Journal of Chromatography, 444 (1988) 325–344
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1334

PREPARATIVE SEPARATION OF PEPTIDE AND PROTEIN SAMPLES BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY WITH GRADIENT ELUTION

II. EXPERIMENTAL EXAMPLES COMPARED WITH THEORY

G. B. COX and P. E. ANTLE
Medical Products Department, E. I. Du Pont de Nemours, Wilmington, DE 19898 (U.S.A.)
and
L. R. SNYDER*
LC Resources Inc., 26 Silverwood Court, Orinda, CA 94563 (U.S.A.)

SUMMARY

Craig simulations of mass-overloaded gradient elution reported in the preceding paper have been extended to the case of non-Langmuir isotherms. Isotherms were selected that appear to be characteristic of peptide and protein samples in reversed-phase high-performance liquid chromatography. The dependence of bandwidth on sample size and gradient conditions was examined by Craig simulation and compared with experimental data for 13 different experimental systems involving four different proteins. There is a good correspondence between simulations and experimental data, and it seems possible to quantitatively predict bandwidth and resolution as a function of small-sample retention data, experimental conditions, and sample size. A systematic approach for designing the preparative or process-scale separation of protein mixtures by reversed-phase gradient elution is proposed.

INTRODUCTION

In the preceding paper[1] we presented a theoretical basis for interpreting mass-overloaded separations of proteins by means of gradient elution. However, this treatment has so far relied exclusively on small-molecule samples for illustration and verification[1,2]. The extension of our model to the case of higher-molecular-weight peptides and proteins is considered here. These compounds are obviously different in some respects, yet the fundamental principles of separation are expected to be the same as for corresponding small-molecule samples (see discussion of ref. 3). This has been demonstrated previously for the case of analytical-scale separations of peptides and proteins by various high-performance liquid chromatography (HPLC) methods: size-exclusion[4], ion-exchange[5], hydrophobic-interaction[6] and reversed-phase chromatography[3,6].

In view of the known properties of peptides and proteins, what special consequences might be expected for their preparative separation by HPLC? That is, how

0021-9673/88/$03.50

TABLE I

SOME POSSIBLE DIFFERENCES IN THE MASS-OVERLOAD SEPARATION OF SAMPLES THAT CONTAIN LARGE BIOMOLECULES

Sample property	*Separation consequences*
High molecular weight	Strong dependence of retention on the amount of organic modifier in the mobile phase Irreversible sorption at beginning of gradient, no separation during sample injection (see ref. 7) Non-Langmuir isotherms Unpredictable column capacity w_s Pore-size effects more important
Poly-electrolyte	Non-Langmuir isotherm Lower values of column capacity w_s More sensitive to silanol effects
Potentialy large differences in α	Unusually high column loads possible (with failure of two-term Langmuir approximation)
Highly viscous solutions	Additional band-broadening of sample, especially during injection
Three-dimensional molecular conformation	Problems and artifacts due to changes in conformation during separation

might the model we have developed for small-molecule chromatography fit these kinds of samples? Some possible differences of this type are summarized in Table I, with a further discussion provided in Results and discussion.

In principle the problems summarized in Table I can complicate the understanding and control of preparative HPLC separations of proteins. However, this need not be the case in every instance, as we will see here. Even where quantitative predictability is not possible, our basic model should prove useful as a qualitative guide. The following discussion is mainly concerned with reversed-phase HPLC, but the principles will be similar for other HPLC methods carried out in a mass-overload gradient elution mode.

THEORY

The complexity of some of the effects enumerated in Table I makes it difficult to properly interpret corresponding experimental observations. We have seen previously that the Craig model can provide reliable insights into different aspects of HPLC in a mass-overload mode[1,8,9]. Here, we will examine some of the effects of Table I by means of Craig simulations.

Non-Langmuir isotherms

We are unaware of any detailed studies of the sorption isoterm for proteins under reversed-phase HPLC conditions*. Strictly Langmuir-isotherm behavior seems unlikely for several reasons cited in Table I. First, reversed-phase HPLC normally

* Huang and Horváth[10] have reported data for di- and tri-peptides which can be fit (approximately) to a Langmuir isotherm over a limited range in sample size.

involves mobile phases with pH between 2 and 8. Under these conditions, a protein molecule acts as a poly-electrolyte, having numerous cationic sites which can selectively interact with the silanols present on silica-based column packings. We have previously observed for angiotensin II and the A-chain of insulin that these compounds appear to interact preferentially with accessible silanols[8], and this accounts for most of the retention of these compounds under typical reversed-phase conditions. The resulting isotherms are bimodal (non-Langmuir), with the result that the apparent column capacity w_s is quite low. For very large samples, the shape of the isotherm should be further complicated by ionic repulsion between adjacent charged molecules (see the discussion of ref. 8).

Second, the Langmuir isotherm assumes a displacement of adsorbed solvent molecules S by an adsorbing solute molecule X:

$$X_m + n\,S_a \rightleftharpoons X_a + n\,S_m \tag{1}$$

Here m and a refer to molecules in the mobile and stationary (absorbed) phase, respectively. In Langmuir sorption, there is unit stoichiometry, *i.e.*, n in eqn. 1 is equal to one. When n is larger than one, the isotherm again departs from Langmuir behavior[8]. Presumably a large molecule such as a protein displaces many individual solvent molecules, so that n is very large[11]. This would imply an isotherm that is quite different from the Langmuir isotherm (eqn. 1 with $n = 1$).

Finally, most column packings have a range in pore diameters such that some pores may have a size similar to that of the protein molecule. This can in turn give rise to sorption sites that are different from those present in wider pores; stronger sorption would normally be expected for molecules that are confined by a narrow pore. As in the case of silanol effects, non-Langmuir sorption should then result.

In the following theoretical treatment we will examine some special cases which may or may not be representative of actual protein HPLC systems. Our initial intention was to simulate the practical consequences of various changes in the sorption isotherm, and to use these as a basis for interpreting protein separations under mass-overload conditions. At the end of this paper we will compare a large body of experimental data (protein samples, mass-overloaded separations with reversed-phase gradient elution) with the immediately following theoretical database obtained from Craig simulations. Our use of various empirical relationships (*e.g.*, eqn. 4) to summarize the interim Craig-simulation data is of limited value *per se*; the value of this approach will only be apparent when comparisons with actual data for protein separations are presented at the end of this paper.

Two-site isotherms. The isotherm for this case can be derived as an extension of the Langmuir isotherm (eqn. 1 of ref. 1):

$$w_{xs} = [w_{s1}\,k_{01}\,C_x/(\psi_1 + k_{01}\,C_x)] + [w_{s2}\,k_{02}\,C_x/(\psi_2 + k_{02}\,C_x)] \tag{2}$$

Here, ws_x is the mass of solute X in the stationary phase, w_{s1} and w_{s2} are w_s values for sites 1 and 2 (*e.g.*, silanols and alkyl chains), k_{01} and k_{02} are the k_0 values of these sites, and ψ_1 and ψ_2 are the phase ratios of sites 1 and 2. The isotherms for angiotensin II and insulin A-chain[8] suggest values of w_s (and ψ) for silanols *vs.* alkyl chains of about 1:50, while the k_0 values are in the ratio of 3:1. Starting with these

representative values, we have derived a numerical solution to the isotherm of eqn. 2 for use in Craig simulations (mass-overloaded gradient elution separations). This algorithm is summarized in Appendix I. It was found accurate within about ±5% over a range of values pertinent to Craig simulations.

