Comparison of QSAR Methods (CoMFA, CoMSIA, HQSAR) of Anticancer 1-N-Substituted Imidazoquinoline-4,9-dione Derivatives

Myung-Eun Suh,* So-Young Park, and Hyun-Jung Lee

Division of Medicinal Chemistry, College of Pharmacy, Ewha Womans University, Seoul 120-750, Korea
Received November 8, 2001

Comparison studies of the Quantitative Structure Activity Relationship (QSAR) methods with new imidazoquinolinedione derivatives were conducted using Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA), and the Hologram Quantitative Structure Activity Relationship (HQSAR). When the CoMFA crossvalidation value, q^2, was 0.625, the Pearson correlation coefficient, r^2, was 0.973. In CoMSIA, q^2 was 0.52 and r^2 was 0.979. In the HQSAR, q^2 was 0.501 and r^2 was 0.924. The best result was obtained using the CoMSIA method according to a comparison of the calculated values with the real *in vitro* cytotoxic activities against human ovarian cancer cell lines.

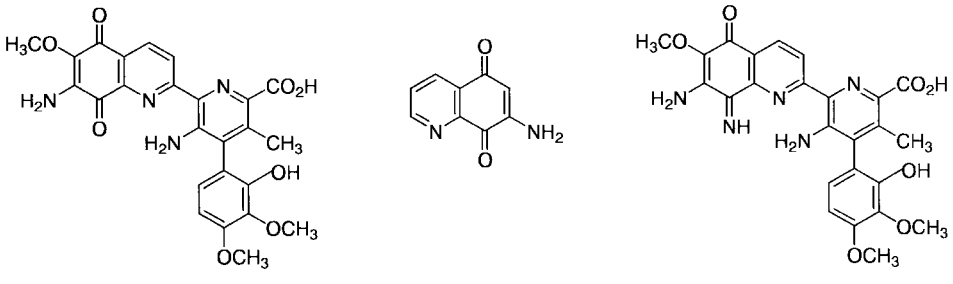
Keywords : 1-N-Substituted Imidazoquinoline-4,9-dione Derivative, Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA), Hologram Quantitative Structure Activity Relationship (HQSAR), Human ovarian cancer cell.

Introduction

Streptonigrin (Figure 1), which has both antitumor and antibiotic activity, was isolated from *Streptomyces flocculus* in 1959.[1] Recently, Johnson[2] reported that the streptonigrin pharmacophore, 7-amino-6-methoxy-5,8-quinoline,[3] in which the pyridyl and its substituted phenyl rings were eliminated, showed more antitumor activity and was less toxic than streptonigrin in avain myeloblastosis virus reverse transcriptase (AMV-RT).[4] Although streptonigrin is one of the more effective anticancer drugs with good selectivity, it has limited use due to serious bone marrow toxicity.[5-7] Therefore, many studies have been conducted to reduce its toxicity.

Studies on the activity of heterocyclic quinones containing nitrogen atoms such as quinolinedione revealed that there is a relationship between the number and position of the nitrogen atoms and its cytotoxicity.[8] Some important Structure and Activity Relationships (SAR) have been reported.[9] The antitumor activity of streptonigrin (**I**) is completely lost when the aminoquinone moiety (**II**) is blocked as in azastreptonigrin (**III**).[8] The methoxy group (quinone ring), the pyridyl and its substituted phenyl rings are not essential for the activity in murine tumors, although they exhibit enhanced activity against human tumor cells.[9] The synthetic analogues without the 7-aminoquinolinequinone moiety (**III**) are also inactive as antitumor agents (Figure 1). The electron withdrawing groups at the 6 and 7 positions of the quinolinediones also contribute to the activity[10] and more condensed heterocyclic quinones have been reported to have increased antitumor activity.[11] Kuo *et al.*[12] also reported that 1-ethyl-2-methylnaphth[2,3-*d*]imidazole-4,9-dione had excellent cytotoxicity on human ovarian cancer cell lines.

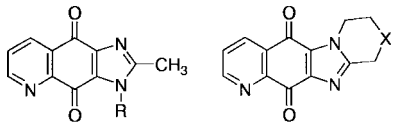
Recently, we synthesized a series of 1-N-substituted imidazoquinoline-4,9-dione derivatives as prodrugs of anticancer agents shown in Table 1.[13] Structure of theses compounds has the required positions to be DNA intercalators according to Moore's theory.[14] Computer aided molecular modeling is able to assist in predicting both the cytotoxic activities and toxicity. Nowadays, many medicinal chemists usually use Quantitative Structure Activity Relationships (QSAR) methods because they might minimize the number of compounds that synthetic chemists need to prepare and the time needed to discover new drug candidates. Essentially, correlating the physicochemical properties of the compounds to their respective cytotoxic activities is believed to provide a useful tool in designing new drugs.



