Protein Affinities for Small Molecules: Conceptions and Misconceptions

IRVING M. KLOTZ AND DONALD L. HUNSTON

Department of Chemistry, Northwestern University, Evanston, Illinois 60201, and the Chemistry Division, Naval Research Laboratory, Washington, D. C. 20375

Reprinted from ARCHIVES OF BIOCHEMISTRY & BIOPHYSICS, Vol. 193, No. 2, April 1979

ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS
Vol. 193, No. 2, April 1, pp. 314–328, 1979

Protein Affinities for Small Molecules: Conceptions and Misconceptions

IRVING M. KLOTZ AND DONALD L. HUNSTON

Department of Chemistry, Northwestern University, Evanston, Illinois, 60201, and the Chemistry Division, Naval Research Laboratory, Washington, D. C. 20375

Received July 31, 1978; revised October 25, 1978

There is much confusion and error in published treatments of data for multiple binding of ligands (e.g., substrates) by proteins (e.g., enzymes). There is a widespread impression that if the equilibrium binding, r, of ligand, A, by a protein with n sites can be fitted to an equation with two hyperbolic terms, i.e.,

$$r = \frac{n_\alpha k_\alpha(A)}{1 + k_\alpha(A)} + \frac{n_\beta k_\beta(A)}{1 + k_\beta(A)} \quad (n_\alpha + n_\beta = n),$$

then k_α and k_β are always the intrinsic binding constants for two sets of sites. Such a conclusion is often incorrect. For example, in many cases, the protein is constituted of identical protomers with initially identical sites for binding ligands, and yet graphical representations of the binding data appear to behave as if the sites are partitioned between two classes. Although the use of a linear combination of hyperbolic terms to represent binding of ligands by macromolecules always yields empirical parameters k_α, $k_\beta \ldots k_\lambda$, they *cannot* correspond to site binding constants when there are interactions between sites. In some circumstances their values may even be imaginary, complex numbers. On the other hand, stoichiometric binding constants can be assigned unambiguously without making any assumption regarding the nature of the interactions among binding sites. These principles are illustrated concretely by analyses of binding measurements for several different proteins containing two to six sites.

Ligand binding by biological macromolecules plays a vital role in a host of biological functions: enzymic reactions and control mechanisms; receptor interactions with neurotransmitters, hormones, and other effectors; immunoglobulin–antigen interactions; control of gene expression; transport, etc. To understand the dependence of a biological response on the concentration of effector, one must have, as a foundation, information on the extent of binding of the effector. Binding data are most directly represented in a graph of moles of bound ligand per mole of macromolecular acceptor, r, as a function of free ligand concentration, A, the latter being presented on a logarithmic scale if it spans a wide range. Figure 1 illustrates such a plot for data (1) on the binding of carbamyl phosphate by aspartate transcarbamylase, a very widely studied allosteric enzyme. From such a graph one can read explicitly how much enzyme–substrate complex is present at a given concentration of free ligand.

In connection with structural questions about oligomeric systems, it has been tempting to use binding data to try to estimate the total number of accessible sites n on the receptor. In essence this corresponds to finding r at infinite free concentration of A. Even when a substantial number of experimental points has been accumulated, as in Fig. 1, a definitive plateau in r, at high (A) is often not discernible. There is no evidence from these data in themselves that one is approaching saturation of the receptor macromolecule. Nevertheless, if the same data are used for one of the reciprocal transform graphs, for example, $r/(A)$ vs r, one can bend perception to wish and choose an intercept

0003-9861/79/040314-15$02.00/0

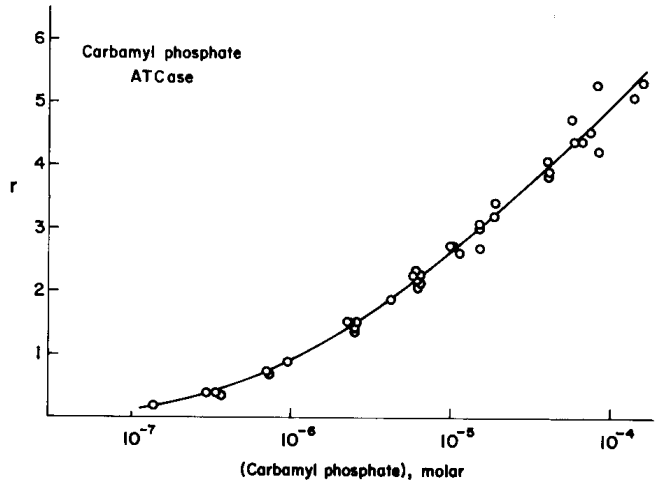

FIG. 1. Binding of carbamyl phosphate by aspartate transcarbamylase (1). Original data kindly supplied by Dr. J. Rosenbusch.

on the r-axis at a value that corresponds with one's predisposition. For example, if the data in Fig. 1, where it is obvious that n is indeterminate, are transformed to the plot of Fig. 2 with the abscissa explicitly numerically labeled, then one can readily be enticed into drawing an intercept at $r = 6$ (the known number of protomers in aspartate transcarbamylase). On the other hand, if the coordinates are kept double blind (Fig. 2A), the indeterminacy of n is manifest. Thus , the r versus log (A) plot provides the best graph for ascertaining the saturation level, if it can be educed from the available data.

A concave curve such as that in Fig. 2 also frequently generates another misconception—the inference that the protein provides two different classes of sites for binding of ligand. With this presumption, binding data are correlated quantitatively by means of an algebraic expression containing two hyperbolic terms, one for each class of independent sites, and the two binding parameters are purported to be the respective site binding constants. For an oligomer constituted of n initially identical protomers, it is unlikely that the n binding sites for a specific ligand are distributed between two different independent classes. Curvature in a graph such as Fig. 2 is commonly a manifestation of interactions between initially identical sites. Under these circumstances, assignments of site binding constants are almost universally incorrect.

These principles will be illustrated concretely by analyses of binding measurements for several different proteins containing two to six sites. For this purpose, we must first state explicitly the general expressions for correlating binding data.

GENERAL FORMS OF EXPRESSION FOR LIGAND BINDING

As has been described recently (2, 3), the multiple equilibria between ligands bound to a macromolecule and free ligand in the bulk solvent can be formulated quantitatively by two different general approaches. One of these, the thermodynamic treatment, defines stoichiometric equilibrium constants, K_i, for the formation of the sequential stoichiometric macromolecule–ligand species PA_1, PA_2, etc., appearing in the equilibria between protein P and ligand A:

$$PA_{i-1} + A = PA_i; \quad K_i = \frac{(PA_i)}{(PA_{i-1})(A)} . \quad [1]$$

The alternative formulation concentrates on individual binding sites on the macromolecule, and defines a site equilibrium constant, k_j, for the equilibrium at each such site ${}_jP$:

