Structural Chemistry, Vol. 9, No. 1, 1998

A Structure–Activity Relationship Study of Lapachol and Some Derivatives of 1,4-Naphthoquinones Against Carcinosarcoma Walker 256

S. Subramanian,[1] M. M. C. Ferreira,[2] and M. Trsic[3,4]

Received April 15, 1997; revised July 2, 1997; accepted July 19, 1997

AM1 semiempirical molecular orbital calculations were carried out on lapachol and several derivatives of 1,4-naphthoquinone in order to investigate possible relationships between electronic structural parameters and activity against carcinosarcoma Walker 256 (W256). It was found that, among the calculated electronic indices, that the HOMO (highest occupied molecular orbital) coefficients for carbon atoms of the side-chain double bond have a significant influence in the activity, while the LUMO (lowest unoccupied molecular orbitals) apparently have no importance. Exploratory data analysis through hierarchical cluster (HCA) and principal component analysis (PCA) showed a clear separation of the active compounds from the inactive ones. The activity against W256 probably involves a mechanism wherein the quinone acts as a reducing agent through the participation of the π-electrons of the side-chain double bond. A classification study with respect to structure-activity using KNN (K-nearest neighbors) and SIMCA (soft independent modeling of class analogy), two established chemometric methods of pattern recognition, have been used to predict activity for a series of lapachol derivatives.

KEY WORDS: MNDO-AM1-SCF-MO; Lapachol; PCA-HCA-KNN-SIMCA.

INTRODUCTION

Among the various natural quinones, the naphthoquinone lapachol [2-hydroxy-3-(3-methyl-2-butenyl)-1,4-naphthoquinone] (Fig. 1), extracted from the heartwood of certain Asian and South American bignoniaceous plants, has been investigated for various biological activities [1]. Lapachol and its analogues are known to possess antitumor, antibiotic, and antimalarial activities. Recently, it has been observed that lapachol also exhibits antiinflammatory and antiulceric action [2, 3]. It has also been shown that this drug is highly active against Walker 256 (W256) carcinosarcoma tumor in rats [1] when parenterally or orally administered. Clinically, lapachol is currently being used in Brazil in the

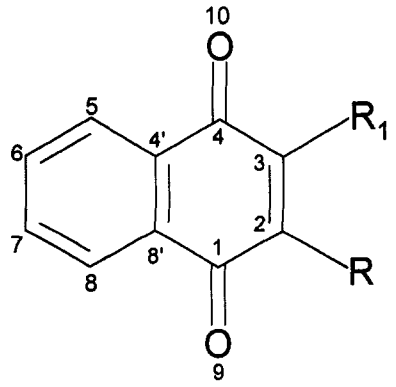


Fig. 1. Structural skeleton of 1,4-naphthoquinone derivatives listed in Table I.

[1]Departamento de Química, CCEN, Universidade Federal da Paraíba, João Pessoa, PB, 58051-900, Brazil.
[2]Departamento de Físico Química, Instituto de Química, Universidade Estadual de Campinas, Campinas SP, 13081-970, Brazil.
[3]Departamento de Química e Física Molecular, Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil.
[4]Correspondence should be directed to M. Trsic at Departamento de Química e Física Molecular, Instituto de Química de São Carlos, Universidade de São Paulo, C.P. 780, 13560-970, São Carlos, SP, Brazil; e-mail: MILAN@IQSC.SC.USP.BR.

1040-0400/98/0200-0047$15.00/0

Table I. Activities of Some 1,4-Naphthoquinone Derivatives Against W256 (Training Set)[a]

Compound number	R	R_1	Activity (TWI, %)
I	OH	$CH_2CH{=}C(Br)_2$	90 [9]
II	OH	$CH_2CH{=}C(Cl)_2$	85 [9]
III	OH	$CH_2CH{=}C(CH_3)_2$	79[b]
IV	OH	$CH_2CH{=}C(CH_3)HCH_2CH_2CH{=}C(CH_3)_2$	68[b]
V	OAc	$CH_2CH{=}C(CH_3)_2$	64 [9]
VI	OH	$CH_2CH{=}C(CH_2OH)CH_3$	58 [6]
VII	CH_3	$CH_2CH{=}C(CH_3)_2$	49 [9]
VIII	OH	$CH_2CH{=}C(C_2H_5)_2$	42 [9]
IX	H	$CH_2CH{=}C(CH_3)_2$	41 [9]
X	OH	$CH_2CH{=}C(Cl)CH_3$	41 [9]
XI	OCH_3	$CH_2CH{=}C(CH_3)_2$	39 [9]
XII	OH	$CH{=}CBrC(CH_3)_2$	34 [9]
XIII	OH	$CH{=}C(CH_3)_2$	25 [9]
XIV	OH	$CH{=}CHCH_2CH_3$	21 [9]
XV	OH	$CH_2CH{=}CH_2$	20 [9]
XVI	OH	$CH{=}CHCH_2CH_2CH_2CH_2CH_3$	20 [6]
XVII	OH	$CH_2CH(OH)C(OH)(CH_3)_2$	20 [9]
XVIII	NH_2	$CH_2CH{=}C(CH_3)_2$	19 [9]
XIX	OH	$CH_2CH_2CH(CH_3)_2$	17 [9]
XX	OH	$CH_2CH_2C(OH)(CH_3)_2$	10 [9]
XXI	OH	$CH_2CH{=}C(COOH)CH_3$	0 [9]
XXII	OH	$CH{=}CHCH(CH_3)_2$	0 [9]
XXIII	OH	$C(CH_3)_2CH{=}CH_2$	0 [9]
XXIV	OH	$CH{=}CHCH_3$	0 [9]
XXV	OH	$CH_2CH_2C(Cl)(CH_3)_2$	0 [9]
XXVI[c]	OH	$CH_2CH{=}C(CH_3)_2$	0 [9]

[a] See Fig. 1 for structures. Compounds with tumor weight inhibition (TWI) ≥ 58% are considered significantly active.
[b] Average value of refs. 6 and 9.
[c] Aromatic ring reduced.

treatment of adenocarcinoma and squamous carcinomas [4, 5].[5]

The high activity of lapachol against W256 prompted Hartwell and Abbott [6] to investigate the efficacy of 67 derivatives of 2-hydroxy-1,4-naphthoquinone. Of these, only the 3,7-dimethyl-2,6-octadienyl derivative showed activity comparable to that of lapachol. Since clinical tests proved lapachol to have undesirable side effects [7], Rao [8] carried out detailed studies on several of its derivatives with structural modifications at the following points: the hydroxyl group at the 2 position, the alkyl side chain at the 3 position, the aromatic ring, and the quinone system (see Fig. 1). Among the compounds investigated, I–VI (see Table I) proved to be significantly active against W256. The present work is an attempt to relate carcinosarcoma activity to the chemical structure of lapachol and its derivatives through quantum mechanical calculations and chemometric methods for this family of compounds.

