Optimization of Alkyl Modifications by Fibonacci Search

Thomas M. Bustard†

Scientific Division, Abbott Laboratories, North Chicago, Illinois 60064. *Received December 3, 1973*

Medicinal chemists are frequently faced with modifying an alkyl, alkoxy, *N*-alkyl, etc., side chain of a lead compound in order to maximize its activity if possible. Many classes of compounds do in fact manifest a maximum activity with respect to alkyl side-chain length. This behavior has been the subject of a recent review in which many examples of physiologically active molecules exhibit maxima with respect to the partition coefficient.[1]

Fibonacci search is a technique which ensures that an optimum, which is known or presumed to exist between certain limits, can be found in the fewest number of function evaluations to within the magnitude of the step size.[2] In the context of drug design, function evaluation is synonomous with the synthesis and testing of a given analog; the step size is a single methyl or methylene group. The fewer the number of compounds made and tested before the optimal one is found, the lower the cost of drug research.

In order to use the method, an appropriate interval is chosen within which a compound of maximal activity is likely to exist, *i.e.*, chain lengths through C_{12} or perhaps alkyl substitutions on the molecule up through 12 carbons. Two points about the center of the interval are chosen to be evaluated. Depending upon which one is greater, a new and smaller interval is defined containing the maximum. Two new points are selected for evaluation about the midpoint of the new interval. However, one of these two points will already have been evaluated in the previous step. This latter step is done repeatedly until the search ends in a predetermined number of evaluations. Note that each step corresponds roughly to halving the remaining interval. The exact manner in which interval lengths and points about the midpoint are chosen depends on the Fibonacci sequence which is conisdered in detail below.

The Fibonacci sequence is 1, 2, 3, 5, 8, 13, 21, . . . , in which each number in the sequence beyond the second is given by the sum of the two previous ones. Let us consider a simple case to illustrate how the search works before applying it to a drug design problem. Consider a function $f(x)$, which we assume reaches a maximum somewhere in the interval $0 \leq x \leq 5$. *Let the step size be 1, so that there are four internal points between 0 and 5. We select two points* $x = 2$ and $x = 3$ at which to evaluate $f(x)$. Note that 5, 3, and 2 are Fibonacci numbers. Now if $f(2) > f(3)$, the maximum must lie in the interval defined by $0 \leq x \leq 3$. On the other hand, if $f(3) > f(2)$, the maximum must lie in the different interval $2 \leq x \leq 5$. We now know that the maximum lies in an interval which contains only two internal points. For concreteness, let us assume that the maximum lies in the interval $2 \leq x \leq 5$. The next Fibonacci number down the sequence is 1, which indicates the next point to evaluate. The point at $x = 3$ has already been evaluated and it lies 1 (a Fibonacci number) unit from the lower limit. The other number must then be 1 unit from the upper limit or at $x = 4$. Now, if $f(4) > f(3)$, the maximum lies in the interval $3 \leq x \leq 5$; if $f(3) > f(4)$, the maximum lies in the interval $2 \leq x \leq 4$. The latter case is represented in Figure 1 in order to make the discussion more obvious. The fact that $f(x)$ may be irregular

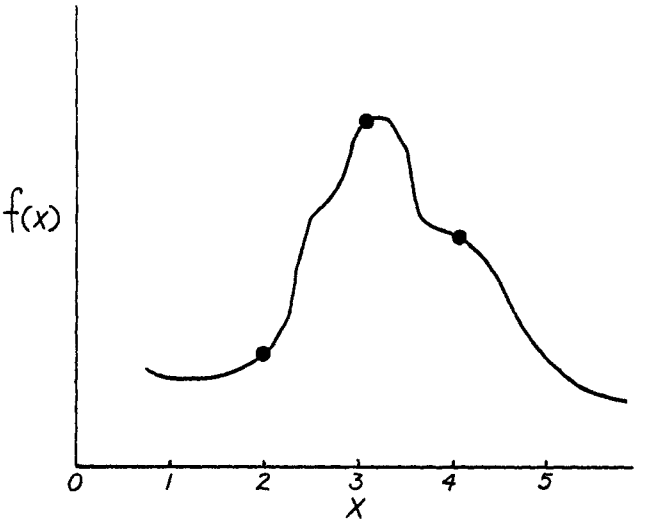


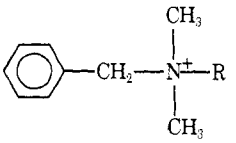
Figure 1. Some arbitrary function which exhibits a maximum somewhere in the interval $0 \leq x \leq 5$.

