Biochemical Pharmacology, Vol. 19, pp. 2193–2220. Pergamon Press. 1970. Printed in Great Britain

FIBRINOLYTIC CONGENERS OF BENZOIC AND SALICYLIC ACID

A MATHEMATICAL ANALYSIS OF CORRELATION BETWEEN STRUCTURE AND ACTIVITY*

CORWIN HANSCH and K. N. VON KAULLA

Department of Chemistry, Pomona College, Claremont, Calif. 91711, and
Department of Medicine, University of Colorado School of Medicine,
Denver, Colo. 80220, U.S.A.

(*Received* 26 *September* 1969; *accepted* 23 *December* 1969)

Abstract—Safe, economical, preventive and therapeutic thrombolytic drugs suitable for the general medical practice do not exist. Discovery of numerous organic anions inducing fibrinolytic activity *in vitro* in humans indicated the potential of a new approach for treatment of thromboembolic disease. Previous attempts to improve the fibrinolytic activity of the organic anions were, although partially successful, rather empirical. In the present study, multiple regression analysis of the correlation between structural features and fibrinolytic activity of these compounds was carried out for establishing guidelines for a more systematic developmental work. Such guidelines resulted and, in the case of salicylates and benzoates, are determined by the relative lipophilic character.

AN INEXPENSIVE oral fibrinolytic (thrombolytic) drug for both therapeutic and preventive use in general medicine does not exist. Only a synthetic agent could possibly be endowed with all properties required for such a drug. It has been known for 80 years that treatment of human or animal serum with chloroform can induce proteolytic activity.[1] These observations have not evoked any systematic investigations into synthetic production of fibrinolysis induction. In the course of our endeavors in this field, it became apparent that there were various synthetic organic anions with large lipophilic moieties, such as derivatives of benzoic, salicylic, γ-resorcylic, anthranilic, and thiophene-3-carboxylic acid, which induce marked fibrinolytic activity in human or animal plasma.[2]

There appears to be a clear relation between certain structural features of the synthetic fibrinolysis inducers and the concentrations required to induce fibrinolytic activity. In this report we discuss the results of multiple regression analysis of the physicochemical characteristics of fibrinolytic, salicylic, and benzoic acids. In our study we have, as in the past,[3] factored substituent effects into three categories: electronic, hydrophobic and steric.

The correlations we have found also reflect the accuracy and sensitivity of the testing system.

* Supported by Grants CA 11110 and HE 9985 from the National Institutes of Health, United States Public Health Service, and a grant from the American Heart Association.

METHOD

The fibrinolytic activity of the compounds was assessed with the "hanging clot" method.[4] In this test, a cylinder-shaped clot from recalcified human citrated plasma is suspended in a buffered (pH 7·4) solution of the compound to be tested. A number of runs were carried out with the "plasma clot" method[2] in which the compounds are dissolved in human citrated plasma which is then clotted and incubated at 37°. Compound concentrations interfering with clot formation cannot be tested with the plasma clot. This test more closely approximates conditions in the living organism. The lowest molarity inducing complete clot dissolution in both methods after 24 hr of incubation at 37° has been used in our calculations.

The log P values marked with a double asterisk (**) in Tables 1 and 3 are experimentally-determined values. The others were calculated from either benzoic acid (1·81) or salicylic acid (2·26) by taking advantage[5, 6] of the additive constitutive character of log P. In the calculation of salicylic acids, π values for substituents from the phenol system were used. The value of 0·50 was used for each CH_3 or CH_2 unit. For each double bond, 0·3 unit was subtracted and for branching, as in isopropyl, 0·2 unit was subtracted. For example, 3-(4-isopropylbenzyl) salicylic acid was calculated as follows: $\log P_{\text{salicylic acid}} + \log P_{\text{benzene}} + \pi_{CH_2} + \pi_{\text{isopropyl}} = 2{\cdot}26 + 2{\cdot}13 + 0{\cdot}50 + 1{\cdot}30 = 6{\cdot}19$. In a number of instances it was not possible to make corrections for groups ortho to each other. Such examples may be in error by 0·1 to 0·2 log unit, which is not serious for our purposes.

