

Available online at www.sciencedirect.com



Analytical Biochemistry 346 (2005) 132-138

ANALYTICAL BIOCHEMISTRY

www.elsevier.com/locate/yabio

# Composition gradient static light scattering: A new technique for rapid detection and quantitative characterization of reversible macromolecular hetero-associations in solution

Arun K. Attri<sup>1</sup>, Allen P. Minton\*

Laboratory of Biochemistry and Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, U.S. Department of Health and Human Services, Bethesda, MD 20892, USA

> Received 7 June 2005 Available online 8 September 2005

## Abstract

A method for rapid detection and quantitative characterization of associations between two proteins in solution is presented, based upon continuous measurement of the Rayleigh light scattering and UV absorbance of a solution whose composition is varied with time in a controlled fashion. Collection and subsequent analysis of data are both rapid and semiautomatic. The results of validation experiments demonstrate that the technique can correctly identify complexes and reliably evaluate equilibrium constants for hetero-association. © 2005 Elsevier Inc. All rights reserved.

Keywords: Protein-protein association equilibria; Rayleigh light scattering

Ouantitative characterization of reversible macromolecular associations in solution is prerequisite to the understanding of how complex biochemical systems respond to changes in composition and environmental variables [1,2]. In general, techniques developed for this purpose are based upon measurement and analysis of the dependence of some average property of the solution on its composition. Perhaps the more reliable of these techniques are those measuring thermodynamic properties such as light scattering, sedimentation equilibrium, and osmotic pressure, that are direct functions of the weight- or number-average mass of solute, since no ad hoc assumptions regarding the contribution of a particular microscopic species to the total measurable signal are required. However, measurements of these properties have traditionally been material and time consuming and labor intensive. We have recently introduced a rapid and precise method for detection and characterization of macromolecular self-association, based on automated measurement and analysis of concentration-dependent changes in Rayleigh (or "static") light scattering [3].<sup>2</sup> We present here an extension of this methodology permitting equally rapid and precise detection and characterization of hetero-associations between two different proteins in solution. Experimental results and analysis are presented for two pairs of proteins: fibrinogen + bovine serum albumin and chymotrypsin + soybean trypsin inhibitor.

#### Materials and methods

## Materials

Bovine fibrinogen (Type IV) and BSA<sup>3</sup> were obtained from Sigma. Chymotrypsin and soybean trypsin inhibitor were obtained from Worthington Biochemical. All proteins were dialyzed extensively against phosphate-buffered saline

<sup>\*</sup> Corresponding author. Fax: +1 301 402 0240.

E-mail address: minton@helix.nih.gov (A.P. Minton).

<sup>&</sup>lt;sup>1</sup> Permanent address: School of Environmental Sciences, Jawaharlal Nehru University, New Delhi 110 067, India.

<sup>0003-2697/\$ -</sup> see front matter @ 2005 Elsevier Inc. All rights reserved. doi:10.1016/j.ab.2005.08.013

<sup>&</sup>lt;sup>2</sup> This publication will be referred to subsequently as AM.

<sup>&</sup>lt;sup>3</sup> Abbreviation used: BSA, bovine serum albumin.



Fig. 1. Schematic of apparatus. (1 and 2) Reservoirs for solutions A and B. (3) Programmable dual-syringe pump. (4) Programmable valves for switching syringes between filling and delivery modes. (5) T-junction for mixing of input streams from the two syringes. (6) Peristaltic pump for delivery of buffer in reservoir (7). (8) Valve for switching apparatus input between peristaltic pump (6) and syringe pump (3). (9) Inline filter. (10) T-junction for splitting solution mixture into parallel streams for concurrent measurement in absorbance detector flow cell (11) and multiangle light scattering detector flow cell (12). Adapted from [3].

(0.05 M Na/K phosphate + 0.15 M NaCl, pH 7.2), and used without further purification. Protein concentrations were determined from the absorbance at 280 nm using the following standard values for absorbance in OD units per centimeter pathlength for a 1 g/L solution: BSA, 0.65 [4]; fibrinogen, 1.20 [5]; chymotrypsin, 2.04 [6]; soybean trypsin inhibitor, 0.94 [7]. Refractive increments were determined as described previously [3] and found to be equal to  $0.185 \pm 0.003$  ml/g at 20 °C for all proteins utilized in the present study. Immediately prior to light-scattering measurement, solutions were prefiltered and centrifuged as described previously [3].

