On the Parabolic Relationship between Drug Potency and Hydrophobicity

JAMES W. MCFARLAND

Pfizer Medical Research Laboratories, Groton, Connecticsd - 06740

Received July 7, DGO

A new derivation of the Hansch-Enjita relationship between drug potency and hydrophobicity is presented. The argument is based on probability concepts and from this vantage shows that the relationship is not, strictly speaking, parabolic, but it may be considered to be so for all practical purposes. This approach to the problem allows one to consider multiple and branched routes of drug travel simultaneously; no change in the basic relationship is incurred thereby. Further analysis of these results show that practically all drug molecules that reach their receptors travel by equivalent most-probable routes; molecules traveling by other routes usually do not result in drug-receptor complexes.

In 1964 Hansch and l'ujita suggested that the probability of a drug penetrating a cell membrane by a passive transport mechanism is a parabolic function of the drug's hydrophobic bonding capacity.¹ These workers also proposed the new extrathermodynamic substituent constant π as a measure of hydrophobicity. The definition of π parallels that for σ in the Hammett relationship:

$$\log (k/l)_{\mathrm{X}} - \log (k/l)_{\mathrm{H}} = \pi \tag{1}$$

In eq 1 $(k/l)_{\rm X}$ is the partition coefficient² for a member of a drug series which bears the substituent X while $(k/l)_{\rm H}$ is the partition coefficient of the unsubstituted parent compound; a standard solvent pair is used to determine these values. The Hansch-Fujita relationship is given either by eq 2 or by eq 3. The potency

$$\log 1/C_X = a\pi_X^2 + b\pi_X + c \tag{2}$$

$$\log 1/C_{\rm X} = a[\log (k/l)_{\rm X}]^2 + b \log (k/l)_{\rm X} + c \quad (3)$$

index, log $1/C_{\rm X}$, is considered to be directly proportional to the logarithm of the probability that the drug will cross one or more cell membranes in a given system. The concentration $C_{\mathbf{X}}$ is that resulting in some standard biological response, e.g., ED₉₀ and LD₅₀. The constants a, b, and c are characteristics of the biological test system. The initial argument for eq 2 and 3 was intuitive,¹ but recently the Hansch school has advanced a kinetic argument based on a model of alternating aqueous and lipid (membrane) phases.² Using a computer they found solutions to a set of complex differential equations for selected values of k/l. Values for log C at an arbitrarily chosen receptor region were obtained. A plot of these data against log k/l gave a set of points which were fitted to a parabolic curve by the method of least squares. The fit was not perfect, and the set of points was skewed slightly away from the maximum of the regression curve, but for all practical purposes it could be said that the parabolic relationship under discussion was confirmed by the mathematical analysis of a suitable model system. While working on a different approach to this problem, I independently struck upon a new derivation of this parabolic relationship which to my mind is easier to grasp since only an understanding of elementary probability concepts and algebra is required. This approach has other advantages which will become apparent further on.

For most drugs, getting to the receptor site is a chance affair. A molecule given at the site of administration normally wanders randomly about in some initial aqueous phase, and in the course of time is absorbed by the lipid phase of some cell membrane. At this point it may either be reabsorbed by the initial aqueous phase or it may penetrate the membrane by being extracted by the aqueous phase on the other side. This process is repeated a number of times until the drug molecule finally reaches the aqueous phase from which the drugreceptor complex can be formed directly in one step. Figure 1 illustrates this condition. The figure is not



Figure 1.—Model of alterating aq phases and lipid-like membranes in a bypothetical biological system. The meanings of the various symbols are explained in the text.

meant to represent actual structures of cell membranes, but rather designates the physical situation important to the passive transport of drugs, *i.e.*, the relationship between the penetration of a drug to its partition coefficient which has been recognized in various forms since the pioneering work of Meyer and Overton.³

In Figure 1, aq_6 represents the aqueous phase where the drug is initially applied; hp_1 represents the first lipid phase the drug enters. At this point the drug can either return to aq_6 or enter aq_2 . The interfaces are coded at the top of the Figure, 1, 2, 3, 4, $\dots n = 2, n = 4, n$, such that aqueous phases are designated by even numbered subscripts while the lipid phases are designated by odd; thus the final aqueous phase from which the drug-receptor complex can be formed will be aq_n where n is an even number. The rate constant which determines how readily an uncharged molecular species at a

⁽¹⁾ C. Hansele and T. Fujita, J. Anov. Chem. Soc., 86, 1616 (1964).

