Journal of Chromatography, 387 (1987) 1–19
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 19 054

PREDICTING BANDWIDTH IN THE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHIC SEPARATION OF LARGE BIOMOLECULES

I. SIZE-EXCLUSION STUDIES AND THE ROLE OF SOLUTE STOKES DIAMETER *VERSUS* PARTICLE PORE DIAMETER

B. F. D. GHRIST
Biomedical Products Dept., E. I. Du Pont de Nemours & Co., Concord Plaza, Wilmington, DE 19898 (U.S.A.)
M. A. STADALIUS
Agricultural Chemicals Dept., E. I. Du Pont de Nemours & Co., Wilmington, DE 19898 (U.S.A.)
and
L. R. SNYDER*
LC Resources, Inc., 26 Silverwood Court, Orinda, CA 94563 (U.S.A.)
(First received August 21st, 1986; revised manuscript received September 4th, 1986)

SUMMARY

Column plate numbers, N, were measured for 12 different proteins as a function of mobile phase flow-rate in two gel filtration systems (either denaturing or non-denaturing conditions). These data were used to extend a previous model that predicts bandwidths in reversed-phase and ion-exchange chromatography. Restriction of diffusion of large molecules within column packing pores is now defined more precisely, with a single relationship describing this effect for both reversed-phase and size-exclusion chromatography (SEC) (and presumably other high-performance liquid chromatography systems). Separations by gel filtration (SEC) are now included in our general model.

A total of 17 flow-rate studies were carried out, involving different proteins, columns and/or mobile phase conditions (denaturing or non-denaturing). Comparisons of plate numbers predicted by the model with experimental values were satisfactory in 15 out of 17 cases. The remaining two cases appear to represent "non-well-behaved" systems, where experimental bandwidths were higher than predicted values by more than 20%. Initial attempts at understanding the origin of these non-ideal effects are described.

INTRODUCTION

Previous papers from this laboratory[1–7] have described the development of a general model for predicting solute retention and bandwidth in the separation of peptides and proteins by reversed-phase or ion-exchange high-performance liquid

0021-9673/87/$03.50

chromatography (HPLC). Our aim has been to establish a fundamental understanding of these separations, so that they can be carried out more effectively. We also hope to use this model for the rational design of new HPLC columns for the analytical and preparative separation of large biomolecules.

The present paper extends this model to gel filtration [size-exclusion chromatography (SEC)], an important HPLC technique for protein separations (*e.g.* ref. 8). Because SEC is based on a simpler separation process than in other HPLC methods, it is possible to use data from SEC separations to derive certain fundamental chromatographic parameters more precisely and less ambiguously, for application both to SEC and to the other HPLC methods. These parameters are applied in the following paper[9,10] to develop a general model that applies to all four of the HPLC methods for protein separation: reversed-phase, chromatography ion-exchange chromatography (IEC), hydrophobic-interaction chromatography (HIC) and SEC in either an isocratic or gradient elution mode.

THEORY

The theory of separations by SEC has been treated by several workers[8,11–13]. Here we will be concerned only with bandwidth relationships. Knox and McLennan[11] have shown for the SEC separation of polystyrenes that bandwidths are predictable from the theory developed for small molecules (so-called Knox equation):

$$h = Av^{1/3} + B/v + Cv \tag{1}$$

Here h is the reduced plate height, v is the reduced mobile-phase velocity, and A, B and C are constants that can be calculated for a given solute and set of conditions. The plate number N for the SEC separation of a given solute is given by the following standard relationships:

$$N = L/H \tag{2}$$

$$h = H/d_p \tag{3}$$

$$v = ud_p/D_m \tag{4}$$

Here, L is column length, H is plate height, d_p is column packing particle diameter and D_m is the solute diffusion coefficient.

Values of A are independent of separation conditions and vary only with how well the column has been packed. Typically, $0.5 \leqslant A \leqslant 1.0$ for well packed columns.

In SEC the parameter B is given[14] as

$$B = 2\gamma \tag{5}$$

where γ is an obstructive factor. Knox and Scott[15] give an expression for C from Giddings[16]

$$C = [k''/(1+k'')^2]/30\ (D_p/D_m) \tag{6}$$

where D_p is the solute diffusion coefficient in the stagnant mobile phase within the

particles of packing, and k'' is the mass of solute inside the particle divided by that outside the particle (similar, but not identical, to the definition of capacity factor k'). This expression for C assumes that v refers to the superficial velocity v':

$$v' = ud_p/D_m x \tag{6a}$$

where x is the fraction of mobile phase that is inside the column and outside the particle pores. The following discussion is based on replacing values of v in Eqn. 1 by values of v'.

Guiochon and Martin[13] define C differently,

$$C = R(1-R)/30\gamma \tag{7}$$

although their derivation also comes from Giddings[16]. Here R is the fraction of the solute outside the particle pores. Eqn. 7 can be rearranged to give

$$C = [k''/(1+k'')^2]/30\gamma \tag{8}$$

from which it is seen that Knox ignores the factor γ (actually, he assumes $\gamma \approx 1$), and Guiochon ignores the factor (D_p/D_m). Guiochon and Martin[13] used their equation for predicting the performance of SEC columns with small particles and protein solutes, but did not compare their equation with experimental values of N.

Walters[12] uses an empirical expression for C

$$C = 0.05/\gamma(D_p/D_m) \tag{9}$$

He showed that protein bandwidths (values of N) in SEC agree (approximately) with experimental values when (D_p/D_m) is calculated by

$$(D_p/D_m) = 1 - 2.10r_{sp} + 2.09r_{sp}^3 - 0.95r_{sp}^5 \tag{10}$$

Here r_{sp} is the ratio of the solute Stokes diameter to the pore diameter of the column packing. Eqn. 10 is derivable for the diffusion of spherical solutes through cylindrical pores[17]. Knox and McLennan[11] also found a decrease in (D_p/D_m) for larger polystyrenes separated by SEC, as a result of the hindered diffusion of large molecules in small pores. Eqn. 10 has been used (*e.g.* ref. 18) to estimate the restricted diffusion of proteins in other HPLC methods.

