Journal of Chromatography, 588 (1991) 125–137
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 643

On-line characterization of polyethylene glycol-modified proteins

Michael Kunitani*, Gavin Dollinger, Deborah Johnson and Lilia Kresin

Analytical Chemistry Department, Cetus Corporation, 1400 Fifty-Third Street, Emeryville, CA 94608 (USA)

(First received March 8th, 1991; revised manuscript received July 16th, 1991)

ABSTRACT

An high-performance liquid chromatographic method has been developed which simultaneously determines three critical physical properties of polyethylene glycol (PEG)-modified proteins: molecular size, polymer distribution and weight composition. With both UV and refractive index (RI) detectors in series, size-exclusion chromatography (SEC) is used to separate the PEG–protein species according to size. The size analysis of these PEG–proteins is predicted to be accurately calibrated with the viscosity radius (universal calibration), which compensates for the shape differences between PEG and protein structures. The heterogeneity of the PEG–protein grafted copolymer is represented by the polymeric term "polydispersity", which describes the size distribution. Separate SEC calibrations of the PEG and the protein used for conjugation allow a determination of the weight composition of the PEG–protein (weight PEG/weight protein) by combining UV and RI chromatograms of a PEG–protein sample. This compositional analysis is validated through independent and direct measurement of the PEG on a PEG–protein via acid hydrolysis and quantitative SEC. Comparisons of compositional analysis of PEG–protein with sodium dodecyl sulfate polyacrylamide gel electrophoresis densitometry demonstrate that gel analysis of some proteins is misleading.

INTRODUCTION

Covalent conjugation of monomethoxy–polyethylene glycol (PEG) to proteins has recently become recognized as a method of dramatically altering a protein's pharmacology and immunogenicity. A number of enzymes, including superoxide dismutase, asparaginase and uricase, have been extensively modified with PEG, resulting in biologically active compounds with longer *in vivo* half-lives [1–5]. Investigators have conjugated PEG to lymphokines such as interleukin-2(IL-2), produced from *Escherichia coli*, in order to mimic some of the effects of native glycosylation [6]. PEG modification of IL-2 has led to increased protein solubility, decreased plasma clearance, increased antitumor potency in mice and reduced immunogenicity in rabbits and mice [7–9].

Many diverse approaches have been taken in the attempt to characterize PEG–proteins. The chemical reactivity of underivatized primary amino groups on lysine has been exploited in order to determine the degree of PEG conjugation [10,11]. This approach has had difficulty in accurately measuring low degrees of PEG modification. It relies upon differentiating between the derivatization of a native protein, which contains numerous lysine residues, and the PEG-modified protein. The result is a measurement of the difference between two relatively large numbers and, when that difference is small, the relative error can be considerable. Perhaps because of this limitation, rigorous validation of the method has been difficult [10]. Another approach to determining the degree of PEG conjugation has been through the densitometric analysis of sodium dodecyl sulfate (SDS) electrophoretic gels [7,11]. The resolution is sufficient to identify various gel band regions, but the gel bands are broad. Quantitative NMR has been used to determine the degree of PEG conjugation to ovalbumin [12]. Overlap of PEG and protein chemical shift resonances will limit the accuracy of this approach. Other approaches, in-

0021-9673/91/$03.50

cluding dynamic light scattering and size-exclusion chromatography, use the hydrodynamic size of the modified protein to infer the degree of PEG modification. With these methods, the relationship between hydrodynamic size and mass is not well defined owing to the difference in partial specific volume between the quasi-random coil configuration of PEG and the compact globular shape of native proteins.

Other chromatographic modes that have been used in characterizing PEG-proteins are reversed-phase, hydrophobic interaction, anion-exchange and cation-exchange high-performance liquid chromatography (HPLC) [11,12]. These four modes of absorptive chromatography all have demonstrated the ability to resolve clearly two or more peaks of PEG-conjugated superoxide dismutase (PEG–SOD) or PEG–ovalbumin. Common to these chromatographic separations are the broad HPLC peaks characteristic of polymeric mixtures. The identity and composition of these peaks are unknown. An approach to characterizing PEG–proteins without relying on surface interactions or molecular size is isoelectric focusing (IEF), which separates molecules on the basis of molecular charge. Presumably, higher degrees of PEG conjugation through lysine derivatization would be represented by bands of lower p*I*. IEF gels of PEG–SOD and PEG-conjugated IL-2 (PEG–IL-2) reveal either a continuous smear or a series of bands of lower p*I* than the native protein [7,11]. The complex array of IEF bands have been attributed to the degree of PEG conjugation and to conformational heterogeneity. Recently, capillary electrophoresis has shown promise in separating PEG–proteins, but is limited to those with relatively low degrees of PEG conjugation [13].

Native proteins, by virtue of a unique primary sequence, have well defined secondary and tertiary structures. In comparison, PEG is a linear polymer which always exists as a distribution of different lengths and random conformations. When PEG is conjugated to a protein, two additional dimensions of heterogeneity are introduced: the number of PEGs per protein molecule and their location. These three dimensions of heterogeneity (length of each PEG, number of PEGs and location of PEGs) all contribute to the effective hydrodynamic size of the PEG–protein conjugate. For example, PEGs may be on the same side or opposite sides of a small protein, resulting in vastly different hydrodynamic diameters for molecules of identical composition and molecular weight. Conversely, a narrow size range of a PEG–protein will contain many distinctly different PEG–protein molecules. Let us consider conjugating PEG to a relatively small protein, ribonuclease, which contains ten lysine residues and one reactive N-terminus. With a PEG of moderate distribution (molecular weight 7000–10 000), the number of combinations of ribonuclease conjugated with three PEGs is $5 \cdot 10^7$. With four PEGs it is $7 \cdot 10^9$. Owing to steric constraints and kinetic factors, not all of these combinations of PEG–proteins will occur. However, it is clear that the immense heterogeneity of PEG–proteins explains the limited resolution and broad peak width seen with the previously mentioned separation techniques.

Polymer chemists have previously characterized compounds such as PEG–proteins and have classified them as grafted copolymers, defined as "a polymer with a small number of long grafts attached to one backbone" [14]. As demonstrated by the previous calculation of theoretically possible PEG–protein heterogeneity, polymers exist as a large family of similar compounds exhibiting group characteristics. This perception diverges dramatically from that of protein chemists, for whom protein microheterogeneity is discrete and, at least in theory, resolvable. Examples of protein HPLC exist, such as that of recombinant IL-2, where variants differing by a single atom are readily resolved on reversed phase [15]. For grafted copolymers such as PEG–proteins, heterogeneity is not discrete, but is a semi-continuous distribution of molecules which can only be described through statistical measurements such as width, average, skew, etc. Analyte properties are not unique to a given mixture, since many other mixtures of similar compounds could have identical group properties and chromatographic behavior.

