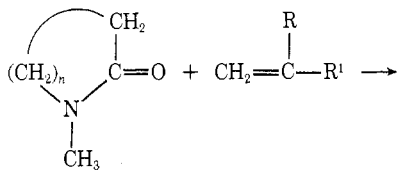
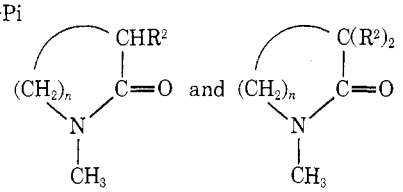
tion with the findings that HMPA is the best dipolar aprotic solvent.[57] A similar solvent effect was encountered in the addition of styrenes to alkylpyridines.[54]

Base-Catalyzed Carbon–Carbon Addition of "Aprotic Solvents"

In the course of an attempt to aralkylate 3-ethylpyridine with styrene using *t*-BuOK as catalyst in "aprotic" solvent *N*-methyl-2-pyrrolidone (NM-2-P), it was found that the solvent itself enters reaction.[58] A similar type of reaction occurred with *N*-methyl-2-piperidone (NM-2-Pi). The general scope of the reaction is represented by



$n = 2$ = NM-2-P
$n = 3$ = NM-2-Pi



R = H; R^1 = CH=CH_2; R^2 = CH_2CH=$CHCH_3$
R = H; R^1 = $C(CH_3)$=CH_2; R^2 = $CH_2CH_2CH(CH_3)_2$, $CH_2CH(CH_3)C_2H_5$ (after hydrogenation)
R = H; R^1 = C_6H_5; R^2 = $CH_2CH_2C_6H_5$
R = H; R^1 = $Si(CH_3)_3$; R^2 = $CH_2CH_2Si(CH_3)_3$
R = CH_3; R^1 = C_6H_5; R^2 = $CH_2CH(CH_3)C_6H_5$

(56) A. Schriesheim, *Amer. Chem. Soc., Div. Petrol. Chem., Prepr.*, **14**, D9 (1969).
(57) J. J. Delpuech, *Tetrahedron Lett.*, 2111 (1965).
(58) H. Pines, S. V. Kannan, and J. Simonik, *J. Org. Chem.*, **36**, 2311 (1971).

The addition reaction takes place at room temperature. In the case of less reactive olefins, such as α-methylstyrene or trimethylvinylsilane, it is essential to carry out the reaction in dimethyl sulfoxide. The yield of products, based on the olefins charged, was about 100%. NM-2-Pi reacts about twice as fast as NM-2-P.

Although DMSO is used extensively as an aprotic solvent, the hydrogens in DMSO are quite labile.[5,59,60] Dimethyl sulfoxide under the influence of base can react with a variety of conjugated dienic hydrocarbons,[61] arylalkenes,[62] and aromatic and heteroaromatic ring compounds.[61,63,64] The net result of these reactions is the formation of methylated analogs of the starting hydrocarbons.

The methylation reactions can be explained by a mechanism using butadiene as an example.[61]

$$CH_3SOCH_3 + t\text{-}BuO^-K^+ \rightleftharpoons \underset{\mathbf{52}}{CH_3SOCH_2^-K^+} + t\text{-}BuOH$$

$$\mathbf{52} + CH_2{=}CHCH{=}CH_2 \rightleftharpoons \underset{\mathbf{53}}{[CH_3SOCH_2CH_2CH{\cdots}CH{\cdots}CH_2]^-K^+}$$

$$\mathbf{53} + BH \longrightarrow \underset{\mathbf{54}}{CH_2SOCH_2CH_2CH{=}CHCH_3} + B^-K^+$$

$$\mathbf{54} + B^- \longrightarrow CH_3SOCH_2\bar{C}HCH{=}CHCH_3 + BH$$

$$CH_3SOCH_2\bar{C}HCH{=}CHCH_3 \longrightarrow CH_3SO^- + CH_2{=}CHCH{=}CHCH_3$$

I wish to thank the collaborators listed in the references for their invaluable contribution to the progress and development of the work.

(59) J. E. Hofmann, R. J. Muller, and A. Schriesheim, *J. Amer. Chem. Soc.*, **83**, 3731 (1961); **85**, 3002 (1963).
(60) C. D. Ritchie and R. E. Uschold, *J. Amer. Chem. Soc.*, **89**, 2960 (1967), and references therein.
(61) P. A. Argabright, J. E. Hofmann, and A Schriesheim, *J. Org. Chem.*, **30**, 3233 (1965).
(62) M. Feldman, S. Danischefsky, and R. Levine, *J. Org. Chem.*, **31**, 4322 (1966).
(63) G. A. Russell and S. A. Weiner, *J. Org. Chem.*, **31**, 248 (1966).
(64) H. Nozaki, Y. Yamamoto, and R. Noyori, *Tetrahedron Lett.*, 1123 (1966).

Protein Interactions with Small Molecules

Irving M. Klotz

Department of Chemistry, Northwestern University, Evanston, Illinois 60201

Received December 19, 1973

Binding of a small molecule by a protein is the pivotal step in a host of biological functions. One of the earliest recognized, and now most familiar, of such interactions is the uptake of O_2 by hemoglobin, where binding inaugurates the transport of molecular oxygen throughout the vascular system. It has been increasingly realized that similar interactions appear at all stages in biochemical and physiological functions, from the cellular level to the organismic.

