REVERSED-PHASE CHROMATOGRAPHIC BEHAVIOR OF PROTEINS IN DIFFERENT UNFOLDED STATES[a]

SHIWEN LIN and BARRY L. KARGER*

Barnett Institute of Chemical Analysis and Department of Chemistry, Northeastern University, Boston, MA 02115 (U.S.A.)

SUMMARY

A series of standard small globular proteins in different unfolded states was studied by gradient reversed-phase liquid chromatography. The retention parameters Z [slope of log capacity factor (k') *vs.* log molar concentration of organic modifier, 1-propanol, in the mobile phase] and log I (the value of log k' at 1 M 1-propanol) were derived from gradient retention data. Each protein in four different conformational states, *i.e.*, folded, chromatographic surface-unfolded, urea-unfolded and disulfide-bridge reduced-unfolded, showed a variation of 10-fold in Z and up to 10^{12}-fold in I values. For the different states of all the proteins studied, the order of Z and I values was as follows: folded $\ll$ surface-unfolded $<$ urea-unfolded $<$ reduced-unfolded. The differences in the values of the coefficients suggest, in agreement with literature reports, that proteins with their disulfide bridges cleaved have the largest degree of unfolding. In addition, the Z and I values and solution refolding kinetics all suggest that chromatographic surface-unfolded proteins have a lower degree of unfolding than their urea-unfolded forms. It was also found that an additional chemical cross-link in lysozyme caused a significant decrease in the first-order rate constant of the surface-induced unfolding process.

INTRODUCTION

Currently reversed-phase liquid chromatography (RPLC) plays an important role in the analysis of peptides and proteins. This significance has prompted numerous studies aimed at a better understanding of the retention process of these biopolymers in order to optimize separations[1–4]. Despite some differences between the RPLC behavior of proteins and low-molecular-weight molecules, recent experience indicates that linear solvent strength (LSS) gradient elution theory developed for simple organic molecules can also be used to characterize the retention behavior of proteins[5].

The LSS model is based on the fact that the retention of solutes in RPLC can be

---

[a] Dedicated to Professor Csaba Horváth on the occasion of his 60th birthday. A dear and loyal friend and an outstanding scientist.

approximated by a logarithmic relation between the capacity factor (k') and volume fraction (φ) of organic solvent in the mobile phase:

$$\log k' = \log k'_0 - S\varphi \tag{1}$$

where k'_0 is the extrapolated capacity factor at zero volume fraction of the organic modifier. The slope S, which is dependent on the organic modifier and the specific solute, can be used to predict gradient retention time and band width[6]. For proteins, the plot of $\log k'$ vs. φ is generally very steep yielding a large value of S. In addition, it has been found that the S value is in part related to the size and contact area of the protein molecule.

Based on ion-exchange chromatography, Geng and Regnier[2] have developed a displacement model for retention in RPLC which can be described by the following equation:

$$P_m + ZM_a \rightleftharpoons P_a + ZM_m \tag{2}$$

where P stands for protein, M for organic solvent, Z is the number of organic solvent molecules needed to displace the adsorbed protein from the surface, and subscripts m, a represent mobile and adsorbed states, respectively. Based on this displacement model, retention can be written as

$$\log k' = \log I - Z \log [D_0] \tag{3}$$

where $[D_0]$ is the molar concentration of the organic modifier and I the value of k' at $[D_0] = 1 M$.

Assuming a constant amount of organic solvent adsorbed (imbibed) to the bonded phase and a constant activity coefficient of the organic modifier in the mobile phase as D_0 is varied, Z can be used as a measure of the contact area of the adsorbed protein. In addition, I represents a measure of the relative binding strength of individual proteins under a fixed mobile phase composition. From the relation of molar concentration $[D_0]$ and volume fraction φ, Z can be related to S by the expression[7]

$$Z = 2.3\varphi S \tag{4}$$

Determination of S and k'_0 from the gradient retention data thus permits calculation of Z and I.

The stoichiometric parameters S and Z have been shown to be related to protein molecular weight in RPLC; however, different dependencies have been found as a function of the extent of denaturation under the mobile phase conditions. For example, Stadalius et al.[8] examined values of S vs. molecular weight (MW) from several studies and obtained a relationship of $S \approx (MW)^{0.44}$. On the other hand, Geng and Regnier[2] found that Z was linearly dependent on the protein molecular weight, when a strong denaturing mobile phase [formic acid–isopropanol (60:40)] was used. Aguilar et al.[9] have also reported S values for peptides and found significant scatter in the S vs. molecular weight correlation. The variable dependence of S and Z on the

molecular weight of proteins implies different degrees of unfolding of proteins on the chromatographic surface.

It is well-known that proteins can alter conformation on chromatographic surfaces, resulting in changes in retention, sometimes with broad asymmetrical peaks[10], or multiple peaks[4,11,12]. It is a general rule that water-soluble peptides and proteins adsorbed on hydrophobic surfaces elute later in an unfolded form than in a folded conformation. Moreover, the S and Z values for various states of unfolding of a protein could be significantly different[7,13]. It is now recognized that Z values of proteins in RPLC are determined in large part by the hydrophobic contact area or the number of the interaction sites established between the solute and the stationary phase during the adsorption process. The hydrophobic contact area is, in turn, dependent on the conformation of the protein on the chromatographic surface.

