The Influence of Binding Domains on the Nature of Subunit Interactions in Oligomeric Proteins

APPLICATION TO UNUSUAL KINETIC AND BINDING PATTERNS

(Received for publication, June 3, 1970)

Athel Cornish-Bowden and D. E. Koshland, Jr.

From the Department of Biochemistry, University of California, Berkeley, California 94720

SUMMARY

The fact that individual protein subunits must have different regions in contact with neighboring subunits in any oligomeric assembly greater than a dimer is considered in relation to the nature of subunit interactions. Binding equations for this model have been derived for various different tetrameric arrangements and compared with the equations given by other models. Methods of distinguishing between models on the basis of binding curves alone are discussed, including the relationships which each model requires between the association constants for successive molecules of ligand. It is found that certain features of a binding curve, such as its symmetry or asymmetry, and the degree of cooperativity between sites, can be used to discriminate between models. Data in the literature which appear to contradict simple models for ligand binding have been re-examined and found to be consistent with the general ligand-induced model for subunit interactions.

Although many enzyme velocity curves follow Michaelis-Menten kinetics, it is clear that significant deviations from the expected hyperbolic curves are observed in many cases. The finding of Bohr (1), that hemoglobin gave a sigmoid curve when fraction saturation (\( f \)) was plotted against ligand concentration (\( X \)), was followed by the discovery of many similar curves in enzymes, particularly those involved in regulation. Varied explanations of this type of behavior have been put forward (2-6). One of these, the model based on ligand-induced conformational changes (3), was shown to apply to a number of enzymes which exhibited positive cooperativity and also led to the prediction that a different type of deviation from Michaelis-Menten kinetics, i.e., negative cooperativity, would be observed. This has now been found in a number of cases (7-9).

Some quite unusual kinetic curves have also been found for proteins which are known to be pure and homogeneous (8-10). These findings, which might at one time have been discarded as "experimental artifacts," have been shown to represent real situations and may provide clues to the molecular design of regulatory enzymes. Accordingly, attempts to explain unusual kinetic and binding curves on the basis of protein structure have been initiated. In this process it has been found that further refinement of the ligand-induced conformational model previously designed to explain cooperative effects will allow an explanation of these anomalous curves and also provide further insight into subunit interactions.

In a protein composed of identical subunits, a subunit in contact with two neighbors (cf. Fig. 1) will inevitably have a different set of amino acid residues in contact with Neighbor 1 from the set in contact with Neighbor 2. This is a necessary consequence of the asymmetry of protein subunits. If these two and other similar sets of amino acid residues are designated as \( p, q, r \), etc., this nomenclature can be used to take into account the different binding domains when the subunit interactions are analyzed. In the previous mathematical derivation of binding equations (3), the complexity of this situation was avoided by using averaged subunit interaction terms, i.e., by assuming that every interaction was of the same type. This is quite a satisfactory approximation in some cases (and, as seen below, it is mathematically precise in certain situations), but in order to cope with the more complex cases it is desirable to extend this mathematics to allow for the diversity of binding domains. As will be seen below, the resulting analysis may provide an explanation for complex kinetic and binding patterns.

THEORETICAL

Subunit Arrangements—In Fig. 2 a schematic representation of some oligomeric proteins with identical subunits is shown. The identical chains are indicated as circles and the binding sites are identified by the letters \( p, q, \) and \( r \) to signify constellations of amino acid residues in the binding region. A pp interaction is isologous in the terminology of Monod, Wyman, and Changeux (2), and a pp interaction is heterologous. Stable dimers are assumed to be isologous (Fig. 2a) since the exposed

![Fig. 1. Schematic representation of a subunit with two neighbors, showing the two different binding domains which must exist with a change of conformation of one subunit.](http://www.jbc.org/content/245/2/6241/F1)

* This work was supported by Grant AM-007765 from the National Institutes of Health.
FIG. 2. Representations of various types of arrangements of subunits in a ring to obtain different types of oligomer. The arrangements on the left utilize isologous interactions only, and must contain an even number of subunits. The arrangements in the center column utilize only heterologous interactions, and may contain any number of subunits larger than 2. The arrangements on the right contain both isologous and heterologous interactions. Each of the arrangements shown in this figure is formally planar, and each subunit can be placed in two orientations only. Subunits containing the symbol ▲ are inverted with respect to the other subunits.

unsatisfied binding sites in a heterologous dimer should lead to further polymerization.

