[CONTRIBUTION NO. 1952 FROM THE GATES AND CRELLIN LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY]

A Simple Model of Molecular Specificity in Enzyme–Substrate Systems. I. Theory and Applications to the System Acetylcholinesterase–Substrate

BY SIDNEY A. BERNHARD[1]

RECEIVED SEPTEMBER 29, 1954

A model of enzyme–substrate specificity which permits approximate calculations of the relative binding energies of substrates as a function of stereochemical and electronic configuration is described. The calculations are made on the basis of the following assumptions. (1) Relative rates of enzyme-catalyzed decompositions are the same as those found in simple homogeneous solution catalysis when the mechanisms of decomposition are similar; the enzyme is thus "non-specific" in the decomposition process. (2) The energy of binding is a function of the degree of complementary structure of the substrate to the structure of the enzyme site. (3) The relative binding energies of substrates are determined experimentally from the parameter K_M (= [E][S]/[ES]), where K_M is an equilibrium constant (*i.e.*, it is not affected by the rate of decomposition of ES). Calculations of substrate specificity involving the interaction with the enzyme acetylcholinesterase are described. The major contributions to the total binding energy are found to arise from London dispersion forces, charge interactions and a weak chemical bond. The first two factors are treated quantitatively.

1. Introduction

A dominant feature of enzyme action is the apparent high degree of molecular specificity between enzyme and substrate required for effective catalytic activity. *A priori*, it might appear that enzyme–substrate systems are ideally suited to a theoretical treatment of the effect of structure on reactivity, particularly for the case of simple structurally well-defined organic substrate. Actually only a few treatments have been attempted, notably those of Pauling and Pressman[2,3a,3b] (for the hapten–antibody system) and Whittaker and Adams (for the "pseudo" and "true" cholinesterase–substrate system).[4a,b,5] Theoretical treatments of specificity in enzyme systems are complicated by the following considerations.

(a) The enzymic process under the simplest conditions can be represented as a two-stage process, *i.e.*

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \qquad (1)$$

$$ES \xrightarrow{k_3} E + \text{products} \qquad (2)$$

The effect of substrate structure must be considered in each stage separately. Stage 1 is a heterogeneous process involving the state of the substrate in solution (S) and in the medium of the enzyme site (ES).

(b) Any theoretical treatment involving the distribution of substrate between solvent and enzyme site (as in stage 1) would, if correct, lead only to relative or absolute values of the equilibrium constant (k_2/k_1). Experimentally, the only measurable parameters are k_3 and K_M, the Michaelis constant,[6] where K_M is equal to $(k_2 + k_3)/k_1$. (In the case of antibody–hapten interaction treated by Pauling and Pressman, $k_3 = 0$ and direct correlation with experiment was possible.)

(c) Theoretical predictions of relative k_3 values involve the difficulties of assessing the effect of structure on reactivity in a highly complex protein medium.

In order to cope with some of these difficulties a special model of a specific system has been defined. It will be shown that actual systems in agreement with this model exist and molecular specificity factors can be calculated.

2. Special Model of an Enzyme–Substrate System

Notation.—Let RX be a series of homologous substrates which undergo enzymatic decomposition to RY. (Molecular specificity will be a function of the structure of R only.) Let RX and R′X represent two such substrates, and subscripts i and j represent two solvent media in which decomposition is possible. Let the decomposition process be represented as

$$C_i + RX_i \xrightarrow{k_i} RY_i + C_i$$

when the medium is homogeneous and

$$RX_i \underset{k_2}{\overset{k_1}{\rightleftharpoons}} RX_j$$

$$E_i + RX_i \xrightarrow{k_3} RY + E_i$$

when the medium is heterogeneous. (C_i and E_j are catalysts effective in the particular solvent denoted by the subscript, and k_i and k_3 are the respective specific rates.)

Let the partition functions Q_R and Q_X denote the unreactive and reactive portions of the substrate, respectively.[7] (Starred superscripts represent activated states.)

$$\frac{k_i}{k_{i'}} = \frac{Q^*_{R_i}Q^*_{X_i}}{Q^*_{R'_i}Q^*_{X'_i}} \frac{Q_{R'_i}Q_{X'_i}}{Q_{R_i}Q_{X_i}} \qquad (2.1)$$

For the over-all heterogeneous process

$$Rx_i + E_i \xrightarrow{k_{i_j}} RY + E_j$$

$$\frac{k_{i_j}}{k'_{i_i}} = \frac{k_3}{k'_3} \frac{Q^*_{R_j}Q^*_{X_i}}{Q^*_{R'_j}Q^*_{X'_j}} \frac{Q^*_{R'_i}Q^*_{X'_i}}{Q^*_{R_i}Q^*_{X_i}} \qquad (2.2)$$

Assumptions: a.—For the substrates under consideration, activation in a homogeneous medium

(1) Naval Medical Research Institute, Bethesda 14, Maryland.

(2) L. Pauling and D. Pressman, THIS JOURNAL, **67**, 1003 (1945).

(3) (a) L. Pauling, D. Pressman and A. L. Grossberg, *ibid.*, **66**, 784 (1944); (b) D. Pressman, A. L. Grossberg, L. H. Pence and L. Pauling, *ibid.*, **68**, 254 (1946).

(4) (a) D. H. Adams, *Biochim. Biophys. Acta*, **3**, 1 (1949); (b) D. H. Adams and V. P. Whittaker, *ibid.*, **3**, 358 (1949).