RESULTS AND DISCUSSION

From the data in Table 1 on derivatives of salicylic acid, equations 1 and 2 have been derived via the method of least squares. In these two equations, n is the number

	n	r	s	
$\log 1/C = 0{\cdot}507\ (\pm 0{\cdot}06) \log P - 0{\cdot}084\ (\pm 0{\cdot}25)$	49	0·929	0·203	(1)
$\log 1/C = -0{\cdot}058\ (\log P)^2 + 0{\cdot}979 \log P - 0{\cdot}208\ (\pm 0{\cdot}19)\sigma + 0{\cdot}220\ (\pm 0{\cdot}14)E_S^0 - 1{\cdot}212\ (\pm 0{\cdot}86)$	49	0·946	0·184	(2)
		$\log P_0 = 8{\cdot}4\ (6{\cdot}3 - 41{\cdot}0)$		

of data points employed in the regression, r is the correlation coefficient, and s is the standard deviation from regression. The figures in parentheses are the 95 per cent confidence intervals. The simple linear relation of equation 1 gives a good correlation for a very diverse set of derivatives of salicylic acid. The more complex equation 2 makes a small but significant reduction in the variance. From equation 2, log P_0 can be calculated by taking the partial derivative with respect to log P. The value of 8·42 for this parameter is high. The confidence intervals[9] on this value are large, so that it is not a very useful limit. Molecules of greater lipophilicity would have to be tested to obtain a better definition of this parameter. It was expected that the electronic or steric effects (or both) of substituents might play a role in the structure–activity relationship. Two different roles were considered for σ. Correlations with the effect of σ on the carboxyl group and on the OH group of salicylic acid were made. In equation 2, the σ value of the substituent in relation to the carboxyl was used, since this gave a better correlation than the value pertaining to the OH group. The

TABLE 1. FIBRINOLYTIC ACTIVITY OF SALICYLIC ACIDS IN THE HANGING CLOT TEST

Derivative	E_s-6*	log P†	$\Sigma\sigma$‡	Obsd log 1/C	Calcd§ log 1/C	Δ log 1/C
Unsubstituted	1·24	2·26**	0·00	0·82	0·98	0·16
5-Nitro	1·24	2·76	0·78	1·16	1·16	0·00
5-Methyl	1·24	2·76	− 0·07	1·16	1·33	0·17
3,6-Dimethyl	0·00	3·26	− 0·24	1·22	1·41	0·19
3-Methyl	1·24	2·76	− 0·07	1·22	1·33	0·11
4-Methyl	1·24	2·76	− 0·17	1·22	1·35	0·13
6-Hydroxy	0·69	2·20**	− 0·36	1·30	0·89	0·42
3-Chloro	1·24	2·95	0·37	1·30	1·36	0·06
5-Ethyl	1·24	3·26	− 0·07	1·30	1·65	0·35
4-Dimethylamino	1·24	2·36	− 0·66	1·40	1·18	0·22
4-Methyl-6-hydroxy	0·69	2·70	− 0·53	1·40	1·27	0·13
5-Bromo	1·24	3·39	0·39	1·40	1·63	0·23
4-Trifluoromethyl	1·24	3·75	0·55	1·52	1·80	0·28
5-Chloro	1·24	3·19	0·37	1·52	1·51	0·00
4-Chloro	1·24	3·30	0·23	1·70	1·61	0·09
5-Iodo	1·24	3·71	0·35	1·70	1·82	0·12
3-Isopropyl-6-methyl	0·00	4·06	− 0·24	1·70	1·85	0·15
3-Methyl-6-isopropyl	− 0·47	4·06	− 0·22	1·70	1·74	0·04
5-*t*-Butyl	1·24	4·06	− 0·12	2·00	2·10	0·10
3,4-Dichloro	1·24	3·94	0·60	2·05	1·89	0·16
5-(2-Cyclopentenyl)	1·24	4·10	− 0·05	2·10	2·10	0·00
3-*sec*-Butyl	1·24	4·06	− 0·07	2·10	2·09	0·01
3-Isopropyl	1·24	3·56	− 0·07	2·10	1·82	0·28
3-Phenyl	1·24	4·39	0·06	2·10	2·22	0·12
3-*t*-Butyl-5-methyl	1·24	4·44	− 0·19	2·16	2·30	0·14
5-*sec*-Butyl	1·24	4·06	− 0·07	2·16	2·09	0·07
3-*n*-Butyl	1·24	4·26	− 0·07	2·16	2·19	0·03
3-*t*-Butyl	1·24	4·06	− 0·12	2·22	2·10	0·12
3-Isopropyl-5-allyl-6-methyl	0·00	5·26	− 0·24	2·22	2·37	0·15
3-*t*-Butyl-6-methyl	0·00	4·44	− 0·29	2·22	2·04	0·18
4-Phenyl	1·24	4·39	− 0·01	2·22	2·24	0·02
5-Phenyl	1·24	4·39	0·06	2·22	2·22	0·00
3-Cyclohexyl	1·24	4·77	− 0·07	2·30	2·42	0·12
3-Benzyl	1·24	4·89	− 0·08	2·30	2·47	0·17
3-Neopentyl	1·24	4·44	− 0·12	2·40	2·28	0·12
3,5-Dibromo	1·24	4·28	0·78	2·40	2·02	0·38
5-Benzyl	1·24	4·89	− 0·08	2·40	2·47	0·07
3-Benzyl-6-hydroxy	0·69	4·83	− 0·44	2·52	2·40	0·12
5-Cyclohexyl-6-hydroxy	0·69	4·71	− 0·42	2·52	2·34	0·18
3-Cinnamyl	1·24	5·09	0·02	2·52	2·53	0·01
3-(3-Chlorobenzyl)	1·24	5·60	0·29	2·52	2·65	0·13
5-Cyclohexyl	1·24	4·77	− 0·07	2·52	2·42	0·10
3-(2-Phenylethyl)	1·24	5·39	− 0·05	2·52	2·65	0·13
3-(4-Chlorobenzyl)	1·24	5·60	0·15	2·70	2·68	0·02
3-(2-Chlorobenzyl)	1·24	5·60	0·15	2·70	2·68	0·02
3-(4-Isopropylbenzyl)	1·24	6·19	− 0·15	2·70	2·92	0·22
3,5-Diiodo	1·24	4·90	0·70	2·70	2·31	0·39
5-(1,1,3,3-Tetramethylbutyl)	1·24	5·46	− 0·12	2·70	2·69	0·01
3-(1,1,3,3-Tetramethylbutyl)	1·24	5·46	− 0·12	3·05	2·69	0·36

