ELSEVIER

Journal of Chromatography A, 766 (1997) 1–14

JOURNAL OF CHROMATOGRAPHY A

Protein retention in hydrophobic interaction chromatography: modeling variation with buffer ionic strength and column hydrophobicity

Tracy W. Perkins, Derek S. Mak, Thatcher W. Root, Edwin N. Lightfoot*

Department of Chemical Engineering, University of Wisconsin, Madison, WI 53706 USA

Received 13 August 1996; revised 11 November 1996; accepted 13 November 1996

Abstract

The variation in protein retention times with protein and surface hydrophobicity and mobile phase composition and concentration has been described with a simple thermodynamic model. Column capacity factors for two proteins have been measured as a function of mobile phase ionic strength for a series of columns with varying levels of hydrophobicity. Application of the model to these data suggests that the protein retention is dominated by the release of water molecules upon adsorption, which is consistent with the entropically driven nature of hydrophobic interactions. The calculated number of water molecules released agrees with estimates based on the reduction in hydrophobic surface area for adsorption.

Keywords: Retention models; Column hydrophobicity; Mobile phase composition; Preferential interaction model; Proteins

1. Introduction

Hydrophobic interaction chromatography is widely used for protein separations because it sorts proteins by differences in the hydrophobic character of their surfaces under mild conditions that reduce losses due to protein denaturation [1–3]. The selectivity of this technique changes with the hydrophobicity of the column and with the composition and concentration of the mobile phase. A better understanding of the interaction of these factors is essential for developing models that can be used to improve the design and operation of adsorptive separation processes.

In hydrophobic interaction chromatography, adsorption is driven by the increase in entropy that accompanies the reduction in the total hydrophobic surface area [4]. As the protein mixture flows through a column, hydrophobic patches on the surfaces of the proteins contact the hydrophobic ligands on the beads. The total wetted area decreases, and water is released. Water molecules in contact with hydrophobic areas, unlike those near charged or polar surfaces or in solution with other water molecules, do not experience a highly favorable enthalpy of interaction. Thus the liberation of water from hydrophobic regions increases the entropy of the system without significantly changing its enthalpy, causing the free energy to decrease on adsorption.

Protein adsorption, and the resulting selectivity of the chromatographic separation process, is adjusted by changing the salt concentration. At high salt concentrations, adsorption is dominated by hydrophobic interactions, and lowering the salt concentrations causes the protein retention times to de-

*Corresponding author.

PII S0021-9673(96)00978-8

crease. As the salt concentration is lowered further, incomplete shielding of charges on the protein and surface can lead to adsorption by ion-exchange mechanisms. At this point, additional reductions in the salt concentration will cause the retention time to increase. The variation in retention times with salt concentration and composition is of considerable interest in both hydrophobic interaction and ion exchange regimes [1,5–7].

Two types of explanations for the effect of salts on protein adsorption have been proposed. One approach, based on *solvophobic theory*, examines the effect of the salt on the properties of the solvent. The interaction of the protein and the surface is considered as a two-step process [8,9]. First a cavity for the protein is formed in the water above the surface site, and then the protein fills the cavity and adsorbs on the surface. The free energy change for this sequence is broken into contributions from cavity formation, solute–solvent interactions (including van der Waals interactions and electrostatic interactions), protein–solvent interactions, and complex–solvent interactions.

This theory captures the competing effects of electrostatic and hydrophobic interactions, allowing it to describe adsorption over the full range of salt concentrations. Under high salt conditions of interest in hydrophobic interaction chromatography, this theory indicates that the variation in the log of the retention time with salt concentration should be proportional to the molal surface tension increment of the salt. Thus salts that increase the surface tension of the solvent will favor the reduction in surface area that accompanies adsorption, resulting in increased retention times. While this result is true for many salts, the correlation between retention times and the molal surface tension increments does not hold for salts that interact strongly with the protein [5,10–12].

An alternative to the solvophobic theory is provided by *preferential interaction analysis*, which focuses on the interactions of solutes with proteins [13–18]. A model-independent thermodynamic analysis is used to relate the effect of salt on the observed equilibrium constant to the change in the distribution of salt ions and water molecules. This analysis predicts that processes that reduce wetted surface area are favored by solutes that are preferentially excluded from the surfaces. Conversely, these processes are impeded by solutes that are accumulated near the surfaces. Like the solvophobic theory, this approach is valid over the full range of solute concentrations. In addition, it allows the effect of solutes on the observed equilibrium constant to be interpreted in terms of a stoichiometric displacement model. As a consequence, the parameters in this model have a more direct physical meaning than the lumped factors used with solvophobic theory.

Preferential interaction analysis has been used to explain the effect of solutes on a variety of assembly processes, including protein stabilization, precipitation, aggregation, ligand binding and adsorption [5,7,14,18–24]. This approach has not been widely applied to adsorptive chromatography. Previous studies, however, have shown that the preferential interaction analysis can be used to explain the thermodynamics of binding on chromatographic media [5,14], retention behavior of proteins under ion-exchange conditions [7], and conformational behavior of proteins on hydrophobic surfaces [19].

The preferential interaction analysis and its extension to adsorptive chromatography is discussed briefly below. Experimental data for several protein–sorbent systems are then evaluated using the preferential interaction model, and the parameters obtained are used to provide a realistic interpretation of the observed retention behavior.

2. Theory

The effect that a protein has on the distribution of solvent and solute molecules is characterized by a measurable quantity known as the preferential interaction coefficient:

$$\Gamma_{32}^{m} \equiv \left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} \tag{1}$$

where m is molal concentration. The subscripts 1, 2 and 3 refer to the solvent, the protein and the solute, respectively. By applying the two-domain model of Timasheff [13,16–18,25], which divides the solution into a bulk region that is unaffected by the presence of the protein and a local region to which the preferential accumulation or exclusion of solutes is

attributed, the preferential interaction coefficient can be expressed by

$$\Gamma^m_{32} = \nu_3 - \frac{m_3}{m_1} \cdot \nu_1 \quad (2)$$

where ν_i is the moles of species i in the vicinity of the protein per mole of protein. This preferential interaction coefficient is zero when the bulk solution and local region have the same composition. Otherwise it is a measure of the protein-induced difference between local and bulk solute concentrations. An analogous development for electrolytes [13] gives:

$$\Gamma^m_{+,2} + \Gamma^m_{-,2} = (\nu_+ + \nu_-) - \frac{n \cdot m_3}{m_1}\nu_1 \quad (3)$$

where n is the total number of ions associated with the electrolyte, and + and − refer to the cations and anions [13–15]. For a protein in an aqueous solution of ammonium sulfate, m_3 is the molal concentration of ammonium sulfate, m_1 is the molal concentration of water (55.51 m), and n is 3.