In a trimer, or any polymer containing an odd number of subunits, in which each subunit interacts only with two neighbors, either all the binding sets may be heterologous (Fig. 2b) or a mixed heterologous-isologous structure (Fig. 2c) is possible, but an all isologous structure is impossible. In a tetramer, or any polymer with an even number of subunits, in which each subunit interacts only with two neighbors, it is possible to have all isologous (Fig. 2d) or all heterologous (Fig. 2e) or mixed heterologous and isologous structures (Fig. 2f).

It will be convenient to refer to the structures of Fig. 2 in two different ways. The notation I_a, H_a, I_2H_2 will be understood to refer, respectively, to the trimeric structures for all isologous (Fig. 2d), all heterologous (Fig. 2e), and mixed, with two isologous and two heterologous binding domains (Fig. 2f and g). These structures can also be characterized more precisely by the actual binding domains, as p̄p̄p̄p̄, p̄p̄p̄p̄, and p̄p̄p̄p̄, respectively. In some cases specification of a particular cyclic sequence of binding domains may be desirable, in order to distinguish a unique structure, e.g. the two I_2H_2 structures shown in Fig. 2, f and g may be referred to as p̄p̄p̄p̄ p̄p̄p̄p̄ and p̄p̄p̄p̄ p̄p̄p̄p̄, respectively. It is clear that the notation with p̄p̄, etc., is the more precise, since the simpler notation will be ambiguous in some cases, but the simpler notation is less cumbersome when there is no ambiguity.

Certain characteristics are common to all structures having each subunit in contact with two neighbors only. All of these structures can be visualized schematically as shown in Fig. 2 as ring structures, even though it is not necessary for the subunits to be exactly coplanar.

When the subunits are each in contact with more than two neighbors, the number of possible arrangements increases. These will be discussed in detail in a later paper. Three "tetrahedral" cases are shown in Fig. 3, in which each subunit makes contact with each of the other 3 subunits in the tetramer. The simplest is the all isologous I_a, or p̄p̄p̄p̄ arrangements (Fig. 3, a, b) which is an extension of the I_a ring arrangement, in which the diagonal interactions between the r sites are included. The second arrangement, I_2H_4 (Fig. 3c), is one of several more complex ways of arrangeing 4 identical subunits, while the third (Fig. 3d) is an adaptation of the I_4 arrangement to a case where there are 2 similar but nonidentical types of subunit, as is observed for hemoglobin. These latter two cases are included to emphasize that the I_a arrangement is not the only possible tetrahedral arrangement, but since they both lead to very complicated binding equations, they will not be examined in detail.

Definition of Terms—The mathematical terms to describe the subunit interactions will utilize the terms defined previously (3, 6), with minor extensions to specify the individual binding domains discussed above. K_i will be used to indicate the equilibrium constant for a protein conformational change, and subscripts, as in K_{123}, to specify the precise conformational
change under consideration, as in Equation 1. The affinity constant $K_X$ will be written with a subscript to designate the particular conformation to which the ligand is bound, e.g., $K_{XB}$ as in Equation 2. The ligand $X$ can be substrate, inhibitor, or activator, and, of course, the affinity constants can apply to any conformation and thus apply to both exclusive and nonexclusive binding. The equations in this article are derived for exclusive binding (i.e., binding to one conformation only), but the mathematics and nomenclature are readily extended to other systems.

The expression for the subunit interaction terms $K_{AA}$, $K_{AB}$, $K_{BB}$, etc. in the earlier discussion (3) assumed an average interaction which did not distinguish between the binding domains. To distinguish between the various interactions, the previous nomenclature can be extended by the inclusion of subscripts, so that the three interactions shown in Fig. 4, a to c can be defined by $K_{s unsubscribe}$, $K_{AA unsubscript}$, and $K_{AB unsubscript}$, respectively. In this nomenclature the first sub-subscript identifies the binding domain for the first conformation listed, and so on. Thus, $K_{AA unsubscript}$ and $K_{AB unsubscript}$ are identical since they both refer to $K_{A unsubscript}$. As described previously (3), the standard state will be taken to be the species for which every subunit is in the $A$ conformation, with no ligand bound, i.e., $A_4$ in the case of a tetramer. The effect of this in writing an expression for the concentration of any other species is that every $K_{AB}$ or $K_{BB}$ term in the numerator will be accompanied by the analogous $K_{AA}$ term in the denominator, and thus, all the $K_{AA}$ terms can be set equal to 1 without loss of generality. This does not imply that all of the subunit interactions in the $A_4$ species are equivalent, only that the $A_4$ form is the standard reference state. This simplification is possible only when no association or dissociation of subunits occurs during ligand binding. It may be noted that polymerization equilibria have been discussed in connection with protein binding behavior (6), and simple extension of the present equations and nomenclature could encompass this possibility.