#### Instrumentation

The apparatus, schematically illustrated in Fig. 1, is essentially identical to that described earlier in AM. A Hamilton Microlab 541C dual-syringe pump (Hamilton, Reno, NV) controlled by a personal computer<sup>4</sup> was programmed to deliver a temporal gradient of solution composition simultaneously to a multiangle 690-nm laser light-scattering detector (DAWN-EOS, Wyatt Technology, Santa Barbara, CA), and a UV–visible absorbance detector (SM3100, Thermo Finnegan, Palm Beach, FL). Scattering measurements were carried out at 20 °C. Scattering and absorbance data were collected together using ASTRA software (Wyatt Technology, Version 4.90), and exported as text files for subsequent analysis.

# Experimental procedures

Dilution experiments conducted to detect and quantify self-association of a single protein were carried out as pre-



Fig. 2. Relative absorbance (solid line) and 90° scattering (dashed line) plotted as functions of elapsed time over the course of a composition gradient proceeding from  $f_A = 0$  to  $f_A = 1$ . In this particular experiment, A is chymotrypsin and B is soybean trypsin inhibitor.

viously described [3]. This protocol was modified as follows for experiments conducted to detect and quantify association between two different macromolecular solutes, referred to as A and B. Let us refer to a solution containing A at w/v concentration  $w_A^o$  as solution A, and a solution containing B at w/v concentration  $w_{\rm B}^{\rm o}$  as solution B. Solutions A and B are placed in reservoirs 1 and 2, respectively, and loaded into the corresponding syringes. A baseline is obtained using buffer loaded into reservoir 7. Then pure B is introduced into the scattering/absorbance detectors until a plateau of signal is obtained. Typically this requires 700-800 µl of solution. Then a temporal gradient of composition is initiated, during which the fraction of solution B  $(f_{\rm B})$  in the solution mixture introduced into the detectors is decreased and the fraction of solution A  $(f_{\rm A} = 1 - f_{\rm B})$  is simultaneously increased in stages, over a time period of 5–20 min, until  $f_A = 1$ . Following each stage, corresponding to an increment of 0.05 in the value of  $f_A$ , the syringe pump pauses for a preselected period to allow the solution mixture to equilibrate.<sup>5</sup> Establishment of the complete gradient of pure B to pure A (or vice versa) requires approximately 1 ml of each solution. The entire experiment typically requires a maximum of 2 ml of each solution, at a concentration of ca. 0.5 mg/ml, or a total of ca. 1 mg of each protein. Typical profiles of absorbance and scattering versus elapsed time, obtained from a mixture of chymotrypsin and soybean trypsin inhibitor, are shown in Fig. 2.

# Data processing

All calculations are performed automatically using scripts and functions written and executed in MATLAB (Mathworks, Natick, MA), that are available upon request. The value of the time- and angle-dependent Rayleigh ratio  $R(\theta,t)$  is calculated from the unprocessed scattering

<sup>&</sup>lt;sup>4</sup> User-written pump control software, running under MS-DOS, is available upon request.

<sup>&</sup>lt;sup>5</sup> The time required for equilibration may be determined by visual comparison of the profiles of absorbance and scattering versus time. If association/dissociation equilibria are not rapid, the attainment of a scattering plateau will lag behind attainment of an absorbance plateau, which reflects only kinetics of mixing and flow.



Fig. 3. R/K calculated from the results of the concentration gradient experiment shown in Fig. 2, plotted against  $A_{280}$  and  $\sin^2(\theta/2)$ .



Fig. 4. Three views of  $\langle R \rangle / K$  obtained for five solutions containing various proportions of BSA (A) and fibrinogen (B), plotted against  $w_{A,tot}$  and  $w_{B,tot}$ . The three views are generated by variation of the azimuth, so that from left to right, the data may be observed from below, along, and above the planar grid generated by the model. Symbols: data. Grid: best fit of Eq. (5), calculated with the following parameters:  $M_A = 64.6 \times 10^3$ ,  $M_B = 343 \times 10^3$ .

data as described in Appendix A of the ASTRA for Windows User's Guide (Wyatt Technology). Data outliers (typically <1% of total data points) are removed via the filtering process described in Appendix A of this paper. For convenience in subsequent processing, the Rayleigh ratio is scaled relative to a precalculated optical constant *K* defined in Appendix B. In Fig. 3, the scaled Rayleigh ratio  $R(\theta,t)/K$  calculated from the gradient shown in Fig. 2 is plotted against  $\sin^2(\theta/2)$  and  $A_{280}(t)$ , eliminating time as an explicit variable.

