

Antitumor Agents. 163.[†] Three-Dimensional Quantitative Structure–Activity Relationship Study of 4'-*O*-Demethylepipodophyllotoxin Analogs Using the Modified CoMFA/ q^2 -GRS Approach

Sung Jin Cho,[‡] Alexander Tropsha,^{*‡} Matthew Suffness,[§] Yung-Chi Cheng,[⊥] and Kuo-Hsiung Lee^{*‡}

Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599, Division of Cancer Treatment, National Cancer Institute, Bethesda, Maryland 20892, and Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06510

Received April 24, 1995[⊗]

Analogs of 4'-*O*-demethylepipodophyllotoxin are considered as potential anticancer agents. We have applied comparative molecular field analysis (CoMFA) and a novel CoMFA/ q^2 -GRS technique recently developed in our group to identify the essential structural requirements for increasing the ability of these compounds to form cellular protein–DNA complex. In addition, a new method to incorporate different types of probe atoms as part of q^2 -GRS routine has been developed. The best final model with 101 compounds using a combination of four different sets of probe atoms and charges [C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1)] yielded a q^2 of 0.584 and the standard error of prediction of 0.660 at 5 principal components. The steric and electrostatic contour plots of the final model were compared with the DNA phosphate backbone environment of the DNA–4'-*O*-demethylepipodophyllotoxin analog complex, which was generated using the X-ray structure of the DNA–nogalamycin complex. The comparison reveals that the CoMFA steric and electrostatic fields are compatible with stereochemical properties of the DNA backbone. The results obtained from this study shall guide our future synthetic efforts.

Introduction

Podophyllotoxin is a plant toxin that inhibits the assembly of microtubules. It was first isolated by Podwysstotzki in 1880 from the North American plant *Podophyllum peltatum* Linnaeus, commonly known as the American mandrake or May apple.² Attempts to use podophyllotoxin in the treatment of human neoplasia were mostly unsuccessful^{3,4} and complicated by side effects such as nausea, vomiting, diarrhea, and damage to normal tissues.⁴ Extensive structure modifications of podophyllotoxin have been performed in order to obtain more potent and less toxic anticancer agents, which resulted in the synthesis of etoposide and teniposide, the semisynthetic glucosidic cyclic acetals of epipodophyllotoxin. These compounds have been shown to be active in the treatment of a number of cancers, including lymphomas; acute leukemia; cancers of the lung, ovary, testis, bladder, and brain; and Kaposi's sarcoma associated with the acquired immunodeficiency syndrome.^{4,5} A large number of etoposide analogs have been synthesized and tested in order to improve the clinical efficacy of etoposide and overcome some problems associated with its use as a drug, such as the development of drug resistance, myelosuppression, and poor oral bioavailability.^{6–8} The efforts toward the development of novel antitumor etoposide analogs were recently reviewed by Zhang and Lee.⁹

Over the years, we have been involved in the synthesis of a number of 4'-*O*-demethylepipodophyllotoxin derivatives^{10–15} possessing various substitutions at 4-position (Figure 1). The rationale behind the synthe-

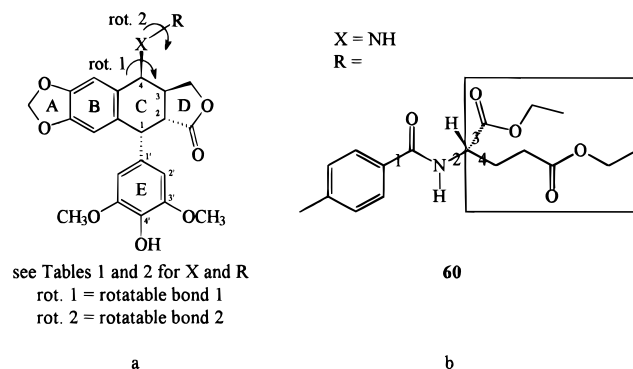


Figure 1. The generalized chemical structure of 4'-*O*-demethylepipodophyllotoxin analogs.

sis of these analogs stems from the mechanism of action of etoposide and its congeners. Unlike podophyllotoxin, the primary mode of action of these compounds is to stabilize the covalent topoisomerase II–DNA cleavage complex, break DNA strands, and cause cell death.^{16,17} We recognized from the early SAR study¹⁸ that the conversion of the primary mode of action from antimicrotubule (of podophyllotoxin) to antitopoisomerase (of etoposide and its congeners) requires three important chemical modifications of podophyllotoxin, demethylation of the 4'-methoxy of the E ring, epimerization at the 4-position, and substitution (glucosidation or others) at the 4-position. Since antitopoisomerase activity varied greatly with the types of substitution at 4-position, we decided to explore this region of etoposide further and found that some of the 4'-*O*-demethylepipodophyllotoxin derivatives showed better antitumor activity profiles than etoposide.^{12,19}

As the number of derivatives increases, the formulation of a useful SAR becomes increasingly difficult. In this paper, we report the development of three-dimensional quantitative structure–activity relationships (3-D

[†] For Part 162, see ref 1.

^{*} To whom correspondence should be addressed.

[‡] University of North Carolina at Chapel Hill.

[§] National Cancer Institute.

[⊥] Yale University School of Medicine.

[⊗] Abstract published in *Advance ACS Abstracts*, March 1, 1996.

Table 1. Ability of Substituted Alkyl and Arylamino Derivatives of 4'-*O*-Demethylepipodophyllotoxin To Form a Cellular Protein-DNA Complex

Compound	R	Cellular Protein-DNA Complex Formation (%)	Ref.	Compound	R	Cellular Protein-DNA Complex Formation (%)	Ref.
	(for reference 10)				(for references 11 and 12)		
Etoposide		100	10	29		164	11
1	β -OH	42.2	10	30		279	11
2	α -OH	3.3	10	31		97	11
3	β -NH ₂	36.4	10	32		140	11
4	α -NH ₂	8.0	10	33		230	11
5	β -NHCH ₂ CH ₂ OH	121.4	10	34		323	11
6	α -NHCH ₂ CH ₂ OH	0.0	10	35		15	11
7	β -NHCH ₂ CH ₂ CH ₃	69.7	10	36		21	11
8	β -NHCH ₂ CH ₂ CH ₃	110.8	10	37		97	11
9	β -NHCH ₂ CH=CH ₂	84.1	10	38		148	11
10	β -NHCH ₂ CH(OH)CH ₃ (R)	167.2	10	39		123	11
11	β -NHCH(CH ₃)CH ₂ OH (R)	161.7	10	40		140	11
12	β -NHCH ₂ CH ₂ CH ₂ OH	89.2	10	41		330	11
13		290	11	42		11	11
14		213	11	43		57	11
15		243	11	44		34	11
16		137	11	45		10	11
17		211	11	46		190	12
18		4	11	47		183	12
19		249	11	48		83	12
20		207	11	49		172	12
21		83	11				
22		129	11				
23		50	11				
24		104	11				
25		150	11				
26		235	11				
27		180	11				
28		47	11				

Table 1 (Continued)

Compound	R	Cellular Protein-DNA Complex Formation (%)	Ref.	Compound	R	Cellular Protein-DNA Complex Formation (%)	Ref.
50		77	12	56		17	12
51		140	12	57		138	12
52		203	12	58		1.9	12
53		183	12	59		6.9	12
54		186	12	60		83	12
55		179	12				

QSARs) of 101 analogs of 4'-*O*-demethylepipodophyllotoxin using the methodology of comparative molecular field analysis (CoMFA)²⁰ and a novel CoMFA/q²-GRS technique recently developed by Cho and Tropsha.²¹ We also describe a new method to incorporate different types of probe atoms using the q²-GRS routine. The final model with 101 compounds using a combination of four different sets of probe atoms and charges [C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1)] resulted in a *q*² value of 0.584 and the standard error of prediction (SDEP) of 0.660 at 5 principal components (PCs).

Biological Activity

Compounds shown in Tables 1 and 2 were tested for their abilities to form intracellular covalent topoisomerase II-DNA complexes. The assay was carried out according to the procedures described previously.¹⁰ The original activity data expressed as the percentage of cellular protein-DNA complex formed (PCPDCF) were transformed by taking the natural logarithm of PCPDCF increased by one so that we could incorporate **6** into the analysis, *i.e.*, ln(PCPDCF + 1). These transformed activities were used in the subsequent CoMFA studies.

Computational Methods

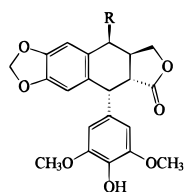
Sybyl molecular modeling software²² was used for structure generation and CoMFA.²⁰ Structure optimization and field-fit minimization were performed using the standard Tripos force field²³ with the maximum iteration cutoff of 1000 steps. The weighted root mean square distance between the X-ray structure and the optimized X-ray structure using the Tripos force field was found to be 0.84 Å, which indicated the adequate accuracy of the force field for our calculations. Sybyl grid (10° increment) and systematic (10° increment and energy option turned on with electrostatic and the maximum energy difference of 0.1 kcal/mol) search methods were used to search for the lowest-energy conformers. All calculations were done on IBM RS6000 Model 340. The default SYBYL settings were used except otherwise noted.