We can relate C to the SEC distribution constant K, defined as

$$K_D = (V_R - V_0)/V_p \tag{11}$$

Here V_R is the retention volume of the solute in an SEC system, V_0 is the volume of mobile phase in the column that is outside of the pores, and V_p is the mobile phase volume inside the pores. The quantity k'' is given as[11]

$$k'' = K_D (V_p/V_0) \tag{12}$$

and (V_p/V_0) is given in terms of x as

$$x = V_0/(V_p + V_0)$$

or

$$(V_p/V_0) = (1 - x)/x \tag{13}$$

Combining eqns. 6, 8, 12, 13 we have finally

$$C = K_D \{1/[K_D + x/(1-x)]\}^2 \, x/30 \, \gamma \, (1-x) \, (D_p/D_m)$$

$$\cong K_D \, x/19.2 \, \{K_D + [x/(1-x)]\} \, (1-x) \, (D_p/D_m) \tag{14}$$

We will use eqn. 14 with the Knox equation (eqn. 1, with v' replacing v) to interpret the experimental data reported in the present study. Specifically we will be interested in the variation of (D_p/D_m) with r_{sp} (*cf.* eqn. 10). Eqns. 1, 5 and 14 allow values of (D_p/D_m) to be determined for different SEC samples (having different r_{sp} values), as in the examples of ref. 11 for polystyrene solutes.

EXPERIMENTAL

Equipment and data handling

All chromatographic runs were carried out using an 8800 Series liquid chromatographic system (DuPont, Wilmington, DE, U.S.A.), with detection at 280 nm. Analog data were digitized and archived with a 760 Series interface (Nelson Analytical, Cupertino, CA, U.S.A.). Data processing was done either manually or with a modified version of the Nelson analytical chromatography software plus Series 220 microcomputer (Hewlett Packard, Fort Collins, CO, U.S.A.). In either case, plate numbers N were calculated from bandwidth measurements at half-height ($W_{\frac{1}{2}}$): $N = 5.54 \, (t_R/W_{\frac{1}{2}})^2$, where t_R is retention time.

Materials and procedure

All separations were carried out at ambient temperature (*ca.* 23°C), with 50-μl injections of protein solutions that contained 1 μg/ml (unless noted otherwise). Two different column-types were used, each having dimensions of 25 × 0.94 cm. The BioSeries GF-250 column (DuPont) has 4-μm particles and 15-nm pores; the BioSeries GF-450 column (DuPont) has 6-μm particles and 30-nm pores. The column packings for these two columns are made from a zirconia-stabilized silica particle that has a diol bonding.

Mobile phases. Two different mobile phases were used, referred to here as "denaturing" and "non-denaturing". The non-denaturing mobile phase was an aqueous solution containing 200 mM ammonium sulfate (Gold-label grade, Aldrich, Milwaukee, WI, U.S.A.), adjusted to pH 8.1 with orthophosphoric acid (Baker, Philipsburg, NJ, U.S.A.). The denaturing mobile phase contained 6 M guanidine hydrochloride (Sigma, St. Louis, MO, U.S.A.) plus 0.1 M sodium phosphate (Baker), adjusted to pH 6.8 with orthophosphoric acid. Each mobile phase was filtered before use through a 0.45-μm filter. All mobile phases were continuously sparged with helium during use. Water was purified using a Milli-Q water purification system (Millipore, Bedford, MA, U.S.A.).

Proteins. All proteins described here were obtained from Sigma and were at

least 90% pure. Sources of the individual proteins were as follows: cyctochrome *c*, equine heart muscle; ribonuclease A, bovine pancreas; lysozyme, chicken egg white; myoglobin, equine heart; carbonic anhydrase, bovine erythrocyte; pepsinogen, porcine stomach; ovalbumin, chicken egg white; phosphoglyceryl kinase, yeast; hemoglobin, bovine erythrocyte; albumin, bovine serum; malate dehydrogenase, porcine heart; alcohol dehydrogenase; equine liver; transferrin, bovine serum; hexokinase, yeast; lactate dehydrogenase, rabbit muscle; glyceraldehyde 3-phosphate dehydrogenase, rabbit muscle; alcohol dehydrogenase, yeast; aldolase, rabbit muscle.

RESULTS AND DISCUSSION

The 17 proteins listed in Table I were run initially, in order to map retention *vs.* solute molecular weight under both denaturing and non-denaturing (native) con-

TABLE I

RETENTION DATA FOR SEC SEPARATION OF NATIVE AND DENATURED PROTEINS ON GF-250 AND GF-450 COLUMNS

Flow-rate of 0.9 ml/min, other conditions as in Experimental section. Nat. = native; Denat. = denatured.

Protein	*M (· 10^{-3})*		*Retention time t_R (min)*			
	Nat.	*Denat.**	*GF-250*		*GF-450*	
			Nat.	*Denat.**	*Nat.*	*Denat.**
Cytochrome *c*	12.5	12.5	13.0	9.0	13.2	11.0
Ribonuclease	13.7	13.7	12.2	10.3	12.2	11.2
Lysozyme	14.0	14.0			13.0	11.5
Myoglobin	17.8	17.8	12.3	9.1	12.1	10.5
Carbonic anhydrase	29.3	29.3	12.0	8.3	11.9	10.0
Pepsinogen	39.4	39.4	10.9	8.2	11.3	9.8
Ovalbumin	45.0	45.0			11.3	9.5
Phosphoglycerate kinase	45.7	45.7	11.2	7.8		
Hemoglobin	65.0	16.1	10.9	9.7	11.7	10.5
Bovine serum albumin	66.0	66.0	10.4	7.7	10.9	9.3
Malate dehydrogenase	66.4	33.2	11.1	8.2	11.2	10.1
Liver alcohol dehydrogenase	80.0	40.0	11.0	8.2	11.2	9.5
Hexokinase	105.0	52.5	10.1	7.8	11.3	10.0
Lactate dehydrogenase	140.0	35.0	9.5	8.1	11.0	9.8
Glyceraldehyde 3-phosphate dehydrogenase	142.0	35.5	10.4	8.9	10.9	9.3
Yeast alcohol dehydrogenase	145.0	36.2	9.1	8.4	10.6	9.4
Aldolase	162	40.5				

* 6 *M* guanidine hydrochloride mobile phase; mol. wt. assumes dissociation of protein into subunits.

