Journal of Chromatography, 216 (1981) 43–62
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

CHROM. 14,099

THEORY OF AFFINITY ELECTROPHORESIS

EVALUATION OF THE EFFECTS OF PROTEIN MULTIVALENCY, IMMOBILIZED LIGAND HETEROGENEITY AND MICRODISTRIBUTION AND DETERMINATION OF EFFECTIVE CONCENTRATION OF IMMOBILIZED LIGAND

V. HOŘEJŠÍ*
Institute of Molecular Genetics, Czechoslovak Academy of Sciences, Vídeňská 1083, Prague 4 (Czechoslovakia)
and
MARIE TICHÁ
Department of Biochemistry, Charles University, Albertov 2030, Prague 2 (Czechoslovakia)
(Received June 9th, 1981)

SUMMARY

Equations are presented which describe the effects of multivalency of protein molecules on electrophoretic mobility in an affinity gel with respect to the possibility of the quantitative evaluation of the strength of protein–ligand interactions under these conditions. At low concentrations of immobilized ligand (c_{im}), fairly exact K and L/n values can be obtained, even for multivalent proteins from linear plots (K = intrinsic dissociation constant of the protein–mobile ligand complex; L = intrinsic dissociation constant of the protein–immobilized ligand complex; n = the number of independent, identical ligand-binding sites in the protein molecule). At higher c_{im} conventional plotting of variables obtained by affinity electrophoresis yields curvilinear plots usually deviating only slightly from straight lines, the evaluation of which yields usually apparent K and L values lower than the true values, *i.e.* the strength of interaction is overestimated under these conditions. The importance of c_{im} for the degree of restriction of multivalent interactions with immobilized ligand is stressed.

No simple means of estimating n by affinity electrophoresis has been found.

For a bivalent protein with two independent non-identical ligand-binding sites, apparent K and L values are obtained which lie between the true K_1, K_2 or L_1, L_2 values characterizing the two sites.

Effective c_{im} values can be obtained by following the dependence of the mobility of a protein in affinity gels containing a fixed c_{im} on the protein concentration. The effects of non-homogeneity of immobilized ligand molecules (*i.e.*, the presence of different types of immobilized ligands within the gel) are quantitatively evaluated.

The effects of steric non-homogeneity of the distribution of immobilized ligand molecules within the gel are considered qualitatively.

The results of some experiments aimed at verifying the equations describing the

0021-9673/81/0000–0000/$02.50

effects of immobilized ligand concentration and microdistribution on the values of K and L for bivalent lectins are presented; experimental approaches for testing the other theoretical conclusions reached are suggested.

Most of the equations derived should also be applicable, after minor modifications, for the quantitative description of affinity chromatography systems.

INTRODUCTION

Affinity electrophoresis has been used in a number of studies both for qualitative detection of ligand-binding proteins and for quantitative purposes, *i.e.*, determination of dissociation constants of complexes of proteins with immobilized or free (mobile) ligands (for a review, see ref. 1). The evaluation of dissociation constants from the data obtained by affinity electrophoresis has so far usually been based on extremely simplified equations, which are valid exactly only under following assumptions:

(1) Immobilization of the ligand is complete; the mobility of the protein-immobilized ligand complex is zero.

(2) The mobility of the protein–mobile ligand complex is the same as that of uncomplexed protein (*i.e.*, interaction with mobile ligand does not affect the mobility of the protein).

(3) The concentration of the protein within the moving zone is much lower than the concentrations of both the immobilized and free ligands.

(4) Complex formation is a very fast reaction; the effects of its kinetics are negligible.

(5) The protein is monovalent with respect to the ligand.

(6) The total concentration of the immobilized ligand is identical with the effective concentration (*i.e.*, all molecules of the immobilized ligand are freely accessible to the protein).

(7) All molecules of the immobilized ligand are equivalent, *i.e.*, the strengths of their interaction with the protein are identical.

(8) The microdistribution of immobilized ligand molecules is homogeneous.

However, in most real systems some of these assumptions obviously do not hold true. The effects of invalidity of assumptions 1–4 were treated in detail in a previous paper on the theory of affinity electrophoresis[2], and the consequences of invalidity of assumptions 5–8 were briefly discussed at a Symposium[3]. The aim of this study was to evaluate theoretically in detail the importance of assumptions 5–8; experimental approaches to testing some of the theoretical conclusions are suggested.

EXPERIMENTAL

Water-soluble glycosyl polyacrylamide copolymers were prepared as described earlier[4]. The sugar contents of the copolymers used in this work (determined spectrophotometrically[5] with reference to the corresponding free sugar) were α-D-galactosyl copolymer 11.1 % and α-D-glucosyl copolymer 13.1 %, if not stated otherwise.

Lectins from the seeds of *Lathyrus sativus*[6] and concanavalin A[7] (con A) were isolated by affinity chromatography on Sephadex G-150.

Affinity electrophoresis in polyacrylamide gel was performed in a discontinuous acidic buffer system[8] according to the standard procedure, omitting the stacking gel layer[9]. The α-D-mannosyl or α-D-glucosyl polyacrylamide copolymers were added to the polymerization mixture used normally for the preparation of the gels in amounts yielding the desired final concentration of immobilized sugar ($0.2 \cdot 10^{-3}$–$3.6 \cdot 10^{-3}$ *M*) in the affinity gels.

The polymerization mixture poured into the glass tubes was overlaid with an aqueous solution containing the same buffer and copolymer concentrations as those present in the gel instead of a water overlay. To determine the *K* values, free D-glucose was added to the above mixture to obtain its final concentration of $2.8 \cdot 10^{-4}$–$2.8 \cdot 10^{-2}$ *M*.

Samples containing 20 μg of the lectin and 10 μg of cytochrome *c* in 30 μl of 20% glycerol solution were applied per tube (5 × 80 mm) and electrophoresis in an acidic buffer system[8] was continued at 7 mA per tube for 1.5 h. Proteins were rendered visible with Amido Black and the migration distances from the top of the gel were measured with an accuracy of ±0.5 mm. To eliminate possible variations of the mobilities in the individual gels, the migration distances of the lectin bands were always measured relative to the mobility of the internal standard cytochrome *c* (which possessed a high mobility in the system used).

Apparent dissociation constants *K* and *L* were determined graphically from plots of $d/(d_0 - d)$ *versus* c and $1/(d_0 - d)$ *versus* $1/c_{im}$, respectively[2,9]. The values of d_0 were estimated on the control gels containing corresponding amounts of α-D-galactosyl copolymer instead of α-D-mannosyl or α-D-glucosyl copolymers.

RESULTS AND DISCUSSION

Effects of multivalency of protein molecules

In this section we shall derive equations describing the mobility of a multivalent protein in affinity gels as a function of c or $1/c_{im}$. The variables used in these equations and their form are always such as to allow direct comparison with the simple equations describing the affinity electrophoresis of monovalent proteins as used normally for the evalution of *K* and *L*[2,9]. Our aim was to assess the applicability of affinity electrophoresis for the estimation of *K* and *L* for multivalent proteins and to evaluate the errors in their estimation under different experimental conditions. The validity of all remaining assumptions (1–4 and 6–8, see Introduction) was assumed during the derivation of all equations.

