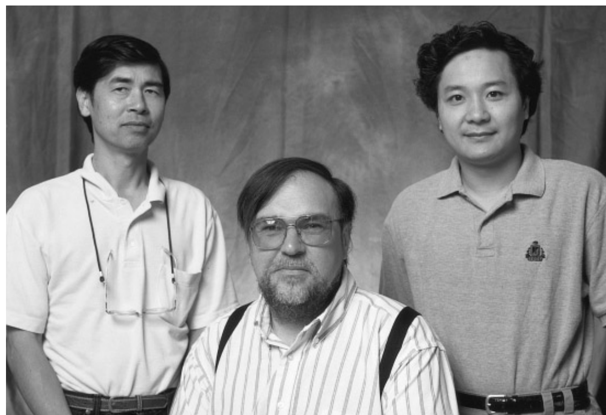
ANALYTICAL BIOCHEMISTRY 240, 155–166 (1996)
ARTICLE NO. 0345

REVIEW

Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions

Jie Wen, Tsutomu Arakawa, and John S. Philo
Protein Chemistry Department, Amgen Inc., Thousand Oaks, California 91320



Tsutomu Arakawa, John S. Philo, and Jie Wen

Techniques of using size-exclusion chromatography (SEC) with on-line light-scattering, uv absorbance, and refractive index detectors to characterize the polypeptide molecular weights of simple proteins or glycoproteins or to determine the stoichiometry of protein complexes are described. Two unique advantages of this approach over conventional SEC are that the molecular weight measurement is independent of elution position and can exclude the contributions from carbohydrates. When a protein or complex contains no carbohydrates, a two-detector method, i.e., light scattering combined with refractive index, can be used to calculate the molecular weight. When a protein contains carbohydrates, a three-detector method is used to calculate the molecular weight of polypeptide alone. Finally, a self-consistent three-detector method is used to determine the stoichiometry of a protein complex containing carbohydrates. Example applications for all these methodologies are described. © 1996 Academic Press, Inc.

I. INTRODUCTION

Proteins function through interactions with other small and large molecules in a highly specific manner and hence knowledge of how they interact with other molecules is fundamental for understanding their functions. In addition, protein molecules often self-associate to oligomers for specific purposes. For newly cloned proteins, especially those being found by large-scale sequencing projects, we often have no idea whether the protein exists in solution as a monomer, dimer, or other oligomer. Therefore, determining the molecular weights of proteins or their complexes is an important step in understanding proteins and their functions. Empirical techniques, such as size-exclusion chromatography (SEC)[1] and sodium dodecyl sulfate–polyacrylamide gel electrophoresis, and absolute techniques, such as mass spectroscopy, light scattering, and analytical ultracentrifugation, have been employed for protein molecular weight determination. Among these, SEC is a simple and fast method for estimating the molecular weight of a protein in its native form based on its elution position. However, there are several problems with this conventional SEC approach. One is that the elution position depends not only on the molecular weight of the protein, but also on its shape. Another problem is that the elution position will change if the protein has any tendency to interact with the column matrix. In addition, when a protein or a protein complex contains carbohydrates, the carbohydrates usually have a disproportionately large effect on its elution position, so SEC may not be able to determine its polypeptide molecular weight. We have seen several examples where recombinant proteins have been incorrectly identified as dimers when the molecular weights were estimated based on elution position alone. In contrast,

[1] Abbreviations used: SEC, size-exclusion chromatography; SEC-LS/UV/RI, size-exclusion chromatography with on-line light-scattering, uv absorbance, refractive index detectors; (LS), intensity of light-scattering signal; (UV), intensity of uv absorbance signal; (RI), intensity of refractive index signal; BSA, bovine serum albumin; RNase, ribonuclease; CHO, Chinese hamster ovary; SCF, stem cell factor; bFGF, basic fibroblast growth factor; TNF, tumor necrosis factor; sTNFR, soluble TNF type I receptor; BDNF, brain-derived neurotrophic factor; NT-3, neurotrophin-3; mAb, monoclonal antibody; HMWH, high-molecular-weight heparin; PBS, phosphate-buffered saline.

0003-2697/96 $18.00
Copyright © 1996 by Academic Press, Inc.
All rights of reproduction in any form reserved.

when SEC is used with on-line light-scattering detection, the molecular weight from this measurement is independent of the elution position, and for glycoproteins we can determine the molecular weight of only the polypeptide component if the extinction coefficient of the polypeptide alone is used in the analysis (the details will be discussed in the following sections). These characteristics make the combination of light scattering with SEC an easy, accurate, and reliable technique. This combination is also now much more readily available with the advent of suitable light-scattering detectors from several manufacturers[2] or through adapting commercial fluorescence detectors to measure 90° light scattering (1).

There are several excellent reviews on the topic of SEC with on-line light scattering (2–4), but none of these focused on studying protein interactions, an area which has been extensively developed during recent years. In this review, we will describe how to use size-exclusion chromatography with on-line light-scattering, uv absorbance, and refractive index detectors (SEC-LS/UV/RI) to determine (a) the molecular weight of simple proteins containing no carbohydrates, (b) the molecular weight of glycoproteins, and (c) most importantly the molecular weight and stoichiometry of protein–protein complexes or protein–carbohydrate complexes. We will illustrate all of these methodologies using example applications and list other related publications at the end of each application section.

II. A PROTEIN CONTAINING NO CARBOHYDRATES

A. Theoretical Background

A typical on-line SEC-LS/UV/RI system uses three detectors in series after an SEC column (2): a laser light-scattering detector, a uv absorbance detector, and a refractive index detector. A SEC instrument can be easily upgraded to a light-scattering/SEC system just by adding two detectors, a light-scattering detector and a refractive index detector.

The basic light-scattering equation is

$$\frac{K^*c}{R(\theta)} = \frac{1}{M}\left[1 + \frac{16\pi^2}{3\lambda^2}\langle r_g^2\rangle \sin^2(\theta/2) + \cdots\right] + 2A_2c,$$

where $R(\theta)$ is the excess intensity of light scattered at angle θ (i.e., the intensity due to the solute). K^* is an optical parameter equal to $[4\pi^2n^2(dn/dc)^2]/(\lambda_0^4N_A)$, c is the solute concentration in mg/ml, λ_0 is the wavelength of the light in vacuum, M is the weight-average molecular weight, $\langle r_g^2$ is the mean square radius of gyration, A_2 is the second virial coefficient, n is the refractive index, dn/dc is the refractive index increment of the solute, and N_A is Avogadro's number. At the low concentrations usually encountered during column chromatography, the virial coefficient term is negligible. In addition, the term $[16\pi^2\langle r_g^2\rangle\sin^2(\theta/2)]/(3\lambda^2)$ will be negligible if we measure the light scattering at small angles, i.e., low-angle laser light scattering (LALLS) (5–8). Furthermore, this term will generally be negligible at all angles for proteins or complexes with $\langle r_g^2\rangle^{1/2} < 15$ nm (1), which is true for a folded polypeptide with $M < 5 \times 10^7$ (this approximation makes the approaches described in this review valid for both single or multiangle on-line light-scattering instruments available commercially). Under these conditions, the above basic light-scattering equation can be simplified to $(K^*c)/R(\theta) = 1/M$. Substituting K^* with $[4\pi n^2(dn/dc)]^2/(\lambda_0^4N_A)$, the intensity of light-scattering signal (LS) is given by

$$(LS) = K_{LS}cM(dn/dc)^2, \quad [1]$$

where K_{LS} is an instrument calibration constant.

