Structure-Activity Relationships of Antifilarial Antimycin Analogues: A Multivariate Pattern Recognition Study

David L. Selwood,* David J. Livingstone, John C. W. Comley, Alan B. O'Dowd, Alan T. Hudson, Peter Jackson, Karamjit S. Jandu, Valerie S. Rose, and Jeremy N. Stables

Wellcome Research Laboratories, Park Langley, Beckenham, Kent, U.K. *Received February 6, 1989*

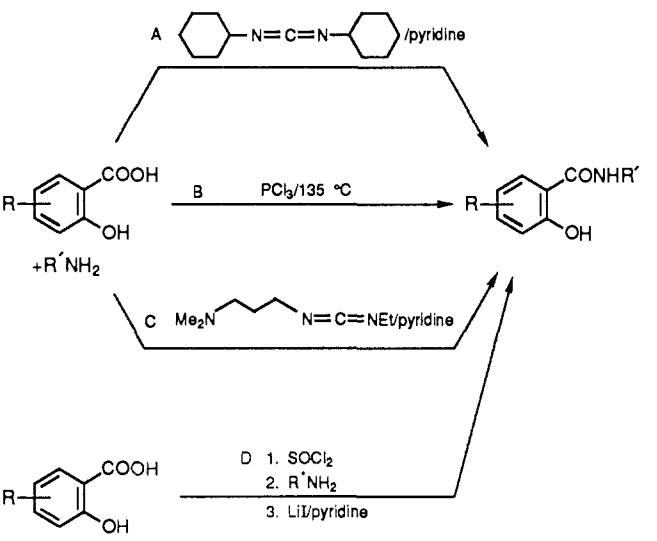
The structure–activity relationships of a series of novel antifilarial antimycin A_1 analogues have been investigated by using computational chemistry and multivariate statistical techniques. The physiochemical descriptors calculated in this way contained information which was useful in the classification of compounds according to their in vitro antifilarial activity. This approach generated a 53 parameter descriptor set, which was reduced with a multivariate pattern recognition package, ARTHUR. Regression analysis of the reduced set yielded several statistically significant regression equations; e.g.-log in vitro activity = 0.017 mp + 0.65 log P - 0.81ESDL$_{10}$ - 7.33 (R = 0.9). With use of this equation, it was possible to make predictions for further untested analogues. The analysis indicated that membrane or lipid solubility is an important determinant in biological activity agreeing with the proposed primary mode of action of the compounds as disrupters of cuticular glucose uptake.

Filariae are tissue-dwelling nematodes, infecting many millions of people worldwide and causing a number of debilitating illnesses. The two most significant diseases in man are onchocerciasis (river blindness), caused by the parasite *Onchocerca volvulus*, and lymphatic filariasis (elephantiasis), casued by the parasites *Brugia malayi* and *Wucheria bancrofti*. In the course of our search for new antifilarial drugs,[1] we have developed a series of analogues of the mammalian electron transport inhibitor antimycin A_1 as potential antifilarials. This paper describes the elucidation of the structure–activity relationships for this series by the application of computer chemical and multivariate statistical methods.

Antimycin A_1 is a potent inhibitor of mammalian electron-transport systems.[2] We have previously shown in these laboratories that several filarial species possess functional electron-transport systems[1] and that the parasites are not homolactate fermenters as previously thought. Antimycin A_1 has moderate activity against adult filarial *Dipetalonema viteae* in vitro (see Table I) but high mammalian toxicity.[2] Our objectives were therefore to develop analogues with greater potency than antimycin but with reduced mammalian toxicity. Though several compounds were synthesized which satisfied these criteria, the structure–activity relationships for the series were unclear; we therefore undertook a multivariate pattern recognition study in order to clarify the situation.

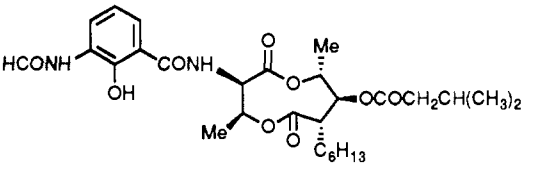
The approach consisted of a combination of the techniques of computational chemistry, in order to generate molecular descriptors, and pattern recognition as a means of analyzing the data.[3] The characterization of chemical structure in terms of calculated properties and "traditional" QSAR descriptors[4] gave rise to a data set containing a large number of descriptors. Analysis of this data, at least in the first instance, is not possible by techniques such as multiple linear regression and thus it was necessary to use other multivariate methods to reduce the dimensionality of the data.[5]

Scheme I

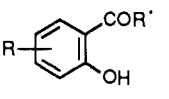


Chemistry

Antimycin A_1 is an amide of 3-formamido-2-hydroxysalicylic acid with the amido portion being a complex dilactone moiety.



antimycin A_1



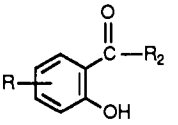
i

We chose to replace the dilactone moiety with simpler lipophilic groups and to vary the substituents on the aromatic ring, proposing analogues with the general structure i. A survey of the literature revealed little work on antimycin analogues other than some *n*-alkyl amides of 3-formamido-2-hydroxybenzoic acid synthesized to examine their mammalian electron-transport activity.[6,7]

(1) See papers by Comley, J. C. W.; O'Dowd, A. B.; Hayes, D.; Selwood, D. L.; Stables, J. N. Proceedings from the Wellcome Foundation Filariasis Seminar 1987. *Trop. Med. Parasitol.* 1988, *39*, 67.
(2) von Jagow, G.; Link, T. A. Use of Specific Inhibitors of the Mitochondrial bc_1 Complex. *Methods in Enzymology*; Academic Press, Inc.: New York, 1986; Vol. 126, p 253.
(3) Hyde, R. M.; Livingstone, D. J. *J. Comput. Aided Mol. Design* 1988, *2*, 145.
(4) Glen, R. C.; Rose, V. S. *J. Mol. Graphics* 1987, *2*, 79.
(5) Sharaf, M. A.; Illman, D. A.; Kowalski, B. R. *Chemometrics*; Wiley: New York, 1986, p 179.
(6) Dickie, P.; Loomans, M. E.; Farley, T. M.; Strong, F. M. *J. Med. Chem.* 1963, *6*, 424.