⁽²⁾ J. T. Penniston, L. Beckett, D. L. Bentley, and C. Hanselt, Mol. Pharmacol., 5, 333 (1969). A number of other symbols have been proposed for the partition coefficient; Hanseh for example uses P. However, because the present discussion leans heavily on probability arguments. P shall be reserved as a symbol for probability. The Hanseh school has recently proposed k/l as an alternate symbol, and it will be used in the present discussion.

 ^{(3) (}a) H. H. Meyer, Acch. Exp. Pathyl. Pharmockol., 42, 109 (1896).
 (b) E. Overt(n. Viertelightwisselie, Nuturfaysch. Gev. Zutrich. 44, 88 (1899).

given concentration will enter a lipid phase is k while the rate constant for the reverse process is l. The partition coefficient is, of course, k/l. It is assumed here that the nature of the intervening aqueous phases are essentially the same as aq_0 , and correspondingly the nature of the intervening lipid phases are essentially the same as lip₁. Thus, regardless where the drug is in the maze of inter- and intracellular fluids and membranes, k and l remain essentially unaltered.

If the rate at which the drug-receptor complex forms is slow with respect to the speed that an equilibrium condition is approached in these intervening phases, then the various lipid barriers will have little influence upon the potency of a drug because the concentration of drug in all aqueous phases will be essentially the same. Thus, under equilibrium conditions, the potency of a drug should be independent of its partition coefficient insofar as getting the drug to the receptor site is the only role the partition coefficient plays. Where the partition coefficient is a direct reflection of a drug's hydrophobic bonding capacity in the drug-receptor complex, then it will once again have an important influence on the drug's potency.

More often than not, however, the drug does not have sufficient time to establish an equilibrium condition in a biological system. Besides getting to its intended target, it also wanders to other corners of the system where it is metabolized or excreted. Only a small fraction of the total drug administered actually reaches the receptor site, and it must do so within a limited time or else the reservoir of drug in aq_0 will be depleted through other losses to the point where it will no longer be able to maintain an adequate concentration at the more distant receptor site aq_n . Under the conditions of determining the potency of a drug, just enough compound is administered to effect a submaximal response. From the foregoing discussion one would expect and frequently observes that the response appears soon after the administration of the drug, *i.e.*, when the drug in the reservoir phase is at its maximum concentration, and then the response fades as the concentration is reduced through various losses. In this situation the standard biological response should occur before a significant amount of drug has left aq_0 . Therefore, it can be assumed that the concentration of drug molecules at the receptor site is quite small with respect to the concentration in the reservoir phase, and it can be calculated by multiplying the concentration in aq_0 times the probability of a drug molecule reaching the receptor site. The concentration in aq_0 is, of course, that which gives the standard response and forms the basis of the potency index, $\log 1/C$. It remains for us to estimate the probability factor. Naturally one should not expect to do this in an absolute sense, but there should be some means to estimate the relative probabilities of the various members of a drug series from the model given in Figure 1. The following is offered as such a method.

The probability of a molecule reaching aq_n from aq_0 is given by eq 4. Eq 4 assumes that there is only one

$$P_{0,n} = P_{0,1} \cdot P_{1,2} \cdot P_{2,3} \cdot P_{3,4} \dots P_{n-2,n-1} \cdot P_{n-1,n} \quad (4)$$

path to the receptor and that there is an exact (although unknown) number of aqueous-lipid interfaces to cross. While it is highly improbable that there is only one pathway available to a drug, it is convenient to develop this equation at this point. Branched and alternate pathways will be discussed later on. As a simplifying assumption we will consider that the probability of a neutral molecule moving from an aqueous phase into a lipid phase is the same throughout the system, such that eq 5 is true.

$$P_{0,1} = P_{2,3} = P_{n-2,n-1} \tag{5}$$

Similarly, the reverse situation is given by eq. 6. By

$$P_{1,2} = P_{3,4} = P_{n-1,n} \tag{6}$$

combining eq 5 and 6 with eq 4, eq 7 is obtained. It

$$P_{0,n} = (P_{0,1})^{n/2} \cdot (P_{1,2})^{n/2} \tag{7}$$

will be assumed here that the probability of a drug transferring from aq_0 to lip_1 is determined at the interface of the phases. Not all drug molecules reaching the interface will succeed in penetrating it; the preference of entering lip_1 or being reflected to aq_0 is determined by the relative values of k and l. When k is larger than l most drug molecules will readily enter lip_1 and a lesser fraction will return to aq_0 ; when l is greater than k the reverse situation obtains. Thus, within a particular time span the number of molecules entering lip_1 will be proportional to k, while the total number of molecules presented the opportunity to enter lip_1 will be proportional to the sum of k and l. Hence, $P_{0,1}$ is given by eq 8. Division of the upper and lower parts of

