of 124, mp 146–148 °C (lit. mp 140–141 °C[17]). 124 was recrystallized from water with no improvement of melting point.

References and Notes

(1) (a) Presented in part at the 175th American Chemical Society National Meeting, Anaheim, CA, Mar 13–17, 1977. (b) Current address: Hercules, Inc., Central Research Department, Wilmington, DE.
(2) (a) J. S. G. Cox, *Nature (London)*, **216**, 1328 (1967); H. Cairns, C. Fitzmaurice, D. Hunter, P. B. Johnson, J. King, T. B. Lee, G. H. Lord, R. Minshull, and J. S. G. Cox, *J. Med. Chem.*, **15**, 583 (1972). (b) For a recent review of antiallergic compounds, see A. L. Oronsky and J. W. F. Wasky, *Annu. Rep. Med. Chem.*, **12**, 70–79 (1977).
(3) D. R. Buckle, B. C. C. Cantello, H. Smith, and B. A. Spicer, *J. Med. Chem.*, **20**, 265 (1977).
(4) J. H. Sellstedt, C. J. Guinosso, A. J. Begany, S. C. Bell, and M. Rosenthale, *J. Med. Chem.*, **18**, 926 (1975).
(5) T. H. Althius, S. B. Kadin, L. J. Czuba, P. F. Moore, and H. J. Hess, 175th National Meeting of the American Chemical Society, Anaheim, CA, Mar 1978, Abstract 2, Medicinal Chemistry Division.
(6) C. R. Willis, K. M. Snader, C. K. Miao, W. L. Mendelson, and L. W. Chakrin, 10th Middle Atlantic Regional Meeting of the American Chemical Society, Philadelphia, Pa., Feb 1976, Abstract J 8.
(7) J. A. Lansing and J. G. White, *J. Am. Chem. Soc.*, **92**, 7187 (1970).
(8) (a) A. K. Kiang, S. F. Tan, and W. S. Wong, *J. Chem. Soc. C*, 2721 (1971); (b) C. R. Willis, W. L. Mendelson, C. K. Miao, and K. M. Snader, to be published.
(9) T. R. Bosin, R. N. Hanson, J. V. Rodricks, R. A. Simpson, and H. Rapoport, *J. Org. Chem.*, **38**, 1591 (1973).
(10) F. L. Scott, D. G. O'Donovan, and J. Reilly, *J. Am. Chem. Soc.*, **75**, 4053 (1953).
(11) W. W. Hartman and O. E. Sheppard, "Organic Syntheses", Collect. Vol. II, Wiley, New York, 1943, p 617.
(12) J. Goose and A. M. J. N. Blair, *Immunology*, **16**, 749 (1969).
(13) L. W. Chakrin, R. D. Krell, J. Mengel, D. Young, C. Zaher, and J. R. Wardell, Jr., *Agents Actions*, **4**(5), 297 (1974).
(14) R. H. Wiley, C. H. Jarboe, H. G. Ellert, H. Krous, E. J. Hittinger, C. L. deSilva, and S. C. Slaymaker, *J. Org. Chem.*, **21**, 686 (1956).
(15) K. Dickore and E. Kuhle, *Angew. Chem., Int. Ed. Engl.*, **4**, 430 (1965).
(16) E. Cohen and B. Klarberg, *J. Am. Chem. Soc.*, **84**, 1994 (1962).
(17) Flurscheim, *J. Prakt. Chem.*, **71**, 538 (1905).

Application of Quantitative Structure–Activity Relationships in the Development of the Antiallergic Pyranenamines[1]

Richard D. Cramer III,* Kenneth M. Snader, Chester R. Willis, Lawrence W. Chakrin, Jean Thomas, and Blaine M. Sutton

Research and Development Division, Smith Kline & French Laboratories, Philadelphia, Pennsylvania 19101.

Received September 21, 1978

QSAR techniques played a major role in development of the antiallergic pyranenamines (I). Graphical analysis of data resulting from an unsuccessful Topliss approach suggested that increased substituent hydrophilicity might enhance potency. The 3-NHAc-4-OH derivative which first resulted was an order of magnitude more potent than any preceding series member, and its deacylated congener is clinical candidate SK&F 78729 (R_1 = $-NH_2$, R_2 = OH, R_3 = H). Further pursuit of hydrophilicity and other strategies suggested by multiple regression yielded 98 pyranenamines, the most active [R_1 = R_3 = $NHCO(CHOH)_2H$, R_2 = H] being 1000 times more potent than any original series member.

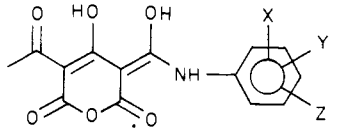
Quantitative structure–activity relationships (QSAR) are equations which express the biological potencies of a series of related compounds as a linear function of their physiochemical properties. A major reason for deriving a QSAR hypothesis is the hope that some aspect of the QSAR can be extrapolated to produce compounds of higher potency. Unfortunately, most of the few examples of successful extrapolation, or "predictive successes",[2] are vulnerable to the following general criticisms: (1) The successful extrapolations are relatively small in magnitude, the potency enhancement in only one instance[3a] being appreciably more than twofold. Most predictions are interpolative. (2) The number of superior compounds associated with any individual successful extrapolation is small. Since an energetic synthesis program is expected to produce compounds of higher potency eventually, regardless of the correctness of any guiding hypothesis, it could be argued that the QSAR success rate is not greater than that produced by chance—or "seat of the pants"—methods alone. (3) The elaborate statistical and computer technologies used to derive QSAR might not be necessary. Alternative but simpler physicochemically based strategies, in particular the "Topliss tree",[3b] seem to point to superior compounds with far less work. (4) "Sooner or later" the compounds embodying the successful extrapolations would have been stumbled upon in any case.

The directed development of the pyranenamine series constitutes a use of QSAR techniques which convincingly counters these criticisms. In retrospect, this development process passed through four sequential phases: (1) progress prior to the use of specialized QSAR techniques; (2) an initial "breakthrough" QSAR prediction using a graphical technique to identify a critical structural property; (3) exploitation of the breakthrough by synthesis of a wide variety of compounds having the desirable property, regression procedures being used in order to explore other structural trends; (4) refinement and confirmation of the SAR understanding embodied in the regression model by synthesis of some less accessible pyranenamines expected to have maximal potency. After an introductory characterization of the biological test system and of the parameters and techniques employed in regression work, the bulk of this paper discusses each of these sequential phases.

Overall Considerations. As discussed within the preceding article,[4] 3-[(arylamino)ethylidene]-5-acyl-

0022-2623/79/1822-0714$01.00/0

pyrantriones, hereafter referred to as pyranenamines (I),



display cromolyn sodium like antiallergic properties in a variety of experimental animal models, in particular the passive cutaneous anaphylactic (PCA) rat. This study is concerned only with the effects of changes in the aromatic substituents -X, -Y, and -Z upon potency in the PCA model.

An extremely direct route of administration of an intravenous injection of completely dissolved[5] compound 30 s before antigen challenge minimized the usual confounding influences of dissolution, transport, and metabolism upon observed potency and favored the development of successful QSAR despite the in vivo character of this test sytem. A less favorable characteristic of these biological data is relatively high uncertainty. Three levels of precision can be distinguished. Measurements of potency at more than one dose were obtained for only a third of the 97 series members, the standard error of such a pI_{50} or log $(1/ED_{50})$ value[6] being roughly ±0.2. The other two-thirds of the series was tested at only a single dose. Whenever the single dose produced a significant response, a pI_{50} value was estimated by extrapolation using the logit transform[7] and the assumption of parallel dose-response curves. Judging from 12 estimates where additional dose-response data were obtained subsequently, these extrapolations are accurate within a standard error of ±0.3. The most serious variability occurs for the third class, 20 compounds which displayed no activity at the single dose. Such compounds were simply assigned a "pI_{50}" 1 unit less than log (1/dose tested). The standard error for such an assignment is unknowable, but if one assumes its value to be ±0.9, the RMS of all the standard errors would be roughly ±0.48.[8]

Regression Studies. Success or failure in a QSAR analysis depends greatly upon an early step, the selection of the structural descriptors to be used. Such descriptor sets should characterize as many as possible of the differences in physical properties which might be responsible for the differences in biological potency. Less important to us is the often-cited objective of keeping the number of descriptor sets as small as possible, in order to lessen the occurrence of the fortuitous or irrelevant correlations[9] which statistics have named "type II" errors. However, statisticians also remind us that decreasing such errors of commission leads to an increasing likelihood of errors of omission. Because the known facts of pharmacology and biological chemistry guarantee that a single correct QSAR exists, whether or not it is obscured hopelessly by limitations in the existing data or in available structural descriptors, and because even an approximately correct QSAR can be exceedingly useful, we would argue that in practical problems it may be more important to perform QSAR analysis in a manner which minimizes type I errors—the rejection or the oversight of a hypothesis which is, in fact, correct—than in a manner which minimizes type II errors.