Craig simulations were carried out as a function of sample size, for the case of the above two-site isotherm. Fig. 1 summarizes one such series of runs (circles), for the case $n_c = 100$, $b = 0.50$ and $k_{0g} = 100$. These results can be compared with the similar plots of Fig. 7 of ref. 1 for a Langmuir isotherm. As in the latter case, values of corrected band-width W_{th} *vs.* sample size w_x/w_s give a straight-line plot when plotted on a log–log basis. However, the slope of this plot in Fig. 1 is equal to 0.35, whereas for Langmuir isotherm systems[1] the slope is 0.85 (dashed lines in Fig. 1).

Eqn. 2 predicts that the two-site case should approach the small k_0, large w_s case (site 2) for large samples, and the large k_0, small w_s case (site 1) for small samples. This is illustrated in Fig. 1, where the dashed plots are for the one-site isotherms corresponding to one site or the other. Deviations from a Langmuir isotherm that are due to large values of n in eqn. 1 were shown in ref. 8 to give rise to isotherms that are similar in shape to two-site isotherms. This suggests that linear plots as in Fig. 1 will be generally observed, but non-Langmuir systems should have slopes of log W_{th} *vs.* log (w_x/w_s) that are less than 0.85. The lower the slope, the greater is the deviation of the isotherm from Langmuir behavior.

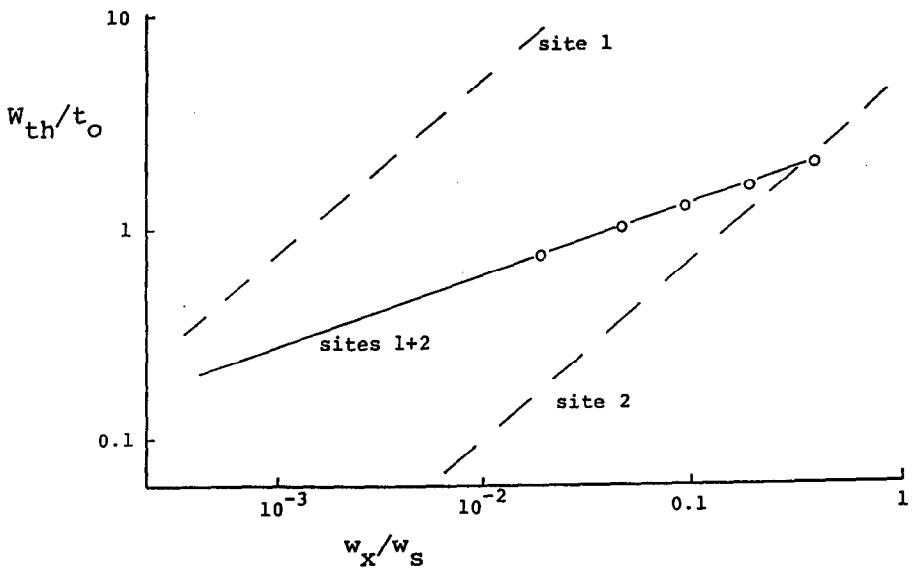

Fig. 1. Dependence of bandwidth W_{th} on sample size w_x/w_s for a two-site isotherm. Craig simulations with $n_c = 100$, $k_{0g} = 100$ and $b = 0.50$. Isotherm described in Appendix I. (O) Data from Craig simulations; (– – –) calculated for one-site case from eqn. 33 of ref. 1. Site 1 (*vs.* site 2) is characterized by a large value of k_0 and a small value of w_s (*e.g.*, silanols).

Effect of gradient steepness on bandwidth. In the previous paper[1] we showed (Craig simulations) that bandwidths for Langmuir-isotherm gradient elution can be represented by the empirical relationship

$$W_{th} = 2.3\ (t_0/b)\ (w_x/w_s)^{0.85} \tag{3}$$

That is, the mass-overload contribution to bandwidth W_{th} is inversely proportional to the gradient-steepness parameter b. This is,in fact, expected, as other (rigorous) relationships for bandwidth in gradient elution show that W_0 and W_{th} are proportional to t_0/b when k_{0g} is large, *e.g.*, eqns. 20 and 30 of ref. 1 and eqn. 43 of ref. 3. The same is true for isocratic elution, since W_{th} from ref. 1 (eqns. 5–7 and 11) can be expressed as

$$W_{th} = 2.4\ t_0\ k_0\ (w_x/ws)^{1/2} \quad \text{(isocratic)} \tag{7a}$$

which with $b = 1/1.15\ \bar{k} = 1/1.15\ k_0$ gives

$$W_{th} = 1.05\ (t_0/b)\ (w_x/w_s)^{1/2}. \tag{7b}$$

A reviewer has pointed out a physical interpretation of these relationships: when $W_{th} \gg W_0$ (either isocratic or gradient elution), the solute band occupies a fixed fraction of the column at the time of elution, regardless of gradient steepness or the value of k_0.

Extending this analogy for the case of non-Langmuir systems such as that of Fig. 1, we might expect that eqn. 3 can be rewritten for both Langmuir and non-Langmuir systems as

$$W_{th} = C\ (t_0/b)\ (w_x/w_s)^z \tag{4}$$

Here C is a constant, and z is the slope of plots as in Fig. 1 ($z = 0.35$ in Fig. 1). We can expect values of z to range from 0.85 downward, depending on the relative deviation of the system from Langmuir isotherm behavior. The example of Fig. 1 suggests that values of z as low as 0.35 are not improbable.

Eqn. 4 was tested by repeating the Craig simulations of Fig. 1 for other values of b. Since C can be calculated equal to 2.8 for the data of Fig. 1, and $z = 0.35$, values of W_{th} can be calculated from eqn. 4 for different values of b and compared with Craig simulations. These results are summarized in Table II; good agreement with eqn. 4 is observed. That is, eqn. 4 accurately predicts the effect of a change in gradient steepness (b) on reversed-phase separations under mass-overloaded conditions.

Band-shape for two-site isotherms. For larger values of w_x/w_s and/or N_0, it was shown in ref. 1 that the resulting bands in gradient elution will assume a "shark-fin" shape. For the case of two-site isotherm systems, greater tailing of the band is expected in both isocratic and gradient elution. For gradient elution this is expected to partly counteract the rounding of the bandtail due to gradient compression. This is illustrated in the representative example of Fig. 2a, for two-site Craig simulation with the same conditions as in Fig. 1 ($w_x/w_s = 0.20$). A corresponding band for a one-site (Langmuir) isotherm is shown for comparison in Fig. 2b (similar conditions).

TABLE II

VERIFICATION OF EQN. 4 FOR BANDWIDTH AS A FUNCTION OF GRADIENT-STEEPNESS PARAMETER b (NON-LANGMUIR SYSTEM)

Conditions as in Fig. 1 (Craig simulations).

b	w_x/w_s	W_{th}	
		Craig	*Eqn. 4*
0.2	0.05	2.7	2.5
	0.10	3.2	3.1
	0.20	3.9	4.0
	0.40	4.6	5.1
2.0	0.10	0.28	0.31
	0.20	0.38	0.40
	0.40	0.54	0.51

Resolution

The preceding relationships allow us to draw some useful conclusions. We have defined resolution R_s in mass-overloaded HPLC[12] as the difference in retention time for small samples of adjacent bands X and Y (t_{gx}, t_{gy}), divided by the width of the second band Y at overload (W_y):

$$R_s = (t_{gy} - t_{gx})/W_y \tag{5}$$

For large samples, as are usually encountered in preparative HPLC, we can approximate W_y by the value of W_{th} for Y. In the preceding paper (ref. 1, eqn. 18) we derived

$$t_{gy} - t_{gx} = (t_0/b)\log[(k_{0g})y/(k_{0g})_x] \tag{6}$$

which can be combined with eqn. 4 ($W_y \approx W_{th}$) to give

$$\begin{aligned} R_s &= \{\log[(k_{0g})_y/(k_{0g})_x]/C\}\ (W_x/w_s)^{-z} \\ &= (\text{constant})\ w_x^{-z} \end{aligned} \tag{7}$$