To date, no direct measurement of Z or S values for proteins in the native (or folded) conformation on a reversed-phase column has been reported. Previous work from this laboratory[4,14] has shown that on a short alkyl (C_4) bonded silica gel column with 1-propanol as organic modifier and at low temperature (4°C), a number of proteins yielded two well-separated peaks corresponding to the folded and an unfolded conformation, respectively. The conversion of the folded conformation to an unfolded one allowed the measurement of the first-order unfolding kinetics of several proteins on the C_4 bonded phase surface[14].

This paper examines the behavior of proteins in various conformational states on a C_4 reversed-phase column and correlates Z and log I values to the extent of unfolding of the protein on the chromatographic surface. The values were derived from gradient elution data generated by varying gradient time. Multiple peaks corresponding to different conformational states of proteins were observed, and the parameters of these conformations were determined within the same gradient run. In addition, chemical cross-links were incorporated into the protein molecules, and the chromatographic behavior of these more rigid species was examined. Solution refolding and surface unfolding kinetics were also measured to provide further information on the changes of solute structure within the chromatographic column.

EXPERIMENTAL

Equipment

The chromatographic system consisted of two Altex 110A pumps with an Altex 420 system control programmer (Beckman, San Ramon, CA, U.S.A.), a fast-scan photodiode array detector (Hewlett Packard, Palo Alto, CA, U.S.A.) and a Model 7125 syringe loading sample injector containing a 20-μl loop (Rheodyne, Cotati, CA, U.S.A.). The chromatograms were collected and stored in an HP 9000 workstation (Hewlett Packard) through HP 7996A operating software (Hewlett Packard).

The C_4 reversed-phase packing material was made using standard bonding procedures[14]. Vydac silica (5.6 μm, 300 Å, 64 m^2/g) (Separations Group, Hesperia, CA, U.S.A.) was bonded with n-butyltrimethoxysilane (Petrarch Systems, Bristol, PA, U.S.A.). The C_4 ligand density was determined by elemental analysis to be 7.1 μmol/m^2, assuming a bonding stoichiometry of two methoxy groups. The 10 cm × 4.6 mm I.D. column was slurry packed in 1-propanol–methanol (30:70, v/v) with methanol as driving solvent.

The column temperature was maintained within $\pm 0.5°C$ by immersing the injector, the column, and the tubing connecting the mobile phase mixer to the inlet of the column in a thermostated water bath (Neslab, Newington, NH, U.S.A.). Mobile phase A was either 0.5% 1-propanol in 10 mM H$_3$PO$_4$, pH 2.1, 2% 1-propanol in 10 mM H$_3$PO$_4$ or 2% 1-propanol in 1 mM hydrochloric acid, pH 3.0 and mobile phase B was either 45% 1-propanol in 10 mM H$_3$PO$_4$, pH 2.1 or 45% 1-propanol in 1 mM hydrochloric acid, pH 3.0. The mobile phases were degassed with helium during all the experiments to remove oxygen dissolved in the solvent. Linear gradients from mobile phase A to B in various gradient times with a flow-rate of 1.0 ml/min were used for the measurement of S. For the determination of the refolding kinetics of ribonuclease A, the flow-rate was varied in the same proportion as the gradient time to maintain the gradient volume constant, see ref. 15.

Reagents

All proteins used in this study were obtained from Sigma (St. Louis, MO, U.S.A.) in the highest available grade and used as received. The proteins and their biological origin are as follows: papain (papaya latex, type IV), lysozyme (chicken egg white, grade I), α-chymotrypsinogen A (bovine pancrease, type II), myoglobin (horse skeletal muscle, type I) and ribonuclease A (bovine pancrease, type IIIA). Reagent-grade phosphoric acid, 1-propanol, and other reagents were obtained from J. T. Baker (Phillipsburg, NJ, U.S.A.).

Preparation of protein samples

The urea-denatured protein samples were prepared as follows: 2.5 mg/ml protein solutions in 8 M urea were heated to 90°C for 2–5 min and cooled to 5°C before injection. The disulfide-bond reduced protein samples were prepared using a standard method from the literature[16]. In particular, a 5-mg/ml protein solution in 8 M urea was incubated overnight with 0.3 M mercaptoethanol at 25°C. Before injection, each of the solutions was diluted two-fold with water to obtain the reduced protein sample in a final concentration of 2.5 mg/ml.

Lys(7)–Lys(41) (intramolecular) cross-linked ribonuclease A was prepared using the method of Lin et al.[17]. Ribonuclease A (55 mg) was dissolved in 100 ml of 50 mM Tris–HCl buffer (pH 8.5). A 10-ml volume of 2,4-difluoro-1,3-dinitrobenzene in methanol–water (2:98) was added at a rate of 0.02 ml/min while the solution was stirred in the dark at room temperature. The solution was stirred for 20 h, and the reaction was quenched by adding hydrochloric acid to reach pH 2. The reaction mixture was then desalted and concentrated on an ultracentrifugation cell, Prep-10, (Amicon, Danvers, MA, U.S.A.) to a final concentration of about 2.5 mg/ml.

