## **Linear Discriminant and Multiple Regression Analyses of Anticoccidial Triazines**

James W. McFarland,\*<sup>\*</sup> Christopher B. Cooper,<sup>†</sup> and Deborah M. Newcomb<sup>†</sup>

*Central Research Division, Pfizer Inc., Groton, Connecticut 06340, and Terre Haute, Indiana 47808. Received November 30,1990* 

Quantitative structure-activity relationships among some anticoccidial 2-(substituted-phenyl)-l,2,4-triazine-3,5-  $(2H,4H)$ -diones were studied by multiple regression analysis (MRA, the Hansch approach) and by linear discriminant analysis (LDA). With MRA the potencies of these compounds are correlated with their reverse-phase HPLC retention times and their <sup>1</sup>H NMR chemical shifts at the 6-position. While the coefficients of the variable terms are significant, the moderate  $R^2$  (0.56) of the correlating equation suggests that predictions made from this analysis are not likely to be accurate. LDA supports the idea that these descriptors are related to potency, but the discriminant function does not lead to good classification. However, when coupled with a graphic display of the results, LDA gives a more immediate sense of the synthetic direction to take when seeking highly potent analogues. It is apparent that other important but not yet identified factors also play a role in determining the potencies of these compounds.

Diclazuril<sup>1</sup> (1) is a 2-(substituted-phenyl)-1,2,4-triazine-3,5 $(2H, 4H)$ -dione (triazine henceforth) that is highly potent in the control of avian coccidiosis.<sup>2</sup> Potent anticoccidial activity in the triazines was originally discovered in these laboratories by Miller et al. $3-7$  The emergence of diclazuril on the world market caused us to reevaluate this earlier work with quantitative structure-activity (QSAR) techniques.

A previous publication mentions the use of Hammett  $\sigma$  and Hansch  $\pi$  substituent constants in the analysis of a small set of simply substituted triazines.<sup>6</sup> The correlation was not sufficiently strong  $(R^2$  approximately 0.5) to make predictions confidently, but it led to the conclusion that lipophilicity was related to potency. This in turn led to compounds with potency comparable (as we now know) to that of diclazuril. Unfortunately, the details of this analysis were not given.

The present work is retrospective in nature and for the first time offers statistical confirmation that both lipophilicity and electronic factors (possibly related to acidity) are important factors governing potency among the triazines. Measured physical properties were used as the analyzed independent variables. The correlation coefficients from the multiple regression analyses (MRA) were not sufficiently high for the associated equations to make accurate predictions of potency, but linear discriminant analysis (LDA) gave valuable insights as to where in the property space potent anticoccidial activity can be found.

**Physical Properties.** A tendency for the more acidic compounds to be the more potent was recognized early.<sup>4</sup> The low water solubility of these triazines forced acidity determinations to be conducted in aqueous solutions containing from 50 to 75%  $N$ , $N$ -dimethylformamide. Values for some compounds could not be determined at all; hence, this approach to measuring an important physical property was limited.

Kluge et al.<sup>8</sup> also pursued the triazine lead and synthesized a set of tricyclic side-chain analogues. In addition, they introduced a novel approach to understanding the structure–activity relationships. The <sup>1</sup>H NMR shift of the proton at the 6-position,  $\delta_6$ , was used as an index of the electron density in the triazine ring and thus indirectly as a measure of acidity. As an indicator of lipophilicity they used *Rm,* a parameter obtained from TLC data. From the set of triazines evaluated there was only a rough indication that compounds with the higher chemical shifts were the more potent. No relationship between potency and *R<sup>m</sup>* could be discerned.

While Kluge et al.<sup>8</sup> achieved only limited success with these descriptors, the concepts appeared potentially useful

for the anlaysis of the wider range of analogues at our disposal. These compounds had both a greater diversity in structure and a larger range in potency: minimum effective concentrations (MECs) in feed from  $0.2$  to  $>500$ ppm. One particular advantage of the NMR technique for measuring "acidity" was that our compounds were all soluble in DMSO- $d_6$ . Because no single TLC system could adequately distinguish all of the compounds considered here, *Rm* cannot be used conveniently as a measure of relative lipophilicity. However, we adapted the idea by substituting reverse-phase HPLC retention times for *Rm.*  Each retention time was transformed into the logarithm of its capacity ratio ( $log k$ ) by the formula

$$
\log k' = \log (t_R/t_0 - 1) \tag{1}
$$

wherein  $t<sub>R</sub>$  is the retention time of the analyte and  $t<sub>0</sub>$  is the column dead time, log *k'* has been shown to be directly proportional to log P.<sup>9</sup>