Eqn. 7 is interesting for two reasons. First we will see later that eqn. 7 provides an easy basis for estimating the optimum sample size in preparative separations of pro-

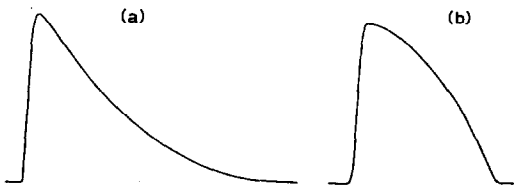

Fig. 2. Elution bands for one-site *vs.* two-site isotherms. Same conditions as in Fig. 1, except w_x/w_s = 0.20 for each case. (a) Two-site isotherm; (b) one-site isotherm.

teins by reversed-phase gradient elution. Second, it is seen that the gradient-steepness parameter b cancels in eqn. 7. This means that separations based on mass-overloaded gradient elution are independent of gradient time and steepness. This is quite unlike the case for analytical-scale (non-overloaded) separations, where resolution generally (but not always) improves as gradient time is increased.

Eqn. 7 also suggests that higher throughput in preparative HPLC will be favored by relatively short (steep) gradients —similar to the finding that preparative throughput is favored by small values of k_0 in isocratic separation[12]. However this conclusion is only valid when either the sample load and/or the column plate number N are large enough so that $w_{xn} \gg 1$ for solute Y (meaning that $W_{th} \gg W_0$).

EXPERIMENTAL

All equipment and procedures are similar to those described in ref. 8.

Instrumentation

The HPLC system comprised an LKB 2150 pump, an LKB 2152 gradient controller, and a Du Pont spectrophotometer detector (fitted with a 1-mm path-length preparative flow-cell).

Columns

All columns were prepared from stainless-steel tubing and were packed by a high-pressure slurry technique. Columns having dimensions of 15 × 0.46 cm I.D. were packed with experimental C_8 bonded-phase packings based upon Zorbax™ PSM 150 (8-μm particles, 15-nm pores), Zorbax PSM 300 (8-μm particles, 30 nm pores), Zorbax PSM 500 (7-μm particles, 50-nm pores), and Zorbax PSM 1000 (8-μm particles, 100-nm pores). Columns having dimensions of 25 × 0.46 cm I.D. were packed with Zorbax Pro-10 C_8 process-grade packing (10-μm particles, 13-nm pores), and columns having dimensions of 8 × 0.62 cm I.D. were packed with Du Pont BioSeries Poly-F (non-porous perfluorocarbon polymer, *ca.* 25-μm particles).

Protein chromatography

Proteins were chromatographed by reversed-phase gradient elution, using linear gradients from 5% (v/v) acetonitrile in water to 70% (v/v) acetonitrile in water; 0.1% (v/v) trifluoroacetic acid was added to each gradient solvent. Gradient times were either 20 or 35 min, and the flow-rate was 1 ml/min.

All protein standards were obtained from Sigma (St. Louis, MO, U.S.A.).

Calculations

Experimental values of bandwidth W were determined as in Fig. 5 of ref. 1. In some cases the actual procedure used was based on other measurements (*e.g.*, width at half-height), but in each case it was confirmed that the final values of W were equivalent to baseline bandwidth values. Values of W for the runs of Figs. 10 and 11 are averages for samples of pure proteins and protein mixtures. In principle, it is expected[8] that the bandwidths for lysozyme in mixtures with carbonic anhydrase would be smaller than in runs with pure lysozyme as sample. However, the actual reduction in W due to this effect was found to be small and within the experimental error of these measurements.

Values of b required in the use of various equations for bandwidth were determined from two experimental runs where only gradient time was varied (small samples). The resulting retention times t_g for the same compound in each run could then be used to determine b as described in eqn. 24 of ref. 3. Values of b for the runs of Figs. 8–11 (same conditions, except for pore size, sample and gradient time) were constant within the experimental error of these measurements ($b = 0.65 \pm 0.04$).

RESULTS AND DISCUSSION

One of the more relevant questions concerning protein separations by reversed-phase gradient elution is: how much sample can be injected for maximum yield of purified product (per unit time)? This is illustrated by the series of separations shown in Fig. 3 for a mixture of lysozyme and carbonic anhydrase [15 × 0.46 cm I.D. column, 15-nm pore size C_8 packing; gradient of 5 to 70% acetonitrile in water (0.1% TFA) in 20 or 35 min]. For the case of adjacent bands X and Y that just touch (see discussion of ref. 12), we can predict the maximum sample size from a knowledge of the width of the second band Y, since

$$W = t_{gy} - t_{gx} \quad \text{(for band Y)} \tag{8}$$

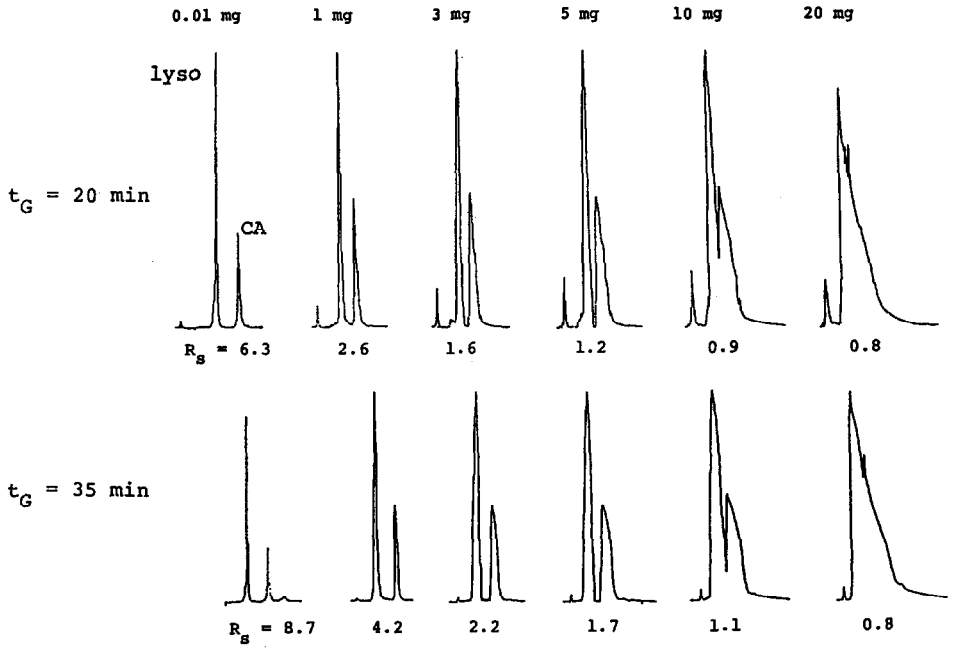

Fig. 3. Separation of lysozyme (lyso) and carbonic anhydrase (CA) as a function of sample size (mg each compound) and gradient time (t_G). Column, 15 × 0.46 cm I.D., 15-nm pores, C_8 (Du Pont); gradient, 5 to 70% acetonitrile in water (0.1% TFA added); flow-rate, 1 ml/min; room temperature.

