1% MeOH gave 14 (0.48 g, 41%), which was crystallized from Me_2CO/MeOH/hexane: mp 204-207 °C; $[\alpha]_D$ +47.5°; UV λ_{max} (MeOH) 242 nm (ϵ 13600); ^1H NMR (Me_2SO-d_6) δ 0.76 (C_{16}-CH_3, d, J = 7 Hz), 1.08 (C_{13}-CH_3, s), 1.40 (C_{10}-CH_3, s), 3.22 (7α-H, m), 4.24 (11α-H, m), 4.75 and 5.07 (C_{21}-H's, d, J = 18 Hz), 5.94 (C_4-H, d, J = 2 Hz), 6.18 (C_2-H, dd, J = 10 and 2 Hz), 7.34 (C_1-H, d, J = 10 Hz). Anal. ($C_{24}H_{32}O_7$·CH_3OH) C, H.

Similar treatment of 17 (0.87 g, 1.56 mmol) gave 16α-methyl-7β,11β,17α,21-tetrahydroxy-1,4-pregnadiene-3,20-dione 17,21-dipropionate (15; 0.24 g, 31%) as white crystals: mp 125-133 °C; $[\alpha]_D$ +32.4°; UV λ_{max} (MeOH) 243 nm (ϵ 13600); ^1H NMR (Me_2SO-d_6) δ 0.81 (C_{16}-CH_3, d, J = 6 Hz), 1.01 (C_{13}-CH_3, s), 1.38 (C_{10}-CH_3, s), 3.12 (7α-H, m), 4.22 (11α-H, m), 5.87 (C_4-H, d, J = 2 Hz), 6.10 (C_2-H, dd, J = 10 and 2 Hz), 7.26 (C_1-H, d, J = 10 Hz). Anal. ($C_{28}H_{38}O_8$) C, H.

In a similar manner, 18 (0.447 g, 0.86 mmol) gave 16β-methyl-7β,11β,17α,21-tetrahydroxy-1,4-pregnadiene-3,20-dione 17,21-dipropionate (16; 0.167 g, 38%) as a noncrystallizable foam.

7α-Fluoro-16α-methyl-11β,17α,21-trihydroxy-1,4-pregnadiene-3,20-dione 17,21-Dipropionate (3b). To a solution of 15 (0.22 g, 0.44 mmol) in CH_2Cl_2 (12 mL) at 0 °C was added *N,N*-diethyl(2-chloro-1,1,2-trifluoroethyl)amine (0.385 mL, 0.455 g, 2.4 mmol). After 1.75 h at 0 °C, the solvent was removed under reduced pressure at 0 °C and the gummy product was purified by preparative TLC (development solvent: Et_2O/hexane, 2:1) to give **3b** (0.095 g, 43%). Crystallization from EtOAc/hexane gave **3b** (0.069 g), mp 157-161 °C.

7α-Fluoro-16α-methyl-11β,17α,21-trihydroxy-1,4-pregnadiene-3,20-dione 21-Acetate (3a). 14 (0.130 g, 0.3 mmol) in CH_2Cl_2 (75 mL) was treated with *N,N*-diethyl(2-chloro-1,1,2-trifluoroethyl)amine (0.286 mL, 0.34 g, 1.8 mmol) for 18 h at 0 °C. The solvent was removed under reduced pressure at 0 °C, and the solid product was purified on preparative TLC (development solvent: EtOAc/$CHCl_3$, 2:5). The product **3a** (0.088 g, 67%) was contaminated with a small amount of the 6,7-dehydro compound **1a**, as shown by the UV spectrum. Further purification on a preparative reverse-phase high-pressure LC column (Whatman Magnum-9 ODS-2), eluting with MeOH/H_2O (3:2), gave, on removal of the MeOH under reduced pressure, **3a** as white crystals, mp 139-143 °C.

7α-Fluoro-16β-methyl-11β,17α,21-trihydroxy-1,4-pregnadiene-3,20-dione 17,21-Dipropionate (3d). To a solution of 16 (0.167 g, 0.33 mmol) in CH_2Cl_2 (10 mL) at 0 °C was added *N,N*-diethyl(2-chloro-1,1,2-trifluoroethyl)amine (0.292 mL, 0.346 g, 1.83 mmol). After 2 h at 0 °C, another portion of the fluorinating agent (0.292 mL) was added, and the reaction mixture stirred for a further 2 h at 0 °C. TLC examination showed very little reaction had occurred, so the reaction mixture was allowed to stand at room temperature for 2 h and then evaporated to dryness under reduced pressure. Separation of this product on preparative TLC (development solvent: $CHCl_3$/EtOAc, 5:1) gave **3d** (0.022 g, 13%), which was crystallized from Me_2CO/hexane to give **3d**, mp 109-113 °C, and 7α-fluoro-16β-methyl-17α,21-dihydroxy-1,4,9-(11)-pregnatriene-3,20-dione 17,21-dipropionate (0.045 g, 28%) as a gum: ^1H NMR ($CDCl_3$) δ 0.73 (C_{13}-CH_3, s), 1.40 (C_{10}-CH_3, s), 4.28 and 4.86 (C_{21}-H's, d, J = 17 Hz), 4.86 (7β-H, d, J = 50 Hz), 5.73 (C_{11}-H, d, J = 6 Hz), 6.10 (C_4-H), 6.23 (C_2-H, dd, J = 10 and 2 Hz), 7.16 (C_1-H, d, J = 10 Hz).

Acid-Catalyzed Elimination of 7α-Halogeno-1,4-pregnadiene-3,20-diones. A solution of 4**b** (1 g, 1.92 mmol) in HCl/dioxane (1.45%, w/v, 100 mL) was stirred at room temperature for 20 h and the solvent was removed under reduced pressure. The resulting gum was chromatographed on a column of silica gel (100 g), eluting with EtOAc/hexane (42:58) to give 1**b** (0.8 g, 86%) as a noncrystallizable foam.

In a comparative experiment, approximately 0.05 M solutions of the 7α-halo compounds **3b**, 4**b**, **5b**, and **6b** in HCl/dioxane (3% w/v) were kept at room temperature and were sampled by removing aliquots and rapidly removing the solvent under a stream of N_2. The resulting gum was examined by TLC and UV spectroscopy and, after no further change was detected, by NMR spectroscopy. Thus, the 7α-iodo and 7α-bromo compounds, **6b** and **5b**, reached equilibrium (approximately 90% 1**b** as judged from the UV maxima at 300 and 250 nm) after 2 h, the 7α-chloro compound 4**b** after 5 h, and the 7α-fluoro compound **3b** was still 50% unchanged after 5 h.

