Available online at www.sciencedirect.com
SCIENCE DIRECT®
Food Chemistry 92 (2005) 701–706
Food Chemistry
www.elsevier.com/locate/foodchem

Application of response surface methodology for extraction optimization of germinant pumpkin seeds protein

Li Quanhong *, Fu Caili

College of Food Science and Nutrition Engineering, China Agriculture University, Beijing 100083, China

Received 19 April 2004; received in revised form 31 August 2004; accepted 31 August 2004

Abstract

Response surface methodology was employed to study the effect of liquid:solid ratio, NaCl concentration and reaction time on the production of protein from germinant pumpkin seeds. Regression analysis was performed on the data obtained. The most relevant variable was liquid:solid ratio. The coefficient determination (R^2) was good for the second-order model. A liquid:solid ratio of 30.2: 1 (v/w), a NaCl concentration of 4.26% and a reaction time of 18.1 min were found to be optimal for protein extraction from germinant pumpkin seeds. By means of additional experiments, the adequacy of this model is confirmed.

Keywords: Response surface methodology; Extraction; Optimization; Germinant; Pumpkin seeds; Protein

1. Introduction

The seeds of plants affiliated to the family of Cucurbitaceae produce a number of proteins and peptides (Wang & Ng, 2003). Pumpkin (*Cucurbita moschata*) has received considerable attention in recent years because of the nutritional and health protective value of the proteins from the seeds. Fluted pumpkin seed flours were used as protein supplements in a variety of local foods (Giami & Bekebain, 1992; Sunday & Issac, 1999). A number of different basic polypeptides, isolated from the soluble and cell wall-derived fractions of seeds of pumpkin (*Cucurbita maxima*), have antifungal activity (Vassiliou, Neumann, Condron, & Polya, 1998). Arginine–glutamate-rich protein (Ng, Parkash, & Tso, 2002) from brown pumpkin seeds was demonstrated to inhibit translation in the rabbit reticulocyte lysate system. In vitro protein digestibility of bread improved when pumpkin seed proteins were added (El-Soukkary, 2001). Preliminary investigations in our laboratories showed that germinant pumpkin seeds reduced blood glucose levels of alloxan-diabetic rats but fresh pumpkin seeds did not (Quanhong, Ze, & Tongyi, 2003; Tongyi, Quanhong, Hong, & Nan, 2003).

The protein content of the pumpkin varies from 24.5% to 36.0% (Jin Bo Dong & Huiru, 1995; Sunday & Issac, 1999; Ping, Qingyan, Ruojing, & Shimao, 2001) in different regions of the world. Arginine–glutamate-rich proteins from fresh seeds of ripe brown pumpkins was extracted in 10 mM Tris–HCl buffer, and pH 7.2 (Ng et al., 2002). Basic antifungal protein from seeds of pumpkin (*Cucurbita maxima*) was extracted with 1 M NaCl in 10 mM phosphate (Na+, pH 8.0) (Vassiliou et al., 1998), but there are few studies on extraction conditions of germinant pumpkin seed protein. The solubility of a protein, as well as its functionality as a nutritional ingredient, may be affected by various parameters, such as pH, temperature, ionic force, salt or solvent type, extraction time, solid–solvent ratio and presence of components causing linking (Mizubuti, Biondo, Souza, Silva, & Ida, 2000).

When many factors and interactions affect desired response, response surface methodology (RSM) is an

---

* Corresponding author.
E-mail address: quanhong_li@hotmail.com (L. Quanhong).

0308-8146/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2004.08.042

effective tool for optimizing the process (Triveni, Shamala, & Rastogi, 2001). As the needed information about the shape of the response surface is applied, RSM is an effective statistical method that uses a minimum of resources and quantitative data from an appropriate experimental design to determine and simultaneously solve a multivariate equation (Kalaimahan & Tapobrata, 1995). Response surface experiments attempt to identify the response that can be thought of as a surface over the explanatory variables' experimental space. It usually uses an experimental design such as central-composite experimental design (CCED) to fit an empirical, full second-order polynomial model. A central-composite experimental design, coupled with a full second-order polynomial model, is a very powerful combination that usually provides an adequate representation of most continuous response surfaces over a relatively broad factor domain (Deming, 1990).

