Journal of Chromatography, 365 (1986) 205–212
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 825

COMPUTER SIMULATION OF PROTEIN ADSORPTION AND CHROMATOGRAPHY

A. A. GORBUNOV*, A. Ye. LUKYANOV, V. A. PASECHNIK and A. V. VAKHRUSHEV
All-Union Research Institute for Highly Pure Biopreparations, Pudozhskaya 7, 197110 Leningrad (U.S.S.R.)

SUMMARY

The adsorption behaviour of protein molecules in the processes of chromatography was studied theoretically. The adsorbent pore model used was a slit of width D, with the protein molecule model being in the form of a spherical particle of radius R. Arranged in a specific pattern on the surface of the model protein molecule were adsorption-active areas (active spots), and adsorption sites were distributed uniformly over the pore walls. The Monte Carlo technique was used to calculate the dependences of the distribution coefficient, K_d, on D, R, the adsorption energy ε, the adsorbent activity σ and the number and arrangement pattern of active areas over the protein model surface. Analytical expressions were obtained for K_d corresponding to specific cases of soluble and membrane proteins.

Separation by chromatography was shown to be possible, in principle, for proteins having various numbers and sizes of exposed active areas. Proteins with no differences other than in the arrangement of active spots were shown to have different K_d values and, as such, are capable of being separated by chromatography. Protein molecules in the process of adsorption were found to exhibit an orientation effect.

The adsorption conditions and adsorbent activity were shown to have a critical influence on protein retention, leading to the possibility of optimizing the separation of proteins by the selection of a suitable adsorbent or eluent composition and temperature variations.

The theoretical results obtained were found to be qualitatively in agreement with experimental data produced by hydrophobic interaction chromatography of proteins.

INTRODUCTION

Chromatographic processes are extensively used to analyse, separate and purify proteins. The method most frequently employed is gel permeation chromatography (GPC), which is realizable in the absence of adsorption interections between the porous material of the column and the molecules being chromatographed. GPC is known to permit the size separation of particles. At the same time, the need frequently arises in practice to separate proteins that are close in size, but where GPC

0021-9673/86/$03.50

techniques are ineffective. Good results are achievable in such instances by using adsorption chromatographic techniques (adsorption, ion exchange, bioaffinity and hydrophobic interaction chromatography).

To be capable of carrying out adsorption chromatographic processes properly, one should know the underlying principles of the adsorption of protein molecules. However, there have been no systematic experimertal or theoretical studies so far on protein adsorption with regard to the molecular structure of proteins. The objective of this work was to develop an approach for describing on a molecular basis the general principles of the adsorption chromatography of protein molecules.

There are several causes of protein adsorption including Van der Waals interactions, hydrogen bonding, interactions by charged groups and hydrophobic and biospecific interactions, etc. We shall consider generalized models in which the nature of the adsorption interactions is left unspecified (the only assumption being for short-range activity of adsorption forces), with the energy of interaction varying over a wide range. Real proteins have a unique spatial structure and are characterized by a definite number and arrangement of adsorption-active areas. We shall consider models in which adsorption-active areas may differ in number, size and arrangement, and investigate the effects that these parameters, and also adsorbent activity, may have on protein adsorption in the process of chromatography.

The proposed theory is thus intended to describe the general principles of the adsorption chromatography of proteins and may be applied to various types of adsorption chromatography. We shall discuss the theoretical results in relation to the experimental data available on the hydrophobic interaction (reversed-phase) chromatography of proteins. Based on the interaction between protein molecules and adsorbents modified with hydrocarbon radicals of specified length or with some other non-polar compounds, this method of chromatography[1] has gained wide acceptance in recent years for biological studies, and there are available in this field some systematic studies[1–10] of which the results are comparable with theory.

MODEL AND CALCULATION METHOD

A slit of width D was used as an adsorbent pore model and a spherical particle of radius $R < D/2$ as a protein molecule model (Fig. 1a). It was assumed that a certain number n of adsorption-active (*e.g.*, hydrophobic) areas were present on the protein molecule surface, and these were modelled as circular spots of radius r. The adsorption sites on the pore walls were assumed to be distributed uniformly with a surface density σ.

The interaction was specified by the parameter ε, *i.e.*, by the change in free energy on a unit area contact being formed between the active parts of protein and adsorbent. The case under consideration was that of short-range adsorption forces with a characteristic radius of action $\Delta \ll R$ (hydrogen bonding, Van der Waals and hydrophobic interactions). It was assumed that the contact formed as the active spots on the protein closed with the pore walls to a distance of $z < \Delta$ (Fig. 1b). The energy due to the interaction of molecule and adsorbent on the formation of the contact was assumed to be equal to $\varepsilon\sigma\tilde{s}$, where $\tilde{s}$ is the total area of all active protein surface elements in contact.