From a biochemical viewpoint our perception of the molecular nature of these interactions is sharpened if we categorize them in terms of the conformational accommodation that plays a dominant role in the functioning of the protein–small molecule complex (Figure 1).[1] In some situations the conformational adaptation of the protein (Figure 1A) is the

I. M. Klotz was born in Chicago and received his bachelor's degree and the Ph.D. (in 1940) from the University of Chicago. He then moved across town to Northwestern University where he is now Morrison Professor of Chemistry and Biology. He is a member of the National Academy of Sciences.

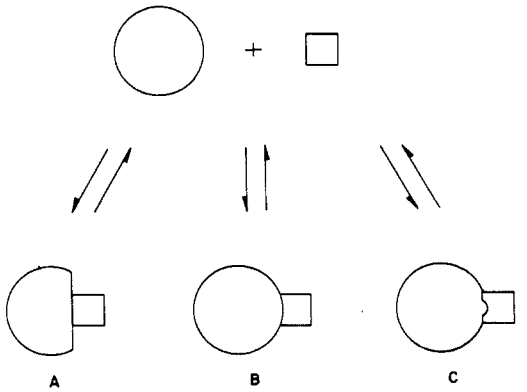


Figure 1. Schematic representation of types[1] of conformational accommodation in complexes of proteins (large circle) with small molecules (small square).

crucial feature; this is typical of interactions serving in a regulatory capacity. Alternatively, the change in conformation of the small ligand may be the center of focus (Figure 1B); this is characteristic of enzymatically catalyzed molecular transformations. Finally there may be a mutual conformational accommodation of both protein and small molecule with both constituents being stabilized by this interaction (Figure 1C); transport functions can readily be fulfilled under these circumstances.

Combinations that participate in a regulatory mechanism generally depend on conformational adaptations[3] in the macromolecule. One of the most interesting examples of these is provided again by hemoglobin, in its interactions with phosphates, particularly 2,3-diphosphoglycerate.[4,5] This small organic anion is bound preferentially by the deoxy form of hemoglobin. As a result the small anion shifts the conformer equilibrium toward the deoxy conformation, whose affinity for O_2 is less than that of the oxy combination. Hence 2,3-diphosphoglycerate decreases oxygen binding. An extraneous phosphate thus can modulate oxygen uptake and release by the protein.

Alternatively, small-molecule binding can affect the combinations between different proteins that, in concert, constitute the biologically active and regulatable species (Tables[6] I and II). One striking example of this phenomenon is provided by the enzyme aspartate transcarbamylase, designated as $\alpha_6\beta_6$ since it is constituted of six catalytic subunits (α) and six regulatory subunits (β). If Zn^{2+} is removed from isolated regulatory subunits this apoprotein is unable to associate with catalytic subunits to generate the $\alpha_6\beta_6$ ensemble.[7] Thus binding of Zn^{2+} plays a vital

Table I
Some Proteins Constituted of Nonidentical Subunits[6]

Protein	Mol wt	Subunit composition
Lactose synthetase	60,000	$\alpha\beta$
Procarboxypeptidase	88,000	$\alpha\beta\gamma$
Fructose diphosphatase	130,000	$\alpha_2\beta_2$
Tryptophan synthetase	148,000	$\alpha_2\beta_2$
Histidine decarboxylase	190,000	$\alpha_5\beta_5$
Aspartate transcarbamylase	310,000	$\alpha_6\beta_6$
Membrane ATPase	385,000	$\alpha_6\beta_6$
TPNH–sulfite reductase	700,000	$\alpha_4\beta_8$

Table II
Ligand-Linked Association or Dissociation of Proteins[6]

Protein	Ligand	Action of ligand
Glucose-6-phosphate dehydrogenase	NADPH	Promotes dissociation
	NADP	Promotes association
Tryptophan synthetase (B component)	Pyridoxal phosphate	Promotes association
Glyceraldehyde-3-phosphate dehydrogenase	ATP or cyclic AMP	Promotes dissociation
Isocitrate dehydrogenase	Citrate or isocitrate	Promotes dissociation
Hemerythrin	Monovalent anions	Promote dissociation
δ-Aminolevulinate dehydratase	Potassium ions	Promote association
Glutamate dehydrogenase	NADPH, GTP	Promote dissociation
	ADP or ATP	Promotes association

role in the assembly and maintenance of the native enzyme. The converse situation is also known (Table II): binding of a small molecule may dissociate a heteromeric subunit assembly to generate the active species. In the cAMP (cyclic adenosine monophosphate) dependent kinases, the enzyme exists normally as a catalytic regulatory ensemble that is *inactive;* the binding of cAMP to the regulatory subunit dissociates the ensemble into the free *active* catalytic subunit and the separated AMP-binding regulatory protein.[8-10] Small-molecule binding can also modulate the state of assembly and biological activity of homomeric subunit ensembles. For example, citrate ion dissociates the inactive aggregated assembly of the enzyme isocitrate dehydrogenase into subunit active form.[11]

The combinations of substrates with enzymes provide numerous examples of the type of interaction represented in Figure 1B. The first step in any enzymic process is the binding of the small-molecule substrate with the macromolecular catalyst. It is in this complex that the transition to the conformation of the activated state is facilitated, with concomitant

(1) When a macromolecule combines with a small molecule, the local environment (and hence chemical potential) of each one changes. Strictly speaking, the chemical potential of one species cannot change without some compensating change in that of the other.[2] Thus a conformational adaptation of the macromolecule must be paralleled by one of the small ligand, *i.e., the conformational accommodation must be mutual.*

(2) I. M. Klotz, *Arch. Biochem. Biophys.*, **116**, 92 (1966).

(3) In principle, all accessible conformers are in equilibrium with each other. When we speak of some ligand-linked conformational accommodation, we are really referring to a shift in equilibrium toward a conformer of a particular spatial structure.