The purpose of the present work was to optimize and to study the effect of liquid:solid ratio, NaCl concentration and reaction time for the production of protein from germinant pumpkin seeds.

2. Materials and methods

2.1. Extraction of protein

Flayed seeds of the brown pumpkin (0.1 kg) were pulverized after culturing to germinate at 35 °C for four days and NaCl solution, at the established volume and concentration, was added and then homogenized at the established time. To the supernatant obtained after centrifugation, five volumes of acetone were added and left for 8 h at 5 °C. Then the precipitation was collected and freeze-dried. The soluble protein content was then determined in duplicate.

2.2. Methods

Association of Official Analytical Chemists AOAC (1990) methods were used to determine seed moisture content, seed protein content and extracted protein content (kjeldahl method, $N \times 6.25$).

2.3. Experimental design

A central-composite experimental design, with three variables, was used to study the response pattern and to determine the optimum combination of variables. The effect of the independent variables X_1 (liquid:solid ratio, R), X_2 (NaCl concentration, C) and X_3 (reaction time, t), at three variation levels (Table 1) in the extraction process, is shown in Table 2. Three replicates (treatments 13–15) at the centre of the design were used to allow for estimation of a pure error sum of squares. Experiments were randomized in order to maximise the effects of unexplained variability in the observed responses due to extraneous factors.

Table 1
Independent variable values of the process and their corresponding levels

Independent variable	Symbol		Levels		
	Uncodified	Codified	−1	0	1
Liquid:solid ratio (v/w)	X_1	x_1	10:1	20:1	30:1
NaCl concentration	X_2	x_2	4%	6%	8%
Reaction time (min)	X_3	x_3	10	20	30

The variables were coded according to the following equation

$$x_i = (Xi - \bar{X}i)/\Delta Xi, \quad (1)$$

where x_i is the dimensionless value of an independent variable, Xi is the real value of an independent variable, $\bar{X}i$ is the real value of an independent variable at the center point, ΔXi is the step change. The specific codes are:

$$x_1 = (R - 20)/10, \quad (2)$$

and

$$x_2 = (C - 6)/2, \quad (3)$$

and

$$x_3 = [t(\text{min}) - 20]/10. \quad (4)$$

2.4. Statistical analyses

The average yield of the duplicate values obtained was taken as the dependent variable or response, Yi. The model proposed for the response is given below:

$$Yi = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3,$$

where Yi is predicted response, β_0 is offset term, β_1, β_2 and β_3 is linear effect terms, β_{11}, β_{22} and β_{33} are squared effects and β_{12}, β_{13} and β_{23} are interaction effects.

The proportion of variance explained by the polynomial models obtained is given by the multiple coefficient of determination, R^2. The significance of each coefficient was determined using the Students t-test and p value. The behaviour of the surface was investigated for the response function $(Yi) = g$ soluble protein from extract/100 g pumpkin seeds, using the regression equation (5). Further, in order to deduce workable optimum conditions, a graphical technique was used (Floros & Chinnan, 1988; Giovanni, 1983) by fixing one variable at a predetermined optimum condition. The optimum condition was verified by conducting experiments under these conditions. Responses were monitored and results compared with model predictions.