We may similarly express the refractive index signal, (RI), as

$$(RI) = K_{RI}c(dn/dc), \quad [2]$$

where K_{RI} is again an instrument calibration constant. For a protein or complex that contains no carbohydrate, the dn/dc is constant (~0.186 ml/g) and nearly independent of its amino acid composition. Hence, we can determine M from the ratio of the two detectors, (LS) and (RI).

$$M = K'(LS)/(RI), \quad [3]$$

where $K' = K_{RI}/[K_{LS}(dn/dc)]$. This is the so-called "two-detector method." It is the method most commonly used and is the method usually provided by the software from the instrument manufacturers, but it is only valid when dn/dc is known, which is generally not true for glycoproteins or their complexes.

To get the instrument calibration constant, K', we normally run ribonuclease (M = 13,690, Calbiochem), ovalbumin (M = 42,750, Sigma), and bovine serum albumin monomer (M = 66,270, Sigma) as calibration standards (some standards of higher molecular weight were used to check the linearity of this method in our early studies) and do a fit of (LS)/(RI) versus molecular weight to a line passing through the origin to determine K', as shown in Fig. 1. After this instrument calibration constant is obtained, the molecular weights of other unknown proteins can be calculated. It should be pointed out that calibration using protein molecular weight standards is not the only method, since K_{LS} can

[2] Precision Detectors, Inc. (Amherst, MA); Viscotek (Houston, TX); Wyatt Technology Corp. (Santa Barbara, CA).

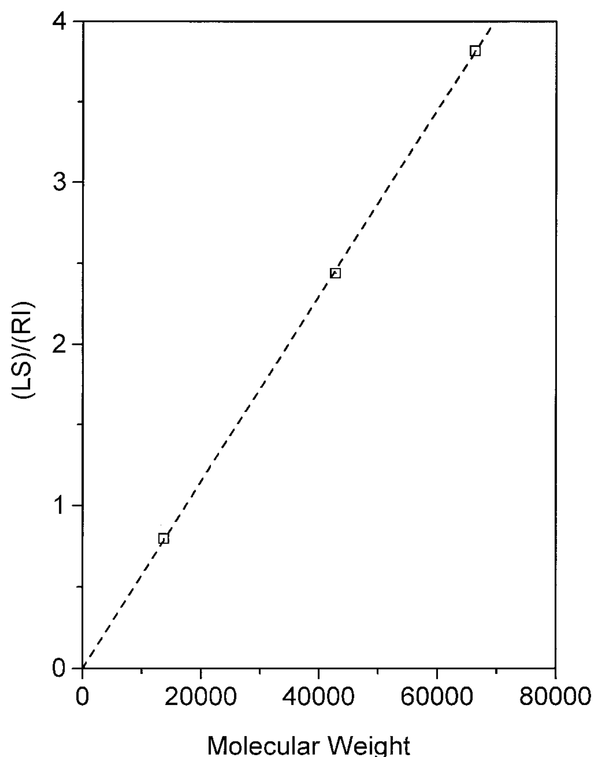


FIG. 1. A typical plot of (LS)/(RI) versus the molecular weights of protein standards (RNase, ovalbumin, and BSA). One hundred microliters of each protein standard was injected onto a SEC column separately to obtain more accurate data (slight overlap may occur for some SEC columns). Typical protein concentrations were 2.0, 1.5, and 1.5 mg/ml for RNase, ovalbumin, and BSA, respectively.

be absolutely calibrated using the intrinsic Rayleigh scattering of a pure solvent such as toluene. The details of the absolute calibration method can be found in Refs. 9 and 10. There are two major reasons why our laboratory uses protein standards to calibrate the instrument. The principal reason is that our RI detector calibration is not very stable with time, and the use of protein standards conveniently calibrates both the RI and LS detectors at once. A second reason is that absolute calibration requires switching between toluene and aqueous buffers, which is inconvenient and can cause precipitation of salts on detector windows. The protein standard calibration method has been used for many years in our laboratory and the data from this method are very reliable. While we generally run each standard separately, it is possible to mix and run standards together as long as they are well resolved by the chromatography.

To obtain the values of LS and RI signals, either peak height or peak area methods may be simply used when the sample is homogeneous. Each method has its own advantages. When the resolution between peaks is poor, we prefer the peak height method. Furthermore, when the sample is heterogeneous, one can calculate the molecular weight distribution, slice-by-slice, through the whole chromatogram and obtain number-average, weight-average, z-average molecular weight, and polydispersity (10–13). For all these calculation methods it is important first to adjust the chromatograms to account for the interdetector delay. The interdetector volume can be conveniently measured using any homogeneous protein giving a narrow peak.

We should also mention that although we have been referring to "the LS signal," for some detectors there may be signals available for scattering at multiple scattering angles. As discussed earlier, except for proteins of high molecular weight, there will not be any significant angular dependence of the scattering. Therefore, it is our general practice to use only the data for scattering at 90° (even though data are available at other angles[3]), both for simplicity and because the 90° data usually have the highest signal/noise ratio. However, for multiangle detectors it is certainly possible to use all the available data for the analysis, and these additional data may improve the overall accuracy of the molecular weight determination in some situations.

B. Applications

One example of a protein containing no carbohydrate is bovine serum albumin (BSA). Its chromatogram is shown in Fig. 2. Commercial BSA from Sigma is a mixture of monomers, dimers, and higher oligomers. The molecular weight of BSA dimer calculated from the two-detector light-scattering method is 132,000, which agrees well with twice the sequence molecular weight of 66,269. As seen in the chromatogram, a nice feature of light scattering is that peak 1 can be clearly distinguished as a dimer of peak 2 even without detailed calculations because the relative intensity of peak 1 to peak 2 signals in the light-scattering data is twice as high as in the RI chromatogram. The reason for this is that the intensity of light scattering is proportional to both concentration and molecular weight.

To illustrate one important advantage of SEC with light scattering over traditional SEC methods, both native and reduced, carboxymethylated ribonucleases (RNases) were subjected to SEC as shown in Fig. 3. Despite the fact that the elution positions for native and reduced RNases are very different, the molecular weights calculated from the two-detector method are the same for both, as expected, because their (LS)/(RI) ratios are the same. This can be clearly seen in Fig. 3C, in which the chromatograms are normalized to the same

[3] All data presented here were obtained with a Wyatt Technology miniDAWN detector, except for the RNase, SCF, and bFGF studies which used a Polymer Laboratories (Amherst, MA) PL-LALLS detector.

