HOA SEN UNIVERSITY
WORLD CLASS EDUCATION

Tạp chí Khoa học Trường Đại học Hoa Sen

https://www.hoasen.edu.vn/qlkh/ tapchidhhs.vn

THE DESIGN AND SCREENING OF COMPOUNDS INHIBITING THE PROLIFERATION OF HEPATIC CANCER CELLS (HEP-G2) WERE CONDUCTED UTILIZING SEVERAL TRITERPENOID DERIVATIVES DERIVED FROM THE PARAMIGNYA TRIMERA PLANT

Tran Tu Uyen*[1] Nguyen Minh Quang[2]

[1]*Faculty of Pharmacy, Ho Chi Minh City University of Technology, Ho Chi Minh*

[2]*Department of Chemical Engineering, Ho Chi Minh City Industrial University, Ho Chi Minh*

Article information

Received: 12/2023

Accepted: 3/2024

Available online: 3/2024

ABSTRACT

Currently, artificial intelligence (AI) technology is ubiquitous across various domains, serving as a highly effective tool to assist humans. The pharmaceutical sector, in general, and drug manufacturing and development, in particular, are capitalizing on this opportunity by employing *in silico* models to discover new drugs. This study is one such endeavor, utilizing an *in silico* model to predict the inhibitory potential against HepG2 cancer cells of Triterpenoid derivatives from the *Paramignya Trimera*. Quantitative Structure-Activity Relationship (QSAR) models were employed to forecast IC50 values for newly designed compounds. Following the screening for drug-like properties, molecular docking was performed on the Bcl-2 receptor of HepG2 liver cancer cells. Among 196 newly designed compounds, after successive screening steps, TPN61 emerged as a potential inhibitory compound against Hep-G2 cancer cells derived from Triterpenoids of the *Paramignya Trimera*.

Hiện nay, công nghệ trí tuệ nhân tạo (AI) được ứng dụng trên nhiều lĩnh vực khác nhau và đóng vai trò là công cụ hỗ trợ con người hiệu quả cao. Ngành dược phẩm nói chung, sản xuất và phát triển thuốc nói riêng đang tận dụng cơ hội này bằng cách sử dụng các mô hình *in silico* để khám phá các loại thuốc mới. Nghiên cứu này là một trong số đó khi sử dụng mô hình *in silico* để dự đoán khả năng ức chế tế bào ung thư Hep-G2 của các dẫn xuất Triterpenoid từ cây *Paramignya Trimera*. Mô hình mối quan hệ tác dụng-cấu trúc định lượng (QSAR) đã được sử dụng để dự báo giá trị IC50 cho các hợp chất được thiết kế mới. Sau khi sàng lọc các tính giống dược, việc docking phân tử được thực hiện trên thụ thể Bcl-2 của tế bào ung thư gan Hep-G2. Trong số 196 hợp chất mới trải qua các bước sàng lọc, TPN61 được lựa chọn là hợp chất có tiềm năng nhất ức chế tế bào ung thư Hep-G2 có nguồn gốc từ Triterpenoids của cây *Paramignya Trimera*.

Keywords: riterpenoid, QSAR, Hep-G2, Paramignya Trimera

1. INTRODUCTION

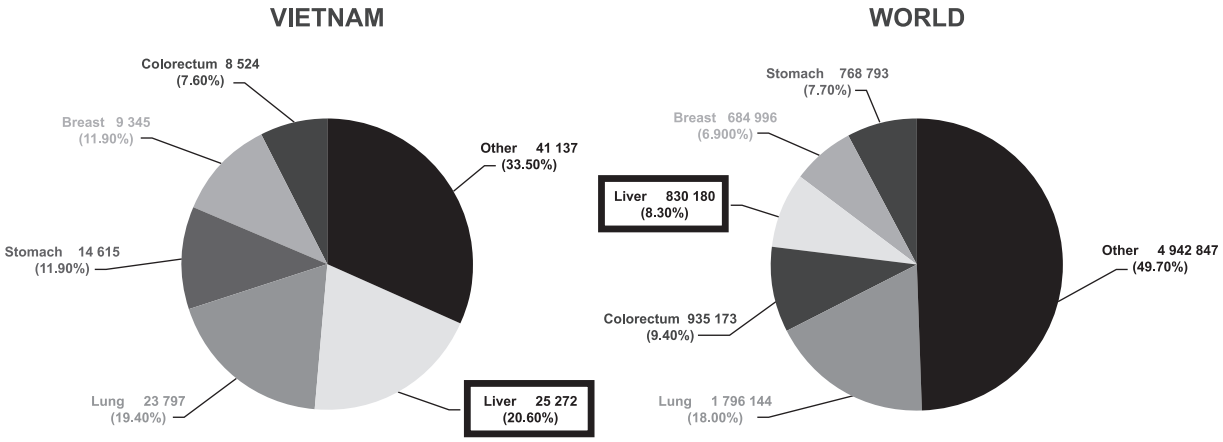
Cancer remains a leading cause of global mortality (Bray et al., 2021). According to the Global Cancer Observatory (Globocan) estimates for 2020 by the World Health Organization (WHO), there were 905 6776 new cases of liver cancer reported in 185 countries, resulting in 830 180 deaths (WHO, 2022). Liver cancer not only claims the lives of hundreds of thousands annually but also imposes a significant economic burden on society. Hence, the identification of new drug derivatives with the potential to combat liver cancer cells is crucial for both cancer patients and humanity at large. With the advancement of artificial intelligence (AI), researchers are leveraging modern technology to conduct studies, and the present research is no exception. This study employs *in silico* models to search for and screen Triterpenoid derivatives with anti-Hep-G2 cancer cell properties.

* Corresponding author: Tran Tu Uyen

Email: tuuyen101000@gmail.com

According to GLOBOCAN, liver cancer is among the top five deadliest cancers, with a high number of new cases and deaths each year in 2020. Figure 1 reveals that liver cancer has the third-highest number of deaths globally, and it's the leading cause of cancer deaths in Vietnam. In the world, liver cancer is the third most common cause of cancer deaths, accounting for 8.3%, behind only lung and colorectal cancer. Liver cancer is the primary cause of death in Vietnam, accounting for 20.6% of all deaths. The data highlights the urgent need for implementing a project to combat liver cancer-related deaths. Liver cancer not only claims hundreds of thousands of lives each year but also has a significant socio-economic impact. Thus, finding derivatives against liver cancer is crucial for the well-being of patients and humanity. (WHO, 2023).