TABLE II

DEPENDENCE OF APPARENT MOLECULAR WEIGHT (M_d) OF DENATURED PROTEINS ON ACTUAL MOLECULAR WEIGHT (M)

M	M_d	M_d/M	
		Eqn. 15	*Kato et al.*[19,20]
500	810	1.6	
1000	2100	2.1	
3000	9900	3.3	3.2–4.3
10 000	54 000	5.4	4.4–5.2
30 000	250 000	8.3	5.5–8.3
100 000	1 300 000	13.5	
300 000	6 300 000	21	

ditions. Subsequently a smaller number of these proteins were run with change in flow-rate, in order to compare band widths with values predicted by our model, and in order to determine values of (D_p/D_m) for comparison with the solute Stokes diameter to pore diameter ratio, r_{sp}.

Protein SEC retention vs. molecular weight

Diffusion coefficients and Stokes diameters for denatured solutes. Our primary interest in the data of Table I was in order to estimate more accurately diffusion coefficients, D_m, and Stokes diameters, d_s, for denatured proteins, for use in our protein-HPLC model (see discussion of following paper, ref. 9). For native proteins, values of D_m can be estimated accurately from values of the protein molecular weight M (*e.g.*, review of ref. 13). Both the D_m value and the SEC distribution constant K_D (for a given protein and column) are determined by the Stokes diameter of the protein. That is, a larger protein molecule diffuses more slowly and has a smaller value of K_D. We should therefore be able to estimate changes in D_m and in Stokes diameter d_s upon denaturation from corresponding changes in K_D for a given protein.

Apparent molecular weight M for denatured proteins. We can express the change in K_D for the denatured protein in terms of an apparent increase in molecular weight: from a value M for the native protein to a value M_d for the denatured protein. We can also measure (M_d/M) from SEC retention data. Kato *et al.*[19,20] have reported SEC retention data for a number of proteins (Toya Soda columns), with the proteins run in both the native and denatured-reduced states*. Their data suggest the values of (M_d/M) shown in Table II. These data can be averaged and extrapolated to allow estimates of M_d for proteins separated under denaturing conditions (*e.g.*, reversed-phase HPLC). This in turn would allow estimates of D_m and the Stokes diameter d_s for any protein, using M_d in place of M for equations that predict D_m from values of M (*e.g.*, as in ref. 13).

* We initially were concerned that values of M_d might be affected by disulfide reduction, giving rise to misleading estimates of D_m and d_s for denatured (but unreduced) proteins. For this reason we carried out the additional experiments of Table I, using non-reduced proteins.

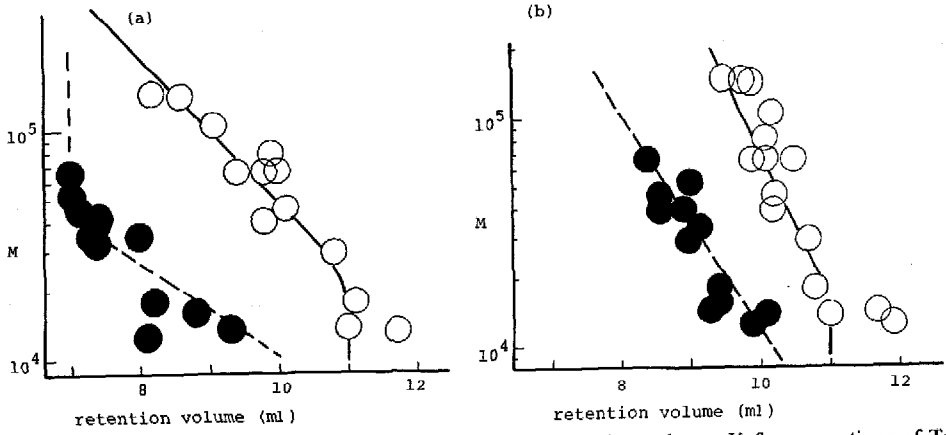

Fig. 1. Relationship of protein molecular weight, M, and retention volume, V, for separations of Table I. (a) Data for GF-250 column, (b) data for GF-450 column. ○, Run under non-denaturing conditions; ●, run under denaturing conditions. ———, Best fit to data; – – –, calculated from solid lines and eqn. 15. Conditions as given in Table I.

The data of Table I can be used to supplement the latter study[19,20]. We did not reduce the disulfide bonds of any of the proteins of Table I (as in refs. 19 and 20), because –S–S– bonds remain intact when a protein is denatured by reversed-phase (or any other denaturing) HPLC separation. The data of Table I are plotted as values of log M *vs.* V_R in Fig. 1a (15-nm pore column) and b (30-nm pore column). The data of Fig. 1 show significant scatter, reflecting the variable tertiary structure of individual proteins (not every native protein can be represented as a perfect sphere). For the denatured proteins, disulfide bridging (and possibly other intramolecular interactions) prevents a totally unconstrained random-coil structure. This means that values of K_D, D_m and d_s (for denatured proteins) are not determined by protein molecular weight alone. However the following treatment ignores these (somewhat minor) effects*.

The data of Fig. 1 suggest that the apparent molecular weight M_d of the denatured protein is related to molecular weight M as

$$M_d = 0.135\, M^{1.4} \tag{15}$$

This is seen in Fig. 1, where the dashed line for denatured proteins (solid circles) is calculated from eqn. 15 plus the empirical solid line through data for the native proteins (open circles)**. Eqn. 15 is physically reasonable in recognizing that denaturation will only affect the tertiary structure of large molecules (no "denaturation" of a single amino acid is possible). Comparison of values of M_d/M from eqn. 15 with values from Kato's data[19,20] is shown in Table II; close agreement exists between the

* We could use meassured K_D values from Fig. 1 to infer something of the detailed three-dimensional configuration of the protein, and to correct values of D_m and d_s estimated from protein molecular weight M alone (for use in our model). We have chosen not to do this.

** Note that a number of the proteins of Table I dissociate under denaturing conditions, with a change in the value of M for the denatured protein (Table I, column 2 *vs.* 3).

two studies, suggesting that reduction of a protein (as in refs. 19 and 20) does not much affect its Stokes diameter.