0022-2623/90/1833-0136$02.50/0

Table I. Structures and Physical Properties of Compounds 1–31



	R_1	R_2	mp, °C	formula	analysis	method of prepn	recryst solvent	% yield
1	3-NHCHO	$NHC_{14}H_{29}$	81–83	$C_{22}H_{36}N_2O_3$	C, H, N	A, E	ethanol/water	38
2	3-NHCHO	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	182–184	$C_{20}H_{14}Cl_2N_2O_4$	C, H, N	B, E	acetic acid	38
3	5-NO_2	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	205–207	$C_{19}H_{12}Cl_2N_2O_5$	C, H, N	D	acetic acid/water	38
4	5-SCH_3	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	143–145	$C_{20}H_{15}Cl_2NO_3S$	C, H, N, Cl, S	B	ethanol	27
5	5-$SOCH_3$	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	165–168	$C_{20}H_{15}Cl_2NO_4S$	C, H, N	B, G	hexane/ethyl acetate	43
6	3-NO_2	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	142–146	$C_{19}H_{12}Cl_2N_2O_5$	C, H, N	C	acetic acid	51
7	5-CN	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	256–258	$C_{20}H_{12}Cl_2N_2O_3 \cdot 1.13H_2O$	C, H, N	B	acetic acid[a]	9
8	5-NO_2	NH-4-(4-$CF_3C_6H_4O$)C_6H_4	199–202	$C_{20}H_{13}F_3N_2O_5$	C, H, N	D	acetic acid	46
9	3-SCH_3	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	151–153	$C_{20}H_{15}Cl_2NO_3S \cdot 0.24H_2O$	C, H, N	B	acetic acid/water	12
10	5-SO_2CH_3	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	195–196	$C_{20}H_{15}Cl_2NO_5S$	C, H, N, S	B	ethanol/water	95
11	5-NO_2	NH-4-(C_6H_5O)C_6H_4	212–215	$C_{20}H_{14}N_2O_5$	C, H, N	B	acetic acid	5
12	5-NO_2	NH-3-Cl-4-(4-ClC_6H_4CO)C_6H_3	246–248	$C_{20}H_{12}Cl_2N_2O_5$	C, H, N	B	acetic acid	35
13	5-NO_2	NH-4-(2-Cl-4-$NO_2C_6H_3O$)C_6H_4	208–211	$C_{19}H_{12}ClN_3O_7$	C, H, N	B	acetic acid[a]	22
14	5-NO_2	NH-3-Cl-4-(4-$CH_3OC_6H_4O$)C_6H_3	159–161	$C_{20}H_{15}ClN_2O_6$	C, H, N	B	acetic acid/water	41
15	3-SO_2CH_3	NH-3-Cl-4-(4-ClC_6H_4O)C_6H_3	178–180	$C_{20}H_{15}Cl_2NO_5S$	C, H, N	B, F	ethanol/water	65
16	5-NO_2	NH-3-Cl-4-(4-ClC_6H_4S)C_6H_3	203–206	$C_{19}H_{12}Cl_2N_2O_4S$	C, H, N	B	acetic acid[a]	28
17	3-NHCHO	NHC_6H_{13}	62–63	$C_{14}H_{20}N_2O_3$	C, H, N	A, E	ether/pentane	60
18	3-NHCHO	NHC_8H_{17}	78–80[b]	$C_{16}H_{24}N_2O_3$	C, H, N	A, E	ethanol/water	38
19	3-$NHCOCH_3$	$NHC_{14}H_{29}$	71–72	$C_{23}H_{38}N_2O_3$	C, H, N	A	ethanol/water	25
20	5-NO_2	$NHC_{14}H_{29}$	90–92	$C_{21}H_{34}N_2O_4$	C, H, N	A	acetic acid	44
21	3-NO_2	$NHC_{14}H_{29}$	67–68	$C_{21}H_{34}N_2O_4$	C, H, N	A	acetic acid	49
22	3-NO_2-5-Cl	$NHC_{14}H_{29}$	81–83	$C_{21}H_{23}ClN_2O_4$	C, H, N	A	acetic acid	61
23	5-NO_2	NH-4-$C(CH_3)_3C_6H_4$	227–229	$C_{17}H_{18}N_2O_4$	C, H, N	C	acetic acid	51
24	5-NO_2	$NHC_{12}H_{25}$	85–87	$C_{19}H_{30}N_2O_4$	C, H, N	B	acetic acid[a]/water	33
25	3-NO_2	$NHC_{16}H_{33}$	79–80[c]	$C_{23}H_{38}N_2O_4$	C, H, N	A	acetic acid/water	55
26	5-NO_2	NH-3-Cl-4-(4-ClC_6H_4NH)C_6H_3	173–175	$C_{19}H_{13}Cl_2N_3O_4$	C, H, N	A	acetic acid[a]/water	10
27	5-NO_2	NH-4-(3-$CF_3C_6H_4O$)C_6H_4	176–178	$C_{20}H_{13}F_3N_2O_5$	C, H, N	D	ethanol/water	43
28	5-NO_2	NH-3-Cl-4-(4-$SCF_3C_6H_4O$)C_6H_3	195–197	$C_{20}H_{12}F_3ClN_2O_5S$	C, H, N	D	ethanol/water	52
29	5-NO_2	NH-3-Cl-4-(3-$CF_3C_6H_4O$)C_6H_3	192–194	$C_{20}H_{12}F_3ClN_2O_5$	C, H, N	B	ethanol	43
30	5-NO_2	NH-4-(C_6H_5CHOH)C_6H_4	178–180	$C_{20}H_{16}N_2O_5$	C, H, N	D	ethanol/water	57
31	5-NO_2	4-ClC_6H_4	170–172	$C_{13}H_8ClNO_4$	C, H, N	D	ethanol	23

[a] Recrystallized twice. [b] Lit.[6] 77–78 °C. [c] Lit.[6] 84.5–85.5 °C. [d] See method in ref 22.

Most of the analogues were simple amides of 2-hydroxybenzoic acids and were prepared by known methods (Scheme I, methods A, B, and C). Method D in Scheme I was developed to facilitate the large-scale synthesis of analogues. The structures, method of preparation, and yields obtained are given in Table I.

Results and Discussion

At the beginning of this study, biological activity data were available for a small set of compounds, 1–16 in Table II. This set may be referred to as the "training set" in that the data for these compounds are used to "train" the analytical methods so as to predict biological activity. The remaining 15 compounds in Table II form a "test" set with biological data being obtained subsequent to the analyses carried out on the training set.