Each pyranenamine has 45 descriptors, the first being the pI_{50}. Descriptors 2-6, 7-11, and 12-16 give the π, $\mathcal{F}$, and $\mathcal{R}$ values, respectively, of substituents in the 2 through 6 positions of the aromatic ring. (In general, for unsymmetrically substituted rings the assumption was made that the more bulky substituent determines the "2" or "3" position when the pyranenamine phenyl ring is bound to the receptor. Some effort was made to ascertain whether an electronic or lipophilic orienting influence might be operating instead of this customary, if tacit, assumption of a bulk-orienting influence, but without useful result.) The values of π, $\mathcal{F}$, and $\mathcal{R}$ were taken from standard compilation[10] or in the case of numerous untabulated π values estimated by adding π values for the largest possible individual fragments, without attempting to correct for effects of intramolecular complexation.[11] Sets 17-21 contain the volumes of substituents in the 2 through 6 position, calculated following Moriguchi,[12] although it should be noted that this method seems to ignore overlap involving 1,3 and more separated atoms, perhaps leading to a systematic underestimation of the volumes of more highly branched and cyclic substituents. Sets 22-24 code for the presence or absence of a hydrogen-bonding substituent in the 3, 4, or 5 position of type -YH-X, where Y must be an electronegative (O, N, or S) atom attached directly to the phenyl ring and X can be any group including an electron pair. A group bearing two hydrogens, such as NH_2, counts only once.

Many of the remaining sets represent mathematical combinations of sets 2-24. Lipophilicity is summed over the 2 and 6 (ortho) positions, over the 3 and 5 (meta) positions, and over the 2-6 positions, no attempt again being made to correct for effects of intramolecular interactions. The last sum, $\sum\pi$ (27) is squared to yield $(\sum\pi)^2$ (28). Ortho and metal summations are also carried out for $\mathcal{F}$ and $\mathcal{R}$, yielding sets 29-32. The overall σ, that is, the electronic effect of all substituents at the 1 position, arises from summation over all $\mathcal{F}$ and $\mathcal{R}$ values using the weights of Norrington and Williams.[13] Squaring this σ (33) yields σ^2 (34). Volume is not only summed over ortho, meta, or all positions but also the square of each of these sums is tabulated, producing sets 35-40. The sum of sets 22-24 appears as "345-HB" (41). The remaining structural descriptors are less conventional. The variable "#NHSO2" (42) enumerates the substituents of type $-NH-SO_2R$ attached to the aromatic ring. One possible type of intramolecular hydrogen bond is indicated in the next column, "#HB-INTRA" being incremented by 1 for each -YHX attached ortho to a substituent having an electronegative atom. The final two parameters are substructural, "4-OCO?" being equal to 1 if the 4 substituent is attached by an -OC=O- moiety and "(-)?" being incremented by 1 for each negatively charged substituent (at physiological pH).

The complete data matrix is tabulated in the Supplementary Material, along with details of the estimates of the untabulated π, $\mathcal{F}$, and $\mathcal{R}$ values. The data necessary to generate the equations of Table II appear in the appropriate columns of Table I, III, or IV.

The only colinearities in the overall correlation matrix having r^2 greater than 0.8 involve bulks of ortho substituents. Regions of the correlation matrix where r^2 between 0.4 and 0.8 frequently occur include "PI-3", "M-PI", "$PI", and "$PI**"; the properties of ortho substituents generally; and the hydrogen bonding and $\mathcal{R}$ values of 3, 4, or 5 substituents. Because of a 45-variable program limit, the relatively colinear sets 2, 6, 7, 11, 12, 16, 26, 30, and 32 were not candidates in the actual regression study. The complete correlation matrix is reproduced in the Supplementary Material, and a partial matrix involving only those variables appearing in Table II is given in Table V.

The programs developed in our laboratories for QSAR studies feature "interactive stepwise regression", a modification of the common Efromyson stepwise regression

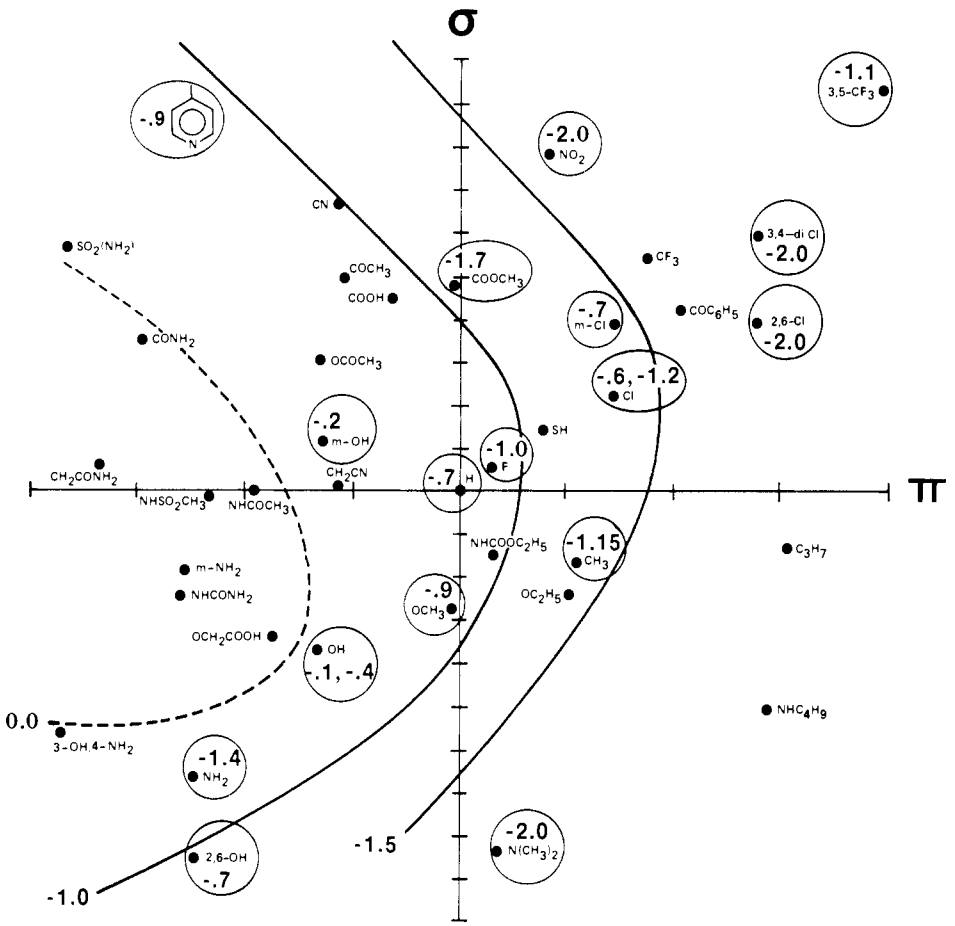


Figure 1. "Relief map" representation of the dependence of potency upon the substituents σ and π for the 19 pyranenamines of Table I, part A.

algorithm,[14] in which at each step the analyst selects the next term for entry from a list of "statistically good" variables. (Conventional stepwise regression instead automatically enters the single "statistically best" but perhaps physically less plausible variable.) Stepwise derivations of the equations presented in Table II appear in the Supplementary Material.

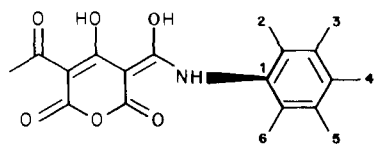
First Phase: "Pre-QSAR" Series Development. Potency (log $1/ED_{50}$) values for the first 19 pyranenamines, synthesized and tested in the PCA assay during the "pre-QSAR" phase of series development, appear in part A of Table I. The majority of these derivatives were inspired by a newly proposed (at that time) decision model, the Topliss operation scheme or "tree".[3] Based in part on the observation that substituent constants are an aspect of the Hansch approach, which is easier to assimilate than is regression analysis, Topliss proposed specific sequences of substituted derivatives to be synthesized, with the next choice at each sequential step being governed by the relative potencies of the preceding pair of compounds. Retrospective studies indeed suggest that use of the "Topliss tree" can halve the number of derivatives needed to achieve an optimal potency. In the pyranenamine series, choice among the "next" appropriate derivatives was complicated by the experimental uncertainty of the potency values, and as a result Table I contains information about most of the nodes in the "Topliss tree", not merely those along an individual branch. For example, the 4-OH derivative, the most potent of the 19, can be reached via the tree only by taking the apparent "wrong turn" at two of three nodes. Strict adherence to the decision model would have produced nothing but derivatives less active than the starting unsubstituted compound.[15]

Thus, the outlook for QSAR studies in the pyranenamine series was at best mixed. Encouraging was the finding that aromatic substituent alteration had produced an eightfold potency range. However, a large variety of substituent types had already been explored, and the lack of success using the Topliss strategy tended to suggest that the customary physicochemical descriptors underlying both the Topliss model and the Hansch approach might prove an unreliable basis for understanding this system.