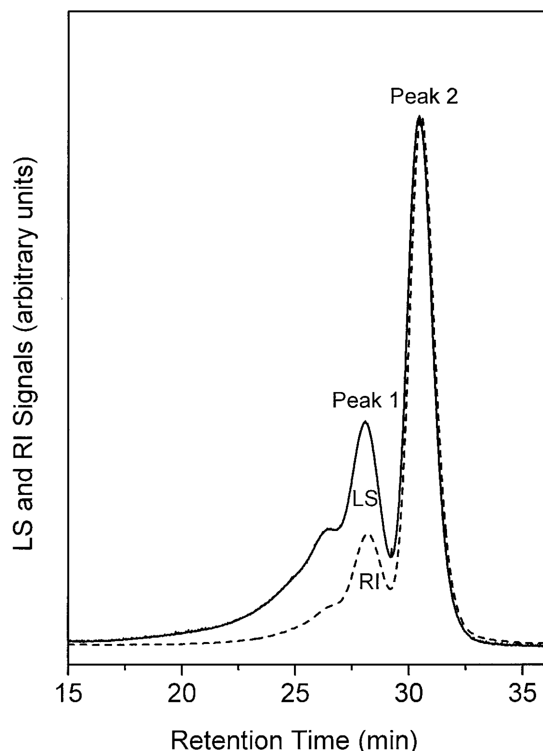


FIG. 2. Chromatogram of BSA that contains monomer, dimer, and other oligomers. One hundred microliters of 4 mg/ml BSA was injected onto a Superose 6 column (Pharmacia) with PBS as eluant at a 0.5 ml/min flow rate. The solid line is the LS signal and the dashed line is the RI signal.

scale. If the molecular weights were calculated based only on their elution positions, we would get a different result, i.e., 13,700 for the native RNase and 41,000 for the reduced RNase. The smaller elution volume of the reduced RNase is due to unfolding. Although many unfolded proteins have a tendency to aggregate, this result indicates that this is untrue for the reduced RNase. As shown by this example, SEC with light-scattering detection may be useful in some instances to provide information regarding the conformation of a protein if the molecular weights derived from light scattering and derived from the elution position are compared.

More applications using this two-detector method to study nonglycosylated proteins can be found in Refs. 14–19.

III. A PROTEIN CONTAINING CARBOHYDRATES

A. Theoretical Background

When a protein or a protein complex contains carbohydrates, its *dn*/*dc* is no longer known or constant because a carbohydrate usually has a different *dn*/*dc* than a polypeptide, and the carbohydrate content is normally unknown (and often varies significantly from molecule to molecule). In such cases, we cannot use the RI detector to determine the concentration, *c*, for Eq. [1]. Fortunately, we can instead use the signal from uv absorbance detector, (UV),

$$(UV) = K_{UV} c \epsilon \quad [4]$$

to determine *c*, where K_{UV} is an instrument calibration constant and ϵ is the extinction coefficient (the absorbance of 1 mg/ml of a glycoprotein or glycoprotein complex at a 1-cm pathlength). By combining Eqs. [1], [2], and [4], we derive

$$M = \frac{K_{RI}^2}{K_{LS} K_{UV}} \frac{(LS)(UV)}{\epsilon (RI)^2}, \quad [5]$$

where *M* and ϵ are the molecular weight and extinction coefficient of the entire protein including carbohydrate. This equation is the basis for the three-detector method. However, in most cases the ϵ is unknown, especially for a protein complex, and therefore Eq. [5] is not commonly used in light-scattering analysis. What may

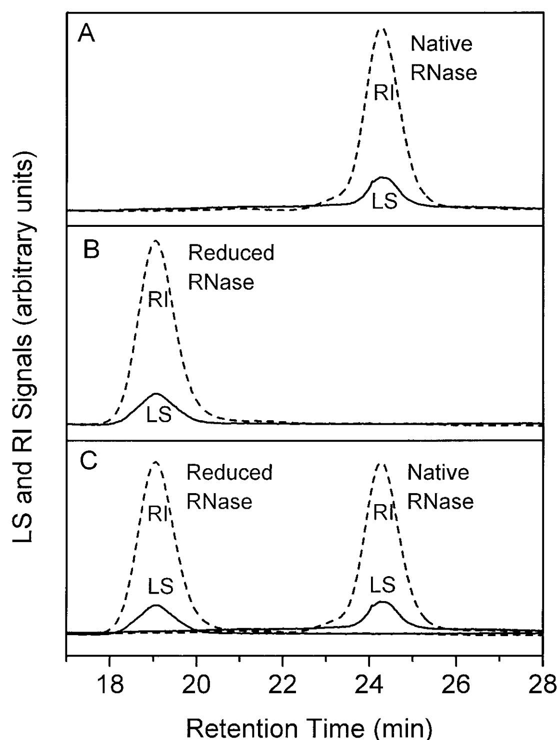


FIG. 3. Chromatograms of native and reduced, carboxymethylated RNase. One hundred microliters of each protein was injected onto a Superdex 75 column (Pharmacia) with PBS as eluant at a flow rate of 0.5 ml/min. (A, 1 mg/ml of native RNase; B, 1 mg/ml of reduced RNase; C, chromatograms in A and B are put into the same scale for comparing the (LS)/(RI) ratios of native and reduced RNase. The lines are the same as defined in the legend to Fig. 2.

be obtained relatively easily is the polypeptide extinction coefficient, ϵ_p, which can be either obtained from experimental data or estimated with reasonable accuracy from the amino acid composition (20). Some caution should always be used when applying calculated extinction coefficients, which may have some errors arising from conformational effects, cofactor binding, or molecular association. Recent data, however, suggest that the differences between folded and unfolded proteins are small and provide a method for calculating extinction coefficients of folded proteins (21). Further, we have seen no evidence of significant extinction coefficient changes caused by molecular associations. If we use ϵ_p and select a wavelength where the carbohydrate does not absorb, it is possible to algebraically eliminate all the contributions from the carbohydrates. To demonstrate this, we reexpress (LS), (RI), and (UV) signals based on the polypeptide concentration, c_p, and the mass and dn/dc of the polypeptide and carbohydrate components. From Eq. [1] we obtain

$$(LS) = K_{LS}c_p\left(\frac{M_p + M_c}{M_p}\right)(M_p + M_c) \times \left(\frac{M_p(dn/dc)_p + M_c(dn/dc)_c}{M_p + M_c}\right)^2 = K_{LS}c_p\,\frac{[M_p(dn/dc)_p + M_c(dn/dc)_c]^2}{M_p}, \quad [6]$$

where subscripts p and c stand for the polypeptide and carbohydrate components. Similarly, Eq. [2] gives

$$(RI) = K_{RI}c_p\left(\frac{M_p + M_c}{M_p}\right)\left(\frac{M_p(dn/dc)_p + M_c(dn/dc)_c}{M_p + M_c}\right) = \frac{K_{RI}c_p[M_p(dn/dc)_p + M_c(dn/dc)_c]}{M_p}. \quad [7]$$

From Eq. [4] we derive

$$(UV) = K_{UV}c_p\left(\frac{M_p + M_c}{M_p}\right)\left(\frac{\epsilon_p M_p + \epsilon_c M_c}{M_p + M_c}\right)$$

but since we have chosen a wavelength where $\epsilon_c = 0$, this simplifies the equation to

$$(UV) = K_{UV}c_p\epsilon_p. \quad [8]$$

Combining Eqs. [6] through [8] and solving for M_p gives

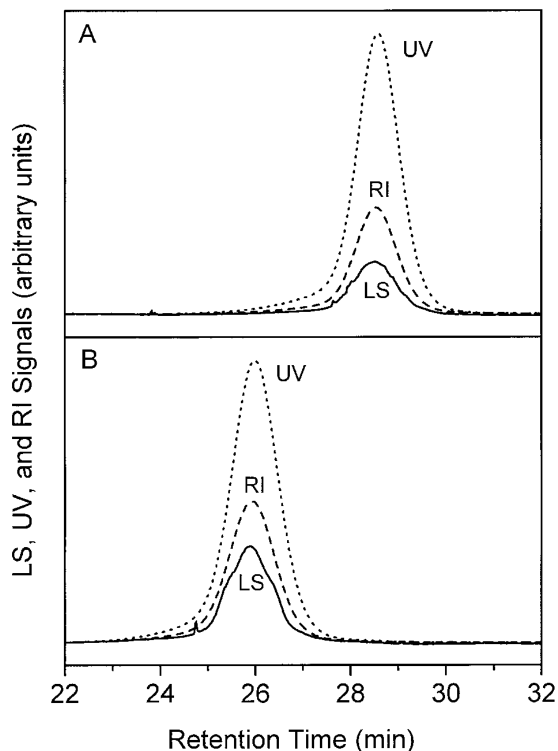


FIG. 4. Chromatograms of *E. coli* and CHO SCF. One hundred microliters of each protein was injected onto a Superdex 200 column (Pharmacia) with PBS as eluant at a flow rate of 0.5 ml/min. A, 3.8 mg/ml of *E. coli* SCF; B, 3.0 mg/ml of CHO SCF. The solid line is the LS signal, the dashed line is the RI signal, and the dotted line is the uv signal.

