Journal of Chromatography, 457 (1988) 13–23
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 20 940

CO-OPERATIVE CLUSTER MODEL FOR MULTIVALENT AFFINITY INTERACTIONS INVOLVING RIGID MATRICES

ROBERT J. YON
School of Biological Sciences and Environmental Health, Thames Polytechnic, Wellington Street, London SE18 6PF (U.K.)

(First received February 23rd, 1988; revised manuscript received July 26th, 1988)

SUMMARY

A rapid-equilibrium model of affinity partitioning of multi-site proteins is described, that applies when long-range translatory movement of matrix-ligand groups is prevented by a rigid matrix. Two essential postulates are made: (1) the matrix ligands are distributed singly or in clusters within spherical bounds of the size of the protein molecule, and may be treated as a Poisson distribution; (2) due to the proximity of matrix ligands, binding of a protein molecule to a cluster is highly co-operative. Equations are derived that allow the predicted partitioning and chromatographic behaviour of this model to be examined and its properties described. Published data on the interaction of aldolase and phosphocellulose are re-interpreted in terms of the new theory.

INTRODUCTION

Early theoretical treatments of affinity chromatography were formulated on the assumption of a single interaction between protein and matrix-ligand, and were consequently applicable only to monomeric proteins or when the matrix-ligand was highly diluted[1–4]. Since then the theory has been adapted to the case of multivalency, *i.e.* where an N-site (*i.e.* presumably oligomeric) protein molecule can establish from 1 to N contacts with matrix-ligands[5–11]. Theory appropriate to oligomeric proteins is of particular importance, for several reasons, including: (1) the great majority of functional proteins, including many important enzymes, are oligomeric in structure; (2) current practice is to use highly-substituted column matrices (*e.g.* dye-ligand matrices) for protein purification, with the consequent increased likelihood of multivalent contacts; (3) since the Gibbs energy of binding is additive with respect to individual occupied sites, extremely tight binding to the matrix is possible in principle without loss in specificity; (4) quantitative affinity chromatography has been recommended for measuring protein–ligand binding constants, and it is important to ensure that the theoretical basis of such use is secure in the case of oligomeric proteins.

In relation to a protein with identical sites, most of the afore-mentioned theoretical treatments of the multivalent case assume, explicitly or implicitly, one or both of the following postulates: (1) a single microscopic equilibrium binding constant

0021-9673/88/$03.50

may be used for each successive interaction of a protein molecule with matrix-ligand; (2) all sites on the protein have access to the same concentration of matrix-ligand. Although they are valid for interactions of soluble ligand with protein, conceptual difficulties arise when these postulates are applied to ligands immobilised on rigid matrices. One difficulty arises from the uneven distribution of matrix-ligand groups at the molecular level, and their inability to diffuse. This uneven distribution and the size of the protein molecule set a limit on the number of matrix-ligand groups that can be bound by any individual protein molecule at a particular locus within the matrix. The equal-access assumption, discussed above, clearly is inappropriate for this situation. A second difficulty is the intuitive expectation that, while the probability of a first contact by a protein molecule within a static cluster of matrix-ligands will be the same as for contact with an isolated matrix-ligand, subsequent contacts within the cluster by the same molecule ought to have a higher probability due to proximity; in other words successive contacts within a cluster should show cooperativity. Of existing theoretical models, only that of Kyprianou and Yon[7] implies such co-operativity. This model, however, is deficient in assuming a single uniform type of cluster. In a recent paper, Hubble[12] discusses the different but related problem of affinity chromatography of proteins with intrinsic co-operative binding properties, *e.g.* allosteric enzymes.

A multivalency model that makes simple assumptions about the clustering of matrix-ligands, and about cooperativity in binding to the matrix, will now be described.

