

Original article:

**RATIONAL DESIGN OF NOVEL COUMARINS:
A POTENTIAL TREND FOR ANTIOXIDANTS IN COSMETICS**

Apilak Worachartcheewan^{1*}, Veda Prachayasittikul², Supaluk Prachayasittikul^{2*},
Visanu Tantivit¹, Chareef Yeeyahya¹, Virapong Prachayasittikul³

¹ Department of Community Medical Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

² Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

³ Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

* Corresponding authors: E-mail: apilak.woa@mahidol.ac.th (A.W), supaluk@g.swu.ac.th (S.P.); Phone: (662) 441-4376; Fax: (662) 441-4380

<http://dx.doi.org/10.17179/excli2019-1903>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT

Coumarins are well-known for their antioxidant effect and aromatic property, thus, they are one of ingredients commonly added in cosmetics and personal care products. Quantitative structure-activity relationships (QSAR) modeling is an *in silico* method widely used to facilitate rational design and structural optimization of novel drugs. Herein, QSAR modeling was used to elucidate key properties governing antioxidant activity of a series of the reported coumarin-based antioxidant agents (**1-28**). Several types of descriptors (calculated from 4 softwares i.e., Gaussian 09, Dragon, PaDEL and Mold² softwares) were used to generate three multiple linear regression (MLR) models with preferable predictive performance ($Q^2_{LOO-CV} = 0.813-0.908$; $RMSE_{LOO-CV} = 0.150-0.210$; $Q^2_{Ext} = 0.875-0.952$; $RMSE_{Ext} = 0.104-0.166$). QSAR analysis indicated that number of secondary amines (nArNHR), polarizability (G2p), electronegativity (D467, D580, SpMin2_Bhe, and MATS8e), van der Waals volume (D491 and D461), and H-bond potential (SHBint4) are important properties governing antioxidant activity. The constructed models were also applied to guide *in silico* rational design of an additional set of 69 structurally modified coumarins with improved antioxidant activity. Finally, a set of 9 promising newly design compounds were highlighted for further development. Structure-activity analysis also revealed key features required for potent activity which would be useful for guiding the future rational design. In overview, our findings demonstrated that QSAR modeling could possibly be a facilitating tool to enhance successful development of bioactive compounds for health and cosmetic applications.

Keywords: Coumarin, antioxidant activity, rational design, QSAR, MLR

INTRODUCTION

Free radicals (or oxidants) are highly reactive molecules containing an unpaired electron, which are generated as by-products of physiological processes and intracellular

pathways (Valko et al., 2007; Winyard et al., 2005). These oxidants are well-known for their harmful potential and deleterious effects in cellular components (i.e., DNA, proteins and lipids). In normal condition, these radicals are scavenged/neutralized by antioxidant

defense mechanism (i.e., endogenous antioxidant molecules and antioxidant enzymes) to prevent cellular oxidative damages. However, the shift of oxidative balance occurs in a condition whereby radicals are overproduced or antioxidant defense mechanism is depleted. This situation leads to excessive accumulation of free radicals and oxidative stress. Oxidative damage involves in pathogenesis and progression of many chronic and aging diseases (i.e., cancer, diabetes mellitus, neurodegenerative diseases, and cardiovascular diseases) (Valko et al., 2007; Winyard et al., 2005). Furthermore, free radicals have been recognized as one of the factors contributing to aging skin (Bogdan Allemann and Baumann, 2008). Antioxidant compounds have been well-recognized for their wide-ranging health applications, especially in cosmeceutical area. Currently, an addition of antioxidants as active ingredients in cosmetics and personal care products has been widely documented (Kusumawati and Indrayanto, 2013; Lupo, 2001). Therefore, discovery of novel potent antioxidant compounds, both from chemical synthesis (Prachayasittikul et al., 2009a; Subramanyam et al., 2017; Worachartcheewan et al., 2012) and natural-derived sources (Elansary et al., 2018; Krishnaiah et al., 2011; Prachayasittikul et al., 2008, 2009b, 2013; Wongsawatkul et al., 2008), has been noted to be an attractive research area, especially in cosmetic applications (Kusumawati and Indrayanto, 2013; Lupo, 2001).

Coumarins, known as benzopyrones, are natural secondary metabolites bearing fused benzene and α -pyrone rings (Witaicenis et al., 2014). Natural-derived coumarins are found in a wide range of plants (Lee et al., 2007; Rodríguez-Hernández et al., 2019; Saleem et al., 2019; Venditti et al., 2019). Coumarins displayed a variety of biological activities including antimicrobial (Arshad et al., 2011), antioxidant (Erzincan et al., 2015), anticancer (Nasr et al., 2014), and anti-inflammatory (Witaicenis et al., 2014) activities. Although synthetic coumarins were banned for oral products due to their potential toxicities, they

are attractive for topical uses due to their high skin penetrating property (Stiefel et al., 2017). Additionally, coumarins are widely used as fragrance ingredient in cosmetics and personal care products because of their sweet herbaceous scent (Ma et al., 2015; Stiefel et al., 2017). Antioxidant property and protective effects against skin photo-aging of coumarins have also been remarked in cosmetic area (Kostova et al., 2011; Lee et al., 2007). Previously, a set of synthesized coumarin derivatives containing 2-methylbenzothiazolines, sulphonamides, and amides were reported to exhibit antioxidant activity with IC_{50} values range of 0.024-2.888 mM (Khoobi et al., 2011; Saeedi et al., 2014). However, deeper understanding of structure-activity relationships (SAR) and mechanism of action is still necessary for an effective rational design of coumarin-based antioxidant agents (Kostova et al., 2011).

Computational approaches have been widely recognized to facilitate and increase success rate of drug development (Nantasenamat and Prachayasittikul, 2015; Prachayasittikul et al., 2015a). Quantitative structure-activity relationship (QSAR) modeling is an *in silico* method to reveal the relationship between chemical structures of the compounds and their biological activities. QSAR modeling provides useful findings such as key features, properties, or moieties that are required for potent activity, which would benefit further rational design of the related compounds. Currently, success stories of QSAR-driven rational design of several classes of promising lead compounds have been documented for anticancer agents (Prachayasittikul et al., 2015b), aromatase inhibitors (Prachayasittikul et al., 2017), and sirtuin-1 activators (Pratiwi et al., 2019). In cosmetic area, QSAR modeling has been employed to improve understanding towards SAR of tyrosinase inhibitors (Gao, 2018; Khan, 2012).

Accordingly, this study aims to construct QSAR models to elucidate SAR of a set of antioxidant coumarin derivatives (**1-28**, Figure 1) originally reported by Khoobi et al. (2011) and Saeedi et al. (2014). Herein, QSAR mod-

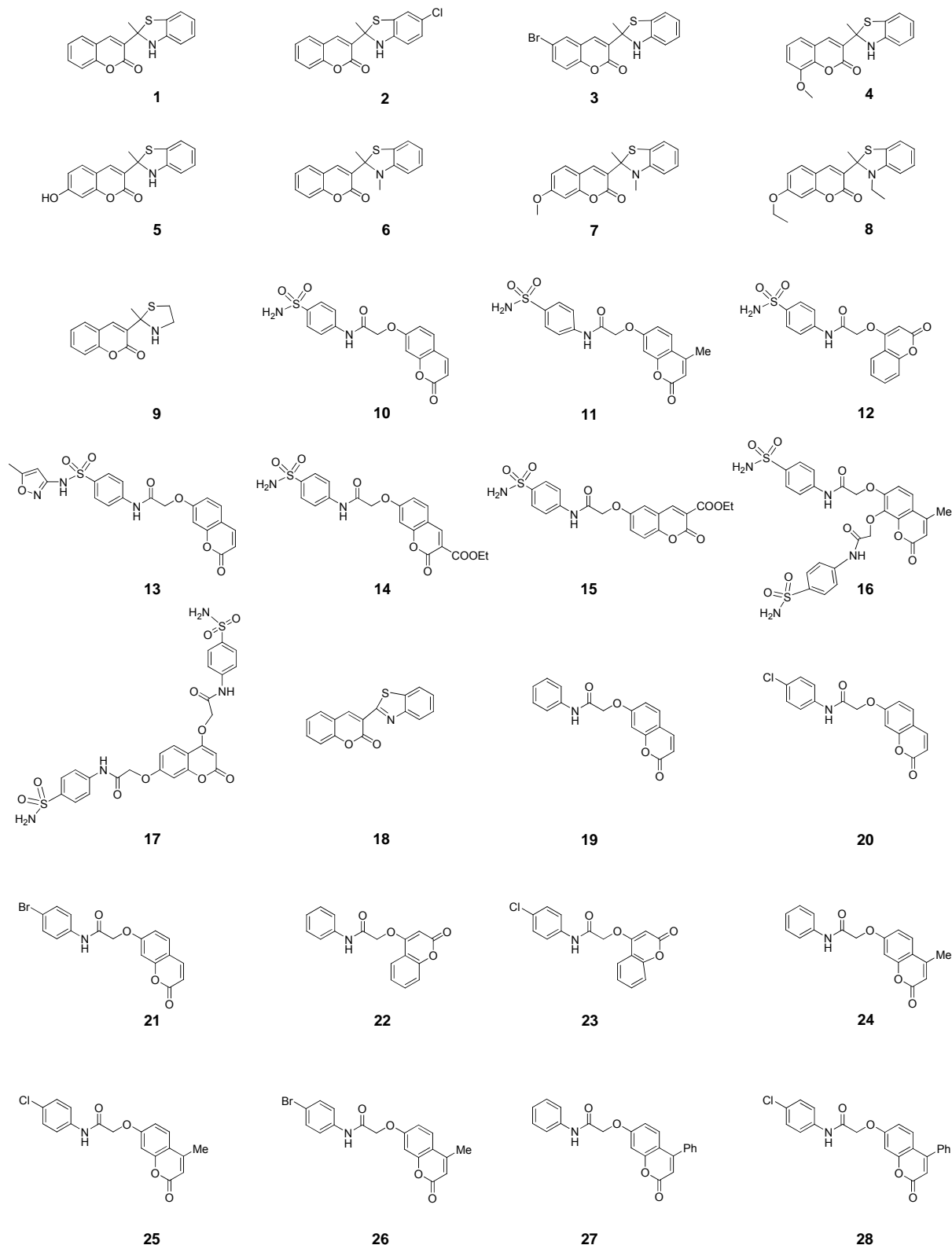


Figure 1: Molecular structures of coumarin derivatives (1-28)

