Proc. Nat. Acad. Sci. USA
Vol. 72, No. 12, pp. 4840–4843, December 1975
Biochemistry

Statistical theory of chromatography: New outlooks for affinity chromatography

(plate theory/distribution profile/moments/variance peak width/fast reactions)

F. C. DENIZOT AND M. A. DELAAGE

Centre de Biochimie et Biologie Moléculaire du Centre National de la Recherche Scientifique 31, chemin Joseph Aiguier, 13274 Marseille Cédex 2, France

Communicated by Henry Eyring, September 22, 1975

ABSTRACT We have developed further the statistical approach to chromatography initiated by Giddings and Eyring, and applied it to affinity chromatography. By means of a convenient expression of moments the convergence towards the Laplace-Gauss distribution has been established. The Gaussian character is not preserved if other causes of dispersion are taken into account, but expressions of moments can be obtained in a generalized form. A simple procedure is deduced for expressing the fundamental constants of the model in terms of purely experimental quantities. Thus, affinity chromatography can be used to determine rate constants of association and dissociation in a range considered as the domain of the stopped-flow methods.

In spite of a long history of experimental research and extensive applications, chromatography was not considered as a random process till 1955, when Giddings and Eyring published their remarkable statistical theory (1). It was characterized by realistic basic assumptions and by an elegant development leading to an elution profile in the form of a modified Bessel function. However, the old theory of Martin and Synge (2) has continued to be used, in part because of the strength of custom, in part because no direct connection between the two theories had been established, which would permit the conversion of the data.

In the present work we have again taken the statistical approach and studied thoroughly the following topics: (*i*) The transformation of the expressions of moments and the convergence of the distribution law towards the Laplace-Gauss distribution, (*ii*) A generalization taking into account other causes of dispersion (diffusion, etc.), which leads to an expression of the fundamental constants of the theory in terms of purely experimental quantities.

These interpretations open a new field of applications for affinity chromatography (3) which makes it possible to obtain rapid kinetic data on noncovalent binding, which are very difficult to obtain otherwise.

Random walk of a molecule in a chromatographic support

$k\,dt$ designates the probability of binding a molecule during elementary time dt: $k'dt$ is the probability of releasing a bound molecule during dt. Elution is obtained when the molecule has spent a definite time t_0 in the free state. The retention time t is a stochastic variable, the law of which is determined as follows, according to Giddings and Eyring (1):

The probability $p(n,t)dt$ of having exactly n stays in the fixed state and a retention time lying between t and $t + dt$ is given by the product of three probabilities of independent events:

n bindings during the time t_0—it is a Poisson process; the probability p_n is given by:

$$p_n = \frac{e^{-kt_0}(kt_0)^n}{n!} \qquad [1]$$

n-1 releasings during the time t—it is still a Poisson process; the probability is given by:

$$\frac{e^{-k't}(k't)^{n-1}}{(n-1)!} \qquad [2]$$

One releasing between t and $t + dt$—the probability is:

$$k'dt$$

Hence the probability $p(n,t)dt$ is:

$$p(n,t)dt = \frac{e^{-kt_0}(kt_0)^n e^{-k't}(k't)^{n-1}k'dt}{n!(n-1)!} \qquad [3]$$

The probability pdt of having a retention time between t and $t + dt$ for any n is obtained by summation with respect to n:

$$pdt = \sum_{n=1}^{\infty} \frac{e^{-kt_0}(kt_0)^n e^{-k't}(k't)^{n-1}k'dt}{n!(n-1)!} \qquad [4]$$

This density is associated with a Dirac distribution at the origin, which expresses that non-retention always has a nonzero probability:

$$\delta e^{-kt_0} \qquad [5]$$

The expression [4] can be written using a modified Bessel function of I_1 type [see ref. (4)].

By plotting p versus t we obtain the elution profile. Fig. 1 shows some examples which will be discussed further.

