The Pharmaceutical and Chemical Journal, 2020, 7(3):75-81

Available online www.tpcj.org

Research Article

ISSN: 2349-7092
CODEN(USA): PCJHBA

Optimization of 3-Dichloromethyl-1,2,2-trimethylcyclopentanecarboxylic Acid Synthesis

Ye. O. Tsapko*[1], O. S. Kryskiv[1], T. O. Titko[2], O. V. Kovalska[1]

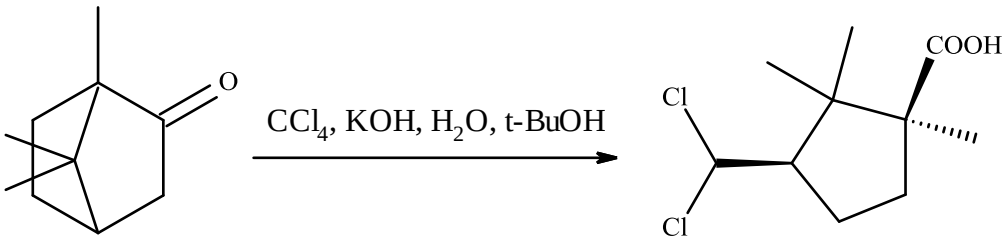
[1]Department of Inorganic Chemistry, National University of Pharmacy, Ukraine
[2]Department of Medicinal Chemistry, National University of Pharmacy, Ukraine
*Corresponding Author: E-mail: evgentsapko@ukr.net

Abstract 3-Dichloromethyl-1,2,2-trimethylcyclopentanecarboxylic acid is used in the synthesis of biologically active substances with a trimethylcyclopentane moiety. The procedure of obtaining of the aforementioned acid has been optimized using mathematical planning of the experiment. The reaction conditions and the amounts of reagents are determined. As a result, technique was simplified, and its duration was decreased.

Keywords 3-dichloromethyl-1,2,2-trimethylcyclopentanecarboxylic acid, methods of mathematical design, synthesis

Introduction

For the first time 3-dichloromethyl-1,2,2-trimethylcyclopentanecarboxylic acid **1** (Scheme) has been synthesized by Meyers [1]. This procedure consists of the chlorination of camphor in tetrachloromethane followed by the splitting by excess of potassium hydroxide in the presence of tert-butanol and water (Scheme). The method is suitable for alcohols and ketones of a certain structure and their halogen derivatives [2].



1

Scheme

We have used the acid **1** to synthesize biologically active substances with the trimethylcyclopentane moiety [3,4]. This synthetic approach allowed us to extend a number of substances derived from camphoric acid. Previous studies have shown that camphoric acid derivatives have hypoglycemic [5], anticonvulsant [6], and diuretic [7] activity.

To optimize the synthesis of the acid **1**, i. e. the amounts of reagents and reaction conditions, mathematical methods have been used. Currently, there are several approaches to the mathematical study of unknown phenomena in scientific research during technological processes for the preparation of synthetic biologically active substances [8 – 12]. The using of active experiment methods [13, 14] allows one to obtain mathematical models describing the

properties of objects of research, in which there is no need to evaluate the processes taking place inside the object. Obtaining a mathematical model is ensured by a clear implementation of the research algorithm and determination the values of the response function of the object. The process model is based upon the statistical processing of the analysis results, while "insignificant" factors are discarded. The design of the experiment makes it possible to simultaneously vary all factors and obtain quantitative estimates of both the main factors and the effects of the interaction between them, and the results obtained are characterized by a smaller error than traditional methods of univariate research [15, 16].

With a large number of independent technological factors, it was necessary to search for that area of factor space in which the yield of the target product would be maximum. To solve this problem, regression patterns obtained during the design of the experiment are used. In this regard, in our opinion, it is advisable to use the factorial experimental method in order to optimize the process. The method allows to carry out a large number of experiments, to realize possible combinations of the main levels of independent variable factors, to establish their optimal levels, and also, much faster than the empirical method, to find and justify the optimal technological parameters.

Materials and Methods

The synthesis of acid **1** was carried out according to the previously described procedure [1], using racemic camphor as the starting reagent. Mathematical calculations were performed using the STATISTICA 10 StatSoft Inc. system and Excel spreadsheet processor of MS Office 2019 Professional Plus.

