3D QSAR/CoMFA and CoMSIA Studies on Antileukemic Steroidal Esters Coupled with Conformationally Flexible Nitrogen Mustards

Agnes Kapou,[†,‡] Nikolas-P. Benetis,[‡] Serdar Durdagi,[‡,§] Sotiris Nikolaropoulos,*[,†] and Thomas Mavromoustakos*[,‡,ǁ]

Department of Pharmacy, University of Patras, 26500 Rio, Patras, Greece, Institute of Organic and Pharmaceutical Chemistry, The National Hellenic Research Foundation, 48 Vas. Constantinou Avenue, 11635 Athens, Greece, Department of Biology Chemistry and Pharmacy, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany, and Department of Chemistry, University of Athens, Zographou, 15784 Athens, Greece

Thirty-eight antileukemic steroidal esters possessing conformationally flexible nitrogen mustards were studied, and the 3D QSAR/CoMFA and CoMSIA methodologies were applied in order to derive the correlation between their structure and the *in vivo* antileukemic activity. These compounds show significantly reduced toxicity and possibly increased bioavailability compared to free nitrogen mustards and therefore constitute potent antileukemic drugs. Both the CoMFA and CoMSIA studies gave similar results indicating that the steric effect and the hydrophobic/hydrophilic balance especially in the steroidal part of the molecules probably determined their bioactivity. Of paramount interest is the observation that the orientation of the alkylating part of the SMEs toward the surface of ring B of the steroidal skeleton was related with increased activity. Concerning the steroidal part, the presence of hydrophobic groups in rings B and D was found to be important for enhanced activity. Enhancement of antileukemic potency is further observed if hydrophilic/H-bond acceptor groups are present at the positions 7 and 17 of the steroidal skeleton. Leapfrog simulations provided novel compounds which lead our future synthetic endeavor for obtaining SMEs with optimum bioactivity.

1. INTRODUCTION

Among the drugs that are used today in chemotherapy against cancer there is a category of compounds that have in common the potency to alkylate biological macromolecules involved in the human biology e.g. DNA, RNA, and several proteins. In particular, their property to alkylate DNA, a process that leads to formation of covalently modified DNA bases,[1] is considered as the main pathway they achieve their antileukemic activity. Nitrogen mustards belong to this category of drugs being therefore very potent antineoplasmatic agents.

Their chemical structure is based on the gas mustard bis(2-chloroethyl) sulfide seen in Figure 1 which is known to damage the bone marrow and the lymph tissue. However, due to their high toxicity, attributed to their low selectivity and the development of resistance by the cancer cells, their use and applications are limited. In order to decrease their toxicity and improve their effectiveness, they have been used in combination, either in the form of a mixture or by coupling with several other classes of organic molecules.

Among these combined molecules, the steroidal esters of nitrogen mustards had displayed a very interesting pharmacological profile. These compounds had been subjected to
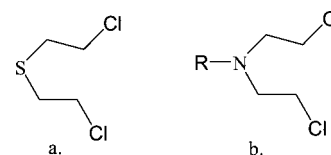

Figure 1. (a) The mustard bis(2-chloroethyl) sulfide. (b) The general structure of nitrogen mustards. The variable group R can be aromatic or nonaromatic.

antileukemic tests *in vitro* as well as against experimental forms of leukemia (P388 and L1210). These tests displayed a great experimental variance of antileukemic activity on BDF1 mice. All the biological data for these complex compounds are shown in Table 1, where also references about their origin are included.[2−18] Hitherto, the mechanism of action of these compounds has not yet been understood. In particular it is not well-known if the role of the steroidal part is simply limited to the transfer of the alkylating part on the site of action, hypothesis which is increasingly shrinking. For this reason, in the past, we have used Differential Scanning Calorimetry (DSC) for evaluating the thermal effects of one of the most active esters inserted in a biological membrane model. The results showed that the steroidal segment ameliorated the thermal effects of the mustard.[19] This was an important finding since the presence of the steroid may be related to the reduced destructive interaction of the alkylating agent with the membrane bilayers and consequently with the reduction of its toxicity. The results depict also the important role of the cell membranes in the mechanism of action of the antileukemic activity of steroidal mustard esters (SMEs).

* Corresponding author phone: +302107274293; fax: + 302107274261; e-mail: tmavrom@chem.uoa.gr (T.M.) and phone: +302610969326; fax: +302610310553; e-mail: snikolar@upatras.gr (S.N.).
† University of Patras.
‡ The National Hellenic Research Foundation.
§ Freie Universität Berlin.
ǁ University of Athens.

ANTILEUKEMIC STEROIDAL ESTERS

J. Chem. Inf. Model., Vol. 48, No. 11, 2008 **2255**

Table 1. Chemical Structures, in Vivo Biological Activity Values (T/C % and T′/C %) and Toxicity (LD$_{10}$) of the Esters of Three Types of Nitrogen Mustards Bound with Simple and Modified Steroids

Comp. No.	Chemical structure	LD$_{10}$	T/C %[a]	T′/C %[b]	Comp. No.	Chemical structure	LD$_{10}$	T/C %[a]	T′/C %[b]
1[2,3]		65	242	242	20		450	126	126
2[4,5]		250	199	199	21		210	139	139
3[6,7]		80	204	204	22[4]		500	201 (1)	222
4[8]		400	141	141	23[4]		550	139	139
5		34	151	151	24[3]		155	140	140
6		240	300 (4)	311	25[4]		540	121	121
7		60	221 (1)	237	26		42	268	268
8		140	231	231	27		14	178	178
9		58	153	153	28		10	133	133
10		220	295 (1)	296	29[3,10]		50	174	174
11		40	199	199	30[11,12]		-	152	152
12[9]		56	333 (3)	343	31[13-16]		36	199	199
13[10]		170	219	219	32[4,5,13,14]		140	226	226
14[11]		34	257	257	33[11,12]		-	154	154
15[12]		400	153	153	34[3]		400	185	185
16[3]		25	314	314	35[3]		610	144	144
17		44	198	198	36[3]		580	130	130
18[3,9]		40	321	321	37[17]		80	132	132
19[9]		70	198	198	38[18]		18	119	119

[a] The survivals are denoted in parentheses. As survivals were considered the animals that survived more than thirty days after the initiation of the treatment. [b] The index T′/C % is described in the text. It comprises the modification of T/C % so that survivals are accounted for.
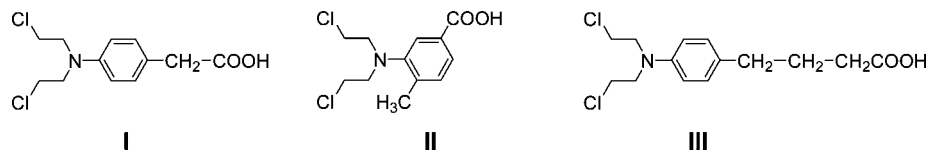
Figure 2. The three different types of alkylating nitrogen mustards groups used as esters with simple or modified steroids.

The present 3D-QSAR (Quantitative Structure Activity Relationships) study aims to reveal the stereoelectronic requirements for antileukemic activity of the synthetic SMEs which ultimately will lead to new potent analogs. In particular, three types of alkylating nitrogen mustards are used esterified with simple and modified steroids.

The two developed 3D-QSAR methodologies, CoMFA[20] (Comparative Molecular Field Analysis) and CoMSIA[21,22] (Comparative Molecular Similarity Indices Analysis), were applied. CoMSIA and its precursor CoMFA are leading to high quality graphical maps which describe mainly the steric and the electrostatic requirements for a biologically active molecule in the 3D Cartesian space. In addition, CoMSIA allows the determination of the 3D distribution of the hydrophobicity of the surface of the active molecules and their capability to contribute to hydrogen bonds with the macromolecular receptor.