$$P_{0,1} = \frac{k}{k+l} \tag{8}$$

the fraction in eq 8 by l gives eq 9. In the reverse

$$P_{0,1} = \frac{k/l}{k/l+1}$$
(9)

situation, $P_{1,0}$ is the probability of a molecule in lip₁ getting into aq_0 , and it is given by eq 10. However,

$$P_{1,0} = 1 - P_{0,1} \tag{10}$$

according to the model given in Figure 1, $P_{1,0}$ is not different from $P_{1,2}$, hence:

$$P_{1,2} = 1 - P_{0,1} \tag{11}$$

Substituting this new expression into eq 7 we have:

$$P_{0,n} = (P_{0,1})^{n/2} (1 - P_{0,1})^{n/2}$$
(12)

$$P_{0,n} = \left(\frac{k/l}{k/l+1}\right)^{n/2} \left(1 - \frac{k/l}{k/l+1}\right)^{n/2} \quad (13)$$

$$P_{0,n} = \left(\frac{k/l}{k/l+1}\right)^{n/2} \left(\frac{1}{k/l+1}\right)^{n/2}$$
(14)

$$P_{0,n} = \frac{(k/l)^{n/2}}{(k/l+1)^n} \tag{15}$$

Thus, according to eq 15, the probability of a drug reaching a receptor site is a function of its partition coefficient and the number of intervening aqueous-lipid interfaces between aq_0 and aq_n .

As eq 15 stands it is difficult to envision even qualitatively how the probability is changing as k/l and n vary. However, some insight is gained by examining Table I which gives solutions to eq 15 for selected values of k/land n. In situations where the drug must pass through

TABLE I **PROBABILITY OF DRUG CROSSING** *n* INTERFACES: $P_{0,n}$

					p_{λ} , p				
λ = 0	$u = 1^{b}$	n = 2	n = 3	<i>n</i> = 4	n = ĵ)) = (i	n = 8	ъ == 10) og 1: 1
1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	– Inf
1.000	0.001	0.001	3.2×10^{-5}	10-6	3.1×10^{-8}	10 - 9	10-12	10^{-15}	-3.00
1.000	0.003	0.003	1.6×10^{-4}	8.9×10^{-6}	4.9×10^{-3}	2.7×10^{-8}	8.0×10^{-11}	2.4×10^{-33}	-2.52
1.000	0.010	0.010	9.7×10^{-4}	1.0×10^{-4}	9.5×10^{-6}	9.4×10^{-7}	9.2×10^{-9}	9.1×10^{-11}	-2.01
1.000	0.029	0.028	4.8×10^{-3}	$-8.0 imes 10^{-4}$	$1.3 imes 10^{-4}$	$2.2 imes 40^{-5}$	6.4×10^{-7}	1.8×10^{-8}	-1.52
1.000	0.091	0.083	0.0238	0.007	0.002	5.6×10^{-4}	4.7×10^{-5}	3.9×10^{-6}	-1.00
1.000	0.231	0.178	0.0747	0.031	0.013	0.006	9.9×10^{-4}	$1.8 imes10^{-4}$	-0.52
1,000	0.500	0.250	0.125	0.063	0.031	0.016	0.1004	0.001	1), 00
1.000	0.750	0.188	0.0812	0.035	0.015	0.008	1.2×10^{-4}	2.3×10^{-4}	0.48
1.000	0.909	0.083	0.0238	0.007	0.002	5.6×10^{-4}	4.7×10^{-5}	3.9×10^{-5}	1.00
1.000	0.968	0.031	5.5×10^{-3}	10×10^{-4}	1.7×10^{-4}	3.0×10^{-5}	9.5×10^{-7}	3.0×10^{-8}	1.48
1.000	0.990	0.010	9.7×10^{-4}	1.0×10^{-4}	9.5×10^{-6}	9.4×10^{-1}	9.2×10^{-9}	9.1×10^{-11}	2.00
1,000	0.996	0.003	1.9×10^{-4}	1.1×10^{-5}	6.3×10^{-7}	3.6×10^{-8}	1.2×10^{-10}	4.0×10^{-33}	2.48
1.000	0.999	0.001	$3.2 imes10^{-5}$	10 - 6	3.1×10^{-8}	1() = 9	1()-)2	10-15	3.00
1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	[m]
	$\begin{array}{c} \hline c = 0 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \\ 1,000 \end{array}$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\kappa = 0$ $u = 1^b$ $\kappa = 2$ 1.000 0.000 0.000 1.000 0.001 0.001 1.000 0.001 0.001 1.000 0.003 0.003 1.000 0.010 0.010 1.000 0.029 0.028 1.000 0.291 0.083 1.000 0.250 0.250 1.000 0.750 0.188 1.000 0.909 0.083 1.000 0.909 0.031 1.000 0.996 0.031 1.000 0.996 0.003 1.000 0.999 0.001	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