$$M_p = \frac{K_{RI}^2}{K_{LS}K_{UV}}\,\frac{(LS)(UV)}{\epsilon_p(RI)^2}. \quad [9]$$

As shown in Eq. [9], all contributions of the carbohydrate are canceled algebraically, and we can measure the polypeptide molecular weight directly as long as the polypeptide extinction coefficient is known. Equation [9] is the actual basis of the three-detector method used in our studies. The instrument calibration constant, $K_{RI}^2/(K_{LS}K_{UV})$, can again be conveniently obtained by running protein standards. The ϵ_p used in our analyses for RNase, ovalbumin, and BSA are 0.706, 0.735, and 0.670, respectively (5).

B. Applications

1. A single glycoprotein. Stem cell factor (SCF) is a dimeric glycoprotein that stimulates hematopoietic progenitor cells in bone marrow (22). Chinese hamster ovary (CHO) and *Escherichia coli*-derived recombinant human SCF were studied using SEC-LS/UV/RI and their chromatograms are shown in Figs. 4A and 4B, respectively. CHO SCF contains >30% carbohydrate

and thus the three-detector method was used. The ϵ_p value was calculated from the amino acid composition derived from the known amino acid sequence (but an experimental value is preferred if one is available). The polypeptide molecular weight from light scattering is 38,000 which agrees quite well with the 37,051 sequence molecular weight of CHO SCF dimer. Although the CHO SCF is glycosylated, the nonglycosylated *E. coli*-derived form is also biologically active. The molecular weight of this protein determined by the two-detector method is 36,000. Using ϵ_p and the three-detector method, we can also obtain the molecular weight of *E. coli* SCF as 38,000. Both values agree, within experimental error, with the sequence molecular weight of *E. coli* SCF dimer, 37,313. In contrast, the molecular weight from the conventional SEC method for *E. coli* SCF was reported as ~57,000 (23), a value which should, in principle, be the same as that from light scattering. In this case, the much higher molecular weight from the conventional SEC method indicates that *E. coli* SCF is highly asymmetric and therefore acts hydrodynamically "bigger" during SEC. Furthermore, for CHO SCF the total molecular weight estimated by the conventional SEC method is 113,000, whereas the true value is 53,100 from sedimentation equilibrium (23). This error by more than a factor of 2 for the conventional SEC method is a consequence of both the asymmetric shape of CHO SCF and the disproportionately large effect of carbohydrates on the hydrodynamic size. This example illustrates the dangers in using the elution position alone to estimate the molecular weight of a protein and how difficult it would be to correctly deduce the dimeric stoichiometry of SCF from its elution position on SEC.

More applications using this three-detector method to study glycoproteins can be found in Refs. 24 and 25.

2. Interactions of proteins with carbohydrate polymers. We can also apply this three-detector method to examine complexes formed by the interactions of proteins with carbohydrates. For example, a number of growth factors bind tightly to highly charged carbohydrates such as heparin or heparan sulfate (26–28), and these interactions are thought to modulate their stability, distribution, and biological activity *in vivo.* We have used the three-detector method to study binding of basic fibroblast growth factor (bFGF) to high-molecular-weight heparin (HMWH, $M_r \approx$16,000) (29). By using this three-detector method to measure the total polypeptide molecular weight of the complexes, we can determine how many bFGF molecules bind to one heparin molecule, i.e., the stoichiometry. Since heparin absorbs little light at 280 nm, the polypeptide molecular weight in the protein complex can be calculated by using the polypeptide extinction coefficient, ϵ_p (see Eq. [9]). The chromatograms of bFGF and HMWH mixtures are shown in Fig. 5 and the polypeptide molecular weights of each peak summarized in Table I. The results indicate that each peak contains an average of six to seven bFGFs. In addition to the molecular weight provided as above, another nice feature of SEC-LS/RI/UV is that all conventional SEC methods may be used at anytime because the data of UV and RI chromatograms are always included in the raw data file along with the LS data. In this particular example, we can use the UV chromatogram to determine the amount of bFGF remaining unbound to heparin and then plot the free bFGF versus ([HMWH]/[bFGF]) as shown in Fig. 6. The slope of the line implies an average of 7.5 ± 0.5 mol bFGF bound per heparin. Since the results obtained from light scattering indicate that the complexes contain six to seven bFGF, we can conclude that there is only one HMWH molecule in the complex and that HMWH has an average of six to seven binding sites for bFGF. More detailed analysis and the stoichiometry of low-molecular-weight heparin with bFGF and other protein–carbohydrate interactions can be found in Refs. 29–32.

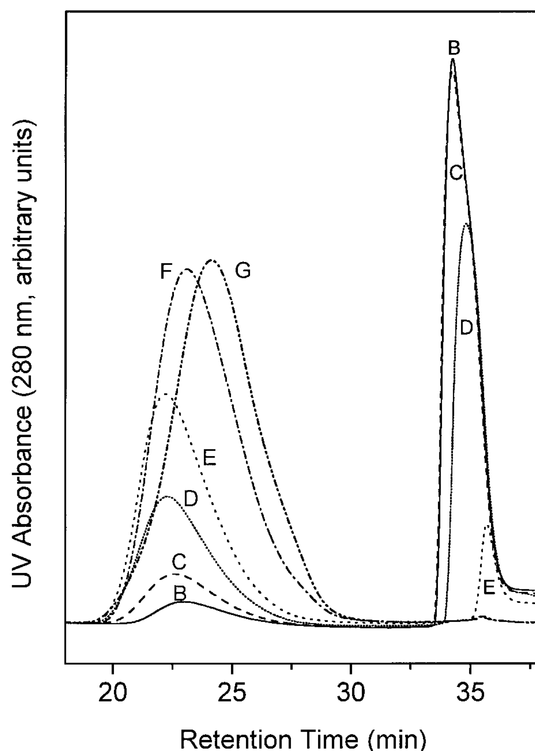


FIG. 5. Chromatograms of bFGF and HMWH mixtures (uv absorbance traces only). All samples contained 117 μM bFGF, and the amount of HMWH varied. The [HMWH]/[bFGF] molar ratio was 0.011:1 (B), 0.022:1 (C), 0.056:1 (D), 0.11:1 (E), 0.22:1 (F), and 0.43:1 (G). Sample A (bFGF control) is not shown here. One hundred microliters of each mixture was injected onto a Superdex 200 column (Pharmacia) with PBS as eluant at a 0.5 ml/min flow rate.

