ISSN 1070-3632, Russian Journal of General Chemistry, 2011, Vol. 81, No. 11, pp. 2320–2327. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © I.G. Tsygankova, S.M. Zhenodarova, 2011, published in Zhurnal Obshchei Khimii, 2011, Vol. 81, No. 11, pp. 1869–1876.

The Structure–Property Correlation for Estimation and Predicting the Apoptosis-Inducing Activity of 3,5-Diaryl-1,2,4-oxadiazoles, the Potential Antitumor Agents

I. G. Tsygankova and S. M. Zhenodarova

Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, ul. Institutskaya 3, Pushchino, Moscow Oblast, 142290 Russia
e-mail: tsygan@iteb.ru

Received November 8, 2010

Abstract—The structure–property correlation relation based on the fragment descriptors of molecular structure was used for the calculation of the apoptosis-inducing activity of 3,5-diaryl-1,2,4-oxadiazoles. A set of the correlation models of similar accuracy, with different numbers of variables was obtained. Analysis of the role of various fragments allowed us to offer compounds with high activity belonging to one chemical class, which can be useful for the creation of new anticancer drugs.

DOI: 10.1134/S107036321111017X

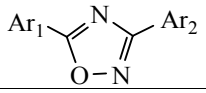
Currently, oncology diseases are the second leading cause of mortality in developed countries [1]. Therefore, despite the notable success in treating these diseases due to the creation of many new drugs, the problem of developing more effective, less toxic and not causing resistance anticancer agents remain in the focus of many research laboratories and pharmaceutical companies [1]. Since the antitumor activity of many drugs correlates with their ability to induce apoptosis [2], the identification and development of new apoptosis inducers is one of the directions at the development of new means for the treatment of cancer. In recent years, at the designing compounds with pro-apoptotic properties were used the results of the study of quantitative structure–activity relations, that allow revealing the structural fragments responsible for the apoptosis-inducing and antiproliferative activity [3–9]. In this study, the structure–property correlation [10–12] was used to evaluate and predict the apoptosis-inducing activity of 3,5-diaryl-1,2,4-oxadiazoles, the potential anticancer agents inhibiting selectively growth of some tumor cells and not affecting healthy cells [13–15]. The experimental values of the apoptosis-inducing activity of these compounds, taken from the literature [14–15], have been determined by the activity of caspases (EC50), cleaving fluorescent substrate in the cell cultures pretreated with the appropriate inducer [16].

The model construction. To construct a structure–property correlation relation we used the fragment representation of the molecular structure: heavy atoms or groups are represented by the fragments of different types. In the set of test compounds (Table 1) there is a repeating core consisting of 1,2,4-oxadiazole and aromatic substituents in 3 and 5 positions. This core is not characterized by a variety of conformations, and despite the presence of various groups in the aromatic substituents it is virtually planar. Activity of some compounds in this set of 1,2,4-oxadiazole derivatives has been defined up to the boundary values only, e.g., to $EC_{50} > 40$ μM. In such cases, we assumed that the concentration is equal to this boundary value. Because active concentration of these compounds varied in the range of several orders of magnitude, we used the logarithm of this characteristic.

The correlation expression was taken as a sum of mono-fragmental contributions associated with an atom of a certain type, and paired contributions, determined by the type of a pair of atoms and the number of bonds separating the atoms [6].

$$P \approx \sum_t n_t P_t + \sum_{t,m,r} k_{t,m-r} P_{t,m-r} + \text{const.} \quad (1)$$

Table 1. Derivatives of 1,2,4-oxadiazoles inducing apoptosis involving caspases.