Sizing has been a traditional characterization of polymers, useful in determining group parameters which could be related to the physical chemical behavior of the PEG–protein mixture. For example, Knauf *et al.* [7] have established that the effective hydrodynamic size of a PEG–protein can be a critical pharmacological measurement. In their study, different numbers and lengths of PEG were conjugated to IL-2 to yield mixtures varying in their

effective hydrodynamic size, as measured by SEC. The pharmacokinetic behavior of the PEG-IL-2 species in rats showed a close correlation between clearance and PEG–IL-2 size. It was suggested that the abrupt reduction in clearance seen above 70 000 dalton was due to the permeability threshold of the kidney, which retains proteins larger than albumin in the plasma. However, the size-exclusion chromatographic (SEC) calibration used by Knauf *et al.* was derived with native protein molecular weight and did not account for the increased hydrodynamic volume of the PEG.

The difficulty in calibrating an SEC system for PEG–proteins lies in the inherent differences in shape between PEG and proteins. On aqueous SEC, PEG chromatographs as a random coil. However, viscometric, calorimetric and several spectroscopic approaches indicate that solutions of aqueous PEG retain a degree of helical character, probably in relatively short sections [16–18]. Further, the molar specific volume of aqueous PEG is smaller than that expected from a totally random coil. Thus the overall structure of aqueous PEG is most likely a quasi-random coil or partially random coil conformation with numerous helical sections. This is very different from the compact globular native conformation of a protein.

For years, many biochemists have assumed the relationship between SEC retention and the logarithm of molecular weight (log MW) to be linear, and have used commercially available protein standards to derive apparent molecular weights of analyte proteins. Under strongly denaturing conditions of high salt (*e.g.*, 6 *M* guanidine hydrochloride), this relationship holds true for most peptides and proteins [19]. Native proteins are defined by their characteristic secondary, tertiary and quaternary structures which may deviate from spherical geometry and thus represent exceptions to the SEC retention to log MW linearity. To account for the dependence of SEC retention on molecular shape and not molecular weight, some researchers have used the intrinsic viscosity-based radius (R_η) instead of the log MW term [20,21]. Called universal calibration, this method is a major improvement over molecular weight calibration and is generally more accurate than that of translational frictional-based Stokes radius (R_s) calibration. Previous researchers have shown that both globular proteins and random coil macromolecules such as pullulans and dextrans can be characterized through universal calibration techniques [22,23]. However, the absolute accuracy of universal SEC calibration for biological molecules is still an issue of debate in the literature, as exceptions have been found [22,24]. The calibration of the SEC system appropriate for PEG–proteins will be explored in this study.

The fundamental polymer characteristics obtained from SEC are the number–average molecular weight ($\bar{M}_n$), the weight–average molecular weight ($\bar{M}_w$) and polydispersity. The number-average molecular weight is traditionally defined as

$$\bar{M}_n = \frac{\sum N_i M_i}{\sum N_i} \tag{1}$$

where N_i is the number of molecules and M_i is the molecular weight of a given chromatographic increment. The number-average molecular weight is an arithmetic mean of the sizing distribution, and assumes that each molecule makes an equal contribution to the polymer property regardless of size or weight.

The weight-average molecular weight is defined as

$$\bar{M}_w = \frac{\sum N_i M_i^2}{\sum N_i M_i} \tag{2}$$

The weight-average molecular weight is a weighted average in which each molecule contributes in accordance with its weight.

The polydispersity is defined as

$$\text{polydispersity} = \frac{\bar{M}_w}{\bar{M}_n} \tag{3}$$

The polydispersity value is characteristic of a polymer distribution. For a Gaussian distribution the polydispersity is 1.0, although in practice its value is nearly always greater. $\bar{M}_n$ and $\bar{M}_w$ are normally calculated with commercially available data reduction programs and calibrated with narrow-range molecular weight standards. The size-exclusion chromatogram is divided into slices (*e.g.*, 0.1-min increments) for which the area (N_i) and molecular weight (M_i) are determined from the area and calibrated mass of each slice.

THEORY, PEG–PROTEIN COMPOSITION

The PEG composition (mass of PEG/mass of protein) can be determined using two different detectors in series: UV at 280 nm (A_{280}) for selectively measuring protein and refractive index (RI) for measuring both protein and PEG. Analysis of copolymer composition has been previously developed for traditional polymers using UV detection at different wavelengths and also the combined UV–RI detection employed here [25–27]. The weight of the protein portion of the PEG–protein, as derived from its UV absorbance, is used to determine the protein's contribution to the refractive index response of the PEG–protein. The remaining refractive index response of the PEG–protein is due to the PEG, and can be converted to a weight value through the appropriate calibrations. In the most general case where both copolymer components have UV and RI responses, the two unknown weights can be determined in a similar fashion from two independent concentration detectors.

The fundamental assumption of this analytical method is that the RI response of the PEG–protein is the sum of the RI responses of the PEG and the protein alone. The RI of a component is characteristic of the electron density of that component. As there is a single covalent bond between two large molecules, protein and PEG, the electron density of each should be essentially unchanged, and the approximation of their RI additivity should be accurate (eqn. 4) [14].

$$n = n_0 + \left(\frac{dn}{dc}\right)_p C_p + \left(\frac{dn}{dc}\right)_x C_x \tag{4}$$

where n_0 is the refractive index of the solvent, $(dn/dc)_p$ and $(dn/dc)_x$ are the specific RI increments for PEG and protein x, respectively, and C_p and C_x are the weight concentrations of the PEG and protein x per unit volume. Rearrangement of eqn. 4 yields

$$\frac{n - n_0}{C_x} \equiv \left(\frac{dn_{px}}{dc_x}\right) = \left(\frac{dn}{dc}\right)_p \frac{C_p}{C_x} + \left(\frac{dn}{dc}\right)_x \tag{5}$$

The left-hand part of eqn. 5 can be redefined as the RI increment for the PEG–protein as a function of the change in the protein portion of this copolymer (dn_{px}/dc_x). Eqn. 5 can be rearranged to give

$$\frac{C_p}{C_x} \equiv \frac{W_p}{W_x} = \frac{\left(\frac{dn_{px}}{dc_x}\right) - \left(\frac{dn}{dc}\right)_x}{\left(\frac{dn}{dc}\right)_p} \tag{6}$$

The volume dimension of the concentration terms, C_p and C_x, cancels to yield the mass ratio (composition) of PEG to protein (W_p/W_x).