els were constructed using multiple linear regression (MLR) algorithm to clearly demon-

strate the linear relationship along with insight SAR analysis. In an attempt to find a robust and validating QSAR models, chemical

descriptors were generated using different four softwares (i.e., Gaussian 09, Dragon, PaDEL and Mold² softwares) to increase a variety of represented physicochemical properties. Consequently, an additional set of structurally modified compounds were rationally designed based on key findings of the constructed models, and their antioxidant activities were predicted to reveal the promising ones with potential for further synthesis and development.

MATERIALS AND METHODS

Data set

A data set of twenty-eight coumarin-based antioxidants (**1-28**, Figure 1) was retrieved from the literature (Khoobi et al., 2011; Saeedi et al., 2014), in which their antioxidant activities are presented in Table 1. All

tested compounds were evaluated by 1,1-diphenyl-2-picrylhydrazyl (DPPH) assay (detailed methodology is provided in original literatures (Khoobi et al., 2011; Saeedi et al., 2014)). The activity was denoted as an IC₅₀ value (mM) which indicates concentration of the compound which can inhibit 50 % of the generated DPPH radicals in experimental setting. As a part of data pre-processing, the unit of IC₅₀ values was converted from mM to M, and the IC₅₀ values were further transformed into pIC₅₀ (–log IC₅₀) by taking the negative logarithm to base 10 as shown in Table 1. The compound with high pIC₅₀ (low IC₅₀) represented the high antioxidant activity. A schematic workflow of QSAR model development is provided in Figure 2.

Table 1: Experimental and predicted antioxidant activities (pIC₅₀) of coumarin derivatives (**1-28**) using multiple linear regression method

Compound	Experimental activity		Predicted activity (pIC ₅₀)		
	IC ₅₀ (mM)	pIC ₅₀	Model 1	Model 2	Model 3
1	0.130	3.887	3.917	3.961	3.893
2	0.466	3.332	3.825	3.863	3.454
3	0.099	4.003	3.789	4.006	3.990
4	0.106	3.973	3.967	3.883	4.120
5	0.024	4.612	4.217	4.530	4.450
6	2.888	2.539	2.746	2.711	2.776
7	2.382	2.623	2.898	2.581	2.670
8	0.681	3.167	2.944	3.195	3.023
9	2.560	2.592	2.362	2.487	2.268
10	1.570	2.804	2.756	2.943	2.934
11	0.660	3.180	3.010	2.936	3.006
12	1.230	2.910	3.100	3.230	2.823
13	0.950	3.022	3.091	2.994	2.959
14	1.750	2.757	2.998	2.647	2.937
15	1.150	2.939	2.916	2.978	2.959
16	0.810	3.092	3.249	3.036	2.818
17	1.080	2.967	2.826	2.942	3.027
18	0.550	3.260	3.217	3.336	3.201
19	2.290	2.640	2.812	2.636	2.853
20	1.140	2.943	2.732	2.685	2.769
21	0.940	3.027	2.930	2.748	2.817
22	1.110	2.955	2.779	2.859	2.879
23	1.230	2.910	3.100	3.004	2.823
24	2.070	2.684	2.976	2.744	2.863
25	2.080	2.682	2.680	2.892	2.798
26	1.680	2.775	2.789	2.886	2.830
27	1.860	2.730	2.715	2.903	2.921
28	0.980	3.009	2.841	2.990	2.858

Compounds **1**, **6**, **15**, **21** and **27** were used as external set. Models 1, 2 and 3 were obtained by Eqs. 2, 3 and 4, respectively.

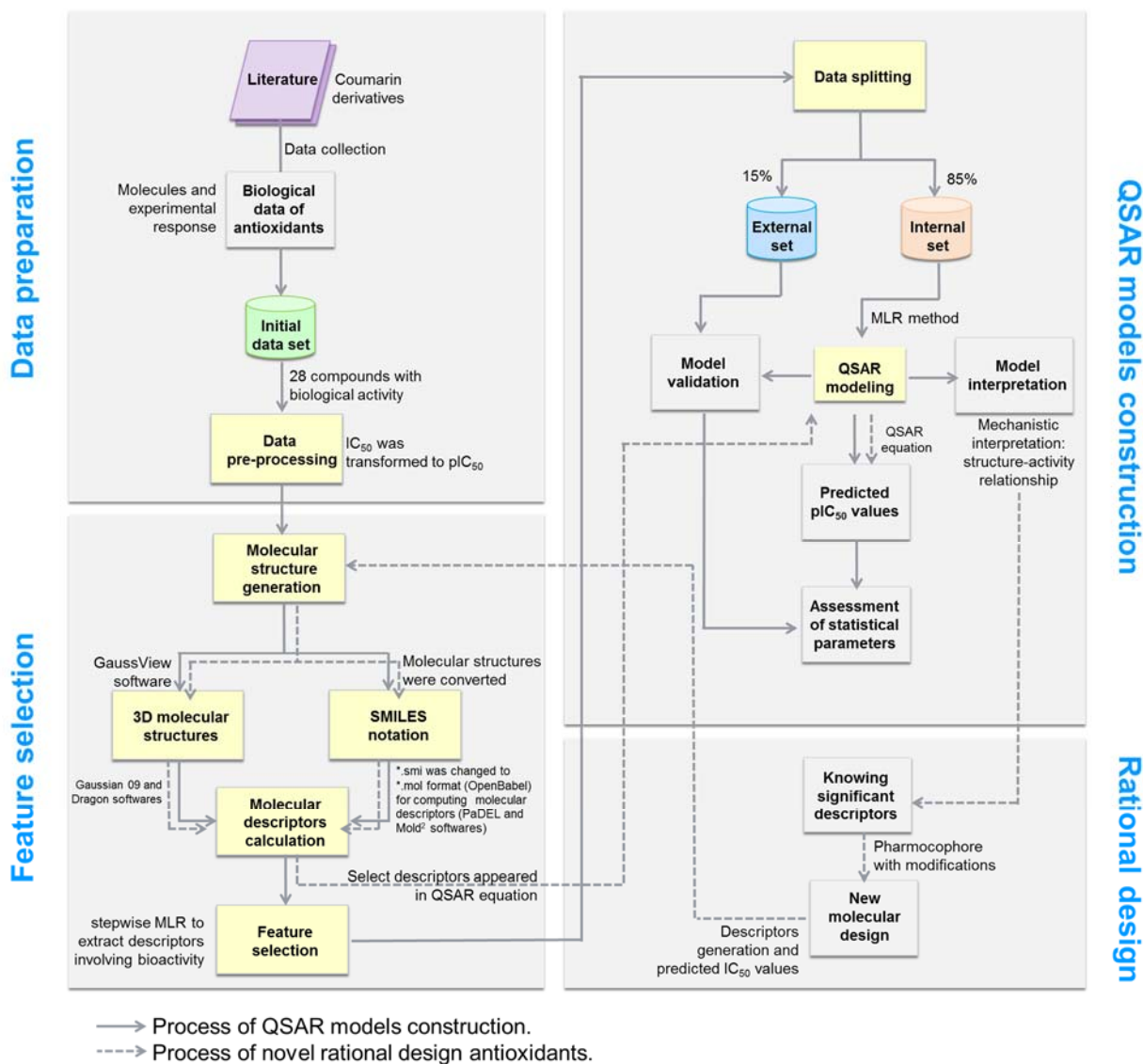


Figure 2: Schematic workflow of QSAR models

Molecular structure optimization

Molecular structures of the coumarin derivatives were constructed by GaussView (Dennington et al., 2003), which were subjected to geometrical optimization by Gaussian 09 (Revision A.02) (Frisch et al., 2009) at the semi-empirical level using Austin Model 1 (AM1) followed by density functional theory (DFT) calculation using Becke's three-parameter hybrid method and the Lee–Yang–Parr correlation functional (B3LYP) together with the 6–31 g(d) basis.

Descriptor calculation and feature selection

The physicochemical properties (i.e., quantum chemical and molecular descriptors) were generated by different calculating softwares including Gaussian 09, Dragon, version 5.5. (Talete, 2007), PaDEL, version 2.20 (Yap, 2011) and Mold², version 2.0 (Hong et al., 2008) softwares. The calculated descriptors as numerical values could be used to represent properties of the compounds, and were further used as predictors (X variables) for QSAR model construction. List of calculated descriptors are shown as follows.