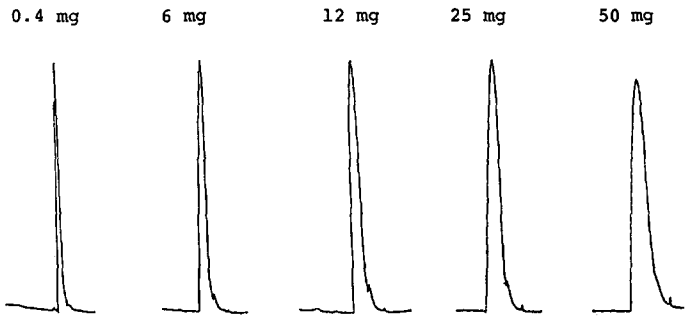

Fig. 4. Gradient runs for bovine serum albumin as a function of sample size. Column, 25 × 0.46 cm I.D. 10-μm particles, 30-nm pores, C_8 (Du Pont); gradient, 5 to 90% acetonitrile in water (0.085 *M* phosphoric acid plus morpholine); flow-rate, 1 ml/min; gradient time, 30 min.

Band width as a function of sample mass w_x

Several common proteins were run on a series of wide-pore or non-porous columns designed for protein separation. In most cases similar gradient conditions were used, resulting in b values of 0.3–0.7 ($\bar{k}$ = 1–3), *i.e.* reasonable experimental conditions. The shapes of the resulting elution bands generally resembled the "shark-fins" observed in Craig simulations of mass-overloaded gradient elution[1]. This is illustrated in Figs. 3 and 4 for some representative examples. In a few cases (particularly for very large samples), the elution bands had somewhat different shapes.

Experimental bandwidths were compared with eqn. 4, which was found to give a generally good description of our results and of additional data from the literature. This is illustrated in Fig. 5 for bovine serum albumin (BSA) at various sample sizes. According to eqn. 4, experimental plots of W_{th} *vs.* sample size w_x should be linear, with a slope that measures the relative deviation of the isotherm from Langmuir behavior. It is seen in Fig. 5 that this plot is linear within experimental error, over a 1000-fold variation in sample size.

Figs. 6–11 summarize additional (log–log) plots of W_{th} *vs.* w_x for a number of gradient elution protein separations. These data are mainly from our own laboratory; the data for Fig. 6 are taken from ref. 13. The plots in Figs. 5–11 are seen in every case to fit eqn. 4 rather well, *i.e.* linear plot within random scatter of the data*. These plots are from 13 different systems, involving four different proteins and a variety of column packings —including stationary phases of diffferent type, and pore diameters and surface areas that vary over an order of magnitude. Table III summarizes the plots of Figs. 5–11.

We do not know the value of w_s for each of the systems of Figs. 5–11, and therefore it is useful to restate eqn. 4 as

$$W_{th} = (C/w_s^z)\,(t_0/b)\,w_x^z$$
$$= C'\,(t_0/b)\,w_x^z \qquad (9)$$

* Such deviations from a best-fit straight line as are noted here (and in later plots of log W_{th} *vs.* log w_x) appear to reflect experimental variability rather than systematic deviations from eqn. 4. Values of W_{th} for small values of w_x are generally less accurate, because they are obtained by subtracting two large quantities (W and W_0, eqn. 11 of ref. 1).

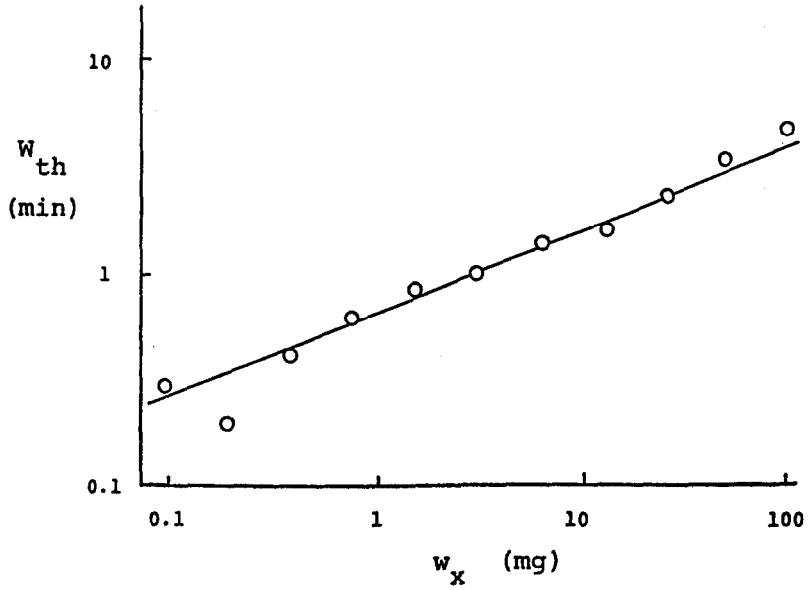

Fig. 5. Correlation of bandwidth W_{th} with sample size w_x. Sample bovine serum albumin; conditions as in Fig. 4.

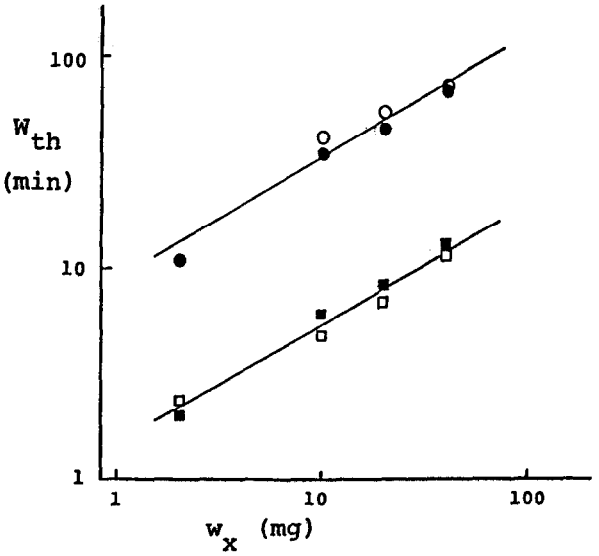

Fig. 6. Correlation of bandwidth W_{th} with sample size w_x (data from ref. 13); ribonuclease (○, ●; W_{th} values × 10) and bovine serum albumin (□, ■) samples. Columns, 16.2-nm pores (●, ■) and 21.6-nm pores (○, □); gradient, 5 to 71% acetonitrile in water (0.1% TFA added); gradient time, 30 min.

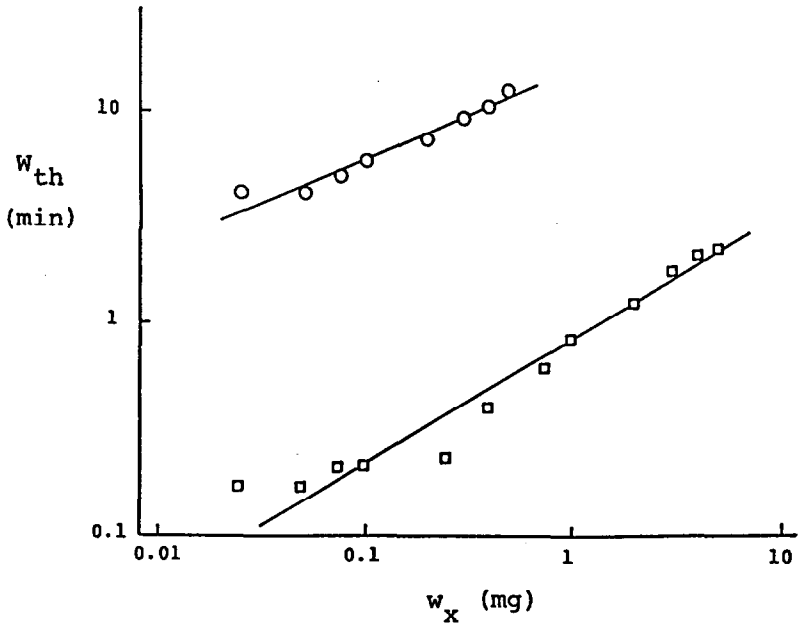

Fig. 7. Correlation of bandwidth W_{th} with sample size w_x; ribonuclease (○; × 10) and carbonic anhydrase (□) samples. Column, 8 × 0.62 cm I.D., non-porous Poly-F (Du Pont); gradient, 10 to 60% acetonitrile in water (0.1% TFA added); flow-rate, 2 ml/min; gradient time, 15 min.

