# *RELATION OF PROTEIN SUBUNIT INTERACTIONS TO THE MOLECULAR SPECIES OBSERVED DURING COOPERATIVE BINDING OF LIGANDS**

By James E. Haber and D. E. Koshland, Jr.

DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CALIFORNIA, BERKELEY

*Read before the Academy, April 26, 1967*†

To understand cooperative effects in proteins and regulatory enzymes it is necessary to correlate the changes in apparent affinity of the ligand with the changes in the protein structure. A model based on sequential changes in the conformations of subunits of the protein can explain ligand binding and enzyme activity in hemoglobin and regulatory proteins.[1-3] In Figure 1, a schematic illustration of the molecular species potentially present in a tetrameric protein in which the ligand can bind to only one of two subunit conformations is shown. Not all of these molecular species are necessarily present in appreciable concentrations, and models with fewer species can, for example, rationalize the $O_2$ binding curve of hemoglobin.[1, 4] The actual molecular species present under different experimental conditions can clarify the nature of the subunit interactions and give clues as to the evolutionary significance of cooperative properties. Some analyses of these relationships are discussed in this paper.

In Figure 2, the conformational changes of an isolated subunit or a monomeric

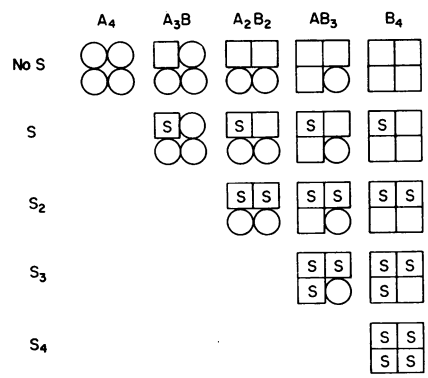



FIG. 1.—Schematic illustration of possible molecular species for a general model involving four subunits, exclusive binding, and two possible conformations per subunit. Species with identical number of subunits in $B$ conformation, vertical columns; identical numbers of ligand bound, horizontal rows. More comprehensive models involving preferential rather than exclusive binding, more than two conformations, etc., can be constructed and analyzed similarly.

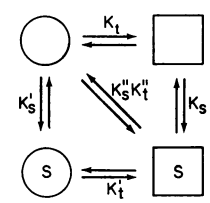



FIG. 2.—Illustration of alternative pathways in a ligand-induced conformation change. $K_sK_t$ pathway designates isomerization to new conformation followed by binding of ligand, $K_s'K_t'$ pathway; binding of ligand followed by conformation change; and $K_s''K_t''$ indicates alternate pathways such as multistep changes in conformation and binding. $K_1K_1$ must equal over-all equuilibrium constants for any pathway, e.g., $K_s'K_t'$.

protein which interacts with a ligand are shown. The initial conformation in the absence of ligand is designated by a circle or the letter $A$. The changed conformation induced by the binding of the ligand is designated by a square or the letter $B$. Three possible pathways for achieving this conformation change are ($a$) an initial

isomerization followed by the binding of substrate to the changed structure (pathway $K_sK_t$), (*b*) an initial binding of substrate followed by a conformational isomerization (pathway $K_s'K_t'$), or (*c*) a concurrent binding and conformational isomerization designated by the diagonal arrow. Thermodynamic considerations require that $K_sK_t = K_s'K_t'$, and any additional pathway, such as the one shown on the diagonal, must give an over-all equilibrium constant equal to this product.

When this monomer is associated with other such monomers in a polymeric protein, a change in conformation in one subunit may produce varying degrees of change in the shape and stability of adjacent subunits. The ligand-induced change from the *A* to the *B* conformation (circle to square, Fig. 3) may leave the