2.1. Application to protein sorption

The two-domain model provides a convenient interpretation of the effect of solutes on processes that can be expressed in terms of the preferential interaction coefficients. For the adsorption process described by

$$p\mathrm{P} + s\mathrm{S} \rightleftarrows c\mathrm{C} \quad (4)$$

where P is the protein, S is the surface site and C is the protein–surface complex, the log of the observed equilibrium constant is

$$\ln K^m_{\mathrm{obs}} = c \ln(m_{\mathrm{C}}) - p \ln(m_{\mathrm{P}}) - s \ln(m_{\mathrm{S}}) \quad (5)$$

Through careful manipulation of partial derivatives, it has been shown [13–18] that the variation of the observed equilibrium constant with the concentration of a nonelectrolyte (or with the mean ionic activity of an electrolyte) is given by the stoichiometrically weighted sum of the preferential interaction coefficients:

$$\text{nonelectrolyte: } SK_{\mathrm{OBS}} \equiv \left[\frac{\partial \ln(K_{\mathrm{OBS}})}{\partial \ln(m_3)}\right]_{T,P,EQ} = c\Gamma^m_{3\mathrm{C}} - p\Gamma^m_{3\mathrm{P}} - s\Gamma^m_{3\mathrm{s}} \quad (6a)$$

$$\text{electrolyte: } SK_{\mathrm{OBS}} \equiv \left(\frac{\partial \ln(K_{\mathrm{OBS}})}{\partial \ln(a_\pm)}\right)_{T,P,EQ} = c(\Gamma^m_{+,\mathrm{C}} + \Gamma^m_{-,\mathrm{C}}) - p(\Gamma^m_{+,\mathrm{P}} + \Gamma^m_{-,\mathrm{P}}) - s(\Gamma^m_{+,\mathrm{S}} + \Gamma^m_{-,\mathrm{S}}) \quad (6b)$$

For electrolytes, $a_\pm$ is the mean ionic activity, which is related to the molal salt concentration by the following equation:

$$a_\pm = (n_+^{n_+} n_-^{n_-})^{1/n} \gamma_\pm m_3 \quad (7)$$

where $\gamma_\pm$ is the mean ionic activity coefficient, n_+ and n_- are the number of cations and anions per unit of salt, and n is the total number of anions and cations per formula unit. Reexpressing these relationships in terms of the two-domain model yields:

$$\text{nonelectrolyte: } SK_{OBS} \equiv \left[\frac{\partial \ln(K_{\mathrm{OBS}})}{\partial \ln(m_3)}\right]_{T,P,EQ} = \Delta\nu_3 - \frac{m_3}{m_1}\Delta\nu_1 \quad (8a)$$

$$\text{electrolyte: } SK_{\mathrm{OBS}} \equiv \left[\frac{\partial \ln(K_{\mathrm{OBS}})}{\partial \ln(a_\pm)}\right]_{T,P,EQ} = (\Delta\nu_+ + \Delta\nu_-) - \frac{n \cdot m_3}{m_1}\Delta\nu_1 \quad (8b)$$

where each $\Delta\nu_i$ is the stoichiometrically weighted change in the number of moles of that species in the local regions of the products and reactants of the process.

2.2. Application to hydrophobic interaction chromatography

The equilibrium constant for adsorption affects the protein's migration velocity relative to the percolating solvent, so changes in the underlying thermodynamics of adsorption are manifested as changes in the retention times. As the mixture of proteins flows through the column, the proteins are slowed by their interactions with the surface. Proteins that spend more time in contact with the surface, and relatively less time in the mobile phase, pass through the column more slowly than solutes that do not adsorb on the column. The flow rate of a particular solute, u_i, is related to the interstitial velocity of the mobile phase, u_m, by

$$u_i = \frac{\epsilon_b \cdot u_m}{\epsilon_b + (1-\epsilon_b)\cdot \epsilon_{p,i} + K^m_{obs}} \quad (9)$$

where ϵ_b is the interstitial void volume of the column, and $\epsilon_{p,i}$ is the accessible porosity of the beads for species i [26]. K^m_{OBS} is the observed equilibrium constant defined by

$$K^m_{obs} = \frac{n_{p,ads}}{n_{p,soln}}[\epsilon_b + (1-\epsilon_b)\cdot \epsilon_{p,i}] \quad (10)$$

where $n_{P,ads}$ and $n_{P,soln}$ are the number of moles of protein adsorbed on the resin and in solution, respectively.

The retention time of solute i is equal to the bed length divided by the solute's velocity. Retention times are readily measured and can be used to calculate the capacity factor of a column, which is given by:

$$k' \equiv \frac{n_{P,ads}}{n_{P,soln}} = \frac{t_{r,i} - t_{m,i}}{t_{m,i}} = \frac{K^m_{obs}}{\epsilon_b + (1-\epsilon_b)\cdot \epsilon_{p,i}} \quad (11)$$

where $t_{r,i}$ is the retention time of a protein that is intermittently adsorbed, and $t_{m,i}$ is the retention time of a solute of the same size that does not bind to the column. For linear chromatography, in which the equilibrium constant and fluid properties are independent of protein concentration [26], the capacity factor is equal to the observed equilibrium constant for the adsorption of protein on the resin divided by the phase ratio of the column. If the phase ratio of the column is unaffected by the composition of the mobile phase, then the observed variation in the retention factor with changes in salt concentration arises from the effect of salt on the observed equilibrium constant for adsorption. Rewriting Eqs. (8b) in terms of the capacity factor and incorporating the variation in the mean ionic activity with the molal salt concentration gives:

$$\left[\frac{\partial \ln(k')}{\partial \ln(m_3)}\right]_{T,P,EQ} = \frac{(\Delta\nu_+ + \Delta\nu_-)}{g} - \frac{n\cdot\Delta\nu_1}{m_1\cdot g} m_3 \quad (12)$$

where

$$g = \left(\frac{\partial \ln m_3}{\partial \ln a_\pm}\right)_{T,P}.$$

Values for g can be calculated from tabulated data for activity coefficients as a function of salt concentration [27], or by applying recently developed thermodynamic models [28].