Derivation of Saturation Curves—A general form of a saturation equation (11) is shown in Equation 5, where $N_X$ is defined as the average number of molecules of ligand $X$ bound per molecule of protein, $n$ is the total number of binding sites, $\bar{Y}$ is the average fractional saturation, and the $\psi$ terms are the parameters which define the curve and are combinations of the molecular parameters $K_B$, $K_X$, $K_{AB}$, $K_{BB}$ etc. The various models for the mechanism of binding will give equations of the same form, but the relationships between the $\psi$ terms will vary.

In general a $\psi$ term represents the association constant for the formation of the protein containing $i$ molecules of ligand from the unliganded protein and free ligand. If we consider a protein with binding domains $pp\,qq\,pp\,qq$ as shown in Fig. 5, there are six different $A_pB_qX_s$ species, II to VII, which can be represented as arising from the unliganded species, $A_p\,I$. If we consider the $\psi$ terms for the saturation curves for all four of the ring tetrahexaenes, $I_4$, $I_6$, $I_8$, $I_{10}$, and also for the simplest tetrahedral tetramer, $I_{10}$. For simplicity the $K_{p unsubscribe}$ and $K_{I unsubscribe}$ terms have been omitted from this table, because they are the same for each $EN$ form. They must, of course, be included when the $\psi$ terms are substituted into Equation 6.

Plotting Saturation Curve—In plotting saturation curves, it is frequently preferable to plot $N_X$ or $\bar{Y}$ or the velocity, $V$, against the logarithm of the ligand concentration rather than against the
Table I

Subunit interaction terms for oligomeric proteins

Subunit interaction terms for combinations of A (circle) and B (square) conformations for various arrangements of identical subunits in a tetrameric protein are shown. The arrangements are listed in the top line of the table and include all of the four possible ring arrangements, and one of the four possible tetrahedral arrangements. Each species is considered relative to the unliganded $A_4$ standard state, and thus all $K'_{AA}$ terms are defined as 1, and do not appear in the table.

<table>
<thead>
<tr>
<th>Subunit interaction terms for combinations of A (circle) and possible ring arrangements, and one of the four possible tetrahedral arrangements. Each species is considered relative to the unliganded $A_4$ standard state, and thus all $K'_{AA}$ terms are defined as 1, and do not appear in the table.</th>
<th>All-isologous $I_4$</th>
<th>Mixed $I_4$</th>
<th>All-isologous $I_4$</th>
<th>All-heterologous $I_4$</th>
<th>All-isologous tetrahedron $I_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subunit interaction terms for combinations of A (circle) and possible ring arrangements, and one of the four possible tetrahedral arrangements. Each species is considered relative to the unliganded $A_4$ standard state, and thus all $K'_{AA}$ terms are defined as 1, and do not appear in the table.</td>
<td>All-isologous $I_4$</td>
<td>Mixed $I_4$</td>
<td>All-isologous $I_4$</td>
<td>All-heterologous $I_4$</td>
<td>All-isologous tetrahedron $I_4$</td>
</tr>
<tr>
<td>$x_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ex.

Binding Domains and Subunit Interactions

Vol. 235, No. 23
**TABLE II**

*Simplified subunit interaction terms for some oligomeric proteins*

Subunit interaction terms for each EX \(_i\) species for the I\(_4\), H\(_4\), and I\(_6\) arrangements of subunits in a tetramer are given. These terms were obtained from those in Table I by summing all of the terms for each EX \(_i\) species. Where possible, the terms \(K_{AB}^2\) and \(K_{BB}^2\) (without sub-subscripts) are used as defined at the top of the table. The \(\psi\) terms for Equation 6 may be obtained from this table by multiplying the listed subunit interaction term by \(K_{AB}^4K_{BB}^4\) for each EX \(_i\). These two constants refer to the binding of i moles of \(X\) to the B conformation and the transition \(A \rightarrow B\) for i subunits. Arrangements having both isologous and heterologous binding sets (I\(_{HH}\), I\(_{HH}\), I\(_{SH}\), I\(_{SH}\), I\(_{SH}\), and the hemoglobin arrangement shown in Fig. 3c) result in subunit interaction terms which are too complex to be simplified, and give \(\psi\) terms which are fully independent of one another, so that Equation 6 must be used to define the binding curve.