Figure 1: Estimated number of deaths by cancer in 2020, World & Viet Nam, both sexes, all ages.



Source: GLOBOCAN

The research applies the construction of in silico models to predict the IC50 of newly designed compounds, aiming to identify compounds with superior inhibitory capabilities against Hep-G2 cells. Specifically, the project focuses on constructing three models: $QSAR_{MLR}$, $QSAR_{PCR}$, and $QSAR_{ANN}$. The use of virtual screening methods contributes to time, cost, and personnel efficiency. The study aims to deliver results that facilitate the screening of compounds in research and the expedited production of new drugs compared to experimental studies. Triterpenoids extracted from the Paramignya Trimera, particularly Escin, have been studied for their potential in inhibiting Hep-G2 cancer cells (Chudzik et al., 2015). Escin, belonging to the Oleanolic Acid group, is the specific Triterpenoid found in Paramignya Trimera. Oleanolic Acid (OA) affects cancer cells through multiple pathways, with the enhancement of the Bcl-2 inhibitory receptor being a mechanism by which OA inhibits Hep-G2 cancer cells. Normally, Bcl-2 is responsible for preventing cell cycle death (apostosis). In this study, Bcl-2 was chosen as the target to inhibit the Bcl-2 receptor, thereby increasing the apoptosis of cancer cells and reducing the growth of cancer cells.(Gu et al., 2022a) Due to the mechanism shown when acting on the Bcl-2 receptor, it affects fewer mediators to increase the apostosis process than other target receptors. As a result, the ability to make a difference is increased.

Therefore, the Bcl-2 receptor is selected as the target pathway for this study. This forms the basis for the research, aiming to discover derivatives with improved anti-cancer cell properties through the construction of in silico models and molecular docking.

Research objectives is development of the *in silico* models ($QSAR_{MLR}$, $QSAR_{PCR}$, $QSAR_{ANN}$) using derivatives obtained from experiments. Design new derivatives based on the Triterpenoid structural framework from natural sources. Predict IC50 values for newly designed derivatives and natural compounds (Escin). Assess drug-likeness based on the Veber rule. Evaluate the binding affinity of screened derivatives with the Bcl-2 receptor on Hep-G2 cells. The research collects Triterpenoid compounds from experimental results on Hep-G2 cells to build a database for constructing in silico models predicting activity for newly designed compounds. This enables the selection of compounds with the best docking results and IC_{50} values. Further details on the steps and theoretical foundations are presented in this report.

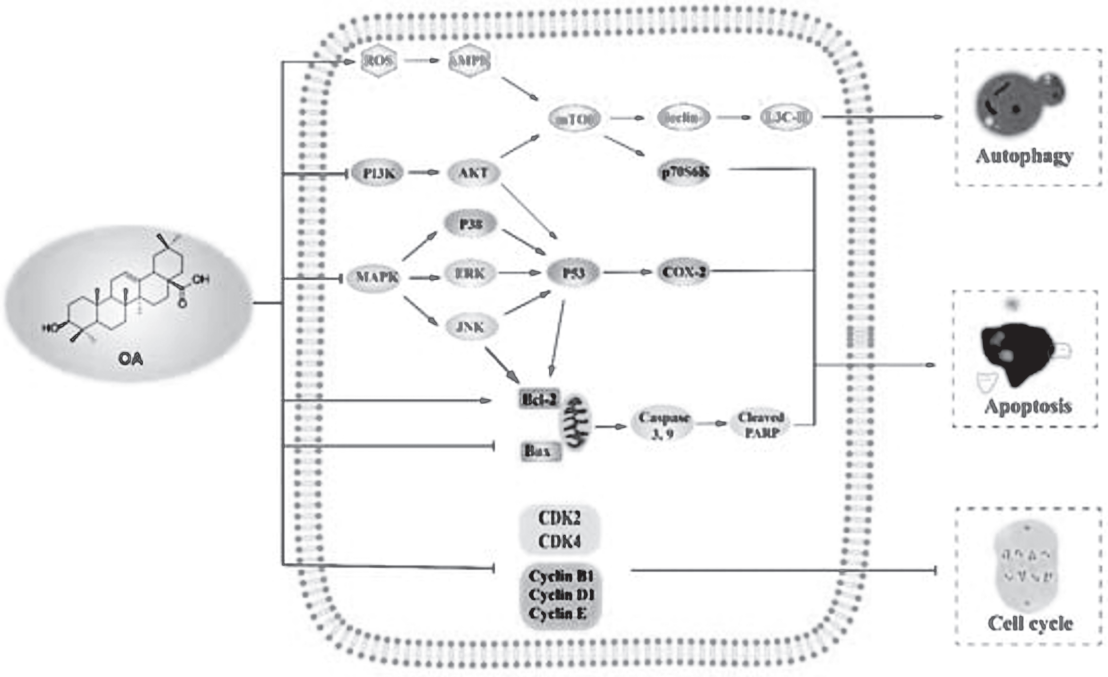


Figure 2: Three anti-tumor mechanisms of OA (Gu et al., 2022a)

Figure 3: *Paramignya Trimera* plant (Phi et al., 2020)

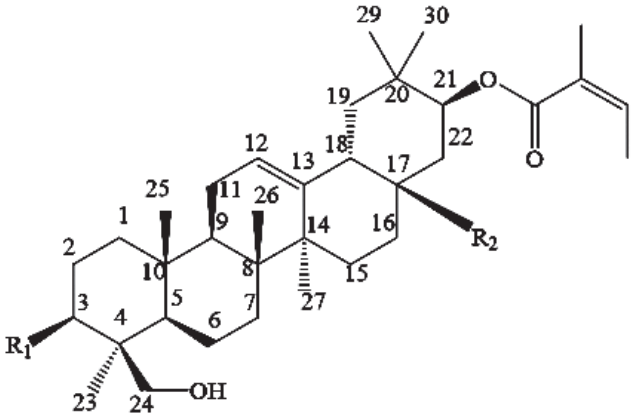


Figure 4: Structure of Escin compound (Triterpenoid from the *Paramignya Trimera*)

2. METHODS

2.1. Data

2.1.1. Data Collection from Experimental Studies

A total of 74 compounds were collected from 19 published articles in reputable journals such as Pubmed, Science Direct, etc., to construct and evaluate the external model. The criteria for inclusion were compounds with a Triterpenoid framework and tested on Hep-G2 cancer cells with reported IC50 values. From the collected data set of 74 compounds, it is divided into 2 sets based on the clustering method, k-means is the training set and the external evaluation set. A training set of 60 compounds was used to develop the QSAR model.