THEORY

Symbols

$[P_t]$	Total concentration of protein
$[P_s]$	Concentration of all soluble protein forms
R	Partitioning ratio, defined as $[P_t]/[P_s]$
[S]	Total concentration of soluble ligand
[M]	Total concentration of accessible matrix-ligand
N	Number of identical ligand-binding sites per protein molecule
r	Radius (nm) of protein molecule
K_S	Microscopic (site) association constant for the binding of soluble ligand
K_M	Microscopic (site) constant for binding of matrix-ligand
$[X_1]$, $[X_2]$,...$[X_N]$	Concentrations of clusters containing 1,2,...N matrix-ligands, respectively
K_1, K_2,...K_N	Stoichiometric (cluster) association constants for concerted binding of protein to clusters of 1,2,...N matrix-ligands, respectively

All concentrations are in mol/l. Microscopic (site) constants and stoichiometric constants are defined according to Klotz[13]. Other symbols will be defined as required.

General description

The present model is an extension of an earlier model[7] in which the concept of concerted binding of a protein to a single type of matrix-ligand cluster was implicitly

proposed. The major present innovation is the assumption that the non-diffusible matrix groups are randomly distributed, thereby enabling the concentrations of "effective clusters" within this distribution to be computed by statistical means. An effective cluster is defined as a group of accessible matrix-ligands within the volume of space (assumed spherical) equivalent to that occupied by a protein molecule. Strictly, the matrix-ligands capable of binding to a protein molecule should be a subgroup within the cluster that satisfies steric requirements imposed by the protein, *e.g.* location at the apices of a tetrahedron. To simplify the mathematics, this restriction will be relaxed in the present treatment. The implications of this relaxation will be discussed later (see Discussion).

Let X_i denote a cluster containing i accessible matrix-ligands, where i takes values from 1 to N. For any one of these cluster-types, protein–ligand complexes involving fewer than i matrix-ligands are neglected. This implies a concerted binding of all i matrix-ligands, *i.e.* a high degree of cooperativity. The justification for this assumption will be discussed later. By reasoning entirely analogous to that described by Nichol *et al.*[3] and Kyprianou and Yon[7] it may be shown that, in an equilibrium batch experiment, the protein will partition between soluble and matrix phases according to the relationship:

$$R = 1 + \sum_{i=1}^{N} \frac{K_i[X_i]}{K_i[P_t] + (1 + K_S[S])^i} \tag{1}$$

This equation may be adapted to frontal-elution chromatography by use of the relationship[10,14] $R = V/V_0$, in which V is the (variable) elution volume of the protein, given by the centroid of the advancing protein front, and V_0 is the elution volume in the absence of any interaction with the matrix, provided that V_0 can be estimated.

Eqn. 1 contains N cluster concentrations $[X_i]$ and N stoichiometric association constants K_i. Expressions for $[X_i]$ and K_i will now be derived.

The cluster-concentrations $[X_i]$

The cluster concentrations may be related to the overall concentration of matrix-ligand, [M], and the dimensions of the protein molecule, as shown next. For mathematical simplicity (see earlier discussion) the protein will be treated as a sphere, and all matrix-ligands "within" this sphere assumed to be capable of binding. This simplification allows the cluster concentrations to be calculated by assuming a Poisson distribution of matrix-ligand groups (see *e.g.* ref. 15), *i.e.* the probability $p(i)$ of finding exactly i groups in a sphere whose average content is m is given by:

$$p(i) = \frac{e^{-m} \cdot m^i}{i!} \tag{2}$$

For a protein of radius r (in nm), and a concentration [M] (in mol/l) of accessible matrix-ligands, the average number of groups contained within the equivalent sphere is readily shown to be:

$$m = 2.52[M]r^3 \tag{3}$$

where the numerical factor 2.52 arises from a combination of the formula for spherical volume (radius in nm) and Avogadro's number. It should be noted that clusters with more than N groups are in theory included in this distribution, *i.e.*, i can take all positive integral values. However, as will be shown later, higher-order clusters are negligible at the accessible matrix-ligand concentrations encountered in practice, therefore clusters with more than N groups are conveniently neglected. The concentrations of effective clusters X_i (eqn. 1) are related to the probabilities $p(i)$ as follows:

$$\frac{i[X_i]}{[M]} = \frac{p(i)}{1 - p(0)} = \frac{[\text{matrix-ligand in } i\text{-clusters}]}{[\text{total matrix ligand}]} \tag{4}$$

