# Design of Experiment (DOE) Approach in Cell Culture Medium Optimization

**Steven C. Peppers, Ph.D., Damon L. Talley, Heather N. Loke and Mathew V. Caple**
**Cell Culture R&D, Sigma-Aldrich Biotechnology**
**PO Box 14508, St. Louis, MO 63178 Phone: (314) 289-8496 Ext. 3183 Email: speppers@sial.com**

## Summary

The global capacity shortfall for production of biopharmaceuticals has given emphasis to the need for optimizing cell culture media for maximal productivity. We have developed a comprehensive approach to medium optimization based on Design of Experiment (DOE) statistical methods. Using a DOE approach makes sense because of the many interactions among medium components, the high intrinsic variance in cell culture assays, and the restriction on numbers of conditions per experiment. Our DOE approach is multidimensional and more efficient than the traditional linear method of media development. This means that customized medium for individual cell lines can be developed within 3-6 months rather than 1-2 years, and therefore at substantial cost savings.

Our DOE approach includes choosing different factorial matrix designs that depend on the level of knowledge already attained. For example, if 30 different medium components are to be screened for productivity effects, we would start with a grouped factorial design. Here the components would first be bundled into groups and each group would be tested at two levels in a partial factorial matrix, such as a $2^{15\text{-}10}$ matrix. Individual components from those groups showing significant effects would be tested next. As the experiments progress during optimization, the concentration of each active component is increased or decreased until the incremental benefit from the change in the component level can no longer be detected. At that point the component would be near its optimal level and would not be included in the next factorial experiment. Late in the optimization scheme, only a few remaining key components are included in full factorials, which may also include central composite design points for more detailed contour analysis.

This DOE approach was used to optimize a medium for a problematic hybridoma cell line, M2. We discovered that diluting our modified version of medium (H4281mod) by 25% into NCTC109 medium gave a significant increase in antibody production. By comparing the two formulations, we found there were 32 compounds present in H4281mod or absent altogether from it. These were "bundled" into 10 separate groups and tested using a $2^{10\text{-}5}$ factorial matrix design. The results showed that 3 of the groups significantly increased IgG production, with two significant interactions among them. In the next experiment, which tested the individual components in these groups, we identified retinol acetate, sodium bicarbonate and sodium acetate as three significant factors influencing productivity of the M2 cells. The performance of the modified serum-free M2 medium was confirmed in a perfusion stirred-tank bioreactor.

## Introduction

Medium development is a complex process, involving the adjustment of numerous interacting components to their final concentrations with a goal of strong consistent support of cell growth and productivity. The traditional approach to medium development, described in great detail by Ham and coworkers (1), is a long series of titrations of each individual medium component, while keeping all other components constant at original levels. At the end of first round of titrations, a working optimum is set for only the component that shows highest activity change. Then the next round of titrations is performed for each component, while keeping the first component at its working optimum. After selecting the working optimum of the second most influential medium component, the next round of titrations is performed, and so forth. During the process, undesired components (e.g., serum) are decreased and other potentially beneficial agents are introduced. Working optimal levels chosen early in the process often must be revised later in the process as other components are individually set.

## Our DOE Approach to Medium Optimization

At Sigma, we wanted to develop a more efficient approach to the development and optimization of medium. The approach we developed is based on design of experiment (DOE) statistical methodology (2,3).

**Two Optimization Approaches Contrasted**

| Traditional *(Linear)* | Multidimensional *(Factorial)* |
|---|---|
| • Based on classical medium development methods | • Based on DOE principles of statistical design and analysis |
| • Labor-intensive approach | • Labor-efficient approach |
| • Fails to address the presence of relying instead upon tests of all components individually | • Recognizes interactions, and defines them early in the process, setting optimum levels accordingly |
| • No defined end-point | • Clearly defined end-point |
| • Long time to market (1-2 years) with higher development costs | • Short time to market (3-6 months) with lower development costs |

*Table 1*

Table 1 contrasts the two approaches. Bottom line advantage in our DOE approach is that medium can be customized to a cell line in a shorter time, bringing the product to market sooner and at less development cost.

**Defining Some DOE terms**

• **Full factorial** ($2^k$) includes all possible combinations of $k$ components, each tested at two levels. For example, a "$2^5$ " factorial would test 5 different factors at 2 levels each in a total of 32 conditions.

• **Center point** (cp) is a single condition whereby all $k$ components in the factorial are set at their midpoints between the lower and upper levels. For example, a "$2^5$+4cp" factorial would be a full factorial of 32 conditions plus 4 replicates at the midpoints of all five factors, totaling 36 conditions.