The selection of the lowest energy conformation of the bioactive conformation of the template molecule and the superimposition of the rest of the molecules on the template compound are the two most critical steps in the 3D QSAR/ CoMFA and CoMSIA studies. Conformational analysis of low energy conformers in combination with distance geometry results using 2D NOESY NMR was applied for the template molecule.[23-27]

The aim of this study is to reveal the stereoelectronic parameters that govern the antileukemic activity of SMEs and use them to design and synthesize novel molecules. The *de novo* design Leapfrog routine of SYBYL[28] was used as an aid for the discovery of new molecules based on 3D QSAR results.

2. RESULTS AND DISCUSSION

2.1. Structure and Biological Activity of SMEs. In the present work, thirty-eight esters of the three nitrogen mustards shown in Figure 2, bound to several modified steroids as seen in Table 1, were studied in order to reveal the relationship between their 3D-structure and their anti-leukemic activity. The three different types of the alkylating mustard groups used in our study are as follows: (**I**) p-N,N-bis(2-chloroethyl)aminophenylacetic acid (PHE), (**II**) p-methyl-m-N,N-bis(2-chloroethyl)aminobenzoic acid (4-Me-CABA), and (**III**) p-N,N-bis(2-chloroethyl)aminophenylbutyric acid (CHL).

2.2. Molecular Modeling and Conformational Analysis. The structures of the studied molecules were subjected to full geometry optimization using a combination of the standard Tripos molecular mechanics force field of the SYBYL molecular modeling package[28] with Powell energy minimization algorithm, Gasteiger-Huckel charges, and 0.001 kcal/(mol.Å) energy gradient convergence criterion.

Since the alkylating agent includes a large number of rotatable bonds, template compound **12** has a very flexible character. Thus, its conformational analysis must be carried
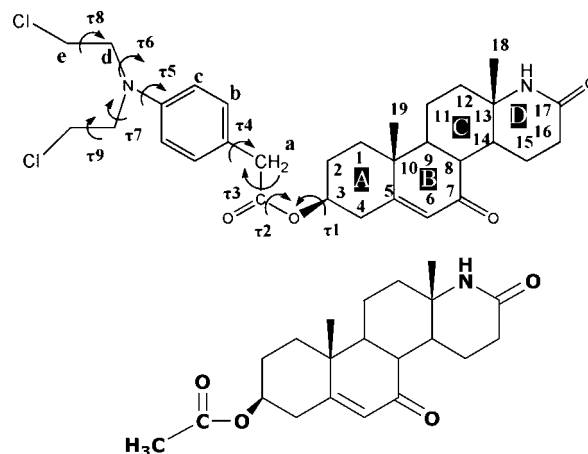

Figure 3. Top: The dihedral angles of the nonsteroidal part of the 3β-hydroxy-17α-aza-D-homo-7,17-dione-5-androsten-p-N,N-bis(2-chloroethyl)aminophenylacetic ester (**12**) which were used as variables for the conformational search. Bottom: The precursor steroidal molecule 3β-acetoxy-17α-aza-D-homo-5-androsten-7,17-dione studied using a combination of NMR spectroscopy and molecular modeling.
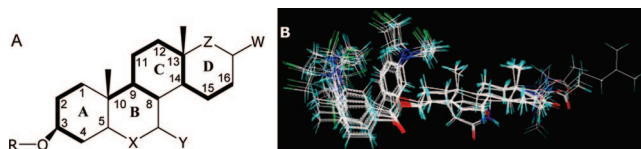

Figure 4. A. The superimposition of all the compounds was performed on the basis of the atoms 1−4 and 8−16 of the template-molecule, compound **12**, which are common in all the steroidal skeletons. **B.** Compounds of the training group are shown after the alignment. R=PHE, 4-Me-CABA, or CHL. X, Y, Z, and W are either polar or hydrophobic substituents. B ring due to its substitution can be a six or seven member ring, and D ring due to its substitution can be five or six member ring.

out very carefully. In our previous studies, the conformational analysis of the template molecule **12** has been reported, using a combination of molecular modeling and experimental studies[30] (Figure 3). This conformation is applied for all figure comparisons.

The conformers of the steroidal esters of the mustards seen in Table 1 were derived from the conformation of the template as it had been explained earlier.[30,31] The Field Fit method on SYBYL was applied associated with energy minimization to match the flexible segments of the synthetic analogs with the template. The procedure was followed by geometry optimization using Powell and Steepest Descent algorithms with tolerance 0.001 kcal·mol^{-1}·Å^{-1} criterion.

The steroidal skeleton that comprises the rigid part of the SMEs is used as the common segment for the alignment of all the other compounds on the lowest energy conformation of **12** seen in Figure 4. For this segment the numbering of the carbon atoms of the common section of all the compounds of the steroid skeleton which were used for the alignment is highlighted in the same figure.

Substituents as it is shown in Figure 4 are focused on rings B and D and constitute hydrophilic and lipophilic segments. These substituents contribute to the expansion of these rings. Thus, ring B can be a six or a seven member ring and ring D can be a five or a six member ring. As it was described earlier R is one of the three defined mustards PHE, 4-Me-CABA, or CHL.

Both CoMFA and CoMSIA exploit the recent developments of mathematical-statistical tools such as PLS (Partial Least Squares) which are able to treat systems having greater number of variables/adjustable parameters and much fewer experimental results describing the activities of the molecules.[32]

After superimposition, the conformations and the experimental bioactivity data for thirty-eight antileukemic steroidal mustard esters (SMEs) were introduced as input in the CoMFA and CoMSIA methods of SYBYL molecular modeling package. In CoMFA, carbon atoms of sp^3 hybridization were used as probes, bearing a $+1$ charge and were put in the intersections of a 3D grid (2 Å). The calculation of fields followed, and 30 kcal/mol cutoff values were used for the computation of the steric and the electrostatic CoMFA fields. CoMSIA was performed using the QSAR option of SYBYL. For the steric and electrostatic fields and hydrophobicity the same probe atoms, grid density, and parameters were used as in the CoMFA model. However, for the donor−acceptor hydrogen-bond fields, the probe atom on the grid was changed to sp^3 hybridized oxygen with charge -1, and the attenuation factor was set to 0.3 for the Gaussian functions.

The CoMFA and CoMSIA models were built using a training set of twenty-six compounds and tested using a test set of twelve compounds. The actual and the predicted by the models values of activity for the twenty-six molecules of the training set are shown in Table 2. The training and test set molecules were chosen among the entire set of molecules to represent all chemical variations in their structures and also in such a way that both groups show variety in activity. Thus, structures have variations in the alkylating groups and in the B and D rings. The size of the test set is restricted by the size of the experimental data available.

In CoMFA, only one model including the two available fields, the steric and the electrostatic fields was possible. In CoMSIA, on the other hand, eight models were constructed with different combinations of the available fields. The relevant data are included in Table 3.

In the CoMFA model the contribution of the two fields, the steric and the electrostatic was comparable, with the steric contribution being slightly greater. In the corresponding CoMSIA (steric-electrostatic) model, the contribution of the steric field was clearly dominating. This inconsistency is due to the different types of functions used for the calculation of the fields in the standard software of the two methods.