Combining eqns. 2 and 4 gives the cluster concentration as:

$$[X_i] = \frac{m^i[M]}{i \cdot i!(e^m - 1)} \tag{5}$$

In practice [M] rarely exceeds 0.1 mM, therefore m is of the order 0.01 for a protein of ordinary radius (say 4 nm). To a good approximation, therefore, the expansion $e^m = 1 + m$ may be used for the exponential term in eqn. 5. Together with eqn. 3 and some rearrangement, this enables eqn. 5 to be written as:

$$[X_i] = \frac{(2.52r^3)^{i-1}\,[M]^i}{i \cdot i!} \tag{6}$$

The stoichiometric constants K_i

For isolated matrix-ligand groups (*i.e.* "clusters" of 1) the stoichiometric association constant is given by

$$K_1 = NK_M \tag{7}$$

where K_M is the intrinsic (site) association constant for binding matrix-ligand. This is the constant measured by standard monovalent methods[2,3] *i.e.* when the concentration of matrix-ligand [M] is sufficiently small for bivalent and higher-order interactions to be ignored. For clusters of 2 or more, the cluster concentration $[X_i]$, and the value of K_M, govern the first contact of the N-valent protein with the cluster. Succeeding contacts within the cluster are governed by the much higher "local" concentration of matrix-groups. It is proposed, therefore, to treat each of the successive constants K_i (where $i > 1$) as the product of K_M^i, a statistical factor ${}^NC_i$ (as in the free-solution case, see ref. 13) and an "enhanced concentration factor" E, defined as

$$E = \frac{\text{apparent matrix-ligand concn. "inside" the cluster}}{\text{cluster concentration}}$$

Since i groups in a sphere of radius r (in nm) is equivalent to a concentration of $i/(2.52r^3)$ mol/l, and the cluster concentration is given by eqn. 6 above, it may be shown that

$$E = \frac{i^2 \cdot i!}{(2.52r^3\,[\mathrm{M}])^i} \tag{8}$$

Hence the association constants K_i are given by

$$K_i = {}^{N}C_i \cdot K_{\mathrm{M}}^{i} \cdot \frac{i^2 \cdot i!}{(2.52r^3\,[\mathrm{M}])^i} \tag{9}$$

where ${}^{N}C_i = N!/[(N-i)!i!]$.

Finally, substitution of eqns. 6, 7 and 9 into 1 yields, as the partitioning equation:

$$R = 1 + \frac{NK_{\mathrm{M}}[\mathrm{M}]}{NK_{\mathrm{M}}[\mathrm{P_t}] + 1 + K_{\mathrm{S}}[\mathrm{S}]} + \sum_{i=2}^{N} \frac{N!i(2.52r^3)^{i-1}(K_{\mathrm{M}}[\mathrm{M}])^i}{i!\{N!K_{\mathrm{M}}^{i}i^2[\mathrm{P_t}] + (2.52r^3[\mathrm{M}])^i(N-i)!(1+K_{\mathrm{S}}[\mathrm{S}])^i\}} \tag{10}$$

This equation implies protein adsorption into a set of mutually-independent adsorption sites.

RESULTS

The following aspects of the predicted behaviour of the theoretical model presented above have been examined by computer modelling.

Effect of protein radius

A novel aspect of the present theory is its introduction of protein size as a determinant of partitioning behaviour through the calculation of cluster concentrations $[\mathrm{X}_i]$, and not, as is more often the case, through gel-filtration effects. This size-dependence is quite sensitive; as shown in eqn. 6, $[\mathrm{X}_i]$ varies as the $(i-1)$th power of r^3, thus a 2.1-fold increase in the radius increases $[\mathrm{X}_2]$ by 10-fold, and $[\mathrm{X}_3]$ by 100-fold. The remaining discussion will focus on a four-site protein of radius 4 nm, typical of several glycolytic and other roughly spherical enzymes *e.g.* aldolase and lactate dehydrogenase.

Predicted cluster concentrations for a typical multi-site protein

The predicted cluster distribution has been examined for a four-site protein of radius 4 nm. Fig. 1 shows the computed cluster concentrations $[\mathrm{X}_i]$ as a function of total accessible matrix-ligand concentration [M], based on use of eqn. 5 with the exponential term unchanged. It has been assumed that matrix-clusters with more than four groups (significant only at very high, and probably unrealistic, values of [M]) are "seen" by the protein as clusters of four. In Fig. 1a the absolute values of $[\mathrm{X}_i]$ are given (note log–log scales) and in (b) $[\mathrm{X}_i]$ is expressed as a fraction of all clusters.