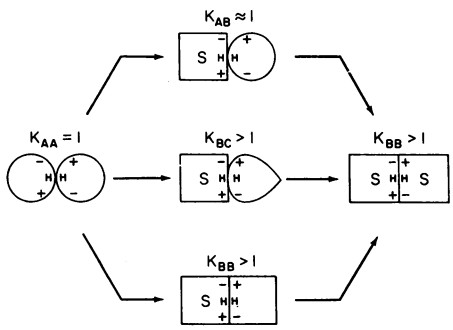



FIG. 3.—Schematic illustration of the way in which conformation change in one subunit can affect stability and/or conformation of neighboring subunit. Initial subunits in *A* conformation (*circles*) held together by hydrophobic bond (H), electrostatic charges (+ and −) being too far apart to interact significantly. (Medium contains positive and negative ions to nullify effects of distant changes.) Ligand (*S*) binding induces change to *B* conformation (designated by *square*). When ligand binds changing conformation of one subunit from *A* to *B*, adjacent subunits may be unchanged (*top line*), partially distorted to conformation *C* (*middle line*), or completely distorted to final conformation *B* (*bottom line*). It is assumed here that adjacent *BB* and *BC* subunits produce added electrostatic interactions which stabilize them relative to the *AA* interaction (hydrophobic only), i.e., $K_{BC}$ and $K_{BB}$ are greater than 1. Since the relative changes in the interactions, not the absolute magnitude of the $K_{AA}$ interaction, are important in cooperative effects, $K_{AA}$ is set at 1 for these purposes (cf. ref. 1).

conformation of an adjacent subunit unchanged, cause it to be partially changed, or may isomerize it completely to the same *B* conformation. Each of these changes may in turn increase, decrease, or leave unchanged the net attraction between subunits. In Figure 3, some illustrative changes are shown using hydrophobic bonds (H) and electrostatic bonds (+ and −) to depict schematically how such shape and energy changes might occur. The nomenclature is that described previously,[1] where $K_{AA}$, $K_{AB}$, $K_{BB}$, etc., are equilibrium constants expressing the strength of interactions between adjacent subunits in *AA*, *AB*, and *BB* conformations. *A* (or a circle) in all cases represents the conformation of the individual subunit when no ligand is bound to the protein. For the present calculations we shall limit our consideration to the situation where only the two conformations, *A* and *B*, are ever present in significant concentrations.[5]

In Figure 4, the molecular species of a tetrameric protein in the "square" geometry at various fractional saturation values ($\bar{Y}$) are shown for one particular case. To calculate the molecular species present, the equation and assumptions for the general model of Figure 1 derived previously (equation (39), ref. 1) were used together with an assigned value of $K_{AB} = 1$. With the indicated combination of $K_s$, $K_t$, $K_{AB}$, and $K_{BB}$, a binding curve with a Hill coefficient of 2.7 is obtained, and only the species lying on the lowest diagonal of Figure 1 are present in any significant concentration. This corresponds to the "simplest sequential model," in which only two conformations of a subunit can exist and only the subunits con-

taining bound ligand change from conformation $A$ to conformation $B$. This model has been discussed extensively in a previous publication.[1] The binding of $O_2$ to hemoglobin has a Hill coefficient of 2.7.[6]

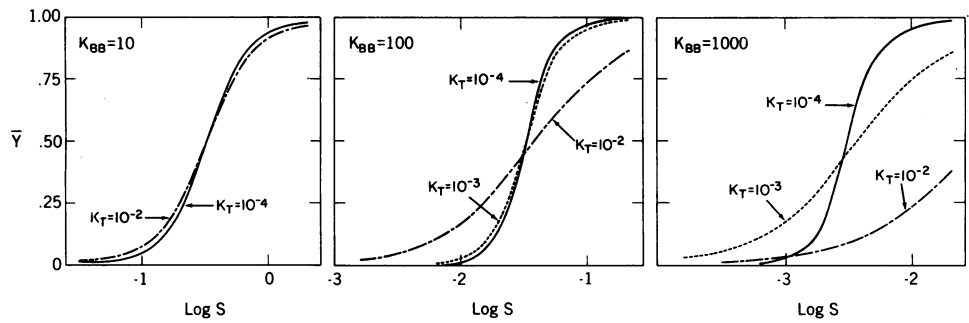
Another extremely simple model which has been shown to fit the saturation curves of hemoglobin is the symmetry model of Monod, Wyman, and Changeux[5]

