Analysis of Macromolecule–Ligand Binding by Determination of Stepwise Equilibrium Constants*

John E. Fletcher,† Arthur A. Spector,‡ and J. Douglas Ashbrook

ABSTRACT: A theoretical analysis of the stepwise equilibrium model for macromolecule–ligand interaction is presented. Conditions for the mathematical equivalence of this model and the Scatchard model are examined. The constants for the stepwise equilibrium model can be approximated from the parameters of the Scatchard model. This provides a practical means for fitting binding data to the stepwise equilibrium model.

The data initially are analyzed in terms of an appro priate Scatchard model, and the constants for the stepwise equilibrium model are estimated from the results of this analysis. These estimates then are refined by a model-fitting procedure which analyzes the data directly in terms of stepwise equilibrium constants. A detailed application involving the binding of palmitic acid to human serum albumin is presented as an illustrative example.

Ligand binding to a macromolecule that contains no subunits can, in general, be described by stepwise equilibrium reactions (Klotz, 1953; Edsall and Wyman, 1958; Tanford, 1961), where

$$P + A \underset{}{\overset{K_1}{\rightleftarrows}} PA = P_1 \quad K_1 = \frac{[P_1]}{[P][A]}$$

$$P_1 + A \overset{K_2}{\rightleftarrows} P_1A = P_2 \quad K_2 = \frac{[P_2]}{[P_1][A]}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$P_{n-1} + A \overset{K_n}{\rightleftarrows} P_{n-1}A = P_n \quad K_n = \frac{[P_n]}{[P_{n-1}][A]}$$

The binding model derived from these equations (Klotz, 1946; Klotz *et al.*, 1946) is

$$\bar{\nu} = \frac{K_1[A] + 2(K_1[A])(K_2[A]) + \ldots + n(K_1[A])(K_2[A])\ldots(K_n[A])}{1 + K_1[A] + (K_1[A])(K_2[A]) + \ldots + (K_1[A])(K_2[A])\ldots(K_n[A])} \quad (1)$$

where $\bar{\nu}$ is the molar ratio of bound ligand to macromolecule and [A] is the concentration of unbound ligand. This model can incorporate many of the newer concepts in macromolecule–ligand interaction, including (a) cooperativity in binding, (b) the creation of new sites in the process of binding, or (c) the possibility that a molecule of bound ligand may become incorporated into a binding site for additional ligand. Therefore, from a theoretical viewpoint, this model is the most satisfactory for analysis of macromolecule–ligand binding data.

A set of equilibrium data can be represented mathematically as a function: $\bar{\nu} = f([A], K_1, \ldots, K_n)$, in which the unknown parameters are $K_1, K_2, \ldots, K_n$. Estimating these parameters involves minimizing an error function $E(\bar{\nu},\hat{\nu})$, where $\hat{\nu}$ is the measured experimental value and $\bar{\nu}$ is the corresponding value obtained after fitting the model. A commonly used procedure involves the weighted least-squares function

$$E(\bar{\nu},\hat{\nu}) = \sum_{i=1}^{l} \omega_i(\bar{\nu}_i - \hat{\nu}_i)^2$$

where ω_i is the relative weight of measurement i, $\hat{\nu}_i$ is the measured value at concentration $[A_i]$, $\bar{\nu}_i$ is the calculated value from the model at concentration $[A_i]$, and l is the total number of observations.

When n is greater than 2 in eq 1, fitting this equation by means of a nonlinear curve-fitting procedure requires reliable starting estimates for the K_i's. Good starting estimates are difficult to predict without prior knowledge of the actual reactions (Tanford, 1961). Such estimates are essential in order to fit eq 1 directly to binding data by means of any curve-fitting procedure. Scatchard (1949) avoided this difficulty by grouping binding sites into classes and calculating an average association constant for each class of sites. For example, if ten binding sites can be grouped into two classes, then only two parameters, N_1k_1 and N_2k_2, need be calculated (Scatchard, 1949). The Scatchard model obtains $\bar{\nu}$ from the formulation

$$\bar{\nu} = \sum_{i=1}^{m} N_i \frac{k_i[A]}{1 + k_i[A]} \quad (2)$$

where m is the number of classes of binding sites, N_i is the number of binding sites in class i, and k_i is the average association constant for that class. A graphical method for obtaining the values of N_i and k_i from the product N_ik_i is available (Scatchard, 1949). In addition, there is a procedure for com-

* From the Laboratory of Applied Studies, Division of Computer Research and Technology, National Institutes of Health, Public Health Service, Department of Health, Education, and Welfare, Bethesda, Maryland 20014 (J. E. F. and J. D. A.), and the Departments of Internal Medicine and Biochemistry, University of Iowa, Iowa City, Iowa 52240 (A. A. S.). *Received July 16, 1970.*

† To whom to address correspondence.

‡ Recipient of a Research Career Development award from the National Institutes of Health (K4-HE-20,338).

puter analysis employing the Scatchard model (Fletcher and Spector, 1968). The difficulty with the Scatchard model is that a number of the fundamental assumptions used in its formulation are open to question, at least in applications involving the binding of large organic ligands to proteins. For example, the Scatchard analysis assumes that all of the binding sites exist initially and compete independently for available ligand. This is inconsistent with data indicating that binding of certain organic ligands to albumin produces conformational changes in which binding sites are altered or formed (Karush and Sonnenberg, 1949; Karush, 1952; Klotz and Ayers, 1953; Markus and Karush, 1958; Lovrien, 1963). Therefore, we thought it important to devise a practical method for applying the more general stepwise equilibrium model to binding data.