Except for very short PEG polymers, the chemical composition per unit length of PEG is constant. Hence $(dn/dc)_p$ for PEG will also be constant over any practical weight range. For most proteins, the principal source of electron density is the amide peptide bond, which is proportional to the molecular weight of the protein. Hence most proteins will also have the same $(dn/dc)_x$ [14]. A protein having an unusual amino acid composition (*i.e.*, an unusually large number of aromatic residues) would have a slightly different $(dn/dc)_x$. However, it is not necessary to rely on literature values of dn/dc for PEG or protein because relative response factors for $(dn/dc)_p$ and $(dn/dc)_x$ can easily be determined experimentally.

Calibration graphs of the appropriate range are constructed both for the PEG alone and for the protein alone, both prior to conjugation. The slopes of these calibration graphs, in refractive index area/weight, will be proportional to the $(dn/dc)_p$ and $(dn/dc)_x$ values in eqn. 6. The (dn_{px}/dc_x) term is determined from the PEG–protein sample itself. The RI detector signal is proportional to the dn_{px} term for the PEG–protein, while the concentration of the protein in the PEG–protein, the dc_x term, is determined from the integrated UV detector area. During the calibration for the $(dn/dc)_x$ for protein alone, the UV response factor at 280 nm (protein UV area/weight) is established. Because the detectors are in series, the protein concentration of the PEG–protein can be determined from the identical injection used to determine the RI response of the PEG–protein sample (eqn. 7).

$$\left(\frac{dn_{px}}{dc_x}\right) = \frac{(\text{PEG–protein RI area})\,K}{\left(\frac{\text{PEG–protein UV area}}{\text{protein UV area/weight}}\right)} \tag{7}$$

where K is a proportionality constant for the RI increment. Therefore, once calibration graphs have been used to establish the RI response of PEG alone

and the UV and RI responses of the protein alone, a single injection of the PEG–protein sample can yield both UV and RI responses to calculate the weight composition of that sample (eqn. 8).

$$\frac{W_p}{W_x} = \frac{\dfrac{\text{PEG–protein RI area}}{\left(\dfrac{\text{PEG–protein UV area}}{\text{protein UV area/weight}}\right)} - \dfrac{\text{protein RI area}}{\text{weight}}}{\dfrac{\text{PEG RI area}}{\text{weight}}} \tag{8}$$

The proportionality constant, K, cancels out in the derivation of eqn. 8. The molar ratio of PEG per protein can be calculated from the weight composition and requires the number-average molecular weight of the PEG:

$$\frac{\text{mol PEG}}{\text{mol protein}} = \frac{(\text{weight composition})(\text{MW protein})}{\bar{M}_n(\text{PEG})} \tag{9}$$

Owing to impurities (*i.e.*, diol and elimination products) in the monomethyl-PEG normally used for conjugation, the $\bar{M}_n$ derived from calibrated SEC and end-group titration may differ from each other.

Previous approaches to the characterization of PEG–proteins usually did not consider the diverse nature of grafted copolymers and, as a result, they often suffered from a lack of accuracy. This paper presents data validating the accuracy of the determination of hydrodynamic size and weight composition for PEG–proteins.

EXPERIMENTAL

Apparatus

The semi-preparative SEC system consisted of an Altex Model 100A pump (Beckman, Fullerton, CA, USA), a WISP 710A autoinjector (Millipore–Waters, Milford, MA, USA), two Zorbax GF-250 columns in series, each 25 cm × 9.4 mm I.D. (DuPont, Wilmington, DE, USA), a Spectraflow 773 UV detector (Kratos, Ramsey, NJ, USA), an ERC-7510 RI detector (Erma, Tokyo, Japan) and a Nelson Analytical 6000-SEC/GPC data acquisition system and software (PE/Nelson, Cupertino, CA, USA). The SEC software calibration option used was a narrow standard peak position calibration method as described by Yau *et al.* [28]. It was employed to calculate $\bar{M}_n$ *and* $\bar{M}_w$ values for experimental PEG samples. Fractions were collected manually.

The analytical SEC system consisted of a Waters Model M6000 pump (Millipore–Waters), a Lo-Pulse LP-21 pulse damper (Scientific Systems, State College, PA, USA), a WISP 710A autoinjector, a Superose 12 column (Pharmacia–LKB, Piscataway, NJ, USA), an Eldex (Menlo Park, CA, USA) 725-101D column heater set at 25°C, a Waters Model 484 UV detector (Millipore–Waters), an ERC-7510 RI detector and a Nelson Analytical 6000-SEC/GPC data acquisition system and software. The eluent for both SEC systems, used at a flow-rate of 0.5 ml/min, consisted of 0.1 *M* sodium sulfate (Fluka, Ronkonkoma, NY, USA) and 0.01 *M* sodium phosphate (pH 7.0) in Type I water from a Technic (Seattle, WA, USA) water system. The hydrolysis equipment consisted of a Thermodyne heating block (Sybron/Thermolyne, Dubuque, IA, USA) and a Savant Speed Vac Concentrator (Savant Instruments, Farmingdale, NY, USA). A Camag TLC Scanner II (Wilmington, NC, USA) set at 590 nm was used to measure the gel bands.

Materials

IL-2 and PEG–IL-2 were obtained from Cetus (Emeryville, CA, USA) [29,30]. Monomethoxypolyethylene glycol used for the preparation of PEG–proteins was obtained from Union Carbide (S. Charleston, WV, USA). The number-average molecular weight of this polymer was *ca.* 6000. Narrow-range PEG and polyethylene oxide (PEO) standards were obtained from American Polymer (Mentor, OH, USA). Thyroglobulin and myoglobin were obtained from Bio-Rad Labs. (Richmond, CA, USA). β-Galactosidase, catalase, aldolase, transferrin, alkaline phosphatase, ribonuclease A, cytochrome *c* and N-(2-hydroxyethyl)piperazine-N′-(3-propanesulfonic acid) (EPPS) buffer were obtained from Sigma (St. Louis, MO, USA). Physical data for these PEG, PEO and protein calibrators are given in Table I. Glutaric anhydride and N-hydroxysuccinimide (NHS) were from Aldrich (Milwaukee, WI, USA). SDS-polyacrylamide gel electrophoresis (PAGE) Daiichi 10 × 10 cm precast, 4–20% gels were obtained from Enprotech (Hyde Park, MA, USA). Ester hydrolysis was performed using 6 *M*

TABLE I

CHARACTERISTICS OF SOLUTES USED FOR SEC CALIBRATION

Solute	MW	$\bar{M}_w/\bar{M}_n$	$[\eta]$ (dl/g)	R_η (Å)	R_s (Å)
PEG standards					
PEO 45000	45 000	1.07	0.610	76	
PEO 21000	21 000	1.12	0.360	49	
PEG 18000	18 000	1.18	0.282	43	
PEG 10750	10 750	1.04	0.202	33	
PEG 10665	10 665	1.05	0.201	32	
PEG 5050	5050	1.04	0.133	22	
PEG 4950	4950	1.10	0.127	22	
PEG 3410	3410	1.08	0.107	18	
PEG 2560	2560	1.08	0.090	15	
PEG 2065	2065	1.05	0.079	14	
PEG 1510	1510	1.06	0.067	12	
Protein standards					
Thyroglobulin	660 000				86
β-Galatosidase	465 000				69
Catalase	220 000				52
Aldolase	158 000				46
Immunoglobulin G	158 000				49[a]
Alkaline phosphatase	86 000				33
Transferrin	81 000				36
Myoglobin	16 900				19
Ribonuclease-A	13 700				18
Cytochrome *c*	13 400				17

[a] Determined by HPSEC as described under Experimental.

hydrochloric acid (J. T. Baker, Phillipsburg, NJ, USA).