$${}_jP + A = {}_jPA; \quad k_j = \frac{({}_jPA)}{({}_jP)(A)} . \quad [2]$$

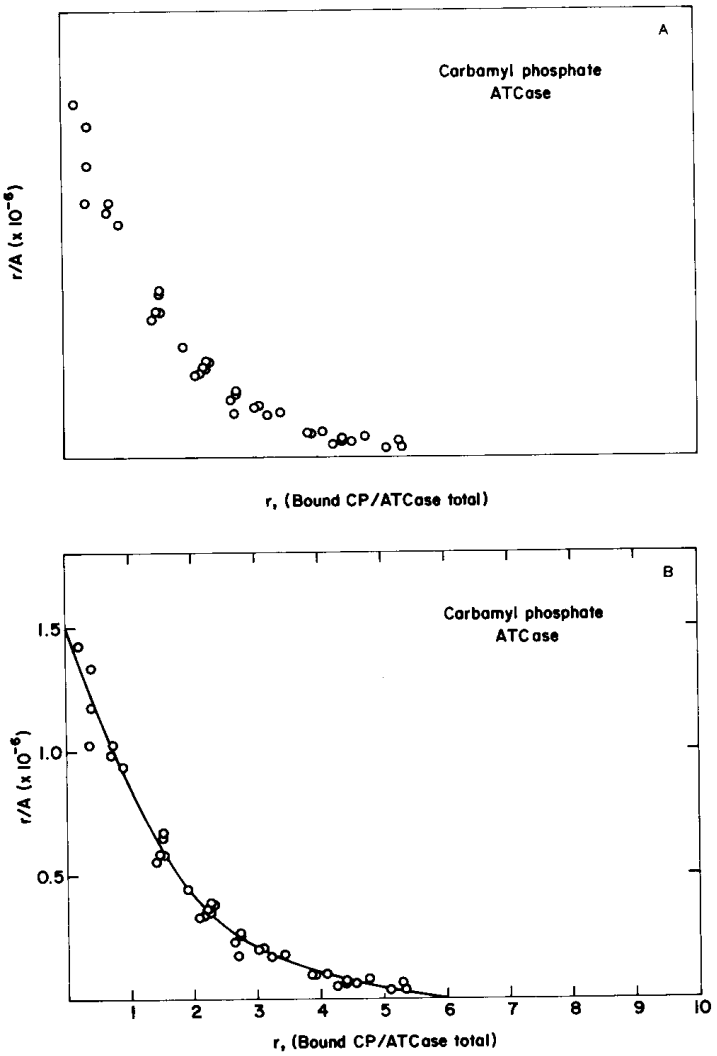

FIG. 2. Single reciprocal plot of data for binding of carbamyl phosphate by aspartate transcarbamylase. In A, numbers have been omitted from both coordinates to illustrate the indeterminacy of any intercept on the abscissa.

An important complication that arises in the site approach, however, is that the affinity of a particular site may be altered if molecules are bound at other sites. In such a situation, more than one site binding constant is needed to describe the behavior at each specific site.

In experimental studies of binding, the quantity normally calculated is r, the moles of bound ligand per mole of total protein. Since the meaning of the various binding constants is often misinterpreted, it is instructive to delineate how r can be formulated in the two different approaches.

In the stoichiometric formulation, the concentration of bound ligand is determined by calculating the concentration of each stoichiometric species (sum of columns in Fig. 3) and multiplying the latter by the coefficient corresponding to its ligand–protein stoichiometry. Following insertion of the appropriate equilibrium constants from Eq. [1], one obtains

$$r = \frac{K_1(A) + 2K_1K_2(A)^2 + \cdots}{1 + K_1(A) + K_1K_2(A)^2 + \cdots}. \quad [3]$$

As illustrated in Fig. 3, the concentration of the stoichiometric species PA_i is the sum of concentrations for all the site species

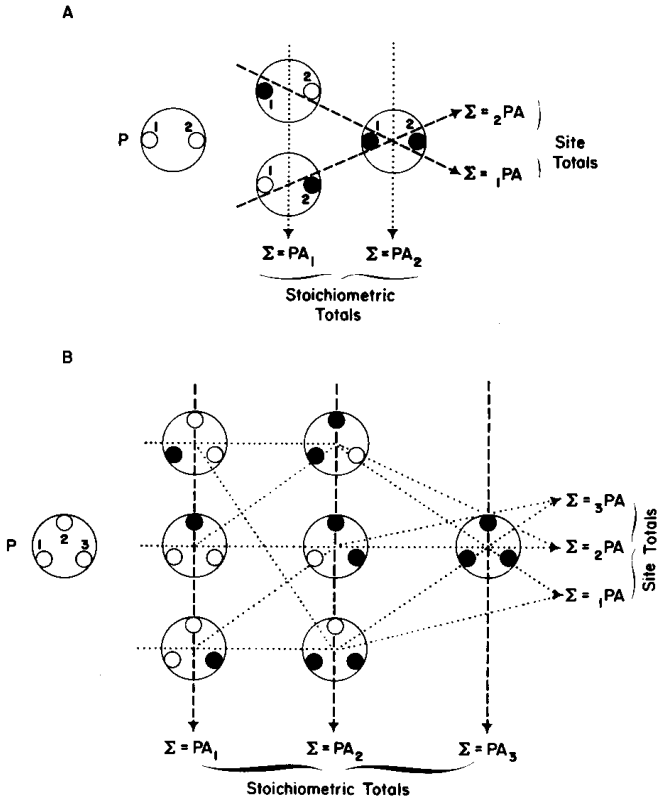

FIG. 3. Comparison of species whose concentrations are summed in stoichiometric and site formulations, respectively. (A) Divalent receptor; (B) trivalent receptor.

with i bound molecules regardless of which specific sites are occupied. Consequently, the values of K_i do not directly provide information about the individual binding affinities of the specific sites, and this is often the information of primary structural importance. Nevertheless, as will be illustrated in this paper, some information about the individual sites can often be obtained by examination of relations between stoichiometric and site constants under different circumstances.

The site approach, in contrast to the stoichiometric one, focuses attention on the individual sites. To calculate r, the concentration of ligand bound at each specific site is first determined by summing over all site species with that particular site occupied (i.e., summing from left to right in Fig. 3). These values are then added for all of the sites to determine the total concentration of bound ligand. In the simple case where the sites have fixed (although different) affinities, i.e., where k_j does not change with the extent of occupancy of other sites by ligand, the well-known expression for r is obtained (4–8):

$$r = \frac{k_1(A)}{1 + k_1(A)} + \frac{k_2(A)}{1 + k_2(A)} + \cdots . \quad [4]$$

Clearly, for this simple situation, the site approach provides direct information about the microscopic aspects of binding.

Unfortunately, Eq. [4] is also widely used for situations in which it is *not* applicable, i.e., when the site affinities *do* change with the extent of occupancy. The site approach can be generalized (3) to give an equation for r that is valid even for interacting sites but the number of site binding constants that must be specified increases dramatically, far beyond the number provided for by Eq. [4] (which is one for each site). Hence, Eq. [4] completely collapses as a framework for describing binding equilibria.

Despite these difficulties, there has been a general tendency to analyze binding data

with an equation containing a linear combination of hyperbolic terms—superficially similar to Eq. [4]—with the number of terms not exceeding the total number of binding sites:

$$r = \frac{k_\alpha(A)}{1 + k_\alpha(A)} + \frac{k_\beta(A)}{1 + k_\beta(A)} + \cdots . \quad [5]$$

Interestingly enough, even though Eq. [4] is valid only for a very special case, Eq. [5] in its most general form (maximum number of terms, with the parameters k_α, k_β, . . . permitted to be complex numbers) contains sufficient parameters to fit the binding data for any binding system, with or without interactions between sites (6–11). Nevertheless, the values obtained for the parameters k_α, k_β, . . . do *not* in general correspond to site binding constants.