Quantum chemical descriptors calculation obtained by low energy conformers from the geometrical optimization using Gaussian 09 composed of the total energy (E_{total}) of the molecule, the highest occupied molecular orbital energy (E_{HOMO}), the lowest unoccupied molecular orbital energy (E_{LUMO}), the total dipole moment (μ) of the molecule, the electron affinity (EA), the ionization potential (IP), the energy difference of HOMO and LUMO ($HOMO-LUMO_{Gap}$), Mulliken electronegativity (χ), hardness (η), softness (S), electrophilicity (ω), electrophilic index (ω_i), and the mean absolute atomic charge (Q_m) (Worachartcheewan et al., 2014). Furthermore, the output files from geometrical optimization of Gaussian 09 were used as the input data for calculating a set of 3,224 molecular descriptors using Dragon software. The calculated descriptors included 22 classes comprising 48 constitutional descriptors, 119 topological descriptors, 47 walk and path counts, 33 connectivity indices, 47 information indices, 96 2D autocorrelation, 107 edge adjacency indices, 64 burden eigenvalues, 21 topological charge indices, 44 eigenvalue based indices, 41 randic molecular profiles, 74 geometrical descriptors, 150 RDF descriptors, 160 3D-MoRSE descriptors, 99 WHIM descriptors, 197 GETAWAY descriptors, 154 functional group counts, 120 atom centered fragments, 14 charge descriptors, 29 molecular properties, 780 2D binary fingerprints, and 780 2D frequency fingerprints.

An additional set of molecular descriptors was calculated by PaDEL software to give 1,444 1D and 2D descriptors, and Mold² software to generate 777 descriptors by encoding the 2D chemical structure information. Before the calculation, the molecular structures were saved to *.smi and then converted to *.mol files using OpenBabel version 2.3.2 (The Open Babel Package 2015). The *.mol files were used as the input data for calculation by PaDEL and Mold² softwares.

Descriptors selection was performed to filter a set of important informative de-

scriptors from a whole set of descriptors. Feature selection was initially performed by stepwise multiple linear regression (MLR) using SPSS statistics 18.0 (SPSS Inc., USA) followed by determination of intercorrelation using Pearson's correlation coefficient using cutoff value of $|r| \geq 0.9$. Any pairs of descriptors with $|r| \geq 0.9$ were defined as highly correlated predictors, and one of them was excluded.

Data splitting

The data set of coumarin derivatives (**1-28**) was randomly selected, in which 85 % (~23 compounds) of the original data set was used as the training and the leave one-out cross-validation (LOO-CV) sets, and 15 % (~5 compounds) was used as the external set. The training set was employed to generate the QSAR models, whereas LOO-CV and external sets were used to evaluate the models. LOO-CV method was performed for internal validation by excluding one sample out from the whole data set to be used as the testing set while the remaining $N-1$ samples were used as the training set (Prachayasittikul et al., 2014). This sampling process was repeated iteratively until every sample in the data set was used as the testing set. The external sets were used to validate the models.

Multivariate analysis

QSAR models were generated using the MLR according to the equation 1.

$$Y = B_0 + \sum B_n X_n \quad (1)$$

where Y is the antioxidant activity (pIC_{50}), B_0 is the intercept, and B_n are the regression coefficients of the descriptors X_n . The MLR method was performed using Waikato Environment for Knowledge Analysis (Weka), version 3.4.5 (Witten et al., 2011).

Models evaluation

Statistical parameters were used to evaluate predictive performance of the constructed QSAR models. The calculated parameters included correlation coefficient (R^2), root mean squared error ($RMSE$), predictivity (Q^2), vari-

ance ratio (*F ratio*), adjusted correlation coefficient (R^2_{Adj}), standard deviation (*s*) and predicted residual sum of squares (PRESS) (Saghaie et al., 2013; Worachartcheewan et al., 2013).

RESULTS AND DISCUSSION

Molecular descriptors selection

Chemical structures of the compounds and their antioxidant activities (Table 1) were used for construction of predictive models. The compounds were geometrically optimized with semi-empirical method AM1 followed by DFT/B3LYP/6–31 g(d) basis using Gaussian 09 to obtain lower-energy conformers. The optimized compounds were extracted to obtain 13 quantum chemical descriptors. These compounds were subsequently used as input files for calculating an additional set of 3,224 molecular descriptors (0D-3D) using Dragon software. The calculated descriptors with constant values and multi-collinearity were determined and removed to give a final

set of 1,489 descriptors. In addition, original molecular structures of compounds were saved as *.smi file format and were converted into *.mol files using OpenBabel version 2.3.2. These *.mol files then were used as input files for descriptors calculation using Mold² and PaDEL softwares to obtain sets of 777 Mold² 2D descriptors and 1,444 PaDEL 0D-2D descriptors, respectively. Consequently, feature selection was performed to select a set of informative descriptors for the whole calculated set. Descriptors showing significant correlation with their bioactivities were selected using stepwise MLR. A set of 14 informative descriptors included 4 Dragon descriptors (i.e., nArNHR, ISH, B04[O-O] and G2p), 6 Mold² descriptors (i.e., D467, D278, D491, D384, D580 and D461), 4 PaDEL descriptors (i.e., SHBint4, SpMin2_Bhe, MATS8e and SssCH2) were obtained. Definition and numerical values of important descriptors are shown in Tables 2 and 3, respectively. Furthermore, the intercor-

Table 2: Definition of descriptors in QSAR models (1-3)

Symbol	Descriptor	Class
nArNHR	Number of secondary amines (aromatic)	Functional group counts
ISH	Standardized information content on the leverage equality	GETAWAY descriptors
B04[O-O]	Presence/absence of O - O at topological distance 4	2D Atom Pairs
G2p	2nd component symmetry directional WHIM index / weighted by polarizability	WHIM descriptors
D467	Geary topological structure autocorrelation length-5 weighted by atomic Sanderson electronegativities	2D descriptors
D278	Total information content order-3 index	2D descriptors
D491	Moran topological structure autocorrelation length-5 weighted by atomic van der Waals volumes	2D descriptors
D384	Sum of topological distance between the vertices S and Cl	2D descriptors
D580	Highest eigenvalue from Burden matrix weighted by electronegativities Sanderson-Scale order-1	2D descriptors
D461	Geary topological structure autocorrelation length-7 weighted by atomic van der Waals volumes	2D descriptors
SHBint4	Sum of E-state descriptors of strength for potential hydrogen bonds of path length 4	2D (Atom type electrotopological state)
SpMin2_Bhe	Smallest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative Sanderson electronegativities	2D (Burden modified eigenvalues)
MATS8e	Moran autocorrelation - lag 8 / weighted by Sanderson electronegativities	2D (Autocorrelation)
SssCH2	Sum of atom-type E-State: -CH2-	2D (Atom type electrotopological state)

Table 3: Values of significant molecular descriptors for QSAR models (1-3)

Compound	nArNHR	ISH	B04[O-O]	G2p	D467	D278	D491	D384	D580	D461	SHBint4	SpMin2_Bhe	MATS8e	SssCH2
1	1	0.832	0	0.164	0.773	0.460	0.518	0	5.406	0.465	4.3184	1.8130	-0.1379	0.0000
2	1	0.898	0	0.164	0.708	0.489	0.588	4	5.410	0.494	4.0816	1.8175	0.3169	0.0000
3	1	0.878	0	0.180	0.722	0.489	0.544	0	5.416	0.463	4.2075	1.8130	-0.2709	0.0000
4	1	0.839	0	0.174	0.774	0.462	0.54	0	5.440	0.356	4.4229	1.8130	-0.2951	0.0000
5	1	0.850	1	0.193	0.805	0.506	0.471	0	5.420	0.186	7.5875	1.8130	0.0848	0.0000
6	0	0.829	0	0.169	0.646	0.426	0.805	0	5.419	0.520	0.0000	1.8187	-0.1210	0.0000
7	0	0.891	1	0.170	0.556	0.427	0.831	0	5.434	0.315	0.0000	1.8187	-0.0071	0.0000
8	0	0.879	1	0.173	0.656	0.423	0.649	0	5.436	0.320	0.0000	1.8212	0.0229	1.4684
9	0	0.914	0	0.179	0.501	0.528	0.933	0	5.390	1.249	3.8389	1.5636	0.1182	2.3687
10	0	0.896	1	0.158	0.515	0.488	0.754	0	5.538	0.403	0.0000	1.8952	0.1915	-0.2690
11	0	0.864	1	0.173	0.479	0.49	0.688	0	5.538	0.544	0.0000	1.9109	0.1872	-0.2700
12	0	0.869	1	0.183	0.652	0.488	0.756	0	5.538	0.316	0.0000	1.8950	0.1419	-0.3336
13	0	0.828	1	0.161	0.503	0.468	0.604	0	5.586	0.515	0.0000	1.8955	0.1400	-0.2867
14	0	0.839	1	0.151	0.493	0.481	0.821	0	5.538	0.373	0.0000	1.9068	0.3121	-0.2161
15	0	0.847	1	0.151	0.544	0.481	0.744	0	5.538	0.326	0.0000	1.9068	0.2481	-0.2215
16	0	0.795	1	0.159	0.553	0.459	0.664	0	5.539	0.329	0.0000	1.9191	0.1378	-1.0815
17	0	0.869	1	0.154	0.595	0.457	0.714	0	5.538	0.355	0.0000	1.9192	0.0588	-0.8724
18	0	0.788	0	0.209	0.501	0.413	0.472	0	5.429	0.471	0.0000	1.8586	-0.2830	0.0000
19	0	0.959	1	0.193	0.509	0.433	0.794	0	5.371	0.619	0.0000	1.8339	-0.1138	-0.1275
20	0	0.923	1	0.171	0.502	0.462	0.827	0	5.371	0.633	0.0000	1.8379	-0.0275	-0.1183
21	0	0.923	1	0.193	0.489	0.462	0.82	0	5.371	0.612	0.0000	1.8395	-0.0511	-0.1139
22	0	0.898	1	0.163	0.652	0.433	0.789	0	5.395	0.484	0.0000	1.8340	-0.1900	-0.1921
23	0	0.897	1	0.179	0.656	0.462	0.826	0	5.395	0.501	0.0000	1.8380	-0.0987	-0.1829
24	0	0.942	1	0.200	0.473	0.440	0.719	0	5.392	0.730	0.0000	1.8342	-0.1244	-0.1285
25	0	0.945	1	0.174	0.465	0.465	0.736	0	5.393	0.726	0.0000	1.8381	-0.0346	-0.1193
26	0	0.945	1	0.187	0.457	0.465	0.731	0	5.393	0.712	0.0000	1.8398	-0.0595	-0.1149
27	0	0.920	1	0.165	0.640	0.383	0.598	0	5.418	0.612	0.0000	1.8350	-0.2016	-0.1498
28	0	0.909	1	0.176	0.647	0.406	0.624	0	5.418	0.629	0.0000	1.8390	-0.1186	-0.1406