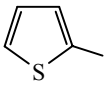
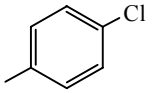
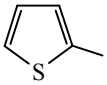
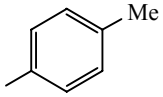
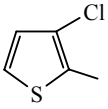
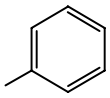
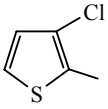
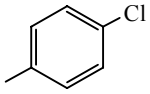
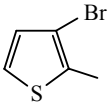
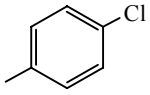
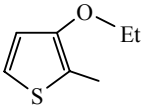
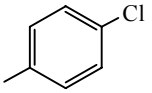
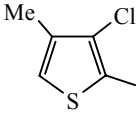
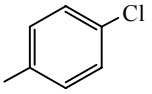
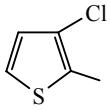
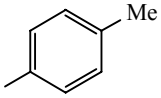
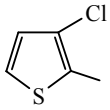
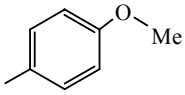
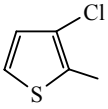
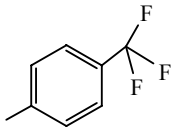
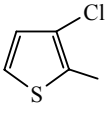
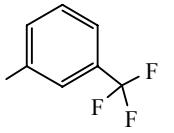
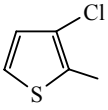
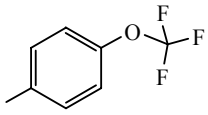
Run no.	Ar_1	Ar_2	Experiment[a]		Calculation log EC_{50}[b]
			EC_{50}, μM[c]	log EC_{50}	
1			40	1.6	1.46
2			40	1.6	1.64
3			2.82	0.45	0.47
4			1.77	0.25	0.30
5			3.32	0.52	0.48
6			40	1.6	1.61
7			40	1.6	1.60
8			3.66	0.56	0.47
9			4.17	0.62	0.62
10			1.21	0.08	0.13
11			0.9	–0.05	0.13
12			0.89	–0.05	–0.07

Table 1. (Contd.)

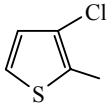
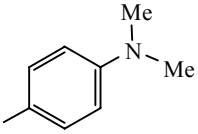
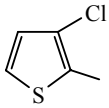
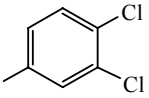
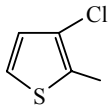
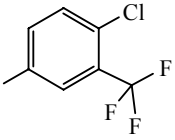
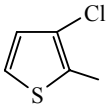
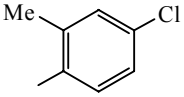
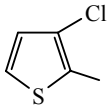
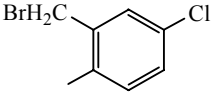
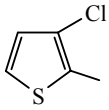
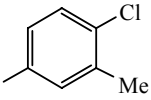
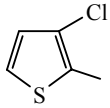
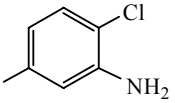
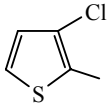
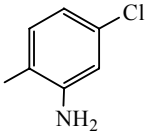
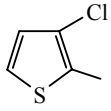
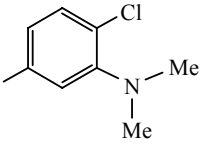
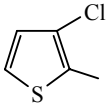
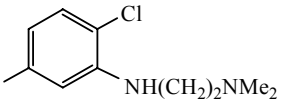
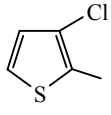
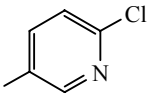
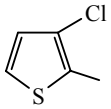
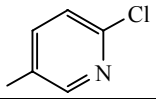
Run no.	Ar_1	Ar_2	Experiment[a]		Calculation log EC_{50}[b]
			EC_{50}, μM[c]	log EC_{50}	
13			40	1.6	1.66
14			1.41	0.15	0.12
15			0.93	–0.03	–0.05
16			0.91	–0.04	0.30
17			4.5	0.65	0.54
18			1.7	0.23	0.30
19			2.4	0.38	0.30
20			3.1	0.49	0.30
21			2.1	0.32	0.30
22			10	1	0.89
23			1.59	0.2	0.31
24			0.75	–0.12	0.00

Table 1. (Contd.)

