Screening of extraction buffer concentration for protease extraction from the leaves of *Syzygium polyanthum* using full factorial design

IOP Conf. Series: Materials Science and Engineering **736** (2020) 022082 doi:10.1088/1757-899X/736/2/022082

Screening of extraction buffer concentration for protease extraction from the leaves of *Syzygium polyanthum* using full factorial design.

Noorhafiza Yahya[1], Zatul Iffah Mohd Arshad[1,3], Siti Kholijah Abdul Mudalip[1], Siti Zubaidah Sulaiman[1], Rohaida Che Man[1], Shalyda Md Shaarani[1] and Mohd Hairul Ab. Rahim[2]

[1]Faculty of Chemical and Natural Resources Engineering, Universiti Malaysia Pahang, Lebuhraya Tun Razak 26300, Gambang, Pahang, Malaysia
[2]Faculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak 26300, Gambang, Pahang, Malaysia
[3]Corresponding author: zatul@ump.edu.my

Abstract. This study aimed to identify the factors that significantly affect the extraction of protease from the leaves of *Syzygium polyanthum*. The extraction of protease from the leaves of *Syzigium polyanthum* was studied at different concentration of substances in the extraction buffer. A 2^4 full factorial experimental design was performed. The effect of four independent operating factors including; A, the concentration of potassium phosphate (KPO_4); B, the concentration of Triton X-100; C, the concentration of glycerol; and D, the concentration of dithiotreitol (DTT) on the protease activity were evaluated statistically. All regression models gave a good fit to the experimental data ($R^2 > 99\%$). The maximum protease activity was detected when the extraction was performed using 3.0 M KPO_4, 1% Triton X-100, 80% glycerol and 1.5 M DTT. The results revealed that the concentration of KPO_4, DTT, and their interaction were statistically significant to affect the protease activity. Thus, the results obtained from this study indicate that full factorial design is useful to screen the significant factors and minimise the number of experiments.

1. Introduction

Syzygium polyanthum is a medium-sized and evergreen herb plant known as 'salam' or 'serai kayu' among the local people in Malaysia and Indonesia, which belongs to the *Myrtaceae* family [1]. The leaves of *Syzigium polyanthum* are edible and aromatic and is commonly used as a spice and flavouring agent in Southeast Asian cuisines, including Malaysia [2]. Recent findings have indicated that the phytochemicals and secondary metabolites isolated from *S. polyanthum* were nutritious and demonstrated positive effects as anti-diarrheal [2] and anti-diabetic agent [1]. Besides, this plant is commonly used by local people as a natural meat tenderizer, and its leaves also have been extensively applied in traditional medicines by Indonesian people due to the medical importance and therapeutic activities [3]. However, no study has been conducted so far on this plant regarding protease extraction and its protease activity.

Syzygium polyanthum is rich in phenolic compounds [4], and may affect the extraction process because these compounds may bind and precipitate enzymes, thus resulting in the loss of protease activity. Several factors should be considered in choosing the best buffer for protease extraction from

the plants, for example, salt, reducing agents, buffering system and stabilizing elements [5]. It is crucial to identify the most significant factors that made up an excellent buffering system for protease extraction so that the protease activity may be retained. This is because each of the factors plays an essential role in the efficiency of the extraction. Therefore, factorial experimental design may be performed to screen the significant factors to increase the efficiency of protease extraction from the leaves of *Syzygium polyanthum*.

According to [6], factorial analysis and experimental design methods are used in many different fields of research to determine the effective parameters and accomplish performance improvement. The factorial design has been developed by mathematicians and statisticians to reduce the number of experiments. Besides, it can also demonstrate the influence of the factors affecting the results of the experiments. The factorial design approach aims to identify the significant factors by examining the changes in the response values when different levels of the factors are used [7]. Full factorial designs are employed to study the system with several factors involved as it also considers the interaction occurs between the factors examined [8].