Table1: The table shows the journal name and publication year of the training set

No	Journal name	Publication years	Ref
TPN2,3,4,5,8	Cytotoxic Triterpenoid Saponins from Lysimachia clethroides	2011	(Liang et al., 2011)
TPN9-13	Cytotoxic triterpenoid glycosides (saikosaponins) from the roots of Bupleurum chinense	2015	(D.-Q. Li et al., 2015)
TPN14	Synthesis and cytotoxic evaluation of novel ester-triazole-linked triterpenoid–AZT conjugates	2014	(Dang Thi et al., 2014)
TPN21-22	Triterpenoid saponins from Xanthoceras sorbifolia Bunge and their inhibitory activity on human cancer cell lines	2012	(Yu et al., 2012)
TPN24	Facile synthesis of triterpenoid saponins bearing β-Glu/Gal-(1→3)-β-GluA methyl ester and their cytotoxic activities	2012	(Gao et al., 2012)
TPN25	Methanolysis of triterpenoid saponin from Ardisia gigantifolia stapf. And structure– activity relationship study against cancer cells	2013	(Mu et al., 2013)
TPN26-30	Anti-inflammatory Triterpenoid Saponins from the Stem Bark of Kalopanax pictus	2011	(Quang et al., 2011)
TPN33,34	Synthesis and Biological Activity of Triterpene–Coumarin Conjugates	2021	(Vega-Granados, et al., 2021)
TPN39	Bioactive oleanane-type saponins from the rhizomes of Anemone taipaiensis	2013	(X.-Y. Wang et al., 2013)
TPN40	New 30-Noroleanane Triterpenoid Saponins from Holboellia coriacea Diels	2016	(Ding et al., 2016)
TPN41,42	Two New Triterpenoid Saponins from the Root of Platycodon grandiflorum. Chemical and Pharmaceutical Bulletin	2013	(Ma et al., 2013)
TPN46-48	Synthesis and Cytotoxic Activity of Novel C-23-Modified Asiatic Acid Derivatives	2020	(Lu et al., 2020)
TPN89,90	Triterpenes with Cytotoxicity from the Leaves of Vernicia fordii. Chemical and Pharmaceutical Bulletin	2013	(Pei et al., 2013)
TPN107-109, 111-113, 117, 120, 121, 129	Synthesis and characterisation of celastrol derivatives as potential anticancer agents. Journal of Enzyme Inhibition and Medicinal Chemistry	2021	(Zhang et al., 2018)
TPN76,181-186, 189-194,196- 200	Synthesis of MeON-Glycoside Derivatives of Oleanolic Acid by Neoglycosylation and Evaluation of Their Cytotoxicity against Selected Cancer Cell Lines	2017	(Du et al., 2021)

An external validation set of 14 compounds was used to evaluate the model's ability to predict biological activity.

Table 2: The table shows the journal name and publication year of the training set

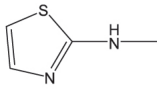
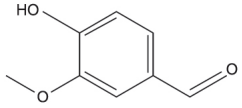
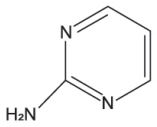
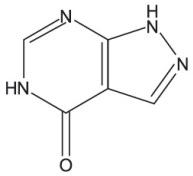
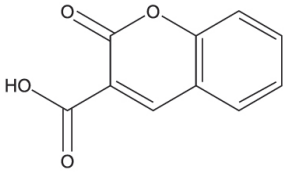
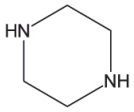
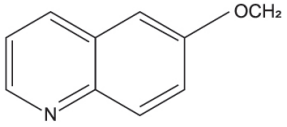
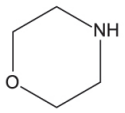
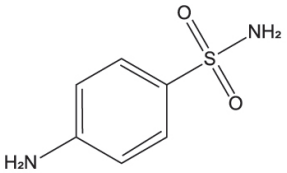
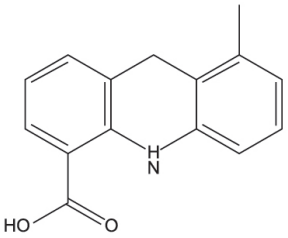
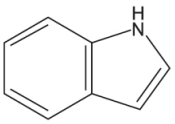
No	Journal name	Publication years	Ref
TPN_N31	Anti-inflammatory Triterpenoid Saponins from the Stem Bark of Kalopanax pictus	2011	(Quang et al., 2011)
TPN_N66	Anticancer Activity of 2α, 3α, 19β, 23β-Tetrahydroxyurs-12-en-28-oic Acid (THA), a Novel Triterpenoid Isolated from Sinojackia sarcocarpa	2011	(O. Wang et al., 2011)
TPN_N88	Synthesis,Characterization, In Vitro Anticancer Potentiality, and Antimicrobial Activities of Novel Peptide– Glycyrrhetinic-Acid-Based Derivatives	2021	(Moustafa et al., 2021)
TPN_N105	Enhanced Water Solubility and Anti-Tumor Activity of Oleanolic Acid through Chemical Structure Modification	2022	(Gu et al., 2022b)
TPN_N110, 114-116, 118, 119, 122, 123, 128	Synthesis of MeON-Glycoside Derivatives of Oleanolic Acid by Neoglycosylation and Evaluation of Their Cytotoxicity against Selected Cancer Cell Lines	2021	(Du et al., 2021)
TPN_N184	Synthesis and characterisation of celastrol derivatives as potential anticancer agents	2017	(H.-J. Zhang et al., 2017)

2.1.2. New compound design

Within the selected structural framework, two positions, R1 and R2 (Figure 2), were identified for attaching functional groups using a multi-level design approach. The design is based on the Escin framework from *Paramignya Trimera*, the R_1 and R_2 positions were selected based on reports that changing the residues at positions C_3 and C_{17} (Fig 4) would result in higher bioavailability and increased activity. (Nistor et al., 2022)

The chosen functional groups for attachment have been previously reported for their anticancer properties, comprising 14 functional groups denoted as T1 to T14.