Structure Generation and Alignment Rules. One of the goals of this study was to test the predictability of CoMFA. The chemically different test compounds (compared to the compounds in the training set) are generally recommended for assessing the true predictability of CoMFA.²⁴ We have analyzed the structure-activity relationships for the total of 102 compounds which were initially broken into two groups: the

training set (Table 1) and the test set (Table 2). The compounds in Table 1 were selected as the training set because they consist of a wide range of structures, including substituted alkyl and arylamino derivatives of 4'-*O*-demethylepipodophyllotoxin. Furthermore, the existence of the X-ray crystal structure of **14** allowed unambiguous determination of the conformation around the rotatable bond **1** (Figure 1a). We selected structurally different benzyl, benzoylamino, and benzoyl derivatives of 4'-*O*-demethylepipodophyllotoxin (Table 2) as test compounds. We have attempted several alignment rules as summarized in Table 3 and described in more detail below.

After compounds in the training (Table 1: etoposide, **1–13**, and **15–60**) and test sets (Table 2: **61–101**) were generated by modifying X-ray crystal structure of **14**, the geometry of all of the compounds (including **14**) was optimized. Charge calculations were performed using the Gasteiger-Huckel method as implemented in Sybyl.²² For the 61 compounds (except for **1–4** and **14**) in the training set, the torsion angle around the rotatable bond linking the C ring and R group was manually modified to fit that of **14** ("rot. 1" in Figure 1a). The lowest-energy conformers of etoposide, **13**, **16**, **29**, **30**, **33**, **35–40**, and **42**, were searched for using Sybyl grid search method ("rot. 2" in Figure 1a was allowed to rotate). For **5–12** the Sybyl systematic search method was used to obtain the lowest-energy conformers. To obtain the lowest-energy conformers of **18–28**, **31**, **32**, and **46–59**, the Sybyl systematic search method was first applied to all rotatable bonds. Sybyl grid search method was then applied to rot. 2 in Figure 1a. Due to a large total number of rotatable bonds, the conformational search for two ester groups (surrounded by a box in Figure 1b) of **60** was first conducted using the Sybyl systematic search method. Four rotatable bonds (indicated by 1, 2, 3, and 4 in Figure 1b) were searched next using the Sybyl systematic search method.

Following the conformational search, the structure of each compound was reoptimized. Four centroids were then defined for the B, C, D, and E rings (Figure 1a). Each compound in the training set (Table 1) was rms-fitted to **14** using these four centroids. It has been suggested that the binding mode of etoposide might be similar to other DNA topoisomerase II inhibiting agents such as daunorubicin and nogalamycin.^{25–27} This prompted us to construct a DNA-etoposide complex using the X-ray structure of a DNA-nogalamycin complex reported by Gao et al.²⁸ This was accomplished by superimposing the B ring of etoposide onto the E ring of nogalamycin in its complex with DNA followed by removing nogalamycin from the complex. Examination of this crude model of a DNA-etoposide complex suggested that in the complexes with DNA the bulky and/or negatively charged functional groups of ligands should point away from the DNA to avoid bad steric

Table 2. Ability of Substituted Aryl, Benzyl, and Benzoylamino and Benzoyl Derivatives of 4'-*O*-Demethylepipodophyllotoxin To Form Cellular Protein–DNA Complex

(for References 13, 14, and 15)

Compound	R	Cellular Protein-DNA Complex Formation (%) ^a	Ref.	Compound	R	Cellular Protein-DNA Complex Formation (%) ^a	Ref.
61		151	13	83		144	14
62		211	13	84		191	14
63		121	13	85		184	14
64		158	13	86		177	15
65		115	13	87		160	15
66		32	13	88		128	15
67		51	13	89		116	15
68		99	13	90		117	15
69		62	13	91		137	15
70		179	13	92		124	15
71		64	13	93		149	15
72		181	14	94		159	15
73		216	14	95		86	15
74		130	14	96		160	15
75		144	14	97		149	15
76		126	14	98		120	15
77		216	14	99		94	15
78		169	14	100		100	15
79		225	14	101		94	15
80		284	14				
81		99	14				
82		159	14				

^a The percentage of cellular protein–DNA complex formed by etoposide is 100.

Table 3. Alignment Rules

alignment	no. of compds included		compd 58	RMS fit	field fit	DNA site minimization	probe atom	charge
	training set	testing set						
1	61		included	x			C (sp ³)	+1
2	61		included	x	x		C (sp ³)	+1
2a	60		not included	x	x		C (sp ³)	+1
2b	60	41	not included	x	x		C (sp ³)	+1
2c	60	41	not included	x	x		C (sp ³)	0
2d	60	41	not included	x	x		H	+1
2e	60	41	not included	x	x		O (sp ³)	-1
2f	60	41	not included	x	x		combination ^a	
3	60		not included	x		x	C (sp ³)	+1

^a Combination of C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1).

and electrostatic interaction, whereas positively charged functional groups should interact favorably with the negatively charged phosphate backbone of DNA. On the basis of this analysis, each compound in the training set was docked into DNA in place of nogalamycin as described above for etoposide. The conformation of each compound was modified by increasing the dihedral angle defined by rotatable bond rot. 2 (Figure 1a) by 180° to eliminate apparently bad contacts with DNA; this modification eliminated bad contacts with DNA without raising the conformer energy. The superimposition of all compounds in the training set in their DNA-bound conformation thus obtained will be referred to as alignment 1 (Table 3).

To obtain maximum steric and electrostatic field overlap between the most active compound (41) and other compounds in the training set (Table 1), the field-fit minimization was applied to compounds generated via alignment 1. The structures of all compounds were then reoptimized with the field-fit option turned off. This will be referred to as alignment 2 (Table 3).

In order to mimic the DNA-bound conformations of the compounds, their B rings (Figure 1a) were rms-fitted to the E ring of nogalamycin in the X-ray structure of the DNA–nogalamycin complex. Nogalamycin and water were removed from the complex, and charges for the compound–DNA complex were computed using Gasteiger–Huckel method. DNA and A, B, C, D, and E rings of each compound were defined as an aggregate, and each complex of the 4'-*O*-demethylepipodophyllotoxin analog with DNA was optimized to remove any bad steric and electrostatic interactions between the DNA phosphate backbone and the R groups of 4'-*O*-demethylepipodophyllotoxin analogs. After removing the DNA, the compounds were reoptimized to fix any distorted internal geometry. This alignment will be referred to as alignment 3 (Table 3).

To generate the test set (Table 2), 41 substituted anilino, benzyl, and benzoyl derivatives of 4'-*O*-demethylepipodophyllotoxin (Table 2) were built, and charges were computed using the Sybyl Gasteiger–Huckel method (following the procedure described above for the alignment 2). The field-fit minimization was utilized to obtain the maximum overlap between the steric and electrostatic fields of these compounds and those of compound 41. The structures of all compounds were then reoptimized with the field-fit option turned off. The final CoMFA models for all compounds (including both training and test sets) were generated using different probe atoms. Alignments 2b–f correspond to CoMFA/q²-GRS with C (sp³, +1), C (sp³, 0), H (+1), O (sp³, -1), and a combination of C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1), respectively (Table 3).

Conventional CoMFA. Conventional CoMFA was performed with the QSAR option of Sybyl. For each cross-validated CoMFA analysis, the minimum σ value was set to 2.0 to expedite calculations. For non-cross-validated CoMFA analyses, the minimum σ value was set to 0. The steric and electrostatic field energies were calculated using sp³ carbon probe atoms with +1 charge. The CoMFA grid spacing was 2.0 Å in all three dimensions within the defined region, which extended beyond the van der Waals envelopes of all molecules by at least 4.0 Å. The CoMFA QSAR equations were calculated with the PLS algorithm. The optimal number of components in the final PLS model was determined by the

standard error of prediction value, obtained from the leave-one-out cross-validation technique.

Modified q²-GRS Routine with Multiple Probe Atoms. Earlier, we have reported the novel q²-GRS routine in CoMFA.²¹ The application of this method leads to reproducible q² values that do not depend on the orientation of the molecular aggregate on the user terminal²¹ unlike standard CoMFA. We have modified the original q²-GRS routine in order to incorporate different types of probe atoms (Figure 2). The modified version of the q²-GRS routine consists of the following steps: (1) a conventional CoMFA is performed initially using an automatically generated region file; (2) the rectangular grid encompassing aligned molecules is then broken into 125 small boxes of equal size, and the Cartesian coordinates of the upper right and lower left corners of each box are calculated; (3) the coordinates calculated from step 2 are used to create region files with four different probe atoms, including C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1); (4) for each of this newly generated region files, a separate CoMFA is performed with the step size of 1.0 Å; (5) the q² values are compared to select the best probe atom for each region; (6) the regions with the q² value greater than the specified cutoff are selected for further analysis; (7) the selected regions are combined to generate a master region file; and (8) the final CoMFA is performed.