DIVALENT (TWO-SITE) SYSTEM

For this situation, the site and stoichiometric equilibrium constants are defined in Chart I. Subscripts to the left of P designate the site(s), subscripts to the right of A represent the moles of bound ligand per mole of protein in that specific complex. The free ligand A in each equilibrium is not shown explicitly in order to simplify the illustration.

For the divalent system, the stoichiometric constants, K_1 and K_2, suffice to completely represent the quantitative binding data in terms of Eq. [3]. Moreover, simple manipulation shows that these constants can be related to the three independent site binding constants

$$K_1 = k_1 + k_2, \quad [6]$$

$$K_2 = \frac{k_1 k_{1,2}}{k_1 + k_2} . \quad [7]$$

It is *not* possible, however, to use the three site-binding constants, k_1, k_2, $k_{1,2}$ (or $k_{2,1}$) with Eq. [4] to represent the binding data; this can readily be seen from the fact that such an equation requires that $r \to 3$ as $(A) \to \infty$, but r can never exceed the stoichiometric valency, in this case 2.

One can still fit binding data to Eq. [5] with two terms, and two parameters k_α and k_β. With some moderate algebraic manipulation (comparison of polynomial forms of Eq. [3], for a divalent system, and of Eq. [5] with two terms), one can obtain relations for k_α and k_β in terms of the stoichiometric binding constants:

$$k_\alpha = \tfrac{1}{2}K_1 \pm \tfrac{1}{2}(K_1^2 - 4K_1K_2)^{1/2}, \quad [8]$$

$$k_\beta = \tfrac{1}{2}K_1 \mp \tfrac{1}{2}(K_1^2 - 4K_1K_2)^{1/2}. \quad [9]$$

It should be noted immediately that when $K_2 > \tfrac{1}{4}K_1$, which would be true if affinities increase with extent of occupancy (2), k_α and k_β have complex values.

It is also possible to derive expressions for k_α and k_β in terms of the site binding constants:

$$k_\alpha = \tfrac{1}{2}(k_1 + k_2) \pm \tfrac{1}{2}[(k_1 + k_2)^2 - 4k_1k_{1,2}]^{1/2}, \quad [10]$$

$$k_\beta = \tfrac{1}{2}(k_1 + k_2) \mp \tfrac{1}{2}[(k_1 + k_2)^2 - 4k_1k_{1,2}]^{1/2}. \quad [11]$$

These provide explicit relationships between the two different types of constants. Nevertheless, k_α is *not* identifiable with any specific site constant, nor is k_β.

One special situation of interest is that in which both sites are *initially identical* in affinities,

$$k_1 = k_2 = k_I, \quad [12]$$

and the second *stage* in occupancy is also characterized by identical site affinities, but different from the first stage, i.e.,

$$k_{1,2} = k_{2,1} = k_{II}. \quad [13]$$

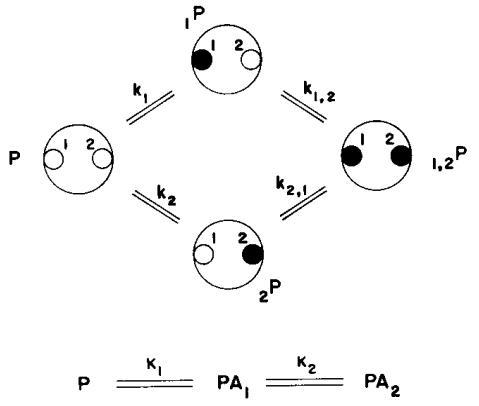

CHART I

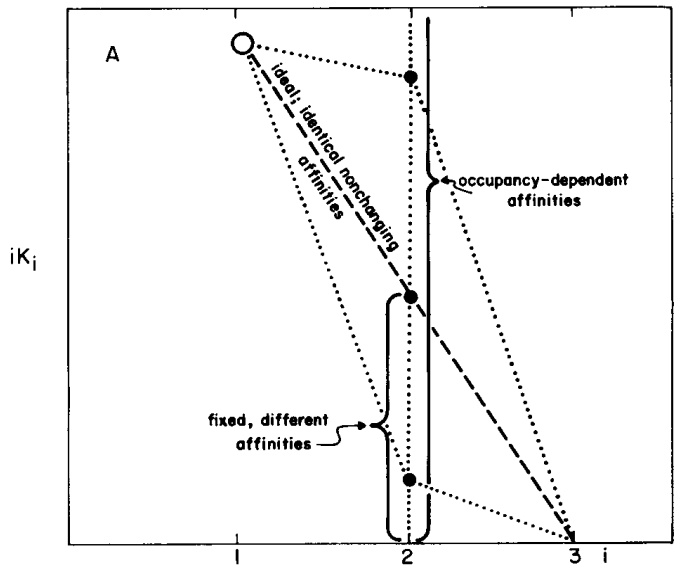

FIG. 4. Affinity profiles for divalent system.

Under these circumstances, one can show readily (from Eqs. [6] and [7]) that the stoichiometric binding constants will automatically lead to the sequential stage affinities k_{I} and k_{II} that characterize the sequential behavior of the sites:

$$K_1 = 2k_{\mathrm{I}}, \qquad [14]$$

$$K_2 = \tfrac{1}{2}k_{\mathrm{II}}. \qquad [15]$$

Here again, although k_α and k_β would have defined values,

$$k_\alpha = k_{\mathrm{I}} + (k_{\mathrm{I}}^2 - k_{\mathrm{I}}k_{\mathrm{II}})^{1/2} \qquad [16]$$

$$k_\beta = k_{\mathrm{I}} - (k_{\mathrm{I}}^2 - k_{\mathrm{I}}k_{\mathrm{II}})^{1/2}, \qquad [17]$$

k_α *cannot* be assigned to k_{I} and k_β *cannot* be assigned to k_{II} (nor vice versa).

Affinity profiles (2, 3), graphs of normalized stoichiometric binding constants in the form of iK_i versus i, for a divalent system reflect the various types of possible relationships among binding sites. For the ideal case of identical, noninteracting sites (Fig. 4), the successive K_i values are linearly related in a graph of iK_i versus i (2, 3). For the divalent system in general, the second stoichiometric parameter, $2K_2$, may be above or below the ideal line. If above, the binding affinity is accentuated with increasing extent of occupancy. If below three different causes are possible: (i) the two specific sites have different but fixed affinities; (ii) site affinities are initially identical but decrease as soon as one site is occupied; (iii) the two specific sites initially have different affinities but these decrease (or they may even increase under some circumstances[1]) as soon as one site is occupied. All three possibilities lead to the same form of affinity profile. Hence, in a two-site system, it is not possible to distinguish "negative cooperativity" from

[1] These three cases can be scrutinized quantitatively by noting that for a two-site system only one interaction parameter, I, is needed, since

$$k_{1,2}/k_2 = k_{2,1}/k_1 = I + 1.$$

If the sites are indexed so that $k_2 \geq k_1$, then α can be defined as

$$\alpha = (k_2/k_1) - 1 \geq 0.$$

Simple manipulation of these relations and Eqs. [6] and [7] shows that the second point in an affinity profile falls above, on, or below the ideal line when the quantity

$$\mathrm{I} - \frac{\alpha^2}{4(1+\alpha)}$$

is positive, zero, or negative, respectively. Since the second term in this quantity is always zero or positive, the second point in an affinity profile will fall below the ideal line when

$$\frac{\alpha^2}{4(1-\alpha)} > \mathrm{I}.$$

fixed but different site affinities with no cooperativity whatsoever. Furthermore, it is even possible to have "positive cooperativity" between initially nonidentical sites and obtain an affinity profile, or a Scatchard plot (11), below the linear one for the ideal case.