TABLE 1
Summary of the Molecular Weights of bFGF/HMWH Complexes[a]

Sample name in figure	Mixture molar ratio [HMWH]/[bFGF]	Peak at 19–29 min: Average M_r	Peak at 19–29 min: Intensity (based on uv)	Free bFGF at ~34 min (based on uv)
A	Control (no HMWH)	No peak	0.0	14
B	0.011:1	101,000	0.46	11
C	0.022:1	102,000	1.0	11
D	0.056:1	86,000	2.6	8.1
E	0.11:1	84,000	4.6	2.2
F	0.22:1	77,000	7.0	0.0
G	0.43:1	44,000	7.2	0.0

[a] All molecular weights in this table reflect only the polypeptide components of the complexes.

IV. STOICHIOMETRY OF A PROTEIN–PROTEIN COMPLEX CONTAINING NO CARBOHYDRATES

A. Theoretical Background

When a protein interacts with other proteins and forms a complex, there are once again two SEC-LS/UV/RI analysis methods depending on whether or not the complex contains carbohydrates. If the complex does not contain carbohydrates, the two-detector method can be directly used to get the total polypeptide molecular weight of the complex and derive its stoichiometry. Such complexes may be either covalent or noncovalent, but in the latter case the affinity of the reversibly interacting components must be sufficiently high to maintain the integrity of the complex during chromatography.

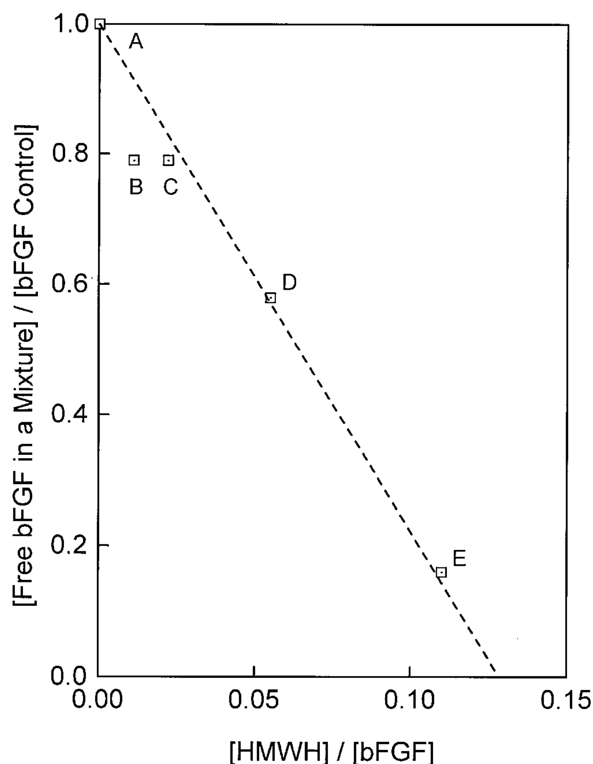


FIG. 6. Relative peak height of the free bFGF peak at ~34 min ([free bFGF in a mixture sample]/[bFGF control]) vs amount of HMWH added ([HMWH]/[bFGF]). The dashed line is the best fit of these data to a straight line, which has an intercept of 1.00 and a slope of 7.5 ± 0.5 mol free bFGF lost per mole of HMWH.

B. Application

Tumor necrosis factor(TNF)-α is a multifunctional cytokine that plays an important role in many inflammatory and immune responses (33, 34). It exists as a trimer in solution (35, 36). Two types of receptors for TNF have been identified and characterized as proteins of 55 kDa (the type I receptor) and 75 kDa (the type II receptor) (37–39). TNF exerts its biological effect when interacting with its cell surface receptors. In this section, we will show the results of using SEC-LS/UV/RI to study the stoichiometry of the complexes of two nonglycosylated proteins, human TNF-α trimer (TNF) and the extracellular domain of the TNF type I receptor (sTNFR) (40). Before studying mixtures of TNF and sTNFR, TNF and sTNFR controls were separately injected onto a Superose 12 column. Their chromatograms are shown in Figs. 7A and 7B. The molecular weight from the two-detector method for each protein alone is 52,000 for TNF and 17,000 for sTNFR. A mixture made at a molar ratio of around three sTNFR per TNF was then studied under the same experimental conditions, giving the chromatogram shown in Fig. 7C. Complexes of sTNFR with TNF elute as a broad distribution from 17.5 to 22 min. The molecular weight calculated by the two-detector method for the peak of this distribution at around 18.5 min is 107,000, indicating that the stoichiometry of this complex is three sTNFR plus one TNF trimer. Mixtures made at 2:1 or 1:1 molar

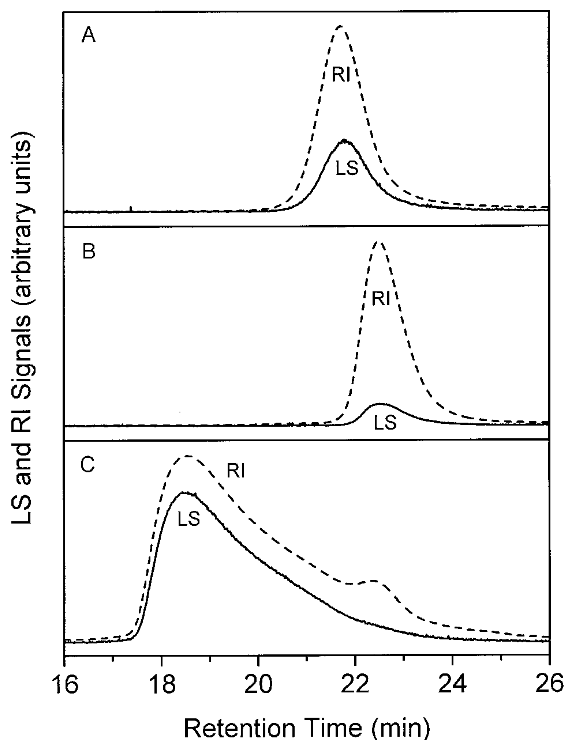


FIG. 7. Chromatograms of TNF, sTNFR, and the mixture of sTNFR and TNF. One hundred microliters of each protein was injected onto a Superose 12 column (Pharmacia) with PBS as eluant at a flow rate of 0.5 ml/min. A, TNF control sample (no sTNFR); B, sTNFR control sample (no TNF); C, a mixture of sTNFR and TNF made at a molar ratio of around three sTNFR per TNF. The lines are the same as defined in the legend to Fig. 2.

ratios were also studied, and in such samples additional types of complexes with either one or two sTNFR per TNF trimer were detected in different percentages in each mixture.

V. STOICHIOMETRY OF A PROTEIN–PROTEIN COMPLEX CONTAINING CARBOHYDRATES

A. Theoretical Background

If a protein complex contains carbohydrates, the three-detector method is required to calculate its polypeptide molecular weight and determine its stoichiometry. In this section, we will focus on how to determine the stoichiometry of such a complex (41).

To use the three-detector method, we must know the polypeptide extinction coefficient of the complex. In most common circumstances, only the polypeptide extinction coefficients of each protein in a complex are known, and thus we need to calculate the polypeptide extinction coefficient of the complex as a whole. The polypeptide extinction coefficient of a complex, ϵ_p, with a known stoichiometry (A_mB_n) can be calculated using the equation

$$\epsilon_p = (m\epsilon_A M_A + n\epsilon_B M_B)/(mM_A + nM_B), \qquad [10]$$

where ϵ_A, ϵ_B, M_A, and M_B are the polypeptide extinction coefficients and molecular weights of protein A or B. After obtaining ϵ_p, we can calculate the polypeptide molecular weight by Eq. [9] and, hence, determine the stoichiometry.