The protein concentrations in the solution and on the surface were assumed

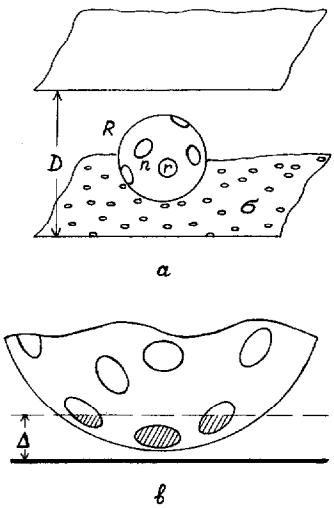

Fig. 1. (a) A model for a protein molecule in an adsorbent pore. (b) A protein molecule fragment at the pore wall (cross-hatched areas are the protein surface elements contributing to the protein wall adsorption interaction). Model parameters shown: pore width D, protein molecule radius R, number, n, of active spots of radius r, adsorbent surface fraction occupied by adsorption sites σ and adsorption forces radius of action Δ.

to be low enough, and the interactions between protein molecules were ignored.

A distribution coefficient, K_d, was calculated, equal to the ratio of the equilibrium protein concentration, C_p, in the chromatographic stationary phase (within adsorbent pores) to the concentration C_0 in the mobile phase. The distribution coefficient is related to the change in the free energy of the protein molecule, ΔF, incident upon a molecule passing from the solution into the sorbent pore, and is given by

$$K_d = \frac{C_p}{C_0} = \exp\left(-\frac{\Delta F}{kT}\right) \tag{1}$$

where k is the Boltzmann constant and T is the absolute temperature. The distribution coefficient can be determined experimentally from the peak positions in chromatograms and also measured in equilibrium adsorption experiments.

The calculation of the distribution coefficient consists in calculating the partition function for the protein molecule in the pore and in the unrestricted space, leading to the following equation:

$$K_d = W^{-1}\int\int d^3r\,d\Omega \exp\left[-\frac{\varepsilon\sigma\tilde{s}(\vec{r},\Omega)}{kT}\right] \tag{2}$$

where $\tilde{s}(\vec{r},\Omega)$ is the total adsorption contact area related to the position and orientation of the protein in the pore, and integration being performed for all possible protein molecule translations and rotations within the pore; W is the volume of the respective translation-and-orientation space.

Integration of eqn. 2 was successful for some specific cases of regular spot arrangement patterns on the protein. The respective equations for K_d will be given later.

For an arbitrary spot arrangement, as a general case, K_d was calculated numerically, using the Monte Carlo technique. With the parameters D, R, r, n, σ, Δ and ε specified, a computer was first used to simulate the model protein surface, then the calculation process was commenced. Using a random number generator, a configuration was specified, *i.e.*, the position and orientation of the protein in the pore, to calculate $\tilde{s}$ and the exponent in eqn. 2. Then another random configuration was utilized, and so on. The process was repeated (up to 25 000 times) until the numerical evaluation of integral 2 converged to a definite value. Also computed were the more probable types of contact between protein molecules and adsorbent.

RESULTS

Two specific cases may be considered, the first being a qualitative chromatographic model for soluble proteins and the other that for proteins included in the cellular membrane composition.

(1) A small number, n, of identical small-sized (such as hydrophobic amino acids) and remotely spaced adsorption-active areas are available on the protein molecule surface. It is assumed that not more than one adsorption-active area may be in contact with the adsorbent (single-site contact). For such a case we were able to obtain an analytical expression for K_d of the form

$$K_d \approx K_0 + \frac{2\Delta}{D} \cdot \frac{S_0 n}{S_M} \cdot \frac{\exp(-E) - 1}{2} \tag{3}$$

where $K_0 = 1 - (2R/D)$ is the distribution coefficient without adsorption; $S_M = 4\pi R^2$ is the protein molecule surface area; $S_0 = 2\pi R\Delta$ is the maximum protein molecule-to-adsorbent contact area; $S = \pi r^2$ is the active group area; and $E = \sigma S\varepsilon/kT$.

(2) The protein surface has large areas continuously covered with adsorption-active groups. The following expression was obtained for this case:

$$K_d \approx K_0 + \frac{2\Delta}{D} \cdot \frac{Sn}{S_M} \cdot \left[\frac{1 - \exp(-E_0)}{E_0} - 1\right] \tag{4}$$

where $E_0 = \sigma S_0\varepsilon/kT$.