Acknowledgment. We thank Drs. T. L. Popper, A. S. Watnick, and Mr. E. L. Shapiro for helpful discussions, J. McGlotten and the staff of the Physical and Analytical Chemistry Research Department for analytical and spectral data, and R. Tiberi for carrying out one of the experiments. We acknowledge here our debt to E. L. Shapiro and M. J. Gentles for the suggestion to use AcOH as a solvent for the hydrogen halide addition reactions.

Structure-Activity Study of Antiulcerous and Antiinflammatory Drugs by Discriminant Analysis

Akio Ogino,* Shingo Matsumura,

Research Laboratories, Nippon Shinyaku Co., Ltd., 14 Nishinosho Monguchi-cho, Kisshoin, Minami-ku, Kyoto, 601

and Toshio Fujita*

Department of Agricultural Chemistry, Kyoto University, Kyoto, 606 Japan. *Received December 3, 1979*

Discriminant analysis was used in the structure-activity study of antiulcerous benzoguanamines, antiinflammatory phenylacetic acids, and aminouracils. The usual discriminant analysis requires the equality of covariance matrix for the multivariate normal distribution between observation groups. When this condition is not fulfilled for some pairs of groups, a modified procedure, the "admissible" discriminant analysis after Anderson and Bahadur, was applied. In this procedure, the model of equal covariance is not the prerequisite for the analysis. As the primary criterion for selecting the best combination of variables in the discriminant functions, we used the number of misclassified compounds which is minimum. The discriminant variables were selected from the physicochemical parameters used to analyze the variation in hydrophobicity due to structural modifications. The potency scores divided into three groups for each of the three series of compounds were predicted with more than 80% accuracy, when the two-group analysis was performed for the most potent and least potent groups omitting the intermediary group.

Discriminant analysis was first applied to structure-activity studies by Martin et al.[1] They demonstrated its usefulness when the potency of a series of drugs is roughly presented, in terms of the response level at a fixed dose or the screening ratings instead of generating a dose-response relationship for each drug. Recently, this procedure has been used with considerable success in structure-ac-

(1) Y. C. Martin, J. B. Holland, C. H. Jarboe, and N. Plotnikoff, *J. Med. Chem.*, 17, 409 (1974).

0022-2623/80/1823-0437$01.00/0

Table I. Structure, Physicochemical Parameters, and Antiulcer Activity of Benzoguanamine Derivatives (I)

substituents[a]					parameters				activity score							
										3-group anal.		2-group anal. 1:2		2:3		1:3	
2	3	4	5	6	log P	$\Sigma\sigma$	E_s	$\mathcal{F}$	obsd[b]	c	d	e	d	f	d	g	d
	Cl				2.33	0.37	0.0	0.0	1 (100)	1	1	1	1			1	1
Cl		Cl			2.07	0.46	−0.97	0.41	1 (100)	1	1	1	1			1	1
Cl			Cl		1.97	0.60	−0.97	0.41	1 (100)	1	1	1	1			1	1
	Br				2.44	0.39	0.0	0.0	1 (96)	1	1	1	1			1	1
Br		Cl			2.09	0.46	−1.16	0.44	1 (96)	1	1	1	1			1	1
F			Br		1.94	0.45	−0.46	0.43	1 (100)	1	2	1	1			1	1
F			Cl		1.81	0.43	−0.46	0.43	1 (89)	2	2	2	1			1	1
Br			Cl		1.85	0.60	−1.16	0.44	1 (81)	1	1	1	1			1	1
Cl			Br		2.04	0.62	−0.97	0.41	1 (86)	1	1	1	1			1	1
		F			1.61	0.06	0.0	0.0	1 (84)	2	2	2	2			1	1
	SCF_3				3.20	0.40	0.0	0.0	1 (87)	1	1	1	1			1	1
	CF_3				2.67	0.43	0.0	0.0	1 (81)	1	1	1	1			1	1
	NO_2				1.57	0.71	0.0	0.0	1 (85)	2	2	2	2			1	1
		Me			1.92	−0.17	0.0	0.0	2 (54)	2	2	2	2	2	2		
		Cl			2.33	0.23	0.0	0.0	2 (69)	1	1	1	1	2	2		
		NO_2			1.60	0.78	0.0	0.0	2 (41)	2	1	2	1	2	2		
Cl				Cl	1.38	0.46	−1.94	0.82	2 (54)	2	3	2	2	2	2		
	SMe				2.16	0.15	0.0	0.0	2 (49)	2	1	2	1	2	2		
Br			Br		2.09	0.62	−1.16	0.44	2 (43)	1	1	1	1	2	2		
OMe			Br		1.53	0.12	−0.55	0.26	2 (49)	2	3	2	2	2	2		
Cl			CF_3		2.27	0.77	−0.97	0.41	2 (51)	1	1	1	1	2	2		
Cl			F		1.38	0.57	−0.97	0.41	2 (78)	2	2	2	2	2	2		
Br			F		1.45	0.57	−1.16	0.44	2 (78)	2	2	2	2	2	2		
	F				1.73	0.34	0.0	0.0	2 (70)	2	2	2	2	2	2		
					1.36	0.0	0.0	0.0	2 (70)	3	3	2	2	3	3		
Cl					1.19	0.23	−0.97	0.41	3 (36)	3	3			3	3	3	3
	SO_2Me				0.29	0.60	0.0	0.0	3 (−6)	3	3			3	3	3	3
Br					1.26	0.23	−1.16	0.44	3 (26)	3	3			3	3	3	3
I					1.41	0.18	−1.40	0.40	3 (−7)	3	3			3	3	3	3
Me					1.26	−0.17	−1.24	−0.04	3 (28)	3	3			3	3	3	3
Et					1.65	−0.15	−1.31	−0.05	3 (20)	3	2			3	2	3	1
OMe					0.58	−0.27	−0.55	0.26	3 (1)	3	3			3	3	3	3
OEt					1.03	−0.24	−0.55	0.22	3 (−5)	3	3			3	3	3	3
NO_2					0.49	0.78	−1.01	0.67	3 (11)	3	3			3	3	3	3

[a] Substituents other than H are indicated. [b] Figures in parentheses are for antiulcer activity in terms of the percent inhibition of the control. [c] Calculated by eq 1a-c. [d] Calculated according to the leave-one-out procedure. [e] Calculated by eq 4a and 9a. [f] Calculated by eq 4b and 9b. [g] Calculated by eq 4c and 9c.