Table 2
CCED three variable with the observed responses and predicted values for protein yield

Treat	Variable levels			Experimental			Predicted (*Yi*)	*Y0* − *Yi*
	x_1	x_2	x_3	I	II	Average (*Y0*)		
1	−1	−1	0	2.92	2.84	2.88	2.77	0.11
2	−1	0	−1	3.67	3.58	3.63	3.45	0.18
3	−1	0	1	4.06	4.16	4.11	4.28	−0.17
4	−1	1	0	4.86	4.85	4.86	4.98	−0.12
5	0	−1	−1	5.88	5.81	5.85	6.14	−0.29
6	0	−1	1	6.62	5.57	6.6	6.54	0.06
7	0	1	−1	6.70	6.79	6.75	6.81	−0.06
8	0	1	1	7.77	7.89	7.83	7.54	0.29
9	1	−1	0	7.90	8.06	7.98	7.86	0.12
10	1	0	−1	7.54	7.64	7.59	7.42	0.17
11	1	0	1	7.50	7.55	7.53	7.71	−0.18
12	1	1	0	7.15	7.24	7.2	7.31	−0.11
13	0	0	0	7.10	7.11	7.11	7.12	−0.01
14	0	0	0	7.13	7.21	7.17	7.12	0.05
15	0	0	0	7.04	7.11	7.08	7.12	−0.04

$R = 0.9951$; $R^2 = 99.03\%$.

The fitted polynomial equation was expressed as surface plots in order to visualise the relationship between the response and experimental levels of each factor and to deduce the optimum conditions. The computer software used for this study was STATISTICA, version 5.5, by Statsoft Inc.

3. Results and discussion

3.1. Fitting the models

The application of RSM yields the following regression equation, which is an empirical relationship between protein yield and the test variable in coded units, as given in the following equation:

$$\begin{aligned}Y = {} & 7.1200 + 1.8525x_1 + 0.4163x_2 + 0.2813x_3 \\ & - 0.6900x_1x_2 - 0.1350x_1x_3 + 0.0825x_2x_3 \\ & - 1.2162x_1^2 - 0.1738x_2^2 - 0.1888x_3^2. \qquad (5)\end{aligned}$$

Each of the observed values, *Y0*, is compared with the predicted value, *Yi* calculated from the model, as depicted in Fig. 1. We can see that *Y0* accords with *Yi*.

The significance of each coefficient was determined using the Student *t* test and *p* value in Table 3. The corresponding variables will be more significant if the absolute *t* value becomes larger and the *p*-value becomes smaller (Nor Aishah Saidina Amin & Didi Dwi Anggoro, 2004). It can be seen that the variable with the largest effect was the linear term of liquid:solid ratio, (x_1), followed by the quadratic of liquid:solid ratio, (x_1x_1); and the interaction effect of liquid:solid ratio and NaCl

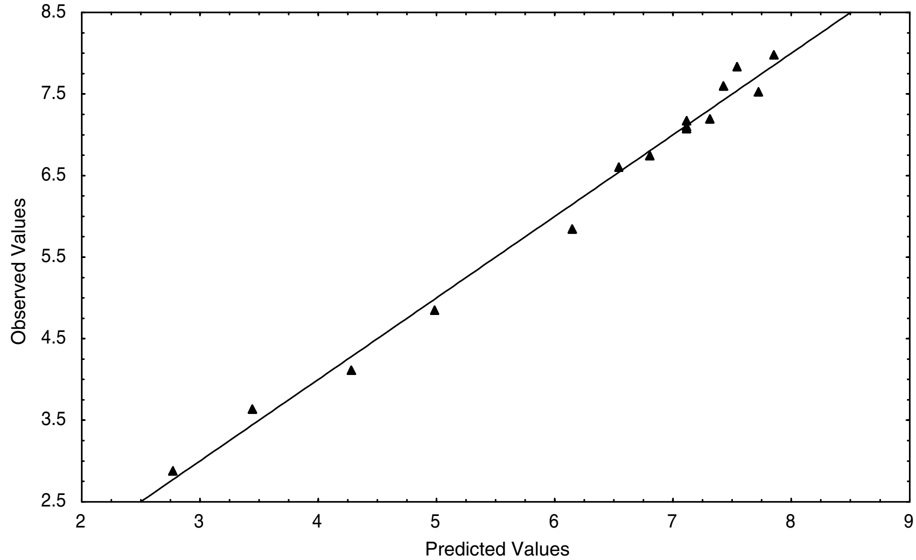

Fig. 1. Comparison between predicted and observed protein yield.