In the general case, a protein molecule will contain n independent, non-identical ligand-binding sites, each characterized by dissociation constants K_i, L_i for a particular mobile or immobilized ligand, respectively. When such a protein is introduced into an affinity gel containing immobilized ligand (c_{im}) and mobile ligand (c), three types of particles will be present at equilibrium: (1) free protein molecules [A]; (2) complexes $[X_{i,j}]$, *i.e.*, protein molecules complexed with i molecules of immobilized ligand ($1 \leqslant i \leqslant n$) and j molecules of mobile ligand ($0 \leqslant j \leqslant n - i$); (3) complexes $[Y_k]$, *i.e.* protein molecules complexed with k molecules of mobile ligand ($1 \leqslant k \leqslant n$). (Thus, $[X_{i,j}]$ complexes are all those containing at least one immobilized ligand molecule regardless of the number of mobile ligand molecules, whereas $[Y_k]$ complexes are those consisting of protein molecules complexed with mobile ligand molecules only).

Owing to the mutual non-identity of the ligand-binding sites there will be $\binom{n}{i} \cdot \binom{n-i}{j}$ different $[X_{i,j}]$ complexes for each i, j pair and $\binom{n}{k}$ different $[Y_k]$ complexes for each k.

As shown earlier[2], the parameter $d/(d_0 - d)$ describing the relative retardation of the protein band in the affinity gel as compared with the control gel is equal to the ratio of the time t_1 spent by the protein molecule in the form of freely mobile particles (free protein or complexes $[Y_k]$) and t_2, *i.e.*, time spent in the form of immobile complexes $[X_{i,j}]$. Thus,

$$\frac{d}{d_0 - d} = \frac{t_1}{t_2} = \frac{[A] + \sum_k [Y_k]}{\sum_i \sum_j [X_{i,j}]} \tag{1}$$

where $[Y_k]$ and $[X_{i,j}]$ represent the sum of all possible complexes of the respective type for given k or i, j pair.

The general case of non-identical ligand-binding sites will now be greatly simplified when all sites will be equivalent (characterized by single K and L values). Then each $[Y_k]$ and $[X_{i,j}]$ will represent single type of particles and eqn. 1 can be rewritten as

$$\frac{d}{d_0 - d} = \frac{t_1}{t_2} = \frac{[A] + \sum_k \binom{n}{k} [Y_k]}{\sum_i \sum_j \binom{n}{i} \binom{n-i}{j} [X_{i,j}]} \tag{1a}$$

where the coefficients $\binom{n}{k}$ and $\binom{n}{i} \cdot \binom{n-i}{j}$ account for the number of originally distinguishable complexes of the respective type (when the sites were non-identical) which contributed to the resulting $[Y_k]$ and $[X_{i,j}]$ in the simplified case of identical sites. Supposing that there are no restrictions of multiple interactions of the protein molecule with immobilized ligand molecules (for details of this phenomenon, see below), equilibrium concentrations of $[Y_k]$ and $[X_{i,j}]$ can be expressed using dissociation constants K, L, the actual free protein concentration $[A]$, c_{im}, c, $[Y_{k-1}]$ and $[X_{i-1,j}]$ or $[X_{i,j-1}]$ as follows:

$$[Y_k] = \frac{[Y_{k-1}]c}{K}$$

$$[X_{i,j}] = \frac{[X_{i,j-1}]c}{K}$$

$$[X_{i,j}] = \frac{[X_{i-1,j}]c_{im}}{L}$$

because $[Y_1]$, $[X_{1,0}]$ and $[X_{1,1}]$ can be expressed as

$$[Y_1] = \frac{[A]c}{K}$$

$$[X_{1,0}] = \frac{[A]c_{im}}{L}$$

$$[X_{1,1}] = \frac{[X_{1,0}]c}{K}$$

all $[Y_k]$ and $[X_{i,j}]$ can be successively expressed using only c, c_{im}, K, L and [A]:

$$[Y_k] = [A]\left(\frac{c}{K}\right)^k$$

$$[X_{i,j}] = [A]\left(\frac{c_{im}}{L}\right)^i\left(\frac{c}{K}\right)^j$$

Thus, eqn. 1a can be written as

$$\frac{d}{d_0 - d} = \frac{[A] + \sum_{k=1}^{n} \binom{n}{k} [A]\left(\frac{c}{K}\right)^k}{\sum_{i=1}^{n} \sum_{j=0}^{n-i} [A]\left(\frac{c_{im}}{L}\right)^i\left(\frac{c}{K}\right)^j \binom{n}{i}\binom{n-i}{j}} \tag{1b}$$

which, after simple rearrangement, yields the basic equation

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^n}{\sum_{i=1}^{n} \binom{n}{i}\left(\frac{c_{im}}{L}\right)^i\left(1 + \frac{c}{K}\right)^{n-i}} \tag{2}$$

which describes the affinity electrophoresis of a multivalent protein under the conditions of unrestricted possibilities of multivalent interactions with an immobilized ligand.

If only interaction with an immobilized ligand is followed, *i.e.*, affinity gels contain variable c_{im} whereas $c = 0$, eqn. 2 becomes

$$\frac{1}{d_0 - d} = \frac{1}{d_0} + \frac{1}{d_0 \sum_{i=1}^{n} \binom{n}{i}\left(\frac{c_{im}}{L}\right)^i} \tag{3}$$

These basic general eqns. 2 and 3 yield for $n = 1$ the following equations, derived previously for the evaluation of K and L in monovalent proteins[2,9]:

$$\frac{d}{d_0 - d} = \frac{L}{c_{im}}\left(1 + \frac{c}{K}\right) \tag{2a}$$

$$\frac{1}{d_0 - d} = \frac{L}{d_0} \cdot \frac{1}{c_{im}} + \frac{1}{d_0} \tag{3a}$$

and for $n = 2$ and 3 eqns. 2b, 2c and 3b, 3c, respectively:

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^2}{\frac{2c_{im}}{L}\left(1 + \frac{c}{K} + \frac{1}{2} \cdot \frac{c_{im}}{L}\right)} \tag{2b}$$

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^3}{\frac{3c_{im}}{L}\left[\left(1 + \frac{c}{K}\right)^2 + \frac{c_{im}}{L}\left(1 + \frac{c}{K}\right) + \frac{1}{3}\left(\frac{c_{im}}{L}\right)^2\right]} \tag{2c}$$

$$\frac{1}{d_0 - d} = \frac{1}{\frac{2d_0 c_{im}}{L}\left(1 + \frac{c_{im}}{2L}\right)} + \frac{1}{d_0} \tag{3b}$$

$$\frac{1}{d_0 - d} = \frac{1}{\frac{3d_0 c_{im}}{L}\left[1 + \frac{c_{im}}{L} + \frac{1}{3}\left(\frac{c_{im}}{L}\right)^2\right]} + \frac{1}{d_0} \tag{3c}$$

