Med Chem Res
DOI 10.1007/s00044-014-1260-1

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Probing the origins of anticancer activity of chrysin derivatives

Apilak Worachartcheewan · Chanin Nantasenamat · Chartchalerm Isarankura-Na-Ayudhya · Virapong Prachayasittikul

Received: 20 June 2014 / Accepted: 11 September 2014

Abstract Chrysin is a derivative of flavonoid, a natural product commonly found in plants. It has been shown to afford a wide variety of pharmacological activities particularly anticancer properties. In this study, 21 chrysin derivatives with anticancer activities against human gastric adenocarcinoma (SGC-7901) and human colorectal adenocarcinoma (HT-29) cell lines were employed for quantitative structure–activity relationship (QSAR) investigation. Molecular structures were geometrically optimized at the B3LYP/6-311++g(d,p) level and their quantum chemical and molecular properties were obtained from Gaussian 09 and Dragon softwares, respectively. Significant descriptors for modeling the anticancer activities of SGC-7901 (i.e., SIC2, Mor11e, P2p, HTp, and R5e+) and HT-29 (i.e., L/Bw, BIC2, and Mor19p) cell lines were deduced from stepwise multiple linear regression (MLR) method. QSAR models were constructed using MLR and their predictivities were verified via internal (i.e., leave one-out cross-validation; LOO-CV) and external sets. The predictive performance was evaluated from their squared correlation coefficients (R^2 and Q^2) and root mean square error (RMSE). Results indicated good correlation between experimental and predicted anticancer activities as deduced from statistical parameters of internal and external sets as follows: $R^2_{Tr} = 0.8778$, $RMSE_{Tr} = 0.0854$, $Q^2_{CV} = 0.7315$, $RMSE_{CV} = 0.1375$, $Q^2_{Ext} = 0.7324$, and $RMSE_{Ext} = 0.1168$ for QSAR models of SGC-7901 while $R^2_{Tr} = 0.8201$, $RMSE_{Tr} = 0.1293$, $Q^2_{CV} = 0.6829$, $RMSE_{CV} = 0.1735$, $Q^2_{Ext} = 0.8486$, and $RMSE_{Ext} = 0.1179$ for QSAR models of HT-29. Furthermore, the obtained QSAR models provided pertinent insights on the structure–activity relationship of investigated compounds where molecular properties such as shape, electronegativities and polarizabilities were crucial for anticancer activity. The knowledge gained from the constructed QSAR models could serve as guidelines for the rational design of novel chrysin derivatives with potent anticancer activity.

Keywords Chrysin · Cytotoxicity · QSAR · Multiple linear regression · Data mining

Introduction

The discovery and development of novel drugs is a challenging and complex endeavor owing to several inherent difficulties including the feasibility of its synthesis as well as inherent flaws in its absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties (Drews, 2007). Nature had always been a good source of inspiration (Newman and Cragg, 2007) owing to the abundance of bioactive components found in these natural products (Cragg *et al.*, 2009). Flavonoids, a naturally occurring polyphenolic compound found in plants, are essential components of natural products that had been found to exert a wide variety of pharmacological activities such as

Electronic supplementary material The online version of this article (doi:10.1007/s00044-014-1260-1) contains supplementary material, which is available to authorized users.

A. Worachartcheewan · C. Nantasenamat (✉)
Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
e-mail: chanin.nan@mahidol.ac.th

C. Nantasenamat · C. Isarankura-Na-Ayudhya · V. Prachayasittikul (✉)
Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
e-mail: virapong.pra@mahidol.ac.th