The fraction of solution A and the time-dependent w/ v concentrations of A and B are calculated from the wavelength- and time-dependent absorbance  $A(\lambda,t)$  according to

$$f_{\rm A}(t) = \frac{A(\lambda, t) - w_{\rm B}^{\rm o}\epsilon_{\rm B}(\lambda)}{w_{\rm A}^{\rm o}\epsilon_{\rm A}(\lambda) - w_{\rm B}^{\rm o}\epsilon_{\rm B}(\lambda)},\tag{1}$$

$$w_{\mathrm{A,tot}}(t) = f_{\mathrm{A}}(t)w_{\mathrm{A}}^{\mathrm{o}},\tag{2a}$$

$$w_{\rm B,tot}(t) = [1 - f_{\rm A}(t)]w_{\rm B}^{\rm o},$$
 (2b)

where  $\epsilon_A(\lambda)$  and  $\epsilon_B(\lambda)$  are the extinction coefficients of A and B, respectively, in inverse w/v concentration units.

When all solute species are small relative to the wavelength of scattering light (i.e., maximum dimension less than ca. 40 nm), there should be no angular dependence of scattering [8], and this lack of angular dependence was confirmed for all of the data obtained in the present study.<sup>6</sup> The processed data are then saved for subsequent modeling as a two-column array of  $\{f_A, \langle R \rangle / K\}$ , where  $\langle R \rangle$  is the mean value of *R* obtained from detectors 4 to 18, excepting outliers as defined in Appendix A.<sup>7</sup>

<sup>&</sup>lt;sup>6</sup> The length of fully extended fibrinogen, the largest protein we have studied, exceeds 40 nm. However, we were unable to detect angular dependence of scattering by fibrinogen in excess of the uncertainty of measurement in our system. This finding may or may not indicate that fibrinogen has a flexible conformation and is not, on average, fully extended in solution.

<sup>&</sup>lt;sup>7</sup> The analysis to follow may also be applied to data exhibiting angular dependence of *R*, except that one models the composition dependence of  $\lim_{\theta \to 0} R(\theta, \{w\})$ , obtained via extrapolation, rather than  $\langle R \rangle(\{w\})$ .

Modeling the composition dependence of the scaled Rayleigh ratio

Consider a solution containing two macromolecular components A and B.<sup>8</sup> In principle the solution may contain monomeric A and B as well as a variety of complexes formed by equilibrium self- and/or hetero-association. A variable associated with a particular species  $A_i B_i$  bears the subscript  $ij^9$  We shall in addition allow for the presence of a certain mass fraction of each protein that is incompetent to form complexes even though it is still soluble. Such incompetent protein may be an artifact of purification or production by recombinant techniques. We shall denote the fractional association competence of each protein by  $f_{A,comp}$ and  $f_{\rm B,comp}$ , respectively. Then the total molar concentrations of competent and incompetent protein are given by

$$c_{X,\text{tot}} = f_{X,\text{comp}} w_{X,\text{tot}} / M_X \tag{3a}$$

and

$$c_{X,\text{inc}} = (1 - f_{X,\text{comp}}) w_{X,\text{tot}} / M_X, \qquad (3b)$$

where X may be either A or B. An equilibrium association scheme is defined by the specification of one or more equilibrium association constants of the form

$$K_{ij} = \frac{c_{ij}}{c_{10}^i c_{01}^j},\tag{4}$$

where  $c_{ij}$  denotes the molar concentration of  $A_i B_j$ , and  $c_{10}$ and  $c_{01}$  refer exclusively to the molar concentrations of competent monomeric A and B, respectively.