This communication describes both the theoretical and practical aspects of analysis by means of the stepwise equilibrium model. First, the mathematical equivalence of this model and the Scatchard model is examined. Second, accurate estimates of the stepwise equilibrium constants are obtained using the computer algorithm mentioned above. Finally, methods for further refinement of the stepwise equilibrium analysis are presented.

Theoretical Development

Assume the values of $\bar{\nu}$ and [A] can be represented by eq 1, written in the form

$$\bar{\nu} = \frac{C_1[\mathrm{A}] + 2C_2[\mathrm{A}]^2 + \ldots + nC_n[A]^n}{1 + C_1[\mathrm{A}] + \ldots + C_n[\mathrm{A}]^n} \tag{3}$$

where $C_1 = K_1$, $C_2 = K_1K_2, \ldots, C_n = K_1 \ldots K_n$, and $C_i > 0$ for all $i = 1, 2, \ldots, n$. By denoting the polynomial denominator of eq 3 by

$$1 + C_1[\mathrm{A}] + C_2[\mathrm{A}]^2 + \ldots + C_n[\mathrm{A}]^n = P_n[\mathrm{A}] \tag{4}$$

we may rewrite eq 3 as

$$\bar{\nu} = \frac{[\mathrm{A}]}{P_n[\mathrm{A}]}\left[\frac{\mathrm{d}}{\mathrm{d}[\mathrm{A}]}\{P_n[\mathrm{A}]\}\right] \tag{5}$$

The polynomial $P_n[\mathrm{A}]$ has n roots (some roots may be repeated and others may be complex), and may be expressed in factored form as

$$P_n[\mathrm{A}] = C_n([\mathrm{A}] - R_1)([\mathrm{A}] - R_2)\ldots([\mathrm{A}] - R_n) \tag{6}$$

where the R_i are the roots of $P_n[\mathrm{A}] = 0$. Cases in which the roots of polynomial eq 4 occur in different ways are now considered.

Case I. Consider first the case where all n roots are real and distinct. Introducing the product notation

$$([\mathrm{A}] - R_1)\ldots([\mathrm{A}] - R_n) = \prod_{i=1}^{n}([\mathrm{A}] - R_i) \tag{7}$$

we obtain from eq 5

$$\bar{\nu}/[\mathrm{A}] = \frac{\sum_{i=1}^{n}\prod_{j\neq i}([\mathrm{A}] - R_j)}{\prod_{j=1}^{n}([\mathrm{A}] - R_j)} \tag{8}$$

Applying the theory of partial fractions, this expression may be expanded in the form (see Appendix B)

$$\bar{\nu} = \frac{[\mathrm{A}]}{[\mathrm{A}] - R_1} + \frac{[\mathrm{A}]}{[\mathrm{A}] - R_2} + \ldots + \frac{[\mathrm{A}]}{[\mathrm{A}] - R_n} \tag{9}$$

Since the polynomial $P_n[\mathrm{A}]$ has all positive coefficients, every real root, R_i, must be negative. Therefore, $k_i = -1/R_i$ is positive and making this substitution in eq 9 we obtain

$$\bar{\nu} = \sum_{i=1}^{n}\frac{k_i[\mathrm{A}]}{1 + k_i[\mathrm{A}]} \tag{10}$$

Note that eq 10 is precisely the form of eq 2, the Scatchard model, with each N_i equal to one. A similar result was obtained by von Muralt (1930) who applied a model like that of eq 1 to determine the titration constants of a multivalent substance. We conclude that, if eq 1 can be fitted to a given set of binding data and $P_n[\mathrm{A}]$ *has n distinct real roots*, then the Scatchard model is mathematically equivalent to eq 1. Furthermore, the roots R_i of $P_n[\mathrm{A}] = 0$ are the negative reciprocals of the association constants, k_i, for the Scatchard model.

Case II. Consider now the case where eq 1 represents the data, but $P_n[\mathrm{A}] = 0$ has *repeated real roots*. Then

$$P_n[\mathrm{A}] = \prod_{i=1}^{m}([\mathrm{A}] - R_i)^p \tag{11}$$

where m is the number of distinct roots and p_i is the multiplicity of root R_i. Applying the partial fractions expansion method (see Appendix B) and expanding in the same manner as case I, eq 3 becomes

$$\bar{\nu} = \sum_{i=1}^{m}p_i\frac{[\mathrm{A}]}{([\mathrm{A}] - R_i)} \tag{12}$$

Again, setting $k_i = -1/R_i$, we obtain the Scatchard form

$$\bar{\nu} = \sum_{i=1}^{m}p_i\frac{k_i[\mathrm{A}]}{1 + k_i[\mathrm{A}]} \tag{13}$$

where p_i is an *integer* greater than or equal to one. This form is generally interpreted biologically to imply m classes of binding sites, p_i sites in each class, and an average association constant k_i representative of the sites in that class.