(5) D. H. Adams and V. P. Whittaker, *ibid.*, **4**, 543 (1950).

(6) L. Michaelis and M. L. Menten, *Biochem. Z.*, **49**, 333 (1913).

(7) The use of partition functions in the following argument is considered desirable in order to illustrate the factorization of the total enzyme–substrate interaction into distinct non-concertive stages, each a function of a particular portion of the substrate molecule.

involves no change in the partition function Q_R

$$Q^*_R/Q_R = 1$$

This assumption is suggested by the observation that in the solvolysis of homologous series of acyl derivatives in homogeneous solution it is often found that

$$k'_i/k_i = k'_j/k_j \qquad (2.3)$$

(even in cases where the activities of the reactants vary widely in both solvents in a manner similar to the activities of the parent hydrocarbons).[8] Appendix 1 lists some examples from the literature.

Therefore, from equation 2.1

$$\frac{k_i}{k'_i} = \frac{Q^*_{X_i} Q_{X'_i}}{Q^*_{X'_i} Q_{X_i}} \qquad (2.4)$$

For cases following assumption a and equation 2.3 the rate of decomposition of any substrate in a particular medium is determined by a constant times the rate in any other medium; *i.e.* $Q_{X_i} = q_i q_X$, where q_i is an effect of solvent only and contains no cross-product terms arising from R contributions, and q_X is a solvent-independent term for a particular RX species (and hence implies that there is no change in activation mechanism with change of medium). This effect is illustrated in Appendix 1.

b.—It is now further assumed that *the distribution of substrate molecules between enzyme site and water solution can be treated as an ordinary distribution between two media, and that the mechanism of the "k_3" step can be related to a particular homogeneous catalyzed mechanism by equation 2.4.*

When considering the site as a medium *in which activation can occur*, all calculations must be made in such a way that X and E (the catalytic group within the enzyme site) are oriented to permit chemical interaction, since the experimentally determined Michaelis constant is a measure of such orientations only. (Enzyme–substrate complexes which do not lead to decomposition do not enter into the Michaelis constant.)

c.—It is assumed that in the specific cases under consideration K_M is equal to k_2/k_1 and thus can be treated as an equilibrium constant. Under these assumptions we have

$$\frac{k_{i_j}}{k'_{i_j}} = \frac{k_i}{k'_i} \times \frac{K_{i_j}}{K'_{i_j}}$$

where K_{i_j} is the distribution coefficient of the substrate between bulk solvent and enzyme site ($K_{i_j} = [S]_E/[S]_W$). This equation is applicable only when the enzyme sites are far from saturation, but this is the situation of greatest interest since the effect of specificity is at a maximum. When the sites are saturated, we have

$$\frac{k_{i_j}}{k'_{i_j}} = \frac{k_3}{k'_3} = \frac{k_i}{k'_i}$$

and according to assumption b no distinction in molecular specificity between the two substrates can be detected. Since $K_M \simeq [E][S]_W/[ES]$, when the sites are far from saturation we have

$$\frac{K_M}{K'_M} = \frac{[S]_W}{[S']_W} \frac{[ES']}{[ES]} = \frac{K'_{i_j}}{K'_{i_j}}$$

and hence

$$\frac{k_i}{k'_i} \times \frac{K'_{i_j}}{K'_{i_j}} = \frac{k_3}{k'_3} \times \frac{K'_M}{K_M} \qquad (2.5)$$

d.—All calculations in the following treatment are based upon the theory of complementary structure proposed by Pauling.[2,3a,b,9]

In order to define the enzyme site stereochemically and electronically, data concerning the binding of competitive inhibitors to the enzyme site can be used to advantage, the relative inhibition constants ($K_I = [E][I]/[EI]$) being taken as a measure of the structure of the site. All binding constants are hereafter referred to as K_S, with no resultant ambiguity when assumption c is applicable. The correlation of K_M with K_I is considered in a following paper.[10]

3. Calculation of Binding Energies and Comparison with Experiment

A treatment of the system acetylcholinesterase–specific substrate is presented below. The choice of system was based primarily on the extensive body of physico-chemical data relating to substrates and inhibitors, and their relatively simple structures. In addition, assumptions b and c appeared to be reasonable for many substrates.

3A. Forces about the Non-reactive Region of the Substrate. The Interaction of Unlike Charges.—A good deal of evidence has been presented in establishing the existence of a negative site on the enzyme complementary to a positive quaternary ammonium ion. K_S values listed in Table I are taken from the work of Bergmann, Nachmansohn, Wilson and co-workers.

TABLE I

Substrate and inhibitor structures	K_S	Reference	K_S/K_S^+
$(CH_3)_2NHCH_2CH_2OH^+$	4.5×10^{-4}	11	
$(CH_3)_2CHCH_2CH_2OH$	1.4×10^{-2}	11	31
Eserine, positively charged acid	1.8×10^{-8}	12, 13	
Eserine, neutral base	4.5×10^{-7}	12, 13	25
$(CH_3)_2NHCH_2CH_2OCOCH_3^+$	1×10^{-3}	11	
$(CH_3)_2CHCH_2CH_2OCOCH_3$	8×10^{-3}	11	8

The esterase (pseudo-cholinesterase) found in blood plasma, which appears to have an active structure similar to acetylcholinesterase but without a negative charge, gives the following K_S values relative to "true" cholinesterase from erythrocytes[5]