(4) R. Benesch and R. E. Benesch, *Biochem. Biophys. Res. Commun.*, **26**, 162 (1967).

(5) A. Chanutin and R. R. Curnish, *Arch. Biochem. Biophys.*, **121**, 96 (1967).

(6) I. M. Klotz, D. W. Darnall, and N. R. Langerman, *Proteins, 3rd Ed.*, in press.

(7) M. E. Nelbach, V. P. Pigiet, Jr., J. C. Gerhart, and H. K. Schachman, *Biochemistry*, **11**, 315 (1972).

(8) G. N. Gill and L. D. Garren, *Biochem. Biophys. Res. Commun.*, **39**, 335 (1970).

(9) E. M. Reimann, C. O. Brostrom, J. D. Corbin, C. A. King, and E. G. Krebs, *Biochem. Biophys. Res. Commun.*, **42**, 187 (1971).

(10) M. Tao, M. L. Salas, and F. Lipmann, *Proc. Nat. Acad. Sci. U. S.*, **67**, 408 (1970).

(11) H. B. LeJohn, B. E. McCrea, I. Suzuki, and S. Jackson, *J. Biol. Chem.*, **244**, 2484 (1969).

remarkable accentuations in rates of molecular transformations. This feature of enzyme–substrate complexes has long been recognized on theoretical grounds.[12-15] It has also been directly demonstrated recently with the increase in resolution of protein structures by X-ray diffraction methods. In lysozyme,[16] for example, the active site has been delineated, on the basis of a structure determination of an enzyme–inhibitor complex, as a crevice running down the entire length of a median of the macromolecule. When a hexasaccharide substrate is bound to lysozyme, the fourth (D) hexose ring, an *N*-acetylmuramic acid residue, adopts a half-chair conformation. This conformation, favored in the enzyme environment, stabilizes the carbonium ion transition state[17] of the reaction pathway, leading to cleavage of the glycosidic bond between the D and E hexose rings. Stabilization of the transition state of the enzyme-bound substrate lowers the free energy of activation, and hence increases the rate of the bond-breaking process.

Many substances combine with proteins to facilitate their transport throughout the vascular system. Under these circumstances there is probably a mutual conformational accommodation (Figure 1C), but this merely stabilizes the complex and does not trigger a physiological process for either the protein or the ligand. Transport proteins may be very specific or very broad, catholic ones. Examples of the former are transferrin, retinol-binding protein, and thyroxine-binding globulin, the first specifically combining with Fe, the second with the visual pigment precursor retinol, and the third with the hormone thyroxine. Broad, nonspecific combinations, in turn, are seen in the many different adducts of serum albumin with metals, metabolites, drugs, etc. These are all of importance in transport processes.

An exploration of any of the interactions represented in Figure 1 needs to examine the distribution of small molecule between binding macromolecule and bulk solvent. This distribution is a manifestation of the stoichiometry of the combinations and the strength of the interactions. The extent of uptake of a small molecule by a macromolecule depends on the number of unoccupied sites available on the latter and on the chemical potential or concentration of the nonbound species of the former. We shall proceed, therefore, to analyze quantitatively these stoichiometric and energetic aspects of binding of small molecules by macromolecules.

Stoichiometric Binding Constants

Recognizing that a given protein macromolecule, P, may have a multiplet of sites for binding a specific ligand A, we express the mass law relationship in terms of a series of multiple equilibria. Two different formulations may be used for this purpose, one being *stoichiometric* in its outlook, the other being *site* oriented.

In the *stoichiometric* formulation one asks, in essence, what are the concentrations of sequential stoichiometric species PA_1, PA_2, etc. For PA_1 we do not distinguish between different single-liganded species with A at different possible sites 1, 2, . . . , or j, but consider the total sum of these as constituents of (PA_1). Similarly for PA_2 we visualize all different species with two A ligands attached to different pairs of sites as contributing to the concentration of (PA_2). These stoichiometric species participate in equilibria 1–3. For these equilibria we may define *stoichiometric equilibrium constants*

$$P + A \rightleftharpoons PA_1 \tag{1}$$

$$PA_1 + A \rightleftharpoons PA_2 \tag{2}$$

$$\vdots \quad\quad \vdots$$

$$PA_{i-1} + A \rightleftharpoons PA_i \tag{3}$$

$$K_1 = \frac{(PA_1)}{(P)(A)} \qquad (PA_1) = K_1(P)(A) \tag{4}$$

$$K_2 = \frac{(PA_2)}{(PA_1)(A)} \qquad (PA_2) = K_1K_2(P)(A)^2 \tag{5}$$

$$K_i = \frac{(PA_i)}{(PA_{i-1})(A)} \qquad (PA_i) = (K_1K_2 \ldots K_i)(P)(A)^i \tag{6}$$

Experimentally we measure r, the moles of bound A per mole of total protein. In terms of the stoichiometric species contributing to r we may write[18-20]

$$r = \frac{(PA_1) + 2(PA_2) + \ldots + i(PA_i) + \ldots + n(PA_n)}{(P) + (PA) + (PA_2) + \ldots + (PA_i) + \ldots + (PA_n)} = \frac{\sum_1^n i(PA_i)}{\sum_0^n (PA_i)} \tag{7}$$

where n is the total number of binding sites on each protein molecule. This equation can also be reduced to an alternative one expressed in terms of the stoichiometric equilibrium constants[19,20]

$$r = \frac{K_1(A) + 2K_1K_2(A)^2 + \ldots}{1 + K_1(A) + K_1K_2(A)^2 + \ldots} = \frac{\sum_1^n i\left(\prod_1^i K_1\right)(A)^i}{1 + \sum_1^n \left(\prod_1^i K_1\right)(A)^i} \tag{8}$$