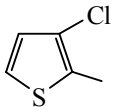
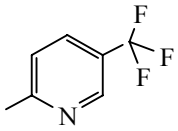
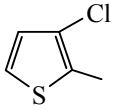
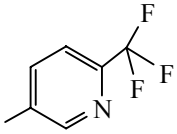
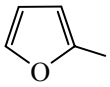
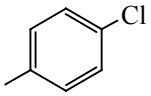
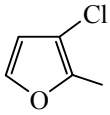
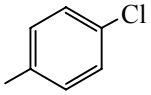
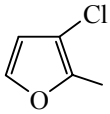
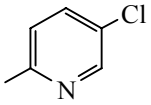
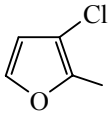
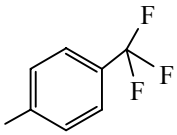
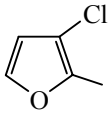
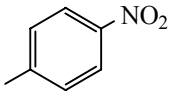
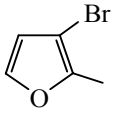
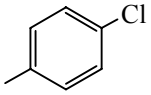
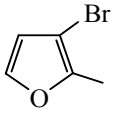
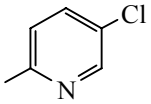
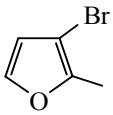
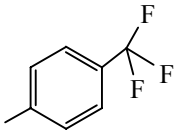
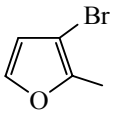
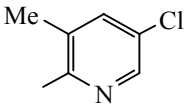
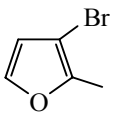
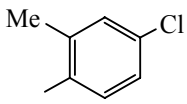
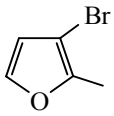
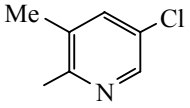
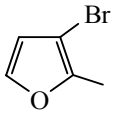
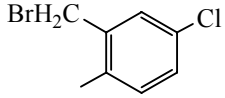
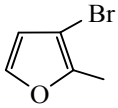
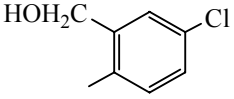
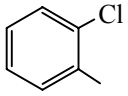
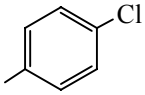
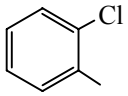
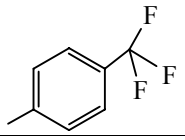
Run no.	Ar_1	Ar_2	Experiment[a]		Calculation log EC_{50}[b]
			EC_{50}, μM[c]	log EC_{50}	
25			1.44	0.16	0.14
26			0.92	–0.04	–0.17
27			17.4	1.24	1.18
28			1.74	0.24	0.02
29			1.11	0.05	0.03
30			0.52	–0.28	–0.15
31			1.97	0.29	0.20
32			1.08	0.03	0.20
33			1.63	0.21	0.12
34			2.01	0.3	0.03
35			0.61	–0.21	–0.05
36			1.1	0.04	0.20

Table 1. (Contd.)

Run no.	Ar_1	Ar_2	Experiment[a]		Calculation log EC_{50}[b]
			EC_{50}, μM[c]	log EC_{50}	
37			0.94	–0.03	0.12
38			2.2	0.34	0.45
39			2.9	0.46	0.20
40			40	1.6	1.73
41			40	1.6	1.56
					r 0.9532
					s 0.15

[a] The experimental values are taken from [14, 15]. [b] For the calculation we used the relation (2). [c] EC_{50} is the concentration at which occurs the splitting of up to 50% of the fluorescent substrate of caspase 3 in the cells of T47D line pre-treated with the corresponding derivative [16].

Here n_t is the number of fragments of t type associated with specific atoms; $k_{t,m-r}$ is the number of pairs of fragments of type t and m with the corresponding atoms separated by r chemical bonds. Types of fragments t or m are determined by chemical composition and structural features of the considered series of compounds. To construct the linear model, the total number of types of fragments should not be very large, since the corresponding fragment contributions P_t and $P_{t,m-r}$ being the parameters of the correlation equation are determined from a limited set of experimentally measured activities by the method of linear regression. The n_t and $k_{t,m-r}$ values describe the molecular structure quantitatively, taking different values for different compounds, and thus perform the role of fragment descriptors. Such fragment descriptors are well suited to represent in the numerical form the molecular structures not featured by stereoisomerism, as occurs in the case of the oxadiazole. The used correlation models with the fragment descriptors as the variables allow estimating activity of compounds without detailed knowledge of its mechanism.

To represent the molecular structure of the oxadiazole derivatives (Table 1) we selected 14 types of fragments associated with different atoms and groups of atoms. To nitrogen atom relate the fragments of four types, designated as follows: to the fragment N1 corresponds the nitrogen atom in the oxadiazole ring, to N2 aromatic nitrogen in the pyridyl substituent, the N3 fragment denotes a nitro group as a whole, the N4 fragment corresponds to the nitrogen atom in amino group. To the carbon atom are assigned three types of fragments: C1 fragment is connected with the carbon atom in the five-membered rings of oxadiazole, thiophene, and furan, C2 is assigned to aromatic carbon atom in the phenyl or pyridyl ring, C3 means an aliphatic carbon atom in the substituents. To the oxygen atom also correspond three types of fragments: O1 denotes the oxygen atom in the five-membered oxadiazole or furan ring, O2 is oxygen in the ether groups of substituent, the O3 fragment encodes the oxygen in a hydroxyl group. To the F, Cl, Br and S atoms is assigned one type of fragments to each, denoted, respectively, as F, Cl, Br and S. Thus, the

indices t and m in Eq. (1) may take 14 different values listed above.