Table 2: Name and Structure of 14 functional groups

No	Compound Name	Structure	No	Compound Name	Structure
T1	Thiazol (Jain et al., 2018; Urban et al., 2012)		T8	Vanillin (Al-Naqeb et al., 2010; Rickard et al., 2021)	
T2	Pyrazole (Bennani et al., 2022)		T9	Triazole (Alam, 2022)	
T3	Pyrimidin (Kang et al., 2012; Urban et al., 2012)		T10	Allopurinol (Y. Li et al., 2016)	
T4	Coumarin (Vega-Granados, Medina-O'Donnell, Rivas, Reyes-Zurita, Martinez, Alvarez de Cienfuegos, et al., 2021)		T11	Piperazine (Al-Ghorbani et al., 2022)	
T5	Quinoline (Borková, 2018)		T12	Morpholine (Arshad et al., 2019)	
T6	Sulfamid (casini et al., 2002)		T12	Mefenamic acid (Patel et al., 2020)	
T7	Imidazol (Sharma et al., 2021)		T8	Indole (Dadashpour & Emami, 2018)	

2.1.3. Target Interaction

The Triterpenoid in the *Paramignya Trimera*, known as Escin and belonging to the Oleanolic Acid group, exerts its effects on cancer cells through various pathways, such as promoting the apoptotic process, regulating the cell cycle, or destroying cancer cells. Typically, Bcl-2 functions to inhibit the apoptotic process in cells. In this study, the chosen target interaction is Bcl-2 with the aim of inhibiting the Bcl-2 receptor, thereby enhancing the apoptotic process of cancer cells (Gu et al., 2022). The Bcl-2 receptor was obtained from the Protein Data Bank (PDB) with the code 4D2M.

2.2. Research Methodology`

The study uses the QSAR model to search for a quantitative relationship between the change in biological activity, specifically IC50, and the change in molecular descriptors. In which $QSAR_{MLR}$ and $QSAR_{PCR}$ are multivariate linear regression and principal component analysis, respectively, helping to build 2 models with the relationship between IC50 and the most influential variables. The $QSAR_{ANN}$ model uses an artificial neural network with higher accuracy and faster thanks to a learning and analysis mechanism like the human brain. The study used three models to find the most objective results.

2.2.1. Multi-level Design

Multi-level design, also known as combinatorial synthesis, is employed to generate a list of designed compounds based on different attachment points on the main structural framework and various substituent groups. Multi-level design considers all possible scenarios when substituting each type of substituent group at each attachment position on the structural framework (Pat Whitcomb, 2022).

2.2.2. $QSAR_{MLR}$ Model

The QSARMLR model predicts the dependent variable Y based on the values of two or more independent variables X. The model is represented as:

$$Y = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \ldots + \beta_k \times_{xk} + \varepsilon \quad (1)$$

Here, Y represents the dependent variable, IC_{50}, where β_0, β_1, β_2, ..., β_k are the regression parameters of the model, x_k is the independent variable (molecular descriptors). k is the number of variables, and ε is the random error. In the context of this research, the dependent variable is the IC_{50} value, and the independent variable consists of the molecular descriptor set. The Regression 2008 software (Steppan D.D. & Yeater P.R., 1998) was utilized for the construction of the $QSAR_{MLR}$ model as reported by Roy et al. (2015).

2.2.3. $QSAR_{PCR}$ Model

The set {A, B}, where A is a data group with m observations and n variables, and B is a vector of dependent variables, comprises unprocessed data gathered for the study. The resulting B does not have a direct relationship with A but is correlated with principal components, representing features of the Principal Component Regression (PCR) model (Roy et al., 2015). The $QSAR_{PCR}$ model was constructed using XLSTAT.

2.2.4. $QSAR_{ANN}$ Model

Artificial Neural Network (ANN) mimics human brain learning processes (Roy et al., 2015). The structure of the $QSAR_{ANN}$ model I(m)-HL(n)-O(k) includes input layer neurons representing the descriptors of the $QSAR_{MLR}$ model, output layer neurons representing IC_{50} values, and hidden layer neurons optimized for the best $QSAR_{ANN}$ model (Rojas R, 1996). The $QSAR_{ANN}$ model was developed using Matlab 2016 software (Matlab, 2016).

2.2.5. Drug-likeness Assessment

Newly designed compounds were screened for drug-likeness to assess their potential to become drugs based on the Veber rule, which requires a number of rotatable bonds ≤ 10 and a polar surface area ≤ 140 $Å^2$ (Veber et al., 2002).

2.2.6. Molecular Docking

The main objective of molecular docking is to understand and predict the molecule's recognition capabilities in terms of structure - finding binding sites, and energy - predicting affinity. The diverse applications of molecular docking include structure-activity relationship studies, optimization, and virtual screening for potential molecules (Stanzione et al., 2021). In this study, MOE 2015 software was used for the molecular docking process. MOE 2015 software was chosen in the study as a tool to perform the molecular docking process due to its utility. MOE 2015 supports users to calculate molecular descriptions in 2D and 3D quickly and has a user-friendly interface. Most importantly, the docking process is simulated by MOE 2015 , which has fast, clear and detailed results.

3. RESULTS AND DISCUSSION

3.1. Database

Comprising 74 compounds collected from experimental studies with reported IC_{50} values to construct the QSAR model, divided into training (60 compounds) and external evaluation (14 compounds) sets. Combined with 196 newly designed compounds using the multi-level molecular design method, denoted as TPN1 to TPN196. The foundational dataset was optimized for structure and molecular descriptor computation, preparing for model construction and external model evaluation.