Case III. The final case to consider is when $P_n[\mathrm{A}] = 0$ has complex as well as real roots. Such complex roots must always occur in conjugate pairs and the terms corresponding to such roots are pure quadratic factors of $P_n[\mathrm{A}]$. They have the form

$[A]^2 + 2\alpha[A] + \beta$, where α and β are real numbers and β is positive. The polynomial $P_n[A]$ thus has the factored form

$$P_n[A] = C_n\left[\prod_{j=1}^{k}([A]^2 + 2\alpha_j[A] + \beta_j)\right]\prod_{i=1}^{n-2k}([A] - R_i) \quad (14)$$

and the partial fractions expansion for eq 1 becomes

$$\bar{\nu} = \sum_{j=1}^{k}\frac{2[A]([A] + \alpha_j)}{([A]^2 + 2\alpha_j[A] + \beta_j)} + \sum_{j=1}^{n-2k}\frac{k_j[A]}{1 + k_j[A]} \quad (15)$$

In the second sum, the k_j's may not all be distinct (*i.e.*, some repeated roots may occur). This form has no biological interpretation as a Scatchard-like binding model because the quadratic terms cannot represent distinct binding sites on the macromolecule. The occurrence of such quadratic factors is a mathematical reflection of the fact that the sequential binding reactions are not independent. However, the absence of such quadratic factors does *not* imply independence of the binding equilibria, but simply that the models 1 and 2 are mathematically indistinguishable.

Relationship of the Stepwise Equilibrium Constants to the Scatchard Association Constants

We have shown that if eq 1 represents a set of macromolecule–ligand binding data, and the polynomial $P_n[A]$ has real roots, then model 1 is mathematically equivalent to model 2. Conversely, suppose the Scatchard model 2 represents the binding data, and we seek the corresponding macroscopic constants C_i, $i = 1, \ldots, n$ in model 1. Suppose only one site exists. Then model 2 is

$$\bar{\nu} = \frac{k_1[A]}{1 + k_1[A]} \quad (16)$$

exactly the same form as eq 1. We may then identify $n = 1$, and $k_1 = K_1$. If there are two classes of sites of differing association constants, but each with only one site, model 2 is

$$\bar{\nu} = \frac{k_1[A]}{1 + k_1[A]} + \frac{k_2[A]}{1 + k_2[A]} \quad (17)$$

Combining terms and equating coefficients with model 1, $n = 2$, $K_1 = k_1 + k_2$, and $K_2 = k_1k_2/(k_1 + k_2)$. In the general case

$$\bar{\nu} = \sum_{i=1}^{m}N_i\frac{k_i[A]}{1 + k_i[A]} \quad (18)$$

will represent the binding data. The total number of binding sites, n, is the sum of the number in each class, or

$$n = N_1 + N_2 + \ldots + N_m \quad (19)$$

If the Scatchard model is rewritten as n consecutive terms and the association constants are relabeled in consecutive order, we have

$$\bar{\nu} = \sum_{j=1}^{n}\frac{\lambda_j[A]}{1 + \lambda_j[A]} \quad (20)$$

where

$$\begin{aligned}\lambda_j &= k_1,\ 1 \le j \le N_1\\ &= k_2,\ N_1 + 1 \le j \le N_2\\ &\ \ \vdots\quad\ \vdots\qquad\ \vdots\\ &= k_m,\ N_{m-1} + 1 \le j \le N_m\end{aligned}$$

Note that in this formulation each sum of the form

$$\sum_{N_{i-1}+1}^{N_i}\frac{\lambda_j[A]}{1 + \lambda_j[A]} = N_i\frac{k_i[A]}{1 + k_i[A]};\ i = 1, \ldots, m \quad (21)$$

represents a single *class* of sites. Then, in a manner analogous to the simple cases given previously we can find

$$\begin{aligned}C_1 &= \sum_{i_1=1}^{n}\lambda_{i_1}\\ C_2 &= \sum_{i_1=1}^{n-1}\sum_{i_2=i_1+1}^{n}\lambda_{i_1}\lambda_{i_2}\\ C_3 &= \sum_{i_1=1}^{n-2}\sum_{i_2=i_1+1}^{n-1}\sum_{i_3=i_2+1}^{n}\lambda_{i_1}\lambda_{i_2}\lambda_{i_3}\\ &\ \ \vdots\\ C_p &= \sum_{i_1=1}^{n-p+1}\cdots\sum_{i_p=i_{p-1}+1}^{n}\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_p},\ 1 \le p \le n\end{aligned} \quad (22)$$

(See Appendix A.) Equations 22 determine the macroscopic constants for eq 3. The equilibrium constant, K_1, is given by C_1 and in general

$$K_{p+1} = \frac{C_{p+1}}{C_p},\ p = 1, 2, \ldots, n - 1 \quad (23)$$

Therefore, all the equilibrium constants for the stepwise equilibrium model 1 can be calculated whenever the association constants for the Scatchard model 2 are known. Special cases of eq 22 have been given earlier (Simms, 1926; Weber, 1927; von Muralt, 1930).