As the most potent member of the pyranenamine series, by the oral as well as the intravenous route of administration, the 4-OH derivative (**25**) was selected for detailed biological evaluation. Meanwhile, following well-known precepts, a variety of derivatives and close congeners ("bioisosteres") of compound **25** were prepared. The results of their PCA testing, shown in part B of Table I, suggest that this "close analogue" strategy yielded compounds which were equivalent, but not markedly superior, in potency to compound **25**. In the absence of the QSAR studies to be described, a reasonable conclusion would have been that compound **25** is the "optimal" pyranenamine, so that the synthetic program which produced it could be terminated.

Initial analysis of the data in Table I, part A, was carried out graphically,[16] by plotting log $(1/C)$ values for the 19 compounds against various physicochemical properties, specifically measures of size (molar refractivity), affinity for polar solvents (π), and intramolecular electronic effect (σ). The most promising graph was a "three-dimensional" plot (Figure 1) which portrays potency, plotted along the Cartesian axis perpendicular to the page, as a joint function of π and σ.[17] To help in communicating the structure–activity relationship which seemed to emerge, potency

Table I. Pyranenamines Synthesized Before QSAR Derivation



substit	no.[a]	pI_{50}	pI_{50},[b] acc	$\mathfrak{F}$-5	$\mathfrak{R}$-5	$\Sigma\pi$	$(\Sigma\pi)^2$	$(\Sigma\sigma)^2$	M-V[c]	345-HBD[d]	#$NHSO_2$[e]	HB-INTRA[f]	4-OCO?[g]
A. Pyranenamines Used for Derivation of First QSAR													
H	11	−0.7	2	0.00	0.00	0.00	0.00	0.00	0.11	0	0	0	0
2-Cl	12	−1.2	2	0.00	0.00	0.71	0.50	0.15	0.11	0	0	0	0
3-Cl	13	−0.7	2	0.00	0.00	0.71	0.50	0.12	0.30	0	0	0	0
4-Cl	14	−0.6	2	0.00	0.00	0.71	0.50	0.07	0.11	0	0	0	0
4-F	15	−1.0	2	0.00	0.00	0.14	0.01	0.01	0.11	0	0	0	0
4-NO_2	17	−2.0	3	0.00	0.00	0.37	0.13	0.69	0.11	0	0	0	0
4-COOMe	19	−1.7	3	0.00	0.00	−0.01	0.00	0.23	0.11	0	0	0	0
4-Me	22	−1.2	2	0.00	0.00	0.56	0.31	0.03	0.11	0	0	0	0
2-OH	23	−0.4	1	0.00	0.00	−0.60	0.36	0.04	0.11	0	0	0	0
3-OH	24	−0.2	1	0.00	0.00	−0.60	0.36	0.00	0.19	1	0	0	0
4-OH	25	−0.1	1	0.00	0.00	−0.60	0.36	0.12	0.11	1	0	0	0
4-OMe	36	−0.9	2	0.00	0.00	−0.06	0.00	0.06	0.11	0	0	0	0
2-NH_2	40	−1.4	1	0.00	0.00	−1.23	1.51	0.32	0.11	0	0	0	0
4-$N(Me)_2$	44	−2.0	3	0.00	0.00	0.18	0.03	0.67	0.11	0	0	0	0
3,4-Cl_2	77	−2.0	3	0.00	0.00	1.42	2.01	0.37	0.30	0	0	0	0
3,5-$(CF_3)_2$	89	−1.1	2	0.38	0.19	1.96	3.84	0.77	0.76	0	0	0	0
2,6-Cl_2	104	−2.0	3	0.00	0.00	1.42	2.01	0.58	0.11	0	0	0	0
2,6-$(OH)_2$	105	−0.7	3	0.00	0.00	−1.34	1.79	0.15	0.11	0	0	0	0
4-pyridyl	112	−0.9	2	0.00	0.00	−0.54	0.29	0.52	0.11	0	0	0	0
B. Bioisosteres and Derivatives of the 4-OH Congener (SK&F 64398; 25)													
4-OCOMe	26	0.2	2	0.00	0.00	−0.64	0.40	0.11	0.11	0	0	0	1
4-OCOEt	27	−0.1	3	0.00	0.00	−0.14	0.01	0.11	0.11	0	0	0	1
4-OCO-*n*-Pr	28	−0.2	1	0.00	0.00	0.44	0.19	0.11	0.11	0	0	0	1
4-OCO-*n*-Bu	29	0.2	2	0.00	0.00	0.94	0.88	0.11	0.11	0	0	0	1
4-OCO-*n*-Am	30	−1.1	1	0.00	0.00	1.44	2.07	0.11	0.11	0	0	0	1
4-OCO-*n*-Hex	31	0.0	2	0.00	0.00	1.94	3.76	0.11	0.11	0	0	0	1
4-OCO-*t*-Bu	32	−0.7	3	0.00	0.00	0.81	0.65	0.11	0.11	0	0	0	1
4-OCOPh	33	−0.7	3	0.00	0.00	1.46	2.13	0.02	0.11	0	0	0	1
4-$OCONH_2$	34	0.0	2	0.00	0.00	−1.05	1.10	0.02	0.11	0	0	0	1
4-OCONH-CH_2Ph	35	0.0	2	0.00	0.00	0.68	0.46	0.02	0.11	0	0	0	1
4-OCH_2-COOH	38	0.3	1	0.00	0.00	−5.00	25.00	0.06	0.11	0	0	0	0
4-SH	39	−0.2	1	0.00	0.00	0.15	0.02	0.02	0.11	1	0	0	0
4-NH_2	43	0.4	2	0.00	0.00	−1.23	1.51	0.43	0.11	1	0	0	0
4-NHCHO	49	0.3	2	0.00	0.00	−0.98	0.96	0.00	0.11	1	0	0	0

[a] Designation for compounds in the text and in the preceding paper in this issue. (Because some of the pyranenamines in the preceding paper could not be included in the QSAR studies, the numbering in the Supplementary Material differs.) [b] Codes the accuracy of the pI_{50}: = 1 for several doses; = 2 for activity at a single dose; = 3 for no activity at a single dose. See Overall Considerations. [c] Volume of all meta substituents (ref 12). [d] Number of 3, 4, or 5 substituents attached to the aryl ring by -YH- where Y is not carbon. [e] Number of 3, 4, or 5 substituents attached by $-NHSO_2-$. [f] Number of "3,4,5-HBD" substituents ortho to one another. [g] Presence or absence of a 4-*O*-acyl moiety.

contour lines were also sketched in. Thus, the completed graph shows potency as a hypothetical function of π and σ in the same manner that a conventional relief map shows ground elevation as a function of longitude and latitude.

A region of highest potency is located around the middle of the left-hand edge of Figure 1, corresponding to substituents possessing a combination of high hydrophilicity and neutral electronic character. If the relationships suggested by the graph are a correct model of the behavior of substituted pyranenamines in the PCA rat, it appears that substituents conferring even greater hydrophilicity should improve potency further. Since the promising area of Figure 1 is very sparsely populated by actual substituents, various multiply substituted derivatives were proposed as synthetic targets.[18]

Two of the recommended compounds were synthesized immediately and proved to be highly active, a 2-hydroxy-5-acetamido derivative (**106**) having a pI_{50} of +0.2, second only to the 4-NH_2 congener, and the isomeric 3-acetamido-4-hydroxyl derivative **82** having a pI_{50} of +0.7 or a potency 2.5 times greater than any other pyranenamine. This result, a striking indication of the correctness of Figure 1, led to a heavy emphasis on hydrophilic substitution in subsequent work.

Phase III: Development of the Hydrophilicity Lead. During the third phase of series development, a large variety of types and combinations of synthetically convenient hydrophilic substituents was explored. From a QSAR specialist's point of view, while increased hydrophilicity was a primary objective, it was also recognized that exploration of varied substituents might uncover dependencies on parameters other than π and σ. For example, the possible effects of substituent size had yet to be greatly explored, or much of the variation in potency which Figure 1 attributes to overall hydrophilicity might instead by caused by specific hydrogen bonding. From the point of view of classical medicinal chemistry, whether or not the tenfold potency enhancement conferred upon the 4-OH derivative by a 3-acetamide substituent relates at all to increased hydrophilicity, its magnitude mandated further exploration of a traditional nature, such as removal of the 4-OH, lengthening and shortening of the acylamido chain, transposing or altering the substituents, or introducing additional substituents. Compounds actually chosen for synthesis tended to be constructive from either point of view.

Given the possibility that more than the two variables π and σ might be influencing PCA rat potency, the graphical approach had to be abandoned, despite its simplicity of construction and clear portrayal of the postulated structure–activity relationship. In addition to the obvious impossibility of portraying more than three variables on a two-dimensional graph, the time required to construct even two-dimensional graphs having more than, say, 30 points is much greater than that to perform equivalent regressions. A further bonus in using regression techniques is the accompanying statistical indices[8] (r, s, and F), which in principle allow an assessment of the probability that an SAR is "real", i.e., not a chance ordering of information which in reality is unrelated.

All the regression equations to be discussed appear in Table II. Their comparison is facilitated by presenting the terms within an equation vertically. Below each equation are given its r^2, s, F, and largest residuals. Complete residuals are listed in the Supplementary Material.