Results and Discussion

The method of steepest ascent combining the design of an experiment with the gradient motion [17, 18] has been used to optimize the synthesis of acid **1**. It was experimentally established that five factors influence the yield of acid **1**: volume of carbon tetrachloride, ml; mass of potassium hydroxide, g; volume of tert-butanol, ml; volume of water, ml; temperature, °C. The mass of camphor in all experiments was 1,5 g. The plan of experiment $2^{(5-2)}$ was drawn up using STATISTICA 12 StatSoft Inc. [19 – 22] (Table 1). The table also shows the yields of acid **1**.

Table 1: Plan of the experiment

	V(CCl_4), ml	**m(KOH), g**	**V(t-BuOH), ml**	**V(H_2O), ml**	**t, °C**	**Yield, %**
1	15.0	10.0	3.0	0.4	25.0	60.8
2	15.0	10.0	6.0	0.4	5.0	46.3
3	15.0	15.0	3.0	0.0	25.0	67.1
4	15.0	15.0	6.0	0.0	5.0	55.0
5	25.0	10.0	3.0	0.0	5.0	52.3
6	25.0	10.0	6.0	0.0	25.0	55.0
7	25.0	15.0	3.0	0. 4	5.0	71.0
8	25.0	15.0	6.0	0.4	25.0	66.7
All Runs						59.3

The factor experiment has been carried to build a mathematical model. The optimization parameter (dependent factor) was the yield of the target product (Y), whereas the volume of carbon tetrachloride, ml (X_1), mass of potassium hydroxide, g (X_2), volume of tert-butanol, ml (X_3), volume of water, ml (X_4), and temperature, °C (X_5) were independent factors. To plan the experiment according this scheme, the upper and lower levels of independent parameters have been established experimentally. The center of the plan and the step of variation were determined based on the average values of the parameters $X_1 – X_5$.

The process of statistical processing of the experimental studies results was divided into two stages:

1) the study of the relationship between the yield of the target acid **1** (dependent variable Y) and independent variables (X_1–X_5) with the determination of the correspondence of the selected linear model to the experimental results;

2) search for optimal conditions of target product obtaining, namely, acid **1**.

During the experiment the parameters $X_1 – X_5$ were checked in order to determine the most significant among them. It was found out that the experimental data are normally distributed and homogeneous. To construct and analyze

empirical models in the experiment, the following criteria were used: selection and optimization of the model, as well as evaluation of the results.

The calculations were performed based on the experimental data using the STATISTICA 12 StatSoft Inc. system [19–22] (Table 2), and as a result, the regression equation (ex. 1) was obtained.

Table 2: The results of the calculation of regression coefficients*

Factor	Regression coefficients	Standard error	t-test	p-Value	Confidence limits -95, % / +95, %
Mean/Intercept	26.9625	6.8131	3.9575	0.05832	-2.3517 / 56.2767
X_1	0.3950	0.1850	2.1351	0.16629	-0.4010 / 1.1910
X_2	2.2700	0.3700	6.1351	0.02555	0.6780 / 3.8620
X_3	-2.3500	0.6167	-3.8108	0.06248	-5.0033 / 0.3033
X_4	9.6250	4.6250	2.0811	0.17290	-10.2748 / 29.5247
X_5	0.3125	0.,0925	3.3784	0.07756	-0.0855 / 0.7150

*Coefficient of multiple correlation: 0.97314; adjusted coefficient of determination: 0,906.

$$Y = 26.9625 + 0.395X_1 + 2.27X_2 – 2.35X_3 + 9.625X_4 + 0.3125X_5 \text{ (equation 1)}$$

Coefficients of interaction between factors (excess effects) were not taken into account, since they are linear combinations of other effects and cannot be estimated. The calculated significance level *p* for the free term and factors $X_1 – X_5$ does not exceed $\alpha = 0.2$, therefore, the obtained coefficients of the regression equation are significant (Table 3).