Experimental Section

Palmitic acid-1-^{14}C binding at 23° to crystalline human serum albumin from which endogenous free fatty acid was removed by the method of Chen (1967) was measured by equilibrium partition (Goodman, 1958). Heptane served as the organic phase, and the aqueous solution contained 0.2 mM albumin, 116 mM NaCl, 0.49 mM KCl, 0.12 mM $MgSO_4$, and 16 mM sodium phosphate (pH 7.4) (Spector *et al.*, 1969). The data consisting of 28 values of the average molar ratio of palmitate bound to albumin, $\bar{\nu}$, at a given unbound palmitate concentration, [A], are the same as those reported previously (Spector *et al.*, 1969). The range of $\bar{\nu}$ values was 0.24–6.49, and the range of [A] values was 4.64×10^{-9} to 1.43×10^{-5} M. As shown in Table I, these data were fitted to 4 different Scatchard binding models. In two cases, computer A and B, the computerized algorithm of Fletcher and Spector (1968)

TABLE I: Scatchard Models for Analysis of Palmitic Acid Binding to Human Serum Albumin.[a]

Scatchard Model	Parameters						Root-Mean-Square Error
	N_1	$k_1 \times 10^{-7}$	N_2	$k_2 \times 10^{-5}$	N_3	$k_3 \times 10^{-2}$	
Computer A	3	2.02	3	5.07	63	5.00	0.37
Computer B	2	5.16	4	9.20	20	8.70	0.39
Graphical A	2	6.00	5	2.00	20	5.00	0.71
Graphical B	3	3.00	18	3.00			6.10

[a] 28 data points were obtained by incubation of palmitate-1-^{14}C with free fatty acid extracted human serum albumin at 23°. The root-mean-square error is the square root of the average square difference between the computed data and experimental data. This parameter is a measure of the average deviation of the data from the fitted curve.

TABLE II: Effect of Number of Parameters on Fitting the Stepwise Equilibrium Model.

Parameter	Estimates	Models			
		$n = 7$	$n = 8$	$n = 9$	$n = 10$
$K_1 \times 10^{-7}$	1.07	1.09	1.11	1.09	1.09
$K_2 \times 10^{-6}$	2.85	1.47	1.40	1.42	1.42
$K_3 \times 10^{-6}$	0.34	0.95	1.01	1.00	1.00
$K_4 \times 10^{-5}$	1.35	1.18	1.08	1.09	1.09
$K_5 \times 10^{-4}$	6.09	2.82	2.93	3.06	3.05
$K_6 \times 10^{-4}$	2.33	2.18	2.65	2.57	2.57
$K_7 \times 10^{-3}$	0.77	1.08	5.86	5.80	5.80
$K_8 \times 10^{-3}$	0.35		2.28	2.08	2.08
$K_9 \times 10^{-2}$	2.02			7.96	7.96
$K_{10} \times 10^{-2}$	1.30				1.30
Root-mean-square error		0.33	0.33	0.33	0.33
Sum of squares		3.124	3.117	3.117	3.117
Condition number of fit		8.0×10^2	5.1×10^3	7.2×10^4	2.5×10^6
Smallest eigenvalue		3.8×10^{-3}	6.9×10^{-4}	5.4×10^{-5}	1.8×10^{-6}

was employed to fit the data to Scatchard models containing three independent classes of binding sites. The models chosen, containing either 3, 3, 63, or 2, 4, 20 sites, represent best fits for these data. This is indicated by the fact that the root-mean-square errors, 0.37 and 0.39, respectively, were the lowest obtained for a Scatchard model. In the other two cases, graphical A and B, the data were fitted according to the graphical procedure (Scatchard, 1949). The three-term model that appeared to fit best by graphical approximation, graphical A, contained 2, 5, and 20 sites. A root-mean-square error of 0.71 was calculated using this model, indicating that it fit the data less well than either computer A or B. The best two-term model that could be estimated graphically, containing 3 and 18 sites, fitted the data least well as indicated by a root-mean-square error of 6.1. Values for the apparent association constants, k_i, are listed in Table I for each of these four Scatchard models.

Preliminary estimates of the first ten stepwise equilibrium constants for palmitate binding to human serum albumin were calculated (Table II). This was done by employing eq 22 and 23 and the parameters obtained with the model computer B. These estimates are listed in Table I. Only the first 15 terms of eq 22 were used for each calculation because the remaining terms have a negligibly small effect on the magnitude of the calculated parameters in the present application. Using these estimates as starting values, stepwise equilibrium models containing seven to ten parameters were each fitted directly to the binding data by means of a least-squares model-fitting procedure (Fletcher and Shrager, 1968). The refined equilibrium constants for each case are listed in Table II. In every case, the refined values for the first six parameters agree very closely. However, the model containing eight parameters was judged to be most appropriate for this set of data. Although the root-mean-square errors were the same in each case, there was a slight improvement in the sum of squares when models containing more than seven parameters were used. However, the sum of squares did not improve further when more than eight parameters were employed. The

TABLE III: Estimated and Refined Stepwise Equilibrium Constants for Palmitic Acid Binding to Human Serum Albumin.