No assumptions regarding the adsorption mechanism have been made, making this a completely general result that could be applied to ion exchange, hydrophobic interaction or reverse phase chromatography. At low salt concentrations where adsorption is governed by electrostatic interactions, the first term on the right hand side of the equation is dominant. Thus the decrease in retention times with increasing salt concentrations is controlled by the displacement of counter-ions, which is consistent with widely used stoichiometric displacement models of ion exchange chromatography.

At the higher salt concentrations used in hydrophobic interaction chromatography, the second term in Eq. (12) is dominant, and the increase in retention times with salt concentrations is controlled by the displacement of water molecules. Two simple cases are considered here: (i) ion and water stoichiometries are independent of the salt concentration and (ii) binding of ions is proportional to the salt concentration. In the first case, integration of Eq. (12) yields

$$\ln(k') = c + \frac{(\Delta\nu_+ + \Delta\nu_-)}{g}\cdot\ln(m_3) - \frac{n\cdot\Delta\nu_1}{m_1\cdot g} m_3 \quad (13)$$

In the second case, however, integration of Eq. (12) gives

$$\ln(k') = c + \left[\frac{(\Delta b_+ + \Delta b_-)}{g} - \frac{n\cdot\Delta\nu_1}{m_1\cdot g}\right]\cdot m_3 \quad (14)$$

where $\nu_i = b_i \cdot m_3$, and Δb is the stoichiometrically weighted change in the ion binding coefficients. Note that in case (ii) the $\ln(m_3)$ term present in Eq. (13) is eliminated, and $\ln(k')$ varies linearly with molal salt concentration. Thus, the proportional binding case is indistinguishable from a constant binding situation in which $(\Delta\nu_+ + \Delta\nu_-) = 0$ unless independent information on the protein–salt interactions is available.

These models are used below to examine the retention of proteins on hydrophobic interaction chromatography columns. Column capacity factors for hen egg white lysozyme and ovalbumin have been measured as a function of mobile phase ionic

strength for a series of columns with varying levels of hydrophobicity, and the variation in protein retention times with mobile phase concentration and column and protein hydrophobicity is evaluated. Then the preferential interaction analysis is applied to retention data for proteins in different salts, showing that the model has broad applicability and can be used to describe or predict protein retention in a variety of systems.

3. Experimental

3.1. Materials

Two globular proteins, hen egg white lysozyme and ovalbumin, were used in these studies. Both proteins were obtained from Sigma (St. Louis, MO, USA). The buffer solutions were prepared from monobasic monohydrate sodium phosphate, dibasic anhydrous sodium phosphate, and ammonium sulfate. All reagents were obtained from Sigma. The buffer solutions contained 0.1 *M* phosphate and ammonium sulfate (0.2, 0.4, 0.6, 0.8 and 1.0 *M*) at a pH of 6.8. Column cleaning solutions included MilliQ water, 0.5 *M* sodium hydroxide solution, and 20% (v/v) ethanol in water.

Five hydrophobic interaction chromatography columns (Pharmacia Biotech, Uppsala, Sweden) were used in these experiments. All of the columns have a nominal bed volume of 1 ml. The hydrophobicity of the columns varies with the ligand used and the ligand density, as shown in Table 1.

3.2. Instrumentation

Protein retention data was obtained with a BioCad 60 (Perseptive Biosystems, Framingham, MA, USA). The BioCad system was configured for a single chromatography column, with a 100 μl sample loop and an analytical flow cell. A UV–Vis detector was used to monitor the absorbance of the effluent at 280 nm.

3.3. Chromatography

For each series of retention measurements, the column was first flushed with 5 ml of low salt buffer (0.1 *M* sodium phosphate, pH = 6.8) at a flow-rate of 1 ml min^{-1}. After equilibrating the column with 25 ml of the high salt buffer (0.1 *M* sodium phosphate, 1 *M* ammonium sulfate, pH = 6.8) at a flow-rate of 1 ml min^{-1}, a 50-μl sample of protein (10 mg of protein per ml of high salt buffer) was injected and the retention time was measured. The column was then equilibrated with 10 ml of the next buffer in the series (0.1 *M* sodium phosphate, 0.8 *M* ammonium sulfate, pH = 6.8), and then the next injection was made. This process was repeated for each buffer in the series. The column was cleaned with 10 ml of MilliQ water, followed by 10 ml of 0.5 *M* sodium hydroxide solution, another 10 ml of MilliQ water, and finally 10 ml of 20% (v/v) ethanol.

The interstitial void volume plus the void volume of column distributors and collectors was determined by measuring the retention time of blue dextran (Sigma, No. D-5751). Using a bed porosity of 36% and a column housing volume of 1.01 ml, the actual bed volume of each column was estimated. To find

Table 1
Summary of ligand densities and hydrophobic areas

Column	Ligand	Ligand density	Ligand areas (Å^2)			Length
		(μmol ml^{-1} of bed)	Total	Side	Tip	(Å)
Phenyl Sepharose HP	Phenyl	25	210	105	80	8.0
Phenyl Sepharose low	Phenyl	20	210	105	80	8.0
Phenyl Sepharose high	Phenyl	40	210	105	80	8.0
Octyl Sepharose	Octyl	5	319	159	48	14.2
Butyl Sepharose	Butyl	50	191	6	48	8.9

Ligand densities listed were provided by Pharmacia Biotech. Ligand areas were estimated from the cavity surface areas reported by Hermann [29].

the accessible intraparticle volume, uptake curves were measured and fitted to a Langmuir model. Langmuir parameters vary with the intraparticle porosity assumed, as does the mobile phase retention time used in the capacity factor calculations. The capacity factor data is measured at low concentrations where the adsorption isotherm is linear, while the adsorption data covers both the linear and nonlinear adsorption regimes, so the capacity factor values and adsorption data have different sensitivities to the unknown intraparticle porosity. Thus a consistent intraparticle porosity can be found by equating the capacity factors predicted from the Langmuir adsorption isotherms to the measured capacity factors:

$$k'_{\mathrm{ret}} = \frac{t_{\mathrm{r},i} - t_{\mathrm{m},i}}{t_{\mathrm{m},i}} \tag{15}$$

$$k'_{\mathrm{ads}} = \frac{S_0}{\epsilon_{\mathrm{b}} + (1 - \epsilon_{\mathrm{b}})\epsilon_{\mathrm{p},i}} \tag{16}$$

where S_0 is the initial slope of the Langmuir adsorption isotherm calculated from the line fit to a Hanes plot (c/q versus c). These intraparticle porosities are consistent with available pore size distribution data for these columns. The bed parameters used are shown in Table 2.