relation matrix between pair of molecular descriptors was performed using Pearson's correlation coefficient (r) (Supplementary Tables 1-3). Cutoff value of $|r| \geq 0.9$ was used to determine the intercorrelation. The results showed that there was no intercorrelation within a set of selected descriptors as displayed by low $|r|$ values ≤ 0.9 , which suggested that each descriptor was independent from other descriptors. Finally, a set of 14 selected descriptors was further employed to construct 3 QSAR models (according to types of software used to calculate descriptor values) for predicting antioxidant activity of the coumarin derivatives.

QSAR models

Descriptors obtained from these softwares have been demonstrated for their successful QSAR modeling such as antioxidant (Alisi et al., 2018; Rastija et al., 2018), antimicrobial (Alyar et al., 2009; Basic et al., 2014; Podunavac-Kuzmanović et al., 2009), anticancer (Sławiński et al., 2017; Suvannang et al., 2018) and antiviral (Duchowicz et al., 2018; Saavedra et al., 2018; Worachartcheewan et al., 2019) activities. Herein, three models were separately constructed based on the types of key descriptors (i.e., model 1 Dragon descriptors, model 2 Mold² descriptors, and model 3 PaDEL descriptors). A set of 14 selected informative descriptors (as independent variables, Table 2) and antioxidant activities (pIC₅₀ values as dependent variables) of the studied compounds were included in the data sets for construction of QSAR models using Eq. (1). Before building the models, the data set of coumarin derivatives (**1-28**) was split into training, LOO-CV, and external sets. The training set was used to construct the model using MLR algorithm whereas both LOO-CV and external sets were utilized for validating the constructed models. Compounds **1**, **6**, **15**, **21** and **27** were randomly selected to be used as external sets, while the remaining 23 compounds in the data sets (i.e., **2-5**, **7-14**, **16-20**, **22-26** and **28**) were employed as training set. As a result, three

QSAR models (models 1-3) were successfully constructed for predicting antioxidant activities (pIC₅₀ values) of the studied coumarin analogs.

Model 1 (Dragon descriptors)

$$\text{pIC}_{50} = 1.2245(\text{nArNHR}) - 4.3187(\text{ISH}) + 0.3944(\text{B04[O-O]}) + 8.1471(\text{G2p}) + 4.9498 \quad (2)$$

$$N_{Tr} = 23, R^2_{Tr} = 0.901, R^2_{Adj} = 0.885, RMSE_{Tr} = 0.154$$

$$N_{LOO-CV} = 23, Q^2_{LOO-CV} = 0.813, RMSE_{LOO-CV} = 0.210, s = 0.238, F \text{ ratio} = 13.043, \text{PRESS} = 1.019$$

$$N_{Ext} = 5, Q^2_{Ext} = 0.952, RMSE_{Ext} = 0.104$$

where N_{Tr} , N_{LOO-CV} and N_{Ext} are the number of compounds of training, LOO-CV and external sets. R^2_{Adj} is the adjusted R^2 .

Four molecular descriptors calculated from Dragon software were used as predictors to construct QSAR model 1 as shown in Eq. (2). Statistical parameters indicating predictive performance of the model are summarized in Table 4. Training set showed $R^2_{Tr} = 0.901$, $RMSE_{Tr} = 0.154$, while the LOO-CV set displayed the $Q^2_{LOO-CV} = 0.813$, $RMSE_{LOO-CV} = 0.210$, and the external set with $Q^2_{Ext} = 0.952$, $RMSE_{Ext} = 0.104$. Comparative plot of experimental versus predicted antioxidant activities (pIC₅₀) is shown in Figure 3a. Residual values were calculated as a difference between experimental and predicted pIC₅₀ values. Small residual values indicated the precision of model prediction. Plot of the residual values is shown in Figure 3b.

Model 2 (Mold² descriptors)

$$\text{pIC}_{50} = 1.4575(\text{D467}) + 7.1429(\text{D278}) - 2.6745(\text{D491}) - 0.1328(\text{D384}) - 2.0251(\text{D580}) - 0.5180(\text{D461}) + 12.1227 \quad (3)$$

$N_{Tr} = 23$, $R^2_{Tr} = 0.958$, $R^2_{Adj} = 0.946$, $RMSE_{Tr} = 0.099$

$N_{LOO-CV} = 23$, $Q^2_{LOO-CV} = 0.888$, $RMSE_{LOO-CV} = 0.169$, $s = 0.203$, $F \text{ ratio} = 21.122$, $PRESS = 0.659$

$N_{Ext} = 5$, $Q^2_{Ext} = 0.875$, $RMSE_{Ext} = 0.170$

Six important descriptors obtained from the Mold² software were used to generate the model 2 as shown in Eq. (3). Statistical analysis of the model is given in Table 4. It was observed that the model provided statistical $R^2_{Tr} = 0.958$, $RMSE_{Tr} = 0.099$ for the training set, $Q^2_{LOO-CV} = 0.888$, $RMSE_{LOO-CV} = 0.169$ for the LOO-CV set, and $Q^2_{Ext} = 0.875$, $RMSE_{Ext} = 0.170$ for the external set. The experimental versus the predicted antioxidant activities of the model 2 were graphically plotted (Figure 3c), while Figure 3d was the plot of residual values.

Model 3 (PaDEL descriptors)

$$\begin{aligned} \text{pIC}_{50} = & 0.2638(\text{SHBint4}) \\ & + 6.9792(\text{SpMin2_Bhe}) \\ & - 0.9768(\text{MATS8e}) \\ & + 0.3238(\text{SssCH2}) \\ & - 10.0346 \end{aligned} \quad (4)$$

$N_{Tr} = 23$, $R^2_{Tr} = 0.950$, $R^2_{Adj} = 0.942$, $RMSE_{Tr} = 0.109$

$N_{LOO-CV} = 23$, $Q^2_{LOO-CV} = 0.908$, $RMSE_{LOO-CV} = 0.150$, $s = 0.170$, $F \text{ ratio} = 44.52$, $PRESS = 0.518$

$N_{Ext} = 5$, $Q^2_{Ext} = 0.885$, $RMSE_{Ext} = 0.166$

Model 3 of Eq. (4) was built using 4 significant descriptors generated by PaDEL software. Statistical evaluation of the model is shown in Table 4. The results showed that the QSAR model displayed $R^2_{Tr} = 0.950$, $RMSE_{Tr} = 0.109$ for the training set, $Q^2_{LOO-CV} = 0.908$, $RMSE_{LOO-CV} = 0.150$ for the LOO-CV set, and $Q^2_{Ext} = 0.885$, $RMSE_{Ext} = 0.166$ for the external set. Comparison of experimental and the predicted antioxidant activities of model 3 are outlined in Figure 3e, while residual values were displayed in Figure 3f.