[5] Current information can be obtained from Prof. Sueli Lins Galdino, Departamento de Antibióticos, Universidade Federal de Pernambuco, 50670-901, Recife, PE, Brazil; phone/FAX (55)(81) 271-8346.

THEORETICAL METHODS

Semiempirical molecular orbital calculations were carried out on 26 compounds (Table I), previously studied experimentally by Hartwell and Abbott [6], Rao [8], and Driscoll *et al.* [9]. The computer program used was MOPAC 6 obtained from Quantum Chemistry Program Exchange (QCPE). Among the semiempirical methods included in that package (MINDO/3, MNDO, AM1, and PM3), AM1 was chosen for all the calculations in this work. The choice was based on the fact that the AM1 method yielded the least average error for bond

lengths l and bond angles θ for the five compounds (lapachol, β-lapachone, 3-allyl-β-lapachone, 3-hydroxy-β-lapachone and 3-bromo-α-xyloidone) for which X-ray data are available in the literature [10], Δl and $\Delta\theta$ being 0.015 Å and 1.22°, on average, respectively.

The molecular indices calculated were E_{HOMO} (energy of the highest occupied molecular orbital), E_{LUMO} (energy of the lowest unoccupied molecular orbital), dipole moment (DM), heat of formation (ΔH_f), and the atomic charges. The frontier orbital coefficients of the atomic centers were also considered. The main idea was to correlate all these descriptors to the activity against W256 through multivariate methods. The data were analyzed through the hierarchical clustering technique (HCA) and principal component analysis (PCA) [11]. SIMCA (soft independent modeling of class analogy) and KNN (K-nearest neighbor), two well-established pattern recognition and classification modeling techniques, were used for model building and prediction. A model for the active and inactive compounds was created and validated using a training set, and in the next step it was used to classify the new objects of a test set by relating them to each class model. A set of sixteen 1,4-naphthoquinone derivatives was used as a test and classified as active/inactive compounds. The results are presented and discussed in the following section. All data analysis was carried out using Pirouette software [12]. Both KNN [13] and SIMCA [14] are similarity-based classification techniques. Classification with KNN is based on distance comparison among samples, and multivariate Euclidean distances between every pair of training samples are calculated. The model is built and the predicted class of a test sample is determined based on the multivariate distance of this sample with respect to K samples in the training set. The SIMCA method builds principal component models for each class in the training set and the classification of a test sample is made by comparing its proximity to the nearest class model.

RESULTS AND DISCUSSION

The experimental data of activity *in vivo* against W256 reported in the literature for 26 derivatives of 1,4-naphthoquinones are shown in Table I. According to Hartwell and Abbott [6] and Driscoll *et al.* [9], compounds with TWI (tumor weight inhibition) value $\geq 58\%$ are considered to be significantly active (I–VI), whereas those between 40% and 58% are deemed to be weakly active (VII–XI). Compounds with TWI below 40% are considered inactive (XII–XXVI).

In a first step, HCA and PCA were carried out for these 26 compounds (I–XXVI), using as descriptors the 25 molecular parameters obtained through theoretical calculations: E_{HOMO}, E_{LUMO}, DM, ΔH_f°, and atomic charges of 21 atoms as shown in Fig. 2. These parameters, with the omission of charges on some less relevant atoms, are shown in Table II. Autoscaling preprocessing (zero mean and unit variance) was applied in order to give each variable equal weight in the analysis. HCA and PCA analyses gave no useful separation of compounds into different classes according to their activities.

As a result of the lack of correlation for the various parameters, attention was then turned to the coefficients of the atomic centers, as $\Sigma\, c_i^2$ ($i = 2s, 2p_x, 2p_y, 2p_z$ for the atomic center considered for second-row atoms) of the frontier orbitals HOMO and LUMO. A detailed ex-

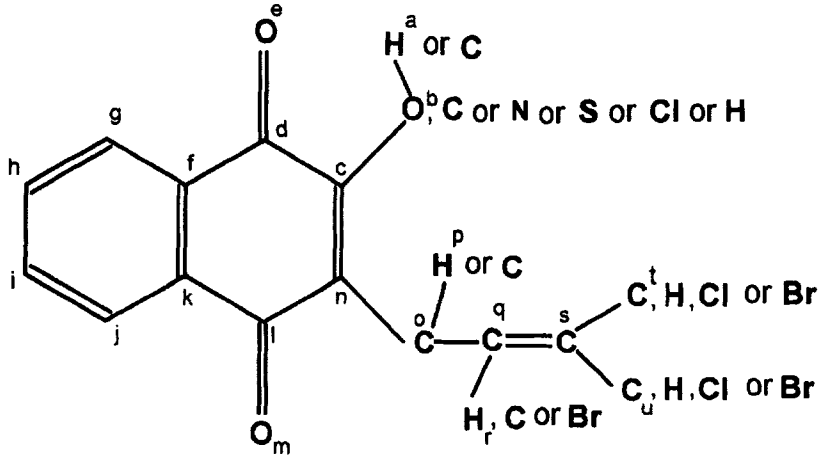


Fig. 2. Labels for the atomic centers in the molecules of 1,4-naphthoquinone derivatives.

Table II. Values of Certain Parameters Obtained Through Theoretical Calculations Using the AM1 Method for the Compounds in Table I