3.2. $QSAR_{MLR}$ Model

The QSAR model was constructed using the multivariate linear regression method, with the equation:

$$IC_{50} = 3.739 + 2.247\times LUMO + 85.94\times PEOE\_RPC- + 24.36\times C21 + 0.156\times vsurf\_DD13 + 5.240\times vsurf\_CW4 - 0.03695\times SlogP\_VSA3 + 0.933\times vsurf\_EWmin1 - 0.131\times SlogP\_VSA4 \quad (2)$$

The model comprises eight variables pertaining to the statistical values of $R^2 = 0.849$, $R^2_{adj} = 0.826$, and $Q^2_{LOO} = 0.789$. The model is deemed satisfactory when the conditions R2 > 0.6, and both R2adj and Q2LOO > 0.5.(Golbraikh & Tropsha, 2002)

3.3. $QSAR_{PCR}$ Model

The QSARPCR model is constructed based on variables derived from the QSARMLR model. The outcomes yielded by the QSARPCR model are as follows: R2 = 0.86, R2adj = 0.831, and Q2LOO = 0.805. The model is considered satisfactory when R2 > 0.6, and both R2adj and Q2LOO > 0.5. (Golbraikh & Tropsha, 2002)

$$IC_{50} = 8.406 + 2.480\times LUMO + 73.319\times PEOE\_RPC- + 25.836\times C21 + 0.135\times vsurf\_DD13 + 3.965\times vsurf\_CW4 - 0.037\times SlopP\_VSA3 + 1.250\times vsurf\_Ewmin1 - 0.130\times SlogP\_VSA4 \quad (3)$$

3.4. $QSAR_{ANN}$ Model and Drug-Likeness

The $QSAR_{ANN}$ model is constructed by selecting variables from the $QSAR_{MLR}$ model as input. The identified network architecture is I(8)-H(6)-O(1), employing the Purelin transfer function, yielding training set R2 = 0.941, test set R2 = 0.915, and evaluation set R2 = 0.912. This indicates that the $QSAR_{ANN}$ model demonstrates superior predictive accuracy compared to both the $QSAR_{MLR}$ and $QSAR_{PCR}$ models. Out of 196 newly designed compounds, 138 compounds meet the Veber rule criteria, signifying their potential as drug candidates.

3.5. Prediction of IC_{50} for Newly Designed Compounds

The study focuses on predicting the biological activity IC50 for the 138 compounds designed to exhibit drug-likeness, including the natural compound Escin. Subsequently, 20 compounds with IC50 values lower than the parent compound are identified, ranked in ascending order based on the QSARANN model's superior predictive capability.

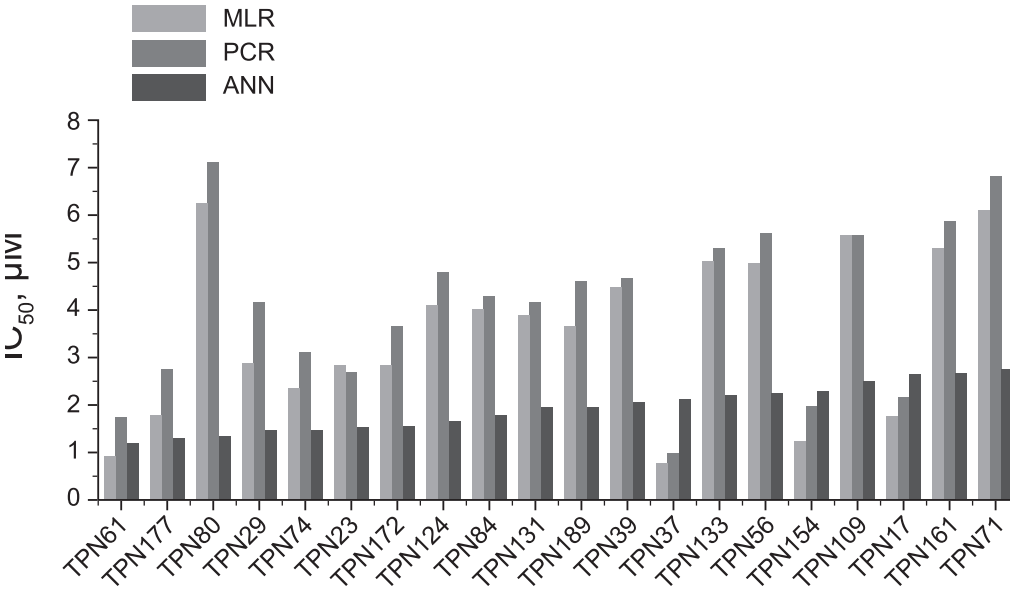


Figure 5: Bar chart illustrating the comparative prediction values of IC50 for the top 20 compounds across three models.

The detailed prediction results from each model for the top 20 compounds are presented in Table 4. Based on the prediction outcomes, the selection of compounds with biological potential is facilitated for subsequent docking studies.

Table 4: Predicted IC50 results from three models for the top 20 Compounds.

No	IC_{50} predicted (µM)			No	IC_{50} predicted (µM)		
	$QSAR_{MLR}$	$QSAR_{PCR}$	$QSAR_{ANN}$		$QSAR_{MLR}$	$QSAR_{PCR}$	$QSAR_{ANN}$
TPN61	0.938	1.764	1.187	TPN189	3.685	4.609	1.957
TPN177	1.818	2.746	1.306	TPN59	4.480	4.687	2.056
TPN23	2.848	2.700	1.532	TPN154	1.229	2.004	2.279
TPN80	6.277	7.156	1.356	TPN37	0.807	0.973	2.094
TPN29	2.905	4.159	1.477	TPN133	5.042	5.325	2.201
TPN74	2.369	3.110	1.496	TPN56	5.013	5.616	2.254
TPN172	2.863	3.642	1.573	TPN109	5.571	5.589	2.502
TPN124	4.110	4.786	1.672	TPN17	1.754	2.150	2.675
TPN84	4.025	4.277	1.817	TPN161	5.299	5.871	2.686
TPN131	3.906	4.175	1.955	TPN71	6.115	6.803	2.719
				Escin	3.391	3.534	2.752

The study found that, under the same calculation conditions, several models predict better IC_{50} values than Escin for over 20 substances. These models were used to select compounds with better biological activity, and a new potential substance was discovered to have superior cancer cell inhibitory ability compared to natural active substances.

The predictive ability of three QSAR models ($QSAR_{MLR}$, $QSAR_{PCR}$, and $QSAR_{ANN}$) was tested for both new molecules and Escin. The analysis of variance for the results showed no significant difference ($F=0.71595 < F_{0.05} = 3.07606$). Therefore, the predictive ability of the three models is consistent with each other and reliable.