The maximum distance between the heavy atoms in the compounds of the series of oxadiazole derivatives is equal to 13 chemical bonds. The total number of descriptors (mono-fragmental and different pairs at different distances) reaches 1379. After exclusion of the descriptors having zero or constant values for all compounds of the series, as well as the correlated descriptors (the correlation coefficient over 0.98), remained 127 descriptors only. It must be emphasized that the matrix of the descriptors values for all compounds of this series is a very incompact one. Only 10 of 127 descriptors have non-zero values for all compounds of the series, while 61 of descriptors (48% of the total) has a non-zero values for less than 10 compounds of the considered set. Such an absence of many contributions presents certain difficulties in interpreting the results. Of a set of 127 descriptors were selected the most effective descriptors for the regression models using the previously described algorithm [6, 17]. The criterion for selection was the requirement that the standard deviation of activity obtained by the estimation with the model and the experimental values was less than 0.19 logarithmic units (~10% difference between the minimum and maximum activity in this set of compounds). Models of such accuracy included 7, 8, 9 or 10 variables. In total, for the collection has been selected 1956 models, which used 67 descriptors.

The collection of selected models includes one with 10 variables, in which the correlation coefficient r between the calculated and experimentally measured activity values is equal to 0.9532, and the standard deviation s = 0.15 logarithmic units (see Table 1). The expression (1) for this model takes the following form:

$$\log (EC_{50}) = (0.474\pm0.305) - (1.151\pm0.364)n_{O2} - (0.712\pm0.338)k_{Cl,N4\text{-}10} - (0.582\pm0.084)k_{Cl,C2\text{-}7} - (0.49\pm0.101)k_{Br,C2\text{-}7} - (0.276\pm0.111)k_{O1,C2\text{-}5} - (0.175\pm0.117)k_{Cl,C1\text{-}7} - (0.115\pm0.042)k_{F,C2\text{-}5} + (0.245\pm0.228)k_{Br,C3\text{-}1} + (0.285\pm0.034)k_{C2,C2\text{-}1} + (1.304\pm0.313)k_{Cl,C3\text{-}11}. \quad (2)$$

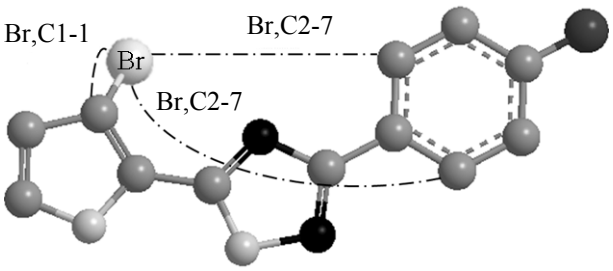
The descriptors O2; C1,N4-10; Cl,C2-7; Br,C2-7; O1,C2-5; Cl,C1-7 and F,C2-5 give negative contributions into the relation (2), that is, the presence of corresponding fragments in the structure a decrease in active concentration. These descriptors are called favorable for the activity of the apoptosis inducers. Conversely, the descriptors with positive parameters in Eq. (2) are unfavorable for this type of activity.

Consider the structural meaning of descriptors selected for the best model. The descriptors O2 (the number of ether oxygen atoms) and C1,N4-10 (the number of pairs "chlorine atom–nitrogen atom of amine group separated by 10 chemical bonds") have the greatest negative parameters in Eq. (2). Note that these descriptors have nonzero values only for a small number of compounds of this series (for 3 and 2 compounds, respectively). In addition, it is seen that with these descriptor well correlates the descriptor Cl,C3-11 (the number of the pairs: chlorine atom–aliphatic carbon atom separated by 11 chemical bonds) with a positive parameter: in compounds with non-zero Cl,C3-11 descriptor almost always (only exception is the compound number 7) there is a contribution or either O2, or C1,N4-10 descriptors. Therefore, to identify separately the role of each of these three descriptors for this set of compounds is impossible.

Negative contributions of the descriptors Cl,C2-7 (number of pairs of chlorine–aromatic carbon atoms separated by seven chemical bonds), and Br,C2-7 (number of pairs of bromine–aromatic carbon atoms separated by seven chemical bonds) to the relation (2) indicate that the presence of a halogen atom or a thienyl or furyl substituent enhances the compound activity. The absolute value of $P_{Cl,C2\text{-}7}$ is larger than $P_{Br,C2\text{-}7}$ indicating that the chlorine atom in the 5-aryl substituent is slightly more preferable than the bromine atom.