Compound number	Heat of formation (kcal/mol)	E_{HOMO} (eV)	E_{LUMO} (eV)	Dipole moment (Debye)	Atomic charge[a]					
					On *b*	On *e*	On *m*	On *o*	On *q*	On *s*
I	−42.559	−9.758	−1.826	4.64	−0.237	−0.281	−0.270	−0.103	−0.067	−0.320
II	−64.598	−9.697	−1.776	4.10	−0.238	−0.282	−0.271	−0.090	−0.134	−0.105
III	−65.803	−9.236	−1.611	2.58	−0.236	−0.287	−0.280	−0.084	−0.177	−0.096
IV	−69.567	−9.043	−1.628	2.71	−0.234	−0.286	−0.282	−0.134	−0.164	−0.122
V	−95.166	−9.339	−1.546	1.11	−0.236	−0.263	−0.272	−0.099	−0.188	−0.085
VI	−111.183	−9.485	−1.684	3.31	−0.236	−0.285	−0.284	−0.100	−0.139	−0.156
VII	−27.345	−9.414	−1.422	0.94	−0.194	−0.277	−0.281	−0.100	−0.187	−0.084
VIII	−77.063	−9.311	−1.623	2.61	−0.234	−0.286	−0.277	−0.087	−0.165	−0.099
IX	−21.158	−9.484	−1.467	1.41	0.158	−0.274	−0.277	−0.104	−0.189	−0.082
X	−65.883	−9.490	−1.676	3.38	−0.237	−0.286	−0.269	−0.090	−0.152	−0.093
XI	−55.881	−9.209	−1.498	1.92	−0.205	−0.280	−0.280	−0.086	−0.176	−0.097
XII	−51.249	−9.681	−1.756	4.14	−0.229	−0.281	−0.266	−0.086	−0.139	−0.069
XII	−57.398	−8.987	−1.601	2.33	−0.232	−0.287	−0.272	−0.142	−0.058	−0.187
XIV	−56.401	−9.222	−1.634	2.55	−0.233	−0.286	−0.273	−0.138	−0.133	−0.141
XV	−49.201	−9.848	−1.639	2.69	−0.237	−0.286	−0.275	−0.091	−0.157	−0.218
XVI	−80.428	−8.975	−1.642	2.08	−0.238	−0.290	−0.277	−0.145	−0.118	−0.139
XVII	−178.362	−9.850	−1.697	3.23	−0.246	−0.284	−0.280	−0.148	−0.013	−0.055
XVIII	−24.841	−8.962	−1.330	1.51	−0.350	−0.285	−0.305	−0.084	−0.187	−0.086
XIX	−91.424	−9.786	−1.618	2.59	−0.235	−0.287	−0.275	−0.115	−0.149	−0.105
XX	−133.295	−9.825	−1.663	3.84	−0.237	−0.286	−0.277	−0.114	−0.179	−0.079
XXI	−142.237	−10.201	−1.845	5.93	−0.241	−0.283	−0.284	−0.108	−0.108	−0.169
XXII	−49.021	−9.848	−1.639	2.69	−0.238	−0.287	−0.279	−0.146	−0.116	−0.084
XXIII	−62.030	−8.975	−1.652	2.30	−0.241	−0.285	−0.268	−0.035	−0.156	−0.217
XXIV	−52.667	−8.956	−1.652	2.26	−0.235	−0.288	−0.278	−0.144	−0.120	−0.193
XXV	−95.570	−9.845	−1.673	3.87	−0.231	−0.283	−0.281	−0.115	−0.155	0.012
XXVI	−97.416	−9.263	−1.650	1.92	−0.235	−0.278	−0.269	−0.084	−0.177	−0.096

[a] See Fig. 2 and Table I.

amination of the LUMO training set coefficients of all samples showed that, except in the case of sample XXVI, there was hardly any variation in the values of Σc_i^2 of any given center shown in Fig. 2. Sample XXVI is the only instance in the training set where the aromatic ring is reduced. Since the LUMO can be expected to have a significant influence on quinone activity, where the reduction mechanism is predominant, the above observation indicates that the reduction of quinones may not be involved in the therapeutic action of the 1,4-naphthoquinone derivatives studied against W256.

On the other hand, an interesting relationship could be observed between the HOMO coefficients and the activity against W256. Table III lists the Σc_i^2 ($i = 2s$, $2p_x$, $2p_y$, and $2p_z$ of the atomic center in question) of HOMO coefficients for each compound included in Table I. It can be seen from Table III that there is a noticeable difference in the Σc_i^2 values for the set of atoms *b*, *c*, *m*, *n*, *o*, *p*, *q*, *s*, *t*, and *u* (see Fig. 2) between active and inactive compounds. Also, compounds for which Σc_i^2 of atoms *q* and *s both* exceed a value of 0.24 are all either significantly active or weakly active against W256. In all inactive compounds, with exception of XXVI, Σc_i^2 for *one* or *both* of these centers has a value < 0.24. The inactivity of compound XXVI can be attributed to the absence of the aromatic ring, which might be relevant for activity. Figure 3 shows the results of PCA carried out using only the Σc_i^2 for atoms *b*–*u* mentioned above as variables. It is gratifying to note that this set of ten variables is capable of separating the compounds in question into active and inactive classes. Note that HCA and PCA complement each other. Figure 3 contains biplots of scores and loadings for the first 2 PCs, which account for 81.44% of total cumulated variance of the data. The first principal component (PC_1) discriminates between active/inactive compounds. Looking at the scores and loadings of bivariate plots, it can be seen that except for descriptor *o*, all the others are heavily loaded on PC_1. Active samples have a high contribution of descriptors *p*–*u* (which include the side-chain double bond and terminal groups) and absence of *b*–*n*. On the other hand, inactive compounds have high

Table III. The HOMO Σc_i^2 Values for Some Relevant Atomic Centers for the Compounds in Table I[a]

Compound number	*b*	*c*	*m*	*n*	*o*	*p*	*q*	*s*	*t*	*u*
I	0.002	0.005	0.001	0.010	0.028	0.036	0.276	0.248	0.198	0.184
II	0.001	0.002	0.001	0.013	0.030	0.044	0.344	0.276	0.144	0.137
III	0.007	0.009	0.003	0.051	0.015	0.020	0.405	0.329	0.024	0.028
IV	0.000	0.001	0.001	0.002	0.038	0.048	0.365	0.341	0.023	0.020
V	0.002	0.005	0.000	0.003	0.018	0.026	0.417	0.329	0.027	0.028
VI	0.010	0.010	0.004	0.035	0.041	0.032	0.376	0.335	0.025	0.025
VII	0.002	0.008	0.004	0.003	0.018	0.026	0.409	0.319	0.024	0.028
VIII	0.001	0.001	0.002	0.018	0.026	0.040	0.399	0.355	0.017	0.021
IX	0.001	0.005	0.003	0.031	0.016	0.023	0.418	0.319	0.024	0.029
X	0.000	0.000	0.001	0.004	0.029	0.040	0.364	0.306	0.018	0.166
XI	0.012	0.013	0.004	0.045	0.016	0.020	0.397	0.327	0.022	0.026
XII	0.033	0.061	0.017	0.073	0.253	0.004	0.229	0.011	0.004	0.022
XIII	0.073	0.150	0.015	0.126	0.253	0.004	0.284	0.015	0.017	0.020
XIV	0.090	0.155	0.023	0.161	0.214	0.006	0.243	0.022	0.006	0.032
XV	0.185	0.241	0.045	0.373	0.026	0.023	0.033	0.006	0.006	0.002
XVI	0.090	0.194	0.015	0.152	0.190	0.000	0.257	0.016	0.026	0.004
XVII	0.084	0.212	0.036	0.311	0.045	0.008	0.067	0.029	0.019	0.003
XVIII	0.301	0.117	0.053	0.386	0.009	0.005	0.051	0.029	0.042	0.004
XIX	0.185	0.231	0.044	0.363	0.033	0.005	0.046	0.008	0.007	0.001
XX	0.176	0.231	0.040	0.349	0.032	0.005	0.041	0.012	0.012	0.001
XXI	0.168	0.221	0.039	0.363	0.025	0.032	0.033	0.011	0.005	0.001
XXII	0.091	0.189	0.017	0.153	0.197	0.001	0.261	0.014	0.001	0.019
XXIII	0.161	0.212	0.032	0.331	0.022	0.007	0.076	0.069	0.005	0.002
XXIV	0.096	0.187	0.018	0.152	0.203	0.002	0.251	0.017	0.023	0.020
XXV	0.185	0.231	0.044	0.374	0.029	0.004	0.043	0.008	0.006	0.001
XXVI	0.007	0.009	0.003	0.041	0.015	0.020	0.409	0.333	0.024	0.028