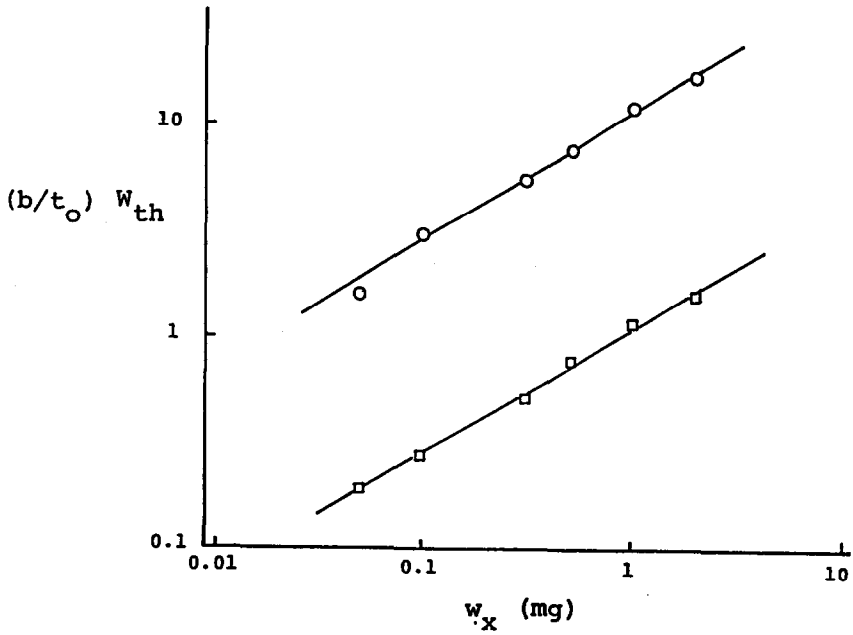

Fig. 8. Correlation of bandwidth W_{th} with sample size w_x; lysozyme (○, × 10) and carbonic anhydrase (□) samples. Column, 15 × 0.46 cm I.D., 100-nm pores, C_8 (Du Pont); gradient, 5 to 70% acetonitrile in water (0.1% TFA added); flow-rate, 1 ml/min; time, 20 min.

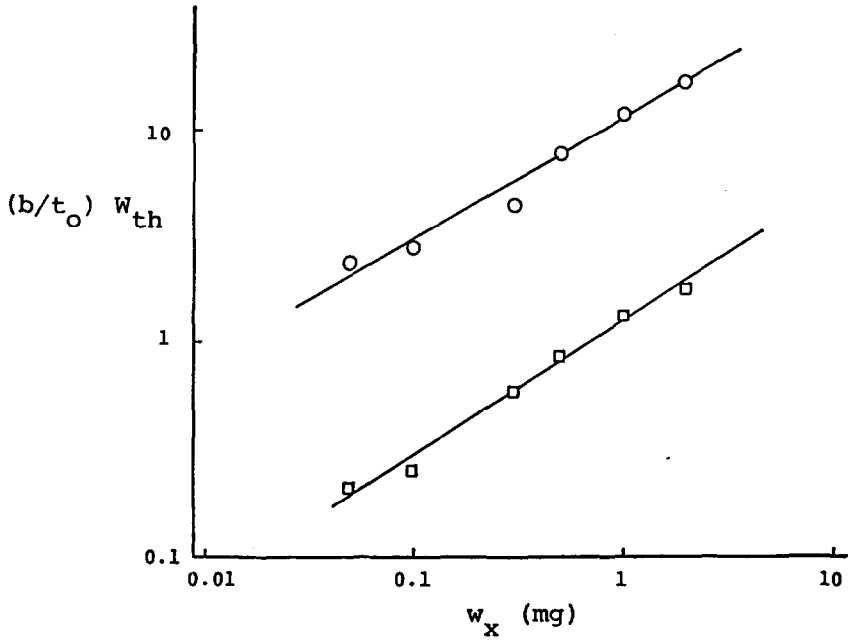

Fig. 9. Correlation of bandwidth W_{th} with sample size w_x; lysozyme (○, × 10) and carbonic anhydrase (□) samples. Same conditions as for Fig. 8, except column had 50-nm pores.

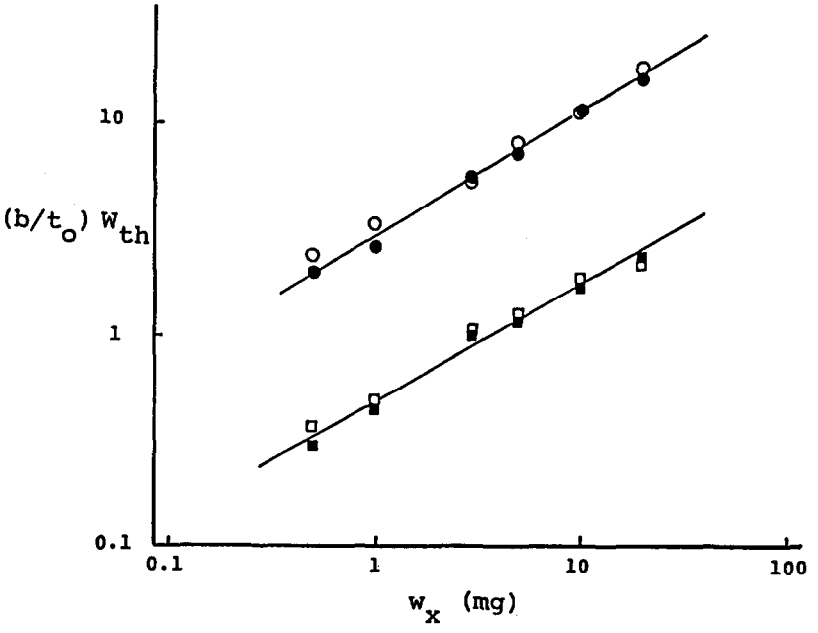

Fig. 10. Correlation of bandwidth W_{th} with sample size w_x; lysozyme (○, ●; × 10) and carbonic anhydrase (□, ■) samples. Same conditions as in Fig. 8, except 30-nm pores and gradient times of 20 min (○, □) or 35 min (●, ■).