3.4. Contact area calculations

The reduction in hydrophobic surface area on adsorption was found by using a simple model of a spherical protein in contact with a flat surface that is studded with ligand "rods". The total contact area is given by the sum of the protein–surface and protein–ligand contact areas, and the reduction in wetted area on adsorption is twice the total contact area. The reduction in exposed hydrophobic area is lower, however, because the protein surface, with a mixture of hydrophilic and hydrophobic patches, is contacting the hydrophobic ligands and substrate [30].

The fraction of the protein surface composed of hydrophobic amino acids was estimated from protein images generated by SYBYL [31], shown in Fig. 1. For ovalbumin and lysozyme, 40% and 33% of their respective surfaces are hydrophobic. Thus the reduction in hydrophobic area is 1.4 times the total contact area for ovalbumin and 1.33 times the total contact area for lysozyme.

The protein–surface contact area, or the area devoid of water molecules due to the presence of protein, is given by

$$A_{\mathrm{S-P}} = \pi \cdot (2 \cdot \sqrt{r_{\mathrm{p}} \cdot r_{\mathrm{w}}} - r_{\mathrm{w}})^2 \tag{17}$$

where r_{w} is the radius of a water molecule and r_{p} is the radius of the protein. For globular proteins, the radius (in Å) varies with molecular weight [32]:

$$r_{\mathrm{p}} = \sqrt{\frac{6.4}{4\pi}(m_{\mathrm{w}}^{0.73})} \tag{18}$$

To determine the protein–ligand contact area, a hexagonal arrangement of the ligands was assumed. The ligand spacing was based on the ligand densities provided by the manufacturer and estimates of the total surface area obtained from our adsorption studies. An algorithm was written to count the protein–ligand contacts for various protein positions and estimate the total ligand area in contact with the protein. Two possible orientations of the ligand were considered: (1) side contact, which occurred only when the ligand was close enough to lie tangent to the protein surface and (2) tip contact. The protein–ligand contact areas for the various protein positions were then averaged. The results of these calculations are summarized in Table 3, which gives the protein–

Table 2
Summary of column volumes and porosities

Column	Bed volume (ml)	Bed porosity, ϵ_{b}	Particle porosity, $\epsilon_{\mathrm{p,HEW}}$	Particle porosity, $\epsilon_{\mathrm{p,OVA}}$
Butyl	0.99	0.36	0.71	0.68
Octyl	1.05	0.36	0.84	0.79
PS-LO	0.98	0.36	0.77	0.73
PS-HP	0.97	0.36	0.69	0.64
PS-HI	0.99	0.36	0.82	0.79

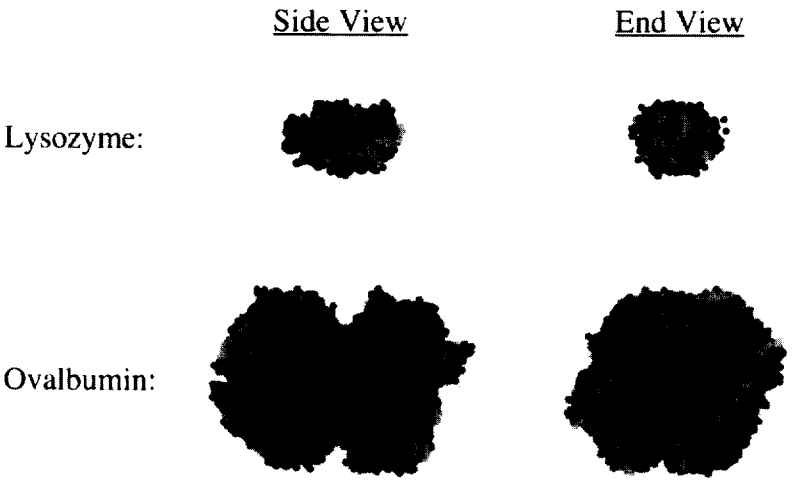

Fig. 1. Images of ovalbumin and hen egg-white lysozyme generated with SYBYL, showing the distribution of hydrophobic and hydrophilic amino acids. Water molecules are black, hydrophilic regions are dark grey, and hydrophobic areas are light grey.

surface and protein–ligand contact areas for ovalbumin and lysozyme on each of the columns, as well as the total reduction in hydrophobic surface area for each protein–column pair.

4. Results and discussion

The preferential interaction analysis was applied to protein retention on hydrophobic columns. First this approach was used to examine the variation in protein retention with column hydrophobicity, and the correlation between the effect of salt on retention and the weted hydrophobic area was explored. The preferential interaction analysis was then used to investigate the effects of changing salt composition. Salts that increase preferential hydration of hydrophobic surfaces were found to have the greatest impact on retention times.

Table 3
Hydrophobic contact areas

Protein	Column	Hydrophobic contact areas (Å^2)			
		Protein–surface	Protein–ligand (average)	Total area	PI model: Area from release of water molecules
OVA	Octyl	529	52	813	1205
	PS-LO	529	120	908	1140
	PS-HP	529	161	966	1800
	PS-HI	529	228	1060	1858
	Butyl	529	251	1092	2208
HEW	Octyl	335	40	500	1017
	PS-LO	335	82	555	1160
	PS-HP	335	109	592	1105
	PS-HI	335	155	652	1033
	Butyl	335	177	681	1346

Comparisons of two estimates of the reduction in hydrophobic area on adsorption. The first is calculated from a simple model using ligand density of the column and protein size. The second uses the area uncovered by the release of water molecules, found using the preferential interaction analysis of retention data.

4.1. Variation in protein retention with column hydrophobicity

To examine the interacting effects of protein and column hydrophobicity and mobile phase salt concentration, the retention times of ovalbumin and hen egg white lysozyme were measured at ammonium sulfate concentrations ranging from 0.2–1.0 *M* on each of the five hydrophobic columns. The column capacity factors, defined in Eq. (11), were then calculated. The preferential interaction model given in Eq. (13) was fitted to the experimental data. Non-linear least squares fitting was done with Engineering Equation Solver [33] to obtain values for c, $(\Delta\nu_+ + \Delta\nu_-)/g$ and $(-n \cdot \Delta\nu_1)/(m_1 \cdot g)$. The parameters obtained for each protein–column combination are given in Table 4. Good agreement between the model and the data was obtained, as shown in Fig. 2. In this system, n is 3, m_1 is 55.15 and g is 1.6, so the model parameters found by fitting the data to Eq. (13) can be used to calculate the number of water molecules and ions released on adsorption. These calculated values are listed in Table 4.