Site Binding Constants

An alternative formulation of the multiple equilibria of P with A focuses on the individual *sites of* binding. If each site has a fixed affinity that does not change with extent of occupancy of other sites by ligands, then the multiple equilibria can be represented as (the lower left subscript in ${}_jP$ referring to the site of binding)

$${}_1P + A \rightleftharpoons {}_1PA \tag{9}$$

$${}_2P + A \rightleftharpoons {}_2PA \tag{10}$$

$$\vdots \quad\quad \vdots$$

$${}_jP + A \rightleftharpoons {}_jPA \tag{11}$$

For each site we can define a *site equilibrium constant*

(12) L. Pauling, *Amer. Sci.*, **36**, 58 (1948).
(13) R. Wolfenden, *Accounts Chem. Res.*, **5**, 10 (1972).
(14) G. E. Lienhard, *Science*, **180**, 149 (1973).
(15) C. DeLisi and D. M. Crothers, *Biopolymers*, **12**, 1689 (1973).
(16) C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips, and V. R. Sarma, *Proc. Roy. Soc., Ser. B*, **167**, 378 (1967).
(17) R. N. Lemieux and G. Huber, *Can. J. Chem.*, **33**, 128 (1955).
(18) I. M. Klotz, *Arch. Biochem.*, **9**, 109, (1946).
(19) I. M., Klotz, F. M. Walker, and R. B. Pivan, *J. Amer. Chem. Soc.*, **68**, 1486 (1946).
(20) I. M. Klotz, *Proteins, 1st Ed.*, **1B**, 727 (1953).

$$k_1 = \frac{(_1\mathrm{PA})}{(_1\mathrm{P})(\mathrm{A})} \quad (12)$$

$$k_2 = \frac{(_2\mathrm{PA})}{(_2\mathrm{PA})(\mathrm{A})} \quad (13)$$

$$\vdots \qquad \vdots$$

$$k_j = \frac{(_j\mathrm{PA})}{(_j\mathrm{P})(\mathrm{A})} \quad (14)$$

For any single site, r_j, the moles of bound A at that site per mole of total protein is given by

$$r_j = \frac{k_j(\mathrm{A})}{1 + k_j(\mathrm{A})} \quad (15)$$

The total moles of bound ligand at all sites is[21-23]

$$r = \sum_1^n r_j = \sum_1^n \frac{k_j(\mathrm{A})}{1 + k_j(\mathrm{A})} \quad (16)$$

Relations between Stoichiometric and Site Binding Constants

Since r of eq 8 is the same experimental quantity as r of eq 16, the *stoichiometric equilibrium constants*, K_i, must be related to the *site equilibrium constants*, k_j. For a two-site binding species, the relationships between K_i and k_j, worked out almost 60 years ago,[24] are

$$K_1 = k_1 + k_2 \quad (17)$$

$$K_2 = \frac{k_1 k_2}{k_1 + k_2} \quad (18)$$

or

$$K_1 K_2 = k_1 k_2 \quad (19)$$

As a prelude to working out the relationships between K_i's and k_j's for multisite binders with fixed unchanging affinities, let us examine the species in a two-site system

Stoichiometric species	PA_0	PA_1	PA_2
Site species	P	$_1PA$	$_{1,2}PA_2$
		$_2PA$	
Total site species	1	2	1

To go from P to $_1PA$ we must put a ligand on site 1; this equilibrium is governed by k_1. To go from P to $_2PA$ we must put a ligand on site 2; this equilibrium is governed by k_2. The stoichiometric equilibrium constant K_1 (eq 4) is thus expressed by eq 20. Correspondingly, to go from P to PA_2 we must put on two ligands in succession, one on site 1 followed by the other on site 2, or *vice versa*. In either event,[25] for the stoichiometric product constant K_1K_2 we find

$$K_1 = \frac{(\mathrm{PA}_1)}{(\mathrm{P})(\mathrm{A})} = \frac{[(_1\mathrm{PA}) + (_2\mathrm{PA})]}{(\mathrm{P})(\mathrm{A})} = k_1 + k_2 \quad (20)$$

$$K_1 K_2 = \frac{(\mathrm{PA}_2)}{(\mathrm{P})(\mathrm{A})^2} = k_1 k_2 = k_2 k_1 \quad (21)$$

For a three-site system the relationships between the *site* species and the *stoichiometric* species may be represented as follows

Stoichiometric species	PA_0	PA_1	PA_2	PA_3
	P	$_1PA$	$_{1,2}PA_2$	$_{1,2,3}PA_3$
Site species		$_2PA$	$_{1,3}PA_2$	
		$_3PA$	$_{2,3}PA_2$	
Total site species	1	3	3	1

The binding of a single ligand to each of the site species $_1PA$, $_2PA$, $_3PA$, is governed by the site constants k_1, k_2, k_3, respectively. The stoichiometric equilibrium constant, K_1, thus can be expressed by

$$K_1 = \frac{(\mathrm{PA}_1)}{(\mathrm{P})(\mathrm{A})} = \frac{[(_1\mathrm{PA}) + (_2\mathrm{PA}) + (_3\mathrm{PA})]}{(\mathrm{P})(\mathrm{A})} = k_1 + k_2 + k_3 \quad (22)$$