An example of actual behavior in a divalent system is illustrated by leucine–isopropylmalate synthase (12). The stoichiometric constants K_1 and K_2, calculated by the computational procedures of Fletcher *et al.* (9, 11) from binding data supplied by Teng-Leary and Kohlhaw (12), are

$$K_1 = 4.8 \times 10^4,$$

$$K_2 = 2.5 \times 10^5.$$

It is immediately evident from an affinity profile (Fig. 5A) that after either of the first two sites is occupied, the affinity in the second sequential step must be markedly accentuated.

If one analyzes the binding data in terms of a linear combination of hyperbolic terms, Eq. [5], then the best values of the parameters k_α and k_β for binding of leucine by isopropylmalate synthase are

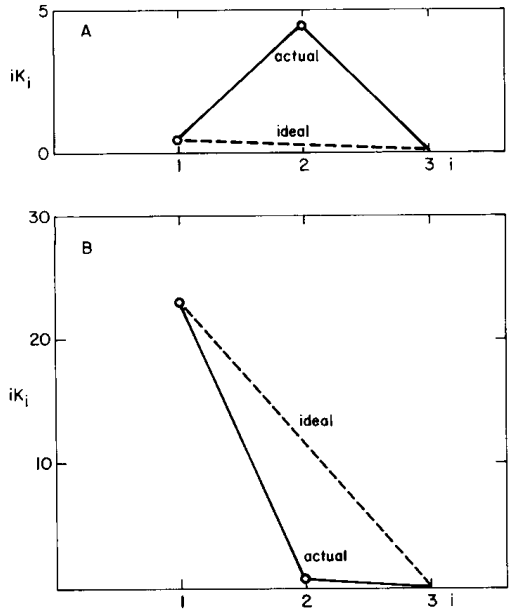

FIG. 5. Affinity profiles for binding of leucine by isopropylmalate synthase (A) and for binding of iron by ovotransferrin (B).

$$k_\alpha = 2.4 \times 10^4 \pm 1.1 \times 10^5\sqrt{-1},$$

$$k_\beta = 2.4 \times 10^4 \mp 1.1 \times 10^5\sqrt{-1}.$$

Thus, k_α and k_β are complex numbers, members of a conjugate pair. Their imaginary values clearly demonstrate that neither k_α nor k_β can be identified with any one of the site binding constants of isopropylmalate synthase.

To assign values to the site binding constants we need additional information. In view of what is known of the structure of this particular enzyme, it can be reasonably assumed that both sites in the original nonliganded protein are identical. Under these circumstances Eqs. [12] and [13] define the binding behavior, and Eqs. [14] and [15] can be used to evaluate the sequential stage affinities:

$$k_{\mathrm{I}} = 0.24 \times 10^5,$$

$$k_{\mathrm{II}} = 5.0 \times 10^5.$$

The first site binding constant then has the value

$$k_1 = 0.24 \times 10^5$$

(see Eq. [12]); however, k_2 is *not* equal to 5.0×10^5. After site 1 is occupied (see Eq. [13])

$$k_{1,2} = 5.0 \times 10^5.$$

Correspondingly,

$$k_2 = 0.24 \times 10^5$$

$$k_{2,1} = 5.0 \times 10^5$$

(see Eqs. [12] and [13]).

A complementary example is illustrated in Fig. 5B, an affinity profile for the binding of ferric ion by divalent ovotransferrin (13). Here K_1 is 23 and K_2 is 0.57. Thus, after either of the first two sites is occupied, the affinity of the residual site, whichever it is, is markedly decreased. Furthermore, for this iron–ovotransferrin system $k_\alpha = 22.4$ and $k_\beta = 0.6$; neither of these values is equal to that of any of the site binding constants, whose values are (13), $k_1 = 10$, $k_2 = 13$, $k_{1,2} = 1.3$, $k_{2,1} = 1$. Ovotransferrin binding of iron falls into the class often called "half-of-sites" reactivity, i.e., the first iron is taken up by either of the two (almost equivalent) unoccupied sites

and then the affinity of the residual site is markedly reduced, regardless of which one it was in the totally unoccupied protein. It should be emphasized, therefore, that if k_α and k_β are used, with Eq. [5], to correlate the binding data in other half-of-sites systems, the values obtained for these parameters are *not* the site binding constants at any stage of uptake of ligand.

TRIVALENT (THREE-SITE) SYSTEM

The appropriate site and stoichiometric constants for this situation in its most general form are defined in Chart II.

For the trivalent system, the three stoichiometric constants can be related (3) to (seven) independent site binding constants by the equations

$$K_1 = k_1 + k_2 + k_3, \qquad [18]$$

$$K_2 = \frac{k_1k_{1,2} + k_1k_{1,3} + k_2k_{2,3}}{k_1 + k_2 + k_3}, \qquad [19]$$

$$K_3 = \frac{k_1k_{1,2}k_{1,2,3}}{k_1k_{1,2} + k_1k_{1,3} + k_2k_{2,3}}. \qquad [20]$$

The variation of r with A can be represented without any ambiguity by the stoichiometric binding equation (3) with terms up to and including the cubic in both numerator and denominator. A computer fit can then be used with any specific set of experimental binding data to obtain best values of the stoichiometric constants K_1, K_2, and K_3. With known values of the stoichiometric constants, however, there is no way to specify the site binding constants since the seven independent latter parameters cannot be fixed by three known experimental constants. Furthermore, a seven-term equation in the format of Eq. [4] cannot be used for it would imply that $r \to 7$ as $A \to \infty$, whereas actually $r \to 3$.

