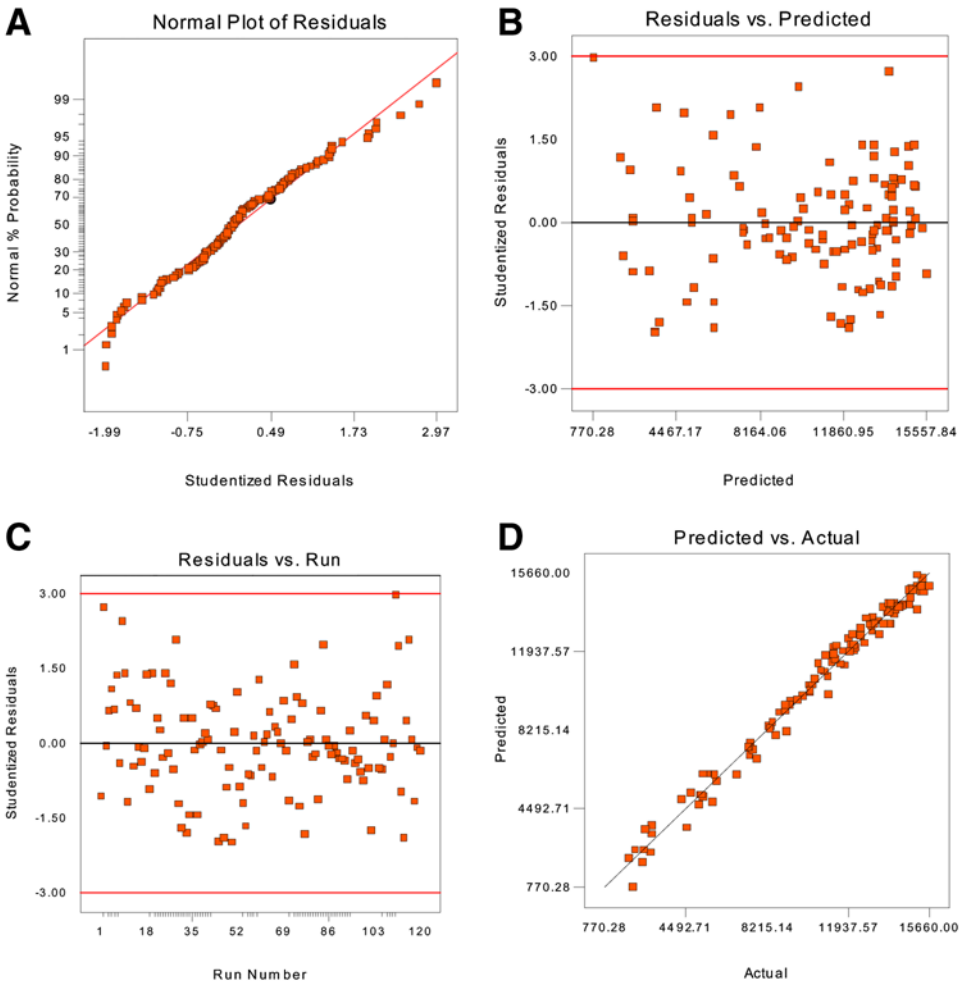

Fig. 4. Residual diagnostics of RCCD model for lipid: **(A)** normality; **(B)** residual vs predicted response; **(C)** residual vs run order; **(D)** predicted response vs actual.

Residual Analysis: Assessment of Model Assumptions

Before accepting any model, the adequacy of the adopted model should be checked by the appropriate statistical method. The model assumptions given next are on the error terms (ε_is) that are mimicked by the residuals, the difference between the observed value of the response variable, and the value predicted by the regression function:

Assumptions on the error terms: $\varepsilon_i \sim N(0, \sigma^2)$ with independent and identical distribution
1. $E(\varepsilon_i) = 0$ (zero mean)
2. Var $(\varepsilon_i) = \sigma^2$ (constant variance)
3. Independent ε_i (no correlation)
4. Normality of ε_i distribution (random variable)