Published online: 24 September 2014

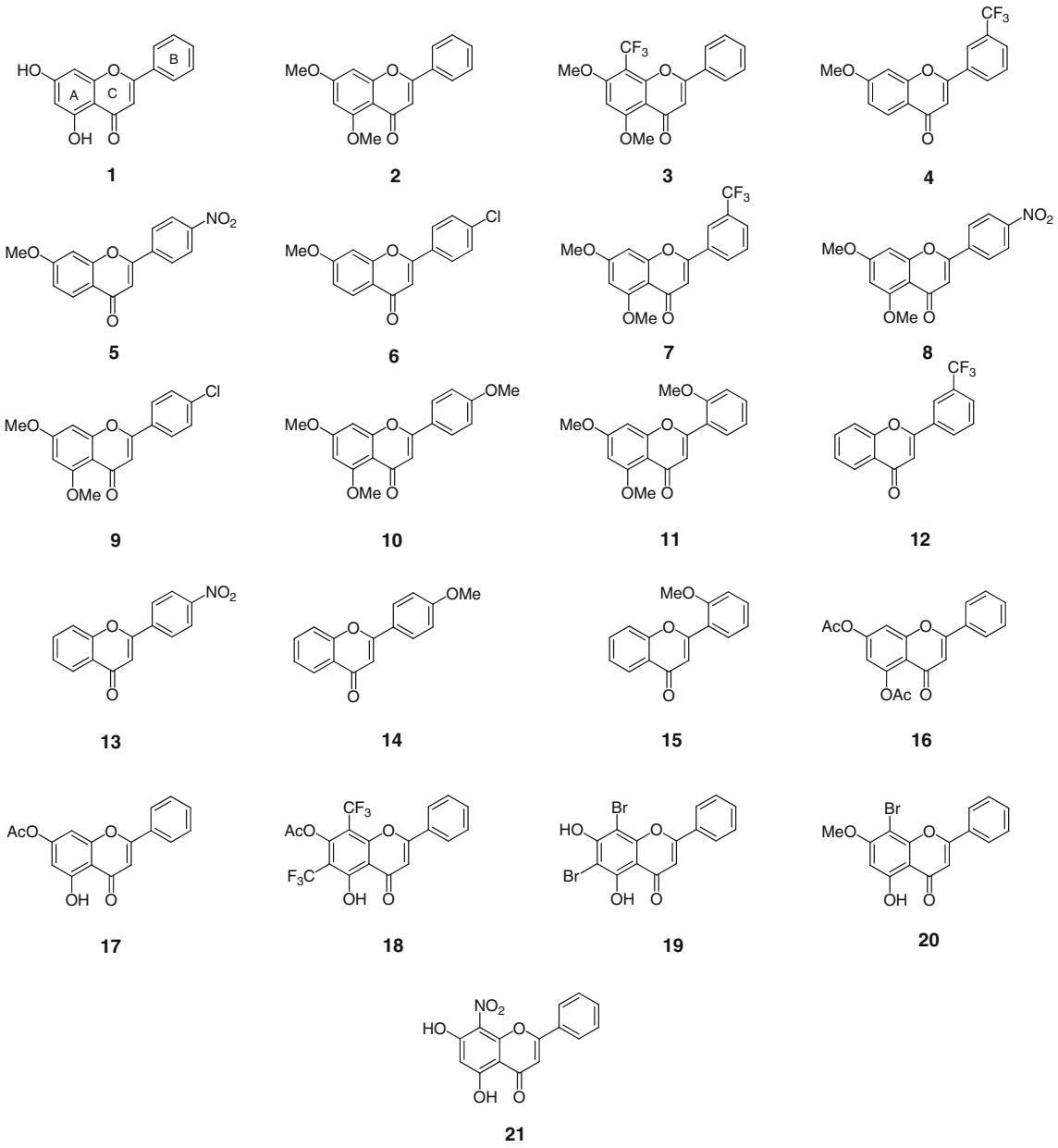

Fig. 1 Chemical structures of 21 chrysin derivatives used in this study

antioxidant (Heim *et al.*, 2002; Pietta, 2000), antimicrobial (Cushnie and Lamb, 2005; Takahashi *et al.*, 2004), anti-inflammatory (Rathee *et al.*, 2009; Serafini *et al.*, 2010), anticancer (Batra and Sharma, 2013; Kandaswami *et al.*, 2005), and antiviral (Ibrahim *et al.*, 2013; Khachatoorian *et al.*, 2012) activities. In fact, there are in excess of 4,000 types of biologically active flavonoid derivatives identified to date that include flavonol, flavanone, isoflavone, flavones, anthocyanin, and catechins sub-classes (Nijveldt *et al.*, 2001).

Chrysin (**1**, Fig. 1), also known as 5,7-dihydroxyflavone with an IUPAC name of 5,7-dihydroxy-2-phenyl-4H-chromen-4-one, belongs to the flavone sub-class of flavonoids. Its chemical structure is essentially based on a three ring nucleus with a phenyl B ring attached to position 2 of the fused bicyclic A and C rings. Chrysin had been shown to afford various biological activities such as antioxidant (Sathiavelu *et al.*, 2009), antimicrobial (Wang *et al.*, 2011), anti-inflammatory (Bae *et al.*, 2011), and anticancer (Zhang *et al.*, 2004a) activities. Particularly, it had been demonstrated to provide anticancer activities as it had been documented to inhibit cancer cells by down-regulating proliferating-cell nuclear antigen (PCNA) in human epithelioid cervix carcinoma (HeLa) (Zhang *et al.*, 2004b) as

well as inducing apoptosis in human macrophage (U937) cell line via caspase activation and Akt inactivation (Woo *et al.*, 2004). Furthermore, the compound was previously shown to provide inhibitory activities against tyrosinase (Kubo *et al.*, 2000), histone deacetylase 8 (Sun *et al.*, 2012), and aromatase (Mohammed *et al.*, 2011). Therefore, wide interest had been placed on the synthesis of chrysin derivatives as to develop novel candidates for further development as anticancer agents such as the synthesis of a series of chrysin derivatives with potent anticancer activities against human gastric adenocarcinoma (SGC-7901) and human colorectal adenocarcinoma (HT-29) cell lines as described by Zheng *et al.* (2003, 2010).

Quantitative structure–activity/property relationship (QSAR/QSPR) constitutes computational approaches for discerning the relationship between a set of molecular features from a series of investigated compounds with their respective biological activity/chemical properties of interest (Nantasenamat *et al.*, 2010). This robust paradigm had been extensively shown to be immensely useful in unraveling the mechanistic insights on the origins of a wide range of biological activities (Nantasenamat *et al.*, 2013a, b; Worachartcheewan *et al.*, 2013a, b) and chemical properties (Nantasenamat *et al.*, 2007a, b).

This study seeks to explore the origins of anticancer activities of chrysin derivatives. QSAR modeling was performed on a set of 21 chrysin derivatives using multiple linear regression (MLR). Important physicochemical features governing the anticancer activity of investigated compounds were identified by feature selection using stepwise MLR. Such QSAR models provide key understanding that could be used in further development of chrysin derivatives with potentially promising anticancer activities.

Materials and methods

Data set

A set of 21 chrysin derivatives were obtained from the work of Zheng *et al.* (2003, 2010). Initially, the data set consisted of 26 compounds but owing to inherent difficulties in obtaining low-energy conformers, therefore, five of these compounds (**22**–**26**) were not considered. The anticancer activity of these compounds was investigated using MTT assay against human gastric adenocarcinoma (SGC-7901) and human colorectal adenocarcinoma (HT-29). Anticancer activities of these investigated compounds were expressed in IC_{50} (μM) and were converted to pIC_{50} as to provide a more uniform distribution of the data points. This was performed by taking the negative logarithm to the base of 10 of IC_{50} values as presented by the following equation:

$$pIC_{50} = -\log_{10}(IC_{50}) \quad (1)$$

The molecular structures and their activity are displayed in Fig. 1, Table 1, respectively.