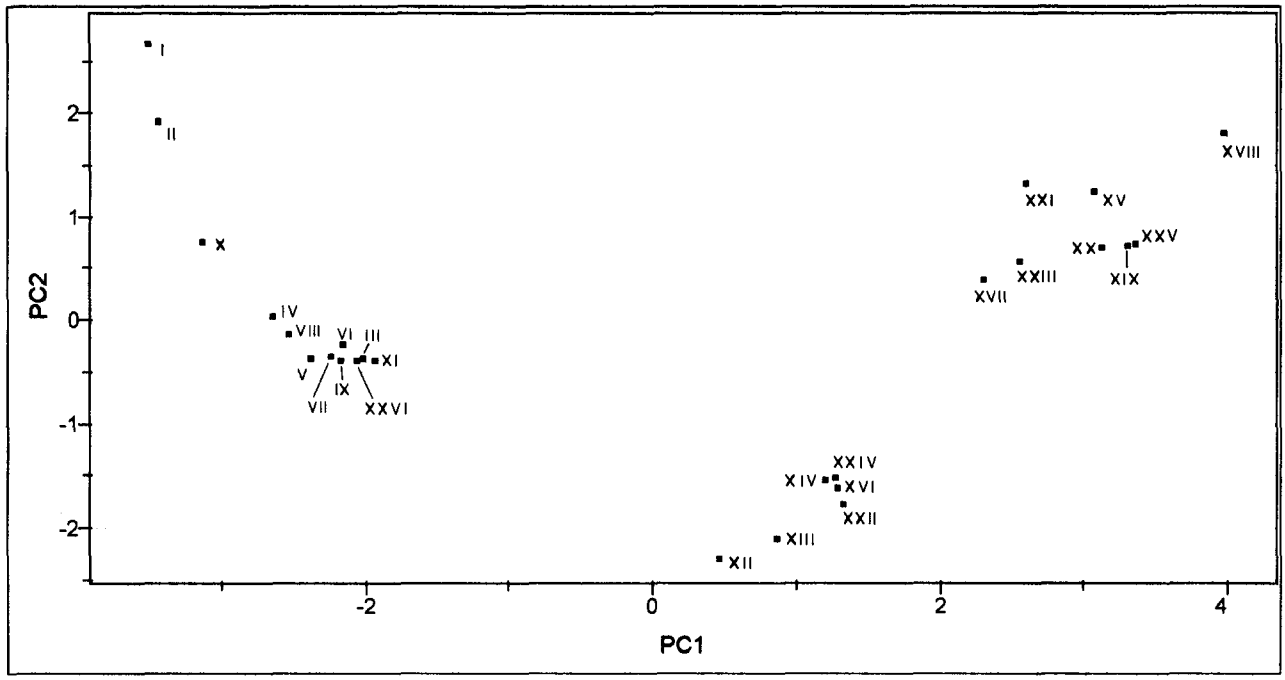
[a] See Fig. 2 and Table I.

contribution from *b–n* and no contribution from *p–q* descriptors on the first PC. The second principal component PC_2 discriminates a small set of inactive compounds which have high contribution of variables *o* and *q* (describing the double bond in the side chain). The only outlier is sample XXVI, experimentally known to be inactive. As mentioned before, its inactivity almost certainly stems from its nonaromatic character. Hierarchical cluster analysis results shown in Fig. 4 confirm the PC analysis presented above, providing a diagnostic of modeling strength. It is thus seen that the chosen set of ten variables is capable of separating the compounds in Table I into two well-defined classes.

LUMO and HOMO diagrams for some representative active and inactive derivatives are presented in Figs. 5 and 6. Comparison of Figs. 5a and 5b leads to the general result that LUMO has no influence on W256 activity since compounds III (active) and XXIII (inactive) both have practically the same LUMO pattern with no density on the side chain. On the other hand, the HOMO of the active compounds are significantly different from those of the inactive ones, an important side-chain density appearing in the active compounds. The participation of the side-chain atoms thus seems important for activity.

The above observation is consistent with some other work in relating to antitumor activity of quinones. According to Adams and Lewis [15], the product obtained through the peroxidation of lapachol, via an exposide intermediate involving the side-chain double bond, is active against KB (human epidermoid carcinoma of the nasopharynx cell culture) (see Fig. 7), Doroshow *et al.* [16], based on some recent *in vitro* experiments on the anticancer action of duxorubicin, have suggested that the quinone is capable of participating in redox reactions as a reducing agent instead of an oxidant, the π-electrons of the side-chain double bonds acting as an electron donor.

Once the data structure is known, classification models can be built using these samples as a training set, and the values of the HOMO coefficients can be used to predict activity against W256 of 1,4-naphtho-