• **Partial factorial** ($2^{k\text{-}p}$) omits a fraction of the combinations that would be present with a full $2^k$ factorial, often with little impact in final analysis. For example, a "$2^{5\text{-}1}$ " factorial would test 5 different factors at 2 levels each in a total of 16 conditions.

• **Grouped factorial** ($2^{g\text{-}p}$) tests g groups of bundled components in a partial factorial design to screen large numbers of factors in a single assay. See the example below.

• **Central composite** design points can be built into full factorials to define more completely the optimal formulation. For example, a 2-component central composite can test the combinations of 2 factors over 5 concentrations each in a total of 12 conditions.

*Table 2*

Table 2 gives a defining description of some different DOE formats we often use. The progression of our experiments depends on our starting point and specific goals of a particular project. We often screen a wide variety of possible medium components by bundling them into groups and performing a grouped factorial experiment. Individual components in the active groups can next be tested using a partial factorial design. Progress from here depends on the calculated main effects and interactions. Some factors will likely be omitted from further testing due to lack of effect. Factors showing positive actions may be tested at higher levels in the next assay, while those with negative actions may be tested at lower levels. As experiments progress, factors no longer showing any effect are dropped from further testing, having their optimal levels estimated by that experiment. Also as fewer agents are tested, we would choose to use higher resolution designs including full factorials. When only 2-4 main agents remain we may also include central composite design points to reveal response curvature and identify more precisely the optimal concentrations.

**Progression of Factorial Optimization Experiments**

Factorial matrix designs are used to progressively increase the concentrations of active components until the response starts to level off. That is, further steps up in concentration do not appear to cause further improvement in productivity.

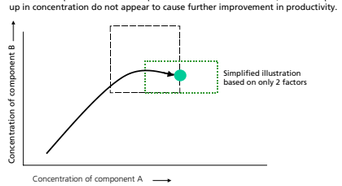


*Figure 1*

Figure 1 is an idealized and simplified depiction of the process of optimization using factorial matrices. Here, the optimal level was chosen as the midpoint of the last factorial performed (green dotted rectangle), where all four corner points gave equal performance. This is simplified because in actuality we do this type of progression for 10-15 components simultaneously.

## Grouped Factorial Design Example

The M2 hybridoma cell line produces an antibody directed against FLAG®, useful for detection and purification of FLAG fusion proteins. The goal of the medium development project was to make anti-FLAG in a bioreactor rather than mouse ascites fluid. In the original medium (RPMI-1640 with 10% FBS) M2 cells grew very slowly with poor viability and gave a low IgG final concentration in spinner flasks. Culturing M2 cells in our hybridoma serum-free medium (H4281) supplemented with 2% FBS greatly helped. However, when cholesterol was substituted for FBS, growth and productivity declined. The DOE approach was applied to optimize the serum-free medium for growth and productivity.

During optimization, we found that a mixture of 75% modified H4281 (H4281mod) and 25% NCTC109 gave a significant increase in antibody production. By comparing the two formulations, we found there were 32 compounds present in NCTC109 that were either present at lower levels in H4281mod or absent altogether from it. These compounds were "bundled" into 10 separate groups and tested using a $2^{10\text{-}5}$+6cp factorial matrix design. Figure 2 is the experimental design for this assay.