The total concentration of immobilised ligand in an affinity matrix is usually readily measurable, and experience has shown that, using present-day immobilisation

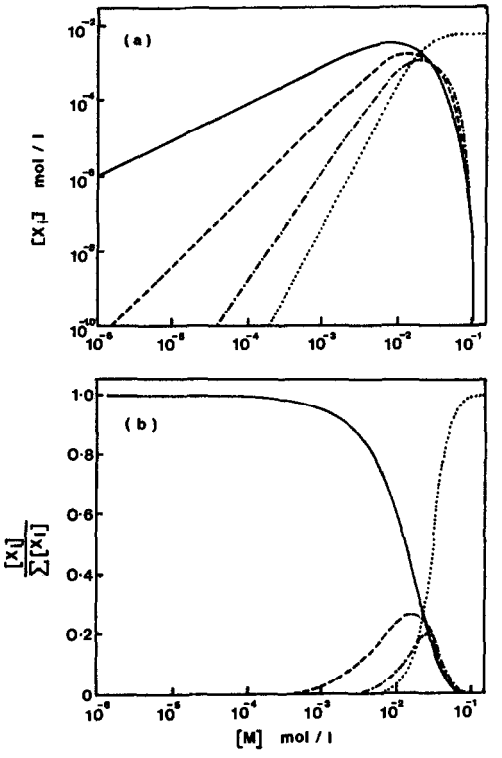

Fig. 1. Dependence of cluster concentrations $[X_i]$ on the total concentration of accessible matrix-ligand groups, [M], for a four-site protein molecule of radius 4 nm. The calculations are based on eqn. 5 in the text; clusters with more than four groups are included with four-clusters. (a) Absolute values of $[X_i]$ (note log–log scale). (b) $[X_i]$ relative to total of all clusters. Key: X_1, ———; X_2, – – –; X_3, ·–·–·–·; X_4, ····.

technology, values greater than 10^{-2} *M* are rarely, if ever, encountered. The fraction of these groups that are accessible constitutes [M]; attempts to estimate its value have invariably indicated that, at most, a few percent of the total may be involved. It seems reasonable, therefore, to take 10^{-3} *M* as a practical upper limit for [M]. Fig. 1 shows that even at the top end of this range, a protein of radius 4 nm "sees" over 95% of all the accessible matrix-ligand as isolated, single ligand-groups. At lower, more realistic [M]-values, single matrix-ligands are overwhelmingly predominant (Fig. 1b). The small concentrations of two-clusters and three-clusters may have significant effects depending on protein concentration (see next paragraph), but clusters of four are nearly always of no practical effect, their concentrations never exceeding about 10^{-10} *M*.

Effect of protein concentration

Inspection of the summated term in eqn. 1 shows that, for clusters of *i* matrix-ligands, this term becomes negligible when the cluster concentration $[X_i]$ is very much smaller than the protein concentration $[P_t]$. That is, only clusters whose concentrations approach or exceed that of the protein are expected to affect the partitioning ratio (or elution volume). Moreover, to be effective cluster concentrations

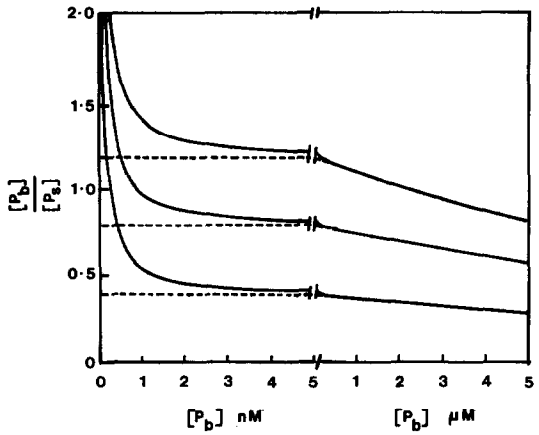

Fig. 2. Scatchard-type partitioning plots for concerted-cluster and monovalent models using the same model parameters. Continuous lines are used for the concerted-cluster model which models a four-site protein of 4 nm radius. Dashed lines are used for the monovalent model, for which the single stoichiometric constant was taken as $4K_M$. The two sets of lines coincide over the right-hand half of the diagram. Parameters for both plots were: [M] = 3 μM, $K_M = 10^5\ M^{-1}$, $K_S = 10^6\ M^{-1}$, [S] = 0 (top curves), [S] = 0.5 μM (middle curves) and [S] = 2 μM (bottom curves). [P_b] = adsorbed protein; [P_s] = protein in solution phase.