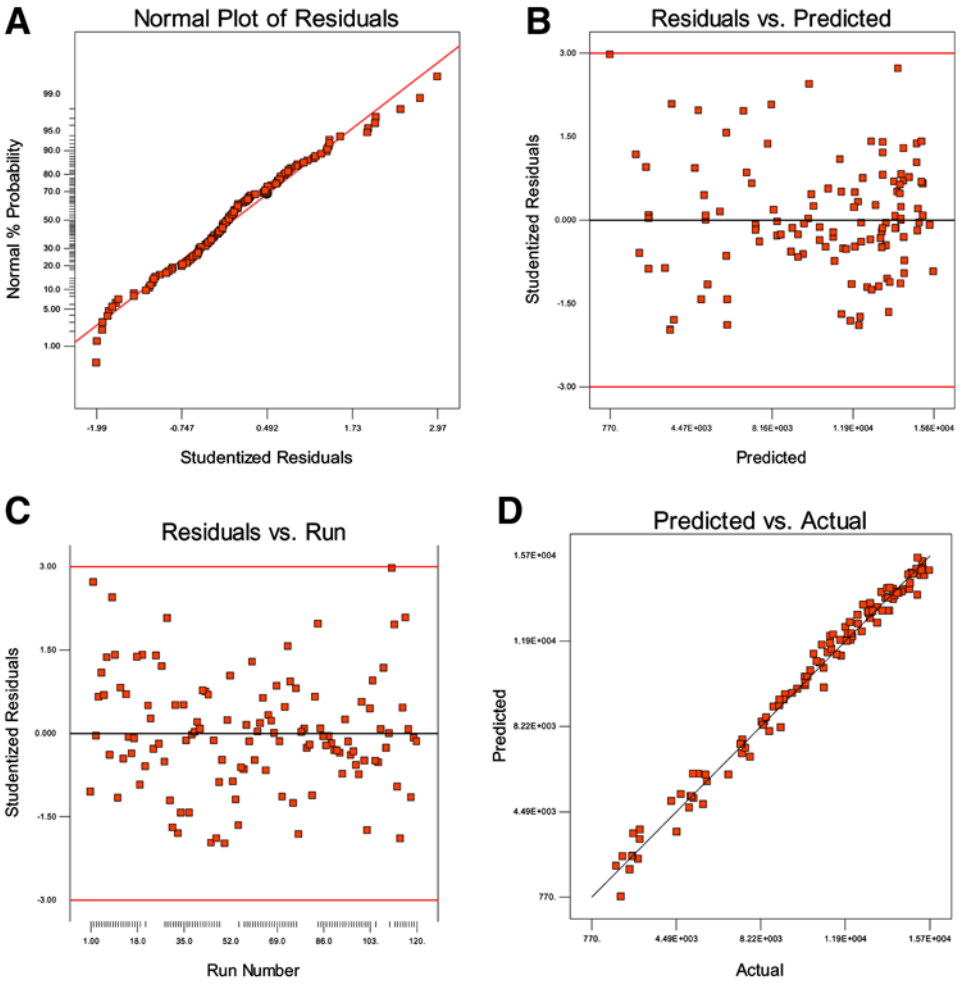

Fig. 5. Residual diagnostics of RCCD model for % lipid: **(A)** normality; **(B)** residual vs predicted response; **(C)** residual vs run order; **(D)** predicted response vs actual.

The major diagnostic method is residual analysis, as shown in Figs. 4 and 5, providing diagnostics for residual behavior. There are several residual plots to test the model assumptions. The primary analysis is to examine a normal probability plot of the studentized residuals, i.e., the number of standard deviations of the actual values from their respective predicted values (Figs. 4A and 5A). The normal probability plot is employed to determine whether the residuals follow a normal distribution, which is the most important assumption for checking statistical modeling and model adequacy.

The next analysis is to look at the residuals plotted vs the predicted responses (Figs. 4B and 5B). There should be no systemic pattern in the plot, and the points should fall within a horizontal band centered at zero. Departure from this may suggest a violation of the constant variance assumption.