Except for the case where n = 1, all probabilities were calculated from eq 15 (see text). ^b Calculated from eq 9 (see text).

at least one lipid phase, then increasing difficulty is met as (i) the partition coefficient becomes very large or very small, and (ii) the number of lipid barriers increases. Where n > 1, a maximum probability is found when k/l is equal to unity (1.00). Where n = 2 or more, the probability of a drug reaching the receptor region drops to nil when the partition coefficient is infinite or is zero. This mathematical conclusion is intuitively satisfying, since if a drug has an infinite partition coefficient, we can see how the drug would be readily taken up by lip₁, but it would never be able to leave that phase; on the other hand, a drug which as a parition coefficient of zero can never get into lip₁ in the first place, and hence is blocked from reaching aq₂.

۱

When the probabilities from Table I are plotted against log k/l the family of curves in Figure 2 is ob-



Figure 2.—The probability of a drug reaching a receptor region (aq_n) as a function of the log partition coefficient $(\log k/l)$ and of n, the number of interfaces separating the site of drug application from the receptor region,

tained. Each member of this family has its maximum value at $\log k/l = 0$, and is reminiscent of a normal distribution curve.

It is of interest to see if the probability argument of the present work gives conclusions similar to those of the Hansch school of thought. As a first step one can convert the probabilities of Table I into their logarthmic form as has been done in Table II. A plot of these numbers against the corresponding values for log k/lresults in the family of curves shown in Figure 3. On



Figure 3.—Same as Figure 2 only the scales of the abscissa and ordinate are both in log units.

inspection one is tempted to believe that the curves are parabolas. But closer analysis reveals that this is not the case. For any particular set of data in Table II where n > 1, a parabola can be fitted to the set of points $\log P_{0,n}, \log k/l$. Regardless of the value of *n* chosen, the correlation coefficient of the regression curve is always r = 0.985, *i.e.*, 97.0% of the variance is accounted for by regression. The "unexplained" 3.0% of the variance is too large to be result of round-off error. However, when one examines the differences between the input values of log $P_{0,n}$ and the values calculated from the regression equation it becomes obvious that the differences themselves are a function of $\log k/l$ and n (see Figure 4). Because these differences are not random it must be concluded that the

TABLE 11									
Logarithm of Probability of Drug Crossing n Interfaces:	Log $P_{0,r}$								

					$-\log P_{0,n}^{a}$					
k/l	n = 0	n = 1	n = 2	n = 3	n = 4	n = 5	n = 6	n = 8	n = 10	$\log k/l$
0.000	0.000	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf
0.001	0.000	-3.00	-3.00	-4.50	-6.00	-7.50	-9.00	-12.00	-15.00	-3.00
0.003	0.000	-2.52	-2.52	-3.79	-5.05	-6.31	-7. 57	-10.10	-12.63	-2.52
0.01	0.000	-2.00	-2.00	-3.01	-4.00	-5.02	-6.03	-8.03	-10.04	-2.00
0.03	0.000	-1.54	-1.55	-2.32	-3.10	-3.87	-4.65	-6.19	-7.74	-1.52
0.1	0.000	-1.04	-1.08	-1.62	-2.17	-2.70	-3.25	-4.34	-5.41	-1.00
0.3	0.000	-0.64	-0.75	-1.13	-1.51	-1.89	-2.22	-3.00	-3.75	-0.52
1	0.000	-0.30	-0.60	-0.90	-1.20	-1.51	-1.81	-2.41	-3.01	0.00
3	0.000	-0.13	-0.73	-1.09	-1.45	-1.82	-2.10	-2.91	-3.63	0.48
10	0.000	-0.04	-1.08	-1.62	-2.17	-2.70	-3.25	-4.34	-5.41	1.00
30	0.000	-0.014	-1.51	-2.26	-3.00	-3.76	-4.52	-6.02	-7.53	1.48
100	0.000	-0.0044	-2.01	-3.01	-4.00	-5.02	-6.03	-8.03	-10.04	2.00
300	0.000	-0.0017	-2.48	-3.72	-4.95	-6.20	-7.46	-9.92	-12.40	2.48
1000	0.000	-0.0004	-3.00	-4.50	-6.00	-7.50	-9.00	-12.00	-15.00	3.00
Inf	0.000	0.000	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	Inf

^a Values of Table I converted into logarithmic form.





