Comparative Molecular Field Analysis of Anticoccidial Triazines

James W. McFarland

Central Research Division, Pfizer Inc., Groton, Connecticut 06340. *Received January 14, 1992*

Comparative molecular field analysis (CoMFA) of 2-(substituted phenyl)-1,2,4-triazine-3,5(2*H*,4*H*)-diones (triazines henceforth) resulted in an excellent correlation of their anticoccidial potencies with their physical properties. Two items about this work are notable: (i) the biological data are from a whole animal infectious disease model; and (ii) for the best results CoMFA required columns of measured "lipophilicity" and "acidity" data in addition to the calculated data in the steric field and electrostatic field columns. CoMFA resulted in a quantitative description of the major steric and electrostatic field effects, and gave significant new insights to factors governing potency. The model was used to "predict" the potencies of diverse triazines not used in making the model itself.

Comparative Molecular Field Analysis (CoMFA) as a means to determine three-dimensional quantitative structure–activity relationships (3D-QSAR) has been rapidly advancing since its introduction in 1988.[1] Until now its successes have dealt largely with enzyme or receptor binding data, and less with intact biological structures.[2-12] What is reported here is apparently the first case of a successful application of CoMFA to results from a whole animal infectious disease model.

The basis for the CoMFA model created here are data from earlier work out of these laboratories.[13] In that work it was shown that the lipophilicities and the acidities of 2-(substituted phenyl)-1,2,4-triazine-3,5(2*H*,4*H*)-diones are important determinants of their anticoccidial potencies. The relative lipophilicities of these triazines were determined by HPLC, the results being expressed as log k'. The relative acidities were estimated from the ^{1}H-NMR chemical shift at position 6 (δ_6) of the triazine ring. While the results of this earlier QSAR study were statistically significant and gave some insights as to factors governing potency, the correlation was not sufficiently strong to allow confident predictions of potency in new triazines (see eq 1). Obviously, there were unrecognized additional contributing factors. The present work reveals some strong possibilities as to what these might be.

$$\log (1/\mathrm{MEC}) = -1.20\ (\pm 0.51)(\log k')^2 + 2.55\ (\pm 0.52)(\log k') + 7.49\ (\pm 2.16)\delta_6 - 56.57 \quad (1)$$

$$n = 54 \qquad R^2 = 0.56 \qquad s = 0.71 \qquad F_{3,50} = 21.30$$

What will be demonstrated here is that, for this case, CoMFA by itself is no better than the earlier multiple regression analysis (MRA);[13] but the combination of CoMFA with log k' and δ_6 gives greatly improved results. Thus, as might be expected when dealing with in vivo biological data, more traditional physical data may profitably augment the steric and electrostatic field data generated by the CoMFA technique.

Anticoccidial Testing

Triazines were tested for anticoccidial activity in a manner described previously.[14] Briefly the method is as follows: chicks are placed on test-drug-treated feed, and then challenged with oocysts of the coccidian parasite *Eimeria tenella*; activity is measured by the degree of lesion control achieved in the ceca of the treated, challenged birds. The lowest drug level in feed preventing

(1) 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.

(2) Marshall, G. R.; Mayer, D.; Naylor, C. B.; Hodgkin, E. E.; Cramer, R. D. Mechanism-Based Analysis of Enzyme Inhibitors of Amide Bond Hydrolysis. In *QSAR: Quantitative Structure-Activity Relationships in Drug Design*, Fauchere, J. L., Ed.; Alan R. Liss, Inc.: New York, 1989; pp 287–295.

(3) Koehler, K. F.; Gasiecki, A. F.; Kramer, S. W.; Shone, R. L.; Collins, P. W.; Tsai, H.; Bianchi, R. Prostaglandin Receptor Models from an Excluded Volume and 3D QSAR Analysis of Omega Chain Enisoprost Analogs. In *Book of Abstracts, 200th American Chemical Society National Meeting*, Washington, DC, Aug 26–31, 1990; American Chemical Society: Washington, DC, 1990; MEDI 75.

(4) Nicklaus, M. C.; Bolen, J. B.; Li, Z-H.; Burke, T. R., Jr. Comparative Molecular Field Analysis (CoMFA) of Tyrosine Kinase Inhibitors. In *Book of Abstracts, 200th American Chemical Society National Meeting*, Washington, DC, Aug 26–31, 1990; American Chemical Society: Washington, DC, 1990; MEDI 139.

(5) Miller, A. B.; Bowen, J. P.; Borghoff, S. J.; Swenberg, J. A. Further Computational and Molecular Modeling Studies of α 2μ-Globulin. In *Book of Abstracts, 200th American Chemical Society National Meeting*, Washington, DC, Aug 26–31, 1990; American Chemical Society: Washington, DC, 1990; MEDI 163.

(6) Allen, M. S.; Yun-C, T.; Trudell, M. L.; Narayanan, K.; Schindler, L. R.; Martin, M. J.; Schultz, C.; Hagen, T. J.; Koehler, K. F.; Codding, P. W.; Skolnick, P.; Cook, J. M. Synthetic and Computer-Assisted Analyses of the Pharmacophore for the Benzodiazepine Receptor Inverse Agonist Site. *J. Med. Chem.* **1990**, *33*, 2343–2357.

(7) Maret, G.; El Tayar, N.; Carrupt, P-A.; Testa, B.; Jenner, P.; Baird, M. Toxication of MPTP (1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine) and Analogs by Monoamine Oxidase. *Biochem. Pharmacol.* **1990**, *40*, 783–792.

(8) Wiese, T. E.; Palomino, E.; Horwitz, J. P.; Brooks, S. C. Comparative Molecular Field Analysis (CoMFA) of the Specific Responses of MCF-7 Cells to A-Ring Substituted Estrogens. In *Book of Abstracts, 201st American Chemical Society National Meeting*, Atlanta, GA, April 14–19, 1991; American Chemical Society: Washington, DC, 1991; MEDI 150.

(9) Randad, R. S.; Abraham, D. J. Allosteric Modifiers of Hemoglobin 3. Design, Synthesis and Evaluation of Hemoglobin Oxygen Dissociating Agents. In *Book of Abstracts, 201st American Chemical Society National Meeting*, Atlanta, GA, April 14–19, 1991; American Chemical Society: Washington, DC, 1991; MEDI 56.

(10) Björkroth, J-P.; Pakkanen, T. A.; Lindroos, J. Comparative Molecular Field Analysis of Some Clodronic Acid Esters. *J. Med. Chem.* **1991**, *34*, 2338–2343.