**Compound Selection.** The  $\delta_6$ s and log k's for 54 triazines of diverse structure are listed in Table I. All compounds with MECs < 1 ppm were chosen for the study. The various structural subclasses and the wide range of potencies are represented by the other analogues selected. Compounds were not added or removed for the purpose of obtaining higher correlation coefficients. In these studies we omitted 2-[3-chloro-4-(4-chlorobenzoyl) phenyl]-1,2,4-triazine-3,5(2H,4H)-dione (55), a potent compound that had undergone metabolism studies. It was shown that this ketone is rapidly and completely metabolized to the alcohol 15 in chickens.<sup>10</sup> Sodium borohydride also readily reduces the ketone to 15. For this reason other easily reduced benzophenones were also omitted. On the other hand, we included the sterically hindered ketones

- (l) Clinacox, Janssen Pharmaceutica.
- (2) Vanparijs, O.; Marsboom, R.; Desplenter, L. *Poultry* Set. 1989, *68,* 489.
- **(3)**  Miller, M. W. U.S. Patent 3,905,971, Sept. 16, 1975.
- **(4)**  Miller, M. W.; Mylari, B. L.; Howes, H. L., Jr.; Lynch, J. E.;
- (5) Lynch, M. J.; Koch, R. C. *J. Med. Chem.* 1979, *22,* 1483. Miller, M. W.; Mylari, B. L.; Howes, H. L., Jr.; Figdor, S. K.;
- Lynch, M. J.; Lynch, J. E.; Koch, R. C. *J. Med. Chem.* 1980, *23* 1083
- (6) 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. *J. Med. Chem.* 1981, *24,* 1337.
- **(7)**  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. *J. Med. Chem.* 1983, *26,* 96.
- (8) Kluge, A. F.; Caroon, J. M.; Unger, S. H. *J. Med. Chem.* 1978, *21,* 529.
- (9) Taylor, P. J. In *Comprehensive Medicinal Chemistry;* Ramsden, C. A., Ed.; Pergamon Press: Oxford, 1990; Vol. 4, p 241.
- (10) Rash, J. J.; Lynch, M. J. *Drug Metabolism Disposition* 1976, 4,59.

f Groton, CT.

<sup>\*</sup> Terre Haute, IN.

## Table I. Triazines: Substituents, MECs, Classes, log k's, <sup>1</sup>H NMR Shifts at Position 6





<sup>a</sup>MEC is expressed as mmol/kg of feed. <sup>b</sup> From eq 5. 'Logarithm of the capacity ratio (see ref 9).  $d$  In DMSO- $d_6$ . 'MEC measured as >250 ppm; assumed it to be one level higher. 'Measured as >500 ppm; assumed MEC to be next level higher.

7 and 34 because they are not reduced by  $NaBH<sub>4</sub>$  and therefore presumably not metabolized in the same manner as 55.

Chemistry. The new compounds (5, 7, and 31) described here were prepared according to the procedures outlined in Scheme I. Diclazuril was hydrolyzed to the amide 56,<sup>11</sup> which in turn was treated with 2 equiv of NaH to provide the diarylmethane  $5.^{12}$  The ketone 55 was

reduced by  $Et_3SiH^{13}$  in  $CF_3CO_2H$  to give the diarylmethane 31. The highly hindered ketone 7 required considerably more effort. The diarylacetonitrile 57<sup>14</sup> was treated with aqueous NaOH under an  $O<sub>2</sub>$  atmosphere to

- (12) Precedence: Smith, H. A.; Hauser, C. R. *J. Am. Chem. Soc.*  1969, *91,* 7774.
- (13) West, C. T.; Donnelly, S. J.; Kooistra, D. A.; Doyle, M. P. *J. Org. Chem.* 1973, *38,* 2675. Also: Doyle, M. P.; West, C. T. *J. Org. Chem.* 1975, *40,* 3835.
- (11) Boeckx, G. M.; Raeymaekers, A. H. M.; Sipido, V. E.P. 0232932, August 19,1987.
- (14) Boeckx, G. M.; Raeymaekers, A. H. M.; Sipido, V. U.S. Patent 4,631,278, December 23, 1986.