At $\varepsilon = 0$, eqns. 3 and 4 yield the well known result

$$K_d^{GPC} = K_0 = 1 - \frac{2R}{D} \tag{5}$$

which implies that under GPC conditions protein retention is dependent on the molecule-to-pore size ratio. The pore size dependence of K_d is universal: from eqns. 2–4 it follows that the value of $K_d - 1$ always varies inversely with D. A simple molecule size dependence of retention exists only in the case of GPC. In adsorption chro-

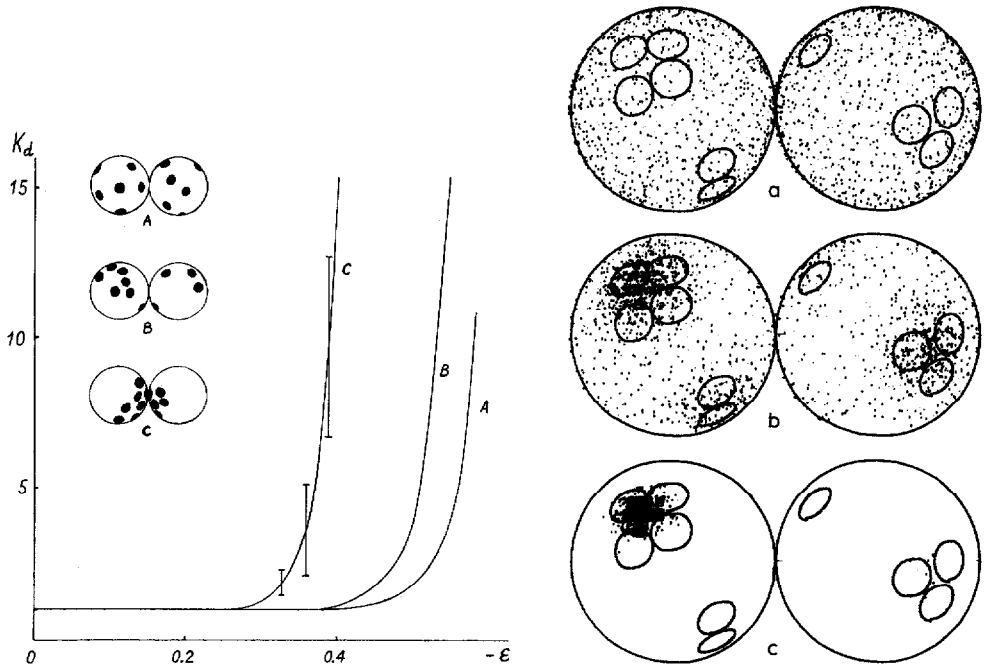

Fig. 2. Dependence of distribution coefficient, K_d on adsorption energy, ε, for three model proteins, A, B and C, with model protein surface charts shown on the left. Parameter values: D = 100 nm; R = 1 nm; Δ = 0.2 nm; r = 0.2 nm; n = 10; σ = 0.5.

Fig. 3. Model protein-to-adsorbent contact probability charts for energy values ε = (a) 0, (b) −0.15 and (c) −0.4. Parameter values: D = 100 nm; R = 1 nm; Δ = 0.2 nm; r = 0.2 nm; σ = 1.

matography there is apparently no definite relationship between K_d and R (molecular weight). As may be seen from eqns. 3 and 4, K_d depends not only on R, but also on the adsorption properties of the protein surface, while there are no correlations between size and surface structures established for real proteins.

It follows from eqns. 3 and 4 that for membrane and soluble proteins K_d is linearly dependent on the number of adsorption-active areas on the protein surface, n, and hence proteins of equal size but differing in the number of active groups on the surface are capable of being separated by adsorption chromatography.

It is of interest that with membrane proteins retention is dependent on the total active protein surface area ns (see eqn. 4), whereas for soluble proteins there is no such universal relationship between K_d and ns (eqn. 3). Soluble protein retention is not only dependent on the total active protein surface and the number of active areas present, but also to a large extent on the size of the active areas.