Table II contains data for both in vitro and in vivo biological tests (see the Experimental Section). Several compounds exhibit potent in vitro activity, notably those of the 2-hydroxy-5-nitro-4-phenoxybenzanilide type; e.g. **3**, **8**, **28**, and a number of others can be considered active in vivo (>70% worm reduction).

The first stage in the analysis of structure–activity relationships for these compounds was to model them with computer graphics and calculate physicochemical descriptors from the models. This is illustrated in Figure 1 with the computer programs used for the various stages identified. Some of the properties calculated were atom specific; Figure 2 shows a plot of compound **3** with the atom numbering as used for the whole series. This figure also shows the orientation which was used for the series, and thus the *X*, *Y*, and *Z* axes may be defined. *X* and *Y* directions are the horizontal and vertical axes of the figure, respectively, with the *Z* direction being orthogonal to the plane of the paper.

A total of 53 parameters were selected for the starting data set as shown in Table III. This data set, of course, contains considerable redundancy and so the next step was to investigate the pairwise interparameter correlation matrix in order to identify features which could be removed.

This task could, in principle, have been carried out manually, but since there are 1378 pairs of correlations in a 53 parameter data set, it was more efficiently achieved with a computer routine. An in-house RS/1 procedure, CORCHOP (see the Experimental Section), enabled the reduction of this data set to just 23 parameters with no interparameter correlations higher than 0.75. The resultant correlation matrix is not shown here but contains only 12 correlations of greater than 0.6 out of a total of 506. A list of the decorrelated parameter set is shown in Table IV. One feature of this procedure is that the variables which are retained may be regarded as "nominated" var-

(7) Farley, T. M.; Strong, F. M.; Bydalek, T. J. *J. Am. Chem. Soc.* **1965**, *87*, 3501.
(8) Hudson, B.; Livingstone, D. J.; Rahr, E. *J. Comput. Aided Mol. Design* In press.

Table II. Biological Activity for Compounds 1–31

compd	in vitro activity: EC_{50},[a] μM	−log in vitro activity	−log activity predicted from eq 3	in vivo activity: % worm reduction[b]
1	7.0	−0.85	−0.90	89
2	2.4	−0.38	0.06	toxic
3	0.04	1.40	1.14	80.6
4	0.48	0.32	0.26	58
5	7.5	−0.88	−0.48	84
6	0.15	0.82	0.77	toxic
7	0.0145	1.84	1.73	17
8	0.095	1.02	0.87	72
9	0.38	0.42	0.42	52
10	1.0	0.00	−0.03	28
11	0.8	0.10	0.11	89
12	0.074	1.13	1.23	14
13	0.12	0.92	0.53	70
14	0.17	0.77	0.62	61
15	0.5	0.30	−0.31	42
16	0.044	1.36	1.28	70
17	>10	−1.0	−3.99	NT
18	2.6	−0.41	−3.03	NT
19	8.0	−0.9	−1.07	NT
20	0.128	0.89	0.01	NT
21	0.152	0.82	−0.34	NT
22	0.0435	1.36	0.49	NT
23	0.59	0.23	0.29	44
24	0.039	1.41	−0.77	0
25	1.1	−0.04	0.51	NT
26	0.37	0.43	1.35	48
27	0.094	1.03	0.51	80
28	0.028	1.55	1.47	85; some toxicity
29	0.085	1.07	1.03	74
30	>10	−1.0	−1.62	39
31	0.33	0.48	0.85	61

[a] EC_{50} = effective concentration at which 50% of the adenine taken up was released into the medium. [b] NT = not tested.

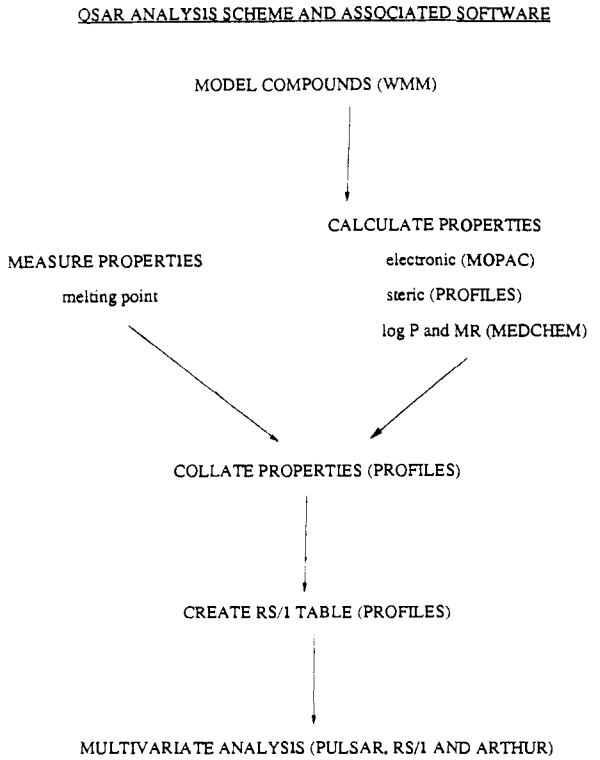


Figure 1. Data generation and analysis protocol.

iables which represent the information content of the sets of correlated variables from which they were chosen. It is important to bear this in mind when seeking a physicochemical explanation for any correlation.

This data set was still too broad to analyze by a method such as multiple regression, but there are techniques which

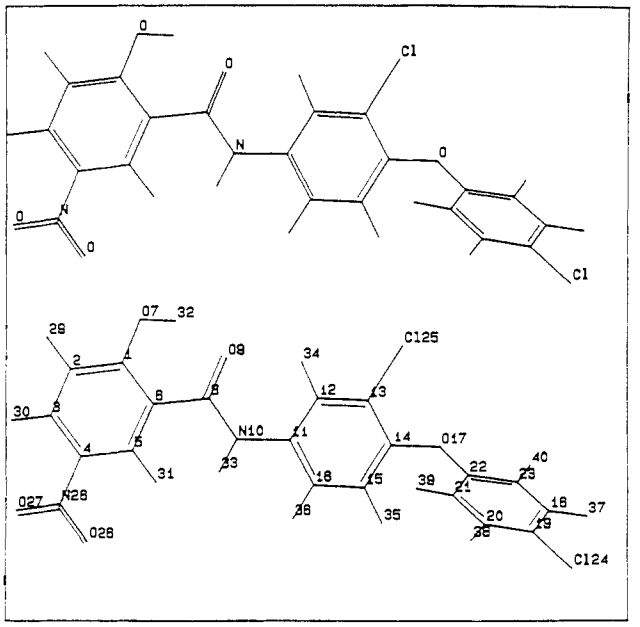


Figure 2. Atom numbering and orientation for the series.