Glu(35)–Trp(108) cross-linked lysozyme was prepared by following the procedure of Imato et al.[18]. Lysozyme (20 mg/ml) was oxidized with I$_2$ (0.6 mol I$_2$ per mol of protein) for 7 h at room temperature. The reaction mixture was then applied to a 90 cm × 4 cm I.D. CM-Sephadex C-25 column, equilibrated with 0.05 M sodium borate–0.05 M sodium carbonate buffer, pH 10, and eluted with a ten-step gradient over 2 l of solution from 0.02 to 0.1 M sodium chloride in the same buffer. The last eluting component corresponding to the ester bond cross-linked lysozyme between the carboxyl group of Glu(35) and the indole C-2 of Trp(108) was collected and concentrated to roughly 2.5 mg/ml before use.

RESULTS AND DISCUSSION

Derivation of Z and log I from gradient data

Operating under linear-solvent strength (LSS) conditions, the relationship between the instantaneous capacity factor k' and time t in a gradient elution can be expressed as:

$$\log k' = \log k'_0 - b(t/t_0) \tag{5}$$

where b is the gradient steepness parameter and t_0 is the migration time of the unretained species through the column. The gradient retention time (t_g) can be derived from eqn. 5 (see ref. 6) as:

$$t_g = (t_0/b)\log(2.3bk'_0 + 1) + t_0 + t_D \tag{6}$$

in which t_D is the gradient delay time, *i.e.*, the time needed for the mobile phase to travel from the mixer to the inlet of the column. The gradient steepness parameter b is related to important gradient and solute parameters as

$$b = \frac{t_0\Delta\varphi S}{t_G} \tag{7}$$

where $\Delta\varphi$ is the gradient range, *i.e.*, the difference in the volume fraction of the organic modifier from the start to the end of the gradient and t_G is the gradient time. By substituting eqn. 7 into eqn. 6 and assuming $2.3bk'_0 \gg 1$, we can obtain an expression of

$$\frac{t_g - t_0 - t_D}{t_G} = -\frac{1}{S\Delta\varphi} \log t_G + \frac{1}{S\Delta\varphi} \log(2.3S\Delta\varphi t_0 k'_0) \tag{8}$$

Since S is independent of t_G, a plot of $(t_g - t_0 - t_D)/t_G$ *vs.* $\log t_G$ will yield a straight line of slope $1/S\Delta\varphi$. The values of S and k'_0 can then be determined from the slope and intercept of the plot. This method is similar to that based on two gradient runs[19].

The solvent displacement stoichiometric parameter Z can be derived from the corresponding S value by the use of eqn. 4:

$$Z = 2.3\bar{\varphi}S \tag{9}$$

where

$$\bar{\varphi} = \varphi_0 + (t_g - t_0 - t_D - 0.3\frac{t_G}{\Delta\varphi S})\frac{\Delta\varphi}{t_G} \tag{10}$$

and $\bar{\varphi}$ is the value of volume fraction of the organic solvent as the solute band passes the center of the column and φ_0 is the volume fraction at the start of the gradient. For 1-propanol, $[D_0] = 1\ M$ at $\varphi = 0.075$. The binding strength, $\log I$, can then be determined by calculating $\log k'$ at this volume fraction in eqn. 1 from corresponding S and $\log k'_0$ values.

Z and log I for folded and unfolded states

The gradient elution chromatograms for four proteins, papain (PAPN), lysozyme (LYSO), α-chymotrypsinogen A (CHTG), and myoglobin (MYOG) are shown in Fig. 1. As described before[4,14], the first peak for PAPN, LYSO and CHTG are ascribed to the folded state of these proteins, and the second peak to an unfolded state. By lengthening the contact time of each protein with the stationary phase, the area of the second peak increased at the expense of the first peak.

It is interesting to note that the first eluted peak of PAPN, LYSO and CHTG could not be observed with the gradient of Fig. 1 using a C_8 bonded phase column of the ligand density of 4.6 μmol/m^2. Evidently, the hydrophobicity of the C_8 phase was sufficient to unfold the proteins rapidly at 5°C. In addition, even on the C_4 bonded phase column, the first peak for CHTG could only be substantially seen when a less acidic (pH $\geqslant$ 3.0) mobile phase was used. Therefore, in Fig. 1, a mobile phase of pH 3 was used for CHTG to obtain a sufficient amount of the folded conformation. These results emphasize that both the stationary and the mobile phase contribute to the unfolding of proteins on the adsorbent surface.
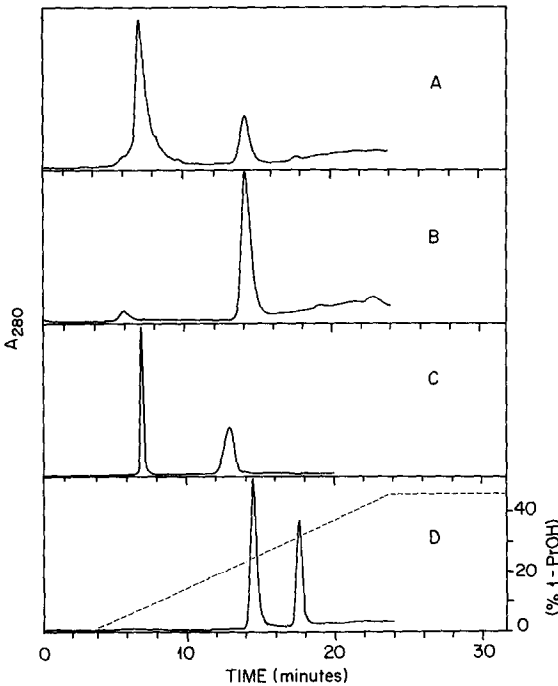