| | $A_4$ | $A_3B$ | $A_2B_2$ | $AB_3$ | $B_4$ | |
|---|---|---|---|---|---|---|
| at $\bar{Y}$ = 0 | 99.96 | .04 | -- | -- | -- | No S |
| at $\bar{Y}$ = 0.5 | 30.99 | .01 | -- | -- | -- | No S |
| | | 12.40 | .03 | -- | -- | $S_1$ |
| | | | 13.02 | .04 | -- | $S_2$ |
| | | | | 12.40 | .12 | $S_3$ |
| | | | | | 30.99 | $S_4$ |
| at $\bar{Y}$ = 0.75 | 10.37 | -- | -- | -- | -- | No S |
| | | 6.37 | .01 | -- | -- | $S_1$ |
| | | | 10.28 | .03 | -- | $S_2$ |
| | | | | 15.04 | .15 | $S_3$ |
| | | | | | 57.74 | $S_4$ |

FIG. 4.—Distribution of molecular species calculated to be present for general model of Fig. 1 assuming the "square geometry" and $K_{AB} = 1$, $K_{BB} = 10$, $K_t = 10^{-4}$ at various values of $\bar{Y}$. Dashes indicate less than 0.01%. Vertical columns for identical subunit conformation; horizontal columns for identical number of ligands, as in Fig. 1.

in which only the species $A_4S_i$ and $B_4S_i$ are assumed to be present,[7] i.e., the symmetry condition requires that all the subunits in the protein exist in the same conformation.

Since both these simple models and the more general model of Figure 2 can give cooperative binding curves, it seemed of interest to vary the values of $K_s$, $K_t$, $K_{BB}$, etc., in the more general model of Figure 1 to see under what circumstances the more simple models are likely to fit the experimental observations. In Figure 5, both the saturation curves and molecular species present for a tetrameric protein in a "square" geometry are shown, with varying values of $K_t$ and $K_{BB}$. An examination of the distribution of molecular species present at the mid point of saturation illustrates the conditions under which each of the simple models is a good approximation. For example, when the increased affinity between $B$ conformation is moderate ($K_{BB} = 10$) and the energy change for a spontaneous change from conformation $A$ to $B$ is relatively large ($K_t$ less than $10^{-3}$), the "simplest sequential model" is an excellent description. As $K_t$ gets larger and less energy is required to convert a molecule from conformation $A$ to conformation $B$, the contribution of species containing unsaturated subunits in the $B$ form (e.g., $B_4S_3$) increases. When $K_t = 10^{-3}$ and $K_{BB} = 100$, for example, significant concentrations of species $i$ with unliganded subunits in the $B$ conformation are found. When $K_{BB}$ is very large, the contribution of "hybrid species" becomes much less significant, and the molecular species corresponding to the symmetry model of Monod, Wyman, and Changeux are the only species present in significant concentrations. Similar calculations for a variety of values leads to the following conclusions: (*a*)




$K_{BB}$ = 10 $K_T$ = $10^{-4}$

| | | | | |
|---|---|---|---|---|
| 99.96 | .04 | -- | -- | -- |
| 31.0 | -- | -- | -- | -- |
| | 12.4 | -- | -- | -- |
| | | 13.0 | -- | -- |
| | | | 12.4 | .1 |
| | | | | 31.0 |

n = 2.7

$K_{BB}$ = 100 $K_T$ = $10^{-4}$

| | | | | |
|---|---|---|---|---|
| 99.96 | .04 | -- | -- | -- |
| 46.22 | -- | -- | -- | -- |
| | 1.85 | -- | -- | -- |
| | | 1.85 | -- | -- |
| | | | 1.85 | 1.9 |
| | | | | 46.2 |

n = 3.8

$K_{BB}$ = 1000 $K_T$ = $10^{-4}$

| | | | | |
|---|---|---|---|---|
| 99.9 | .04 | -- | -- | -- |
| 40.3 | -- | -- | -- | -- |
| | .2 | -- | -- | .2 |
| | | .2 | -- | 2.4 |
| | | | .2 | 16.1 |
| | | | | 40.3 |

