Journal of Medicinal Chemistry

Volume 18, Number 10 October 1975

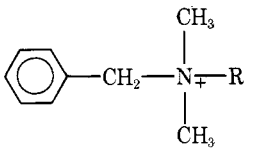
Non-Computer Approach to Structure–Activity Study. An Expanded Fibonacci Search Applied to Structurally Diverse Types of Compounds

Norman J. Santora* and King Auyang

Department of Medicinal Chemistry, William H. Rorer, Inc., Fort Washington, Pennsylvania 19034. *Received April 14, 1975*

The Fibonacci search technique, first applied to a structure–activity study by Bustard, has been expanded to allow the analysis of a broad class of structural types of compounds. The compounds are first arranged in order of increasing value of a molecular property of the analogs such as log P, $\Sigma\pi$, $\Sigma\sigma$, or R_m. A successful Fibonacci search of the compounds will find the most active analog in a small, predetermined number of steps. Examples are given where insight as to mechanism of action is indicated by the combination of various parameters such as log P and pK_a. Additional examples illustrate the use of Fibonacci search to establish the parabolic dependence of the biological activity of lipophilicity and σ, where such dependency had not been observed initially. This technique allows the treatment of a variety of structurally diverse types of compounds simultaneously. It is to be stressed that Fibonacci search can be applied to structure–activity studies without the use of a computer.

Fibonacci search is a non-computer technique which can be used to locate the most biologically active compound in a series of analogs between set limits in a predetermined number of steps. The method was first applied by Bustard,[1] who considered a series of alkylammonium compounds 1 previously reported by Hansch and Clayton.[2] In this model the step size was a single methylene unit from C_7 through C_{20}.



1

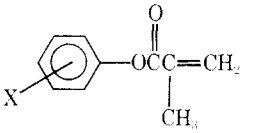
The method described in this paper arranges the compounds in order of increasing value of some molecular property such as log P, $\Sigma\pi$, $\Sigma\sigma$, and R_m; the step size is simply moving from one compound to the next in order of increasing magnitude of the molecular property being considered. It is to be noted that these steps involve varying magnitudes of the molecular property in contrast to Bustard's constant increase of one methylene unit; moreover, our approach can be applied to structurally diverse compounds as well as alkylated side chains. In order to illustrate the Fibonacci search method used in this paper, consider the data of Woodside et al.[3] on **2** in Table I.

The compounds are first rearranged in order of increasing values of the molecular property being considered; in this example the compounds have been rearranged in order of increasing values of log P. Use is made of Table II, column 1, in order to choose an appropriate number of compounds. In this example the data supply eight derivatives; however, nine compounds are needed in order to have seven internal points (Table II, column 1). Consequently, a fictitious point is added at the end of the series. Wilde[4] states that the expedient of adding fictitious points in order to be able to perform a Fibonacci search is safe, simple, and quite effective. Note that the first point number is always designated as 0. The remaining compounds are numbered consecutively through 8 as shown in Table I, column 1.

The first two analogs chosen for synthesis are selected by referring to Table II, columns 1 and 2. Since there are seven internal points in this example, one would choose for synthesis the analogs corresponding to point numbers 3 and 5. The biological activity of these analogs is determined. Assuming that one of these analogs is greater in activity than the other, a new set of limits is established. The number of internal points within the limits is predetermined by Table II, column 1. The activity (3.76) of the analog corresponding to point number 5 is greater than the activity (3.49) of the analog corresponding to point number 3. Consequently, the new set of limits for the maximum is between points 8 and 3. The internal number of points between these limits is 4, which corresponds to the interval predetermined by Table II, column 1. The limit is expressed mathematically: $3 \leq \max \leq 8$.