Methods

PEG–protein conjugates were prepared using the method of Katre *et al.* [8] by first making an active ester of PEG, and then allowing it to react with protein in EPPS buffer (pH 8.5). Glutaric anhydride was used to make methoxy PEG–glutarate, and the active ester prepared from this was N-hydroxysuccinimide ester [5]. In order to achieve various degrees of protein conjugation, the activated PEG ester was added to the protein solutions at different molar excesses. The resulting solutions were chromatographed on the semi-preparative SEC system, two GF 250 columns in series, to isolate populations of PEG proteins enriched for certain compositions. This separation also excluded from further analysis the unconjugated protein and PEG. The mass of protein injected on the columns varied, but injection volume of the sample was $\leqslant 200$ μl. The polymeric nature of PEG–protein usually precluded the isolation of fractions containing a unique species. Fractions were collected manually. The analytical SEC system was then used to analyze PEG–proteins for weight composition. Although it provided lower resolution than the semi-preparative system, the Superose 12 column was chosen for the analytical system because it gave quantitative recoveries for the unconjugated proteins used in this study.

In order to make a direct measurement of the amount of PEG in each fraction, samples were hydrolyzed by incubation with an equal volume of 6 *M* hydrochloric acid at 85°C for 2 h. This step precipitated the protein fraction. It was followed by vacuum drying, extraction of the free PEG by reconstitution in eluent, centrifugation and removal of the supernatant. Hydrolyzed samples were injected onto the Superose 12 column in parallel with their non-hydrolyzed counterparts. The non-

hydrolyzed sample was used to determine the weight composition of the sample and the protein concentration; the hydrolyzed samples gave the amount of PEG.

Calibration graphs were obtained from analyses of standard solutions (in the range 0.1–1.5 mg/ml) of PEG alone and of protein alone in Type I water. Linear regression analysis of the composite, integrated peak areas in relation to mass for each calibrator yielded a slope (area/mg) which was used to calculate the weight composition using eqn. 8.

SDS-PAGE analyses of the PEG–proteins were loaded with an average protein load of 4.5 μg (15 μl load volume) per well and run at 30 mA constant current per gel. The running buffer was 0.025 *M* Tris–0.19 *M* glycine–0.1% SDS. The staining of the gel was done with a solution of 0.09% Coomassie Blue in 25% ethanol–8% glacial acetic acid in a microwave oven for 3 min. After cooling, the gels were destained in 25% ethanol–8% glacial acetic acid overnight.

RESULTS AND DISCUSSION

SEC calibration for PEG–proteins

As PEG is a homopolymer with a quasi-random structure, it would be expected to have an SEC calibration linear with log MW. Fig. 1a shows that, under the SEC conditions employed, calibration with narrow molecular weight range standards of PEG and PEO did yield a linear calibration for peak apex retention volume *vs.* log MW over a wide range of molecular weights. Protein standards covering a broad molecular weight range, chromatographed on the identical SEC system, gave the peak apex retentions plotted in Fig. 1a. This calibration demonstrates reasonable agreement with the retention volume *vs.* log MW calibration, but the deviations from linearity are greater than those for PEG and PEO. A dramatic difference is seen between the protein and the PEG or PEO calibration graphs in the displacement and dissimilarity of their respective slopes. The explanation for this difference is that the quasi-random coil conformation of aqueous PEG has a much larger effective volume than that of compact native proteins. Thus a commonly used calibration for the SEC of proteins, namely apparent molecular weight, will be unsatisfactory for PEG–proteins. In Fig. 1, a given retention volume indicates a protein molecular weight nearly ten times larger than that of a PEG of similar hydrodynamic volume. The calibration of a PEG–protein will lie between these two extremes and will vary as a function of the amount, length and placement of the

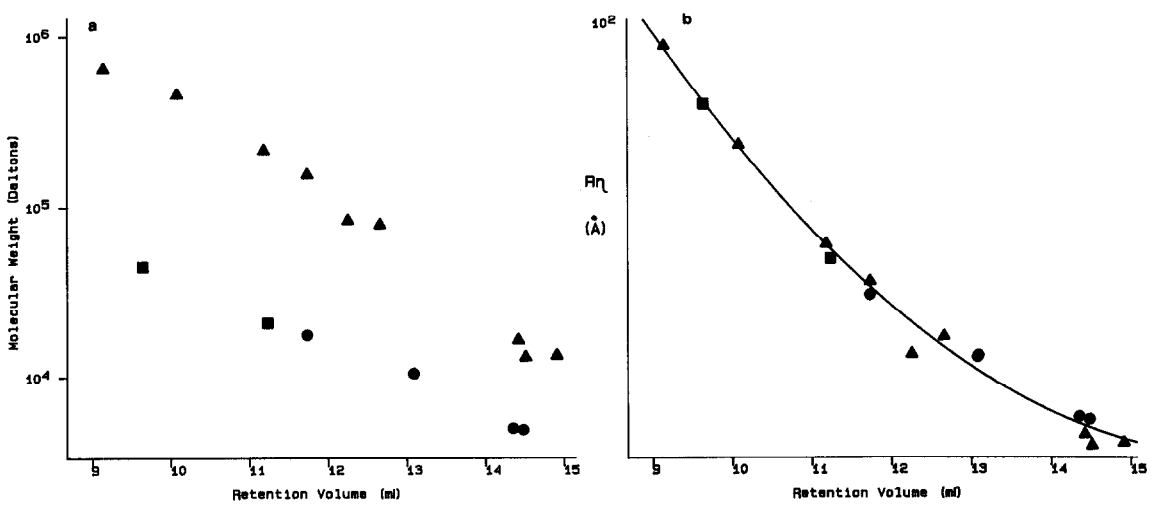

Fig. 1. (a) Peak apex retention volumes of (■) polyethylene oxides (MW 45 000 and 21 000), (●) polyethylene glycols (MW 18 000, 10 750, 10 665, 5050 and 4950) and (▲) proteins [thyroglobulin (660 000), β-galactosidase (465 000), catase (220 000), aldolase (158 000), alkaline phosphatase (86 000), transferrin (81 000), myoglobin (16 900), ribonuclease A (13 700) and cytochrome *c* (13 400)], plotted as a function of log (molecular weight). For conditions of SEC, see Experimental. (b) Chromatographic data identical with those in (a) plotted as a function of R_η.