Substrate	$K_S$plasma/$K_S$erythrocyte[5]
Choline	29
Acetylcholine	7

(8) (a) "Tables of Chemical Kinetics," Nat. Bureau of Standards, Washington, Circular 510, 1951; (b) C. K. Ingold and W. S. Nathan, *J. Chem. Soc.*, 222 (1936); (c) E. W. Timm and C. N. Hinshelwood, *ibid.*, 862 (1938); (d) H. A. Smith, THIS JOURNAL, **61**, 254, 1176 (1939); (e) H. A. Smith and H. S. Levenson, *ibid.*, **61**, 1172 (1939); (f) H. S. Levenson and H. A. Smith, *ibid.*, **62**, 1556, 2324 (1940); (g) H. A. Smith and J. H. Steele, *ibid.*, **63**, 3466 (1941); (h) H. A. Smith and C. H. Reichardt, *ibid.*, **63**, 605 (1941); (i) H. A. Smith and R. R. Myers, *ibid.*, **64**, 2362 (1942).

(9) L. Pauling, D. H. Campbell and D. Pressman, *Physiol. Revs.*, **23**, 203 (1943).

(10) S. A. Bernhard, THIS JOURNAL, **77**, 1973 (1955).

(11) I. B. Wilson, *J. Biol. Chem.*, **197**, 215 (1952).

(12) D. Nachmansohn and I. B. Wilson, *Adv. in Enzymology*, **12**, 259 (1951).

(13) I. B. Wilson and F. Bergmann, *J. Biol. Chem.*, **185**, 479 (1950).

It appears that the maximum interaction free energy for the pure ion–ion (electrostatic) interaction between enzyme and substrate is −2.1 kcal./mole ($\Delta F = -RT \ln 30$) at an ionic strength of about 0.2 M. The electrostatic energy (corrected for the ion atmosphere assuming the Debye–Huckel theory)[14] is given by

$$\Delta F = \frac{\epsilon^2 N Z_A Z_B}{Dr} \times \frac{e^{-\kappa(a_i - r)}}{1 + \kappa a_i} \qquad (3.1)$$

where r is the distance between the two charges, a is the distance of closest approach of the ion atmosphere to the central ion, and κ is the Debye–Huckel parameter. For the interaction of small ions we have $a_i \sim r$.[15] However, the structure of the enzyme site may be such that the positive ion cannot approach the distance $r \sim a_i$ due to steric factors. Under such conditions the identification of r with the sum of the van der Waals ionic radii of the positive and the negative group on the enzyme is not valid. The model we have assumed in this case is that of a small negative ion fixed to the surface of the enzyme and surrounded by an atmosphere of small ions and charged substrate (the charged group of the substrate having the same van der Waals contact radius as the small ions). When the substrate is small (no steric considerations) equation 3.1 with $a_i - r = 0$ is assumed valid. The problem of the effective dielectric constant (D) has been handled by use of the Schwarzenbach function.[16] Other treatments tend here to give larger values of the effective dielectric constant and these larger values in turn lead to charge separations (r) smaller than the sum of the van der Waals contact radii of the charged group on the substrate and a carboxylate oxygen.

When steric factors arise ($a_i - r \neq 0$), the entire equation (3.1) is to be considered.

Assuming that at the limiting (maximum) interaction energy (−2.1 kcal.), the condition $r \sim a_i$ holds, we may derive the equation

$$-2.1 = \frac{-2.6 \times 10^{-6}}{Dr(1 + \kappa r)}$$

and solving for D and a with the Schwarzenbach function, we obtain

$$r = 5.6 \times 10^{-8} \text{ cm.}, \quad D = 23$$

essentially in agreement with the values of Whittaker and Adams,[5] and Nachmansohn and Wilson[12] determined by means of similar equations.

As was pointed out by the above authors, this value of r is in excellent agreement with the sum of the van der Waals ionic radii of the tetramethylammonium ion and a negative oxygen (on the enzyme), namely, 3.5 + 1.6 = 5.1 Å.

We have here assumed a model for two spherically symmetric charges in a uniform dielectric medium. Such a model cannot be a perfect representation of enzyme and charged substrate. If the isolated negative charge on the enzyme (probably a carboxylate ion) protrudes from the surface, and since the substrates all have nearly spherical charge distributions the model is not as gross an exaggeration of the true situation as it might at first seem. Since K_S for some substrates must be near the equilibrium value (k_2/k_1) and k_3 is very large, equilibration appears to be faster than diffusion through the protein medium would allow. The relative values of r, calculated from the model for different substrates, seem reasonable although the absolute values may show an indeterminate discrepancy.

A good fit at this site exists with choline, dimethylethanolammonium ion and eserine. In all examples listed in Table I the effect of dispersion forces has been eliminated by taking the ratio of $K_{\text{ion-ion}}/K_{\text{ion-neutral molecule}}$ where the neutral molecule can be expected to make the same dispersion energy contribution as the corresponding ion. (See following section.)

When the ratio of $K_{\text{ion-ion}}$ $K_{\text{ion-neutral molecule}}$ is taken for the substrate series A-CH_2-CH_2-O-C-CH_3, where A is either an alkyl substituted ammonium or methyl group, the electrostatic energy drops to about −1.2 kcal./mole, indicating poor fit at the site.