The size of the studentized residual should be independent of its predictive value, which means that the spread should be about the same across all levels of the predicted values. Graphs of residual vs run order (number) reveal any time-based effects or sequential component (Figs. 4C and 5C). Actual vs predicted displays the real response data plotted against the predicted responses (Figs. 4D and 5D). Points above or below the diagonal line mean areas of over- or underprediction. Graphs of residual vs factors also need to be examined to determine whether there might be variance changes (dispersion effect) with any factor or level of factor. There were no significant violations of the model assumptions found in this residual analysis, as shown in Figs. 4 and 5. Therefore, the modeling can be used for further studies such as optimization, factor analysis, or simulation without any bias.

Modeling

Figures 6 and 7 show models with contour and three-dimensional formation of lipid and % lipid, respectively, representing model testing for the dynamic fermentation process separated by different fermentation times (24, 120, and 240 h). Toggled (flagged) points in the isoresponse contour plot represent tentative optimal points for the corresponding response at specific times, and dots on the plot indicate design points created by RCCD.

The unique characteristic of this experimental design is that it is a combination of response surface method ([RSM] RCCD) with repeated measures technique, which is able to show statistical effects and the dynamic nature of the process simultaneously from the single experimental design. The RSM combined with repeated measures is a novel experimental design consisting of multistage response surfaces explained by medium component factors (ice cream and buffer) along with a process factor (fermentation time). One can predict a response within experimental range by using conceptional model graphs built from experimental data and optimize the process by using a nonlinear programming algorithm (simplex optimization).

Optimization

Numerical optimization was carried out to calculate the optimal factor combination (Table 11). Once the optimal factor settings are found, the model is used to estimate the value of the responses at those settings (Table 12). Numerical optimization can be represented by a general nonlinear algorithm with constraints applied to the main objective function, which is a desirability function for multiple responses. In numerical optimization, the desired goals (constraints) for each response and factor, such as maximize, minimize, target, within range, or none, are selected along with weight and importance that can be assigned to each goal. A weight for each goal can adjust the shape of the desirability function, and the importance of each