Results

CoMFA/q²-GRS of 4'-*O*-Demethylepipodophyllotoxin Analogs Included in the Training Set. The results obtained after performing CoMFA/q²-GRS are summarized in Table 4. For each alignment, the predictability of the CoMFA model was initially assessed by conventional CoMFA. The q²-GRS routine was then applied to optimize the standard CoMFA model. Initial conventional CoMFA run using alignment 1 produced a q² of 0.370 and SDEP of 1.051 with 4 PCs (Table 4). Various different q² cutoff values were tried to remove irrelevant variables (noise) associated with the decreasing of the lattice spacing. The highest q² value (0.448; 4 PCs) and lowest SDEP value (0.984; 4 PCs) were obtained with the q² cutoff value of 0.1 for alignment 1 (Table 4).

The conventional CoMFA run using alignment 2 gave a slightly lower q² (0.338; 4 PCs) and higher SDEP (1.077; 4 PCs) values compared to those of alignment 1 (Table 4). Among the four different q² cutoffs that were used to optimize the initial PLS model, the q² cutoff of 0.1 gave the highest q² (0.503; 4 PCs) and lowest SDEP (0.933; 4 PCs) values for alignment 2 (Table 4). The q² values obtained using alignment 2 generally gave higher q² values compared to those of alignment 1. The subsequent non-cross-validated PLS run for the alignment 2 at 0.1 q² cutoff yielded the conventional r² of 0.820 and the standard error of estimate (SDEE) of 0.562 (Table 5). Using this PLS model, the activity of each compound was calculated and compared with the

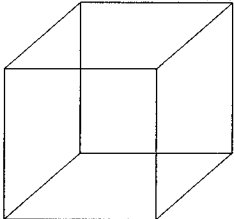
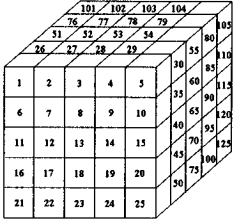
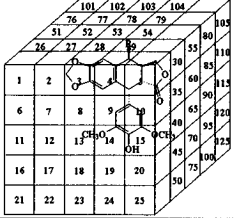
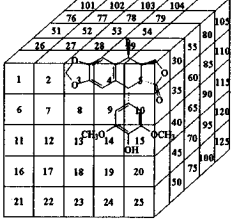
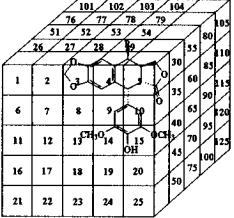
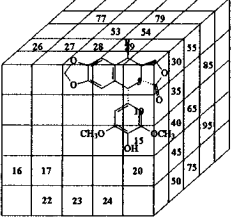
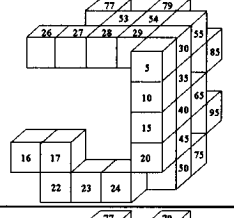
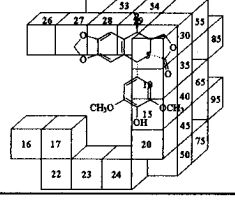
Step 1:	Perform CoMFA using an automatically created region file.	Probe atom: C.3 Charge: +1 Step size: 2.0	
Step 2:	Use the region file generated from Step 1 to calculate the Cartesian coordinates of the upper right and lower left corners of 125 small region files.		
Step 3:	Use the coordinates calculated from Step 2 to create region files with different probe atoms and charges.	Set 1: C.3, +1, 1.0 Set 2: C.3, 0, 1.0 Set 3: H, +1, 1.0 Set 4: O.3, -1, 1.0	
Step 4:	Perform CoMFA using region files generated from Step 3.	Set 1: C.3, +1, 1.0 Set 2: C.3, 0, 1.0 Set 3: H, +1, 1.0 Set 4: O.3, -1, 1.0	
Step 5:	Select the best probe atom and charge for each region using the q^2 value as a guide.	Set 1: Regions 1-20 Set 2: Regions 21-39 Set 3: Regions 40-70 Set 4: Regions 71-125	
Step 6:	Select those regions which the q^2 value is greater than a specified cutoff value.	Set 1: 5, 10, 15-17, 20 Set 2: 22-24, 26-29, 30, 35 Set 3: 40, 45, 50, 53-55, 65 Set 4: 75, 77, 79, 85, 95	
Step 7:	Create a master region file.	Set 1: 5, 10, 15-17, 20 Set 2: 22-24, 26-29, 30, 35 Set 3: 40, 45, 50, 53-55, 65 Set 4: 75, 77, 79, 85, 95	
Step 8:	Perform CoMFA using the master region file created from Step 7.	Set 1: 5, 10, 15-17, 20 Set 2: 22-24, 26-29, 30, 35 Set 3: 40, 45, 50, 53-55, 65 Set 4: 75, 77, 79, 85, 95	

Figure 2. Modified version of q^2 -GRS that can incorporate different probe atoms.

actual value (Figure 3). It was found that the worst prediction (residual = -2.34) was made for compound **58**, whereas the activity of compound **59** which is the salt form of compound **58** was predicted much better

(residual = -0.06). Since the carboxylic acid moiety of compound **58** is most likely to be deprotonated in solution, it was encouraging to find out that the model could distinguish between the protonated and deprotonated

Table 4. q^2 and Standard Error of Prediction (Numbers in Parentheses) Values Obtained after Performing CoMFA/ q^2 -GRS with Different q^2 Cutoff Values^a

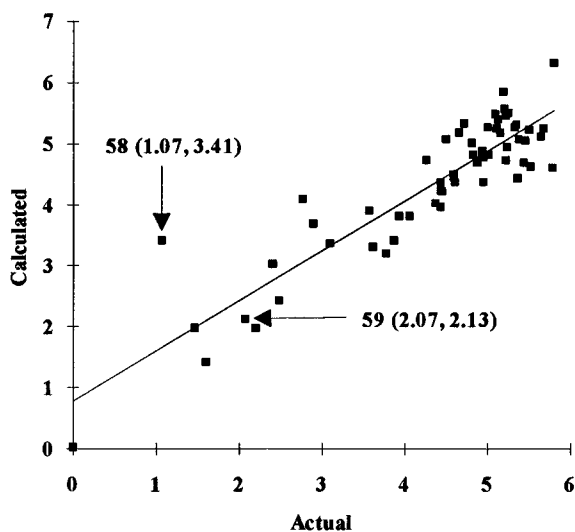
alignment	q^2 cutoff	no. of lattice points	no. of components					
			3	4	5	6	7	8
1	none ^b	1859	0.277 (1.116)	0.370 (1.051)	0.250 (1.157)	0.254 (1.165)	0.163 (1.246)	0.016 (1.364)
	0.1	1980	0.364 (1.047)	0.448 (0.984)	0.304 (1.115)	0.322 (1.110)	0.251 (1.178)	0.194 (1.234)
	0.2	1440	0.393 (1.023)	0.338 (1.077)	0.320 (1.102)	0.324 (1.109)	0.313 (1.129)	0.238 (1.200)
	0.3	180	0.336 (1.070)	0.387 (1.037)	0.393 (1.041)	0.355 (1.083)	0.324 (1.119)	0.261 (1.182)
2	none ^b	2184	0.279 (1.115)	0.338 (1.077)	0.311 (1.109)	0.203 (1.204)	0.013 (1.352)	-0.194 (1.502)
	0.1	3240	0.446 (0.977)	0.503 (0.933)	0.421 (1.017)	0.387 (1.056)	0.345 (1.101)	0.284 (1.163)
	0.2	1980	0.426 (0.995)	0.454 (0.979)	0.387 (1.047)	0.336 (1.099)	0.276 (1.158)	0.251 (1.189)
	0.3	720	0.355 (1.054)	0.442 (0.989)	0.431 (1.008)	0.358 (1.081)	0.298 (1.141)	0.229 (1.207)
2a	0.1	180	0.317 (1.085)	0.429 (1.001)	0.415 (1.022)	0.358 (1.080)	0.313 (1.128)	0.289 (1.159)
	0.1	3240	0.518 (0.866)	0.576 (0.819)	0.530 (0.871)	0.491 (0.915)	0.485 (0.930)	0.450 (0.970)
3	none ^b	2184	0.384 (0.979)	0.452 (0.932)	0.465 (0.929)	0.455 (0.947)	0.444 (0.966)	0.378 (1.031)
	0.1	2700	0.462 (0.915)	0.518 (0.874)	0.523 (0.878)	0.482 (0.923)	0.501 (0.914)	0.467 (0.955)
	0.2	1800	0.471 (0.908)	0.496 (0.894)	0.496 (0.902)	0.504 (0.903)	0.469 (0.944)	0.450 (0.969)
	0.3	900	0.459 (0.918)	0.500 (0.890)	0.509 (0.890)	0.523 (0.885)	0.503 (0.913)	0.455 (0.965)
	0.4	540	0.393 (0.972)	0.416 (0.962)	0.470 (0.925)	0.510 (0.898)	0.522 (0.895)	0.548 (0.879)

^a The numbers in bold represent the q^2 values for the optimal number of components and the lowest standard error of prediction values. ^b The results of conventional CoMFA.