(11) Carroll, F. I.; Gao, Y.; Rahman, M. A.; Abraham, P.; Parham, K.; Lewin, A. H.; Boja, J. W.; Kuhar, M. J. Synthesis, Ligand Binding, QSAR, and CoMFA Study of 3β-(*p*-Substituted phenyl)tropane-2β-carboxylic Acid Methyl Esters. *J. Med. Chem.* **1991**, *34*, 2719–2725.

(12) Diana, G. D.; Kowalczyk, P.; Treasurywala, A. M.; Olgesby R. C.; Pevear, D. C.; Dutko, F. J. CoMFA Analysis of the Interactions of Antipicornavirus Compounds in the Binding Pocket of Human Rhinovirus-14. *J. Med. Chem.* **1992**, *35*, 1002–1008.

(13) McFarland, J. W.; Cooper, C. B.; Newcomb, D. M. Linear Discriminant and Multiple Regression Analyses of Anticoccidial Triazines. *J. Med. Chem.* **1991**, *34*, 1908–1911.

(14) Chappel, L. R.; Howes, H. L.; Lynch, J. E. The Site of Action of a Broad-Spectrum Aryltriazine Anticoccidial, CP-25,415. *J. Parasitol.* **1974**, *60*, 415–420.

0022-2623/92/1835-2543$03.00/0

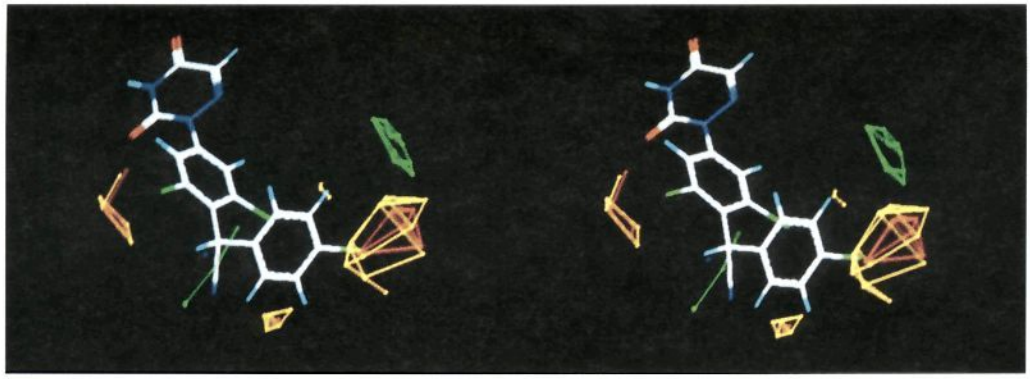


Figure 1. Diclazuril (1) embedded in the CoMFA steric field from model C. Orange polyhedra, planes, and lines indicate intersections in CoMFA steric field that contain the highest 5% of field values adding to anticoccidial potency; the corresponding yellow items indicate the next highest 5%. The blue dot (near the green line) represents the highest 5% of field values detracting from potency, while the green line together with the green polyhedron indicate the next highest 5%.

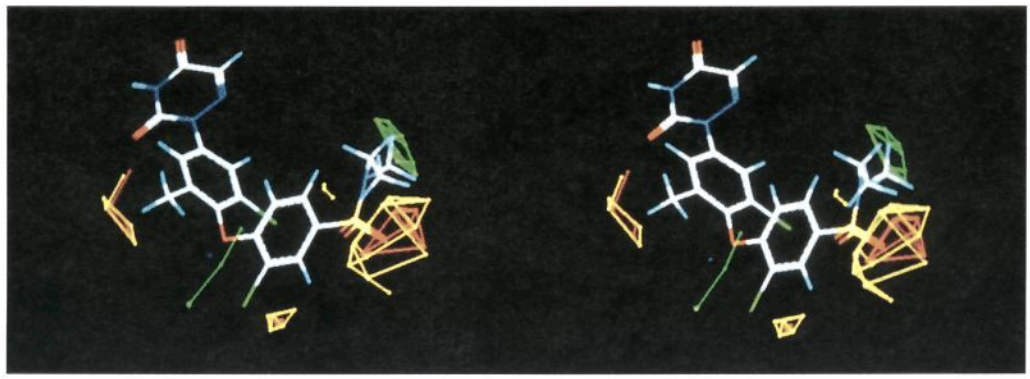


Figure 2. Compound 18 embedded in the CoMFA steric field from model C. Colors of dots, lines, planes, and polyhedra are interpreted as described in the legend to Figure 1.

lesions is reported as the minimum effective concentration (MEC). The results of this procedure are usually reproducible to within one 2-fold dilution, i.e. within ±0.3 in terms of log (1/MEC).

Data and Analytical Methods

CoMFA was performed by using the QSAR option of SYBYL version 5.41 (Tripos Associates, 1699 S. Hanley Road, Suite 303, St. Louis, MO 63144). Except where noted, default values were used. The CoMFA grid spacing was 2.0 Å in the x, y, and z directions. The grid region was generated automatically by the software, and was large enough to include all the molecules used to generate the model with an extension of 4.0 Å in each direction. The probe atom used to generate the interaction energies at each grid intersection had the van der Waals properties of sp^3 carbon with a charge of +1.0. At points where repulsive steric interaction energies exceeded 30 kcal/mol, these energies were truncated to just 30 kcal/mol. At these same locations, the electrostatic interaction energies were set to the mean value of the corresponding energies for the other molecules at the same point.

To avoid swamping the effects of data related to log k', $(\log k')^2$, and δ_6 by data from the much more numerous CoMFA items, the data in all columns were normalized, and columns not contributing significantly to the overall variance in the descriptors were removed as indicated below. For log k', $(\log k')^2$, and δ_6 normalization was accomplished in the ususal way: for an entry in each column, the mean of that column was subtracted and the result was divided by the column's standard deviation. For the "steric" CoMFA columns the mean for all entries in all columns was substracted from the value of each entry; then that result was divided by the standard deviation for all entries in all columns. This is justified by the fact that all "steric" columns contain data of the same type. The "electrostatic" CoMFA columns were treated analogously. These CoMFA columns individually will have different means and standard deviations. Those that have standard deviations too small to be important can be removed from the analysis by using the MINIMUM SIGMA option. Applying this technique to the present case, and setting MINIMUM SIGMA to 2.0 reduced the number of grid intersections for consideration from 2160 to the 135 that were richest in variance. The data from these 135 intersections, log k', $(\log k')^2$, and δ_6 were then analyzed by partial least squares (PLS). Because they are so many, the steric and electrostatic terms are not presented here, but they are expressed graphically as the polyhedra defining the regions of favorable and unfavorable interactions in Figures 1–3.