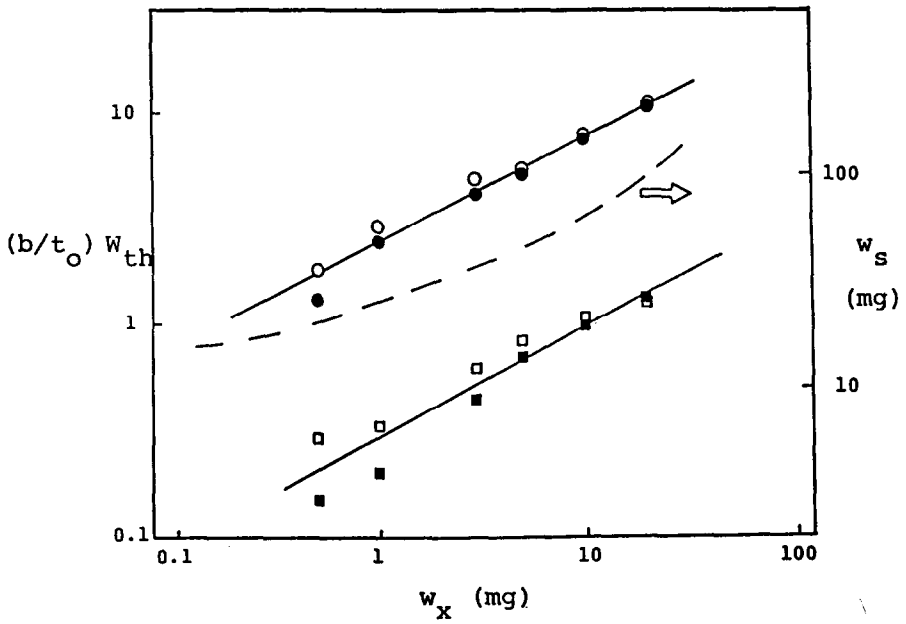

Fig. 11. Correlation of bandwidth W_{th} with sample size w_x; lysozyme (○, ●; × 10) and carbonic anhydrase (□, ■) samples Same conditions as in Fig. 8, except 15-nm pores and gradient times of 20 min (○, □) or 35 min (●, ■); w_s values (– – –) for lysozyme calculated from eqn. 24 of ref. 1.

Here, C' is a constant that should be smaller for larger column capacities (values of w_s). When values of b are not known for a particular system, eqn. 9 can be simplified further:

$$W_{th} = C'' w_x^z \tag{10}$$

Table III summarizes values of C', C'' and z for the mass-overload separations of Figs. 5–11.

Bandwidth as a function of sample and column. The parameter z of eqns. 4, 9 and 10 is seen in Table III to be relatively constant ($z = 0.54 \pm 0.08$) for these four proteins and seven different columns. The average value of z (equal 0.54) is also intermediate between values for a Langmuir isotherm ($z = 0.85$) and the markedly non-Langmuir isotherm ($z = 0.35$) upon which the correlation of Fig. 1 is based. If similar values of z are found for other proteins and reversed-phase systems, it suggests a similar retention mode and isotherms of similar shape. We do not believe, however, that the two-site isotherm discussed earlier necessarily governs protein retention in reversed-phase HPLC. The reason is that the tailing bands (Fig. 2a) predicted for this type of isotherm are not observed experimentally (see Figs. 3 and 4).

Constant values of z simplify the prediction of separation as a function of sample and separation conditions. Thus, for this case only two experimental runs are required in order to define bandwidth W as a function of sample size and gradient

TABLE III

SUMMARY OF CORRELATIONS OF BANDWIDTH WITH SAMPLE SIZE

Data of Figs. 5–11, fit to eqn. 9 or 10.

Figure	*Column characteristics**	*Conditions***	*Sample****	*z*	*C'*	*C''*
5	25/0.46/13/10/C_8	5–90/1/30§	BSA	0.39		0.64
6	?/?/16.2/?/C_{18}§§	5–71/?/30	RNase	0.58		0.9
			BSA	0.55		1.4
7	8/0.62/0/20§§§	10–60/2/15	RNase	0.41		0.6
			CA	0.59		2.2
8	15/0.46/100/10/C_8	5–70/1/20†	Lyso	0.60	0.29	
			CA	0.58	0.29	
9	15/0.46/50/10/C_8	5–70/1/20†	Lyso	0.56	0.32	
			CA	0.64	0.30	
10	15/0.46/30/10/C_8	5–70/1/20†	Lyso	0.59	0.30	
			CA	0.55	(0.50)	
11	15/0.46/15/10/C_8	5–70/1/20†	Lyso	0.50	0.25	
			CA	0.52	0.30	
			Mean ± S.D.	0.54 ± 0.08	0.29 ± 0.02	

* Length (cm) / I.D. (cm) / pore diameter (nm) / particle diameter (μm) / bonded phase.

** Gradient, % acetonitrile / flow-rate (ml/min) / gradient time (min); usually 2 < pH < 3 by adding TFA except Fig. 5.

*** BSA = bovine serum albumin; RNase = ribonuclease; CA = carbonic anhydrase; Lyso = lysozyme.

§ Morpholine and 0.085 *M* phosphoric acid added to the mobile phase.

§§ Column dimensions (and some other conditions) not given[11].

§§§ Non-porous fluorocarbon particles (Poly-F, Du Pont).

† Gradient times of 20 and 35 min; average data, see Figs. 10 and 11.

conditions (value of *b*): one run with a small sample, and a second run with significant overloading (preferably so that $W \gg W_0$). In principle, additional runs allow a value of *z* to be determined for each system, but the data of Figs. 5–11 suggest that this approach must be used with caution because of the considerable experimental scatter observed in these plots*.

Values of C' (eqn. 9) are directly related to column capacity w_s; the larger C' is, the smaller will be w_s. For the four columns of similar type and varying pore diameter (Figs. 8–11), values of C' for both lysozyme and carbonic anhydrase (CA) are remarkably constant: $C' = 0.29 \pm 0.02$ (excepting one deviant data point, for CA and the 30-nm pore column). This suggests that values of w_s are relatively independent of both pore diameter and protein molecular size (lysozyme, mol. wt. 14000; carbonic anhydrase, mol. wt. 29300).

Other studies (*eg.*, Fig. 2 of ref. 13) show the apparent column capacity for

* This uncertainty in the values of individual data points creates even larger uncertainty in the slope *z* derived from these data. We therefore recommend the assumption $z = 0.54$ until further data are accumulated.

the separation of ribonuclease and bovine serum albumin increasing as the pore diameter of the column packing increases from 8.5 to 16.2 to 21.6 nm. Related studies of column capacity as a function of protein size and pore diameter of the packing have also been reported for ion-exchange HPLC[14]. There, column capacity was found to be largest for pores of intermediate size; smaller proteins gave the largest values of w_s for smaller pores, and larger proteins exhibited the largest values of w_s for large pores. These latter effects can be rationalized in terms of two effects: increasing access to pores of larger diameter, and decreasing surface areas for larger-pore packings. The constant C' values for the plots of Figs. 8–11 are therefore somewhat puzzling; further work aimed at clarifying this picture is underway in our laboratory.

Effect of gradient conditions on bandwidth and resolution. Eqn. 9 predicts that log–log plots of $W_{th}(b/t_0)$ *vs.* w_s should give a single curve for all values of b, *i.e.* corresponding to varying gradient time or other gradient conditions. This is observed in Figs. 10 and 11 for gradient times of 20 (open circles and squares, respectively) and 35 min (closed circles and squares, respectively), other conditions equal. Eqn. 7 also predicts that mass-overloaded separation should be independent of gradient steepness; this seemed to us a surprising conclusion. The separations in Fig. 3 test this prediction for two series of separations (varying sample size) where gradient time has been varied by a factor of 1.75.

The separations of Fig. 3 show a similar resolution of the two bands (lysozyme and carbonic anhydrase) as gradient time is varied, for sample sizes of 10 or 20 mg (each protein). However, resolution is greater for the 35-min runs when the sample size is smaller. This agrees with our preceding discussion, since (a) we have assumed that $W_{th} \gg W_0$ (*i.e.* sample size is large) and (b) resolution normally improves with decreasing gradient steepness (larger value of t_G) for a small sample. Since we will usually inject as large a sample as possible in preparative separations, this suggests that preparative separations by gradient elution should be run with relatively steep gradients (for maximum throughput). This is analogous to the observation in isocratic separations that smaller k_0 values ($k_0 \approx 1$) favor higher sample throughput[12].