Fig. 1. Reversed-phase chromatographic behavior of (A) papain (PAPN), (B) α-chymotrypsinogen A (CHTG), (C) lysozyme (LYSO) and (D) myoglobin (MYOG). Column: 10 cm × 4.6 mm I.D., C_4 bonded phase on Vydac silica (5.6 μm, 300 Å). Sample: 2.5 mg/ml protein in HPLC-grade water, 20 μl injection. Mobile phase A: 0.5% 1-propanol in 10 mM H$_3$PO$_4$, pH 2.1 (LYSO), 2% 1-propanol in 10 mM H$_3$PO$_4$, pH 2.1 (PAPN, MYOG), 2% 1-propanol in 1 mM hydrochloric acid, pH 3.0 (CHTG). Mobile phase B: 45% 1-propanol in 10 mM H$_3$PO$_4$, pH 2.1 (LYSO, PAPN, MYOG), 45% 1-propanol in 1 mM HCl, pH 3.0 (CHTG). Gradient time: A to B in 20 min. Flow-rate: 1.0 ml/min. Column temperature: 5°C. 1-PrOH = 1-Propanol.

It is also to be noted that the first peak of MYOG was eluted at a time corresponding to the unfolded peaks of other proteins and the second peak was eluted even later. The on-line UV spectrum collected from the photodiode array detector identified the second peak as the prosthetic heme group (λ_{max} = 400 nm) released from myoglobin upon unfolding and the first peak as apomyoglobin (APMY) (λ_{max} = 280 nm). This result is consistent with that found by others[12].

Plots of $(t_g - t_0 - t_D)/t_G$ vs. log t_G for each folded and unfolded state of the above proteins were linear ($r^2 \geqslant 0.98$). The values of S and log k'_0 were derived from the slope and intercept of these plots, with the relative standard deviation of 10% (n = 6). Care was also taken to avoid potential errors in accuracy[19]. From S and log k'_0, the values of Z and log I were determined as described previously. Table I presents these latter values along with S in order to compare with the results of other workers who reported this coefficient[8,9].

Before proceeding, it is useful to note that while the displacement model has been shown to be applicable to RPLC[2], it does require that the adsorbed amount of the organic modifier on the bonded phase be maintained approximately constant over the whole mobile composition range of interest. Since the range over which the different conformational states of the solutes elute is between roughly 5 and 25% (v/v)

TABLE I

Z, S AND LOG I VALUES OF PROTEINS IN DIFFERENT CONFORMATIONAL STATES

For chromatographic conditions, see Fig. 1.

Proteins and states	MW	Z		S	log I
Heme	618				
In water		10	$\pm$ 1[a]	17 $\pm$ 1	2.3 $\pm$ 0.3
In urea		10	$\pm$ 1	17 $\pm$ 1	2.3 $\pm$ 0.3
RNase A	13 000				
Surface-unfolded		11	$\pm$ 1	36 $\pm$ 4	2.5 $\pm$ 0.7
Urea-unfolded		16	$\pm$ 2	38 $\pm$ 4	4.9 $\pm$ 0.4
Disulfide-reduced		21	$\pm$ 3	44 $\pm$ 4	7.0 $\pm$ 0.7
LYSO	14 300				
Folded		2.6	$\pm$ 0.1	25 $\pm$ 2	−0.6 $\pm$ 0.2
Surface-unfolded		20	$\pm$ 1	46 $\pm$ 3	4.9 $\pm$ 0.6
Disulfide-reduced		35	$\pm$ 3	57 $\pm$ 5	11.6 $\pm$ 1.2
APMY	17 000				
Surface-unfolded		22	$\pm$ 2	45 $\pm$ 4	6.5 $\pm$ 1.1
Urea-unfolded		34	$\pm$ 1	52 $\pm$ 5	11.2 $\pm$ 1.2
PAPN	21 000				
Folded		4.3	$\pm$ 0.1	26 $\pm$ 2	−0.3 $\pm$ 0.3
Surface-unfolded		23	$\pm$ 1	43 $\pm$ 2	6.1 $\pm$ 0.5
Urea-unfolded		33	$\pm$ 3	53 $\pm$ 5	10.7 $\pm$ 1.1
CHTG	25 000				
Folded		3.9	$\pm$ 0.3	34 $\pm$ 2	−1.3 $\pm$ 0.1
Surface-unfolded		27	$\pm$ 2	53 $\pm$ 5	7.9 $\pm$ 1.1
Urea-unfolded		38	$\pm$ 3	61 $\pm$ 5	9.6 $\pm$ 0.1

[a] $\pm$ Standard deviation (n = 6).

propanol, some change in the amount of adsorbed or imbibed propanol may be expected. Consequently, the Z values for different conformational states should be viewed in a qualitative fashion, rather than as quantitative differences in contact area.