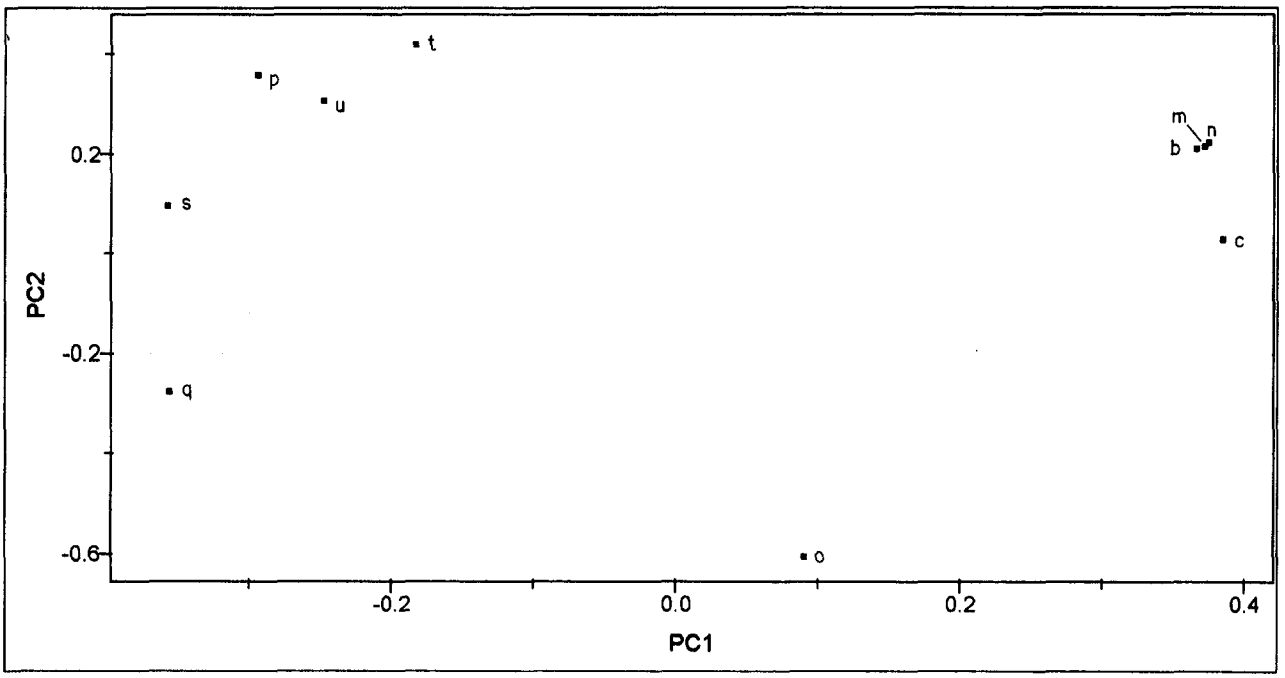


Fig. 3. (a) Plot of the first two PC score vectors (PC_1 and PC_2) for 26 1,4-naphthoquinone derivatives (training set). (b) Plot of the first two PC loading vectors (PC_1 and PC_2) for 26 1,4-naphthoquinone derivatives (training set).

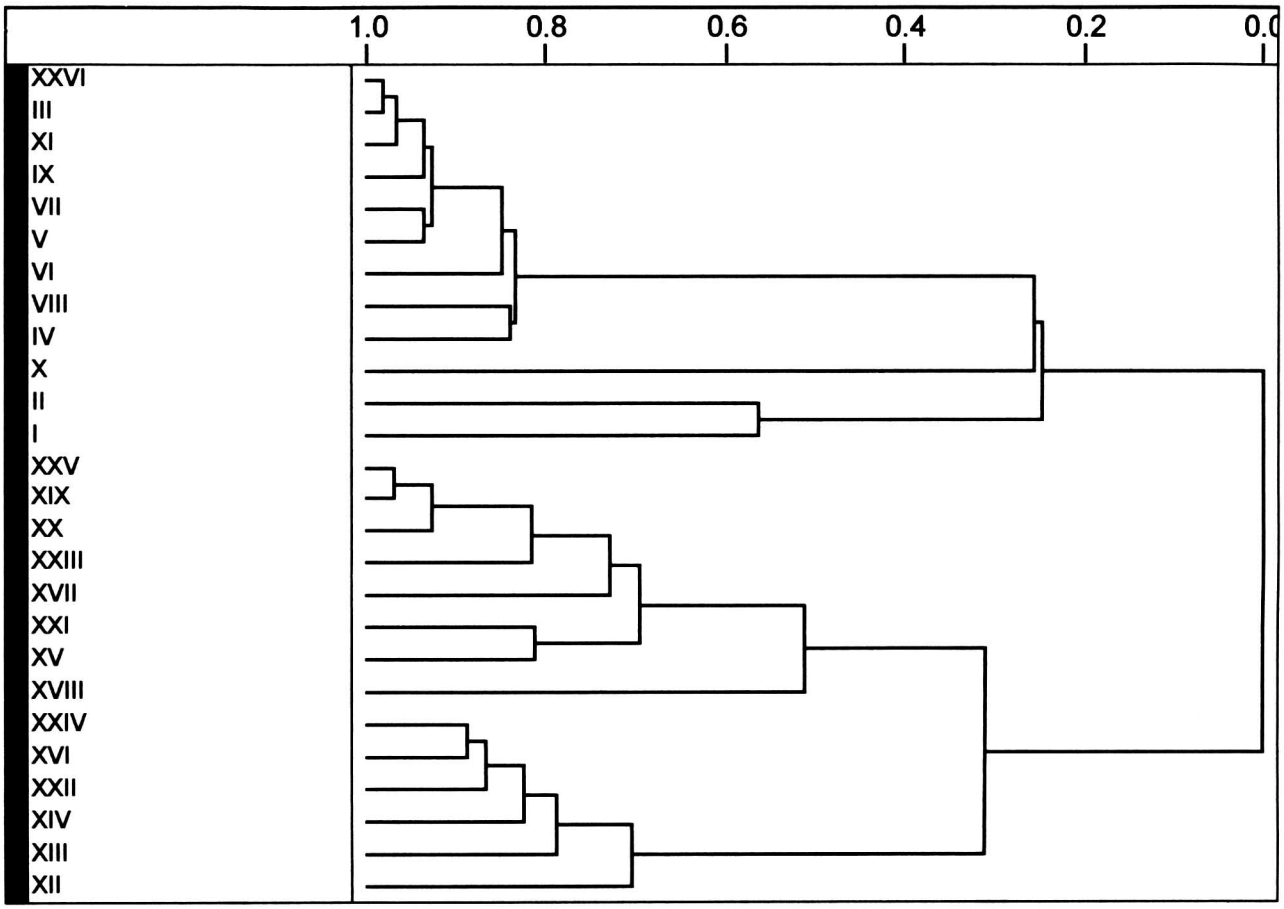


Fig. 4. Hierarchical cluster analysis plot for 26 1,4-naphthoquinone derivatives (training set).

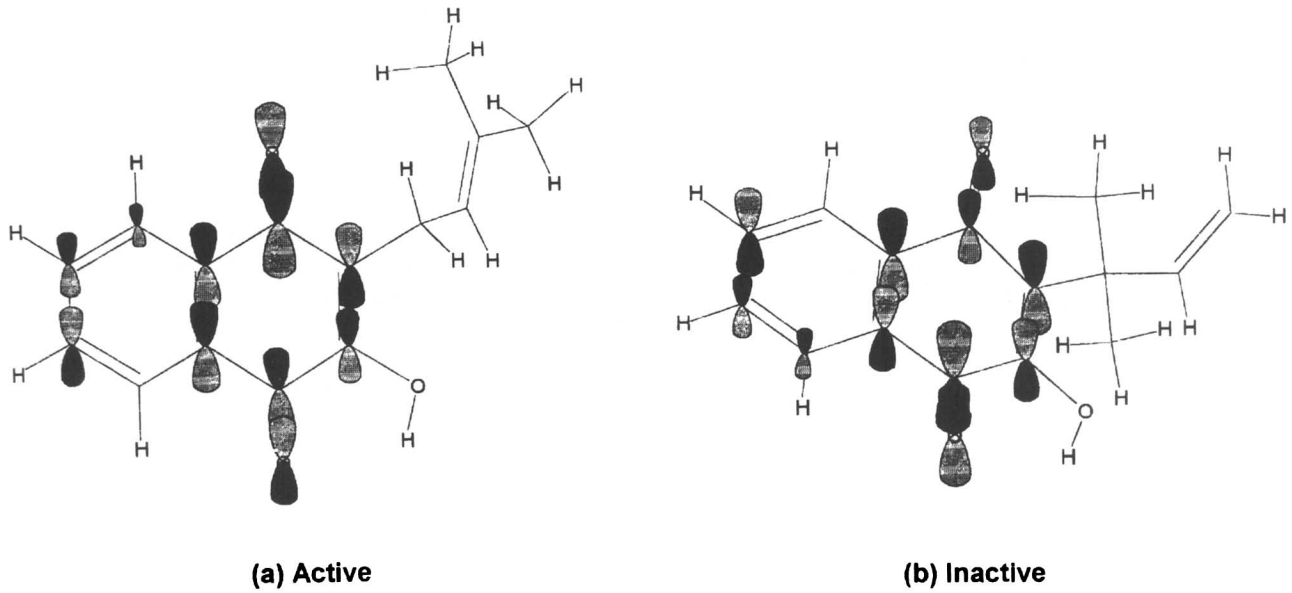


Fig. 5. (a) LUMO diagram of lapachol (III). (b) LUMO diagram of 1,1-dimethylallyl lawsone (XXIII).