Predicting values of D_m. The diffusion coefficients D_m of a large number of native proteins have been empirically fitted to the expression (ref. 21, cited in ref. 13)

$$D_m = (8.34 \cdot 10^{-10})T/\eta M^{1/3} \tag{16}$$

where T is absolute temperature (K) and η is mobile phase viscosity in cPoise. Eqn. 16 with $M_d = M$ allows extension of this relationship for prediction of D_m values for denatured proteins as well. Previously[2] we have estimated the temperature dependence of η as

$$\eta = (298/T)^6 \eta_{25} \tag{16a}$$

which combines with eqn. 16 to give a more easily used relationship:

$$D_m = 9 \cdot 10^{-6}(T/298)^7[2.5M^{-1/3} + 62/M]/\eta_{25} \tag{16b}$$

Here η_{25} refers to the value of η at 25°C. The value of D_m in water at 25°C will also be required:

$$D_{mw} = 10^{-5}\,[2.5M^{-1/3} + 62/M] \tag{16c}$$

Eqn. 16b accurately describes the D_m values of native proteins (*i.e.*, is equivalent to eqn. 16); Eqn. 16b also closely approximates the Wilke–Chang equation for small molecules ($M < 1000$).

Stokes diameters d_s. The diffusion coefficient D_m is related to the Stokes diameter d_s of the solute molecule as[13]

$$D_m = kT/3\pi\eta d_s \tag{17}$$

Here k is the Boltzmann constant. Using available values of D_m (cm^2/s) and d_s (nm) for various native proteins[22], eqn. 17 becomes

$$d_s = 3.4 \cdot 10^{-6}/D_{mw} \tag{18}$$

where D_{mw} is the value of D_m in water at 25°C. Eqn. 18 is also expected to apply for denatured proteins.

Plate number measurements as a function of flow-rate

Tables III and IV summarize N values measured for several proteins as a function of flow-rate, under both denaturing and native conditions (15-nm pore and 30-nm pore columns, respectively).

Determining (D_p/D_m) vs. r_{sp}. For higher flow-rates, F, the C term of eqn. 1 becomes of primary importance, allowing better estimates of C (and D_p/D_m) from experimental bandwidth data. We have therefore used N values for the highest flow-rates in Tables III and IV (F = 4.5 ml/min) in order to estimate values of

TABLE III

EXPERIMENTAL *vs.* CALCULATED VALUES OF N (SEC separation)

Non-denaturing conditions, GF-250 column. Mobile phase is 0.2 *M* phosphate buffer.

Solute	*Flow-rate (ml/min)*	N ($\cdot 10^{-3}$)	
		Expt.	*Calc.**
Uracil	0.3	11.3	23.3
	0.9	21.6	32.8
	1.8	23.0	31.7
	4.5	21.4	25.5
Caffeine	0.3	13.4	27.7
	0.9	21.4	32.9
	1.8	20.3	29.6
	4.0	19.1	22.7
Uridine	0.3	16.5	29.2
	0.7	25.9	33.1
	0.9	27.0	32.6
	1.8	23.7	28.7
	3.6	21.0	23.3
	4.5	19.4	21.6
	5.4	18.2	20.2
Ribonuclease A	0.3	20.7	28.3
	0.7	19.9	21.4
	1.8	10.7	14.4
	4.5	6.5	9.2
Ribonuclease A (0.4 *M* buffer)	0.3	22.0	28.3
	0.7	19.4	21.4
	1.8	12.1	14.4
	4.5	9.2	9.2
Myoglobin	0.3	25.1	27.5
	4.5	7.0	8.5
Carbonic anhydrase	0.3	24.5	25.8
	0.7	19.6	18.7
	0.9	20.7	16.8
	1.8	11.6	12.2
	4.5	6.0	7.4
Pepsinogen	0.3	23.5	24.9
	4.5	6.4	6.9
Phosphoglycerate kinase	0.3	22.5	24.2
	0.7	18.3	17.3
	1.8	9.5	11.0
	4.5	5.7	6.5
Hemoglobin**	0.3	8.5	22.9
	4.5	2.8	5.7
Hemoglobin (0.4 *M* buffer)	0.3	9.3	22.9

(*Continued on p. 10*)

TABLE III (*continued*)

Solute	*Flow-rate (ml/min)*	*N* (· 10^{-3})	
		Expt.	*Calc.**
Bovine serum albumin	0.3	18.7	20.6
	0.7	13.8	16.2
	0.9	15.5	14.4
	1.4	8.7	11.6
	2.9	6.0	7.7
Liver alcohol dehydrogenase	0.3	20.0	21.9
	4.5	4.3	5.1
Transferrin	0.3	14.7	22.2
	4.5	3.6	5.3
Glyceraldehyde 3-phosphate dehydrogenase	0.3	14.7	19.2
	0.7	9.7	12.6
	1.8	7.2	7.1
	4.5	4.2	3.8
Yeast alcohol dehydrogenase	0.3	15.6	20.6
	0.7	11.4	14.0
	1.8	8.0	8.5
	4.5	4.4	4.7
Aldolase**	0.3	5.7	19.6
	4.5	3.3	4.2
Aldolase (0.4 *M* buffer)	0.3	7.0	

* Calculated as described in present paper; M and K_D values from data of Table I; 25 × 0.94 cm I.D. columns; d_p = 4 μm for GF-250 column, 6 μm for GF-450; 14-nm pore for GF-250, 29-nm pore for GF-450 (1 nm decrease due to diol phase); Knox A value = 1.0; x = 0.63; T = 25°C; viscosity of mobile phase = 0.95 cPoise for 0.2 M phosphate (non-denaturing) and 1.62 for 6 M guanidine (denaturing).

** Non-well-behaved protein.

(D_p/D_m) from eqns. 1, 2 and 14*. These values of (D_p/D_m) = ρ^* are plotted in Fig. 2 *vs.* values of r_{sp}. The ρ^* values of Fig. 2 fall reasonably close to the best (solid) curve through these various data, for both native (open circles) and denatured (solid circles) proteins. Since these derived values of ρ^* are based on estimates of D_m from eqn. 16b, and d_s from eqn. 18, the correlation of Fig. 2 also confirms these two equations.

The dashed curve of Fig. 2 is that predicted by eqn. 10, and it might be argued that the data of Fig. 2 fit either curve (dashed or solid) within experimental error. Our preference for the solid curve is governed by two considerations: (1) larger values of ρ^* (for the same value of r_{sp}) are more likely to correspond to "ideal" behavior (on which the present model is based); (2) the same dependence of ρ^* *vs.* r_{sp} should apply for reversed-phase HPLC, and in the following paper[9] we will show that re-

* The value of A for these two columns was assumed equal to 1.0, but derived values of C are not very sensitive to the value of A, because v' is large. A value of A = 1.0 was later confirmed by analyzing N value data at low flow-rates.