The equivalence between regression and traditional methods of data analysis is perhaps less appreciated than it should be. For example, a regression equation which can be derived from the data in Table I, part A, provides, in compact form, much the same information as does Figure 1. As shown in Table II (eq A), this equation has two structurally dependent terms, corresponding to the two dimensions π and σ of Figure 1. The negative coefficient of its π term suggests that increases in substituent hydrophilicity will tend to augment potency, equivalent to the right to left upward slope of the "hill" in Figure 1. However, the large confidence intervals of this coefficient imply that our hypothesis of a relationship between hydrophilicity and potency was actually very dubious statistically, despite its productivity. The negative coefficient of the "$SIG**" term in eq A connotes a potency which decreases as σ deviates markedly from zero, just as suggested in Figure 1.

The unusually low value of r^2 for eq A requires comment. Since one should never expect s to be much less than the experimental variability of the biological measurements, the only circumstances allowing achievement of a high r^2 are a very low experimental variability or a very large range of biological potencies. In fact, the r^2 of 0.46 is actually about as high as one might hope to obtain from the data of Table I, since its accompanying s of 0.47 is comparable to the experimental variability of 0.48 previously estimated for the overall data set.[8] Fortunately, comparison of the $F_{2,16} = 7.3$ with tabulated values shows that the likelihood of obtaining a correlation even this "poor" by chance alone is less than 1%.

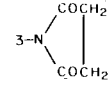
Hydrophilic groups explored during this third phase of pyranenamine development are listed in Table III. Several features of these data are inconsistent with the initial hypothesis that biological potency depends solely on π and σ: (1) The 3,5-$(NHCOMe)_2$ derivative **94** is an order of magnitude more active than any other series member despite unremarkable π and σ values. (2) The potencies of the increasingly lipophilic 3-NHCOMe, 3-NHCOEt, and 3-NHCOPr subseries (**51**, **55**, and **59**) and of the lengthy 4-OCOR subseries (**26**–**33**) would be expected to decline, rather than remain constant, in the absence of additional trends. These observations appear to be satisfactorily explained by eq B in Table II, derived from all data in Tables I and III, except for exclusion of the 4-OCH_2COOH substituent because of the structural ambiguity introduced by its two possible protomeric forms.

Lipophilicity, a statistically dubious causative factor in eq A, has with the additional data emerged as the dominant structural influence on potency, according to the F test of the $\sum\pi$ term. At this point in series development, potency appeared to double with each log unit increase in hydrophilicity, over a range of 4 log units. Since many more log units of substituent hydrophilicity were still potentially accessible, most readily by employing charged substituents, a major unresolved question was: "Will a substituent hydrophilicity be encountered which is optimal, with further increases in hydrophilicity perhaps depressing potency?"

Electronic effects on potency are expressed by the second and third terms in eq B. Strongly electron-donating or -withdrawing effects upon the 1 position of the phenyl ring [the $(\sum\pi)^2$ term] continue to be deleterious to potency. However, the "$\mathcal{F}$-5" term suggests that an inductively electron-withdrawing substituent (high value of the Swain–Lupton $\mathcal{F}$) in the 5 position may very substantially enhance potency. A 5 substituent is defined only when there are substituents having at least equivalent bulk at

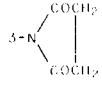
Table II. Regression Equations Showing the Evolution of the Pyranenamine QSAR[a]

A. after 19 compds[c]	B. after 61 compds[d]		C. after 98 compds[e]	
$pI_{50} = -0.72$	$pI_{50} = -0.75$		$pI_{50} = -0.59$	
$-0.14\ (\pm 0.29) \times \Sigma\pi$	$-0.30\ (\pm 0.12) \times \Sigma\pi$	(23)	$-0.33\ (\pm 0.11) \times \Sigma\pi$	(34)
$-1.35\ (\pm 0.98) \times (\Sigma\sigma)^2$	$-1.5\ (\pm 0.67) \times (\Sigma\sigma)^2$	(16)	$-0.034\ (\pm 0.016) \times (\Sigma\pi)^2$	(16)
	$+2.0\ (\pm 2.0) \times$ ℱ-5	(4)	$+4.3\ (\pm 1.6) \times$ ℱ-5	(31)
	$+0.39\ (\pm 0.22) \times$ #345-HBD	(12)	$+1.3\ (\pm 0.85) \times$ ℜ-5	(9)
	$-0.63\ (\pm 0.33) \times$ #NHSO2	(14)	$-1.7\ (\pm 0.62) \times (\Sigma\sigma)^2$	(28)
	$+0.78\ (\pm 0.46) \times$ M-V	(11)	$+0.73\ (\pm 0.22) \times$ #345-HBD	(45)
	$+0.72\ (\pm 0.31) \times$ 4-OCO?	(21)	$-0.86\ (\pm 0.34) \times$ #HB-INTRA	(25)
			$-0.69\ (\pm 0.28) \times$ #NHSO2	(24)
			$+0.72\ (\pm 0.35) \times$ 4-OCO?	(16)
$r^2 = 0.48, s = 0.47, F_{2,16} = 7.3$	$r^2 = 0.77, s = 0.40, F_{7,53} = 25.1$		$r^2 = 0.75, s = 0.48, F_{9,88} = 28.7$	

residuals > s[b]			residuals > s[b]			residuals > s[b]					
substit	calcd pI_{50}	obsd − calcd	substit	calcd pI_{50}	obsd − calcd	substit	calcd pI_{50}	obsd − calcd	substit	calcd	obsd − calcd
3,5-$(CF_3)_2$	−2.07	+0.97	4-NH_2	−0.54	−0.94	3,5-$(NHCOEt)_2$	1.28	−1.22	2-NHCHO	−0.32	−0.62
4-OH	−0.80	+0.70	3-NHCHO	+0.21	+0.91	4-NHCOMe	0.44	+1.14	3-NHC(=NH)NH_2	+0.70	+0.60
4-COOMe	−1.03	−0.67	4-NHCHOMe	+0.02	+0.72	3-NHCHO	0.39	+1.09	4-SO_2NH_2	−0.70	−0.60
3,4-Cl	−1.43	−0.57	4-COOMe	−0.99	+0.70	3-OCH_2COO-	0.20	+0.90	4-OCO-*n*-Bu	−0.40	−0.60
3-OH	−0.64	+0.44	2-NHCHO	−0.39	−0.69	4-OCO-*n*-Hex	−0.82	−0.82		−0.59	−0.59
			4-NHCO-*i*-Pr	−0.29	−0.69	3-SO_2NH_2	−0.48	−0.78			
			4-OCO-*n*-Hex	−0.69	−0.69	3-NHMe	+0.07	+0.77			
			2-NH_2	−0.76	+0.64	2-COOEt	−1.24	−0.74			
			4-OCO-*n*-Bu	−0.39	−0.59	4-COOMe	−0.96	+0.74	3-(NHCO)-*i*-Pr-4-OH	+0.11	−0.59
			4-OCO-*n*-Am	−0.54	+0.56	3,5-$[NHCO(CHOH)_2H]_2$	+2.27	−0.73	3-$NHSO_2$Me-4-OH	−0.23	+0.57
			4-SO_2NH_2	−0.65	−0.55	3-NHC(=NH)NHTos	+0.08	+0.71	3-NHCOCOO-	+0.95	−0.55
			3,4-Cl	−1.5	+0.50	3-$NHSO_2NH_2$	−0.18	−0.68	3-NHCOMe-4-OMe	−0.43	−0.53
			3-SO_2NH_2	−0.13	−0.43	2-NH_2	−0.75	+0.64	3-OH	+0.33	+0.53
			3-NH_2	+0.12	−0.38	3-CONHMe	−0.43	−0.63	4-OCO-*n*-Am	−0.60	+0.50
						3-$NHSO_2$Ph	−0.73	−0.63	4-NHCOEt	+0.29	+0.49
						4-$NHSO_2NH_2$	−0.13	−0.63	4-$NHSO_2$Me	−0.21	+0.49
						4-NH_2	−0.22	−0.62	3-$NHCO(CHOH)_2H$	+0.84	−0.46

[a] The terms of the equations are listed vertically. For each term are given, left to right, the coefficient and its 95% confidence interval; the name of the structural descriptor; and, for equations B and C, the F test in parentheses. Below each equation appear its r^2, s, and F overall. [b] Only the largest residuals are listed. The complete list appears in the Supplementary Material. [c] Listed in part A of Table I. [d] Listed in Tables I and III (4-OCH_2COOH excluded). [e] All compounds in Tables I, III, and IV.