Conservation of mass is expressed by

$$c_{\rm A,tot} = \sum_{i,j} i c_{ij} = \sum_{i,j} i K_{ij} c_{10}^i c_{01}^j,$$
(5a)

$$c_{\rm B,tot} = \sum_{i,j} j c_{ij} = \sum_{i,j} j K_{ij} c_{10}^i c_{01}^j.$$
(5b)

The variables  $f_{A,comp}$ ,  $f_{B,comp}$ ,  $M_{10}$ ,  $M_{01}$ , and the common logarithms of the various  $K_{ij}$  are specified to be parameters of the model.<sup>10</sup> For each value of  $f_A$  the corresponding value of  $\langle R \rangle / K$  is calculated according to the following sequence:

- 1. The values of  $w_{A,tot}$  and  $w_{B,tot}$  are calculated using Eqs. (2a) and (2b), with independently determined (fixed) values of  $w_A^o$  and  $w_B^o$ .
- 2. Given test values of  $f_{A,comp}$  and  $f_{B,comp}$ , the values of c<sub>A,tot</sub>, c<sub>A,inc</sub>, c<sub>B,tot</sub> and c<sub>B,inc</sub> are calculated using Eqs. (3a) and (3b).
- 3. Given test values of  $M_{10}$ ,  $M_{01}$ , and the log  $K_{ij}$ , the values of  $c_{10}$  and  $c_{01}$  are obtained by either analytical or numerical solution of Eqs. (5a) and (5b).

- 4. The values of all  $c_{ij}$  are calculated using Eq. (4).
- 5. The value of  $\langle R \rangle / K$  is then calculated according to (cf. Eq. (B.4))

$$\langle R \rangle / K = M_{10}^2 c_{\text{A,inc}} + M_{01}^2 c_{\text{B,inc}} + \sum_{i,j} M_{ij}^2 c_{ij}.$$
 (6)

Best-fit values of two solute component model parameters were determined by nonlinear least-squares fitting of a model of the type described above or in AM to the appropriate data set(s) [9]. Compound models were also constructed to enable simultaneous fitting of multiple data sets by models containing both global parameters (common to all data sets) and local parameters (applying to fewer than all data sets). A 95% confidence limit of best-fit model parameters was determined via sum-of-squares profiling [10] combined with the Fisher F test for equality of variances [9].

#### Results

#### Fibrinogen + bovine serum albumin

As a test of the accuracy and precision of our measurements and the correctness of our calculations, control experiments were carried out on mixtures of two proteins, fibrinogen and bovine serum albumin, that are known not to self- or hetero-associate under the conditions of the experiment [11]. In Fig. 4, the values of  $\langle R \rangle / K$  obtained from five dilution experiments, carried out as described in Appendix C, are plotted against  $w_{A,tot}$  and  $w_{B,tot}$ . Also plotted in this figure is the best-fit of the function

$$R/K = M_{\rm A} w_{\rm A,tot} + M_{\rm B} w_{\rm B,tot},\tag{7}$$

which is the special case of Eq. (B.4) in the absence of selfor hetero-association.11

As may be seen in Fig. 4, the data obtained from dilution experiments carried out on all five solution mixtures may be well accounted for by Eq. (7). Combination of Eqs. (2) and (7) indicates that in the absence of self- and hetero-association, scattering should depend linearly on  $f_A$ :

$$R/K = M_{\rm B}w_{\rm B}^{\rm o} + (M_{\rm A}w_{\rm A}^{\rm o} - M_{\rm B}w_{\rm B}^{\rm o})f_{\rm A}.$$
(8)

Experimental observation of a significantly nonlinear dependence of R/K against  $f_A$  (or R/K against t, when  $f_A$ varies linearly with time) is thus a qualitative indicator of interactions between two macrosolutes in solution. Examples of such nonlinear dependence are presented below. Although modeling is not required to detect interactions, it is required to characterize the stoichiometry and strength of those interactions.

#### Chymotrypsin (A) + soybean trypsin inhibitor (B)

The value of  $\langle R \rangle / K$ , calculated from the set of data shown in Fig. 3, is plotted as a function of  $f_A$  in Fig. 5. Also

Models for the composition dependence of the Rayleigh ratio in a solution containing a single solute component were described in AM.

<sup>&</sup>lt;sup>9</sup> For example, the w/v concentration of monomeric A will be denoted by  $w_{10}$ , and the molar mass of A<sub>2</sub>B by  $M_{21}$ . <sup>10</sup> Use of the logarithms ensures that the best-fit value of each  $K_{ij}$  will be

positive definite.

<sup>&</sup>lt;sup>11</sup> Eq. (7) holds even if A and/or B are heterogeneous, in which case the values of  $M_{\rm A}$  and/or  $M_{\rm B}$  are concentration-independent weight averages.