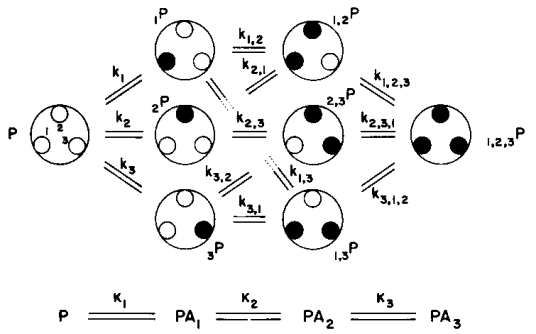

CHART II

For special cases, of general interest, site binding affinities can be assigned once the stoichiometric constants have been evaluated. For example, if a trimeric protein is constituted of identical subunits, all three subunits should have identical initial affinities for (substrate) ligand, i.e.,

$$k_1 = k_2 = k_3 \equiv k_{\mathrm{I}}. \qquad [21]$$

One possible mode of interaction between sites could result in a changed affinity for the second sequential stage in uptake of ligand converting PA_1 to PA_2:

$$k_{1,2} = k_{1,3} = k_{2,1} = k_{2,3} = k_{3,1} = k_{3,2} \equiv k_{\mathrm{II}}. \qquad [22]$$

Let us assume that uptake of a third ligand occurs with the same affinity as in the second stoichiometric step, so that

$$k_{1,2,3} = k_{2,3,1} = k_{3,1,2} \equiv k_{\mathrm{II}}. \qquad [23]$$

Under these circumstances it follows from Eqs. [18]–[20] that

$$K_1 = 3k_{\mathrm{I}}, \qquad [24]$$

$$K_2 = k_{\mathrm{II}}, \qquad [25]$$

$$K_3 = \tfrac{1}{3}k_{\mathrm{II}}. \qquad [26]$$

The affinity profile for ligand binding in this situation is illustrated in Fig. 6. Two of many possible profiles are shown by the solid lines, one in which the affinity in the second stoichiometric step is accentuated, the other in which it is attenuated. In both cases, the line connecting $i = 2$, 3, and 4 is linear, that is, the corresponding stoichiometric constants are related "ideally" (i.e., by the statistics of the situation only), as is required by the specifications in Eqs. [22] and [23].

The stoichiometric binding equation (3) for the circumstances described by Eqs. [24]–[26] is

$$r = \frac{3k_{\mathrm{I}}(A) + 6k_{\mathrm{I}}k_{\mathrm{II}}(A)^2 + 3k_{\mathrm{I}}k_{\mathrm{II}}^2(A)^3}{1 + 3k_{\mathrm{I}}(A) + 3k_{\mathrm{I}}k_{\mathrm{II}}(A)^2 + k_{\mathrm{I}}k_{\mathrm{II}}^2(A)^3}. \qquad [27]$$

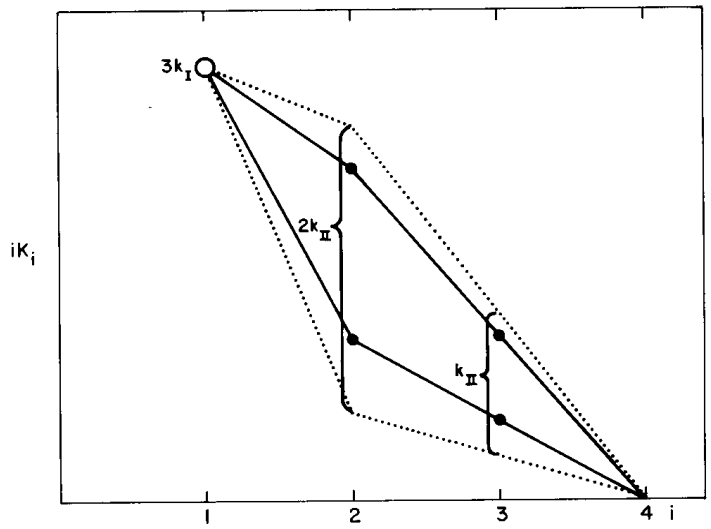

FIG. 6. Affinity profiles for trivalent system.

With two sequential classes of sites such as those specificed by Eqs. [21] to [23], one might be inclined to correlate the binding data and evaluate the class binding constants by a linear combination of hyperbolic terms, i.e., Eq. [5]. Ostensibly one could write

$$r = \frac{k_\alpha A}{1 + k_\alpha A} + \frac{2k_\beta A}{1 + k_\beta A} \qquad [28]$$

and fit this equation to the experimental data. The coefficient, 2, is used in the second term to indicate two equal-affinity classes. Superficially one might think that k_α should be assigned to k_I and k_β to k_{II}. This is absolutely incorrect, however. Relations between the k's of Eq. [28] and the sequential class constants can be derived by rearranging Eqs. [27] and [28] into their respective linear polynomial forms:

$$\text{Stoichiometric:} \quad r + 3(r-1)k_I A + 3(r-2)k_I k_{II} A^2 + (r-3)k_I k_{II}^2 A^3 = 0, \qquad [29]$$

$$\text{Linear hyperbolic:} \quad r + [(r-1)k_\alpha + (r-2)k_\beta]A + (r-3)k_\alpha k_\beta A^2 = 0. \qquad [30]$$

It is immediately apparent why Eq. [30], and hence [28], fails. First, it lacks the cubic term in A, and thus (in view of Eq. [29]), requires that k_I or k_{II} always be zero. Furthermore, even if we restrict ourselves to a comparison of the linear and square terms in Eqs. [29] and [30], we find that k_α and k_β depend on r as well as k_I and k_{II}:

$$k_\alpha = \frac{3}{2}k_I \pm \left[\frac{9}{4}k_I^2 - 3\frac{(r-2)^2}{(r-1)(r-3)}k_I k_{II}\right]^{1/2}, \qquad [31]$$

$$k_\beta = 3\frac{(r-1)}{(r-2)}k_I - \frac{(r-1)}{(r-2)}\left\{\frac{3}{2}k_I \pm \left[\frac{9}{4}k_I^2 - 3\frac{(r-2)^2}{(r-1)(r-3)}k_I k_{II}\right]^{1/2}\right\}. \qquad [32]$$

In other words, k_α and k_β are not unique; their values depend on the best-fit criterion used in the curve-fitting procedure. No matter what criterion is used, moreover, *their values certainly cannot be identified with particular site binding constants.*

An example of actual behavior in a tervalent system is available from the studies of Ridge *et al.* (14) of the binding of trifluorodihydroxypropyl phosphonate (an analog of carbamyl phosphate) by asparate transcarbamylase. From their published data we have computed the following stoichiometric constants:

$$K_1 = 2.89 \times 10^4,$$

$$K_2 = 0.0295 \times 10^4,$$

$$K_3 = 0.0875 \times 10^4.$$

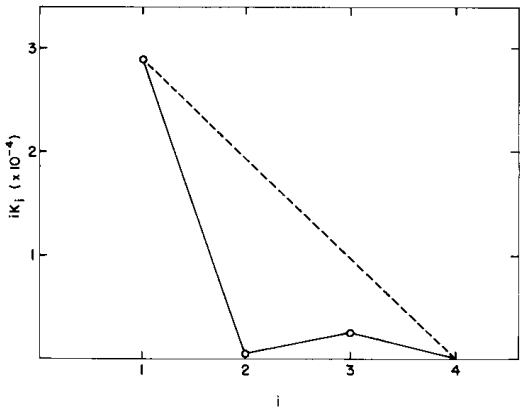

FIG. 7. Affinity profile for binding of trifluorodihydroxypropyl phosphonate by aspartate transcarbamylase.

The corresponding affinity profile is illustrated in Fig. 7. It is immediately evident that after one site (of the three initially identical ones) is occupied, the remaining unoccupied sites drop markedly in affinity.[2] In contrast, a small rise in affinity occurs in the occupation by ligand of the last open site.