Descriptors calculation

Molecular structures of investigated compounds were drawn using the GaussView software package (DenningtonII *et al.*, 2003). Initial geometrical optimization of the crudely drawn 3D molecular structures were computed using the semi-empirical Austin Model 1 (AM1) method followed by further refinement at the density functional theory (DFT) level using Becke's three parameter hybrid method and the Lee–Yang–Parr (B3LYP) functional together with the 6-31g(d) basis set (B3LYP/6-31g(d)) using Gaussian 09 software (Frisch *et al.*, 2009), respectively. Quantum chemical descriptors were obtained from single point calculation at B3LYP/6-311++g(d,p) level using low-energy conformers as described above. This comprises of the dipole moment (μ), the total energy, the highest occupied molecular orbital energy (HOMO), and the lowest unoccupied molecular orbital (LUMO). Furthermore, an additional set of 7 quantum chemical descriptors (Parr *et al.*, 1978; Parr *et al.*, 1999; Parr and Pearson, 1983; Karelson *et al.*, 1996; Thanikaivelan *et al.*, 2000) comprising the energetic difference of HOMO and LUMO (HOMO–LUMO), Mulliken's electronegativity (χ), hardness (η), electrophilicity (ω), softness (S), electrophilicity index (ω_i), and mean absolute atomic charge (Q_m) were calculated from the following equations:

$$\chi = (\mathrm{HOMO} + \mathrm{LUMO})/2 \quad (2)$$

$$\eta = (\mathrm{LUMO} - \mathrm{HOMO})/2 \quad (3)$$

$$\omega = (\mathrm{HOMO} + \mathrm{LUMO}/2)^2/2\eta \quad (4)$$

$$S = 1/2\eta \quad (5)$$

$$\omega_i = \mu^2/2\eta \quad (6)$$

The mean absolute atomic charge (Q_m) was calculated from Mulliken population analysis according to the following equation:

$$Q_m = \sum_{a=1}^{N} |Q_a|/N, \quad (7)$$

where $|Q_a|$ and N are the absolute value of charges on all atoms and the total number of atoms present in the compound, respectively.

As to obtain additional set of molecular descriptors, the final optimized molecular structures from Gaussian were used as input to Dragon software (version 5.5, Talete SRL,

Table 1 Molecular descriptors and their values with anticancer activity of chrysin derivatives (**1–21**)

Compounds	Descriptors								SGC-7901		HT-29	
	SIC2	Mor11e	P2p	HTp	R5e+	L/Bw	BIC2	Mor19p	IC_{50} (μM)	pIC_{50}	IC_{50} (μM)	pIC_{50}
1	0.770	−0.071	0.181	7.502	0.066	3.750	0.708	0.229	5.800	−0.763	3.100	−0.491
2	0.733	−0.438	0.232	8.255	0.041	2.970	0.685	0.312	3.700	−0.568	2.000	−0.301
3	0.767	−0.251	0.206	8.711	0.038	2.090	0.720	0.375	5.900	−0.771	2.300	−0.362
4	0.820	0.025	0.188	8.238	0.077	5.970	0.764	0.260	5.780	−0.762	6.190	−0.792
5	0.801	−0.202	0.200	8.308	0.047	5.070	0.736	0.231	3.570	−0.553	5.860	−0.768
6	0.797	−0.020	0.201	9.117	0.047	4.900	0.737	0.176	6.890	−0.838	4.370	−0.640
7	0.791	−0.318	0.207	8.368	0.063	5.250	0.743	0.326	2.460	−0.391	3.800	−0.580
8[a]	0.777	−0.518	0.205	8.457	0.042	5.110	0.721	0.288	4.860	−0.687	5.810	−0.764
9[a]	0.772	−0.401	0.207	9.310	0.041	4.880	0.721	0.244	6.080	−0.784	5.220	−0.718
10	0.701	−0.521	0.175	8.660	0.056	4.590	0.660	0.259	12.050	−1.081	10.160	−1.007
11	0.711	−0.490	0.241	9.061	0.052	3.030	0.669	0.537	12.240	−1.088	14.900	−1.173
12[a]	0.759	0.194	0.156	7.572	0.093	6.270	0.700	0.240	13.280	−1.123	14.000	−1.146
13[a]	0.753	−0.034	0.147	7.845	0.072	7.380	0.683	0.213	6.250	−0.796	11.790	−1.072
14	0.738	−0.105	0.120	7.872	0.084	6.880	0.682	0.169	14.920	−1.174	15.560	−1.192
15	0.735	−0.112	0.210	8.332	0.062	3.480	0.679	0.486	7.980	−0.902	9.8800	−0.995
16	0.738	−0.231	0.339	9.502	0.041	1.690	0.688	0.355	4.500	−0.653	2.600	−0.415
17[a]	0.808	−0.335	0.185	8.366	0.043	3.740	0.749	0.312	3.600	−0.556	3.800	−0.580
18	0.800	0.391	0.212	9.470	0.071	2.840	0.750	0.526	8.600	−0.934	2.900	−0.462
19[a]	0.813	−0.420	0.193	9.825	0.040	3.740	0.747	0.449	5.100	−0.708	4.000	−0.602
20	0.831	−0.379	0.198	8.654	0.066	2.000	0.771	0.471	2.500	−0.398	1.900	−0.279
21	0.816	−0.712	0.219	7.977	0.139	2.290	0.745	0.280	2.800	−0.447	2.300	−0.362

[a] Compounds belonging to the external test set

Milan, Italy, 2007). This resulted in the calculation of large set of 3,224 descriptors that can be classified according to the following 22 categories including 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.