The CoMFA model that takes into account steric and electrostatic interactions gave an acceptable r_{cv}^2 value of 0.559. If identical interactions were assumed for CoMSIA, then a smaller value of (0.479) was observed. Although relatively small, the contribution of the H-bond acceptor field could not be ignored due to the N and O atoms of the steroidal part. In that case one could include three fields, the steric, the hydrophobic, and the H-bond acceptor, and

Table 2. Experimental Values of the Biological Activity and the Theoretical Values of the Activity Which Were Computed on the Basis of the CoMFA and CoMSIA Models Obtained from the 26 Compounds of the Training Set

training set compd no.	T′/C*		
	exp values	CoMFA predictions	CoMSIA predictions
1	242	264	257
2	199	182	170
3	204	199	192
5	151	127	135
7	237	250	255
8	231	223	216
11	199	191	214
12	343	346	335
13	219	226	228
15	153	159	194
16	314	313	325
18	321	307	286
19	198	186	183
20	126	143	120
22	222	199	181
23	139	140	138
24	140	140	156
25	121	132	125
29	174	190	193
30	152	155	165
31	199	193	174
33	154	129	139
34	185	173	193
36	130	144	150
37	132	150	139
38	119	142	142

exclude the electrostatic field and the H-bond donor field. Finally, since the consideration of the electrostatic field had no negative consequences on r_{cv}^2 while it resulted in a slight increase of r_{pred}^2 it was also included in the CoMSIA model. Therefore, a model (CoMSIA2) with four fields, the steric, the hydrophobic, the electrostatic, and the hydrogen-bond (H-bond) acceptor fields, was used. A short justification for this choice will be given further, pointing out first that the correlation was statistically significant for both CoMFA and for CoMSIA models with the above selected fields as it is indicated by the high r_{conv}^2 values. In addition, the statistically significant r_{cv}^2 and r_{bs}^2 emphasize the significant reliability of both models.

Preliminary results of CoMSIA models indicated that the hydrophobic field dominated, then the steric field, and the H-bond acceptor field. From the results of the CoMFA and CoMSIA in Table 3, it is shown that hydrophobic interactions are the dominant. However, it was observed that the ability of the prediction of the potency of the compounds of the test group was improved by the additional consideration of the steric, electrostatic, and the hydrogen-bond acceptor fields, since their incorporation gave a slightly increased value of the r_{pred}^2.

As seen in Figure 5 the experimental *in vivo* activities of the diagram are accurately reproduced by the predictions.

2.3. CoMFA Fields. For the graphical presentation of the QSAR results according to the CoMFA model the usual equipotential 3D surfaces of the products STDEV*COEFF were used with standard coloring.

Steric field. Green contour plots indicate regions where bulky groups are related with increased bioactivity, while yellow contour plots indicate regions where such bulky groups are not favoring bioactivity.

Table 3. Summary of the Results of the CoMFA and CoMSIA Analyses[a]

	CoMFA	CoMSIA1 (st,el,hy, acc,don)	CoMSIA2 (st,el,hy, acc)	CoMSIA3 (st,el,hy)	CoMSIA4 (st,hy, acc)	CoMSIA5 (st,hy)	CoMSIA6 (st,el)	CoMSIA7 (st)	CoMSIA8 (hy)
r_{cv}^2	0.559	0.541	0.547	0.580	0.549	0.583	0.479	0.421	0.586
s_{press}	47.368	47.096	46.806	45.019	46.701	44.870	50.171	54.250	44.732
components	6	5	5	5	5	5	5	6	5
r_{conv}^2	0.946	0.897	0.895	0.899	0.893	0.897	0.897	0.914	0.893
SE	16.521	22.304	22.480	22.102	22.768	22.275	22.312	20.909	22.769
r_{pred}^2	0.674	–	0.731	0.720	–	0.714	–	–	0.701
r_{bs}^2	0.968	–	0.928	–	–	–	–	–	–
SD_{bs}	11.311	–	17.202	–	–	–	–	–	–
Field	Fraction								
steric	0.574	0.291	0.295	0.308	0.303	0.321	0.836	1.000	–
electrostatic	0.426	0.039	0.041	0.055	–	–	0.164	–	–
hydrophobic	–	0.534	0.551	0.637	0.573	0.679	–	–	1.000
H-bond donor		0.028	–	–	–	–	–	–	–
H-bond acceptor	–	0.109	0.113	–	0.124	–	–	–	–

[a] In the parentheses of the column headings the abbreviations of the used fields are indicated. st = steric; el = electrostatic; hy = hydrophobic; acc = H-bond acceptor; don = H-bond donor; r_{cv}^2 = cross-validated r^2; r_{conv}^2 = conventional r^2; r_{bs}^2 = bootstrapping r^2.
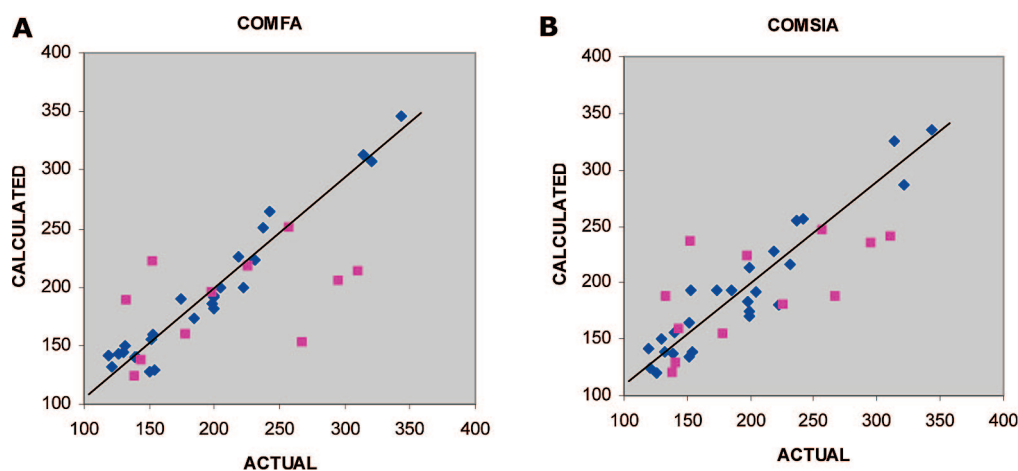

Figure 5. Experimental biological activity values plotted versus the predicted values of bioactivity which they were obtained from the selected models: **A.** CoMFA and **B.** CoMSIA. The blue rhombs and the magenta squares refer to the training set and the test sets, respectively.
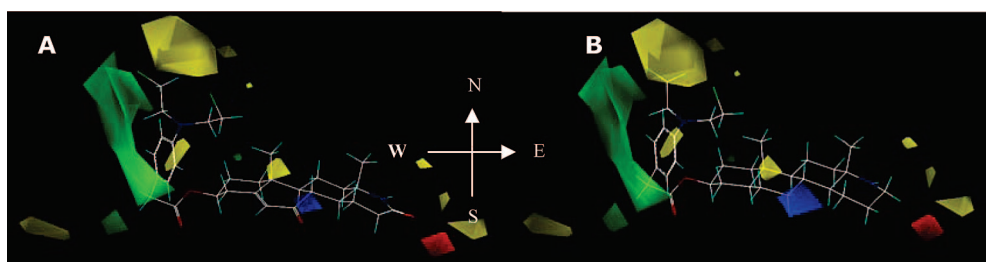

Figure 6. The steric/electrostatic contour map of the CoMFA. **A.** The very active template molecule **12** occupies the green colored contour in the region of the alkylating agent (N−W). The oxygen atom of the lactamic group of ring D is oriented toward the corresponding red colored contour (S-E). **B.** The inactive derivative **33** occupies a section of the yellow colored contour in the region of the alkylating agent (N) and does not possess any substituents with oxygen atom at the 17-position of ring D.

Figure 6 shows that the northern steroidal $ClCH_2CH_2-$ group of compounds **12** and **33** is oriented toward the green polyhedron where the bulky groups are favored according to the obtained model. Meanwhile, the other $ClCH_2CH_2-$ group is located relatively away from the yellow colored contour. In addition, the two relatively small yellow colored contours in the region of position 17 indicate that significantly bulky substituents decrease bioactivity as for example in the case of the inactive cholesterol derivative **5**.