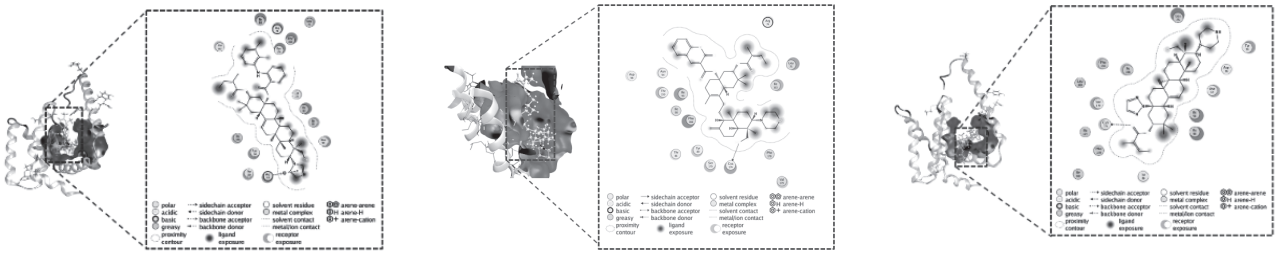
3.6. Molecular Docking

The top 20 compounds with the most favorable IC50 values were subjected to molecular docking studies on the Bcl-2 receptor to assess their inhibitory potential against Hep-G2 cancer cells. Compounds exhibiting strong binding are considered with RMSD < 2 Å and Binding Energy < -7 kcal/mol. Finding 6 compounds with the best binding ability on the Bcl-2 receptor and better than Escin in nature is truly meaningful. Specifically, the better the compound binds to Bcl-2, the more likely it is to inhibit the cell's ability to inhibit apoptosis, thereby increasing this process. When the process of apoptosis is increased, cancer cells will be eliminated. Therefore, it slows down the patient's ability to develop liver cancer. Among these, the six compounds with the best results are presented in Table 5.

Table 5: Molecular docking results for the top 6 compounds with the best performances

Compounds	Binding Energy (kcal/mol)	RMSD (Å)
TPN17	-7.158	1.539
TPN37 (Fig. 5a)	-7.817	1.696
TPN61 (Fig. 5b)	-7.933	1.915
TPN109 (Fig. 5c)	-7.869	1.279
TPN177	-7.367	1.846

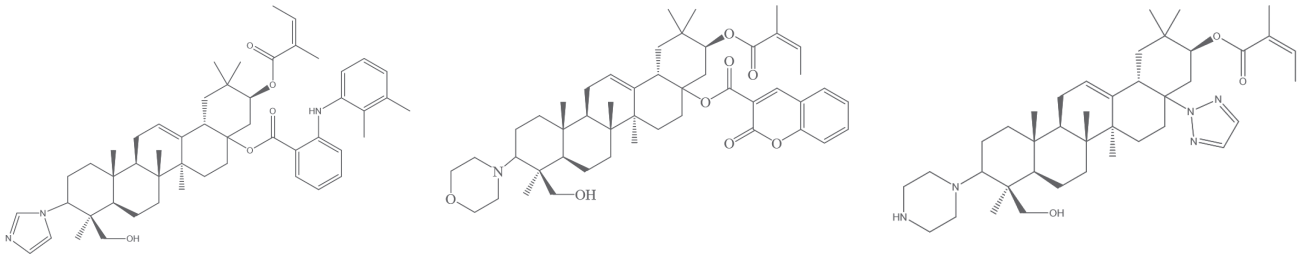
The simulated docking images of the top three compounds, namely TPN37, TPN61, and TPN109, with the best outcomes are depicted in Figure 6.



6a 6b 6c

Figure 6: Docking simulations of compounds TPN37, TPN61, and TPN109, denoted as 6a, 6b, and 6c, respectively, on the Bcl-2 receptor with the code 4D2M.

TPN37 (Figure 6a) forms a binding interaction with the Bcl-2 receptor through a π-cation bond with the amino acid ARG154 (distance = 3.56 Å, energy = -1.6 kcal/mol). TPN61 (Figure 6b) engages with the Bcl-2 receptor through a Hydro bond with CYS174 (distance = 3.37 Å, energy = -1.2 kcal/mol). TPN109 (Figure 6c) establishes a Hydro bond with the Bcl-2 receptor (distance = 3.23 Å, energy = -0.8 kcal/mol).



7a 7b 7c

Figure 7: Structural formulas of TPN37, TPN61, and TPN109, presented as 7a, 7b, and 7c, respectively.

Based on Figure 7, it is evident that potential molecules feature R1 and R2 groups such as pyrazole, mefenamic acid, morpholine, coumarin, piperazine, and triazole.

Considering the results, TPN61 exhibits a predicted IC50 (μM) of 0.938; 1.764, and 1.187 across the QSARMLR, QSARPCR, and QSARANN models, respectively. Notably, QSARANN provides the most accurate prediction, with a value lower than the natural compound. Docking analysis reveals a binding energy of -7.933 kcal/mol and an RMSD of 1.915 Å for TPN61. Compound TPN61 (Figure 4b) binds to the Bcl-2 receptor through a Hydro bond with amino acid CYS174 (distance D = 3.37 Å, energy E = -1.2 kcal/mol). Consequently, TPN61 is selected as the compound with the most promising potential for inhibiting Hep-G2 cancer cells. The research has great practical significance as it proposes a potential compound against liver cancer cells. This helps save time, human resources and funds for further research. Compounds selected and developed to become cancer treatment drugs must have the ability to inhibit cancer cells as a prerequisite. Since then, to become an official cancer treatment drug, it needs to undergo many studies on its effectiveness on patients and unwanted side effects. Therefore, to speed up the process of screening potential compounds, research has been applied well with the support of software and large simulated biological data repositories.

4. CONCLUSION

The research successfully achieved the set research goals when successfully building 3 QSAR models, designing and screening the 20 most potential compounds that met Veber's rule. From there, based on the results of the molecular docking process, 6 compounds with the best binding ability were obtained. Among them, TPN61 has just been predicted IC50 and has the best binding ability. The study has proposed an important result for future research in finding compounds against liver cancer cells from *Paramignya Trimera.*