Model Building

The template molecular model was diclazuril (1) and included hydrogen atoms. It was created by the SYBYL option BUILD, and its conformation was minimized using the ENERGY MINIMIZE option. No attempt was made to establish whether or not this was the global minimum. Charges were then generated for each atom using the Gasteiger–Hückel method.[15] The asymmetric atom (that α to the cyano group) was arbitrarily oriented in the R configuration (i.e., R^*). A molecular model for each subsequent compound was created in a like manner except that the conformation of each new model was matched to

(15) Anonymous. *Sybyl Molecular Modeling Software Version 5.4, Theory Manual*; Tripos Associates, Inc.: St. Louis, 1991; p 2070.

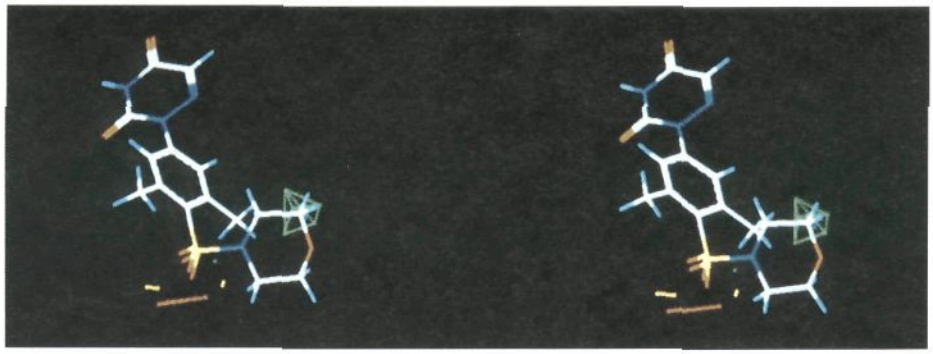


Figure 3. Compound **33** embedded in the CoMFA electrostatic field from model C. Colors of dots, lines, and polyhedra are interpreted as described in the legend to Figure 1.

that of diclazuril as closely as possible. This was accomplished through bond rotation, avoiding any high energy processes such as bond bending and bond stretching, and by observing the following rules in order of preference:

(1) Corresponding atoms of the 2-phenyl-1,2,4-triazine-3,5(2*H*,4*H*)-dione systems were overlapped exactly.

(2) When there was a choice of substitution in the 3 and 5 positions of the phenyl ring, substituents were placed according to the sequence rule of Cahn, Ingold, and Prelog:[16] the higher ranked substituent was assigned to the position more distal to the 3-carbonyl group in the triazine ring, i.e. anti.

(3) Second phenyl rings attached through a linking atom to position 4 of the first sometimes could not be overlapped exactly with the corresponding phenyl in the template. In these cases the second phenyl ring was rotated so that each of the atoms ortho to the point of attachment ("shoulders") were equidistant to the corresponding atoms in the template. An exception was compound **60**; here the tricyclic nature of the entire moiety attached to the triazine ring did not permit this rule to be applied reasonably. Only the first two rules were used.

(4) Chloro substituents at the 2-position in the second phenyl ring (compounds **18**, **27**, **30**, and **32**) were placed proximally to the 3-carbonyl group in the triazine ring, i.e. syn.

(5) Substituents at the 3-position in the second phenyl ring (compounds **11**, **25**, and **46**) were oriented distally to the 3-carbonyl group in the triazine ring, i.e. anti. This and the previous rule are arbitrary, but apparently are not critical to the analysis: orienting the substituents in the 3-position in the syn mode gave results comparable to those as when the anti mode was employed.

(6) Like **1**, the benzyl alcohols **15** and **41**, and the sulfoxide **10** are asymmetric at the atom linking the two phenyl rings. In these cases the oxygen atom was oriented as closely as possible in the same direction as the cyano group in **1**.

At the end of this process there were obtained three-dimensional structures for 54 triazines, each matching the template structure in a reasonable manner. None of these exhibited any obvious unfavorable intramolecular nonbonded contacts. In 3D-QSAR several other approaches to superimposing analog structures on a reference structure have been proposed. Cramer et al.[1] have used least squares fitting to certain key atoms; this seems to work well with rigid structure like steroids. Some, employing the Distance Geometry method, favor the assumption that the global minimum energy conformation of the most potent compound in a data set is the ideal reference structure. However, it is clear that if this assumption fails in modeling the biological activity, then less-than-ideal reference structures will be explored until a good model is achieved.[17-19] This is a perfectly sensible option in light of recent observations on significant differences between the conformations of enzyme-bound and unbound ligands.[20]

Because 3D-QSAR analyses are complicated enough, it seemed preferable in the present case to start simply. What could be more simple than to assume that one low-energy conformation of the most potent compound is as good a reference structure as any other? If the initial model fails, others could be tried later. As it turned out the approach outlined above gave such good results that there was no motivation to continue. Clearly, while this assumption seems to work for a series of semirigid triazines, it would be rash to assume that it could be broadly applied to all classes of compounds.

Physical Properties

In addition to the CoMFA steric and electrostatic fields generated by the SYBYL QSAR program, measures of lipophilicity and acidity of the triazines were used. For compounds reported in Table I, relative lipophilicities were estimated by the retention times of the compounds in a standard HPLC system and are expressed as log *k's*. Relative electron densities were used as surrogates for pK_a's, and were estimated by the chemical shift, δ_6, of the proton in the 6-position of the triazine ring. Both sets of values have been reported previously.[13]

log *k's* were not determined for compounds **59** to **71** reported in Table V. log *k's* for these compounds were estimated from the correlation between log *k'* and CLOGP[21] of 53 of the 54 compounds in Table I. Com-

(16) Cahn, R. S.; Ingold, C.; Prelog, V. Specification of Molecular Chirality. *Angew Chem., Int. Ed. Engl.* **1966**, *5*, 385–415.

(17) Ghose, A. K.; Crippen, G. M.; Revankar, G. R.; McKernan, P. A.; Smee, D. F.; Robins, R. K. Analysis of the in Vitro Antiviral Activity of Certain Ribonucleosides against Parainfluenza Virus Using a Novel Computer Aided Receptor Modeling Procedure. *J. Med. Chem.* **1989**, *32*, 746–756.

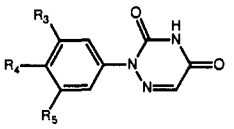
(18) Ghose, A. K.; Crippen, G. M. Modeling the Benzodiazepine Receptor Binding Site by the General Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationship Method REMOTEDISC. *Mol. Pharmacol.* **1990**, *37*, 725–734.