I **II** **III**

Figure 1. Structures of streptonigrin (**I**), aminoquinone moiety (**II**), and azastreptonigrin (**III**).

Table 1. Fuctional groups of the Imidazoquinoline-4,9-dione derivatives



A B

Structure No.	R	Structure No.	R (or X)
A1	$-CH_3$	**A10**	$-CH_2CH_2CH_2CH_3$
A2	$-p-C_6H_5NO_2$	**A11**	$-CH_2CH_2Cl$
A3	$-p-C_6H_5Cl$	**A12**	$-CH_2C_6H_5$
A4	$-p-C_6H_5Br$	**A13**	$-CH_2CH_2OH$
A5	$-p-C_6H_5OC_2H_5$	**A14**	$-CH(CH_3)_2$
A6	$-p-C_6H_5CH_3$	**A15**	$-C_6H_5$
A7	$-p-C_6H_5CF_3$	**B1**	$-CH_2-$
A8	$-CH_2CH_3$	**B2**	-O-
A9	$-CH_2CH_2CH_3$	**B3**	-S-

Quantitative Structure-Activity Relationships (QSAR) based on the 2 dimensional (2D) or 3D structures of the ligands alone, and involves three methods: Hologram QSAR (HQSAR), Comparative Molecular Field Analysis (CoMFA), and Comparative Molecular Similarity Indices Analysis (CoMSIA). Until 10 years ago, the majority of descriptors used in such correlation studies were substituted parameters representing specific properties of a functional group. This type of model is classified as 2D QSAR since the descriptors do not capture any 3D information concerning the ligands and the specific conformation of the molecules. A new 2D technique, namely, HQSAR has recently been introduced.[15,16] In this method, the chemical structure is converted to a characteristic molecular fingerprint based on enumerating the presence of certain types of molecular fragments. This numerical representation of the molecules is used as the QSAR descriptor.

Recently, more advanced techniques have been used in the attempt to model the receptor environment from the perspective of the ligand structure. Quantitative Structure-Activity Relationship (QSAR) studies incorporate the three-dimensional information for the ligands and provides a more detailed analysis of ligand-receptor interactions. The CoMFA program[17] places the drug molecules with a steric or an electrostatic probe at evenly spaced grid points. The CoMSIA program[18] is known as one of the new 3D QSAR descriptors. In CoMSIA, both the steric and electrostatic features, hydrogen bond donor, hydrogen bond acceptor and hydrophobic fields are considered. The 3D method provides the conformation or spatial orientation of molecules. In addition, they supply good information for designing new compounds or potential drug candidates. The biological cytotoxicity of 1-N-substituted Imidazoquinoline-4,9-dione derivatives *in vitro* were compared with their predicted values from the QSAR methods-CoMFA, CoMSIA, and HQSAR.

Computational Methods

Molecular 3D Structure Building. Structures of the entire 1-N-substituted Imidazoquinoline-4,9-dione analogue set were built using the Sybyl 6.5 version Molecular Modeling Software. Structural energy minimization was performed using the standard Tripos molecular mechanics force field and Gasteiger-Hueckel charge, with a 0.001 kcal/mol energy gradient convergence criterion on a Silicon Graphics IRIS O2 R 5000 computer system.

Methods of QSAR Analysis. Low energy conformation was investigated using a systematic and grid conformational search. All the structures generated were aligned in a 3D lattice by fitting them with imidazoquinoline-4,9-dione as a common structure.

In this report, the r^2 and q^2 values were measured. The r^2 value is the Pearson correlation coefficient, which is the correlation between the calculated activities and the observed cytotoxic activities. The q^2 is the predicted value based on a leave-one-out (LOO) cross-validation method.[19] The three QSAR methods used are ligand-based QSAR techniques. In this study, the CoMFA, the CoMSIA, and HQSAR modules in Sybyl (version 6.5, Tripos Inc.) were employed.

$$q^2 = 1 - \frac{\sum_{i=1}^{N} (y_{i\ \text{observed}} - y_{i\ \text{predicted}})^2}{\sum_{i=1}^{N} (y_{i\ \text{observed}} - y_{i\ \text{observed}})^2}$$

(1) Comparative Molecular Field Analysis (CoMFA): Comparative Molecular Field Analysis (CoMFA) is one of the more famous 3D QSAR methods. It provides steric and electrostatic values in addition to ClogP values. ClogP means the hydrophobic parameters of the ligands.