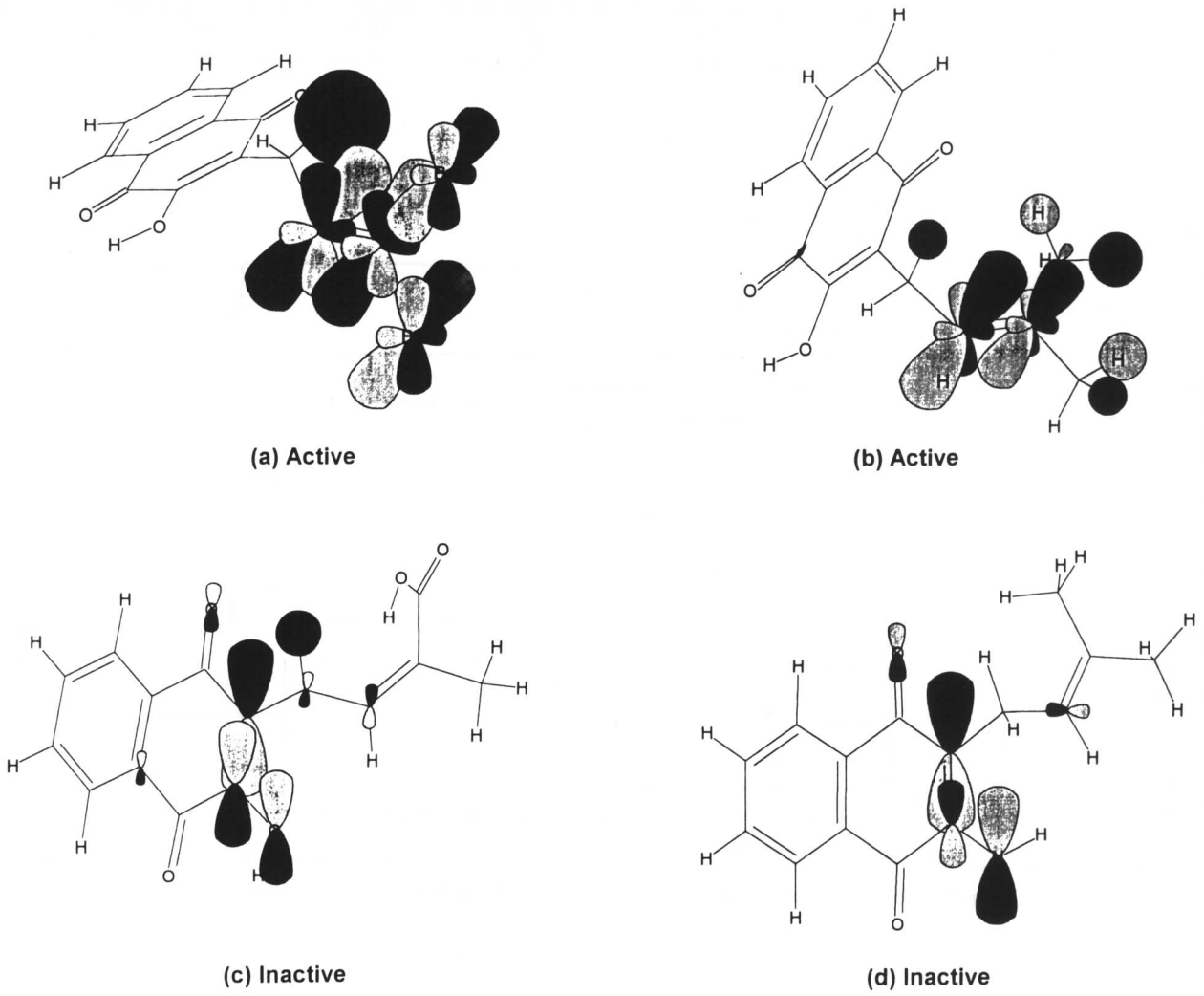


Fig. 6. HOMO diagrams of some representative 1,4-naphthoquinone derivatives active and inactive against W256. (a) Compound I, dibromoallyl-lawsome (active); (b) compound III, lapachol (active); (c) compound IV, allyl-lawsone (inactive); (d) compound XVIII, 2-amine-3-(3-methyl-2-butenyl)-1,4-naphthoquinone (inactive).

quinone derivatives. In this paper, two different techniques of classification have been used: KNN and SIMCA. Sample XXVI, being an outlier, has been excluded from the model.

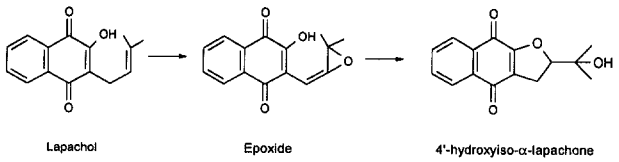


Fig. 7. Formation of 4′-hydroxyiso-α-lapachone from lapachol.

AM1 calculations were done for 16 selected 1,4-naphthoquinone derivatives (test set) listed in Table IV for which experimental data of W256 activity are not available, and Table V shows standard MO parameters for this set. Table VI includes the HOMO Σc_i^2 values for the relevant atomic centers for this test set. The latter are used for the statistical analyses.

Results with KNN Method

The model was built using three nearest neighbors in the training set, with zero misses and the active/inactive predicted classes for the test sample listed in Ta-

Table IV. Compounds Selected for Theoretical Prediction of Activity Against W256 (Test Set)[a]

Compound number	R	R_1
XXVII	$OCH_2CH=CH_2$	$CH_2CH=C(CH_3)_2$
XXVIII	OH	$CH_2C(CH_3)=CH_2$
XXIX	OH	$CH_2C(CH_3)=C(CH_3)_2$
XXX	OH	$CH_2C(CH_3)=CH(CH_3)$
XXXI	NH_2	$CH_2CH=CBr_2$
XXXII	NH_2	$CH_2CH=CCl_2$
XXXIII	NH_2	$CH_2CH=C(CH_3)CH_2CH_2CH=C(CH_3)_2$
XXXIV	SH	$CH_2CH=C(CH_3)_2$
XXXV	SH	$CH_2CH=C(Br)_2$
XXXVI	SH	$CH_2CH=C(Cl)_2$
XXXVII	SH	$CH_2CH=C(CH_3)CH_2CH_2CH=C(CH_3)_2$
XXXVIII	Cl	$CH_2CH=C(CH_3)_2$
XXXIX	Cl	$CH_2CH=C(Br)_2$
XL	Cl	$CH_2CH=C(Cl)_2$
XLI	Cl	$CH_2CH=C(CH_3)CH_2CH_2CH=C(CH_3)_2$
XLII[b]	OH	$CH_2CH=C(CH_3)_2$

[a] See Fig. 1 for structures.
[b] $C_c=C_n$ reduced to C_c-C_n (see Fig. 2).

ble VII. Only sample XXVIII, which is expected to be inactive (see below), has not been predicted as such.