It is obvious that this is a circular argument. On the one hand, we want to use Eq. [10] to calculate the polypeptide extinction coefficient of the complex and then use Eq. [9] to determine the corresponding molecular weight and the stoichiometry; on the other hand, the polypeptide extinction coefficient of the complex cannot be calculated from Eq. [10] until the stoichiometry of the complex is known. To solve this conundrum, a self-consistent method has been developed. In this method we first assume various possibilities for the stoichiometry of the complex. For each assumed stoichiometry we then calculate its corresponding theoretical molecular weight from those of its components and also its experimental molecular weight from Eqs. [9] and [10]. Finally, we select the stoichiometry with the best consistency between the experimental and theoretical molecular weights as the correct stoichiometry for the complex.

B. Applications

One important group of protein interactions is that leading to oligomerization of cell surface receptors, which is believed to be the key initiator of signal transduction for many cytokine and growth factor receptors (42, 43). For example, when brain-derived neurotrophic factor (BDNF) and neurotrophin-3 (NT-3) bind to their corresponding receptors, TrkB and TrkC, activation of the receptor tyrosine kinase and receptor autophosphorylation occur, and this leads to nerve growth, differentiation, or survival (44, 45). To understand the binding stoichiometry, the interactions of the extracellular domains of the TrkB and TrkC neurotrophin receptors (sTrkB and sTrkC) with BDNF and NT-3 were studied using the technique described above. When BDNF was injected onto a Superdex 200 column with Dulbecco's phosphate-buffered saline (PBS) as eluant, it showed no elution peak under these conditions because of its interaction with the column matrix. High-ionic-strength buffer may elute BDNF, but such buffers may interfere with the interaction of BDNF with sTrkB. Fortunately, the complex of BDNF with sTrkB elutes in PBS, and thus PBS was used for this study. sTrkB is a glycoprotein and elutes with no indication of interaction with the column (Fig. 8A). The molecular weight determined from the three-detector method for sTrkB is 44,000, in agreement with the sequence molecular weight (Table 2), indicating that sTrkB exists as a monomer in solution. The chromatogram of a sTrkB/

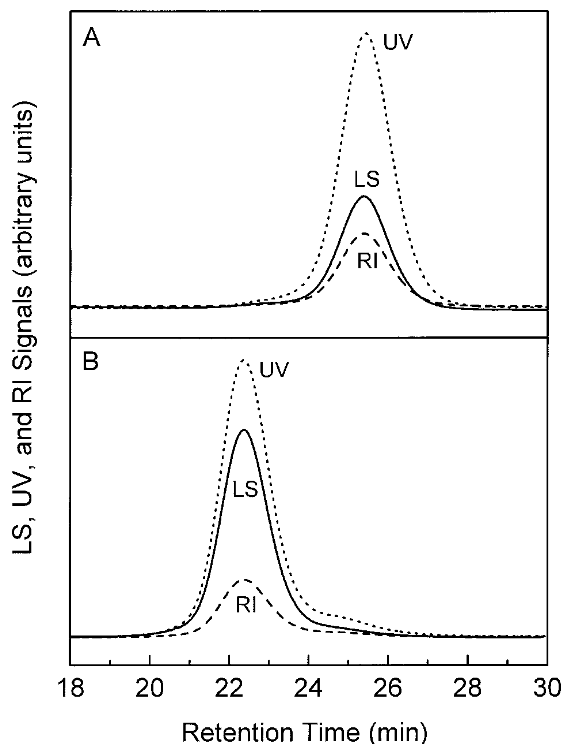


FIG. 8. Chromatograms of sTrkB and the mixture of sTrkB and BDNF. One hundred microliters of each protein was injected onto a Superdex 200 column (Pharmacia) with PBS as eluant at a flow rate of 0.5 ml/min. A, sTrkB control sample (no BDNF); B, a mixture of sTrkB and BDNF made at a molar ratio of around two sTrkB per BDNF. The lines are the same as defined in the legend to Fig. 4.

BDNF mixture made at a molar ratio of two sTrkB per BDNF dimer is shown in Fig. 8B. To calculate the molecular weight and stoichiometry of the sTrkB/BDNF complex, it is necessary to calculate its polypeptide extinction coefficient as discussed under Theoretical Background. However, the extinction coefficient cannot be calculated until the stoichiometry of the complex is known. Therefore, we assume that either one or two sTrkB bind to one BDNF dimer and calculate the corresponding experimental and theoretical molecular weights. The results are summarized in Table 2. Experimental molecular weights were calculated from Eqs. [9] and [10], and theoretical molecular weights were calculated from the sequence molecular weights of each component under each assumed stoichiometry. Obviously, the experimental molecular weight agrees with the theoretical weight under the assumption that two sTrkB bind to one BDNF dimer. Therefore, we conclude that the stoichiometry, 2sTrkB:1BDNF dimer, is the correct one for the sTrkB/BDNF complex. Using the same method, sTrkC, NT-3, and a mixture of sTrkC/NT-3 were studied, and the results are summarized in the second part of Table 2. The results indicate that NT-3 can also dimerize sTrkC. Other techniques were also used to study these interactions and the results agree with each other quite well (46).

For "orphan" receptors with unknown ligands, monoclonal antibodies (mAbs) directed against the receptor are often used as "artificial ligands" (47, 48) that presumably activate by dimerizing the receptor. In the second example, we will study such interactions: the interactions of sHer2, which is the extracellular domain of erbB2/Her2 tyrosine kinase receptor expressed on breast cancer cells, with monoclonal antibodies prepared against it. One hundred-microliter samples of sHer2, mAb35, mAb52, mAb58, and mAb42b controls, as well as mixtures of sHer2 with each mAb, were injected into the SEC-LS/UV/RI system separately. The results are summarized in Table 3 and the chromatograms of one set of samples (sHer2 and mAb35) are shown in Fig. 9. The molecular weights of mAbs were calculated using a typical average extinction coefficient value for mAbs of 1.4 ml/(mg · cm) since their amino acid compositions have not been determined. The molecular weights thus determined varied around 140,000, with at least some of this variation probably being due to the uncertain extinction

TABLE 2

Summary of sTrkB/BDNF and sTrkC/NT-3 Results[a]

Samples	Assumed stoichiometry	ϵ [ml/(mg · cm)]	Experimental molecular weight from LS ±5%	Theoretical molecular weight from sequence	Correct assumption?
BDNF dimer		1.6	27,000	27,300	
NT-3 dimer		2.17	28,000	27,500	
sTrkB		1.15	44,000	44,100	
sTrkC		1.19	45,000	44,700	
sTrkB + BDNF mixture	1sTrkB:1BDNF dimer	1.32	110,000	71,000	No
	2sTrkB:1BDNF dimer	1.26	115,000	116,000	Yes
sTrkC + NT-3 mixture	1sTrkC:1NT-3 dimer	1.56	118,000	72,000	No
	2sTrkC:1NT-3 dimer	1.42	130,000	117,000	Yes

[a] All extinction coefficients and molecular weights in this table reflect only the polypeptide components of the complexes.