Table 3
Significance of regression coefficient for protein yield

Variables	Regression coefficient	Standard error	Computed t value	Significance level, p value
Constant	7.1200	0.1552	45.8780	0.0000001
x_1 (liquid:solid ratio)	1.8525	0.0950	19.4924	0.0000066
x_2 (NaCl concentration)	0.4163	0.0950	4.3799	0.0071551
x_3 (reaction time)	0.2813	0.0950	2.9594	0.0315394
x_1x_1	−1.2162	0.1399	−8.6942	0.0003330
x_2x_2	−0.1738	0.1399	−1.2421	0.2692931
x_3x_3	−0.1888	0.1399	−1.3493	0.2351214
x_1x_2	−0.6900	0.1344	−5.1338	0.0036644
x_1x_3	−0.1350	0.1344	−1.0045	0.3612535
x_2x_3	0.0825	0.1344	0.6138	0.5662155

concentration (x_1x_2). The factor t test value (19.4924) and p value ($p = 0.0000066$) corresponds to x_1, while the t test values for x_1x_1 and x_1x_2 are smaller at 8.6942 and 5.1338, respectively, but the p values are still significant at $p = 0.0003330$ and 0.0036644, respectively.

The multiple coefficients of correlation $R = 0.9951$ indicate a close agreement between experimental and predicted values of the protein yield. The total determination coefficient, $R^2 = 99.03\%$ implies that the sample variations of 99.03% for protein are attributable to the independent variables, namely liquid:solid ratio, NaCl concentration and reaction time.

3.2. Optimization of the process

The 3D surface plots were drawn to illustrate the main and interactive effects of the independent variables on the dependent one. These graphs were drawn by imposing a constant value (i.e., the central points of the interval taken into consideration to one independent variable).

The effects of liquid:solid ratio, NaCl concentration and reaction time on response i.e. yield are shown (Table 3) by the coefficients of second order polynomials. The response surfaces based on these coefficients are shown in Figs. 2–4 with one variable kept at optimum level and varying the other two within the experimental range. In general, exploration of the response surfaces indicated a complex interaction between the variables.

Fig. 2 shows the effect of NaCl concentration and of the liquid:solid ratio on protein production. A quadratic effect of liquid:solid ratio and a linear effect of NaCl concentration on the response were observed. Fig. 3 shows the effect of liquid:solid ratio and reaction time; liquid:solid ratio exerted a quadratic effect on protein production, whereas reaction time had a linear effect. Fig. 4 depicts the influence of NaCl concentration and reaction time; it can be seen as a linear effect for both NaCl concentration and reaction time. Therefore, increases of NaCl concentration and reaction time both resulted in a higher protein yield.

If the partial derivative of Eq. (5) is zero, three equations can be constructed as follows:

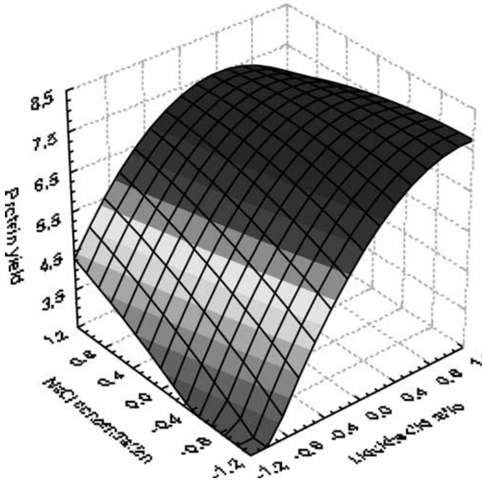

Fig. 2. 3D graphic surface optimization of protein yield versus liquid:solid ratio and NaCl concentration.