In our previous work [9], we studied only the best extraction protocol and 1-methyl-8-oxyquinolinium betaine (QB) solvent-based extraction was identified to be the best extraction protocol to extract protease from the leaves of *Syzygium polyanthum*. However, currently, there is no study conducted on screening of extraction buffer with different concentration of substances in the (QB) solvent-based extraction of protease from the leaves of *S. polyanthum*. Therefore, this paper presents a screening of extraction buffer with different concentration of substances in the (QB) solvent-based extraction that may affect the protease activity by using full factorial experimental design. In this study, four independent substance concentrations in QB solvent method was used including; A, the concentration of potassium phosphate (KPO_4); B, the concentration of Triton X-100; C, the concentration of glycerol; and D, the concentration of dithiotreitol (DTT).

2. Materials and method

2.1. Materials

All the chemicals were bought from Sigma-Aldrich, R & M Chemicals and Acros Organics. *Syzigium polyanthum* leaves were obtained from a tree located at Kuantan, Pahang, Malaysia. The fresh leaves were cleaned, and they were cut into small pieces. A total of 150 mg of cut leaves sample was used in the extraction procedure.

2.2. Extraction of crude protease

The concentration of the substances used in the extraction buffer was very important to increase the extraction efficiency in producing protease with high protease activity. In this study, the extraction buffer was prepared with a mixture of potassium phosphate, Triton X-100, glycerol, dithiotreitol (DTT) and distilled water. All four independent factors (A, the concentration of potassium phosphate (2.0 – 3.0 M); B, the concentration of Triton X-100 (1.0 -1.5 %); C, the concentration of glycerol (80.0 – 90.0 %); and D, the concentration of DTT (1.0 – 1.5 M)) were combined randomly by the statistical software Design Expert 7.1.6 (Stat-Ease, Minneapolis, United States) based on full factorial experimental design rules shown in Table 1. The extraction method conducted in this study was following QB solvent-based extraction method [10] with a slight modification. The extraction buffer, 1 mL, pH 7.8, was used to extract 150 mg of *Syzygium polyanthum* leaves samples in 15 mL centrifuge tube. The mixture was vortexed briefly for 30 sec to ensure complete homogenization of the samples and the extraction buffer. After mixing by brief vortex, the tube that contained the lysate was centrifuged at 12000 rpm for 20 minutes at 4°C by using refrigerated centrifuge (5810 R, Eppendorf, Germany). An aliquot of supernatant (crude enzyme) was transferred to a new tube, and the protease activity was determined by protease activity assay procedure described in section 2.3.

2.3. Protease assay

Protease activity assay was carried out by using casein as the substrate [11]. An amount of 0.1 mL crude enzyme extract was mixed with 0.5 mL of 0.65% (w/v) casein in 0.05 M potassium phosphate buffer, pH 7.5, in 15 mL centrifuge tube. The mixture was swirled slowly for 10 sec and then incubated at 37°C in a water bath for 10 min. The enzymatic reaction was stopped by adding 0.5 mL of 110 mM trichloroacetic acid (TCA) into the tube, and the solution was further incubated for 30 min at 37°C in a water bath. Then, a syringe equipped with a 0.45 µL filter was used to filter the incubated solution into a 1.5 mL microcentrifuge tube. Both 0.65 mL of sodium carbonate solution and 0.15 mL of 0.5 M of Folin-Ciocalteau reagent were added into a new 15 mL centrifuge tube containing 0.25 mL of filtrate. The mixture was incubated again in a water bath at 37°C for a further 30 min. All of the samples were done in triplicates. The absorbance of the enzyme samples was measured at 660 nm using UV-Vis Spectrophotometer (Cary 50 Bio, Varian, United States). One unit of protease activity was defined as the amount of enzyme that liberated 1 µmol of tyrosine per minute under the assay conditions. A standard curve was made for a tyrosine concentration range from 0 – 0.55 µmol. The protease activity of the crude extract was calculated by using the following Equation (1), and the protease activity was defined in the unit of U/ mL.

$$PA = \frac{x \times Vt}{Ve \times Vs \times T} \qquad (1)$$

Which;
PA: protease activity (U/ mL)
x: amount of tyrosine (µmol)
Vt: volume of total solution of reaction (mL)
Ve: volume of enzymatic solution (mL)
Vs: volume of sample in the cuvette (mL)
T: time for enzymatic reaction (10 min)

2.4. Two-level (2^4) full factorial design

A 2^4 full factorial design was conducted to screen the factors that affect the efficiency of protease extraction. The following four factors, namely the concentration of potassium phosphate (KPO_4 : 2.0 – 3.0 M), the concentration of Triton X-100 (1.0 – 1.5 %), the concentration of glycerol (80.0 – 90.0 %) and the concentration of DTT (1.0 – 1.5 M) were investigated. A two-level factorial design was a statistical-based method that involves simultaneous adjustment of experimental factors by varying the factors at two levels coded as -1 (low level) and +1 (high level). The low level represents the lowest range of the factors while high level presents the highest range of the factors. Table 1 demonstrates the factors and levels employed in 2^4 full factorial design. The design suggested a set of 16 experiments to screen the significant factors affecting the protease activity (Table 2).