General expression for moments

Let us calculate the rth moment of the distribution of t: With $r > 0$ the Dirac distribution at the origin does not contribute. Thus:

$$E(t^r) = \int_0^{\infty} t^r p\,dt = \sum_{n=1}^{\infty} \frac{e^{-kt_0}(kt_0)^n k'^n}{n!(n-1)!} \int_0^{\infty} t^{n-1+r} e^{-k't}dt$$

After integration:

$$E(t^r) = \sum_{n=1}^{\infty} \frac{e^{-kt_0}(kt_0)^n (n-1+r)!}{k'^r n!(n-1)!} \qquad [6]$$

This series can be expressed otherwise:

$$E(t^r) = \frac{e^{-kt_0}}{k'^r} \frac{\partial^{r-1}}{\partial(kt_0)^{r-1}} (e^{kt_0}(kt_0)^r) \qquad [7]$$

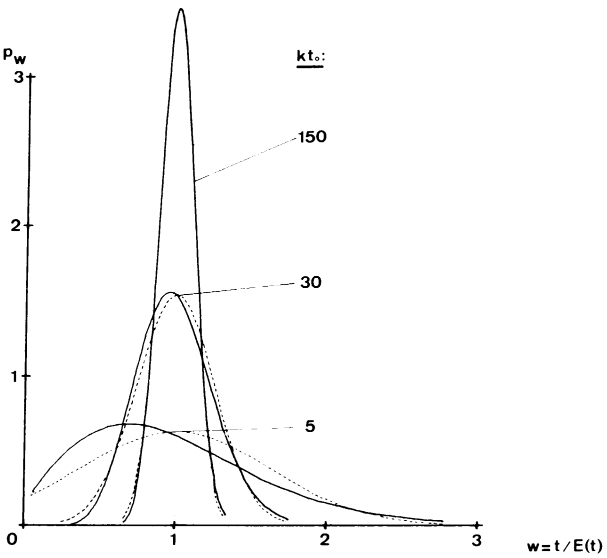

FIG. 1. Elution profiles (full lines) in reduced coordinates showing the effect of the mean number of exchanges (kt_0) on the thickness of the peaks and the tendency towards the Gaussian distribution (broken lines). The curves were computed by numerical summation of the series [13]: first, the value of n giving the highest term was calculated as being the integer the nearest to $\sqrt{kk'tt_0}$, let this be n_0; then the terms were added in the order n_0, $n_0 + 1$, $n_0 - 1$, $n_0 + 2$, $n_0 - 2$, etc. . . until they became negligible.

The $(r - 1)$th derivative can be written in the form of a finite sum, then, the exponential factors cancel each other, yielding:

$$E(t^r) = \frac{r!(r-1)!}{k'^r} \sum_{j=0}^{r-1} \frac{(kt_0)^{r-j}}{j!(r-j)!(r-j-1)!} \qquad [8]$$

The expression [8] offers a great advantage over those previously derived (1, 5), having a finite number of terms without an exponential factor. The approximation used by Giddings and Eyring gives the same result as [8] for the first two moments, a discrepancy appears only with the third one.

Distribution profiles

The first moment, average of the distribution is:

$$E(t) = \frac{kt_0}{k'} \qquad [9]$$

It represents, in experimental terms, the position of the peak. The second moment is:

$$E(t^2) = \frac{kt_0(kt_0 + 2)}{k'^2} \qquad [10]$$

Hence the variance σ^2 is

$$\sigma^2 = \frac{2kt_0}{k'^2} \qquad [11]$$

σ, the standard deviation, represents the width of the peak. As expected, $E(t)$ is proportional to t_0, therefore to the length of the column, and also to k/k', the equilibrium constant of binding. σ varies as the square root of the length of the column.

All the possible profiles can be represented by a network of curves depending on one parameter. Let us consider the dimensionless ratio: $w = t/E(t)$.

Its distribution law has for density p_w:

$$p_u = \frac{kt_0}{k'} p \qquad [12]$$

$$p_u = \sum_{n=1}^{\infty} \frac{e^{-kt_0}(kt_0)^{n+1} e^{-kt_0 w}(kt_0 w)^{n-1}}{n!(n-1)!} \qquad [13]$$

The only parameter remaining is the mean exchange number kt_0. The representation p_w is equivalent to taking the end of the first void volume as origin and the retention volume as unity. Fig. 1 shows the different plots obtained for different values of kt_0. The curves are asymmetric for low values of kt_0 and become symmetric for high values. Fig. 1 also shows the tendency towards the Laplace-Gauss law: for each curve, the Gaussian distribution is shown which has the same mean and the same standard deviation. They become very similar as kt_0 increases.