TABLE IV

EXPERIMENTAL *vs.* CALCULATED VALUES OF N (SEC separation)

Denaturing conditions, GF-450 column. Mobile phase is 6 M guanidine hydrochloride.

Solute	*Flow-rate (ml/min)*	N ($\cdot 10^{-3}$)	
		Expt.	*Calc.**
Uridine	0.3	18.0	21.9
	0.8	17.8	19.1
	2	14.9	14.3
	5	7.6	10.1
Ribonuclease A	0.3	11.4	10.9
	0.8	7.3	7.5
	2	4.5	4.5
	5	2.7	2.6
Carbonic anhydrase	0.3	9.6	9.2
	0.8	6.2	6.0
	1	5.5	5.3
	2	3.6	3.5
	5	1.9	1.9
Phosphoglycerate kinase	0.3	6.3	8.0
	0.8	4.2	5.1
	2	3.2	2.8
	5	1.9	1.4
Bovine serum albumin	0.3	7.6	7.4
	0.8	4.9	4.6
	2	2.9	2.5
	5	1.6	1.3
Glyceraldehyde 3-phosphate dehydrogenase	0.3	5.6	9.0
	0.8	3.8	5.9
	2	3.1	3.4
	5	1.4	1.9

* See conditions given in Table III (footnote).

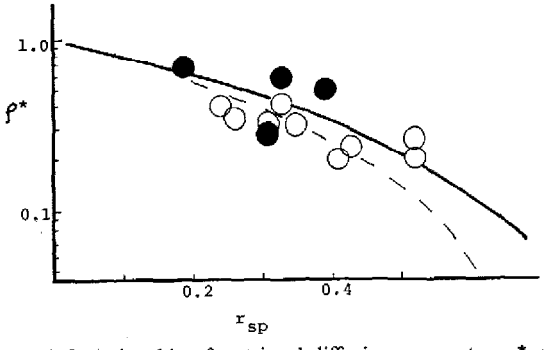

Fig. 2. Relationship of restricted-diffusion parameter, ρ^*, and ratio, r_{sp}, of solute Stokes diameter to pore diameter. ρ^* Values calculated from present model plus N values for 4.5 ml/min. ○, GF-250 data under non-denaturing conditions; ●, GF-450 data, under denaturing conditions. ———, Eqn. 15; – – –, eqn. 10.

TABLE V

SUMMARY OF EXPERIMENTAL *vs.* PREDICTED PLATE NUMBERS FOR SEC RUNS OF TABLES III AND IV

Compound	*M or M_d* (· 10^{-3})	K_D	*Conditions**	R_n**	*S.D.***
Uracil	0.11	1.5	GF-250	0.68	0.35
Caffeine	0.19	1.7	GF-250	0.67	0.35
Uridine	0.24	1.3	GF-250	0.84	0.21
Ribonuclease	13.7	0.98	GF-250	0.83	0.18
Myoglobin	17.8	1.00	GF-250	0.87	0.14
Carbonic anhydrase	29.3	0.93	GF-250	1.00	0.14
Pepsinogen	39.4	0.69	GF-250	0.94	0.07
Phosphoglycerate kinase	45.7	0.76	GF-250	0.93	0.10
Hemoglobin§	65	0.69	GF-250	0.44	0.56
Bovine serum albumin	66	0.58	GF-250	0.87	0.17
Liver alcohol dehydrogenase	80	0.71	GF-250	0.88	0.13
Transferrin	80	0.58	GF-250	0.67	0.33
Glycerylaldehyde 3-phosphate dehydrogenase	142	0.58	GF-250	0.92	0.17
Yeast alcohol dehydrogenase	145	0.29	GF-250	0.86	0.16
Aldolase§	162	0.33	GF-250	0.54	0.52
Uridine	0.24	1.0	GF-450	0.84	0.18
Ribonuclease A	83§§	0.67	GF-450	1.01	0.04
Carbonic anhydrase	242§§	0.50	GF-450	1.03	0.04
Phosphoglycerate kinase	450§§	0.46	GF-450	1.03	0.24
Bovine serum albumin	753§§	0.30	GF-450	1.12	0.14
Glycerylaldehyde 3-phosphate dehydrogenase	316§§	0.37	GF-450	0.73	0.30

* GF-250 refers to column, run under non-denaturing conditions; GF-450 is run under denaturing conditions; note that GPDH dissociates into four sub-units under denaturing conditions.

** Average value of ratio $R_n = N_{expt}/N_{calc}$.

*** S.D. = $(\Sigma R_n^2)/n$; standard deviation of values of R_n *vs.* $R_n = 1.00$.

§ Non-well-behaved separations.

§§ Values of M_d.

versed-phase data fit the solid curve of Fig. 2 more closely than the dashed curve. The solid curve of Fig. 2 is approximatable by the function

$$\rho^* = 1 - 1.83r_{sp} + 1.21r_{sp}^3 - 0.38r_{sp}^5 \qquad (19)$$

Fit of data of Tables III and IV to present model. Experimental N values in Tables III and IV are compared there with N values calculated from the present model (including eqn. 19; see ref. 9 for added details). We have summarized these

comparisons in Table V in terms of mean values $\bar{R}_n$ and standard deviations (S.D.); $R_n = N_{expt}/N_{calc}$ for each protein.

In most cases, reasonable agreement was found between these two sets of data (experimental *vs.* calculated). Thus for all but 2 of the 17 protein flow-rate studies summarized in Table V, it was found that $0.7 \leqslant R_n \leqslant 1.1$, with S.D. values $\leqslant 0.3$. The average error in predicted bandwidth values is $\pm 8\%$ (1 S.D.), which includes any uncertainty in measurement of the experimental values. This is somewhat better than the average precision we have so far found for the present model and the other three HPLC methods[9]. The data of Table V suggest that our model can predict values of N for SEC protein separations within about 15% (1 S.D.) for "well-behaved" systems.

Two protein data sets from Table III show poorer correlation: hemoglobin and aldolase, whose experimental N values are significantly lower than predicted by the model. We will discuss this further in a following section. We believe that wider bands or lower N values (experimental *vs.* model) are an indication of "non-well-behaved" systems; *i.e.*, involving other effects than are expected for "ideal" chromatography. These non-ideal effects can involve phenomena associated with the chemistry or biochemistry of the system, as discussed by others (*e.g.* ref. 23) and further reviewed below.