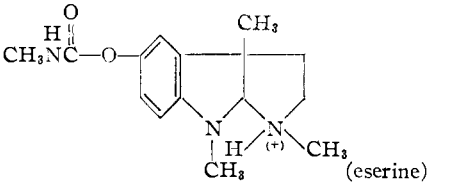
$$\Delta F = -1.2 = \frac{\epsilon^2 N Z_A Z_B}{Dr} \frac{e^{\kappa(a_i - r)}}{(1 + \kappa a_i)} = \frac{-2.6 \times 10^{-6}}{Dr} e^{\kappa(a_i - r)}$$

Solving again with the aid of the Pauling and Pressman approximation of the Schwarzenbach function[2] (in this case, $D = 6r - 12$)

$$r = 6.6 \text{ Å.}$$

The correction for the ion atmosphere should be made on the basis of the closest approach of sodium ions, rather than $(CH_3)_4N^+$, since essentially only the buffer electrolyte accounts for the ionic-strength contribution. However, little error is introduced in the exponential of the above equation and none at all in the denominator by assuming $a(\text{solvated } Na^+) = a((CH_3)_4N^+)$.

By means of atomic models in this Laboratory (similar to Fisher–Hirschfelder models) the eserine molecule was constructed.



(eserine)

The distance between the positive nitrogen and the carbonyl carbon was found to be 1.4 Å. greater than that for the corresponding distance in a fully extended model of acetylcholine. (Binding at the carbonyl carbon will be discussed in a later section.) This indicates that even when completely extended, the acetylcholine molecule is not long enough to be a good complementary model of the enzyme site, since the eserine molecule meets the requirements for a good fit (see Table I). On the basis of electrostatic energy calculations acetylcholine appears to be 1.0 Å. too short. More evidence bearing on this misfit will be presented in the following section. Recently, Friess and McCarville[17] have calculated

(14) See for example, E. S. Amis, "Kinetics of Chemical Change in Solution," The Macmillan Co., New York, N. Y., 1949.

(15) G. Scatchard, *Chem. Revs.*, **10**, 229 (1932).

(16) G. Schwarzenbach, *Z. physik. Chem.*, **A176**, 133 (1936).

(17) S. L. Friess and W. J. McCarville, THIS JOURNAL, **76**, 1363 (1954).

the distance between charge and carbonyl carbon to be 1.1 Å. greater in eserine than in acetylcholine.

Dispersion Forces.—Calculations of the binding energies of substrates to the enzyme resulting from a complementary fit were made on the basis of the treatment of Pauling and Pressman.[2] The final relationship employed in the present paper is

$$W_B = \frac{-4.0 \times 10^5 [R_B]}{(r_B + 2)^6} \text{ (cal. mole}^{-1}\text{)} \qquad (3.3)$$

where $[R_B]$ and r_B are the mole refractions (in cm.^{3}) and van der Waals contact radii (in Å.) of the substituent B, respectively. $[R_B]$ was calculated from the refractive index obtained from the Landolt–Börnstein tables.[18] r_B is taken from Pauling.[19] Dispersion energies in the present paper, unless otherwise specified, signify $\Delta W = W_B - W_H$ where W_H is the calculated dispersion energy due to the interaction of a hydrogen atom with the enzyme. The derivation of equation 3.3 is given in Appendix 2.

The following inferences may be drawn from Table II. Replacement of hydrogen by methyl at the quaternary N atom results uniformly in an increase of binding energy of $-RT \ln 6.7 = -1.2$ kcal., 700 cal. more than would be predicted from the dispersion energy contribution, for both the quaternary ammonium ions (compounds two, three and four, Table II) and the ethanolammonium ions (compounds six, seven and eight). As was pointed out by Wilson,[11] there is no increase in binding energy on substitution of the fourth methyl group (compounds one and five), indicating that the enzyme surface can engulf only three of the tetrahedral nitrogen orbitals.

TABLE II

EFFECT OF DISPERSION FORCES

Structure	$K_S \times 10^3$
$N(CH_3)_4^+$	1.6
$HN(CH_3)_3^+$	1.4
$H_2N(CH_3)_2^+$	11.0
$H_3NCH_3^+$	63
$(CH_3)_3NCH_2CH_2OH^+$	0.45
$(CH_3)_2NHCH_2CH_2OH^+$	0.45
$CH_3NH_2CH_2CH_2OH^+$	6.3
$H_3NCH_2CH_2OH$	21
$N(C_2H_5)_4^+$	0.45
$(CH_3)_2NHC_3H_7\text{-}n^+$	0.86
$(CH_3)_3NC_6H_5^+$	0.072
$n\text{-}C_6H_{13}N(CH_3)_2CH_2CH_2OH^+$	0.41
$(CH_3)_3NCH_2CH_2OCOCH_3^+$	0.45
$(CH_3)_2NHCH_2CH_2OCOCH_3^+$	1.0
$(CH_3)_2CHCH_2CH_2OCOCH_3$	8
$CH_3CH_2OCOCH_3$	500
$CH_2CH_2OCOCH_2Cl$	30

From similar calculations it is found that substitution of larger groups for hydrogen on the ammonium center results in a discrepancy of 700 cal. between equation 3.2 and experiment. If the binding energies of two larger substituents are compared, there is good agreement between equation 3.2 and experiment. This is illustrated in Table IV where an average value of 1.2×10^{-3} M was used for K_S of $(CH_3)_4N^+$ and $(CH_3)_3\overset{+}{N}H$.