Convergence towards Laplace-Gauss distribution

The curves presented in Fig. 1 strongly suggest the convergence of the distribution [4] towards a Laplace-Gauss distribution. We are going to demonstrate this important point:

From the moments $E(t^r)$ of the distribution we can write the series expansion of its characteristic function $\phi(\theta)$:

$$\phi(\theta) = 1 + \sum_{r=1}^{\infty} \frac{(i\theta)^r}{r!} E(t^r) \qquad [14]$$

Replacing the moments by their expression [8], it follows:

$$\phi(\theta) = 1 + \sum_{r=1}^{\infty} \frac{(i\theta)^r (r-1)!}{k'^r} \sum_{j=0}^{r-1} \frac{(kt_0)^{r-j}}{j!(r-j)!(r-j-1)!}. \qquad [15]$$

Now we are going to substitute for t the associated reduced variable u:

$$u = \frac{t - E(t)}{\sigma} = \frac{k't}{\sqrt{2kt_0}} - \sqrt{\frac{kt_0}{2}} \qquad [16]$$

This transformation being linear the new characteristic function $\phi_u(\theta)$ is easily deduced:

$$\phi_u(\theta) = e^{-i\theta\sqrt{\frac{kt_0}{2}}} \left[1 + \sum_{r=1}^{\infty} \left(\frac{ik'\theta}{\sqrt{2kt_0}} \right)^r \times \frac{(r-1)!}{k'^r} \sum_{j=0}^{r-1} \frac{(kt_0)^{r-j}}{j!(r-j)!(r-j-1)!} \right] \qquad [17]$$

Rearranging the terms and letting $s = r - j$ one can write $\phi_u(\theta)$ as:

$$\phi_u(\theta) = e^{-i\theta\sqrt{\frac{kt_0}{2}}} \times \left[1 + \sum_{j=0}^{\infty} \frac{\left(i\theta\sqrt{\frac{kt_0}{2}}\right)^j}{j!(kt_0)^j} \sum_{s=1}^{\infty} \frac{\left(i\theta\sqrt{\frac{kt_0}{2}}\right)^s (s+j-1)!}{s!(s-1)!} \right]. \qquad [18]$$

The series with the s index has been already encountered in Eq. [6]; it can be treated in the same way, except for the $j = 0$ value for which the corresponding term must be calculated separately. The exponential factor is thus cancelled:

$$\phi_u(\theta) = 1 + \sum_{j=1}^{\infty} \frac{(-1)^j}{2^j} \frac{\theta^{2j}}{j!} \sum_{h=0}^{j-1} \frac{(-i)^h (j-1)! j! 2^{h/2}}{h!(j-h)!(j-1-h)!(kt_0)^{h/2} \theta^h} \qquad [19]$$

for $kt_0 \to \infty$ only the terms corresponding to $h = 0$ do not vanish, the series [19] tends, term by term, towards a limit series:

$$\lim_{kt_0 \to \infty} \phi_u(\theta) = \sum_{j=0}^{\infty} \frac{(-1)^j \theta^{2j}}{2^j (j)!} = e^{-\theta^2/2} \qquad [20]$$

which is the characteristic function of a Laplace-Gauss distribution. In practice one can consider that for kt_0 values higher than 100, for instance, the law [4] can be replaced by the associated Gaussian distribution:

$$\frac{k'}{2\sqrt{\pi k t_0}} e^{-\frac{(k't - kt_0)^2}{4kt_0}} \qquad [21]$$

Correspondence with the theory of Martin and Synge

The plate theory of Martin and Synge (2) has widely been accepted and many results are expressed in terms of "height of the theoretical plate." We have to orient both theories with respect to each other and to indicate the way of connecting them.