As can be seen from Table I, the prosthetic heme group has a much smaller Z value ($Z = 10$) than the unfolded parent protein APMY ($Z = 22$). This result is expected when the differences in the sizes of the two species are considered. On the other hand, compared to small peptides in the same molecular weight range (MW ≈ 600)[8], the Z (or S) value for heme is still large. This is undoubtedly due to the relatively large contact area of the porphyrin moiety with the hydrophobic surface.

It is interesting to note that the S values for the unfolded conformation of PAPN, LYSO and CHTG are above 40 and are close to the values reported by others[8] for these particular proteins under more unfolding-favored conditions (C_8-bonded silica as stationary phase, and/or acetonitrile as organic solvent at room temperature). By using less destablizing conditions ($5°C$, C_4-bonded stationary phase, 1-propanol), we were able to maintain both unfolded and folded conformations of these proteins in the chromatographic process.

Table I shows that the Z and log I values for each of the folded proteins are much smaller than those for the corresponding unfolded species. Low Z values have also been found in hydrophobic interaction chromatography (HIC) of proteins under mild conditions[13]. The results demonstrate that proteins adsorbed in a folded conformation have less contact area with the hydrophobic surface than in the unfolded state. In the folded form, the proteins appear to act as weakly binding small molecules. When the protein unfolds on the surface, the contact area increases, leading to stronger binding and an increase in the values of each of the parameters.

It is also interesting to note in Table I that the I values for the folded proteins are only in the range of 0.05–0.5 and are 10^{11}–10^{12} times smaller than the I values for unfolded proteins. This substantial difference in binding strength for the two conformers emphasizes the important point that the chromatographic surface can act as a means of amplifying structural differences in macromolecules. Finally, it is seen that the I values for the folded proteins are even smaller than that of heme, a small molecule. The reason for this result, in spite of the larger S values for the folded proteins, is probably due to the quite hydrophobic porphyrin moiety of heme.

Urea-unfolded and reduced-unfolded states

In order to investigate further the relationship of Z and log I to the conformation of the protein on the chromatographic surface, we next measured the appropriate values for urea-unfolded and disulfide-reduced proteins. Fig. 2 presents the chromatograms for the urea-unfolded proteins (PAPN, CHTG and MYOG) obtained under the same column conditions as in Fig. 1. The Z and log I values are listed in Table I. (For urea-treated LYSO, a complex elution pattern was observed and measurements were not possible.) Note in Fig. 2 that the first peak corresponding to the folded form does not appear in the elution profile of the urea-treated PAPN and CHTG, and the second peak corresponding to the unfolded form is shifted to longer retention. Note also that the urea-unfolded APMY appeared as a broader peak with larger retention than the surface-unfolded APMY.

As can be seen from Table I, the Z values were larger for each of the urea-unfolded proteins (PAPN, CHTG and APMY) than for their surface-unfolded
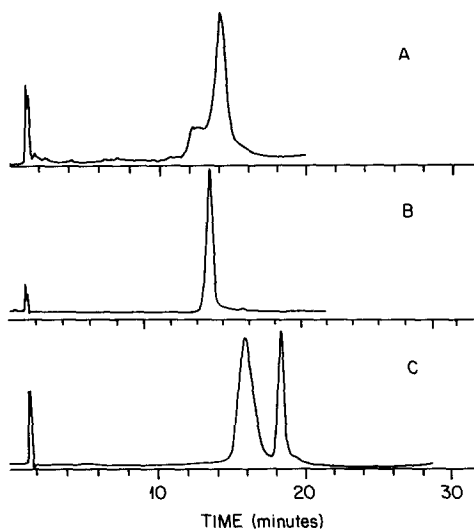
Fig. 2. Chromatographic behavior of urea-unfolded proteins. (A) PAPN; (B) CHTG; (C) MYOG. Conditions are same as in Fig. 1.

form. On the other hand, as expected, the values for heme did not change with the urea treatment. The above results suggest that the urea-treated proteins are more unfolded and adsorb on the chromatographic surface with a larger contact area than the surface-induced unfolded state of these proteins. It is important from the separation point of view to recognize that urea can alter the conformation and retention of proteins, although at times complex peak shape can be observed[20].

The next step was to investigate the influence of disulfide bonds on the adsorption behavior of these globular proteins on the C_4 reversed-phase support. Upon reduction of the disulfide bridge(s), proteins can be completely unfolded to a random coil structure. For LYSO, the reduced species was eluted later in the gradient than the surface-unfolded species and with larger Z and log I values, see Fig. 3. The
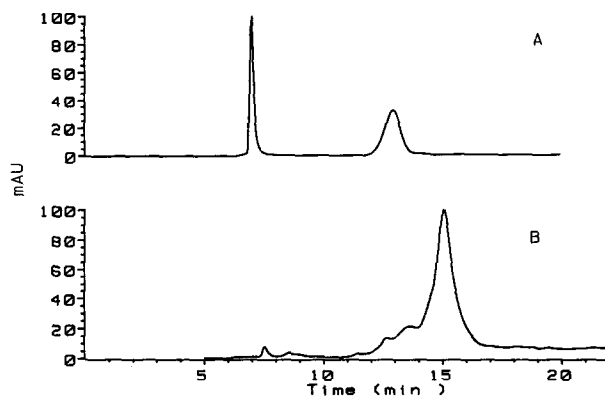

Fig. 3. Chromatographic behavior of (A) LYSO and (B) reduced LYSO. Conditions are same as in Fig. 1C.

longer retention of the reduced LYSO is in agreement with other published results[21], in which it is suggested that the stronger interaction is due to the more exposed hydrophobic residues on the protein.