Electrostatic Field. The red colored contours indicate regions in which the presence of the negative charge increases activity, while blue colored contours show regions where the negative charge is not favored. The contribution of the oxygen to the activity due to the electrostatic field in compound **12** is eminent. Such contribution is lacking for the inactive analog **33**. The blue polyhedron in the region of ring B is due to the favoring effect of the lactamic group (N atom).
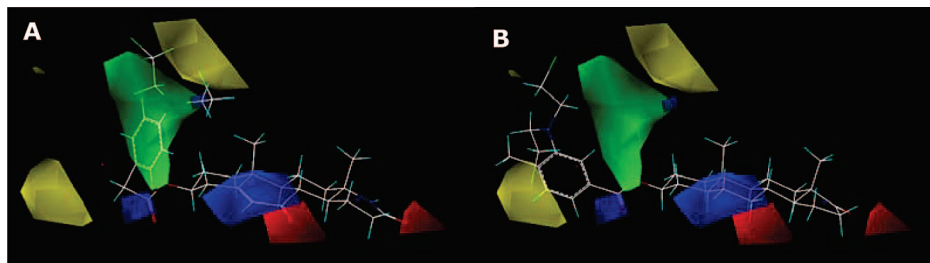
ANTILEUKEMIC STEROIDAL ESTERS

J. Chem. Inf. Model., Vol. 48, No. 11, 2008 **2259**

Figure 7. The steric/electrostatic contour map of the CoMSIA model shown together with the conformations for **A.** the very active derivative **12** and **B.** the inactive derivative **36**.
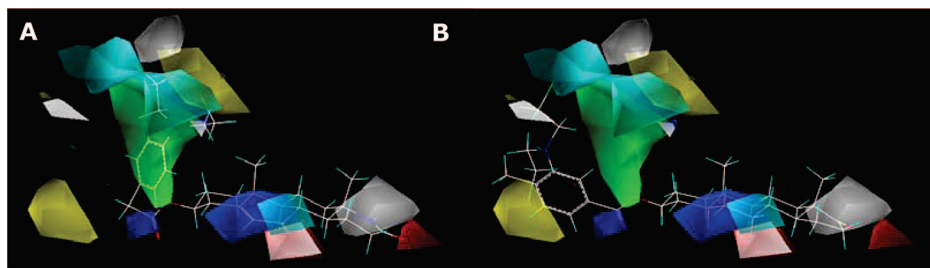

Figure 8. The contour map of the CoMSIA model that involves the steric and electrostatic fields along with hydrophobicity field shown together with the conformations for **A.** the very active derivative **12** and **B.** the inactive derivative **36.**
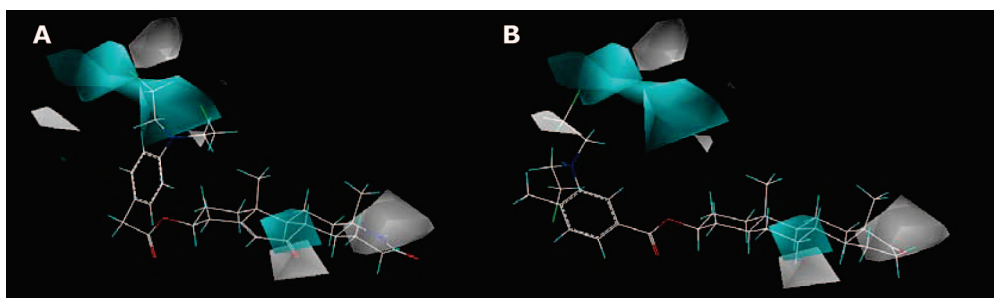

Figure 9. The contour map describing the hydrophobic field along with **A.** the highly bioactive template compound **12** and **B.** the inactive derivative **36**.
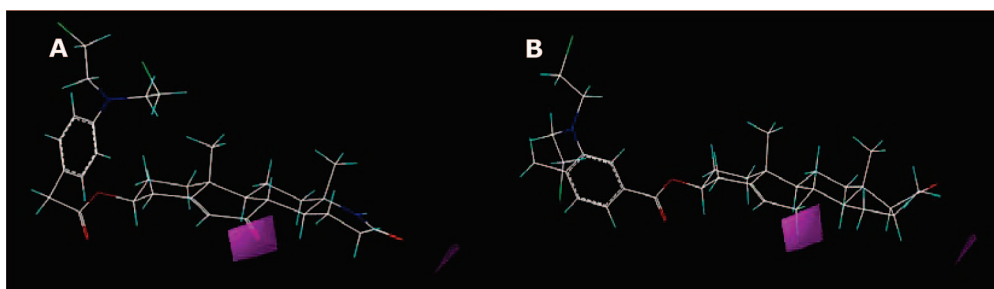

Figure 10. The contour map of the hydrogen-bond-acceptor field of the CoMSIA model together with the conformer for **A.** the strongly bioactive derivative **12** and **B.** the inactive derivative **20**.

2.4. CoMSIA Fields. *Steric-Electrostatic Fields.* The CoMSIA steric and electrostatic fields in Figure 7 are analogous to those in the CoMFA model (Figure 6). In addition, in CoMSIA is emphasized that the partial positive charge in ring B due to the conjugated systems C=CH-C=O increases activity. Such an increase is observed also for the conjugated system C=CH-C(O)-NH (see for example compounds **9** and **10**).

Hydrophobicity Field. According to the CoMSIA model, as shown in Figures 8 and 9, hydrophobicity is favored in ring B (cyan colored contour), while the O atom of position 7, C=O or the lactamic group C=CH-C(O)-NH coincides with the gray colored contour emphasizing the favoring effect of a hydrophilic substituent. The gray colored contour in the

region of rings B and D is correlated with the increase in bioactivity when SME possesses a hydrophilic group and/ or H-bond acceptor group at positions 7 and 17.

Hydrogen-Bond-Acceptor Field. In Figure 10 it is observed that in the highly active template molecule **12**, the magenta colored contour in the region of position 7 coincides with the oxygen atom of the -C=O group. On the contrary, in the inactive derivative **20** there is not any hydrogen-bond acceptor substituent in that region.

2.5. Predictions for the Test Set. The CoMFA and the selected CoMSIA model (CoMSIA 2) which comprised steric, electrostatic, hydrophobic, and H-bond acceptor fields were used for the prediction of the activity of the twelve compounds of the test set. The CoMFA and CoMSIA model

Table 4. Experimental Values of the Bioactivity and the Theoretical Values of the Test Group As Predicted by the CoMFA and CoMSIA Models

test set compd no.	T/C*		
	exp values	CoMFA	CoMSIA
4	141	95	129
6	311	213	241
9	153	222	237
10	296	205	235
14	257	250	247
17	198	195	223
21	139	124	120
26	268	153	188
27	178	159	155
28	133	188	187
32	226	218	180
35	144	137	159

predictions of the activity of the twelve compounds provided quite satisfactory results as more than half of the compounds were predicted quite precisely (Table 4).

The bioactivity of compounds **14**, **17**, **21**, **27** and **35** is predicted satisfactorily both by CoMFA and CoMSIA. In addition, bioactivity of compound **4** is predicted satisfactorily by the CoMSIA model and the activity of compound **32** is predicted satisfactorily by the CoMFA model.