REFERENCES

1. Alam, M. M. (2022). 1,2,3-Triazole hybrids as anticancer agents: A review. Archiv Der Pharmazie, 355(1), 2100158. https://doi.org/10.1002/ardp.202100158
2. Al-Ghorbani, M., Gouda, M. A., Baashen, M., Alharbi, O., Almalki, F. A., & Ranganatha, L. V. (2022). Piperazine Heterocycles as Potential Anticancer Agents: A Review. Pharmaceutical Chemistry Journal, 56(1), 29–37. https://doi.org/10.1007/s11094-022-02597-z
3. Al-Naqeb, G., Ismail, M., Bagalkotkar, G., & Adamu, H. A. (2010). Vanillin rich fraction regulates LDLR and HMGCR gene expression in HepG2 cells. Food Research International, 43(10), 2437–2443. https://doi.org/10.1016/j.foodres.2010.09.015
4. Arshad, F., Khan, M. F., Akhtar, W., Alam, M. M., Nainwal, L. M., Kaushik, S. K., Akhter, M., Parvez, S., Hasan, S. M., & Shaquiquzzaman, M. (2019). Revealing quinquennial anticancer journey of morpholine: A SAR based review. European Journal of Medicinal Chemistry, 167, 324–356. https://doi.org/10.1016/j.ejmech.2019.02.015
5. Bennani, F. E., Doudach, L., Karrouchi, K., El rhayam, Y., Rudd, C. E., Ansar, M., & El Abbes Faouzi, M. (2022). Design and prediction of novel pyrazole derivatives as potential anti-cancer compounds based on 2D-QSAR study against PC-3, B16F10, K562, MDA-MB-231, A2780, ACHN and NUGC cancer cell lines. Heliyon, 8(8), e10003. https://doi.org/10.1016/j.heliyon.2022.e10003
6. Borková, M. L. (2018). Triterpenoids With Anticancer Properties and Their mechanism of Action.
7. Casini, A., Scozzafava, A., Mastrolorenzo, A., & Supuran, C. (2002). Sulfonamides and Sulfonylated Derivatives as Anticancer Agents. Current Cancer Drug Targets, 2(1), 55–75. https://doi.org/10.2174/1568009023334060

8. Dadashpour, S., & Emami, S. (2018). Indole in the target-based design of anticancer agents: A versatile scaffold with diverse mechanisms. European Journal of Medicinal Chemistry, 150, 9–29. https://doi.org/10.1016/j.ejmech.2018.02.065
9. Dang Thi, T. A., Kim Tuyet, N. T., Pham The, C., Thanh Nguyen, H., Ba Thi, C., Doan Duy, T., D'hooghe, M., & Van Nguyen, T. (2014). Synthesis and cytotoxic evaluation of novel ester-triazole-linked triterpenoid–AZT conjugates. Bioorganic & Medicinal Chemistry Letters, 24(22), 5190–5194. https://doi.org/10.1016/j.bmcl.2014.09.079
10. Ding, W., Li, Y., Li, G., He, H., Li, Z., & Li, Y. (2016). New 30-Noroleanane Triterpenoid Saponins from Holboellia coriacea Diels. Molecules, 21(6), 734. https://doi.org/10.3390/molecules21060734
11. Du, Z., Li, G., Zhou, X., & Zhang, J. (2021). Synthesis of MeON-Glycoside Derivatives of Oleanolic Acid by Neoglycosylation and Evaluation of Their Cytotoxicity against Selected Cancer Cell Lines. Molecules, 26(3), 772. https://doi.org/10.3390/molecules26030772
12. Gao, J., Li, X., Gu, G., Liu, S., Cui, M., & Lou, H.-X. (2012). Facile synthesis of triterpenoid saponins bearing β-Glu/Gal-(1→3)-β-GluA methyl ester and their cytotoxic activities. Bioorganic & Medicinal Chemistry Letters, 22(7), 2396–2400. https://doi.org/10.1016/j.bmcl.2012.02.032
13. Golbraikh, A., & Tropsha, A. (2002). Beware of q2! Journal of Molecular Graphics and Modelling, 20(4), 269–276. https://doi.org/10.1016/S1093-3263(01)00123-1
14. Gu, Z., Lin, S., Yan, W., Chen, D., Zeng, Z., Chen, L., Li, Y., & He, B. (2022a). Enhanced Water Solubility and Anti-Tumor Activity of Oleanolic Acid through Chemical Structure Modification. International Journal of Molecular Sciences, 23(21), 13291. https://doi.org/10.3390/ijms232113291
15. Gu, Z., Lin, S., Yan, W., Chen, D., Zeng, Z., Chen, L., Li, Y., & He, B. (2022b). Enhanced Water Solubility and Anti-Tumor Activity of Oleanolic Acid through Chemical Structure Modification. International Journal of Molecular Sciences, 23(21), 13291. https://doi.org/10.3390/ijms232113291
16. Jain, S., Pattnaik, S., Pathak, K., Kumar, S., Pathak, D., Jain, S., & Vaidya, A. (2018). Anticancer Potential of Thiazole Derivatives: A Retrospective Review. Mini-Reviews in Medicinal Chemistry, 18(8), 640–655. https://doi.org/10.2174/1389557517666171123211321
17. Kang, X., Hu, J., Gao, Z., Ju, Y., & Xu, C. (2012). Synthesis, anti-proliferative and proapoptotic activity of novel oleanolic acid azaheterocyclic derivatives. MedChemComm, 3(10), 1245. https://doi.org/10.1039/c2md20051a
18. Li, D.-Q., Wu, J., Liu, L.-Y., Wu, Y.-Y., Li, L.-Z., Huang, X.-X., Liu, Q.-B., Yang, J.-Y., Song, S.-J., & Wu, C.-F. (2015). Cytotoxic triterpenoid glycosides (saikosaponins) from the roots of Bupleurum chinense. Bioorganic & Medicinal Chemistry Letters, 25(18), 3887–3892. https://doi.org/10.1016/j.bmcl.2015.07.053
19. Li, Y., Cao, T.-T., Guo, S., Zhong, Q., Li, C.-H., Li, Y., Dong, L., Zheng, S., Wang, G., & Yin, S.-F. (2016). Discovery of Novel Allopurinol Derivatives with Anticancer Activity and Attenuated Xanthine Oxidase Inhibition. Molecules, 21(6), 771. https://doi.org/10.3390/molecules21060771
20. Liang, D., Hao, Z.-Y., Zhang, G.-J., Zhang, Q.-J., Chen, R.-Y., & Yu, D.-Q. (2011). Cytotoxic Triterpenoid Saponins from Lysimachia clethroides. Journal of Natural Products, 74(10), 2128–2136. https://doi.org/10.1021/np2004038
21. Lu, Y., Chen, M., Liu, F., Xu, Z., Tian, X., Xie, Y., & Huang, C. (2020). Synthesis and Cytotoxic Activity of Novel C-23-Modified Asiatic Acid Derivatives. Molecules, 25(16), 3709. https://doi.org/10.3390/molecules25163709
22. Ma, G., Guo, W., Zhao, L., Zheng, Q., Sun, Z., Wei, J., Yang, J., & Xu, X. (2013). Two New Triterpenoid Saponins from the Root of Platycodon grandiflorum. Chemical and Pharmaceutical Bulletin, 61(1), 101–104. https://doi.org/10.1248/cpb.c12-00713
23. Moustafa, G. O., Shalaby, A., Naglah, A. M., Mounier, M. M., El-Sayed, H., Anwar, M. M., & Nossier, E. S. (2021). Synthesis, Characterization, In Vitro Anticancer Potentiality, and Antimicrobial Activities of Novel Peptide–Glycyrrhetinic-Acid-Based Derivatives. Molecules, 26(15), 4573. https://doi.org/10.3390/molecules26154573
24. Mu, L.-H., Huang, C.-L., Zhou, W.-B., Guo, D.-H., & Liu, P. (2013). Methanolysis of triterpenoid saponin from Ardisia gigantifolia stapf. And structure–activity relationship study against cancer cells. Bioorganic & Medicinal Chemistry Letters, 23(22), 6073–6078. https://doi.org/10.1016/j.bmcl.2013.09.029