Equilibrium Constants	Estimated				Refined			
	Computer		Graphical		Computer		Graphical	
	A	B	A	B	A	B	A	B
$K_1 \times 10^{-7}$	0.62	1.07	1.21	0.94	1.11	1.11	1.11	1.09
$K_2 \times 10^{-6}$	2.12	2.85	3.08	3.26	1.38	1.40	1.39	1.39
$K_3 \times 10^{-6}$	0.77	0.34	0.98	1.25	1.02	1.01	1.03	1.02
$K_4 \times 10^{-5}$	1.33	1.35	0.40	3.31	1.09	1.08	1.11	1.07
$K_5 \times 10^{-4}$	4.86	6.09	2.01	16.99	2.90	2.93	2.82	3.13
$K_6 \times 10^{-4}$	1.68	2.33	1.02	10.62	2.59	2.65	2.73	2.36
$K_7 \times 10^{-3}$	0.44	0.77	4.23	72.71	6.40	5.86	6.13	7.03
$K_8 \times 10^{-3}$	0.20	0.35	0.38	52.18	1.92	2.28	1.95	1.58

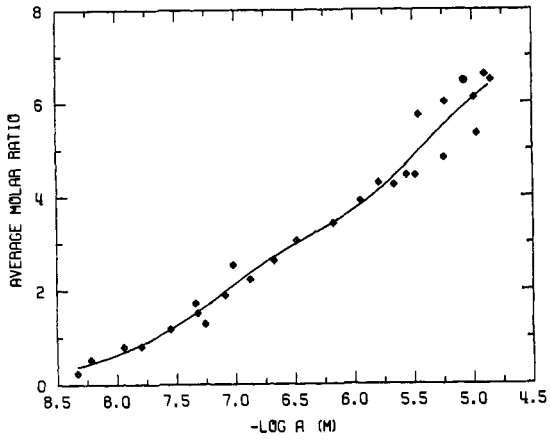

FIGURE 1: Graph for the eight-parameter stepwise equilibrium model.

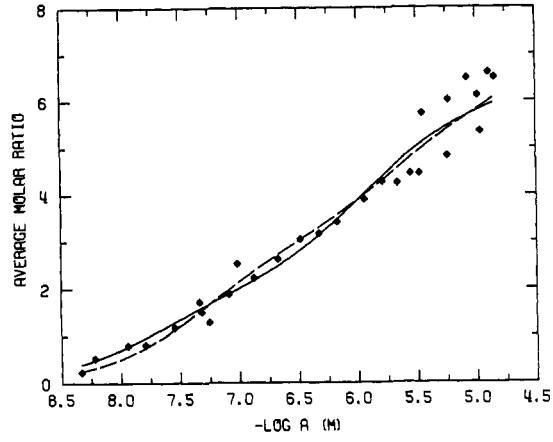

FIGURE 2: Graph for the Scatchard models. The solid curve represents the 2, 4, 20 model; the dashed curve represents the 3, 3, 63 model.

spectral condition number[1] increased by a factor of ten when the number of parameters was raised from seven to eight, and eight to nine. In contrast, a 100-fold increase in the condition number occurred when the number of parameters was raised from nine to ten. The smallest eigenvalue decreased monotonically by a factor of ten as the number of parameters was increased. This suggests an increasing inability of the model-fitting procedure to resolve all the parameters from the given range of data. The model usually selected as best-fitting contains the fewest number of parameters with a relatively small condition number and a smallest eigenvalue $>1 \times 10^{-5}$. These criteria are met most satisfactorily by the model containing eight parameters.

Table III contains estimates for the first eight parameters of the stepwise equilibrium model calculated from each of the four Scatchard models listed in Table I. The significant digits for the estimate of each parameter differ somewhat using the respective Scatchard models. However, there is general agreement in the order of magnitude of each estimated parameter, at least for K_1 through K_6. The estimates obtained with each of the four Scatchard models were refined by the least-squares-fitting procedure. The refined values in each case exhibit almost complete agreement in spite of the somewhat different values used as starting estimates. Identical sums of squares (3.12) and root-mean-square errors (0.33) were obtained in each case. Hence, we conclude that as long as the Scatchard parameters reasonably reproduce the experimental data, they are adequate for determining starting estimates for directly fitting the stepwise equilibrium model.

A plot of the resulting fit of the data for the 8 parameter stepwise equilibrium model is shown in Figure 1. Similar plots for the 2, 4, 20, and 3, 3, 63 Scatchard models are shown in Figure 2.

It is evident from Figure 2 that both Scatchard models are adequate representations of the experimental data. Further comparison of Figures 1 and 2 demonstrates that the stepwise equilibrium model produces a better fit than either of the Scatchard models.

[1] The spectral condition number is the largest eigenvalue divided by the smallest eigenvalue. It is a measure of the reliability of matrix inversion process used in obtaining the parameter estimates. Best results are obtained when the condition number is relatively low.

Discussion

We have shown that the Scatchard model 2, known to be an approximation of the more general stepwise equilibrium model 1, is also an exact mathematical equivalent to eq 1 whenever the polynomial equation 4 has only real roots. This fact provides a means for simplifying the fitting of binding data to the stepwise equilibrium model. This simplification is of particular value when the macromolecule contains more than two binding sites. Under these conditions, reliable starting estimates for the stepwise equilibrium constants previously were difficult to predict. The present analysis shows that starting estimates can be generated by making a preliminary analysis in terms of a Scatchard model. A computerized algorithm for obtaining the best-fitting Scatchard model is available for this purpose. However, as demonstrated in Table III, a Scatchard model obtained by the relatively crude method of graphical estimation (Scatchard, 1949) also is adequate for generating starting parameter estimates. In order to obtain more precise values for the stepwise equilibrium constants, it is necessary to refine the starting estimates. This can be done by fitting the stepwise equilibrium model directly to the data by means of a least-squares procedure (Fletcher and Shrager, 1968).