Secondary-retention effects. Proteins can be retained in SEC systems by processes other than exclusion (*cf.* discussion of Kato in ref. 24). These processes, which include ion-exchange and hydrogen-bonding with silanols on the silica surface of the packing, should be more important for smaller solute molecules, which presumably have easier access to the silica surface. Secondary-retention, when it occurs, often leads to excessive broadening of the solute band; *i.e.*, to values of n that are lower than predicted. Indications of such effects are found in Table III for two of the three small sample-molecules: uracil and caffeine, each of which show small $\bar{R}_n$ values (0.68 and 0.67) and large S.D. values (S.D. = 0.35)*. These two compounds also show K_D values greater than one (K_D = 1.5 and 1.7), which is further proof of secondary retention.

It has been found for the GF-250 column that increased buffer concentrations effectively minimize secondary retention for proteins; thus proteins show stable K_D values if the buffer concentration is kept $\geqslant 0.2\ M$ (used in runs of Table III). However increases in buffer concentration above 0.2 M phosphate normally have little further effect on retention. We nevertheless investigated the effect of increased buffer concentration on the N values for the non-well-behaved proteins of Table III. Experimental data are included in Table III for a mobile phase with 0.4 M buffer, *vs.* the 0.2 M buffer used initially. Examination of these 0.4 M runs shows no significant increase in N for these non-ideal compounds in Table IV. We therefore conclude that the present non-well-behaved systems are not caused by secondary retention.

Sample reaction during separation. Another possible cause of "non-well-behaved" behavior is the alteration of protein conformation during separation. That is, if denaturation occurs during separation, the effect is generally either the production of multiple peaks (*e.g.* ref. 25) or of a badly distorted and broadened peak (*e.g.*

* Values of the standard deviation S.D. of R_n values are relative to $R_n = 1$, not the mean value of R_n. The scatter of R_n values around the mean is usually considerably smaller than are values of S.D.

TABLE VI

NON-WELL-BEHAVED PROTEINS OF TABLE III; EFFECT OF INCREASING COLUMN LENGTH ON PLATE NUMBER, N

Conditions as in Table III.

Protein	*Flow-rate (ml/min)*	*Plate number N ($\cdot 10^{-3}$)**		
		Column 1	*Column 2*	*Columns 1 and 2*
Carbonic anhydrase**	0.9	14.8	13.6	24.5 (86%)***
Hemoglobin	0.3	7.3	8.3	7.3 (47%)
Transferrin	0.3	15.8	12.9	29.4 (102%)
Aldolase	0.3	5.9	6.4	7.1 (58%)

* Different columns used here *vs.* in Table III.
** Well-behaved protein (for reference).
*** N divided by N for columns 1 plus 2, expressed as %.

refs. 23 and 26). This effect can be distinguished from other cases in that a change in column length, L, will not produce the expected proportional increase in N when sample denaturation is occurring. Anomalous N *vs.* L comparisons due to sample reaction will be most obvious at lower flow-rates, where N values are normally greater, and where the extent of sample reaction is greatest.

The two "non-well-behaved" proteins of Table III were each run on two individual columns at a different flow-rate: 0.3 ml/min, then run on the two columns in series. The results of these experiments are summarized in Table VI. It is seen that the "well-behaved" proteins carbonic anhydrase and transferrin give plate numbers for the two columns in series that are close (86–102%) to the values predicted for normal chromatography (100%). However the two "non-well-behaved" proteins, hemoglobin and aldolase, show strong indications of sample denaturation or a similar effect. Their N values on the two columns in series are only 47–58% of the predicted values. We should also note that the hemoglobin peak tailed badly at lower flow-rates (asymmetry value A_s equal 4.3 at 0.3 ml/min, *vs.* average values of A_s of *ca.* 1.5 for the remaining proteins of Tables III and IV). This is further support for the reaction or denaturation of this compound during separation.

The latter effects for hemoglobin and aldolase can also be explained by impure samples, or "microheterogeneity" of the protein molecule. Preliminary experiments with hemoglobin did not support the latter possibility, however.

Knox *et al.*[9] have noted an effect similar to that of Table VI for hemoglobin and aldolase. Thus for the separation of a 200 000-Dalton polystyrene fraction, Knox found at low flow-rates that N did not increase in proportion to column length. Rather, plate numbers were roughly constant for 5, 10 and 25-cm columns of the same kind, and peak shape became poorer as column length increased. This was interpreted in terms of a second, slow retention process involving unraveling of the polymer molecule so as to allow penetration into pores where exclusion normally occurs. We do not believe this is the case in the examples of Table VI, because these molecules have access to the entire network of pores within the particle (Zorbax® packings have a quite narrow pore size distribution; see further comments in ref. 9).

Systematic development of an SEC protein separation

There is a substantial literature on this subject (*e.g.*, refs. 8 and 24), and not much qualitative advice can be added to what most practical workers already know. For increasing the resolution of a particular compound within a given sample, the general approach is:

(1) Select a column pore size that positions the band of interest near the middle of the fractionation range of the column ($K_D \approx 0.5$).

(2) Adjust the mobile phase composition[24] so as to eliminate secondary-retention or ion-exclusion effects (*e.g.*, use a 200 m*M* phosphate buffer with the present GF-250 and GF-450 columns).

(3) Decrease flow-rate for increased resolution.

(4) Increase column length for increased resolution.

With the present model available, we can make quantitative predictions about the relative value of each of these options for a given case. We will illustrate this with general results for hypothetical separations.

The calibration plot corresponding to the GF-250 data of Table I will be assumed. The quality of the separation can be assessed from the value of N for a protein of given molecular weight, since SEC retention does not change with separation conditions (for a given solute and column). Typical separation conditions were chosen for the following calculations (non-denaturing conditions). Fig. 3 summarizes these results, as plots of N *vs.* separation time for three different proteins: (a) molecular weight equal 30 000, (b) 100 000 and (c) 300 000. The resulting K_D values (0.89, 0.50 and 0.13, respectively) span most of the fractionation range. Fig. 3 shows the result of varying both flow-rate F and the number n of GF-250 columns connected in series (equivalent to varying column length). The various plots of N *vs.* separation time are similar for the three proteins, with maximum plate numbers developed at lower flow-rates (longer times) for larger proteins.