Fig. 5. Lower panel:  $\langle R \rangle / K$  plotted as a function of  $f_A$  for a composition gradient of chymotrypsin ( $w_A^{\circ} = 0.52g/L$ ) + soybean trypsin inhibitor ( $w_B^{\circ} = 0.22g/L$ ) in PBS, pH 7.2. Symbols: experimental data. Dotted curve represents the best fit of a simple 1–1 association model (Eq. (7a) only). Dashed curve represents the best fit of the equilibrium model described by Eqs. (9a) and (9b), with  $f_{A,comp}$  constrained to be equal to 1. Best-fit parameter values and 95% confidence limits are: log  $K_{site} = 5.9 ~(\pm 0.1)$ ,  $M_A = 25.0 ~(\pm 0.7) \times 10^3$ ,  $M_B = 17.2 ~(\pm 1.1) \times 10^3$ . Solid curve represents the best fit of the relaxed equilibrium model, allowing for the possibility of  $f_{A,comp} < 1$ . Best-fit parameter values and 95% confidence limits are: log  $K_{site} = 6.5 ~(\pm 0.3)$ ,  $M_A = 24.9 ~(\pm 0.4) \times 10^3$ ,  $M_B = 21.4 ~(\pm 1.7) \times 10^3$ ,  $f_{A,comp} = 0.7 ~(\pm 0.1)$ . Upper panel: Best fit residuals of the constrained (dashed curve) and relaxed (solid curve) models.

plotted are the calculated best fits of three models. The first model postulates only a simple 1–1 association:  $A + B \rightleftharpoons AB$ . The second and third models allow for the association of one molecule of chymotrypsin with each of two independent sites on soybean trypsin inhibitor with approximately equal affinity, in accordance with the following equilibrium scheme.

 $\mathbf{A} + \mathbf{B} \rightleftharpoons \mathbf{A}\mathbf{B} \quad K_{11}, \tag{9a}$ 

$$2\mathbf{A} + \mathbf{B} \rightleftharpoons \mathbf{A}_2 \mathbf{B} \quad K_{21},\tag{9b}$$

where

 $K_{11} = 2K_{\rm site}$ 

and

 $K_{21} = K_{\text{site}}^2$ 

where  $K_{\text{site}}$  is an intrinsic equilibrium association constant for binding to either individual site.<sup>12</sup> These two models differ because in the first, the value of  $f_{A,\text{comp}}$  is constrained to equal unity, while in the second,  $f_{A,comp}$  is allowed to vary to achieve a best fit. Values of the best-fit parameters and their respective 95% confidence limits are given in the caption to Fig. 5. The best-fit stoichiometry and affinity of the interaction between chymotrypsin and soybean trypsin inhibitor agree well with a previous characterization of this interaction obtained under very similar experimental conditions via analysis of sedimentation equilibrium [12]. While some of the improvement in the fit of model 3 over that of model 2 may be attributable to added flexibility conferred by the floating of  $f_{A,comp}$ , the best-fit value of  $M_{\rm B}$  obtained using model 3 is in significantly better agreement with published values of the molar mass of soybean trypsin inhibitor [7]. More importantly, it should be noted that the results plotted in Fig. 5 are qualitatively incompatible with the simple 1–1 association model, indicating that the method presented here allows rapid and unambiguous discrimination between appropriate and inappropriate models prior to evaluation of best-fit parameters within appropriate models.

## Discussion

Very recently, Bajaj et al. [13] reported the development of a combination optical flow cell for the simultaneous measurement of light scattering and the absorbance of a protein solution. In principle, this combination cell could be utilized in place of the separate light-scattering and absorbance detectors utilized here to accomplish similar objectives. However, the combination cell of Bajaj et al. [13] is three times less sensitive than our present absorbance detector with respect to the determination of concentration, due to the shorter optical path length. Additionally, the combination cell provides light-scattering data at only 2 angles in contrast to the 15 angles provided by the DAWN-EOS utilized in the present work. As a result, scattering data obtained from the combination optical cell are less precise, and the accuracy with which scattering data may be extrapolated to zero scattering angle, if necessary, is questionable.