Use is made of Table II, columns 1 and 2, in order to select a pair of points down the Fibonacci sequence (1, 2, 3, 5, 8, 13, . . .) from the number of internal points between the set of limits just determined. These numbers are added to the lower limit point number in order to generate two new point number analogs for synthesis. Note that this process always generates a point number previously considered. In this example, the number of internal points between the limits, $3 \leq \max \leq 8$, is 4. Table II, columns 1 and 2, provides the next number pair down the Fibonacci sequence, 2 and 3. These numbers are added to the lower limit point number 3 to provide the next two point number analogs: $3 + 2 = 5$ and $3 + 3 = 6$. Thus, the two new analogs to be considered are those corresponding to point numbers 5 and

Table I. Data from Woodside[3] et al. on Either Bacteriostatic or Fungistatic Minimum Inhibitory Concentration of 2 Arranged in Order of Increasing Log P Values



2

Point no.	X	Log P	pC
0	H	1.99	2.89
1	4-Cl	2.69	3.25
2	2-Cl	2.75	3.08
3	2,4-Cl	3.45	3.49
4	2,4,5-Cl_3	4.21	3.84
5	2,4,6-Cl_3	4.21	3.76
6	2,4,5,6-Cl_4	4.97	3.97
7	Br_5	6.74	3.64
8			

Table II. Numerical Relationships within the Fibonacci Search

No. of internal points	Points corresponding to the analogs to be made first	Total no. of analogs required to locate max
2	1 and 2	2
4	2 and 3	3
7	3 and 5	4
12	5 and 8	5
20	8 and 13	6
33	13 and 21	7
54	21 and 34	8

6. Note that the analog corresponding to point number 5 has been considered in the previous pair of analogs.

The latter process is repeated until the Fibonacci search is completed in the predetermined number of steps, Table II, column 3, as reported in Table III. Thus, the Fibonacci search for the data in Table I has determined correctly that the most active analog corresponds to point number 6.

We have found also that the addition of the electronic parameter, σ, or the acidic dissociation constant, pK_a, to the lipophilic parameter, log P or π, can give results superior to those considering log P or π alone. These results indicate that the Fibonacci search technique, as described in this paper, can be used to give insight as to the possible mechanism of biological activity. It is possible that such results may be due to curve smoothing or some other factor unrelated to biologic activity; however, the retrospective examples we have reported on required the same combinations of parameters in Hansch analyses in order to give statistically meaningful results. We are currently at work on this problem and hope to resolve the question in a future publication.

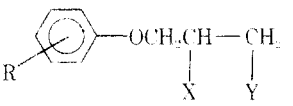
In order to illustrate the use of this technique, consider the phenoxy ethers 3 studied by Berger[5] and reported by Hansch and Lien.[6] The data for the Fibonacci search are presented in Table IV. A Fibonacci search on these analogs as a function of log P alone identified the compound corresponding to point number 19 as being the most active. Although the latter analog is not the most potent analog, it is still very active. However, the most active compound is unambiguously determined when the analogs are arranged in order of increasing $\Sigma(\log P + \sigma)$. Table V displays the analogs in the latter order. A Fibonacci search of these analogs identifies correctly the analog corresponding to point number 19 as the most active derivative. Note that if point numbers 19 and 20 are interchanged, the resulting Fibonacci search still leads to the same compound.

The Fibonacci search technique can be applied to a set of structurally diverse compounds if they show nonspecific mechanism of activity. This point is illustrated by considering six different structural classes of compounds studied individually by Hansch et al.[7] as hypnotics. Table VI displays these compounds listed in order of increasing log P values, and Table VII gives the structural types and frequency of usage of the compounds studied. The most active compound, 24, was found; however, when a different set of data was studied, the most active compound was not found. But the compound was among the most active. This again illustrates that if more than one maximum exists, the Fibonacci search technique will find one of them.

Table III. Fibonacci Search for the Data in Table I

Point no.	pC	Point no.	pC	Limits
3	3.49	5	3.76	3 ≤ max ≤ 8
5	3.76	6	3.97	5 ≤ max ≤ 8
6	3.97	7	3.64	5 ≤ max ≤ 7

Table IV. Data of Berger et al.[5] on Minimum Fungistatic Concentration of 3, Reported by Hansch and Lien,[6] Arranged in Order of Increasing Value of Log P