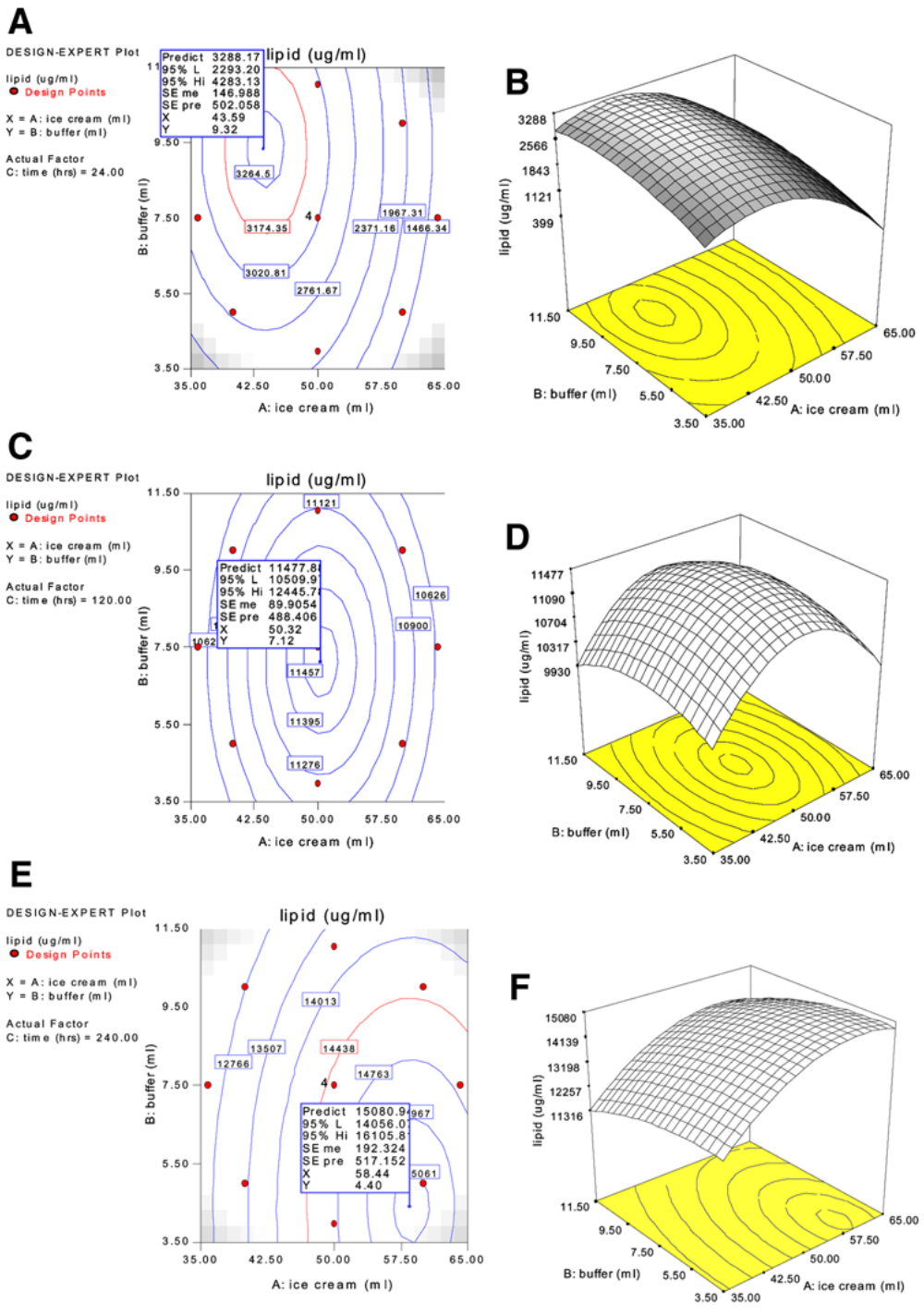

Fig. 6. Model graph of RCCD for lipid: **(A,B)** 24 h; **(C,D)** 120 h; **(E,F)** 240 h.