We have extended a previously published method for the rapid quantitative characterization of equilibrium macromolecular self-associations to treat the more general case of macromolecular hetero- and self-associations. Validation studies have shown that the simultaneous measurement and analysis of light scattering and absorbance from composition gradients provide a rapid guide to the identification of equilibrium association schemes and the evaluation of equilibrium constants characterizing those schemes, providing that both reactants and products are detectable within the accessible range of solute concentrations.

As pointed out above, the presence of detectable protein-protein association is manifested as a significantly nonlinear relation between the gradient of composition and the gradient of light scattering, which is observable without any requirement for modeling. It follows that with the addition of robotics for sequential introduction of multiple samples,

<sup>&</sup>lt;sup>12</sup> Although the present data do not exclude the possibility that the two chymotrypsin binding sites on the inhibitor may differ in intrinsic association constant, the assumption of identical site affinities is subjectively preferred, in the spirit of "Occam's Razor," as the simplest model consistent with the data.

the present method can provide a moderately high throughput method for routine assay of reversible macromolecular association in solution. Modeling may then be employed to further characterize any associations so detected.

#### Acknowledgments

This research was supported by the Intramural Research Program, NIH, NIDDK, and by the appointment of A.K.A. to the NIDDK Research Participation Program, administered by the Oak Ridge Institute of Science and Education. We thank Dr. Peter McPhie, NIDDK, for critical reading of manuscript drafts.

# Appendix A. Removal of outliers from the scattering data

For each time point collected during the experiment, the function

$$R(t)/K = a_0(t) + a_1(t)\sin^2(\theta/2)$$
(A.1)

is fitted by linear least squares to values of  $R_i/K$  obtained from detectors 8 through 16, corresponding to scattering angles  $\theta_{(8-16)}$  between 60° and 142° in water, which are observed to provide the least noisy data. Then using the bestfit values of  $a_0$  and  $a_1$ , the squared residual corresponding to detectors 4–18<sup>13</sup> is calculated according to

$$\delta_i^2(t) = [R_i(t)/K - a_0(t) - a_1(t)\sin^2(\theta_i/2)]^2.$$
(A.2)

We now define a mean square residual characterizing the data obtained at intermediate scattering angles (i.e., low noise):

$$MSR(t) = \frac{1}{9} \sum_{i=8}^{16} \delta_i^2(t).$$
 (A.3)

Data filtering is then accomplished by eliminating each data point for which  $\delta_i^2(t) > 3 \text{ MSR}(t)$ .

# Appendix B. Composition dependence of Rayleigh scattering in a solution of self- and/or hetero-associating macromolecules

The value of the Rayleigh ratio for a mixture of dilute (thermodynamically ideal) solutes behaving as point scatterers is given by [8]

$$R(\{w\}) = 4\pi^2 n_o^2 \lambda_o^{-4} N_A^{-1} \sum_{i,j} \left(\frac{\mathrm{d}n}{\mathrm{d}w_{ij}}\right)^2 M_{ij} w_{ij}, \tag{B.1}$$

where  $n_o$  is the refractive index of solvent,  $\lambda_o$  is the wavelength of the scattering light in vacuum,  $N_A$  is Avogadro's number, and  $dn/dw_{ij}$  is the specific refractive increment of species *ij*. The contribution of each component to the total refractive index of the solution is, to within experimental precision, independent of the state of association of the component,

$$n = n_{\rm o} + \frac{\mathrm{d}n}{\mathrm{d}w_{\rm A}} w_{\rm A,tot} + \frac{\mathrm{d}n}{\mathrm{d}w_{\rm B}} w_{\rm B,tot}, \qquad (B.2)$$

where  $dn/dw_A$  and  $dn/dw_B$  denote the specific refractive increments of pure A and B, respectively.

It follows from Eq. (B.2) that

$$\frac{\mathrm{d}n}{\mathrm{d}w_{ij}} = f_{\mathrm{A},ij}\frac{\mathrm{d}n}{\mathrm{d}w_{\mathrm{A}}} + f_{\mathrm{B},ij}\frac{\mathrm{d}n}{\mathrm{d}w_{\mathrm{B}}},\tag{B.3}$$

where  $f_{A,ij}$  and  $f_{B,ij}$  (=1 –  $f_{A,ij}$ ) denote the corresponding mass fractions of A and B in species  $A_iB_j$ . We have ascertained experimentally that the specific refractive increment at 690 nm of each of the proteins used in the present study is approximately equal to a standard value of 0.185 cm<sup>3</sup>/g. Under these circumstances,  $dn/dw_{ij}$  is identical to within experimental uncertainty for all possible complexes, and Eq. (B.1) simplifies to

$$R(\{w\})/K = \sum_{i,j} M_{ij} w_{ij}, \tag{B.4}$$

where

$$K = 4\pi^2 n_{\rm o}^2 \lambda_{\rm o}^{-4} N_{\rm A}^{-1} \left(\frac{{\rm d}n}{{\rm d}w}\right)^2.$$
(B.5)