Factors in 2^(10-5) + 6 cp Factorial Matrix Experiment

| Run | A:A | B:B | C:C | D:D | E:E | F:F | G:G | H:H | J:J | K:K |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | (1) | 1 | (1) | 1 | 1 | (1) | 1 | (1) | 1 |
| 2 | (1) | (1) | 1 | 1 | (1) | 1 | (1) | (1) | 1 | 1 |
| 3 | (1) | 1 | (1) | (1) | 1 | (1) | 1 | 1 | (1) | 1 |
| 4 | (1) | 1 | (1) | 1 | 1 | 1 | 1 | (1) | 1 | (1) |
| 5 | (1) | 1 | (1) | (1) | (1) | (1) | (1) | (1) | 1 | (1) |
| 6 | (1) | (1) | 1 | (1) | 1 | (1) | 1 | (1) | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 | (1) | (1) | 1 | (1) | 1 | 1 |
| 8 | (1) | 1 | (1) | 1 | (1) | 1 | (1) | 1 | (1) | 1 |
| 9 | (1) | (1) | 1 | (1) | (1) | (1) | (1) | 1 | (1) | (1) |
| 10 | 1 | (1) | (1) | 1 | 1 | 1 | 1 | (1) | 1 | (1) |
| 11 | (1) | (1) | (1) | (1) | (1) | 1 | 1 | 1 | 1 | 1 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | (1) | (1) | (1) | 1 | 1 | 0 | (1) | 1 | 1 | 1 |
| 15 | (1) | 1 | 1 | (1) | (1) | 1 | 1 | (1) | (1) | 1 |
| 16 | 1 | 1 | (1) | 1 | 1 | (1) | 1 | (1) | (1) | (1) |
| 17 | (1) | (1) | (1) | 1 | (1) | (1) | 1 | (1) | (1) | (1) |
| 18 | 1 | (1) | 1 | 1 | 1 | (1) | (1) | (1) | 1 | (1) |
| 19 | 1 | (1) | (1) | 1 | (1) | 1 | (1) | 1 | 1 | (1) |
| 20 | (1) | (1) | (1) | (1) | 1 | 1 | (1) | (1) | (1) | (1) |
| 21 | 1 | 1 | (1) | (1) | (1) | 1 | 1 | 1 | (1) | (1) |
| 22 | 1 | (1) | 1 | (1) | (1) | 1 | 1 | (1) | 1 | 1 |
| 23 | 1 | 1 | 1 | 1 | (1) | 1 | (1) | (1) | (1) | (1) |
| 24 | (1) | 1 | 1 | (1) | 1 | 1 | (1) | 1 | 1 | (1) |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 | 1 | (1) | (1) | 1 | (1) | 1 | 1 | 1 | 1 | 1 |
| 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28 | (1) | 1 | 1 | 1 | (1) | (1) | 1 | 1 | (1) | (1) |
| 29 | (1) | 1 | 1 | 1 | (1) | (1) | (1) | (1) | (1) | 1 |
| 30 | 1 | 1 | 1 | (1) | (1) | (1) | (1) | 1 | 1 | 1 |
| 31 | 1 | (1) | (1) | (1) | (1) | (1) | (1) | 1 | (1) | 1 |
| 32 | 1 | (1) | 1 | 1 | (1) | (1) | 1 | 1 | (1) | 1 |
| 33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 34 | 1 | 1 | 1 | (1) | 1 | (1) | 1 | (1) | (1) | (1) |
| 35 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 36 | 1 | 1 | (1) | (1) | 1 | (1) | (1) | (1) | 1 | 1 |
| 37 | (1) | (1) | 1 | 1 | 1 | 1 | 1 | 1 | (1) | (1) |
| 38 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Figure 2*

Figure 2 The columns (A through K) are the groups of components and the rows (1 through 38) are the individual spinner flasks. A negative one "(1)" indicates that the particular group is present at the lower level. Likewise, a positive one "1" indicates the group is present at the higher level. For example, run #1 is a spinner that cantains A at the high level , B at the low level, etc. A zero "0" indicates the group is present at the midpoint between (1) and 1. To bolster the statistical resolution and to check for nonlinear types of interactions, six replicate centerpoints were included in this experiment. Needless to say, the procedure of setting up and performing this type of assay requires diligent planning and careful execution.

**Growth and Productivity in Grouped Factorial**

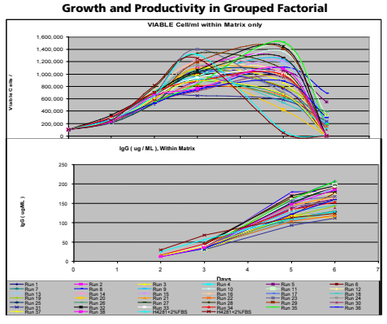


*Figure 3*

Figure 3 shows the range of growth and productivity results from individual spinner flasks prior to DOE analysis. The two spinners that dropped down to near zero viable cell density on day 5 were controls outside the matrix and contained H4281 with 2% FBS. The range of scatter in growth and final IgG concentration suggested that we would likely find some statistically significant and meaningful differences among the treatments.Figure 2