Steric restriction of multiple interactions with immobilized ligands

As stated above, the general eqns. 2 and 3 are valid only if there are no restrictions on the multiple interactions of the protein molecule with immobilized ligand molecules. However, when c_{im} is sufficiently low, the mean distance between neighbouring immobilized ligand molecules may be much greater than the dimensions of the protein molecule (distance between ligand-binding sites) and, therefore, simultaneous interaction of the protein with two or even more immobilized ligand molecules becomes very unlikely. Quantitative evaluation of this phenomenon in a general case seems to be very complex owing to involvement of factors such as geometry and flexibility of the protein molecule. However, if only monovalent and restricted bivalent interactions of an n-valent protein with immobilized ligand are considered (tri- and higher-valent simultaneous interactions are neglected owing to their strongly decreasing probability), an empirical parameter p ($0 < p < 1$) can be defined, which denotes the fraction of immobilized ligand molecules the nearest neighbour of which lies within the range of protein molecule dimensions and which are therefore available for the formation of bivalent complexes with immobilized ligand. Under these conditions of restricted bivalent interaction with immobilized ligand, the same type of reasoning as in deriving eqns. 2 and 3 yields the equations

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^2}{\frac{nc_{im}}{L}\left[1 + \frac{c}{K} + p \cdot \frac{(n-1)}{2} \cdot \frac{c_{im}}{L}\right]} \tag{4}$$

and

$$\frac{1}{d_0 - d} = \frac{1}{\frac{d_0 nc_{im}}{L}\left(1 + p \cdot \frac{(n-1)}{2} \cdot \frac{c_{im}}{L}\right)} + \frac{1}{d_0} \tag{5}$$

When, in addition to restricted bivalent interactions, trivalent interactions of an n-valent protein with immobilized ligand are also taken into account (neglect of higher interactions), the following equations are approximately valid:

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^3}{\frac{nc_{im}}{L}\left[\left(1 + \frac{c}{K}\right)^2 + \frac{n-1}{2} \cdot p \cdot \frac{c_{im}}{L}\left(1 + \frac{c}{K}\right) + p^2 \cdot \frac{(n-1)(n-2)}{6} \cdot \left(\frac{c_{im}}{L}\right)^2\right]}$$

$$\frac{1}{d_0 - d} = \frac{1}{\frac{d_0 nc_{im}}{L}\left[1 + \frac{(n-1)}{2} \cdot p \cdot \frac{c_{im}}{L} + p^2 \cdot \frac{(n-1)(n-2)}{6} \cdot \left(\frac{c_{im}}{L}\right)^2\right]} + \frac{1}{d_0} \tag{7}$$

Of course, for $n = 2$ and $p = 1$, eqns. 4 and 5 yield the formerly derived eqns. 2b and 3b, respectively; for $n = 1$, p is automatically zero and eqns. 4 and 5 yield the simple eqns. 2a and 3a. For $n = 3$ and $p = 1$, eqns. 4 and 5 become eqns. 2c and 3c, respectively. Whereas for $n = 2$ or 3 eqns. 4 and 5 are only approximate (due to neglect of multiple interactions of higher order with immobilized ligands), especially for $p \to 1$, when the probability of multiple interactions strongly increases eqns. 8 and 9, *i.e.*, eqns. 4 and 6 for $n = 2$ and 3, respectively, should be exact for any p, as all possible kinds of "multiple" interactions (*i.e.*, double or triple) are included:

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^2}{\frac{2c_{im}}{L}\left(1 + \frac{c}{K} + p \cdot \frac{c_{im}}{2}\right)} \tag{8}$$

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K}\right)^3}{\frac{3c_{im}}{L}\left[\left(1 + \frac{c}{K}\right)^2 + p \cdot \frac{c_{im}}{L}\left(1 + \frac{c}{K}\right) + \frac{p^2}{3}\left(\frac{c_{im}}{L}\right)^2\right]} \tag{9}$$

Eqns. 2b and 2c are in fact identical with eqns. 8 and 9, respectively, for $p = 1$.

However, it should be noted that p is a function of c_{im} (see below), so that although c_{im} and thus also p can be kept constant during plotting $d/(d_0 - d)$ *vs.* c (eqns. 4, 6 and 8), when plotting $1/(d_0 - d)$ *vs.* $1/c_{im}$ (in order to determine L, eqns. 5 and 7) p is dependent on c_{im}.

The extent of multivalent interactions with immobilized ligand molecules and their effects on the results of affinity electrophoresis depend substantially on the parameter p, characterizing the degree of restriction of bivalent interactions. The value of p should be unambiguously determined by the concentration of immobilized ligand, c_{im}, and the size and geometry of the protein molecule (see the following section).

When $p \to 0$, which is an experimentally important and interesting case, even bivalent interactions of an n-valent protein with immobilized ligand are negligible and the general eqns. 2 and 3 are greatly simplified to

$$\frac{d}{d_0 - d} = \frac{L}{nc_{im}}\left(1 + \frac{c}{K}\right) \tag{10}$$

$$\frac{1}{d_0 - d} = \frac{L}{nd_0} \cdot \frac{1}{c_{im}} + \frac{1}{d_0} \tag{11}$$

Dependence of p on c_{im} and dimensions of protein molecules

A simple equation for determining p can be derived under the following simplified conditions. Immobilized ligand molecules (concentration c_{im}) are distributed statistically within the gel and are completely fixed in their positions (*i.e.*, devoid of any vibrational, rotational and similar limited movements). If the concentration of bivalent protein molecules (containing two ligand-binding sites with maximum distance D) is much lower than c_{im}, then the probability that after complexing the protein molecule with one immobilized ligand at least one other immobilized ligand molecule will be found within the distance D from that first protein-complexed immobilized ligand molecule (*i.e.*, only in such a case can the second ligand-binding site also be complexed) is according to Poisson's law:

$$P = 1 - e^{-\frac{4}{3}\pi D^3 c_{im}} \tag{12}$$

where c_{im} must be expressed as the absolute number of particles per unit volume. Obviously, according to its definition, this P value can be taken as a rough first-order approximation of p ($P = p$). However, two major effects are expected to affect this theoretical p value as given by eqn. 12:

(1) Immobilized ligand molecules are certainly not totally fixed; they can oscillate within some range depending mainly on the rigidity of the macromolecular carrier (*e.g.*, gel matrix) used as a framework for their immobilization. This incomplete fixation will clearly tend to increase the p value compared with eqn. 12 owing to an increase in the effective D value and effective local c_{im} value.

(2) Eqn. 12 assumes that the protein molecule is absolutely flexible; however, the rigidity of real protein molecules will prevent complexation with the second ligand molecule if it is located within a distance less than D from the first complexed ligand molecule but in a sterically unfavourable position, *e.g.*, too close to the first molecule,

In fact, there should be only a relatively limited region of favourable positions of the second ligand molecule with respect to the first ligand, depending on the properties of the protein molecule. This factor should decrease the p value as predicted by eqn. 12.

It is difficult to guess the combined effects of these two major opposing factors on the resulting effective value of p. It may be feasible to use the coefficient p_{eff} defined as

$$p_{eff} = fp$$

where p is a theoretical value calculated from known c_{im} and D values and the empirical coefficient f accounts for all effects neglected by eqn. 12; however, the D values is known only in a few cases of thoroughly characterized proteins and the uncertainty in c_{im} is discussed in the following sections.