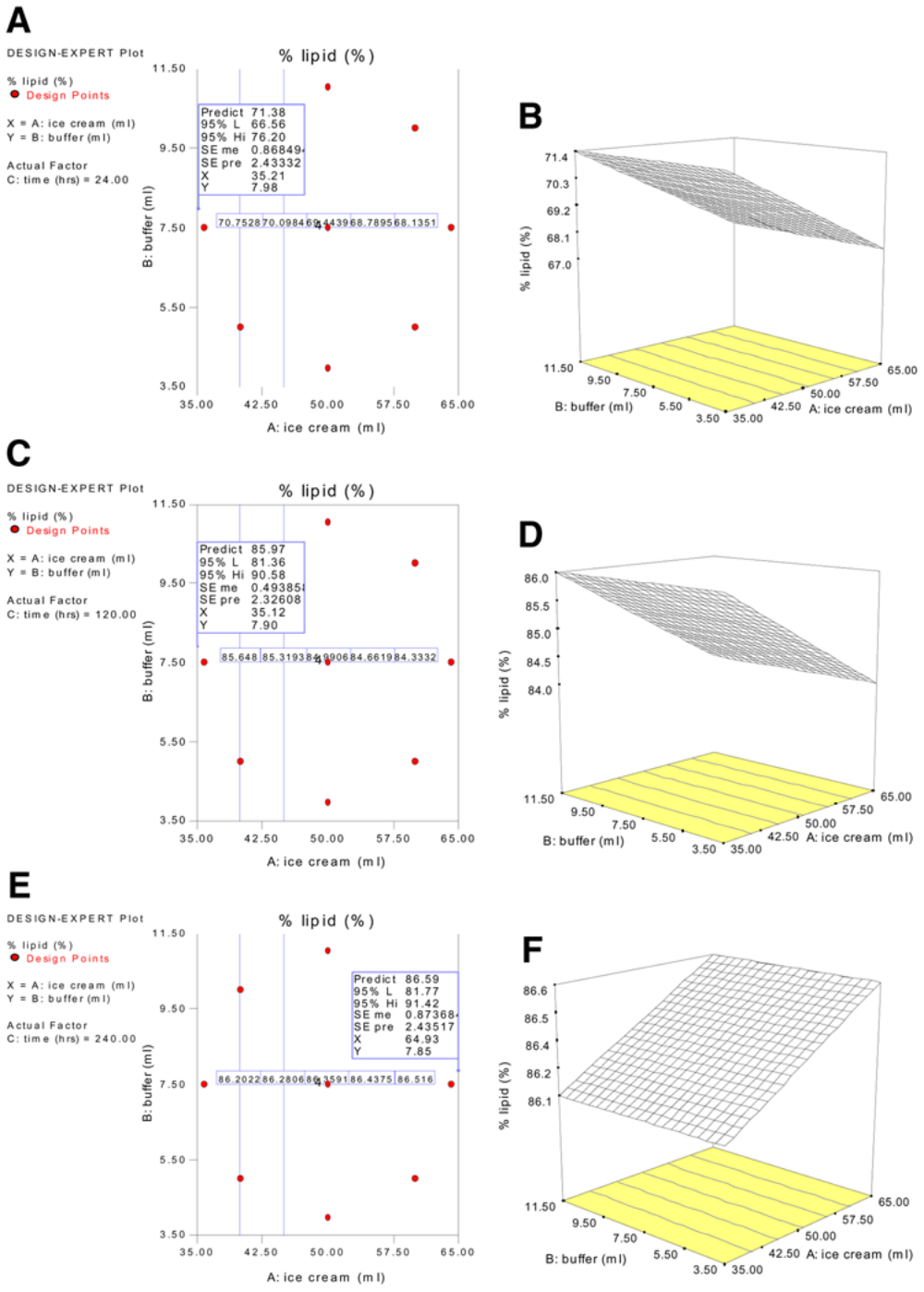

Fig. 7. Model graph of RCCD for % lipid: **(A,B)** 24 h; **(C,D)** 120 h; **(E,F)** 240 h.

Table 11
Numerical Optimization for RCCD

a. Constraints

Name	Goal	Lower limit	Upper limit	Lower weight	Upper weight	Importance
Ice cream (mL)	Within range	40	60	1	1	3
Buffer (mL)	Within range	5	10	1	1	3
Time (h)	Within range	24	240	1	1	3
DCW (μg/mL)	Within range	2695	17990	1	1	3
Lipid (μg/mL)	Maximize	1855	15660	1	1	5
d lipid (μg/[mL · d])	Within range	0	3315	1	1	3
% Lipid (%)	Maximize	62.08	90.16	1	1	4
q Lipid (d^{-1})	Within range	0.00	0.58	1	1	3
Lipid productivity (μg/[mL · h])	Within range	58.54	120.35	1	1	3

b. Solutions

Ice cream (mL)	Buffer (mL)	Time (h)	DCW (μg/mL)	Lipid (μg/mL)	d Lipid (μg/[mL · d])	% Lipid (%)	q Lipid (d^{-1})	Lipid productivity (μg/[mL · h])	Desirability
56.68	5.03	213.76	17,075.6	14,922.8	462.27	87.76	0.02	86.34	0.93
52.83	5.00	211.95	16,944.6	14,829.4	462.74	87.83	0.018	85.61	0.93
56.39	5.79	212.90	17,049.5	14,898.2	470.72	87.79	0.02	85.76	0.93

DCW, dry cell weight.

Table 12
Point Optimization and Confidence Interval for RCCD

a. Selected optimal conditions

Factor	Name	Level	Low level	High level
A	Ice cream (mL)	56.68	40	60
B	Buffer (mL)	5.03	5	10
C	Time (mL)	213.76	24	240

b. Prediction

Responses	Prediction	SEM	95% CI low	95% CI high	SE Prediction	95% PI low	95% PI high
DCW (μg/mL)	16,441.57	147.90	16,148.47	16,734.67	588.66	15,274.98	17,608.16
Lipid (μg/mL)	14,358.38	124.61	14,111.43	14,605.32	495.97	13,375.48	15,341.27
d Lipid (μg/[mL · d])	462.26	71.63	320.39	604.13	453.08	–435.12	1359.64
% Lipid (%)	87.76	0.43	86.92	88.61	2.31	83.18	92.34
q Lipid (d^{-1})	0.02	0.01	0.01	0.03	0.04	–0.05	0.09
Lipid productivity (μg/[mL · h])	82.66	1.85	78.99	86.32	7.53	67.74	97.58

SEM, standard error of mean; CI, confidence interval; PI, prediction interval; DCW, dry cell weight.

goal can change the relation to other goals. The goals are combined into an overall desirability function, which is an objective function of optimization, with its outcome ranging from zero (beyond the goal limits) to one (matching the exact goal).