Ribonuclease A (RNase A) is another well-defined globular protein. Injection of urea-denatured and disulfide reduced RNase A yielded similar peak shapes to that of the unmodified RNase A shown previously[15], although the retention for the former two species was longer (see Fig. 4). In agreement with earlier results[10], the on-line absorbance ratio determined at 288 and 254 nm (A_{288}/A_{254}) for the folded conformation of RNase A was close to unity whereas the ratio was less than one for an unfolded state. The A_{288}/A_{254} for each of the late eluted peak maxima of Fig. 4A–C, was found to be 0.55, 0.50 and 0.69, respectively. On the other hand, the A_{288}/A_{254} ratio measured at the maxima of each of the early eluting shoulders was 0.91, 0.88 and 0.71, respectively. These absorbance ratios suggest that the late eluted peak of each of the RNase A samples corresponds to an unfolded conformation[10]. The shoulder in the case of the urea-unfolded RNase A (see Fig. 4B), in analogy to the surface unfolded results, is assumed to be the refolded form.

It is known that unfolded RNase A can gradually refold in solution once the denaturant is removed or diluted[22]. In the present case, the denaturant urea was washed away from the adsorbed RNase A at the beginning of the chromatographic process and the conformational stress of the hydrophobic surface was released upon desorption. Therefore, favorable conditions existed for RNase A to refold in the mobile phase upon desorption. However, the shoulder in the case of reduced RNase A (see Fig. 4C) had an A_{288}/A_{254} ratio similar to that of the major peak. While not proven, it is possible that this shoulder may represent an intermediate in the refolding pathway of the reduced RNase A. More work is required to identify this band in Fig. 4C.
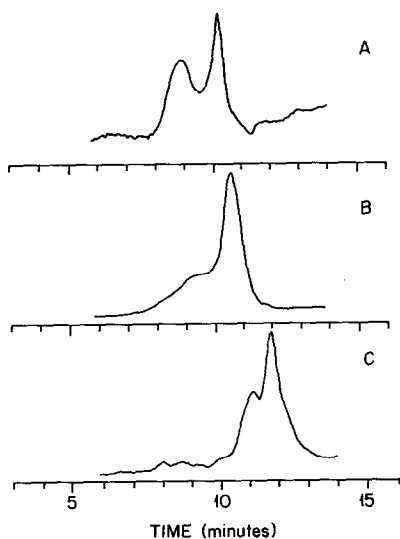

Fig. 4. Chromatographic behavior of (A) RNase A; (B) urea-unfolded RNase A; (C) reduced RNase A. Conditions are same as in Fig. 1C.

The results of Z and log I reveal a similar pattern for RNase A as that of the other proteins studied. The urea-unfolded form had larger Z and log I values than the surface-unfolded form, supporting the argument that urea unfolds the protein to a larger extent with a resultant larger contact area and stronger interaction with the stationary surface. The even larger Z and log I values of the reduced-unfolded RNase A indicate that the reduced species is the most unfolded with the largest contact area. From the results with all the above studied proteins, the order of binding strength for different states can be summarized as: folded $\ll$ surface-unfolded $<$ urea-unfolded $<$ reduced-unfolded.

Effect of additional cross-link on surface unfolding

In contrast to the reduction of the internal disulfide bridges, modification of proteins with additional cross-links would be expected to yield a more rigid structure with restricted conformation. Two well-characterized cross-linked proteins —dinitrophenyl cross-linked [Lys(7)–Lys(41)]RNase A and the ester bond cross-linked [Glu(35)–Trp(108)]LYSO— were used in this study. X-Ray crystallographic studies have shown that the dinitrophenyl cross-link in the 7–41 positions of RNase A is blocked in the enzymatic pocket[23]. Furthermore, the ester bond cross-link between the carboxyl of Glu(35) and the indole C-2 of Trp(108) in the 35–108 positions of LYSO is also buried in the cleft of this enzyme[18]. The three-dimensional structures of these cross-linked proteins are almost identical to their corresponding intact folded structures except for a very small movement of the two cross-linked residues.

The Z and log I values for the cross-linked and unmodified LYSO and RNase A are listed in Table II. For LYSO, there is essentially no difference in the values for the two folded species. Evidently, the cross-link in the active site cleft does not influence the surface position of contact which is believed to be opposite the catalytic site[24,25]. However, in the unfolded state there is a small but meaningful difference in the two forms, with the species with the extra cross-link displaying a smaller Z and I value, as might be expected[21]. In agreement with these results, Perry and Witzel[26] have recently found that T_4-lysozyme engineered with an additional disulfide bridge eluted earlier than its non-cross linked mutant on a C_8 reversed-phase column.