give the benzophenone 58.<sup>16</sup> This benzophenone was then transformed into 7 by methods previously described.<sup>7</sup>

**Multiple Regression Analysis.** MRA was performed on the 54 triazines described in Table I. In the equations below, the coefficients are given with their standard deviations in parentheses. Equation 2 shows that potency is modestly but significantly correlated with  $\log k'$  ( $p =$ 0.000 000 3). By itself  $\delta_6$  is a weak descriptor, but eq 3 is statistically significant  $(p = 0.015)$ . The combined descriptors lead to eq 4 in which the role of  $\delta_6$  is much enhanced (p = 0.0011, Student's *t* test). Addition of the log *k a* term to give eq 5 resulted in further improvement. Table II is the correlation matrix for the descriptors.

 $log (1/MEC) = 1.41 \ (\pm 0.24) log k' + 0.95$  (2)

$$
n = 54 \qquad r^2 = 0.40 \qquad s = 0.81 \qquad F_{1,52} = 34.58
$$

 $log (1/MEC) = 7.58 (\pm 3.01)\delta_6 - 56.62$  (3)

$$
n = 54 \qquad r^2 = 0.11 \qquad s = 0.99 \qquad F_{1,52} = 6.33
$$

 $log(1/MEC)$  =

1.42 ( $\pm$ 0.22) log k' + 7.74 ( $\pm$ 2.25) $\delta$ <sub>6</sub> - 58.52 (4)

$$
n = 54 \qquad R^2 = 0.51 \qquad s = 0.74 \qquad F_{2,51} = 26.82
$$

$$
\log (1/\text{MEC}) = -1.20 \ (\pm 0.51) \log k^2 +
$$
  
2.55 (\pm 0.52) log k' + 7.49 (\pm 2.16) $\delta_6$  - 56.57 (5)

 $n = 54$   $R^2 = 0.56$  $s = 0.71$   $F_{3.50} = 21.30$ 

Equation 5 appears to be about the best that can be done with the data at hand. The optimum value for log  $k$ ' is 1.06, which is consistent with the values for the more

**Table II.** Correlations among the Descriptors



Figure 1. Triazines: anticoccidial activity as a function of log  $k'$  and NMR shift.

potent compounds. Clearly, additional factors that contribute to the potency of these triazines have not yet been recognized.

While eq 5 is highly significant statistically, it is unsatisfactory in at least two ways: (i) the standard deviation is such that within the 95% confidence limits one could expect an actual MEC to be off from the predicted value by a factor of up to 26 in either direction and (ii) as shown in the column of residuals (res) in Table I, the potencies of those compounds with MECs  $\leq$ 2 ppm are consistently underestimated while the potencies of the others are mostly overestimated. Thus, eq 5 has a systematic bias and is not a reliable guide toward additional potent analogues.

**Linear Discriminant Analysis.** We next analyzed the data using LDA to see if we could gain insights into potency-determining factors that might have been overlooked by using MRA alone. To this end we arbitrarily grouped the 54 triazines in Table I into four potency classes: class 0 (MEC from 100 to >500 ppm), class 1 (10 to 100 ppm), class 2 (1 to 10 ppm), and class 3 (0.1 to 0.5 ppm).  $\log k'$ and  $\delta_6$  were used as variables, and LDA was performed with the SAS procedures STEPDISC and DISCRIM. The results are displayed graphically in Figure 1.

The aim of STEPDISC is to find independent variables that allow the maximum separation of the *centroids* of the various groups. The variables so discovered are thus identified as descriptors that probably influence biological activity. DISCRIM, on the other hand, computes discriminant functions for *classifying* the observations into their respective groups.

LDA by STEPDISC supports the conclusion drawn from MRA (above): *jointly*  $\log k'$  and  $\delta_6$  have a marked

<sup>(15)</sup> Selikson, S. J.; Watt, D. S. *J. Org. Chem.* **1975,** *40,* 267.