does not affect the convergence, but the search can only be applied over a unimodal (single maximum) range. If there is more than one maximum, the search will usually find one of them.

There are many search techniques available for use in this type of problem.[3] The Fibonacci method is the most efficient for general functions.[2] It has the additional advantage that "biological noise" in pharmacological assays is minimized by choosing points which are not close to one another in the early stages of the search. Dichotomous search[3] is another efficient technique, but pairs of points must be as close as is possible for maximum efficiency. Clearly, an error in relative pharmacological activity at an early part of this search would lead to the wrong part of the interval to be investigated in future steps.

Whenever a side-chain optimization problem occurs, there are usually limits set on its length. If the Fibonacci technique is being used, then limits can only have certain values. In the previous example, we needed four internal points between the limits 0 and 5, or more generally x_0 and x_5. The next Fibonacci number is 8, which requires seven internal points between x_0 *and* x_8. Table I illustrates how the search is related to the number of internal points and how the first two analogs are to be chosen.

Consider a series of alkylammonium compounds 1 reported by Hansch and Clayton[1] as an example to which we can apply the Fibonacci search. The data are summarized in Table II. In this series, the limits are between C_7 and C_{20} with 12 internal points.



1

The Fibonacci search proceeds as in Table III. Each row in this table represents the outcome of successively smaller magnitude searches. The first considers 12 internal points, the second considers seven internal points and so forth as in Table I. The Fibonacci numbers show where in the interval the next point should be chosen. In the second row, for example, points 8 and 10 are chosen to be three and five units away from the end points of the interval (C_{12}, C_{20}). Note that after the first pair of points is

†Address correspondence to this author at 15020 N. 24th Ave., Phoenix, Ariz. 85023.

Table I. Numerical Relationships within the Fibonacci Search

No. of internal points	Points which correspond 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

Table II. Data of Cutler, *et al.*, of compounds **1** as reported by Hansch and Clayton

No.	R	pC[a]	No.	R	pC[a]
0	C_7		7	C_{14}	5.12
1	C_8	3.00	8	C_{15}	5.14
2	C_9	3.02	9	C_{16}	5.16
3	C_{10}	3.57	10	C_{17}	4.70
4	C_{11}	4.06	11	C_{18}	4.71
5	C_{12}	4.61	12	C_{19}	4.73
6	C_{13}	5.10	13	C_{20}	

[a] pC is the negative logarithm of the MIC against *Clostridium welchii*.

Table III. Fibonacci Search for the Data in Table II

Point no.	Activity	Point no.	Activity	Limits
5 (C_{12})	4.61	8 (C_{15})	5.14	$C_{12} \leq$ max $\leq C_{20}$
8 (C_{15})	5.14	10 (C_{17})	4.70	$C_{12} \leq$ max $\leq C_{17}$
7 (C_{14})	5.12	8 (C_{15})	5.14	$C_{14} \leq$ max $\leq C_{17}$
8 (C_{15})	5.14	9 (C_{16})	5.16	$C_{15} \leq$ max $\leq C_{17}$

chosen, only one additional point is required in each successive step.

Observe that there were 12 analogs in this series but that the optimal one could be found by making only five analogs. Moreover, most of the compounds indicated lie in the neighborhood of the maximum. None of the other compounds in the series were required to locate the maximum.

The search can actually extend over as large a chain length as desirable, which may be practically determined by the number of compounds which are feasible to make and test. Competing physical processes such as chain folding, intramolecular hydrophobic bonding, micelle formation, etc., impose limits on the length of alkyl side chains to be considered. The maximum number of analogs to be made before an optimal one is found is listed in Table I for various search intervals.