In overview, three constructed models provided satisfactory results as indicated by their statistical parameters such as R^2 , Q^2 , $RMSE$, $F \text{ ratio}$ and $PRESS$ values. The R^2 and Q^2 of the obtained QSAR models were considered as acceptable values when $R^2 > 0.6$ and $Q^2 > 0.5$ (Golbraikh and Tropsha, 2002; Nantasenamat et al., 2010). These parameters of all constructed models were in acceptable range (model 1: $R^2_{Tr} = 0.901$, $Q^2_{LOO-CV} = 0.813$ and $Q^2_{Ext} = 0.952$, model 2: $R^2_{Tr} = 0.958$, $Q^2_{LOO-CV} = 0.888$ and $Q^2_{Ext} = 0.875$, model 3: $R^2_{Tr} = 0.945$, $Q^2_{LOO-CV} = 0.908$ and $Q^2_{Ext} = 0.885$). In addition, low values of $RMSE$, s and $PRESS$, but high value of $F \text{ ratio}$ indicated that the models were significant ($RMSE_{Tr} = 0.154$, $RMSE_{LOO-CV} = 0.210$, $RMSE_{Ext} = 0.104$, $s = 0.238$, $PRESS = 1.019$ and $F \text{ ratio} = 13.043$ in model 1, $RMSE_{Tr} = 0.099$, $RMSE_{LOO-CV} = 0.169$, $RMSE_{Ext} = 0.170$, $s = 0.203$, $PRESS = 0.659$ and $F \text{ ratio} = 21.122$ in model 2, and $RMSE_{Tr} = 0.109$, $RMSE_{LOO-CV} = 0.151$, $RMSE_{Ext} = 0.164$, $s = 0.170$, $PRESS = 0.518$ and $F \text{ ratio} = 44.52$ in model 3) (Frimayanti et al., 2011; Rastija et al., 2018). The statistical (Table 4) and graphical (Figure 3) results showed that the QSAR models (models 1-3) gave a reliable agreement of the experimental and the predicted antioxidant values. Furthermore, the plots of experimental activity and residual values (Figures 3b, 3d and 3f) displayed the distribution of residuals on both sides of the zero values indicating that there are no systemic error in the models (Jalali-Heravi and Kyani, 2004). Therefore, the QSAR models 1-3 could be possibly used and reliable for predicting the antioxidant activity of coumarin derivatives. Considering the correlation coefficient (Q^2) of external set, it was shown that the Dragon descriptors gave the highest quality of the prediction for external test set (model 1: $Q^2_{Ext} = 0.952$) followed by the PaDEL descriptors (model 3: $Q^2_{Ext} = 0.885$) and the Mold² descriptors (model 2: $Q^2_{Ext} = 0.875$).

Table 4: Summary of statistical results in predicting antioxidant activity of coumarin derivatives (1-28)

Model	Training set			LOO-CV set		External set	
	R^2_{Tr}	$RMSE_{Tr}$	Q^2_{LOO-CV}	$RMSE_{LOO-CV}$	F	Q^2_{Ext}	$RMSE_{Ext}$
1	0.901	0.154	0.813	0.210	19.564	0.952	0.104
2	0.958	0.099	0.888	0.169	21.122	0.875	0.170
3	0.950	0.109	0.908	0.150	44.520	0.885	0.166

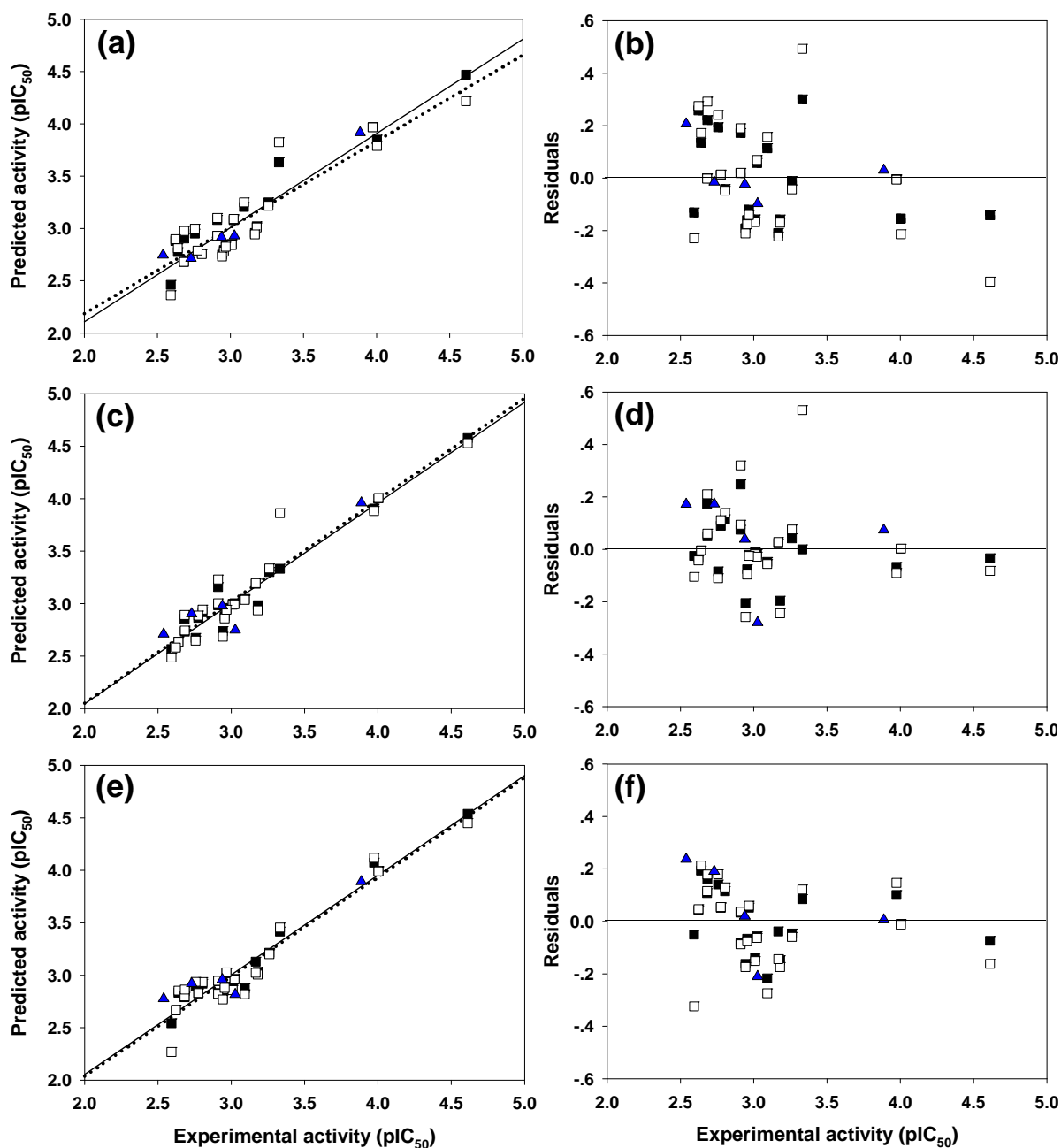


Figure 3: Plots of the experimental and the predicted activities of model 1 (3a), model 2 (3c), and model 3 (3e) for the training set (\square ; regression line is represented as solid line), the leave-one out cross-validated (\blacksquare ; regression line is represented as dotted line), and the external (\blacktriangle) sets. Distribution of the experimental activity and residual values (the difference between experimental and predicted activities) of model 1 (3b), model 2 (3d) and model 3 (3f)

Structure-activity relationship (SAR)

Regression coefficient values of the key descriptors (as independent variables) in QSAR models define the degree or weight of their influence on dependent variables (antioxidant activity: pIC₅₀ values). According to the linear QSAR equations, high values of descriptors with positive regression coefficient and low values of descriptors with negative regression coefficient are required for potent antioxidant activity (high pIC₅₀ values).

The descriptors in model 1 including nArNHR, B04[O-O] and G2p displayed positive values of regression coefficient involved in the increased antioxidant activity as a positive effect, while ISH descriptor having negative value of regression coefficient involved in the decreased activity as a negative effect. The important descriptors are ranked according to their regression coefficient values as G2p>nArNHR>B04[O-O]>ISH with corresponding values of 8.1471, 1.2245, 0.3944, and -4.3187, respectively. The descriptors in model 2 including D467 and D278 with positive regression coefficient values displayed the positive effect, whereas D491, D384, D580 and D461 descriptors exerted the negative effect on the bioactivity. The order of descriptors are D278>D467>D384>D461>D580>D491 with corresponding values of 7.1429, 1.4575, -0.1328, -0.5180, -2.0251, and -2.6745, respectively. For model 3, SHBint4, SpMin2_Bhe and SssCH2 descriptors showed the positive effect on activity, but MATS8e descriptor displayed the negative effect. The order of important descriptors are SpMin2_Bhe>SssCH2>SHBint4>MATS8e with values of 6.9792, 0.3238, 0.2638 and -0.9768, respectively.

To gain insights into SAR, coumarin derivatives (**1-28**, Figure 1) are categorized into 3 groups according to their core structures (i.e., thiazole group I (**1-9** and **18**), sulfonamides group II (**10-17**) and amides group III (**19-28**) for effective SAR analysis. Thiazoles group I (**1-9** and **18**) showed antioxidant activity (Table 1) with pIC₅₀ range of 2.539-4.612. The most potent and the least potent compounds of benzothiazoles group I were **5**

(pIC₅₀ = 4.612) and **6** (pIC₅₀ = 2.539), respectively. Among group II compounds (**10-17**), compound **11** was the most active (pIC₅₀ = 3.180), and **14** was the least active compound. For group III of amides **19-28**, compound **21** displayed the most potent activity (pIC₅₀ = 3.027) and compound **19** exhibited the lowest activity with pIC₅₀ of 2.640.

It should be noted that coumarins substituted by thiazole at 3-position displayed better activity when compared with those substituted by sulfonamides and amides at 7- or 4- or 6- position (group II and III). The antioxidant activity (Table 1) is shown as the following trend: group I **5**>**3**>**4**>**1**>**2**>**18**>**8**>**7**>**9**>**6**; group II **11**>**16**>**13**>**17**>**15**>**12**>**10**>**14**; group III **21**>**28**>**22**>**20**>**23**>**26**>**27**>**24**>**25**>**19**.