Feature descriptors selection

Constant values present in the set of 3,224 molecular descriptors were subjected to removal, and the remaining 1,428 molecular descriptors were combined with the set of 11 quantum chemical descriptors. Afterward, these combined sets of descriptors were subjected to further rounds of feature selection using stepwise MLR (SPSS statistics 18.0, SPSS Inc., USA). Furthermore, the independence of each descriptors were evaluated from the intercorrelation matrix among the selected significant descriptors as performed using Pearson's correlation coefficient as implemented in SPSS statistics 18.0 (SPSS Inc., USA).

Model validation

Validation of the QSAR model was done by sampling the data into two major subsets comprising of (i) internal and (ii) external sets. Training and leave-one-out cross-validation (LOO-CV) sets constituted the internal set, whereas six randomly selected compounds served as the external test set.

Multivariate analysis

MLR was used for development of QSAR models as briefly explained by the following equation:

$$Y = m_1x_1 + m_2x_2 + \cdots + m_nx_n + b, \quad (8)$$

where Y is the anticancer activity as expressed by pIC_{50} values, m is the regression coefficient values for descriptors, x is the descriptor, and b represents the y-intercept value. MLR was performed using the Waikato

Environment for Knowledge Analysis (WEKA), version 3.4.5 (Witten *et al.*, 2011).

Statistical analysis

Predictive performance of the QSAR models was evaluated using statistical parameters comprising of correlation coefficient (R^2 and Q^2) and root mean square error (RMSE) for the three data subsets: training, internal, and external test sets. In addition, the Fischer's (F) ratio was used to evaluate the statistical significance of the predictive models by considering the ratio of explained and unexplained variance. Moreover, a metric proposed by Eriksson and Johansson (1996) involving the difference of R^2 and Q^2 was also used to verify that the obtained QSAR models were not due to chance correlations. Thus, models were considered robust if the margin of R^2-Q^2 was less than 0.2–0.3 units.

Results and discussion

QSAR model of anticancer activities

The anticancer activity of the investigated set of chrysin derivatives was determined using the MTT assay against SGC-7901 and HT-29 cell lines. These analogs carry functional groups such as halogens (F, Cl, Br), methoxy (OCH_3), and nitrogen dioxide (NO_2) on rings A, B, and C (**1–21**, Fig. 1). The 5-fluorouracil served as the reference compound affording IC_{50} values of 5.28 and 9.56 μg/mL for SGC-7901 and HT-29 cells, respectively. The substituent and position on ring A and B of chrysin structure displayed significantly improved activity to inhibit growth of cancer cells. Almost all compounds (Table 1) were shown to afford effective anticancer activity when compared to 5-fluorouracil. As such, this study explores the origins of the anticancer activity via QSAR modeling.

Quantum chemical and molecular descriptors had previously been shown to be useful in the construction of QSAR/QSPR models, particularly, for representing the inherent molecular features of investigated compounds as well as being used for understanding their bioactivities (Ishihara *et al.*, 2006; Ishihara *et al.*, 2007; Nantasenamat *et al.*, 2007a, b; Nantasenamat *et al.*, 2013a, b; Worachartcheewan *et al.*, 2013a, b, Worachartcheewan *et al.*, 2014). The investigated set of 21 chrysin derivatives was described by quantum chemical and molecular descriptors. Particularly, quantum chemical descriptors were obtained from single point calculation at B3LYP/6-311++g(d) level using low-energy conformers geometrically optimized initially at AM1 level followed by refined optimization at B3LYP/6-31g(d) level. Owing to the inherently large volume of 3,235 descriptors (11 quantum chemical and 3,224 molecular descriptors), there was a need to select a smaller subset as to reduce the computational time as well as the multi-collinearity and redundancy in the descriptors. Therefore, feature selection was applied by firstly subjecting the 3,224 Dragon descriptors with constant values for removal. The remaining 1,428 Dragon descriptors were then integrated with the 11 descriptors from Gaussian 09. Further feature selection was carried out using stepwise MLR for the identification of significant descriptors corresponding to their activity. After using stepwise MLR, 5 descriptors comprising of SIC2, Mor11e, P2p, HTp, and R5e+ were used for constructing QSAR model of SGC-7901 cell line while 3 descriptors consisting of L/Bw, BIC2, and Mor19p were used for computing the QSAR model of HT-29 cell line. Values and definitions of these significant descriptors are presented in Tables 1, 2, respectively, for the SGC-7901 and HT-29 data sets. It can be seen that each of the important descriptors obtained from stepwise MLR was independent from one another as deduced from low correlation coefficient among the set of descriptors (i.e., SIC2, Mor11e, P2p, HTp, and R5e+) for the SGC-7901 data set (Table 3) as well as among the set

Table 2 Definitions of significant descriptors

Model	Activity	Symbol	Definition	Type
1	pIC_{50} of SGC-7901	SIC2	Structural information content (neighborhood symmetry of 2-order)	Information indices
		Mor11e	3D-MoRSE-signal 11/weighted by atomic Sanderson electronegativity	3D-MoRSE descriptors
		P2p	2nd component shape directional WHIM index/weighted by atomic polarizabilities	WHIM descriptors
		HTp	*H* total index/weighted by atomic polarizabilities	GETAWAY descriptors
		R5e+	*R* maximal autocorrelation of log 5/weighted by atomic Sanderson electronegativity	GETAWAY descriptors
2	pIC_{50} of HT-29	L/Bw	Length-to-breadth ratio by WHIM	Geometrical descriptors
		BIC2	Bond information content (neighborhood symmetry of 2-order)	Information indices
		Mor19p	3D-MoRSE-signal 19/weighted by atomic polarizabilities	3D-MoRSE descriptors

Table 3 Intercorrelation between significant descriptors for predicting anticancer activity against SGC-7901 cell lines