Table 3: The calculation of the significance levels of factors $X_1 – X_5$*

Factor	Effect	Standard error	t-test	p-Value	Confidence limits -95, % / +95, %	Coefficients	Standard error coefficients	Confidence limits -95, % / +95, %
Mean/ Intercept	59.2750	0.9250	64,0811	0.000243	55.2950 / 63.2550	59.2750	0.9250	55.25 / 63.2550
X_1	3.9500	1.8500	2,1351	0.1662	-4.0099 / 11.9099	1.9750	0.9250	-2.0050 /5.9550
X_2	11.3500	1.8500	6,1351	0.0256	3.3901 / 19.3099	5.6750	0.9250	1.6951 / 9.6550
X_3	-7.0500	1.8500	-3,8108	0.6248	-15.0099 / 0.9099	-3.5250	0.9250	-7.5050 / 0.4550
X4	3.8500	1.8500	2,0811	0.1729	-4.1099 / 11.8099	1.9250	0.9250	-2.0550 / 5.9050
X5	6.2500	1.8500	3,3784	0.0776	-1.7099 / 14.2099	3.1250	0.9250	-0.8550 / 7.1050

*Coefficient of multiple correlation: 0.97314; adjusted coefficient of determination: 0.906.

The resulting value of the F-test is not statistically significant. It confirms that factors $X_1 – X_5$ significantly affect the response of yield, % in an additive way (without interactions). The calculated value of the F-test is less than the tabular critical value; therefore, the differences between the means obtained under different experimental conditions can be explained by a simple additive model with two variables. Thus, the model is adequate (Table 4).

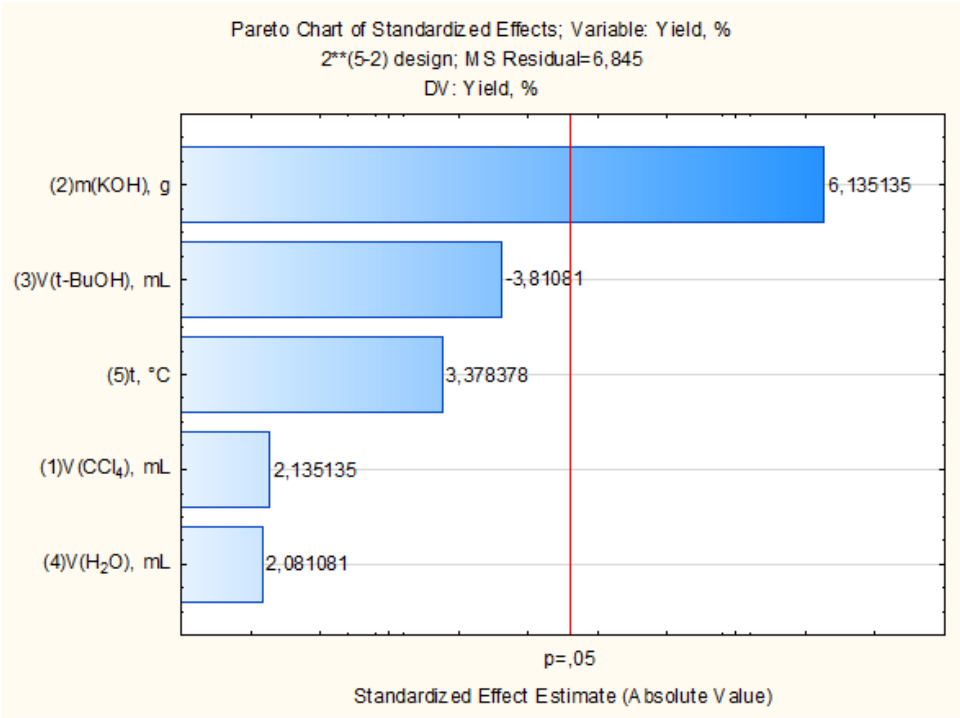
Table 4: Checking the adequacy of the resulting model*

Factor	Sum-of-squares (SS)	Degrees of freedom (df)	Mean squares (MS)	F-test	p-Value
X_1	31.2050	1	31.2050	4.5588	0.1663
X_2	257.6450	1	257.6450	37.6399	0.02555
X_3	99.4050	1	99.4050	14.5223	0.06248
X_4	29.6450	1	29.6450	4.3309	0.1729
X_5	78.1250	1	78.1250	11.4134	0.07756
Error	**13.6900**	**2**	**6.8450****		
Total SS	**509.7150**	**7**			