must be substantially higher than this when in competition with soluble ligand S, *i.e.* when the soluble ligand concentration is such that K_S[S] makes a substantial contribution to the denominator of eqn. 1. At the assumed upper limit for [M] (about 10^{-3} M), the sum of the cluster concentrations [X_2] and [X_3] is in the micromolar range, so partitioning or chromatography experiments that use protein concentrations in the micromolar range (as many do) are likely to show the effects of such higher-order clusters. The corollary to this is, of course, that in most cases [M] is expected to be well below 10^{-3} M, so that unless [P_t] is well below the micromolar range, the effects of higher-order clusters will pass unnoticed, *i.e.* the partitioning will appear effectively monovalent. This point is well illustrated in the Scatchard plots of Fig. 2, which show that, above a certain value of [P_t], predictions of the present theory and of monovalent theory[3] are indistinguishable.

Predicted values of stoichiometric association constants

Fig. 3a shows the dependence of the stoichiometric constants K_i on [M], for $K_M = 10^4\ M^{-1}$, $K_M = 10^6\ M^{-1}$ (a fairly typical value and about the minimum for K_M consistent with effective affinity chromatography[16,17]), and $K_M = 10^8\ M^{-1}$. Note the log–log scales used. The most striking feature of this prediction is the relatively enormous increase on going from K_{i-1} to K_i, compared to what one would expect from purely statistical considerations[13]. The difference is due to the postulated "enhanced concentration factor", E (see Theory), the corresponding values of which, for $i=2$, 3 and 4, are given in Fig. 3b (note log–log scales). Under conditions least favourable to large E (*i.e.* [M] approaching 10^{-3} M), its size is several hundreds for two-clusters, and tens of thousands for three-clusters. More likely conditions (higher K_M and/or lower [M]) both lead to enormously increased values of E.

Two consequences flow from these predictions: (1) Since, for $i>1$, each K_i is at

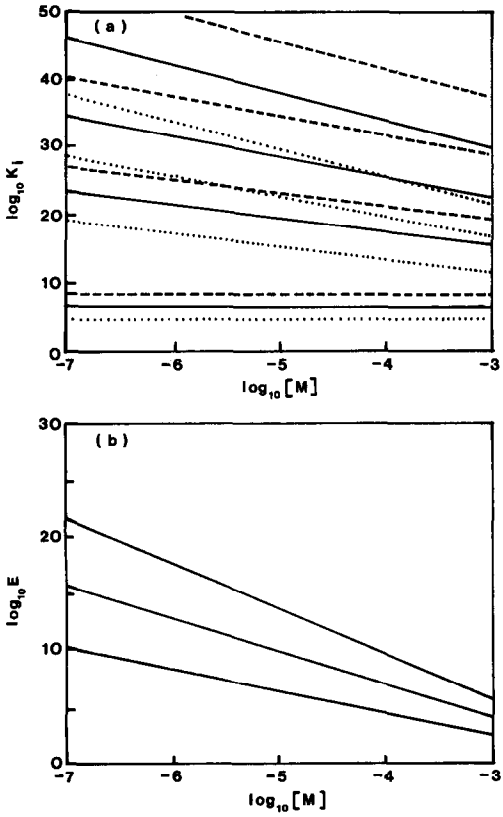

Fig. 3. Dependence of (a) stoichiometric cluster constants K_i and (b) "enhanced concentration" factor, E, on the total concentration of accessible matrix-ligand groups, [M]. The calculations are based on eqns. 7–9 in the text, applied to a four-site protein of radius 4 nm, with K_M set to $10^4\ M^{-1}$ (···), $10^6\ M^{-1}$ (———) or $10^8\ M^{-1}$ (– – –). Reading upwards for each K_M set, the lines are for K_1, K_2, K_3 and K_4 respectively. Note log–log scales.