Replacing, in the plate theory, volume by retention time as a variable, we obtain for the first two moments according to the plate theory:

mean: $E(t) = kt_0/k'$, the same result as in the statistical theory for the standard deviation:

$$\sigma = \frac{t_0}{k'} \sqrt{\frac{k(k + k')}{P}} \qquad [22]$$

where P is the number of plates.

We can define the number of plates P, which results in the identity of the first two moments in both theories, then:

$$P = \frac{(k + k')t_0}{2} \qquad [23]$$

Hence, for a well-retained substance ($k \gg k'$) the number of plates is approximately equal to $kt_0/2$, which corresponds to half the mean number of exchanges. This can be better understood when we consider the form of the plate distribution with high retention:

$$\frac{k'}{2}\left(\frac{k't}{2}\right)^{P-1} \frac{e^{(-k't/2)}}{(P - 1)!} dt \qquad [24]$$

[24] has the same form as the conditional law $p(n,t)/p_n$ which is obtained in statistical theory by fixing the number of exchanges equal to P instead of considering it as a random variable. (See Eqs. [1] and [3].) P, however, cannot be equal to the mean actual number of exchanges kt_0; it must assume a lower value to obtain a correct variance.

The correspondence between the two theories cannot be maintained for higher order moments.

Generalization, taking into account other causes of dispersion

The interpretation of experimental data will require that other causes of dispersion than the random character of binding should be taken into account. The main causes are: diffusion in the free state and heterogeneity of the velocity of the solvent: zero inside the beads of the chromatographic support or on their surface, maximum between them. Consequently, it is no longer valid to consider t_0 time as a constant; it is, like t, a random variable. It must be noticed that t and t_0 are *not independent* random variables, as supposed by McQuarrie (5). Only the elution time t', not the retention time t, can be determined directly:

$$t' = t_0 + t \qquad [25]$$

We do not intend to develop a detailed model for the t_0 distribution, nor even to assign a particular law to it. Let us assume, however, that at least the first two moments exist. Eq. [4] is here considered as the conditional law of t, t_0 being fixed. The qth moment of t' is then easily deduced, and expressed as a function of the different moments of the t_0 law:

$$E(t'^q) = E(t_0{}^q) + \sum_{r=1}^{q} \frac{q!(r - 1)!}{(q - r)!k'^r} \sum_{j=0}^{r-1} \frac{k^{r-j} E(t_0{}^{q-j})}{j!(r - j)!(r - 1 - j)!} \qquad [26]$$

From [26] we deduce:

$$\text{mean} = E(t') = E(t_0)\left(1 + \frac{k}{k'}\right) \qquad [27]$$

variance:

$$\sigma'^2 = \frac{2k}{k'^2}E(t_0) + \left(1 + \frac{k}{k'}\right)^2 \sigma_0{}^2 \qquad [28]$$

where $\sigma_0{}^2$ is the variance of t_0.

Experimental determination of k and k'

The above equations open the way to the experimental determination of k and k': A first experiment with the substance under study gives the experimental values $E(t')$ and σ', deduced from the position and the width of the peak, respectively, or by numerical integration.

A second experiment with another molecule, having similar hydrodynamic properties, but no affinity for binding sites, gives $E(t_0)$ and σ_0.

Another way, especially suitable for affinity chromatography, consists in performing the first chromatography with a small amount of the labeled molecule, alone, and the second with the same sample, plus, in the solvent, a large amount of nonradioactive substance in order to saturate the binding sites.

Then, the Eqs. [27] and [28] are solved with respect to k and k':

$$k = \frac{2E(t_0)[E(t') - E(t_0)]^2}{\sigma'^2 E^2(t_0) - \sigma_0{}^2 E^2(t')} \qquad [29]$$

$$k' = \frac{2E^2(t_0)[E(t') - E(t_0)]}{\sigma'^2 E^2(t_0) - \sigma_0{}^2 E^2(t')} \qquad [30]$$