<table>
<thead>
<tr>
<th>Enzyme species</th>
<th>I(_4)</th>
<th>H(_4)</th>
<th>I(_6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>EX (_2)</td>
<td>(4K_{AB}^2)</td>
<td>(4K_{AB}^2)</td>
<td>(4K_{AB}^3)</td>
</tr>
<tr>
<td>EX (_3)</td>
<td>(2K_{AB}^2 + K_{BB}^2)</td>
<td>(2K_{AB}^2 + 2K_{AB}^4)</td>
<td>(2[K_{AB}^2K_{BB}^2,K_{AB}^2K_{BB}^2,K_{AB}^2K_{BB}^2,K_{AB}^2K_{BB}^2])</td>
</tr>
<tr>
<td>EX (_4)</td>
<td>(K_{BB}^4)</td>
<td>(K_{BB}^4)</td>
<td>(K_{BB}^6)</td>
</tr>
</tbody>
</table>

**Definitions**

<table>
<thead>
<tr>
<th>K(_{AB})</th>
<th>(\sqrt{K_{AB}K_{AB}})</th>
<th>(\sqrt{K_{AB}K_{AB}})</th>
<th>(\sqrt{K_{AB}K_{AB}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>K(_{BB})</td>
<td>(\sqrt{K_{BB}K_{BB}})</td>
<td>(K_{BB})</td>
<td>(\sqrt{K_{BB}K_{BB}})</td>
</tr>
</tbody>
</table>
**RESULTS**

**Limitations of Constants**

The $\psi_i$ terms derived as above for $\psi_4$ for the $IIHH$ structure are shown in Table I. If the terms in Table I for the $IIHH$ and $IIHH$ structures are examined carefully, it may be seen that in each case the four values are completely independent of one another, so that there are no restrictions on the relative values of the four $\psi_i$ terms. A similar conclusion is reached when the $\psi_i$ terms for other structures containing both isologous and heterologous binding are examined. Thus, it appears that, in general, mixed structures must be described by the most general form of the binding equation, Equation 3, or Equation 6 in the special case of tetrameric proteins.

For the remaining structures, which contain either all isologous or all heterologous binding sets, the $\psi_i$ terms for the saturation curves are not independent. For example, in the $H_4$ case, it is found that $K_{A_{i}B}$ and $K_{B_{i}A}$, invariably occur together as a product, and thus can be represented as a single constant, $K_{AB} = K_{AB_{i}}K_{BA_{i}}$. Similar combinations of constants are also possible for the all isologous structures, $I_4$ and $I_6$. For these cases the expressions for the $\psi_i$ terms can be simplified as shown in Table II. It is immediately apparent that the binding equation for the $H_4$ structure is identical with that for the "square" case derived by Koshland, Nemethy, and Filmer (3). The $I_4$ case also simplifies to give the same equation in the special case where $K_{BB\epsilon} = K_{BB_{\epsilon}}$. The $I_4$ and $I_6$ structures give equations identical with the previously derived (3) square and tetrahedral cases, except for the $\psi_i$ terms. It should be emphasized that although the forms of these equations are similar, the $K_{AB}$ and $K_{BB_{\epsilon}}$ constants are defined differently, and this may be important when detailed evaluation of these constants becomes possible.

**Comparison of Saturation Curves for Most Likely Cases of Tetramers with Identical Subunits**

From the previous discussion it can be seen that the $I_{4}, H_{4},$ and $I_{6}$ structures of a tetramer give a restricted form of the binding equation, i.e., the constants $\psi_i$, $\phi_i$, and $\psi_4$ of Equation 6 cannot assume all conceivable values in relation to one another. As we shall discuss elsewhere, these structures, together with possibly $I_{12}$ and $I_{12}H_{4}$ tetrahedral structures, are the most likely to occur as the main components of a tetrameric protein. Accordingly, we shall examine the characteristics of the saturation curves expected for each of these models.