Table III. Pyranenamines Synthesized during the Third Phase, in Pursuit of Hydrophilicity

substit	no.[a]	pI_{50}	pI_{50},[b] acc	$\mathfrak{F}$-5	$\mathfrak{R}$-5	$\Sigma\pi$	$(\Sigma\pi)^2$	$(\Sigma\sigma)^2$	M-V[c]	345-HBD[d]	#$NHSO_2$[e]	#HB-INTRA[f]	4-OCO?[g]
3-SO_2NH_2	20	0.3	2	0.00	0.00	−1.82	3.31	0.22	0.53	0	0	0	0
4-SO_2NH_2	21	−0.1	3	0.00	0.00	−1.82	3.31	0.36	0.11	0	0	0	0
3-NH_2	41	0.5	1	0.00	0.00	−1.23	1.51	0.45	0.23	1	0	0	0
2-NHCHO	47	0.3	2	0.00	0.00	−0.98	0.96	0.01	0.11	0	0	0	0
3-NHCHO	48	−0.7	3	0.00	0.00	−0.98	0.96	0.03	0.41	1	0	0	0
2-NHCOMe	50	−0.7	3	0.00	0.00	−0.97	0.94	0.02	0.11	0	0	0	0
3-NHCOMe	51	0.7	2	0.00	0.00	−0.97	0.94	0.03	0.57	1	0	0	0
4-NHCOMe	52	−0.7	2	0.00	0.00	−0.97	0.94	0.00	0.11	1	0	0	0
3-$NHCONH_2$	53	0.3	2	0.00	0.00	−1.30	1.69	0.00	0.50	1	0	0	0
4-$NHCONH_2$	54	0.2	2	0.00	0.00	−1.30	1.69	0.05	0.11	1	0	0	0
3-NHCOEt	55	0.7	1	0.00	0.00	−0.47	0.22	0.02	0.72	1	0	0	0
4-NHCOEt	56	−0.2	1	0.00	0.00	−0.47	0.22	0.00	0.11	1	0	0	0
3-NHCO-*i*-Pr	57	0.5	2	0.00	0.00	0.03	0.00	0.02	0.87	1	0	0	0
4-NHCO-*i*-Pr	58	0.4	2	0.00	0.00	0.03	0.00	0.00	0.11	1	0	0	0
3-NHCO-*n*-Pr	59	0.4	2	0.00	0.00	0.03	0.00	0.04	0.87	1	0	0	0
	61	0.0	2	0.00	0.00	−1.60	2.56	0.76	0.80	0	0	0	0
3-$NHSO_2Me$	62	0.1	2	0.00	0.00	−1.18	1.39	0.03	0.68	1	1	0	0
4-$NHSO_2Me$	63	0.7	3	0.00	0.00	−1.18	1.39	0.00	0.11	1	1	0	0
3-$NHSO_2Ph$	64	0.1	1	0.00	0.00	0.45	0.20	0.02	1.23	1	1	0	0
4-$NHSO_2Ph$	65	−0.7	3	0.00	0.00	0.45	0.20	0.00	0.11	1	1	0	0
3-NH_2-4-OH	78	0.1	1	0.00	0.00	−1.90	3.61	0.32	0.23	2	0	1	0
3-NHCOMe-4-OH	82	0.7	1	0.00	0.00	−1.57	2.46	0.03	0.57	2	0	1	0
3,4-(N=CMe-O-)$_2$	83	0.1	3	0.00	0.00	0.37	0.13	0.19	0.22	0	0	0	0
3-NHCOMe-4-OMe	84	0.1	2	0.00	0.00	0.99	0.98	0.00	0.36	1	0	1	0
3-NHCOEt-4-OH	86	0.4	1	0.00	0.00	−1.07	1.14	0.04	0.72	2	0	1	0
3-NHCO-*i*-Pr-4-OH	87	0.7	1	0.00	0.00	−0.57	0.32	0.05	0.87	2	0	1	0
3,5-$(NHCOMe)_2$	94	1.9	1	0.28	−0.26	−1.94	3.76	0.14	1.02	2	0	0	0
3-NHCOMe-6-OH	106	0.2	2	0.00	0.00	−1.57	2.46	0.00	0.57	1	0	0	0

[a–g] See corresponding footnotes in Table I.

the 2 and/or 3 position, so this trend is based on just two examples, the 3,5-$(CF_3)_2$ ($\mathcal{F}$-5 = 0.38) and 3,5-$(NHCOMe)_2$ ($\mathcal{F}$-5 = 0.28) derivatives. Further exploration of a larger range of variation in $\mathcal{F}$-5 was clearly indicated.

The hypothesis of a "directional" electronic effect acting only at the 5 position is not as inconsistent with classical physical organic chemistry as might at first appear. It must be remembered that the σ scale is defined by *intramolecular* phenomena, whereas receptor binding is *intermolecular*. Because of symmetry, one would not expect the intramolecular electronic effects of a substituent on the 5 position to differ from effects of a 3 substituent, but the intermolecular environment of an aromatic ring bound to a receptor might be very different at different positions.

The next pair of terms in eq B ascribes desirability to certain types of hydrogen-bonding groups attached to the meta and/or para positions. The "#345-HBD" term describes an increase in potency of 0.42 log unit for every group of general type -HYR attached to the 3, 4, or 5 position, where Y may be N, O, or S and R may be anything including H or a lone pair. (At this point, actual -YHR were OH, SH, NH_2, NHCOR′, and $NHSO_2R'$, where R′ = H, alkyl, NH_2, or aryl.) This trend might be attributed to actual hydrogen bonding between pyranenamine and its receptor, although it is not easy to understand how any number of hydrogen bond donors in any position might exert equivalent and additive effects. Note that, even though the trends involving $\sum\pi$ and #345-HBD portrayed by eq B (and C) physicochemically parallel one another, colinearity between these variables for all 98 compounds is actually quite small (r = -0.35), and thus the two trends are statistically independent. The second hydrogen-bonding variable "#$NHSO_2$" is substructural and suggests that the $-NHSO_2R$ group is a -YHR which does not have suitable properties. Together these trends raised the question: "What other types and arrangements of hydrogen-bonding groups will and will not enhance potency?"

The "M-V" term defines a surprising situation; an increase in the volume of meta (3 or 5) substituents appears to *increase* potency. One might question, in retrospect, whether the volumes of such variously shaped substituent types *should* have commensurable effects on potency. Previously published QSAR correlations involving size have been based mostly either on small and nearly symmetric groups or else on highly flexible groups such as higher alkyl, whereas the semirigidity of many of these large groups probably constrains them to occupy different regions of "receptor space". Nevertheless, considerable QSAR effort was expended, attempting to define an "optimal group size".

Finally, the substructural 4-OCO? term, from its F test, the second most important overall to eq B, indicates that acyl derivatives of 4-OH are five times as potent as would be implied by the physical properties of the esters themselves. This trend is consistent with the speculation that these acyl derivatives behave as biologically equivalent prodrugs, hydrolyzing in vivo to produce the 4-OH derivative itself.

At this time, SK&F 78729, the 3-NH_2-4-OH derivative **78**, replaced compound **25** as the clinical lead. It is interesting that this structure exemplifies the general type of structural modification initiated by the graphical QSAR study and, in fact, is simply the deacylation product of the 3-NHAc-4-OH derivative **82** whose high potency first validated the study. However, the superior promise of compound **78** resided primarily in its properties in secondary test systems which are only weakly related to the PCA assay.

Final Phase of Pyranenamine Development. Most of the remaining pyranenamines, listed in Table IV, were synthesized to answer specific SAR questions of greatly enhanced potency. As detailed above, key QSAR questions concerned the possibility of a hydrophilicity optimum, the need to identify potency-enhancing hydrogen-bonding groups, and the reality of the $\mathcal{F}$-5 and M-V terms. To help answer these questions, considerable synthetic effort went into the preparation of particular exotic substituents, such as $-NHC(=NH)NH_2$, -CONH (barbiturate), and $-NHCO(CHOH)_2H$.

The regression equation which seems to best describe the QSAR for all 98 pyranenamines is eq C in Table II. Compared with eq B, additional terms relating to lipophilicity, electronic effects, and hydrogen bonding have appeared, while the volume-related term has disappeared.

The new lipophilicity term, $(\sum\pi)^2$, when taken with the $\sum\pi$ term, defines a parabolic relationship between potency and hydrophilicity, with a remarkably hydrophilic optimum, roughly -5, for the sum of substituent π values.[19] It should be cautioned that only one of the compounds has an estimated π substantially less than -5 [3,5-$(NHCOCO^-)_2$], and all of the half-dozen π estimates for groups more hydrophilic than π = -2.5 are based on strict group additivity, no correction being made for intramolecular interactions, and thus are probably too hydrophilic.[20] It is also possible that negative charge, rather than extreme hydrophilicity, is the property of extremely hydrophilic groups which is deleterious to potency. If nine of the compounds (**37**, **30**, **45**, **46**, **69**, **70**, **76**, **99**, and **100**) having the most hydrophilic of the uncertain π estimates are deleted, a similar equation lacking a $(\sum\pi)^2$ is obtained (see Supplementary Material). Nevertheless, the existence of an unusually hydrophilic optimum within these potency data seems at least qualitatively probable.

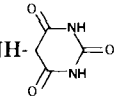
The electronic aspects of eq C extend those of eq B. The $(\sum\sigma)^2$ term continues to indicate the desirability of an overall σ near 0,[21] while the $\mathcal{F}$-5 term strongly sustains the value of inductively withdrawing substituents at the 5 position. The $\mathcal{R}$-5 term buttresses the $\mathcal{F}$-5 term by its indication that resonance-withdrawing characteristics further promote potency.