(19) Viswanadhan, V. N.; Ghose, A. K.; Hanna, N. B.; Matsumoto, S. S.; Avery, T. L.; Revankar, G. R.; Robins, R. K. Analysis of the in Vitro Antitumor Activity of Novel Purine-6-sulfenamide, -sulfinamide, and -sulfonamide Nucleosides and Certain Related Compounds Using a Computer-Aided Receptor Modeling Procedure. *J. Med. Chem.* **1991**, *34*, 526–532.

(20) Jorgensen, W. L. Rusting of the Lock and Key Model for Protein–Ligand Binding. *Science* **1991**, *254*, 954–955.

(21) Calculated log *P*'s (CLOGP) by MEDCHEM, version 3.54, Daylight Chemical Information Systems, Inc., 18500 Von Karman Avenue, Suite 450, Irvine CA 92715.

Table I. Anticoccidial Triazines: Potencies Predicted in Models A, B, and C for the 54 Compounds Used in Forming the Models



				log (1/MEC)						
no.	R_5	R_4	R_3	obsd[a]	model A[b]	res[c]	model B[d]	res[c]	model C[e]	res[c]
1	Cl	$CH(CN)C_6H_4$-4-Cl	Cl	3.61	2.70	0.91	2.68	0.93	3.16	0.45
2	Cl	SC_6H_4-4-Cl	Cl	3.30	2.79	0.51	2.22	1.08	2.74	0.56
3	Cl	SC_6H_4-4-Cl	Me	3.28	2.53	0.75	2.47	0.81	2.84	0.44
4	Cl	SC_6H_4-4-Ac	Me	3.19	2.30	0.89	3.34	−0.15	3.51	−0.32
5	Cl	$CH_2C_6H_4$-4-Cl	Cl	3.18	2.24	0.94	2.62	0.56	2.73	0.45
6	Cl	$SO_2C_6H_4$-4-Cl	Cl	2.94	2.75	0.19	1.89	1.05	2.49	0.45
7	Cl	COC_6H_4-4-Cl	Cl	2.90	2.88	0.02	1.80	1.10	2.19	0.71
8	Me	SC_6H_4-4-Cl	Me	2.86	2.19	0.67	2.23	0.63	2.61	0.25
9	Me	$SO_2C_6H_4$-4-Cl	Me	2.59	2.09	0.50	1.95	0.64	2.38	0.21
10	Me	$S(O)C_6H_4$-4-Cl	Me	2.58	1.43	1.15	2.22	0.36	2.21	0.37
11	Me	OC_6H_3-3-Me-4-SMe	Me	2.27	1.99	0.28	1.90	0.37	2.11	0.16
12	Cl	OC_6H_4-4-Ac	Me	2.27	1.75	0.52	3.13	−0.86	2.96	−0.69
13	Cl	OC_6H_4-4-CH(OH)Me	Me	2.27	1.13	1.14	2.74	−0.47	2.15	0.12
14	Cl	SC_6H_4-4-Cl	H	2.26	2.15	0.11	2.11	0.15	2.13	0.13
15	Cl	$CH(OH)C_6H_4$-4-Cl	H	2.26	1.27	0.99	2.11	0.15	1.71	0.55
16	Me	OC_6H_4-4-SMe	Me	2.25	1.86	0.39	1.96	0.29	2.17	0.08
17	Me	OC_6H_4-4-SMe	H	2.23	1.64	0.59	1.51	0.72	1.58	0.65
18	Cl	OC_6H_3-2-Cl-4-SO_2NMeEt	Me	2.11	2.30	−0.19	2.29	−0.18	2.27	−0.16
19	Cl	OC_6H_4-4-Cl	Me	1.98	2.23	−0.25	2.13	−0.15	2.36	−0.38
20	Cl	OC_6H_4-4-SMe	H	1.98	1.95	0.03	1.82	0.16	1.84	0.14
21	Me	OC_6H_4-4-SO_2Me	H	1.97	0.52	1.45	1.65	0.32	0.97	1.00
22	H	$SO_2C_6H_4$-4-Cl	H	1.96	1.70	0.26	1.53	0.43	1.48	0.48
23	Cl	H	Cl	1.81	1.65	0.16	0.41	1.40	1.13	0.68
24	Cl	OC_6H_4-4-I	H	1.77	2.12	−0.35	1.87	−0.10	1.92	−0.15
25	Cl	O(naphth-2-yl-6-Br)	H	1.77	2.24	−0.47	1.88	−0.11	1.77	0.00
26	H	$SO_2C_6H_4$-4-Br	H	1.71	1.79	−0.08	1.55	0.16	1.56	0.15
27	Cl	OC_6H_3-2,4-Cl_2	H	1.71	2.13	−0.42	1.65	0.06	1.67	0.04
28	Cl	$SO_2C_6H_4$-4-Cl	Me	1.71	2.40	−0.68	2.08	−0.37	2.50	−0.79
29	Me	OC_6H_4-4-Br	H	1.70	1.79	−0.09	1.50	0.20	1.60	0.10
30	Me	OC_6H_3-2,4-Cl_2	H	1.69	1.82	−0.13	1.33	0.36	1.39	0.30
31	Cl	$CH_2C_6H_4$-4-Cl	H	1.64	2.08	−0.44	2.48	−0.84	2.54	−0.90
32	Cl	OC_6H_3-2-Cl-4-SO_2NH-c-C_3H_5	Me	1.51	2.07	−0.56	2.36	−0.85	2.11	−0.60
33	Me	$SO_2N(CH_2CH_2)_2O$	Me	1.39	0.76	0.63	1.71	−0.32	1.21	0.18
34	Me	COC_6H_4-4-Cl	Me	1.38	2.01	−0.63	1.85	−0.47	1.85	−0.47
35	CF_3	Br	H	1.35	1.98	−0.63	1.59	−0.24	1.70	−0.35
36	CF_3	F	H	1.26	1.48	−0.22	1.45	−0.19	1.16	0.10
37	Cl	H	Me	1.20	1.10	0.10	0.73	0.47	1.09	0.11
38	Cl	$SO_2N(CH_2CH_2)_2O$	Cl	1.13	1.63	−0.50	1.70	−0.57	1.58	−0.45
39	Cl	$SO_2N(CH_2CH_2)_2O$	H	1.09	1.06	0.03	1.55	−0.46	1.12	−0.03
40	H	SO_2Ph	H	1.04	1.09	−0.05	1.48	−0.44	1.03	0.01
41	Me	CH(OH)Ph	H	1.01	0.24	0.77	1.55	−0.54	0.77	0.24
42	CF_3	H	H	0.93	1.42	−0.49	1.50	−0.57	1.47	−0.54
43	Me	H	Me	0.86	0.50	0.36	0.42	0.44	0.63	0.23
44	CN	H	H	0.85	0.45	0.40	1.15	−0.30	0.68	0.17
45	Me	OC_6H_4-4-Cl	H	0.74	1.77	−1.03	1.45	−0.71	1.57	−0.83
46	Cl	S(naphth-2-yl)	H	0.48	2.08	−1.60	1.99	−1.51	1.65	−1.17
47	Cl	H	H	0.25	0.77	−0.52	0.32	−0.07	0.49	−0.24
48	Cl	OMe	Cl	0.06	1.47	−1.41	−0.04	0.10	0.26	−0.20
49	H	NO_2	H	−0.03	1.19	−1.22	−0.09	0.06	0.24	−0.27
50	H	OPh	H	−0.25	1.22	−1.47	0.96	−1.21	0.64	−0.89
51	H	SO_2Me	H	−0.27	0.22	−0.49	0.55	−0.82	−0.03	−0.24
52	OMe	H	OMe	−0.30	0.04	−0.34	−0.29	−0.01	−0.18	−0.12
53	H	H	H	−0.42	0.00	−0.42	−0.03	−0.39	−0.32	−0.10
54	OMe	OMe	OMe	−0.55	−0.48	−0.07	−0.44	−0.11	−0.93	0.38