Table 5. Summary of CoMFA/ q^2 -GRS Results

	alignment 2 ^{a,b}	alignment 2a ^{a,b}	alignment 2f ^a
optimal number of components	4	4	5
probe atom	C (sp ³ , +1)	C (sp ³ , +1)	combination ^c
cross-validated R^2	0.503	0.576	0.584
CVR ² -GRS cutoff	0.1	0.1	0.2
number of lattice points	3240	3240	2340
standard error of estimate	0.562	0.450	0.402
R^2	0.820	0.872	0.845
F values	63.805 ^d	93.736 ^e	103.870 ^f
prob. of $R^2 = 0$	0.000 ^d	0.000 ^f	0.000 ^f
contributions			
steric	0.499	0.516	0.528
electrostatic	0.501	0.484	0.472
"predictive" R^2		0.237 ^g	

^a See Table 3 for alignment rules. ^b See Table 4. ^c Combination of C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1) probes. ^d n1 = 4, n2 = 56. ^e n1 = 4, n2 = 55. ^f n1 = 5, n2 = 95. ^g "Predictive" $R^2 = (SD - \text{"press"})/SD$, where SD is the sum of the squared deviations between the affinities of molecules in the test set and the mean affinity of the training set molecules and "press" is the sum of the squared deviations between predicted and actual affinity values for every molecule in the test set.²⁹

**Figure 3.** Actual vs calculated $\ln(\text{PCPDCF} + 1)$ using alignment 2 (see Table 6).

nated form. The PLS analysis was performed again after excluding compound **58** from the model (Table 4: alignment 2a). The highest q^2 and lowest SDEP values of 0.576 and 0.819, respectively, occurred at 4 PCs.

For alignment 3, the lowest SDEP of 0.874 occurred at 0.1 q^2 cutoff and 4 PCs. The highest q^2 value, however, occurred at 0.4 q^2 cutoff and 8 PCs.

Among the three different alignments tried, align-

ment 2a (Table 3) yielded the highest q^2 and lowest SDEP values with conventional r^2 of 0.872 (Tables 4 and 5). The actual, calculated, and residual activities of etoposide, **1–57**, **59**, and **60** (alignment 2a, Table 5), are shown in Table 6. The plot of actual vs calculated activities of alignment 2a is shown in Figure 4.

CoMFA/ q^2 -GRS of 4'-O-Demethylepipodophyllotoxin Analogs Included in the Test Set. The PLS model obtained for alignment 2a was used to predict the activities of 41 compounds included in the test set (Table 2). The predicted activities for compounds **61–101** are shown in Table 6, and the plot of actual vs predicted activities is shown in Figure 5. The average absolute error of 0.42 (Table 6) and the predictive r^2 of 0.237 (Table 5, alignment 2a) were obtained. Due to the poor predictability of this model, we decided to combine both test and training sets excluding compound **58** (cf. Table 4, alignment 2a). The results of CoMFA/ q^2 -GRS with different q^2 cutoffs and probe atoms for 101 compounds are shown in Table 7. Initial conventional CoMFA (sp³ carbon with +1 charge) gave a q^2 of 0.404 and SDEP of 0.790 at 5 PCs (Table 7, alignment 2b). The subsequent application of the q^2 -GRS routine raised q^2 (0.581) and lowered SDEP (0.655). In order to explore the effect of different probe atoms on the predictability of the PLS model, probe atoms other than the sp³ carbon with a +1 charge were examined. However, employing different probe atoms generally did not significantly

Table 6. CoMFA Actual, Calculated, and Predicted Activities^a for Training and Test Set Molecules

compound	actual	alignment 2a ^{b,c}				alignment 2f ^{b,c}		compound	actual	alignment 2a ^{b,c}				alignment 2f ^{b,c}	
		calcd	residual	predicted	residual	calcd	residual			calcd	residual	predicted	residual	calcd	residual
etoposide	4.61	4.47	0.15			4.60	0.01	52	5.32	5.30	0.02			5.24	0.08
1	3.77	3.21	0.55			3.37	0.40	53	5.22	4.78	0.44			5.00	0.21
2	1.46	2.07	-0.61			2.13	-0.67	54	5.23	4.77	0.46			4.72	0.51
3	3.62	3.28	0.34			3.27	0.35	55	5.19	5.75	-0.56			5.57	-0.37
4	2.20	2.01	0.19			1.96	0.24	56	2.89	3.73	-0.84			3.78	-0.89
5	4.81	4.90	-0.09			4.91	-0.10	57	4.93	4.73	0.20			4.67	0.26
6	0.00	0.29	-0.29			0.38	-0.38	59	2.07	2.45	-0.39			2.31	-0.24
7	4.26	4.70	-0.44			4.55	-0.29	60	4.43	4.49	-0.06			4.58	-0.15
8	4.72	5.26	-0.55			5.22	-0.51	61	5.02			5.20	-0.18	5.04	-0.02
9	4.44	4.25	0.19			4.15	0.29	62	5.36			5.38	-0.02	5.26	0.10
10	5.12	5.39	-0.27			5.40	-0.27	63	4.80			5.02	-0.22	4.85	-0.04
11	5.09	5.36	-0.27			5.26	-0.17	64	5.07			5.07	0.00	4.97	0.10
12	4.50	5.03	-0.53			4.97	-0.47	65	4.75			5.10	-0.35	4.98	-0.22
13	5.67	5.18	0.49			5.04	0.64	66	3.50			3.70	-0.20	3.52	-0.02
14	5.37	5.16	0.21			5.15	0.22	67	3.95			5.08	-1.13	4.85	-0.90
15	5.50	5.18	0.32			5.11	0.38	68	4.61			5.29	-0.68	5.05	-0.44
16	4.93	4.95	-0.02			4.70	0.23	69	4.14			4.82	-0.68	4.72	-0.58
17	5.36	4.94	0.41			4.84	0.51	70	5.19			5.22	-0.03	5.14	0.05
18	1.61	1.27	0.34			1.17	0.43	71	4.17			5.35	-1.18	5.16	-0.99
19	5.52	4.67	0.86			4.56	0.96	72	5.20			4.81	0.39	5.28	-0.08
20	5.34	5.26	0.08			5.17	0.16	73	5.38			4.45	0.93	5.12	0.26
21	4.43	4.27	0.16			4.30	0.13	74	4.88			4.53	0.35	5.13	-0.25
22	4.87	4.78	0.09			4.69	0.18	75	4.98			4.77	0.21	5.00	-0.02
23	3.93	3.96	-0.03			4.05	-0.12	76	4.84			4.67	0.17	5.18	-0.33
24	4.65	5.11	-0.46			5.00	-0.34	77	5.38			4.57	0.81	5.24	0.14
25	5.02	4.87	0.15			4.77	0.25	78	5.14			4.51	0.63	5.24	-0.10
26	5.46	4.97	0.50			5.01	0.45	79	5.42			4.38	1.04	5.04	0.38
27	5.20	5.41	-0.21			5.14	0.06	80	5.65			4.54	1.12	5.27	0.38
28	3.87	3.41	0.46			3.43	0.44	81	4.61			4.39	0.22	4.95	-0.34
29	5.11	5.10	0.01			5.01	0.09	82	5.07			4.60	0.47	5.25	-0.18
30	5.64	4.99	0.64			4.95	0.69	83	4.98			3.79	1.19	4.58	0.40
31	4.59	4.67	-0.09			4.69	-0.10	84	5.26			4.63	0.63	5.30	-0.04
32	4.95	5.06	-0.12			5.03	-0.08	85	5.22			4.65	0.57	5.24	-0.02
33	5.44	4.73	0.71			4.64	0.80	86	5.18			4.77	0.41	4.99	0.19
34	5.78	4.99	0.79			4.90	0.88	87	5.08			4.81	0.27	5.03	0.05
35	2.77	4.26	-1.49			4.26	-1.49	88	4.86			4.74	0.12	4.82	0.04
36	3.09	3.60	-0.51			3.69	-0.60	89	4.76			4.79	-0.03	4.91	-0.15
37	4.59	4.50	0.09			4.32	0.26	90	4.77			4.67	0.10	4.89	-0.12
38	5.00	5.19	-0.18			5.11	-0.11	91	4.93			4.73	0.20	4.91	0.02
39	4.82	4.99	-0.17			4.97	-0.15	92	4.83			4.61	0.22	4.80	0.03
40	4.95	5.03	-0.09			4.80	0.15	93	5.01			4.66	0.35	4.84	0.17
41	5.80	6.46	-0.66			6.38	-0.58	94	5.07			4.45	0.62	4.68	0.39
42	2.48	2.36	0.12			2.13	0.36	95	4.47			4.72	-0.25	4.79	-0.32
43	4.06	3.67	0.39			3.74	0.32	96	5.08			4.66	0.42	4.79	0.29
44	3.56	3.65	-0.10			3.74	-0.19	97	5.01			4.73	0.28	4.87	0.15
45	2.40	2.99	-0.59			3.08	-0.69	98	4.80			4.68	0.12	4.76	0.04
46	5.25	5.51	-0.26			5.42	-0.17	99	4.55			4.69	-0.14	4.68	-0.13
47	5.22	5.46	-0.24			5.40	-0.19	100	4.61			4.64	-0.03	4.58	0.04
48	4.43	4.20	0.23			4.25	0.18	101	4.55			4.72	-0.17	4.62	-0.07
49	5.15	5.09	0.07			5.09	0.06								
50	4.36	4.32	0.04			4.27	0.09								
51	4.95	4.50	0.45			4.74	0.21								
												sum	17.14		
											average absolute error		0.42		