25. Nistor, G., Trandafirescu, C., Prodea, A., Milan, A., Cristea, A., Ghiulai, R., Racoviceanu, R., Mioc, A., Mioc, M., Ivan, V., & Șoica, C. (2022). Semisynthetic Derivatives of Pentacyclic Triterpenes Bearing Heterocyclic Moieties with Therapeutic Potential. Molecules, 27(19), 6552. https://doi.org/10.3390/molecules27196552
26. Patel, S. S., Tripathi, R., Chavda, V. K., & Savjani, J. K. (2020). Anticancer Potential of Mefenamic Acid Derivatives with Platelet-Derived Growth Factor Inhibitory Property. Anti-Cancer Agents in Medicinal Chemistry, 20(8), 998–1008. https://doi.org/10.2174/1871520620666200415100614
27. Pei, Y.-H., Kwon, O.-K., Lee, J.-S., Cha, H.-J., Ahn, K.-S., Oh, S.-R., Lee, H.-K., & Chin, Y.-W. (2013). Triterpenes with Cytotoxicity from the Leaves of Vernicia fordii. Chemical and Pharmaceutical Bulletin, 61(6), 674–677. https://doi.org/10.1248/cpb.c13-00055
28. Phi, T. C. M., Chu, H. H., Trieu Le, N., & Nguyen, D. B. (2020). Phylogenetic relationship of Paramignya trimera and its relatives: An evidence for the wide sexual compatibility. Scientific Reports, 10(1), 21662. https://doi.org/10.1038/s41598-020-78448-2
29. Quang, T. H., Ngan, N. T. T., Minh, C. V., Kiem, P. V., Nhiem, N. X., Tai, B. H., Thao, N. P., Tung, N. H., Song, S. B., & Kim, Y. H. (2011). Anti-inflammatory Triterpenoid Saponins from the Stem Bark of Kalopanax pictus. Journal of Natural Products, 74(9), 1908–1915. https://doi.org/10.1021/np200382s
30. Rickard, B. P., Ho, H., Tiley, J. B., Jaspers, I., & Brouwer, K. L. R. (2021). E-Cigarette Flavoring Chemicals Induce Cytotoxicity in HepG2 Cells. ACS Omega, 6(10), 6708–6713. https://doi.org/10.1021/acsomega.0c05639
31. Sharma, P., LaRosa, C., Antwi, J., Govindarajan, R., & Werbovetz, K. A. (2021). Imidazoles as Potential Anticancer Agents: An Update on Recent Studies. Molecules, 26(14), 4213. https://doi.org/10.3390/molecules26144213
32. Urban, M., Vlk, M., Dzubak, P., Hajduch, M., & Sarek, J. (2012). Cytotoxic heterocyclic triterpenoids derived from betulin and betulinic acid. Bioorganic & Medicinal Chemistry, 20(11), 3666–3674. https://doi.org/10.1016/j.bmc.2012.03.066
33. Vega-Granados, K., Medina-O'Donnell, M., Rivas, F., Reyes-Zurita, F. J., Martinez, A., Alvarez De Cienfuegos, L., Lupiañez, J. A., & Parra, A. (2021). Synthesis and Biological Activity of Triterpene–Coumarin Conjugates. Journal of Natural Products, 84(5), 1587–1597. https://doi.org/10.1021/acs.jnatprod.1c00128
34. Vega-Granados, K., Medina-O'Donnell, M., Rivas, F., Reyes-Zurita, F. J., Martinez, A., Alvarez de Cienfuegos, L., Lupiañez, J. A., & Parra, A. (2021). Synthesis and Biological Activity of Triterpene–Coumarin Conjugates. Journal of Natural Products, 84(5), 1587–1597. https://doi.org/10.1021/acs.jnatprod.1c00128
35. Wang, O., Liu, S., Zou, J., Lu, L., Chen, L., Qiu, S., Li, H., & Lu, X. (2011). Anticancer Activity of 2α, 3α, 19β, 23β-Tetrahydroxyurs-12-en-28-oic Acid (THA), a Novel Triterpenoid Isolated from Sinojackia sarcocarpa. PLoS ONE, 6(6), e21130. https://doi.org/10.1371/journal.pone.0021130
36. Wang, X.-Y., Gao, H., Zhang, W., Li, Y., Cheng, G., Sun, X.-L., & Tang, H.-F. (2013). Bioactive oleanane-type saponins from the rhizomes of Anemone taipaiensis. Bioorganic & Medicinal Chemistry Letters, 23(20), 5714–5720. https://doi.org/10.1016/j.bmcl.2013.08.006
37. WHO. (2023). Cancer Today. https://gco.iarc.who.int/today/
38. Yu, L., Wang, X., Wei, X., Wang, M., Chen, L., Cao, S., Kang, N., & Qiu, F. (2012). Triterpenoid saponins from Xanthoceras sorbifolia Bunge and their inhibitory activity on human cancer cell lines. Bioorganic & Medicinal Chemistry Letters, 22(16), 5232–5238. https://doi.org/10.1016/j.bmcl.2012.06.061
39. Zhang, H.-J., Zhang, G.-R., Piao, H.-R., & Quan, Z.-S. (2018). Synthesis and characterisation of celastrol derivatives as potential anticancer agents. Journal of Enzyme Inhibition and Medicinal Chemistry, 33(1), 190–198. https://doi.org/10.1080/14756366.2017.1404590