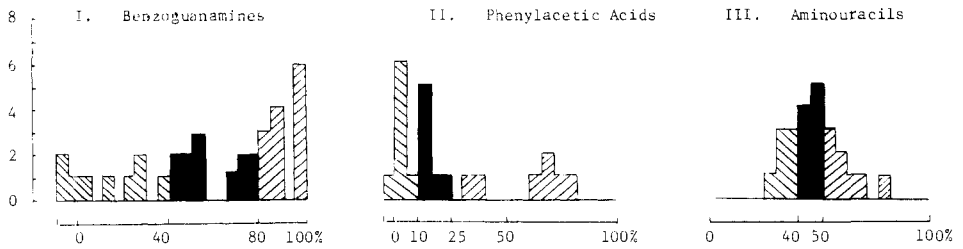


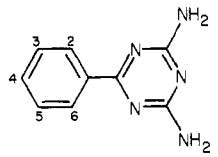
Figure 1. Frequency distribution of the response level and boundary selection for preestablished classification. The abscissa is the percent inhibition value, and the number of compounds was counted at every 5% interval. In each series of compounds, the histogram is divided into three parts (from right to left): highest, intermediary, and lowest potency groups.

tivity studies of antitumor benzoquinone and naphthoquinone derivatives[2] and thymidylate synthetase inhibiting quinazolines.[3] More recently, the procedure has been also applied to classify sets of drugs according to their therapeutic categories.[4,5] We have used discriminant analysis in structure–activity studies of a series of antiulcerous and antiinflammatory drugs. The procedure is, to a certain degree, a useful technique for discriminating the known activity ratings of congeneric drugs in terms of the linear combinations of various physicochemical parameters and for predicting the potency of untested compounds. We here report examples of the analyses for three series of drugs, I, II, and III.

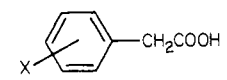
Experimental Section

Compounds. We analyzed the antiulcerous activity of 34 benzoguanamine derivatives (I), the antiinflammatory activity of 22 phenylacetic acids (II), and the activity of 24 aminouracil derivatives (III) as shown in Tables I, II, and III, respectively. The benzoguanamine derivatives[6] and phenylacetic acids[7] were

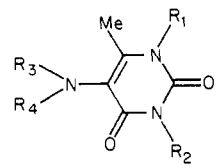
(2) (a) E. M. Hodnett, G. Prakash, and J. Amirmoazzami, *J. Med. Chem.*, 21, 11 (1978); (b) G. Prakash and E. M. Hodnett, *J. Med. Chem.*, 21, 369 (1978).
(3) B.-K. Chen, C. Horrath, and J. R. Bertino, *J. Med. Chem.*, 22, 483 (1979).
(4) S. Dove, R. Franke, O. L. Mndshojan, W. A. Schkuljev, and L. W. Chashakjan, *J. Med. Chem.*, 22, 90 (1979).
(5) D. R. Henry and J. H. Block, *J. Med. Chem.*, 22, 465 (1979).
(6) F. Ueda and K. Ohata, in preparation.
(7) Y. Tamura, Y. Yoshimoto, K. Kunimoto, S. Tada, T. Tomita, T. Wada, E. Seto, M. Murayama, Y. Shibata, A. Nomura, and K. Ohata, *J. Med. Chem.*, 20, 709 (1977).



(I) benzoguanamines



(II) phenylacetic acids



(III) aminouracils

synthesized in our laboratory. The aminouracil derivatives[8] are the gift of Professor Shigeo Senda of the Gifu College of Pharmacy.

Antiulcerous Activity. Pyloric ligation was performed under ether on ten Wistar–King rats (200–250 g) which had fasted for 24 h. Ten milligrams of each compound was suspended in 1 mL of 0.9% saline solution containing a trace of Tween 80. The suspension was administered intraperitoneally to each of the ten rats in a dose of 20 mg/kg. The ten control animals were treated only with saline solution. Ulcerous injury in the forestomach was observed macroscopically 18 h later. The severity of the injury was expressed in an ulcer index according to the criteria described by Ohata et al.[6] The effect of the drugs was judged in terms of averaged ulcer indices, and the percent inhibition value against the injury was expressed relative to the averaged index of the control group. The error involved in the value was less than ±10%.

Antiinflammatory Activity. Antiinflammatory activity was examined by the method of Winter et al.[9] Ten male Wistar rats (120–150 g) were used as a group. The hind paw volume was measured by displacement in a mercury bath. Phenylacetic acid and aminouracil derivatives, at doses of 100 and 200 mg/kg, respectively, were administered orally as 1% suspensions in 0.5% aqueous sodium carboxymethylcellulose solution. Thirty minutes later, 0.1 mL of 1% carrageenin in 0.9% saline solution was injected subcutaneously into the plantar surface of the hind paw. Three hours later, the paw volume was measured again. The increase in paw volume in drug-treated rats was averaged for the group. The inhibitory effect was represented as the percent value relative to the average volume of the control. The error in the percent value was less than ±10%.

Physicochemical Parameters. The partition coefficient of the un-ionized form, P, was determined using the 1-octanol/water system.[10] Since the pK_a values of phenylacetic acids (II) and those of the conjugate acids of benzoguanamine (I) and aminouracil derivatives (III) are located at around 5.5, 3.0, and 4.0,[10] respectively, the pH of the water phase in the partitioning system was adjusted to 1.1, 7.4, and 7.4 with suitable buffer solution. The σ, σ°, and σ^* values were used as the electronic parameters of substituents in compounds I, II, and III, respectively. Except for those evaluated from the pK_a' values determined in 50% aqueous ethanol for some phenylacetic acids, the electronic parameters were taken from the literature.[10-13] As the hydrophobic parameter, the π values of substituents were used, as well as the log P of the whole molecule for the aminouracil derivatives (III). Most π values are those for the monosubstituted benzene derivatives,[14] except for some values estimated using the additivity principle.[10,15] For the proximity effects of the ortho substituents[16] in the benzoguanamine derivatives (I), we used the Taft–Kutter–Hansch[17] E_s constant and the Swain–Lumpton–Hansch $\mathcal{F}$ constant[14] for the steric and proximity electronic effects, respectively. The reference of the E_s constant was shifted to that of H, i.e., $E_s(H) = 0$.