Table III. Linear Discriminant Analysis: Probability To Enter Variable in STEPDISC Procedure

variable	classes	n	$p$ (to enter) <sup>a</sup>	decision
log k'	$0$ to $3$	54	0.0001	accept
$\delta_6$	$0$ to $3$	54	0.0008 <sup>b</sup>	accept
log k'	$0$ to $2$	46	0.0001	accept
$\delta_6$	0 to 2	46	$0.2686^{b}$	reject

° Most likely variable entered first and then the other is considered. 'After log *k'*has been accepted.

influence on anticoccidial potency. Table III shows the probabilities to enter for the independent variables contemplated here. When all four classes are considered together (classes 0 to 3) log *k'* is identified as the variable more likely to be a significant determinant of potency; once it is accepted, then  $\delta_6$  can be entered with a high level of confidence ( $p = 0.0008$ ). However, when the most potent class of triazines (class 3) is omitted from the analysis, then only log *k'* is a significant variable. This observation suggests that  $\log k$ <sup>'</sup> dominates as a determinant of anticoccidial potency throughout the series, but  $\delta_6$  is important only in distinguishing compounds in class 3 from those in the other groups, i.e., once optimum lipophilicity obtains, only then does  $\delta_6$  make a difference. The centroids of the various classes are marked with  $(+)$  in Figure 1.

DISCRIM did not lead to a reliable classification of the triazines. Only 34 out of 54 triazines are correctly classified. It is clear from the spread of the data in Figure 1 that one should not expect better than this. Nevertheless, in analyzing compound potency by *class* we have a much clearer idea of the properties necessary for high potency. In MRA we might have expected that a high  $\delta_6$  value would compensate for a low log  $k'$ , but LDA shows that this is not likely. Indeed, in Figure 1 there are several examples demonstrating this. Therefore, to produce potent analogues, one must first assure proper lipophilicity and then proper "acidity". From contemplating Figure 1, we would not expect to obtain a class 3 compound every time, but the chances are greatly improved.

## **Experimental Section**

Data Analysis. MRA was conducted with RS/1 software (BBN Software Products Corporation, Cambridge MA). LDA was performed with the STEPDISC and DISCRIM procedures (SAS Institute Inc., Cary, NC).

Determination of HPLC Retention Times. Methanol solutions of the various triazines were examined by reverse-phase high pressure liquid chromatography using a C18 Nova Pak column (4.6 mm  $\times$  15 cm, Waters Associates). The mobile phase was acetonitrile-1 % aqueous acetic acid (1:1) at 1.0 mL/min. In this system peaks were detected by UV absorption at 280 nm. Dead time was determined by injecting a methanol solution of  $NaNO<sub>2</sub>$  and reading the eluent at 230 nm.

Determination of 'H NMR Chemical Shifts at Position  $6$  ( $\delta$ <sup>a</sup>). NMR instrument: Bruker 300 MHz. Solvent: DMSO- $d$ <sup>a</sup>. Concentration: 2 mg/0.5 mL.

Anticoccidial Activity. The triazines described in this article were evaluated for anticoccidial activity by the method of Chappel

et al.<sup>16</sup> The potencies of the various compounds are expressed as the minimum effective concentration (MEC) in ppm, i.e., mg of triazine per kg of feed.

2-[3,5-Dichloro-4-[(4-chlorophenyl)methyl]phenyl]-l,2,4 triazine-3,5(2H,4H)-dione (5). A solution of 118 mg (0.28 mmol) of 56<sup>11</sup> and 1 mL of DMF was added by syringe to a suspension of NaH (25 mg, 0.61 mmol) in 2 mL of DMF stirred magnetically at room temperature. The reaction mixture was heated to 100 °C; after 4 h it was allowed to cool to room temperature and was then treated with 50 mL of 1 N HC1. The precipitated product was purified by silica gel chromatography  $\overline{\text{CH}_2\text{Cl}_2-\text{MeCN}}$ , 9:1) to give 30 mg (28%) of 5: mp 182 °C; *<sup>l</sup>K* NMR (DMSO-dg) *&* 4.32  $(s, 2 H, CH<sub>2</sub>)$ , 7.17 (d, 2 H, Ar), 7.31 (d, 2 H, Ar), 7.60 (s, 1 H, C-6), 7.74 (s, 2 H, Ar); MS (EI) *m/e* 381 (molecular ion).