All Heterologous Tetramer, $H_4$.—The equation for the all heterologous tetramer is identical in form with the "square" case of Koshland et al., (3), and consequently it gives identical saturation curves. Three representative cases are shown in Fig. 7, (a) a positively cooperative enzyme, (b) a Michaelis-Menten enzyme, and (c) a negatively cooperative enzyme. There are four general characteristics of curves for this model. (a) The curve is symmetrical about the half-saturation point; (b) there can be either one or three points of inflection; (c) the value of the substrate concentration at half-saturation is given by $1/K_{2}K_{4}K_{BB_{\epsilon}}$; and (d) the shape of the curve, or in other words its cooperativity, is determined by the value of $K_{AB}/\sqrt{K_{BB_{\epsilon}}}$. It should be noted that while these are necessary conditions for the existence of a protein in the $H_4$ state, they are not sufficient, because the same characteristics can be observed with any of the other structures if the parameters for those models fortuitously yield the same relationship. The possibility that constants which can assume all possible values would fortuitously fit the restrictions of the $H_4$ case might seem unlikely. However, as will be seen below, this is in fact a very real possibility.

All Isologous Ring Tetramer, $I_4$.—The equation for the $I_4$ case differs from that of the $H_4$ case only in the expression for $\psi$. 

---

**Fig. 6.** a and c, theoretical plots of $N_{A} \times \log \left(\frac{X}{X_{0}}\right)$; b and d, the curves of a and c replotted as $N_{A} \times \log \left(\frac{X}{X_{0}}\right)$. Illustration of the fact that diagnostic features of a binding curve are more evident in a logarithmic plot than in a linear plot. The point of inflection indicated by an arrow in a is not perceptible in b, and the very pronounced point of inflection at the half-saturation point of c is barely visible in d. The symmetry seen in a and c is not detectable in b and d.
Consequently, we might expect that the two models would result in rather similar curves, and that is what is found. All of the first three characteristics listed for \( H_1 \) apply to \( I_4 \) also, but in addition to \( \frac{K_{BB}^2}{\sqrt{K_{BS}^2}} \), the quantity \( (K_{BB}^2 + K_{BS}^2)/2\sqrt{K_{BB}^2 K_{BS}^2} \), which we shall call \( \delta \), also affects the shape of the curve. This quantity can have any value greater than or equal to 1. When \( \delta = 1 \), the equation for the saturation curve is identical with the equation in the \( H_4 \) case, but increasing \( \delta \) above 1 tends to produce two new points of inflection placed symmetrically about the half-saturation point of the curve, or, if the curve already has three points of inflection, to accentuate the flattened portion of the curve.

It must be noted that although it is mathematically possible to distinguish between the \( I_4 \) and \( H_4 \) cases by saturation curves alone, it is much less likely that they can be distinguished experimentally. This results from the definition of \( \delta \), because relatively large differences between \( K_{BB} \) and \( K_{BS} \) produce quite small values of \( \delta \). Thus, if \( K_{BB} \) and \( K_{BS} \) differ by a factor of 100, the value of \( \delta \) is only 5.05, a value which produced only minor departures from the curve for the \( H_4 \) case.

All Isotologous Tetrahedral Tetramer, \( I_4 \).—The equation for the \( I_4 \) case is similar to that for the tetrahedral case derived earlier \((3)\), being different only in the expression for \( \varphi_2 \). The following characteristics apply to the binding curves for this case. (a) The curve is symmetrical about the half-saturation point; (b) there can be one, three, or seven points of inflection; (c) the value of the ligand concentration at half-saturation is given by \( 1/K_B K_{sB} \), and (d) the shape of the curve is defined by the value \( K_{BB}/\sqrt{K_{BS}^2} \) and by the value of the parameter \( \delta \) which is defined by Equation 7. Just as in the \( I_4 \) case, \( \delta \) has a minimum value of 1, and when \( \delta = 1 \), the equation reduces to the equation for the

\[
\delta = \frac{K_{BB}^2}{3K_{BB}^2} \left( \frac{K_{BB}^2 + K_{BS}^2}{K_{BB}^2 + K_{BS}^2 + K_{BB}^2 + K_{BS}^2} \right) = \frac{K_{BS}^2}{K_{BB}^2 + K_{BS}^2 + K_{BB}^2 + K_{BS}^2} \quad (7)
\]

averaged tetrahedral case of Koshland et al. \((3)\). Increasing the value of \( \delta \) tends to produce or accentuate flattening of the curve in the region of half-saturation, and to decrease it in other parts of the curve. In other words, as \( \delta \) is increased, the number of points of inflection tends to change from one to seven to three.

In physical terms, the \( I_4 \) structure is a special case of \( I_4 \) in which the contribution of the \( r \) sites to the behavior is negligible. Increasing \( \delta \) tends to make the curve for \( I_4 \) resemble the curve for \( I_4 \). Above a limiting value of \( \delta \)

\[
\delta > \frac{K_{BB}^2}{2K_{BS}^2} + \frac{2\sqrt{K_{BB}^2 K_{BS}^2}}{K_{BB}^2}
\]

the \( I_4 \) curve becomes identical with the \( I_4 \) curve.