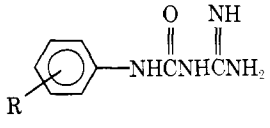
3

Point no.	X	Y	R	Log P	Log 1/C
0	OH	OH	2-Me	1.38	2.26
1	OH	OH	4-Cl	1.40	2.31
2	OH	OH	2,6-Cl_2	1.88	2.37
3	OH	OH	2,4-Cl_2	1.99	2.61
4	OH	H	2-Cl	2.05	2.84
5	OH	OH	4-Cl, 2-Me	2.08	2.33
6	OH	H	2-Me	2.14	2.46
7	OH	H	4-Cl	2.16	2.81
8	H	OH	2-Me	2.34	2.79
9	H	OH	4-Cl	2.36	3.07
10	OH	OH	2,6-Cl_2, 4-Me	2.40	3.10
11	OH	OH	4-Cl, 3,5-Me_2	2.42	3.24
12	OH	H	2,6-Cl_2	2.64	3.04
13	OH	H	4-Cl, 3-Me	2.67	3.30
14	OH	H	2,4-Cl_2	2.75	3.35
15	OH	OH	4-Cl, 2,6-Me_2	2.76	2.76
16	OH	H	4-Cl, 2-Me	2.84	3.30
17	OH	H	2,6-Cl_2, 4-Me	3.16	3.47
18	OH	H	4-Cl, 3,5-Me_2	3.18	3.68
19	H	OH	2,6-Cl_2, 4-Me	3.36	3.67
20	H	OH	4-Cl, 3,5-Me_2	3.38	3.93
21	OH	H	4-Cl, 2,6-Me_2	3.52	3.51

Table V. Data of Berger et al.[5] on Minimum Fungistatic Concentration of 3, Reported by Hansch and Lien,[6] Arranged in Order of Increasing Value of $\Sigma(\text{Log } P + \sigma)$

Point no.	X	Y	R	Σ(log $P + \sigma$)	Log $1/C$
0	OH	OH	2-Me	1.24	2.26
1	OH	OH	4-Cl	1.63	2.31
2	OH	H	2-Me	2.00	2.46
3	OH	OH	4-Cl, 2-Me	2.17	2.33
4	H	OH	2-Me	2.20	2.79
5	OH	H	2-Cl	2.26	2.84
6	OH	OH	2,6-Cl_2	2.30	2.37
7	OH	H	4-Cl	2.39	2.81
8	OH	OH	2,4-Cl_2	2.43	2.61
9	OH	OH	4-Cl, 3,5-Me_2	2.51	3.24
10	H	OH	4-Cl	2.59	3.07
11	OH	OH	2,6-Cl_2, 4-Me	2.65	3.10
12	OH	OH	4-Cl, 2,6-Me_2	2.71	2.76
13	OH	H	4-Cl, 3-Me	2.83	3.30
14	OH	H	4-Cl, 2-Me	2.93	3.30
15	OH	H	2,6-Cl_2	3.06	3.04
16	OH	H	2,4-Cl_2	3.19	3.35
17	OH	H	4-Cl, 3,5-Me_2	3.27	3.68
18	OH	H	2,6-Cl_2, 4-Me	3.41	3.47
19	H	OH	4-Cl, 3,5-Me_2	3.47	3.93
20	OH	H	4-Cl, 2,6-Me_2	3.47	3.51
21	H	OH	2,6-Cl_2, 4-Me	3.61	3.67

The nature of the Fibonacci search technique is dependent upon an extremum between limits. In practice, however, it will find one among several maxima or minima.[8] This property of the Fibonacci search enables it to elicit whether a parabolic relationship on log P or π is operative. For example, Goodford et al.[9] carried out a Hansch analysis on 19 arylamidinoureas 4 as antimalarials. They arrived at