* From reference 7.

† See reference 8.

‡ σ was chosen with respect to the COOH group.

§ Calculated by using equation 2.

** Experimental values.

steric effects of substituents in both the 3- and 6-positions were investigated using Taft's E_s parameter. Values for halogen and hydroxyl groups were from the recent modification in the definition of aromatic E_s values[7]. Only E_s-6 resulted in an improved correlation. In using E_s-3, not all of the derivatives could be included in the study, since suitable E_s values were lacking. However, as near as we can ascertain with the data at hand, steric effects of substituents in the 3-position are negligible. The negative sign with the σ term in equation 2 would indicate that electron-releasing substituents promote activity. However, the confidence intervals on this term are wide and the effect is, at best, small. The positive sign with the E_s term in equation 2 indicates that large groups decrease fibrinolytic activity; the exact meaning of this is not clear. The effect of a large group in the ortho position might be associated with its effect on log P, the ionization of the carboxyl group, or the interaction of the carboxyl group with a protein moiety. Since the roles for E_s and σ are so small in equation 2, and since log P_0 is so high, equation 2 offers little over equation 1 to help us in the design of better fibrinolytics. The intercept in equations such as 1 is a most important parameter for the drug designer. This is the value of log $1/C$ when log $P = 0$, and provides a useful way of comparing sets of congeners. When such sets are measured by the same test, the intercept is an expression of the intrinsic activity of isolipophilic compounds.