In the initially ligand-free catalytic trimer, all three monomers are presumably identical. In view of Eqs. [18] and [21], we may specify, therefore, their initial site binding constants, all being identical:

$$k_1 = k_2 = k_3 = k_I = \tfrac{1}{3}K_1 = 0.96 \times 10^4.$$

No matter which one is occupied in the first stoichiometric step, each of the residual open sites then has a changed binding constant. If Eq. [22] is applicable, then from Eq. [25] we have

$$k_{II} = 0.0295 \times 10^4.$$

Finally, if the last open binding site, whether 1, 2, or 3 originally, has the same (further changed) affinity, k_{III}, for ligand, then from Eqs. [18]–[20] it follows that

$$k_{III} = 3K_3 = 0.262 \times 10^4.$$

It should be emphasized, however, that the values 0.96×10^4, 0.0295×10^4, and 0.262×10^4 are *not* site binding constants for sites 1, 2, and 3, respectively. No single value of a site-binding constant can be assigned to site 1 (or 2 or 3). The affinity at each site depends on whether the other sites are open or occupied.

Ridge *et al.* (14) have correlated their binding data by means of a two-term hyperbolic equation (5). From their parameters we calculate

$$k_\alpha = 2.0 \times 10^4,$$

$$k_\beta = 0.057 \times 10^4.$$

Despite their rough similarity in magnitude[3] to K_i or to $k_I - k_{III}$, k_α and k_β should *not* be identified with any stoichiometric or site binding constant.

TETRAVALENT (FOUR-SITE) SYSTEM

The appropriate stoichiometric and site binding constants for this system are defined in Chart III.

For a tetravalent system the four stoichiometric constants can be related (3) to (fifteen) independent site binding constants by the equations

$$K_1 = k_1 + k_2 + k_3 + k_4, \qquad [33]$$

$$K_2 = \frac{k_1k_{1,2} + k_1k_{1,3} + k_1k_{1,4} + k_2k_{2,3} + k_2k_{2,4} + k_3k_{3,4}}{k_1 + k_2 + k_3 + k_4}, \qquad [34]$$

$$K_3 = \frac{k_1k_{1,2}k_{1,2,3} + k_1k_{1,2}k_{1,2,4} + k_1k_{1,3}k_{1,3,4} + k_2k_{2,3}k_{2,3,4}}{k_1k_{1,2} + k_1k_{1,3} + k_1k_{1,4} + k_2k_{2,3} + k_2k_{2,4} + k_3k_{3,4}}, \qquad [35]$$

$$K_4 = \frac{k_1k_{1,2}k_{1,2,3}k_{1,2,3,4}}{k_1k_{1,2}k_{1,2,3} + k_1k_{1,2}k_{1,2,4} + k_1k_{1,3}k_{1,3,4} + k_2k_{2,3}k_{2,3,4}}. \qquad [36]$$

[2] This behavior could be called "third-of-sites reactivity," in analogy to "half-of-sites reactivity."

[3] In the formulation $r = \sum\{k_\lambda(A)/[1 + k_\lambda(A)]\}$ each term makes a contribution to r of between 0 and 1.

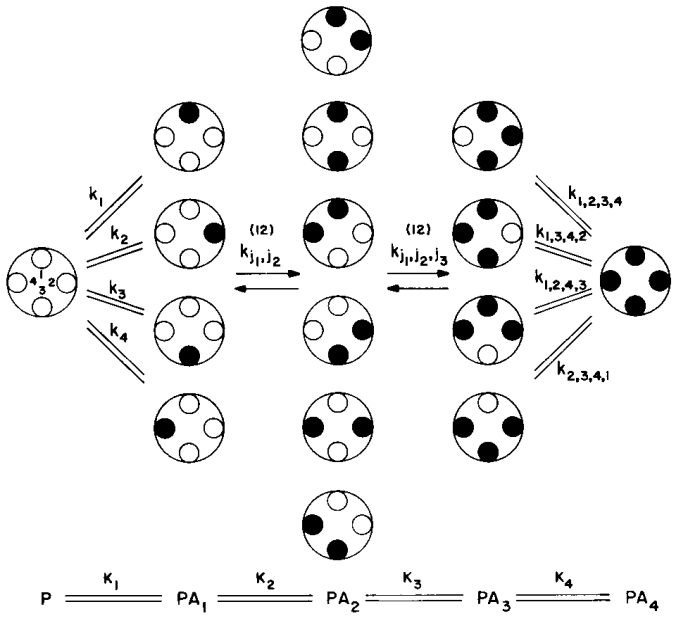

CHART III

The variation of r with A can be represented without any ambiguity by the stoichiometric binding equation (3) with terms up to and including $(A)^4$ in both numerator and denominator, as was shown by Adair (15) for tetravalent hemoglobin.

A special case of interest for a tetravalent system is one in which the site binding constants fall into the following categories:

$$k_1 = k_2 = k_3 = k_4 = k_{1,2} = k_{1,3} = k_{1,4} = k_{2,3} = k_{2,4} = k_{3,4} = k_{\mathrm{I}}, \quad [37]$$

$$k_{1,2,3} = k_{1,2,4} = k_{1,3,4} = k_{2,3,4} = k_{1,2,3,4} = k_{\mathrm{II}}. \quad [38]$$

This would be applicable, for example, to a tetrameric protein with four identical subunits, arranged perhaps in a tetrahedral array, in which the first two sequential stoichiometric steps involved binding by sites of identical affinity, k_{I}, and in which the third and fourth sequential steps reflected identical site affinities of k_{II}, where $k_{\mathrm{II}} \neq k_{\mathrm{I}}$. Under these circumstances one can show readily from Eqs. [33]–[36] that

$$K_1 = 4k_{\mathrm{I}}, \quad [39]$$

$$K_2 = \tfrac{3}{2}k_{\mathrm{I}}, \quad [40]$$

$$K_3 = \tfrac{2}{3}k_{\mathrm{II}}, \quad [41]$$

$$K_4 = \tfrac{1}{4}k_{\mathrm{II}}. \quad [42]$$

The affinity profile for such a situation is illustrated in Fig. 8. Two of many possible specific profiles are shown by the solid lines, one representing relationships when the affinity is hyperideal, $(k_{\mathrm{II}} > k_{\mathrm{I}})$, the other when it is hypoideal. For each, the line connecting the two stoichiometric steps 1 and 2, or that connecting 3 and 4, intersects the abscissa at $n + 1$, and corresponds to ideal behavior for the corresponding stoichiometric constants.

If one were to correlate the binding data for a system in which equations [37] and

The concentration of (A) at which a particular term contributes 0.5 is $(A_{0.5})_\lambda = k_\lambda^{-1}$. If K_1 is significantly larger than the subsequent K_i's, $(A_{0.5})_\alpha$ can be near the midpoint in the uptake of the first ligand. This can result in a k_α that is similar in value to K_1. A similar situation can also exist for other constants. The magnitude and spacing of the K_i's, however, depend on all the site affinities and any interactions that are present. Thus, any superficial similarity in values that may be found is a result of the mathematical form of the equations and is no justification for identifying k_λ with a particular site affinity.