^a Activities are expressed as $\ln(\text{PCPDCF} + 1)$, where PCPDCF is the Percentage of Cellular Protein-DNA Complex Formed. ^b See Table 3 for alignment rules. ^c See Table 5 for the summary of CoMFA/q²-GRS results.

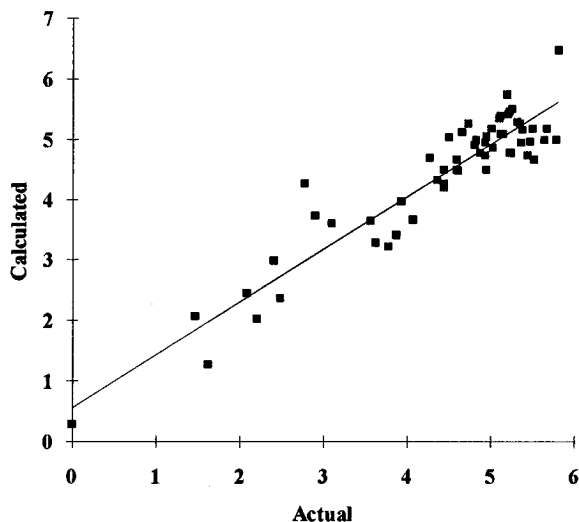


Figure 4. Actual vs calculated $\ln(\text{PCPDCF} + 1)$ using alignment 2a (see Table 6).

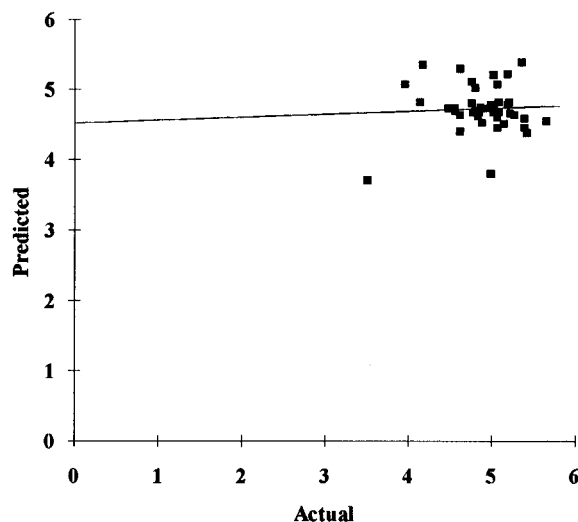


Figure 5. Actual vs predicted $\ln(\text{PCPDCF} + 1)$ using alignment 2a (see Table 6).

Table 7. q^2 and Standard Error of Prediction (Numbers in Parentheses) Values Obtained After Performing CoMFA/ q^2 -GRS with Different q^2 Cutoff Values and Probe Atoms for 101 Compounds (Alignments 2b–f)^a

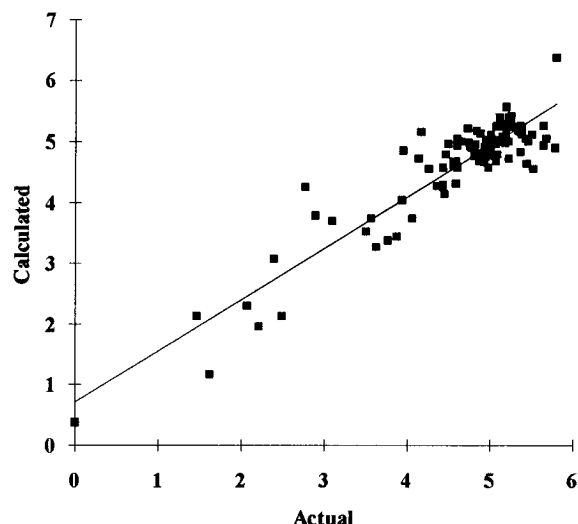
alignment	probe atom	charge	q^2 cutoff	no. of lattice points	no. of components		
					3	4	5
2b ^b	C (sp ³)	+1	none ^c	2184	0.272 (0.864)	0.356 (0.816)	0.404 (0.790)
			0.1	2520	0.424 (0.768)	0.538 (0.692)	0.568 (0.672)
			0.2	2160	0.427 (0.767)	0.543 (0.688)	0.574 (0.668)
			0.3	1620	0.435 (0.761)	0.533 (0.696)	0.542 (0.692)
			0.4	900	0.581 (0.655)	0.492 (0.725)	0.525 (0.705)
2c ^c	C (sp ³)	0	0.1	2160	0.515 (0.705)	0.535 (0.694)	0.511 (0.716)
			0.2	1980	0.532 (0.693)	0.530 (0.697)	0.436 (0.768)
			0.3	1980	0.532 (0.693)	0.530 (0.697)	0.436 (0.768)
			0.4	720	0.504 (0.713)	0.516 (0.708)	0.486 (0.733)
2d ^b	H	+1	0.1	2880	0.415 (0.774)	0.535 (0.694)	0.569 (0.671)
			0.2	1800	0.432 (0.763)	0.534 (0.695)	0.541 (0.693)
			0.3	1080	0.385 (0.794)	0.500 (0.719)	0.526 (0.705)
			0.4	180	0.438 (0.759)	0.481 (0.733)	0.464 (0.749)
2e ^b	O (sp ³)	-1	0.1	3240	0.407 (0.780)	0.526 (0.701)	0.564 (0.675)
			0.2	2160	0.418 (0.772)	0.522 (0.703)	0.530 (0.701)
			0.3	1440	0.429 (0.765)	0.536 (0.693)	0.548 (0.687)
			0.4	720	0.464 (0.741)	0.526 (0.700)	0.528 (0.703)
2f ^b	combination of C (sp ³ , +1), C (sp ³ , 0), H (+1), and O (sp ³ , -1)		0.1	3420	0.412 (0.776)	0.538 (0.692)	0.578 (0.665)
			0.2 ^d	2340	0.427 (0.766)	0.549 (0.684)	0.584 (0.660)
			0.3	1980	0.440 (0.758)	0.541 (0.689)	0.552 (0.685)
			0.4	1080	0.555 (0.675)	0.520 (0.705)	0.517 (0.711)

^a The numbers in bold represent the q^2 values for the optimal number of components. ^b See Table 3 for alignment rules. ^c The results of conventional CoMFA. ^d This region file contains 13 boxes (each with 180 probe atoms). Three boxes contain O (sp³, -1) probe atoms, five boxes contain C (sp³, 0), one box contains H (+1), and four boxes contain C (sp³, +1).

improve the predictability of the PLS model. Alignments 2c, 2d, and 2e yielded the highest q^2 values of 0.535 (SDEP, 0.694), 0.569 (SDEP, 0.671), and 0.564 (SDEP, 0.675), respectively (Table 7).

Since the receptor atoms interacting with the molecules in the real active site environment are of various chemical types, we decided to incorporate in the analysis simultaneously four types of probe atoms as described in computational methods (Table 3, alignment 2f). Out of four different cutoffs used to generate the PLS model, a 0.2 q^2 cutoff gave the highest q^2 and lowest SDEP of 0.584 and 0.660, respectively (Table 7), which were better than any other alignments examined before (cf. Tables 4 and 7). The resulting region file contained 13 boxes with 2340 lattice points, including three boxes with O (sp³, -1) probes, five boxes with C (sp³, 0) probes, one box with H (+1) probes, and four boxes with C (sp³, +1) probes. Non-cross-validated CoMFA results for alignment 2f are shown in Table 5 ($R^2 = 0.845$, SDEE = 0.402, $F = 103.870$). The actual, calculated, and residual activities for alignment 2f are shown in Table 6. The plot of actual vs calculated activities of alignment 2f is shown in Figure 6.