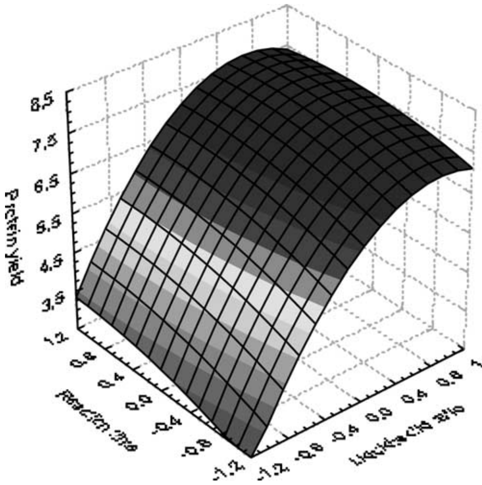

Fig. 3. 3D graphic surface optimization of protein yield versus liquid:solid ratio and reaction time.

$$1.8525 - 2.4324x_1 - 0.6900x_2 - 0.1350x_3 = 0, \tag{6}$$

$$0.4163 - 0.6900x_1 - 0.3476x_2 + 0.0825x_3 = 0, \tag{7}$$

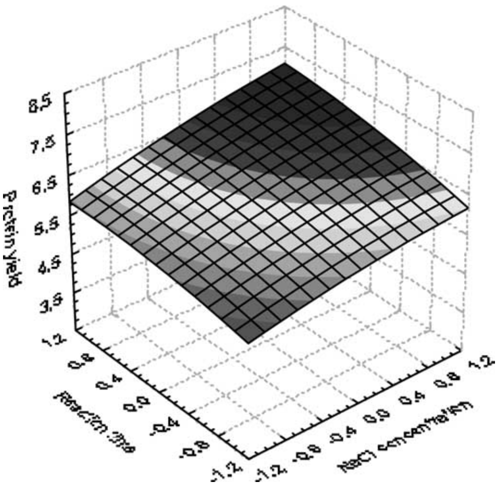

Fig. 4. 3D graphic surface optimization of protein yield versus NaCl concentration and reaction time.

Table 4
Optimum of condition (based on graphical optimisation), predicted and experimental value of response at that condition

Optimum condition	Coded levels	Actual levels	
Liquid:solid ratio	1.02	30.2: 1	
NaCl concentration	−0.87	4.26	
Responses	Predicted value	Experimental value	
		Mean[a]	Range
Yield	7.84	7.83 ± 0.11	7.77 − 7.99

[a] Mean value of five determinations.

$$0.2813 - 0.1350x_1 + 0.0825x_2 - 0.3776x_3 = 0. \tag{8}$$

Using Eqs. Eqs. (2)–(4) and Eqs. (6)–(8) the following results can be obtained:

$$x_1 = 1.02,\ x_2 = -0.87,\ x_3 = -0.19,$$

$$X_1 = 30.2 : 1,\ X_2 = 4.26\%,\ X_3 = 18.1.$$

3.3. Verification of results

The suitability of the model equation for predicting the optimum response values was tested using the recommended optimum conditions. This set of conditions was determined to be optimum by a RSM optimization approach, which was also used to experimentally validate and predict the value of the responses using model equations. The experimental values were found to be in agreement with the predicted ones (Table 4).

4. Conclusions

The production of protein from germinant pumpkin seeds was optimized using 'Statsoft Statistica' version 5.5 software. The three independent variables involved in the optimisation are liquid:solid ratio (x_1), NaCl concentration (x_2) and reaction time (x_3). The Student t test and p value indicated that the variable with the largest effect was the liquid:solid ratio (x_1). This is followed by the quadratic effect of liquid:solid ratio (x_1x_1) and interaction effect of liquid:solid ratio and NaCl concentration (x_1x_2). From the RSM results, the optimal protein yield of 7.84 g of soluble protein from extract/100 g germinant pumpkin seeds was obtained. The adequacy of this model is confirmed by means of additional experiments

Acknowledgements

We gratefully acknowledge the financial support received in the form of a research Grant (Project No: 31070664) from the National natural Science Foundation of China.