Affinity electrophoresis of proteins containing two different, independent ligand-binding sites

Finally, we shall consider the case of a protein with two different independent ligand-binding sites (the first site characterized by dissociation constants K_1 and L_1 and the second site by K_2 and L_2 for a particular ligand). Using the same approach as that leading to eqns. 2 and 3, eqns. 13 and 14 are obtained, which are valid under the conditions of unrestricted bivalent interactions with immobilized ligands ($p = 1$):

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K_1}\right)\left(1 + \frac{c}{K_2}\right)}{\frac{c_{im}(L_1 + L_2)}{L_1 L_2}\left(1 + \frac{L_2K_1 + L_1K_2}{K_1K_2(L_1 + L_2)} \cdot c + \frac{c_{im}}{L_1 + L_2}\right)} \tag{13}$$

$$\frac{1}{d_0 - d} = \frac{1}{\frac{d_0(L_1 + L_2)}{L_1L_2} \cdot c_{im}\left(1 + \frac{c_{im}}{L_1 + L_2}\right)} + \frac{1}{d_0} \tag{14}$$

When bivalent interactions with immobilized ligand are negligible ($p = 0$), the following equations are valid:

$$\frac{d}{d_0 - d} = \frac{\left(1 + \frac{c}{K_1}\right)\left(1 + \frac{c}{K_2}\right)}{\frac{c_{im}(L_1 + L_2)}{L_1L_2}\left[1 + \frac{c(K_1L_2 + L_1K_2)}{K_1K_2(L_1 + L_2)}\right]} \tag{15}$$

$$\frac{1}{d_0 - d} = \frac{1}{d_0 c_{im}\left(\frac{1}{L_1} + \frac{1}{L_2}\right)} + \frac{1}{d_0} \tag{16}$$

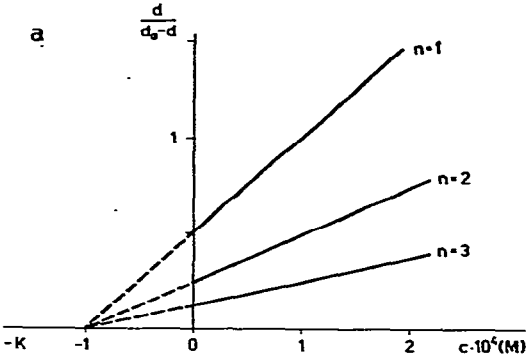
Eqns. 13, 14, 15 and 16 become eqns. 2b, 3b, 10 and 11, respectively ($n = 2$), for $K_1 = K_2$ and $L_1 = L_2$.

Experimental consequences of eqns. 2–16

(1) From the above equations, it is obvious that there is no simple means of determining the number of ligand-binding sites (n) of a protein molecule from the results of affinity electrophoresis.

(2) When a protein is subjected to affinity electrophoresis, usually $d/(d_0 - d)$ is plotted against c (to determine K) or $1/(d_0 - d)$ against $1/c_{im}$ (to determine L). Both of these plots are straight lines for monovalent proteins[2]. The general eqns. 2 and 3 describing the behaviour of multivalent proteins are generally non-linear with respect to these variables. Therefore, it is of interest to evaluate the degree of non-linearity caused by multivalency of a protein and to assess the applicability of plotting the above variables for approximate determination of K and L and to determine expected errors in the determination of these parameters.

Importantly, eqns. 10 and 11 (which describe the affinity electrophoresis of multivalent proteins under the conditions of $p = 0$, *i.e.*, at sufficiently low c_{im} and/or with small dimensions of the protein molecule) are linear. As illustrated in Fig. 1,



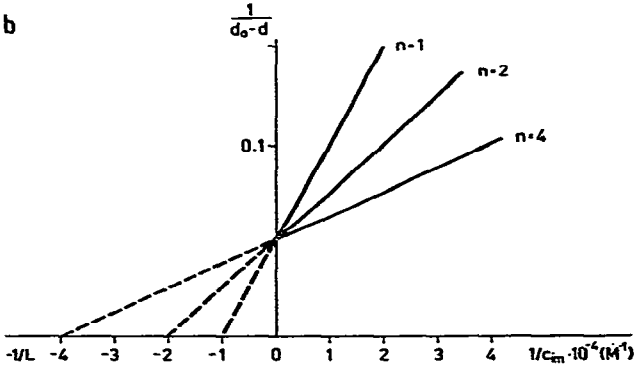


Fig. 1. Graphical representation of (a) eqn. 10, (b) eqn. 11 for a hypothetical protein occurring in stable mono-, di-, tri- and tetravalent forms and characterized by $K = L = 10^{-4}$ M for the free and immobilized ligand used. The values $c_{im} = 2 \cdot 10^{-4}$ M and $d_0 = 20$ mm were used for construction of the plots. Note that K is independent of n.

under these conditions eqn. 10 can be used for the determination of K regardless of n; however, instead of L the value L/n is obtained from both eqns. 10 and 11. The true value of L cannot be found without a knowledge of independently determined n. Thus, K and L/n values can be determined exactly from linear plots even for multivalent proteins, provided that multivalent interactions with immobilized ligand molecules are negligible. As stated above, for proteins in the ordinary molecular weight range this condition is better met the lower is c_{im}. The strength of the interaction with immobilized ligands (as defined by the value of L/n) determines the range of c_{im} applicable in the affinity gels to obtain a reasonable retardation of the protein. Usually c_{im} values are used such as to produce $0.2\ d_0 < d < 0.75\ d_0$. Thus, c_{im} of the approximate order L/n are used and, consequently, the stronger is the interaction the lower are the c_{im} values that can (and even must) be used, which favours meeting the condition $p \rightarrow 0$, *i.e.*, suppression of multivalent interactions with immobilized ligands. It may be useful to calculate the numerical value of p using eqn. 12 for some typical values of c_{im} and D (Table I).

TABLE I

CALCULATED VALUES OF p FOR TYPICAL VALUES OF c_{im} AND D

D (nm)	c_{im} (M)	p
5	10^{-2}	0.969
5	10^{-3}	0.270
5	10^{-4}	0.031
10	10^{-4}	0.223
10	10^{-5}	0.025
10	10^{-6}	0.003
35	10^{-6}	0.102
35	10^{-7}	0.011
35	10^{-8}	0.001

The value $D = 5$ nm corresponds roughly to the distance between sugar binding sites in the con A molecule[10], and $D = 10$ and 35 nm correspond approximately to the maximum distances between ligand-binding sites in typical IgG and IgM molecules, respectively[11]. Thus, although these values of p must be considered only as approximate (owing to the neglect of some steric factors discussed above) it is obvious that for con A-like molecules multiple interactions (or, more exactly, bivalent interactions) with immobilized ligand molecules will be nearly negligible at $c_{im} < 10^{-4}\ M$ for IgG at $c_{im} < 10^{-5}\ M$ and for IgM for $c_{im} < 10^{-6}$–$10^{-7}\ M$. A clear test for the detection of multivalent interactions should be the experimentally found form of the plots of $d/(d_0 - d)$ *vs.* c or $1/(d_0 - d)$ *vs.* $1/c_{im}$; only if these plots are good straight lines will multivalent interactions not occur. However, in practice it may be difficult to decide from experimental data whether the plot is or is not a straight line. This is illustrated in Fig. 2, which shows the graphical representation of eqns. 4, 5, 7, 8 and 9 for defined values of K, L, c_{im}, p and n. These plots clearly demonstrate the following:

(a) Multivalent interactions with immobilized ligands cause an apparent increase in the strength of interactions with immobilized ligands.