In CoMFA analysis, the ligands are placed in a 3D lattice and then the steric and electrostatic fields of the ligands at the various grid points of the lattice are calculated. The resulting field matrix is analyzed by the Partial Least Squares (PLS). The 3D lattice was set up as a $22 \times 16 \times 19$ Å^3 lattice with a 1 Å grid spacing for both the steric and electrostatic fields, the default truncation cutoff was set as 30 kcal/mol.

(2) Comparative Molecular Similarity Indices Analysis (CoMSIA): Comparative Molecular Similarity Indices Analysis (CoMSIA) is known as one of the newer 3D QSAR descriptors. CoMSIA was developed at BASF Ludwigshafen, Germany by Klebe *et al.* This technique is most commonly used in drug discovery to find the common features that are important in binding to the relevant biological receptor.

In CoMSIA, both steric and electrostatic features, hydrogen bond donor, hydrogen bond acceptor and hydrophobic fields are considered.

The equation used to calculate the similarity indices[20] is

$$A^q_{F.K.(j)} = -\Sigma\, W_{probe,k}\, W_{ik}\, e^{-\alpha r^2_{iq}}$$

where A is the similarity index at grid point q, summed over all atoms, i, of the molecule j. $W_{probe,k}$ is the probe atom with a radius 1 Å charge +1, hydrophobicity +1, hydrogen bond donating +1, hydrogen bond accepting +1. W_{ik} is the actual value of the physicochemical property, k, of atom i. r_{iq} is the mutual distance between the probe atom at grid point q and

atom i of the test molecule. α is the attenuation factor. A larger value results in a steeper Gaussian function and a strong attenuation of the distance-dependent effects of molecular similarity.

(3) Hologram QSAR (HQSAR): Hologram QSAR is a unique QSAR method. This method does not require the exact 3D information for the ligands. In this study, the molecule is hashed to a molecular fingerprint that encodes the frequency of the occurrence of various molecular fragment types. In other words, the fragment size controls both the minimum and maximum length of the fragments to be included in the hologram fingerprint. Molecular holograms are produced by generating all the linear and branched fragments, which range in size from 4 to 7 atoms.

In the SYBYL HQSAR mode (version 6.5 Tripos Inc.), fragments can be distinguished based on the atoms, bonds, connections, number of hydrogen atoms and chirality parameters. HQSAR works by identifying the patterns of the substructual fragments related to cytotoxic activity in sets of bioactive molecules. The cytotoxic activity of each fragment allows a prediction of the cytotoxic effect of the molecules. 12 default hologram lengths that have been found to yield predictive models on a number of test data sets are provided.

Results and Discussion

The training set was composed of 18 synthesized compounds (**A1-15, B1-3**). The CoMFA, CoMSIA, and HQSAR were used to estimate the activities against Human Ovarian cancer cell lines as a dependent column. In the case of HQSAR, the fragment information was composed of atoms, bonds and connections. The best hologram length was found to be 59.

The results showed a CoMFA q^2 value of 0.625 and an r^2 value of 0.973 (Table 2). In CoMSIA, the q^2 value was 0.520 and the r^2 value was 0.979. In HQSAR, the q^2 value was 0.501 and the r^2 was 0.924. In Table 2, all the crossvalidation values, q^2, are available because the crossvalidated Pearson correlation coefficient, r^2, has some accuracy if $q^2 > 0.5$.

The Pearson correlation coefficient, r^2 shows how much the predicted activity approximated the cytotoxic activity *in vitro*. The best r^2 value was 0.979, which means that the CoMSIA results have a 97.9% precision level compared with the cytotoxic results in the test. Therefore, CoMSIA is quite reliable for predicting the antitumor activities in the 1-N-substituted imidazoquinoline-4,9-dione derivatives.