While the hanging clot test is an excellent test with which to prospect for a set of congeners having high intrinsic activity (large positive intercepts in equations such as 1), it is not a good test for designing the best practical drug. Such a drug will have to enter the plasma system of a person, where its reaction with plasma proteins will compete with that producing the fibrinolytic action. The data in Table 2 are from the

TABLE 2. FIBRINOLYTIC ACTIVITY OF SALICYLIC ACIDS IN PLASMA CLOT TEST

Derivative*	log $1/C$†	Derivative*	log $1/C$†
Unsubstituted	0·70	3-(2-Chlorobenzyl)	2·00
6-Hydroxy	1·22	3-(3-Chlorobenzyl)	1·96
6-Hydroxy-3-methyl	1·00	3-(4-Chlorobenzyl)	2·00
3-Isopropyl-6-methyl	1·60	3-(4-Isopropylbenzyl)	1·92
3-*t*-Butyl	1·82	3-Benzyl-6-hydroxy	2·00
5-Cyclohexyl	1·89	3-Cinnamyl	1·96
3-Benzyl	1·75		

* See Table 1 for substituent constants.
† C is the lowest active molar concentration.

plasma test, a test in which the fibrinolytic process occurs in undiluted serum (containing essentially all plasma proteins minus fibrinogen). From these data, equations 3 and 4 have been formulated.

	n	r	s	
$\log 1/C = 0{\cdot}293\ (\pm\ 0{\cdot}09) \log P + 0{\cdot}377$	13	0·903	0·194	(3)
$\log 1/C = -\ 0{\cdot}079\ (\log P)^2 + 0{\cdot}931 \log P$	13	0·940	0·161	(4)
$-\ 0{\cdot}773\ (\pm\ 1{\cdot}12)$		$\log P_0 = 5{\cdot}9\ (4{\cdot}9 - 30)$		

Two facts stand out from a comparison of equations 3 and 4 with equations 1 and 2: in the first place, the slope and intercept of equation 3 are less favorable. With plasma present, only $\frac{3}{5}$ the increase in activity is obtained for each additional increment in log P. The log P_0 value of equation 4 sets a lower limit on activity that can be obtained by increasing log P. The confidence interval on the intercept of equation 4 is quite broad. However, the value of the intercept is most probably somewhat lower than that of equation 1, indicating lower intrinsic activity for the salicylates in the presence of serum proteins.

While the constants in equations 3 and 4 are not as well established as those of equations 1 and 2, because of the much smaller group of compounds so far tested, the results do rather clearly indicate the lower ceiling on activity one can obtain by increasing lipophilic character.

Table 3 summarizes data for a set of congeneric benzoic acids correlated by equations 5 and 6. For the set of benzoic acids in Table 3, the addition of a term in σ to either

TABLE 3. FIBRINOLYTIC ACTIVITY OF BENZOIC ACIDS IN THE HANGING CLOT TEST

Derivative	log P	$\Sigma\sigma$	Obsd log $1/C$	Calcd* log $1/C$	Δ log $1/C$
4-Nitro	1·83‡	0·78	0·70	0·62	0·08
3-Methyl-4-nitro	2·33	0·71	1·00	0·94	0·06
4-Ethyl	2·81	− 0·15	1·00	1·18	0·18
Pentafluoro	2·64	0·86	1·00	1·10	0·10
4-Chloro	2·68‡	0·23	1·05	1·12	0·07
2-Isopropyl-4-propoxy-6-methyl	4·69	− 0·59	1·16	1·41	0·25
4-Isopropyl	3·11	− 0·15	1·22	1·29	0·07
4-Iodo	2·95‡	0·28	1·30	1·24	0·06
4-Allyl-3,5-diiodo	5·19	− 0·41	1·30	1·29	0·01
4-Propyl	3·31	− 0·15	1·40	1·35	0·05
4-*t*-Butyl	3·49	− 0·20	1·52	1·39	0·13
4-Alloxy-3-propyl	4·09	− 0·34	1·52	1·46	0·06
4-Alloxy-3,5-dichloro	4·25	0·47	1·52	1·46	0·06
4-Alloxy-3,5-diethyl	4·59	− 0·41	1·52	1·43	0·09
4-*t*-Butyl-2-bromo	4·24	0·03	1·52	1·46	0·06
2,4-Dichloro†	3·27	0·46	0·82	1·33	0·51

* Calculated by using equation 6.
† See Table 1.
‡ This point was not included in the derivation of constants.

