A DESIGN OF EXPERIMENT (DOE) APPROACH TO INVESTIGATE SYNERGY EFFECTS AMONG NATURAL PRODUCTS

[1]TAEWAN KIM, [2]HYEONAE JANG, [3]SANGMUN SHIN

[1,2,3]Department of Industrial & Management Systems Engineering, Dong-A University,
37, Nakdong-daero 550beon-gil, Saha-gu, Busan 49315, Republic of Korea
E-mail: [1]twkim4231@naver.com,[2]jha3562@dau.ac.kr, [3]sshin@dau.ac.kr

Abstract - In order to find a significant synergy effects among the combinations of a number of natural products, many researchers have been investigated possible combinations of existing natural products. Because of current foods and lifestyles, an obesity issue has considered as a significant problem and research focus. However, in much literature, many existing approaches may not demonstrate systemic designed experiments and may not consider statistical correlations. The main purposes of this research is to propose a design of experiment (DoE) method and to demonstrate its associated statistical analysis approaches. First, we focused on the inhibition of adipocyte differentiation among a number of selected natural products (i.e., Astragaloside I, Ginsenoside Rd, 3,4-Dicaffeoylquinic acid, and Corilagin) based on existing literature. Second, in order to investigate synergy effects including correlations among a number of natural products, we then proposed a simplex centroid design for this experiment. Third, based on the results of the DoE method and its associated analysis, we further investigated the optimal combinations among the selected natural products. Finally, the experiment results (i.e., DoE, statistical correlation, and optimization) provided a significant possible combination of the selected natural products.

Keywords - Mixture centroid design, Desirability function, Natural product,Hedgehog signaling,Astragaloside I, Ginsenoside Rd, 3,4-Dicaffeoylquinic acid, Corilagin

I. INTRODUCTION

Since current life styles and nutrition patterns in many countries, most medical associations including United States and Europe declared the obesity as a significant disease because obese patients may suffer from a many disabilities and associated different diseases [1][2][3].Many researchers has been investigated a number of obesity problems including their functional and bio-medical causes, many treatments to the level of obesity, possible materials to increase and decrease obesity, and metabolisms of many phytochemicals. It is significantly difficult to identify the cause and relationships between the obesity including metabolisms and many input factors (i.e., foods, phytochemicals, drugs, personal life styles, etc.). In much literature, many researchers investigated a number of natural products which can possibly reduce the level of obesity. However, there are another difficulty related to a large number of possible candidates (i.e., natural products), which may provide a high dimensionality problem. In addition, each natural product provides both synergy and harmful effects because it contains many biological and chemical components.

In order to find a significant synergy effects among the combinations of a number of natural products, many researchers have been investigated possible combinations of existing natural products. However, in much literature, many existing approaches may not demonstrate systemic designed experiments and may not consider statistical correlations. The main purposes of this research is to propose a design of experiment (DOE) method and to demonstrate its associated statistical analysis approaches.

First, we focused on the inhibition of adipocyte differentiation among a number of selected natural products (i.e., Astragaloside I, Ginsenoside Rd, 3,4-Dicaffeoylquinic acid, and Corilagin) based on existing literature.Second, in order to investigate correlation synergy effects, we then proposed a simplex centroid design for this experiment. Third, based on the results of DoE method and analysis, we further investigate the optimal combinations among the selected natural products. Finally, the experiment results (i.e., DoE, statistical correlation, and optimization) can provide a significant possible combination of the selected natural products.

II. EXPERIMENT

According to JM Suh (2006), hedgehog signaling has an effect in inhibiting adipocyte differentiation and activating osteoblast [5].Therefore, if we stimulate hedgehog signaling, we maybe able to prevent or cure obesity. Generally,the more the Gli1 protein, the transcription factor of the hedgehog signaling, is created, the less the level of adipocyte differentiation is. To measure the transcription of Gli1 protein, SHH-Light2 cells were made by transfecting NIH-3T3-L1 cells with thymidine kinase promoter-driven Renilla luciferase reporter. and then SHH-Light2 cells were seeded for 24hours.The SHH-Light2 cells that were incubated were treated with a number of combinations of natural products and then seeded again for 24 hours. After we harvested the cells using lysis buffer,finally, Luciferase and Renilla were measured with the Veritas luminometer included in the dual luciferase kit.Through this process, we can investigate Gli1 expressionby measuring the amount of light emitted by Luciferase. We can also know the cell growth rate by measuring the amount of light emitted by Renilla. Based on measured values, we calcula-

tedthe cell growth rate over transcriptional activity of Gli1by calibrating the luciferase value to the Renilla value.