Table 2. Summary of the CoMFA, CoMSIA, HQSAR output

	CoMFA	CoMSIA	HQSAR
Opt. Number of components	4	6	4
Crossvalidation q^2	0.625	0.52	0.501
Conventional r^2	0.973	0.979	0.924
Standard error of estimate	0.11	0.10	–
F value	(n1 = 4, n2 = 11) 98.244	(n1 = 6, n2 = 9) 70.669	–
Probe atom	C (sp^3, +1)		

Table 3. Relative contributions of the CoMFA

Relative contributions	
CoMFA (steric)	0.500
CoMFA (electrostatic)	0.262
ClogP	0.237

Table 4. Relative contributions of the CoMSIA

Relative contributions	
CoMSIA(steric)	0.128
CoMSIA(electrostatic)	0.192
CoMSIA(hydrophobic)	0.530
CoMSIA(acceptor-donner) Steric	0.033
CoMSIA(acceptor-donner) Electrostatic	0.118

(1) CoMFA: Table 3 shows the relative contributions to the CoMFA analysis. The optimun value of the crossvalidated r^2 for 10 components was 0.625 for 4 components. In this analysis, the standard estimation error was 0.11. r^2 was 0.973 and the F value was 98.244 (n1 = 4, n2 = 11).

(2) CoMSIA: Table 4 shows the relative contributions to the CoMSIA. The optimum value of the crossvalidated r^2 for 10 components was 0.520 for 6 components. In this the study, the standard error of estimation was 0.10, r^2 was 0.979 and F value was 70.669 (n1 = 6, n2 = 9).

(3) HQSAR: In the HQSAR method. r^2 was 0.924 and q^2 was 0.501.

The results of the CoMFA, CoMSIA and HQSAR analysis are shown in Table 5 with a comparison of the predicted activities with the actual cytotoxic activities. In this study, 18 compounds were analyzed. However, compounds **A2** and **B1** were omitted in the CoMFA and CoMSIA, and compounds **A8** and **A11** were excluded in the HQSAR method. This is because the results of the factor analysis show that these compounds had factors that exhibited an inaccurate influence on the QSAR methods.

In the electrostatic CoMFA map (Figure 2), the red color showed that groups in that region with greater electronegativity could confer better activity. In the steric CoMFA map (Figure 2), the large green colored area around the substituted group of the template molecule indicated that a bulky group at the position could enhance the cytotoxicity.

In the electrostatic and steric map (Figure 3), the red color indicated greater electronegativity. In the hydrophobic and hydrogen bond CoMSIA map (Figure 4), the hydrophobic region is yellow, the hydrophilic region is gray and the purple color indicated that hydrogen bonding acceptor groups at that region could confer better activities.

In the E_CoMFA, E_CoMSIA and E_HQSAR values (the E denotes the error.), each value indicates how well the CoMFA, CoMSIA, and HQSAR values approximate the biological values, which were tested with human ovarian cells. In the E_CoMFA, compounds **A1**, **A3**, **A7**, **A9** and

Table 5. The results of the CoMFA, CoMSIA and HQSAR analysis, and the error values with a comparison with each biological values

No.	Poval ($-\log IC_{50}$)	ClogP	Predicted Activity			E_CoMFA	E_CoMSIA	E_HQSAR
			CoMFA	CoMSIA	HQSAR			
A1	0.54	−0.26	0.53	0.64	0.51	0.01	0.10	0.02
A2	–	–	–	–	−0.56	–	–	0.10
A3	1.05	2.34	1.06	0.99	1.03	0.01	0.06	0.03
A4	1.00	2.49	1.12	1.03	1.09	0.12	0.03	0.07
A5	0.66	2.24	0.54	0.64	0.77	0.12	0.02	0.11
A6	1.00	2.12	1.10	1.07	1.09	0.10	0.06	0.10
A7	0.44	2.51	0.45	0.47	0.46	0.01	0.02	0.01
A8	0.85	0.27	0.75	0.69	–	0.10	0.16	–
A9	0.68	0.80	0.67	0.65	0.49	0.01	0.03	0.19
A10	0.70	1.33	0.75	0.64	0.51	0.05	0.06	0.20
A11	−0.30	0.34	−0.28	−0.36	–	0.02	0.05	–
A12	0.74	1.31	0.65	0.75	0.79	0.01	0.01	0.04
A13	−0.29	−1.09	−0.21	−0.29	0.20	0.08	0.00	0.49
A14	0.22	0.58	0.30	0.42	0.15	0.08	0.20	0.07
A15	1.10	1.62	0.95	1.04	0.87	0.15	0.06	0.23
B1	–	–	–	–	1.62	–	–	0.12
B2	1.70	−1.27	1.58	1.68	1.59	0.12	0.02	0.13
B3	1.70	−0.38	1.89	1.71	1.67	0.19	0.01	0.07

Poval (-log(oval)) : Log value of activity against human ovarian cancer cell lines. ClogP: hydrophobic parameter. E_CoMFA: E_CoMFA is found the difference between Poval and CoMFA. (E is abbreviation of Error.). E_CoMSIA: E_CoMSIA is calculated the difference between Poval and CoMSIA. E_HQSAR: E_HQSAR is analyzed the difference between Poval and HQSAR.