Results with SIMCA Method

The model used two PCs for each class and the predicted classes are listed in Table VII. Figure 8 is a visual classification representation in a biplot of class distances. The *F*-statistic threshold lines (defined in terms of a probability and chosen here as 95%) divide the plot into four quadrants. Samples in the NW quadrant belong to CS1 (active class) and those in quadrant SE to CS2 (inactive class). Samples located on the SW quadrant can be a member of either class and samples

Table V. Values of Certain Parameters Obtained Through Theoretical Calculations Using the AM1 Method for the Compounds in Table IV

Compound number	Heat of formation (kcal/mol	E_{HOMO} (eV)	E_{LUMO} (eV)	Dipole moment (Debye)	Atomic charge[a] On *b*	On *e*	On *m*	On *o*	On *q*	On *s*
XXVII	−36.746	−9.289	−1.494	1.56	−0.212	−0.275	−0.277	−0.091	−0.179	−0.094
XXVIII	−56.053	−9.600	−1.626	2.66	−0.239	−0.287	−0.275	−0.085	−0.106	−0.220
XXIX	−71.068	−8.946	−1.593	2.63	−0.240	−0.287	−0.278	−0.079	−0.122	−0.097
XXX	−64.693	−9.227	−1.611	2.65	−0.238	−0.286	−0.279	−0.085	−0.116	−0.154
XXXI	−2.960	−9.243	−1.511	3.10	−0.354	−0.280	−0.302	−0.102	−0.077	−0.333
XXXII	−24.923	−9.225	−1.465	2.73	−0.347	−0.280	−0.299	−0.089	−0.142	−0.119
XXXIII	−29.309	−8.950	−1.362	1.39	−0.341	−0.285	−0.306	−0.123	−0.169	−0.115
XXXIV	−16.560	−8.900	−1.598	1.88	0.094	−0.291	−0.279	−0.104	−0.184	−0.089
XXXV	−6.801	−9.215	−1.818	3.86	0.212	−0.285	−0.272	−0.122	−0.068	−0.319
XXXVI	−15.714	−9.155	−1.771	3.45	0.209	−0.286	−0.272	−0.109	−0.131	−0.108
XXXVII	−20.291	−8.909	−1.613	2.02	0.195	−0.290	−0.281	−0.133	−0.166	−0.119
XXXVIII	−24.210	−9.478	−1.637	2.30	0.028	−0.257	−0.272	−0.105	−0.187	−0.084
XXXIX	−0.476	−9.995	−1.831	3.30	0.038	−0.243	−0.262	−0.123	−0.078	−0.314
XL	−22.460	−9.936	−1.777	2.73	0.034	−0.245	−0.261	−0.110	−0.144	−0.099
XLI	−28.089	−9.150	−1.633	2.34	0.031	−0.252	−0.273	−0.133	−0.167	−0.118
XLII	−86.618	−9.311	−1.057	2.44	−0.296	−0.259	−0.277	−0.114	−0.198	−0.087

[a] See Fig. 2 and Table IV.

Table VI. The HOMO Σc_i^2 Values for Some Relevant Atomic Centers for the Compounds in Table IV[a]

Compound number	*b*	*c*	*m*	*n*	*o*	*p*	*q*	*s*	*t*	*u*
XXVII	0.005	0.001	0.001	0.032	0.018	0.029	0.412	0.329	0.026	0.027
XXVIII	0.033	0.022	0.009	0.078	0.014	0.001	0.307	0.399	0.001	0.002
XXIX	0.002	0.003	0.001	0.023	0.011	0.017	0.368	0.350	0.019	0.028
XXX	0.005	0.011	0.001	0.027	0.010	0.014	0.362	0.371	0.000	0.037
XXXI	0.308	0.103	0.060	0.429	0.010	0.023	0.021	0.007	0.005	0.005
XXXII	0.304	0.111	0.059	0.417	0.011	0.026	0.024	0.009	0.005	0.003
XXXIII	0.292	0.109	0.053	0.397	0.008	0.026	0.012	0.006	0.004	0.003
XXXIV	0.594	0.063	0.027	0.227	0.005	0.001	0.036	0.021	0.003	0.002
XXXV	0.609	0.059	0.027	0.252	0.003	0.004	0.013	0.006	0.004	0.003
XXXVI	0.624	0.058	0.028	0.252	0.003	0.003	0.015	0.006	0.003	0.002
XXXVII	0.593	0.063	0.027	0.240	0.001	0.000	0.006	0.004	0.000	0.000
XXXVIII	0.004	0.006	0.002	0.033	0.005	0.023	0.405	0.325	0.023	0.029
XXXIX	0.000	0.001	0.000	0.001	0.026	0.029	0.282	0.240	0.212	0.184
XL	0.001	0.000	0.000	0.006	0.029	0.040	0.352	0.263	0.151	0.140
XLI	0.000	0.000	0.001	0.001	0.034	0.048	0.365	0.342	0.023	0.023
XLII[b]	0.012	0.010	0.002	0.033	0.018	0.023	0.418	0.324	0.025	0.026

[a] See Fig. 2 and Table IV.
[b] $C_c{=}C_n$ reduced to $C_c{-}C_n$ (see Fig. 2).

in the NE quadrant are members of neither class. It can be seen that many samples are very close to the threshold and might be considered borderline. The derivative XXVIII is *close* to the active class, but not *inside* it (NE quadrant). In all probability, it should prove inactive since the side-chain C=C carries only one methyl substituent. It is interesting to note that both in the training and the test sets all nonhalogenated derivatives that belong to class 1 have *at least two* methyl or substituted methyl groups attached to the C=C in the side chain. This appears to be one of the essential conditions for W256 activity. Compounds XXXIV–XXXVII are located in the NE quadrant, which means that they are neither sharply included in active or inactive classes, although very close to the latter. All these derivatives have a common structural feature—a thiol group in place of a hydroxyl group at C-2. The presence of the C-2 hydroxyl group is believed to be another prerequisite for W256 activity. So the results obtained are reasonable since in the modeling there is a lack of representation of the thiol derivatives and, more likely, these four compounds belong to an inactive class subset.