TABLE 3

Summary of sHer2 and mAbs' Light-Scattering Results[a]

Samples	Assumed stoichiometry	ϵ [ml/(mg·cm)]	Expt *MW* from LS ± 5%	Theor *MW* from sequence	Correct assumption?
sHer2		0.85	69,000		
mAb 35		1.4	139,000		
mAb 52		1.4	151,000		
mAb 58		1.4	142,000		
mAb 42b		1.4	136,000		
sHer + mAb35 mixture	1sHer2:1mAb35	1.24	237,000	208,000	No
	2:1	1.14	261,000	277,000	Yes
	3:1	1.08	275,000	346,000	No
	1:2	1.31	226,000	347,000	No
	1:3	1.41	208,000	486,000	No
sHer + mAb52 mixture	1sHer2:1mAb52	1.24	252,000	220,000	No
	2:1	1.14	275,000	289,000	Yes
	3:1	1.08	289,000	358,000	No
	1:2	1.31	240,000	371,000	No
	1:3	1.41	223,000	522,000	No
sHer + mAb58 mixture	1sHer2:1mAb58	1.24	252,000	211,000	No
	2:1	1.14	272,000	280,000	Yes
	3:1	1.08	289,000	348,000	No
	1:2	1.31	237,000	353,000	No
	1:3	1.41	220,000	522,000	No
sHer + mAb42b mixture	1sHer2:1mAb42b	1.24	246,000	205,000	No
	2:1	1.14	266,000	274,000	Yes
	3:1	1.08	281,000	343,000	No
	1:2	1.31	232,000	341,000	No
	1:3	1.41	214,000	477,000	No

[a] All extinction coefficients and molecular weights in this table reflect only the polypeptide components of the complexes.

coefficient values used for the calculation. The molecular weight of sHer2 was calculated with the extinction coefficient estimated from the amino acid sequence, and the result agrees well with its sequence molecular weight, indicating that sHer2 exists as a monomer in solution. The mixtures made with excess sHer2 all showed a peak eluting earlier than the mAb or sHer2 control. This indicates that complexes were formed in all mixtures. Different possibilities were assumed for the stoichiometry of each complex, and the corresponding experimental and theoretical molecular weights were thus calculated (Table 3). For all complexes, the experimental molecular weights are most consistent with the theoretical weights under the assumption that each mAb binds two sHer2 molecules for all of the different mAbs tested, showing that these antibodies do dimerize sHer2.

More applications using this self-consistent three-detector methods to study protein interactions can be found in Refs. 49–52.

VI. SUMMARY

SEC-LS/UV/RI has been developed as a fast, accurate, and reliable technique to characterize the molecular weight of a protein or the stoichiometry of a protein complex. Two unique characteristics of SEC-LS/RI/UV, in addition to having all the capabilities of conventional SEC, are (i) that its molecular weight measurement is independent of elution position and (ii) that the molecular weight calculated may exclude carbohydrates. The two-detector method was used to calculate the molecular weights of proteins containing no carbohydrate, including BSA dimer and native or reduced RNase. The molecular weight of BSA dimer calculated by this method agrees with its sequence molecular weight. The reduced RNase was identical, in molecular weight, to the native RNase, which was a result very different from that derived from the conventional SEC elution position method. In addition, when the proteins are not glycosylated, this two-detector method can even be applied to protein–protein complexes. sTNFR and TNF were shown mainly to form a 3:1 complex when mixed at approximately three sTNFR per TNF trimer. In this calculation, no information on amino acid sequence or extinction coefficient was required. This method may therefore be most useful for studying interactions between signal transduction mediators which play important roles in cytoplasmic signaling since they are not glycosylated. The three-detector method was used to determine the molecular weight of a protein con-

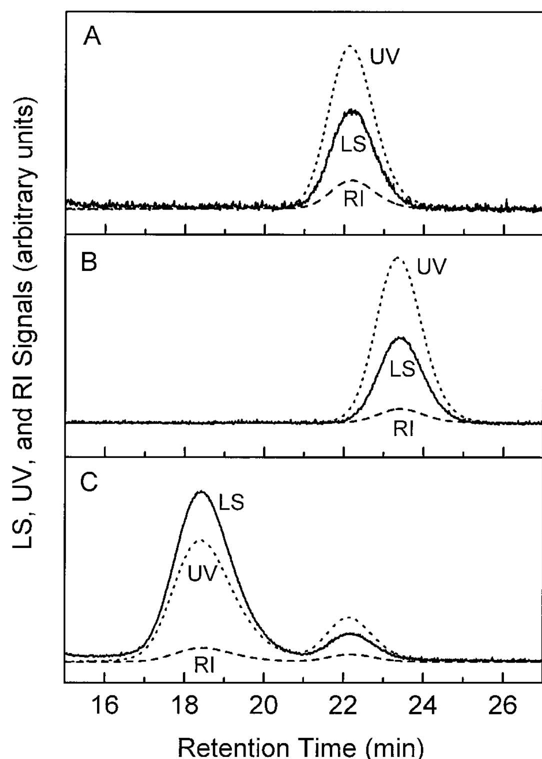


FIG. 9. Chromatograms of mAb35, sHer2, and the mixture of sHer2 and mAb35. One hundred microliters of each protein was injected onto a Superdex 200 column (Pharmacia) with PBS as eluant at a flow rate of 0.5 ml/min. A, mAb35 control sample (no sHer2); B, sHer2 control sample (no mAb35); C, a mixture of mAb35 and sHer2 made at a molar ratio of around 2.7 sHer2 per mAb35. The lines are the same as defined in Fig. 4.

taining carbohydrates. The molecular weight of CHO SCF was thus determined as 38,000 and HMWH was shown to bind six to seven bFGF molecules per HMWH. To determine the stoichiometry of a protein–protein complex containing carbohydrates, a self-consistent three-detector method was developed. The stoichiometry of both sTrkB/BNDF and sTrkC/NT-3 complexes was shown to be 2:1, and anti-Her2 mAbs were shown to dimerize sHer2.

ACKNOWLEDGMENTS

We thank Keith Langley and Jette Wypych for providing us with SCF, Mary Rosendahl for sTNFR, Jane Talvenheimo and Andrew Welcher for sTrkB and sTrkC, Tom Horan for sHer2, and Yoshiko Kita, Julia Tseng, and Margery Nicolson for anti-Her2 mAbs.