From these plots it appears that an increase in column length is more effective generally than is a decrease in flow-rate. In the case of columns with larger particles ($d_p = 4$ μm for the GF-250 column), flow-rate changes should prove relatively more effective than in Fig. 3. The present model allows quantitative predictions of separation *vs.* experimental conditions for any column whose physical characteristics are known.

It is generally assumed that large proteins will be separated less effectively than small molecules, and that the plate numbers for typical large-molecule separations will be relatively small. This is certainly not the case for SEC separations, where plate numbers of 50 000–100 000 can be developed in relatively short times, even for large molecules (Fig. 3). The data of Table II confirm this possibility, where experimental plate numbers equal to 20 000 (single column) are seen for several of these proteins at low flow-rates.

CONCLUSIONS

A general model for predicting HPLC separations of protein samples has been extended to the specific case of gel filtration (SEC). Experimental plate numbers N were obtained for 12 proteins (molecular weights of 13–160 kDaltons) as a function of mobile phase flow-rate, for both denaturing and non-denaturing elution condi-

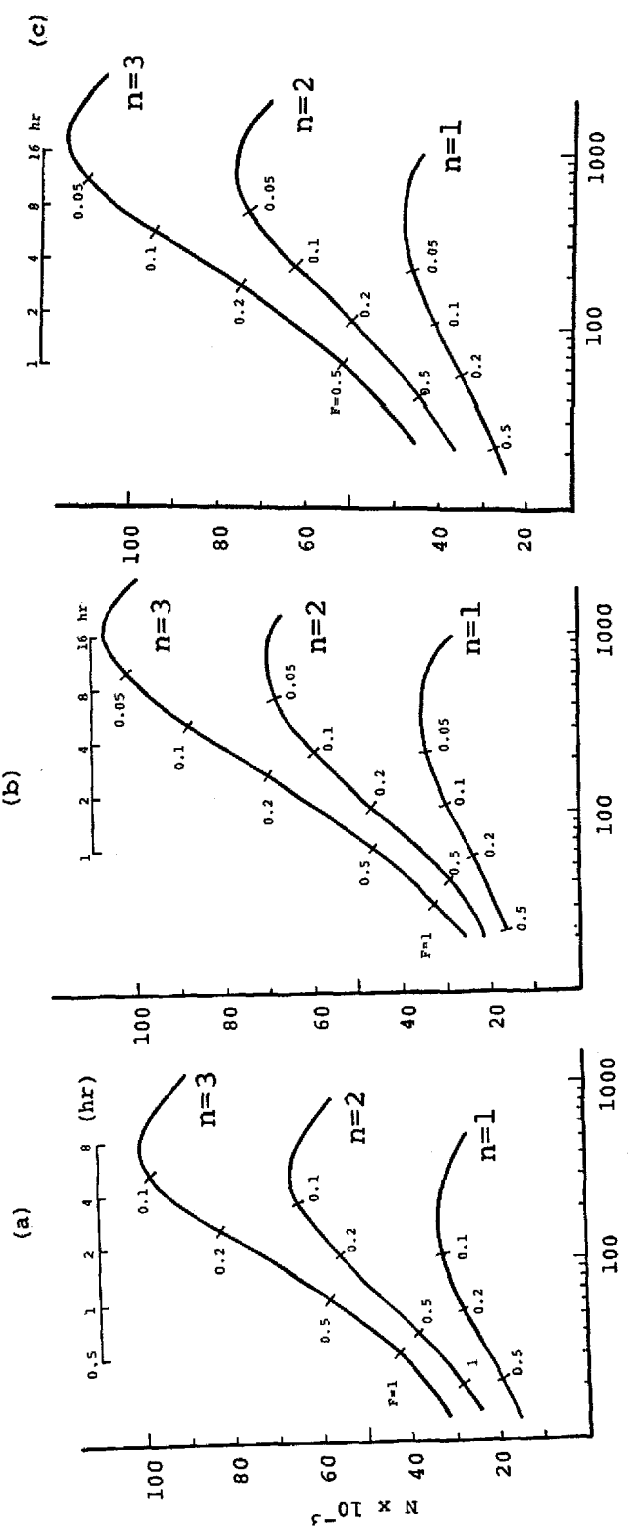

Fig. 3. SEC Separation as a function of experimental conditions. Predictions of present model. Assumes GF-250 column and conditions of Table III unless noted otherwise.

tions. In 15 out of 17 cases, good agreement was found between experimental N values and values predicted by the model: ±8% (1 S.D.) overall in predicted values of band width. The remaining 2 "non-well-behaved" systems, which gave lower experimental plate numbers than predicted, were further investigated. Secondary-retention effects appeared not to be responsible for this excess band-broadening.

The present experimental data base was also used to determine differences in protein diffusion coefficients inside and outside of the pores (D_p/D_m), as a function of molecular size. These results could be generalized, permitting definition of the relationship between (D_p/D_m) and the ratio r_{sp} = (solute Stokes diameter)/(column packing pore diameter). This relationship is shown in the following paper[9] to apply also to the reversed-phase separation of various peptides and proteins on 15- and 30-nm pore packings. Equations for the diffusion coefficient D_m and Stokes diameter d_s as a function of protein molecular weight (for either denatured or non-denatured molecules) were also derived and confirmed experimentally.

LIST OF SYMBOLS

Present and following paper, noted as I and II, respectively; e.g., eqn. II-5 refers to eqn. 5 from ref. 9

b	gradient-steepness parameter (eqn. II-16)
a', b'	constants in eqn. II-7
A, B, C	parameters in Knox equation (eqn. I-1; also eqns. I-5 and 14)
c	concentration of salt in the mobile phase
$\bar{c}$	in ion-exchange gradient elution, value of c at column midpoint when solute band has reached this position (see ref. 16)
c_f	in ion-exchange gradient elution, final value of c in the gradient
c_0	in ion-exchange gradient elution, initial value of c in the gradient
d_c	column inside-diameter (cm)
d_p	column packing particle size (cm)
d_s	solute Stokes diameter (nm, eqn. I-18)
D_m	solute diffusion coefficient in bulk mobile phase (cm²/s, eqn. I-16b)
D_{mw}	value of D_m in water at 25°C (eqn. I-16c)
D_p	solute diffusion coefficient in mobile phase inside pore (cm²/s, eqn. II-8)
D_p	average diffusion coefficient (cm²/s) inside particle (eqns. II-13 and 14)
D_s	solute diffusion coefficient in stationary phase (cm²/s, eqn. II-8)
D_s/D_p	diffusion in stationary phase relative to diffusion in adjacent mobile phase (inside pore) (eqns. II-23 and 24)
F	mobile phase flow-rate (ml/min)
h	reduced plate height (Eqns. I-1 and 3)
H	plate height (cm, eqn. I-2)
J	peak compression factor (see ref. 3)
$\bar{k}$	value of k' in gradient elution when band has reached column midpoint
k'	solute capacity factor
k''	mass of sample inside particle divided by mass outside particle (eqn. I-12)
k_0	value of k' at the beginning of gradient elution
k_z	value of k' at the end of gradient elution