2.1. Mixture design

The mixture design proposed by Scheffe is use to examine interactions of input factors (i.e., natural products). Unlike other experiments,the relative proportions among the mixtures are more important than its absolute amount. The values of input factors are their proportions in the mixture, which can be between 0 and 1. And the proportions of all mixtures in each mixture experimental design point must be 100%. Therefore, the mixture design is represented by the following equation(1).

$$\sum_{i=1}^{k} x_i = 1 \quad \text{and} \quad 0 \le x_i \le 1 \qquad (1)$$

Where x_i is the proportion of component $i\,(i = 1, 2, \cdots, k)$ and k is the numbers of components. [6]We call the area of the experimental design that satisfies these constraints as a simplex of k-1 dimension.Fig.1 represents a simplex design area of a straight line when k = 2, and Fig.2 represents a triangular two-dimensional simplex design area when k = 3.If k = 4, it has a simplex design area in the form of a tetrahedron.Because of the constraint that each component is not independent, the proportions of other factors also change when the ratio of one component changes.

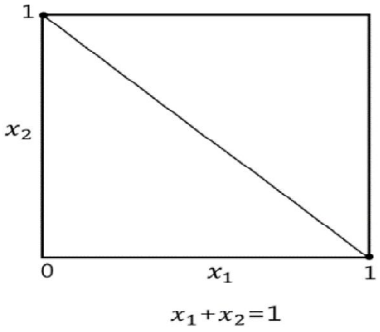

Fig.1. When k = 2, the 1-dimensional simplex design space [9].

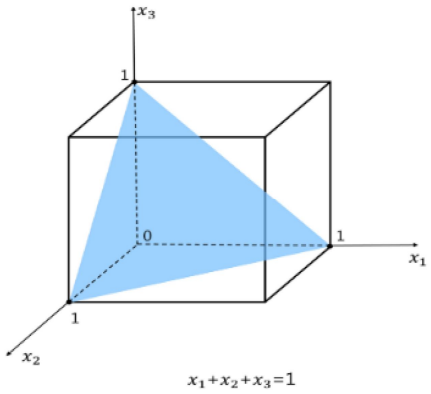

Fig.2. When k = 3, the 2-dimensional simplex design space[9].

In the simplex centroid design $2^k - 1$ observations are taken, one on each of the following: the $k$pure components, the $\binom{k}{2}$ binary mixtures with equal proportions, the$\binom{k}{3}$ ternary mixtures with equal proportions and the k-nary mixture with equal proportions. This corresponds to the points $(x_1, x_2, \ldots, x_k)$ of the equation (1) obtained by making the k permutations of (1, 0, 0, …, 0), the $\binom{k}{2}$ permutations of $\left(\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\right)$, the $\binom{k}{3}$ permutations of $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \ldots, 0)$, …, and the point $(\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k})$. The design thus consist of the centroid of the equation (1) and the centroids of all the lower dimensional simplexes it contains [10]. In this paper, we proposed simplex centroid design with added axial points when the number of factor is 4 as shown in Fig.3 below and this experimental design was provided 29 runs.