	SIC2	Mor11e	P2p	HTp	R5e
SIC2	1.000				
Mor11e	0.255	1.000			
P2p	−0.129	−0.158	1.000		
HTp	−0.097	0.231	0.599	1.000	
R5e	0.396	−0.206	−0.290	−0.436	1.000

Table 4 Intercorrelation between significant descriptors for predicting anticancer activity against HT-29 cell lines

	L/Bw	BIC2	Mor19p
L/Bw	1.000		
BIC2	−0.008	1.000	
Mor19p	−0.623	−0.015	1.000

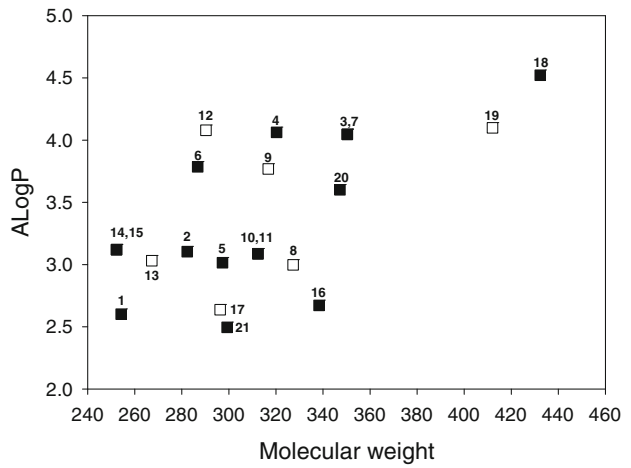

Fig. 2 Plot of the molecular weight versus ALogP of 21 chrysin derivatives. Compounds belonging to the internal set are shown as *black squares* while compounds in the external set are shown as *white squares*

of descriptors (i.e., L/Bw, BIC2, and Mor19p) for the HT-29 data set (Table 4). Therefore, sets of 5 and 3 descriptors were further employed in the construction of QSAR models for predicting the anticancer activities of SGC-7901 and HT-29, respectively.

Validation of QSAR modeling

Compounds **8**, **9**, **12**, **13**, **17**, and **19** were randomly selected from the original data set to serve as the external test set while the remaining 15 compounds (**1–7**, **10–11**, **14–16,** and **20–21**) was utilized as the internal test set. The general chemical space of the internal and external sets is depicted in Fig. 2 as plot of molecular weight versus ALogP. It is observed that all compounds are drug-like having MW <500 Da and ALogP <5. Two predictive models were generated for the two distinct cancer cell lines as described below.

Linear equations derived from the QSAR model for predicting the anticancer activity against SGC-7901 is shown below:

$$\begin{aligned}\text{SGC-7901 } \text{pIC}_{50} = {} & 5.4129\,(\text{SIC2}) + 2.9888\,(\text{P2p}) \\ & -0.3873\,(\text{Mor11e}) - 0.1930\,(\text{HTp}) \\ & -3.3596(\text{R5e+}) - 3.7742\end{aligned} \tag{9}$$

$N_{CV} = 15$, $N_{Ext} = 6$, $R^2_{Tr} = 0.8778$, $\text{RMSE}_{Tr} = 0.0854$, $Q^2_{CV} = 0.7315$, $\text{RMSE}_{CV} = 0.1375$, $Q^2_{Ext} = 0.7324$, $\text{RMSE}_{Ext} = 0.1168$, F value $= 4.896$, critical F value $= 3.482$, $R^2_{Tr} - Q^2_{CV} = 0.1463$, and $= R^2_{Tr} - Q^2_{Ext} = 0.1454$

The predictive performance for predicting the pIC_{50} value against SGC-7901 was shown to afford good results as deduced from the correlation coefficient and RMSE of $R^2_{Tr} = 0.8778$ and $\text{RMSE}_{Tr} = 0.0854$ for the training set, $Q^2_{CV} = 0.7315$ and $\text{RMSE}_{CV} = 0.1375$ for the LOO-CV set while $Q^2_{Ext} = 0.7324$ and $\text{RMSE}_{Ext} = 0.1168$ were obtained for the external set. These models exhibited $R^2 > 0.6$ and $Q^2 > 0.5$, which suggest its robust performance (Tropsha *et al.*, 2003). In addition, R^2-Q^2 margins for LOO-CV and external sets were 0.1463 and 0.1454, respectively, which were significantly less than 0.3, thereby indicating high reliability of the constructed models (Eriksson and Johansson, 1996). The predicted activity values against SGC-7901 cell line are presented in Table 5 and a plot of the experimental versus predicted values is shown in Fig. 3a.

Linear equation obtained from the QSAR model for predicting the anticancer activity against HT-29 cell is shown below:

$$\begin{aligned}\text{HT-29 } \text{pIC}_{50} = {} & 4.5185\,(\text{BIC2}) - 0.1851\,(\text{L/Bw}) \\ & -1.2682\,(\text{Mor19p}) - 2.7660\end{aligned} \tag{10}$$

$N_{CV} = 15$, $N_{Ext} = 6$, $R^2_{Tr} = 0.8201$, $\text{RMSE}_{Tr} = 0.1293$, $Q^2_{CV} = 0.6289$, $\text{RMSE}_{CV} = 0.1735$, $Q^2_{Ext} = 0.8486$, $\text{RMSE}_{Ext} = 0.1179$, F ratio $= 7.898$, critical F value $= 3.587$, $R^2_{Tr} - Q^2_{CV} = 0.1372$, and $R^2_{Tr} - Q^2_{Ext} = -0.0285$

Similarly, good predictive results were attained for the HT-29 data set as corroborated by statistical results in which the training set provided $R^2_{Tr} = 0.8201$ and $\text{RMSE}_{Tr} = 0.1293$. The LOO-CV set afforded values of $Q^2_{CV} = 0.6829$ and $\text{RMSE}_{CV} = 0.1735$ while the external set gave $Q^2_{Ext} = 0.8486$ and $\text{RMSE}_{Ext} = 0.1179$. Similarly, these models also had $R^2 > 0.6$ and $Q^2 > 0.5$ as well as affording low margins of R^2-Q^2 for both LOO-CV and external sets with corresponding values of 0.1372 and −0.0285, respectively. The predicted activity values against HT-29 cells are presented in Table 5, and a plot of the experimental versus predicted activities is presented in Fig. 3b.