Table III. Initial Physicochemical Descriptor Set

partial atomic charges (ATCH) for atoms 1–10
electrophilic (ESDL) and nucleophilic (NSDL) superdelocalizabilities for atoms 1–10
dipole moment and vectors (X, Y, and Z) of the dipole moment
van der Waal's volume, surface area, molecular weight, moments of inertia (X, Y, and Z) and principal ellipsoid axes (1, 2, and 3)
three parameters describing substituent dimensions in the X, Y, and Z directions and the coordinates of the center of the substituent
calculated log P, melting point, and the sums of the F and R substituent constants

Table IV. Decorrelated Parameter Set (23 Variables)

Electronic
three atom charges (atoms 2–4)
dipole moment and the x, y, and z components
two electrophilic superdelocalizabilities (atoms 5 and 10)
two nucleophilic superdelocalizabilities (atoms 2 and 10)
Bulk
van der Waal's volume
molecular weight
moment of inertia (in the X axis)
principal ellipsoid axis (in the Y axis)
Substituent
substituent length (Y and Z axes)
coordinates of the center of the substituent (X and Z coordinates)
sums of the substituent F and R values
Other
melting point
log P

may be used to investigate a data matrix such as this.[3,5]

The training set of 16 compounds described by 23 parameters may be regarded as 16 points in a 23-dimensional space, with the distances between points in this space (the Euclidian distance) giving an indication of the chemical similarity between compounds (in terms of all 23 descriptors). The implicit assumption in all structure–activity studies is that the more similar two compounds are with respect to their chemical and physical properties the more similar they will be in terms of their biological properties. Thus, methods which allow a two-dimensional

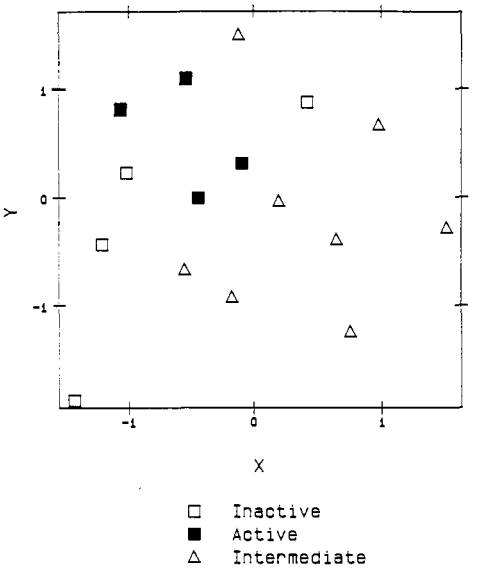


Figure 3. Nonlinear map of the training set based on 23 physicochemical properties.

representation of a higher dimensional space,[9] such as 23 dimensions, might be expected to reveal patterns in the chemical data which match patterns in the biological data.

Figure 3 shows the result of the application of one such technique, nonlinear mapping,[8,9] to this data set. A nonlinear map is generated in such a way that the distances between points (compounds) in the two-dimensional plot is a good representation of the distances between the points in the higher dimensional space. In this figure it can be seen that the four most active compounds are clustered together, thus indicating that the descriptor set of 23 properties contains some information which is "useful" in classifying compounds as active. It can also be seen from this figure that most of the compounds with intermediate activity fall into one region of the plot and most of the inactives fall into another area. Such a plot can be used to make qualitative predictions of activity.

The next stage in the analysis of this data set was to identify those properties which are most successful in terms of explanation of the biological data. In order to achieve this, a feature selection routine (SELECT), which is part of the ARTHUR[10] package, was employed. The SELECT method[11] first of all identifies the feature which is best correlated with the data of interest, i.e. the biological data. The remaining properties are decorrelated from this first selected feature, and the next best descriptor, in terms of correlation to biological data, is selected. This procedure continues until all the features have been ordered or until some lower limit of correlation (selected by the user) has been achieved.

A list of the first 10 variables chosen by SELECT is shown in Table V. It should be recognized that this selection procedure is equivalent to a forward-stepping regression analysis and there is thus a danger that variables may be identified as important due to chance correlations. Topliss,[12] for example, has studied chance correlations using data sets composed of random numbers. While it is desirable to limit the probability of chance correlations occurring, the main danger of these artefacts is the failure to recognize them. Both chemical synthesis and biological testing are resource demanding, and unsuccessful predictions due to chance effects can prove expensive. The artificial device of dividing a known data set into training and test sets is one means by which such mistakes may be avoided, as are the techniques of cross-validation[13] and bootstrapping.[14]

Table V. The First 10 Variables Chosen by SELECT from the Decorrelated Data Set of 23 Descriptors

order	variable
1	melting point
2	log P
3	$ESDL_{10}$
4	DIPV-Z (*Z* dipole vector)
5	$ESDL_5$
6	sum of *R*
7	S8-1CZ (*Z* coordinate of substituent center)
8	DIPV-Y (*Y* dipole vector)
9	$NSDL_2$[a]
10	$ATCH_2$[b]

[a] $NSDL_2$ = nucleophilic superdelocalizaability for atom 2. [b] $ATCH_2$ = partial atomic charge for atom 2.

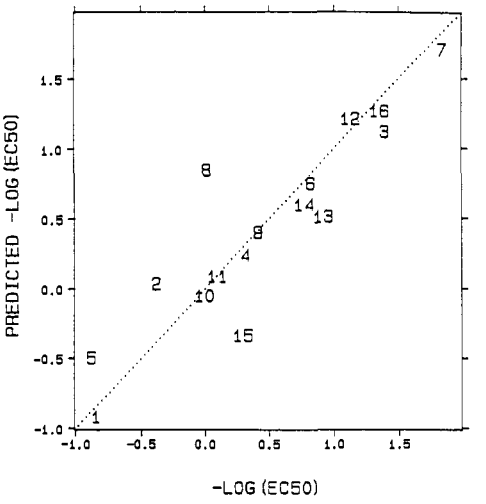


Figure 4. Plot of -log EC_{50} values predicted by eq 3 vs observed values for the training set.