Equation 3.3 is insufficient when considering the substitution of N–R for N–H in amines since the breaking of the particularly strong N–H $\cdots$ OH_2 hydrogen bond in the hydrogen compound is not considered. In a previous investigation,[2] the difference in binding between N–H and N–R was found to be 500 to 1000 cal. greater than that calculated from equation 3.3.

From Table III it may be inferred that although the surface of the enzyme can accommodate the groups trimethylammonium, phenyldimethylammonium, *n*-propylmethylammonium, and dimethylethanolammonium, substitution of *n*-hexyl for methyl in choline produces no significant increase in binding. The *n*-hexyl group apparently occupies the unbound corner of the tetrahedron. Similarly, tetraethylammonium does not bind as strongly as would be expected on the basis of a complementary protein structure, indicating an unsymmetrical structure for the enzyme site (at least one axis longer than the others such that the maximum interaction between surface and substrate is realized for both trimethylammonium and phenyldimethylammonium ions).

TABLE III

CALCULATED DISPERSION ENERGY ΔW AND EXPERIMENTAL FREE ENERGY DIFFERENCES $\Delta\Delta F$ DUE TO CHANGE OF NITROGEN SUBSTITUENT

A	B	Cpd. in Table III	ΔW, cal./mole	$\Delta\Delta F$, cal./mole (expl. mean)
Hydrogen	Methyl	1–7	− 500	−1200
Methyl	*n*-Propyl	2, 10	− 900	− 900
Methyl	Ethanolammonium	1–7	− 800	− 700
Methyl	Phenyl	1, 2, 11	−2000	−1900
Methyl	*n*-Hexyl	5, 6, 12	−2200	− 50
Methyl	Ethyl	1, 2, 9	− 450	− 200

Ion-dipole Forces.—In the study of the interaction of this enzyme with low molecular weight substrates, dipolar and ion-dipole forces are of secondary importance. The largest force operative in this particular case would be due to the interaction of the positive charge of the substrate with the enzyme. The maximum interaction energy is given approximately by[2]

$$W = -\frac{3\epsilon^2[\text{Renz.}]}{4\pi N D r^4} \sim 200 \text{ cal.}$$

This small interaction energy would vary little with r.

3B. Forces Near the Reactive Center. The Carbonyl Carbon–Enzyme Interactions.—Wilson and Bergmann[12,20,21] have presented evidence for a strong interaction between a basic group on the enzyme and the electrophilic carbonyl carbon atom. Table IV lists K_S values for some relevant carbonyl compounds.

From the results in Table IV it is possible to calculate the binding of the –CO–CH_3 group by two independent methods.

(a) $\Delta F_{(-COCH_3)} = -RT \ln \left(\frac{K_{C_6H_5N(CH_3)_3^+}}{K \text{ (cpd. 4.6)}} \right) = -2.7$ kcal.

(18) Landolt-Börnstein, "Physikalisch-chemische Tabellen," J. Springer, Berlin, 1923, and 2nd suppl., 1931.

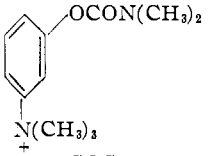
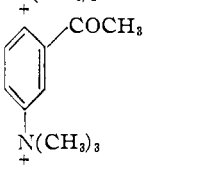
(19) L. Pauling, "Nature of the Chemical Bond," Cornell Univ. Press, Ithaca, N. Y., 1936.

(20) I. B. Wilson and F. Bergmann, *J. Biol. Chem.*, **186**, 693 (1950).

(21) I. B. Wilson, *Biochim. Biophys. Acta*, **9**, 473 (1953).

TABLE IV

THE BINDING OF CARBONYL COMPOUNDS TO ACETYLCHOLINESTERASE

	Compound	K_S	Ref.
4.1	Nicotinamide	4×10^{-3}	20
4.2	Ethyl nicotinate	6×10^{-5}	20
4.3	Nicotinic acid (mol. acid)	6×10^{-5}	20
4.4	β-Acetylpyridine	1.3×10^{-4}	20
4.5		5×10^{-8}	13
4.6		8×10^{-7}	22
4.7	Ethyl acetate	0.5 (± 0.2)	9
4.8	Ethyl chloroacetate	0.03 (± 0.015)	9

No interaction between the quaternary ion and the carbonyl group of compound 4.6 is here assumed.

(b) $\Delta F_{(-COCH_3)} = -RT \ln K$ (cpd. 4.4) $+ \Delta W_{pyridine}$
$= -5.2 \quad + 2.4 = -2.8$ cal.

(ΔW = dispersion energy for the binding of pyridine to a protein surface). The subtraction of the dispersion energy due to the binding of the aromatic nucleus seems reasonable on the basis of the close agreement of the two free energy calculations. A similar calculation for the contribution of $-\overset{O}{\overset{\|}{C}}-OC_2H_5$ in ethyl nicotinate results in $\Delta F_{(-COOC_2H_5)} = -5.9 + 2.4 = -3.5$ kcal.

The maximum value of $\Delta F_{(-COOC_2H_5)}$ in ethyl acetate is -0.4 kcal., assuming no binding at all at the methyl group.

There appears to be no other way of accounting for these very large binding energy differences than by postulating chemical bond formation between the electrophilic carbonyl carbon and an electron donor group on the enzyme. This explanation was proposed originally by Wilson.[10] The binding energies, from the data in Table IV, are in all cases consistent with the electrophilic character of the carbonyl carbon.