Table 5 Summary of the experimental and predicted pIC_{50} values along with their residuals for both SGC-7901 and HT-29 data sets

Compound	pIC_{50} (SGC-7901)			pIC_{50} (HT-29)		
	Experimental	Predicted	Residual	Experimental	Predicted	Residual
1	−0.763	−0.656	−0.107	−0.491	−0.563	0.072
2	−0.568	−0.716	0.148	−0.301	−0.680	0.379
3	−0.771	−0.707	−0.064	−0.362	−0.102	−0.260
4	−0.762	−0.593	−0.169	−0.792	−0.726	−0.066
5	−0.553	−0.515	−0.038	−0.768	−0.655	−0.113
6	−0.838	−0.745	−0.093	−0.640	−0.547	−0.093
7	−0.391	−0.603	0.212	−0.580	−0.848	0.268
8[a]	−0.687	−0.529	−0.158	−0.764	−0.819	0.055
9[a]	−0.784	−0.756	−0.028	−0.718	−0.721	0.003
10	−1.081	−1.147	0.066	−1.007	−0.946	−0.061
11	−1.088	−0.870	−0.218	−1.173	−0.853	−0.320
12[a]	−1.123	−1.049	−0.074	−1.146	−1.068	−0.078
13[a]	−0.796	−1.002	0.206	−1.072	−1.316	0.244
14	−1.174	−1.190	0.016	−1.192	−1.159	−0.033
15	−0.902	−0.952	0.050	−0.995	−0.944	−0.051
16	−0.556	−0.477	−0.079	−0.415	−0.422	0.007
17[a]	−0.653	−0.639	−0.014	−0.580	−0.470	−0.110
18	−0.934	−1.235	0.301	−0.462	−0.623	0.161
19[a]	−0.708	−0.665	−0.043	−0.602	−0.652	0.050
20	−0.398	−0.452	0.054	−0.279	−0.234	−0.045
21	−0.447	−0.330	−0.117	−0.362	−0.378	0.016

[a] Compounds belonging to the external test set

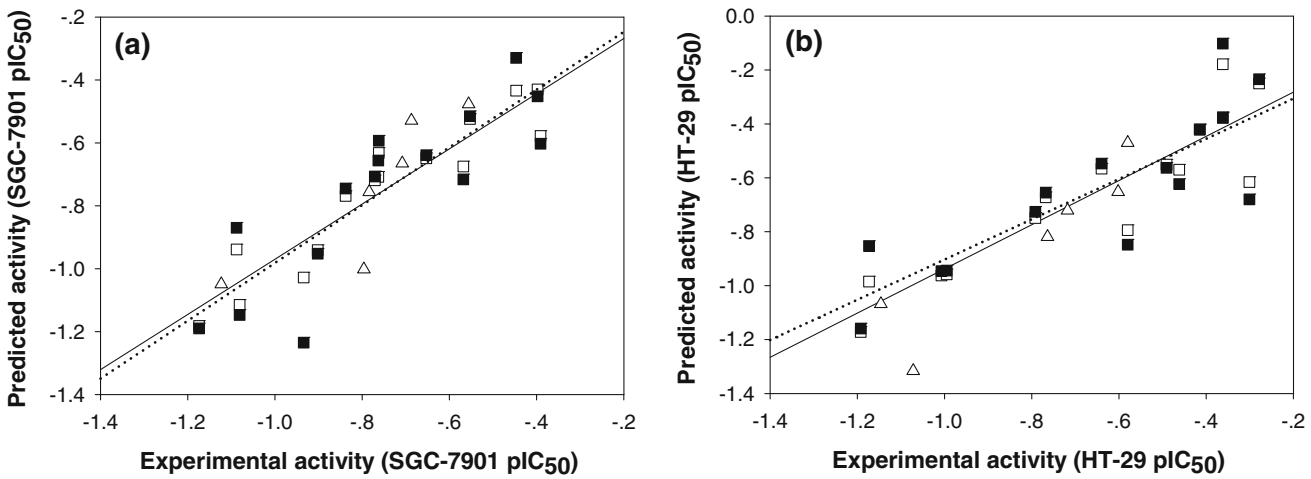

Fig. 3 Plot of the experimental versus predicted anticancer (pIC_{50}) activities. Shown are data samples from the training set (*white squares* and *solid line*), leave-one-out cross-validation set (*black squares* and *dotted line*), and external test set (*white triangle* for external test set) for SGC-7901 (**a**) and HT-29 (**b**) data sets

Interpretation of QSAR model

As observed from the obtained MLR model against SGC-7901 in Eq. (9), the most significant descriptors were found to be involved in atomic electronegativities (i.e., Mor11e and R5e+) and atomic polarizabilities (i.e., P2p and HTp) as shown in Table 2. Furthermore, one of the descriptors accounts for the shape of the molecule (i.e., SIC2). A closer analysis of the relative importance of descriptors as displayed in Eq. (9) and obtained by calculating the absolute standardized regression coefficient (Duchowicz *et al.*, 2014) revealed the following order of significance:

SIC2 > P2P > HTp > Mor11e > R5e+ with corresponding values of 0.803, 0.525, 0.522, 0.380, and 0.285, respectively. It may be implied that the interaction between cancer cells and investigated compounds is reliant on molecular shape followed by polarizability and electronegativity of the molecules.