References

Amin, Nor Aishah Saidina., & Anggoro, Didi Dwi. (2004). Optimization of direct conversion of methane to liquid fuels over Cu loaded W/ZSM-5 catalyst. *Fuel*, *83*, 487–494.

Association of Official Analytical Chemists (1990). Official methods of analysis. Washington, DC: Association of Official Analytical Chemists.

Deming, S. N. (1990). Quality by design-part 5. *Chemtech, 20*, 118–120.

Dong, Jin Bo., & Huiru. (1995). Differences between various-maturity seeds of cucurbita ficifolia and their germinative characteristic in germination. *Acta Agriculture Shanghai*, *11*(2), 77–80.

El-Soukkary, F. A. H. (2001). Evaluation of pumpkin seed products for bread fortification. *Plant Foods for Human Nutrition, 56*, 365–384.

Floros, J. D., & Chinnan, M. S. (1988). Computer graphics assisted optimization for product and process development. *Food Technology, 42*, 71–78.

Giami, S. Y., & Bekebain, D. A. (1992). Proximate composition and functional properties of raw and processed full-fat fluted pumpkin (*Telfairia occidentalis*) seed flour. *Journal of the Science of Food and Agriculture, 59*, 321–325.

Giovanni, M. (1983). Response surface methodology and product optimization. *Food Technology, 37*(11), 41–45.

Kalaimahan, Theodore, & Tapobrata, Panda (1995). Application of response surface methodology to evaluate the influence of temperature and initial pH on the production of β-1,3-glucanase and carboxymethylcellulase from *Trichoderma hzrzianum*. *Enzyme and microbial technology, 11*(17), 1043–1049.

Mizubuti, Ivone Yurika., Biondo, Oswaldo., Jr., Souza, Luiz Waldemar de Oliveira., Silva, Rui SeÑ rgio dos Santos Ferreira da., & Ida, Elza Iouko. (2000). Response surface methodology for extraction optimization of pigeon pea protein. *Food Chemistry*, *70*, 259–265.

Ng, T. B., Parkash, A., & Tso, W. W. (2002). Purification and characterization of moschins, arginine–glutamate-rich proteins with translation-inhibiting activity from brown pumpkin (*Cucurbita moschata*) seeds. *Protein Expression and Purification, 26*, 9–13.

Ping, Wang, Qingyan, Zhao, Ruojing, Wang, & Shimao, Cui (2001). The study on change of main nutrition during maturation process in seed-pumpkin seed. *Acta Horticulturae Sinica, 28*(1), 47–51.

Quanhong, Li, Ze, Tian, & Tongyi, Cai (2003). Study on the hypoglycemic action of pumpkin extract in diabetic rats. *Acta Nutrimenta Sinica, 25*(1), 34–36.

Sunday, Y. Giami, & Isaac, Isichei (1999). Preparation and properties of flours and protein concentrates from raw, fermented and germinated fluted pumpkin (*Telfairia occidentalis* Hook) seeds. *Plant Foods for Human Nutrition, 54*, 67–77.

Tongyi, Cai, Quanhong, Li, Hong, Yan, & Nan, Li (2003). Study on the hypoglycemic action of pumpkin seed protein. *Journal of Chinese Institute of Food Science and Technology, 3*(1), 7–11.

Triveni, R., Shamala, T. R., & Rastogi, N. K. (2001). Optimised production and utilization of exopolysaccharide from *Agrobacterium radiobacter*. *Process Biochemistry, 36*, 787–795.

Vassiliou, A. G., Neumann, G. M., Condron, R., & Polya, Gideon M. (1998). Purification and mass spectrometry – assisted sequencing of basic antifungal proteins from seeds of pumpkin (*Cucurbita maxima*). *Plant Science, 134*, 141–162.

Wang, H. X., & Ng, T. B. (2003). Isolation of cucurmoschin, a novel antifungal peptide abundant in arginine, glutmate and glycine residues from black pumpkin seeds. *Peptides, 24*, 969–972.