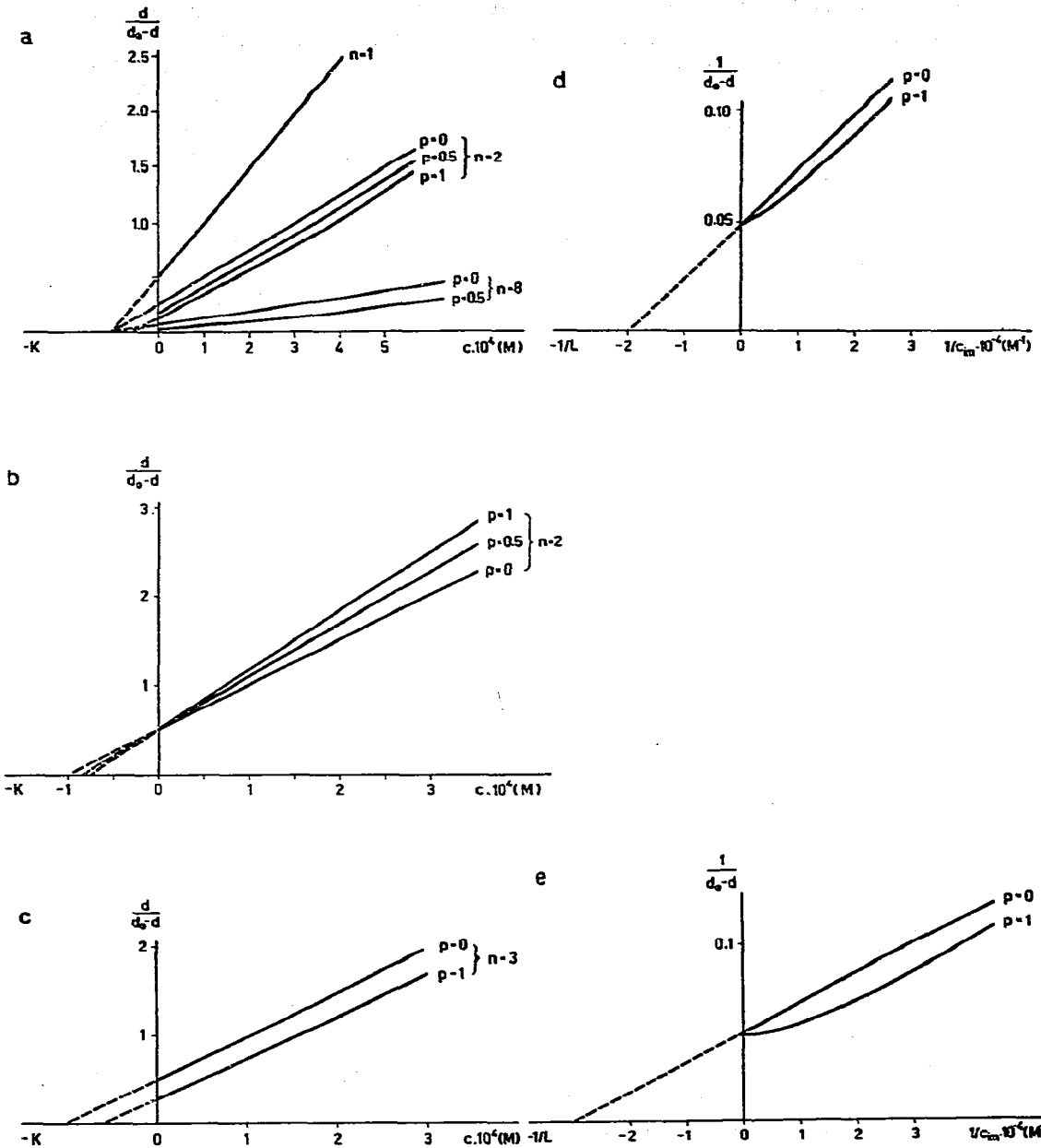


Fig. 2. Graphical representations of (a) eqn. 4, (b) eqn. 8, (c) eqn. 9, (d) eqn. 5 and (e) eqn. 7 for hypothetical proteins with $K = L = 10^{-4}$ *M* and the values of *n* and *p* as indicated. The value of c_{im} is (a) $2 \cdot 10^{-4}$ *M*, (c) $0.66 \cdot 10^{-4}$ *M*, or was chosen so to produce $d/(d_0 - d) = 0.5$ for (b) $c = 0$ (*i.e.*, $c_{im} = 10^{-4}$ *M* for the curve with $p = 0$, $c_{im} = 0.88 \cdot 10^{-4}$ *M* for the curve with $p = 0.5$ and $c_{im} = 0.73 \cdot 10^{-4}$ *M* for the curve with $p = 1$). The $-K$ values are determined from the plots as intercepts of the broken lines with the negative abscissa (a, b and c) and the $1/L$ values as intercepts of the $1/(d_0 - d)$ vs. $1/c_{im}$ plots with the negative abscissa (d and e). Note the apparent decrease in *K* at $p > 0$ and the markedly curvilinear character of the $1/(d_0 - d)$ *vs.* $1/c_{im}$ plot for $p > 0$, resulting in an uncertain estimation of the apparent *L*.

(b) The curves expected for multivalent (in fact, bivalent) interactions with immobilized ligands are very similar to straight lines; detection of a non-linear character from experimental points would be difficult, at least when using c_{im} to produce optimal retardation ($d \approx 0.3\ d_0$). The departure from linearity is more obvious at higher c_{im}; however, at such c_{im} values exact measurement of very small d values is difficult.

(c) The occurrence of bivalent interactions always leads more or less to an overestimation of the strength of interactions with free ligands (underestimation of K). The effects of multivalency would probably be much more marked when, in addition to bivalent interactions, higher interactions are also taken into account; however, in such a case experimental complications might occur, such as irreversible precipitation of the protein near the application site. Nevertheless, it should be stressed that when $p \rightarrow 0$, the K value estimated by affinity electrophoresis should be very close to its true value. To test experimentally the validity of the condition $p \rightarrow 0$ (resulting in the absence of bivalent interactions with immobilized ligands) during typical affinity electrophoresis experiments, we measured K using the plot of $d/(d_0 - d)$ *vs.* c for con A, which is bivalent at the acidic pH used in our experiments. We used affinity gels with different c_{im} (immobilized α-D-mannosyl or α-D-glucosyl residues, $2 \cdot 10^{-4}\ M < c_{im} < 3.6 \cdot 10^{-3}\ M$; the occurrence of bivalent interactions at higher c_{im} (p increases with increasing c_{im}) would be manifested by lower apparent K values estimated on affinity gels with higher c_{im} values. However, in all instances identical values of $K = 4.1 \cdot 10^{-3}\ M$ were found for free D-glucose from linear $d/(d_0 - d)$ *vs.* c plots (graph not shown). This result is in agreement with the above guess that for con A p should be low for c_{im} of the order 10^{-3}–$10^{-4}\ M$, which should virtually eliminate the possibility of bivalent interactions with immobilized ligand molecules.