Relationships between Intrinsic Binding Constants

The restrictions on the permissible binding curves for the various models can be conveniently expressed as relationships between the intrinsic binding constants \( K_1 \) of the protein. These intrinsic constants are the sequential association constants corrected for statistical factors, i.e. \( K_1' \) = \( K_1/4 \) = \( \varphi_1/4 \), \( K_2' = 2K_2/3 = 2\varphi_2/3 \varphi_4 \), \( K_3' = 3K_3/2 = 3\varphi_3/2 \varphi_4 \), \( K_4' = 4K_4 = 4\varphi_4 \). The use of intrinsic constants has the advantage that cooperative effects become very obvious when these constants are compared, i.e. a Michaelis-Menten protein has \( K_1' = K_2' = K_3' = K_4' \), a positively cooperative protein has \( K_1' < K_2' < K_3' < K_4' \), etc. The required relationships between the constants are shown in Table III for the two simplest sequential cases \((3)\), the various extended sequential cases discussed in this paper, and three cases of the symmetry or concerted model of Monod, Wyman, and Changeux. It can be seen that although the restrictions overlap considerably, so that in many cases the values of the constants would not determine the model unequivocally (no matter how accurately the constants might be determined), there are nonetheless important differences between the models which can be used for diagnostic purposes. In the first place, it can be seen that all of the sequential cases allow negative cooperativity (\( K_1' > K_2' > K_3' > K_4' \)) as well as the more familiar positive cooperativity, but that the concerted model does not permit negative cooperativity, even if it is generalized to include any number of symmetrical conformations. Secondly, the condition for a symmetrical binding curve is \( K_1'K_4' = K_2'K_3' \), and it may be seen that this condition is required for all of the simplest sequential cases except for the mixed structures \( IIH \) and \( IIIH \). It is allowed, but not required, for the concerted model with nonexclusive binding, and is forbidden for the concerted model with exclusive binding. Thus, we find that if the experimental data exhibit significant negative cooperativity, the concerted model with exclusive binding can be unequivocally excluded. If the binding curve shows significant departures from symmetry about the half-saturation point, then those cases of the sequential model which require such symmetry can be excluded.

---

1 In this discussion we shall refer to the Monod, Wyman, and Changeux model as the concerted model to avoid confusion between the symmetry of the binding curves and the symmetry of the protein structure postulated in the model.
Table III

Relationships between molecular models and intrinsic constants

In some cases the relationships are identical, e.g. for the simplest sequential square case and the $H_t$ case, so that these two models are not in principle distinguishable on the basis of binding curves alone. Similarly, the relationship for the $I_t$ case is a special case of the relationship for the $I_t$ case, so that in a choice between these two models the binding curve might rule out the $I_t$ case but could never rule out the $I_t$ case. The requirement for a symmetrical binding curve, $K_1' = K_3'$, is part of the requirement for all of the sequential model cases listed, except for the $IIHI$ and $IIIH$ cases.

<table>
<thead>
<tr>
<th>Model</th>
<th>Relationship between intrinsic constants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplest sequential model</td>
<td>$0.333 \leq K_2/K_3 &lt; 2.25$</td>
</tr>
<tr>
<td>Square (averaged interactions)</td>
<td>$K_1' = K_2' = K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>Tetrahedral (averaged interactions)</td>
<td>$K_2' = K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>$I_t$</td>
<td>$0.333 &lt; K_2'/K_3' &lt; 2.25$</td>
</tr>
<tr>
<td>$IIHH$ and $IIH$</td>
<td>No restrictions on allowed values</td>
</tr>
<tr>
<td>$H_t$</td>
<td>$0.333 &lt; K_2'/K_3' &lt; 2.25$</td>
</tr>
<tr>
<td>Symmetry model</td>
<td>$K_1' &lt; K_2' = K_3' = K_4'$</td>
</tr>
<tr>
<td>Exclusive binding</td>
<td>$K_1' &lt; K_2' &lt; K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>Nonexclusive binding, two</td>
<td>$K_1' &lt; K_2' = K_3' = K_4'$</td>
</tr>
<tr>
<td>conformations</td>
<td>$K_1' &lt; K_2' &lt; K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>Generalized, any number of</td>
<td>$K_1' = K_2' = K_3' = K_4'$</td>
</tr>
<tr>
<td>conformations</td>
<td>$K_1' &lt; K_2' &lt; K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>Types of cooperativity defined</td>
<td>$K_1' &lt; K_2' &lt; K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>No interaction between sites</td>
<td>$K_1' &lt; K_2' &lt; K_3' &lt; K_4'$</td>
</tr>
<tr>
<td>(Michaelis-Menten)</td>
<td>$K_1' &gt; K_2' &gt; K_3' &gt; K_4'$</td>
</tr>
<tr>
<td>Simple positive cooperativity</td>
<td>$K_1' &gt; K_2' &gt; K_3' &gt; K_4'$</td>
</tr>
<tr>
<td>Simple negative cooperativity</td>
<td>$K_1' &gt; K_2' &gt; K_3' &gt; K_4'$</td>
</tr>
</tbody>
</table>