Correspondingly, to go from P to PA_2 we must put on two ligands in succession to form site-species $_{1,2}PA_2$, also two to form $_{1,3}PA_2$ and then two to form $_{2,3}PA_2$. Each of these site-filling processes is governed by a product site equilibrium constant, k_1k_2, k_1k_3, and k_2k_3, respectively. Consequently, we can show that

$$K_1K_2 = \frac{(\mathrm{PA}_2)}{(\mathrm{P})(\mathrm{A})^2} = \frac{[(_{1,2}\mathrm{PA}_2) + (_{1,3}\mathrm{PA}_2) + (_{2,3}\mathrm{PA}_2)]}{(\mathrm{P})(\mathrm{A})} = k_1k_2 + k_1k_3 + k_2k_3 \quad (23)$$

Finally, to go from P to PA_3 we must put on three ligands in succession. The order of filling the three sites does not matter since the final result in each path is the same, $_{1,2,3}PA_3$. Thus we can take the ordinal sequential path, in which the overall equilibrium is governed by the site constant $k_1k_2k_3$. In any event we can show that

$$K_1K_2K_3 = \frac{(\mathrm{PA}_3)}{(\mathrm{P})(\mathrm{A})^3} = \frac{(_{1,2,3}\mathrm{PA}_3)}{(\mathrm{P})(\mathrm{A})^3} = k_1k_2k_3 \quad (24)$$

Corresponding analyses of the equilibria for attachment of ligands in succession to obtain the various site species that are constituents of the stoichiometric class PA_i leads to the following relationships between stoichiometric and site equilibrium constants[26]

$$K_1 = k_1 + k_2 + \ldots + k_n = \sum_{j_1=1}^{n} k_{j_1} \quad (25)$$

$$K_1K_2 = k_1k_2 + k_1k_3 + \ldots + k_1k_n + k_2k_3 + k_2k_4 + \ldots = \sum_{j_1=1}^{n-1} \sum_{j_2=j_1+1}^{n} k_{j_1}k_{j_2} \quad (26)$$

$$\vdots$$

$$K_1K_2\ldots K_i = \sum_{j_1=1}^{n-i+1} \sum_{j_2=j_1+1}^{n-i+2} \cdots \sum_{j_i=j_{i-1}+1}^{n} k_{j_1}k_{j_2}\ldots k_{j_i} \quad (27)$$

$$\vdots$$

$$K_1K_2\ldots K_n = k_1k_2\ldots k_n \quad (28)$$

The stoichiometric formulation is always applicable to the correlation of experimental binding data since it is in essence a thermodynamic analysis. It implicitly includes any cooperative or antagonistic interactions between successively bound ligands, due to conformational changes, electrostatic interactions, etc. In contrast the fixed-affinity, specific-site representation visualizes each of the binding sites as per-

(21) G. Scatchard, *Ann. N. Y. Acad. Sci.*, **51**, 660 (1949).
(22) G. Scatchard, I. H. Scheinberg, and S. H. Armstrong, Jr., *J. Amer. Chem. Soc.*, **72**, 535, 540 (1950).
(23) F. Karush, *J. Amer. Chem. Soc.*, **72**, 2705 (1950).
(24) E. Q. Adams, *J. Amer. Chem. Soc.*, **38**, 1503 (1916).
(25) For thermodynamic purposes the path used does not matter so long as the initial state for each path is the same and the final state of each path is identical.
(26) An alternative formulation of eq 25-28 derived from a purely formal algebraic analysis of eq 8 and 16 has been described by J. E. Fletcher., A. A. Spector, and J. D. Ashbrook, *Biochemistry*, **9**, 4580 (1970).

manently fixed in the macromolecules and possessing a characteristic independent affinity for a particular ligand. Conformational accommodations associated with successive ligand binding, which can lead to the abolition, creation, or modification of binding sites, is not encompassed in the fixed-affinity specific-site model.

An illustration of the type of difficulty that may be encountered in a site-constant analysis is provided by a very simple two-site association, the uptake of H^+ ions by aminoethylmercaptide ion, $H_2NCH_2CH_2S^-$. A stoichiometric analysis in parallel with a site analysis may be represented in eq 29 and 30, the H^+ ion uptake being implicit in each step.

$$P^- \overset{K_1}{\rightleftharpoons} PH \overset{K_2}{\rightleftharpoons} PH_2^+ \qquad (29)$$

$$\begin{array}{ccc} & HS(CH_2)_2NH_2 & \\ {}^{k_1}\swarrow & & \searrow^{k_2'} \\ {}^-S(CH_2)_2NH_2 & & HS(CH_2)_2NH_3^+ \\ {}_{k_2}\searrow & & \swarrow_{k_1'} \\ & {}^-S(CH_2)_2NH_3^+ & \end{array} \qquad (30)$$

The H^+-uptake curve throughout the entire pH range can be described in terms of the two stoichiometric equilibrium constants K_1 and K_2. On the other hand, it cannot be fitted by the two site parameters k_1 and k_2 for the S^- and NH_2 sites, respectively. Two additional parameters, k_1' and k_2' (only one of which is independent) are necessary. This situation arises because the uptake of H^+ by the amino group in ${}^-S(CH_2)_2NH_2$ is stronger than in $HS(CH_2)_2NH_2$. A corresponding statement can be made for the S^- group in a comparison of H^+ uptake of ${}^-S(CH_2)_2NH_2$ with ${}^-S(CH_2)_2NH_3^+$. In essence then there is an interaction between sites which can change each site equilibrium constant as a ligand is bound at a different site.