The preferential interaction model was then used to examine the salt dependence of the capacity factors. By inserting the parameters obtained for Eq. (13) into Eq. (12), SK_{OBS} or $[\partial \ln(k')]/[\partial \ln(m_3)]_{T,P}$ is seen to vary linearly with salt concentration. The slope of SK_{OBS} versus salt concentration is $(-n \cdot \Delta\nu_1)/(m_1 \cdot g)$ and the intercept is $(\Delta\nu_+ + \Delta\nu_-)/g$ Thus the slope is proportional to the number of water molecules released upon adsorption of the protein, and the intercept is proportional to the change in the binding of ions.

For all of the protein–substrate pairs studied, adsorption leads to the release of significantly more water molecules (100–200) than salt ions (0.5–3). This result is consistent with the view that hydrophobic interactions are driven by the increase in entropy resulting from the release of ordered water molecules. No correlation was found between the number of ions released and the hydrophobicity of the substrate. Instead the number of ions released appears to be constant for each protein, with a value of 1.6±1.3 for ovalbumin and 1±0.4 for lysozyme. Note that the number of water molecules and ions released corresponds to a salt concentration of 0.16 *M*. This value is lower than the bulk salt concentrations used, confirming that salt ions are being preferentially excluded from the water layer in contact with the protein or column surfaces. The exclusion of salt ions, or the preferential hydration of the surfaces, is expected for adsorption processes that are favored by the addition of salt.

The slope of SK_{OBS} versus salt concentration, which is proportional to the number of water molecules released $(-\Delta\nu_1)$, was found to increase with the total hydrophobic area of the column and also with the hydrophobic area of the protein, as shown in Fig. 3. These trends closely follow the estimated

Table 4
Parameters describing variation in retention with ammonium sulfate concentration

Protein	Column	$\frac{-n \cdot \Delta\nu_1}{m_1 \cdot g}$	$\frac{(\Delta\nu_+ + \Delta\nu_-)}{g}$	c	$-\Delta\nu_1$	$-(\Delta\nu_+ + \Delta\nu_-)$
OVA	Octyl	4.1	−0.4	−3.1	120	0.7
	PS-LO	3.8	−0.3	−2.7	111	0.4
	PS-HI	5.7	−1.5	−4.6	169	2.4
	PS-HP	6.2	−0.5	−3.2	184	0.8
	Butyl	7.1	−1.8	−6.6	211	2.9
HEW	Octyl	3.4	−0.6	−3.1	101	1.0
	PS-LO	3.9	−0.6	−2.7	114	0.9
	PS-HI	3.7	−0.7	−2.5	108	1.1
	PS-HP	3.5	−0.2	−1.0	104	0.4
	Butyl	4.4	−0.8	−3.8	130	1.2

Summary of preferential interaction model parameters and the corresponding changes in binding of water molecules and salt ions on adsorption. The uncertainties for $\Delta\nu_1$ are ±15 for ovalbumin and ±6 for lysozyme. For $(\Delta\nu_+ + \Delta\nu_-)$, the uncertainties are ±0.4 for ovalbumin and ±0.2 for lysozyme.

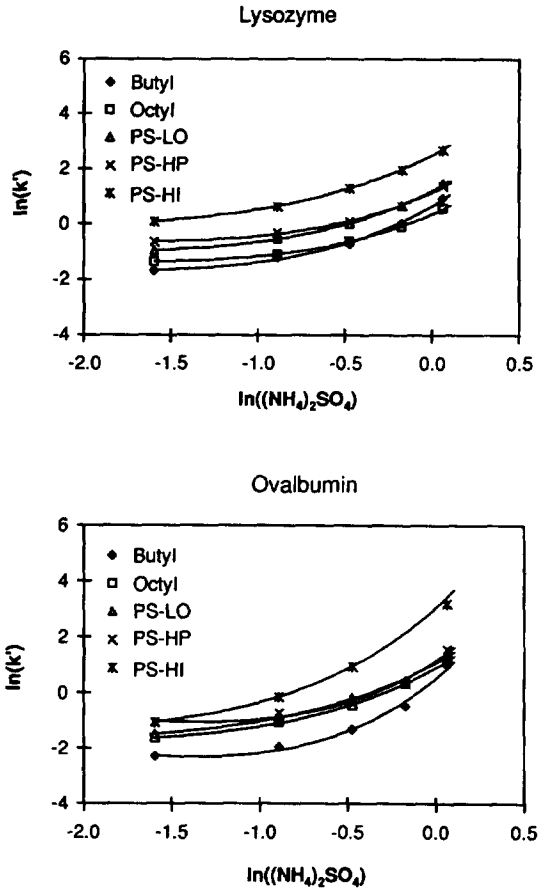

Fig. 2. Effect of column hydrophobicity on the variation of capacity factors with molal ammonium sulfate concentration for hen egg white lysozyme and ovalbumin.

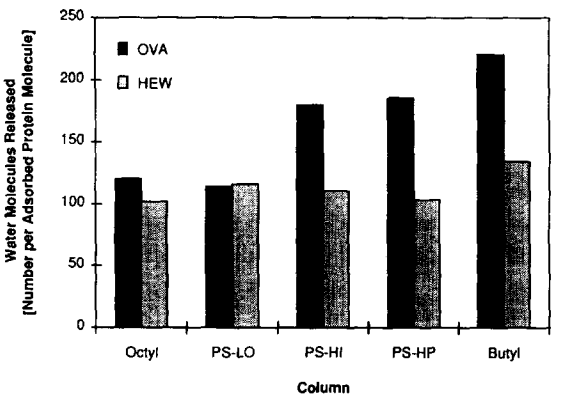

Fig. 3. Calculated number of water molecules released on binding of ovalbumin and lysozyme as a function of column.