* Excluding the case where all the parameters are equal.

Lactate Dehydrogenase Data of Anderson and Weber

The most complete data available in the literature for a protein which appears to deviate from the Adair equation are those obtained by Anderson and Weber (12) for the binding of NADH to the purified isozymes of bovine muscle lactate dehydrogenase. One of these isozymes, pure $M_4$ lactate dehydrogenase, is of particular interest because it consists of 4 identical subunits and shows an asymmetric curve (Fig. 8). Since measurements were made over about 90% of the titration range of the enzyme, at an amply large number of substrate concentrations, this seemed to be a promising case to study in detail. Anderson and Weber noted the unusual shape of the curve and deduced that it could not be explained by the Adair equation. The seriousness of this for protein studies was recognized by Weber and Anderson (13) who suggested that a combination of fast and slow relaxations between protein species might explain the data. Since the latter leads to a conflict with the law of microscopic reversibility, it seemed desirable to determine whether any simpler explanations were possible.

The points in Fig. 8 are plotted from careful measurements of the coordinates of the points in Fig. 3a of the original paper, and the line which is drawn through them corresponds as closely as possible to the line drawn by Anderson and Weber. Two principal points may be made about this line. (a) It shows two rather sharp changes in slope, and (b) it does not approach $N_s = 0$ at low substrate concentration, nor does it approach $N_s = 4$ at high substrate concentration. While the second point could perhaps be accounted for by postulating the presence of small amounts of impurities, the first seems to be quite irreconcilable with the normal mechanism for substrate binding described by the Adair equation.

In order to determine how closely the data could be described by the Adair equation, the curve-fitting procedure described elsewhere (14) was used and the closeness of fit was assessed by summing the squares of the deviations between the observed and calculated values of $N_s$. The results are shown in Table IV, where a comparison is shown between the fits obtained using the "non-Adair" curve of Anderson and Weber, and some other models, including the simplest sequential model. Since we are dealing with a nonlinear system, it is not strictly valid to apply standard statistical tests to these results in order to assess their significance. Moreover, the true number of parameters needed to describe the ad hoc curve drawn by Anderson and Weber may well be more than the five given in the table. However, it is clear from inspection of the results that the general four parameter equation fits the data considerably better than the equation required by the simpler models, and that the ad hoc curve is only moderately better than the four parameter curve. Applying an F test to the results, an F value of 1.65 is obtained, which indicates that the four parameter equation is a good fit for the experimental data at the 95% confidence level, and thus agrees...
TABLE IV
Fit of various models to lactate dehydrogenase data

Results obtained using the curve-fitting procedure to fit various theoretical equations to the data given by Anderson and Weber (1970) for the binding of NADH to $M_1$ lactate dehydrogenase are given. The experimental curve (a) is asymmetric, but deviates from symmetry in the opposite sense from the deviations from symmetry required by the concerted model with exclusive binding to one conformation (d). Thus, this model fits the data more poorly than the sequential models (c) which require symmetrical binding curve. The concerted model with nonexclusive binding (e), with an extra parameter, permits deviations from symmetry in either sense, and fits the data better than the sequential model.