Discriminant Analysis. In discriminant analysis, the groups into which each member is classified should be preestablished. The classification could be done ideally by natural grouping, i.e., from the frequency distribution of the response level.[18] In this work, the distribution of the percent inhibition values for each series of drugs is not so as to naturally classify into groups of equal size. Thus, the group boundaries were simply selected to make the groups approximately of equal size. We divided each series of drugs into three groups: the most active (first), intermediary (second) and least active (third) as shown in Figure 1. The divisions still allow each group in each of the series to contain more than six members. This is compatible with a criterion put forward by Martin.[18] We first made the three-group analysis. Next, the two-group analysis was performed against each combination of the two out of three groups.

With calculations programed for computer use,[19] discriminant functions, $Z(i)$, were derived for the ith groups as linear combinations of physicochemical discriminant variables. By substituting the values of the physicochemical variables, the values of all Z functions were calculated for each compound. A compound was assigned to the ith group when the value of the $Z(i)$ function was larger than that of the other functions. For the two-group analysis, the discriminant functions were represented in the form of $Z(i)$–$Z(j)$. In this case, each compound was classified into the ith or jth group, depending upon the sign of the discriminant function. The program is not for the stepwise selection of variables but for the computation of discriminant functions with any possible combination of variables. In this respect, it is similar to the BMD 05M program.[20] We examined various combinations of variables within the scope that the number of variables does not exceed one-fifth the number of observations.[18]

The best set of discriminant functions was selected according to the following criteria: (1) A combination of variables which minimizes the number of misclassified compounds is best. In the three-group analysis, misclassification between the first and third groups is inferior to misclassification between neighboring groups. (2) If more than one combination gives the same number of misclassified compounds, the combination of the least number of independent variables is selected. (3) If more than one combination of the least number of independent variables gives the same minimum number of misclassification, the combination of independent variables among which collinearities are minimum is selected.

Results

Discriminant Functions. The best sets of discriminant functions selected for the three-group model are shown as eq 1a–c, 2a–c, and 3a–c.

benzoguanamines

$$Z(1) = -17.40 + 14.00 \log P + 7.65\sum\sigma + 6.82\mathcal{F} \quad (1a)$$

$$Z(2) = -12.23 + 11.75 \log P + 5.83\sum\sigma + 6.69\mathcal{F} \quad (1b)$$

$$Z(3) = -4.56 + 7.01 \log P + 1.09\sum\sigma + 7.17\mathcal{F} \quad (1c)$$

(8) S. Senda, K. Hirota, and K. Banno, *J. Med. Chem.*, 15, 471 (1972).
(9) C. A. Winter, E. A. Risley, and G. W. Nuss, *Proc. Soc. Exp. Biol. Med.*, 111, 544 (1962).
(10) S. Inoue, A. Ogino, M. Kise, M. Kitano, S. Tsuchiya, and T. Fujita, *Chem. Pharm. Bull.*, 22, 2064 (1974).
(11) O. Exner, "Advances in Linear Free Energy Relationships", N. B. Chapman and J. Shorter, Eds., Plenum, London, 1972, p 1.
(12) J. Shorter, "Advances in Linear Free Energy Relationships", N. B. Chapman and J. Shorter, Eds., Plenum Press, London, 1972, p 71.
(13) M. S. Tute, *Adv. Drug Res.*, 6, 1 (1971).
(14) C. Hansch, A. Leo, H. Unger, K. H. Kim, D. Nikaitani, and E. J. Lien, *J. Med. Chem.*, 16, 1207 (1973).
(15) A. Leo, C. Hansch, and D. Elkins, *Chem. Rev.*, 71, 525 (1971).
(16) T. Fujita and T. Nishioka, *Prog. Phys. Org. Chem.*, 12, 49 (1976).
(17) E. Kutter and C. Hansch, *J. Med. Chem.*, 12, 647 (1969).
(18) Y. C. Martin, "Quantitative Drug Design", Marcel Dekker, New York, 1978, pp 136 and 242.
(19) (a) HITAC 8250, NDOS, HSAP: "Statistical Computation Manual" Hitachi Seisakusho, Tokyo, 1973, p 84. (b) M. Goto, "Tahenryo Deta no Kaisekiho (The Multivariate Analysis)", Kagakujohosha, Osaka, 1973, p 157.
(20) W. J. Dixon, "BMD–Biomedical Computer Program", University of California Press, Berkeley, Calif., 1971, p 196.

phenylacetic acids

$$Z(1) = -92.27 + 136.57 \log P - 23.51(\log P)^2 + 20.84\sum\sigma^\circ \quad (2a)$$

$$Z(2) = -100.69 + 147.13 \log P - 25.88(\log P)^2 + 23.76\sum\sigma^\circ \quad (2b)$$

$$Z(3) = -88.34 + 140.10 \log P - 25.15(\log P)^2 + 25.98\sum\sigma^\circ \quad (2c)$$

aminouracils

$$Z(1) = -93.94 + 132.08 \log P - 37.14(\log P)^2 + 28.79\sum\pi(R_1 + R_2) \quad (3a)$$

$$Z(2) = -91.95 + 137.13 \log P - 38.10(\log P)^2 + 25.90\sum\pi(R_1 + R_2) \quad (3b)$$

$$Z(3) = -62.50 + 115.23 \log P - 32.14(\log P)^2 + 20.80\sum\pi(R_1 + R_2) \quad (3c)$$

The best discriminant functions for the two-group analysis are shown as eq 4a–c, 5a–c, and 6a–c.