Column saturation capacity, w_s

Eqn. 24 of the preceding paper[1] has been recommended for the determination of column capacity w_s. The use of this equation for the lysozyme data of Fig. 11 (dashed curve) shows that w_s appears to increase with sample size w_x. A similar increase in apparent values of w_s for larger samples has been seen previously for the isocratic separation of angiotensin-II[15]; this is a general characteristic of systems exhibiting non-Langmuir retention. As a result it is not possible to determine the actual (maximum) value of w_s for systems such as those of Figs. 5–11 from eqn. 24 of ref. 1. Rather, it is necessary to determine the value of w_x at which column saturation and breakthrough of the elution band (at t_0) occurs (see the further discussion of Fig. 5b and c of ref. 1).

Apparent values of w_s (for a given value of w_x) from eqn. 24 of ref. 1 can be used for other purposes, as in the estimation of blockage effects (see following section).

Blockage effects

In ref. 1 we discussed the shift in retention of an early-eluting band X due to

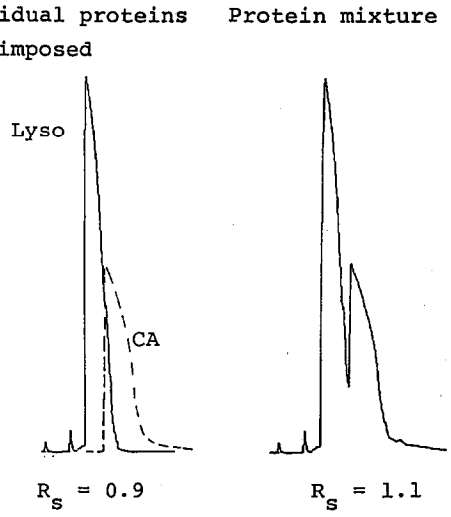

Fig. 12. Illustration of blockage effects for separation of mixtures of lysozyme (Lyso) and carbonic anhydrase (CA). Conditions as in Fig. 11. See text for details (20-min gradient, 10 mg of each protein injected).

the mass of a later band (or bands) Y. Thus the more strongly retained band Y appears to "block" part of the column, reducing the retention of band X. This effect was observed in the present studies, as illustrated in Fig. 12 for elution of 10-mg samples of lysozyme and carbonic anhydrase, either in admixture or as individual proteins. For the separation of the individual solutes (first chromatogram in Fig. 12). the superimposed bands suggest that resolution will be marginal: R_s = 0.9. The actual mixture, however, shows a significantly better resolution (R_s = 1.1) as a result of blockage of X by Y. That is, the lysozyme band is shifted to shorter retention

TABLE IV

COMPARISON OF EXPERIMENTAL AND PREDICTED BLOCKAGE EFFECTS IN THE SEP ARATION OF LYSOZYME–CARBONIC ANHYDRASE MIXTURES

Runs of Figs. 10 and 11, see details summarized in Table VI.

Pore diameter (nm)	*Gradient time (min)*	w_x *(mg)*	*Shift in lysozyme* t_g *(*Δt*, min)*	
			Experimental	*Calculated**
15	20	5	0.4	0.2
		10	0.3	0.3
		20	0.4	0.4
	35	5	0.4	0.3
		10	0.5	0.5
		20	0.6	0.8
30	20	5	0.2	0.2
		10	0.3	0.3
		20	0.3	0.4

* See Appendix II.

times due to the presence of a large sample (10 mg) of the carbonic anhydrase.

We can predict the shift in retention of band X due to the presence of band Y as discussed in ref. 1. Details of this calculation for lysozyme in the runs of Figs. 10 and 11 are summarized in Appendix II. The results are shown in Table IV; good agreement (±0.1 min, 1 S.D.) is found between experimental and calculated shifts (Δt) in the retention of the lysozyme band.

Prediction of resolution

A few simple measurements allow us to estimate sample resolution as a function of sample size and gradient conditions. We will illustrate this for the examples of Fig. 3 (lysozyme and carbonic anhydrase mixtures). We first run two standard gradients having t_G values differing by a factor of *ca.* 2. The conditions chosen for this sample are given in Fig. 11. These two runs allow us to optimize gradient conditions, for optimum band-spacing, maximum resolution and minimum runtime[16,17]. Or, we can improve our initial resolution by changing other separation conditions. Assuming that our two initial runs are in fact optimal, we can obtain values of t_g for each compound (lysozyme and carbonic anhydrase) in each run (20- and 35-min gradient in the present example), as well as values of W_0 for carbonic anhydrase (the second band): t_g = 13.69 and 20.68 min for lysozyme (20- and 35-min gradients); t_g = 15.88 and 24.29 for carbonic anhydrase. Experimental values of W_0 for carbonic anhydrase are 0.40 min (20-min run) and 0.43 min (35-min run). These t_g values allow the calculation of b values for carbonic anhydrase in each run (3): b = 0.68 (20-min run) and 0.39 (35-min run).

We next repeat one of the analytical-scale runs (t_G = 20 min) with a larger sample: w_x = 5 mg. In this case, we have chosen the weight of carbonic anhydrase in the sample, but we could have used the total sample weight; the following calculations would proceed in the same way, giving the same final predictions of resolution. The bandwidth W for the carbonic anhydrase band in the latter run is found equal to 2.58 min. By means of eqn. 11 of ref. 1 we calculate W_{th} equal to 2.54 min. Given that t_0 = 1.6 min, we then obtain C' = 0.26 from eqn. 9.

Values of W_{th} for carbonic anhydrase as a function of gradient time (or b) and sample weight w_x can now be calculated from eqn. 9. Corresponding values of W are then obtainable from eqn. 11 of ref. 1. For example, for the case of the 35-min run and w_x = 10 mg, we find that W_{th} = 3.70 min and W = 3.73 min. Now, the quantity ($t_{gy} - t_{gx}$) is obtained and corrected for blockage effects (Δt) to give the retention time difference Δt_g for a given value of w_x: $\Delta t_g = [(t_{gy} - t_{gx}) + \Delta t]$. For our 10-mg sample, the value of Δt (Table IV) is calculated equal to 0.5 min, giving a final value of Δt_g = 4.1 min. The resolution is then calculated from eqn. 5 (corrected for blockage) as R_s = 1.1. This is seen to match the actual value in Fig. 3 (R_s = 1.1).

The similar prediction of resolution for the other cases of Fig. 3 shows good agreement (±0.2 units) with experimental values of R_s when $w_x > 5$ mg, *i.e.* for relatively large samples. Agreement was poorer for smaller samples, but this is of less practical interest because these runs are sub-optimal in terms of sample throughput.

A practical aproach to developing a preparative separation of a protein sample

The present study suggests a straightforward approach to method development for the reversed-phase gradient elution separation of proteins samples in a mass-

TABLE V

SYSTEMATIC APROACH TO DEVELOPING A PREPARATIVE HPLC SEPARATION OF PROTEIN SAMPLES, BASED ON REVERSED-PHASE GRADIENT ELUTION

Step	*Comment*
(1) Develop coresponding analytical-scale separation	Choose best conditions: maximum α values for bands of interest
(2) Repeat separation with different gradient time (small sample)	Determine value of b for bands of interest
(3) Repeat run from step a with a large sample (same small column)	Determine bandwidth W_{th} and calculate value of C' (assume $z = 0.54$, eqn. 9)
(4) Calculate maximum sample size w_x for touching bands; vary gradient b value	Determine optimum sample and gradient steepness for maximum throughput
(5) Calculate sample size for different column diameters; keep b constant	Scale-up separation for desired final throughput

overload (preparative) mode. Our recommendations are summarized in Table V; they parallel the procedure of the preceding section for predicting resolution as a function of sample size and gradient conditions.