least E times greater than the corresponding K_{i-1}, we are justified in neglecting protein molecules interacting by fewer than i contacts with a cluster of i matrix-ligands, a basic assumption of the present model. That is, cluster interactions are indeed highly concerted. (2) The enormous K_i values suggest that one should expect "irreversible" binding to clusters to be quite common. This will likely arise from extremely small "off" rate constants for protein-cluster interactions, since the chance of all i contacts being simultaneously severed is small. "Irreversible" cluster interactions permit a major simplification of eqn. 10. Since other terms in the numerator of eqn. 1 are negligible compared to $K_i[P_t]$, one may cancel K_i so that the summated term becomes $[X_i]/[P_t]$, for $i > 1$. For most reasonable values of [M] only two-clusters are likely to be significant (see Fig. 1), so the partitioning ratio turns out as

$$R = 1 + \frac{NK_M[M]}{NK_M[P_t] + 1 + K_S[S]} + \frac{2.52r^3[M]^2}{4[P_t]} \qquad (11)$$

In attempts to fit experimental data to the cluster model by the least-squares method, it has been found in several cases that this simpler equation gives as good a fit to the data, and generates the same parameter values as the more complex eqn. 10. This was the case for a re-interpretation of published data on the interaction of aldolase and phosphocellulose[20].

DISCUSSION

The equations presented above rest mainly on two assumptions: (1) interactions within a cluster are highly cooperative; (2) cluster concentrations follow a Poisson-type distribution. The idea that proximity leads to highly cooperative interactions is not new in biochemistry; it occurs for example in quantitative treatments of DNA annealing and of the helix-coil transition in proteins, and is here extended to multivalent affinity interactions. Invoking the Poisson distribution, *i.e.* assuming a truly random distribution of matrix-ligand groups is a more uncertain proposition, as a number of factors may operate to produce non-random distributions. For example, if the matrix-ligand is a dye, there is evidence that stacking of dye molecules may occur at high dye concentrations[19]. Moreover, the assumption that all matrix-ligands within the bounds of a cluster can bind to protein ignores the restrictive requirements of the spatial distribution (*e.g.* tetrahedral) of protein binding-sites. Strictly, the model as currently formulated applies to matrix-ligands that lack long-range translatory motion, but have considerable local freedom (short-range translation and rotation). Conceivably, more elaborate models of geometrical probability will be found to replace the Poisson distribution for other cases. Notwithstanding these difficulties, the concerted-cluster model represents the first attempt to address the problem of restricted mobility in an affinity matrix. A need for some such model is likely to arise soon, as modern trends in affinity chromatography (in particular the development of high-density, high-performance columns) will inevitably call for greater use of rigid, high-flow matrices (silica, controlled-pore glass, etc.) with consequent limitations in matrix-ligand mobility.

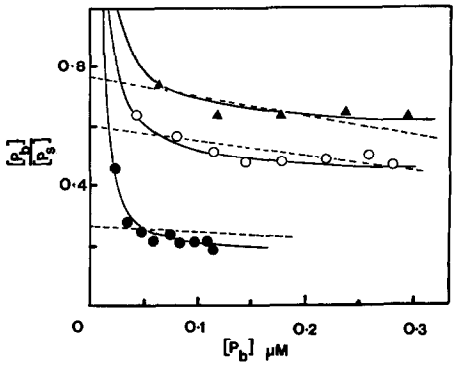

Fig. 4. Scatchard-type plots of the interaction of aldolase and phosphocellulose, with phosphate as soluble ligand. The data are from Table II of ref. 6. Concentrations of phosphate were: ▲, 0.3 m*M*; ○, 1.0 m*M*; ●, 5.0 m*M*. Fitted curves are for the concerted cluster model (———) and the reacted-site-probability model (– –). For other details, see ref. 20. Diagram by permission of the Biochemical Society.