For RNase A, the unfolded forms of the two species reveal identical Z and I values. Only when the reduced forms are compared is a significant difference observed between the two species. Four disulfide bridges in RNase A were destroyed upon reduction, while the artificial Lys(7)–Lys(41) cross-link remained in the reduced (7–41) RNase A. This cross-link restricted the unfolding and reduced the contact area of (7–41)RNase A with the stationary phase, resulting in a smaller value of Z and I.

The results of Table II suggest that in certain cases it may not be possible to observe differences in folded forms of protein variants where disulfide scrambling takes place. Sometimes examining the surface-unfolded forms may be helpful (LYSO), but this is not always true (RNase A). Another sensitive approach is to determine the rate of surface unfolding or solution refolding upon desorption. Kinetic parameters can prove to be highly useful in characterizing variants as shown in the following sections.

Surface unfolding kinetics of cross-linked LYSO

In order to study the effect of a cross-link in protein molecules on the

TABLE II

Z, S AND I VALUES vs. NUMBER OF CROSS-LINKS OF PROTEINS

For chromatographic conditions, see Fig. 1.

Protein	Cross-links	Z	S	log I
Folded (35–108)LYSO	5	2.5 ± 0.1	22 ± 2	-0.3 ± 0.1
Folded LYSO	4	2.6 ± 0.1	25 ± 2	-0.6 ± 0.2
Unfolded (35–108)LYSO	5	16 ± 1	40 ± 4	4.6 ± 0.7
Unfolded LYSO	4	20 ± 1	45 ± 4	4.9 ± 0.6
(7–41)RNase A	5	10 ± 1	34 ± 2	2.6 ± 0.7
RNase A	4	11 ± 1	36 ± 2	2.5 ± 0.7
Reduced (7–41)RNase A	1	16 ± 1	39 ± 1	3.0 ± 1.1
Reduced RNase A	0	21 ± 3	44 ± 2	4.5 ± 0.7

surface-induced unfolding kinetics, we measured chromatographically the first-order unfolding rate constant of the cross-linked (35–108)LYSO as well as that of LYSO. As previously[14], the area of the first peak, corresponding to the concentration of the folded conformation, decreased with contact time of the protein with the chromatographic surface. The area of the second peak correspondingly increased. By following the reduction in the area of the first peak with on-column incubation time, the first-order unfolding rate constant k_u could be determined[14]. At 20°C the first order unfolding rate constant was found to be $9.1 \cdot 10^{-4}$ s^{-1} for the cross-linked (35–108)LYSO and $6.7 \cdot 10^{-3}$ s^{-1} for unmodified LYSO. The significantly slower unfolding rate constant for the cross-linked LYSO variant corresponds to a higher activation energy of 1.2 kcal/mol. A solution study of (35–108)LYSO in 1-propanol–water (4.6 M or 35%) also showed that the cross-linked LYSO was more stable towards thermal unfolding[23]. Since the Glu(35)–Trp(108) cross-link spans the catalytic cleft of LYSO[18,27], it is reasonable to conclude that the opening of this cleft is part of the unfolding process on the surface.

The assumption that the Glu(35)–Trp(108) region participates in the unfolding process is also in agreement with the results of chromatographic and surface spectroscopic studies. Using a series of lysozymes obtained from different biological species, it was shown, as already noted, that the protein adsorbed on a hydrophobic patch that is on the opposite side of the catalytic cleft[24]. The same conclusion was arrived at with computer modeling of the hydrophobic sites on the surface of LYSO[25], and the on-column intrinsic fluorescence of LYSO on a C$_4$ reversed-phase surface also suggested the opening of the catalytic cleft[28]. The cross-linking of Glu(35) and Trp(108) restricted the opening of the cleft and therefore increased the activation energy of the unfolding process.

Based on these results, the kinetics of surface unfolding may be a useful tool for evaluating protein conformation resulting from cross-linking. In particular, it may be possible to differentiate in a sensitive manner species that have formed incorrect disulfide bridges. In combination with retention patterns[7], the surface unfolding kinetics as determined by chromatography or by on-line intrinsic fluorescence of adsorbed species may represent a powerful analytical approach.

Refolding kinetics of RNase A upon desorption

As noted in Fig. 4, the shoulder for urea-denatured RNase A was lower than that of surface-unfolded RNase A. Since on-line spectroscopy revealed that in both cases the shoulder was due to refolding in the mobile phase upon desorption, the first-order rate constants in each case were measured as described previously[15]. At 20°C, the refolding rate constant k_r for the surface-unfolded RNase A and urea-unfolded RNase A were $1.04 \cdot 10^{-2}$ s^{-1} and $4.33 \cdot 10^{-3}$ s^{-1}, respectively, corresponding to a difference of 0.51 kcal/mol in the activation energy of the refolding process. The slower refolding of urea-unfolded RNase A may be a result of the more unfolded structure of this denatured protein on the chromatographic surface.