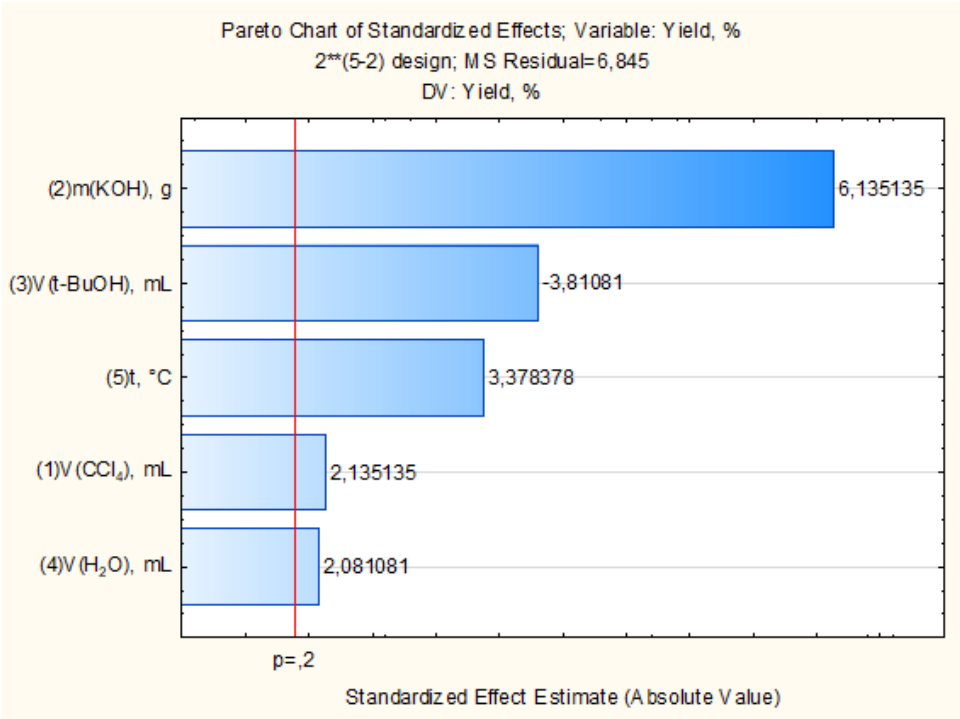
*Coefficient of multiple correlation: 0.97314; ** MS Residual.

For the confidence probability P = 0,95 only one variable is significant, namely, the mass of potassium hydroxide, mg (X_2) (Fig. 1a); for a confidence probability P = 0,8 all variables (X_1 - X_5) will be significant (Fig. 1b). If the

criteria (parameters of synthesis) is placed in the decreasing sequence of significance, the following range is obtained: $X_2 > X_3 > X_5 > X_1 > X_4$ (Fig. 1b).



a)



b)

Figure 1: Pareto charts of independent factors $X_1 – X_5$ for confidence probability P = 0.95 (a) and P = 0.8 (b).

Model rating:

in general, multiple correlation (R = 0.97314) is quite strong;

in the table of residuals, the degree of conformity of the model in this experiment is obvious (Table 5); the significance of the coefficients for the variables $X_1 – X_5$ is shown.

Table 5: Observed values, predicted values and residuals

Case or Run	Observed	Predicted	Residuals
1	55.000	53.250	1.750
2	52.300	54.050	-1.750
3	71.000	69.250	1.750
4	67.100	67.700	0.600
5	46.300	46.900	-0.600
6	55.000	54.400	0.600
7	60.800	60.200	0.600
8	66.700	68.450	-1.750

Establishment of optimal parameters for obtaining of acid 1. From the foregoing, it is clear that the optimization of synthesis of the acid **1** is the targeted search for the values of influencing factors, wherein the extremum of the optimality criterion is achieved, taking into account the restrictions imposed on all influencing factors and response functions.The solution of the tasks depends on the area of change of independent and dependent variables, on the starting point of the search, on the type of regression equation.For the planned experiments it was required to achieve the optimum yield of the target product with the following restrictions on the main technological parameters determined experimentally: 15 ml ≤ X_1 ≤ 25 ml; 10 g ≤ X_2 ≤ 15 g; 3 ml ≤ X_3 ≤ 6 ml; 0 ml ≤ X_4 ≤ 0,4 ml; 5°C ≤ X_5 ≤ 25 °C (Table 6).