CoMFA Fields. The CoMFA steric and electrostatic fields obtained using a combination of C (sp³, +1), C (sp³, 0), H (+1), and O (sp³, -1) probe atoms (alignment 2f) and the structure of etoposide, **6**, and **41** are shown in Figures 7–9. The field values were calculated by multiplying the β -coefficient and standard deviation of columns in the QSAR table (stdev*coeff). The green (sterically favorable) and yellow (sterically unfavorable) contours shown in Figure 7 represent 80% and 20% level contributions, respectively. The simulated etoposide-DNA intercalated structure (Figures 7) is constructed as described in the computational methods section. This simulated etoposide-DNA complex is shown with the CoMFA steric field in order to compare the location of CoMFA fields with respect to the DNA structure. The green contour region faces away from the DNA (away from the direction of intercalation) whereas the yellow

**Figure 6.** Actual vs calculated $\ln(\text{PCPDCF} + 1)$ using alignment 2f (see Table 6).

contour region is localized near or over the backbone of DNA (Figure 7). The R group of etoposide nicely extends into the green region, whereas the R group of the inactive compound (**6**) extends into the yellow contour region. The inactivity of this compound is probably due to an unfavorable interaction between DNA backbone and the R group.

The contours generated using the electrostatic field display a much more complicated picture. Because both positive and negative charge probe atoms are used, the interpretation of electrostatic contour plots requires the knowledge of the identity of probe atoms used in CoMFA subregions. (The structure of DNA is not shown in order to better display CoMFA electrostatic field.) Three rectangular boxes shown in Figure 8 represent negatively charged probe atom regions [O (sp³, -1)], and the red (20% contribution level) and blue (80% contribution level) contours correspond to positive and negative charge favored regions, respectively. Five rectangular

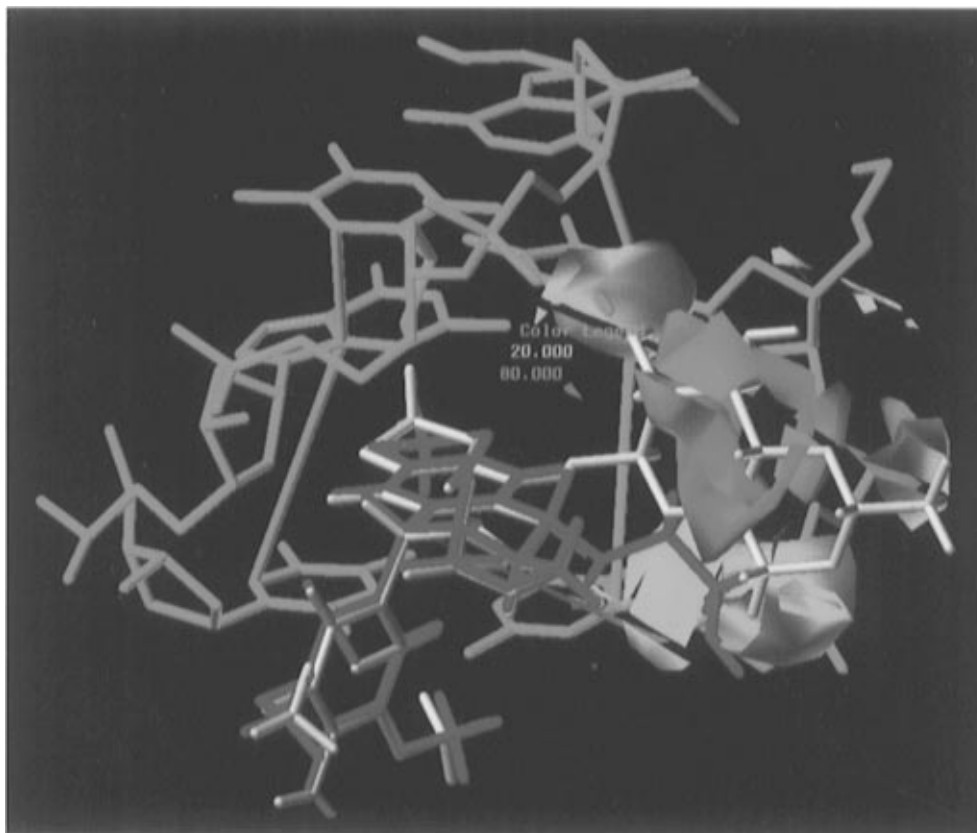


Figure 7. The CoMFA steric stdev*coeff contour plot of alignment 2f. Green regions represent a contribution level of 80%, sterically favored areas. Yellow regions represent a contribution level of 20%, sterically disfavored areas. Etoposide and compound **6** are depicted as white and purple sticks, respectively.

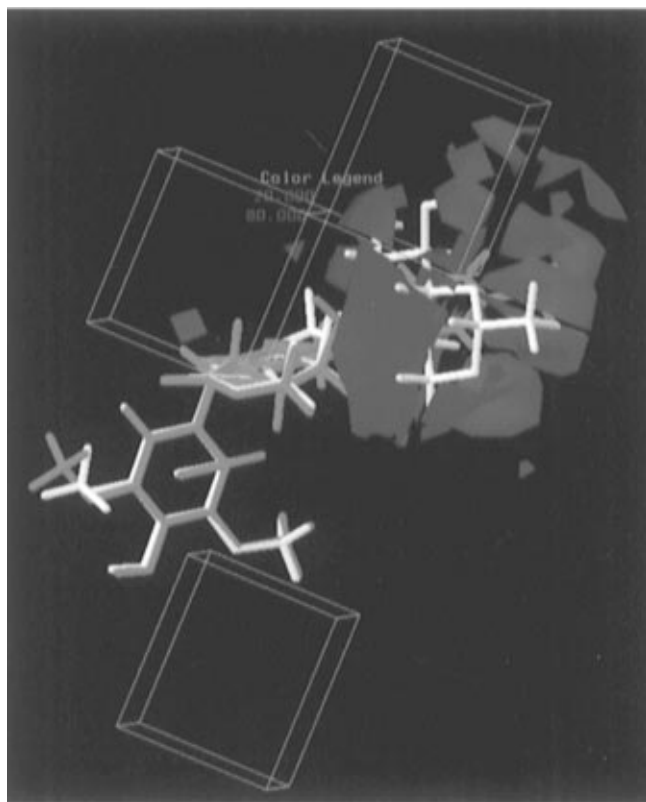


Figure 8. The CoMFA electrostatic stdev*coeff contour plot of alignment 2f. Three rectangular boxes represent the regions with negatively charged probe atoms (O.3). Blue regions within these boxes represent a contribution level of 80%, negative charge favored areas. Red regions within these boxes represent a contribution level of 20%, positive charge favored areas. Etoposide and compound **41** are depicted as white and green sticks, respectively.

boxes shown in Figure 9, however, represent positively charged probe atom regions [C (sp^3 , +1) and H (+1)], and the red (20% contribution level) and blue (80% contribution level) contours correspond to negative and positive charge favored regions, respectively. The rest of lattice boxes [five boxes containing C (sp^3 , 0) probes] exhibited no electrostatic contours because the probe atom does not have a charge. The positive charge favored region seems to face away from the direction of intercalation, to lie along the minor groove, and to be separated by the negative charge favorable region that seems to be near and over the backbone of DNA.

Discussion

The CoMFA approach has rapidly become one of the most powerful tools for 3-D QSAR studies. Studies done by us²¹ and others^{20,29,30} revealed that CoMFA results can be extremely sensitive to a number of factors such as alignment rules, overall orientation, lattice placement, step size, and probe atom type. A recently described CoMFA/ q^2 -GRS approach addressed the problems related to overall orientation, lattice placement, and step size.²¹ In addition, we found that a simple modification of the q^2 -GRS process (Figure 2) can easily incorporate different types of probe atoms and charges into the CoMFA model, which would be difficult with the conventional CoMFA implementation. Since the CoMFA approach, initially, stems from the analysis of the physicochemical factors involved in the drug-receptor interaction, describing steric and electrostatic molecular fields with multiple probe types seems logical.