For the completely general case of n binding sites whose affinities change with extent of occupancy by ligands, it is necessary to define $2^n - {}^1n$ different site binding constants.[27] Thus for a three-site protein, there are 12 site binding constants, whereas there are only three stoichiometric binding constants. If for a three-site system one uses an equation of the form of (16) with three terms in it, the three k_j parameters that fit the data are *not* the site binding constants and, in general, have no unique relationship to them.[27] The three stoichiometric constants (of eq 8), however, can be related to the site binding constants.[27]

Graphical Representations

Having formulated the equilibrium relationships in a multistep binding process, we are in a position to apply the analysis to specific ligand–protein complexes. Prior to such a treatment it is advantageous to scan the entire range of experimental data. For this purpose graphical presentations are usually most suitable.

The most straightforward set of coordinates would seem to be r *vs.* (A). An example of such a graph, based on an extensive study of the binding of leucine by α-isopropylmalate synthase,[28] is illustrated in Figure 2. This plot illustrates the major disadvantage of a graph with a linear scale for the concentration of free ligand, (A); the data are very compressed at the left end of the scale. Furthermore, this defect becomes magnified the more extensive is the experimental study, for the wider the range of concentrations covered experimentally (about a 1000-fold in Figure 2), the greater the compression of information at the dilute end of ligand concentration.

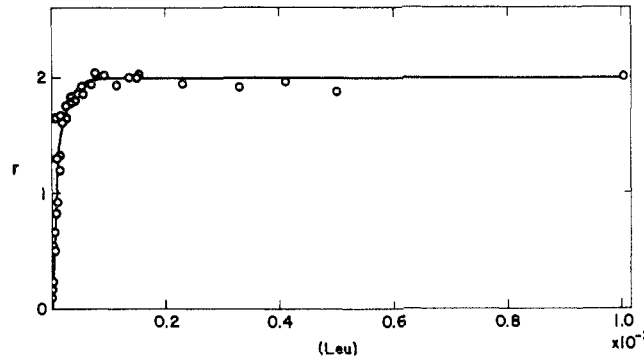


Figure 2. Graphical representation of data for binding of leucine by α-isopropylmalate synthase[28] using linear scales for the experimental variables r and (A). Additional data points, besides those published, were kindly provided by Dr. G. B. Kohlhaw.

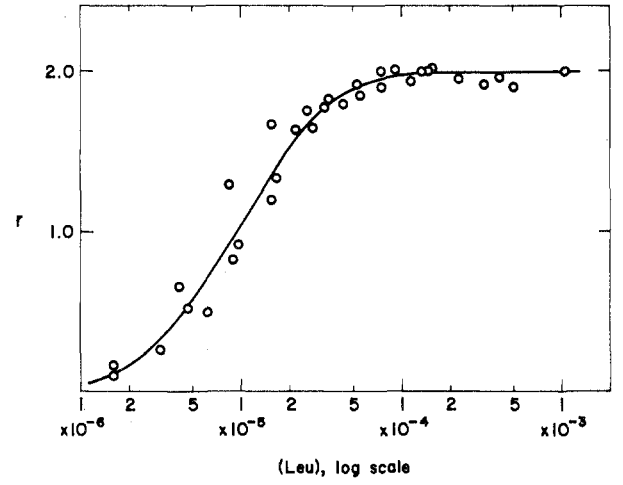


Figure 3. Semilogarithmic representation of data for binding of leucine by α-isopropylmate synthase. Additional data points provided by Dr. G. B. Kohlhaw.

To spread the observations out more uniformly one should use a logarithmic scale[19,29] for (A). The marked improvement in depiction of the binding of leucine by isopropylmalate synthase is illustrated in Figure 3. Clearly the compression of information at low ligand concentrations (Figure 2) has been eliminated, and the observations over a thousandfold range in concentration of (A) are evenly spread.

The study illustrated in Figure 3 is one of the extremely few examples (outside of O_2 binding by oxygen-carrying proteins) of actual attainment of saturation within the range of experimental measurement. This was feasible because the binding constant for leucine–isopropylmalate synthase is very high (*ca.* 10^5) and saturation is reached below 10^{-4} M leucine. More typical behavior is shown in, for example, the binding of copper by serum albumin[30]

(27) I. M. Klotz and D. L. Hunston, in preparation.
(28) E. Teng-Leary and G. B. Kohlhaw, *Biochemistry*, **12**, 2980 (1973).
(29) C. J. Thompson and I. M. Klotz, *Arch. Biochem. Biophys.*, **147**, 178 (1971).
(30) I. M. Klotz and H. G. Curme, *J. Amer. Chem. Soc.*, **70**, 939 (1948).

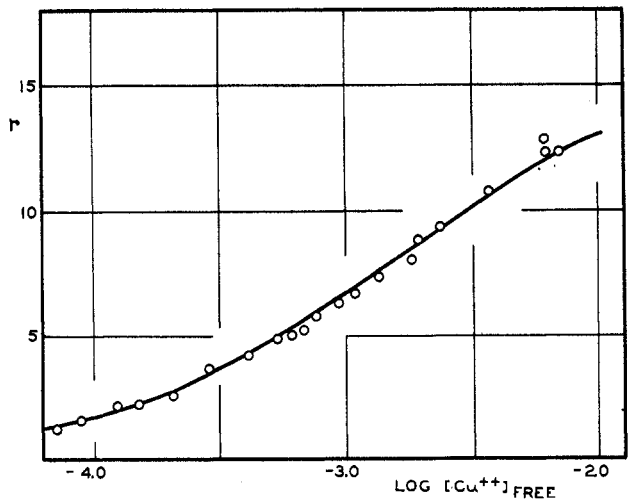


Figure 4. Binding of Cu^{2+} by bovine serum albumin.[30]

(Figure 4). Despite the fact that a 100-fold range in concentration of free Cu^{2+} has been spanned, a definitive saturation plateau has not been reached although there is a hint that one is being approached. In most binding studies, furthermore, the range in concentration of free ligand is not very large, very frequently less than tenfold. The limited scope of the data is readily visualized in an r–log (A) graph but is often obscured, as we shall illustrate in more detail below, in some alternative graphical representations, for example, graphs of $r/(A)$ *vs.* r.