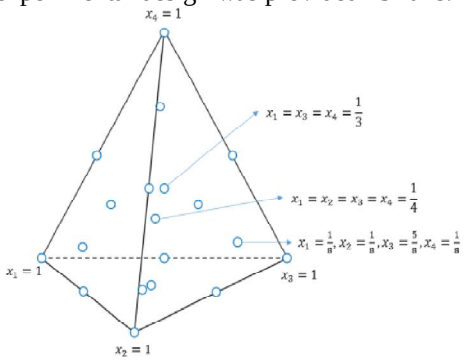

Fig.3. Simplex centroid design
(The number of factor is 4, add axial point)

2.2. Cox Response Trace Plot

In the case of mixture experiments, the main effect of each component on the response can't be calculated.In this situation, the cox response trace plot is especially useful when weunderstand the main effect of each component on the reaction. The cox response trace plot is consist of X-axis which means 'deviation from reference blend in proportion' and Y-axis which means the response we are interested in.If the slope increases when the value of X-axis increases from zero, we suppose this component increases the value of the response. On the other hand, if the slope decreases when the value of X-axis increases from zero, we suppose this component reduce the value of the response.

2.3. Analysis of variance

Analysis of variance (ANOVA) is a method of analyzing data based on F statistics.First, we calculate the sum of squares (SS) based on the experimental results, and then it disintegrated into the SS of each factor.Finally, we compare the SS of the error with the SS of each factor to identify which factors have significant effect more than the error[7].

2.4. Optimization approach

Desirability function is a kind of utility function and it optimize the value of response by transforming the estimation function derived through experiment into the desirability function. Desirability function can

perform a single response optimization or multiple response optimization [8].

The desirability function can be applied three quality characteristics. First,N-type(nominal the best characteristic) is good when the response match a specific target value.Secondly, S-type(the smaller the better characteristics) is good when the smaller the response value is the better.Finally, The L-type(the larger the better characteristics) is good when the bigger the response value is the better. In this paper, we used the L-type desirability function because of the response(Y)was defined the value that subtracting the experimental result of the experimental group from the experimental result of the control group.The desirability function of L-type is represented as Fig.4and its equation is as shown in the following equation (2).

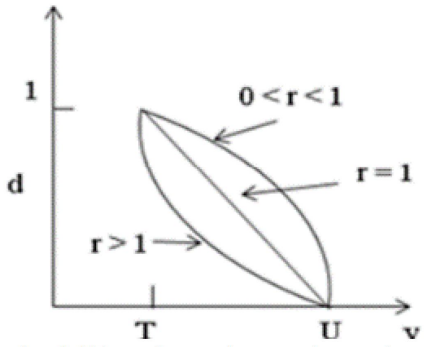

Fig.4. Desirability function when the characteristic is L-type

$$d_i = \begin{cases} 1 & y < T \\ \left(\frac{U-y}{U-T}\right)^r & T \le y \le U \\ 0 & y > U \end{cases} \qquad (2)$$

Where,'T' means the target value to which the reaction should be satisfied. 'L' is the lower limited and 'U' is the upper limited.'d' is the individual satisfaction and has a value between 0 and 1. If y is closer to the target, d will have the larger.

In the case of multiple optimization, we take the geometric mean of 'd's and judge the degree of achievement of quality characteristic by overall satisfaction.And we can assign a weight between 0.1 and 10 depending on the importance of the responses. If the weight is smaller than 1, desirability is insensitive when the response is far from the target value. Conversely, if the weight is larger than 1, the desirability is more sensitive when the reaction is far from the target value.

III. RESULTS AND DISCUSSION

3.1. Experimental design

We selected four natural products(i.e., Astragaloside I, Ginsenoside Rd, 3,4-Dicaffeoylquinic acid, and Corilagin)based on existing literature. The characteristics of Astragaloside I and Ginsenoside Rd are non-toxic and high-active, while 3,4-Dicaffeoylquinic acid and Corilagin are non-active and high-toxic.

The experiment was designedsimplex centroid design with four axial points and it provides 19 design points.We performed 6 repetition each design point.We calculated the mean of each design point and used it.

First, we normalized the response value of the control group to 1 and normalized the response value of the experimental group using the values used for standardization of the control group.Y(response) can defined that thecell growth rate over transcriptional activity of Gli1 in the control group minus the cell growth rate over transcriptional activity of Gli1 in the experimental group, for the analysis.From now on, Y1 represents the response value of the experiment when treatedToxicity 1ug/ml with the cells and Y2 represents the response value of the experiment when treatedActivity 1ug/ml with the cells.