Table 1. Factors and levels employed in 2^4 full factorial design.

Factors	Symbols	Unit	Low-level (-1)	High-level (+1)
Concentration of KPO_4	A	M	2.0	3.0
Concentration of Triton X-100	B	%	1.0	1.5
Concentration of glycerol	C	%	80.0	90.0
Concentration of DTT	D	M	1.0	1.5

Table 2. Experimental design of 2^4 full factorial design.

Run	Factors			
	A: Concentration of KPO_4, M	B: Concentration of Triton X-100, %	C: Concentration of glycerol, %	D: Concentration of DTT, M
1	2.0	1.0	80.0	1.0
2	3.0	1.0	80.0	1.0
3	2.0	1.5	80.0	1.0
4	3.0	1.5	80.0	1.0
5	2.0	1.0	90.0	1.0
6	3.0	1.0	90.0	1.0
7	2.0	1.5	90.0	1.0
8	3.0	1.5	90.0	1.0
9	2.0	1.0	80.0	1.5
10	3.0	1.0	80.0	1.5
11	2.0	1.5	80.0	1.5
12	3.0	1.5	80.0	1.5
13	2.0	1.0	90.0	1.5
14	3.0	1.0	90.0	1.5
15	2.0	1.5	90.0	1.5
16	3.0	1.5	90.0	1.5

3.0 Result and discussion

3.1 Screening of factors affecting the activity of the protease extracted

The effect of four independent factors including the concentration of potassium phosphate, the concentration of Triton X-100, the concentration of glycerol and the concentration of DTT, on the protease activity were screened by using 2^4 full factorial design. Based on the results as tabulated in Table 3, the observed values of protease activity were ranged from 0.433 U/ mL to 0.507 U/ mL. The maximum protease activity was detected when the extraction was performed using 3.0 M KPO_4, 1% Triton X-100, 80% glycerol and 1.5 M DTT. The significant effect of the factors on the response was determined by normal probability and Pareto chart at 5% significance level.

Table 3. Experimental design and results of full factorial design

Run	Factors				Protease Activity (U/ mL)
	Concentration of KPO_4 (M)	Concentration of Triton X 100 (%)	Concentration of glycerol (%)	Concentration of DTT (M)	
1	2.0	1.0	80.	1.0	0.433
2	3.0	1.0	80.	1.0	0.447
3	2.0	1.5	80.	1.0	0.433
4	3.0	1.5	80.	1.0	0.465
5	2.0	1.0	90.	1.0	0.448
6	3.0	1.0	90.	1.0	0.448
7	2.0	1.5	90.	1.0	0.442
8	3.0	1.5	90.	1.0	0.462
9	2.0	1.0	80.	1.5	0.469
10	3.0	1.0	80.	1.5	0.507
11	2.0	1.5	80.	1.5	0.485
12	3.0	1.5	80.	1.5	0.455
13	2.0	1.0	90.	1.5	0.472
14	3.0	1.0	90.	1.5	0.463
15	2.0	1.5	90.	1.5	0.484
16	3.0	1.5	90.	1.5	0.486

3.2 Normal probability plot of standardised effects

The significance of main and interaction factors on the protease activity of the protease extracted was evaluated through a normal probability plot of the standardised effect showed in Figure 1. For the insignificant factors, their effects are normally distributed with mean zero and variance in which they are usually distributed along the linear [12]. Meanwhile, for the significant factors, their effects are distributed with non-zero means and are deviated away from the linear [13]. As the factors deviate further from the line, the more significance is the effects. The results show that the significant factors are the concentration of KPO_4 (A), concentration of DTT (D), the interaction between KPO_4 and DTT (AD) and the interaction between Triton X-100 and glycerol (BC). The other interaction effects like ABC, ABD, ACD, BCD and ABCD are negligible even though they have also deviated from the straight line [8]. Therefore, A, D, AD and BC were the only significant effects for this design. The sequence of main and interaction factors concerning increasing of effect on protease activity was identified to be BC < AD < A < D.