It might be argued that since urea-unfolded RNase A desorbed at a slightly higher concentration of 1-propanol in the gradient elution, the slower refolding could simply be the effect of the propanol. However, from the previous calibration of k_r *vs.* percentage propanol[15], such a small difference in the concentration of 1-propanol at desorption (14.8% for urea-unfolded RNase A and 14.0% for the surface-unfolded RNase A) could not cause a 2-fold change in the refolding rate constant. It can be concluded that the solution kinetics upon desorption may be a function of the state of the molecule on the chromatographic surface.

CONCLUSIONS

It has been demonstrated that a protein can exist on a chromatographic surface with different extents of unfolding. These states lead to changes in Z and log I and hence retention patterns. Particularly striking were the large differences in I between the folded and unfolded state, indicating the amplifying nature of the support surface for structure variations. This amplification could also be observed in the kinetics of unfolding on the surface. Indeed, the measurement of the rate of unfolding of adsorbed species either by chromatography or intrinsic fluorescence in which precisions of rate constants of better than 10% relative standard deviation are possible[29] may represent in appropriate cases a significant method for characterization of protein variants. Finally, it needs to be recognized that sample pretreatment can have a significant effect on the ultimate chromatographic separation. This point should be kept in mind as separation procedures are being developed.

ACKNOWLEDGEMENTS

REFERENCES

1 L. R. Snyder and M. A. Stadalius, in Cs. Horváth (Editor), *High Performance Liquid Chromatography —Advances and Perspectives*, Vol. 4, Academic Press, New York, 1986, p. 195.
2 X. Geng and F. E. Regnier, *J. Chromatogr.*, 296 (1984) 15.
3 M. T. W. Hearn and M. I. Aguilar, *J. Chromatogr.*, 397 (1987) 47.

4 S. A. Cohen, K. P. Benedek, S. Dong, Y. Tapuhi and B. L. Karger, *Anal. Chem.*, 56 (1984) 217.
5 M. A. Stadalius, M. A. Quarry, T. H. Mourey and L. R. Snyder, *J. Chromatogr.*, 358 (1986) 17.
6 J. L. Glajch, M. A. Quarry, J. F. Vasta and L. R. Snyder, *Anal. Chem.*, 58 (1986) 280.
7 M. Kunitani, D. Johnson and L. R. Snyder, *J. Chromatogr.*, 371 (1986) 313.
8 M. A. Stadalius, H. S. Gold and L. R. Snyder, *J. Chromatogr.*, 296 (1984) 31.
9 M. I. Aguilar, A. N. Hodder and M. T. W. Hearn, *J. Chromatogr.*, 327 (1985) 115.
10 S. A. Cohen, K. Benedek, Y. Tapuhi, J. C. Ford and B. L. Karger, *Anal. Biochem.*, 144 (1985) 275.
11 M. T. W. Hearn, A. N. Hodder and M. I. Aguilar, *J. Chromatogr.*, 327 (1985) 47.
12 R. H. Ingraham, S. Y. M. Lau, A. K. Taneja and R. S. Hodges, *J. Chromatogr.*, 327 (1985) 77.
13 S-L. Wu, A. Figueroa and B. L. Karger, *J. Chromatogr.*, 371 (1986) 3.
14 K. Benedek, S. Dong and B. L. Karger, *J. Chromatogr.*, 317 (1984) 227.
15 X. M. Lu, K. Benedek and B. L. Karger, *J. Chromatogr.*, 159 (1986) 19.
16 F. H. White, *Methods Enzymol.*, 11 (1967) 481.
17 S. H. Lin, Y. Konishi, M. E. Denton and H. A. Sheraga, *Biochemistry*, 23 (1984) 5504.
18 T. Imato, F. Hartdegen and J. A. Rupley, *J. Mol. Biol.*, 80 (1973) 637.
19 M. A. Quarry, R. L. Grob and L. R. Snyder, *Anal. Chem.*, 58 (1986) 907.
20 K. D. Nugent, W. G. Burton and L. R. Snyder, *J. Chromatogr.*, 443 (1988) 381.
21 A. S. Acharya, *J. Protein Chem.*, 5 (1986) 199.
22 L. N. Lin and J. F. Brandts, *Biochemistry*, 22 (1983) 553.
23 S. Segawa and M. Sugihara, *Biopolymers*, 23 (1984) 2489.
24 J. L. Fausnaugh and F. E. Regnier, *J. Chromatogr.*, 359 (1986) 131.
25 D. Horsley, J. Herron, V. Hlady and J. D. Andrade, in J. L. Brash and T. A. Horbett (Editors), *Proteins at Interfaces —Physicochemical and Biochemical Studies*, American Chemical Society, Washington, DC, 1987, p. 290.
26 J. L. Perry and R. Witzel, *Science (Washington, D.C.)*, 226 (1984) 555.
27 C. R. Beddell, C. C. F. Blake and S. J. Oatley, *J. Mol. Biol.*, 97 (1975) 643.
28 X. M. Lu, A. Figueroa and B. L. Karger, *J. Am. Chem. Soc.*, 110 (1988) 1978.
29 P. Oroszlan, R. Blanco, X-M. Lu, D. Yarmush and B. L. Karger, *J. Chromatogr.*, 500 (1990) 481.