The difference between eqns. 3 and 4 signifies that the distribution coefficient also depends on the mutual arrangement of active spots on the protein surface. We investigated this effect in greater detail, using the Monte Carlo method to model protein adsorption. Fig. 2 shows the dependence of K_d on the adsorption energy ε for three model proteins differing in the arrangement of active groups only (the number of such groups and their size being identical). As can be seen from Fig. 2, the

distribution coefficients of such proteins differ considerably where the attraction energies involved are high enough. The reason for this effect lies in the local active spot density of the surfaces of the three model proteins being at variance. As shown by Fig. 2, protein A has active spots distributed in a more uniform pattern whereas with protein C there are areas on the protein surface where active spots are completely absent, and other areas where they are arrayed with a greater density than in A and B. As they are areas having a maximal density of active spots that contribute most to partition function 2, protein C is adsorbed more readily than proteins A and B (Fig. 2) as it has the more heterogeneous surface. It has thus been proved that separation by chromatography is possible for proteins differing in no other respect but the arrangement of active groups.

Interestingly, proteins with non-uniform patterns of arrangement of active groups on their surfaces are oriented in a specific fashion relative to the adsorbing surface.

Fig. 3 gives adsorbent contact probability charts for a model protein with a surface having clusters of active areas of different number. The density of points in any area on the surface charts (Fig. 3) is proportional to the probability of this protein surface area contacting the adsorbent. At low adsorption energies, protein-to-adsorbent contacts occur whatever the orientation of the protein molecule (Fig. 3a). As the attraction energy ($-\varepsilon$) is increased, the contact probability distribution will initially become similar to the active spot distribution (Fig. 3b). At higher energies, the protein will have only one side adhering to the adsorbent, *viz.*, that side where the active-group density is at its maximum (Fig. 3c).

The data in Figs. 2 and 3 show that for the adsorption of proteins having a non-uniform pattern of distribution of active spots the essential factor is not the mean density of such spots on the protein surface, but the presence of regions where they are "condensed". These regions act as a determinant not only in binding proteins to adsorbents, but apparently also in numerous biological functions of proteins, *viz.*, their mutual recognition, ability to form multi-enzyme complexes and intracellular organelles, interactions of proteins with cellular membranes, etc.

We shall now consider the dependence of K_d on the degree of adsorbent modification. It follows from eqns. 3 and 4 that K_d increases with increasing σ (active adsorbent surface fraction) in a near-exponential relationship. Fig. 4 shows two K_d *versus* σ plots calculated numerically for two model proteins differing in adsorption-capable group number. As can be seen, at low σ values the distribution coefficient is almost independent of σ and identical for the two proteins: there is no protein sorption. With σ increasing above a certain critical value, adsorption begins, starting with the protein having the larger number of active groups, with the other protein following suit. At high σ values, $K_d \gg 1$ for both proteins, implying almost irreversible adsorption. It therefore follows from Fig. 4 that by varying the degree of adsorbent modification one can select optimal conditions for separating the two proteins, when one of the proteins is not yet adsorbed, having $K_d \approx 1$, while the other is reversibly adsorbed and has $K_d > 1$. Such optimal conditions will, of course, vary with various proteins to be chromatographed. It is therefore practicable to have a set of adsorbents with different adsorption activities, or to be able to vary the protein adsorption conditions (change the ε parameter) by using multi-component eluents or gradient chromatographic programmes. The K_d *versus* ε relationship (Fig. 2) is similar to the K_d

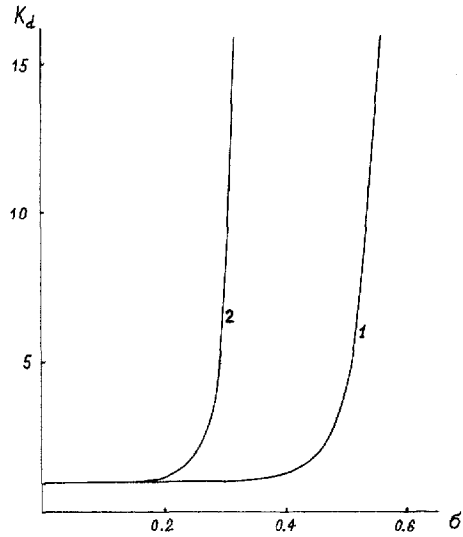

Fig. 4. Dependence of distribution coefficient, K_d, on adsorbent activity, σ, for two model proteins with a random arrangements of adsorption-active areas. Adsorption energy $\varepsilon = -0.075$. n = (1) 20 and (2) 50; values of other parameters as in Fig. 2.

versus σ relationship, and for this reason optimal protein separation is achievable by varying the chromatographic conditions through changes in ε, with the adsorbent remaining unchanged, and also by changing σ.

DISCUSSION

The regularities of protein adsorption chromatography, using hydrocarbon-modified sorbents, have been studied experimentally[1–10]. The interactions between protein molecules and adsorbent were varied in these studies by various methods: adding non-polar organic components or detergents to the eluent[1–3,6,10], varying the temperature[10], or changing the length[1–10] and amount[3] of alkyl chains linked to the sorbent matrix.