REFERENCES

1. Dollinger, G., Cunico, B., Krunitani, M., Johnson, D., and Jones, R. (1992) *J. Chromatogr.* 592, 215–228.
2. Takagi, T. (1990) *J. Chromatogr.* 506, 409–416.
3. Stuting, H. H., Krull, I. S., Krzysko, S. C., and Barth, H. G. (1992) *LC·GC* 7, 402–415.
4. Wyatt, P. J. (1993) *Anal. Chim. Acta* 272, 1–40.
5. Maezawa, S., and Takagi, T. (1983) *J. Chromatogr.* 280, 124–130.
6. Takagi, T. (1985) *in* Progress in HPLC (Parvez, H., Kato, Y., and Parvez, S., Eds.), Vol. 1, pp. 27–41, VNU Science Press, Utrecht, The Netherlands.
7. Stuting, H. H., and Krull, I. S. (1990) *Anal. Chem.* 62, 2107–2114.
8. Krull, I. S., Stuting, H. H., and Krzysko, S. C. (1988) *J. Chromatogr.* 442, 29–52.
9. Pittz, E. P., Lee, J. C., Bablouzian, B., Townend, R., and Timasheff, S. N. (1973) *in* Methods in Enzymology, Vol. XXVII (Hirs, C. H. W., and Timasheff, S. N., Eds.), pp. 209–256, Academic Press, San Diego.
10. Wyatt, P. J., Jackson, C., and Wyatt, G. K. (1988) *Am. Lab.* 20, 86–89.
11. Cantor, C. R., and Schimmel (1980) *in* Biophysical Chemistry, Part II, pp. 628–629, Freeman, San Francisco.
12. Jackson, C., and O'Brien, J. P. (1995) *Macromolecules* 28, 5975–5977.
13. Sobieszek (1995) *Biochemistry* 34, 11855–11863.
14. Philo, J. S., Rosenfeld, R., Arakawa, T., Wen, J., and Narhi, L. O. (1993) *Biochemistry* 32, 10812–10818.
15. Zhang, M. Z., Wen, J., Arakawa, T., and Prestrelski, S. J. (1995) *Pharm. Res.* 12, 1447–1452.
16. Narhi, L. O., Rosenfeld, R., Wen, J., Arakawa, T., Prestrelski, S. J., and Philo, J. S. (1993) *Biochemistry* 32, 10819–10825.
17. Gombotz, W. R., Pankey, S. C., Phan, D., Drager, R., Donaldson, K., Antonsen, K. P., Hoffman, A. S., and Raff, H. V. (1994) *Pharm. Res.* 11, 624–632.
18. Kenney, W. C., Haniu, M., Herman, A. C., Arakawa, T., Costigan, V. J., Lary, J., Yphantis, D. A., and Thomason, A. R. (1994) *J. Biol. Chem.* 269, 12351–12359.
19. Samal, B. B., Arakawa, T., Boone, T. C., Jones, T., Prestrelski, S. J., Narhi, L. O., Wen, J., Strearns, G. W., Crandall, C. A., Pope, J., and Suggs, S. (1995) *Biochim. Biophys. Acta Gene Struct. Expr.* 1260, 27–34.
20. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* 182, 319–326.
21. Pace, C. N., Vajdos, F., Fee, L., Grimsley, and Gray, T. (1995) *Protein Sci.* 4, 2411–2423.
22. Zsebo, K. M., Wypych, J., McNiece, I. K., Lu, H. S., Smith, K. A., Karkare, S. B., Sachdev, R. K., Yuschenkoff, V. N., Birkett, N. C., Williams, L. R., Satyagal, V. N., Tung, W., Bosselman, R. A., Mendiaz, E. A., and Langley, K. E. (1990) *Cell* 63, 195–201.
23. Arakawa, T., Yphantis, D. A., Lary, J. W., Narhi, L. O., Lu, H. S., Prestrelski, S. J., Clogston, C. L., Zsebo, K. M., Mendiaz, E. A., Wypych, J., and Langley, K. E. (1991) *J. Biol. Chem.* 266, 18942–18948.
24. Arakawa, T., Kameyama, K., and Takagi, T. (1992) *Anal. Biochem.* 203, 53–57.
25. Lu, H. S., Chang, D., Philo, J. S., Zhang, K., Narhi, O. L., Liu, N., Zhang, M., Sun, J., Wen, J., Yanagihara, D., Karunagaran, D., Yarden, Y., and Ratzkin, B. (1995) *J. Biol. Chem.* 270, 4784–4791.
26. Burgess, H., and Maciag, T. (1989) *Annu. Rev. Biochem.* 58, 575–606.
27. Folkman, J., and Klagsbrun, M. (1987) *Science* 235, 442–447.
28. Prestrelski, S. J., Fox, G. M., and Arakawa, T. (1992) *Arch. Biochem. Biophys.* 213, 314–319.

29. Arakawa, T., Wen, J., and Philo, J. S. (1994) *Arch. Biochem. Biophys.* 308, 267–273.
30. Wen, J., Arakawa, T., Wypych, J., Langley, K. E., Schwartz, M. G., and Philo, J. S (1997) *in* Techniques in Protein Chemistry VIII (Marshak, D. R., Ed.), Academic Press, San Diego, in press.
31. Arakawa, T., Holst, P., Narhi, L. O., Philo, J. S., Wen, J., Prestrelski, S. J., Zhu, X., Rees, D. C., and Fox, G. M. (1995) *J. Protein Chem.* 14, 263–274.
32. Kato, A., Kameyama, K., and Takagi, T. (1992) *Biochim. Biophys. Acta* 1159, 22–28.
33. Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, N., and Williamson, B. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3666–3670.
34. Beutler, B., and Cerami, A. (1988) *Annu. Rev. Immunol.* 7, 625–655.
35. Erc, M. J., and Sprang. S. R. (1989) *J. Biol. Chem.* 264, 17595–17605.
36. Arakawa, T.. and Yphantis, D. A. (1987) *J. Biol. Chem.* 262, 7484–7485.
37. Loetscher, H., Pan, Y.-C. E., Lahm, H.-W., Gentz, R., Brockhaus, M., Tabuchi, H., and Lesslauer, W. (1990) *Cell* 61, 351–359.
38. Schall, T. J., Lewis, M., Koller, K. J., Lee, A., Rice, G. C., Wong, G. H. W., Gatanaga, T., Granger, G. A., Lentz, R., Raab, H., Kohr., W. J., and Goeddel, D. V. (1990) *Cell* 61, 361–370.
39. Smith, C. A., Davis, T., Anderson, D., Solam, L., Beckmann, M. P., Jerzy, R., Dower, S. K., Cosman, D., and Goodwin, R. G. (1990) *Science* 248, 1019–1023.
40. Wen, J., Philo, J. S., Arakawa, T., Rosendahl, M., and Narhi, L. O. (1996) *Protein Science* 5 (Suppl. 1), 89.
41. Wen, J., Arakawa, T., Talvenheimo, J., Welcher, A., Horan, T., Kita, T., Tseng, J., Nicolson, M., and Philo, J. S. (1996) *in* Techniques in Protein Chemistry VII (Marshak, D. R., Ed.), pp. 23–31, Academic Press, San Diego.
42. Blume-Jensen, P., Claesson-Welsh, L., Sieghahn, A., Zsebo, K. M., Westermark, B., and Heldin, C.-H. (1991) *EMBO J.* 10, 4121–4128.
43. Li, W., and Stanley, E. R. (1991) *EMBO J.* 10, 277–288.
44. Glass, D. J., Nye, S. H., Hantzopoulos, P., Macchi, M. J., Squinto, S. P., Goldfarb, M., and Yancopoulos, G. D. (1991) *Cell* 66, 405–413.
45. Lamballe, F., Klein, R., and Barbacid, M. (1991) *Cell* 66, 967–979.
46. Philo, J., Talvenheimo, J., Wen, J., Rosenfeld, R., Welcher, A., and Arakawa, T. (1994) *J. Biol. Chem.* 269, 27840–27846.
47. Harwerth, I.-M., Wels, W., Schlegel, J., Muller, M., and Hynes, N. E. (1993) *Br. J. Cancer* 68, 1140–1145.
48. Srinivas, U., Tagliabue, E., Campiglio, M., Menard, S., and Colnaghi, M. I. (1993) *Cancer Immunol. Immunother.* 36, 397–402.
49. Philo, J. S., Aoki, K. H., Arakawa, T., Narhi, L. O., and Wen, J. (1996) *Biochemistry* 35, 1681–1691.
50. Philo, J. S., Wen, J., Wypych, J., Schwartz, M. G., Mendiaz, E. A., and Langley, K. E. (1996) *J. Biol. Chem.* 271, 6895–6902.
51. Horan, T., Wen, J., Narhi, L., Parker, V., Arakawa, T., and Philo, J. (1996) *Biochemistry* 35, 4886–4896.
52. Horan, T., Wen, J., Arakawa, T., Liu, N., Brankow, D., Hu, S., Ratzkin, B., and Philo, J. S. (1995) *J. Biol. Chem.* 270, 24604–24608.