Despite the advantages of graphs with a logarithmic scale for (A), it is very common to find binding data presented with a scale of $1/(A)$, in particular in graphs of $r/(A)$ *vs.* r. These coordinates were used first by Scatchard[21] in analyzing protein binding and correspond to the variables recommended earlier by Eadie[31] for establishing Michaelis–Menten parameters in enzyme kinetics. Any scale using $1/(A)$ of neccesity compresses data very strongly as (A) becomes large.[29] As a result, in a plot using $1/(A)$ it is tempting and feasible to put a line (or curve) through the data at high concentrations that is concordant with one's predilections but which would be clearly unacceptable if translated to a logarithmic graph.

These problems with reciprocal graphs can be illustrated with data taken from a careful study of the binding of acetylcholine to a protein from electroplax membranes.[32] Three graphical representations of the data are illustrated in Figure 5. In Figure 5a, the coordinates are $r/(A)$ *vs.* r, but numerical values along the abscissa have been blotted out to eclipse subjective influences on a decision as to the intercept on the r axis. From points shown it is clearly impossible to decide whether a curve connecting them would ever cross the x axis. When the coordinates of the abscissa are labeled explicitly (Figure 5b), one is greatly tempted to draw a curve consisting of two linear sections, the r intercept of the first being sighted near 1, attributed to a single strong binding site,[33] and the r intercept of the second being placed at 10, the presumed total number of binding sites. However, when the same data are plotted on a semilogarithmic graph (Figure 5c) it is obvious that an inflection point, if it has been reached at all, must occur at $r \geq 7$. Since the inflection point in an r–log (A) graph of binding by noninteracting equivalent sites appears half-way toward saturation, the saturation plateau must be near 14. Thus the true intercept on the abscissa of Figure 5b must actually be at a much higher value than that shown, in fact well beyond the terminus of the scale published.

(31) G. S. Eadie, *J. Biol. Chem.*, 146, 85 (1942).
(32) E. DeRobertis, G. S. Lunt, and J. L. La Torre, *Mol. Pharmacol.* 7, 97, (1971).
(33) Actually if there are two classes of binding sites the r intercept of the first line of Figure 5b is not really the number of sites in the strongest binding class but a more complicated function that includes the binding constants of both classes. See ref 34.
(34) I. M. Klotz and D. L. Hunston, *Biochemistry*, 10, 3065 (1971).

Evaluation of Binding Parameters

The most precise method of establishing binding parameters is to fit the data to the function $r = g[(A), K_1, K_2 \ldots K_n]$ using nonlinear least-squares curve-fitting procedures. With the general availability of high-speed computers, the tedium in making the necessary computations has been removed.

An analysis of the binding of palmitic acid by serum albumin[26] provides an excellent illustrative example of nonlinear least-squares fitting to data. In this example the best fit with the fewest parameters is obtained for $n = 8$ and the values of K_1 to K_8 shown (implicitly) in Figure 6. This *affinity profile* illustrates the relationships among the stoichiometric binding constants.

From an affinity profile it is simple to establish whether successive binding steps act cooperatively or antagonistically (and even to calculate the stepwise interaction energies[27]). If the binding sites are completely independent and equivalent in intrinsic affinity, then the (ideal) stoichiometric binding constants are related by eq 31[18,20]

$$(K_i)_{\text{ideal}} = \frac{n - i + 1}{i} K \tag{31}$$

From this relation it follows that

$$i(K_i)_{\text{ideal}} = K(n + 1) - Ki \tag{32}$$

in other words that $i(K_i)_{\text{ideal}}$ is linear in i. If the slope of the line connecting any actual K_{i-1} with K_i is greater than the slope of the ideal line, *accentuating interactions* must be present, that is, the uptake of ligand in the ith stoichiometric step accentuates the binding affinity. Contrariwise, if the slope of the line connecting two points is less than that of the ideal line, then *attenuating interactions* are associated with the uptake of ligand in the ith stoichiometric step.

The results of most binding investigations, however, have not been subjected to computer least-squares fitting. Graphical procedures have generally been resorted to to obtain n and the binding constants.

If saturation is actually attained, there is no difficulty in establishing n. In practice, however, most interactions are not so strong. Consequently it is necessary to extrapolate observed values of r to (A) $\rightarrow \infty$, a hazardous procedure. In practice, this is attempted by plots in which $1/(A)$, by itself or as the function $r(1/(A))$, is used as one of the coordinates. If one is reasonably certain that the binding sites are equivalent and noninteracting, eq 8 can be reduced to a simple hyperbolic form[18,20] and graphs of $1/r$ *vs.* $1/(A)$ or of $r/(A)$ *vs.* r may give reliable extrapola-

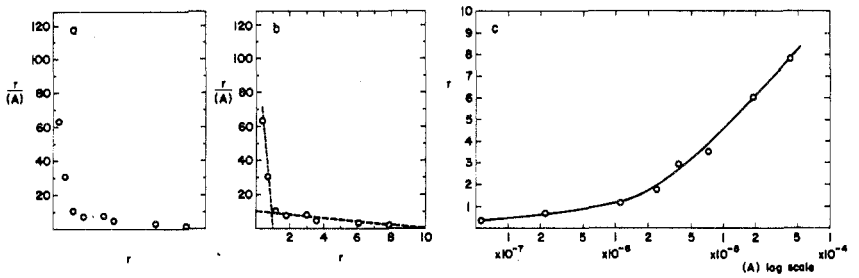


Figure 5. Three plots of data for binding of acetylcholine by a protein from electroplax membranes:[32] (a) $r/(A)$ *vs.* r but with numerical values omitted from abscissa; (b) same as plot a with numerical values explicitly shown along coordinate axes; (c) a graph with logarithmic scale for (A).