Both models do not predict satisfactorily the activity of compounds **6**, **9**, **10**, **26**, and **28** which possess an extended lactamic ring B. The lactamic ring B may cause two major effects that are contributing to the poor prediction: (a) extension of the ring but more importantly (b) extension of the conjugated system due to the amide bond. This may have tremendous consequences on the stereoelectronic properties of the relatively poorly predicted compounds **6**, **9**, **10**, **26**, and **28.**

2.6. LeapFrog (LF) Predictions. LF is a designed *de novo* drug discovery program for predicting novel hits for medicinal chemists by repeatedly making structural changes and then either keeping or discarding the results.[34] The basic information about LF is explained in the Experimental Section 4.2. As a starting compound in the procedure in order to derive novel optimal compounds, the template compound **12** was used with restraining -N(CH$_2$CH$_2$Cl)$_2$ (NM groups) shown in the Experimental Section since NM groups are a necessary part for the activity of molecules. From the three available alternative modes (OPTIMIZE, DREAM, and GUIDE) we have chosen OPTIMIZE that suggests improvement of the existing leads. Since no information regarding the receptor was available in the literature, the available CoMFA contour maps were used to generate a hypothetical receptor cavity which was used to create the site points. The new ligand structures were evaluated on their binding energy. Only the ones that exhibited better binding energy than **12** were considered for further evaluation. This procedure was repeated for 1000 moves, and twenty of the new optimized compounds were collected and presented in Table 5. The proposed activity order of the derived compounds due to their different stereoelectronic parts was arranged in descending order.

The compounds presented in Table 5 were used as a starting database for generating novel compounds with higher predicted affinity values relative to the experimental value of **12**. These results are shown in Table 6.

Preliminary experimental tests for LF36 and LF21 compounds showed T′/C values of 350.3 and 365.2, respectively. As it can be observed from the predicted data shown in Table 6 (349 for LF36 and 369 for LF21) their difference in activity ranges from 0.37−1.04% indicating that the constructed model is highly reliable.

Some observations existing in Table 6 are important to report. The optimized hits by LF indicate that a G substitution group exerting an inductive effect (-I) such as -Cl enhances the bioactivity of the compound. G substitution groups exerting an inductive effect (-I) like -Cl and -CN can also enhance bioactivity when amide bonds adjacent to the carbonyl of D ring and in the connection of the steroidal moiety with mustards exist. The same interpretation can be given for the B substituent group exerting an inductive effect (-I) like the formyl group when amide bonds adjacent to the carbonyl of D ring and in the connection of the steroidal moiety with mustards exist. In addition, the K substitution group -CHCF$_3$ which exerts an inductive effect (-I) enhances activity when M substitution with an NH group exists.

3. CONCLUSIONS

The bioactive form of the studied SMEs with the flexible nitrogen mustard segments necessitated detailed conformational analysis in order to be amenable to 3D-QSAR studies. For this purpose the putative bioactive conformations of the molecules were obtained based on molecular modeling calculations in combination with 2D NMR spectroscopy.
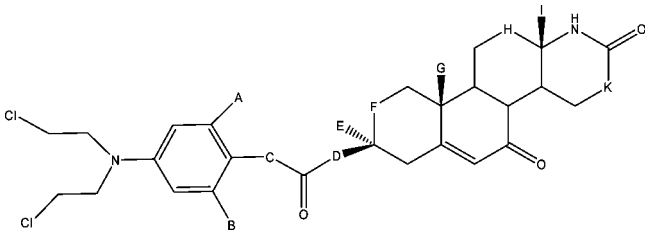
The attachment of steroid components on the mustard molecules, which are efficient antileukemic agents, reduces their toxicity giving SMEs with preserved activity and far less side effects, appropriate for pharmaceutical use. The DSC results in an earlier study showed that the steroidal segment reduced the extent of thermal effects of the alkylating factor in lipid bilayers. The new findings show that substitutions in rings B and D influence considerably the activity of mustards conjugates. Such findings necessitate additional studies using DSC necessary to prove possible QSAR relationships between bioactivity and reduction of the thermal effects of the alkylating segment.

CoMFA and CoMSIA 3D QSAR analyses of thirty-eight studied SMEs based on the highly bioactive molecule 3β-hydroxy-17α-aza-D-homo-7,17-dione-5-androsten-p-N,N-bis(2-chloroethyl)aminophenylacetic ester gave similar results indicating that the steric effect and the hydrophobic/hydrophilic balance especially in the steroidal part of the molecules determined activity. The r_{cv}^2 values of CoMFA and CoMSIA are in acceptable criterion for statistical validity (>0.5) and allow for the assumption of a significant QSAR.
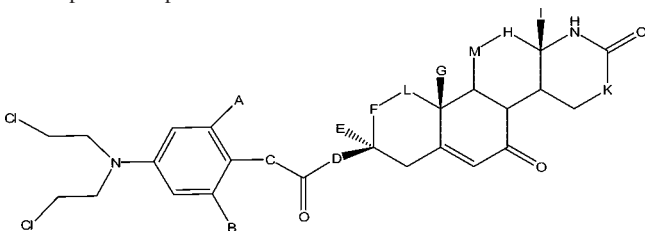
The orientation of the alkylating part of the SMEs toward the surface of ring B in the steroidal skeleton was related with increased activity provided that a minimum distance between the alkylating part and the steroidal skeleton was maintained so that the contact with the position 19-methyl group was avoided.

The presence of two hydrophilic/H-bond acceptor groups in rings B and D, particularly in positions 7 and 17 simultaneously, is correlated with increased activity.

The *de novo* synthesis of the described active derivatives resulted using LF are highly reliable as it is shown for two representative structures evaluated *in vivo*.

Table 5. Compounds Collected with Lowest Scoring Binding Energy by the LF Routine Comparison with the Template Compound **12**

compd	A	B	C	D	E	F	G	H	I	K
12	H	H	CH₂	O	H	CH₂	CH₃	CH₂	CH₃	CH₂
LF1	H	H	**CHEt**	O	H	CH₂	CH₃	CH₂	CH₃	CH₂
LF2	H	H	**CHNHCH₃**	**NH**	H	CH₂	CH₃	CH₂	CH₃	CH₂
LF3	H	H	**NH**	O	H	**NH**	CH₃	CH₂	CH₃	**NH**
LF4	H	H	CH₂	O	**NHCONH₂**	CH₂	CH₃	CH₂	CH₃	CH₂
LF5	H	H	**NH**	O	H	CH₂	CH₃	CH₂	CH₃	CH₂
LF6	H	H	CH₂	O	H	CH₂	**Cl**	CH₂	CH₃	CH₂
LF7	H	H	CH₂	O	H	CH₂	CH₃	CH₂	**Cl**	CH₂
LF8	H	H	CH₂	O	H	CH₂	**NH₂**	CH₂	CH₃	CH₂
LF9	H	H	CH₂	O	H	CH₂	CH₃	**CO**	Cl	CH₂
LF10	H	H	CH₂	O	H	CH₂	CH₃	CH₂	**NH₂**	CH₂
LF11	H	H	CH₂	O	H	CH₂	**CN**	CH₂	CH₃	CH₂
LF12	H	H	CH₂	O	H	**NH**	**CH₂NHCONH₂**	CH₂	CH₃	**NH**
LF13	H	H	**NH**	O	**NHOH**	CH₂	**CN**	CH₂	CH₃	CH₂
LF14	H	**3pyrrole**	CH₂	O	H	CH₂	CH₃	CH₂	CH₃	CH₂
LF15	**CHO**	H	CH₂	O	H	CH₂	**CH₂CH₂OH**	CH₂	CH₃	CH₂
LF16	H	**3pyrrole**	CH₂	O	H	CH₂	**Cl**	CH₂	CH₃	CH₂
LF17	H	H	CH₂	O	H	CH₂	**CH₂NHCONH₂**	CH₂	CH₃	**NH**
LF18	**Butyl**	H	CH₂	O	H	CH₂	**CH₂NHCONH₂**	CH₂	CH₃	**NH**
LF19	H	H	CH₂	**CH₂**	H	CH₂	**Cl**	CH₂	CH₃	CH₂
LF20	H	H	**CHEt**	**CH₂**	H	CH₂	CH₃	CH₂	CH₃	CH₂