PEGs conjugated to the protein.

The use of universal calibration is one way to address this calibration difference between PEG and proteins, and to provide a means of measuring the copolymeric species of PEG–proteins. Further, universal calibration of a PEG–protein yields a number or weight-averaged R_η which is a meaningful measure of its pharmacological behavior [7]. For the PEG and PEO calibrators, R_η was calculated from $[\eta]$ supplied by the vendor, using eqn. 10 as described in ref. 20.

$$R_\eta = \frac{30M[\eta]}{\pi N} \tag{10}$$

where R_η is the viscosity radius, M is the molecular weight, $[\eta]$ is the intrinsic viscosity of the molecule and N is Avogadro's number. These data are given in Table I and the results are plotted in Fig. 1b. For water-soluble, globular proteins, no systematic difference has been found between R_s (frictional-based) and R_η (viscosity-based radius) [20]. Thus, as a first approximation, literature values of R_s were used for the protein calibrators in Fig. 1a and substituted for R_η in Fig. 1b [21]. Fig. 1b is the universal calibration obtained for PEG, PEO and proteins. A third-degree polynomial fit shows a well behaved curve ($R^2 = 0.991$), with all of the calibrators sharing a common line. Although universal calibration for PEG–proteins is not rigorously validated in this paper, it is clearly an improvement over the erroneous native protein molecular weight calibration used by previous researchers [7,11]. As universal calibration utilizing both PEG and proteins is well behaved, universal calibration is likely also to be an accurate and pharmacologically relevant measure of PEG–protein size, as the viscosity radius, R_η, effectively compensates for the differences between aqueous PEG and globular protein structures.

The data in Fig. 1b show that proteins diverge from a common calibration more frequently than do PEGs, perhaps owing to residual secondary retention artifacts. It is also possible that the actual R_η values for these calibration proteins may vary from the values used for Fig. 1a, owing to a change in the protein tertiary structure when solubilized in an eluent differing from that used to determine the literature R_s value. Sensitive on-line HPLC viscometers, which have recently become commercially available, would be useful in establishing the R_η of protein and PEG calibrators and of analyte samples with the specific SEC system used. Use of an HPLC viscometric detector would also have allowed us to determine the actual R_η values for various PEG–proteins and thus validate the use of universal calibration for the SEC of PEG–proteins. Polydispersity is a unitless ratio of $\bar{M}_w/\bar{M}_n$ and is immune from issues of calibration accuracy.

Tables I and II list the R_η values for native and PEG conjugated versions of ribonuclease, myoglobin, and IgG. Using universal calibration, SEC analysis shows approximately a 2–3-fold increase in R_η on PEG conjugation of these proteins. The length of the PEG and the degree of conjugation can be controlled to provide PEG–proteins of a desired size range [1–5].

Weight composition

The determination of the weight of PEG per weight of protein as described in the theory section uses two detectors in series to provide sufficient information to solve for the two unknown values, PEG and protein weight concentration. The uncon-

TABLE II

VALIDATION OF WEIGHT COMPOSITION BY HYDROLYSIS FOR VARIOUS PEG–PROTEINS

PEG–protein	Weight composition	PEG expected (mg)	PEG found (mg)	% of expected	R_η (Å)
PEG–ribonuclease	1.69	0.35	0.33	94	58
PEG–myoglobin	0.45	0.17	0.16	90	39
	1.29	0.49	0.44	94	57
PEG–IgG	0.25	0.13	0.12	92	102

jugated PEG and protein used in the construction of the PEG–protein are chromatographed separately over the appropriate concentration ranges as independent calibration graphs. Peak symmetry, quantitative recovery and linearity of the calibration graphs assure that the PEG and the protein are not chemically interacting with the column support and are being retained only through sizing mechanisms. For proteins, this type of calibration graph also establishes the lack of chromatographic artifact from concentration-dependent aggregates. Calibration graphs for the PEG are monitored only with RI detection, as underivatized PEG has no useful UV absorbance at 280 nm. Both UV and RI detectors are used in series to monitor protein calibration graphs. From the calibration graphs, the response factors in RI area/weight is calculated for PEG and for protein. The slopes of these calibration graphs, in RI area/weight, will be proportional to the dn/dc values for the particular SEC set-up. The arbitrary area units are specific for the data system used and also for the column, eluent and detector combination used. Changes in eluent should be monitored for changes in the dn/dc proportionality constant for both PEG and protein. For the four proteins used in this study, the RI calibration graph slopes were similar to one another with a relative standard deviation (R.S.D.) of 6%. This result supports the expectation that protein dn/dc values will be similar. These protein calibration graphs were also highly linear, with R^2 ranging from 0.9984 to 0.9999, and averaging 0.9995 for the UV and RI calibration graphs. Determination of protein concentrations from absorbance at 280 nm will be more accurate than weighing if an accurate molar absorptivity (ε'_{280}) is available, owing to possible errors from salt and hydration.

In this study, experiments were designed to establish that the RI detector was accurately measuring polymer mass. To validate this aspect of the analysis, calibration graphs for five of the PEG and PEO narrow range molecular weight calibrators were chromatographed and the resulting RI areas plotted *versus* weight of polymer injected (Fig. 2). All five calibration graphs are highly linear (R^2 = 0.999), and have virtually identical slopes. This shows that PEGs of any length can be used for RI calibration and that the PEG dn/dc is constant. However, calibrating with the PEG used for PEG–protein conjugation has the advantage of establishing the