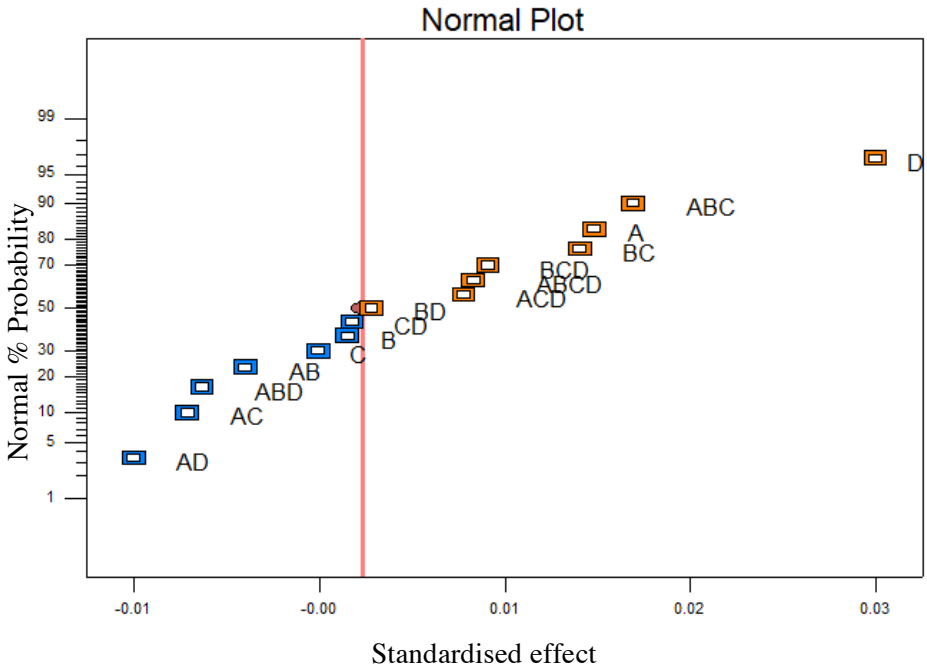

Figure 1. Normal probability plot of standardized effect for protease activity.

3.3 Pareto chart

The results demonstrated by the normal probability plot of the standardized effect for protease activity (Figure 1) were in agreement with a Pareto chart shown in Figure 2. The horizontal line in the Pareto chart indicates the minimum statistically significant effect. Meanwhile, the vertical column lengths are proportional to the degree of significance for each effect. The main and interaction factors that exceed the horizontal line is considered significant. The sequence of main and interaction factors with respect to the increasing effect on protease activity was in agreement with that of indicated by the normal probability plot of standardized effects, that is BC < AD < A < D.

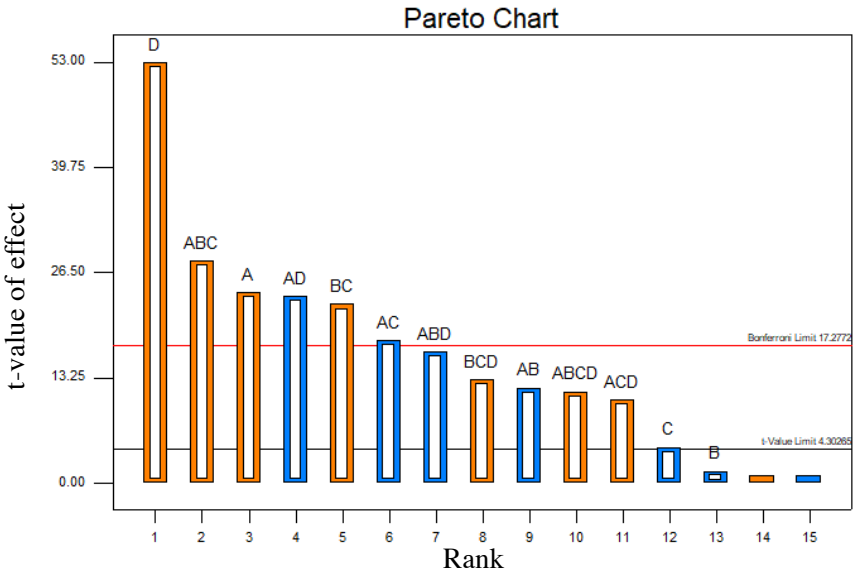

Figure 2. Pareto chart of standardised effect for protease activity.