The descriptor Cl,C1-7 gives a small negative contribution in the case of the compounds containing chlorine atom in the 3-aryl substituent, because only in this case there is a chain of 7 bonds to a carbon atoms of each of five-membered ring. Such a structure can be found in both active (e.g., nos. 16, 24, 29, and 32, Table 1) and inactive (e.g., nos. 1, 7, and 40) compounds of the considered set.

The descriptor O1,C2-5 (the number of pairs: oxygen atom of the five-membered ring-aromatic carbon atoms separated by five chemical bonds) gives

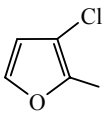
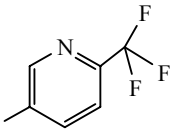
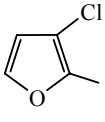
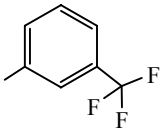
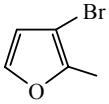
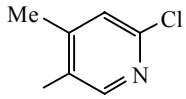
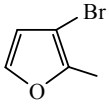
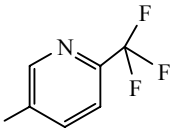
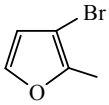
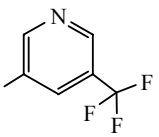


Molecular structure of 3-(4-chlorophenyl)-5-(3-bromothiophen-2-yl)-1,2,4-oxadiazole. Dashed lines connect the atoms in the pairs corresponding to the descriptor Br,C1-1 and Br,C2-7.

two contributions from the oxygen oxadiazole group in all compounds of the xonsidered set, and an additional contribution when the compounds has a furyl substituent. The presence of CF_3 group is also favorable for the activity, giving three negative contribution of the C2,F-5 descriptor to the expression (2).

The descriptor C2,C2-1 (the number of aromatic CC bonds) has several positive contributions in all the compounds of the consideration set. The largest number of aromatic CC bonds are found in the low-active compounds, no. 40 and no. 41, that is, where two aromatic substituents are phenyl derivatives. Descriptor C2,C2-1 can also serve as an indicator of the pyridyl substituent in the 3 position of 1,2,4-oxadiazole. In such compounds the value $k_{C2,C2-1}$ by 2 units less than in compounds with a phenyl substituent, and compounds with pyridyl substituent tend to have lower values of EC_{50}, that is, they are more active (compare compounds nos. 4 and 23; nos. 4 and 24; nos. 10 and 26; nos. 28 and 29, nos. 34 and 35). The exceptions are the pairs no. 10 and no. 25, no. 32 and no. 33, in which the meta-position of nitrogen atom to the substituent may affect their activity: it can be seen that the compounds are more active when the N atom in *ortho*-position to the substituent (cf. nos. 23 and 24, nos. 25 and 26).

Table 2. Derivatives of 3-aryl-5-aryl-1 ,2,4-oxadiazoles, designed on the basis of calculation results, and their predicted activity

Run no.	Ar_1	Ar_2	log EC_{50}[a]	EC_{50}, μM
1			–0.45	0.36
2			0.12	1.33
3			0.08	1.21
4			–0.26	0.55
5			–0.26	0.55

[a] The values calculated by the relation (2).

We can assume that the positive contribution from the descriptor Br,C3-1, which is present only in compounds nos. 17 and 38, points that the presence of $BrCH_2$ group in the 3-phenyl substituent of oxadiazole is undesirable.

The high frequency of occurrence of a descriptor in the models of the entire collection in our opinion can evidence the importance of an appropriate structural element for the activity. Among the above descriptors, four descriptors, C2,C2-1; O2; C1,N4-10 and Cl,C3-11 occur in more than 50% models. The descriptor Br,C1-1 also occurs frequently. On detailed inspection, it turns out to be close in meaning with the Br,C2-7 descriptor of the best model. Both these descriptors indicate the presence of Br atoms in a particular situation: in the studied set of compounds only the bromine atom in thienylnom or furyl residues contains a carbon atom of the five-membered ring at a distance of one bond, and an aromatic carbon atom at a distance of seven bonds (see the figure). Thus, the importance of the bromine atom in 5-aryl substituent is confirmed by the analysis of the entire collection of models. Another two descriptor of the best model, C2,F-5 and Cl,C2-7 occur in not less than 30% of the models of collection, which also shows the importance of associated structural elements: chlorine atom in the 5-aryl substituent and the fluorinated groups in the 3-aryl substituent.