The activity ratings calculated from these sets of discriminant functions are shown in Tables I–III.

benzoguanamines

$$\text{1st vs. 2nd: } Z(1)-Z(2) = \log P + 0.658\sum\sigma + 0.253\mathcal{F} - 2.276 \quad (4a)$$

$$\text{2nd vs. 3rd: } Z(2)-Z(3) = \log P + 0.719\sum\sigma - 1.573 \quad (4b)$$

$$\text{1st vs. 3rd: } Z(1)-Z(3) = \log P + 0.434\, E_s - 1.270 \quad (4c)$$

phenylacetic acids

$$\text{1st vs. 2nd: } Z(1)-Z(2) = (\log P)^2 - 8.623\sum\sigma^\circ - 10.906 \quad (5a)$$

$$\text{2nd vs. 3rd: } Z(2)-Z(3) = (\log P)^2 - 5.600\sum\sigma^\circ - 6.067 \quad (5b)$$

$$\text{1st vs. 3rd: } Z(1)-Z(3) = (\log P)^2 + 0.291\sum\sigma^\circ - 8.827 \quad (5c)$$

aminouracils

$$\text{1st vs. 2nd: } Z(1)-Z(2) = (\log P)^2 - 4.417 \log P + 1.401\sum\pi(R_1 + R_2) + 0.741 \quad (6a)$$

$$\text{2nd vs. 3rd: } Z(2)-Z(3) = (\log P)^2 - 3.593 \log P - 0.782\sum\pi(R_1 + R_2) + 4.633 \quad (6b)$$

$$\text{1st vs. 3rd: } Z(1)-Z(3) = (\log P)^2 - 19.333\sum\pi(R_1 + R_2) + 43.000 \quad (6c)$$

Efficacy of the Analysis. In order to obtain a sharper criterion for the efficacy of discrimination, the leave-one-out technique was applied. Each compound was first left out and then reclassified according to discriminant functions derived from other compounds by using the same combination of independent variables as the original functions. The results, which would simulate the predictability of the potency score of untested compounds more closely than the original discriminant functions, are also shown in Tables I–III. If the number of misclassified compounds does not increase significantly, the original analysis can be regarded as being stable and reliable.

With no manipulation of activity data, the total (prior) probability of misclassification, P_{mis}°, can be given by eq 7, where $P^\circ(i)$ is the prior probability of finding compounds in the ith group. The difference between this value and the ratio of the number of misclassified compounds to that of the total represents the efficacy of the discrimination. Results are summarized in Table IV. By original sets of discriminant functions, the chance of misclassification decreases ~25–50%, leading to about ~75–100% accuracy in the discrimination of activity ratings for each series of drugs. The leave-one-out procedure generally increases the number of misclassified compounds. In the three-group models, the ratio of the correctly predicted compounds decreases ~15–20% from that of the original analysis. For the two-group cases, the efficacy, as well as the stability, of the original analysis is highest in the discrimination between first and third groups, while they are lower between neighboring groups for each series of drugs.

$$P_{mis}^\circ = \sum_i P^\circ(i)[1 - P^\circ(i)] \quad (7)$$

Admissible Discriminant Functions. The classification procedure with the usual discriminant analysis is derived from the model of multivariate normal distribution of observations within each of the scored groups, such that the covariance matrix is the same for all groups.[20] The common covariance matrix value to all of the groups is required to define the linear discriminant functions. We have tested the model of "equal covariance" by means of the method of Dempster.[21] This method can only apply between two groups. The equality for three population groups was examined by testing the null hypothesis for the equality for all possible pairs of two groups. In fact, the equality was significant at the 95% level for most of the pairs of two groups in the three series of drugs, regardless of the types of analysis. Only for the pair of the first and third groups of benzoguanamines was the hypothesis abandoned in deriving eq 4c for the two-group analysis. In this situation, the "admissible" discriminant procedure developed by Anderson and Bahadur[22] can be used as an alternative.

In the usual discriminant procedure between two population groups, the discriminant functions are derived under conditions selecting a hyperplane dividing the parameter space into two regions so as to minimize the total posterior probability of misclassification, P_{mis}, evaluated according to eq 8, where $P(j/i)$ is the posterior probability

$$P_{mis} = \sum_{i,j=1,2}^{i \neq j} P^\circ(i)P(j/i) \quad (8)$$

of assigning compounds which should belong to the ith group to those of the jth group. $P(j/i)$ values are estimated from the generalized Mahalanobis distance between the centroid of the space for the ith group and that for the jth group. For the pair of first and third groups of benzoguanamines in deriving eq 4c, the conditions for the minimum probability of misclassification should be different from those based on the equicovariance model. The "admissible" procedure is to search for such conditions to derive the "most reasonable" discriminant functions. The equality of covariance matrices between groups is not required, but the "best" covariance matrix common to two groups is estimated in terms of a linear combination of unequal covariance matrices of the two groups by an iterative procedure. It is applicable only to the two-group classification problem.

Even though the application of the usual discriminant analysis was approved for most pairs of groups in this work, the level of significance at which the null hypothesis for

(21) A. P. Dempster, *Ann. Math. Stat.*, **35**, 190 (1964).
(22) T. W. Anderson and R. R. Bahadur, *Ann. Math. Stat.*, **33**, 420 (1962).

Table II. Physicochemical Parameters and Antiinflammatory Activity of Phenylacetic Acids (II)

substituents			parameters		activity score									
						3-group anal.		2-group anal. 1:2		2:3		1:3	
3	4	5	log P	$\Sigma\sigma^\circ$	obsd[a]	b	c	d	c	e	c	f	c
Ph	H	OMe	3.30	0.16	1 (68)	2	2	2	2			1	1
c-C_6H_{11}	H	OMe	3.66	−0.01	1 (35)	1	1	1	1			1	1
2-Cl-Ph	H	OMe	3.66	−0.06[g]	1 (75)	1	1	1	1			1	1
4-Cl-Ph	H	OMe	3.71	0.23[g]	1 (66)	1	1	1	2			1	1
2,4-Cl_2-Ph	H	OMe	4.08	0.05[g]	1 (60)	1	1	1	1			1	1
4-F-Ph	H	OMe	3.50	0.06[g]	1 (74)	1	1	1	1			1	1
H	*i*-Bu	H	3.38	−0.17[g]	1 (32)	1	2	1	2			1	1
4-OMe-Ph	H	OMe	3.28	0.10[g]	2 (13)	2	2	2	2	2	2		
i-Pr	H	OMe	2.75	−0.14[g]	2 (10)	2	2	2	2	2	2		
O-*i*-Bu	H	Me	3.33	−0.03	2 (13)	2	2	1[h]	1	2	2		
Ph	H	OH	2.75	0.14	2 (14)	2	2	2	2	2	3		
4-Me-Ph	H	OMe	3.72	0.12[g]	2 (23)	1	1	1	1	2	2		
4-Cl-Ph	H	OH	3.42	0.20[g]	2 (14)	2	1	2	2	2	2		
OEt	H	H	1.88	0.02[g]	2 (17)	3	3	2	2	3	3		
Me	H	OMe	1.94	−0.01	3 (6)	3	3			3	3	3	3
OBu	H	Me	3.38	−0.03	3 (3)	2	2			2	2	1	1
OMe	H	H	1.45	0.06	3 (−3)	3	3			3	3	3	3
Me	H	H	1.89	−0.07	3 (0)	3	3			3	3	3	3
F	H	H	1.59	0.35	3 (0)	3	3			3	3	3	3
Cl	H	H	2.21	0.37	3 (0)	3	3			3	3	3	3
Br	H	H	2.36	0.38	3 (0)	3	2			3	3	3	3
NO_2	H	H	1.34	0.70	3 (0)	3	3			3	3	3	3

[a] Figures in parentheses are the percent inhibition values for the edema growth. [b] Calculated from eq 2a-c. [c] Calculated according to the leave-one-out procedure. [d] Calculated by eq 5a and 10a. [e] Calculated by eq 5b and 10b. [f] Calculated by eq 5c and 10c. [g] Evaluated from a linear relationship between σ° and pK_a' values (in 50% EtOH). [h] The calculated score by eq 10a is 2.