3.4 Regression model for protease activity

Table 4 presents the regression coefficients, standard deviation ($SD_{coefficient}$) and probability values (p) values, for the main and interaction factors obtained by the statistical software. A second-order polynomial model in coded factors which correlates protease activity with each main factor and interaction factors was suggested and given by the following Equations (2) and (3):

$$PA = 0.46 + 6.000\times10^{-3}A - 3.740\times10^{-4}B - 1.125\times10^{-3}C + 0.013D - 3.000\times10^{-3}AB - 4.500\times10^{-3}AC - 5.875\times10^{-3}AD + 5.625\times10^{-3}BC + 7.000\times10^{-3}ABC - 4.125\times10^{-3}ABD + 2.625\times10^{-3}ACD + 3.250\times10^{-3}BCD + 2.875\times10^{-3}ABCD \quad (2)$$

After neglecting the insignificant factors at 5% significance level (factors with p values more than 0.05) as tabulated in Table 4, a reduced model of Equation 3 was produced:

$$PA = 0.46 + 6.000\times10^{-3}A + 0.013D - 5.875\times10^{-3}AD + 5.625\times10^{-3}BC \quad (3)$$

According to Hua *et al*. (2011), the positive sign of the coefficient signifies the synergistic effect of the factor on the response; meanwhile, the negative sign of the coefficient indicates the antagonistic effect of the factor on the response. In equation 3, the positive coefficients of [KPO_4] (A), [DTT] (D) and [Triton X-100]*[glycerol] (BC) showed that they had a positive effect on maximising the protease activity. In contrast, [KPO_4]*[DTT] (AD) interaction had a negative effect on protease activity.

Table 4. Estimated effects and coefficients of the regression models for protease activity.

Factors	Coefficient	$SD_{coefficient}$	p-value probability
Model	0.46	2.500×10^{-4}	0.0020
A	6.000×10^{-3}	2.500×10^{-4}	0.0017
B	-3.740×10^{-4}	2.500×10^{-4}	0.2724
C	-1.125×10^{-3}	2.500×10^{-4}	0.0500
D	0.013	2.500×10^{-4}	0.0004
AB	-3.000×10^{-3}	2.500×10^{-4}	0.0069
AC	-4.500×10^{-3}	2.500×10^{-4}	0.0031
AD	-5.875×10^{-3}	2.500×10^{-4}	0.0018
BC	5.625×10^{-3}	2.500×10^{-4}	0.0020
ABC	7.000×10^{-3}	2.500×10^{-4}	0.0013
ABD	-4.125×10^{-3}	2.500×10^{-4}	0.0037
ACD	2.625×10^{-3}	2.500×10^{-4}	0.0089
BCD	3.250×10^{-3}	2.500×10^{-4}	0.0059
ABCD	2.875×10^{-3}	2.500×10^{-4}	0.0075

*A: [KPO_4]; B: [Triton X-100] (%); C: [glycerol] (%); D: [DTT]; Coefficient: Regression coefficient; $SD_{coefficient}$: Standard deviation of coefficient; All variables are in coded units

3.5 Determination of model adequacy

The more reliable way to evaluate the adequacy of the model is by the application of analysis of variance (ANOVA) so that the fitted model provides an adequate approximation to the true system [14]. The factors were screened with a confidence level of 95%, and the results of ANOVA were recorded in Table 5.

Table 5. Statistical analysis for protease activity.

R^2	0.9997
Adjusted R^2	0.9977
Predicted R^2	0.9800
Adequate Precision	79.109

Based on the results tabulated in Table 5, the multiple correlation coefficient [15] or also known as the coefficient of determination [16], R^2 value obtained was 0.9997. This indicates satisfactory goodness of fit for the model because the model was able to comprehend 99.97% of the data variability. Besides, the predicted R^2 of 0.9800 is in a reasonable agreement with adjusted R^2 of 0.9977, which implies that this model is significant. The effect of the factors on the response could be explained accurately using this model. In general, adequate precision is a measure of signal to noise ratio, which allows the judgement to be made to perceive if this model is adequate to predict the response [14]. In this study, the adequate precision value for this model was 79.109.