K	ion-exchange equilibrium distribution constant (eqn. II-17)
K_D	SEC distribution constant (eqn. I-11)
L	column length (cm)
m	in ion exchange, the effective charge on the solute molecule divided by the charge on the mobile phase counter-ion
M	solute molecular weight
M_d	apparent molecular weight of a denatured protein as measured by SEC (eqn. I-15)
N	plate number (eqn. I-2)
r	ion-exchange gradient parameter (see ref. 5)
r_{sp}	solute Stokes diameter divided by pore diameter of column packing
R	fraction of solute outside of pores
R_n	ratio of experimental value of N to value of N predicted by model
$\bar{R}_n$	average value of R_n for a given solute (separation conditions varied)
R_w	ratio of experimental value of bandwidth to value predicted by model
$\bar{R}_w$	average value of R_w for a given solute
S	solute parameter in reversed-phase gradient elution (see ref. 2); also similar parameter in HIC (eqn. II-18)
S.D.	standard deviation of values of R_n or R_w *vs.* 1.00; measure of accuracy of predicted values of N or band width
t_g	retention time in gradient elution (min)
t_G	gradient time (min)
u	mobile phase velocity (cm/s, eqn. II-3)
u'	superficial mobile phase velocity (cm/s, eqn. II-4)
V_D	"dwell" or "hold-up" volume for HPLC system; volume from mobile phase mixer to column inlet (ml)
V_m	column dead-volume (ml, eqn. II-6)
V_0	volume of mobile phase outside of pores (ml)
V_p	volume of mobile phase inside pores
V_R	solute retention volume (ml)
x	fraction of mobile phase outside of column-packing pores
γ	tortuosity factor, equal to 0.64
$\Delta\varphi$	change in φ during gradient elution
η	mobile phase viscosity (cPoise, eqn. I-16a)
η_{25}	value of η at 25°C
ν	reduced mobile phase velocity (eqn. I-4)
ν'	superficial reduced mobile phase velocity (eqn. II-5); used for all equations in present model (both SEC and other HPLC methods)
ρ	restricted-diffusion parameter for small molecules (eqn. II-13a)
ρ^*	restricted-diffusion parameter, equal to (D_p/D_m) for large molecules
φ	in reversed-phase HPLC, the volume fraction of organic solvent in the mobile phase; in HIC, see eqn. II-18
λ	HIC mobile phase parameter (eqn. II-19)

REFERENCES

1 L. R. Snyder, M. A. Stadalius and M. A. Quarry, *Anal. Chem.*, 55 (1983) 1412A.
2 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 296 (1984) 31.
3 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 327 (1985) 27.
4 M. A. Stadalius, M. A. Quarry and L. R. Snyder, *J. Chromatogr.*, 327 (1985) 93.
5 R. W. Stout, S. I. Sivakoff, R. D. Ricker and L. R. Snyder, *J. Chromatogr.*, 353 (1986) 439.
6 J. L. Glajch, M. A. Quarry, J. F. Vasta and L. R. Snyder, *Anal. Chem.*, 58 (1986) 280.
7 M. G. Kunitani, D. J. Johnson and L. R. Snyder, *J. Chromatogr.*, 371 (1986) 313.
8 W. W. Yau, J. J. Kirkland and D. D. Bly, *Modern Size-Exclusion Chromatography*, Wiley-Interscience, New York, 1979.
9 M. A. Stadalius, B. F. D. Ghrist and L. R. Snyder, *J. Chromatogr.*, 387 (1987) 21.
10 L. R. Snyder and M. A. Stadalius, in Cs. Horváth (Editor), *High-performance Liquid Chromatography. Advances and Perspectives*, Vol. 4, Academic Press, New York, 1987.
11 J. H. Knox and F. McLennan, *J. Chromatogr.*, 185 (1979) 289.
12 R. R. Walters, *J. Chromatogr.*, 249 (1982) 19.
13 G. Guiochon and M. Martin, *J. Chromatogr.*, 326 (1985) 3.
14 R. W. Stout, J. J. DeStefano and L. R. Snyder, *J. Chromatogr.*, 282 (1983) 263.
15 J. H. Knox and H. P. Scott, *J. Chromatogr.*, 282 (1983) 297.
16 J. C. Giddings, *Dynamics of Chromatography*, Marcel Dekker, New York, 1965.
17 C. N. Satterfield, C. K. Colton and W. H. Pitcher, *AIChE J.*, 19 (1973) 628.
18 M. T. W. Hearn and B. Grego, *J. Chromatogr.*, 296 (1984) 61.
19 Y. Kato, K. Komiya, H. Sasaki and Y. Hashimoto, *J. Chromatogr.*, 190 (1980) 297.
20 Y. Kato, K. Komiya, H. Sasaki and T. Hashimoto, *J. Chromatogr.*, 193 (1980) 458.
21 M. E. Young, P. A. Carroad and R. L. Bell, *Biotechnol. Bioeng.*, 22 (1980) 947.
22 J. M. Brewer, A. J. Pesce and R. B. Ashworth, *Experimental Techniques in Biochemistry*, Prentice Hall, New York, 1974.
23 M. T. W. Hearn, A. N. Hodder and M.-I. Aguilar, *J. Chromatogr.*, 327 (1985) 47.
24 Y. Kato, *LC, Liq. Chromatogr. HPLC Mag.*, 1 (1983) 540.
25 S. A. Cohen, K. P. Benedek, S. Diong, Y. Tapui and B. L. Karger, *Anal. Chem.*, 26 (1984) 217.
26 K. A. Cohen, K. Schnellenberger, K. Benedek, B. L. Karger, B. Grego and M. T. W. Hearn, *Anal. Biochem.*, 140 (1984) 223.