4

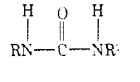
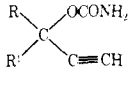
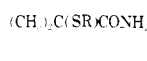
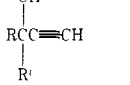
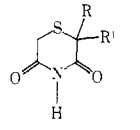
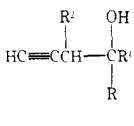
the following equation: B.A. = $0.28\pi + 0.86\sigma - 0.07$. They stated that the range covered by the compounds was not great enough to define a quadratic relationship; however, a Fibonacci search of these compounds located the most active analog (Table VIII). The latter result is irrefutable evidence that a parabolic relationship of activity with π does indeed exist! In the same regard, Hansch and Fujita[10] analyzed only 8 of 14 diethylphenyl phosphates in their toxicity to houseflies and found that the activity was primarily dependent upon σ, whereas its dependence upon π was very slight. A Fibonacci search of these compounds (Table VIII) located the most active derivative. This is another example that the activity is actually parabolically dependent upon π. Of course, if the effect of π is slight, it could be left out of a Hansch equation. A more important advantage of the Fibonacci search technique as applied to this problem is that all 14 of the compounds were used, whereas Hansch used only 8, stating that one compound was left out because of possible difference in mode of action. We are in the process of examining other cases where linear relationships between lipophilic character and biological activity have been reported in order to ascertain whether a parabolic dependence on π or log P more accurately describes the activity.

Table VI. Data of Hansch et al.[7] Arranged in Order of Increasing Values of Log P

Point no.	Compd no.[a]	Log P	Log $1/C$[b]
0	163	−0.10	1.84
1	157	0.32	2.28
2	164	0.40	2.06
3	103	0.50	2.40
4	126	0.59	2.74
5	162	0.80	2.58
6	112	0.88	2.41
7	125	0.89	2.86
8	166	0.90	2.23
9	136	1.05	2.44
10	127	1.09	3.11
11	159	1.12	2.68
12	116	1.18	2.65
13	129	1.21	3.28
14	158	1.32	2.91
15	134	1.35	2.50
16	113	1.38	2.79
17	128	1.39	3.31
18	169	1.40	2.35
19	104	1.50	2.97
20	161	1.52	2.70
21	141	1.56	2.39
22	139	1.65	2.70
23	120	1.68	2.92
24	130	1.71	3.32
25	137	1.85	2.70
26	171	1.90	2.38
27	140	1.95	2.54
28	122	2.00	3.20
29	132	2.01	3.13
30	142	2.03	2.30
31	138	2.15	2.67
32	131	2.21	3.00
33	108	2.27	2.67
34	115	2.38	2.65

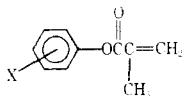
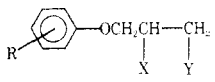
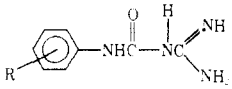
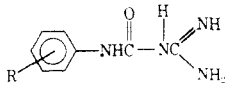
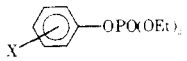
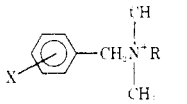
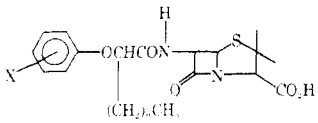
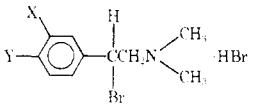
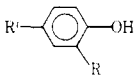
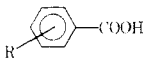
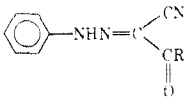
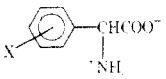
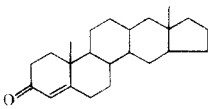
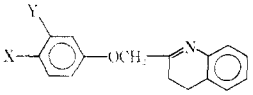
[a] These are the identifying numbers of the compounds studied by Hansch et al.[7] [b] C represents the moles of drug per kilogram of test animal producing "hypnosis".