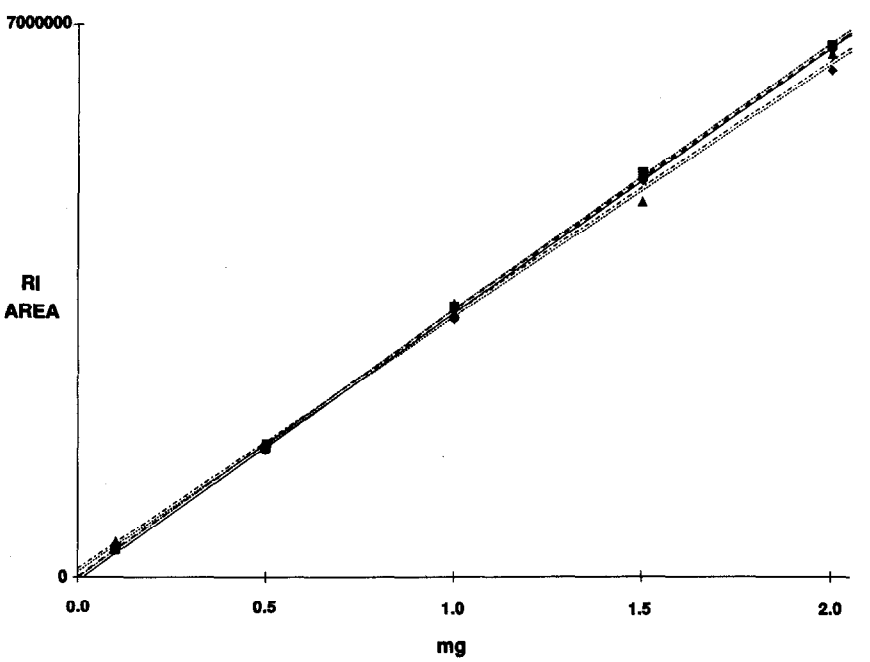

Fig. 2. Calibration graphs of peak area *versus* weight injected for various size PEG and PEO standards (■ = 45 000; ◆ = 21 000) and PEG standards (▲ = 10 750, □ = 3410).

distribution ($\bar{M}_n$,$\bar{M}_w$ and polydispersity) of the PEG actually employed.

The weight composition of a PEG–protein sample is calculated from the three slopes derived from the calibration graphs obtained when protein UV, protein RI and PEG RI responses are measured. Using the identical SEC system, a sample of PEG–protein whose mass is consistent with the PEG and protein calibration graphs is chromatographed, and the resulting area values are used to calculate the weight composition using eqn. 8. Linearity can be checked by constructing calibration graphs of the appropriate range, or it can be assumed, and the slope calculated from a one-point standard calibration passing through zero concentration. The mass of protein injected is calculated on-line with the use of the UV detector area so that a knowledge of the actual protein concentration in the amount injected is not critical.

Another aim of this study was to establish the accuracy of this approach to weight compositional analysis by independently measuring the mass of the PEG and protein moieties. Various PEG–proteins were prepared and semipreparatively fractionated to yield samples conjugated to different extents. For the examples presented here, glutaric acid was used as a linker, providing an easily hydrolyzable ester bond between the PEG and the protein. Acid hydrolysis was used to cleave the conjugated PEG from the corresponding protein. The protein concentration in the PEG–protein conjugate was determined from a calibration graph for the PEG–protein, which was calibrated off-line by UV absorbance of the standards at 280 nm. Literature molar absorptivities were then used to calculate the protein concentrations of the standards. Non-hydrolyzed aliquots of the PEG–protein samples were analyzed by the UV–RI method to establish their weight composition. The expected amount of PEG in a given volume was calculated for each sample. After acid hydrolysis, the free PEG was measured by SEC with RI detection. The PEG recovery was calculated by dividing the amount of PEG found by the amount expected and multiplying by 100. The results for PEG–ribonuclease, PEG–myoglobin and PEG–IgG are given in Table II. The uniformly high recoveries of hydrolyzed PEG for the various PEG–protein samples establish the validity and general utility of this approach.

As the glutaric acid linker is not hydrolyzed off the protein but is measured as PEG–protein mass prior to hydrolysis, one expects the recovery to be 2% low, the ratio of glutaric acid molecular weight to the PEG number-average molecular weight. Errors in the protein molar absorptivities will lead to proportional errors in the weight composition as it will affect the protein calibration graphs. However, when calculating the recovery of hydrolyzed PEG, inaccuracy in molar absorptivity will be inversely compensated for by the calculation of the expected mass of PEG. Thus, errors in molar absorptivities are immaterial to the calculation of PEG recovery. As a check on chromatographic artifacts, acid hydrolysis of a PEG sample was performed in parallel, and no effect on the PEG was seen. However, base hydrolysis of PEG–proteins did yield chromatographic artifacts, perhaps from base–catalyzed protein reactions.

Calculation of the molar composition is a useful PEG–protein characterization and is easily done with eqn. 9. Accurate determination of $\bar{M}_n$ requires some additional analysis. During the process of conjugating PEG to a protein, the PEG undergoes several chemical and purification steps. There are many opportunities to alter the size distribution of the PEG through kinetic or physical chemical differences. Thus, $\bar{M}_n$ should be determined for the acid-hydrolyzed PEG to establish precisely the PEG size distribution on the PEG–protein itself. In this study, the PEG that was hydrolyzed from the IL-2 protein was compared with the starting material PEG and was found to have the same size distribution.

In order to characterize further a PEG–protein population by weight composition, preparations of PEG–IL-2 were sequentially fractionated (preparatively and analytically) by SEC to yield samples which were enriched in species of a given weight composition. Analysis of all the fractions revealed that their weight compositions varied over a broad range (Table III). Nearly quantitative recoveries of hydrolyzed PEG were seen for all of these samples. This establishes that weight composition analysis is valid over any practical PEG to protein ratio, provided that the sample size and peak resolution are sufficient for accurate baseline integration. Because the PEG–protein mixture is essentially a continuous distribution of conformers, baseline

TABLE III

VALIDATION OF WEIGHT COMPOSITION BY HYDROLYSIS FOR PEG–IL-2 FRACTIONS ISOLATED BY SEMI-PREPARATIVE SEC

PEG–IL-2 sample	Weight composition	PEG expected (mg)	PEG found (mg)	% of expected
A	0.53	0.44	0.44	100
B	0.85	0.79	0.77	98
C	1.01	0.22	0.21	94
D	1.16	0.29	0.28	95
E	1.20	0.43	0.42	98
F	1.32	1.64	1.60	97
G	1.36	1.05	0.99	95
H	1.41	1.68	1.70	101
I	1.70	4.59	4.41	96
J	1.70	10.27	9.96	97
K	1.82	2.28	2.22	97
L	2.18	2.59	2.44	94
M	2.35	2.23	2.23	100
N	26.50	15.92	15.76	99