The theoretical analysis shows that the Scatchard model can be fitted to data for which its parameters cannot be interpreted as binding constants. This is true because the Scatchard constants are approximating the negative reciprocals of the roots of eq 4 and the number of times that each root is repeated. Therefore, the Scatchard model can adequately fit data in spite of the fact that the model may have no true interpretation as a binding model. Thus we conclude that the mechanism of the binding process need not necessarily satisfy the Scatchard assumptions even though the data are adequately fitted by a Scatchard model.

The only case in which the stepwise equilibrium model is not mathematically equivalent to the Scatchard model is when eq 4 has complex roots. In most applications, data of sufficient range and accuracy to determine the precise nature of the polynomial equation 4 cannot be obtained. Therefore, the presence or absence of complex roots cannot be established conclusively. In view of this uncertainty, the only practical approach is to proceed as if the polynomial equation 4 has only real roots and obtain starting estimates for fitting the stepwise equilibrium model as described above. This procedure can be justified because the roots of a polynomial are continuous functions of its coefficients. Hence, the estimates derived from an approximating Scatchard model provide, in either the real or complex case, estimates for the coefficients of $P_n[A]$ and, therefore, for the roots of $P_n[A]$. Of course, in the complex case, we are approximating the magnitude of the complex root. The coefficients estimated in this manner guarantee a model which "resembles" the experimental data; therefore, it is sufficiently close for refinement of the parameters to more accurate values.

Goodman (1958) analyzed the binding of palmitic acid to human serum albumin by graphical approximation in terms of a 2, 5, 20 Scatchard model. Using a computerized algorithm (Fletcher and Spector, 1968), we have obtained best fits for palmitate binding to human albumin with 3, 3, 63; 3, 3, 20; and 2, 4, 20 Scatchard models (Spector *et al.*, 1969). In each of these cases, the values for the apparent association constants were of the order of $k_1 = 10^7$, $k_2 = 10^6$, and $k_3 = 10^3$. Spectrophotometric studies of laurate binding to human serum albumin have been interpreted to confirm the presence of two classes of high-energy binding sites: a primary class of two sites, and a secondary class of five sites (Zakrzewski and Goch, 1968).

Each of the Scatchard models selected for the palmitate–human serum albumin data satisfactorily fit the experimental points (Goodman, 1958; Spector *et al.*, 1969). However, the assumptions implicit in the Scatchard model, such as the preexistence and independence of all binding sites, appear to be inconsistent with certain information concerning the effects of large organic anions on albumin (Karush, 1952; Markus and Karush, 1957, 1958; Lovrien, 1963; Reynolds *et al.*, 1968). On the other hand, the stepwise equilibrium model is sufficiently general to account for cooperativity, ligand-induced conformational change, and other phenomena that may be associated with binding. Therefore, a more realistic interpretation of the binding process probably can be obtained with this model. For example, the stepwise equilibrium constants indicate that human albumin does not contain a single class of two or three high-energy palmitate binding sites. The constant for binding of the first mole of palmitate is ten times greater than that for binding of the second mole. In contrast, the second and third association constants are almost identical, suggesting that these two binding sites may be similar. The presence of only a single very strong palmitate binding site is more consistent with interpretations made for the association of other large organic ligands with human albumin. For example, human albumin also contains a single strong binding site for thyroxine (Sterling, 1964) and coumarin (O'Reilly, 1967).

One supposed advantage of the Scatchard analysis is that the total number of macromolecule binding sites can be estimated by extrapolation even when the range of experimental data is limited. As demonstrated in Table II, this is not possible when the stepwise equilibrium model is employed. The highest value of $\bar{\nu}$ in our data was 6.49, and the binding model became ill-conditioned when more than nine parameters were used. While the Scatchard analysis may give some indication of the presence of additional binding sites, this information is of questionable utility. The values estimated for sites outside the data range, *i.e.*, n_3 and k_3, are very poorly determined (Goodman, 1958; Spector *et al.*, 1969). Indeed, Karush and Sonnenberg (1949) have questioned whether a true value of the total number of binding sites can be defined by extrapolation. Our analysis further supports this opinion because we demonstrate that the Scatchard constants do not necessarily have meaning as binding constants. This is particularly true when eq 4 has complex roots.

Appendix A

Derivation of Formulas Relating the Constants in Models 1 and 2. We assume $C_n > 0$ and rewrite the polynomial equation 4 as

$$P_n[A] = C_n\left[[A]^n + \frac{C_{n-1}}{C_n}[A]^{n-1} + \ldots + \frac{C_1}{C_n}[A] + \frac{1}{C_n}\right] \quad \text{(A1)}$$

Let $r_1, r_2, \ldots, r_n$ be the roots of the equation $P_n[A] = 0$. Then

$$P_n[A]/C_n = [([A] - r_1)([A] - r_2)\ldots([A] - r_n)] \quad (A2)$$

Multiplying this expression out and collecting terms, one obtains the usual polynomial relations

$$\frac{1}{C_n} = (-1)^n r_1 r_2 \ldots r_n$$

$$\frac{C_{n-1}}{C_n} = (-1)^{n-1}(r_1 + r_2 + \ldots r_n)$$

$$\frac{C_{n-2}}{C_n} = (-1)^{n-2}[(r_1r_2 + r_1r_3 + \ldots + r_1r_n) + \ldots + (r_{n-1}r_n)] \quad (A3)$$

$$\frac{C_{n-3}}{C_n} = (-1)^{n-3}[(r_1r_2r_3 + r_1r_3r_4 + \ldots + r_1r_{n-1}r_n) + \ldots + (r_{n-2}r_{n-1}r_n)]$$

$$\vdots \qquad\qquad \vdots$$

$$\frac{C_1}{C_n} = (-1)^1[(r_1r_2\ldots r_{n-1}) + (r_1\ldots r_{n-2}r_n) + \ldots + (r_2\ldots r_n)]$$