The CoMFA/ q^2 -GRS²¹ and modified q^2 -GRS (Figure 2) have been used initially to obtain 3D-QSAR of 60 analogs of 4'-*O*-demethylepipodophyllotoxin. Various

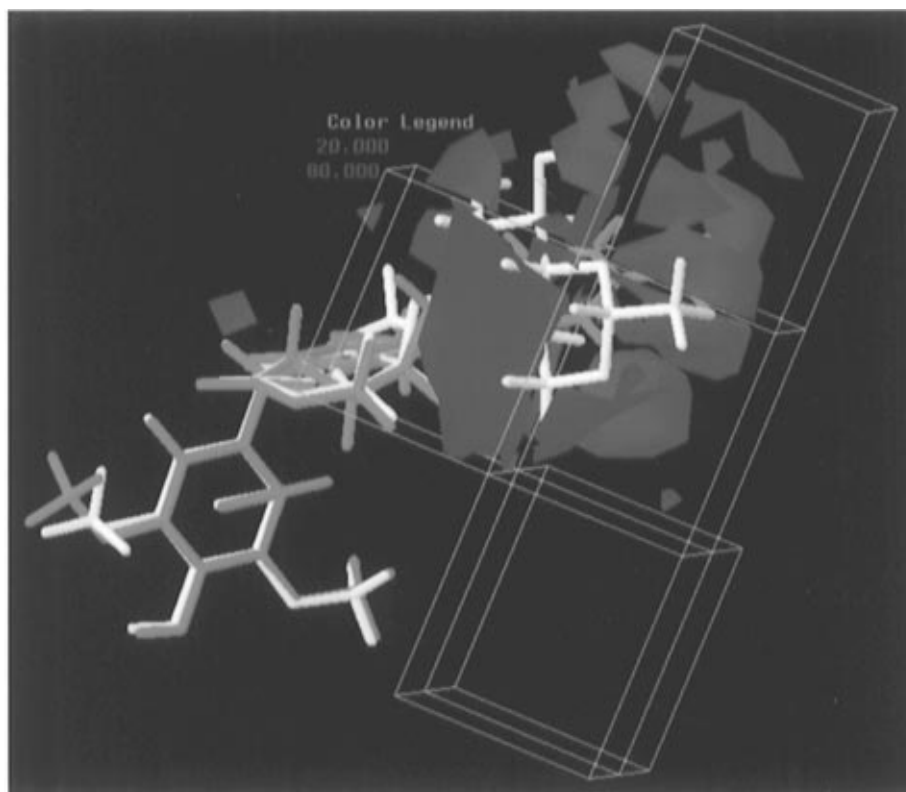


Figure 9. The CoMFA electrostatic $\text{stdev} \times \text{coeff}$ contour plot of alignment 2f. Five rectangular boxes represent the regions with positively charged probe atoms (C.3 and H). Blue regions within these boxes represent a contribution level of 80%, positive charge favored areas. Red regions within these boxes represent a contribution level of 20%, negative charge favored areas. Etoposide and compound **41** are depicted as white and green sticks, respectively.

alignments were examined (Table 3). Alignments 1, 2, 2a, and 3 were initially used to derive PLS models. Among these, alignment 2a yielded the highest q^2 (0.576) and lowest SDEP (0.819) values (Table 4). However, the predictive R^2 of this alignment was found to be low, suggesting the inadequacy of our training set (Table 5). This result reflects the common problem of QSAR methods which are generally good at interpolating the data but have moderate success in extrapolating the data. Therefore, we have decided to combine both training and test sets (Tables 1 and 2) in order to develop a single 3D-QSAR model. Four different probe atoms were used to derive the final PLS models (Table 7; alignments 2b–e). In addition, CoMFA was performed with all four probe atoms using the modified q^2 -GRS routine (Table 7, alignment 2f). Among these, alignment 2f yielded the highest q^2 (0.584) and lowest SDEP (0.660). This alignment was selected to construct the steric and electrostatic contours (Figures 7–9).

Our examination of the contour plots (Figures 7–9) suggests that they agree well with the composite pharmacophore model (Figure 10) proposed by MacDonald et al.²⁵ for DNA topoisomerase II inhibitors. According to this model, topoisomerase II activity, which is exhibited by various inhibitors such as daunorubicin, amsacrine, and etoposide, is due to three structurally distinct domains: DNA intercalating moiety, the minor groove binding site, and the molecular region that can accommodate a number of structurally diverse substituents, which might also bind to the minor groove (Figure 10). One interesting aspect of this composite pharmacophore model is that, in order to be active, compounds do not have to interact with all three domains.²⁵ Among the three binding domains, the minor groove binding domain (e.g., the E ring of etoposide) is the only region

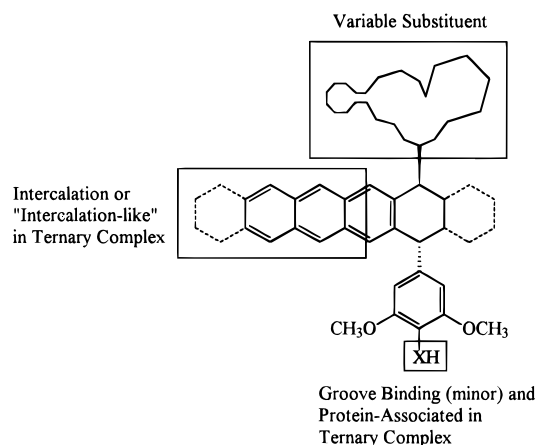


Figure 10. A composite pharmacophore model for the expression of topoisomerase activity.

that has been studied extensively.^{5,18,25} Our CoMFA study of 101 analogs of 4'-*O*-demethylepipodophyllotoxin is an important step toward understanding the less explored variable substituent domain of this composite pharmacophore model. We are currently investigating the importance of the third domain, the intercalating region (Figure 10). We believe the 4'-*O*-demethylepipodophyllotoxin analogs are excellent model inhibitors of topoisomerase II, and a large number of analogs synthesized and tested by our group provides an ideal condition for studying the nature of interactions among the variable substituent domain, DNA, and topoisomerase II.

Examination of the steric and electrostatic contour plots of alignment 2f reveals a number of important characteristics of the active compounds in Tables 1 and 2 which agree with the hypothesis that the variable

substituent domain binds to the DNA minor groove.^{25,27} The sterically unfavorable (yellow) contours (Figure 7) surround the DNA backbone, indicating that any R group that extends into this region should cause a bad steric interaction with the DNA backbone as shown for the R group of compound **6**. Such unfavorable interaction is detrimental to the cellular protein linked DNA complex formation. In contrast, compounds that extend into sterically favorable contours (Figure 7) are devoid of any bad steric interaction with the DNA backbone. In addition, electrostatic contour plots show that active compounds should have positive charged functional groups near the minor groove of DNA.

On the basis of these results, we propose that, in order for the different functional groups of compounds in Tables 1 and 2 to interact with DNA minor groove, these functional groups have to be sterically and electrostatically compatible with the environment of the DNA minor groove. This means that such compounds must lack any unnecessary steric interaction with the DNA backbone and require a positive charge around the R group to complement the negative rich environment of the DNA backbone, as in the case of minor groove binding compounds.^{31,32}

It is still not clear how these inhibitors affect the DNA topology. However, the fact that the changes in the DNA topology have been shown to influence the rates of initiation and elongation for both replication and transcription suggests that the formation of the covalent topoisomerase II–DNA cleavage complex might also be influenced by the DNA topology changes.³³ According to MacDonald et al.,²⁵ a drug-induced DNA topology change might cause DNA to assume the form of one of the intermediates in the catalytic cycle of topoisomerase II and act as a “transition state analog” of DNA. Since enzymes bind more strongly to transition states than ground states, a drug-bound DNA should interact stronger with topoisomerase II, eventually forming the stable covalent drug–DNA–topoisomerase II cleavage complex.

In summary, we have investigated the variable substituent domain of the proposed composite pharmacophore using 3-D QSAR of 101 analogs of 4'-O-demethylepipodophyllotoxin via the CoMFA/q²-GRS approach. We have found that the CoMFA results agree well with the pharmacophore model proposed earlier.²⁵ Currently, the CoMFA model obtained from this study is being used in the design of novel 4'-O-demethylepipodophyllotoxin analogs.

Acknowledgment. This investigation was supported by a grant from the American Cancer Society (No. DHP-13H) awarded to K. H. Lee. A. Tropsha acknowledges the software grant from Tripos Inc.