	n	r	s	
$\log 1/C = 0{\cdot}190\ (\pm\ 0{\cdot}11) \log P + 0{\cdot}589\ (\pm\ 0{\cdot}39)$	15	0·720	0·186	(5)
$\log 1/C = -\ 0{\cdot}158\ (\log P)^2 + 1{\cdot}307 \log P$	15	0·904	0·119	(6)
$-\ 1{\cdot}248\ (\pm\ 0{\cdot}94)$		$\log P_0 = 4{\cdot}14\ (3{\cdot}9 - 4{\cdot}8)$		

equation 5 or 6 did not significantly improve the correlation. With the benzoic acids, the importance of the $(\log P)^2$ is much greater than for the salicylic acids. Log P_0 is much better defined at a much lower level. For the benzoic acids of low log P, we find activity quite comparable to that of salicylic acids of the same log P. This is illustrated in Table 4. However, the parallel activity between the two series stops when

TABLE 4. COMPARISON OF ACTIVITY OF SALICYLIC AND BENZOIC ACIDS HAVING COMPARABLE LOG P VALUES

Compound	log P	log $1/C$
3-Methyl-4-nitrobenzoic acid	2·33	1·00
Salicylic acid	2·26	0·82
4-Isopropylbenzoic acid	3·11	1·22
5-Ethylsalicylic acid	3·26	1·30
4-*t*-Butylbenzoic acid	3·49	1·52
5-Bromosalicylic acid	3·39	1·40

compounds with log P values of about 4 are reached. Increasing the lipophilic character of the benzoic acids beyond 4 actually results in loss of activity. Our interpretation of this is that the benzoic acids are not as free to diffuse randomly into the plasma clot and in this way find the critical sites of action.[10] Making the benzoic acids more hydrophobic slows this process until at log P of about 4 this becomes the rate-determining factor. This parabolic relationship between activity and log P is quite well known for many other systems.

It has been pointed out[11] that fibrinolytic activity may be enhanced or inhibited by the same compound, depending on the concentration employed. It is also known[11] that fibrinolytic activity is quite dependent on the fibrin content of the clot. With these facts in mind, it is of interest to consider the studies of Kochmann[8] on the inhibition and promotion of the swelling of purified horse fibrin by various organic compounds. Equations 7 and 8 are derived from the data in Table 5. The swelling or inhibition of

TABLE 5. ACTION OF ORGANIC COMPOUNDS ON FIBRIN

Compound	log P	Inhibition of swelling		Promotion of swelling	
		Obsd log $1/C$	Calcd log $1/C$	Obsd log $1/C$	Calcd log $1/C$
Chloroform	1·97**	2·18	1·90	2·80	2·56
Pentanol	1·34	1·41	1·36	2·02	2·06
Ether	0·77**	0·72	0·87	1·92	1·60
Butanol	0·84	0·60	0·93	1·51	1·66
Propanol	0·34**	0·48	0·51	0·60	1·25
Ethanol	− 0·16	0·30	0·08	0·90	0·84
Methanol	− 0·66	0·00	− 0·35	0·90	0·43
Ethyl carbamate	− 0·15**	− 0·30	0·09	0·60	0·84

** See Table 1.

Inhibition of fibrin swelling

$$\log 1/C = 0{\cdot}856\ (\pm\ 0{\cdot}32)\ \log P + 0{\cdot}215\ (\pm\ 0{\cdot}31) \qquad n = 8,\ r = 0{\cdot}937,\ s = 0{\cdot}299 \qquad (7)$$

Promotion of fibrin swelling

$$\log 1/C = 0{\cdot}818\ (\pm\ 0{\cdot}41)\ \log P + 0{\cdot}986\ (\pm\ 0{\cdot}40) \qquad n = 8,\ r = 0{\cdot}895,\ s = 0{\cdot}383 \qquad (8)$$

swelling was measured against dried fibrin which was placed in 0·01 N hydrochloric acid. The organic compounds were tested in various dilutions of 0·01 N hydrochloric acid. Considering the difficulty of quantitatively determining the degree of swelling, the correlations with equations 7 and 8 are quite reasonable. The slopes of equations 7 and 8 are, as near as one can tell, identical. The difference in the two equations is in the intercepts. Here we find that the higher concentrations inhibit swelling (lower intercept) while the lower concentrations promote swelling. In each case activity for compounds of quite different structure follows hydrophobic character as defined by log P. While isolated dried fibrin is rather far removed from its natural condition, these results do suggest that fibrinolytic activity of a given blood sample might be greatly influenced by the amount of lipoid material in the blood. Indeed, it is known[11] that various fats do inhibit fibrinolysis. Further study of this point in the easy-to-handle hanging clot test may be of considerable value in treating clotting disorders; that is, varying amounts of natural lipids could be incorporated into the clot and the effect on fibrinolysis with the synthetic agents noted.