Table III. Structure, Physicochemical Parameters, and Antiinflammatory Activity of Aminouracil Derivatives (III)

substituents			parameters			activity score								
							3-group anal.	2-group anal. 1:2		2:3		1:2	
R_1	R_2	NR_3R_4	log P	$\Sigma\pi(R_1 + R_2)$	$\Sigma\sigma^*(R_3 + R_4)$	obsd[a]	b	c	d	c	e	c	f	c
Me	Ph	NH-Pr	1.20	2.52	0.37	1 (64)	1	1	1	1			1	1
Ph	Me	NMe_2	1.19	2.52	0.0	1 (65)	1	1	1	1			1	1
Me	Ph	NH-*i*-Bu	1.48	2.52	0.36	1 (58)	2	2	2	2			1	1[g]
allyl	Ph	NMe_2	1.44	3.06	0.0	1 (54)	1	1	1	1			1	1
Ph	allyl	NMe_2	1.98	3.06	0.0	1 (75)	1	1	1	1			1	1
Ph	Et	NMe_2	2.29	2.98	0.0	1 (51)	1	1	1	2			1	1
Ph	Et	NEt_2	2.78	2.98	−0.20	1 (53)	1	3	1	2			1	1[g]
Et	Ph	NEt_2	2.20	2.98	−0.20	1 (57)	1	1	1	2			1	1
Me	Ph	NMe_2	1.30	2.52	0.0	2 (44)	1	1	1	1	2	2[g]		
Me	Ph	NEt_2	2.11	2.52	−0.20	2 (45)	2	2	2	2	2	2		
Me	Ph	NHBu	1.69	2.52	0.36	2 (48)	2	2	2	2	2	2		
Ph	Me	NHPr	1.70	2.52	0.37	2 (45)	2	2	2	2	2	2		
c-C_6H_{11}	Me	morpholino	2.39	2.82	0.67	2 (44)	2	2	2	2	2	2		
Ph	Me	$N(allyl)_2$	2.59	2.52	0.26	2 (46)	2	3	2	2	3	3		
4-Cl-Ph	Me	NMe_2	2.32	3.23	0.0	2 (41)	1	1	1	1	2	2		
4-OMe-Ph	Me	NMe_2	1.96	2.50	0.0	2 (41)	2	2	2	2	2	2		
Ph	Me	NEt_2	2.19	2.52	−0.20	2 (46)	2	2	2	2	2	2		
Me	Me	NMe_2	0.99	1.12	0.0	3 (37)	3	3			3	3	3	3
Ph	H	NMe_2	1.40	1.96	0.0	3 (35)	3	2			3	2[h]	3	3
Me	Ph	piperidino	2.76	2.52	−0.14	3 (27)	3	2			3	2[h]	3	1[h]
Me	Ph	morpholino	0.69	2.52	0.67	3 (35)	3	1			3	3	1	1
Me	Ph	pyrrolidino	2.26	2.52	−0.26	3 (32)	2	2			2	2	3	1[h]
Ph	CH_2CH_2OH	NMe_2	1.12	1.82	0.0	3 (34)	3	3			3	3	3	3
CH_2CH_2OH	Ph	NMe_2	0.85	1.82	0.0	3 (31)	3	3			3	3	3	3

[a] Figures in parentheses are the percent inhibition values for the edema growth. [b] Calculated by eq 3a-c. [c] Calculated according to the leave-one-out procedure. [d] Calculated by eq 6a and 11a. [e] Calculated by eq 6b and 11b. [f] Calculated by eq 6c and 11c. [g] Misclassified according to the leave-one-out procedure based on the corresponding admissible discriminant functions. [h] Correctly classified according to the leave-one-out procedure based on the corresponding admissible discriminant functions.

the equality of covariance is accepted varies depending upon the selection of independent variables. It is required to examine the null hypothesis after the best combination of variables is determined. In other words, it is impossible by the usual discriminant analysis to guarantee that the procedure is admissible in terms of the equicovariance standard until the final set of discriminant functions is derived. Thus, we performed the admissible discriminant analysis regardless of whether the equicovariance model is accepted or abandoned for each pair of groups of the

Table IV. Efficacy of Discriminant Analysis with Equations 1-6[a]

	3-group anal.			2-group anal.								
				I			II			III		
	I	II	III	1:2	2:3	1:3	1:2	2:3	1:3	1:2	2:3	1:3
total prior probability of misclassification, %	65.9	66.6	66.3	49.5	49.0	48.4	50.0	49.8	49.8	49.8	49.2	49.8
posterior misclassification ratio, %	20.6 (38.2)	18.2 (31.8)	16.7 (37.5)	24.0 (32.0)	4.8 (14.3)	0.0 (4.5)	21.4 (35.7)	13.3 (13.3)	6.7 (6.7)	17.6 (35.3)	12.5 (25.0)	6.7 (20.0)
correct classification, %	79.4 (61.8)	81.8 (68.2)	83.3 (62.5)	76.0 (68.0)	95.2 (85.7)	100.0 (95.5)	78.6 (64.3)	86.7 (86.7)	93.3 (93.3)	82.4 (64.7)	87.5 (75.0)	93.3 (80.0)

[a] The figures in parentheses are from the leave-one-out technique.