According to the significant descriptors in models 1-3, secondary (*sec*-) amine, polarizability, electronegativity and H-bond displayed positive effect in the antioxidant activity. This is noted in the most potent coumarin **5** bearing *sec*-amine (part of aromatic thiazole), and 7-OH group (on the coumarin ring) with H-bond and polarizability properties. On the other hand, tertiary (*tert*-) amine **6** without 7-OH group exerted the lowest activity among the coumarin derivatives **1-28**. This could be implied that the *sec*-amine (-NH-) and OH as H-bond and polarizing group are important for the better activity.

In a series of compounds **1-28**, only the *sec*-amines (**1-5**) had nArNHR = 1. The most potent compound **5** had higher values of nArNHR = 1, SHBint4 = 7.5875, G2p = 0.193, D467 = 0.805, but lower values of van der Waals volume (D491 = 0.471 and D461 = 0.186) when compared with the least potent compound **6** (nArNHR = 0, SHBint4 = 0.000, G2p = 0.169, D467 = 0.646, van der Waals volume: D491 = 0.805 and D461 = 0.520). Apparently, the *tert*-amines (**6-8**) and other amides, sulfonamides (**10-28**) as well as non-aromatic *sec*-amine **9** had nArNHR = 0.

Because each descriptor represents certain characteristic/property of the compound, QSAR equations are useful to guide effective rational molecular design of novel bioactive compounds with preferable activity (De et al.,

2017; Mitra et al., 2011; Prachayasittikul et al., 2017; Pratiwi et al., 2019). To improve the antioxidant activity of these derivatives (**1-28**), 8 compounds (i.e., **2, 3, 4, 5, 11, 13, 18** and **21**) were selected as parent compounds for structural modification. A series of novel analogs were rationally designed based on key influencing properties (which were revealed by descriptors presented in models 1-3). The modification was conducted by substitution of diverse types of functional groups with electronegativity, polarizability and H-bond properties (i.e., electron donating and electron withdrawing groups such as OH, NH₂, SH, OCH₃, CN, CF₃ and halogen) on various positions of the coumarin ring. The new structurally modified compounds are shown in Supplementary Table 4). Values of molecular descriptors of these compounds are provided in Supplementary Tables 5-7). The prediction showed that the most potent compounds (pIC₅₀) of each modified series were **2b, 2e** (4.503, 4.502), **3n** (6.340), **4g** (6.445), **5h** (7.016), **11a** (4.526), **13d** (4.356), **18c** (4.197) and **21d** (5.634) as shown in Figure 4. All highly potent modified compounds were ranked as **5h>4g>3n>21d>11a>2b=2e>13d>18c**. The top three modified compounds (ranked as **5h>4g>3n**) shared a common feature of 4-amino coumarin moiety

(Figure 4). This moiety involves in H-bonding, polarizability, and electronegativity properties of the compounds which are essential for potent predicted activity. Apparently, these most potent compounds displayed the highest SHBint4 values of 16.903, 13.660, and 13.064 for **5h, 4g**, and **3n**, respectively.

The following discussed compounds are shown in Supplementary Table 4. Compound **2** (pIC₅₀ = 3.332) was modified by substitution of R1 (6-position) and R2 (5-position) groups (H, F, Cl, CF₃, CN, NO₂) on the core structure to give a new series of compounds **2a-2h** with improved activity (pIC₅₀ = 3.167-4.503), except for compounds **2f** and **2h**. Compounds **2b** (5-CF₃, 6-H) and **2e** (5-CF₃, 6-NO₂) were the most potent compounds with comparable pIC₅₀ values of 4.503 and 4.502, respectively. These could be due to the enhancing effect of 5-CF₃ (R2) and 6-NO₂ (R1) groups that provided high polarizability (G2p: **2b** = 0.1820, **2e** = 0.1730), electronegativity (SpMin2_Bhe: **2b** = 1.847, **2e** = 1.847), and potential H-bond (SHBint4: **2b** = 5.150, **2e** = 5.480), thus, improved antioxidant activity of the modified compounds comparing with their parent **2** (G2p = 0.164, SpMin2_Bhe = 1.8175, SHBint4 = 4.0816).

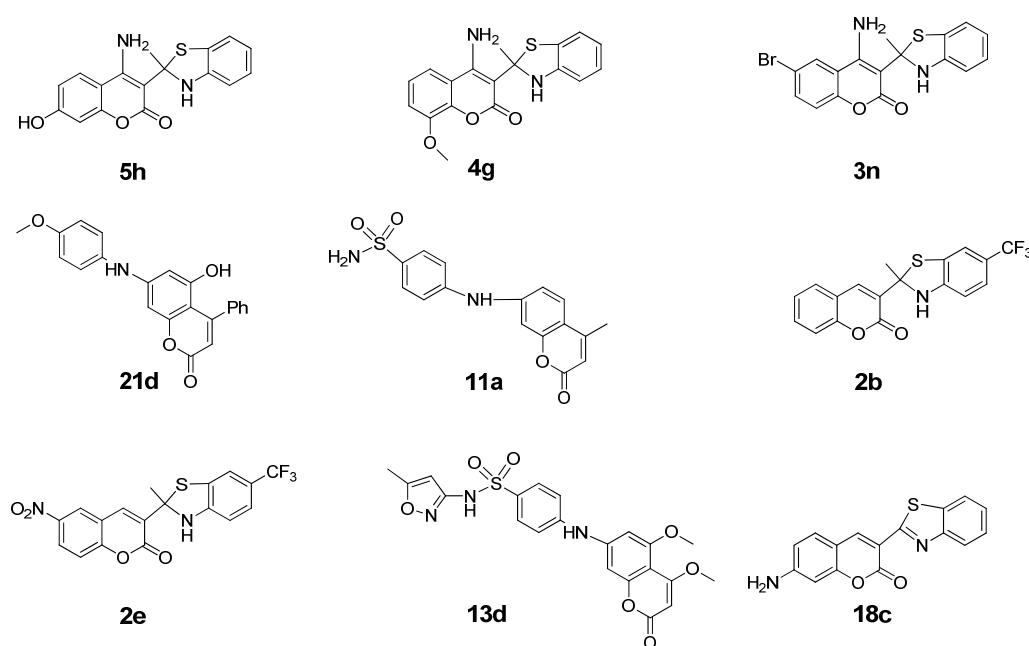


Figure 4: The most potent compounds in each rationally modified series

Similarly, modified compounds **3a-3p** were obtained by substitution of R1, R2, R3 at positions 7, 6, and 4 of the parent compound **3** (R1 = H, OH, OCH₃, NH₂, N(CH₃)₂, SH, SC₆H₅; R2 = F, Cl, Br, CF₃; R3 = H, OH, OCH₃, NH₂, N(CH₃)₂). Compound **3n** (R1 = H, R2 = Br, R3 = NH₂) was predicted as the most potent one with pIC₅₀ value of 6.340. Compounds **3f-3p** (pIC₅₀ = 4.129-6.340) were mostly active than the parent compound **3** (pIC₅₀ = 4.003). It should be noted that these compounds (**3f-3p**) contained the same type of substituted R2 (Br) group as their parent **3** (6-Br). Therefore, higher H-bonding potential (SHBint4 = 13.064) and electronegativity (D467 = 0.736) play an important role in improving the activity of compound **3n** when compared with its parent compound (**3**: SHBint4 = 4.2075, D467 = 0.722). The higher SHBint4 and D467 of **3n** may result from the property of the substituted NH₂ group in forming H-bond with the interacting species. In addition, the substitution with high electronegativity Br group (R2) at position 6 is required for improving activity of the compounds as noted for compound **3f-3p** and the parent compound **3**.

Structural modification improved activity of most of the compounds in series **4** (**4a-4i**: pIC₅₀ = 4.033-6.445), except for compounds **4d** and **4h**. The results showed that **4g** (R1 = 7-H, R2 = 4-NH₂) was the most potent compound (pIC₅₀ = 6.445). Comparing with the parent compound **4** (pIC₅₀ = 3.973, SHBint4 = 4.442, G2p = 0.174), the most potent derivative **4g** displayed higher potential H-bond (SHBint4 = 13.660) and higher polarizability (G2p = 1.960). It is reasonable to explain that 4-NH₂ (R2) of **4g** participates in H-bond forming and polarization through a keto group of the coumarin ring as shown by the resonant ionic form (**4A**) in Figure 5. On the other hand, compound **4c** (R1 = 7-NH₂, R2 = 4-H) exerted lower activity with lower descriptors values (pIC₅₀ = 0.4772, SHBint4 = 7.432, G2p = 0.1830) comparing with the most potent compound **4g**.

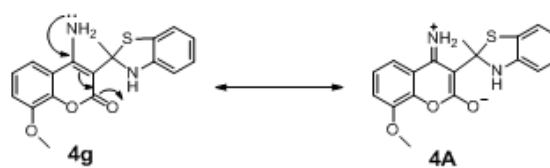


Figure 5: The resonant ionic form of **4A**

The most potent coumarin **5** was modified to give derivatives **5a-5l**, in which compounds **5f-5i** displayed the improved antioxidant activity (pIC₅₀ = 4.745-7.016). Compound **5h** was the most potent one (pIC₅₀ = 7.016) with higher H-bond potential (SHBint4 = 16.903), but lower electronegativity (MATS8e = 0.063) when compared with its parent **5** (pIC₅₀ = 4.612, SHBint4 = 7.5875, MATS8e = 0.0848).