Summing up all the said above about the role of the descriptors in the best model, we propose the following modification of the structure of 3,5-diaryl-1,2,4-oxadiazoles to increase the activity: (1) using of

Cl- or Br-furyl substituent in the 5-position, (2) introducing CF_3 groups in the Ar_2 substituent, (3) preference for pyridyl substituent in 3 position to reduce the undesired positive contributions of the C2,C2-1 descriptor. Within the chemical class of compounds shown in Table 1, we can recommend the new compounds with relatively high activity (Table 2).

Thus, using fragment descriptors and linear models obtained by regression method, it was possible to calculate the activity of the apoptosis inducers involving the caspase, with good accuracy. Analysis of the role of descriptors and their corresponding elements of molecular structure, allowed us to propose new compounds, which, according to the models will have high activity. It should be noted, however, that the set of compounds the inducers of apoptosis, which were studied experimentally [14, 15], contains insufficient systematic variants in molecular structure to get a model that would reveal the role of ether oxygen atom, amino group and aliphatic groups.

REFERENCES

1. Varmus, H., S*cience*, 2006, vol. 312, no. 5777, p. 1162.
2. Vial, J.-Ph., Belloc, F., Dumain, P., Bernard, S., Lacombe, F., Boisseau, M.R., Reifers, J., and Bernard, Ph., *Leukemia Res.,* 1997, vol. 21, no. 2, p. 163.
3. Dessalev, N., *Med. Chem. Res*., 2007, vol. 16, nos. 7–9, p. 449.
4. Matysiak, J., *QSAR Comb. Sci*., 2008, vol. 27, no. 5, p. 607.
5. Chang, J., Lei, B., Li, J., Li, Sh., Shen, Y., and Yao, X., *QSAR Comb. Sci*., 2008, vol. 27, nos. 11–12, p. 1318.
6. Tsygankova, I.G. and Zhenodarova, S.M., *Zh. Obshch. Khim*., 2008, vol. 78, no. 9, p. 1529.
7. Saiz-Urra, L., Pérez-Castillo, Y., González, M.P., Ruiz, R.M., Cordeiro, M.N.D.S., Rodriguez-Borges, J.E., and Garcia-Mera, X., *QSAR Comb. Sci*., 2009, vol. 28, no. 1, p. 98.
8. Singh, R., Jain, A., Ravichandran, V., Mourya, V., and Agraval, R.K., *Med. Chem. Res*., 2009, vol. 18, no. 7, p. 523.
9. Tsygankova, I.G. and Zhenodarova, S.M., *Zh. Obshch. Khim*., 2011, vol. 81, no. 5, p. 825.
10. Golovanov, I.B. and Zhenodarova, S.M., *Zh. Obshch. Khim*., 2003, vol. 73, no. 1, p. 90.
11. Golovanov I.B. and Tsygankova, I.G., *Zh. Obshch. Khim*., 1999, vol. 69, no. 8, p. 1275.
12. Golovanov, I.B. and Tsygankova, I.G., *Quantitative Structure Activity Relationship (QSAR),* 2000, vol. 19, p. 554.
13. Jessen, K.A., English, N.M., Maliartchouk, S., Archer, S.P., Qui, L., Brand, R., Kuemmerle, J., Zhang, H.-Zh., Gehisen, K., Dreve, J., Tseng, B., Cai, S.X., and Kasibhatla, S., *Mol. Cancer Ther.,* 2005, vol. 4, no. 5, p. 761.
14. Zhang, H.-Z., Kasibhatla, S., Kuemmerle, J., Kemnitzer, W., Ollis-Mason, K., Qiu, L., Crogan-Grundy, C., Tseng, B., Dreve, J., and Cai, S.X., *J. Med. Chem*., 2005, vol. 48, no. 16, p. 5215.
15. Kemnitzer, W., Kuemmerle, J., Zhang, H.-Z., Kasibhatla, S., Tseng, B., Dreve, J., and Cai, S.X., *Biorg. Med. Chem. Lett*., 2009, vol. 19, no. 15, p. 4410.
16. Cai, S.X., Nguyen, B., Jia, S., Herich, J., Guastelle, J., Reddy, S., Tseng, B., Dreve, J., and Kasibhetla, S., *J. Med. Chem*., 2003, vol. 46, no. 12, p. 2474.
17. Tsygankova, I.G., *Current Computer-Aided Drug Design*, 2008, vol. 4, no. 2, p. 132.