Hydrogen bonding, or some other characteristic of the previously defined -YHR group, has become the most important feature of the series (relative F test). The "#345-HBD" term describes a fourfold potency enhancement for each -YHR substituent. However, the "#HB-INTRA" term constrains the #345-HBD term by indicating that a -YHR, which is capable of forming an intramolecular hydrogen bond (five- or six-membered ring) with a heteroatom-containing ortho substituent, will not enhance potency,[22] and, as in eq B, the "NHSO2" term effectively excludes $-NHSO_2R$ groups from this potency-enhancing -YHR class.

The remaining term, "4-OCO?", continues to attribute unusual behavior to acyl derivatives of the 4-OH pyranenamine. However, the M-V effect within eq B has not persisted strongly. Although the next of the remaining 40 variables to enter eq C would favor increased substituent volume, particularly of meta substituents, the associated increase in r^2 and decrease in s would be only 0.01.

It is interesting that the s values of eq A, B, and C remained stable as the series developed, while the r^2 values increased from 0.48 to 0.77 as the range of pyranenamine potencies expanded. A more detailed understanding of the strengths and weaknesses of these QSAR is provided by

Table IV. Compounds Synthesized during the Final Phase

substit	no.[a]	pI_{50}	pI_{50},[b] acc	$\mathcal{I}$-5	$\mathcal{R}$-5	$\Sigma\pi$	$(\Sigma\pi)^2$	$(\Sigma\sigma)^2$	M-V[c]	345-HBD[d]	#$NHSO_2$[e]	#HB-INTRA[f]	4-OCO?[g]
4-Br	16	−0.7	3	0.00	0.00	0.86	0.74	0.07	0.11	0	0	0	0
2-COOEt	18	−0.5	3	0.00	0.00	0.51	0.26	0.29	0.11	0	0	0	0
3-OCH_2COOH	37	−0.7	3	0.00	0.00	−5.00	25.00	0.00	0.58	0	0	0	0
3-NHMe	42	−0.7	3	0.00	0.00	−0.47	0.22	0.13	0.39	1	0	0	0
3-NHC=NHNHSO$_2$Ph-*p*-Me	45	−0.7	3	0.00	0.00	0.00	0.00	0.08	1.71	1	0	0	0
3-NHC=$NHNH_2$	46	0.1	1	0.00	0.00	−4.50	20.25	0.15	0.55	1	0	0	0
3-$NHCOCH_2CH_2COOEt$	60	0.3	1	0.00	0.00	−0.48	0.23	0.03	1.11	1	0	0	0
3-$NHSO_2NH_2$	66	0.5	2	0.00	0.00	−1.52	2.31	0.03	0.62	1	1	0	0
4-$NHSO_2NH_2$	67	0.5	1	0.00	0.00	−1.52	2.31	0.00	0.11	1	1	0	0
3-NHCOOEt	68	0.4	1	0.00	0.00	0.17	0.03	0.00	0.78	1	0	0	0
3-NHCOCOO-	69	1.5	1	0.00	0.00	−5.00	25.00	0.02	0.63	1	0	0	0
3-$NHCO(CHOH)_2H$	70	1.3	1	0.00	0.00	−3.10	9.61	0.03	0.84	1	0	0	0
3-$CH_2NHCOCH_3$	71	−0.7	3	0.00	0.00	−0.70	0.49	0.00	0.72	0	0	0	0
3-$CONH_2$	72	−0.7	3	0.00	0.00	−1.49	2.22	0.08	0.30	0	0	0	0
3-CONHMe	73	0.2	2	0.00	0.00	−1.27	1.61	0.12	0.57	0	0	0	0
3-CONHEt	74	−0.7	3	0.00	0.00	−0.77	0.59	0.12	0.72	0	0	0	0
3-$CON(Me)_2$	75	−0.7	3	0.00	0.00	−0.08	0.00	0.12	0.72	0	0	0	0
3-CONH- 	76	0.2	2	0.00	0.00	−2.77	7.67	0.12	1.09	0	0	0	0
3-OH-4-NH_2	79	−0.5	1	0.00	0.00	−1.90	3.61	0.36	0.19	2	0	1	0
3-NHMe-4-OH	80	−0.7	3	0.00	0.00	−1.14	1.29	0.51	0.39	2	0	1	0
3-CH_2NHMe-4-OH	81	−0.7	3	0.00	0.00	−1.42	2.01	0.12	0.54	1	0	1	0
3-CONHMe-4-OH	85	0.1	2	0.00	0.00	−1.94	3.76	0.00	0.57	1	0	1	0
3-$NHSO_2Me$-4-OH	88	−0.8	2	0.00	0.00	−1.85	3.42	0.03	0.69	2	1	1	0
3-OH-5-NH_2	90	0.2	1	0.02	−0.68	−1.90	3.61	0.02	0.31	2	0	0	0
3-NHCOMe-5-OH	91	1.7	1	0.29	−0.64	−1.64	2.68	0.02	0.65	2	0	0	0
3-NHCOMe-5-NH_2	92	1.0	1	0.02	−0.68	−2.20	4.84	0.00	0.69	2	0	0	0
3,5-$(NH_2)_2$	93	0.3	1	0.02	−0.68	−2.46	6.05	0.19	0.35	2	0	0	0
3,5-$(NHCOEt)_2$	95	2.5	1	0.18	−0.26	−0.02	0.00	0.03	1.33	2	0	0	0
3,5-(NHCO-*n*-$Pr)_2$	96	1.3	1	0.14	−0.28	0.06	0.00	0.01	1.64	2	0	0	0
3,5-$(NHCOOEt)_2$	97	0.6	1	0.14	−0.28	0.34	0.12	0.01	1.45	2	0	0	0
3,5-$(NHCOCOOEt)_2$	98	1.7	1	0.36	−0.21	0.02	0.00	0.31	1.80	2	0	0	0
3,5-$(NHCOCOO-)_2$	99	1.5	1	0.28	−0.26	−10.00	100.00	0.14	1.16	2	0	0	0
3,5-$[NHCO(CHOH)_2H]_2$	100	3.0	1	0.28	−0.26	−6.20	38.44	0.14	1.57	2	0	0	0
3,5-$(NHSO_2Me)_2$	101	0.5	1	0.25	−0.20	−2.36	5.56	0.12	1.26	2	2	0	0
3,5-$(NHSO_2Ph)_2$	102	−0.7	3	0.21	−0.18	0.90	0.81	0.08	2.35	2	2	0	0
3,4,5-$(OH)_3$	103	0.4	1	0.29	−0.64	−2.01	4.04	0.05	0.27	3	0	2	0
3-OH-6-NH_2	107	−0.5	2	0.00	0.00	−1.90	3.61	0.25	0.19	1	0	0	0

[a-g] See corresponding footnotes in Table I.

Table V. Correlation Matrix (r) for Structural Properties Given in Tables I, III, and IV

	pI_{50}	$\mathcal{F}$-5	$\mathcal{R}$-5	$\Sigma\pi$	$(\Sigma\pi)^2$	$(\Sigma\sigma)^2$	M-V	#345-HBD	#$NHSO_2$	#HB-INTRA	4-OCO?
pI_{50}	1.00	0.46	0.44	−0.48	0.33	−0.39	0.49	0.59	−0.02	0.04	−0.06
$\mathcal{F}$-5		1.00	−0.49	−0.20	0.34	0.16	0.56	0.59	0.18	0.06	−0.12
$\mathcal{R}$-5			1.00	0.27	−0.18	0.14	−0.28	−0.60	−0.03	−0.11	0.12
$\Sigma\pi$				1.00	−0.80	0.14	−0.20	−0.35	0.01	−0.12	0.29
$(\Sigma\pi)^2$					1.00	−0.01	0.22	0.19	−0.04	−0.04	−0.08
$(\Sigma\sigma)^2$						1.00	−0.12	−0.19	−0.14	0.03	−0.08
M-V							1.00	0.52	0.32	−0.00	−0.00
#345-HBD								1.00	0.25	0.49	−0.33
#$NHSO_2$									1.00	−0.03	−0.10
#HB-INTRA										1.00	−0.12
4-OCO?											1.00

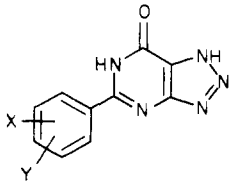
the residuals, in particular those for the final eq C. The largest deviations are too great to be produced by biological variability, so their existence must have structural origin. However, a rationale for the outlying behavior of such structurally nonremarkable substituents, such as the 3,5-$(NHCOEt)_2$, 4-NHCOMe, 3-NHCHO, and 3-OCO_2COO^- derivatives, is far from obvious, while, at the same time, such unusual functionalities as $NHC(=NH)NH_2$, $-N(COCH_2)_2$, and CONH (barbiturate) are relatively well handled by eq C.