[a] MEC is expressed as mmol/kg of feed; data from ref 13. [b] From eq 5 in ref 13. [c] Difference between the obsd log (1/MEC)s and those in immediate left column. [d] Computed by PLS using only the steric and electrostatic CoMFA field columns. [e] Computed by PLS using the log k', $(\log k')^2$, δ_6, and the steric and electrostatic CoMFA field columns.

pound **10** was omitted from this regression analysis because of a missing fragment value. The correlation equation is:

$$\log k' = 0.27\ (\pm 0.014)\text{CLOGP} - 0.61 \qquad (2)$$

$$n = 53 \quad r^2 = 0.88 \quad s = 0.16 \quad F = 374$$

The δ_6's for compounds **59** to **71** were also not determined. Because each of these compounds has one or more close analogs among the original 54 congeners, δ_6's for these were estimated from the structurally closest measured analog or the mean when more than one applied. In the case of compound **60** there was an actual suitable literature value available.[22]

Results

The observed and the predicted potencies from three different models are reported in Table I. Model A is from

(22) Kluge, A. F.; Caroon, J. M.; Unger, S. H. Tricyclic Aryl-Substituted Anticoccidial Azauracils. *J. Med. Chem.* **1978**, *21*, 529–536.

Table II. Comparison of CoMFA Models

descriptors	optimum components	R^2 cross-validated	R^2 conventional
$(\log k')^2$, $\log k'$, δ_6	3	0.488	0.561
CoMFA field only	2	0.469	0.658
all	4	0.613	0.801

Table III. Statistics on CoMFA Model C

statistic	value
standard error of estimate	0.482
R^2	0.801
F	49.225
probability of $R^2 = 0.0$	0.000
number of compounds	54
number of components	4
X-V groups	0

Table IV. Contributions of Descriptors to CoMFA Model C

descriptor	fraction
$\log k'$	0.229
$(\log k')^2$	0.112
δ_6	0.087
CoMFA steric field	0.293
CoMFA electrostatic field	0.278

the same MRA reported previously,[13] and is used as the basis for comparing the other models. Model B is derived from PLS analysis of the CoMFA fields alone. Model C results from combining the $(\log k')^2$, $\log k'$, and δ_6 columns with the CoMFA columns, and analyzing them by PLS. Table II compares the models by the optimum number of principal components from PLS analyses, and in terms of R^2 by cross-validation (54 X–V groups) and by conventional means (zero X-V groups).[1]

The first line in Table II gives the results of the PLS analysis of the measured physical data reported earlier.[13] As expected, the conventional R^2 is the same as that given by MRA, and the optimum number of components is three. However, PLS now allows us to get the cross-validation (X-Vd) R^2 as well. Thus, we now have a more robust estimate of the predictive value of the analysis as opposed to how well the data are fitted.[1] The X-Vd R^2 is only somewhat less than the conventional R^2.

PLS analysis of the CoMFA columns alone (model B) results in an optimum number of components of just two, and the conventional R^2 is somewhat better than that given by model A. However, there is no improvement in the X-Vd R^2. At this stage it cannot be said that CoMFA offers advantages over MRA.

In the last line of Table II we see that the combination of models A and B gives a much better result, i.e. model C. Both the X-Vd and conventional R^2 are greatly improved, and only four components were required to do this. Table III presents a summary of statistics on model C. Its conventional R^2 is respectable, accounts for 80% of the variance in the biological data, and is statistically highly significant.

Table IV shows the relative contributions of the various inputs to model C. In agreement with the earlier MRA,[13] $\log k'$ is much more important than either $(\log k')^2$, or δ_6, while the latter two make significant but lesser contributions. In addition, we are now able to quantitate some steric and electrostatic effects. Table IV only shows their aggregate contributions (29% and 28%, respectively), but by examining Figures 1–3 we get more specific ideas as to just where these effects take place and to what degree.

Figure 1 shows diclazuril (1) embedded in the CoMFA steric field created for model C. The yellow and orange polyhedra represent those regions that interact favorably with substituents on the triazine of interest; the orange color locates intersections giving the highest 5% of beneficial interactions, while the yellow color locates the next highest 5%. The blue dot (a single intersection) and the green line and green polyhedron represent regions that interact unfavorably; the blue and the green are interpreted analogously to the orange and yellow colors, except the result is negative.

It is satisfying that several easily discerned qualitative SAR are reflected in Figure 1: (SAR i) substitution at the 3 position generally enhances potency (cf. **42**, **44**, **47** with **53**); (SAR ii) a greater increase in potency is usually observed when both the 3 and 5 position are substituted, i.e. the second substitution results in a greater increase than the first (cf. **23** with **47** and **53**);[23] and (SAR iii) substitution at the 4 position of the terminal phenyl ring also enhances potency (cf. **22** and **40**).[24] Additional examples of the above SAR's can be gleaned from the broader range of compounds described in the original literature.[23-26]

Perhaps more important is the recognition of heretofore unsuspected SAR: (SAR iv) for a chiral molecule like diclazuril (1), arrangement of the substituents in the R^* configuration is compatible with good potency, but an S^* arrangement is detrimental, i.e., a non-hydrogen substitutent pointing towards the blue dot and green line in Figure 1 detracts from potency; (SAR v) near the 4-position of the terminal phenyl ring, there is a region of unfavorable interaction close to one that is favorable; and (SAR vi) ortho substitution in the second phenyl ring favors potency. Support justifying these latter effects is more subtle than the ones previously mentioned.