**Grouped Factorial to test NCTC109 Components**

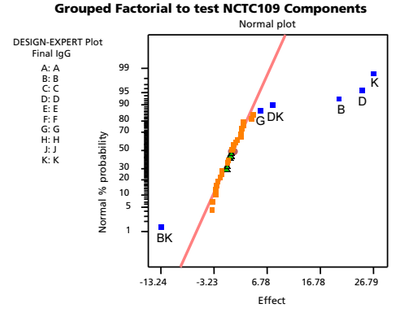


*Figure 4*

Figure 4 is a normal probability plot of the group effects for IgG production. Each square depicts the effect by one group or interaction upon the final IgG concentration. Effects are defined as the average of all conditions at the "1" level minus the average of all conditions at the "(1)" level. The triangles in the figure depict random error from replicated centerpoints. To produce a normal probability plot, all effects (x-axis) are ranked by magnitude and plotted in series using probablility coordinates (y-axis). Those effects that are not significantly different from random variance will tend to form a straight line in this plot. Those effects that are significantly greater or less than the average will appear on the plot as squares in the upper right or lower left corners, respectively . In this experiment, the main effects from groups K, D and B are clearly positive and promote IgG production. In addition, the interaction between groups D and K is positive (synergistic), while the interaction between B and K is negative (antagonistic). Analysis of variance (not shown) reveals that the main effect of group G is marginally positive (p=0.057) as well.

**Interaction between groups B and K**

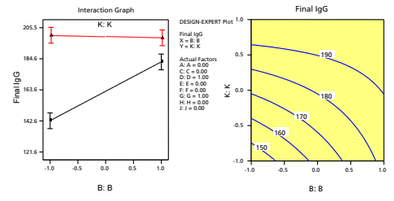


*Figure 5*

Figure 5 including an interaction graph (left) and a contour plot (right), explores the DK interaction while holding groups B and G at their higher levels. The interaction graph shows the effect of increasing group B while K is at the low level (black lower line) and while K is at the high level (red upper line). Confidence brackets (p=0.05) for these four points are also shown. The contour plot shows lines of constant IgG values, predicted by the least-squares model. One can see that adding group B alone would improve productivity, but B makes little or no difference if group K is already present at the higher level.

**Cube plot of groups B, D and K**

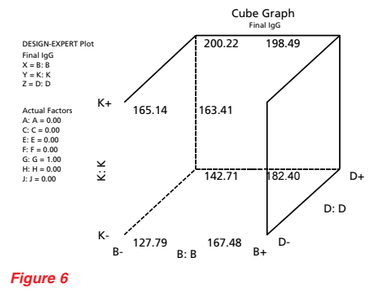


*Figure 6*

Figure 6 is a cube plot (xyz coordinates) that includes the three most important groups, B, D and K, while keeping group G at the high level. Next to the corners are the average IgG concentrations at the eight combinations of conditions. For example, at the lower left corner 127.79 ug/ml is the average IgG concentration when B, D, and K are at their lower level (124.98 ug/ml if group G is also low). The maximum concentration (200.22 ug/ml) occurs when D and K (and G) is high and B is low.

The next experiment (not shown) tested the 12 individual components from groups B, D, G and K. We identified retinol acetate, sodium bicarbonate and sodium acetate as the three most significant factors influencing productivity of the M2 cells.

**Improved Performance of M2 Hybridoma Cells**

| Growth Attribute | Culture Media in Spinner Flasks: DMEM +10%FBS | H4281 +2%FBS | H4281 M2-Optimized |
|---|---|---|---|
| Lag from frozen state | 3-4 wks | 1-2 days | n.d. |
| Viability During Log Growth | 60-70% | 80-90% | 75-85% |
| Growth rate (pop.doublings) | 0.8 pd/day | 1.6 pd/day | 1.5 pd/day |
| Maximum Viable Density | 5 x $10^5$/ml | 1.2 x $10^6$/ml | 1.0 x $10^6$/ml |
| Final Cell-Days in Spinner Flask | 0.9 x $10^6$ | 3.1 x $10^6$ | 4.0 x $10^6$ |
| **Final IgG Conc**. in Spinner Flask | **40 ug/ml** | **130 ug/ml** | **160 ug/ml** |

*Table 3*

Table 3 outlines the progress we have made in optimizing medium for the M2 cell line. The bottom line descripes a productivity increase in moving from serum-containing hybridoma medium to the optimized serum-free medium. In fact, the final IgG level in the new serum-free medium is four-times the level in original RPMI-1640 plus 10% FBS. Altogether, we made ten changes in our H4281 formulation for this particluar cell line. Our experience with other hybridoma and CHO cell lines suggest that customizing media formulations for individual cell lines will often give more than a 50% boost in productivity.

**Verifying M2 Medium in 5-L Perfusion Bioreactor**

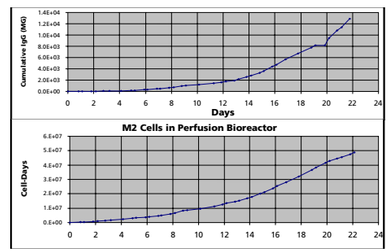


*Figure 7*

Figure 7 shows our first trial of the M2 optimized medium in a 5-liter stirred tank bioreactor set up with hollow fiber perfusion. After adjusting bioreactor conditions during the first few days the rate of antibody production was sufficient to prove feasibility. There is no apparent difference between the bioreactor-produced IgG and the original ascites-produced IgG, as determined by 19 different functional and biochemical assays (not shown). Further trials are scheduled during the development stage of this project to determine optimal reactor conditions for production in 30-liter perfusion bioreactors.