Table 6. Compounds Collected with Lowest Scoring Binding Energy by the LF Routine and Further Assessed by the Predictions of Activity versus the Experimental Activity of the Template Compound **12**[a]


compd	B	C	E	F	G	H	K	L	M	pred.
12	H	CH₂	H	CH₂	CH₃	CH₂	CH₂	CH₂	CH₂	333
LF21	H	CH₂	H	CH₂	**Cl**	CH₂	CH₂	CH₂	CH₂	369
LF22	H	**NH**	**NHCH₃**	CH₂	**CN**	CH₂	CH₂	CH₂	CH₂	369
LF23	H	**NH**	H	**NH**	**Cl**	CH₂	**NH**	CH₂	CH₂	367
LF24	H	**NH**	H	**NH**	**CN**	CH₂	**NH**	CH₂	CH₂	366
LF25	H	**NH**	H	CH₂	**CN**	CH₂	**NH**	CH₂	CH₂	366
LF26	H	CH₂	H	CH₂	CH₃	CH₂	**CHCF₃**	CH₂	**NH**	366
LF27	**CHO**	CH₂	H	**CHCH₃**	CH₃	**NH**	CH₂	CH₂	CH₂	366
LF28	**CHO**	CH₂	H	**CHCH₃**	CH₃	CH₂	CH₂	**NH**	**NH**	366
LF29	H	**NH**	H	CH₂	**CN**	CH₂	**NH**	**NH**	CH₂	364
LF30	**CHO**	CH₂	H	**CHCH₃**	CH₃	**NH**	**NH**	CH₂	CH₂	364
LF31	**CHO**	CH₂	H	CH₂	CH₃	**NH**	**NH**	CH₂	CH₂	363
LF32	H	**CHCH₂OH**	H	CH₂	**Cl**	CH₂	**NH**	CH₂	CH₂	363
LF33	H	CH₂	H	CH₂	**CN**	CH₂	**NH**	CH₂	**NH**	362
LF34	H	**NH**	H	CH₂	**CN**	CH₂	CH₂	CH₂	CH₂	361
LF35	H	CH₂	H	**NH**	**CN**	CH₂	**NH**	CH₂	CH₂	360
LF36	H	**NCOCH₃**	H	CH₂	CH₃	CH₂	CH₂	CH₂	CH₂	349

[a] Only the molecules that had higher predicted score than template compound **12** are included. A is hydrogen, D is oxygen, I is a methyl group.

## 4. EXPERIMENTAL SECTION

4.1. Synthesis of the Tested Compounds. The synthesis of the final SMEs, which are reported in the present study, has been published at previous publications of our research team. Herein we report a novel synthetic procedure for the preparation of the final nitrogen mustards, which significantly improves the total yields of the final compounds. In the following reaction schemes this novel procedure is outlined, followed by the steps undertaken in this 3D QSAR investigation of the SMEs.

Scheme 1. General Procedure for the Preparation of the Final Nitrogen Mustards[a]

CHL:	R = H	X = (CH₂)₃	p-NH₂
PHE:	R = H	X = CH₂	p-NH₂
4-Me-CABA:	R = CH₃	X = -	m-NH₂

$$NM = -N(CH_2CH_2Cl)_2$$
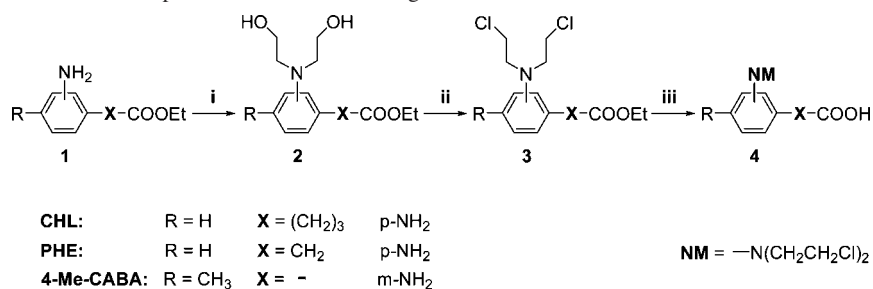
[a] Reagents and conditions: i) ethylene oxide/CH₃COOH, autoclave, rt, 12−24 h; ii) 1. MeSO₂Cl or (MeSO₂)₂O, Et₃N, and DMAP, 12−24 h; 2. LiCl/DMF, reflux, 4−6 h; iii) aq. HCl, neutralization.

Scheme 2. General Procedure for the Preparation of the Final Steroidal Alcohols[a]

CHL:	R = H	X = (CH₂)₃	p-NH₂
PHE:	R = H	X = CH₂	p-NH₂
4-Me-CABA:	R = CH₃	X = -	m-NH₂

$$NM = -N(CH_2CH_2Cl)_2$$
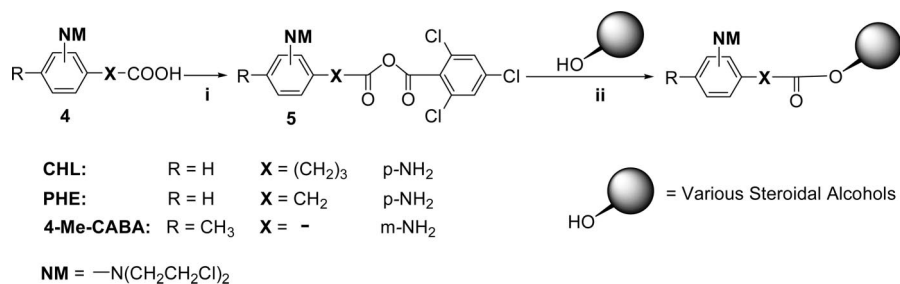
[a] Reagents and conditions: i) 2,4,6-trichlorobenzoylchloride(TCB-Cl)/Et₃N/toluene, reflux, 1−2 h; ii) 4-DMAP/toluene/reflux under Ar, 1.5−3 h.

4.2. Biological Evaluation. The details of biological test are reported elsewhere and are shown on the numbers of compounds included in Table 1. For the toxicity measurements and the *in vivo* antileukemic activities, BDF1 and BALB/C male and female rats were used. Toxicity was quantified using two toxicity indices. The index LD_{10} was the lethal dose for 10% of the animals with a single intraperitoneal (i.p.) injection. The index LD_{50} was the mean lethal dose of the compounds under investigation.

The evaluation of the anticancer activity of the compounds against leukemia P388 was performed as follows: The animals were initially infected by an i.p. injection of 10^6 leukemic cells in order to develop the disease. The animals were then divided into groups of six. In addition, one control group consisting of eight infected animals was also formed to which the inactive compound (placebo) corn oil was given. A dose of 1/2 LD_{10} of each compound in corn oil was injected into the other groups periodically on days one, four, and seven, also i.p.[29]

The antileukemic efficiency of each of the thirty-eight compounds was estimated using the term T/C %, which designates the quotient of the mean time of survival of the animals under therapy T (in days) to the mean time of survival of the control animals (C) × 100. The values of T/C % and LD_{10} are presented in Table 1 (the number of animals that were cured during therapy can be found as survivals in parentheses in the column T/C % of the table). As cured were considered the animals that survived at least thirty days after the initiation of the treatment. However, the index T/C % did not contain the animal survivals. In order to account for this additional fact, the following approach was used and T′/C % was introduced to show the results. For each compound for which there were animal survivals, the product of the number of the animals under treatment which did not survive for more than thirty days by the mean value of their survival was computed. Then, the product of

thirty days with the number of animals that survived for at least thirty days was also computed. Finally, the sum of the two products was computed and was divided by six, which was the number of animals per group under treatment. The obtained digit comprised the value of the corrected therapeutic survival time T′ which was further divided by C in order to give T′/C. The value of this quotient multiplied by 100 gave the modified due to the survivals in vivo activity values T′/C %.