Since we have been discussing a *molecular property* which is roughly proportional to the number of CH_2 or CH_3 groups present, the technique applies not only to straight chains at one particular site but generally to multiple substitutions, chain branching, and cyclic substituents. The only requirement is that there be a maximum within the chosen interval.

In summary, this search method will find the most active analog of any series provided that the biological activity exhibits a maximum with respect to some molecular property. The exact form of the relationship is immaterial. In the case of alkyl substitutions, the molecular property involved is most likely to be the partition coefficient. However, the molecular property need not be considered explicitly since properties such as the partition coefficient or molecular refractivity are roughly proportional to the number of CH_2 or CH_3 groups present.

References

(1) C. Hansch and J. M. Clayton, *J. Pharm. Sci.*, **62**, 1 (1973).
(2) L. C. W. Dixon, "Nonlinear Optimisation," Crane, Russak and Co., New York, N. Y., 1972, Chapter 1.
(3) D. J. Wilde, "Optimum Seeking Methods." Prentice-Hall. Englewood Cliffs, N. J., 1964, Chapter 2.

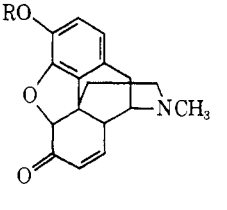
Preparation of Morphine-6-^3H and Its Isotopic Stability in Man and in Rat

Jack Fishman,* Baiba Norton, MaryLou Cotter, and Elliot F. Hahn

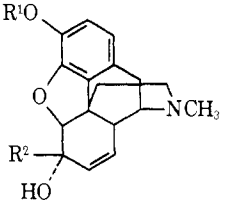
Institute for Steroid Research, Montefiore Hospital and Medical Center, Bronx, New York 10467. *Received February 15, 1974*

Morphine labeled with radioactive isotopes is essential to the detailed study of the biological disposition and biochemistry of this premier narcotic alkaloid. In order to permit tracing and identification of all its transformation products, the isotope in the labeled material must be located in a biologically stable position. To allow for the investigation of specific protein binding of morphine and in particular to aid in the search for the putative narcotic receptor, radioactive substrates with high specific activity are required. In addition, the recent advent of radioimmunoassay for morphine creates the need for high specific activity radiolabeled material.[1,†] Presently available labeled morphines have severe limitations by the above criteria. The initial radiolabeled alkaloid, morphine-*N*-$^{14}CH_3$,[2] suffers loss of the isotope in the course of N-demethylation, a significant biotransformation of morphine.[3] The specific activity of the material is also limited by the nature of the isotope to a maximum of about 50 mCi/mmol. Biological stability limitations apply also to the recently described morphine-*N*-CT_3 which has been prepared with a specific activity of 68 mCi/mmol.[4] Randomly labeled morphine-^3H has been obtained *via* tritium exchange by several procedures including acid catalysis[5] and microwave irradiation.[6] While apparently biologically stable, these substrates have reported maximum specific activities of less than 10 mCi/mmol rendering them of limited value for the purposes detailed above. Selective tritium introduction into the morphine molecule has been achieved in the preparation of morphine-2-^3H,[7,8] but the specific activity of the product is only of the order of 12 mCi/mmol.

Chemistry. It appeared to us that the best route to high specific activity radiolabeled morphine required the specific introduction of tritium by means of a one-step



Ia = H
b = Ac
c = CH_3



IIa, R^1 = Ac; R^2 = H
b, R^1 = Ac; R^2 = ^3H
c, R^1 = H; R^2 = ^3H
d, R^1 = CH_3; R^2 = ^3H

reaction on a suitable substrate. The reduction of the 6-keto derivative of morphine, morphinone (Ia),[9] with a labeled reagent to effect introduction of tritium into the 6β

† In the absence of high specific activity morphine, dihydromorphine-7,8-^3H (sp act. 388 mCi/mmol) was used as a substitute.