The amides **11**, **13**, and **21** were structurally modified by removing oxyketo group from the parent core structures to give *sec*-amines. The amine derivatives **11a-11f** with improved activity (pIC₅₀ = 3.402-4.526) were achieved from compound **11** (pIC₅₀ = 3.180). When compared with its parent (amide **11**: nArNHR = 0, D467 = 0.479, SHBint4 = 0.000, D491 = 0.688, D461 = 0.544 and MATS8e = 0.1872), improved activity of the most potent amine **11a** (pIC₅₀ = 4.526) was governed by its higher nArNHR = 1, electronegativity (D467 = 0.690), and H-bond (SHBint4 = 2.048), but lower van der Waals volume (D491 = 0.342 and D461 = 0.022) and electronegativity (MATS8e = -0.058). This could suggest that the improved activity of compounds requires a smaller size *sec*-amine (**11a**) with lower van der Waals volume but higher potential H-bonding. Similarly, a series of *sec*-amines **13a-13d** (R1 = CH₃, OH, OCH₃, R2 = CH₃, OCH₃) were obtained (pIC₅₀ = 3.341-4.356) from the amide **13** (pIC₅₀ = 3.022). Compound **13d** (pIC₅₀ = 4.356) was the most potent amine with higher nArNHR = 1, D467 = 0.640, SHBint4 = 4.368, but lower van der Waals volume D461 = 0.455 and MATS8e = -0.125 comparing to its parent compound **13** (nArNHR = 0, D467 = 0.503, SHBint4 = 0.000, D461 = 0.515 and MATS8e = 0.1400). The amide **21** (pIC₅₀ =

3.027) was modified to give analogs with improved antioxidant effect i.e., *sec*-amines analogs **21a-21h** (pIC_{50} = 3.549-5.634). The most potent modified amine **21d** (pIC_{50} = 5.634) displayed higher $nArNHR$ = 1, electronegativity D_{467} = 0.657, H-bond SHB_{int4} = 9.823, electronegativity $SpMn2-Bhe$ = 1.843 and $SssCH_2$ = 0, but lower ISH = 0.869, van der Waals volume D_{491} = 0.350 and D_{461} = 0.436, and $MATS_{8e}$ = -0.219 when compared with the parent **21** ($nArNHR$ = 0, D_{467} = 0.489, SHB_{int4} = 0.000, $SpMin2-Bhe$ = 1.8395, $SssCH_2$ = -0.1139, ISH = 0.923, van der Waals volume D_{491} = 0.82 and D_{461} = 0.612, and $MATS_{8e}$ = -0.0511).

Benzothiazole coumarin **18** was structurally modified to provide compounds **18a-18f** (R_1 = OH, OCH₃, NH₂, N(CH₃)₂, SH, SC₆H₅ at position 7). Comparing with the parent compound **18** (pIC_{50} = 3.260), compounds **18b-18e** were more potent antioxidants (pIC_{50} = 3.758-4.197). Improved activity of the most potent compound **18c** (pIC_{50} = 4.197) was governed by higher electronegativity (D_{467} = 0.595), total information content (D_{278} = 0.473) and H-bond (SHB_{int4} = 2.894), but lower van der Waals volume (D_{491} = 0.377 and D_{461} = 0.299) when compared to parent compound **18** (D_{467} = 0.501, D_{278} = 0.413, SHB_{int4} = 0.000, D_{491} = 0.472 and D_{461} = 0.471).

It should be noted that the most potent modified compounds had higher values of H-bonding descriptor (SHB_{int4} = 2.048-16.903, Supplementary Table 7) when compared with their parent compounds (SHB_{int4} = 0.000-7.5875, Table 3). Thus, SHB_{int4} might be the important descriptor in governing the potent antioxidant activity.

CONCLUSION

Understanding SAR is important for improving bioactivities and pharmacokinetic properties in development of potent and safe cosmetic products. Herein, a set of coumarin derivatives (**1-28**) with antioxidant activity was used to construct three QSAR models (1-3) using three different descriptor types and MLR method. Results of statistical evaluation

showed that three generated QSAR models provide good reliability and comparable predictive performance (Q^2_{LOO-CV} = 0.813-0.908; $RMSE_{LOO-CV}$ = 0.150-0.210; Q^2_{Ext} = 0.875-0.952; $RMSE_{Ext}$ = 0.104-0.166). In addition, good correlation obtained from model prediction suggests that the selected significant descriptors were shown to be good representatives for revealing correlation between chemical structures of the compounds (i.e., $nArNHR$, H-bonding, polarizability, van der Waals volume and electronegativity properties) and their antioxidant activities. An application of the constructed models was demonstrated by rationally designed an additional set of 69 structurally modified coumarins based on key descriptors, in which their antioxidant activities were predicted using the obtained QSAR models (1-3). Most of the rationally designed compounds displayed more improved antioxidant activity when compared with their parents. Particularly, the top three newly designed compounds (**5h**, **4g** and **3n**) showing high H-bonding (SHB_{int4}) descriptor values which may play part in governing the most improved antioxidant activity. Finally, a set of newly designed promising coumarin analogs were highlighted for their potential to be further developed as potent antioxidants. Insights SAR findings also provided beneficial guidelines for the rational design of novel coumarin-based compounds with potent antioxidant effect for cosmetic applications.

Supplementary information

Supplementary information is available on the EXCLI Journal website.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Thailand Research Fund (Grant No. MRG6180053) and the annual budget grant from Mahidol University (B.E. 2562-2563).

REFERENCES

- Alisi IO, Uzairu A, Abechi SE, Idris SO. Evaluation of the antioxidant properties of curcumin derivatives by genetic function algorithm. *J Adv Res.* 2018;12:47-54.
- Alyar S, Karacan N. Synthesis, characterization, antimicrobial activity and structure-activity relationships of new aryldisulfonamides. *J Enzyme Inhib Med Chem.* 2009;24:986-92.
- Arshad A, Osman H, Bagley MC, Lam CK, Mohamad S, Zahariluddin AS. Synthesis and antimicrobial properties of some new thiazolyl coumarin derivatives. *Eur J Med Chem.* 2011;46:3788-94.
- Basic J, Kalinic M, Ivkovic B, Eric S, Milenkovic M, Vladimirov S, et al. Synthesis, QSAR analysis and mechanism of antibacterial activity of simple 2'-hydroxy chalcones. *Dig J Nanomater Biostruct.* 2014;9:1537-46.
- Bogdan Allemann I, Baumann L. Antioxidants used in skin care formulations. *Skin Therapy Lett.* 2008;13:5-9.
- De B, Adhikari I, Nandy A, Saha A, Goswami BB. *In silico* modelling of thiazolidine derivatives with antioxidant potency: Models quantify the degree of contribution of molecular fragments towards the free radical scavenging ability. *J Mol Struct.* 2017;1138:17-26.
- Dennington II R, Keith T, Millam J, Eppinnett K, Hovell WL, Gilliland R. *GaussView, Version 3.09.* Shawnee Mission, KS: Semichem Inc., 2003.
- Duchowicz PR, Bacelo D, Fioressi S, Palermo V, Ibezim NE, Romanelli GP. QSAR studies of indoyl aryl sulfides and sulfones as reverse transcriptase inhibitors. *Med Chem Res.* 2018;27:420-8.
- Elansary HO, Szopa A, Kubica P, Ekiert H, Ali HM, Elshikh MS, et al. Bioactivities of traditional medicinal plants in Alexandria. *Evid Based Complement Altern Med.* 2018;2018:1-13.
- Erzincan P, Saçan MT, Yüce-Dursun B, Daniş Ö, Demir S, Erdem SS, et al. QSAR models for antioxidant activity of new coumarin derivatives. *SAR QSAR Environ Res.* 2015;26:721-37.
- Frimayanti N, Yam ML, Lee HB, Othman R, Zain SM, Rahman NA. Validation of quantitative structure-activity relationship (QSAR) model for photosensitizer activity prediction. *Int J Mol Sci.* 2011;12:8626-44.
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. *Gaussian 09, Revision A.02.* Wallingford, CT: Gaussian Inc., 2009.
- Gao H. Predicting tyrosinase inhibition by 3D QSAR pharmacophore models and designing potential tyrosinase inhibitors from Traditional Chinese medicine database. *Phytomedicine.* 2018;38:145-57.
- Golbraikh A, Tropsha A. Beware of q^2 !. *J Mol Graph Model.* 2002;20:269-76.
- Hong H, Xie Q, Ge W, Qian F, Fang H, Shi L, et al. Mold(2), molecular descriptors from 2D structures for chemoinformatics and toxicoinformatics. *J Chem Inf Model.* 2008;48:1337-44.
- Jalali-Heravi M, Kyani A. Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds: a PCA-MLR-ANN approach. *J Chem Inf Comput Sci.* 2004;44:1328-35.
- Khan MT. Novel tyrosinase inhibitors from natural resources - their computational studies. *Curr Med Chem.* 2012;19:2262-72.
- Khoobi M, Emami S, Dehghan G, Foroumadi A, Ramazani A, Shafiee A. Synthesis and free radical scavenging activity of coumarin derivatives containing a 2-methylbenzothiazoline motif. *Arch Pharm.* 2011;344:588-94.
- Kostova I, Bhatia S, Grigorov P, Balkansky S, Parmar VS, Prasad AK, et al. Coumarins as antioxidants. *Curr Med Chem.* 2011;18:3929-51.
- Krishnaiah D, Sarbatly R, Nithyanandam R. A review of the antioxidant potential of medicinal plant species-Food Bioprod Process. 2011;89:217-33.
- Kusumawati I, Indrayanto G. Natural antioxidants in cosmetics. *Stud Nat Prod Chem.* 2013;40:485-505.
- Lee BC, Lee SY, Lee HJ, Sim GS, Kim JH, Kim JH, et al. Anti-oxidative and photo-protective effects of coumarins isolated from *Fraxinus chinensis*. *Arch Pharm Res.* 2007;30:1293-301.
- Lupo MP. Antioxidants and vitamins in cosmetics. *Clin Dermatol.* 2001;19:467-73.
- Ma Q, Xi H, Ma H, Meng X, Wang Z, Bai H, et al. Simultaneous separation and determination of 22 coumarin derivatives in cosmetics by UPLC-MS/MS. *Chromatographia.* 2015;78:241-9.
- Mitra I, Saha A, Roy K. Chemometric QSAR modeling and in silico design of antioxidant NO donor phenols. *Sci Pharm.* 2011;79:31-57.
- Nantasenamat C, Isarankura-Na-Ayudhya C, Prachayasittikul V. Advances in computational methods to predict the biological activity of compounds. *Expert Opin Drug Discov.* 2010;5:633-54.