Table VII. Statistical Data for the Compounds in Table IV (Test Set)

Compound number	Predicted activity against W256	
	KNN	SIMCA[a]
XXVII	Yes	Yes
XXVIII	Yes	0
XXIX	Yes	Yes
XXX	Yes	Yes
XXXI	No	0
XXXII	No	No
XXXIII	No	No
XXXIV	No	0
XXXV	No	0
XXXVI	No	0
XXXVII	No	0
XXXVIII	Yes	Yes
XXXIX	Yes	Yes
XL	Yes	Yes
XLI	Yes	Yes
XLII	Yes	Yes

[a] Samples which did not fit in either of the two classes are assigned to a class 0.

Finally, all the compounds belonging to class 1, both in the training and the test sets, have HOMO Σc_i^2 values above a threshold value of 0.24 for both the side-chain carbon centers *q* and *s*. This appears to be an essential criterion for activity against W256. Future experiments with the compounds of the test set should confirm or disprove this prediction.

CONCLUSION

In this study of the structure–activity relationship for W256, the 1,4-naphthoquinone derivatives appear to participate as reducing agents involving the alkyl side-

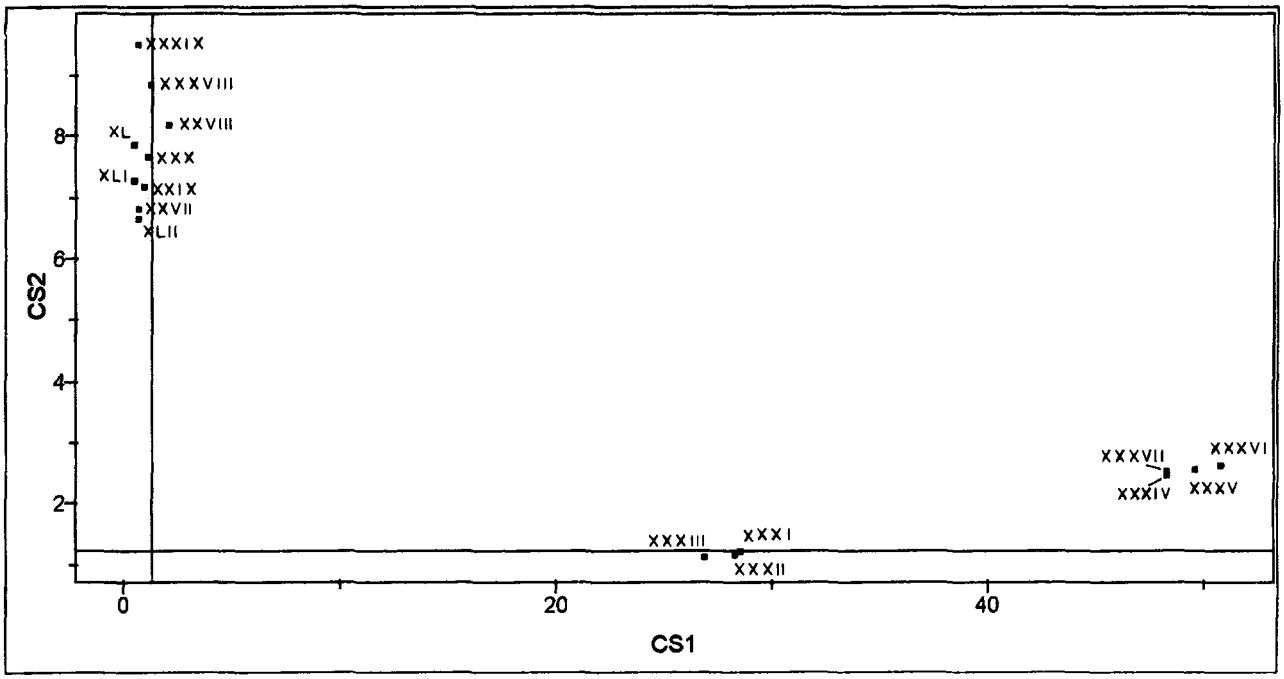


Fig. 8. Class distances plot for active (CS1) and inactive (CS2) classes for test samples.

chain double bond. The results obtained were used to predict the activity of new derivatives against W256.

ACKNOWLEDGMENTS

The authors thank CAPES, CNPq, FAPESP, and FINEP (Brazilian agencies) for financial support for this work. They are also grateful to Dr. Benício B. Neto of the Federal University of Pernambuco, Recife, for helpful discussions.

REFERENCES

1. Morrison, R. K.; Brown, D. E.; Oleson, J. J.; Cooney, D. A. *Toxicol. Appl. Pharmacol.* **1970,** *17*, 1-11.
2. De Almeida, E. R.; Da Silva Filho, A. A.; Dos Santos, E. R.; Lopes, C. A. C. *J. Ethnopharmacol.* **1990,** *29*, 239-241.
3. Goel, R. K., Pathak, N. K.; Biswas, M.; Pandey, V. B.; Sanyal, A. K. *J. Pharm. Pharmacol.* **1987,** *39*, 138-142.
4. Santana, C. F.; Lins, L. J. P.; Asfora, J. J.; Melo, A. M.; Gonçalves de Lima, O.; D'Albuquerque, I. L. *Rev. Inst. Antibiot.* **1980/1981,** *20*, 61-66.
5. Almeida, E. R.; Mello, A. C.; Santana, C. F.; Silva Filho, A. A.; Santos, E. R.; *Rev. Port. Farm.* **1988,** *38*(3), 21-23.
6. Hartwell, J. L.; Abbot, B. J. *Adv. Pharmacol. Chemother.* **1969,** *7*, 117-209.
7. Block, J. B.; Serpick, A. A.; Miller, W.; Wiernik, P. H. *Cancer Chemother. Rep. Part 2* **1974,** *4*(4), 27-28.
8. Rao, K. V. *Cancer Chemother. Rep. Part 2* **1974,** *4*(4), 11-17.
9. Driscoll, J. D.; Hazard, G. F.; Wood, Jr., H. B.; Goldin, A. *Cancer Chemother. Rep. Part 2* **1974,** *4*(2), 1-27.
10. Pereira, M. A. Ph.D. Thesis. Instituto de Física e Química de Sãon Carlos, Universidade de São Paulo, São Carlos, São Paulo, Brazil, 1989.
11. Malinowski, E. R. *Factor Analysis in Chemistry*. Wiley: New York, 1991.
12. Pirouette Multivariate Data Analysis for IBM PC Systems, Version 2.0. Infometrix: Seattle, WA, 1996.
13. Kowalski, B. R.; Bender, C. F. *J. Am. Chem. Soc.* **1972,** *94*, 5632.
14. Wold, S. *Pattern Recognition* **1976,** *8*, 127-139.
15. Adams, J. H.; Lewis, J. R. *J. Chem. Res. (M)* **1978,** 0186-0192.
16. Doroshow, J. H.; Arkman, S.; Chu, F.-F.; Esworthy, S. *Pharmac. Ther.* **1990,** *47*, 359-370.