In a search for experimental evidence to support co-operativity and clustering in affinity partitioning, published data on the interaction of aldolase and phosphocellulose have been re-examined; a communication on this work has been published[20]. These data were originally cited in support of a theory of multivalency based on reacted site probability[6]. In the re-examination, the values of K_M and [M] were estimated by non-linear regression of the data in Table II of ref. 6, using both the concerted-cluster and the reacted-site-probability models. Fig. 4 shows the results plotted in the same format as Fig. 2. By the criterion of the smaller sum-of-squares of residual errors, the fit to the concerted-cluster model was better by an order of magnitude. For other details see ref. 20. Of several published sets of data examined, the aldolase–phosphocellulose data have shown the most pronounced upwards curvature at low protein concentrations, in Scatchard-type graphs. This behaviour is predicted by the cluster model, but not by the monovalent or reacted-site-probability models. While experimental data which show a steep increase in [bound protein]/[free protein] at low total protein concentrations clearly fit the cluster model better, data which follow a more near-linear relationship tend to fit all three (monovalent, cluster, reacted-site probability) models about equally well (by the least-squares criterion). In such cases, the cluster model approximates the monovalent model in predicted behaviour (see Fig. 2), *i.e.* clusters are predicted to be negligible. Upon examination, several other sets of published data on affinity interactions appear to exhibit this near-linear relationship. Before it can be decided whether or not the concerted-cluster theory is applicable to these systems, the experiments should be repeated with the emphasis on a considerably lower range of protein concentrations, where the steep increase in Scatchard-type graphs, denoting significant concentrations of high-affinity adsorption sites, may be evident.

ACKNOWLEDGEMENTS

I thank Dr. John Gates, School of Mathematics, Statistics and Computing, Thames Polytechnic, for suggesting use of the Poisson distribution. The School of Biological Sciences and Environmental Health, Thames Polytechnic, is a member of the London Centre for Biotechnology.

REFERENCES

1 P. Andrews, B. J. Kitchen and D. J. Winzor, *Biochem. J.*, 135 (1973) 897.
2 B. M. Dunn and I. M. Chaiken, *Proc. Natl. Acad. Sci. U.S.A.*, 71 (1974) 2382.
3 L. W. Nichol, A. G. Ogston, D. J. Winzor and W. H. Sawyer, *Biochem. J.*, 143 (1974) 435.
4 K.-I. Kasai and S.-I. Ishii, *J. Biochem. (Tokyo)*, 77 (1975) 261.
5 I. M. Chaiken, *Anal. Biochem.*, 97 (1979) 1.
6 L. W. Nichol, L. D. Ward and D. J. Winzor, *Biochemistry*, 20 (1981) 4856.
7 P. Kyprianou and R. J. Yon, *Biochem. J.*, 207 (1982) 549.
8 H. W. Hethcote and C. DeLisi, in I. M. Chaiken, M. Wilchek and I. Parikh (Editors), *Affinity Chromatography and Biological Recognition*, Academic Press, New York, 1983, p. 119.
9 D. J. Winzor, in P. D. G. Dean, W. S. Johnson and F. A. Middle (Editors), *Affinity Chromatography: A Practical Approach*, IRL Press, Oxford, 1985, Ch. 6, p. 149.
10 P. J. Hogg and D. J. Winzor, *Arch. Biochem. Biophys.*, 234 (1984) 55.
11 H. E. Swaisgood and I. M. Chaiken, *Biochemistry*, 25 (1986) 4148.
12 J. Hubble, (1987), *Biotech. Bioeng.*, 30 (1987) 208.
13 I. M. Klotz, *Acc. Chem. Res.*, 7 (1974) 162.

14 G. A. Gilbert, *Nature (London)*, 212 (1966) 296.
15 C. Chatfield, *Statistics for Technology*, Chapman & Hall, London, 2nd edn., 1978, p. 69.
16 R. J. Yon, *Biochem. J.*, 185 (1980) 211.
17 R. K. Scopes, *Anal. Biochem.*, 165 (1987) 235.
18 R. J. Yon, *Biochem. Soc. Trans.*, 16 (1988) 542.
19 Y-C. Liu, R. Ledger and E. Stellwagen, *J. Biol. Chem.*, 259 (1984) 3796.
20 R. J. Yon, *Biochem. Soc. Trans.*, 16 (1988) 543.
21 P. J. Hogg and D. J. Winzor, *Arch. Biochem. Biophys.*, 240 (1985) 70.