References

- (1) Kitanaka, S.; Yasuda, I.; Kashiwada, Y.; Hu, C. Q.; Bastow, K. F.; Bori, I. D.; Lee, K. H. Antitumor Agents 162. Cell-Based Assays for Identifying Novel DNA Topoisomerase Inhibitors: Studies on Constituents of *Fatsia japonica*. *J. Nat. Prod.* **1995**, *58*, 1647–1654.
- (2) Podwyssotzki, V. *Arch. Exp. Pathol. Pharmacol.* **1980**, *13*, 29.
- (3) Kelly, M. G.; Hartwell, J. L. The Biological Effects and the Chemical Composition of Podophyllin, a Review. *J. Natl. Cancer Inst.* **1954**, *14*, 967–1010.
- (4) Jardin, I. Podophyllotoxins. In *Anticancer Agents Based on Natural Product Models*; Cassady, J. M., Douros, J. D., Eds.; Academic Press: New York, 1980; pp 319–351.
- (5) Sackett, D. L. Podophyllotoxin, Steganacin, and Combretastatin: Natural Products That Bind at the Colchicine Site of Tubulin. *Pharm. Ther.* **1993**, *59*, 163–228.
- (6) Van Maanen, J. M. S.; Retel, J.; De Vries, J.; Pinedo, H. M. Mechanism of Action of Antitumor Drug Etoposide: A Review. *J. Natl. Cancer Inst.* **1988**, *80*, 1526–1533.
- (7) Hainsworth, J. D.; Williams, S. D.; Einhorn, L. H.; Birch, R.; Greco, F. A. Successful Treatment of Resistant Germinal Neoplasms with VP-16 and Cisplatin: Results of a Southeastern Cancer Study Group Trial. *J. Clin. Oncol.* **1985**, *3*, 666–671.
- (8) Shah, J. C.; Chen, J. R.; Chow, D. Preformulation Study of Etoposide: Identification of Physicochemical Characteristics Responsible for the Low and Erratic Oral Bioavailability of Etoposide. *Pharm. Res.* **1989**, *6*, 408–412.
- (9) Zhang, Y.; Lee, K. H. Recent Progress in the Development of Novel Antitumor Etoposide Analogs. *Chin. Pharm. J.* **1994**, *46*, 319–369.
- (10) Lee, K. H.; Imakura, Y.; Haruna, M.; Beers, S. A.; Thurston, L. S.; Dai, H. J.; Chen, C. H. Antitumor Agents, 107. New Cytotoxic 4-Alkylamino Analogues of 4'-Demethylepipodophyllotoxin as Inhibitors of Human DNA Topoisomerase II. *J. Nat. Prod.* **1989**, *52*, 606–613.
- (11) Wang, Z. Q.; Kuo, Y. H.; Schnur, D.; Bowen, J. P.; Liu, S. Y.; Han, F. S.; Chang, J. Y.; Cheng, Y. C.; Lee, K. H. Antitumor Agents. 113. New 4 β -Arylamino Derivatives of 4'-O-Demethylepipodophyllotoxin and Related Compounds as Potent Inhibitors of Human DNA Topoisomerase II. *J. Med. Chem.* **1990**, *33*, 2660–2666.
- (12) Zhang, Y. L.; Guo, X.; Cheng, Y. C.; Lee, K. H. Antitumor Agents. 148. Synthesis and Biological Evaluation of Novel 4 β -Amino Derivatives of Etoposide with Better Pharmacological Profiles. *J. Med. Chem.* **1994**, *37*, 446–452.
- (13) Lee, K. H.; Beers, S. A.; Mori, M.; Wang, Z. Q.; Kuo, Y. H.; Li, L.; Liu, S. Y.; Chang, J. Y.; Han, F. S.; Cheng, Y. C. Antitumor Agents. 111. New 4-Hydroxylated and 4-Halogenated Anilino Derivatives of 4'-Demethylepipodophyllotoxin as Potent Inhibitors of Human DNA Topoisomerase II. *J. Med. Chem.* **1990**, *33*, 1364–1368.
- (14) Zhou, X. M.; Wang, Z. Q.; Chang, J. Y.; Chen, H. X.; Cheng, Y. C.; Lee, K. H. Antitumor Agents. 120. New 4-Substituted Benzylamine and Benzyl Ether Derivatives of 4'-O-Demethylepipodophyllotoxin as Potent Inhibitors of Human DNA Topoisomerase II. *J. Med. Chem.* **1991**, *34*, 3346–3350.
- (15) Zhou, X. M.; Wang, Z. Q.; Chen, H. X.; Cheng, Y. C.; Lee, K. H. Antitumor Agents. 125. New 4 β -Benzoylamino Derivatives of 4'-O-Demethyl-4-desoxypodophyllotoxin and 4 β -Benzoyl Derivatives of 4'-O-Demethylpodophyllotoxin as Potent Inhibitors of Human DNA Topoisomerase II. *Pharm. Res.* **1993**, *10*, 214–219.
- (16) Osheroff, N.; Zechiedrich, E. L.; Gale, K. C. Catalytic Function of DNA Topoisomerase II. *BioEssays* **1991**, *13*, 269–275.
- (17) Alton, P. A.; Harris, A. L. Annotation. *Br. J. Haematol.* **1993**, *85*, 241–245.
- (18) Liu, S. Y.; Hwang, B. D.; Haruna, M.; Imakura, Y.; Lee, K. H.; Cheng, Y. C. Podophyllotoxin Analogs: Effects on DNA Topoisomerase II, Tubulin Polymerization, Human Tumor KB Cells, and Their VP-16-Resistant Variants. *Mol. Pharmacol.* **1989**, *36*, 78–82.
- (19) Zhang, Y. L.; Tropsha, A.; McPhail, A. T.; Lee, K. H. Antitumor Agents. 152. *In Vitro* Inhibitory Activity of Etoposide Derivative NPF A Against Human Tumor Cell Lines and a Study of Its Conformation by X-ray Crystallography, Molecular Modeling, and NMR Spectroscopy. *J. Med. Chem.* **1994**, *37*, 1460–1464.
- (20) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- (21) Cho, S. J.; Tropsha, A. Cross-Validated R² Guided Region Selection for Comparative Molecular Field Analysis (CoMFA): A Simple Method to Achieve Consistent Results. *J. Med. Chem.* **1995**, *38*, 1060–1066.
- (22) The Sybyl program (version 6.0) is available from Tripos Associates, 1699 South Hanley Road, St. Louis, MO 63144.
- (23) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
- (24) DePriest, S. A.; Mayer, D.; Naylor, C. B.; Marshall, G. R. 3D-QSAR of Angiotensin-Converting Enzyme and Thermolysin Inhibitors: A Comparison of CoMFA Models Based on Deduced and Experimentally Determined Active-site Geometries. *J. Am. Chem. Soc.* **1993**, *115*, 5372–5384.
- (25) MacDonald, T. L.; Lehnert, E. K.; Loper, J. T.; Chow, K. C.; Ross, W. E. On the Mechanism of Interaction of DNA Topoisomerase II with Chemotherapeutic Agents. In *DNA Topoisomerase in Cancer*; Potmesil, M., Kohn, K. W., Eds.; Oxford University Press: New York, 1991; pp 119–214.
- (26) Chow, K. C.; MacDonald, T. L.; Ross, W. E. DNA Binding by Epipodophyllotoxins and N-Acyl Anthracyclines: Implications for Mechanism of Topoisomerase II Inhibition. *Mol. Pharmacol.* **1988**, *34*, 467–473.
- (27) Baguley, B. C. DNA Intercalating Anti-Tumor Agents. *Anti-Cancer Drug Des.* **1991**, *6*, 1–35.

- (28) Gao, Y. G.; Liaw, Y. C.; Robinson, H.; Wang, A. H. Binding of the Antitumor Drug, Nogalamycin and Its Derivatives to DNA: Structural comparison. *Biochemistry* **1990**, *29*, 10307–10316.
- (29) Waller, C. L.; Oprea, T. I.; Giolitti, A.; Marshall, G. R. Three-Dimensional QSAR of Human Immunodeficiency Virus (I) Protease Inhibitors. 1. A CoMFA Study Employing Experimentally-Determined Alignment Rules. *J. Med. Chem.* **1993**, *36*, 4152–4160.
- (30) Debnath, A. K.; Hansch, C.; Kim, K. H.; Martin, Y. C. Mechanistic Interpretation of the Genotoxicity of Nitrofurans (Antibacterial Agents) Using Quantitative Structure-Activity Relationships and Comparative Molecular Field Analysis. *J. Med. Chem.* **1993**, *36*, 1007–1016.
- (31) Fairley, T. A.; Tidwell, R. R.; Donkor, I.; Naiman, N. A.; Ohemeng, K. A.; Lombardy, R. J.; Bentley, J. A.; Cory, M. Structure, DNA Minor Groove Binding, and Base Pair Specificity of Alkyl- and Aryl-Linked Bis(amidinobenzimidazoles) and Bis-(amidinoindoles). *J. Med. Chem.* **1993**, *36*, 1746–1753.
- (32) Pullman, B.; Pullman, A. Structural Factors Involved in the Binding of Netropsin and Distamycin A to Nucleic Acids. *Stud. Biophys.* **1981**, *86*, 95–102.
- (33) Osheroff, N.; Zechiedrich, E. L.; Gale, K. C. Catalytic Function of DNA Topoisomerase II. *BioEssays* **1991**, *13*, 269–275.

JM9503052