Numerical optimization seeks a point that maximizes this desirability function. All goals become combined into one desirability function that is selected from various responses and factors. Searching for the optimal point begins at a random starting factor combination and continues up the steepest slope to a maximum. There may be several maximums, owing to curvature of the response surface and their combination into the desirability function. Table 11 introduces constraints for the responses and factors, and optimal points based on the desirability function.

Point prediction is used to make predictions for responses at any factor combination (setting). In Table 12, the average response is estimated with 95% confidence at the optimal factor settings. Also given is a 95% prediction interval (PI) for an individual observation of the response at the optimal factor settings. The standard error of the estimated average response (SE Mean) is always smaller than the standard error of the estimated individual response (SE Pred), making for a smaller margin of error and a more precise interval estimate. These values tell one what to expect for an individual verification or confirmation experiment. Expectations of the process can be managed by referring to these predictions.

The final optimal factor settings for fermentation were calculated based on the quadratic response surface model fitted to RCCD. From the optimization method, optimal factor combinations (56.68 mL of ice cream or 56.68% ice cream [v/v], 5.03 mL of buffer, and 213.76 h of fermentation time) with the high desirability of 0.93 were selected for further verification, robustness, and formulation studies, as shown in Tables 11 and 12.

Verification by Mixture Design

Adequacy of the RCCD model equation for predicting optimization of lipid production was tested at the selected optimal conditions. They are not identical to the ones from RCCD optimization but modified by a proportion of each component to fit the mixture design points. Experimental data from the mixture design were compared with predicted values from RCCD model equations in Table 13. Both data were reasonably close, indicating that the model was adequate for fermentation process optimization, as displayed in Tables 11 and 12. In other words, the experimental values from the crossed design points (mixture design combined with process factor) for verification studies were within the predicted range of 95% PI from the point estimation, which confirms the function of RCCD to predict and optimize the lipid production process. Robustness and formulation studies by crossed design will be discussed in another article *(15)*.

Table 13
Verification for Optimization

Model vs experiment	Ice cream (mL)	Buffer (mL)	Time (h)	DCW (µg/mL)	Lipid (µg/mL)	*d* Lipid (µg/[mL · d])	% lipid (%)	*q* Lipid (d^{-1})	Lipid productivity (µg/[mL · h])
Prediction	56.68	5.03	213.76	17,075.60	14,922.80	462.27	87.76	0.02	86.34
Verification 1	51.67	6.67	216.00	17,435.00	15,105.00	–175.00	86.63	–0.01	87.23
Verification 2	51.67	6.67	216.00	17,870.00	15,375.00	–65.00	86.04	0.00	88.84
95% PI low	NA	NA	NA	15,274.98	13,375.48	–435.12	83.18	–0.05	67.74
95% PI high	NA	NA	NA	17,608.16	15,341.27	1359.64	92.34	0.09	97.58

NA, not applicable; DCW, dry cell weight.

Acknowledgments

We thank the Environment Sciences Program for financial support, Dr. Debra Ingram (Arkansas State University) for statistical discussions, Dr. Carmen Scholz (University of Alabama at Huntsville) for NMR analysis, and Yarnell Ice Cream Company for providing raw material.