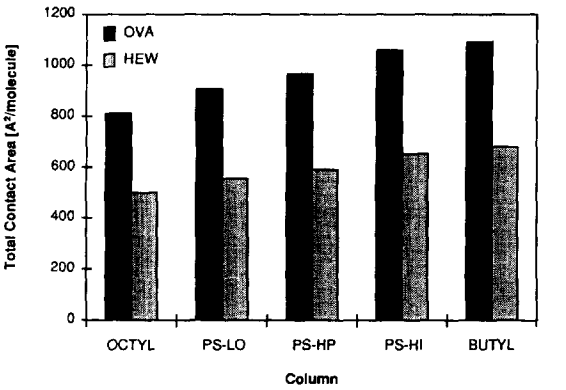

Fig. 4. Variation in the protein–column contact area (estimated from protein size and ligand density) with column hydrophobicity for ovalbumin and lysozyme.

reduction in wetted hydrophobic area, shown in Fig. 4, suggesting that the reduction in exposed area might be used to predict the effect of salt on protein retention.

To examine these trends quantitatively, the calculated number of water molecules released was used to provide a second estimate of the reduction in wetted surface area on adsorption. The water molecules were assumed to form a close-packed monolayer, with each molecule having a radius of 1.5 Å [29] and effectively occupying an area of 7.8 $Å^2$. These estimates of the hydrophobic contact areas are listed in Table 3, along with the values estimated from protein size and ligand density.

The correlation between the two estimates of the hydrophobic contact area is shown in Fig. 5. The contact areas determined from the preferential interaction model are 1.4±0.1 times higher than the estimated values based on protein size and ligand density. This difference may indicate that more than a monolayer of water is desorbed from each hydrophobic surface during adsorption, or that the protein spreads once it is in contact with the surface. It may also reflect the sensitivity of the contact area estimates to the value used for the protein radius. The uncertainty in the estimates of the protein radii make quantitative interpretation of the relationship between the contact area and the number of water molecules released difficult, but the larger protein has larger contact areas from both estimate techniques on all of the columns (Fig. 5). The variability in contact areas

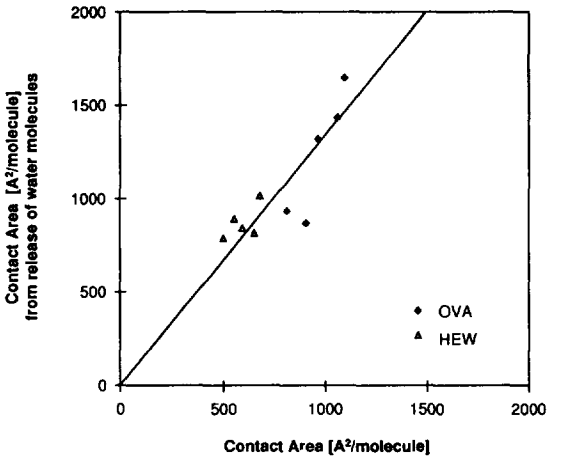

Fig. 5. Comparison of the hydrophobic contact area calculated from the chromatographic data to the value estimated from the ligand densities and protein sizes.

seen for each protein across the five columns reflects the difficulty in evaluating the column ligand densities. Nevertheless, a semi-empirical model based on these trends can still be used to predict the effects of changing the protein or column hydrophobicity.

The results presented thus far indicate that the observed variation in retention times with salt concentration is consistent with an adsorption process that is dominated by the release of water molecules from hydrophobic surfaces. The correlation between the number of water molecules released and the reduction in hydrophobic area suggests that the variation in $\ln(k')$ with salt concentration can be predicted from an estimate of the reduction in wetted hydrophobic area. Fig. 5 relates contact area estimates to the parameter $\Delta\nu_1$ for protein adsorption in ammonium sulfate solutions. For other salts, different water-ordering effects may affect the variation in $\Delta\nu_1$ with contact area. Thus we now consider the effect of salt type on protein retention.

4.2. Variation in protein retention with salt species

To examine the application of the preferential interaction model to systems using different salts, data reported by Fausnaugh and Regnier [10] for hen egg white lysozyme, ovalbumin and conalbumin adsorbing on TSKgel in ammonium sulfate, sodium sulfate, sodium tartrate and magnesium sulfate were used. Their original analysis of these data, based on solvophobic theory, showed that the variation in the retention times with salt composition could not be fully explained by changes in the molal salt increments. For sodium tartrate and magnesium sulfate the authors suggested that ion binding might be responsible for the observed deviations.

As seen earlier, if the change in the binding of ions and water is independent of the salt concentration, then integration of Eq. (12) gives the form of the preferential interaction model used previously (Eq. (13)). With binding that is proportional to the salt concentration, however, integration of Eq. (12) gives a slightly modified form shown in Eq. (14). In the case of proportional binding, $\ln(k')$ varies linearly with molal salt concentration. The data for magnesium sulfate and sodium tartrate displayed this linear behavior, while the data for sodium sulfate and ammonium sulfate did not. Therefore, we fit the data for sodium sulfate and ammonium sulfate with Eq. (13) and the data for magnesium sulfate and sodium tartrate with Eq. (14).

For sodium sulfate and ammonium sulfate, the number of water molecules and salt ions released was calculated, and the results are summarized in Table 5. For these calculations, g is 1.68 for ammonium sulfate and 1.64 for sodium sulfate, and n is 3 for both salts. The release of water molecules and salt ions on adsorption increases with the surface area of the protein, as shown in Figs. 6 and 7. The stoichiometrically weighted change in the number of ions is the same for these two salts. Thus the observed differences in protein retention are due to the effect of the salt on the amount of water released from the hydrophobic regions of the protein and sorbent, rather than to differences in ion binding.

For magnesium sulfate and sodium tartrate, the variation in retention time with salt concentration is linear. Thus the data is well described by the preferential interaction model with modifications for concentration dependent ion binding (Eq. (14)). The change in ion binding on adsorption and the release of water molecules cannot be found independently,

Table 5
Retention parameters for lysozyme, ovalbumin and conalbumin in various salts

Protein	Salt	$\frac{-n \cdot \Delta\nu_1}{m_1 \cdot g}$	$\frac{(\Delta\nu_+ + \Delta\nu_-)}{g}$	c	$-\Delta\nu_1$ or $-\Delta\nu_{1,\mathrm{eff}}$	$\Delta\nu_+ + \Delta\nu_-$
HEW	Na_2SO_4	8.0	−1.8	−5.5	247	−3.0
	$(NH_4)_2SO_4$	7.2	−3.1	−6.5	217	−5.1
	$MgSO_4$	4.2		−3.7	219	
	Na_2Tart	5.9		−3.3	127	
OVA	Na_2SO_4	21.6	−10.1	−19.3	671	−17.0
	$(NH_4)_2SO_4$	12.8	−6.2	−12.6	387	−10.2
	$MgSO_4$	14.0		−17.4	730	
	Na_2Tart	9.2		−8.0	195	
CON	Na_2SO_4	28.4	−15.2	−26.8	884	−25.6
	$(NH_4)_2SO_4$	23.5	−17.7	−24.9	711	−29.1
	$MgSO_4$	13.2		−17.0	686	
	Na_2Tart	10.5		−9.8	223	