- Nantasenamat C, Prachayasittikul V. Maximizing computational tools for successful drug discovery. *Expert Opin Drug Discov.* 2015;10:321-9.
- Nasr T, Bondock S, Youns M. Anticancer activity of new coumarin substituted hydrazide-hydrazone derivatives. *Eur J Med Chem.* 2014;76:539-48.
- Podunavac-Kuzmanović SO, Cvetković DD, Barna DJ. QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against *Pseudomonas aeruginosa*. *Int J Mol Sci.* 2009;10:1670-82.
- Prachayasittikul S, Buraparungsang P, Worachartcheewan A, Isarankura-Na-Ayudhya C, Ruchirawat S, Prachayasittikul V. Antimicrobial and antioxidative activities of bioactive constituents from *Hydnophytum formicarum* Jack. *Molecules.* 2008;13:904-21.
- Prachayasittikul S, Worachartcheewan A, Lawung R, Ruchirawat S, Prachayasittikul V. Activities of thiotetrahydropyridines as antioxidant and antimicrobial agents. *EXCLI J.* 2009a;8:107-14.
- Prachayasittikul S, Suphapong S, Worachartcheewan A, Lawung R, Ruchirawat S, Prachayasittikul V. Bioactive metabolites from *Spilanthes acmella* Murr. *Molecules.* 2009b;14:850-67.
- Prachayasittikul V, Prachayasittikul S, Ruchirawat S, Prachayasittikul V. High therapeutic potential of *Spilanthes acmella*: a review. *EXCLI J.* 2013;12:291-312.
- Prachayasittikul V, Pingaew R, Worachartcheewan A, Nantasenamat C, Prachayasittikul S, Ruchirawat S, et al. Synthesis, anticancer activity and QSAR study of 1,4-naphthoquinone derivatives. *Eur J Med Chem.* 2014;84:247-63.
- Prachayasittikul V, Worachartcheewan A, Shoombuatong W, Songtawee N, Simeon S, Prachayasittikul V, et al. Computer-aided drug design of bioactive natural products. *Curr Top Med Chem.* 2015a;15:1780-800.
- Prachayasittikul V, Pingaew R, Anuwongcharoen N, Worachartcheewan A, Nantasenamat C, Prachayasittikul S, et al. Discovery of novel 1,2,3-triazole derivatives as anticancer agents using QSAR and *in silico* structural modification. *Springerplus.* 2015b;4:571.
- Prachayasittikul V, Pingaew R, Worachartcheewan A, Sitthimonchai S, Nantasenamat C, Prachayasittikul S, et al. Aromatase inhibitory activity of 1,4-naphthoquinone derivatives and QSAR study. *EXCLI J.* 2017;16:714-26.
- Pratiwi R, Prachayasittikul V, Prachayasittikul S, Nantasenamat C. Rational design of novel sirtuin 1 activators via structure-activity insights from application of QSAR modeling. *EXCLI J.* 2019;18:207-22.
- Rastija V, Molnar M, Siladi T, Masand VH. QSAR analysis for antioxidant activity of dipicolinic acid derivatives. *Comb Chem High Throughput Screen.* 2018;21:204-14.
- Rodríguez-Hernández KD, Martínez I, Agredano-Moreno LT, Jiménez-García LF, Reyes-Chilpa R, Espinoza B. Coumarins isolated from *Calophyllum brasiliense* produce ultrastructural alterations and affect in vitro infectivity of *Trypanosoma cruzi*. *Phyto-medicine.* 2019;61:152827.
- Saavedra LM, Romanelli GP, Duchowicz PR. Quantitative structure-activity relationship (QSAR) analysis of plant-derived compounds with larvicidal activity against Zika *Aedes aegypti* (Diptera: Culicidae) vector using freely available descriptors. *Pest Manag Sci.* 2018;74:1608-15.
- Saeedi M, Goli F, Mahdavi M, Dehghan G, Faramarzi MA, Foroumadi A, et al. Synthesis and biological investigation of some novel sulfonamide and amide derivatives containing coumarin moieties. *Iran J Pharm Res.* 2014;13:881-92.
- Saghaie L, Shahlaei M, Fassihi A. Quantitative structure activities relationships of some 2-mercaptoimidazoles as CCR2 inhibitors using genetic algorithm-artificial neural networks. *Res Pharm Sci.* 2013;8:97-112.
- Saleem H, Zengin G, Locatelli M, Mollica A, Ahmad I, Mahomoodally FM, et al. *In vitro* biological propensities and chemical profiling of *Euphorbia milii* Des Moul (Euphorbiaceae): a novel source for bioactive agents. *Ind Crops Prod.* 2019;130:9-15.
- Sławiński J, Szafranski K, Pogorzelska A, Żołnowska B, Kawiak A, Macur K, et al. Novel 2-benzylthio-5-(1,3,4-oxadiazol-2-yl)benzenesulfonamides with anticancer activity: Synthesis, QSAR study, and metabolic stability. *Eur J Med Chem.* 2017;132:236-48.
- Stiefel C, Schubert T, Morlock GE. Bioprofiling of cosmetics with focus on streamlined coumarin analysis. *ACS Omega.* 2017;2:5242-50.
- Subramanyam C, Nayab Rasool S, Janakiramudu DB, Rasheed S, Uday Sankar A, Naga Raju C. Synthesis and bioactivity evaluation of some novel sulfonamide derivatives. *Phosphorus Sulfur Silicon Relat Elem.* 2017;192:845-9.
- Suvannang N, Preeyanon L, Malik AA, Schaduangrat N, Shoombuatong W, Worachartcheewan A, et al. Probing the origin of estrogen receptor alpha inhibition via large-scale QSAR study. *RSC Adv.* 2018;8:11344-56.
- Talete srl. DRAGON for Windows, Software for molecular descriptor calculations, version 5.5. Milano, Italy, 2007.

- Valko M, Leibfritz D, Moncol J, Cronin MT, Mazur M, Telser J. Free radicals and antioxidants in normal physiological functions and human disease. *Int J Biochem Cell Biol.* 2007;39:44-84.
- Venditti A, Serrilli AM, Di Cecco M, Ciaschetti G, Bianco A. Coumarins and other components of *Daphne oleoides* Schreb. subsp. *oleoides* from Majella National Park. *Biochem Syst Ecol.* 2019;83:39-46.
- Winyard PG, Moody CJ, Jacob C. Oxidative activation of antioxidant defence. *Trends Biochem Sci.* 2005;30:453-61.
- Witaicenis A, Seito LN, da Silveira Chagas A, de Almeida LD Jr, Luchini AC, Rodrigues-Orsi P, et al. Antioxidant and intestinal anti-inflammatory effects of plant-derived coumarin derivatives. *Phytomedicine.* 2014;21:240-6.
- Witten IH, Frank E, Hall MA. *Data mining: practical machine learning tools and techniques.* San Francisco, CA: Morgan Kaufmann, 2011.
- Wongsawatkul O, Prachayasittikul S, Isarankura-Na-Ayudhya C, Satayavivad J, Ruchirawat S, Prachayasittikul V. Vasorelaxant and antioxidant activities of *Spilanthes acmella* Murr. *Int J Mol Sci.* 2008;9:2724-44.
- Worachartcheewan A, Prachayasittikul S, Pingaew R, Nantasenamat C, Tantimongcolwat T, Ruchirawat S, et al. Antioxidant, cytotoxicity, and QSAR study of 1-adamantylthio derivatives of 3-picoline and phenylpyridines. *Med Chem Res.* 2012;21:3514-22.
- Worachartcheewan A, Nantasenamat C, Isarankura-Na-Ayudhya C, Prachayasittikul V. QSAR study of amidino bis-benzimidazole derivatives as potent anti-malarial agents against *Plasmodium falciparum*. *Chem Pap.* 2013;67:1462-73.
- Worachartcheewan A, Nantasenamat C, Owasirikul W, Monnor T, Naruepantawart O, Janyapaisarn S, et al. Insights into antioxidant activity of 1-adamantylthiopyridine analogs using multiple linear regression. *Eur J Med Chem.* 2014;73:258-64.
- Worachartcheewan A, Songtawee N, Siriwong S, Prachayasittikul S, Nantasenamat C, Prachayasittikul V. Rational design of colchicine derivatives as anti-HIV agents via QSAR and molecular docking. *Med Chem.* 2019;15:328-40.
- Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem.* 2011;32:1466-74.