# Appendix C. Protocol for preliminary experiments on BSAfibrinogen mixtures

Solutions containing BSA (A) at w/v concentration  $w_A^o$  and fibrinogen (B) at concentration  $w_B^o$  were prepared. Three solution mixtures were then prepared by combining the two pure solutions in the ratio A:B = 3:1, 1:1, and 1:3, resulting in a total of five solutions with  $f_A = 1$ , 0.75, 0.5, 0.25, and 0. Each of these solutions has absorbance

$$A_{\rm o}(\lambda, f_{\rm A}) = f_{\rm A} w_{\rm A}^{\rm o} \epsilon_{\rm A}(\lambda) + (1 - f_{\rm A}) w_{\rm B}^{\rm o} \epsilon_{\rm B}(\lambda).$$
(C.1)

A conventional dilution experiment, described in AM, was performed on each of the five solutions, and the total concentrations of A and B at each time point were calculated according to

$$w_{\rm A,tot} = f_{\rm A} w_{\rm A}^{\rm o} A(\lambda) / A_{\rm o}(\lambda), \qquad (C.2)$$

$$w_{\rm B,tot} = (1 - f_{\rm A}) w_{\rm B}^{\rm o} A(\lambda) / A_{\rm o}(\lambda).$$
(C.3)

#### References

- E.M. Phizicky, S. Fields, Protein–protein interactions: methods for detection and analysis, Microbiol. Rev. 59 (1995) 94–123.
- [2] P.A. Srere (Ed.), Heterologous protein-protein interactions, Methods 19 (1999) (special issue), 191–349.
- [3] A.K. Attri, A.P. Minton, New methods for measuring macromolecular interactions in solution via static light scattering: basic methodology and application to nonassociating and self-associating proteins, Anal. Biochem. 337 (2005) 103–110.
- [4] C. Tanford, J. Roberts, Phenolic hydroxyl ionization in proteins. I. Bovine serum albumin, J. Am. Chem. Soc. 74 (1952) 2509–2515.
- [5] N.M. Kaye, P. Jolles, Characterization of the amino acids of bovine fibrinogen involved in the fibrinogen-thrombin interaction of the blood clotting process, Mol. Cell. Biochem. 20 (1978) 173–182.

<sup>&</sup>lt;sup>13</sup> Signals from detectors 1–3 are not processed.

- [6] K. Morimoto, G. Kegeles, Dimerization and activity of chymotrypsin at pH 4, Biochemistry 6 (1967) 3007–3010.
- [7] Y.V. Wu, H.A. Scheraga, Studies of soybean trypsin inhibitor, Biochemistry 1 (1962) 698–705.
- [8] K.A. Stacey, Light-Scattering in Physical Chemistry, Academic Press, New York, 1956.
- [9] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, Cambridge, 1987.
- [10] H.A. Saroff, Evaluation of uncertainties for parameters in binding studies: the sum-of-squares profile and Monte Carlo estimation, Anal. Biochem. 176 (1989) 161–169.
- [11] G. Rivas, J.A. Fernández, A.P. Minton, Direct observation of the self-association of dilute proteins in the presence of inert macromolecules at high concentration via tracer sedimentation equilibrium: theory, experiment, and biological significance, Biochemistry 38 (1999) 9379–9388.
- [12] U. Quast, E. Steffen, The soybean trypsin inhibitor is a doubleheaded inhibitor, Hoppe-Seyler's Z. Phys. Chem. 356 (1975) 617– 620.
- [13] H. Bajaj, V.K. Sharma, D.S. Kalonia, Determination of second virial coefficient of proteins using a dual-detector cell for simultaneous measurement of scattered light intensity and concentration in SEC-HPLC, Biophys. J. 87 (2004) 4048–4055.