More general initial conditions

This case can be described by adding to the expression of elution time t' a random term t_1 representing the input time:

$$t' = t + t_0 + t_1 \qquad [31]$$

We only assume that the distribution of the input time t_1 has at least two moments. The origin on the time axis can be chosen in such a way that $E(t_1) = 0$. t_1 is independent of t and t_0 and the generalized expression of σ'^2, and is nearly obtained by adding $\sigma_1{}^2$, the variance of t_1, to expression [28]:

$$\sigma'^2 = \frac{2k}{k'^2}E(t_0) + \left(1 + \frac{k}{k'}\right)^2 \sigma_0{}^2 + \sigma_1{}^2 \qquad [32]$$

σ_0^2 itself cannot be determined directly. If the same input distribution is used the observed variable is $t'_0 = t_0 + t_1$ and σ_0^2 is calculated from the observed variance σ'^2_0 as:

$$\sigma_0^2 = \sigma'^2_0 - \sigma_1^2 \qquad [33]$$

Finally k and k' are determined by replacing σ'^2 by $\sigma'^2 - \sigma_1^2$ in Eqs. [29] and [30].

Discussion

Affinity chromatography is usually designed for separation purposes: the affinity between the molecule to be retained and the ligand covalently attached to the support is high and elution is obtained by modifying the solvent. Here we deal with low affinity systems, with the aim of determining not only affinity but also the rate constants of association and dissociation from the experimental parameters k and k'. The first approach consists in keeping k' as the time dissociation rate constant and obtaining the second order rate constant of association by dividing k by the estimated concentration of binding sites. The weakness of this procedure is obvious, since there is a nonhomogenous distribution of binding sites: empty spaces of the mobile phase occur together with a higher density of sites inside the support or on its surface. In the case of a poorly diffusing molecule with a high association rate constant the probability of being trapped is different when the molecule has just left a binding site and when it is present in the middle of the mobile phase. This phenomenon is worsened by the heterogeneity of the flow velocity and only partially balanced by the measurement of the t_0 distribution. The following consideration gives an idea, what the conditions are that will make this effect be of critical importance.

Let us assign a value to the mean binding time $1/k$, which gives to the diffusion distribution of the free molecule a standard deviation of the same order of magnitude as the mean thickness of the mobile phase e; that gives:

$$k = \frac{4D}{e^2}$$

where D is the diffusion coefficient.

As an example, for a small molecule $D = 10^{-5}$ cm^2/sec and with an $e = 10^{-3}$ cm, we obtain for the critical value of k, 40 sec^{-1}. If the experimental value of k is lower than the estimated critical value the actual rate constants of the reaction between the partners can be correctly determined. The question whether rate constants determined with one immobilized partner are identical to those between free partners is beyond the scope of this paper, but in most cases they will be very similar.

Curiously, chromatography (a rather slow process) offers the opportunity to determine rate constants which are in the range of stopped-flow methods. A similar situation occurs in experiments of sedimentation velocity of reacting systems, where the interpretation of sedimentation profiles can lead to the determination of chemical rate constants (6, 7).

In a more general way, this method can be applied outside the field of chromatography: any transport process with random fixation can be treated in the same way: e.g., the displacement of a molecule in the blood capillary with interaction with some receptor in the wall. A statistical interpretation could provide a better understanding of the propagation of hormonal and pharmacological effects.

We are greatly indebted to Dr. C. Goridis for his help in writing the manuscript.

1. Giddings, J. C. & Eyring, H. (1955) *J. Phys. Chem.* **59,** 416–421.
2. Martin, A. J. P. & Synge, R. L. M. (1941) *Biochem. J.* **35,** 1358–1368.
3. Cuatrecasas, P. (1972) *Adv. Enzymol.* **36,** 29–89.
4. Abramowitz, M. & Stegun, I. A. (1965) in *Handbook of Mathematical Functions* (Dover Publications, Inc., New York), p. 375.
5. McQuarrie, D. A. (1963) *J. Chem. Phys.* **38,** 437–445.
6. Cann, J. R. & Oates, D. C. (1973) *Biochemistry* **12,** 1112–1119.
7. Claverie, J. M., Dreux, H. & Cohen, R. (1975) *Biopolymers* **14,** 1685–1700.