Step 1. If the time and effort required for separating and/or purifying a given sample is to be minimized, it is important that sample resolution first be optimized for a small sample. See ref. 12 for general guidelines.

Step 2. Two small-sample runs in which only gradient time is varied are needed, if gradient steepness is to be optimized for maximum sample throughput. This will allow us to estimate values of W as a function of gradient steepness b. A small-diameter column should be used at this stage.

Step 3. A third run (same column) is required with a large sample (sufficient to make $W \gg W_0$), in order to determine the parameter C' or C'' which will be used to estimate bandwidth as a function of sample size and gradient steepness.

Step 4. Resolution is now calculated for various sample sizes and values of b. Choose a combination of w_x and b that provides $R_s = 1$ and minimum run time. If the first-eluting compound (*e.g.*, lysozyme in Fig. 3) is of interest, a larger sample can be injected for maximum recovery of purified product in each run[12]. The predicted separation should be confirmed experimentally with the initial small-diameter column used in steps 1–3.

Step 5. Increase column diameter and length as required for adequate sample throughput, keeping b constant and increasing sample size in proportion to total column volume (scale-up). Use the same column packing as in steps 2–4.

Further examples and a more detailed description of this approach will be provided in future publications from this laboratory.

CONCLUSIONS

Computer simulations based on the Craig model were carried out for the mass-overloaded gradient elution separation of samples that follow either Langmuir or non-Langmuir sorption isotherms. These simulations suggest a simple, empirical

relationship (eqn. 9) that describes bandwidth and sample resolution as a function of sample weight, gradient conditions and various (measurable) separation parameters.

Experimental data for a number of reversed-phase gradient elution protein separations have been compared with our initial Craig simulations. Bandwidths for these systems are in good agreement with eqn. 9. Log–log plots of corrected bandwidths W_{th} *vs.* sample size w_x are linear as predicted, and their slopes are proportional to t_0/b. Plots of $W_{th}(t_0/b)$ are approximately constant; this suggests sorption isotherms of similar type for these 19 systems involving four different proteins, seven different column types, and varying gradient conditions. We tentatively propose that other protein separations by reversed-phase HPLC will exhibit similar behavior.

If the presently found relationships prove to be general, it is possible to predict preparative HPLC separations of protein samples (by reversed-phase gradient elution) as a function of sample size and gradient conditions. Only two or three experimental runs are required for these predictions. This can considerably simplify the design of preparative and process separations for maximum throughput of purified material, as outlined here.

GLOSSARY OF SYMBOLS

For a glossary of symbols see ref. 1.

APPENDIX I

Algorithm for craig simulations with two-site isotherm

Eqn. 2 is assumed with the following values: $w_{s1} = 0.004$, $w_{s2} = 0.20$, $k_{01} = 3$, $k_{02} = 1$, $\psi_1 = 0.20$, $\psi_2 = 0.004$. The fraction R of solute in the mobile phase is desired as a function of the total weight of solute w_x in the mobile plus stationary phase of each Craig stage.

(1) define $m = 1.04 - 0.28 \log (k_0 + 1)$
(2) if $k_0 C_x < 0.001$ then $R = 1/(k_0+1)$
(3) if $k_0 C_x > 6$ then $R = (w_x - w_s)/w_x$
(4) if $0.001 < k_0 C_x < 6$ then $R = (0.833/m)\, w_x - [(1-m)/m]$.

APPENDIX II

Detailed calculation of blockage effects: example of Table V

Certain experimental information must be collected before the approach outlined in ref. 1 can be followed. This is illustrated for the blockage of lysozyme by carbonic anhydrase in the runs of Figs. 10 and 11. It is first necessary to obtain t_g values for the two compounds under consideration (lysozyme and carbonic anhydrase) with two different gradient times (20 and 35 min here, other conditions the same). These data allow us to determine a value of b for each run[3]; here we use the average values of b for all the runs of Figs. 8–11: $b = 0.65$ for $t_G = 20$ min, and b

TABLE VI

DETAILED CALCULATION OF BLOCKAGE EFFECTS FOR SEPARATION OF LYSOZYME–CARBONIC ANHYDRASE MIXTURES

Data of Figs. 10 and 11; see discussion of Appendix II.

Pore diameter (nm)	w_x *(mg)*	w_x *(mg)*	x
15	5	35	0.18
	10	52	0.24
	20	85	0.37
30	5	35	0.16
	10	52	0.22
	20	85	0.27

= 0.37 for t_G = 35 min. A value of k_y/k_x is needed next, for use in the blockage calculations. This is given by eqns. 19 and 20 of ref. 1 as

$$\log(k_y/k_x) = (b/t_0)\,(t_{gy} - t_{gx}) \qquad \text{(A1)}$$

The value of t_0 was 1.6 min in the present systems. Values of k_y/k_x from eqn. A1 are: 7.1 (15-nm pore column) and 9.4 (30-nm pore column). Eqn. 26 of ref. 1 next allows the calculation of blockage-factors f_{yx} for each column: 1.24 (15-nm pore column) and 1.13 (30-nm pore column). Apparent values of w_s must be determined next for each value of w_x, using eqn. 24 (Table VI). Now, the fractional blockage x (Table VI) and retention-time shift Δt_g can be determined from eqns. 25 and 28. Final values of Δt are given in Table IV.

REFERENCES

1 L. R. Snyder, G. B. Cox and P. E. Antle, *J. Chromatogr.*, 444 (1988) 303.
2 J. E. Eble, R. L. Grob, P. E. Antle and L. R. Snyder, *J. Chromatogr.*, 405 (1987) 51.
3 L R. Snyder and M. A. Stadalius, in Cs. Horváth (Editor), *High-Performance Liquid Chromatography—Advances and Perspectives*, Vol. 4, Academic Press, New York, 1986, p. 195.
4 B. F. D. Ghrist, M. A. Stadalius and L. R. Snyder, *J. Chromatogr.*, 387 (1987) 1.
5 R. W. Stout, S. I. Sivakoff, R. D. Ricker and L. R. Snyder, *J. Chromatogr.*, 353 (1986) 439.
6 M. A. Stadalius, B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 387 (1987) 21.
7 J. M. Di Bussolo and J. R. Gant, *J. Chromatogr.*, 327 (1985) 67.
8 J. E. Eble, R. L. Grob, P. E. Antle and L. R. Snyder, *J. Chromatogr.*, 384 (1987) 45.
9 J. E. Eble, R. L. Grob, P. E. Antle, G. B. Cox and L. R. Snyder, *J. Chromatogr.*, 405 (1987) 31.
10 J.-X. Huang and Cs. Horváth, *J. Chromatogr.*, 406 (1987) 275.
11 X. Geng and F. E. Regnier, *J. Chromatogr.*, 296 (1984) 15.
12 L. R. Snyder, G.B. Cox and P. E. Antle, *Chromatographia*, 24 (1987) 82.
13 Y. S. Kim, B. W. Sands and J. L. Bass, *J. Liq. Chromatogr.*, 10 (1987) 839.
14 M. A. Rounds, W. Kopaciewicz and F. E. Regnier, *J. Chromatogr.*, 362 (1986) 187.
15 J. E. Eble, R. L. Grob, P. E. Antle and L. R. Snyder, *J. Chromatogr.*, 384 (1987) 45.
16 J. W. Dolan and L. R. Snyder, *Chromatography*, 2 (1987) 49.
17 B. F. D. Ghrist, B. S. Cooperman and L. R. Snyder, in preparation.