Figure 4.—The difference between values of log P in Table II and those calculated from a regression parabola used to correlate the data.

relationship between log $P_{0,n}$ and log k/l is not, strictly speaking, parabolic, however, for all practical situations it may as well be.

At this point the question of branched and alternate pathways can be taken up again. When more than one route of travel is open to a drug molecule then the probability of a drug molecule reaching its target will be the sum of the probabilities of the various routes available (see eq 16). In eq 16 the left hand superscript of each

$$P_{0,n} = ({}^{1}P_{0,n} + {}^{2}P_{0,n} + {}^{3}P_{0,n} \dots {}^{j}P_{0,n})$$
(16)

probability term within the parenthesis designates one identifiable route. However, each ${}^{j}P_{0,n}$ term in eq 16 is a function of log k/l as illustrated by the family of curves in Figure 2. Each member of this family is

similar to a normal distribution curve with its maximum at log k/l = 0.00. The sum of such curves is still another curve of the same form,⁴ and hence the presence of branched and alternate routes to the receptor site does not alter the basic relationship of $P_{0,n}$ to log k/l. The normal distribution curve therefore approximates this relationship for the probability considering all routes as well as an individual route.

In real situations there may be several independent routes involving the same number of interfaces, n. These routes of equal probability can be considered degenerate in analogy to quantum mechanics' degenerate states, *i.e.*, states of equal probability. Thus, for molecules traveling in a set of degenerate routes, the probability of reaching a receptor is independent of the path taken. More important, however, routes involving the fewest interfaces are the routes of highest probability. Hence, these minimum pathways will bear the greatest traffic as far as drug molecules reaching the receptors are concerned. Obviously, the minimum pathway could be comprised of several degenerate routes, and this set of routes would be the primary access a drug molecule would have to the receptors. Only a negligible number of molecules will reach the receptors by routes having

(4) It would be difficult to prove this assertion mathematically, but it can be readily demonstrated to be true for the case at hand. A typical regression equation is that for n = 2 (see eq i).

$$\log P_{0,2} = -(0.261 \pm 0.013) (\log k/l)^2 -$$

n =

$$(0.002 \pm 0.0221) (\log k/l) - 0.804$$
 (i)
13 $r^2 = 0.970$ $s = 0.149$ $F_{2,10} = 191.8$ $P < 0.0005$

Let us now consider the probabilities for three routes consisting of different values of n, say n = 2, n = 4, and n = 6. By summing these probabilities for different values of log k/l, and taking the logarithms thereof, we can eventually arrive at the regression equation eq ii.

$$\log (P_{0,2} + P_{0,4} + P_{0,8}) = -(0.271 \pm 0.016) (\log k/l)^2 - (0.002 \pm 0.026) (\log k/l) - 0.739 \quad (ii)$$

$$n = 13$$
 $r^2 = 0.960$ $s = 0.177$ $F_{2,10} = 146.8$ $P < 0.0005$

Thus the data leading to eq ii fit a parabola only slightly less better than the data leading to eq i. In fact the difference is probably due to the greater round-off error inherent in eq ii. Most interesting are the differences butween the "observed" and calculated values given by both equations. The differences associated with eq i show the same regular fourth-order dependence on log k/l as those associated with eq i. These differences are observed for all values of n (see Figure 4). The addition of higher valued terms in n (e.g., $P_{0,8}$) to eq ii will have very little effect on the outcome since such terms make a negligible contribution to the overall probability relative to the lower ordered terms.

six or more interfaces above those of the minimum pathway.