For example, compound **12** gave three survivals. This means that out of the six animals under experimental treatment, three animals survived for at least thirty days. The mean value of survivals of the other three treated animals was 28.3 days. The mean value of the survival days of the control animals was 8.5 days. The T/C % without consideration of the survivals could thus be computed as

$$\frac{28.3}{8.5} \times 100 = 333 \qquad (1)$$

The new T′/C % is obtained as

$$\frac{\dfrac{(28.3 \times 3) + (30 \times 3)}{6}}{8.5} \times 100 = 343 \qquad (2)$$

In analogy, the values of T′/C % were computed for compounds **6**, **7**, **10**, and **22** for which there were survivals as well.

4.3. Computational Methodologies. The following steps have been applied throughout the 3D QSAR/CoMFA and CoMSIA analysis.

Step 1. 3β-Hydroxy-17α-aza-D-homo-7,17-dione-5-androsten-p-N,N-bis(2-chloroethyl)aminophenylacetic ester (compound **12**, in Table 1) was chosen as the template molecule since it has the highest bioactivity in the data set. The conformation of both the template compound **12** (top

ANTILEUKEMIC STEROIDAL ESTERS

J. Chem. Inf. Model., Vol. 48, No. 11, 2008 **2263**

structure in Figure 3) and its precursor molecule 3β-acetoxy-17α-aza-D-homo-5-androsten-7,17-dione (bottom structure in Figure 3) in solution had been determined earlier using a combination of 1D and 2D NMR spectra and molecular modeling studies.[30,31]

Step 2. The rest of the studied molecules were restrained to adopt the same conformation as the template compound **12**.

Step 3. All the obtained conformations of the rest of the molecules were superimposed with respect to the optimal template conformation. This is a critical step before CoMFA and CoMSIA can be applied.

Step 4. The set of thirty-eight compounds under study was divided into to a training and a test set; the training and test sets included twenty-six and twelve compounds, respectively. Thus, the training set contains more than twice the structures than the test set; a legitimate ratio usually used in CoMFA and CoMSIA studies. The molecules in the test set are selected to represent the whole data set in terms of their molecular structure and variety of activity.

Step 5. The CoMFA and CoMSIA models were constructed based on the training set and were evaluated by statistical quality indices.

Step 6. The test set was used for validation of the obtained model.

Step 7. Steroelectronic contour maps were generated.

Step 8. Novel molecules with predicted better activity were proposed using steroelectronic contour maps and the *de novo* design using the LEAPFROG option of the SYBYL molecular modeling package that uses CoMFA results.

The second generation *de novo* drug discovery module LF of the SYBYL molecular modeling package was used as an aid in the present work to design a series of new potentially active antileukemic compounds. LF performs by repeating various structural changes on a prechosen lead molecule and then keeps only the results satisfying the steric and chemical restraints of the binding site and discards the unsuccessful analogues depending on the binding energy results.[33−36] As an initial basic procedure of LF a virtual receptor site was generated using the CoMFA model inferred from the PLS result of the steroid mustards. The template molecule **12** was used as the starting structure with restraining the NM groups shown in the Experimental Section since NM groups are a necessary part for the activity of molecules. Binding energy calculations in LF were performed by steric, electrostatic, and hydrogen bonding enthalpies of ligand-cavity binding using the Tripos force field.[28]

The mode OPTIMIZE, suggesting improvements to the existing lead, was used in the present work, with the additional constraint of minimizing synthetic difficulties during the seeking procedure. JOIN, FUSE, WEED, and CROSSOVER modules were performed after the initial run of 100 moves, and the derived ligands that had the best binding energy were used for the repeating cycle of 1000 moves.

Abbreviations: SME, Steroidal Mustard Esters; r_{cv}^2, cross-validated r^2; r^2, square correlation coefficient; QSAR, Quantitative Structure Activity Relationship; CoMFA, Comparative Molecular Field Analysis; CoMSIA, Comparative Molecular Similarity Indices Analysis.

REFERENCES AND NOTES

(1) Ferguson, L. R.; Denny, W. A. Anticancer drugs: an underestimated risk or an underutilized resource in mutagenesis. *Mutat. Res.* **1995**, *331*, 1–16.

(2) Catsoulacos, P.; Camoutsis, C.; Wampler, G. Effect of a D5-Homo-Aza-Steroidal Ester in P388 and L1210 Murine Leukemias. *Oncology* **1982**, *39*, 59–60.

(3) Arsenou, E. S.; Fousteris, M. A.; Koutsourea, A. I.; Papageorgiou, A.; Karayianni, G.; Mioglou, E.; Iakovidou, Z.; Mourelatos, D.; Nikolaropoulos, S. The allylic 7-ketone at the steroidal skeleton is crucial for the antileukemic potency of chlorambucil's active metabolite steroidal esters. *Anti-Cancer Drugs* **2004**, *15*, 983–990.

(4) Fousteris, M.; Koutsourea, A. I.; Arsenou, E.; Papageorgiou, A.; Mourelatos, D.; Nikolaropoulos, S. Antileukemic and Cytogenetic Effects of Modified and Non-modified Esteric Steroidal Derivatives of 4-Methyl-3-bis(2-chloroethyl)amino benzoic acid (4-Me-CABA). *Anticancer Res.* **2002**, *22*, 2293–2300.

(5) Anastasiou, A.; Catsoulacos, P.; Papageorgiou, A.; Margariti, E. On the Formation of Homo-azasteroidal Esters of N, N -bis-(2-chloroethyl) aminobenzoic Acid Isomers and their Tumor Activity. *J. Heterocycl. Chem.* **1994**, *31*, 367–374.

(6) Pairas, G.; Catsoulacos, P.; Papageorgiou, A.; Boutis, L. Further Studies on the Antineoplastic Activity of Homo-aza-Steroidal Esters. *Oncology* **1986**, *43*, 344–348.

(7) Pairas, G.; Catsoulacos, P. On the synthesis of homo-aza-steroidal-esters. *Eur. J. Med. Chem.-Chim. Ther* **1986**, *21*, 525–526.

(8) Karayianni, V.; Mioglou, E.; Iakovidou, Z.; Mourelatos, D.; Fousteris, M.; Koutsourea, A.; Arsenou, E.; Nikolaropoulos, S. A new approach for evaluating in vivo anti-leukemic activity using the SCE assay. An application on three newly synthesised anti-tumour steroidal esters. *Mutat. Res.* **2003**, *535*, 79–86.

(9) Karayianni, V.; Papageorgiou, A.; Mioglou, E.; Iakovidou, Z.; Mourelatos, D.; Fousteris, M.; Koutsourea, A.; Arsenou, A.; Arsenou, E.; Nikolaropoulos, S. 7-keto-hybrid steroidal esters of nitrogen mustard: cytogenetic and antineoplastic effects. *Anti-Cancer Drugs* **2002**, *13*, 637–643.

(10) Papageorgiou, A.; Boutis, L.; Nikolaropoulos, S.; Catsoulacos, P. Potential antitumor agents, steroidal amidoester with an alkylating moiety. *Oncology* **1987**, *44*, 128–132.