Table 6: Limitations of dependent factors according to the results of calculations of the STATISTICA 12 package

Variable	Valid N	Mean	Minimum	Maximum	Standard deviation
X_1	8	20.0	15.0	25.0	5.34
X_2	8	12.5	10.0	15.0	2.67
X_3	8	4.5	3.0	6.0	1.60
X_4	8	0.2	0.0	0.4	0.21
X_5	8	15.0	5.0	25.0	10.69
Yield, %	8	59.3	46.3	71.0	8.53

The application of the simplex method to solve the obtained regression equation (MS Excel 2019, MathCad 14), given the limitations of the factors (Table 6), have not given satisfactory results.

Based upon the experimental results, it can be assumed that using of the method of steepest ascent will be effective, since the obtained linear model is adequate and is not sharply asymmetric with respect to the coefficients. Calculations on this algorithm were performed using the Excel spreadsheet processor package MS Office 2019 Professional Plus (Ver. 1905) [22] (Table 7). At step 2 the value of the yield of the target product (Y) reached its maximum (having a chemical meaning) value (Table 7), so there was no point in further calculations, the optimal conditions for the process for obtaining of acid **1** were determined. The maximum yield can be achieved under the following conditions: V (CCl_4) = 24 ml (X_1); m (KOH) = 17 g (X_2); V (-tert-BuOH) = 11 ml (X_3); V (H_2O) = 4 ml (X_4); t = 19°C (X_5). The average yield of acid **1** at the given temperature and amounts of reagents was 74.7%.

Table 7: The results of the calculations for method of steepest ascent in MS Excel 2019

	V(CCl_4), ml	m(KOH), g	V (t-BuOH), ml	V (H_2O), ml	t, °C	Yield, %
b_i	0.40	2.27	-2.35	9.63	0.31	
Δx_i	1	1	1	0.1	1	
bi × Δxi	0.40	2.27	-2.35	0.963	0.31	
γ	0.174	2.27	-1.03	0.42	0.14	
Δx_i	2.10	2.27	3.31	1.85	2.13	
Base	20.0	12.50	4.50	0.2	15	59.3
Step 1	22	15	8	0.2	17	75.9
Step 2	24	17	11	0.4	19	92.6
Step 3	26	19	14	0.6	21	109.2

We have also simplified the isolation procedure of the target product. In the original procedure [1] after the reaction the solvent is distilled off in vacuo, the residue is dissolved in water, the non-polar impurities are washed with diethyl ether, the solution is acidified, the target product is extracted with ether portions, the combined ether extracts are evaporated, and the product is crystallized from alcohol. We have excluded the solvent distillation procedure and extraction of impurities. The extraction of the target product has been replaced by filtration. Thus, we managed to avoid the using of precursor (diethyl ether), reduce the laboriousness and the time of procedure. The decrease in the duration of isolation of acid **1** also has a positive effect on the yield, because staying of acid **1** in a highly alkaline environment leads to the hydrolysis of the dichloromethyl group and, consequently, decreases the yield.

Conclusions

1. The procedure of obtaining of 3-dichloromethyl-1,2,2-trimethylcyclopentanecarboxylic acid was optimized using the mathematical design of the experiment.
2. Using the method of steepest ascent, the maximum yield 74,7% was established under the following conditions: V (CCl_4) = 24 ml; m (KOH) = 17 g; V (t-BuOH) = 11 ml; V (H_2O) = 4 ml; t = 19°C.
3. The laboriousness, cost and duration of the procedure of the synthesis were reduced.

References

[1]. Meyers, C. Y., Malte A. M. (1970) Pat. US 3896164. – Publ. 22.07.75.

[2]. Meyers, C. Y., Kolb, V. M. (1978) Facile and Selective Chlorination-Cleavage of Some Cyclanones and Cyclanols with the $CC1_4$-KOH-t-BuOH Reagent. In Situ Conversion of Estrones and Estradiols into Dichlarodoisynolic Acids. The Journal of Organic Chemistry, 43(10):1985-1990.

[3]. Tsapko, Ye. O. (2015) Search of Synthetic Approaches to 1,3,4-Oxadiazoles with 1,2,2-Trimethylcyclopentane Moiety. The Pharma Innovation Journal, 4(9):27-31.