The binding of compound 4.5 has been measured as a function of pH and has been found to fall off in acid solution. Wilson and Bergmann have determined the pK_A of the enzyme group (pK_A = 7.2), and have suggested the imidazole nitrogen as a possible donor.[23,23]

The difference in binding energy ($\Delta F = -1.7$ kcal.) between ethyl acetate and ethyl chloroacetate adds further support to the chemical bond formation hypothesis. This difference cannot be accounted for by the sum total of all other pertinent forces which might be operative. Moreover, a hydrogen bond with the enzyme moreover would favor the binding of ethyl acetate.

3C. The Activation Process.—Comparison of specific rates of enzyme-catalyzed hydrolysis of a series of substrates is complicated by the variability of the enzyme preparation. Wilson[11] has studied the relative rates for a series of substrates with the same acetylcholinesterase preparation. These data (V_{max}) together with specific rates of hydroxyl and hydronium ion catalyzed hydrolysis of two of the substrates are listed in Table V. The striking fact illustrated is that the enzyme-catalyzed rates vary no more considerably than the homogeneous rates with change of substrate structure. It would be difficult to ascribe a mechanism to the enzyme catalysis solely on the basis of these data; however, the evidence is suggestive of a hydroxyl mechanism.

TABLE V

RATES OF ESTER HYDROLYSIS AT 25° IN WATER

Substrate	V_{max},[10]	k_{OH}[a]	k_H[a] × 10⁵
Ethyl acetate	12	0.12[26]	4.8[26]
Ethyl chloroacetate	13	...	..
Isoamyl acetate	11	...	..
Dimethylaminoethyl acetate	38	...	..
Acetylcholine	100	1.20[25]	3.0[25]

[a] In the units of l. mole^{-1} sec.$^{-1}$.

The results given in Table V are in good agreement with the assumption (b) that the specificity of the process lies in the binding stage. Little change in specific rate of hydrolysis is to be expected from substitution on the β-alkoxyl carbon atom. For the case of *meta*- and *para*-methyl and nitro-substituted benzyl acetates, no significant variation of rate of acid hydrolysis is observed, whereas the maximum variation in basic hydrolysis (benzylacetate and *m*-nitrobenzyl acetate) is a factor of only three.[8c,24] Substitution at the β-carbon should produce much smaller effects in basic hydrolysis. Variation of rate with variation of charge should be expected.[25,26]

5. Conclusions

The simple model of enzyme specificity outlined in this paper has permitted calculations of relative binding energies in good agreement with experiment in specific instances. Moreover, it presents an approach toward a quantitative physical-chemical description of enzyme specificity.

Qualitatively, the model often distinguishes reasonable from unreasonable mechanisms. Thus the interpretation of the total binding specificity in terms of small physical forces alone would lead to the prediction of a dipole-dipole or ion-dipole energy of greater than 3 kcal./mole due to carbonyl–enzyme interaction in the case of β-acetylpyridine (Table IV), a value 10 times greater than expected. The assumption that dispersion forces alone account for the difference in binding energy between amine and methylamine to the enzyme leads to a dispersion energy greater than the heat of vaporization of methane (2.2 kcal./mole), after the dispersion energy loss due to the loss of eight water molecules has been subtracted.

The essential feature of the treatment is the subdivision of the problem into three distinct steps, the

(22) F. Bergmann and A. Shimoni, *Biochim. Biophys. Acta*, **9**, 473 (1953).

(23) I. B. Wilson and F. Bergmann, *J. Biol. Chem.*, **186**, 683 (1950).

(24) E. Tommila and C. N. Hinshelwood, *J. Chem. Soc.*, 1801 (1938).

(25) J. Buttersworth, D. D. Eley and G. S. Stone, *Biochem. J.*, **53**, 30 (1953).

(26) J. E. Potts and E. S. Amis, THIS JOURNAL, **71**, 2112 (1949).

enzyme-R (physical) and enzyme-X (chemical) binding energies, and the activation process. These steps are here assumed not to be concertive. The first step may be estimated by approximate treatments of physical forces, and the second step may be approached by the method of Hammett (*i.e.*, the empirical evaluation of electronic substituent effects). This approach has not been applied here due to lack of data regarding the pertinent substituent constants (σ). The work of Nath and Rhydon[27] presents a notable example of such an approach to enzymic problems. Recently Friess and McCarville[16] have investigated this electronic effect quantitatively with a series of inhibitors of acetylcholinesterase. Buttersworth, Eley and Stone[24] have reported specific rates of acid- and base-catalyzed hydrolysis of acetylcholine; however, the homogeneous solution kinetic data would have to include a series of substrates before a comparison with the present model could be undertaken.

Acknowledgments.—The author wishes to express his thanks to Professor Linus Pauling for stimulating discussions of molecular specificity and to Dr. I. B. Wilson for discussion of his experimental data and results. The work was done wholly under a fellowship from the American Cancer Society upon recommendation of the Committee on Growth of the National Research Council (U.S.A.).

(27) R. L. Nath and H. N. Rhydon, *Biochem. J.*, **57**, 1 (1954).