Table 6. Summary of the error values of the CoMFA, CoMSIA, and HQSAR analysis

	E_CoMFA	E_CoMSIA	E_HQSAR
Mean	0.07	0.05	0.12
Standard-deviations	0.06	0.05	0.12
High	0.15	0.20	0.49
Low	0.01	0.00	0.01

A12 had lower values than any other compound. The difference was found to be 0.01. In the E_CoMSIA. compound **A13** showed no difference between the calculated and the biological value tested with human ovarian cells. In addition. compounds **A5**. **A7**, **A12**, **A17** and **A18** had E_CoMSIA values that were similar to the biological values. Therefore. the CoMSIA method was the most accurate method. With E_HQSAR. compound **A7** had the closest match to the biological value tested with human ovarian cells with an approximate value of 0.01.

When E_CoMFA, E_CoMSIA, and E_HQSAR were calculated to obtain a mean value. the E_CoMSIA mean of 0.05 was the lowest (Table 6). In the table, the standard deviations of E_CoMFA and E_CoMSIA are similar. The highest standard deviation was from E_HQSAR. In conclusion. when 1N-subsituted Imidazoquinoline-4,9-dione derivatives are synthesized, CoMSIA is quite useful for predicting the cytotoxic activities.

In the electrostatic and steric CoMFA map (Figure 2) and the electrostatic and steric CoMSIA map (Figure 3), the large yellow colored area around the substituted group of the template molecule indicated that the less bulky group in this position. 1N-position of imidazoquinoline-4,9-dione. could result in the more powerful antitumor activities. In the hydrophobic and hydrogen bonding CoMSIA map (Figure 4), hydrogen boding acceptors in the 1N and 3N-position of imidazoquinoline-4,9-dione would confer better activities. According to this result, in this study. we designed and synthesized the new series of pyrroloquinoline-4,9-dione (**P-III**) which were possessing aliphatic groups (propyl. methoxyethyl, hydroxyethyl. ect.) at 1N-position and ethoxycarbonyl group at 3-position. And then CoMFA and CoMSIA were used to predict the cytotoxicity of these unknown compounds. The predicted activities were compared to the cytotoxicity against human ovarian cnacer cell lines (SK-OV-3) and the results were presented that CoMSIA analysis provided the best result (Tables 7 and 8). Even though the cytotoxic activities are not absolutely correct. predicting the cytotoxic activity using QSAR methods is important as the exact mechanisms and effects of these compounds in the human body are unknown.

Conculsion

The 3D QSAR analysis, CoMFA, CoMSIA and HQSAR were available for the imidazoquinolinedione derivatives to predict their biological activity. The biological activities of the unknown samples. with pyrroloquinolinedione derivatives are easily predicted with 3D QSAR analysis. It is also useful to make a plan to synthesize new compounds. pyrroloquinolinedione-4,9-dione derivatives, with good biological activities.

Acknowledgment. This work was supported by the Faculty Research Grant from Ewha Womans University and the Ministry of Science and Technology (Grant # 99-N6-01-A-02 in 1999). We are grateful for the financial support.