(3) The validity of equations describing the effects of multivalency would be tested best by using mono- to n-valent derivatives of the same protein, *e.g.*, monoclonal immunoglobulins and their fragments or oligomers as used for similar purposes in quantitative affinity chromatographic systems[12].

(4) The curvilinear character of the plot of $d/(d_0 - d)$ *vs.* c predicted by eqn. 2 (Fig. 2) was recently observed during affinity electrophoresis of *Dolichos biflorus* lectin on affinity gels containing immobilized hog blood group substance[13].

(5) The presence of two types of ligand-binding sites in a bivalent protein molecule would not be revealed experimentally by affinity electrophoresis; in this instance the results would be the same as for a protein with two identical sites characterized by K and L values intermediate to K_1, K_2 and L_1, L_2, respectively (Fig. 3). The only indication of the presence of two different ligand-binding sites instead of two identical sites is the non-linearity of eqn. 15 with respect to c; however, this non-linearity may be difficult to establish experimentally (Fig. 3).

Effective concentration of immobilized ligands

Immobilization of the ligand within the affinity gel can be achieved in different ways[1] and any of these ways may lead to a portion of total number of immobilized ligand molecules which for steric reasons are not accessible to the interaction with the protein migrating through the affinity gel. This may occur especially when the ligand is immobilized by means of its macromolecular derivative physically entrapped within the gel network. Such a macromolecular carrier of the ligand may have a

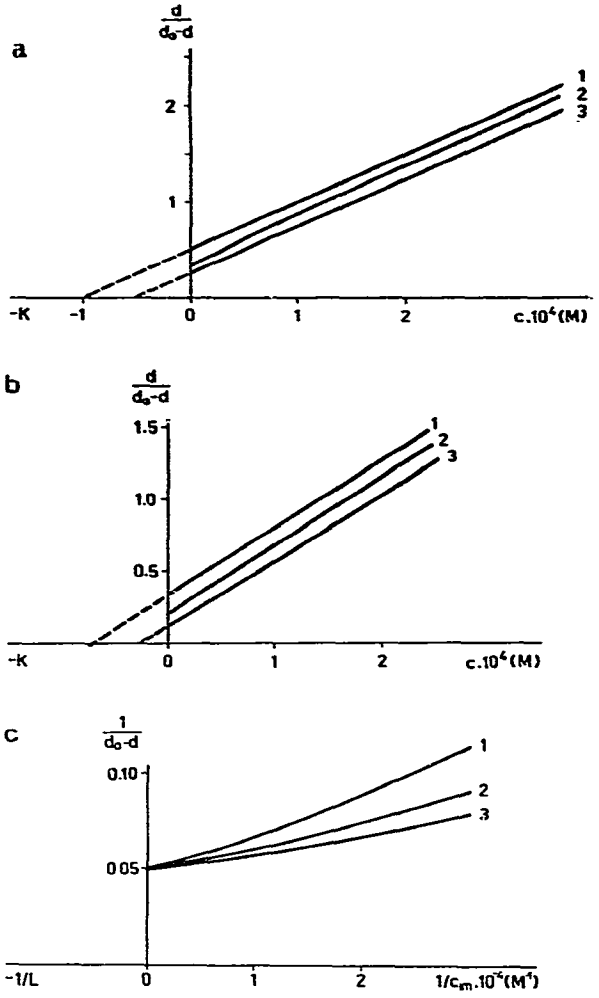


Fig. 3. Comparison of $d/(d_0 - d)$ *vs.* c and $1/(d_0 - d)$ *vs.* $1/c_{im}$ plots for different bivalent proteins with identical sites and a hypothetical bivalent protein with two different binding sites. (a) Curve 1 = bivalent protein with identical sites ($K = L = 10^{-4}$ *M*); curve 3 = bivalent protein with identical sites ($K = L = 10^{-5}$ *M*); curve 2 = bivalent protein with different sites ($K_1 = L_1 = 10^{-4}$ *M*, $K_2 = L_2 = 10^{-5}$ *M*). In all instances $c_{im} = 10^{-4}$ *M*, $p = 0$ (*i.e.*, curve 2 represents eqn. 15). (b) As in (a) but $p = 1$ (*i.e.*, curve 2 represents eqn. 13). (c) As in (b) but $c = 0$ (*K* is not considered) (*i.e.*, curve 2 represents eqn. 14). Note in all instances the intermediate apparent *K* and *L* values obtained from curves 2 compared with curves 1 and 3.

conformation in which some ligand molecules can be "buried" and thus inaccessible to the protein. The knowledge of the effective c_{im} value is essential for the estimation of L in the simple case of a monovalent protein or L/n for multivalent proteins.

In a previous paper[2] eqn. 17 was derived, which relates the mobility of the protein on an affinity gel of c_{im} to the concentration of protein (eqn. 10c in ref. 2).

$$\frac{d}{d_0} = \frac{1}{a} \cdot \frac{dc_{im}}{d_0 - d} - \frac{L}{a} \qquad (17)$$

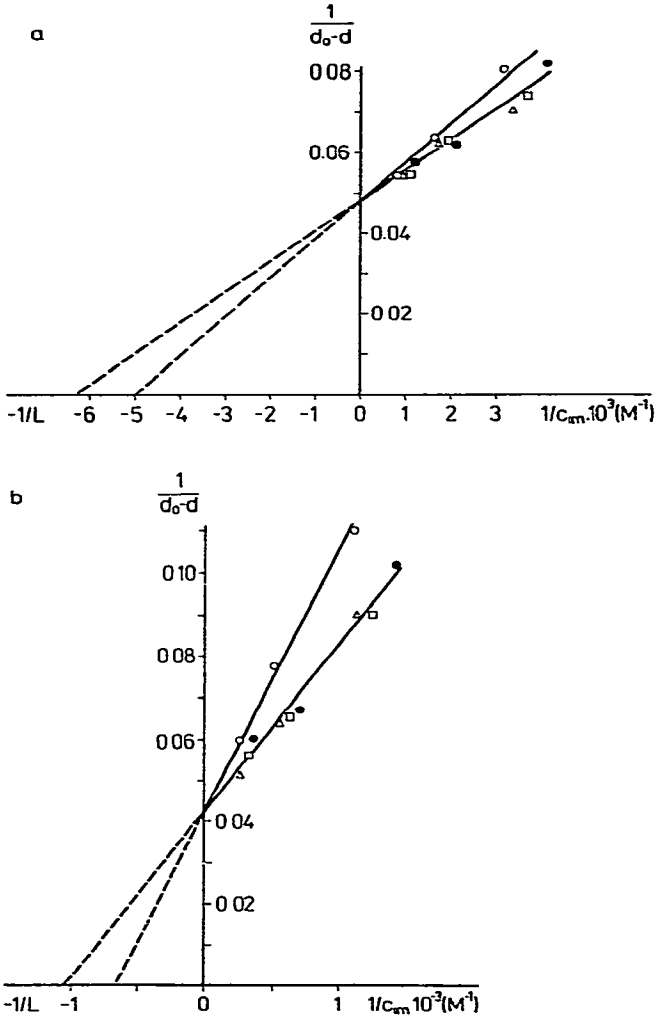


Fig. 4. Affinity electrophoresis of (a) con A and (b) *Lathyrus sativus* lectin on affinity gels prepared from α-D-mannosyl polyacrylamide copolymers with different mannose contents. ●, Copolymer containing 1.65% of mannose; □, copolymer containing 3% of mannose; △, copolymer containing 6.3% of mannose; ○, copolymer containing 16.8% of mannose.