For example, when the value of the experimental group is 1.5, we calculate the response(Y) value = (control group - experimental group) =-0.5 and wecanpredict to be 50% higher than the control group after treated the combination. Therefore, the negative value is better than zero or positive value.The level of the control factors of each experimental point and the experimental results are as shown in Table 1.

Table 1: Experiment results (4 mixtures, 19 experimental runs)

SO	EO	Input factors				Responses	
		x_1	x_2	x_3	x_4	Y_1	Y_2
		Astragaloside I	Ginsenoside Rd	3,4-Dicaffeoylquinic acid	Corilagin	Toxicity 1ug/ml)	Activity 1ug/ml)
1	5	100	0	0	0	-0.122491	-0.39364
2	17	0	100	0	0	-0.119129	-0.52532
3	8	0	0	100	0	0.080917	-1.06091
4	4	0	0	0	100	0.321803	-1.34570
5	13	50	50	0	0	-0.124574	-0.58298
6	16	50	0	50	0	0.113341	-1.35905
7	15	50	0	0	50	0.115448	-1.18742
8	7	0	50	50	0	-0.169190	-0.71801
9	10	0	50	0	50	0.274123	-1.21210
10	2	0	0	50	50	0.188035	-1.22738

11	19	33.333	33.333	33.333	0	0.153131	-1.28148
12	9	33.333	33.333	0	33.333	0.238366	-1.36864
13	18	33.333	0	33.333	33.333	0.214586	-1.29777
14	11	0	33.333	33.333	33.333	0.229629	-1.13915
15	1	25	25	25	25	0.072449	-0.84672
16	3	62.5	12.5	12.5	12.5	0.235585	-0.99093
17	6	12.5	62.5	12.5	12.5	0.040875	-0.67254
18	14	12.5	12.5	62.5	12.5	0.024672	-0.98633
19	12	12.5	12.5	12.5	62.5	0.231710	-1.15137

3.2. Analysis results

Through the Cox response trace plot, we can predicted the main effects.

In case of Toxicity 1ug/ml as shown in Fig.5, two natural products (i.e., AstragalosideI(x_1), Ginsenoside Rd(x_2) and 3,4-Dicaffeoylquinic acid(x_3)) and the other natural product (i.e., Corilagin(x_4))provideresponse (Y1) reduction effect and response(Y1) increment effect, respectively.In case of Activity 1ug/ml as shown in Fig.6, a natural product (i.e., gin(x_4)) and the other natural products (i.e., Astragaloside I(x_1) and Ginsenoside Rd(x_2)) provide response (Y1) reduction effect and response(Y1) increment effect, respectively. And the

slope of a natural product(i.e., 3,4-Dicaffeoylquinic acid(x_3)) is close to zero.

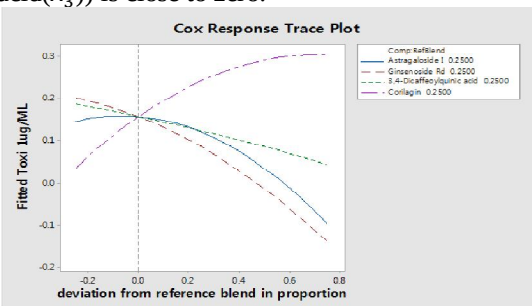

Fig.5. Cox response trace plot (Toxicity 1ug/ml)

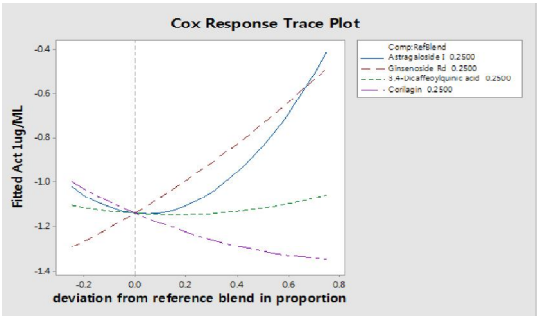

Fig.6. Cox response trace plot (Activity 1ug/ml)

The results of analysis to Toxicity 1ug / ml showed that the p-value of the model was 0.028 and the R-square was 79.42% as shown in Table.2. Normally, we consider that a term of a negative coefficient takes aresponse (Y1) reduction effect. There was no significant term in the second-order interaction, but the p-value of GinsenosideRd(x_2)*Corilagin(x_4) was 0.053, almost close to 0.05 and the coefficientwas -0.271.