n = 3.4

$K_{BB}$ = 10 $K_T$ = $10^{-3}$

| | | | | |
|---|---|---|---|---|
| 99.6 | .4 | -- | -- | -- |
| 30.4 | .1 | -- | -- | -- |
| | 12.2 | .3 | -- | -- |
| | | 12.8 | .4 | -- |
| | | | 12.2 | 1.2 |
| | | | | 30.4 |

n = 2.7

$K_{BB}$ = 100 $K_T$ = $10^{-3}$

| | | | | |
|---|---|---|---|---|
| 99.5 | .4 | -- | -- | -- |
| 38.0 | .2 | -- | -- | -- |
| | 1.5 | .3 | .1 | -- |
| | | 1.5 | .4 | 2.3 |
| | | | 1.5 | 15.0 |
| | | | | 38.0 |

n = 3.2

$K_{BB}$ = 1000 $K_T$ = $10^{-3}$

| | | | | |
|---|---|---|---|---|
| 49.7 | .2 | .2 | .2 | 49.7 |
| 5.9 | -- | -- | -- | 5.9 |
| | -- | -- | -- | 23.5 |
| | | -- | -- | 35.2 |
| | | | -- | 23.5 |
| | | | | 5.9 |

n = 1.2

FIG. 5.—Saturation curves and molecular species present for the model of Fig. 1 with various values of $K_{BB}$ and $K_t$ at $\bar{Y} = 0$ and $\bar{Y}$ approximately 0.5. In all cases $K_{AB} = K_{AA} = 1$ and format is similar to Fig. 4; $n$ values are Hill coefficients for case indicated. Molecular species present at given $\bar{Y}$ value do not depend on $K_s$.

Many combinations of subunit interactions ($K_{BB}$, $K_{AB}$, etc.), conformational transitions ($K_t$), and ligand affinity ($K_s$) lead to the simplest models, i.e., the symmetry model or the simplest sequential model. Consequently there is a reasonable probability that the properties of some proteins can be explained using simplifying postulates such as "only subunits containing bound ligand have undergone conformation changes" or "all subunits change conformation simultaneously." (*b*) Many combinations of these same parameters will not fit the assumptions of the simplest models and the more complex general model will be needed in these cases to explain the molecular species present. This is particularly true if nonidentical subunits are present. (*c*) Curves of equal cooperativity can be obtained in a

variety of ways with different types of subunit interactions. (*d*) Cooperativity is obtained by a *change* in subunit interactions. For convenience, the standard state is defined for the conformations of the nonliganded protein ($K_{AA} = 1$), and cooperativity occurs when the new conformations have different interactions ($K_{AB}$ and/or $K_{BB}$ not equal to 1). The absolute value of $K_{AA}$ which reflects the true strength of interaction of the subunits in the absence of ligand does not enter directly into the calculations on cooperativity. It may, however, affect the manner in which this cooperativity is expressed. For example, a change to conformation $B$ with reduced subunit interactions ($K_{AB} < 1$) may be just sufficient to cause dissociation if the absolute value of $K_{AA}$ is low but may only lead to changes within the polymer if the absolute magnitude of $K_{AA}$ is high.

Since it is clear that more than one model can give identical saturation curves,[1] additional measurements will be needed to determine the subunit changes occurring in a particular case. The fast reaction techniques of Kirschner and Eigen,[8] the spin resonance probes of Ogawa and McConnell,[9] the crystallographic-kinetic measurements of Theorell,[10] and various other tools are being developed to aid in this problem. The comparison of $\bar{B}$ (the fraction of subunits in the $B$ conformation) with $\bar{Y}$ (the fraction of sites occupied by ligand) seems to be a particularly attractive test. Such a comparison was indeed applied by Ogawa and McConnell to hemoglobin in which $\bar{Y}$ was determined by spectral methods and $\bar{B}$ by their spin