Anal. Calcd for  $C_{16}H_{10}C_{18}N_3O_2$  (MW 382.63): C, 50.22; H, 2.63; N, 10.88. Found: C, 50.47; H, 2.55; N, 10.67.

2-[3,5-Dichloro-4-(4-chlorobenzoyl)phenyl]-l,2,4-triazine-3,5(2H,4H)-dione (7). A mixture of 30.0 g of 57<sup>14</sup> (0.0878) mol) in 310 mL of  $H_2O$  and 750 mL of EtOH was treated with 440 mL of 2 N NaOH (0.878 mol) at room temperature under a stream of air. The resulting blue solution was stirred vigorously for 18 h. The solution was adjusted to pH 5.0 with 6 N HC1 and was extracted with  $CH_2Cl_2$  (4  $\times$  700 mL). The combined extracts were concentrated, then redissolved in 500 mL of  $CH_2Cl_2$ , washed with H<sub>2</sub>O (1  $\times$  500 mL), and dried (Na<sub>2</sub>SO<sub>4</sub>). The dried organic layer was evaporated to give 28.4 g of a yellow-brown solid. The residue was purified by preparative high performance liquid chromatography [gradient elution: hexanes- $CH_2Cl_2$  (5:1) to hexanes-CH<sub>2</sub>Cl<sub>2</sub> (1:1)] to furnish 21.0 g (72%) of 58: mp 151-153 °C: <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  7.38–7.78 (dd, 4 H, Ar), 8.25 (s, 2 H, Ar); <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 123.50, 129.73, 130.81, 133.30, 141.79, 142.89, 148.54, 190.5 (C= $O$ ).

This material was converted to the triazine 7 according to literature procedures:<sup>7</sup> mp 108–110 °C; <sup>1</sup>H NMR (DMSO- $\bar{d}_6$ )  $\delta$ 7.68 (d, 2 H, Ar), 7.74 (s, 1 H, C-6), 7.81 (d, 2 H, Ar), 7.84 (d, 2 H, Ar); MS (EI) *m/e* 395 (molecular ion).

Anal. Calcd for  $C_{16}H_8Cl_3N_3O_3$  (MW 396.62): C, 48.42; H, 2.02; N, 10.58. Found: C, 48.17; H, 2.00; N, 10.48.

2-[3-Chloro-4-[(4-chlorophenyl)methyl]phenyl]-l,2,4-triazine-3,5(2H,4H)-dione (31). To a stirred solution of 55 (1.0 g, 2.76 mmol) in  $CF<sub>3</sub>CO<sub>2</sub>H$  (3.2 mL) at room temperature was added  $Et<sub>3</sub>SiH$  (0.97 mL, 6.1 mmol) in one portion by syringe. The reaction solution was stirred for 48 h and was then poured into 50 mL of H20 to produce a colorless precipitate. The solids were taken up in EtOAc. The organic solution was washed with saturated aqueous NaHCO<sub>3</sub>, dried over anhydrous MgSO<sub>4</sub>, filtered, and evaporated. The residue was purified by silica gel chromatography (CH<sub>2</sub>Cl<sub>2</sub>-CH<sub>3</sub>CN, 20:1) to give 569 mg (59%) of 31 as colorless crystals: mp 153-154 °C; <sup>1</sup>H NMR (DMSO-d<sub>6</sub>)  $\delta_6$  4.10  $(s, 2 H, CH<sub>2</sub>)$ , 7.23 (d, 2 H, Ar), 7.36 (d, 2 H, Ar), 7.44 (s, 2 H, Ar), 7.62 (d, 2 H, C-6), 12.43 (s, 1 H, NH); MS (EI) *m/e* 347 (molecular ion).

Anal. Calcd for  $C_{16}H_{11}Cl_2N_3O_2$  (MW 348.19): C, 55.19; H, 3.18; N, 12.07. Found: C, 55.00; H, 3.05; N, 12.10.

Acknowledgment. We thank Kyle T. Blair, Paul Bowles, Jerold W. Hargis, David A. Koss, Maria L. Muzzi, and Richard D. Sweet for valuable technical assistance given to this work.

<sup>(16)</sup> Chappel, L. R.; Howes, H. L., Jr.; Lynch, J. E. *J. Parasitol.*  1974, *60,* 415.