Appendix 1.—Solvolysis of Acyl Derivatives

Reaction	Solvent
$RCO_2C_2H_5 + H_3O^+ \xrightarrow{k_1} RCO_2H + C_2H_5OH + H_3O^+$	Acetone (56 wt. %)–H_2O
$RCO_2H + CH_3OH + H^+ \xrightarrow{k_2} RCO_2CH_3 + H_3O^+$	Methanol
$RCO_2H + C_2H_5OH + H^+ \xrightarrow{k_3} RCO_2C_2H_5 + H_3O^+$	Ethanol
$RCONH_2 + H_3O^+ \xrightarrow{k_4} RCO_2H + NH_4^+$	H_2O
$RCO_2C_2H_5 + OH^- \xrightarrow{k_5} RCO_2^- + C_2H_5OH$	85% Ethanol–H_2O
$RCO_2C_2H_5 + OH^- \xrightarrow{k_6} RCO_2^- + C_2H_5OH$	Acetone (56 wt. %)–H_2O
$RCONH_2 + OH^- \xrightarrow{k_7} RCO_2^- + NH_3$	H_2O
$RC_6H_4CO_2CH_3 + H_3O^+ \xrightarrow{k_8} RC_6H_4CO_2H + CH_3OH + H_3O^+$	Ethanol[a]–H_2O
$RC_6H_4CO_2CH_3 + H_3O^+ \xrightarrow{k_9} RC_6H_4CO_2H + CH_3OH + H_3O^+$	Acetone[a]–H_2O
$RC_6H_4CO_2CH_3 + OH^- \xrightarrow{k_{10}} RC_6H_4CO_2^- + CH_3OH$	Acetone (56 wt. %)–H_2O
$RC_6H_4CO_2CH_3 + OH^- \xrightarrow{k_{11}} RC_6H_4CO_2 + CH_3OH$	Acetone[a]–H_2O

[a] 40 ml./100 ml. solution.

All rates were obtained from "Tables of Chemical Kinetics," (Nat. Bureau of Standards Circular 510, Washington, 1951). Original references are given under 8b-i.

Specific rates were rounded to 10% (the estimated comparability of data). All data are at 25° except k_1 (63.2°) and k_8 (80°). Reference specific rates are listed at the top of the columns to illustrate the numerical spread. Specific rates are in l. mole^{-1} sec.$^{-1}$. Propionyl derivatives were taken as reference points (k_0) in the acyl series since acetyl derivatives appear to be exceptional.

RELATIVE RATES OF SOLVOLYSIS IN VARIOUS SOLVENTS

k_0 =	3.74×10^{-5}	5.73×10^{-2}	1.01×10^{-2}	4.3×10^{-5}	3.63×10^{-3}	2.24×10^{-3}	3.8×10^{-5}		2.05×10^{-5}	2.09×10^{-5}	2.87×10^{-3}	5.85×10^{-4}
R	k_1/k_0	k_2/k_0	k_3/k_0	k_4/k_0	k_5/k_0	k_6/k_0	k_7/k_0	$RC_6H_4CO_2CH_3$	k_8/k_0	k_9/k_0	k_{10}/k_0	k_{11}/k_0
CH_3	1.2	1.1	1.5	0.9	1.7	2.0	1.0	R	1.00	1.00	1.00	1.00
C_2H_5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	p-NH_2	...	...	0.028	0.022
n-C_3H_7	0.50	0.50	0.50	0.45	0.45	0.40	0.40	m-CH_3	...	...	.6	.7
i-C_3H_7	.37	.37	.40	.50	.22	.25	.35	p-CH_3	1.1	0.9	.40	.44
n-C_4H_9	.50	.50	.50		.50	.30	...	p-OH	0.50	.50		
i-C_4H_9	.15	.12	.12	..	.12	.10	...	o-NO_2	.070	.075		
t-C_4H_9	.035	.036	.0024	..	.007	.01		m-NO_2	.10	.10		
C_5H_{11}	.50	.50	....	.45	....	....		p-Cl	.9	.9		
$(C_2H_5)_2CH$	.032	.009	....	..	.0040	.0040	...	p-Br	.9	.9		
n-C_6H_{13}	.45	.50										
$(n$-$C_3H_7)_2CH$	.03	.008										
$C_6H_5CH_2$	.40	.45	.60									
$C_6H_{11}CH_2$	.13	.13										

Appendix 2.—Dispersion-Forces—Derivation of Equation 3.3.—The approximate treatment of London results in the following equation correlating the interaction energy (W) with the "instantaneous electric dipoles" of two atoms, groups or molecules

$$W = -3/2 \frac{\alpha_A \alpha_B}{r^6_{AB}} \times \frac{I_A I_B}{I_A + I_B}$$

where I_A and I_B are the average energies between ground and first excited states, α the electronic po-

larizabilities and r_{AB} the distance between the groups.

Pauling and Pressman[2] assume from the similar ionization energies of several molecules

$$I_A = I_B = 14 \text{ e. v.}$$

and substituting $[R] = 4/3\pi N \alpha$, where [R] is the mole refraction

$$W = \frac{-38{,}000}{r^6_{AB}} [R_A][R_B] \text{ (cal. mole}^{-1}) \qquad (1)$$

with r_{AB} in Å.

From data on crystals and gases it is found that the electronic repulsion between the groups results in a repulsion energy of 30–50% of the energy calculated from equation 1. Pauling and Pressman assume the net energy to be 0.6 of the maximum. This constant factor appears to be justified in our present treatment where r does not vary to any large extent.

In the interaction of a group with the enzyme surface the water molecules formerly in contact with the group will be lost. In the present paper it is assumed that for a group (B) of approximately 2 Å. radius substituted to molecules of the type B–CH_2–CH_2, eight water molecules are lost /2 Å. radius unit in the interaction, and the dispersion energy due to the contact of eight water molecules must hence be subtracted.