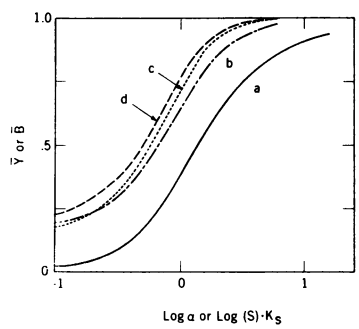



FIG. 6.—Comparison of $\bar{B}$ vs. ($S$) for cases with identical $\bar{Y}$ vs. ($S$) curves. Saturation curve [$\bar{Y}$ vs. ($S$)] for all cases is curve (*a*). $\bar{B}$ vs. ($S$) curves are: (*b*) general sequential model (Fig. 1) when $K_t = 0.05$, $K_{AB} = 1$, and $K_{BB} = 10$; (*c*) same except $K_t = 0.001$ and $K_{BB} = 100$, and (*d*) Monod, Wyman, and Changeux model for $L = 5$ and $c = 0$. The simplest sequential model will give identical $\bar{Y}$ vs. ($S$) and $\bar{B}$ vs. ($S$) curves (curve *a*).

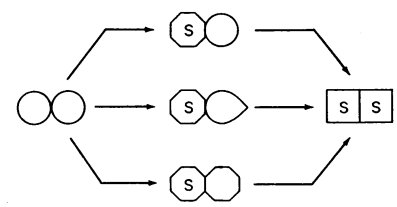



FIG. 7.—Schematic illustration of possible conformational changes when the constraints of the unsaturated subunit prevent the induction of the $B$ conformation in the saturated subunit. The hexagon designates this intermediate conformation, and, as in Fig. 3, the adjacent subunits to which no ligand is bound can be unchanged, partially changed, or completely changed to this new state.

resonance probes. The curves in Figure 6 are a good illustration of this. It is seen that the $\bar{B}$ versus ($S$) curves are not identical for several cases, all of which have identical $\bar{Y}$ versus ($S$) curves. Thus the $\bar{B}$ curves for a very large energetic change between conformations ($K_t = 10^{-4}$) and moderately weak changes in subunit interactions ($K_{BB} = 2.1$) are clearly distinguishable from the case of $K_{BB} = 1000$, $K_t = 6 \times 10^{-4}$, assuming the more general sequential model of Figure 1. How-

ever, it is noted that $\bar{B}$ curves for three of the cases are very similar, and therefore, highly accurate data would be needed to distinguish between them. From analogous computations on a number of systems, it can be concluded that $\bar{B}$ versus ($S$) curves can frequently distinguish between models which give identical $\bar{Y}$ versus ($S$) curves, but in other cases the $\bar{B}$ curves will be sufficiently similar to make differentiation difficult.

It is not, *a priori*, necessary to limit a sequential model to exclusive binding or to two conformations. In Figure 3, for example, three intermediate situations are shown in which the subunit to which ligand is bound is immediately changed to the $B$ conformation, whereas the adjacent subunit could be unchanged, partially distorted, or completely isomerized. It is also possible, however, that the subunit containing ligand is altered to some intermediate conformation when the adjacent subunit is unsaturated (Fig. 7). Presumably, in these cases, the constraints imposed by the unoccupied subunit prevent the first ligand from inducing a full change to conformation $B$. Indeed, evidence for such intermediate states has been obtained for rabbit muscle glyceraldehyde-phosphate dehydrogenase which shows negative cooperativity.[3] The necessity of considering additional conformations requires the introduction of additional terms such as $K_{AC}$, $K_{CC}$, $K_{CD}$, etc. This appreciably increases the difficulty of determining a unique model, but fortunately the theoretical concepts and mathematical tools for describing the protein action are the same as those developed for the simpler sequential models.