With use of the variables chosen by SELECT, regression equations were calculated for the training set as shown below:

$$-\log EC_{50} = \underset{(3.67)}{0.013}mp - \underset{(2.91)}{1.97} \quad (1)$$

$$R = 0.70 \quad F = 13.55 \quad SE = 0.58$$

$$-\log EC_{50} = \underset{(5.72)}{0.016}mp + \underset{(3.46)}{0.56}\log P - \underset{(4.70)}{6.14} \quad (2)$$

$$R = 0.86 \quad F = 18.0 \quad SE = 0.45$$

$$-\log EC_{50} = \underset{(6.87)}{0.017}mp + \underset{(4.44)}{0.65}\log P - \underset{(2.26)}{0.81}ESDL_{10} - \underset{(5.84)}{7.33} \quad (3)$$

$$R = 0.90 \quad F = 17.5 \quad SE = 0.38$$

where mp = melting point, log *P* = log of the calculated partition coefficient, EC_{50} = concentration of drug (μM) required for reduction of adenine release by 50%, and $ESDL_{10}$ = electrophilic superdelocalizability at atom 10.

Values of *t* statistics are shown in parentheses for the regression coefficients; $n = 16$ for all equations. While the standard errors are quite large, it should be recognized that the activity data spans nearly 3 orders of magnitude for the training set. Figure 4 shows a plot of predicted versus observed -log EC_{50} values for the training set based on eq

(9) Kowalski, B. R.; Bender, C. F. *J. Am. Chem. Soc.* **1973**, *95*, 686.
(10) ARTHUR; infometrix Inc., Seattle, WA 98121.
(11) Kowalski, B. R.; Bender, C. F. *Pattern Recognit.* **1976**, *8*, 1.
(12) Topliss, J. G.; Edwards, R. P. *J. Med. Chem.* **1979**, *22*, 1238.
(13) Geisser, S. *J. Am. Stat. Assoc.* **1975**, *70*, 328.
(14) Diaconis, P.; Efron, B. *Sci. Am.* **1984**, 116.

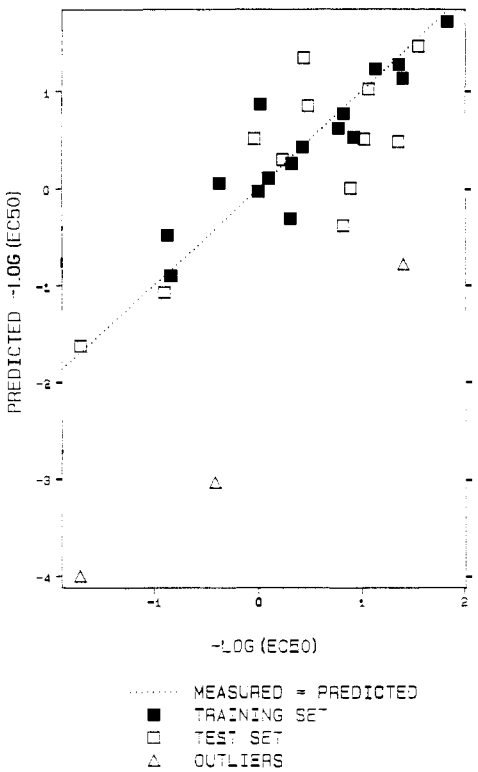


Figure 5. Plot of -log EC_{50} values predicted by eq 3 vs observed values for the test and training sets.

3. It can be seen that this equation performs well in these training set predictions. Use of this equation for test set predictions is shown in Table II and in Figure 5, with the training set compounds indicated by filled squares. As might be expected from inspection of this figure, the regression equation for the combined set gives a much poorer fit as shown below:

$$-\log EC_{50} = \underset{(4.14)}{0.008mp} + \underset{(5.54)}{0.44 \log P} - \underset{(1.33)}{0.30ESDL_{10}} - \underset{(5.65)}{3.93} \quad (4)$$

$$R = 0.77 \quad F = 13.29 \quad SE = 0.61 \quad n = 31$$

It can be seen from this equation that not only is the term in $ESDL_{10}$ no longer significant, as shown by its t statistic, but also that the regression coefficient for melting point has a much different value to the previous coefficients (see eq 1–3). This deterioration in fit is mostly caused by compounds **17**, **18**, and to a lesser extent, **24**, which are shown on Figure 5 as triangles. Chemically, these compounds are distinct from the bulk of the training set in that they have an n-alkyl side chain as opposed to a side chain of the phenoxy ether type. The regression equations are therefore not suitable for prediction of activity for compounds with radically different structure from the training set. Omission of these three compounds results in an equation which is very similar to eq 3 as shown below:

$$-\log EC_{50} = \underset{(7.15)}{0.014mp} + \underset{(7.77)}{0.68 \log P} - \underset{(3.08)}{0.54ESDL_{10}} - \underset{(7.79)}{6.83} \quad (5)$$

$$R = 0.86 \quad F = 23.3 \quad SE = 0.45 \quad n = 28.$$

Thus, with the exception of these three compounds, it was possible to make quite reasonable predictions of EC_{50} values for the test set from the equations derived from the training set. It is interesting to note here that despite the generation of a large number of calculated steric and electronic properties, only one of these has been found to be useful in the prediction of EC_{50} values. The large data matrix originally generated may contain useful chemical information which is simply not related to this particular biological activity. On the other hand, it is possible that the starting data matrix contains little extra information over and above that given by log P and melting point. The fact that this matrix was readily reduced from 53 parameters to just 23 certainly lends support to the notion that this data contains a much smaller amount of "real" information than its size might suggest.

The importance of partition coefficient and melting point in these regression equations is suggestive that the in vitro activity is related to lipid or membrane solubility. Interestingly, Valvani and Yalkowsky[15] have shown that for several different series of compounds aqueous solubility (log S_w^c) can be predicted well by equations of the form shown below:

$$\log S_w^c = -\log P - 0.01mp + C \quad (6)$$

where C is a constant and regression coefficients for log P fall in the range 0.88–1.1. The log P and melting point terms of eq 3 and 5 are of this form but of opposite sign.