Regarding (SAR iv), it should be noted that early results in this series suggested that potency should increase with triazine acidity.[23] There was a consensus then that, other things being equal, triazines with sulfones as links between the phenyl rings should be more potent than those with sulfides as links. This expectation was met with uniform disappointment. In part this can be rationalized by the fact that the sulfones would be less lipophilic than the corresponding sulfides,[26] but we now see that in addition there is no conformation sulfones can achieve that will avoid the unfavorable interaction region between the two phenyl rings. Further evidence on this point is given in Predictions.

(SAR v) is not apparent from Figure 1. The earlier workers noted increased potency when the second phenyl ring was substituted in the 4-position by chloro.[26] Acetyl and methylthio also appear to be beneficial to potency. However, when larger substituents were employed, potency began to decline (see **18** and **32**). Figure 2 shows the di-

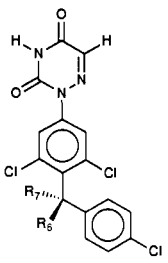
(23) Miller, M. W.; Mylari, B. L.; Howes, H. L., Jr.; Lynch, J. E.; Lynch, M. J.; Koch, R. C. Anticoccidial Derivatives of 6-Azauracil. 2. High Potency and Long Plasma Life of N1-Phenyl Structures. *J. Med. Chem.* **1979**, *22*, 1483–1487.

(24) Carroll, R. D.; Miller, M. W.; Mylari, B. L.; Chappel, L. R.; Howes, H. L., Jr.; Lynch, M. J.; Lynch, J. E.; Gupta, S. K.; Rash, J. J.; Koch, R. C. Anticoccidial Derivatives of 6-Azauracil. 5. Potentiation by Benzophenone Side Chains. *J. Med. Chem.* **1983**, *26*, 96–100.

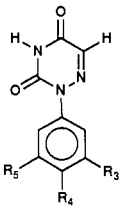
(25) Miller, M. W.; Mylari, B. L.; Howes, H. L., Jr.; Figdor, S. K.; Lynch, M. J.; Lynch, J. E.; Koch, R. C. Anticoccidial Derivatives of 6-Azauracil. 3. Synthesis, High Activity, and Short Plasma Half-Life of 1-Phenyl-6-azauracils Containing Sulfonamide Substituents. *J. Med. Chem.* **1980**, *23*, 1083–1087.

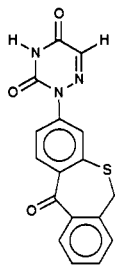
(26) Miller, M. W.; Mylari, B. L.; Howes, H. L., Jr.; Figdor, S. K.; Lynch, M. J.; Lynch, J. E.; Gupta, S. K.; Chappel, L. R.; Koch, R. C. Anticoccidial Derivatives of 6-Azauracil. 4. A 1000-fold Enhancement of Potency by Phenyl Sulfide and Phenyl Sulfone Side Chains. *J. Med. Chem.* **1981**, *24*, 1337–1342.

Table V. Triazines for Prediction, Their Calculated log k's and Their Inferred ^{1}H-NMR Chemical Shifts at Position 6



	R_5	R_7
1R*	CN	H
1S*	H	CN
59	CN	CH_3





60

no.	R_3	R_4	R_5	CLOGP[a]	calcd log k'[b]	δ_6 (ppm)[c]	ref
1R*[d]	Cl	$CH(CN)C_6H_4$-4-Cl	Cl	4.901	0.78	7.749	27
1S*	Cl	$CH(CN)C_6H_4$-4-Cl	Cl	4.901	0.78	7.749	27
59	Cl	$C(Me)(CN)C_6H_4$-4-Cl	Cl	5.300	0.82	7.749	27
60		(see structure above)		3.668	0.38	7.650[e]	22
61	Cl	Cl	H	3.418	0.31	7.670	23
62	CF_3	Cl	H	3.588	0.36	7.718	23
63	I	H	I	4.238	0.53	7.682	23
64	Cl	$SO_2N(CH_2CH_2)_2CH_2$	H	2.533	0.07	7.754	25
65	Cl	$SO_2N(CH_2CH_2)_2CHOH$	H	0.446	−0.49	7.754	25
66	Cl	$SO_2N(CH_2CH_2)_2S$	Me	2.457	0.05	7.741	25
67	Cl	SO_2NMe_2	Me	1.839	−0.11	7.741	25
68	Cl	SO_2Me	Cl	0.657	−0.43	7.775	26
69	Cl	$SO_2C_6H_4$-4-Cl	H	3.146	0.24	7.754	26
70	Et	SC_6H_4-4-Cl	Me	6.654	1.19	7.669	26
71	Cl	SC_6H_4-4-t-Bu	H	6.393	1.12	7.660	26

[a] Calculated log P's (CLOGP), see ref 21. [b] From the equation: log k' = 0.27CLOGP − 0.61 (see text). [c] Value of NMR chemical shift at position 6 of closest analog in original 54 compounds of ref 13 or mean when more than one applies. [d] Reference compound, diclazuril; arbitrarily selected R configuration as template for overlaying other compounds in CoMFA. [e] Actual literature value (ref 22).

phenyl ether 18 embedded in the steric CoMFA field, and leads to a plausible explanation for these results: the orange and yellow polyhedra show a region for favorable interaction close to the 4-position of the second phenyl ring; this suggests that substituents occupying this region would lead to increased potency; on the other hand, as the size of the substituent increases, as with 18, unfavorable interactions in the region of the blue and green polyhedra begin to dominate. Hence, 18 is not as potent as hoped.

Traditional SAR methods led Miller et al.[26] to conclude that substitution in the ortho position of the terminal phenyl ring was detrimental to potency. Upon examining Figure 2 closely, one arrives at the opposite opinion: there are small orange and yellow polyhedra near the region in question. Hence, (SAR vi) seems more likely to be true.

In discussing the electrostatic CoMFA field for model C it should be borne in mind that a substituent may add to and detract from activity owing to different aspects it may affect. For example, the 4-sulfonyl group likely does lower the pK_a and thus contributes to factors enhancing potency, but as already mentioned it also has a steric interaction that is detrimental to potency. Figure 3 shows morpholino 4-sulfonyl analog 33 embedded in the CoMFA electrostatic field. There is a region of favorable interaction (orange and yellow polyhedra) near the 4-position of the phenyl ring (SAR vii). This appears to represent attractive forces directly between this part of the drug and the receptor, and probably reflects the potency enhancing effects that electron withdrawing substituents in that area would have: for example, the cyano group of diclazuril (1) and the sulfonyl groups found in various analogs.