The results of calculations of the type presented here and others of an analogous nature lead to the following conclusions: (*a*) The assumption of a localized deformation of protein structure induced by ligand which may be transmitted with varying efficiencies to other parts of the protein molecule can be expressed mathematically in a simple way which relates equilibrium constants and free energies to protein structure. This mathematical approach is readily applied to a wide variety of models. (*b*) Under many circumstances the more general sequential model reduces to two quite simple extremes, i.e., the symmetry model of Monod, Wyman, and Changeux (in which all the subunits are present in the same conformation), or the simplest sequential model (in which only the subunit with bound ligand undergoes a conformation change). From the nature of the interactions it would not be expected that all enzymes will reduce to such simple cases, but the finding that they represent good first approximations in many cases should be helpful in a qualitative way and as a first step in the analysis of more complex two ligand situations (cf. refs. 2 and 4). (*c*) Theoretical calculations are desirable in the design of experiments, since it is seen that added experiments of the $\bar{B}$ versus ($S$) or $\bar{Y}_B$ versus ($S$) type will distinguish between mechanisms in some cases and will not in others. The mathematical approaches are useful in calculating the molecular species expected to be present in a given situation, a prediction which can be checked with some of the developing physical and chemical probes. (*d*) The demonstration that (1) the general model of Figure 1 can reduce in effect to the simpler models, (2) that curves of equal cooperativity can be obtained in diverse ways, and (3) that conformations intermediate between initial and final states can exist makes it seem improbable that any highly simple model will apply to all proteins. On the contrary, it suggests that evolution may have operated like a giant computer, varying subunit interactions until a protein with the appropriate

kinetic and control properties evolved. The basic change would appear to be the conformation change in the monomer or subunit. In a polymeric protein a conformation change in one monomer causes a perturbation of adjacent monomers with consequent amplification of the response to a given ligand. Over evolutionary time the structures with the most favorable subunit interactions were selected. The specific models which fit individual proteins may then reflect individual variations on a general theme, i.e., the utilization of changes in subunit interactions to tailor the kinetic and binding properties of proteins for the advantage of the organism.

*Summary.*—Calculations were performed on "sequential" models of conformational changes in proteins to explain cooperative effects. It is seen that certain values of subunit interactions and conformation energy transitions often lead to molecular species predicted by simplified models for cooperative effects, whereas other values require the more generalized models. The calculations also reveal that distinction between various models having identical saturation curves, i.e., $\bar{Y}$ or $S$, can in some cases be achieved by analyzing other measurable properties of the protein, e.g., fraction of changed subunit conformations $\bar{B}$, as a function of ligand concentration. The computations suggest that variations in subunit interactions provide a potent tool for the selection over evolutionary time of proteins with favorable kinetic and control properties.

* Support of this research by the National Institutes of Health and the National Science Foundation is gratefully acknowledged.

† Presented under the title "Protein Flexibility in Enzyme Action and Enzyme Control."

[1] Koshland, D. E., Jr., G. Nemethy, and D. Filmer, *Biochemistry*, **5**, 365 (1966).

[2] Kirtley, M. E., and D. E. Koshland, Jr., *J. Biol. Chem.*, **242**, 4192 (1967); Koshland, D. E., Jr., and M. E. Kirtley, *Natl. Cancer Inst. Monograph*, vol. 27, in press.

[3] Koshland, D. E., Jr., A. Conway, and M. E. Kirtley, in "Regulation of Enzyme Activity and Allosteric Interactions," *Federation of European Biochem. Soc. Symp., 4th, Oslo, June 1967* (Oslo Univ. Press, in press), vol. 1.

[4] Monod, J., J. Wyman, and J. Changeux, *J. Mol. Biol.*, **12**, 88 (1965).

[5] The same type of quantitative approach can also be applied to situations involving more than two conformations (cf. refs. 2 and 3).

[6] Hill, A. V., *J. Physiol. (London)*, **40**, 190 (1910).

[7] $B_4S_i$ Species will be used to designate all species in which all four subunits are in the $B$ conformation regardless of the number of $S$ molecules bound, i.e., $B_4$, $B_4S$, $B_4S_2$, etc. The same will be true for $A_4S_i$ species, $A_3BS_i$ species, etc.

[8] Kirschner, K., M. Eigen, R. Bittman, and B. Voigt, these PROCEEDINGS, **56**, 1661 (1966).

[9] Ogawa, S., and H. M. McConnell, these PROCEEDINGS, **58**, 19 (1967).

[10] Theorell, H., A. Ehrenberg, and C. de Zalenski, *Biochem. Biophys. Res. Commun.*, **27**, 309 (1967).