Although it was possible to make reasonable predictions of activity for the test set using eq 3, it will be noticed that there are a number of relatively inactive compounds which were subsequently made and tested. The reason for this is that in vivo activity is not simply related to in vitro activity. Some of the test set compounds, therefore, were synthesized in order to probe the in vivo system even though the expected in vitro activity might be low.

Conclusions. The computer chemistry methods described have enabled the QSAR's for this series to be determined. Although many parameters were examined, only three, log P, melting point, and superdelocalizability, were shown to be important in determining activity in vitro. The regression equations obtained were useful for the prediction of activity of compounds closely related to the training set but not for compounds of rather different chemical structure. The dependence of in vitro activity on high log P values is consistent with our subsequent studies on the mode of action of this series. These studies[16] showed that although active compounds were inhibitors of filarial electron transport, their primary effects were most likely as membrane disruptors, leading to a nonspecific inhibition of glucose uptake.

Experimental Section

Chemistry. Melting points were determined on an Electrothermal melting point apparatus and are uncorrected. "Flash" chromatography was performed according to the method of Still.[17]

Noncommercially available 2-hydroxybenzoic acids were either prepared by literature methods or by application of standard methods. Thus 2-hydroxy-3-nitrobenzoic acid was prepared by the method of Meldrum.[18] 2-Hydroxy-3-(methylthio)- and 2-hydroxy-5-(methylthio)benzoic acids were prepared by carboxylation of the phenols.[19] 5-Cyano-2-hydroxybenzoic acid was prepared by the method of Jones.[20] 5-Chloro-2-hydroxy-3-

(15) Valvani, S. C.; Yalkowsky, S. H., Yalkowsky, S. H.; Sinkula, A. A., Valvani, S. C., Eds.; *Physical Chemical Properties of Drugs*; Medicinal Research Series, Marcel Dekker, Inc.: New York and Basel, 1980; Vol. 10, p 201.
(16) O'Dowd, A. B.; Hayes, D. J.; Balsara, R. R.; Selwood, D. L. *Trop. Med. Parasitol.* **1988**, *39*, 79. See also: Comley, J. C. W.; O'Dowd, A. B.; Hudson, A. T.; Jackson, P.; Jandu, K. S.; Livingstone, D.; Rose, S.; Selwood, D. L.; Stables, J. N.; *Trop. Med. Parasitol.* **1989**, *40*, 87.
(17) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923.
(18) Meldrum, A. N.; Hirve, N. W. *J. Indian Chem. Soc.* **1928**, *5*, 95.
(19) Ford, R. E.; Knowles, P.; Lunt, E.; Marshall, S. M.; Summers, A. J. H. British patent GB 1,561,350, 1976.

nitrobenzoic acid was prepared by nitration of 5-chloro-2-hydroxybenzoic acid.

Substituted amines were either commercially available or prepared by literature methods.[21-24] In most cases the compounds were stored in the nitro form and reduced directly before use in the amide preparation.

Method A. To the appropriate substituted 2-hydroxybenzoic acid (10 mmol) in dry pyridine (100 mL) was added the required amine (10 mmol) and dicyclohexylcarbodiimide (12 mmol). This mixture was heated at 90 °C for 2 h. On cooling, the dicyclohexylurea was filtered off. The pyridine was removed on a rotary evaporator. The residue was treated with 10% hydrochloric acid (100 mL) and stirred for half an hour. The precipitate was filtered, washed with water (2 × 50 mL), and dried in vacuo. If necessary, the crude material was chromatographed. The product was recrystallized from the solvent indicated in Table I.

Method B. To the appropriate, substituted 2-hydroxybenzoic acid (12 mmol) in chlorobenzene (50 mL) was added the required amine (12 mmol) and phosphorus trichloride (7 mmol) plus 2 drops of dimethylformamide. The mixture was refluxed under nitrogen for 2 h. After cooling, ethyl acetate (100 mL) was added. This mixture was washed with saturated sodium bicarbonate (2 × 50 mL) and water (2 × 50 mL). The organic layer was dried ($MgSO_4$), and the solvents were removed on a rotary evaporator. If necessary, the crude material was chromatographed. The product was recrystallized from the solvent indicated in Table I.

Method C was the same as method A except 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide was used instead of dicyclohexylcarbodiimide.

Method D. To the appropriate, substituted 2-methoxybenzoic acid (10 mmol) was added an excess of thionyl chloride and dimethylformamide (1 drop). This mixture was refluxed for 1 h. On cooling, the thionyl chloride was removed on a rotary evaporator. The residue was taken up in dry pyridine (50 mL) and the required amine (10 mmol) in dry pyridine (50 mL) was added dropwise. The temperature was maintained below 40 °C during the addition. After addition, the mixture was heated to 40 °C for 2 h. On cooling, the pyridine was removed on a rotary evaporator. The residue was treated with 1 M hydrochloric acid (50 mL) and stirred for half an hour. The mixture was extracted with ethyl acetate (2 × 50 mL), the extract was washed with water (2 × 25 mL) and dried ($MgSO_4$). The solvent was removed on a rotary evaporator. If necessary, the crude material was chromatographed.

Removal of the Methoxy Group. To the above product (1.7 mmol) in pyridine (50 mL) was added lithium iodide (2.5 mmol). This mixture was refluxed for 2 h. On cooling, the pyridine was removed on a rotary evaporator. The residue was taken up in ethyl acetate (25 mL) and 1 M hydrochloric acid (25 mL) was added. The mixture was warmed; the organic layer was separated and washed with water (2 × 20 mL) and dried. The solvents were removed on a rotary evaporator. If necessary, the crude material was chromatographed. The product was recrystallized from the solvent indicated in Table I.

Method E. The 2-hydroxy-3-nitrobenzamide (10 mmol) was hydrogenated over 10% palladium on carbon (0.1 mmol) in ethanol (50 mL) at atmospheric pressure. This mixture was filtered through Celite, and the solvent was removed on a rotary evaporator to give 3-amino-2-hydroxybenzamide. To a cooled solution of the crude product and dicyclohexylcarbodiimide (11 mmol) in toluene (50 mL) was added dropwise formic acid (11 mmol). This mixture was stirred at room temperature for 1 h. The solvent was removed on a rotary evaporator. The residue was taken up in diethyl ether (50 mL), washed with saturated sodium bicarbonate (2 × 20 mL) and water (2 × 20 mL), and dried ($MgSO_4$). The solvent was removed on a rotary evaporator. If necessary, the crude material was chromatographed. The product was recrystallized as indicated in Table I.