Surprisingly, there is also a region in the electrostatic field which disfavors potency: (SAR viii). In Figure 3 a blue dot and a green polyhedron are near the oxygen atom of the morpholine system. Again these compounds were prepared originally with the hope that the electron withdrawing properties of the sulfonyl group would lead to greater potency, and again disappointing results were obtained.[25] Apart from the unfavorable steric interactions the sulfonyl group would experience (SAR iv), it appears the oxygen atom of the morpholine group also has an unfavorable electrostatic interaction.

With the CoMFA electrostatic field giving such consistent results, one might understandably raise the question as to why then include the descriptor δ_6 to estimate the relative pK_a's. The answer is that the Gasteiger-Hückle method calculates essentially the same electrostatic charge at all positions of the triazine ring regardless of the substitution on the phenyl ring. The G–H method is not as sensitive to substitution effects as is the NMR method. Thus, δ_6 is a better indicator of relative pK_a.

Predictions

Model C accounts for 80% of the variance in log (1/MEC) and "predicts" the potencies of the 54 triazines that were used to generate the model with reasonable accuracy. The range in potency covers 4 orders of magnitude while the largest residue is only one order of magnitude. It was of interest to see how well model C would predict the potencies of triazines outside of the original 54. Table V lists physical data for 13 triazines that were not used to make the models, 14, if the S* enantiomer of diclazuril is counted. The log k's of these compounds were not determined by experiment; they were estimated from the correlation between CLOGP and log k' of the original compounds. Similarly, the chemical shifts of these compounds were also not determined; they were estimated from the chemical shifts of close analogs found in the original 54.

The resolution of the optical isomers of diclazuril has not been published; therefore, it is not known what the potencies of the enantiomers are. However, as shown in

Table VI. Anticoccidial Triazines: Potencies of 13 Compounds Not Used in Forming Models A, B, and C but Predicted by Them

	log (1/MEC)						
no.	obsd[a]	model A[b]	res[c]	model B[d]	res[c]	model C[e]	res[c]
1R*[f]	3.61	2.70	0.91	2.68	0.93	3.16	0.45
1S*		2.70		2.11		2.82	
59	2.62	2.76	−0.14	2.25	0.37	2.63	−0.01
60	1.04	1.53	−0.49	1.31	−0.27	1.25	−0.21
61	0.63	1.55	−0.92	0.34	0.29	0.77	−0.14
62	1.56	2.00	−0.44	1.49	0.07	1.59	−0.03
63	1.17	1.98	−0.81	0.66	0.51	1.60	−0.43
64	0.17	1.69	−1.52	1.76	−1.59	1.71	−1.54
65	−0.11	−0.03	−0.08	1.84	−1.95	0.67	−0.78
66	0.53	1.54	−1.01	2.32	−1.79	2.03	−1.50
67	0.76	1.12	−0.36	1.48	−0.72	1.31	−0.55
68	0.13	0.35	−0.22	0.89	−0.76	0.38	−0.25
69	2.30	2.05	0.25	1.75	0.55	1.91	0.39
70	0.79	2.20	−1.41	2.25	−1.46	2.62	−1.83
71	0.51	2.16	−1.65	2.48	−1.97	2.40	−1.89

[a] From values given in refs in Table V expressed in mmol/kg of feed. [b] Calcd from eq 5 in ref 13. [c] Difference between the obsd log (1/MEC)'s and those in immediate left column. [d] Computed by PLS using only the steric and electrostatic CoMFA field columns. [e] Computed by PLS using the log k′, log k'^2, δ_6, and the steric and electrostatic CoMFA field columns. [f] Reference compound, diclazuril, arbitrarily selected *R* configuration as template for overlaying other compounds in CoMFA.

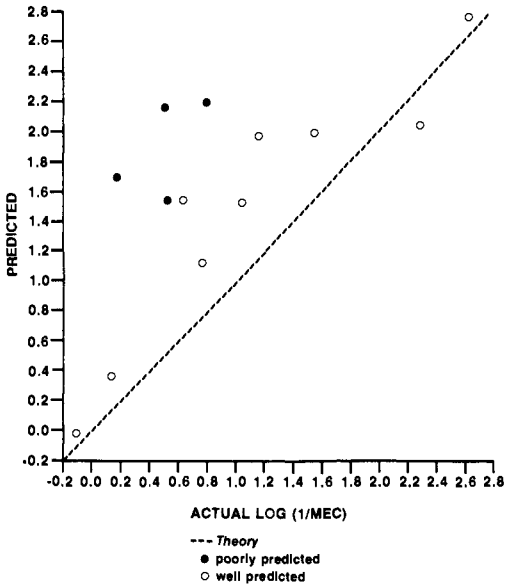


Figure 4. Actual versus predicted potencies of compounds **59** to **71** according to model A; (○) well predicted (see text); (●) poorly predicted; dashed line shows theoretical values.

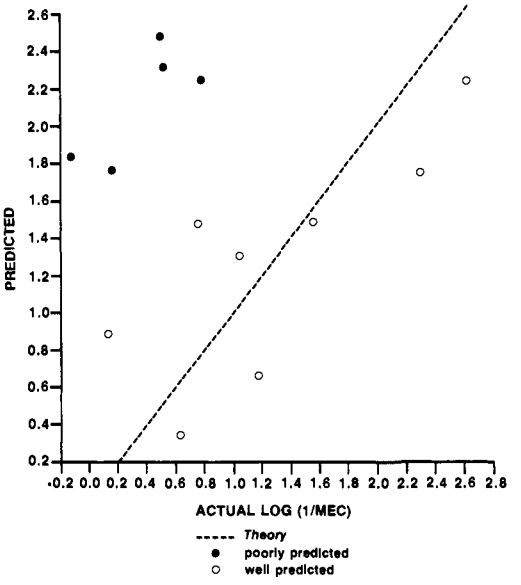


Figure 5. Actual versus predicted potencies of compounds **59** to **71** according to model B; symbols and line interpreted as in legend to Figure 4.

Table VI, model C predicts the *S** enantiomer to be less potent than the *R**. The difference in the estimated potencies of the enantiomers is small, within the error of prediction; therefore, it would probably not be worth the effort to resolve them. But if anyone should, the prediction is that each enantiomer will be active, but one will be about twice as potent as the other.