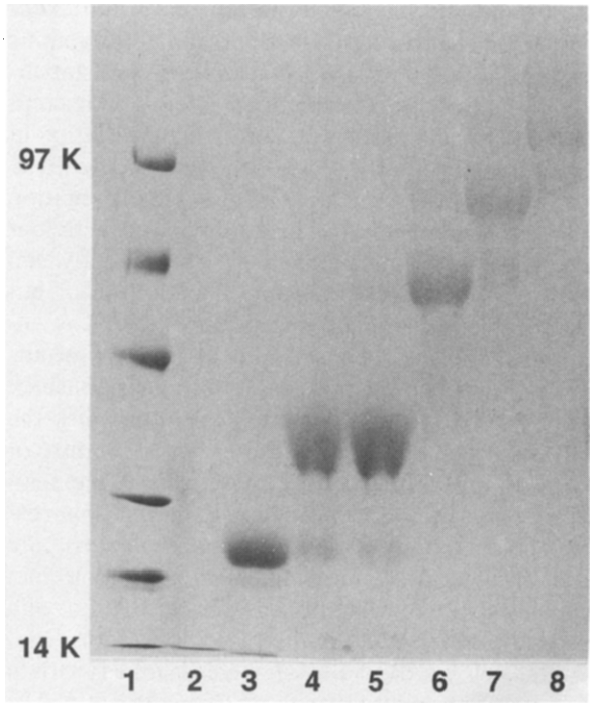

Fig. 3. Coomassie Blue-stained SDS-PAGE gel of PEG–IL-2 samples isolated by semi-preparative SEC. Molar composition (mol PEG/mol IL-2): lane 1 = MW standards; lane 2 = 0 PEG; lane 3 = 1.0 PEG; lane 4 = 2.0 PEG; lane 5 = 3.1 PEG; lane 6 = 4.1 PEG; lane 7 = 4.9 PEG; lane 8 = 6.0 PEG. K = kilodalton.

drops between fused peaks represent arbitrary divisions of related mixtures and may not be equally reflected in both the RI and UV chromatograms. Compositional accuracy of a sample will be assured only if the entire chromatographic peak area is integrated.

Comparison with gels

SDS-PAGE has been a popular method to characterize PEG–protein distributions [7,11]. The assumptions made were that the higher molecular weight gel band ladders represented one molar incremental additions of PEG to the protein, and that the staining density was proportional to protein concentration in each of the bands. In this study, the validity of these assumptions was tested using UV–RI SEC methodology.

Some of the PEG–IL-2 samples used for the hydrolysis validation were also analyzed by SDS-PAGE (Fig. 3). As the weight composition increased, the gel lanes generally increased in apparent hydrodynamic size. However, the 2.0 and 3.1 molar composition PEG–IL-2 samples were not resolved, but co-migrated in a broad band. An investigator without the aid of compositional analysis would incorrectly identify the SDS-PAGE bands of PEG–IL-2 based purely on the one molar ladder assumptions. This anomalous PAGE migration of the PEG–IL-2 samples was also observed in the SEC analysis of similar samples (Fig. 4). Comparison of the molar composition with retention time and observed R_η of these samples shows a similar pattern; the 2.1 and 3.0 molar PEG–IL-2 species were not resolved, whereas other samples followed the expected retention trend.

The lack of resolution between the 2.1 and 3.0 molar composition PEG–IL-2 species is not a general phenomenon, as the other proteins used in this study, ribonuclease, myoglobin and IgG, did not display similar behavior. One explanation is that conjugation of the third PEG–NHS molecule to IL-2 might occur close to either the first or the second PEG molecule in the IL-2 tertiary structure, presenting a minimum increase in effective hydrodynamic size. Another possibility is that the 2-PEG–

IL-2 molecules might have an altered tertiary protein structure, resulting in a larger hydrodynamic size. Although these PEG conjugations are statistical and yield complex mixtures of molecular combinations, there may be enough kinetic variation in lysine reactivity in IL-2 to direct a selective PEG conjugation. The kinetics of PEG–NHS conjugation coupled with the lysine locations and reactivities within the tertiary structure for IL-2 apparently lead to PEG–IL-2 species varying in composition but having similar PAGE migrations and SEC retentions. The exact location and distribution of the PEG–lysine conjugations is not well established. This PAGE and SEC anomaly demonstrates the importance of recognizing the polymeric nature of PEG–IL-2. PEG–protein samples which are relatively pure by SDS-PAGE, having similar apparent molecular weights (hydrodynamic volumes), are actually composed of millions of different molecules which vary in all three dimensions of PEG–protein heterogeneity, including molar composition.

The gel lanes in Fig. 3 were scanned with a densitometer and the results are shown in Fig. 5. The amount of protein applied to each lane was calculated from A_{280} of the in-solution samples, and the gel band response factors were expressed as scan area per milligram of IL-2 protein. As the degree of PEG conjugation increased, the response factor dramatically decreased to approximately one tenth of its original value. It is thought that the PEG molecules shield the IL-2 molecules from dye adsorption. Studies with other PEG–proteins have shown this trend to vary widely from protein to protein. Some proteins, such as myoglobin, do not display decreases in gel response factor with increasing degree of PEG conjugation (data not shown). The PEG–IL-2 SDS-PAGE is therefore an example showing the potential for error in both gel band identification and in estimating PEG distribution.

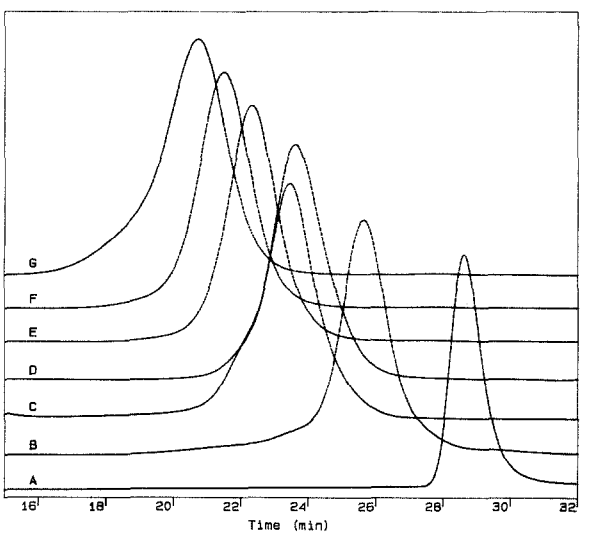

Fig. 4. SE-HPLC traces for PEG–IL-2 samples. Molar compositions and (in parentheses) R_η in Å for trace (A) 0 PEG (18); (B) 1.0 PEG (35); (C) 2.1 PEG (49); (D) 3.0 PEG (44); (E) 4.1 PEG (53); (F) 4.9 PEG (59); (G) 6.0 PEG (67).