Whatever the method used for varying adsorption interaction, there was always a critical effect to be observed: a dramatic increase in retention occurred on reaching certain conditions with a subsequent slight variation. Most workers observed this effect while comparing the same proteins being chromatographed by the use of sorbents modified with hydrocarbons of varying length[1–3,8,9]. The existence of a critical alkyl chain length essential for adsorption to commence, a fact revealed in the studies under discussion[1–10], was initially accounted for based on a hypothesis assuming the presence of hydrophobic voids and pockets[1,2] inside some protein molecules, which can be penetrated only by hydrocarbon chains of predetermined length. It was found later, however, that similar critical effects also occurred when the degree of sorbent modification was increased by the use of hydrocarbons of equal length[2] or when organic compounds were added to the eluent[6,10].

In our model, the parameters used as analogues of experimentally varied quan-

tities are ε and σ. The general form of the K_d *versus* ε and σ plots obtained theoretically (Figs. 2 and 4) is qualitatively in agreement with the respective experimental curves. Specifically, a critical effect is also present in the K_d *versus* ε and σ plots (Figs. 2 and 4), and this was achieved in our model without resorting to the hydrophobic pockets hypothesis. This effect is a result of exponential weighing of adsorbed protein molecule conformations (see eqn. 2) and, as a consequence, must always occur in the adsorption chromatography of any proteins.

There is observable experimentally a certain predetermined sequence of protein separation, characteristic of hydrophobic interaction chromatography, independent of the sorbent and eluent used, and apparently dependent on the properties of proteins themselves. The search for correlations between chromatographic retention and protein properties had attracted much attention in experimental work. It has been attempted to find these correlations by analogy with the hydrophobic chromatography of low-molecular-weight compounds where some correlations had been found [specifically, retention *versus* molecular weight (size) correlations for linear oligomer homologues[6,11,12] or retention *versus* total hydrophobicity for small peptides][6]. No such correlations have been found possible for proteins[5–10].

Based on the results of this theoretical study, one essential factor in determining retention in hydrophobic interaction chromatography of proteins is the presence of hydrophobic groups exposed on the protein molecule surfaces. Thus, protein retention must depend, first and foremost, on the number of such groups or on the hydrophobic protein surface area. Another factor that is equally important, as this work has shown, is the mutual arrangement of the exposed hydrophobic groups. We have been led to suppose that a correlation should be sought between retention and the structural characteristics of protein molecule surface, which characteristics may be obtainable by X-ray analysis[13–16].

REFERENCES

1 S. Shaltiel and Z. Er-el, *Proc. Natl. Acad. Sci. USA*, 70 (1973) 778.
2 S. Shaltiel, in R. Epton (Editor), *Chromatography of Synthetic and Biological Polymers*, Vol. 2, Ellis Horwood, Chichester, 1978, p. 13.
3 H. P. Jenissen and L. M. G. Heilmeyer, *Biochemistry*, 14 (1975) 754.
4 M. Wilchek and T. Miron, *Biochem. Biophys. Res. Commun.*, 72 (1976) 108.
5 B. H. J. Hofstee and N. F. Otillio, *J. Chromatogr.*, 159 (1977) 57.
6 M. J. O'Hare and E. C. Nice, *J. Chromatogr.*, 171 (1979) 209.
7 E. Keshavarz and M. Nakai, *Biochim. Biophys. Acta*, 576 (1979) 269.
8 G. Halperin, M. Breitenbach, M. Tauber-Fikelstein and S. Shaltiel, *J. Chromatogr.*, 215 (1981) 211.
9 E. C. Nice, M. W. Capp, N. Cooke and M. J. O'Hare, *J. Chromatogr.*, 218 (1981) 569.
10 R. A. Barford, B. J. Sliwinski, A. C. Breyer and H. L. Rothbart, *J. Chromatogr.*, 235 (1982) 281.
11 B. L. Karger, J. R. Gant, A. Hartkopf and P. H. Weiner, *J. Chromatogr.*, 128 (1976) 65.
12 G. E. Berendsen and L. De Galan, *J. Chromatogr.*, 196 (1980) 21.
13 B. Lee and F. M. Richards, *J. Mol. Biol.*, 55 (1971) 379.
14 A. Shrake and J. H. Rupley, *J. Mol. Biol.*, 79 (1973) 351.
15 C. Chothia, *Nature (London)*, 254 (1975) 304.
16 H. Meirovitch and H. A. Sheraga, *Macromolecules*, 13 (1980) 1406.