Table VII. Structural Types of Compounds Examined in Table VI

Structural types used	Frequency of usage	Structural types used	Frequency of usage
	5		8
	5		6
	3		8

In summary, the Fibonacci search technique as described in this report has the ability to study a variety of com-

Table VIII. Summary of Fibonacci Searches Studied

Structure	Type of act.	No. of compds	No. of compds to find max	Parameters	Ref
	Bactericidal and antifungal	8	4	Log P	3
	Antifungal	22	6	Log P and σ	5
	Antimalarial	22	6	π	9
	Antimalarial	14	5	π	9
	Toxicity to houseflies	14	5	π	10
	Bactericidal	35	7	Log P	11
	Bactericidal	22	6	π	12
	Adrenergic blocking agents	22	6	π	13
	Bactericidal	14	5	Log P	14
	Hapten antibody interaction	22	6	π	15
	Uncoupling of oxidative phosphorylation	35	7	π and pK_a	16
	Enzymic oxidation	14	5	π	17
	Hemolytic	14	5	R_m	18
	Neuramidase inhibition	14	5	σ	19
Polyene alkyl esters	Antibiotic	20	7	R_f	20

pounds in combination or separately. The inclusion of σ and pK_a parameters to the lipophilic parameter may provide insight as to mechanism of action, and a successful search clearly demonstrates the parabolic relationship toward the molecular property being considered. The method is fast and easy to use and, although retrospective cases were studied, the technique is quite easily amenable to use in the beginning stages of an investigation. It is to be emphasized that this is a rapid, non-computer technique.

Table VIII summarizes several of the cases in which we have applied the Fibonacci search.

Acknowledgment. The authors wish to thank Dr. Arthur Cammarata for his encouragement and many helpful discussions.

References and Notes

(1) T. M. Bustard, *J. Med. Chem.*, 17, 777 (1974).
(2) C. Hansch and J. M. Clayton, *J. Pharm. Sci.*, **62**, 1 (1973).
(3) R. Woodside, M. Zief, and G. Sumrell, *Antibiot. Chemother.*, **9**, 470 (1959).
(4) D. J. Wilde, "Optimum Seeking Methods", Prentice-Hall, Englewood Cliffs, N.J., 1964, Chapter 2.

(5) F. M. Berger, C. V. Hubbard, and B. J. Ludwig, *Appl. Microbiol.*, 1, 146 (1953).
(6) C. Hansch and E. J. Lien, *J. Med. Chem.*, 14, 653 (1971).
(7) C. Hansch, A. R. Steward, S. M. Anderson, and D. Bentley, *J. Med. Chem.*, 11, 1 (1968).
(8) L. C. W. Dixon, "Nonlinear Optimisation", Crane, Russack and Co., New York, N.Y., 1972, Chapter 1.
(9) P. J. Goodford, F. E. Norrington, W. H. G. Richards, and L. P. Walls, *Br. J. Pharmacol.*, 48, 650 (1973).
(10) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, 86, 1616 (1964).
(11) S. Ross, C. E. Kwartter, and J. H. Baile, *J. Colloid Sci.*, 8, 385 (1953).
(12) C. Hansch and A. R. Steward, *J. Med. Chem.*, 7, 691 (1964).
(13) C. Hansch and E. J. Lien, *Biochem. Pharmacol.*, 17, 709 (1968).
(14) E. Klarmann, V. A. Shternov, and L. W. Gates, *J. Am. Chem. Soc.*, 55, 2576 (1933).
(15) E. Kutter and C. Hansch, *Arch. Biochem. Biophys.*, 135, 126 (1969).
(16) K. H. Büchel and W. Draber, "Biological Correlations—The Hansch Approach", American Chemical Society, Washington, D.C., 1972, Chapter 10.
(17) C. Hansch and R. Kerley, *J. Med. Chem.*, 13, 957 (1970).
(18) G. L. Biagi, M. C. Guerra, and A. M. Barbaro, *J. Med. Chem.*, 13, 944 (1970).
(19) M. S. Tute, *J. Med. Chem.*, 13, 48 (1970).
(20) T. Bruzzese, M. Cambieri, and F. Recusani, *J. Pharm. Sci.*, 64, 463 (1975).