Table 2: Results of analysis of variance and coefficient (Toxicity 1ug/ml)

Terms	Coefs	SE Coefs	T	P	VIF
Astragaloside I	-0.0981	0.0929	*	*	2.41
Ginsenoside Rd	-0.1401	0.0929	*	*	2.41
3,4-Dicaffeoylquinic acid	0.0437	0.0929	*	*	2.41
Corilagin	0.3022	0.0929	*	*	2.41
Astragaloside I *Ginsenoside Rd	0.364	0.408	0.89	0.395	1.86
Astragaloside I * 3,4-Dicaffeoylquinic acid	0.799	0.408	1.96	0.082	1.86
Astragaloside I *Corilagin	0.213	0.408	0.52	0.614	1.86
Ginsenoside Rd * 3,4-Dicaffeoylquinic acid	-0.271	0.408	-0.66	0.523	1.86
Ginsenoside Rd *Corilagin	0.907	0.408	2.23	0.053	1.86
3,4-Dicaffeoylquinic acid *Corilagin	0.047	0.408	0.11	0.911	1.86

Terms	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	0.322046	0.322046	0.035783	3.86	0.028
Linear	3	0.226353	0.128530	0.042843	4.62	0.032
Quadratic	6	0.095693	0.095693	0.015949	1.72	0.223
Astragaloside I*Ginsenoside Rd	1	0.004883	0.007399	0.007399	0.80	0.395
Astragaloside I*3,4-Dicaffeoylquinic acid	1	0.037865	0.035659	0.035659	3.85	0.082
Astragaloside I*Corilagin	1	0.001533	0.002531	0.002531	0.27	0.614
Ginsenoside Rd*3,4-Dicaffeoylquinic	1	0.005451	0.004098	0.004098	0.44	0.523

acid						
Ginsenoside Rd*Corilagin	1	0.045839	0.045959	0.045959	4.96	0.053
3,4-Dicaffeoylquinic acid*Corilagin	1	0.000122	0.000122	0.000122	0.01	0.911
Residual Error	9	0.083460	0.083460	0.009273		
Total	18	0.405506				
S = 0.0962981	R-square = 79.42					

Analysis of Activity 1ug / ml showed that the p-value of the model was 0.027 and the R-square was 79.65% as shown in Table.3. In case of Activity 1ug/ml, we also consider that a term of a negative coefficient takes aresponse (Y2) reduction effect.As the p-value of AstragalosideI(x_1)*3,4-Dicaffeoylquinic acid(x_3) in the second order interaction is 0.016 and the coefficient is -2.443, we consider that this interaction is reaction to decreasethe response (Y2).

Table 3: Results of analysis of variance and coefficient (Activity 1ug/ml)

Terms	Coefs	SE Coefs	T	P	VIF
Astragaloside I	-0.415	0.189	*	*	2.41
Ginsenoside Rd	-0.487	0.189	*	*	2.41
3,4-Dicaffeoylquinic acid	-1.058	0.189	*	*	2.41
Corilagin	-1.349	0.189	*	*	2.41
Astragaloside I*Ginsenoside Rd	-0.920	0.830	-1.11	0.296	1.86
Astragaloside I*3,4-Dicaffeoylquinic acid	-2.443	0.830	-2.94	0.016	1.86
Astragaloside I*Corilagin	-1.156	0.830	-1.39	0.197	1.86
Ginsenoside Rd*3,4-Dicaffeoylquinic acid	0.241	0.830	0.29	0.778	1.86
Ginsenoside Rd*Corilagin	-1.134	0.830	-1.37	0.205	1.86
3,4-Dicaffeoylquinic acid*Corilagin	0.386	0.830	0.46	0.653	1.86