Model C predicts the potencies of the α-methyl analog **59**[27] and compound **60**[22] accurately. These results are excellent support for this work, because these compounds were not prepared or tested in our laboratories, and are therefore unbiased observations. In addition, the lower potency of **59** compared to **1** is consistent with (SAR iv), i.e. the idea that there is an unfavorable steric interaction in the region where the two phenyl rings join. However, four of the 13 triazines are not well predicted at all. See

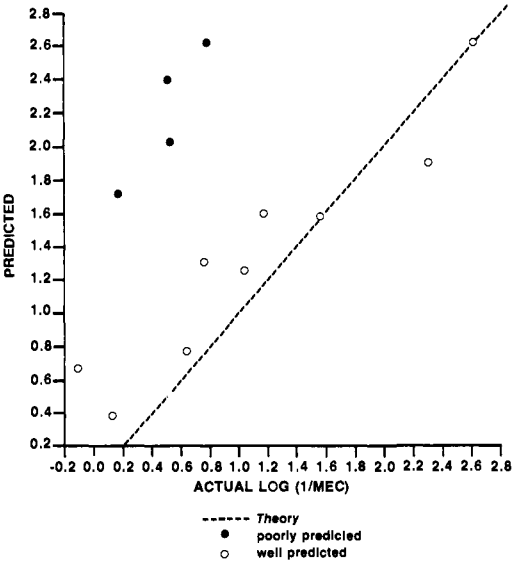


Figure 6. Actual versus predicted potencies of compounds **59** to **71** according to model C; symbols and line interpreted as in legend to Figure 4.

(27) Boeckx, G. M.; Raeymaekers, A. H. M.; Sipido, V. Anti-protozoal α-Aryl-4-(4,5-dihydro-3,5-dioxo-1,2,4-triazin-2(3*H*)-yl)-benzeneacetonitrile Derivatives, Pharmaceutical Compositions, and Method of Use Therefore. U.S. Patent 4,631,278 (December 23, 1986).

Figures 4–6. Four compounds badly predicted in models A and C, and four out of the five compounds mispredicted in model B are the same compounds, and they are each predicted to be more potent than they actually are. In each case the criterion for good prediction is that the calculated potency falls within ±0.964 of the actual, i.e. within two standard errors for model C. The errors of prediction in these models are in the preferred direction for discovery research: weakly active compounds may be wrongly classified as potent, but truly potent compounds are not identified as being weak.

Given the X-Vd R^2 for model C, it is perhaps acceptable that only 9 of the 13 compounds were well predicted. Forecasting the potencies of compounds not used in formulating the model employed is rare in the literature. The results of this attempt are similar to those given by Cramer et al. in the original CoMFA paper,[1] and those of Selwood et al.[28] It seems advisable that more work should be done along these lines in order to build a reference base. In the absence of much other similar work it is difficult to say whether the current efforts at prediction are better or worse than could be expected.

Discussion

If nothing else, the results of this study are gratifying in that the potencies of the anticoccidial triazines have been correlated with their physical properties to account for 80% of the variance, an outcome superior to that of the previous effort.[13] In addition, new and heretofore unsuspected SAR's have been uncovered. All of the findings are consistent with what is known about this class of compounds. This was accomplished with biological data obtained from a whole animal infectious disease model.

However, some significance must be attached to the fact that to achieve these results it was necessary to use: (i) δ_6 and log k' which are believed to be related to pK_a and log P, respectively; and (ii) interaction energies from specific steric and electrostatic regions in the CoMFA fields. It is plausible that the successful use of this combination relates to the in vivo nature of the biological data.

Both pK_a and log P are properties associated with transport and distribution (pharmacokinetics) in tissues and biological fluids, and therefore will have an important influence on the potencies of compounds in whole animals.[29] Further, factors governing the distribution of drugs usually do not have significant stereospecific components as is the case with pK_a and log P. Thus, it is perhaps not surprising that the earlier MRA study was able to succeed as well as it did.[13]

The polyhedra in Figures 1–3 reflect the CoMFA inputs into the model, and are reminiscent of features encountered in a receptor. Thus, the CoMFA input is doing what it was designed to do: model the interaction between a drug and its receptor. This latter aspect will also have an impact on potency. Contrary to the situation described above, drug–receptor interactions usually do have strong dependencies on stereospecific features such as those supplied by CoMFA.

Early in the history of QSAR, Hansch recognized lipophilic, electronic, and steric effects as the major factors contributing to drug potency.[30] In the work that followed, the lipophilic aspects were most reliably treated. In many systems, the electronic effects were often also readily discerned. While there were some notable successes with steric effects, they were for the most part the least well handled.

The current study to some extent recapitulates this situation: in the earlier work based on MRA, the lipophilic contribution was the most easily identified while the electronic effect as modeled by δ_6 made a small but noticable contribution.[13] There was no reasonable way to provide for steric effects. Thus, the MRA approach reached its limit. By adding CoMFA to the analysis it was possible to introduce steric effects and to augment the electronic contribution with additional interactions. Thus, it appears that an understanding of the SAR amongst anticoccidial triazines depends on recognizing that important effects are taking place on two levels: the pharmacokinetic and the drug–receptor levels. The former appears to be readily treated by the MRA approach, while the latter needs a more sophisticated technique like CoMFA.

QSAR problems are multivariate in nature. It is usually difficult to sort out the various factors contributing to biological activity. Since the original Hansch–Fujita paper,[31] numerous methods and techniques have been introduced to enhance our abilities to do this. CoMFA is one of the most important.

Supplementary Material Available: Atomic coordinates, connection tables, and fractional charges for compounds **1**, **2**, **5**, **7**, **9**, **10**, **15**, **18**, **25**, **33**, and **60** (22 pages). Ordering information is given on any current masthead page.

(28) Selwood, D. L.; Livingstone, D. J.; Comley, J. C. W.; O'Dowd, A. B.; Hudson, A. T.; Jackson, P.; Jandu, K. S.; Rose, V. S.; Stables, J. N. Structure–Activity Relationships of Antifilarial Antimycin Analogues: A Multivariate Pattern Recognition Study. *J. Med. Chem.* **1990**, *33*, 136–142.

(29) Gaillot, J.; Bruno, R.; Montay, G. Distribution and Clearance Concepts. In *Comprehensive Medicinal Chemistry*; Taylor, J. B., Ed.; Pergamon Press: Oxford, 1990; Vol. 5, pp 71–109.

(30) Hansch, C. A Quantitative Approach to Biochemical Structure–Activity Relationships. *Acc. Chem. Res.* **1969**, *2*, 232–239.

(31) Hansch, C.; Fujita, T. ρ-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **1964**, *86*, 1616–1626.