The real roots of $P_n[A]$ have been shown to correspond to the apparent association constants of the Scatchard model. That is, $k_i = -1/r_i$; $i = 1, 2, \ldots, n$. Making this substitution in the first equation of A3, we obtain

$$\frac{1}{C_n} = (-1)^n(-1)^n\left\{\frac{1}{k_1}\frac{1}{k_2}\cdots\frac{1}{k_n}\right\}$$

Then

$$C_n = (k_1k_2k_3\ldots k_n) \quad (A4)$$

Making the same substitutions in the remaining eq A3, and using eq A4 we obtain

$$C_1 = \sum_{i_1=1}^{n} k_{i_1}$$

$$C_2 = \sum_{i_1=1}^{n-1}\sum_{i_2=i_1+1}^{n} k_{i_1}k_{i_2} \quad (A5)$$

$$\vdots \qquad\qquad \vdots$$

$$C_p = \sum_{i_1=1}^{n-p+1}\sum_{i_2=i_1+1}^{n-p}\cdots\sum_{i_p=i_{p-1}+1}^{n} k_{i_1}k_{i_2}\ldots k_{i_p}$$

for $1 \leq p \leq n$. Therefore all the macroscopic constants for model 1 are completely determined whenever model 2 represents a given set of data.

Appendix B

Partial Fractions Expansion of a Rational Function. The application of partial fractions methods to the rational functions used here is quite simple because the numerator is the product of the derivative of the denominator and the independent variable. For this reason, we choose to expand the form

$$\bar{\nu}/[A] = \frac{\frac{d}{d[A]}[P_n[A]]}{P_n[A]} \quad (B1)$$

Using eq 6, we obtain

$$\frac{dP_n[A]}{d[A]} = \sum_{i=1}^{n}\prod_{j\neq i}([A] - R_j) \quad (B2)$$

Using eq B2 and eq 6 in eq B1 and canceling common factors in the numerator and denominator we obtain

$$\bar{\nu}/[A] = \sum_{i=1}^{n}\frac{1}{([A] - R_i)}$$

or

$$\bar{\nu} = \sum_{i=1}^{n}\frac{[A]}{([A] - R_i)} \quad (B3)$$

Note that at this point no assumptions regarding the nature of the roots were necessary.

If all roots R_i are real and distinct, eq B3 is the desired form. If a real root R_i is repeated p_i times, then we may group all equal roots in the form

$$\bar{\nu} = \sum_{j=1}^{m} p_j\frac{[A]}{[A] - R_j} \quad (B4)$$

If there are complex roots, they must occur in conjugate pairs, and when grouped as pairs

$$\frac{[A]}{[A] - R_k} + \frac{[A]}{[A] - \bar{R}_k} = \frac{2[A]^2 - [A](R_k + \bar{R}_k)}{[A]^2 - [A](R_k + \bar{R}_k) + |R_k|^2}$$

which we rewrite as

$$\frac{2[A]([A] + \alpha_k)}{[A]^2 + 2\alpha_k[A] + \beta_k}$$

Then B1 has the form

$$\bar{\nu} = \sum_{j=1}^{m}\frac{2[A]([A] + \alpha_j)}{[A]^2 + 2\alpha_j[A] + \beta_j} + \sum_{j=1}^{n-2m}\frac{[A]}{[A] - R_j} \quad (B5)$$

Equations B3, B4, and B5 therefore represent all possibilities for the model 1.

References

Chen, R. F. (1967), *J. Biol. Chem. 242*, 173.

Edsall, J. T., and Wyman, J. (1958), Biophysical Chemistry, Vol. 1, New York, N. Y., Academic, pp 623–626.

Fletcher, J. E., and Shrager, R. I. (1968), *Technol. Rep. No. 1, DCRT*, Bethesda, Md., National Institutes of Health.