It is of interest to compare the slopes of equations 7 and 8 with that of equation 1. The slopes of the former are somewhat higher, indicating a greater sensitivity of the swelling to the lipophilic character of the drug. Other such comparisons can be made, as in equations 9–12, which correlate the binding of organic compounds to

Type of compound	Macromolecule		n	r	Ref.	
Miscellaneous	Bovine albumin	$\log 1/C = 0{\cdot}75 \log P + 2{\cdot}30$	42	0·960	3	(9)
Barbiturates	Bovine albumin	$\log (B/F) = 0{\cdot}51 \log P - 1{\cdot}22$	17	0·896	3	(10)
Barbiturates	Homogenized brain	$\log (B/F) = 0{\cdot}52 \log P - 1{\cdot}44$	5	0·973	3	(11)
Penicillins	Human serum	$\log (B/F) = 0{\cdot}49\pi - 0{\cdot}63$	79	0·924	12	(12)

macromolecules. Equations 9–12 show that the binding of many drugs to various kinds of protein can be quantitatively correlated with log P or π. The dependence of binding on hydrophobic character seems to be surprisingly uniform. In many examples besides those of equations 9–12, it has been found[13] that slopes are in the narrow range of 0·5 to 0·7, regardless of the kind of macromolecules studied. The nonspecific binding of lipophilic compounds by the many different components of the blood means that a very carefully designed program will be needed to find a clinically useful fibrinolytic agent. The guidelines for the synthesis of fibrinolytic compounds apparent at this point in our study are: (1) A careful testing of the widest possible selection of functional groups such as OH, carbamate, etc. should be undertaken. From known π values,[5, 6] one can select one or two examples of each class with log P of about 2. (2) Equation 4 indicates that the most to be gained from adding lipophilic character will be reached

at log P near 5. For some sets of congeners this may be lower (e.g. the benzoic acids). Several derivatives of each of the most active functions with log P values between 2 and 5 should then be tested in the plasma clot test to ascertain rough log P_0 values for each set. (3) For the best set of congeners (i.e. having highest activity at log P_0), a fairly large group should be tested to establish firmly the structure–activity relationship. (4) Toxicity studies should be run with this set of congeners and a toxicity structure–activity equation established. By using the two equations, drugs with the maximum tolerable toxicity and maximum activity can be designed and tested. (5) The best of these drugs can then be tested in animals for thrombolytic activity.

REFERENCES

1. W. J. DENYS and H. DE MARBAIX, *Cellule* **5**, 197 (1889).
2. K. N. VON KAULLA, in *The Chemical Control of Fibrinolysis* (Ed. J. SHOR), in press. Interscience, New York (1969).
3. C. HANSCH, *Acts chem. Res.* **2**, 232 (1969).
4. K. N. VON KAULLA, *J. med. Chem.* **8**, 164 (1965).
5. C. HANSCH and S. M. ANDERSON, *J. org. Chem.* **32**, 2583 (1967).
6. T. FUJITA, J. IWASA and C. HANSCH *J. Am. chem. Soc.* **86**, 5175 (1964).
7. E. KUTTER and C. HANSCH, *J. med. Chem.* **12**, 647 (1969).
8. M. KOCHMANN, *Biochem. Z.* **136**, 49 (1923).
9. C. HANSCH, A. R. STEWARD, S. M. ANDERSON and D. BENTLEY, *J. med. Chem.* **11**, 1 (1968).
10. J. T. PENNISTON, L. BECKETT, D. BENTLEY and C. HANSCH, *Molec. Pharmac.* **5**, 333 (1969).
11. K. N. VON KAULLA, *Chemistry of Thrombolysis*: *Human Fibrinolytic Enzymes*, Thomas, Springfield, Ill. (1963).
12. A. E. BIRD and A. C MARSHALL, *Biochem. Pharmac.* **16**, 2275 (1967).
13. C. HANSCH, in *Drug Design* (Ed. E. J. ARIËNS), Vol. I, in press. Academic Press, New York.