Summary of preferential interaction model parameters and the corresponding changes in binding of water molecules and salt ions. Results are based on fits of Fausnaugh and Regnier's retention data on TSKgel Phenyl-5PW [10]. In addition to the salt listed, the solutions contained 0.01 *M* potassium phosphate buffer at pH of 7.0.

so for each protein the effective number of water molecules released is calculated:

$$-\frac{n \cdot \Delta\nu_1^{\mathrm{eff}}}{m_1 \cdot g} = \left[\frac{(\Delta b_+ + \Delta b_-)}{g} - \frac{n \cdot \Delta\nu_1}{m_1 \cdot g}\right] \qquad (19)$$

Since the number of ions released for a concentration-dependent ion-binding salt is proportional to m_3, these ions plus a certain portion of the water released will have the molality of the bulk solution. These ions and water molecules will have no hydrophobic effect. Only the additional ion-free water will cause the hydrophobic effect. The number of ion-free water molecules released on adsorption is given by

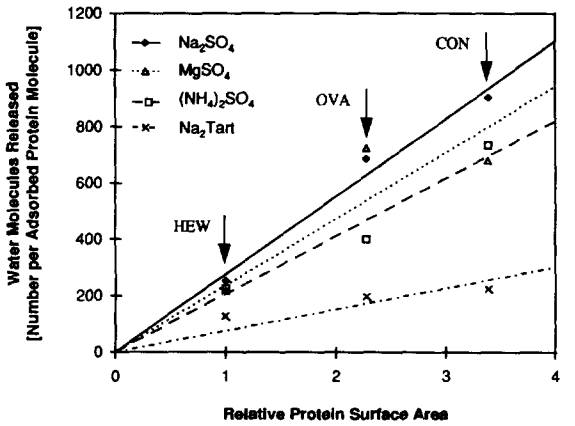

Fig. 6. Variation in the number of water molecules released on adsorption on TSKgel with protein size and salt for hen egg white lysozyme, ovalbumin and conalbumin.

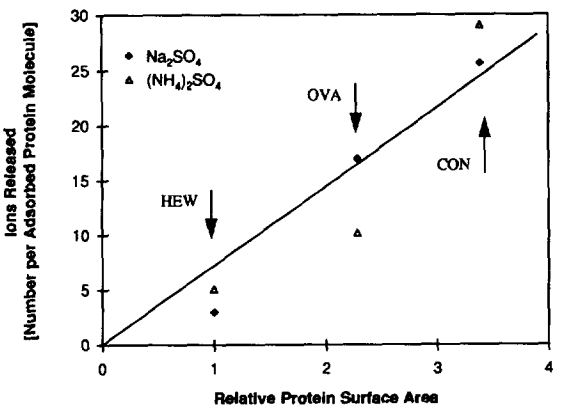

Fig. 7. Variation in the number of ions released on adsorption on TSKgel with protein size and salt for hen egg white lysozyme, ovalbumin and conalbumin.

$$\Delta\nu_1^{\text{eff}} = \Delta\nu_1 - \frac{m_1}{n}(\Delta b_+ + \Delta b_-)$$
$$= \Delta\nu_1 - \frac{m_1}{m_3}(\Delta\nu_+ + \Delta\nu_-) \quad (20)$$

where the second term is the number of water molecules needed to release the bound ions at the bulk concentration, m_3. For magnesium sulfate n is 2 and g is 1.88, and for sodium tartrate n is 3 and g is 1.15. The effective number of water molecules released on adsorption is plotted in Fig. 6.

Fig. 6 shows that the number of water molecules released on adsorption increases with the size of the protein, and hence the amount of exposed hydrophobic area eliminated on adsorption. The number of water molecules released also increases with the salt's ability to enhance preferential hydration at hydrophobic surfaces. The effectiveness of cations for inducing preferential hydration is greater for Na^+ than for Mg^{2+} [16,17]. From their positions in the lyotropic series [9], preferential hydration is greater with Na^+ than NH_4^+. For the anions, the water-ordering ability of sulfate is greater than that of tartrate. Thus the water-ordering is highest for sodium sulfate, and decreases in the order of magnesium sulfate, ammonium sulfate and sodium tartrate. This trend holds for the observed release of water molecules: magnesium sulfate, ammonium sulfate and sodium tartrate provide only 86%, 74% and 27%, respectively, of the water ordering ability of sodium sulfate. The ratio of preferential interaction coefficients for sodium sulfate and magnesium sulfate with bovine serum albumin [17] is 83%, so the values based on this analysis are similar to those obtained by direct measurement of the preferential interaction coefficients. The trends observed for these salts can be used to predict the variation in SK_{OBS} for other proteins on this TSKgel column.

4.3. Modeling protein retention in HIC

The preferential interaction model provides a thermodynamic framework that allows simple, quantitative predictions of the effect of salt on protein retention to be made. Such predictions require: (1) an estimate of the hydrophobic area of the protein, (2) estimates of the accessible surface area and the ligand density of the substrate, (3) values of

$$g = \left(\frac{\partial \ln m_3}{\partial \ln a_\pm}\right)_{T,P}$$

from data references or theory, (4) the present correlations for $\Delta\nu_1$ and $(\Delta\nu_+ + \Delta\nu_-)$ and (5) a single measured retention factor. With the estimates of the hydrophobic areas of the protein and the substrate, the reduction in wetted hydrophobic area on adsorption can be calculated. This contact area can be used with Fig. 5 to get a value for $\Delta\nu_1$ in ammonium sulfate. The number of water molecules released on adsorption with other salts can be found by multiplying this $\Delta\nu_1$ by the relative slopes of the lines in Fig. 6. Similarly, data like that in Table 4 or Fig. 7 can be used to estimate the number of ions released with ammonium sulfate or sodium sulfate. With this information and a value for g, the parameters used in Eq. (12) can be calculated, and an expression for the linear variation in SK_{OBS} with salt concentration can be obtained. With a single measured retention time, this equation for SK_{OBS} can be integrated to generate the preferential interaction model that describes the variation in retention times with salt concentration. This model can then be used to predict the retention behavior of the protein at any other salt concentration. In a future paper [34], this approach will be used to predict protein capacity factors in other protein–salt–substrate systems.