The Role of Anionic, Imidic, and Amidic Forms in Structure–Activity Relationships. Correlation of Electronic Indices and Bacteriostatic Activity in Sulfonamides

Augusto Rastelli,* Pier G. De Benedetti, Giovanna Gavioli Battistuzzi,

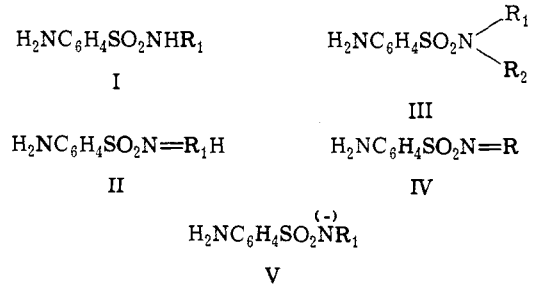
Istituto di Chimica Fisica dell'Università di Modena, 41100 Modena, Via Campi 183, Italy

and Albano Albasini

Istituto di Chimica Farmaceutica e Tossicologica dell'Università di Modena, 41100 Modena, P. le S. Eufemia 19, Italy.
Received March 17, 1975

The problem of structure–activity relationships in sulfonamide type compounds is tackled on the ground that both bacteriostatic activities and structural indices must be referred to the specific individual forms assumed by sulfa drugs in the active solutions. The frequency value of the symmetric stretching mode of the sulfonyl group ν_s (SO_2) is chosen as a suitable electronic index and measured for the individual active forms in aqueous and Me_2SO solutions. The *linear* correlation that exists between bacteriostatic parameter and vibration frequency (over the complete range of data at present available) proves a strict relationship between electronic structure and bacteriostatic activity in this class of drugs. Furthermore, it justifies the assumption used for the calculation of the bacteriostatic activity of the anionic form; i.e., in equilibrium with a very active species (the anion) a less active species (the neutral form) gives a negligible contribution or does not contribute at all to the total activity. The results can be summarized as follows: *the lower the frequency of the symmetric stretching mode of the SO_2 group of any active species of sulfonamide type compounds, the higher its bacteriostatic activity.* The existence of a clear structure–activity correlation demonstrates that the whole class of compounds, whatever their form, has a single mechanism of action, while incontrovertible deviations from the general trend indicate differences or complications in the mechanism itself, but does not demonstrate that the group on which the structural index is localized plays a dominant role in the biological process. The usefulness of pK_a and NH_2 proton chemical shift of precursor amine as indirect indices of the electronic structure of the anionic forms is explored on extensive sets of available data.

Sulfa drugs antagonized by *p*-aminobenzoic acid (PABA) contain the common moiety *p*-$NH_2C_6H_4SO_2$ which can be considered the minimal structural condition for their bacteriostatic action; the substituents bonded to the SO_2 group therefore serve only to determine the actual values of bacteriostatic activity. In *p*-aminobenzenesulfonamides there is a substituted NH_2 group attached to the SO_2 group; according to the number and type of substituents on this NH_2 group one can distinguish the following structural situations I–V.



The isomeric forms I (amide) and II (imide), which can coexist in aqueous solution, have acid properties and give the common anion V in alkaline solution. These structures (I, II, and V) have markedly different electronic properties, as is demonstrated by their spectroscopic behavior[1–4] which can in turn be used as a tool for analytical purposes. It shows, for example, that in aqueous solutions 2-*p*-aminophenylsulfonamidopyridine (sulfapyridine) and -thiazole (sulfathiazole) assume forms II and V, their ratio depending on the pH of the solution; N^1-phenylsulfanilamide and 2-*p*-aminophenylsulfonamidopyrimidine (1) (sulfapyrimidine), on the other hand, assume forms I and V. *p*-Aminobenzenesulfonylguanidine (sulfaguanidine) has been assigned structure IV by comparison with some ring *N*-methyl derivatives[4] which must necessarily have that structure, whereas structure III is obviously assumed by all the N^1-methyl or N^1-acetyl derivatives of sulfa drugs pertaining to forms I or II.

Although these structures have long been the subject of research,[5–8] their role in structure–activity relationships does not appear to have been fully understood. In fact, even if the anion structure V is commonly claimed to be the most active, the neutral form I has usually been preferred for measuring[9–15] or calculating[16,17] the structural indices. If it is important that the index chosen as representative of electronic structure be related to realistic events (energetically) in which the drug may participate and that, when localized on a particular position or functional group on a molecule, it should reflect a chemical event actually taking