Method F. 3'-Chloro-4'-(4-chlorophenoxy)-2-hydroxy-5-(methylsulfonyl)benzanilide (10). To **4** (420 mg, 1 mmol) in acetone (5 mL) at room temperature was added 30% hydrogen peroxide solution (0.2 mL) followed by ammonium molybdate (0.25 g, 0.2 mmol) in water (0.5 mL). The reaction was stirred for 20 h, water (100 mL) was added, the aqueous mixture was extracted with ethyl acetate (50 mL), and the extract washed with water (10 mL) and dried (Na_2SO_4). The solvent was removed on a rotary evaporator to give a white foam, which was crystallized from ethanol/water to give **10**: 285 mg; 63%; mp 195–196 °C.

Method G. 3'-Chloro-4'-(4-chlorophenoxy)-2-hydroxy-5-(methylsulfinyl)benzanilide (5). To **4** (420 mg, 1 mmol) in dichloromethane (60 mL) at 0 °C was added *m*-chloroperbenzoic acid (172 mg, 1 mmol) in dichloromethane (20 mL). The reaction was stirred at 0 °C for 1 h. Water (100 mL) was added and the reaction was extracted with ethyl acetate (100 mL). The extract was washed with water and dried (Na_2SO_4). The solvent was removed on a rotary evaporator to give a white foam. The crude product was "flash" chromatographed (cyclohexane/ethyl acetate; 3:1) and recrystallized from hexane/ethyl acetate to give **5** (190 mg, 44%), mp 165–168 °C.

Method H. Acylation of 2-Hydroxy-3-nitro-*N*-tetradecylbenzamide. 2-Hydroxy-3-nitro-*N*-tetradecylbenzamide (1.0 g, 2.65 mmol) was hydrogenated over 10% palladium on carbon (11 mg, 0.1 mmol) in ethanol (50 mL) at atmospheric pressure. This mixture was filtered through Celite and the solvent was removed on a rotary evaporator to give 3-amino-2-hydroxy-*N*-tetradecylbenzamide. To the crude product in benzene (20 mL) and pyridine (0.5 mL) was added acetic anhydride (0.297 g, 2.91 mmol), and the mixture was stirred at room temperature for 2 h. To this mixture was added diethyl ether (50 mL). The organic phase was washed with 1 M hydrochloric acid (2 × 20 mL) and water (2 × 20 mL) and dried. The solvent was removed on a rotary evaporator. The crude material was flash chromatographed (cyclohexane/ethyl acetate; 3:1). The product was recrystallized from ethanol/water to yield 260 mg, 25%; mp 71–72 °C.

Parasitology. In Vitro Testing. In vitro macrofilaricidal activity against *D. viteae* females was quantified by using the leakage of incorporated radiolabeled adenine method.[25] In this assay, macrofilaridal viability was determined after 120-h exposure to a range of drug concentrations, descending from 10 μM, in Eagles minimal essential medium (no serum), and the EC_{50} level for adenine leakage was calculated.

In Vivo Testing. Sixty to 100 day old infections of *D. viteae* in hamsters and *B. pahangi* in gerbils were used as sources of adult worms for transplant studies. The parasites were isolated, and the assay was performed as described by Court et al.[26] Briefly, gerbils were anesthetized using halothane and carrier oxygen. *D. viteae* (eight females, seven males) were implanted into a subcutaneous pouch behind the shoulder and *B. pahangi* (eight females, seven males) were implanted into the peritoneal cavity of the same gerbil. Animals were allowed to recover for 5 days and then given 5 daily doses of the test compound either by the subcutaneous or oral route. Animals were autopsied 35 days after the final dose and the worm recoveries were compared with those from untreated controls.

Computational Chemistry and Data Analysis. The compounds were modeled with an in-house molecular modeling package (Wellcome Molecular Modeling) and the geometry was optimized by molecular mechanics with a modified and extended version of White's force field.[27] Steric properties, both whole molecule and substituent, were calculated with an in-house package, PROFILES,[4] and electronic properties were calculated by the quantum mechanics program MOPAC.[28] log *P* values were

(20) Jones, H.; Fordice, M. W.; Greenwald, R. B.; Hannah, J.; Jacobs, A.; Ruyle, W. V.; Walford, G. L.; Shen, T. Y. *J. Med. Chem.* **1978**, *21*, 1100.
(21) Patchett, A. A.; Mrozik, H. H.; Hoff, D. R. U.S. patent US 3,914,418, 1975. See also: Hann, M. *Ger. Offen.* 2311, 638, 1973.
(22) Ullmann, F.; Mallet, E. *Ber. Dtsch. Chem. Ges.* **1898**, *31*, 1694.
(23) Goodyear Tire & Rubber Co. British Patent, 834,510, 1960.
(24) Yagupol'skii, L. M.; Marenets, M. S. *Zh. Obshch. Khim.* **1954**, *24*, 887.
(25) Comley, J. C. W.; Rees, M. J.; O'Dowd, A. B. *Trop. Med. Parasitol.* **1988**, *39*, 221.
(26) Court, J. P.; Stables, J. N.; Lees, G. M.; Martin-Short, M. R.; Rankin, R. *J. Helminthol.* **1988**, *62*, 1.
(27) White, D. N.; Bovill, M. *J. Chem. Soc., Perkin Trans. 2* **1977**, 1610.
(28) Stewart, J. J. P. *QCPE* Program No. 455 (version 3.2) (University of Indiana, Bloomington, IN).

calculated using the MEDCHEM[29] software suite and the substituent constants F and R were obtained from the Hansch and Leo listing.[30]

The calculated properties were collated along with the measured melting point data with PROFILES and a data table created for the data management package RS/1.[31] Data analysis and graphical display were carried out with an in-house package, PULSAR, which is a unified driver system for RS/1; in-house regression routines; and the pattern recognition suite, ARTHUR.[10] This procedure is shown in Figure 1. Initial data reduction was carried out with an RS/1 procedure, CORCHOP,[32] which sorts parameters in terms of the number and magnitude of their correlations with other parameters in the set. This routine suggests variables for deletion from the set; the objective is to produce a smaller data set with interparameter correlations below a set limit (selected by the user) but which retains as much as possible of the "information content" of the set.