CONCLUSIONS

Conjugation of a monomethoxy-PEG polymer to a protein does not create a unique species, but a family of compounds called grafted copolymers which are highly heterogeneous. The normal definition of purity used by protein biochemists is therefore inappropriate for the description of PEG–proteins, as are the methods normally used to characterize them. An SEC analysis based on the premise of polymeric heterogeneity has been developed and validated for PEG–proteins. This method quantitatively describes a PEG–protein mixture in statistical, non-unique terms, thus taking into account its heterogeneous nature. By using UV and RI detectors in series, three characteristics of PEG–proteins can be measured from a single chromatographic run: molecular size (R_η), polydispersity and weight composition. The viscosity-based universal SEC calibration compensates for the difference in

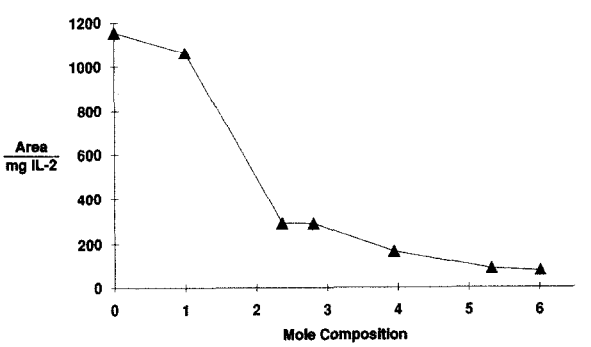

Fig. 5. Densitometric analysis of the SDS-PAGE gel in Fig. 3. Normalized response factors (integrated area per weight protein) as a function of molar composition as determined by UV–RI SEC.

partial specific volume of aqueous PEG and native, globular proteins, allowing meaningful SEC calibration of the PEG–protein hybrid. Both the PEG and PEO standards and the protein standards follow universal calibration for SEC, suggesting that the PEG–protein copolymer will also be accurately calibrated by this method. The width of the polymeric distribution is described by the traditional polymer term, polydispersity. The weight (or molar) composition is obtained through UV and RI detector calibration of the SEC system with the PEG and protein separately, before conjugation. Comparison of the SEC analysis with traditional biochemical analysis by SDS-PAGE shows that the latter can produce errors in band identification and quantification. To prevent erroneous gel interpretation, investigators are advised to calibrate their particular PEG–protein for SDS-PAGE or replace it with the UV–RI SEC compositional analysis.

ACKNOWLEDGEMENT

We thank Heatherbell Fong for her editorial aid in the preparation of the manuscript.

REFERENCES

1 F. F. Davis, A. Abuchowski, T. Van Es, N. C. Palczuk, K. Savoca, R. H.-L. Chen and P. Pyatak, *Biomedical Polymers*, Academic Press, New York, 1980.
2 K. Y. Park, A. Abuchowski, S. Davis and F. Davis, *Anticancer Res.*, 1 (1981) 373.
3 R. H.-L. Chen, A. Abuchowski, T. Van Es, N. C. Palczuk and F. F. Davis, *Biochim. Biophys. Acta*, 660 (1981) 293.
4 C. O. Beauchamp, S. L. Gonias, D. P. Menapace and S. V. Pizzo, *Anal. Biochem.*, 131 (1983) 25.
5 A. Abuchowski, G. M. Kazo, C. R. Verhoest, T. Van Es, D. Kafkewitz, M. L. Nucci, A. T. Viau and F. F. Davis, *Cancer Biochem. Biophys.*, 7 (1984) 175.
6 R. J. Goodson and N. V. Katre, *Biotechnology*, 8 (1990) 343.
7 M. J. Knauf, D. P. Bell, P. Hirtzer, Z.-P. Luo, J. D. Young and N. V. Katre, *J. Biol. Chem.*, 263 (1988) 15064.
8 N. V. Katre, M. J. Knauf and W. J. Laird, *Proc. Natl. Acad. Sci. U.S.A.*, 84 (1987) 1487.
9 N. V. Katre, *J. Immunol.*, 144 (1990) 209.
10 S. J. Stocks, A. J. M. Jones, C. W. Ramsey and D. E. Brooks, *Anal. Biochem.*, 154 (1986) 232.
11 P. McGoff, A. Baziotis and R. Maskiewicz, *Chem. Pharm. Bull.*, 36 (1988) 3079.
12 C.-J. Jackson, J. Charlton, K. Kuzminski, G. Lang and A. Sehon, *Anal. Biochem.*, 165 (1987) 114.
13 R. Cunico, V. Gruhn, L. Kresin, D. Nitecki and J. Wiktorowicz, *J. Chromatogr.*, 555 (1991) 467.
14 P. Kratochvil, *Classical Light Scattering from Polymer Solutions*, Elsevier, Amsterdam, 1987, pp. 209 and 314.
15 M. Kunitani, P. Hirtzer, D. Johnson, R. Halenbeck, A. Boosman and K. Koths, *J. Chromatogr.*, 359 (1986) 391.
16 E. A. Bekturov and Z. K. Bakauova, *Synthetic Water-Soluble Polymers in Solution*, Hüthig & Wepf, Basle, 1981, p. 142.
17 P. Molyneux, *Water-Soluble Synthetic Polymers: Properties and Behavior*, Vol. 1, CRC Press, Boca Raton, FL, 1983, p. 19–45.
18 W. Melander, A. Nahum and Cs. Horváth, *J. Chromatogr.*, 185 (1979) 129.
19 W. O. Richter, B. Jacob and P. Schwandt, *Anal. Biochem.*, 133 (1983) 288.
20 K. Horiike, H. Tojo, T. Yamano and M. Nozaki, *J. Biochem.*, 93 (1983) 99.
21 M. Potschka, *Anal. Biochem.*, 162 (1987) 47.
22 P. Dubin and J. Principi, *Macromolecules*, 22 (1989) 1891.
23 R. Frigon, J. Leypoldt, S. Uyeji and L. Henderson, *Anal. Chem.*, 55 (1983) 1349.
24 M. le Marie, A. Viel and J. Moller, *Anal. Biochem.*, 177 (1989) 50.
25 S. Mori and T. Suzuki, *J. Liq. Chromatogr.*, 4 (1981) 1685.
26 J. R. Runyon, D. E. Barnes, J. F. Rudd and L. H. Tung, *J. Appl. Polym. Sci.*, 13 (1969) 2359.
27 X. Zhonge, Y. Ping, Z. Jingguo, J. Erfang, W. Meiyan and L. J. Fetters, *J. Appl. Polym. Sci.*, 37 (1989) 3195.
28 W. W. Yau, J. S. Kirkland and D. D. Bly, *Modern Size Exclusion Chromatography*, Wiley-Interscience, New York, 1979, pp. 289 and 317.
29 S. Wolfe, G. Dorin, F. Smith and A. Lim, *Eur. Pat. Appl.*, WO88/08849 (1988).
30 K. Koths, J. Thomson, M. Kunitani, K. Wilson and W. Hanisch, *US Pat.*, 4 569 790 (1986).