In the enzyme extraction process, designing a suitable homogenizing medium and providing optimized conditions are highly significant, as they affect maximum protease extraction and enhanced activity. Various reducing agents, activators, and detergents are thereby used by many researchers working on plant proteases to improve the extraction process and stability of the enzyme [17], [18]. The addition of additives like potassium phosphate salt and DTT could provide excellent stability to the protease by maintaining the ionic strength of the buffer and reducing the disulphide bonds in the molecule [9].

Conclusion

The extraction of protease from the leaves of *Syzygium polyanthum* was analyzed by using 2^4 full factorial design. The effect of four potential factors like the concentration of KPO_4, the concentration of Triton X-100, the concentration of glycerol and the concentration of DTT that are affecting the protease activity were studied. The independent factors like [KPO_4], [DTT] and the interaction factors [KPO_4]*[DTT] and [Triton X-100]*[glycerol] were found to have a significant effect on protease activity. The sequence of main and interaction factors with respect to increasing of effect on protease activity was found to be [Triton X-100]*[glycerol] < [KPO_4]*[DTT] < [KPO_4] < [DTT]. High R^2 showed that the model obtained could give a reasonably good estimate of response for the system in the range studied. The finding reveals that full factorial design is applicable in studying the effect of many factors by minimising the number of experiments to run.

Acknowledgement

The authors thank the Faculty of Chemical Engineering and Natural Resources of Universiti Malaysia Pahang (UMP), for the fellowships, laboratory facilities and financial support in terms of internal grant RDU 160334, PGRS 180388.

References

[1] Widyawati T, Purnawan W W, Atangwho I J, Yusoff N A, Ahmad M and Asmawi M Z 2015 *IJPSR* **6** 1698-1704
[2] Malik A and Ahmad A R 2013 *Int. Res. J. Pharm* **4** 106-108
[3] Har L W and Ismail I S 2012 *Int. J. Med. Arom. Plants* **2** 219-228
[4] Abd Rahim E N A, Ismail A, Omar M N, Rahmat U M and Wan Ahmad W A N 2018 *Pharmacogn J*. **10** 110-119
[5] Cable J *https://bitesizebio.com/7893/ how-to-design-the-perfect-protein-purification- buffer/*, accessed on 27 April 2019
[6] Taherdoost H, Sahibuddin S and Jalaliyoon N 2014 *Advances In Applied And Pure Mathematics* **375**
[7] Japón-Luján R and Luque De Castro M D 2007 *J. Agric. Food Chem*. **55** 4584-4588.
[8] Montgomery D C, Runger G C and Hubele N F 2006 *Engineering Statistics* (New York: Wiley)
[9] Yahya N, Mohd Arshad Z I, Abdul Mudalip S K, Sulaiman S Z, Che Man R, Md Shaarani S and Ab. Rahim M H 2019 *Indones J. Chem*. **19** 503-510
[10] Ni M, Dehesh K, Tepperman J M and Quail P H 1996 *Plant Cell* **8** 1041–1059
[11] Cupp-Enyard C 2008 *J. Vis. Exp*. **19** 899
[12] Lee K, Morad N, Teng T and Poh B 2012 *APCBEE Procedia* **1** 59-65
[13] Chang S H, Teng T T and Ismail N 2011 *J. Env. Management* **92** 2580-2585
[14] Myers R H and Montgomery D C 2002 *Response Surface Methodology* (New York: Wiley)
[15] Ahmad M N, Liew S L, Yarmo M A and Said M 2012 *Biosci, Biotechnol, Biochem* **76** 1438-1444
[16] Ullah M 2012 *Coefficient of Determination: A model Selection Criteria,* retrieved from http://itfeature.com
[17] Xiao B, Heath R, Saiu P, Leiper F, Leone P and Jing C 2007 *Nature* **449** 496-500
[18] Secades P, Alvarez B and Guijarro J A 2001 *Applied and Environmental Microbiology* **67** 436-44