(29) CLOGP, Medchem Software Version 3.52; Pomona College Medicinal Chemistry Project: Claremont, California, 1987.
(30) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley-Interscience: New York, 1979.
(31) RS/1, BBN Software Products, U.K. Ltd., Staines, Middlesex, U.K.
(32) Livingstone, D. J.; Rahr, E. *Quant. Struct.-Act. Relat.* In press.

Acknowledgment. We thank G. M. Lees, S. E. McCaig, R. Rankin, and M. J. Rees for technical support and Drs. W. E. Gutteridge, D. C. Jenkins, and J. F. Batchelor for helpful discussions. This study was supported by the World Health Organization Onchocerciasis Chemotherapy Project of the Onchocerciasis Control Program in the Volta River Basin, OCT contract no. 82005.

Supplementary Material Available: Analyses for compounds 1-31 and for intermediates, parameter values (53) for all 31 compounds, and a complete interparameter correlation matrix are included (16 pages). Ordering information is given on any current masthead page.

Synthesis and Structure-Activity Relationships of New 7-[3-(Fluoromethyl)piperazinyl]- and -(Fluorohomopiperazinyl)quinolone Antibacterials

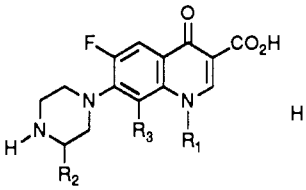
C. B. Ziegler, Jr.,* P. Bitha,* N. A. Kuck, T. J. Fenton, P. J. Petersen, and Y.-i Lin

American Cyanamid Company, Medical Research Division, Lederle Laboratories, Pearl River, New York 10965.
Received February 6, 1989

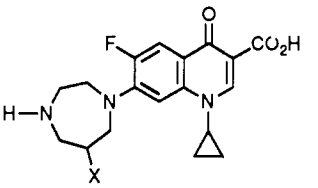
Some novel 6-fluoro-7-substituted-1,4-dihydro-4-oxoquinoline-3-carboxylic acids have been prepared. At the N-1 position "standard" substitution was employed with the ethyl, cyclopropyl, and *p*-fluorophenyl groups being used. At C-7 the introduction of some novel piperazines was made. Most notably, 2-(fluoromethyl)piperazine (**10**) and hexahydro-6-fluoro-1*H*-1,4-diazepine (**16**, fluorohomopiperazine) at the quinolone C-7 position produced products with similar in vitro antibacterial activity as the ciprofloxacin reference. The in vivo efficacy of 1-cyclopropyl-6-fluoro-7-[3-(fluoromethyl)piperazinyl]-1,4-dihydro-4-oxoquinoline-3-carboxylic acid (**20**) was excellent with better oral absorption than ciprofloxacin (**2**).

The emergence of norfloxacin[1] (**1**) as a broad-spectrum, orally active quinolone antibacterial represented a new generation of increased-potency drug.[2] Subsequent members in this structurally similar class employed the C-7 piperazine moiety and some N-1 alkyl groups. Ciprofloxacin (**2**[3]) is an extremely potent N-1 varient. Both compounds are marketed in the U.S., Europe, and Japan.

Among the host of quinolone N-1 substituents synthesized since norfloxacin, the ethyl, cyclopropyl, and fluorophenyl moieties are certainly among the best in terms of antibacterial efficacy. At C-7, the piperazinyl and amino-substituted pyrrolidinyl groups have been utilized with optimal results. As this field evolves, active compounds in the quinolone-3-carboxylic acid class are reported with novel C-5,[4] C-8,[5] and C-2[6] substitution.



1: R_1 = Et; R_2 = H; R_3 = H
2: R_1 = cyclopropyl; R_2 = H; R_3 = H
3: R_1 = Et; R_2 = CH_3; R_3 = F
4: R_1 = 2,4-difluorophenyl; R_2 = CH_3; R_3 = H
5: R_1 = Et, cyclopropyl, or 4-fluorophenyl; R_2 = CH_2OH, CH_2F, or CHF_2; R_3 = H or F



6: X = H, OH, or F

We have been active in preparing quinolone C-7 structural analogues.[7] Particularly, we have invested consid-

(1) Synthesis: Koga, H.; Itoh, A.; Murayama, S.; Suzue, S.; Irkura, T. *J. Med. Chem.* **1980**, *23*, 1358. Review of antibacterial, pharmacokinetic, and therapeutic properties: Holmes, B.; Brogden, R.; Richards, D. *Drugs* **1985**, *30*, 482.
(2) Nalidixic Acid (1-ethyl-1,4-dihydro-7-methyl-4-oxo-1,8-naphthyridine-3-carboxylic acid) was the first member of this totally synthetic antibacterial class. For a comprehensive review of this and all other earlier structurally similar analogues, see: Albrecht, R. *Prog. Drug Res.* **1977**, *21*, 9.
(3) Wise, R.; Andrews, J. M.; Edwards, L. *J. Antimicrob. Agents Chemother.* **1983**, *23*, 559. Felmingham, D.; O'Hare, M. D.; Robbins, R. A.; Williams, A. H.; Cremer, A. W.; Ridgeway, G. L.; Gruneberg, R. N. *Drugs Exp. Clin. Res.* **1985**, *11*, 316.
(4) Domagala, J. M.; Hagen, S. E.; Heifetz, C. L.; Hutt, M. P.; Mich, T. F.; Sanchez, J. P.; Trehan, A. K. *J. Med. Chem.* **1988**, *31*, 506. Moran, D. B.; Ziegler, C. B., Jr.; Dunne, T. S.; Kuck, N. A.; Lin, Y.-i *J. Med. Chem.* **1989**, *32*, 1313.
(5) Sanchez, J. P.; Domagala, J. M.; Hagen, S. E.; Heifetz, C. L.; Hutt, M. P.; Nichols, J. B.; Trehan, A. K. *J. Med. Chem.* **1988**, *31*, 983. Veda, H.; Miyamoto, H. (Otsuka Pharm. Co., Ltd.) *Drugs Future* **1988**, *10*, 412.
(6) Chu, D. T. W.; Fernandes, P. B.; Pernet, A. G. *J. Med. Chem.* **1986**, *29*, 1531. Chu, D. T. W.; Fernandes, A. K.; Claiborne, L. S.; Pernet, A. G. 27th ICAAC, **1987**, *250*.

0022-2623/90/1833-0142$02.50/0