Fletcher, J. E., and Spector, A. A. (1968), *Comput. Biomed. Res. 2*, 164.
Goodman, D. S. (1958), *J. Amer. Chem. Soc. 80*, 3892.
Karush, F. (1952), *J. Phys. Chem. 56*, 70.
Karush, F., and Sonnenberg, M. (1949), *J. Amer. Chem. Soc. 71*, 1369.
Klotz, I. M. (1946), *Arch. Biochem. 9*, 109.
Klotz, I. M. (1953), *Proteins 1B*, 748.
Klotz, I. M., and Ayers, J. (1953), *Discuss. Faraday Soc. 13*, 189.
Klotz, I. M., Walker, F. M., and Pivan, R. B. (1946), *J. Amer. Chem. Soc. 68*, 1486.
Lovrien, R. (1963), *J. Amer. Chem. Soc. 85*, 3677.
Markus, G., and Karush, F. (1957), *J. Amer. Chem. Soc. 79*, 3264.
Markus, G., and Karush, F. (1958), *J. Amer. Chem. Soc. 80*, 89.
O'Reilly, R. A. (1967), *J. Clin. Invest. 46*, 829.
Reynolds, J., Herbert, S., and Steinhardt, J. (1968), *Biochemistry 7*, 1357.
Scatchard, G. (1949), *Ann. N. Y. Acad. Sci. 51*, 660.
Simms, H. S. (1926), *J. Amer. Chem. Soc. 48*, 1239.
Spector, A. A., John, K., and Fletcher, J. E. (1969), *J. Lipid Res. 10*, 56.
Sterling, K. (1964), *J. Clin. Invest. 43*, 1721.
Tanford, C. (1961), Physical Chemistry of Macromolecules, New York, N. Y., Wiley, pp 526–546.
von Muralt, A. L. (1930), *J. Amer. Chem. Soc. 52*, 3518.
Weber, H. H. (1927), *Biochem. Z. 189*, 381.
Zakrzewski, K., and Goch, H. (1968), *Biochemistry 7*, 1835.

Transfer of *N*-Acetylglucosamine from Uridine Diphosphate *N*-Acetylglucosamine to 3,15α-Dihydroxyestra-1,3,5(10)-trien-17-one by Human Adult and Fetal Kidney Homogenates*

R. G. Cable, H. Jirku, and M. Levitz†

ABSTRACT: 3,15α-Dihydroxyestra-1,3,5 (10)-trien-17-one (15α-OHE_1) and its 3-sulfate are found in human body fluids principally as conjugates of *N*-acetylglucosamine. The biosynthesis of these novel conjugates was achieved. Human adult and fetal kidney homogenates fortified with uridine diphosphate *N*-acetylglucosamine converted [^3H]15α-OHE_1 into 3-hydroxy-17-oxoestra-1,3,5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside (15α-OHE_1GNAc). Hydrolysis by a pure preparation of jack bean meal β-*N*-acetylhexosaminidase yielded 15α-OHE_1, strongly suggesting the β configuration for the glycosidic bond. Reduction of 15α-OHE_1GNAc with $NaBH_4$ followed by methylation with diazomethane and subsequent acid hydrolysis yielded 3-methoxyestra-1,3,5(10)-triene-15α, 17β-diol, thereby establishing C-15 as the position of attachment of the sugar moiety. In a similar fashion 15α-OHE_1-3 sulfate was converted into 3-sulfato-17-oxoestra-1,3,-5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside. Attempts to demonstrate significant *N*-acetylglucosaminyl transferase activity in fresh human fetal liver and adult liver obtained post mortem have been unsuccessful.

3,15α-Dihydroxyestra-1,3,5(10)-trien-17-one and estra-1,3,5(10)-triene-3,15α,17β-triol have been detected in fetal tissues and maternal urine in human pregnancy (Schwers *et al.*, 1965a,b,c; Knuppen *et al.*, 1965; Lisboa *et al.*, 1967), and in the bile and urine of a nonpregnant subject given labeled estrone sulfate (Jirku and Levitz, 1969). Of particular interest is that these steroids, 15α-OHE_1[1] and 15α-OHE_2, are excreted principally in the form of *N*-acetylglucosaminides. The *N*-acetylglucosaminides are found as single conjugates and as double conjugates, the second conjugating moiety being sulfate (Jirku and Levitz, 1969, 1970). Considering that only two other laboratories have isolated and identified steroid *N*-acetylglucosaminides (Layne *et al.*, 1964; Arcos and Lieberman, 1967), these conjugates are novel. The purposes of the present study were to see whether the biosynthesis of *N*-acetylglucosaminides of 15α-hydroxyestrogens could be

* From the Department of Obstetrics and Gynecology, New York University Medical Center, New York, New York 10016. *Received June 25, 1970.* This investigation was supported by grants from the National Cancer Institute, National Institutes of Health (Grant 5 R01 CA 2071-17), and from the American Cancer Society (Grant P-206K). A preliminary report was given at the 51st Meeting of the Endocrine Society, New York, N. Y., June 1969, abstract number 252.

† Research Career Development Awardee, 5 K3-HD-18,422-08, of the National Institute of Child Health and Human Development. To whom correspondence should be addressed.

[1] Abbreviations used are: 15α-OHE_1, 3,15α-dihydroxyestra-1,3,5-(10)-trien-17-one; 15α-OHE_1-3S, 3-sulfato-15α-hydroxyestra-1,3,5(10)-trien-17-one; 15α-OHE_1GNAc, 3-hydroxy-17-oxoestra-1,3,5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside; 15α-OHE_1SGNAc, 3-sulfato-17-oxoestra-1,3,5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside; 15α-OHE_2, estra-1,3,5(10)-triene-3,15α,17β-triol; 15α-OHE_2GNAc, 3,17β-dihydroxyestra-1,3,5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside; 15α-OHE_2SGNAc, 3-sulfato-17β-hydroxyestra-1,3,5(10)-trien-15α-yl 2′-acetamido-2′-deoxy-β-D-glucopyranoside; GNAc, *N*-acetylglucosaminide; SGNAc, sulfo-*N*-acetylglucosaminide; UDPGNAc, uridine diphosphate *N*-acetylglucosamine; HBV, holdback volume; TEAMS, triethylammonium sulfate.