The pyranenamine found to be the most potent of all, the 3,5-$[NHCO(CHOH)_2H]_2$ derivative **100**, exemplifies the characteristics that eq C promotes. The two highly polar $-NHCO(CHOH)_2H$ groups give an estimated substituent total π value of −6.2, not far from the optimum π. There are two −YHR groups, which being meta are assumed not to form an intramolecular hydrogen bond. The 5 substituent has a positive $\mathcal{F}$ value of +0.28, somewhat offset by a negative $\mathcal{R}$ value of −0.25, producing with the 3 substituent an overall negligibly positive $\sum\sigma^2$ of 0.08. This combination of properties yields a predicted pI_{50} of 2.3 when inserted into eq C and an actual pI_{50} of +3.0. Biological activity at this nanomolar administered blood level is exceptional for any type of pharmacological agent.

Discussion

The contributions of QSAR to the development of the pyranenamines were substantial during all phases of the program. As discussed above, an immediate potency enhancement of almost an order of magnitude was produced by the first graphical QSAR, specifically the 3-NHAc-4-OH pyranenamine, whose de-N-acylated derivative became the clinical lead (**78**). Continued pursuit of these and other trends ultimately led to compound **100**, a thousand times more active in the PCA rat assay than any member of the original series.

Another predictive QSAR "success story" of comparable magnitude reported by Woolridge[2] involves the same biological test system but a structurally very different lead molecule (II). Interestingly, similar QSAR trends were



II

encountered, hydrophilicity and electron-withdrawing character (preferably rather specifically oriented) being desirable. This similarity is consistent with the postulate of similar, if not identical, receptors for any molecule having this action. The potency enhancement of 100 000× reported by Wooldridge is greater than the 1000× enhancement achieved with the pyranenamines, although the final potencies are similar.

It is fashionable to draw speculative conclusions about drug–receptor interactions from the physicochemical trends revealed by a QSAR equation. An unusual feature of the equations in Table II is the tendency for potency to increase, rather than decrease, with increasing hydrophilicity and the remarkably hydrophilic optimum π. The additional and parallel strong tendency for potency increases to be associated with particular hydrogen-bonding moieties seems, because of its geometric specificity, to reflect receptor-site interactions rather than transport. These two factors, unusual hydrophilicity dependence and its parallelism to probably receptor-mediated hydrogen bonding, suggest the possibility that π dependency for this series and test system may relate more to receptor binding than to transport. There may be a "hydrophilic pocket" for the aryl ring, one or more of whose hydrophilic groups are well-positioned for interaction with −YHR groups attached to the 3, 4, or 5 position of the aryl ring.

The potency conferred by 5 substituents which are electron-withdrawing, particularly via an inductive mechanism, can be rationalized by the further assumption that a hydrophilic moiety adjacent to the 5 position is also electropositive, for example, a protonated lysine, arginine, or histidine residue. However, recognizing that an unsymmetrically substituted aryl ring can by "flipping" assume two possible orientations with respect to the receptor site, the resulting ambiguities being settled for QSAR purposes by an assumption that bulk will be the dominant influence on orientation, one might ask why certain strongly electron-withdrawing meta substituents, i.e., SO_2NH_2 or Cl, are not strongly enough attracted by this electropositive center to interact with the receptor as a "5" substituent rather than a "3" substituent?

In conclusion, these results provide a very satisfactory "success story" to help counter the criticisms of many previous QSAR "predictions" listed in the beginning of this article. Specifically: (1) The extent of the potency enhancement, from an original potency range spanning perhaps two orders of magnitude, immediately to three orders, and ultimately to five orders of magnitude, is hardly trivial. (2) The enhancements in potency produced by the use of QSAR are entirely too consistent, across the series and over time, to be attributed to chance. (3) By producing useful QSAR from data which were not successfully handled by the Topliss decision model, computer and statistical techniques in this study demonstrated a superiority in analytical power.[23] (4) Although many members of this series might, sooner or later, have been prepared without the influence of QSAR studies, the most active compounds include unusual substituents. For instance, are there any other examples of a 3,5-$[NHCO(CHOH)_2H]_2$ aryl substitution, which in this series was the

most potent combination encountered?

The exceptional degree of success encountered in using QSAR techniques on the pyranenamine series must reflect an appreciable amount of good fortune. None of the other 20-odd QSAR studies in these laboratories have been so dramatically productive, and at leeast a third of these studies have altogether failed, that is, produced correlations and extrapolative predictions which were subsequently found to be completely erroneous. Nevertheless, the outstanding contribution which the QSAR approach made to the development of the pyranenamine series argues strongly for its trial in any program which seeks to optimize the aromatic substituent on a "lead" molecule.

Supplementary Material Available: Complete parameter values and their correlation matrix, calculation of π, $\mathcal{F}$, and $\mathcal{R}$ values not appearing in the literature, and the stepwise derivation of eq C and of an equation excluding nine more hydrophilic pyranenamines, along with their complete residuals and residual plots (32 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Presented in part at the 175th National Meeting of the American Chemical Society, Anaheim, CA, Mar 13-17, 1978. Altered version to appear in a forthcoming ACS Symposium Series Book on "Drugs Affecting the Respiratory System".

(2) Reviewed in C. Hansch, M. Yoshimoto, and M. H. Doll, *J. Med. Chem.*, **19**, 1089 (1976); R. D. Cramer III, *Ann. Rep. Med. Chem.*, 11, 301 (1976); and J. G. Topliss and J. Y. Fukunaga, *ibid.*, 13, 292 (1978). Recent tests of predictions include C. Grieco, C. Silipo, A. Vittoria, and C. Hansch, *J. Med. Chem.*, **20**, 596 (1977); C. Hansch and J. Y. Fukunaga, *CHEMTECH*, **7**, 120 (1977); and ref 16.

(3) (a) K. R. H. Wooldridge, *Med. Chem., Proc. Int. Symp., Main Lect., 5th, 1975*, 427 (1976). I thank J. G. Topliss for referring me to this work. (b) J. G. Topliss, *J. Med. Chem.*, **15**, 1006 (1972); Y. C. Martin and W. J. Dunn III, *J. Med. Chem.*, **16**, 578 (1973); J. G. Topliss and Y. C. Martin, *Drug Des.*, **5**, 1, (1975).

(4) K. M. Snader, L. W. Chakrin, R. D. Cramer III, Y. M. Gelernt, C. K. Miao, D. H. Shah, J. W. Venslavsky, C. R. Willis, and B. M. Sutton, *J. Med. Chem.*, **22**, preceding paper in this issue.

(5) A cosolvent, typically Me_2SO, was added in amount sufficient to produce complete solution.

(6) The log $(1/ED_{50})$ transform puts biological data into a more readily analyzable form and is theoretically reasonable by analogy with the familiar equilibrium relation between reagent concentrations and free energy difference:

$$\Delta G = -2.303RT \log (C_{products}/C_{reactants})$$

(7) G. Redl, R. D. Cramer III, and C. E. Berkoff, *Chem. Soc. Rev.*, **3**, 273 (1974).

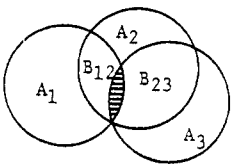
(8) An outstanding discussion of the interrelationships among the r, s, and F properties of a regression equation, filling an apparent void in the QSAR literature, is given by Y. C. Martin in "Quantitative Drug Design: A Critical Introduction", Marcel Dekker, New York, 1978, Chapters 5 and 7. Regression studies in which individual pyranenamine data were weighted inversely as their estimated pI_{50} accuracies gave results similar to those of Table II.

(9) J. G. Topliss and R. J. Costello, *J. Med. Chem.*, **15**, 1066 (1972).

(10) C. Hansch, A. Leo, S. Unger, K. Kim, D. Nikaitami, and E. Lien, *J. Med. Chem.*, **16**, 1207 (1973).

(11) Details of the calculation of untabulated π, $\mathcal{F}$, and $\mathcal{R}$ values appear in the Supplementary Material. The effects of intramolecular hydrogen bonding in increasing the actual log P values of the more hydrophilic pyranenamines are probably large but not calculable. Unfortunately, efforts to measure these values experimentally were not successful.

(12) I. Moriguchi, Y. Kanada, and K. Kumatsu, *Chem. Pharm. Bull.*, **24**, 1799 (1976). The errors which might arise from neglect of 1,3 overlap are surprisingly complex. Consider the planar projection of a three atom chain:



The volume of this assembly would be computed following Moriguchi as: $A_1 + A_2 + A_3 - B_{12} - B_{23}$. The shaded area where A_1 and A_3 themselves overlap (within A_2) is subtracted twice (being included in both B_{12} and B_{23}) and thus the volume of the total assembly would be underestimated. However, if A_1 and/or A_3 become larger, there would arise A_1,A_3 overlap still not recognized in the Moriguchi sum but now *outside* of A_2, resulting in a countertendency to *overestimate* the assembly volume. When one furthermore recognizes possible overlap among 1,4 and more distant atoms (as might occur in cyclic systems), it would appear that a carefully structured computer program is almost mandatory for reliable and self-consistent volume calculations even using this simple sphere-overlap model of volume. Another practical approach involves "analogue modeling" via the displacement volume of CPK models [A. Leo, C. Hansch, and P. Y. C. Jow, *J. Med. Chem.*, **19**, 611 (1976)].