(11) Mourelatos, D.; Papageorgiou, A.; Boutis, L.; Catsoulacos, P. Comparative study on cytogenetic damage induced by homo-aza-steroidal esters in human lymphocytes. *Mutat. Res.* **1995**, *334*, 19–22.

(12) Catsoulacos, P.; Camoutsis, C.; Papageorgiou, A.; Adamiac-Margariti, E. Cytostatic effect of Homo-aza-steroidal Esters *in Vivo* and *in Vitro*. Structure-Activity Relationships. *Anticancer Res.* **1992**, *12*, 1617–1620.

(13) Nikolaropoulos, S.; Tsavdaridis, D.; Arsenou, E.; Papageorgiou, A.; Karaberis, E.; Mourelatos, D. Synergistic antineoplastic and cytogenetic effects by the combined action of two homo-aza-steroidal esters of nitrogen mustards on P388 and L1210 leukaemias *in vivo* and *in vitro*. *Anticancer Res.* **2000**, *20*, 2745–2752.

(14) Papageorgiou, A.; Nikolaropoulos, S. S.; Arsenou, S. E.; Karaberis, E.; Mourelatos, D.; Kotsis, A.; Chrysogelou, E. Enhanced cytogenetic and antineoplastic effects by the combined action of two esteric steroidal derivatives of nitrogen mustards. *Chemotherapy* **1999**, *45*, 61–67.

(15) Nikolaropoulos, S.; Arsenou, E. S.; Papageorgiou, A.; Mourelatos, D. Antitumour and cytogenetic effects of steric (ASE) and amidic (ASA) steroidal derivatives of p-bis (2-chloroethyl) amino phenylacetic acid (CAPA). A. comparative study. *Anticancer Res.* **1997**, *17*, 4525–4530.

(16) Mourelatos, D.; Mylonaki, E.; Papageorgiou, A.; Boutis, L.; Paradelis, A.; Anastasiou, A.; Catsoulacos, P. Comparative study of SCE induction and cytostatic effects by homo-azasteroidal esters of N, N-bis(2-chloroethyl) aminobenzoic acid in human lympocytes. *Mutat. Res.* **1995**, *346*, 129–133.

(17) Papageorgiou, A.; Boutis, L.; Nikolaropoulos, S.; Catsoulacos, P. Potential antitumor agents, steroidal amidoester with an alkylating moiety. *Oncology* **1987**, *44*, 128–132.

(18) Papageorgiou, A.; Catsoulacos, P.; Mourelatos, D.; MioglouKalouptsi, E.; LakovidouKritsi, Z.; Boutis, L.; Kotsis, A. A comparative study

of the cytogenetic and antineoplasmatic effects induced by steroidal amidoesters of p-N, N-bis (2-chloroethyl) aminophenylbutyric acid (chlorambucil). *Cancer J.* **1996**, *9*, 203–206.

(19) Kapou, A.; Nikolaropoulos, S.; Siapi, E.; Mavromoustakos, T. Effects of steroidal carriers of alkylating agents on the phase transition in DPPC membrane bilayers. *Thermochim. Acta* **2005**, *429*, 53–56.

(20) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *18*, 5959–5967.

(21) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.

(22) Klebe, G. Comparative molecular similarity indices analysis: CoMSIA. *Perspect Drug Discovery Des.* **1998**, *12*, 87–104.

(23) Mavromoustakos, T.; Kapou, A.; Benetis, N. P.; Zervou, M. Conformational Analysis using 2D NMR Spectroscopy Coupled with Computational Analysis as an Aid in the Alignment Procedure of 3D QSAR Studies. *Drug Des. Rev. Online* **2004**, *3*, 235–245.

(24) Durdagi, S.; Kapou, A.; Kourouli, T.; Andreou, T.; Nikas, S. P.; Nahmias, V. R.; Papahatjis, D. P.; Papadopoulos, M. G.; Mavromoustakos, T. The Application of 3D-QSAR Studies for Novel Cannabinoid Ligands Substituted at the C1′ Position of the Alkyl Side Chain on the Structural Requirements for Binding to Cannabinoid Receptors CB1 and CB2. *J. Med. Chem.* **2007**, *50*, 2875–2885.

(25) Durdagi, S.; Papadopoulos, M.; Papahadjis, D.; Mavromoustakos, T. Combined 3D QSAR and Molecular Docking Studies to Reveal Novel Cannabinoid Ligands with Optimum Binding Activity. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6754–6763.

(26) Durdagi, S.; Heribert, R.; Papadopoulos, M. G. Comparative Molecular Dynamics Simulations of the Potent Synthetic Classical Cannabinoid Ligand AMG3 in Solution and at Binding Site of the CB1 and CB2 Receptors. *Bioorg. Med. Chem.* **2008**, *16*, 7377–7387.

(27) Kapou, A.; Benetis, N.-P.; Avlonitis, N.; Calogeropoulou, T.; Koufaki, M.; Scoulica, E.; Nikolaropoulos, S.; Mavromoustakos, T. 3D-Quantitative Structure-Activity Relationships of Synthetic Antileishmanial Ring-Substituted Ether Phospholipids. *Bioorg. Med. Chem.* **2007**, *15*, 1252–1265.

(28) Sybyl molecular modeling software packages, ver. 6.8, 2001, Tripos Inc., St Louis, MO 63144.

(29) Kleine, I.; Platonova, G. N. *Experimental Evaluation of anticancer drugs in the USA and USSR and clinical correlation*; Goldin, A.; Kleine, I.; Sofina, Z. P.; Syrkin, A. B.; Eds.; National Cancer Institute Monograph 55, NIH publication No. 80-193, NCI: Bethesda, MD, 1980; Vol. 25, p 20205.

(30) Kapou, A.; Fousteris, M. A.; Nikolaropoulos, S.; Zervou, M.; Grdadolnik, S. G.; Zoumpoulakis, P.; Kyrikou, I.; Mavromoustakos, T. 2D NMR and conformational analysis of a prototype anti-tumour steroidal ester. *J. Pharm. Biomed. Anal.* **2005**, *38*, 428–434.

(31) Kapou, A.; Mavromoustakos, T.; Grdadolnik, S. G.; Nikolaropoulos, S. *Conformational Analysis of Bioactive Compounds in Drug Discovery and Design: Medical Aspects, Part III*; Matsoukas, J., Mavromoustakos, T., Eds.; IOS Press: 2002; Vol. 16, pp 7–173.

(32) Dunn, W. J., III; Wold, S.; Edlund, U.; Hellberg, S.; Gasteeger, J. Multivariate Structure Activity Relationships between Data from a Battery of Biological Test and an Ensemble of Structure Descriptors. The PLS method. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131–136.

(33) Nair, C. P.; Sobiha, M. E. CoMFA based de novo design of pyridazine analogs as PTP1B inhibitor. *J. Mol. Graphics Modell.* **2007**, *26*, 117–123.

(34) Makhija, M. T.; Kasliwal, R. T.; Kulkarni, V. M.; Neamati, N. De novo design and synthesis of HIV-1 integrase inhibitors. *Bioorg. Med. Chem.* **2004**, *12*, 2317–2333.

(35) Durdagi, S.; Mavromoustakos, T.; Papadopoulos, M. G. 3D QSAR CoMFA/CoMSIA, molecular docking and molecular dynamics studies of fullerene-based HIV-1 PR inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, in press.

(36) Durdagi, S.; Mavromoustakos, T.; Chronakis, N.; Papadopoulos, M. G. Computational design of novel fullerene analogues as potential HIV-1 PR inhibitors: Analysis of the binding interactions between fullerene inhibitors and HIV-1 PR residues using 3D QSAR, molecular docking and molecular dynamics simulations. *Bioorg. Med. Chem.* **2008**, in press.

CI800240M