Until now the time element has not been seriously considered in these deliberations. Since the main interest in work of this kind lies in the relative potencies of drugs as measured in a standard biological test, it is fair to limit our discussion to only those tests in which the observation period is the same in each instance. Thus, in the measurement of potency, time is considered to be held constant. In looking at the time variable one quickly realizes that the observation period is normally set by previous experience in dealing with drugs of the type being studied. So to begin with it is known that the test system allows sufficient time for at least some drug molecules to reach the receptors. Now if there are no membranes in the biological system, and a drug is administered as a concentrated dose in a small region at one end of the system, the time it takes for one of these added molecules to reach the other end is defusion controlled, and is a function of the molecules' average velocity. At a constant temperature, molecules of the same mass but of different chemical structures will have the same average velocities (there may be exceptions of course): but even if one molecule has twice the mass of another it will take it only 1.414 times as long to travel the same distance. Hence, other things being equal, the lighter molecules will have a slightly higher probability of reaching a particular distant point than the heavier ones in a given period.

However, these differences are negligible compared to those introduced when membranes and partition coefficients are considered, as is well known. The introduction of membranes to the system does not really alter a drug's rate of travel, what it does is to effectively increase the distance it must cover. It may be repelled at an interface; in order to penetrate it, a molecule must wander about randomly until, by chance, it strikes the interface again. The molecule may have to repeat this process many times. If it is not very successful, it will never reach the receptor to help effect the biological response; after all, there is a time limit.

At this point one may appreciate the above argument on equivalent most-probable routes. Those molecules which travel by the routes of greatest probability will be overwhelmingly represented at the receptor. Almost all others are wasted. What is important in determining the final outcome is the particular pathway the drug molecule travels, and the probability that a particular molecule penetrates an interface upon collision.

Suppose two molecules with different partition coefficients happen to follow exactly the same path, and are repelled at various interfaces exactly the same number of times: each molecule arrives at the receptor region at approximately the same time (we make a small correction for the different masses). An interesting event. but surely the molecule whose partition coefficient is least favorable has overcome greater odds to achieve its status. In a particular test system practically all molecules which reach the receptor travel by equivalent most-probable routes and arrive at approximately the same time. The problem therefore reduces to determining the probability that particular molecules will follow the most favorable routes, and this of course is done by the arguments leading to eq.15. Higher doses of drugs with unfavorable partition coefficients must be administered in order to offer more chances for enough molecules to get through to the receptor in time to effect a response. The foregoing argument may not apply to all possible types of drugs, but it is probably valid for many common cases.

Acknowledgments.—Thanks are due to Dr. L. H. Conover of these laboratories for his encouragement in preparing this work for publication.

Mass Spectral Analysis in the Identification of Human Metabolites of Warfarin¹

WILLIAM F. TRAGER,* RICHARD J. LEWIS, AND WILLIAM A. GARLAND

Department of Pharmacentical Chemistry, University of California San Francisco, San Francisco, California – 94122, and Research Service, Veterans Administration Hospital, San Francisco, California – 94121

Received April 13, 1970

The mass spectra of warfarin [3-(α -acetonylbenzyl)-4-hydroxyconmarin], a denterium labeled analog, 5 hydroxylated analogs, and a reduced side chain analog are discussed and mechanistic rationalizations are given for the major fragmentation processes observed. These data together with nv, tlc, and synthetic data are used to identify and establish 6- and 7-hydroxywarfarin and the two diastereoisomers of 3-[α -(2-hydroxypropyl)benzyl]-4-hydroxyconmarin as metabolites of warfarin in normal man.

Warfarin [3-(α -acetonylbenzyl)-4-hydroxycoumarin, 1a] is an oral anticoagulant commonly employed in this country. Among the problems encountered in its clinical use is the occasionally inordinate difficulty in the maintenance of a stable degree of anticoagulation; as a consequence hemorrhage can result from the same drug dose that previously produced acceptable hypoprothrombinemia. In addition the magnitude of the anticoagulant effect appears to be extremely sensitive to the influence of other drugs. These effects can usually be correlated with changes in the plasma clearance rate of the coumarin anticoagulants. For example² several

^{*} To whom correspondence should be addressed.

⁽¹⁾ This investigation was supported in part by The University of California Academic Senate Grant 10, San Francisco Division, and in part under a Grant-in-Aid of the American Heart Association, supported by the Alameda, San Francisco, San Mateo, and Santa Clara Heart Associations. The authors are grateful for the technical assistance of Mr. Ken Chan and to Endo and Abbott Labo.atories for supplying warfarin.

^{(2) (}a) H. M. Solomon and J. J. Schrogie, J. Phirmacol. Exp. Ther., 154, 660 (1967); (b) L. K. Garretton, J. M. Perel, and P. G. Dayton, J. Agree. Med. Astr., 207, 2053 (1):69).