<table>
<thead>
<tr>
<th>Model used to fit data</th>
<th>Number of parameters</th>
<th>Sum of squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)  &quot;Non-Adair&quot;</td>
<td>5</td>
<td>0.099</td>
</tr>
<tr>
<td>(b) Unrestricted Adair equation</td>
<td>4</td>
<td>0.168</td>
</tr>
<tr>
<td>(c) Tetrahedral</td>
<td>2</td>
<td>0.655</td>
</tr>
<tr>
<td>(d) Concerted, exclusive binding</td>
<td>2</td>
<td>1.11</td>
</tr>
<tr>
<td>(e) Concerted, nonexclusive</td>
<td>3</td>
<td>3.356</td>
</tr>
<tr>
<td>(f) Mixture of five species*</td>
<td>12</td>
<td>0.151</td>
</tr>
</tbody>
</table>

* This line refers to the ad hoc curve drawn by Anderson and Weber which is not known to be described by any theoretical equation.

* At least five parameters (i.e. at least one more than the four Adair parameters) are needed to define a non-Adair curve.

* These results refer to the tetrahedral case of the simplest sequential model, but essentially the same fit was obtained with the square cases, and with the more general $H_1, I_1$, and $I_0$ models described in this paper.

* Of the five species, one was considered to be inactive, one was fully saturated over the whole range of NADH concentrations, and the other three individually obeyed the simplest sequential "square" model.

with the conclusion from visual inspection that the results do not permit rejection of the Adair equation.*

DISCUSSION

If the ligand-induced model is extended to the situation in which the subunit interactions are not averaged, but the binding domains are considered in detail, it is seen that the saturation curves become slightly more complex. These curves still remain in general very similar to those derived by Koshland et al. (3), and only in the case of highly accurate data is there any distinction between the more accurate curves and those which assume averaged interactions. Since highly accurate data are available in some cases, and since the use of other types of data, e.g. $B$ versus $T$, will undoubtedly lead to better tests of the more detailed subunit interactions, it is important to analyze the differences between the theoretically complete equations and the simpler averaged equations.

It is seen that the all isologous and all heterologous examples,

* G. Weber (personal communication) has pointed out that the most pronounced deviation from the Adair curve occurs in the vicinity of $N_S = 1$, and that ideally the $F$ test should be applied to the data in this region, rather than to the whole set of data. If the points in this part of the curve are significantly more accurate than the average for the whole curve, then it is quite possible that this deviation is more significant than our test would suggest. Thus, our analysis emphasizes the fit over the entire curve; Anderson and Weber's analysis emphasizes the lack of fit in a local region.

regardless of the number of subunits, give restricted saturation equations in which the individual association constants are related to one another, e.g. by the relationship $K'_1K'_4 = K'_2K'_3$, and thus are not completely independent. On the other hand, mixed structures which contain both isologous and heterologous bonds lead to the general form of the binding equatino in which the individual association constants are independent of one another. Of course, it is always possible that any of the more general models can fortuitously give the characteristics of a simpler model, in which case other types of diagnostic tests are needed to distinguish between models.

Certain characteristics of the binding curve itself can distinguish definitively between models. For example, a significantly asymmetric binding curve, where $K'_1K'_4 \neq K'_2K'_3$, indicates that none of the simpler cases of the sequential model can apply, whereas a curve showing negative cooperativity indicates that the symmetry model can be ruled out. In some cases, points of inflection in the binding curve give useful information. For any pure protein, which must obey the Adair equation (Equation 3), the even numbered points of inflection must occur in the vicinity of integral values of $N_X$, i.e. any plateau regions of the binding curve must occur close to integral values of $N_X$. No such requirement applies to mixtures of species however, and so the occurrence of a plateau region at a significantly noninteger value of $N_X$ would be strong evidence that the protein was impure.

In this paper we have explored only the theoretical equations for cases in which individual subunits can exist in two conformations. The reason for this limitation is that the simplest cases illustrate the principles and mathematical techniques for comparing detailed considerations of binding domain interactions with less detailed approaches using averaged interactions. However, it is clear that the detailed considerations outlined here will also be important in the more complex cases, e.g. those in which a change in conformation of one subunit leads to partial distortions of a neighboring subunit (7, 9). As in the present examples, the detailed considerations of binding domains will sometimes lead to different saturation equations from those obtained with averaged interactions. The basic principles for consideration of binding domains will be the same, although the detailed mathematics may be different for the more complex cases.

The extension of previous procedures to encompass the accurate consideration of the actual binding domains involves significant changes in the theoretical binding equations in some cases and little or no change in others. Since accurate curve-fitting procedures (14) and additional diagnostic tests (15) are now available, this extension will becoming of increasing importance and may explain binding and kinetic behavior in regulatory proteins which seem to be anomalous when compared with models which use averaged interactions.

Acknowledgment—We acknowledge the valuable and constructive suggestions made by Dr. Gregorio Weber.

REFERENCES