Substitution in equation 1 gives

$$W = \frac{-38000 \times 0.6 ([R_{enz.}] - 8[R_{H_2O}])[R_B]}{(r_{enz.} + r_B)^6}$$

$$= \frac{-4.0 \times 10^6 [R_B]}{(r_B + 2)^6} \text{ (cal. mole}^{-1})$$

assuming $[R_{enz.}] = 47.2$ ml.,[28] and $r_{enz.} = 2.0$ Å.

Note Added in Proof

During the process of publication of this paper, there appeared a paper by Bergmann and Segal[29] on the relationship of quaternary ammonium salts to the anionic sites of true (electric eel) and pseudo (plasma) cholinesterase. On the basis of inhibition studies with inhibitors of the type $(CH_3)_3NC_nH_{2n+1}^{(+)}$ and $^{(+)}(CH_3)_3NC_nH_{2n}N(CH_3)_3^{(+)}$ these authors were led to the following conclusions: (1) Contrary to previous findings, interaction with pseudo cholinesterase involves one negative charge on the enzyme. (2) Interaction with true cholinesterase involves two negative sites on the enzyme which are located sufficiently close to be considered as one doubly negative site.

We wish to differ with the interpretation of these experiments. Consider the binding of tetramethylammonium ion (Table II, this paper) to true cholinesterase. K_S for this ion was calculated from the data of Wilson[11] and found to be 1.6×10^{-3} M, based on an I_{50} (the concentration of inhibitor which results in 50% inhibition of hydrolysis) of 1.6×10^{-2} M. The I_{50} reported in the new paper is essentially in agreement with this, however, by use of other values of the constants for the binding of substrate and self-inhibition by substrate, K_S is here[29] calculated as 6.3×10^{-4} M. We shall therefore normalize our values (from Table II) by means of a multiplication factor (0.4). This does not alter any previous calculations since all calculations were based on ratios. The total free energy of binding of the tetramethylammonium ion to true cholinesterase becomes $\Delta F = -RT \ln K_S = -4.5$ kcal. mole^{-1}. Subtracting from this value the true electrical contribution (-2.1 kcal. mole^{-1}) determined in section 3A we have -2.4 kcal. mole^{-1}, the resultant contribution of all other forces. Since the dispersion energy contribution alone would be about -2 kcal. (for three CH_3 groups and a small portion of exposed nitrogen), no other force can contribute appreciably to the free energy change. Suppose now, that the same ion were to interact with a spatially similar surface which was uncharged. The free energy change anticipated due to dispersion energy *only* would be about -2 kcal. mole^{-1}, exactly the value found by Bergmann and Segal for the binding of tetramethylammonium ion to pseudo cholinesterase, thus confirming the original hypothesis of an uncharged site in the pseudo compound. The contrary conclusions of the above paper resulted from the assignment of the entire binding free energy of tetramethylammonium ion to a coulombic type interaction (p. 695, par. 1 of ref. 29).

We can now turn to figures 1 and 2 of the paper of Bergmann and Segal. Plots of I_{50} *vs.* n for the ions $(CH_3)_3NC_nH_{2n+1}^{(+)}$ and $(CH_3)_3NC_nH_{2n}(CH_3)_3^{(+)}$ with true cholinesterase are shown. In figure 2 the free energy of binding increment becomes smaller with increase in n beyond $n = 3$ or 4 for the monoacidic bases, in agreement with our calculations (Table III, this paper). This effect, presumably due to large steric repulsions, would become even more marked with diacidic bases having the same value of n. For the diacidic bases (fig. 1, ref. 29) two effects will make large contributions to the total binding free energy increment, *viz.*, the steric effect and the location of the second positive charge. For the case of n small (≤ 3), the second charge will be close enough to the anionic site to contribute significantly to the binding. As n increases, the separation of the second charge from the site will begin to increase and steric repulsions will increase, both effects tending to counterbalance the gain in attractive dispersion energy. As n increases to even larger values, bending of the aliphatic chain may occur and repulsions may decrease by binding to a new non-specific surface. This in turn, may bring the second charge closer to the anionic charge and result in increasing free energy increments at very large n. On this basis a sigmoidal curve would be predicted for the binding constants of the dibasic acids to true cholinesterase, as a function of chain length. This explanation is in better agreement with the experimental facts than the assumption of a divalent anionic site, as long as the total free energy change with n in the region where the increment is at a maximum ($n = 6$ to 10) can be accounted for. From the data of figure 1, ref. 29, $\Delta F_{n=10} - \Delta F_{n=6}$ is calculated to be -3.9 kcal. mole^{-1}. Allowing for -500 cal. mole^{-1}/CH_2 (section 3A), the dispersion energy change contributes -2 kcal. mole^{-1}. The remaining -1.9 kcal. can be readily accounted for by the interaction of the second charged group, since the maximum calculated in section 3A was -2.1 kcal. mole^{-1} in a much higher dielectric medium. In the regions $n < 6$, $n > 10$, the ΔF increment ≥ -500 cal. mole^{-1}. That steric factors are appreciable at values of n below 6 is shown by the much less marked inhibition by the diacidic bases relative to the monoacidic bases with the same value of n.

PASADENA, CALIFORNIA

(28) A. N. Winchell, "The Optical Properties of Organic Compounds," Univ. of Wisconsin Press, Madison, 1943.

(29) F. Bergmann and R. Segal, *Biochem. J.*, **58**, 692 (1954).