[4]. Tsapko, Ye. O. (2015) The Alkylation Reaction of Aromatic Acids Hydrazides with (±)-cis-3-Dichloromethyl-1,2,2-trimethylcyclopentancarboxylic Acid. Журналорганічноїтафармацевтичноїхімії, 3(51):46-50.

[5]. Tsapko, Ye. O., Grytsenko, I. S., Maloshtan, L. M. et al. (2013) Дослідження гіпоглікемічної активності та гострої токсичності похідних (±)-камфорної кислоти з хіназолін-4-оновим фрагментом. Актуальні питання фармацевтичної і медичної науки та практики, 13(3):106-108.

[6]. Цапко, Є. О.,Гриценко, І. С.,Самура, Б. А.,Ніколаєв, В. О. (2008) Синтез та біологічна активність похідних іміду (±)-камфорної кислоти з фрагментами хіназолін-4(3Н)-ону та 1,2,3,4-тетрагідрохіназолін-4-ону. Лекарства человеку, современные проблемы создания и апробации лекарственных средств: мат. XXV юбилейной научно-практ. конф., Харьков, 388-393.

[7]. Цапко, Є. О., Гриценко, І. С., Красільнікова, О. А.,Бушля, Н. Е. (2007) Синтез та діуретична активність похідних похідних (±)1,2,2-триметил-3-(2-R-4-оксо-1,2,3,4-тетрагідрохіназолін-3-іл-карбамоїл)-циклопентан-1-карбонової кислоти. Вісник фармації, 3(51):12-17.

[8]. Jiang, W., Xu, C., Jiang, S. et al. (2017) Establishing a Mathematical Equations and Improving the Production of L-tert-Leucine by Uniform Design and Regression Analysis. Biotechnology and Applied Biochemistry., 181(4):1454–1464.

[9]. Andraos, J. (2012) The Algebra of Organic Synthesis: Green Metrics, Design Strategy, Route Selection, and Optimization. CRC Press., 1210.

[10]. Bornscheuer UT. (2017) The fourth wave of biocatalysis is approaching. Phil. Trans. R. Soc. A, 376: 20170063.

[11]. Tkach, V., Ivanushko, Y., Lukanova, S. et al. (2018) The Mathematical Description for CoO(OH)-Assisted Hydroxylamine Electrochemical Determination in Neutral Media. Iranian Journal of Chemistry and Chemical Engineering, 37(2):111-116.

[12]. Sultana, S., Arefin, A.,Shahruzzaman, M. (2019). Development of a mathematical model for alkylation of m-cresol with α-methylstyrene using anhydrous aluminium chloride. Bangladesh Journal of Scientific and Industrial Research, 54(3):195-202.
[13]. Pukelsheim, F. (2006) Optimal Design of Experiments. Society for Industrial and Applied Mathematics, 479.
[14]. Grachev, Yu. P., Plaksin, Y. M., (2005) Matematicheskiye metody planirovaniya eksperimenta, Moscow: De Li Print, 296.
[15]. Prochaska, Ch., Theodore, L. (2018) Introduction to Mathematical Methods for Environmental Engineers and Scientists. John Wiley & Sons, 498.
[16]. Montgomery, D. C. (2012) Statistical Quality Control. 7th Edition. Kindle Edition, 768.
[17]. Eriksson, L., Johansson, E., Kettaneh-Wold, N. et al. (2008) Design of Experiments: Principles and Application. Umetrics Academy, 425.
[18]. Yan, X., Gang, Su X. (2009) Theory and Computing: Linear Regression Analysis, 1-348,TIBCO Statistica™: Quick Reference Software Release 13.3. June 2017 (Electronic Version), 205 p.
[19]. Халафян, А. А.,Боровиков, В. П.,Калайдина, Г. В. (2017) Теория вероятностей, математическая статистика и анализ данных. Основы теории и практика на компьютере. Statistica. Excel: учебное пособие. Москва : URSS, ЛЕНАНД, 317.
[20]. Халафян, А. А. (2013) Промышленная статистика: Контроль качества, анализ процессов, планирование экспериментов в пакете STATISTICA. Издательство: Либроком, 384.
[21]. Rowe, P. (2016) Essential Statistics for the Pharmaceutical Sciences. Second Ed. John Wiley & Sons, 409.
[22]. Schmuller, J. (2016) Statistical Analysis with Excel For Dummies. 4th Ed. Publisher: For Dummies, 552.