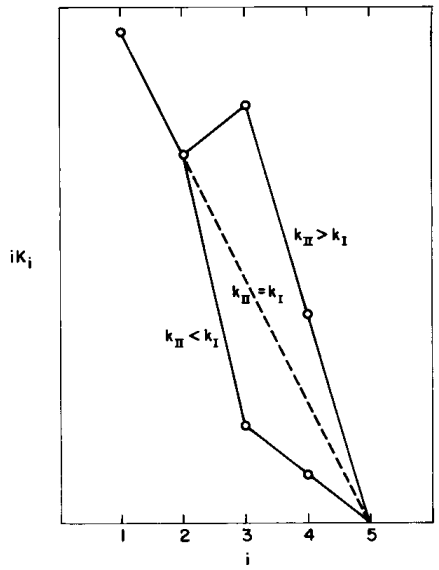

FIG. 8. Affinity profiles for a tetrameric system when initially identical site affinities change after half-of-site sites are occupied.

[38] are applicable by a linear combination of hyperbolic terms, (Eq. [5]), the following two-term equation would be deemed appropriate:

$$r = \frac{2k_\alpha A}{1 + k_\alpha A} + \frac{2k_\beta A}{1 + k_\beta A}. \quad [43]$$

If this relation is rearranged to a linear polynomial form it becomes

$$r + (r - 2)(k_\alpha + k_\beta)A + (r - 4)k_\alpha k_\beta A^2 = 0. \quad [44]$$

The stoichiometric binding equation for the circumstances at hand can be obtained readily from Eq. [3] combined with the constraints specified by Eqs. [39]–[42]:

$$r = \frac{4k_I A + 12k_I^2 A^2 + 12k_I^2 k_{II} A^3 + 4k_I^2 k_{II}^2 A^4}{1 + 4k_I A + 6k_I^2 A^2 + 4k_I^2 k_{II} A^3 + k_I^2 k_{II}^2 A^4}. \quad [45]$$

In a linear polynomial format this becomes

$$r + 4(r - 1)k_I A + 6(r - 2)k_I^2 A^2 + 4(r - 3)k_I^2 k_{II} A^3 + (r - 4)k_I^2 k_{II}^2 A^4 = 0. \quad [46]$$

Comparison of Eqs. [44] and [46] shows that the former is inadequate since it lacks the A^3 and A^4 terms that must be present, and hence, it requires that k_I or k_{II} always be zero. Again, if we restrict ourselves to a comparison of the coefficients of the linear and square terms in Eqs. [44] and [46] we find that k_α and k_β depend on r as well as on k_I and k_{II}. Clearly k_α and k_β do not even have unique values and certainly cannot be identified with particular site binding constants.

Another special case of interest for a tetravalent ligand–binding protein is that in which the initially identical subunits interact in a pairwise fashion, that is, when one is occupied its unoccupied pair-partner changes in affinity but the unoccupied pair is unaffected. To analyze this situation we index the pairs as 1 and 3, and 2 and 4 (see Chart III). Initially all sites have identical site binding constants, which we shall designate k_I. For each site, however, affinity changes to k_{II} when the partner is occupied. Specifically, then, we can assign the following values to each site binding constant:

$$\begin{aligned}
k_1 &= k_2 = k_3 = k_4 = k_I, \\
k_{1,2} &= k_{1,4} = k_{2,1} = k_{2,3} = k_{3,2} = k_{3,4} = k_{4,1} = k_{4,3} = k_I, \\
k_{1,3,2} &= k_{1,3,4} = k_{2,4,1} = k_{2,4,3} = k_I, \\
k_{1,3} &= k_{2,4} = k_{3,1} = k_{4,2} = k_{II}, \\
k_{1,2,3} &= k_{1,2,4} = k_{1,4,2} = k_{1,4,3} = k_{2,3,1} = k_{2,3,4} = k_{3,4,1} = k_{3,4,2} = k_{II}, \\
k_{1,2,3,4} &= k_{1,3,4,2} = k_{1,2,4,3} = k_{2,3,4,1} = k_{II}.
\end{aligned} \quad [47]$$

If we define $k_{j_1,j_2,\cdots j_l}$ as the site binding constant for site j_l when sites $j_1, j_2, \cdots j_{l-1}$ are occupied then Eqs. [47] can be represented more concisely as follows:

$$\begin{aligned}
k_{j_1,j_2\cdots j_l} &= k_{II} \quad \text{if} \quad |j_l - j_m| = 2 \quad \text{for} \quad m \text{ between 1 and } (l-1), \\
k_{j_1,j_2\cdots j_l} &= k_I \quad \text{if} \quad |j_l - j_m| \neq 2.
\end{aligned} \quad [48]$$

With these specifications, the stoichiometric constants (Eqs. [33]–[36]) for a tetramer with pairwise interactions become

$$K_1 = 4k_{\mathrm{I}},$$

$$K_2 = k_{\mathrm{I}} + \tfrac{1}{2}k_{\mathrm{II}},$$

$$K_3 = \frac{k_{\mathrm{I}}k_{\mathrm{II}}}{k_{\mathrm{I}} + \tfrac{1}{2}k_{\mathrm{II}}}, \qquad [49]$$

$$K_4 = \frac{k_{\mathrm{II}}}{4}.$$

Some affinity profiles for different relative values of $k_{\mathrm{II}}/k_{\mathrm{I}}$ are illustrated in Fig. 9. Pairwise interactions are clearly distinguishable from the situation in which affinity changes after any two of the four mutually unoccupied sites are filled by ligand (Fig. 8).

An example of actual behavior of a tetrameric protein is illustrated by the binding of acetylcoenzyme A to pyruvate carboxylase. From experimental data supplied by Frey and Utter (16), we have computed the following stoichiometric constants:

$$K_1 = 2.08 \times 10^5,$$

$$K_2 = 1.64 \times 10^4,$$

$$K_3 = 3.67 \times 10^5,$$

$$K_4 = 1.30 \times 10^4.$$

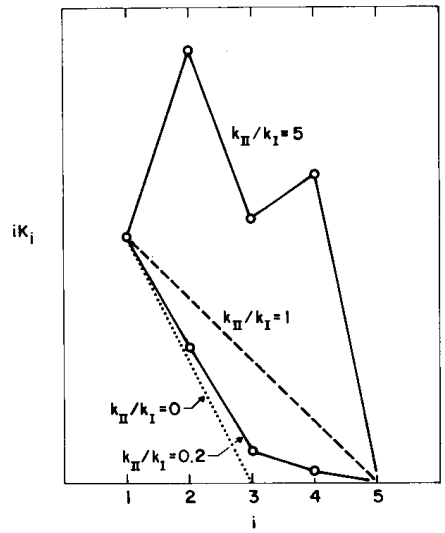

FIG. 9. Affinity profiles for a tetrameric system in which initially identical sites interact in a pairwise fashion.