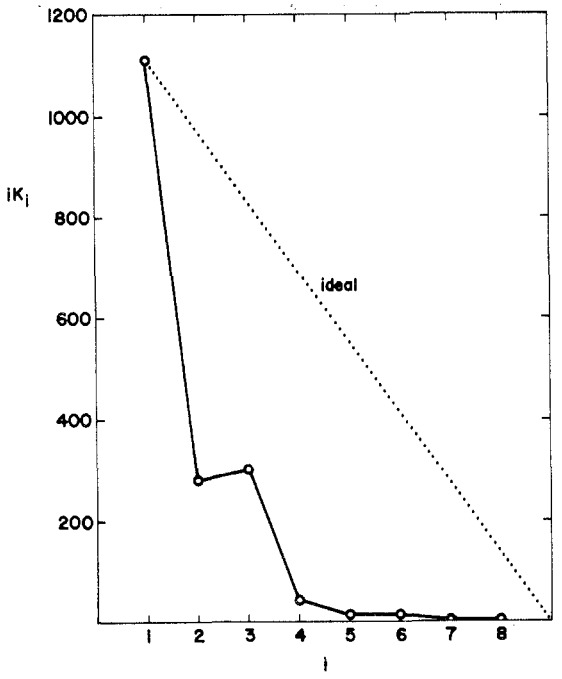


Figure 6. Affinity profile of stoichiometric binding constants for binding of palmitic acid by human serum albumin.[26] The dotted line illustrates the linear decline in iK_i for the limiting ideal situation where all the binding sites are equivalent and independent.

tions to n since they are linearized representations of the hyperbolic form. In practice one can rarely be confident that the sites are noninteracting, even if they are identical (as in homomeric subunit quaternary structures), for the addition of even the first ligand introduces asymmetry into the quaternary structure. Since any scale in $1/(A)$ (or $r/(A)$) compresses data severely as (A) becomes large, there is a strong temptation to draw a straight line through the data. Furthermore, even when it is obvious that the data do not fit a linear relation in a graph of $r/(A)$ *vs.* r, there is an almost overpowering temptation to fit the observations to a curve that intersects the abscissa at some point that one has been predisposed to consider reasonable. If the same data are plotted in a graph in which the axes are numerically *un*marked, as in the two examples illustrated in Figure 7,[35,36] it is readily apparent that the binding observations in themselves are unable to establish the point of interception with the x coordinate.

Conclusion

As this Account indicates, the distribution of a ligand between bulk solvent and a binding macromolecule can be described most economically and definitively in terms of stepwise stoichiometric equilibria. For a molecule with n binding sites, the number of stoichiometric binding constants is simply n. Interactions between sites, whether they accentuate or attenuate binding affinities, manifest themselves in K_i, since the stoichiometric equilibrium constant implicitly reflects the properties of its constituent site equilibrium constants.[27] In particular an *affinity profile*, a graph of iK_i *vs.* i, reveals at a glance the nature of the interactions. Whether or not there are interactions between sites, one can evaluate the first binding constant, K_1, and with this make energetic comparisons between different types of ligands or different classes of proteins.

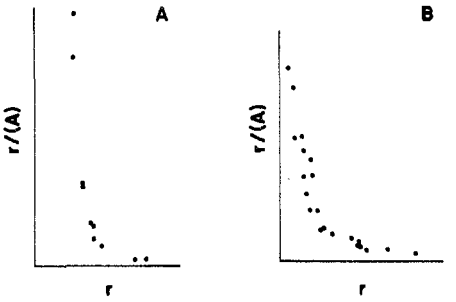


Figure 7. Binding data plotted as $r/(A)$ *vs.* r but with numerical values removed from coordinates: (A) binding of a hapten by an immoglobulin,[35] (B) binding of CTP by aspartate transcarbamylase.[36] In each example the authors conclude an intercept occurs on the abscissa at that point at which we have terminated the axis.

Of course from K_i's alone one obtains no direct insight into the molecular nature of specific individual binding sites. A thermodynamic analysis in itself cannot reveal molecular features, although it does establish the constraints within which molecular interpretations must operate. Molecular details of the interactions of proteins with small molecules can be disclosed by a variety of spectroscopic probes. Interactions of atomic groupings with electromagnetic radiation reflect the nature of the constituent atoms, their environments, and their dynamic properties. Different frequency ranges in the electromagnetic spectrum provide different types of information. Studies with spectroscopic probes of all types are needed to complement stoichiometric and energetic ones if we are to establish a broad foundation for understanding protein interactions with small molecules.[37]

(35) R. G. Painter, H. J. Sage, and C. Tanford, *Biochemistry*, 11, 1338 (1972).

(36) S. Matsumoto and G. G. Hammes, *Biochemistry*, 12, 1388 (1973).

(37) Much of the material described in this review was developed in investigations supported in part by grants from the National Science Foundation (No. GB 35206) and the National Institute of General Medical Sciences, U. S. Public Health Service (No. GM-09280).