Terms	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	1.35490	1.35490	0.150544	3.91	0.027
Linear	3	0.83352	0.65747	0.219157	5.70	0.018
Quadratic	6	0.52138	0.52138	0.086897	2.26	0.131
Astragaloside I*Ginsenoside Rd	1	0.02899	0.04724	0.047245	1.23	0.296
Astragaloside I*3,4-Dicaffeoylquinic acid	1	0.33574	0.33320	0.333199	8.66	0.016
Astragaloside I*Corilagin	1	0.07003	0.07454	0.074539	1.94	0.197
Ginsenoside Rd*3,4-Dicaffeoylquinic acid	1	0.00418	0.00324	0.003238	0.08	0.778
Ginsenoside Rd*Corilagin	1	0.07415	0.07180	0.071798	1.87	0.205
3,4-Dicaffeoylquinic acid*Corilagin	1	0.00830	0.00830	0.008296	0.22	0.653
Residual Error	9	0.34617	0.34617	0.038464		
Total	18	1.70107				
S = 0.196122	R-square = 79.65					

To find out the optimal combination of decreasingthe responses, we optimized Toxicity 1ug/ml and Activity 1ug/ml respectively by applying the desirability function.Both of the responses were optimized to minimize as the quality characteristics were L-type. And we assigned the target of Toxicity 1ug/ml and Activity 1ug/ml value to -2 widely. Because of minimum value of experimental responses we conducted including Toxicity 1ug/ml orActivity 1ug/ml was -1.3684.The upper limited of Toxicity 1ug/ml assigned as the 3 sigma of the control group. We also assigned the upper limited of Activity 1ug/ml as the -3 sigma of the control group.As a result optimization of Toxicity 1ug/ml, it was judged that when deRd(x_2) (83.8384%) and 3,4-Dicaffeoylquinic id(x_3) (16.1616%) were mixed, the response (Y1) was decreased by 14.71% compared with the control group as shown in Fig.7.

As a result optimization of Activity 1ug/ml, it was judged that when AstragalosideI(x_1) (36.3636%) and 3,4-Dicaffeoylquinic acid(x_3) (63.6364%) were mixed, the response (Y2) was decreased by 138.97% compared with the control group as shown in Fig.8.

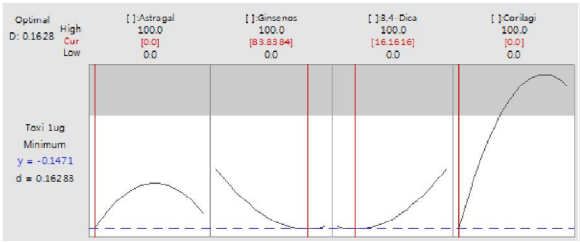

Fig.7. Optimization result based on desirability function (Toxicity 1ug/ml)

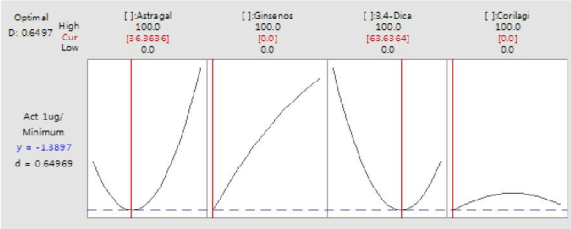

Fig.8. Optimization result based on desirability function (Activity 1ug/ml)

We also optimized Toxicity 1ug/ml and Activity 1ug/ml at the same time while assign various weights as shown in Table 4. First, when Toxicity 1ug/ml and Activity 1ug/ml were assigned the same weight, most of the mixing ratios were GinsenosideRd(x_2) and 3,4-Dicaffeoylquinic acid(x_3), and a small amount of Astragaloside I(x_1) was mixed. Secondly, when we assigned a larger weight to Toxicity 1ug/ml than Activity 1ug/ml, mixing ratio was predicted that GinsenosideRd(x_2) and 3,4-Dicaffeoylquinic acid(x_3) should be mixed. On the other hand, when we assigned a larger weight to B than to A, mixing ratio was predicted that AstragalosideI(x_1) and 3,4-Dicaffeoylquinic acid(x_3) should be mixed..When we assigned more weight to Toxicity 1ug/ml than Activity 1ug/ml, we found that both the activity and the toxic response were negative.