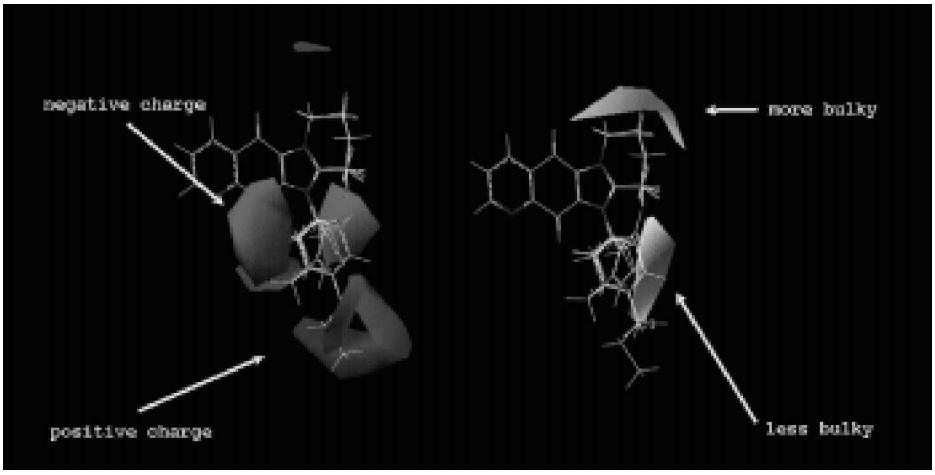


Figure 2. The electrostatic and steric CoMFA map. Red color is negative charge region, blue is positive charge region, green is more bulky region, and yellow is less bulky region.

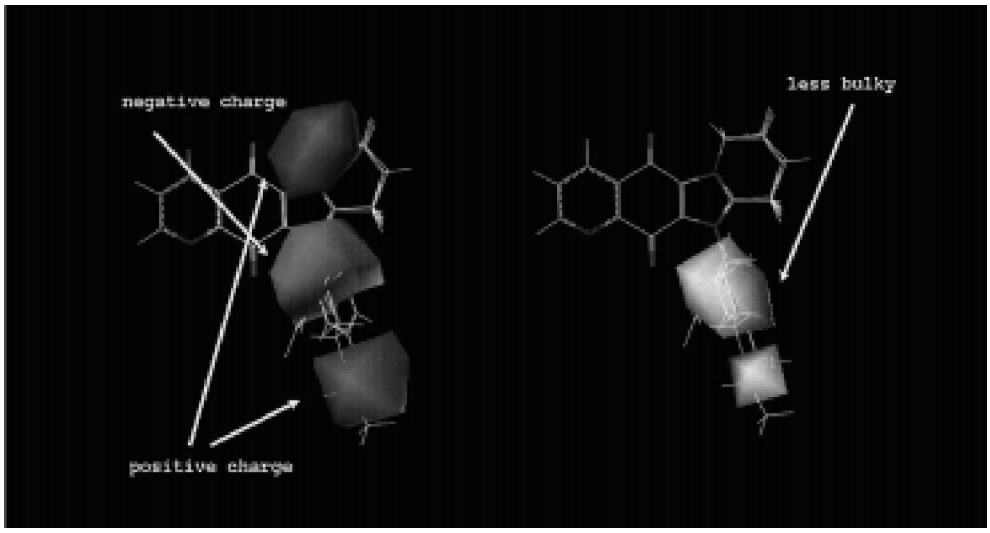


Figure 3. The electrostatic and steric CoMSIA map. Red color is negative charge region, blue is positive charge region and yellow is less bulky region.

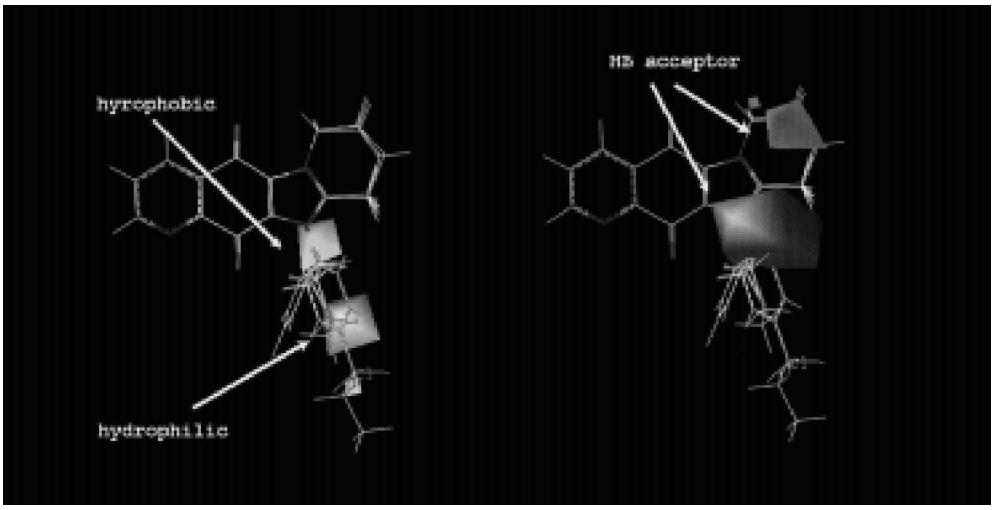
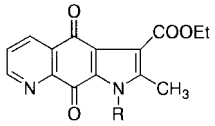


Figure 4. The hydrophobic and hydrogen bonding CoMSIA map. Hydrophobic region is yellow, hydrophilic region is gray and hydrogen acceptor region is purple.