Simple rearrangement of this equation yields directly eqn. 17a, which can be used for the calculation of the effective c_{im}:

$$\frac{d}{d_0 - d} = \frac{1}{c_{im}} \cdot \frac{da}{d_0} + \frac{L}{c_{im}} \tag{17a}$$

Thus, the effective c_{im} can be obtained from the experimentally determined dependence of mobility d *versus* protein concentration a at a fixed c_{im} by plotting $d/(d_0 - d)$

vs. da/d_0; the effective c_{im} is obtained as the tangent of this linear plot and in addition the "true" L value (relating to effective c_{im}) is obtained as the intercept with the abscissa, provided that the protein is monovalent. A similar relationship has been described for the estimation of the effective c_{im} value in quantitative affinity chromatography[14].

However, eqn. 17a implies a very simplified situation, *i.e.*, the presence of a fraction of fully accessible immobilized ligand molecules and the remainder completely inaccessible molecules. In reality, there will be a full range of immobilized ligand molecules of different accessibilities, yielding complexes with the protein of widely differing L and in some instances even chemically different types of immobilized ligand may be present that interact differently with the protein. It is therefore of interest to establish the effect of this heterogeneity of immobilized ligand molecules. This situation can be described as follows: an affinity gel contains ligand λ_1 of concentration $(c_{im})_1$; its complexing with monovalent protein is characterized by a dissociation constant L_1; λ_2 [concentration $(c_{im})_2$], etc., and generally λ_i [concentration $(c_{im})_i$]. The total concentration of all ligands is $c_{im} = \sum_{i=1}^{n} (c_{im})_i$; thus, $(c_{im})_i = f_i c_{im}$, where f_i denotes the fraction of all ligand molecules corresponding to λ_i and $\sum_i f_i = 1$. The c_{im} value is assumed to be known (an effective value). Then, the following simple equation will describe the equilibrium state, assuming that $a \ll (c_{im})_i$ for every i:

$$\frac{d}{d_0 - d} = \frac{1}{\sum_{i=1}^{n} \frac{(c_{im})_i}{L_i}} = \frac{1}{\sum_{i=1}^{n} \frac{f_i}{L_i}} \cdot \frac{1}{c_{im}}$$

$$\frac{d}{d_0 - d} = L_{eff} \cdot \frac{1}{c_{im}} \qquad (18)$$

Obviously, heterogeneity of immobilized ligand molecules does not change the character of the $1/(d_0 - d)$ *vs.* $1/c_{im}$ plot, but instead of the true L an effective value L_{eff} is obtained; its numerical value will depend on the degree of heterogeneity of the immobilized ligand molecules and on the proportions of individual species of the ligand.

Experimental consequences of eqns. 17a and 18

Eqn. 17a is experimentally important, as it allows the simple estimation of the effective c_{im} from the observable dependence of d on a. This relationship should lead to the simultaneous estimation of the true L and c_{im} values; the true L value is important for the quantitative evaluation of protein–immobilized ligand interactions, whereas the effective c_{im} (and its comparison with the total c_{im}) provides an interesting insight into steric conditions within the gel. It should be noted that these values (true L and effective c_{im}) are meaningful only if one type of immobilized ligand greatly predominates; when there is extensive heterogeneity of the immobilized ligand with respect to the strength of interaction with the protein, these values again represent some mean values and it is not possible to obtain any more detailed information about them (eqn. 18).

For the experimental estimation of the effective c_{im} it is necessary to know only the concentration of protein in the moving zone; thus it is necessary to prevent any concentration or dilution effects during electrophoresis, such as stacking or diffusion. The simplest experimental design would possess the following features: the use of a continuous electrophoretic system (a discontinuous system would produce stacking) in a porous gel (to avoid any complicating sieving effects), and relatively large sample volume to prevent excessive dilution by diffusion, at least in the central part of the zone, *i.e.*, conditions similar to those in frontal affinity chromatography[14]. However, it is necessary to account for "sharpening" of the interacting protein bands on affinity gels, which occurs when the sample enters the gel and leads to concentration of the protein. The degree of sharpening, *i.e.*, the ratio of the band width (w) in the affinity gel ($c_{im} < 0$, $c = 0$) to the original width (w_0) of the sample layer for a monovalent protein is

$$f = \frac{w}{w_0} = \frac{d}{d_0} = \frac{L}{c_{im} + L} \tag{19}$$

The degree of protein concentration is of course, $1/f$; *e.g.*, for $c_{im} = 2L$, which yields $d = 1/3d_0$, the concentration of the protein within the zone moving in the gel is three times higher than the concentration in the sample layer. If this affinity-concentrating effect is to be avoided, the sample itself must be present from the very beginning in a medium of composition identical with that of the affinity gel. This can easily be achieved by dissolving the monovalent protein in a polymerization mixture normally used for the preparation of the affinity gel (buffered acrylamide and bisacrylamide solution containing a soluble macromolecular derivative of the ligand and a catalyst). After polymerization, the protein is thus present from the beginning in the affinity gel and after applying the voltage no concentrating effect can occur; the band width and therefore also the protein concentration should remain relatively constant (except for both edges of the band, where diffusion will cause a decrease in a); d (as a function of c_{im}) could be measured from the position of either the leading or the tailing edge.

From the experimental point of view, monovalent proteins interacting relatively strongly with the immobilized ligand ($L = 10^{-4}$–10^{-6} M) would be most suitable, such as lysozyme [interaction with immobilized (D-GlcNAc)$_n$], some nucleases (interaction with immobilized nucleotide derivatives[15]) or monovalent Fab fragments of monoclonal immunoglobulins (interaction with immobilized hapten[12]). Unfortunately, no suitably reliable monovalent lectins (our preferred test-proteins) are available at present, so that we could not use these in affinity electrohoresis for the purpose of the estimation the effective c_{im} in the glycosyl affinity gels. We are currently studying the application of some of the mentioned systems for the estimation of effective c_{im} values.

Eqn. 18 again demonstrates that even if the effective c_{im} can be determined, the L value may be poorly defined owing to heterogeneity of immobilized ligands, whereas K can be determined exactly.