Table 4: Weight-Based optimization results

scenario	Input setting				Response	
	Astragaloside I	Ginsenoside Rd	3,4-Dicaffeoylquinic acid	Corilagin	Toxicity 1ug/ml	Activity 1ug/ml
Target(Minimize)	0~100				-2	-2
Toxicity 1ug/ml : Activity 1ug/ml = 1:1	2.1123	27.2727	70.615	0	-0.0476	-0.8843
Toxicity 1ug/ml : Activity 1ug/ml = 10:1	0	73.7374	26.2626	0	-0.1443	-0.5903
Toxicity 1ug/ml : Activity 1ug/ml = 5:1	0	66.6566	34.3434	0	-0.1381	-0.6288
Toxicity 1ug/ml : Activity 1ug/ml = 1:10	27.2727	0	72.7273	0	0.1636	-1.3675
Toxicity 1ug/ml : Activity 1ug/ml = 1:5	21.0997	4.1528	74.7475	0	0.127	-1.2848

CONCLUSIONS

we investigate the inhibition of adipocyte differentiation among a number of selected natural products (i.e., Astragaloside I, Ginsenoside Rd, 3,4-Dicaffeoylquinic acid, and Corilagin). The results of analysis of variance, Toxicity 1ug/ml and Activity 1ug/ml model were all significant.(α=0.05)As a result of individual optimization, it was concluded that the responsevalues of both Toxicity 1ug/ml and Activity 1ug/ml were negative. Especially, the results of optimization by changing the weights showed that when the toxicity material(i.e., 3,4-Dicaffeoylquinic acid(x_3)) is mixed with the acticity material(i.e., Astragaloside I(x_1) or Ginsenoside Rd(x_2)), the response may be reduce (i.e., the cell growth rate over transcriptional activity of Gli1 is increased.). Through this discovery, it will be useful to treat or prevent obesity that is difficult to cure the root cause

ACKNOWLEDGMENTS

This work was supported by the Bio-Synergy Research Project (NRF2013M3A9C4078156) of the Ministry of Science, ICT and Future Planning through the National Research Foundation.

REFERENCES

[1] Fontaine KR, Cheskin LJ, Barofsky I. "Health related quality of life in obese persons seeking treatment." J Fam Pract, vol. 43, pp. 265-270, 1996

[2] Barofsky I, Fontaine KR, Cheskin LJ, "Pain in the obese: impact on health-related quality of life." Ann Behav Med, vol. 19, pp. 408-410, 1998

[3] Fontaine KR, Bartlett SJ, Barofsky I. "Health related quality of life among obese persons seeking and not currently seeking treatment." Int J Eat Disord, vol. 27, pp. 101-105, 2000

[4] Scheff H,"Experiments with mixtures", J. R. Stat. Soc. Series BStat. Methodol, vol. 20, pp. 344-360, 1958

[5] JM Suh, Xiaohuan Gao, Jim McKay, Renee McKay, Zack Salo and Jonathan M. Graff, "Hedgehog signaling plays a conserved role in inhibiting fat formation",CELL METABOLISM, vol. 3, pp. 25-34, 2006

[6] MC Kim, TH Le, Cheng Bao, JT Kim, HS Chun, SM Shin and HJ Lee, "Robust optimization for the simultaneous enhancement of nitric oxide inhibition and reduction of hepatotoxicity from green teacatechins" Food SciBiotechnol, vol.26, no.6, pp. 1725-1734, 2017

[7] JC Park and SM Shin, "A Simulation-based Optimization of Design Parameters for Cooling System of Injection Mold by using ANOVA with Orthogonal Array", Journal of the Korean Society of Manufacturing Process Engineers, vol. 11, no. 5, pp. 121-128, 2012

[8] George Derringer and RonaldSuich, "Simultaneous optimization of several response variables", Journal of Quality Technology, vol. 12, no. 4, pp. 214–219, 1980

[9] Douglas C. Montgomery. Design and Anallysis of Experimnets. Jonh Wiley & Sons, Inc 6th Edition. 2005.

[10] Scheffe, "The Simplex-Centroid Design for Experiments with Mixtures", Journal of the Royal Statistical Society, vol. 25, no. 2, pp. 235-263, 1963

★★★