References

1. Kessler, B. and Witholt, B. (1999), in *Encyclopedia of Bioprocess Technology, volume 4,* Flickinger, M. C. and Drew, S. W., eds., John Wiley & Sons, New York, pp. 2024–2040.
2. Kubasek, N. K. and Silverman, G. S. (1997), *Environmental Law,* Prentice Hall, Upper Saddle River, NJ.
3. Swift, G. (1993), *Acc. Chem. Res.* **26,** 105–110.
4. Brandl, H., Bachofen, R., Mayer, J., and Wintermantel, E. (1995), *Can. J. Microbiol.* **41(Suppl. 1),** 143–153.
5. Kim, D. Y. and Rhee, Y. H. (2003), *Appl. Microbiol. Biotechnol*. **61,** 300–308.
6. Kim, M.-N., Lee, A.-R., Yoon, J.-S., and Chin, I.-J. (2000), *Eur. Polym. J.* **36,** 1677–1685.
7. Sang, B.-I, Hori, K., Tanji, Y., and Unno, H. (2002), *Appl. Microbiol. Biotechnol*. **58,** 241–247.
8. Ayorinde, F. O., Saeed, K. A., Price, E., Murrow, A., Collins, W. E., McInnis, F., Pollack, S. K., et al. (1998), *J. Ind. Microbiol. Biotechnol.* **21,** 46–50.
9. Yu, P. H. F., Chura, H., Huang, A. L., Lo, W. H., and Ho, K. P. (1999), *Water Sci. Technol.* **40(1),** 365–370.
10. Ahn, W. S., Park, S. J., and Lee, S. Y. (2000), *Appl. Environ. Microbiol.* **66(8),** 3624–3627.
11. Fuchtenbusch, B., Wullbrandt, D., and Steinbuchel, A. (2000), *Appl. Microbiol. Biotechnol*. **53,** 167–172.
12. Hong, K., Leung, Y. C., Kwok, S. Y., Law, K. H., Lo, W. H., Chua, H., and Yu, P. H. (2000), *Appl. Biochem. Biotechnol*. **84–-86,** 381–390.
13. Wong, A. L., Chua, H., and Yu, P. H. F. (2000), *Appl. Biochem. Biotechnol.*. **84–86,** 843–857.
14. Yu, P. H., Chua, H., Huang, A. L., Lo, W., and Chen, G. Q. (1998), *Appl. Biochem. Biotechnol*. **70(72),** 603–614.
15. Lee, K.-M. and Gilmore, D. F. (2005), *Process Biochem*. **40,** 226–249.
16. Park, S. J., Park, J. P., and Lee, S. Y. (2002), *Biotechnol. Lett*. **24,** 185–189.
17. Reddy, C. S. K., Ghai, R., and Kalia, V. C. (2003), *Bioresour. Technol*. **87,** 137–146.
18. Ashby, R. D. and Foglia, T. A. (1998), *Appl. Microbiol. Biotechnol*. **49,** 431–437.
19. Cromwick, A.-M., Foglia, T., and Lenz, R. W. (1996), *Appl. Microbiol. Biotechnol.* **46,** 464–469.
20. Haaland, P. D. (1989), *Experimental Design in Biotechnology,* Marcel Dekker, New York.
21. Montgomery, D. C. (2004), *Design and Analysis of Experiments,* 6th ed., John Wiley & Sons, New York.
22. Myers, R. H. and Montgomery, D. C. (2002), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments,* 2nd ed., John Wiley & Sons, New York.
23. Kuehl, R. O. (2000), *Design of Experiments*: *Statistical Principle of Research Design and Analysis,* 2nd ed., Brooks/Cole, Pacific Crove, CA.
24. Oehlert, G. W. (2000), *A First Course in Design and Analysis of Experiments,* W. H. Freeman and Company, New York.
25. Wu, C. F. J. and Hamada, M. (2000), *Experiments: Planning, Analysis, and Parameter Design Optimization,* John Wiley & Sons, New York.
26. Dean, A. and Voss, D. (1999), *Design and Analysis of Experiments,* Springer-Verlag, New York.
27. Taylor, B. W. III. (1999), *Introduction to Management Science,* 6th ed., Prentice Hall, Upper Saddle River, NJ.
28. Mason, R. L., Gunst, R. F., and Hess, J. C. (1989), *Statistical Design and Analysis of Experiments with Applications to Engineering and Science,* John Wiley & Sons, New York.
29. Floros, J. D. and Chinan, M. S. (1988), *Food Technol.* **42(20),** 72–78.
30. Bayne, C. and Rubin, I. B. (1986), *Practical Experimental Designs and Optimization Methods for Chemists,* VCH, Deerfield, FL.
31. Hendrix, C. D. (1980), *Chemtech* August:488–497.
32. Matavulj, M. and Molitoris, H. P. (1992), *FEMS Microbiol. Rev*. **103,** 323–331.