Table V. Statistics of Discriminant Analysis with Equations 1-6

		3-group anal.			2-group anal.								
					I			II			III		
		I	II	III	1:2	2:3	1:3	1:2	2:3	1:3	1:2	2:3	1:3
total posterior probability of misclassification, %		34.4	31.8	28.9	31.6	13.3	6.4	23.3	20.6	4.2	20.0	11.7	14.2
F values and the level of significance (% in parens) in discriminating between groups	1:2	1.77 (75.0)	2.13 (75.0)	3.45 (95.0)	1.77 (75.0)			3.29 (90.0)			3.45 (95.0)		
	2:3	7.51 (99.5)	2.82 (90.0)	6.33 (99.0)		11.81 (99.5)			4.51 (95.0)			6.33 (99.0)	
	1:3	17.04 (99.5)	11.82 (99.5)	6.27 (99.0)			101.5[a] (99.9)			18.83 (99.5)			7.89 (99.0)

[a] Statistical values were estimated by means of the method of Beherens and Fisher (ref 19b).

Table VI. Squared Correlation Matrix for the Parameters Used in the Analysis of Benzoguanamine Derivatives

	$\log P$	$(\log P)^2$	$\Sigma\sigma$	E_s	$\mathcal{J}$
$\log P$	1.000				
$(\log P)^2$	0.927	1.000			
$\Sigma\sigma$	0.035	0.041	1.000		
E_s	0.065	0.106	0.009	1.000	
$\mathcal{J}$	0.056	0.075	0.153	0.590	1.000

Table VII. Squared Correlation Matrix for the Parameters Used in the Analysis of Phenylacetic Acids

	$\log P$	$(\log P)^2$	$\Sigma\sigma^\circ$
$\log P$	1.000		
$(\log P)^2$	0.986	1.000	
$\Sigma\sigma^\circ$	0.165	0.140	1.000

Table VIII. Squared Correlation Matrix for the Parameters Used in the Analysis of Aminouracil Derivatives

	$\log P$	$(\log P)^2$	$\Sigma\pi(R_1 + R_2)$
$\log P$	1.000		
$(\log P)^2$	0.976	1.000	
$\Sigma\pi(R_1 + R_2)$	0.329	0.288	1.000

three series of drugs, including the pair of the first and third groups of benzoguanamines.

The best "admissible" functions were derived by using the computer program developed by Goto and co-workers[19b] as eq 9a–c, 10a–c, and 11a–c.

benzoguanamines

$$Z(1)\text{–}Z(2) = \log P + 0.617\sum\sigma + 0.251\mathcal{F} - 2.256 \quad (9a)$$

$$Z(2)\text{–}Z(3) = \log P + 0.762\sum\sigma - 1.562 \quad (9b)$$

$$Z(1)\text{–}Z(3) = \log P + 0.692E_s - 0.812 \quad (9c)$$

phenylacetic acids

$$Z(1)\text{–}Z(2) = (\log P)^2 - 7.203\sum\sigma^\circ - 11.390 \quad (10a)$$

$$Z(2)\text{–}Z(3) = (\log P)^2 - 8.341\sum\sigma^\circ - 6.364 \quad (10b)$$

$$Z(1)\text{–}Z(3) = (\log P)^2 - 0.637\sum\sigma^\circ - 9.822 \quad (10c)$$

aminouracils

$$Z(1)\text{–}Z(2) = (\log P)^2 - 4.434 \log P + 1.510\sum\pi(R_1 + R_2) + 0.426 \quad (11a)$$

$$Z(2)\text{–}Z(3) = (\log P)^2 - 3.668 \log P - 0.745\sum\pi(R_1 + R_2) + 4.854 \quad (11b)$$

$$Z(1)\text{–}Z(3) = (\log P)^2 - 15.302\sum\pi(R_1 + R_2) + 36.245 \quad (11c)$$

These admissible functions classify the compounds entirely in the same manner as their counterparts, eq 4a–c, 5a–c, and 6a–c. The only exception is a compound which belongs to the second group of phenylacetic acids. This is correctly classified by this procedure while misclassified by the usual analysis as shown in Table II. The results of the leave-one-out procedure are also equivalent to those from the usual analysis, except for the two-group analyses of aminouracil derivatives as shown in Table III.

Discussion

By dividing into three groups, we expected to make the classification more finely than the two-group analysis, where a series of drugs is usually divided just into high- and low-potency groups. As shown in Table IV, the efficacy of discrimination by means of the three-group analysis is moderate as the whole, the predictability being around ~60–80% for three series of drugs. The highest efficacy is shown between highest and lowest potency groups in terms of the two-group analysis. In each series of drugs, more than 30% decrease in the chance of misclassification leading to more than 80% accuracy in the prediction of the activity scores is obtained by this procedure between highest and lowest potency groups. The leave-one-out technique also shows the highest stability for this procedure.

Since the error of the original biological data is about 10% for each series, the boundary is not very sharp between neighboring groups, so that there should be an error region. Leaving out the compounds in the error region, it is reasonable that the analysis for the highest and lowest potency groups gives the best results. The total posterior misclassification probability as well as F values testing the level of significance in discriminations by means of the usual procedure are shown in Table V. The values for the discrimination of first vs. third groups of benzoguanamines are estimated by means of the method of Beherens and Fisher[19b] based on the result from the admissible analysis. The overall general trend in the misclassification probability in this table coincides with that in the misclassification ratio actually observed for the series of analyses, especially with the value from the leave-one-out technique in Table IV. The level of significance is acceptable between first and third, as well as between second and third, groups in each series of drugs, being mostly higher than 95%. It is lowest between first and second groups, which corresponds to the largest number of misclassified compounds between these two groups among the total misclassified.

Table IX. Significance of log P and $(\log P)^2$ Terms

compd series	variables used	no. of misclassified compds: total	no. of misclassified compds: between 1st & 3rd
II			
3-group anal. (eq 2)	$\log P$, $(\log P)^2$	4	0
	$\log P$	5	1
	$(\log P)^2$	4	1
III			
3-group anal. (eq 3)	$\log P$, $(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	4	0
	$\log P$, $\Sigma\pi(R_1 + R_2)$	6	1
	$(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	6	1
III			
2-group anal. (1st:2nd, eq 6a)	$\log P$, $(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	3	
	$\log P$, $\Sigma\pi(R_1 + R_2)$	4	
	$(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	4	
III			
2-group anal. (2nd:3rd, eq 6b)	$\log P$, $(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	2	
	$\log P$, $\Sigma\pi(R_1 + R_2)$	3	
	$(\log P)^2$, $\Sigma\pi(R_1 + R_2)$	3	

In this work, the results from the usual discriminant analysis are almost entirely the same as those from the admissible procedure. However, it would not be always the case. The most reasonable procedure for the three potency groups, the boundaries of which have an error region, is to discriminate the highest and lowest potency groups with the two-group analysis. First, the usual discriminant procedure can be performed and the best combination of variables are selected so as to minimize the number of misclassified compounds. Then, for the discriminant functions derived in this manner, the model of the equal covariance is examined between groups. If the model is abandoned, the admissible procedure is applied. One can use, alternatively, the admissible procedure from the start of the analysis without worrying about the model of equal covariance. Lastly, the level of significance in the discrimination should be examined.