Microdistribution of immobilized ligand molecules

The affinity polyacrylamide gels can be prepared either by direct copolymeri-

zation-mediated incorporation of a suitable unsaturated derivative of the ligand (*e.g.*, an allyl or acryloyl derivative) or by incorporation of a macromolecular derivative of the ligand into the polyacrylamide gel network[1]. Presumably, the former way should lead to an approximately homogeneous spatial distribution of immobilized ligand molecules, whereas affinity gels prepared in the latter way should contain "islet-like" distributed ligand molecules, because the macromolecule used for the immobilization usually contains many ligand molecules; outside these islets there will be large regions completely devoid of any ligand. During the evaluation of effects of this inhomogeneous distribution of immobilized ligands, several factors have to be taken into account: the size of ligand-carrier macromolecule and the density of its substitution with ligand molecules; the conformation of the macromolecular carrier, determining the accessibility of the ligand molecules to the protein molecules; and the size of protein molecule. All of these factors may influence both the effective c_{im} and the possibility of multivalent interactions in the case of a multivalent protein. It is obvious that local clustering of immobilized ligand molecules may support the possibility of local multivalent interactions, *i.e.*, increase the p value; on the other hand the ligand-free regions outside the ligand clusters will behave as a completely non-interacting gel. These two effects should clearly act in opposite directions and it is difficult to assess on a qualitative basis which one will prevail. If the macromolecular carrier possesses some compact conformation, most of the ligand molecules can be buried in the interior of the macromolecule, so that the effective c_{im} is much lower than the total c_{im} and multivalent interactions are negligible. Owing to the obvious complexity of these effects, we were not able to find any quantitative relationship to describe them and therefore we carried out only preliminary experimental tests on this problem.

We performed affinity electrophoresis of *Lathyrus sativus* lectin and concanavalin A on affinity gels prepared by incorporation of α-D-mannosyl polyacrylamide copolymers with various sugar contents, *i.e.*, differing in the number of carbohydrate units per macromolecule (Fig. 4). We did not find any observable dependence of the apparent strength of interaction on the sugar content of the copolymers; thus the effective c_{im} was probably not affected by the degree of "islet-ness" of the ligand distribution under these conditions. Only for the copolymer with the highest sugar content was the apparent strength of interaction slightly decreased, indicating that a relatively larger proportion of ligand molecules was inaccessible to interaction with the protein, presumably for steric reasons. At least in this instance the effect of a decreased effective c_{im} probably prevailed over the increased local possibility of bivalent interactions. A more exact evaluation of this experiment, however, must await the direct determination of the effective c_{im} as discussed above.

CONCLUSIONS

These and previous results[2] represent the theoretical basis of affinity electrophoresis. Thus, the effects of impairment of several assumptions (assumptions 1–8 in the Introduction) originally restricting the range of applicability of this method were evaluated in detail. However, several points still remain to be clarified:

(1) It should be possible to formulate a general equation of affinity electrophoresis that would include all factors dealt with so far separately. However, such an equation would be certainly extremely complex and cumbersome for practical use. It

is nearly always possible to arrange the experimental conditions so that some of the simpler equations can be used adequately.

(2) A theoretical description of the affinity electrophoresis of a multivalent (or at least bivalent) protein with cooperative binding sites is lacking at present and it might be interesting to work it out in the future.

(3) It is necessary to verify experimentally the equations describing the affinity electrophoresis of multivalent proteins under different conditions and to estimate the effective c_{im} in suitable systems as suggested above.

(4) It is of great importance to examine thoroughly the conditions that would finally permit the determination of the number of ligand-binding sites in a protein molecule by affinity electrophoresis.

These theoretical conclusions, together with some technical improvements of the method[1], should contribute to the wider applicability of affinity electrophoresis.

SYMBOLS

- a Total concentration of protein within the moving zone (considered as constant within the zone; the effects of diffusion are neglected).
- [A] Equilibrium concentration of free protein (M).
- c Concentration of mobile ligand in the affinity gel (M).
- c_{im} Concentration of immobilized ligand in the affinity gel (M).
- d Distance travelled by the protein band from the start during the whole electrophoretic experiment (time t_0) in the affinity gel (mm).
- d_0 Distance travelled by the protein band from the start during the whole electrophoretic experiment (time t_0) in the control (non-interacting) gel (mm).
- K Intrinsic (microscopic) dissociation constant of the protein–mobile ligand complex (*i.e.*, dissociation constant of the binding site–ligand complex) (M).
- L Intrinsic (microscopic) dissociation constant of the protein–immobilized ligand complex (M).
- n Number of ligand-binding sites in the protein molecule.
- p Parameter characterizing the fraction of immobilized ligand molecules available for steric reasons for the formation of protein complexes with two molecules of immobilized ligand simultaneously ($0 < p < 1$)
- $X_{i,j}$ Complex of a protein with i molecules of immobilized and j molecules of free ligand, respectively ($1 < i < n$, $0 < j < n - i$)
- Y_k Complex of a protein with k molecules of free ligand ($1 < k < n$).

Note several differences in the symbols used in this and the previous paper[2], *e.g.*, c_{im} (instead of c_i), L (instead of K_i). Also, the term "mobile ligand" is used instead of "free ligand" to avoid confusion when discussing complexes of protein with that ligand (previously we used the potentially confusing term "protein-free ligand complex"; in fact, a ligand in a complex is no longer "free").

ACKNOWLEDGEMENTS

We are indebted to Dr. J. Hořejší for mathematical advice, to Dr. V. Matoušek for suggesting us the use of eqn. 12 for the calculation of p, to Drs. F. Franěk and J. Turková for careful critical reading of the manuscript, to Drs. P. Sedmera and F.

Tureček for helpful discussions and suggestions, to Mr. J. Josek for help during the preparation of the manuscript and to Mrs. L. Vyhnálková for careful typing of the manuscript.

REFERENCES

1 V. Hořejší, *Anal. Biochem.*, 112 (1981) 1–12.
2 V. Hořejší, *J. Chromatogr.*, 178 (1979) 1–13.
3 V. Hořejší, in J.-M. Egly (Editor), *Proceedings of the Third International Symposium on Affinity Chromatography and Molecular Interactions*, INSERM Symposium Series, INSERM, Paris, 1979, pp. 391–398.
4 V. Hořejší, P. Smolek and J. Kocourek, *Biochim. Biophys. Acta*, 538 (1978) 293–298.
5 M. Dubois, K. A. Gilles, J. K. Hamilton, P. A. Rebers and F. Smith, *Anal. Chem.*, 28 (1956) 350–356.
6 G. Louženský, *Isolation of Hemagglutinins from Seeds of the Sweet Pea (Lathyrus sativus L. var. Karcagi)*, *Thesis*, Faculty of Natural Sciences, Charles University, Prague, 1970.
7 B. B. L. Agrawal and I. J. Goldstein, *Biochim. Biophys. Acta*, 147 (1967) 262–267.
8 R. A. Reisfeld, V. S. Lewis and D. E. Williams, *Nature (London)*, 195 (1962) 281–283.
9 V. Hořejší, M. Tichá and J. Kocourek, *Biochim. Biophys. Acta*, 499 (1977) 290–300.
10 G. M. Edelman, B. A. Cunningham, G. N. Reeke, J. W. Becker, M. J. Waxdal and J. L. Wang, *Proc. Nat. Acad. Sci. U.S.*, 69 (1972) 2580–2584.
11 H. Metzger, *Advan. Immunol.*, 12 (1970) 57–116.
12 D. Eilat and I. M. Chaiken, *Biochemistry*, 18 (1979) 790–794.
13 C. Borrebaeck and M. E. Etzler, *FEBS Lett.*, 117 (1980) 237–240.
14 K. Kasai and S. Ishii, *J. Biochem.*, 84 (1978) 1051–1060.
15 I. M. Chaiken and H. C. Taylor, *J. Biol. Chem.*, 251 (1976) 2044–2048.