## Discussion

To our knowledge, this is the first comprehensive plan for media development that fully incorporates DOE tools. Scattered literature reporting the use of DOE procedures in medium development have mostly focused solely on one or two aspects of the process, usually screening a few factors at a time (4,5). Some have screened factors using techniques based on the Plackett-Burman design (6,7). This design fails to recognize the problem of interactions during analysis and suffers from low statistical resolution (2). Factorial designs in bioreactor trials have recently been reported for medium or process control optimization (8,9).

This DOE approach to medium optimization is more efficient than the traditional linear process. The time required to customize a culture medium for a particular cell line depends upon the defining criteria, such as the starting point of the medium and the goals needed (e.g., 50% increase in productivity). Generally, an optimized medium can be formulated in less than half the amount of time using the DOE approach compared with the linear method. This means the product gets to market faster and with lower development costs. It is this efficiency gain that makes fast customization of medium extremely attractive to boost productivity, especialy in today's worldwide shortage of biopharmaceutical production facilities.

## Materials and Methods

**Materials.** All materials were from Sigma-Aldrich Corporation (St. Louis, MO) unless otherwise stated.

**Spinner Assays.** Spinner flasks (125-ml capacity) from Bellco Glass (Vineland, NJ) or Techne Inc. (Princeton, NJ) were used. Only one manufacturer of spinner was used in any given experiment. Cells were seeded into the spinner vessels at either 5x$10^4$ or 1x$10^5$ cells/ml on day zero. Total cells/ml in each spinner was determined daily using a CASY®-1 cell counter (Scharfe Systems, Reutlingen, Germany). Percent viability was determined using trypan blue and a hemacytometer. The concentration of viable cells/ml was then calculated as percent viability times total cells/ml. Cell-days, the integral area under a cell growth curve, was approximated using a trapezoidal point-to-point summation method.

**IgG Concentrations.** The concentration of immunoglobulin G secreted into the medium was determined by HPLC using a Protein-G binding column for capture and 280nm absorbance for quantitation.

**Analysis Software.** Matrix analysis and graphical representation of results were generated with Design-Expert®, version 6.0.2 (Stat-Ease, Inc., Minneapolis, MN, USA).

## References

1. McKeehan, W., McKeehan, K, and Ham, R. (1981), The relation ship between defined low-molecular-weight substances and undefined serum-derived factors in the multiplication of untrans formed fibroblasts. The Growth Requirements of Vertebrate Cells In Vitro, Chapter 15, (C. Waymouth, R. Ham and P. Chapple, eds.), Cambridge Univ. Press (New York), pp. 223-243.

2. Montgomery, D., Design and Analysis of Experiments, Fourth Edition. John Wiley and Sons, New York (1997).

3. Cornell, J., How to Run Mixture Experiments for Product Quality. The ASQC Basic References in Qualtiy Control: Statistical Techniques (S. Shapiro and E. Mykytka, eds.) Volume 5. Quality Press, Milwaukee, WI (1990).

4. Liu, C.-H., et al., Factorial designs combined with the steepest ascent method to optimize serum-free media for CHO cells. Enzyme and Microbial Technol., 28, 314-321 (2001).

5. Ikonomou, L., et al., Design of an efficient medium for insect cell growth and recombinant protein production. In Vitro Cell. Dev. Biol.-Animal, 37, 549-559 (2001).

6. Ganne, V., and Mignot, G., Application of statistical design of experiments to the optimization of factor VIII expression by CHO cells. Cytotechnol., 6, 233-240 (1991).

7. Castro, P., et al., Application of statistical design to the optimization of culture medium for recombinant interferon-gamma production by Chinese hamster ovary cells. Appl. Microbiol. Biotechnol., 38, 84-90 (1992).

8. Moran, E., et al., A systematic approach to the validation of process control parameters for monoclonal antibody production in fed-batch culture of a murine myeloma. Biotechnol. Bioeng., 69, 242-255 (2000).

9. Delacruz, N., et al., Factorial design and analysis for rapid development of an antibody production process in CHO cells. Animal Cell Technology: From Target to Market (E. Lindner-Olsson, et al., eds.), 279-285 (2001).

099-137-499