5. Conclusions

A thermodynamic description of the effects of solutes on assembly processes has been used to examine the retention behavior of proteins on hydrophobic interaction chromatography columns. When used in conjunction with the two-domain model of Timasheff [13,16–18], this approach relates the observed changes in retention times to the release of ions and water molecules on adsorption.

Application of this model to retention data for hen egg white lysozyme and ovalbumin on a series of hydrophobic interaction chromatography columns

indicates that the protein retention in these systems is dominated by the release of adsorbed water molecules, which is consistent with the entropically driven nature of hydrophobic interactions. The number of water molecules released was calculated and was found to increase with the size of the protein and the hydrophobicity of the resin. For ammonium sulfate, the number of water molecules released corresponds to 1.4 times the number predicted for liberation of a monolayer from the hydrophobic surfaces eliminated during adsorption.

Analysis of protein retention data as a function of mobile phase composition indicates that this description can be used to characterize the effect of salts and ion binding. The simplest analysis, in which the stoichiometrically weighted change in the binding of ions and water molecules is independent of salt concentration, shows that the variation in retention time with salt concentration reflects the displacement of ordered water molecules. The number of water molecules released on adsorption increases with protein size and with the salting-out ability of the salt. With slight modifications, the same analysis can be applied to systems in which ion binding is proportional to the salt concentration. This analysis, which is applicable to data for protein retention in sodium tartrate and magnesium sulfate, shows that the observed variation in the retention times with salt concentration can be described by considering the reduction in surface area that results both in displacement of water molecules and in reduction of the binding of ions. Regardless of the form of the model used, the release of water molecules increases with the surface area of the protein and the water-ordering ability of the salt.

This model provides a promising method for interpreting the variation in retention times resulting from the complex interplay of mobile phase composition, concentration, and protein and column hydrophobicity. Simple correlations between the reduction in wetted hydrophobic area on adsorption and the number of water molecules and ions released have been found, and the preferential interaction model provides a means of linking these values to the observed retention behavior. This technique can be used to predict retention behavior in a new system from an estimate of the contact area, knowledge of the salt characteristics, and a single retention measurement.

Acknowledgments

Funding for this work was provided by NSF Grant CTS 9204436. Access to the BioCad 60 was provided by the Bioseparations Facility at UW-Madison.

References

[1] B.F. Roettger and M.R. Ladisch, Biotech. Adv., 7 (1989) 15.
[2] Z. El Rassi and Cs. Horvarth, J. Liquid Chromatogr., 9(15) (1986) 3245.
[3] H. Engelhardt and U. Schon, J. Liquid Chromatogr., 9(15) (1986) 3225.
[4] W. Norde, Adv. Colloid Interface Sci., 25 (1986) 267.
[5] T. Arakawa, Arch. Biochem. Biophys., 248(1) (1986) 101.
[6] X. Geng, L. Guo and J. Chang, J. Chromatogr., 507 (1990) 1.
[7] J.G.E.M. Fraaije, R.M. Murris, W. Norde and J. Lyklema, Biophys. Chem., 40 (1991) 303.
[8] W.R. Melander and Cs. Horvath, Theory of Reversed-Phase Chromatography, in Cs. Horvath, (Ed.), High Performance Liquid Chromatography, New York, Academic Press, 1980, p. 201.
[9] W. Melander and Cs. Horvath, Arch. Biochem. Biophys., 183 (1977) 200.
[10] J.L. Fausnaugh and F.E. Regnier, J. Chromatogr., 359 (1986) 131.
[11] T. Arakawa and S.N. Timasheff, Biochemistry, 23 (1984) 5912.
[12] N.P. Golovchenko, I.A. Kataeva and V.K. Akimenko, Enzyme Microb. Technol., 14 (1992) 327.
[13] M.T. Record, Jr. and C.F. Anderson, Biophys. J., 68 (1995) 786.
[14] J-H. Ha, M.W. Capp, M.D. Hohenwalter, M. Baskerville and M.T. Record, J. Mol. Biol., 228 (1992) 252.
[15] M.T. Record, Jr, J.-H. Ha and M.A. Fisher, Methods Enzymol., 208 (1991) 291.
[16] S.N. Timasheff, Annu. Rev. Biomol. Struct., 22 (1993) 67.
[17] T. Arakawa and S.N. Timasheff, Biochemistry, 21, (1982) 6545.
[18] S.N. Timasheff and T. Arakawa, J. Crystal Growth, 90 (1988) 39.
[19] S.-L. Wu, A. Figueroa and B.L. Karger, J. Chromatogr., 371 (1986) 3.
[20] T. Arakawa, Biopolymers, 26 (1987) 45.
[21] T. Arakawa, R. Bhat and S.N. Timasheff, Biochemistry, 29 (1990) 1914.

[22] C.F. Anderson and M.T. Record, Jr., J. Phys. Chem., 97 (1993) 7116.
[23] M.T. Record, Jr., T.M. Lohman and P. de Haseth, J. Mol. Biol., 107 (1976) 145.
[24] P.R. Wills and D.J. Winzor, Biopolymers, 33 (1993) 1627.
[25] J.A. Schellman, Biphys. Chem., 45 (1993) 273.
[26] A.M. Athalye, S.J. Gibbs and E.N. Lightfoot, J. Chromatogr., 589 (1992) 71.
[27] D. Dobos, Electrochem. Data, Elsevier Science, New York, 1975.
[28] A. Kolker and J.J. de Pablo, AIChE J., 41 (1995) 1563.
[29] R.B. Hermann, J. Phys. Chem., 76(19) (1972) 2754.
[30] E. Heftmann, (Ed.), J. Chromatog. Library, Elsevier, New York, 1983. Vol. 22A: Chromatography Fundamentals and Applications of Chromatographic and Electrophoretic Methods. Part A: Fundamentals and Techniques, R.P. Bywater and N.V.B. Marsden, Hydrophobic Interactions in Gels, A309-A313.
[31] Tripos Associates, Inc., SYBYL (6.1). St. Louis: Tripos, 1994.
[32] T.E. Creighton, Proteins: Struct. Mol. Prop., 2nd Ed., W.H. Freeman, NY, 1984.
[33] S.A. Klein and F.L. Alvarado, Engineering Equation Solver. Middleton, WI: F-Chart Software, 1996.
[34] T.W. Perkins, T.W. Root and E.N. Lightfoot, in preparation.