Even in the "best" discriminant functions, some of the independent variable terms do not represent a significant improvement above the 95% level over the corresponding functions minus the variable term by F test. Even so, we prefer to apply that criteria of the minimum number of misclassified compounds to select the best combination of variables.

Tables VI–VIII show the degrees of collineality between variables examined for discriminant functions. The collinearity between log P and $(\log P)^2$ for phenylacetic acids and aminouracils is very high. One might think that one of these two variables is redundant for discrimination of phenylacetic acids and aminouracils in eq 2, 3, 6, and 11. Table IX examines the effects of these two variables on correct classification, showing that both variables are necessary to minimize the number of misclassified compounds for corresponding analyses.

To derive discriminant functions, one can use various sets of physicochemical parameters. Here, we used the experimentally determined log P value as one parameter. Others were selected from parameters which have been used to correlate the log P or π values for each of the series, as shown in eq 12,[23] 13,[24] and 14a–c.[10] In these equations, benzoguanamines

$$\log P = \underset{(0.038)}{0.913}\sum\pi(\text{monosubstituted benzene}) + \underset{(0.087)}{0.510}\sum\sigma + \underset{(0.069)}{0.602}E_s(\text{ortho}) - \underset{(0.165)}{0.974}\mathcal{F}(\text{ortho}) + \underset{(0.043)}{1.466}$$

$$n = 35;\ s = 0.133;\ r = 0.978 \qquad (12)$$

phenylacetic acids

$$\log P = \underset{(0.057)}{0.954}\pi(\text{monosubstituted benzene}) + \underset{(0.143)}{0.294}\sigma^\circ + \underset{(0.059)}{0.012}$$

$$n = 20;\ s = 0.079;\ r = 0.993 \qquad (13)$$

aminouracils

$$\pi_{R_1} = \underset{(0.212)}{0.815}\pi(\text{monosubstituted benzene}) - \underset{(0.616)}{0.698}\sigma^* - \underset{(0.312)}{1.020}$$

$$n = 8;\ s = 0.182;\ r = 0.976 \qquad (14a)$$

$$\pi_{R_2} = \underset{(0.342)}{0.762}\pi(\text{monosubstituted benzene}) - \underset{(1.103)}{1.142}\sigma^* - \underset{(0.538)}{0.445}$$

$$n = 9;\ s = 0.345;\ r = 0.940 \qquad (14b)$$

$$\sum\pi(R_3 + R_4) = \underset{(0.441)}{0.674}\sum\pi(\text{aliphatic}) - \underset{(0.758)}{0.866}\sum\sigma^*(R_3 + R_4) - \underset{(0.817)}{0.801}$$

$$n = 9;\ s = 0.262;\ r = 0.921 \qquad (14c)$$

n is the number of compounds, s is the standard deviation, r is the correlation coefficient, and the figures in parentheses are the 95% confidence intervals. It is generally acknowledged that log P is a very important parameter, especially for activities in vivo. Other effects should also participate in the variation of activity. The significance of parameters other than log P in discriminant functions may be that they adjust differences in various physicochemical effects between the partitioning process and in vivo behavior.

In this work, we limited our approach to utilizing the linear elementary discriminant procedures. When the equicovariance model does not hold between populations, we could use the quadratic discriminant analysis. The nonelementary discriminant functions may allow a mechanistic interpretation more clearly than the present procedure, especially for the simultaneous classification of three potency groups. Comparisons of the present results with those from these alternative procedures, as well as appropriate pattern-recognition techniques, will offer us useful information about application methodology of multivariate analysis to structure–activity studies, which will be reported elsewhere.

Acknowledgment. We are grateful to Professor Ikuo Moriguchi and Dr. Yvonne Martin for their helpful discussions and to Drs. Masashi Goto and Takaji Teranishi for their invaluable advice on admissible discriminant analysis. We also thank one of the reviewers for his suggestions to improve the presentations.

(23) A. Ogino, unpublished work.
(24) T. Fujita, *Prog. Phys. Org. Chem.*, in press.

Convulsant and Anticonvulsant Barbiturates. 1. Molecular Conformations from Classical Potential-Energy Calculations

G. P. Jones[1] and P. R. Andrews*[2]

Department of Physical Biochemistry, John Curtin School of Medical Research, Australian National University, Canberra City, A.C.T. 2601, Australia. *Received September 19, 1979*

Conformational energy calculations are reported for a series of convulsant and anticonvulsant barbiturates derived from 5-ethyl-5-(1'-methylbutyl)barbituric acid (pentobarbital) by minor structural changes to the butyl side chain. A number of low-energy conformations are identified for each barbiturate. In each case substantial barriers to rotation exist between the alternative conformations, and the magnitudes of these barriers suggest that the barbiturates may be conformationally restricted even at physiological temperatures. Fully extended conformations, with both side chains perpendicular to the plane of the barbiturate ring, are favored. In the 1'-methyl derivatives, conformations with the 1'-methyl group located directly above the barbiturate ring are equally low in energy.

Minor structural changes to barbiturates and related drugs frequently result in dramatic switches between convulsant and anticonvulsant activity,[3-6] but no structural explanation for these differences is yet available. In an effort to obtain definitive structure–activity relationships for the convulsant and anticonvulsant barbiturates, we have undertaken a theoretical and experimental study of

(1) Address: Waite Institute, University of Adelaide, Glen Osmond, S.A. 5064, Australia.
(2) Corresponding address: Victorian College of Pharmacy, Parkville, Victoria, 3052, Australia.
(3) H. Downes, R. S. Perry, R. E. Ostlund, and R. Karler, *J. Pharmacol. Exp. Ther.*, 175, 692 (1970).
(4) P. K. Knoefel, *J. Pharmacol. Exp. Ther.*, 84, 26 (1945).
(5) L. Velluz, J. Mathieu, and R. Jequier, *Ann. Pharm. Fr.*, 9, 271 (1951).
(6) P. R. Andrews, G. P. Jones, and D. Lodge, *Eur. J. Pharmacol.*, 55, 115 (1979).

0022-2623/80/1823-0444$01.00/0