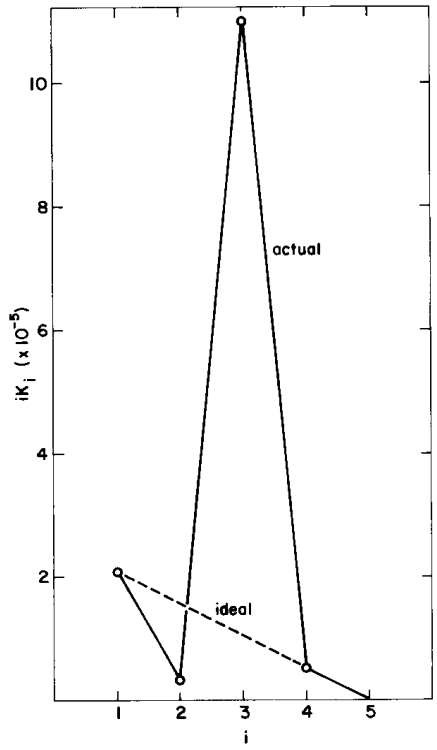

FIG. 10. Affinity profile for binding of acetylcoenzyme A by pyruvate carboxylase.

The corresponding affinity profile is shown in Fig. 10. Most prominent is the very steep rise in affinity in the third stoichiometric stage. To some extent this behavior is similar to that in Fig. 8 for $k_{\mathrm{II}} > k_{\mathrm{I}}$. Thus, pyruvate carboxylase markedly increases its affinity for acetylcoenzyme A after any two of the original open sites are occupied. However, the actual curve, Fig. 10, does not otherwise fit the model situation of Fig. 8, but rather reveals some weaker interactions in steps 2 and 4 of the stoichiometric stages.

CONCLUSION

For pentavalent or hexavalent systems, representations analogous to Charts I through III became much too crowded to illustrate clearly interrelationships between species and equilibrium constants. A more concise chart, such as that shown in Chart IV, does enable one to visualize the distribution of the 192 site binding constants among the successive stoichiometric steps. For a hexavalent system the six stoichiometric

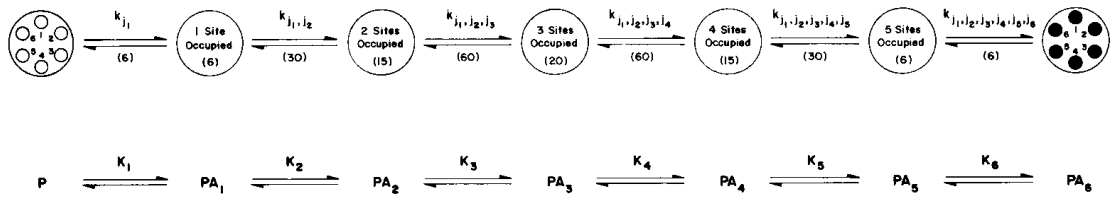

CHART IV

constants can be related to the (sixty-three) independent site constants (3), but the general equations obtained are much too cumbersome and the site constants are indeterminate. When interactions between sites are not all-pervasive, the general equations can be reduced to simpler forms. Several special cases have been examined recently (17) for a hexavalent system, and the actual behavior of aspartate transcarbamylase has been compared with these limiting boundary models.

Binding valencies above six are too unwieldy to describe with specific details, unless special interactions occur that simplify the algebra considerably in a manner similar to that illustrated in the preceding discussions. Nevertheless, completely general equations in concise format have been developed (3) and these make it possible to extract particular relationships for a macromolecule with any number of binding sites.

In any event, the central contribution of binding measurements to the interpretation of the influence of an effector on a biological response is to provide information on the extent of occupancy of receptor sites by ligands. Such information is most explicitly displayed by a graph of moles bound as a function of the free concentration of effector on a logarithmic scale (e.g., Fig. 1). Such a graph provides the essential molecular information for correlation with macromolecular, cellular, or physiological behavior.

If one wants an analytical expression to complement the graphical representation of extent of binding versus concentration, then a number of algebraic options are available. Two of these which have possible molecular connotations have been elaborated upon in this discussion. However, if the parameters obtained are to be associated with binding constants under all circumstances, then *only one* formulation, the stoichiometric one (Eq. [3]), is universally applicable.

The stoichiometric binding constants, K_i, reflect the nature of interactions between sites with increasing occupancy by ligand. These can readily be visualized in affinity profiles, as illustrated for several examples in this paper. From a molecular point of view, an important insight revealed from these examples is that some modification in site affinities occurs at *each* stoichiometric step. Contrary to common practice, there is no basis for interpreting curvature in reciprocal transform graphs of binding data as evidence of two classes of sites with fixed binding constants. The same data lead to affinity profiles that clearly reveal much more complicated behavior. These thermodynamic constructs, therefore, provide a sieve through which any postulated molecular model must be able to pass.

ACKNOWLEDGMENTS

We are grateful to Dr. J. E. Fletcher for his assistance in setting up the computerized procedures for the computation of ligand binding constants and to Dr. E. E. Wells for his help in the use of a Grafacon digitizer interfaced to a TI960A minicomputer. We are also indebted to Drs. G. B. Kohlhaw, J. P. Rosenbusch, W. H. Frey, and M. F. Utter for their kindness in supplying us with original binding data. This investigation was supported in part by a grant (No. GM-09280) from the National Institute of General Medical Sciences.

REFERENCES

1. SUTER, P., AND ROSENBUSCH, J. (1976) *J. Biol. Chem.* **251**, 5986–5991.
2. KLOTZ, I. M. (1974) *Accts. Chem. Res.* **7**, 162–168.
3. KLOTZ, I. M., AND HUNSTON, D. L. (1975) *J. Biol. Chem.* **250**, 3001–3009.
4. ADAMS, E. Q. (1916) *J. Amer. Chem. Soc.* **38**, 1503–1510.

5. SIMMS, H. S. (1926) *J. Amer. Chem. Soc.* **48**, 1239–1261.
6. VON MURALT, A. L., (1930) *J. Amer. Chem. Soc.* **52**, 3518–3523.
7. WYMAN, J. (1948) *Advan. Protein Chem.* **4**, 407–531.
8. SCATCHARD, G. (1949) *Ann. N. Y. Acad. Sci.* **51**, 660–672.
9. FLETCHER, J. E., SPECTOR, A. A., AND ASHBROOK, J. D. (1970) *Biochemistry* **9**, 4580–4587.
10. WYMAN, J., AND PHILLIPSON, P. E. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 3431–3434.
11. FLETCHER, J. E. (1977) *J. Phys. Chem.* **81**, 2374–2378.
12. TENG-LEARY, E., AND KOHLHAW, G. B. (1973) *Biochemistry* **12**, 2980–2986.
13. DONOVAN, J. W., BEARDSLEE, R. A., AND ROSS, K. D. (1976) *Biochem. J.* **153**, 631–639.
14. RIDGE, J. A., ROBERTS, M. F., SCHAFFER, M. H., AND STARK, G. R. (1976) *J. Biol. Chem.* **251**, 5966–5975.
15. ADAIR, G. S. (1925) *J. Biol. Chem.* **63**, 529–545.
16. FREY, W. H., AND UTTER, M. F. (1977) *J. Biol. Chem.* **252**, 51–56.
17. KLOTZ, I. M., AND HUNSTON, D. L. (1977) *Proc. Nat. Acad. Sci. USA* **74**, 4959–4963.