In regards to the MLR model of HT-29 as described in Eq. (10), the interaction between cancer cells and compounds mostly involved the molecular shape (i.e., L/Bw and BIC2) followed by atomic polarizabilities (i.e., Mor19p). Furthermore, it was found that the relative importance of descriptors as shown in Eq. (10) was L/Bw > BIC2 > Mor19p with corresponding absolute standardized regression coefficient values of 0.919, 0.467, and 0.435, respectively. The interaction between cancer cells and the tested compounds may involve molecular shape followed by polarizability of the molecules.

Both QSAR models for predicting the anticancer activities against SGC-7901 and HT-29 cell lines were firstly accounted for by their molecular shape and further by their electronegativities and polarizabilities. There are evidences to support that generally the interaction of the cellular target with drugs or compounds mostly involved the conformation or shape of molecules (Ishihara *et al.*, 2006; Ishihara *et al.*, 2007). This coincides with the selected descriptors of chrysin derivatives where their molecular shape and electron distribution as described by their electronegativities and polarizabilities were important in explaining the observed anticancer activities.

Conclusion

This study carried out a QSAR investigation of chrysin derivatives as to unravel the origins of their anticancer activity against human gastric cancer (SGC-7901) and human colon adenocarcinoma (HT-29). Particularly, essential physicochemical parameters giving rise to anticancer activities of chrysin derivatives toward SGC-7901 included SIC2, Mor11e, P2p, HTp, and R5e+ while pertinent properties against HT-29 included L/Bw, BIC2, and Mor19p. Models were validated by means of internal and external sets in which good predictive performance were achieved in this study as observed from high correlation coefficient and low RMSE. Structure–activity relationship revealed ideal characteristics for anticancer activities in chrysin derivatives and they were found to mostly involve the following molecular features: shape, electronegativities, and polarizabilities. QSAR models developed in this study could be useful in rationalizing the origins of anticancer activities and could serve as guidelines for the future design of new chrysin derivatives for potential therapeutic applications.

Acknowledgments This research project is supported by the annual budget grant of Mahidol University (B.E. 2556–2558). A. W. is thankful for Mahidol University Talent Management Program. Partial support is gratefully acknowledged from Office of the Higher Education Commission and Mahidol University under the National Research Universities Initiative.

References

Bae Y, Lee S, Kim SH (2011) Chrysin suppresses mast cell-mediated allergic inflammation: involvement of calcium, caspase-1 and nuclear factor-κB. Toxicol Appl Pharmacol 254:56–64

Batra P, Sharma AK (2013) Anti-cancer potential of flavonoids: recent trends and future perspectives. 3 Biotech 3:439–459

Cragg GM, Grothaus PG, Newman DJ (2009) Impact of natural products on developing new anti-cancer agents. Chem Rev 109:3012–3043

Cushnie TP, Lamb AJ (2005) Antimicrobial activity of flavonoids. Int J Antimicrob Agents 26:343–356

DenningtonII R, Keith T, Millam J, Eppinnett K, Hovell WL, Gilliland R (2003) GaussView, Version 3.09. Semichem Inc, Shawnee Mission, KS

Drews J (2007) Drug discovery: a historical perspective. Science 287:1960–1964

Duchowicz PR, Bennardi DO, Bacelo DE, Bonifazi EL, Rios-Luci C, Padrón JM, Burton G, Misico RI (2014) QSAR on antiproliferative naphthoquinones based on a conformation-independent approach. Eur J Med Chem 77:176–184

Eriksson L, Johansson E (1996) Multivariate design and modeling in QSAR. Chemometr Intell Lab Syst 34:1–19

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.1. Wallingford, Connecticut

Heim KE, Tagliaferro AR, Bobilya DJ (2002) Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. J Nutr Biochem 13:572–584

Ibrahim AK, Youssef AI, Arafa AS, Ahmed SA (2013) Anti-H5N1 virus flavonoids from *Capparis sinaica* Veill. Nat Prod Res 27:2149–2153

Ishihara M, Yokote Y, Sakagami H (2006) Quantitative structure-cytotoxicity relationship analysis of coumarin and its derivatives by semiempirical molecular orbital method. Anticancer Res 26:2883–2886

Ishihara M, Kawase M, Westman G, Samuelsson K, Motohashi N, Sakagami H (2007) Quantitative structure-cytotoxicity relationship analysis of phenoxazine derivatives by semiempirical molecular-orbital method. Anticancer Res 27:4053–4057

Kandaswami C, Lee LT, Lee PP, Hwang JJ, Ke FC, Huang YT, Lee MT (2005) The antitumor activities of flavonoids. In Vivo 19:895–909

Karelson M, Lobanov VS, Katritzky AR (1996) Quantum-chemical descriptors in QSAR/QSPR studies. Chem Rev 96:1027–1044

Khachatoorian R, Arumugaswami V, Raychaudhuri S, Yeh GK, Maloney EM, Wang J, Dasgupta A, French SW (2012) Divergent antiviral effects of bioflavonoids on the hepatitis C virus life cycle. Virology 433:346–355

Kubo I, Kinst-Hori I, Chaudhuri SK, Kubo Y, Sánchez Y, Ogura T (2000) Flavonols from *Heterotheca inuloides*: tyrosinase inhibitory activity and structural criteria. Bioorg Med Chem 8:1749–1755

Mohammed HA, Ba LA, Burkholz T, Schumann E, Diesel B, Zapp J, Kiemer AK, Ries C, Hartmann RW, Hosny M, Jacob C (2011) Facile synthesis of chrysin-derivatives with promising activities as aromatase inhibitors. Nat Prod Commun 6:31–34