(13) S. G. Williams and F. E. Norrington, *J. Am. Chem. Soc.*, **98**, 508 (1976).

(14) N. R. Draper and H. Smith, "Applied Regression Analysis", Wiley, New York, 1966, Chapter 11.

(15) A more recent alternative, preparation of five derivatives and comparing their potency order with tabulated orders [J. G. Topliss, *J. Med. Chem.*, **20**, 463 (1977)], fails also. The order actually observed is not among those tabulated, probably because of experimental error in one or more of the original potency measurements which is excessive compared to the π and σ relationships which actually exist.

(16) Similar procedures have since been described: F. Darvas, *J. Med. Chem.*, **17**, 799 (1974); I. T. Harrison, W. Kurz, I. J. Massey, and S. H. Unger, *ibid.*, **21**, 588 (1978).

(17) Another description of Figure 1 would be "potency as third dimension on a Craig plot"; P. N. Craig, *J. Med. Chem.*, **14**, 680 (1971).

(18) The proposal to synthesize additional multiply substituted arylpyranenamines was at that time contrary to experience, the 3,4-Cl_2, 2,6-Cl_2, 3,5-$(CF_3)_2$, and 2,6-$(OH)_2$ derivatives all being inferior to their monosubstituted congeners.

(19) Partition coefficients were determined experimentally for six series members. The substituent pattern, identification number, experimental log P, and tabulated π (ref 10) are: 4-OH (25), 1.87 (-0.67); 3-OH (24), 1.90 (-0.67); 2-OH (23), 1.78 (-0.67); 4-NHAc (52), 1.26 (-0.97); 4-NH_2 (43), 1.17 (-1.23); intramolecularly H-bonded 3-NHAc-4-OH (82), 1.54 (-1.64). The "log P" for the unsubstituted arylpyranenamine is calculated to be 2.52 with 95% CI of ±0.67, based on the above data including (82) and assuming a slope of 1. Thus, the "optimal log P" for this series of molecules is about -2.5.

(20) Intramolecular hydrogen bonding has a substantial effect on calculated π values. From the data in ref 19, the lipophilicity enhancing effect of hydrogen bonding between a 3-NHAc and a 4-OH substituent is +0.66 log unit.

(21) Similar "optimal σ" dependencies are reported in ref 16b and by K. H. Buchel and W. Draber, *Adv. Chem. Ser.*, 114, 141 (1972).

(22) The effect of intramolecular H bonding in increasing lipophilicity and thereby lowering potency makes a contribution to the "±HB-INTRA" term, since the $\sum\pi$ estimates did not consider such H bonding. For the 3-NHAc-4-OH derivative, the actual log P value is known (ref 19 and 20) and the size of this effect on the calculated pI_{50} can be calculated as -0.22. Apparently, the majority of the "±HB-INTRA" coefficient of -0.86 has another mechanistic source.

(23) This lack of success using the Topliss model can be attributed to its intentional bias toward the more frequent

trend of $+\pi/+\sigma$ and away from the $-\pi/-\sigma^2$ relationship actually derived, and also to the presence of data points which are seriously aberrant even with respect to the $-\pi/-\sigma^2$ relationship. Although the failure of this Topliss tree provides a counterexample to the previous findings of consistent success in retrospective studies using the technique (ref 3), we note that no overall inefficiency in series development seems to have resulted from its use and that the resulting data were satisfactory for the more complex QSAR analyses described in this paper. Consequently, we tend to regard this study as somewhat supportive of use of the Topliss approach, provided that the Topliss models are supplemented by formal QSAR methods as more data are available.

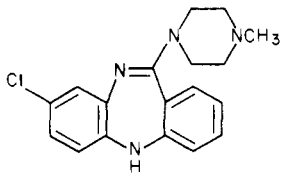
10-(Alkylamino)-4*H*-thieno[3,4-*b*][1,5]benzodiazepines. A Novel Class of Potential Neuroleptic Agents

Jeffery B. Press,* Corris M. Hofmann, Nancy H. Eudy, William J. Fanshawe, Ivana P. Day, Eugene N. Greenblatt, and Sidney R. Safir

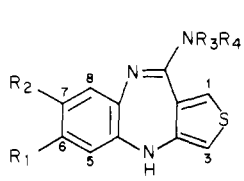
Cardiovascular-CNS Disease Research Section, Medical Research Division of American Cyanamid Company, Lederle Laboratories, Pearl River, New York 10965

An intensive investigation of the structural requirements for CNS activity of the title compounds was undertaken. A synthesis of the precursor dihydro-10*H*-thieno[3,4-*b*][1,5]benzodiazepin-10-ones was achieved and three routes for their conversion to the title compounds were developed. The compounds were tested for neuroleptic activity by means of the blockade of *d*-amphetamine lethality in aggregated mice and/or effects on locomotor activity in rats. Antidepressant activity was examined using inhibition of tetrabenazine-induced depression in mice. Most of the compounds were found to be potent neuroleptic agents with several exhibiting additional antidepressant activity.

As part of an intensive program of developing novel psychotropic agents, we became interested in potential neuroleptic agents which exhibit low extrapyramidal symptoms (EPS). Reports that clozapine (1) was the first



1



2

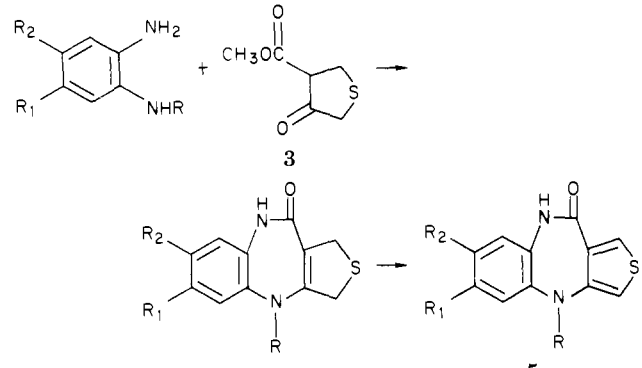
tricyclic antipsychotic drug that did not cause EPS[1] encouraged us to undertake a program of synthesis and investigation of a related tricyclic system, 4*H*-thieno[3,4-*b*][1,5]benzodiazepine[2] (2), as a source of agents with an improved therapeutic profile.

Chemistry. The thieno[3,4-*b*][1,5]benzodiazepin-10(9*H*)-one system (5)[3] was prepared as outlined in Chart I. The appropriately substituted *o*-phenylenediamine was condensed with methyl tetrahydro-4-oxo-3-thiophenecarboxylate[4] (3) in refluxing toluene or xylene to give a 1,3,4,9-tetrahydrothieno[3,4-*b*][1,5]benzodiazepin-10(9*H*)-one (4). Oxidation of 4 could be effected by fusion with elemental sulfur, treatment with sulfuryl chloride in refluxing chloroform, or, preferably, treatment with *N*-chlorosuccinimide in pyridine at 60 °C to give the fully aromatic thiophene derivative 5. When the above sequence was performed with unsymmetrical *o*-phenylenediamines ($R_1 \neq R_2$ or R = CH_3), separation of the isomeric products was most easily accomplished via fractional crystallization of the oxidized derivatives 5.

The desired title compounds 2 were prepared from the lactams 5 in one of three ways as summarized in Chart II. In method A, the principal method used for the preparation of the final compounds 2, lactam 5 was converted by the action of phosphorus pentasulfide in pyridine to a thiolactam 6. Alkylation of 6 with methyl iodide or dimethyl sulfate and base gave the methyl thioether 7, which reacted smoothly with a variety of amines to give the desired 2.

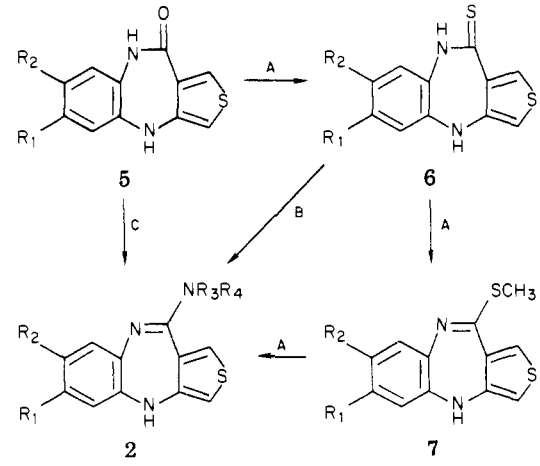
Alternatively, reaction of thiolactam 6 directly with the appropriate amine (e.g., *N*-methylpiperazine) gave 2 in moderate yield (method B). A more useful procedure is that reported by Fryer et al.[5] and is depicted as method

Chart I. Synthesis of Thieno[3,4-*b*][1,5]benzodiazepin-10(9*H*)-ones



R = H, CH_3

Chart II. Synthesis of 10-(Alkylamino)-4*H*-thieno-[3,4-*b*][1,5]benzodiazepines



method A, P_2S_5/pyr; $(CH_3)_2SO_4$/KOH; R_3R_4NH/heat
method B, R_3R_4NH/heat
method C, $R_3R_4NH/TiCl_4$

0022-2623/79/1822-0725$01.00/0