Table 7. The results of the CoMFA and CoMSIA analysis, and the error values with a comparison with each biological values of unknown compounds



P-III

Name	R	Poval ($-\log IC_{50}$)	Predictd activity (CoMFA)	E_CoMFA	Predicted activity (CoMSIA)	E_CoMSIA
P-IIIc	$-(CH_2)_2CH_3$	0.85	1.25	0.40	0.82	0.02
P-IIIe	-cyclopropyl	0.27	0.84	0.55	0.43	0.16
P-IIIf	$-C_2H_4OCH_3$	1.05	0.88	0.17	0.70	0.35
P-IIIg	$-(CH_2)_2OH$	0.11	1.27	0.16	0.37	0.26
P-IIIh	$-CH_2-C_6H_5$	0.77	0.97	0.20	0.84	0.07
P-IIIi	-furfurylethyl	0.39	1.27	0.88	0.54	0.15

Poval ($-\log IC_{50}$) : Log value of activity against human ovarian cancer cell lines. E_CoMFA: E_CoMFA is found the difference between Poval and CoMFA. (E is abbreviation of Error.). E_CoMSIA: E_CoMSIA is calculated the difference between Poval and CoMSIA.

Table 8. Summary of the error values by the CoMFA and CoMSIA of unknown compounds

	E_CoMFA	E_CoMSIA
Mean	0.39	0.17
Standard-deviations	0.28	0.12
High	0.88	0.35
Low	0.16	0.02

References

1. Rao, K. V.; Cullen, W. P. *Antibiot. Ann.* **1968**, *950*, 1959.
2. Johnson, F.; Shaikh, I. A.; Grollmann, A. P. *J. Med. Chem.* **1986**, *29*, 1329.
3. Inouye, Y.; Take, Y.; Oogose, K. *J. Antibiotics* **1987**, *40*, 105.
4. Rao, K. V. *Cancer Chemother. Rep.* **1974**, part 2, *4*, 11.
5. Hackethal, C. A.; Golbey, P. B.; Tan, C. T. C.; Karnofsky, D. A.; Burchenal, J. H. *Antibiotic Cheother.* **1961**, *11*, 178.
6. Humphrey, E. W.; Blank, N.; Medrek, T. J. *Cancer Chemother. Rep.* **1961**, *12*, 99.
7. Rivers, S. L.; Wittington, R. M.; Spencer, H. H.; Pento, M. E. *Cancer* **1966**, *19*, 1377.
8. Kremer, W. B.; Laszlo, J. *Cancer Chemotheraphy Rep.* **1967**, *51*, 19.
9. Foye, W. O. *Cancer Chemotherapeutic Agents*; American Chemical Society press: Washington, U.S.A., 1995; 645.
10. Lown, J. W.; Joshus, A. V.; Lee, J. S. *Biochemistry* **1982**, *21*, 419.
11. Yamashita, Y.; Tsubata, Y.; Suzuki, T.; Miyashi, T.; Mukai, T.; Tanaka, Y. *Chem. Lett.* **1990**, *1990*, 445.
12. Kuo, S. C.; Ibuka, T.; Huang, L. J.; Lednica, D. *J. Med. Chem.* **1996**, *39*, 1447.
13. Suh, M. E.; Kang, M. J.; Yoo, H. W.; Park, S. Y.; Lee, C. O. *Bioorg. & Med. Chem.* **2000**, *8*, 2079.
14. Moore, M. H.; Hunter, W. H.; Kennard, O. *J. Mol. Biol.* **1987**, *206*, 693.
15. *HQSAR*., Version 1.0; Tripos. Inc.: St. Louis, MO, 1999.
16. Winkler, D. A.; Burden, T. R. *Quant. Struct. Act. Relat.* **1998**, *17*, 224.
17. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
18. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 24, 4130.
19. So, S. S.; Karplus, M. *J. Med. Chem.* **1997**, *40*, 4360.
20. Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; Wiley: New York, 1980.