Nantasenamat C, Isarankura-Na-Ayudhya C, Naenna T, Prachayasittikul V (2007a) Quantitative structure-imprinting factor relationship of molecularly imprinted polymers. Biosens Bioelectron 22:3309–3317

Nantasenamat C, Isarankura-Na-Ayudhya C, Tansila N, Naenna T, Prachayasittikul V (2007b) Prediction of GFP spectral properties using artificial neural network. J Comput Chem 28:1275–1289

Nantasenamat C, Isarankura-Na-Ayudhya C, Prachayasittikul V (2010) Advances in computational methods to predict the biological activity of compounds. Expert Opin Drug Discov 5:633–654

Nantasenamat C, Worachartcheewan A, Prachayasittikul S, Isarankura-Na-Ayudhya C, Prachayasittikul V (2013a) QSAR modeling of aromatase inhibitory activity of 1-substituted 1,2,3-triazole analogs of letrozole. Eur J Med Chem 69:99–114

Nantasenamat C, Li H, Mandi P, Worachartcheewan A, Monnor T, Isarankura-Na-Ayudhya C, Prachayasittikul V (2013b) Exploring the chemical space of aromatase inhibitors. Mol Divers 17:661–677

Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. J Nat Prod 70:461–477

Nijveldt RJ, van Nood E, van Hoorn DE, Boelens PG, van Norren K, van Leeuwen PA (2001) Flavonoids: a review of probable mechanisms of action and potential applications. Am J Clin Nutr 74:418–425

Parr RG, Pearson RG (1983) Absolute hardness: companion parameter to absolute electronegativity. J Am Chem Soc 105:7512–7516

Parr RG, Donnelly RA, Levy M, Palke WE (1978) Electronegativity: the density functional viewpoint. J Chem Phys 68:3801–3807

Parr RG, Szentpaly Lv, Liu S (1999) Electrophilicity Index. J Am Chem Soc 121:1922–1924

Pietta PG (2000) Flavonoids as antioxidants. J Nat Prod 63: 1035–1042

Rathee P, Chaudhary H, Rathee S, Rathee D, Kumar V, Kohli K (2009) Mechanism of action of flavonoids as anti-inflammatory agents: a review. Inflamm Allergy Drug Targets 8:229–235

Sathiavelu J, Senapathy GJ, Devaraj R, Namasivayam N (2009) Hepatoprotective effect of chrysin on prooxidant-antioxidant status during ethanol-induced toxicity in female albino rats. J Pharm Pharmacol 61:809–817

Serafini M, Peluso I, Raguzzini A (2010) Flavonoids as anti-inflammatory agents. Proc Nutr Soc 69:273–278

Sun LP, Chen AL, Hung HC, Chien YH, Huang JS, Huang CY, Chen YW, Chen CN (2012) Chrysin: a histone deacetylase 8 inhibitor with anticancer activity and a suitable candidate for the standardization of Chinese propolis. J Agric Food Chem 60:11748–11758

Takahashi T, Kokubo R, Sakaino M (2004) Antimicrobial activities of eucalyptus leaf extracts and flavonoids from *Eucalyptus maculata*. Lett Appl Microbiol 39:60–64

Thanikaivelan P, Subramanian V, Raghava Rao J, Unni Nair B (2000) Application of quantum chemical descriptor in quantitative structure activity and structure property relationship. Chem Phys Lett 323:59–70

Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci 22:69–77

Wang J, Qiu J, Dong J, Li H, Luo M, Dai X, Zhang Y, Leng B, Niu X, Zhao S, Deng X (2011) Chrysin protects mice from *Staphylococcus aureus* pneumonia. J Appl Microbiol 111:1551–1558

Witten IH, Frank E, Hall MA (2011) Data mining: practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, San Francisco

Woo KJ, Jeong YJ, Park JW, Kwon TK (2004) Chrysin-induced apoptosis is mediated through caspase activation and Akt inactivation in U937 leukemia cells. Biochem Biophys Res Commun 325:1215–1222

Worachartcheewan A, Nantasenamat C, Isarankura-Na-Ayudhya C, Prachayasittikul V (2013a) QSAR study of amidino bis-benzimidazole derivatives as potent anti-malarial agents against *Plasmodium falciparum*. Chem Pap 67:1462–1473

Worachartcheewan A, Nantasenamat C, Isarankura-Na-Ayudhya C, Prachayasittikul V (2013b) Predicting antimicrobial activities of benzimidazole derivatives. Med Chem Res 22:5418–5430

Worachartcheewan A, Nantasenamat C, Owasirikul W, Monnor T, Naruepantawart O, Janyapaisarn S, Prachayasittikul S, Prachayasittikul V (2014) Insights into antioxidant activity of 1-adamantylthiopyridine analogs using multiple linear regression. Eur J Med Chem 73:258–264

Zhang S, Yang X, Morris ME (2004a) Flavonoids are inhibitors of breast cancer resistance protein (ABCG2)-mediated transport. Mol Pharmacol 65:1208–1216

Zhang T, Chen X, Qu L, Wu J, Cui R, Zhao Y (2004b) Chrysin and its phosphate ester inhibit cell proliferation and induce apoptosis in Hela cells. Bioorg Med Chem 12:6097–6105

Zheng X, Meng WD, Xu YY, Cao JG, Qing FL (2003) Synthesis and anticancer effect of chrysin derivatives. Bioorg Med Chem Lett 13:881–884

Zheng X, Zhao FF, Liu YM, Yao X, Zheng ZT, Luo X, Liao DF (2010) Synthesis and preliminary biological evaluation of chrysin derivatives as potential anticancer drugs. Med Chem 6:6–8